Benchmarking the reactive transport code SCEPTER v1.0.2

Yoshiki Kanzaki[1] and Christopher T. Reinhard[1]

[1]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA

Correspondence: Yoshiki Kanzaki (ykanzaki3@gatech.edu) and Christopher T. Reinhard (chris.reinhard@eas.gatech.edu)

Abstract. One-dimensional reactive transport codes are powerful tools for examining a range of geologic, biogeochemical, and agronomic phenomena. The reactive transport code SCEPTER (Soil Cycles of Elements simulator for Predicting TERrestrial regulation of greenhouse gases) has been recently developed for simulating a range of processes controlling soil biogeochemistry in managed lands, with a particular emphasis on soil pH management and enhanced weathering as a carbon sequestration strategy. While much of the basic framework implemented in SCEPTER is structurally and parametrically akin to existing reactive transport codes, its behavior has not been systematically benchmarked against other longstanding reactive transport models. Here, we quantitatively evaluate the performance of SCEPTER relative to a range of other reactive-transport models through a series of benchmarking experiments designed to assess the capacity of the code to simulate: (1) soil hydrology and fluid transport; (2) charge balance during cation exchange; and (3) mineral dissolution/precipitation, with (2) and (3) accompanied by diffusive/advective fluid transport and equilibria for aqueous speciation and gas dissolution into pore fluids. We show that the performance of SCEPTER is functionally identical to all other hydrological and reactive transport codes across the simulated benchmark conditions and discuss the emerging need for a reactive transport model benchmarking procedure that is fit for the purpose of predictive modeling of soil pH management in agricultural lands.

1 Introduction

Theoretical models of fluid transport coupled with biogeochemical reaction networks are indispensable for understanding complex, open geochemical systems where elements can move through multiple phases via biophysical transport processes with biogeochemical reactions occurring across phases. Tracking and predicting the mechanics of natural terrestrial weathering and element cycling through the Earth's "critical zone", the dynamics of marine sediment diagenesis, and the environmental impacts of human activity across scales all benefit enormously from the insights and predictive frameworks provided by reactive transport models (RTMs). Systematic documentation, benchmarking, uncertainty analysis, and validation of RTMs are

thus critical for developing robust scientific understanding and for integrating theoretical reactive transport frameworks into decision-making processes and environmental management strategies.

Conducting benchmark experiments for individual (or grouped) reactive transport codes represents an important first step toward more comprehensive model validation and evaluation of structural and parametric model uncertainty (cf. Class et al., 2009; Steefel et al., 2015b; Tian et al., 2019). Ideally, a well-developed benchmark problem can be used to evaluate the structural capabilities, numerical implementation, and accuracy of a given RTM against an accepted set of results across multiple independent code bases. Although this does not constitute a validation of any particular model for any specific application — in essence it represents an evaluation of inter-model precision rather than model predictive accuracy — it can be used to assess whether a given RTM or configuration adheres to accepted or established norms of behavior and performance in a well-defined context (e.g., Steefel et al., 2015a). Benchmarking problems with established results are thus extremely valuable for initial assessment of newly developed RTM codes and for evaluation of major developments or updates to more well-established RTM code bases.

Here, we present a series of benchmark exercises designed to evaluate the performance of the one-dimensional reactive transport code Soil Cycles of Elements simulator for Predicting TERrestrial regulation of greenhouse gases (SCEPTER) relative to the core functionalities of existing RTM frameworks. In Sec. 2 we describe the basic mechanics of SCEPTER as relevant to the benchmarking exercises presented here, describe modifications to the model implemented for the purposes of benchmarking, and outline the general procedure for each benchmarking simulation. In Sec. 3 we compare the results of our benchmarking simulations with relevant results from other reactive transport codes. In Sec. 4 we discuss the results in the broader context of model documentation, benchmarking, and validation, particularly in the context of predictive simulation of soil pH management in agricultural lands and intentional modification of soil inorganic carbon cycling. Lastly, in Sec. 5 we provide a short summary and some recommendations for future work.

2 Methods

The physical and biogeochemical architecture and numerical implementation at the core of SCEPTER have been extensively documented previously (Kanzaki et al., 2022, 2024), and the basic framework for physical transport and coupled geochemical reactions in SCEPTER is similar to that of a number of existing reactive transport codes (Steefel et al., 2015a; Kanzaki et al., 2022). However, SCEPTER has been designed specifically for use in understanding the behavior of managed agricultural soils. In particular, SCEPTER is equipped with a range of features for simulating soil charge balance and modifications to soil management, such as implementation of a wide range of physical mixing styles (bio-diffusion by soil fauna and/or manual tilling), explicit tracking of particle size distributions in solid phases added to the upper boundary of the model domain, and a protocol for in-silico simulation of conventional agronomic soil measurements (Kanzaki et al., 2024). These functionalities

are not commonly implemented in other reactive transport codes and are thus not examined here, although benchmarking these functionalities represents an important area for future work (see Sec. 4).

Here, we focus on the core functionalities commonly included in reactive transport codes (e.g., Steefel et al., 2015a). Specifically, we examine the performance of SCEPTER in simulating: (1) soil hydrology and fluid flow; (2) cation exchange and charge balance between pore fluid and the soil exchange complex; and (3) geochemical reaction and mineral dissolution/-precipitation. Soil hydrology (e.g., soil water content and infiltration rate) is a key factor determining the dynamics of soil solution chemistry and gas transport, and is thus critical for both natural weathering (e.g., Clow and Drever, 1996; Stonestrom et al., 1998; Maher, 2010) and carbon cycling in agricultural soils (e.g., Cipolla et al., 2021; Baek et al., 2023). It has also been known for decades that cation exchange and transient storage on the soil exchange complex can cause significant time lags in cation transport through surface and subsurface flow paths (Haynes and Goh, 1980; Appelo, 1994), and this process may have importance for the dynamics of carbon removal through enhanced weathering in cropland soils (Kanzaki et al., 2025). Lastly, the dissolution and secondary formation of minerals act to control input and output fluxes of base cations, trace elements, and inorganic carbon to the soil and are thus expected to govern instantaneous and time-integrated process rates in natural systems and managed lands.

2.1 Updates to relevant mechanics in the SCEPTER model

2.1.1 Soil hydrology

Soil hydrology is described using the mass conservation of water under the assumption of constant water density (e.g., Ireson et al., 2023):

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} + S \,, \tag{1}$$

where t and z are time (y) and soil depth (m), respectively, and θ, q, and S are the soil water content ($\mathrm{m^3\ m^{-3}}$), the vertical water flux ($\mathrm{m\ y^{-1}}$), and additional source/sink terms ($\mathrm{m^3\ m^{-3}\ y^{-1}}$), respectively. Darcy's law describes the water flux q as being proportional to the gradient of the hydraulic head H (m):

$$q = -K\frac{\partial H}{\partial z} = -K\left(\frac{\partial \psi}{\partial z} - 1\right), \tag{2}$$

where K is the hydraulic conductivity ($\mathrm{m\ y^{-1}}$) and ψ is the pressure head (m). Combining Eqs. (1) and (2), the governing equation for soil water content is given as the Richardson-Richards equation (or Richards equation) (Richardson, 1922; Richards, 1931):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left(K\left(\frac{\partial \psi}{\partial z} - 1\right)\right) + S \,. \tag{3}$$

Solving Eq. (3) requires that the hydraulic properties of the soil be parameterized and related to each other. By default, we adopt the following formulations of Schaap and Leij (2000):

$$\sigma = \left(1 + (\alpha\psi)^n\right)^{-m} , \tag{4}$$

$$\theta = \theta_\mathrm{r} + (\theta_\mathrm{s} - \theta_\mathrm{r})\,\sigma , \tag{5}$$

$$K = K_\mathrm{s}\sigma^L \left(1 - \left(1 - \sigma^{1/m}\right)^m\right)^2 , \tag{6}$$

where σ is the soil water saturation ratio ($\mathrm{m}^3\ \mathrm{m}^{-3}$), θ_r and θ_s are the residual and saturation θ, respectively, and K_s is the hydraulic conductivity (K) at water saturation. The parameters α (m^{-1}), n, m and L correspond to specific hydraulic properties of the soil, with $m = 1 - \frac{1}{n}$ (Schaap and Leij, 2000). Note that when $L = 0.5$ Eq. (6) is identical to the Mualem–van Genuchten model (van Genuchten, 1980; Schaap and Leij, 2000).

Richards equation is solved by the same time-implicit finite difference method as that adopted for solving the other governing equations described in Kanzaki et al. (2022). The upper boundary condition is given as a specified water flux, while a free-drainage boundary is specified for the lower boundary (Ireson et al., 2023). If enabled, the code can use the soil water saturation (σ) and flux (q) obtained by solving the Richards equation in the governing equations for aqueous and gaseous phases in each time integration step (see Kanzaki et al., 2022, 2024). Currently, the option to use Richards equation for soil water dynamics can be implemented when the code is compiled, at which point users can specify hydraulic properties (θ_r, θ_s, α, n, m and L) in the input file `h2odynpars.in`. The parameter values used in the benchmark experiment for soil hydrology in this study are used as default (see Sec. 2.2.1 and Code Availability).

2.1.2 Activity coefficients and redox equilibrium

The chemical activity of a given aqueous species i ($\{i\}$) can differ from the concentration ($[i]$, $\mathrm{mol\ L}^{-1}$) as a result of ionic interactions, particularly for concentrated solutions with elevated dissolved solids. Calculation of activity coefficients ($\gamma_i \equiv \{i\}/[i]$) in SCEPTER is based on Debye-Hückel theory, and includes options to use the formulations adopted for WATEQ (Truesdell and Jones, 1974), also referred to as the "B-dot" model (Bethke, 1996):

$$\log\gamma_i = -\frac{Az_i^2\sqrt{I}}{1 + Ba_i\sqrt{I}} + b_i I , \tag{7}$$

where z_i is the charge of species i, I is ionic strength ($= \frac{1}{2}\sum_i z_i^2[i]$, $\mathrm{mol\ L}^{-1}$), and a_i (Å) and b_i ($\mathrm{L\ mol}^{-1}$) are species-specific parameters (though a_i is somewhat related to the size of i, both a_i and b_i are in practice fitting parameters calibrated using laboratory data). The parameters A ($\mathrm{L}^{0.5}\ \mathrm{mol}^{-0.5}$) and B ($\mathrm{L}^{0.5}\ \mathrm{mol}^{-0.5}\ \mathrm{Å}^{-1}$) in Eq. (7) are given as follows (Truesdell and

Jones, 1974):

$$A = \frac{1.82483 \times 10^6 \rho^{1/2}}{(\epsilon T)^{3/2}} \ , \tag{8}$$

$$B = \frac{50.2916 \rho^{1/2}}{(\epsilon T)^{1/2}} \ , \tag{9}$$

where T is temperature (K), ρ is water density ($\mathrm{g\ cm^{-3}}$), and ϵ is the dielectric constant of water. By default, the parameters ρ and ϵ are calculated as a function of temperature based on Wagner and Pruß (2002) and Malmberg and Maryott (1956), respectively.

