Available online at www.sciencedirect.com

Talanta 63 (2004) 907–916

Talanta

www.elsevier.com/locate/talanta

LJUNGSKILE: a program for assessing uncertainties in speciation calculations

A. Ödegaard-Jensen[a,*], C. Ekberg[a], G. Meinrath[b,c,d]

[a] *Nuclear Chemistry, Department of Materials and Surface Chemistry, Chalmers University of Technology, S-41296 Gothenburg, Sweden*
[b] *RER Consultants, Schießstattweg 3a, D-94032 Passau, Germany*
[c] *Institute of Inorganic Chemistry, Technische Universität Bergakademie Freiberg, D-09596 Freiberg, Germany*
[d] *Institute of Geology, Technische Universität Bergakademie Freiberg, D-09596 Freiberg, Germany*

Received 11 August 2003; received in revised form 22 December 2003; accepted 24 December 2003

Abstract

Speciation calculations are often the base upon which further and more important conclusions are drawn, e.g., solubilities and sorption estimates used for retention of hazardous materials. Since speciation calculations are based on experimentally determined stability constants of the relevant chemical reactions, the measurement and experimental uncertainty in these constants will affect the reliability of the simulation output. The present knowledge of the thermodynamic data relevant for predicting the behaviour of a complex chemical system is quite heterogeneous. In order to predict the impact of these uncertainties on the reliability of a simulation output requires sophisticated modelling codes. In this paper, we will present a computer program, LJUNGSKILE, which utilises the thermodynamic equilibrium code PHREEQC to statistically calculate uncertainties in speciation based on uncertainties in stability constants. A short example is included.

Keywords: LJUNGSKILE; Speciation calculations; Latin hypercube sampling

1. Introduction

Aquatic systems form an important compartment of the geochemical cycle of metals. Interactions of metal ions with water constituents, e.g., hydroxide, carbonate, phosphate, formic and acetic acid, or fulvic and humic substances, decide the fate of metal ions in the aqueous compartment. It has been recognized since some time that the mere amount of an element in a water does not allow a detailed understanding of an element's behaviour and effects. The specific coordination of the metal ion strongly modifies the chemical behaviour of a given element. A rationalisation of the behaviour of metal ions in aquatic systems therefore required an understanding of their specific chemical form in a given medium. This finding has given rise to speciation analysis [1,2].

The clue to speciation analysis for metal ions with fast equilibria is the Gibbs free energy of formation ΔG_R°.

$$\Delta G_R^{\circ} = -RT \ln K^{\circ} \tag{1}$$

with R denoting the gas constant and T the absolute temperature in Kelvin. Eq. (1) directly links the thermodynamic formation constant K° with the Laws of thermodynamics. Thus, a formation constant K° is a constant of nature. Hence, a constant K° determined under more or less convenient conditions in a laboratory can be applied to calculate the relevance of the same reaction in more complicated situations and at concentration levels that were inaccessible otherwise. For most inorganic solution species, other methods for determination of their amount at trace level concentrations do not exist.

A large number of collections holding such data are available (e.g., [3–8]). The necessary numerical transformations readily get too complicated for easily hand-made calculations and computers have been used quite early for such calculations [9]. Large and versatile computer codes exist

---

[*] Corresponding author. Tel.: +46-31-772-2804; fax: +46-31-772-2931.
E-mail address: arvidoj@chem.chalmers.se (A. Ödegaard-Jensen).

0039-9140/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2003.12.053

(e.g., [10–13]). These systems are used, e.g., to simulate and predict the composition of aqueous solutions in natural waters like sea waters [14], salt brines and plumes [15]. In the vastly overwhelming amount of cases, the output of speciation calculations consists of values originating from using the given stability constants without any uncertainty associated. However, in most cases, e.g., where the uncertainties in the stability constants are relatively large, such results are rather useless [16,17].

Quantitative expression of doubt is of importance as soon as conclusions of a numerical simulation/prediction become a basis for important decisions. The need for a common basis for quantitative and comparable expression of doubt is a necessary element of mutual recognition, trust and transparency. ISO in conjunction with other institutions (among those IUPAC and IUPAP) has issued normative criteria for the quantitative expression of uncertainty in measurement [16]. The urgent need for quantitative expression of uncertainty in thermodynamic data has been outlined previously [15,18–21] in the context of uranium mining remediation, acid mine water drainage and safety assessment of nuclear waste repositories.

Expression of doubt in the output of a speciation calculation has far-reaching consequences. Determination of the number of significant species in solution, spotting the species with largest contribution of uncertainty in data from chemical analysis, detection of less important species are only a few possible situations where speciation calculations with associated uncertainties will be helpful. Mean value based argumentation may consequently result in wrong conclusions, especially if larger amounts of species are included into a calculation due to international agreements and normative requirements [22–25], scientific conclusions with an impact outside the academic sphere will not any longer be acceptable without a traceable statement of uncertainty.

This paper describes the LJUNGSKILE program designed for easy calculation of speciation uncertainties in complex chemical systems. It uses the public domain PHREEQC code [12] as a backbone.

This paper describes the LJUNGSKILE program (Version 2.0) designed for easy calculation of speciation uncertainties in complex chemical systems. A manual holding a general description of the code as well as the code itself is available from the WWW [26]. This manuscript presents the code, illustrates the changes between the Version LJUNGSKILE 1.0 and the Version LJUNGSKILE 2.0 and gives an application example. The discussion assumes that the reader has a basic understanding of PHREEQC input conventions and a LJUNGSKILE 1.0 manual at hand [27].

2. Algorithmic and statistical considerations

Since thermodynamic data with a traceable uncertainty representing the complete uncertainty budget of its determination is almost non-existent, approximations to a reasonable amount of measurement uncertainty is unavoidable. It is assumed that the uncertainties in a formation constant Y can adequately be described by a normally distributed probability density $Y \pm u$ where the value u in the "$\pm u$" represents the region of approximately 68% confidence.

