



HBGC123D: a high-performance computer model of coupled hydrogeological and biogeochemical processes[☆]

Jin P. Gwo^{a,*}, Eduardo F. D'Azevedo^b, Hartmut Frenzel^c, Melanie Mayes^c, Gour-Tsyh Yeh^d, Philip M. Jardine^c, Karen M. Salvage^e, Forrest M. Hoffman^c

^aDepartment of Civil & Environmental Engineering, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA

^bComputer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6203, USA

^cEnvironmental Sciences Division, Oak Ridge National Laboratory,¹ Oak Ridge, TN 37831-6203, USA

^dDepartment of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816, USA

^eDepartment of Geological Sciences and Environmental Studies, State University of New York, P. O. Box 6000, Binghamton, NY 13902-6000, USA

Received 11 February 2000; received in revised form 27 September 2000; accepted 29 September 2000

Abstract

Groundwater flow and transport models have been used to assist management of subsurface water resources and water quality. The needs of more efficient use of technical and financial resources have recently motivated the development of more effective remediation techniques and complex models of coupled hydrogeological and biogeochemical processes. We present a high-performance computer model of the coupled processes, HBGC123D. The model uses a hybrid Eulerian–Lagrangian finite element method to solve the solute transport equation and a Newton's method to solve the system of nonlinear, mixed kinetics and equilibrium reaction equations. Application of the model to a laboratory soil column with multispecies tracer injection suggests that one may use the model to derive important parameters of subsurface solute fate and transport. These parameters may be used for predictive purpose in similar field problems. To this end, we present a three-dimensional, hypothetical bioremediation simulation on an aquifer contaminated by CoNTA. The simulation suggests that, using oxygen alone to stimulate the biodegradation of the contaminant, one may reduce the waste to 40% in 10 years. Using a refined mesh of this three-dimensional model, we also conduct a performance study of HBGC123D on an array of SGI Origin 2000 distributed shared-memory processors. Both the computational kernels and the entire model show very good performance up to 32 processors. The CPU time is essentially reduced by 20-fold using 64 processors. This result suggests that HBGC123D may be a useful tool in assisting environmental restoration efforts such as waste site characterization and remediation. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Hydrobiogeochemistry; Coupled processes; Bioremediation; Shared-memory parallel computing; Flow and transport

1. Introduction

Environmental concerns regarding the quantity and quality of drinking water sources in the subsurface had captured the attention of this nation long before the Love Canal incidence in the 1970s (ABC News Close Up: The Killing Ground, March 29, 1979). Over the next few decades, added to the list of concerns were issues such as over-pumping that resulted in lowering of groundwater table in the American high plain states

[☆]Code available from: <http://www.iamg.org/CGEditor.htm> or from <http://hbgc.esd.ornl.gov>.

*Corresponding author. Tel.: +1-410-455-3323; fax: +1-410-455-3559.

E-mail address: jgwo@umbc.edu (J.P. Gwo).

¹Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the US Department of Energy, under contract DE-AC05-00OR22725.

(Zwingle and Richardson, 1993), leaking of gasoline into groundwater from storage facilities (e.g., Jutras et al., 1997), salinization and contamination of groundwater due to irrigation and fertilization (e.g., Wopereis et al., 1998; Donoso et al., 1999), and so on. The sustainability and quality of groundwater resources over an expanding economy and growing population becomes not only an engineering and technical issue but also a societal issue of resource management and allocation (e.g., Wells drying up in New Jersey hills, *New York Times*, August 2, 1999). Groundwater flow and transport models have been used to assist management and decision-making in this respect. In their early years, these models mostly dealt with simple scenarios under assumptions such as homogeneous aquifers, a single representative chemical component, simple fluid–solid phase interactions, etc. The complex web of chemical species and complicated coupled processes of fluid flow, transport, and chemical interactions were largely sidestepped due to our limited computational power and scientific understanding of the complexity. Under these circumstances, decisions regarding environmental management, restoration and protection may become overly conservative and result in inefficient use of both technical and financial resources.

In the past two decades, complex reactive flow and transport models with varying degrees of sophistication have been developed (e.g., Grove and Wood, 1979; Valocchi et al., 1981; Jennings et al., 1982; Schulz and Reardon, 1983; Walsh et al., 1984; Cederberg, 1985; Kirkner et al., 1985; Lichtner, 1985; Bryant et al., 1986; Lewis et al., 1987; Hostetler et al., 1989; Liu and Narashimhan, 1989a, b; Yeh and Tripathi, 1991; Zyvoloski et al., 1992; Lichtner, 1992; Steefel and Lasaga, 1994; Steefel and Yabusaki, 1996; Lichtner and Seth, 1996). More recently, the emergence of complex groundwater reactive flow models has been largely driven by managerial needs of efficiently utilizing technical and financial resources in restoring the health of the subsurface environment (e.g., Cheng and Yeh, 1998; Clement et al., 1998; Hunter et al., 1998; Schafer et al., 1998; Salvage and Yeh, 1998; Tebes-Stevens et al., 1998; Yeh et al., 1998a; Chilakapati and Yabusaki, 1999; Chilakapati et al., 2000). It also coincides with a growing wealth of scientific understanding of subsurface coupled processes and new remediation techniques such as the application of surfactants (Knox et al., 1999), biodegradation of organics and metals (e.g., Jutras et al., 1997), electroremediation (e.g., Bannerjee et al., 1988), air sparging (e.g., Gierke et al., 1999), and so on. Common among these remediation technologies is the need for greater understanding of the coupled processes and the ability to identify and utilize key specific processes in the complex subsurface environment. Demands for complex hydrobiogeochemistry models that are able to quickly address design, monitoring, and implementation issues are therefore on the rise. Because

of the highly sophisticated nature of these models and the large CPU needs, high performance becomes a natural direction of their numerical implementation (e.g., Gwo and Yeh, 1996, 1997).

This paper presents a high-performance hydrobiogeochemistry model HBGC123D (HydroBioGeoChemistry in one, two, and three dimensions). The history of the model traces back to the transport model FEMWASTE developed at Oak Ridge National Laboratory, Tennessee (Yeh and Ward, 1981). The major processes that may be simulated by HBGC123D include (1) solute transport, (2) heat transfer, (3) mixed, heterogeneous, chemical kinetics and equilibrium, and (4) coupled reactive transport. HBGC123D is a Eulerian–Lagrangian finite element model that accommodates one-, two-, and three-dimensional problem domains. It is also a shared-memory parallel code that utilizes the OpenMP directives and may be run on shared-memory machines such as SGI Origin 2000, SGI multiprocessor Unix workstations, Cray PVPs and other machines on which OpenMP directives are available. A distributed memory version of the code is under development as of the writing of this manuscript. Serial version of the software is freely available at <http://hbgc.esd.ornl.gov> (HydroBioGeoChem 123D (HBGC123D)) where one may find the model documentation and many more benchmark applications other than those presented in this paper. Binary and partial source code for other platforms such as Win 32 are also available at the web site. The parallel version is undergoing rigorous field validation tests and will be made available at the same web site.

