

## **A general paradigm to model reaction-based biogeochemical processes in batch systems**

Yilin Fang

Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, Pennsylvania, USA

Gour-Tsyh Yeh

Department of Civil and Environmental Engineering, University of Central Florida, Orlando, Florida, USA

William D. Burgos

Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, Pennsylvania, USA

Received 30 August 2002; revised 25 November 2002; accepted 25 November 2002; published 2 April 2003.

[1] This paper presents the development and illustration of a numerical model of reaction-based geochemical and biochemical processes with mixed equilibrium and kinetic reactions. The objective is to provide a general paradigm for modeling reactive chemicals in batch systems, with expectations that it is applicable to reactive chemical transport problems. The unique aspects of the paradigm are to simultaneously (1) facilitate the segregation (isolation) of linearly independent kinetic reactions and thus enable the formulation and parameterization of individual rates one reaction by one reaction when linearly dependent kinetic reactions are absent, (2) enable the inclusion of virtually any type of equilibrium expressions and kinetic rates users want to specify, (3) reduce problem stiffness by eliminating all fast reactions from the set of ordinary differential equations governing the evolution of kinetic variables, (4) perform systematic operations to remove redundant fast reactions and irrelevant kinetic reactions, (5) systematically define chemical components and explicitly enforce mass conservation, (6) accomplish automation in decoupling fast reactions from slow reactions, and (7) increase the robustness of numerical integration of the governing equations with species switching schemes. None of the existing models to our knowledge has included these scopes simultaneously. This model (BIOGEOCHEM) is a general computer code to simulate biogeochemical processes in batch systems from a reaction-based mechanistic standpoint, and is designed to be easily coupled with transport models. To make the model applicable to a wide range of problems, programmed reaction types include aqueous complexation, adsorption-desorption, ion-exchange, oxidation-reduction, precipitation-dissolution, acid-base reactions, and microbial mediated reactions. In addition, user-specified reaction types can be programmed into the model. Any reaction can be treated as fast/equilibrium or slow/kinetic reaction. An equilibrium reaction is modeled with an infinite rate governed by a mass action equilibrium equation or by a user-specified algebraic equation. Programmed kinetic reaction rates include multiple Monod kinetics, nth order empirical, and elementary formulations. In addition, user-specified rate formulations can be programmed into the model. No existing models to our knowledge offer these simultaneous features. Furthermore, most available reaction-based models assume chemical components a priori so that reactions can be written in basic (canonical) forms and implicitly assume that fast equilibrium reactions occur only for homogeneous reactions. The decoupling of fast reactions from slow reactions lessens the stiffness typical of these systems. The explicit enforcement of mass conservation overcomes the mass conservation error due to numerical integration errors. The removal of redundant fast reactions alleviates the problem of singularity. The exclusion of irrelevant slow reactions eliminates the issue of exporting their problematic rate formulations/parameter estimations to different environment conditions. Taking the advantage of the nonuniqueness of components, a dynamic basis-species switching strategy is employed to make the model numerically robust. Backward basis switching allows components to freely change in the simulation of the chemistry module, while being recovered for transport simulation. Three example

problems were selected to demonstrate the versatility and robustness of the model.

*INDEX TERMS:* 1099 Geochemistry: General or miscellaneous; 1831 Hydrology: Groundwater quality; 1899 Hydrology: General or miscellaneous; 4807 Oceanography: Biological and Chemical: Chemical speciation and complexation; 4825 Oceanography: Biological and Chemical: Geochemistry; *KEYWORDS:* geochemical modeling, biochemical modeling, reaction network, matrix decomposition, basis species switching, batch systems

**Citation:** Fang, Y., G.-T. Yeh, and W. D. Burgos, A general paradigm to model reaction-based biogeochemical processes in batch systems, *Water Resour. Res.*, 39(4), 1083, doi:10.1029/2002WR001694, 2003.

## 1. Introduction

[2] While the coupling of hydrologic transport and chemical reaction models is an active area of research, the development of chemical reaction batch models has received much less attention. Chemical reactions can be divided into two classes [Rubin, 1983]: (1) those that are “sufficiently fast” and reversible, so that local equilibrium may be assumed and (2) those that are “insufficiently fast” and/or irreversible, where the local equilibrium formulation is inappropriate. The modeling of equilibrium chemistry has been well established. Some models for equilibrium chemistry have come into wide use, such as WATEQ [Truesdell and Jones, 1974], MINEQL [Westall et al., 1976], PHREEQE [Parkhurst et al., 1980], EQ3/6 and their derivatives [see Wolery, 1992, and references therein]. However, more complex types of biochemical and geochemical reactions exist in the subsurface or surface waters that cannot be described by equilibrium chemistry alone [see Friedly and Rubin, 1992, and references therein]. Therefore the modeling of equilibrium chemistry coupled with the kinetic chemistry has become a necessity. There are quite a few models, either stand-alone or being embedded in transport codes, that can model mixed kinetic and equilibrium reactions (e.g., model name unknown [Lin and Benjamin, 1990]; KEMOD [Yeh et al., 1995]; OS3D [Steeffel and Yabusaki, 1996]; BIOKEMOD [Salvage and Yeh, 1998]; RAFT [Chilakapati, 1995; Chilakapati et al., 2000]). All of these chemistry modules have had varied scopes, and most have limited capabilities in terms of rate equations for kinetic reactions and reaction types for equilibrium reactions.

[3] A generic chemistry module to describe geochemical and biochemical processes in batch systems is needed to improve reactive transport models. Since qualitative geochemical and biochemical processes must be conceptualized quantitatively as a reaction network [Yeh et al., 2001a], a reaction-based biogeochemical model is conceivably the most generic approach to modeling these processes. This paper presents the development and verification of a generic numerical model (BIOGEOCHEM) using demonstrative examples. The objective is to provide a general paradigm for modeling reactive chemicals in batch systems, with the expectation that this paradigm is applicable to reactive transport systems.

[4] To make the model applicable to as wide a range of problems as possible, BIOGEOCHEM embodies a complete suite of reactions including aqueous complexation, adsorption-desorption, ion-exchange, redox, precipitation-dissolution, acid-base reactions, and microbial mediated reactions. Any reaction can be treated as fast/equilibrium or slow/kinetic reaction. An equilibrium reaction is modeled with an infinite rate governed by a mass action equilibrium equation or by a user-specified algebraic equation. A kinetic

reaction is modeled with a finite rate with microbial mediated enzymatic kinetics, an elementary rate, or a user-specified rate equation. None of the existing models has encompassed this wide array of scopes.

[5] To make the model numerically robust, a dynamic basis-species switching strategy is employed taking the advantage of the nonuniqueness of components. Components are free to change in the simulation of the chemistry module, while being recovered for transport simulation by backward basis switching.

[6] BIOGEOCHEM can be coupled with any hydrologic transport model in the same way that PHREEQ or MINTEQ can be coupled to transport models [Lin and Benjamin, 1990; Steefel and Yabusaki, 1996]. A subsequent paper will present a reactive biogeochemical transport model (HYDROBIOGEOCHEM) based on the same paradigm described in this paper. Instead of performing the matrix decomposition on species balance equations, one performs the decomposition on species transport equations in transport systems. It is expected that this matrix decomposition approach can be used by others to modify their reactive transport models to improve the design capabilities. For this reason, we will first review some existing reactive transport models that could have taken advantage of this generic paradigm before we outline what will be included in this paper.

[7] Reactive chemical transport models have had varied scopes. Conventional solute transport models often ignore chemical speciation in the aqueous phase. Much attention has been given to heterogeneous reactions with partitioning between aqueous and sorbed chemicals represented by linear ( $K_d$  approach) or nonlinear (Freundlich and/or Langmuir) isotherms [Yeh and Tripathi, 1991; Davis et al., 2000]. Most models cannot account for the complete set of biogeochemical processes (biogeochemical processes are “coupled organic and inorganic reaction processes” [Brun and Engesgaard, 2002]) and they cannot be easily extended to include mixed equilibrium/kinetic reactions or mixed chemical/microbial reactions using mechanistic rate formulations. From a geochemical point of view, approaches that use empirical partitioning reactions can be considered as reactive chemical transport because they are dealing with adsorption/desorption phenomena. However, from a modeling point of view, we hesitate to classify such models as reactive transport because a “true” reactive chemical transport models should be based on the principles of thermodynamics (for fast/equilibrium reactions) and chemical kinetics (for slow/kinetic reactions). Many conventional transport models [e.g., van der Zee and van Riemsdijk, 1987; Bosma and van der Zee, 1993; Tompson, 1993; Toride et al., 1993; Brusseau, 1994] that have been proclaimed to be reactive transport models perhaps should not be categorized as such.

[8] True reactive chemical transport models have been extensively documented. Many models couple simulation of transport with equilibrium geochemistry [e.g., *Miller and Benson*, 1983; *Cederberg et al.*, 1985; *Hostetler and Erickson*, 1989; *Narasimhan et al.*, 1986; *Liu and Narasimhan*, 1989a, 1989b; *Griffioen*, 1993; *Yeh and Tripathi*, 1991; *Cheng*, 1995; *Parkhurst*, 1995; *Parkhurst and Appelo*, 1999]. Some models couple transport with kinetic geochemistry for certain geochemical processes like precipitation-dissolution [e.g., *Lichtner*, 1996; *Steeffel and Yabusaki*, 1996; *Suarez and Simunek*, 1996], adsorption [e.g., *Theis et al.*, 1982; *Szecsody et al.*, 1998], redox [*Lensing et al.*, 1994; *Saiers et al.*, 2000], or biodegradation [*MacQuarrie et al.*, 1990; *Chen et al.*, 1992; *Chang et al.*, 1993; *Cheng and Yeh*, 1994; *Wood et al.*, 1994].

[9] Models coupling transport with mixed equilibrium/kinetic reactions have appeared since the mid-1990s [e.g., *Steeffel and Lasaga*, 1994; *McNab and Narasimhan*, 1994, 1995; *Salvage et al.*, 1996; *Yeh et al.*, 1996; *Abrams et al.*, 1998; *Chilakapati et al.*, 1998; *Tebes-Stevens et al.*, 1998; *Salvage and Yeh*, 1998; *Yeh et al.*, 2001b; *Brun and Engesgaard*, 2002]. For most models chemical components must be selected a priori such that only limited reaction network can be considered [e.g., *Parkhurst*, 1995; *Parkhurst and Appelo*, 1999]. Most models have implicitly assumed that equilibrium reactions occur only among aqueous species so that transport of components can be manually decoupled from fast reactions [*Steeffel and Lasaga*, 1994; *Lichtner*, 1996; *Tebes-Stevens et al.*, 1998; *Parkhurst and Appelo*, 1999]. However, in a complicated reactive system, the identification of component species may not be so easy when there are many parallel kinetic reactions [*Friedly and Rubin*, 1992; *Chilakapati et al.*, 1998; *Yeh et al.*, 2000]. Under such circumstances, matrix methods may be better employed to define component species and derive governing equations to model mixed equilibrium/kinetic reactions. There appears to be few general purpose transport models that can simulate a generic reaction including both biochemical and geochemical reactions, and mixed equilibrium/kinetic reactions [*Fang and Yeh*, 2002]. Instead, recent reactive biogeochemical transport models either add geochemical reactions to biodegradation transport models or add simple biodegradation reactions to geochemical transport models [*Brun and Engesgaard*, 2002].

[10] This paper is organized as follows. In section 2, a reaction network used to represent a biogeochemical system is defined with a simple example. Possible difficulties in using primitive or DAE (mixed differential/algebraic equation) approaches to modeling complex systems are heuristically exposed. In section 3, a formal decomposition of the reaction network, which results in a new numerical approach to setting up and integrating differential equations, will be presented using two sets of results (one from a geochemist's point view of components, the other with a totally different designation of components). Advantages of using this generic paradigm are discussed. In section 4, reaction rate formulations for kinetic reactions are discussed. In section 5, the numerical solution of the governing equations is outlined. In section 6, the general application of BIOGEOCHEM is presented using three demonstrative examples. In section 7, conclusions and discussions from the current research are summarized.

## 2. Reaction Network

### 2.1. Definition

[11] Reactive systems are completely defined by specifying reaction networks [*Yeh et al.*, 2001a]. From a mathematical point of view, a system of  $M$  ordinary differential equations can be written for  $M$  chemical species in a reactive, well-mixed batch system as

$$\frac{dC_i}{dt} = r_{i|N}, i \in M \quad (1)$$

where  $C_i$  is the concentration of the  $i$ th chemical species,  $t$  is time, and  $r_{i|N}$  is the production-consumption rate of the  $i$ th species due to  $N$  biogeochemical reactions. The determination of  $r_{i|N}$  and associated parameters is a primary challenge in biogeochemical modeling. There are two general models to formulate  $r_{i|N}$ : ad hoc and reaction-based models, distinctions between which have been discussed extensively [*Yeh et al.*, 2001a]. It should be noted that equation (1) can be easily extended to transport systems by replacing the ordinary differential equation with a transport equation:  $\frac{\partial C_i}{\partial t} + L(C_i) = r_{i|N}$  where  $L$  is the advection-dispersion/diffusion operator. Thus the framework provided in this paper is also applicable to transport systems except that, instead of equation (1), one employs the set of transport equations. A subsequent paper will present the development of a biogeochemical transport model (HYDROBIOGEO-CHEM) using the same paradigm described here [*Fang and Yeh*, 2002].

[12] In an ad hoc model, the production-consumption rate of a species is described with an empirical function

$$\frac{dC_i}{dt} = r_{i|N} = f_i(C_1, C_2, \dots, C_M; p_1, p_2, \dots) \quad (2)$$

where  $f_i$  is the empirical function for the  $i$ th species and  $p_1, p_2, \dots$  are rate parameters used to fit experimental data. The ad hoc formulations and their associated fitting parameters may be applicable only to the specific experiment tested, because in these ad hoc approaches, the contribution of all geochemical and biochemical processes is lumped and the contribution from individual reactions is not explicitly modeled. Many widely used models (e.g., WASP [*Ambrose et al.*, 1988] and QUAL2E [*Barnwell and Brown*, 1987], and all WASP-based and QUAL2E-based water quality modeling) have taken this approach. It is not difficult to see that calibrated rate parameters using the ad hoc approach are only applicable to the environmental condition under study because the contribution from individual reactions are not segregated. As a result, new calibrations have to be performed for every studies under different environmental conditions when a different set of reactions is the contributing chemical processes.

[13] In a reaction-based model, the production-consumption rates of  $M$  species are described by

$$\frac{dC_i}{dt} = r_{i|N} = \sum_{k=1}^N (\nu_{ik} - \mu_{ik})R_k, i \in M \quad (3)$$

where,  $\nu_{ik}$  is the reaction stoichiometry of the  $i$ th species in the  $k$ th reaction associated with the products,  $\mu_{ik}$  is the



reaction stoichiometry of the  $i$ th species in the  $k$ th reaction associated with the reactants, and  $R_k$  is the rate of the  $k$ th reaction. Equation (3) is a statement of mass balance for any species  $i$  in a batch system. It simply states that the rate of change of mass of any species is due to all reactions that produce or consume that species. This formulation can be extended to transport systems by replacing the ordinary differential equation with the transport equation:  $\frac{\partial C_i}{\partial t} + L(C_i) = r_i|_N = \sum_{k=1}^N (n_{ik} - m_{ik})R_k$ . This partial differential equation is then a statement of mass balance for any species at a point in a transport system. It simply states that the rate of change of mass of any chemical species at any point is due to the net transport to or from the point and due to all chemical reactions occurring at the point.

[14] In a reaction-based approach, the contribution from all individual reactions is explicitly modeled. Thus a properly formulated and parameterized rate equation may still find its application to a wide range of environmental conditions. This is so because rate equations are description of reactions and because under a different environmental condition, only the reaction-rate equations (not the lumped rate equations) associated with the reactions that are operative under the conditions are used. Thus a reaction-based approach is superior to an ad hoc approach. A reaction based approach using mechanistic rate formulations would be the ultimate goal of biogeochemical modeling, though it would be difficult to obtain.

## 2.2. Example

[15] Let us consider a simple system with the following reaction network for nitrotriacetic acid (NTA), cobalt(II), and microbial biomass, B (simplified from the one modeled by Yeh et al. [1998]):



an equilibrium reaction with  $R_1 = \infty$ ;



an equilibrium reaction with  $R_2 = \infty$ ;



a kinetic reaction, with its rate  $R_3$  that must be formulated;



irreversible microbial degradation reaction with its rate  $R_4$  that must be formulated, where  $R_1$ ,  $R_2$ ,  $R_3$ , and  $R_4$  denote the reaction number in the system and  $R_1$ ,  $R_2$ ,  $R_3$ , and  $R_4$  denote the corresponding rates of reaction. This notation is used throughout the paper. It is noted that we have considered the rate of an equilibrium reaction infinity as in  $R_1$  and  $R_2$ .

[16] The concept that the rate of an equilibrium reaction is infinity (not indeterminate) is very important. All reactions in reality are kinetics and take time, however small, to reach equilibrium. In the limit, it takes zero time to reach equilibrium (the “zero time” is, of course, a mathematical abstraction). With this abstraction, one uses the thermodynamic approach to model a fast reaction resulting in the

concept of infinite rate, rather than uses the chemical kinetic approach to model a fast reaction resulting in the concept of indeterminate rate. Consider a simple system of three species subject to one fast reaction as follows:  $A + B \rightleftharpoons C$ . Using the reaction-based approach, we can write the following three ordinary differential equations (ODEs) governing the evolution of three species- concentrations for this reactive system:  $\frac{d[A]}{dt} = -R$ ;  $\frac{d[B]}{dt} = -R$ ; and  $\frac{d[C]}{dt} = R$  where  $[A]$ ,  $[B]$ , and  $[C]$  are the concentrations of species A, B, and C, respectively, and  $R$  is the reaction rate. From a kinetic point of view, when the reaction is fast and quickly achieves to equilibrium, we can formulate  $R$  as the asymptotic approximation of the elementary rate with the backward rate constant approaching infinity. With this asymptotic formulation, the three ODEs become:  $\frac{d[A]}{dt} = k^b([C] - K^e[A][B])$ ;  $\frac{d[B]}{dt} = k^b([C] - K^e[A][B])$ ; and  $\frac{d[C]}{dt} = -k^b([C] - K^e[A][B])$ , where  $k^b$  and  $K^e$  are the very large (in the limit infinity) backward rate constant and the fixed equilibrium constant, respectively. Theoretically, one can solve these three ODEs to yield the equilibrium concentrations,  $[A]_e$ ,  $[B]_e$ , and  $[C]_e$ , given the initial concentrations,  $[A]_o$ ,  $[B]_o$ , and  $[C]_o$ . [At equilibrium, the rate computed with  $R = -\lim_{k^b \rightarrow \infty} k^b([C]_e - K^e[A]_e[B]_e)$  is indeterminate.] In practice, this asymptotic approach is extremely difficult to solve numerically because the system is very stiff or infinitely stiff in the limit. Alternatively, from a thermodynamic point of view, we do not have to formulate the reaction rate  $R$ . Instead, we simply postulate that a fast reaction reaches equilibrium instantaneously. The reaction is then described with a thermodynamically consistent expression, for example:  $R = \infty \exists [C] = K^e[A][B]$  (or other thermodynamically consistent algebraic expressions). (Note this expression is not the same as to formulating  $R$  as  $R = -\lim_{k^b \rightarrow \infty} k^b([C] - K^e[A][B])$  a priori).  $R = \infty$  is a mathematical abstraction.