The concentration of a species in solution can be calculated from the thermodynamic data and the law of mass action. For a reaction

$$\mathrm{A} + \mathrm{B} \Leftrightarrow \mathrm{C} + \mathrm{D} \qquad (2)$$

with equilibrium constant K follows for the species of interest D

$$\log[\mathrm{D}] = \log K - \log[\mathrm{C}] + \log[\mathrm{A}] + \log[\mathrm{B}] \qquad (3)$$

where [ ] denote concentrations in units of $\mathrm{mol\,dm^{-3}}$ and K is the valid constant under given conditions, e.g., the desired temperature and ionic strength. Here, one must realise that in addition to the equation systems described above the activities of the different species also have to be included. This makes the situation even more complicated. It is a simple task to calculate quantity [D]. The numeric get a bit more involved if several species play a role that are coupled by one or more constituents. Only simple situations can be treated without resorting to least-squares optimisation codes like PHREEQC or EQ3/6 [12,13].

Once uncertainties in the thermodynamic data gets involved, it is always possible to resort to Monte Carlo (MC) methods. That is, to repeat the calculations many times by randomly drawing input values from the specified uncertainty range of each constant involved in the calculation [28].

This method has been implemented in the LJUNGSKILE code [26]. The code comes in two parts. The first part does the number crunching while the second part transforms the output into different graphical displays. The public domain speciation code PHREEQC is used to calculate the composition of a solution specified by the user. A water composition can be specified and the values and associated uncertainties of the species of interest can be specified. Since the database format of PHREEQC is easily manipulated, few restrictions exist. A pure Monte Carlo approach is possible but inefficient. Sampling from distributions requires proper consideration of the tails of a distribution. The majority of samples will be drawn from the centre of a distribution. Hence, a large number of samples has to be dawn to ensure that the tails of the distribution will be represented in the sample. Sampling from multidimensional distributions will require a further increase of the number of samples, because all tails have to be considered appropriately simultaneously. Instead, Latin hypercube sampling (LHS) has been chosen [29]. LHS is a stratified sampling technique. The distribution is split into several strata of equal probability. Hence, all sections of a distribution are equally represented. Details of the LHS have already been discussed elsewhere [30].

A Latin hypercube sampling is used favourably if many other quantities enter the calculation process. In fact, LHS requires that each of the distribution of values of these quantities is likewise stratified in the same manner. Furthermore, the minimum number of strata is given by the number of quantities entering the calculation.

LHS allows to calculate a statistics from a small number of runs. If *n* quantities enter the calculation, a minimum of *n* runs have to be performed. Latin Hypercube samples can naturally have a larger number of samples than the number of variables. If more than only an estimate of the standard deviation is desired, a large number of runs is necessary. In the LJUNGSKILE implementation, the user specifies the number of runs.

3. Program description

3.1. The LJUNGSKILE code

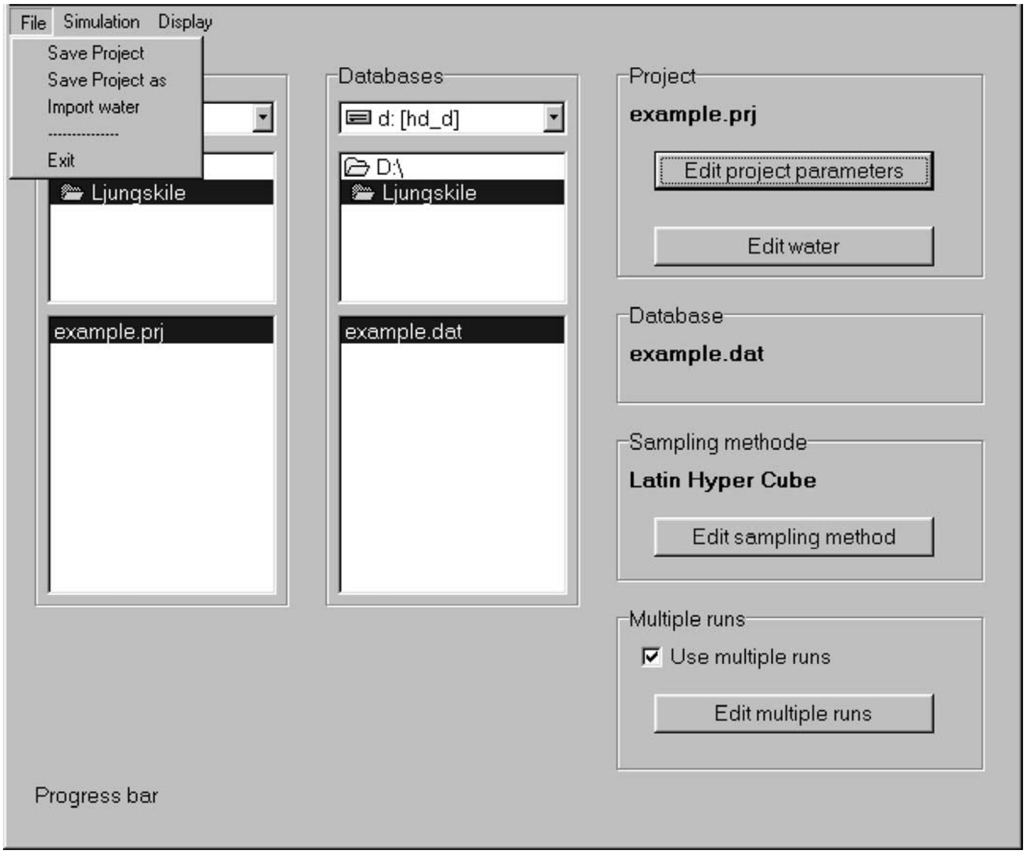
The LJUNGSKILE program is written in the programming language C^{++} using the Borland C++ builder (Borland). The program is divided into two major parts. One part containing the chemical calculations and one containing the statistical methods used. The basic idea is to make many runs with PHREEQC and in each run the stability constants of interest are changed slightly within their respective confidence interval according to the statistical method used (see below). Then, after each such cycle, if this option was selected, one of the factors may be changed by given increments to a given level, e.g., changing pH, and thus create a speciation diagram versus pH with uncertainty bands for each species concentration.

3.2. Chemical calculations

The basic approach of the LJUNGSKILE program is to use different statistical sampling techniques to determine the effect of uncertain stability constants on a speciation calculation. The chemical speciation calculations are made by the well-known thermodynamic equilibrium program, PHREEQC. The chemical choices the user has to make are given in Fig. 1: selecting species and formation constants (as a project file ∗.prj) whose formation constants will be varied and a database (∗.dat) holding information on thermodynamic properties of other species.

