

Computing Multi-Species Chemical Equilibrium with an Algorithm Based on the Reaction Extents

Juan Manuel Paz-García^{a,*}, Björn Johannesson^a, Lisbeth M. Ottosen^a, Alexandra B. Ribeiro^b,
José Miguel Rodríguez-Maroto^c

^aDepartment of civil Engineering, Technical University of Denmark, Brovej, Building 118, Dk 2800 Kgs. Lyngby, Denmark.

^bDepartment of Environmental Sciences and Engineering, Faculty of Sciences and Technology, New University of Lisbon, Caparica, Portugal

^cDepartment of Chemical Engineering, Faculty of Sciences, University of Malaga, Campus de Teatinos, Malaga, Spain.

Abstract

A mathematical model for the solution of a set of chemical equilibrium equations in a multi-species and multiphase chemical system is described. The computer-aid solution of model is achieved by means of a Newton-Raphson method enhanced with a line-search scheme, which deals with the non-negative constraints. The residual function, representing the distance to the equilibrium, is defined from the chemical potential (or Gibbs energy) of the chemical system. Local minimums are potentially avoided by the prioritization of the aqueous reactions with respect to the heterogeneous reactions. The formation and release of gas bubbles is taken into account in the model, limiting the concentration of volatile aqueous species to a maximum value, given by the gas solubility constant.

The reaction extents are used as state variables for the numerical method. As a result, the accepted solution satisfies the charge and mass balance equations and the model is fully compatible with general reactive transport models.

Keywords: Chemical equilibrium, Speciation, Reaction extent, Reactive transport

1. Introduction

In a system at chemical equilibrium, the reactions take place at equal rates in their direct and reverse directions and, therefore, the concentrations of the reacting substances (reactants and products) do not change with time. The chemical equilibrium assumption (CEA) consists of considering that all the species in the system have time enough to reach the equilibrium state. As a result, the time derivative of the chemical equations becomes zero and a set of algebraic chemical concentration equations can be used to mathematically describe such a system. Despite that, the equations describing multi-species chemical equilibrium (MSCE) systems are strongly nonlinear and the analytical solution cannot be found except for the most simple cases, being necessary the use of computer-aid approaches.

Numerical models based on the CEA may be used, for example, to: (1) fit experimental data to chemical equilibrium state in a numerical speciation process, (2) to estimate trace compounds concentrations based on the analytical measurements of species with significantly higher concentrations, (3) to study the response of a chemical system in equilibrium with respect to external changes, (4) to study weathering mechanisms in geochemical systems, and (5) to model chemical interactions between the pore

solution and the solid matrix in porous materials (Bethke, 2008).

The reliability of the CEA depends on the time scale for the speciation problem. A representative example would be a reactive transport model under the local CEA in which the driving forces are chemical or electrical potential gradients. In those cases, the transport rates of the main transport mechanisms (diffusion, electromigration and electroosmosis) are slow enough to accept the equilibrium assumption except for those very slow heterogeneous reactions (Jacobs & Probst, 1996; Morel & Hering, 1993).

Aqueous reversible reactions frequently have high kinetic rates in both directions, towards the products and towards the reactants, making the CEA acceptable in most cases. Acids and bases dissociation or the self-ionization of water, for example, can reach the equilibrium state in the order of microseconds or even faster. Aqueous complexation reactions are also usually quite fast, normally achieving the equilibrium in the range from micro- to milliseconds. In the case of heterogeneous reactions, such as adsorption/desorption, ion-exchange or precipitation/dissolution reactions, their kinetic rates are commonly slower than the above mentioned aqueous homogeneous reactions and their rates may compete with the transport rates in some cases.

Some models for reactive transport through porous media taking into consideration the kinetics of the chemical system have been presented, as e.g. (Johannesson, 2009; Lichtner, 1985; Steefel & Van Cappellen, 1990). Neverthe-

*Corresponding author

Email address: jugra@byg.dtu.dk (Juan Manuel Paz-García)

Nomenclature

List of symbols

a	chemical activity (-)
\mathbf{a}	chemical activities vector (-)
A	Davies constant (kg mol^{-1}) ^{1/2}
b	Setschenow constant
f	residual function
\mathbf{f}	residual function vector
ΔG	Gibbs energy
Δh	numerical increment
I	Ionic strength
\mathbf{J}	Jacobian matrix
K	equilibrium constant
K_H	Henry's constant
\mathbf{k}	equilibrium constant matrix
m	molal concentration (mol kg^{-1})
M	number of reactions (-)
\mathbf{M}	stoichiometric matrix (-)
n	molal amount (mol)
\mathbf{n}	molal amounts vector (mol)
N	number of species (-)
P	pressure (atm)
R	universal gas constant ($\text{J K}^{-1} \text{mol}^{-1}$)
T	absolute temperature (K)
x	reaction extent (mol)
\mathbf{x}	reaction extents vector (mol)
z	ionic charge (-)
Greek letters	
γ	activity factor (-)
λ	line-search scaling factor (-)
μ	chemical potential (J mol^{-1})
ν	stoichiometric coefficient (-)
Subscripts	
eq	equilibrium
i	chemical species
init	initial
max	maximal
min	minimal
r	equilibrium reaction

less, models under the chemical equilibrium assumption are more extensively used, as e.g. (Al-Hamdan & Reddy, 2008; Javadi & Al-Najjar, 2007; Rubin, 1983; Vereda-Alonso et al., 2007). Rate-controlled equilibrium models, which are considered more realistic, include a set of feasible reactions capable to reach the equilibrium in the time scale of the transport process and a parallel solution of the kinetic equations for the slow and/or the irreversible reactions (Koukkari & Pajarre, 2006; Ribeiro et al., 2005; Steyer et al., 2005).

The most accepted mathematical algorithm for MSCE calculation is based on the Morel's Tableau stoichiometric method (Morel & Hering, 1993; Morel & Morgan, 1972), which basically consists in dividing the set of chemical species in a number of master species (or components) and secondary species (or compounds). Secondary species are

defined as a combination of master species by means of stoichiometric chemical equilibrium equations. In (Reed, 1982) and (Bethke, 2008), comprehensive compilations of different proposed models for the mathematical solution of chemical equilibrium problems can be found. This literature review will not be repeated here.

Several computer programs for MSCE have been released. Some of the most used are PHREEQC (Parkhurst & Appelo, 1999), PHREEQC*i* (Charlton et al., 1997), WATEQ4F (Ball & Nordstrom, 1991), MINTEQ (Peterson et al., 1987), EQ3/6 (Wolery, 1992) and GEMS (Remy, 2004). Coupling these kind of stand-alone programs with more general codes for different purposes is possible. Nevertheless, exporting and importing data between different programs normally involves the participation of the program user, as well as significant computational time, making simulations slow or, sometimes, unfeasible. Consequently, the implementation of tailor-made codes is almost always necessary for more complex problems (Rodriguez-Maroto & Vereda-Alonso, 2009).

