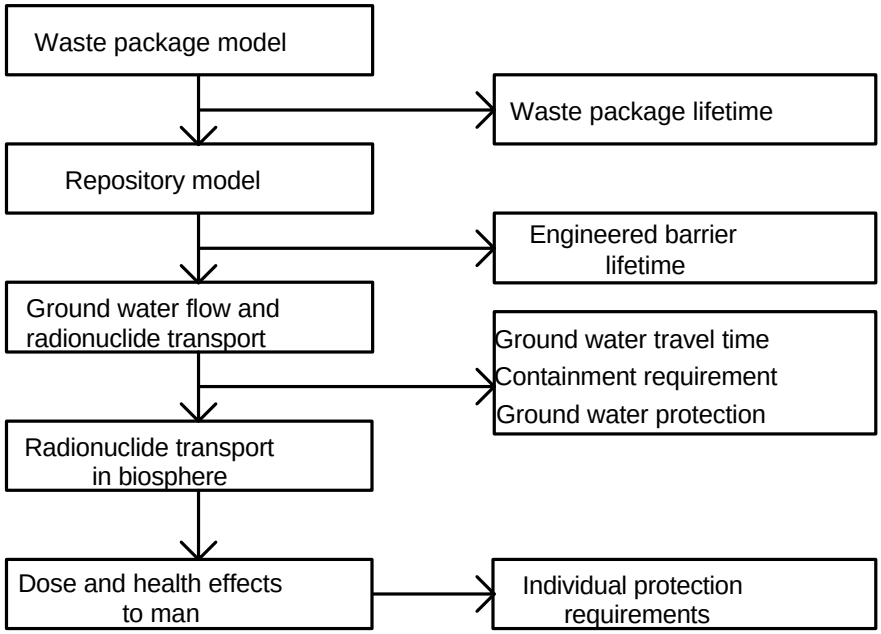
**Freiberg Online Geoscience Vol. 2**

FOG is an electronic journal registered under




# SENSITIVITY ANALYSIS AND SIMULATION UNCERTAINTIES IN PREDICTIVE GEOCHEMICAL MODELLING

## A CASE STUDY

**Christian Ekberg**

**Chalmers University of Technology**
**Department of Nuclear Chemistry**
**Göteborg, Sweden**

This report is slightly modified from the doctoral thesis of Ch. Ekberg, defended in public at Chalmers University of Technology Göteborg, Sweden on Wednesday, March 31, 1999. The faculty opponent was Dr. Norman Eisenberger, US Nuclear Regulatory Commission.

The report has been edited for publication in FOG by Dr. G. Meinrath, Institute of Geology, TU Bergakademie Freiberg and RER Consultants Passau, Germany.

# SENSITIVITY ANALYSIS AND SIMULATION UNCERTAINTIES IN PREDICTIVE GEOCHEMICAL MODELLING

Christian Ekberg

## Abstract

Computer simulations have become increasingly popular in many different areas over the years, owing mainly to more effective and cheaper machines. In many cases, the trend seems to be that computer simulations are replacing experiments, at least in areas in which experiments are very difficult (expensive) or impossible. One such area is that of attempting to foresee what will happen in the future. Such analyses are very important for a durable construction such as a repository for spent nuclear fuel, for example. In the modelling effort, several computer codes are used and input data are often used without scrutiny. However, this work shows that even the rather simple task of calculating the solubility of a solid phase in a given water is encumbered with the effects of different uncertainties. These uncertainties may make the calculated solubility vary by several orders of magnitude. Thus the input to the more complex codes, simulating processes in connection with the repository, will also be affected.

This report presents some computer programs for uncertainty and sensitivity analysis of solubility calculations. They are then illustrated by numerical simulations and estimation of uncertainty intervals for a case at the Äspö site in Sweden.

Some of the input data treated as uncertain parameters are the stability constants for the reactions between the metal ion concerned and the elements present in the selected water or the rock. Stability constants and the enthalpies and entropies of reaction for the thorium-water-acetylacetone-phosphate system have been determined experimentally. In addition to the values determined for these entities, uncertainty intervals are also estimated. A complexing mechanism for the thorium-phosphates at pH 8 is also suggested.

Keywords: uncertainty analysis, sensitivity analysis, SENVAR, UNCCON, MINVAR, thorium hydrolysis, thorium phosphate complexes, solvent extraction, potentiometric titrations

# Table of contents

# 1. Introduction

The art of foreseeing the future with the aid of computers has become more and more popular as the speed and memory of the machines have increased. Furthermore, the desire to know what happens in systems in which measurements are impossible or impractical has brought about the development of many computational models. Regardless of the aims of these computer models, they all suffer the same drawback: uncertainty. In the past, when the desired but unobtainable knowledge was guessed at by clairvoyants using more or less obscure methods, it was of course most difficult to make any system analysis: would the future be brighter if the shaman tossed the bones higher or was it more important if he used his right or left hand? Today, as most of the predictions of the future are made with computers that produce deterministic results, it is possible to make uncertainty analysis of their results. There exist several stages where different uncertainties may enter a simulation attempt [LIL 99], see Figure 1.1.

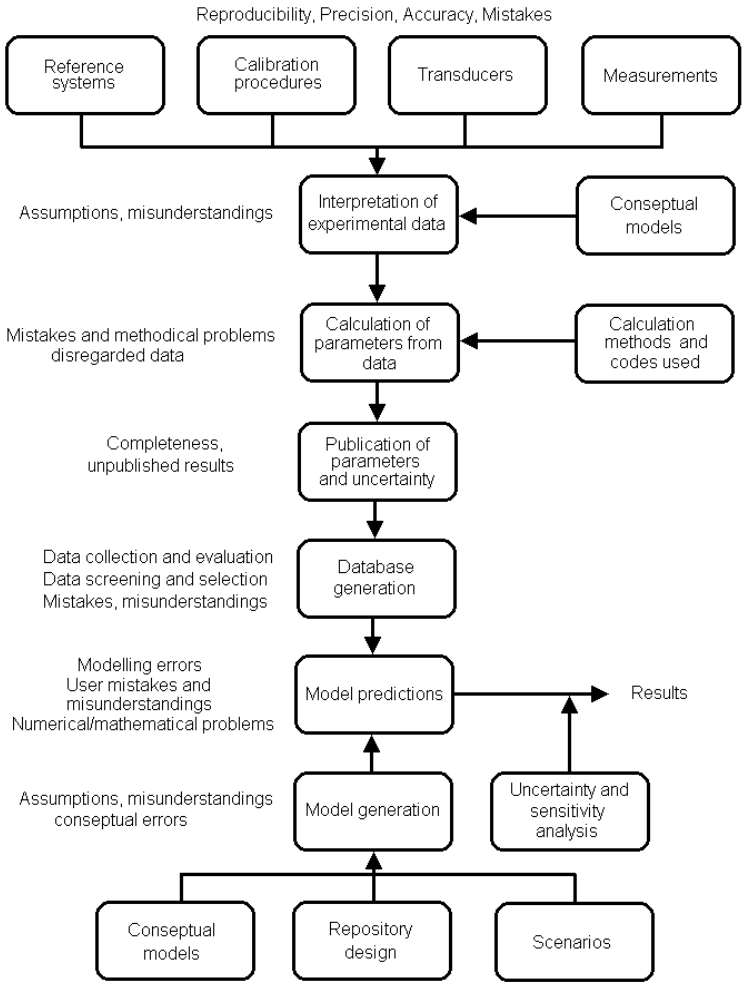



***Figure 1.1:*** *Factors affecting a simulation attempt, illustrated using a repository for spent nuclear fuel*

The modelling of some phenomenon are often the result of at least two paths, as seen in Figure 1.1. One of them originating from the reality, or at least the assumed reality, and another one originating from the efforts to theoretically describe a system. Clearly different uncertainties enter the modelling effort at different stages, which will be discussed in more detail in this report. Regarding the conceptual uncertainties there is a great problem in the fact that different models may describe different properties of a system. In this case, several arguments exist as to whether to believe in one model or another and it is often impossible to decide which is the better one. Often there is not much to do about such uncertainties except to compare the different results and hope that, with time, the predictions will converge to the same, or almost the same, value. Or one may keep the different results as describing different properties of the studied object, e.g. the wave particle dualiy of light. At present, the expressions "uncertainty analysis" and "sensitivity analysis" mean investigation of how changes in the data entered into a model implemented as a computer program will affect the results. They are usually made in two steps; the sensitivity analysis aims at determining which parameters are the important ones, and these are usually very few, and the uncertainty analysis shows how much the results are changed when the input data are changed to a certain level. There exist several methods to perform uncertainty analysis [EKB 95:1], and the source of uncertainties may be traced, perhaps through other computer calculations, back to some experiment that may need refinement.

In the construction of such a durable creation as a repository for spent nuclear fuel, information on what might happen in the future and how the not-yet-built repository will perform is of great interest. There are a number of ideas in different parts of the world about how construct such a repository. The method planned for Sweden is to place the fuel in bedrock several hundred meters below the ground surface. According to present plans by SKB, the fuel will be stored in copper or copper/steel canisters which will be placed in cylindrical holes drilled in tunnel floors in the rock. The remaining space in the holes will be filled with compacted sodium bentonite clay. The tunnels and shafts in the rock will be filled with a mixture of sand and sodium bentonite clay [KBS 83] [PAS 93].

Governmental licensing of a repository must be based on a judgement of its allowability from several aspects, the most important being safety and radiation protection. The necessary judgement must be based on an assessment of the performance of the repository in different time frames ranging from the present to a time $10^6$ years in the future. Further, in connection with the safety analysis of a repository, there are many disciplines to take into account, e.g. hydrology, geology, rock mechanics and geochemistry. For most of these disciplines, computer simulations are used.

To narrow the discussion further we will concentrate on the chemistry in and around a hypothetical repository. Several computer programs exist to aid geochemists in predicting the chemistry in the vicinity of the repository, e.g. EQ3-6 [WOL 92], PHREEQE [PAR 80], CRACKER [EMR 98:1] and PASSIPHIC [BÖR 98]. Input data to many such programs are

thermodynamic data often in the form of stability and solubility constants. The results from these programs may, in many cases, be used as input data to more complex codes for performance assessment, e.g. CALIBRE [WOR 95] and CRYSTAL [WOR 91]. In many cases, for performance assessment models, one important type of input data is the solubilities of the solid phases deemed to be the solubility controlling ones. These solubilities may be determined either experimentally or they may be calculated. In both cases there exist several uncertainties, but in this report only the effect of uncertainties in calculated solubilities will be more thoroughly discussed.

Calculated solubilities depend on stability and solubility constants which are usually determined from experiments. The literature often gives small intervals, if any at all, for the uncertainty in the measurements, and the absurdity of those intervals becomes apparent when several determinations are compared. Although the uncertainties in each experiment are small, i.e. the precision in the study is good, the accuracy of the obtained values may be small. There are many possible explanations for this phenomenon, of which systematic errors are a great contributor. Thus the obtained (mean) value may deviate from the "true" value with a value far greater than the estimated standard deviation of the experiment. If an uncertainty analysis is made, it is important that the input uncertainties are as reasonable as possible, i.e. taking all the reasonable values in to account, in order to avoid erroneous results.

## 1.1 Objective

The aim of this report is to investigate the effect of different uncertainties on the calculated solubilities of solid phases. These are:

Uncertainties in thermodynamic data, e.g. stability constants and enthalpies of reaction.

Uncertainties in water composition, of different origins

Conceptual uncertainties

In the case of thermodynamic data uncertainties, an effort has been made to establish thermodynamic data for the thorium-phosphate system with as reasonable uncertainy intervals as possible. This knowledge, together with that reported in many other studies, may serve as a basis for further calculations of reasonable uncertainty intervals of calculated solubilities.

## 2. Uncertainties in safety assessment - an overview

The following sections may give a hint of the vast number of uncertainties that are associated with the function of an underground repository for spent nuclear fuel. At the end of this chapter the discussion is narrowed down to the uncertainties that are explicitly dealt with in this report.

The methods for understanding and treating the uncertainties associated with a repository for spent nuclear fuel vary widely. The common aim is to ensure that the nuclear waste is stored safely for the time needed to render it harmless in terms of radiation. Some of the crucial questions in the decision / construction process may be [JOH 87]:

1. How to select a site
2. How to investigate a site
3. How to design a repository
4. Whether to authorise its construction and operation or not

In many cases, computer simulations are made to visualise the changes in and around the repository. It should however be recognised that the calculated results are only one of many possible predictions of what will really happen, i.e. all such calculations are encumbered with uncertainties. The more important sources of uncertainties may be grouped in the following way [CRA 87:1]:

1. Future states of the disposal system, including its environment
2. Models used to simulate these future states
3. Data and parameters used in the modelling effort

Once the uncertainties are grouped, it may prove difficult to determine where the borders between different groups should be drawn, but it is worthwhile to make an effort in order to attain a good overview and be sure that as many uncertainties as possible are taken into account.

Future states of the disposal system are generally referred to as scenarios. This includes such events as earthquakes, climate changes and human intrusions.

Uncertainties associated with models used to simulate future states of the repository arise mainly from the fact that not all mathematical models are good representations of reality. In some cases, several different models may be used to describe the same phenomenon. Then there obviously exist a conceptual (how it should be done) uncertainty. In addition, models may be incorrectly implemented in a computer program. One way of using a model incorrectly in a computer program is to extend the calculations to a region in which the model and/or the data are no longer valid.

The uncertainties associated with data and parameters are probably the most easily quantifiable ones since they may be derived from error propagation analysis or some statistical method that is applied to the results of a computer simulation.

## 2.1 Future states of the disposal system

One of the most complex contributors to the total uncertainty of the behaviour of a disposal system is the question of future events at the site or its surroundings. The identification of these events is often called scenario analysis and is vital not only to the safety evaluation of a repository but may also work as a guide for data collection and site selection. The events may be classified in the following way [INT 83].

1. Naturally occurring geologic events
2. Events caused by the actions of humans
3. Events caused by the repository system

It is easy to identify a great number of different events that may occur in the future. Unfortunately, it is also impossible to take any action against the main part of these. This statement stresses the importance of discussions of the likelihood and severity of each event.

Many approaches can be used to solve this problem. One is to try to quantify the effects of a certain event which, may then give rise to new events. The effect of each event may be the subject of computer simulations and speculation. A common means of evaluating them is to visualise the situation by placing the different events in boxes and entering the effects as lines connecting the boxes [SUM 93]. Unfortunately, this usually results in very complex figures and is, despite its simplicity, cumbersome to work with. However, development of computer programs for visualisation, and importance levels have made this method more easily handled [CHA 95] and it is now a tool for the governmental licensing process in Sweden.

It is important to remember that, while in most cases it is impossible to foresee the future, in the case of nuclear fuel repositories, it is essential that the attempt is made, i.e. development of reasonable scenarios instead of trying to establish a fixed future. To give a perspective of the time span involved, one can conclude that the length of time discussed in these scenarios is further into the future than Stone-Age man is into the past. Thus history may not provide sufficient information for it to be used to predict many of the possible events in the future. Especially those which concern the activity of humans.

### 2.1.1 Geological and climatic events

The possibility for many natural geological events to occur may depend on the choice of place for the repository. This would hold for example for earthquakes and volcanic eruptions. Most other processes that change the properties of the rock barrier are, in this context, very slow,

and they may thus be divided into different intervals: the present, from now to about 10000 years and beyond 10000 years in the future
[INT 83].

At present, most uncertainties lie in parameters and modelling. These may include spatial variation of parameters and choice of model. The uncertainties thus introduced are found in separate sections in this report.

For Swedish conditions, one very likely event within the next 50000 years is the occurrence of a new ice age, i.e. a time when the surface of the repository is covered with thick land ice. Some effects of this event are more or less possible to foresee, for example the possible occurence of pluvial lakes, stresses caused by ice load and changes in sea level. On the other hand, it is difficult to predict the hydraulic conductivity and the flow paths in the rock during or after a glaciation. Attempts have been made to calculate the water flow through the Äspö site in Sweden for the different phases of a glaciation. These calculations show that the flowpaths of the groundwater will be highly affected and that oxidising water may reach the repository [VOS 98] [GLY 98]. These effects may be a reason to place the repository in a place where no glaciation is likely to occur. However, there are factors which are of much greater importance for the performance of the repository than glaciation, for example seismic activity.

In the longer time scale, many events are possible. There may be meteor downfall (this is possible today, too), volcanic eruptions, earthquakes, landslides and several ice ages [CAM 78]. The method available for predicting these events is the past. However, ice ages may be predicted from astronomical calculations [MIL 41]. Stretches in the rock and continental drift are in many cases possible to foresee, but the effect of these events on the repository and the surrounding rock may be vast or nonexistent [CAM 78]. In any case, the effort should be in determining to what extent an event calls for extra precaution.

### 2.1.2 Events induced by humans

The repository may be breached for many reasons, intentional or otherwise. The most likely is perhaps the desire to reclaim elements in the fuel or ores close to the repository. Another possibility is accidental breaching owing to a loss of informatinon about the repository. On a short scale, this would probably be the result of a collapse of the society in that country or on earth as a whole [PRO 90] but, on a larger scale, it is higly probable that no information remains after an ice age. Barring such occurences, unintentional breaching may be prevented and the explorers may have information about the location of the repository.

To foresee what will happen if the repository is intentionally violated is almost impossible since the effects depend on the technological level of that time. However, it is currently being discussed whether retrievability of the fuel is advantagues or not. A safer technique for transforming or disposing of the waste may be discovered a short time after the closure of the repository, while a good technique for opening it safely may not yet exist.

The event of unintentional breaching of the repository is more easily handled. The recommended approach is to use an over estimating direct-release analysis combined with expert opinion [PRO 90].

### 2.1.3 Features, events and processes (FEPs) caused by the repository

The repository system may be described by a system of FEPs. The sometimes called system uncertainties arise in the procedure of establishing these FEPs, since the completeness of the description selection can not be assured. As an alien part of the rock, the repository may interact in several ways with the host rock. As a whole, these interactions and their effects are more a continuing process than an event. However, many of these processes have already been quantified and analysed [CHA 95]. The next question after whether it is mechanically possible to build the repository is how to build it to achieve as little negative interaction as possible with the rock itself. Unfortunately, there still exist some uncertainties, for example how the pressure will be compensated with the swelling pressure of the bentonite and whether it is necessary to use concrete in the construction of the HLW repository.

## 2.2 Modelling future events and states

If an event is deemed likely to occur, it is important to try to estimate its effect. Such estimations are often made by computer simulation. However, simulations are often subject to the drawback that there is no or very few data to use as input to the assumed model. It is also important to remember that no simulation is true, i.e. it does not completely describe what actually happens. It may further be noted that, in some cases, a change in only one input parameter may change the result considerably [SUN 86]. It may therefore be advantageous to make a sensitivity analysis of a computer model in order to determine a ranking list of the parameters.

Modelling features in nature include several steps which may be illustrated as a flowsheet such as the one in Figure 2.1. This figure was originally developed by Andersson et al. [AND 95] and then modified by the author of this report.

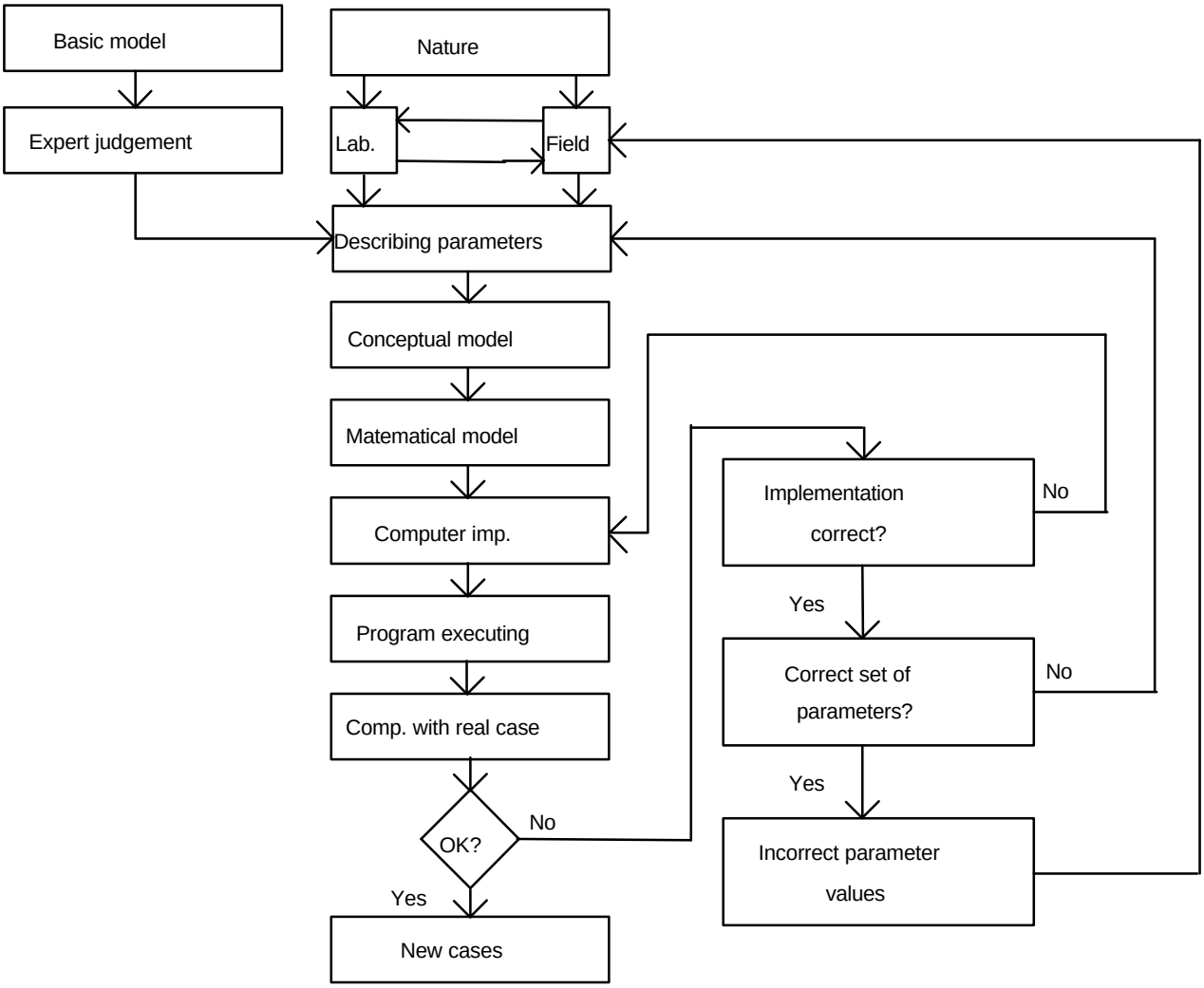



***Figure 2.1**: Flowsheet of the construction of computer simulations [AND 95]*

The starting point is usually nature, but in many cases there are also some basic laws describing natural phenomena, e.g. Gibbs phase rule and its implementation, that must be taken into account. The visualisation of reality is usually made by measurements in the field or a laboratory. These values will then act as parameters that describe the system. From those parameters, a conceptual model of the investigated phenomenon is constructed. To make this model usable, it is usually incorporated in a computer program in which the model is expressed mathematically. The program is then executed and the result compared with the natural phenomena, if possible. If the result approximates these phenomena fairly well, a new case is tested. If reality is not reasonably reflected, several steps must be taken to identify where the error was made. These steps include checking whether the computer code is correct and whether the conceptual model is correct and checking the adequacy of the parameters and the correctness of their numerical values. Clearly, there are several steps in this chain that are uncertain, and there are several methods to identify them. Some of these are described in the following sections.

The modelling of the performance of a repository for spent nuclear fuel may be divided according to Figure 2.2 [CAM 88]. Each of the boxes in the left hand column in Figure 2.2 represents complex systems that may be modelled with greater or lesser simplifications. The right hand column shows the expected results of the modelling. It should be noted that a regulatory criterion for the results from the "Groundwater flow and radionuclide transport" exist in the USA but not in Sweden. Some programs take all of Figure 2.2 into account, e.g [AHN 90], but in such cases many of the phenomena are simplified even more than in the original models.

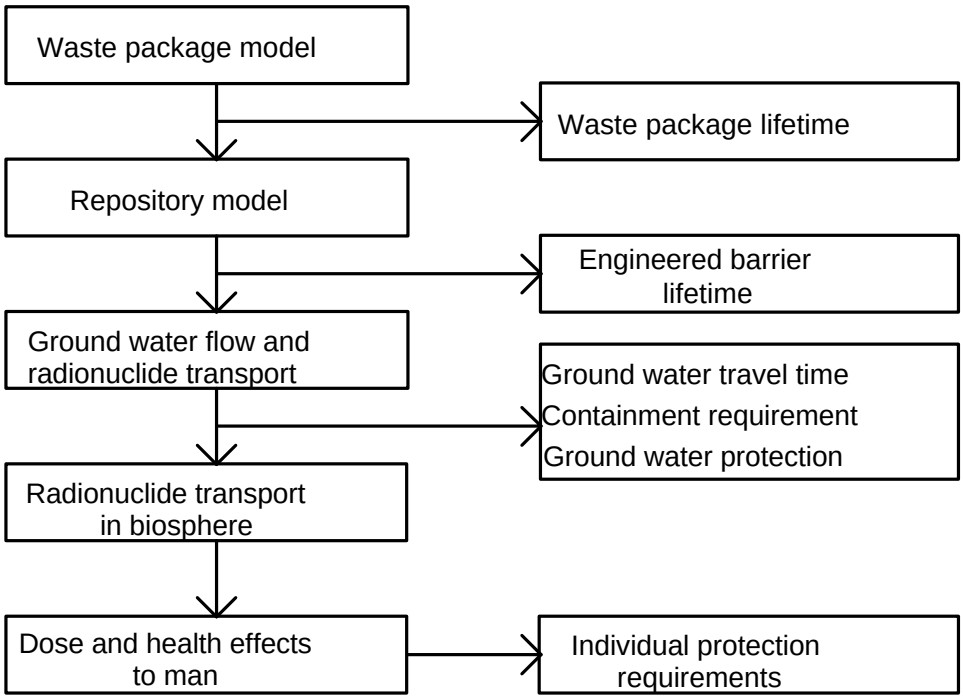



***Figure 2.2:*** *Illustration of consequence modelling sequence [CAM 88].*

There is also the problem of validating the results of the models. In some cases, it is impossible to validate a model [NOR 92]. Thus it is sometimes more important to increase the understanding of the system being modelled, for example by sensitivity analysis, than to try to validate it against field data. However, these data may be very uncertain, thus making the validation procedure difficult.

## 2.3 Evaluation of models

If there exists a possibility to verify and validate a computer model, it is important that this is done in order to ascertain the correctness of the chosen approach. Unfortunately, this is not always done properly, which may result in serious misjudgements when making the decisions for which the modelling was performed.

### 2.3.1 Verification

A model is verified chiefly to confirm that the computer code functions as intended. This includes the choice of the mathematical expressions and databases used, together with the range in which they are valid. It is also possible to examine the model with respect to the sensitivity to changes in the input data. Simulations with various databases, and comparisons of their results are thus needed. However, it is essential to remember that different databases are often derived from the same experiments and are thus equipped with the same faults. There exists no general method to verify computer codes and hence it must be the responsibility of the programmer to verify the model in the best possible way. Nordstrom [NOR 92] gives a suggestion for how this should be done. Since this procedure is not wholly general, however, there will still be custom made verification processes, the reliability of which experts may argue.

### 2.3.2 Validation

Validation of a model is essentially performed to examine whether experimental data can be reasonably reproduced or not. In some cases, it is also desirable to make comparisons with the results of other models or use the judgement of independent experts. There have been some international attempts to validate models, such as the CHEMVAL project [REA 91] and INTRAVAL [INR 92]. The latter is concerned mainly with the transport of radioactive substances and hydrology and the former with geochemistry. It is noteworthy that one conclusion of the INTRAVAL project is that most of the present performance assessment migration models may be overly conservative in their results.

Many models, e.g. geochemical models and computational fluid dynamics, are almost impossible to validate against experiments since there is no complete knowledge of the system but only approximations. Therefore, the only method of validation may in this case be invalidation, i.e. showing that the model gives unreasonable results rather than showing that the results are correct. This approach may be a very powerful tool in the validation of models

The lack of organised validations of models seems to apply to most disciplines, and great effort should be made in constructing validation strategies before the computer models are being used in any real simulation.

## 2.4 Data and parameter uncertainties

Data and parameter uncertainties are perhaps the types that are most easily quantifiable. These uncertainties may arise from several sources, such as [CRA 87:2]:

1a. measurement errors
1b. misinterpretation of data
2. paucity of data
3. spatial variation of parameters
4. assumptions regarding behaviour of the system, conceptual uncertainties

### 2.4.1 Measurement errors and misinterpretation of data

A lack of precision in the use of an equation may be caused either by measurement errors or a misinterpretation of data. It may be very difficult to determine which is the cause. In addition there may be numerical problems in the calculation process. Moffat [MOF 88] gives an example of misinterpretation errors, in which the parameters in the equation used should be bulk averages while local values measured with a probe are used instead. This leads, in this case, to unrealistic values of the calculated heat transfer. The main cause of this problem is that there is no clear definition of which values should be used. The result may therefore be more or less inaccurate. It is clearly most vital to understand what is needed in the equations and how it should be measured.

Pure measurement errors can sometimes be found by recalibration of the instrument. However, this calibration may itself be incorrect. It is sometimes easy to detect the possible uncertainties in the calibration since new calibrations are usually performed before each experiment. Another case of "measurement errors" is that in which the result of the measurement is not correctly read or transmitted.

Errors whose origin lies in a misinterpretation of which data should be used in an equation are often difficult to locate. In most cases, it may not be possible to detect these errors, unless a clearly impossible answer has been obtained. There may be many small factors that make the measurement incorrect. To use the data obtained without knowing how they were measured may be fatal. However, if all data are regarded with a certain amount of suspicion and their origin is investigated thoroughly, the effect of these errors can be reduced significantly. If the number of data are great it is possible to use a sensitivity analysis to select the most important ones, which are more thoroughly investigated.

### 2.4.2 Paucity of data

Paucity of data is perhaps the most common source of a lack of accurracy in the results of many models since many vital parameters may be missing. In some cases, not enough measurements have been made to make a good estimate of the true values. In other cases, however, there may be no available data at all. Very few possibilities are then open to a modeller to make a model. One common way to solve this problem may be to perform a sensitivity analysis of the model and thus determine whether or not the parameter concerned is important. If this approach is not possible, it may be possible to seek an expert opinion on the subject. In many cases, the model result is not obviously wrong although the parameter estimation may be wrong. It may then be almost impossible to find and correct a value that may be of great importance in later simulations.

### 2.4.3 Spatial variation of data

Spatial variation of data is one of the most important contributors to uncertainties in the determination of the geology of a disposal system. In a rock, minerals may appear almost at

random, thus making accurate predictions of their distribution nearly impossible. In addition, taking samples from the rock may change its properties, e.g. holes are left in the rock and the sampling may create new fractures or dilute the groundwater with drilling water. It is sometimes argued that the water comes into contact mainly with the relatively few fracture filling minerals, but this is impossible to state in all given circumstances, and thus the entire composition of the rock may be important.

Another example of spatial variations in the rock is hydraulic parameters such as conductivity and porosity [GEL 76]. These variations are propagated to the calculated groundwater flow velocity through the use of Darcy's law and are therefore of great interest. It may be noted that such parameters as conductivity may, owing to its spatial variability, be treated as stochastic parameters [CRA 87:1]. This is not usually the case, however, since in most computer programs an average is used, with the drawback that a large error may be induced in the calculations.

Some changes in time are rapid enough to appear to be changes in space instead of time. An example may be the chemical fluctuations in an aqueous system [SCH 88]. This may occur for example for oscillating reactions when two measurements are made at some distance apart in the solution. This distance must be greater than the diffusion distance for the duration of the measurement. The measurement of these fluctuations may be an example of misinterpretation errors, as described in section 2.1, if the phenomenon is not known to the performer of the experiments.

### 2.4.4 Assumptions regarding the behaviour of the system

If no or very few experiments are made, it may be necessary to guess how a system will behave. This guess is often based on some assumptions that may or may not be correct.

If the geology, hydrology and chemistry of a repository are well determined and as many interactions as possible are investigated, some parameters may still be guessed at rather than measured, e.g. the distribution and composition of the minerals in the rock. However, it is important to bear in mind the uncertainties associated with assumptions regarding a system. Some of these assumtions may still be used. Unfortunately, it may be that their origin has been forgotten causing them later to be treated as facts.

## 2.5 Treatment of data and parameter uncertainties

There are several methods for treating data and parameter uncertainties. Those most commonly used are probably the following:

1. statistical methods
2. interpolation techniques
3. differential analysis techniques

In the uncertainty calculations presnted in this report only statistical methods have been used. The other two thechniques are included only for the sake of completeness. All of these methods naturally require the uncertainty interval for each input data. The quantification of such intervals is further discussed in section 5.3.5.

### 2.5.1 Statistical methods

Statistical methods are perhaps the best developed and most widely used of the techniques for treatment of uncertainties in data. Most of the experimental applications involve a random error that influences the experimentally determined data. The errors may be estimated by repetition of the experiment under the same conditions. However, this is not the case when experiments or events are simulated by the aid of computers. A computer code is based on a theoretical model of what happens in the experiment, and the models are usually deterministic, i.e. a given set of input variables will always give the same output value. Therefore, in performing an uncertainty analysis with statistical methods, the design selected should not include replication [HAR 83].

Statistical methods may be divided into two different subgroups:

1. experimental design methods
2. sampling methods

#### *2.5.1.1 Experimental design methods*

The main concept in experimental design methods is to use a specific design to select a specific set of input variables and their mutual order. These methods are used both for computer simulations and laboratory experiments. A typical experimental design method is the factorial experimental design. These designs vary all input variables at the same time, which makes it different from the old "one-factor-at-a-time" designs. The efficiency of factorial designs in the estimation of interactions between input variables is well documented, and the theories found in many textbooks, e.g. [BOX 87, BAT 88]. Factorial experimental design makes use of all data in the estimation of the effect of each input variable. It is clear, however, as the number of input variables increases, that the number of computer runs needed will increase rapidly. One way to eliminate this problem is to assume that all high order parameter dependencies are zero. This will give only the main effects of the input variables, but those are usually the only ones of interest.

Iterated fractional factorial design may be used to minimise the number of calculations and not lose a great deal of information [AND 93]. This method creates a small number of groups and then assigns each variable randomly to one of the groups. Fractional factorial design combines the symmetrical properties of factorial designs and the statistical sampling of the Monte Carlo (MC) methods. Experimental design methods are usually better suited for sensitivity analysis than uncertainty analysis.

#### *2.5.1.2 Sampling methods*

Sampling methods are based on treating the model input parameters as random variables with a given probability distribution and, if necessary, correlations, i.e. the first order interactions. The parameters are then selected by the use of some sampling procedure. The computer program is executed for each set of inputs, and the result of each set may be used to calculate the distribution function for the output [WIL78].

There are three different methods for designing the sampling procedure: Monte Carlo ( MC ), stratified sampling and Latin Hypercube Sampling (LHS) [MCK 79]. There are several reasons for preferring some kind of random sampling method [IMA 87]:

1. If done properly, MC methods can be designed to avoid some inputs for which the model is not valid.
2. The MC approach varies all input parameters simultaneously, thoroughly exploring the input space, and can be made very efficient.
3. If the probability distributions assigned to the inputs are meaningful, then statistical estimates of output percentiles, means and variances can be made.

Two of these techniques, MC and LHS, will be discussed in more detail in Appendix 1. A general conclusion is that the most effective is LHS. Sampling methods provide a simple way to investigate the effect of uncertainties in the input of a model, but there is one major drawback. It takes many computer runs to make the result statistically significant. The main advantage of the MC-based simulations are that the result may be analysed with common statistical tools. In most cases, it is possible to include the statistical calculations in the code as seen in the following sections.

### 2.5.2 Interpolation Techniques

Interpolation techniques are often used to estimate a complete surface structure from spatially distributed data if no model exists of how the data should be distributed. This is usually done by using a linear estimator as shown below.

$$\hat{o}(x_0) = \sum_{i=1}^{n} l_i o(x_i) \qquad (2.1)$$

where $\hat{o}(x_0)$is an estimated value at a points between data points, $o(x_1).....o(x_n)$ are measured values and $\lambda_1...\lambda_n$ are the weights of each measurement. In many cases the sum of the $\lambda_i$ is equal to one and thus the estimated value is a weighted mean. There exist several methods to select $\lambda_i$. One such technique is kriging, as described by [MAT 69], [MAI 93]. The main principle of this method is to select $\lambda_i$ to create an estimate being unbiased with minimum variance. Kriging is used mostly by mining engineers and geologists, since it is common in those disciplines that it is not possible to obtain the properties at any given location.

