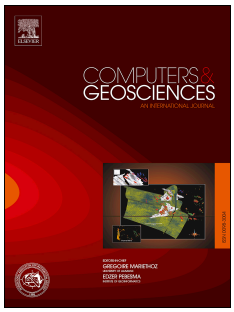
TReacLab: An object-oriented implementation of non-intrusive splitting methods to couple independent transport and geochemical software

Daniel Jara, Jean-Raynald de Dreuzy, Benoit Cochepin

Accepted Manuscript



TReacLab: An object-oriented implementation of non-intrusive splitting methods to couple independent transport and geochemical software

Daniel Jara, Jean-Raynald de Dreuzy, Benoit Cochepin

PII: S0098-3004(17)30251-0

DOI: 10.1016/j.cageo.2017.09.005

Reference: CAGEO 4019

To appear in: *Computers and Geosciences*

Received Date: 6 March 2017

Revised Date: 1 September 2017

Accepted Date: 7 September 2017

Please cite this article as: Jara, D., Dreuzy, J.-R.d., Cochepin, B., TReacLab: An object-oriented implementation of non-intrusive splitting methods to couple independent transport and geochemical software, *Computers and Geosciences* (2017), doi: 10.1016/j.cageo.2017.09.005.

TReacLab: an object-oriented implementation of non-intrusive splitting methods to couple independent transport and geochemical software

Daniel Jara[1], Jean-Raynald de Dreuzy[1], Benoit Cochepin[2]

[1]Géosciences Rennes, UMR CNRS 6118, Campus de Beaulieu, University of Rennes 1, Rennes, France

[2]ANDRA, 1/7 Rue Jean Monnet, 92298 Châtenay-Malabry, France

Abstract

Reactive transport modeling contributes to understand geophysical and geochemical processes in subsurface environments. Operator splitting methods have been proposed as non-intrusive coupling techniques that optimize the use of existing chemistry and transport codes. In this spirit, we propose a coupler relying on external geochemical and transport codes with appropriate operator segmentation that enables possible developments of additional splitting methods. We provide an object-oriented implementation in TReacLab developed in the MATLAB environment in a free open source frame with an accessible repository. TReacLab contains classical coupling methods, template interfaces and calling functions for two classical transport and reactive software (PHREEQC and COMSOL). It is tested on four classical benchmarks with homogeneous and heterogeneous reactions at equilibrium or kinetically-controlled. We show that full decoupling to the implementation level has a cost in terms of accuracy compared to more integrated and optimized codes. Use of non-intrusive implementations like TReacLab are still justified for coupling independent transport and chemical software at a minimal development effort but should be systematically and carefully assessed.

Keywords: Porous media; Reactive transport; Operator splitting; Object-oriented programming.

Corresponding author: daniel.jara.heredia@gmail.com

1. Introduction

The fate of chemical species in geological media results from the interaction of physical transport and chemical reactivity (Steefel *et al.*, 2005). Understanding how they interact requires field and laboratory studies as well as numerical models. Numerical models are important for building predictive scenarios where experiments are limited spatially and temporally, as in long-term nuclear waste disposal assessment (Marty *et al.*, 2014; Thouvenot *et al.*, 2013; Trotignon *et al.*, 2007). On the physical transport side, extensive work in applied mathematics and computational science has provided widely-used software for single and multi-phase flows as well as transport of chemical species such as MODFLOW (McDonald and Harbaugh, 1988), MT3DMS (Zheng and Wang, 1999), HYDRUS (Kool and Van Genuchten, 1991), COMSOL (COMSOL, 2010), FEFLOW (Diersch, 1996), MRST (Lie, 2014), and TOUGH2 (Pruess *et al.*, 1999). On the chemistry side, geochemical software have implemented a wide range of chemical functions and reactions, including equilibrium aqueous speciation, equilibrium mineral dissolution/precipitation, gas phase exchange, ion exchange, redox reactions, and kinetic reactions. Some of these software are PHREEQC (Parkhurst and Appelo, 1999), GEMS (Kulik *et al.*, 2013), CHEPROO (Bea *et al.*, 2009), MINTEQ (Peterson *et al.*, 1987), CHESS (Van der Lee, 2002), and Geochemist's Workbench (Bethke, 2007).

To combine physical and chemical reactivity, couplers have been developed between transport and geochemical codes such as PHAST for coupling HST3D and PHREEQC (Parkhurst *et al.*, 2004), HP1 for HYDRUS and PHREEQC (Šimůnek *et al.*, 2006), PHT3D

for MT3DMS and PHREEQC (Prommer *et al.*, 1999), HYTEC for RT1D/R2D2/METIS and CHESS (van der Lee *et al.*, 2003), OpenGeoSys-GEMS (Kulik *et al.*, 2013) and iCP for COMSOL and PHREEQC (Nardi *et al.*, 2014), UTCHEM-IPhreeqc and UTCHEM-EQBATCH (Kazemi Nia Korrani *et al.*, 2015, 2016), multicomponent transport software-IPhreeqc (Muniruzzaman and Rolle, 2016), FEFLOW-IPhreeqc (MIKE(DHI), 2016), Lattice Boltzmann transport software-IPhreeqc (Patel *et al.*, 2013). Most of the previously cited codes have embedded the coupling method with the geochemical and transport methods to enhance global performance and reliability. Here, in order to gain flexibility, we propose in our code TReacLab a complementary development in the form of an ensemble of Operator Splitting methods (OS) with a generic set of interfaces to transport and reaction operators. In this context, OS decouples chemistry from transport as opposed to global implicit solvers, which have been proven to be more accurate but less flexible (Hammond *et al.*, 2012; Hammond *et al.*, 2014; Mayer, 2000; Steefel, 2009; Zhang, 2012). TReacLab is designed as an open toolbox where additional OS techniques can be implemented and benchmarked. Other transport and geochemical codes may also be used at the minimal cost of developing the necessary interfaces. TReacLab is written in MATLAB based on a series of abstract classes using object-oriented programming (Commend and Zimmermann, 2001; Register, 2007; Rouson *et al.*, 2011).

After recalling in section 2 the reactive transport and OS formalism used, we present in section 3 our OS implementation. We especially show how to implement alternative OS methods and how to connect other transport and geochemical codes. Methods are assessed and discussed on the basis of 3 benchmarks in section 4.

2. Numerical model

2.1. Reactive transport equation

The reactive transport equation can be written in a general way as (Saaltink *et al.*, 1998):

$$\frac{\partial \theta c}{\partial t} = ML(c) + \theta S_e^t r_e + \theta S_k^t r_k + Q, \quad (1)$$

where c is the vector of concentrations for N_s chemical species in the system. θ is a diagonal matrix containing the porosity or volumetric content of the phase. M is a diagonal matrix that specifies whether a species is mobile or immobile. Its diagonal elements are 1 or 0 accordingly. S_k^t and S_e^t are the transposed stoichiometric matrix for kinetic and equilibrium reactions, respectively. r_e and r_k ($ML^{-3}T^{-1}$) are the reaction rates of the N_e equilibrium and N_k kinetic reactions, respectively. Q is the external sink/source term ($ML^{-3}T^{-1}$). L is the transport operator ($ML^{-3}T^{-1}$), which includes advection and diffusion. In the following, we consider only single-phase flow:

$$L(c) = \nabla \cdot [\mathbf{D}\nabla c - \theta v c]. \quad (2)$$

$\mathbf{D}$ (L^2T^{-1}) is the effective dispersion-diffusion tensor (Bear, 1972). The velocity v (LT^{-1}) is computed in a pre-processing phase, which can be decoupled from the reactive transport problem as long as hydraulic properties are not modified by the chemical reactivity. The chemical system can be generically written as the combination of the N_e equilibrium reactions:

$$\phi_e(c) = 0, \quad (3)$$

and of the N_k kinetically-controlled reactions:

$$r_k = \emptyset_k(c)\,. \tag{4}$$

The reactive transport problem is thus made up of the N_s mass balance equation (1) and of the $N_e + N_k$ equilibrium and kinetic equations (3) and (4). Its unknowns are the concentrations c and the reaction rates r_e and r_k. The chemical equilibrium system (3) is composed of the conservation equation and of the mass action law, relating reactants and products (Apoung-Kamga *et al.*, 2009; Molins *et al.*, 2004):

$$S_e log(c) = log(K), \tag{5}$$

where K is the vector of equilibrium constants.

Components u are generally introduced when considering equilibrium reactions (Saaltink *et al.*, 2011):

$$u = Uc, \tag{6}$$

where U is the component matrix (Fang *et al.*, 2003; Friedly and Rubin, 1992; Hoffmann *et al.*, 2012; Kräutle and Knabner, 2005; Steefel *et al.*, 2005). They are N_s - N_e linear combinations of chemical species that are not modified by equilibrium reactions (Molins *et al.*, 2004; Morel and Hering, 1993):

$$US_e^t r_e = 0. \tag{7}$$

The component matrix is not unique. However, its application to equation (1) always leads to a reduced system without the equilibrium rates but with the components u (Molins *et al.*, 2004; Saaltink *et al.*, 1998):

$$\frac{\partial u}{\partial t} = UML(c) + U\theta S\, r_k + UQ. \tag{8}$$

The reactive transport problem is then made up of the $2N_s - N_e + N_k$ equations (3-6) and (8) for the same number of unknowns u, c and r_k.

Under the assumption that solid species are not transported and all species have the same diffusion coefficient (i.e. $UML(c) = UL(u)$). Equation (8) classically gives the two following formulations TC and CC (Amir and Kern, 2010):

TC: $$\frac{\partial u}{\partial t} = L(u_a) + U\theta S\, r_k + UQ. \tag{9}$$

CC: $$\frac{\partial u_a}{\partial t} + \frac{\partial u_f}{\partial t} = L(u_a) + U\theta S\, r_k + UQ. \tag{10}$$

where $u_a = UMc$ and $u_f = U(I - M)c$ are the aqueous and fixed components. In the TC formulation, the fixed species concentration are deducted from the solution in the total component concentration (T) and the solute concentration (C). In the CC formulation, the total component concentration is divided in aqueous and fixed components.

2.2. Usual first-order sequential non-iterative and iterative approaches

In this section, we show how the reactive transport problem can be solved using independent transport and chemical solvers. We distinguish the sequential non-iterative and iterative approaches respectively based on TC and CC formulations. For the sequential non-iterative approach, we extract from the TC formulation, the transport operator in which we keep the sink/source term:-

$$\frac{\partial u}{\partial t} = L(u_a) + UQ. \tag{11}$$

The chemical operator derives from equations (3-6), and (8). Note that it does not contain any source/sink term, as it has been included in the transport equation:

$$\begin{gathered} \frac{\partial u}{\partial t} = US_k^t r_k \\ r_k = \emptyset_k(c) \\ u = Uc \\ \emptyset_e(c) = 0. \end{gathered} \tag{12}$$

