A NOVEL REACTIVE TRANSPORT FRAMEWORK FOR FLUID-ROCK INTERACTION ANALYSIS: COMPUTATIONAL APPROACH, APPLICATIONS AND BENCHMARKS

Inauguraldissertation der Philosophisch-naturwissenschaftlichen Fakultät der Universität Bern

Vorgelegt von

LEONARDO HAX DAMIANI

aus Brasilien

Leiter der Arbeit:

Prof. Dr. Sergey Churakov

Institut für Geologie, Universität Bern

Co-Leiter:

PD Dr. Georg Kosakowski

Labor für Endlagersicherheit, Paul Scherrer Institut

A NOVEL REACTIVE TRANSPORT FRAMEWORK FOR FLUID-ROCK INTERACTION ANALYSIS: COMPUTATIONAL APPROACH, APPLICATIONS AND BENCHMARKS

Inauguraldissertation der Philosophisch-naturwissenschaftlichen Fakultät der Universität Bern

Vorgelegt von

LEONARDO HAX DAMIANI

von Brasilien

Leiter der Arbeit:
Prof. Dr. Sergey Churakov
Institut für Geologie, Universität Bern

Co-Leiter:
PD Dr. Georg Kosakowski
Labor für Endlagersicherheit, Paul Scherrer Institut

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen

Bern, den 22. Oktober 2021

Der Dekan:
Prof. Dr. Zoltan Balogh

Contents

Acknowledgments

Dear reader,

This thesis and its research was a long, challenging, and active journey that taught me a lot, in all possible aspects. It allowed me to work and get to know many interesting people. Therefore, I would like to thank:

Prof. Dr. Sergey Churakov for the systematic approach when solving problems and giving me the freedom to explore different ideas related to reactive transport modeling. The constructive and positive critical point of view was very much appreciated. Also, the runs (i.e., meetings) near PSI forest and the table tennis games at UniBern were a joy.

Dr. Georg Kosakowski for mentoring me with an outstanding work excellence and guiding me through the field of reactive transport modeling and computational simulations. My sincere gratitude for the recommendations, analytical thinking, patience, fruitful (and enduring) discussions, correction of the manuscripts, and encouragement to overcome several challenges during the course of my PhD study. In addition, many thanks for the bike rides (and endless bike-related chats), the family-like warmth (along with Monika and the kids), and making all of this possible by giving me the opportunity.

Dr. Nikolaos Prasianakis for the positive atmosphere, exciting machine learning and neural network discussions, and helping me keep the momentum and motivation through difficult moments.

Dr. Martin Glaus, Dr. Luc Van Loon, and Dr. Agnès Vinsot for their commitment when conducting the laboratory experiments and dedication when explaining things.

Dr. Dmitrii Kulik, Dr. George-Dan Miron, and Dr. Allan Leal for the many eye-opening discussions about GEMS and Reaktoro, and for sharing their geochemical and computational knowledge.

Dr. Thomas Gimmi, Dr. Wilfried Pfingsten, Dr. Alina Yapparova, Dr. Ravi Patel, Dr. Andrés Idiart and Dr. Vanessa Montoya for the meaningful feedback throughout the development of this thesis.

LES members and PSI colleagues for the work and non-work-related conversations and the countless cups of coffee (or mate) during the Znüni and Zvieri over the years. Many thanks to Alexandre, Annamaria, Andrea, Andrej, Amir, Ashish, Benjamin, Cornelia, Dario, Fulvio, Georgia, Guoqing, Latina, Mohamed, Petar, Philipp, Pietro, René, Robin, Simon E., Simon R., and Typhaine.

My parents — Vilmar and Cleonice — who have always been supportive, present (even with the Atlantic Ocean in between), and strong pillars throughout my life.

My sister — Cristiane — and friends for sharing good (and bad) times, laughs and friendly ears.

Thank you!

Financial support

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Training Programme of the European Atomic Energy Community (EURATOM) (H2020- NFRP-2014/2015) under grant agreement n° 662147 (CEBAMA).

Chapter 1

Introduction

1.1 Reactive transport modelling

Reactive transport modelling investigates coupled physical and chemical processes at different spatial and temporal scales by means of computer models (Steefel et al., 2005). Computational models of this kind rely on rigorous physical and chemical models that help to explain the behavior of the system (or parts of it), as well as to understand both the coupling between processes and their interactions.

Reactive transport models have been widely applied in broad context of scientific research and industrial applications, including pollutant migration and groundwater transport (Bear and Verruijt, 1987), nuclear waste disposal (Xu et al., 2008; Marty et al., 2009; Seigneur et al., 2019), heat transfer and thermal energy storage (Sheikholeslami et al., 2019), in-situ bioremediation and biogeochemical reactions (Barry et al., 2002; Gharasoo et al., 2012), engine combustion (Glarborg et al., 2018), CO_2 sequestration (Xu et al., 2011), oil reservoir prediction (Trangenstein, 1987), and diagenetic processes (Hensen et al., 1997).

Coupled geochemical modelling and mass transport has been one of the most dynamic research areas, combining geology, physical chemistry, and computer science. One of the main challenges is connected to the computational costs of geochemical modelling, which often impose limits on coupled transport modelling. Over the last 60 years — and especially in recent decades thanks to increases in computational power — thermodynamic modelling and reactive transport simulations have been the subject of ongoing research and have become standard tools in geochemistry (Millington, 1959; Hooyman, 1961; Ingri et al., 1967; Zeleznik and Gordon, 1968; Meissner et al., 1969; Zeggeren, 1970; Bos and Meershoek, 1972; Helgeson and Kirkham, 1974a,b, 1976; Crerar, 1975; Nordstrom et al., 1979; W. E. Dibble, 1981; Reed, 1982; Lichtner, 1985; Kharaka et al., 1988; Yeh and Tripathi, 1991; Wolery, 1992; Steefel and Lasaga, 1994; Appelo et al., 1998; Xu et al., 2006; Steefel and Maher, 2009; Kosakowski and Watanabe, 2014; Leal et al., 2016; Bilke et al., 2019; Damiani et al., 2020; Idiart et al., 2020).

Moreover, modern computational technologies, such as machine learning (ML) and artificial intelligence (AI) have significantly contributed to the efficiency of reactive transport modelling by making possible a more detailed description of the processes (Christo et al., 1996; Jatnieks et al., 2016; Guerillot and Bruyelle, 2016; Laloy and Jacques, 2019; Guérillot and Bruyelle, 2020; Leal et al., 2020; Prasianakis et al., 2020; Wei et al., 2020; Su et al., 2020; Yekta et al., 2021;

Lapeyre et al., 2021).

Application-driven implementations of different conceptual models and numerical approaches resulted in the development of powerful simulation packages, each having its specific advantages, functionalities, and limitations (Steefel et al., 2015; Damiani et al., 2020). Nevertheless, it is possible to notice that improvements and developments in a specific field often lead to compromises and commitments that set back the flexibility and customization of a code.

Reactive transport modelling can have a high computational cost, when it is applied to solve a very large system of coupled nonlinear equations in order to describe the coupled processes and combine spatial and temporal discretizations (Liu et al., 2016; Berkowitz et al., 2016; Hunt and Sahimi, 2017). Nonetheless, reactive transport modelling has proven to be a uniquely useful tool when it comes to analyzing fluid-rock interactions — for example, in the prediction of in-situ conditions in deep geological repositories used for nuclear waste disposal (Tsang et al., 2012; Jenni et al., 2017) — because it enables the evolution of the system to be predicted over a very long timeframe.

Attempts to overcome the limitations inherent in specific reactive transport modelling areas include, for example, the spatial dimension and temporal scale (Hoch et al., 2012; Li et al., 2017; Ge et al., 2020), the geochemical models and thermodynamic databases (Johnson et al., 1992; Thoenen et al., 2014; Lothenbach et al., 2019; Leal et al., 2017), the type of coupling between physical (transport) and chemical processes (Steefel et al., 2015), pore-scale properties and upscaling (Narsilio et al., 2009; Prasianakis et al., 2017), the numerical approach, robustness, usability, and high-performance computing (HPC) compatibility (Keyes et al., 2013; Nardi et al., 2014; Su et al., 2017).

Nevertheless, the field of reactive transport modelling field still faces a number of unresolved challenges (Keyes et al., 2013), such as understanding pore-scale properties and processes. One popular approach involves using X-ray micro-tomography in order to analyze the heterogeneities on the surface of porous media and make 3D models (Lai et al., 2015; Flukiger and Bernard, 2009; Godinho and Withers, 2018). However, tomographic methods are not yet accurate enough to resolve nanometric pores in clays and cement. Additionally, attempts to expand pore-scale processes to cover a continuum approach, which usually involves several orders of magnitude difference in the temporal and spatial scale, presents a number of difficulties and imposes inherent limitations on the description of the pore-medium properties (e.g., tortuosity, surface area). Such simulations are heavily dependent on the availability of experimental data for porous media properties (e.g., porosity, tortuosity, surface charge density), which may not be fully available or represented in an appropriate parametric form (Bossart et al., 2002; Mäder et al., 2017). Additionally, the software might not be compatible with modern HPC architectures, while the necessary computational power may be too expensive or even unavailable (Sohrabi et al., 2019).

Most codes tend to support functionalities relevant to a specific research domain. For this reason, initiatives aimed at developing state-of-the-art, general-purpose reactive transport modelling frameworks are essential for different research fields and thus highly in demand.

1.2 Objectives

In this context, the objectives for this thesis are:

- The development of a general-purpose reactive transport framework that:
  - couples physical and chemical processes accurately and with a high degree of complexity, including diffusive and advective transport and mineral precipitation and dissolution reactions
  - supports a flexible description of the temporal and spatial properties
  - is easy to use, useful, and allows a high level of customization without the need for advanced programming knowledge
  - provides a modular approach, in order to facilitate the usage of different chemical and transport solvers, as well as a straightforward abstraction of physical and chemical processes, along with their initial and boundary conditions
  - uses state-of-the-art computational tools and enables the exploration of new computational techniques in order to improve the performance and efficiency of the models
  - is validated with known benchmarks and experimental data
- The application of the developed framework to investigate:
  - the impact of electrochemical migration and electric potential on reactive transport
  - the Hydrogen Transfer (HT) experiment:
    * explore the mass transport mechanism between the borehole and the Opalinus Clay
    * explore the impact of (bio)geochemical reactions and the presence of microbes on the overall evolution of the chemical composition, pressure, and aqueous content in the borehole
    * verify and test hypotheses concerning how the system is affected by different physical and chemical processes
    * analyze long-term HT experimental data

1.3 Development of a reactive transport framework

In view of the overall objectives of this thesis, the aim is to implement an open-source, general-purpose reactive transport framework that goes beyond standard solvers by providing usage flexibility, state-of-the-art simulation methods, and computational efficiency using modern computer architectures. Even though multiple reactive transport codes are available, new computational technologies permit a much broader, more efficient, and more realistic way of describing and coupling the physical and chemical phenomena related to reactive transport modelling.

Hence, a newer, more modern, and more flexible reactive transport code that carries out numerical simulations without rewriting extensive codes could potentially have an enormous positive impact. It should provide the flexibility necessary to test parameters and hypotheses rapidly and predict the physical and chemical evolution of fluid-rock systems in a robust, accurate, and problem-oriented way. Python was used for the development of this framework. It is a

performant programming language, which is intuitive and self-explanatory to nonprofessional programmers. Furthermore, it allows the use of object-orientation concepts in order to produce a modular and reusable reactive transport modelling tool.

During the development of this framework, the design concepts were carefully planned to provide a flexible, customizable, and realistic way of describing the processes in different temporal and spatial scales. Furthermore, in light of the above-mentioned objectives, the developed reactive transport framework was subjected to stress scenarios, in order to validate and test its design concepts.

1.3.1 Transport of solutes in porous media

Diffusion, advection-dispersion, and electromigration are physical mass transport processes driven by the concentration gradient, pressure gradient, and electrostatic force (Steefel and Maher, 2009), respectively. The volumetric concentration c ($\mathrm{mol} \cdot \mathrm{m}^{-3}$) of a substance dissolved in water is:

$$c = \frac{m}{v} \tag{1.1}$$

where m is the mol amount and v is the volume of the solution. The flux J ($\mathrm{mol} \cdot \mathrm{m}^{-2} \cdot \mathrm{s}^{-1}$)is defined as the amount that passes through a selected area A (m^2) during a time interval Δt (s):

$$J = \frac{m}{A\Delta t} \tag{1.2}$$

The advective-dispersive flux in porous media is described as:

$$J_{adv} = \theta V c - D_m \frac{\partial c}{\partial t} \tag{1.3}$$

where J_{adv} is the advective-dispersive flux, θ is the porosity, V is the linear velocity of the moving front ($\mathrm{m} \cdot \mathrm{s}^{-1}$), and D_m is the mechanical dispersion coefficient ($\mathrm{m}^2 \cdot \mathrm{s}^{-1}$). D_m is defined as:

$$D_m = \alpha V \tag{1.4}$$

where α is the dispersivity or dispersion length, which can also be considered a diffusion process (Bear, 1972).

In the presence of a concentration gradient, the solutes are transported by diffusion and follow Fick's first law (Steefel et al., 2015). The diffusive flux is proportional to the substance's concentration gradient and is known as Fick's first law. The diffusive flux is described as:

$$J_{dif} = -D_e \frac{\partial c}{\partial t} \tag{1.5}$$

where $\mathrm{D_e}$ is the effective diffusion coefficient in porous media ($\mathrm{m}^2 \cdot \mathrm{s}^{-1}$). To account for the complexity of the pore space, following Appelo et al. (2010), the effective diffusion coefficient is often derived as:

$$De = D_0 \cdot (\theta \frac{\delta}{\tau^2}) = D_0 \cdot (\theta \frac{1}{G}) \tag{1.6}$$

where D_0 is the solute's diffusion coefficient in bulk water ($\mathrm{m}^2 \cdot \mathrm{s}^{-1}$), τ is the tortuosity ($\mathrm{m} \cdot \mathrm{m}^{-1}$) and δ is the constrictivity ($\mathrm{m} \cdot \mathrm{m}^{-1}$). Tortuosity is the ratio of the diffusion paths within the pore

space and the straight distance. Constrictivity is the ratio of the diameter of the diffusing particle to the pore diameter. Tortuosity and constrictivity are often combined into G, the so-called geometric factor. G can be related to the accessible porosity (θ_a) via Archie's law (Archie, 1942; Epstein, 1989; Van Loon et al., 2007):

$$\frac{\delta}{\tau^2} = \frac{1}{G} = {\theta_a}^n \tag{1.7}$$

where n is an empirical factor. D_0 may be experimentally measured or estimated using the Stokes-Einstein relation (Reid et al., 1977). D_0 is the intrinsic mobility of the solutes, which means that it is independent of the concentration gradient, as in:

$$D_0 = \frac{k_b T}{6\pi\mu R_0} \tag{1.8}$$

where $\mathrm{k_b}$ is the Boltzmann constant ($\mathrm{J \cdot K^{-1}}$), T is the temperature (K), μ is the solvent viscosity ($\mathrm{N \cdot s \cdot m^{-2}}$), and R_0 is the diffusing particle radius (m). Combining equations 1.7 and 1.8, $\mathrm{D_e}$ is:

$$D_e = \frac{k_b T \theta}{6\pi\mu R_0 G} \tag{1.9}$$

For charged particles in the presence of an electric field, the electromigration flux must be taken into account, since it has a major impact on ion transport. In a multi-component system with charged species in a porous media, the Fickian diffusion is insufficient to describe the electrostatic interaction between the ions during the transport (Liu et al., 2011). For example, it was found that the effect of charged mineral surfaces is not negligible and needs to be taken into account for radionuclide diffusion in clay-based materials (Van Loon et al., 2007; Glaus et al., 2010; Tournassat and Appelo, 2011; Gimmi and Kosakowski, 2011). The Nernst-Planck equation, which describes the diffusive transport of a charged ion in the presence of an electric potential, is:

$$J_{elec} = -D_e \cdot (\frac{cFz}{RT}\nabla\Phi) \tag{1.10}$$

where F is the Faraday constant ($\mathrm{C \cdot mol^{-1}}$), R is the gas constant ($8.31\mathrm{J \cdot K^{-1} \cdot mol^{-1}}$), z is the charge (C), and Φ is the electrical potential (V). The relation between the concentration of charged ions and the electric field is described, as in Samson and Marchand (1999), via the Poisson equation:

$$\nabla^2\Phi + \frac{F}{\epsilon}(\sum_{i=1}^{m} z_i c_i + w) = 0 \tag{1.11}$$

where ϵ is the dielectric constant (or relative permitivity) of the material ($\mathrm{C \cdot V^{-1} \cdot m^{-1}}$) and w is the charge density of the medium ($\mathrm{C \cdot m^{-1}}$). This equation 1.11 is necessary to describe, for example, what occurs when an electric field is applied as a driving force in combination with the concentration gradient (Maes et al., 2004). Assuming that both the electroneutrality condition and the null current condition are satisfied, respectively:

$$\sum_{i=1}^{m} z_i c_i = 0 \tag{1.12}$$

$$F \sum_{i=1}^{m} z_i J_{i_{tot}} = 0 \tag{1.13}$$

where c_i is the concentration of the i_{th} species among the total number of species m, in the multi-species system. One can describe the electric potential Φ in terms of concentration of the charged species, as expressed in Liu et al. (2011) and Steefel and Maher (2009):

$$\nabla\Phi = -\frac{\sum_{i=1}^{m} z_i D_i \nabla c_i}{\frac{F}{RT} \sum_{i=1}^{m} z_i^2 D_i c_i} \tag{1.14}$$

the electrochemical flux can be expressed, when combining equations 1.10 and 1.14, as:

$$J_{elec} = -\sum_{k=1}^{m} D_{ik} \nabla c_i \tag{1.15}$$

where D_{ik} is the cross diffusion term and is defined as:

$$D_{ik} = D_i \delta_{ik} - \frac{z_i z_k D_i D_k c_i}{\sum_{k=1}^{m} z_k^2 D_k c_k} \tag{1.16}$$

where δ is the Kronecker symbol (1 if $i = j$ or 0 otherwise). Combining J_{adv}, J_{dif} and J_{elec}, we obtain the total flux $J_{i_{tot}}$ for species i:

$$J_{i_{tot}} = J_{i_{adv}} + J_{i_{dif}} + J_{i_{elec}} = \theta V c_i - D_{e_i}(\frac{\partial c_i}{\partial t} + \frac{c_i F z_i}{RT} \nabla\Phi) \tag{1.17}$$

As the dispersive flux term in equation 1.3 is analogous to the diffusive flux term in equation 1.5, for the sake of simplification, it is considered as an additional diffusion flux and combined into D_e in equation 1.17.