3.2.1. Selection of project file

The project file holds all relevant information of the task to be treated by LJUNGSKILE code. If the project parameters are set, the water is defined, the sampling method selected and the multiple run feature chosen or ignored, then LJUNGSKILE is put to work by clicking “Simulation” in the menu. Before the code starts, the project file will have to



Project
example.prj

Database
example.dat

Sampling methode
Latin Hyper Cube

Multiple runs
☑ Use multiple runs

Fig. 1. LJUNGSKILE main window (screenshot).

be saved under a file name given by the user. If a project file is created from scratch, no name exists and a window will pop up ("Cannot create file"). The project file must be saved under a suitable name using "Save as" in the File menu.

3.2.2. Selection of water

The user gives the chemical composition of a water together with other quantities such as pH, pe and temperature. This information is then written in the input file to PHREEQC. Its filename is set by default to "PHREEQEC.IN" and cannot be changed. It is noteworthy that if the water is not charge balanced on input it will become so by the addition of an inert positive or inert negative element, Ip and Im, respectively (see below).

3.2.3. Selection of database

The so-called thermodynamic databases are actually collections of some mean values for formation constants and solubility products of a certain number of chemical compounds. The format is depending on the very program designed to read this data file. We will nevertheless stick with the common term 'database' in this manual. The choice of a certain database may influence the calculated result considerably. Hence, the user may generate an own database. As long as the database is in agreement with the PHREEQC format, the LJUNGSKILE output will be based on the selected database. In some cases the desired species are not present in the general database and therefore there is a possibility to change the database used. Note, that the species Im and Ip must be included into any self-created database in the way they are given in the example database.

3.2.4. Selection of sampling method

The sampling methods offered are the Monte Carlo design and the Latin hypercube design. The Latin hypercube design is more efficient if a larger number of thermodynamic constants have been selected to vary in the project, while MC sampling is appropriate for just a handful of thermodynamic constants to vary. However, it is always helpful to have the possibility to compare different approaches and, hence, the MC approach is provided as a choice.

3.2.5. Selection of multiple run parameters

Sometimes it is also interesting to investigate the evolution of a situation as a function of a parameter. The parameter, its starting and stopping value and the step width can be specified here. Note, that the multiple step feature has to be used wisely. Especially if solid phase dissolution and/or precipitation occur, the solution variations may become considerable. The LJUNGSKILE program accounts for charge balancing by adding 'virtual' species Im and Ip (see below) but if large amounts of these virtual species are to be added to achieve convergence, the numerical results may be void.

If the project parameters are set and the water is defined, the sampling method selected and the multiple run feature chosen or ignored, then LJUNGSKILE is put to work by clicking "Simulation" in the menu bar. PHREEQC runs in a DOS box, controlled by the LJUNGSKILE code.

3.3. PHREEQC settings

As described above the water composition is given by the user together with pH, pe and temperature. If the given water composition results in an charge imbalance this is taken care of by the LJUNGSKILE program. The method used is to use the inert elements Ip and Im which are positive and negatively charged, respectively. This is made by first adding 10^{-6} M of Ip and then use the flag CHARGE. In case PHREEQC does not converge, Im is used in the same manner and thus charge balance is obtained. Thus, it is simple to change the concentration of one of the elements present in the water without affecting the other water properties. This is a necessity since in a speciation diagram there is generally, at least, one factor changing along the *x*-axis.

Other settings in the PHREEQC input file is the use of the PRINT keyword. The flags used here are: reset false, totals true, species. This is to make the result file clearer and more structured (interested readers are directed to the PHREEQC manual [12]).

3.4. Some comments on statistical and programming methods

Statistics is looking for objective criteria to extract information from larger amounts of data. Data may be considered to be raw information in a qualitative or quantitative way. Data producers have the obligation to present all pertinent information that would impact on the use of it, to the extent possible. Of course, every possible use of data cannot be envisioned when it is produced, but the details of its production, its limitations and quantitative estimates of its reliability always can be presented. In an ideal situation, a reasonable measure of the doubt to be associated with given data will always be available.

Decisions need to be made despite the fact that informations are doubtful. Within a computer program, handling and even creating uncertainty seems a contradiction to the foremost abilities of a computer being a deterministic machine for invariable execution of instructions [31].

However, the computer is a very helpful tool—and its helpfulness is the larger as there is no other tool to perform the task—to handle doubt and uncertainty: (a) because it is a deterministic machine, (b) because it is fast and (c) because suitable algorithms to simulate randomness exist.

The use of computers for theoretical speciation calculations started rather early with the computer programs of Lars Gunnar Sillén in Sweden [32,33]. Several major achievements have been included into Sillén's HALTAFALL modelling codes: application of suitable minimisation routines for linear equations, suitable input formats for chemical information and programming techniques to accommodate all code in the tiny storage area of the early computers. Compared to these early applications, modern desktop machines represent the computing power of several computer centres in Sillén's time. Since then, storage capacity, CPU clock rates and algorithms have improved dramatically. The elements, however, have to be brought to work.

The sampling in the LJUNGSKILE program is made by either simple Monte Carlo sampling or Latin hypercube sampling [29]. The basic step for LHS is to make a cumulative distribution function (cdf) for the sample based on the mean value and the standard deviation given for the respective variable. Here, the user may select how precise this cdf itself should be by selecting how many points are to be included in its construction. The samples are then drawn from this empirical cdf. For the MC sampling, there is no extra work, since this method samples complete randomly within the given interval/distribution.

In the current version of the LJUNGSKILE program, there are two distributions and two sampling methods to select from. The distributions are the uniform distribution and the normal distribution. The sampling intervals are given slightly different for the two distributions. The uniform distribution requires a highest and a lowest value while the normal distribution uses a mean and a standard deviation.

