The Pennsylvania State University

The Graduate School

Department of Civil and Environmental Engineering

REACTIVE CHEMICAL TRANSPORT UNDER MULTIPHASE SYSTEM

A Thesis in

Civil Engineering

by

Yilin Fang

© 2003 Yilin Fang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2003

We approve the thesis of Yilin Fang.

Date of Signature

Gour-Tsyh Yeh Professor of Civil Engineering Thesis Advisor Co-Chair of Committee	
William D. Burgos Associate Professor of Environmental Engineering Co-Chair of Committee	
Christopher J. Duffy Professor of Civil Engineering	
Derek Elsworth Professor of GeoEnvironmental Engineering	
Andrew Scanlon Professor of Civil Engineering Head of the Department of Civil and Environmental Engineering	

ABSTRACT

This thesis presents the development of BIOGEOCHEM, a numerical model to simulate biochemical and geochemical reactions in a batch system and HYDROBIOGEOCHEM, a numerical model to simulate HYDROlogic transport and BIOchemical and GEOCHEMical reactions under nonisothermal multiphase systems in 2-dimensions and 3-dimensions.

HYDROBIOGEOCHEM is a coupled hydrologic transport and biogeochemical reaction code, reaction module of which is BIOGEOCHEM, to predict the spatio-temporal distributions of all the important chemical species. The formulation of BIOGEOCHEM and HYDROBIOGEOCHEM is new in that it is based on a general paradigm which uses the well accepted diagonalizetion-decomposition procedure. The unique features of the general paradigm are that it can 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.

The new formulation in HYDROBIOGEOCHEM uses Gauss-Jordan decomposition to diagonalize the governing matrix equations for hydrologic transport to reduce primary dependent variables (PDVs), resulting in mobile components and mobile kinetic variables as PDVs. Methods for coupling biogeochemical reaction and hydrologic transport, such as the sequential iteration approach (SIA) and predictor corrector approach are incorporated into the code to make the model versatile.

The governing transport equation can be written in conservative form and nonconservative form. Five different numerical schemes based on the two forms of equations are incorporated in HYDROBIOGEOCHEM to better understand different physical processes. They are (1) FEM on advective form of equation, (2) FEM on conservative form of equation, (3) Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on advective form of equation for boundary elements, (4) hybrid Lagrangian-Eulerian FEM, and (5) Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on conservative form of equation for boundary elements.

Heat transfer is considered in HYDROBIOGEOCHEM to account for the temperature variations that may impact hydrologic transport by affecting the hydrologic and chemical conditions in the subsurface system. Weak coupling will be applied to solve a system of chemical transport and heat transfer to save computing resources.

HYDROBIOGEOCHEM is designed to assess migration of subsurface contamination and help design remediation technologies. The prospective field application of the model is demonstrated by examples.

TABLE OF CONTENTS

LIST OF FIGURES……vii

LIST OF TABLES……..ix

ACKNOWLEDGEMENTS……...….x

CHAPTER 1. NTRODUCTION…….……1

1.1. BACKGROUND AND LITERATURE REVIEW …….... 1
1.2. OBJECTIVE AND FORMAT…….…….. 6
REFERENCES ……...8

CHAPTER 2. A GENERAL PARADIGM TO MODEL REACTION-BASED BIOGEOCHEMICAL PROCESSES IN BATCH SYSTEMS…….…….12

ABSTRACT ……..12
2.1. INTRODUCTION ……..... 15
2.2. REACTION NETWORK……... 19
2.3. DECOMPOSITION OF REACTION NETWORK ……...30
2.4. REACTION RATE FORMULATIONS ...…….47
2.5. NUMERICAL SOLUTION …..……50
2.6. EXAMPLE SIMULATIONS …….……54
2.7. CONCLUSION AND DISCUSSION ……... 66
APPENDIX 2.A ……69
APPENDIX 2.B …… 78

ACKNOWLEDGEMENTS……...83
REFERENCES…….. 84

CHAPTER 3. REACTIVE TRANSPORT UNDER MULTIPHASE SYSTEM: – MODEL FORMUATION AND NUMERICAL SCHEMES .....……..93

ABSTRACT……...93
3.1. INTRODUCTION…….. 95
3.2. PARADIGM REVIEW …..……97
3.3. GOVERNING EQUAITONS FOR TRANSPORT ……101
3.4. EXAMPLE SIMULATIONS ……...114
3.5. DISCUSSION AND CONCLUSIONS ……... 127
APPENDIX 3.A MATRIX EQUATION FOR DIFFERENT NUMERICAL SCHEMES……...129

ACKNOWLEDGEMENTS……135
REFERENCES…… 136

CHAPTER 4. 3-DIMENSIONAL NON-ISOTHERMAL REACTIVE TRANSPORT UNDER MULTIPHASE SYSTEM…….…….....143

ABSTRACT…….143
4.1. INTRODUCTION …….. 145
4.2. GOVERNING EQUATIONS FOR HYDROLOGIC TRANSPORT ……. 146
4.3. NUMERICAL IMPLEMENTATOIN ……..150
4.4. GOVERNING EQUATIONS FOR HEAT TRANSFER …… 153
4.5. EXAMPLE SIMULATIONS ……156
4.6. CONCLUSION …… 159
APPENDIX 4.A ADVECTIVE FORM EQUATONS FOR HYDROLOGIC TRANSPORT……..160
APPENDIX 4.B INITIAL AND BOUNDARY CONDITIONS FOR HYDROLOGIC TRANSPORT……...163
APPENDIX 4.C MATRIX EQUATIONS FOR DIFFERENT NUMERICAL SCHEMES……...167
APPENDIX 4.D ADVECTIVE FORM EQUATIONS FOR HEAT TRANSFER……..……...173

ACKNOWLEDGEMENTS…….174
REFERENCES…… 175

CHAPTER 5. SUMMARY AND FUTURE WORK…….. 191

5.1. SUMMARY……..191
5.2. FUTURE WORK ……192
REFERENCES…… 193

LIST OF FIGURES

Figure 2.1. Solution of the reaction network in table 2.1: Comparison of this study with results of *Chilakapati et al*.[1998]. (C_0 and C_1 are the initial concentrations of Co(II)EDTA(aq) and microorganisms, respectively……...……..55

Figure 2.2. Concentration curves for major species directly involving in the mineral dissolution…….……61

Figure 2.3. Concentration curves for main complexed and adsorbed species ……..61

Figure 2.4. Concentration curves for ion-exchanged species ……...62

Figure 3.1. Concentration profile of chemical C at t = 8 days: (a) α_L = 0.1 cm, (b) α_L = 10 cm, (c) α_L = 50 cm……117

Figure 3.2. Concentration profile for problem 3.4.2……118

Figure 3.3. Gaseous species concentration at selected times ……...……...121

Figure 3.4. Problem domain ……122

Figure 3.5. Concentration profile of different species in aqueous phase ……126

Figure 3.6. Contours of gaseous hydrocarbon (top, unit in Pa) and aqueous hydrocarbon (bottom, unit in mg/L) in the unsaturated and saturated zone…….126

Figure 4.1. Concentration profile comparison with Sun et al. (1999) at selected time steps.The solid line is from Sun et al., circle is from HYDROBIOGEOCHEM…….183

Figure 4.2. Temperature distribution after 8 days……..…….184

Figure 4.3. Problem domain (after UTCHEM, 2000)……..……...184

Figure 4.4. B_{NAPL} distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m……185

Figure 4.5. T_{NAPL} distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m…...186

Figure 4.6. X_s distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m……. 187

Figure 4.7. $O_{2(aq)}$ distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m…….188

Figure 4.8. $NO_3^-{}_{(aq)}$ distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m….189

Figure 4.9. Species concentration distribution with respect to time…….190

LIST OF TABLES

Table 2.1. Reaction network for the example problem in Section 2.6.1 (after *Chilakapati et al.* [1998]) ……..55

Table 2.2. Reaction network for the example problem in Section 2.6.2 ……...58

Table 2.3. Reaction network for the example problem in Section 2.6.3 …….. 65

Table 2.4. Initial and simulated species concentration ……. 66

Table 2.B.1. Matrix decomposition for the reaction network of the transformation of EDTA ……78

Table 2.B.2. Matrix decomposition for the complex (multiple reaction types) mixed (equilibrium and kinetic) reaction system ……80

Table 3.1. R^2 values based on the analytical and calculated results for different values of dispersivity…….....115

Table 3.2. Processes considered in problem 3.4.2…….. 118

Table 3.3. Physical and Chemical Parameters and values for Problem 3.4.2 ……119

Table 3.4. R^2 values based on the analytical and calculated results for selected time steps……..119

Table 3.5. Physical and Chemical Parameters and values for Problem 3.4.3……...…...120

Table 3.6. Physical and Chemical Parameters and values for Problem 3.4.4……..123

Table 3.7. Processes considered in problem 3.4.4……... 124

Table 4.1. Processes considered in problem 4.1…….. 178

Table 4.2. System parameters and values…… 178

Table 4.3. Flow Parameters and Values…...……179

Table 4.4. Biogeochemical Processes……. 180

Table 4.5. Reaction Rates…… 181

Table 4.6. Simulation parameters...……. 182

ACKNOWLEDGEMENTS

I am grateful for the chance to study under the guidance of Dr. Gour-Tsyh Yeh and his ongoing financial support through my graduate study. Dr. Yeh's way of research has impressed me and will greatly help me in my future career. His ideas and input has enlightened me in this work.

I would like to thank Dr. William D. Burgos for the chance to work with him. I would also like to thank my other committee members, Dr. Christopher J. Duffy and Dr. Derek Elsworth for their time and effort on this thesis.

I would like to thank the Department of Civil and Environmental Engineering at University of Central Florida. UCF has provided me computing resources and other facilities during my thesis research there for almost two years.

I am indebted to my family for their support and encouragement that have enabled me to successfully complete this work. Especially, I am grateful to my husband Tao for his love and encouragement during my study. I would also like to thank my daughter Michelle who has brightened my life.

CHAPTER 1.
INTRODUCTION

This thesis presents the numerical modeling of biochemical and geochemical reactions in batch systems and non-isothermal reactive chemical transport under multiphase systems. Included in the transport model, HYDROBIOGEOCHEM, (a numerical model to simulate HYDROlogic transport and BIOchemical and GEOCHEMical reactions under nonisothermal multiphase systems) are transport processes including advection, dispersion, and molecular diffusion, and chemical processes including aqueous complexation, adsorption-desorption, ion-exchange, oxidation-reduction, precipitation-dissolution, acid-base reactions, and microbial mediated reactions. Chemical processes are modeled with BIOGEOCHEM, a batch numerical speciation model. Reaction rates considered in the BIOGEOCHEM is the most comprehensive so far. Because of this broad modeling scope, HYDROBIOGEOCHEM has a prospective to be applied to a wide range of subsurface contamination problems, and to provide a powerful tool for mechanistically understanding biogechemical processes under transport in natural systems and proper designing of remediation technologies.

1.1 BACKGROUND AND LITERATURE REVIEW

Increasing human activities have induced potential adverse effect on the environment, particularly the contamination of the groundwater. General sources of groundwater contamination include accidental spills and leaks, mining, salt intrusion and

others (Chareneau, 2000). Once the contaminants are introduced into the subsurface, they will pose a potential threat to large volumes of groundwater flowing through porous media under the surface of the earth.

Typical porous medium contains solid phase, water phase and gas phase, and there may be a number of natural chemical species in each phase. In the case of contamination, the medium also contains foreign chemicals – contaminants, which can end up in different phases during their migration processes as the result of the complex interplay among hydrological, physical, geochemical, and biochemical processes in the subsurface environment.

Once in contact with soil, a contaminant composed of several chemical species enters the soil and migrates downward through the unsaturated zone under gravity. In some circumstances, the contaminant may be dissolved as a portion of the aqueous phase. When the source of contaminant is associated with leaking underground tanks, it will be in the form of nonaqueous phase liquid (NAPL) because the majority of these tanks store petroleum products with low solubility in water. Ttrichloroethylene (TCE), for example, has a solubility of 1100 ppm in water at 20^{o}C (Pinder and Abriola, 1986). In addition, light components of the contaminant may vaporize into the soil air. Some contaminant will be trapped in the matrix of the porous media, while others will eventually reach the water table and migrate as part of the groundwater.

The migration of a chemical contaminant in the subsurface is a very complex process since its release to the subsurface environment. It will be controlled by hydrological processes, transport processes such as advection and dispersion, and biogeochemical processes such as precipitation-dissolution, adsorption, and aqueous

speciation reactions (Viswanathan et al., 1998). Therefore, the assessment of subsurface contamination and proper design of remediation technologies requires accurate model of the transport of contaminants in subsurface system.

A fundamental question to the transport of the chemical species under the multiphase system is how the concentrations of each chemical species within different phases relate to each other, the understanding of which is known as reactive chemical transport (Cheng, 1995, and reference therein). The simplest and most common approach is the local equilibrium assumption (Charbeneau, 2000). This approach assumes that the rate of mass transport through the porous media within a phase is slow compared to the rate of mass transfer between phases in contact locally, i.e., the species in different phases are in thermodynamic equilibrium. However, experimental evidence has shown that this assumption is not always tenable (Friedly and Rubin, 1992, and references therein).

Various ways have been attempted by researchers to solve the reactive transport problems. Hunter et al. (1998) classified and reviewed models of reactive transport developed till 1996 in subsurface environments. Most of these biogeochemical models are based on local equilibrium assumption due to computational limitations and the lack of database in kinetic rate constants. This assumption may provide a good approximation for many homogeneous speciation, acid-base and adsorption reactions; however, it may not reflect realistic spatio-temporal distributions of chemical species in subsurface systems when dealing with slow reactions, such as mineral dissolution/precipitation and oxidation/reduction (Hunter et al., 1998).

In the past few years, there have been considerable efforts to deal with the kinetically controlled, multi-component transport in the development of reactive chemical

transport in the subsurface systems. We have witnessed the rapid evolution of models for the analysis of reactive chemical transport with the growing thermodynamic and kinetic databases. These models include PHREEQC (Parhurst, 1995), an ion-association aqueous model designed for low-temperature aqueous geochemical calculations; RAFT (Chilakapati, 1995; Chilakapati et al. 2000), a tool for arbitrarily complex coupled kinetic-equilibrium heterogeneous reaction networks; OS3D (Steefel and Yabusaki, 1996), a 3D model designed for multicomponent, multispecies chemistry, considering multiscale heterogeneities; HYDROGEOCHEM (Cheng and Yeh, 1998), a model designed for the simulation of reactive multispecies-multicomponent chemical transport through saturated-unsaturated media under non-isothermal conditions; BIORXNTRN (Hunter et al., 1998), a model focusing on the comprehensive kinetic reaction network for the multi-component biogeochemical dynamics of groundwater systems; BIOKEMOD (Salvage and Yeh, 1998), a model designed for simulating any mixture of geochemical and mocriobiological reactions in subsurface systems; TRANQUI (Xu, 1999), a 2D model dealing with thermo-hydro-geochemical problems for single phase saturated-unsaturated porous media flow systems; LEHGC2.0 (Yeh et. al, 2001a), a model for coupled fluid flow and reactive chemical transport; and some others. There were several articles on the application and development of reactive chemical transport models in a special issue of Journal of Hydroloy (Volume 209).

Even though there is a large reservoir of numerical models available as mentioned above, numerical efficiency for modeling reactive chemical transport is still of major concern because of the complex nature of transport and reaction processes and their interactions. Usually, flow and transport can be treated in separate steps. However, from

a physical point of view, the chemicals contained in the subsystem affect the solution density and induce changes in the flow field, which becomes dependent not only on the hydrogeologic parameters of the system, but also on the chemical concentration. This interdependence produces a coupling between the flow and transport equations. Secondly, as the chemical part of reactive transport models becomes more complex (considering mixed equilibrium and kinetic reactions), a challenge is posed on numerical formulations that can solve the resulting governing equations efficiently.

Method to couple chemistry codes to transport models varies according to the particular codes. Some chemical codes have been extended to include solute transport, for example, the geochemical code PHREEQE has been expanded to account for one-dimensional advection, and equilibrium code MINTEQ2 has been coupled with the two-dimensional advective-dispersive transport code PLUME2D, as summarized by Hunter et al. (1998). Some chemical codes are iteratively coupled with transport codes, such as HYDROGEOCHEM (Yeh and Tripathi, 1991).

Yeh and Tripathi (1989) reviewed and discussed a number of approaches used in different transport models and compared the computation effort of each approach with regard to the formulation of governing equations and the types of reactions considered. Steefel and MacQuarrie (1996) also reviewed several approaches for accuracy and computational efficiency. However, the comparison of different approaches is still an active topic in the literature (for example, Saaltink et al., 2001). In the meantime, there are new techniques being proposed other than the three general categories: operator splitting, the global implicit method and sequential iteration methods, such as ‘selective coupling method’ in FEHM (Robinson et al., 2000).

While reactive chemical transport in a single flow system is still a subject of intensive research, some researchers have already initiated investigations in multiphase systems (for example, Lichtner and Seth, 1996; Viswanathan et al., 1998; Xu and Pruess, 2001). Limits of applicability of some of the models were addressed by Wu and Pruess (2000), for example, the assumption of isothermal system in the models. However, temperature variations may impact transport by affecting the hydrologic and chemical conditions. Therefore, heat transfer should be considered in the models to account for the temperature effect.

Of all the models cited above, most have the limitation of assuming that all biogeochemical reactions can be easily written in basic (canonical) forms, which is not obvious when there are many parallel kinetic reactions (Yeh et al., 2000; Yeh et al., 2001b). And this limitation will certainly affect the generality of these models. On the other hand, the rate formulation for biogeochemical reactions is a primary challenge in biogeochemical modeling. *Ad hoc* approach dominates in current literature. It is not able to capture key features of a natural system via several measurable parameters. Therefore, a reaction-based formulation, though difficult to attain, is in need to alleviate the problem.

1.2 OBJECTIVE AND FORMAT

This thesis is to develop a non-isothermal, reactive chemical transport model on a reaction-based formulation under the multiphase system. The ultimate goal is to develop the capability of the model to calculate the spatio-temporal distributions of all the

important chemical species, reaction rates in the system. The organization of the thesis is as follows: A theoretical description of the batch model BIOGEOCHEM will be presented first. Followed by the derivation of the governing equations for reactive transport based on a new formulation and heat transfer under multiphase system, the techniques involved in transport calculations will be discussed. The thesis is made up of three journal articles. Chapter 2 presents the development, and application of the batch biogeochemical reaction model BIOGEOCHEM; this paper has been accepted for publication by *Water Resources Research*. Chapter 3 presents the formulation and different numerical schemes of HYDROBIOGEOCHEM, a reactive transport model; this paper is intended for submission to *Water Resources Research*. Chapter 4 presents the formulation of a three dimensional non-isothermal reactive transport model, 3DHYDROBIOGEOCHEM; this paper is intended for submission to *the Journal of Hydrology.* Chapter 5 summarizes the work presented in this thesis and outlines the limits of the study.

REFERENCES

Charbeneau, R.J., 2000. *Groundwater Hydraulics and Pollutant Transport.* Prentice Hall, New Jersey.

Cheng, H.P., 1995. Development and application of a three-dimensional finite element model of subsurface flow, heat transfer, and reactive chemical transport. Dissertation, The Pennsylvania State University.

Cheng, H.P. and Yeh, G.T., 1998. Development and demonstrative application of a 3-D numerical model of subsurface flow, heat transfer, and reactive chemical transport: 3DHYDROGEOCHEM. J. Contami. Hydrol. 34, 47-83.

Chilakapati, A., 1995. RAFT: A simulator for ReActive Flow and Transport of groundwater contaminants. PNL Report 10636, Pacific Northwest Laboratory, Richland, WA.

Chilakapati, A., Yabusaki, S., Szecsody, J. and MacEvoy, W., 2000. Groundwater flow, multicomponent transport and biogeochemistry: development and application of a coupled process model. J. Contami. Hydrol. 43, 303-325.

Friedly, J.C., Rubin, J., 1992. Solute transport with multiple equilibrium-controlled or kinetically controlled chemical reactions. Water Resources Research, 28, 1935-1953.

Hunter, K.S., Wang, Y., and Frind, E.O., 1998. Modeling the effect of chemical heterogeneity on acidification and solute leaching in overburden mine spoils. J. Hydrol. 209, 166-185.

Lichtner, P.C., Seth, M., 1996. Multiphase-multicomponent nonisothermal reactive transport in partially saturated porous media. In: Proceedings of the International Conference on Deep Geological Disposal of Radioactive Waste, Canadian Nuclear Society, September 16-19, Winnnipeg, Manitoba, Canada, 3-133-42pp.

Parkhurst, D.L., 1995. User's guide to PHREEQC – a computer program for speciation, reaction-path, advective transport, and inverse geochemical calculations. US Geological Survey Water Resources Investigations Report 95-4227.

Pinder, G.F. and Abriola, L.M., 1986. On the simulation of nonaqueous phase organic compounds in the subsurface. Water Resour. Res. 22, 109S-119S.

Robinson, B.A., Viswanathan, H.S. and Valocchi, A.J., 2000. Efficient numerical techniques for modeling multicomponent ground-water transport based upon simultaneous solution of strongly coupled subsets of chemical components. Advances in Water Resources. 23, 307-324.

Saaltink, M.W., Carrera, J. and Ayora, C., 2001. On the behavior of approaches to simulate reactive transport. J. Contami. Hydrol. 48, 213-235.

Salvage, K., Yeh, G.T., 1998. Development and application of a numerical model of kinetic and equilibrium microbiological and geochemical reactions (BIOKEMOD). J. Hydrol. 209, 27-52.

Steefel, C.I., MacQuarrie, K.T., 1996. Approaches to modeling reactive transport in porous media. In: Lichtner, P.C., Steefel, C.T., Oelkers, E.H. (Eds), Reactive Transport in Porous Media, Reviews in Mineralology 34. Mineral. Soc. Am. 83, 83-129.

Steefel, C.I., Yabusaki, S.B., 1996. OS3D/GIMRT, Software for modeling multi-

component-multidimensional reactive transport, user's manual and programmer's guide. PNL-11166, Pacific Northwest Lagoratory, Richland, WA.

Viswanathan, H.S., Robinson, B.A., Valocchi, A.J., and Triay, I.R., 1998. A reactive transport model of neptunium migration from the potential repository at Yucca Mountain. J. Hydrol. 209, 251-280.

Wu, Y.S. and Pruess, K., 2000. Numerical simulation of non-isothermal multiphase tracer transport in heterogeneous fractured porous media. Advances in Water Resour. 23, 699-723.

Xu, T., Samper, J., Ayora C., Manzano, M. and Custodio, E., 1999. Modeling of non-isothermal multi-component reactive transport in field scale porous media flow systems. J. Hydro. 214, 144-164.

Xu, T. and Pruess, K., 2001. Modeling multiphase non-isothermal fluid flow and reactive transport in variably saturated fractured rocks: 1. methodology. American journal of science, Vol. 301(1), 16-33.

Yeh, G.T., Tripathi, V.S., 1989. A critical evaluation o frecent developments of hydrogeochemical transport models of reactive multichemical components. Water Resour. Res. 25(1), 93-108.

Yeh, G.T., Tripathi, V.S., 1991. HYDROGEOCHEM: A coupled model of hydrologic transport and geochemical equilibrium in reactive multicomponent systems. ORNL-6371, Oak Ridge National Laboratory, Environmental Sciences Division.

Yeh, G.T., Burgos W.D., Fang, Y. and Zachara J.M., 2000. Modeling biogeochemical kinetics: issues and data needs. XIII International Conference on Computational Methods in Water resources, 2000.

Yeh, G.T., Siegel, M.D. and Li, M.H., 2001a. Numerical modeling of coupled variably saturated fluid flow and reactive transport with fast and slow chemical reactions. J. Contami. Hydrol. 47, 379-390.

Yeh, G.T., Burgos, W.D., and Zachara, J.M., 2001b. Modeling and measuring biogeochemical reactions: system consistency, data needs, and rate formulation. Advances in Environmental Research. 5, 219-237.

CHAPTER 2.
A GENERAL PARADIGM TO MODEL REACTION-BASED BIOGEOCHEMICAL PROCESSES IN BATCH SYSTEMS

Yilin Fang[1] , Gour-Tsyh Yeh[2*] and William D. Burgos[1]

[1]Department of Civil and Environmental Engineering
The Pennsylvania State University
University Park, PA 16802

[2]Department of Civil and Environmental Engineering
University of Central Florida
Orlando, FL 32816-2450

ABSTRACT

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, users-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 users-specified algebraic equation. Programmed kinetic reaction rates include multiple Monod kinetics, n-th order empirical, and elementary formulations. In addition, users-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 non-uniqueness 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.

Keywords: Geochemical Modeling; Biochemical Modeling; Reaction Network; Matrix Decomposition; Basis Species Switching; Batch Systems.

2.1 INTRODUCTION

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 [Steefel 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.

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.

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

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

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.

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 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.

A 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 Erikson, 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; Steefel 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).

Models coupling transport with mixed equilibrium/ kinetic reactions have appeared since the mid-1990s (e.g., Steefel and Lagasa, 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; Yeh and Salvage, 1998; Yeh et al., 2001b; Brun and Engesgaard, 2002). For most models chemical components must be selected *a prior* such that only limited reaction network can be considered (e.g., Parkurst, 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 (Steefel 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).

2.1.1 OUTLINE

This paper is organized as follows. In section 2.2.2, 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 2.2.3, 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 2.2.4, reaction rate formulations for kinetic reactions are discussed. In section 2.2.5, the numerical solution of the governing equations is outlined. In section 2.2.6, the general application of BIOGEOCHEM is presented using three demonstrative examples. In section 2.2.7, conclusions and discussions from the current research are summarized.

2.2 REACTION NETWORK

2.2.1 Definition

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

$$\frac{dC_i}{dt} = r_i|_N, \ i \in M \tag{1}$$

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 Eq. (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 Eq. (1), one employs the set of transport equations. A subsequent paper will present the development of a biogeochemical transport model (HYDROBIOGEOCHEM) using the same paradigm described here (Fang and Yeh, 2002).

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

$$\frac{dC_i}{dt} = r_i|_N = f_i(C_1, C_2, ..., C_M; p_1, p_2, ...) \tag{2}$$

where f_i is the empirical function for the *i*-th species and p_1, p_2, ...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 study under different environmental conditions when a different set of reactions is the contributing chemical processes.

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

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

where, ν_{ik} is the reaction stoichiometry of the i-th species in the k-th reaction associated with the products, μ_{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: $\frac{\partial C_i}{\partial t} + L(C_i) = r_i|_N = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k$.

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.

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.

2.2.2 Example

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 in Yeh et al., 1998):

(R1) $H + NTA \leftrightarrow HNTA$, an equilibrium reaction with $R_1 = \infty$;

(R2) $CoNTA \leftrightarrow Co + NTA$, an equilibrium reaction with $R_2 = \infty$;

(R3) $CoNTA + H \leftrightarrow Co + HNTA$, a kinetic reaction, with its rate R_3 that must be formulated;

(R4) $HNTA + H \rightarrow B$, irreversible microbial degradation reaction with its rate R_4 that must be formulated.

where R1, R2, R3, and R4 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 R1

and R2.

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 $\leftrightarrow$ 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: $\frac{d[A]}{dt} = k^b\left([C] - K^e[A][B]\right);$ $\frac{d[B]}{dt} = k^b\left([C] - K^e[A][B]\right);$ and $\frac{d[C]}{dt} = -k^b\left([C] - K^e[A][B]\right),$ 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 $R = -\lim_{k^b \Rightarrow \infty} k^b\left([C]_e - K^e[A]_e[B]_e\right)$ 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 $R = - \lim_{k^b \Rightarrow \infty} k^b \left([C]_e - K^e [A]_e [B]_e \right)$ *a priori*]. $R = \infty$ is a mathematical abstraction.

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 $R = - \lim_{k^b \Rightarrow \infty} k^b \left([C]_e - K^e [A]_e [B]_e \right)$ 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.

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 $R = - \lim_{k^b \Rightarrow \infty} k^b \left([C]_e - K^e [A]_e [B]_e \right)$. 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 $R = - \lim_{k^b \Rightarrow \infty} k^b \left([C]_e - K^e [A]_e [B]_e \right)$ but must be computed with $R = \frac{d[C]}{dt} = \frac{[C]_e - [C]_o}{\Delta t = 0} = \infty$. Therefore, the argument that the rate of an equilibrium reaction is indeterminate is flawed when a thermodynamic approach is used to model the reaction.

To distinguish the conceptual difference between kinetic modeling and thermodynamic modeling of a fast reaction, the equation, $R = - \lim_{k^b \Rightarrow \infty} k^b \left([C]_e - K^e [A]_e [B]_e \right)$ is called a mass-action kinetic equation while the

equation, $R = \infty \ni [C] = K^e [A] [B]$, 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.

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 R1 through R4 can be written as:

$$\frac{d[H]}{dt} = - R_1 - R_3 - R_4 \tag{4.1}$$

$$\frac{d[NTA]}{dt} = - R_1 + R_2 \tag{4.2}$$

$$\frac{d[HNTA]}{dt} = R_1 + R_3 - R_4 \tag{4.3}$$

$$\frac{d[Co]}{dt} = R_2 + R_3 \tag{4.4}$$

$$\frac{d[CoNTA]}{dt} = - R_2 - R_3 \tag{4.5}$$

$$\frac{d[B]}{dt} = R_4 \tag{4.6}$$

2.2.3 Difficulties in Solving Eqs. (4.1) through (4.6)

An analytical solution of Eqs. (4.1) through (4.6) is generally impossible, which makes numerical integration attractive. The majority of existing reactive chemical models has taken either a primitive approach (i.e., integrate Eq. (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 Eqs. (4.1) through (4.6). 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 Eqs. (6.1) and (6.2)], an ordinary differential equation governing the slow/kinetic reaction [refer to Eq. (6.3)], and three component equations governing three mass conservation of TOT_H, TOT_{Co}, and TOT_{NTA} [refer to Eqs. (6.4), (6.5), and (6.6)]. 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).

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 Eqs. (4.1) through (4.5), 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 Eqs. (4.1) through (4.6) 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 Eqs. (6.4) through (6.6), 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.

Third, if more than one equilibrium reactions are involved, there is no way to define the subtraction or addition of infinity. For example, in Eq. (4.2), 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 these irrelevant 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 coincidently come up with true rate formulations and parameters. Therefore, it is desirable to exclude these irrelevant reactions with a systematic approach.

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.

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.

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, et al., 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.

2.3 DECOMPOSITION OF REACTION NETWORK

Eqs. (4.1) through (4.6) 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 users- 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 (i) no more than one fast reaction is allowed in any reduced ODE and (ii) 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 non-removed 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.

2.3.1 Description of reaction network decomposition

As stated in section 2.2.2, to make possible the formulation of rate equations one reaction by one reaction and to enable robust and efficient numerical integration of Eq.

(3), a systematic approach rather than the primitive or DAE approaches is necessary. The approach can be achieved with a diagonalization of Eq. (3). Eq. (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; Steefel 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 2.2.2 is used to illustrate the decomposition and demonstrate how the difficulties presented earlier can be overcome.

The system of Eqs. (4.1) through (4.6) written in matrix form is

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

The decomposition of Eq. (5) is not unique. Step-by-step decomposition procedures for this simple example are given in Appendix 2.A. For complex systems, the decomposition procedures are provided within the BIOGEOCHEM preprocessor (description in section 2.2.3.2). Two different decompositions are described and compared below.

2.3.1.1 Decomposition I

One possible decomposition, when choosing Co, H, and NTA as chemical components, results in the following set of 6 equations for the 6 species (Appendix 2.A):

$$\frac{d([HNTA]+[B])}{dt} = R_1 + R_3 : \frac{d([HNTA]+[B])}{dt} \approx R_1 = \infty \qquad \exists\ [HNTA] = K_1^e [H][NTA] \tag{6.1}$$

$$\frac{d[CoNTA]}{dt} = -R_2 - R_3 : \frac{d[CoNTA]}{dt} \approx -R_2 = -\infty \qquad \exists\ [Co][NTA] = K_2^e [CoNTA] \tag{6.2}$$

$$\frac{d[B]}{dt} = R_4 \tag{6.3}$$

$$\frac{d([H]+[HNTA]+2[B])}{dt} = 0 \Rightarrow TOT_H = [H]+[HNTA]+2[B] = \text{const} \tag{6.4}$$

$$\frac{d([NTA]+[HNTA]+[CoNTA]+[B])}{dt} = 0 \qquad \Rightarrow TOT_{NTA} = [NTA]+[HNTA]+[CoNTA]+[B] = \text{const} \tag{6.5}$$

$$\frac{d([Co]+[CoNTA])}{dt} = 0 \Rightarrow TOT_{Co} = [Co]+[CoNTA] = \text{const} \tag{6.6}$$

The two variables ([HNTA] + [B]) and [CoNTA] in the differential operator in Eqs. (6.1) and (6.2) are defined as equilibrium variables. The variable [B] in Eq. (6.3) is defined as the kinetic variable. The three variables, ([H] + [HNTA] + 2[B]), ([NTA] + [HNTA] + [CoNTA] + [B]), and ([Co] + [CoNTA]), in Eqs. (6.4) through (6.6) are defined as component variables. It should be noted that R_3 appears in both Eqs. (6.1) and (6.2) because R3 is linearly dependent on R1 and R2.

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.

Now let us discuss how the diagonalization procedure (i.e., decomposition of

reaction matrix) overcomes the difficulties stated in section 2.2.2. **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 Eq. (6.1), 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 R1, which results in the last equation in Eq. (6.1), 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 2.2.2.2. Similarly, $R_2 = \infty$ results in the last equation in Eq. (6.2). 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 Eqs. (4.1) through (4.6) since either more than one infinite rate appears in one equation [e.g., R_1 and R_2 in Eq. (4.2)] or one infinite rate appears in more than one equation [e.g., R_1 in Eqs. (4.1), (4.2), and (4.3) or R_2 in Eqs. (4,2), (4,4), and (4.5)].