This is still a system of $2N_s$ - N_e + N_k equations for the same number of unknowns. This decoupled system can be solved with the classical sequential non-iterative approach using an explicit integration of temporal derivatives (herein, we assume forward Euler). The solution at time step *n+1* can be obtained from the solution at time step *n*, with the following successive application of the transport and chemical operators in a sequential approach:

$$u^* = u_n + \Delta t(L(u_a) + UQ)$$

$$\begin{cases} u^* = Uc_{n+1} \\ \emptyset_s(c_{n+1}) = 0 \end{cases} \quad (13)$$

$$r_{k_{n+1}} = \emptyset_k(c_{n+1})$$

$$u_{n+1} = u^* + \Delta t U S_k^t r_{k_{n+1}}$$

The transport operator (11) is applied to the components. Then the chemical operator is applied with the updated mobile components for speciation between fixed and solute concentrations. In the specific case where chemical reactions are all at equilibrium and no kinetics is involved, a TC formulation is used to fully decouple (de Dieuleveult *et al.*, 2009). In such case the decoupling does not then rely on operator splitting, but on a block Gauss-Seidel method.When the stability conditions of the explicit integration are too much constraining, implicit schemes should be used instead within a sequential iterative approach (Carrayrou *et al.*, 2004; de Dieuleveult and Erhel, 2010; Yeh and Tripathi, 1989):

$$u_{n+1} = u_n + \Delta t\left[L\left(u_{a_{n+1}}\right) + \theta S_k^t r_{k_{n+1}} + UQ\right]$$

$$\begin{cases} u_{n+1} = Uc_{n+1} \\ \emptyset_s(c_{n+1}) = 0 \end{cases} \quad (14)$$

$$r_{k_{n+1}} = \emptyset_k(c_{n+1}).$$

Classical Picard's method have been extensively used to solve such kind of problems:

$$u_{n+1}^{k+1} = u_n + \Delta t\left[L\left(u_{a_{n+1}}^{k+1}\right) + \theta S_k^t r_{k_{n+1}}^{\ k} + UQ\right]$$

$$\begin{cases} u_{n+1}^{k+1} = Uc_{n+1}^{k+1} \\ \emptyset_e\left(c_{n+1}^{k+1}\right) = 0 \end{cases} \tag{15}$$

$$r_{k_{n+1}}^{\ k+1} = \emptyset_k\left(c_{n+1}^{k+1}\right),$$

where *k* is the index of the Picard iteration method instantiated by:

$$u_{n+1}^{k=1} = u_n$$

$$r_{k_{n+1}}^{\ k=1} = r_{k_n}. \tag{16}$$

We recall the necessity to check the consistency of the temporal integration scheme with the Operator Splitting method chosen. With this decomposition, explicit first-order scheme naturally leads to sequential non-iterative approach. The implicit first-order scheme requires a sequential iterative approach. Other choices are possible and might reduce errors depending on the chemical system (Barry *et al.*, 1996). As it should be possible to test and benchmark them at a reduced development cost, we use a generic decoupling formalism that can be used to implement a broad range of schemes.

2.3. Generic operator splitting implementation

The reactive transport system can be generically split in two operators. Using the formalism of Gasda *et al.* (2011), equation (1) can be written as:

$$\frac{\partial Z}{\partial t} = \mathcal{L}_1 Z + \mathcal{L}_2 Z, \qquad Z(t=0) = Z_0, \qquad 0 \leq t \leq T, \tag{17}$$

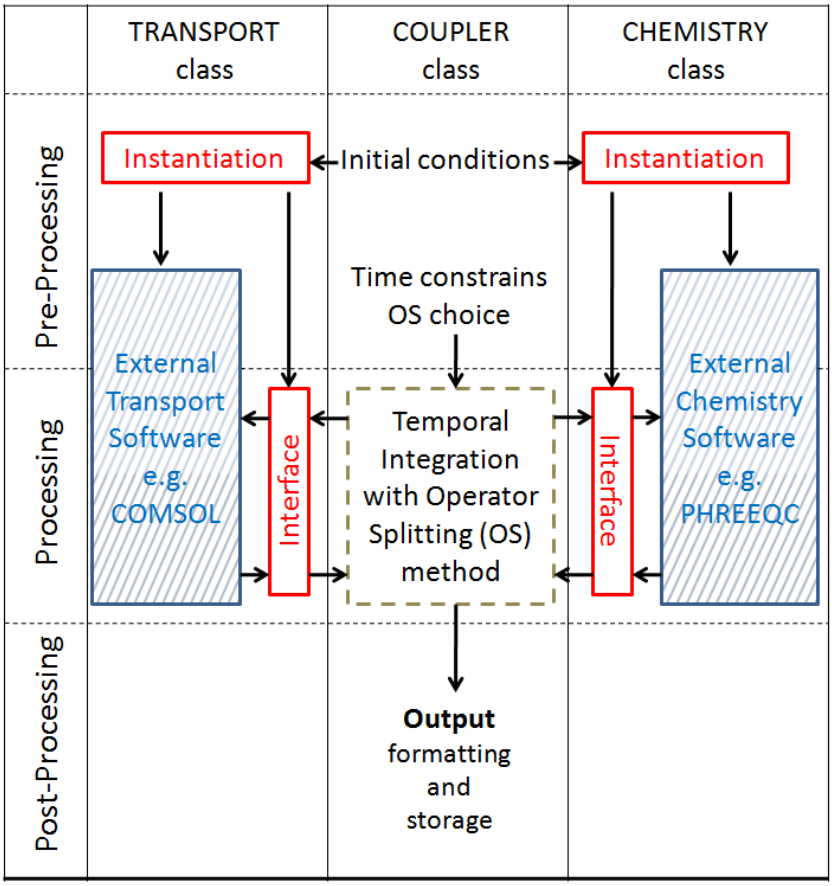
where Z is the unknown, $\mathcal{L}_1$ and $\mathcal{L}_2$ can be equation (11) and (12), respectively. Other decomposition are possible, e.g. the transport operator can be subdivided into an advection and a diffusion-dispersion operator (Clement *et al.*, 1998), or one operator might contain advection-reaction and the other diffusion (Liu and Ewing, 2005). Each operator will be solved separately for a splitting time step $\Delta t = t^{n+1} - t^n$ using adapted numerical methods.

The generic operator splitting methods implemented into the Toolbox are the sequential splitting, additive splitting, Strang splitting, symetrically weighted splitting, and alternating method (Appendix A). Assuming exact integration of the operators and homogeneous boundary conditions in equation (18), the first two have a first-order temporal truncation error, and the following three a second-order one (Hundsdorfer and Verwer, 2013). Since the operators are usually solved using numerical methods, the global order of such approaches might be modified because of the order of the numerical methods used for each operator (Barry *et al.*, 1996; Csomós and Faragó, 2008). The alternating splitting increases the order of the sequential splitting if the time steps are small enough (Simpson and Landman, 2008; Valocchi and Malmstead, 1992).

3. Operator splitting implementation and software organization

We provide in TReacLab an object-oriented toolbox for the non-intrusive operator splitting methods of the previous section. TReacLab is organized along three main components for coupling transport and reactivity, and proceeds in three pre-processing, processing and post-processing phases (Figure 1). These three components correspond to the three

well-identified coupler, transport and chemistry classes. The three classes are fully segmented and exchange information through interfaces. Segmentation ensures that any of the three coupler, transport and chemistry classes can be replaced without modifications of any of the two other ones. The solution of the reactive transport problem after spatial discretization eventually consists in the temporal integration with the chosen OS technique, which iteratively calls transport and geochemical solvers through interfaces (Figure 1, middle row). This is the core of the simulation that we identify as the processing phase. It is generic and does not require at run time any further specification of transport, reactivity and coupler methods. Standard error management techniques are used to stop the algorithm when any of the integration method of the three classes fails, stopping the running process and returning adapted error messages.



174

175 *Figure 1: General software organization of TReacLab with the three coupler, transport and*
176 *chemistry classes in columns, and the three pre-processing, processing, and post-processing*
177 *phases in rows. Generic components represented in black are the organization and the*
178 *coupler class. External software for transport and chemistry are represented in blue with*
179 *hatched line (cannot be modified). Red boxes highlight the instantiation and interface*
180 *methods that must be developed when connecting new transport or chemistry software.*

181 The processing phase can be generic because all specifications of the coupler, transport, and
182 chemistry classes are performed in a pre-processing phase (Figure 1, first row). The
183 pre-processing phase consists in the instantiation of the coupler, transport and chemical
184 classes, in the preparation of the interfaces that will transfer information and in the
185 specifications of the initial conditions. As detailed in Appendix B, instantiations are code
186 dependent. Instantiation can be done externally for example with the definition of a transport

or chemical problem through the graphical user interface of software like COMSOL or PHREEQC. It can also be done internally by a method within TReacLab specifying the inputs and parameters to existing interfaces like IPhreeqc (Charlton and Parkhurst, 2011), PhreeqcRM (Parkhurst and Wissmeier, 2015), or COMSOL livelink (COMSOL, 2010). Even when instantiation is complex, it remains independent for each of the three classes. Cross-dependencies and feedback between transport and reactivity like density-driven flows with reacting species are not supported at this stage, although they may be important in some applications like CO_2 sequestration (Abarca *et al.*, 2013).

Pre-processing phase specifies the initial conditions and transfers them to the coupler in charge of starting the numerical integration. Post-processing is generic and only consists in formatting and storing output concentrations and solver performances (Figure 1, bottom row). Specifications are all restricted to the instantiation of the software and interface in the pre-processing phase while processing and post-processing remain fully generic. Connections between specific algorithms and generic structures are done by interfaces. Appendix B provides a detailed description of the transport and chemistry classes, defining the interfaces to the external codes.

4. Examples and benchmarks

The three following examples validate the methods and illustrate the implementation presented in sections 2 and 3. The three of them are based on a 1D hydraulically homogeneous system with steady-state flow and uniform dispersion (equation (2)).The examples are compared visually against analytical solution or well-know numerical software. Moreover, we show a convergence study for the first case being the reference solution the numerical solution with finest time resolution.

The four examples display evolving degrees of complexity both in terms of chemical systems and in terms of software called for transport and reactivity, software versions are given in Table 1. The first example is a single-species transport with first-order decay. The transport solver is COMSOL and the chemical solver is a simple analytical solution. This example is used to assess the different coupling algorithms implemented and to check the implementation of the interface with COMSOL. The second example is an equilibrium precipitation/dissolution chemical system in a 1D hydraulically homogeneous system. Chemical solver is IPhreeqc. Several solvers have been compared for the transport solver, both to check IPhreeqc interface implementation and to evaluate the effect of the transport solver. The third example is the most advanced in terms of chemistry. Chemical reactions are partly in equilibrium and partly kinetically controlled. They involve precipitation and dissolution reactions. The chemical code is PhreeqcRM. It is used in combination with COMSOL as transport solver. The last problem face a 2D unsaturated system where transport is modeled by Richards equation and solved by COMSOL. Chemistry is solved by PhreeqcRM. These four test cases have been chosen to check the implementation and assess the coupling methods developed. They are also simple enough from the development point of view to be taken as starting points to model more advanced chemical systems and transport conditions.

Software	Version
MATLAB	R2013b
COMSOL	4.3b
PHREEQC	3.3.7
IPhreeqc	3.3.7
PhreeqcRM	3.3.9

Table 1: Software versions.

4.1. Single-species transport with first-order decay

A single-species transport with first-order decay using different OS methods is compared to an analytical solution (Van Genuchten and Alves, 1982). The reactive transport system contains a single solute species of concentration c :

$$\frac{\partial c}{\partial t} = L(c) - kc, \tag{18}$$

where L is given by equation (2). Equation (18) can straightforwardly be separated into transport and chemistry operators corresponding to the two right-hand side terms.

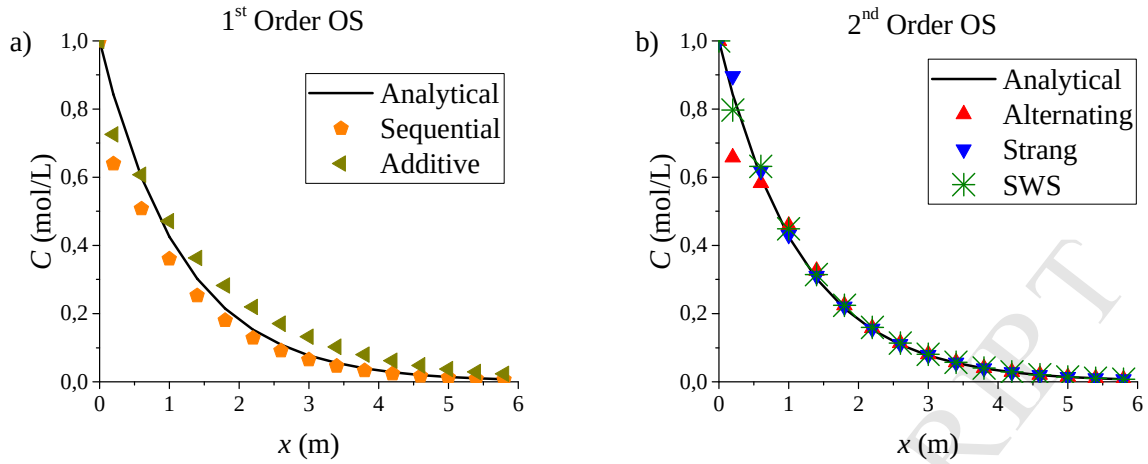
At time 0, the solute concentration is 0 in the domain ($c(x, t=0) = 0$). The concentration at the left boundary is constant and equal to 1 mol/m^3 ($c(x = 0, t) = 1$ mol/m^3). The boundary condition on the right side of the domain is a perfectly absorbing condition ($c(x = x_{max}, t) = 0$).