In summary, numerical strategies for computing chemical equilibrium problems are classified into two main groups: stoichiometric and non-stoichiometric algorithms (van Baten & Szczepanski, 2011; Blomberg & Koukkari, 2011; Brassard & Bodurtha, 2000; Carrayrou et al., 2002). Stoichiometric algorithms converge on the solution of a set of simultaneous mass balance and mass action equations at each iteration, while non-stoichiometric algorithms aim a direct minimization of the Gibbs energy functions of the chemical species constrained by mass balances.

The most popular stoichiometric algorithm uses the concentration of the master species as the independent state variable while the residual function to be minimized in the numerical procedure is defined from the mass balance on the chosen master species and the saturation index of the existing solid components. The mass balance of protons is replaced by the charge balance equation, in this approach, for two reasons: (1) to avoid considering the water as a chemical species in the setup, and (2) to assure charge neutrality in the equilibrated solution. This method has shown accurate results and has been the base for some of the chemical equilibrium models listed above.

In this work, we propose a stoichiometric method for MSCE, based on the Tableau's concept, iterating on the reaction extents and submitted to the non-negative composition constraints. Despite this method shows a higher tendency of falling into the so-called local minima, some advantages may be remarked: (1) A model based on the reaction extent is compatible with rate-controlled models dealing simultaneously with chemical equilibrium and chemical kinetics, and (2) for a given value of the reaction extent, the mass and the charge conservation principles are always satisfied increasing the compatibility with general reactive-transport models.

A Newton-Raphson (NR) method, enhanced with a line-search (LS) with backtracking scheme, is used for the solution of the nonlinear system of equations (Press et al.,

1992; Wolery & Walters, 1975). The model includes heterogeneous precipitation/dissolution reactions, as well as the possibility of the release of bubbles of volatile aqueous species to the surrounding atmosphere. Logarithm scale is used in order to (1) convert the exponential mass action equation into algebraical equations, (2) assure the non-negative constrain, and (3) to minimize the error induced by the programming language when using values with significant differences in their orders of magnitude, as common in multi-species chemical systems.

2. Mathematical model

2.1. Stoichiometry of the chemical system

The analytical description of the chemical equilibrium problem is carried out by a set of nonlinear algebraic equations formulated from the mass balance and the action mass equations describing the chemical system (Bethke, 2008; Rawlings & Ekerdt, 2002; Stumm & Morgan, 1970).

The total number of species, N , is divided into the M secondary species and the $(N - M)$ master species that define them. The chosen master species must represent all the chemical elements in the system (Rodriguez-Maroto & Vereda-Alonso, 2009).

For a set of N chemical species, the M stoichiometric equations are defined as:

$$\sum_{i=1}^N \nu_{r,i} n_i = 0 \quad , \quad r = 1, 2, \dots, M \quad (1)$$

where n_i (mol) denotes the total amount of the i^{th} chemical species in the system and $\nu_{r,i}$ is the stoichiometric coefficient for the i^{th} species in the r^{th} reaction. Herein, the stoichiometric equations are written in the form of dissociation or dissolution reactions, being the stoichiometric coefficients of the reactants defined with negative values. By definition, each secondary species participate only in the r^{th} stoichiometric equation that describes it.

In the following, we will use the chemical system resulting from the dissolution of CaCO_3 (s), CaCl_2 (s) and $\text{Ca}(\text{OH})_2$ (s) in HCl solution, summarized in Table 1 as an example to explain the MSCE model and for the test examples simulations shown in Section 4. Despite the relative simplicity of this chemical system, it is complete enough to illustrate the mathematics and features of the proposed model. It involves both gaseous and solids phases.

The species CO_2^* is used to represent the combination of carbonic acid H_2CO_3 (aq) and free carbon dioxide CO_2 (aq). In aqueous solution carbonic acid exists in equilibrium with carbon dioxide and the concentration of H_2CO_3 is much lower than the concentration of CO_2 . It is very difficult to analytically distinguish both species and individual equilibrium constant are not reliable. Consequently, the equilibrium constant used here a the combination of both species into a single one, with the stoi-

Table 1: Chemical system resulting from the dissolution of CaCO_3 (s), CaCl_2 (s) and $\text{Ca}(\text{OH})_2$ (s) in HCl solutions. Equilibrium constants for the given stoichiometry when using molal concentrations are obtained from (Parkhurst & Appelo, 1999).

Reaction	$\log(K_{\text{eq}})$
$\text{H}_2\text{O} \rightleftharpoons \text{H}^+ + \text{OH}^-$	-14.00
$\text{CaOH}^+ \rightleftharpoons \text{Ca}^{2+} + \text{H}^+$	-1.22
$\text{CaCl}^+ \rightleftharpoons \text{Ca}^{2+} + \text{Cl}^-$	0.29
$\text{CaCl}_2 \rightleftharpoons \text{Ca}^{2+} + 2\text{Cl}^-$	0.64
$\text{HCl} \rightleftharpoons \text{H}^+ + \text{Cl}^-$	0.71
$\text{CaCO}_3 \rightleftharpoons \text{Ca}^{2+} + \text{CO}_3^{2-}$	-3.22
$\text{CO}_2^* + \text{H}_2\text{O} \rightleftharpoons 2\text{H}^+ + \text{CO}_3^{2-}$	-16.68
$\text{HCO}_3^- \rightleftharpoons \text{H}^+ + \text{CO}_3^{2-}$	-10.33
$\text{CaHCO}_3^+ \rightleftharpoons \text{Ca}^{2+} + \text{H}^+ + \text{CO}_3^{2-}$	-11.43
CaCO_3 (s) $\rightleftharpoons \text{Ca}^{2+} + \text{CO}_3^{2-}$	-8.48
CaCl_2 (s) $\rightleftharpoons \text{Ca}^{2+} + 2\text{Cl}^-$	11.77
$\text{Ca}(\text{OH})_2$ (s) $\rightleftharpoons 2\text{Ca}^{2+} + 2\text{OH}^-$	-5.19

chiometry of the carbonic acid (Greenwood & Earnshaw, 1984).

A matrix of stoichiometric coefficients (denoted as stoichiometric matrix, \mathbf{M}), with M rows and N columns is formed from the set of stoichiometric equations describing the system, as shown in Fig. 1.

2.2. Extent of the reaction

When N species react with each other to reach the chemical equilibrium state, a conservation of mass equation, function of the extent of the reaction x_r , can be used to describe the chemical transformation. The reaction extent (or progress) is the extensive quantity describing the progress of a chemical reaction equal to the number of chemical transformations (McNaught & Wilkinson, 1997). In order to reach the chemical equilibrium state, any reversible reaction can proceed toward the products or toward the reactants, meaning that x_r can be either a positive or a negative value.