### 2.5.3 Differential analysis techniques

The differential analysis techniques are based mainly on a Taylor expansion of the simulated properties and the associated partial derivatives [IMA 85]. The general idea is to treat the dependent variable of interest as a function, f, of the independent variables, $x_1$.....$x_k$. This function is then expanded as a Taylor series about some vector, $X_o = (x_{1o}....x_{ko})$, of base case values for variables $X = (x_1.....x_k)$. This series is usually truncated after the the first order derivative:

$$f(X) = f(X_o) + \sum_j \frac{df(X_o)}{dx_j}(x_j - x_{jo}) \qquad (2.2)$$

The first order Taylor series described above generates a linear model as described by:

$$Y = b_0 + \sum_j b_j x_j \qquad (2.3)$$

This model may then be used for uncertainty or sensitivity analysis, i.e. calculating only the effects on the linear equation rather than for the complete model.

## 2.6 Sensitivity analysis

If a computer model is used to simulate some complex process, it is of great interest for both the programmer and the user to make a sensitivity analysis of the system and the program. The sensitivity analysis is also important for verification of the model.

A sensitivity analysis usually means that the change in the result owing to changes in the input is investigated. Such an investigation has several important benefits. i) First, the correctness of expert predictions may be tested. ii) Second, unimportant variables or unnecessary model complexity may be revealed. iii) Third, the input data may be ranked with respect to their influence on the result. As a result of the second and third advantages, the number of variables used in an uncertainty analysis can be decreased significantly. The uncertainty analysis may thus be performed more rapidly, saving computer time for complex codes. However, one might argue that this kind of parameter screening before the actual uncertainty analysis removes important second order interactions, e.g. coupled effects.

There are many ways to perform a sensitivity analysis. Some of the most important ones are described below.

### 2.6.1 Response surface methods

Response surface replacement for computer simulations is generally based on some experimental design to select input values for the computer program. The method of least squares is then used to estimate the parameters in a linear equation approximating the actual response surface, see Equation 2.2. Generally, the derivatives in this equation are dependent on

x but, in the approximation mentioned above, they are assumed to be constant, thus yielding a multidimensional plane. Linear models are ordinarily expressed with an error term added to represent stochastic variation. Computer models, however, usually produce deterministic output and thus differences between the linear equation and the computer model may be the result of lack of fit rather than stochastic variations.

When a linear model is created, it may be used to predict the sensitivity not only of each parameter but also of an uncertainty analysis. In the latter case, Monte Carlo sampling (cf. 2.6.2), of the input values is needed to estimate the effect on the dependent variable. It is also possible to obtain expectation values and variances for the result directly from Equation 2.3 in the following way.

$$E(Y) = b_0 + \sum_j b_j E(x_j) \tag{2.4}$$

$$Var(Y) = \sum_j b_j^2 Var(X_j) + 2\sum_i \sum_j b_i b_j Cov(x_i, x_j) \tag{2.5}$$

The expectation value and the variance may then be used to make a description of the system response.

### 2.6.2 Monte Carlo methods

The Monte Carlo approach may be used directly with the simulation by choosing the input values from the interval of the variables used. In some cases, the distribution in this interval may also be taken into account. The approach may be to hold one variable at a fixed level over ten simulations, for example, and then to fix variable number two and continue in this way until all the variables are used. The variances are then compared and the variable that gives the smallest variance in the result when held fixed is deemed to be the most important, the next smallest to be the second most sensitive and so on. This method, however, is somewhat unstable since there must exist a significant difference in the sensitivity of the variables in order to obtain a reproducible answer, i.e. the result must not be dependent on the seed to the random number generator. Such a disadvantage may be avoided by using a fixed random matrix made in the beginning of the simulation. The number of computer runs is also very large for simple MC selection, so this method must be modified by some other sampling procedure to be time efficient.

### 2.6.3 Differential analysis

It is possible to make a sensitivity analysis from a differential analysis, see Section 2.5.3. The coefficients in such a Taylor expansion may be normalised and thus used to develop a ranking of the variables of importance. Further, estimations of the expectation value and the variance may be derived from Equations 2.4 and 2.5, respectively, since these are properties of linear models and may thus be used in both linear regression models and Taylor series [IMA 85].

# 3. Uncertainties in solubility calculations

The solubilities of different solid phases are common input data to, for example, performance assessment models. Usually these solubilities are given as fixed values, which may be greatly or slightly wrong depending on their origin. Therefore also the predictions made by the performace assessment model will be uncertain. The method to avoid this is to give the input data to a model as an interval rather than as a fixed value. The model will then give an interval as a result. For the case of solubilities there are two methods to obtain the data, either by experiments or by calculations . Even in the latter case experiments are needed to give more basic data such as stability constants for different chemical species. The uncertainties in solubility calculations have been investigated with the aid of some computer programs which have been written for different purposes depending on which uncertainty to evaluate.
The uncertainties considered in this report are:

Uncertainties in thermodynamic data such as stability constants and enthalpies of reaction. This is done by the SENVAR program package.

Uncertainties in water composition using the UNCCON package. The water composition uncertainties may originate from two sources. Either there are measurement uncertainties or the water composition may be estimated using a rock-water interaction program. In the latter case the rock composition is uncertain and the effect of this uncertainty has also been investigated using the MINVAR package.

There may also be several conceptual models for the calculation of solubilities under different groundwater conditions, each of them giving a slightly different answer. These conceptual uncertainties will be described below as a calculation exercise.

## 3.1 Common features of the programs

The computer programs developed within the scope of this report all have some features in common, the most important being that they are all written in the programming language C. However, in some cases, sub-programs written in another computer language are used. The chemical "calculation motor" in all the program packages are the PHREEQE program [PAR 80].

All the programpackages use a well defined and detailed directory structure developed by Emrén [EMR 98:1] in order to make it easy to find the appropriate files and to maintain a good overview. In addition, most of the programs described below are in reality program packages but is is only the name of the master program that is described in the section headings. The sub-programs are directed and run from the master program. Each of the programs may be run separately, but it is vital to remember that some information from preceding programs may be needed in order to make a single run. Therefore, this approach is not recommended unless the manual for the program or the source code has been thoroughly examined. However, if done properly, this procedure makes it fairly easy to change some part

of the program without extensive work and also makes it possible to run each part separately. In this way, each part of the calculation can be investigated, allowing for detailed examination.

## 3.2 Rock composition uncertainties, MINVAR

The MINVAR package is designed to calculate how the composition of a water is affected by changes in the mineral composition of the host rock. For these purposes, there is a need for a program that calculates the composition of a water that is in contact with a rock with a given mineral abundance. Several codes with nearly the desired properties exist, and it is thus possible to use any of them as sub-programs with some changes in the source code of MINVAR.

The assumptions made on the mineral uncertainty are that each mineral abundance may be given by a mean value and a minimum value, thus defining a symmetrical interval in which the abundances are assumed to be uniformly distributed. A Latin Hypercube Sampling is made within these intervals. This is particularly simple in this case since the distribution assumed makes the strata that are equal in probability also equal with respect to percentage.

The calculations then follow the typical LHS procedure, i.e. the mineral abundances in the input files to the simulation program CRACKER [EMR 98:1] are changed according to the LHS matrix. When all the "samples" are taken, some statistical estimators for the different elements are calculated, e.g. the mean concentration, the standard deviation and a 95% confidence interval based on a log-normal distribution assumption for the element concentrations.

It is further possible to display the concentration fluctuations for each chemical element and water parameter, e.g pH and pe, in order to determine the accuracy of the normal assumption mentioned above. Presently the graphs displayed are a calculated frequency function and a cumulative distribution function (cdf) for the pH, pe, and the concentration of different elements in the water. However, since one of the advantages of LHS is that few samples are needed to cover the variable space, it may be difficult to obtain any relevant information from the frequency function since the few samples may give local fluctuations in the function values.

The output produced by MINVAR may, in addition to illuminating the desired problem, be used as indata to the UNCCON program, see Section 4.1.2, which calculates how the solubility of a solid phase varies with the water composition.

### 3.2.1 Program description

The subprograms that constitute the MINVAR package may, roughly, belong to two groups: one statistical sampling and evaluation part and one simulation part. For the simulations, any program that produces a water composition from a rock composition may be used with minor modifications in the statistical part. However, for practical purposes, as described above, the CRACKER program has been used here. However, there is another positive effect of this selection and that is that being an equilibrium program, CRACKER probably overestimates the

local fluctuations in the fracture. This produces concervative results which is desired in for example a performance assessment.

#### *3.2.1.1 The CRACKER program*

The simulation part consists at present of the CRACKER program [EMR 92], which calculates the composition of a water flowing through a crack with a given mineral composition. This is done by dividing the rock surface into small hexagonal parts, each such part containing one mineral. The sizes of these diffusion cells are dependent on the flow rate, and their purpose is to be able to assume that equilibrium exists within each cell. These hexagonal mineral grains are then distributed randomly across the surface, maintaining the percental distribution. The water flowing through the rock is then assumed to be in equilibrium in each cell, and the cells are mixed according to Figure 3.1.

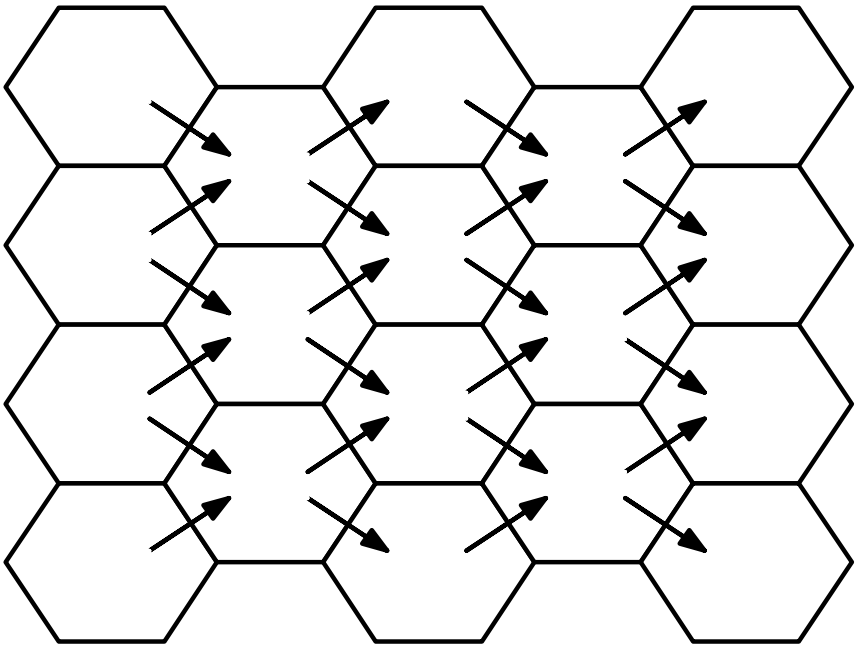


*Figure 3.1, mixing scheme for CRACKER*

After a number of length steps, it is possible to assume that the water has reached a steady state in the sense that, if all the waters from the last row are mixed, the change in two successive rows is negligible. The water composition thus obtained is assumed to be the one existing in a natural fracture system with the given mineral abundances.

This approach was tested with waters from the Äspö site in Sweden [EMR 98:2]. It was found that, when a reasonable mineral set is used, the calculated element concentrations were in good agreement with the measured ones.

### 3.2.1.2 Statistical sampling and evaluation

The statistical sampling and evaluation part is easily defined as the MINVAR package without CRACKER. Since the simulation part is incorporated in the MINVAR package, it will be found in the flowchart of the program below, but the description is found in the section above.

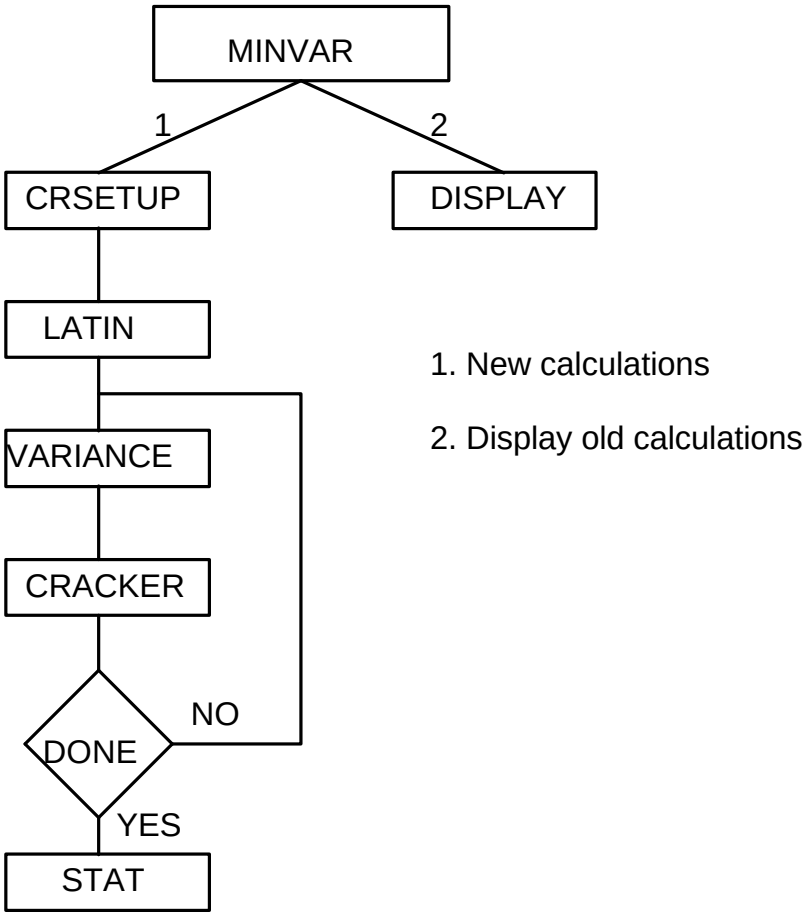



***Figure 3.2:*** *Flowchart of the MINVAR package.*

During a typical calculation, the programs are run by the MINVAR master program. A short description of each program may be necessary to understand the flowsheet properly.

MINVAR: This is the master program which is invoked when a calculation is started. There are mainly two options: to start a new calculation or display an old one.

CRSETUP: If a new calculation is selected, the necessary information is given to this program by a menu system. The choices that must be made are for example: the length and width of the crack, what water to start with and the rock composition, together with the associated uncertainties as intervals.

LATIN: The frame for the latin hypercube sampling is made here, i.e. the intervals for each mineral are divided into equal parts and one value is

selected randomly from each part, whereupon the values are mixed randomly, thus producing a matrix with random values.

VARIANCE: This program changes the infiles for CRACKER according to the random matrix created by LATIN. This is done one time per iteration cycle.

STAT: The statistical estimators are calculated here. The result file contains the mean concentration, standard deviation and a 95% confidence interval for each element plus pH and pe.

DISPLAY: To make the results visual, this program is included to show the calculated frequency function and the cumulative distribution function (cdf). These plots may be drawn for any of the entities included in the calculations.

For more information, see [EKB 99:1] for the source code and handling instructions.

## 3.3 Water composition uncertainties, UNCCON

The UNCCON program investigates how uncertainties in water composition affect the solubility of a solid phase. Each element concentration is given as a uniformly distributed interval, indicated by the user, from which the samples are taken with the LHS technique. Other factors that can be selected are different iteration criteria, e.g. number of LHS intervals, and which database to use. A simple sensitivity analysis is also included. The basis for this is simply to give one of the investigated parameters a high and a low value, while the others are kept at their mean value, and then run PHREEQE once for each case. This is repeated for all the parameters, and the difference in solubility for each case is compared. The parameter that gives rise to the greatest difference is considered to be the most important one and so on. This approach will obviously give only a rough estimate of the internal ranking since only first order effects are taken into account. However, the sensitivity analysis is usually used only as an indicator in this case, and no further calculations depend on its result.

The result of an UNCCON uncertainty calculation consists of two parts, first, a file containing some statistical estimators such as the mean solubility, minimum and maximum solubility and a confidence interval for the mean and, secondly, a part consisting of plots of the calculated frequency function and the cumulative distribution function (cdf) for the solubility. As there are relatively few samples, owing to the LH sampling, the plots will be somewhat jagged, but they will at least give a hint of the distribution of the solubility.

### 3.3.1 Program description

The different subprograms of the UNCCON program package are hierachically distributed according to Figure 3.3.

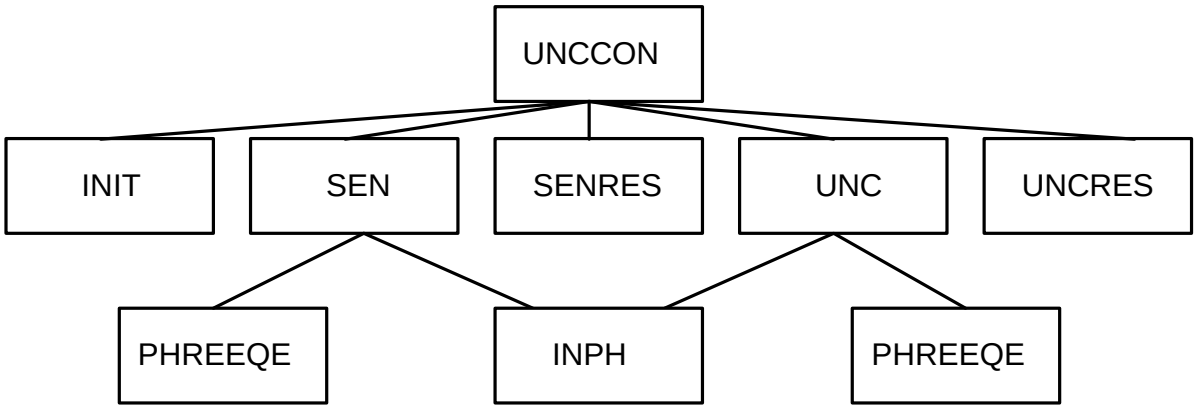



***Figure 3.3:*** *Hierarchical diagram of the UNCCON package*

A short description of the different programs are given below and for a more thorough description see [SAM 96].

UNCCON The master program. Starts and controls the other programs. It also saves results from both the sensitivity and uncertainty analysis.

INIT A program in which the user determines the conditions for the sensitivity and uncertainty analysis.

SEN The main program in the sensitivity analysis. Handles the input from INIT, runs PHREEQE and creates a result file to SENRES.

SENRES Evaluates the results from SEN and produces a result file containing the elements in order of importance.

INPH Creates a PHREEQE infile from data given by UNC or SEN.

UNC The main program in the uncertainty calculations. Runs PHREEQE and produces a result file to UNCRES.

UNCRES Calculates the mean solubility, standard deviation and a confidence interval for the mean and writes this together with the largest and smallest solubility in a result file.

## 3.4 Thermodynamic uncertainties, SENVAR

If the water composition is well known and the calculations are performed at the same temperature at which the stability constants were measured, the greatest source of uncertainties in the results is in the determination of the different stability and solubility constants.

To illustrate the problem, consider the following simplified description of the dissolution of $U(OH)_4(s)$ in pure water. The reactions describing the system are:

$$U(OH)_4(s) \Leftrightarrow U^{4+} + 4OH^- \quad \text{(R3.1)}$$
$$U^{4+} + OH^- \Leftrightarrow UOH^{3+} \quad \text{(R3.2)}$$
$$U^{4+} + 2OH^- \Leftrightarrow U(OH)_2^{2+} \quad \text{(R3.3)}$$
$$U^{4+} + 3OH^- \Leftrightarrow U(OH)_3^{+} \quad \text{(R3.4)}$$
$$U^{4+} + 4OH^- \Leftrightarrow U(OH)_4 \quad \text{(R3.5)}$$

An equilibrium expression may be written for each of the reactions 1-5, for example:

$$\ln(\beta_2) = \ln(a_{U(OH)_2^{2+}}) - \ln(a_{U^{4+}}) - 2\ln(a_{OH^-}) \quad (3.1)$$

where $\beta_2$ is the equilibrium constant for reaction R3.3 and a is the activity of the species concerned. Since a natural system is far more complicated, a couple of hundred equilibrium equations may be needed. Clearly, a change in an important stability constant may change the result of the calculation significantly. If the system of equations, such as Equation 3.1, is solved, the solubility will be expressed as a sum of $\ln(\beta_i)$. Therefore, according to the central limit theorem, if the parameters in the sum are changed within a uniform interval, the result will be normally distributed. Thus, it is expected that the solubilities will follow a log-normal distribution. Further, it must be noted that other parameters, such as initial pH, pe and temperature, may have a significant effect on the calculations.

The SENVAR package was developed to investigate the effect of uncertainties in thermodynamical data. In its current form, it handles, in addition to the constants mentioned above, the investigation of the effect of uncertainties in enthalpies of reaction, see section 3.5.

### 3.4.1 General features of the program

The SENVAR package is a combination of a statistical sampling and evaluation program. The solubilities are, as for the UNCCON program, calculated by the thermodynamic equilibrium program PHREEQE.

The calculation frame, which is given by the user, is constituted by the choice of solid phase, water composition and database. It is also necessary to give some iteration criteria. The calculations are then made in two steps, first a preliminary sensitivity analysis and then an uncertainty analysis, the results of which are also used for the stepwise regression, which

serves as the final sensitivity analysis. For a more detailed description of the program, see [EKB 95:2] and Paper I (cf. Appendix 6) for flow sheets.

### 3.4.2 Sensitivity analysis

The preliminary sensitivity analysis may be performed in two ways, either by using a variance analysis or by a binary search tree. The main advantage of the variance analysis is that a preliminary ranking of the parameters is made, but the calculations take a great deal of time. The binary search tree, on the other hand, is fast but does not rank the parameters. The concepts of the two approaches are explained below.

#### 3.4.2.1 Variance analysis

The variance analysis is made by holding one of the investigated parameters at a fixed value while the others vary for a given number of iterations, e.g. 20. The variance in solubility for these iterations is then calculated, and the next parameter of interest is held at a fixed value. The species that gives the smallest variance when held constant is deemed the most important and so on.

At the beginning of the calculations, a random matrix is created. It contains random values for the different parameters, each sampled within a given uncertainty range.

There is one row for each investigated parameter. These rows combined forms a matrix with as many rows as there are investigated parameters and as many columns as the selected iteration number. In the first iteration, the first parameter is held at its mean value and the others receive values according to the first column of the matrix. In the next iteration, the values are taken from the second column and so on for the given number of iterations. The second parameter is then held at a fixed value and the others receive values according to the columns of the random matrix. Evidently this approach will give the same variance for the unimportant parameters, thus making the selection criterion simple. The selection criterion is usually that the difference between two successive variances must be at least one thousandth of the last one. The important parameters are then transferred to the uncertainty analysis.

#### 3.4.2.2 Binary search tree

The theory behind the binary search tree is rather simple, and the approach is more commonly used in optimisation problems. The inputs to the model are seen as a vector containing the different parameter values. It is known, *a priori*, that only a few of these are important. Therefore, by using a binary search tree, the number of iterations needed to identify the important parameters may be less than the total number of parameters.

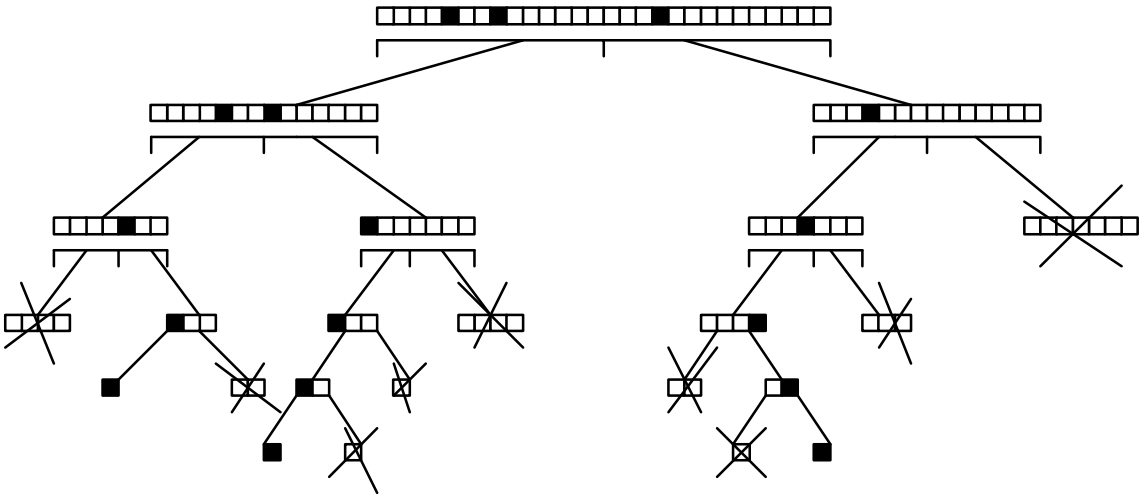


***Figure 3.4:*** *Binary search tree for a 28-cell input vector with three important parameters (filled boxes).*

The method illustrated in Figure 3.4 may be described in the following way. The calculations are made two times, one with every investigated parameter at their maximum value, and one with the minimum value. The results are then compared toinvestigate whether there is a significant difference. If so, the input vector is divided into two parts, each becoming the base for further calculations. The same approach is used at the next level of the search tree, except that at this level it is the values in half of the original vector that are changed while the rest are held fixed. If there is no significant change this time, it is concluded that there are no important parameters in that part of the vector and it is not further investigated. If the change is significant, the new vector is divided into two parts and the method described above is applied to both parts. Finally, all the important parameters are identified. Consider the case in which two iterations are made at each level, i.e. high and low values for the parameters. The number of iterations needed to investigate which parameters influence the result may then be significantly reduced if the method indicated in Figure 3.4 is used. As seen for the case described above, the number of iterations was only slightly reduced compared with the "one factor at a time" approach. However, as the number of parameters grows, the greater the profit will be of using a binary search tree. For the cases described in this section, the number of parameters is usually about two hundred and there are between one and six important ones. Thus the approach is very effective in the sense that it requires only a few iterations, usually reduced to about a third of the number of parameters.

### 3.4.3 Uncertainty analysis

In the uncertainty analysis, the Monte Carlo sampling is made without any restrictions, i.e. the values of the parameters are selected randomly within each interval at each iteration. This approach makes it possible to detect any synergistic effects between two or more species. The calculations are usually made with 1000 to 2000 samples, thus giving enough values to cover the parameters' space sufficiently well. Every solubility calculation is saved in a file which is later used for the final sensitivity analysis.

The results of the uncertainty calculations are a plot showing the calculated density function of the solubilities and some statistics. The calculated statistics are the mean, the variance, the skewness, a 95% confidence interval for the mean based on the solubilities being log-normally distributed and a 95% confidence interval for the solubility population based on an arbitrary distribution, see Appendix in Paper I (cf. Appendix 6). These data may then be used as input to, for example, a transport model.

The final sensitivity analysis is made from the solubility calculations in the uncertainty analysis, thus requiring no further PHREEQE runs. The program that performs these calculations is the STEPR program [LIL 95]. It is assumed that the sensitivity to one parameter is described by the regression coefficients, $a_0$ through $a_n$, in a linear model, such as:

$$Y=a_0+a_1x_1+.....a_nx_n \qquad (3.2)$$

where $x_1$ through $x_n$ represent the input parameters, e.g. the logarithm of the stability constants or the entalpies of reaction, and Y represents the output result, i.e. solubilities. Such an approach yields very small regression errors in the cases used here and may therefore be usable.

### 3.4.4 Program description

The directory structure is, as mentioned before, similar in all the program packages described in this reprt. The directory structure of the SENVAR package is illustrated below.

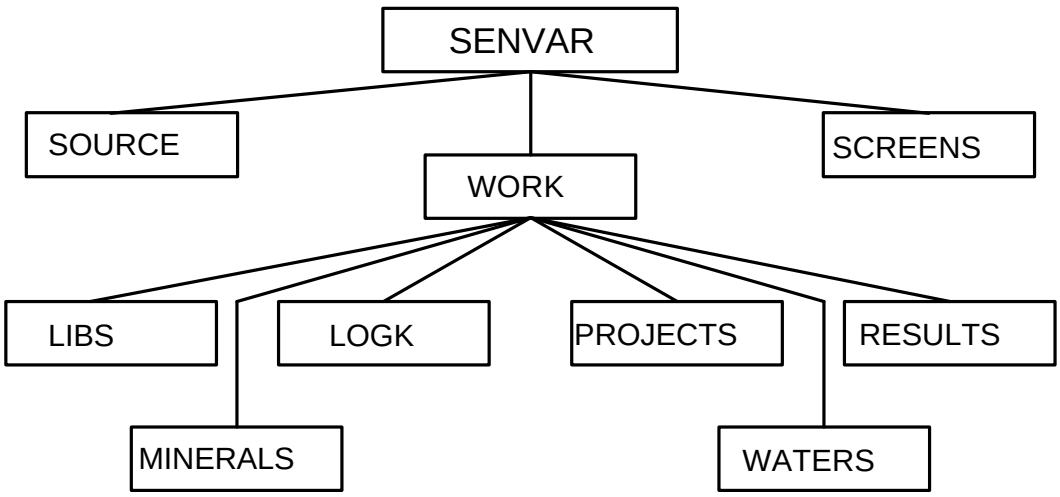



***Figure 3.5:*** *Directory structure of the SENVAR package.*

The contents of the directories are as follows:

SOURCE: source codes of the programs together with include files

WORK: executable programs and temporary result files

LIBS: databases

MINERALS: available solid phases (New ones may be copied from other directories or databases.)

LOGK: lists of the mean value of the stability constants in the selected database

PROJECTS: description files for the different projects

WATERS: files containing the necessary information on the water used in the project and executable infiles for PHREEQE

RESULTS: files containing the calculated solubilities in the sensitivity analysis, result files from the sensitivity analysis, calculated solubilities and matrix files from the uncertainty analysis and results from the stepwise regression.

To visualise the function of the programs, the directory structure and program description are accompanied by a hierarchical diagram for the program package, see Figure 3.6. It should be noted that a program called HELPIT [EMR 92] is also present in the package but, as with STEPR and PHREEQE, it is not described in this report.

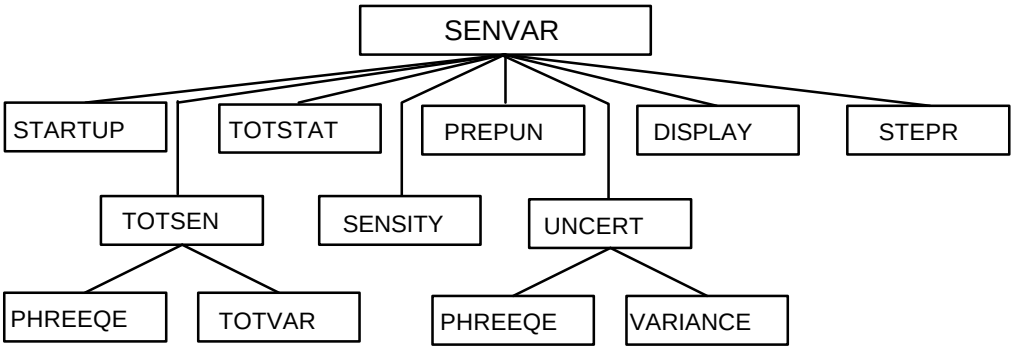



***Figure 3.6:**. Hierarchical diagram of the SENVAR package.*

The general function of each program is listed below, and a more detailed description and the source code may be found in [EKB 95:2] and [EKB 95:3], respectivly.

SENVAR: The master program which controls and distributes tasks to the other programs. Preparations for the other programs are also made by this program.

STARTUP: This is the program in which the user may design every run separately.

Several choices are made from different menus, thus giving the essential information for the calculation programs to work.

TOTSEN: The main program in the sensitivity analysis based on variance analysis, which handles the results from PHREEQE and extracts data from this to make a result file, which is used by the evaluation program, TOTSTAT.

PHREEQE: A thermodynamic equilibrium code used to calculate the speciation in a water when equilibrated with a solid phase.

TOTVAR: This program is called one time each iteration, in the variance analysis, to change the stability constants and a solubility constant in the database for PHREEQE.

TOTSTAT: The evaluation of the variance analysis is done by this program. The main feature is that the names of the most important species are transferred to a separate file for later use in the uncertainty analysis.

SENSITY: This is the program which performs and analyses the sensitivity analysis based on the binary search tree approach. The results are, as for TOTSTAT, presented as a file for use in the uncertainty analysis.

PREPUN: The indata file to UNCERT is made from the results of the sensitivity analysis. Each of the interesting species together with the water parameters are given an uncertainty interval. This may be changed by the user at a later stage.

UNCERT: The main program in the uncertainty calculations which handles the results from PHREEQE and extracts data from this to make a result file, which is used by the evaluation program, DISPLAY.

VARIANCE: This program is used to change the PHREEQE database, according to some specified criteria, i.e. which parameters are to be changed and to what extent.

DISPLAY: The distribution functions for the different solubilities are shown by this program. In addition, some useful statistical estimators are also calculated.

STEPR: STEPR calculates how some of the different stability constants and pH, pe and temperature are related to each other in the sense of their initial uncertainties'

contribution to the uncertainty in the solubility. This is done by stepwise linear regression.

HELPIT: A program invoked when there are convergence problems in the PHREEQE calculations. The main function is to include the "KNOBS" keyword in the PHREEQE infile and run PHREEQE one more time.

## 3.5 Enthalpies of reaction

When an attempt is made to simulate the reactions in groundwater in the vicinity of a repository for spent nuclear fuel, it is important to remember that the temperature will not be the same as in the rest of the rock owing to the heat development in the fuel. Therefore, if solubilities are to be calculated, for example, the model discussed above is not sufficient. It may be necessary to assume that the stability constants in the database are correct at the temperature of measurement and concentrate efforts on the enthalpies of reactions instead. The correlation between the stability constant and the enthalpy of reaction is given by Equation 3.3, which is formally derived in Appendix 5.

$$\left(\frac{\partial \ln(b)}{\partial \frac{1}{T}}\right)_P = \frac{\Delta H}{R} \qquad (3.3)$$

where $\beta$ is the stability constant, T the temperature, R the molar gas constant and $\Delta H$ the enthalpy change of the reaction concerned. Clearly, there is a linear dependence of $\Delta H$ on the $\ln(\beta)$ if the temperature dependence of the enthalpy and the entropy are neglected. It is thus expected that the solubilities will be log-normally distributed for the same reason as explained above. If the enthalpy of the reaction is encumbered with some uncertainty, the new value of $\beta$ will be wrong and, consequently, the calculated solubility will also be wrong. It is therefore important to know how influential the different enthalpies are on the calculated solubility in order to ascertain that the important species have a well determined $\Delta H$ value.

### 3.5.1 Distinctive features of the model

The calculations of the effect of uncertainties in enthalpies of reaction are made with a modified version of the SENVAR package. The assumptions and statistical calculations are as described above for the first version of SENVAR.

# 4. Results of uncertainty assessment in solubility calculations

This section deals with the results of the different programs developed for investigation of the effect of different input uncertainties on solubility calculations. The results are presented in sections named after the program used to produce the result. As previously stated these are: MINVAR investigating the effect of rock composition uncertainties on the groundwater composition, UNCCON calculating the effect of water composition variations on the solubility of a solid phase, and SENVAR estimating the effect of uncertainties in thermodynamic entities, such as stability constants and enthalpies of reaction, on the calculated solubility of a solid phase. In addition, the effect of some conceptual uncertainties are discussed and examplified.

The aim of the calculations shown here was to use approximately the same conditions in the different calculations, e.g. the same water and databases. The data are usually taken from the Äspö site in the southeast of Sweden, thus the results are not general. Specific results for any site may be obtained by complementary calculations by anyone interested since it is possible to acquire the computer programs used and documentations from the author of this report. Further, the results are divided into sections depending on the program that produced them. After each section, comments are given on the results. All figures presenting either cumulative distribution functions (cdf) or probability density functions (pdf) of the investigated property also include a line representing a normal distribution.

In addition, - or + is included in the figures, for the SENVAR cases, to indicate whether the Pearson test, see Appendix 4, rejected the hypothesis that the sample points are normally distributed with 95% confidence. However, although cases may be rejected as normally distributed, the conclusion drawn for cases close to normality still holds.

