Mine Water and the Environment (2002) 21: 24–35 © IMWA Springer-Verlag 2002

Technical Article

Thermodynamic Prediction in the Mine Water Environment

Günther Meinrath[1] and Peter M. May[2]

[1]RER Consultants Passau, Schießstattweg 3a, D-94032 Passau, Germany and the Inst of Geology, Univ of Mining and Technology Freiberg, Freiberg, Germany; [2]AJ Parker Cooperative Research Ctr for Hydrometallurgy, Chemistry, Murdoch Univ, Murdoch WA 6150, Australia; e-mail: rer@panet.de

Abstract. Reliable prognosis by geochemical modeling requires thermodynamic data with a stated measure of reliability. Due to the complexity of experimental and numerical techniques commonly applied to extract thermodynamic quantities of chemical reactions, a meaningful method for evaluation of such a reliable measure is currently not available. Some reasons for this situation are exemplified in a simulated solubility study of Fe(III) as a function of pH, where the effects of correlation, nonlinearity and sample size are discussed. To improve the accuracy of thermodynamic quantities for chemical reactions, researchers are attempting to generate internally consistent databases and introduce metrological concepts into the evaluation of thermodynamic data. The concept of measurement uncertainty is discussed, using a cause-and-effect diagram for the Fe(III) simulation example.

Key Words: correlation, geochemical modeling, metrology, nonlinear parameter optimization, speciation, statistics, thermodynamic data, thermodynamic database

Introduction

The ever increasing availability of cheap computing power experienced during the past decade has rapidly promoted the field of geochemical modeling. Efficient transport algorithms and stable speciation codes now allow implemention of complex simulation models for detailed inquiry into the likely hydrogeological and geochemical fate of environmentally sensitive sites. Geochemical modeling is largely an interdisciplinary activity combining information technology, physics, chemistry, hydrogeology, geology, geochemistry, rheology and others. In order to achieve reliable understanding and insight, the current limitations (and ongoing activities to overcome or circumvent these limitations) must be communicated among these participating disciplines. Meanwhile, it has become apparent that the advances due to high-speed computers are only one side of the coin. The quality of the input data available for predictive simulation is the other, darker side. Only an admittedly rather limited discussion of activities dealing with this issue can be presented here.

Speciation – an outline

In the 1980s, availability of instrumental techniques for elemental analysis, i.e. Atomic Absorption Spectroscopy (AAS) and Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES), allowed routine determination in the µg per kg to ng per kg range of a metal ion concentrations in environmental matrices. However, it soon became apparent that more detailed information on the specific form of a substance in the matrix was necessary in order to quantify/predict the effects of that substance.

Here, the terms effect and specific form are meant quite generally. Effect may be the sorption tendency of a substance on rock or the poisonous consequences of it; specific form refers to the chemical species that occur, which may be anionic, neutral or cationic. The chemical speciation includes the redox state of a metal ion, which often determines molecular-level characteristics (e.g., methylmercury is much more toxic than mercury vapor whereas the arsenates and arsenites have a higher toxicity than the abundant monomethyl arsonic acid or dimethyl arsinic acid).

It is important to appreciate that the meaning of the term speciation depends on the framework of interest in a pragmatic way. As a consequence, there are many more methods to determine speciation than can be considered here, where the focus is on environments affected by mining and surface alteration. A wide field for speciation studies has been opened up by analytical chemistry, especially in the development of coupled techniques such as Inductively Coupled Plasma–Mass Spectrometry (ICP-MS) or High Pressure Liquid Chromatography–Mass Spectrometry (HPLC-MS). Such techniques allow a sensitive, and often also selective, determination of specific compounds in complex matrices.

In many cases, however, the species of interest are kinetically quite labile and exist only in a dynamic equilibrium. Techniques with separation times on the order of seconds are not capable of resolving very fast equilibria, which are established on a time scale of µs or ns. Any attempt to describe natural aqueous systems by a mechanistic kinetic approach is thus likely to be in vain (Perkins et al. 1997). Species formed in nature by fast, usually ionic, equilibria are difficult to identify and even more difficult to quantify. Nevertheless, techniques have been devised to study even such labile species as $FeOH^{2+}$, $CaSO_4°$, etc., despite the fact that there is no chance to isolate or concentrate such entities. The key to any description of labile equilibria are the appropriate fundamental thermodynamic relationships.

Currently, however, it seems that aquatic chemistry is struck down by its own success. Thermodynamic data on species like $CaCl^+$, AlF_2^+, CH_3COOH, and many thousands more, get assembled into large databases. These databases are coupled to speciation codes by the use of fast computers and then linked to complex transport models with the goal of predicting long-term behaviour of plumes and sites and even sites that may be created in the future.

The main caveat is that most of the information available on species, commonly thermodynamic data, was not determined or processed with such far-reaching applications as geochemical transport modeling in mind. Chemists are primarily interested in structure, energy, and the relationship among the elements of the periodic table. Thermodynamic data have been serving this important purpose of chemical understanding of complex processes for more than 100 years. It should be remembered that chemists had a three-dimensional idea of compounds before an atomic model had even been suggested and that a symbolic language for complex interatomic interactions existed long before the idea of chemical bonding was substantiated by quantum mechanics. On the other hand, long-term prediction had hardly been considered by aquatic chemists until about 1980.

Thus, the availability and applicability of chemical data in the geochemical modeling framework offers new challenges. In order to tackle this challenge, the present position, the current activities, and the future needs have to be carefully pondered.

Thermodynamics and kinetics

Chemical reasoning is founded on two pillars of basic chemistry from the 19th century. The first pillar is the Law of Mass Action (Guldberg and Waage 1865; 1867). The second pillar is provided by the three Laws of Thermodynamics (Brdicka and Dvorak 1982). By combining the two, the amount of energy, ΔG, provided by a process and available for work under isobaric and isothermic conditions is given by

$$\Delta G = \Delta H - T\,\Delta S \qquad (1)$$

where ΔG is the difference in Gibbs' energy, ΔH is the enthalpy difference, and ΔS is the entropy difference over the process at a given absolute temperature T in Kelvin. Eq. 1 can be considered in a variety of ways, but such a discussion is outside the scope of the present work. The Law of Mass Action is given for a general reaction

$$a\,A + b\,B <==> c\,C + d\,D \qquad (2)$$

of reactants A and B with respective stoichiometric coefficients a and b, and products C and D with respective stoichiometric coefficients c and d (Eq. 3a)

$$\frac{\{C\}^c \quad \{D\}^d}{\{A\}^a \quad \{B\}^b} = K \qquad (3a)$$

The brackets indicate activities of species A, B, C, and D, usually in the molal scale. The link between Eqs. 1 and 3a is given by

$$\Delta G°_R = -\,RT \ln K° \qquad (4)$$

Here, the subscript R indicates reaction 2 and superscript ° refers to a standard state, defined according to the nature of species A, B, C, and D. Since Eq. 1 expresses a fundamental law of nature, Eq. 4 links the equilibrium constant K of reaction 3a directly to this fundamental law of nature. Obviously, the thermodynamic datum K° under conditions defined by the standard state under which Eq. 3a is referred to is a fundamental constant of nature, too.