3.5. The LJUNGSKILE display program (LDP)

LDP is intended to graphically display the results obtained for the LJUNGSKILE program. It is not a sophisticated graphics program but kept comparatively simple. LDP gets its information mainly from the respective project file in the \results\ directory. The LJUNGSKILE program copies these files to the results directory and renames the project file into an ∗.ldp file. Thus, LDP is set by default to open an ∗.ldp file and searches its respective input files in the same directory. From the different output files LDP can create a variety of diagrams. The user, under standard circumstances, will see just one ∗.ldp file in the respective result subdirectory. In fact, the diagram displayed may vary considerably depending on the task given to the LJUNGSKILE program.

The basic choice in the diagram options has been made by the user already when running the LJUNGSKILE code by simulating either a single run or a multiple run. The LDP code recognises from the ∗.ldp file whether or not the single run option has been chosen.

3.6. Single run display

A single run is displayed as a kind of modified box plot. An example is given in Fig. 2. The square centre shows the median value. The box around the median value gives the

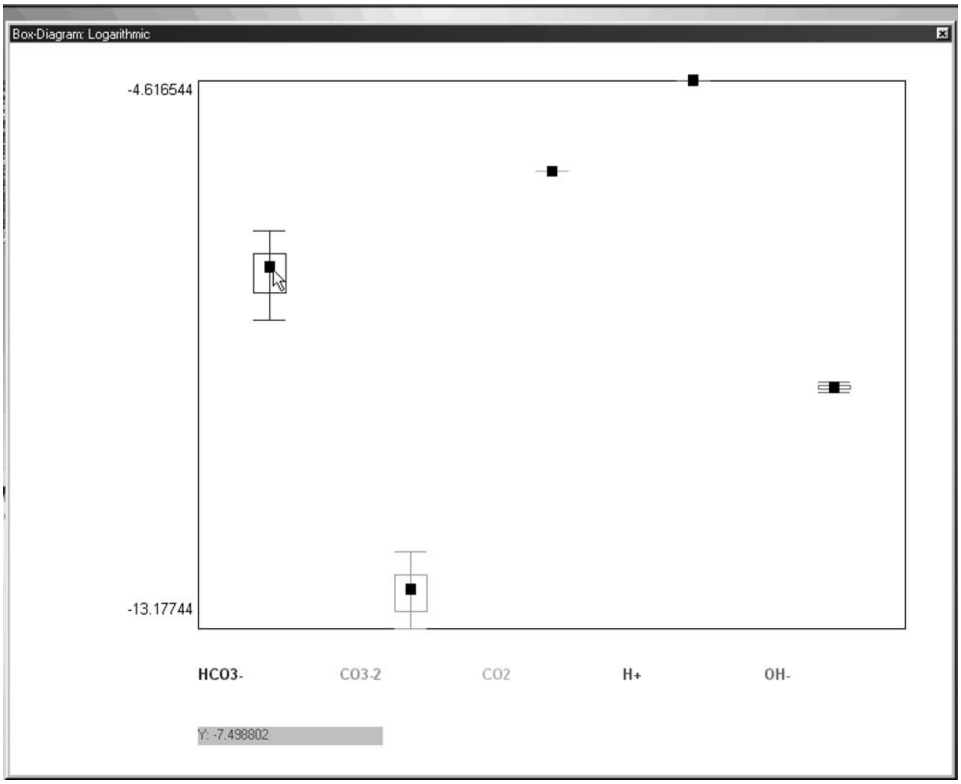


Fig. 2. Logarithmic diagram of calculated species concentrations in the CO_2/H_2O system at pH 4.7. Filled square gives median, the open box gives the 68% confidence range and the bars indicate the 95% confidence range. On screen the LDP the diagram is shown in colour. In the field below the diagram the position of the cursor is given in terms of the ordinate scale. From left to right, the species are HCO_3^-, CO_3^{2-}, $CO_{2(aq)}$, H^+ and OH^- (screenshot).

68% confidence range. The error bar-like whiskers give the 95% confidence range. All four display options are available for a Single run LJUNGSKILE simulation. However, in most cases only the display of concentrations in logarithmic form makes sense. The data can be saved as an ASCII file. However, in most situations the authors have found it more suitable to use the full data set in PHREEQEC.OUT file in a suitable graphics program. The 'Species' menu is also available but the 'Confidence' menu options do not work with a single run data set.

Fig. 2 gives a Single Run logarithmic concentration diagram of CO_2, HCO_3^-, CO_3^{2-}, H^+ and OH^- at pH 4.7. In this project, the CO_2 partial pressure has been fixed to 0.035%, while the amount of H^+ was fixed by the selected pH. Hence, both species $CO_2(aq)$ and H^+ do not show an uncertainty range.

3.7. Multiple run displays

If LJUNGSKILE has calculated a multiple run problem, e.g., a speciation as a function of pH, then the initial LDP output is a concentration diagram as function of pH. The uncertainties are displayed as approximate 68% confidence percentiles. The user has several options to modify the LDP diagram.

3.8. Diagram options

The following illustrations are given with the purpose to illustrate the features of the LDP program with a given output at hand. For a more detailed guideline on how to generate results, please refer to Section 4 below. The diagrams are modified by the features provided in the menu items 'Diagram', 'Species' and 'Confidence'.

A diagram type is selected from the 'Diagram' menu: (a) concentration, (b) logarithmic concentration, (c) percent distribution, and (d) total concentration.

The 'Species' menu allows the user to include/exclude certain species from being displayed in the LDP diagram.

The 'Confidence' menu allows to select four different confidence levels if more than 36 LJUNGSKILE cycles have been calculated. These confidence levels are 68, 90, 95 and 99%. If less then 37 LHS cycles have been calculated in the LJUNGSKILE runs, 'standard deviation' and 'range' can be chosen.

The concentration presentations are calculated from the ∗.out files saved by the LJUNGSKILE program in the path\results\∗.prj directory. A species distribution is calculated from the phrout.∗ files. The program recalculates the species distribution for each run from the concentration information, evaluates the empirical distribution function and chooses the value closest to the 0.16 percentile, the median and value closest to the 0.84 percentile value as uncertainty limits and centre, respectively. It is obvious that a small number of runs provides a more variable uncertainty value than a larger number of runs. It should be kept in mind that the LJUNGSKILE code wants to visualise the uncertainties in sets of formation constants. For several reasons, these visualisations can be approximations only. The experience with LJUNGSKILE calculations shows, however, the necessity of an attitude of great carefulness that should be taken towards speciation calculations.