The model also includes an option to implement a Davies formulation where $a_i = 1/B$ and $b_i = 0.2$, eliminating the need to provide species-specific values for a_i and b_i in Eq. (7) (e.g., Bethke, 1996; Zeebe and Wolf-Gladrow, 2001). The current version of the SCEPTER code uses the Davies formulation by default, but in what follows we implement the B-dot formulation for benchmark experiments in order to render the simulation protocol as consistent as possible across models (Sec. 2.2). In addition, prior versions of the SCEPTER code implement redox coupling and transformations for, e.g., $\mathrm{Fe(II)/Fe(III)}$ and $\mathrm{S(-II)/S(VI)}$, through kinetics (Kanzaki et al., 2022), while here we enable redox coupling through chemical equilibrium in order to facilitate benchmark comparison across models (Sec. 2.2).

2.1.3 Mineral dissolution/precipitation kinetics

The rate laws describing mineral dissolution and precipitation are fundamental in determining the fate and transport of elements in reactive transport models. By default, mineral precipitation/dissolution kinetics for a solid phase j are formulated in SCEPTER according to the general form:

$$R_j = S_j k_j \left(1 - \Omega_j\right) \ . \tag{10}$$

Here, R_j ($\mathrm{mol\ m^{-3}\ y^{-1}}$) is the rate of dissolution ($R_j > 0$) or precipitation ($R_j < 0$) of solid phase j, Ω_j is the thermodynamic saturation state of j in pore fluid, k_j is a rate coefficient ($\mathrm{mol\ m^{-2}\ y^{-1}}$) that is a function of pore fluid pH and soil CO_2 (for carbonates), and S_j is the surface area of the species per unit bulk soil/rock volume ($\mathrm{m^2\ m^{-3}}$) that is exposed to pore fluid. The SCEPTER code relates S_j to the molar amount of solid phase j per unit bulk soil/rock volume (m_j; $\mathrm{mol\ m^{-3}}$) via a range of user-specified scaling options, and by default scales reactive surface area as a function of porosity and hydraulic radius (see Kanzaki et al., 2022).

We implement the following rate laws for mineral dissolution/precipitation in our benchmark simulations:

$$R_j = k_j(1 - \Omega_j) \ , \tag{11}$$

$$R_j = k_j(\phi_j/\phi_j^0)^{2/3}(1 - \Omega_j) \ , \tag{12}$$

$$R_j = \max\left[A_j \left[k_j^{\mathrm{acid}}\{\mathrm{H^+}\}^{n_j^{\mathrm{acid}}} + k_j^{\mathrm{alk}}\{\mathrm{H^+}\}^{n_j^{\mathrm{alk}}}\right](\phi_j/\phi_j^0)^{2/3}(1 - \Omega_j), 0\right] \ , \tag{13}$$

where R_j and Ω_j are given as above, ϕ_j is the volume fraction of j in bulk soil, with ϕ_j^0 denoting the initial value of ϕ_j, A_j ($\mathrm{m^2\ m^{-3}}$) is the surface area of j, and k_j ($\mathrm{mol\ m^{-3}\ y^{-1}}$) is the rate constant for dissolution/precipitation of j. In the full formulation (Eq. 13) the rate constant is parameterized as a function of pH, where $\{\mathrm{H^+}\}$ denotes activity of $\mathrm{H^+}$, k_j^{acid} and k_j^{alk} ($\mathrm{mol\ m^{-2}\ y^{-1}}$) represent the rate constants in acidic and alkaline pH ranges, respectively, and n_j^{acid} and n_j^{alk} represent the dependence on pH in the corresponding pH ranges. The default formulation in SCEPTER is very similar to Eq. (13), but it differs in that a pH-independent term is also included in the rate law (e.g., Palandri and Kharaka, 2004), surface area changes linearly with solid phase concentration rather than non-linearly (although the $2/3$ scaling exponent is a user option; see Kanzaki et al., 2022), and $\mathrm{H^+}$ is tracked with respect to concentration rather than activity (Kanzaki et al., 2022).

2.2 Benchmark experiments

Our benchmark experiments on (1) soil hydrology, (2) cation exchange, and (3) mineral dissolution/precipitation are based on experimental setups used by Ireson et al. (2023), Appelo (1994), and Mayer et al. (2015), respectively. More details on individual benchmarks are provided below.

2.2.1 Soil hydrology

We implement the "infiltration" benchmark exercise of Ireson et al. (2023), developed to benchmark the openRE model for solving the Richards equation (Eq. 3) against the Hydrus-1D model (e.g., Šimůnek et al., 2005). The benchmark simulation considers a $1.5\ \mathrm{m}$ deep soil column and a 10-year time series of daily precipitation as model input (Fig. 1a). Soil hydraulic properties are specified according to the values given in Table 1, along with a uniform initial ψ of $-3.59\ \mathrm{m}$ and a SCEPTER model domain of 30 layers. Input data as well as the results from openRE and Hydrus-1D for this benchmark problem are available in the openRE repository (see Code Availability of Ireson et al., 2023).

2.2.2 Cation exchange

Appelo (1994) introduces a widely used theoretical formulation of cation exchange along with activity modification linked to the surface potential of the exchange complex, and conducts multiple cation exchange experiments accompanied with solute advection and dispersion using the widely used PHREEQM software. Here, we implement an example simulation using PHREEQC v3.0 (Parkhurst and Appelo, 2013) in which a $1.1\ \mathrm{meq\ L^{-1}}$ cation exchanger initially equilibrated with a solution containing $1\ \mathrm{mM\ Na}$, $0.2\ \mathrm{mM\ K}$ and $1.2\ \mathrm{mM\ NO_3}$ is flushed with a $0.6\ \mathrm{mM\ CaCl_2}$ solution, driven by advection and dispersion with a Péclet number of 40. The experimental setup is freely accessible as an example script of PHREEQC v3.0 along with the software (Parkhurst and Appelo, 2013). Another reactive transport software OpenGeoSys (Kolditz et al., 2012) adopts this cation exchange experiment as its benchmark exercise as well (v$\geq 6.4.1$; Naumov et al., 2021), allowing us to compare

cation exchange behavior across all three models. Although this benchmark has previously been implemented in v1.0 of the SCEPTER code and compared with results from PHREEQC v3.0 (Kanzaki et al., 2024), we provide it here for completeness and in order to evaluate consistency between all three reactive-transport frameworks and between v1.0 and v1.0.2 of the SCEPTER reactive transport code. As discussed above, extensive description of the parameterization of cation exchange and charge balance at the soil exchange complex is given in Kanzaki et al. (2024).

2.2.3 Mineral dissolution/precipitation

Mayer et al. (2015) introduced a series of benchmark simulation exercises in which four different well-established reactive transport models — CrunchFlow (Steefel, 2009), Flotran (Lichtner, 2007), HP1 (Jacques et al., 2008), and MIN3P (Mayer et al., 2002) — are used to simulate acid rock drainage (ARD). The benchmark consists of three sets of numerical experiments (ARD-B1 to ARD-B3) with staged levels of complexity and involving different varieties of primary and secondary minerals during the generation of ARD and its neutralization. Simulated gaseous/aqueous species are fixed throughout ARD-B1 to ARD-B3, and include O_2 and CO_2 as gaseous/aqueous species and a comprehensive suite of Al, Ca, K, Cl, Si, S(VI)/S($-$II) and Fe(II)/Fe(III) aqueous species, with (uniform) initial and boundary conditions for gaseous/aqueous species remaining unchanged across all experiments. Initial concentrations of total dissolved Al, Ca, K, C, Cl, Si, and Fe are given as 2.59×10^{-8}, 1.43×10^{-2}, 9.00×10^{-3}, 2.49×10^{-3}, 1.14×10^{-3}, 1.93×10^{-3}, and 1.45×10^{-4} mol L^{-1}, respectively, while the initial total S concentration and soil partial pressure of O_2 (pO_2, atm) are tuned so that initial soil water pH and pe are fixed at 7.0 and -2.5, respectively.

Boundary conditions are established as fixed concentrations at the upper boundary and are set as 1.28×10^{-8}, 1.90×10^{-3}, 8.70×10^{-3}, 1.14×10^{-4}, 1.99×10^{-4}, and 5.00×10^{-7} mol L^{-1} for total dissolved Al, Ca, K, Cl, Si, and Fe, respectively. The total S concentration is tuned so that the boundary soil water pH is fixed at 5.0. Boundary conditions for gaseous O_2 and CO_2 are set as atmospheric partial pressures (p): $pO_2 = 0.21$ and $pCO_2 = 3.17\times10^{-4}$ atm, respectively. The calculation of activity coefficients is based on Eq. (7) with the parameters a_i and b_i given in the Supplementary Material of Mayer et al. (2015) for all aqueous species considered in the benchmark simulation. Additional details on all aqueous species included in the benchmark simulation and their thermodynamic constants can be found in the Supplementary Material of Mayer et al. (2015).

Throughout the benchmark experiments ARD-B1 to ARD-B3 the parameters relevant to aqueous and gaseous transport remain unchanged, characterized by a porosity of 0.5, free-phase diffusion coefficients in water and soil gas of 2.4×10^{-9} and 2.1×10^{-5} m^2 s^{-1}, respectively, and a longitudinal dispersivity of 5.0×10^{-4} m. Soil water content and flux are assumed to be at steady state, with a water table located around the middle (2.5 m) of the 5 m model domain (Fig. 2 of Mayer et al., 2015) and steady state q given as 0.3 m y^{-1}. The effect of tortuosity on soil fluid diffusion is formulated as a function of soil fluid saturation based on Millington (1959). The 5 m model domain is divided into 101 layers with equal thicknesses for numerical implementation of a finite difference method (Kanzaki et al., 2022). All simulations are conducted for 10 model years. The

steady-state water saturation profile, as well as the results of all ARD benchmark experiments from CrunchFlow, Flotran, HP1, and MIN3P, are available in the Supplementary Material of Mayer et al. (2015).