The objective of this research is to demonstrate the capability of HBGC123D, using laboratory soil column experiment data and hypothetical bioremediation problems. We also present and discuss the performance issues of HBGC123D on a distributed shared-memory machine, an SGI Origin 2000 at the National Center for Supercomputing Applications (NCSA).

2. Mathematical formulations

The HBGC123D model is characterized by a set of three governing equations: (1) subsurface flow, or the Richard's equation, (2) the advection–dispersion equation for solute transport and heat transfer, and (3) the mixed kinetics–equilibrium reaction equations that describe the rates of geochemical reactions and microbial biodegradation. These equations are solved by a Eulerian–Lagrangian finite element method for the flow and transport system and by a Newton's method for the biogeochemical subsystem. This section briefly describes the governing equations, the numerical methods used, and the high-performance implementation of the model on shared-memory parallel computers.

2.1. Flow and transport equations

The fluid flow and the solute transport systems in a hydrobiogeochemistry model are often represented by the following three mass balance equations (Yeh, 1987; Yeh et al., 1998a):

$$\frac{\partial \theta}{\partial t} = \nabla \cdot (\mathbf{K} \cdot \nabla(h+z)) + Q, \quad (1)$$

$$\begin{aligned} \frac{\partial \rho \theta T_j}{\partial t} + \nabla \cdot (\rho \theta C_j \mathbf{v}) - \nabla \cdot (\rho \theta \mathbf{D} \cdot \nabla C_j) \\ = R_j^c - A_j^c - \rho \theta T_j \left(\alpha \frac{\partial \rho}{\partial t} \right) + M_j^c, \end{aligned} \quad (2)$$

$$\frac{\partial W_j}{\partial t} = R_j^s - \Lambda_j^s - W_j \alpha \frac{\partial \rho}{\partial t} + M_j^s, \quad (3)$$

where t is the time (T), θ the water content, h the pressure head (L), z the elevation (L), \mathbf{v} the fluid velocity vector (L/T), Q the external fluid source/sink ($L^3/L^3/T$), \mathbf{K} the hydraulic conductivity tensor (L/T), ρ the fluid density (M/L^3), T_j the total equilibrium concentration of aqueous component j (M/M), W_j the total equilibrium concentration of absorbent component j (M/M), C_j the total dissolved concentration of aqueous component j (M/M), \mathbf{D} the dispersion tensor (L^2/T), R_j the production/consumption rate of component j ($M/L^3/T$), M_j the external source/sink of component j ($M/L^3/T$), A_j the decay rate of aqueous component j ($M/L^3/T$), p the fluid pressure (M/LT^2), and α the compressibility of the medium (LT^2/M). Superscripts c and s suggest affiliations with the aqueous and solid phases, respectively. Assuming that the fluid flow in the subsurface is laminar and isothermal, the momentum balance equations of fluid flow can be simplified to the well-known Darcy's equation (Bear, 1988)

$$\mathbf{v} = -\mathbf{K} \cdot \nabla(h+z). \quad (4)$$

Transport of kinetic species may also be represented by Eq. (2) by setting the T_j 's and C_j 's in Eq. (2) equal to the concentrations of individual species (Yeh et al., 1998a). Similarly, the heat transfer equation can be represented by Eq. (2) by substituting the T_j 's and C_j 's with temperature, substituting the dispersivity tensor with heat conductivity, including heat capacity in the first and second terms of Eq. (2), and disregarding the consumption/reduction and decay terms (Cheng and Yeh, 1998). Eqs. (1)–(4) implicitly suggest that the fluid flow and the solute/heat transport be treated as two de-coupled processes in HBGC123D. Effects of solute concentrations on fluid movement, e.g., density-dependent flow, are assumed insignificant, suggesting one of the limitations of the code. However, solute transport and heat transfer processes may be simulated as coupled processes in HBGC123D. One may or may not choose to use the van't Hoff relationship (e.g. see Bejan, 1988) to

relate kinetic rate coefficients and stability constants with temperature. Eqs. (1)–(4), with additional initial and boundary conditions (e.g., Yeh et al., 1998a), completely specify the flow and transport subsystem of a hydrobiogeochemistry model.

2.2. Mixed kinetics and equilibrium reactions

Rate expressions based on collision theory are used to handle kinetic geochemical reactions in HBGC123D. A chemical reaction can be generally represented by a stoichiometric equation of the following form (Yeh et al., 1998a):



where c_j and s_j are the reactants and products, respectively; ν_j and μ_j are the reactant and product stoichiometric coefficients, respectively; and M_x is the set of chemical species of interest. A generic rate equation corresponding to Eq. (5) may then be written as

$$\Omega = k_f \prod_{j \in M_x} (\gamma_j c_j)^{\nu_j} - k_b \prod_{j \in M_x} (\gamma_j s_j)^{\mu_j}, \quad (6)$$

where Ω , k_f , and k_b are the reaction rate, forward rate constant, and backward rate constant, respectively, and γ_j is the activity coefficient of species j according to the Davies formula (Davies, 1962; also see Yeh et al., 1998a for details). Homogeneous acid–base, reduction–oxidation and other complexation reactions in the aqueous phase, heterogeneous reactions including acid–base, reduction–oxidation, ion exchange, adsorption–desorption with simple surface complexation model, the constant capacitance model, or the triple layer model (e.g., Stumm and Morgan, 1981), and precipitation–dissolution may be simulated by HBGC123D. Detailed representation and discussion of these reactions may be found elsewhere (e.g., Yeh et al., 1998a, b; Salvage and Yeh, 1998). Change of media permeability as a result of heterogeneous biogeochemical reactions is currently not considered by HBGC123D.

Microbial growth rates are represented by a modified Monod kinetic equation in HBGC123D as follows (Yeh et al., 1998a):

$$\Omega = \Gamma I_g \left[\frac{S}{(k_s I_s + S)} \right] \left[\frac{A}{k_a + A} \right] \left[\frac{N}{k_n + N} \right] L, \quad (7)$$

where Γ is the microbial growth rate constant; K_s , K_a , and K_n are the half-saturation constants of substrate, electron acceptor, and nutrient, respectively; I_g and I_s are the inhibition factors for the growth rate and the half-saturation constant of the substrate, respectively; L is the metabolic lag coefficient; and S , A , and N are the concentrations of substrate, electron acceptor, and nutrient, respectively. Within HBGC123D, either the method of operator splitting (Chilakapati and

Yabusaki, 1999) or strong coupling of solute transport and biogeochemical reactions (Yeh and Tripathi, 1991) may be used to solve the coupled Eqs. (1)–(7). The latter formulation in HBGC123D is similar to the sequential iteration approach identified in Steefel and MacQuarrie (1996), which iterates between the transport and reaction equations and attempts to correct the errors that may occur in the classic operator-splitting approach.