Parameters are derived from Steefel and MacQuarrie (1996) and given in Table 2. The solver for transport is COMSOL and an analytical solution is used for the first-order decay. Solute concentration progressively invades the domain from the left boundary with a smooth profile resulting from the combination of dispersion and decay (Figure 2). Second-order methods perform much better than first-order methods as expected. Errors are more pronounced at the inlet boundary condition on the left side of the domain where the concentration is higher

(Steefel and MacQuarrie, 1996; Valocchi and Malmstead, 1992). The sequential splitting method with the transport operator performed first overestimates the amount of reaction for the whole domain since it considers that all incoming solute is getting in without decay for the full first time step. If the sequence of operators is exchanged, namely first chemistry is solved, and then transport is solved, the amount of reaction is underestimated. The second-order alternating splitting, which alternates between transport-chemistry and chemistry-transport steps, shows strong improvement with compensations between overestimation in the first application of the chemical operator and underestimation in the second application of the chemical operator (Simpson and Landman, 2008; Valocchi and Malmstead, 1992).

Parameter	**Value**
v [m/y]	100
D [m^2/y]	20
k [y^{-1}]	100
x_{max} [m]	6
Δx [m]	0.4
Δt [y]	4 10^{-3}

Table 2: Parameters for the single-species transport with first-order decay benchmark. v is the velocity, D is the dispersion coefficient, k is the decay rate, x_{max} is the length of the 1D column, Δx is the grid size, and Δt is the time step.



256 *Figure 2: Comparison of first- and second-order OS for the single-species transport with*
257 *first-order decay at t = 0.5 y. Parameters are given in Table 2. Analytical solution is derived*
258 *from Van Genuchten and Alves (1982).*

259 The error at time t = 0.5 y is taken as the quadratic relative difference over the domain of the
260 finest time step of the numerical solution and the numerical solutions for the corresponding
261 time step, c_{NF} and c_N respectively:

$$\|e\|_2 = \sqrt{\sum_{i=1}^{N_x} \left(\frac{c_{NF}^i(t) - c_N^i(t)}{c_{NF}^i(t)} \right)^2}. \qquad (19)$$

262 Table 3 displays the values for evolving time steps and shows that all methods converge with
263 the time. The reference finest time step for each method has been Δt = 2 10^{-4} s (i.e. $c_{NF}^i(t)$
264 value). While all methods perform well, the sequential method is more accurate than the
265 additive one and second-order methods are overall more accurate than first-order methods.
266 The performance on convergence arranged on descending order is given by Strang,
267 symmetrically weighted splitting, alternating, sequential and additive.

		1st order		2nd order		
Δx = 0.4 m	Δt (y)	**Additive**	**Sequential**	**Alternating OS**	**Strang**	**SWS**
	$4\ 10^{-3}$	1,107	0,1667	0,075	0,032	0,049
	$2\ 10^{-3}$	0,514	0,079	0,032	0,026	0,028
	$4\ 10^{-4}$	0,114	0,019	0,029	0,031	0,029

Table 3: Error $\|e\|_2$ of equation (19) for the single-species transport with first-order decay with different OS methods and splitting time steps.

4.2. Calcite dissolution

Calcite dissolution and dolomite formation has become a classical benchmark for reactive transport problems with sharp precipitation/dissolution fronts (Beyer *et al.*, 2012; Engesgaard and Kipp, 1992; Prommer *et al.*, 1999). Progressive introduction of magnesium calcium in a domain at equilibrium between calcium carbonate in solution and calcite ($CaCO_3$) dissolves the calcite and precipitates dolomite ($CaMg(CO_3)_2$). This chemical system has been modeled with the physical and chemical parameters given by Table 4, Table 5, and Table 6. Chemical concentrations are initially homogeneous. At the initial time (t = 0), the chemical system is destabilized with the introduction of magnesium instead of calcium at the upper boundary condition (x = 0), inducing the dissolution/precipitation process. The boundary condition at the downstream limit (x_{max}) is a simple outflow of the solutes.

Here, we show how transport solvers can be applied and validate our interface to IPhreeqc. IPhreeqc performs the computation of components, aqueous speciation, precipitation and dissolution reactions (Charlton and Parkhurst, 2011). The database used is 'NAPSI_290502(260802).dat'. Transport is solved either with COMSOL Multiphysics (COMSOL, 2012), with a finite difference spatial discretization and forward Euler time

integration, derived from built-in pdepe function of MATLAB (Skeel and Berzins, 1990). Transport and chemistry are coupled through the simple sequential approach of equations (A.1)-(A.3). PHREEQC is independently run as 1D reactive transport solver for general comparison.

Parameter	Value
v [m/s]	10^{-5}
D [m^2/s]	6.7 10^{-8}
x_{max} [m]	0.25
Δx [m]	0.01
Δt [s]	50

Table 4: Physical parameters for the calcite dissolution benchmark. v is the average velocity, D is the dispersion coefficient, x_{max} is the maximum length of the column, Δx is the grid size, and Δt is the time step.

Chemical Component and Species	Initial value	Boundary value at x=0
Ca [mol/L]	1.23 10^{-4}	0
C [mol/L]	1.23 10^{-4}	0
Cl [mol/L]	0	2 10^{-3}
Mg [mol/L]	0	10^{-3}
pH [-]	9.91	7
Calcite [mol/L]	2 10^{-4}	-
Dolomite [mol/L]	0	-

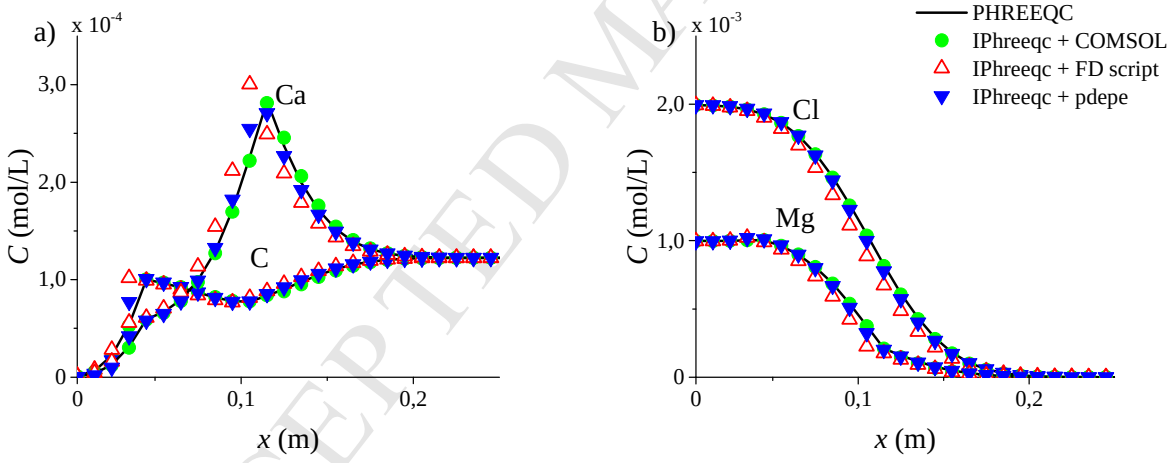
Table 5: Calcite dissolution benchmark initial and boundary values for aqueous components and mineral species. In PHREEQC, components are called elements.

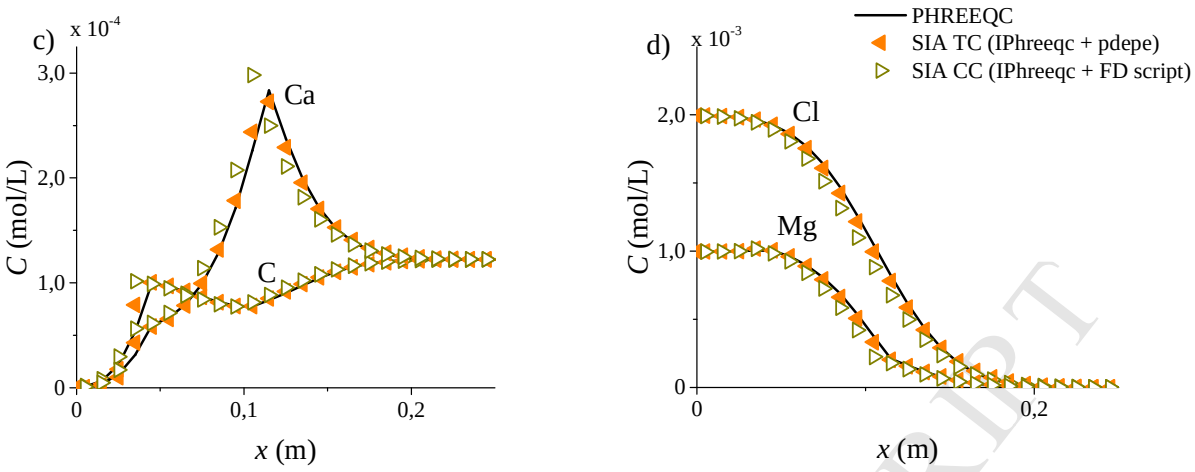
295

Homogeneous reactions	**log (K)**
$2H^+ + 2e^- \leftrightarrow H_2$	-3.1055
$2H_2O - 4H^+ - 4e^- \leftrightarrow O_2$	-85.9862
$HCO_3^- + 9H^+ + 8e^- - 3H_2O \leftrightarrow CH_4$	27.8493
$H_2O - H^+ \leftrightarrow OH^-$	-13.9995
$H^+ - H_2O + HCO_3^- \leftrightarrow CO_2$	6.3519
$HCO_3^- - H^+ \leftrightarrow CO_3^{2-}$	-10.3289
$Ca^{2+} - H^+ + HCO_3^- \leftrightarrow CaCO_3$	-7.1048
$Ca^{2+} + HCO_3^- \leftrightarrow CaHCO_3^+$	1.1057
$Ca^{2+} + H_2O - H^+ \leftrightarrow CaOH^+$	-12.78
$Mg^{2+} - H^+ + HCO_3^- \leftrightarrow MgCO_3$	-7.3492
$Mg^{2+} + HCO_3^- \leftrightarrow MgHCO_3^+$	1.0682
$Mg^{2+} + H_2O - H^+ \leftrightarrow MgOH^+$	-11.44
Homogeneous reactions	
Calcite $CaCO_3 \leftrightarrow Ca^{2+} - H^+ + HCO_3^-$	1.849
Dolomite $CaMg(CO_3)_2 \leftrightarrow Ca^{2+} + Mg^{2+} - 2H^+ + 2HCO_3^-$	4.118

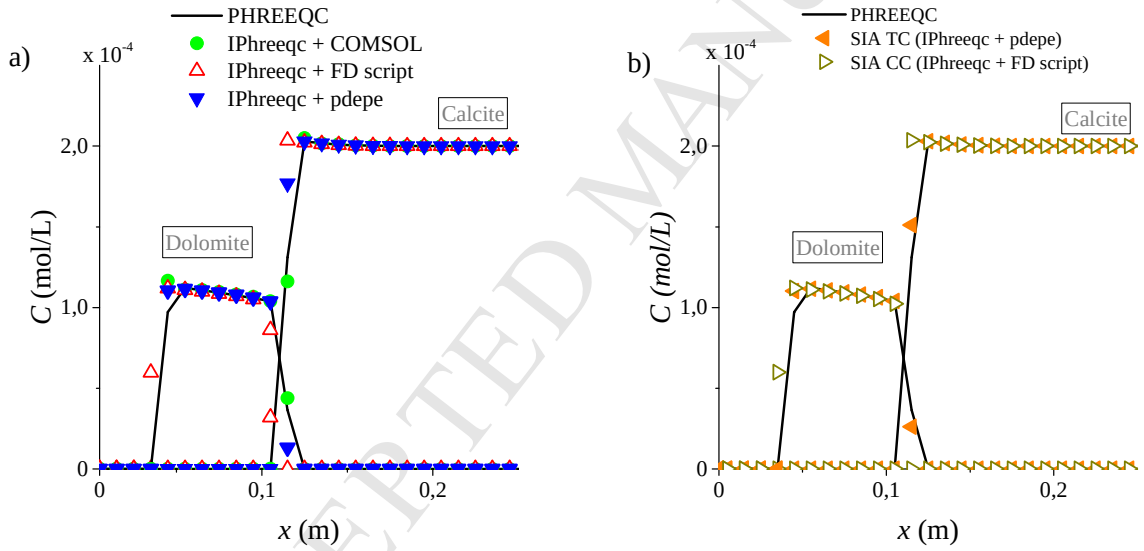
Table 6: Chemical system of the calcite dissolution benchmark. The upper part comprises the homogeneous equations and the lower part the heterogeneous reactions. The first column shows the equilibrium reactions and the second one the logarithms of equilibrium constants.