Simulation ARD-B1 is the simplest of the benchmark simulations and only considers acid rock drainage in mine tailings by simulating oxidation of pyrite with soil O_2 transported from the atmosphere — e.g., the only mineral involved is pyrite. We adopt Eq. (12) with $k_j = 3 \times 10^{-10}$ mol L^{-1} s^{-1} as the rate law for pyrite dissolution. The initial condition for pyrite is given as a uniform distribution set at a soil volume fraction of 0.2 %. It is important to note that boundary conditions are not required for solid species in any of the benchmark simulations explored here given that the transport of solid phases is not considered in the original simulations of Mayer et al. (2015). We return to a consideration of this phenomenon in Sec. 4.

Simulation ARD-B2 adds complexity to ARD-B1 by additionally considering pH buffering by dissolution of carbonates and hydroxides and the precipitation of secondary minerals including calcite and gibbsite as primary minerals and ferrihydrite, jarosite, and gypsum as potential secondary minerals. The dissolution rate law and initial concentration of pyrite remain unchanged in ARD-B2. The dissolution/precipitation rate law for calcite is given as Eq. (12) with $k_j = 1 \times 10^{-8}$ mol L^{-1} s^{-1}, while those of the other mineral phases are given as Eq. (11) with the same k_j value of 1×10^{-8} mol L^{-1} s^{-1}. Initial conditions for calcite and gibbsite are given as uniform distributions throughout the model soil column at soil volume fractions of 0.177 and 0.083 %, respectively.

Finally, simulation ARD-B3 adds further complexity to ARD-B2 by including pH buffering by silicate weathering through the addition of K-feldspar and muscovite as primary minerals and amorphous silica as a potential secondary phase, and switches gibbsite from a primary mineral to a potential secondary mineral. The dissolution/precipitation rate laws and initial conditions of the minerals simulated in ARD-B2 remain unchanged in ARD-B3 except that the initial gibbsite concentration is now zero as gibbsite is now specified as a secondary rather than a primary mineral phase. The dissolution rate laws of K-feldspar and muscovite are given as Eq. (13), with the following parameter values (Mayer et al., 2015); $A_j = 10$ m^2 L^{-1}, $k_j^{\mathrm{acid}} = 10^{-9.93}$ mol m^{-2} s^{-1}, $k_j^{\mathrm{alk}} = 10^{-16.5}$ mol m^{-2} s^{-1}, $n_j^{\mathrm{acid}} = 0.5$ and $n_j^{\mathrm{alk}} = -0.45$ for K-feldspar, and $A_j = 30$ m^2 L^{-1}, $k_j^{\mathrm{acid}} = 10^{-12.6}$ mol m^{-2} s^{-1}, $k_j^{\mathrm{alk}} = 10^{-13.5}$ mol m^{-2} s^{-1}, $n_j^{\mathrm{acid}} = 0.08$ and $n_j^{\mathrm{alk}} = -0.10$ for muscovite. Initial uniform concentrations of K-feldspar and muscovite are set to soil volume fractions of 2.68 and 7.31 %, respectively.

3 Results

3.1 Soil hydrology

Soil water storage ($= \int \theta dz$, m) and drainage from the model domain (q at depth of 1.5 m) are shown in Fig. 1b, c for openRE, Hydrus-1D and SCEPTER when forced with the same 10-year daily precipitation record (Fig. 1a). All three models show systematic increases in soil water storage and drainage following periods of intense rainfall. The results of SCEPTER are

essentially indistinguishable from those of openRE and Hydrus-1D, indicating that the optional implementation of Richards equation in SCEPTER is fully consistent with existing 1D hydrological transport codes.

3.2 Cation exchange

As also discussed in Kanzaki et al. (2024, 2025), the processes controlling cation exchange in soil may in many cases be critically important for the time dynamics of inorganic carbon cycling and mitigation of anthropogenic soil acidity. The simulations presented here indicate that replacement of K and Na by Ca requires variable pore volumes in all three models for each cation due to the different affinities of individual cations for cation exchangeable sites (Fig. 2). However, we find negligible differences in the breakthrough dynamics of $\mathrm{CaCl_2}$ between PHREEQC v3.0, OpenGeoSys v6.4.1 and SCEPTER v1.0.2. This indicates that the representation of charge balance of key cations at the soil exchange complex is consistent across the three model frameworks under the conditions simulated here. This result is further supported by data-model comparison of simulated cation exchange and empirical measurements from soil mesocosm experiments (Kanzaki et al., 2024).

3.3 Mineral dissolution/precipitation

3.3.1 ARD-B1

Chemical profiles after a 10-year ARD-B1 simulation by SCEPTER show a significant drop in soil $p\mathrm{O_2}$ within $2\ \mathrm{m}$ (Fig. 3d) caused by consumption through oxidation of pyrite (Fig. 3b). Pyrite oxidation also leads to corresponding drops in pH and pe (Fig. 3a). Vertical solute concentration profiles (Fig. 3c) emerge from the combined effects of solute supply from pyrite oxidation, aqueous speciation as impacted by pH and pe in soil pore fluids, and transport reflecting soil water content and vertical fluid flux. All of the reactive transport codes examined here perform similarly, with Flotran and HP1 showing very slight elevations in dissolved $\mathrm{SO_4}$ and $\mathrm{Fe(II)}$ at the bottom of the model domain relative to the other models. The results from SCEPTER are virtually indistinguishable from those of CrunchFlow and MIN3P with respect to solution chemistry, the abundance of solid, gaseous, and aqueous species, and the timeseries of dissolved $\mathrm{SO_4}$ and $\mathrm{Fe(II)}$ (Fig. 3, 4).

The simulated integrated release fluxes of $\mathrm{SO_4}$ and $\mathrm{Fe(II)}$ and consumption of $\mathrm{O_2}$ species are also very similar across all reactive transport codes (Table 2). Integrated $\mathrm{SO_4}$ release is virtually identical across CrunchFlow, Flotran, MIN3P, and SCEPTER, with $\mathrm{SO_4}$ export from HP1 very slightly elevated relative to the other models (Table 2). As implied by the solute profiles and timeseries results (Fig. 3, 4), differences in integrated $\mathrm{Fe(II)}$ release from the system are clustered, with CrunchFlow, MIN3P, and SCEPTER yielding nearly identical results while HP1 and Flotran show slight elevations in $\mathrm{Fe(II)}$ flux over the 10-year model integration period. The column-integrated $\mathrm{O_2}$ consumption is slightly lower in SCEPTER, though the magnitude of this difference is small relative to the overall range in $\mathrm{O_2}$ consumption rates across the reactive transport

codes included here (Table 2). Overall, the solute dynamics and tracer production/consumption are extremely similar across all reactive transport codes for simulation ARD-B1.

3.3.2 ARD-B2

The effect of adding calcite and gibbsite as primary minerals and introducing secondary mineral precipitation in ARD-B2 is clear from the difference in chemical profiles at 10 model years (Fig. 5) relative to those of ARD-B1 (Fig. 3). In particular, the profile of soil solution pH is much more complex, reflecting buffering to different values with depth as a result of the additional mineral equilibria — a calcite buffer at $\mathrm{pH} \sim 6$, a gibbsite buffer at $\mathrm{pH} \sim 4$, a ferrihydrite buffer at $\mathrm{pH} \sim 3$, and a jarosite-buffered region above $1\ \mathrm{m}$ with a $\mathrm{pH} \sim 2$ (Fig. 5a–c). The profile of soil redox potential (pe) is largely controlled by pyrite oxidation by O_2 and thus the general features of the soil solution pe profile are not strongly impacted by the addition of calcite and gibbsite except for small steps caused by step-wise changes in soil solution pH (Fig. 5a, b, e).

The inclusion of a more diverse suite of primary and secondary minerals has significant impacts on key solute and gas species. Because of the production of secondary minerals containing SO_4 and $\mathrm{Fe(III)}$ (Fig. 5c) — ferrihydrite ($\mathrm{Fe(OH)_3}$), jarosite ($\mathrm{KFe_3(SO_4)_2(OH)_6}$), and gypsum ($\mathrm{CaSO_4 \cdot 2H_2O}$) — the concentrations of total dissolved SO_4 and Fe are significantly reduced (Fig. 5c). In addition, the dissolution of primary calcite from pyrite-derived acid production produces a complex soil pCO_2 profile — the resulting CO_2 is released to the atmosphere via gas diffusion through shallower depths, but accumulates below the water table, reaching a peak of almost $1\ \mathrm{atm}$, and is then advected downward with pore fluid (Fig. 5e).

Although there are some slight differences between the results of the reactive transport codes — in particular for dissolved $\mathrm{Fe(II)}$ near the redox front and for soil pCO_2 just below the depth of the water table (Fig. 5d, e) — overall the codes agree very closely with respect to depth profiles of major gaseous, aqueous, and solid species. The simulated integrated release fluxes of SO_4 and $\mathrm{Fe(II)}$ and the integrated O_2/CO_2 fluxes all agree closely across models, with the CrunchFlow, MIN3P, and SCEPTER codes again aligning very closely with respect to integrated solute fluxes (Table 3). Once again, the overall solute dynamics and tracer production/consumption are extremely similar across all reactive transport codes for simulation ARD-B2, as are the depth profiles and abundances of key primary and secondary mineral phases.

3.3.3 ARD-B3

Further including dissolution of K-feldspar and muscovite and removing gibbsite as a primary mineral results in a higher pH at depth but a lower pH around 1 to $1.5\ \mathrm{m}$ depths (compare Figs. 6a and 5a). These differences reflect the combined buffering effects of primary mineral dissolution and secondary mineral precipitation with the modified mineral assemblage (Fig. 6b, c). Dissolution of K-feldspar and muscovite are driven by acid production through pyrite oxidation (Fig. 7), resulting in more

effective acid neutralization in ARD-B3 relative to ARD-B2. However, buffering at jarosite solubility decreases the pore fluid pH at shallower depths despite the additional buffering capacity of K-feldspar and muscovite (Mayer et al., 2015).