## 4.1 MINVAR

The results of the MINVAR package are mainly a source for obtaining a rough estimate of the uncertainties in water composition, thus serving as input to UNCCON. The results are a 95% confidence interval for the concentration of the different elements in the water, together with pH and pe. Further, the calculated frequency function together with the cumulative distribution function may be displayed. However, since the sampling was based on LHS, the plots look somewhat jagged owing to the iteration time for each calculation step and thus the small sample size used. Despite this, they may give a rough estimate of the distribution.

### 4.1.1 Calculation settings

The calculation settings follow as close as possible the idea of this report, i.e. all calculations should be made with assumptions and settings as close as possible to the conditions at the Äspö site in Sweden. The CRACKER calculations were made for a fracture 20 diffusion cells wide and 100 diffusion cells long, see section 4.1.2, with the mineral composition shown in Table 4.1.

***Table 4.1:*** *Estimated mineral composition at Äspö together with assumed uncertainty intervals [TUL 95].*

| Mineral | Abundance, (%) |
|---|---|
| Chlorite | 40±10 |
| Calcite | 25±10 |
| Epidote | 20±10 |
| Fluorite | 10±5 |
| Hematite | 5±5 |
| Quartz | 2±2 |
| Illite | 2±2 |
| Montmorillionite | 2±2 |
| Pyrite | 2±2 |

The water used serves as a typical groundwater from the Äspö site, according to Table 4.2. This water is used as input data to the CRACKER simulations. If the resulting water composition is in fair agreement with this water it is assumed that the rock composition used is representative for the inverstigated site. Naturally, it is also possible to use a very well known rock composition and starting the calculations with pure water. This approach has not been used here but is described elsewhere [EMR 98:2].

***Table 4.2:*** *Composition of reference water from the Äspö site, borehole KAS 02, level 530-535m [SME 92, NIL 92]. Concentrations in mM.*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Ca | 47.2 | | Cl | 181 | | U[a] | 5.45E-07 |
| Mg | 1.73 | | $C_{tot}$ | 0.164 | | Sr | 0.399 |
| Na | 91.3 | | $S_{tot}$ | 5.83 | | Li | 0.144 |
| K | 0.207 | | F | 7.90E-02 | | $N_{tot}$ | 3.52E-03 |
| Fe[a] | 4.37E-03 | | Br | 0.501 | | pH | 8.1 |
| Mn | 5.28E-03 | | P | 1.61E-04 | | pe[b] | -4.37 |
| Al[a] | 1.00E-03 | | Si | 0.146 | | Temp (°C) | 15 |

[a] The analysis did not contain any value for this species. The concentration was estimated on the basis of other Äspö ground water samples [EMR 95].

[b] The pe value was adjusted from the measured value (-5.42) with regard to additions of Al, Fe, U, and the equilibrium between $SO_4^{2-}/SO_3^{2-}$ in the solution [EMR 95].

Clearly, the mineral composition given in Table 4.1 does not account for many of the features of the selected water, e.g. uranium content, but the water seems to be fairly stable in contact with the mean rock composition, i.e. the simulated water composition agrees reasonably well with the experimentally determined composition. This indicates that the selected set is at least a possible representation of the minerals presently in contact with the selected water.

### 4.1.2 Results

The results of a MINVAR calculation primarily give confidence intervals for the concentrations of the different elements in the selected water together with a cumulative distribution function for each element. In this report, the results presented are limited to those elements that showed a significant dependence on the rock composition.

***Table 4.3:*** *Minimum and maximum concentrations together with 95% confidence intervals for the mean concentrations.*

| Parameter | Min. value | Max. value | Confidence interval for the mean (95%) |
|---|---|---|---|
| pH | 8.0 | 8.8 | 8.6±0.05 |
| pe | -5.2 | -4.3 | 5.0±0.06 |
| Al | 8.2E-07 | 5.2E-06 | 2.0E-06±4.6E-07 |
| Fe | 3.4E-09 | 1.0E-08 | 8.7E-09±3.5E-09 |
| Si | 6.1E-05 | 1.4E-04 | 8.1E-05±8.8E-06 |
| C | 3.2E-05 | 1.2E-04 | 4.5E-05±4.5E-06 |

The confidence intervals presented in Table 4.3 are calculated under the assumption that the element concentrations are normally distributed. If the concentrations presented in Table 4.3 are compared with the measured values from Äspö, Table 4.2, it is clear that the mineral composition used was not entirely correct. This fact raises the question of whether to believe in experimentally determined or calculated mineral abundancies. As it is doubtful whether the model used to calculate the concentrations is completely verified and validated it is not possible to make such a conclusion at this stage.

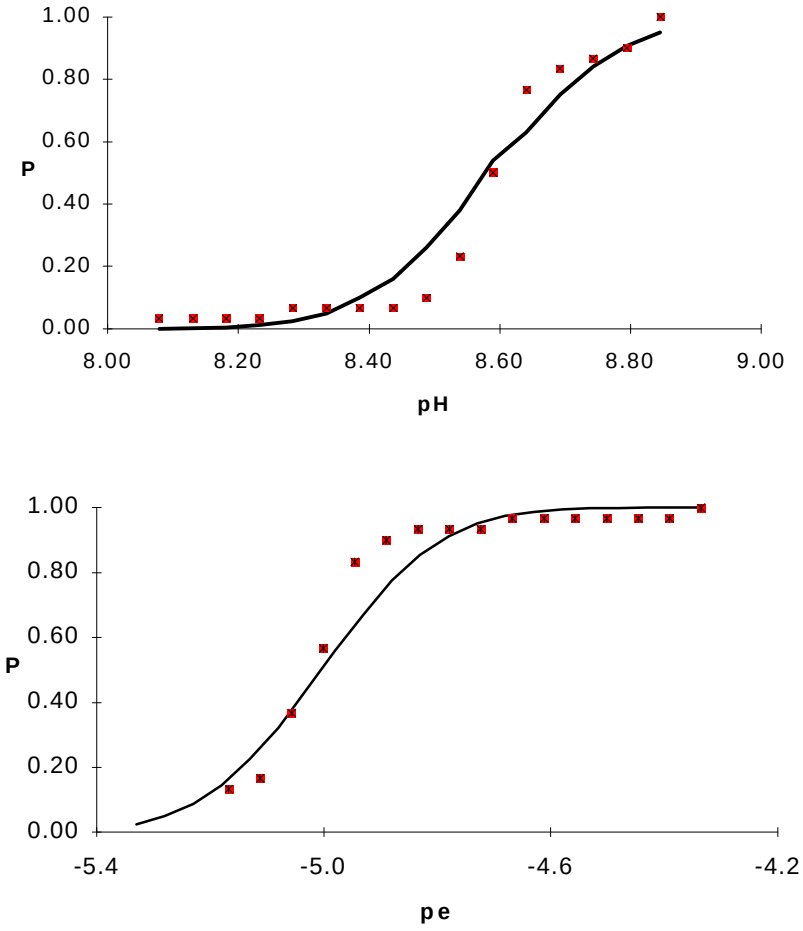



***Figure 4.1:*** *Cumulative distribution functions for pH and pe. The x-values are the sampled results and the solid line is the cdf for a log-normal distribution*

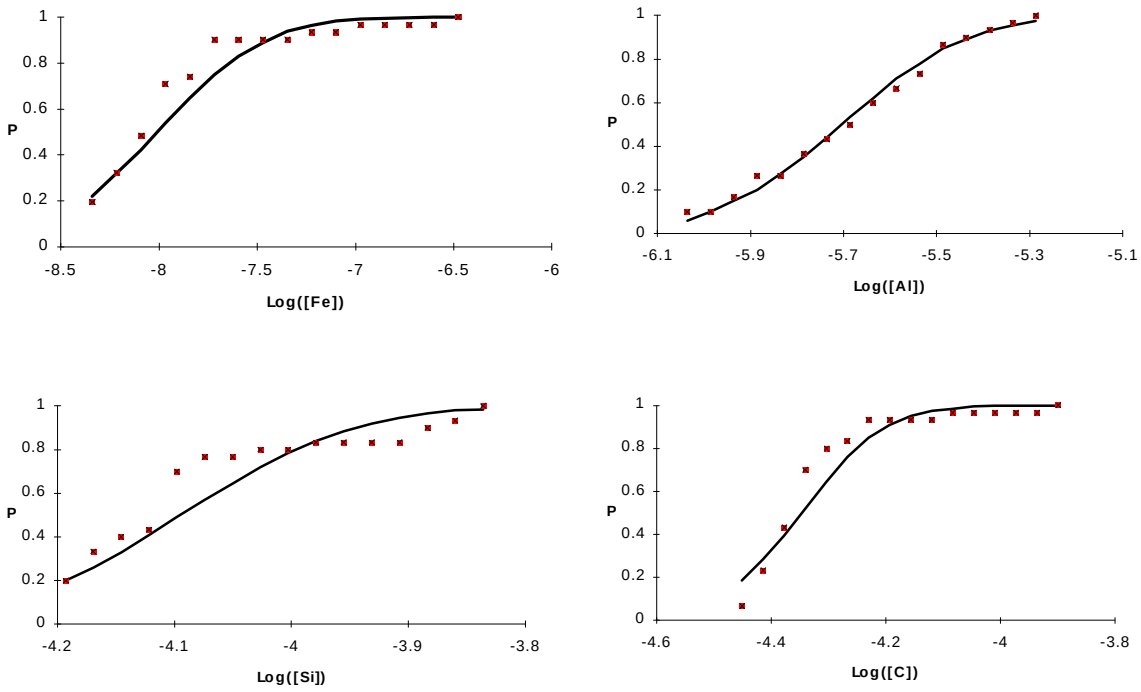



***Figure 4.2:*** *Cumulative distribution functions for some important elements. The x values are the sampled results and the solid line is the cdf for a log-normal distribution.*

The solid line in figures 4.1 and 4.2 is the cumulative distribution function (cdf) for a log-normal distribution with the calculated mean logarithm of the concentration (μ) and standard deviation (σ), according to:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \qquad (4.1)$$

In order not to evaluate this integral, and instead use a table of values, x must be normally distributed with zero mean and a variance of one, N(0,1), i.e. x will assume values according to:

$$x = \frac{y - m}{s} \qquad (4.2)$$

where y is the logarithm of the concentration values, i.e. the x-axes in Figures 4.1 and 4.2, and μ and σ are the mean of the logarithm of the concentration and the standard deviation, respectively. This will transform Equation 4.1 into:

$$P(x) = \frac{1}{\sqrt{2p}} \int_{-\infty}^{x} e^{-t^2/2} dt \qquad (4.3)$$

which may be evaluated using a standard table of values.

Clearly, the log-normal assumption made earlier does not hold for a critical examination since the tails in the distributions are often missing at one of the end points. However, the agreement in the remaining region may be considered sufficiently good. The deviation from normality does not affect the UNCCON calculations because, to be somewhat conservative, the distributions assumed in that program are uniform with the minimum and maximum values from MINVAR as endpoints for the interval.

It is interesting to see from Table 4.3 that the water composition does not change a great deal when the mineralogic composition of the rock is changed. However, in these calculations, the only mechanisms for changing the water composition are mineral dissolution or precipitation. In nature, such factors as sorption and matrix diffusion may play an important role. Thus the real concentration intervals may be significantly larger when these factors, together with measurement errors and changes in the water from the sampling point to the laboratory, are taken into account.

## 4.2 UNCCON

The results presented here were derived from two different cases. In the first case, the Äspö water from Table 4.2 was used with an uncertainty interval originating from measurement errors and random fluctuations [SAM 96]. The other water was from the MINVAR calculations. The values shown in Table 4.3 were used, with the minimum and maximum values used as limits for a uniform interval. The fact that almost all of them might be log-normally distributed has not been taken into account in order to produce a conservative result. The calculations were made with a sample size of fifty points.

The selection of solid phases in this study may seem arbitrary since not all of them are stable under the conditions used, e.g. $UO_3 \cdot 2H_2O$ which is a U(VI) mineral. In addition the selection of amorphous or crystalline phases is not clear since what probably exist in nature is a mixture, making simulation difficult. However, these minerals have been suggested to be among the solubility limiting one close to the repository [PRO 90]. Although the choice of solubility limiting phase is a great source of uncertainties in real solubility estimates, this question is not the scope of this report. Rather, this work is focused on the calculations themselves.

### 4.2.1 Results

The results of the UNCCON program are plots of the cumulative distribution function for the calculated solubility, the calculated distribution function, the maximum and minimum solubility and a confidence interval for the mean solubility. This interval assumes that the solubilities are log-normally distributed. According to Figures 4.3 and 4.4, such an assumption seems fairly good for some of the elements. The results of the two different cases are presented in parallel to make comparison convenient.

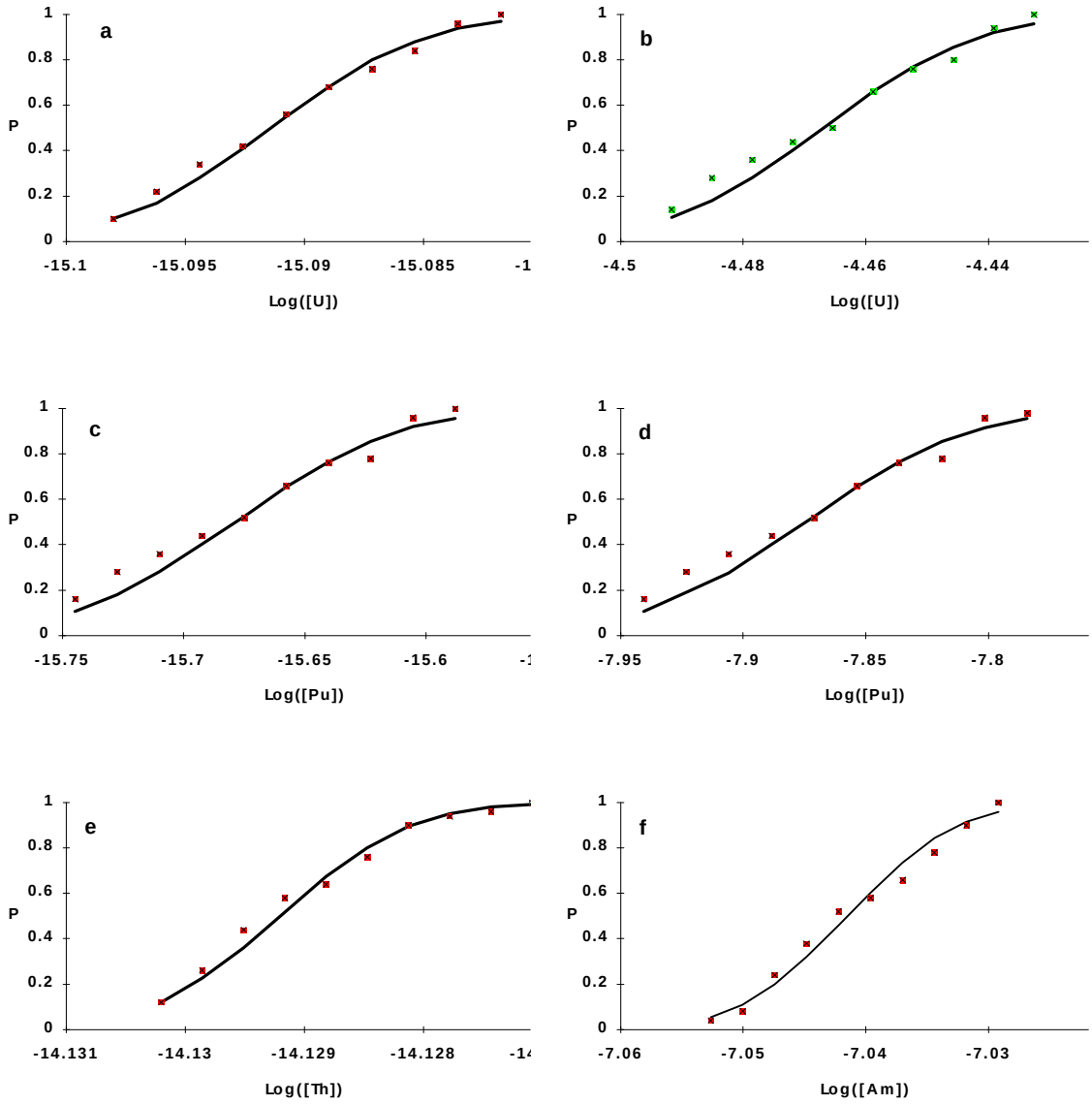



***Figure 4.3:*** *Äspö case, cumulative distribution function for the solubility of: a) $UO_{2(c)}$, b) $UO_3$¥$2H_2O$, c) $PuO_2$, d) $Pu(OH)_4$, e) $ThO_{2(c)}$, f) $Am_2(CO_3)_3$. The x values are the sampled results and the solid line is the cdf for a log-normal distribution*

The log-normal assumption seems to be better in the case of the water from Äspö than for the MINVAR case. Unfortunately, the uncertainty intervals are not generally equally large, and thus it is not possible to assume a log-normal distribution for the solubilities with respect to uncertainties in groundwater composition.

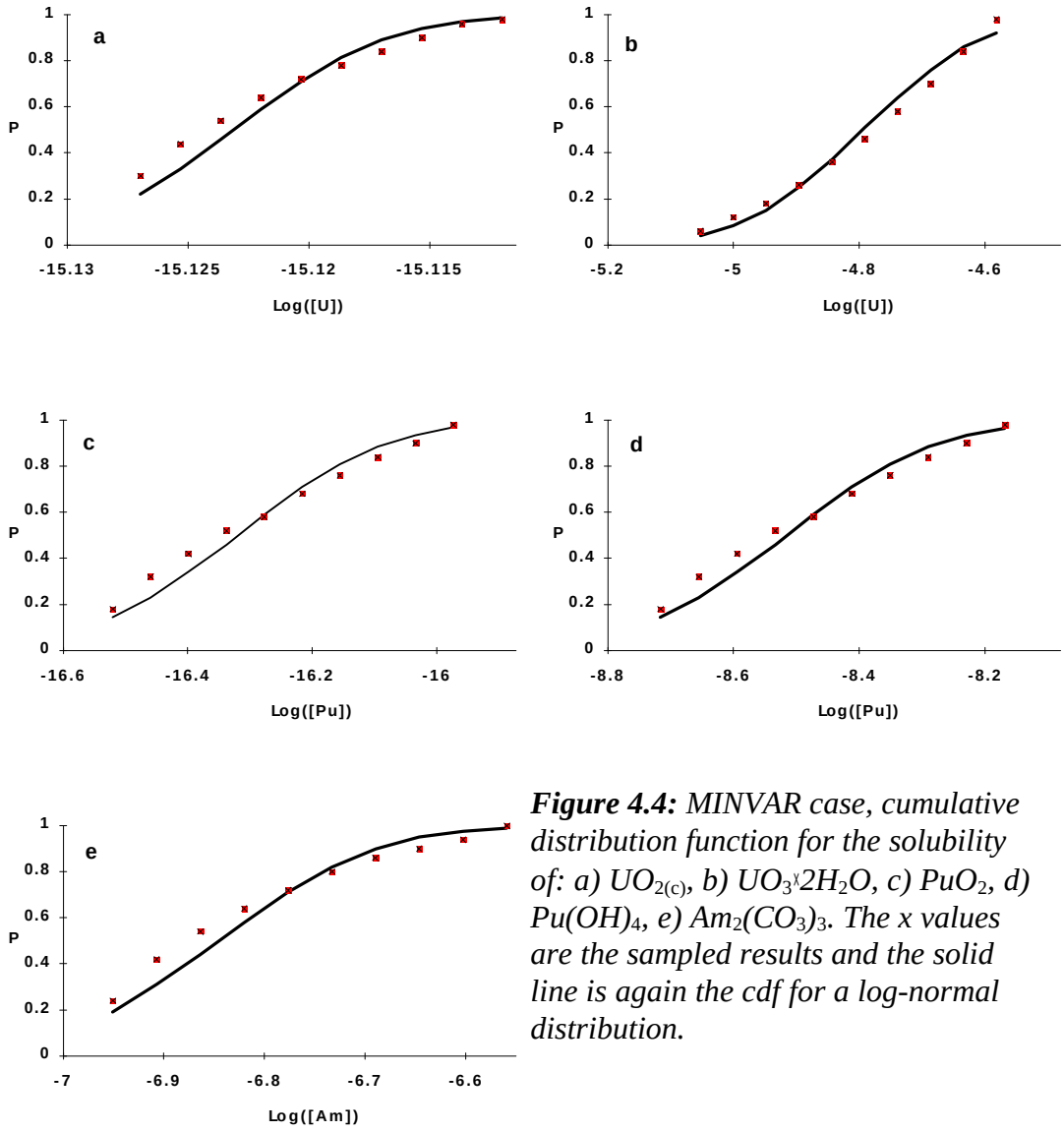



***Figure 4.4:*** *MINVAR case, cumulative distribution function for the solubility of: a) $UO_{2(c)}$, b) $UO_3 \cdot 2H_2O$, c) $PuO_2$, d) $Pu(OH)_4$, e) $Am_2(CO_3)_3$. The x values are the sampled results and the solid line is again the cdf for a log-normal distribution.*

The values in Table 4.4 are not entirely the same as in Paper V (cf. Appendix 6) because the calculations presented in that paper were made with an erroneous database. The errors have been corrected and the calculations remade, and some may be found in [EKB 98].

***Table 4.4****: Statistics for the case with Äspö water*

| Solid phase | Confidence interval for the mean solubility (M), 95% | Min ; Max solubility (M) |
|---|---|---|
| $UO_3 \cdot 2H_2O$ | 3.40E-05±3.1E-06 | 3.18E-05 ; 3.69E-05 |
| $UO_2$(c) | 8.10E-16±1.9E-17 | 7.95E-16 ; 8.28E-16 |
| $PuO_2$(c) | 2.10E-16±5.3E-17 | 1.73E-16 ; 2.58E-16 |
| $Pu(OH)_4$ | 1.33E-08±3.4E-09 | 7.09E-09 ; 1.98E-08 |
| $Am_2(CO_3)_3$ | 9.09E-08±2.9E-09 | 8.50E-08 ; 1.06E-08 |

***Table 4.5:*** *Statistics for the case with MINVAR water*

| Solid phase | Confidence interval for the mean solubility (M), 95% | Min ; Max solubility (M) |
|---|---|---|
| $UO_3 \cdot 2H_2O$ | 1.61E-05±1.3E-05 | 7.85E-06 ; 2.62E-05 |
| $UO_2(c)$ | 7.53E-16±1.7E-17 | 7.44E-16 ; 7.73E-16 |
| $PuO_2(c)^a$ | 4.78E-17±3.3E-17 | 2.62E-17 ; 1.06E-16 |
| $Pu(OH)_4{}^a$ | 3.05E-09±2.1E-09 | 2.54E-09 ; 8.90E-09 |
| $Am_2(CO_3)_3$ | 1.43E-07±9.1E-08 | 2.00E-07 ; 2.65E-07 |

a, Due to the large standard deviation the interval is in this case 80% of the mean value

The solubility is not generally higher or lower in any of the cases, but the confidence interval for the mean is larger for all of the phases presented in the MINVAR case except for $UO_2(c)$. The reason for this may be that the dissolution of U(IV) involves elements which are not varied in this case, and thus the change in uranium solubility will not be large.

## 4.3 SENVAR

The calculations presented in this section deals with the effect of uncertainties in thermodynamic data such as stability constants and enthalpies of reaction. The results are based on Papers II and III (cf. Appendix 6). The water used is the same as for the MINVAR calculations, see Table 4.2. The same water composition was used for the high temperature case, but elevated to 80°C. The preliminary sensitivity analysis was made with the variance analysis approach with a width of the uncertainty interval for the stability constants of 4 log units in the room temperature calculations. In the 80°C case, the preliminary sensitivity analysis was made with the binary search tree. The uncertainty interval may seem arbitrary but the actual lwidth of the interval is not of greater importance since it is only the internal sensitivity analysis which is based on the results. The order of importance will not change if another uncertainty interval is selected within a reasonable interval. The pH, pe and temperature were also included in both cases, with an uncertainty interval of one unit. In the uncertainty analysis for the room temperature case, the interval was one log unit wide for both the stability constants and the water parameters, i.e. pH, pe, and temperature. This value was derived from Gibbs energy of formation as given by NEA [NEA 92]. For the case with elevated temperature, the enthalpies of reaction for the formation of the important species were changed within a four kcal/ mole wide interval due to a lack of relevant data. The sample size was in both cases 1000, derived from uniform distributions of the parameters. It may seem unrealistic to select the same, somewhat arbitrary, uncertainty limits for all species, but it is very difficult to find reasonable values in the literature, except in the book mentioned above, which in turn is a motivation for making uncertainty analyses. The values in one study are often given with some rather small error limits but when compared to the results of another study the first results may be very different from the interval given in the second study. This is probably a consequence of different experimental techniques and, hence, must widen the uncertainty interval significantly.

However, papers reviewing other studies are not easy to find, and more effort should thus be made to produce consistent and well determined uncertainty values.

### 4.3.1 Room temperature case

As mentioned earlier, the results of the uncertainty analysis using the SENVAR package are a plot of the calculated frequency functions and some statistics, such as mean solubility and different confidence intervals, both for the mean and the solubility populations. As seen in Figures 4.5 and 4.6, almost every calculation case shows the expected log-normal distribution or is at least close to such a distribution by visual inspection.

The solid lines in Figures 4.5 and 4.6 represent a log-normal distribution based on the mean and the variance obtained from the sample calculations. These values are also used to calculate the confidence intervals for the mean solubility given in Table 4.6. In some of the cases, it may not be possible, on good grounds, to assume a log-normal distribution of the solubilities, and the confidence interval for the solubility population should therefore be used instead, see Table 4.7. The intervals presented in Tables 4.6 and 4.7 are calculated for the logarithms and thus only approximated to be symmetrical in the linear scale.

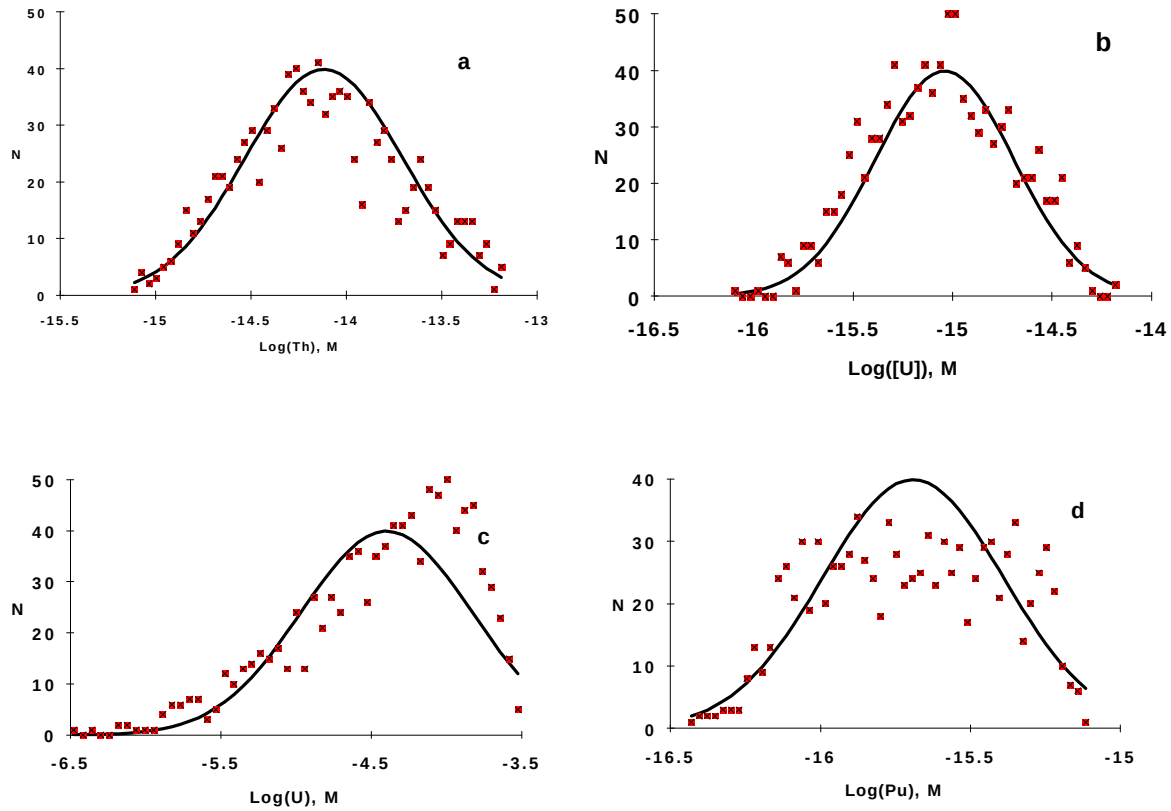



***Figure 4.5:*** *Calculated density functions for: a) $ThO_{2(c)}$, b) $UO_2$, c) $UO_3 \cdot 2H_2O$, d) $PuO_2$. The x values are the sampled results; the solid line is the pdf for a log-normal distribution.*

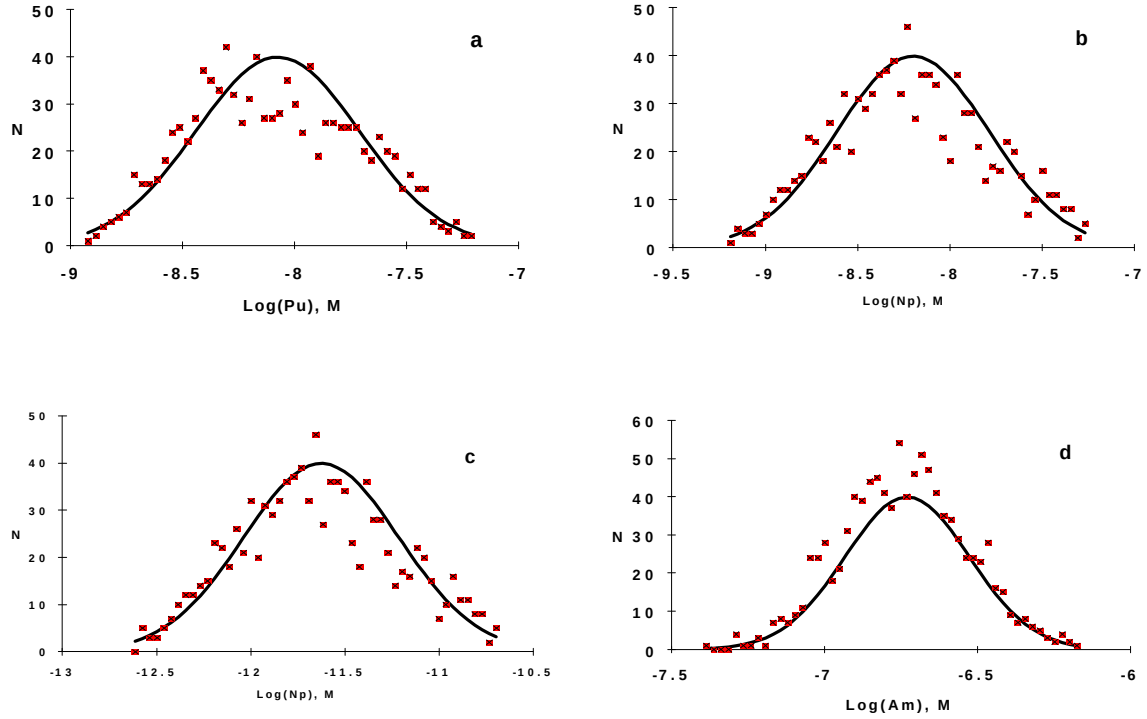



***Figure 4.6:*** *Calculated density functions for: a) $Pu(OH)_4$, , b) $Np(OH)_4$, c) $NpO_2$, d) $Am_2(CO_3)_3$. The x values are the sampled results; the solid line is the pdf for a log-normal distribution.*

***Table 4.6, Statistical results of the uncertainty analysis***

| Solid Phase | Mean solubility (M) | Confidence interval for the mean solubility (M), 95% | Min;Max solubility (M). |
|---|---|---|---|
| $ThO_2(c)$ | 7.61E-15 | 7.61E-15±9.84E-16 | 8.44E-16 ; 7.09E-14 |
| $UO_2(c)$ | 9.05E-16 | 9.05E-16±9.11E-17 | 8.78E-17 ; 7.17E-15 |
| $UO_3 \cdot 2H_2O$ | 4. 01E-05 | 4. 01E-05±8.73E-06 | 3.83E-07 ; 3.43E-04 |
| $PuO_2(c)$ | 2.04E-16 | 2.04E-16±1.66E-17 | 3.95E-17 ; 8.14E-16 |
| $Pu(OH)_4$ | 8.40E-09 | 8.40E-09±9.02E-10 | 1.30E-09 ; 6.66E-08 |
| $NpO_2(c)$ | 2.36E-12 | 2.36E-12±3.06E-13 | 2.65E-13 ; 2.20E-11 |
| $Np(OH)_4$ | 6.31E-09 | 6.31E-09±8.16E-10 | 7.08E-10 ; 5.87E-08 |
| $Am_2(CO_3)_3$ | 1.87E-07 | 1.87E-07±9.09E-09 | 4.35E-08 ; 7.08E-07 |

The formula used for the calculation of the intervals presented in Table 4.7 make use of the Chebychev inequality and can be found in Paper II.

***Table 4.7, 95% confidence interval for the solubility population***

| Solid phase | Confidence interval (M), 95% |
|---|---|
| $PuO_2(c)$ | 2.04E-16±9.2E-16 |
| $UO_3(c)$ | 4.01E-05±4.8E-04 |

The uncertainty intervals for the solubility population, as given in Table 4.7, may be seen only as an upper bound for the solubility. The reason for presenting the interval is to give some information on what happens when it is not possible to assign a log-normal distribution to a sample but when an arbitrary distribution must be used instead. The results shown in Table 4.6 and Figures 4.5 and 4.6 are sufficient for use in further calculations in which the solubility of a solid phase is needed as input. However, to understand the system and determine which parameters should be investigated more thoroughly to minimise the uncertainty in the stability constants, a sensitivity analysis is needed. Table 4.8 lists the species which stability constants and other parameters that were found to have the largest influence on the results of the solubility calculations. Since all parameters are given equally large uncertainty intervals, the regression coefficients obtained may be used directly to determine the internal ranking.

***Table 4.8****: The species which stability constants and the other parameters that were selected by the sensitivity analysis.*

| Solid phase | Important species (stability constants) or property | Internal ranking (Reg. coeff.) |
|---|---|---|
| $ThO_2$(c) | $Th(OH)_4$, $ThO_2$(*c*) | 2.30, 2.30 |
| $UO_2$(c) | $UO_2$(c), $U(OH)_4$, pe, $HCO_3^-$ | 2.30, 1.10, 1.02, 0.31 |
| $UO_3 \cdot 2H_2O$ | $UO_3 \cdot 2H_2O$, $UO_2^{2+}$, pe<br>$HCO_3^-$, $(UO_2)_2CO_3OH^-$ | 2.17, -2.17, -2.15<br>-1.03, 1.02 |
| $PuO_2$(c) | $PuO_2$(c), pe, $Pu(OH)^{2+}$<br>$NaHCO_3$ | 2.31, 0.36, 0.20<br>-0.19 |
| $Pu(OH)_4$(s) | $Pu(OH)_4$(s), $Pu(CO_3)_3^{2-}$<br>$Pu(OH)^{2+}$, $NaHCO_3$ | 2.26, 1.50<br>0.55, -0.21 |
| $NpO_2$(c) | $NpO_2$(c), $Np(OH)_4$ | 2.31, 2.30 |
| $Np(OH)_4$(s) | $Np(OH)_4$(s), $Np(OH)_4$ | 2.31, 2.30 |
| $Am_2(CO_3)_3$(s) | $Am_2(CO_3)_3$(s), pe, $CaCO_3$<br>$Am(CO_3)_2^-$ | 1.13, 0.72, 0.62<br>0.58 |

One method for investigating the stability of the approach adopted for sensitivity determination is to compare the results from the stepwise regression with the preliminary sensitivity analysis. When the variance analysis method is used, this is a simple task as the parameters are then also ranked, see Table 4.9.