[17] An important question is then: is it an appropriate abstraction? To answer this question, we need to find out: whether  $R$  is infinity or not (i.e.,  $R \stackrel{?}{=} \infty$ ) assuming that the reaction reaches equilibrium instantaneously? Now one of the system equations is  $[C] = K^e[A][B]$ . The other two equations can be derived with algebraic manipulations of the three ODEs to yield two component equations:  $[A] + [C] = [A]_o + [C]_o$  and  $[B] + [C] = [B]_o + [C]_o$ . Given the initial concentrations,  $[A]_o$ ,  $[B]_o$ , and  $[C]_o$ , we solve these three algebraic equations to yield three concentrations at equilibrium,  $[A]_e$ ,  $[B]_e$ , and  $[C]_e$ . The reaction rate is finally computed with  $R = \frac{d[C]}{dt} = \frac{[C]_e - [C]_o}{\Delta t = 0} = \infty$  (keep in mind that the reaction reaches equilibrium instantaneously, so  $\Delta t = 0$ ), not with  $R = -\lim_{k^b \rightarrow \infty} k^b([C]_e - K^e[A]_e[B]_e)$  that has never entered into the picture in the thermodynamic approach. Indeed, the reaction rate for an equilibrium reaction is infinity, and the mathematical abstraction is an appropriate one. In other words, when we model a fast reaction with a thermodynamic expression, the rate corresponding to this reaction is infinity when it is used in reaction-based mass balance equations. This mathematical abstraction of treating a fast rate as infinity is also valid and consistent for a system with mixed equilibrium and kinetic reactions as shall be addressed in the subsequent sections. As we shall see later, the abstraction of an infinite rate is equivalent to saying that the concentration of an equilibrium variable (the definition of the equilibrium variable under mixed kinetic and

equilibrium reactions will be given later) reaches its equilibrium value instantaneously at any time. The conceptualization of a fast reaction as having a rate of infinity, when this rate is used in chemical kinetic equations, is thus justified.

[18] In contrast, it has been argued by many authors [e.g., *Lichtner*, 1996] that the rate for an equilibrium reaction is indeterminate. This argument has resulted from the calculation of  $R$  using  $R = -\lim_{k^b \rightarrow \infty} k^b ([C]_e - K^e [A]_e [B]_e)$ . This argument is logical if a fast reaction is modeled with the chemical kinetic approach using the asymptotic approximation of an elementary rate. However, the argument is incorrect when the fast reaction is modeled with the thermodynamic approach. Since the rate  $R$  has not been formulated a priori when a consistent thermodynamic expression is directly used to model the fast reaction, the rate cannot be computed with  $R = -\lim_{k^b \rightarrow \infty} k^b ([C]_e - K^e [A]_e [B]_e)$ , but must be computed with  $R = \frac{d[C]}{dt} = \frac{[C]_e - [C]_a}{\Delta t} = \infty$ . Therefore the argument that the rate of an equilibrium reaction is indeterminate is flawed when a thermodynamic approach is used to model the reaction.

[19] To distinguish the conceptual difference between kinetic modeling and thermodynamic modeling of a fast reaction, the equation,  $R = -\lim_{k^b \rightarrow \infty} k^b ([C]_e - K^e [A]_e [B]_e)$ , is called a mass action kinetic equation while the equation,  $R = \infty \exists [C]_e = K^e [A]_e [B]_e$  is called a mass action equilibrium equation throughout this paper. The concept of infinite rates for equilibrium reactions is used to facilitate the discussion on the difficulties encountered in solving a set of ODEs involving mixed fast/slow reactions in the following section.

[20] The species charges of the reactions R1 through R4 are not included for simplicity of presentation. In this system, the total number of species  $M = 6$ , and the total number of reactions  $N = 4$ . From equation (3), the mass balance equations for the reaction network reactions R1 through R4 can be written as

$$\frac{d[H]}{dt} = -R_1 - R_3 - R_4 \quad (4a)$$

$$\frac{d[NTA]}{dt} = -R_1 + R_2 \quad (4b)$$

$$\frac{d[HNTA]}{dt} = R_1 + R_3 - R_4 \quad (4c)$$

$$\frac{d[Co]}{dt} = R_2 + R_3 \quad (4d)$$

$$\frac{d[CoNTA]}{dt} = -R_2 - R_3 \quad (4e)$$

$$\frac{d[B]}{dt} = R_4 \quad (4f)$$

### 2.3. Difficulties in Solving Equations (4a) Through (4f)

[21] An analytical solution of equations (4a) through (4f) is generally impossible, which makes numerical integration

attractive. The majority of existing reactive chemical models has taken either a primitive approach (i.e., integrate equation (3) directly) or a mixed differential and algebraic equations (DAEs) approach. In a DAE approach, most models, with a few exception [e.g., *Friedly and Rubin*, 1992; *Chilakapati et al.*, 1998], do not have the capability to automate the setup of DAEs based on equations (4a) through (4f). Instead, the set of DAEs is manually and directly obtained based on the reactive system [e.g., *Liu et al.*, 2001]. For the reactive network R1 through R4, one can easily set up two mass action equilibrium equations governing the fast reactions (refer to equations (6a) and (6b)), an ordinary differential equation governing the slow/kinetic reaction (refer to equation (6c)), and three component equations governing three mass conservation of  $TOT_H$ ,  $TOT_{Co}$ , and  $TOT_{NTA}$  (refer to equations (6d), (6e), and (6f)). This system of DAEs is then numerically integrated simultaneously. The manual and direct setup of mass action equilibrium equations, ordinary differential equations, and mass conservation equations is easy for this simple system. For a generic system, with many fast and slow reactions including parallel reactions, a manual setup will be difficult even for experienced modelers and/or chemists. Under such circumstances, a better way is to automate the generation of a new but equivalent governing set of equations [*Friedly and Rubin*, 1992; *Chilakapati et al.*, 1998; *Yeh et al.*, 2001a]. Another problem is that when fast reactions are not written in basic form (a reaction is defined as basic in this paper when all reactants are component species and includes only one product species), they are not easy to decouple from slow reactions. Clearly, a systematic approach is needed to achieve the mission of decoupling (i.e., eliminating fast reactions from simultaneous solution).

[22] In a primitive approach, direct numerical integration of the system of ODEs is performed using very large backward and forward rate constants (with their ratio defining the equilibrium constant) to mimic the rates of fast/equilibrium reactions. Several numerical difficulties can be encountered in this approach. First, the reaction rates of  $N$  reactions can, in general, range over several orders of magnitude. The time step size used in numerical integration is dictated by the largest reaction rate among  $N$  reactions. If at least one of the reactions is fast/equilibrium, its rate is infinitely large. As a result, the time step size must be infinitely small, which makes integration impractical. For example, in equations (4a) through (4e),  $R_1$  and  $R_2$  are infinitely large, which dictates the use of an infinitely small time step for numerical integration. Second, physics dictates that the number of linearly independent reactions must be less than the number of species. This implies that there must be one or more chemical components whose mass must be conserved during the reactions. For example, there are three linearly independent reactions in the above reaction network. Thus there must be three components ( $N_C = M - N_I$ ). If Co, H, and NTA are selected (at the discretion of a user) as the three chemical components, then the  $TOT_{Co}$ ,  $TOT_H$ , and  $TOT_{NTA}$ , defined in next section, must be invariant. The direct numerical integration of equations (4a) through (4f) yields the solution of all six species assuming one can afford the use of an infinitely small time step size. Because of numerical errors, there are no assurances that  $TOT_{Co}$ ,  $TOT_H$ , and  $TOT_{NTA}$ , as defined in equations (6d) through

(6f), respectively, are invariant when the primitive set of equations are integrated directly. Similarly, in case of transport systems, even without spatial discretization errors, there are no assurances that  $TOT_{Co}$ ,  $TOT_H$ , and  $TOT_{NTA}$  obtained from solving the system of primitive PDEs (partial differential equations) will satisfy the component transport equations.

[23] Third, if more than one equilibrium reactions are involved, there is no way to define the subtraction or addition of infinity. For example, in equation (4b), both  $R_1$  and  $R_2$  are infinitely large, so how can one interpret  $(-R_1 + R_2)$ ? Fourth, redundant fast/equilibrium reactions must be removed from consideration, otherwise the system would become singular or over-constrained. Redundant reactions can easily be detected and excluded from consideration manually when the system is simple and components are chosen a priori. When there are many fast/slow reactions, redundancies are not easy to detect and a systematic way must be employed to detect them. Fifth, the inclusion of irrelevant slow/kinetic reactions makes their rate formulations and parameter determinations meaningless because they are insensitive to the system. Because of their insensitivity, any rate formulation and parameter estimate are as good as any others. As a result, their rate formulations and parameter estimates are not true descriptions of these rates. Hence these rate formulations and parameters are not portable to other conditions, when they are relevant, unless one is lucky enough to have coincidentally come up with true rate formulations and parameters. Therefore it is desirable to exclude these irrelevant reactions with a systematic approach.

[24] Sixth, which is the most important difficulty, even if all reactions are slow/kinetic, their rates are coupled via the concentration-versus-time curves of all species. They cannot be formulated and parameterized one reaction by one reaction independently of each other.

[25] While a DAE approach can overcome the first difficulty of system stiffness, it greatly increases the computational burdens when there are numerous fast reactions or if the model is coupled with transport when hundreds or thousands of DAE problems must be solved. A DAE approach is definitely able to alleviate the second difficulty of mass balance error as it also explicitly enforces mass conservation. The problem is that the number of species that must be included in a component equation is not obvious and may be specified incorrectly when many parallel reactions are involved. The issues related to the third, fourth, and fifth difficulties are eliminated from a DAE approach only if all reactions are written in basic form. For a realistic system it would be difficult to write all reactions in basic form, thus the DAE approach will be demanding to address these issues. Finally, a DAE approach definitely cannot resolve the sixth difficulty related to kinetic rate coupling.

[26] In spite of the difficulties outlined above, the majority of literature has taken the primitive approach of directly integrating the system of ODEs [e.g., *McNab and Narasimhan*, 1993, 1994; *Chilakapati*, 1995; *Saiers et al.*, 2000] or the DAE approach [e.g., *Miller and Benson*, 1983; *Chilakapati et al.*, 1998; *Liu et al.*, 2001]. Neither the primitive nor the DAE approaches will pose computational burdens when the system is small, for example, when modeling laboratory experiments [e.g., *Lin and Benjamin*, 1990; *Szecsody et al.*, 1994, 1998]. However, for large problems or when the reactive chemical model is intended

to couple with hydrologic transport, a systematic approach that can overcome the above difficulties is needed. This paper presents the development of such a model using the well-known Gauss-Jordan matrix decomposition.

### 3. Decomposition of Reaction Network

[27] Equations (4a) through (4f) govern the evolution of six species and contain both fast and slow reactions on the right hand sides. Obviously, the above system of equations is very stiff if fast reactions are modeled with mass action kinetic equations because the rate constants are very large and infinity in the limit. To reduce the stiffness of the system, a direct thermodynamic approach may be employed to model the two fast reactions; then two of the above ODEs may be replaced by two thermodynamically consistent equilibrium expressions (mass action equilibrium equations or user-specified algebraic equations). But difficulties still exist. First, we do not know which two ODEs should be replaced (i.e., we do not know how to define the subtraction or addition of infinity). Second, even if we know which two to replace, the remaining ODEs still contain fast reactions and the system is still stiff. Therefore a systematic way must be provided to reduce the system of equations such that (1) no more than one fast reaction is allowed in any reduced ODE and (2) any fast reaction is not allowed to appear in more than one reduced ODE. These two constraints can be met if and only if all fast reactions are linearly independent. The use of Gauss-Jordan reduction on the matrix made up of only fast reactions can determine if all fast reactions are linearly independent. During reduction process, any reaction detected to be dependent reaction must be removed and the constraints can be satisfied. With these constraints, the number of reduced ODEs that contain fast rates will be equal to the number of nonremoved fast reactions after the reduction. Any reduced ODE that contains one and the only one linearly independent fast reaction may still contain other slow reactions that are linearly dependent on this fast reaction. However, because the rates of all slow reactions are conceptualized infinitely small compared to the rate of the fast reaction, these rates can be discarded and the reduced ODE can be replaced by the corresponding equilibrium expression. This resolves the problem of not knowing which equation is to be replaced. The replacement can be done for every reduced ODE in which the number of fast reactions is only one. After the completion of the replacement, the remaining ODEs will no longer contain any fast reactions, and the problem of stiffness is resolved. The reduction can be achieved via the Gauss-Jordan decomposition of the reaction matrix which will be presented in the following sections.

#### 3.1. Description of Reaction Network Decomposition

[28] As stated in section 2, to make possible the formulation of rate equations one reaction by one reaction and to enable robust and efficient numerical integration of equation (3), a systematic approach rather than the primitive or DAE approaches is necessary. The approach can be achieved with a diagonalization of equation (3). Equation (3) written in matrix form can be decomposed based on the type of biogeochemical reactions via Gauss-Jordan column reduction of the reaction matrix [*Chilakapati*, 1995; *Steeffel and*



MacQuarrie, 1996; Yeh *et al.*, 2001a]. Each column of the reaction matrix is made up of reaction stoichiometries of a reaction. To perform column reduction is to determine a pivot element, and use a matrix row operation to convert the column containing the pivot element into a unit column. The example from section 2 is used to illustrate the decomposition and demonstrate how the difficulties presented earlier can be overcome.

[29] The system of equations (4a) through (4f) written in matrix form is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{pmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{pmatrix} \quad (5)$$

The decomposition of equation (5) is not unique. Step-by-step decomposition procedures for this simple example are given in Appendix A. For complex systems, the decomposition procedures are provided within the BIOGEOCHEM preprocessor (description in section 3.2). Two different decompositions are described and compared below.

### 3.1.1. Decomposition I

[30] One possible decomposition, when choosing Co, H, and NTA as chemical components, results in the following set of six equations for the six species (Appendix A):

$$\frac{d([HNTA] + [B])}{dt} = R_1 + R_3 : \frac{d([HNTA] + [B])}{dt} \approx R_1 = \infty$$

$$\exists [HNTA] = K_1^c [H][NTA] \quad (6a)$$

$$\frac{d[CoNTA]}{dt} = -R_2 - R_3 : \frac{d[CoNTA]}{dt} \approx -R_2 = -\infty$$

$$\exists [Co][NTA] = K_2^c [CoNTA] \quad (6b)$$

$$\frac{d[B]}{dt} = R_4 \quad (6c)$$

$$\frac{d([H] + [HNTA] + 2[B])}{dt} = 0 \Rightarrow \text{TOT}_H = [H] + [HNTA] + 2[B] = \text{const} \quad (6d)$$

$$\frac{d([NTA] + [HNTA] + [CoNTA] + [B])}{dt} = 0 \quad (6e)$$

$$\Rightarrow \text{TOT}_{NTA} = [NTA] + [HNTA] + [CoNTA] + [B] = \text{const}$$

$$\frac{d([Co] + [CoNTA])}{dt} = 0 \Rightarrow \text{TOT}_{Co} = [Co] + [CoNTA] = \text{const} \quad (6f)$$

The two variables ( $[HNTA] + [B]$ ) and  $[CoNTA]$  in the differential operator in equations (6a) and (6b) are defined as equilibrium variables. The variable  $[B]$  in equation (6c) is defined as the kinetic variable. The three variables, ( $[H] + [HNTA] + 2[B]$ ), ( $[NTA] + [HNTA] + [CoNTA] + [B]$ ), and ( $[Co] + [CoNTA]$ ), in equations (6d) through (6f) are defined as component variables. It should be noted that  $R_3$  appears in both equations (6a) and (6b) because  $R_3$  is linearly dependent on  $R_1$  and  $R_2$ .

[31] This first decomposition yields the above set of equations which, one may argue, can be easily set up manually from a geochemist's point of view. However, it may be easy to miss species B in defining the total H and total NTA even for this simple system. For complex systems, it may not be obvious to generate such a system manually, and a systematic approach using the matrix decomposition will accomplish this task.

[32] Now let us discuss how the diagonalization procedure (i.e., decomposition of reaction matrix) overcomes the difficulties stated in section 2. First, it is noted that either none or one and only one fast reaction appears in any of the above six ODEs. Since the rate of a fast reaction is infinitely large,  $R_3$  is small compared to  $R_1 = \infty$  in equation (6a), the ODE is reduced to  $\frac{d([HNTA] + [B])}{dt} \approx R_1 = \infty$ . The infinity rate is interpreted as a reaction that can reach equilibrium instantly. Thus  $\frac{d([HNTA] + [B])}{dt} \approx R_1 = \infty$  states the use of thermodynamics (not chemical kinetics) to model  $R_1$ , which results in the last equation in equation (6a), and the equilibrium variable, ( $[HNTA] + [B]$ ), reaches its equilibrium value instantaneously at any time. It is clear that the equilibrium variable ( $[HNTA] + [B]$ ) for this system of mixed fast and slow reactions is analogous to the variable  $[C]$  in the simple system described in the second paragraph in section 2.2. Similarly,  $R_2 = \infty$  results in the last equation in equation (6b). With the decomposition of reaction networks and the modeling of fast reactions with thermodynamically consistent expressions, no more ODEs involve infinite rates, and the simulation time step size is dictated only by slow kinetic rates, which makes the system much less stiff than in the primitive approach. These reductions cannot be applied to the primitive approach using equations (4a) through (4f) since either more than one infinite rate appears in one equation (e.g.,  $R_1$  and  $R_2$  in equation (4b)) or one infinite rate appears in more than one equation (e.g.,  $R_1$  in equations (4a), (4b), and (4c) or  $R_2$  in equations (4b), (4d), and (4e)).