Figure 3 and Figure 4 display aqueous and mineral equivalent concentrations at time $t = 10^4$ s. As magnesium and chloride get in the domain (Figure 3b and Figure 3d), calcite progressively dissolves and is replaced by dolomite as expected (Figure 4). Some of the calcium remains in solution and is flushed out (Figure 3a and Figure 3c). Because of the subsequent absence of calcium in solution, dolomite dissolves again with some increase of calcium in solution (Figure 3a and Figure 3c). The three different transport solvers give the same tendency as the PHREEQC solution.





306 *Figure 3: Aqueous concentration profiles at time t = 10^4s.*



307 *Figure 4: Dolomite and calcite equivalent concentration profiles with open and filled symbols*

308 *respectively at time t = 10^4 s.*

309 Although COMSOL leads to good results, it is more than one order of magnitude slower than

310 the two other transport methods (Table 7). We checked that this large difference in

311 performances does not come from the numerical method but from the large time required for

312 COMSOL to start and stop when called numerous times externally. While this might not be

an issue for large transport problems for which limitations will rather come from transport operator, it is a constrain for smaller tests and benchmarks.

Software Coupling	Time
IPhreeqc + COMSOL	668 s
IPhreeqc + FD script	24 s
IPhreeqc + pdepe	40 s

Table 7: Time performance for the calcite dissolution benchmark using a sequential operator splitting.

Whatever the coupling method, the consistency with PHREEQC is overall good. Although COMSOL uses, as default, implicit time integration schemes for solving the transport equation instead of the required explicit method, it still compares well with PHREEQC and the other software couplings. Indeed, the sequential non iterative method requires an explicit time integration for transport (equation (13)). It is not the case for COMSOL which uses (as default) a backward differentiation formula temporal integration scheme, which order varies with the internal time step adaptation (COMSOL, 2012). It thus introduces an additional error in the coupling scheme (de Dieuleveult *et al.*, 2009). However, by using such stable and accurate temporal integrations, it enhances the robustness of the transport scheme.

4.3. Mixed equilibrium-kinetic system

We simulate the progressive increase of dissolved species in an atmospheric water infiltrating a granitic bedrock. This test case is derived from Nardi *et al.* (2014). The hydraulic properties of the system are found in Table 8. The infiltrating water has much lower concentrations of dissolved species than the resident water. It interacts with five minerals (Table 9). It is in equilibrium with calcite. The four other minerals k-feldspar, illite, albite and pyrite are subject to kinetically controlled dissolution with rates ranging from 10^{-13} to 10^{-11} mol/s. All

parameters and rate laws of the simulation are provided in the PHREEQC file of iCP (Nardi *et al.*, 2014). The infiltrating water dissolves calcite to maintain equilibrium, increasing both the concentration of calcium and the pH of the solution. Other minerals also dissolve and increase the concentrations of Al and K in solution, however at a much slower rate because of the kinetic control of the reactions. pH is eventually buffered by the dissolution of illite and pyrite.

Parameter	**Value**
v [m/s]	$2.78\ 10^{-6}$ m/s
D [m/s]	$5.55\ 10^{-9}$ m^2/s
x_{max} [m]	0.08 m
Δx [m]	10^{-3} m
Δt [s]	720 s

Table 8: Parameters for mixed equilibrium-kinetic benchmark. v is the average velocity, D is the dispersion coefficient, x_{max} is the maximum length of the column, Δx is the grid size, and Δt is the splitting time.

342

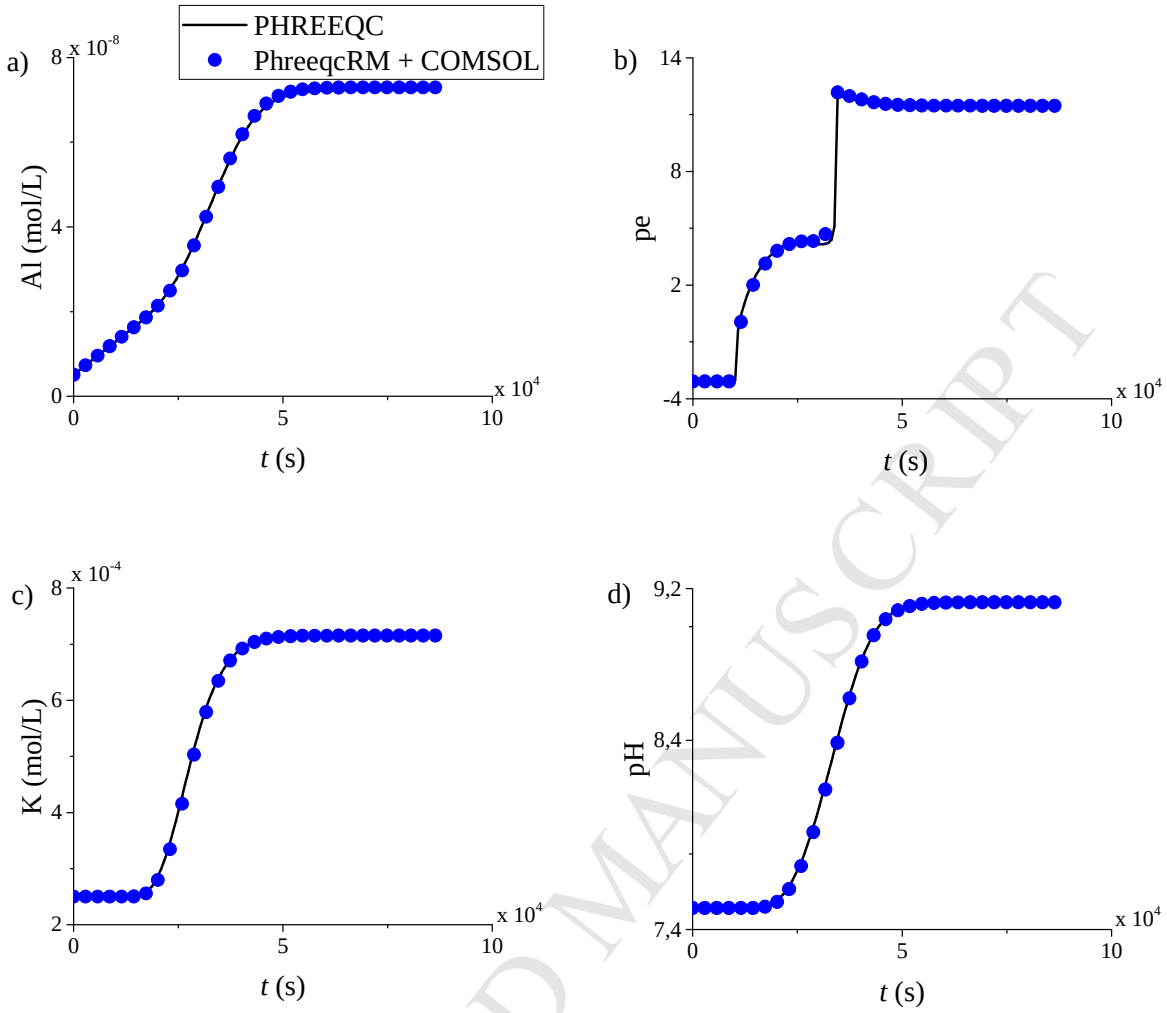
Chemical Component and Species	Initial value	Boundary value
Ca [mol/L]	$1.4\ 10^{-3}$	$3\ 10^{-4}$
C [mol/L]	$4.9\ 10^{-3}$	$1.9\ 10^{-4}$
Cl [mol/L]	$1.1\ 10^{-2}$	$9\ 10^{-4}$
Mg [mol/L]	$7.4\ 10^{-4}$	$2\ 10^{-4}$
Mn [mol/L]	$3.4\ 10^{-6}$	0
S [mol/L]	$9.6\ 10^{-4}$	$4.8\ 10^{-4}$
Na [mol/L]	$1.3\ 10^{-2}$	$3\ 10^{-4}$
K [mol/L]	$2.5\ 10^{-4}$	$7.1\ 10^{-4}$
Fe [mol/L]	$7.2\ 10^{-6}$	$5.4\ 10^{-5}$
Sr [mol/L]	0	$6.8\ 10^{-7}$
Si [mol/L]	$2\ 10^{-4}$	$2.5\ 10^{-6}$
Al [mol/L]	$5.1\ 10^{-9}$	10^{-8}
P [mol/L]	$3.8\ 10^{-6}$	0
Br [mol/L]	$1.7\ 10^{-5}$	0
F [mol/L]	$3.1\ 10^{-5}$	$1.6\ 10^{-5}$
pH [-]	7.5144	7.3
pe [-]	-3.0836	13.6
Calcite [mol/L]	6.065	-
K-feldspar [mol/L]	0.239	-
Illite [mol/L]	0.144	-
Albite [mol/L]	0.289	-
Pyrite [mol/L]	1.17	-

343 *Table 9: Aqueous components and mineral species for mixed equilibrium-kinetic benchmark.*

344 To simulate this set of reactions, we have chosen PhreeqcRM to assess the flexibility of
345 TReaCLab. Transport is simulated with COMSOL to benefit from the accurate transport
346 solver, it uses a variable order (between 1 and 5) backward differentiation formula. In the

presence of both kinetically controlled and equilibrium reactions, both the quality of the transport and reactive integrations and coupling issues may be critical. We choose a simple sequential OS method with the successive integration of transport and reactivity. The results obtained by the coupling of COMSOL and PhreeqcRM are close to the solution given by PHREEQC alone for the dissolved species and kinetically dissolving minerals (Figure 5). The time step of the coupled PhreeqcRM and COMSOL integration has been taken smaller than the characteristic mesh scale transport time and reactive time at least for the kinetical reaction to ensure accurate integrations. The most difficult quantity to get accurately is the calcium concentration because calcite is at equilibrium. The time step must be reduced to recover a steeper reactive front (Figure 6).