Finally, the mass transport equation for each i_{th} transported aqueous species, based on Bear and Bachmat (1990), is in a general form:

$$\frac{\partial c_i}{\partial t} = -\nabla J_{i_{tot}} + Q_i \tag{1.18}$$

where Q_i is the source/sink term. Combining the equations 1.11, 1.17, and 1.18 gives the Poisson-Nernst-Planck (PNP) system of coupled equations, which is necessary to describe the transport of species in the presence of an electric potential. Such an electric potential can be generated, for example, by minerals with charged surfaces or by the movement of charged species.

Transport solver

The reactive transport framework presented in this thesis focuses on continuum scale models. It describes the transport of species in an aqueous phase for a constant temperature in a fully saturated medium and a non-compressible solvent, i.e., the volume of the solution does not change in relation to composition, pressure, or other processes. Such conditions can be assumed, for example, in cases of diffusion-dominated transport in clay and cementitious material, and are properly defined and described in chapters 2, 3, and 4. Pore-scale properties are averaged or upscaled to continuum scale using the notion of a representative elementary volume, known as REV (Kanit et al., 2003).

The transport equations are solved with the FEniCS library (Alnaes et al., 2015), a collection of numerical tools for solving partial differential equations (PDE) based on Finite Element (FE) numerical methods. In order to solve a PDE in FEniCS, it is necessary to define the PDE in its weak formulation and then to describe it using the Unified Form Language (UFL) (Alnaes et al., 2012), which is part of the FEniCS framework. The weak formulation, which is numerically easier to solve, is obtained by multiplying the regular PDE with a smooth, continuous, and differentiable test function (v) and integrating the product over the simulation domain. For example, the corresponding weak formulation of equation 1.11 is described as follows:

$$\int_{\Omega} \nabla\Phi \cdot \nabla v_{\Phi} dx - \int_{\Omega} \frac{F}{\epsilon} (\sum_{i=1}^{m} z_i c_i + w) \cdot v_{\Phi} dx = 0 \tag{1.19}$$

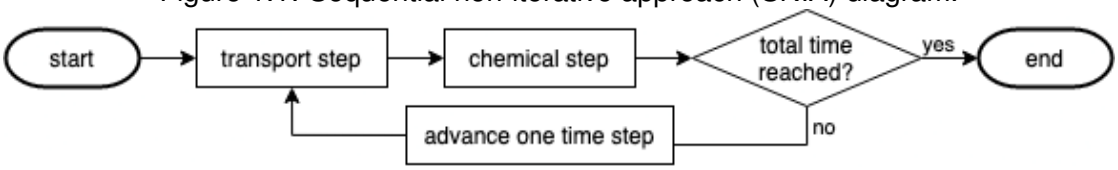
Full access to the PDE level implemented in FEniCS provides the flexibility necessary for the development of a general-purpose reactive transport framework. For example, it is straightforward and requires little effort to change the numerical method and to add new physical processes, as well as source/sink terms.

FEniCS provides HPC portability via PETSc (Balay et al., 2021a,b; Dalcin et al., 2011), an open-source parallel numerical software for solving PDEs, and offers spatial discretization in 1D, 2D, and 3D. It also enables the use of meshes generated by third-party software, like GMESH (Geuzaine and Remacle, 2009). FEniCS supports different boundary conditions, predefined solvers (e.g., minimal residual method, generalized minimal residual method, PETSc built-in LU solver, and biconjugate gradient stabilized method), and preconditioners (e.g., incomplete LU factorization and PETSc algebraic multigrid). Furthermore, FEniCS makes it possible to easily adopt different families of finite elements by directly changing the weak formulation (e.g., arbitrary order continuous and discontinuous Lagrange). FEniCS is available in both C++ and Python.

1.3.2 Chemical solver

For the calculation of the chemical equilibrium condition, GEMS3K (Kulik et al., 2013) and Reaktoro (Leal et al., 2017) were used. Both codes use the Gibbs Energy Minimization (GEM) algorithm in order to calculate the equilibrium condition and to enable the description of non-ideal and ideal multicomponent solution phases, ion exchange, multiple activity models, and several equations of states for multicomponent fluid systems. The GEM algorithm problem is based on the

Figure 1.1: Sequential non-iterative approach (SNIA) diagram.



total mole amounts of chemical elements, pressure, and temperature. It consists of finding, based on the thermodynamic data, the equilibrium phase assembly, equilibrium speciation, component activities, saturation indexes, and other properties by minimizing the Gibbs energy of the system of interest. Alternatively, a very common approach among popular chemical solvers, such as Phreeqc (Parkhurst and Appelo, 2013) and Orchestra (Meeussen, 2003), is the law of mass action (LMA). The LMA calculates the equilibrium condition of the chemical system by solving a system of mass balance equations (Reed, 1982). The extended LMA (xLMA), presented in (Leal et al., 2016), improves the traditional LMA by adopting concepts of the GEM algorithm in order to determine the stable phases and improve performance.

The large computational resources needed for such chemical equilibrations, as discussed in Kulik et al. (2013) and Leal et al. (2020), may result in these calculations taking much longer than the solution of the transport equations (as described in section 1.3.1). This problem is known to represent the bottleneck in reactive transport simulations. Therefore, an efficient reactive transport framework requires a fast, robust, and accurate chemical solver (Leal et al., 2016).

1.3.3 Coupling of chemical and transport processes

The sequential non-iterative approach (SNIA) (Walter et al., 1994; Carrayrou et al., 2004; Hundsdorfer and Verwer, 1995; Appelo et al., 1998) is used for coupling transport and chemical solvers. For each time step and node, the transport equations are solved, the changed concentrations are updated, and a chemical equilibrium state is calculated. This process is repeated until the simulation time reaches the total time. At each time step iteration, a chemistry step follows a transport step. Figure 1.1 presents the SNIA diagram.

When implementing the SNIA approach, the implicit time step scheme is adopted, which is mathematically described as:

$$\frac{\partial c_i}{\partial t} \approx \frac{c_i^n - c_i^{n-1}}{\Delta t} \tag{1.20}$$

where Δt is the time step (s), c_i and c_{n-1} are, respectively, the concentration of species i at times n and $n-1$. The Δt is adjusted according to the Von Neumann stability stability criteria in order to avoid negative concentrations (Charney et al., 1950).

One way to accelerate the chemical step is to implement a multi-threading approach, as described in Kosakowski and Watanabe (2014). Multi-threading can be implemented in the developed reactive transport framework with the help of the python *threading* library. When using multi-threading, multiple instances of the geochemical solver calculate the chemical equilibrations

in parallel. This technique was used in 3 for the reactive transport in Opalinus Clay. Another approach for accelerating the chemical step is described in subsection 1.3.4 and was used in 4.

1.3.4 Machine learning as chemical solver

The usage of modern machine learning algorithms (Hinton et al., 2006) to mimic the role of traditional chemical solvers in reactive transport modelling makes it possible to address complex problems in more detail, while improving performance without sacrificing accuracy (Tu, 1996). Tensorflow (Abadi et al., 2015) and Jupyter notebooks (Kluyver et al., 2016) were used to create an infrastructure that makes it possible to investigate the impact of hyperparameters (parameters that control the network's learning process), as well as to create deep neural networks (DNN) (Schmidhuber, 2015) that emulate the chemical solver's speciation. Similar to a conventional chemical solver, the neural network's input is the independent components and the output is the aqueous speciation and mole amounts of mineral phases. Tensorflow offers several building blocks for different machine learning and numerical computation algorithms. In addition to deep learning, which represents its most popular algorithms, it also offers K-means clustering, Random Forests, Support Vector Machines, Gaussian mixture model clustering, and linear/logistic regression. It runs on a traditional central processing unit (CPU), but also on GPU computational architectures with no code changes required. The usage of a neural network requires a preliminary training phase, which can be highly computing-intensive and which demands a careful definition and analysis of hyperparameters. Moreover, a large data set is needed to provide the information necessary during training, and some of the information found in the traditional chemical solver is not present (e.g., thermodynamic properties). Nevertheless, thanks to the development conducted in this work, one has all the necessary tools bundled into a unique framework that investigates several machine learning techniques. It starts from the geochemical solver, which produces the training dataset, uses machine learning to train and test a DNN, and then goes on to run the complete reactive transport problem with both the geochemical solver and the DNN. This automated procedure, with all its ingredients, allows strategies to be developed for testing and validating machine learning algorithms in an unprecedented way.

1.4 Deep geological disposal of radioactive waste

Nuclear waste is produced in different sectors of the modern economy, for example, by nuclear power plants, medicine, industry, and research facilities. Nuclear waste and its disposal can potentially threaten our environment and humankind due to its high radiotoxicity. Therefore, the development and implementation of a safe disposal strategy is of great importance and represents an ethical commitment on the part of modern society towards future generations (NEA, 1995).

The concept of the deep geological disposal of radioactive waste, accepted by many countries as the most reliable and sustainable solution, relies on a combination of geological and engineered barriers that retard the transport of radionuclides (NAGRA, 2002). Depending on its half-life, all radionuclides will decay to stable nuclei given enough time. Some radionuclides have a half-life on the order of seconds, while that of others lasts millions of years. Depending on its activity and inventory, the waste is classified into one of two groups: 1) Low and Intermediate-Level Waste

(L/ILW); 2) High-Level Waste (HLW) (NAGRA, 2008). L/ILW consists of waste from radioactive operations (e.g., activated components of the reactor vessel, protective medical clothing, and laboratory tools). HLW is highly radioactive and primarily composed of spent fuel from nuclear power plants. It may need 10^5 to 10^6 years to decay to the level of natural uranium ore. Therefore, a nuclear waste disposal concept must be capable of protecting the environment for extended periods on a geological timescale.

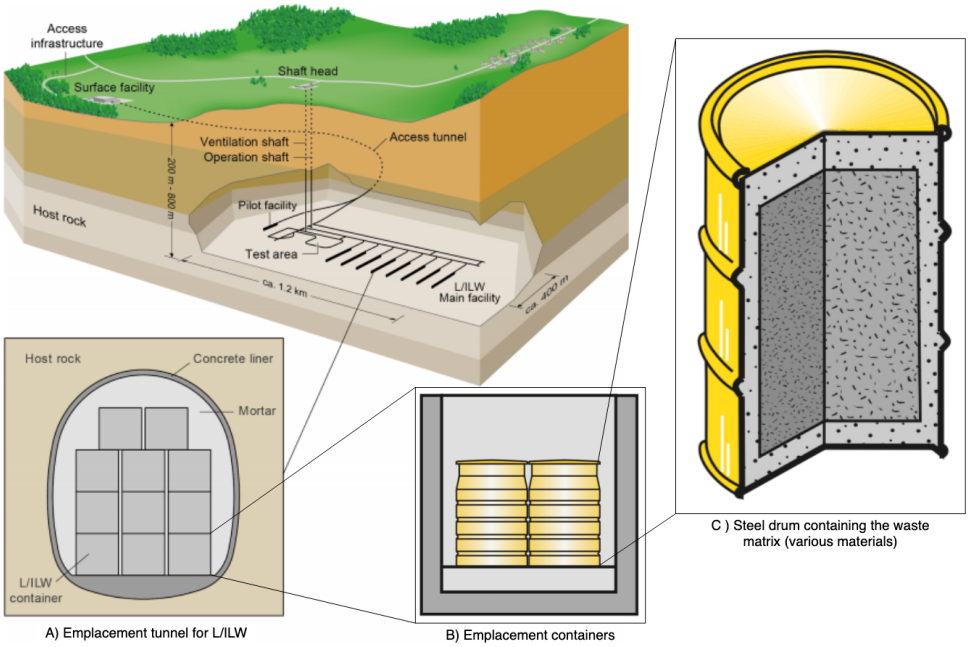
The arguments for the safety case of deep geological disposal are based on four main principles (NAGRA, 2002; Mazurek et al., 2008): site selection and characterization (NAGRA, 2008; Nussbaum et al., 2011); modelling (Kulik et al., 2013; Marty et al., 2015); rock laboratories (Kickmaier, 2002); and natural analogues (Alexander et al., 2015). In order to gain a deep and comprehensive understanding of long-term repository evolution, these topics are the subject of intensive scientific research and analysis (Bossart et al., 2017). Among other things, the safety-relevant characteristics of a repository are addressed: for example, how will the steel endure over the years and what is the impact of anoxic corrosion of the steel? (Saheb et al., 2010; Necib et al., 2017; Mancini et al., 2020) How can mineral dissolution and precipitation compromise the long-term stability of the repository? (Poonoosamy et al., 2015) What is the influence and solubility of radionuclides and their interactions with cementitious materials? (Kulik and Kersten, 2001) Which host rock offers natural advantages for radionuclides retention (e.g., low permeability, swelling and sorption mechanisms)? (Pape et al., 1999; Skipper et al., 2006; Van Loon et al., 2007; Bradbury and Baeyens, 2009; Wigger and Van Loon, 2018; Al Reda et al., 2020; Poonoosamy et al., 2020) How does microbial activity inside the host rock affect the overall physical and chemical processes? (Lovley and Chapelle, 1995; Stroes-Gascoyne et al., 2007a)

Opalinus Clay is a potential host rock for waste disposal in Switzerland. It is an argillaceous claystone with low permeability and high sorption capacity. Figure 1.2 presents a possible layout of the Swiss L/ILW concept for the deep geological repository, which includes surface facilities, tunnels for access, ventilation and operation shafts, a pilot facility, a test area, and the L/ILW facilities. In the L/ILW concept, for example, the wastes are conditioned into the waste matrix and placed inside drums (1.2C). These drums are then inserted into a concrete container (1.2B) and placed inside the tunnels (1.2A).

Laboratory experiments and underground rock laboratories (URL), such as the Mont Terri underground rock laboratory (Thury and Bossart, 1999), are of great importance for the hydrogeological, geochemical, and geotechnical characterization of the materials, as well as for understanding processes, testing model concepts, and estimating parameters (Pearson et al., 2003; Van Loon et al., 2004; Glaus et al., 2015; Blechschmidt and Vomvoris, 2017; Van Loon et al., 2018; Mäder et al., 2017; Jenni et al., 2014; Stroes-Gascoyne et al., 2007b). The safety assessment imposed by regulatory requirements can only be performed once the physical and chemical processes that change the repository system over periods of hundreds of thousands of years have been thoroughly understood (De Windt et al., 2004; Gaucher and Blanc, 2006; Trotignon et al., 2007; Koroleva et al., 2011; Kosakowski and Berner, 2013; Mazurek and de Haller, 2017).

Reactive transport models represent the most reliable tool for analyzing and predicting the evolution of deep geological nuclear waste repositories. Reactive transport models increase confidence that the repositories satisfy the safety requirements, effectively retaining the radionuclides

Figure 1.2: Multibarrier concept for the Swiss geological L/ILW repository (combined from Kosakowski et al. (2014) and NAGRA (2002))



and thus the harmful radioactivity from nuclear waste.

1.5 Outline of the thesis

This thesis consists of three chapters derived from publications in peer-reviewed journals. Chapter 2 presents the governing equations, a description of the developed innovative reactive transport modelling approach, and the benchmarks used to validate it. This chapter has been published as *Damiani, L. H., Kosakowski, G., Glaus, M. A., & Churakov, S. V. (2020). A framework for reactive transport modelling using FEniCS–Reaktoro: governing equations and benchmarking results. Computational Geosciences, 24(3), 1071–1085. https://doi.org/10.1007/s10596-019-09919-3*

In chapter 3, we apply a new approach to investigate the interaction of gases in Opalinus Clay at the Hydrogen Experiment (HT) at the Mont Terri rock laboratory (Vinsot et al., 2017). This chapter has been published in Environmental Geotechnics, for a special themed issue of Physical-Chemical Coupling, as *Damiani, L. H., Kosakowski, G., Vinsot, A., & Churakov, S. V. (2021). Hydrogen gas transfer between a borehole and claystone: Experiment and geochemical model https://doi:10.1680/jenge.21.00061.*

Chapter 4 presents a machine learning assisted framework that uses deep neural networks instead of the traditional chemical solver in reactive transport modelling. The ML framework has enormous potential to tackle much more complex problems thanks to its computational advantages. We discuss the advantages and limitations of the geochemical ML framework and

present the application results inspired by a popular reactive transport benchmark. This chapter is currently being prepared for submission to a peer-reviewed journal as *Damiani, L. H., Haller, R. L. I., Kosakowski, G., Miron, G. D., Kulik, D. A., Churakov, S. V. & Prasianakis, N. (in preparation) Machine learning assisted geochemical reactive transport modelling: a unified modular computational framework*

In Chapter 5, the summary and potential directions for future research are discussed.