Second, there is no reaction rate appearing in Eqs. (6.4), (6.5), and (6.6). 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 Eqs. (6.1) through (6.6) 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.

Fourth, R3 appears in both Eqs. (6.1) and (6.2). 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, than 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 Eqs. (6.1) and (6.2). 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 2.2.2 have now been resolved; thus the advantages of the diagonalization-decomposition approach over the

primitive approach should become clear.

Finally, the mass-action equilibrium equations, the last equations in Eqs. (6.1) and (6.2), can be decoupled from Eqs. (6.3) through (6.6) by eliminating [HNTA] and [CoNTA] (the procedure of elimination is given in section 2.3.2.2 and a simple example is given in Section 2.A.3 in Appendix 2.A). In other words, after the substitution of the last equations of Eqs. (6.1) and (6.2) into Eqs. (6.3) through (6.6), [H], [Co], [NTA], and [B] can be obtained by solving three linear algebraic equations [Eqs. (6.4) through (6.6)] and one ODE [Eq. (6.3)]. Then [HNTA] and [CoNTA] can be simply calculated using the last equations of Eqs. (6.1) and (6.2), 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.

2.3.1.2 Decomposition II

A second decomposition, when NTA, CoNTA, and B are chosen as the components, yields the following six equations (Appendix 2.A).

$$\frac{d([NTA]-[Co])}{dt} = -R_1 - R_3 : \frac{d([NTA]-[Co])}{dt} \approx -R_1 = -\infty \qquad \exists\ [HNTA] = K_1^e [H][NTA] \tag{7.1}$$

$$\frac{d[Co]}{dt} = R_2 + R_3 : \frac{d[Co]}{dt} \approx R_2 = \infty \qquad \exists\ [Co][NTA] = K_2^e [CoNTA] \tag{7.2}$$

$$\frac{d([H]-[NTA]+[Co])}{dt} = -R_4 \tag{7.3}$$

$$\frac{d(-[H]+2[NTA]+[HNTA]-2[Co])}{dt} = 0 \qquad \Rightarrow\ Tot_{HNTA} = [HNTA]+2[NTA]+[HNTA]-2[Co] = \text{const} \tag{7.4}$$

$$\frac{d([Co]+[CoNTA])}{dt} = 0 \Rightarrow Tot_{CoNTA} = [CoNTA]+[Co] = const \tag{7.5}$$

$$\frac{d([B]+[H]-[NTA]+[Co])}{dt} = 0$$
$$\Rightarrow Tot_B = [B]+[Co]+[H]-[NTA] = const \tag{7.6}$$

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 [Eq. (6.3)]. 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 non-uniqueness of the matrix decomposition may allow employment of appropriate governing equations for easier formulations of rate equations (Burgos, et al., 2002a, 2002b) 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., Eq. (6.3)] or a combination of species [e.g., Eq. (7.3)] depending on the reaction matrix decomposition.

Recall that in Decomposition I, the mass-action equilibrium equations, Eqs. (6.1) and (6.2), were decoupled from Eqs. (6.3) through (6.6) 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 Eqs. (7.1) and (7.2) from Eqs. (7.3) through (7.6) 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 2.3.2.2.

2.3.1.3 Other Advantages of Decomposition

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 Eqs. (6.3) or (7.3), 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, Eq. (4.6) and (6.3) are quite different. In Eq. (4.6), [B] denotes the concentration of the species B due to all reactions (in this case just R3) while in Eq. (6.3), [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 new paradigm.

2.3.2 BIOGEOCHEM Preprocessor

2.3.2.1 Determines the Number of Linearly Independent Reactions

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 is 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.

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.

All reactions are initially indexed with 0 in the BIOGEOCHEM preprocessor. During the reduction process, the index of every linearly dependent reaction is overwritten with a non-zero 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 Eq. (A4.3), R3 depends on R1 (keep in mind, R1 has been

already chosen an independent reaction) because R_3 and R_1 show up in the same equation. Similarly, R3 depends on R2 because both R_3 and R_2 show up in Eq. (A4.5). Thus, R3 depend on both R1 and R2. Because R2 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.

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 non-zero 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 users-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.

2.3.2.2 Selects Master and Secondary Species

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 users-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 users-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 reactions are modeled with users-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 involves more than one product species.

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 users-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^e values of fast reactions. Users must specify one species as a master species for this reaction.

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 users-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 users-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 non-zero 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 non-pivoting 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 2.A using a simple example.

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 users-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.

2.3.2.3 Generates Three Subsets of Governing Equations

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

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

where **B** is the reduced unit matrix, **D** is the diagonal matrix representing a submatrix of the reduced reaction-matrix with size of N_I x N_I reflecting the effects of N_I linearly independent reactions on the production-consumption rate of all kinetic-variables, **K** is a submatrix of the reduced reaction-matrix with size of N_I x 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 x N_I, and $\mathbf{0}_2$ is a zero matrix representing a submatrix of the reduced reaction-matrix with size N_C x N_{KD}. An example illustrating this form of the equation is given for Decomposition I in Appendix 2.A.

The decomposition of Eq. (3) to Eq. (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 users-specified algebraic equation (cf. Section 2.4); 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 -

$$\frac{dE_i}{dt} = D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j,\ k \in N_E,\ i \in M:\ \frac{d}{dt}\left(\frac{E_i}{D_{kk}}\right) \approx R_k = \infty$$
$$\exists \text{ a thermodynamically consistent equation} \tag{9}$$

Governing Equations for (N_I - N_E) Kinetic Variables –

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

and *Mass Conservation Equations for N_C Chemical Components –*

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

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 Eq. (9) is called an equilibrium variable, the variable E_i in Eq. (10) is called a kinetic variable, and the variable E_i in Eq. (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.

The unique feature of the diagonalization is that there is only one linearly independent kinetic reaction appearing on the right-hand side of Eq. (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., 2002a, 2002b). Without a careful design of experiments to exclude linearly dependent kinetic reactions, only lumped kinetic rates can be formulated and characterized (Burgos et al., 2002a, 2002b).

2.4. REACTINO RATE FORMULATIONS

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

$$\sum_{i=1}^{M} \mu_{ik} G_i \Leftrightarrow \sum_{i=1}^{M} \nu_{ik} G_i, \; k \in N \tag{12}$$

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

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

in which

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

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, γ_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 (Steefel and van Cappellen, 1998, Yeh et al. , 2001a). To make BIOGEOCHEM completely general, for any non-elementary kinetic reaction, its rate may be users-specified as follows:

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

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,… are rate parameters used to fit experimental data. Equation (15) looks like Eq. (2) at first glance, but they render different meanings. In an ad hoc approach, Eq. (2) represents the empirical rate of change of a species due to all reactions. In a reaction-based approach, Eq. (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.

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

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

Conceptually, the use of Eqs. (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 2.2) and its reactants and products satisfy the mass-action equilibrium equation instantaneously. The last equation in Eq. (16) is not a rate formulation because the concept of infinite 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, Eq. (16), at equilibrium. Even at the slightest deviation from the equilibrium, the rate as defined by Eq. (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 [Eqs. (4.1) to (4.6) for the example given in Section 2.2] infinitely stiff. Rather, it is better to consider the rate of a fast reaction infinity as a mathematical abstraction [i.e. Eq. (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 equation or a users-specified algebraic equation.

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

$$\begin{aligned} \frac{dE_i}{dt} &= D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j, \; k \in N_E, \; i \in M: \\ \frac{dE_i}{dt} &\approx D_{kk} R_k = \infty \;\; \exists \; F_k(C_1, C_2, ..., C_M; p_1, p_2, ...) = 0 \end{aligned} \tag{18}$$

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.

2.5 NUMERICAL SOLUTION

After the reaction matrix is decomposed in BIOGEOCHEM, equations of the type of Eq. (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 Eq. (16) can be easily eliminated). On the other hand, users-appointed species in equations of the type of Eq. (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 variable), one set of users-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 reaction, 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).

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 Westall et al. (1976) and Steefel and MacQuarrie (1996).

Since the concentration of some species may be several orders of magnitude smaller than others in a component equation, mass non-balance or non-convergence 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 the swap 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.

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

$$x_i = \alpha_i \prod_j c_j^{a_{ij}} \tag{19}$$

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 α_i given by

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

is the modified stability constant.

Suppose c_m is the component to be switched with x_k , then from Eq. (19), for i = k

$$x_k = \alpha_k \prod_{j \neq m} c_j^{a_{kj}} \, c_m^{a_{km}} \tag{21}$$

from which we can obtain

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

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

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

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

$$x_i' = \alpha_i' \prod_j c_j'^{\, a_{ij}'} \tag{24}$$

in which for i = k:

$$\begin{gathered} x_i' = c_m ; \quad \alpha_i' = \alpha_k' = \alpha_k^{-\frac{1}{a_{km}}} = \frac{1}{\alpha_k^{a_{km}'}} ; \\ a_{ij}' = a_{kj}' = \frac{-a_{kj}}{a_{km}} = -a_{kj} a_{km}' \quad \text{and} \quad c_j' = c_j \quad \text{if} \quad j \neq m ; \\ a_{ij}' = a_{km}' = \frac{1}{a_{km}} \quad \text{and} \quad c_j' = x_k \quad \text{if} \quad j = m \end{gathered} \tag{25}$$

for $i \neq k$:

$$\begin{gathered} x_i' = x_i ; \quad \alpha_i' = \alpha_i \alpha_k^{\frac{-a_{im}}{a_{km}}} = \alpha_i (\alpha_k')^{a_{im}} ; \\ a_{ij}' = a_{ij} - \frac{a_{kj} a_{im}}{a_{km}} = a_{ij} + a_{im} a_{kj}' \quad \text{and} \quad c_j' = c_j \quad \text{if} \quad j \neq m ; \\ a_{ij}' = a_{im}' = \frac{a_{im}}{a_{km}} = a_{im} a_{km}' \quad \text{and} \quad c_j' = x_k \quad \text{if} \quad j = m \end{gathered} \tag{26}$$

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.

2.6 EXAMPLE SIMULATIONS

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.

2.6.1 Mixed Microbiological and Chemical Kinetics

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 2.1. 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_2 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.

Table 2.1. Reaction network for the example problem in Section 2.6.1
(after *Chilakapati et al.* [1998])

Reaction	#	Reaction constants
$Co(II)(aq) + S_{neg} \leftrightarrow S_{neg}\text{-}Co$	(R1)	$K_1^e = 12.0$
$Co(II)EDTA(aq) + S_{pos} \leftrightarrow S_{pos}\text{-}Co(II)EDTA$	(R2)	$K_2^e = 25.0$
$Fe(III)EDTA(aq) + S_{pos} \leftrightarrow S_{pos}\text{-}Fe(III)EDTA$	(R3)	$K_3^e = 9.0$
$EDTA(aq) + S_{pos} \leftrightarrow S_{pos}\text{-}EDTA$	(R4)	$K_4^e = 25.0$
$Co(III)EDTA(aq) + S_{pos} \leftrightarrow S_{pos}\text{-}Co(III)EDTA$	(R5)	$K_5^e = 2.5$
$S_{pos}\text{-}Co(II)EDTA \leftrightarrow Co(II)(aq) + S_{pos}\text{-}EDTA$	(R6)	$k_6^f = 1.0\ h^{-1}$, $k_6^b = 1.0 \times 10^{-3}\ L\ mM^{-1}h^{-1}$
$S_{pos}\text{-}EDTA \leftrightarrow Fe(III)EDTA(aq) + S_{pos}$	(R7)	$k_7^f = 2.5\ h^{-1}$, $k_7^b = 0.0\ L\ mM^{-1}h^{-1}$
$Co(II)EDTA(aq) \leftrightarrow Co(III)EDTA(aq)$	(R8)	$k_8^f = 1.0 \times 10^{-3}\ h^{-1}$, $k_8^b = 0.0\ h^{-1}$
$Fe(III)EDTA(aq) + 6O_2 \rightarrow 3CO_2 + Biomass$	(R9)	user specified reaction rate as expressed by R_9 in Eq. (B1.9) in Table 2.B.1, where $\mu_1 = 2.5 \times 10^{-4}\ h^{-1}$, $k_{11} = 1.0 \times 10^{-5}\ mM\ L^{-1}$, and $k_{21} = 1.0 \times 10^{-5}\ mM\ L^{-1}$
$EDTA(aq) + 6O_2 \rightarrow 3CO_2 + Biomass$	(R10)	user specified reaction rate as expressed by R_{10} in Eq. (B1.9) in Table 2.B.1, where $\mu_2 = 0.025\ h^{-1}$, $k_{12} = 1.0 \times 10^{-5}\ mM\ L^{-1}$, and $k_{22} = 1.0 \times 10^{-5}\ mM\ L^{-1}$

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], species initially present in the system are Co(II)EDTA (0.032 mM), dissolved O_2 (0.256 mM), microorganisms (0.02 mM), the charged surface site S_{pos} (0.016 mM) and S_{neg} (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 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 Table 2.B.1 in Appendix 2.B.

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 2.1. 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 BIOGEOCHEMl. For description and discussion of the experimental data see Chilakapati et al. [1998].

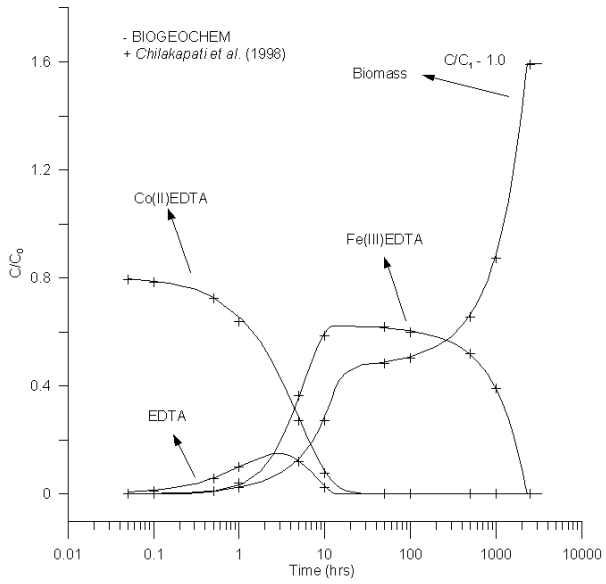


Figure 2.1. Solution of the reaction network in table 2.1: Comparison of this study with results of *Chilakapati et al.*[1998]. (C_0 and C_1 are the initial concentrations of Co(II)EDTA(aq) and microorganisms, respectively)

2.6.2 Complexation, Adsorption, Ion-exchange, and Dissolution in a System of Mixed Equilibrium and Kinetic Reactions

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 2.2 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 2.2); 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 2.2). 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).

Table 2.2 Reaction network for the example problem in Section 2.6.2

Reaction	#	Reaction constants
Mineral Dissolution and Surface Site Formation Reactions		
$M \leftrightarrow C_1 - 3C_2$	(R1)	$k_1^f = 0.05$
$M \leftrightarrow S_1$	(R2)	User specified partition between bulk and surface metal ions as expressed by equation (B2.1) in the Appendix 2.B.
Aqueous Complexation Reactions		
$C_3 \leftrightarrow C_4 + C_5$	(R3)	$\text{Log } k_3^f = 2.03$, $\text{Log } k_3^b = 20.00$
$C_6 + C_5 \leftrightarrow C_7$	(R4)	$\text{Log } K_4^e = 12.32$
$C_2 + C_5 + C_6 \leftrightarrow C_8$	(R5)	$\text{Log } K_5^e = 15.93$
$C_6 \leftrightarrow C_2 + C_9$	(R6)	$\text{Log } K_6^e = -12.60$
$C_1 + C_5 \leftrightarrow C_{10}$	(R7)	$\text{Log } k_7^f = 25.00$, $\text{Log } k_7^b = -2.57$
$C_1 + C_2 + C_5 \leftrightarrow C_{11}$	(R8)	$\text{Log } K_8^e = 29.08$
$C_1 + C_5 \leftrightarrow C_2 + C_{12}$	(R9)	$\text{Log } K_9^e = 19.65$
$C_1 + C_5 \leftrightarrow 2C_2 + C_{13}$	(R10)	$\text{Log } K_{10}^e = -36.30$
$C_1 \leftrightarrow C_2 + C_{14}$	(R11)	$\text{Log } K_{11}^e = -2.19$
$C_1 \leftrightarrow 2C_2 + C_{15}$	(R12)	$\text{Log } K_{12}^e = -5.67$

$C_1 \leftrightarrow 3C_2 + C_{16}$	(R13)	Log K_{13}^e = -13.60
$C_1 \leftrightarrow 4C_2 + C_{17}$	(R14)	Log K_{14}^e = -21.60
$2C_1 \leftrightarrow 2C_2 + C_{18}$	(R15)	Log K_{15}^e = -2.95
$C_2 + C_4 + C_5 \leftrightarrow C_{19}$	(R16)	Log K_{16}^e = 21.40
$C_4 \leftrightarrow C_2 + C_{20}$	(R17)	Log K_{17}^e = -9.67
$C_4 \leftrightarrow 2C_2 + C_{21}$	(R18)	Log K_{18}^e = -18.76
$C_4 \leftrightarrow 3C_2 + C_{22}$	(R19)	Log K_{19}^e = -32.23
$C_2 + C_5 \leftrightarrow C_{23}$	(R20)	Log K_{20}^e = 11.03
$2C_2 + C_5 \leftrightarrow C_{24}$	(R21)	Log K_{21}^e = 17.78
$3C_2 + C_5 \leftrightarrow C_{25}$	(R22)	Log K_{22}^e = 20.89
$4C_2 + C_5 \leftrightarrow C_{26}$	(R23)	Log K_{23}^e = 23.10
$C_2 + C_{27} \leftrightarrow C_{28}$	(R24)	Log K_{24}^e = 14.00
Adsorption-Desorption Reactions		
$S_1 \leftrightarrow S_2 + C_2$	(R25)	Log K_{25}^e = -11.60
$S_1 + C_2 \leftrightarrow S_3$	(R26)	Log K_{26}^e = 5.60
$S_1 + 3C_2 + C_5 \leftrightarrow S_4$	(R27)	Log K_{27}^e = 30.48
$S_1 + C_1 + C_2 + C_5 \leftrightarrow S_5$	(R28)	Log k_{28}^f = 40.00, Log k_{28}^b = 2.37
$S_1 + C_2 + C_4 + C_5 \leftrightarrow S_6$	(R29)	Log k_{29}^f = 30.00, Log k_{29}^b = 1.51
$S_1 - C_2 + C_4 \leftrightarrow S_7$	(R30)	Log k_{30}^f = -0.99, Log k_{30}^b = 1.70
$S_1 + C_2 + C_5 + C_6 \leftrightarrow S_8$	(R31)	Log k_{31}^b = 1.19, Log k_{31}^f = 25.0
Ion-Exchange Reactions		
C_{29} + 2 site-$C_{30} \leftrightarrow$ site-$C_{29} + 2C_{30}$	(R32)	Log k_{32}^b = -0.5, Log k_{32}^f = -0.75
C_6 + 2 site-$C_{30} \leftrightarrow$ site-$C_6 + 2C_{30}$	(R33)	Log k_{33}^e = 0.6

The slow reactions were assumed to be the dissolution of the mineral, and the adsorption of C_5-complexes and C_4 (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.

From the above conceptualization of the reaction network, this simulation includes a total of 41 species (M = 41), excluding the species C_{28} because its activity is assumed 1.0, 25 equilibrium reactions ($N_e = 25$) and 8 kinetic reactions ($N_k = 8$). Initial concentrations of C_6, M, C_3, and C_{30} are 2×10^{-3} M, 2.36×10^{-5} M, 8.51×10^{-6} M, and 0.1552 M, respectively; and initial concentrations of site-C_{30}, site-C_{29}, and site-C_6 are 0.0651 M, 0.1463 M, and 0.1562 M, respectively. All other species are initially 0.0 M except for C_2, whose concentration is fixed at 3.16×10^{-5} M. The BIOGEOCHEM preprocessor indicated that all reactions in this system are linearly independent, i.e., $N_I = 33$, $N_E = N_e = 25$, $N_{KI} = N_K = N_k = 8$. 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 users-specified algebraic equation (25 total), 8 kinetic variable equations, and 8 mass conservation equations ($N_C = M - N_I = 41 - 33$). An example of one decomposition of the reaction matrix presented in Table 2.2 is shown in Table 2.B.2 in Appendix 2.B.

The 24 mass-action equilibrium equations are decoupled from the other 17 equations (8 mass conservation, 8 kinetic-variable, and one users-specified algebraic equation). In other words, 24 equations are substituted into 17 equations to eliminate 24 secondary variables (C_7, C_8, C_9, C_{11} through C_{27}, S_2, S_3, S_4, and site-C_{30}). The resulting 17 equations are then solved simultaneously for 17 master variables (C_1 through C_6, C_{10}, C_{28}, C_{29}, C_{30}, S_1, S_5 through S_8, M, and site-C_{29}).

The time-variant dissolution of the mineral is described by the concentration variations of six species (Figs. 2.2 and 2.3). C_3, C_{16} and M in Fig. 2.2 represent the species that are directly involved in the dissolution according to the reaction network. The formation of species C_{10} dominates the aqueous speciation in the system and is

ultimately the driving force for mineral dissolution. The last two species in Fig. 2.3 represent the two main adsorbed species in the system. The simulation of the interaction of these six species demonstrates the ability of BIOGEOCHEM to model a complex (multiple reaction types) mixed (equilibrium and kinetic) reaction system.

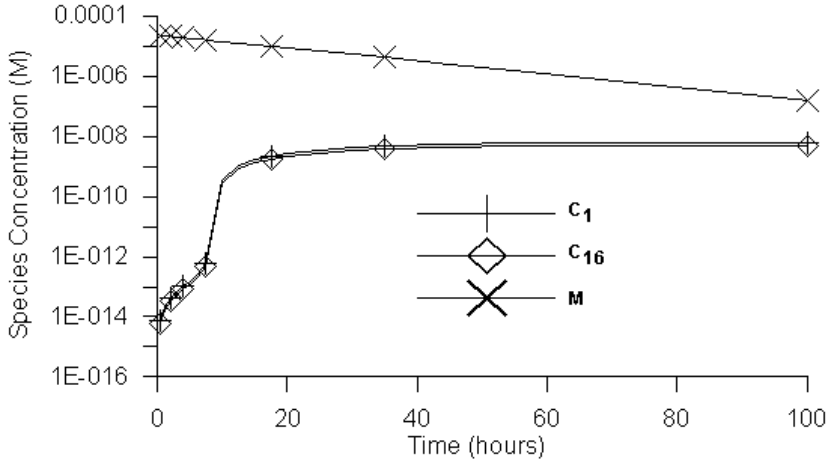


Figure 2.2. Concentration curves for major species directly involving in the mineral dissolution.

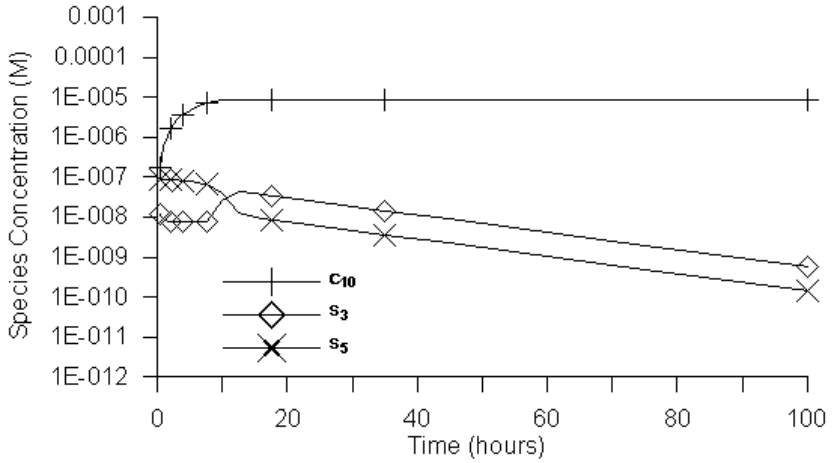


Figure 2.3. Concentration curves for main complexed and adsorbed species

For the first 10 hours, there is a decline in the mineral concentration (Fig. 2.2) which corresponds to an equivalent increase in the concentration of C_{10} (Fig. 2.2). Dissolved species C_1 quickly combines with C_5 to form the complexed species. This

continues until C_{10} begins to reach its equilibrium value (at around 15 h) due to the limiting of total concentration of C_5 (Fig. 2.2). At this time, the aqueous species C_1 and C_{16} that are formed no longer combine directly into the complexed species but remain free and increase to their approximate equilibrium values (Fig. 2.2).

The two main sorbed species, S_3 and S_5, also undergo considerable changes during the first 20 h. Initially S_5 is the main adsorbed species (Fig. 2.3). But when the production of C_{10} begins to be limited by the available free C_5, its high production rate drives its continued formation by scavenging other sources of C_5 species. As a result, the adsorbed species (S_5) begins to decrease to form the aqueous complexed species (C_{10}) after ca. 12 h. S_3 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 2.2 because of the assumption that it is an irreversible reaction.

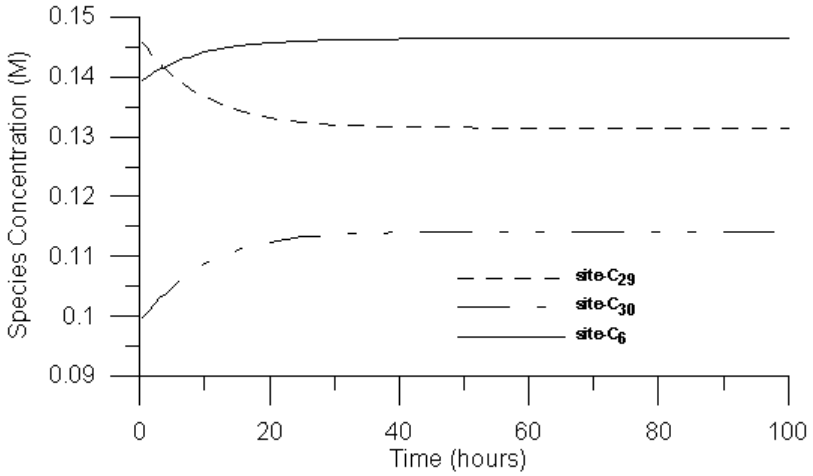


Figure 2.4. Concentration curves for ion-exchanged species

Reactions R32 and R33 are weakly coupled to the rest of the system because the amounts of C_6 and its complexed species are insignificant. About 10% of site-C_6 was exchanged immediately into the solution because of the equilibrium reaction R33. There

are significant amounts of site-C_{29} and C_{30} in the system, R32 proceeds from right to left till it reaches equilibrium as shown in Fig. 2.4.

2.6.3 Basis switching

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 2.3).

Based on this reaction network there are a total of 31 species (M = 31), excluding the species $H_2O(l)$ because its activity is assumed 1.0, and 24 equilibrium reactions (N_e = 24). The BIOGEOCHEM preprocessor indicated that all 24 equilibrium reactions in this system are linearly independent, i.e., $N_E = N_e = 24$. Hence, a formal matrix decomposition of the reaction network should yield 24 mass-action equilibrium equations and 7 mass conservation equations ($N_C = M - N_I = 31 - 24$).

Without basis switching, the simulation stopped while checking mass balances. BIOGEOCHEM displayed information in the output file that mass balance 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 $Al(OH)_4^{-1}$ 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 4.263×10^{-5} M while the calculated TOT Al^{3+} is 3.684×10^{-4} M; thus it is seen that there is a large error in mass balance. The simulated concentration of species Al^{3+} is 3.015×10^{-22} M, which is close to 0, while the calculated

concentration of $Al(OH)_4^-$ is 3.680×10^{-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+}.

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 2.4, with 3 of the 4 possible precipitated species actually formed within the system. The simulated concentration of $Al(OH)_4^-$ is 5.64×10^{-8} M, while the calculated concentration of Al^{3+} is 2.80×10^{-10} M. It is seen, from Table 2.4, 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 $Al(OH)_4^-$.

Table 2.3. Reaction network for the example problem in Section 2.6.3.

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

Table 2.4. Initial and simulated species concentration

Species	C_0 (M)	Log(C_{eq})	Species	C_0 (M)	Log(C_{eq})
Ca^{2+}	6.335×10^{-1}	-2.160	HCO_3^-	0.000	-2.000
$CO3^{2-}$	6.365×10^{-1}	-5.478	H_2CO_3	0.000	-2.292
Al^{3+}	4.263×10^{-5}	-9.553	$CaCO_3$	0.000	-5.260
$SO4^{2-}$	3.177×10^{-2}	-1.582	$CaHCO_3^+$	0.000	-3.471
H^+	2.056×10^{-2}	-6.536	HSO_4^-	0.000	-6.443
Fe^{3+}	1.234×10^{-5}	-13.77	$CaSO_4$	0.000	-2.283
Na^+	3.043×10^{-5}	-1.522	$AlSO_4^+$	0.000	-8.896
OH^-	0.000	-7.359	$Al(SO4)_2^-$	0.000	-9.398
$FeOH^{2+}$	0.000	-9.952	$FeSO_4^+$	0.000	-12.576
$Fe(OH)_2^+$	0.000	-7.211	$Fe(SO4)_2^-$	0.000	-13.198
$Fe(OH)_3$	0.000	-7.670	$NaSO_4^-$	0.000	-3.454
$Fe(OH)_4^-$	0.000	-10.069	$CaCO_3$ (s)	0.000	-0.207
$Al(OH)^{2+}$	0.000	-8.542	$Al(OH)_3$ (s)	0.000	-4.371
$Al(OH)_2^+$	0.000	-7.521	$Fe(OH)_3$ (s)	0.000	-4.912
$Al(OH)_4^-$	0.000	-7.249	$CaSO_4$ (s)	0.000	$-\infty$
$Al(OH)_3$	0.000	-8.090			

2.7 CONCLUSION AND DISCUSSION

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 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 users-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 users-specified algebraic equation. Any slow reaction can be modeled with microbial-mediated enzymatic kinetics, empirical n-th order rates, an elementary rate, or a users-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., 2002a, 2002b).

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.

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 el al., 2001a). The individual rate formulation/parameterization is more descriptive of geochemical and biochemical reactions (Burgos et al., 2002a, 2002b).

To make the model numerically robust, a dynamic basis-species switching strategy due to the non-uniqueness 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.

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 2.A

2.A.1 Decomposition I

Let's start with the following matrix, which is a repeat of Eq. (5)

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

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.

In the first column of the reaction matrix in Eq. (5), there are three non-zero 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 HNTA) is chosen as a pivot element. Pivoting on Row 3 (HNTA) gives

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{Bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix} \tag{A1}$$

In the second column of the Column 1-reduced reaction matrix in Eq. (A1), there

are three non-zero 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, chose the fifth row as the pivot element. Pivoting on Row 5 (CoNTA) gives

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

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 users-selected component list and treated it as a product species.

There are only two no-zero rows (Rows 3 and 5) in the third column in the Column 2-reduced reaction in Eq. (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.

In the fourth column of the reduced reaction matrix in Eq. (A2) there are four

non-zero 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 chose Row 6 as the pivot element. Pivoting on Row 6 (B) gives

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

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

Expand the matrix in Eq. (A3) and the following equations are obtained:

$$\frac{d[H]}{dt} + \frac{d[HNTA]}{dt} + 2\frac{d[B]}{dt} = 0 \tag{A4.1}$$

$$\frac{d[NTA]}{dt} + \frac{d[HNTA]}{dt} + \frac{d[CoNTA]}{dt} + \frac{d[B]}{dt} = 0 \tag{A4.2}$$

$$\frac{d[HNTA]}{dt} + \frac{d[B]}{dt} = R_1 + R_3 \tag{A4.3}$$

$$\frac{d[Co]}{dt} + \frac{d[CoNTA]}{dt} = 0 \tag{A4.4}$$

$$\frac{d[CoNTA]}{dt} = -R_2 - R_3 \tag{A4.5}$$

$$\frac{d[B]}{dt} = R_4 \tag{A4.6}$$

Equations (A4.1), (A4.2) and (A4.4) 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 (A4.3) and (A4.5) each contains one equilibrium reaction. They are replaced by the mass-action equilibrium equations as shown in the last equation of Eqs. (6.1) and (6.2) in section 2.3.

Eq. (A3) is re-arranged so that it can be written in the form of Eq. (8) in section 2.3.2.3 as

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

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

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

we can easily see that

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}; \ \frac{d\mathbf{C}}{dt} = \begin{Bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{Bmatrix}; \ \textit{and} \ \ \mathbf{R} = \begin{Bmatrix} R_1 \\ R_2 \\ R_4 \\ R_3 \end{Bmatrix} \tag{A5.1}$$

and

$$\mathbf{D}=\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{bmatrix};\ \mathbf{K}=\begin{bmatrix}1\\ -1\\ 0\end{bmatrix};\ \mathbf{0}_1=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix};\ \textit{and}\ \ \mathbf{0}_2=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix} \tag{A5.2}$$

2.A.2 Decomposition II

An alternative diagonalization is to choose HNTA, CoNTA, and B as components. Again repeat Eq. (5)

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

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

$$\begin{bmatrix}1 & 0 & 1 & 0 & 0 & 0\\ -1 & 1 & 0 & 0 & 0 & 0\\ 1 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}\begin{Bmatrix}\mathrm{d[H]/dt}\\ \mathrm{d[NTA]/dt}\\ \mathrm{d[HNTA]/dt}\\ \mathrm{d[Co]/dt}\\ \mathrm{d[CoNTA]/dt}\\ \mathrm{d[B]/dt}\end{Bmatrix}=\begin{bmatrix}-1 & 0 & -1 & -1\\ 0 & 1 & 1 & 1\\ 0 & 0 & 0 & -2\\ 0 & 1 & 1 & 0\\ 0 & -1 & -1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{Bmatrix}R_1\\ R_2\\ R_3\\ R_4\end{Bmatrix} \tag{A6}$$

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

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{Bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix} \tag{A7}$$

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

$$\begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & -2 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{Bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \end{Bmatrix} = \begin{bmatrix} -1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{Bmatrix} \tag{A8}$$

Expand the matrix and the following equations can be obtained:

$$\frac{d[NTA]}{dt} - \frac{d[Co]}{dt} = -R_1 - R_3 \tag{A9.1}$$

$$\frac{d[Co]}{dt} = R_2 + R_3 \tag{A9.2}$$

$$-\frac{d[H]}{dt} + 2\frac{d[NTA]}{dt} + \frac{d[HNTA]}{dt} - 2\frac{d[Co]}{dt} = 0 \tag{A9.3}$$

$$\frac{d[H]}{dt} - \frac{d[NTA]}{dt} + \frac{d[Co]}{dt} = -R_4 \tag{A9.4}$$

$$\frac{d[Co]}{dt} + \frac{d[CoNTA]}{dt} = 0 \tag{A9.5}$$

$$\frac{d[H]}{dt} - \frac{d[NTA]}{dt} + \frac{d[Co]}{dt} + \frac{d[B]}{dt} = 0 \tag{A9.6}$$

Equations (A9.3), (A9.5) and (A9.6) define Total HNTA, Total CoNTA, and Total B, respectively. Equations (A9.1) and (A9.2) contain equilibrium reactions, they are replaced by the mass-action equilibrium equations as shown in the last equation of Eqs. (7.1) and (7.2) in section 2.3.

