



**US Army Corps
of Engineers®**
Engineer Research and
Development Center

System-Wide Water Resources Program

A Generic Reaction-Based BioGeoChemical Simulator (RBBGCS), Version 1.0

Hwai-Ping Cheng, Stacy E. Howington, Matthew W. Farthing,
Christian J. McGrath, and Jing-Ru C. Cheng

May 2010

A Generic Reaction-Based BioGeoChemical Simulator (RBBGCS), Version 1.0

Hwai-Ping Cheng, Stacy E. Howington, and Matthew W. Farthing

*Coastal and Hydraulics Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199*

Christian J. McGrath

*Environmental Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199*

Jing-Ru C. Cheng

*Information Technology Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199*

Final report

Approved for public release; distribution is unlimited.

Abstract: This report presents a generic reaction-based biogeochemical simulator (RBBGCS) that was developed as part of the advancement of the subsurface reactive transport capability in the Adaptive Hydrology/Hydraulics (ADH) model. RBBGCS has been incorporated into ADH to model subsurface reaction transport. The simulator can also be coupled with other transport models to perform reactive transport modeling in surface and subsurface systems. RBBGCS can model geochemical/biogeochemical reactions that are equilibrium-controlled (fast reversible), instantaneous (fast irreversible), and kinetic (slow reversible or irreversible). It has a preprocessor that automatically and systematically produces reaction-based differential-algebraic equations (DAE) as the reaction governing equations and a solver that solves the set of governing equations for the concentration distribution of chemical species. It allows both user-specified empirical equation and formulation based on the collision theory to be used to describe reaction equilibrium and reaction rate(s). The numbers of chemical species, biogeochemical reactions, and porous medium phases that may be defined for a modeled system are unrestricted, limited only by the computational resources that are available.

This report describes the development of RBBGCS, including a nine-step preprocessor to generate reaction-based DAE systems and solution techniques to solve the DAE system. The preprocessor constructs a valid reaction network and produces the associated governing equations, which can save modelers a significant amount of time when modeling complex reaction systems. The solution technique section details (1) the computational procedures in RBBGCS, (2) the DAE system when man-induced sources exist, (3) Newton's method to solve DAE systems, (4) implementation of the constraint equations in DAE systems, and (5) treatment for zero-order reactions. Multiple test examples are presented to verify and demonstrate RBBGCS' capabilities in solving complex geochemical/biogeochemical reaction problems. RBBGCS development serves as a guide for continued model development for coupling with ADH and other transport models to perform reactive transport simulation.

DISCLAIMER: The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products. All product names and trademarks cited are the property of their respective owners. The findings of this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

DESTROY THIS REPORT WHEN NO LONGER NEEDED. DO NOT RETURN IT TO THE ORIGINATOR.

Contents

Contents.....	iii
Figures and Tables.....	v
Preface.....	vii
1 Introduction.....	1
2 A Generic Biogeochemical Simulator.....	4
2.1 Governing equations of biogeochemical reaction systems.....	4
2.2 Nine-step reaction-based DAE preprocessor	9
Step 1. Detect and remove redundant equilibrium-controlled reactions	11
Step 2. Detect and remove conflicting instantaneous reactions	13
Step 3. Examine consistency and conduct sorting for instantaneous reactions	14
Step 4. Detect and remove irrelevant kinetic reactions	16
Step 5. Determine the number of linearly independent reactions.....	17
Step 6. Rearrange the reaction order.....	17
Step 7. Decompose the reaction network from step 6.....	18
Step 8. Express the decomposed reaction network in equation form.....	19
Step 9. Identify secondary species for effective biogeochemical modeling.....	22
2.3 Solution techniques in RBBGCS	26
2.3.1 Computational procedures	26
2.3.2 The DAE system when sources exist.....	27
2.3.3 Newton's method.....	28
2.3.4 Implementation of constraint equations.....	30
2.3.5 Treatment for zero-order reactions	32
2.3.6 Accounting for species with fixed concentrations.....	34
3 Verification and Demonstration Examples.....	36
3.1 Verification of preprocessing.....	36
Problem 1. 7-species system.....	36
Problem 2. Mixed microbiological and abiotic reactions	41
Problem 3. Complexation, adsorption, ion-exchange, and dissolution in a system of mixed equilibrium-controlled and kinetic reactions	42
Problem 4. Sorting instantaneous reaction systems	42
3.2 Verification of solution technique	47
3.3 Verification of using constraint equations for instantaneous reactions.....	54
Problem 5. One instantaneous reaction system	54
Problem 6. Correlated instantaneous reaction system	56
3.4 Verification of using the LR method for zero-order reactions.....	58
Problem 7. Zero-order reaction system	58

3.5 Capability of solving reaction systems of mixed types	60
<i>Problem 8. Mixed-type reaction system (1)</i>	60
<i>Problem 9. Hypothetical PCE reaction system</i>	62
<i>Problem 10. Hypothetical organic waste system</i>	65
4 Summary	71
References	72
Appendix A: RBBGCS Input Guide	73
Report Documentation Page	

Figures and Tables

Figures

Figure 1. An algorithm to check consistency and conduct sorting for instantaneous reactions.	15
Figure 2. Flow chart of the computation in RBBGCS.	27
Figure 3. Flow chart of solving nonlinear DAE systems with the Newton's method	30
Figure 4. Flow chart of finding and using the correct constraint equation combination within a time-step in RBBGCS	32
Figure 5. Relationship between zero-order reaction rate and reactant concentration in the linear regularization (LR) method	34
Figure 6. Variation in the computed concentration for selected species comparing simulations associated with the two alternative choices of component species	54
Figure 7. Numerical solution of Reaction system 5 when different time-step sizes are used.	55
Figure 8. Numerical solutions of Reaction system 5 from FIR and SR	56
Figure 9. Numerical solutions of Reaction system 5 from FIR and SR	56
Figure 10. The computed concentration time series of Reaction system 6 using constraint equations for instantaneous reactions	57
Figure 11. The computed concentration profiles of Reaction system 7 between using the LR and the switching-based methods for the zero-order reaction $A + 2B \rightarrow C + 2D$	59
Figure 12. The computed concentration profiles of Reaction system 7 between using 0.1 and using 2 as the time-step size for simulation	59
Figure 13. The computed concentration trends for P and Q in Reaction system 8 when fast reactions are simulated in different ways	61
Figure 14. Impact of time-step size on the computed concentration profiles of N, P, and Q in Reaction system 8 when fast reactions are simulated in different ways	63
Figure 15. The computed concentration profiles of the 18 species in Reaction system 9 when the time-step sizes are set to 120, 720, and 1,800 hours	65
Figure 16. The hypothetical reaction system of Test Problem 10 (left), and adsorption of dissolved organic waste and biodegradation residual onto sorption sites S1= and S2= from both water column and interstitial water (right).	66
Figure 17. The concentration evolution of species in the aqueous phases: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.1 mole/dm ² for both S1= and S2=.	68
Figure 18. The concentration evolution of species associated with sorption site 1: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.1 mole/dm ² for both S1= and S2=.	68
Figure 19. The concentration evolution of species associated with sorption site 2: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.1 mole/dm ² for both S1= and S2=.	68

Figure 20. The concentration evolution of species in the aqueous phases: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.001 and 0.002 mole/dm ² for S1= and S2=, respectively.....	69
Figure 21. The concentration evolution of species associated with sorption site 1: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.001 and 0.002 mole/dm ² for S1= and S2=, respectively.....	69
Figure 22. The concentration evolution of species associated with sorption site 2: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.001 and 0.002 mole/dm ² for S1= and S2=, respectively.....	70

Tables

Table 1. Reaction system for Problem 1.	36
Table 2. Summary of Problem 1 Decomposition Results.....	37
Table 3. Reaction system for Problem 2.	41
Table 4. Summary of Problem 2 Decomposition Results.....	42
Table 5. Reaction system for Problem 3.	43
Table 6. Summary of Problem 3 Decomposition Results.....	44
Table 7. Reaction system for Problem 4.....	44
Table 8. Reaction system for Problem 5	54
Table 9. Reaction system for Problem 6.	57
Table 10. Reaction system for Problem 7.....	58
Table 11. Reaction system for Problem 8.....	60
Table 12. Definition of various treatments used to simulate reactions in Problem 8	60
Table 13. Reaction system for Problem 9.....	63
Table 14. Reaction system for Problem 10.....	66

Preface

This report summarizes initial efforts undertaken in developing a generic reaction-based biogeochemical simulator (RBBGCS) that was developed as part of the advancement of the subsurface reactive transport capability in the Adaptive Hydrology/Hydraulics (ADH) model. This development effort was performed by the U.S. Army Engineer Research and Development Center (ERDC), Vicksburg, MS. Funding was provided under the System-Wide Water Resources Program (SWWRP). Appreciation is extended to all those who assisted in the development and review of this report.

Principal investigators for this study were Dr. Hwai-Ping Cheng, Dr. Stacy E. Howington, and Dr. Matthew W. Farthing, Hydrologic Systems Branch, Coastal and Hydraulics Laboratory (CHL); Christian J. McGrath, Environmental Processes Branch, Environmental Laboratory (EL); and Dr. Jing-Ru C. Cheng, DoD Supercomputing Resource Center (DSRC), Information Technology Laboratory (ITL), ERDC. Dr. Cheng, Dr. Howington, and Dr. Farthing conducted their portion of the study under the general supervision of Earl V. Edris, Chief, Hydrologic Systems Branch, CHL; Bruce A. Ebersole, Chief, Flood and Storm Protection Division, CHL; and Dr. William D. Martin, Director of CHL.

Dr. Carlos Ruiz, EL, and Dr. Gaurav Savant, CHL, reviewed this report and provided valued comments.

COL Gary E. Johnston was Commander and Executive Director of ERDC. Dr. Jeffery P. Holland was Director.

This report should be cited as follows:

Cheng, H-P., S. E. Howington, M. W. Farthing, C. J. McGrath, and J-R. C. Cheng. 2010. *System-wide water resources program: A Generic Reaction-Based Biogeochemical Simulator (RBBGCS), Version 1.0*. ERDC TR-10-5. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

1 Introduction

Computer simulation of flow and transport is an essential component in the rigorous analysis of water supply, contamination, environmental cleanup, and ecosystem restoration. Simulation of coupled equations of transport and geochemical/biogeochemical reactions (for convenience, “biogeochemical reactions” is used to represent both geochemical and biogeochemical reactions from this point on) based on the principle of mass conservation is the only quantitative approach for integrating multiple, complex, environmental processes into an internally consistent conceptual model with which to assess water quality and to design engineered solutions for remedial alternatives. A reactive transport model in a more general sense treats a multi-component, multi-species system in which a number of equilibrium-controlled and perhaps kinetic and instantaneous reactions occur simultaneously. The naturally complicated reactive transport problem requires more specialized solution techniques, including the best known and most commonly used operator-splitting or other numerical methods (Yeh et al 2001).

No matter what solution technique is used in solving reactive transport equations, it is vital to model biogeochemical reactions and solve for concentration distribution among species in accurate and efficient means. Bethke (2008) discussed the conceptual model, mathematical formulations, and numerical solution for reaction processes and systems of various kinds. His discussion demonstrates the capabilities of a reaction model to determine the model’s applicability which, as a result, influences the performance of reactive transport modeling when the reaction model is integrated with mass transport.

Biogeochemical reactions have been modeled frequently using an *ad hoc* approach, in which production/decay rates are empirical functions of the concentrations of species (McNaught and Wilkinson 1997) with rate parameters fit to experimental data (Neuhaus and Maier 1996). Such approaches may provide useful monitoring and management tools when calibrated to specific site conditions, but extension to different conditions requires complete re-calibration. With a better understanding of complex biogeochemical reactions and their mathematical formulations (Chilakapati et al. 1998; Yeh et al. 2001), mechanistic reaction models have a potential for broader applicability (Steefel and Cappellen 1998).

The Adaptive Hydrology/Hydraulics (ADH, <https://adh.usace.army.mil/>) model is a modular, parallel, adaptive finite-element model for multi-dimensional flow and transport. It simulates variably-saturated groundwater flow in porous media, overland flow, flow through hydraulic structures, and flow in riverine and estuarine systems. ADH was developed at the Engineer Research and Development Center through multiple collaborations. The System-Wide Water Resources Program (SWWRP, <https://swwrp.usace.army.mil/DesktopDefault.aspx>) funded the enhancement of the subsurface reactive transport capability in ADH. In this project, a generic reaction-based biogeochemical simulator (RBBGCS) was newly developed and incorporated into the existing ADH subsurface transport module to perform subsurface reactive transport simulation.

RBBGCS was designed to be general enough to account for a wide spectrum of biogeochemical reactions, e.g., biodegradation, acid-base reactions, aqueous complexation, adsorption-desorption, precipitation-dissolution, ion-exchange, oxidation-reduction, partitioning, and volatilization. While some reactions are slow and the others are fast relative to the temporal components of the model (Rubin 1983), e.g., modeling time-step and the residence time within a model element/cell, RBBGCS can model reaction systems that are combinations of equilibrium-controlled, instantaneous, and kinetic reactions. RBBGCS allows fast reversible reactions to be simulated as equilibrium-controlled reactions and characterized using mass action equations or user-specified algebraic equations. It also allows a reaction rate to be represented as a function of the computed species concentrations and given system parameters. These features give modelers great flexibility in evaluating alternative reaction mechanisms, constructing reaction networks, and conducting numerical simulations for batch reaction systems.

The governing equations of RBBGCS that can take into account mixed fast and slow reaction system are formulated in a reaction-based differential-algebraic equation (DAE) form (Chilakapati et al. 1998). They are produced automatically and systematically by a preprocessor, which is a modified version of Fang et al. (2003), to also account for instantaneous reactions. In this DAE formulation, fast reversible reactions are represented by algebraic equations that define reaction equilibrium, and fast irreversible reactions, by zero-concentration constraint equations that characterize instantaneous reactions. As a result, the requirement of using small time steps to resolve fast reactions is relieved.

Because ADH will require major modifications when it incorporates the complete suite of RBBGCS to perform reactive transport, a two-phase approach was taken. In Phase 1, all biogeochemical reactions are expressed as rate-controlled in ADH reactive transport, where large rate constants are used for fast reactions, and one mobile aqueous phase and one immobile solid phase are included (completed in 2009). The Phase 1 product is ADH reactive transport Version 1.0, which is sufficient for modeling many subsurface contamination problems (e.g., PCE/TCE, RDX, nutrient cycles, etc.) associated with DoD military facilities/sites. In Phase 2, RBBGCS will be fully implemented in the ADH reactive transport module to account for multiple species in multiple phases (e.g., aqueous, solid, and gas), phase mobility (i.e., a phase can be either mobile or immobile), and different reaction types (i.e., equilibrium-controlled, instantaneous, and kinetic) without using small time steps to resolve fast reactions.

Numerical aspects of RBBGCS are described in Chapter 2, including the governing equation, the preprocessor, and the solution techniques employed to solve the DAE form of governing equations generated by the preprocessor. In Chapter 3, RBBGCS' capabilities are demonstrated and verified with multiple test examples. Final remarks on the development of RBBGCS and an outline of tasks for future advancements are offered in Chapter 4. The RBBGCS input guide is given in Appendix A.

2 A Generic Biogeochemical Simulator

In modeling biogeochemical reactions, each reaction can be considered fast or slow regarding the modeling time step. In theory, each reaction is a kinetic reaction of which the reaction rate is non-zero when it occurs. Given a fixed system condition, a fast reversible reaction can be treated as an equilibrium-controlled reaction because it reaches equilibrium so quickly. Likewise, a fast irreversible reaction can be treated as an instantaneous reaction because it reaches its end state within a short time. The argument of what is the rate for an equilibrium-controlled reaction has been controversial. It has been argued that the rate of an equilibrium-controlled reaction “can be mathematically abstracted as infinity” for the convenience of decoupling equilibrium-controlled reactions from kinetic reactions (Fang et al. 2003). It has also been argued that the rate of an equilibrium-controlled reaction is indefinite (Lichtner 1996). This controversy might not have been aroused at all because, by definition, an equilibrium-controlled reaction should not have been associated with a rate. We can associate a rate to a fast reversible reaction only if we treat it as a kinetic reaction. When we model a fast reversible reaction as a kinetic reaction, small time steps are essential to resolve the evolution of concentration distribution among the reactants and products of the reaction. Moreover, the reaction rate approaches zero when equilibrium comes close. By adopting this concept, *we consider reaction rates undefined for all fast reactions throughout this report if they are treated as equilibrium-controlled or instantaneous reactions* to make the modeling of mixed fast and slow reactions consistent.

2.1 Governing equations of biogeochemical reaction systems

Given M species and N_R reactions in a biogeochemical reaction system, each biogeochemical reaction can be expressed as [Yeh 2000]

$$\sum_{i=1}^M \mu_{ik} S_i \longleftrightarrow \sum_{i=1}^M \nu_{ik} S_i \quad (1)$$

where:

k = the ID of the reaction, $k \in [1, N_R]$;

S_i = the ID of the i^{th} species, $i \in [1, M]$;

- μ_{ik} = the stoichiometric coefficient of the i^{th} species that is a reactant of the k^{th} reaction;
- ν_{ik} = the stoichiometric coefficient of the i^{th} species that is a product of the k^{th} reaction.

The rates of the concentration change for the M species can be represented by a set of M ordinary differential equations as

$$\left(\frac{d(D_i G_i)}{dt} \right) = r_i|_{N_R} \quad i \in [1, M] \quad (2)$$

where:

D_i = density property of the medium phase in which the i^{th} species dwells;

- it is the water content associated with the phase [(water volume)/(bulk volume)] if the medium phase is an aqueous phase;
- it is the product of the bulk density [(solid mass)/(bulk volume)] and the surface area density associated with the phase [(solid surface area)/(solid mass)] if the medium phase is a solid phase;
- it is the gas content associated with the phase [(gas volume)/(bulk volume)] if the medium phase is a gas phase.

G_i = concentration of the i^{th} species in the system;

- it is expressed in (species amount)/(water volume) if the species dwells in an aqueous phase;
- it is expressed in (species amount)/(solid surface area) if the species dwells in a solid phase;
- it is expressed in (species amount)/(gas volume) if the species dwells in a gas phase.

t = time [t];

$r_i|_{N_R}$ = production/transformation rate of the i^{th} species due to the N_R reactions in the system [$M/t/(L^3 \text{ bulk volume})$].

Equation 2 can be further written as

$$\left(\frac{d(D_i G_i)}{dt} \right) = r_i|_{N_R} = \sum_{k=1}^{N_R} (v_{ik} - \mu_{ik}) \cdot R_k \quad i \in [1, M] \quad (3)$$

where:

R_k = reaction rate of the k^{th} reaction [(species amount)/time/(bulk volume)].

It is noted that R_k can be expressed as Equation 4 when the k^{th} reaction is an elementary kinetic reaction, where the collision theory is applicable, or as Equation 5, which is an empirical formula based on observations/experimental results when the reaction is not of an elementary type.

$$R_k = k_k^f \prod_{i=1}^M (A_i)^{\mu_{ik}} - k_k^b \prod_{i=1}^M (A_i)^{\nu_{ik}} \quad k \in [1, N_R] \quad (4)$$

where:

k_k^f = forward rate constant of the k^{th} reaction
 k_k^b = backward rate constant of the k^{th} reaction
 A_i = concentration (or activity) of the i^{th} species.

$$R_k = f_k(\mathbf{A}, \mathbf{a}) \quad k \in [1, N_R] \quad (5)$$

where:

f_k = the empirical formula used to describe the rate of the k^{th} reaction
 $\mathbf{A}$ = the vector representing the M species concentrations (or activities)
 $\mathbf{a}$ = the vector representing the coefficients used to describe the rate of the k^{th} reaction.

Equation 3 can be written in matrix form as

$$\mathbf{I} \frac{d\mathbf{G}}{dt} = \mathbf{vR} \quad (6)$$

where:

- $\mathbf{I}$ = identity matrix of size M
- $\mathbf{G}$ = vector representing the M ($D_i G_i$) [(species amount)/(bulk volume)]
- $\mathbf{v}$ = reaction stoichiometry matrix with $-\mu_{ik}$ and v_{ik} as its components
- $\mathbf{R}$ = vector representing the N_R reaction rates [(species amount)/time/(bulk volume)].

Equation 6 is the so-called primitive form (PF) of rate equation that can be used to solve for the evolution of the distribution of species concentration in biogeochemical reaction systems. This formulation has been widely used in many reactive transport models (e.g., RT3D, <http://bioprocess.pnl.gov/rt3d.htm>). It is straightforward to construct the PF rate equations as well as to incorporate them as a source/sink term into multi-species transport equations. However, the following seven drawbacks exist when this primitive approach is used to solve biogeochemical reaction problems (Fang et al. 2003):

1. Time steps must be sufficiently small to resolve fast reactions: When equilibrium-controlled or instantaneous reactions exist, the time-step size must be vanishingly small, which makes integration impractical;
2. Conservation of the total mass of component species is not guaranteed due to numerical errors, attributed to truncation;
3. There is no way to define the subtraction or addition of undefined reaction rate if a species is involved in more than one equilibrium-controlled or instantaneous reaction;
4. If redundant equilibrium-controlled reactions are defined, the computational system becomes singular or over-constrained (note: a redundant equilibrium-controlled reaction is a fast reversible reaction that can be represented by other equilibrium-controlled reactions and thus is not an independent equilibrium-controlled reaction from the mathematical point of view);
5. Irrelevant kinetic (slow) reactions may exist, which makes their rate formulations and parameter determinations meaningless (note: an irrelevant kinetic reaction can be represented by some equilibrium-controlled reactions and thus is pointless to be taken into account);
6. If conflicting instantaneous (fast irreversible) reactions exist, the computational system becomes singular or over-constrained (note: a conflicting instantaneous reaction is a fast irreversible reaction that can be

represented by some equilibrium-controlled reactions and thus causes a conflict in constructing a reaction network);

7. All the PF reaction rates are coupled via the concentration-time curves of all species: reaction rates cannot be formulated and parameterized one reaction by one reaction independently of one another. For example, in the hypothetical reaction systems given in Section 2.2 below, each rate change of species concentration in Equations 7 through 12 is associated with more than one reaction rate. It requires some arrangements to represent each reaction rate as a specific combination of the time derivatives of species concentration. Without such arrangements, the formulation and parameterization of reaction rate cannot be validated with measured concentrations.

Drawbacks 1 through 3, associated with the primitive approach, can be removed by using the DAE approach, in which fast reactions are decoupled from slow reactions. Drawbacks 4 through 6 can be overcome if redundant equilibrium-controlled reactions, irrelevant kinetic reactions, and conflicting instantaneous reactions are detected and removed while setting up reaction equations manually. The traditional DAE approach cannot solve the problems mentioned in the 7th drawback. To overcome all of the drawbacks associated with the primitive approach, Fang et al. (2003) proposed a reaction-based preprocessor which uses Gauss-Jordan matrix decomposition to detect as well as remove redundant equilibrium-controlled reactions and irrelevant kinetic reactions in a systematic manner in forming the reaction equation in the DAE form. As a result, this preprocessor produces three types of equations:

- Equilibrium-controlled equations, one for each independent equilibrium-controlled reaction;
- Rate equations, one for each relevant kinetic reaction;
- Conservation equations of total mass, one for each component species.

Here we modified their preprocessor to also detect and remove conflicting instantaneous reactions. Our modified preprocessor includes the following nine steps:

- Step 1. Detect and remove redundant equilibrium-controlled reactions;
- Step 2. Detect and remove conflicting instantaneous reactions;
- Step 3. Examine consistency and conduct sorting for instantaneous reactions;

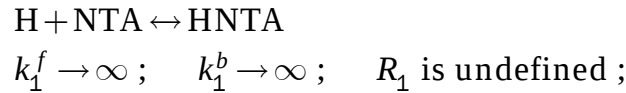
- Step 4. Detect and remove irrelevant kinetic reactions;
- Step 5. Determine the number of linearly independent reactions;
- Step 6. Rearrange the reaction order;
- Step 7. Decompose the reaction network from Step 6;
- Step 8. Express the decomposed reaction network in equation form;
- Step 9. Identify secondary species for effective biogeochemical modeling.

In Section 2.2, we demonstrate how these nine steps are implemented.

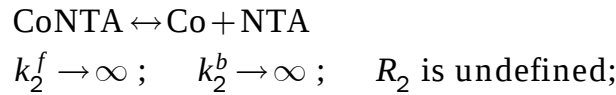
2.2 Nine-step reaction-based DAE preprocessor

Suppose that a modeler conceptualizes a biogeochemical reaction system with the following essential reactions:

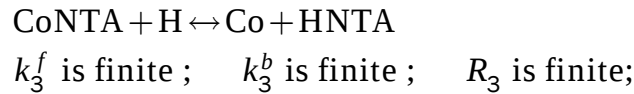
(R1) An equilibrium-controlled (fast reversible) reaction:



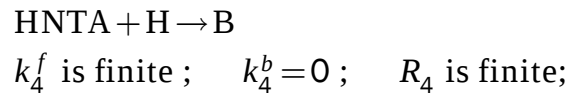
(R2) An equilibrium-controlled (fast reversible) reaction:



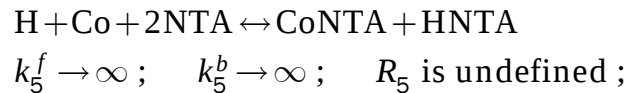
(R3) A kinetic (slow reversible) reaction:



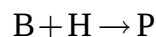
(R4) A kinetic (slow irreversible) reaction:



(R5) An equilibrium-controlled (fast reversible) reaction:

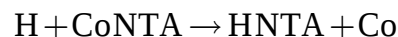


(R6) An instantaneous (fast irreversible) reaction:



$$k_6^f \rightarrow \infty; \quad k_6^b = 0; \quad R_6 \rightarrow \infty \text{ when all reactants exist};$$

(R7) An instantaneous (fast irreversible) reaction:



$$k_7^f \rightarrow \infty; \quad k_7^b = 0; \quad R_7 \rightarrow \infty \text{ when all reactants exist};$$

This system involves seven species: H, NTA, HNTA, Co, CoNTA, B, and P. For simplicity, all species are treated as nonionic. Of the seven reactions, three are equilibrium-controlled, i.e., (R1), (R2), (R5); two are instantaneous, i.e., (R6) and (R7); and two are kinetic, i.e., (R3) and (R4). However, only two of the three equilibrium-controlled reactions are independent, i.e., any one of the three equilibrium-controlled reactions can be represented by the combination of the other two. The kinetic reaction (R3) and the instantaneous reaction (R7) actually have identical reactants and products, and both reactions are equivalent to the combination of two equilibrium-controlled reactions, (R1) and (R2). In other words, (R3) is an irrelevant kinetic reaction, and (R7) is a conflicting instantaneous reaction. Both (R3) and (R7) should be excluded in biogeochemical computation to prevent the computational system from being singular or over-constrained. This example was designed to demonstrate how the nine-step preprocessing (1) detects and removes redundant equilibrium-controlled, conflicting instantaneous, and irrelevant kinetic reactions, and (2) constructs the reaction-based DAE systematically through Gauss-Jordan matrix decomposition. It may be easy for most modelers to identify redundant equilibrium-controlled, conflicting instantaneous, and irrelevant kinetic reactions in a simple reaction system like this. However, when a reaction system contains many species and reactions, it may become a tedious task even for an experienced geochemist. In this case, using a preprocessor to help construct a valid reaction network and generate the associated governing equation in DAE form will save the modeler a significant amount of time and effort.