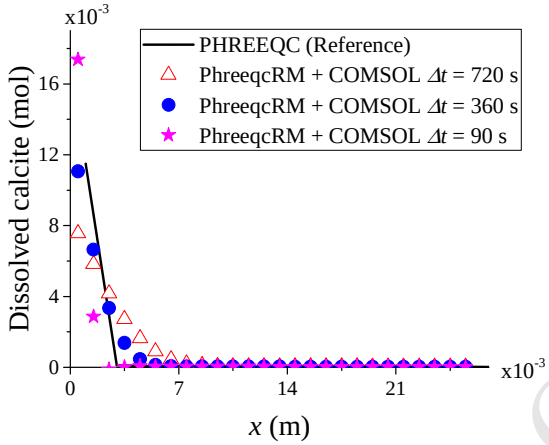
This more advanced test shows that the computational load should be well balanced between the coupler, transport and chemistry methods. While coupling is the critical component in cases of equilibrium reactions and may even require highly integrated coupling strategies like global implicit methods (Hoffmann *et al.*, 2010; Saaltink *et al.*, 2001), it is not the case for kinetically controlled reactions. In this case of mixed equilibrium kinetic reaction, elementary coupling and accurate transport and reactive solvers can be efficient with small enough time steps where sharp reaction fronts are involved.



364 *Figure 5: Comparison of results between the coupling of PhreeqcRM and COMSOL and*

365 *PHREEQC observed for the mixed equilibrium-kinetic benchmark at the output of the column.*

366

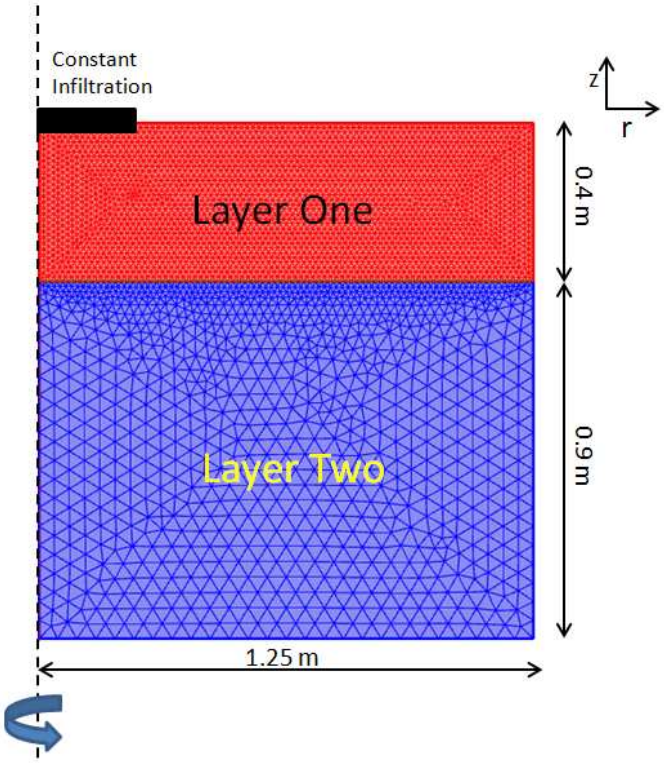


367

368 *Figure 6: Quantity of dissolved calcite with PhreeqcRM and COMSOL for two different*
369 *splitting time steps Δt = 720 s, 360 s and 90 s. PHREEQC independently is used as reference.*

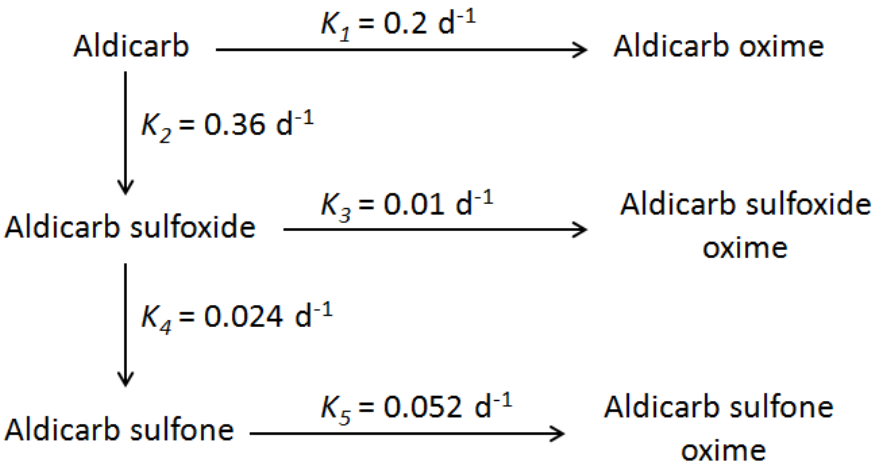
370 **4.4. Pesticide infiltration**

371 The following benchmark concerns the infiltration in an unsaturated soil column of a
372 carbamate insecticide (Aldicarb) (MIKE(DHI), 2016; Multiphysics, 2008; Šimůnek *et al.*,
373 1994; Wissmeier and Barry, 2011). The soil column is a 2D axisymmetric cylinder made up
374 of two layers with a smaller hydraulic conductivity in the upper layer but higher saturation.
375 Transport is modeled by Richards' equation and solved by COMSOL (Figure 7). Aldicarb is
376 transported downwards and sideways from the infiltration (top of the column from r = 0 m to
377 r = 0.25 m). Chemistry is described by first-order decay chain reactions (Figure 8), being only
378 mobile Aldicarb, Aldicarb sulfoxide and Aldicarb sulfone (i.e. the other species are fix
379 species). These system of ordinary differential equations is solved by PhreeqcRM.



380

381 *Figure 7: Soil column geometry and mesh.*



382

383 *Figure 8:Aldicarb reaction chain.*

384 The simulation time is 8 days with a splitting time step of 0.05 days. The number of nodes is
385 3936 nodes. Figure 9a and Figure 9b show the concentration in the soil column of Aldicarb
386 and Aldicarb sulfone, respectively. Aldicarb disappears fast from the domain since its kinetic
387 constant are fast in comparison to the kinetic constants of the daughter species. Therefore,
388 Aldicarb (and also Aldicarb oxime) are presented close to the infiltration condition. On the
389 contrary, the other daughter species (Aldicarb sulfoxide, aldicarb sulfone, aldicarb sulfoxide
390 oxime, aldicarb sulfone oxime) have a similar distribution in the domain. Figure 9c and 9d
391 show the concentration of Aldicarb and Aldicarb sulfone when r = 0 m for the different OS
392 methods and COMSOL alone. It is possible to see a good agrement between all the methods,
393 although a discrepancy between the methods and COMSOL is observable. The discrepancy is
394 related to the OS error and the chosen integration time scheme for the chemistry step.

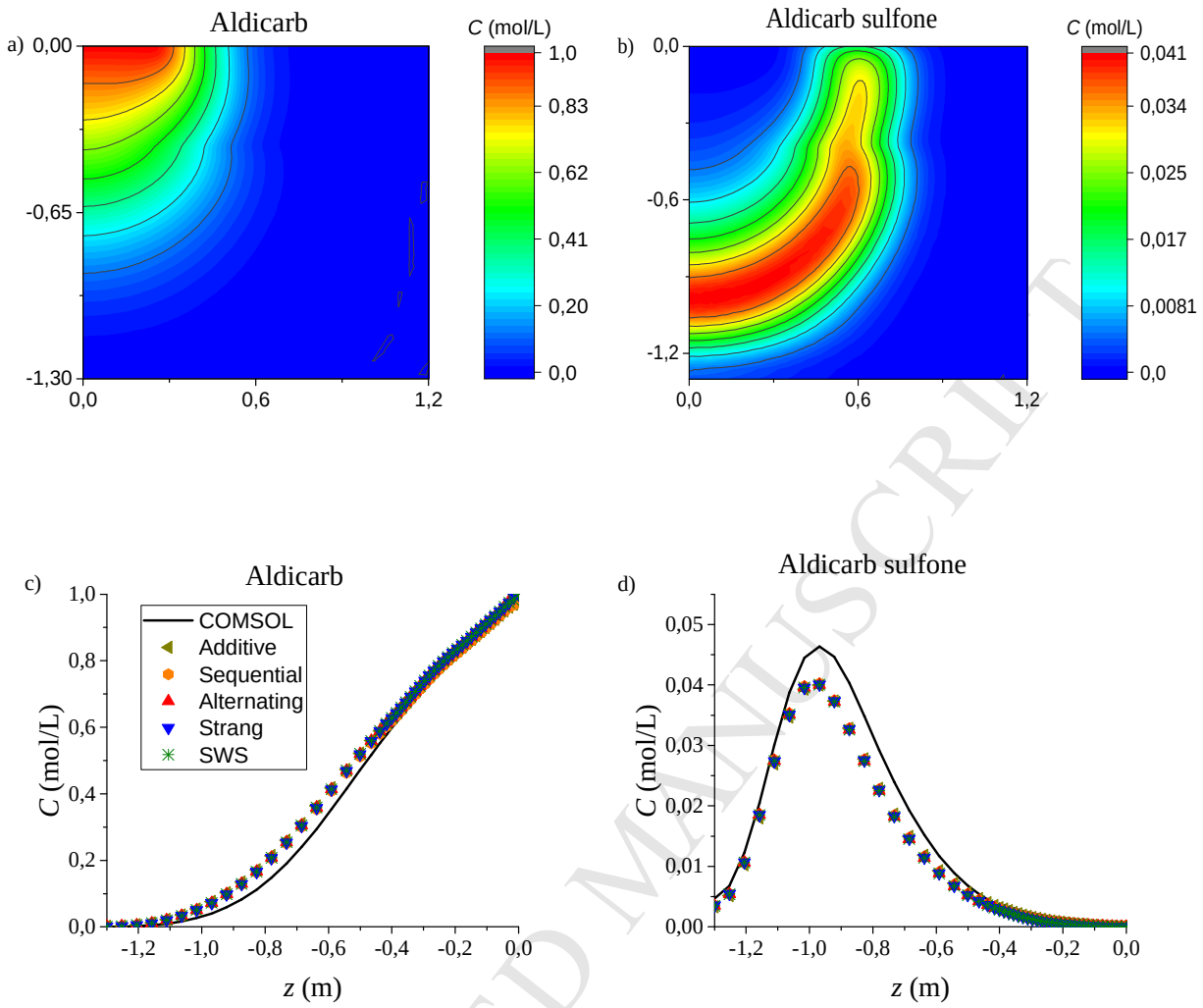


Figure 9: a) Aldicarb contour plot after 8 days, b) Aldicarb oxime contour plot after 8 days, c) Concentration aldicarb at r = 0 m for all the methods and Comsol, d) Concentration aldicarb oxime at r = 0 m for all the methods and Comsol.

5. Discussion

As shown by many previous studies and by the four examples of the previous section, reactive transport problems can be solved by a wide diversity of transport, chemistry, and operator splitting methods. No method is currently accepted as systematically more accurate and efficient than any other. Integration of the transport and chemistry operators in PHREEQC using more appropriate splitting with advection-reaction on one side and diffusion-reaction on

the other side leads to better resolution of chemical fronts as shown in the second and third cases (Parkhurst and Appelo, 1999), but such front can be obtained by smaller time steps. It is not only the integration but also the successive improvements of the methods that lead to significantly more accurate schemes. TReacLab results remain however close, displaying the same overall behavior both on solute and mineral concentrations. The interest of fully segmented reactive transport implementations like in TReacLab is not motivated by the accuracy and should not be used when other more integrated and optimized software are appropriate and freely available.

Despite their lower accuracy, fully segmented implementations may be useful in situations where flexibility is essential. It is the case when extensive modeling work has been performed in independent software environments for transport or chemistry, and extensions to reactive transport problems are required. Transport and chemistry solvers are then imposed and should be coupled with as few specific developments as possible. For example, COMSOL and PHREEQC have been interfaced here and in several other works because of their complementarity (Nardi *et al.*, 2014; Nasir *et al.*, 2014; Wissmeier and Barry, 2011). It is possible to specify advanced geometrical configurations in COMSOL through a convenient graphical user interface (Azad et al., 2016). PHREEQC provides advanced capacities for modeling complex geochemical systems with extensive database of reactions (Charlton and Parkhurst, 2011; Parkhurst and Wissmeier, 2015). In such cases, building the structure of the model may be the first and dominant issue in developing simulation capacities. That is when codes like TReacLab can provide practical bridges for reactive transport systems. The examples of section 4 however shows that they must be used with great care. Especially, the called software may have different temporal integration schemes than the explicit and implicit methods required by the SNIA and SIA coupling methods as discussed for the higher-order schemes of COMSOL in section 4. Using codes like COMSOL may enhance robustness at a

certain cost of accuracy. Thus, implementation capacity does not guarantee validity. Validity must be carefully checked and argued with other comparable cases or with appropriate convergence analysis.