From Eq. 4, -RT ln K° is an energy. Energies are not directly observable. In this case, the energy manifests itself as an activity ratio of the species A, B, C, and D according to Eq. 2. Generally, only the concentration ratios of A, B, C, and D can be observed, from which the activities must be derived.

$$\frac{[C]^c \quad [D]^d}{[A]^a \quad [B]^b} = K' \qquad (3b)$$

The square brackets denote concentrations. Except in extremely diluted solutions, $K' \neq K°$. The task of describing the relationship between K' and K° for arbitrary ionic equilibria has challenged chemists for almost a century and the problem is still not resolved satisfactorily (May 2000).

To establish a relationship, activity coefficients, γ, have been defined:

$$\{A\} = \gamma_A [A] \qquad (5)$$

Thus, the relationship between K' and K° based on the reaction of Eq. 2 is

$$K^\circ = K' \frac{\gamma_C^c \, \gamma_D^d}{\gamma_A^a \, \gamma_B^b} \qquad (6)$$

Different approaches have been taken to evaluate the activity coefficients. Most of these approaches are directly related to the Debye-Hückel theory (Debye and Hückel 1923), where for very dilute solutions, a limiting law has been established:

$$\lg \gamma_C = -A\, z_C^2 \sqrt{\frac{1}{2}\sum_i m_i z_i^2} \qquad (7)$$

The term on the right hand side under the root is commonly referred to as ionic strength, I, which sums up all ions in solution multiplied by their respective squared charge in order to calculate the activity coefficient of charged species C with formal charge z_C^2. For a typical electrolyte solution to be treated in mine water studies, the range of validity of the Debye-Hückel Eq. 7 is much too limited. As a rule of thumb, Eq. 7 may be applied accurately up to only $I = 0.001$ mol kg^{-1}. For higher ionic strengths, extensions (Davies 1961), approximations (Debye and Hückel 1923) or phenomenological equations (Pitzer 1991) have been derived.

Either way, thermodynamic data describe macroscopic equilibria. In nature, processes are not usually at equilibrium. Often, the term steady state seems better suited. One reason is the lack of unequivocal experimental methods to assess whether an equilibrium state has been attained. Metastable states are well known in nature; the calcite-aragonite-vaterite system is a prominent example (Mucci 1983; Wolf et al. 1996). Changes in a system take place due to its displacement from equilibrium. Otherwise, the system is inert. Disequilibria are a central driving force in nature. As disequilibria tend towards equilibrium, the processes initiated by the disequilibrium can be described kinetically. Chemical kinetics follow laws that to a first approximation can be described using the following general form

$$\frac{\partial c}{\partial t} = k \prod_i \{X_i\}^n \qquad (8)$$

where k is a rate constant, while the X_i are time-dependent concentrations of species participating in the kinetic process described by Eq. 8. The exponent may differ for each component X_i and need not necessarily be an integer quantity. The sum of all exponents n in these steps defines the overall order of the kinetic process. Even seemingly simple chemical reactions are often complex processes on the molecular level and it is not unusual to find that even a simple reaction is separated into several distinct kinetic steps. The numerical complexity of the applicable differential equations may then require simplification of any mathematical model (Okino and Mavrovouniotis 1998).

Contrary to what is often said, the rate law Eq. 8 cannot be inferred from the Law of Mass Action; it is strictly an empirical relationship. Seemingly analogous chemical reactions may follow completely different rate laws. To describe complex equilibria by the respective rate laws, an enormous amount of information must be determined. In order to determine a rate constant and the order of a chemical process, conditions must be found where the process of interest can be studied without interference of competing processes. Often, such conditions cannot be realized. For example, it is almost impossible to study exchange processes in the coordination sphere of iron(III) by acetate ions without interference by hydrolysis in the pH range 4-9, which is the common pH range of natural aqueous systems.

While thermodynamic constants K° are independent of the concentration range of the reaction itself, rate constants k may be highly sensitive to total concentrations. Moreover, if the mechanism of a reaction changes, the rate law will change too. Last but not least, kinetic rate parameters are not fundamental constants of nature. When determining a rate constant under a given conditions, transferability of the information extracted from experimental data to different conditions is not justified. Natural aqueous systems are multicomponent systems. The species concentrations are generally quite low and the variations in composition large. The use of kinetic data therefore is easily challenged, whereas the use of thermodynamic data is justified by fundamental thermodynamic relationships, even though the validity of the assumption of thermodynamic equilibrium may be difficult to defend. Nevertheless, under certain conditions, the use of more or less empirical kinetic models may be appropriate. A typical situation occurs with acid mine drainage, for which a rate-determining step has been identified and adequately described (Lowson 1982; Posner 1953; Singer and Stumm 1970).

Statistical Evaluation of Thermodynamic Data

All experimental quantities are affected by the limited precision of the measurement process. Hence, conclusions on the system under study and decisions on the basis of the findings depend on the uncertainty in the data. Dealing with such situations is a fundamental problem of all sciences. Nevertheless, the tool to tackle this situation, statistics, is by no means easily applied in chemistry (Thompson 1997). This is particularly true of thermodynamic data, where the predominant errors tend to be of a systematic (rather than random) nature.

Thermodynamic data for geochemically relevant reactions are tabulated in thermodynamic databases. Often, several determinations of a particular equilibrium constant have been reported, sometimes with considerable discrepancy (Meinrath et al. 2000; Millero et al. 1995). To discuss a few relevant issues, the hydrolysis reaction of Fe(III) shall be taken as an example. Iron plays an important role in acid mine drainage and so this choice is not entirely arbitrary.