For convenience in this demonstration, the density property is set to unity, i.e., $D_i = 1$, for the aqueous phase associated with all species. The rate equations for the seven species are:

$$\frac{d[H]}{dt} = -R_1 - R_3 - R_4 - R_5 - R_6 - R_7 \quad (7)$$

$$\frac{d[NTA]}{dt} = -R_1 + R_2 - 2R_5 \quad (8)$$

$$\frac{d[HNTA]}{dt} = R_1 + R_3 - R_4 + R_5 + R_7 \quad (9)$$

$$\frac{d[Co]}{dt} = R_2 + R_3 - R_5 + R_7 \quad (10)$$

$$\frac{d[CoNTA]}{dt} = -R_2 - R_3 + R_5 - R_7 \quad (11)$$

$$\frac{d[B]}{dt} = R_4 - R_6 \quad (12)$$

$$\frac{d[P]}{dt} = R_7 \quad (13)$$

Equations 7 through 13 can be written in matrix form as

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \\ d[P]/dt \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & -2 & 0 & 0 \\ 1 & 0 & 1 & -1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1 & 0 & 1 \\ 0 & -1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \\ R_5 \\ R_6 \\ R_7 \end{bmatrix} \quad (14)$$

Step 1. Detect and remove redundant equilibrium-controlled reactions

We first rearrange the order of reactions such that all of the given equilibrium-controlled reactions are placed before the other types of reaction. Because (R1), (R2), and (R5) are the given equilibrium-controlled reactions, we swap (R5) and (R3) in Equation 14 (red boxes), which results in Equation 15. In Equation 15, the first three columns of the coefficient matrix on the right-hand side are associated with the three given equilibrium-controlled reactions.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \\ d[P]/dt \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & -2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_5 \\ R_4 \\ R_3 \\ R_6 \\ R_7 \end{bmatrix} \quad (15)$$

We then use column elimination to determine the rank of the columns associated with equilibrium-controlled reactions (Press et al. 1992). As shown below, the rank of the three “equilibrium-controlled reaction” columns is two (the number of columns with non-zero entries), which means there are two independent equilibrium-controlled reactions and one of the three given equilibrium-controlled reactions is redundant and must be removed from the reaction network.

$$\begin{bmatrix} -1 & 0 & -1 \\ -1 & 1 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(R1) (R2) (R5) (R1) (R2) (R5)

If we choose (R5) to be the redundant equilibrium-controlled reaction and remove it from the reaction network, Equation 15 becomes

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \\ d[P]/dt \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_4 \\ R_3 \\ R_6 \\ R_7 \end{bmatrix} \quad (16)$$

Step 2. Detect and remove conflicting instantaneous reactions

We now rearrange the reaction order again by moving all instantaneous reaction columns immediately to the right of the equilibrium-controlled reaction columns. Here (R7) and (R6) are swapped with (R4) and (R3), respectively, in this process. Equation 16 thus becomes

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \\ d[P]/dt \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_7 \\ R_6 \\ R_3 \\ R_4 \end{bmatrix} \quad (17)$$

We then examine, via column elimination, the rank of the sub-matrix that contains the columns associated with the linearly independent equilibrium-controlled reactions and one column that is associated with an instantaneous reaction. This examination is conducted for all instantaneous reactions, one at a time. Below shows the examination of the two columns associated with instantaneous reactions.

$$\begin{bmatrix} -1 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

(R1) (R2) (R7) (R1) (R2) (R7) (R1) (R2) (R6)

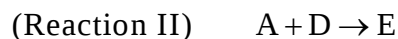
It is obvious that (R7), which can be represented as the sum of (R1) and (R2), is a conflicting instantaneous reaction and must be removed. On the other hand, (R6) is not a conflicting instantaneous reaction and will stay in the reaction network. After removing (R7), Equation 17 becomes

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \\ d[P]/dt \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \\ R_5 \\ R_6 \\ R_7 \end{bmatrix} \quad (18)$$

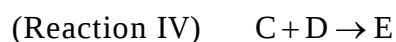
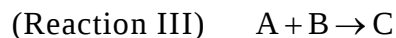
Step 3. Examine consistency and conduct sorting for instantaneous reactions

In this step, we first check whether there exists any inconsistency among the given instantaneous reactions. The following two rules are employed for examination.

Rule #1: A species must not appear on the reactant side of different instantaneous reactions. In the two instantaneous reaction below, for example, Species A appears on the reactant side of both reactions, and it is undefined how much of Species C and E will be produced when Species B and D are present in excess.



Rule #2: No loop can exist among instantaneous reactions. In other words, an instantaneous reaction cannot be both an upstream and a downstream reaction of another instantaneous reaction. One reaction is considered upstream of another reaction when one of its product species is a reactant of the other reaction. It is considered downstream of another reaction when one of its reactant species is a product of the other reaction. For example, (Reaction III) below is upstream of (Reaction V) because it indirectly produces Species E which is a reactant of (Reaction V). In the meantime, (Reaction III) is also downstream of (Reaction V) because one of its reactants, Species A, is a product of (Reaction V). This loop of instantaneous reaction is impossible to solve.





When Rule #1 or #2 is violated, the preprocessor will output a warning message reporting the inconsistency and stop the preprocessing. The modeler needs to remove inconsistency by removing inadequate instantaneous reactions from the input file before rerun the preprocessor.

When consistency among instantaneous reactions is assured, a sorting to re-order the instantaneous reactions is conducted. This sorting makes the instantaneous reactions in a sequence such that these reactions are handled from upstream down in the biogeochemical computation, which makes the computation more efficient. When two instantaneous reactions are not related in this upstream-downstream sense, the computation order will not matter. Figure 1 shows the algorithm implemented in RBBGCS to check consistency with the aforementioned two rules and to execute sorting.

```

Enqueue each reaction to Q
While exist Q do
  For each rxn  $q_i \in Q$  do
    insert reactants to R
    if  $r$  is counted twice then
      error message 1 and quit
    Endif
    insert products to P
  End for
  For each rxn  $q_i \in Q$  do
    if  $\exists r \in P$  then
      Enqueue  $q_i$  to  $Q'$ 
    else
      Enqueue  $q_i$  to Z
    Endif
  End for
  If  $Q' = Q$  then
    error message 2 and quit
  Else
     $Q = Q'$ , Empty P and R
  End if
Endwhile

```

Figure 1. An algorithm to check consistency and conduct sorting for instantaneous reactions.

Our demonstration example contains only one instantaneous reaction after Step 2 is implemented. That instantaneous reaction, i.e., (R6), is retained after consistency checking and sorting.

Step 4. Detect and remove irrelevant kinetic reactions

Here we examine, via column elimination, the rank of the sub-matrix that contains the columns associated with the linearly independent equilibrium-controlled reactions, consistent and sorted instantaneous reactions, and one column that is associated with a kinetic reaction. This examination is conducted for all the given kinetic reactions, one at a time. Below we examine whether the kinetic reaction (R3) is dependent on equilibrium-controlled and instantaneous reactions.

$$\begin{array}{c} \begin{bmatrix} -1 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ (R1)(R2)(R6) \text{ (R3)} \end{array}$$

As shown above, the rank is three indicating that (R3) can be represented by (R1), (R2), and/or (R6) (in the case here, $(R3) = (R1) + (R2)$). Therefore, it is an irrelevant kinetic reaction that can be removed from the reaction network. On the other hand, (R4) is not an irrelevant kinetic reaction and should be included in the reaction network. After removing (R3), Equation 18 becomes

$$\begin{array}{c} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \\ d[P]/dt \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_6 \\ R_4 \end{bmatrix} \end{array} \quad (19)$$

Step 5. Determine the number of linearly independent reactions

Suppose there are N_K relevant kinetic reactions after Step 4. The number of independent kinetic reactions can be determined by examining the rank of the sub-matrix that contains the N_K columns associated with the relevant kinetic reactions. If the rank is N_{KI} , N_{KI} is the number of independent kinetic reactions, and $N_{KD} (= N_K - N_{KI})$ is the number of dependent kinetic reactions. Then, the total number of independent reactions (N_I) is equal to the sum of the number of independent equilibrium-controlled reactions (N_{EQ}), the number of consistent instantaneous reactions (N_{INS}), and N_{KI} . That is,

$$N_I = N_{EQ} + N_{INS} + N_{KI} \quad (20)$$

As a result, the number of component species (N_C) can be computed with

$$N_C = M - N_I \quad (21)$$

In our demonstration example here, we have $M = 7$, $N_{EQ} = 2$, $N_{INS} = 1$, and $N_{KD} = 0$. Therefore, $N_I = 4$ and $N_C = 3$.

Step 6. Rearrange the reaction order

In this step, we rearrange the right hand side of Equation 19 to make the reaction sequence follow the order below:

1. Independent equilibrium-controlled reactions;
2. Independent instantaneous reactions;
3. Independent kinetic reactions;
4. Dependent kinetic reactions.

The reaction sequence for our demonstration example is: (R1), (R2), (R6), and (R4). And Equation 19 is now written as

$$\begin{bmatrix}
 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1
 \end{bmatrix}
 \begin{bmatrix}
 d[H]/dt \\
 d[NTA]/dt \\
 d[HNTA]/dt \\
 d[Co]/dt \\
 d[CoNTA]/dt \\
 d[B]/dt \\
 d[P]/dt
 \end{bmatrix}
 =
 \begin{bmatrix}
 -1 & 0 & -1 & -1 \\
 -1 & 1 & 0 & 0 \\
 1 & 0 & 0 & -1 \\
 0 & 1 & 0 & 0 \\
 0 & -1 & 0 & 0 \\
 0 & 0 & -1 & 1 \\
 0 & 0 & 1 & 0
 \end{bmatrix}
 \begin{bmatrix}
 R_1 \\
 R_2 \\
 R_6 \\
 R_4
 \end{bmatrix} \quad (22)$$

Step 7. Decompose the reaction network from step 6

In the decomposition of a reaction network, we first pick N_C suspected component species, followed by implementing the Gauss-Jordan elimination with full pivoting to decompose the reaction network to obtain the following matrix equation.

$$\begin{bmatrix}
 \mathbf{A}_{11} & \mathbf{A}_{12} \\
 \mathbf{A}_{21} & \mathbf{A}_{22}
 \end{bmatrix}
 \frac{d\mathbf{G}}{dt}
 =
 \begin{bmatrix}
 \mathbf{D} & \mathbf{K} \\
 \mathbf{O}_1 & \mathbf{O}_2
 \end{bmatrix}
 \begin{bmatrix}
 \mathbf{R}_I \\
 \mathbf{R}_{KD}
 \end{bmatrix} \quad (23)$$

where $\mathbf{A}_{11}$ is an $(N_I \times N_I)$ coefficient matrix; $\mathbf{A}_{12}$ is an $(N_I \times N_C)$ coefficient matrix; $\mathbf{A}_{21}$ is an $(N_C \times N_I)$ coefficient matrix; $\mathbf{A}_{22}$ is an $(N_C \times N_C)$ coefficient matrix; $\mathbf{D}$ is an $(N_I \times N_I)$ coefficient matrix; $\mathbf{K}$ is an $(N_I \times N_{KD})$ coefficient matrix; $\mathbf{O}_1$ is an $(N_C \times N_I)$ coefficient matrix; $\mathbf{O}_2$ is an $(N_C \times N_{KD})$ coefficient matrix; $\mathbf{R}_I$ is a rate vector representing the N_I reaction rates associated with independent reactions; $\mathbf{R}_{KD}$ is a rate vector representing the N_{KD} reaction rates associated with dependent kinetic reactions.

Fang et al. (2003) describe detailed step-by-step decomposition procedures in Appendix A, so we do not repeat the lengthy description here. Although they did not include instantaneous reactions, the essential details concerning decomposition were clearly explained in their paper. As they pointed out, the decomposition is not unique, depending on the component elements selected. From a geochemist point of view, H , NTA , and Co are most likely to be selected as component species. After decomposition in this case, Equation 22 will become

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[HNTA]/dt \\ d[Co]/dt \\ d[CoNTA]/dt \\ d[B]/dt \\ d[P]/dt \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_6 \\ R_4 \end{bmatrix} \quad (24)$$

Step 8. Express the decomposed reaction network in equation form

After decomposition, four types of equations can be produced. They are

- N_{EQ} mass action or user-specified algebraic equations representing independent equilibrium-controlled reactions;
- N_{INS} constraint equations representing independent instantaneous reactions;
- N_{KI} rate equations of kinetic variables representing independent kinetic reactions;
- N_C conservation equations of the total mass of component species.

Equation 24 is then equivalent to the following differential and algebraic equations that serve as the governing equations of the biogeochemical reaction model of our demonstration example.

1. Two mass action equations: for (R1) and (R2)

$$K_1^{eq} = \frac{k_1^f}{k_1^b} = \frac{[HNTA]}{[H][NTA]} \quad (25a)$$

$$K_2^{eq} = \frac{k_2^f}{k_3^b} = \frac{[Co][NTA]}{[CoNTA]} \quad (25b)$$

2. One constraint equation: for (R6)

$$[B]=0 \quad (\text{or } [H]=0) \quad (25c)$$

3. One rate equation: for (R4)

$$\frac{d[B]}{dt} + \frac{d[P]}{dt} = R_4 \quad (25d)$$

4. Three conservation equations: for total *NTA*, total *H*, and total *Co*

$$\frac{d[NTA]}{dt} + \frac{d[HNTA]}{dt} + \frac{d[CoNTA]}{dt} + \frac{d[B]}{dt} + \frac{d[P]}{dt} = 0 \quad (25e)$$

$$\frac{d[H]}{dt} + \frac{d[HNTA]}{dt} + 2\frac{d[B]}{dt} + 3\frac{d[P]}{dt} = 0 \quad (25f)$$

$$\frac{d[Co]}{dt} + \frac{d[CoNTA]}{dt} = 0 \quad (25g)$$

As mentioned previously, the decomposition is not unique. If we select *CoNTA*, *B*, and *HNTA* to be the component species, the DAE system generated by the preprocessor contains the following equations.

1. Two mass action equations: for (R1) and (R2)

$$K_1^{eq} = \frac{k_1^f}{k_1^b} = \frac{[HNTA]}{[H][NTA]} \quad (26a)$$

$$K_2^{eq} = \frac{k_2^f}{k_3^b} = \frac{[Co][NTA]}{[CoNTA]} \quad (26b)$$

2. One constraint equation: for (R6)

$$[B] = 0 \quad (\text{or } [H] = 0) \quad (26c)$$

3. One rate equation: for (R4)

$$-\frac{d[H]}{dt} + \frac{d[NTA]}{dt} - \frac{d[Co]}{dt} - \frac{d[P]}{dt} = R_4 \quad (26d)$$

4. Three conservation equations: for total *CoNTA*, total *B*, and total *HNTA*

$$\frac{d[Co]}{dt} + \frac{d[CoNTA]}{dt} = 0 \quad (26e)$$

$$\frac{d[H]}{dt} - \frac{d[NTA]}{dt} + \frac{d[Co]}{dt} + \frac{d[B]}{dt} + 2\frac{d[P]}{dt} = 0 \quad (26f)$$

$$-\frac{d[H]}{dt} + 2\frac{d[NTA]}{dt} + \frac{d[HNTA]}{dt} - 2\frac{d[Co]}{dt} - \frac{d[P]}{dt} = 0 \quad (26g)$$

If $CoNTA$, B , and P are chosen as the component species, the DAE system resulting from decomposition will include:

1. Two mass action equations: for (R1) and (R2)

$$K_1^{eq} = \frac{k_1^f}{k_1^b} = \frac{[HNTA]}{[H][NTA]} \quad (27a)$$

$$K_2^{eq} = \frac{k_2^f}{k_3^b} = \frac{[Co][NTA]}{[CoNTA]} \quad (27b)$$

2. One constraint equation: for (R6)

$$[B] = 0 \quad (\text{or } [H] = 0) \quad (27c)$$

3. One rate equation: for (R4)

$$\frac{d[NTA]}{dt} + \frac{d[HNTA]}{dt} - \frac{d[Co]}{dt} = R_4 \quad (27d)$$

4. Three conservation equations: for total $CoNTA$, total B , and total P

$$\frac{d[Co]}{dt} + \frac{d[CoNTA]}{dt} = 0 \quad (27e)$$

$$-\frac{d[H]}{dt} + 3\frac{d[NTA]}{dt} + 2\frac{d[HNTA]}{dt} - 3\frac{d[Co]}{dt} + \frac{d[B]}{dt} = 0 \quad (27f)$$

$$\frac{d[H]}{dt} - 2\frac{d[NTA]}{dt} - \frac{d[HNTA]}{dt} + 2\frac{d[Co]}{dt} + \frac{d[P]}{dt} = 0 \quad (27g)$$

By examining Equations 25a through 27g, we observe the following when different choices are made for the component species.

- The two mass action equations for equilibrium-controlled reactions and the constraint equation for the instantaneous reaction remain the same;

- The rate of the kinetic reaction (R_4) can be represented by different combinations of the time derivatives of species concentration, as given in Equations 25d, 26d, and 27d;
- The contents of the total mass of a component species may depend on the selection of the other component species, as shown in the comparison of Equations 26e through 26g and 27e through 27g.

The second observation above is extremely useful for experimentalists to compute the rate of each kinetic reaction, which is essential in the analysis of reaction mechanisms. It is common that not all species in any biogeochemical systems can be analyzed quantitatively. Given available measured species concentrations, the aforementioned decomposition using different choices of component species greatly increases the chance of determining the true rate of a kinetic reaction. For example, when the concentration variation in time of H , Co , NTA , B , and P can be measured, both Equations 25d and 26d can be used to determine and verify R_4 . Equation 27d cannot be used to compute R_4 because the concentration of $HNTA$ is not available. When only the concentrations of H , Co , and NTA are available, it looks like none of Equations 25d, 26d, and 27d can be used. However, since $[HNTA] = K_1^{eq}[H][NTA]$, the time derivative of $HNTA$ can be computed based on the concentration profiles of H and NTA , and Equation 27d can thus be used to compute R_4 .

When there are multiple kinetic reactions in the system, one may need to use one set of component species in the decomposition to compute the rate of one kinetic reaction and use a different set of component species to determine the rate of another reaction. It all depends on what species concentrations are available.

Step 9. Identify secondary species for effective biogeochemical modeling

In many cases, equilibrium-controlled reactions can be described with mass action equations, where the concentrations (or activities) of a species can be represented by the concentrations (or activities) of the other species through equilibrium constant and stoichiometries. That species can be a reactant or a product. Suppose there are $N_{EQ}^{secondary}$ species associated with equilibrium-controlled reactions, which can be represented by the selected component species. The concentrations of these $N_{EQ}^{secondary}$ species can be computed then based on the concentrations of the selected component species. In other words, we can represent these $N_{EQ}^{secondary}$ species with the

selected component species when we solve the DAE system, and use the computed concentrations of component species to compute the $N_{EQ}^{secondary}$ species. As a result, we are solving for $(M - N_{EQ}^{secondary})$, rather than M , concentrations in the DAE system.

These $N_{EQ}^{secondary}$ species are called secondary species. As defined by Fang et al. (2003), a secondary species must satisfy the following conditions.

- It is associated with an equilibrium-controlled reaction;
- It is not an ion-exchanged species;
- It is not a precipitated species;
- Its concentration (or activity) can be represented with a mass action equilibrium equation that is formulated based on the equilibrium constant approach in thermodynamics.

The set of secondary species can be obtained by applying Gauss-Jordan row elimination to the matrix whose rows are composed of reaction stoichiometries of the linearly independent equilibrium-controlled reactions that are modeled with mass action equilibrium equations. For example, suppose that H , NTA , and Co are component species, $HNTA$ and $CoNTA$ are qualified to be secondary species in our demonstration example, the reaction stoichiometries of (R1) and (R2) can be expressed in matrix form as

$$\begin{array}{cccccc}
 \begin{array}{|c|c|c|c|c|c|c|}
 \hline
 -1 & -1 & 1 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 1 & 0 & 1 & -1 & 0 & 0 \\
 \hline
 \end{array} & \begin{array}{l} \text{(R1)} \\ \text{(R2)} \end{array} \\
 H & NTA & HNTA & Co & CoNTA & B & P
 \end{array}$$

The matrix above contains the stoichiometric coefficients of the two reactions, where negative numbers are used to represent reactant species and positive numbers represent product species. We now diagonalize the matrix by selecting pivot elements that are species to be used to eliminate mass action equilibrium equations, i.e., the secondary species. A pivot element can be any species in an equilibrium-controlled reaction except the component species, the species associated with user-specified algebraic equations to represent equilibrium-controlled reaction, the ion-exchanged species, the precipitated species, or the species that have already been used as pivot elements. In the matrix above, $HNTA$ and

CoNTA are pivot elements, and the matrix is already diagonalized. Therefore, the concentrations of *HNTA* and *CoNTA* can be represented as

$$K_1^{eq} = [H]^{-1}[NTA]^{-1}[HNTA]^1 \Rightarrow [HNTA] = \frac{K_1^{eq}}{[H]^{-1}[NTA]^{-1}} \quad (28)$$

$$K_2^{eq} = [Co]^1[NTA]^1[CoNTA]^{-1} \Rightarrow [CoNTA] = \left(\frac{K_2^{eq}}{[Co]^1[NTA]^1} \right)^{-1} \quad (29)$$

To remove secondary species in our DAE system, we first rearrange Equation 24 by moving the two secondary species, i.e., *HNTA* and *CoNTA*, to follow the other species on the left-hand side of the equation, as shown in Equation 30.

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 3 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[P]/dt \\ d[Co]/dt \\ d[B]/dt \\ d[CoNTA]/dt \\ d[HNTA]/dt \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_6 \\ R_4 \end{Bmatrix} \quad (30)$$

We then reorder the equations in the DAE system by moving the two mass action equations associated with secondary species to follow the conservation equations, as shown in Equation 31.

$$\begin{array}{l} \text{row associated with CoNTA} \Rightarrow \\ \text{row associated with HNTA} \Rightarrow \end{array} \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 3 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[P]/dt \\ d[Co]/dt \\ d[B]/dt \\ d[CoNTA]/dt \\ d[HNTA]/dt \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} R_1 \\ R_2 \\ R_6 \\ R_4 \end{Bmatrix} \quad (31)$$

Based on Equations 28 and 29, we have

$$\frac{d[HNTA]}{dt} = K_1^{eq} \left(\frac{[NTA] \cdot d[H]}{dt} + \frac{[H] \cdot d[NTA]}{dt} \right) \quad (32)$$

$$\frac{d[CoNTA]}{dt} = \frac{1}{K_2^{eq}} \left(\frac{[NTA] \cdot d[Co]}{dt} + \frac{[Co] \cdot d[NTA]}{dt} \right) \quad (33)$$

Substituting Equations 32 and 33 into the first five equations yields

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ K_1^{eq} \cdot [NTA] & 1 + \frac{[Co]}{K_2^{eq}} + K_1^{eq} \cdot [H] & 1 & \frac{[NTA]}{K_2^{eq}} & 1 \\ 1 + K_1^{eq} \cdot [NTA] & K_1^{eq} \cdot [H] & 3 & 0 & 2 \\ 0 & \frac{[Co]}{K_2^{eq}} & 0 & 1 + \frac{[NTA]}{K_2^{eq}} & 0 \end{bmatrix} \begin{Bmatrix} d[H]/dt \\ d[NTA]/dt \\ d[P]/dt \\ d[Co]/dt \\ d[B]/dt \end{Bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} R_6 \\ R_4 \end{Bmatrix} \quad (34)$$

Similar to what is done in Step 8, Equation 34 is equivalent to the following equations.

1. One constraint equation: for (R6)

$$[B] = 0 \quad (\text{or } [H] = 0) \quad (35a)$$

2. One rate equation: for (R4)

$$\frac{d[B]}{dt} + \frac{d[P]}{dt} = R_4 \quad (35b)$$

3. Three conservation equations: for total *NTA*, total *H*, and total *Co*

$$\begin{aligned} & \left(K_1^{eq} \cdot [NTA] \right) \frac{d[H]}{dt} + \left(1 + \frac{[Co]}{K_2^{eq}} + K_1^{eq} \cdot [H] \right) \frac{d[NTA]}{dt} + \frac{d[P]}{dt} \\ & + \left(\frac{[NTA]}{K_2^{eq}} \right) \frac{d[Co]}{dt} + \frac{d[B]}{dt} = 0 \end{aligned} \quad (35c)$$

$$\left(1 + K_1^{eq} \cdot [NTA] \right) \frac{d[H]}{dt} + \left(K_1^{eq} \cdot [H] \right) \frac{d[NTA]}{dt} + 3 \frac{d[P]}{dt} + 2 \frac{d[B]}{dt} = 0 \quad (35d)$$

$$\left(\frac{[Co]}{K_2^{eq}} \right) \frac{d[NTA]}{dt} + \left(1 + \frac{[NTA]}{K_2^{eq}} \right) \frac{d[Co]}{dt} = 0 \quad (35e)$$

One can thus solve Equations 35a through 35e for the concentrations of *H*, *NTA*, *P*, *Co*, and *B*. Then, the concentrations of the *HNTA* and *CoNTA* can be calculated by substituting the concentrations of *H*, *Co*, and *NTA* back into Equations 28 and 29.