The total amount of each species at the equilibrium is given by the mass balance equation along the chemical reaction path, in the form of:

$$n_{i,\text{eq.}} = n_{i,\text{init.}} + \sum_{r=1}^M \nu_{r,i} x_r \quad ; \quad i = 1, 2, \dots, N \quad (2)$$

In vector/matrix notation, Eq. (2) is:

$$\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_N \end{bmatrix}_{\text{eq.}} = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_N \end{bmatrix}_{\text{init.}} + \mathbf{M}^T \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} \quad (3)$$

or

$$\mathbf{n}_{\text{eq.}} = \mathbf{n}_{\text{init.}} + \mathbf{M}^T \mathbf{x} \quad (4)$$

	H ⁺	OH ⁻	Ca ²⁺	CO ₃ ²⁻	Cl ⁻												
H ₂ O	1	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0
CaOH ⁺	0	1	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
CaCl ⁺	0	0	1	0	1	0	0	-1	0	0	0	0	0	0	0	0	0
CaCl ₂	0	0	1	0	2	0	0	0	-1	0	0	0	0	0	0	0	0
HCl	1	0	0	0	1	0	0	0	0	-1	0	0	0	0	0	0	0
CaCO ₃	0	0	1	1	0	0	0	0	0	0	-1	0	0	0	0	0	0
CO ₂ * + H ₂ O	2	0	0	1	0	0	0	0	0	0	0	-1	0	0	0	0	0
HCO ₃ ⁻	1	0	0	1	0	0	0	0	0	0	0	-1	0	0	0	0	0
CaHCO ₃ ⁺	1	0	1	1	0	0	0	0	0	0	0	0	-1	0	0	0	0
Ca(OH) ₂ (s)	0	2	1	0	0	0	0	0	0	0	0	0	0	-1	0	0	0
CaCl ₂ (s)	0	0	1	0	2	0	0	0	0	0	0	0	0	0	-1	0	0
CaCO ₃ (s)	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	-1

Figure 1: Stoichiometric matrix, \mathbf{M} , for the modeled chemical system.

The vector \mathbf{x} is used as the independent state variable in the numerical method.

2.3. Mass action equation

The chemical potential, μ_i (J mol⁻¹), is defined as the Gibbs energy of the species i , per mole unit, at a given temperature and pressure, as:

$$\mu_i = \mu_i^\ominus - RT \ln a_i \quad (5)$$

where a_i (dimensionless) is the chemical activity, μ_i^\ominus is the standard chemical potential, R (≈ 8.314 J K⁻¹ mol⁻¹) is the universal gas constant and T (K) is the absolute temperature. The chemical potential for the r^{th} reaction, μ_r , also referred as partial molar Gibbs energy, is obtained from Eq. (6).

$$\mu_r = \sum_{i=1}^N \nu_{r,i} \mu_i^\ominus + RT \sum_{i=1}^N \nu_{r,i} \ln a_i \quad (6)$$

At the equilibrium, the chemical potential or the reaction is zero. Defining the Standard Gibbs energy and the equilibrium constant for the r^{th} reaction, respectively as:

$$\Delta G_r^\ominus = \sum_{i=1}^N \nu_{r,i} \mu_i^\ominus \quad (7)$$

$$K_r = \exp\left(-\frac{\Delta G_r^\ominus}{RT}\right) \quad (8)$$

the mass action equation equilibrium equation, Eq. (9), is easily obtained:

$$\sum_{i=1}^N \nu_{r,i} \ln a_i - \ln K_r = 0 \quad (9)$$

When the concentration of species in the system is not at chemical equilibrium state ($a_i \neq a_{i,\text{eq}}$), Eq. (9) will be:

$$f_r(\mathbf{x}) = \sum_{i=1}^N (\nu_{r,i} \ln a_i) - \ln K_r \quad (10)$$

where $f_r(\mathbf{x})$ is the residual function that tends to zero when the system approaches to the equilibrium, as $f_r(\mathbf{x}) = \mu_r(\mathbf{x})/RT$. Thus, the function $f_r(\mathbf{x})$ represents the distance to the equilibrium state for the r^{th} reversible reaction.

Let \mathbf{a} and \mathbf{k} be the vectors resulting from the natural logarithm of the activity and equilibrium constants of the M chemical reactions in the system. For the numerical procedure both decimal and natural logarithm can be used. Nevertheless, even though the decimal logarithm is a more intuitive choice when defining the equilibrium constants, the natural logarithm is computed faster in most programming languages. For this reason, equilibrium constants are converted to the natural logarithm scale for the proposed model.

Eq. (10) in vector/matrix notation is:

$$\begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_M(\mathbf{x}) \end{bmatrix} = \mathbf{M} \begin{bmatrix} \ln a_1(\mathbf{x}) \\ \ln a_2(\mathbf{x}) \\ \vdots \\ \ln a_N(\mathbf{x}) \end{bmatrix} - \begin{bmatrix} \ln K_1 \\ \ln K_2 \\ \vdots \\ \ln K_M \end{bmatrix} \quad (11)$$

or

$$\mathbf{f}(\mathbf{x}) = \mathbf{M} \mathbf{a}(\mathbf{x}) - \mathbf{k} \quad (12)$$

The vector-function $\mathbf{f}(\mathbf{x})$ compiles the value of $f_r(\mathbf{x})$ for the complete set of M reactions. The goal of the numerical method is to obtain a vector \mathbf{x} that drives the system to the equilibrium, e. i. $\mathbf{f}(\mathbf{x}) \approx 0$. The root finding procedure is addressed by the minimization of the norm of the vector $\mathbf{f}(\mathbf{x})$, which is equivalent to a least squares technique.

2.4. Chemical activities

Different approaches for the calculation of the chemical activities can be used when implementing the method. In the simulations presented here, the Davies and Setschenow equations, described later in this section, have been used for ionic and non-ionic aqueous species respectively. The activity of the solvent and solid species is considered 1.

In the present model, the activity values of each aqueous chemical species are obtained from the amount $n_i(\mathbf{x})$ calculated in each numerical iteration step. The activity is defined from the molal concentration m_i (mol kg^{-1}), the activity factor γ_i (dimensionless), and the standard concentration m_i^\ominus (mol kg^{-1}) necessary to ensure that the activity is also dimensionless.

$$a_i = \gamma_i \frac{m_i}{m_i^\ominus} \quad (13)$$

The Davies equation, Eq. (14), is an empirical extension of the Debye-Hückel equation. The Davies equation is considered to be a reasonable approximation even at relatively high ionic concentrations.

$$\ln \gamma_i = Az_i^2 \left(\frac{\sqrt{I}}{1 + \sqrt{I}} - 0.3I \right) \quad (14)$$

where $A = -1.172$ (kg mol^{-1})^{1/2}, z_i (mol mol^{-1}) is the ionic charge and I (mol kg^{-1}) is the ionic strength of the electrolyte media, calculate by:

$$I = \frac{1}{2} \sum_{i=1}^N m_i z_i^2 \quad (15)$$

For the case of non-ionic aqueous species, the activity factor is obtained using the Setschenow relation, as

$$\ln \gamma_i = b_i I \quad (16)$$

where b_i is the Setschenow coefficient. In the simulations presented in this work, $b_i = 0.1$ is used for all non-charged aqueous species, as done in (Parkhurst & Appelo, 1999).