As in ARD-B2, soil solution pe is largely determined by pyrite oxidation by $\mathrm{O_2}$ and thus the general features of the redox profile remain unchanged except for small steps reflecting soil solution pH. Despite the lack of ferrihydrite formation total concentrations of dissolved Fe remain low (Fig. 6d), as enhanced jarosite formation at the top of the soil column due to $\mathrm{K^+}$ release mostly from K-feldspar offsets the lack of ferrihydrite precipitation (Fig. 6c). The accumulation of soil $\mathrm{CO_2}$ is significantly reduced near the water table across all models, resulting at least in part from significantly higher pore fluid pH and correspondingly higher solubility (Fig. 6e).

One of the more notable differences between reactive transport codes highlighted by Mayer et al. (2015) is a difference in $\mathrm{Fe(II)}$ and $\mathrm{SO_4}$ breakthrough at a depth of $2.5\ \mathrm{m}$, such that HP1 and Flotran show transient elevations in dissolved concentrations relative to CrunchFlow and MIN3P (Fig. 8). As with benchmark simulations ARD-B1 and ARD-B2, the results from SCEPTER are very close to those of CrunchFlow and MIN3P (Fig. 8). Interestingly, the integrated impacts of this slight elevation in dissolved $\mathrm{Fe(II)}$ and $\mathrm{SO_4}$ in HP1 and Flotran do not dramatically impact the time-integrated fluxes of $\mathrm{Fe(II)}$ and $\mathrm{SO_4}$ after 10 model years (Table 4). In any case, the depth profiles of major gaseous, aqueous, and solid species, the timeseries of dissolved $\mathrm{Fe(II)}$ and $\mathrm{SO_4}$ transport, and the cumulative fluxes of major solutes and gas species are all very similar across models for benchmark simulation ARD-B3, with the results of SCEPTER again aligning most closely to CrunchFlow and MIN3P.

4 Discussion

The suite of benchmark simulations provided here demonstrates that SCEPTER performs similarly to existing, well-established hydrologic and reactive transport codes with respect to fluid transport, cation exchange, and coupled biogeochemistry of key gaseous, aqueous, and solid phases. Soil water storage and drainage in SCEPTER with Richards equation enabled are identical to results from openRE and HYDRUS, while cation exchange during a benchmark breakthrough simulation is virtually identical to both PHREEQC and OpenGeoSys. Slight differences in solute concentrations and transport are observed across the five models examined here in the more complex coupled reactive transport experiments (ARD-B1 to ARD-B3), but given the level of process complexity in the models and the level of complexity in the benchmark simulations the results are essentially identical across the reactive transport codes examined here.

The acid rock drainage benchmark simulation of Mayer et al. (2015) was chosen as a benchmark here because it evaluates key acid-base processes — including silicate/carbonate transformations, pore fluid pH dynamics, and $\mathrm{CO_2}$ transport — coupled with hydrologic transport in a near-surface environment on anthropogenic timescales. As such, it provides one of the only existing frameworks for evaluating some of the key features of anthropogenic soil pH management in agricultural systems

that has been implemented across multiple well-established reactive transport codes. In this context, it is clear that the performance of SCEPTER is functionally identical to the other reactive transport codes. However, there are a range of key processes important for tracking the climate impacts of soil pH management in agricultural lands that cannot be assessed by this or any other existing benchmark exercise. These include (but are not limited to) solid phase transport, evolution of reactive surface area and particle size distributions of anthropogenic feedstocks, simulation of conventional soil geochemical and agronomic measurements from reactive transport model results, organic carbon cycling, nutrient amendment and cycling, dynamic crop uptake of base cations and nutrients, and the representation of other greenhouse gases (e.g., CH_4 and N_2O).

For instance, functionality allowing for multiple styles of solid phase transport is likely to be important for predictive representation of agricultural pH management. Transport of solid phases can include upward advection through uplift of parent rock and erosion of soil at the surface, and mixing of solid species in surface layers of soil through either natural bioturbation or intentional soil mixing through discing, tilling, and plowing. Although the timescale of solid phase advection in natural settings is usually assumed to be long enough that it can be ignored, in croplands agricultural management can lead to erosion rates that are often significantly faster than those in natural weathering environments (Montgomery, 2007; Nearing et al., 2017; Thaler et al., 2021; Quarrier et al., 2023). In addition, soil mixing practice in agricultural systems can span a wide range of intensities and mixing styles, from moldboard (inversion) ploughing to no-till, or hybrid approaches that depend on crop rotation and year (e.g., Köller, 2003). These practices have variable but clear effects on geochemical fluxes and stocks in managed soils (Abdalla et al., 2013; Huang et al., 2018; Guenet et al., 2021; Li et al., 2023; Meng et al., 2024), including impacts to local erosion rates (Melland et al., 2016). As a result, agricultural extension recommendations for soil pH management practice often vary according to tilling style (e.g., Ketterings et al., 2006).

In addition, the representation of specific (reactive) surface area (SSA) and how it evolves over time can also be critical, particularly for practices employing feedstocks that dissolve relatively slowly (e.g., basalt). This will likely need to include the ability to track particle size distributions of feedstocks from (measured) starting distributions as they dissolve (Přikryl et al., 2017; Ghanbarian et al., 2021; Lewis et al., 2021; Kanzaki et al., 2022), along with robust parameterization of shielding and reaction inhibition by formation of secondary mineral phases (Velbel, 1993; Nugent et al., 1998; Zhu et al., 2010; Daval et al., 2018). Optimizing practice in light of these impacts will be crucial for deployment of enhanced weathering in practice, given the need to account for upstream carbon emissions associated with feedstock grinding (Moosdorf et al., 2014; Beerling et al., 2020; Zhang et al., 2023; Li et al., 2024). Validation of these process representations will benefit strongly from the incorporation of key weathering tracers into reactive transport codes and isotope-specific benchmark exercises formulated in the context of managed soils out of steady state, particularly the systematics of strontium (Sr) and lithium (Li) isotopes.

There is also a practical need to evaluate varying techniques for simulating conventional soil geochemical and agronomic measurements based on key outputs from reactive transport codes. Soil pH provides one important illustrative example (e.g., Kanzaki et al., 2024). Reactive transport codes commonly simulate pore fluid pH, and model frameworks that include cation exchange have the potential to simulate the pool of exchangeable H^+ on the soil exchange complex. However, in an agronomic

context "soil pH" is an aggregate measurement that depends crucially on the laboratory buffer being used and the solid/fluid ratio employed during measurement. These can vary considerably from lab to lab, across commercial agronomic service labs and local cooperative extensions (Kissel and Sonon, 2008; Fernandez and Hoeft, 2009; Donohue, 2023) . The abundances of other exchangeable cations, soil organic pools, and nutrients are often similarly operational when measured in the field, and particularly in the context of soil management in agricultural or forested lands. As a result, there is a clear need for transparent, adequately benchmarked tools and protocols for inverting reactive transport model results for conventional soil measurements. Benchmarking this functionality across models will be especially critical for validation of reactive transport models against many common field agronomic and soil geochemical data.

In any case, we suggest that there is a clear need for simulation benchmarks designed specifically to evaluate the behavior of reactive transport codes in the context of soil pH management and enhanced weathering in agricultural systems. This could involve setting a baseline level of process representation for benchmarking of full modeling systems, or could employ a modular approach in which individual key components of different modeling systems — e.g., inorganic geochemical reaction and physical transport, organic carbon and/or nutrient subcycling, crop uptake and growth — are benchmarked separately and for specific practical applications. A well-designed model benchmarking procedure could also have the added benefit of better informing field trial design for robust model validation against empirical data — a process with particular salience given the need to effectively augment soil health and food security in coming decades and ongoing compensatory claims against fossil fuel emissions based on enhanced weathering practices.

5 Conclusions

This study presents a series of benchmark experiments of the reactive transport code SCEPTER in order to examine its performance in simulating soil hydrology, cation exchange, and key mineral dissolution/precipitation reactions coupled to physical transport. We find that the performance of SCEPTER is functionally identical to the existing models examined here — soil water dynamics in the SCEPTER reactive transport code are consistent with openRE and Hydrus-1D, cation dynamics through interaction with the soil complex, advection, and dispersion are consistent with results from PHREEQC and OpenGeoSys, and multi-phase reactive mass transport in partially saturated porous media during the generation and attenuation of acid rock drainage are consistent with MIN3P, CrunchFlow, Flotran and HP1. A key avenue for future work will be the design and implementation of benchmarking protocols appropriate for assessing the behavior of existing and newly developed reactive transport codes specifically in the context of soil pH management and greenhouse gas cycling in managed lands.

Code availability. The source codes of the model are available at GitHub (https://github.com/cdr-laboratory/SCEPTER) under the GNU General Public License v3.0. The specific version of the model used in this paper is tagged as “v1.0.2” and has been assigned a doi

(https://doi.org/10.5281/zenodo.16885370, Kanzaki and Reinhard, 2025). A readme file on the web provides the instructions for executing the simulations.

Author contributions. YK and CTR designed and implemented the model. YK and CTR designed the experiments. YK conducted the experiments and analysed the results. All authors contributed to the writing of the paper.

Competing interests. The authors declare no competing interests.

Acknowledgements. The authors wish to acknowledge support from the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology. CTR acknowledges funding from the U.S. Department of Energy and the Grantham Foundation.

References

Abdalla, M., Osborne, B., Lanigan, G., Forristal, D., Williams, M., Smith, P., and Jones, M. B.: Conservation tillage systems: a review of its consequences for greenhouse gas emissions, Soil use and management, 29, 199–209, https://doi.org/10.1111/sum.12030, 2013.

Appelo, C.: Cation and proton exchange, pH variations, and carbonate reactions in a freshening aquifer, Water resources research, 30, 2793–2805, https://doi.org/10.1029/94WR01048, 1994.

Baek, S. H., Kanzaki, Y., Lora, J. M., Planavsky, N., Reinhard, C. T., and Zhang, S.: Impact of climate on the global capacity for enhanced rock weathering on croplands, Earth's Future, 11, e2023EF003 698, https://doi.org/10.1029/2023EF003698, 2023.