Figure 1 shows a calculated solubility of Fe(III) as a function of pH. The thermodynamic data from the JESS database given in Table 1 have been used. This choice is somewhat arbitrary but justified by the fact that this database is available over the worldwide-web (May and Murray 2000), together with a detailed guide (May and Murray 2001). Figure 1 gives the solubility of Fe(III) in equilibrium with amorphous iron hydroxide, $Fe(OH)_3(s)$, having a solubility product lg $K_{sp}=3.20$. For comparison, thermodynamic data from a few other sources are given in Table 1. Please note that there is no intention here to present a concise review of Fe(III) hydrolysis data for geochemical modeling. The interested reader is directed to Millero et al. (1995) and Khoe et al. (1986) and literature cited therein. From Figure 1, the contribution of the polynuclear species $Fe_2(OH)_2^{4+}$ and $Fe_3(OH)_4^{5+}$ can be seen to be almost negligible. The solid $Fe(OH)_3(s)$ is the most soluble Fe(III) hydroxide phase. A reduction in the concentration of $Fe^{3+}_{(aq)}$ due to a less soluble phase further reduces the relative contribution of the polynuclear species. Nevertheless, Table 1 indicates that there are 20 entries in the relevant bibliography of the JESS database (Daniele et al. 1994; Khoe et al. 1986).

A sceptical recipient of information provided by thermodynamic databases and compilations may wonder about the dependability and reliability of thermodynamic values published in the literature. Indeed, the reliability of thermodynamic data has been recognized as a key limitation of the quality of speciation calculations (May and Murray 1991). A major criticism with available thermodynamic data is a lack of proper statistical assessment. Most thermodynamic data are reported in the literature without even the faintest statistical corroboration (Meinrath et al. 1999). If statistics are invoked at all, classical linear statistics is applied in most cases. It is, therefore, almost impossible to know to what extent a proposed chemical interpretation of experimental data is based on fact or on subjective interpretation (Chalmers 1993).

This obvious deficiency of existing thermodynamic data has an equally obvious reason: the complexity of the model functions do not easily lend themselves to a statistical analysis. The model function for the solubility of Fe(III) as a function of pH (Figure 1) is given by Eq. 9 below, with the constants K_s and K_{nm}

$$Fe(III)_{total} = K_s\{H^+\}^3 + K_sK_{11}\{H^+\}^2 + K_sK_{12}\{H^+\} + K_{13} + \frac{K_sK_{14}}{\{H^+\}} + K_s{}^2K_{22}\{H^+\}^4 + K_s{}^3K_{34}\{H^+\}^5$$

Table 1: Selected hydrolysis constants of Fe(III) from different compilations

$Fe(OH)^{2+}$	$Fe(OH)_2^+$	$Fe(OH)_3^\circ$	$Fe_2(OH)_2^{4+}$	$Fe_3(OH)_4^{5+}$	$Fe(OH)_4^-$	$Fe(OH)_{3(s)}$	Source
61[a]	9[a]	2[a]	20[a]	5[a]	3[a]		JESS
-2.19	-5.67	-12.0	-	-	-21.6		B/M
-2.5	-6.5	-12.0	-3.0	-6.1	-21.6		CIV
-2.19	-5.67	-12.56	-2.95	-6.3	-		PQC
-2.19	-4.59	--	-2.85	-6.3	-		NIST
-2.05	-6.35	-13.45	-2.90	-6.04	-21.43	3.20	JESS[b]

[a] number of literature references in the JESS database for a particular species; [b] the weights given in JESS database indicate that only $FeOH^{2+}$ can be considered to be well characterised. B/M: Baes and Mesmer 1976; JESS: Joint Expert Specation System (May and Murray 1991a; 2000); NIST: Martell et al.1993; PQC: PhreeqC Default Database (Parkhurst 1995); CIV: Chemval (1992)

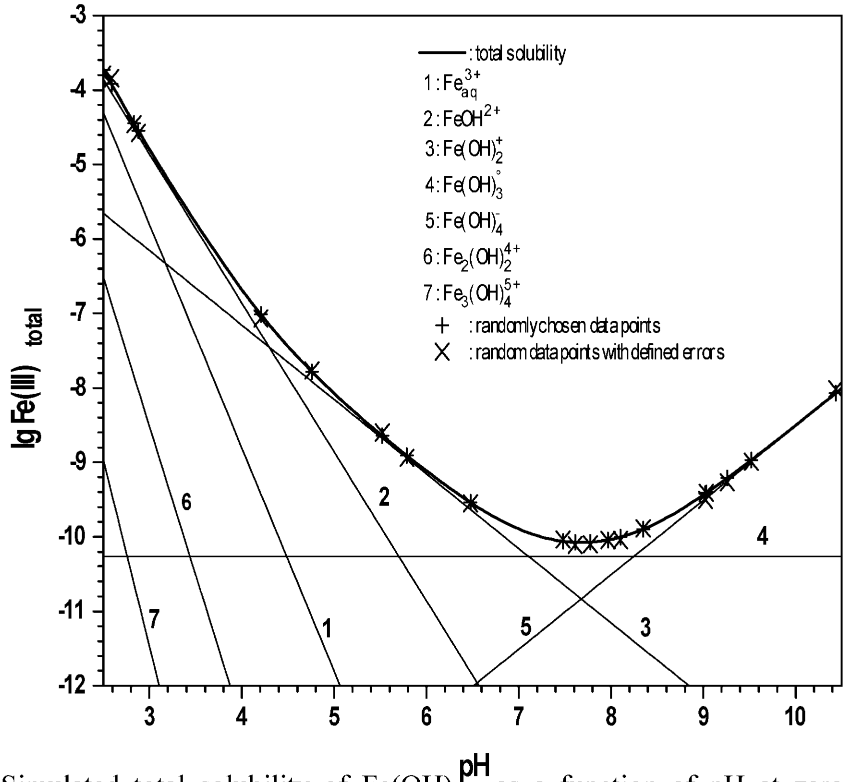

Figure 1. Simulated total solubility of $Fe(OH)_{3(s)}$ as a function of pH at zero ionic strength together with calculated concentrations of the individual solution species. The locations of 20 randomly drawn data points are included.