2.3. Numerical methods

A hybrid Eulerian–Lagrangian finite element method is used to solve the heat transfer and solute transport equations. The system of nonlinear mixed kinetics and equilibrium reaction equations are solved by Newton's method. HBGC123D consists of four major computational modules: (1) heat transfer, (2) Lagrangian particle

tracking, (3) Eulerian finite element calculation, assembly, and solution of the system of transport equations, and (4) the biogeochemistry module. The computational flow of HBGC123D is shown in Fig. 1.

To solve the system of heat transfer and solute transport equations, HBGC123D uses line elements in one dimension, triangles and quadrilaterals in two dimensions, and hexahedral, pentahedral, and tetrahedral elements in three dimensions. First order linear basis functions are used in all cases. HBGC123D is flexible in accepting an unstructured mesh of conforming elements. The mesh can consist of a mixture of hexahedral, pentahedral and tetrahedral elements and need not be logically rectangular. The transport equation is solved either by upwind weighted finite elements or by backward Lagrangian particle tracking. The upwind option leads to assembly of nonsymmetric linear systems, whereas symmetry is preserved with the

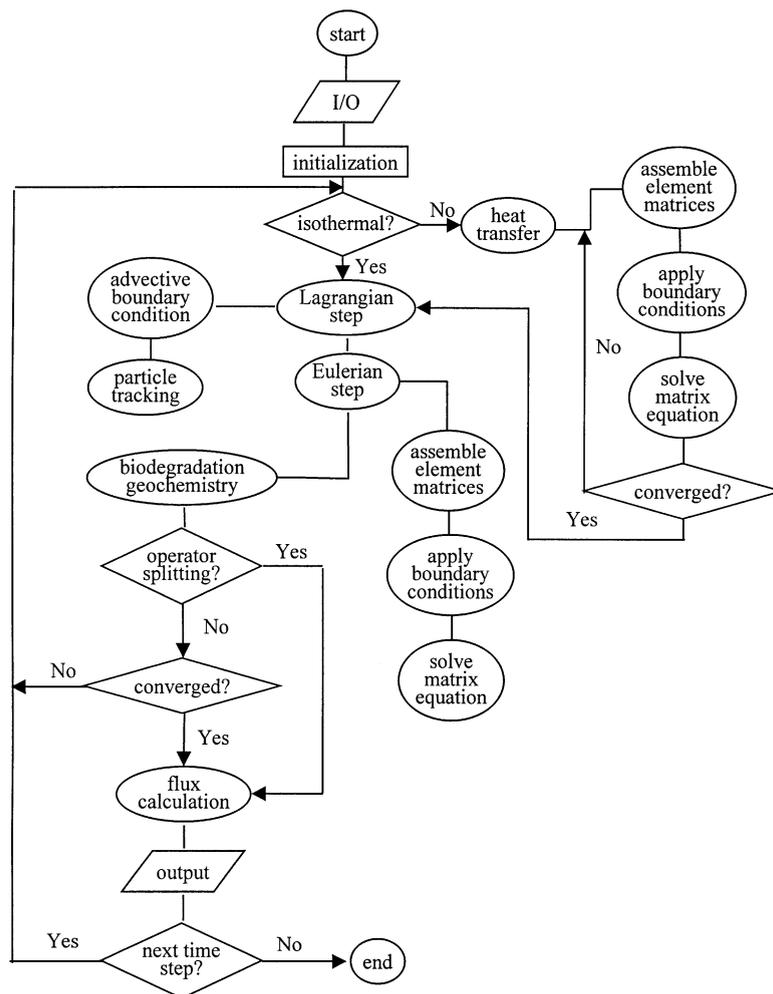


Fig. 1. Computational flow of HBGC123D.

Lagrangian particle tracking option. The Lagrangian option is appropriate for advection dominated transport and allows large time steps to be taken (Yeh et al., 1998a).

The chemistry module is based on earlier work by Yeh et al. (1998b), Salvage and Yeh (1998), and Yeh and Tripathi (1991) on hydrobiogeochemistry. The module assumes heat transfer and transport are already performed and computations are based solely on conditions at a specific node. The same chemistry module is used in all one-, two- and three-dimensional simulations. Mixed kinetic and equilibrium reactions are modeled as a system of non-linear algebraic equations. Newton's method with explicit Jacobian matrices is used to enhance convergence of the nonlinear iterations. The process might produce very ill-conditioned matrices and may require full row and column pivoting in LU decomposition.

2.4. Shared-memory parallel implementation

A parallel version of HBGC123D is available for shared memory multiprocessor systems such as the SGI Origin 2000. OpenMP (Dagum and Menon, 1998) compiler directives such as "parallel do" are added to most time consuming loops. Runtime profiles of HBGC123D highlight the following computational kernels, all of which can be easily and effectively parallelized:

- finite element assembly,
- solution of sparse linear systems by iterative methods,
- Lagrangian particle tracking,
- solution of systems of non-linear equations in the chemistry module.

In the finite element assembly process, each element produces a dense matrix (e.g., 8×8 for hexahedrons or 4×4 for tetrahedrons) that is scatter-summed into a global sparse matrix. To enable concurrent updates, the elements are first grouped (colored) into maximal independent sets. Two elements are dependent if they share at least one common node. Two dependent elements must be assigned distinct colors. Elements within an independent set have no nodes in common and their individual matrices may be assembled or scatter-summed into the global sparse matrix without any potential for conflict. For simplicity, the processing for source, sink and boundary conditions are performed in serial.

BiCGSTAB or conjugate gradients (CG) preconditioned with simple preconditioners (Barrett et al., 1994) such as diagonal scaling is available to solve the sparse linear system in parallel. Matrix vector multiply can be viewed as performing multiple independent dot product

operations and may easily fit in the OpenMP framework. Other options for linear solver such as band solver or GMRES iterative solver preconditioned by SSOR, ILU and matrix polynomials are not parallelized because of the sequential steps involved in the SSOR and ILU preconditioners.

Particle tracking in parallel can easily be implemented for the Lagrangian transport option. The trajectory is constructed by a piecewise linear path using an "element-by-element" algorithm. The initial flow velocity on entry to the element is used to predict a linear path through the element. The exit face and time of flight are computed using this simple model and tracking continues in the adjacent neighboring element until either the available time is exhausted or the particle exits the computational domain. HBGC123D precomputes a neighbor connectivity list that holds information about which elements share a common face to enable fast lookup of neighboring elements. The concentration of a fictitious particle is evaluated by interpolation or application of time-dependent boundary conditions. Due to numerical round-off, extra care is needed for handling trajectories near corner nodes and non-planar faces on hexahedra. Tetrahedral and pentahedral elements with planar faces may be used in place of hexahedra to avoid this problem. On distributed-memory machines, subregions of nodes or particles are assigned to individual processors, and a fictitious particle may access an element that does not belong to the subregion that a processor is responsible for. Because of the shared-memory nature of the OpenMP implementation in HBGC123D, this does not pose a serious problem. However, it may increase the memory access time on distributed shared-memory machines such as the SGI Origin 2000. For truly distributed memory machines such as the Cray T3E and the IBM SP, access of memory across processor boundaries will complicate implementation of particle tracking algorithms.