Beerling, D. J., Kantzas, E. P., Lomas, M. R., Wade, P., Eufrasio, R. M., Renforth, P., Sarkar, B., Andrews, M. G., James, R. H., Pearce, C. R., Mercure, J.-F., Pollitt, H., Holden, P. B., Edwards, N. R., Khanna, M., Koh, L., Quegan, S., Pidgeon, N. F., Janssens, I. A., Hansen, J., and Banwart, S. A.: Potential for large-scale CO_2 removal via enhanced rock weathering with croplands, Nature, 583, 242–248, https://doi.org/10.1038/s41586-020-2448-9, 2020.

Bethke, C.: Geochemical Reaction Modeling: Concepts and Applications, Oxford University Press, ISBN 0-19-509475-1, 1996.

Cipolla, G., Calabrese, S., Noto, L. V., and Porporato, A.: The role of hydrology on enhanced weathering for carbon sequestration II. From hydroclimatic scenarios to carbon-sequestration efficiencies, Advances in Water Resources, 154, 103 949, https://doi.org/10.1016/j.advwatres.2021.103949, 2021.

Class, H., Ebigbo, A., Helmig, R., Dahle, H. K., Nordbotten, J. M., Celia, M. A., Audigane, P., Darcis, M., Ennis-King, J., Fan, Y., Flemisch, B., Gasda, S. E., Jin, M., Krug, S., Labregere, D., Beni, A. N., Pawar, R. J., Sbai, A., Thomas, S. G., Trenty, L., and Wei, L.: A benchmark study on problems related to CO_2 storage in geologic formations: Summary and discussion of the results, Computational Geosciences, 13, 409–434, https://doi.org/10.1007/s10596-009-9146-x, 2009.

Clow, D. and Drever, J.: Weathering rates as a function of flow through an alpine soil, Chemical Geology, 132, 131–141, https://doi.org/10.1016/S0009-2541(96)00048-4, 1996.

Daval, D., Calvaruso, C., Guyot, F., and Turpault, M.-P.: Time-dependent feldspar dissolution rates resulting from surface passivation: Experimental evidence and geochemical implications, Earth and Planetary Science Letters, 498, 226–236, https://doi.org/10.1016/j.epsl.2018.06.035, 2018.

Donohue, S. J.: Soil Testing and Plant Analysis, in: Virginia Agronomy Handbook, chap. 8, pp. 121–136, Virginia Cooperative Extension, https://www.pubs.ext.vt.edu/content/dam/pubs_ext_vt_edu/424/424-100/spes-299.pdf, 2023.

Fernandez, F. G. and Hoeft, R. G.: Managing Soil pH and Crop Nutrients, in: Illinois Agronomy Handbook, chap. 8, pp. 91–112, University of Illinois Cooperative Extension, http://extension.cropsciences.illinois.edu/handbook/pdfs/chapter08.pdf, 2009.

Ghanbarian, B., Hunt, A. G., Bittelli, M., Tuller, M., and Arthur, E.: Estimating specific surface area: Incorporating the effect of surface roughness and probing molecule size, Soil Science Society of America Journal, 85, 534–545, https://doi.org/10.1002/saj2.20231, 2021.

Guenet, B., Gabrielle, B., Chenu, C., Arrouays, D., Balesdent, J., Bernoux, M., Bruni, E., Caliman, J.-P., Cardinael, R., Chen, S., et al.: Can N2O emissions offset the benefits from soil organic carbon storage?, Global Change Biology, 27, 237–256, https://doi.org/10.1111/gcb.15342, 2021.

Haynes, R. and Goh, K.: Some observations on surface soil pH, base saturation and leaching of cations under three contrasting orchard soil management practices, Plant and Soil, 56, 429–438, https://doi.org/10.1007/BF02143036, 1980.

Huang, Y., Ren, W., Wang, L., Hui, D., Grove, J. H., Yang, X., Tao, B., and Goff, B.: Greenhouse gas emissions and crop yield in no-tillage systems: A meta-analysis, Agriculture, Ecosystems & Environment, 268, 144–153, https://doi.org/10.1016/j.agee.2018.09.002, 2018.

Ireson, A. M., Spiteri, R. J., Clark, M. P., and Mathias, S. A.: A simple, efficient, mass-conservative approach to solving Richards' equation (openRE, v1.0), Geoscientific Model Development, 16, 659–677, https://doi.org/10.5194/gmd-16-659-2023, 2023.

Jacques, D., Šimůnek, J., Mallants, D., and van Genuchten, M. T.: Modelling coupled water flow, solute transport and geochemical reactions affecting heavy metal migration in a podzol soil, Geoderma, 145, 449–461, https://doi.org/10.1016/j.geoderma.2008.01.012, 2008.

Kanzaki, Y. and Reinhard, C. T.: cdr-laboratory/SCEPTER: v1.0.2, https://doi.org/10.5281/zenodo.16885370, 2025.

Kanzaki, Y., Zhang, S., Planavsky, N. J., and Reinhard, C. T.: Soil Cycles of Elements simulator for Predicting TERrestrial regulation of greenhouse gases: SCEPTER v0.9, Geoscientific Model Development, 15, 4959–4990, https://doi.org/10.5194/gmd-15-4959-2022, 2022.

Kanzaki, Y., Chiaravalloti, I., Zhang, S., Planavsky, N. J., and Reinhard, C. T.: In silico calculation of soil pH by SCEPTER v1.0, Geoscientific Model Development, 17, 4515–4532, https://doi.org/10.5194/gmd-17-4515-2024, 2024.

Kanzaki, Y., Planavsky, N. J., Zhang, S., Jordan, J., Suhrhoff, T. J., and Reinhard, C. T.: Soil cation storage is a key control on the carbon removal dynamics of enhanced weathering, Environmental Research Letters, 20, 074 055, https://doi.org/10.1088/1748-9326/ade0d5, 2025.

Ketterings, Q. M., Reid, W. S., and Czymmek, K. J.: Lime guidelines for field crops in New York, Dept. Crop and Soil Sciences Extension Series E06-02. Cornell Univ., Ithaca, NY, 2006.

Kissel, D. E. and Sonon, L.: Soil Test Handbook for Georgia, Tech. rep., University of Georgia Cooperative Extension, https://aesl.ces.uga.edu/publications/soil/sthandbook.pdf, 2008.

Kolditz, O., Bauer, S., Bilke, L., Böttcher, N., Delfs, J. O., Fischer, T., Görke, U. J., Kalbacher, T., Kosakowski, G., McDermott, C. I., Park, C. H., Radu, F., Rink, K., Shao, H., Shao, H. B., Sun, F., Sun, Y. Y., Singh, A. K., Taron, J., Walther, M., Wang, W., Watanabe, N., Wu, Y., Xie, M., and Zehner, B.: OpenGeoSys: an open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (THM/C) processes in porous media, Environmental Earth Sciences, 67, 589–599, https://doi.org/10.1007/s12665-012-1546-x, 2012.

Köller, K.: Techniques of soil tillage, in: Soil Tillage in Agroecosystems, edited by El Titi, A., pp. 1–25, CRC Press, Boca Raton, FL, https://doi.org/10.1201/9781420040609, 2003.

Lewis, A. L., Sarkar, B., Wade, P., Kemp, S. J., Hodson, M. E., Taylor, L. L., Yeong, K. L., Davies, K., Nelson, P. N., Bird, M. I., Kantola, I. B., Masters, M. D., DeLucia, E., Leake, J. R., Banwart, S. A., and Beerling, D. J.: Effects of mineralogy, chemistry and physical properties of basalts on carbon capture potential and plant-nutrient element release via enhanced weathering, Applied Geochemistry, 132, 105 023, https://doi.org/10.1016/j.apgeochem.2021.105023, 2021.

Li, Z., Zhang, Q., Li, Z., Qiao, Y., Du, K., Yue, Z., Tian, C., Leng, P., Cheng, H., Chen, G., et al.: Responses of soil greenhouse gas emissions to no-tillage: A global meta-analysis, Sustainable Production and Consumption, 36, 479–492, https://doi.org/10.1016/j.spc.2023.02.003, 2023.

Li, Z., Planavsky, N. J., and Reinhard, C. T.: Geospatial assessment of the cost and energy demand of feedstock grinding for enhanced rock weathering in the coterminous United States, Frontiers in Climate, 6, 1380 651, https://doi.org/10.3389/fclim.2024.1380651, 2024.

Lichtner, P. C.: FLOTRAN user's manual: two-phase non-isothermal coupled thermal-hydrologic-chemical (THC) reactive flow and transport code, version 2, Los Alamos National Laboratory, 2007.

Maher, K.: The dependence of chemical weathering rates on fluid residence time, Earth and Planetary Science Letters, 294, 101–110, https://doi.org/10.1016/j.epsl.2010.03.010, 2010.

Malmberg, C. G. and Maryott, A. A.: Dielectric constant of water from 0° to 100°C, Journal of Research of the National Bureau of Standards, 56, 1–24, https://doi.org/10.6028/jres.056.001, 1956.

Mayer, K. U., Frind, E. O., and Blowes, D. W.: Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions, Water Resources Research, 38, 1174, https://doi.org/10.1029/2001WR000862, 2002.

Mayer, K. U., Alt-Epping, P., Jacques, D., Arora, B., and Steefel, C. I.: Benchmark problems for reactive transport modeling of the generation and attenuation of acid rock drainage, Computational Geosciences, 19, 599–611, https://doi.org/10.1007/s10596-015-9476-9, 2015.

Melland, A., Antille, D., and Dang, Y.: Effects of strategic tillage on short-term erosion, nutrient loss in runoff and greenhouse gas emissions, Soil Research, 55, 201–214, https://doi.org/10.1071/SR16136, 2016.

Meng, X., Meng, F., Chen, P., Hou, D., Zheng, E., and Xu, T.: A meta-analysis of conservation tillage management effects on soil organic carbon sequestration and soil greenhouse gas flux, Science of the Total Environment, 954, 176 315, https://doi.org/10.1016/j.scitotenv.2024.176315, 2024.

Millington, R.: Gas diffusion in porous media, Science, 130, 100–102, https://doi.org/10.1126/science.130.3367.100.b, 1959.