Table 2: Parameter estimates, approximate standard deviations (1σ) and variance-covariance matrices[a] of the parameters K_{11}, K_{12}, K_{13}, K_{14} and K_s obtained for a synthetic data set generated with normally-distributed noise. Only the upper triangle of the variance-covariance matrix is given. Correlations are given in the lower section.

lg K_{11}	lg K_{12}	lg K_{13}	lg K_{14}	lg K_s	
$-2.1_8 \pm 0.2_8$	$-6.4_3 \pm 0.2_2$	$-13.5_6 \pm 0.2_3$	$-21.8_1 \pm 0.2_3$	$3.2_7 \pm 0.2_3$	
0.08172	0.06284	0.0666	0.0657	-0.0658	K_{11}
+0.985	0.04971	0.05146	0.05123	-0.05118	K_{12}
+0.984	+0.974	0.05608	0.0537	-0.0541	K_{13}
+0.989	+0.998	+0.976	0.0540	-0.0535	K_{14}
-0.995	-0.992	-0.987	-0.995	0.0535	K_s

[a] calculated from a SIMPLEX analysis on basis of Spendley's approximation (Spendley 1969)

defined as in Table 1. Since a system at I=0 is treated to avoid the additional complications of ionic strength correction, activities are assumed equal to concentrations and pH = $-\log\{H^+\}$. Despite the fact that Eq. 9 is composed of additive terms, all terms except the first on the right hand side are nonlinear with respect to parameters. More to the point, the correlations among the parameters are non-zero. Using classical statistics (evaluation of residual sum of squares and confidence regions by the Gauss-Legendre algorithm), the statistical properties of the fit of a given experimental data set with a given model function can be estimated. To illustrate such an approach, a synthetic data set of 20 data points has been generated. These points, given in Figure 1 by crosses (+) are contaminated by normally distributed random noise with standard deviation σ = ±0.036 in pH and 3% relative in total Fe(III) concentration. The small magnitude of the generated noise can be judged by comparing the uncontaminated data (+) to the noise-affected data (X) in Figure 1. The resulting least-squares estimates of the parameters K_{11}, K_{12}, K_{13}, K_{14} and K_s are given in Table 2, together with the variance-covariance matrix.

The parameters K_{22} and K_{34} cannot be independently determined from the synthetic solubility data because their relative contributions are too low. Their values have accordingly been fixed at the given values. The square roots of the diagonal elements of the variance-covariance matrix give the marginal confidence ranges. Due to the large statistical correlation among the parameters, the small scatter in the data results in appreciable estimates of uncertainty for the parameters. Figure 2 gives the 68% and 95% confidence ellipse for parameters K_{11} and K_{12} together with the marginal confidence ranges and the calculated mean values. Since the data are simulated, the true value is known (cf. Table 1) and can be compared with the regressed values.

The main axes of the ellipse are strongly tilted as a result of the correlation between the parameters of +0.98. The differences between the calculated and the true values are evident. This difference is termed bias and is exacerbated by nonlinearity in the model equation (Box 1971).

A comparison of the data in Table 2 and Figure 1 shows that the estimates of uncertainty in the parameters almost cover the range amongst the different data sources given in Table 1. A comparison of the synthetic data from Figure 1 with experimental data (i.e., Figure 6 from Millero et al. (1995)) shows that experimentally observed scatter is much larger than those given by Figure 1. The linear confidence ranges given in Table 2 and Figure 2 might seem exaggerated in comparison to the small scatter in the synthetic data. The approximations inherent in an interpretation of nonlinear model functions by the linear Gauss-Legendre algorithm might be suspected of causing an overestimation of the confidence limits. Therefore, parameter- and model-free computer-intensive statistics were invoked to assess the data without such approximations (DiCiccio and Romano 1988; Efron 1979; Efron and Tibshirani 1991; Meinrath 1999; 2000). The probability densities of the five optimized parameters, determined by a balanced bootstrap method, are shown in Figure 3 with additional parameters in Table 3. The non-parametric bootstrap results give a clear message: the probability densities are very broad, with 68% confidence ranges between 0.4 and 1.3.

There seems to be a discrepancy between the obvious coincidence of the synthetic experimental data points and the theoretical curve and the statistically assessed large confidence ranges. However, the statistical analysis is not answering how good is the agreement between an optimum parameter model and some experimentally obtained data points. Instead, the analysis gives an indication of the likely range where the parameters should be expected from repeated experiments. This expectation is not only influenced by the agreement between model and the data points but also by the complexity of the model function, the correlation of the parameters, and the number and range of data points. For single parameter models, the

Table 3: Medians, upper and lower percentiles for parameters K_{1n} (n=1-4) and K_s from a balanced bootstrap for 20 and 40 randomly generated data points

	lg K_{11}	lg K_{12}	lg K_{13}	lg K_{14}	lg K_s	
median	-2.3_1	-6.5_2	-13.6_5	-21.9_2	3.3_7	20 data
Upper .68	$+0.8_0$	$+0.6_6$	$+0.7_0$	$+0.7_0$	$+0.4_3$	
Lower .68	$-1._4$	-0.3_8	-0.4_5	-04_5	-0.7_0	
median	-2.3_8	-6.6_6	-13.8_3	-22.0_4	3.5_0	40 data
Upper .68	$+0.7_1$	$+0.6_5$	$+0.6_5$	$+0.6_7$	$+0.2_9$	
Lower .68	-0.3_9	-0.2_7	-0.2_9	-0.2_9	-0.6_6	

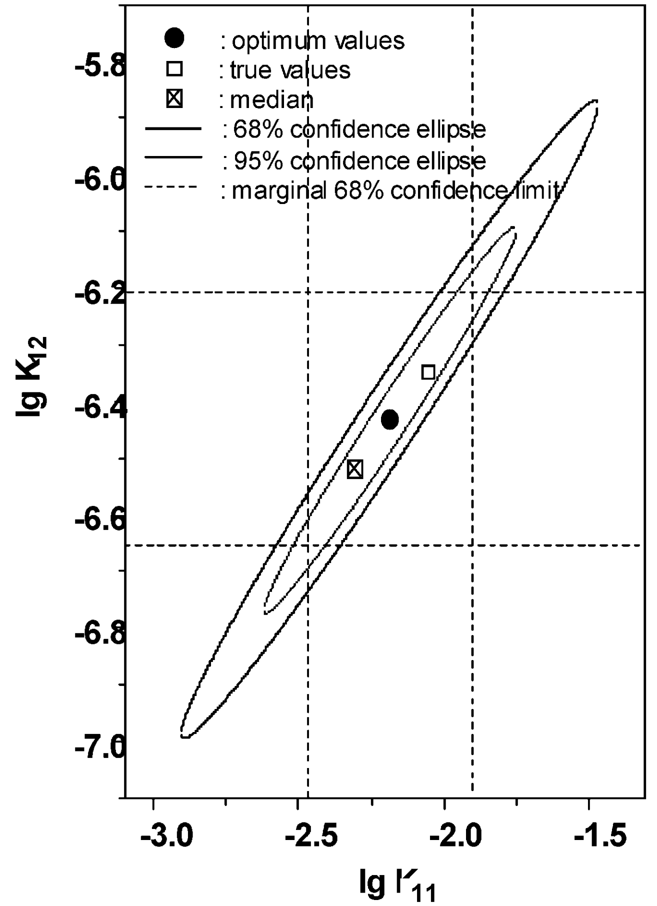

Figure 2. 68% and 95% confidence ellipses in parameter space (K_{11},K_{12}) obtained from a linearization of the model function. The true values, the optimum values from the linearization approach and the medians from a balanced bootstrap analysis are given. For comparison, the marginal confidence regions are included.

degrees of freedom (d.f.) is given by the difference between the number of available data points minus one (for the single parameter). With multiple parameters, the model function can be assessed in an overall way (i.e., by testing hypotheses on different parameter combinations), but there is no corresponding concept for the individual parameters of a model. Nevertheless, increasing the d.f. by doubling the amount of data results in sharper probability densities (given as dashed curves in Figure 3). It should be noted that despite the considerable uncertainties in the individual parameters, the resulting fitted solubility curves are quite similar. But the large correlations cause considerable flexibility in representing a set of total concentrations by different parameter combinations.