[33] Second, there is no reaction rate appearing in equations (6d), (6e), and (6f). This signifies that there are three components, which naturally results from the fact that the rank of the reaction matrix is three. (The rank of a reaction matrix is equal to the number of linearly independent reactions,  $N_I$ . The number of components,  $N_C$ , is equal to the number of species,  $M$ , minus the number of linearly independent reactions (i.e.,  $N_C = M - N_I = 6 - 3 = 3$ )). Again, the components are determined systematically, and the question of how many species are used to define a component is automated. The solution of the diagonalized set of equations (6a) through (6f) explicitly enforces the mass conservation of the components Co, H, and NTA. Third, because none or one and only one infinite rate appears in each equation, the issue of subtraction or addition of infinite rates is moot.

[34] Fourth, R3 appears in both equations (6a) and (6b). This indicates that this reaction is linearly dependent on R1 and R2. If this reaction were also assumed to be a fast reaction, it must be removed from consideration because two infinite rates are not allowed to appear in any ODE in a diagonalized system. If R3 is not removed, then three infinite rates state three mass action equilibrium equations, but the mass action equilibrium equation for R3 is a combination of those for R1 and R2. If all three mass action equilibrium equations were used, the system would become singular or over-constrained, which would necessitate the removal of the dependent fast reaction R3. This automatic removal of fast reactions that are linearly dependent on other fast reactions circumvents the problem of singularity in the system. Fifth, if R3 is a slow reaction as assumed in this case, since R3 is linearly dependent on only two fast reactions R1 and R2, its rate appears only in the equations where rates  $R_1$  and  $R_2$  appear after decomposition. Because  $R_3$  is small compared to  $R_1$  and  $R_2$ , it disappears from equations (6a) and (6b). Also, it never appeared in any of the other four equations after decomposition. Thus  $R_3$  does not appear in any of the final six governing equations and hence is irrelevant to the system, and is automatically removed from the system by the BIOGEOCHEM preprocessor. Five difficulties presented in section 2 have now been resolved; thus the advantages of the diagonalization-decomposition approach over the primitive approach should become clear.

[35] Finally, the mass action equilibrium equations, the last equations in equations (6a) and (6b), can be decoupled from equations (6c) through (6f) by eliminating [HNTA] and [CoNTA] (the procedure of elimination is given in section 3.2.2 and a simple example is given in section A3 in Appendix A). In other words, after the substitution of the last equations of equations (6a) and (6b) into equations (6c) through (6f), [H], [Co], [NTA], and [B] can be obtained by solving three linear algebraic equations (equations (6d) through (6f)) and one ODE (equation (6c)). Then [HNTA] and [CoNTA] can be simply calculated using the last equations of equations (6a) and (6b), respectively. In general, equilibrium reactions are decoupled from kinetic reactions. This offers great advantages over DAE approaches when the system involves a large number of fast/equilibrium reactions.

### 3.1.2. Decomposition II

[36] A second decomposition, when HNTA, CoNTA, and B are chosen as the components, yields the following six equations (Appendix A).

$$\frac{d([NTA] - [Co])}{dt} = -R_1 - R_3 : \frac{d([NTA] - [Co])}{dt} \approx -R_1 = -\infty$$

$$\exists [HNTA] = K_1^e [H][NTA] \quad (7a)$$

$$\frac{d[Co]}{dt} = R_2 + R_3 : \frac{d[Co]}{dt} \approx R_2 = \infty$$

$$\exists [Co][NTA] = K_2^e [CoNTA] \quad (7b)$$

$$\frac{d([H] - [NTA] + [Co])}{dt} = -R_4 \quad (7c)$$

$$\frac{d(-[H] + 2[NTA] + [HNTA] - 2[Co])}{dt} = 0 \quad (7d)$$

$$\Rightarrow \text{Tot}_{HNTA} = [HNTA] + 2[NTA] + [HNTA] - 2[Co] = \text{const}$$

$$\frac{d([Co] + [CoNTA])}{dt} = 0 \Rightarrow \text{Tot}_{CoNTA} = [CoNTA] + [Co] = \text{const} \quad (7e)$$

$$\frac{d([B] + [H] - [NTA] + [Co])}{dt} = 0 \quad (7f)$$

$$\Rightarrow \text{Tot}_B = [B] + [Co] + [H] - [NTA] = \text{const}$$

[37] While the mass conservation of Co, NTA, and H may be intuitive to geochemists, the mass conservation of HNTA, CoNTA, and B may be a surprise. These alternative mass conservation equations are nonetheless mathematically as correct as the mass conservation of Co, NTA, and H. This is beside the point. The question then becomes why bother with the second decomposition? In the first intuitive decomposition, the reaction rate  $R_4$  is measured by the evolution of the [B] versus time curve (equation (6c)). If the concentration of microbial biomass, [B], is difficult to measure, the second decomposition may offer a way to determine the reaction rate  $R_4$  by measuring the  $([H] - [NTA] + [Co])$ -versus-time curve if the concentration of H (or pH), NTA, and Co can be measured easier. Under such circumstances, the second decomposition offers an attractive way of formulating reaction rate  $R_4$ . In any event, the nonuniqueness of the matrix decomposition may allow employment of appropriate governing equations for easier formulations of rate equations [Burgos *et al.*, 2002, 2003] and/or for robust numerical computation. The variable that can be used to measure a kinetic reaction is defined as a kinetic variable. A kinetic variable may consist of just one species (e.g., equation (6c)) or a combination of species (e.g., equation (7c)) depending on the reaction matrix decomposition.

[38] Recall that in Decomposition I, the mass action equilibrium equations, equations (6a) and (6b), were decoupled from equations (6c) through (6f) by eliminating [HNTA] and [CoNTA]. In this alternative decomposition, since HNTA and CoNTA are chosen as the component species, they are not eligible species for elimination. Instead, the decoupling of equations (7a) and (7b) from equations (7c) through (7f) is achieved by eliminating two species (e.g., [NTA] and [Co], or [Co] and [H]) out of the three product species ([H], [NTA], and [Co]). While the elimination of [HNTA] and [CoNTA] in Decomposition I is straightforward, the elimination of [NTA] and [Co], for example, is not as straightforward. The detailed procedure of decoupling mass action equilibrium equations from kinetic-variable and component equations by eliminating a subset of product species is given in section 3.2.2.

### 3.1.3. Other Advantages of Decomposition

[39] An even more significant advantage is that the decomposition decouples all kinetic rates such that rate formulations/parameters can be determined one reaction at a time, independent of all other kinetic reactions (when parallel kinetic reactions are not present). For example, in equations (6c) or (7c), there is only one kinetic rate on its right hand side. Thus the slope of [B]-versus-time or  $([H] - [NTA] + [Co])$ -



versus-time defines the reaction rate  $R_4$ . For the former case, the kinetic variable is a single species B while for the latter case, the kinetic variable consists of three species. Conceptually, equation (4f) and (6c) are quite different. In equation (4f), [B] denotes the concentration of the species B due to all reactions (in this case just R3) while in equation (6c), [B] denotes the concentration of a kinetic variable (in this case just B) that is measured by a linearly independent kinetic reaction and possibly other linearly dependent kinetic reactions (none in this case). In general, a kinetic variable may consist of more than one species depending on the complexity of the system. The rate of a linearly independent kinetic reaction is measured by a kinetic variable not by a single species. The rate of a kinetic reaction is measured by a single species only when it is the only contributing reaction to the production/consumption of the species. The distinction between kinetic variables and time-variant species concentrations must always be borne in mind when one models geochemical and biochemical processes with this general paradigm.

### 3.2. BIOGEOCHEM Preprocessor

#### 3.2.1. Determines the Number of Linearly Independent Reactions

[40] Let us start with a reaction network that consists of  $N_e$  equilibrium and  $N_k$  kinetic reactions. For a complex system, it will not be obvious to know if all  $N_e$  reactions are linearly independent. Nor is it obvious to know if any of the  $N_k$  kinetic reaction are linearly dependent on only fast reactions. Thus from the above discussion, the first step in the decomposition is to determine the rank, i.e., the number of linearly independent reactions, of the matrix whose columns are made of the stoichiometric coefficients of  $N_e$  equilibrium reactions. The rank of a matrix can be determined with column reduction using any standard matrix operation package [Press et al., 1992]. The reduction is carried out column by column. If no pivot element can be found for a column, the equilibrium reaction corresponding to this column is considered a dependent reaction and is redundant. Suppose the rank of the matrix is  $N_E$ , which should be less than or equal to  $N_e$ , we choose the first  $N_E$  columns in which a pivot element can be found as the  $N_E$  linearly independent reactions. The remaining ( $N_e - N_E$ ) equilibrium reactions are discarded from further consideration in the reaction network. It does not matter which  $N_E$  equations are chosen when all fast reactions are modeled with thermodynamically consistent equilibrium expressions. Second, we determine which of the  $N_k$  kinetic reactions is linearly dependent on only the chosen  $N_E$  equilibrium reactions. To do this, we check the rank of the  $N_E + 1$  matrix (defined as the matrix whose columns are made of the stoichiometric coefficients of  $N_E$  equilibrium reactions and one kinetic reaction). When the rank of this matrix is equal to  $N_E$ , i.e., no pivot element can be found in the ( $N_E + 1$ )th column, the kinetic reaction in the matrix is linearly dependent on only  $N_E$  equilibrium reactions, and this kinetic reaction is removed from the reaction network. Continuing this process one by one for all  $N_k$  kinetic reactions, we are left with  $N_K$  kinetic reactions, which is less than or equal to  $N_k$ .

[41] Starting with  $N_e$  fast reactions and  $N_k$  slow reactions, we end up with  $N_E$  linearly independent equilibrium reactions and  $N_K$  kinetic reactions. Among those  $N_K$  kinetic reactions, some are linearly independent reactions while

some may be linearly dependent on at least one other kinetic reaction. Under such circumstances, only linearly independent kinetic reactions can be segregated. In other words, a subset of the  $N_K$  kinetic reactions (referred to as  $N_{KI}$ ) consists of linearly independent kinetic reactions while the remaining reactions (referred to as  $N_{KD}$  where,  $N_K - N_{KI} = N_{KD}$ ) are linearly dependent kinetic reactions. Of course, the selection of  $N_{KI}$  linearly independent kinetic reactions is not unique. Let us denote  $N_I$  as the number of linearly independent reactions, i.e.,  $N_I = N_E + N_{KI}$ . The remaining task for the BIOGEOCHEM preprocessor is to formally decompose the reaction matrix which is made of  $N_E$  linearly independent equilibrium reactions and  $N_K$  kinetic reactions.

[42] All reactions are initially indexed with 0 in the BIOGEOCHEM preprocessor. During the reduction process, the index of every linearly dependent reactions is overwritten with a nonzero value. A redundant reaction is indexed with 1, an irrelevant reaction with 2, and all other relevant linearly dependent reactions with 3. The indexing array is used to determine which linearly independent reactions the dependent reactions depend on, and is also used to indicate if the dependent is redundant, irrelevant, or relevant. For example, in equation (A4c), R3 depends on R1 (keep in mind, R1 has been already chosen an independent reaction) because R3 and R1 show up in the same equation. Similarly, R3 depends on R2 because both R3 and R2 show up in equation (A4e). Thus R3 depend on both R1 and R2. Because R3 is a slow reaction and it depends on only fast reactions, it is irrelevant and the preprocessor will give us the index of 2 for this reaction.

[43] The decomposition of the reaction matrix is not unique [Westall et al., 1976; Yeh et al., 2000, 2001a]. In order to obtain a decomposition that contains intuitively obvious or recognizable quantities, we observe in the Gauss-Jordan column decomposition that (1) when a row is chosen as a pivot element, its corresponding species is a product species, (2) the species corresponding to the row that has never been chosen as a pivot element is a component species so that one can exert control of components based on one's understanding of the problem, and (3) a linearly dependent reaction will appear only in the rows that contain the linearly independent reaction (each row has one) that this reaction depends on after completion of the decomposition. Based on observations (1) and (2), we choose the row that contains the least number of nonzero entries as the pivot element in any column reduction. Based on observation (3) and because all  $N_E$  equilibrium reactions are linearly independent, an equilibrium reaction appears in only one row and any row will not have more than one equilibrium reaction, if the reduction is performed firstly for  $N_E$  columns corresponding to the  $N_E$  reactions. In this  $N_E$  set, each reduced ODE (not the original ODE) that has one and only one equilibrium reaction is simplified by discarding slow kinetic rates in the equation. The simplified ODE is then replaced by the corresponding mass action equilibrium equation or a user-specified algebraic equation. The number of ODEs replaced is thus equal to  $N_E$ . The remaining ODEs will not have any fast reaction rates in them.

#### 3.2.2. Selects Master and Secondary Species

[44] Before decomposition, users can judiciously select a set of components. After decomposition, the BIOGEOCHEM preprocessor will automatically rectify illegitimate

components that a user may have wrongfully selected, and return a complete set of components ( $N_C = M - N_I$ ). This automatic procedure is performed by removing an incorrectly selected component during a column reduction step. If a pivot element can be found from rows other than those corresponding to user-selected components, then all selected components are still legitimate up to this point. Otherwise, a row corresponding to the selected-component must be chosen as the pivot element and this selected-component must be de-listed from the set of components. All species that have been chosen as the pivot elements form the set of product species (the number of product species is equal to  $N_I$ ). The remaining species make up the set of component species. After completion of the decomposition, the BIOGEOCHEM preprocessor will keep those user-selected components that are legitimate and remove those that are illegitimate. The component species are classified as master species (term used in PHREEQE) or basis species (term used in MINEQL). A product species can be classified as a master or a secondary species. In order to minimize the number of simultaneous equations in biogeochemical modeling, all fast reactions that are modeled with mass action equilibrium equations should be decoupled from the kinetic reactions. In other words, each fast reaction modeled with the mass action equilibrium equation can be used to eliminate one product species from simultaneous consideration. An eliminated product species is termed a secondary species. Since there are  $N_E$  linearly independent equilibrium reactions, there may be up to  $N_E$  secondary species (Note: the number of secondary species will be equal to  $N_E$  if and only if all  $N_E$  reactions are modeled with the mass action equilibrium equations. If some fast reaction are modeled with user-specified algebraic equations or some fast reactions are ion-exchange reactions or precipitation-dissolution reactions, then the number of secondary species is less than  $N_E$  because these algebraic equations and the mass action equilibrium equations for ion-exchange and precipitation-dissolution reactions are not eliminated.) The choice of the set of secondary species among  $N_I$  product species is not unique when at least one of the fast reactions involve more than one product species.

[45] To obtain a suitable set of secondary species, we perform Gauss-Jordan row decomposition of the matrix whose rows are made up of reaction stoichiometries of those fast reactions that are modeled with mass action equilibrium equations. The fast reactions that are modeled with user-specified algebraic equations should not be included in the matrix decomposition for the selection of secondary species, because the inclusion of these reactions will not allow the linear combination of the log  $K^c$  values of fast reactions. Users must specify one species as a master species for this reaction.

[46] The BIOGEOCHEM preprocessor performs the final step of finding and eliminating the secondary species after the components and kinetic variables are defined and after the redundant equilibrium reactions and irrelevant kinetic reactions are removed. The procedure is achieved by diagonalizing the matrix whose rows are made of the reaction stoichiometries of fast reactions (excluding those fast reactions that are modeled with user-specified algebraic equations). The species that will be used to eliminate a mass action equilibrium equation is chosen as a pivot element. Any

species in a fast reaction can be chosen as a pivot element except the component species, the species that have been reserved in user-specified algebraic equations, and the species that have already been found as pivot elements. The procedure of choosing a pivot element is automated in the BIOGEOCHEM preprocessor. In the automation, either the first eligible column or the eligible column that contains the least number of nonzero entries is chosen as the pivot element in any row reduction. If a product species that is chosen as a pivot element is not a precipitated species or an ion-exchanged species, it will be eliminated from simultaneous consideration and thus is a secondary species; otherwise, the species will not be eliminated and remains as a master species and the corresponding mass action equilibrium equation (used to model precipitation-dissolution or ion-exchanged reaction) is not eliminated. Those product species that have not been chosen as pivot elements are master species. The element in the diagonalized matrix that corresponds to the pivoting species is normalized so that its activity can be expressed explicitly as the function of the activities of component species and other nonpivoting species. Meanwhile, the log equilibrium constant of each pivoting species (secondary species) can be determined as the linear combination of the log equilibrium constants of all the original independent equilibrium reactions. The detail of these procedures is given in Appendix A using a simple example.

[47] Careful decisions must be made as to which equilibrium reactions are to be treated as dependent reactions (keep in mind that when some reactions are dependent on each other, the selection of independent reactions versus dependent reactions is not unique) and hence as redundant reactions. When all fast reactions are modeled with mass action equilibrium equations, it does not matter which fast reaction is removed. However, when at least one fast reaction is formulated with a user-specified algebraic equation, one must consider the importance of each fast reaction and remove the least important one. If one wants to exclude certain reactions from the list of candidate eliminations, one simply arranges these among the first  $N_E$  reactions. This is the point that a user must be aware of and guard against.

### 3.2.3. Generates Three Subsets of Governing Equations

[48] With the above discussion, equation (3) is decomposed into the following equation via the Gauss-Jordan elimination [Chilakapati, 1995; Steefel and MacQuarrie, 1996; Yeh *et al.*, 2001a]

$$\mathbf{B} \frac{d\mathbf{C}}{dt} = \begin{bmatrix} \mathbf{D} & \mathbf{K} \\ \mathbf{0}_1 & \mathbf{0}_2 \end{bmatrix} \mathbf{R} \quad (8)$$

where  $\mathbf{B}$  is the reduced unit matrix,  $\mathbf{D}$  is the diagonal matrix representing a submatrix of the reduced reaction-matrix with size of  $N_I \times N_I$  reflecting the effects of  $N_I$  linearly independent reactions on the production-consumption rate of all kinetic-variables,  $\mathbf{K}$  is a submatrix of the reduced reaction-matrix with size of  $N_I \times N_{KD}$  reflecting the effects of  $N_{KD}$  dependent kinetic reactions,  $\mathbf{0}_1$  is a zero matrix representing a submatrix of the reduced reaction-matrix with size  $N_C \times N_I$ , and  $\mathbf{0}_2$  is a zero matrix representing a submatrix of the reduced reaction-matrix with size  $N_C \times N_{KD}$ . An example illustrating this form of the equation is given for Decomposition I in Appendix A.