Montgomery, D. R.: Soil erosion and agricultural sustainability, Proceedings of the National Academy of Sciences, 104, 13 268–13 272, https://doi.org/10.1073/pnas.0611508104, 2007.

Moosdorf, N., Renforth, P., and Hartmann, J.: Carbon dioxide efficiency of terrestrial enhanced weathering, Environmental science & technology, 48, 4809–4816, https://doi.org/10.1021/es4052022, 2014.

Naumov, D., Bilke, L., Fischer, T., Wang, W., Lu, R., Grunwald, N., Meisel, T., Buchwald, J., Garibay, J., Rink, K., Nagel, T., Lehmann, C., Zill, F., Yoshioka, K., Kern, D., Shao, H., Beese, S., Randow, J., Hassanzadegan, A., Silbermann, C., and Thieda, J.: ufz/ogs: 6.4.1, https://doi.org/10.5281/zenodo.5562297, 2021.

Nearing, M. A., Xie, Y., Liu, B., and Ye, Y.: Natural and anthropogenic rates of soil erosion, International Soil and Water Conservation Research, 5, 77–84, https://doi.org/10.1016/j.iswcr.2017.04.001, 2017.

Nugent, M., Brantley, S. L., Pantano, C. G., and Maurice, P.: The influence of natural mineral coatings on feldspar weathering, Nature, 395, 588–591, https://doi.org/10.1038/26951, 1998.

Palandri, J. L. and Kharaka, Y. K.: A compilation of rate parameters of water–mineral interaction kinetics for application to geochemical modeling, Open-File Report 2004–1068, U.S. Geological Survey, Menlo Park, CA, USA, https://doi.org/10.3133/ofr20041068, 2004.

Parkhurst, D. L. and Appelo, C. A. J.: Description of input and examples for PHREEQC version 3: A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations, Tech. rep., U.S. Geological Survey, https://doi.org/10.3133/tm6A43, 2013.

Přikryl, J., Jha, D., Stefánsson, A., and Stipp, S.: Mineral dissolution in porous media: An experimental and modeling study on kinetics, porosity and surface area evolution, Applied Geochemistry, 87, 57–70, https://doi.org/10.1016/j.apgeochem.2017.05.004, 2017.

Quarrier, C. L., Kwang, J. S., Quirk, B. J., Thaler, E. A., and Larsen, I. J.: Pre-agricultural soil erosion rates in the midwestern United States, Geology, 51, 44–48, https://doi.org/10.1130/G50667.1, 2023.

Richards, L. A.: Capillary conduction of liquids through porous mediums, Physics, 1, 318–333, https://doi.org/10.1063/1.1745010, 1931.

Richardson, L. F.: Weather Prediction by Numerical Process, Franklin Classics, https://doi.org/10.1017/CBO9780511618291, 1922.

Schaap, M. G. and Leij, F. J.: Improved prediction of unsaturated hydraulic conductivity with the Mualem-van Genuchten model, Soil Science Society of America Journal, 64, 843–851, https://doi.org/10.2136/sssaj2000.643843x, 2000.

Steefel, C. I.: CrunchFlow Software for Modeling Multicomponent Reactive Flow and Transport: User's Manual, Lawrence Berkeley National Laboratory, Earth Sciences Division, Berkeley, CA, USA, version dated October 12, 2009, 2009.

Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., Lagneau, V., Lichtner, P., Mayer, K. U., Meeussen, J., Molins, S., Moulton, D., Shao, H., Šimůnek, J., Spycher, N., Yabusaki, S. B., and Yeh, G. T.: Reactive transport codes for subsurface environmental simulation, Computational Geosciences, 19, 445–478, https://doi.org/10.1007/s10596-014-9443-x, 2015a.

Steefel, C. I., Yabusaki, S. B., and Mayer, K. U.: Reactive transport benchmarks for subsurface environmental simulation, Computational Geosciences, 19, 439–443, https://doi.org/10.1007/s10596-015-9499-2, 2015b.

Stonestrom, D. A., White, A. F., and Akstin, K. C.: Determining rates of chemical weathering in soils—Solute transport versus profile evolution, Journal of Hydrology, 209, 331–345, https://doi.org/10.1016/j.advwatres.2021.103949, 1998.

Thaler, E. A., Larsen, I. J., and Yu, Q.: The extent of soil loss across the US Corn Belt, Proceedings of the National Academy of Sciences, 118, e1922375 118, https://doi.org/10.1073/pnas.1922375118, 2021.

Tian, H., Yang, J., Xu, R., Lu, C., Canadell, J. G., Davidson, E. A., Jackson, R. B., Arneth, A., Chang, J., Ciais, P., Gerber, S., Ito, A., Joos, F., Lienert, S., Messina, P., Olin, S., Pan, S., Peng, C., Saikawa, E., Thompson, R. L., Vuichard, N., Winiwarter, W., Zaehle, S., and Zhang, B.: Global soil nitrous oxide emissions since the preindustrial era estimated by an ensemble of terrestrial biosphere models: Magnitude, attribution, and uncertainty, Global Change Biology, 25, 640–659, https://doi.org/10.1111/gcb.14514, 2019.

Truesdell, A. H. and Jones, B. F.: WATEQ, a computer program for calculating chemical equilibria of natural waters, Journal of Research of the U.S. Geological Survey, 2, 233–248, 1974.

van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Science Society of America Journal, 44, 892–898, https://doi.org/10.2136/sssaj1980.03615995004400050002x, 1980.

Velbel, M. A.: Formation of protective surface layers during silicate-mineral weathering under well-leached, oxidizing conditions, American Mineralogist, 78, 405–414, 1993.

Wagner, W. and Pruß, A.: The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use, Journal of Physical and Chemical Reference Data, 31, 387–535, https://doi.org/10.1063/1.1461829, 2002.

Zeebe, R. E. and Wolf-Gladrow, D. A.: CO_2 in Seawater: Equilibrium, Kinetics, Isotopes, vol. 65 of *Elsevier Oceanography Series*, Gulf Professional Publishing (Elsevier), Amsterdam, ISBN 0-444-50946-1, 2001.

Zhang, B., Kroeger, J., Planavsky, N., and Yao, Y.: Techno-economic and life cycle assessment of enhanced rock weathering: a case study from the Midwestern United States, Environmental science & technology, 57, 13 828–13 837, https://doi.org/10.1021/acs.est.3c01658, 2023.

Zhu, C., Lu, P., Zheng, Z., and Ganor, J.: Coupled alkali feldspar dissolution and secondary mineral precipitation in batch systems: 4. Numerical modeling of kinetic reaction paths, Geochimica et Cosmochimica Acta, 74, 3963–3983, https://doi.org/10.1016/j.gca.2010.04.012, 2010.

Šimůnek, J., van Genuchten, M. T., and Sejna, M.: The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media, University of California-Riverside Research Reports, 3, 1–240, 2005.

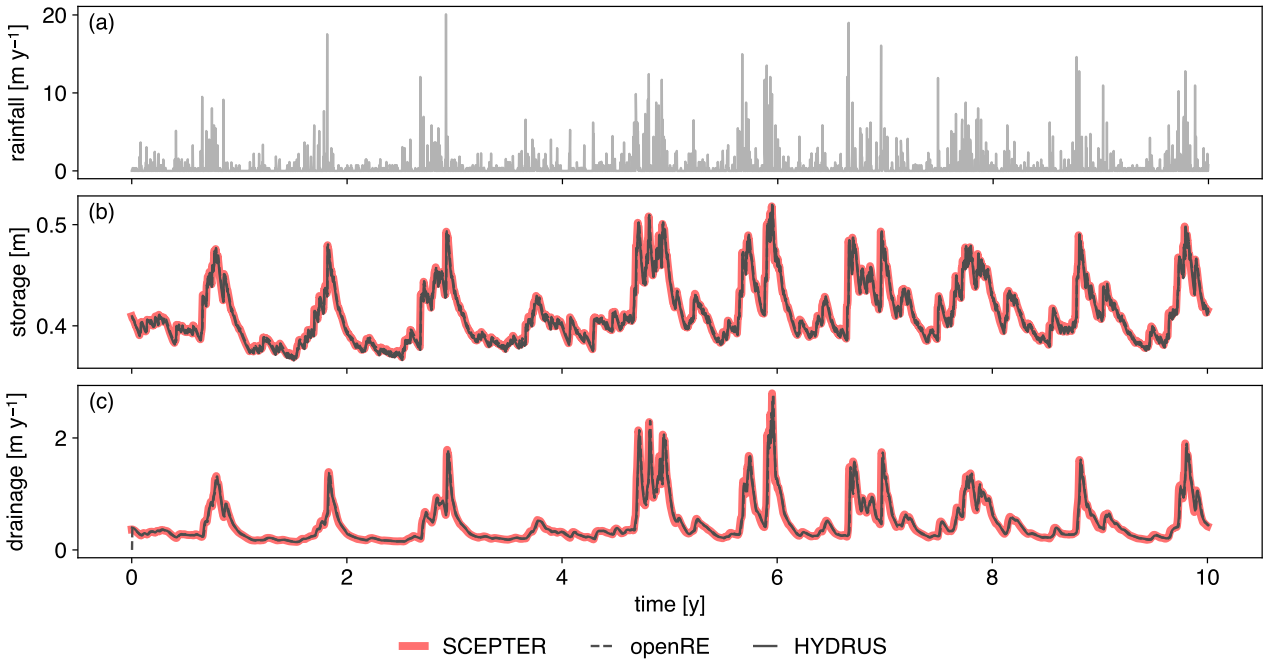

Figure 1. 10 year long simulations of (b) soil water storage and (c) drainage, in response to (a) daily precipitation, by openRE, Hydrus-1D and SCEPTER.