The validity of the Davies and Setschenow equations is limited to values of ionic strength equal to or lower than 0.5. Over this limit, spurious results may be obtained. Some alternative theories, such as the Specific Ion Interaction Theory (SIT) (Guggenheim & Turgeon, 1955) and the Pitzer activity coefficients (Pitzer, 1973) could be used to increase the range of validity of the activity coefficients.

2.5. Precipitation/dissolution heterogeneous equilibrium

Heterogeneous precipitation and dissolution reactions have some key differences with respect to the aqueous complexation. First, the activity of the solids is set to the unity value. Consequently, in the mass action equation for precipitation/dissolution reactions, the excess of solids in saturated solutions does not affect the concentration of aqueous species.

The existence of solid species is limited by the so-called saturation index (SI), which determines if the electrolyte solution is saturated or undersaturated with respect to the solid defined by a precipitation/dissolution reaction. Indeed, the equilibrium constant for a precipitation/dissolution reaction is typically denoted as solubility

product. The value of the residual function $f_r(\mathbf{x})$ is used as an indicator of the saturation index, in the form:

$$\begin{cases} f_r(\mathbf{x}) < 0 & \text{undersaturated} \\ f_r(\mathbf{x}) \geq 0 & \text{saturated} \end{cases} \quad (17)$$

If a solution is undersaturated, the corresponding solid would be completely dissolved and it will not participate in the equilibrium process. Therefore, its contribution in the residual function must be ignored. The numerical procedure, for any iteration, consists of:

1. Identify the precipitation/dissolution reactions that are undersaturated, i.e. $f_r(\mathbf{x}) < 0$
2. Force the reaction extent for undersaturated reactions to be the maximum (complete dissolution).
3. Ignore the contribution of $f_r(\mathbf{x}) < 0$ to the calculation of the norm of the residual.

In conditions close to the equilibrium state, the property $f_r(\mathbf{x})$ can oscillate around the value of zero, and so the reaction can oscillate around the saturated/undersaturated status during the speciation process. Consequently, this strategy has to be followed iteration-by-iteration.

Furthermore, using the density and the molecular mass, the volume fraction occupied by the solution and the solids can be measured. As a result, the porosity of the porous media can be calculated, which is a crucial parameter in the case of reactive-transport modeling through these porous media.

2.6. Vapor-liquid heterogeneous reactions

Vapor-liquid heterogeneous reactions are difficult to couple with the previously described aqueous complexation and mineral dissolution reactions, because many extra assumptions are necessary. If the time scale is large enough to assume the equilibrium condition for the vapor-liquid equilibrium, and in the case of very dilute systems, the Henry's law may be used. Henry's law states that, at a constant temperature, the amount of a given gas dissolved in a volume of liquid is directly proportional to the partial pressure of that gas in equilibrium with that liquid, by means of the so-called Henry's constant.

$$P_{i(g)} = K_H m_{i(aq)} \quad (18)$$

This kind of approach is used to calculate the concentration of gaseous species in equilibrium with the liquid system in many geochemical speciation models, as e.g. in (Parkhurst & Appelo, 1999).

For the case of chemical equilibrium models designed to be coupled with reactive-transport in porous media, some extra considerations have to be taken. For example, depending on the time scale of the speciation problem, assuming vapor-liquid equilibrium between the atmospheric species and the aqueous species can lead to unrealistic results. This is the case of the carbonic acid equilibrium. If we consider a porous material in contact

with the atmosphere, being the atmospheric partial pressure of the carbon dioxide a constant value, the concentration of the species $\text{CO}_2(\text{aq})$, HCO_3^- , CO_3^{2-} among others, would be fixed. Consequently, the system would be strongly buffered by the carbonic acid equilibrium. Rate constants for the reversible reaction between the gaseous and aqueous CO_2 equilibrium is slow with respect to the aqueous complexation reactions. Thus, the buffered capacity of the atmospheric CO_2 in the porous system would be overestimated if the equilibrium assumption is accepted.

According to this, a different approach is taken in this model. First, the system is assumed to be at constant pressure (normally 1 atm). The Henry's constant is used to define the maximum amount of the volatile species aqueous concentration that would produce the release of gas bubbles from the system. For example, the concentration of gaseous carbonic acid would be given by:

$$P_{\text{CO}_2(\text{g})} = K_{\text{H}} m_{\text{CO}_2(\text{aq})} \quad (19)$$

Mathematically, this assumption is equivalent to consider a maximum limit value for the aqueous concentration of volatile compounds. For the case of $\text{CO}_2(\text{g})$ at room temperature and $P_{\text{CO}_2} = P_{\text{total}} = 1$ (atm), the maximum $\text{CO}_2(\text{aq})$ concentration would be given by:

$$m_{\text{CO}_2(\text{aq})}|_{\text{max}} = \frac{1}{K_{\text{H}}} = 3.36 \times 10^{-2} \text{ (mol kg}^{-1}\text{)} \quad (20)$$

3. Numerical implementation

3.1. Newton-Raphson method for non-linear systems

The value of the reaction extents vector \mathbf{x} that assures global equilibrium is obtained by an iterative procedure based on a NR method with a line-search technique to satisfy the non-negative restriction (Press et al., 1992). The numerical procedure is summarized in the pseudocode shown in Fig. (2).

The NR method for a non-linear matrix system of equations indicates the next iteration value for the unknowns, \mathbf{x}_{new} , obtained from the present value, \mathbf{x}_{old} , by adopting a numerical increment, $\delta\mathbf{x}$, toward the direction that decreases the global residual function.

$$\mathbf{x}_{\text{new}} = \mathbf{x}_{\text{old}} + \delta\mathbf{x} \quad (21)$$

The Taylor's series expansion of the residual function $\mathbf{f}(\mathbf{x})$ is used.

$$\mathbf{f}(\mathbf{x}_{\text{new}}) = \mathbf{f}(\mathbf{x}_{\text{old}}) + \mathbf{J} \delta\mathbf{x} + \mathbf{O}(\delta\mathbf{x})^2 \quad (22)$$

where $\mathbf{O}(\delta\mathbf{x})^2$ represents the error related to terms with order greater than 2 and \mathbf{J} is the Jacobian matrix of partial derivatives defined as:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_M} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial x_1} & \cdots & \frac{\partial f_M}{\partial x_M} \end{bmatrix} \quad (23)$$