Depending on the complexity of reactions and number of species, the chemistry module is the most expensive computational kernel. At a specific node, a Newton step requires the construction of an $m \times m$ Jacobian matrix, where m depends on the number of species. LU factorization of the ill-conditioned Jacobian matrix requires $O(m^3)$ operations. Fortunately, this computation is highly parallelizable since the computations are performed after the transport calculation and each reaction depends solely on conditions at the node.

3. Benchmark applications

Two benchmark HBGC123D applications are used to demonstrate the capability of the computer model. The first application uses HBGC123D to obtain important model parameters from laboratory multispecies tracer

injections on undisturbed soil columns. Parameters thus obtained are crucial in reducing uncertainty of model predictions for field applications. The second application is a hypothetical bioremediation problem that uses oxygen supplied from two ditches surrounding a contaminated site to stimulate in situ biodegradation. This application demonstrates one possible remediation scenario that may be studied and examined using HBGC123D.

3.1. Laboratory soil column experiment

The Oak Ridge Reservation (ORR) in Oak Ridge, TN, is one of the several US Department of Energy facilities, where large quantities of low-level radioactive waste have been stored underground in previous decades. This waste may include organically complexed contaminants (co-contaminants) like CoEDTA. Investigations on the ORR have shown that co-contaminants may be transported into the groundwater from the waste burial ground (Means et al., 1978). In order to improve our understanding of the transport behavior of co-contaminants on the field-scale, several laboratory experiments were conducted (Mayes et al., 2000). These experiments were performed on undisturbed soil columns obtained from the ORR.

The first example application of HBGC123D demonstrates the simulation of one of these laboratory column experiments. The cylindrical soil column has a length of 14.5 cm and a radius of 4.2 cm. This is represented by a one-dimensional mesh with 60 elements and 61 nodes. There are seven chemical components and 13 chemical species, including the non-reactive Br^- , that are involved in six chemical reactions. These reactions, their kinetic rate coefficients and stability constants, are shown in Table 1. The experiment was carried out over a period of 2415.5 h. Co(II)EDTA^{2-} was injected and adsorbed to amorphous Mn(IV) oxides in the solid phase. Subsequently, the Co(II)EDTA^{2-} was oxidized to Co(III)EDTA^- . These reactions may result in the release of aqueous Mn^{2+} . Bromide was used as a non-reactive tracer in order to distinguish physical non-equilibrium processes from chemical reactions. To this end, two flow interrupts were performed in the course of the experiment. The first one started at 359.6 h and lasted until 527.6 h. The second one started at 1447.5 h and lasted until 1615.5 h. During these flow interrupts, no water or solute was injected. For the remainder of the experiment the inflow rate was 5 ml/h. This yields a Darcy velocity of 9.02×10^{-3} dm/h that is used directly as input of the model.

Fig. 2 shows the measured and the simulated effluent concentrations for Co(III)EDTA^- , Mn^{2+} , and Br^- . All simulation results match the experimental data fairly well. The experimental data for Br^- indicate that there is some physical non-equilibrium since there is a slight

Table 1
Geochemical reactions of laboratory soil column experiment

Stoichiometric equation ^a	Rate coefficient/ stability constant ^b
2Co(III)EDTA^- $-\text{Co(II)EDTA}^{2-} - 2\text{H}^+ + \text{Mn}_{(\text{aq})}^{2+} \Leftrightarrow$ $\text{Co(II)EDTA}^{2-} + 2\text{H}^+ + \text{MnO}_2$	$\log k_f = -25.36$
$\text{H}_2\text{O} \Leftrightarrow \text{H}^+ + \text{OH}^-$ $-\text{H}^+ + >\text{SOH} \Leftrightarrow >\text{SO}^-$	$\log k_b = 6.72$ $\log K = -14.00$ $\log k_f = -5.90$ $\log k_b = -2.00$
$\text{Co(III)EDTA}^- \Leftrightarrow \text{Co(III)EDTA}_{(\text{ads})}^-$	$\log k_f = -0.87$ $\log k_b = -1.96$
$\text{Co(II)EDTA}^{2-} \Leftrightarrow \text{Co(II)EDTA}_{(\text{ads})}^{2-}$	$\log k_f = 0.50$ $\log k_b = -0.38$
$\text{Mn}_{(\text{aq})}^{2+} \Leftrightarrow \text{Mn}_{(\text{ads})}^{2+}$	$\log k_f = 0.00$ $\log k_b = -2.00$

^a $>\text{SOH}$ represents solid surface group with associated OH^- molecule.

^b k_f = forward reaction rate coefficient; k_b = backward reaction rate coefficient; K = stability constant.

disturbance following the first flow interrupt. The physical non-equilibrium is likely caused by the existence of two distinct flow regions, a fracture region with fast flow and a matrix region with slower flow. HBGC123D is not capable of modeling fractured flow systems. The double-porosity medium is therefore represented by a single continuum model, and the flow interrupt effect cannot be seen in the simulation. Single continuum models have been used, at times, to represent dual-continuum media (e.g., Lee et al., 1992), with the assumption that the temporal and spatial scale may result in complete mixing of the two pore regions. Our single-continuum conceptualization here is primarily a result of the current capabilities of HBGC123D. Nevertheless, the simulation should provide valuable insights to the coupled hydrogeological and geochemical processes in the soil.

After the first flow interrupt, the measured Co(III)EDTA^- concentrations are lower than the simulated ones. This is probably due to passivation of the MnO_2 surfaces (Mayes et al., 2000). The simple, linear adsorption model (Table 1) is not adequate in representing the process and is a result of the lack of experimental information regarding the nature of the solid phase and, particularly, the oxide surfaces (Mayes et al., 2000). Interaction of Co(II)EDTA^{2-} with other surface sites may cause the dissociation of Co(II)EDTA^{2-} and the release of Co^{2+} (data not shown). This interaction, not accounted for in the simulation, may reduce the amount of Co(II)EDTA^{2-} available to react with MnO_2 . As a result, following the

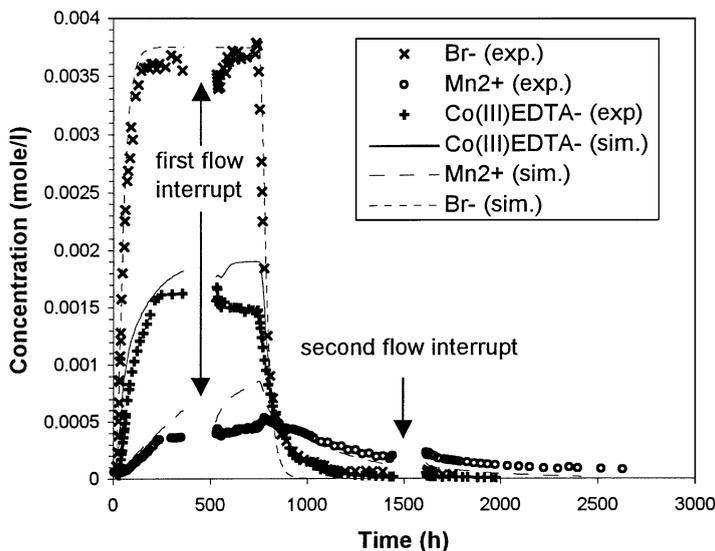


Fig. 2. Comparison of experimentally measured (“exp.”) and simulated (“sim.”) concentrations of Co(III)EDTA^- , Mn^{2+} , and Br^- for laboratory column example. Experimentally measured concentrations refer to effluent of laboratory column, simulated ones to last node of model domain.

flow interrupt, the simulated Mn^{2+} concentrations are higher than the measured.