2.3 Solution techniques in RBBGCS

In this section, we discuss the following topics:

- Computational procedures in RBBGCS
- The DAE system when sources exist
- Newton's method to solve DAE systems
- Implementation of the constraint equations in DAE systems
- Treatment for zero-order reactions
- Accounting for species with fixed concentrations

2.3.1 Computational procedures

When both fast and slow reactions exist in a system, the given initial concentrations may not represent the initial equilibrium condition if the “fast reaction” effect is not taken into account, particularly in hypothetical or preliminary model development or before in-situ remediation methods are initiated. This is because both equilibrium-controlled and instantaneous reactions will take place as soon as all of the species are put together in the reaction system, and a new concentration distribution among the species involved in these reactions will be reached immediately. When this happens, the estimation of the reaction rates of kinetic reactions using the given initial concentrations can be erroneous, and the computation may not converge or may converge to a wrong solution. To overcome this issue, we compute internally consistent, initial concentrations by setting the reaction rate to zero for all kinetic reactions in solving the DAE system, where the given initial concentrations are employed as input. The end result is thus a representation of the initial equilibrium condition that accounts for the fast reaction effect.

Figure 2 depicts the flow chart showing the RBBGCS procedures of computing the evolution of concentration distribution among species within a batch reaction system. As shown in the flow chart, the user provides in the input file's basic information about species, reactions and phases, as well as the initial concentrations for all species and time-series sources for species when applicable. The needed basic information about species, reactions, and phases is defined in the RBBGCS Input Guide (Appendix A). The aforementioned preprocessor then uses Gauss-Jordan matrix decomposition to process the input species and reaction information to produce the reaction-based DAE system as discussed in Section 2.2.

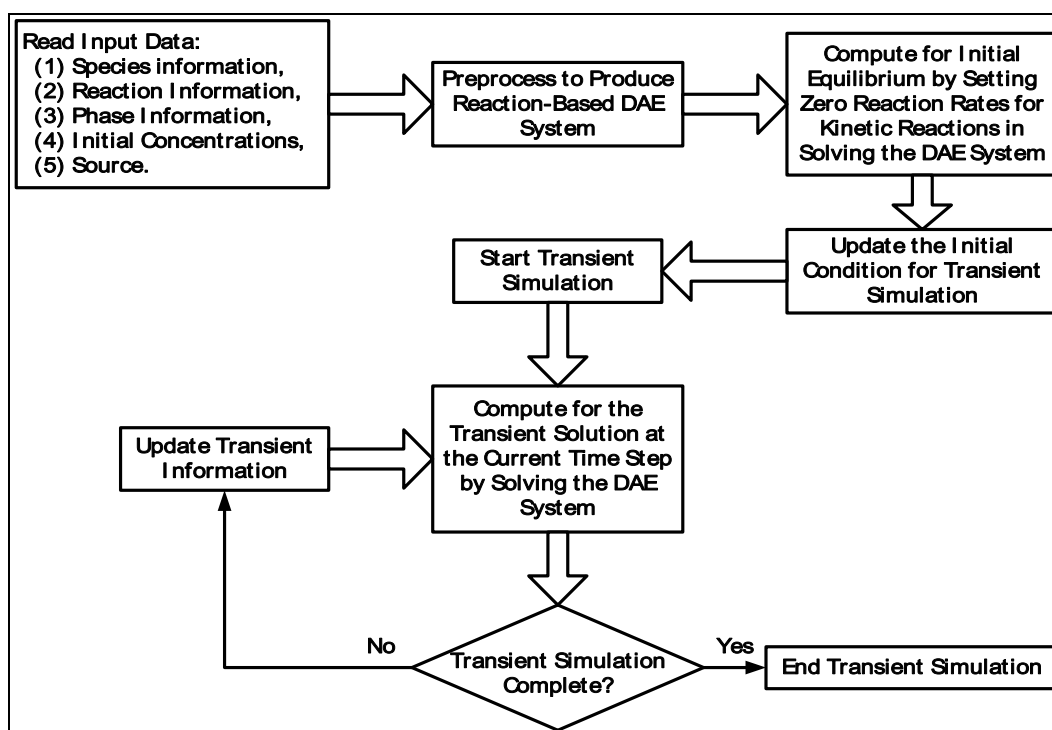


Figure 2. Flow chart of the computation in RBBGCS.

The Newton's method is employed then to solve the nonlinear DAE system (Section 2.3.3). The initial equilibrium condition for the subsequent transient simulation is computed by ignoring all kinetic reactions. This is achieved by simply setting reaction rates to zero for all kinetic reactions when the DAE system is solved as stated previously. In the transient simulation, concentration-dependent reaction rates are calculated using the given rate parameters and the concentrations from the previous Newton iteration.

2.3.2 The DAE system when sources exist

RBBGCS allows the user to model the sources of species with specified time series. When sources exist during transient simulation, the DAE system must account for the source term in the rate equations and conservation equations. For example, Equations 25a through 25g compose the DAE system of our demonstration example in Section 2.2, when H , NTA , and Co are selected to be component species and there exists no source for any species. When sources are taken into account, Equations 25a through 25c remain unchanged in the DAE system. But, Equations 25d through 25g are modified into

$$\frac{d[B]}{dt} + \frac{d[P]}{dt} = R_4 + Source_B + Source_P \quad (36a)$$

$$\begin{aligned} \frac{d[NTA]}{dt} + \frac{d[HNTA]}{dt} + \frac{d[CoNTA]}{dt} + \frac{d[B]}{dt} + \frac{d[P]}{dt} \\ = Source_{NTA} + Source_{HNTA} + Source_{CoNTA} + Source_B + Source_P \end{aligned} \quad (36b)$$

$$\begin{aligned} \frac{d[H]}{dt} + \frac{d[HNTA]}{dt} + 2\frac{d[B]}{dt} + 3\frac{d[P]}{dt} \\ = Source_H + Source_{HNTA} + 2 \cdot Source_B + 3 \cdot Source_P \end{aligned} \quad (36c)$$

$$\frac{d[Co]}{dt} + \frac{d[CoNTA]}{dt} = Source_{Co} + Source_{CoNTA} \quad (36d)$$

As a result, the DAE system with sources taken into account is composed of Equations 25a through 25c and Equations 36a through 36d.

2.3.3 Newton's method

The reaction-based DAE system (with or without the secondary species considered) accounts for two types of equation: algebraic and ordinary differential equations. Algebraic equations include those used to describe equilibrium-controlled and instantaneous reactions, e.g., Equations 25a through 25c. Ordinary equations include the rate equations for kinetic reactions and total mass conservation equations of component species, e.g., Equations 25d through 25g. The residual of an algebraic equation can be expressed in function form as

$$F = f\left([C_1]^{(t+\Delta t)}, [C_2]^{(t+\Delta t)}, \dots, [C_M]^{(t+\Delta t)}, a_1, a_2, \dots, a_{Na}\right) \quad (37)$$

where:

- F = residual of the algebraic equation;
- f = residual function associated with the algebraic equation;
- $[C_i]^{(t+\Delta t)}$ = concentration of the i^{th} species at the current time;
- a_j = value of the j^{th} parameter used for describing the algebraic equation;
- N_a = number of parameters used for describing the algebraic equation.

The residual of an ordinary differential equation can be expressed as

$$G = \sum_{i=1}^M \left(\alpha_i \cdot \frac{d[C_i]}{dt} \right) - g([C_1], [C_2], \dots, [C_{NCS}], d_1, d_2, \dots, d_{Nd}) \quad (38)$$

where:

- G = residual of the differential equation;
- g = right-hand side (source/sink) of the differential equation;
- α_i = coefficient associated with the time derivation of the i^{th} species;
- d_j = value of the j^{th} parameter used for describing the differential equation;
- N_d = number of parameters used for describing the differential equation.

Applying backward difference in time to Equation 38 yields

$$G = \sum_{i=1}^M \left(\alpha_i \cdot \frac{[C_i]^{(t+\Delta t)}}{\Delta t} \right) - \sum_{i=1}^M \left(\alpha_i \cdot \frac{[C_i]^{(t)}}{\Delta t} \right) - g([C_1]^{(t+\Delta t)}, [C_2]^{(t+\Delta t)}, \dots, [C_{NCS}]^{(t+\Delta t)}, d_1, d_2, \dots, d_{N_d}) \quad (39)$$

where $[C_i]^{(t)}$ represents the concentration of the i^{th} species at the previous time.

The Newton method is used to solve the nonlinear DAE system composed of $(N_{EQ} + N_{INS})$ algebraic equations and $(N_{KI} + N_C)$ ordinary differential equations, where the Jacobians are estimated numerically.

Figure 3 depicts the flow chart of solving the DAE system with the Newton's method, in which the Jacobian matrix equation is solved with a full-pivoting direct solver. The Newton iteration loop is highlighted with yellow arrows.

In the examination of convergence in

Figure 3, a convergent solution is obtained when Equation 40 is satisfied.

$$\frac{|C_{i,n+1} - C_{i,n}|}{RTOL \cdot \max(|C_{i,n+1}|, |C_{i,n}|) + ATOL} < 1 \quad \text{for } i \in [1, M] \quad (40)$$

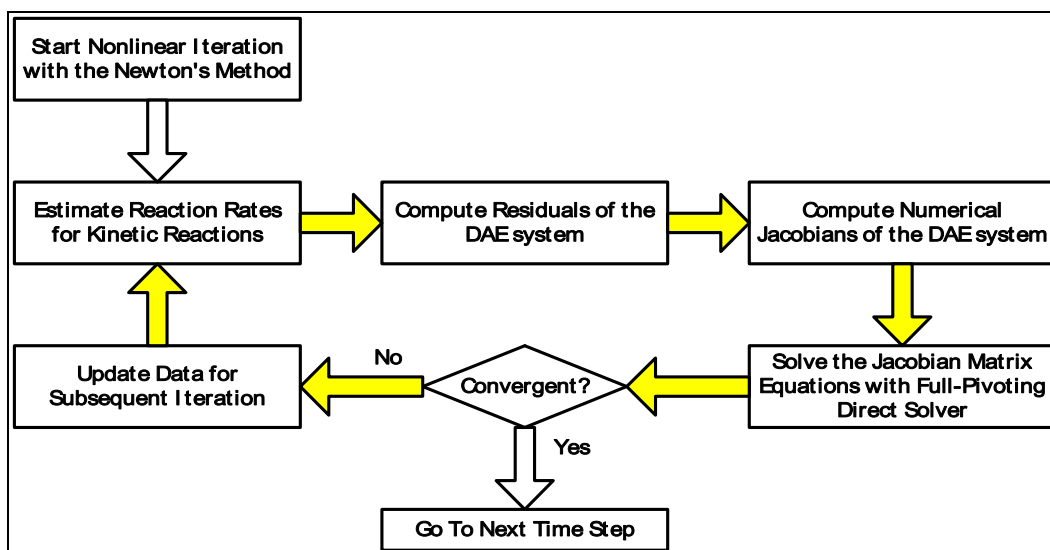


Figure 3. Flow chart of solving nonlinear DAE systems with the Newton's method.

where:

$C_{i,n+1}$ = concentration of the i^{th} species computed at the current iteration;

$C_{i,n}$ = concentration of the i^{th} species computed at the previous iteration;

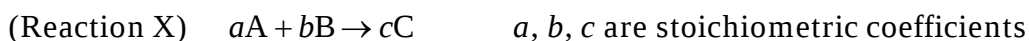
$ATOL$ = absolute error-related parameter used to determine non-linear convergence;

$RTOL$ = relative error-related parameter used to determine non-linear convergence.

Both $ATOL$ and $RTOL$ are user-specified input (Appendix A).

2.3.4 Implementation of constraint equations

When all of the reactants of an instantaneous reaction exist, the reaction will proceed at an “infinitely” high speed until one of the reactants is completely consumed. For example, if (Reaction X) below is an instantaneous reaction, we will have $[A] = 0$, or $[B] = 0$, or $[A] = [B] = 0$, anytime we examine the system.



Therefore, the constraint equation for (Reaction X) is $[A] = 0$ or $[B] = 0$. When a DAE system which includes (Reaction X) is solved, we can first assume $[A] = 0$. If the computed concentration of B is non-negative, we

have found the solution. Otherwise, we set $[B] = 0$ as the constraint equation for (Reaction X) in solving the DAE system. When there exist N_{INS} instantaneous reactions, the number of possible constraint equation combinations is equal to the product of the number of reactants in each instantaneous reaction, i.e.,

$$N_{CCE} = \prod_{i=1}^{N_{INS}} (N_i^{react}) \quad (41)$$

where:

N_{CCE} = number of possible constraint equation combinations
 N_i^{react} = number of reactants associated with the i^{th} instantaneous reaction.

For example, (Reaction X) and (Reaction Y) below are two instantaneous reactions that may occur in a biogeochemical reaction system, where there are two and three reactants in (Reaction X) and (Reaction Y), respectively.

(Reaction X) $aA + bB \rightarrow cC$ a, b, c are stoichiometric coefficients
 (Reaction Y) $dD + eE + fF \rightarrow gG$ d, e, f, g are stoichiometric coefficients

The number of possible constraint equation combinations is equal to 6 ($= 2 \times 3$). These possible combinations are listed below.

	Constraint Equation for (Reaction X)		Constraint Equation for (Reaction Y)
(Combination 1)	$[A] = 0$	&	$[D] = 0$
(Combination 2)	$[A] = 0$	&	$[E] = 0$
(Combination 3)	$[A] = 0$	&	$[F] = 0$
(Combination 4)	$[B] = 0$	&	$[D] = 0$
(Combination 5)	$[B] = 0$	&	$[E] = 0$
(Combination 6)	$[B] = 0$	&	$[F] = 0$

In RBBGCS, all possible combinations are identified and stored before solving the DAE system. At the first solution of the DAE system, i.e., computing the initial equilibrium condition, RBBGCS takes into account one combination at a time until the correct solution is found. The correct solution corresponds to when all of the computed concentrations are non-

negative. The correct combination that leads to the correct solution is then stored and used as the starting value for the first time-step computation in the transient simulation. If this combination yields negative concentration(s), RBBGCS will use the next combination from the stored information until the new correct combination is found. RBBGCS keeps updating the correct combination and using it as the first guess in the computation of the subsequent time-step to find correct solutions more quickly. Figure 4 depicts the flow chart of locating the correct constraint equation combination within a time-step (bounded by dash lines). The loop going through possible combinations to find the correct solution is highlighted with yellow arrows.

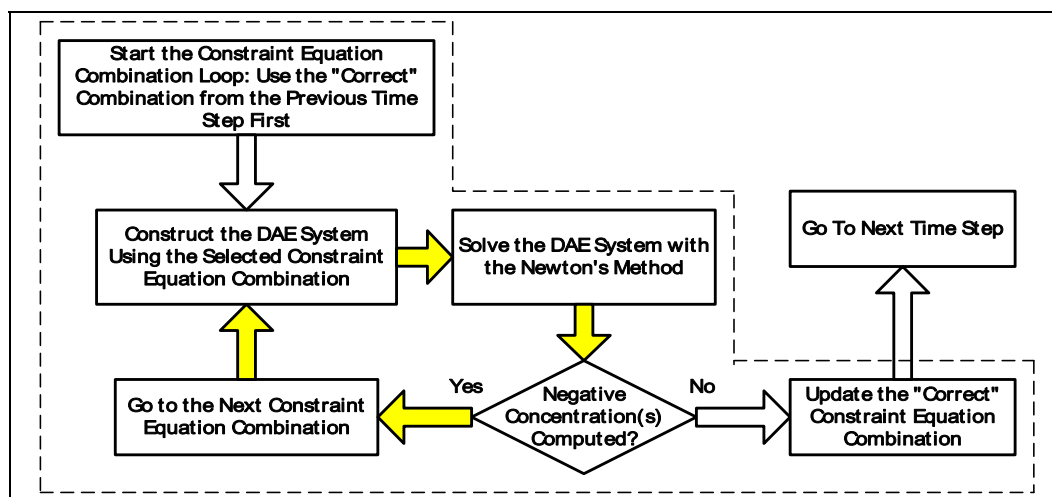
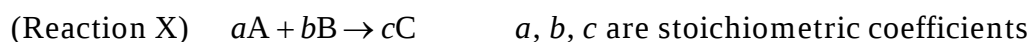


Figure 4. Flow chart of finding and using the correct constraint equation combination within a time-step in RBBGCS.

2.3.5 Treatment for zero-order reactions

A zero-order reaction, by definition, has a rate that is independent of reactant concentration(s). Variation of reactant concentrations will not change the rate of the reaction as long as all reactant concentrations are positive. The reaction rate remains constant until one of the reactants is completely consumed. Suppose (Reaction X) below is a zero-order reaction and the only reaction in a closed system.



The rate law for a zero-order reaction is

$$R = -\frac{1}{a} \frac{d[A]}{dt} = -\frac{1}{b} \frac{d[B]}{dt} = \frac{1}{c} \frac{d[C]}{dt} = R_0 \quad (42)$$

where:

R = reaction rate

R_0 = reaction rate coefficient with units of concentration/time.

Equation 42 can be integrated to yield what is often called the integrated zero-order rate law below.

$$[A]^{(t+\Delta t)} - [A]^{(t)} = -aR_0 \cdot \Delta t \quad (43a)$$

$$[B]^{(t+\Delta t)} - [B]^{(t)} = -bR_0 \cdot \Delta t \quad (43b)$$

$$[C]^{(t+\Delta t)} - [C]^{(t)} = cR_0 \cdot \Delta t \quad (43c)$$

Equations 43a through 43c are true only when both $[A]^{(t+\Delta t)}$ and $[B]^{(t+\Delta t)}$ remain non-negative. In what we call a switching-based (SB) method, these three equations are applicable if both $[A]^{(t)}$ and $[B]^{(t)}$ are positive. However, when Δt is too large, this SB method will generate negative values of $[A]^{(t+\Delta t)}$ or $[B]^{(t+\Delta t)}$ and produce non-physical solutions. In some models, $[A]^{(t+\Delta t)}$ or $[B]^{(t+\Delta t)}$ is set to zero when its computed value is negative to continue the numerical simulation. However, this reset introduces artificial mass.

To handle this situation, we propose a linear regularization (LR) method. As shown in Figure 5, $C_{\min}$ is a given small positive value that is close to zero. The reaction rate of a zero-order reaction is R_0 when the concentrations of all of the reactants are greater than or equal to $C_{\min}$, i.e., under condition 1. Otherwise, the reaction rate is set to $(f_{n+1} \cdot R_0)$, where f_{n+1} is a value computed at the $(n + 1)^{th}$ Newton's iteration based on the reactant concentrations that are computed at the n^{th} Newton's iteration and lower than $C_{\min}$, as defined in Equation 44.

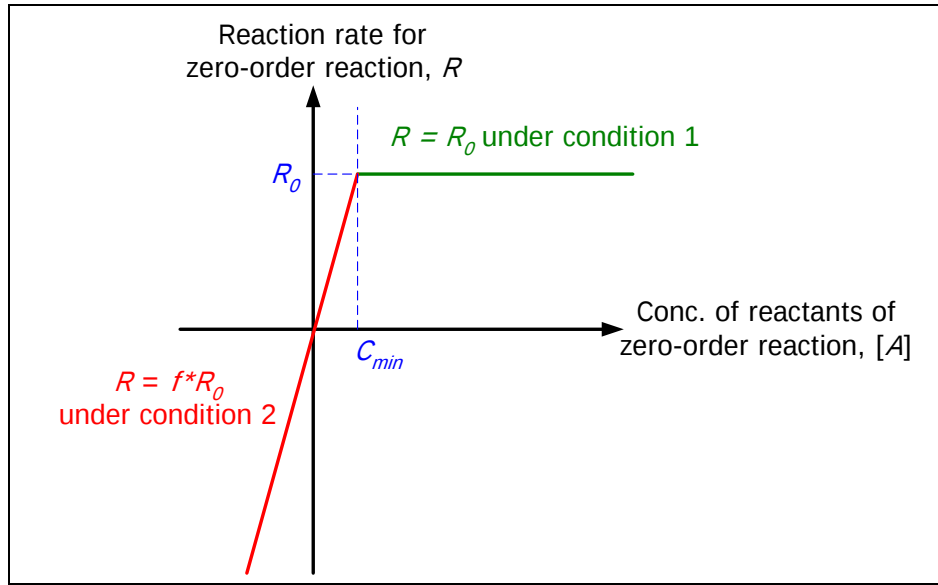


Figure 5. Relationship between zero-order reaction rate and reactant concentration in the linear regularization (LR) method.

$$f = \underset{\substack{i=1, \\ \mu_{ik} \neq 0, \\ [C_i^n] < C_{\min}}}{\text{Minimum}}^M \left(\frac{[C_i^n]}{\mu_{ik} \cdot C_{\min}} \right) \quad (44)$$

where:

M = the total number of species

μ_{ik} = the stoichiometric coefficient of the i^{th} species associated with the k^{th} reaction on the reactant side

$[C_{i,n}]$ = the concentration of the i^{th} species that is computed at the n^{th} Newton's iteration.

In this proposed LR method, the reaction rate of a zero-order reaction is adjusted linearly based upon the reactant concentration from the previous Newton's iteration when the concentration value is lower than $C_{\min}$, i.e., under condition 2. This LR method computes an equivalent reaction rate that can be used to avoid negative concentrations when time stepping is too large. In the next section, we will demonstrate how effective this LR method works when compared with the SB method.

2.3.6 Accounting for species with fixed concentrations

In some scenarios, the concentrations of some species are controlled or can be considered unchanged during the reaction period of time. For

instance, the partial pressure of oxygen in the air can be considered fixed at 0.2 atmosphere (atm) all of the time. In this case, RBBGCS simply assigns these species as component species and overwrites the corresponding conservation equations with the specified concentrations assigned to these species.

3 Verification and Demonstration Examples

Nine test problems are used to verify and demonstrate RBBGCS' various capabilities. Problems 1 through 4 are used to verify the preprocessor of RBBGCS. Problem 2 is also used to verify the solution technique in RBBGCS. Problems 5 and 6 are employed to verify the approach of using constraints equations to account for instantaneous reactions. Problem 7 is used to verify the LR method in handling zero-order reactions. Problem 8 considers a system composed of 12 species and five reactions of mixed types. Problem 9 accounts for a hypothetical TCE reaction system that includes biodegradation, adsorption-desorption, and irreversible transformation. The convergence criterion are set to 10^{-10} and 10^{-6} for *ATOL* and *RTOL*, respectively, in solving the DAE system with the Newton's method.

3.1 Verification of preprocessing

Problem 1. 7-species system

The first RBBGCS test case uses the scenario discussed previously to illustrate the nine-step preprocessing (Section 2.2). The reaction system is detailed in Table 1. As explained in Section 2, the rate of a fast reaction is considered undefined if it is treated as an equilibrium-controlled or an instantaneous reaction.

Table 1. Reaction system for Problem 1.

Reaction System 1						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	$H + NTA \leftrightarrow HNTA$	E	NA**	NA	K_1^{eq}	undefined
(R2)	$CoNTA \leftrightarrow Co + NTA$	E	NA	NA	K_2^{eq}	
(R3)	$CoNTA + H \leftrightarrow Co + HNTA$	K	k_3^f	k_3^b	NA	computed
(R4)	$HNTA + H \rightarrow B$	K	k_4^f	0	NA	
(R5)	$H + Co + 2NTA \leftrightarrow CoNTA + HNTA$	E	NA	NA	K_5^{eq}	undefined
(R6)	$B + H \rightarrow P$	I	∞	0	NA	undefined
(R7)	$H + CoNTA \rightarrow HNTA + Co$	I	∞	0	NA	

NTA = Nitrilotriacetic acid (chelating agent); Co = cobalt.

* E = equilibrium-controlled; I = instantaneous; K = kinetic.

** NA = not applicable.

As discussed in Step 8 of Section 2.2, the decomposition result depends on the choice of component species. We list here the decomposition results associated with three choices of the set of component species. It is noteworthy that there exist two secondary species associated with Choice 1, but there are no secondary species associated with Choices 2 and 3. For comparison, Table 2 summarizes key RBBGCS preprocessing output associated with the three choices. Regardless of the choice of component species, the computational system contains two algebraic equations associated with independent equilibrium-controlled reactions, one constraint equation independent instantaneous reaction, one rate equation independent kinetic reaction, and three conservation equations.

Table 2. Summary of Problem 1 Decomposition Results.