[49] The decomposition of equation (3) to equation (8) effectively reduces a set of  $M$  simultaneous ODEs into three subsets of equations: the first contains  $N_E$  infinite-rate equations, each represents an equilibrium reaction that can be formulated with either a mass action equilibrium equation or a user-specified algebraic equation (see section 4); the second contains  $(N_I - N_E)$  simultaneous ODEs representing the rate of change of the kinetic variables; and, the third contains  $N_C$  linear algebraic equations representing mass conservation of the chemical components. These equation subsets are defined as

Infinite-rate equations for  $N_E$  equilibrium reactions

$$\frac{dE_i}{dt} = D_{kk}R_k + \sum_{j \in N_{KD(k)}} D_{ij}R_j, k \in N_E, i \in M: \frac{d}{dt} \left( \frac{E_i}{D_{kk}} \right) \approx R_k = \infty$$

$\exists$  a thermodynamically consistent equation (9)

Governing equations for  $(N_I - N_E)$  kinetic variables

$$\frac{dE_i}{dt} = D_{kk}R_k + \sum_{j \in N_{KD(k)}} D_{ij}R_j, k \in N_{KI}, i \in M \quad (10)$$

Mass conservation equations for  $N_C$  chemical components

$$\frac{dE_i}{dt} = 0; \text{ denoting } T_j = E_i, j \in N_C, i \in M \quad (11)$$

where  $E_i$  is the linear combination of species concentration resulting from the matrix decomposition and  $N_{KD(k)}$  is the subset of linearly dependent kinetic reactions, which depends on the  $k$ th linearly independent reaction. For this generic system, the variable,  $E_i$ , in equation (9) is called an equilibrium variable, the variable  $E_i$  in equation (10) is called a kinetic variable, and the variable  $E_i$  in equation (11) is called a component variable and is normally denoted with  $T_j$ . It should be emphasized here that only linearly independent reactions can be segregated, not all reactions.

[50] The unique feature of the diagonalization is that there is only one linearly independent kinetic reaction appearing on the right-hand side of equation (10) when parallel kinetic reactions are not present. The significance of this unique feature is that all linearly independent kinetic reactions are segregated, thus it enables the formulation and parameterization of linearly independent reaction rates one reaction at a time, independent of all other kinetic reactions; if linearly dependent kinetic reactions are not present in the system. When an experiment is conducted to study kinetics, it must be controlled such that  $N_{KD(k)} = 0$ . Otherwise, kinetic reactions cannot be completely segregated [Yeh *et al.*, 2000a] and the independent formulation of kinetic reaction rates cannot be achieved [Burgos *et al.*, 2002, 2003]. Without a careful design of experiments to exclude linearly dependent kinetic reactions, only lumped kinetic rates can be formulated and characterized [Burgos *et al.*, 2002, 2003].

#### 4. Reaction Rate Formulations

[51] A rate equation must be specified in order to quantitatively describe a general biogeochemical reaction. A general reaction can be written as



where  $G_i$  is the chemical formula of the  $i$ th species involved in  $k$  reactions. For an elementary kinetic reaction the rate law is given by collision theory [Smith, 1981; Atkins, 1986] as

$$R_k = \left( k_k^f \prod_{i=1}^M (C_i)^{\mu_{ik}} - k_k^b \prod_{i=1}^M (C_i)^{\nu_{ik}} \right), k \in N_K \quad (13)$$

in which

$$k_k^f = K_k^f \prod_{i=1}^M (\gamma_i)^{\mu_{ik}} \text{ and } k_k^b = K_k^b \prod_{i=1}^M (\gamma_i)^{\nu_{ik}} \quad (14)$$

where  $R_k$  is the reaction rate,  $k_k^f$  is the concentration-based forward rate constant, and  $k_k^b$  is the concentration-based backward rate constant of the  $k$ th kinetic reaction,  $C_i$  is the concentration of the  $i$ th species,  $K_k^f$  is the activity-based forward rate constant,  $\gamma_i$  is the activity coefficient of the  $i$ th species, and  $K_k^b$  is the activity-based backward rate constant. When a kinetic reaction cannot be modeled with an elementary rate, it may be formulated based on either empirical or mechanistic approaches [Steeffel and van Cappellen, 1998; Yeh *et al.*, 2001a]. To make BIOGEOCHEM completely general, for any nonelementary kinetic reaction, its rate may be user-specified as follows:

$$R_k = R_k(C_1, C_2, \dots, C_M; k_1, k_2, \dots) \quad (15)$$

where  $R_k$  is the prescribed rate law by the user written as a function of the concentrations of species participating in the reaction and a number of parameters;  $C_i$  is the concentration of the  $i$ th species and  $k_1, k_2, \dots$  are rate parameters used to fit experimental data. Equation (15) looks like equation (2) at first glance, but they render different meanings. In an ad hoc approach, equation (2) represents the empirical rate of change of a species due to all reactions. In a reaction-based approach, equation (15) represents the empirical rate (if one cannot come up with a pathway) or mechanistically derived rate (if one can come up with a pathway, for example enzymatic kinetics) of a single reaction. Hence, for a reaction-based model, even when an empirical rate formulation is used, the formulation is theoretically descriptive of the specific chemical reaction and therefore may be applicable to a wider range of environmental conditions. Only when the rate of a species is due to only one reaction are equations (2) and (15) conceptually the same.

[52] If the reaction is a fast reaction, it is assumed that the reaction instantaneously reaches equilibrium and its rate is mathematically abstracted as infinity stating a thermodynamically consistent expression, for example, the mass action equilibrium equation as

$$\frac{dE_i}{dt} = D_{kk}R_k + \sum_{j \in N_{KD(k)}} D_{ij}R_j, k \in N_E, i \in M: \frac{dE_i}{dt} \approx D_{kk}R_k = \infty$$

$$\exists K_k^e = \frac{\left( \prod_{i=1}^M (A_i)^{\nu_{ik}} \right)}{\left( \prod_{i=1}^M (A_i)^{\mu_{ik}} \right)} \quad (16)$$

where  $K_k^e$  is the equilibrium constant of the  $k$ th reaction and  $A_i$  is the activity of the  $i$ th species. For precipitation/



dissolution reactions, the activity of solid species is assumed to be constant and equal to unity. The mass action equilibrium equation can be simply treated as a thermodynamic hypothesis [Goldberg, 1991]. It does not have to be conceived as the kinetic rate expression that follows mass action kinetics as described with the following equation

$$R_k = - \lim_{K_k^b \rightarrow \infty} K_k^b \cdot \left( \prod_{i=1}^M (A_i)^{\nu_{ik}} - K_k^e \prod_{i=1}^M (A_i)^{\mu_{ik}} \right); K_k^e = \frac{K_k^f}{K_k^b} = \text{fixed constant} \quad (17)$$

[53] Conceptually, the use of equations (16) and (17) to describe an equilibrium reaction are quite different. Equation (16) states that the rate of a fast reaction is infinity as a mathematical abstraction (Recall this is justified in the paragraph under section 2) and its reactants and products satisfy the mass action equilibrium equation instantaneously. The last equation in equation (16) is not a rate formulation because the concept of finite rate does not exist in thermodynamics. On the other hand, equation (17) states that the rate of a fast reaction is formulated as an asymptote of the mass action kinetics (a kinetic reaction modeled with an elementary rate expression with both forward and backward rate constants infinitely large while keeping their ratio at a fixed value, i.e., the equilibrium constant). (Keep in mind equation (16) is called mass action equilibrium equation while equation (17) is called mass action kinetic equation.) The rate of the fast kinetic reaction is indeterminate since its reactants and products satisfy the mass action equilibrium equation, equation (16), at equilibrium. Even at the slightest deviation from the equilibrium, the rate as defined by equation (17) is infinity. This is precisely the reason that the fast reaction rate should not be formulated as the asymptotic approximation of a mass action kinetics because it would make the set of ODEs (equations (4a) to (4f) for the example given in section 2) infinitely stiff. Rather, it is better to consider the rate of a fast reaction infinity as a mathematical abstraction (i.e., equation (16)). The mathematical abstraction serves a dual function. First, it facilitates the comparison of the magnitude of various reaction rates appearing in the reduced set of ODEs. Second, it states that the thermodynamic approach is taken to model the equilibrium reactions using either a mass action equilibrium equations or a user-specified algebraic equations.

[54] To make BIOGEOCHEM completely general, a user-specified algebraic equation can be used to describe a fast reaction when it cannot be modeled with a mass action equilibrium equation

$$\frac{dE_i}{dt} = D_{kk}R_k + \sum_{j \in N_{KD(k)}} D_{ij}R_j, k \in N_E, i \in M: \quad (18)$$

$$\frac{dE_i}{dt} \approx D_{kk}R_k = \infty \exists F_k(C_1, C_2, \dots, C_M; p_1, p_2, \dots) = 0$$

where  $F_k$  is an implicit function of species concentrations with a number of parameters. For example, the linear ( $K_d$  approach) and nonlinear (Freundlich) isotherms describing

heterogeneous reactions with partitioning between aqueous and adsorbed chemicals fall into this category.

## 5. Numerical Solution

[55] After the reaction matrix is decomposed in BIOGEOCHEM, equations of the type of equation (16) are manipulated so that only one secondary species can appear in any of mass action equilibrium equation to facilitate computation (i.e., equations of the type of equation (16) can be easily eliminated). On the other hand, user-appointed species in equations of the type of equation (18) cannot be easily eliminated in general. Thus for computational efficiency in the Newton-Raphson technique [Westall *et al.*, 1976; Yeh *et al.*, 1995], the numbers of equations are kept to a minimum in BIOGEOCHEM, involving one set of linear algebraic equations (mass conservation equations), one set of nonlinear ordinary differential equations (for kinetic variables), one set of user-specified algebraic equations (for the fast reactions that cannot be modeled with mass action equilibrium equations), one set of nonlinear algebraic equations for equilibrium ion exchanged reactions, and one set of nonlinear algebraic equations for equilibrium precipitation reactions. Full pivoting is used to solve the matrix equation resulting from the Newton-Raphson iteration [Yeh *et al.*, 1998].

[56] The governing set of linear algebraic equations, nonlinear algebraic equations, and nonlinear ODEs are solved with a significantly modified version of the code BIOKEMOD [Salvage and Yeh, 1998]. An iteration loop of basis switching is incorporated within a Newton-Raphson iteration loop to improve the robustness and efficiency of computation. The idea of basis switching can be found in the works by Westall *et al.* [1976] and Steefel and MacQuarrie [1996].

[57] Since the concentration of some species may be several orders of magnitude smaller than others in a component equation, mass nonbalance or nonconvergence may occur for the component. In other words, the computed species concentrations that compose the mole balance equation cannot add up to the given total component concentration to within the error tolerance. This normally occurs when the concentration of a component species is very low and the concentration of a secondary species that is a member of the component is very high. The component species is explicitly included in solving the mass balance equation. Because of its low concentration, it may still satisfy the mass balance to within the error tolerance even though its percentage error may not be very small. On the other hand, the concentration of the secondary species is calculated using the simulated component species (and possibly other master species) concentrations. When the reduced log K value of the secondary species is large and the percentage error in the concentration of the component species is not that small, the calculated concentration of the secondary species may have a significant absolute error that is much larger than the concentrations of the component species and other master species consisting of the component. This error naturally causes a significant mass balance error. To resolve this problem, basis switching is used to swap the low-concentration component species with the high-concentration secondary species (i.e., the component species become a new secondary species and the secondary

species becomes a new component species). With this species switching, the concentrations of the new component species can be expected to be much higher than that of the new secondary species. As a result, the contribution to the mass balance is mainly due to the new component species. Since the new component species concentration is simulated to explicitly satisfy the mass balance equation, it will not cause much mass balance error. The mass balance error is due mainly to the new secondary species. However, because the concentration of the new secondary species is much smaller than that of the new component species, even a large percentage error in its concentration will not cause much mass balance error. Therefore the mass balance is likely satisfied with the new component and secondary species.

[58] The mathematical algorithm of basis switching is summarized as follows. From mass action equilibrium equations, we have

$$x_i = \alpha_i \prod_j c_j^{a_{ij}} \quad (19)$$

where  $c_j$  is the concentration of the  $j$ th component or species,  $x_i$  is the concentration of the  $i$ th secondary species,  $a_{ij}$  is the stoichiometric coefficient of the  $j$ th species in the  $i$ th mass action equilibrium equation and  $\alpha_i$  given by

$$\alpha_i = K_i^e \frac{\prod_j \gamma_j^{a_{ij}}}{\gamma_i} \quad (20)$$

is the modified stability constant.

[59] Suppose  $c_m$  is the component to be switched with  $x_k$ , then from equation (19), for  $i = k$

$$x_k = \alpha_k \prod_{j \neq m} c_j^{a_{kj}} c_m^{a_{km}} \quad (21)$$

from which we can obtain

$$c_m = x_k^{\frac{1}{a_{km}}} \left[ \alpha_k \prod_{j \neq m} c_j^{a_{kj}} \right]^{-\frac{1}{a_{km}}} \quad (22)$$

Substituting equation (22) into equation (19) for  $i \neq k$ , we have

$$x_i = \alpha_i \alpha_k^{-\frac{a_{im}}{a_{km}}} \prod_{j \neq m} c_j^{a_{ij} - \frac{a_{kj} a_{im}}{a_{km}}} x_k^{\frac{a_{im}}{a_{km}}} \quad (23)$$

Equations (22) and (23) can be described in a general form as

$$x'_i = \alpha'_i \prod_j c_j^{a'_{ij}} \quad (24)$$

in which for  $i = k$ :

$$\begin{aligned} x'_i &= c_m; \alpha'_i = \alpha'_k = \alpha_k^{-\frac{1}{a_{km}}} = \frac{1}{\alpha_k^{a_{km}}}; \\ a'_{ij} &= a'_{kj} = \frac{-a_{kj}}{a_{km}} = -a_{kj} a'_{km} \text{ and } c'_j = c_j \text{ if } j \neq m; \\ a'_{ij} &= a'_{im} = \frac{1}{a_{km}} \text{ and } c'_j = x_k \text{ if } j = m \end{aligned} \quad (25)$$

for  $i \neq k$ :

$$\begin{aligned} x'_i &= x_i; \alpha'_i = \alpha_i \alpha_k^{-\frac{a_{im}}{a_{km}}} = \alpha_i (\alpha'_k)^{a_{im}}; \\ a'_{ij} &= a_{ij} - \frac{a_{kj} a_{im}}{a_{km}} = a_{ij} + a_{im} a'_{kj} \text{ and } c'_j = c_j \text{ if } j \neq m; \\ a'_{ij} &= a'_{im} = \frac{a_{im}}{a_{km}} = a_{im} a'_{km} \text{ and } c'_j = x_k \text{ if } j = m \end{aligned} \quad (26)$$

[60] After basis switching, the same calculation is carried out with the new components and secondary species. Although the choice of component species is not unique, any choice is equivalent; thus results for the system will not be changed after basis switching. By switching one species at each time, BIOGEOCHEM can switch as many times as needed. In order to keep the system consistent, BIOGEOCHEM performs forward and backward basis switching. After backward basis switching, the components are the same as the original ones. Because the components must be consistent in hydrologic transport (i.e., once the set of components is chosen, the same set must be used throughout transport simulations), the backward basis switching is necessary when the reaction chemical module is coupled with a hydrologic transport model.

## 6. Example Simulations

[61] A total of three example problems are employed to demonstrate the application of BIOGEOCHEM. The first example is used to partially verify the model using a comprehensive reactive-chemical and biodegradation problem that was modeled by others [Chilakapati *et al.*, 1998]. In addition, this problem involves nonlinear Monod kinetics as well as nonlinear elementary kinetics. A successful simulation of this problem would demonstrate the capability of the model to address various types of nonlinear cases. The second example is used to demonstrate the capability of BIOGEOCHEM to simulate generic reactive chemical problems involving mixed fast/equilibrium and slow/kinetic reactions of simultaneous geochemical processes including aqueous complexation, adsorption-desorption, ion exchange, and precipitation-dissolution. The third problem is used to exemplify the need for basis species switching to deal with problems that are difficult to converge.

### 6.1. Mixed Microbiological and Chemical Kinetics

[62] This case has been studied by Chilakapati *et al.* [1998]. The main purpose of the case was to study biogeochemical reactions that affect the transformation of Co(II) ethylenediaminetetraacetic acid (EDTA) complexes. The reaction network is listed in Table 1. It was simplified by Chilakapati *et al.* [1998] from the full reaction network of over 64 reactions studied by Szecsody *et al.* [1994]. Reactions R1–R5 are fast adsorption/desorption reactions. R6 and R7 represent iron dissolution as a two-step ligand promoted process. R8 is an oxidation reaction. R9 and R10 are biodegradation reactions, where Fe(III) EDTA and EDTA are the electron donors and O<sub>2</sub> is the terminal electron acceptor. Some of the reactants and products are not included in reactions R7 and R8 because these species are not the contributing factors in the simulation.