3.2. A hypothetical bioremediation problem

This sample application demonstrates the use of HBGC123D for a three-dimensional problem. The steady-state flow problem was solved with 3DFEM-WATER (Yeh, 1987). The size of the domain is $100\text{ m} \times 40\text{ m} \times 7.2\text{ m}$ (Fig. 3). Initially, a region of high contaminant concentrations (Co^{2+} and CoNTA^-) extends from $x = 40$ to 65 m and $y = 20$ to 30 m . The initial microbe distribution is uniformly at $1.36 \times 10^{-4}\text{ mol/l}$. There is an extraction well at $x = 54\text{ m}$, $y = 0\text{ m}$. Pumping at the well results in water flowing into the problem domain from the two ditches at $x = 0$ and 100 m . This water contains dissolved oxygen that is necessary for the microbes that may degrade NTA aerobically.

The mesh comprises 2079 nodes and 1600 elements and there are 17 reactions and 24 species. The reactions include complexation, adsorption, and microbial biodegradation. The stoichiometric equations, rate coefficients and stability constants of the reactions are listed in Table 2. Solute transport is calculated for seven equilibrium components and two kinetic species. The simulation time was 10 years with 7300 time steps of 0.5d. Fig. 4 shows the concentrations of the contaminant (adsorbed CoNTA) and the microbes in a horizontal cross-section of the model domain after a simulation time of five years. It can be clearly seen that the regions where the fresh water (from the remediation

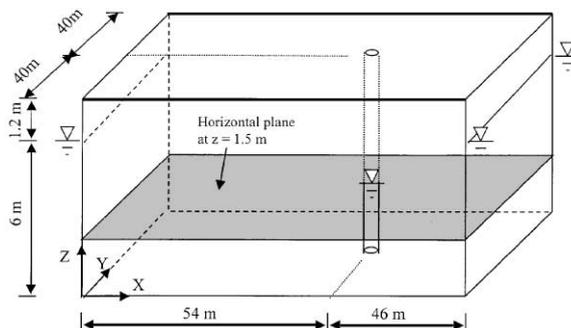


Fig. 3. Computational domain of hypothetical, three-dimensional bioremediation problem.

ditches) flows relatively faster exhibit much lower concentrations of the contaminant. The microbial cells are concentrated on the edge of the contaminated region because both the substrate (the contaminant) and the electron acceptor (oxygen) are available. The total mass of the contaminant in the model domain has been reduced to about 40% of the initial mass at the end of the simulation time of 10 years.

3.3. Parallel performance of HBGC123D

To test the parallel performance of HBGC123D, we refined the finite element mesh used in Section 3.2 from consisting of 2079 nodes and 1600 elements to 14,637 nodes and 12,800 elements. This refinement is necessary to test the performance of the computational kernels on

Table 2
Biogeochemical reactions of three-dimensional bioremediation problem

Stoichiometric equation	Rate coefficient/stability constant ^a
$3\text{H}^+ + \text{NTA}^{3-} \rightleftharpoons \text{H}_3\text{NTA}$	$\log K = 14.90$
$2\text{H}^+ + \text{NTA}^{3-} \rightleftharpoons \text{H}_2\text{NTA}^-$	$\log K = 13.30$
$\text{H}^+ + \text{NTA}^{3-} \rightleftharpoons \text{HNNTA}^{2-}$	$\log K = 10.30$
$\text{Co}^{2+} + \text{NTA}^{3-} \rightleftharpoons \text{CoNTA}^-$	$\log K = 11.70$
$\text{Co}^{2+} + 2\text{NTA}^{3-} \rightleftharpoons \text{Co}(\text{NTA})_2^{4-}$	$\log K = 14.50$
$\text{Co}^{2+} + \text{NTA}^{3-} + \text{H}_2\text{O} \rightleftharpoons \text{CoOHNTA}^{2-} + \text{H}^+$	$\log K = 0.50$
$\text{Co}^{2+} + \text{H}_2\text{O} \rightleftharpoons \text{CoOH}^+ + \text{H}^+$	$\log K = -9.70$
$\text{Co}^{2+} + 2\text{H}_2\text{O} \rightleftharpoons \text{Co}(\text{OH})_2 + \text{H}^+$	$\log K = -22.90$
$\text{Co}^{2+} + 3\text{H}_2\text{O} \rightleftharpoons \text{Co}(\text{OH})_3 + \text{H}^+$	$\log K = -31.50$
$\text{H}_2\text{CO}_3^*(\text{g}) \rightleftharpoons \text{HCO}_3^- + \text{H}^+$	$\log K = -6.35$
$\text{H}_2\text{CO}_3^*(\text{g}) \rightleftharpoons \text{CO}_3^{2-} + 2\text{H}^+$	$\log K = -16.68$
$\text{NH}_4^+ \rightleftharpoons \text{NH}_3 + \text{H}^+$	$\log K = -9.30$
$\text{H}_2\text{O} \rightleftharpoons \text{H}^+ + \text{OH}^-$	$\log K = -14.00$
$\text{H}^+ + \text{Buffer}^- \rightleftharpoons \text{Hbuffer}$	$\log K = 6.00$
$\text{CoNTA}_{(\text{aq})}^- \rightleftharpoons \text{CoNTA}_{(\text{ads})}^-$	$\log k_f = -0.70$
	$\log k_b = 0.00$
$\text{Co}_{(\text{aq})}^{2+} \rightleftharpoons \text{Co}_{(\text{ads})}^{2+}$	$\log k_f = -1.00$
	$\log k_b = -1.20$
$\text{HNNTA}^{2-} + 1.62\text{O}_2 + 1.272\text{H}_2\text{O} + 2.424\text{H}^+ \rightleftharpoons 17.3707\text{C}_5\text{H}_7\text{O}_2\text{N} + 3.12\text{H}_2\text{CO}_3^* + 0.424\text{NH}_4^+$	$\log \lambda = -2.682$
	$\mu_{\text{max}} = 0.092$
	$K_S = 1.64 \cdot 10^{-7}$
	$K_A = 6.25 \cdot 10^{-6}$