2.A.3 Elimination of Secondary Species

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

(RA1): $A + C \leftrightarrow 3D$, with equilibrium constant K_1^e

(RA2): $A + 2B \leftrightarrow 2C$, with equilibrium K_2^e

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 of these two reactions is first generated as

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

In the first row in Eq. (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 non-zero rows, it is chosen as the pivot element for the first row reduction. Pivoting on species D gives

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

It is noted that Eq. (A11) is identical to Eq. (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.

In the second row in Eq. (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

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

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 Eq. A(12)]. After the normalization, we have

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

Expanding Eq. (A13), we have

$$(D) = (K_1^e)^{\frac{1}{3}} (K_2^e)^{\frac{1}{6}} (A)^{\frac{1}{2}} (B)^{\frac{1}{3}} \tag{A14}$$

and

$$(C) = (K_2^e)^{\frac{1}{2}} (A)^{\frac{1}{2}} (B) \tag{A15}$$

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

APPENDIX 2.B

The decomposed set of equations for Examples 1 and 2, respectively, are given in Table 2.B.1 and 2.B.2, respectively.

Table 2.B.1. Matrix decomposition for the reaction network of the transformation of EDTA

Mass-action Equilibrium Equations	
$(\mathrm{Co(II)(aq)}) = \frac{1}{K_1^e}\ \frac{(S_{neg} - \mathrm{Co})}{(S_{neg})}$	(B1.1)
$(S_{pos} - \mathrm{Co(II)EDTA}) = K_2^e\ (\mathrm{Co(II)EDTA(aq)})\ (S_{pos})$	(B1.2)
$(S_{pos} - \mathrm{Fe(III)EDTA}) = K_3^e\ (S_{pos})(\mathrm{Fe(III)EDTA(aq)})$	(B1.3)
$(\mathrm{EDTA(aq)}) = \frac{1}{K_4^e}\frac{(S_{pos})}{(S_{pos} - \mathrm{EDTA})}$	(B1.4)
$(\mathrm{Co(III)EDTA(aq)} = \frac{1}{K_5^e}\frac{(S_{pos} - \mathrm{Co(III)EDTA})}{(S_{pos})}$	(B1.5)
Kinetic-Variable Equations	
$\frac{d([\mathrm{Co(II)(aq)}] + [S_{neg} - \mathrm{Co}])}{dt} = R_6;\ R_6 = k_6^f[S_{pos} - \mathrm{Co(II)EDTA}] - k_6^b[\mathrm{Co(II)(aq)}][S_{pos} - \mathrm{EDTA}]$	(B1.6)
$\frac{d([\mathrm{EDTA}] + [S_{pos} - \mathrm{EDTA}) - [\mathrm{Co(II)(aq)}] + [S_{neg} - \mathrm{Co}])}{dt} = -R_7 - R_{10}$ $R_7 = k_7^f[S_{pos} - \mathrm{EDTA}] - k_7^b[S_{pos}][\mathrm{Fe(III)} - \mathrm{EDTA(aq)}],\ R_{10} = \frac{\mu_2\ [\mathrm{EDTA(aq)}]\ [O_2]\ [\mathrm{Biomass}]}{(k_{12} + [\mathrm{EDTA(aq)}])\ (k_{22} + [O_2])}$	(B1.7)
$\frac{d([\mathrm{Co(III)EDTA(aq)}] + [S_{pos} - \mathrm{Co(III)EDTA}])}{dt} = R_8;\ R_8 = k_8^f[\mathrm{Co(II)EDTA(aq)}] - k_8^b[\mathrm{Co(III)EDTA(aq)}]$	(B1.8)
$\frac{d[CO_2]}{dt} = 3R_9 + 3R_{10}$ $R_9 = \frac{\mu_1\ [\mathrm{Fe(III)EDTA(aq)}]\ [O_2]\ [\mathrm{Biomass}]}{(k_{11} + [\mathrm{Fe(III)EDTA(aq)}])\ (k_{21} + [O_2])},\ R_{10} = \frac{\mu_2\ [\mathrm{EDTA(aq)}]\ [O_2]\ [\mathrm{Biomass}]}{(k_{12} + [\mathrm{EDTA(aq)}])\ ([k_{22} + [O_2])}$	(B1.9)
Mass Conservation Equations	
$\mathrm{TOT}_{S_{neg}} = [S_{neg}] + [S_{neg} - \mathrm{Co}] = \mathrm{Initial}\ \mathrm{TOT}_{S_{neg}}$	(B1.10)

$$TOT_{Co(II)EDTA} = [Co(II)(aq)] + [S_{neg}] - Co] + [Co(II)EDTA(aq)] + [S_{pos} - Co(II)EDTA] + [Co(III)EDTA(aq)] + [S_{pos} - Co(III)EDTA] = Initial\ TOT_{Co(II)EDTA} \quad (B1.11)$$

$$TOT_{S_{pos}} = [S_{pos}] + [S_{pos} - Co(II)EDTA] + [S_{pos} - Fe(III)EDTA] + [S_{pos} - EDTA] + [S_{pos} - Co(III)EDTA] = Initial\ TOT_{S_{pos}} \quad (B1.12)$$

$$TOT_{Fe(III)EDTA} = [-Co(II)(aq)] - [S_{neg} - Co] + [Fe(III)EDTA(aq)] + [S_{pos} - Fe(III)EDTA] + [EDTA(aq)] + [S_{pos} - EDTA] + 1/3[CO_2] = Initial\ TOT_{Fe(III)EDTA} \quad (B1.13)$$

$$TOT_{O_2} = [O_2] + 2[CO_2] = Initial\ TOT_{O_2} \quad (B1.14)$$

$$TOT_{Biomass} = [Biomass] - 1/3[CO_2] = Initial\ TOT_{Biomass} \quad (B1.15)$$

Note: 1. A species inside a parenthesis in Table 2.B.1 denotes the activity of the species.

2. A species inside a bracket in Table 2.B.1 denotes the concentration of the species.

3. Activity = activity coefficient × concentration

Table 2.B.2. Matrix decomposition for the complex (multiple reaction types) mixed (equilibrium and kinetic) reaction system

Users-Specified Algebraic Equations	
$\frac{S_A N_s}{N_A} M_{mineral}[M] = [S_1] + [S_2] + [S_3] + [S_4] + [S_5] + [S_6] + [S_7] + [S_8]$ (B2.1)	
This users' specified equation is modified from Stumm and Morgan, 1996, in which S_A is the unit surface area ($m^2\ g^{-1}$) of mineral, N_S is the surface site density (mol sites m^{-2}), N_A is Avogadro's number (mol sites mol^{-1}), $M_{mineral}$ is the molecular weight of mineral (g mol^{-1}).	
Mass-action Equilibrium Equations	
$(C_7) = K_4^e\ (C_5)\ (C_6)$ (B2.2)	$(C_8) = K_5^e\ (C_2)\ (C_5)\ (C_6)$ (B2.3)
$(C_9) = K_6^e \frac{(C_6)}{(C_2)}$ (B2.4)	$(C_{11}) = K_8^e (C_1)(C_2)(C_5)$ (B2.5)
$(C_{12}) = K_9^e \frac{(C_1)(C_5)}{(C_2)}$ (B2.6)	$(C_{13}) = K_{10}^e \frac{(C_1)(C_5)}{(C_2)^2}$ (B2.7)
$(C_{14}) = K_{11}^e \frac{(C_1)}{(C_2)}$ (B2.8)	$(C_{15}) = K_{12}^e \frac{(C_1)}{(C_2)^2}$ (B2.9)
$(C_{16}) = K_{13}^e \frac{(C_1)}{(C_2)^3}$ (B2.10)	$(C_{17}) = K_{14}^e \frac{(C_1)}{(C_2)^4}$ (B2.11)
$(C_{18}) = K_{15}^e \frac{(C_1)^2}{(C_2)^2}$ (B2.12)	$(C_{19}) = K_{16}^e\ (C_2)\ (C_4)\ (C_5)$ (B2.13)
$(C_{20}) = K_{17}^e \frac{(C_4)}{(C_2)}$ (B2.14)	$(C_{21}) = K_{18}^e \frac{(C_4)}{(C_2)^2}$ (B2.15)
$(C_{22}) = K_{19}^e \frac{(C_4)}{(C_2)^3}$ (B2.16)	$(C_{23}) = K_{20}^e\ (C_2)\ (C_5)$ (B2.17)
$(C_{24}) = K_{21}^e\ (C_2)^2\ (C_5)$ (B2.18)	$(C_{25}) = K_{22}^e\ (C_2)^3\ (C_5)$ (B2.19)
$(C_{26}) = K_{23}^e\ (C_2)^4\ (C_5)$ (B2.20)	$(C_{27}) = \frac{1}{K_{24}^e} \frac{1}{(C_2)}$ (B2.21)
$(S_2) = K_{25}^e \frac{(S_1)}{(C_2)}$ (B2.22)	$(S_3) = K_{26}^e\ (C_2)\ (S_1)$ (B2 23)
$(S_4) = K_{27}^e\ (C_2)^3\ (C_5)\ (S_1)$ (B2.24)	$(\text{site-}C_{30}) = \frac{1}{K_{33}^{e\ 0.5}} \frac{(C_{30})(\text{site-}C_6)^{0.5}}{(C_6)^{0.5}}$ (B2.25)
Kinetic-Variable Equations	

$$\frac{d\,[C_{10}]}{dt} = R_7;\ R_7 = k_7^f[C_1][C_5] - k_7^b[C_{10}] \tag{B2.26}$$

$$\frac{d}{dt}([S_1] + [M] + [S_2] + [S_3] + [S_5] + [S_6] + [S_7] + [S_4] + [S_8] = -R_1,\ R_1 = k_1^f[M] \tag{B2.27}$$

$$\frac{d\,[C_3]}{dt} = -R_3,\ R_3 = k_3^f[C_3] - k_3^b[C_4][C_5] \tag{B2.28}$$

$$\frac{d\,[S_5]}{dt} = R_{28},\ R_{28} = k_{28}^f[S_1][C_1][C_2][C_5] - k_{28}^b[S_5] \tag{B2.29}$$

$$\frac{d\,[S_6]}{dt} = R_{29},\ R_{29} = k_{29}^f[S_1][C_2][C_4][C_5] - k_{29}^b[S_6] \tag{B2.30}$$

$$\frac{d\,[S_7]}{dt} = R_{30},\ R_{30} = k_{30}^f\frac{[S_1][C_4]}{[C_2]} - k_{30}^b[S_7] \tag{B2.31}$$

$$\frac{d\,[S_8]}{dt} = R_{31},\ R_{31} = k_{31}^f[S_1][C_2][C_5][C_6] - k_{31}^b[S_8] \tag{B2.32}$$

$$\frac{d\,[C_{29}]}{dt} = -R_{32},\ R_{32} = k_{32}^f[C_{29}][\text{site-}C_{30}]^2 - k_{32}^b[\text{site-}C_{29}][C_{30}]^2 \tag{B2.33}$$

Mass Conservation Equations

$$TOT_{C_6} = [C_6] + [C_7] + [C_8] + [C_9] + [S_8] + [C_{29}] - 0.5\,[\text{site-}C_{30}] = \text{Initial } TOT_{C_6} \tag{B2.34}$$

$$TOT_{C_4} = [C_4] + [C_3] + [C_{19}] + [C_{20}] + [C_{21}] + [C_{22}] + [S_6] + [S_7] = \text{Initial } TOT_{C_4} \tag{B2.35}$$

$$\begin{aligned} TOT_{C_2} = {} & [C_2] - 3[S_1] + [C_8] - [C_9] + [C_{11}] - [C_{12}] - 2[C_{13}] - [C_{14}] - 2[C_{15}] - 3[C_{16}] - 3[M] \\ & - 4[C_{17}] - 2[C_{18}] + [C_{19}] - [C_{20}] - 2[C_{21}] - 3[C_{22}] + [C_{23}] + 2[C_{24}] + 3[C_{25}] + 4[C_{26}] \\ & - [C_{27}] - 4[S_2] - 2[S_3] - 2[S_5] - 2[S_6] - 4[S_7] - 2[S_8] = \text{Initial } TOT_{C_2} \end{aligned} \tag{B2.36}$$

$$\begin{aligned} TOT_{C_5} = {} & [C_5] + [C_7] + [C_8] + [C_{10}] + [C_{11}] + [C_{12}] + [C_{13}] + [C_3] + [C_{19}] + [C_{23}] + [C_{24}] + [C_{25}] + [C_{26}] \\ & + [S_5] + [S_6] + [S_4] + [S_8] = \text{Initial } TOT_{C_5} \end{aligned} \tag{B2.37}$$

$$\begin{aligned} TOT_{C_1} = {} & [C_1] + [S_1] + [C_{10}] + [C_{11}] + [C_{12}] + [C_{13}] + [C_{14}] + [C_{15}] + [C_{16}] + [M] + [C_{17}] + 2[C_{18}] \\ & + [S_2] + [S_3] + 2[S_5] + [S_6] + [S_7] + [S_4] + [S_8] = \text{Initial } TOT_{C_1} \end{aligned} \tag{B2.38}$$

$$TOT_{C_{30}} = [C_{30}] + [\text{site-}C_{30}] = \text{Initial } TOT_{C_{30}} \tag{B2.39}$$

$$TOT_{\text{site-}C_{29}} = [C_{29}] + [\text{site-}C_{29}] = \text{Initial } TOT_{C_{29}} \tag{B2.40}$$

$$TOT_{\text{site-}C_6} = -[C_{29}] + 0.5[\text{site-}C_{30}] + [\text{site-}C_6] = \text{Initial } TOT_{\text{site-}C_6} \tag{B2.41}$$

Note: 1. A species inside a parenthesis in Table 2.B.2 denotes the activity of the

species.

2. A species inside a bracket in Table 2.B.2 denotes the concentration of the species.

3. Activity = activity coefficient × concentration

ACKNOWLEDGEMENTS. This research was supported, in part, by National Science Foundation (NSF) Grant No.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 No.DE-FG02-98ER62691 with the Pennsylvania State University.

REFERENCES

Atkins, P. W., *Physical Chemistry*. Oxford University Press, 1986.

Abrams, R. H. and Loague K., 1998. Development and testing of a compartmentalized reaction network model fro redox oznes in contaminated aquifers.

Ambrose, R.B. et al. 1988. WASP4, A Hydrodynamic and WaterQuality Model--Model theory,User's Manual, and Programmer's Guide. U.S. Environmental Protection Agency, Athens, A. EPA/600/3-87-039.

Barnwell, T. O. and Brown, L. C., 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual, EPA/600/3-87/007.

Bosma, W. and van der Zee, 1993.Transport of reacting solute in a one-dimensional, chemically heterogeneous porous medium. Water Resources Research, 29(1): 117-131.

Brun, A. and Engesgaard, P., 2002. Modelling of transport and biogeochemical processes in pollution plumes: literature review and model development. Journal of Hydrology 256:211-227

Brusseau, M., 1994. Transport of reactive contaminants in heterogeneous porous media. Reviews of Geophysics, 32(3): 285-313.

Burgos W.D., Royer R.A., Fang Y.L., Yeh G.T., Fisher A.S., Jeon B.H. and Dempsey B.A. (2002a) Theoretical and experimental considerations related to reaction-based modeling: A case study using iron(III) oxide bioreduction. *Geomicrobiology J.* **19**(2), 1-35

Burgos W. D., Fang Y. L., Royer R. A., Yeh G. T, Stone J. J., Jeon B. H. and Dempsey B. A. (2002b) Reaction-based modeling of quinone-mediated bacterial iron(III)

reduction. Submitted.

Cederberg, G. A., Street, R. L., and Leckie, J. O., 1985. A Groundwater Mass Transport and Equilibrium Chemistry Model for Multicomponent Systems. Water Resources Research, 21: 1095-1104.

Chang, J. R., G. T. Yeh, and T. E. Short. 1993. Modeling two-dimensional subsurface flow, fate and transport of microbes and chemicals. Proc. Int. Symp. on Eng. Hydrology, San Francisco, July 25-30, 1993, p.611-616.

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 Resources Researc*h, 28(7): 1833-1847, July 1992.

Cheng, J. R. and Yeh, G. T. , 1994. Modeling three-dimensional subsurface flow, fate and transport of microbes and chemicals (3DFATMIC). Proc. X-th Int. Conf. on Numerical Methods in Water Resources, Heidelberg, Germany, July 19-22, 1994, 217-224.

Cheng, H.-P., 1995. Development and Application of a Three-Dimensional Finite Element Model of Subsurface Flow, Heat Transfer, and Reactive Chemical Transport. Ph.D. Thesis, Department of Civil and Environmental Engineering, The Pennsylvania State University. University Park, PA.

Chilakapati, A., 1995. RAFT: A Simulator for ReActive Flow and Transport of Groundwater Contaminants. PNL Report 10636.

Chilakapati, A., Ginn, T. and Szecsody, J., 1998. An Analysis of Complex Reaction Networks in Groundwater Modeling. Water Resources Research, 34(7): 1767-1780.

Chilakapati, A., Yabusaki, S., Szecsody, J. and MacEvoy, W., 2000. Groundwater flow, multicomponent transport and biogeochemistry: development and application of a coupled process model. J. Contami. Hydrol. 43: 303-325.

Davis, J., Kent D., Coston, J., Hess, K. and Joye, J., 2000. Multispecies reactive tracer test in an aquifer with spatially variable chemical conditions. Water Resources Research, 36(1): 119-134.

Fang, Y. and G. T. Yeh, 2002. Numerical modeling of reactive chemical transport under multiphase systems. Proc. XIV International Conference on Computational Methods in Water Resources. Delft, The Netherlands, June 23-28, 2002.

Friedly, J. and Rubin, J., 1992. Solute Transport with multiple equilibrium-controlled or kinetically controlled chemical reactions. Water Resources Research, 28(6):1935-1953.

Goldberg, D. E., 1991. Chemistry Foundations. Schuam=s Outline of Theory and Problems. Schaum's Outline Series, McGraw-Hill, Inc., New York 372 pp

Griffioen, J., 1993. Multicomponent Cation Exchange Including Alkalinization/Acidification Following Flow Through a Sandy Sediment. Water Resources Research, 29(9):3005-3019.

Hostetler, J. C. and Erickson, R. L. ,1989. FASTCHEM Package 5. Report EA-5870-CCM, Electric Power Research Institute.

Lensing, H. J., Voyt, M., and Herrling, B., 1994. Modeling of Biologically Mediated Redox Processes in the Subsurface. Journal of Hydrology, 159: 125-143.

Lichtner, P. C., 1996. Continuum Formulation of Multicomponent-Multiphase Reactive Transport. In: P. C. Lichtner, C. I. Steefel, and E. H. Oelkers (Editors), Reactive Transport in Porous Media, Reviews in Mineralogy, Volume 34. Mineralogical

Society of America, Washington, D.C., pp. 1-79.

Lin, C.F. and Benjamin, M.M.,1990. Dissolution Kinetics of Minerals in the Presence of Sorbing and Complexing Ligands. Environmental Science Technology, Vol. 24, 126-134.

Liu, C. and Narasimhan, T., 1989a. Redox-controlled multiple-species reactive chemical transport 1. model development. Water Resources Research, 25(5): 869-882.

Liu, C. and Narasimhan, T., 1989b. Redox-controlled multiple-species reactive chemical transport 2. Verification and application. Water Resources Research, 25(5): 869-882.

Liu C., Kota S., Zachara J. M., Fredrickson J. K., and Brinkman C. K. (2001) Kinetic analysis of the bacterial reduction of geothite. Environmental Science & Technology 35(12), 2482-2490.

MacQuarrie, K.T.B., Sudicky, E.A., Frind, E.O., 1990. Simulation of biodegradable organic contaminants in groundwater: 1. Numerical formulation in principal directions. Water Resour. Res. 26, 207-222.

McNab, J., W. and Narasimhan, T., 1993. A multiple species transport model with sequential decay chain interactions in heterogeneous subsurface environments. Water Resources Research, 29(8): 2737-2746.

McNab, J., W.W. and Narasimhan, T.N., 1994. Modeling reactibe transport of organic compounds in groundwater using a partial redox disequilibrium approach. Water Resources Research, 30(9): 2619-2635.

McNab, J., W. W. and Narasimhan, T.N., 1995. Reactive transport of petroleum hydrocarbon constituents in a shallow aquifer: modeling geochemical interactions between organic and inorganic species. Water Resources Research,

31(8): 2027-2033.

Miller, C.W.,and Benson, L.V., 1983. Simulation of solute transport in a chemical reactive heterogeneous system: Model development and application, Water Resources Research, 19(2):381-391.

Narasimhan, T. N., White, A.F., and Tokunaga, T., 1986. Groundwater contamination from an inactive uranium mill tailings pile: 2. Application of a dynamic mixing model. Water Resources Research, 22(13):1820-1834.

Parkhurst, D.L., Thorstenson, D.C., and Plummer, L.N., 1980, PHREEQE--A computer program for geochemical calculations: U.S. Geological Survey Water-Resources Investigations Report 80-96, p. 195. (Revised and reprinted August, 1990.)

Parkhurst, D. L., 1995. User's Guide to PHREEQC, A Computer Model for Speciation, Reaction-Path, Advective-Dispersive Transport and Inverse Geochemical Calculations. U.S. Geological Survey, Water Resources Investigations Report 95-4227.

Parkhurst, D.L. and Appelo, C.A.J., 1999. User's guide to PHREEQC (version 2)--A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations: U.S.Geological Survey Water-Resources Investigations Report 99-4259,312 p.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T., 1992. Numerical Recipes in FORTRAN. Cambridge University Press, Cambridge.

Rubin, J., 1983. Transport of reacting solutes in porous media: relation between mathematical nature of problem formulation and chemical nature of reactions, Water Resours. Res., 19: 493-502.

Saiers, J.E., Guha, H., Jardine, P., and Brooks, S., 2000. Development and evaluation of a

mathematical model for the transport and oxidation-reduction of CoEDTA. Water Resources Research, 31(11):3151-3165.

Salvage, K.M., Yeh, G.T., Cheng, H.P. and Cheng, J.R., 1996. Development of a model of subsurface hydrologic transport and biogeochemical reactions (HYDROBIOGEOCHEM), Computational Methods in Water Resources XI. Computation Methods in Surface Flow and Transport Problesm, pp. 517-524.

Salvage, K., Yeh G.T., 1998. Development and Application of a Numerical Model of Kinetic and Equilibrium Microbiological and Geochemical Reactions (BIOKEMOD). J. Hydrol., 209, 27-52.

Smith, J. M., Chemical Engineering Kinetics. R. R. Donnelley & Sons Company, 1981.

Steefel, C.I., Lasaga, A.C., 1994. 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.

Steefel, C.I., MacQuarrie, K.T., 1996. Approaches to modeling reactive transport in porous media. In: Lichtner, P.C., Steefel, C.T., Oelkers, E.H. (Eds.), Reactive Transport in Porous Media, Reviews in Mineralology 34. Mineral. Soc. Am. 83,83-129

Steefel, C.I., Yabusaki, S.B., 1996. OS3D/GIMRT, Software for Modeling Multi-Component-Multidimensional Reactive Transport, User's Manual and Programmer's Guide. PNL-11166, Pacific Northwest Laboratory, Richland, WA.

Steefel, C.I., van Cappellen, P., 1998. Preface: reactive transport modeling of natural systems. J. Hydrol. 209, 1-7.

Stumm, W. and Morgan, J.J., Aquatic Chemistry, Third Edition, John Wiley & Sons, Inc., New York, 533-534, 1996.

Suarez, D. and Šimunek, J., 1996. Solute Transport Modeling Under Variably Saturated Water Flow Conditions. In: P. C. Lichtner, C. I. Steefel, and E. H. Oelkers (Editors), Reactive Transport in Porous Media, Reviews in Mineralogy, Volume 34. Mineralogical Society of America, Washington, D.C., pp. 229-268.

Szecsody, J. E., Zachara, J.M., and Bruckhart, P., 1994. Adsorption-dissolution reactions affecting the distribution and stability of Co(II)-EDTA in Fe-oxide coated sand. Environmental Science and Technology, 28:1706-1716.

Szecsody, J.E., Zachara, J.M., Chilakapati, A., Jardine, P.M. and Ferrency, A.S., 1998. Importance of flow and particle-scale heterogeneity on Co(II/III)EDTA reactive transport. Journal of Hydrology, 209(1): 112-136.

Tebes-Stevens, C., Valocchi, A.J., VanBriesen, J.M. and Rittmann, B.E., 1998. Multicomponent transport with coupled geochemical and microbiological reactions: model description and example simulations. Journal of Hydrology, 209: 8-26.

Theis, T. L., Kirkner, D.J., and Jennings, A.A., 1982. Multi-Solute Subsurface Transport Modeling for Energy Solid Wastes. Tech. Progress Rept. C00-10253-3, Department of Civil Engineering, University of Notre Dame, Norte Dame, IN.

Tompson, A.F.B., 1993. Numerical simulation of chemical migration in physically and chemically heterogeneous porous media. Water Resources Research, 29(11): 3709-3726.

Toride, N., Leij, F.J. and Genuchten, M.T.v., 1993. A comprehensive set of analytical solutions for nonequilibrium solute transport with first-order decay and zero-order production. Water Resources Research, 29: 2167-2182.

Truesdell, A.H., and Jones, B.F., 1974. WATEQ, a Computer Program for Calculating

Chemical Equilibria of Natural Water: Journal of Research of the United States Geological Survey, v.2, p. 233-248.

Westall, J.C., Zachary, J.L., and Morel, M.F.,1976. MINEQL: A Computer Program for the Calculation of Chemical Equilibrium Composition of Aqueous System. Technical Note 18, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA, 91 pp.

Wolery, T.J., 1992. EQ3/6, A Software Package for Geochemical Modeling of Aqueous Systems: Package Overview and Installation Guide: UCRL-MA-110662 PT 1, Lawrence Livermore National Laboratory, Livermore, California.

Wood, B. D., Dawson, C.N., Szecsody, J.E. and Streile, G.P., 1994. Modeling contaminant transport and biodegradation in a layered porous media system, *Water Resources Research* , 30(6): 1833-1845.

Yeh, G.T., Tripathi, V.S., 1991. A model for simulating transport of reactive multispecies components: Model development and demonstration. Water Resources Research, 27: 3075-3094.

Yeh, G.T., Iskra G, Zachara J.M., Szecsody J.E., and Streile G.P., 1995. KEMOD: a Mixed Chemical Kinetic and Equilibrium Model of Aqueous and Solid Phase Geochemical Reactions. PNL-10380, Pacific Northwest Laboratory.

Yeh, G. T., Salvage, K., and Choi, W.H., 1996. Reactive Multispecies-Multicomponent Chemical Transport Controlled by both Equilibrium and Kinetic Reactions. Proc. XI-th Int. Conf. on Numerical Methods in Water Resources, Cancun, Mexico, July 22-26, 1996. p. 585-592

Yeh, G. T. and Salvage, K.M. 1998. Modeling reactive chemical transport. Proc. XII-th Int. Conf. on Numerical Methods in Water Resources, Crete, Greece, June 15-19,

1998. p.271-278.

Yeh, G.T. Salvage, K.M., Gwo, J.P., Zachara, J.M., and Szecsody, J.E. 1998. HYDROBIOGEOCHEM: A Coupled Model of HYDROlogic Transport and Mixed BIOGEOCHEMical Kinetic/Equilibrium Reactions in Saturated-Unsaturated Media. Technical Report. Department of Civil and Environmental Engineering, Pennsylvania State University, PA 16801.

Yeh, G.T., Burgos W.D., Fang, Y.L. and Zachara J.M., 2000. Modeling biogeochemical kinetics: issues and data needs. XIII International Conference on Computational Methods in Water resources, 2000.

Yeh, G.T., Burgos W.D., and Zachara J.M., 2001a. Modeling and measuring biogeochemical reactions: System consistency, data needs, and rate formulation. *Adv. Environ. Res*., 5, 219-237, 2001.

Yeh, G. T., Siegel, M.D., and Li, M.H., 2001b. Numerical Modeling of Coupled Fluid Flows and Reactive Transport Including Fast and Slow Chemical Reactions. J. Contaminant Hydrology. 47:379-390.

Yeh, G. T. and Fang, Y., 2002. BIOGEOCHEM 2-0: A Numerical Model to Simulate BIOGEOCHEMical Reactions under Multiple Phase Systems. Technical Report. Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816.

van der Zee, S.E.A.T.M. and van Riemsdijk, W.H., 1987. Transport of reactive solute in spatially variable soil systems. Water Resources Research, 23(11): 2059-2069

CHAPTER 3.
REACTIVE TRANSPORT UNDER MULTIPHASE SYSTEM: – MODEL FORMULATION AND NUMERICAL SCHEMES

Yilin Fang[1] , Gour-Tsyh Yeh[2*] and William D. Burgos[1]

[1]Department of Civil and Environmental Engineering
The Pennsylvania State University
University Park, PA 16802

[2]Department of Civil and Environmental Engineering
University of Central Florida
Orlando, FL 32816-2450

ABSTRACT

This paper presents the development of a numerical model, HYDROBIOGEOCHEM, modeling reactive transport processes in the subsurface system composed of three fluid phases (gas, aqueous, and NAPL if present) and the solid soil matrix. The formulation of this model is based on the general paradigm for modeling reactive chemicals in batch systems described in model BIOGEOCHEM, so that BIOGEOCHEM can be easily coupled with HYDROBIOGEOCHEM. The uniqueness of the general paradigm lies in that it can (1) facilitate the segregation of kinetic reactions and, thus, enable the formulation and paramertization of individual rates one reaction by one reaction; (2) enable the inclusion of as many types of reactions as possible; (3) accomplish automation in decoupling fast reactions from slow reactions; (4) perform systematic operations to remove redundant fast reactions and irrelevant kinetic reactions; (5) achieve the reduction of problem stiffness as much as possible; (6) increase the robustness of numerical integration of the governing equations; and (7) systematically define chemical components and explicitly enforce mass conservation. Based on the conservative form and advective form of the governing equations, five different numerical schemes were used to solve the governing transport equations. They are: (1) FEM on advective form of equation; (2) FEM on conservative form of equation; (3)

Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on advective form of equation for boundary elements; (4) hybrid Lagrangian-Eulerian FEM; and (5) Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on conservative form of equation for boundary elements. The advantages, drawbacks of these schemes, and the application of HYDROBIOGEOCHEM are illustrated by four examples.

Keywords: Biogeochemical Modeling; Reaction Network; Matrix Decomposition; Reactive Chemical Transport; Multiphase System; Conservative Form; Advective Form.

3.1 INTRODUCTION

3.1.1 Background

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 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. The insufficiency of these models to describe the complex biogeochemical reactions in subsurface environments was itemized in Zhu et al. (2001).

"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 Erikson, 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; Steefel 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).

Models coupling transport with mixed equilibrium/ kinetic reactions have appeared since the mid-1990s (e.g., Steefel and Lagasa, 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; Yeh and Salvage, 1998; Yeh et al., 2001a; Brun and Engesgaard, 2002). For most models chemical components must be selected *a prior* such that only limited reaction network can be considered (e.g., Parkurst, 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 (Steefel 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).

Most of past modeling focus was on the chemical processes between groundwater

and the solid matrix in the subsurface system. For the few multiphase reactive transport models, such as models in Lichtner and Seth (1996), Viswanathan et al. (1998), Xu and Pruess (2001), their formulation are limited to that of water phase reactive transport only.

In this study, a numerical model, HYDROBIOGEOCHEM, based on a new formulation is presented. A comprehensive suite of reaction types and reaction rates are incorporated into the model. Reaction types include aqueous complexation, adsorption-desorption, ion-exchange, oxidation-reduction, precipitation-dissolution, acid-base reactions, and microbial mediated reactions. Phase transfer and decay of species will not be separately considered in hydrologic transport equations, instead, they are treated as reactions in HYDROBIOGEOCHEM. An equilibrium reaction is modeled with an infinite rate governed by a mass-action equilibrium equation or by a users-specified algebraic equation. Programmed kinetic reaction rates include multiple Monod kinetics, n-th order empirical, and elementary formulations.

3.1.2 Outline

This paper is organized in the following order. In section 3.2, the general paradigm for model BIOGEOCHEM is briefly reviewed. Conservative form of the governing equations for hydrologic transport is derived in section 3.3.1. Section 3.3.2 derives the non-conservative form or advective form of the governing transport equations. Boundary conditions for the two forms of transport equations are given in section 3.3.3, followed by the options of different numerical implementations in section 3.3.4. Section 3.4 presents examples used to verify the model HYDROBIOGEOCHEM and show its potential application.

3.2 PARADIGM REVIEW

Before deriving the governing equations for hydrologic transport, the general paradigm for modeling reactive chemicals in batch systems described in Fang et al.