In Fig. 3a and b, the use of the 'Species' menu is illustrated. In some cases, this menu allows to investigate the output of a LJUNGSKILE simulation in more detail by excluding a subset of species from LDP calculations and display. An example in Fig. 3a gives the speciation of Fe(II) and Fe(III) hydroxo species as a function of pH. The following species are included: Fe^{2+}, $FeOH^+$ and $Fe(OH)_2$, Fe^{3+}, $FeOH^{2+}$, $Fe(OH)_2{}^+$, $Fe(OH)_3$ and $Fe_2(OH)_2{}^{4+}$. Under given conditions, Fe(II) dominates the speciation distribution. The hydrolysis of a divalent cation like Fe(II) occurs in the alkaline pH range while the free Fe^{2+} ion prevails in the acidic to neutral pH range.

At the selected pe of 5, some Fe(III) is also in solution but at much lower amounts. In such a situation, the 'Species' menu is useful to eliminate the Fe(II) species from the calculations. The result is shown in Fig. 3b.

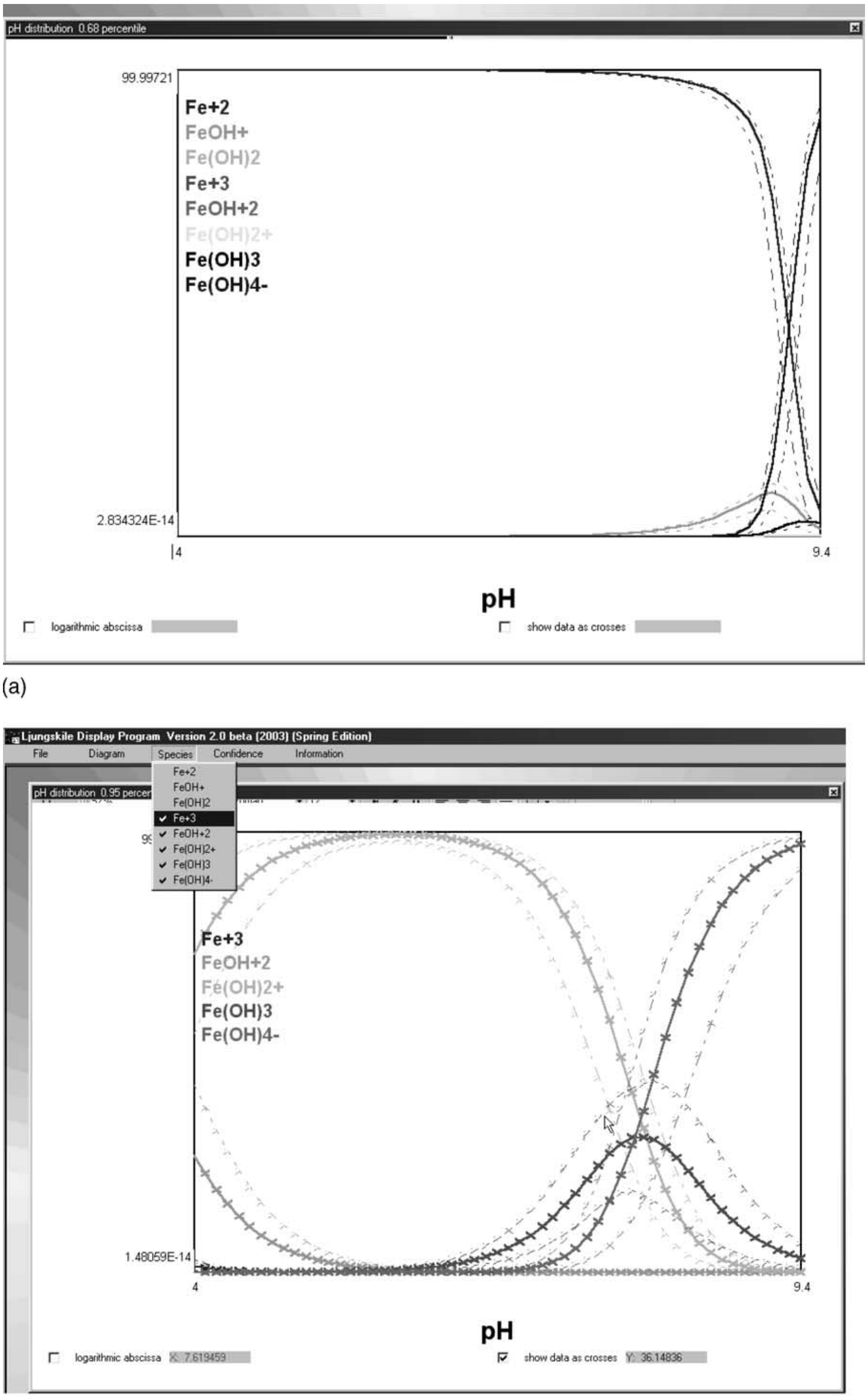
The species are deselected in the 'species' menu by clicking the hook before the desired species. This process is reversible. In Fig. 3b, the species Fe^{2+}, $FeOH^+$ and $Fe(OH)_2$ have been deselected. The names of species still considered in the diagram are shown. A line is colour-coded and drawn in the same colour as the respective species name.

For more than 36 runs, the 'Confidence' menu provides four confidence limits: 68, 90, 95 and 99%. Otherwise, only standard deviation (S.D.) or the range an be selected by the user. In case of Fig. 3b, 225 LHS cycles have been calculated and the four confidence ranges can be selected. In Fig. 3b, a confidence range of 95% is chosen as indicated in the diagram header.

The diagram has two further options: The user can choose the presentation of the abscissa in logarithmic format or linearly directly on the diagram itself. A logarithmic presentation does not make sense if the abscissa variable is pH. The second option 'show data as crosses' allows to display the location of the points calculated in the LJUNGSKILE run in the shape of crosses. This option has been activated in Fig. 3b.

Information about the LJUNGSKILE project file applied in simulating Fig. 3b is available from the 'Information' menu.