^a k_f = forward reaction rate coefficient; k_b = backward reaction rate coefficient; K = stability constant; λ = microbial death/decay constant; K_S = half-saturation constants for substrate; K_A = half-saturation constants for electron acceptor; μ_{max} = microbial growth rate constant.

a large array of processors. It essentially increases the total number of equations by approximately 7-fold and the finite element matrix calculations and assembly by 8-fold. Using this same configuration with 10 time steps of 0.5 d, we ran 7 HBGC123D simulations on the NCSA SGI Origin 2000 distributed shared-memory machines, with up to 64 threads or processors. The runs were submitted to the regular batch queue on the machines. Timing of four computational kernels, together with the total execution time, are shown in Fig. 5a. The speedup of the kernels and the entire code are plotted in Fig. 5b. Our test suggests that the performance difference between the OpenMP and the sequential, non-OpenMP codes on one processor is less than 3%. The speedup provided in Fig. 5b is therefore relative to the OpenMP one processor performance, or the parallel speedup. We use a diagonally preconditioned conjugate gradient matrix solver, which turns out to be very efficient and uses only a small fraction of the total computation time. The time of this kernel, however, slightly increases from 32 processors to 64 processors and its speedup is the worst among the four kernels. The other three computational kernels, Lagrangian particle tracking, element matrix calculation and assembly, and the biogeochemistry system of nonlinear equations, also show linear scaling behavior up to eight processors. The less ideal

scaling behavior of these kernels at larger number of processors may be caused by higher thread overhead and increasing memory access time. The SGI Origin 2000 is a distributed shared-memory machine on which memory access time may not be uniform when a relatively large number of processors are used. Overall the most expensive kernel is the biogeochemistry system of nonlinear equations. The kernel, however, has linear speedup up to 32 processors, suggesting the highly parallel nature of the biogeochemistry module. The total execution time decreases from 1270 s using one processor to 157 s on 8 processors and then to 59 s on 64 processors, or essentially from more than 20 min to under 1 min.

Various degrees of complexity may be encountered for industrial applications of hydrobiogeochemistry codes such as HBGC123D. These applications often have much larger study areas and therefore require many more finite elements. The data requirement may be more intensive and the conceptual models more complex. Efforts involving remediation scenario design and data collection may be extensive. Model calibration and prediction are often required to iterate with these efforts. Multiple, if not numerous, model simulations are not unusual and performance of models may become a critical issue in the entire remediation effort. As of the

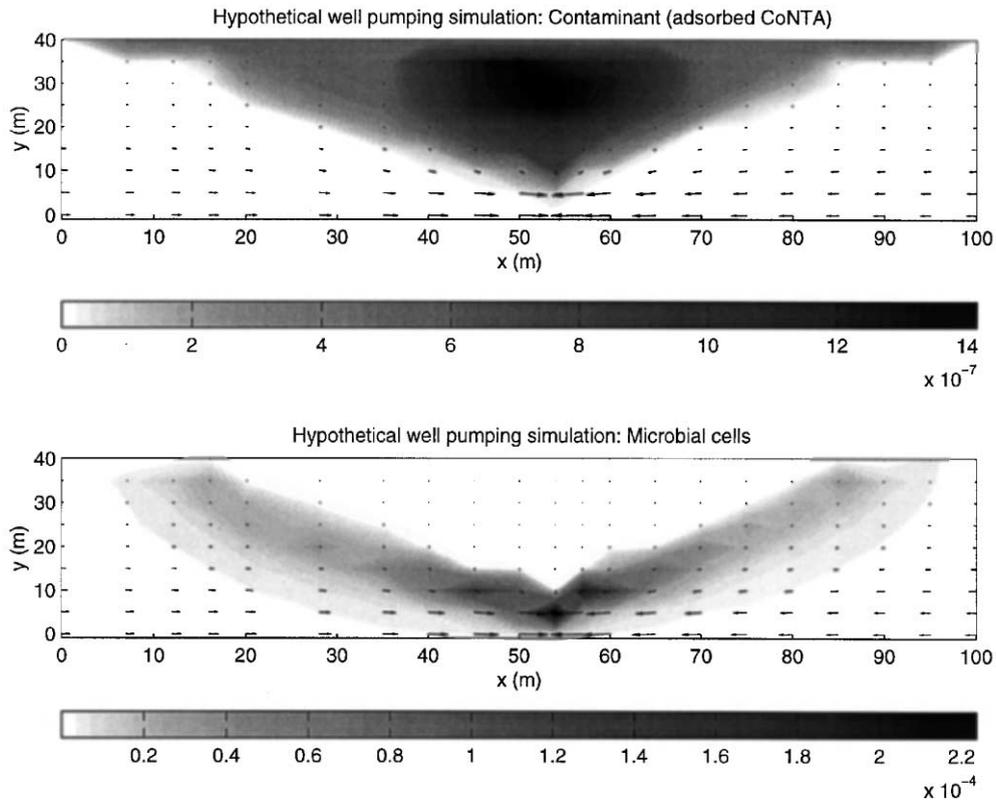


Fig. 4. Simulated concentrations for hypothetical well pumping example in horizontal cross-section with $z = 1.5$ m after simulation time of 5 years. Top: contaminant (adsorbed CoNTA), bottom: microbial cells. Units in figures are mol/l. Also shown are vectors representing horizontal velocities.

writing of this paper, HBGC123D is being used for two large-scale research and industrial bioremediation projects (NABIR Field Research Center Oak Ridge, Tennessee, Oak Ridge: DNAPL in Fractured Bedrock). Additional performance issues are being addressed as we are using HBGC123D to examine multiple remediation designs at these industrial sites. Nonetheless, the performance of HBGC123D may be problem dependent for industrial applications. Because the biogeochemistry kernel is the most expensive but is also highly parallel, applications that simulate many biogeochemical reactions and, therefore, require extensive biogeochemistry calculations but have a relatively small number of components and kinetic mobile species may see similar performance as those shown in this paper. However, for industrial applications for which biogeochemistry and hydrogeological transport calculations may need similar CPU times, it is likely that the performance of HBGC123D on shared-memory parallel computers will deteriorate with a relative large number of processors.