Another targeted use of TReacLab concerns the development and test of new coupling methods or strategies. Operator splitting can be performed with various methods including for example adaptative time stepping (Belfort *et al.*, 2007; Gasda *et al.*, 2011). Global implicit approaches that separate geochemical and transport software might also be more widely tested providing the Jacobian of the chemical operator and taking into account current limitations such as the difficulties to model precipitation/dissolution reactions (Amir and Kern, 2010). TReacLab may then be used as a platform where interfaces to chemical and transport operators are available and have been tested and documented for other coupling methods.

These applications are possible because TReacLab is a fully free and open software that can be directly accessed and downloaded (https://github.com/TReacLab/TReacLab). The free and open use of TReacLab has been dominant in its development and in the choices made for its organization. The repository thus provides two main directories with sources and examples respectively. Sources are organized in four main categories for chemistry, transport, coupler and utilitaries. At the root of the chemistry, transport and coupler directories are the virtual classes as main entries. Examples of instantiations are provided in the subdirectories. Additional developments may take advantage of the documented examples provided at the different levels of the software.

6. Conclusion

We provide in the TReacLab code a fully segmented implementation of the coupling of independent geochemical and transport software. Coupling is based on a general expression of

the split-operator strategy with a set of classical methods. TReacLab should facilitate the development of reactive transport simulation capacities for independent reactive and transport software. Systematic comparison to the well-established PHREEQC model for uniform 1D reactive transport cases shows a good agreement with TReacLab. Systematic comparison against COMSOL for the 2D problem shows that full decoupling at the implementation level has a cost in accuracy. Sharp dissolution fronts of thermodynamically controlled reactions especially are generally smoothed in split-operator strategies. Steeper fronts might be recovered with smaller splitting time steps at larger computational costs. Beyond the implementation and the simulation capacity, consistency and validity of the numerical models should be systematically assessed. TReacLab can be freely accessed and used to promote the development of coupling methods and to provide additional modeling capacity for reactive transport coupling in geological media.

Appendix A: Implemented operator splitting methods

We detail the mathematical formulation for the sequential splitting (Geiser, 2009) :

$$\frac{\partial Z^1}{\partial t} = \mathcal{L}_1 Z^1, \qquad Z^1(x,t^n) = Z(x,t^n), \qquad t^n \leq t \leq t^{n+1}, \tag{A.1}$$

$$\frac{\partial Z^2}{\partial t} = \mathcal{L}_2 Z^2, \qquad Z^2(x,t^n) = Z^1(x,t^{n+1}), \qquad t^n \leq t \leq t^{n+1}, \tag{A.2}$$

$$Z(x,t^{n+1}) = Z^2(x,t^{n+1}), \tag{A.3}$$

the additive splitting (Faragó *et al.*, 2008a; Faragó *et al.*, 2008b) :

$$\frac{\partial Z^1}{\partial t} = \mathcal{L}_1 Z^1, \qquad Z^1(x,t^n) = Z(x,t^n), \qquad t^n \leq t \leq t^{n+1}, \tag{A.4}$$

$$\frac{\partial Z^2}{\partial t} = \mathcal{L}_2 Z^2, \qquad Z^2(x,t^n) = Z(x,t^n), \qquad t^n \le t \le t^{n+1}, \tag{A.5}$$

$$Z(x,t^{n+1}) = Z^1(x,t^{n+1}) + Z^2(x,t^{n+1}) - Z(x,t^n), \tag{A.6}$$

467 the Strang splitting (Strang, 1968) :

$$\frac{\partial Z^1}{\partial t} = \mathcal{L}_1 Z^1, \qquad Z^1(x,t^n) = Z(x,t^n), \qquad t^n \le t \le t^{n+1/2}, \tag{A.7}$$

$$\frac{\partial Z^2}{\partial t} = \mathcal{L}_2 Z^2, \qquad Z^2(x,t^n) = Z^1\left(x,t^{n+1/2}\right), \qquad t^n \le t \le t^{n+1}, \tag{A.8}$$

$$\frac{\partial Z^1}{\partial t} = \mathcal{L}_1 Z^1, \qquad Z^1\left(x,t^{n+1/2}\right) = Z^2(x,t^{n+1}), \qquad t^{n+1/2} \le t \le t^{n+1}, \tag{A.9}$$

$$Z(x,t^{n+1}) = Z^1(x,t^{n+1}), \tag{A.10}$$

and the symmetrically weighted splitting (SWS) (Csomós *et al.*, 2005) :

$$\frac{\partial Z^1}{\partial t} = \mathcal{L}_1 Z^1, \qquad Z^1(x,t^n) = Z(x,t^n), \qquad t^n \leq t \leq t^{n+1}, \tag{A.11}$$

$$\frac{\partial Z^2}{\partial t} = \mathcal{L}_2 Z^2, \qquad Z^2(x,t^n) = Z^1(x,t^{n+1}), \qquad t^n \leq t \leq t^{n+1}, \tag{A.12}$$

$$\frac{\partial Z^{2*}}{\partial t} = \mathcal{L}_2 Z^{2*}, \qquad Z^{2*}(x,t^n) = Z(x,t^n), \qquad t^n \leq t \leq t^{n+1}, \tag{A.13}$$

$$\frac{\partial Z^{1*}}{\partial t} = \mathcal{L}_1 Z^{1*}, \qquad Z^{1*}(x,t^n) = Z^{2*}(x,t^{n+1}), \qquad t^n \leq t \leq t^{n+1}, \tag{A.14}$$

$$Z(x,t^{n+1}) = \frac{Z^2(x,t^{n+1}) + Z^{1*}(x,t^{n+1})}{2}. \tag{A.15}$$

The alternating splitting algorithm (Valocchi and Malmstead, 1992) is based on a sequential splitting. It is defined by two successive splitting time steps with a permutation of the operator sequence between the splitting time steps.

Appendix B: Complementary notes on software organization

We successively describe the general toolbox organization, the coupler, transport and chemistry classes. We concretely show how operator splitting methods can be introduced and how other transport and geochemical codes can be connected.

B.1 Coupling methods

The coupler is at the center of TReacLab as it performs the temporal integration and calls the transport and chemistry solvers through the OS algorithm. In the pre-processing phase, it gets the initial conditions and the temporal constrains of the integration. It is also in charge of

storing the required results before formatting and outputting them in the post-processing phase. Because the coupler is at the core of the toolbox, its methods remain generic. Interactions with the transport and chemistry solvers are also fully generic thanks to template interfaces calling external software and managing the exchange of information. Calling external software relies on the so-called Solve_Engine method for both transport and chemistry software. Solve_Engine takes as inputs the concentration data and the time step over which the integration must be performed. It returns the updated concentrations, a flag to check the success of the integration and an error message in case of failure to activate and inform the error management procedure mentioned in the former section. The coupler is based on a fixed structure of concentration data. Whatever the structure of concentrations in the transport and chemical codes, the structure of concentrations within the coupler is always the same. It consists in a matrix with in columns chemical species and in rows the position within the domain (Figure B.1). The size of the matrix is equal to the number of cells times the number of chemical species and components passed through the coupler. Chemical species include solutes and fixed species. As this is the sole link between the chemical code and the coupler also in charge of temporary results storage for the post-processing, it must transfer all quantities necessary for the algorithm and for the later extraction. The format of the matrix is set in the pre-processing phase and it is fixed for the whole simulation. TReacLab does not support yet any modification of species number to transfer between codes. Even if some solute species are absent over some time of the simulation, they will be transferred. This choice does not limit the capacity of the software as long as the chemical system is known from the beginning but might have some consequences on its performance in cases where solute composition strongly evolves. The choice of generality and flexibility, here like in other places, has a cost in efficiency. All modifications of concentration format are eventually performed in the interfaces between the coupler and the transport and chemistry solvers

(Figure 1).

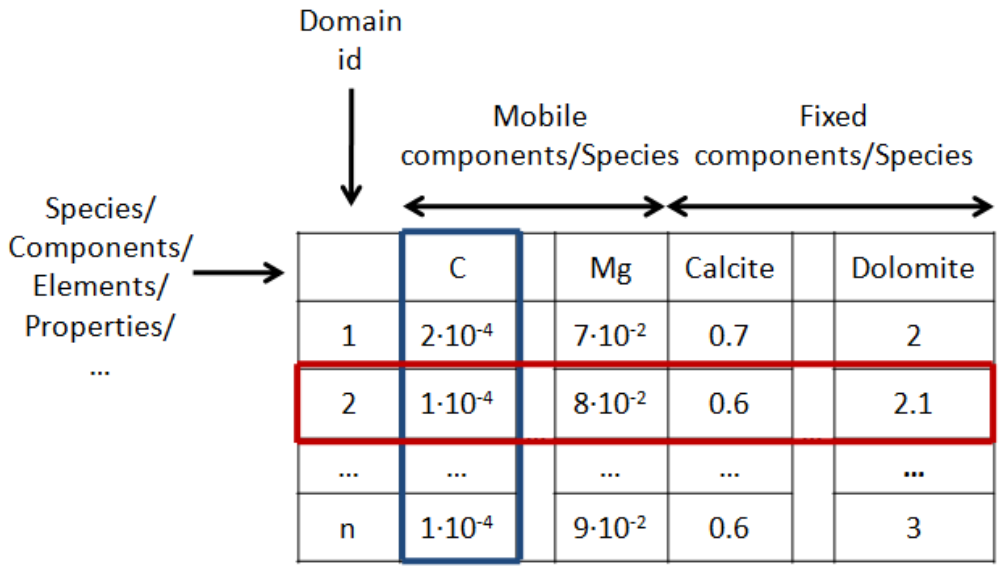


Figure B.1: Concentration format internal to the coupler class. To ensure generality, this structure of concentration is always the same and does not depend on the external transport and chemistry software. Species concentration are given in columns and are passed to the transport software as such. Concentrations at given locations are stored in rows with both mobile and fixed species. They are transferred either line per line or globally to the chemistry software. Fixed species are transferred from the chemistry code to the coupler to enable their possible use in the post-processing phase for results and outputs.

Thanks to the template methods calling the transport and chemical solvers and to the generic concentration format, operator splitting methods can be simply implemented. These are not more than a combination of simple calls of solvers passing and updating concentration information. Several sequential non-iterative techniques have thus been implemented, as detailed in section 2.3.

519 Specifications of the coupler are thus the name of the coupling method necessary to switch to
520 the corresponding method in the coupler class, the temporal constrains of the integration and a
521 vector of additional parameters. Temporal constrains of the integration are not only the initial
522 and final times of the integration but also the times at which the solution must be stored. All
523 time related parameters are stored into a time class. Additional parameters may be tolerances
524 for example when using sequential iterative approaches. Instantiation of the coupler class thus
525 consists in providing the identifier of the chosen coupling technique, the time constrains in the
526 time class (initial time, final time, time to save the results, OS time step) and the additional
527 parameters possibly needed by the algorithm.

528 **B.2 Geochemical solver**

529 Geochemical codes widely differ by their principles, the type of reactivity they consider and
530 their input/output formats and parameters. We propose to normalize some of their interface to
531 simplify exchanges with the coupler. In any case of equilibrium or kinetic reactions or of a
532 mixed combination of them, geochemical codes steadily take concentrations, reaction
533 constants, rate parameters, reaction times, as inputs and return output concentrations. All
534 specifications linked to the choice of components, primary and secondary species should be
535 set in the geochemical code or in the interface so that the geochemical solver does not have to
536 be modified and the coupler remains generic. Whether components are used or not, the
537 definition of the chemical system is not unique. Even when components are used, several
538 alternative and reliable definitions can be chosen (Fang *et al.*, 2003; Hoffmann *et al.*, 2012;
539 Molins *et al.*, 2004). Numerical and conceptual consistencies between the transport and
540 chemical systems should thus be ensured externally before any implementation.