Chapter 2

A framework for reactive transport modeling using FEniCS–Reaktoro: governing equations and benchmarking results

https://doi.org/10.1007/s10596-019-09919-3

Computational Geosciences (2020) 24:1071–1085
https://doi.org/10.1007/s10596-019-09919-3

ORIGINAL PAPER

A framework for reactive transport modeling using FEniCS–Reaktoro: governing equations and benchmarking results

Leonardo Hax Damiani[1,2] · Georg Kosakowski[1] · Martin A. Glaus[1] · Sergey V. Churakov[1,2]

Received: 19 February 2019 / Accepted: 8 November 2019 / Published online: 21 January 2020

Abstract
Reactive transport codes are widely applied in geoscience to predict or reconstruct spatial and temporal evolution of geochemical systems. To provide an accurate description of natural systems at different spatial and temporal scales, the reactive transport code has to deal with coupling of different physical and chemical phenomena. Many reactive transport codes have been developed in the past and each of these codes has specific strengths and limitations. Here, we present a new versatile reactive transport framework based on the FEniCS equations solver and the chemical solver Reaktoro. This development was motivated by the need for an advanced open-source tool allowing user-friendly modeling environment and, at the same time, full control over the numerical methods. Unlike most of the currently available codes, the developed FEniCS–Reaktoro framework offers full flexibility in setting up the reactive transport simulations of arbitrary complexity in terms of process couplings, simulation domain geometry and the boundary conditions applied. The simulations are setup using a simple high-level scripting language intuitively linked to the equation based model definition without the need of advanced programming skills. The chemical solver Reaktoro allows thermodynamic modeling of multicomponent multiphase system with several fluids and solid phases, including highly non-ideal solid solutions. The coupling of transport and chemistry is implemented using the sequential non-iterative approach (SNIA) in which the transport of the aqueous components and the chemical reactions are solved in two consequent steps. The flexibility and results of the FEniCS–Reaktoro framework are demonstrated against several widely accepted reactive transport benchmarks.

Keywords Reactive transport · Electrochemical transport · Multicomponent diffusion · Finite element method · Porous media · Gibbs energy minimization · Operator splitting approach

1 Introduction

Reactive transport modeling is widely applied to study the temporal and special evolution of geochemical conditions in natural system in various geoscientific fields [1]. To provide an accurate description of natural systems at different spatial and temporal scales, the reactive transport code has to deal with the coupling of different physical and chemical phenomena, such as heat transfer [2], advective fluid transport in porous and fractured media [3], diffusive and electrochemical migration of chemical species [4], electromigration [5, 6], mechanical stress [7], and heterogeneous chemical reactions [8, 9].

✉ Leonardo Hax Damiani
Leonardo.Hax@psi.ch

Extended author information available on the last page of the article.

To allow for full control in the system description, a general purpose code should thus be:

- Able to solve chemical systems of realistic complexity
- Able to model physical processes for arbitrary complex geometries
- Numerically accurate and robust
- Easily scalable in terms of process coupling and system complexity
- Compatible with HPC architecture
- Well documented and preferably available as open-source

Table 1 compares the capabilities of popular reactive transport codes currently available to the scientific community. Each code listed in Table 1 has specific advantages and limitations related to the specific implementation of the modeling approach. These are analyzed in detail in the following sections.

Table 1 Reactive transport codes comparison

Code	Fully open-source	Numerical method[1]	Dimensions	Mass transport PDE[2]	Electrochemical migration	Chemical method[3]	Compatible with HPC	Numerical approach[4]
CrunchFlow [10]	Yes	FV	1,2,3D	BI	Yes	LMA	No	GI,OS
Comsol-Phreeqc iCP [11]	No	FE	1,2,3D	UD	No	LMA	Yes	OS
Comsol-Phreeqc [12]	No	FE	1,2,3D	UD	Yes	LMA	No	OS
Comsol-GEMS [13]	No	FE	1,2,3D	UD	Yes	GEM	Yes	OS
CORE^{2D} [14]	No	FE	1,2D	BI	No	LMA	No	OS
CSMP++GEM [15]	No	FE, FV	1,2,3D	BI	No	GEM	Yes	OS
DuCOM-Phreeqc [16]	No	FE	1D	BI	Yes	LMA	No	OS
Frachem [2]	No	FE, FV	1,2D	BI	No	LMA	No	OS
GeoSysBRNS [9]	No	FE	1,2,3D	BI	No	LMA	Yes	OS
HP1/HPx [17]	No	FE	1,2,3D	BI	No	LMA	Yes	OS
Matlab-IPhreeqc [6]	No	FE,FV	1,2D	BI	Yes	LMA	No	OS
MIN3P [18]	No	FV	1,2,3D	BI	Yes	LMA	No	GI
The Geochemist's Workbench [19]	No	FD	1,2D	BI	No	LMA	Yes	OS
OGS-Chemapp [20]	No	FE	1,2,3D	BI	No	GEM	No	OS
OGS-GEM [21]	Yes	FE	1,2,3D	BI	No	GEM	Yes	OS
OGS-IPhreeqc [22]	Yes	FE	1,2,3D	BI	No	LMA	Yes	OS
Orchestra [23]	No	MC	1,2D	UD	Yes	LMA	Yes	OS
Pflotran [24]	Yes	FV	1,2,3D	BI	No	LMA	Yes	GI, OS
Phreeqc [25]	Yes	MC	1D	BI	Yes	LMA	No	OS
Proost [26]	No	FE	1,2,3D	BI	No	LMA	No	OS
ReactMiCP [27]	Yes	FE	1D	BI	No	LMA	Yes	OS
ReacTran [28]	Yes	FD	1,2,3D	UD	No	RR	Yes	GI
Toughreact [29]	No	FV	1,2,3D	BI	No	LMA	Yes	OS
TReacLab[5] [30]	No	FD	1,2D	UD	Yes	LMA	No	OS

[1]FE denotes finite elements, FV denotes finite volumes, FD denotes finite differences, and MC denotes mixing cell

[2]BI denotes built-in and UD denotes user-defined

[3]LMA denotes law of mass action, GEM denotes Gibbs energy minimization and RR denotes reaction rates

[4]GI denotes global implicit and OS denotes operator splitting

[5]Since the purpose of TReacLab is to provide a generic operator splitting interface, its definitions vary according to the transport and chemical solvers used

The selection of the simulation tool should be based on a careful evaluation of code functionality to assure that the simulation tools meet the user's requirements. In addition to a broad range of physical phenomena, a number of recently developed codes also support parallelism and provide HPC compatibility [31–33]. The numerical performance and scalability of the codes are essential criteria when realistic 2D and 3D natural systems are considered.

Chemical solvers based on law of mass action (LMA) (e.g., equilibrium constant of chemical reactions) are most commonly used to calculate equilibrium speciation in reactive transport codes due to the simplicity of the algorithm [34]. The LMA solvers have limitations when multicomponent phases (e.g., solid solutions, non-ideal liquid, and gaseous phases) need to be considered. The Gibbs energy minimization (GEM) algorithm [8] is the method of choice for the simulations of multiphase systems and is based on the standard chemical potentials. Both LMA and GEMS have pros and cons when practical applications are considered. Both approaches are based on different algorithms and assumptions but are equivalent from the theoretical point of view and must, therefore, provide the same result. Recently, Leal et al. [35] developed an innovative approach referred to as xLMA method which combines the basic algorithms used in the LMA and GEM approach resulting in substantial improvement of the stability and performance of the chemical solver.

Depending on the code design, the mass transport equations can be built-in into the code or defined by the user via an external interface or by a high-level scripting language. Most commonly, the partial differential equations (PDE) are hard-coded within the core routines and not accessible or changeable by the user, as in [25]. In contrast, in a more flexible approach, some codes provide a scripting or graphical interface that allows the user to define the mass transport equation, as in [23, 36].

Coupling of reactions and transport can be implemented using the operator splitting (OS) approach [37, 38], or the global implicit (GI) method [39, 40]. In practice, both methods become intractable for the complex system due to the complexity of numerical implementation and increasing computer memory requirements. The pros and cons of each approach have been discussed extensively in the literature, as in [41–43].

The characteristic time scale of physical and chemical phenomena or spatial heterogeneities in natural systems can span over several orders of magnitude [44]. Numerical simulations of such systems are very challenging because the appropriate integration time and the spatial discretization of the domain are typically limited by the fastest small-scale phenomena. Numerical codes typically hit the performance limits when pore-scale phenomena [45] controls the processes at the field scale [46, 47]. Particularly challenging are the simulations of systems with strong chemical gradients such as nuclear waste disposal sites or geochemical installations [48, 49].

Transport equations can be solved by different numerical methods: finite elements (FE) [50–52], finite differences (FD), or finite volumes (FV) [53]. FE is well suited for dealing with complex geometries and provides extra stability by using small elements where the function is changing rapidly and bigger elements elsewhere. The FE method has disadvantages as far as the mass and the energy conservation are concerned. The FV method ensures local mass and energy conservation but requires more efforts to describe irregular geometries. For the sake of completeness, it is worth mentioning an alternative approach based on lattice Boltzmann method [54]. The particular strength of the LB method is the exceptional stability of the LB equations in a wide range of hydrodynamic regimes and a straightforward implementation of a parallelization algorithm.

The ability of a code to allow the description of materials properties by different constitutive relationships, i.e., porosity [55], tortuosity [56], diffusivity [57], and boundary conditions [58], is critical for successful implementation of system setup. The possibility to allow user-controlled parametrization and consideration of more specific processes and interaction, e.g., anisotropic diffusion [59] or porosity–permeability relation [60], are highly desirable features which are not commonly available in most of the codes.

The numerical performance of the code and the compatibility with high-performance computing (HPC), as in [11, 32, 61], deliver better performance when treating a large system of equations necessary for describing the physical and chemical phenomena. The numerical approaches take advantage of the usage of HPC during the multithreaded node-oriented chemical calculation, in the OS approach, or by an efficient solution of the large equation system, in the GI approach. Since the speed of the clock is stagnated at 2–3 GHz, the further increase of code performance needs to be improved by taking advantage of multicore core architecture CPUs and massively parallel HPC facilities. The multi-threading is a simple way of improving the code performance but is limited essentially to the shared memory architecture. Most of the modern HPC facilities rely on distributed memory architecture. Use of such a resources prerequisite decomposition of the simulated domain and explicit data exchange between subdomains.

The usage of commercial closed source (or proprietary) codes, for example, [19, 36], often involves costs for licensing and subscriptions that easily exceed the costs of research projects and make them accessible only to a limited number of researchers. Commercial codes' usability is based on a predefined application programming interface (API) which typically limits their flexibility and transparency. Nonetheless, for the most part, they offer good quality

documentation and support. The open-source codes, like [21, 28], are typically developed at research institutes and universities as part of scientific projects. The development is often discontinued and many times the projects are short on documentation and reduced support. Most of the open-source codes have a potential to benefit from collaborative development and intensive benchmarking by users, as in the SeS Bench (Subsurface Environmental Simulation benchmark) [62]. Exemplary, the potential of the NumFOCUS https://numfocus.org/, which has been recognized by research and industry communities in multiple projects [63–66].

2 FEniCS–Reaktoro framework

In this work, we present a new framework for reactive transport simulations called FEniCS–Reaktoro, able to deal with complex multicomponent heterogeneous systems. It follows an object-oriented (OO) approach and uses Python scripting interface for the system setup and coupling of the PDE solver FEniCS [67] with the chemical solver Reaktoro [68]. It allows fast and flexible system setup with a moderate requirement of programming skills. The framework is suited for systems with different spatial and temporal scales such as modeling of a short-term laboratory scale and medium-term field-scale reactive transport experiments, or large-scale reservoir modeling. In the following section, we present a detailed description of the tools used, design, architecture, and capabilities of the FEniCS–Reaktoro framework. We also describe the equations used in the benchmarks as well as the description of the model setup. Numerical benchmarks and applications are presented to demonstrate the versatility of the FEniCS–Reaktoro framework. The combination of such features is not currently available in another open-source reactive transport code, which makes FEniCS–Reaktoro a unique and thriving modeling platform.

2.1 FEniCS

FEniCS is a popular open-source computing framework for solving PDE with several FE-based families. The significant advantage of using FeniCS, compared to other existing software packages, is the flexibility of an automated and efficient solver for the differential equations. The user has essentially to define the equations, and FEniCS takes care of numeric methods necessary to solve the mathematical problem. With FEniCS, the PDE is expressed in the weak form in a math-like human-readable way called Unified Form Language (UFL). For more details on the weak formulation and UFL, we refer to [69]. FEniCS provide the user with full control on the PDE level and allow a quick change of the FE family, equation definition (as well as addition of extra terms into the equation), equation solver, and use other geometries and dimensions (1D, 2D, or 3D). FEniCS offers different inbuilt boundary conditions such as Dirichlet, Neumann, Robin, several solvers (i.e., minimal residual method, generalized minimal residual method, PETSc library built-in LU solver, and biconjugate gradient stabilized method), and preconditioners (i.e., incomplete LU factorization, and PETSc algebraic multigrid). FEniCS provides portability to HPC [70] without rewriting existing codes based on the PETSc [71]. The documentation and examples are available at https://www.fenicsproject.com and it is available in C++ and Python.

2.2 Reaktoro

Reaktoro is an open-source software for modeling chemically reactive systems [68]. It implements efficient numerical methods for solving chemical equilibrium and kinetics for multiphase systems that are particularly suitable for demanding applications such as reactive transport modeling. Reaktoro takes advantage of the GEM method to define the stable phases in equilibrium and is able to import files contaiing thermodynamic systems from GEM–Selektor [8] and Phreeqc [25].

Reaktoro can model non-ideal aqueous solutions using Davies, Debye–Huckel, Pitzer, and HKF activity models. For gases, Reaktoro implements cubic equations of state such as Peng–Robinson, Redlich–Kwong, Soave–Redlich–Kwong, and some models specific for certain mixtures, such as [72] for $H_2O - CO_2$ and [73] for $H_2O - CO_2 - CH_4$. Reaktoro can also model the non-ideal thermodynamic behavior of solid solutions using, for example, Redlich–Kister solid solution model.

Reaktoro also offers support for the CEMDATA, which is a cement specific database [74], the PSI-Nagra database [75], the SUPCRT92 [76], and the Phreeqc [25] databases. Reaktoro allows the flexibility to model high temperatures and pressures (up to 1000 °C and 5000 bar).

Reaktoro is implemented in C++ and offers a Python interface. Its source code, documentation, and examples are available at https://www.reaktoro.com.

3 Mathematical setup

The reference model discussed in the following section describes the single-phase transport of aqueous species in fully saturated porous media with a realistic description of the chemical system. The model includes the following:

– Advection, diffusion, and electrochemical migration of aqueous species in an incompressible flow (aqueous

phase). Including the transport across pores confined by charged mineral surfaces [77, 78] by the Poisson–Nernst–Planck equation
– The multiphase chemical systems, including ternary or higher order non-ideal solid solutions, in which precipitation and dissolution of minerals can alter material properties [79]

The description of the reactive mass transport in a general form, based in [80], is as follows:

$$\frac{\partial c_i}{\partial t} = -\nabla\big(V c_i - D_i(\nabla c_i + \frac{z_i F c_i}{\mathrm{R\,T}}\nabla\Phi)\big) + Q_i \tag{1}$$

and

$$\frac{\partial c_i}{\partial t} = \Gamma(c_1 ... c_m) \tag{2}$$

where c_i is the volumetric concentration (mol/m^3) of the ith transported species, t is time (s), V is the pore velocity (m/s), D_i is the diffusion coefficient (m^2/s), z_i is the charge (C), F is the Faraday constant (C/mol), Φ is the electrical potential (V), R is the gas constant (J/mol/K), T is the temperature (K), Q_i is the source/sink term, $\Gamma(c_1 ... c_m)$ is the chemical change rate term (mol/s), and m is the total number of species in the multi-species system. In the right side of Eq. 1, the first term refers to the advective transport, the second term to the aqueous diffusion, the third to the electrochemical migration, and the last to the source/sink term. The Poisson equation that describes Φ, as in [81], is as follows:

$$\nabla^2\Phi + \frac{F}{\epsilon}(\sum_{i=1}^{m} z_i c_i + w) = 0 \tag{3}$$

where ϵ is the dielectric constant (or relative permitivity) of the material (C/V/m) and w is the charge density of the medium (C/m). The combination of Eqs. 1 and 3 is known as the Poisson–Nernst–Planck (PNP) equations. Such equation system also needs a definition of initial and boundary values for concentrations and electric potential, which are defined according each model's definition. For solving transient transport with FE methods (as implemented in FEniCS), we discretize the above equation system in time and transform it into the so-called weak form. We choose the OS approach with an implicit time discretization scheme due to its stability when coupling transport and chemical solvers using OS [82] in combination with small time step length. The implicit time step scheme, as in [10], is as follows:

$$\frac{\partial c_i}{\partial t} \approx \frac{c^n - c^{n-1}}{\Delta t} \tag{4}$$

where c^n is the current unknown concentration, c^{n-1} is the known concentration from previous time step, and Δt is the time step size. In the OS approach, the transport of the aqueous components and the chemical reactions are solved in two sequential steps. It allows maximal flexibility for coupling the transport and chemical solvers but requires a time step size that respects the Courant–Friedrichs–Lewy (CFL) condition [83], the von Neumann criteria [84], and the rate-limiting step (RLS) [85] for kinetically controlled reactions. The time step size needs to account correctly the transport rates as well as the equilibrium (fast) and kinetic (slow) reaction rates, and properly address the different orders of magnitude between them (from microseconds up to several days) [86]. Thus, it could be expected thousands or millions of time step iterations from the initial condition until some time of interest according to each model.

The weak formulation is achieved by multiplying the regular PDE form with a test function (v_i) and integrating over the domain. Such test function needs to be smooth, continuous, and differentiable inside the domain. Therefore, the solution of the system becomes numerically easier, if compared to solving the system based on its regular PDE form, also known as strong form.

In this work, all the weak formulations presented are defined for the Lagrange family of FE basis with linear interpolation, and we assume either a constant pore velocity or zero, for cases without advective transport. Alternatively, one could use the discontinuous Galerkin (DG) family of elements [87, 88], which can be done in FEniCS by describing the weak formulation in the appropriate DG form, as in [89, 90].