(2003) will be briefly reviewed in the following.

The conceptualization of reaction network is a prerequisite to modeling geochemical and biochemical processes. A reaction network is defined by specifying chemical species and biogeochemical reactions that produce them (Yeh et al., 2001a). From a mathematical point of view, a system of M_s ordinary differential equations can be written for M_s chemical species in a reactive, well-mixed batch system as

$$\frac{dC_i}{dt} = r_i|_N, \ i \in M_s \tag{1}$$

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, 2001b).

In a reaction-based model, the production-consumption rates of M_s species are described by

$$\frac{dC_i}{dt} = r_i|_N = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k, \ i \in M_s \tag{2}$$

where, v_{ik} is the reaction stoichiometry of the i-th species in the k-th reaction associated with the products, μ_{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 (2) 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: $\frac{\partial C_i}{\partial t} + L(C_i) = r_i|_N = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k$. 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.

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.

With the above discussion, Eq. (2) is decomposed into the following equation via the Gauss-Jordan elimination (Chilakapati, 1995; Steefel and MacQuarrie, 1996; Yeh et al., 2001b)

$$\mathbf{B}\frac{dC}{dt} = \begin{bmatrix} \mathbf{D} & \mathbf{K} \\ \mathbf{0}_1 & \mathbf{0}_2 \end{bmatrix} R \tag{3}$$

where **B** is the reduced unit matrix, **D** is the diagonal matrix representing a submatrix of the reduced reaction-matrix with size of N_I x N_I reflecting the effects of N_I linearly independent reactions on the production-consumption rate of all kinetic-variables, **K** is a submatrix of the reduced reaction-matrix with size of N_I x 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 x N_I, and $\mathbf{0}_2$ is a zero matrix representing a submatrix of the reduced reaction-matrix with size N_C x N_{KD}.

The decomposition of Eq. (2) to Eq. (3) 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 users-specified algebraic equation; 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 -

$$\frac{dE_i}{dt} = D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j, \; k \in N_E, \; i \in M_s : \frac{d}{dt}\left(\frac{E_i}{D_{kk}}\right) \approx R_k = \infty \qquad (4)$$

$$\exists \text{ a thermodynamically consistent equation}$$

Governing Equations for (N_I - N_E) Kinetic Variables –

$$\frac{dE_i}{dt} = D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j, \; k \in N_{KI}, \; i \in M_s \qquad (5)$$

and *Mass Conservation Equations for N_C Chemical Components* –

$$\frac{dE_i}{dt} = 0; \quad \text{denoting } T_j = E_i, \; j \in N_C, \; i \in M_s \qquad (6)$$

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. The variable, E_i, in Eq. (4) is called an equilibrium variable, the variable E_i in Eq. (5) is called a kinetic variable, and the variable E_i in Eq. (6) is called a component variable and is normally denoted with T_j.

The formulation of the model HYDROBIOGEOCHEM in the present study is based on the general paradigm for modeling reactive chemicals in batch systems described in model BIOGEOCHEM, so that BIOGEOCHEM can be easily coupled with HYDROBIOGEOCHEM. The uniqueness of the general paradigm lies in that it can (1) facilitate the segregation of kinetic reactions and, thus, enable the formulation and paramertization of individual rates one reaction by one reaction; (2) enable the inclusion

of as many types of reactions as possible; (3) accomplish automation in decoupling fast reactions from slow reactions; (4) perform systematic operations to remove redundant fast reactions and irrelevant kinetic reactions; (5) achieve the reduction of problem stiffness as much as possible; (6) increase the robustness of numerical integration of the governing equations; and (7) systematically define chemical components and explicitly enforce mass conservation.

3.3 GOVERNING EQUATIONS FOR TRANSPORT

According to custom definition, species in each fluid phase is considered to be mobile and every solid or solid associated species is considered to be immobile. In the present model, each mobile species (including aqueous-phase species, mobile NAPL phase species, and air-phase species) concentration has the unit of moles per mass of fluid, while immobile species has the unit of moles per unit mass of rock matrix, excluding precipitated species concentration, which has a unit of moles per bulk volume. With the concentration unit defined, governing equation for transport of mobile chemical species can be derived from the principle of mass balance as

$$\frac{D}{Dt}\int_{\nu}\rho_{\alpha}\theta_{\alpha}c_i^{\alpha}d\nu = -\int_{\Gamma}n\bullet(\rho_{\alpha}\theta_{\alpha}c_i^{\alpha}\mathbf{V}_{fs,\alpha})d\Gamma - \int_{\Gamma}\mathbf{n}\bullet\mathbf{J}_i^{\alpha}d\Gamma + \int_{\nu}r_i d\nu + \int_{\nu}m_i^{\alpha}d\nu \qquad (7)$$

where ν is the material volume containing constant amount of media (L^3); Γ is the surface enclosing the material volume ν (L^2); $\mathbf{n}$ is the outward unit vector normal to the surface Γ; ρ_{α} is the fluid density of phase α, α denoting water, gas or oil, (M/L^3); θ_{α} is the volume fraction of fluid α within the representative elementary volume (REV), (L^3/L^3), $\theta_{\alpha}=\phi S_{\alpha}$; ϕ is the effective porosity; S_{α} is the fluid saturation; c_i^{α} is the concentration of the i-th chemical species in fluid phase α (M/M of fluid); $\mathbf{J}_i^{\alpha}$ is the surface flux of the i-th chemical species with respect to the velocity of fluid α [$(M/T)/L^2$];

and $\mathbf{J}_i^{\alpha} = \rho_{\alpha}\theta_{\alpha}c_i^{\alpha}(\mathbf{V}_{i,\alpha} - \mathbf{V}_{f,\alpha})$; $\mathbf{V}_{i,\alpha}$ is the local velocity of the i-th chemical species in phase α (L/T); $\mathbf{V}_{f,\alpha}$ is the local fluid velocity of phase α (L/T); $\mathbf{V}_{f,\alpha} = \mathbf{V}_{fs,\alpha} + \mathbf{V}_s$; $\mathbf{V}_{fs,\alpha}$ is the fluid velocity of phase α relative to the solid surface (L/T); $\mathbf{V}_s$ is the solid velocity (L/T); m_i^{α} is the external source/sink rate of the i-th species in phase α per unit bulk volume [(M/L^3)/T]; r_i^{α} is the production/consumption rate of the i-th species in phase α due to all chemical reactions (decay of the i-th species is also considered as a reaction)[(M/ L^3 bulk volume)/T] and it can be expressed by

$$r_i = \sum_{k=1}^{N}(\nu_{ik} - \mu_{ik})R_k \tag{8}$$

where ν_{ik} is the stoichiometric coefficient of the i-th species in the k-th reaction as a product; μ_{ik} is the stoichiometric coefficient of the i-th species in the k-th reaction as a reactant.

By the Reynolds transport theorem (Owczarek, 1964) and Gaussian divergence theorem, neglecting the deformation of the porous media, the following equation is obtained:

$$\frac{\partial \rho_{\alpha}\theta_{\alpha}c_i^{\alpha}}{\partial t} + \nabla \bullet (\rho_{\alpha}c_i^{\alpha}\mathbf{V}_{\alpha}) + \nabla \bullet \mathbf{J}_i^{\alpha} = \sum_{k=1}^{N}(\nu_{ik} - \mu_{ik})R_k + m_i^{\alpha} \tag{9}$$

The surface flux can be expressed in terms of the gradient of c_i^{α} (Nguyen et al., 1982):

$$\mathbf{J}_i^{\alpha} = -\rho_{\alpha}\theta_{\alpha}\mathbf{D}_{\alpha} \bullet \nabla c_i^{\alpha} \tag{10}$$

and the dispersion tensor $\mathbf{D}_{\alpha}$ of phase α, according to Millington and Quirk (1961) (from Gerke et al., 1998), can be given in the standard form as

$$\theta_{\alpha}\mathbf{D}_{\alpha} = a_T|\mathbf{V}_{\alpha}|\delta + (a_L - a_T)\mathbf{V}_{\alpha}\mathbf{V}_{\alpha}/|\mathbf{V}_{\alpha}| + a_{m\alpha}\theta_{\alpha}\tau_{\alpha}\delta \tag{11}$$

where D_{α} is the dispersion coefficient tensor of phase i, (L^2/T); a_T is the transverse disffusivity (L); a_L is the longitudinal diffusion coefficient (L^2/T); $a_{m\alpha}$ is the molecular

diffusion coefficient in pure phase α (L^2/T); τ_α is the tortuosity factor of the medium for fluid phase α; δ is the Kronecker delta tensor; and $\mathbf{V}_\alpha$ is the Darcy velocity of phase α (L/T), it is related with fluid velocity by $\mathbf{V}_\alpha = \mathbf{V}_{f,\alpha}\theta_\alpha$.

$$\frac{\partial \rho_\alpha \theta_\alpha c_i^\alpha}{\partial t} + \nabla \bullet (\rho_\alpha c_i^\alpha \mathbf{V}_\alpha) - \nabla \bullet (\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla c_i^\alpha) = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k + m_i^\alpha \tag{12}$$

Immobile species are not subject to hydrological transport, the mass balance equation for these species can be obtained from Eq. (12) by dropping advection, dispersion/diffusion, and external source/sink terms and replacing $(\rho_\alpha \theta_\alpha c_i^\alpha)$ with $(\rho' c_i)$:

$$\frac{\partial \rho' c_i}{\partial t} = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k \tag{13}$$

where $\rho' = 1.0$ for precipitated species and ρ' = the bulk density for other immobile species.

3.3.1 Governing Equations for Primary Dependent Variables (PDVs)

It is noted that Eq. (1) can be easily extended to transport systems by replacing the ordinary differentiation 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 Fang et al. (2003) is also applicable to transport systems except that, instead of Eq. (1), one employs the set of transport equations.

Decompose the matrix equation made up of the transport equations of mobile species and the mass balance of immobile species, the following set of governing equations are obtained:

· *Transport Equation for Mobile Components* –

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right] = M_j, \quad j \in N_m \tag{14}$$

· *Transport Equation for Mobile Kinetic Variables –*

$$\frac{\partial E_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right]$$

$$= M_j + D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j, \; k \in N_{KI}, \; j \in M_m \tag{15}$$

· *Infinite-Rate Equations for N_E Equilibrium Reactions* -

$$\frac{d E_j}{dt} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right]$$

$$= M_j + D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j, \; k \in N_E, : \frac{d}{dt}\left(\frac{E_j}{D_{kk}} \right) \approx R_k = \infty \tag{16}$$

$$\exists \text{ a thermodynamically consistent equation}$$

where T_j is the total concentration of the j-th component, T_j^α is the concentration of the j-th component j in phase α; E_j is the total concentration of the j-th kinetic-variable; E_j^α is the concentration of the j-th kinetic variable in fluid phase α; N_m is the number of mobile components, M_m is the number of mobile kinetic variables; and M_{im} is the number of immobile kinetic variables. The number of immobile components is $N_{im} = N_C - N_m$; the number of immobile kinetic variable is $M_{im} = M - M_m$.

Governing equations for immobile components and kinetic variables are in the same form as Eqs. (14) and (15) by dropping advection, dispersion/diffusion, and external source/sink terms in Eqs. (14) and (15).

For the convenience of presentation, Eqs. (14) and (15) will be expressed by the following equation in a general form:

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right] = M_j + RR_j \tag{17}$$

where $RR_j = 0$ for components, and $RR_j = D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j$ for kinetic variables.

To solve Eq. (17) numerically, T_j^{α} is expressed in terms of T_j to make T_js as PDVs and the following equation is obtained:

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\frac{T_j^{\alpha}}{\rho_{\alpha} \theta_{\alpha} T_j} \rho_{\alpha} \theta_{\alpha} \mathbf{V}_{f,\alpha} T_j \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_{\alpha} \theta_{\alpha} \mathbf{D}_{\alpha} \bullet \nabla \left(\frac{T_j^{\alpha}}{\rho_{\alpha} \theta_{\alpha} T_j} T_j \right) \right) = M_j + RR_j \tag{18}$$

which is conservative form of the transport equations.

3.3.2 Conversion to advective Form

Differential Eq. (18) is the conservative or primitive form of the governing equations for hydrologic transport. They are derived from the law of mass conservation. They can also be converted to the non-conservative or advective form by expanding the second term of Eq. (18), i.e., divergence of advection flux. Though conservative form and advective form are mathematically equivalent, the numerical solutions for them have different characteristics. The conservative form offers advantages in preserving conservative properties in numerical implementations. On the other hand, the advective form offers some advantages in preserving the transport property in numerical implementations by the finite element method (FEM) (Huyakorn et al., 1985), and facilitates the implementation of Lagrangian-Eulerian approach.

However, for FEM in advective form, mass conservation property cannot be satisfied, while for Lagrangian–Eulerian methods, (1) mass conservative property cannot be satisfied, (2) boundary conditions are not easy to implement. We shall see later on the different features of the numerical implementation for these two forms.

Eq. (18) is used for illustration in the following derivation. Expanding the second term of Eq. (18) yields

$$\frac{\partial T_j}{\partial t}+\sum_{\alpha=1}^{3}\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha}\nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\right)+\sum_{\alpha=1}^{3}\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha}\right)$$
$$-\sum_{\alpha=1}^{3}\nabla\bullet\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha}\right)\right]=M_j+RR_j \tag{19}$$

Express T_j^{α} in terms of T_j to make T_js as PDVs:

$$\frac{\partial T_j}{\partial t}+\sum_{\alpha=1}^{3}\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}T_j\nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\right)+\sum_{\alpha=1}^{3}\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}T_j\right)$$
$$-\sum_{\alpha=1}^{3}\nabla\bullet\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}T_j\right)\right]=M_j+RR_j \tag{20}$$

Expanding the third term on the left hand side of Eq. (20), we obtain

$$\sum_{\alpha=1}^{3}\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}T_j\right)$$
$$=T_j\sum_{\alpha=1}^{3}\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}\right)+\sum_{\alpha=1}^{3}\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j \tag{21}$$

Expanding the fourth term on the right hand side of Eq. (20), we obtain

$$\sum_{\alpha=1}^{3}\nabla\bullet\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}T_j\right)\right]$$
$$=\sum_{\alpha=1}^{3}\nabla\bullet\left\{\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j+\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}\right)\bullet\nabla T_j\right\}$$
$$=\sum_{\alpha=1}^{3}\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}\right)\right]\bullet\nabla T_j+T_j\sum_{\alpha=1}^{3}\nabla\bullet\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}\right)\right]$$
$$+\sum_{\alpha=1}^{3}\nabla\bullet\left[\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^{\alpha}}{\rho_\alpha\theta_\alpha T_j}\right)\bullet\nabla T_j\right] \tag{22}$$

Substituting Eqs. (21), (22) into Eq. (20) and combining similar terms, the following equation is obtained

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3}\left[\rho_\alpha\theta_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\mathbf{V}_{f,\alpha} - \rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right]\bullet\nabla T_j$$
$$-\sum_{\alpha=1}^{3}\nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\right) \qquad (23)$$
$$+\sum_{\alpha=1}^{3}\left[\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\nabla\bullet(\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}) + \rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j} - \nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j$$
$$= M_j + RR_j$$

Combining Eq. (23) and the continuity equation of flow

$$\frac{\partial\rho_\alpha\theta_\alpha}{\partial t} = -\nabla\bullet(\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}) + \rho_\alpha^* Q_\alpha,\ \alpha = \text{water, gas or oil} \qquad (24)$$

yields the following equation

$$\frac{\partial T_j}{\partial t} + L(T_j) + KT_j + \sum_{\alpha=1}^{3}\frac{\rho_\alpha^* Q_\alpha T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j = M_j + RR_j \qquad (25)$$

where ρ_i^* is the density of source/sink fluid, (M of liquid/L^3 of liquid), equal to ρ_i if sink/withdrawal (negative Q) or $\rho_{i,in}$ if source/injection (positive Q); Q is the source or sink due to artificial injection or withdrawal of water, unit of volume of liquid per volume of porous media per time (L^3/L^3 T). K is a first order term and L( ) is an advection-dispersion operator:

$$K = \sum_{i=1}^{3}\left[\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right) - \frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\frac{\partial\rho_\alpha\theta_\alpha}{\partial t} - \nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right] \qquad (26)$$

$$L(T_j) = \mathbf{V}_e\bullet\nabla T_j - \nabla\bullet[\mathbf{D}_e\bullet\nabla T_j],\ j\in N_m \qquad (27)$$

where $\mathbf{V}_e$ and $\mathbf{D}_e$ are effective velocity and effective dispersion coefficient respectively, and they are given by

$$\mathbf{V}_e = \sum_{\alpha=1}^{3}\left[\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right) - \rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^i}{\rho_\alpha\theta_\alpha T_j}\right] \tag{28}$$

$$\mathbf{D}_e = \sum_{\alpha=1}^{3}\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right) \tag{29}$$

The source/sink term can be defined by

$$M_j = \sum_{i=1}^{3}\left(\rho_\alpha^* T_j^{\alpha*} Q\right) \tag{30}$$

where $T_j^{\alpha*}$ is the total concentration of the j-th component in the α-th source/sink liquid (M/M of liquid), and $T_j^{\alpha*} = \dfrac{T_j^i}{\rho_\alpha\theta_\alpha}$ in the case of sink/withdrawl, or $T_j^{\alpha*} = T_{j,in}^{\alpha}$ -- input concentration in the case of source/injection.

3.3.3 Initial and Boundary Conditions

Initially, the concentration of each chemical species must be given throughout the region of interest:

$$\rho_\alpha\theta_\alpha c_i = C_{i0} \text{ at } t = 0,\ i \in M_s \tag{31}$$

where C_{i0} is the initial concentration of the i-th chemical species, unit of mass of chemical species per volume of porous media (M/L^3).

For all species that are subject to hydrological transport, boundary conditions are specified. Types of boundary conditions incorporated are: Dirichlet, Neumann, Cauchy, or variable.

A Dirichlet condition is used to deal with a prescribed concentration as:

$$\rho_\alpha\theta_\alpha c_i = C_{iD} \quad \text{on } B_D, i \in M_s \tag{32}$$

Where C_{iD} is the prescribed Dirichlet concentration for the i-th chemical species, unit of mass of chemical species per volume of porous media (M/L^3).

A boundary is often treated as a Neumann boundary when the flux due to the

gradient of species concentration is known, or as a Cauchy boundary when the total flux is given:

$$-\mathbf{n}\bullet(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla c_i)=q_{iN,\alpha}\quad\text{on } B_N,\ i\in M_s \tag{33}$$

$$\mathbf{n}\bullet(\rho_\alpha\mathbf{V}_\alpha c_i-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla c_i)=q_{iC,\alpha}\quad\text{on } B_C,\ i\in M_s \tag{34}$$

where q_{iN} is the normal Neumann flux, unit of mass of the i-th chemical species per unit area of porous media per time ($M/L^2/T$), q_{iC} is the normal Cauchy flux, unit of mass of the i-th chemical species per unit area of porous media per time($M/L^2/T$), **n** is an outward unit vector normal to the boundary.

A flow-through boundary is normally the soil-air interface or soil-water interface, and the appropriate condition is variable depending on the direction of transport across the boundary. For this type of boundary, a variable-type boundary is imposed. When the flow is directed out of the region, the boundary is depicted by the Neumann condition with zero gradient flux; when the flow is directed into the region, the boundary is depicted by the Cauchy condition with given total flux. Therefore, the variable boundary condition can be written as

$$\begin{aligned}&-\mathbf{n}\bullet(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla c_i)=q_{iV,\alpha}=0\quad\text{if } \mathbf{V}_\alpha\bullet\mathbf{n}>0\text{ on } B_V,\ i\in M_s\\&\mathbf{n}\bullet(\rho_\alpha\mathbf{V}_\alpha c_i-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla c_i)=q_{iV,\alpha}\quad\text{if } \mathbf{V}_\alpha\bullet\mathbf{n}<0\text{ on } B_V,\ i\in M_s\end{aligned} \tag{35}$$

Expressed in the form of total phase concentration, Eqs. (31) to (35) can be written as:

$$T_j=T_{j_0}\quad\text{at } t=0,\ j\in(N_c+M) \tag{36}$$

$$-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jN,\alpha}\quad\text{on } B_N,\ j\in(N_m+M_m) \tag{37}$$

$$\mathbf{n}\bullet\left(\rho_\alpha\mathbf{V}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jC,\alpha}\quad\text{on } B_C,\ j\in(N_m+M_m) \tag{38}$$

$$-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jV,\alpha}=0\quad\text{if }\mathbf{V}_\alpha\bullet\mathbf{n}>0\text{ on }B_V,\ j\in(N_m+M_m)\tag{39}$$

$$\mathbf{n}\bullet\left(\rho_\alpha\mathbf{V}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jV,\alpha}\quad\text{if }\mathbf{V}_\alpha\bullet\mathbf{n}<0\text{ on }B_V,\ j\in(N_m+M_m)\tag{40}$$

When applying FEM to the advective form of the governing equations, boundary conditions should be reformulated to be consistent with Eq.(25) (regarding to the effective velocity and effective dispersion tensor).

Express Eq. (37) in terms of T_j, we obtain

$$\begin{aligned}&-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j\right)\right)\\&=-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\left(\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)T_j+\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\right)=q_{jN,\alpha}\end{aligned}\tag{41}$$

Add $\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\dfrac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j$ to both sides of Eq. (37), we obtain

$$\begin{aligned}&\mathbf{n}\bullet\left[\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j-\mathbf{n}\bullet\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\\&=\mathbf{n}\bullet\mathbf{V}_{f,\alpha}\frac{T_j^\alpha}{T_j}T_j+q_{jN,\alpha},\ \text{on }B_N,\ j\in(N_m+M_m)\end{aligned}\tag{42}$$

Similarly, express Eq. (38) in terms of T_j, we obtain

$$\mathbf{n}\bullet\left(\rho_\alpha\mathbf{V}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j\right)=q_{jC,\alpha}\tag{43}$$

Expand the second term on the left hand side of Eq. (43) as in Eq. (41) and combine similar terms, the following equation is obtained

$$\mathbf{n}\bullet\left[\rho_\alpha \mathbf{V}_\alpha \frac{T_j^i}{\rho_\alpha\theta_\alpha T_j} - \rho_\alpha\theta_\alpha \mathbf{D}_\alpha \bullet \nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right] T_j - \mathbf{n}\bullet\rho_\alpha\theta_\alpha \mathbf{D}_\alpha \frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j \tag{44}$$
$$= q_{jC,\alpha} \text{ on } B_C,\ j\in(N_m + M_m)$$

Eq. (39) and (40) are special case of Eqs. (37) and (38), respectively, and they can be easily written out as

$$\mathbf{n}\bullet\left[\rho_\alpha\theta_\alpha \mathbf{V}_{f,\alpha} \frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j} - \rho_\alpha\theta_\alpha \mathbf{D}_\alpha \bullet \nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right] T_j - \mathbf{n}\bullet\rho_\alpha\theta_\alpha \mathbf{D}_\alpha \frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j$$
$$= \mathbf{n}\bullet\mathbf{V}_{f,\alpha}\frac{T_j^\alpha}{T_j} T_j \text{ if } \mathbf{V}_\alpha\bullet\mathbf{n} > 0 \text{ on } B_V,\ j\in(N_m + M_m) \tag{45}$$

$$\mathbf{n}\bullet\left[\rho_\alpha \mathbf{V}_\alpha \frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j} - \rho_\alpha\theta_\alpha \mathbf{D}_\alpha \bullet \nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right] T_j - \mathbf{n}\bullet\rho_\alpha\theta_\alpha \mathbf{D}_\alpha \frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j = q_{jV,\alpha} \tag{46}$$
$$\text{if } \mathbf{V}_\alpha\bullet\mathbf{n} < 0 \text{ on } B_V,\ j\in(N_m + M_m)$$

3.3.4 Numerical Implementation

Finite element method (FEM) is used to solve the transport equations. There are five numerical schemes included in HYDROBIOGEOCHEM, based on the conservative form and advective form of the governing equations. They are: scheme 1, FEM on advective form of equation with upstream weighting; scheme 2, FEM on conservative form of equation with upstream weighting; scheme 3, Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on advective form of equation for boundary elements; scheme 4, hybrid Lagrangian-Eulerian FEM; and scheme 5, Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on conservative form of equation for boundary elements. Scheme 4 is different from traditional hybrid Lagrangian-Eulerian FEM in that it first approximates the time derivative term with finite difference method and multiplies both sides of the governing equation with the time interval. The Galerkin FEM is then applied to the resulted residual equation and we'll obtain the linearized equation as Eq.

(A.18) in appendix 3.A, where the numerical solutions of the transport equation by FEM for all five schemes are presented. We call the traditional Lagrangian-Eulerian FEM method as scheme 6 in the example simulation section.

Three approaches for solving coupled hydrologic transport and biogeochemical reaction problems are included in HYDROBIOGEOCHEM. These approaches are fully implicit sequential iteration approach, operator splitting approach, and predictor-corrector approach. In all three approaches, only the mobile PDVs are solved in the transport module. The immobile PDVs are solved in the chemistry module. For the simplicity of presentation, without losing generality, the transport equation can be expressed as

$$\frac{\partial T}{\partial t} = L(T) + R(T) \tag{47}$$

where L( ) is an advective-dispersive transport operator and R( ) is a reaction operator.

For the fully implicit sequential iteration approach, Eq. (47) becomes:

$$\frac{T^{n+1,r+1} - T^{n}}{\Delta t} = L(T^{n+1,r+1}) + R(T^{n+1,r+1}) \tag{48}$$

and the reaction term will be linearized by a Taylor series approximation:

$$R(T^{n+1,r+1}) = R(T^{n+1,r}) + \sum_{j=1}^{N_m+M_m} \left(\frac{\partial R}{\partial T_j}\right)^{n+1,r} \left(T_j^{n+1,r+1} - T_j^{n+1,r}\right) \tag{49}$$

Where n = time step level, and r = iteration step.

In the fully implicit approach, the biogeochemical reaction subsystem of equations and the hydrologic transport subsystem of equations are solved sequentially and iteratively. The concentrations of all immobile PDVs are computed in the biogeochemical subsystem using all species concentrations from the prior iteration. The concentrations of the mobile PDVs are not solved in the reaction subsystem. They are determined in the hydrologic module using the species concentrations and the reaction term evaluated from species concentrations of the current iteration. These concentrations then passed to the reaction module as given constants.

In the operator splitting approach, the concentrations of mobile PDVs are solved for in both subsystems of equations. The mobile PDVs are first subjected to transport without considering the reaction term to yield an intermediate value of their concentrations. These transported concentrations are then subjected to the biogeochemical reactions, yielding final values of the concentrations for the current time step. In this approach, Eq. (47) is solved in two steps:

$$\frac{T^{n+1/2} - T^{n}}{\Delta t} = L(T^{n+1/2}) \tag{50}$$

$$\frac{T^{n+1} - T^{n+1/2}}{\Delta t} = R(T^{n+1}) \tag{51}$$

where the superscripts (n), (n+1/2), and (n+1) denote the values at the prior time, at an intermediate point in the current computation, and at the new time, respectively.

In the predictor-corrector approach the concentrations of mobile PDVs are also solved for in both subsystems of equations. The only difference of this approach from the operator splitting approach lies in whether reaction terms be considered in the transport equation. In this approach, Eq.(47) is solved in the following two steps:

$$\frac{T^{n+1/2} - T^{n}}{\Delta t} = L(T^{n+1/2}) + R(T^{n}) \tag{52}$$

$$\frac{T^{n+1} - T^{n+1/2}}{\Delta t} = R(T^{n+1}) - R(T^{n}) \tag{53}$$

where the superscripts (n), (n+1/2), and (n+1) denote the values at the prior time, at an intermediate point in the current computation, and at the new time, respectively.

In the operator splitting approach and the predictor-corrector approach, the coefficients in the L( ) operator are determined from previous iterate. No iteration between the transport subsystem and the reaction subsystem is implemented. However, when the PDVs have immobile components, these coefficients in the L( ) operator are dependent on the PDVs to be simulated. This implementation will give a less accurate solution.

Compared to the fully implicit approach, the operator splitting approach and predictor-corrector approach are less computationally intensive and more robust, but they sacrifice the accuracy of the solution (Salvage, 1998).

3.4 Example Simulations

3.4.1 1-D Solute Transport Verification

The numerical schemes have been compared with analytical solutions for simple cases with one adsorption reaction. This example solves the transient simulation of chemical transport in a horizontal beam. The domain of interest has a size of 150 cm × 1 cm. The flow beam contains water ($S_w = 1$) with an initial concentration of chemical C(x ,0) = 0 mmol/cm^3 of bulk volume. It also contains soil with a bulk density of 1.5 g/cm^3 and sorbed chemical S, initial concentration of which is S(x,0) = 0 mmol/cm^3 of bulk volume. A steady flow field is assumed with a constant Darcy velocity V = 1 cm/day and an effective moisture content of 0.2. The molecular diffusion coefficient is 0 cm^2 /day. The boundary condition at x = 0 is flow in boundary condition and the incoming concentration of chemical C is 5.0mmol/cm^3. At x = 150 cm is natural boundary condition. The domain is discretized with 100 equal size elements (1.5 cm × 1 cm each). A constant time step size of 0.1 day is used for the simulation. The simple equilibrium reaction considered is

$$C \leftrightarrow S,\ K_d = 0.1333 \tag{54}$$

where K_d is partition coefficient.

Five schemes presented in this paper are simulated separately and are compared with analytical solution given by Lindstrom and Freed (1967) in figure 3.1 and table 3.1 for concentration profile of chemical C given after 8 days. Scheme 6 is the traditional form of Lagrangian Eulerian FEM.

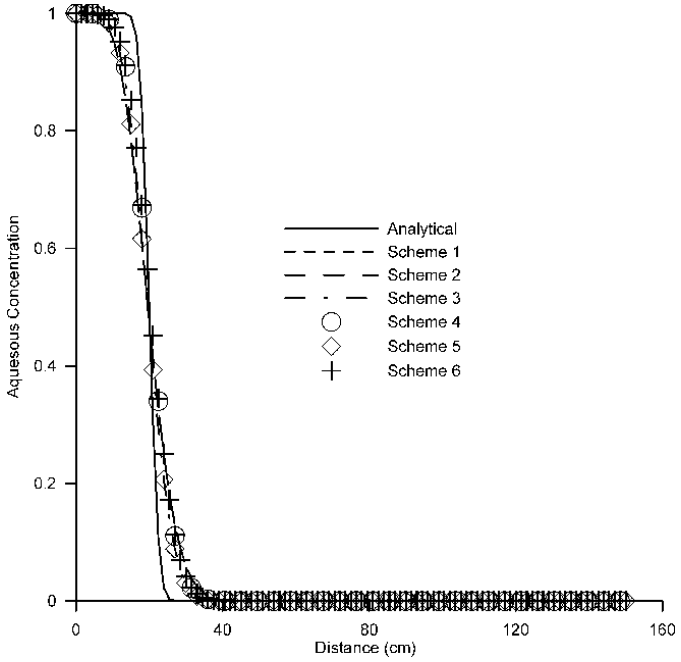
The grid Peclet number ($Pe = \frac{V\Delta x}{D} = \frac{V\Delta x}{\alpha_L V} = \frac{\Delta x}{\alpha_L}$) in the simulation corresponds to the three dispersivity, α_L = 0.1 cm, α_L = 10 cm, and α_L = 50 cm are Pe = 15, Pe = 0.15, and Pe = 0.03, respectively. Grid Courant number is less than 1. Differences of the different schemes are apparent from Figure 3.1.

It is seen from table 3.1 that for large Peclet number, scheme 4 gives more accurate result even though it shows some amount of numerical dispersion. However, for small Peclet number, accuracy of scheme 4 and 6 are bad because of the inaccurate treatment of boundary fluxes in the Lagrangian step of these two schemes. Although schemes 1 and 2 present high accuracy, they are susceptible to numerical dispersion, and limitations on the size of the Courant number. Therefore, for dispersion dominant problems with flux boundary conditions, we recommend scheme 3 and 5. For advection dominant problem, scheme 3 and 5 are also recommended because they ensure mass conservation.

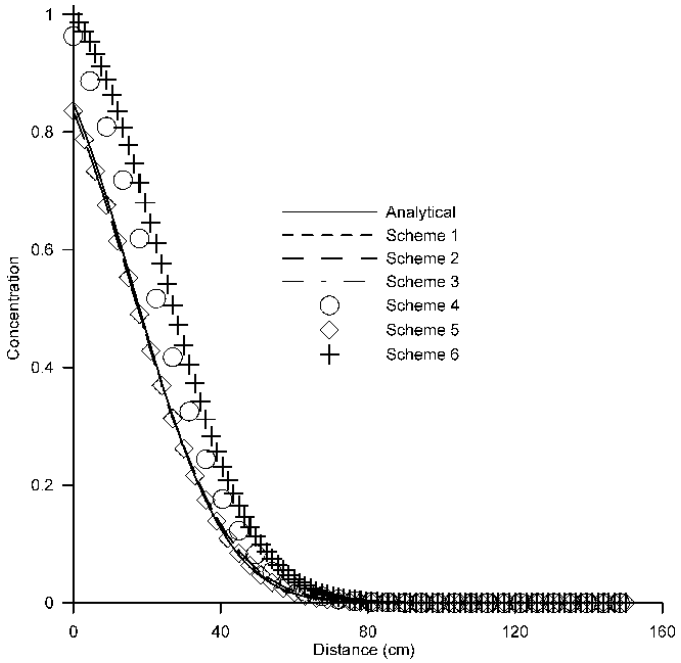
Table 3.1. R^2 values based on the analytical and calculated results for different values of dispersivity.