4. Summary and conclusion

This paper presents a high performance hydrobiogeochemistry model, HBGC123D, for multidimensional and multispecies solute fate and transport studies in the subsurface. HBGC123D uses a hybrid Eulerian–Lagrangian finite element method to solve the solute transport equations and a Newton’s method to solve the system of nonlinear, mixed kinetics and equilibrium reaction equations. Two benchmark applications are demonstrated. The application on a multispecies tracer injection suggests that HBGC123D may be used to simulate coupled hydrogeological and geochemical processes in laboratory soil columns. The suite of heterogeneous geochemical reactions and the parameters thus obtained may be used later for field subsurface remediation efforts. The application on a hypothetical, three-dimensional bioremediation problem demonstrated the capability of HBGC123D in assisting field bioremediation designs and scenario studies. The performance of HBGC123D on the SGI Origin 2000,

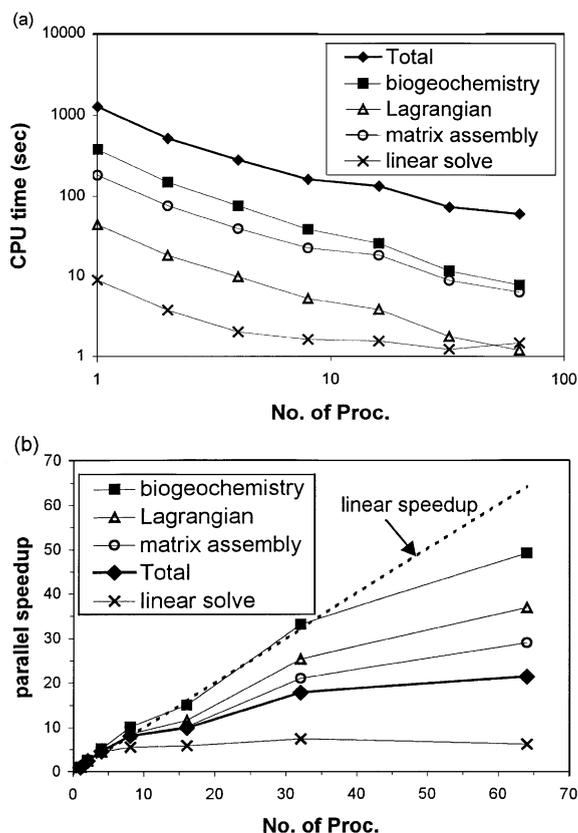


Fig. 5. Performance of HBGC123D on NCSA SGI Origin 2000 using up to 64 processors: (a) CPU times vs. number of processors, and (b) parallel speedup vs. number of processors.

albeit not ideal, further suggests that HBGC123D may be a useful tool to assist site characterization and environmental restoration efforts that require timely input from a simulation model. This input may be used to plan for more cost-effective data collection activities, to assist multi-objective cost-benefit analysis, and to identify effective, novel technology for in situ bioremediation.

Acknowledgements

This research is supported by the Environmental Technology Partnership (ETP) Program of the Office of Biological and Environmental Research, US Department of Energy. The authors would like to thank Dr. Paul Bayer, who is a contract officer for the DOE's ETP program, for financially supporting this research. The authors are also indebted to the National Center for Supercomputing Applications (NCSA) at University of Illinois, Urbana-Champaign, which provided computing

facilities for the timing exercise of HBGC123D. We are especially grateful to Dr. Kwai Wong of the Joint Institute for Computational Science at the University of Tennessee, who assisted our access to the computing facilities at NCSA. This research was also supported in part by an appointment to the Oak Ridge National Laboratory Postgraduate Research Program administered jointly by the Oak Ridge Institute for Science and Education and Oak Ridge National Laboratory. The authors are also grateful to the reviewers of the manuscript, who provided valuable input for the content and writing of the paper.

References

- Bannerjee, S., Horng, J., Ferguson, J.F., Nelso, P.O., 1988. Field-scale feasibility study of electro-kinetic remediation. CR 811762-01, US Environmental Protection Agency, Washington, DC.
- Barrett, R., Berry, M., Chan, T., Demmel, J., Donato, J., Dongarra, J., Eijkhour, V., Pozo, R., Romine, C., van der Vorst, H., 1994. Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods. SIAM, Philadelphia, PA, 112pp.
- Bear, J., 1988. Dynamics of Fluids in Porous Media. Dover, New York, 764pp.
- Bejan, A., 1988. Advanced Engineering Thermodynamics. Wiley, New York, 758pp.
- Bryant, S.L., Schechter, R.S., Lake, L.W., 1986. Interactions of precipitation/dissolution waves and ion exchange in flow through permeable media. AI Chemical Engineering Journal 32, 751–764.
- Cederberg, G.A., 1985. A groundwater mass transport and equilibrium chemistry model for multicomponent systems. Water Resources Research 21, 1095–1104.
- Cheng, H.P., Yeh, G.T., 1998. Development and demonstrative application of a 3-D numerical model of subsurface flow, heat transfer, and reactive chemical transport: 3DHYDRO-GEOCHEM. Journal of Contaminant Hydrology 34, 47–83.
- Chilakapati, A., Yabusaki, S., 1999. Nonlinear reactions and nonuniform flows. Water Resources Research 35, 2427–2438.
- Chilakapati, A., Yabusaki, S., Szecsody, J., MacEvoy, W., 2000. Groundwater flow, multicomponent transport and biogeochemistry: development and application of a coupled process model. Journal of Contaminant Hydrology 43, 303–325.
- Clement, T.P., Sun, Y., Hooker, B.S., Petersen, J.N., 1998. Modeling multispecies reactive transport in groundwater. Groundwater Monitoring Review 18, 79–92.
- Dagum, L., Menon, R., 1998. OpenMP: an industry standard API for shared-memory programming. IEEE Computational Science and Engineering 5, 46–55.
- Davies, C.W., 1962. Ion Association. Butterworths, Washington, DC, 190pp.
- Donoso, G., Cancino, J., Magri, A., 1999. Effects of agricultural activities on water pollution with nitrates and pesticides in the Central Valley of Chile. Water Science Technology 39, 49–60.