3.1 Transport solver

Table 2 summarizes the equations implemented and presented in this work:

3.1.1 Poisson–Nernst–Planck equation

In a strongly diffusion dominated transport materials, e.g., low permeability bentonite or Opalinus clay [91], it is possible to neglect the advective term. Assuming the source/sink term zero and combining (1) and 4, we achieve the following:

$$\frac{c^n - c^{n-1}}{\Delta t} = -\nabla\big(- D_i(\nabla c_i + \frac{z_i F c_i}{\mathrm{R\,T}}\nabla\Phi)\big) \tag{5}$$

Table 2 Description of the equations implemented for the benchmarks presented in this work

Transport equation description	Eq.
Poisson–Nernst–Planck (PNP)	Eqs. 7 and 6
Nernst–Planck equations (assuming electroneutrality and null current conditions)	Eq. 13
Advection–diffusion equation	Eq. 15

after algebraic manipulation and the introduction of the porosity term (θ), we present the weak formulation, respectively, of Eqs. 3 and 5:

$$\int_\Omega \nabla\Phi \cdot \nabla v_\Phi \mathrm{d}x - \int_\Omega \frac{F}{\epsilon}(\sum_{i=1}^{m} z_i c_i + w) \cdot v_\Phi \mathrm{d}x = 0 \tag{6}$$

$$\int_\Omega \theta c_i^n v_i \mathrm{d}x = \int_\Omega \theta c_i^{n-1} v_i - \theta \Delta t D_i (\nabla c_i^n + \frac{z_i F c_i^n}{\mathrm{R\,T}} \nabla\Phi) \cdot \nabla v_i \mathrm{d}x \tag{7}$$

where v_Φ is the test function associated with the electric potential.

3.1.2 Nernst–Planck equation (concentration-based description of the electric potential)

Assuming that the electroneutrality condition is as follows:

$$\sum_{i=1}^{m} z_i c_i + w = 0 \tag{8}$$

and the null current condition:

$$F \sum_{i=1}^{m} z_i J_i = 0 \tag{9}$$

are satisfied, one can describe Φ in terms of concentration of species, as expressed in [92]:

$$\nabla\Phi = -\frac{\sum_{i=1}^{m} z_i D_i \nabla c_i}{\frac{F}{\mathrm{R\,T}} \sum_{i=1}^{m} z_i^2 D_i c_i} \tag{10}$$

Combining (7) and (10):

$$\frac{\partial c_i}{\partial t} = -\sum_{k=1}^{m} D_{\mathrm{ik}} \nabla c_i \tag{11}$$

where D_{ik} is the cross-diffusion term and is defined as follows:

$$D_{\mathrm{ik}} = D_i \delta_{\mathrm{ik}} - \frac{z_i z_k D_i D_k c_i}{\sum_{k=1}^{m} z_k^2 D_k c_k} \tag{12}$$

where δ is the Kronecker symbol (1 if $i = j$ or 0 otherwise). Finally, we introduce the porosity term and present the weak formulation of Eq. 11:

$$\int_\Omega \theta c_i^n v_i \mathrm{d}x = \int_\Omega \left(\theta c_i^{n-1} v_i - \theta \Delta t \sum_{k=1}^{m} D_{ik} \nabla c_k^n\right) \cdot \nabla v_i \mathrm{d}x \tag{13}$$

3.1.3 Advection–diffusion equation

Equation 5 becomes the known advection–diffusion equation, when the electrochemical migration and the source/sink term are neglected, as in [92]:

$$\frac{\partial c_i}{\partial t} = -\nabla \cdot \left(V c_i - D_i (\nabla c_i)\right) \tag{14}$$

The weak formulation for a porous medium is then [93]:

$$\int_\Omega \theta c_i^n v_i \mathrm{d}x = \int_\Omega \left(\theta c_i^{n-1} - \theta \Delta t (V c_i - D_i \nabla c_i^n)\right) \cdot \nabla v_i \mathrm{d}x \tag{15}$$

3.2 Chemical solver

The thermodynamic solver calculates the chemical term $\Gamma(c_1 ... c_m)$. It describes the equilibrium and kinetic reactions that occur in the fully saturated multicomponent system with an arbitrary number of mineral and gases phases. The chemical equilibrium is calculated with the GEM approach at a given temperature, pressure, and bulk composition, with volume treated as a free parameter. Setup of the chemical system comprises the definition and thermodynamic data for possible solid/liquid/gas phases, temperature, pressure, and the total bulk elemental composition of the system to equilibrate. The newly equilibrated equilibrium state is then used to extract the updated species concentrations.

Reaktoro allows to calculate Lasaga-type kinetics rates [94]. If one needs to describe a special kinetically controlled rate, FEniCS–Reaktoro framework allows one to implement it into the reactive transport loop directly.

In all the benchmarks presented in the following section, the thermodynamic database set up uses the PSI–Nagra database [75].

3.3 Transport and chemical coupling

The chemical and the transport processes are coupled using the sequential non-iterative approach (SNIA) as conceptualized in Fig. 1.

Within each time step, the PDEs for transport of the aqueous components are solved first. The new composition of the chemical system, at each FE, node is passed to Reaktoro for the calculation of thermodynamic equilibria or kinetically controlled chemical reactions. After the chemical equilibration step, the concentrations for each transported species are passed back to each FE node for the next time step.

The transport PDEs are implemented using volumetric concentrations while the chemical solver uses absolute amounts of moles elements. The coupling between transport and chemical solver involves the conversion of

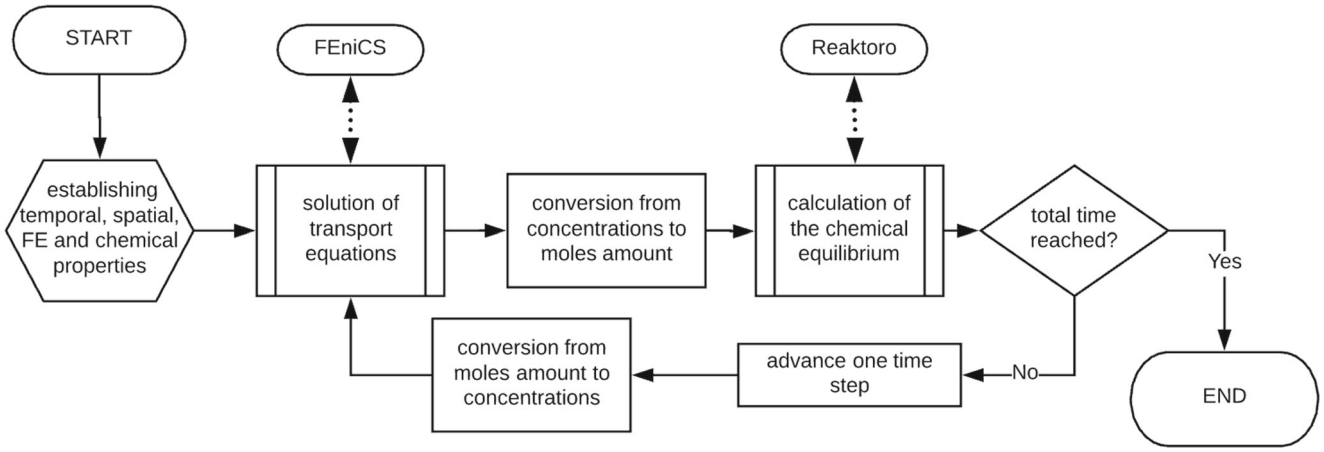


Fig. 1 FEniCS–Reaktoro operator splitting SNIA approach

the volumetric concentrations transported to total moles amounts needed for the chemical solver and backward.

In Appendix A, we present details of the main core class responsible for the transport chemical coupling as well as the UML class and sequence diagram [95].

4 Benchmarking results

In this section, we present four 1D benchmarks that illustrate the code capability to simulate: diffusion dominated flow and advection–diffusion flow, effects of electrochemical migration, mineral precipitation, and dissolution reactions. The results are compared with results from other codes and validated against experimental data. The first benchmark describes the cross-diffusion of ions due to electrochemical coupling. In the second benchmark, simulations of the electrochemical diffusion of ions through an uncharged membrane are compared with experimental data. The third benchmark presents the diffusion of ions through a charged membrane under an external electric field. The fourth benchmark deals with mineral precipitation and dissolution in porous media. Table 3 summarizes the physical processes and the equations used to describe them for each benchmark case.

4.1 Benchmark 1: Coupled electrochemical migration diffusion

This benchmark was developed by [96], inspired by [97]. It describes the multicomponent diffusion of HNO_3 from a pH = 4 solution to a pH = 6 solution, with a constant NaCl background concentration. The left side (at $x = 0$) has a fixed concentration, while on the right side (at $x = 1.0$) an open boundary condition is applied. There are no chemical reactions taking place in this simulation. Table 4 presents the details of the benchmark 1.

Our results show a perfect agreement with the results obtained from Phreeqc and CrunchFlow, as presented in Fig. 2.

4.2 Benchmark 2: coupled electrochemical diffusion through an uncharged membrane

Benchmark 2 represents a simple laboratory experiment similar to benchmark 1. The experiment determines the electrochemical flux of KCl between two reservoirs with an initially identical concentration of KCl. The electrochemical flux of KClis triggered by the HNO_3 concentration gradient imposed between the two reservoirs. The experimental set up consists of two adjacent reservoirs separated by a cellulose acetate membrane. Cellulose acetate represents an uncharged porous medium. An electroneutral and non-sorbing tracer, HTO (tritiated water), is used to measure the

Table 3 Benchmark descriptions and equations used to describe the transport mechanisms

Benchmark	Description	Eq.
1	Coupled electrochemical migration diffusion	Eq. 13
2	Coupled electrochemical migration diffusion through an uncharged membrane	Eq. 13
3	Coupled electrochemical migration diffusion through a charged membrane with external electric field	Eqs. 6 and 7
4	Mineral precipitation and dissolution based on the advective inflow of $MgCl_2$through a calcite column	Eq. 15

Table 4 Chemical composition, diffusion coefficients, and boundary conditions for the electrochemical effect benchmark

Species	Diffusion coefficient $[m^2 \times s^{-1}]$	Boundary condition at $x=0$ $[mol \times m^{-3}]$	Initial condition at $t=10$ $[mol \times m^{-3}]$
H^+	9.311×10^{-9}	0.001×10^{-3}	1.0×10^{-4}
NO_3^-	1.902×10^{-9}	0.001×10^{-3}	1.0×10^{-4}
Na^+	1.334×10^{-9}	1.0×10^{-4}	1.0×10^{-4}
Cl^-	2.032×10^{-9}	1.0×10^{-4}	1.0×10^{-4}

geometric transport properties (tortuosity) of the membrane. The diffusion of HNO_3 induces an additional flux of KCl to maintain the charge balanced at all times. In the experiment, fluxes of K and Cl against their concentration gradient are observed. This is the so-called "uphill fluxes." It is when the movement of electrolyte originate from the different mobilities of the involved ions, and initially, such fluxes are controlled by the electrochemical migration term. The differences compared to benchmark 1 are as follows: (1) constantly stirred reservoirs on both sides instead of a constant concentration on the boundary; (2) separation of the solutions by a porous membrane.

To model this experiment properly, a volume adjustment factor was introduced in the weak formulation to account for the different cross-sectional areas and volume sizes of the nodes representing different domains (membrane and tanks). This allows to use a simple 1D FE discretization and no additional discretization of the reservoirs are needed. The membrane porosity is 0.44, the diameter is 23 mm, and the width is 0.13 mm. The domain is discretized into 29 cells, and the volume adjustment factor (α) for membrane cells is 4.156×10^{-4} while for tanks cells are 1.0. The updated version of Eq. 13 becomes as follows:

$$\int_\Omega \alpha\theta c_i^n v_i \mathrm{d}x = \int_\Omega \left(\alpha\theta c_i^{n-1} v_i - \alpha\theta\,\Delta t \sum_{k=1}^{m} D_{\mathrm{ik}} \nabla c_k^n\right) \cdot \nabla v_i \mathrm{d}x \tag{16}$$

The diffusion coefficients, as well as the chemical composition, are presented in Table 5. The initial condition for the membrane at $t = 0$ is assumed to be the same as the left reservoir.

Figures 3 and 4 present the evolution of the concentration in each reservoir over time. Our results show a very good agreement with Phreeqc, Flotran, and the experimental data. The small differences for temporal evolution of HTO can be explained by the different numerical methods and to the numerical precision among the compared codes.

The electromigration term fully controls the transport phenomena observed in this benchmark [98, 99] and, therefore, the traditional Fickian approach is not sufficient to adequately describe the fluxes observed experimentally, as also discussed in [100].

4.3 Benchmark 3: Coupled electrochemical diffusion through a charged membrane with external electric field

An external electric field is widely used in laboratory experiments [101] to accelerate the transport of charged aqueous species and assist measurements of ion transport properties in a timely manner. In benchmark 3, inspired by [81], the diffusion of ions through a charged membrane is modeled in the presence of an external electric field. The chemical composition and diffusion coefficients are presented in Table 6. The membrane's porosity is 1.0 and a charge of 50 mol m^{-3} is equally distributed at every membrane node. The magnitude of the applied membrane charge is comparable with the typical surface charge of clay and cement minerals [102].

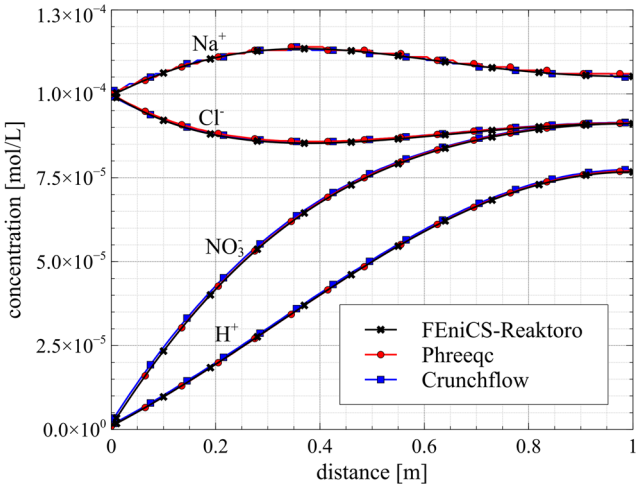


Fig. 2 Comparison of the electrochemical benchmark results, inspired in [97] and [96], after 1-h diffusion time

Table 5 Chemical composition, diffusion coefficients, and left and right reservoirs concentrations

Species	Diffusion coefficient $[m^2 \cdot s^{-1}]$	Right reservoir $[mol \times L^{-1}]$	Left reservoir $[mol \times L^{-1}]$
H^+	9.0×10^{-9}	0.001×10^{-3}	1.0×10^{-3}
NO_3^-	1.9×10^{-9}	0.001×10^{-3}	1.0×10^{-3}
K^+	$2.0 \cdot 10^{-9}$	$1.0 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$
Cl^-	$2.0 \cdot 10^{-9}$	$1.0 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$
HTO	$2.3 \cdot 10^{-9}$	$3.44 \cdot 10^{-10}$	$1.0 \cdot 10^{-15}$

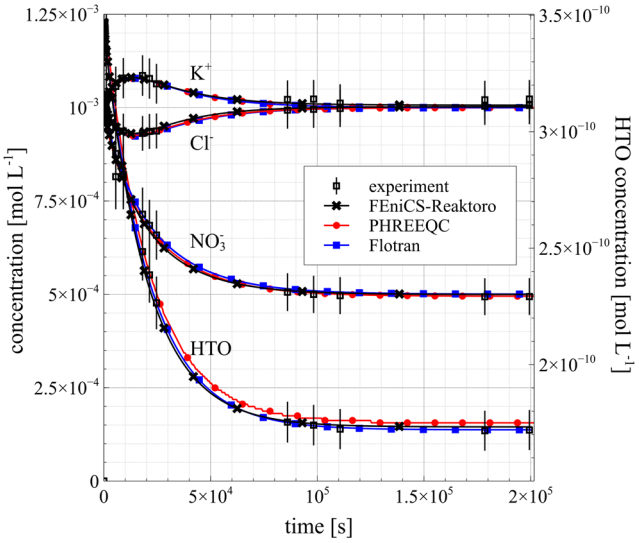


Fig. 3 Left upstream reservoir evolution of reservoir concentration with time for charged species diffusion through a porous uncharged membrane

The external electric gradient of 13 V m^{-1} is oriented from left to right. The dielectric constant of the membrane and the temperature are set to 7.0832×10^{-10} C V^{-1} m^{-1} and 300 K, respectivelly. The 1 cm thick membrane is discretized into 50 equidistant cells. This benchmark represents the stationary problem obtained from Eq. 7 by setting the change in concentration over time to zero:

$$0 = \int_{\Omega} -\theta D_i (\nabla c_i^n + \frac{z_i F c_i^n}{RT} \nabla \Phi) \cdot \nabla v_i \mathrm{d}x \tag{17}$$

Figure 5 presents the comparison between our results and the values calculated by [81] for the stationary case of

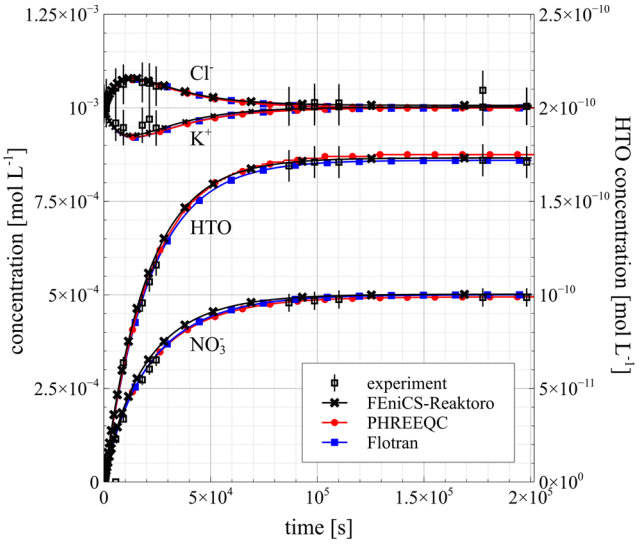


Fig. 4 Right downstream reservoir evolution of reservoir concentration with time for charged species diffusion through a porous uncharged membrane

Table 6 Chemical Composition, diffusion coefficient and boundary conditions

Species	Diffusion coefficient $[m^2 \cdot s^{-1}]$	Left boundary condition $[mmol \cdot L^{-1}]$	Right boundary condition $[mmol \cdot L^{-1}]$
SO_4^{2-}	$3.0 \cdot 10^{-10}$	50.0	41.725
Mg^{2+}	$3.0 \cdot 10^{-10}$	25.0	0.0
Na^{+}	$3.0 \cdot 10^{-10}$	0.0	8.45
K^{+}	$5.0 \cdot 10^{-10}$	0.0	25.0

coupled ion diffusion through a charged membrane in the presence of an external electric field.