***Table 4.9:*** *Results of the preliminary sensitivity analysis.*

| Solid phase | Important species (stability constants) or properties (in order of importance) |
|---|---|
| $ThO_2(c)$ | $Th(OH)_4$, $ThO_2(c)$ |
| $UO_2(c)$ | $UO_2(c)$, $HCO_3^-$, $OH^-$ |
| $UO_3 \cdot 2H_2O$ | $(UO_2)_2CO_3OH^-$, $UO_2^{2+}$, $UO_3 \cdot 2H_2O$ |
| $PuO_2(c)$ | $PuO_2(c)$, pe, $Pu(OH)^{2+}$ |
| $Pu(OH)_4(s)$ | $Pu(CO_3)_3^{2-}$, $Pu(OH)_4(s)$ |
| $NpO_2(c)$ | $NpO_2(c)$, $Np(OH)_4$ |
| $Np(OH)_4(s)$ | $Np(OH)_4(s)$, $Np(OH)_4$ |
| $Am_2(CO_3)_3(s)$ | $Am_2(CO_3)_3(s)$, $Am(CO_3)_2^-$, $AmCl^{2+}$ |

The parameters included in Table 4.9 are only the ones that gave rise to a distinctively reduced variance when held fixed. This selection is therefore smaller than in the final sensitivity analysis, see Table 4.8. However, it is clear that there are no significant differences between the different calculation methods. It is thus possible to conclude that there has been no fatal error in the calculations.

The frequency functions may also be plotted in the same figure, thus giving a rough estimate of the relative uncertainty intervals.

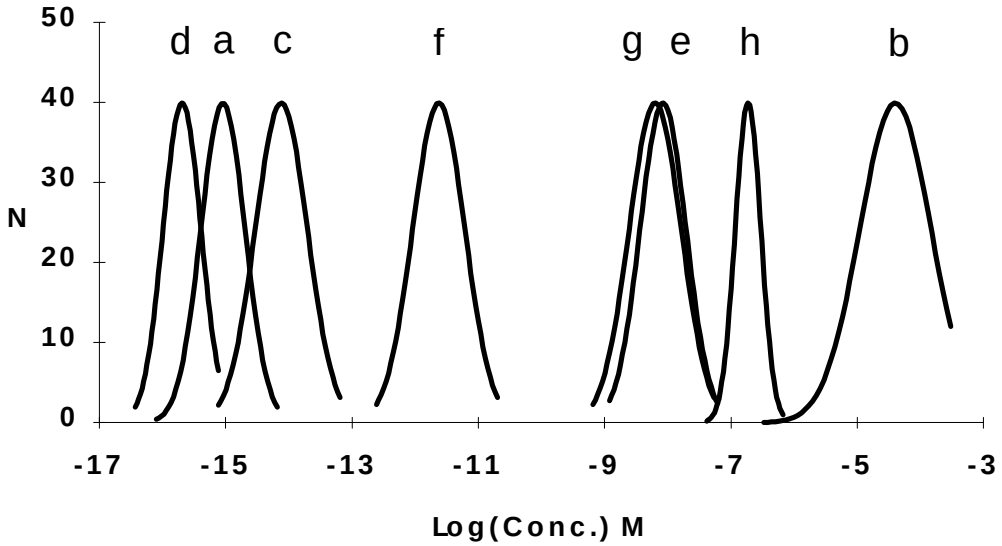



***Figure 4.7****: Plot of the solubility distributions on the same scale for: (a) $UO_{2(c)}$,(b) $UO_3 \cdot 2H_2O$, (c) $PuO_2$, (d) $Pu(OH)_4$, (e) $ThO_{2(c)}$, (f) $NpO_2$ , (g) $Np(OH)_4$, (h) $Am_2(CO_3)_3$.*

It is clearly seen in Figure 4.7 that the solubility intervals are approximately equally large, except for $Am_2(CO_3)_3$ and $UO_3 \cdot 2H_2O$. In the other cases, the width of the log(solubility) interval is approximately 1.9, i.e. the difference between the highest and lowest solubility obtained in the calculations is 1.9 log units. The same behaviour is noted when the uncertainty intervals for the important species are doubled, i.e. the interval for the logarithm of the

solubility is doubled. If this were generally true, the propagation of errors in stability constants would be fairly easy to estimate under the assumption that all parameters have equal uncertainty. Unfortunately, as seen above, this is not generally the case.

### 4.3.2 80°C case

Since the calculations made here are performed by the same program as was used for the room temperature case, the format of the output is the same. However, it should be noted that, in this case, it is the enthalpies of formation which are treated as uncertain parameters rather than the stability constants themselves. The water parameters, pH, pe and temperature are also included.

The + or - at the upper left part of each picture indicate whether the Pearson chi-square test rejected (-) the hypothesis that the sampled solubilities actually follow a normal distribution.

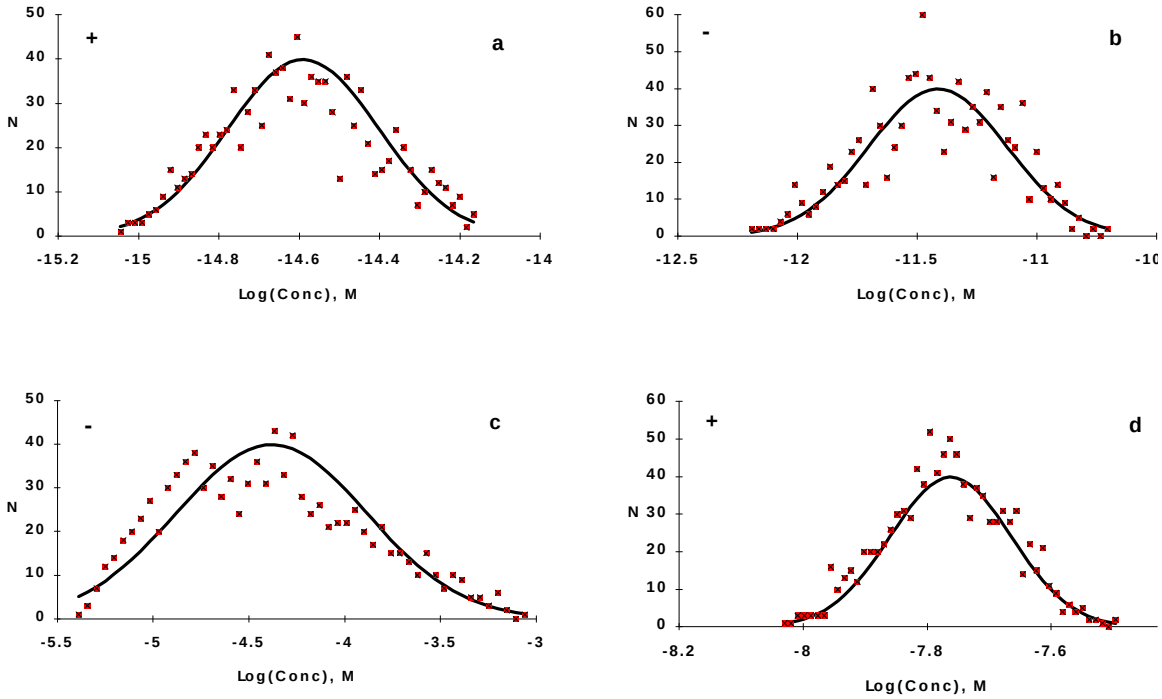



***Figure 4.8:*** *Calculated density functions for the 80°C case for: a) $ThO_2$, b) $UO_{2(c)}$, c) $UO_3 \cdot 2H_2O$, d) $PuO_2$. The x values are the sampled result; the solid line is the pdf for a log-normal distribution. The sign in the upper left corner indicates whether a goodness of fit test rejected (-) or not (+).*

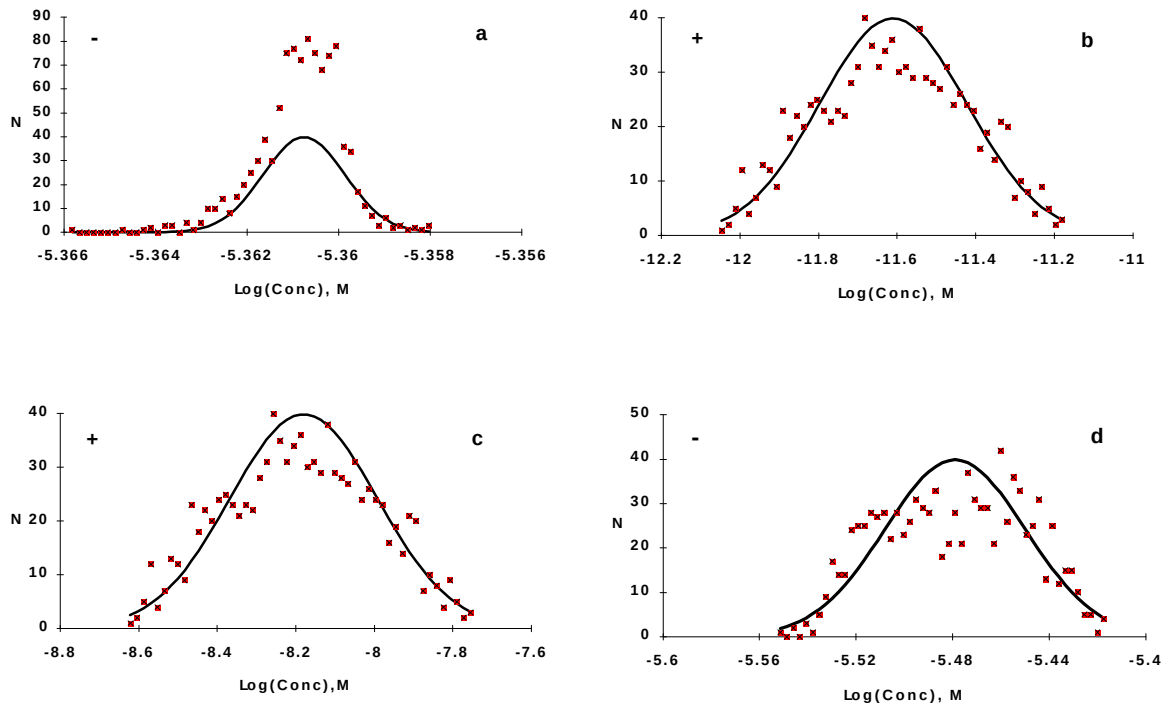



***Figure 4.9:*** *Calculated density functions for the 80°C case for: a) $Pu(OH)_4$, b) $NpO_2$, c) $Np(OH)_4$, d) $Am_2(CO_3)_3$. The x values are the sampled results; the solid line is the pdf for a log-normal distribution. The sign in the upper left corner indicates whether a goodness of fit test rejected (-) or not (+).*

The assumption that the solubilities are log-normally distributed is more correct in this case, as seen in Figures 4.8 and 4.9, than in the case in which the stability constants were varied, except for $Pu(OH)_4$. Table 4.10 also shows that the width of the solubility interval for this phase differs significantly from the other phases investigated.

***Table 4.10****: Statistical results from the uncertainty analysis.*

| Solid Phase | Mean solubility (M) | Confidence interval for the mean solubility (M), 95% | Min;Max solubility (M) |
|---|---|---|---|
| $ThO_2(c)$ | 2.57E-15 | 2.57E-15±1.14E-16 | 9.38E-16 ; 7.12E-15 |
| $UO_3 \cdot 2H_2O$ | 4.15E-05 | 4.15E-05±7.24E-06 | 4.54E-06 ; 9.65E-04 |
| $UO_2(c)$ | 3.85E-12 | 3.85E-12±2.97E-13 | 6.90E-13 ; 2.10E-11 |
| $PuO_2(c)$ | 1.72E-08 | 1.72E-08±3.54E-10 | 9.56E-09 ; 3.27E-08 |
| $Pu(OH)_4$ | 4.36E-06 | 4.36E-06±7.31E-10 | 4.32E-06 ; 4.40E-06 |
| $NpO_2(c)$ | 2.48E-11 | 2.48E-11±1.10E-13 | 9.38E-13 ; 6.88E-12 |
| $Np(OH)_4$ | 6.60E-09 | 6.60E-09±2.92E-10 | 2.49E-09 ; 1.83E-08 |
| $Am_2(CO_3)_3$ | 3.32E-06 | 3.32E-06±1.88E-08 | 2.83E-06 ; 3.85E-06 |

In this case, it is not possible to make comparisons with the preliminary sensitivity analysis that was made with the binary search tree. However, since the results for $Pu(OH)_4$ showed somewhat odd behaviour, a variance analysis calculation was made for this case. This

calculation shows that the only important parameter by far is the pe. This is not the result of the stepwise regression and, since the different sensitivity analysis approaches give different results, the results may be assumed to be suspicious. Some of the thermodynamic data may for example be inconsistent.

***Table 4.11:*** *The species which stability constants and the other parameters that were selected by the sensitivity analysis.*

| Solid phase | Important species (stability constants) or properties | Internal ranking (reg. coeff. ) |
|---|---|---|
| $ThO_2(c)$ | $ThO_2(c)$; $Th(OH)_4$ | 0.26; 0.25 |
| $UO_3 \cdot 2H_2O$ | $UO_2^{2+}$; $UO_3 \cdot 2H_2O$; $UO_2(OH)_2$ | -0.69; 0.68; 0.10 |
| $UO_2(c)$ | pe; $UO_2(c)$**;** $Fe(OH)_3$ | 1.58; 0.26; -0.16 |
| $PuO_2$ | $PuO_2$; $PuOH^{2+}$ | 0.13; 0.12 |
| $Pu(OH)_4(s)$ | $Pu(OH)_4(s)$; $PuOH^{2+}$ | 0.02; 0.01 |
| $NpO_2(c)$ | $Np(OH)_4$; $NpO_2(c)$ | 0.26; 0.26 |
| $Np(OH)_4(s)$ | $Np(OH)_4(s)$; $Np(OH)_4$ | 0.26; 0.26 |
| $Am_2(CO_3)_3$ | $Am_2(CO_3)_3$; $NaCO_3^-$ $AmSO_4^+$ | -0.05; -0.02 ; -0.01 |

It is interesting in the 80° case to compare the different frequency functions in the same figure. There is obviously a great difference between the different solid phases. It should be remembered when examining Figures 4.8 and 4.9 that the input uncertainty range was equally great in the different cases. This is, as was shown earlier, best illustrated by plotting all the empirical distribution functions in the same figure.

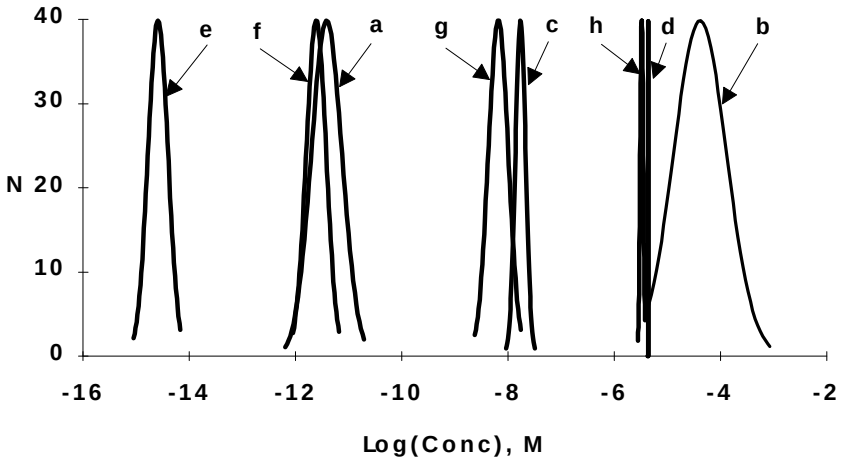



***Figure 4.10:*** *Plot of the solubility distributions for: a) $UO_{2(c)}$, b) $UO_3{}^x2H_2O$, c) $PuO_2$, d) $Pu(OH)_4$, e) $ThO_{2(c)}$, f) $NpO_2$,, g) $Np(OH)_4$,h) $Am_2(CO_3)_3$.*

The distributions in Figure 4.10 are calculated log-normal distributions based on the mean solubility and variance obtained in the uncertainty analysis.

## 4.4 Conceptual model uncertainties

In many cases, the greatest uncertainty does not lie in the effect of input parameter uncertainties, but rather in the choice of model to use in the calculations. As the modelled

system becomes more complicated, there may exist several methods to obtain the desired result. Unfortunately, there does not always exist a method to determine which conceptual model is the correct one. (Unfortunately, even a correct model may give erroneous results if used with incorrect data.) This difficulty to select only one conceptual model will result in different answers to the same question. In the following sections, the rather simple task of calculating the solubility of a solid phase in a selected groundwater has been used to demonstrate the effect of conceptual uncertainties. The problem is how to model the interactions with the rock through which the water containing, for example, an actinide flows. Four different approaches were tested, and uncertainty calculations were made for each one. The uncertainties induced in such calculations due to uncertainties in thermodynamical data are not included since they constitute a rather different problem and are thus treated elsewhere, see previous sections. All of the models have the common factor that they may be motivated by different hypothetical experimental ways to mirror reality.

### 4.4.1 Method 1, Isolated dissolution

The term isolated dissolution derives from the fact that the dissolution procedure was performed without taking the minerals of the rock into account. One method to do this experimentally may be to take a sample of water from a borehole, dissolve the desired solid phase in it and measure the concentration of the desired element. Naturally, several uncertainties exist, such as how the water has changed from the sampling site to the actual experiment.

In the case of simulation instead of experiments, one of the input data to a thermodynamic equilibrium program is the water composition. This water composition is supposed to be known, measured or calculated, together with uncertainty intervals.

In the case presented here, the water composition was obtained by rock water interaction calculations. MINVAR program was used to sample mineral abundances from the same mineral set as for the case desrcibed above, see Table 4.1. The sampling technique used was Latin Hypercube Sampling (LHS), see Appendix 1. The resulting mineral sets were used as inputs to the CRACKER program [EMR 92] which calculated simulated groundwater compositions. From the results, a mean water composition and uncertainty intervals for the properties were calulated. These results were propagated to the UNCCON program where the effect on the calculated solubilities was obtained. The LHS technique was also applied in this case.

### 4.4.2 Method 2, One mineral

The one mineral case may be conceptually described to be similar to the isolated dissolution case, but this time there is also a mineral present in the water. The explanation for such a method is the fact that fracture walls are sometimes covered with one single mineral. If such an area is encountered, where the solubility of the selected element is low, the concentration further downstream will be controlled by the solubility in the presence of that mineral. Minerals selected here were taken from Table 4.1, one at a time, and no weight was applied for the statistical calculations, i.e. every mineral is encountered once.

The method for simulating this is to use a thermodynamic equilibrium program and equilibrate the sampled water with the selected mineral and the solid actinide phase at the same time. When this is done for each mineral in the list, the mean solubility and the standad deviation are calculated. Clearly, there will be no good distribution data from this approach, but the width of the solubility interval is still interesting.

### 4.4.3 Method 3, Simulated water pumping

In the third method, an attempt is made to simulate the real situation in a fracture. Several minerals are distributed along the fracture, for example, according to Figure 4.11. Groundwater flowing through the fracture reacts with the mineral grains.

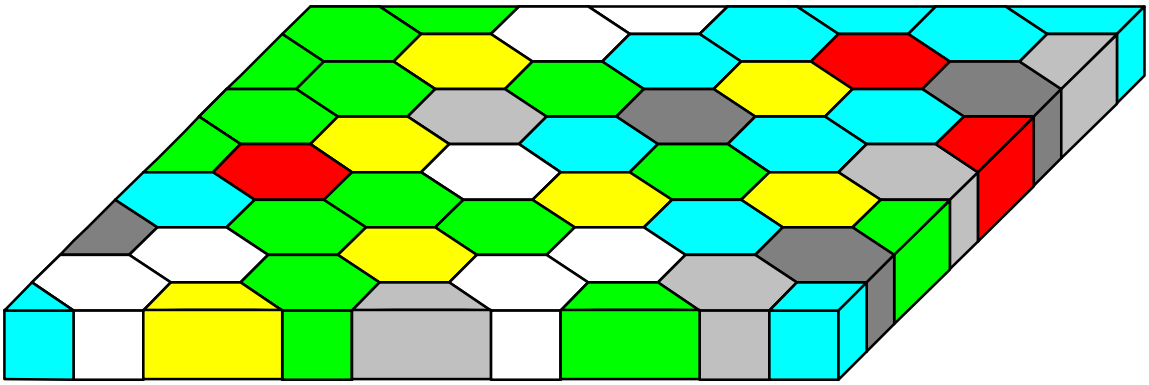


***Figure 4.11****: Example of mineral grain distribution across a fracture surface according to the CRACKER model [EMR 92]*

In the simulations considered here, it has been assumed that a solid phase, e.g. $Pu(OH)_4(s)$, has precipitated onto the surface. The dissolved element, e.g. Pu, is then able to equilibrate with the solid phase, while the groundwater is modified by reactions with the supporting mineral surface. The reactions may locally change pH values and concentrations of ligand species, e.g. carbonate. Due to the spatial variability in the chemistry of the surface, ground water properties will vary, causing "flowpaths" of high and low solubility for the element of interest. Finally, it is supposed that the water flowing through the fracture will have reached some kind of steady state at which all the local changes even out. At that point, water samples from several cells are mixed to simulate the pumping procedure and what actinide concentration is likely to be found by water pumping at a distance from the failed repository. Naturally, several simplifications have been made here. One simplification of major importance is that these calculations are made for one solubility determining phase. It is not certain that this phase will actually be the solubility limiting one, but it has been selected for demonstration purposes.

The calculations were performed using the MINVAR program and, instead of using only the mineral for equilibration in each cell, the desired actinide solid phase is also incorporated.

### 4.4.4 Method 4, Random sampling

This method is similar to the simulated pumping method except that no pumping is simulated. Instead, the full range of solubilities in the fracture is investigated. This method is a complement to method two in the sense that the solubilities are given by one mineral at a time, but the water composition is allowed to change along a fracture. Solubility values are sampled from some 3000 locations along this simulated fracture.

### 4.4.5 Results

The calculations presented here were all done using the same database and the same equilibrium program, and thus the different results may not be assigned to the use of different input data. The calculation motor was the PHREEQE code, and the database was Hatches 5.0 [BON 92]. It must be noted that the choice of the solubility limiting phase and the uncertainty thus introduced are not within the scope of this report and are therefore not included here. The ones selected are at least among those supposed to be solubility limiting ones. The results are mainly statistical estimators, such as mean and variance of the solubilities. However, some additional information on how the different approaches behave may be obtained from the distribution functions, see Figures 4.12, 4.13 and 4.14.

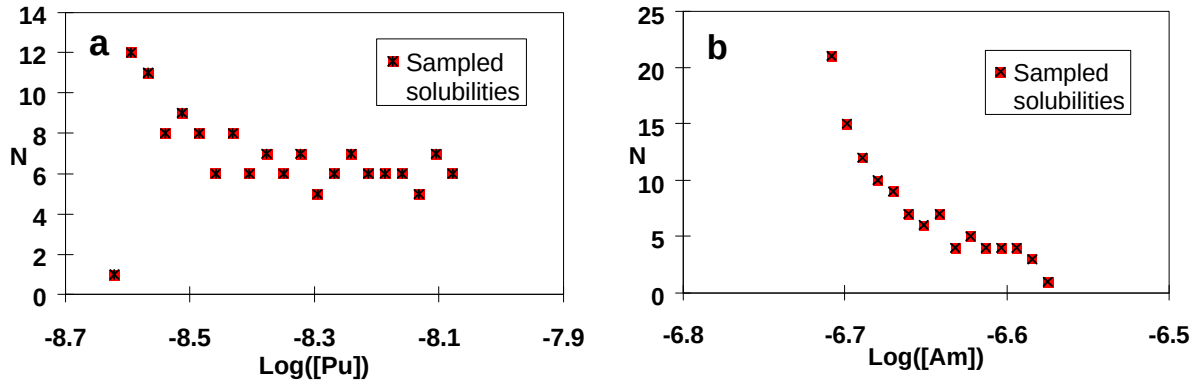



***Figure 4.12:*** *Distribution functions for the isolated dissolution case: a),$Pu(OH)_4$, b), $Am_2(CO_3)_3$.*

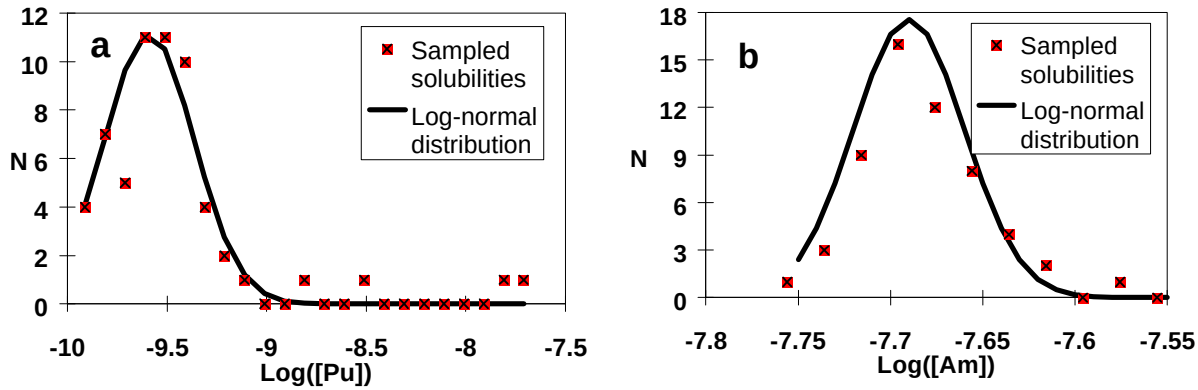



***Figure 4.13:*** *Distribution functions for the simulated water pumping case a), $Pu(OH)_4$, b), $Am_2(CO_3)_3$.*

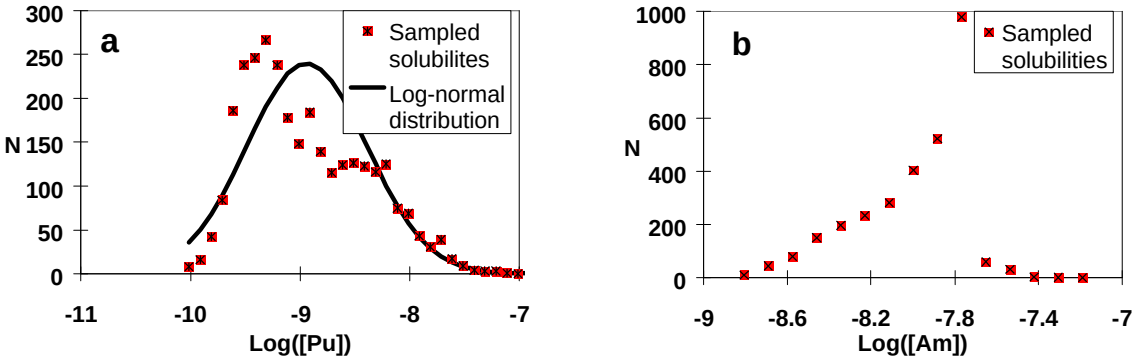



***Figure 4.14:*** *Distribution functions for the simulated random sampling case, a), $Pu(OH)_4$, b), $Am_2(CO_3)_3$.*

As seen in the figures above, it is not obvious how the solubilities will be distributed for the different methods of solubility calculation. There are some similarities in each method but it is not possible to draw any prior conclusion about the distribution. This implies that the use of the mean as the most probable value is not valid for some of the cases, i.e. the isolated dissolution case. The distribution together with the maximum and minimum solubility thus gives the best information. For each of the cases, the minimum and maximum solubility is determined together with a confidence interval for the mean. In some cases, however, it may be difficult to fit the results to a well known distribution. Hence, no attempt has been made to give a confidence interval based on the distributions, but rather with a standard deviation calculated on a logarithmic scale and then recalculated to a linear scale. This will naturally not provide a symmetric interval but the results are more easily read and the bias will shift the results to the higher end, thus making the results more conservative. In the isolated dissolution case two different approaches for obtaining water composition uncertainties have been used. In the first case, the water composition and uncertainties obtained from the MINVAR calculations were used, and the second case used an actual sampled water from the Äspö site in Sweden. The uncertainties are in this case statistical ones originating from different results at different times and measurement uncertainties caused by the analysis method used [SAM 96].

***Table 4.12:*** *Minimum and maximum solubility for the different calculation cases*

| Solid phase | $Pu(OH)_2CO_3$ | $Pu(OH)_4$ | $Am_2(CO_3)_3$ |
|---|---|---|---|
| Case | min ; max solubility (M) | min ; max solubility (M) | min ; max solubility (M) |
| isolated dissolution sim. water unc. | 4.60E-12 ; 6.63E-12 | 2.54E-09 ; 8.90E-09 | 2.00E-07 ; 2.65E-07 |
| isolated dissolution meas. water unc. | 1.02E-11 ; 1.93E-11 | 7.09E-09 ; 1.98E-08 | 8.50E-08 ; 1.06E-07 |
| one mineral | 1.41E-11 ; 2.19E-11 | 1.09E-08 ; 5.39E-08 | 1.72E-09 ; 1.56E-08 |
| simulated water pumping | 1.31E-12 ; 9.65E-12 | 1.23E-10 ; 2.83E-08 | 1.77E-08 ; 2.69E-08 |
| random sampling | 1.13E-12 ; 4.43E-11 | 9.72E-11 ; 2.75E-08 | 3.46E-09 ; 8.51E-08 |

Table 4.12 shows that the variability between the different methods is in some cases greater than their internal uncertainty except for the simulated random sampling. The latter covers almost the entire uncertainty space. Calculations have been made which show that the uncertainties in calculated solubility caused by measurement uncertainties of the water composition, with the isolated dissolution model, are similar to the results obtained with simulated groundwater composition uncertainties, see Paper V. Some deviation may occur since the simulated water does not perfectly match the sampled one. However, this uncertainty will be present for cases where real *in situ* sampling can not or will not be done. Uncertainty intervals of this magnitude or even far greater have also been reported for a solubility calculation with the same water and database [EMR 99].

***Table 4.13:*** *Confidence intervals one standard deviation wide for the calculation cases*

| Solid phase | $Pu(OH)_2CO_3$ | $Pu(OH)_4$ | $Am_2(CO_3)_3$ |
|---|---|---|---|
| Case | min ; max solubility (M) | min ; max solubility (M) | min ; max solubility (M) |
| isolated dissolution sim. water unc. | 5.16E-12±5.17E-13 | 4.43E-09±2.03E-09 | 2.18E-07±2.06E-08 |
| isolated dissolution meas. water unc. | 5.16E-12±5.17E-13 | 1.35E-08±9.00E-09 | 1.35E-08±9.00E-09 |
| one mineral | 1.63E-11±2.44E-12 | 1.40E-08±2.18E-08 | 6.77E-09±5.81E-09 |
| simulated water pumping | 2.02E-12±9.08E-13 | 5.04E-10±8.36E-10 | 2.18E-08±1.97E-09 |
| random sampling | 4.29E-12±3.05E-12 | 1.15E-09±8.38E-10 | 1.48E-08±1.37E-08 |

The great variation term in the simulated water pumping case for $Pu(OH)_4$, seen in Table 4.13, originates from relatively few samples and thus the outliers give a relatively large contribution to the variance. When comparing the numbers given in Table 4.13 with the corresponding distribution functions shown above, it is obvoius that the numerical values do not mirror the "real" situation. Care must thus be taken when using simply a mean value together with its standard deviation without knowledge of the underlying distribution.

## 4.5 Summary of results

To gain a good overview of the results of the calculations made on the uncertainty analysis of solubility calculations, it may be advantageous to see the results together. This will also show the relative magnitude of the effect of the different uncertainties. Clearly, results shown in Table 4.14 are valid for specific cases and may not be generalised but only used as an indication of the magnitude of the uncertainty intervals generated by the different input uncertainties.

***Table 4.14:*** *Results of calculations with uncertain stability constants and enthalpies of reaction*

| Case | Solid phase | Solubility interval (M) | Approx. distr. |
|---|---|---|---|
| Stability constants, 25 $^{o}$C | $Pu(OH)_4$ | 1.30E-09 ; 6.66E-08 | log-normal |
| Enthalpies of reaction, 80 $^{o}$C | $Pu(OH)_4$ | 4.32E-06 ; 4.40E-06 | log-normal |
| Stability constants, 25 $^{o}$C | $Am_2(CO_3)_3$ | 4.35E-08 ; 7.08E-07 | log-normal |
| Enthalpies of reaction, 80 $^{o}$C | $Am_2(CO_3)_3$ | 2.83E-06 ; 3.85E-06 | log-normal |

With the uncertainty intervals selected, the contribution to the solubility uncertainty from the uncertainties in enthalpies of reaction is not important compared to the uncertainties in solubility constants. Naturally, this may change if either the uncertainty intervals for the input parameters or the temperature is changed significantly. The input uncertainties for the case of uncertainties in water composition are taken from two sources: first, the composition given by the MINVAR program based on mineral uncertainties given in Table 4.1, and, second, a water composition uncertainty based on detection and sampling uncertainties in the field [SAM 96].

***Table 4.15:*** *Results of calculations with uncertain water composition from different sources.*

| Case | Solid phase | Solubility interval (M) | Approx. distr. |
|---|---|---|---|
| sim. water comp. | $Pu(OH)_4$ | 2.54E-09 ; 8.90E-09 | -- |
| measured water comp. | $Pu(OH)_4$ | 7.09E-09 ; 1.98E-08 | log-uniform[a] |
| sim. water comp. | $Am_2(CO_3)_3$ | 2.00E-07 ; 2.65E-07 | -- |
| measured water comp. | $Am_2(CO_3)_3$ | 8.50E-08 ; 1.06E-07 | log-uniform[a] |

a, a uniform and a log-uniform distribution are naturally similar

In this case, it is difficult to draw any conclusion about the different methods for obtaining the water composition uncertainty, c.f. Table 4.15. However, it should be noted that the uncertainties in solubility are almost as large as for the case with uncertainties in stability constants. The dashes in the distribution column indicate that it has not been possible to assign any distribution. The results of the calculations on conceptual uncertainties are shown in Table 4.16. It is to note that the case with the isolated dissolution is the same as the simulated water composition in the case described above.

***Table 4.16:*** *Solubility intervals for the different conceptual models ( concentrations in M)*

| | isolated dissolution | one mineral | simulated water pumping | simulated random sampling |
|---|---|---|---|---|
| Solid phase | min ; max | min ; max | min ; max | min ; max |
| $Pu(OH)_4$ | see Table 4.15 | 1.09E-08 ; 5.39E-08 | 1.23E-10 ; 2.83E-08 | 9.72E-11 ; 2.75E-08 |
| $Am_2(CO_3)_3$ | see Table 4.15 | 1.72E-09 ; 1.56E-08 | 1.77E-08 ; 2.69E-08 | 3.46E-09 ; 8.51E-08 |

The uncertainty interval varies significantly between the different methods, as does the distribution of the solubility, see Paper VI. It is not obvious which method should be selected and thus the modeller may influence the result more than is desired. This has been proved for the case of calculating the solubility of $Pu(OH)_4$ with the same database and water [EMR 99]. The authors report differences of several orders of magnitude.

It may be concluded that the effect of uncertainties in stability constants in the cases investigated was the greatest contributer to uncertainties in calculated solubility. However, another factor that is not easy to quantify, i.e. conceptual uncertainties, is something of a "wild card" in most uncertainty analyses. It ought to be mentioned that those uncertainties are very difficult to minimise, and thus it is important that as many approaches as possible are investigated in order to minimise the risk of omitting an important aspect of a problem.

# 5. Experimental

Several techniques exist for the determination of stability constants for water systems. Often, only one method is used in an investigation but, since different methods have different origins, they ought to be used as complements to one another. For the studies presented in this report, two different methods have been used: solvent extraction and potentiometric titrations. These methods complement each other as the potentiometric titration has advantages in regions where there is great difficulty in finding a value for the stability constant by the solvent extraction technique, and vice versa.

## 5.1 Solvent extraction

The principle of extracting an element or substance from one solution to another is rather old. Buchholz [BUC 05] described experiments where extraction of uranylnitrate from a water phase to an ether phase was performed. He was also able to strip the uranium compound from the organic phase back to the water phase again. Bucholz further conclude that the extraction is dependent on the volume ratio between the organic and aqueous phase and the concentration of the element in the aqueous phase.

The main parts of an extraction system are the two almost immiscible phases and, in many cases, an extraction reagent. There exist several kinds of extraction reagents, e.g. chelating reagents and ion exchange reagents. The extraction mechanism varies for each type of reagent and knowledge of the mechanisms is essential for some investigations, e.g. determinations of stability constants.

The extraction process is usually performed by contacting the desired phases in a test tube, e.g. shaking. The test tube is then often centrifuged to ensure a good phase separation, whereafter samples are taken from each phase and measured. Such a method is cumbersome if many samples are needed, and methods for performing the mixing and separation automatically have therefore been developed, e.g. the AKUFVE technique [RYD 69]. The AKUFVE technique will be described in more detail below.

