Introduction of a new platform for parameter estimation of kinetically complex environmental systems

CrossMark

Mehdi Gharasoo [a, *], Martin Thullner [b], Martin Elsner [a, c]

[a] Helmholtz Zentrum München, Institute of Ground Water Ecology, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany
[b] Department of Environmental Microbiology, Helmholtz Centre for Environmental Research-UFZ, Permoserstraße 15, 04318 Leipzig, Germany
[c] Institute of Hydrochemistry, Technical University of Munich, Marchioninistr. 17, 81377 Munich, Germany

ABSTRACT

A modeling framework (ReKinSim - Reaction Kinetics Simulator) is introduced, within which biogeochemical reactions in environmental systems can be described and inversely fitted to experimental data. Three key features of this simulation environment are: (1) a generic mathematical tool for solving sets of unlimited, arbitrary, non-linear ordinary differential equations; (2) no limitation to the number or type of reactions or other influential dynamics (e.g., isotope fractionation or small-scale mass-transfer limitations); (3) an easy to use and flexible module for nonlinear data-fitting. It allows users to easily define any kinetic model by a set of biogeochemical reactions relevant to the experimental application and to obtain the values of the kinetic parameters by fitting of the model to data. By allowing users to include the environmentally related processes and solving them along with the chemical kinetics, ReKinSim helps the user to elucidate the extent that these processes are controlled by factors other than kinetics. The novelty of the presented program primary lays in its unique combination of flexibility, computational efficiency and user-friendliness. ReKinSim's usability is showcased by four case studies of varying complexity, and compared against a set of currently available modeling tools.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Over the past few decades significant improvements have been made in understanding the processes controlling the biogeochemical cycling of elements in environmental systems. Recent developments in the field of computational technology and algorithms have further helped scientists to simulate complicated biochemical processes and to enhance the exchange of information between recent scientific questions and models (e.g., Cirpka and Valocchi, 2007; Eckert et al., 2013; Hunkeler et al., 2009; König et al., 2017; Rosenzweig et al., 2014; Thullner et al., 2008). For a more quantitative analysis of such processes in natural or man-made environmental systems, it is important to acquire a fundamental understanding of the kinetics of degradation mechanisms and their dependency on environmental factors (Or et al., 2007; Thullner et al., 2005, 2007; Van Loosdrecht et al., 1990). Such mechanisms and their interplay with other biogeochemical processes are described by a sequence of reactions/transformations which are then translated into mathematical formulations. Novel conclusions can be drawn when hypothetical descriptions of such mechanisms (usually in the form of models) are reconciled with experimental evidence. To this end, computational tools play a key role in stating the hypothetical functional relationships in the form of mathematical models and minimizing the error between suggested formulations and the observed data. For such a purpose, a tool for kinetic modeling and data fitting must (i) carry sophisticated algorithms equipped with the most recent updates and improvements, (ii) not be computationally elaborate, and (iii) be sufficiently versatile for implementation.

The current advances in kinetics modeling have often led to stand-alone platforms which are not modular, with outcomes that cannot be easily exported and linked to other codes (e.g., AQUASIM, Reichert, 1994). Furthermore, some of the programs suffer from the lack of an easily-operated interface and cannot be handled conveniently by scholars who are not strong modelers/informaticians. In special cases, the integration of numerous add-ons and often superfluous options has made some of the programs difficult to operate. Versatility is another issue in many of programs (e.g., ASM3, Iacopozzi et al. (2007), Global KinTek Explorer, Johnson et al.

(2009), and Rayleigh Tool, Höhener and Atteia (2014) and Morrill et al. (2006)) that are confined to the solution of a few specific reaction mechanisms and thus are difficult to employ for a wider range of environmentally associated problems with more complicated dynamics. Addressing the above limitations while keeping the advantages of previously presented programs, we felt the need for a new program which better tackles environmentally related research questions. To that end, a newly developed computational tool (ReKinSim - Reaction Kinetics Simulator) is brought forward in this paper that is accurate, adjustable, user-friendly, and reasonably quick. The presented tool (ReKinSim) has been designed with the purpose of offering an easy-to-operate while flexible platform, free from extra mathematical complexities that are often confusing and sometimes misleading. Flexibility is another key attribute of ReKinSim, so there are no limitations in defining the kinetic mechanisms of interest or the actual number of chemicals in a given system. ReKinSim is coded in MATLAB language and the use of MATLAB numerical modules contributes greatly to its performance, extensibility, and competence. The potential relevance of the presented tool for biochemical/environmental systems is demonstrated and verified through several examples. Some case studies have previously been presented and served as challenging test cases using other optimization tools, including AQUASIM and Global KinTek Explorer. We revisited these case studies to compare the results from ReKinSim with those reported. The rest of the case studies were designed to test ReKinSim and are, in comparison, more complicated and challenging.