[63] Included in this simulation are a total of 15 species ( $M = 15$ ), 5 equilibrium reactions ( $N_e = 5$ ) and 5 kinetic reactions ( $N_k = 5$ ). According to Chilakapati *et al.* [1998],

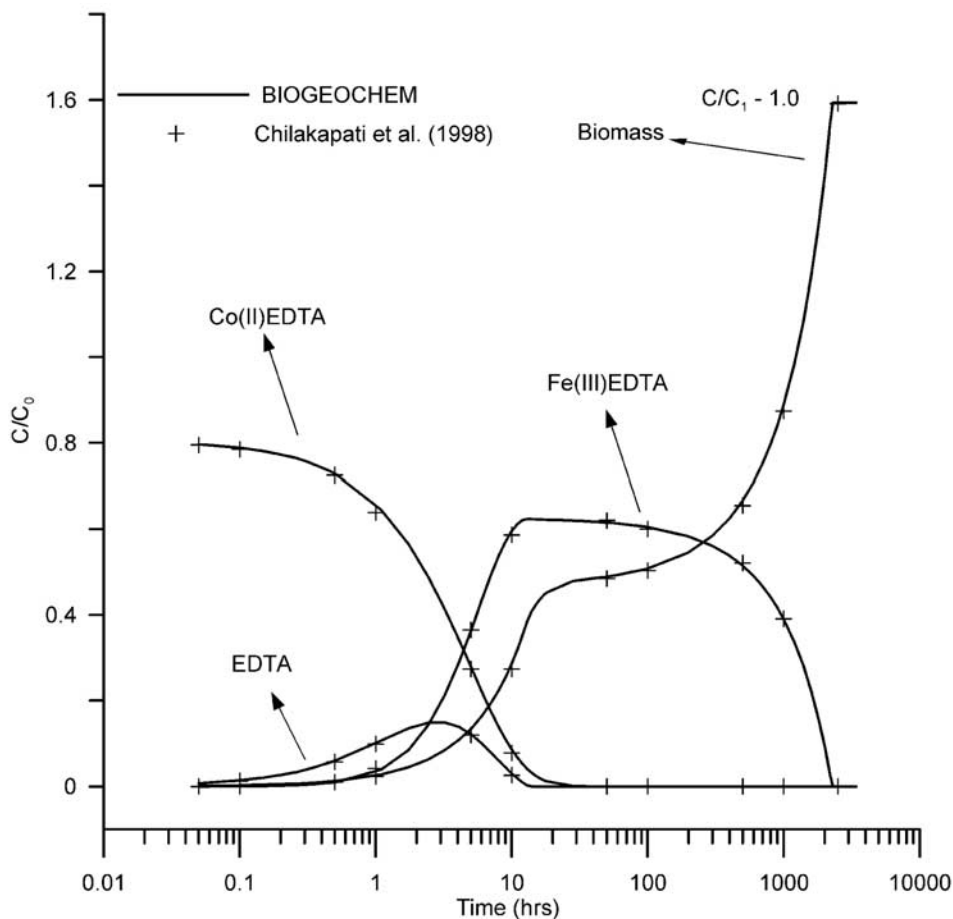
**Table 1.** Reaction Network for the Example Problem in Section 6.1<sup>a</sup>

Reaction	Reaction Number	Reaction Constants
$\text{Co(II)(aq)} + \text{S}_{\text{neg}} \rightleftharpoons \text{S}_{\text{neg-Co}}$	(R1)	$K_1^c = 12.0$
$\text{Co(II)EDTA(aq)} + \text{S}_{\text{pos}} \rightleftharpoons \text{S}_{\text{pos-Co(II)EDTA}}$	(R2)	$K_2^c = 25.0$
$\text{Fe(III)EDTA(aq)} + \text{S}_{\text{pos}} \rightleftharpoons \text{S}_{\text{pos-Fe(III)EDTA}}$	(R3)	$K_3^c = 9.0$
$\text{EDTA(aq)} + \text{S}_{\text{pos}} \rightleftharpoons \text{S}_{\text{pos-EDTA}}$	(R4)	$K_4^c = 25.0$
$\text{Co(III)EDTA(aq)} + \text{S}_{\text{pos}} \rightleftharpoons \text{S}_{\text{pos-Co(III)EDTA}}$	(R5)	$K_5^c = 2.5$
$\text{S}_{\text{pos-Co(II)EDTA}} \leftrightarrow \text{Co(II)(aq)} + \text{S}_{\text{pos-EDTA}}$	(R6)	$k_6^f = 1.0 \text{ h}^{-1}, k_6^b = 1.0 \times 10^{-3} \text{ L mM}^{-1}\text{h}^{-1}$
$\text{S}_{\text{pos-EDTA}} \leftrightarrow \text{Fe(III)EDTA(aq)} + \text{S}_{\text{pos}}$	(R7)	$k_7^f = 2.5 \text{ h}^{-1}, k_7^b = 0.0 \text{ L mM}^{-1}\text{h}^{-1}$
$\text{Co(II)EDTA(aq)} \leftrightarrow \text{Co(III)EDTA(aq)}$	(R8)	$k_8^f = 1.0 \times 10^{-3} \text{ h}^{-1}, k_8^b = 0.0 \text{ h}^{-1}$
$\text{Fe(III)EDTA(aq)} + 6\text{O}_2 \rightarrow 3\text{CO}_2 + \text{Biomass}$	(R9)	user specified reaction rate as expressed by $R_9$ in equation (B9) in section B1.3, where $\mu_1 = 2.5 \times 10^{-4} \text{ h}^{-1}, k_{11} = 1.0 \times 10^{-5} \text{ mM L}^{-1},$ and $k_{21} = 1.0 \times 10^{-5} \text{ mM L}^{-1}$
$\text{EDTA(aq)} + 6\text{O}_2 \rightarrow 3\text{CO}_2 + \text{Biomass}$	(R10)	user specified reaction rate as expressed by $R_{10}$ in equation (B9) in section B1.3, where $\mu_2 = 0.025 \text{ h}^{-1}, k_{12} = 1.0 \times 10^{-5} \text{ mM L}^{-1},$ and $k_{22} = 1.0 \times 10^{-5} \text{ mM L}^{-1}$

<sup>a</sup>After Chilakapati et al. [1998].

species initially present in the system are Co(II)EDTA (0.032 mM), dissolved O<sub>2</sub> (0.256 mM), microorganisms (0.02 mM), the charged surface site S<sub>pos</sub> (0.016 mM) and S<sub>neg</sub> (0.0011 mM). If we suppose all the reactions are linearly independent, then there must be 5 components. However, the BIOGEOCHEM preprocessor determined there are 6 components. In other words, there are only 9 linearly independent reactions, i.e., the rank of the reaction matrix turned out to be 9. Because all five equilibrium reactions are linearly independent, none

of these fast reactions is redundant. The decomposition indicated that none of the five slow reactions depends on only fast reactions, thus all kinetic reactions are relevant to the system. The decomposition also revealed that one of the three kinetic reactions (R7, R9, and R10) can be considered linearly dependent on the other two and possibly other fast reactions. Without loss of generality, we consider R10 a linearly dependent reaction. The decomposition revealed that R10 depends on R4 (a fast reaction), R7 and R9. R10 is



**Figure 1.** Solution of the reaction network in Table 1: Comparison of this study with results of Chilakapati et al. [1998]. C<sub>0</sub> and C<sub>1</sub> are the initial concentrations of Co(II)EDTA(aq) and microorganisms, respectively.



**Table 2.** Reaction Network for the Example Problem in Section 6.2

Reaction	Reaction Number	Reaction Constants
<i>Mineral Dissolution and Surface Site Formation Reactions</i>		
$M \rightleftharpoons C_1 - 3C_2$	(R1)	$k_1^f = 0.05$
$M \rightleftharpoons S_1$	(R2)	User specified partition between bulk and surface metal ions as expressed by equation (B16) in Appendix B.
<i>Aqueous Complexation Reactions</i>		
$C_3 \rightleftharpoons C_4 + C_5$	(R3)	$\text{Log } k_3^f = 2.03, \text{Log } k_3^b = 20.00$
$C_6 + C_5 \rightleftharpoons C_7$	(R4)	$\text{Log } K_4^c = 12.32$
$C_2 + C_5 + C_6 \rightleftharpoons C_8$	(R5)	$\text{Log } K_5^c = 15.93$
$C_6 \rightleftharpoons C_2 + C_9$	(R6)	$\text{Log } K_6^c = -12.60$
$C_1 + C_5 \rightleftharpoons C_{10}$	(R7)	$\text{Log } k_7^f = 25.00, \text{Log } k_7^b = -2.57$
$C_1 + C_2 + C_5 \rightleftharpoons C_{11}$	(R8)	$\text{Log } K_8^c = 29.08$
$C_1 + C_5 \rightleftharpoons C_2 + C_{12}$	(R9)	$\text{Log } K_9^c = 19.65$
$C_1 + C_5 \rightleftharpoons 2C_2 + C_{13}$	(R10)	$\text{Log } K_{10}^c = -36.30$
$C_1 \rightleftharpoons C_2 + C_{14}$	(R11)	$\text{Log } K_{11}^c = -2.19$
$C_1 \rightleftharpoons 2C_2 + C_{15}$	(R12)	$\text{Log } K_{12}^c = -5.67$
$C_1 \rightleftharpoons 3C_2 + C_{16}$	(R13)	$\text{Log } K_{13}^c = -13.60$
$C_1 \rightleftharpoons 4C_2 + C_{17}$	(R14)	$\text{Log } K_{14}^c = -21.60$
$2C_1 \rightleftharpoons 2C_2 + C_{18}$	(R15)	$\text{Log } K_{15}^c = -2.95$
$C_2 + C_4 + C_5 \rightleftharpoons C_{19}$	(R16)	$\text{Log } K_{16}^c = 21.40$
$C_4 \rightleftharpoons C_2 + C_{20}$	(R17)	$\text{Log } K_{17}^c = -9.67$
$C_4 \rightleftharpoons 2C_2 + C_{21}$	(R18)	$\text{Log } K_{18}^c = -18.76$
$C_4 \rightleftharpoons 3C_2 + C_{22}$	(R19)	$\text{Log } K_{19}^c = -32.23$
$C_2 + C_5 \rightleftharpoons C_{23}$	(R20)	$\text{Log } K_{20}^c = 11.03$
$2C_2 + C_5 \rightleftharpoons C_{24}$	(R21)	$\text{Log } K_{21}^c = 17.78$
$3C_2 + C_5 \rightleftharpoons C_{25}$	(R22)	$\text{Log } K_{22}^c = 20.89$
$4C_2 + C_5 \rightleftharpoons C_{26}$	(R23)	$\text{Log } K_{23}^c = 23.10$
$C_2 + C_{27} \rightleftharpoons C_{28}$	(R24)	$\text{Log } K_{24}^c = 14.00$
<i>Adsorption-Desorption Reactions</i>		
$S_1 \rightleftharpoons S_2 + C_2$	(R25)	$\text{Log } K_{25}^c = -11.60$
$S_1 + C_2 \rightleftharpoons S_3$	(R26)	$\text{Log } K_{26}^c = 5.60$
$S_1 + 3C_2 + C_5 \rightleftharpoons S_4$	(R27)	$\text{Log } K_{27}^c = 30.48$
$S_1 + C_1 + C_2 + C_5 \rightleftharpoons S_5$	(R28)	$\text{Log } k_{28}^f = 40.00, \text{Log } k_{28}^b = 2.37$
$S_1 + C_2 + C_4 + C_5 \rightleftharpoons S_6$	(R29)	$\text{Log } k_{29}^c = 30.00, \text{Log } k_{29}^b = 1.51$
$S_1 - C_2 + C_4 \rightleftharpoons S_7$	(R30)	$\text{Log } k_{30}^f = -0.99, \text{Log } k_{30}^b = 1.70$
$S_1 + C_2 + C_5 + C_6 \rightleftharpoons S_8$	(R31)	$\text{Log } k_{31}^b = 1.19, \text{Log } k_{31}^f = 25.0$
<i>Ion Exchange Reactions</i>		
$C_{29} + 2 \text{ site-}C_{30} \rightleftharpoons \text{site-}C_{29} + 2C_{30}$	(R32)	$\text{Log } k_{32}^b = -0.5, \text{Log } k_{32}^f = -0.75$
$C_6 + 2 \text{ site-}C_{30} \rightleftharpoons \text{site-}C_6 + 2C_{30}$	(R33)	$\text{Log } K_{33}^c = 0.6$

relevant because it depends on at least one of the kinetic reactions (two in this case). Had R10 depended on only fast reactions, it would be irrelevant. An example of matrix decomposition of the reaction network is shown in Appendix B, section B1.

[64] Comparison of the simulation results obtained with the present study (solid line) and with *Chilakapati et al.* [1998] for the steady disappearance of Co(II)EDTA, the production/consumption of EDTA and Fe(III)EDTA, and the growth of microorganisms are shown in Figure 1. These results demonstrate that the simulations calculated using BIOGEOCHEM were essentially identical to the simulation calculated by *Chilakapati et al.* [1998]. Identical model simulations serve to partially verify the performance and accuracy of BIOGEOCHEM. For description and discussion of the experimental data see *Chilakapati et al.* [1998].

## 6.2. Complexation, Adsorption, Ion Exchange, and Dissolution in a System of Mixed Equilibrium and Kinetic Reactions

[65] This example demonstrates the generic flexibility of BIOGEOCHEM. It is a fictitious system involving aqueous complexation, adsorption, ion exchange, and mineral dissolution reactions. Each type of reaction includes both fast/

equilibrium and slow/kinetic reactions. Table 2 lists the conceptualized reaction network. Initially, species  $C_3$  (an organic complex) is in contact with mineral  $M$ , which is a metal-hydroxide. Through dissolution a portion of the mineral becomes solubilized (Reaction R1 in Table 2); the dissolution is assumed to be an irreversible reaction. Adsorbing sites are formed on the surface of the mineral  $M$  (R2). The surface site commonly undergoes ionization reactions upon contact with water (R25 and R26). Species  $C_3$  dissociates into  $C_4$  and  $C_5$  (R3 in Table 2). Species  $C_4$ ,  $C_5$ ,  $C_6$  and dissolved species  $C_1$  react to form various complexed species (R4 through R23). In turn, some of these complexed species are adsorbed onto the surface of the mineral  $M$  (R27 through R31). The system also includes an ion exchange site and three ions ( $C_6$ ,  $C_{29}$  and  $C_{30}$ ) compete for the site (R32 and R33). The chemical system is quite complex, although some of the reactions are made up in order to demonstrate the generality of BIOGEOCHEM. For example, R2 (partition between bulk and surface metal ions) is used to reflect the mass balance of surface sites. Some of the reaction constants are obtained from *Szecsody et al.* [1994] and *Yeh et al.* [1995].

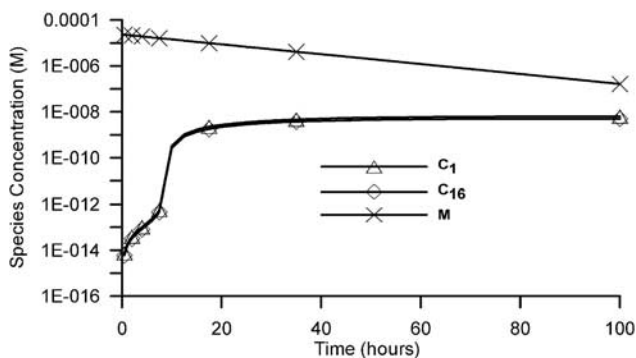
[66] The slow reactions were assumed to be the dissolution of the mineral, and the adsorption of  $C_5$ -complexes and

C<sub>4</sub> (R1 and R28–31, respectively). Sorbed species were modeled using simple surface complexation approaches built into BIOGEOCHEM. To further demonstrate the capability of BIOGEOCHEM to model kinetic aqueous complexation and ion exchange reactions, R3, R7 and R32 are treated as kinetic reactions.

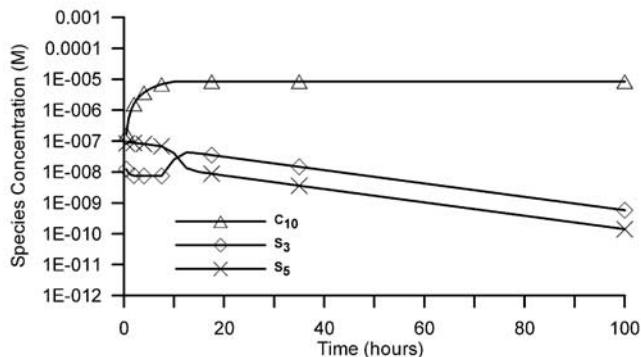
[67] From the above conceptualization of the reaction network, this simulation includes a total of 41 species (M = 41), excluding the species C<sub>28</sub> because its activity is assumed 1.0, 25 equilibrium reactions (N<sub>e</sub> = 25) and 8 kinetic reactions (N<sub>k</sub> = 8). Initial concentrations of C<sub>6</sub>, M, C<sub>3</sub>, and C<sub>30</sub> are  $2 \times 10^{-3}$  M,  $2.36 \times 10^{-5}$  M,  $8.51 \times 10^{-6}$  M, and 0.1552 M, respectively; and initial concentrations of site-C<sub>30</sub>, site-C<sub>29</sub>, and site-C<sub>6</sub> are 0.0651 M, 0.1463 M, and 0.1562 M, respectively. All other species are initially 0.0 M except for C<sub>2</sub>, whose concentration is fixed at  $3.16 \times 10^{-5}$  M. The BIOGEOCHEM preprocessor indicated that all reactions in this system are linearly independent, i.e., N<sub>I</sub> = 33, N<sub>E</sub> = N<sub>e</sub> = 25, N<sub>KI</sub> = N<sub>K</sub> = N<sub>k</sub> = 8. In other words, none of the 25 equilibrium reactions nor any of the 8 kinetic reactions are removed from consideration. Thus a formal matrix decomposition of the reaction network should yield 24 mass action equilibrium equations and one user-specified algebraic equation (25 total), 8 kinetic variable equations, and 8 mass conservation equations (N<sub>C</sub> = M - N<sub>I</sub> = 41 - 33). An example of one decomposition of the reaction matrix presented in Table 2 is shown in Appendix B, section B2.

[68] The 24 mass action equilibrium equations are decoupled from the other 17 equations (8 mass conservation, 8 kinetic-variable, and one user-specified algebraic equation). In other words, 24 equations are substituted into 17 equations to eliminate 24 secondary variables (C<sub>7</sub>, C<sub>8</sub>, C<sub>9</sub>, C<sub>11</sub> through C<sub>27</sub>, S<sub>2</sub>, S<sub>3</sub>, S<sub>4</sub>, and site-C<sub>30</sub>). The resulting 17 equations are then solved simultaneously for 17 master variables (C<sub>1</sub> through C<sub>6</sub>, C<sub>10</sub>, C<sub>28</sub>, C<sub>29</sub>, C<sub>30</sub>, S<sub>1</sub>, S<sub>5</sub> through S<sub>8</sub>, M, and site-C<sub>29</sub>).

[69] The time-variant dissolution of the mineral is described by the concentration variations of six species (Figures 2 and 3). C<sub>1</sub>, C<sub>16</sub> and M in Figure 2 represent the species that are directly involved in the dissolution according to the reaction network. The formation of species C<sub>10</sub> dominates the aqueous speciation in the system and is ultimately the driving force for mineral dissolution. The last two species in Figure 3 represent the two main adsorbed species in the system. The simulation of the interaction of these six species demonstrates the ability of BIOGEO-



**Figure 2.** Concentration curves for major species directly involving in the mineral dissolution.



**Figure 3.** Concentration curves for main complexed and adsorbed species.

CHEM to model a complex (multiple reaction types) mixed (equilibrium and kinetic) reaction system.

[70] For the first 10 hours, there is a decline in the mineral concentration (Figure 2) which corresponds to an equivalent increase in the concentration of C<sub>10</sub> (Figure 3). Dissolved species C<sub>1</sub> quickly combines with C<sub>5</sub> to form the complexed species. This continues until C<sub>10</sub> begins to reach its equilibrium value (at around 15 h) due to the limiting of total concentration of C<sub>5</sub> (Figure 3). At this time, the aqueous species C<sub>1</sub> and C<sub>16</sub> that are formed no longer combine directly into the complexed species but remain “free” and increase to their approximate equilibrium values (Figure 2).