	Scheme 1	Scheme 2	Scheme 3	Scheme 4	Scheme 5	Scheme 6
Pe = 0.1	0.9660	0.9660	0.9752	0.9756	0.9752	0.9355
Pe = 10	0.9995	0.9995	0.9995	0.9542	0.9995	0.8286
Pe = 50	0.9999	0.9999	0.9998	0.4030	0.9998	-1.3554

(a)


Aquesous Concentration
Distance (cm)
Analytical
Scheme 1
Scheme 2
Scheme 3
Scheme 4
Scheme 5
Scheme 6
0
0.2
0.4
0.6
0.8
1
0
40
80
120
160

(b)


Concentration
Distance (cm)
Analytical
Scheme 1
Scheme 2
Scheme 3
Scheme 4
Scheme 5
Scheme 6
0
0.2
0.4
0.6
0.8
1
0
40
80
120
160

(c)

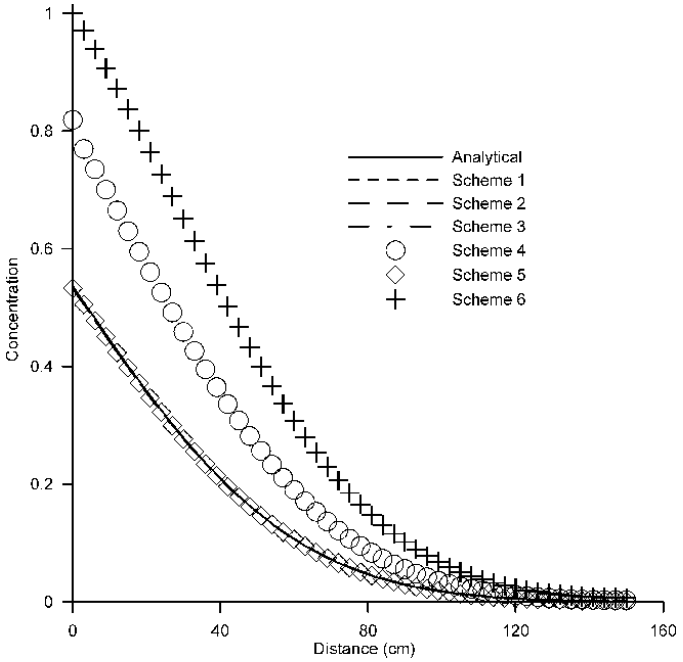


Figure 3.1 Concentration profile of chemical C at t = 8 days: (a) α_L = 0.1 cm, (b) α_L = 10 cm, (c) α_L = 50 cm.

3.4.2 1-D Transport: Chemical Decay and Productoin

The problem domain is the same as example 3.4.1. The boundary condition is the Dirichlet boundary condition, which involves a constant total concentration of 1.0 mol/L at x = 0, corresponding to constant concentration of c = 2.5 mol/kg of water. Processes considered in this problem are listed in table 3.2. Simulation parameters are given in table 3.3. A constant time step size of 0.05 day is used for the simulation. Analytical solution of the problem is given by van Genuchten and Alves (1982). The concentration profiles of species with different schemes are given and compared with the analytical solutions at selected times in figure 3.2 and table 3.4. They all show good agreement with the analytical solutions before the change of boundary conditions (t = 4 d). But after the change of boundary condition, scheme 4 shows the best accuracy, scheme 3 and 5 are next. This is due to the big concentration gradient generated after 4 days. Decreasing time

steps around 4 days, scheme 3 and 5 may give better results.

Table 3.2. Processes considered in problem 3.4.2.

Process	#	Type
c ↔ s	(R1)	Equilibrium reaction described by linear isotherm s = k c
Decay of c	(R2)	First-order kinetics, rate $R_2 = -\mu_w c$
Decay of s	(R3)	First-order kinetics, rate $R_3 = -\mu_s s$
Production of c	(R4)	Zero-order kinetics, rate $R_4 = \gamma_w$
Production of s	(R5)	Zero-order kinetics, rate $R_5 = \gamma_s$

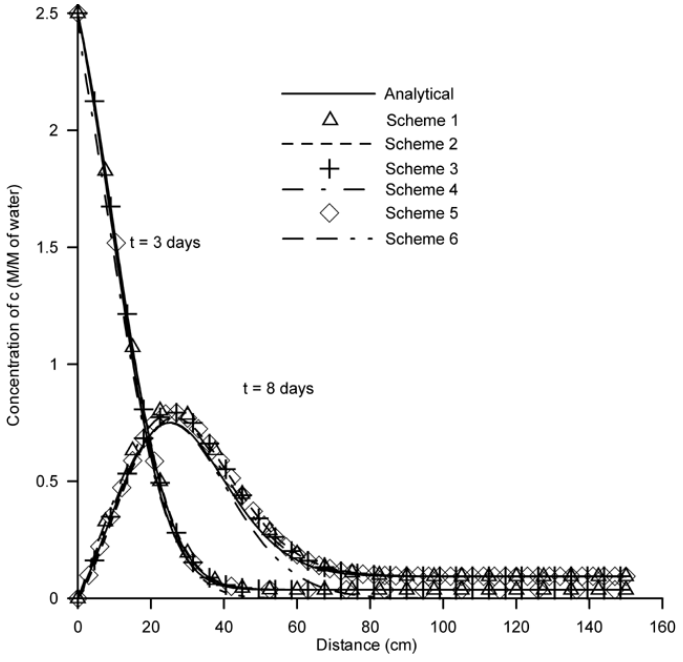


Figure 3.2. Concentration profile for problem 3.4.2.

Table 3.3. Physical and Chemical Parameters and values for Problem 3.4.2.

Parameters	Value
Effective porosity	$\varphi = 0.2$
Bulk density	$\rho_b = 1.5$ g/L
Longitudinal dispersivity	$\alpha_L = 10.0$ cm
Vertical dispersivity	$\alpha_T = 0.0$ cm
Tortuosity	$\tau = 1.0$
Aqueous diffusion coefficient	$a_{mw} = 0.0$ cm^2/d
Darcy velocity	V = 1.0 cm/d
Boundary condition	c = constant at x = 0 (0 < t < 4 d)
Initial species concentration	0.0 of bulk volume
Partition coefficient	k = 0.1333
first-order decay constant in water phase	$\mu_w = 4.0 \times 10^{-3}$ g/cm^3 d
first-order decay constant in solid phase	$\mu_s = 2.25 \times 10^{-4}$ g/cm^3 d
zero-order production rate constant in water phase	$\gamma_w = 2.0 \times 10^{-3}$ mol/cm^3 d
zero-order production rate constant in solid phase	$\gamma_s = 3.0 \times 10^{-3}$ mol/cm^3 d

Table 3.4. R^2 values based on the analytical and calculated results for selected time steps

	Scheme 1	Scheme 2	Scheme 3	Scheme 4	Scheme 5	Scheme6
t = 3 d	0.9987	0.9987	0.9989	0.9983	0.9989	0.9965
t = 8 d	0.9741	0.9892	0.9867	0.9915	0.9867	0.8945

3.4.3 1-D Diffusive Transport: Gas-Liquid Reaction

This problem is to verify the applicability of HYDROBIOGEOCHEM in multiphase system. Transport of hydrocarbon in the unsaturated zone due to phase partitioning is simulated and compared with analytical solution. Problem domain is identical to problem 3.4.1. Initial species concentration and boundary conditions are given in table 3.5.

Table 3.5 Physical and Chemical Parameters and values for Problem 3.4.3.

Parameters	Value
Effective porosity	$\varphi = 0.37$
Bulk density	$\rho_b = 1.5$ g/L
water volumetric content	$\theta_w = 0.10$
air volumetric content	$\theta_w = 0.27$
Longitudinal dispersivity	$\alpha_L = 0.0$ cm
Transverse dispersivity	$\alpha_T = 0.0$ cm
Tortuosity	$\tau = 0.34$
Aqueous diffusion coefficient	$a_{mw} = 8.64 \times 10^{-3}$ cm^2/d
Gaseous diffusion coefficient	$a_{ma} = 86.4$ cm^2/d
Boundary condition	G = constant at x = 0
Initial species concentration	0.0 mole per bulk volume

Processes included are:

· partitioning between the water and air phases for the hydrocarbon

$$G \leftrightarrow C,\ H = 1.05 \tag{55}$$

· adsorption

$$C \leftrightarrow S,\ K_d = 0.3 \tag{56}$$

where H is the Henry's law coefficient. G is in gaseous phase, C is in aqueous phase, and S is sorbed species.

Parameters used for this problem is shown in table 3.5. Most of the values are from Barhr (1987). A constant time step size of 1.0 day is used for the simulation.

Simulated results for gaseous species concentration with scheme 5 for selected times are compared with the analytical solution given in Baehr (1987) in Figure 3 and they show very good agreement. Other schemes are not compared for this example because for Dirichlet boundary conditions, there is not much difference among different numerical schemes as has already shown in problem 3.4.2.

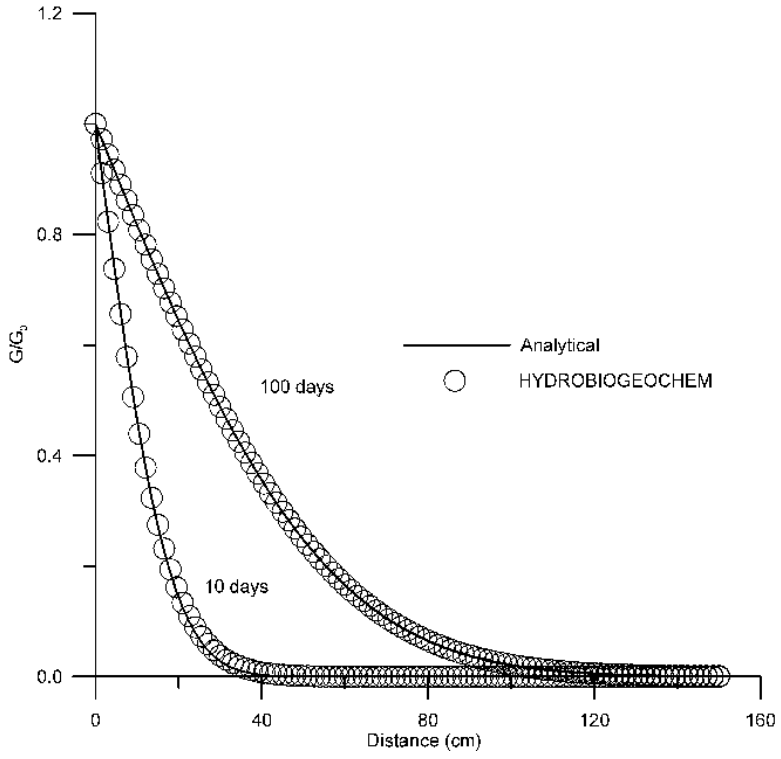


Figure 3.3 Gaseous species concentration at selected times

3.4.4 2-D Transport: Biodegradation of Hydrocarbon

The problem domain consists of two zones and is described in Figure 3.4. Zone 1 is unsaturated, zone 2 is saturated. The properties of these two zones are listed in Table 3.6. Parameters for this example are mainly from Borden and Bedient (1986). Groundwater flow was assumed to be uniform and horizontal. Boundary conditions are variable boundary condition on the upstream, influent concentrations given in table 3.6, and no flux boundary condition on the downstream. At the soil surface, partial pressure of oxygen is fixed (table 3.6). Initial species concentrations are given in table 3.6. Only diffusion is considered in the gas phase.

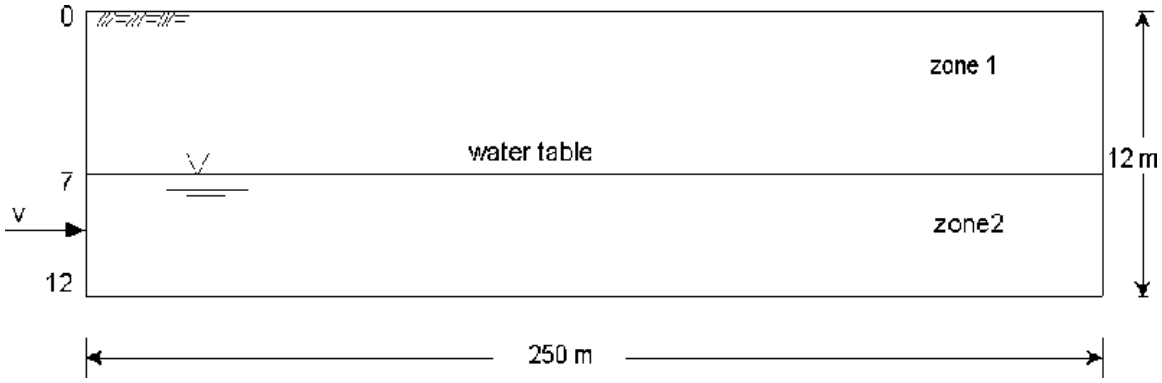


Figure 3.4 Problem domain

Monod kinetics for biodegradation is described by the following function

$$\frac{dM}{dt} = M\,k\,Y\left(\frac{H}{K_h + H}\right)\left(\frac{O}{K_o + O}\right) \tag{57}$$

and the microbial decay is represented by the first order decay rate

$$\frac{dM}{dt} = -\,b\,M_t \tag{58}$$

where

H hydrocarbon concentration;

O oxygen concentration;

M microbial concentration;

k maximum hydrocarbon utilization rate per unit mass microorganisms;

Y microbial yield coefficient;

K_h hydrocarbon half saturation constant;

K_o oxygen half saturation constant;

b microbial decay rate.

Table 3.6. Physical and Chemical Parameters and values for Problem 3.4.4.

Parameters	Value
Thickness of zone 1	7.0 m
Thickness of zone 2	5.0 m
Effective porosity	$\varphi = 0.262$
Bulk density	$\rho_b = 1.6$ g/L
water volumetric content in zone 1	$\theta_w = 0.125$
air volumetric content in zone 1	$\theta_a = 0.137$
Longitudinal dispersivity	$\alpha_L = 9.1$ m
Transverse dispersivity	$\alpha_T = 0.05$ m
Tortuosity	$\tau = 0.048$
Gaseous diffusion coefficient	$a_{ma} = 0.01$ m^2/d
Groundwater velocity (darcy velocity)	$v_w = 0.02$m/d
Partial pressure of oxygen at soil surface	21176 Pa

Initial hydrocarbon concentration	0.0 mg/L of bulk volume
Initial aqueous oxygen concentration	3.0 mg/L of bulk volume
Initial microbial biomas,	0.001 mg/L
Influent hydrocarbon concentration	4.5 mg/L
Influent oxygen concentration	3.0 mg/L
Maximum hydrocarbon utilization rate	1.7 day^{-1}
Microbial yield coefficient	0.58 mole cell/ mole hydrocarbon
Hydrocarbon half saturation constant	K_h = 0.13 mg/L of bulk volume
Oxygen half saturation constant	K_o = 0.10 mg/L of bulk volume
Microbial decay coefficient	b = 0.01 day
Ratio of oxygen to hydrocarbon consumed	F = 3.0
Henry's law coefficient for oxygen	2290 Pa L/mg
Henry's law coefficient for hydrocarbon	841 Pa L/mg
Adsorption constant for microbial biomass	K_d = 650

Considered in this problem are 6 species and 7 processes as shown in table 3.7.

Table 3.7. Processes considered in problem 3.4.4.

Process	#	Type
$O_{2(g)} \leftrightarrow O_{2(w)}$	(R1)	Equilibrium reaction described by Henry's Law
$H_{(g)} \leftrightarrow H_{(w)}$	(R2)	Equilibrium reaction described by Henry's Law

$M \leftrightarrow M_{(s)}$	(R3)	Equilibrium reaction described by linear isotherm
$H_{(w)} + F\ O_{2(w)} \leftrightarrow M$	(R4)	Kinetic reaction described by Monod kinetics
$H_{(w)} + F\ O_{2(w)} \leftrightarrow M_{(s)}$	(R5)	Kinetic reaction described by Monod kinetics
Decay of M	(R6)	1-st order rate decay
Decay of $M_{(s)}$	(R7)	1-st order rate decay

* F is ratio of oxygen to hydrocarbon consumed, H is hydrocarbon, M is microbe.

A constant time step of 1.0 day is used in the simulation. Simulation results after 750 days are shown in Figures 3.5 and 3.6. Figures 3.5 are concentration profiles for important aqueous species. Compared to the conservative tracer, hydrocarbon is removed from the system by biodegradation (reactions R4 and R5 in table 3.7). Close to the contaminant source (influent at the upstream boundary) we can see a nearly complete removal of oxygen even though there is supply of oxygen from the atmosphere (table 3.6). And this becomes a factor that limits the rate of biodegradation of contaminant. Figure 3.6 shows the concentration distribution of hydrocarbon in the saturated and unsaturated zones. From figure 3.6 we can see the accumulation of the gaseous contamination in the unsaturated zone above the aquifer because only diffusion is considered in gaseous phase. Once oxygen is depleted, the gaseous contaminant will dissolve into the groundwater and pose potential threat to the groundwater again. Therefore, for a better effect of bioremediation, a sufficient supply of electron acceptor, oxygen for this problem, must be provided in the system.

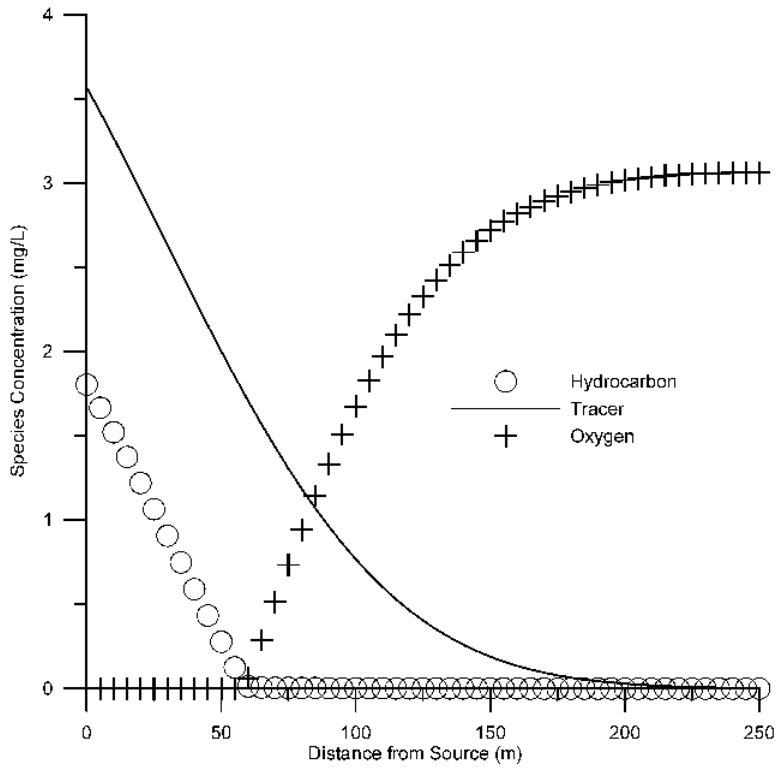


Figure 3.5. Concentration profile of different species in aqueous phase

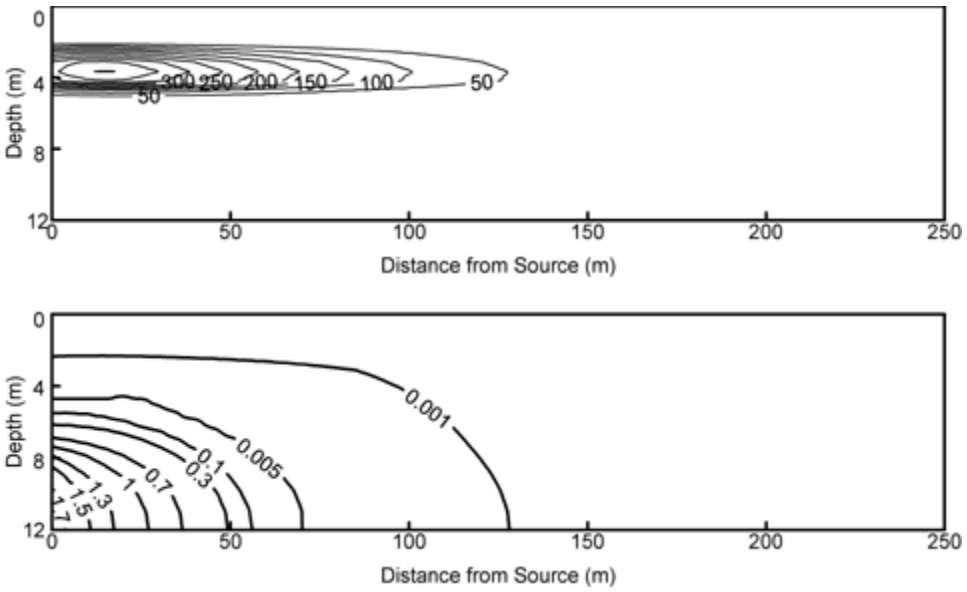


Figure 3.6. Contours of gaseous hydrocarbon (top, unit in Pa) and aqueous hydrocarbon (bottom, unit in mg/L) in the unsaturated and saturated zone.

3.5 DISCUSSION AND CONCLUSIONS

This paper presented a new formulation for multiphase reactive transport in the subsurface system. The primary dependent variables to be solved are no longer manually selected; rather they are systematically defined from the Gauss-Jordan decomposition procedure in this formulation. Any number of species and a complete set of mixed equilibrium and kinetic reactions can be easily handled in the system. The equilibrium expressions for fast reactions and rate expressions for slow reactions can be of virtually any form.

A computer model, HYDROBIOGEOCHEM, is developed on this formulation. Five different numerical schemes corresponding to advective and conservative forms of transport equations are compared and discussed.

There are advantages and drawbacks of different numerical schemes based on the conservative and advective form of the hydrologic transport governing equations. Conventional FEM with upstream weighting are susceptible to numerical dispersion, and limitations on the size of the Courant number. Lagrangian-Eulerian FEM in the Lagrangian step fails to conserve mass and have difficulty in treating boundary fluxes when characteristics intersect the boundary of the domain (Celia and Binning, 1992; Healy and Russell, 1993). The mass conservation errors result mainly from (1) the nonconservative form of the advection dispersion equation; (2) inaccurate tracking of the characteristics lines; (3) errors in the evaluation of transport properties at the feet of the characteristic lines; and (4) approximate treatment of boundary fluxes when characteristics intersect inflow or outflow boundaries (Healy and Russell, 1993; Oliveira and Baptista, 1998). The combination of FEM for boundary elements and Lagrangian-Eulerian FEM can overcome the drawback of Lagrangian-Eulerian FEM. It is shown that the combination method offers good result for dispersion dominant problems.

Both conservative and advective forms of governing equations are equivalent

mathematically. Numerically however, there are reasons to favor one form over the other because of the characteristics of different numerical schemes: (1) conservative form preserves conservative properties while advective form preserves transport properties in numerical implementations; (2) advective form facilitates Lagrangian-Eulerian implementation, which allows larger time steps.

APPENDIX 3.A MATRIX EQUATIONS FOR DIFFERENT NUMERICAL SCHEMES

- **Scheme 1: FEM on Advective Form of Equation**

The upstream finite element is employed to solve the hydrologic transport equation in advective form. The transport equation is discretized as follows:

$$[M]\left\{\frac{dT}{dt}\right\}+([L]+[V]+[K]+[Q])\{T\}=\{S\}+\{B\}+\{R\} \tag{A.1}$$

where [M] is the mass matrix; [L] is the stiff matrix resulting from dispersion/diffusion; [V] is the stiff matrix resulting from the velocity term; [K] is the stiff matrix resulting from the first order rate term; [Q] is the stiff matrix resulting from the source term; {S} is the load vector from external sources; {B} is the load vector from boundary sources; {R} is the load vector from the biogeochemical reactions; $\left\{\frac{dT}{dt}\right\}$ is the column vector containing the values of $\frac{dT}{dt}$; $\{T\}$ is the column vector containing the nodal values of *T*, and

$$M_{ij}=\sum_{e\in M_e}\int_{R_e}N_\alpha^e N_\beta^e dR \tag{A.2}$$

$$L_{ij}=\sum_{e\in M_e}\sum_{\alpha=1}^{3}\int_{R_e}\left(\nabla N_\alpha^e\right)\bullet\rho_\ell\theta_\ell\mathbf{D}_\ell\frac{T_\ell}{\rho_\ell\theta_\ell T}\bullet\left(\nabla N_\beta^e\right)dR \tag{A.3}$$

$$V_{ij}=\sum_{e\in M_e}\sum_{\ell=1}^{3}\int_{R_e}W_\alpha^e\left(\rho_\ell\mathbf{V}_\ell\frac{T_\ell}{\rho_\ell\theta_\ell T}-\rho_\ell\theta_\ell\mathbf{D}_\ell\bullet\nabla\frac{T_\ell}{\rho_\ell\theta_\ell T}\right)\nabla N_\beta^e dR \tag{A.4}$$

$$K_{ij}=\sum_{e\in M_e}\sum_{\ell=1}^{3}\int_{R_e}N_\alpha^e\left[\rho_\ell\mathbf{V}_\ell\bullet\nabla\frac{T_\ell}{\rho_\ell\theta_\ell T}-\frac{T_\ell}{\rho_\ell\theta_\ell T}\frac{\partial\rho_\ell\theta_\ell}{\partial t}-\nabla\bullet\left(\rho_\ell\theta_\ell\mathbf{D}_\ell\bullet\nabla\frac{T_\ell}{\rho_\ell\theta_\ell T}\right)\right]N_\beta^e dR \tag{A.5}$$

$$Q_{ij} = \sum_{e \in M_e} \sum_{\ell=1}^{3} \int_{R_e} N_\alpha^e \left(\rho_\ell^* Q_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) N_\beta^e dR \tag{A.6}$$

$$B_i = - \sum_{e \in N_{Se}} \sum_{\alpha=1}^{3} \int_{B_e} N_\alpha^e n \bullet \left(\rho_\ell \mathbf{V}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell} - \rho_\ell \theta_\ell D_\ell \frac{T_\ell}{\rho_\ell \theta_\ell} \right) dB \tag{A.7}$$

$$S_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left(\sum_{\ell=1}^{3} Q_\ell T_\ell^{in} \right) dR \tag{A.8}$$

$$R_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e RR \, dR \tag{A.9}$$

Where R_e = region of element e.

M_e = set of elements that have a local side α-β coinciding with the global side i-j.

N_α^e = α-th local base function of element e.

W_α^e = α-th local weighting function of element e.

B_e = the length of boundary segment e.

N_{Se} = set of boundary segments having a local node α coinciding with global node i.

- **Scheme 2: FEM on Conservative Form of Equation**

The upstream finite element is employed to solve the hydrologic transport equation in conservative form. The transport equation is discretized as follows:

$$[M]\left\{\frac{dT}{dt}\right\} + ([L]+[V])\{T\} = \{S\} + \{B\} + \{R\} \tag{A.10}$$

where [M] is the mass matrix; [L] is the stiff matrix resulting from dispersion/diffusion; [V] is the stiff matrix resulting from the velocity term; {S} is the load vector from external sources; {B} is the load vector from boundary sources; {R} is the load vector from the biogeochemical reactions, and

$$M_{ij} = \sum_{e \in M_e} \int_{R_e} N_\alpha^e N_\beta^e dR \tag{A.11}$$

$$L_{ij} = \sum_{e \in M_e} \sum_{\alpha=1}^{3} \int_{R_e} (\nabla N_\alpha^e) \bullet \rho_\ell \theta_\ell \mathbf{D}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \bullet (\nabla N_\beta^e) dR \tag{A.12}$$

$$V_{ij} = -\sum_{e \in M_e} \sum_{\ell=1}^{3} \int_{R_e} (\nabla W_\alpha^e) \bullet \left(\rho_\ell \mathbf{V}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} - \rho_\ell \theta_\ell \mathbf{D}_\ell \bullet \nabla \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) N_\beta^e dR \tag{A.13}$$

$$B_i = -\sum_{e \in N_{Se}} \sum_{\alpha=1}^{3} \int_{B_e} N_\alpha^e n \bullet \left(\rho_\ell \mathbf{V}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell} - \rho_\ell \theta_\ell D_\ell \frac{T_\ell}{\rho_\ell \theta_\ell} \right) dB \tag{A.14}$$

$$S_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left(\sum_{\ell=1}^{3} Q_\ell T_\ell^{in/out} \right) dR \tag{A.15}$$

$$R_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e RR \, dR \tag{A.16}$$

Where R_e = region of element e.

M_e = set of elements that have a local side α-β coinciding with the global side i-j.

N_α^e = α-th local base function of element e.

W_α^e = α-th local weighting function of element e.

B_e = the length of boundary segment e.

N_{Se} = set of boundary segments having a local node α coinciding with global node i.

- **Scheme 3: Hybrid Lagrangian-Eulerian FEM for Interior Elements + FEM on Advective Form of Equation for Boundary elements**

The transport equation for interior elements is discretized as follows:

$$[M]\left\{\frac{DT}{D\tau}\right\} + ([L]+[K]+[Q])\{T\} = \{S\} + \{R\} \tag{A.17}$$

where [M], [L], [K], [Q], {S}, and {R} are the defined the same as those in Eq. (A.1). $\left\{\frac{DT}{D\tau}\right\}$ is a column vector containing the values of $\frac{DT}{D\tau}$; τ is the backward tracking time.

The transport equation for boundary elements is the same as Eq. (A.1).

- **Scheme 4: Hybrid Lagrangian-Eulerian FEM**

To overcome the inaccuracy caused by tracking time at the boundary node, we'll first approximate the time derivative with finite difference method and then multiply the governing equation with the time interval. With similar procedure of deriving governing equations in conventional FEM, applying Galerkin FEM to the resulted equation, following linearized equation is obtained:

$$[M]\{\Delta T\} + ([L'] + [K'] + [Q'])\{T\} = \{S'\} + \{B'\} + \{R'\} \tag{A.18}$$

where [M] is the mass matrix, $\{\Delta T\}$ is a vector composed of the difference of values of T between the new and old time steps. [L′] is the stiff matrix resulting from dispersion, [K′] is the stiff matrix resulting from the first order rate term, [Q′] is the stiff matrix resulting from the source term, {S′} is the load vector from external sources, {B′} is the load vector from boundary sources, {R′}is the load vector from the biogeochemical reactions, and

$$M_{ij} = \sum_{e \in M_e} \int_{R_e} N_\alpha^e N_\beta^e dR \tag{A.19}$$

$$K'_{ij} = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left\{ \sum_{\ell=1}^{3} \left[\rho_\ell \mathbf{V}_\ell \bullet \nabla \frac{T_\ell}{\rho_\ell \theta_\ell T} - \frac{T_\ell}{\rho_\ell \theta_\ell T} \frac{\partial \rho_\ell \theta_\ell}{\partial t} - \nabla \left(\rho_\ell \theta_\ell \mathbf{D}_\ell \bullet \nabla \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) \right] \right\} \Delta\tau \, N_\beta^e dR \tag{A.20}$$

$$L'_{ij} = \sum_{e \in M_e} \int_{R_e} \left(\nabla N_\alpha^e\right) \cdot \left(\sum_{\ell=1}^{3} \rho_\ell \theta_\ell \mathbf{D}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) \Delta\tau \cdot \left(\nabla N_\beta^e\right) dR \tag{A.21}$$

$$Q'_{ij} = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left(\sum_{\ell=1}^{3} \rho_\ell^* Q_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) \Delta\tau \, N_\beta^e dR \tag{A.22}$$

$$B'_i = \sum_{e \in N_{Se}} \int_{B_e} N_\alpha^e \mathbf{n} \bullet \left(\sum_{\ell=1}^{3} \rho_\ell \theta_\ell \mathbf{D}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \bullet \nabla T \right) \Delta\tau \, dB \tag{A.23}$$

$$S'_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left(\sum_{\ell=1}^{3} Q_\ell T_\ell^{in} \right) \Delta\tau \, dR \tag{A.24}$$

$$R'_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e RR \, \Delta\tau \, dR \tag{A.25}$$

Where R_e = region of element e.

M_e = set of elements that have a local side α-β coinciding with the global side i-j.

N_α^e = α-th local base function of element e.

B_e = the length of boundary segment e.

N_{Se} = set of boundary segments having a local node α coinciding with global node i.

- **Scheme 5: Hybrid Lagrangian-Eulerian FEM for Interior Elements + FEM on Conservative Form of Equation for Boundary elements**

In this scheme, the transport equation for interior elements is discretized as Eq. (A.17). The transport equation for boundary elements is the same as Eq. (A.10).

- **Scheme 6: Traditional Hybrid Lagrangian-Eulerian FEM**

In this scheme, the transport equation for interior elements is discretized as

follows:

$$[M]\left\{\frac{DT}{D\tau}\right\}+([L]+[K]+[Q])\{T\}=\{S\}+\{R\}+\{B\} \tag{A.26}$$

where [M], [L], [K], [Q], {S}, {R}, and {B} are the same as those in Eq. (A.1).

ACKNOWLEDGEMENTS. This research was supported, in part, by National Science Foundation (NSF) Grant No.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 No.DE-FG02-98ER62691 with the Pennsylvania State University.

REFERENCES

Abrams, R. H. and Loague K., 1998. Development and testing of a compartmentalized reaction network model fro redox oznes in contaminated aquifers.

Baehr, A. L., 1987, Selective transport of hydrocarbons in the unsaturated zone due to aqueous and vapor phase partitioning, water resources research, v. 23(10), p. 1926-1938.

Borden, R.C., and P.B. Bedient, 1987, Transport of dissolved hydrocarbons influenced by oxygen-limited biodegradation 1. theoretical development, water resources research, v. 22(13), p. 1973-1982.

Bosma, W. and van der Zee, 1993.Transport of reacting solute in a one-dimensional, chemically heterogeneous porous medium. Water Resources Research, 29(1): 117-131.

Brun, A. and Engesgaard, P., 2002. Modelling of transport and biogeochemical processes in pollution plumes: literature review and model development. Journal of Hydrology 256:211-227

Brusseau, M., 1994. Transport of reactive contaminants in heterogeneous porous media. Reviews of Geophysics, 32(3): 285-313.

Cederberg, G. A., Street, R. L., and Leckie, J. O., 1985. A Groundwater Mass Transport and Equilibrium Chemistry Model for Multicomponent Systems. Water Resources Research, 21: 1095-1104.