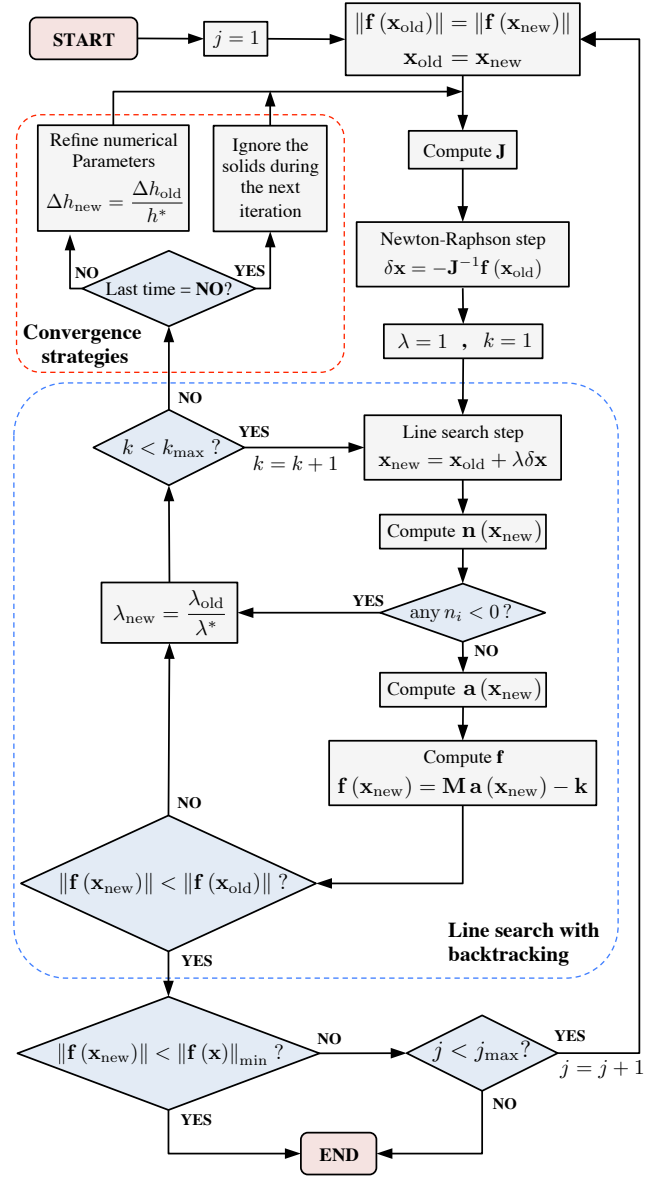


Figure 2: Flowchart for the NR method

The next step in the NR method adopted will be given from the fact that the target is to obtain $\mathbf{f}(\mathbf{x}_{\text{new}}) = \mathbf{0}$. Ignoring the term $\mathbf{O}(\delta\mathbf{x})^2$, the increment in the extent of all considered reactions is, therefore, obtained as:

$$\delta\mathbf{x} = -\mathbf{J}^{-1}\mathbf{f}(\mathbf{x}_{\text{old}}) \quad (24)$$

3.2. Line-search with backtracking enhancement

A line-search with backtracking enhancement is used in the NR method. Line-search iterative methods usually require an important computational effort due to a large number of iterations until convergence. As mentioned before the vector $\delta\mathbf{x}$ indicates the direction in which the residual decreases but it carries no information about the magnitude of the increment in the extent of the reactions.

If the step is too long, the system may get far from the equilibrium, failing in minimizing the residual. In addition to this, the system is restricted to the non-negative physical constraint for the concentrations or amounts of all chemical species (van Baten & Szczepanski, 2011).

In the LS method, a scalar factor λ is used to control the magnitude of the increment calculated by the NR method. Eq. (25) is used instead of Eq. (21).

$$\mathbf{x}_{\text{new}} = \mathbf{x}_{\text{old}} + \lambda \delta \mathbf{x} \quad (25)$$

The optimal λ value is obtained by a try-and-error algorithm, following a decreasing sequence given by the geometric progression according to Eq. (26), with an initial value of $\lambda_{\text{init}} = 1$.

$$\lambda_{\text{new}} = \lambda_{\text{old}} / \lambda^* \quad (26)$$

3.3. Numerical computation of the Jacobian

Due to the strong non-linearity of the system of equations, the analytical calculation of the Jacobian matrix is unfeasible. Therefore, an incomplete Newton-Raphson scheme is used, based on a numerical computation of the Jacobian matrix, achieved by central differences vector differentiation, as:

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_r} = \frac{\mathbf{f}(\mathbf{x} + \Delta x_r) - \mathbf{f}(\mathbf{x} - \Delta x_r)}{2\Delta x_r} \quad (27)$$

The increment Δx_r used in the calculation of the Jacobian plays an important role in the convergence capability of the problem. The reaction extent of the different equilibrium equations may differ significantly, and therefore a constant Δx_r would find unwanted local minimums ruining the iterative approach. In order to avoid this problem, the value Δx_r used is scaled to the value of the vector \mathbf{x} during any iteration:

$$\Delta x_r = |x_r| \Delta h \quad (28)$$

where Δh is a value between Δh_{init} and Δh_{min} .

Incomplete Newton-Raphson schemes have linear convergence rates, while analytical calculated Jacobian would converge quadratically. In addition to this, the numerical calculation of the Jacobian matrix usually requires a significant amount of computational time. In Algorithm (1) we propose a vectorized implementation for computing the numerical Jacobian. In this context, vectorized implementation means that no loops are used for the calculation of the Jacobian, using only matrix algebra operations. Therefore, instead of using a loop for the calculation of the M vectors resulting from the solution of Eq. (27) for the M derivatives, the problem is solved using a single matrix operation. Eq. (4) in matrix notation would be:

$$\begin{bmatrix} \mathbf{n}_1 & \cdots & \mathbf{n}_M \end{bmatrix}_{\text{eq.}} = \begin{bmatrix} \mathbf{n}_1 & \cdots & \mathbf{n}_M \end{bmatrix}_{\text{init.}} + \mathbf{M}^T \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_M \end{bmatrix}$$

or

$$\mathbf{N}_{\text{eq.}} = \mathbf{N}_{\text{init.}} + \mathbf{M}^T \mathbf{X} \quad (29)$$

As the Jacobian matrix is computed several times during the process, using the proposed vectorized algorithm, the entire method is between 10 to 100 times faster than compared to using computational “for” or “while” loops.

3.4. Convergence strategies

The main disadvantage of using the reaction extent as the state variable is that there is a moderate to high risk of falling into a local (relative) minimum, as illustrated in Fig 3. The probability of finding local minimums depends on the distance to the equilibrium from the initial set of concentrations. A good initial estimation of the state variable \mathbf{x} is needed (Brassard & Bodurtha, 2000).

In dynamic problems, such as reactive transport processes, if small enough time increments are used, the distance to the equilibrium is small and an initial estimation of $\mathbf{x} = \mathbf{0}$ is usually valid. When using the model for the speciation of compounds in trace concentration, an initial guess can be estimated based on the concentration of the major species. If there is no information about the equilibrium concentrations at the beginning of the calculation, an initial guess is suggested based on the vector \mathbf{k} , with the equilibrium constants at logarithm scale, which would indicate the tendency of the species toward the dissolution or the dissociation.

Even when using a good initial guess, it may happen that the NR with LS method fails at finding the equilibrium in some specific cases. In order to reduce the risk of finding local minima, some extra considerations have been taken. In this context, the “convergence strategies” block shown in Fig. 2, includes two different procedures, which are alternatively used when the LS method reaches the maximum value of iterations without reducing the residual function.