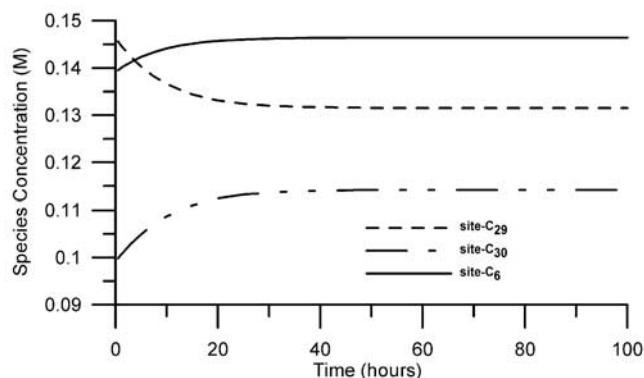
[71] The two main sorbed species, S<sub>3</sub> and S<sub>5</sub>, also undergo considerable changes during the first 20 h. Initially S<sub>5</sub> is the main adsorbed species (Figure 3). But when the production of C<sub>10</sub> begins to be limited by the available free C<sub>5</sub>, its high production rate drives its continued formation by scavenging other sources of C<sub>5</sub> species. As a result, the adsorbed species (S<sub>5</sub>) begins to decrease to form the aqueous complexed species (C<sub>10</sub>) after ca. 12 h. S<sub>3</sub> increases as a result for approximately the next 3 h. After that time, both species concentrations decrease because the available sites decrease due to the continuous dissolution of M. Given sufficient time, M will completely dissolve as indicated in Figure 2 because of the assumption that it is a irreversible reaction.

[72] Reactions R32 and R33 are weakly coupled to the rest of the system because the amounts of C<sub>6</sub> and its complexed species are insignificant. About 10% of site-C<sub>6</sub> was exchanged immediately into the solution because of the equilibrium reaction R33. There are significant amounts of site-C<sub>29</sub> and C<sub>30</sub> in the system, R32 proceeds from right to left till it reaches equilibrium as shown in Figure 4.

### 6.3. Basis Switching

[73] This example demonstrates the need for basis switching when component concentrations become very low and cause mass balance errors. The problem is defined by 20 aqueous complexation reactions and four precipitation-dissolution reactions involving 27 aqueous species and 4 precipitated species (reaction network presented in Table 3).

[74] Based on this reaction network there are a total of 31 species (M = 31), excluding the species H<sub>2</sub>O(l) because its activity is assumed 1.0, and 24 equilibrium reactions (N<sub>e</sub> = 24). The BIOGEOCHEM preprocessor indicated that all 24 equilibrium reactions in this system are linearly independ-



**Figure 4.** Concentration curves for ion exchanged species.

ent, i.e.,  $N_E = N_c = 24$ . Hence, a formal matrix decomposition of the reaction network should yield 24 mass action equilibrium equations and 7 mass conservation equations ( $N_C = M - N_I = 31 - 24$ ).

[75] Without basis switching, the simulation stopped while checking mass balances. BIOGEOCHEM displayed information in the output file that mass balances of the third component, i.e.,  $Al^{3+}$ , was not maintained. A false convergence was reached due to the low concentration of  $Al^{3+}$ . The absolute error of  $Al(OH)_4^-$  concentration computed based on that of  $Al^{3+}$  was too large. When this concentration was substituted into the mass balance equation for the component  $Al^{3+}$ , it resulted in a large mass balance error. Hence the converged  $Al^{3+}$  concentration is a false value, and the simulation stopped. The input concentration for TOT  $Al^{3+}$  was  $4.263 \times 10^{-5}$  M while the calculated TOT  $Al^{3+}$  is  $3.684 \times 10^{-4}$  M; thus it is seen that there is a large error in mass balance. The simulated concentration of

species  $Al^{3+}$  is  $3.015 \times 10^{-22}$  M, which is close to 0, while the calculated concentration of  $Al(OH)_4^-$  is  $3.680 \times 10^{-4}$  M, which is relatively very large compared to other species in the mass balance equation. Thus the species  $Al(OH)_4^-$  is the major contributor to the calculated TOT  $Al^{3+}$ .

[76] When the basis switching was enabled,  $Al(OH)_4^-$ , which was the most abundant species in the mass balance equation for  $Al^{3+}$ , was selected automatically by BIOGEOCHEM to replace  $Al^{3+}$  as the master species for aluminum. Simulation results are shown in Table 4, with 3 of the 4 possible precipitated species actually formed within the system. The simulated concentration of  $Al(OH)_4^-$  is  $5.64 \times 10^{-8}$  M, while the calculated concentration of  $Al^{3+}$  is  $2.80 \times 10^{-10}$  M. It is seen from Table 4 that the species  $Al(OH)_4^-$ , that has been chosen as the component species by BIOGEOCHEM, has the largest concentration among all species composing the component  $Al^{3+}$ . Thus the contribution to mass balance is mainly from the species  $Al(OH)_4^-$ ; however, the mass balance error due to the species  $Al(OH)_4^-$  is small, because the concentration of species  $Al(OH)_4^-$  is directly simulated to satisfy the mass balance equation. The contribution to mass balance error by species  $Al^{3+}$  is small because its concentration (which is indirectly calculated based on the concentrations of component species) is relatively small compared to that of the component species  $Al(OH)_4^-$ .

## 7. Conclusion and Discussion

[77] This paper proposed a generic framework to model biogeochemical processes. To use this generic paradigm for modeling reactive chemicals, the system must be translated mathematically into a reaction network. A computer model

**Table 3.** Reaction Network for the Example Problem in Section 6.3

Reaction	Reaction Number	Reaction Constants
$H_2O(\ell) \rightleftharpoons H^+ + OH^-$	(R1)	$\text{Log } K_1^c = -14.0$
$Fe^{3+} + H_2O(\ell) \rightleftharpoons FeOH^{2+} + H^+$	(R2)	$\text{Log } K_2^c = -2.19$
$Fe^{3+} + 2H_2O(\ell) \rightleftharpoons Fe(OH)_2^+ + 2H^+$	(R3)	$\text{Log } K_3^c = -5.67$
$Fe^{3+} + 3H_2O(\ell) \rightleftharpoons Fe(OH)_3(aq) + 3H^+$	(R4)	$\text{Log } K_4^c = -12.56$
$Fe^{3+} + 4H_2O(\ell) \rightleftharpoons Fe(OH)_4^- + 4H^+$	(R5)	$\text{Log } K_5^c = -21.6$
$Al^{3+} + H_2O(\ell) \rightleftharpoons AlOH^{2+} + H^+$	(R6)	$\text{Log } K_6^c = -5.00$
$Al^{3+} + 2H_2O(\ell) \rightleftharpoons Al(OH)_2^+ + 2H^+$	(R7)	$\text{Log } K_7^c = -10.20$
$Al^{3+} + 4H_2O(\ell) \rightleftharpoons Al(OH)_4^- + 4H^+$	(R8)	$\text{Log } K_8^c = -23.00$
$Al^{3+} + 3H_2O(\ell) \rightleftharpoons Al(OH)_3(aq) + 3H^+$	(R9)	$\text{Log } K_9^c = -17.20$
$H^+ + CO_3^{2-} \rightleftharpoons HCO_3^-$	(R10)	$\text{Log } K_{10}^c = 10.33$
$2H^+ + CO_3^{2-} \rightleftharpoons H_2CO_3$	(R11)	$\text{Log } K_{11}^c = 16.68$
$Ca^{2+} + CO_3^{2-} \rightleftharpoons CaCO_3(aq)$	(R12)	$\text{Log } K_{12}^c = 3.22$
$Ca^{2+} + H^+ + CO_3^{2-} \rightleftharpoons CaHCO_3^+$	(R13)	$\text{Log } K_{13}^c = 11.44$
$H^+ + SO_4^{2-} \rightleftharpoons HSO_4^-$	(R14)	$\text{Log } K_{14}^c = 1.99$
$Ca^{2+} + SO_4^{2-} \rightleftharpoons CaSO_4(aq)$	(R15)	$\text{Log } K_{15}^c = 2.30$
$Al^{3+} + SO_4^{2-} \rightleftharpoons AlSO_4^+$	(R16)	$\text{Log } K_{16}^c = 3.50$
$Al^{3+} + 2SO_4^{2-} \rightleftharpoons Al(SO_4)_2^-$	(R17)	$\text{Log } K_{17}^c = 5.00$
$Fe^{3+} + SO_4^{2-} \rightleftharpoons FeSO_4^+$	(R18)	$\text{Log } K_{18}^c = 4.04$
$Fe^{3+} + 2SO_4^{2-} \rightleftharpoons Fe(SO_4)_2^-$	(R19)	$\text{Log } K_{19}^c = 5.42$
$Na^+ + SO_4^{2-} \rightleftharpoons NaSO_4^-$	(R20)	$\text{Log } K_{20}^c = 0.07$
$Ca^{2+} + CO_3^{2-} \rightleftharpoons CaCO_3(s)$	(R21)	$\text{Log } K_{21}^c = 8.48$
$Al^{3+} + 3H_2O(\ell) \rightleftharpoons Al(OH)_3(s) + 3H^+$	(R22)	$\text{Log } K_{22}^c = -9.11$
$Fe^{3+} + 3H_2O(\ell) \rightleftharpoons Fe(OH)_3(s) + 3H^+$	(R23)	$\text{Log } K_{23}^c = -4.89$
$Ca^{2+} + SO_4^{2-} \rightleftharpoons CaSO_4(s)$	(R24)	$\text{Log } K_{24}^c = 4.58$

**Table 4.** Initial and Simulated Species Concentration

Species	$C_0$ (M)	$\text{Log}(C_{eq})$
$Ca^{2+}$	$6.335 \times 10^{-1}$	-2.160
$CO_3^{2-}$	$6.365 \times 10^{-1}$	-5.478
$Al^{3+}$	$4.263 \times 10^{-5}$	-9.553
$SO_4^{2-}$	$3.177 \times 10^{-2}$	-1.582
$H^+$	$2.056 \times 10^{-2}$	-6.536
$Fe^{3+}$	$1.234 \times 10^{-5}$	-13.77
$Na^+$	$3.043 \times 10^{-5}$	-1.522
$OH^-$	0.000	-7.359
$FeOH^{2+}$	0.000	-9.952
$Fe(OH)_2^+$	0.000	-7.211
$Fe(OH)_3$	0.000	-7.670
$Fe(OH)_4^-$	0.000	-10.069
$Al(OH)_2^+$	0.000	-8.542
$Al(OH)_3$	0.000	-7.521
$Al(OH)_4^-$	0.000	-7.249
$Al(OH)_3$	0.000	-8.090
$HCO_3^-$	0.000	-2.000
$H_2CO_3$	0.000	-2.292
$CaCO_3$	0.000	-5.260
$CaHCO_3^+$	0.000	-3.471
$HSO_4^-$	0.000	-6.443
$CaSO_4$	0.000	-2.283
$AlSO_4^+$	0.000	-8.896
$Al(SO_4)_2^-$	0.000	-9.398
$FeSO_4^+$	0.000	-12.576
$Fe(SO_4)_2^-$	0.000	-13.198
$NaSO_4^-$	0.000	-3.454
$CaCO_3$ (s)	0.000	-0.207
$Al(OH)_3$ (s)	0.000	-4.371
$Fe(OH)_3$ (s)	0.000	-4.912
$CaSO_4$ (s)	0.000	$-\infty$



was developed to diagonalize the reaction network and numerically solve the diagonalized system of kinetic-variable equations, mass conservation equations, and a subset of nonlinear algebraic equations (governing fast ion exchange and precipitation-dissolution reactions, and some fast reactions that are modeled with user-specified algebraic equations). The model is designed to have the most generic capability for modeling biogeochemical processes. It can simulate both equilibrium and kinetic reactions involving aqueous complexation, adsorption, ion exchange, precipitation/dissolution, oxidation-reduction, acid-base reactions, and microbial-mediated reactions. Any fast reaction can be modeled with an infinite rate governed by a mass action equilibrium equation or by a user-specified algebraic equation. Any slow reaction can be modeled with microbial-mediated enzymatic kinetics, empirical nth order rates, an elementary rate, or a user-specified rate equation. The conceptualization of reaction networks and the specification of rate formulations and parameters should be carried out iteratively in collaboration with modelers and experimentalists who understand the system [Burgos *et al.*, 2002, 2003].

[78] The selection of chemical components and kinetic variables are automated within the BIOGEOCHEM preprocessor. In order to facilitate numerical integration, the set of ordinary differential equations governing the production/consumption of all species are decomposed into three subsets: mass action equilibrium equations representing fast equilibrium reactions, kinetic-variable equations representing slow kinetic reactions, and mass conservation equations representing chemical components. Basis switching is included to enhance the robustness of the model.

[79] The general paradigm addresses all the questions and difficulties arising from primitive or DAE approaches. This paradigm is new only in concept as the use of the diagonalization-decomposition procedure from mathematical and numerical perspectives is well accepted. This procedure facilitates the elimination of fast/equilibrium reactions from slow/kinetic reactions to reduce the number of unknowns that must be solved simultaneously, explicitly enforces mass conservation, removes redundant fast/equilibrium reactions, and excludes irrelevant slow/kinetic reactions in the reaction network. The decoupling of fast reactions from kinetic reactions alleviates the stiffness of the system. The explicit enforcement of mass conservation overcomes the mass conservation error due to numerical integration errors. The removal of redundant reactions circumvents the problem of singularity. The exclusion of irrelevant slow/kinetic reactions greatly improves computational efficiency and avoids problematic export of meaningless rate formulations and parameter estimations to other system in which these reactions are relevant. Finally and most importantly, the diagonalization of slow/kinetic reactions allows the formulation and parameterization of individual rate equations rather than the optimization of rate formulation/parameters for all reactions simultaneously [Yeh *et al.*, 2001a]. The individual rate formulation/parameterization is more descriptive of geochemical and biochemical reactions [Burgos *et al.*, 2002, 2003].

[80] To make the model numerically robust, a dynamic basis-species switching strategy due to the nonuniqueness of components is employed. Backward basis switching allows components to freely change in the simulation of

chemistry module, while being recovered for transport simulation. Three example problems were selected to demonstrate the versatility and robustness of the model.

[81] BIOGEOCHEM is a stand-alone batch model, which can also be coupled with a hydrologic transport model [Fang and Yeh, 2002]. It can apply to systems of high complexity and can serve as a tool for the planning of batch experiments, assessing system consistency and minimum data needs for reaction based modeling [Yeh *et al.*, 2001a]. The model described was tested with over 15 examples [Yeh and Fang, 2002], although only 3 problems are demonstrated in this paper to illustrate the versatility and flexibility of the model.

## Appendix A

### A1. Decomposition I

[82] Let start with the following matrix, which is a repeat of equation (5)

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[\text{H}]/dt \\ d[\text{NTA}]/dt \\ d[\text{HNNTA}]/dt \\ d[\text{Co}]/dt \\ d[\text{CoNTA}]/dt \\ d[\text{B}]/dt \end{Bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix}$$

First, choose suspected component species. From a geochemist's point of view, let's choose H, NTA, and Co as component species, i.e., species that have the low priority to be chosen as pivot elements. The remaining species can therefore have higher priority to be chosen as pivot elements.

[83] In the first column of the reaction matrix in equation (5), there are three nonzero rows (first, second, and third rows, respectively). The first and second rows are on the low priority list for choosing pivot elements because their corresponding species, H and NTA, respectively, have been selected as components. Therefore the third row (corresponding to HNNTA) is chosen as a pivot element. Pivoting on row 3 (HNNTA) gives

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[\text{H}]/dt \\ d[\text{NTA}]/dt \\ d[\text{HNNTA}]/dt \\ d[\text{Co}]/dt \\ d[\text{CoNTA}]/dt \\ d[\text{B}]/dt \end{Bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix} \quad (\text{A1})$$

[84] In the second column of the column 1 reduced reaction matrix in equation (A1), there are three nonzero rows (second, fourth, and fifth rows, respectively). The second and fourth rows (corresponding to NTA and Co, respectively) are on the low priority list for the selection of pivot elements because NTA and Co have been selected as components. (If the third row were not zero, it could not be chosen as a pivot element because it had already been chosen as the pivot element in the reduction of first column.) Therefore choose the fifth row as the pivot element. Pivoting on row 5 (CoNTA) gives

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{pmatrix} \quad (A2)$$

If the fifth row in column 2 of the column 1 reduced reaction matrix in (A1) were 0, then one would have to choose either row 2 or row 4 as the pivot element. This would signify that the user had selected incorrect components. Under such circumstances, the preprocessor would have chosen either NTA (row 2) or Co (row 4) as the pivot element and removed it from the user-selected component list and treated it as a product species.

[85] There are only two no-zero rows (rows 3 and 5) in the third column in the column 2 reduced reaction in equation (A2). However, these two rows (corresponding to HNTA and CoNTA) have already been chosen as the pivot elements in column 1 and column 2 reductions, respectively. They cannot be chosen as pivoting elements, hence no row can be found as the pivot element for this third column. This signifies that the third reaction is a linearly dependent reaction and it depends on the first and second reactions, i.e., R3 is linearly dependent on R1 and R2.

[86] In the fourth column of the reduced reaction matrix in equation (A2) there are four nonzero rows (rows 1, 2, 3, and 6). Row 3 is not eligible as a pivoting element because it has already been chosen in the first column reduction. Rows 1 and 2 are on the low priority list. Therefore choose row 6 as the pivot element. Pivoting on row 6 (B) gives

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{pmatrix} \quad (A3)$$

The above procedure completes the reduction of the reaction matrix and the decomposition of the unit matrix.

[87] Expand the matrix in equation (A3) and the following equations are obtained:

$$\frac{d[H]}{dt} + \frac{d[HNTA]}{dt} + 2\frac{d[B]}{dt} = 0 \quad (A4a)$$

$$\frac{d[NTA]}{dt} + \frac{d[HNTA]}{dt} + \frac{d[CoNTA]}{dt} + \frac{d[B]}{dt} = 0 \quad (A4b)$$

$$\frac{d[HNTA]}{dt} + \frac{d[B]}{dt} = R_1 + R_3 \quad (A4c)$$

$$\frac{d[Co]}{dt} + \frac{d[CoNTA]}{dt} = 0 \quad (A4d)$$

$$\frac{d[CoNTA]}{dt} = -R_2 - R_3 \quad (A4e)$$

$$\frac{d[B]}{dt} = R_4 \quad (A4f)$$

Equations (A4a), (A4b), and (A4d) define Total H, Total NTA, and Total Co, respectively. Since the total amount of H, NTA, and Co are reaction invariant, they are defined as component species. Equations (A4c) and (A4e) each contains one equilibrium reaction. They are replaced by the mass action equilibrium equations as shown in the last equation of equations (6a) and (6b) in section 3.