Ions are forced by the external electric field to move towards the oppositely charged electrode. The concentration of counter-charged ions near the charged surface will be enriched to compensate the electric potential while same-charge ions will be repelled, as observed in clay materials [103]. In a highly compacted medium, as the nanopores in clay interlayers, the counter-ions can be entirely excluded from the pore space between charged mineral surfaces [78]. The obtained results are in perfect agreement with the simulations of [81].

4.4 Benchmark 4: Mineral precipitation and dissolution due to injection of $MgCl_2$ into a calcite column

This benchmark, originally proposed by [104], is used for the validation of several reactive transport codes [13, 15, 105–107]. The setup consist of a 0.5-m long 1D calcite column which is infiltrated with $MgCl_2$ through the left side boundary. The initial system porosity is 0.32 and the advective velocity of pore fluid is 9.375×10^{-6}m s^{-1}. A detailed description of the chemical and transport properties

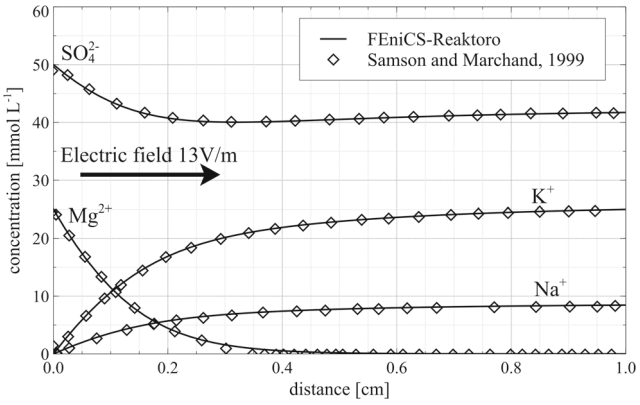


Fig. 5 Results for the stationary diffusion through a charged membrane in the presence of an external electric field

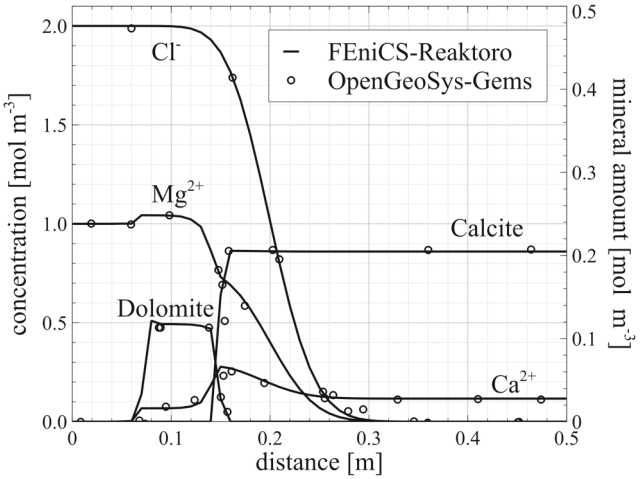


Fig. 6 Comparison of a mineral dissolution-precipitation benchmark where a calcite column is flushed with $MgCl_2$ causing calcite dissolution and dolomite precipitation

of the system can be found in [106]. The constant inflow of $MgCl_2$ causes calcite dissolution and temporary dolomite precipitation close to the reaction front as well as an increase of Mg^{2+} in solution. Figure 6 presents the results obtained after 21,000 s of simulation and demonstrates a good agreement between the results of FEniCS–Reaktoro and OpenGeoSys–GEMS.

5 Conclusion

The development of FEniCS–Reaktoro framework was driven by the need for an advanced open-source tool that could provide a flexible modeling environment and, at the same time, allow full control over the numerical methods. FEniCS–Reaktoro was developed using an object-oriented approach and permits the solution of user-defined single or multiphase fluid transport problems in saturated porous media coupled to thermodynamic and kinetic chemical equilibria between a fluid and numerous solid phases. An intuitive math equations alike scripting language allows quick system setup of arbitrary complexity regarding process couplings, domain geometry, and the boundary conditions without need of deep programming skills. All in all, FEniCS–Reaktoro is a goal-oriented and easy to learn framework for reactive transport modeling that combines the computing engine from two state-of-the-art platforms for describing physical and chemical processes.

Several peculiar features of the FEniCS–Reaktoro include the ability to tackle the transport equations at a PDE level allowing adjustments, changes, and implementation of new terms, in a straightforward way. Another aspect of FEniCS–Reaktoro is the combination of electrochemical migration transport, based on the Nernst–Planck equation with the ability to deal with complex multiphase chemical systems including solid solutions. FEniCS–Reaktoro is compatible with distributed memory HPC architecture offered with PETSc-based parallelization.

The chemical solver, Reaktoro, is probably the most advanced chemical solver available and combines the advantages of existing LMA and GEM thermodynamic solvers. Recently, the standalone numerical kernel of Reaktoro has been significantly improved in numerical performance and convergence speed thanks to machine learning–based algorithms, which help to obtain an initial guess close to the optimal solution.

The open-source FEniCS–Reaktoro framework for reactive transport simulations was described in detail and validated against several conceptually different experimental and numerical benchmarks. Overall, an excellent agreement between the simulation results and the benchmark could be achieved. FEniCS–Reaktoro could reproduce the essential transport relevant processes and chemical reactions consistently demonstrating an enormous potential for further applications.

Acknowledgments We thank Phillip Krejci, Dr. Thomas Gimmi, and Dr. Dmitrii Kulik for helpful discussions during the development of the FEniCS–Reaktoro framework. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Swiss Government.

Funding information This work was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 15.0186-2.Authors receive partial financial support from Nagra. The research leading to these results has received funding from the European Union's Horizon 2020 Research and Training Programme of the European Atomic Energy Community (EURATOM) (H2020 - NFRP - 2014 / 2015) under grant agreement n^o662147 (CEBAMA).

Appendix A

The core class of FEniCS–Reaktoro encapsulates two parts: (1) the initialization routine and (2) the reactive transport loop. The initialization routine prepares the underlying structure and defines multiple properties: temporal and spatial; initial conditions, boundary conditions, and constants; single diffusion coefficient or species-dependent diffusion coefficients; the equation system; and the chemical systems. The reactive transport loop is where the transport and chemical equations are solved within a one-time loop until total simulation time is reached.

The finite element mesh and the chemical system are defined during the initialization routine before the reactive transport loop. This process includes the definition of (1) all species and phases involved in the chemical reactions; (2)

governing equations necessary to describe the transport of such species; and (3) initialization of the cells necessary for the domain representation using FE infrastructure provided by FEniCS.

Figures 7 and 8 present the class and sequence diagram, respectively, of the FEniCS–Reaktoro framework using the UML language.

The usage of FEniCS–Reaktoro coupled code to run a reactive transport simulations requires the generation of the following:

- The chemical system file: the chemical system file can be generated using (a) the GEM-Selektor v.3 graphical user interface, (b) the Phreeqc graphical user interface, or (c) Reaktoro's manual chemical system editor
- The mesh file: the mesh file can be generated using (a) GMSH or (b) FEniCS, depending on the complexity of the geometry
- The 4 files included in the user-defined input package: main input file, transport file, boundary conditions file and diffusion coefficients file

The detailed composition of the 4 files from the user-defined input package is as follows:

- Main input file: contains the temporal and spatial properties, finite element family and degree, type of transport equation used, chemical composition (composed by external chemical definition files), and constants (optional)
- Transport file: contains the description of the physical processes, through the definition of the weak formulation of the PDE
- The boundary conditions file: defines the boundary conditions of the simulation
- Diffusion coefficients file: defines the diffusion coefficient of the species (single or multicomponent diffusion are supported).

We point to the framework's repository located at bitbucket.org/lhdamiani/fenics-reaktoro, where one can find instructions, licensing information, documentation, and demos.

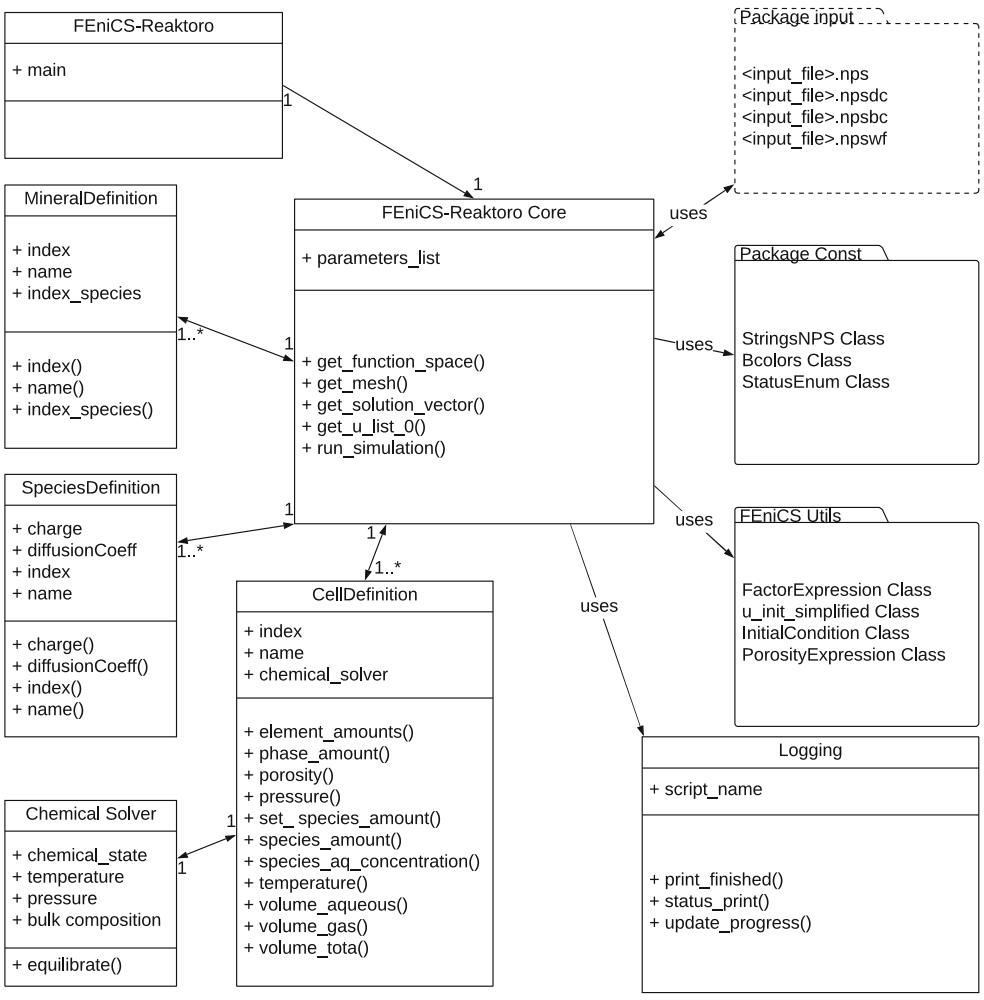


Fig. 7 The UML class diagram description of FEniCS–Reaktoro framework

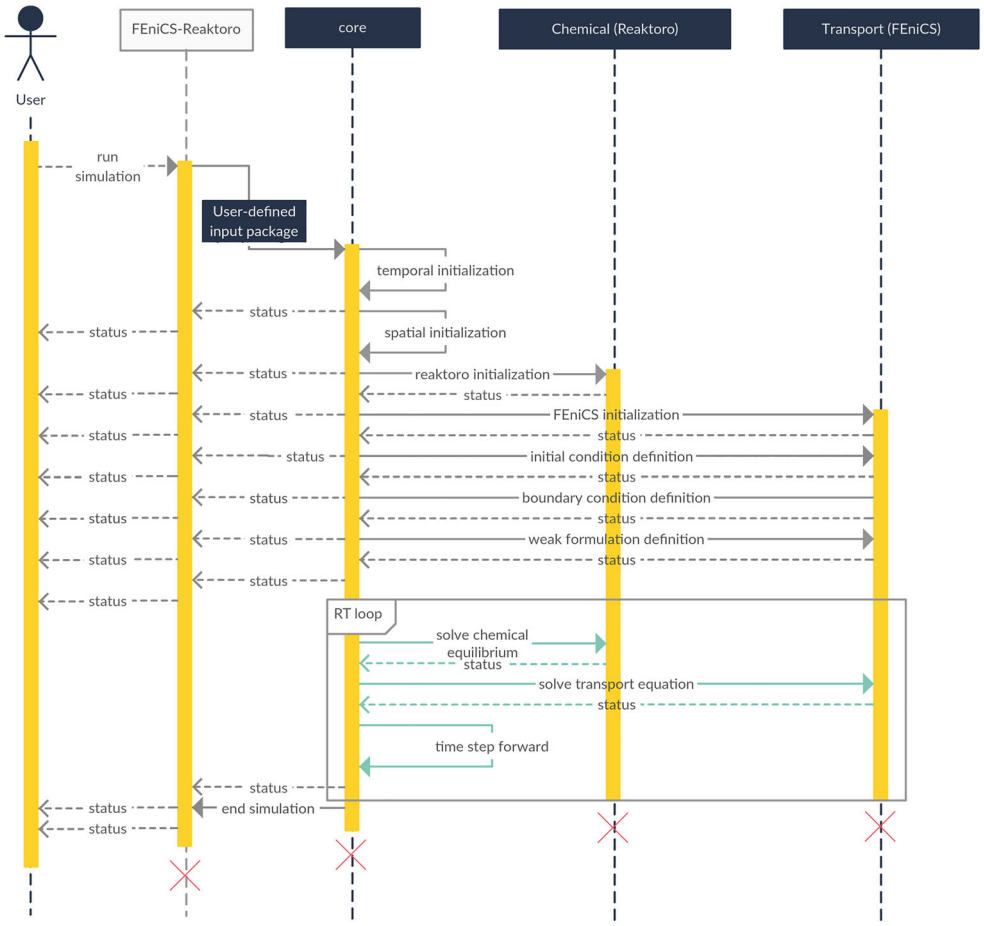


Fig. 8 The FEniCS–Reaktoro framework UML sequence diagram depicting the interaction between the components

References

1. Steefel, C.I., DePaolo, D.J., Lichtner, P.C.: Reactive transport modeling: an essential tool and a new research approach for the Earth sciences. Earth Planet Sci. Lett. **240.3-4**, 539–558 (2005)
2. Bächler, D., Kohl, T.: Coupled thermal-hydraulic-chemical modelling of enhanced geothermal systems. Geophys. J Int. **161.2**, 533–548 (2005)
3. Darland, J.E., Inskeep, W.P.: Effects of pore water velocity on the transport of arsenate. J. Am. Chem. Soc. (1997)
4. Lichtner, P.C.: Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. Geochim. Cosmochim. Acta **49.3**, 779–800 (1985)
5. Tournassat, C., Steefel, C.I.: Ionic transport in nano-porous clays with consideration of electrostatic effects. Rev. Min. Geochem. **80**, 287–329 (2015)
6. Muniruzzaman, M., Rolle, M.: Modeling multicomponent ionic transport in groundwater with IPhreeqc coupling Electrostatic interactions and geochemical reactions in homogeneous and heterogeneous domains. Adv. Water Res. **98**, 1–15 (2016)
7. Konda, S.S.M., et al.: Chemical reactions modulated by mechanical stress: extended Bell theory. J. Chem. Phys. **135.16**, 164103 (2011)
8. Dmitrii, A. et al.: Kulik GEM-selektor geochemical modeling package: revised algorithm and GEMS3k numerical kernel for coupled simulation codes. Computat. Geosci. **17.1**, 1–24 (2013)
9. Centler, F. et al.: GeosysBRNS-a exible multidimensional reactive transport model for simulating biogeochemical subsurface processes. Comput. Geosci.-UK. **36.3**, 397–405 (2010)
10. Steefel, C., Molins, S.: CrunchFlow software for Modeling Multicomponent Reactive Flow and Transport (2016)
11. Nardi, A. et al.: Interface COMSOL-PHREEQC (iCP), an efficient numerical framework for the solution of coupled multiphysics and geochemistry. Comput. Geosci.-UK. **69**, 10–21 (2014)
12. Guo, B. et al.: A COMSOL-PHREEQC interface for modeling the multi-species transport of saturated cement-based materials. Construct. Build Mater. **187**, 839–853 (2018)
13. Azad, V.J. et al.: A COMSOL-GEMS interface for modeling coupled reactive-transport geochemical processes. Comput. Geosci. **92**, 79–89 (2016)
14. Samper, J. et al.: Core^{2D}. A code for non-isothermal water flow and reactive solute transport. Users Manual version 2 (2000)
15. Yapparova, A. et al.: Reactive transport modelling of dolomitisation using the new CSMP++GEM coupled code: governing equations, solution method and benchmarking results. Transport Porous Med. **117.3**, 385–413 (2017)
16. Elakneswaran, Y., Ishida, T.: Development and verification of an integrated physicochemical and geochemical modelling framework for performance assessment of cement-based materials. J. Adv. Concr. Technol. **12.4**, 111–126 (2014)
17. Jacques, D., Simunek, J.: User manual of the multicompenent variably-saturated ow and transport model hp1 (2005)
18. Mayer, K.U.: A numerical model for multicomponent reactive transport in variably saturated porous media. PhD thesis (1999)
19. Bethke, C.M.: Geochemical Reaction Modeling: Concepts and Applications. Oxford University Press, Oxford, pp. 416 (1996)
20. Li, D. et al.: OpenGeoSys-ChemApp: a coupled simulator for reactive transport in multiphase systems and application to CO_2