Key RBBGCS Preprocessing Output	Component Species Selected		
	< Choice 1> H, NTA, Co	< Choice 2> CoNTA, B, HNTA	< Choice 3> CoNTA, B, P
No. of Independent Equilibrium-controlled Reactions	2	2	2
No. of Independent Instantaneous Reactions	1	1	1
No. of Independent Kinetic Reactions	1	1	1
No. of Conservation Equations	3	3	3
Secondary Species	HNTA, CoNTA	None	None

< **Choice 1** > H, NTA, and Co are the selected component species:

(3) DECOMPOSITION RESULTS

```

(3-1) GROUP 1 EQUATIONS:
NO. OF MASS ACTION EQUATIONS FOR INDEPENDENT EQUILIBRIUM-CONTROLLED REACTIONS = 2
=====
THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:
INDEPENDENT EQUILIBRIUM REACTION ID  REACTION ID AS GIVEN IN THE INPUT FILE
-----
          1              1
          2              2

(3-2) GROUP 2 EQUATIONS:
NO. OF CONSTRAINT EQUATIONS FOR INDEPENDENT INSTANTANEOUS REACTIONS = 1
=====
THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:
INDEPENDENT INSTANTANEOUS REACTION ID  REACTION ID AS GIVEN IN THE INPUT FILE
-----
          3              6

(3-3) GROUP 3 EQUATIONS:
NO. OF KINETIC VARIABLE EQUATIONS FOR INDEPENDENT KINETIC REACTIONS = 1
=====
< FOR KINETIC VARIABLE EQUATION 1 >
THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:
NON-ZERO  CORRESPONDING SPECIES ID
COLUMN ID  COEFFICIENT  AS GIVEN IN THE INPUT FILE  SPECIES NAME
-----
          6          1.000          6          B
          7          1.000          7          P
THE NON-ZERO COEFFICIENTS ON THE RHS MATRIX ARE:

```

COLUMN ID	NON-ZERO COEFFICIENT	CORRESPONDING REACTION ID AS GIVEN IN THE INPUT FILE	REACTION NAME
4	1.000	4	REACTION_4

(3-4) GROUP 4 EQUATIONS:

NO. OF MASS CONSERVATION EQUATIONS FOR COMPONENT SPECIES = 3

=====

< FOR MASS CONSERVATION EQUATION 1 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES NTA

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

COLUMN ID	NON-ZERO COEFFICIENT	CORRESPONDING SPECIES ID AS GIVEN IN THE INPUT FILE	SPECIES NAME
2	1.000	2	NTA
3	1.000	3	HNTA
5	1.000	5	CoNTA
6	1.000	6	B
7	1.000	7	P

=====

< FOR MASS CONSERVATION EQUATION 2 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES H

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

COLUMN ID	NON-ZERO COEFFICIENT	CORRESPONDING SPECIES ID AS GIVEN IN THE INPUT FILE	SPECIES NAME
1	1.000	1	H
3	1.000	3	HNTA
6	2.000	6	B
7	3.000	7	P

=====

< FOR MASS CONSERVATION EQUATION 3 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES Co

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

COLUMN ID	NON-ZERO COEFFICIENT	CORRESPONDING SPECIES ID AS GIVEN IN THE INPUT FILE	SPECIES NAME
4	1.000	4	Co
5	1.000	5	CoNTA

(3-5) SECONDARY SPECIES:

NO. OF SECONDARY SPECIES = 2

=====

SECONDARY SPECIES 1 IS DERIVED FROM REACTION 1 AND IS ASSOCIATED WITH SPECIES 3 :

THE NON-ZERO COEFFICIENTS ON THE REACTANT SIDE ARE:

SPECIES ID	NON-ZERO COEFFICIENT	SPECIES NAME
1	1.000	H
2	1.000	NTA

THE NON-ZERO COEFFICIENTS ON THE PRODUCT SIDE ARE:

SPECIES ID	NON-ZERO COEFFICIENT	SPECIES NAME
3	1.000	HNTA

=====

SECONDARY SPECIES 2 IS DERIVED FROM REACTION 2 AND IS ASSOCIATED WITH SPECIES 5 :

THE NON-ZERO COEFFICIENTS ON THE REACTANT SIDE ARE:

SPECIES ID	NON-ZERO COEFFICIENT	SPECIES NAME
5	1.000	CoNTA

THE NON-ZERO COEFFICIENTS ON THE PRODUCT SIDE ARE:

SPECIES ID	NON-ZERO COEFFICIENT	SPECIES NAME
2	1.000	NTA
4	1.000	Co

===== < THIS IS THE END OF REACTION-BASED PREPROCESSING RESULTS > =====

< Choice 2 > CoNTA, B, and HNTA are the selected component species

(3) DECOMPOSITION RESULTS

(3-1) GROUP 1 EQUATIONS:

NO. OF MASS ACTION EQUATIONS FOR INDEPENDENT EQUILIBRIUM-CONTROLLED REACTIONS = 2

=====

THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:

INDEPENDENT EQUILIBRIUM REACTION ID REACTION ID AS GIVEN IN THE INPUT FILE

```
-----
              1              1
              2              2
-----
```

(3-2) GROUP 2 EQUATIONS:

NO. OF CONSTRAINT EQUATIONS FOR INDEPENDENT INSTANTANEOUS REACTIONS = 1

=====

THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:

INDEPENDENT INSTANTANEOUS REACTION ID REACTION ID AS GIVEN IN THE INPUT FILE

```
-----
              3              6
-----
```

(3-3) GROUP 3 EQUATIONS:

NO. OF KINETIC VARIABLE EQUATIONS FOR INDEPENDENT KINETIC REACTIONS = 1

=====

< FOR KINETIC VARIABLE EQUATION 1 >

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
  1      1.000      1      H
  2     -1.000      2     NTA
  4      1.000      4      Co
  7      1.000      7      P
-----
```

THE NON-ZERO COEFFICIENTS ON THE RHS MATRIX ARE:

NON-ZERO CORRESPONDING REACTION ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE REACTION NAME

```
-----
  4     -1.000      4     REACTION_4
-----
```

(3-4) GROUP 4 EQUATIONS:

NO. OF MASS CONSERVATION EQUATIONS FOR COMPONENT SPECIES = 3

=====

< FOR MASS CONSERVATION EQUATION 1 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES CoNTA

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
  4      1.000      4      Co
  5      1.000      5     CoNTA
-----
```

< FOR MASS CONSERVATION EQUATION 2 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES B

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
  1      1.000      1      H
  2     -1.000      2     NTA
  4      1.000      4      Co
  6      1.000      6      B
  7      2.000      7      P
-----
```

< FOR MASS CONSERVATION EQUATION 3 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES HNTA

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
  1     -1.000      1      H
  2      2.000      2     NTA
  3      1.000      3     HNTA
  4     -2.000      4      Co
  7     -1.000      7      P
-----
```

(3-5) SECONDARY SPECIES:

NO. OF SECONDARY SPECIES = 0

===== < THIS IS THE END OF REACTION-BASED PREPROCESSING RESULTS > =====

< **Choice 3** > CoNTA, B, and P are the selected component species

(3) DECOMPOSITION RESULTS

(3-1) GROUP 1 EQUATIONS:

NO. OF MASS ACTION EQUATIONS FOR INDEPENDENT EQUILIBRIUM-CONTROLLED REACTIONS = 2

=====

THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:

INDEPENDENT EQUILIBRIUM REACTION ID REACTION ID AS GIVEN IN THE INPUT FILE

```
-----
      1      1
      2      2
-----
```

(3-2) GROUP 2 EQUATIONS:

NO. OF CONSTRAINT EQUATIONS FOR INDEPENDENT INSTANTANEOUS REACTIONS = 1

=====

THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:

INDEPENDENT INSTANTANEOUS REACTION ID REACTION ID AS GIVEN IN THE INPUT FILE

```
-----
      3      6
-----
```

(3-3) GROUP 3 EQUATIONS:

NO. OF KINETIC VARIABLE EQUATIONS FOR INDEPENDENT KINETIC REACTIONS = 1

=====

< FOR KINETIC VARIABLE EQUATION 1 >

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
      2      -1.000      2      NTA
      3      -1.000      3      HNTA
      4       1.000      4      Co
-----
```

THE NON-ZERO COEFFICIENTS ON THE RHS MATRIX ARE:

NON-ZERO CORRESPONDING REACTION ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE REACTION NAME

```
-----
      4       1.000      4      REACTION_4
-----
```

(3-4) GROUP 4 EQUATIONS:

NO. OF MASS CONSERVATION EQUATIONS FOR COMPONENT SPECIES = 3

=====

< FOR MASS CONSERVATION EQUATION 1 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES CoNTA

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
      4       1.000      4      Co
      5       1.000      5      CoNTA
-----
```

=====

< FOR MASS CONSERVATION EQUATION 2 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES B

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
      1      -1.000      1      H
      2       3.000      2      NTA
      3       2.000      3      HNTA
      4      -3.000      4      Co
      6       1.000      6      B
-----
```

=====

< FOR MASS CONSERVATION EQUATION 3 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES P

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

NON-ZERO CORRESPONDING SPECIES ID
COLUMN ID COEFFICIENT AS GIVEN IN THE INPUT FILE SPECIES NAME

```
-----
      1       1.000      1      H
      2      -2.000      2      NTA
      3      -1.000      3      HNTA
      4       2.000      4      Co
      7       1.000      7      P
-----
```

(3-5) SECONDARY SPECIES:

NO. OF SECONDARY SPECIES = 0

===== < THIS IS THE END OF REACTION-BASED PREPROCESSING RESULTS > =====

Problem 2. Mixed microbiological and abiotic reactions

Table 3 lists the reaction system for the example problem presented in Section 6.1 of Fang et al. (2003). This reaction system includes 15 species and 10 reactions. When S_{neg} (cation exchange site), Co(II)EDTA(aq) , S_{pos} (anion exchange site), Fe(III)EDTA(aq) , O_2 , and Biomass are the selected component species, RBBCGS yields the decomposition result the same as stated in Appendix B1 of Fang et al. (2003). In this case, the reaction-based DAE system resulting from matrix decomposition contains five equilibrium-controlled equations, four kinetic variable equations, and six conservation equations.

Table 3. Reaction system for Problem 2.

Reaction System 2						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	$\text{Co(II)(aq)} + S_{\text{neg}} \leftrightarrow (S_{\text{neg}} - \text{Co})$	E	NA**	NA	$10^{1.08}$	undefined
(R2)	$\text{Co(II)EDTA(aq)} + S_{\text{pos}} \leftrightarrow (S_{\text{pos}} - \text{Co(II)EDTA})$	E	NA	NA	$10^{1.4}$	
(R3)	$\text{Fe(III)EDTA(aq)} + S_{\text{pos}} \leftrightarrow (S_{\text{pos}} - \text{Fe(III)EDTA})$	E	NA	NA	$10^{0.95}$	
(R4)	$\text{EDTA(aq)} + S_{\text{pos}} \leftrightarrow (S_{\text{pos}} - \text{EDTA})$	E	NA	NA	$10^{1.4}$	
(R5)	$\text{Co(III)EDTA(aq)} + S_{\text{pos}} \leftrightarrow (S_{\text{pos}} - \text{Co(III)EDTA})$	E	NA	NA	$10^{0.4}$	
(R6)	$(S_{\text{pos}} - \text{Co(II)EDTA}) \leftrightarrow \text{Co(II)(aq)} + (S_{\text{pos}} - \text{EDTA})$	K	10^0	10^{-3}	NA	computed
(R7)	$(S_{\text{pos}} - \text{EDTA}) \leftrightarrow \text{Fe(III)EDTA(aq)} + S_{\text{pos}}$	K	$10^{0.4}$	10^{-10}	NA	
(R8)	$\text{Co(II)EDTA(aq)} \leftrightarrow \text{Co(III)EDTA(aq)}$	K	10^{-3}	10^{-10}	NA	
(R9)	$\text{Fe(III)EDTA(aq)} + 6\text{O}_2 \rightarrow 3\text{CO}_2 + \text{Biomass}$	K	NA	0	NA	Monod+
(R10)	$\text{EDTA(aq)} + 6\text{O}_2 \rightarrow 3\text{CO}_2 + \text{Biomass}$	K	NA	0	NA	Monod

EDTA = ethylene-diamine-tetraacetic acid (chelating agent) or H_4xEDTA^- ; S_{neg} represents negatively charged surface sites for cation sorption; S_{pos} = positively charged surface site (anion sorption).

* E = equilibrium-controlled; I = instantaneous; K = kinetic.

** NA = not applicable.

+ Monod = forward reaction rate is estimated using the Monod-type equation.

Table 4 summarizes the preprocessing result.

Table 4. Summary of Problem 2 Decomposition Results.

Key RBBGCS Preprocessing Output	Component Species Selected: S_{neg} (cation exchange site), S_{pos} (anion exchange site), Co(II)EDTA(aq), Fe(III)EDTA(aq), O_2 , and Biomass
No. of Independent Equilibrium-controlled Reactions	5
No. of Independent Instantaneous Reactions	0
No. of Independent Kinetic Reactions	4
No. of Conservation Equations	6
Secondary Species	$(S_{\text{pos}} - \text{Co(II)EDTA})$, $(S_{\text{pos}} - \text{Fe(III)EDTA})$

Problem 3. Complexation, adsorption, ion-exchange, and dissolution in a system of mixed equilibrium-controlled and kinetic reactions

Table 5 lists the reaction system for the example problem presented in Section 6.2 of Fang et al. (2003). This reaction system includes 42 species and 33 reactions, where (R1) and (R2) are associated with mineral dissolution and surface site formation, (R3) through (R24) are aqueous complexation reactions, (R25) through (R31) are adsorption-desorption reactions, and (R32) and (R33) are ion-exchange reactions. When we select C1, C2, C4, C5, C6, C28, C30, (site-C6), and (site-C29) as the component species, RBBGCS yields the decomposition results identical to those stated in Appendix B2 of Fang et al. (2003). Here the preprocessor produces a reaction-based DAE system containing 25 equilibrium-controlled equations, 8 kinetic variable equations, and 9 conservation equations.

Table 6 compares the preprocessing results when two sets of component species are employed. As shown in Table 6, using Choice 1 results in 20 secondary species, but there are no secondary species associated with Choice 2.

Problem 4. Sorting instantaneous reaction systems

Here we demonstrate how the screening algorithm proposed in Step 3 of Section 2.2 can effectively check consistency and conduct sorting for instantaneous reactions. Reaction Systems 4a through 4d in Table 7 that contain various instantaneous reactions are tested with RBBGCS's preprocessor that incorporates the screening algorithm.

Table 5. Reaction system for Problem 3.

Reaction System 3						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	$M \rightarrow C_1 - 3C_2$	K	$10^{-1.3}$	0	NA**	computed
(R2)	$M \leftrightarrow S_1$	E	NA	NA	NA	partition+
(R3)	$C_3 \leftrightarrow C_4 + C_5$	K	$10^{2.03}$	10^{20}	NA	computed
(R4)	$C_6 + C_5 \leftrightarrow C_7$	E	NA	NA	$10^{12.32}$	undefined
(R5)	$C_2 + C_6 + C_5 \leftrightarrow C_8$	E	NA	NA	$10^{15.93}$	
(R6)	$C_6 \leftrightarrow C_2 + C_9$	E	NA	NA	$10^{-12.6}$	
(R7)	$C_1 + C_5 \leftrightarrow C_{10}$	K	10^{25}	$10^{-2.57}$	NA	computed
(R8)	$C_1 + C_2 + C_5 \leftrightarrow C_{11}$	E	NA	NA	$10^{29.08}$	undefined
(R9)	$C_1 + C_5 \leftrightarrow C_2 + C_{12}$	E	NA	NA	$10^{19.65}$	
(R10)	$C_1 + C_5 \leftrightarrow 2C_2 + C_{13}$	E	NA	NA	$10^{-36.3}$	
(R11)	$C_1 \leftrightarrow C_2 + C_{14}$	E	NA	NA	$10^{-2.19}$	
(R12)	$C_1 \leftrightarrow 2C_2 + C_{15}$	E	NA	NA	$10^{-5.67}$	
(R13)	$C_1 \leftrightarrow 3C_2 + C_{16}$	E	NA	NA	$10^{-13.6}$	
(R14)	$C_1 \leftrightarrow 4C_2 + C_{17}$	E	NA	NA	$10^{-21.6}$	
(R15)	$2C_1 \leftrightarrow 2C_2 + C_{18}$	E	NA	NA	$10^{-2.95}$	
(R16)	$C_2 + C_4 + C_5 \leftrightarrow C_{19}$	E	NA	NA	$10^{21.4}$	
(R17)	$C_4 \leftrightarrow C_2 + C_{20}$	E	NA	NA	$10^{-9.67}$	
(R18)	$C_4 \leftrightarrow 2C_2 + C_{21}$	E	NA	NA	$10^{-18.16}$	
(R19)	$C_4 \leftrightarrow 3C_2 + C_{22}$	E	NA	NA	$10^{-32.23}$	
(R20)	$C_2 + C_5 \leftrightarrow C_{23}$	E	NA	NA	$10^{-11.03}$	
(R21)	$2C_2 + C_5 \leftrightarrow C_{24}$	E	NA	NA	$10^{17.78}$	
(R22)	$3C_2 + C_5 \leftrightarrow C_{25}$	E	NA	NA	$10^{20.89}$	
(R23)	$4C_2 + C_5 \leftrightarrow C_{26}$	E	NA	NA	$10^{23.1}$	
(R24)	$C_2 + C_{27} \leftrightarrow C_{28}$	E	NA	NA	10^{14}	
(R25)	$S_1 \leftrightarrow S_2 + C_2$	E	NA	NA	$10^{-11.6}$	
(R26)	$S_1 + C_2 \leftrightarrow S_3$	E	NA	NA	$10^{5.6}$	
(R27)	$S_1 + 3C_2 + C_5 \leftrightarrow S_4$	E	NA	NA	$10^{30.48}$	
(R28)	$S_1 + C_1 + C_2 + C_5 \leftrightarrow S_5$	K	10^{40}	$10^{2.37}$	NA	computed
(R29)	$S_1 + C_2 + C_4 + C_5 \leftrightarrow S_6$	K	10^{30}	$10^{1.51}$	NA	
(R30)	$S_1 - C_2 + C_4 \leftrightarrow S_7$	K	$10^{-0.99}$	$10^{1.7}$	NA	
(R31)	$S_1 + C_2 + C_5 + C_6 \leftrightarrow S_8$	K	10^{25}	$10^{1.19}$	NA	
(R32)	$C_{29} + 2(\text{site} - C_{30}) \leftrightarrow (\text{site} - C_{29}) + 2C_{30}$	K	$10^{-0.75}$	$10^{-0.5}$	NA	
(R33)	$C_6 + 2(\text{site} - C_{30}) \leftrightarrow (\text{site} - C_6) + 2C_{30}$	E	NA	NA	$10^{0.6}$	undefined

M = mineral; Ci = aqueous species, $i \in [1,30]$; Sj = adsorbed species, $j \in [1,8]$; site-C₆, site-C₂₉, site-C₃₀ = ion-exchanged species

* E = equilibrium-controlled; I = instantaneous; K = kinetic

** NA = not applicable

+ partition = equilibrium equation is established using a partition relationship

Table 6. Summary of Problem 3 Decomposition Results.

Key RBBGCS Preprocessing Output	Component Species Selected:	
	< Choice 1 > C ₁ , C ₂ , C ₄ , C ₅ , C ₆ , C ₂₈ , C ₃₀ , (site-C ₆), and (site-C ₂₉)	< Choice 2 > C ₂₈ , S ₅ , S ₆ , S ₇ , S ₈ , (site- C ₆), (site-C ₂₉), (site-C ₃₀), M
No. of Independent Equilibrium-controlled Reactions	25	25
No. of Independent Instantaneous Reactions	0	0
No. of Independent Kinetic Reactions	8	8
No. of Conservation Equations	9	9
Secondary Species	20	0

Table 7. Reaction system for Problem 4.

ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
Reaction System 4a						
(R1)	A + B → C	I	∞	0	NA**	undefined
(R2)	A + D → E	I	∞	0	NA	
(R3)	F → G	I	∞	0	NA	
Reaction System 4b						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	A + B → C	I	∞	0	NA	undefined
(R2)	C + D → E	I	∞	0	NA	
(R3)	E + F → A + G	I	∞	0	NA	
Reaction System 4c						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	A + B → C	I	∞	0	NA	undefined
(R2)	C + D → E	I	∞	0	NA	
(R3)	F → G	I	∞	0	NA	
Reaction System 4d						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	E + F → G	I	∞	0	NA	undefined
(R2)	A + B → C	I	∞	0	NA	
(R3)	C + D → E	I	∞	0	NA	

* I = instantaneous; E = equilibrium-controlled; K = kinetic.

** NA = not applicable.

In Reaction System 4a, Species A appears as a reactant of two instantaneous reactions, which is against Rule #1 for consistency. When RBBGCS detects this, it will write out the following warning message and stop simulation.

CONFLICTING INSTANTANEOUS REACTIONS IDENTIFIED.

=== ERROR MESSAGE #1 IN SUBROUTINE SORT_FIR CONCERNING FAST IRREVERSIBLE REACTIONS:
SPECIES ICS APPEARS MORE THAN ONCE ON THE REACTANT SIDE!

ICS = 1

< CHECK REACTIONS THAT CONTAIN THE SPECIES >

=== SIMULATION STOPS HERE ===

In Reaction System 4b, a loop ($A \rightarrow C \rightarrow E \rightarrow A$) forms amongst the three instantaneous reactions, which is against Rule #2 for consistency. As soon as RBBGCS detects this, it prints out the following warning message and stops simulation.

CONFLICTING INSTANTANEOUS REACTIONS IDENTIFIED.

=== ERROR MESSAGE #2 IN SUBROUTINE SORT_FIR CONCERNING FAST IRREVERSIBLE REACTIONS:
A LOOP EXISTS AMONG FAST IRREVERSIBLE RXNS!

REACTION NETWORK NEEDS TO BE REVISED!

CHECK THE FOLLOWING REACTIONS:

1

2

3

=== SIMULATION STOPS HERE ===

On the other hand, Reaction Systems 4c and 4d do not violate the two rules for consistency. The decomposition results for them are as follows.

< For Reaction System 4c >

(3) DECOMPOSITION RESULTS

(3-1) GROUP 1 EQUATIONS:

NO. OF MASS ACTION EQUATIONS FOR INDEPENDENT EQUILIBRIUM-CONTROLLED REACTIONS = 0

(3-2) GROUP 2 EQUATIONS:

NO. OF CONSTRAINT EQUATIONS FOR INDEPENDENT INSTANTANEOUS REACTIONS = 3

=====

THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:

INDEPENDENT INSTANTANEOUS REACTION ID REACTION ID AS GIVEN IN THE INPUT FILE

1	1
2	2
3	3

(3-3) GROUP 3 EQUATIONS:

NO. OF KINETIC VARIABLE EQUATIONS FOR INDEPENDENT KINETIC REACTIONS = 0

(3-4) GROUP 4 EQUATIONS:

NO. OF MASS CONSERVATION EQUATIONS FOR COMPONENT SPECIES = 4

=====

< FOR MASS CONSERVATION EQUATION 1 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES D

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

COLUMN ID	COEFFICIENT	AS GIVEN IN THE INPUT FILE	SPECIES NAME
-----------	-------------	----------------------------	--------------

4	1.000	4	D
5	1.000	5	E

=====

< FOR MASS CONSERVATION EQUATION 2 >

NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES B

THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:

COLUMN ID	COEFFICIENT	AS GIVEN IN THE INPUT FILE	SPECIES NAME
-----------	-------------	----------------------------	--------------

2	1.000	2	B
3	1.000	3	C
5	1.000	5	E

```

=====
< FOR MASS CONSERVATION EQUATION 3 >
NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES A
THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:
      NON-ZERO    CORRESPONDING SPECIES ID
COLUMN ID  COEFFICIENT  AS GIVEN IN THE INPUT FILE  SPECIES NAME
-----
1          1.000        1          A
3          1.000        3          C
5          1.000        5          E
=====
< FOR MASS CONSERVATION EQUATION 4 >
NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES G
THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:
      NON-ZERO    CORRESPONDING SPECIES ID
COLUMN ID  COEFFICIENT  AS GIVEN IN THE INPUT FILE  SPECIES NAME
-----
6          1.000        6          F
7          1.000        7          G

```

(3-5) SECONDARY SPECIES:
NO. OF SECONDARY SPECIES = 0

===== < THIS IS THE END OF REACTION-BASED PREPROCESSING RESULTS > =====

< For Reaction System 4d >

(3) DECOMPOSITION RESULTS

(3-1) GROUP 1 EQUATIONS:
NO. OF MASS ACTION EQUATIONS FOR INDEPENDENT EQUILIBRIUM-CONTROLLED REACTIONS = 0

(3-2) GROUP 2 EQUATIONS:
NO. OF CONSTRAINT EQUATIONS FOR INDEPENDENT INSTANTANEOUS REACTIONS = 3

```

=====
THEY ARE ASSOCIATED WITH THE FOLLOWING REACTIONS:
INDEPENDENT INSTANTANEOUS REACTION ID  REACTION ID AS GIVEN IN THE INPUT FILE
-----
1          1
2          2
3          3

```

(3-3) GROUP 3 EQUATIONS:
NO. OF KINETIC VARIABLE EQUATIONS FOR INDEPENDENT KINETIC REACTIONS = 0

(3-4) GROUP 4 EQUATIONS:
NO. OF MASS CONSERVATION EQUATIONS FOR COMPONENT SPECIES = 4

```

=====
< FOR MASS CONSERVATION EQUATION 1 >
NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES D
THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:
      NON-ZERO    CORRESPONDING SPECIES ID
COLUMN ID  COEFFICIENT  AS GIVEN IN THE INPUT FILE  SPECIES NAME
-----
4          1.000        4          D
5          1.000        5          E
6         -1.000        6          F
=====
< FOR MASS CONSERVATION EQUATION 2 >
NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES A
THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:
      NON-ZERO    CORRESPONDING SPECIES ID
COLUMN ID  COEFFICIENT  AS GIVEN IN THE INPUT FILE  SPECIES NAME
-----
1          1.000        1          A
3          1.000        3          C
5          1.000        5          E
6         -1.000        6          F
=====
< FOR MASS CONSERVATION EQUATION 3 >
NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES B
THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:
      NON-ZERO    CORRESPONDING SPECIES ID
COLUMN ID  COEFFICIENT  AS GIVEN IN THE INPUT FILE  SPECIES NAME
-----
2          1.000        2          B
3          1.000        3          C
5          1.000        5          E
6         -1.000        6          F

```

```

=====
< FOR MASS CONSERVATION EQUATION 4 >
NOTE: THIS IS CORRESPONDING TO COMPONENT SPECIES G
THE NON-ZERO COEFFICIENTS ON THE LHS MATRIX ARE:
      NON-ZERO    CORRESPONDING SPECIES ID
COLUMN ID  COEFFICIENT  AS GIVEN IN THE INPUT FILE  SPECIES NAME
-----
      6         1.000         6         F
      7         1.000         7         G