[88] Equation (A3) is rearranged so that it can be written in the form of equation (8) in section 3.2.3 as

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{pmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{pmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{bmatrix} \begin{pmatrix} R_1 \\ R_2 \\ R_4 \\ R_3 \end{pmatrix} \quad (A5)$$

Comparing equation (A5) and equation (8) which is repeated here for the convenience of direct comparison,

$$\mathbf{B} \frac{d\mathbf{C}}{dt} = \begin{bmatrix} \mathbf{D} & \mathbf{K} \\ \mathbf{0}_1 & \mathbf{0}_2 \end{bmatrix} \mathbf{R}$$

we can easily see that

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix};$$

$$\frac{d\mathbf{C}}{dt} = \begin{pmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{pmatrix}; \text{ and } \mathbf{R} = \begin{pmatrix} R_1 \\ R_2 \\ R_4 \\ R_3 \end{pmatrix} \quad (A5')$$

and

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \mathbf{K} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}; \quad (\text{A5''})$$

$$\mathbf{0}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \text{ and } \mathbf{0}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

## A2. Decomposition II

[89] An alternative diagonalization is to choose HNTA, CoNTA, and B as components. Again repeat equation (5)

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[\text{H}]/dt \\ d[\text{NTA}]/dt \\ d[\text{HNTA}]/dt \\ d[\text{Co}]/dt \\ d[\text{CoNTA}]/dt \\ d[\text{B}]/dt \end{Bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix}$$

The reduction and decomposition procedure is similar to decomposition I as follows. In the first column in equation (5), pivoting on H (row 1) gives

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[\text{H}]/dt \\ d[\text{NTA}]/dt \\ d[\text{HNTA}]/dt \\ d[\text{Co}]/dt \\ d[\text{CoNTA}]/dt \\ d[\text{B}]/dt \end{Bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix} \quad (\text{A6})$$

In the second column in equation (A6), pivoting on NTA (row 2) gives

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[\text{H}]/dt \\ d[\text{NTA}]/dt \\ d[\text{HNTA}]/dt \\ d[\text{Co}]/dt \\ d[\text{CoNTA}]/dt \\ d[\text{B}]/dt \end{Bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix} \quad (\text{A7})$$

No pivot element can be found in the third column. This signifies that R3 is linearly dependent on R1 and R2. In the fourth column in (A7), pivoting on Co (row 4) gives

$$\begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & -2 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[\text{H}]/dt \\ d[\text{NTA}]/dt \\ d[\text{HNTA}]/dt \\ d[\text{Co}]/dt \\ d[\text{CoNTA}]/dt \\ d[\text{B}]/dt \end{Bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix} \quad (\text{A8})$$

Expand the matrix and the following equations can be obtained:

$$\frac{d[\text{NTA}]}{dt} - \frac{d[\text{Co}]}{dt} = -R_1 - R_3 \quad (\text{A9a})$$

$$\frac{d[\text{Co}]}{dt} = R_2 + R_3 \quad (\text{A9b})$$

$$-\frac{d[\text{H}]}{dt} + 2\frac{d[\text{NTA}]}{dt} + \frac{d[\text{HNTA}]}{dt} - 2\frac{d[\text{Co}]}{dt} = 0 \quad (\text{A9c})$$

$$\frac{d[\text{H}]}{dt} - \frac{d[\text{NTA}]}{dt} + \frac{d[\text{Co}]}{dt} = -R_4 \quad (\text{A9d})$$

$$\frac{d[\text{Co}]}{dt} + \frac{d[\text{CoNTA}]}{dt} = 0 \quad (\text{A9e})$$

$$\frac{d[\text{H}]}{dt} - \frac{d[\text{NTA}]}{dt} + \frac{d[\text{Co}]}{dt} + \frac{d[\text{B}]}{dt} = 0 \quad (\text{A9f})$$

[90] Equations (A9c), (A9e), and (A9f) define Total HNTA, Total CoNTA, and Total B, respectively. Equations (A9a) and (A9b) contain equilibrium reactions, they are replaced by the mass action equilibrium equations as shown in the last equation of equations (7a) and (7b) in section 3.

## A3. Elimination of Secondary Species

[91] To illustrate the procedure of eliminating secondary species, consider the following reaction network with two fast reactions:



with equilibrium constant  $K_1^c$



with equilibrium  $K_2^c$ . Let us assume that A and B have been chosen as the component via column reduction of the



reaction matrix. Reactions RA1 and RA2 are two linearly independent reactions, two species can be eliminated from simultaneous consideration. To illustrate the procedure of elimination, a matrix made up this two reactions is first generated as

$$\begin{bmatrix} -1 & 0 & -1 & 3 \\ -1 & -2 & 2 & 0 \end{bmatrix} \begin{Bmatrix} \log(A) \\ \log(B) \\ \log(C) \\ \log(D) \end{Bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} \log K_1^c \\ \log K_2^c \end{Bmatrix} \quad (\text{A10})$$

[92] In the first row in equation (A10), since A and B are component species, the first and second columns corresponding to species A and B, respectively, cannot be chosen as the pivoting element. The third and fourth columns corresponding to species C and D are the potential candidates. Because the fourth column (Species D) has fewer nonzero rows, it is chosen as the pivot element for the first row reduction. Pivoting on species D gives

$$\begin{bmatrix} -1 & 0 & -1 & 3 \\ -1 & -2 & 2 & 0 \end{bmatrix} \begin{Bmatrix} \log(A) \\ \log(B) \\ \log(C) \\ \log(D) \end{Bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} \log K_1^c \\ \log K_2^c \end{Bmatrix} \quad (\text{A11})$$

It is noted that equation (A11) is identical to equation (10) for this case. This is so because all rows in column 4 (corresponding to species D, the chosen pivot) are zeros, hence pivoting on species D has not changed the matrix.

[93] In the second row in equation (A11), the first and second columns corresponding to Spaces A and B cannot be chosen as the pivot element as in the first row reduction. Column 4 corresponding to Species D has already been pivoted. The third column corresponding to Species C is the only candidate for pivoting. Pivoting on Species C (column 3) gives

$$\begin{bmatrix} -3/2 & -1 & 0 & 3 \\ -1 & -2 & 2 & 0 \end{bmatrix} \begin{Bmatrix} \log(A) \\ \log(B) \\ \log(C) \\ \log(D) \end{Bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} \log K_1^c \\ \log K_2^c \end{Bmatrix} \quad (\text{A12})$$

[94] The next step is to normalize the finally reduced matrix with respect to the pivot elements (column 4 of the first row and column 3 of the second row in equation A(12)). After the normalization, we have

$$\begin{bmatrix} -1/2 & -1/3 & 0 & 1 \\ -1/2 & -1 & 1 & 0 \end{bmatrix} \begin{Bmatrix} \log(A) \\ \log(B) \\ \log(C) \\ \log(D) \end{Bmatrix} = \begin{bmatrix} 1/3 & 1/6 \\ 0 & 1/2 \end{bmatrix} \begin{Bmatrix} \log K_1^c \\ \log K_2^c \end{Bmatrix} \quad (\text{A13})$$

[95] Expanding equation (A13), we have

$$(D) = (K_1^c)^{\frac{1}{3}} (K_2^c)^{\frac{1}{6}} (A)^{\frac{1}{3}} (B)^{\frac{1}{6}} \quad (\text{A14})$$

and

$$(C) = (K_2^c)^{\frac{1}{2}} (A)^{\frac{1}{2}} (B) \quad (\text{A15})$$

The elimination of Species D and C using equation (A14) and (A15) is as easy as “eating rice” (Chinese proverb).

## Appendix B

[96] The decomposed set of equations for examples 1 and 2, respectively, are given in sections B1 and B2, respectively.

### B1. Matrix Decomposition for the Reaction Network of the Transformation of EDTA

#### B1.1. Mass Action Equilibrium Equations

$$(\text{Co(II)(aq)}) = \frac{1}{K_1^c} \frac{(S_{\text{neg-Co}})}{(S_{\text{neg}})} \quad (\text{B1})$$

$$(S_{\text{pos-Co(II)EDTA}}) = K_2^c (\text{Co(II)EDTA(aq)}) (S_{\text{pos}}) \quad (\text{B2})$$

$$(S_{\text{pos-Fe(III)EDTA}}) = K_3^c (S_{\text{pos}}) (\text{Fe(III)EDTA(aq)}) \quad (\text{B3})$$

$$(\text{EDTA(aq)}) = \frac{1}{K_4^c} \frac{(S_{\text{pos}})}{(S_{\text{pos-EDTA}})} \quad (\text{B4})$$

$$(\text{Co(III)EDTA(aq)}) = \frac{1}{K_5^c} \frac{(S_{\text{pos-Co(III)EDTA}})}{(S_{\text{pos}})} \quad (\text{B5})$$

#### B1.2. Kinetic-Variable Equations

$$\frac{d([\text{Co(II)(aq)}] + [S_{\text{neg-Co}}])}{dt} = R_6; \quad (\text{B6})$$

$$R_6 = k_6^f [S_{\text{pos-Co(II)EDTA}}] - k_6^b [\text{Co(II)(aq)}] [S_{\text{pos-EDTA}}]$$

$$\frac{d([\text{EDTA}] + [S_{\text{pos-EDTA}}] - [\text{Co(II)(aq)}] + [S_{\text{neg-Co}}])}{dt} = -R_7 - R_{10}$$

$$R_7 = k_7^f [S_{\text{pos-EDTA}}] - k_7^b [S_{\text{pos}}] [\text{Fe(III)-EDTA(aq)}], \quad (\text{B7})$$

$$R_{10} = \frac{m_2 [\text{EDTA(aq)}] [\text{O}_2] [\text{Biomass}]}{(k_{12} + [\text{EDTA(aq)}]) (k_{22} + [\text{O}_2])}$$

$$\frac{d([\text{Co(III)EDTA(aq)}] + [S_{\text{pos-Co(III)EDTA}}])}{dt} = R_8; \quad (\text{B8})$$

$$R_8 = k_8^f [\text{Co(II)EDTA(aq)}] - k_8^b [\text{Co(III)EDTA(aq)}]$$

$$\frac{d[\text{CO}_2]}{dt} = 3R_9 + 3R_{10}R = \frac{\mu_1 [\text{Fe(III)EDTA(aq)}] [\text{O}_2] [\text{Biomass}]}{(k_{11} + [\text{Fe(III)EDTA(aq)}]) (k_{21} + [\text{O}_2])},$$

$$R_{10} = \frac{\mu_2 [\text{EDTA(aq)}] [\text{O}_2] [\text{Biomass}]}{(k_{12} + [\text{EDTA(aq)}]) (k_{22} + [\text{O}_2])} \quad (\text{B9})$$

**B1.3. Mass Conservation Equations**

$$\text{TOT}_{\text{S}_{\text{neg}}} = [\text{S}_{\text{neg}}] + [\text{S}_{\text{neg-Co}}] = \text{Initial TOT}_{\text{S}_{\text{neg}}} \quad (\text{B10})$$

$$\begin{aligned} \text{TOT}_{\text{Co(II)EDTA}} &= [\text{Co(II)(aq)}] + [\text{S}_{\text{neg-Co}}] + [\text{Co(II)EDTA(aq)}] \\ &+ [\text{S}_{\text{pos-Co(II)EDTA}}] + [\text{Co(III)EDTA(aq)}] \\ &+ [\text{S}_{\text{pos-Co(III)EDTA}}] = \text{Initial TOT}_{\text{Co(II)EDTA}} \end{aligned} \quad (\text{B11})$$

$$\begin{aligned} \text{TOT}_{\text{S}_{\text{pos}}} &= [\text{S}_{\text{pos}}] + [\text{S}_{\text{pos-Co(II)EDTA}}] + [\text{S}_{\text{pos-Fe(III)EDTA}}] \\ &+ [\text{S}_{\text{pos-EDTA}}] + [\text{S}_{\text{pos-Co(III)EDTA}}] \\ &= \text{Initial TOT}_{\text{S}_{\text{pos}}} \end{aligned} \quad (\text{B12})$$

$$\begin{aligned} \text{TOT}_{\text{Fe(III)EDTA}} &= -[\text{Co(II)(aq)}] - [\text{S}_{\text{neg-Co}}] \\ &+ [\text{Fe(III)EDTA(aq)}] + [\text{S}_{\text{pos-Fe(III)EDTA}}] \\ &+ [\text{EDTA(aq)}] + [\text{S}_{\text{pos-EDTA}}] + 1/3[\text{CO}_2] \\ &= \text{Initial TOT}_{\text{Fe(III)EDTA}} \end{aligned} \quad (\text{B13})$$

$$\text{TOT}_{\text{O}_2} = [\text{O}_2] + 2[\text{CO}_2] = \text{Initial TOT}_{\text{O}_2} \quad (\text{B14})$$

$$\text{TOT}_{\text{Biomass}} = [\text{Biomass}] - 1/3[\text{CO}_2] = \text{Initial TOT}_{\text{Biomass}} \quad (\text{B15})$$

[97] Note that a species inside a parenthesis in equations (B1)–(B15) denotes the activity of the species. A species inside a bracket in equations (B1)–(B15) denotes the concentration of the species. Activity is activity coefficient  $\times$  concentration.

**B2. Matrix Decomposition for the Complex (Multiple Reaction Types) Mixed (Equilibrium and Kinetic) Reaction System**

**B2.1. User-Specified Algebraic Equations**

$$\frac{S_A N_S}{N_A} M_{\text{mineral}} [\text{M}] = [\text{S}_1] + [\text{S}_2] + [\text{S}_3] + [\text{S}_4] + [\text{S}_5] + [\text{S}_6] + [\text{S}_7] + [\text{S}_8] \quad (\text{B16})$$

[98] This user-specified equation is modified from *Stumm and Morgan* [1996], in which  $S_A$  is the unit surface area ( $\text{m}^2 \text{g}^{-1}$ ) of mineral,  $N_S$  is the surface site density (mol sites  $\text{m}^{-2}$ ),  $N_A$  is Avogadro's number (mol sites  $\text{mol}^{-1}$ ),  $M_{\text{mineral}}$  is the molecular weight of mineral ( $\text{g mol}^{-1}$ ).

**B2.2. Mass Action Equilibrium Equations**

$$(\text{C}_7) = K_4^e (\text{C}_5)(\text{C}_6) \quad (\text{B17})$$

$$(\text{C}_8) = K_5^e (\text{C}_2)(\text{C}_5)(\text{C}_6) \quad (\text{B18})$$

$$(\text{C}_9) = K_6^e \frac{(\text{C}_6)}{(\text{C}_2)} \quad (\text{B19})$$

$$(\text{C}_{11}) = K_8^e (\text{C}_1)(\text{C}_2)(\text{C}_5) \quad (\text{B20})$$

$$(\text{C}_{12}) = K_9^e \frac{(\text{C}_1)(\text{C}_5)}{(\text{C}_2)} \quad (\text{B21})$$

$$(\text{C}_{13}) = K_{10}^e \frac{(\text{C}_1)(\text{C}_5)}{(\text{C}_2)^2} \quad (\text{B22})$$

$$(\text{C}_{14}) = K_{11}^e \frac{(\text{C}_1)}{(\text{C}_2)} \quad (\text{B23})$$

$$(\text{C}_{15}) = K_{12}^e \frac{(\text{C}_1)}{(\text{C}_2)^2} \quad (\text{B24})$$

$$(\text{C}_{16}) = K_{13}^e \frac{(\text{C}_1)}{(\text{C}_2)^3} \quad (\text{B25})$$

$$(\text{C}_{17}) = K_{14}^e \frac{(\text{C}_1)}{(\text{C}_2)^4} \quad (\text{B26})$$

$$(\text{C}_{18}) = K_{15}^e \frac{(\text{C}_1)^2}{(\text{C}_2)^2} \quad (\text{B27})$$

$$(\text{C}_{19}) = K_{16}^e (\text{C}_2)(\text{C}_4)(\text{C}_5) \quad (\text{B28})$$

$$(\text{C}_{20}) = K_{17}^e \frac{(\text{C}_4)}{(\text{C}_2)} \quad (\text{B29})$$

$$(\text{C}_{21}) = K_{18}^e \frac{(\text{C}_4)}{(\text{C}_2)^2} \quad (\text{B30})$$

$$(\text{C}_{22}) = K_{19}^e \frac{(\text{C}_4)}{(\text{C}_2)^3} \quad (\text{B31})$$

$$(\text{C}_{23}) = K_{20}^e (\text{C}_2)(\text{C}_5) \quad (\text{B32})$$

$$(\text{C}_{24}) = K_{21}^e (\text{C}_2)^2 (\text{C}_5) \quad (\text{B33})$$

$$(\text{C}_{25}) = K_{22}^e (\text{C}_2)^3 (\text{C}_5) \quad (\text{B34})$$

$$(\text{C}_{26}) = K_{23}^e (\text{C}_2)^4 (\text{C}_5) \quad (\text{B35})$$

$$(\text{C}_{27}) = \frac{1}{K_{24}^e} \frac{1}{(\text{C}_2)} \quad (\text{B36})$$

$$(S_2) = K_{25}^e \frac{(S_1)}{(C_2)} \quad (B37)$$

$$(S_3) = K_{26}^e (C_2)(S_1) \quad (B38)$$

$$(S_4) = K_{27}^e (C_2)^3 (C_5)(S_1) \quad (B39)$$

$$(\text{site-C}_{30}) = \frac{1}{K_{33}^{e, 0.5}} \frac{(C_{30})(\text{site-C}_6)^{0.5}}{(C_6)^{0.5}} \quad (B40)$$

### B2.3. Kinetic Variable Equations

$$\frac{d[C_{10}]}{dt} = R_7; R_7 = k_7^f [C_1][C_5] - k_7^b [C_{10}] \quad (B41)$$

$$\frac{d}{dt} ([S_1] + [M] + [S_2] + [S_3] + [S_5] + [S_6] + [S_7] + [S_4] + [S_8]) = -R_1, \quad (B42)$$

$$R_1 = k_1^f [M]$$

$$\frac{d[C_3]}{dt} = -R_3, R_3 = k_3^f [C_3] - k_3^b [C_4][C_5] \quad (B43)$$

$$\frac{d[S_5]}{dt} = R_{28}, R_{28} = k_{28}^f [S_1][C_1][C_2][C_5] - k_{28}^b [S_5] \quad (B44)$$

$$\frac{d[S_6]}{dt} = R_{29}, R_{29} = k_{29}^f [S_1][C_2][C_4][C_5] - k_{29}^b [S_6] \quad (B45)$$

$$\frac{d[S_7]}{dt} = R_{30}, R_{30} = k_{30}^f \frac{[S_1][C_4]}{[C_2]} - k_{30}^b [S_7] \quad (B46)$$

$$\frac{d[S_8]}{dt} = R_{31}, R_{31} = k_{31}^f [S_1][C_2][C_5][C_6] - k_{31}^b [S_8] \quad (B47)$$

$$\frac{d[C_{29}]}{dt} = -R_{32}, R_{32} = k_{32}^f [C_{29}][\text{site-C}_{30}]^2 - k_{32}^b [\text{site-C}_{29}][C_{30}]^2 \quad (B48)$$

### B2.4. Mass Conservation Equations

$$\text{TOT}_{C_6} = [C_6] + [C_7] + [C_8] + [C_9] + [S_8] + [C_{29}] - 0.5[\text{site} - C_{30}] \\ = \text{Initial TOT}_{C_6} \quad (B49)$$

$$\text{TOT}_{C_4} = [C_4] + [C_3] + [C_{19}] + [C_{20}] + [C_{21}] + [C_{22}] + [S_6] + [S_7] \\ = \text{Initial TOT}_{C_4} \quad (B50)$$

$$\text{TOT}_{C_2} = [C_2] - 3[S_1] + [C_8] - [C_9] + [C_{11}] - [C_{12}] \\ - 2[C_{13}] - [C_{14}] - 2[C_{15}] - 3[C_{16}] - 3[M] \\ - 4[C_{17}] - 2[C_{18}] + [C_{19}] - [C_{20}] - 2[C_{21}] \\ - 3[C_{22}] + [C_{23}] + 2[C_{24}] + 3[C_{25}] + 4[C_{26}] \\ - [C_{27}] - 4[S_2] - 2[S_3] - 2[S_5] - 2[S_6] \\ - 4[S_7] - 2[S_8] = \text{Initial TOT}_{C_2} \quad (B51)$$

$$\text{TOT}_{C_5} = [C_5] + [C_7] + [C_8] + [C_{10}] + [C_{11}] + [C_{12}] + [C_{13}] \\ + [C_3] + [C_{19}] + [C_{23}] + [C_{24}] + [C_{25}] + [C_{26}] + [S_5] \\ + [S_6] + [S_4] + [S_8] = \text{Initial TOT}_{C_5} \quad (B52)$$

$$\text{TOT}_{C_1} = [C_1] + [S_1] + [C_{10}] + [C_{11}] + [C_{12}] + [C_{13}] + [C_{14}] \\ + [C_{15}] + [C_{16}] + [M] + [C_{17}] + 2[C_{18}] + [S_2] + [S_3] \\ + 2[S_5] + [S_6] + [S_7] + [S_4] + [S_8] = \text{Initial TOT}_{C_1} \quad (B53)$$

$$\text{TOT}_{C_{30}} = [C_{30}] + [\text{site-C}_{30}] = \text{Initial TOT}_{C_{30}} \quad (B54)$$

$$\text{TOT}_{\text{site-C}_{29}} = [C_{29}] + [\text{site-C}_{29}] = \text{Initial TOT}_{C_{29}} \quad (B55)$$

$$\text{TOT}_{\text{site-C}_6} = -[C_{29}] + 0.5[\text{site-C}_{30}] + [\text{site-C}_6] \\ = \text{Initial TOT}_{\text{site-C}_6} \quad (B56)$$

[99] Note that in equations (B16)–(B56) a species inside parentheses denotes the activity of the species. A species inside a bracket denotes the concentration of the species. Activity is activity coefficient  $\times$  concentration.