Celia, M.A. and P. Binning, 1992. A mass conservative numerical solution for two-phase flow in porous media with applications to unsaturated flow. Water Resources Research, Vol. 28 (10), 2819-2828.

Chang, J. R., G. T. Yeh, and T. E. Short, 1993. Modeling two-dimensional subsurface flow, fate and transport of microbes and chemicals. Proc. Int. Symp. on Eng. Hydrology, San Francisco, July 25-30, 1993, p.611-616.

Chen, Y., L. M. Abriola, P. J. J. Alvarez, P. J. Anid and T. M. Vogel, 1992. Modeling

transport and biodegradation of benzene and toluene in sandy aquifer material: comparisons with experimental measurements, *Water Resources Researc*h, 28(7): 1833-1847, July 1992.

Cheng, J. R. and Yeh, G. T. , 1994. Modeling three-dimensional subsurface flow, fate and transport of microbes and chemicals (3DFATMIC). Proc. X-th Int. Conf. on Numerical Methods in Water Resources, Heidelberg, Germany, July 19-22, 1994, 217-224.

Cheng, H.-P., 1995. Development and Application of a Three-Dimensional Finite Element Model of Subsurface Flow, Heat Transfer, and Reactive Chemical Transport. Ph.D. Thesis, Department of Civil and Environmental Engineering, The Pennsylvania State University. University Park, PA.

Chilakapati, A., Ginn, T. and Szecsody, J., 1998. An Analysis of Complex Reaction Networks in Groundwater Modeling. Water Resources Research, 34(7): 1767-1780.

Davis, J., Kent D., Coston, J., Hess, K. and Joye, J., 2000. Multispecies reactive tracer test in an aquifer with spatially variable chemical conditions. Water Resources Research, 36(1): 119-134.

Fang, Y. and G. T. Yeh, 2002. Numerical modeling of reactive chemical transport under multiphase systems. Proc. XIV International Conference on Computational Methods in Water Resources. Delft, The Netherlands, June 23-28, 2002.

Fang, Y., Yeh, G.T., and Burgos, W.T., 2003. A generic paradigm to model reaction-based biogeochemical processes. Water Resources Research, 39(4) (in press).

Friedly, J. and Rubin, J., 1992. Solute Transport with multiple equilibrium-controlled or kinetically controlled chemical reactions. Water Resources Research, 28(6):1935-1953.

Gerke, H.H., Molson, J.W., and Frind, E.O., 1998. Modelling the effect of chemical heterogeneity on acidification and solute leaching in overburden mine spoils. J.

Hydrol. 209: 166-185.

Healy, R.W. and Russel, T.F., 1993. A finite-volume Eulerian-Lagrangian localized adjoint method for solution of the advection-dispersion equation, Water Resources Research, Vol. 29(7), 2399-2413.

Hostetler, J. C. and Erickson, R. L. , 1989. FASTCHEM Package 5. Report EA-5870-CCM, Electric Power Research Institute.

Huyakorn, P.S., Mercer J.W., and Ward, D.S, 1985. Finite element matrix and mass balance computational schemes for transport in variably saturated porous media. Water Resources Research, 21(3), 346-358.

Lapidus, L., and Amundson, N.R., 1952, Mathematics of adsorption in beds. VI. The effect of longitudinal diffusion in ion-exchange and chromatographic columns: Journal of Physical Chemistry, v. 56, p. 984-988.

Lensing, H. J., Voyt, M., and Herrling, B., 1994. Modeling of Biologically Mediated Redox Processes in the Subsurface. Journal of Hydrology, 159: 125-143.

Lichtner, P. C., 1996. Continuum Formulation of Multicomponent-Multiphase Reactive Transport. In: P. C. Lichtner, C. I. Steefel, and E. H. Oelkers (Editors), Reactive Transport in Porous Media, Reviews in Mineralogy, Volume 34. Mineralogical Society of America, Washington, D.C., pp. 1-79.

Lichtner, P.C., Seth M., 1996. Multiphase-multicomponent nonisothermal reactive transport in partially saturated porous media. In: Proceedings of the International Conference on Deep Geological Disposal of Radioactive Waste, Canadian Nuclear Society, September 16-19, Winnnipeg, Manitoba, Canada, 3-133-42pp.

Lindstrom, F.T., Haque, R., Freed, V.H. and Boersma, L., 1967, Theory on the movement of some herbicides in soils--Linear diffusion and convection of chemicals: Environmental Science and Technology, v. 1, p. 561-565.

Liu, C. and Narasimhan, T., 1989a. Redox-controlled multiple-species reactive chemical transport 1. model development. Water Resources Research, 25(5): 869-882.

Liu, C. and Narasimhan, T., 1989b. Redox-controlled multiple-species reactive chemical transport 2. Verification and application. Water Resources Research, 25(5): 869-882.

MacQuarrie, K.T.B., Sudicky, E.A., Frind, E.O., 1990. Simulation of biodegradable organic contaminants in groundwater: 1. Numerical formulation in principal directions. Water Resour. Res. 26, 207-222.

McNab, J., W.W. and Narasimhan, T.N., 1994. Modeling reactibe transport of organic compounds in groundwater using a partial redox disequilibrium approach. Water Resources Research, 30(9): 2619-2635.

McNab, J., W. W. and Narasimhan, T.N., 1995. Reactive transport of petroleum hydrocarbon constituents in a shallow aquifer: modeling geochemical interactions between organic and inorganic species. Water Resources Research, 31(8): 2027-2033.

Miller, C.W.,and Benson, L.V., 1983. Simulation of solute transport in a chemical reactive heterogeneous system: Model development and application, Water Resources Research, 19(2):381-391.

Millington, R.J., Quirk, J.M., 1961. Permeability of porous solids, Trans. Faraday Soc., 7, 1200-1207.

Narasimhan, T. N., White, A.F., and Tokunaga, T., 1986. Groundwater contamination from an inactive uranium mill tailings pile: 2. Application of a dynamic mixing model. Water Resources Research, 22(13):1820-1834.

Nguyen, V.V., Gray, W.G., Pinder, G.F., Botha, J.F., and Gerar, D.A., 1982. A theoretical investigation on the transport of chemicals in reactive porous media. Water Resour. Res. 18, 1149-1156.

Oliveira, A. and A. M. Baptista, 1998. On the role of tracking on Eulerian-Lagrangian solutions of the transport equation. Advances in Water Resources, Vol. 21, pp. 539-554.

Owczarek, J.A. *Fundamentals of Gas Dynamics*. International Textbook Company, Scranton, PA., 668pp, 1964.

Parkhurst, D. L., 1995. User's Guide to PHREEQC, A Computer Model for Speciation, Reaction-Path, Advective-Dispersive Transport and Inverse Geochemical Calculations. U.S. Geological Survey, Water Resources Investigations Report 95-4227.

Parkhurst, D.L. and Appelo, C.A.J., 1999. User's guide to PHREEQC (version 2)--A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations: U.S.Geological Survey Water-Resources Investigations Report 99-4259, 312 p.

Saiers, J.E., Guha, H., Jardine, P., and Brooks, S., 2000. Development and evaluation of a mathematical model for the transport and oxidation-reduction of CoEDTA. Water Resources Research, 31(11):3151-3165.

Salvage, K.M., Yeh, G.T., Cheng, H.P. and Cheng, J.R., 1996. Development of a model of subsurface hydrologic transport and biogeochemical reactions (HYDROBIOGEOCHEM), Computational Methods in Water Resources XI. Computation Methods in Surface Flow and Transport Problesm, pp. 517-524.

Salvage, K.M., 1998. Reactive contaminant tranpsport in variably saturated porous media: biogeochemical model development, verification, and application. . Ph.D. Thesis, Department of Civil and Environmental Engineering, The Pennsylvania State University. University Park, PA.

Steefel, C.I., Lasaga, A.C., 1994. 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.

Steefel, C.I., Yabusaki, S.B., 1996. OS3D/GIMRT, Software for Modeling Multi-Component-Multidimensional Reactive Transport, User's Manual and Programmer's Guide. PNL-11166, Pacific Northwest Laboratory, Richland, WA.

Suarez, D. and Šimunek, J., 1996. Solute Transport Modeling Under Variably Saturated Water Flow Conditions. In: P. C. Lichtner, C. I. Steefel, and E. H. Oelkers (Editors), Reactive Transport in Porous Media, Reviews in Mineralogy, Volume 34. Mineralogical Society of America, Washington, D.C., pp. 229-268.

Szecsody, J.E., Zachara, J.M., Chilakapati, A., Jardine, P.M. and Ferrency, A.S., 1998. Importance of flow and particle-scale heterogeneity on Co(II/III)EDTA reactive transport. Journal of Hydrology, 209(1): 112-136.

Tebes-Stevens, C., Valocchi, A.J., VanBriesen, J.M. and Rittmann, B.E., 1998. Multicomponent transport with coupled geochemical and microbiological reactions: model description and example simulations. Journal of Hydrology, 209: 8-26.

Theis, T. L., Kirkner, D.J., and Jennings, A.A., 1982. Multi-Solute Subsurface Transport Modeling for Energy Solid Wastes. Tech. Progress Rept. C00-10253-3, Department of Civil Engineering, University of Notre Dame, Norte Dame, IN.

Tompson, A.F.B., 1993. Numerical simulation of chemical migration in physically and chemically heterogeneous porous media. Water Resources Research, 29(11): 3709-3726.

Toride, N., Leij, F.J. and Genuchten, M.T.v., 1993. A comprehensive set of analytical solutions for nonequilibrium solute transport with first-order decay and zero-order production. Water Resources Research, 29: 2167-2182.

van Genuchten M. Th. and W. J. Alves, 1982. Analytical solutions of the one-dimensional convective-dispersive solute transport equation. USDA ARS Technical Bulletin Number 1661. U.S. Salinity Laboratory, 4500 Glenwood Drive, Riverside, CA 92501

Viwanathan, H.S., Robinson, B.A., Valocchi, A.J., and Triay, I.R., 1998. A reactive transport model of neptunium migration from the potential repository at Yucca Mountain. J. Hydrol, Vol. 209, 251-280.

Wood, B. D., Dawson, C.N., Szecsody, J.E. and Streile, G.P., 1994. Modeling contaminant transport and biodegradation in a layered porous media system, *Water Resources Research* , 30(6): 1833-1845.

Xu, T. and Pruess, K., 2001. Modeling multiphase non-isothermal fluid flow and reactive transport in variably saturated fractured rocks: 1. methodology. American journal of science, Vol. 301(1), 16-33.

Yeh, G.T., Tripathi, V.S., 1991. A model for simulating transport of reactive multispecies components: Model development and demonstration. Water Resources Research, 27(12): 3075-3094.

Yeh, G. T., Salvage, K., and Choi, W.H., 1996. Reactive Multispecies-Multicomponent Chemical Transport Controlled by both Equilibrium and Kinetic Reactions. Proc. XI-th Int. Conf. on Numerical Methods in Water Resources, Cancun, Mexico, July 22-26, 1996. p. 585-592

Yeh, G.T., Burgos W.D., Fang, Y.L. and Zachara J.M., 2000. Modeling biogeochemical kinetics: issues and data needs. XIII International Conference on Computational Methods in Water resources, 2000.

Yeh, G. T., Siegel, M.D., and Li, M.H., 2001a. Numerical Modeling of Coupled Fluid Flows and Reactive Transport Including Fast and Slow Chemical Reactions. J. Contaminant Hydrology. 47:379-390.

Yeh, G.T., Burgos W.D., and Zachara J.M., 2001b. Modeling and measuring biogeochemical reactions: System consistency, data needs, and rate formulation. *Adv. Environ. Res*., 5, 219-237, 2001.

van der Zee, S.E.A.T.M. and van Riemsdijk, W.H., 1987. Transport of reactive solute in spatially variable soil systems. Water Resources Research, 23(11): 2059-2069

Zhu, C., Hu, F.Q., and David, S. B., 2001. Multi-component reactive transport modeling of natural attenuation of an acid groundwater plume at a uranium mill tailings site. Journal of Contaminant Hydrology, Vol. 52, 85-108.

CHAPTER 4.
THREE DIMENSIONAL NON-ISOTHERMAL REACTIVE TRANSPORT UNDER MULTIPHASE SYSTEM

Yilin Fang[1] , Gour-Tsyh Yeh[2*] and William D. Burgos[1]

[1]Department of Civil and Environmental Engineering
The Pennsylvania State University
University Park, PA 16802

[2]Department of Civil and Environmental Engineering
University of Central Florida
Orlando, FL 32816-2450

ABSTRACT

In this paper, a three dimensional numerical model, 3DHYDROBIOGEOCHEM, is presented for modeling non-isothermal reactive chemical transport under multiphase systems. The model is a coupled hydrologic transport and biogeochemical reaction code, which is used to predict the spatio-temporal distributions of all the important chemical species. To reduce primary dependent variables (PDVs), Gauss-Jordan decomposition is applied to the governing matrix equations for transport, resulting in mobile components and mobile kinetic variables as PDVs. Methods for coupling reaction and transport, such as sequential iteration approach (SIA) and operator splitting approach are incorporated into the code to make the model versatile. Weak coupling will be applied to solve a system of chemical transport and heat transfer to save computing resources. Based on the conservative form and advective form of the hydrologic transport equations and heat transfer equations, five different numerical schemes are incorporated in the model to solve the governing transport equations. The model developed is a practical tool for

assessing migration of subsurface contamination and proper designing of remediation technologies. Examples are presented to demonstrate the application of the new model.

Keywords: Reactive Chemical Transport; Numerical Model; Biogeochemistry; Multiphase System; Heat Transfer.

4.1 INTRODUCTION

Transport of chemical species with a variety of biogeochemical processes is mathematically described by a set of linear partial differential equations (PDEs) and/or nonlinear algebraic equations (AEs). A variety of approaches solving the system composed of PDEs and AEs for transport have been documented in recent literature (see Yeh and Tripathi, 1989; Liu and Narasimhan, 1989; Steefel and MacQuarrie, 1996; and references therein). Researchers have also made great effort and progress at mathematical formulations for the physical and chemical processes included in the reactive transport model. Except for those transport models summarized by Brun and Engesgaard (2002), there are models simulating reactive transport in multiphase system (Viswanathan et al., 1998; Wu and Pruess, 2000; Xu and Pruess, 2001) as well. The conclusion from reviewing the present transport models is that they are becoming more and more complex by considering the important physical and chemical processes simultaneously due to the requirement of a comprehensive understanding of natural systems.

Although research in reaction-based reactive transport model in multiphase system is relatively new, it has received increasing attention because of its significance in natural resource recovery, waste storage and environmental remediation (Wu and Pruess, 2000). However, the physical and chemical complexities of fracture-matrix systems underground, such as the inherent heterogeneity and the variety of chemical reactions, have given rise to the difficulties in mathematical formation and solving of a reactive transport in multiphase system.

The main objective of the present study is to present a non-isothermal, reactive transport formulation in multiphase systems and the numerical implementations. The new

numerical model is represented by a hydrologic transport code coupling with a reaction speciation code BIOGEOCHEM (Fang et al., 2003), which is based on the new paradigm proposed by Yeh et al. (2001a).

4.2 GOVERNING EQUATIONS FOR HYDROLOGIC TRANSPOT

According to custom definition, species in each fluid phase is considered to be mobile and every solid or solid associated species is considered to be immobile. In the present model, each mobile species (including aqueous-phase species, NAPL-phase species, and air-phase species) concentration has the unit of moles per mass of fluid, while immobile species has the unit of moles per unit mass of rock matrix, excluding precipitated species concentration, which has a unit of moles per bulk volume. With the concentration unit defined, governing equation for transport of mobile chemical species can be derived from the principle of mass balance as

$$\frac{D}{Dt}\int_{\nu}\rho_{\alpha}\theta_{\alpha}c_i^{\alpha}d\nu = -\int_{\Gamma} n\bullet(\rho_{\alpha}\theta_{\alpha}c_i^{\alpha}\mathbf{V}_{fs,\alpha})d\Gamma - \int_{\Gamma}\mathbf{n}\bullet\mathbf{J}_i^{\alpha}d\Gamma + \int_{\nu} r_i d\nu + \int_{\nu} m_i^{\alpha}d\nu \qquad (1)$$

where ν is the material volume containing constant amount of media (L^3); Γ is the surface enclosing the material volumeν (L^2); **n** is the outward unit vector normal to the surface Γ; ρ_{α} is the fluid density of phase α, α denoting water, gas or oil, (M/L^3); θ_{α} is the volume fraction of fluid α within the representative elementary volume (REV), (L^3/L^3), $\theta_{\alpha}=\phi S_{\alpha}$; ϕ is the effective porosity; S_{α} is the fluid saturation; c_i^{α} is the concentration of the i-th chemical species in fluid phase α (M/M of fluid); $\mathbf{J}_i^{\alpha}$ is the surface flux of the i-th chemical species with respect to the velocity of fluid α [(M/T)/L^2];

and $\mathbf{J}_i^{\alpha} = \rho_{\alpha}\theta_{\alpha} c_i^{\alpha} (\mathbf{V}_{i,\alpha} - \mathbf{V}_{f,\alpha})$; $\mathbf{V}_{i,\alpha}$ is the local velocity of the i-th chemical species in phase α (L/T); $\mathbf{V}_{f,\alpha}$ is the local fluid velocity of phase α (L/T); $\mathbf{V}_{f,\alpha} = \mathbf{V}_{fs,\alpha} + \mathbf{V}_s$; $\mathbf{V}_{fs,\alpha}$ is the fluid velocity of phase α relative to the solid surface (L/T); $\mathbf{V}_s$ is the solid velocity (L/T); m_i^{α} is the external source/sink rate of the i-th species in phase α per unit bulk volume [(M/L^3)/T]; r_i^{α} is the production/consumption rate of the i-th species in phase α due to all chemical reactions (decay of the i-th species is also considered as a reaction)[(M/ L^3 bulk volume)/T] and it can be expressed by

$$r_i = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k \tag{2}$$

where ν_{ik} is the stoichiometric coefficient of the i-th species in the k-th reaction as a product; μ_{ik} is the stoichiometric coefficient of the i-th species in the k-th reaction as a reactant.

By the Reynolds transport theorem (Owczarek, 1964) and Gaussian divergence theorem, neglecting the deformation of the porous media, the following equation is obtained:

$$\frac{\partial \rho_{\alpha}\theta_{\alpha} c_i^{\alpha}}{\partial t} + \nabla \bullet (\rho_{\alpha} c_i^{\alpha} \mathbf{V}_{\alpha}) + \nabla \bullet \mathbf{J}_i^{\alpha} = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k + m_i^{\alpha} \tag{3}$$

The surface flux can be expressed in terms of the gradient of c_i^{α} (Nguyen et al., 1982):

$$\mathbf{J}_i^{\alpha} = -\rho_{\alpha}\theta_{\alpha} \mathbf{D}_{\alpha} \bullet \nabla c_i^{\alpha} \tag{4}$$

and the dispersion tensor $\mathbf{D}_{\alpha}$ of phase α, according to Millington and Quirk (1961) (from Gerke et al., 1998), can be given in the standard form as

$$\theta_{\alpha} \mathbf{D}_{\alpha} = a_T |\mathbf{V}_{\alpha}| \delta + (a_L - a_T) \mathbf{V}_{\alpha} \mathbf{V}_{\alpha} / |\mathbf{V}_{\alpha}| + a_{m\alpha} \theta_{\alpha} \tau_{\alpha} \delta \tag{5}$$

where D_{α} is the dispersion coefficient tensor of phase i, (L^2/T); a_T is the transverse

disffusivity (L); a_L is the longitudinal diffusion coefficient (L^2/T); $a_{m\alpha}$ is the molecular diffusion coefficient in pure phase α (L^2/T); τ_α is the tortuosity factor of the medium for fluid phase α; δ is the Kronecker delta tensor; and $\mathbf{V}_\alpha$ is the Darcy velocity of phase α (L/T), it is related with fluid velocity by $\mathbf{V}_\alpha = \mathbf{V}_{f,\alpha}\theta_\alpha$.

$$\frac{\partial \rho_\alpha \theta_\alpha c_i^\alpha}{\partial t} + \nabla \bullet (\rho_\alpha c_i^\alpha \mathbf{V}_\alpha) - \nabla \bullet (\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla c_i^\alpha) = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k + m_i^\alpha \tag{6}$$

Immobile species are not subject to hydrological transport, the mass balance equation for these species can be obtained from Eq. (6) by dropping advection, dispersion/diffusion, and external source/sink terms and replacing $(\rho_\alpha \theta_\alpha c_i^\alpha)$ with $(\rho' c_i)$:

$$\frac{\partial \rho' c_i}{\partial t} = \sum_{k=1}^{N} (\nu_{ik} - \mu_{ik}) R_k \tag{7}$$

where $\rho' = 1.0$ for precipitated species and ρ' = the bulk density for other immobile species.

4.2.1 Governing Equations for Primary Dependent Variables (PDVs)

Decompose the matrix equation made up of the transport equations, the following set of governing equations are obtained:

· *Transport Equation for Mobile Components* –

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right] = M_j, \quad j \in N_m \tag{8}$$

· *Transport Equation for Mobile Kinetic Variables* –

$$\frac{\partial E_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right] = M_j + D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j, \quad k \in N_{KI}, \quad j \in M_m \tag{9}$$

· *Infinite-Rate Equations for N_E Equilibrium Reactions* -

$$\frac{dE_j}{dt} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{E_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right]$$

$$= M_j + D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j, \; k \in N_E, : \frac{d}{dt}\left(\frac{E_j}{D_{kk}} \right) \approx R_k = \infty \qquad (10)$$

$$\exists \text{ a thermodynamically consistent equation}$$

where T_j is the total concentration of the j-th component, T_j^α is the concentration of the j-th component j in phase α; E_j is the total concentration of the j-th kinetic-variable; E_j^α is the concentration of the j-th kinetic variable in fluid phase α; N_m is the number of mobile components, M_m is the number of mobile kinetic variables; and M_{im} is the number of immobile kinetic variables. The number of immobile components is $N_{im} = N_C - N_m$, where N_C is the total number of components.

Governing equations for immobile components and kinetic variables are in the same form as Eqs. (8) and (9) by dropping advection, dispersion/diffusion, and external source/sink terms in Eqs. (8) and (9).

For the convenience of presentation, Eqs. (8) and (9) will be expressed by the following equation in a general form:

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right] = M_j + RR_j \qquad (11)$$

where $RR_j = 0$ for components, and $RR_j = D_{kk} R_k + \sum_{j \in N_{KD(k)}} D_{ij} R_j$ for kinetic variables.

To solve Eq. (11) numerically, T_j^α is expressed in terms of T_j to make T_js as

PDVs and the following equation is obtained:

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \nabla \bullet \left(\frac{T_j^{\alpha}}{\rho_\alpha \theta_\alpha T_j} \rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} T_j \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{T_j^{\alpha}}{\rho_\alpha \theta_\alpha T_j} T_j \right) \right) = M_j + RR_j \quad (12)$$

Differential Eq. (12) is the conservative or primitive form of the governing equation for hydrologic transport. The equation is derived from the law of mass conservation. It can also be converted to the mathematically equivalent advective form by expanding the second term of Eq. (12). The derivation is given in Appendix 4.A.

Due to the effect of reactions, true transport velocity (appendix 4.A) is used in the formulation to reach a right solution (Yeh, 2001b).

4.2.2 Initial and Boundary Conditions

To solve the advection dispersion equation (ADE), initial and boundary conditions should be known. Initially, the concentration of each chemical species must be given throughout the region of interest.

For all species that are subject to hydrological transport, boundary conditions are specified. Types of boundary conditions incorporated are: Dirichlet, Neumann, Cauchy, or variable. When applying FEM to the advective form of the governing equations, boundary conditions should be reformulated to be consistent with the corresponding ADEs. Initial and boundary conditions for hydrologic transport are presented in appendix 4.B.

4.3 NUMERICAL IMPLEMENTATION

The ADEs along with the initial and boundary conditions are solved with finite element method. Five schemes are provided in the model: (1) FEM on advective form of equation; (2) FEM on conservative form of equation; (3) Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on advective form of equation for boundary elements; (4) hybrid Lagrangian-Eulerian FEM; and (5) Hybrid Lagrangian-Eulerian FEM for interior elements + FEM on conservative form of equation for boundary elements (Ref. Chapter 3 in this thesis). FEM discretizations for the different schemes are presented in appendix 4.C.

The reaction terms in the transport equations presented above result in a set of non-linear partial differential equations. These equations are computed by the fully implicit sequential iteration approach (SIA), which solves the transport and reaction subsystems of equations sequentially and iteratively, the operator splitting approach, and predictor-corrector approach, which solve the two subsystems of equations sequentially with no iterate between them. In all three approaches, only the mobile PDVs are solved in the transport module. The immobile PDVs are solved in the chemistry module. For the fully implicit approach, the convergent rate would be very slow; while with the latter approaches, accuracy would be sacrificed (Salvage, 1998).

Without losing generality, the transport equations can be expressed in the following form:

$$\frac{\partial T}{\partial t} = L(T) + R(T) \tag{13}$$

where L( ) is an advective-dispersive transport operator and R( ) is a reaction operator. For the fully implicit sequential iteration approach, Eq. (13) becomes:

$$\frac{T^{n+1,r+1} - T^{n}}{\Delta t} = L(T^{n+1,r+1}) + R(T^{n+1,r+1}) \tag{14}$$

and the reaction term will be linearized by a Taylor series approximation:

$$R(T^{n+1,r+1}) = R(T^{n+1,r}) + \sum_{j=1}^{N_m+M_m} \left(\frac{\partial R}{\partial T_j}\right)^{n+1,r} \left(T_j^{n+1,r+1} - T_j^{n+1,r}\right) \tag{15}$$

Where n = time step level, and r = iteration step.

In the fully implicit approach, the concentrations of all immobile PDVs are computed in the biogeochemical subsystem using all species concentrations from the prior iteration. The concentrations of the mobile PDVs are not solved in the reaction subsystem. They are determined in the hydrologic module using the species concentrations and the reaction term evaluated from species concentrations of the current iteration. These concentrations then passed to the reaction module as given constants.

In the operator splitting approach, the concentrations of mobile PDVs are solved for in both subsystems of equations. The mobile PDVs are first subjected to transport without considering the reaction term to yield an intermediate value of their concentrations. These transported concentrations are then subjected to the biogeochemical reactions, yielding final values of the concentrations for the current time step. In this approach, Eq. (13) is solved in two steps:

$$\frac{T^{n+1/2} - T^{n}}{\Delta t} = L(T^{n+1/2}) \tag{16}$$

$$\frac{T^{n+1} - T^{n+1/2}}{\Delta t} = R(T^{n+1}) \tag{17}$$

where the superscripts (n), (n+1/2), and (n+1) denote the values at the prior time, at an intermediate point in the current computation, and at the new time, respectively.

In the predictor-corrector approach the concentrations of mobile PDVs are also

solved for in both subsystems of equations. The only difference of this approach from the operator splitting approach lies in whether reaction terms be considered in the transport equation. In this approach, Eq.(13) is solved in the following two steps:

$$\frac{T^{n+1/2}-T^{n}}{\Delta t}=L(T^{n+1/2})+R(T^{n}) \tag{18}$$

$$\frac{T^{n+1}-T^{n+1/2}}{\Delta t}=R(T^{n+1})-R(T^{n}) \tag{19}$$

In the operator splitting approach and the predictor-corrector approach, the coefficients in the L( ) operator are determined from previous iterate. No iteration between the transport subsystem and the reaction subsystem is implemented. However, when the PDVs have immobile components, these coefficients in the L( ) operator are dependent on the PDVs to be simulated. This implementation gives less accurate solution.

4.4 GOVERNING EQUATIONS FOR HEAT TRANSFER

Assuming thermal equilibrium so that temperature in the water, air, NAPL and solid phases are the same locally at any REV, the governing equation for heat transfer can be modified from Yeh and Luxmoore (1983) as

$$\left(\sum_{\alpha=1}^{3}C_{\alpha}\rho_{\alpha}+C_{m}\rho_{b}\right)\frac{\partial T^{h}}{\partial t}+\nabla\bullet\left[\left(\sum_{\alpha=1}^{3}C_{\alpha}\rho_{\alpha}\mathbf{V}_{\alpha}\right)T^{h}\right]+\left(\sum_{\alpha=1}^{3}C_{\alpha}\frac{\partial\rho_{\alpha}\theta}{\partial t}\right)T^{h} \tag{20}$$
$$-\nabla\bullet(\lambda\mathbf{I}\bullet\nabla T^{h})=SSH$$

where α stands for water, air, and NAPL. ρ_{α} is density of fluid phase α; θ_{α} is volume fraction of fluid phase α; C_{α} is heat capacity of fluid phase α; ρ_b is bulk density of dry medium; C_m is heat capacity of dry medium; T^h is temperature; I is identity tensor; λ is

apparent thermal conductivity; and SSH is source/sink of heat that may arise from artificial injection and withdrawal as well as chemical reactions

For the case of injection, with the given injection rate Q, the volume fraction of each phase in the injecting fluid, and the temperature of the injecting fluid, the heat injected into the system can be computed. Similarly, for the case of withdrawal, we specify the pumping rate Q at the pumping point where the volume fractions of all three phases are known from the flow simulation. Thus the heat withdrawal would be

$$SSH = \sum_{\alpha=1}^{3} C_{\alpha}\rho_{\alpha}\theta_{\alpha}QT^{h} \tag{21}$$

For the case of chemical reactions, SSH is given by

$$SSH = \sum_{k=1}^{N} \Delta H_{k} \tag{22}$$

where N is the number of reactions, and ΔH_k is the heat release rate of the k-th reaction.

4.4.1 Initial and Boundary Conditions for Heat Transfer

The initial conditions for heat transfer can be given as

$$T^{h} = T_{0}^{h}(x, y, z) \text{ in } R \tag{23}$$

Four types of boundary conditions are considered in the present study:

(1) Dirichlet boundary condition, on which the temperature is prescribed as a function of time on the boundaries:

$$T^{h} = T_{D}^{h}(x_{b}, t) \text{ on } B_{D} \tag{24}$$

where $T_D{}^h$ is the prescribed temperature on the boundary.

(2) Neumann boundary condition, on which the conductive heat flux is given:

$$-n \bullet (\lambda \mathbf{I} \bullet \nabla T^{h}) = q_{Tn}(x_{b}, t) \text{ on } B_{N} \tag{25}$$

where q_{TN} is the given conductive flux through the boundary

(3) Cauchy boundary condition, on which the total heat flux is given:

$$-\mathbf{n}\bullet\left[\lambda\mathbf{I}\bullet\nabla T^{h}-(\sum_{\alpha=1}^{3}C_{\alpha}\rho_{\alpha}\mathbf{V}_{\alpha})T^{h}\right]=q_{Tc}(x_{b},t)\ \text{on}\ B_{C} \tag{26}$$

where q_{TC} is the given total flux through the boundary

(4) Variable boundary condition

$$\mathbf{n}\bullet\nabla T^{h}=0\ \text{ if }\ \mathbf{n}\bullet\mathbf{V}_{\alpha}>0 \tag{27a}$$

$$-\mathbf{n}\bullet(\lambda\mathbf{I}\bullet\nabla T^{h}-C_{\alpha}\rho_{\alpha}\mathbf{V}_{\alpha}T^{h})=C_{\alpha}\rho_{\alpha}\mathbf{V}_{\alpha}T_{v}^{h}(x_{b},t)\ \text{ if }\ n\bullet\mathbf{V}_{\alpha}\le 0 \tag{27b}$$

where $\mathbf{V}_{\alpha}$ represents velocity of water, air, or NAPL phase, T_v^h is the specified temperature associated with each phase passing through the boundary. While Eq. (27a) is for the inward flow scenario, Eq. (27b) is for the outgoing flow scenario.

Like hydrologic transport, we can also derive the advective form (appendix 4.D) for heat transfer and use the five numerical schemes based on the conservative form and advective form of equation to solve heat transfer problem.

4.4.2 The Coupling of Chemical Transport and Heat Transfer

Weak coupling (Cheng, 1995) will be applied to solve for the system of chemical transport and heat transfer to save computing resources.

The temperature-dependent equilibrium constant K will be computed from the van't Hoff relationship as

$$K_{eq}=K_{eq,0}\exp\left[\frac{\Delta H_{rxn}}{R}\left(\frac{1}{T_{0}^{h}}-\frac{1}{T^{h}}\right)\right] \tag{28}$$

Where K_{eq} and $K_{eq,0}$ are the temperature-dependent and reference equilibrium constant respectively; ΔH_{rxn} is the enthalpy of the associated chemical reaction, which is unchanged with temperature; T^h is the temperature of interest in degrees of Kelvin; T_0^h is the reference temperature associated with $K_{eq,0}$; R is ideal gas constant.

The temperature-dependent kinetic constant will be computed from the Arrhenius equation (Steefel and Lasaga, 1994; Xu and Pruess, 1999):

$$k = k_0 \exp\left[\frac{-E_a}{R}\left(\frac{1}{T^h} - \frac{1}{T_0^h}\right)\right] \tag{29}$$

where E_a is the activation energy, k and k_0 are the temperature-dependent and reference rate constant respectively.

A user can also specify a function by fitting experiment data to consider the temperature-dependency of the constants.

4.5 EXAMPLE SIMULATIONS

Three examples are presented in this section. Problem 4.5.1 is used to verify the hydrologic transport part of 3DHYDROBIOGEOCHEM by comparing with a one dimensional analytical solution. Problem 4.5.2 is used to verify the heat transfer part of 3DHYDROBIOGEOCHEM by comparing with a one dimensional analytical solution. Problem 4.5.3 is the application of 3DHYDROBIOGEOCHEM to an imaginary problem, considering thermal effect on the hydrologic transport.

4.5.1 One Dimensional Solute Transport Verification

This example is modified from Sun et al. (1999). The problem domain is a three dimensional aquifer (100 m × 1 m × 1 m). The source location is on the left side of the domain. Considered in the system are one contaminant species and its daughter products produced by sequential degradation (first order rate), processes of which are shown in table 4.1. System parameters for this problem are shown in table 4.2.