1. Refine numerical parameters

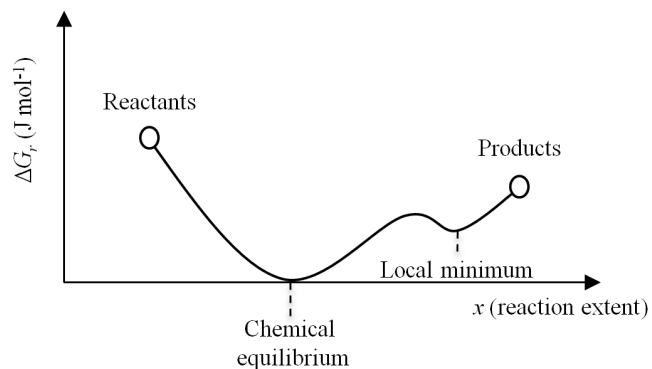


Figure 3: Chemical equilibrium as a function of the reaction extent

Algorithm 1 Pseudo-code for the numerical calculation of the Jacobian matrix

1. Create a matrix $\mathbf{X}^{+\Delta\mathbf{X}}$ (size M by M) including the current value of the state variable \mathbf{x} summed to a diagonal matrix with the finite increment for the numerical calculation of the derivatives:

$$\mathbf{X}^{+\Delta\mathbf{X}} = \mathbf{X} + \text{diag}(\Delta\mathbf{X})$$

where

$$\mathbf{X}^{+\Delta\mathbf{X}} = \begin{bmatrix} (x_1 + \Delta x_1) & \cdots & x_1 \\ \vdots & \ddots & \vdots \\ x_M & \cdots & (x_M + \Delta x_M) \end{bmatrix}$$

2. Compute a matrix \mathbf{N} with the molar amount of species, as a function of $\mathbf{X}^{+\Delta\mathbf{X}}$:

$$\mathbf{N}_{\text{eq.}} = \mathbf{N}_{\text{init.}} + \mathbf{M}^T \mathbf{X}^{+\Delta\mathbf{X}}$$

3. Compute the matrix \mathbf{A} formed from the activities, using the matrix $\mathbf{N}_{\text{eq.}}$ and the same agreements for the activity factors.
4. Using $\mathbf{K} = [\mathbf{k} \ \cdots \ \mathbf{k}]$, compute the residual matrix function, $\mathbf{F}(\mathbf{X}^{+\Delta\mathbf{X}})$ as:

$$\mathbf{F}(\mathbf{X}^{+\Delta\mathbf{X}}) = \mathbf{M} \mathbf{A} - \mathbf{K}$$

5. Repeat steps 1-4 using $\mathbf{X}^{-\Delta\mathbf{X}} = \mathbf{X} - \text{diag}(\Delta\mathbf{X})$
6. Compute the Jacobian matrix as:

$$\mathbf{J} = \frac{\mathbf{F}(\mathbf{X}^{+\Delta\mathbf{X}}) - \mathbf{F}(\mathbf{X}^{-\Delta\mathbf{X}})}{2\Delta\mathbf{X}}$$

where the latter operation is done “term-by-term” instead of by standard matrix division, using the full matrix of increments in the form:

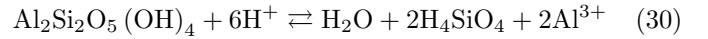
$$\Delta\mathbf{X} = \begin{bmatrix} \Delta x_1 & \cdots & \Delta x_M \\ \vdots & \ddots & \vdots \\ \Delta x_1 & \cdots & \Delta x_M \end{bmatrix}$$

If the norm of the residual function is not minimized for a certain number of iterations k_{max} , the incremental parameter for the numerical calculation of the Jacobian is reduced ten times, i.e. $\Delta h_{\text{new}} = \Delta h_{\text{old}}/h^*$. If the value Δh_{min} is reached and, even though, the residual $\|\mathbf{f}(\mathbf{x})\|$ is not reduced, the local minimum is accepted as a valid result.

2. *Ignore precipitation/dissolution reactions*

The second procedure consists of ignoring the precipitation/dissolution reactions during one iteration. This innovative method is based on the concept that, even when the equilibrium assumption is accepted, the kinetic of the heterogeneous reactions is almost always lower than the homogeneous aqueous reactions.

As an example of this kind of behavior, Eq. (30) shows the dissolution reaction of kaolinite as a function of the master species H^+ , H_2O , H_4SiO_4 and Al^{3+} . According to this stoichiometry and the Le Châtelier’s principle, an external increment in the pH in the system containing kaolinite should produce the precipitation of the mineral, what is characterized with a release of protons which will counteract the pH change. On the other hand, the pH also affects to the species H_4SiO_4 and Al^{3+} . H_4SiO_4 is an acid and can dissociate releasing protons, and Al^{3+} reacts with hydroxides to form different aluminum hydroxides. Consequently, the kaolin is forced to precipitate, due to the pH increase, but it is also forced to dissolve in order to counteract the consumption of H_4SiO_4 and Al^{3+} , due to the same pH change. This kind of situations can, in some instances, lead to the development of a local minimum.



If one takes into account the kinetic rates of the reaction, it can be assured that the aqueous complexation reactions of H_4SiO_4 and Al^{3+} are faster than the precipitation/dissolution reaction. So, these two aqueous species would react first in order to counteract the pH changes and so the kaolinite would be forced to the dissolution. This is congruent with the experimental results observed in (Carroll & Walther, 1990).

4. Test examples

A test example is shown based on the chemical system described in Table 1. In this test example, the concentration in equilibrium of calcite, CaCO_3 , and portlandite, $\text{Ca}(\text{OH})_2$, in 1 kg of HCl solution at different concentrations is computed. For the simulations, the minerals are assumed in equilibrium with 1 kg of pure water. HCl is progressively added to the solution, to cover a range from 0 until 0.6 mol per kg of water. The equilibrium concentration of aqueous species, the amount of precipitated solids and the pH are shown with respect to the amount of acid added to the system.