A displayed data set can be saved as ASCII file to be converted to a presentation graphics by a suitable graphics program (e.g., ORIGIN). By default the extension ∗.psa is assigned to these ASCII data sets. Only the data of those species selected for the graphical display are saved in the coordinate system displayed when the data is saved. Hence, the saved data corresponds closely to the selections the user made.



(a)

☐ logarithmic abscissa

☐ show data as crosses

(b)

☐ logarithmic abscissa X: 7.619459

☑ show data as crosses Y: 36.14836

Fig. 3. (a) Species distribution of an Fe system as a function of pH. The speciation diagram is governed by the Fe(II) species while the Fe(III) speciation is not visible. Please note that (a) and (b) represent gray scale screen captures. The intention of (a) and (b) is to show the modification in the display upon selecting/deselecting certain species. The details are of less relevance here. In regular use, the information from LDP is exported and processed in an external graphics program (cf. Fig. 6). (screenshot). (b) The same system as given in (a) but without the Fe(II) species. Now, the distribution of the Fe(III) species is visible. The confidence range shown is 95% as indicated in the diagram header. The Fe(II) species are deselected as given by missing hooks in the ‘Species’ menu. The ‘show data as crosses’ field is checked and the pH values for which the speciation is calculated is given as crosses. The both fields below the diagram display the position of the cursor relative to the coordinates of the diagram (screenshot).

4. Example

In the following illustrating example, the solubility of goethite in a natural water will be calculated using several features of the LJUNGSKILE code. After starting the LJUNGSKILE code, the project file ‘example.prj’ should be selected in the ‘Projects’ frame and ‘example.dat’ in the ‘Databases’ frame. In the ‘Project’ frame, the selected project ‘example_Fe.prj’ is given in bold letters. This project file will serve only as a template for a new project file. Hence, a new project file ‘example_Fe.prj’ should be created opening the menu item ‘Save as’ in the File menu.

Subsequently, the ‘Edit project parameters’ button should be clicked in the ‘Project’ frame. The ‘project parameter’ window opens. The content still belongs to the previous project. It is removed by right-clicking each item. A window opens where among several other choices the item ‘Remove species’ will be selected. This process is to be repeated until one last species is left. The last species cannot be removed but must be renamed. By clicking in the field, the content can be edited into ‘Fe+2’. The next two field holds ‘0’ while upon right-clicking the last column ‘Distribution’ the already known field allows to select the item ‘Distribution’ with the three choices ‘Normal’, ‘Uniform’ and ‘Master species’. Here, ‘Master species’ is selected.

Clicking in the empty space below this first row allows to generate new lines which should be filled according to Table 1.

Finally, the solid phase name is changed to ‘Goethite’. Please note that the names must correspond to the phase name given in the data base and are case sensitive. The amount of solid phase is ‘1E−4’. The ‘$CO_2(g)$’ field is checked and the atmospheric CO_2 partial pressure ‘0.00035’ is specified. Pressing the ‘Close edit parameter’ button lets the window disappear.

Upon clicking the ‘Edit water’ button a window opens where the detailed composition of a water can be defined. A ‘water description’, e.g., ‘Iron water’ should be given. The pH does not matter here, because a multiple run simulation will be performed. In case of a single run or a run with another varying entity, e.g., sulphate concentration, the pH must be specified here. The ‘pe’ is 5, ‘temperature’ is set to 12 °C. The elements are specified in the same way as the species previously. Finally, the water composition in Table 2 should have been specified. This water corresponds to a dilute ground water.

Table 1
Fe species, formation constants and uncertainties

Species[a]	Formation constant log *K*	Uncertainty[b]	Distribution
Fe+2	0	0	Master
Fe+3	−13.03	0.2	Normal
FeOH+2	−15.22	0.2	Normal
Fe(OH)2+	−18.7	0.1	Normal
Fe(OH)3	−26.63	0.1	Normal
Fe(OH)4−	−34.63	0.2	Normal
Fe2(OH)2+4	−29.01	0.2	Normal

[a] Note the designation of species in PHREEQC convention. For instance, a species Fe^{2+} is denoted as Fe+2.

[b] The meaning of the term ‘uncertainty’ may vary with the distribution chosen by the user. If the distribution ‘Normal’ is chosen, uncertainty means a 68% confidence limit. If a ‘Uniform’ distribution is chosen, uncertainty means the interval between maximum and minimum value, centered around the specified formation constant.

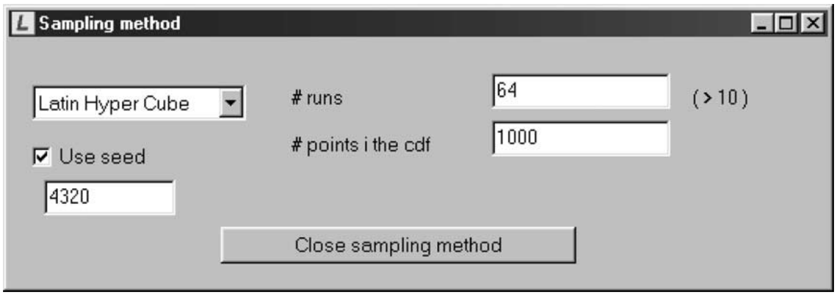
Table 2
Composition of a natural water

Element	Amount (mol dm^{-3})
Cl	2.05E−3
Na	2.1E−2
S	1.2E−3
Ca	1.3E−4
Mg	1.03E−5
N	3.12E−5
Fe	2.5E−6
C	2.34E−5

After pressing the ‘Exit’ button, the project should be saved. The ‘Sampling method’ button opens a window where the following details should be specified (cf. Fig. 4). Then, the window should be closed by pressing the ‘Close sampling method’ button.

After checking the ‘Use multiple run’ option, the following parameters should be specified (cf. Fig. 5).