(3-5) SECONDARY SPECIES:
NO. OF SECONDARY SPECIES = 0

===== < THIS IS THE END OF REACTION-BASED PREPROCESSING RESULTS > =====

```

It is noted in Reaction System 4d that (R2) is upstream to (R3), and (R3) to (R1). Our screening algorithm can handle instantaneous chain reactions given in arbitrary order.

3.2 Verification of solution technique

As demonstrated in Problem 1 (Section 3.1), different choices of the set of component species will produce different reaction-based DAE systems. With tight convergent criteria set for solving these DAE systems, the numerical solutions must converge to the true solution regardless of component species choice. Otherwise, the decomposition or the solution technique is not correctly implemented. To also verify that the solution technique is adequate, we solve the reaction system of Problem 3 (Section 3.1) with two choices of the set of component species and compare the simulation results. We use the nine component species mentioned in the section of Problem 3 (Section 3.1) for the first choice, and let the preprocessor automatically select component species for the second choice. Without specifying any component species in the input file, the preprocessor picks C28, S5, S6, S7, S8, (site-C6), (site-C29), (site-C30), and M as the nine component species for the second choice. The DAE systems associated with the two choices are composed of 25 equilibrium-controlled equations, eight rate equations, and nine conservation equations. As discussed in Step 8 of Section 2.2, the equilibrium-controlled equations for both DAE systems will be the same, but the rate and the conservation equations will be different.

The rate equations for the first choice are:

$$\theta \frac{d[C_{10}]}{dt} = R_7 \quad (45a)$$

$$\theta \frac{d[C_3]}{dt} = -R_3 \quad (45b)$$

$$\rho_b S_A \frac{d[S_5]}{dt} = R_{28} \quad (45c)$$

$$\rho_b S_A \frac{d[S_6]}{dt} = R_{29} \quad (45d)$$

$$\rho_b S_A \frac{d[S_7]}{dt} = R_{30} \quad (45e)$$

$$\rho_b S_A \frac{d[S_8]}{dt} = R_{31} \quad (45f)$$

$$\theta \frac{d[C_{29}]}{dt} = -R_{32} \quad (45g)$$

$$\rho_b S_A \left(\frac{d[S_1]}{dt} + \frac{d[S_2]}{dt} + \frac{d[S_3]}{dt} + \frac{d[S_4]}{dt} + \frac{d[S_5]}{dt} + \frac{d[S_6]}{dt} + \frac{d[S_7]}{dt} + \frac{d[S_8]}{dt} + \frac{d[M]}{dt} \right) = R_1 \quad (45h)$$

where:

θ = moisture content;

ρ_b = bulk density;

S_A = surface area density of the solid phase (e.g., soil).

The conservation equations for the first choice are:

$$\theta \left(\frac{d[C_6]}{dt} + \frac{d[C_7]}{dt} + \frac{d[C_8]}{dt} + \frac{d[C_9]}{dt} + \frac{d[C_{29}]}{dt} \right) + \rho_b S_A \left(\frac{d[S_8]}{dt} - \frac{1}{2} \frac{d[(site - C_{30})]}{dt} \right) = 0 \quad (45i)$$

$$\theta \left(\frac{d[C_{27}]}{dt} + \frac{d[C_{28}]}{dt} \right) = 0 \quad (45j)$$

$$\theta \left(\begin{aligned} &\frac{d[C_3]}{dt} + \frac{d[C_5]}{dt} + \frac{d[C_7]}{dt} + \frac{d[C_8]}{dt} + \frac{d[C_{10}]}{dt} \\ &+ \frac{d[C_{11}]}{dt} + \frac{d[C_{12}]}{dt} + \frac{d[C_{13}]}{dt} + \frac{d[C_{19}]}{dt} \\ &+ \frac{d[C_{23}]}{dt} + \frac{d[C_{24}]}{dt} + \frac{d[C_{25}]}{dt} + \frac{d[C_{26}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{d[S_4]}{dt} + \frac{d[S_5]}{dt} + \frac{d[S_6]}{dt} + \frac{d[S_7]}{dt} \right) = 0 \quad (45k)$$

$$\theta \frac{d[C_{30}]}{dt} + \rho_b S_A \frac{d[(site - C_{30})]}{dt} = 0 \quad (45l)$$

$$-\theta \frac{d[C_{29}]}{dt} + \rho_b S_A \left(\frac{d[(site - C_6)]}{dt} - \frac{1}{2} \frac{d[(site - C_{30})]}{dt} \right) = 0 \quad (45m)$$

$$\theta \left(\begin{aligned} &\frac{d[C_2]}{dt} + \frac{d[C_8]}{dt} - \frac{d[C_9]}{dt} + \frac{d[C_{11}]}{dt} - \frac{d[C_{12}]}{dt} \\ &- 2 \frac{d[C_{13}]}{dt} - \frac{d[C_{14}]}{dt} - 2 \frac{d[C_{15}]}{dt} - 3 \frac{d[C_{16}]}{dt} \\ &- 4 \frac{d[C_{17}]}{dt} - 2 \frac{d[C_{18}]}{dt} + \frac{d[C_{19}]}{dt} - 1 \frac{d[C_{20}]}{dt} \\ &- 2 \frac{d[C_{21}]}{dt} - 3 \frac{d[C_{22}]}{dt} + \frac{d[C_{23}]}{dt} + 2 \frac{d[C_{24}]}{dt} \\ &+ 3 \frac{d[C_{25}]}{dt} + 4 \frac{d[C_{26}]}{dt} - \frac{d[C_{27}]}{dt} \end{aligned} \right) - \rho_b S_A \left(\begin{aligned} &3 \frac{d[S_1]}{dt} + 4 \frac{d[S_2]}{dt} + 2 \frac{d[S_3]}{dt} + 2 \frac{d[S_5]}{dt} \\ &+ 2 \frac{d[S_6]}{dt} + 4 \frac{d[S_7]}{dt} + 2 \frac{d[S_8]}{dt} + 3 \frac{d[M]}{dt} \end{aligned} \right) = 0 \quad (45n)$$

$$\theta \frac{d[C_{29}]}{dt} + \rho_b S_A \frac{d[(site - C_{29})]}{dt} = 0 \quad (45o)$$

$$\theta \left(\begin{aligned} &\frac{d[C_1]}{dt} + \frac{d[C_{10}]}{dt} + \frac{d[C_{11}]}{dt} + \frac{d[C_{12}]}{dt} \\ &+ \frac{d[C_{13}]}{dt} + \frac{d[C_{14}]}{dt} + \frac{d[C_{15}]}{dt} + \frac{d[C_{16}]}{dt} \\ &+ \frac{d[C_{17}]}{dt} + 2 \frac{d[C_{18}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\begin{aligned} &\frac{d[S_1]}{dt} + \frac{d[S_2]}{dt} + \frac{d[S_3]}{dt} + \frac{d[S_4]}{dt} + 2 \frac{d[S_5]}{dt} \\ &+ \frac{d[S_6]}{dt} + \frac{d[S_7]}{dt} + \frac{d[S_8]}{dt} + \frac{d[M]}{dt} \end{aligned} \right) = 0 \quad (45p)$$

$$\theta \left(\frac{d[C_3]}{dt} + \frac{d[C_4]}{dt} + \frac{d[C_{19}]}{dt} + \frac{d[C_{20}]}{dt} + \frac{d[C_{21}]}{dt} + \frac{d[C_{22}]}{dt} \right) + \rho_b S_A \left(\frac{d[S_6]}{dt} + \frac{d[S_7]}{dt} \right) = 0 \quad (45q)$$

The rate equations for the second choice are:

$$\theta \frac{d[C_3]}{dt} = -R_3 \quad (46a)$$

$$\theta \frac{d[C_{10}]}{dt} = R_7 \quad (46b)$$

$$\theta \left(2 \frac{d[C_6]}{dt} + 2 \frac{d[C_7]}{dt} + 2 \frac{d[C_8]}{dt} + 2 \frac{d[C_9]}{dt} + 2 \frac{d[C_{29}]}{dt} + \frac{d[C_{30}]}{dt} \right) = -R_{31} \quad (46c)$$

$$\theta \frac{d[C_{29}]}{dt} = -R_{32} \quad (46d)$$

$$\theta \left(\begin{aligned} & \frac{3}{4} \frac{d[C_1]}{dt} + \frac{1}{4} \frac{d[C_2]}{dt} - \frac{1}{4} \frac{d[C_3]}{dt} + \frac{1}{4} \frac{d[C_4]}{dt} \\ & - \frac{1}{4} \frac{d[C_5]}{dt} + \frac{1}{4} \frac{d[C_6]}{dt} - \frac{1}{4} \frac{d[C_7]}{dt} + \frac{1}{4} \frac{d[C_{10}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{11}]}{dt} - \frac{1}{4} \frac{d[C_{13}]}{dt} + \frac{1}{2} \frac{d[C_{14}]}{dt} + \frac{1}{4} \frac{d[C_{15}]}{dt} \\ & - \frac{1}{4} \frac{d[C_{17}]}{dt} + \frac{d[C_{18}]}{dt} - \frac{1}{4} \frac{d[C_{21}]}{dt} - \frac{1}{2} \frac{d[C_{22}]}{dt} \\ & - \frac{1}{4} \frac{d[C_{23}]}{dt} + \frac{1}{4} \frac{d[C_{25}]}{dt} + \frac{1}{2} \frac{d[C_{26}]}{dt} - \frac{1}{4} \frac{d[C_{27}]}{dt} \\ & + \frac{1}{4} \frac{d[C_{29}]}{dt} + \frac{1}{8} \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{-1}{4} \frac{d[S_2]}{dt} + \frac{1}{4} \frac{d[S_3]}{dt} + \frac{1}{4} \frac{d[S_4]}{dt} \right) = -\frac{1}{2} R_{28} \quad (46e)$$

$$\theta \left(\begin{aligned} & \frac{d[C_1]}{dt} + \frac{d[C_2]}{dt} - \frac{d[C_3]}{dt} + \frac{d[C_4]}{dt} \\ & - 2 \frac{d[C_5]}{dt} + \frac{d[C_6]}{dt} - \frac{d[C_7]}{dt} - \frac{d[C_{10}]}{dt} \\ & - 2 \frac{d[C_{12}]}{dt} - 3 \frac{d[C_{13}]}{dt} - \frac{d[C_{15}]}{dt} - 2 \frac{d[C_{16}]}{dt} \\ & - 3 \frac{d[C_{17}]}{dt} - \frac{d[C_{21}]}{dt} - 2 \frac{d[C_{22}]}{dt} - \frac{d[C_{23}]}{dt} \\ & + \frac{d[C_{25}]}{dt} + 2 \frac{d[C_{26}]}{dt} - \frac{d[C_{27}]}{dt} - \frac{d[C_{29}]}{dt} + \frac{1}{2} \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(-\frac{d[S_2]}{dt} + \frac{d[S_3]}{dt} + \frac{d[S_4]}{dt} \right) = -2R_1 \quad (46f)$$

$$\theta \left(\begin{aligned} & \frac{3}{2} \frac{d[C_1]}{dt} + \frac{1}{2} \frac{d[C_2]}{dt} - \frac{1}{2} \frac{d[C_3]}{dt} + \frac{3}{2} \frac{d[C_4]}{dt} \\ & - 2 \frac{d[C_5]}{dt} + \frac{3}{2} \frac{d[C_6]}{dt} - \frac{1}{2} \frac{d[C_7]}{dt} + \frac{d[C_9]}{dt} \\ & - \frac{1}{2} \frac{d[C_{10}]}{dt} - \frac{d[C_{12}]}{dt} - \frac{3}{2} \frac{d[C_{13}]}{dt} + \frac{d[C_{14}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{15}]}{dt} - \frac{1}{2} \frac{d[C_{17}]}{dt} + 2 \frac{d[C_{18}]}{dt} + \frac{d[C_{20}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{21}]}{dt} - \frac{3}{2} \frac{d[C_{23}]}{dt} - \frac{d[C_{24}]}{dt} - \frac{1}{2} \frac{d[C_{25}]}{dt} \\ & - \frac{1}{2} \frac{d[C_{27}]}{dt} + \frac{3}{2} \frac{d[C_{29}]}{dt} + \frac{3}{4} \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{-1}{2} \frac{d[S_2]}{dt} + \frac{1}{2} \frac{d[S_3]}{dt} - \frac{1}{2} \frac{d[S_4]}{dt} \right) = -R_{30} \quad (46g)$$

$$\theta \left(\begin{aligned} & \frac{-3}{2} \frac{d[C_1]}{dt} - \frac{1}{2} \frac{d[C_2]}{dt} + \frac{3}{2} \frac{d[C_3]}{dt} - \frac{1}{2} \frac{d[C_4]}{dt} \\ & + 2 \frac{d[C_5]}{dt} - \frac{3}{2} \frac{d[C_6]}{dt} + \frac{1}{2} \frac{d[C_7]}{dt} - \frac{d[C_9]}{dt} \\ & + \frac{1}{2} \frac{d[C_{10}]}{dt} + \frac{d[C_{12}]}{dt} + \frac{3}{2} \frac{d[C_{13}]}{dt} - \frac{d[C_{14}]}{dt} \\ & - \frac{1}{2} \frac{d[C_{15}]}{dt} + \frac{1}{2} \frac{d[C_{17}]}{dt} - 2 \frac{d[C_{18}]}{dt} + \frac{d[C_{19}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{21}]}{dt} + \frac{d[C_{22}]}{dt} + \frac{3}{2} \frac{d[C_{23}]}{dt} + \frac{d[C_{24}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{25}]}{dt} + \frac{1}{2} \frac{d[C_{27}]}{dt} - \frac{3}{2} \frac{d[C_{29}]}{dt} - \frac{3}{4} \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{1}{2} \frac{d[S_2]}{dt} - \frac{1}{2} \frac{d[S_3]}{dt} + \frac{1}{2} \frac{d[S_4]}{dt} \right) = -R_{29} \quad (46h)$$

The conservation equations for the second choice are:

$$\theta \left(\begin{aligned} & \frac{3}{2} \frac{d[C_1]}{dt} + \frac{1}{2} \frac{d[C_2]}{dt} - \frac{1}{2} \frac{d[C_3]}{dt} + \frac{1}{2} \frac{d[C_4]}{dt} \\ & - \frac{d[C_5]}{dt} + \frac{1}{2} \frac{d[C_6]}{dt} - \frac{1}{2} \frac{d[C_7]}{dt} + \frac{1}{2} \frac{d[C_{10}]}{dt} \\ & + \frac{d[C_{11}]}{dt} - \frac{1}{2} \frac{d[C_{13}]}{dt} + \frac{d[C_{14}]}{dt} + \frac{1}{2} \frac{d[C_{15}]}{dt} \\ & - \frac{1}{2} \frac{d[C_{17}]}{dt} + 2 \frac{d[C_{18}]}{dt} - \frac{1}{2} \frac{d[C_{21}]}{dt} - \frac{d[C_{22}]}{dt} \\ & - \frac{1}{2} \frac{d[C_{23}]}{dt} + \frac{1}{2} \frac{d[C_{25}]}{dt} + \frac{d[C_{26}]}{dt} - \frac{1}{2} \frac{d[C_{27}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{29}]}{dt} + \frac{1}{4} \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{-1}{2} \frac{d[S_2]}{dt} + \frac{1}{2} \frac{d[S_3]}{dt} + \frac{1}{2} \frac{d[S_4]}{dt} + \frac{d[S_5]}{dt} \right) = 0 \quad (46i)$$

$$\theta \left(\frac{d[C_6]}{dt} + \frac{d[C_7]}{dt} + \frac{d[C_8]}{dt} + \frac{d[C_9]}{dt} + \frac{d[C_{29}]}{dt} + \frac{1}{2} \frac{d[C_{30}]}{dt} \right) - \rho_b S_A \frac{d[S_8]}{dt} = 0 \quad (46j)$$

$$-\theta \left(\frac{d[C_{29}]}{dt} + \frac{1}{2} \frac{d[C_{30}]}{dt} \right) + \rho_b S_A \frac{d[(site - C_6)]}{dt} = 0 \quad (46k)$$

$$\theta \frac{d[C_{29}]}{dt} + \rho_b S_A \frac{d[(site - C_{29})]}{dt} = 0 \quad (46l)$$

$$\theta \frac{d[C_{30}]}{dt} + \rho_b S_A \frac{d[(site - C_{30})]}{dt} = 0 \quad (46m)$$

$$\theta \left(\frac{d[C_{27}]}{dt} + \frac{d[C_{28}]}{dt} \right) = 0 \quad (46n)$$

$$\theta \left(\begin{aligned} & \frac{-3}{2} \frac{d[C_1]}{dt} - \frac{1}{2} \frac{d[C_2]}{dt} + \frac{3}{2} \frac{d[C_3]}{dt} - \frac{1}{2} \frac{d[C_4]}{dt} \\ & + 2 \frac{d[C_5]}{dt} - \frac{3}{2} \frac{d[C_6]}{dt} + \frac{1}{2} \frac{d[C_7]}{dt} - \frac{d[C_9]}{dt} \\ & + \frac{d[C_{10}]}{dt} + \frac{d[C_{12}]}{dt} + \frac{3}{2} \frac{d[C_{13}]}{dt} - \frac{d[C_{14}]}{dt} \\ & - \frac{1}{2} \frac{d[C_{15}]}{dt} + \frac{1}{2} \frac{d[C_{17}]}{dt} - 2 \frac{d[C_{18}]}{dt} + \frac{d[C_{19}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{21}]}{dt} + \frac{d[C_{22}]}{dt} + \frac{3}{2} \frac{d[C_{23}]}{dt} + \frac{1}{2} \frac{d[C_{24}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{25}]}{dt} + \frac{1}{2} \frac{d[C_{27}]}{dt} - \frac{3}{2} \frac{d[C_{29}]}{dt} - \frac{3}{4} \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{1}{2} \frac{d[S_2]}{dt} - \frac{1}{2} \frac{d[S_3]}{dt} + \frac{1}{2} \frac{d[S_4]}{dt} + \frac{d[S_6]}{dt} \right) = 0 \quad (46o)$$

$$\theta \left(\begin{aligned} & -2 \frac{d[C_1]}{dt} - \frac{d[C_2]}{dt} - 2 \frac{d[C_4]}{dt} + 2 \frac{d[C_5]}{dt} - 2 \frac{d[C_6]}{dt} \\ & - \frac{d[C_8]}{dt} - \frac{d[C_9]}{dt} - \frac{d[C_{11}]}{dt} + \frac{d[C_{12}]}{dt} + 2 \frac{d[C_{13}]}{dt} \\ & - \frac{d[C_{14}]}{dt} + \frac{d[C_{16}]}{dt} + 2 \frac{d[C_{17}]}{dt} - 2 \frac{d[C_{18}]}{dt} - \frac{d[C_{19}]}{dt} \\ & - \frac{d[C_{20}]}{dt} + \frac{d[C_{22}]}{dt} + \frac{d[C_{23}]}{dt} - \frac{d[C_{25}]}{dt} - 2 \frac{d[C_{26}]}{dt} \\ & + \frac{d[C_{27}]}{dt} - 2 \frac{d[C_{29}]}{dt} - \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{d[S_1]}{dt} + 2 \frac{d[S_2]}{dt} + \frac{d[M]}{dt} \right) = 0 \quad (46p)$$

$$\theta \left(\begin{aligned} & \frac{3}{2} \frac{d[C_1]}{dt} + \frac{1}{2} \frac{d[C_2]}{dt} - \frac{1}{2} \frac{d[C_3]}{dt} + \frac{3}{2} \frac{d[C_4]}{dt} \\ & - 2 \frac{d[C_5]}{dt} + \frac{3}{2} \frac{d[C_6]}{dt} - \frac{1}{2} \frac{d[C_7]}{dt} + \frac{d[C_9]}{dt} \\ & - \frac{1}{2} \frac{d[C_{10}]}{dt} - \frac{d[C_{12}]}{dt} - \frac{3}{2} \frac{d[C_{13}]}{dt} + \frac{d[C_{14}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{15}]}{dt} - \frac{1}{2} \frac{d[C_{17}]}{dt} + 2 \frac{d[C_{18}]}{dt} + \frac{d[C_{20}]}{dt} \\ & + \frac{1}{2} \frac{d[C_{21}]}{dt} - \frac{3}{2} \frac{d[C_{23}]}{dt} - \frac{d[C_{24}]}{dt} - \frac{1}{2} \frac{d[C_{25}]}{dt} \\ & - \frac{1}{2} \frac{d[C_{27}]}{dt} + \frac{3}{2} \frac{d[C_{29}]}{dt} + \frac{3}{4} \frac{d[C_{30}]}{dt} \end{aligned} \right) + \rho_b S_A \left(\frac{-1}{2} \frac{d[S_2]}{dt} + \frac{1}{2} \frac{d[S_3]}{dt} - \frac{1}{2} \frac{d[S_4]}{dt} + \frac{d[S_7]}{dt} \right) = 0 \quad (46q)$$

Given the initial concentrations for all 42 species, variations of the computed concentration values of selected species from the simulations associated with the two choices are calculated. The indistinguishable simulation results (Figure 6) confirm that the solution technique used in RBBGCS is adequate. The given initial concentrations are: C2 = 3.16×10^{-5} M (M = molar), C3 = 8.16×10^{-6} M, C6 = 2×10^{-3} M, C28 = 1.0 M, C30 = 0.1552, (site-C6) = 0.1562 M, (site-C29) = 0.1463 M, (site-C30) = 0.0651 M, M (mineral) = 2.36×10^{-5} M, and 0.0 M for the other species.

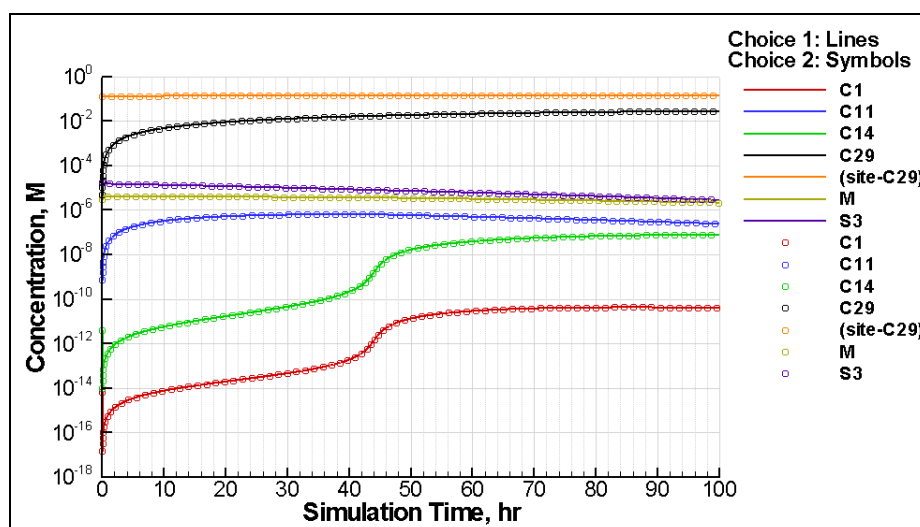


Figure 6. Variation in the computed concentration for selected species comparing simulations associated with the two alternative choices of component species.

3.3 Verification of using constraint equations for instantaneous reactions

Problem 5. One instantaneous reaction system

We first consider one hypothetical instantaneous reaction, as listed in Table 8, where consistent units of time, length, and mass can be used. The initial condition and source rate of the simulation are defined in Table 8 also. Figure 7 shows the concentration profiles of the three species, i.e., A, B, and C, from a simulation when the proposed constraint-equation approach is used to represent instantaneous reactions. It is noted that the concentration values of the three species at time zero do not match the given initial condition. This is because the initial concentrations shown in Figure 7 are the initial equilibrium concentrations, as explained in Section 2.3.1. In Figure 7a, the numerical solutions from runs with various time- step sizes are compared. It is observed from this plot that the

Table 8. Reaction system for Problem 5.

Reaction System 5						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	$A + 2B \rightarrow C$	I	∞	0	NA**	undefined
Given initial condition: $[A] = 1^\dagger$; $[B] = 1$; $[C] = 1$ at time = 0.						
Source rate: 0.001 for A during time = 0 through 60; 0.0015 for B during time = 30 through 100.						

* E = equilibrium-controlled; I = instantaneous; K = kinetic.

** NA = not applicable.

† A normalized concentration is used here.

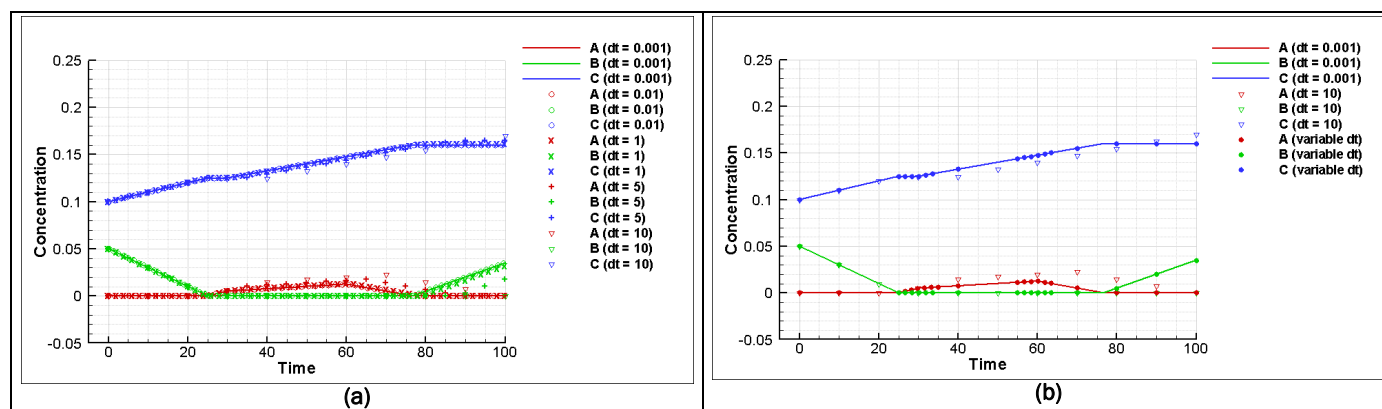


Figure 7. Numerical solution of Reaction system 5 when different time-step sizes are used.

numerical solutions associated with all the five time-step sizes are identical during time = 0 through 30, but start to differ after time = 30, when the time-step size is 5 or 10. This divergence is because RBBGCS takes the arithmetic mean of the source rates associated with the previous and current times as the source rate in the computation of the current time step. This arithmetic-mean source rate will not represent the source profile accurately when the time-step size gets too large. This problem can be overcome by using variable time steps to capture source profiles. For example, when we use 0.01 as the time-step size for time ranges 25–35 and 55–60, but 5, 10, or 15 for the other quiescent time periods, we will also receive an accurate solution (Figure 7b).

Figure 8 compares numerical solutions from the proposed constraint-equation approach (labeled FIR in the plots) and the approach that treats the reaction using rate equations with various high k^f values and $k^b = 0$ (labeled SR in the plots). In the legend of Figure 8, SR_f2 denotes $k^f = 10^2$. Based on our numerical experiment (not shown here), convergent solutions can be obtained when $k^f = 10^6$ if the time-step size is not greater than 0.001. Therefore, we use 0.001 as the time-step size for all simulation runs included in Figure 8. It is evident (Figure 8) that the SR solution does not match the FIR solution well until k^f approaches 10^6 (plotted as SR_f6). This indicates that when an instantaneous reaction is treated as a kinetic reaction, it is essential to use a sufficiently large forward rate constant to yield accurate solutions. As a consequence, the time-step size must be small enough to produce convergent solutions.

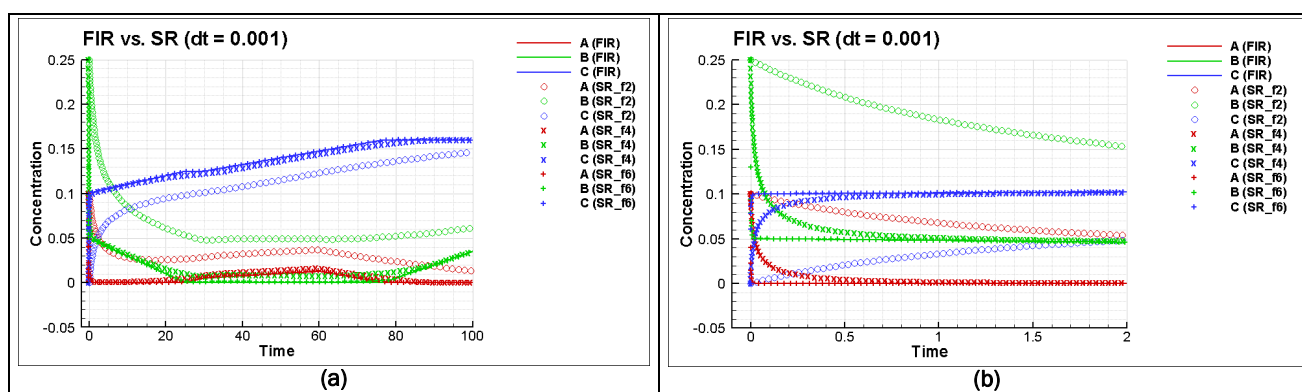


Figure 8. Numerical solutions of Reaction system 5 from FIR (using constraint equations for instantaneous reactions) and SR (treat instantaneous reactions as kinetic reactions with large forward and zero backward rate constants): (a) time = 0 to 100; (b) time = 0 to 2.

According to the definition of instantaneous reaction, the concentration profiles should be identical when we use 0.05, 0.15, and 0.05 as the given initial concentrations of A, B, and C, respectively, i.e., the 2nd set of IC. However, when instantaneous reactions are represented using rate equations, we will observe an initial condition effect if k_f is not large enough. Figure 9 demonstrates this initial condition effect, where the 1st set of IC represents the given initial condition listed in Table 8.

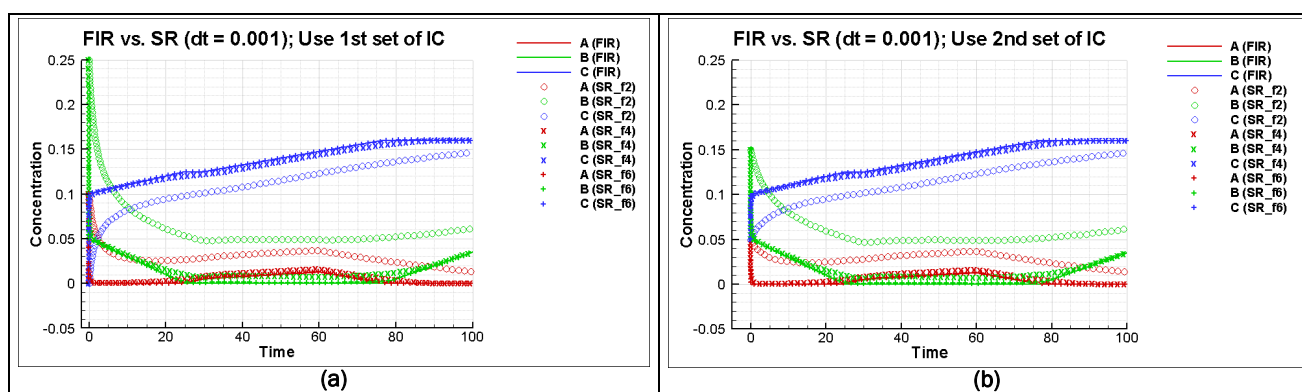


Figure 9. Numerical solutions of Reaction system 5 from FIR (using constraint equations for instantaneous reactions) and SR (treat instantaneous reactions as kinetic reactions with large forward and zero backward rate constants): (a) use 1st set of IC; (b) use 2nd set of IC.

Problem 6. Correlated instantaneous reaction system

Here we consider a reaction system that includes two instantaneous reactions, where the two reactions are correlated, as listed in Table 9. The given initial condition and constant source rate are also given in Table 9. Figure 10 depicts the concentration profiles of all of the seven species (A, B, C, D, E, F, and G), where 0.01 is the time-step for computation.

Table 9. Reaction system for Problem 6.