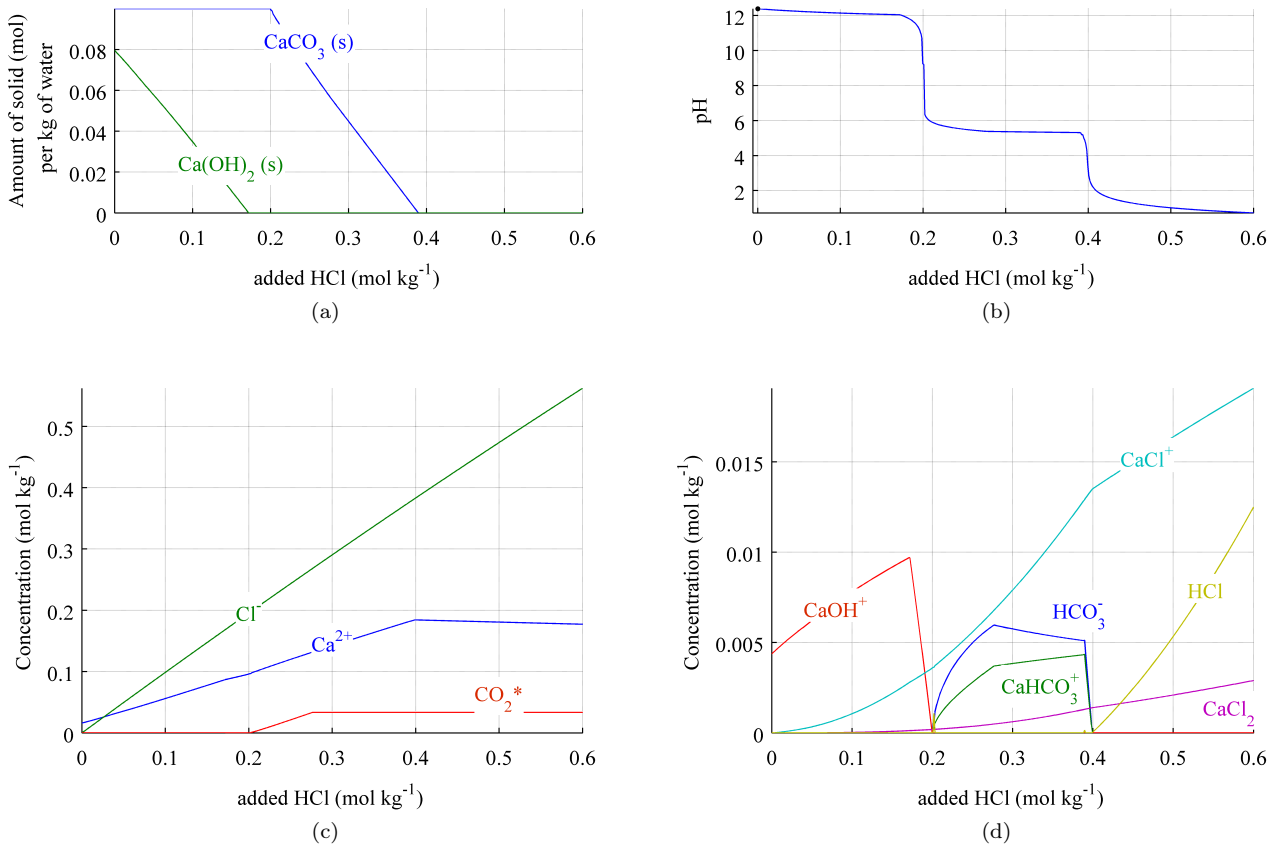


Figure 4: Simultaneous dissolution of portlandite and calcite versus added HCl

The chosen test example shows some of the main features of the model presented. It shows the concentration at the equilibrium in a very common multi-species and multiphase system. Two different solid phases coexist and dissolve in a different range of pH. Furthermore, moderate high concentrations are used to reach the saturation value of aqueous carbon dioxide associated with the release of gaseous carbon dioxide bubbles.

In Fig. 4, the simultaneous dissolution of a mixture of 0.1 mol (10 g) of calcite and 0.1 mol (7.4 g) of Portlandite is shown. In Fig. 4a, the amount of solids is shown with respect to the molal concentration of the hydrochloric acid solution. It can be seen that the dissolution process occurs in a sequential stage. Portlandite is more soluble than calcite. In pure water, about 20% of the initial portlandite was dissolved, while most of the calcite stays in mineral phase. When increasing the concentration of the acid, portlandite tends to dissolve. The portlandite mineral is completely dissolved when concentration of the acid is around 0.2 mol kg⁻¹. Afterwards, calcite starts to dissolve to counteract the addition of extra acid in the media.

The dissolution of calcite and portlandite leads to high concentrations of ionic calcium, Ca^{2+} , as shown in Fig. 4c. When both minerals are completely dissolved, the concentration of Ca^{2+} reaches a maximum. After that moment,

increasing the hydrochloric acid concentration leads to the formation of the complexes CaCl^+ and CaCl_2 with a slight decrease of the aqueous Ca^{2+} . Figure 4c also shows the limit value assigned to the aqueous carbon dioxide as a consequence of the restriction of maximum concentration. In these conditions, formation of $\text{CO}_2(\text{g})$ bubbles is predicted when the concentration of acid is equal to or higher than approximately 0.28 mol kg⁻¹.

The pH value, shown in Fig. 4b, stays approximately constant while the mineral phases are present, showing the buffer capacity of the two solids considered. There are two pH drops related to the moment in which each of the mineral phases completely dissolves. Portlandite keeps the media in alkaline conditions, as typically seen in cement-based materials. Calcite maintains the media in a slightly acid environment (pH = 5).

Fig. 4d shows the concentration of the other aqueous species of interest in the system. According to the results shown here, when coexisting portlandite and calcite, the main secondary aqueous species, apart from OH^- , is the complex CaOH^+ . When only calcite is precipitated, there are bicarbonate complexes in the system. As mentioned before, CaCl^+ , CaCl_2 and HCl increases their concentration in the process, and they are the main secondary species when the mineral phases are completely dissolved.

Table 2: Numerical parameters used in the test simulations

Parameter	value
λ^*	2
$\Delta h_{\text{init.}}$	2^{-25}
Δh_{min}	2^{-250}
h^*	2^{25}
k_{max}	10
j_{max}	250
$\ \mathbf{f}(\mathbf{x})\ _{\text{min}}$	10^{-10}

Table 2 collects the numerical parameters used in the simulations presented here. A total of 500 simulations were done to meet the entire range of HCl concentrations. For an algorithm implemented in Matlab® 2012, and run in a 2.3 GHz Intel Core i5 computer with 4 GB 1333 MHz DDR3, the computing time was 107.13 s.

Conclusions

A method for computing chemical equilibrium speciation problems has been described, based on the extent of each chemical reaction of the system. The numerical solution method is based on a quasi-NR method in which the Jacobian matrix is formed by a numerical technique rather than being calculated analytically. In this context, a vectorized implementation of the Jacobian matrix is proposed in order to significantly reduce the computational time. A line search approach is used to deal with the non-negative concentrations constraint of all the chemical species.

The presented algorithm is designed to work in a multi-species and multiphase chemical systems under the assumption of chemical equilibrium. Problems associated with local minimums are potentially avoided by the prioritization of the aqueous reactions with respect to the solid dissolution or precipitation reactions. This prioritization method is based on the assumption that that the kinetic rates of reactions involving only aqueous species are faster than those involving solid mineral phases.

The formation of gaseous species is included in the model by allowing free release of gas bubbles of gaseous species when those bubbles have a higher pressure than the surrounding liquid media and vapor phase. This feature limits the concentration of volatile aqueous species to a maximum value given by the gas solubility constant.

A test example has been used to illustrate the main features of the model presented here. The use of the extent of the reactions as the state variable makes the presented algorithm suitable for more general problems, in which a thorough control of the electrical balance is required.