After closing the window, the project can be started by clicking LJUNGSKILE ’s ‘Simulation’ menu item. The numerical operations are performed in a DOS window by the public domain code PHREEQEC [12]. LJUNGSKILE



Latin Hyper Cube

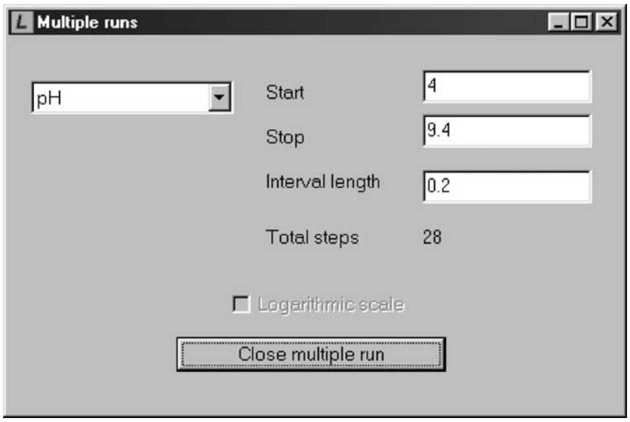
runs 64 (> 10)

points i the cdf 1000

☑ Use seed

4320

Fig. 4. LJUNGSKILE ‘Sampling method window’ (screenshot).



pH

Start 4

Stop 9.4

Interval length 0.2

Total steps 28

☐ Logarithmic scale

Fig. 5. LJUNGSKILE 'Multiple run' windows (screenshot).

provides the input information to PHREEQEC modified according to the LHS procedure. To complete the simulations may take some time. As a rule of thumb, the product of CPU clock speed in GHz and time in minutes is 30.

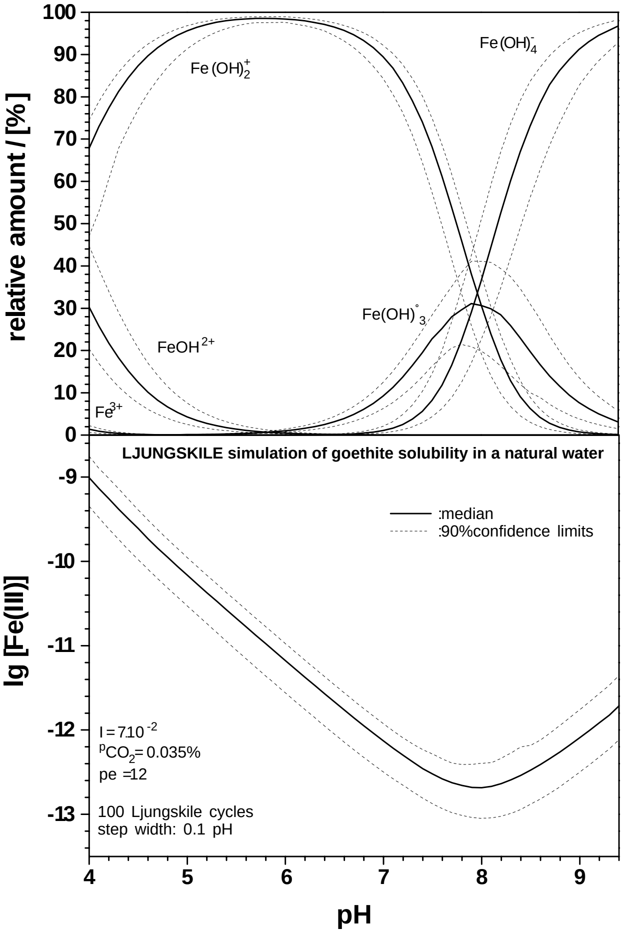


Fig. 6. Solubility curve of $Fe(OH)_{3am}$ (below) and speciation of Fe(III) hydrolysis (top) in the pH range 4–9.4 at $I = 7 \times 10^{-2}$ and atmospheric CO_2 partial pressure. The curves are obtained from LJUNGSKILE simulations on basis of the 'iron.prj' project file.

After successfully performing the LHS calculations, LJUNGSKILE code moves the created output files in sub-directory of the \results\ sub-directory in the same directory where the LJUNGSKILE code is installed. Upon clicking the 'Display' menu option in the LJUNGSKILE main window, LDP starts automatically displaying the last calculated project results.

From these data, Fig. 6 has been generated with ORIGIN using the ASCII output option of LDP.

5. Conclusions

The introduction of uncertainties in computer calculations is getting more and more attention recently. Decision makers now realise that the deterministic results generally presented for a given set of conditions for a given problem, not have to be true. The trueness of a value is an essential epistemological concept in science. Science is based on the assumption that true values exist. Measurements, however, are unable to reproduce a true value without doubt. A measurement is a comparison and, hence, can produce relative measures only. In order to make measurements comparable, these measurements must relate to the same references. This is the metrological concept of traceability [34]. The traceability concept requires that each comparison is accompanied by a statement of measurement uncertainty. Otherwise, comparisons of data obtained at different location or different time are meaningless. In chemistry, the analytical determination of values for a quantity of interest is generally a complicated procedure which, before the advent of high-speed desktop computers, was almost impossible to trace statistically. Modern computer-intensive statistics have changed the situation [18,35]. ISO, in cooperation with other international institutions like IUPAC, has issued a guideline for expression of uncertainty in measurement [16]. This guideline has given rise to proposals for expression of uncertainty in chemical thermodynamic data [36]. In the sequel, methods for the progression of measurement uncertainty information in complex procedures have been proposed [37]. The need for such methods is well documented [15,20,38,39]. There are a series of causes for measurement uncertainty. Rather, the deviation from reality can be significant only due to the fact that one of the input parameters value were not determined with the desired accuracy. It has been shown previously that uncertainties in stability constants [40], enthalpies [41] and water composition [42] to varying degree affects the calculated solubility of any given solid phase, especially actinides and similar elements where good data are scarce. For these purposes several computer programs have been developed, e.g., SENVAR, MINVAR and UNCCON [43,44]. However, it is clear that in a broader perspective the uncertainties in speciation calculations have even wider effects since the speciation often is used as input to other estimations ranging from solvent extraction industry, reprocessing of spent nuclear fuel and basic sorption studies. In this paper, we have shown a computer program, LJUNGSKILE 2.0, which calculates the uncertainties in speciation calculations. It is an easy to use program with good graphical capabilities allowing to investigate the effects of uncertainties in stability constants for any given situation. Is has also been shown, in a simple test case, that the program may also be used for solubility calculations depending on one variable, e.g., pH or sulphate concentration. In addition the partial pressure of CO_2 may be added.

Acknowledgements

This work was partly founded by the Swedish Nuclear Power Inspectorate.