Initially, all the species concentration c = 0. The column is connected to a continuous source at x = 0 with a constant $c_{10} = 1$, $c_{20} = 0$, $c_{30} = 0$, and $c_{40} = 0$.

Figure 4.1 shows the concentration profiles of the four after 100 d, 200 d, 300 d, and 400 d. The concentration profiles are compared with the analytical solution given by Sun et al. (1999). Except some spreading at earlier time steps, the results agree with the analytical solution well.

4.5.2 One Dimensional Heat Transfer Verification

This example solves the transient simulation of heat in a horizontal beam. The domain of interest has a size of 150 cm × 1 cm × 1 cm. The flow beam contains water ($S_w = 1$) with an initial temperature $T^h(x,0) = 0$ mmol/cm^3 of bulk volume. A steady flow field is assumed with a constant darcy velocity V = 1 cm/day and an effective moisture content of 0.2. The boundary condition at x = 0 is Dirichlet boundary condition with a constant temperature of 498 ^oK. The boundary at x = 150 cm is an open boundary. The domain is discretized with 100 equal size elements (1.5 cm × 1 cm × 1 cm each). A constant time step size of 0.1 day is used for the simulation. The heat capacity of water is 1.0×10^{20}. Heat capacity of dry medium is 5.0×10^{19}, and the apparent thermal conductivity is 1.0×10^{21}.

Figure 4.2 shows the temperature in the column after 8 days. It compares well with the analytical solution.

4.5.3 Three Dimensional Multiphase Non-isothermal Reactive Transport

This problem simulates biodegradation of benzene and toluene by attached biomass in the aquifer. The dimension of the problem domain is 250 m × 60 m × 15 m. This problem is modified from the LNAPL simulation example in UTCHEM technical documentation (2000). Some of the simulation parameters are taken from UTCHEM.

Point source of benzene and toluene from a gasoline spill is located in the range of 25 m $\leq y \leq$ 35 m, 85 m $\leq x \leq$ 90 m at z = 15 m, about one third from the upstream of the aquifer (figure 4.3). The contaminant source lasted for 4 days. Physical and biogeochemical parameters and values are shown in table 4.3 through table 4.6. Only longitudinal flow for water and NAPL phase are considered. Air phase is assumed to be immobile. The domain is discretized into $50 \times 12 \times 3$ elements (Figure 4.3).

Initial conditions are given in table 4.6. Upstream boundary is influent boundary, downstream boundary is natural boundary. Initial temperature in the system is 298 ^{o}K. Temperature in the influent flow is 498 ^{o}K. Because of the unavailable form of temperature dependency of kinetic constants, we assume they are not dependent on temperature. Only the octanol water partition coefficients depend on temperature. They are calculated from the van't Hoff relationship. Initial time step in the simulation is 10^{-4} day. Percent change in the time step size is 0.05. Maximum time step is 0.01 day. The first time the time step size changed to initial value is after 4 days when the spill stopped.

Major species concentration at the cross section of x = 90 m are shown in Figures 4.4 to 4.8. $O_{2(aq)}$ and $NO_3^-{}_{(aq)}$ distribution with respect to time is shown in Figure 4.8. We can see that $O_{2(aq)}$ is depleted at the completion of the contaminant spill. As we can expect, T_{NAPL} is degraded more rapidly than B_{NAPL} (it has a higher concentration in the contaminant source) because it can use both $O_{2(aq)}$ and $NO_3^-{}_{(aq)}$ as electron donor. The presence of preferential electron donor is important to the design of bioremediation processes.

Though we only considered biodegradation process in the attached biomass, HYDROBIOGEOCHEM can simulate the processes in non-attached biomass as well. The problem is just a demonstration of the capability of HYDROBIOGEOCHEM.

Parameters and values presented here may not be readily used for real system simulation.

4.6 CONCLUSION

In the present study, a new three dimensional numerical model for reactive chemical transport under multiphase systems is developed based on the new transport formulation, named 3DHYDROBIOGEOCHEM. In this formulation, the governing equation for transport is first decomposed into three subsets: transport for components; transport for kinetic variables; and mass action equations. The purpose of decomposition is to reduce PDVs to alleviate the coupling of transport and fast biogeochemical reaction processes.

APPENDIX 4.A ADVECTIVE FORM EQUATIONS FOR HYDROLOGIC TRANSPORT

Eq. (11) in section 2 is used for illustration in the following derivation. Expanding the second term of Eq. (11) yields

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \nabla \bullet (\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha}) + \sum_{\alpha=1}^{3} \rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha} \right) \right] = M_j + RR_j \tag{A.1}$$

Express T_j^α in terms of T_j to make T_js as PDVs:

$$\frac{\partial T_j}{\partial t} + \sum_{\alpha=1}^{3} \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} T_j \nabla \bullet (\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha}) + \sum_{\alpha=1}^{3} \rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} T_j \right) - \sum_{\alpha=1}^{3} \nabla \bullet \left[\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} T_j \right) \right] = M_j + RR_j \tag{A.2}$$

Expanding the third term on the left hand side of Eq. (A.2), we obtain

$$\sum_{\alpha=1}^{3} \rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} T_j \right) = T_j \sum_{\alpha=1}^{3} \rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} \right) + \sum_{\alpha=1}^{3} \rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} \bullet \nabla T_j \tag{A.3}$$

Expanding the fourth term on the right hand side of Eq. (A.2), we obtain

$$
\begin{aligned}
&\sum_{\alpha=1}^{3}\nabla\bullet\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j\right)\right]\\
&=\sum_{\alpha=1}^{3}\nabla\bullet\left\{\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j+\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\bullet\nabla T_j\right\}\\
&=\sum_{\alpha=1}^{3}\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]\bullet\nabla T_j+T_j\sum_{\alpha=1}^{3}\nabla\bullet\left[\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]\\
&+\sum_{\alpha=1}^{3}\nabla\bullet\left[\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\bullet\nabla T_j\right]
\end{aligned}
\tag{A.4}
$$

Substituting Eqs. (A.3), (A.4) into Eq. (A.2) and combining similar terms, the following equation is obtained

$$
\begin{aligned}
&\frac{\partial T_j}{\partial t}+\sum_{\alpha=1}^{3}\left[\rho_\alpha\theta_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\mathbf{V}_{f,\alpha}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right]\bullet\nabla T_j\\
&-\sum_{\alpha=1}^{3}\nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\right)\\
&+\sum_{\alpha=1}^{3}\left[\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\right)+\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}-\nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j\\
&=M_j+RR_j
\end{aligned}
\tag{A.5}
$$

Combining Eq. (A.5) and the continuity equation of flow

$$
\frac{\partial\rho_\alpha\theta_\alpha}{\partial t}=-\nabla\bullet\left(\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\right)+\rho_\alpha^{*}Q_\alpha,\ \alpha=\text{water, gas or oil}
\tag{A.6}
$$

yields the following equation

$$
\frac{\partial T_j}{\partial t}+L(T_j)+KT_j+\sum_{\alpha=1}^{3}\frac{\rho_\alpha^{*}Q_\alpha T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j=M_j+RR_j
\tag{A.7}
$$

where ρ_i^{*} is the density of source/sink fluid, (M of liquid/L^3 of liquid), equal to ρ_i if

sink/withdrawal (negative Q) or $\rho_{i,in}$ if source/injection (positive Q); Q is the source or sink due to artificial injection or withdrawal of water, unit of volume of liquid per volume of porous media per time (L^3/L^3 T). K is a first order term and L( ) is an advection-dispersion operator:

$$K = \sum_{i=1}^{3}\left[\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha} \bullet \nabla\left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j}\right) - \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j}\frac{\partial \rho_\alpha \theta_\alpha}{\partial t} - \nabla \bullet \left(\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j}\right)\right] \tag{A.8}$$

$$L(T_j) = \mathbf{V}_e \bullet \nabla T_j - \nabla \bullet [\, \mathbf{D}_e \bullet \nabla T_j \,], \quad j \in N_m \tag{A.9}$$

where $\mathbf{V}_e$ and $\mathbf{D}_e$ are effective or true transport velocity and effective or true dispersion coefficient respectively, and they are given by

$$\mathbf{V}_e = \sum_{\alpha=1}^{3}\left[\rho_\alpha \theta_\alpha \mathbf{V}_{f,\alpha}\left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j}\right) - \rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \frac{T_j^i}{\rho_\alpha \theta_\alpha T_j}\right] \tag{A.10}$$

$$\mathbf{D}_e = \sum_{\alpha=1}^{3}\left(\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j}\right) \tag{A.11}$$

The source/sink term can be defined by

$$M_j = \sum_{i=1}^{3}\left(\rho_\alpha^* T_j^{\alpha *} Q\right) \tag{A.12}$$

where $T_j^{\alpha *}$ is the total concentration of the j-th component in the α-th source/sink liquid (M/M of liquid), and $T_j^{\alpha *} = \frac{T_j^i}{\rho_\alpha \theta_\alpha}$ in the case of sink/withdrawl, or $T_j^{\alpha *} = T_{j,in}^{\alpha}$ -- input concentration in the case of source/injection.

APPENDIX 4.B INITIAL AND BOUNDARY CONDITIONS FOR HYDROLOGIC TRANSPORT

Initially, the concentration of each chemical species must be given throughout the region of interest:

$$\rho_\alpha \theta_\alpha c_i = C_{i0} \text{ at } t = 0,\ i \in M_s \tag{B.1}$$

where C_{i0} is the initial concentration of the i-th chemical species, unit of mass of chemical species per volume of porous media (M/L^3).

For all species that are subject to hydrological transport, boundary conditions are specified. Types of boundary conditions incorporated are: Dirichlet, Neumann, Cauchy, or variable.

A Dirichlet condition is used to deal with a prescribed concentration as:

$$\rho_\alpha \theta_\alpha c_i = C_{iD} \quad \text{on } B_D,\ i \in M_s \tag{B.2}$$

Where C_{iD} is the prescribed Dirichlet concentration for the i-th chemical species, unit of mass of chemical species per volume of porous media (M/L^3).

A boundary is often treated as a Neumann boundary when the flux due to the gradient of species concentration is known, or as a Cauchy boundary when the total flux is given:

$$-\mathbf{n} \bullet (\rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla c_i) = q_{iN,\alpha} \quad \text{on } B_N,\ i \in M_s \tag{B.3}$$

$$\mathbf{n} \bullet (\rho_\alpha \mathbf{V}_\alpha c_i - \rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla c_i) = q_{iC,\alpha} \quad \text{on } B_C,\ i \in M_s \tag{B.4}$$

where q_{iN} is the normal Neumann flux, unit of mass of the i-th chemical species per unit area of porous media per time ($M/L^2/T$), q_{iC} is the normal Cauchy flux, unit of mass of the i-th chemical species per unit area of porous media per time($M/L^2/T$), **n** is an

outward unit vector normal to the boundary.

A flow-through boundary is normally the soil-air interface or soil-water interface, and the appropriate condition is variable depending on the direction of transport across the boundary. For this type of boundary, a variable-type boundary is imposed. When the flow is directed out of the region, the boundary is depicted by the Neumann condition with zero gradient flux; when the flow is directed into the region, the boundary is depicted by the Cauchy condition with given total flux. Therefore, the variable boundary condition can be written as

$$\begin{aligned} &-\mathbf{n}\bullet(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla c_i)=q_{iV,\alpha}=0 \quad \text{if } \mathbf{V}_\alpha\bullet\mathbf{n}>0 \text{ on } B_V,\ i\in M_s \\ &\mathbf{n}\bullet(\rho_\alpha\mathbf{V}_\alpha c_i-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla c_i)=q_{iV,\alpha} \quad \text{if } \mathbf{V}_\alpha\bullet\mathbf{n}<0 \text{ on } B_V,\ i\in M_s \end{aligned} \tag{B.5}$$

Expressed in the form of total phase concentration, Eqs. (B.1) to (B.5) can be written as:

$$T_j=T_{j0} \quad \text{at } t=0,\ j\in(N_c+M) \tag{B.6}$$

$$-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jN,\alpha} \quad \text{on } B_N,\ j\in(N_m+M_m) \tag{B.7}$$

$$\mathbf{n}\bullet\left(\rho_\alpha\mathbf{V}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jC,\alpha} \quad \text{on } B_C,\ j\in(N_m+M_m) \tag{B.8}$$

$$-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jV,\alpha}=0 \quad \text{if } \mathbf{V}_\alpha\bullet\mathbf{n}>0 \text{ on } B_V,\ j\in(N_m+M_m) \tag{B.9}$$

$$\mathbf{n}\bullet\left(\rho_\alpha\mathbf{V}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha}\right)=q_{jV,\alpha} \quad \text{if } \mathbf{V}_\alpha\bullet\mathbf{n}<0 \text{ on } B_V,\ j\in(N_m+M_m) \tag{B.10}$$

When applying FEM to the advective form of the governing equations, boundary conditions should be reformulated to be consistent with Eq.(A.7) (regarding to the effective velocity and effective dispersion tensor).

Express Eq. (B.7) in terms of T_j, we obtain

$$
\begin{aligned}
&-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j\right)\right)\\
&=-\mathbf{n}\bullet\left(\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\left(\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)T_j+\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\right)=q_{jN,\alpha}
\end{aligned}
\tag{B.11}
$$

Add $\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\dfrac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j$ to both sides of Eq. (B.7), we obtain

$$
\begin{aligned}
&\mathbf{n}\bullet\left[\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j-\mathbf{n}\bullet\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\\
&=\mathbf{n}\bullet\mathbf{V}_{f,\alpha}\frac{T_j^\alpha}{T_j}T_j+q_{jN,\alpha},\ \text{on } B_N\ ,\ j\in(N_m+M_m)
\end{aligned}
\tag{B.12}
$$

Similarly, express Eq. (B.8) in terms of T_j, we obtain

$$
\mathbf{n}\bullet\left(\rho_\alpha\mathbf{V}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}T_j\right)=q_{jC,\alpha}
\tag{B.13}
$$

Expand the second term on the left hand side of Eq. (B.13) as in Eq. (B.11) and combine similar terms, the following equation is obtained

$$
\begin{aligned}
&\mathbf{n}\bullet\left[\rho_\alpha\mathbf{V}_\alpha\frac{T_j^i}{\rho_\alpha\theta_\alpha T_j}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j-\mathbf{n}\bullet\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\\
&=q_{jC,\alpha}\ \text{ on } B_C\ ,\ j\in(N_m+M_m)
\end{aligned}
\tag{B.14}
$$

Eq. (B.9) and (B.10) are special case of Eqs. (B.7) and (B.8), respectively, and they can be easily written out as

$$
\begin{aligned}
&\mathbf{n}\bullet\left[\rho_\alpha\theta_\alpha\mathbf{V}_{f,\alpha}\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}-\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\bullet\nabla\left(\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\right)\right]T_j-\mathbf{n}\bullet\rho_\alpha\theta_\alpha\mathbf{D}_\alpha\frac{T_j^\alpha}{\rho_\alpha\theta_\alpha T_j}\bullet\nabla T_j\\
&=\mathbf{n}\bullet\mathbf{V}_{f,\alpha}\frac{T_j^\alpha}{T_j}T_j\ \text{ if } \mathbf{V}_\alpha\bullet\mathbf{n}>0\ \text{ on } B_V\ ,\ j\in(N_m+M_m)
\end{aligned}
\tag{B.15}
$$

$$\mathbf{n} \bullet \left[\rho_\alpha \mathbf{V}_\alpha \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} - \rho_\alpha \theta_\alpha \mathbf{D}_\alpha \bullet \nabla \left(\frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} \right) \right] T_j - \mathbf{n} \bullet \rho_\alpha \theta_\alpha \mathbf{D}_\alpha \frac{T_j^\alpha}{\rho_\alpha \theta_\alpha T_j} \bullet \nabla T_j = q_{jV,\alpha} \tag{B.16}$$

$$\text{if } \mathbf{V}_\alpha \bullet \mathbf{n} < 0 \text{ on } B_V \text{ , } j \in (N_m + M_m)$$

APPENDIX 4.C MATRIX EQUATIONS FOR DIFFERENT NUMERICAL SCHEMES

- **Scheme 1: FEM on Advective Form of Equation**

The upstream finite element is employed to solve the hydrologic transport equation in advective form. The transport equation is discretized as follows:

$$[M]\left\{\frac{dT}{dt}\right\}+([L]+[V]+[K]+[Q])\{T\}=\{S\}+\{B\}+\{R\} \tag{C.1}$$

where [M] is the mass matrix; [L] is the stiff matrix resulting from dispersion/diffusion; [V] is the stiff matrix resulting from the velocity term; [K] is the stiff matrix resulting from the first order rate term; [Q] is the stiff matrix resulting from the source term; {S} is the load vector from external sources; {B} is the load vector from boundary sources; {R} is the load vector from the biogeochemical reactions; $\left\{\frac{dT}{dt}\right\}$ is the column vector containing the values of $\frac{dT}{dt}$; $\{T\}$ is the column vector containing the nodal values of *T*, and

$$M_{ij}=\sum_{e\in M_e}\int_{R_e}N_\alpha^e N_\beta^e dR \tag{C.2}$$

$$L_{ij}=\sum_{e\in M_e}\sum_{\alpha=1}^{3}\int_{R_e}\left(\nabla N_\alpha^e\right)\bullet\rho_\ell\theta_\ell\mathbf{D}_\ell\frac{T_\ell}{\rho_\ell\theta_\ell T}\bullet\left(\nabla N_\beta^e\right)dR \tag{C.3}$$

$$V_{ij}=\sum_{e\in M_e}\sum_{\ell=1}^{3}\int_{R_e}W_\alpha^e\left(\rho_\ell\mathbf{V}_\ell\frac{T_\ell}{\rho_\ell\theta_\ell T}-\rho_\ell\theta_\ell\mathbf{D}_\ell\bullet\nabla\frac{T_\ell}{\rho_\ell\theta_\ell T}\right)\nabla N_\beta^e dR \tag{C.4}$$

$$K_{ij} = \sum_{e \in M_e} \sum_{\ell=1}^{3} \int_{R_e} N_\alpha^e \left[\rho_\ell \mathbf{V}_\ell \bullet \nabla \frac{T_\ell}{\rho_\ell \theta_\ell T} - \frac{T_\ell}{\rho_\ell \theta_\ell T} \frac{\partial \rho_\ell \theta_\ell}{\partial t} - \nabla \bullet \left(\rho_\ell \theta_\ell \mathbf{D}_\ell \bullet \nabla \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) \right] N_\beta^e dR \quad (C.5)$$

$$Q_{ij} = \sum_{e \in M_e} \sum_{\ell=1}^{3} \int_{R_e} N_\alpha^e \left(\rho^*_\ell Q_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) N_\beta^e dR \quad (C.6)$$

$$B_i = -\sum_{e \in N_{Se}} \sum_{\alpha=1}^{3} \int_{B_e} N_\alpha^e n \bullet \left(\rho_\ell \mathbf{V}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell} - \rho_\ell \theta_\ell D_\ell \frac{T_\ell}{\rho_\ell \theta_\ell} \right) dB \quad (C.7)$$

$$S_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left(\sum_{\ell=1}^{3} Q_\ell T_\ell^{in} \right) dR \quad (C.8)$$

$$R_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e RR \, dR \quad (C.9)$$

Where R_e = region of element e.

M_e = set of elements that have a local side α-β coinciding with the global side i-j.

N_α^e = α-th local base function of element e.

W_α^e = α-th local weighting function of element e.

B_e = the length of boundary segment e.

N_{Se} = set of boundary segments having a local node α coinciding with global node i.

- **Scheme 2: FEM on Conservative Form of Equation**

The upstream finite element is employed to solve the hydrologic transport equation in conservative form. The transport equation is discretized as follows:

$$[M]\left\{\frac{dT}{dt}\right\}+([L]+[V])\{T\}=\{S\}+\{B\}+\{R\} \tag{C.10}$$

where [M] is the mass matrix; [L] is the stiff matrix resulting from dispersion/diffusion; [V] is the stiff matrix resulting from the velocity term; {S} is the load vector from external sources; {B} is the load vector from boundary sources; {R} is the load vector from the biogeochemical reactions, and

$$M_{ij}=\sum_{e\in M_e}\int_{R_e}N_\alpha^e N_\beta^e dR \tag{C.11}$$

$$L_{ij}=\sum_{e\in M_e}\sum_{\alpha=1}^{3}\int_{R_e}\left(\nabla N_\alpha^e\right)\bullet\rho_\ell\theta_\ell\mathbf{D}_\ell\frac{T_\ell}{\rho_\ell\theta_\ell T}\bullet\left(\nabla N_\beta^e\right)dR \tag{C.12}$$

$$V_{ij}=-\sum_{e\in M_e}\sum_{\ell=1}^{3}\int_{R_e}\left(\nabla W_\alpha^e\right)\bullet\left(\rho_\ell\mathbf{V}_\ell\frac{T_\ell}{\rho_\ell\theta_\ell T}-\rho_\ell\theta_\ell\mathbf{D}_\ell\bullet\nabla\frac{T_\ell}{\rho_\ell\theta_\ell T}\right)N_\beta^e dR \tag{C.13}$$

$$B_i=-\sum_{e\in N_{Se}}\sum_{\alpha=1}^{3}\int_{B_e}N_\alpha^e n\bullet\left(\rho_\ell\mathbf{V}_\ell\frac{T_\ell}{\rho_\ell\theta_\ell}-\rho_\ell\theta_\ell D_\ell\frac{T_\ell}{\rho_\ell\theta_\ell}\right)dB \tag{C.14}$$

$$S_i=\sum_{e\in M_e}\int_{R_e}N_\alpha^e\left(\sum_{\ell=1}^{3}Q_\ell T_\ell^{in/out}\right)dR \tag{C.15}$$

$$R_i=\sum_{e\in M_e}\int_{R_e}N_\alpha^e RR\,dR \tag{C.16}$$

Where R_e = region of element e.

M_e = set of elements that have a local side α-β coinciding with the global side i-j.

N_α^e = α-th local base function of element e.

W_α^e = α-th local weighting function of element e.

B_e = the length of boundary segment e.

N_{Se} = set of boundary segments having a local node α coinciding with global

node i.

- **Scheme 3: Hybrid Lagrangian-Eulerian FEM for Interior Elements + FEM on Advective Form of Equation for Boundary elements**

The transport equation for interior elements is discretized as follows:

$$[\mathrm{M}]\left\{\frac{\mathrm{DT}}{\mathrm{D}\tau}\right\}+([\mathrm{L}]+[\mathrm{K}]+[\mathrm{Q}])\{\mathrm{T}\}=\{\mathrm{S}\}+\{\mathrm{R}\} \tag{C.17}$$

where [M], [L], [K], [Q], {S}, and {R} are the defined the same as those in Eq. (C.1). $\left\{\frac{DT}{D\tau}\right\}$ is a column vector containing the values of $\frac{DT}{D\tau}$; τ is the backward tracking time.

The transport equation for boundary elements is the same as Eq. (C.1).

- **Scheme 4: Hybrid Lagrangian-Eulerian FEM**

To overcome the inaccuracy caused by tracking time at the boundary node, we'll first approximate the time derivative with finite difference method and then multiply the governing equation with the time interval. With similar procedure of deriving governing equations in conventional FEM, applying Galerkin FEM to the resulted equation, following linearized equation is obtained:

$$[\mathrm{M}]\{\Delta\mathrm{T}\}+([\mathrm{L}']+[\mathrm{K}']+[\mathrm{Q}'])\{\mathrm{T}\}=\{\mathrm{S}'\}+\{\mathrm{B}'\}+\{\mathrm{R}'\} \tag{C.18}$$

where [M] is the mass matrix, $\{\Delta\mathrm{T}\}$ is a vector composed of the difference of values of T between the new and old time steps. [L′] is the stiff matrix resulting from dispersion,

[K′] is the stiff matrix resulting from the first order rate term, [Q′] is the stiff matrix resulting from the source term, {S′} is the load vector from external sources, {B′} is the load vector from boundary sources, {R′}is the load vector from the biogeochemical reactions, and

$$M_{ij} = \sum_{e \in M_e} \int_{R_e} N_\alpha^e N_\beta^e dR \tag{C.19}$$

$$K'_{ij} = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left\{ \sum_{\ell=1}^{3} \left[\rho_\ell \mathbf{V}_\ell \bullet \nabla \frac{T_\ell}{\rho_\ell \theta_\ell T} - \frac{T_\ell}{\rho_\ell \theta_\ell T} \frac{\partial \rho_\ell \theta_\ell}{\partial t} - \nabla \left(\rho_\ell \theta_\ell \mathbf{D}_\ell \bullet \nabla \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) \right] \right\} \Delta\tau \, N_\beta^e dR \tag{C.20}$$

$$L'_{ij} = \sum_{e \in M_e} \int_{R_e} \left(\nabla N_\alpha^e \right) \cdot \left(\sum_{\ell=1}^{3} \rho_\ell \theta_\ell \mathbf{D}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) \Delta\tau \cdot \left(\nabla N_\beta^e \right) dR \tag{C.21}$$

$$Q'_{ij} = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left(\sum_{\ell=1}^{3} \rho_\ell^* Q_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \right) \Delta\tau \, N_\beta^e dR \tag{C.22}$$

$$B'_i = \sum_{e \in N_{Se}} \int_{B_e} N_\alpha^e \mathbf{n} \bullet \left(\sum_{\ell=1}^{3} \rho_\ell \theta_\ell \mathbf{D}_\ell \frac{T_\ell}{\rho_\ell \theta_\ell T} \bullet \nabla T \right) \Delta\tau \, dB \tag{C.23}$$

$$S'_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e \left(\sum_{\ell=1}^{3} Q_\ell T_\ell^{in} \right) \Delta\tau \, dR \tag{C.24}$$

$$R'_i = \sum_{e \in M_e} \int_{R_e} N_\alpha^e RR \, \Delta\tau \, dR \tag{C.25}$$

Where R_e = region of element e.

M_e = set of elements that have a local side α-β coinciding with the global side i-j.

N_α^e = α-th local base function of element e.

B_e = the length of boundary segment e.

N_{Se} = set of boundary segments having a local node α coinciding with global node i.

- **Scheme 5: Hybrid Lagrangian-Eulerian FEM for Interior Elements + FEM on Conservative Form of Equation for Boundary elements**

In this scheme, the transport equation for interior elements is discretized as Eq. (C.17). The transport equation for boundary elements is the same as Eq. (C.10).

APPENDIX 4.D ADVECTIVE FORM EQUATIONS FOR HEAT TRANSFER

Multiply both sides of the continuity equation of flow, Eq. (A.6), by $C_\alpha T$ (α represents water, air and NAPL), we get

$$C_\alpha T^h \frac{\partial \rho_\alpha \theta_\alpha}{\partial t} = - C_\alpha T^h \nabla \bullet (\rho_\alpha V_\alpha) + C_\alpha T^h \rho_\alpha^* q_\alpha \tag{D.1}$$

Substitution of Eqs.(D.1) into Eq. (20) in section 4.4 yields

$$\begin{aligned} &\left(\sum_{\alpha=1}^{3} C_\alpha \rho_\alpha \theta_\alpha + C_m \rho_b \right) \frac{\partial T^h}{\partial t} + \sum_{\alpha=1}^{3} (C_\alpha \rho_\alpha V_\alpha) \bullet \nabla T^h - \nabla \bullet (\lambda \mathbf{I} \bullet \nabla T^h) \\ &= \sum_{\alpha=1}^{3} (C_\alpha \rho^*_\alpha q_\alpha) T^{hs} - \sum_{\alpha=1}^{3} (C_\alpha \rho^*_\alpha q_\alpha) T^h \end{aligned} \tag{D.2}$$

Where T^{hs} is the source temperature.

Eq. (D.2) can also be written in the following form for the convenience of implementing Langrangian-Elerian finite element method:

$$\left(\sum_{\alpha=1}^{3} C_\alpha \rho_\alpha \theta_\alpha + C_m \rho_b \right) \frac{DT^h}{Dt} - \nabla \bullet (\lambda \mathbf{I} \bullet \nabla T^h) + \sum_{\alpha=1}^{3} (C_\alpha \rho^*_\alpha q_\alpha) T^h = \sum_{\alpha=1}^{3} (C_\alpha \rho^*_\alpha q_\alpha) T^{hs} \tag{D.3}$$

and the tracking velocity is expressed as

$$V_T = \frac{\sum_{\alpha=1}^{3} C_\alpha \rho_\alpha \mathbf{V}_\alpha}{\sum_{\alpha=1}^{3} C_\alpha \rho_\alpha \mathbf{V}_\alpha + C_m \rho_b} \tag{D.4}$$

ACKNOWLEDGEMENTS. This research was supported, in part, by National Science Foundation (NSF) Grant No.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 No.DE-FG02-98ER62691 with the Pennsylvania State University.

REFERENCES

Brun, A., Engesgaard, P., Modelling of tranpsort and biogeochemical processes in pollution plumes: literature review and model development, Journal of Hydrology, 256, 211-227 (2002).

Burgos W. D., Royer R. A., Fang Y., Yeh G., Fisher A. S., Jeon B. H., and Dempsey B. A., 2002. Theoretical and experimental considerations related to reaction-based modeling : A case study using iron(III) oxide bioreduction. *Geomicrobiology Journal* **19**(2), 1-35.

Burgos W. D., Fang Y., Royer R. A.,Yeh G., Stone, J.J., Jeon B. H., and Dempsey B. A. , 2003. Reaction-based modeling of quinine-mediated bacterial iron(III) reduction. Geochmica Cosmochimica Acta (in press).

Chen, Y., Abriola, L.M., Alvarez, P.J.J., Anid P.J. and Vogel, T.M., 1992. Modeling transport and biodegradatino of benzene and toluene in sandy aquifer material: comparisons with experimental measurements, Water Resources Res., 28(7): 1833-1847.

Cheng, H.-P., 1995. Development and Application of a Three-Dimensional Finite Element Model of Subsurface Flow, Heat Transfer, and Reactive Chemical Transport. Ph.D. Thesis, Department of Civil and Environmental Engineering, The Pennsylvania State University. University Park, PA.

Fang, Y., Yeh, G.T., and Burgos, W.T., 2003. A generic paradigm to model reaction-based biogeochemical processes. Water Resources Research (in press).

Gerke, H.H., Molson, J.W., and Frind, E.O., 1998. Modelling the effect of chemical heterogeneity on acidification and solute leaching in overburden mine spoils. J. Hydrol. 209: 166-185.

Liu, C.W., Narasimhan, T.N.,Redox-controlled-multiple-species reactive chemical transport, model development, Water Resour. Res. 25(5),869-82 (1989).

Millington, R.J., Quirk, J.M., 1961. Permeability of porous solids, Trans. Faraday Soc., 7, 1200-1207.

Nguyen, V.V., Gray, W.G., Pinder, G.F., Botha, J.F., and Gerar, D.A., 1982. A theoretical investigation on the transport of chemicals in reactive porous media. Water Resour. Res. 18, 1149-1156.

Owczarek, J.A. *Fundamentals of Gas Dynamics*. International Textbook Company, Scranton, PA., 668pp, 1964.

Salvage, K.M., 1998. Reactive contaminant tranpsport in variably saturated porous media: biogeochemical model development, verification, and application. . Ph.D. Thesis, Department of Civil and Environmental Engineering, The Pennsylvania State University. University Park, PA.

Steefel, C.I., and Lasaga, A.C., 1994. A coupled model for transport of multiple chemical species and kinetic precipitaton/dissolution reactions with applications to reactive flow in single phase hydrothermal system. American Journal of Science, 294, 529-592.

Steefel, C.I., MacQuarrie, K.T., Approaches to modeling reactive transport in porous media. In: Lichtner, P.C., Steefel, C.T., Oelkers, E.H. (Eds.), Reactive Transport in Porous Media, Reviews in Mineralology 34. Mineral. Soc. Am. 83,83-129 (1996).

Stumm W. and Morgan J. J. (1996) *Aquatic Chemistry*. John Wiley and Sons.

Sun, Y., Petersen, J.N., and Clement, T.P., 1999. Analytical solutions for multiple species reactive transport in multiple dimensions. Journal of Contaminant Hydrology, 35,

429-440.

Viswanathan, H.S., Robinson, B.A., Valocchi, A.J., and Triay, I.R., A reactive transport modelof neptuniun migration from the potential repository at Yucca Mountain. J. Hydrol. 209: 251-280 (1998).

Wu, Y.S. and Pruess, K., 2000. Numerical simulation of non-isothermal multiphase tracer transport in heterogeneous fractured porous media. Advances in water resources, 23: 699-723.

Xu, T. and Pruess, 2001. Modeling multiphase non-isothermal fluid flow and reactive geochemical transport in variably saturated fractured rocks: 1. methodology. American Journal of Science, 301(1), 16-33.

Yeh G.T. and Luxmoore, R.J., 1983. MATTUM: a multidimensional model for simulating moisture and thermal transport in unsaturated porous media, Report ORNL-5888, Oak Ridge National Laboratory, Oak Ridge, Tenn. 37831.

Yeh, G.T. and Tripathi, V.S., A Critical evaluation of recent developments in hydrogeochemical transport models of reactive multichemical components, Water Resour. Res. 25(1), 93-108 (1989).

Yeh, G.T., Burgos, W.D, Zachara, J.M., 2001a. Modeling and measuring biogeochemical reactions: system consistency, data needs, and rate formulations, Adv. Envrion. Res. 5, 219-237.

Yeh, G.T., Physics and numerics of reaction-based transport modeling, invited presentation at NABIR workshop on biogeochemical modeling, Oct. 21-23 (2001b).

- Technical Documentation for UTCHEM-9.0, a three-dimensional chemical flood simulator, Volume II, 256pp. The University of Texas at Austin, Austin, Texas 78712.

Table 4.1. Processes considered in problem 4.1.