Trends in measurement of thermodynamic data

Despite the obvious and urgent need for a dependable assessment of the reliability and transferability of published thermodynamic data, existing methods, at best, meet only the most rudimentary requirements (Meinrath 2000b). No recipe of procedures and requirements has so far been formulated that allows the information in a given experimental data set to be assessed unambiguously. Such requirements must ensure comparability and consistency among all data reported in the literature.

There a two major directions of current research that are mutually fruitful. The first direction constructs a framework and respective tools to ensure consistency of existing data. There is probably no alternative to such an approach because of the enormous amount of already published thermodynamic data that cannot be properly assessed due to missing information (May and Murray 1991; 2001). One of the most common and serious problems concerns the loss of information that occurs in tabulating standard thermodynamic parameters without details of the functions used to derive them. With aqueous solutions in particular, large errors are likely to arise in modeling calculations based on equations for the effect of ionic strength that differ from those employed in determining the standard values.

The second direction is achieved by the increasing emphasis being given to quality assurance and comparability of analytical data (Ellison et al. 1997; Meinrath 2001; ECN 2000). Thermodynamic data is evaluated from analytical measurements. Therefore, the current developments in quality assurance and quality control of analytical data should also encompass the determination of thermodynamic data.

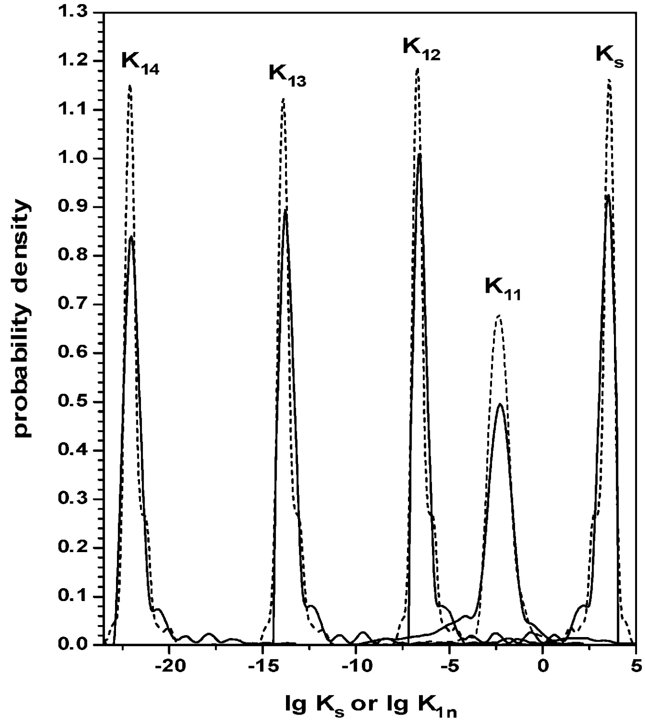

Figure 3. Comparison of empirical probability densities for the thermodynamic parameters K_{1n} (n=1-4) and K_s for a sample size of 20 (solid curves) and 40 (dashed curves) data points.

The major stipulations in this regard are due to the world metrological organisation, the Bureau International des Poids et Mésures (BIPM 1993) and the Mutual Recognition Agreement (MRA) (Richter 1999). The MRA is an international agreement between the national metrological institutes of 38 developed countries on the mutual acceptance of results of chemical measurements. The MRA allows confidence in procedures and methodologies even in cases of discordant interests. Such agreements are strongly political but are based on the conviction that objective criteria and rules can be defined that offer more benefit than insisting on one's own sovereignty. Chemical measurements, in fact, play an increasingly important role in a globalized market place (Bièvre et al. 1997; Bièvre and Taylor 2000; Ellison 1997). Thermodynamic data are applied in modeling of mine water characteristics, nuclear waste disposal (Ekberg et al. 1997), environmental restoration (Nitzsche et al. 2000), bio-inorganic systems (Königsberger and Tran-Ho 1997; May 1995), etc. There is no doubt that thermodynamic data are being applied in making important decisions that are likely to affect many people, sometimes over a time scale of 10^4 years (Bièvre et al. 1997) and even more.

It has been accepted, albeit with some reluctance, that the rules of metrological science can be applied to chemical measurements (King 1997) and, consequently, that there is no alternative but to develop and apply metrological rules to analytical measurements (King 1997b; Meinrath 2000b). Otherwise, there seems little doubt that unacceptable discrepancies in available compilations of thermodynamic data will persist and seriously hamper future scientific and technological development (Ekberg et al. 1997; May and Murray 1991c, 2001; Meinrath et al. 2000; Meinrath 2000b; Nitzsche et al. 2000).

Metrology is based on traceability and the concept of measurement uncertainty (BIPM 1993, BIPM 1994; EURACHEM/CITAC 1995). Traceability is defined as the "property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties" (BIPM 1994). It is evident from this definition that confidence in the concept of metrology depends on the selection of reference materials. The concept of measurement uncertainty rests upon the identification of all sources of variation that might significantly affect the measurement. It is obvious that this relies heavily on the responsibility of the experimenter, who must identify all contributing sources of variability (Ellison et al. 1997; May et al. 1982). Otherwise, one can improve the results by simply ignoring contributions of uncertainty. The price to pay is then a lack of comparability–the present situation.

To support the identification and quantification of sources of variability in the results of chemical analysis, cause-and-effect diagrams are helpful. Such cause-and-effect diagrams are even more versatile in chemical thermodynamics because the number of techniques for studying chemical equilibria is limited and so characteristic diagrams can be developed for particular methods. For a solubility study as simulated in Figure 1, a cause-and-effect diagram is given as Figure 4.