References

[1] W. Stumm, J.J. Morgan, Aquatic Chemistry, 2nd ed., Wiley Interscience, New York, USA, 1981.
[2] F.R. Hartley, C. Burgess, R.M. Alcock, Solution Equilibria, Ellis Horwood Ltd., Chichester, UK, 1980.
[3] J. Bjerrum, G. Schwarzenbach, L.G. Sillèn, Stability Constants of Metal Ion Complexes, Special Publication 6, Royal Chemical Society, London, UK, 1957.
[4] L.G. Sillèn, A.E. Martell, Stability Constants of Metal Ion Complexes, Special Publications 25, Royal Chemical Society, London, UK, 1971.
[5] G. Anderegg, Critical Survey of Stability Constants of EDTA Complexes, IUPAC Chem Data Ser 14, Pergamon Press, Oxford, UK, 1977.
[6] A.M. Bond, G.T. Hefter, Critical Survey of Stability Constants and Related Thermodynamic Data for Fluoride Complexes in Aqueous Solution, IUPAC Chem Data Ser 27, Pergamon Press, Oxford, UK, 1980.
[7] R.M. Smith, A.E. Martell, Critical Stability Constants, vol. 5, Plenum Press, New York, USA, 1982.
[8] R.M. Smith, A.E. Martell, Critical Stability Constants, vol. 6, Plenum Press, New York, USA, 1989.
[9] L. G Sillén, Acta Chem. Scand. 18 (1964) 1085.
[10] J.C. Westall, J.R. Zahary, F.M.M. Morel, MINEQL: A Computer Program for the Calculation of Chemical Equilibrium Composition of Aqueous Systems, MIT Tech Note 18, 1976.
[11] E. Östhols, The C-LETAGROP Program System (Version 1.4.1), Rep TRITA-OOK-2051, Royal Institute of Technology, Stockholm, 1994.
[12] D.L. Parkhurst, Water Resources Investigation Report 95-4227, US Geological Survey, Lakewood, CO, USA, 1995.
[13] T.J. Wolery, EQ3/6, A Software Package for Geochemical Modeling of Aqueous Solutions: Package Overview and Installation Guide (Version 7.0), REPORT UCRL-MA-110662 PT 1, LLNL, Livermore, CA, USA, 1992.
[14] J.W. Morse, Q. Wang, Mar. Chem. 57 (1997) 187.
[15] O. Nitzsche, G. Meinrath, B. Merkel, J. Contam. Hydrol. 44 (2000) 223.
[16] BIPM, IEC, IFCC, ISO IUPAC, IUPAP and OIML, Guide to the Expression of Uncertainty in Measurement, Published in the Name of BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML by International Standardization Organisation, Geneva/CH, 1993.
[17] C. Ekberg, Mine Water Environ. 21 (2002) 45.
[18] G. Meinrath, Fresenius J. Anal. Chem. 368 (2000) 574.
[19] C. Ekberg, A.T. Emrén, Monatsh. Chem. 132 (2001) 1171.
[20] G. Meinrath, S. Hurst, R. Gatzweiler, Fresenius J. Anal. Chem. 368 (2000) 561.
[21] G. Meinrath, B. Merkel, P. Schneider, B. Delakowitz, in: B. Merkel, B. Planer-Friedrich, Wolkersdorfer (Eds.), Uranium in the Environment, Springer, Berlin/D, 2002, pp. 351–368.
[22] R.I. Wielgosz, Anal. Bioanal. Chem. 374 (2002) 767.
[23] BIPM, Reconnaissance Mutuelle, Bureau International des Poids et Mésures, Sèvres/F, 1999.
[24] ISO, General Requirements for the Competence of Testing and Calibration Laboratories, ISO 17025 International Standardization Organisation, Geneva/CH, 2000.
[25] W. Richter, Fresenius J. Anal. Chem. 365 (1999) 569.
[26] http://www.geo.tu-freiberg.de/software/LJUNGSKILE.
[27] C. Ekberg, A. Ödegaard-Jensen, G. Meinrath, LJUNGSKILE 1.0: A Computer Program for Investigation of Uncertainties in Chemical Speciation, SKI Report 03:3, I SRN SKI-R-03/3-SE, Swedish Nuclear Power Inspectorate, 2003.
[28] R.Y. Rubinstein, Simulation and the Monte Carlo Method, Wiley and Sons, New York, USA, 1981.
[29] M.D. McKay, W.J. Conover, R.J. Beckman, Technometrics 21 (1979) 239.
[30] G. Meinrath, C. Ekberg, A. Landgren, J.O. Liljenzin, Talanta 51 (2000) 231.
[31] D.E. Knuth, The Art of Computer Programming, Vol. II: Seminumerical Algorithms, 2nd ed., Addisson-Wessley, Reading, USA, 1981.
[32] L.G. Sillén, Acta Chem. Scand. 16 (1962) 159.
[33] L.G. Sillén, Acta Chem. Scand. 18 (1964) 1085.
[34] International Vocabulary of Basic and General Terms in Metrology, Beuth Verlag Berlin, Germany, 1994.
[35] G. Meinrath, Chemom. Intell. Lab. Syst. 51 (2000) 175.
[36] G. Meinrath, Fresenius J. Anal. Chem. 369 (2001) 690.
[37] C. Ekberg, G. Meinrath, B. Strömberg, J. Chem. Thermodyn. 35 (2003) 55.
[38] R.A. Chalmers, Talanta 40 (1993) 121.
[39] K. Heydorn, T. Anglov, Accred. Qual. Assur. 7 (2002) 153.
[40] C. Ekberg, I. Lundén, J. Stat. Comput. Simul. 57 (1997) 271.
[41] C. Ekberg, G. Skarnemark, A.T. Emrén, I. Lundén, Mater. Res. Soc. Symp. Proc. 412 (1996) 889.
[42] C. Ekberg, A.T. Emrén, A. Samuelsson, Mater. Res. Soc. Symp. Proc. 465 (1996) 735.
[43] C. Ekberg, A.T. Emrén, Comput. Geosci. 22 (1996) 867.
[44] C. Ekberg, S. Börjesson, A.T. Emrén, A. Samuelsson, Comput. Geosci. 26 (2000) 121.