The main principle of the extraction theory is that the desired element ion can be bound to an organic or inorganic ligand giving an uncharged complex. The ligand must be soluble to some degree in both the organicand the aqueous phase. The complex will be distributed between them and the distribution is dependent on the hydrophilic/hydrophobic property of the complex, i.e. the more hydrophobic the complex, the greater the concentration in the organic phase. In the remainder of this report, it is assumed that the organic and aqueous phases are completely immiscible, i.e. no mutual mixing exists. This is not gererally true but, for the practical purposes discussed here, it is a valid assumption [ALB 88:1].

The distribution of the element of interest between two immiscible liquids can then be expressed in terms of the *distribution ratio*, D. By also using the fact that the concentration of a radioactive element, M, is proportional to the specific radioactivity, Equation 5.1 is obtained.

$$D_M = \frac{[M]_{tot,org}}{[M]_{tot,aq}} = \frac{\text{total concentration of M in the organic phase}}{\text{total concentration of M in the aqueous phase}}$$

$$= \frac{\text{Specific radioactivity in the organic phase } (Bq/mL)}{\text{Specific radioactivity in the aqueous phase } (Bq/mL)} = \frac{S_{org}}{S_{aq}} \qquad (5.1)$$

The distribution ratio, $D_M$, is not a constant since it is dependent on, among other things, the ligand concentration, [L].

In these studies, the D value is used to determine the stability constants described by:

$$qM + rL == M_qL_r \qquad (R5.1)$$

$$b_{qr} = \frac{g_{M_qL_r}[M_qL_r]}{g_M^q g_A^r [M]^q [L]^r} \qquad (5.2)$$

were $\gamma_i$ is the activity coefficient for species i.

The subsequent argument concerns mono acidic extraction reagents and is thus not generally transferable to other extraction mechanisms. Assuming that:

i) only mononuclear species exist and
ii) only one species is extractable, the uncharged complex $ML_n$.

Then the distribution of the uncharged complex between the phases is given by:

$$l_n = \frac{g_{ML_n,o}[ML_n]_{org}}{g_{ML_n}[ML_n]_{aq}} \qquad (5.3)$$

It is now possible to derive an expression for the D-value using the stability constants.

$$D_M = \frac{[ML_n]_{org}}{[M^{n+}] + [ML^{(n-1)}] + \ldots\ldots\ldots + [ML_k^{(n-k)}]} =$$

$$= \frac{\frac{l_n}{g_{ML_{n,o}}} g_{ML_n}[ML_n]_{aq}}{[M^{n+}] + [ML^{(n-1)+}] + \ldots + [ML^{(n-k)}]} = \frac{\frac{l_n}{g_{ML_{n,o}}} g_M g_L^n b_n [M^{n+}][L^-]^n}{\sum_{i=0}^{k} \frac{g_M g_L^i b_i}{g_{ML_i}} [M^{n+}][L^-]^i} \qquad (5.4)$$

where $\beta_0$ is equal to one by definition. The right hand side of Equation 5.4 may be simplified by division of the activity of the free metal ion:

$$D_M = \frac{\frac{l_n g_{ML_n}}{g_{ML_{n,o}}} \frac{g_L^n b_n}{g_{ML_n}} \left[L^-\right]^n}{\sum_{i=0}^{k} \frac{g_L^i b_i}{g_{ML_i}} \left[L^-\right]^i} \qquad (5.5)$$

In a more general system in which the hydrolysis may be important, the hydroxide species are also included in the aqueous phase together with other species of interest, as seen in the following sections.

In order to make the evaluation of stability constants a more straightforward process, e.g. to avoid using the activity coefficients, some assumptions are made.

1. The activity coefficient for the uncharged complex in the organic phase is constant during the experiments. This may be obtained by tracer concentrations of the metal and fairly high concentrations of the ligand [WIN 79]. Then this activity coefficient may be a part of λ which is then a system dependent constant.

2. The activity coefficients for the aqueous species are also assumed to be constant during the experiment, owing to a high background electrolyte, thus simpilfying evaluation.

It is clear that the stability constants obtained by the above reasoning will be system dependent, since what is actually determined is the quotent shown in Equation 5.6.

$$\frac{g_L^i b_i}{g_{ML_i}}, i = 1 \ldots n \qquad (5.6)$$

The only way to avoid this is to either calculate the activity coefficients or justify an assumption that they are all equal to 1. With such assumptions, Equation 5.5 may be expressed as:

$$D_M = \frac{l_n b_n \left[L^-\right]^n}{\sum_{i=0}^{k} b_i \left[L^-\right]^i} \qquad (5.7)$$

If the system is well represented by an equation such as 5.7, stability constants for the different species may be determined by fitting of the equation to experimental data. This procedure together with an uncertainty analysis of the results, are explained in more detail below.

### 5.1.1 The AKUFVE apparatus and method

The solvent extraction studies presented within the scope of this report are aimed at the determination of stability constants for the thorium-phosphate system. In doing this, the stability constants for the thorium-acetylacetonates and the thorium hydrolysis are also determined. The originally metallic parts of the AKUFVE system described in Figure 5.1 were made of Pd-passified titanium. The main purpose of using titanuim is that it is fairly inert and resistant to both acids and bases. Some of the experiments presented here were performed using a centrifuge made of PEEK.

The AKUFVE-LISOL system used was developed from the original system by Albinsson [ALB 88:2]. The main principle of the system is shown in Figure 5.1.

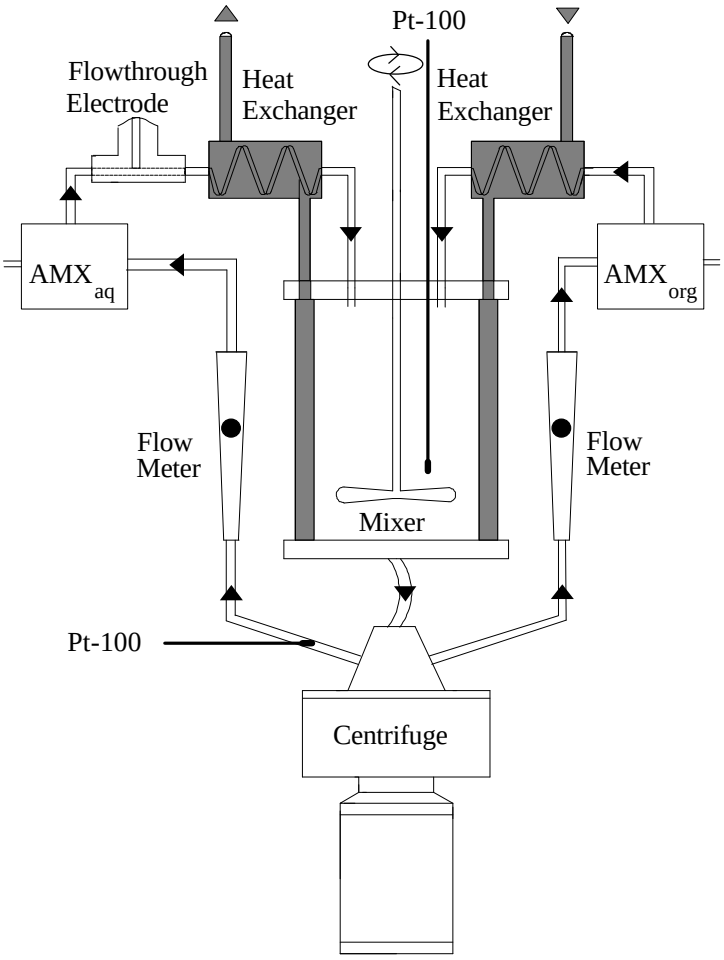


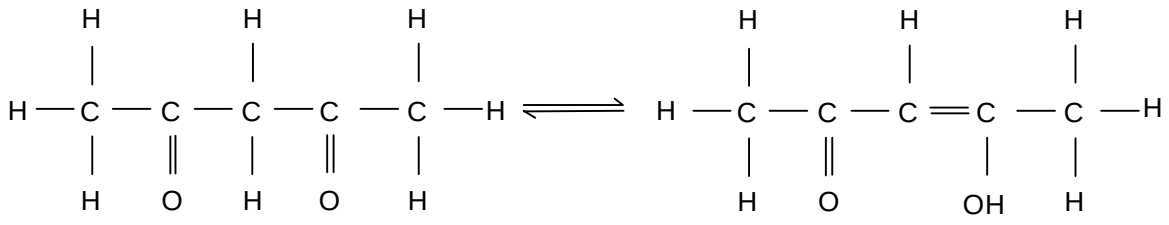

***Figure 5.1:*** *The AKUFVE-LISOL system*

As seen in Figure 5.1, the principle is simple and the equipment consists mainly of two parts: a mixing chamber and a centrifuge. The two immiscible phases are contacted in the mixer and then flow down to the centrifuge for separation. The centrifuge works as a separator as well as a pump. The separated phases are pumped back to the mixing chamber. Samples can be withdrawn from the two flow loops or they may pass a detector of some kind to make on-line measurements.

Since investigations often require good control of temperature and sometimes need to be performed at different temperatures, the temperature is controlled by a circulating liquid flow. The liquid exchanges heat with the solutions in the heat exchangers before entering the mixing chamber. The temperature is monitored by two Pt-100 thermoelements. In earlier experiments, only one sensor inserted in the mixing chamber was used. However, the temperature at the separation is not the same as in the mixing chamber owing to the energy input from the centrifuge and at higher temperatures due to loss of heat. A second Pt-100 sensor was thus placed at the outlet of the aqueous phase at the centrifuge. At 25° C, the temperature difference between the two sensors was about 0.9° C for the water flow rates and centrifuge speeds used.

### 5.1.2 The chemical system

The basic chemical system in the extraction studies was toluene and an aqueous phase. The aqueous phase consisted of 1.0 M $NaClO_4$. A high concentration of an inert background electrolyte minimises the problem of changing activity coefficients in the aqueous phase, as described previously. Perchlorate also has a very weak tendency to form complexes and is therefore assumed not to affect the speciation in the systems. Two thorium isotopes were used in the investigation, $^{234}$Th ($\beta$-, $E_{\beta,max}$ = 0.2 MeV, $t_{1/2}$ = 25.7d) and $^{232}$Th ($\alpha$, $E_\alpha$ = 4.013 MeV and 3.954 MeV, $t_{1/2} = 1.405*10^{10}$ a). $^{232}$Th was used as a hold-back carrier to make the original solution about $10^{-5}$ M with respect to Th. This procedure reduces the effect of sorption losses and trace contaminents [CHO 80]. The $^{234}$Th was prepared according to [ALB 99], and the $^{232}$Th was obtained from dissolving $Th(NO_3)_4$ in perchloric acid, evaporating it to dryness and again dissolving it in perchloric acid. This evaporation and dissolution was repeated four times. The nitrate left in the final solution was deemed not to influence the behaviour of the extraction since the complexes studied are so much stronger. This is not the case for the potentiometric titrations of the thorium-hydroxide complexes, as discussed below. The extraction agent in the thorium studies was acetylacetone (2,4-pentandione). Acetylacetone is often abbreviated as HAa, and this nomenclature is used in this report also. HAa is a colourless liquid with a molecular weight of 100.13 g/mole, a boiling point of 139°C and a melting point of -23°C. The chemical structures of the keto- and enolic forms are shown in Figure 5.2.



Keto form Enolic form

***Figure 5.2:*** *Structural formula of HAa (2,4-pentanedione)*

As seen in Figure 5.2, acetylacetone in the enolic form is a bidentate complexing agent occupying two sites of the bound molecule. This makes it ideal for extracting an ion with a valence state of four and a coordination number of eight, as in the case of thorium.

It is known that acetylacetone decomposes to acetic acid and acetone at high pH [COM 87], which is not desired in the solvent extraction experiments. To investigate this decomposition, spectrophotometric measurements were performed at a pH of about 12.6. This is actually higher than that used in the experiments but, if no significant change was observed within 40 minutes (the time at which the pH in the experiments were higher than 9) at pH 12.6, there should have been no change at lower pH regions. The spectrum obtained is shown in Figure 5.3.

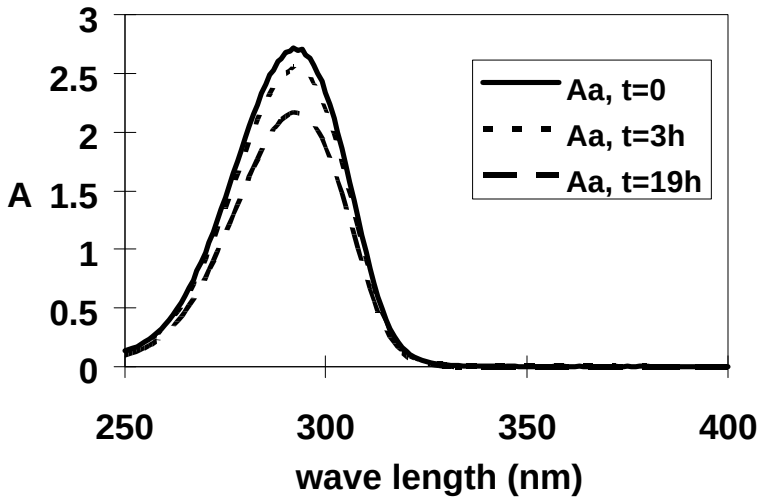



***Figure 5.3:*** *Absorbance spectum for acetylacetone at different times.*

The change in acetylacetone concentration is about 6% after three hours at pH 12.6. Since the experiments were performed at a lower pH and for a shorter period of time, it is safe to conclude that the change in acetylacetone concentration during the experiment is negligible.

Schematically, the chemical system may be described by Figure 5.4.

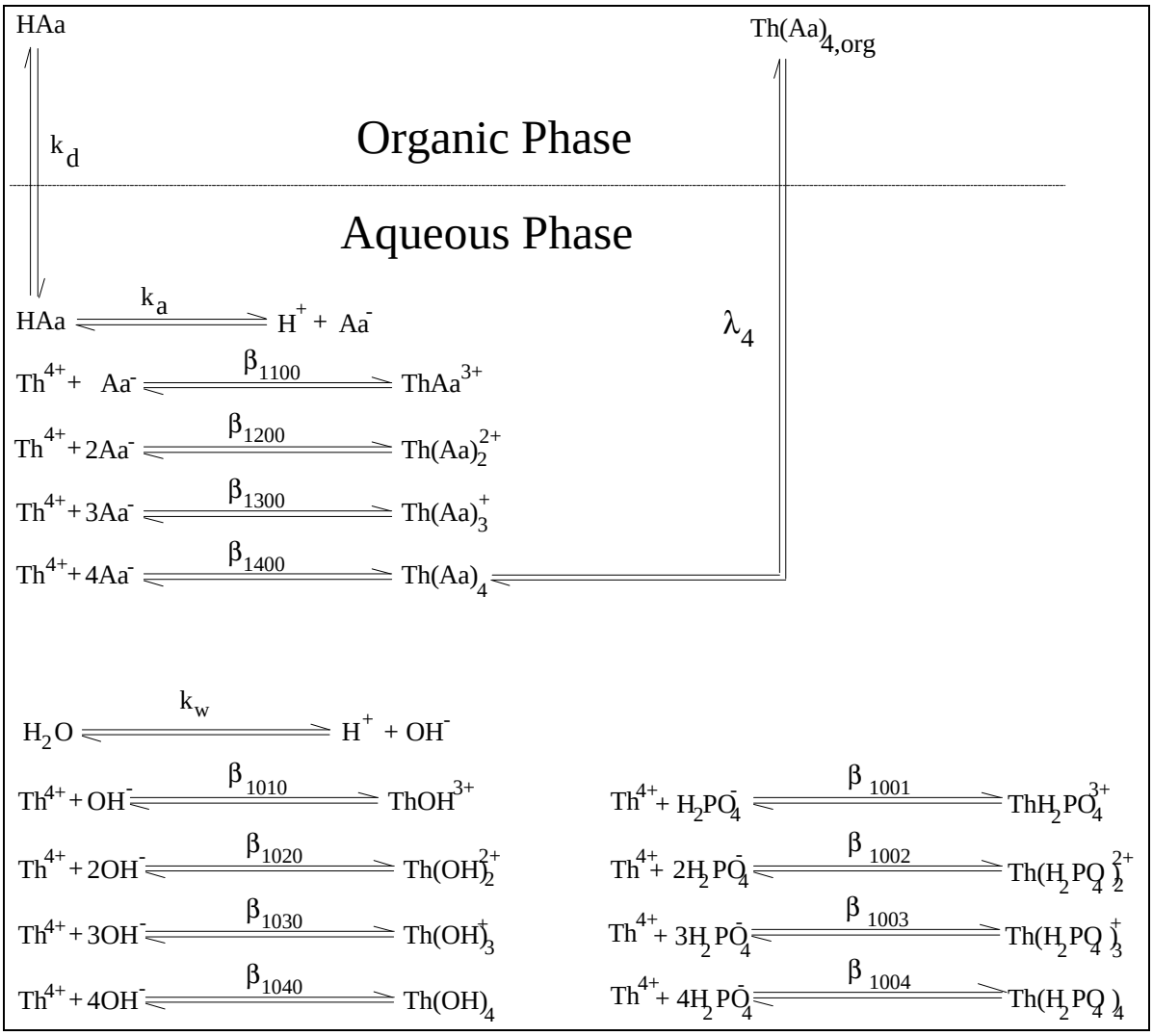



***Figure 5.4:*** *The chemical system used in the phosphate investigation with some modifications*

The system described in Figure 5.4 is the one assumed in this report. Naturally, several simplifications have been made. Among these are:

1. The only species to extract to the organic phase is $ThAa_4$. The other uncharged complexes probably also extract, but to a degree that is negligable compared to $ThAa_4$.

2. There probably exist mixed complexes between thorium and different phosphate species but, with the current method, it is not possible to detect them in a satisfactory way. There may also be mixed complexes between thorium hydroxides and phosphates, but no evidence has been seen for their existence in any greater amounts.

3. The formation of poly-nuclear complexes usually occurs at a higher metal concentration than was used in these solvent extraction studies. However, an existence of poly-nuclear thorium hydroxides, mainly the (4,8) and (6,15) species, has been noted in the potentiometric titration cases.

### 5.1.3. Experimental procedure

The experiments were carried out in two steps at each temperature. The aim of the first step was to determine the stability constants for the thorium-acetylacetonate complexes and for the thorium hydrolysis. The second step was to actually determine the stability constants for the thorium-phosphate complexes.

To determine the constants for the system, thorium-water-acetylacetone experiments were conducted with phases containing, apart from the inert electrolyte, thorium ($<10^{-5}$ M) and 0.1 M acetylacetone in toluene. Since the thorium stock solution was dissolved in 1 M $HClO_4$ the initial pH of the solution was about 2 and was increased during the experiment by additions of 1.0 M NaOH. Samples were taken from each phase, and their activities were measured using a liquid scintillation counter. Calculations have shown that the effect of trace inpurities in the $^{232}Th(NO_3)_4$ on the liquid scintillation measurements were negligable compared to the activity of the $^{234}Th$. The experimental procedure gave extraction curves similar to the one shown in Figure 5.5.

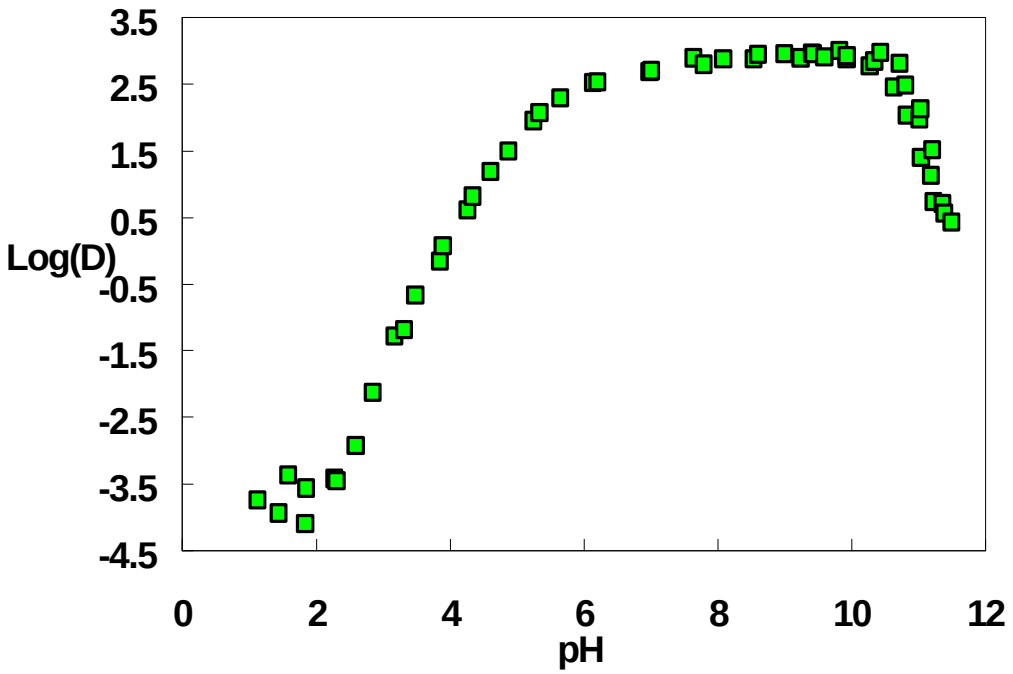



***Figure 5.5:*** *A typical extraction curve for the thorium-water-acetylacetonate system*

The extraction curves were evaluated as described in section 5.3.1. Stability constants for the thorium-phosphate system were obtained using the system described above, but this time with the pH of the solution held at a fixed value. The phosphate concentration was then increased by additions of a mixture of $NaH_2PO_4$ and $Na_2HPO_4$, which was mixed in proportions to buffer the pH in the desired region. However, after some addition of phosphate, the total activity in the solutions decreased owing to formation of a solid phase.

## 5.2. Potentiometric titrations

The main principle behind potentiometric titrations is observation of the change in, e.g. pH, as one solution is titrated with another. One of the more basic cases of potentiometric titrations is the addition of a well known base to a solution containing an acid in order to determine the acidic content. Potentiometric titrations for the determination of stoichiometric and thermodynamic properties of elements in solution are usually based on the Nernst equation [NER 89], Equation 5.8.

$$\mathrm{E} = \frac{RT}{nF} \ln \frac{\mathrm{a}_1}{\mathrm{a}_2} \tag{5.8}$$

where E is the potential between two vessels, R the molar gas constant, T the temperature in kelvin, F the Faradays constant, ν the charge of the ion and a the activity of the determined ion in different vessels. This equation was later generalised to a pair of oxidation states in the same solution and was formulated according to Equation 5.9.

$$\mathrm{E} = \mathrm{E}^\circ + \frac{RT}{nF} \ln \frac{\mathrm{a}_1}{\mathrm{a}_2} \tag{5.9}$$

where E° is the potential at some standard state and $a_i$ is now the acitvity of the different states in the solution. In the following text, Equation 5.9 will be used and denoted the Nernst equation.

There are two methods of fundamental fundamental relevance on the use of potentiometric titrations to determine the properties of a solution. The first was developed by Bodländer [BOD 01] at the very beginning of the 20th century. The main features of this method were to obtain the stoichiometric constants in a reaction between a metal (M) and a ligand (L), similar to reaction RX, see Appendix 2.

However, it is usually also desirable to obtain a value for the stability constant in Equation 5.2. A suitable potentiometric titration method was introduced by Bjerrum [BJE 41] in the 1940s. In his method, Bjerrum used the average ligand number defined by:

$$\bar{\mathrm{n}} = \frac{\left[\mathrm{L}^-\right]_{\mathrm{bound}}}{\left[\mathrm{M}\right]_{\mathrm{tot}}} \tag{5.10}$$

which, expressed as a function of the ligand concentration, is called the formation function. For the case of a ligand releasing one proton in the complexing action and the assumption that the deprotonising of the ligand is negligable compared to the other contributors of protons, the derivation of the formation function will be as follows.

$$\left[\mathrm{H}^+\right]_{\mathrm{m}} = \left[\mathrm{H}^+\right]_0 - \left[\mathrm{OH}^-\right]_{\mathrm{tit}} + \left[\mathrm{ML}^{u-1}\right] + 2\left[\mathrm{ML}_2{}^{u-2}\right] + \ldots \tag{5.11}$$

where the suffix m means measured, subscript 0 denotes the starting point and subscript tit is added during the titration.This becomes, with the use of the stability constants described above

$$\left[\mathrm{H}^{+}\right]_{\mathrm{m}} - \left(\left[\mathrm{H}^{+}\right]_{0} - \left[\mathrm{OH}^{-}\right]_{\mathrm{tit}}\right) = \\ = \frac{b_1}{g_{\mathrm{ML}}} g_{\mathrm{M}} g_{\mathrm{L}} \left[\mathrm{M}^{u+}\right]\left[\mathrm{L}^{-}\right] + 2\frac{b_2}{g_{\mathrm{ML}_2}} g_{\mathrm{M}} g_{\mathrm{L}}^{2} \left[\mathrm{M}^{u+}\right]\left[\mathrm{L}^{-}\right]^{2} + \ldots \tag{5.12}$$

Thus the amount of L, that is coordinated to M will be

$$\left[\mathrm{L}^{-}\right]_{\mathrm{bound}} = \frac{b_1}{g_{\mathrm{ML}}} g_{\mathrm{M}} g_{\mathrm{L}} \left[\mathrm{M}^{u+}\right]\left[\mathrm{L}^{-}\right] + 2\frac{b_2}{g_{\mathrm{ML}_2}} g_{\mathrm{M}} g_{\mathrm{L}}^{2} \left[\mathrm{M}^{u+}\right]\left[\mathrm{L}^{-}\right]^{2} + \ldots \tag{5.13}$$

A mass balance for the total metal concentration yields

$$\left[\mathrm{M}^{u+}\right] = \frac{\left[\mathrm{M}\right]_{\mathrm{tot}}}{\sum_{\mathrm{j}=0}^{\mathrm{n}} \frac{b_{\mathrm{j}}}{g_{\mathrm{ML\,j}}} g_{\mathrm{M}} g_{\mathrm{L}}^{\mathrm{j}} \left[\mathrm{L}^{-}\right]^{\mathrm{j}}} \tag{5.14}$$

Equation 5.13 together with Equation 5.14 yield the following expression for the average ligand number

$$\bar{\mathrm{n}} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{i} \frac{b_{\mathrm{i}}}{g_{\mathrm{ML}_{\mathrm{i}}}} g_{\mathrm{M}} g_{\mathrm{L}}^{\mathrm{i}} \left[\mathrm{L}^{-}\right]^{\mathrm{i}}}{\sum_{\mathrm{j}=0}^{\mathrm{n}} \frac{b_{\mathrm{j}}}{g_{\mathrm{ML}_{\mathrm{j}}}} g_{\mathrm{M}} g_{\mathrm{L}}^{\mathrm{j}} \left[\mathrm{L}^{-}\right]^{\mathrm{j}}} \tag{5.15}$$

where the average ligand number may be calculated as:

$$\bar{\mathrm{n}} = \frac{\left[\mathrm{H}^{+}\right]_{\mathrm{m}} - \left(\left[\mathrm{H}^{+}\right]_{0} - \left[\mathrm{OH}^{-}\right]_{\mathrm{tit}}\right)}{\left[\mathrm{M}\right]_{\mathrm{tot}}} \tag{5.16}$$

The stability constants are then obtained by fitting of the stability constants. Thus, there is the same problem as for the solvent extraction case. Unfortunately, in potentiometric titrations, the assumption of activity coefficients close to one is more doubtful since it is common to work with higher metal concentrations. However, it may still be possible to assume that they are constant during the experiment, thus making it possible to determine pseudo stability constants according to Equation 5.6. One property is noteworthy, however: it is not expected from the reasoning above that the stability constants obtained by solvent extraction will be the same as for potentiometric titrations. Still, hopefully, they are in the same region.

### 5.2.1 The potentiometric titration apparatus and method

The titrations were made using an automatic byrette and measuring system [ABU, Radiometer, Denmark] which was controlled by a personal computer. The programs used were supplied in part by the manufacturer of the measuring system and in part developed at the Department of Nuclear Chemistry [JAC 98]. The main feature is that the potential between a reference electrode and a glass electrode is measured several times from one addition of titrand to another. The time between additions is given by the user, as is the number of samples during this time. The different measurements made after each addition are displayed to ascertain that a steady state has been reached. The values thus displayed are in turn the average of a given number of measurements, and the associated standard deviation may also be obtained. The temperature is also measured.

### 5.2.2 The chemical system used and experimental procedure

Two different investigations were performed, determinations of thorium-acetylacetonate formation constants and thorium hydrolysis constants. The experimental setup was similar to that described above. The thorium used was originally made by using the procedure described in section 5.1.2. However, studies using an ion chromatograph, showed that almost no nitrate was driven off. This fact does not disturb the measurements of the acetylacetone complexes to any great extent, but the precence of nitrate was unsatisfactory for the hydrolysis studies. Thus, in order to eliminate the nitrate in the thorium solution, $ThO_2$ was used. The $ThO_2$ was boiled in concentrated perchloric acid for one day, after which a drop of HF was added to catalyse the dissolution. The solution was then evaporated to dryness and the precipitate dissolved in 1.0 M perchloric acid. However, the resulting solution did not have the desired acidity, i.e. the acidity given by the acid and the resulting pH was about one unit higher. Such behaviour may be explained by the following reaction having taken place at the end of the evaporating step.

$$Th(ClO_4)_4 + nH_2O == ThOH_n(ClO_4)_{(4-n)} + nHClO_4 \qquad (R5.2)$$

where the perchloric acid formed is driven off during the evaporation. The effect of this phenomenon is that the pH of the thorium solution is not known *a priori* but must be determined. Gran titrations [GRA 50], [GRA 52] were made on a solution containing the thorium solution and an added amount of $HClO_4$. This gave the value of n=1 in Reaction R5.2. However, owing to evident problems in using a thorium solution which is not acidic, a slightly different approach was used. The thorium solution was made according to the procedure described above, with the difference that it was not evaporated to complete dryness. Then the acid content was determined by Gran titrations.

To obtain a higher concentration precision in the titrand, the density of the participating solutions was measured using pyknometers. The density of 1.0 M $NaClO_4$ is 1078.1 kg/m$^3$, and the density of 0.1 M acetylacetone in 1.0 M $NaClO_4$ is 1078.6 kg/m$^3$. In the determination of the stability constants for the thorium-acetylacetonate stability constants, the titrations were

made both from the acidic and basic side. In the acidic case, the titrand consisted of 53.4g 0.1M acetylacetone in 1.0M $NaClO_4$ together with 0.5ml 1.0M Th in 1.0M $HClO_4$. This was titrated with 1.0 M NaOH, with respect to hydroxide, in 1.0 M $NaClO_4$.The titrations from the basic side used a similar titrand except that 1.5ml 1.0 M NaOH in 1.0 M $NaClO_4$ was added to the solution specified above. This was titrated with 1.0 M $HClO_4$ with total additions up to 1.8 ml.

The thorium hydolysis studies were made with a titrand constisting of 53.4g 1.0M $NaClO_4$ together with 0.5ml 1.0M Th in acidic solution. Since the concentration of the perchloric acid in the thorium solution was not known, Gran titrations were performed in each experiment.

## 5.3. Methods of evaluation

The commonly used method for evaluating the results in the case of both potentiometric titrations and solvent extraction, has been the fitting of parameters in some expression to the obtained data. There exist several methods to obtain this fitting and assigning the corresponding uncertainties, as described in section 3.4.3.

### 5.3.1 Solvent extraction

The solvent extraction experiments were performed in two stages. The first one with the system containing only thorium and acetylacetone together with the solvents and the in second one, phosphate was added during the experiment. These different approaches cause the objective function in the fitting process to be slightly different.

The first experimental stage aimed at determining the stability constants for the thorium-acetylacetonates and the thorium-hydroxides. The expression for the distribution of thorium between the organic and aqueous phase is described by:

$$D_{Th} = \frac{l_4 b_{1400}\left[Aa^-\right]^4}{1+\sum_{i=1}^{4} b_{1i00}\left[Aa^-\right]^i + \sum_{j=1}^{4} b_{10j0}\left[OH^-\right]^j} \qquad (5.17)$$

where the notations are according to Figure 5.4.

If the extraction curve for this system is investigated, it is clear that the fitting procedure may be divided into two parts, see Figure 5.6.

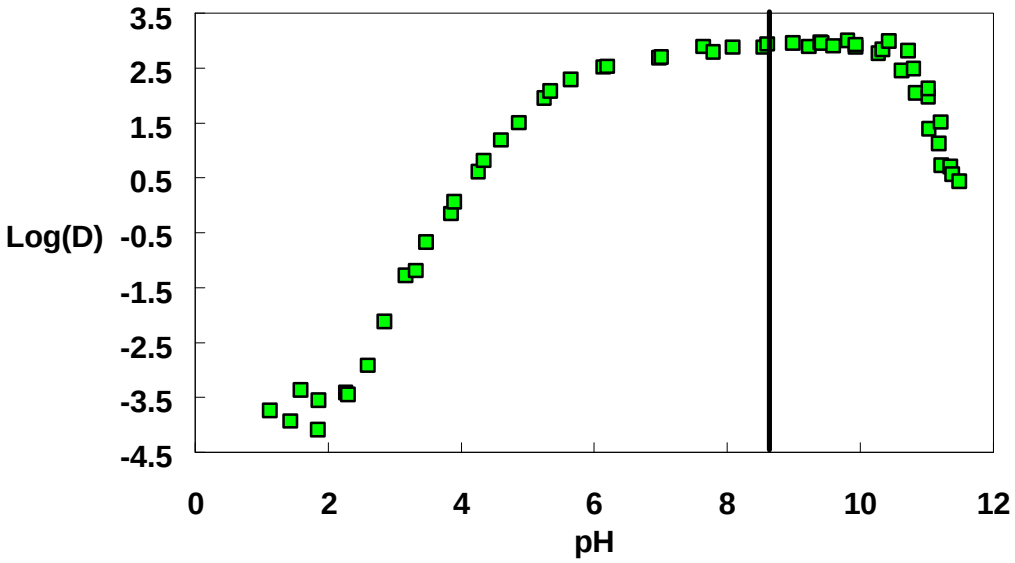



***Figure 5.6:*** *Extraction curve for the thorium-acetylacetonate system, [HAa]=0.1M in toluene, aqueous phase:1.0 M $NaClO_4$, 35°C, three experiments*

As shown in Figure 5.6, the plot may be divided somewhere on the plateau. In the left part, the hydrolysis is negligible compared to the acetylacetone complexation, and the distribution function may thus be simplified in such a way that it is only necessary to determine the stability constants for the thorium-acetylacetonates. In the next step, all the data may be used to determine the remaining hydrolysis constants.

There is, however, a drawback using solvent extraction with this system. The points at the far left of Figure 5.6 indicate the lowest distribution ratio detectable in this case. The concentration of the acetylacetonate in the water phase does not reach sufficiently low levels to make the concentration of the free thorium dominant in the solution as compared with the other complexes and thus the term 1 in the denominator of Equation 5.17 may be omitted. It is now possible to divide the stability constants by an arbitrary constant, and thus the task of obtaining the stability constant has an infinite number of solutions, see Appendix 3. However the ratio between two successive constants is always the same, so if it is possible to obtain the value for the stability constant for the first complex, the rest are fixed at the desired value. In this work, the first stability constant was determined by potentiometric titrations, see section 5.3.2.

The extraction curve for the thorium-phosphate-acetylacetone system is similar to that shown in Figure 5.7.