storage formation in Northern Germany. Acta Geotech. **9.1**, 67–79 (2014)
21. Kosakowski, G., Watanabe, N.: OpenGeoSys-Gem: a numerical tool for calculating geochemical and porosity changes in saturated and partially saturated media. Phys. Chem. Earth **70–71**, 138–149 (2014). Mechanisms and Modelling of Waste-Cement and Cement-Host Rock Interactions
22. He, W. et al.: A parallelization scheme to simulate reactive transport in the subsurface environment with OGS#IPhreeqc 5.5.7-3.1.2. Geosci. Model Dev. **8.10**, 3333–3348 (2015)
23. Meeussen, J.C.L.: ORCHESTRA: An object-oriented framework for implementing chemical equilibrium models. Environ. Sci. Tech. **37.6**, PMID: 12680672, 1175–1182 (2003). eprint: https://doi.org/10.1021/es025597s
24. Lichtner, P.C., et al.: PFLOTRAN User Manual A Massively Parallel Reactive Flow and Transport Model for Describing Surface and Subsurface Processes (2015)
25. Parkhurst, D.L., 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. U.S. Geological Survey Techniques and Methods, book 6, chapter A43, 497 p. U.S. Geological Survey Techniques and Methods, book 6, chapter A43, 6-43A (2013)
26. Gamazo, P. et al.: PROOST: object-oriented approach to multiphase reactive transport modeling in porous media. J. Hydroinform. **18**(2), 310–328 (2016)
27. Georget, F., Prévost, J.H., Huet, B.: A reactive transport simulator for variable porosity problems. Computat. Geosci. **21.1**, 95–116 (2017)
28. Soetaert, K., Meysman, F.: Reactive transport in aquatic ecosystems Rapid model prototyping in the open source software R. Environ. Model. Softw. **32**, 49–60 (2012)
29. Xu, T. et al.: TOUGHREACT-a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: applications to geothermal injectivity and CO_2 geological sequestration. Comput. Geosci.-UK. **32.2**, 145–165 (2006)
30. Jara, D., de Dreuzy, J.-R., Cochepin, B.: TReacLab: an object-oriented implementation of non-intrusive splitting methods to couple independent transport and geochemical software. Comput. Geosci-UK. **109**, 281–294 (2017)
31. Su, D., Ulrich Mayer, K., MacQuarrie, K.T.B.: Parallelization of MIN3p-THCm: a high performance computational framework for subsurface flow and reactive transport simulation. Environ. Model. Softw. **95**, 271–289 (2017)
32. Trebotich, D. et al.: High-resolution simulation of pore-scale reactive transport processes associated with carbon sequestration. Comput. Sci. Eng. **16.6**, 22–31 (2014)
33. James, J. et al.: Beisman ParCrunchFlow: an efficient, parallel reactive transport simulation tool for physically and chemically heterogeneous saturated subsurface environments. Computat. Geosci. **19.2**, 403–422 (2015)
34. Smith, W.R., Missen, R.W.: Chemical reaction equilibrium analysis : theory and algorithms. Wiley, New York, pp. 364 (1982)
35. Leal, M.M. et al.: Computational methods for reactive transport modeling: an extended law of mass-action, xLMA, method for multiphase equilibrium calculations. Adv. Water Res. **96**, 405–422 (2016)
36. Comsol Multiphysics, The Platform for Physics-Based Mod- eling and Simulation. Comsol Inc., Burlington, (2013)
37. Bell, L.S.J., Binning, P.J.: A split operator approach to reactive transport with the forward particle tracking Eulerian Lagrangian localized adjoint method. Adv Water Res. **27.4**, 323–334 (2004)
38. Saaltink, M.W., Carrera, J., Ayora, C.: On the behavior of approaches to simulate reactive transport. J. Contam. Hydrol. **48.3-4**, 213–35 (2001)
39. Carrayrou, J. et al.: Comparison of numerical methods for simulating strongly nonlinear and heterogeneous reactive transport problems-the MoMaS benchmark case. Computat. Geosci. **14.3**, 483–502 (2010)
40. Hoffmann, J., Kräutle, S., Knabner, P.: A parallel global-implicit 2-D solver for reactive transport problems in porous media based on a reduction scheme and its application to the MoMaS benchmark problem. Computat. Geosci. **14.3**, 421–433 (2010)
41. Yeh, G.T., Tripathi, V.S.: A critical evaluation of recent developments in hydrogeochemical transport models of reactive multichemical components. Water Resour. Res. **25.1**, 93–108 (1989)
42. Walter, A.L. et al.: Modeling of multicomponent reactive transport in groundwater: 1. Model development and evaluation. Water Resour. Res. **30.11**, 3137–3148 (1994)
43. Samper, J., Xu, T., Yang, C.: A sequential partly iterative approach for multicomponent reactive transport with CORE2D. Computat. Geosci. **13.3**, 301–316 (2009)
44. Li, L. et al.: Expanding the role of reactive transport models in critical zone processes. Earth-Sci. Rev. **165**, 280–301 (2017)
45. Ravi, A. et al.: Patel A three-dimensional lattice Boltzmann method based reactive transport model to simulate changes in cement paste microstructure due to calcium leaching. Con- Str. Build. Mater. **166**, 158–170 (2018)
46. Fowler, D. et al.: Atmospheric composition change: ecosystems-atmosphere interactions. Atmos. Environ. **43.33**, 5193–5267 (2009)
47. Dentz, M. et al.: Mixing, spreading and reaction in heterogeneous media: a brief review. J. Contam. Hydrol. **120-121**, 1–17 (2011)
48. Ian, G. et al.: Microbiology in nuclear waste disposal: interfaces and reaction fronts. FEMS Microbiol. Rev. **20.3-4**, 545–556 (1997)
49. Kosakowski, G., Berner, U.: The evolution of clay rock/cement interfaces in a cementitious repository for low and intermediate level radioactive waste. Phys. Chem. Earth Parts A/B/C **64**, 65–86 (2013)
50. Kolditz, O. et al.: OpenGeoSys: an open-source initiative for numerical simulation of thermo - hydro - mechanical / chemical (THM/c) processes in porous media. Environ. Earth Sci. **67.2**, 589–599 (2012)
51. Choo, J., Lee, S.: Enriched Galerkin finite elements for coupled poromechanics with local mass conservation. Comput. Methods Appl. Mech Eng. **341**, 311–332 (2018)
52. Lehmann, C., Kolditz, O., Nagel, T.: The FEM Simulation Software OpenGeoSys, vol. 6, pp. 29–48. Springer, Cham (2018)
53. Sharma, P.K., Joshi, N., Ojha, C.P.: Reactive transport through porous media using finite-difference and finite-volume methods. J. Hydraul. Eng. **18.1**, 11–19 (2012)
54. Prasianakis, N.I. et al.: Deciphering pore-level precipitation mechanisms. Sci. Rep. **7.1**, 1–9 (2017)
55. Zhao, C.-b., Schaubs, P., Hobbs, B.: Effects of porosity heterogeneity on chemical dissolution-front instability in fluid-saturated rocks. J. Cent South Univ. **24.3**, 720–725 (2017)
56. Hatanaka, A. et al.: The impact of tortuosity on chloride ion diffusion in slag-blended cementitious materials. J. Adv. Concr. Technol. **15.8**, 426–439 (2017)
57. Navarre-Sitchler, A. et al.: Evolution of porosity and diffusivity associated with chemical weathering of a basalt clast. J. Geophys. Res. **114.F2**, F02016 (2009)
58. Ma, R. et al.: Assessment of controlling processes for field-scale uranium reactive transport under highly transient flow conditions. Water Resour. Res. **50.2**, 1006–1024 (2014)

59. Van Loon, L.R. et al.: Anisotropic diffusion in layered Argillaceous rocks: a case study with Opalinus clay (2004)
60. Hommel, J., Coltman, E., Class, H.: Porosity-permeability relations for evolving pore space: a review with a focus on (bio-)geochemically altered porous media. Transport Porous Med. **124.2**, 589–629 (2018)
61. Trinchero, P. et al.: Implications of grain-scale mineralogical heterogeneity for radionuclide transport in fractured media. Transport Porous Med. **116.1**, 73–90 (2017)
62. Steefel, C.I., Yabusaki, S.B., Ulrich Mayer, K.: Reactive transport benchmarks for subsurface environmental simulation. Computat. Geosci. **19.3**, 439–443 (2015)
63. Hunter, J.D.: Matplotlib: a 2D graphics environment. Comput. Sci. Eng. **9.3**, 90–95 (2007)
64. Perez, F., Granger, B.E.: IPython: A system for interactive scientific computing. Comput. Sci. Eng. **9.3**, 21–29 (2007)
65. McKinney, W.: Data structures for statistical computing in Python (2010)
66. Oliphant, T.E.: Guide to NumPy. 2nd USA: Createspace Independent Publishing Platform (2015)
67. Martin, S. et al.: Alnaes The FEniCS Project Version 1.5. Arch. Numer. Softw. **3.100**, 9–23 (2015)
68. Allan, M.M. et al.: An overview of computational methods for chemical equilibrium and kinetic calculations for geochemical and reactive transport modeling. Pure Appl. Chem. **89.5**, 597–643 (2017)
69. Alnæs, M.S., et al.: Unified form language: a domain speci fic language for weak formulations of partial differential equations. ACM Trans. Math. Softw. **40.2**, 9:1-9:37 (2014)
70. Hoffman, J.J.J., Jansson, N.: FEniCS-HPC automated predictive high-performance finite element computing with applications in aerodynamics. Parallel Process. Appl. Math. **9573**, 356–365 (2016)
71. Abhyankar, S. et al.: PETSc/TS: A modern scalable ODE/DAE Solver Library. arXiv:1806.01437 (2018)
72. Spycher, N., Pruess, K., Ennis-King, J.: CO_2-H_2O mixtures in the geological sequestration of CO_2. I. Assessment and calculation of mutual solubilities from 12 to 100 °C and up to 600 bar. Geochim. Cosmochim. Acta **67.16**, 3015–3031 (2003)
73. Spycher, N.F., Reed, M.H.: Fugacity coefficients of H_2, CO_2, CH_4, H_2O and of H_2O- CO_2-CH_4 mixtures: a virial equation treatment for moderate pressures and temperatures applicable to calculations of hydrothermal boiling. Geochim. Cosmochim. Acta **52.3**, 739–749 (1988)
74. Matschei, T., Lothenbach, B., Glasser, F.P.: Thermodynamic data for hydrated solids in Portland cement system CaO-Al2o3-Sio2-CaSO4-CaCO3- Fe2O3-MgO-H2O. Cem. Concr. Res. **37**, 1379–1410 (2007)
75. Thoenen, T., et al.: The PSI/Nagra Chemical Thermodynamic Database 12/07 Nuclear Energy and Safety Research Department Laboratory forWaste Management (LES) (2014)
76. Johnson, J.W., Oelkers, E.H., Helgeson, H.C.: SUPCRT92: a software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000 °c. Comput. Geosci. **18.7**, 899–947 (1992)
77. Wigger, C., Van Loon, L.R.: Importance of interlayer equivalent pores for anion diffusion in clay-rich sedimentary rocks. Environ. Sci. Tech. **51.4**, 1998–2006 (2017)
78. Van Loon, L.R., Glaus, M.A., Müller, W.: Anion exclusion effects in compacted bentonites: towards a better understanding of anion diffusion. Appl. Geochem. **22.11**, 2536–2552 (2007)
79. Altmann, S., et al.: Diffusion-driven transport in clayrock formations. Appl. Geochem. **27.2**, 463–478 (2012). arXiv: 1-1
80. Bear, J., Bachmat, Y.: Introduction to Modeling of Transport Phenomena in Porous Media. Springer, Dordrecht (1990)
81. Samson, E., Marchand, J.: Numerical solution of the extended Nernst-Planck model. J Colloid Interface Sci. **215**, 1–8 (1999)
82. Daus, A.D., Frind, E.O.: An alternating direction Galerkin technique for simulation of contaminant transport in complex groundwater systems. Water Resour. Res. **21.5**, 653–664 (1985)
83. Courant, R., Friedrichs, K., Lewy, H.: ÜBer die partiellen Differenzengleichungen der mathematischen Physik. Mathematische Annalen **100.1**, 32–74 (1928)
84. Isaacson, E., Keller, H.B.: Analysis of numerical methods dover books on mathematics. Dover Publications (1994)
85. Murdoch, J.R.: What is the rate-limiting step of a multistep reaction? J. Chem. Educ. **58.1**, 32 (1981). eprint: https://doi.org/10.1021/ed058p32
86. Zhang, Y.: Geochemical kinetics. Princeton University Press, pp. 631 (2008)
87. Arnold, D.N. et al.: Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems (2002)
88. Riviére, B.: Discontinuous Galerkin methods for solving elliptic and parabolic equations : theory and implementation. SIAM, Society for Industrial and Applied Mathematics, pp. 190 (2008)
89. Zhang, C., Zarrouk, S.J.: Rosalind Archer. A mixed finite element solver for natural convection in porous media using automated solution techniques. Comput. Geosci.-UK. **96**, 181–192 (2016)
90. Houston, P., Sime, N.: Automatic symbolic computation for discontinuous Galerkin finite element methods. SIAM J. Sci Comput. **40.3**, C327–C357 (2018)
91. Berner, U., Kulik, D.A., Kosakowski, G.: Geochemical impact of a low-pH cement liner on the near field of a repository for spent fuel and high-level radioactive waste. Phys. Chem. Earth **64**, 46–56 (2013)
92. Steefel, C.I., Maher, K.: Fluid-rock interaction: a reactive transport approach. Rev. Mineral. Geochem. **70.1**, 485–532 (2009)
93. Quintard, M., Whitaker, S.: Transport in ordered and disordered porous media: volume - averaged equations, closure problems, and comparison with experiment. Chem. Eng. Sci. **48.14**, 2537–2564 (1993)
94. Lasaga, A.C.: Kinetic Theory in the Earth Sciences, pp. 811. arXiv:1011.1669v3 (1998)
95. Rumbaugh, J., Jacobson, I., Booch, G.: The unified modeling language reference manual. Addison-Wesley, pp. 550 (1999)
96. Rasouli, P. et al.: Benchmarks for multicomponent diffusion and electrochemical migration. Computat. Geosci. **19.3**, 523–533 (2015)
97. Lichtner, P.C.: Principles and Practice of Reactive Transport Modeling (1994)
98. Martin, A. et al.: Seeming steady-state uphill diffusion of $^{22}Na+$ in compacted montmorillonite. Environ. Sci. Technol. **47.20**, 11522–11527 (2013)
99. Glaus, M.A. et al.: Cation diffusion in the electrical double layer enhances the mass transfer rates for Sr2+, Co2+ and Zn2+ in compacted illite. Geochim. Cosmochim. Acta **165**, 376–388 (2015)
100. Massimo, R. et al.: Nernst-Planck-based description of transport, coulombic interactions, and geochemical reactions in porous media: modeling approach and benchmark experiments. Water Resour. Res. **54.4**, 3176–3195 (2018)
101. Maes, N. et al.: Determination of the diffusion coefficient of ionic species in boom clay by electromigration: feasibility study. Radiochim. Acta **82.s1**, 183–190 (1998)
102. Berner, U.: Radionuclide concentration limits in the cementitious near-field of an ILW repository (2003)

103. Jenni, A. et al.: In situ interaction between different concretes and Opalinus clay. Phys. Chem. Earth **70-71**, 71–83 (2014)
104. Engesgaard, P., Kipp, L.K.: A geochemical transport model for redox-controlled movement of mineral fronts in groundwater ow systems: a case of nitrate removal by oxidation of pyrite. Water Resour. Res. **28.10**, 2829–2843 (1992)
105. Prommer, H., Barry, D.A., Zheng, C.: MODFLOW/ MT3DMS - Based reactive multicomponent transport modeling. Ground Water **41.2**, 247–257 (2002)
106. Shao, H., et al.: Modeling reactive transport in non-ideal aqueous-solid solution system. Appl. Geochem. **24.7**, 1287–1300 (2009)
107. He, W.: Code verification: Engesgaard benchmark. OpenGeoSys Tutorial: Computational Hydrology III: OGS # IPhreeqc Coupled Reactive Transport Modeling. Springer International Publishing, Berlin, pp. 31–35 (2018)

Affiliations

Leonardo Hax Damiani[1,2] · Georg Kosakowski[1] · Martin A. Glaus[1] · Sergey V. Churakov[1,2]

Georg Kosakowski
georg.kosakowski@psi.ch

Martin A. Glaus
martin.glaus@psi.ch

Sergey V. Churakov
sergey.churakov@psi.ch

[1] Laboratory for Waste Management, Paul Scherrer Institut, 5232 Villigen, Switzerland

[2] Institute of Geological Sciences, University of Bern, 3012 Bern, Switzerland

Chapter 3

Hydrogen gas transfer between a borehole and claystone: Experiment and geochemical model

https://doi.org/10.1680/jenge.21.00061

Mont Terri HT experiment: Modelling of gas transfer between a borehole and claystone

Leonardo Hax Damiani[1,3], Georg Kosakowski[1*], Agnès Vinsot[2], Sergey V. Churakov[1,3]

[1] *Laboratory for Waste Management, Paul Scherrer Institut, Villigen Switzerland*

[2] *Andra, CMHM, Bure, France*

[3] *Institute of Geological Sciences, University of Bern, Bern, Switzerland*

[*] *corresponding author: georg.kosakowski@psi.ch*

Revised version for submission to *Environmental Geotechnics* special issue on *"Physical-Chemical Coupling in Environmental Geotechnics"*

Keywords

Geochemistry, Multiphase science, Numerical methods

Abstract

The Hydrogen Transfer (HT) experiment, located at the Mont Terri underground rock laboratory in Switzerland, is an in situ experimental study of the interactions and transport of hydrogen injected into a borehole installed within Opalinus Clay, a claystone formation.