Reaction System 6						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	A + 2B → 2C + D	I	∞	0	NA**	undefined
(R2)	C + 2E → 2F + G	I	∞	0	NA	
Given initial condition: $[A] = [B] = [C] = [D] = [E] = [F] = [G] = 0.01$ † at time = 0.						
Constant source rate: 10 ⁻⁵ for A; 2x10 ⁻⁴ for B; 6x10 ⁻⁴ for E; 0 for the others						

* I = instantaneous; E = equilibrium-controlled; K = kinetic.

** NA = not applicable.

† An arbitrary concentration is used here.

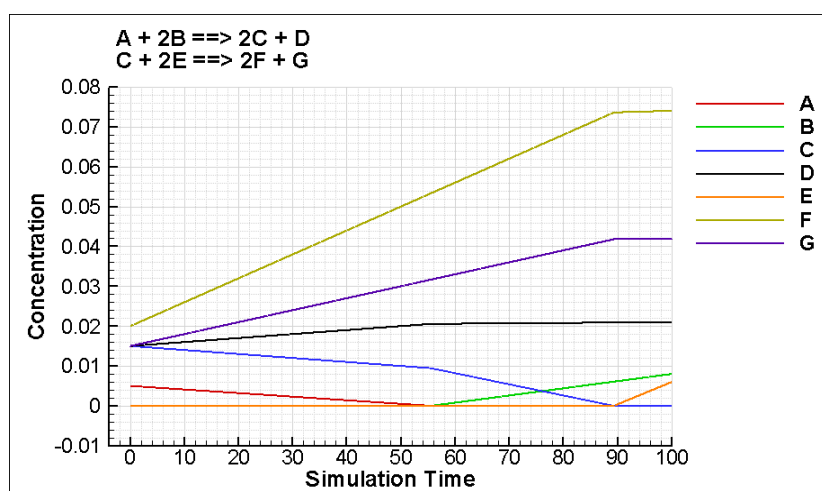


Figure 10. The computed concentration time series of Reaction system 6 using constraint equations for instantaneous reactions.

The constraint equations are $[A] = 0$ or $[B] = 0$ for (R1), and $[C] = 0$ or $[E] = 0$ for (R2). As discussed in Section 2.3.4, there exist four possible combinations of constraint equation for this reaction system, and the correct numerical solution results from using one of the four combinations in the DAE system. In this simulation, RBBGCS first computes the true initial condition based on the given initial concentrations. As a result, the true initial concentrations are 0.005 for A, 0 for B, 0.015 for C, 0.015 for D, 0 for E, 0.02 for F, and 0.015 for G (Figure 10), where $[B] = 0$ and $[E] = 0$ are the constraint equations used. In this computation of true initial condition, 0.005 of A and 0.01 of B react to produce 0.01 of C and 0.005 of D via (R1). Meanwhile, 0.005 of C and 0.01 of E react to produce 0.01 of F and 0.005 of G via (R2). As the simulation proceeds, we have A, B, and E added into the system at different rates as specified in Table 9. After time = 55.56 it is B, instead of A, appearing in the system because A is completely consumed at time = 55.56 and the source rate of B is greater than double that of the source rate of A. Likewise after time = 89.28, it is E,

rather than C, existing in the system. The constraint equations used in the DAE system for Problem 6 can be summarized as follows.

Time = 0 to 55.56, $[B]=0$ and $[E]=0$;
 Time = 55.56 to 89.28, $[A]=0$ and $[E]=0$;
 Time = 89.28 to 100, $[A]=0$ and $[C]=0$.

By using the mechanism presented in Section 2.3.4, RBBGCS automatically locates the correct combination of constraint equation to represent instantaneous reactions.

3.4 Verification of using the LR method for zero-order reactions

Problem 7. Zero-order reaction system

Table 10 lists the zero-order reaction that is employed to test the LR method discussed in Section 2.3.5, where the reaction rate is fixed at R_0 as long as both A and B exist. The given initial concentration and source rate information for simulation runs are also listed in Table 10. We compare in Figure 11 the numerical results from using the LR method versus using the SB method with two sets of R_0 and S_B (1st set: $R_0 = 2 \cdot 10^{-3}$ and $S_B = 10^{-3}$; 2nd set: $R_0 = 5 \cdot 10^{-3}$ and $S_B = 2 \cdot 10^{-3}$), and two time-step sizes (0.1 and 2). It is observed that the two approaches produce matching results when R_0 , S_B , and time-step size are small (left plot, Figure 11). On the other hand, the results from the two approaches differ significantly when R_0 , S_B , and time-step size are large (right plot, Figure 11). Moreover, the SB method may generate negative concentration, which is non-physical behavior.

Figure 12 compares the computed concentration using a time-step of 0.1 versus using 2 for this scenario. It is observed that time-step size, R_0 and S_B have no effect to the LR method. But, the deviation between the two simulation runs using the SB method increases with time-step size, as well as with R_0 and S_B .

Table 10. Reaction system for Problem 7.

Reaction System 7						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	$A + 2B \rightarrow C + 2D$	K	NA**	NA	NA	R_0
Given initial condition: $[A]=[B]=[C]=0.01^\dagger$ and $[D]=0$ at time = 0. Source rate: S_B for B during time = 42 to 50; 0 for the others.						

* E = equilibrium-controlled; I = instantaneous; K = kinetic.

** NA = not applicable.

† An arbitrary concentration is used here.

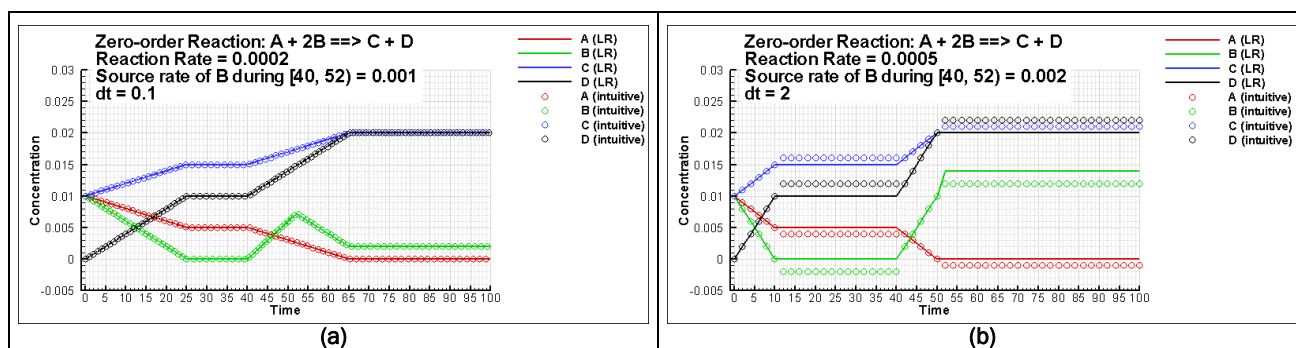


Figure 11. The computed concentration profiles of Reaction system 7 between using the LR and the switching-based methods for the zero-order reaction $A + 2B \rightarrow C + 2D$: (a) $R_0 = 2 \times 10^{-3}$, $S_B = 10^{-3}$, $dt = 0.1$; (b) $R_0 = 5 \times 10^{-3}$, $S_B = 2 \times 10^{-3}$, $dt = 2$.

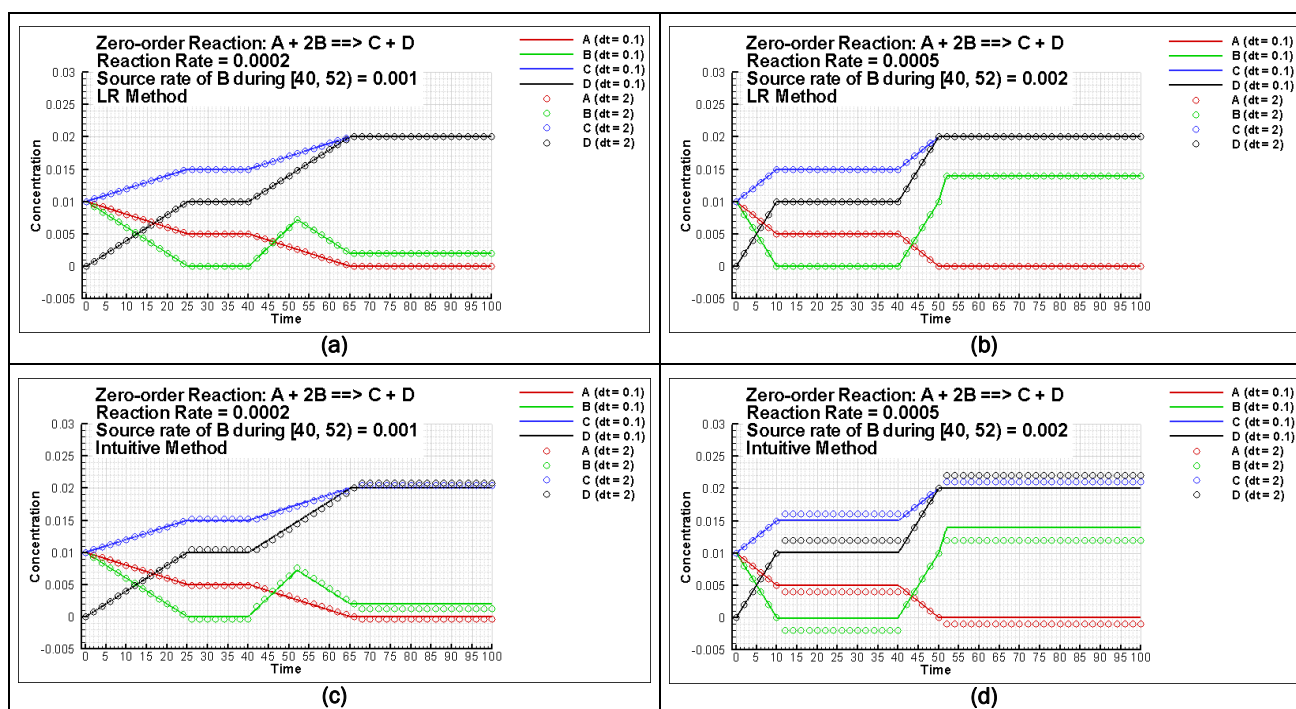


Figure 12. The computed concentration profiles of Reaction system 7 between using 0.1 and using 2 as the time-step size for simulation: (a). $R_0 = 2 \times 10^{-3}$, $S_B = 10^{-3}$, LR method; (b) $R_0 = 5 \times 10^{-3}$, $S_B = 2 \times 10^{-3}$, LR method; (c) $R_0 = 2 \times 10^{-3}$, $S_B = 10^{-3}$, switching-based method; (d) $R_0 = 5 \times 10^{-3}$, $S_B = 2 \times 10^{-3}$, switching-based method.

3.5 Capability of solving reaction systems of mixed types

Problem 8. Mixed-type reaction system (1)

Problem 8 is a hypothetical reaction system containing five reactions (two equilibrium-controlled, one instantaneous, and two kinetic) and 12 species (A, B, C, D, E, F, G, H, M, N, P, and Q), as shown in Table 11. Here we examine the numerical results when fast reactions, including equilibrium-controlled and instantaneous reactions, are treated differently in the DAE system for simulation. We use SR, SR&FIR, SR&FRR, SR&FIR&FRR to represent these different treatments (Table 12).

Table 11. Reaction system for Problem 8.

Reaction System 8						
ID	Formula	Type*	k^f	k^b	K^{eq}	Rate
(R1)	$A + 2B \rightarrow C$	K	10	0	NA**	computed
(R2)	$D + 2E \leftrightarrow F$	K	1	10	NA	
(R3)	$C + 2F \leftrightarrow 2G + H$	E	NA	NA	10	undefined
(R4)	$H + 2M \leftrightarrow N$	E	NA	NA	10	
(R5)	$N + 2P \rightarrow Q$	I	∞	0	NA	undefined

Given initial condition†: 0.01 for all species at time = 0.
 Constant Source rate: 10^{-5} for A; 2×10^{-5} for B; 3×10^{-5} for C; 4×10^{-5} for D; 5×10^{-5} for E; 6×10^{-5} for F; 6×10^{-5} for G; 5×10^{-5} for H; 4×10^{-5} for M; 5×10^{-5} for N; 5×10^{-5} for P; 10^{-5} for Q.

* E = equilibrium-controlled; I = instantaneous; K = kinetic.

** NA = not applicable.

† An arbitrary concentration unit is used here.

Table 12. Definition of various treatments used to simulate reactions in Problem 8.

ID	Treatment		
	Equilibrium-controlled Reactions	Instantaneous Reactions	Kinetic Reactions
SR	rate equations	rate equations	rate equations
SR&FIR	rate equations	zero-concentration constraint equations	rate equations
SR&FRR	algebraic equilibrium equations	rate equations	rate equations
SR&FIR&FRR	algebraic equilibrium equations	zero-concentration constraint equations	rate equations

Figure 13 depicts the concentration profiles of P and Q when the aforementioned four treatments are used to deal with Reaction system 8 in the DAE approach, where the time-step size is 1.0. As shown in the Figures 13a and 13b, when equilibrium-controlled reactions are simulated with rate equations using $k_f = 10^6$ and $k^b = 10^5$ (i.e., SR and SR&FIR), there exist significant discrepancies in the computed concentrations of P and Q.

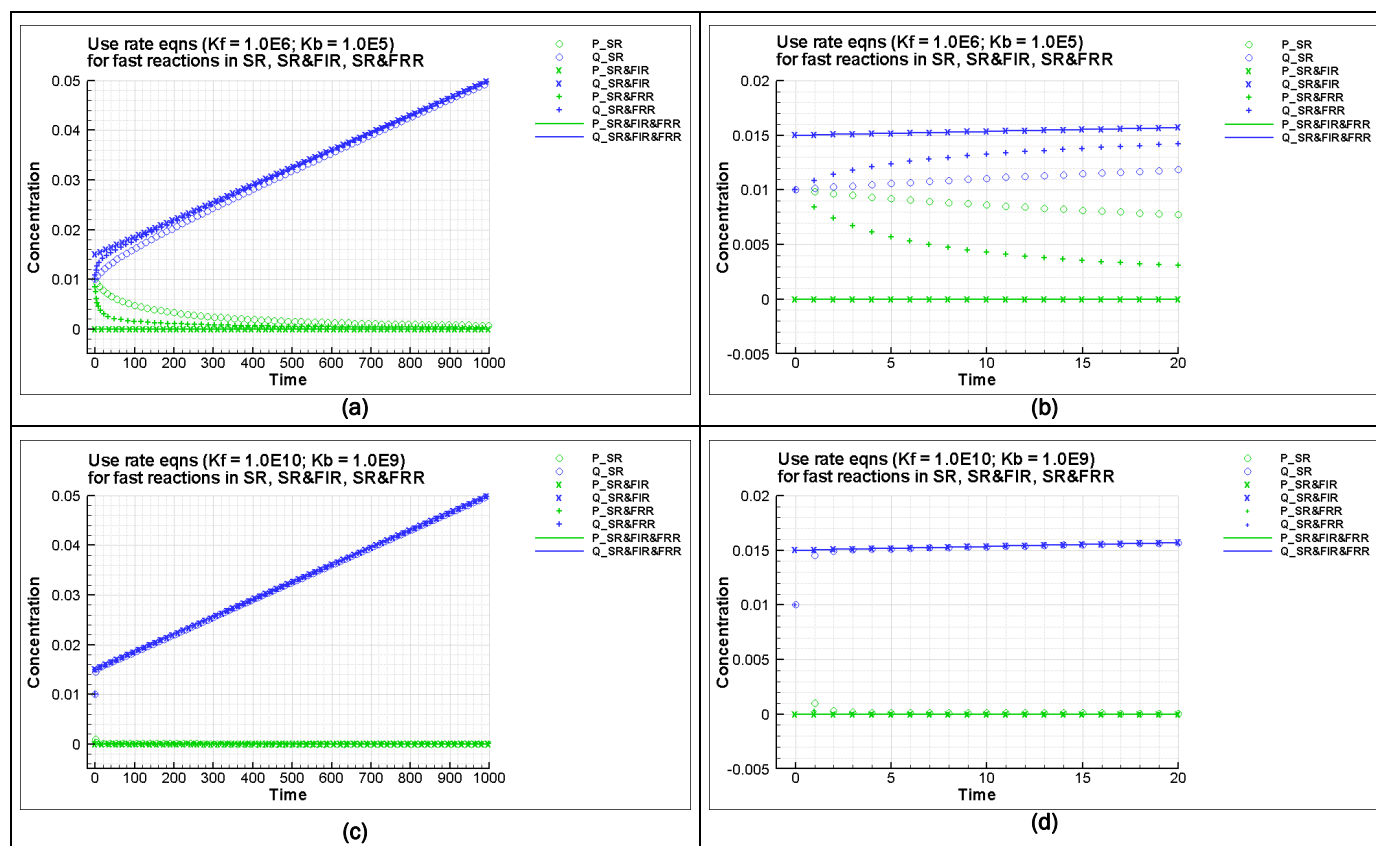


Figure 13. The computed concentration trends for P and Q in Reaction system 8 when fast reactions are simulated in different ways: use $k_f = 10^6$ and $k_b = 10^5$ in the upper plots; $k_f = 10^{10}$ and $k_b = 10^9$ in the lower plots for fast reactions in SR, SR&FIR, and SR&FRR. Plots to the right are expanded views of the left plots for earliest times and concentrations

On the other hand, when $k_f = 10^{10}$ and $k_b = 10^9$ are used in the rate equations representing fast reactions (Figures 13c and 13d), all four treatments synchronize after time = 2. This indicates that *the forward and backward rate constants used must be large enough such that rate equations can be used to represent fast reactions adequately*. The matching results from the four treatments when large rate constants are used also help to verify the RBBGCS routines involved in the computation associated with those treatments.

In a primitive RT formulation, instantaneous reactions require some approximation. For example, in a limited linear-rate (LLR) approximation, the equivalent rate of an instantaneous reaction is computed as the minimum of the reactant concentration at the previous time divided by the product of its stoichiometric coefficient and time step. For example, the equivalent rate of (R5) in Table 11 is computed as

$$R_5^{equivalent} = \frac{\min\left(\frac{[N]^t}{1}, \frac{[P]^t}{2}\right)}{\Delta t} \quad (47)$$

where $R_5^{equivalent}$ is the equivalent rate of (R5), Δt is time step, $[N]^t$ and $[P]^t$ are the concentrations of N and P at the previous time, and 1 and 2 are the stoichiometric coefficients associated with N and P, respectively, in (R5).

The above approximation is used in the RT3D model (Clement 1997) to obtain equivalent rates for instantaneous reactions. The approach is straightforward and yields adequate numerical solutions when the time-step size is small. But it may produce large errors when the time-step size is large. Figure 14 compares the impact of time-step size on the concentration profiles of N, P, and Q, as fast reactions are handled differently. When fast reactions are represented with equilibrium and constraint equations in the DAE approach (Figure 14a), the impact from time-step size is minimal. When rate equations with $k_f = 10^{10}$ and $k^b = 10^9$ are used to simulate fast reactions in the DAE approach (Figure 14b), the time-step size impact is still negligible. However, the numerical error increases with time-step size when the LRA is used to handle instantaneous reactions in the primitive approach (Figure 14c). Two factors contribute to this phenomenon. First, there are sources associated with the two reactants of Instantaneous Reaction (R5), i.e., N and P, each at a rate of 5×10^{-5} as shown in Table 11. Second, N is involved in equilibrium-controlled reaction (R4). The combined contribution from sources and (R4) usually increases with time-step size. It is thus insufficient to use only the concentrations at the previous time for computing the consumption/production associated with instantaneous reactions.

Problem 9. Hypothetical PCE reaction system

Table 13 lists the 13 reactions in this PCE reaction system, including four equilibrium-controlled reactions associated with sorption-desorption, four instantaneous reactions associated with clean-up treatment, and five irreversible kinetic reactions associated with biodegradation. This hypothetical system contains four species in the solid phase (PCE_{sorbed} , TCE_{sorbed} , DCE_{sorbed} , and VC_{sorbed}) and 14 species in the aqueous phase (PCE, TCE, DCE, VC, Ethane, CO_2 , CCP, CCT, CCD, CCV, CCR_{PCE} , CCR_{TCE} , CCR_{DCE} , and CCR_{VC}). CCP, CCT, CCD, and CCV are clean-up chemicals that are added into the reaction system and interact with PCE, TCE, DCE, and VC,

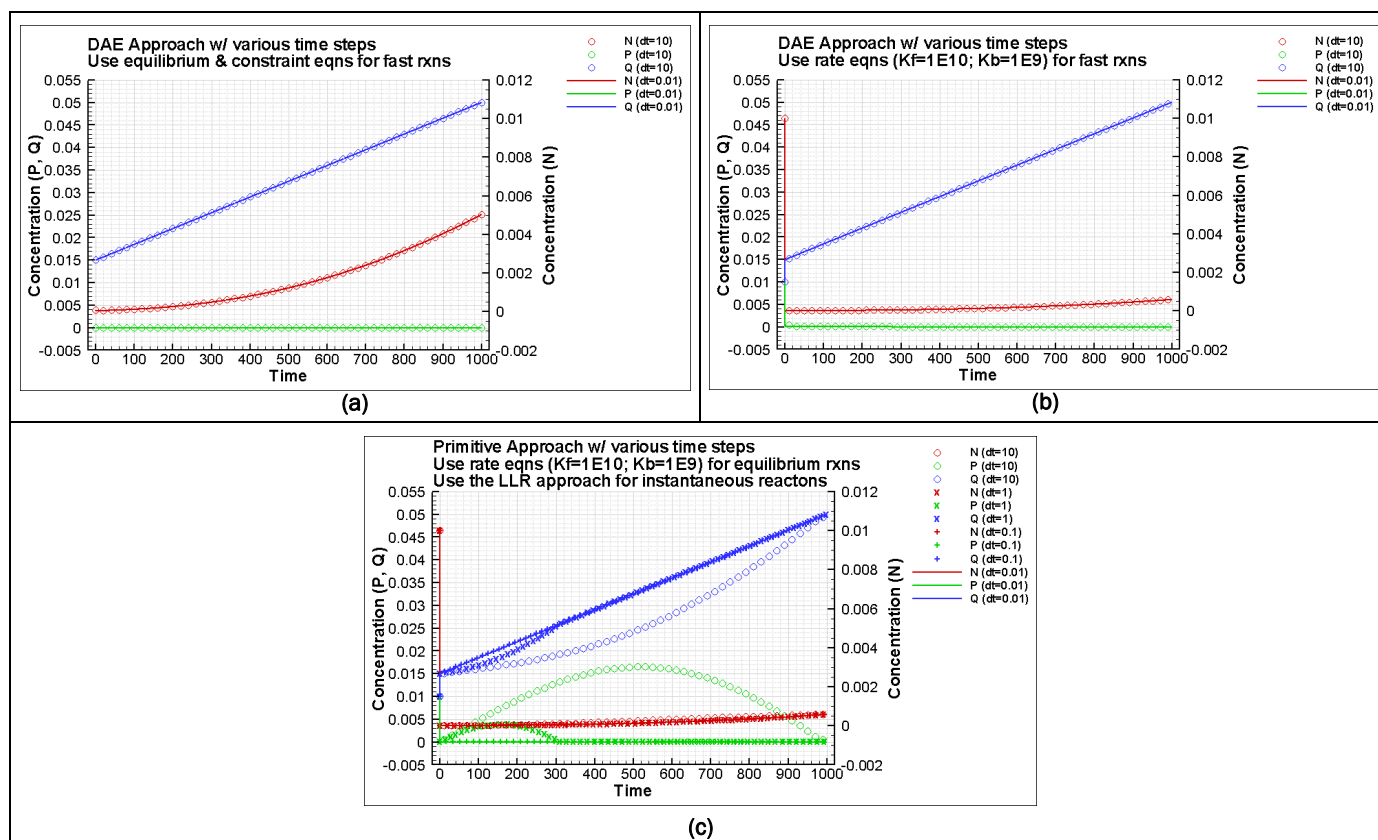


Figure 14. Impact of time-step size on the computed concentration profiles of N, P, and Q in Reaction system 8 when fast reactions are simulated in different ways: (a) use equilibrium and constraint equations for fast reactions in the DAE approach; (b) use rate equations ($k_f = 10^{10}$ and $k_b = 10^9$) for fast reactions in the DAE approach; (c) use rate equations ($k_f = 10^{10}$ and $k_b = 10^9$) for equilibrium-controlled reactions and the limited linear-rate (LLR) approach for instantaneous reactions in the primitive approach

respectively, to produce CCR_{PCE} , CCR_{PCE} , CCR_{PCE} , and CCR_{PCE} immediately. CCR_{PCE} , CCR_{TCE} , CCR_{DCE} , and CCR_{VC} are assumed to be environmentally benign. The rate constant and equilibrium constant values in Table 13 are from Glynn et al. (2001).

Figure 15 plots the concentration profiles of all species from a simulation of 10 years. In the figure, results from runs using 120, 720, and 1,800 hours (or 5, 30, and 75 days) as the time-step size are compared. As shown in the figure, the differences between using 120 and 720 hours as the time-step size are limited. Although the differences become differentiable when the time-step size increases to 1,800 hours, using 1,800 hours as the time-step size for simulation might still be acceptable for estimating clean-up efficiency from an engineering point of view.

Table 13. Reaction system for Problem 9.

Reaction System 9						
ID	Formula	Type*	k^f [1/yr]	k^b	K^{eq} [mL/g] ⁺	Rate
(R1)	$PCE \leftrightarrow PCE_{sorbed}$	E	NA**	NA	1.2	undefined
(R2)	$TCE \leftrightarrow TCE_{sorbed}$	E	NA	NA	0.4	
(R3)	$DCE \leftrightarrow DCE_{sorbed}$	E	NA	NA	0.1	
(R4)	$VC \leftrightarrow VC_{sorbed}$	E	NA	NA	0.005	
(R5)	$PCE + CCP \rightarrow CCR_{PCE}$	I	∞	0	NA	undefined
(R6)	$TCE + CCT \rightarrow CCR_{TCE}$	I	∞	0	NA	
(R7)	$DCE + CCD \rightarrow CCR_{DCE}$	I	∞	0	NA	
(R8)	$VC + CCV \rightarrow CCR_{VC}$	I	∞	0	NA	
(R9)	$PCE \rightarrow TCE$	K	1.44	0	NA	computed
(R10)	$TCE \rightarrow DCE$	K	1.08	0	NA	
(R11)	$DCE \rightarrow VC$	K	0.72	0	NA	
(R12)	$VC \rightarrow Ethane$	K	0.60	0	NA	
(R13)	$Ethane \rightarrow CO_2$	K	0.96	0	NA	computed
Arbitrary initial condition: 0.2 mM each for PCE_{sorbed} and TCE_{sorbed} and zero for the other species at time = 0. Constant Source rate: 0.1 mM/yr for CCP during time = 2 to 3 yr; 0.05 mM/yr for CCT during time = 4 to 5 yr; 0.05 mM/yr for CCD during time = 6 to 7 yr; 0.05 mM/yr for CCV during time = 9 to 10 yr.						

* E = equilibrium-controlled; I = instantaneous; K = kinetic.

** NA = not applicable.

⁺ $K^{eq} = (\theta \cdot K^f) / (\rho_b \cdot K^b)$.

In the upper left plot of Figure 15, we see the concentration drop due to the addition of clean-up chemicals: during 2 to 3 years for PCE (red), during 4 to 5 years for TCE (green), during 6 to 7 years for DCE (blue), and during 9 to 10 years for VC (dark khaki). In the upper right plot, we see the production of the environmentally safe product species associated with the addition of the clean-up chemicals from the instantaneous reactions (R5) through (R8). The bottom left plot shows the decrease of sorbed PCE and sorbed TCE due directly to desorption, i.e., (R1) and (R2), and indirectly to biodegradation and clean-up reactions. The bottom right plot shows that concentrations of all clean-up chemicals stay at zero concentration during this 10-year simulation. It is because they react with contaminant species in the PCE biodegradation chain as soon as they are added into the reaction system.

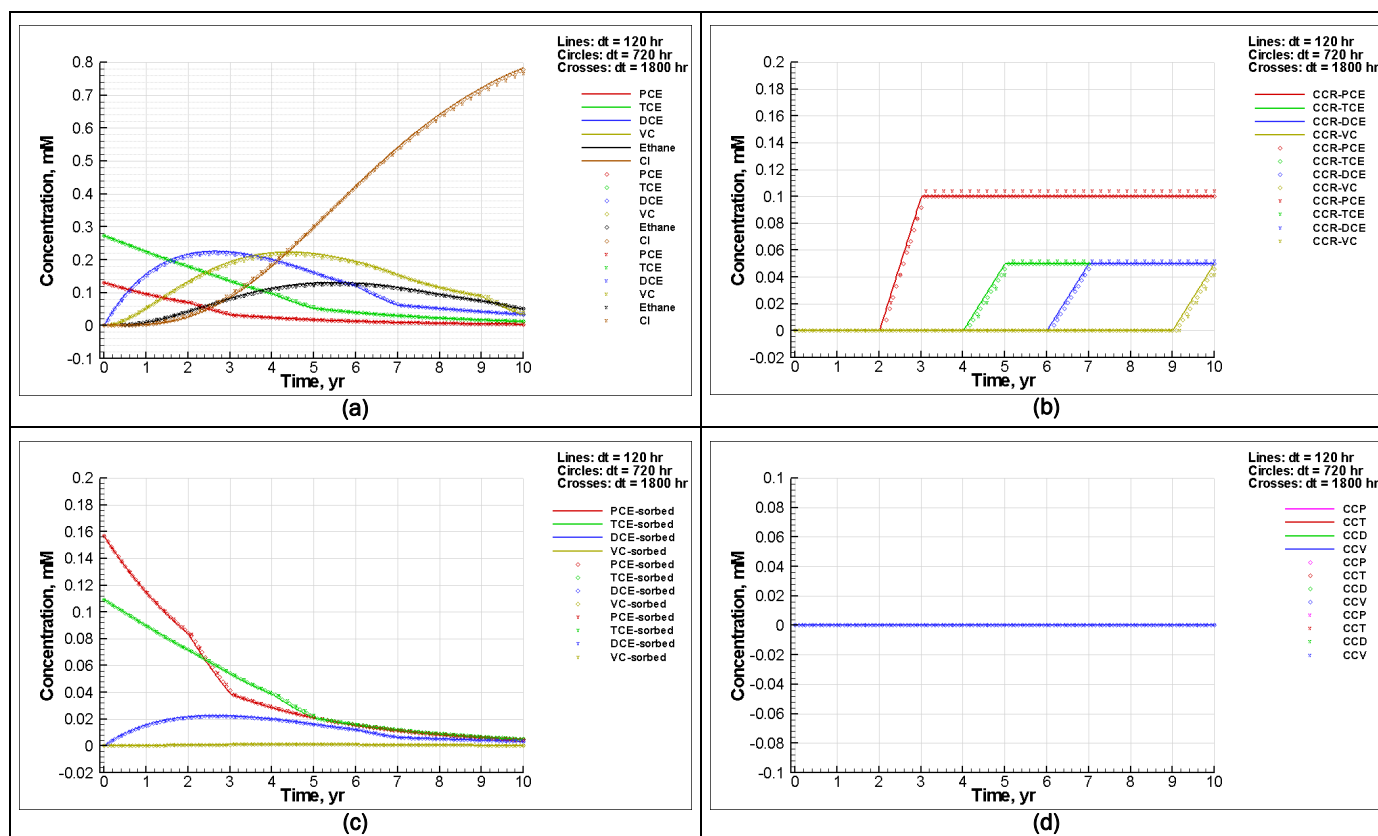


Figure 15. The computed concentration profiles of the 18 species in Reaction system 9 when the time-step sizes are set to 120, 720, and 1,800 hours

Problem 10. Hypothetical organic waste system

This problem was designed to demonstrate RBBGCS's capability of handling multi-phase, multi-sorption site reaction systems. It is also to demonstrate how the fate of dissolve organic waste in surface water systems may be simulated. Figure 16 depicts the hypothetical reaction network of this problem. The hypothetical reaction system for this problem includes four phases: a gas phase representing air, one solid phase representing bed sediment, and two aqueous phases representing the water column and the interstitial water between bed sediment grains. There are 10 reactions and 12 species considered. The 12 species are dissolved oxygen (DO), dissolved organic waste (DOW), dissolved residual (RS) from the aerobic biodegradation of DOW, two sorption sites (S1= and S2=) on the surface of the bed sediment, four adsorbed species (S1=OW, S2=OW, S1=RS, and S2=RS) on the surface of the bed sediment, dissolved organic waste in the interstitial water (DOW_I), dissolved residual in the interstitial water (RS_I), and oxygen in the air. Table 14 lists the 10 reactions taken into account in

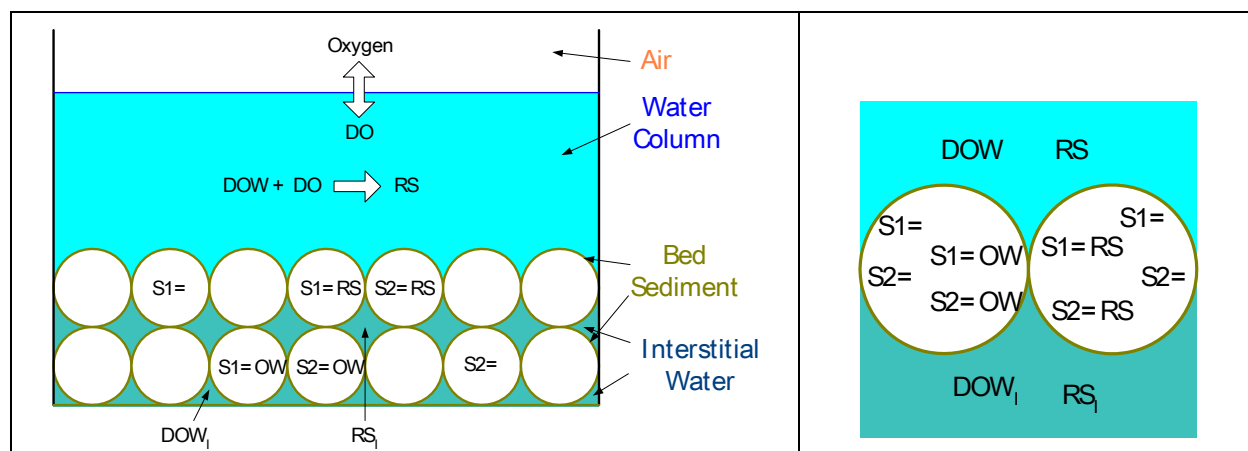


Figure 16. The hypothetical reaction system of Test Problem 10 (left), and adsorption of dissolved organic waste and biodegradation residual onto sorption sites S1= and S2= from both water column and interstitial water (right).

Table 14. Reaction system for Problem 10.

Reaction System 10				
ID	Formula	k^f	k^b	Rate
(R1)	$DO \leftrightarrow \text{Oxygen}$	0.048	0.0024	$\theta \cdot k^f \cdot [DO] - k^b \cdot P_{\text{oxygen}}$
(R2)	$DOW + DO \rightarrow RS^{**}$	0.0014	0	$\theta \cdot k^f \cdot [DOW]$
(R3)	$DOW + S1 \rightleftharpoons S1 = OW$	1.0	0.1	$k^f \cdot (\theta [DOW]) \cdot (\rho_b S_A [S1 =]) - k^b \cdot (\rho_b S_A [S1 = OW])$
(R4)	$DOW + S2 \rightleftharpoons S2 = OW$	0.001	1.0×10^{-4}	$k^f \cdot (\theta [DOW]) \cdot (\rho_b S_A [S2 =]) - k^b \cdot (\rho_b S_A [S2 = OW])$
(R5)	$RS + S1 \rightleftharpoons S1 = RS$	0.01	0.01	$k^f \cdot (\theta [RS]) \cdot (\rho_b S_A [S1 =]) - k^b \cdot (\rho_b S_A [S1 = RS])$
(R6)	$RS + S2 \rightleftharpoons S2 = RS$	0.1	0.1	$k^f \cdot (\theta [RS]) \cdot (\rho_b S_A [S2 =]) - k^b \cdot (\rho_b S_A [S2 = RS])$
(R7)	$DOW_I + S1 \rightleftharpoons S1 = OW$	1.0	0.1	$k^f \cdot (\theta_I [DOW_I]) \cdot (\rho_b S_A [S1 =]) - k^b \cdot (\rho_b S_A [S1 = OW])$
(R8)	$DOW_I + S2 \rightleftharpoons S2 = OW$	0.001	1.0×10^{-4}	$k^f \cdot (\theta_I [DOW_I]) \cdot (\rho_b S_A [S2 =]) - k^b \cdot (\rho_b S_A [S2 = OW])$
(R9)	$RS_I + S1 \rightleftharpoons S1 = RS$	0.01	0.01	$k^f \cdot (\theta_I [RS_I]) \cdot (\rho_b S_A [S1 =]) - k^b \cdot (\rho_b S_A [S1 = RS])$
(R10)	$RS_I + S2 \rightleftharpoons S2 = RS$	0.1	0.1	$k^f \cdot (\theta_I [RS_I]) \cdot (\rho_b S_A [S2 =]) - k^b \cdot (\rho_b S_A [S2 = RS])$

* E = equilibrium-controlled; I = instantaneous; K = kinetic.
 ** Hemond, H .F. and E .J. Fechner-Levy. 2000. Chemical Fate and Transport in the Environment. Academic Press, 433 pp.
 NA = not applicable.

the biogeochemical reaction model. Among the 10 reactions, (R1) is the volatilization reaction of dissolved oxygen (DO); (R2) represents the biodegradation of dissolved organic waste (DOW) under an aerobic condition, where RS is the residual after biodegradation; (R3) and (R4) are adsorption of DOW onto sorption sites S1= and S2=, respectively, on bed sediments; (R5) and (R6) are adsorption of RS onto sorption sites S1= and S2=, respectively; (R7) and (R8) are adsorption of the dissolved organic waste in the interstitial water (DOW_I) onto sorption sites S1= and S2=, respectively; (R9) and (R10) are adsorption of the biodegradation residual in the interstitial water (RS_I) onto sorption sites S1= and S2=, respectively. Among all the adsorption/desorption reactions, the adsorption of DOW onto S1= (i.e., (R3) and (R7)) is the fastest and DOW onto S2= (i.e., (R4) and (R8)) the slowest.