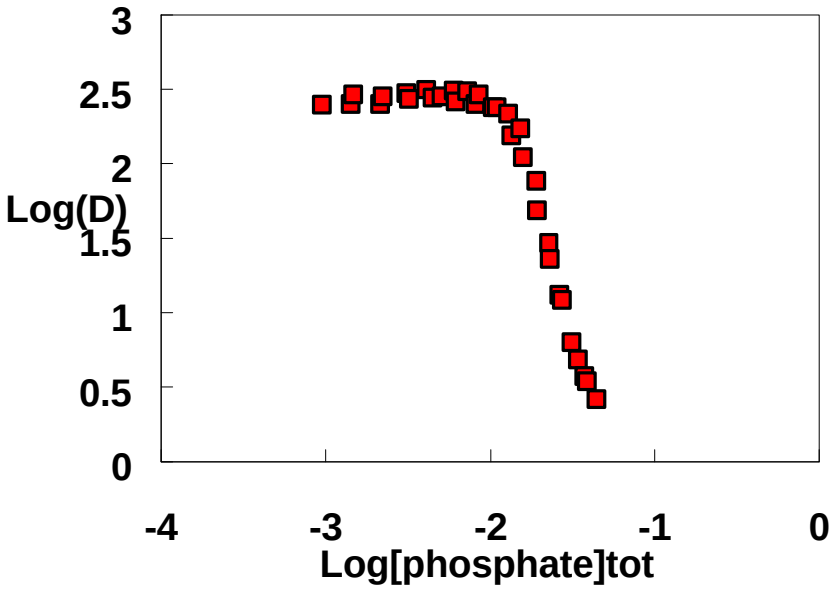



***Figure 5.7:*** *Typical extraction curve for the thorium-phosphate-acetylacetone system, [HAa]=0.1M in toluene, aqueous phase:1.0 M $NaClO_4$, 15°C, pH=8, two experiments*

However, it is not clear which are the complexing species and thus how the expression for the distribution function should be defined. In this report, the argument is that the phosphate species participating in complexes with thorium at the investigated pH, i.e. pH 7, 8 and 9, are $H_2PO_4^-$. This may seem in contradiction to other measurements, e.g. [ENG 94] and [MOS 67], who conclude that $HPO_4^{2-}$ participated in the complex formation. The data by Moskvin *et al.* is discussed by Engkvist and Albinsson in [ENG 94]. However, Engkvist and Albinsson write that they do not have a satisfactory explanation for their own results. In the light of later investigations, the results they find difficult to explain may be explained using the assumption made in this report. They concluded that there was a slope of three in the extraction curve, thus indicating that there were three phosphates on each thorium. However, upon examination of later results, the slope seem to be four which indicates four phosphates on each thorium. The data obtained by Engkvist and Albinsson have been plotted together with the theoretical curve obtained by using the constants presented in this report, and the agreement is strikingly good, see Figure 6.9. Another problem was illuminated in the paper by Engkvist and Albinsson. They had problem explaining that the stripping of thorium from the organic phase began at about a one order of magnitude higher phosphate concentration for pH 9 than for pH 8. This was not observed in recent measurements up to a phosphate concentration of 0.6 M. At higher concentrations precipitates were formed which made it impossible to draw any conclusions at these high concentrations. As seen in Figure 5.8, no stripping should be expected at pH 9 until there is a much higher phosphate concentration, which fits the interpretation given in this report. On the other hand, if the $HPO_4^{2-}$ complexes are dominant, strip at pH 9 should occur earlier, as seen in Figure 5.8.

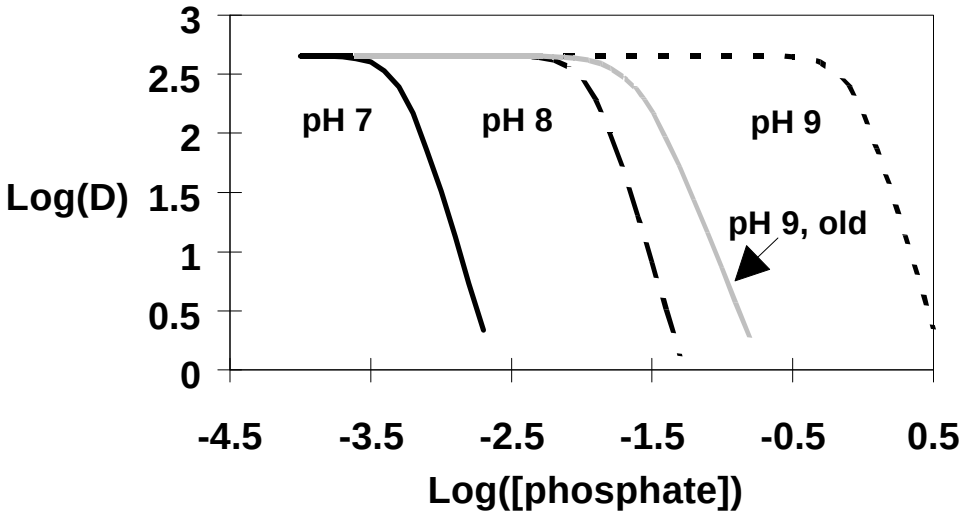



***Figure 5.8:*** *Calculated extraction behaviour at different pH. Old - only $HPO_4^{2-}$-complexes considered in the fitting procedure.*

The results of the current work follow the theoretical lines in Figure 5.8 fairly well. At least considerably better than for the old assumption with $HPO_4^{2-}$ complexation. A recent reexamination of the data of Engkvist and Albinsson shows that the strip from the organic phase at pH 9 presented in their paper may have been caused by precipitation problems. These conclusions have led to the following expression for the distribution coefficient for the thorium-phosphate-acetylacetone system at pH 8.

$$D_{\mathrm{Th}} = \frac{l_4 b_{1400}\left[\mathrm{Aa}^-\right]^4}{1+\sum_{i=1}^{4} b_{1i00}\left[\mathrm{Aa}^-\right]^i + \sum_{j=1}^{4} b_{10j0}\left[\mathrm{OH}^-\right]^j + \sum_{k=1}^{4} b_{100k}\left[\mathrm{H_2PO_4}^-\right]^k} \qquad (5.18)$$

The acetylacetonate data and the hydroxide data were determined earlier and thus only the phosphate data are determined at this stage. The reactions assumed in Equation 5.18 are the ones shown in Figure 5.4.

### 5.3.2 Potentiometric titrations

Potentiometric titrations were performed for two different systems, the thorium-acetylacetone system and the thorium-hydroxide system. The hydroxide is naturally present in the acetylacetone system as well but in so much smaller quantities than acetylacetonate that it will not influence the measurements of the acetylacetonate complexes. The method used to obtain the desired stability constants was a fitting procedure in this case as well.

A typical potentiometric curve for the acetylacetone system is shown in Figure 5.9. This curve was evaluated using Equation 5.19.

$$\left[\mathrm{H}^+\right]_{\mathrm{m}} - \left(\left[\mathrm{H}^+\right]_0 - \left[\mathrm{OH}^-\right]_{\mathrm{tit}}\right) = b_{1100}\left[\mathrm{Th}^{4+}\right]\left[\mathrm{Aa}^-\right] + 2b_{1200}\left[\mathrm{Th}^{4+}\right]\left[\mathrm{Aa}^-\right]^2 + \ldots \qquad (5.19)$$

where subscript m means measured, subscript 0 denotes starting point and subscript tit gives quantities added during the titration. For convenience, the expression on the left side in Equation 5.19 is called Y, according to Equation 5.20.

$$Y = \left[H^+\right]_m - (\left[H^+\right]_0 - \left[OH^-\right]_{tit}) \qquad (5.20)$$

Y was calculated from the experiment and the right side was used for the fitting of the stability constants. However, as indicated earlier, only the first and second stability constants were evaluated using potentiometric titrations.

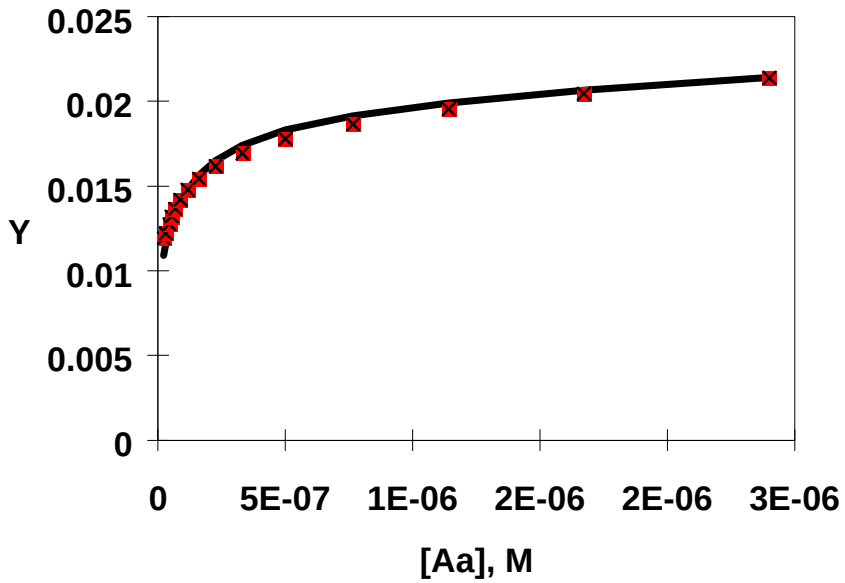



***Figure 5.9****: Typical curve for the potentiometric titration on the Th-Aa system,where Y is the left hand side of Equation 5.19. The x values are experimental points; the curve is fitted.*

The desired stability constants were obtained by fitting Equation 5.19 to the data obtained, selecting the weight in the fitting in such a way that the largest weight was given to the points determining the first stability constant. The reason for this is that the first constant is the one that is most difficult to determine by solvent extraction thus it is the most important one to determine accurately by potentiometric titrations.

For the hydroxide stability constants, the approach was slightly different. Instead of fitting the average ligand number to the data, the quotient shown in Equation 5.21 was used.

$$\frac{\bar{n}}{1-\bar{n}} = b_{1010}\left[OH^-\right] \qquad (5.21)$$

Thus a linear equation is obtained in which the slope is the value of the first hydroxide stability constant. This procedure is motivated for several reasons, one of the most important being the formation of polynuclear hydroxide species as the hydroxide concentration becomes higher.

This is the case at least in the titrations performed here, where the thoruim concentration is high, about 0.01M.

For some of the hydroxide species, e.g. (1,2), (4,8), and (6,15), a different approach was used. The experiments and evaluation were conducted at the Australian Nuclear Safety and Technolgy Organisation (ANSTO) site, chiefly by Dr. Paul Brown, see Paper VII. The potentiometric data were analysed using the computer program MINIQUAD [BRO 83]. MINIQUAD separately and independently minimises both $\{[H]_{T(calc.)} - [H]_{T(obs.)}\}^2$ and $\{[M]_{T(calc.)} - [M]_{T(obs.)}\}^2$ as defined by the mass balance equations and experimental observations. It is an augmented version of the original program [SAB 74, GAN 76] providing the following enhanced features:

(i) numerical refinement of the analytical proton excess at the beginning of a titration, allowing a titration to be commenced at any pH value irrespective of the extent of reaction;
(ii) optional numerical refinement of the relationship between pH values and hydrogen ion concentrations using equation (5.22).

$$[H^+] = \frac{10^{-pH}}{l} \qquad (5.22)$$

(iii) optional numerical refinement of negative formation constants; and
(iv) two automated model (as opposed to species) selection procedures in addition to the 'manual' method as described by Gans *et al.* [GAN 76].

### 5.3.3 Enthalpy and enthropy of reaction

The evaluation of the enthalpy of reaction is similar regardless of which method has been used for the determination of the stability constants. The derivation of the expression used is shown below.

$$G = H - TS \qquad (5.23)$$

At the standard state the stability constant are related to the free energy accoding to:

$$\Delta G^0 = -RT * \ln b \qquad (5.24)$$

Using Equation 5.23 for the right and left side of the desired reaction, at a standard state, inserting Equation 5.24 and dividing by T yields:

$$-R \ln 10 \log b = \frac{\Delta H^0}{T} - \Delta S^0 \qquad (5.25)$$

Assuming that the enthalpy and entropy are temperature independent in the temperature range investigated, they may be obtained by linear regression of the stability contants at the different temperatures. At the same time the fitting errors may be obtained.

Another approach is to use all degrees of freedom and fit three parameters to the data at three temperatures. Then it is assumed that both the enthalpy and the enthropy are temperature dependent according to Equations 5.26 and 5.27.

$$\Delta H^0_T = \Delta H^0_{T_0} + \int_{T_0}^{T} \Delta c_p dT \tag{5.26}$$

$$\Delta S^0_T = \Delta S^0_{T_0} + \int_{T_0}^{T} \frac{\Delta c_p}{T} dT \tag{5.27}$$

This gives the following set of equations, if the $\Delta c_p$ is assumed to be temperature independent, for the case presented in this report, i.e. the experiments are made at 15°C , 25°C and 35°C, and 25°C is selected as the standard temperature. These equations may be solved using matrix operations.

$$\Delta H^0_{T_2} + \Delta c_p((T_1 - T_2) - T_1 \ln(\frac{T_1}{T_2})) - T_1 \Delta S^0_{T_2} = -RT_1 \ln(b_1) \tag{5.28}$$

$$\Delta H^0_{T_2} - T_2 \Delta S^0_{T_2} = -RT_2 \ln(b_2) \tag{5.29}$$

$$\Delta H^0_{T_2} + \Delta c_p((T_3 - T_2) - T_3 \ln(\frac{T_3}{T_2})) - T_3 \Delta S^0_{T_2} = -RT_3 \ln(b_3) \tag{5.30}$$

where subscripts 1, 2 and 3 denote temperatures and temperature 2 is selected as the standard state. However, it is important to note that all degrees of freedom are now used and thus no excess information is available to give any uncertainty intervals for the fitted parameters. The choice of method is thus dependent on the confidence in the experimental results.

### 5.3.5 Uncertainty analysis

There exist several methods for making uncertainty analysis of fitted parameters, for example bootstrap, jackknife [EFR 93] and chi-square methods [EKB 99:2]. In this work the Jackknife and the Chi-square methods have been used and they will be outlined below. Both methods have the advantage that they may be performed *a posteriori*, which in this case means that they may be used regardless of the fitting algorithm.

### 5.3.5.1 The Jackknife estimate

The jackknife procedure of estimating parameter uncertainties may be introduced as a resampling method [Ekb 99:2]. Since the jackknife does not require that the errors, $\varepsilon_i$, follow a certain distribution function, jackknife is representative of non-parametric statistics. Resampling is done by using the m originally measured data pairs $(x_i,y_i)$ as adequate representation of the true but unknown distribution function. The properties of this distribution function are explored by creating sub-samples from the m data pairs. In the jackknife approach, each sub-sample contains m-1 of the $(x_i,y_i)$ data pairs. An algorithmic application of the jackknife for estimation of variance may be as follows:

a) obtain the n parameters p* by fitting the model function $f(x_i,p_n)$ to the m data points $(x_i,y_i)$.
b) create a sub-sample by omitting the k-th data pair $(x_k,y_{m,k})$, re-fit the sub- sample and obtain n sub-sample parameters $p_k$.
c) repeat step b) m times by successively omitting the *m*-th data pairs $(x_m,y_{m,m})$
d) calculate the jackknife estimate of variance $s^2$ by:

$$\sigma^2 = \frac{m-1}{m}\sum_{k=1}^{m}(p_k - p^*)^2 \qquad (5.31)$$

In the more general case, the mean standard error is used instead of the variance, but the empirical variance is more appropriate for the applications described in this report. If there are many experimental data points (m is large), the jackknife may be abbreviated by randomly choosing j (j<m) data pairs for which sub-samples are considered. Of course, m must be replaced by j in Equation 5.31. Although the Jackknife procedure may look like a rough-and-ready statistical tool, it has a sound statistical basis as reviewed e.g. in [EFR 93]. It should be noted that jackknife estimates of variance are too high. The jackknife approach requires m+1 (or j+1) replications of the fitting procedure. Considering the abundance of high-speed desk-top computers nowadays, jackknifing even large data sets is done relatively fast. Modifying existing software for a new data set usually takes more time than the computations themselves. There is one great advantage of the full jackknife approach in the correlation matrix that results from the calculations: for each data point omitted, a set of fitted parameters is obtained. It is thus possible to detect the effect of possible outlier data and their effect on the estimated parameters.

### 5.3.5.2 The Chi-square estimate

The basic idea behind this method is the assumption that the difference between the measured data points $y_{m,i}$ and $y_i$ is normally distributed according to:

$$N_i = \frac{y_{m,i} - y_i}{k_i} \tag{5.32}$$

where $k_i$ is a weight factor. The sum of squared normal distributions

$$c^2 = \sum_i N_i^2 \tag{5.33}$$

is said to be $c^2$ distributed. This sum, divided by its degrees of freedom (df), gives a measure of the overall quality of the fit. The degree of freedom is defined by:

$$df = m - n \tag{5.34}$$

where m is the number of data points and n the number of parameters fitted.

This experimentally obtained $c^2$ may be compared to the expectation value of the $c^2$ distribution, which is simply the degree of freedom itself. The variance of the $c^2$ is two times the expectation value which may be used to estimate the variance for the fitted parameters. The procedure is simple, it is necessary only to change the parameter value in the model and calculate $c^2$. When it has reached two times its original value, one variance has been added to the investigated parameter. A standard deviation, $s_i$, for each parameter is then assigned as the square root of the variance.

The weighting parameter, $k_i$, in Equation 5.32 is usually chosen as the variance of a data point $y_i$ [KOR 75]. If this quantity - which often requires considerable additional experimental effort - is not available, then $k_i$ is set to unity.

# 6. Experimental results

One of the primary reasons for making this experimental analysis, in addition to trying to shed light on areas not deeply investigated, is to gain a better understanding of the uncertainties associated with determinations of thermodynamic data. Thus, a rather conservative approach has been taken when estimating the uncertainties presented here. They may seem large in some cases but it ought to be remembered that it is only the uncertainties in the evaluation method that are calculated since, in these cases, they are the largest. In addition, as stated earlier, different methods have been used to obtain the desired parameters, thus trying to optimise method vs. stability constant as well as possible.

Throughout this work, the correlation between the activity and the concentration of $H^+$ has been calculated according to Equation 6.1.

$$-\log([H+]) = pH + 0.23 \qquad (6.1)$$

This equation is valid for ion strength 1.0 [FAN 96], and all experiments presented in this section were performed at that ion strength.

## 6.1 The acetylacetone system

Both potentiometric titrations and solvent extraction were used to obtain the full range of the stability constants for the Th-Aa system. The parameters were then obtained by fitting as described in section 5.3.2.

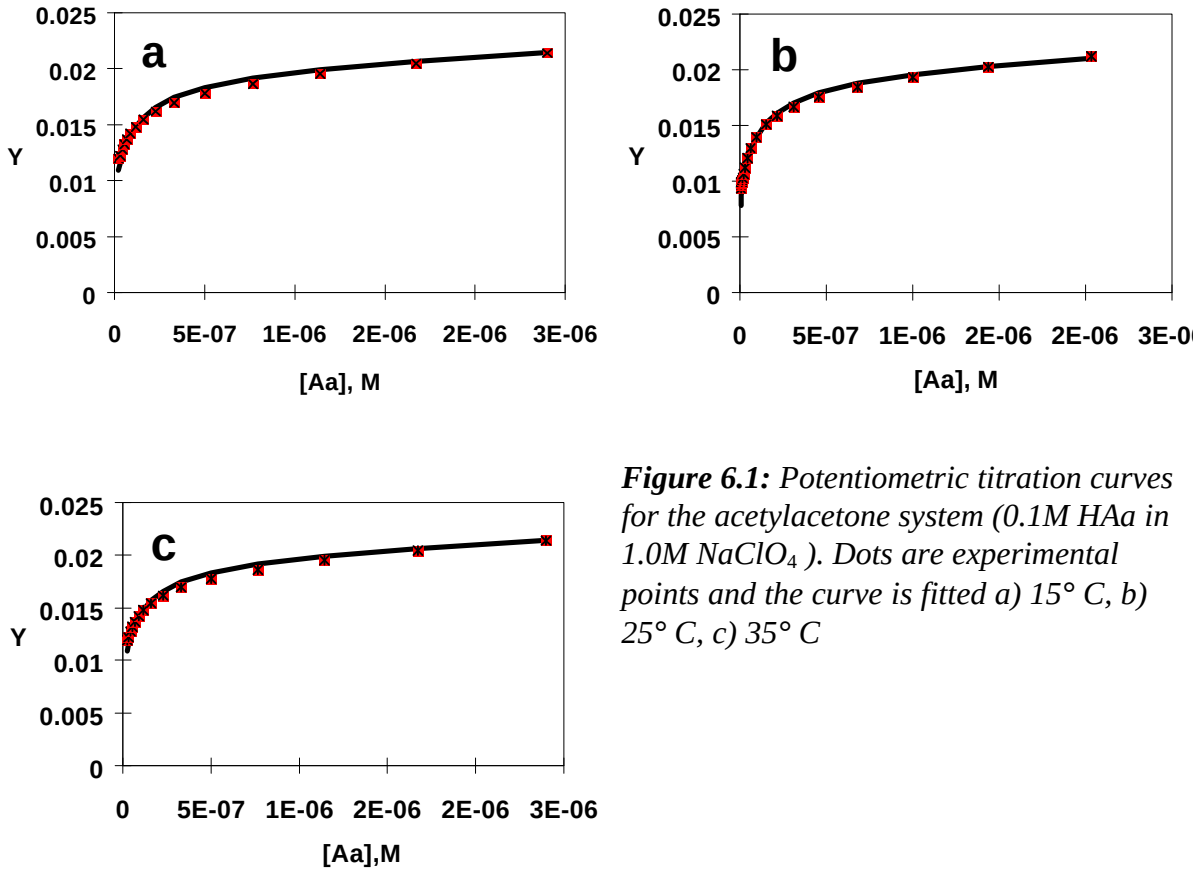



***Figure 6.1:*** *Potentiometric titration curves for the acetylacetone system (0.1M HAa in 1.0M $NaClO_4$ ). Dots are experimental points and the curve is fitted a) 15° C, b) 25° C, c) 35° C*

Although the curves obtained in Figure 6.1 used both the first and second stability constants in the fitting, only the values for the first ones were used as a final result. This is mainly due to the greater confidence in the values for the stability constants obtained by solvent extraction for the (1,2), (1,3) and (1,4) complexes. However, the results for the second stability constant were in good agreement with the solvent extraction results, e.g. 16.2±0.2 for the potentiometric titrations at 25° C as compared to 16.7±0.6 for the solvent extraction case. It must be noted that the uncertainty in the stability constant obtained by solvent extraction is greater than for the potentiometric titrations since it is only the fitting errors that are given. The other uncertainties in the potentiometric titrations are probably greater due to effects of activity coefficients, see section 5.3.

The solvent extraction experiments were evaluated by using the left part of the extraction curve to extract the second, third and fourth stability constants. The whole set of data was then used for the determination of the third and fourth stability constant for the hydroxide system, as described below. The value of the distribution of the uncharged acetylacetone complex between the organic and aqueous phase is given in Table 6.1 together with the dissociation constant for acetylacetone.

***Table 6.1:*** *Parameters used in the evaluation*

| Temp. / Entity | 15°C | 25°C | 35°C | ref. |
|---|---|---|---|---|
| $\lambda_4$ | 280±16 | 450±19 | 850±17 | 427, [ENG 92], 25°C |
| $Log(K_{a,\ HAa})$ | -9.11 | -9.00 | -8.88 | [LIL69] |

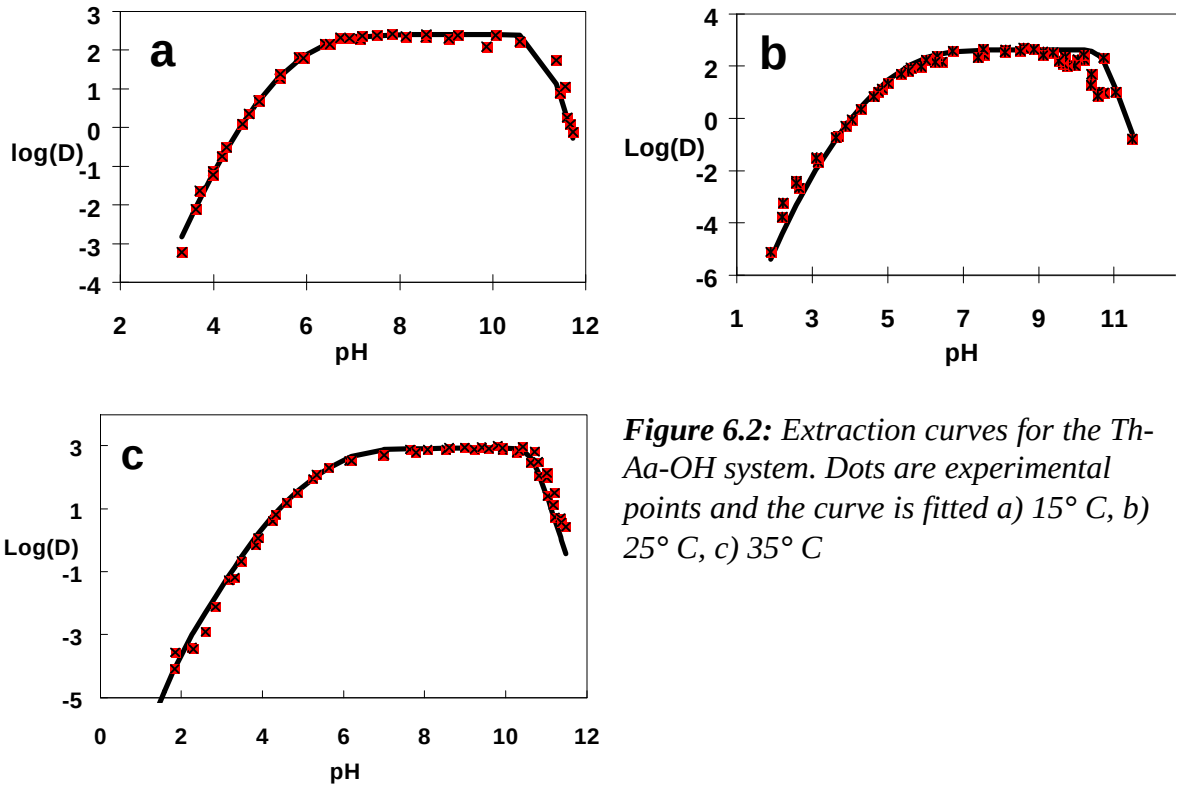



***Figure 6.2:*** *Extraction curves for the Th-Aa-OH system. Dots are experimental points and the curve is fitted a) 15° C, b) 25° C, c) 35° C*

The great deviation of the fitted curve from some of the data at the right part of Figure 6.2, b was caused by experimental errors in one of the experiments. However, since the experiment was correct up to a certain point, it has been included in the figure and used for obtaining the acetylacetone data. It was not used for the hydroxide data, however.The stability constants obtained for the thorium-acetylacetone complexes are shown in Table 6.2.

***Table 6.2:*** *log($b_i$)for the Th-Aa system*

| Temp. / Complex | 15°C | 25°C | 35°C | [ENG 92], 25°C | [RYD 53], 25°C |
|---|---|---|---|---|---|
| $ThAa^{3+}$ | 9.4±0.1 | 9.0±0.2 | 8.8±0.1 | 8.2±0.6 | 7.7 |
| $ThAa_2^{2+}$ | 16.5±0.3 | 16.7±0.6 | 17.1±0.5 | 15.4±0.6 | 14.9 |
| $ThAa_3^{+}$ | 22.2±0.5 | 22.8±0.6 | 23.5±0.5 | 22.6±1.9 | 20.8 |
| $ThAa_4$ | 26.7±0.4 | 27.4±0.2 | 27.9±0.4 | 25.8±1.2 | 25.7 |

The results obtained in this study are in good agreement with the literature for an experiment performed at similar conditions, i.e. [ENG 92] and in fair agreement with a study made under slightly different conditions [RYD 53], as seen in Table 6.2. It is noteworthy that if the values obtained by Rydberg are adjusted by a constant, i.e. addition of 1.1 to the logarthm of the stability constant, the values are in better agreement with the ones obtained in this study. The motivation for such an addition is the problem of determining the first stability constant by the solvent extraction technique, as described in Appendix 3. The stability constants at the different temperatures may be used to calculate other thermodynamic entities, e.g. enthalpies ($\Delta H^0$) and enthropies ($\Delta S^0$) of the formation reaction and heat capacity ($\Delta c_p$). This may, as explained in Section 5.3.3, be accomplished in two ways. Either $\Delta H^0$ and $\Delta S^0$ are temperature independent in the temperature region investigated or they may have a temperature dependence of $\Delta c_p$, which in turn is assumed to be temperature independent in the selected region. The two approaches give results according to Table 6.3.

***Table 6.3:*** *$DH^0$, $DS^0$ and $Dc_p$ for the formation of the different Th-Aa complexes*

| Entity / Complex | $\Delta H^0$ ,linear (kJ/mol) | $\Delta H^0$ (kJ/mol) | $\Delta S^0$ ,linear (J/mol) | $\Delta S^0$ (J/mol) | $\Delta c_p$ (kJ/mol) |
|---|---|---|---|---|---|
| $ThAa^{3+}$ | -55.4±11.1 | -54.4 | -12.1±37.6 | -10.0 | 3.54 |
| $ThAa_2^{2+}$ | 50.7±10.7 | 51.8 | 491.3±36.2 | 493.5 | 3.74 |
| $ThAa_3^{+}$ | 110.2±7.0 | 110.9 | 807.3±23.6 | 808.7 | 2.44 |
| $ThAa_4$ | 102.0±7.8 | 101.3 | 865.9±26.3 | 864.4 | -2.71 |

Clearly the values obtained by solving the equation system, Equations 5.28,, 5.29, and 5.30, lie within the uncertainty intervals given by the linear approach. However visualising the results may shed some light on which approach to believe in, see Figure 6.3.

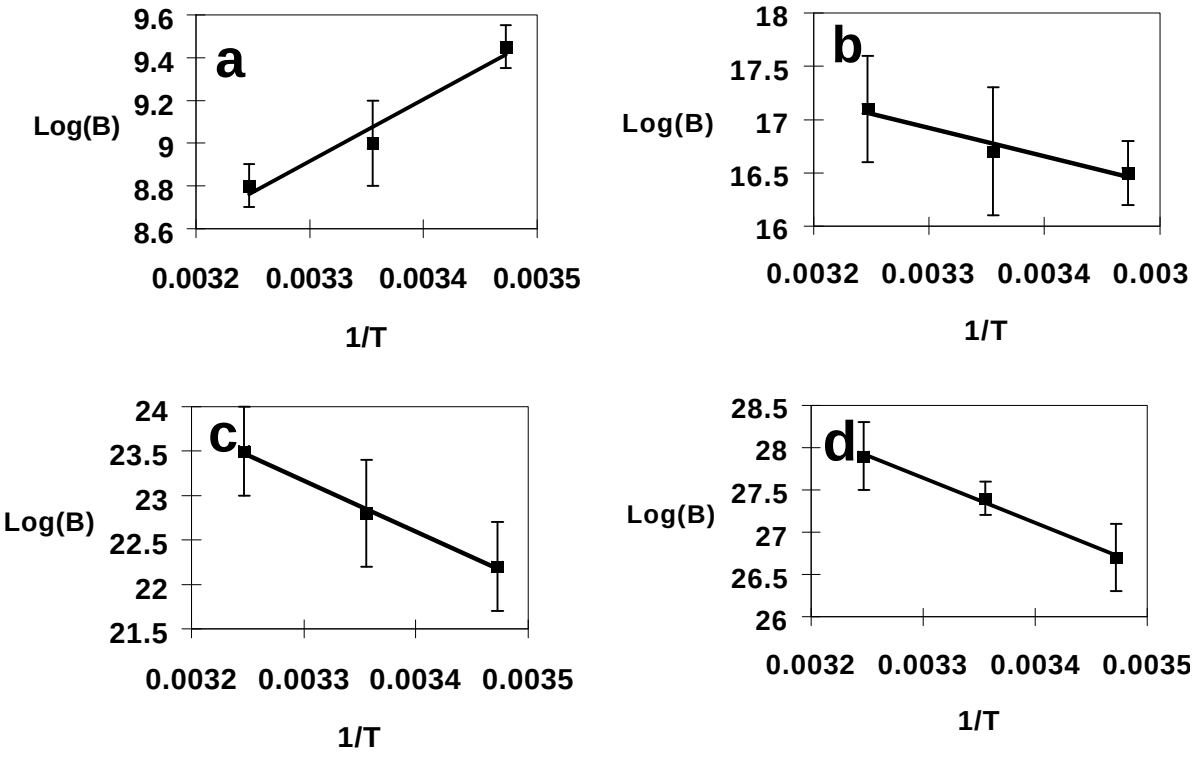



***Figure 6.3:*** *log($b_i$) vs. 1/T for a), $ThAa^{3+}$ b), $ThAa_2^{2+}$ c), $ThAa_3^{+}$ d), $ThAa_4$*

As seen in Figure 6.3, the error bars for one standard deviation all fall on the straight line. Thus it is not possible to conclude that there is a temperature dependence visible for the enthaly and the enthropy in the selected region.

## 6.2 The hydroxide system

The data for the hydroxide system were obtained in collaboration with Dr. P. Brown and Dr. J. Comarmond at ANSTO in Australia since their method of evaluation is better for e.g. determination of poly-nuclear species, as described in Section 5.3.2. It can also be noted that, as seen in Paper VII, the different laboratories obtained almost the same values for the stability constants measured in both laboratories.

The (1,1) complex constants were obtained by potentiometric titrations and evaluated using both linear curve fit and the ANSTO approach, see Section 5.3.2. Only the linear curve fit is discussed here but the value given for the (1,1) stability constant is an average of the different results.

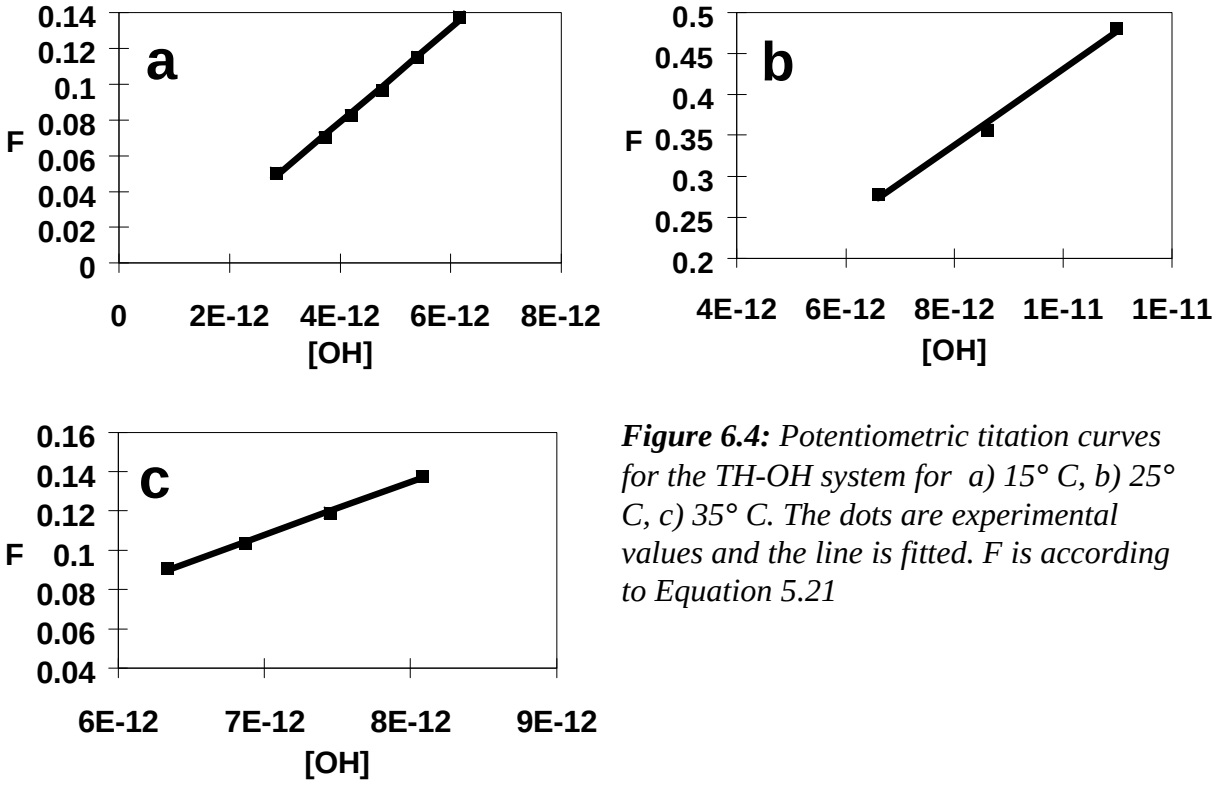



***Figure 6.4:*** *Potentiometric titation curves for the TH-OH system for a) 15° C, b) 25° C, c) 35° C. The dots are experimental values and the line is fitted. F is according to Equation 5.21*

The linear approximation is good, as seen in Figure 6.4. This indicates that it is actually the (1,1) complex that is dominant in that region.