541 While solute concentrations are instantiated by the coupler and systematically passed to the
542 geochemical solver, equilibrium and kinetic constants are considered as constant. They are

543 defined once for all in the pre-processing phase. For example in PHREEQC, chemical
544 reactions and constants are already defined in databases like 'Phreeqc.dat' or 'llnl.dat'.
545 Initialization of mineral quantities is done at the beginning of the simulation when setting the
546 initial conditions through the coupler. The interface between the coupler and the geochemical
547 solver is made up of the Solve_Engine that calls the geochemical solver and the methods that
548 modify the concentration format. By default, the geochemical solver is instantiated and stored
549 for each of the nodes of the computational grid for the whole domain of the simulation. Any
550 data that are not passed to the coupler is, in general, kept in the instances of the geochemical
551 code. Another option is provided by software that allow simultaneous computations for
552 several independent batches like it is for example the case of PHREEQC. In such cases only
553 one instance of the geochemical solver is necessary. Exchanges of data between the coupler
554 and the geochemical solver are defined in the pre-processing phase and remain fixed for the
555 whole duration of the simulation. It is precisely at this stage that components are derived
556 through the algebraic operations of equation (6) and passed to the coupler. The coupler does
557 not manage the transformation of concentration and species but just their transfer between the
558 transport and geochemical solvers. The use of components does not fundamentally change the
559 calling sequence of the geochemical operator but modifies its interface to the coupler.
560 Components may be specified by the geochemical code like in PHREEQC or by the user in
561 the pre-processing phase by loading the matrix of U (equation (6)). In this latter case,
562 components are defined by the user in the pre-processing phase and are computed by the
563 interface that adapts the information to be passed through the coupler to the transport solver.

564 Connection of a new geochemical code requires essentially four operations. First, a new
565 daughter class of the template chemistry class must be defined. It can be built up using, as
566 template, one of the examples provided and described in the section 4. Second, an interface
567 must be created to filter the required information given from the coupler to the Solve_Engine

568 method. Third, an instantiation procedure should be provided whether it is internal or external
569 to TReacLab. Fourth, the template Solve_Engine calling function of the geochemical solver
570 must be written and optionally tested before being effectively used in reactive transport
571 problems.

572 **B.3 Transport solver**

573 Despite the diversity of the transport mechanisms and numerical schemes to solve them, we
574 provide here a basic interface designed mostly to address transport in a generic way. As
575 previously stated, this approach assumes that transport parameters are not modified by the
576 species concentration. This absence of feedback currently precludes density driven flows as
577 well as permeability and porosity modifications due to precipitation or dissolution. TReacLab
578 might be extended in this direction on the basis of slow evolutions of porosity or density. The
579 transport operator relies on concentration independent parameters. We detail in the following
580 the interaction between the coupler and the transport classes with the exchange of data and the
581 instantiation of the transport solver. We will conclude this section with the development
582 required to connect other transport codes.

583 While geochemical codes operate on species concentration on a given computational node,
584 transport codes operate on a given species concentration over all the domain. In terms of data
585 structure, each of the columns of the concentration array are successively transferred to the
586 geochemical code and each of the rows (or linear combinations of rows) are given to the
587 transport code (Figure B.1). The transport operator is thus iteratively called for each of the
588 species or components explicitly specified in the interface between the coupler and the
589 transport solver (Figure 1). The time range over which temporal integration should be
590 performed and the identifiers of the transported species are also transferred to the transport
591 solver. Species identification is essential when considering species sensitive diffusion

coefficient. The transport solver returns the updated concentration field at the final time of the time range, an indicator of success or failure of the integration and a message to document algorithm failures. The basic exchange of concentrations with the imposed integration times are the sole requirements for the coupler to proceed.

All other parameters of the transport code should be set in the pre-processing phase, which may become an important part of the eventual reactive transport code. In fact it does not cover only the flow and transport parameters but more broadly the full structure of the domain, of the computational grid, and of the boundary conditions. As for the geochemical code, the transport code can be instantiated internally or externally. In case of internal definition, it should contain at least the flow and transport properties, the morphology of the domain and the structure of the computational grid (coordinates of the computational nodes). A default set of classes is provided for 1D problems as templates for the morphology (domain definitions), the computational grid (identification and coordinates of nodes and edges), the boundary conditions (nature and values for boundary conditions) and the hydraulic and transport properties. We recall as also said in section 2.2 that some operator splitting techniques might impose limitations on the transport solver in terms of integration scheme or in terms of time step (de Dieuleveult *et al.*, 2009). Both the OS technique and the transport integration should be chosen consistent.

Operations on the transport class are thus decomposed between the pre-processing and the processing phases. Specifications of the operator with all necessary parameters is performed in the pre-processing phase. Only generic exchanges of concentrations are needed in the processing phase. Additional information would generally be needed externally to identify the location of the computational nodes. More advanced information from the definition of the domain, parameters and boundary conditions will be generally defined in the transport code

rather than in TReacLab. For example, Comsol or Modflow have their own grid definitions. They are complete and efficient. It may be straightforwardly extracted and cross-referenced with the results of TReacLab as long as the cell numbers correspond, a basic but necessary requirement. This choice is motivated by both the generality and the simplicity of TReacLab. It also highlights that TReacLab remains a coupler that transfers information and does not process in any way the relation of concentrations between cells like a transport operator does.

The methodological choice of handling the spatial dimension of the problem within the transport operator is not only operational. It is also ensuring the capacity to connect a wide range of transport codes with their own logic and structure. For example, the multi-physics software COMSOL has its own mesh generator methods and internal structures that should not be duplicated in TReacLab but interfaced. Connecting other codes would thus require reduced work as long as they can already be called from the same environment of development (here MATLAB) on a discretized time basis. More in details, any new transport code would require: 1) the development of the main calling function Solve_Engine to call it from the coupler 2) the adaptation of the concentration format in the interface methods that match the concentrations to the internal data structure of the external code, 3) the instantiation of the transport class and 4) the access to the coordinates of the computational nodes for outputs purposes. As for the geochemical code, implementation of the interface should be checked before any full reactive transport coupling. This can be completed within TReacLab by using an idle process instead of the geochemical code.

Acknowledgements: We acknowledge ANDRA and the ANR project H2MNO4 under the number ANR-12-MONU0012-01 for their financial support, Jocelyne Erhel and David L. Parkhurst for constant and fruitful discussions, and Javier Molinero for initial exchanges.

639

640 **REFERENCES**

641 Abarca, E., Nardi, A., Grandia, F., Molinero, J., 2013. Feedback between reactive transport
642 and convective flow during CO2 migration in a saline aquifer, EGU General
643 Assembly Conference Abstracts, p. 7707.

644 Amir, L., Kern, M., 2010. A global method for coupling transport with chemistry in
645 heterogeneous porous media. Computational Geosciences 14, 465-481.

646 Apoung-Kamga, J.-B., Have, P., Houot, J., Kern, M., Semin, A., 2009. Reactive Transport in
647 Porous Media. ESAIM: Proceedings 28, 227 - 245.

648 Azad, V.J., Li, C., Verba, C., Ideker, J.H., Isgor, O.B., 2016. A COMSOL–GEMS interface
649 for modeling coupled reactive-transport geochemical processes. Computers &
650 Geosciences 92, 79-89.

651 Barry, D.A., Miller, C.T., Culligan-Hensley, P.J., 1996. Temporal discretisation errors in non-
652 iterative split-operator approaches to solving chemical reaction/groundwater transport
653 models. Journal of Contaminant Hydrology 22, 1-17.

654 Bea, S.A., Carrera, J., Ayora, C., Batlle, F., Saaltink, M.W., 2009. CHEPROO: A Fortran 90
655 object-oriented module to solve chemical processes in Earth Science models.
656 Computers & Geosciences 35, 1098-1112.

657 Bear, J., 1972. Dynamics of fluids in porous media. Elsevier, New York.

658 Belfort, B., Carrayrou, J., Lehmann, F., 2007. Implementation of Richardson extrapolation in
659 an efficient adaptive time stepping method: applications to reactive transport and
660 unsaturated flow in porous media. Transport in Porous Media 69, 123-138.

661 Bethke, C.M., 2007. Geochemical and biogeochemical reaction modeling. Cambridge
662 University Press.

663 Beyer, C., Li, D., De Lucia, M., Kühn, M., Bauer, S., 2012. Modelling CO2-induced fluid–
664 rock interactions in the Altensalzwedel gas reservoir. Part II: coupled reactive
665 transport simulation. Environmental Earth Sciences 67, 573-588.

666 Carrayrou, J., Mosé, R., Behra, P., 2004. Operator-splitting procedures for reactive transport
667 and comparison of mass balance errors. Journal of Contaminant Hydrology 68, 239-
668 268.

669 Charlton, S.R., Parkhurst, D.L., 2011. Modules based on the geochemical model PHREEQC
670 for use in scripting and programming languages. Computers & Geosciences 37, 1653-
671 1663.

672 Clement, T.P., Sun, Y., Hooker, B.S., Petersen, J.N., 1998. Modeling Multispecies Reactive
673 Transport in Ground Water. Ground Water Monitoring & Remediation 18, 79-92.

674 Commend, S., Zimmermann, T., 2001. Object-oriented nonlinear finite element
675 programming: a primer. Advances in Engineering Software 32, 611-628.

676 COMSOL, A., 2010. COMSOL Multiphysics-LiveLink for Matlab User's Guide, comsol 4.1
677 edition.

678 COMSOL, A., 2012. 4.3 User's Guide. Comsol.

679 Csomós, P., Faragó, I., 2008. Error analysis of the numerical solution of split differential
680 equations. Mathematical and Computer Modelling 48, 1090-1106.

681 Csomós, P., Faragó, I., Havasi, Á., 2005. Weighted sequential splittings and their analysis.
682 Computers & Mathematics with Applications 50, 1017-1031.

683 de Dieuleveult, C., Erhel, J., 2010. A global approach to reactive transport: application to the
684 MoMas benchmark. Computational Geosciences 14, 451-464.

685 de Dieuleveult, C., Erhel, J., Kern, M., 2009. A global strategy for solving reactive transport
686 equations. Journal of Computational Physics 228, 6395-6410.

687 Diersch, H.-J.G., 1996. Interactive, graphics-based finite-element simulation system
688 FEFLOW for modeling groundwater flow, contaminant mass and heat transport
689 processes. WASY Institute for Water Resource Planning and System Research Ltd.,
690 Berlin, Germany.

691 Engesgaard, P., Kipp, K.L., 1992. A geochemical transport model for redox-controlled
692 movement of mineral fronts in groundwater flow systems: A case of nitrate removal
693 by oxidation of pyrite. Water Resources Research 28, 2829-2843.

694 Fang, Y., Yeh, G.-T., Burgos, W.D., 2003. A general paradigm to model reaction-based
695 biogeochemical processes in batch systems. Water Resources Research 39, 1083.

696 Faragó, I., Gnandt, B., Havasi, Á., 2008a. Additive and iterative operator splitting methods
697 and their numerical investigation. Computers & Mathematics with Applications 55,
698 2266-2279.

699 Faragó, I., Thomsen, P.G., Zlatev, Z., 2008b. On the additive splitting procedures and their
700 computer realization. Applied Mathematical Modelling 32, 1552-1569.

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

703 Gasda, S.E., Farthing, M.W., Kees, C.E., Miller, C.T., 2011. Adaptive split-operator methods
704 for modeling transport phenomena in porous medium systems. Advances in Water
705 Resources 34, 1268-1282.