The entire volume of the reaction system was 1.0 L. The volume of water in water column was 900 mL, and the volume of interstitial water 20 mL. The mass of bed sediment was 450 g. Therefore, the equivalent water contents of the water column (i.e., θ) and the interstitial water (i.e., θ_i) were 0.9 and 0.02, respectively; and the equivalent bulk density of the bed sediment was 2,500 g/L ($= 450/(1000-900-20)$). The specific surface area associated with sorption sites S1= and S2= was 1.0 dm²/g each. The partial pressure of oxygen in the air was fixed at 0.2 atm. Initially, both of the concentrations of S1= and S2= were set to 0.1 mole/dm³; the concentration of DOW was set to 10⁻² M; and the concentrations of the other species were set to zero.

A simulation of 1,000 minutes was conducted. A sensitivity analysis indicated that using a constant time-step of 0.1 minute would provide accurate numerical solutions. Figures 17 through 19 depict the evolution of species concentration with time during the simulation. These figures show how the dissolution of oxygen through (R1) triggers the biodegradation of DOW and the distributions of organic waste and biodegradation residual among the water column, bed sediment, and interstitial water via adsorption/desorption. From Figure 17, [DOW] drops fast in the first minute (right plot) due to quick adsorption onto S1=, which is supported by the quick increase of [S1=OW] (right plot, Figure 18). The continuous and small increase of [RS] and [RS_I] indicates the slow biodegradation of DOW.

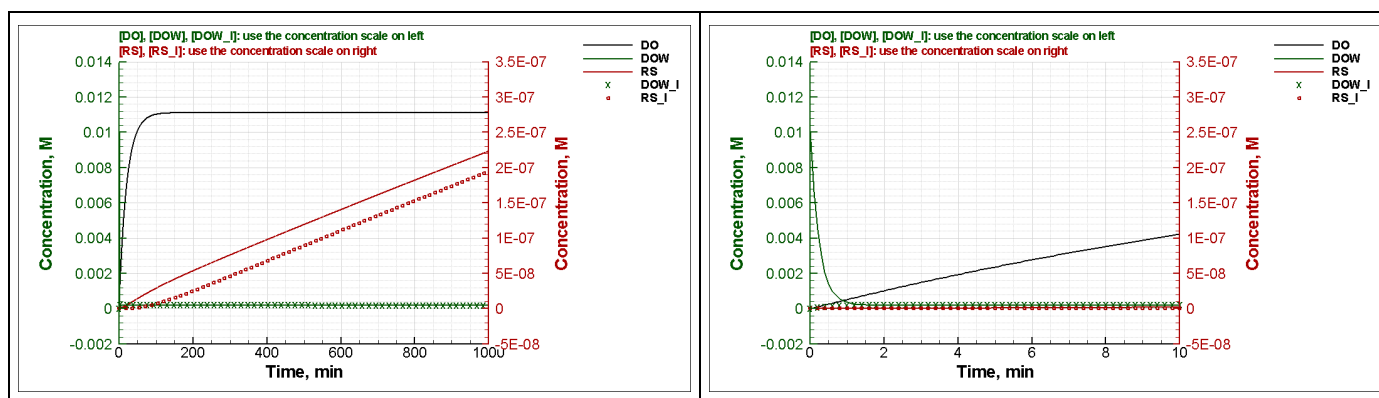


Figure 17. The concentration evolution of species in the aqueous phases: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.1 mole/dm² for both S1= and S2=.

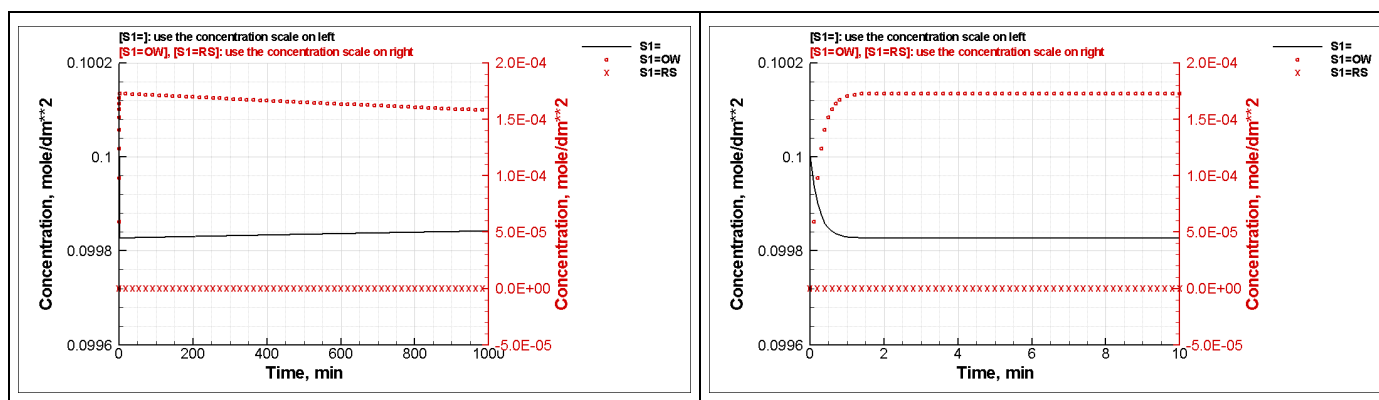


Figure 18. The concentration evolution of species associated with sorption site 1: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.1 mole/dm² for both S1= and S2=.

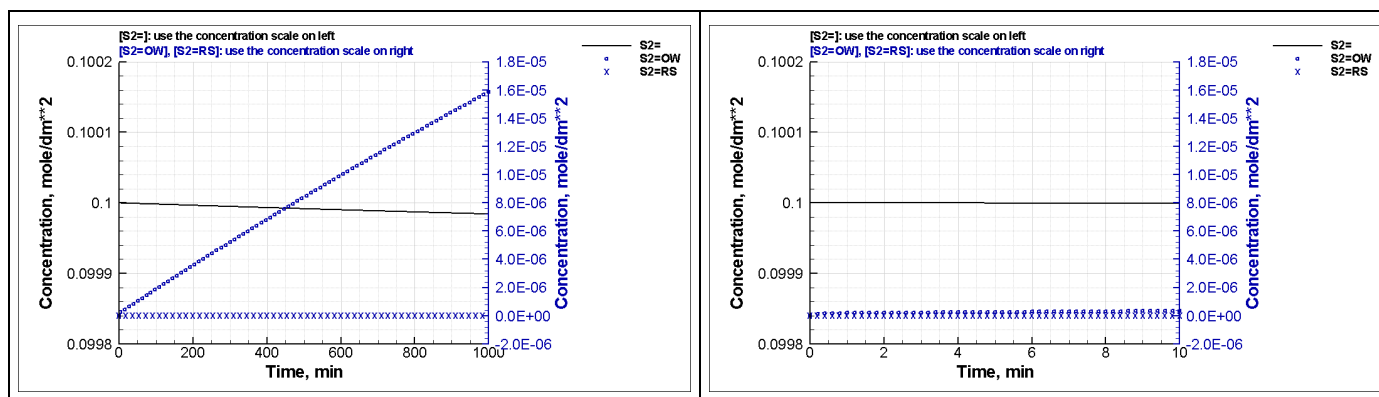


Figure 19. The concentration evolution of species associated with sorption site 2: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.1 mole/dm² for both S1= and S2=.

Moreover, [DO] does not vary after time = 250 min, indicating that (R1) has reached equilibrium though the other reactions have not. The stable change of the other species concentrations also supports this. In Figure 18, the slow drop of [S1=OW] after the quick increase in the beginning of the simulation indicates the desorption of organic waste from the bed sediment into the water column to proceed biodegradation, which is due to the continuous dissolution of oxygen from air. Figure 19 shows an increase of [S2=OW] throughout the simulation, indicating that the slow adsorption of organic wastes onto S2= also has an effect on the desorption of S1=OW.

Figures 20 through 22 depict species concentrations when the capacities change to 0.001 and 0.002 mole/dm² for S1= and S2=, respectively. These figures, when compared with Figures 17 through 19, demonstrate the influence of sorption capacity on concentration distribution.

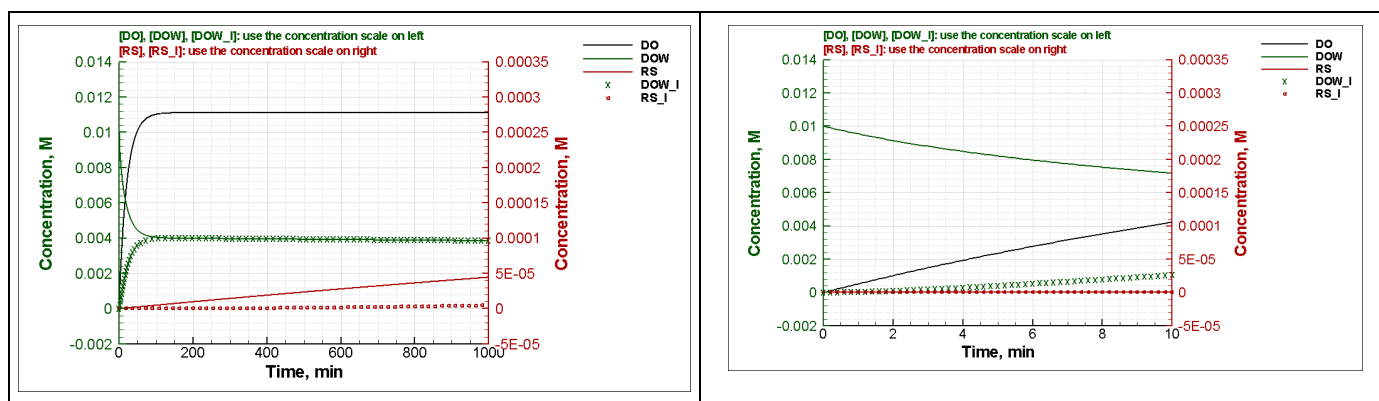


Figure 20. The concentration evolution of species in the aqueous phases: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.001 and 0.002 mole/dm² for S1= and S2=, respectively.

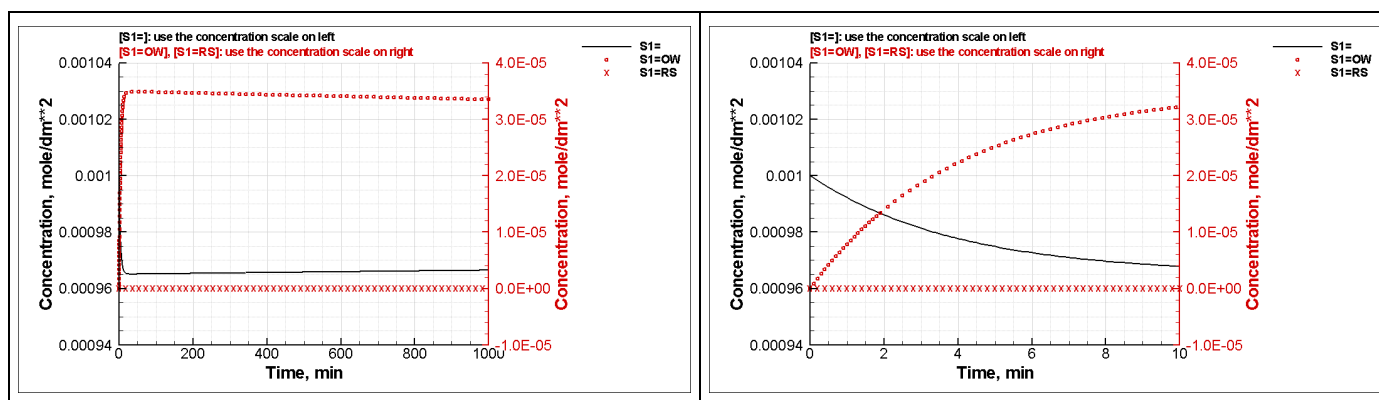


Figure 21. The concentration evolution of species associated with sorption site 1: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.001 and 0.002 mole/dm² for S1= and S2=, respectively.

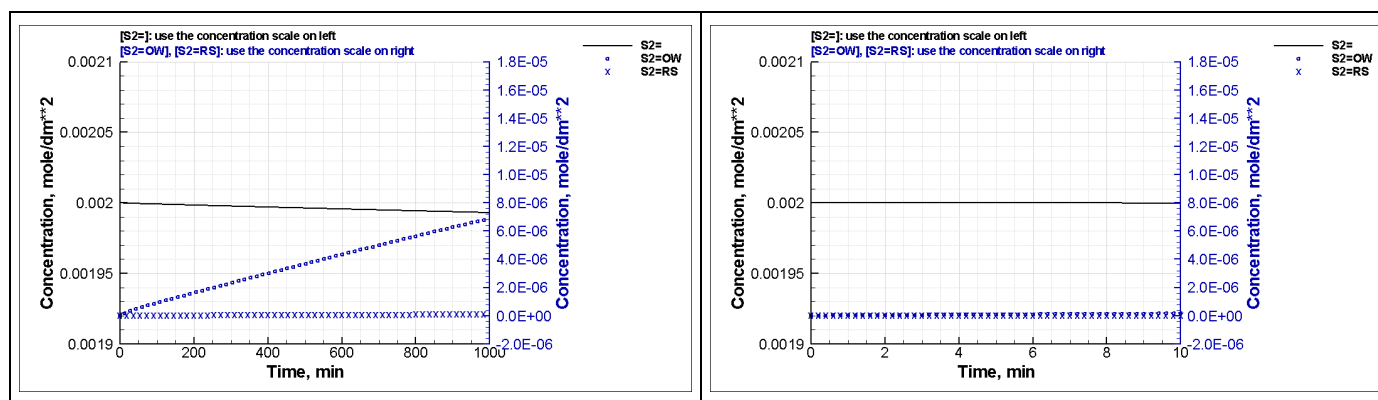


Figure 22. The concentration evolution of species associated with sorption site 2: from time = 0 to 1000 minutes (left) and from time = 0 to 10 minutes (right) when the capacities are 0.001 and 0.002 mole/dm² for S1= and S2=, respectively.

4 Summary

A newly developed reaction-based biogeochemical simulator, RBBGCS, is presented in this report. RBBGCS was designed to be a generic biogeochemical modeling tool that can be used to simulate the evolution of species concentration distribution within non-transport or batch systems. Although it is currently incorporated into ADH to perform reactive transport, it can be coupled with any transport models for reactive transport modeling in surface and subsurface systems. We detailed the two components of RBBGCS: a nine-step preprocessor to generate a reaction-based DAE system automatically and systematically, and the solution techniques used to solve the DAE system. These solution techniques include using the Newton's method plus a full-pivoting director solver to solve the non-linear DAE system, choosing an adequate set of constraint equations to deal with instantaneous reactions, and using the LR method to handle zero-order reactions. We employed various test examples to verify and demonstrate RBBGCS' capabilities in simulating reaction systems of different levels of complexity. RBBGCS is capable of effectively and efficiently simulating reaction systems that contain equilibrium-controlled, instantaneous, and kinetic reactions.

Future advancements may include the following.

1. Develops a graphic user interface (GUI) to help the user translate a system of equations from a conceptual model to the DAE;
2. Improves the time integration scheme used for reaction computation, e.g., adaptive time steps;
3. Accounts for temperature effect;
4. Accounts for the change of phase attribute due to reactions.

References

- Bethke, C. M. 2008. *Geochemical and Biogeochemical Reaction Modeling*, 2nd edition, Cambridge University Press, 564 pp.
- Chilakapati A., T. Ginn, and J. E. Szecsody. 1998. An analysis of complex reaction networks in groundwater modeling. *Water Resources Research* 34:1767–1780.
- Clement, T. P. 1997. *RT3D: A modular computer code for simulating reactive multi-species transport in 3-dimensional groundwater systems*. Technical Report PNNL-11720, Pacific Northwest National Laboratory.
- Fang, Y., G-T. Yeh, and W. D. Burgos. 2003. A general paradigm to model reaction-based biogeochemical processes in batch systems. *Water Resources Research* 39(4):1083. doi:10.1029/2002WR001694.
- Glynn, G. M., S. E. Howington, W. L. Murphy, J. P. Hallberg, M. Zakikhani, R. S. Maier, and B. A. Abraham. 2001. *Groundwater flow and contaminant transport modeling near the Peele Dixie well field*. ERDC TR-01-7. Vicksburg, MS: U.S. Army Engineer Research and Development Center.
- Lichtner, P. C. 1996. Continuum formulation of multicomponent-multiphase reactive transport. In *Reactive transport in porous media*, ed. P. C. Lichtner, C. I. Steefel, and E. H. Oelkers, 1–79. Washington, DC: Reviews in Mineralogy 34, Mineralogical Society of America.
- McNaught A. D., and A. Wilkinson. 1997. *IUPAC Compendium of Chemical Terminology*, The Gold Book, 2nd ed., Blackwell Science.
- Neuhaus, G. W., and W. F. Maier. 1996. A general integral method for the determination of empirical reaction rates and parameters. *Chemical Engineering and Technology* 19:249–262.
- Press, W. H, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical recipes in Fortran*, 2nd ed., Cambridge University Press, 963 pp.
- Rubin, J. 1983. Transport of reacting solutes in porous media: Relation between mathematical nature of problem formulation and chemical nature of reactions. *Water Resources Research* 19:493–502.
- Steefel, C. I., and V. P. Cappellen. 1998. Preface: Reactive transport modeling of natural systems. *Journal of Hydrology* 209:1–7.
- Yeh, G-T., W. D. Burgos, and J. M. Zachara. 2001. Modeling and measuring biogeochemical reactions: System consistency, data needs, and rate formulation. *Advances in Environmental Research* 5:219–237.
- Yeh, G-T. 2001. *Computational subsurface hydrology-reactions, transport, and fate*. Norwell, MA: Kluwer Academic Publishers.

Appendix A: RBBGCS Input Guide

When the user runs the executable of the stand-alone RBBGCS, he/she will be asked to provide a super file name. The super file is a plain text file that includes five lines as given in the table below. Details of the contents of both the input and auxiliary input files are given in the Sections A.1 and A.2. The seven aqueous species example used in Section 2.2 to explain the nine-step preprocessor is taken as an example to demonstrate the contents included in the input and auxiliary input files.

Line ID	Field 1	Field 2	Description
1	INPU	Input_file_name	Name of the input file that defines the reaction network of the biogeochemical system
2	INPS	Auxiliary_input_file_name	Name of the auxiliary input file that contains data that is specified in the ADH bc file
3	OUTP	Preprocessing_output_file_name	Name of the output file that contains the preprocessing result
4	CONC	Concentration_output_file_name	Name of the output file that contains the initial and file concentration distributions of the specified transient simulation
5	ENDO	---	End of the super file

A.1 Cards in the Input File

File Comments			
Field	Type	Value	Description
1	char	TI	Card group identifier
2	char	TLE	Card type
3	char	String of size 130	Comments to describe the input file
Time Stepping			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	DTR	Card type
3	real	+	Time stepping for reaction computation
Phase Names			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	PNM	Card type
3	int	+	Number of phases (NPHASE)
Followed by NPHASE lines, where each line contains the following two fields for a phase			
Field	Type	Value	Description
1	int	+	ID of the phase being input
2	char	string of size 30	Name of the phase
Phase Attributes			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	PAT	Card type
3	int	+	ID of the phase being input
4	int	+	Attribute of this phase: Gas.
		1	This is an aqueous phase
		2	This is a solid phase
		3	This is a gas phase
Phase Mobility			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	PMO	Card type
3	int	+	ID of the phase being input
4	int	+	Indicator of the mobility of the phase: Gas.
		1	This is a mobile phase
		2	This is an immobile phase

Species Names			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SNM	Card type
3	int	+	Number of species (NCS)
4	int	+	Number of mobile species
5	int	0, +	Number of immobile species
Followed by NCS lines, where each line contains the following two fields for a phase			
Field	Type	Value	Description
1	int	+	ID of the species being input
2	char	string of size 30	Name of the species
Species Attributes			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SAT	Card type
3	int	+	ID of the chemical species being input
4	int	+	ID of the phase that the species is associated with
5	int	+	Type of the species
		1	This is an aqueous species
		2	This is a sorbed or sorbing species
		3	This is an ion-exchanged species
		4	This is a precipitated species
		5	This is a gas species
		...	To be defined as needed
6	real	-, 0, +	Electric charge of the species
Species Activity Coefficient			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SAC	Card type
3	int	+	ID of the chemical species being input
4	int	-, 0, +	ID of the model used for computing the activity coefficient of the species
		-1	Activity is set to 1, e.g., for precipitated species
		0	Activity coefficient is set to 1
		1	Activity coefficient is computed using the Güntelberg approximation (Table 3.3, p.135, "Aquatic Chemistry" by W. Stumm & J. J. Morgan, 2 nd ed., 1981)
		2	Activity coefficient is computed using the Davies approximation (Table 3.3, p.135, "Aquatic Chemistry" by W. Stumm & J. J. Morgan, 2 nd ed., 1981)
		3	Activity coefficient is computed using the extended Debye-Hückel approximation (Tables 3.3-3.4, p.135, "Aquatic Chemistry" by W. Stumm & J. J. Morgan, 2 nd ed., 1981)

		4	Activity coefficient is computed using the 1 st user-defined approximation (the user-defined approximation must be coded into the model)
		5	Activity coefficient is computed using the 2 nd user-defined approximation (the user-defined approximation must be coded into the model)
		...	To be defined as needed
5	int	0, +	Number of parameters used for computing the activity coefficient of the species (NACP) Note: it is greater than zero when user-defined approximations are considered

Followed by one line that includes NACP fields to represent the NACP parameters used for computing activity coefficient when user-defined approximations are considered

Field	Type	Value	Description
1	real	-, 0, +	1 st parameter
...	real	-, 0, +	...
NACP	real	-, 0, +	NACP th parameter

1st-Order Decay Parameters

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SDY	Card type
3	int	+	ID of the chemical species being input
4	real	+	1 st -order decay constant for the species

Suspected Component Species

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SCA	Card type
3	int	+	ID of the suspected component species being input

Fixed Concentration Species

Note: This species must be a component species, such that the conservation equation associated with the species can be replaced by the equation of fixed concentration.

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SFC	Card type
3	int	+	ID of the fixed-concentration species being input
4	real	+	Fixed concentration for the designated species

Ion-Exchange Site Names

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	IEN	Card type
3	int	+	Number of ion-exchange sites (NIESITE)

Followed by NIESITE lines, where each line contains the following two fields for a phase

Field	Type	Value	Description
1	int	+	ID of the ion-exchange site being input

2	char	string of size 30	Name of the ion-exchange site
---	------	-------------------	-------------------------------

Ion-Exchange Site Attributes

Note: It is assumed that every ion-exchange site is fully occupied by possible ion-exchanges species

< Example > Three ion-changed species: ($\equiv S - Na$), ($\equiv S_2 - Ca$), ($\equiv S_2 - Mg$)

$$\rightarrow SiteCapacity = C(\equiv S - Na) * 1 + C(\equiv S_2 - Ca) * 2 + C(\equiv S_2 - Mg) * 2$$

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	IEA	Card type
3	int	+	ID of the ion-exchange site being input
3	int	+	Number of ion-exchanged species involved in the ion-exchange sites (NCS_IE)

Followed by NCS_IE lines, where each line contains the following two fields for a phase

Field	Type	Value	Description
1	int	+	ID of the ion-exchanged species that is designated as one of the NCS_IE species involved in the site
2	int	+	ID of the aqueous phase species that is associated with the ion-exchanged species specified in Field 1

Reaction Names

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	RNM	Card type
3	int	+	Number of reactions (NRX)

Followed by NRX lines, where each line contains the following two fields for a phase

Field	Type	Value	Description
1	int	+	ID of the reaction being input
2	char	string of size 30	Name of the reaction

Reaction Attributes

Note: every reaction, whichever type it is, must be further described with a RBRR1, RBRR2, RBRE1, or RBRE2 card.

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	RAT	Card type
3	int	+	ID of the reaction being input
4	int	+	1 st attribute of the reaction
		1	This is an equilibrium-controlled (fast reversible) reaction
		2	This is a kinetic (slow reversible/irreversible) reaction
		3	This is an instantaneous (fast irreversible) reaction
5	int	+	1 st attribute of the reaction
		1	This is an aqueous complexation/speciation reaction
		2	This is a sorption/desorption reaction
		3	This is an ion-exchange reaction
		4	This is a precipitation-dissolution reaction
		5	This is a volatilization reaction
		6	Otherwise

Reaction Stoichiometry			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	RST	Card type
3	int	+	ID of the reaction being input
4	int	+	Number of species involved in the reaction (NCS_RX)

Followed by NCS_RX lines, where each line contains the following two fields for a phase

Field	Type	Value	Description
1	int	+	ID of the species corresponding to one of the NCS_RX species involved in the reaction
2	real	+	Stoichiometric coefficient associated with the species specified in Field 1
3	real	-, +	Indicator of whether the species specified in Field 1 is a reactant or a product species in the reaction
		-1.0	This species is on the reactant side
		1.0	This species is on the product side

Reaction Rate Parameters (Case 1: based on the collision theory)

Note: RR1 card is used for an instantaneous (fast irreversible) reaction, where the logarithm values of the forward and backward rate coefficients are usually set to 10.0 and -10.0, respectively.

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	RR1	Card type
3	int	+	ID of the reaction being input
4	real	-, 0, +	Logarithm of the forward reaction rate coefficient of the reaction
5	real	-, 0, +	Logarithm of the backward reaction rate coefficient of the reaction

Reaction Rate Parameters (Case 2: using an empirical formula)

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	RR2	Card type
3	int	+	ID of the reaction being input
4	int	+	ID of the empirical formula used to compute reaction rate for the reaction (the empirical formula must be coded into the model)
5	int	+	Number of parameters used in the empirical formula for the reaction (NPRX)

Followed by one line that contains NPRX fields to represent the NPRX parameters

Field	Type	Value	Description
1	real	-, 0, +	1 st parameter
...	real	-, 0, +	...
NPRX	real	-, 0, +	NPRX th parameter

Equilibrium Parameters (Case 1: based on the collision theory)

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	RE1	Card type
3	int	+	ID of the reaction being input

4	real	-, 0, +	Logarithm of the equilibrium constant of the reaction
---	------	---------	---

Equilibrium Parameters (Case 2: using an empirical formula)

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	RE2	Card type
3	int	+	ID of the reaction being input
4	int	+	ID of the empirical formula used to describe equilibrium for the reaction (the empirical formula must be coded into the model)
5	int	+	Number of parameters used in the empirical formula for the reaction (NPRE)

Followed by one line that contains NPRE fields to represent the NPRE parameters

Field	Type	Value	Description
1	real	-, 0, +	1 st parameter
...	real	-, 0, +	...
NPRE	real	-, 0, +	NPRE th parameter

Nonlinear Convergence Parameters

Non-linear convergence is obtained if the following inequality exists for all species:

$$\frac{|C^{n+1} - C^n|}{RTOL \cdot \max[|C^{n+1}|, |C^n|] + ATOL} < 1$$

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	NLC	Card type
3	int	+	Maximum number of iteration allowed for the non-linear iteration loop used to compute biogeochemistry
4	rel	+	Absolute error-related parameter used to determine non-linear convergence (ATOL)
5	rel	+	Relative error-related parameter used to determine non-linear convergence (RTOL)

Concentration Constraint

Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	CCN	Card type
3	rel	+	Minimum chemical species concentration value below which the chemical concentration is considered zero
4	rel	+	Maximum chemical species concentration value allowed for computing reaction rates

End of File

Field	Type	Value	Description
1	char	EN	Card group identifier
2	char	DDD	Card type

According to the card description above, the input file for the seven species example used in Section 2.2 may appear to be as follows.