The stability constants for the (1,2), (4,8) and (6,15) complexes were determined at ANSTO by potentiometric titrations, mainly because of their evaluation method which makes it possible to detect poly-nuclear complexes with good accuracy. In addition the determination of the stability constant for the (1,2) complex had to be determined at a lower thorium concentration than used at Chalmers, $<10^{-2}$ M. At ANSTO, a thorium concentration of $<10^{-3}$ M was used, as seen in Paper VII (cf. Appendix 6). There have been, however, problems reported with determination of the stability constant for the (1,1) complex.

Solvent extraction was used to determine the stability constants for the (1,3) and (1,4) complexes, as seen in Figure 6.2. However, there was great difficulty in obtaining a value for the third stability constant, as is evident from its large standard deviation in Table 6.4. This is a result of the large scatter of data on the right part of Figure 6.2. As stated earlier, when Figure 6.2b is examined more closely, it is clear that only some of the obtained data were used for the evaluation of the Th-OH stability constants. This is due to experimental error that occurred in the other experiment. Hence, those used are the ones assumed most correct. The presence of poly nuclear complexes and precipitation of the $Th(OH)_4$ species were neglected because of the low Th concentration, i.e. well below $10^{-5}$ M [ENG 93]. The validity of this statement is discussed below.

***Table 6.4:*** *Log($b_i$)for the Th-OH system*

| Temp. / Complex | 15°C | 25°C | 35°C | Method |
|---|---|---|---|---|
| $ThOH^{3+}$ | -3.6±0.2 | -3.3±0.2 | -3.17±0.1 | Pot. |
| $Th(OH)_2^{2+}$ | -8.8±0.1 | -8.6±0.1 | -8.36±0.2 | Pot. |
| $Th(OH)_3^{+}$ | -13.9±2.8 | -14.3[a] | -12.74±3.5 | Sx. |
| $Th(OH)_4$ | -22.0±0.4 | -19.4±0.5 | -17.8±0.4 | Sx. |
| $Th_4(OH)_8^{8+}$ | -20.2±0.3 | -19.1±0.2 | -18.0±0.1 | Pot. |
| $Th_6(OH)_{15}^{9+}$ | -41.4±0.3 | -39.5±0.2 | -36.6±0.3 | Pot. |

a, Calculated by linear interpolation from the values at 15° C and 35° C

The values for the stability constants presented in Table 6.4 give the following figures for the evaluation of the enthalpy, entropy and the heat capacity.

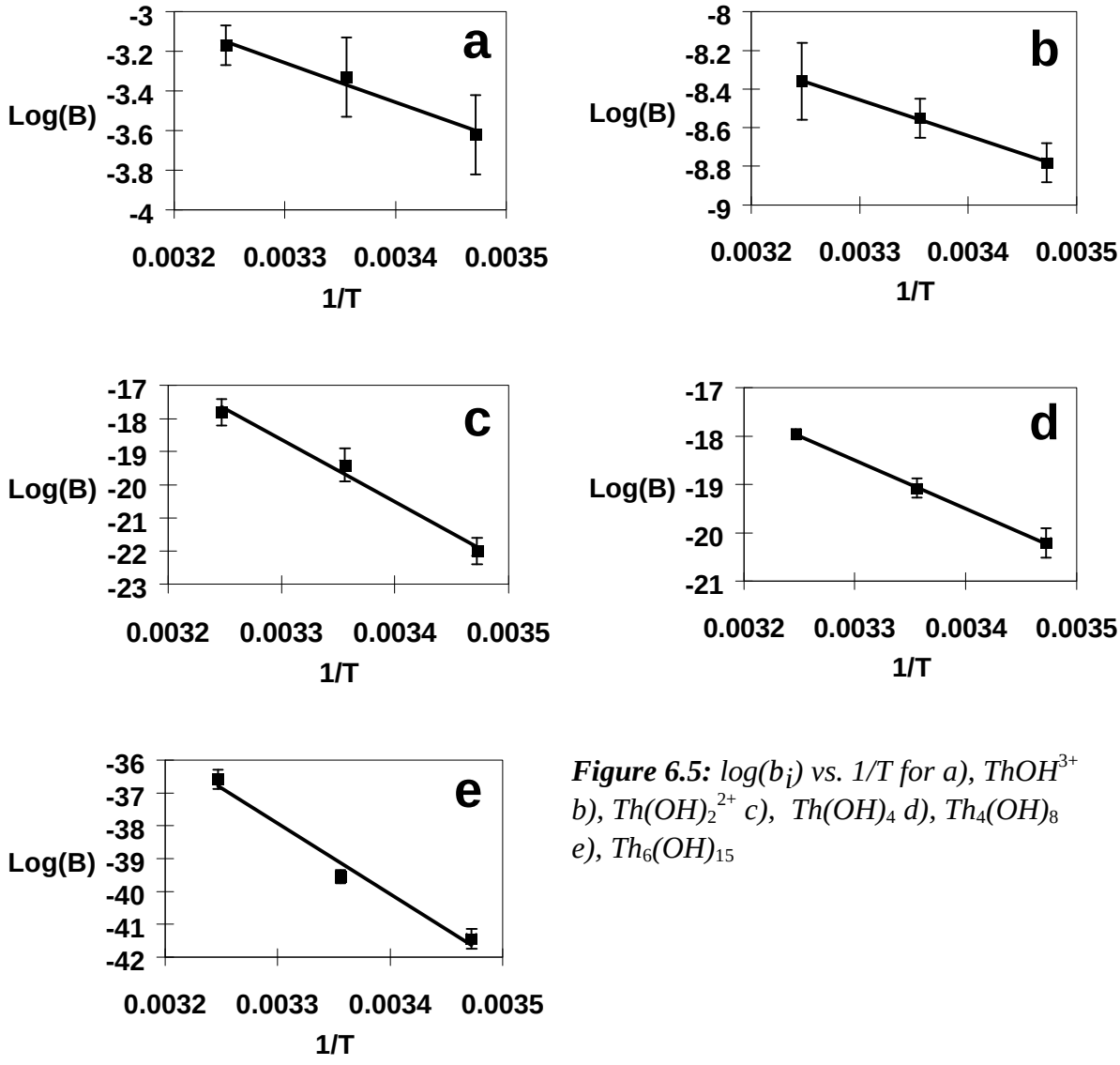



***Figure 6.5:*** *log($b_i$) vs. 1/T for a), $ThOH^{3+}$ b), $Th(OH)_2^{2+}$ c), $Th(OH)_4$ d), $Th_4(OH)_8$ e), $Th_6(OH)_{15}$*

The values obtained from the fitting procedures are given in Table 6.5.

***Table 6.5:*** *$DH^0$, $DS^0$ and $Dc_p$ for the formation of the different Th-OH complexes*

| Entity / Complex | $\Delta H^0$ ,linear (kJ/mol) | $\Delta H^0$ (kJ/mol) | $\Delta S^0$ ,linear (J/mol, K) | $\Delta S^0$ (J/mol,K) | $\Delta c_p$ (kJ/mol) |
|---|---|---|---|---|---|
| $ThOH^{3+}$ | 38.3±5.6 | 37.8 | 64.1±18.9 | 63.0 | -1.95 |
| $Th(OH)_2^{2+}$ | 35.7±12.7 | 35.6 | -44.1±4.3 | -44.3 | -0.44 |
| $Th(OH)_3^{+}$ | 270.7 | ---- | 634.0 | ---- | ---- |
| $Th(OH)_4$ | 357.4±42.2 | 353.3 | 822.5±141.1 | 814 | -1.46 |
| $Th_4(OH)_8$ | 191.0±3.2 | 191.3 | 276.0±10.7 | 276.6 | 1.11 |
| $Th_6(OH)_{15}$ | 411.5±60.0 | 417.3 | 631.7±201.3 | 643.4 | 20.79 |

The stability constants presented in Table 6.4 give a speciation of the Th-OH complexes at different pH according to Figure 6.6.

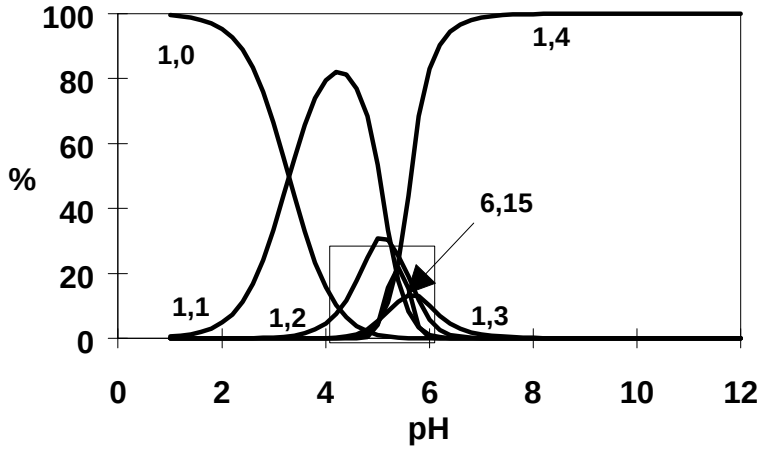



***Figure 6.6:*** *Distribution of the Th-OH system at [Th]=$10^{-5}$ M, T=25° C.*

It may seem suspicious that there is a cluster of species at about pH=5.5. This is however in agreement with other studies, e.g. [GRE 91], even though their study gave slightly different values for the stability constants. The square indicated in Figure 6.6 has been magnified for sake of clarity.

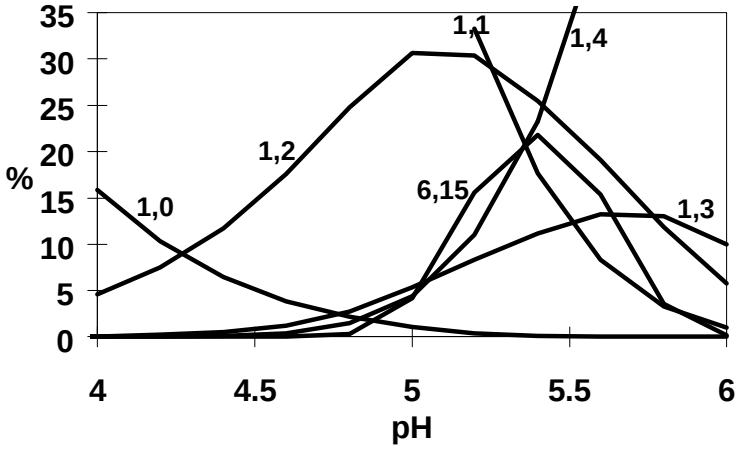



***Figure 6.7:*** *Distribution of the Th-OH system at [Th]=$10^{-5}$M,T=25°C; detail of Figure 6.6*

It is clear in from Figures 6.6 and 6.7 that the assumption of no poly-nuclear species forming in the solvent extraction experiments is valid. At least in the region where the stability constants for the (1,3) and (1,4) complexes are determined.

## 6.3 The phosphate system

The main problem in the evaluation of the phosphate system was the large scatter of the data. Compared with the other systems investigated in this report, this scatter gives rise to rather large uncertainty intervals in the determination of the stability constants, as seen in the figures and tables below. In this case, the solvent extraction method was the only one used and thus it was necessary to evaluate all data from the graphs shown in Figure 6.8.

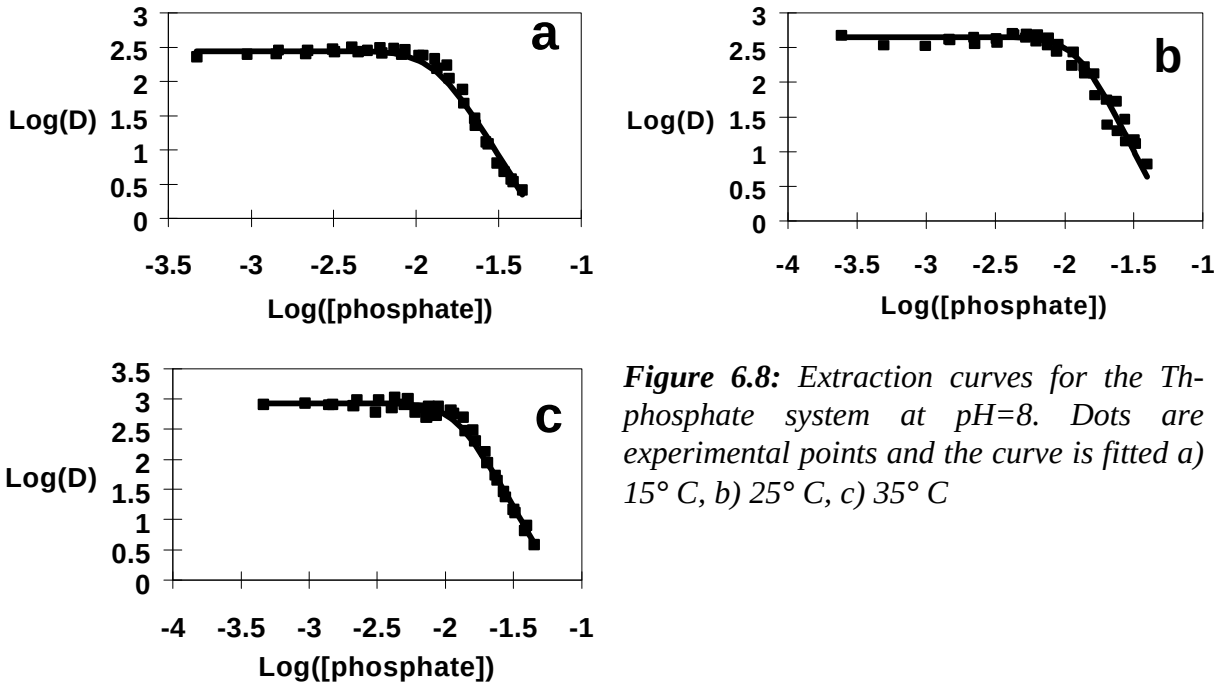



***Figure 6.8:*** *Extraction curves for the Th-phosphate system at pH=8. Dots are experimental points and the curve is fitted a) 15° C, b) 25° C, c) 35° C*

The scatter of data is most apparent in the 25°C case, why it has not been possible to estimate the values for the first stability constant for this case, see Table 6.8.

***Table 6.8:*** *Stability constants obtained for the species, based on reactions R1-R4 in 1.0M $NaClO_4$*

| Species | 15°C, Log(β)±σ | 25°C, Log(β)±σ | 35°C, Log(β)±σ | [ELY 90], 25°C |
|---|---|---|---|---|
| $ThH_2PO_4^{3+}$ | 8.8±5.1 | 7.9[a] | 7.1±5.4 | ---- |
| $Th(H_2PO_4)_2^{2+}$ | 15.5±2.1 | 16.2±2.4 | 17.0±2.4 | 9.65 |
| $Th(H_2PO_4)_3^{+}$ | 19.5±1.4 | 19.6±1.6 | 19.3±1.3 | ---- |
| $Th(H_2PO_4)_4$ | 28.1±0.2 | 28.9±0.3 | 29.4±0.1 | 11.86 |

a, Calculated by linear interpolation from the values at 15° C and 35° C

As seen from Table 6.8, it was difficult to get a value for the first stability constants also for the 15° C and 35° C cases. However, the values given are the results of a minimisation and are thus better than nothing. The large confidence intervals are a result of the fact that the influence of that parameter is very small in the region where good data exist. The enthalpy and

entropy presented in Table 6.9 are thus only presented for the linear evaluation for the constants where only two measured points exist. To check the assumptions made here, i.e. concerning which species participate in the complex formation, the data from Engkvist *et al.* have been plotted together with the theoretical line obtained from the stability constants presented here, see Figure 6.9.

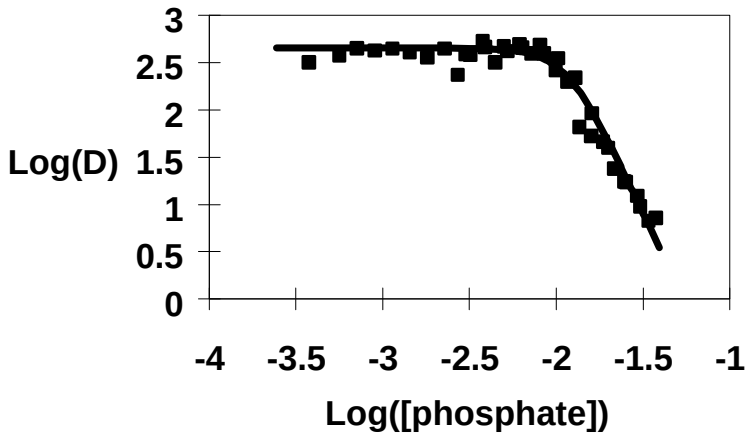



***Figure 6.9:*** *Comparison between data from Engkvist et al. at 25°C, pH 8 (dots) and the model adopted in this report (solid line)*

It is clear from Figure 6.9 that the fit is good and thus enhances the probability that it actually is $H_2PO_4^-$ that participates in the complex formation with thorium, at least at pH 8.

Elyahyaoui *et al.* [ELY 90], cf. Table 6.8, give data for the (1,2) and (1,4) species only. However, the values for the stability constants for these species are lower than those obtained in this work. This may be due to a strong complex with $PO_4^{3-}$ determined by [ELY 90]. Fraction of $PO_4^{3-}$ at the pH investigated is about $10^{-12}$ % of the total phosphate concentration. Quantification of a $ThPO_4^+$ complex is then difficult. Existence of this complex, with the stability constant given, can not be verified by our experiments. If that species had been present in accordance with its stability constant, the strip from the organic phase would have occurred at much lower phosphate concentrations than obtained. Further calculations have been made using all the stability constants from [ELY 90] and the result is shown in Figure 6.10.

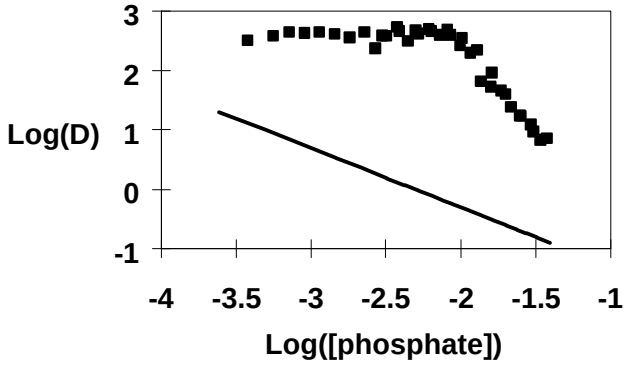



***Figure 6.10:*** *Experiments data from this report ( dots), together with a calculated extraction behaviour using data from [ELY 90] (solid line) at T=25° C .*

There is no agreement between the constants from [ELY 90] and the experiments presented here as seen in Figure 6.10. However, as stated earlier, an interesting fact shown in [ELY 90] is that they were only able to find the (1,2) and (1,4) species. This finding is in accordance with the data presented here. The data shown in Table 6.8 indicate that the (1,2) and (1,4) complexes dominate - probably due to the complexing mechanism. If the complexing behaviour of a molecule fairly similar to $H_2PO_4^-$, namely HDEHP ( Di(2-ethyl hexyl)phosphoric acid) is examined, the structure found in Figure 6.11 is obtained.

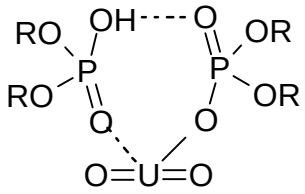



***Figure 6.11****: Structural formula of a Uranyl(VI)-HDEHP complex [RYD 92]*

It must, however, be noted that HDEHP readily form dimers in non-polar solution [MAR 69]. Thus there may be a slight difference in the structural formula for $H_2PO_4^-$ complexation, since the formation constant for formation of dimers is 0.12 [ELY 90]. Thus dimer formation at the concentrations used in this work is negligible. The suggested structural formula for $Th(H_2PO_4)_2^{2+}$ is shown in Figure 6.12.

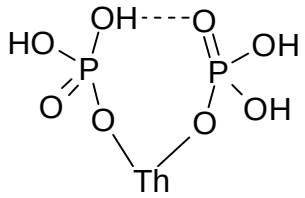
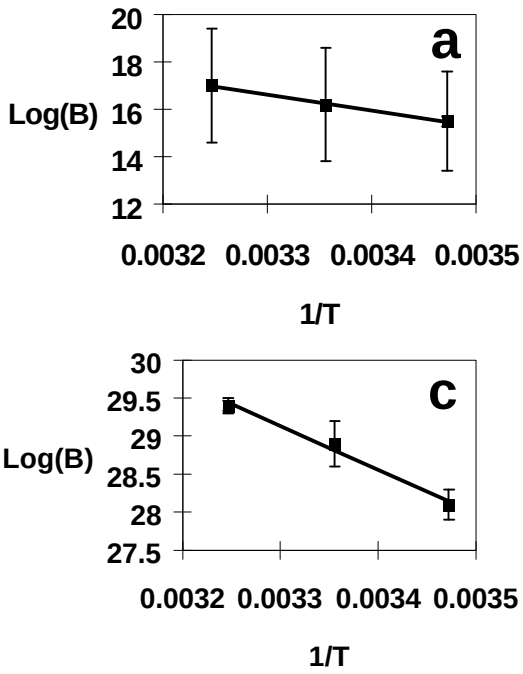



***Figure 6.12:*** *Structural formula of $Th(H_2PO_4)_2^{2+}$*

Naturally the evidence for a complex such as in Figure 6.12 must be found by other methods than speculation and solvent extraction, e.g. spectroscopic methods.

The enthalpy and entropy were calculated using the same assumptions as described earlier. However, owing to the problems in obtaining data at all the investigated temperatures, only three cases are shown in Figure 6.13. In the other case, only two values were obtained, thus making the linear fitting trivial.



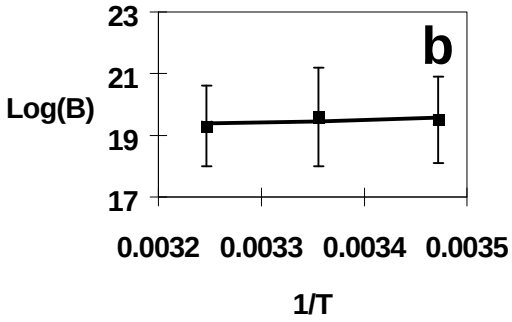







***Figure 6.13:*** *log($b_i$) vs. 1/T for a), $Th(H_2PO_4)_2^{2+}$ b), $Th(H_2PO_4)_3^{+}$ c), $Th(H_2PO_4)_4^{\circ}$*

Clearly, if a straight line is fitted through two points, there are no degrees of freedom left to give an uncertainty interval. Some of the data in Table 6.9 thus lack uncertainties. They are included with the value obtained with the two endpoints and it might thus be argued that the accuracy is not good. However, these are the values obtained. Whether one wishes to trust them is up to the user and depends on hers or his intentions.

***Table 6.9:*** *$DH^0$, $DS_0$ and $Dc_p$ for the formation of the different Th-phosphate complexes*

| Entity / Complex | $\Delta H^0$ ,linear (kJ/mol) | $\Delta H^0$ (kJ/mol) | $\Delta S^0$ ,linear (J/mol) | $\Delta S^0$ (J/mol) | $\Delta c_p$ (kJ/mol) |
|---|---|---|---|---|---|
| $ThH_2PO_4^{3+}$ | -144.3 | ---- | -332.4 | ---- | ---- |
| $Th(H_2PO_4)_2^{2+}$ | 127.2±7.2 | 130.3 | 737.0±24.7 | 747.5 | 1.72 |
| $Th(H_2PO_4)_3^{+}$ | -16.5±20.0 | -21.2 | 316.9±66.9 | 303.7 | -7.44 |
| $Th(H_2PO_4)_4$ | 110.6±12.5 | 106.7 | 922.8±42.2 | 911.4 | -4.88 |

The stability constants given in Table 6.8, yield a distribution of the thorium phosphate species at pH= 8, according to Figure 6.14. It must be noted that this figure only gives a distribution of the species concerned, i.e. it does not give a representation of the total thorium distribution.

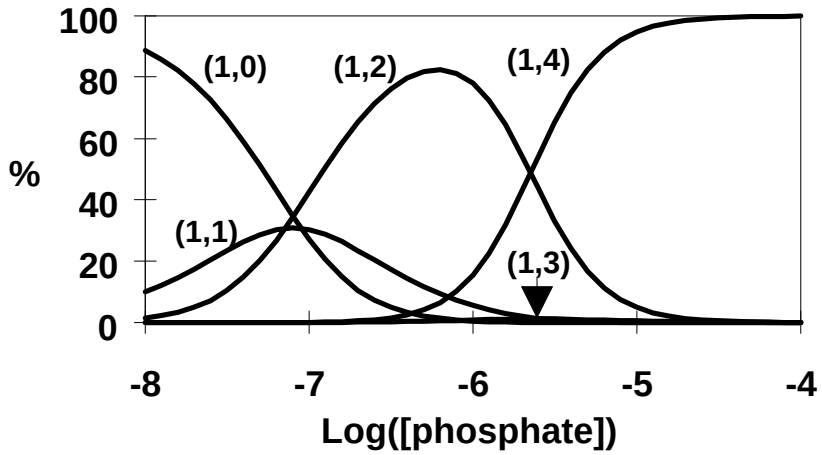



***Figure 6.14:*** *Distribution of the Th-phosphate system at [Th]=$10^{-7}$M and pH=8, hydroxide species omitted*

Another method of showing speciation is a logarithm diagram. In some cases, this actually gives a better view of how a system looks. Such a diagram is shown for the phosphate system at pH 8 in Figure 6.15. The dominant hydroxide species is also included to make the thorium concentration correct over the whole range.

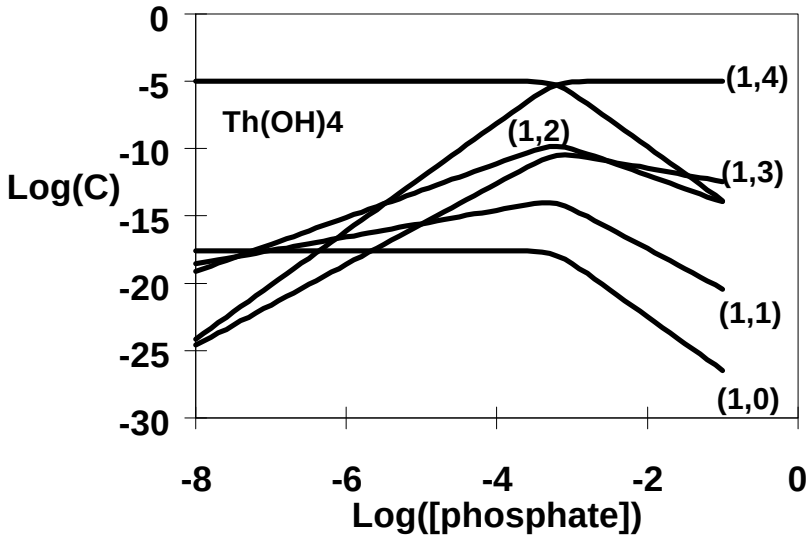



***Figure 6.15****: Speciation of the Th-phosphate system at [Th]=$10^{-5}$ M and pH=8*

It is clear from examination of Figures 6.14 and 6.15 that the determination of the first phosphate complex is almost impossible due to their low abundance at phosphate concentrations from about $10^{-4}$ M and upwards. Another problem is that acetylacetone is a very good complexing agent for thorium. Thus the phosphate concentration must be high to be able to compete with the acetylacetone (1,4) complex in order to achieve any strip from the organic phase. This may be solved with a weaker complexing agent or perhaps with a lower concentration of the extractant. However, given that the constants obtained here are correct, there would still be a problem in determining the stability constant for the (1,1) complex unless a very dilute phosphate solution could be added.

## 6.4 Uncertainties in stability constants

When discussing uncertainties in stability constants it may be interesting to make a visualisation on how, for example, a speciation figure will look as the uncertainty intervals for the stability constants are taken into account. Ideally it is easy to imagine that the effect will only be thin bands around the lines in the speciation figures. Unfortunately this is not generally the case. To illustrate this, a calculation of uncertainty intervals have been added to the original figures (6.14 and 6.15) to give Figures 6.16 and 6.17.

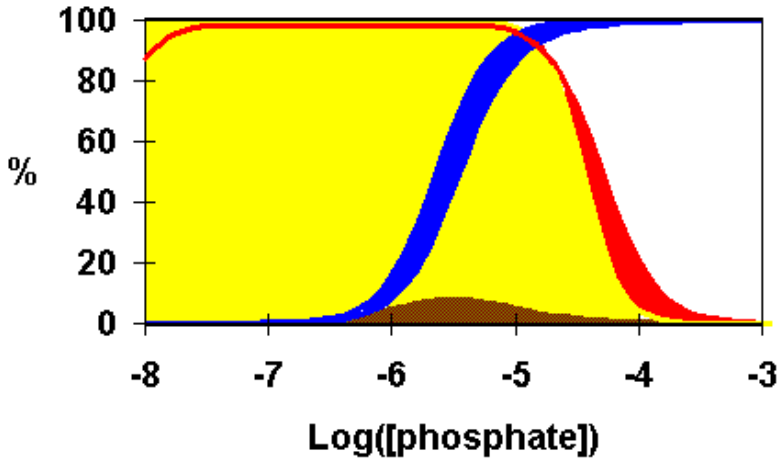



***Figure 6.16:*** *Distribution of the Th-phosphate system at [Th]=$10^{-7}$M and pH=8 with uncertainty intervals, (1,1)-yellow, (1,2)-red, (1,3)-brown, (1,4)-blue, the red part is partly covered by the yellow one.*

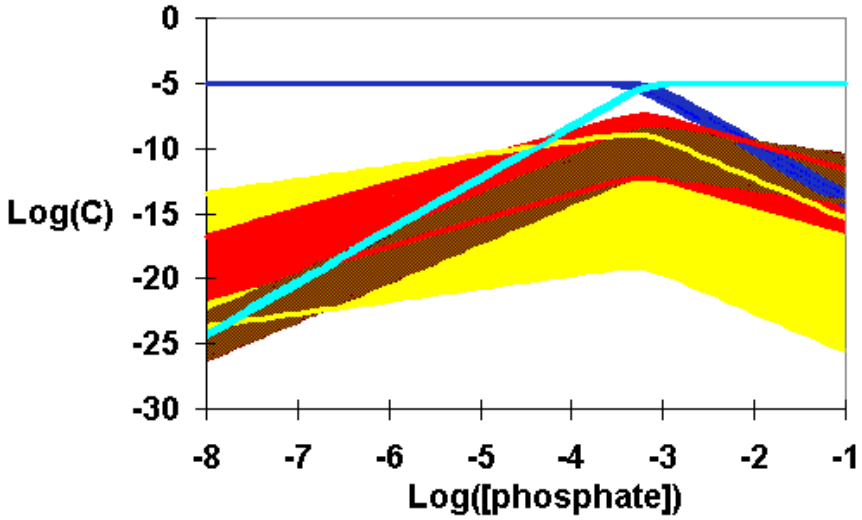



***Figure 6.17:*** *Speciation of thorium at pH 8 with uncertainty intervals, (1,1)-yellow, (1,2)-red, (1,3)-brown, (1,4)-cyan, $ThOH_4^\circ$-blue*

A conclusion to be drawn from Figures 6.16 and 6.17 is that speciation diagrams may be useful when dealing with systems in which there are small uncertainties in the stability constants. For many systems, however, they give a far too optimistic impression. The concentrations should instead be given as bands, as in Figure 6.17, which gives a realistic impression of the situation.

Further, it is clear from Figure 6.16 that a percentual distribution of species is not a good way of displaying the effect of uncertainties in stability constants. The only conclusion that can be drawn from Figure 6.16 is that nothing much may be said about the distributions of the species.

As described above, the determined stability constants are given with one standard deviation uncertainty interval. These intervals have been determined assuming that the stability constants are independent. This might be true for the determination of the hydolysis constants, but not for the others. There is a mutual dependence that may need some thought.

When constants are determined by fitting, they should all be investigated at one time to find the global minimum for the difference between experiments and the fitted equation using some norm, e.g. least squares. Then it is possible to give confidence limits to the fitted parameters by some method, e.g. as described above. If one variable is fixed at its, say, highest limit and the fitting process is made once more, new results may be obtained thus making it possible to create a correlation matrix. In some cases the correlation is called covariance but these entities are similar in repect to what they describe. In the cases presented here it is possible to make a correlation matrix for the phosphates and the acetylacetone complexes, see Table 6.10. For the phosphate system it was not possible to investigate this correlation for more than the two highest complexes since the influence of the first two species was too small. Because the highest goodness-of-fit was obtained for the 35° C case this situation has been used for the correlaion analysis.

***Table 6.10:*** *Correlation between stability constants*

| | $ThAa_2^{2+}$ | $ThAa_3^{+}$ | $ThAa_4$ | $Th(H_2PO_4)_3^{+}$ | $Th(H_2PO_4)_4$ |
|---|---|---|---|---|---|
| $ThAa_2^{2+}$ | ----- | 1.4 | 1.2 | | |
| $ThAa_3^{+}$ | 0.9 | ---- | 0.9 | | |
| $ThAa_4$ | 0.8 | 1.3 | ---- | | |
| $Th(H_2PO_4)_3^{+}$ | | | | ---- | 2.6 |
| $Th(H_2PO_4)_4$ | | | | 0.1 | ---- |

The correlation given in Table 6.10 shows by which factor the logarithm for the other constants vary as the constant at the head of the column of the table is changed. The first stability constant for the acetylacetone system is not used in the correlation calculations since it has been obtained using another method. As described in Appendix 3, inclusion of the first stability constant in the correlation calculation would only result in a correlation factor of 1 between all constants.

The results shown in Table 6.10 show that it is not possible to treat the stability constants as independent variables in an uncertainty analysis. The reason for that it is done so in this report is that the correlations beween the stability constants are not generally known, thus making an effort pointless. The assumption of independent stability constants probably overestimates the effect of uncertainties in them.

## 7. Uncertainty calculations in the Th-phosphate system

Calculations using available data for the thorium phosphate system sometimes show that the solubility of thorium in some groundwaters is unusually high due to the strong complexation with phosphates. This has not been observed experimentally, which is why thorium phosphates are sometimes excluded from the calculations to obtain more realistic results [ENG 94].

Effect of the previous phosphate data vs. the data presented in this report was investigated using the SENVAR program package. In these calculations both the stability constants for the hydrolysis and the phosphate complexation have been changed. The water selected was a phosphate-rich groundwater presented in KBS 3 [KBS 83]; to further enhance the effect of phosphate complexation, the solid phase selected for equilibration was $Th(HPO_4)_{2(s)}$. This is not the solubility limiting phase in this system but was selected to demonstrate the effect of thorium phosphates. The resulting distribution functions are shown in Figure 7.1. Width of the different intervals depend on the uncertainty intervals selected for the stability constants. In case a, the uncertainty intervals were the ones obtained in this report , while the uncertaity intervals for case b were 0.5 log units wide.

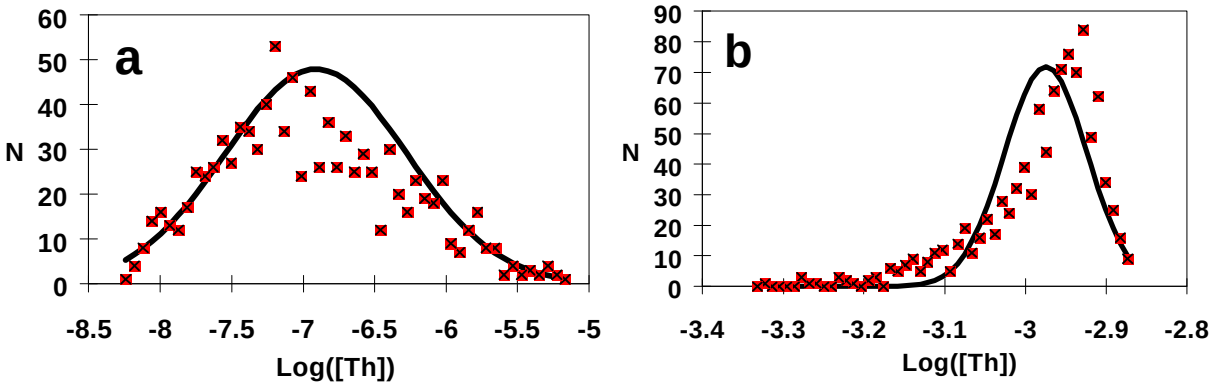



***Figure 7.1:*** *Distribution functions for the calculated solubility of $Th(HPO_4)_2$ using a) stability constants obtained in this report, b) data given in Hatches 5.0 [BON 92],[MOS 67]*

It is seen from Figure 7.1 that there is a large deviation in the calculated mean solubility for the different data sets. It is important to remember that only the data for the thorium hydroxides and phosphates that were changed and, in the case of the hydroxides, the change was almost within the uncertainty limits. The dominating species in case a) is $Th(OH)_4^\circ$ and in case b) $Th(HPO_4)_2^\circ$ and $ThHPO_4^{2+}$.