A cause-and-effect diagram summarizes the uncertainty contributions affecting a measurement (Ellison & Barwick 1998). Measurement uncertainty is devided into two types (BIPM 1993). Type-A uncertainty is obtained from the experimental data itself and can be evaluated by statistical methods, while Type-B uncertainty comes from other sources associated with auxiliary data or separate experimentation. In the given situation, Type-A uncertainty contributions, such as different types of correlation, are intertwined with each other and are given in the right-hand-side box. The Type-B uncertainties are summarized in the fishbone part of the cause-and-effect diagram. In order to assess the information on the true (but unknown) value of the thermodynamic parameter(s) enclosed in the experimental data, a rather complex procedure for statistical analysis needs to be implemented (Meinrath 2000, 2001).

A cause-and-effect diagram takes into account variability from the main elements of a solubility determination: analysis of the metal concentration, analysis of the ligand concentration, pH measurement (if pH sensitive equilibria are being studied), ionic strength correction and temperature. Such a diagram has recently been applied to the thermodynamic data for a carbonate complexation reaction where the determination of the free carbonate concentration had to be integrated (Meinrath 2001). Such diagrams can be designed for each experimental method and show at a glance what uncertainty components have been included into the evaluation of a thermodynamic constant (Meinrath and Lis 2002). If the magnitudes of these uncertainty components are given, the data can be reevaluated at a later time, i.e. in order to verify whether relevant uncertainty components have been neglected in an evaluation. If similar approaches are applied in the evaluation of thermodynamic data from other experimental methods (such as Meinrath and Lis 2002), thermodynamic data from different methods become more easily comparable.

Filtration is an often significantly underestimated as a contributor to measurement uncertainty. Solubility is defined as the amount of substance in solution at stable or metastable equilibrium with a solid phase after filtration through 0.45 μm pore size. This pore size is not necessarily sufficiently small to exclude all particulates, especially in systems like iron(III). Even though the analysis of the sample after filtration may be very accurate and precise, an assessment of the contribution of measurement uncertainty from filtration can be very tricky. Here, educated and reasonable estimates are often the only viable way to adequately account for experimental variability.

Since natural aqueous systems have near-neutral pH values, hydrolytic reactions always have to be taken into account. Establishing a measurement uncertainty for the pH, however, is by no means trivial (May et al. 1982; Meinrath and Spitzer 2000; Sigel et al. 1991). It is now accepted that current IUPAC recom- mendations on pH measurement are metrologically questionable (Baucke et al. 1993). The forthcoming IUPAC recommendation include a multi-point calibration procedure (IUPAC 2002). The IUPAC recommenda-

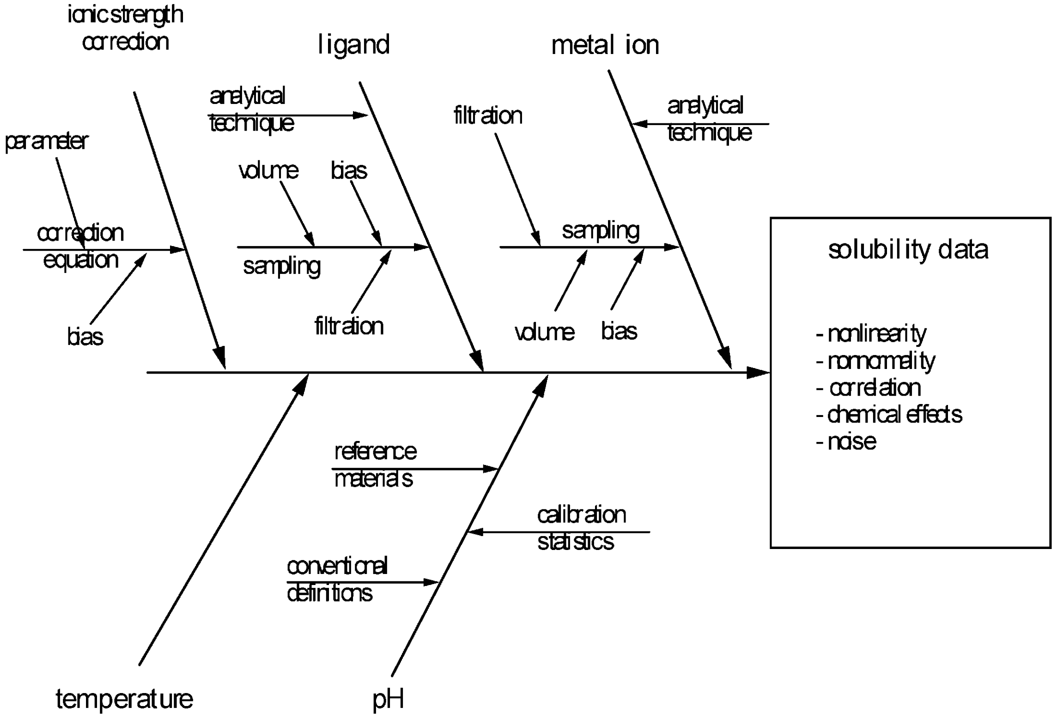

Figure 4. Schematic cause and effect diagram including the major sources of measurement uncertainty for solubility studies.

tion does not include natural waters (as has been the case with previous recommendations on pH), despite the fact that pH is an important parameter affecting reliability of geochemical transport prognosis (Meinrath and Nitzsche 2000).

Ionic strength correction is another often ignored source of uncertainty. Different equations can be applied to account for such effects, but none is satisfactory. Simple equations, i.e. the Davies equation or the extended Debye-Hückel equation, have a small number of parameters but are unable to deal with electrolyte solutions of moderate or high concentration. Other approaches, i.e. the Pitzer approach (Pitzer 1991), may be able to account for concentrated systems but need a large number of experimentally determined parameters. Currently, Pitzer equations are almost exclusively calculated without taking into account the unavoidable measurement uncertainty in each parameter. It is, however, evident that the uncertainty contributions from Pitzer's equations are by no means negligible (Marshall et al. 1995). Hence, the experimenter's task is to identify the break-even point between uncertainty due to misfit and uncertainty due to over-parametrization. Such a point of view is especially important for mine waters, which can have ionic strengths where a classical electrolyte solution model is no longer feasible (Nordstrom et al. 2000). To avoid a further extension of this paper, ionic strength effects have not been included in the simulated solubility study of Fe(III).

Temperature effects also play a role, especially with metal-ligand interactions where covalent bonding becomes significant. For predominantly ionic interactions, the enthalpy changes are difficult to measure due to their weakness. Nevertheless, a larger number of enthalpies for ionic metal-ligand interactions have been reported. It is likely that much of these data do not reflect enthalpy changes as much as statistical correlation (Krug et al. 1976).