A python-based model was developed to analyse and model the experimental data, for diffusion of dissolved gases and solutes in claystone pore water, for thermodynamic modelling of gas-water-solid phase equilibria in the injection interval and for simulations of chemical equilibria and reaction kinetics in claystone and injection interval.

The developed model reproduces the temporal evolution of gas pressure, composition and solute concentrations measured in situ with a minimum set of adjustable parameters. The effective diffusion coefficients for dissolved gases in Opalinus Clay derived by the modelling of experimental data were found to be very close to values measured in other experimental studies. It was discovered that an accurate description of the temporal variations in hydrogen injection and temporal changes in the inflow of formation water is essential for modelling of microbial mediated hydrogen consumption in the injection interval.

Notation

B molal concentration of biomass for Michaelis-Menten kinetics (mol/kg)

B_{max} max. biomass concentration for Michaelis-Menten kinetics (mol/kg)

b_{high} model variant with low biomass concentration

b_{low} model variant with high biomass concentration

$b_{reference}$ model with reference biomass concentration

C volumetric concentration (mol/m^3)

D_e effective diffusion coefficient (m²/s)

D_{e_high} effective diffusion coefficient, model with high values: (m²/s)

D_{e_low} effective diffusion coefficient, model with low values (m²/s)

D_p pore diffusion coefficient (m²/s)

D_w diffusion coefficient in water (m²/s)

f source/sink term ((mol/m^3)/s)

g_{eo} geometry factor

$k_{1/2}$ half-rate constants for Michaelis-Kinetics (mol/kg)

k_{Bd} first order death rate of the bacteria for Michaelis-Menten kinetics (s^{-1})

m material dependent exponent

m_x molal concentration for species x (mol/kg)

n amount (mol)

P pressure (Pa)

pCO_2 partial pressure for carbon dioxide (CO_2) (bar)

R molar gas constant ((J/K)/mol)

T, t_c temperature (°C)

T, t_K temperature (K)

t time (s)

V volume (m^3)

Y yield factor for Michaelis-Menten kinetics (mol B/mol sulfate (SO_4^{2-}))

η_w water viscosity ((kg/m)/s)

ϕ porosity

ϕ^m material dependent tortuosity

μ_{max} specific degradation rate for Michaelis-Menten kinetics (s^{-1})

∇x gradient operator with respect to the variable x

$\frac{\partial x}{\partial y}$ partial derivative of variable x with respect to variable y

Introduction

Deep geological disposal of radioactive waste is foreseen in several countries for long-term isolation of radioactive waste from the environment (Alexander and McKinley, 2007; IAEA, 2011). Some countries plan to use claystone formations as the host rock for their disposals, for example, France (Callovo-Oxfordian clay rock) and Switzerland (Opalinus Clay rock). In various disposal concepts, massive steel disposal casks are used for the initial isolation of high-level nuclear waste. The anoxic corrosion of steel may release large amounts of hydrogen (H_2). The microbial degradation of organic matter, if present, may produce methane (CH_4) and carbon dioxide (CO_2). Transport of such gases in claystone is also of great interest to other types of deep underground space usage. Because of the low-permeability matrix, claystone caprocks are often the seal for oil and gas reservoirs (Grunau, 1987; Selley and Sonnenberg, 2014), for carbon dioxide sequestration (Benson and Cole, 2008; Gaus et al., 2005) or the storage of hydrogen in the geological underground (Crotogino, 2016).

The Hydrogen Transfer (HT) experiment, located at the Mont Terri Rock Laboratory, is implemented to investigate the in situ diffusion of various dissolved gases in Opalinus Clay (Vinsot et al., 2014, 2017). The experimental setup is demonstrated in Figure 1. The studies

are focused on (a) the diffusion of inert gases and hydrogen in the Opalinus Clay; (b) the interaction of hydrogen with the clay rock; and (c) the potential for microbial hydrogen consumption. Experimental data suggest a very fast decrease of the hydrogen amount in the injection interval, which could not be explained by the hydrogen diffusion into the claystone alone and was accounted for by the biotic hydrogen consumption involving hydrogen oxidation, sulfate reduction, and Fe(III) reduction (Vinsot et al., 2014). Appelo and Vinsot (2012) modelled the HT experiment using the PHREEQC code (Parkhurst and Appelo, 2013) and successfully reproduced key experimental observations (Vinsot et al., 2014, 2017). The previous modelling approach leaves potential for improvement regarding the representation of the borehole tilt, the calculation of gas fluxes between the injection reservoir and the rock, and resolution of concentration gradients near the borehole/rock interface.

This work presents a new model for the HT experiment, which couples a reactive thermodynamic model describing the chemical conditions in the injection reservoir with a reactive transport model for diffusion of dissolved gases in the Opalinus Clay. It was successfully applied to re-produce the general evolution of measured pressure, gas phase composition and water composition. Furthermore, different conceptualizations of modelled processes were tested in order to improve the understanding of the processes and process couplings that influence the experiment.

HT experiment

The HT experimental setup in Figure 1 is extensively described in Vinsot et al. (2014, 2017). A borehole of 0.076 m diameter was drilled under an argon atmosphere perpendicular to the anisotropic Opalinus Clay bedding, with a 48^{o} inclination. A 5 m long injection interval section with a total volume of 0.0095m^{3} (9.5 litres) was isolated with packers. It was equipped with temperature and pressure sensors and circulation tubes to allow injection or extraction of gases

and water (Vinsot et al., 2017). Measurement and circulation lines are made of stainless steel for gas and PEEK (polyether ether ketone) for water to prevent chemical interaction with the fluids. The installation includes a stainless-steel inner tube coated with PFA (perfluoroalkoxy alkanes) of 40 mm diameter surrounded by a ceramic filter screen of 14.75 mm thickness and porosity of 0.42.

The experiment started on the 21st of April 2009, when the drilling was finished. Until now, three experimental phases were conducted. Vinsot et al. (2017) reports the first two phases, a calibration phase (2009-2010) and the first hydrogen injection phase (2011-2013). The measurement and modelling for the second hydrogen injection phase (2015-2017) are published for the first time in this paper. More details on the chronology of the tests are given in the electronic Supplement.

Model implementation

The new model developed in this work uses a newly implemented python interface to couple FEniCS (Alnæs et al., 2015), a library for solving partial differential equations with Finite-Element methods, and the Gibbs Energy based thermodynamic solver GEMS3K (Kulik et al., 2012). The latter was used for thermodynamic modelling of borehole gas and liquid phases and Opalinus Clay pore-water composition. A 2D Finite Element (FE) model was applied for diffusive transport of dissolved gases and solutes in the claystone. The thermodynamic pressure and equilibrium chemical composition of gas, fluid and mineral phases in the injection interval were calculated based on the minimisation of Helmholtz Energy with GEMS3K. The thermodynamic model for injection interval and the transport model for host rocks were coupled via flux terms for dissolved gases and solutes. The composition of the water phase in the borehole provides a concentration boundary for the transport model, while diffusive fluxes into and out of the injection interval were calculated from transport model results based on

concentration gradients at the borehole wall. In this approach, the transport of dissolved gases and other species in the rock pore-water is characterised by transient and relatively large-scale concentration gradients in the claystone induced by fast and strongly changing concentrations at the borehole/rock interface. Post-processing of the modelling results was implemented with Jupyter notebooks (Kluyver et al., 2016), allowing a well-documented and reproducible data analysis and graphical representation of the simulation results.

Modelled processes

Diffusive transport in claystone

Transport of dissolved gases in the Opalinus Clay was implemented based on the partial differential equation for the diffusion in a water-saturated porous medium

$$\frac{\partial \phi C}{\partial t} = \nabla\left(\phi D_p \nabla C\right) + f \qquad (1)$$

where ϕ denotes the porosity; C denotes volumetric concentration (mol/m^3); D_p is the species-dependent pore diffusion coefficient (m^2/s), which is related to the effective diffusion coefficient $D_e = \phi\ D_p$; and f is a source/sink term ((mol/m^3)/s).

The influence of pore space geometry and temperature on diffusion can be considered via a material and temperature-dependent geometry factor g_{eo}, which includes a tortuosity ϕ^m in the porous medium and a correction term to a specific temperature t_K in Kelvin.

The tortuosity can be expressed via the porosity ϕ and a material-dependent exponent m

$$g_{eo} = \phi^m \cdot \frac{t_K \cdot 0.891}{298.0 \cdot \eta_w} \qquad (2)$$

The influence of temperature t_K =288.65 K (t_c =15.6 °C) for the investigated system was implemented via a temperature-dependent water viscosity η_w following Appelo and Postma (2005). Compared to 20 °C, the effective diffusion coefficient is reduced by factor 0.77 at 15.6 °C, similar to corrections proposed by Van Loon et al. (2005).

Finally, the effective diffusion coefficient for all species at the desired temperature was calculated based on the species diffusion coefficient in water (at reference temperature) and the geometry factor. Please note that the pore diffusion coefficient D_p is in this definition $\frac{D_w \cdot g_{eo}}{\phi}$

$$D_e = D_w \cdot g_{eo} = \phi \cdot D_p \quad (3)$$

The temperature and the porosity of the claystone are known. Therefore, the only unknown parameter in this definition of the effective diffusion coefficient is the tortuosity exponent *m*. Table 1 lists water diffusion coefficients, effective diffusion coefficients for dissolved gases, and gas diffusion coefficients used in this study. For dissolved gases species-specific diffusion coefficients were used. All other charged and neutral species were modelled using the same diffusion coefficient to ensure local charge conservation.

The Opalinus Clay is an anisotropic medium with different diffusion coefficients parallel and perpendicular to bedding (Gimmi et al., 2014; Van Loon, Soler, et al., 2004). The 3D natural geometry of the experiment was simplified based on the radial symmetry of the experimental layout. For the given geometry it is possible to use a 2D model domain perpendicular to the borehole and parallel to the rock bedding for transport calculations in Opalinus Clay.

Thermodynamic equilibration of borehole

Thermodynamic and geochemical conditions in the borehole and the Opalinus Clay pore-water were calculated using the python interface to the chemical equilibrium solver GEMS3K.

From the implemented fluid models (equation of states) described in Wagner et al. (2012), the ideal gas law was used.

$$\frac{V}{n} = \frac{RT}{P} \quad (4)$$

V denotes the volume (m^3); *n*, the amount of gas (mol); *T*, the temperature (K); *P*, the pressure (Pa); and *R*, the molar gas constant ($(J/K)/mol$).

GEMS3K uses the Gibbs Energy minimisation (GEM) method for calculation of chemical equilibria for a fixed bulk composition, temperature, and pressure while the volume is the free variable. Since the borehole and the attached gas circulation system have a fixed volume, the thermodynamic equilibrium for this system was calculated by minimising the Helmholtz energy based on constant bulk composition, temperature, and volume (Dimian et al., 2014). In the Helmholtz energy minimisation, the pressure is the free variable. The solution was obtained iteratively, adjusting the pressure for GEM calculations until the equilibrated system's volume is consistent with the desired volume (cf. supplement figure S1).

Temporal changes in the borehole composition were implemented by adding or removing components from the gas or aqueous phase during each time step, which changes the bulk composition in the borehole. For interaction with the surrounding rock, dissolved gases and other dissolved species were removed or added. Water and dissolved species were added to mimic the inflow of Opalinus Clay water.

It is assumed that hydrogen consumption in the borehole is mainly controlled by hydrogen-fueled, microbial mediated sulfate reduction (Appelo and Vinsot, 2012; Bagnoud, Chourey, et al., 2016; Bagnoud, Leupin, et al., 2016; Boylan et al., 2019; Leupin et al., 2017; Vinsot et al., 2014, 2017).

The implemented sulfate reduction mimics the kinetic model in the technical note by Appelo and Vinsot (2012). Boylan et al. (2019) proposed that bisulfide (HS^-) produced from sulfate reduction reacts with iron (II) (Fe^{2+}) to form iron (II) sulfide (FeS). Sulfate and iron (II) concentration might be controlled by the celestite and siderite solubility, respectively. The overall reaction can be written as:

$$SO_4^{2-} + Fe^{2+} + 4H_2 = FeS + 4H_2O \quad (5)$$

The temporal rate of reduction of sulfate is described by Michaelis-Menten kinetics following Appelo and Postma (2005)

$$\mathrm{d}\frac{m_{SO_4^{2-}}}{\mathrm{d}t} = -\mu_{max}\frac{B}{Y} \quad (6)$$

which is coupled to the temporal change of bacterial mass B

$$\frac{\mathrm{d}B}{\mathrm{d}t} = Y\frac{d\, m_{SO_4^{2-}}}{\mathrm{d}t} - k_{Bd}B \quad (7)$$

t is the time (s), m is the molal concentration of sulfate and hydrogen (mol/kg), $\mu_{max} = 2.0 \times 10^{-6}$ is the specific degradation rate (s^{-1}), B is the molal concentration of the actual biomass (mol/kg) with an initial value of $B = 1.0\times10^{-5}$ (mol). $Y = 0.2$ is the yield factor (mol B/mol SO_4^{2-}), $k_{1/2} = 1.0\times10^{-5}$ are the half-rate constants for sulfate and hydrogen (mol/kg), and $k_{Bd} = 1.0\times10^{-9}$ is the first order death rate of the bacteria (s^{-1}). All parameter values were taken from Appelo and Vinsot (2012).

Thermodynamic setup

A unified setup in terms of (aqueous, gas and mineral) phases definition based on the Thermoddem DB (Blanc et al., 2012) has been used for all the thermodynamic calculations. Some of the gases of interest, specifically N_2 and CH_4, are considered non-reactive species in the gas and water phases. Reactivity of H_2 in the borehole is kinetically controlled, as described further before.

It should be stressed that the same thermodynamic setup is used throughout both models, equilibration of the borehole and the reactive transport in Opalinus Clay, and the models rely on the GEM algorithm to determine the stability and composition of phases from the bulk elemental composition provided.

The thermodynamic setup for reactive transport in Opalinus Clay is based on the Opalinus Clay pore-water model by Pearson et al. (2011). The pore-water composition is described by equilibria with mineral phases and a cation exchanger phase, which is implemented as an ideal solid solution phase following Kulik (2010). Pearson et al. (2011) compared Opalinus Clay water models that use different combinations of clay minerals for equilibration. Scoping

calculations suggest that the combination of kaolinite, illite, celestite, dolomite, and some minor amount of daphnite reported in Table 2 gives a pore-water composition that is close to the measured composition of HT borehole water in terms of major cations and sulfate concentrations. A comparison of water properties between the model and those used by Vinsot et al. (2014) is given in Table 3. More details on the setup are provided in the Supplement.

The initial phase composition and the major speciation of the aqueous water phase in the injection interval are given Table 2 and Table 3, respectively. The system was created by equilibrating the equivalent of 10 ml of Opalinus Clay with 0.16 litres of Opalinus Clay pore-water and 6.762 mol of argon. The system's total initial volume is about 9.6 litres, of which about 9.4 litres is occupied by the gas phase, 0.16 litre by the liquid phase, and 0.04 litres by several mineral phases. Also, siderite was added to be the source of iron for microbial mediated sulfate reduction during hydrogen injection.

This setup gives an equilibrated borehole water composition different to that of the Opalinus Clay pore-water, as the liquid to solid ratio is much higher in the borehole. It was assumed that celestite is not present in the borehole to avoid a buffering of sulfate concentrations during microbial-mediated hydrogen consumption.

Mass transfer across the borehole wall

The borehole wall is a boundary between two compartments: a gas-filled injection interval and a water-saturated claystone. The models for both compartments are coupled via the concentration boundary condition for the transport model based on the thermodynamic model for the injection interval, while the mass balance in the injection interval model was adjusted based on the mass fluxes across the borehole wall provided by the transport model.

Diffusive fluxes of dissolved gases between the Opalinus Clay and the injection interval are calculated based on the concentration gradients in Opalinus Clay at the interfaces, which are available from the solution of the transport equations. Positive fluxes are directed from the

borehole towards the Opalinus Clay, whereas negative fluxes are directed from the Opalinus Clay towards the borehole.

The borehole wall is assumed to be covered by a thin water film (Figure 2) which is in equilibrium with the gas phase in injection interval. It is assumed that biochemical reactions (i.e., hydrogen consumption by sulfate-reducing microorganisms) occur preferably in the liquid film at the borehole wall, and then the induced changes in species and gas concentrations set a concentration boundary condition at the surface of Opalinus Clay.

Gravitational forces let the water in the film flow downward towards the bottom of the borehole, where water accumulates. The water film is replenished by the (slow) advective flux of pore-water from claystone towards the borehole due to a time-varying large-scale pressure gradient. Mixing of water at the walls with water at the bottom is fast, therefore thermodynamic equilibrium of all water and gas in the injection interval is assumed in the models.

Other processes related to experimental setup and operation

The model also includes other processes or events to improve the description of experimental measurements. Gas leakage from the circulation system and an air ingress, due to a major leak in the gas circulation system, are the most relevant.

The gas circulation module is not entirely tight and showed gas losses of about 0.022 mmol/day during a tightness test with 2.2 vol% hydrogen in an argon atmosphere for a 0.11 MPa pressure. This rate corresponds to a pressure drop of 0.05 kPa/day and a volume loss of 0.5 ml gas/day at standard temperature and pressure conditions (Vinsot et al., 2014). Observed relative gas losses for hydrogen were significantly higher than for argon. In line with the observations, this gas loss was implemented as a diffusive process, i.e., fractions of total losses for each gas were calculated based on a weighted average with mole fraction and gas diffusion coefficients from Table 1 as weights. In addition, the total loss rate was linearly scaled at each time step by the

quotient between the actual borehole pressure and a reference pressure of 0.11 MPa, the pressure at which the tightness test for the gas circulation module was conducted.