It may thus be concluded that, with the stability constants obtained in this study, the phosphate complexes do not enhance the calculated solubility of thorium in most Swedish groundwaters.

# 8. Summary and conclusions

The work presented here consists of two parts, the first dealing with how to handle uncertainties in solubility calculation and the second dealing with the determination of stability constants for the Th-Aa-OH-phosphate system. In the second part, an effort has been made to discuss the uncertainties associated with such determinations.

It is clear from the studies presented here that even such a simple task as calculating the solubility of a solid phase may not be easy and the results may be affected by large uncertainties. The input uncertainties investigated had of course different significances when changed within reasonable intervals. The most important factor in the uncertainties in calculated solubilities for groundwater conditions are the conceptual uncertainties, i.e. how to actually perform the calculation. As these are the most difficult to quantify, the remedy is simply to make as many calculations as possible with different approaches and then try to rule some of them out. The remaining difference is the uncertainty interval.

If the method for calculation has been decided, the uncertainties in the input data are the next problem. One of the most important parameters to determine accurately is, not surprisingly, the stability constants of the species deemed important in the sensitivity analysis, i.e. stability constants that actually influence the calculations. The sensitivity analysis may be found necessary since it can not always be judged intuitively which stability constants are important. It is not necessarily the dominating ones. The change in some important stability constants by 0.5 log unit may result in an uncertainty in the solubility of about two orders of magnitude. Thus minimisation of the uncertainties in the stability constants is important to reduce the uncertainty of a calculated solubility. At present, it is unusual to find reasonable uncertainties in stability constants in the literature.

If the calculations are made for temperatures other than the one used for the measurement of the stability constants, the enthalpy of reaction for the species deemed important is crucial. An error range in the $\Delta H$ of 4 kcal/mole for the important species has been shown to give solubilities that differ by as much as one order of magnitude for calculations at 80°C.

Proceeding down the ladder of uncertainties, the next uncertain parameter is the water composition. This composition may be obtained in two ways, either by simulation of rock-water interactions or by measurements. In the first case, the uncertainty in rock composition, apart from conceptual uncertainties, will give rise to uncertainties in water composition. In the latter case, it is the measurement uncertainties that enter the calculations. Both these cases were investigated and it was found that, compared with the other uncertainties, the water composition did not represent great contribution.

In the experimental part, stability constants for the Th-Aa-OH-phosphate system were determined. This was done for three different temperatures, and the enthalpy and entropy of reaction were thus also determined. Since there are three measurements at different temperatures, calculation of the heat capacity is also possible, although the significance of these results may not be great owing to a lack of degrees of freedom.

The acetylacetonate system and the hydroxide system are evaluated using the standard system descriptions. This is not so for the Th-phosphate system, however. The common idea is that complexation at pH 8 occurs with $HPO_4^{2-}$ only. In this work, it has been assumed that the complexation with $H_2PO_4^-$ dominates. This theory was validated by experiments at pH 7 and pH 9. Old experiments were also used in the validation procedure.

The impact of the new data for the Th-phosphate system is that thorium phosphate complexes do not dominate under natural conditions at pH 8. This was shown by calculations made with earlier data and comparison with the new ones.

The correlation between the determined stability constants was also investigated for some cases, and it was found that these correlations are not negligible. This means that, if one stability constant is changed, others also change owing to the dynamics of the studied system. This effect of this is that it is not possible to treat stability constants as independent variables in an uncertainty analysis. However, it is possible to argue that assuming that the stability constants are independent only widens the uncertainty interval for the calculated solubility, thus making the result more conservative.

Finally, it is clear that, for reliable predictions to be made by computer simulations, the uncertainties in the input parameters must be thoroughly investigated and given as intervals rather than as single values, thus producing an interval as the result of the simulation. In addition, the selection of conceptual model may be crucial.

# 9. Acknowledgement

There are many people involved in the process of making this report. Some are directly related to the paperwork (hopefully) just read, and some have made their contribution by making my time pleasant.

In the first group, my thoughts go to one of my supervisors Allan Emrén, and Anna Nyström-Claeson for being directly responsible for my starting this work at all. As with so many good things in life, a multitude is best. I have during this time had the oppurtunity of having severeal people guiding me in different areas of nuclear chemistry. One of them, Allan Emrén is already mentioned for starting me in this field and being my primary supervisor. He has also been a great support in theorethical subjects. For the experimental work I am indebted to Ingemar Engkvist for starting me off and to Yngve Albinsson for skilful guidiance through the experimental jungle and always being there for more or less silly questions and many laughs. Finally, in the supervisor section, I wish to express my gratitude to Jan-Olov Liljenzin for his expertise in a multitude of subjects and many funny parties.

In the final phase of this work, i.e. the writing of this report, I have used the help of many people. First, I am greatly indebted to Gunnar Skarnemark for giving me so much of his time in the reading of this report and to give sound criticism to this work and for giving me an oppurtunity to learn more about the heaviest elements. The supervisors, mentioned above have made a great job in making a rough skeletton to a final report, for which I am grateful.

There are also some "external" people that have given their aid and I would like to mention Dr. Stig Wingefors, Dr. Bo Strömberg and Madelene Lindqvist as being major contributants. Dr. Sanne Börjesson is also acknowledged for making my time at the department pleasant and for reading and making suggestions about the theis in spite of time pressure on a new job. I would also express my gratitude to prof. Ingemar Grenthe for aways promptly answering my e-mail questions, regardless where on earth he was.

A part of the work presented in this report have been made by Anders Samuelsson and Arvid Ödegaard-Jensen and without their help I may still be busy programming.

Being not a linguist I have had help with correcting and writing english and german, thus I would like to acknowledge the help of Michael Koucky and Janet Vesterlund. Further, I would like to thank everyone at the Department of Nuclear Chemistry for being a cosy bunch, and together with whom it is not possible to be bored. Singeling some out would be unfair and to mention them all and their specialities and features would take too much room.

Finally, I would like to thank the people close to me, in my spare time, for being a great support. Some which ought to be mentioned are Cecilia, Barbro, Lasse, Kerstin, Rüdiger, Arvid, Henrik, Torulf, Harald, Nisse, Anders, Christina, Madelene, Claes and Eva. All those not mentioned but certainly not forgotten. Thank you.

This work was funded by the Swedish Nuclear Inspectorate (SKI).

# 10. References


AHN 90 Ahn, J., Suzuki, A., "Sensitivity Analysis for a High-Level Radioactive Waste Repository with the Bounding Fracture-Transport Model.", Radioactive Waste Management and the Nuclear Fuel Cycle, Vol.14(4), p. 257-273, 1990.

ALB 88:1 Albinsson, Y., "Development of the AKUFVE-LISOL Techniques. Solvent Extraction Studies of Lanthanide Acetylacetonates.", PhD thesis, Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden 1988.

ALB 88:2 Albinsson, Y., Ohlsson, L.-E., Persson, H., Rydberg, J., "LISOL, a Technique for Liquid Scintillation On-line Measurements", Appl. Radiat. Isot., Vol. 39, No. 2, pp. 113- 120, 1988.

ALB 99 Albinsson, Y., Ekberg, C., Holgersson, S., Landgren, A., Ramebeck, H., "A Method for Preparation and Purification of $^{234}$Th.", in preparation, 1999.

AND 95 Andersson, K., Molander, S., "Kompendium i miljösystemanalys", Technical Environmental Planning, Chalmers Univerity of Technology, 1995.

AND 93 Andres, T. H., Hajas, W. C., "Using Iterated Fractional Factorial Design to Screen Parameters in Sensitivity Analysis of a Probabilistic Risk Assessment Model", Proceedings of the Joint International Conference on Mathematical Methods and Supercomputing in Nuclear Applications, 19-23 April, Karlsruhe, Germany, 1993.

BAT 88 Bates,D.M., Watts, D.G., "Nonlinear Regression Analysis and its Applications" Wiley Series in Probability and Mathematical Statistics, J. Wiley & Sons, Chichester/UK (1988).

BJE 41 Bjerrum, J., "Metal Ammine Formation in Aqueous Solution", PhD thesis, Haase and Son, Copenhagen, Denmark, 1941.

BOD 01 Bodländer, G., Fittig, R., "Das Verhalten von Molekularverbindungen bei der Auflösung II", Z. physik. Chem., Vol. 39, 1901

BON 92 Bond, K. Moreton, A. D., Heath, T. G., "Hatches 5.0 - The HARWELL/ NIREX Database", Harwell Laboratory, Oxfordshire, ORA, UK, 1992.

BOX 87 Box, G. E. P., Draper, N. R., "Empirical model-building and responce surfaces.", John Wiley & Sons Inc., New York, USA, 1987.

BRO 83 Brown, P.L., Ellis, J., Sylva, R.N., "The Hydrolysis of Metal Ions. Part 5. Thorium(IV)." J. Chem. Soc. Dalton Trans. Vol. 31 (1983)

BUC 05 Buchholz, C. F., "Verhalten des Schwefeläthers zu einer gesättigten Auflösung des Salpetersauren Urans und des Wassers zu dem Uranhaltigen Schwefeläther." Neues allgem. J. Chemie, Vol. 4, p. 157, 1805.

BÖR 98 Börjesson, K. S., Emrén, A. T., "PASSIPHIC: A Program for Solubility Calculations Involving Pure Minerals and Solid Solutions", Computers & Geosciences, Vol. 24, No. 9, p. 839, 1998.

CAM 78 Campbell, J. E., et al., "Risk Methology for Geologic Disposal of Radioactive Waste: Interim Report", Rep. NUREG/CR-0458, Sandia National Laboratories, Albuquerque, NM, 1978.

CAM 88 Campbell, J. E., Cranwell, R. M., "Performance Assessment of Radioactive Waste Repositories", Science, Vol 239, p. 1389-1392, 1988.

CHA 95 Chapman, N. A., Andersson, J., Robinson, P., Skagius, K., Wene, C.-O., Wiborgh, M., Wingefors, S., "Systems Analysis, Scenario Construction and Consequence Analysis Definition for SITE-94.", SKI report 95:26, Swedish Nuclear Power Inspectorate (SKI), Stockholm, 1995.

CHO 80 Choppin, G. R., Rydberg, R., "Nuclear Chemistry, Theory and Applications", Pergamon Press, 1980.

COM 87 Combes, M. A., "Nouvelle réaction du chlorure d'aluminium synthéses dans la série grasse, chaptre VI: action du sodium et homologues supérieurs de l'acétylacétone.", Ann. Chim. Phy., p. 245, 1887

CRA 87:1 Cranwell, R. M., Martin, D. J., Peterson, M., Bragg, K., Matsunaga, T., "Overview of the treatment of uncertainties arising in radioactive waste disposal assessments", Proceedings of a NEA Workshop on Uncertainty Analysis for Performance Assessments of Radioactive Waste Disposal Systems, OECD/NEA,1987.

CRA 87:2 Cranwell, R. M., "Uncertainties in performance assessments of high level waste repositories". Trans. American Nuclear Society, 55:308, 1987

EFR 93 Efron, B., Tibshirani, R.J., "An Introduction to the Bootstrap", Monographs on Statistics and Applied Probability 57, Chapman & Hall, London/U.K., 436 (1993)

EKB 95,1 Ekberg, C., "Uncertainties in Safety Analysis, A Literature Survey", SKI report 95:17, Swedish Nuclear Power Inspectorate (SKI), Stockholm, 1995.

EKB 95, 2 Ekberg, C., “SENVAR: a Code for Handling Chemical Uncertainties in Solubility Calculations”, SKI report 95:18, Swedish Nuclear Power Inspectorate (SKI), Stockholm, 1995.

EKB 95, 3 Ekberg, C., "Handling instructions and source codes for the SENVAR package", KKR 95-08-22, Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden, 1995

EKB 98 Ekberg, C., Emrén, A. T., "Uncertainties in solubility calculations aimed at a repository for spent nuclear fuel.", Proceedings of the 4th International Conference on Probabilistic Safety Assessment and Management, 13.18 September, New York City, USA, 1998.

EKB 99:1 Ekberg, C., "MINVAR and UNCCON-handling instructions and source code", Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden, in preparation.

EKB 99:2 Meinrath, G., Ekberg, C., Landgren, A., Liljenzin, J.-O., "Methods for assigning uncertainties to fitted parameters.", submitted

ELY 90 Elyahyaoui, A., Brillard, L., Boulhassa, S., Hussonnois, M., Guillaumont, R., "Complexes of Thorium with Phosphoric Acid.", Radiochimica Acta, Vol. 49, nr. 39-44, p.39, 1990.

EMR 92 Emrén, A. T. , "CRACKER - A program coupling chemistry and transport" Version 92-11, SKI TR 93:6, The Swedish Nuclear Power Inspectorate (SKI), Box 27106, Stockholm, Sweden, 1992.

EMR 95 Emrén, A. T., Personal communication, Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden, 1995.

EMR 98:1 Emrén, A. T., "CRACKER a Program Coupling Inhomogeneous Chemistry and Transport", Computers & Geosciences, under review

EMR 98:2 Emrén, A. T., "Modelling of groundwater chemistry at Äspö Hard Rock Laboratory (SITE-94)", SKI report 96:32, Swedish Nuclear Power Inspectorate (SKI), Stockholm, 1998.

EMR 99 Emrén A. T., Arthur R., Glynn P. D., McMurry J., "The modeler's influence on calculations or performance assessments: A case study of $Pu(OH)_4$ solubility in Hypothetical Environments." SKI technical report, in preparation 1999

ENG 92 Engkvist, I, Albinsson, Y., "Hydrolysis Studies of Thorium Using Solvent Extraction Technique.", Radiochimica Acta, Vol. 58/59, s. 109-112, 1992.

ENG 93 Engkvist, I., "Diffusion and Hydrolysis of some Actinides", PhD thesis, Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden, 1993

ENG 94 Engkvist, I., Albinsson, Y., "Thorium Complexation with Phosphates", Proceedings of the Migration-93 Conference, Charleston, South Carolina, USA, pp. 139-142, 1994.

FAN 96 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 $Na^+/H^+/OH^-/HCO_3^-/ClO_4^-$/H2O at 25°C", Jounal of Solution Chemistry, Vol. 25, No. 4, 1996

GAN 76 Gans, P., Sabatini, A., Vacca, A., "An improved computer program for the computation of formation constants from potentiometric data.", Inorganica Chimica Acta, 18, 237-239, 1976.

GEL 76 Gelhar, L. W., "Effects of hydraulic conductivity variations on ground water flows", paper presented at Proceedings, Second International IAHR Symposium on Stochastic Hydraulics, Int. Assoc. of Hydraul. Res., Lund, Sweden, 1976.

GLY 98 Glynn, P., Voss, C.,"Geochemical charaterization of Simpevarp ground waters near the Äspö Hard Rock Laboratory", SKI report 96:29, Swedish Nuclear Power Inspectorate (SKI), Stockholm, in press, 1998

GRA 50 Gran, G., "Determination of Equivalent Point in Potentiometric Titrations", Acta Chemica Scandinavia, 4, p. 559-557, 1950.

GRA 52 Gran, G., "Determination of Equivalent Point in Potentiometric Titrations. Part II", International Congress on Analytical Chemistry, 77, p. 661-671, 1952.

HAR 83 Harper, W., V., "Sensitivity/uncertainty analysis techniques for nonstochastic computer codes.", ONWI-444, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, OH, 1983.

IMA 80 Iman, R. L., Conover, W., J., "Small Sample Sensitivity Analysis Techniques for Computer Models, With an application to Risk Assessment", Communication in Statistics, A9, 17, p. 1749-1874, 1980.

IMA 85 Iman, R. L., Helton, J. C., "Comparison of Uncertainty and Sensitivity Analysis for Computer Models.", Report NUREG/CR-3904, Sandia National Laboratories, Albuquerque, NM, 1985.

IMA 87 Iman, R. L., "Modeling Inputs to Computer Models Used in Risk Assessment.", Trans. American Nuclear Society, 55:309-310, 1987

INT 83 "Workshop on uncertainty analysis of postclosure nuclear waste isolation system performance.", INTERA Environmental Consultants, Inc., ONWI-419, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, OH, 1983.

INR 92 "The International INTRAVAL Project, Phase 1", The Swedish Nuclear Power Inspectorate (SKI), Stockholm, Sweden, 1992.

JAC 98 Jacobsson, A.-M., "ABU2: users manual and program description.", Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden, in preparation, 1998.

JOH 87 Johnston, P. D., Thompson, B. G. T., "The aims and uses of uncertainty analysis in the context of development and regulation of underground disposal of radioactive waste", Proceedings of a NEA Workshop on Uncertainty Analysis for Performance Assessments of Radioactive Waste Disposal Systems, OECD/NEA, 1987.

KBS 83 Final Storage of Spent Nuclear Fuel - KBS-3. Swedish Nuclear Fuel and Waste Management Co. (SKB), Stockholm, Sweden, 1983

KOR 75 Korin, B. P., "Statistical Concepts for the Social Sciences", Winthorp Inc., Cambridge/USA, 1975.

LIL 69 Liljenzin, J. -O., "The Dissociation Constant for Acetylacetone in 1 M Sodium Perchlorate Solution at 15-45°C", Acta Chem. Scand., vol. 23, p. 3592, 1969.

LIL 95 Liljenzin, J. -O. personal comunication, Department of Nuclear Chemistry, Chalmers University od Technology, Göteborg, Sweden. 1995.

LIL 99 Liljenzin, J.-O., personal comunication, Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden. 1999

MAI 93 Maidment, D. R. (ed.), "Handbook of Hydrology", McGraw-Hill, 1993.

MAR 69 Marcus, Y., Kertes, A.S., "Ion Exchange and Solvent Extraction of Metal Complexes", John Wiley and Sons Ltd., London, U.K., 1969.

MAT 69 Matheron, G., "Le Krigeage Universal", Cah. Centre Morphol. Math., 1, Fontainebleau, France, 1969

MCK 79 McKay, M. D., Beckman, R. J., Conover, W. J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code", Technometrics, Vol. 21, p 239-245, 1979.

MIL 41 Milankovitch, M. M., "Kanon der Erdberstrahlung und seine Anwendung auf das Eiszeitproblem.", Acad. Roy. Serbe, Ed. spec. 133, p 633, 1941.

MOF 88 Moffat, Robert, J. "Identifying the true value - the first step in uncertainty analysis", Instrument Society of America. Test Measurement Division (ISA 34), NC, USA. p 255-258, 1988.

MOS 67 Moskvin, A., Essen, L. N., Bukhtiyarova, T. N., Russ. J. Inorg. Chem. Vol. 12, 1794, 1967.

NEA 92 Nuclear Energy Agency (Grenthe, I., Fuger, J., Konigs, J. M., Muller, A. B., Nguyen-Trung, C., Wanner, H.,) "Chemical Thermodynamics of Uranium", NEA/OECD, 1992.

NER 89 Nernst, W., Z.. Phys. Chem., No. 2, p 613 and No. 4, p 129, 1889.

NIL 92 Nilsson, A-C., "Groundwater chemistry monitoring at Äspö during 1991", Swedish Nuclear Fuel and Waste Management Co. (SKB), progress report, 1992

NOR 92 Nordstrom, D. K., "On the Evaluation and Application of Geochemical Models.", Proceedings of the fifth CEC Natural Analogue Working Group Meeting EUR 15176, Toledo, Spain, 5-9 Oct., 1992.

PAR 80 Parkhurst, D.L., Thorstenson, D.C., Plummer, L.N. "PHREEQE, A Computer Program for Geochemical Calculations", USGS/WRI p 80-96, 1980.

PAS 93 "Project on Alternative Systems Study (PASS), Final report", Swedish Nuclear Fuel and Waste Management Co. (SKB), Report TR 93-04 Stockholm, Sweden, 1993.

PRO 90 "SKI Project - 90", Swedish Nuclear Power Inspectorate (SKI), TR 91-23, Stockholm, 1991.

REA 91 Read, D., Boyd, T. W., " The CHEMVAL international exercise: a geochemical benchmark within the MIRAGE project - Second phase", Radioactive Waste Management and Disposal, CEC Publ., p. 574-584, 1991.

RYD 53 Rydberg, J., "Studies on the extraction of metal complexes. X: Additional results on the complex formation between thorium and acetylacetone.", Arkiv för Kemi, band 5, nr. 39, 1953.

RYD 69 Rydberg, J., Acta Chem. Scand. 23, 647, 1969 b)Reinhardt, H. and Rydberg, J., ibid. 23, 2773, 1969 c) Andersson, C., Andersson, S.O., Liljenzin, J.O., Reinhardt, H. and Rydberg, J. ibid. 23, 2781, 1969 d) Johansson, H., and Rydberg, J., ibid. 23, 2797, 1969.

RYD 92 Rydberg, J., Musikas, C., Choppin, G. R., "Principles and Practices of Solvent Extraction", Marcel Dekker Inc., New York, USA, 1992.

SAB 74 Sabatini, A.,Vacca, A., Gans, P., "MINIQUAD. General computer program for the computation of formation constants from potentiometric data.", Talanta, 21, 53-77, 1974.

SAM 96 Samuelsson, A., "UNCCON: A computer program for uncertainty analysis of the effect on the solubility of solid phases, from uncertainties in the concentration of groundwater elements", Diploma thesis, Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden, 1996.

SCH 88 Schecher, W. D., Driscoll, C. T., "An Evaluation of the Equilibrium Calculations Within Acidification Models: The Effect of Uncertainty in Measured Chemical Components", Water Resources Research, vol. 24, no.4, p. 533-540, 1988.

SME 92 Smellie, J.,Laaksoharju, M., "The Äspö Hard Rock Laboratory: Final evaluation of the hydrochemical pre-investigations in relation to existing geologic and

hydraulic conditions", Swedish Nuclear Fuel and Waste Management Co. (SKB), TR 92-31 Stockholm, Sweden, 1992.

SUM 93 Sumerling, T. J., Zuidema, P., Grogan, H., van Dorp, F., "Scenario development for safety demonstration for deep geological disposal in Switzerland", High Level Radioactive Waste Management, Proceedings of the 4th Annual Internatinal Conference, Las Vegas, USA, 1993.

SUN 86 Sundström, B., "Development of uncertainty analyses for the transport of radionuclides from a repository for high-level radioactive waste. Stage one.", Swedish Nuclear Power Inspectorate (SKI), TR 87-01, Stockholm, Sweden, 1986.

TUL 95 Tullborg, E-L., personal communication, TERRALOGICA AB, Gråbo, 1995.

VOS 98 Voss, C., Tirén, S., Glynn, P., "Hydrogeology of Äspö Island, Simpevarp", SKI report 96:13, Swedish Nuclear Power Inspectorate (SKI), Stockholm, Sweden, in press, 1998

WIL 78 Williams, W.H.; "A Sampler on Sampling"; Wiley Series in Probability and Mathematical Statistics, J.Wiley & Sons, Chichester, UK, 1978

WIN 79 Wingefors, S., "Kompendium i extraktionskemi", Department of Nuclear Chemistry, Chalmers University of Technology, Göteborg, Sweden, 1979.

WOL 92 Wolery, T. J., "EQ3/6, A software package for geochemical modeling of aqueous systems: Package overview and installation guide (Version 7.0).", UCRL-MA-110662 PT 1, LLNL, Livermore, USA, 1992.

WOR 91 Worgan, K. J., Robinson, P. C., "CRYSTAL: A Model of a Fractured Rock Geosphere for Performance assessment within SKI Project-90." ,SKI Report 91:13, Swedish Nuclear Power Inspectorate (SKI), Stockholm, Sweden, 1991.

WOR 95 Worgan, K. J., Robinson, P. C., " The CALIBRE Source-Term Code: Technical Documentation for Version 2.", SKI Report 95:17, Swedish Nuclear Power Inspectorate (SKI), Stockholm, Sweden, 1995.

# Appendix 1

## Sampling techniques

For many purposes in statistical analyses of the output of computer simulations, the sampling techniques are very popular. This is mainly because these techniques are rather simple to use and evaluate. Three methods are usually discussed in this context: Monte Carlo (MC) sampling, stratified sampling, and Latin Hypercube Sampling (LHS). However, there exist some hybrids for example iterated fractional factorial design [AND 93], but since this is not a pure sampling method it will not be treated here. In each of the cases described, let $X=\{x_1,....,x_n\}$ be the input vector to the tested program, i.e. the input data which are included in the uncertainty analysis, and let $F_i$ be the distribution function for each input.

### Monte Carlo sampling

In simple Monte Carlo sampling, one value is sampled from each input distribution, producing one input vector for which the model is executed. This is repeated until enough results are obtained to give good statistically certain results.

The main drawback is that many computer runs are usually needed while still just barely covering the input space.

### Stratified sampling

In stratified sampling, each distribution interval is portioned into m equal parts and one sample is taken from each part, thus producing an nxm input matrix where each column is an input vector to the computer program. This method increases the possibility that the whole sample space will be represented. However, in some cases values with low probability may be over represented. Naturally, the intervals may be taken with equal probability instead, but this then approaches the Latin Hypercube sampling, shown below.

### Latin Hypercube sampling

Latin Hypercube sampling appears to be similar to extended stratified sampling, but has one large difference. The values in each row of the nxm input matrix are mixed randomly in order to produce m totally randomised input vectors. The evaluation of statistical estimators using the LHS technique is described in detail by Iman [IMA 80].

It has been experimentally proven that LHS is often more then 50 times more effective than MC sampling. Effective meaning that there is a need for 50 times more computer runs to obtain equal results from MC than LHS. Further, it is seen that LHS ensures that a larger part of the input space is covered, Figure 1. There, samples are drawn from a uniform distribution on the interval [0,1].

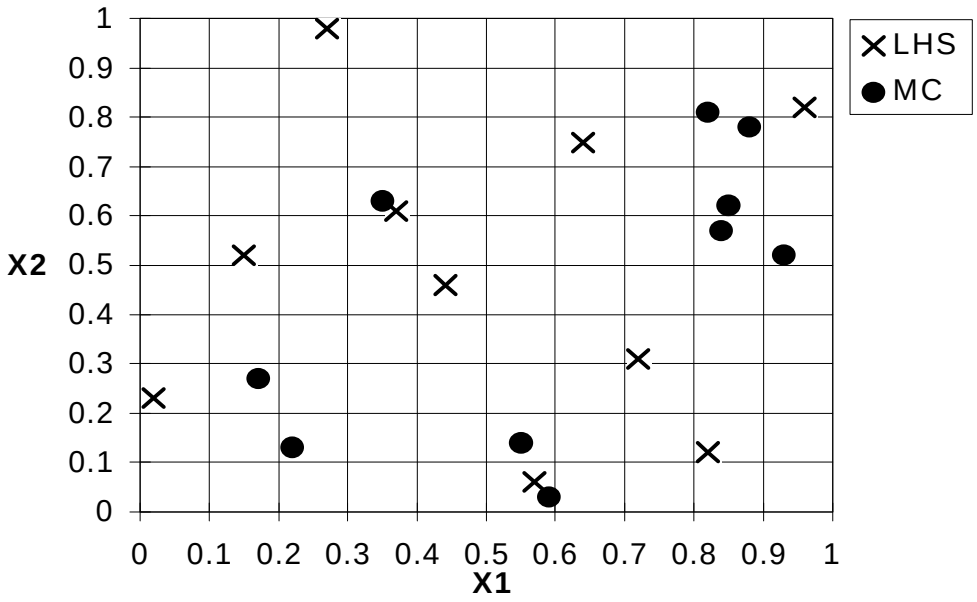



***Figure 1:*** *Comparison of input space coverage by MC and LHS techniques*

Clearly, the input space is not as well covered with MC as with LHS. One way to see this is that, when using LHS, every row and column will contain one sample. This is not the case with MC. The example shown in Figure 1 is only two dimensional. As the dimension increases, the power of the LHS will be more clear. However, in many cases, MC sampling is still used because of the simplicity of its implementation.

# Appendix 2

## Bodländers method

One fundametal property of chemistry is the stoichiometry of a reaction. In the beginning of the 20th century, Bodländer described a method to obtain these constants by the use of potentiometric titration. This method is outlined below.
Assume the following reaction

$$qM + rL == M_qL_r \quad \text{(R1)}$$

The stability constant for this reaction is:

$$b_{qr} = \frac{[M_qA_r]}{[M]^q * [A]^r} \quad (1)$$

If two different solutions are used with different concentrations of the metal, the ratio between the metal concentrations is:

$$\frac{[M]_1}{[M]_2} = \left(\frac{[M_qA_r]_1 * [A]_2^r}{[M_qA_r]_2 * [A]_1^r}\right)^{1/q} \quad (2)$$

This ratio may be obtained using the Nernst equation, see Section 5.2. By holding the complex concentration fixed at the same level in the two solutions, Equation 2 is reduced to:

$$\frac{[M]_1}{[M]_2} = \left(\frac{[A]_2}{[A]_1}\right)^{r/q} \quad (3)$$

In the same way one might keep the free ligand concentration at similar levels and obtain Equation 4.

$$\frac{[M]_1}{[M]_2} = \left(\frac{[M_qA_r]_1}{[M_qA_r]_2}\right)^{1/q} \quad (4)$$

It is then possible to obtain the stoichiometric constants r and q.

# Appendix 3

## Detection limits in solvent extraction.

Solvent extraction is in many cases a powerful tool in the determination of stability constants for different reactions. However, in some cases, it may prove difficult to determine the first stability constant in consecutive reactions. The cause is often a practical problem such as selection of extractant concentration or method of evaluation. An attempt to obtain stability constants by fitting the distribution function to the experimental data, by least square fit or similar methods, may result in a multitude of stability constants. Such a case is illustrated below.

Assume the following expression for the distribution coefficient:

$$D = \frac{l_4 b_4 \left[M^{4+}\right]\left[L^{-}\right]^4}{\left[M^{4+}\right] + b_1\left[M^{4+}\right]\left[L^{-}\right] + \ldots + b_4\left[M^{4+}\right]\left[L^{-}\right]^4} \qquad (1)$$

where M is a metal with a valence state of 4, L a ligand with a valence state of -1 and $\lambda_4$ is the distribution coefficient of the uncharged complex between the organic and aqueous phase. If the the free metal ion never reaches a level greater than about 0.1 percent of the dominating species in the solution, its concentration will be negligible compared with the other terms in Equation A1. Thus it is possible to remove that term in the sum. It is then possible to divide all the $\beta_i$ with a nearly arbitrary constant, and it is thus possible to obtain almost any values for the stability constants. However, the quotient between them will be fixed, so if the first constant can be determined by some other means, the rest are fixed. As mentioned earlier, the constant with which the $\beta_i$ are divided may not exceed a value close to the value of $\beta_i$ itself, as shown below.

Assume the following reaction:

$$M^{p+} + L^{-} == ML^{(p-1)+} \qquad (R1)$$

where p > 1.

The stability constant for this reaction is.

$$b_1 = \frac{\left[ML^{p-1}\right]}{\left[M^{p+}\right] * \left[L^{-}\right]} \qquad (2)$$

To determine the stability constant for a complex it must be about one percent of the dominating one, depending on ,among other things, the fitting algorithm and the original data. The limit at which the free metal ion starts to dominate in the solution is when the concentrations of it and the first complex are equal. Thus it is possible to determine the largest obtainable value for the first stability constant. In this case, it yields:

$$\log(b_1) = -\log(\left[L^-\right]) \quad (3)$$

Therefore, in order for the free metal ion to be dominant in a solution, a ligand concentration of lower than -log($\beta_1$) must be obtained. In order to determine a stability constant by fitting procedures, it is thus desirable to come close to this concentration. This can be solved with a better solvent, but then it is probable that extreme values at the extraction maximum will occur, thus making determinations in that region very uncertain.

# Appendix 4

## Testing goodness-of-fit

Testing goodness-of-fit for values assumed to follow some particular distribution is important in that such a test gives an answer as to how good the assumption is. Such information is important when further conclusions are to be drawn from the sampled values.

The goodness-of-fit is often performed using the Pearson chi square statistic, $D^2$, see Equation 1.

$$D^2 = \sum_{i=1}^{k} \frac{(x_i - E[x_i])^2}{E[x_i]} \quad (1)$$

where $x_i$ are the sampled values, k is the number of samples and $E[x_i]$ are the theoretical values according to the desired distribution. If the null hypothesis is that the sample values follow the assumed distribution, the chi square test rejects if the value of $D^2$ is large. The rejection limit is given by the sample size and the confidence level for the rejection. This limit is often obtained from a table listing confidence level versus degrees of freedom, i.e. sample size minus one.

However, it is important to remember that a chi square test gives large significance to the sample points in the "tails" of the distribution. Thus if there is a small difference in a value with low probability, the value of $D^2$ will be significantly increased. The confidence intervals calculated in this thesis may therefore still be valid in spite of the fact that the Pearson test rejected the normal assumption.

## Appendix 5

In many thermodynamic equilibrium programs the dependence of the stability constants on the remperature is obtained by the folowing equation:

$$\frac{\P\log(b)}{\P\frac{1}{T}} = -\frac{\Delta H^0}{R\ln(10)} \tag{1}$$

This equation is derived according to the reasoning below.
Generally, Equation 2 is true.

$$G = H - TS \tag{2}$$

Dividing by T and taking the partal derivative with respect to T yields:

$$\left(\frac{\P\left(\frac{G}{T}\right)}{\P T}\right)_p = H*\left(\frac{-1}{T^2}\right) + \left(\frac{\P H}{\P T}\right)_p * \frac{1}{T} - \left(\frac{\P S}{\P T}\right)_p \tag{3}$$

Together with:

$$\frac{\P H}{\P T} = c_p \tag{4}$$

and

$$\frac{\P S}{\P T} = \frac{c_p}{T} \tag{5}$$

and variable transforamtion one obtains:

$$\left(\frac{\partial\left(\frac{G}{T}\right)}{\partial\left(\frac{1}{T}\right)}\right)_P = H \tag{6}$$

which after taking the difference between two states and selecting a standard state yields:

(7)

At the standard state the stability constant, $\beta$, is related to the Gibbs free energy accoding to:

$$\Delta G^0 = -RT*\ln b \tag{8}$$

Using equations 7 and 8 and compensating for the fact that usually it is not lnβ that is known, but rather logβ, gives:

$$\frac{\P \log(b)}{\P \frac{1}{T}} = -\frac{\Delta H^0}{R \ln(10)} \qquad (9)$$

## Appendix 6: List of papers

This report is based on the following publications:

Paper I
Ekberg, C., Emrén, A.T., "SENVAR: A code for handling chemical uncertainties in solubility calculations", Computers & Geosci., Vol. 22, 867, 1996.

PaperII
Ekberg, C., Lundén, I., "Uncertainty analysis for some actinides under groundwater conditions.", Journal of Statistical Computation and Simulation, vol. 57 (1-4), p. 271-284, 1997.

PaperIII
Ekberg, C., Skarnemark, G., Emrén, A. T., Lundén, I., "Uncertainty and sensitivity analysis of solubility calculations at elevated temperatures", Mat. Res. Soc. Symp. Proc., Vol. 412, 889, 1996.

Paper IV
Ekberg, C., Börjesson, S., Emrén, A. T., Samuelsson, A., "MINVAR and UNCCON, Computer Programs for Uncertainty Analysis of Solubility calculations in Geological Systems.", submitted to Computers & Geosci., 1998.

Paper V
Ekberg, C., Emrén, A. T., Samuelsson, A., "The Effect of Mineral Variability on the Solubility of some Actinides: an Uncertainty Analysis", Mat. Res. Soc. Symp. Proc., Vol. 465, 735, 1996.

Paper VI
Ekberg, C., Emrén, A.T., "Conceptual Uncertainties in Solubility Calculations in Groundwater Systems: A Calculation exercise.", presented at the SAMO-98 conference and submitted to Radioactive Waste Management and Environmental Restoration, 1998.

Paper VII
Ekberg, C., Albinsson, Y., Brown, P. and Comarmond, M., J., "Studies on the Complexation Behaviour of Thorium(IV). 1. Hydrolysis.", submitted to J. Solution Chem., 1999.

Paper VIII
Ekberg, C., Albinsson, Y., Brown, P. and Comarmond, M., J., "Studies on the Complexation Behaviour of Thorium(IV). 2. Phosphates.", submitted to J. Solution Chem., 1999.