Each experimental data point of any solubility study is affected by these measurement uncertainties. Measurement uncertainty adds to the statistical effects that have been discussed in some detail. The uncertainty ranges given in Tables 2 and 3 must be considered as lower limits. Experimental errors that are systematic in nature worsen matters greatly. The statistical analysis summarized in Figures 2 and 3 does not consider contributions from measurement uncertainty but only contributions to the uncertainty in the value of a parameter from the statistical analysis. For a data set obtained by experiment, measurement uncertainty contributions must be added. Finally, Fe(III) in natural aqueous systems is mostly not available as defined solution complexes but is often bound to particulates (Palmer 1987), which contributes further uncertainty.

Conclusions

The thermodynamic quantities that characterize chemical reactions are fundamental constants of nature. Thus, thermodynamic data are universally valid and independent of time and space. Other data, i.e. kinetic rate constants, are not similarly well founded as a basis for predictive simulation, i.e. in geochemical transport prognosis. Despite their universality, however, thermodynamic data of chemical reactions do not lend themselves to easy determination. Without the availability of computers, experimental ingenuity and graphical evaluation techniques had to be devised to extract information about thermodynamic quantities (Hartley et al. 1980). Digital data treatment techniques have revolutionised this situation but now the consequences of data quality are becoming a crucial issue. This is increasingly relevant the more the thermodynamic data are applied under circumstances where the agreement between the simulation and reality is crucial. In other words, concepts ensuring comparability and accuracy become essential whenever the thermodynamic data are used for predictive simulation.

In an example, the evaluation of thermodynamic constants from simulated solubility data of Fe(III) has shown that the uncertainties in the evaluated constants may be considerable despite close agreement between experimental data points and the theoretical curve. Parameter correlation and bias due to nonlinearity are the main reasons for this situation. The sample size and the scatter in the data are other contributions to the total width of the probability distributions. Such effects are often not considered in statements about the accuracy of existing thermodynamic data. Thus, these data are not easily compared and their reliability is very difficult to assess. Not surprisingly, reproducibility of thermodynamic data is often poor (Hefter 1982; May and Murray 2001; Ekberg and Emrén 2001; Meinrath et al. 2000).

Attempting to cope with the demand for thermodynamic data having a precisely stated measure of dependability, researchers are attempting to assure consistency among thermodynamic data (May and Murray 2001; Olbricht et al. 1994) and introduce metrological concepts into chemical thermodynamics (Meinrath 2000b; 2001). Cause and effect diagrams provide a unified scheme to communicate and evaluate contributions of measurement uncertainty to a thermodynamic quantity but require complex data evaluation methods. It should finally be noted that the assessment of algorithm and software reliability has consequently become an important field of metrology (Cox et al. 2001).

References

Baes CF, Mesmer RE (1976) The Hydrolysis of the Cations, J.Wiley and Sons, London, UK

Baucke FGK, Naumann R, Alexander-Weber Ch (1993) Multiple-point calibration procedure with linear regression as a proposed standardization procedure for high-precision pH measurement. Anal Chem 65: 3244–3251

De Bièvre P, Kaarls R, Peiser HS, Rasberry SD, Reed WP (1997) Protocols for traceability in chemical analysis. part I: definitions and terminology. Accred Qual Assur 2: 168–179

De Bièvre P, Taylor PDP (2000) "Demonstration" vs. "designation" of measurement competence: the need to link accreditation to metrology. Fresenius J Anal Chem 368: 567-573

BIPM (1993) Guide to the Expression of Uncertainty in Measuerement. BIPM Paris, France

BIPM (1994) International Vocabulary of Basic and General Terms in Metrology. DIN Berlin/FRG

Brdicka R, Dvorak J (1982) Fundaments of Physical Chemistry. Deutscher Verlag der Wissenschaften, Berlin, Germany

Box MJ (1971) Bias in nonlinear estimation, J Royal Statist Soc B33: 171 – 201

Chalmers RA (1993) Space Age Analysis, Talanta 40: 121–126

Chemval (1992) CHEMVAL Project: Guide to the CHEMVAL Thermodynamic Data Base. UKDOE Report No. DOE/HMIP/RR/92.094

Cox MG, Dainton MP, Harris PM (2001) Software Specifications for Uncertainty Calculation and Associated Statistical Analysis. NPL Report CMSC 10/01 NPL Teddington, UK

Daniele PG, Rigano C, Sammartano S, Zelano V (1994): Ionic strength dependence of formation constants–XVIII, the hydrolysis of iron(III) in aqueous KNO_3 solutions, Talanta 41: 1577–1582

Davies R (1961) Ion Associations. Butterworths London, UK

Debye P, Hückel E (1923) Zur Theorie der Elektrolyte, Phys Z 24: 185-206

DiCiccio TJ, Romano JP (1988) A review of bootstrap confidence intervals (with discussion), J Royal Statist Soc B50: 338-370

ECN (2000) General Requirements for the Competence of Testing and Calibration Laboratories: ISO/IEC 17025, European Committee for Standardization Brussels, Belgium

Efron B (1979) Computers and the theory of statistics: thinking the unthinkable, SIAM Review 21: 460–480

Efron B, Tibshirani R (1991) Statistical data analysis in the computer age, Science 253: 390–395

Ekberg C, Emrén AT (2001) Uncertainties in solubility calculations. Monatsh Chemie 132: 1171 - 1179

Ekberg C, Emrén AT, Samuelson A (1997) The effect of mineral solubility on the solubility of some actinides: an uncertainty analysis, MRS Symp Ser 465: 735–742

Ellison S, Wegscheider W, Williams A (1997) Measurement uncertainty. Anal Chem 69: A607–A613

Ellison SLR, Barwick VJ (1998) Using validation data for ISO measurements uncertainty estimation. Part 1. principles of an approach using cause and effect analysis Analyst 123: 1387-1392

EURACHEM/CITAC (1995) Quantifying Uncertainty in Analytical Measurement, EURACHEM Teddington, UK

EURACHEM/CITAC (1998) Quality Assurance for Research and Development and Non-Routine Analysis EURACHEM Teddington, UK

Guldberg CM, Waage P (1865) Studier over Affiniteten; Forh. Vid. Selsk. Christiana, Christiana (Oslo) Univ p. 35–45

Guldberg CM, Waage P (1867) Etudes sur les Affinités Chimiques, Brögger and Christie, Oslo, Norway

Hartley FR, Burgess C, Alcock RM (1980) Solution Equilibria, Ellis Horwood Pub, Chichester, UK

Hefter GT (1982) Critical Survey of Stability Constants and Related Thermodynamic Data of Fluoride Complexes in Aqueous Solution. IUPAC Pergamon Press, London, UK