Process	#	Rate
c1 ↔ c2	(R1)	$R_1 = k_1 c_1$
c2 ↔ c3	(R2)	$R_2 = k_2 c_2$
c3 ↔ c4	(R3)	$R_3 = k_3 c_3$
c4 ↔ϕ*	(R4)	$R_4 = k_4 c_4$

* ϕ represent dead species.

Table 4.2. System parameters and values

Parameter	Value
Pore velocity	v_x = 0.2 m/d, v_y = 0 m/d, v_z = 0 m/d
Dispersivity	α_x = 1.5 m, α_y = 0.45 m, α_z = 0.15 m
Decay rate of species 1	k_1 = 0.05 /d
Decay rate of species 2	k_2 = 0.02 /d
Decay rate of species 3	k_3 = 0.01 /d
Decay rate of species 4	k_4 = 0.005 /d

Table 4.3 Flow Parameters and Values

Parameter	Value
darcy velocity (water), v (m/d)	v_x =0.1, v_y = 0.0, v_z = 0.0
darcy velocity (NAPL), v (m/d)	v_x =0.001, v_y = 0.0, v_z = 0.0
porosity, φ	0.38
water density, ρ_w (g/cm ^{3})	1.0
NAPL density, ρ_n (g/cm ^{3})	0.87
bulk soil density, ρ_b (g/cm ^{3})	1.64
longitudinal dispersivity, a_L (m)	5
transverse dispersivity, a_T (m)	0.625
water saturation, S_w	0.5
NAPL saturation, S_n	0.3
flow rate of contaminant source point (m^3/d)	0.95
benzene concentration in the source, (mol/L)	0.125
toluene concentration in the source, (mol/L)	0.56

Table 4.4 Biogeochemical Processes

Process	#	Type
$B_{NAPL} \leftrightarrow B_{aq}$	(R1)	Phase partition, K_{ow} = 134.9
$T_{NAPL} \leftrightarrow T_{aq}$	(R2)	Phase partiton, K_{ow} = 537
$B_{aq} \leftrightarrow B_s$	(R3)	Kinetic adsorption
$T_{aq} \leftrightarrow T_s$	(R4)	Kinetic adsorption
$O_{2(aq)} \leftrightarrow O_{2(s)}$	(R5)	Kinetic adsorption
$NO_3^-{}_{(aq)} \leftrightarrow NO_3^-{}_{(s)}$	(R6)	Kinetic adsorption
$B_s + 7.5\ O_{2(s)} \leftrightarrow Y_b\ X_s$	(R7)	Biodegradation in attached biomass
$T_s + 9.0\ O_{2(s)} \leftrightarrow Y_t\ X_s$	(R8)	Biodegradation in attached biomass
$T_s + 7.2\ NO_3^-{}_{(s)} \leftrightarrow Y_t\ X_s$	(R9)	Biodegradation in attached biomass
Endogenous decay of biomass	(R10)	First order kinetics

Note: B-benzene, T-toluene, X - biomass, $-_{NAPL}$ - species in NAPL phase, $-_{aq}$ species in aqueous phase, $-_s$ - species in attached biomass

Table 4. 5 Reaction Rates

Process #	Reaction Rate	
(R3)	$r_3 = \frac{\beta\kappa X_s}{m_c}(B_{aq} - B_s)$	(30)
(R4)	$r_4 = \frac{\beta\kappa X_s}{m_c}(T_{aq} - T_s)$	(31)
(R5)	$r_5 = \frac{\beta\kappa X_s}{m_c}(O_{2(aq)} - O_{s(s)})$	(32)
(R6)	$r_6 = \frac{\beta\kappa X_s}{m_c}(NO^-_{3(aq)} - NO^-_{3(s)})$	(33)
(R7)	$r_7 = \frac{\mu_{max} X_s}{Y_b}\left(\frac{B_s}{K_b + B_s}\right)\left(\frac{O_{2(s)}}{K_o + O_{s(s)}}\right)$	(34)
(R8)	$r_8 = \frac{\mu_{max} X_s}{Y_t}\left(\frac{T_s}{K_t + T_s}\right)\left(\frac{O_{2(s)}}{K_o + O_{s(s)}}\right)$	(35)
(R9)	$r_9 = \frac{\mu_{max} X_s}{Y_t}\left(\frac{T_s}{K_t + T_s}\right)\left(\frac{NO^-_{3(s)}}{K_n + NO^-_{3(s)}}\right)$	(36)
(R10)	$r_{10} = -bX_s$	(37)

where (UTCHEM, 2000)

β = surface area of a single microcolony (L^2);

κ = mass transfer coefficient ($L\ T^{-1}$);

m_c = mass of cells in a single microcolony;

μ_{max} = maximum specific growth rate (T^{-1});

Y = yield coefficient; mass of cells produced per mass of substrate biodegraded;

K = substrate half-saturation coefficient ($M\ L^{-3}$);

b = endogeneous decay coefficient (T^{-1}).

Table 4.6 Simulation parameters

Parameter	Value
Physiochemical parameters	
Microcolony surface area,(m^2 /microcolony)	1.19×10^{-10}
m_c = mass of cells in a single microcolony (g/microcolony)	1.0×10^{-16}
Mass transfer coefficient for benzene, k_b (m^2 /d)	8.52×10^{2}
Mass transfer coefficient for toluene, k_t (m^2 /d)	7.89×10^{2}
Mass transfer coefficient for oxygen, k_o (m^2 /d)	7.92×10^{-1}
Mass transfer coefficient for nitrate, k_n (m^2 /d)	6.52×10^{-1}
Microbial parameters (modified from Chen et al ., 1992 in UTCHEM, 2000)	
Initial attached biomass concentration, X_s (M/L)	3.36×10^{-6}
Maximum specific growth rate on benzene, $\mu_{max,b}$ (d^{-1})	6.806
Maximum specific growth rate on toluene, $\mu_{max,t}$ (d^{-1})	8.118
Yield coefficient for benzene, Y_b (mole cells/mole benzene)	0.3456
Yield coefficient for toluene, Y_t (mole cells/mole toluene)	0.4077
Half-saturation coef. of benzene for oxygen respiration, K_{bo} (mole/L)	5.9×10^{-5}
Half-saturation coef. of toluene for oxygen respiration, K_{bo} (mole/L)	7.176×10^{-5}
Half-saturation coef. of toluene for nitrate respiration, K_{to} (mole/L)	1.066×10^{-4}
Half-saturation coef. of oxygen for benzene biodeg , K_{ob} (mole/L)	1.188×10^{-6}
Half-saturation coef. of oxygen for toluene biodeg , K_{ot}(mole/L)	1.188×10^{-7}
Half-saturation coef.of nitrate for toluene biodeg., K_{nt} (mole/L)	1.59×10^{-5}
Endogenous decay coefficient, b (d^{-1})	0.1

Initial $O_{2(aq)}$ concentration (mol/L)	2.5×10^{-4}
Initial $NO_3^-{}_{(aq)}$ concentration (mol/L)	1.6×10^{-4}
Influent $O_{2(aq)}$ concentration (mol/L)	2.5×10^{-4}
Influent $NO_3^-{}_{(aq)}$ concentration (mol/L)	1.6×10^{-4}

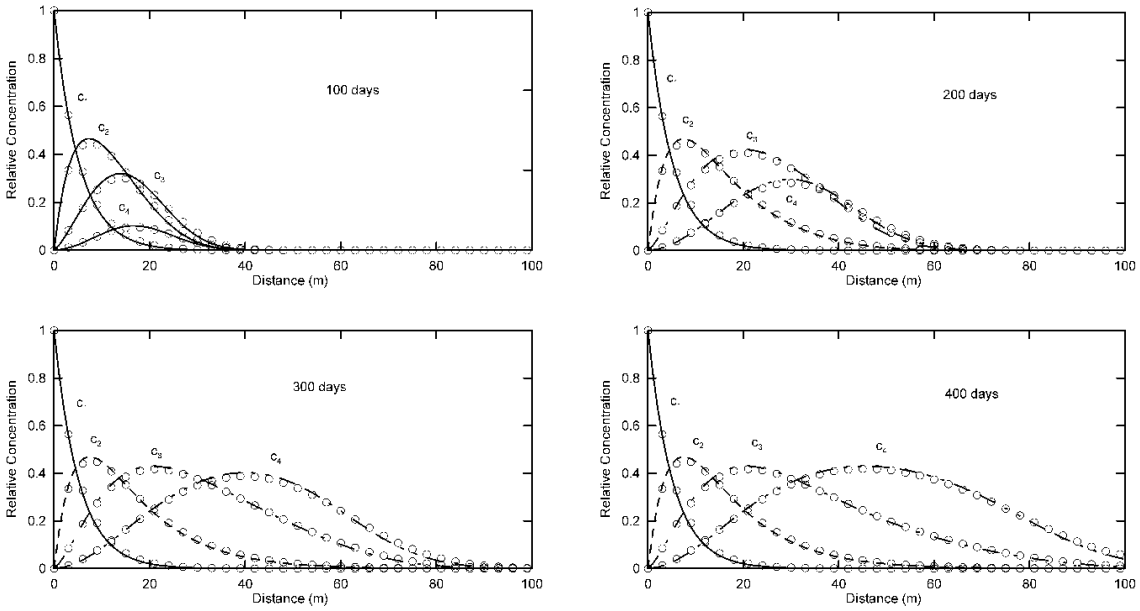


Figure 4.1. Concentration profile comparison with Sun et al. (1999) at selected time steps. The solid line is from Sun et al., circle is from HYDROBIOGEOCHEM.

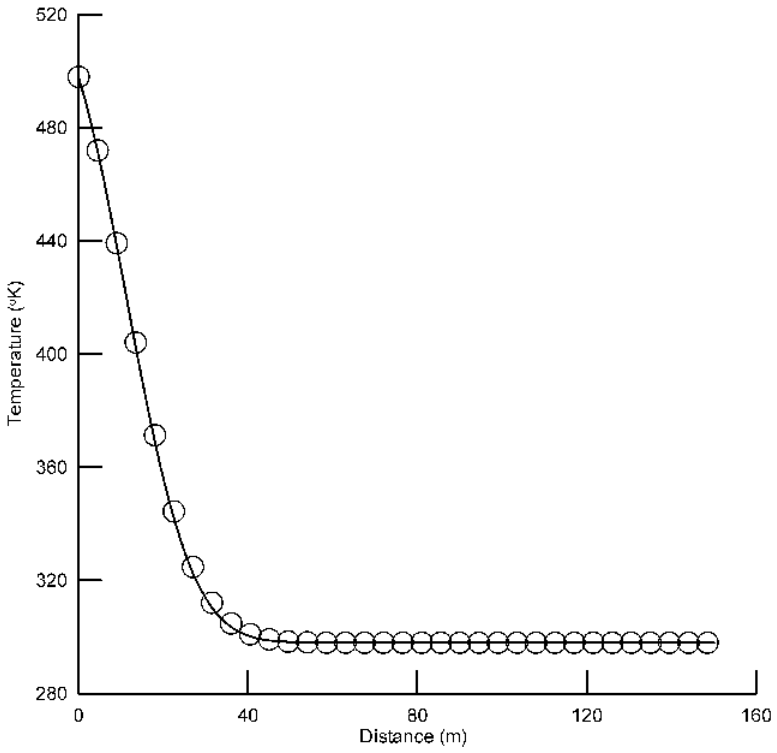


Figure 4.2 Temperature distribution after 8 days.

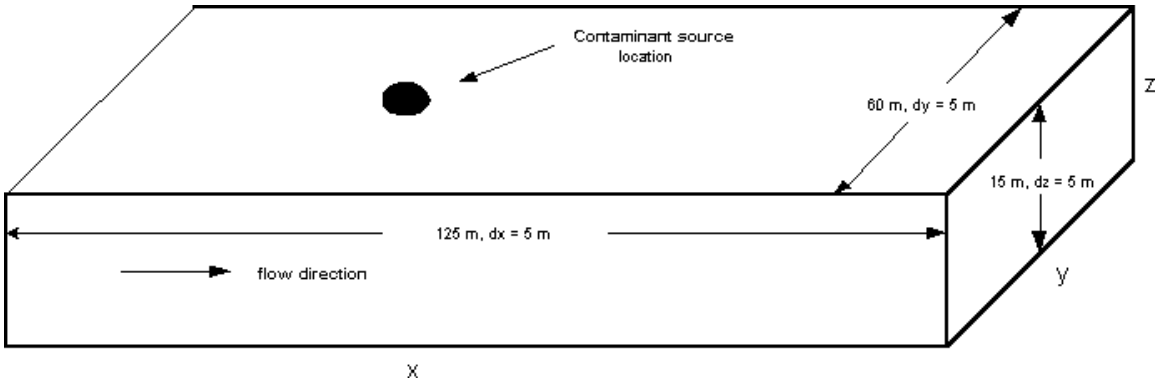
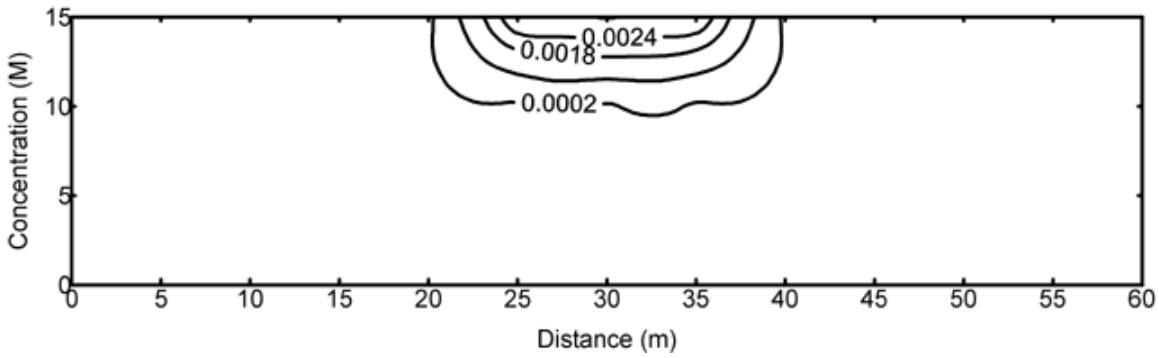


Figure 4.3 Problem domain (after UTCHEM, 2000)

(a)



(b)

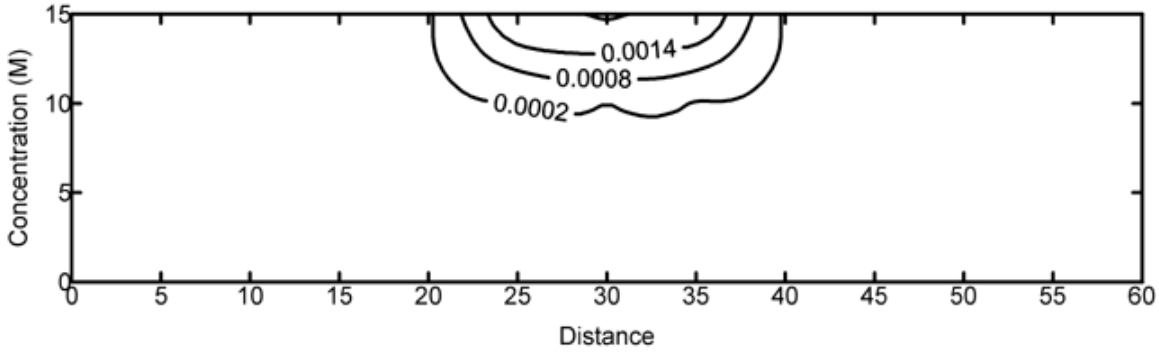
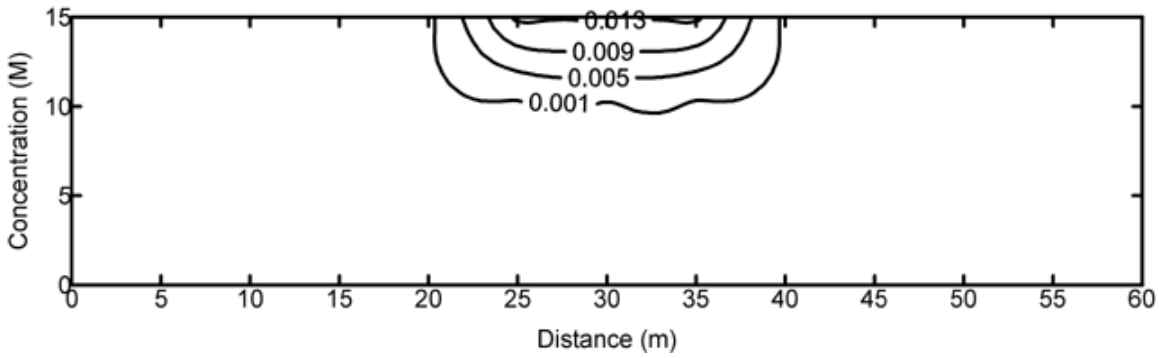


Figure 4.4 B_{NAPL} distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m

(a)



(b)

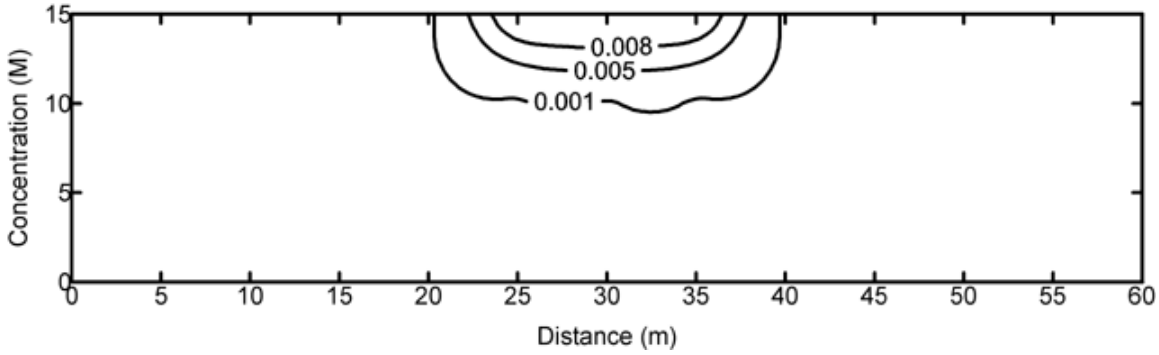
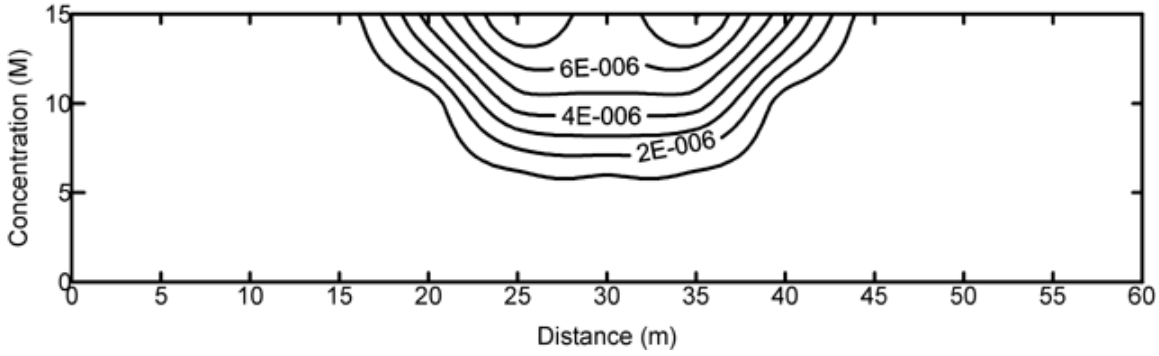


Figure 4.5 T_{NAPL} distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m

(a)



(b)

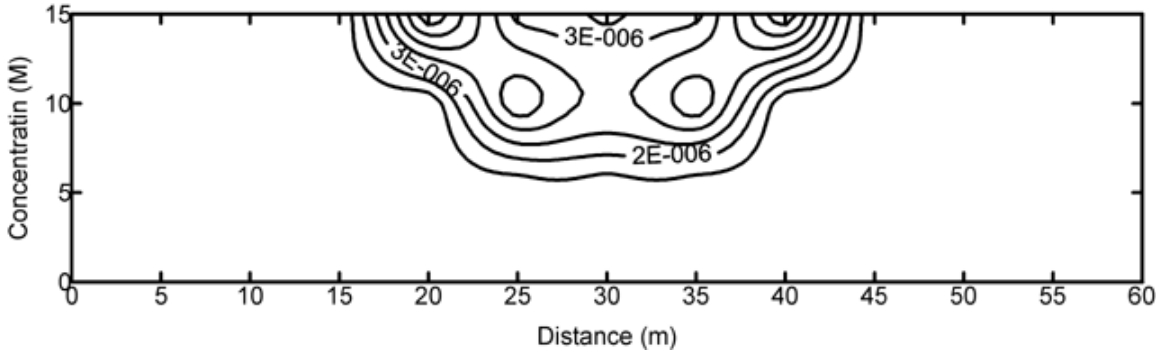
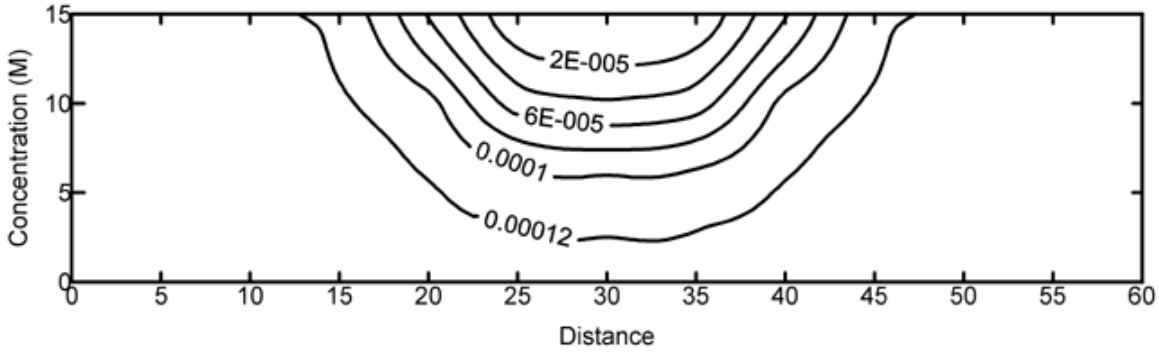


Figure 4.6 X_s distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m

(a)



(b)

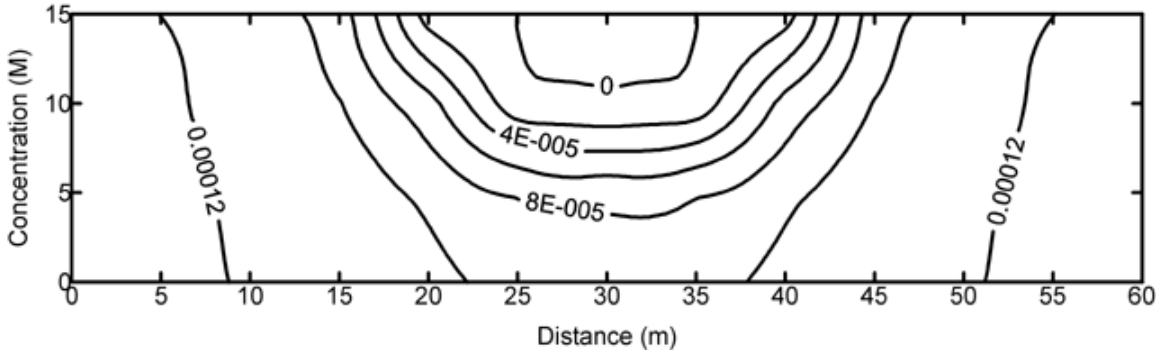
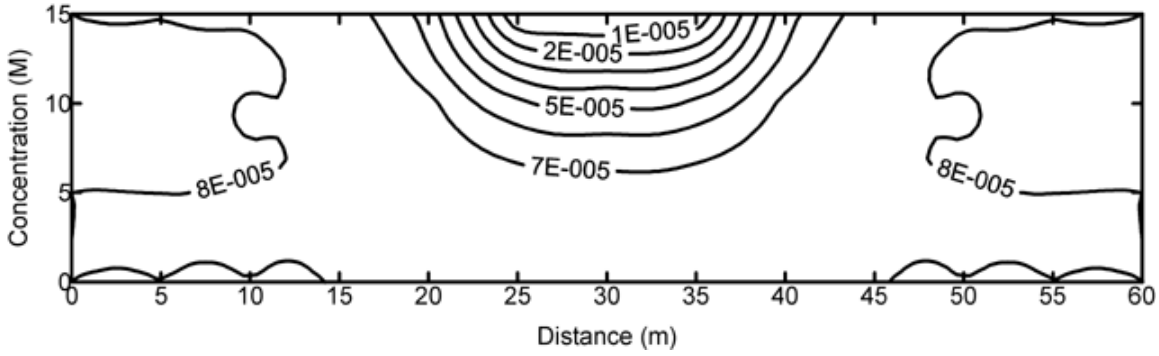


Figure 4.7 $O_{2(aq)}$ distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m

(a)



(b)

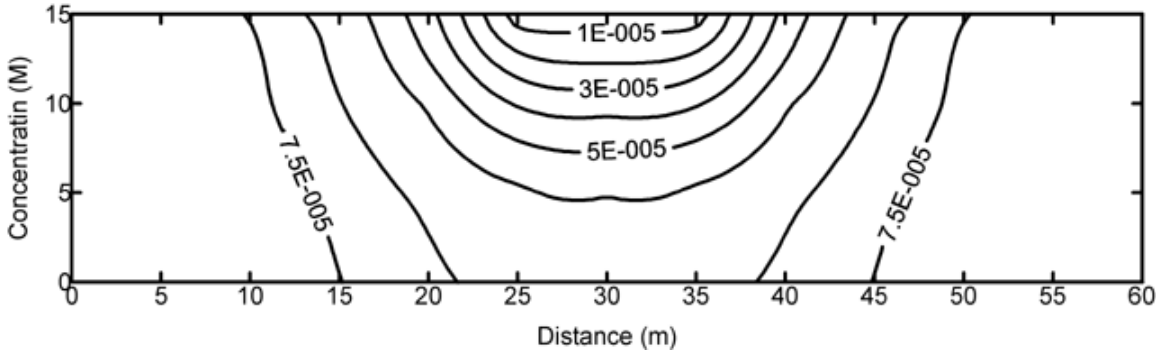


Figure 4.8 $NO_3^-{}_{(aq)}$ distribution after (a) t = 30d and (b) after t = 150 d at x = 90 m

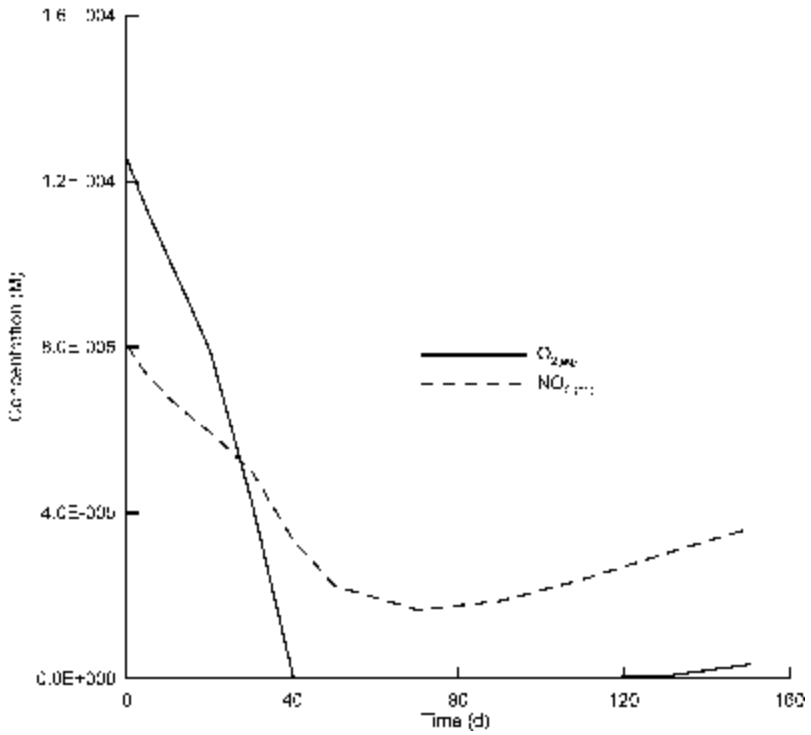


Figure 4.9 Species concentration distribution with respect to time

CHAPTER 5.
SUMMARY AND FUTURE WORK

5.1 SUMMARY

This thesis is made up of three computer codes, namely (1) BIOGEOCHEM, a numerical model to simulate chemical speciation in batch systems; (2) 2DHYDROBIOGEOCHEM, a 2 dimensional non-isothermal reactive chemical transport code under multiphase system; and (3) 3DHYDROBIOGEOCHEM, a 3 dimensional non-isothermal reactive chemical transport code under multiphase system. The formulations of these codes are based on the same framework, which employs decomposition-diagonalization procedure to derive governing equations to solve the physical and chemical processes in the system that are of concern. BIOGEOCHEM is thus easily combined with the hydrologic transport codes.

The unique aspects of the formulation in this work from those in the literature are that it can: (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.

Different numerical implementations incorporated in HYDROBIOGEOCHEM are tested and compared with each other. Users should be careful as to which numerical scheme to choose to better understand the complex processes in the subsurface system even though they are mathematically equivalent.

HYDROBIOGEOCHEM is a general reactive transport code. It is designed to be applicable to a wide range of physical and biogeochemical problems, including multiphase hydrological transport, heat transfer, and biogeochemical processes, such as aqueous complexation, adsorption-desorption, ion-exchange, oxidation-reduction, precipitation-dissolution, acid-base reactions, and microbial mediated reactions.

5.2 FUTURE WORK

HYDROBIOGEOCHEM is a hydrologic transport and heat transfer code only. It is not coupled with a multiphase flow module. It can read flow field from other flow simulators; however, it can not reflect the dynamic feedback effect of biogeochemical processes in the system on hydrological flow field. This dynamic feedback is important in the real-world system since biogeochemical processes will significantly change the magnitude and structure of local porosity and permeability (Glassley et al. 1998). This issue is addressed by a lot of researchers recently when applying the bioremediation technique (Taylor and Jaffee, 1990; cf., Glassley et al., 1998), because the change of local porosity and permeability in the system will affect the satisfactory efficacy of bioremediation.

Because of the heterogeneity of the porous media under the surface of the earth, to use the conventional hydrologic transport theory will not accurately predict the contaminant distribution in the subsurface system (Bai and Roegiers, 1997). The multi porosity (Bai et al., 1993; Bai and Roegers, 1997) and multi-pore-region (Gwo et al., 1995) theories are thus

conceptualized to better interpret the complex phenomenon in the heterogeneous system. HYDROBIOGEOCHEM can be expected to communicate with these models in the future.

HYDROBIOGEOCHEM is lack of field study experience. In order to be applicable to real-world systems, HYDROBIOGEOCHEM should be combined and verified with field experiments.

REFERENCES

Bai, M., Elsworth, D., and Roegers, J.-C., 1993. Multi-porosity/multi-permeability approach to the simulation of naturally fractured reservoirs. Water Resources Research, 29(6), 1621-1633.

Bai, M. and Roegers, J.-C., 1997. Triple-porosity analysis of solute transport. Journal of Contaminant transport, 28, 247-266.

Glassley, W.E., Bourcier, W.L., and Johnson, J.W., 1998. Review of existing reactive transport software. UCRL-ID-129846, Lawrence Livermore National Laboratory.

Gwo, J.P., Jardine, P.M., Wilson, G.V., and Yeh, G.T., 1995. A multi-pore-region concept to modeling mass transfer in subsurface media. Journal of Hydrology, 164, 217-237.

Taylor, S.W., and P.R. Jaffé, 1990. Biofilm Growth and the Related Changes in Physical Properties of a Porous Media. 1.- Experimental Investigation. *Water Resources Research*, 26(9), 2153-2159.

VITAE

Yilin Fang

Education

1999 – Current Ph.D. candidate, Department of Civil and Environmental Engineering, The Pennsylvania State University

1993 – 1998 D.E., Department of Mechanics, Hohai University, P.R.China

1989 – 1993 B.E., Department of Hydraulic Construction, Gezhouba Institute of Hydro-Electric Engineering

Experience

1999 – Current Graduate research assistant in the Department of Civil and Environmental Engineering, The Pennsylvania State University

1998 – 1999 Lecturer in the Department of Engineering Mechanics, Tongji University, P.R.China.

1995 – 1998 Graduate research assistant in the Department of Mechanics, Hohai University, P.R.China.

1993 – 1995 Graduate research assistant in the Department of Hydro-Electric Enginerring, Hohai University, P.R.China.

Selected Publications

Fang, Y., G.T. Yeh and W.D. Burgos, 2003. A General Paradigm to Model Reaction-Based Biogeochemical Processes, *Water Resources Res*earch, 39(4) (in press).

Burgos, W.D., R.A. Royer, Y. Fang, G.T. Yeh, A.S. Fisher, B.H. Jeon and B.A.Dempsey. 2002. Theoretical and Experimental Considerations Related to Reaction-Based Modeling: A Case Study Using Iron(III) Oxide Bioreduction.*Geomicrobiology Journal*. 19:253-287.

Burgos, W.D., Y. Fang, R.A. Royer, G.T. Yeh, J.J. Stone, B.H. Jeon and B.A. Dempsey. 2003. Reaction-Based Modeling of Quinone-Mediated Bacterial Iron(III) Reduction. *Geochimica Cosmochimica Acta* (in press).

Fang, Y. and G.T. Yeh. Numerical modeling of reactive chemical transport under multiphase systems, *XIV International* Conference on Computational Methods in Water Resources, *2002*, volume I: 671-678

Yeh, G.T. and Y. Fang. Numerical model of coupled fluid flow and reactive geochemical transport, *XIV International* Conference on Computational Methods in Water Resources, *2002*, Volume I: 607-614

Yeh, G.T., W.D. Burgos, Y. Fang and J.M. Zachara. Modeling biogeochemical kinetics: issues and data needs, presented at *XIII International Conference on computational Methods in Water Resources.*