[100] **Acknowledgments.** This research was supported, in part, by National Science Foundation (NSF) grant EAR-0196048 with University of Central Florida and, in part, by the Natural and Accelerated Bioremediation Research Program (NABIR), Office of Biological and Environmental Research (OBER), U.S. Department of Energy (DOE) grant DE-FG02-98ER62691 with the Pennsylvania State University.

### References

- Abrams, R. H., K. Loague, and D. B. Kent, Development and testing of a compartmentalized reaction network model for redox zones in contaminated aquifers, *Water Resour. Res.*, 34(6), 1531–1542, 1998.
- Ambrose, R. B., et al., WASP4, a hydrodynamic and waterquality model—Model theory, user's manual, and programmer's guide, *EPA/600/3-87-039*, U.S. Environ. Prot. Agency, Washington, D. C., 1988.
- Atkins, P. W., *Physical Chemistry*, Oxford Univ. Press, New York, 1986.
- Barnwell, T. O., and L. C. Brown, The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and user manual, *EPA/600/3-87/007*, U.S. Environ. Prot. Agency, Washington, D. C., 1987.
- Bosma, W., and S. E. A. T. M. van der Zee, Transport of reacting solute in a one-dimensional, chemically heterogeneous porous medium, *Water Resour. Res.*, 29(1), 117–131, 1993.
- Brun, A., and P. Engesgaard, Modelling of transport and biogeochemical processes in pollution plumes: literature review and model development, *J. Hydrol.*, 256, 211–227, 2002.
- Brusseau, M., Transport of reactive contaminants in heterogeneous porous media, *Rev. Geophys.*, 32(3), 285–313, 1994.
- Burgos, W. D., R. A. Royer, Y. L. Fang, G. T. Yeh, A. S. Fisher, B. H. Jeon, and B. A. Dempsey, Theoretical and experimental considerations related to reaction-based modeling: A case study using iron (III) oxide bioreduction, *Geomicrobiol. J.*, 19(2), 1–35, 2002.
- Burgos, W. D., Y. L. Fang, R. A. Royer, G. T. Yeh, J. J. Stone, B. H. Jeon, and B. A. Dempsey, Reaction-based modeling of quinone-mediated bacterial iron (III) reduction, *Geochim. Cosmochim. Acta*, in press, 2003.
- Cederberg, G. A., R. L. Street, and J. O. Leckie, A groundwater mass transport and equilibrium chemistry model for multicomponent systems, *Water Resour. Res.*, 21, 1095–1104, 1985.
- Chang, J. R., G. T. Yeh, and T. E. Short, Modeling two-dimensional subsurface flow, fate and transport of microbes and chemicals, paper presented at International Symposium of Engineering Hydrology, Am. Soc. Civ. Eng., San Francisco, 25–30 July 1993.
- Chen, Y., L. M. Abriola, P. J. J. Alvarez, P. J. Anid, and T. M. Vogel, Modeling transport and biodegradation of benzene and toluene in sandy



- aquifer material: Comparisons with experimental measurements, *Water Resour. Res.*, 28(7), 1833–1847, 1992.
- Cheng, H.-P., Development and application of a three-dimensional finite element model of subsurface flow, heat transfer, and reactive chemical transport, Ph.D. thesis, Dep. of Civ. and Environ. Eng., Pa. State Univ., University Park, 1995.
- Cheng, J. R., and G. T. Yeh, Modeling three-dimensional subsurface flow, fate and transport of microbes and chemicals (3DFATMIC), paper presented at Xth International Conference on Numerical Methods in Water Resources, Univ. of Heidelberg, Heidelberg, Germany, 19–22 July 1994.
- Chilakapati, A., RAFT: A simulator for reactive flow and transport of groundwater contaminants, *PNL Rep. 10636*, Pac. Northwest Lab., Richland, Wash., 1995.
- Chilakapati, A., T. Ginn, and J. Szecsody, An analysis of complex reaction networks in groundwater modeling, *Water Resour. Res.*, 34(7), 1767–1780, 1998.
- Chilakapati, A., S. Yabusaki, J. Szecsody, and W. MacEvoy, Groundwater flow, multicomponent transport and biogeochemistry: Development and application of a coupled process model, *J. Contam. Hydrol.*, 43, 303–325, 2000.
- Davis, J., D. Kent, J. Coston, K. Hess, and J. Joye, Multispecies reactive tracer test in an aquifer with spatially variable chemical conditions, *Water Resour. Res.*, 36(1), 119–134, 2000.
- Fang, Y., and G. T. Yeh, Numerical modeling of reactive chemical transport under multiphase systems, paper presented at XIV International Conference on Computational Methods in Water Resources, Delft Univ. of Technol., Delft, Netherlands, 23–28 June 2002.
- Friedly, J., and J. Rubin, Solute transport with multiple equilibrium-controlled or kinetically controlled chemical reactions, *Water Resour. Res.*, 28(6), 1935–1953, 1992.
- Goldberg, D. E., Chemistry foundations, in *Schaum's Outline of Theory and Problems*, 372 pp., McGraw-Hill, New York, 1991.
- Griffioen, J., Multicomponent cation exchange including alkalization/acidification following flow through a sandy sediment, *Water Resour. Res.*, 29(9), 3005–3019, 1993.
- Hostetler, J. C., and R. L. Erickson, FASTCHEM Package 5, *Rep. EA-5870-CCM*, Battelle Pac. Northwest Lab., Richland, Wash., 1989.
- Lensing, H. J., M. Voyt, and B. Herrling, Modeling of biologically mediated redox processes in the subsurface, *J. Hydrol.*, 159, 125–143, 1994.
- Lichtner, P. C., Continuum formulation of multicomponent-multiphase reactive transport, in *Reactive Transport in Porous Media*, edited by P. C. Lichtner, C. I. Steefel, and E. H. Oelkers, *Rev. Mineral.*, 34, 1–79, 1996.
- Lin, C. F., and M. M. Benjamin, Dissolution kinetics of minerals in the presence of sorbing and complexing ligands, *Environ. Sci. Technol.*, 24, 126–134, 1990.
- Liu, C., and T. Narasimhan, Redox-controlled multiple-species reactive chemical transport, 1, Model development, *Water Resour. Res.*, 25(5), 869–882, 1989a.
- Liu, C., and T. Narasimhan, Redox-controlled multiple-species reactive chemical transport, 2, Verification and application, *Water Resour. Res.*, 25(5), 883–910, 1989b.
- Liu, C., S. Kota, J. M. Zachara, J. K. Fredrickson, and C. K. Brinkman, Kinetic analysis of the bacterial reduction of goethite, *Environ. Sci. Technol.*, 35(12), 2482–2490, 2001.
- MacQuarrie, K. T. B., E. A. Sudicky, and E. O. Frind, Simulation of biodegradable organic contaminants in groundwater, 1, Numerical formulation in principal directions, *Water Resour. Res.*, 26, 207–222, 1990.
- McNab, J. W., and T. Narasimhan, A multiple species transport model with sequential decay chain interactions in heterogeneous subsurface environments, *Water Resour. Res.*, 29(8), 2737–2746, 1993.
- McNab, W. W., Jr., and T. N. Narasimhan, Modeling reactive transport of organic compounds in groundwater using a partial redox disequilibrium approach, *Water Resour. Res.*, 30(9), 2619–2635, 1994.
- McNab, W. W., Jr., and T. N. Narasimhan, Reactive transport of petroleum hydrocarbon constituents in a shallow aquifer: Modeling geochemical interactions between organic and inorganic species, *Water Resour. Res.*, 31(8), 2027–2033, 1995.
- Miller, C. W., and L. V. Benson, Simulation of solute transport in a chemical reactive heterogeneous system: Model development and application, *Water Resour. Res.*, 19(2), 381–391, 1983.
- Narasimhan, T. N., A. F. White, and T. Tokunaga, Groundwater contamination from an inactive uranium mill tailings pile, 2, Application of a dynamic mixing model, *Water Resour. Res.*, 22(13), 1820–1834, 1986.
- Parkhurst, D. L., User's guide to PHREEQC, A computer model for speciation, reaction-path, advective-dispersive transport and inverse geochemical calculations, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, 95-4227, 1995.
- Parkhurst, D. L., and C. A. J. Appelo, User's guide to PHREEQC (version 2)—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, 99-4259, 312 pp., 1999.
- Parkhurst, D. L., D. C. Thorstenson, and L. N. Plummer, PHREEQC—A computer program for geochemical calculations, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, 80-96, 195 pp., 1980.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in FORTRAN*, Cambridge Univ. Press, New York, 1992.
- Rubin, J., Transport of reacting solutes in porous media: Relation between mathematical nature of problem formulation and chemical nature of reactions, *Water Resour. Res.*, 19, 493–502, 1983.
- Saiers, J. E., H. Guha, P. Jardine, and S. Brooks, Development and evaluation of a mathematical model for the transport and oxidation-reduction of CoEDTA, *Water Resour. Res.*, 31(11), 3151–3165, 2000.
- Salvage, K., and G. T. Yeh, Development and application of a numerical model of kinetic and equilibrium microbiological and geochemical reactions (BIOKEMOD), *J. Hydrol.*, 209, 27–52, 1998.
- Salvage, K. M., G. T. Yeh, H. P. Cheng, and J. R. Cheng, Development of a model of subsurface hydrologic transport and biogeochemical reactions (HYDROBIOGEOCHEM), in *Computational Methods in Water Resources XI*, vol. 2, *Computation Methods in Surface Flow and Transport Problems*, pp. 517–524, Comput. Mech., Boston, Mass., 1996.
- Smith, J. M., *Chemical Engineering Kinetics*, 3rd ed., McGraw-Hill, New York, 1981.
- Steefel, C. I., and A. C. Lasaga, A coupled model for transport of multiple chemical species and kinetic precipitation/dissolution reactions with application to reactive flow in single phase hydrothermal systems, *Am. J. Sci.*, 294, 529–592, 1994.
- Steefel, C. I., and K. T. MacQuarrie, Approaches to modeling reactive transport in porous media, in *Reactive Transport in Porous Media*, edited by P. C. Lichtner, C. T. Steefel, and E. H. Oelkers, *Rev. Mineral.*, 34, 83–129, 1996.
- Steefel, C. I., and P. van Cappellen, Preface: Reactive transport modeling of natural systems, *J. Hydrol.*, 209, 1–7, 1998.
- Steefel, C. I., and S. B. Yabusaki, OS3D/GIMRT, Software for modeling multi-component-multidimensional reactive transport, user's manual and programmer's guide, *PNL-11166*, Pac. Northwest Lab., Richland, Wash., 1996.
- Stumm, W., and J. J. Morgan, *Aquatic Chemistry*, 3rd ed., pp. 533–534, John Wiley, New York, 1996.
- Suarez, D., and J. Šimunek, Solute transport modeling under variably saturated water flow conditions, in *Reactive Transport in Porous Media*, *Rev. Mineral.*, vol. 34, edited by P. C. Lichtner, C. I. Steefel, and E. H. Oelkers, pp. 229–268, Mineral. Soc. of Am., Washington, D. C., 1996.
- Szecsody, J. E., J. M. Zachara, and P. Bruckhart, Adsorption-dissolution reactions affecting the distribution and stability of Co(II)-EDTA in Fe-oxide coated sand, *Environ. Sci. Technol.*, 28, 1706–1716, 1994.
- Szecsody, J. E., J. M. Zachara, A. Chilakapati, P. M. Jardine, and A. S. Ferency, Importance of flow and particle-scale heterogeneity on Co(II)/III-EDTA reactive transport, *J. Hydrol.*, 209(1), 112–136, 1998.
- Tebes-Stevens, C., A. J. Valocchi, J. M. VanBriesen, and B. E. Rittmann, Multicomponent transport with coupled geochemical and microbiological reactions: Model description and example simulations, *J. Hydrol.*, 209, 8–26, 1998.
- Theis, T. L., D. J. Kirkner, and A. A. Jennings, Multi-solute subsurface transport modeling for energy solid wastes, *Tech. Progress Rep. C00-10253-3*, Dep. of Civ. Eng., Univ. of Notre Dame, Notre Dame, 1982.
- Tompson, A. F. B., Numerical simulation of chemical migration in physically and chemically heterogeneous porous media, *Water Resour. Res.*, 29(11), 3709–3726, 1993.
- Toride, N., F. J. Leij, and M. T. van Genuchten, A comprehensive set of analytical solutions for nonequilibrium solute transport with first-order decay and zero-order production, *Water Resour. Res.*, 29, 2167–2182, 1993.
- Truesdell, A. H., and B. F. Jones, WATEQ, A computer program for calculating chemical equilibria of natural water, *U.S. Geol. Surv. J. Res.*, 2, 233–248, 1974.
- van der Zee, S. E. A. T. M., and W. H. van Riemsdijk, Transport of reactive solute in spatially variable soil systems, *Water Resour. Res.*, 23(11), 2059–2069, 1987.
- Westall, J. C., J. L. Zachary, and M. F. Morel, MINEQL: A computer program for the calculation of chemical equilibrium composition of aqueous system, *Tech. Note 18*, 91 pp., Dep. of Civ. Eng., Mass. Inst. of Technol., Cambridge, 1976.

- Wolery, T. J., EQ3/6, a software package for geochemical modeling of aqueous systems: Package overview and installation guide, *UCRL-MA-110662 PT 1*, Lawrence Livermore Natl. Lab., Livermore, Calif., 1992.
- Wood, B. D., C. N. Dawson, J. E. Szecsody, and G. P. Streile, Modeling contaminant transport and biodegradation in a layered porous media system, *Water Resour. Res.*, *30*(6), 1833–1845, 1994.
- Yeh, G. T., and Y. Fang, BIOGEOCHEM 2-0: A numerical model to simulate biogeochemical reactions under multiple phase systems, technical report, Dep. of Civ. and Environ. Eng., Univ. of Central Fla., Orlando, 2002.
- Yeh, G. T., and K. M. Salvage, Modeling reactive chemical transport, paper presented at XIIth International Conference on Numerical Methods in Water Resources, Inst. of Chem. Eng. And High Temperature Chem. Processes-Found. For Res. and Technol., Crete, Greece, 15 –19 June 1998.
- Yeh, G. T., and V. S. Tripathi, A model for simulating transport of reactive multispecies components: Model development and demonstration, *Water Resour. Res.*, *27*, 3075–3094, 1991.
- Yeh, G. T., G. Iskra, J. M. Zachara, J. E. Szecsody, and G. P. Streile, KEMOD: A mixed chemical kinetic and equilibrium model of aqueous and solid phase geochemical reactions, *PNL-10380*, Pac. Northwest Lab., Richland, Wash., 1995.
- Yeh, G. T., K. Salvage, and W. H. Choi, Reactive multispecies-multicomponent chemical transport controlled by both equilibrium and kinetic reactions, paper presented at XIth Internatioanl Conference on Numerical Methods in Water Resources, Mexican Inst. of Water Technol., Canquan, Mexico, 22 –26 July 1996.
- Yeh, G. T., K. M. Salvage, J. P. Gwo, J. M. Zachara, and J. E. Szecsody, HYDROBIOGEOCHEM: A coupled model of hydrologic transport and mixed biogeochemical kinetic/equilibrium reactions in saturated-unsaturated media, technical report, Dep. of Civ. and Environ. Eng., Pa. State Univ., University Park, 1998.
- Yeh, G. T., W. D. Burgos, Y. L. Fang, and J. M. Zachara, Modeling biogeochemical kinetics: issues and data needs, paper presented at XIII International Conference on Computational Methods in Water Resources, Univ. of Calgary, Calgary, Alberta, Canada, 2000.
- Yeh, G. T., W. D. Burgos, and J. M. Zachara, Modeling and measuring biogeochemical reactions: System consistency, data needs, and rate formulation, *Adv. Environ. Res.*, *5*, 219–237, 2001a.
- Yeh, G. T., M. D. Siegel, and M. H. Li, Numerical modeling of coupled fluid flows and reactive transport including fast and slow chemical reactions, *J. Contam. Hydrol.*, *47*, 379–390, 2001b.

---

W. D. Burgos and Y. Fang, Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA 16802, USA.

G.-T. Yeh, Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816-2450, USA. (gyeh@mail.ucf.edu)