Acknowledgments

Authors acknowledge The Danish Agency of Science Technology and Innovation. This research was funded as part of the project “Fundamentals of Electrokinetics in

In-homogeneous Matrices”, with project number 274-08-0386. J.M. Rodríguez-Maroto acknowledges the financial support from the Ministerio de Ciencia e Innovación of the Spanish Government through the project ERHMES, CTM2010-16824, and FEDER funds.

References

- Al-Hamdan, A. Z., & Reddy, K. R. (2008). Electrokinetic remediation modeling incorporating geochemical effects. *J. Geotech. Geoenviron.*, *134*, 91–105.
- Ball, J. W., & Nordstrom, D. K. (1991). *User’s manual for WATEQ4F, with revised thermodynamic data base and test cases for calculating speciation of major, trace, and redox elements in natural waters*. US Geological Survey Denver, CO.
- van Baten, J., & Szczepanski, R. (2011). A thermodynamic equilibrium reactor model as a cape-open unit operation. *Computers and Chemical Engineering*, *35*, 1251–1256.
- Bethke, C. M. (2008). *Geochemical and biogeochemical reaction modeling*. (2nd ed.). New York: Cambridge University Press.
- Blomberg, P. B. A., & Koukkari, P. S. (2011). A systematic method to create reaction constraints for stoichiometric matrices. *Computers and Chemical Engineering*, *35*, 1238–1250.
- Brassard, P., & Bodurtha, P. (2000). A feasible set for chemical speciation problems. *Comput. Geosci.*, *26*, 277–291.
- Carrayrou, J., Mose, R., & Behra, P. (2002). New efficient algorithm for solving thermodynamic chemistry. *AIChE J.*, *48*, 894–904.
- Carroll, S., & Walther, J. (1990). Kaolinite dissolution at 25, 60, and 80 c. *Am. J. Sci.*, *290*, 797–810.
- Charlton, S. R., Macklin, C. L., & Parkhurst, D. L. (1997). Phreeqci—a graphical user interface for the geochemical computer program phreeqc. *Water Resources Investigation Report*, *97*, 4222.
- Greenwood, N. N., & Earnshaw (1984). *Chemistry of the Elements*. Pergamon press Oxford etc.
- Guggenheim, E., & Turgeon, J. (1955). Specific interaction of ions. *Trans. Faraday Soc.*, *51*, 747–761.
- Jacobs, R. A., & Probst, R. F. (1996). Two-dimensional modeling of electroremediation. *AIChE J.*, *42*, 1685–1696.
- Javadi, A., & Al-Najjar, M. (2007). Finite element modeling of contaminant transport in soils including the effect of chemical reactions. *J. Hazard. Mater.*, *143*, 690–701.
- Johannesson, B. (2009). Ionic diffusion and kinetic homogeneous chemical reactions in the pore solution of porous materials with moisture transport. *Comput. Geotech.*, *36*, 577–588.
- Koukkari, P., & Pajarre, R. (2006). Introducing mechanistic kinetics to the lagrangian gibbs energy calculation. *Computers and Chemical Engineering*, *30*, 1189–1196.
- Lichtner, P. C. (1985). Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. *Geochim. Cosmochim. Acta*, *49*, 779–800.
- McNaught, A. D., & Wilkinson, A. (1997). *IUPAC. Compendium of Chemical Terminology. (the "Gold Book")*. (2nd ed.). Oxford: Blackwell Scientific Publications.
- Morel, F., & Hering, J. (1993). *Principles and applications of aquatic chemistry*. Wiley-Interscience.
- Morel, F., & Morgan, J. (1972). Numerical method for computing equilibria in aqueous chemical systems. *Environmental Science & Technology*, *6*, 58–67.
- Parkhurst, D. L., & Appelo, C. A. J. (1999). *User’s guide to PHREEQC (version 2) - A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations*. U.S. Department of the Interior, Water-Resources Investigations Reports (99-4259).
- Peterson, S. R., Hostetler, C. J., Deutsch, W. J., & Cowan, C. E. (1987). *MINTEQ user’s manual*. Technical Report Pacific Northwest Lab., Richland, WA (USA); Nuclear Regulatory Commission, Washington, DC (USA). Div. of Waste Management.

- Pitzer, K. (1973). Thermodynamics of electrolytes I. Theoretical basis and general equations. *J. Phys. Chem.*, *77*, 268–277.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). Root finding and nonlinear sets of equations. In *Numerical recipes in C: The art of Scientific Computing* chapter 9. Cambridge University Press.
- Rawlings, J. B., & Ekerdt, J. G. (2002). *Chemical reactor analysis and design fundamentals*. Madison, Wisconsin: Nob Hill Publishing.
- Reed, M. H. (1982). Calculation of multicomponent chemical equilibria and reaction processes in systems involving minerals, gases and an aqueous phase. *Geochim. Cosmochim. Acta*, *46*, 513–528.
- Remy, N. (2004). Geostatistical earth modeling software: User's manual. *Stanford Center for Reservoir Forecasting (SCRF), Stanford University, California, USA*, .
- Ribeiro, A. B., Rodriguez-Maroto, J. M., Mateus, E. P., & Gomes, H. (2005). Removal of organic contaminants from soils by an electrokinetic process: the case of atrazine. Experimental and modeling. *Chemosphere*, *59*, 1229–1239.
- Rodriguez-Maroto, J. M., & Vereda-Alonso, C. (2009). Electrokinetic modeling of heavy metals. In K. R. Reddy, & C. Camesselle (Eds.), *Electrochemical Remediation Technologies for Polluted Soils, Sediments and Groundwater* chapter 25. (pp. 537–562). Wiley Online Library.
- Rubin, J. (1983). Transport of reacting solutes in porous media: Relation between mathematical nature of problem formulation and chemical nature of reactions. *Water Resour. Res.*, *19*, 1231–1252.
- Steeffel, C., & Van Cappellen, P. (1990). A new kinetic approach to modeling water-rock interaction: The role of nucleation, precursors, and ostwald ripening. *Geochim. Cosmochim. Acta*, *54*, 2657–2677.
- Steyer, F., Flockerzi, D., & Sundmacher, K. (2005). Equilibrium and rate-based approaches to liquid–liquid phase splitting calculations. *Computers and Chemical Engineering*, *30*, 277–284.
- Stumm, W., & Morgan, J. J. (1970). *Aquatic Chemistry; An Introduction Emphasizing Chemical Equilibria in Natural Waters*. (2nd ed.). New York: Wiley-Interscience.
- Vereda-Alonso, C., Heras-Lois, C., Gomez-Lahoz, C., Garcia-Herruzo, F., & Rodriguez-Maroto, J. M. (2007). Ammonia enhanced two-dimensional electrokinetic remediation of copper spiked kaolin. *Electrochim. Acta*, *52*, 3366–3371.
- Wolery, T. J. (1992). *EQ3/6: A software package for geochemical modeling of aqueous systems: package overview and installation guide (version 7.0)*. Lawrence Livermore National Laboratory.
- Wolery, T. J., & Walters, L. J. (1975). Calculation of equilibrium distributions of chemical species in aqueous solutions by means of monotone sequences. *Mathematical Geology*, *7*, 99–115.