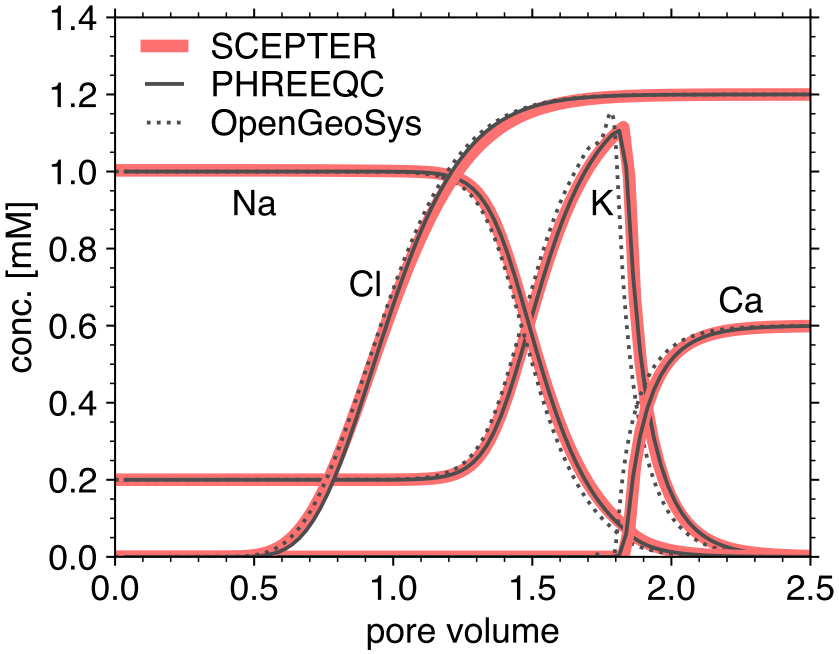

Figure 2. Simulation of $0.6\ \mathrm{mM}\ CaCl_2$ solution breaking through $1.1\ \mathrm{meq\ L^{-1}}$ cation exchanger initially equilibrated with solution containing $1\ \mathrm{mM}$ Na, $0.2\ \mathrm{mM}$ K and $1.2\ \mathrm{mM}\ NO_3$ via advection and dispersion with a Péclet number of 40 by PHREEQC, OpenGeoSys and SCEPTER.

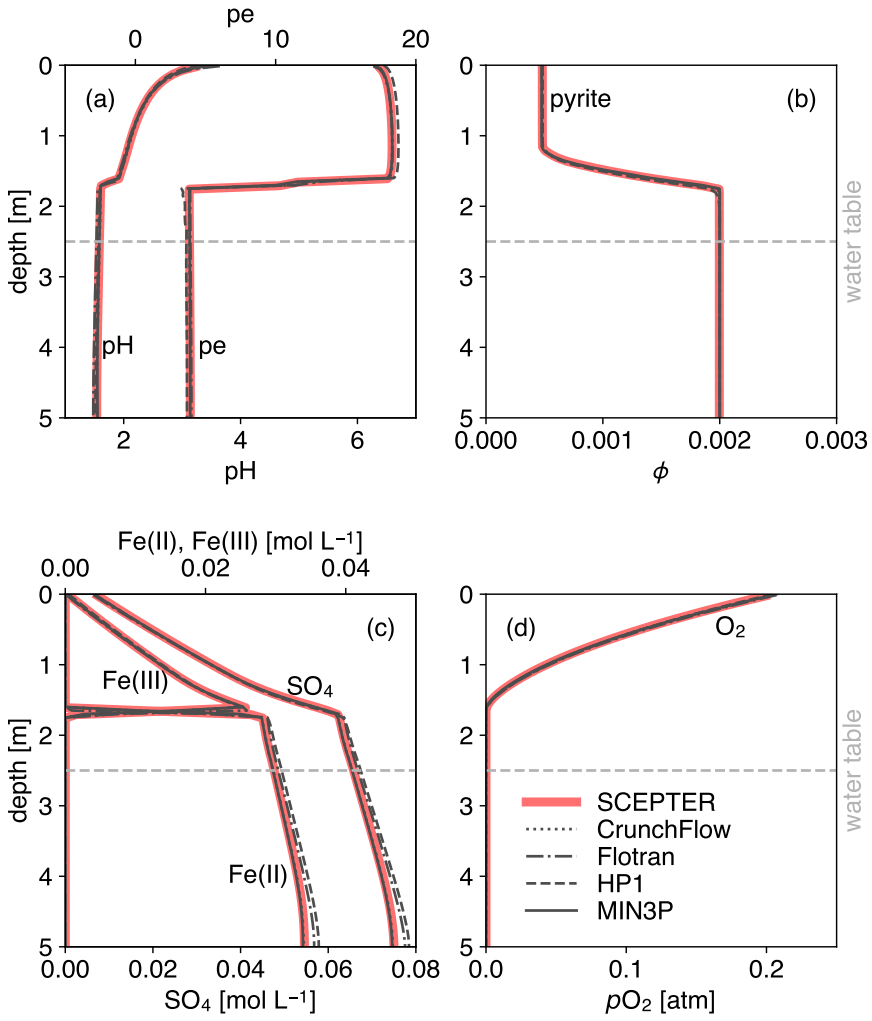

Figure 3. Depth profiles of (a) soil solution pH and redox potential (pe), (b) soil volume fraction (ϕ) of pyrite, (c) concentrations of total Fe(II), Fe(III) and SO_4 species and (d) partial pressure of soil gas O_2 (pO_2), at 10 model years of ARD-B1 simulation by CrunchFlow, Flotran, HP1, MIN3P and SCEPTER. See Sec. 2.2.3 for the details on experimental setup of ARD-B1.

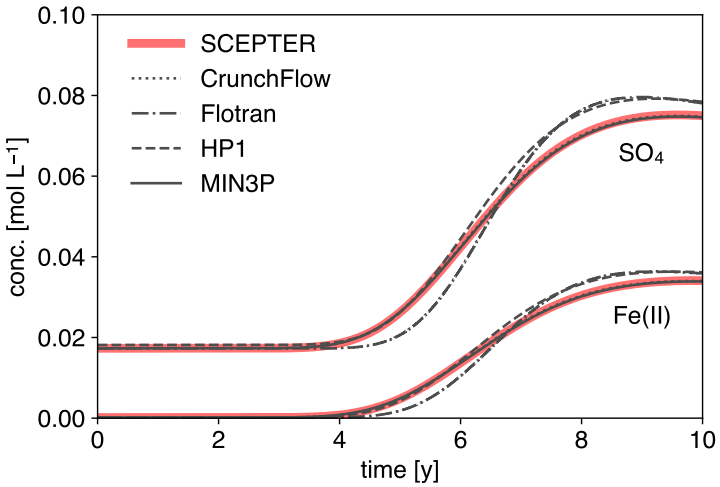

Figure 4. Time evolution of total $\mathrm{Fe(II)}$ and $\mathrm{SO_4}$ concentrations at the soil bottom through ARD-B1 simulation by CrunchFlow, Flotran, HP1, MIN3P and SCEPTER. See Sec. 2.2.3 for the details on experimental setup of ARD-B1.

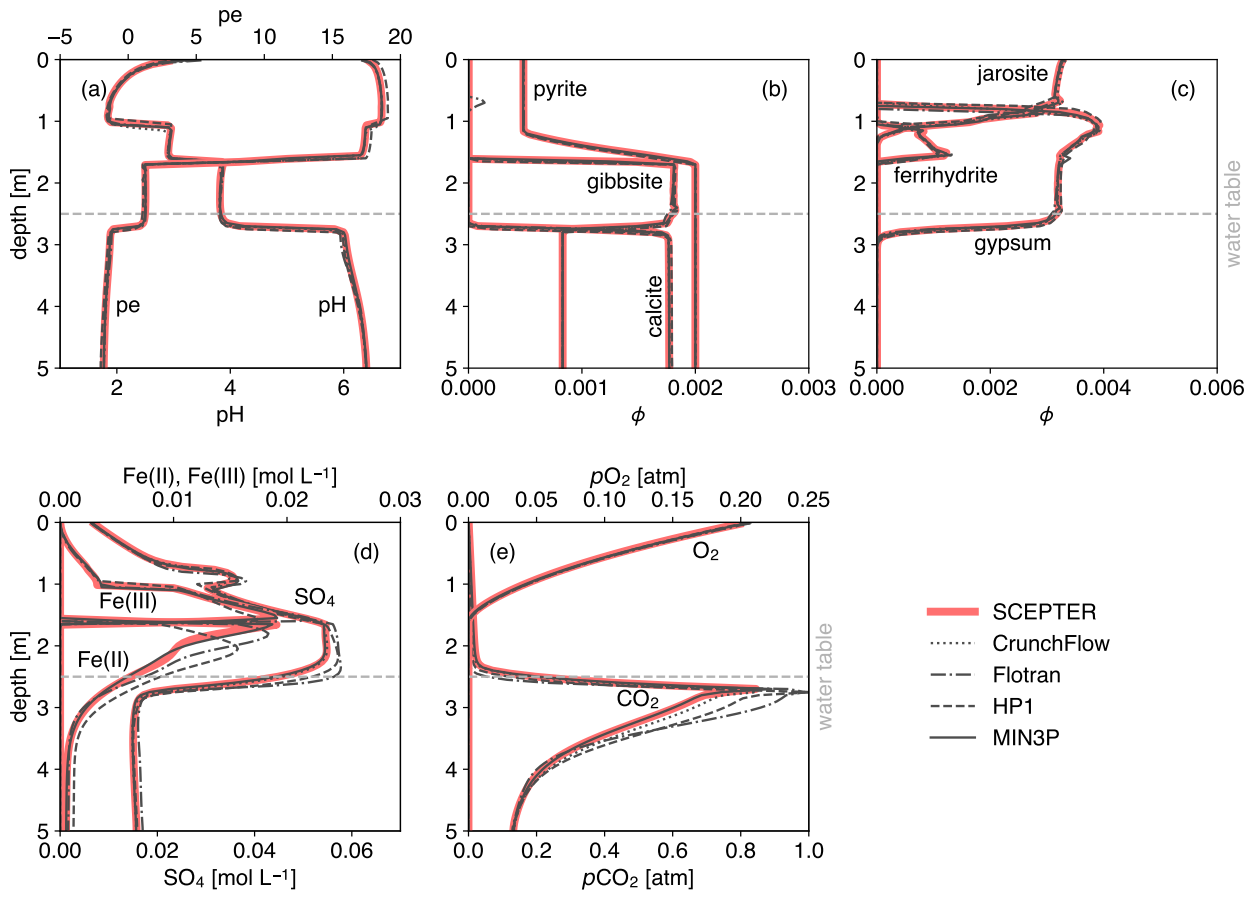

Figure 5. Depth profiles of (a) soil solution pH and pe, (b) soil volume fractions (ϕ) of pyrite, gibbsite and calcite, (c) soil volume fractions (ϕ) of jarosite, gypsum and ferrihydrite, (d) concentrations of total Fe(II), Fe(III) and SO_4 species and (e) partial pressures of soil gas O_2 (pO_2) and CO_2 (pCO_2), at 10 model years of ARD-B2 simulation by CrunchFlow, Flotran, HP1, MIN3P and SCEPTER. See Sec. 2.2.3 for the details on experimental setup of ARD-B2. Note that the results for total Fe(II), Fe(III) and SO_4 concentrations from CrunchFlow are not available in the Supplementary Material of Mayer et al. (2015) and thus are not shown in (d). These should be essentially identical to those of MIN3P (see Fig. 5 in Mayer et al., 2015).