- Gierke, J.S., Christopher, L.W., Hutzler, N.J., 1999. Field test of air sparging coupled with soil vapor extraction. American Chemical Society Symposium 725, 153–166.
- Grove, D.B., Wood, W.W., 1979. Prediction and field verification of subsurface water quality changes during artificial recharge. Lubbock, Texas, Ground Water 17, 250–257.
- Gwo, J.P., Yeh, G.-Y., 1996. A parallel 3-dimensional HYDROGEOCHEM and an application to a proposed waste disposal site at the Oak Ridge National Laboratory, EOS Transactions. American Geophysics Union 77 (46), F203.
- Gwo, J.P., Yeh, G.-T., 1997. Application of a parallel 3-dimensional hydrogeochemistry HPF code to a proposed waste disposal site at the Oak Ridge National Laboratory. In: Tentner, A. (Ed.), Proceedings on High Performance Computing, Grand Challenges in Computer Simulation. Society for Computer Simulation International, Atlanta, pp. 128–133.
- Hostetler, C.J., Erikson, R.L., Fruchter, J.S., Kincaid, C.T., 1989. FASHEM package, Vol. 1, Overview and application to a chemical transport problem. Final Report for EPRI EA-5870, Project 2485-2, Pacific Northwest Laboratory, Richland, WA.
- Hunter, K.S., Wang, Y., Van Cappellen, P., 1998. Kinetic modeling of microbially-driven redox chemistry of subsurface environments: coupling transport, microbial metabolism and geochemistry. Journal of Hydrology 209, 53–80.
- HydroBioGeoChem 123D (HBGC123D). <http://hbgc.esd.ornl.gov>.
- Jennings, A.A., Kirkner, D.J., Theis, T.L., 1982. Multicomponent equilibrium chemistry in groundwater quality models. Water Resources Research 18, 1089–1096.
- Jutras, E.M., Smart, C.M., Rupert, R., Pepper, I.L., Miller, R.M., 1997. Field-scale biofiltration of gasoline vapors extracted from beneath a leaking underground storage tank. Biodegradation 8, 31–42.
- Kirkner, D.J., Jennings, A.A., Theis, T.L., 1985. Multicomponent mass transport with chemical interaction kinetics. Journal of Hydrology 76, 107–117.
- Knox, R. C., Shau, B. J., Sabatini, D. A., Harwe, J. H., 1999. Field demonstration studies of surfactant-enhanced solubilization and mobilization at Hill Air Force Base Utah. American Chemical Society Symposium 725, 49–63.
- Lee, R.R., Ketelle, R.H., Bownds, J.M., Rizk, T.A., 1992. Aquifer analysis and modeling in a fractured, heterogeneous medium. Ground Water 30, 589–597.
- Lewis, F.M., Voss, C.I., Rubin, J., 1987. Solute transport with equilibrium aqueous complexation and either sorption or ion exchange: simulation methodology and applications. Journal of Hydrology 90, 81–115.
- Lichtner, P.C., 1985. Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. Geochimica Cosmochimica Acta 49, 779–800.
- Lichtner, P.C., 1992. Time-space continuum description of fluid/rock interaction in permeable media. Water Resources Research 28, 3135–3155.
- Lichtner, P.C., Seth, M.S., 1996. User's manual for MULTIFLO: Part II. MULTIFLO 1.0 and GEM 1.0, multi-component-multiphase reactive transport model. CNWRA 96-010, Southwest Research Institute, Texas.
- Liu, C.W., Narasimhan, T.N., 1989a. Redox-controlled multiple-species reactive chemical transport, 1, model development. Water Resources Research 25, 869–882.
- Liu, C.W., Narasimhan, T.N., 1989b. Redox-controlled multiple-species reactive chemical transport, 2, Verification and application. Water Resources Research 25, 883–910.
- Mayes, M.A., Jardine, P.M., Larsen, I.L., Brooks, S.C., Fendorf, S.E., 2000. Multispecies contaminant transport in undisturbed columns of weathered, fractured saprolite. J. Contaminant Hydrol., in press.
- Means, J.L., Crerar, D.A., Borsik, M.P., 1978. Adsorption of Co and selected actinides by Mn and Fe oxides in soils and sediments. Geochimica et Cosmochimica Acta 42, 1763–1773.
- NABIR Field Research Center Oak Ridge, Tennessee. <http://www.esd.ornl.gov/BCV-FieldSite/>.
- Oak Ridge: DNAPL in Fractured Bedrock. <http://web.ead.anl.gov/techcon/projects/dnapl/>.
- Salvage, K.M., Yeh, G.T., 1998. Development and application of a numerical model of kinetic and equilibrium microbiological and geochemical reactions (BIOKEMOD). Journal of Hydrology 209, 27–52.
- Schafer, D., Schafer, W., Kinzelbach, W., 1998. Simulation of reactive processes related to biodegradation in aquifers: 1. Structure of the three-dimensional reactive transport model. Journal of Contaminant Hydrology 31, 167–186.
- Schulz, H.D., Reardon, E.J., 1983. A combined mixing cell/analytical model to describe two-dimensional reactive solute transport for unidirectional groundwater flow. Water Resources Research 19, 493–502.
- 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. American Journal of Science 294, 529–592.
- Steefel, C.I., MacQuarrie, K.T.B., 1996. Approaches to modeling of reactive transport in porous media. In: Lichtner, P.C., Steefel, C.I., Oelkers, E. H. (Eds.), Reactive Transport in Porous Media, Reviews in Mineralogy, Vol. 34, pp. 83–129.
- Steefel, C.I., Yabusaki, S.B., 1996. OS3D/GIMRT, software for multicomponent-multidimensional reactive transport, user manual and programmer's guide. PNL-11166, Pacific Northwest National Laboratory, Richland, WA.
- Stumm, M.W., Morgan, J.J., 1981. Aquatic Chemistry: An Introduction Emphasizing Chemical Equilibria in Natural Waters. Wiley, New York, 780pp.
- Tebes-Stevens, C., Valocchi, A.J., VanBriesen, J.M., Rittman, B.E., 1998. Multicomponent transport with coupled geochemical and microbiological reactions: model description and example simulations. Journal of Hydrology 209, 1–7.
- The National Center for Supercomputing Applications (NCSA). <http://www.ncsa.uiuc.edu>.
- Valocchi, A.J., Street, R.L., Roberts, P.V., 1981. Transport of ion-exchanging solutes in groundwater: chromatographic theory and field simulation. Water Resources Research 17, 1517–1527.
- Walsh, M.P., Bryant, S.L., Lake, L.W., 1984. Precipitation and dissolution of solids attending flow through porous media. AI Chemical Engineering Journal 30, 317–328.

- Wopereis, M.C.S., Ceuppens, J., Boivin, P., Ndiaye, A.M., Kane, A., 1998a. Preserving soil quality under irrigation in the Senegal River Valley. *Netherlands Journal of Agricultural Science* 46, 97–107.
- Yeh, G.T., 1987. FEMWATER: a finite element model of WATER flow through saturated–unsaturated porous media—First Revision. ORNL-5567/R1, Oak Ridge National Laboratory, Tennessee.
- Yeh, G.T., Iskra, G., Zachara, J.M., Szecsody, J.E., 1998. Development and verification of a mixed chemical kinetic and equilibrium model. *Advances in Environmental Research* 2, 24–56.
- Yeh, G.T., Salvage, K.M., Gwo, J.P., Zachara, J.M., Szecsody, J.E., 1998a. HYDROBIOGEOCHEM: a coupled model of HYDROlogic transport and mixed BIOGEOCHEMical kinetic/equilibrium reactions in saturated–unsaturated media. ORNL/TM-13668, Oak Ridge National Laboratory, Tennessee.
- 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., Ward, D.S., 1981. FEMWASTE: a finite element model of WASTE transport through saturated–unsaturated porous media. ORNL-5601, Oak Ridge National Laboratory, Tennessee.
- Zwingle, E., Richardson, J., 1993. Ogallala aquifer: wellspring of the high plains. *National Geographic Magazine*, March 1993, pp. 80–109.
- Zyvoloski, G., Dash, Z., Kelkar, S., 1992. FEHMN 1.0: finite element heat and mass transfer code. LA-12062-MS, Rev 1., Los Alamos National Laboratory, New Mexico.