```

TI TLE  NTA REACTION SYSTEM (7 species)
TI TLE  Last Update: June 01, 2009
RB DTR  1.0E-2
RB PNM  1
      1  MOBILE_AQUEOUS_PHASE
RB PAT  1 1
RB PMO  1 1
RB SNM  7 7 0
      1  H
      2  NTA
      3  HNTA
      4  Co
      5  CONTA
      6  B
      7  P
RB SAT  1 1 1 0.0E0
RB SAT  2 1 1 0.0E0
RB SAT  3 1 1 0.0E0
RB SAT  4 1 1 0.0E0
RB SAT  5 1 1 0.0E0
RB SAT  6 1 1 0.0E0
RB SAT  7 1 1 0.0E0
RB SCA  1
RB SCA  2
RB SCA  4
RB SAC  1 -1 0
RB SAC  2 -1 0
RB SAC  3 -1 0
RB SAC  4 -1 0
RB SAC  5 -1 0
RB SAC  6 -1 0
RB SAC  7 -1 0
RB SDY  1 0.0E0
RB SDY  2 0.0E0
RB SDY  3 0.0E0
RB SDY  4 0.0E0
RB SDY  5 0.0E0
RB SDY  6 0.0E0
RB SDY  7 0.0E0
RB RNM  7
      1  EQUILIBRIUM-CONTROLLED_REACTION_1
      2  EQUILIBRIUM-CONTROLLED_REACTION_2
      3  KINETIC_REACTION_1
      4  KINETIC_REACTION_2
      5  EQUILIBRIUM-CONTROLLED_REACTION_3
      6  INSTANTANEOUS_REACTION_1
      7  INSTANTANEOUS_REACTION_2
RB RAT  1 1 1
RB RAT  2 1 1
RB RAT  3 2 1
RB RAT  4 2 1
RB RAT  5 1 1
RB RAT  6 3 1
RB RAT  7 3 1
RB RST  1 3
      1 1 -1
      2 1 -1
      3 1 1
RB RE1  1 0.0E0
RB RST  2 3
      5 1 -1
      4 1 1
      2 1 1
RB RE1  2 0.0E0
RB RST  3 4
      1 1 -1
      5 1 -1
      3 1 1
      4 1 1
RB RR1  3 -1.0E0 -1.0E0
RB RST  4 3
      1 1 -1
      3 1 -1
      6 1 1
RB RR1  4 -1.0E0 -1.0E0
RB RST  5 5

```

```

1 1 -1
4 1 -1
2 2 -1
3 1 1
5 1 1
RB RST 6 3
1 1 -1
6 1 -1
7 1 1
RB RR1 6 10.0E0 -10.0E0
RB RST 7 4
1 1 -1
5 1 -1
3 1 1
4 1 1
RB RR1 7 10.0E0 -10.0E0
RB RE1 5 0.0E0
RB NLC 500 1.0E-10 1.0E-6
RB CCN 1.0E-12 1.0E0
EN DDD

```

A.2 Cards in the Auxiliary Input File

File Comments			
Field	Type	Value	Description
1	char	TI	Card group identifier
2	char	TLE	Card type
3	char	String of size 130	Comments to describe the input file
Total Simulation Time			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	DTT	Card type
3	real	+	Simulation time for stand-alone reaction computation
Bulk Density			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	BDN	Card type
3	real	+	Bulk density
Phase Attribute Property			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	PAT	Card type
3	int	+	ID of the phase being input
4	real	+	Attribute property of this phase: It is porosity associated with the phase if it is an aqueous phase [L ³ pore volume/L ³ bulk volume]; It is surface area density associated with the phase if it is a solid phase [L ² surface area/M dry bulk]; It is porosity associated with the phase if it is a gas phase [L ³ pore volume/L ³ bulk volume].

Initial Concentrations of Species			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SIC	Card type
3	int	+	ID of the species being input
4	real	+	The given initial concentration of the species
Source Rate of Species			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	SRC	Card type
3	int	+	ID of the species being input
4	int	+	Series ID number for the source rate associated with the species
X-Y Series			
Field	Type	Value	Description
1	char	RB	Card group identifier
2	char	XYS	Card type
3	int	+	ID of the series being input
4	int	+	Number of data points included in the series (NPOINT)
Followed by NPOINT lines, where each line contains the following two fields for a data point			
Field	Type	Value	Description
1	real	X	X value of the data point
2	real	Y	Y value of the data point
End of File			
Field	Type	Value	Description
1	char	EN	Card group identifier
2	char	DDD	Card type

The auxiliary input file for the seven species example used in Section 2.2 may appear to be as follows.

```

TI TLE  NTA REACTION SYSTEM (7 species)
TI TLE  Last Update: June 01, 2009
RB DTT  100.0E0
RB BDN  1.5E0
RB PAT  1 0.3E0
RB SIC  1 1.0E-2
RB SIC  2 1.0E-2
RB SIC  3 1.0E-2
RB SIC  4 1.0E-2
RB SIC  5 1.0E-2
RB SIC  6 1.0E-2
RB SIC  7 1.0E-2
EN DDD

```

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) May 2010		2. REPORT TYPE Final report		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE A Generic Reaction-Based BioGeoChemical Simulator (RBBGCS), Version 1.0				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Hwai-Ping Cheng, Stacy E. Howington, Matthew W. Farthing, Christian J. McGrath, and Jing-Ru C. Cheng				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Engineer Research and Development Center Coastal and Hydraulics Laboratory, Environmental Laboratory, and Information Technology Laboratory 3909 Halls Ferry Road Vicksburg, MS 39180-6199				8. PERFORMING ORGANIZATION REPORT NUMBER ERDC TR-10-5	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Headquarters, U.S. Army Corps of Engineers Washington, DC 20314-1000				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This report presents a generic reaction-based biogeochemical simulator (RBBGCS) that was developed as part of the advancement of the subsurface reactive transport capability in the Adaptive Hydrology/Hydraulics (ADH) model. RBBGCS has been incorporated into ADH to model subsurface reaction transport. The simulator can also be coupled with other transport models to perform reactive transport modeling in surface and subsurface systems. RBBGCS can model geochemical/biogeochemical reactions that are equilibrium controlled (fast reversible), instantaneous (fast irreversible), and kinetic (slow reversible or irreversible). It has a preprocessor that automatically and systematically produces reaction-based differential-algebraic equations (DAE) as the reaction governing equations and a solver that solves the set of governing equations for the concentration distribution of chemical species. It allows both user-specified empirical equation and formulation based on the collision theory to be used to describe reaction equilibrium and reaction rate(s). The numbers of chemical species, biogeochemical reactions, and porous medium phases that may be defined for a modeled system are unrestricted, limited only by the computational resources that are available. (Continued)					
15. SUBJECT TERMS Biogeochemical simulator Geochemical reaction Reaction-based Differential-algebraic equation Numerical modeling Reactive transport					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 91	19a. NAME OF RESPONSIBLE PERSON
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED			19b. TELEPHONE NUMBER (include area code)

14. ABSTRACT (Concluded).

This report describes the development of RBBGCS, including a nine-step preprocessor to generate reaction-based DAE systems and solution techniques to solve the DAE system. The preprocessor constructs a valid reaction network and produces the associated governing equations, which can save modelers a significant amount of time when modeling complex reaction systems. The solution technique section details (1) the computational procedures in RBBGCS, (2) the DAE system when man-induced sources exist, (3) Newton's method to solve DAE systems, (4) implementation of the constraint equations in DAE systems, and (5) treatment for zero-order reactions. Multiple test examples are presented to verify and demonstrate RBBGCS' capabilities in solving complex geochemical/biogeochemical reaction problems. RBBGCS development serves as a guide for continued model development for coupling with ADH and other transport models to perform reactive transport simulation.