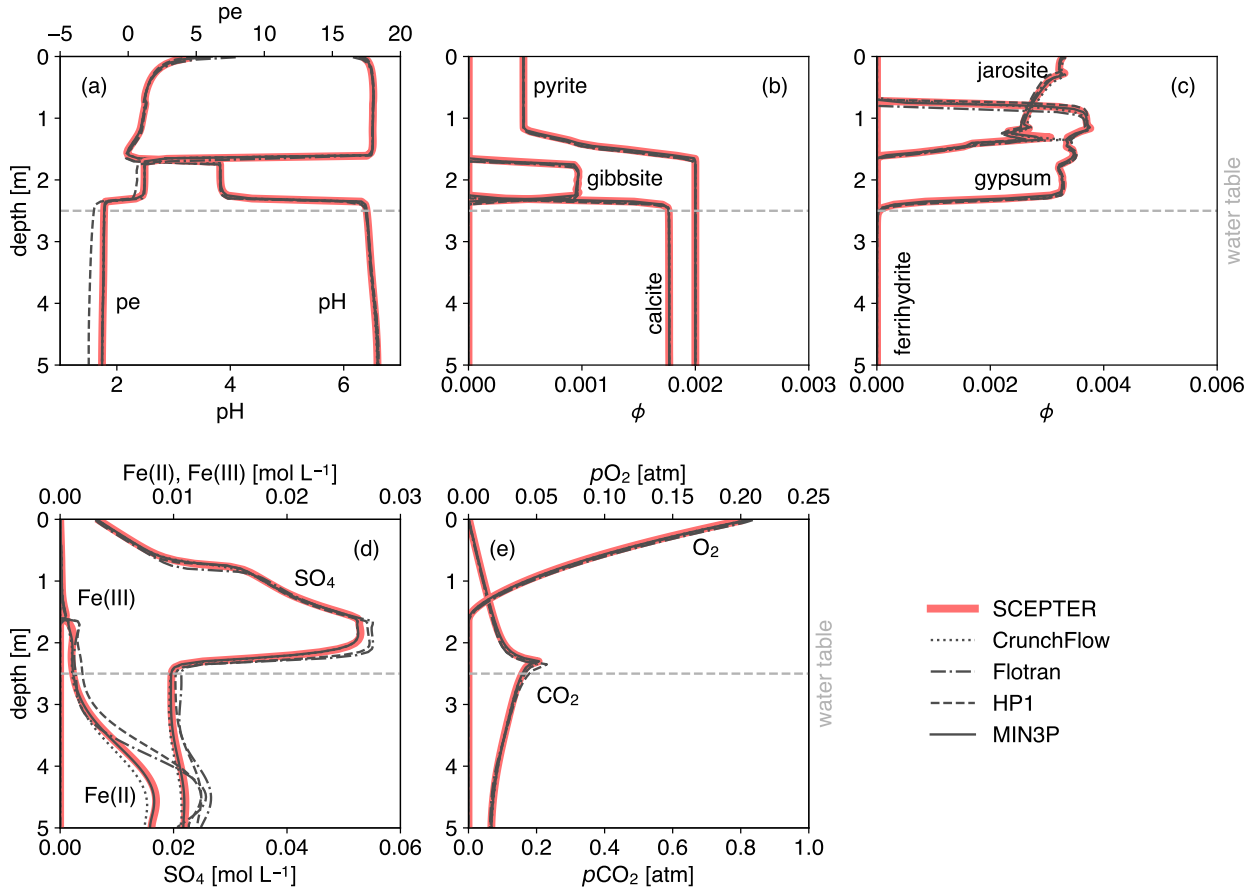

Figure 6. Depth profiles of (a) soil solution pH and pe, (b) soil volume fractions (ϕ) of pyrite, gibbsite and calcite, (c) soil volume fractions (ϕ) of jarosite, gypsum and ferrihydrite, (d) concentrations of total Fe(II), Fe(III) and SO_4 species and (e) partial pressures of soil gas O_2 (pO_2) and CO_2 (pCO_2), at 10 model years of ARD-B3 simulation by CrunchFlow, Flotran, HP1, MIN3P and SCEPTER. See Sec. 2.2.3 for the details on experimental setup of ARD-B3.

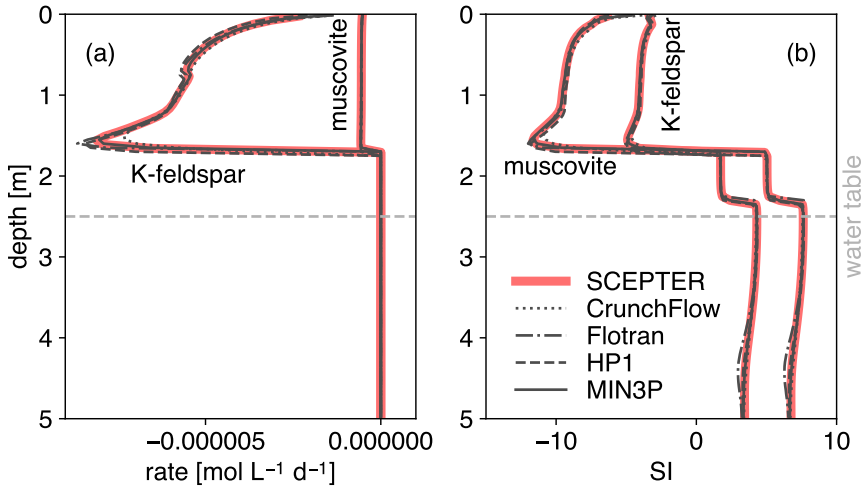

Figure 7. Depth profiles of (a) reaction rates and (b) saturation index ($\mathrm{SI} = \log \Omega$) of K-feldspar and muscovite, at 10 model years of ARD-B3 simulation by CrunchFlow, Flotran, HP1, MIN3P and SCEPTER. In (a) reaction rate is negative and positive when a mineral dissolves and precipitates, respectively, as in Mayer et al. (2015) (cf. Eq. 13). See Sec. 2.2.3 for the details on experimental setup of ARD-B3.

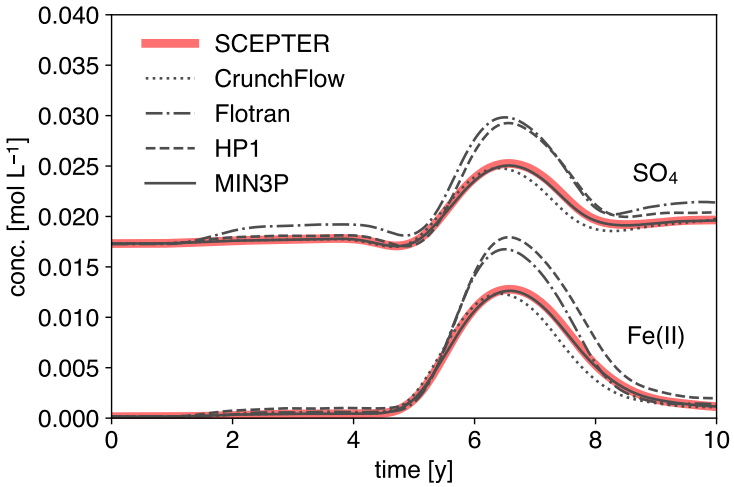

Figure 8. Time evolution of total $\mathrm{Fe(II)}$ and $\mathrm{SO_4}$ concentrations at $2.5\ \mathrm{m}$ soil depth through ARD-B3 simulation by CrunchFlow, Flotran, HP1, MIN3P and SCEPTER. See Sec. 2.2.3 for the details on experimental setup of ARD-B3.

Table 1. Soil hydraulic properties for benchmarking simulation of fluid transport.

Parameter	Description	Units	Value
θ_{r}	residual soil water content	$\mathrm{m}^3\ \mathrm{m}^{-3}$	0.131
θ_{s}	saturation soil water content	$\mathrm{m}^3\ \mathrm{m}^{-3}$	0.396
α	inverse air-entry pressure	m^{-1}	0.423
K_{s}	saturation hydraulic conductivity	$\mathrm{m\ d}^{-1}$	0.0496
L	pore connectivity factor	–	0.5
n	pore size distribution factor	–	2.06
m	[derived from n]	–	0.515

Table 2. Cumulative flux ($\mathrm{mol\ m^{-2}}$) for experiment ARD-B1 over 10 model years.

Model	O_2-ingress	SO_4-release	$Fe(II)$-release
CrunchFlow	344.0	116.7	37.3
Flotran	345.9	116.4	39.6
HP1	349.0	123.0	39.8
MIN3P	344.1	116.5	37.2
SCEPTER	342.3	116.9	37.3

Table 3. Cumulative flux ($\mathrm{mol\ m^{-2}}$) for experiment ARD-B2 over 10 model years.

Model	O_2-ingress	SO_4-release	$Fe(II)$-release	CO_2-egress
CrunchFlow	346.6	51.3	0.8	93.1
Flotran	346.1	52.7	1.0	92.9
HP1	350.0	53.1	1.6	95.1
MIN3P	343.8	51.4	0.8	92.8
SCEPTER	343.3	51.5	0.8	94.5

Table 4. Cumulative flux ($\mathrm{mol\ m^{-2}}$) for experiment ARD-B3 over 10 model years.

Model	O_2-ingress	SO_4-release	$Fe(II)$-release	CO_2-egress
CrunchFlow	344.3	53.6	3.6	85.0
Flotran	345.8	53.9	3.3	84.6
HP1	350.0	54.6	4.9	87.2
MIN3P	343.5	53.9	3.8	84.5
SCEPTER	343.2	53.9	3.8	85.9