706 Geiser, J., 2009. Decomposition methods for differential equations: theory and applications.
707 CRC Press.

708 Hammond, G., Lichtner, P., Lu, C., Mills, R., 2012. Pflotran: reactive flow & transport code
709 for use on laptops to leadership-class supercomputers. Groundwater reactive transport
710 models, 141-159.

711 Hammond, G.E., Lichtner, P.C., Mills, R.T., 2014. Evaluating the performance of parallel
712 subsurface simulators: An illustrative example with PFLOTRAN. Water Resources
713 Research 50, 208-228.

714 Hoffmann, J., Kräutle, S., Knabner, P., 2010. A parallel global-implicit 2-D solver for reactive
715 transport problems in porous media based on a reduction scheme and its application to
716 the MoMaS benchmark problem. Computational Geosciences 14, 421-433.

717 Hoffmann, J., Kräutle, S., Knabner, P., 2012. A general reduction scheme for reactive
718 transport in porous media. Computational Geosciences 16, 1081-1099.

719 Hundsdorfer, W., Verwer, J.G., 2013. Numerical solution of time-dependent advection-
720 diffusion-reaction equations. Springer Science & Business Media.

721 Kazemi Nia Korrani, A., Sepehrnoori, K., Delshad, M., 2015. Coupling IPhreeqc with
722 UTCHEM to model reactive flow and transport. Computers & Geosciences 82, 152-
723 169.

724 Kazemi Nia Korrani, A., Sepehrnoori, K., Delshad, M., 2016. A Mechanistic Integrated
725 Geochemical and Chemical-Flooding Tool for Alkaline/Surfactant/Polymer Floods.
726 SPE Journal 21, 32-54.

727 Kool, J., Van Genuchten, M.T., 1991. Hydrus: One-dimensional Variably Saturated Flow and
728 Transport Model, Including Hysteresis and Root Water Uptake; Version 3.3. US
729 Salinity Laboratory.

730 Kräutle, S., Knabner, P., 2005. A new numerical reduction scheme for fully coupled
731 multicomponent transport-reaction problems in porous media. Water Resources
732 Research 41, W09414.

733 Kulik, D.A., Wagner, T., Dmytrieva, S.V., Kosakowski, G., Hingerl, F.F., Chudnenko, K.V.,
734 Berner, U.R., 2013. GEM-Selektor geochemical modeling package: revised algorithm
735 and GEMS3K numerical kernel for coupled simulation codes. Computational
736 Geosciences 17, 1-24.

737 Lie, K., 2014. An introduction to reservoir simulation using MATLAB: User guide for the
738 Matlab reservoir simulation toolbox (MRST), SINTEF ICT, Department of Applied
739 Mathematics, Oslo, Norway.

740 Liu, J., Ewing, R.E., 2005. An Operator Splitting Method for Nonlinear Reactive Transport
741 Equations and Its Implementation Based on DLL and COM, In: Zhang, W., Tong, W.,
742 Chen, Z., Glowinski, R. (Eds.), Current Trends in High Performance Computing and
743 Its Applications: Proceedings of the International Conference on High Performance
744 Computing and Applications, August 8–10, 2004, Shanghai, P.R. China. Springer
745 Berlin Heidelberg, Berlin, Heidelberg, pp. 93-102.

746 Marty, N.C.M., Munier, I., Gaucher, E.C., Tournassat, C., Gaboreau, S., Vong, C.Q., Giffaut,
747 E., Cochepin, B., Claret, F., 2014. Simulation of Cement/Clay Interactions: Feedback
748 on the Increasing Complexity of Modelling Strategies. Transport in Porous Media 104,
749 385-405.

750 Mayer, K., 2000. MIN3P V1. 0 User Guide. University of Waterloo, Department of Earth
751 Sciences 26.

752 McDonald, M.G., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference
753 ground-water flow model, In: Survey, U.S.G. (Ed.), Denver, Colorado.

754 MIKE(DHI), 2016. piChem: A FEFLOW Plugin for Advanced Geochemical Reactions, User
755 Guide.

756 Molins, S., Carrera, J., Ayora, C., Saaltink, M.W., 2004. A formulation for decoupling
757 components in reactive transport problems. Water Resources Research 40, W10301.

758 Morel, F.M., Hering, J.G., 1993. Principles and applications of aquatic chemistry. John Wiley
759 & Sons.

760 Multiphysics, C., 2008. Pesticide transport and reaction in soil. Earth Science Module Model
761 Library.

762 Muniruzzaman, M., Rolle, M., 2016. Modeling multicomponent ionic transport in
763 groundwater with IPhreeqc coupling: Electrostatic interactions and geochemical
764 reactions in homogeneous and heterogeneous domains. Advances in Water Resources
765 98, 1-15.

766 Nardi, A., Idiart, A., Trinchero, P., de Vries, L.M., Molinero, J., 2014. Interface COMSOL-
767 PHREEQC (iCP), an efficient numerical framework for the solution of coupled
768 multiphysics and geochemistry. Computers & Geosciences 69, 10-21.

769 Nasir, O., Fall, M., Evgin, E., 2014. A simulator for modeling of porosity and permeability
770 changes in near field sedimentary host rocks for nuclear waste under climate change
771 influences. Tunnelling and Underground Space Technology 42, 122-135.

772 Parkhurst, D.L., Appelo, C., 1999. User's guide to PHREEQC (Version 2): A computer
773 program for speciation, batch-reaction, one-dimensional transport, and inverse
774 geochemical calculations.

775 Parkhurst, D.L., Kipp, K.L., Engesgaard, P., Charlton, S.R., 2004. PHAST, a program for
776 simulating ground-waterflow, solute transport, and multicomponent geochemical
777 reactions. USGS Techniques and Methods 6, A8.

778 Parkhurst, D.L., Wissmeier, L., 2015. PhreeqcRM: A reaction module for transport simulators
779 based on the geochemical model PHREEQC. Advances in Water Resources 83, 176-
780 189.

781 Patel, R., Perko, J., Jacques, D., De Schutter, G., Ye, G., Van Breugel, K., 2013. Lattice
782 Boltzmann based multicomponent reactive transport model coupled with geochemical
783 solver for scale simulations, 5th International Conference on Computational Methods
784 for Coupled Problems in Science and Engineering. International Center for Numerical
785 Methods in Engineering (CIMNE), pp. 806-817.

786 Peterson, S., Hostetler, C., Deutsch, W., Cowan, C., 1987. MINTEQ user's manual. Pacific
787 Northwest Lab., Richland, WA (USA); Nuclear Regulatory Commission, Washington,
788 DC (USA). Div. of Waste Management.

789 Prommer, H., Davis, G., Barry, D., 1999. PHT3D—A three-dimensional biogeochemical
790 transport model for modelling natural and enhanced remediation. Contaminated Site
791 Remediation: Challenges Posed by Urban and Industrial Contaminants. Centre for
792 Groundwater Studies, Fremantle, Western Australia, 351-358.

793 Pruess, K., Oldenburg, C., Moridis, G., 1999. TOUGH2 user's guide version 2. Lawrence
794 Berkeley National Laboratory.

795 Register, A.H., 2007. A guide to MATLAB object-oriented programming. CRC Press.

796 Rouson, D., Xia, J., Xu, X., 2011. Scientific software design: the object-oriented way.
797 Cambridge University Press.

798 Saaltink, M.W., Ayora, C., Carrera, J., 1998. A mathematical formulation for reactive
799 transport that eliminates mineral concentrations. Water Resources Research 34, 1649-
800 1656.

801 Saaltink, M.W., Carrera, J., Ayora, C., 2001. On the behavior of approaches to simulate
802 reactive transport. Journal of Contaminant Hydrology 48, 213-235.

803 Saaltink, M.W., Yakirevich, A., Carrera, J., Ayora, C., 2011. Fluid flow, solute and heat
804 transport equations. Geochemical Modeling of Groundwater, Vadose and Geothermal
805 Systems, 83.

806 Simpson, M.J., Landman, K.A., 2008. Theoretical analysis and physical interpretation of
807 temporal truncation errors in operator split algorithms. Mathematics and Computers in
808 Simulation 77, 9-21.

809 Šimůnek, J., Jacques, D., Van Genuchten, M.T., Mallants, D., 2006. Multicomponent
810 geochemical transport modeling using HYDRUS-1D and HP1. J. Am. Water Resour.
811 Assoc 42, 1537-1547.

812 Šimůnek, J., Vogel, T., Van Genuchten, M.T., 1994. The SWMS-2D code for simulating
813 water and solute transport in two dimensional variably saturated media–Version 1.2.
814 Research Report 132, US Salinity Lab., Agric. Res. Serv. USDA, Riverside,
815 California, USA.

816 Skeel, R.D., Berzins, M., 1990. A method for the spatial discretization of parabolic equations
817 in one space variable. SIAM journal on scientific and statistical computing 11, 1-32.

818 Steefel, C., 2009. CrunchFlow software for modeling multicomponent reactive flow and
819 transport. User's manual. Earth Sciences Division. Lawrence Berkeley, National
820 Laboratory, Berkeley, CA. October, 12-91.

821 Steefel, C.I., DePaolo, D.J., Lichtner, P.C., 2005. Reactive transport modeling: An essential
822 tool and a new research approach for the Earth sciences. Earth and Planetary Science
823 Letters 240, 539-558.

824 Steefel, C.I., MacQuarrie, K.T., 1996. Approaches to modeling of reactive transport in porous
825 media. Reviews in Mineralogy and Geochemistry 34, 85-129.

826 Strang, G., 1968. On the construction and comparison of difference schemes. SIAM Journal
827 on Numerical Analysis 5, 506-517.

828 Thouvenot, P., Bildstein, O., Munier, I., Cochepin, B., Poyet, S., Bourbon, X., Treille, E.,
829 2013. Modeling of concrete carbonation in deep geological disposal of intermediate
830 level waste, EPJ Web of Conferences. EDP Sciences, p. 05004.

831 Trotignon, L., Devallois, V., Peycelon, H., Tiffreau, C., Bourbon, X., 2007. Predicting the
832 long term durability of concrete engineered barriers in a geological repository for
833 radioactive waste. Physics and Chemistry of the Earth, Parts A/B/C 32, 259-274.

834 Valocchi, A.J., Malmstead, M., 1992. Accuracy of operator splitting for advection-dispersion-
835 reaction problems. Water Resources Research 28, 1471-1476.

836 Van der Lee, J., 2002. CHESS Software for Geochemistry. Hydrology and Environmental
837 Science, École des Mines de Paris, Fontainebleau, France.

838 van der Lee, J., De Windt, L., Lagneau, V., Goblet, P., 2003. Module-oriented modeling of
839 reactive transport with HYTEC. Computers & Geosciences 29, 265-275.

840 Van Genuchten, M.T., Alves, W., 1982. Analytical solutions of the one-dimensional
841 convective-dispersive solute transport equation. United States Department of
842 Agriculture, Economic Research Service.

843 Wissmeier, L., Barry, D.A., 2011. Simulation tool for variably saturated flow with
844 comprehensive geochemical reactions in two- and three-dimensional domains.
845 Environmental Modelling & Software 26, 210-218.

846 Yeh, G.T., Tripathi, V.S., 1989. A critical evaluation of recent developments in
847 hydrogeochemical transport models of reactive multichemical components. Water
848 Resources Research 25, 93-108.

849 Zhang, F., 2012. Groundwater reactive transport models. Bentham Science Publishers.

850 Zheng, C., Wang, P.P., 1999. MT3DMS: a modular three-dimensional multispecies transport
851 model for simulation of advection, dispersion, and chemical reactions of contaminants
852 in groundwater systems; documentation and user's guide. DTIC Document.

853

Highlights

- Object-oriented implementation of non-intrusive couplings for reactive transport
- Validation for 3 equilibrium and kinetically controlled 1D cases with PhreeqC and one 2D case with COMSOL.
- Illustration of implementation flexibility with different transport and reaction codes