IUPAC Subcommittee on pH Measurement (2002) The measurement of pH - definitions, standards and procedures. Report of the working group on pH. Pure Appl Chem, in press

Khoe GH, Brown PL, Sylva RN, Robins RG (1986) The hydrolysis of metal ions, part 9, Iron(III) in perchlorate, nitrate, and chloride media (1 mol dm^{-3}), J Chem Soc Dalton Trans 1901–1906

King B (1997) Metrology and analytical chemistry: bridging the cultural gap, Metrologia 34: 41–47

King B (1997) Traceability of chemical analysis. Analyst 122: 197–204

Königsberger E, Tran-Ho LC (1997) Solubility of substances related to urolithasis–experiment and computer-modeling, Current Top Solution Chem 2: 183–202

Krug RR, Hunter WG, Grieger RA (1976) Enthalpy-entropy compensation. 1. Some fundamental statistical problems associated with the analysis of van't Hoff and Arrhenius data, J Phay Chem 80: 2335–2341

Lowson R (1982) Aqueous oxidation of pyrite by molecular oxygen. Chem Rev 82: 461–497

Marshall SL, May PM, Hefter GT (1995) Least-squares analyses of osmotic coefficient data at 25 °C according to Pitzer's equation. 1:1 electrolytes. J Chem Eng Data 40: 1041-1052

Martell AE, Smith RM, Motekaitis RJ (1993): The NIST Table of Critical Stability Constants, NIST Gaithersburg, USA

May PM, Williams DR, Linder PW, Torrington RG (1982) The use of glass electrodes for the determination of formation constants–I. A definitive method for calibration, Talanta 29: 249-256

May PM, Murray K (1991) JESS–a joint expert speciation system I. Raison d'Etre. Talanta 38: 1409-1417

May PM, Murray K (1991) JESS, a thermodynamic expert speciation system, II: the thermodynamic database, Talanta 38: 1419-1426

May PM (1995) Modeling metal-ligand interaction in blood plasma. in: Handbook of Metal-ligand Interactions in Biological Fluids. 1. Bioinorganic Chemistry (G Berthon ed) 2: 1184-1194

May PM (2000) A simple general and robust function for equilibria in aqueous electrolyte solutions to high ionic strength and temperature. Chem Comm 2000: 1265-1266

May PM, Murray K (2000) http:\\jess.murdoch.edu.au

May PM, Murray K (2001) Database of chemical reactions designed to achieve thermodynamic consistency automatically, J Chem Eng Data, 46: 1035-1040

Meinrath G (2000) Computer-intensive methods for uncertainty estimation in complex situations. Chemometrics Intell Lab Syst 51: 175–187

Meinrath G (2000b) Comparability of thermodynamic data – a metrological point of view, Fresenius J Anal Chem 368: 574–584

Meinrath G (2001) Measurement uncertainty of thermodynamic data. Fresenius J Anal Chem 369: 690-697

Meinrath G, Ekberg C, Landgren A, Liljenzin JO (1999): Assessment of uncertainty in parameter evaluation and prediction. Talanta 51: 231-246

Meinrath G, Hurst S, Gatzweiler R (2000) Aggravation of licensing procedures by doubtful thermodynamic data, Fresenius J Anal Chem 368: 561-566

Meinrath G, Nitzsche O (2000) Impact of measurement uncertainty in chemical quantities on environmental prognosis by geochemical transport modeling, Isotopes Environ Health Stud 36: 195-210

Meinrath G, Spitzer P (2000) Uncertainties in determination of pH, Mikrochim Acta 135: 155-168

Meinrath G., Lis S. (2002) Application of cause-and-effect diagrams to the interpretation of UV-vis spectroscopic data, Anal Bioanal Chem 372: 333-340

Millero FJ, Yao W, Aicher J (1995) The speciation of Fe(II) and Fe(III) in natural waters, Mar Chem 50: 21–39

Mucci A (1983) The solubility of calcite and aragonite in seawater at various salinities, temperatures and one atmosphere total pressure, Am J Sci 283: 780-799

Nitzsche O, Meinrath G, Merkel B (2000) Database uncertainty as a limiting factor in reactive transport prognosis, J Contam Hydrol 44: 223-237

Nordstrom DK, Alpers CN, Ptacek CJ, Blowes DW (2000) Negative pH and extremely acidic mine waters from Iron Mountain, California, Environ Sci Technol 34: 254-258

Okino MS, Mavrovouniotis ML (1998) Simplification of mathematical models of chemical reaction systems, Chem Rev 98: 391-408

Olbricht W, Chatterjee N, Miller K (1994) Bayes estimation: a novel approach to derivation of internally consistent thermodynamic data for minerals, their uncertainty, and correlation. Part I: Theory, Phys Chem Miner 21: 36-49

Palmer W (1987) Über die Partikelgröße und Konzentration des Eisens(III) in Flußwässern. Report KfK-4289, Nuclear Research Center Karlsruhe, Karlsruhe, Germany

Parkhurst DL (1995) User's guide to PHREEQC–a computer program for speciation, reaction-path, advective transport, and inverse geochemical calculations, WRI Report 95-4227, U.S. Geol Survey, Denver, USA

Perkins EH, Gunter WD, Nesbitt MW, St Arnaud Lc (1997) Critical review of acidic drainage from mine waste rock, 4th Int Conf Acidic Rock Drainage, Vancouver, Canada, pp 587-601

Pitzer KS (1991) Ion interaction approach: theory and data correlation. In: Pitzer KS (Ed) Activity Coefficients in Electrolyte Solutions, 2nd edit, CRC Press, Boca Raton FL, USA

Posner AM (1953) Trans Faraday Soc 49: 389

Richter BW (1999) Comparability and recognition of chemical measurement results–an international goal. Fresenius J Anal Chem 365: 569–573

Sigel H, Zuberbühler A, Yamauchi O (1991) Comments on potentiometric pH titrations and the relation between pH-meter reading and hydrogen ion concentration, Anal Chim Acta 255: 63–72

Singer PC, Stumm W (1970) Acid mine drainage: the rate determining step, Science 167: 1121-1123

Spendley R (1969) Nonlinear least squares fitting using using a modified SIMPLEX minimisation method, in: Fletcher R (ed) Optimization, Inst of Mathematics and its Applications, London,UK p. 259

Thompson M (1997) Statistics–the curse of the analytical classes. Analyst 119: 127N

Wolf G, Lerchner H, Schmidt H, Gamsjäger H, Königsberger E, Schmidt P (1996) Thermodynamics of $CaCO_3$ phase transitions, J Therm Anal 46: 353-359