The overall aim of this contribution is to introduce a comprehensive tool for modeling of complex reaction mechanisms, enabling environmental engineers and biochemists to understand and quantify reaction kinetics/dynamics associated with their systems of interest. In this respect, ReKinSim can in principle be employed as an exploratory tool (e.g. using trial and error methods) to unravel the kinetics of the processes that are suspected to take place in a system. However, its main usage is to estimate the values of kinetic parameters by fitting a reaction mechanism to a set of experimental data. ReKinSim is developed with simplicity of usage in mind and unlike some other similarly oriented approaches (e.g., Ferrai et al., 2010; Fang et al., 2009; Kneis et al., 2017; Iacopozzi et al., 2007; Morrill et al., 2006; Regnier et al., 2002) it provides a simpler framework with a more approachable and less complicated set of options that lets scientists focus more on experimental and theoretical features of their research, and less on mathematical and numerical aspects. It is the goal of the program ReKinSim to allow users to specify models as freely as possible, to keep the complexities to a minimum, and to estimate with acceptable precision the values of kinetic parameters.

2. Fundamentals of using ReKinSim

The definition of reaction kinetics in ReKinSim is straightforward. The conceptual kinetic model, i.e. the definition of the influential processes (e.g., biogeochemical reactions) and the kinetic description of them in form of ordinary differential equations (ODEs), leads to a set of equations that need to be supplied to ReKinSim by user. For predictive reaction modeling only the relevant set of ODEs is required, which then can be solved by ReKinSim provided the values of all kinetic/dynamic parameters are already declared. For parameter estimation, in addition to a clear description of the mechanisms in form of ODEs, a reasonable number of fitting parameters must be specified for which the model equations are fitted to experimental data. It is worth mentioning that ReKinSim is a kinetic modeling tool and does not provide solutions to dynamics in a spatially resolved domain, meaning that processes like hydrodynamic flow and solute transport are exempted.

Nevertheless, for inclusion of such transport mechanisms ReKinSim code-based module can be easily coupled with any transport code via operator splitting techniques (Valocchi and Malmstead, 1992).

ReKinSim is delivered in two modes: code-based and with a graphical user interface (GUI). The GUI mode facilitates ReKinSim application for the majority of users with little knowledge of programming. More advanced users can benefit further from the code-based mode that allows user to couple or integrate it into other codes.

2.1. Computer algorithms

The ReKinSim program is written in MATLAB. The code-based mode needs the MATLAB environment and the Optimization Toolbox to run. The GUI mode is stand-alone and for running only requires the proper MATLAB libraries which are patched with the program and are alternatively available for download on MathWorks website. The internal numerical solver is a multi-variable-step, variable-order solver based on the numerical differentiation formulas of orders 1 to 5, and therefore can handle relatively stiff problems (Shampine and Reichelt, 1997). The nonlinear regression method for fitting a model to data is based on least-squares minimization algorithms and can be chosen between trust-region-reflective or Levenberg-Marquardt. The trust-region-reflective algorithm is based on the interior-reflective Newton method described in Coleman and Li (1996). For this algorithm, the nonlinear system of equations cannot be under-determined, that is, the number of equations must be at least as many as the number of fitting parameters. On the other hand, the Levenberg-Marquardt algorithm cannot handle bound constraints, as described in Levenberg (1944); Marquardt (1963). As a consequence, under-determined problems with parameters that are constrained to a certain range cannot be solved with ReKinSim.

2.2. Definition of mechanisms in ReKinSim

A MATLAB function file allows the direct input of the kinetic mechanisms into ReKinSim in the form of a system of ODEs. The file thus contains the differential equations that are needed for fitting. The ReKinSim method of input is structurally different and more straightforward than other previously mentioned programs. For example, KinTek Explorer (Johnson et al., 2009) uses a simple text file with a specific syntax to input the mechanisms, from which the program then automatically derives the differential equations. As a result, KinTek Explorer is limited to the elementary reactions only and cannot cover the more complex environmentally related dynamics. Input of kinetic mechanisms into AQUASIM (Reichert, 1994) is more versatile, though it requires the error-prone procedure of determining the stoichiometry matrix from the system of differential equations.

Apart from trivial reaction mechanisms such as reversible and irreversible first, second or higher order reactions, ReKinSim's versatility in solving different, arbitrary types of ordinary equations allows the inclusion of more complex dynamics into the kinetic models, such as Bioavailability limitations (Best, 1955), isotope fractionation and enrichment (Elsner, 2010), absorption/desorption (e.g., Hill-Langmuir equation, Hill, 1910; Langmuir, 1918), biomass growth (e.g, logistic function, McKendrick and Pai, 1912), equations of chemostats/retentostats, etc. Therefore, its applicability goes beyond the examples shown later in Section 3, and ReKinSim can in principle handle any mechanism of interest as long as its description can be written in the form of a first-order differential equation,