The experiment was also affected by two leakage events. The first occurred during the calibration phase after 1.5 years (day 565), and, more importantly, the second one (day 1233) took place during the semi-continuous injection of hydrogen. After the first leak, the borehole was flushed with Argon. The second event is characterized by a sudden measured increase in nitrogen and drop of argon concentrations in the borehole gas phase. There was no pressure change recorded that can be associated with this event. For this event, a total of 45 mmol of gas had to be removed from the borehole model. As for the gas losses from the gas module, single gas fractions were removed accordingly to their mole/fraction and gas diffusion coefficients. The total removed amount was replaced with nitrogen to maintain the borehole pressure.

Results and discussion

Despite the apparent simplicity of the HT experimental setup, the gas pressure measurements, water height, gas phase and water compositions in the injection interval showed a rather complex evolution over the monitoring period. This is reflected in short and long-term variations in gas fluxes across the borehole wall (cf. Figure S3 in the supplement). Accordingly, the underlying model matching the experimental observations has to include several coupled processes and becomes quite complex. This section discusses the modelling results based on the major processes: (a) the influx of rock pore-water and the amount of water in the borehole; (b) the diffusion of dissolved gases in the claystone; and (c) the biochemical reaction in the borehole.

Influx of rock pore-water

The pressure evolution in the injection interval and estimation of the inflow variation of Opalinus Clay water are shown in Figure 3. No water extraction from the borehole occurred during the first experimental phase (the calibration phase). The inflow of water caused an increase in gas pressure during the first year. Several hundred mL of water were removed manually from the borehole to prevent the water level from rising above the injection line in the test interval (see appendix of Vinsot et al., 2017). After one year (day 364), an automatic water extraction system was installed. The water height in the injection interval (middle part of Figure 3) was estimated from the differential pressure between gas pressure and water pressure measured 'below' the borehole. It was attempted to keep the water height constant below 10 cm by extracting the water. Due to the complex geometry of the inclined injection interval (Figure 1) with build-in infrastructure, there is no simple translation from water height into water volume. Based on the geometric considerations, on average, 1 cm height should correspond to about 37 ml. For water levels close to the bottom, i.e. below 6 cm, the tilt will change the relation to about 14 ml/cm. Figure 3 shows that the water height varies significantly during the first calibration phase of the experiment. Within this study, inflow rates were varied manually to fit pressure evolution, which averaged at 11-12 ml/day.

With the start of the first hydrogen injection phase after day 778, the water height stabilises, and water inflow is estimated at about 18 ml/day, assuming that the long-term average water extraction rate corresponds to the average water inflow rate.

A closer look at the daily extraction rates reveals quite some scatter in the recorded data points caused by the measurement error for the water height. The water height in the borehole is calculated from the difference of two pressure measurements, each with a measurement accuracy of 0.4 kPa (or a pressure head of 4 cm H_2O). Averaging the extraction rates with a rolling average shows systematic changes in the extraction rate that are not reflected in the

water height. Therefore, a piecewise linear fit of the cumulative water extraction curves slope was applied, and the fitted extraction rates were used as inflow rates in the model. The measured extraction rates were adopted in the model as is, except if the borehole's water volume would go below the lower limit of 380 ml (equivalent to about 14 cm), then the extraction rates were adjusted to zero. This reference approach to calculate water inflow and extraction is indicated as model V in Figure 4.

The alternative model (model C in Figure 4) assumed a constant inflow of 11 ml/day up to day 556, then 12 ml/day up to day 780 and 18 ml/day with a constant extraction of 18 ml/day after day 780. This implementation ensures a constant residual water level in the borehole and results in a constant water exchange rate.

Figure 4 shows for both extraction approaches the results for pressure, water in the borehole, and evolution of the gas phase in the borehole system. The calculated pressure evolution in the upper left figure panel reproduces the experimental values quite well, although there is a general tendency to overestimate pressure slightly. The use of measured water extraction rates in model V can reproduce some short-term pressure variations, while model C with constant extraction seems to fit slightly better the long-term average pressure evolution. Despite the difference between models V and C the influence on long term gas composition in the borehole is small.

The most recognizable effects on gas pressure are related to the time-varying inflow of rock pore-water into the borehole combined with time-varying water extraction. Furthermore, pressure variations are caused by in- and out-diffusion of gases between the borehole and the adjacent claystone, a long-term small gas loss from the gas circulation system, the injection of hydrogen, or experimental manipulations, like the addition of evacuated sampling cylinders, for example at day 2059, which causes a visible pressure decrease. Gas leakage from the gas circulation system has a significant influence on the long–term pressure evolution in the

borehole. For example, a relatively small increase of this gas loss rate can easily compensate for model's V high pressure values.

Diffusion of gases in claystone

The evolution of gas composition could be reproduced very well for major components like argon, methane, and nitrogen (Figures 4 and 5). The effective diffusion coefficients for dissolved gases in Opalinus Clay were fitted by adjusting manually the geometry factor (Equation 2) in terms of the exponent *m*. The effect of changing the geometry factor for Opalinus Clay rock is demonstrated in Figure 5.

A value of m=1.2 would result in an effective diffusion coefficient for HTO (tritiated water tracer) of 3.1×10^{-11} m/s (D_w=2.24×10^{-9}, ϕ=0.16, m=1.2, T=15.6 ºC). This value for D_e is slightly lower than values obtained in previous studies. A temperature corrected value of $4.2 \pm 0.4 \times 10^{-11}$m/s ($\phi$=0.17, T=14 ºC) was measured for HTO in samples from Mont Terri in the laboratory (Van Loon, Soler, et al., 2004) and the value of 4.0×10^{-11} m/s (ϕ=0.15, T=14 ºC) was estimated within an in situ migration experiment (Van Loon, Wersin, et al., 2004).

In addition to the reference model with m=1.2, Figure 5 contains results for m=1.0 (D_{e_high} = 4.4×10^{-11} m/s for HTO) and for m=1.4 (D_{e_low} = 2.1×10^{-11} m/s for HTO). Generally, the results for the two additional model variants envelope the experimentally measured concentrations in the gas phase. In terms of gas composition, the agreements for D_{e_low} (m=1.4) seems better during the calibration phase, but for this variant, gas pressures in the borehole drop significantly below the measured values in the long term. Besides, the volume fractions and the in-diffusion of methane and nitrogen are underestimated in the long term. Visualizations of typical concentration profiles for dissolved gases are included in the supplement.

On day 566, hydrogen injection experiments started. In a first test, the injection interval was filled with a mixture of 85 vol% argon, 5 vol% neon (Ne), 5 vol% Helium (He), and 5 vol %

hydrogen. The model considered that the measured fractions for neon, helium and hydrogen in the injection interval were about 20% less.

It was supposed that neon and helium behave like other inert gases, but unexpectedly, their concentrations dropped faster than expected (Figure 4). The effective diffusion coefficients for helium and neon in Opalinus Clay are not high enough to explain the concentration decrease, which occurs only during the injection of hydrogen. After the first hydrogen injection is finished, the slopes of modelled and measured concentration curves for helium and neon seem to match. This indicates that the implemented diffusion model and its parameterisation describes the long-term behaviour of these two gases.

The exact reason for the high losses of helium and neon during hydrogen injection phases is not clear, but might be related to the hydrogen injection procedure. In addition, there is a considerable jump in neon and helium concentrations during the gas leak event after 1233 days. This event has been included in the modelling process, and fractions of both gases were removed from the borehole in relation to their partial pressures and gas diffusion coefficients.

Figure 4 and 6 show the trace amount of carbon dioxide in the gas phase. For the implemented thermodynamic setup, the carbon dioxide concentration in borehole water and in the inflowing water from Opalinus Clay are mainly controlled by equilibrium with minerals and only to a very small degree by diffusive transport in Opalinus Clay.

Hydrogen injection variations

The fast drop in the hydrogen concentration relative to the one of inert gases (Figures 4) was interpreted as the consequence of its consumption by microbial mediated sulfate reduction (Vinsot et al., 2014). On day 926, a single hydrogen injection (14.3 mmol) was performed. On the day 1038, a semi-continuous hydrogen injection started. The hydrogen content in the borehole gas phase was kept constant at a value of about 4.5 % until day 1553. A second hydrogen injection phase started on day 2402, in which several hydrogen pulse injections were

conducted over about two months. Then, on day 2485, a semi-continuous injection started, which set the hydrogen fraction in the gas phase to about 2 % and, after about 1.5 years, on day 3029, the injection was stopped. It should be noted, that measured hydrogen concentrations are comparable between online measurements and gas samplings, although the online measurements (H_2 specific sensor and Raman spectroscopy) where more affected by rapid fluctuations (Vinsot et al., 2014).

Figure 6 (upper left sub-figure) compares measured and modelled hydrogen volume fractions for different model variants. In Equation 7, the growth of biomass is only dependent on the availability of sulfate. Model variants not shown in the paper indicated that the growth of biomass must be limited to a value of B_{max} in order to be able to model the semi-steady state hydrogen injection. Six model variants have been considered which differ in the maximum biomass amount in the borehole (low: 1.0×10^{-3} mol, reference: 1.2×10^{-3} mol and high: 1.5×10^{-3} mol) and in the way the hydrogen is injected during the semi-continuous injection phases (mimic experimentally recorded injection rates or constant rate injection).

The amount of biomass directly controls the hydrogen consumption rate (see Equations 6 and 7). A smaller amount of biomass causes a slower decrease of hydrogen during the pulse injection of hydrogen gas and gives higher hydrogen concentrations during the semi-continuous hydrogen injection. In addition to hydrogen consumption in the injection interval, hydrogen can also diffuse into the claystone. The model assumes that microbes do not consume hydrogen in the claystone because the microbes are expected to be in a dormant state and have only a reduced activity due to limited space and nutrients availability (Leupin et al., 2017). The predicted diffusive hydrogen flux into the adjacent claystone was at least two orders of magnitude slower than the microbial hydrogen consumption rate for all model variants considered in the study.

In general, all model variants reproduce the fast reduction of hydrogen concentrations during pulse injection. In addition, the slight pressure change during hydrogen injection into the borehole is well reproduced. The models with lower biomass amounts fit the measured behaviour better, while the models with high biomass amounts result in too fast hydrogen consumption.

The hydrogen consumption rate directly depends on sulfate availability as implemented via the Michaelis-Menten rate law (Equations 6) and is directly reflected in the amount of mackinawite precipitated and the dissolved amount of siderite (Figure 6, lower left panel). The consumption of sulfate during hydrogen consumption is visible in the drop of sulfate concentrations measured and modelled in the borehole water after the beginning of the hydrogen injection phase (Figure 6, middle left panel). The decrease in sulfate concentration depends not only on the hydrogen consumption rates, but also on the amount of water in the borehole related to the sulfate replenishment due to inflow and water extraction.

The water pH in the borehole did not change considerably, although the sulfate reduction should cause acidification of borehole water. In the models, the pH and alkalinity were buffered mainly by the precipitation of dolomite and calcite (Figure 6, lower left panel), which is also reflected in variations of measured and modelled calcium and magnesium concentrations (Figure 6 middle right panel). There is a slight increase in calculated CO_2 in the gas phase, as the CO_2 released by siderite dissolution is not entirely consumed by carbonate formation (Figure 6, lower right panel).

In the reference model, the hydrogen injection is implemented as a time-dependent rate function extracted from the experimental recordings of injection rate. The injection rates show fluctuations that are reflected in the hydrogen concentrations during the semi-continuous hydrogen injection. The trend of experimental data from samples and online hydrogen measurements are not explicitly reproduced during the first semi-continuous injection period.

An alternative approach using constant injection rates ensures a constant hydrogen concentration (molar fraction) in the borehole's gas phase. The modelled injection rates were manually adjusted such that the cumulative injected amount matches the experimental values.

For both injection approaches, the differences between the observed and modelled system response to the hydrogen injection are related to the injection protocol. For hydrogen injection into the gas circulation system, the injection device contains a 540 ml syringe filled with hydrogen. During the initial ten days of continuous injection, the hydrogen injection rate was adapted manually to keep a constant 6 kPa (60 mbar) hydrogen partial pressure in the circulation system. Nevertheless, during this time, a valve stayed open, causing hydrogen gas to diffuse out of the syringe and into the gas circulation system, while other gases could enter the syringe and dilute the hydrogen inside. Specifically, helium and neon could accumulate in the syringe during continuous and semi-continuous operation due to their high gas diffusion coefficients. This could explain part of the unexpected fast concentration drop for these gases. As a consequence of the open valve, the apparent hydrogen injection rates measured (based on injected hydrogen volume) are too low for a hydrogen filled syringe, while they would be too high for a syringe in which the gas composition inside of the syringe is already nearly equilibrated with the gas circulation system.

After the ten days period, an automatic semi-continuous injection procedure was used. This operation phase alternates short time intervals (minutes to hours) during which hydrogen was injected at a constant flow rate with more extended periods without injection.

Conclusions

This work presents a new conceptualisation and numerical analysis of the experimental data collected in the HT experiment at the Mont Terry Underground Rock Laboratory. More specifically, the developed model describes the relationships between transport processes and

chemical interactions controlling the behaviour of gases, both inert and reactive, in the borehole and the surrounding Opalinus Clay.

The modelling study demonstrates that fluxes of non-reactive gases across the borehole wall are constrained mainly by diffusive transport of dissolved gases in Opalinus Clay. The borehole's gas phase composition and pressure control the gas concentrations in the water phase at the borehole's wall in the injection interval. The concentration difference in the water phase between the borehole wall and Opalinus Clay is the driving force for the diffusive transport of gases in the Opalinus Clay. Advective transport of dissolved gases is of secondary importance, while water accumulation in the borehole increases the borehole gas pressure.

The model supports the hypothesis that sulfate reducing microbes consume hydrogen. The experimentally observed reduction in sulfate concentrations is controlled by an interplay between the inflow of Opalinus Clay water, water extraction by the automatic extraction system and the rate of sulfate reduction. Sulfate reduction should cause acidification of borehole water. The pH, borehole's water composition and CO_2 in the gas phase are buffered by mineral reactions, showing a good agreement between the thermodynamic model and the experimental observations.

Given the uncertainty in estimating the pore-water composition in tight claystone (Gaucher et al., 2009; OECD, 2000; Wersin et al., 2009) and the time variability in sampled pore-waters from the HT experiment (Vinsot et al., 2017), the presented model reproduces well the observed behaviour major gases and sulfate concentrations in the water phase.

Large-scale hydraulic gradients control water flow magnitude near the HT borehole. Little is known about the hydraulic conditions in the vicinity of the HT borehole. Specifically it is not clear, if water entering the HT borehole is confined to specific layers or spots with higher permeability, or if water flow is more or less homogeneously distributed from all directions into the borehole. The characteristics of water inflow might have consequences for

representation the Opalinus Clay pore-water in thermodynamic models and the basic underlying modelling assumption that the borehole can be treated as a homogeneous well-mixed system.

Acknowledgments

We kindly acknowledge funding for this work from the Mont Terri consortium under contract No. 2017/170.

Data and code availability

The data and the code that support the findings of this study are openly available in Zenodo at doi:10.5281/zenodo.5344121. The numerical model was run with FEniCS V2019.1.0 (https://www.fenicsproject.org) and xGEMS V-JUL2020 (https://bitbucket.org/gems4/xgems.git).

References

Alexander, W.R. and McKinley, L.E. (2007), *Deep Geological Disposal of Radioactive Waste*, Elsevier, Amsterdam Boston.

Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., et al. (2015), "The FEniCS project version 1.5", *Archive of Numerical Software*, Vol. 3 No. 100, available at:https://doi.org/10.11588/ANS.2015.100.20553.

Appelo, C.A.J. and Postma, D. (2005), *Geochemistry, Groundwater and Pollution*, A.A.Balkema Publishers, Amsterdam, Netherlands.

Appelo, C.A.J. and Vinsot, A. (2012), *HT(Hydrogen Transfer) Experiment: Model for Transport of Gases Ar, H_2, He, Ne, N_2, CO_2 and Alkanes with Reactions for Hydrogen in Opalinus Clay*. Mont Terri Project, St. Ursanne, Switzerland, Mont Terri Technical Note TN 2012-58.

Bagnoud, A., Chourey, K., Hettich, R.L., de Bruijn, I., Andersson, A.F., Leupin, O.X., Schwyn, B., et al. (2016a), "Reconstructing a hydrogen-driven microbial metabolic network in Opalinus Clay rock", *Nature Communications*, 7(1): article 12770, https://doi.org/10.1038/ncomms12770

Bagnoud, A., Leupin, O.X., Schwyn, B. and Bernier-Latmani, R. (2016b), "Rates of microbial hydrogen oxidation and sulfate reduction in Opalinus Clay rock", *Applied Geochemistry*, Pergamon, Vol. 72, pp. 42–50. https://doi.org/10.1016/j.apgeochem.2016.06.011

Benson, S.M. and Cole, D.R. (2008), "CO_2 sequestration in deep sedimentary formations", *Elements*, Vol. 4 No. 5, pp. 325–331, https://doi.org/10.2113/gselements.4.5.325

Blanc, P., Lassin, A., Piantone, P., Azaroual, M., Jacquemet, N., Fabbri, A. and Gaucher, E.C. (2012), "Thermoddem: A geochemical database focused on low temperature water/rock interactions and waste materials", *Applied Geochemistry*, Vol. 27 No. 10, pp. 2107–2116, https://doi.org/10.1016/j.apgeochem.2012.06.002

Boylan, A.A., Perez-mon, C., Guillard, L., Burzan, N., Loreggian, L., Maisch, M., Kappler, A., et al. (2019), "H_2-fuelled microbial metabolism in Opalinus Clay", *Applied Clay Science*, Elsevier, Vol. 174 No. March, pp. 69–76, https://doi.org/10.1016/j.clay.2019.03.020.

Crotogino, F. (2016), "Larger Scale Hydrogen Storage", *Storing Energy*, Elsevier, Amsterdam, the Netherlands, pp. 411–429.

Dimian, A.C., Bildea, C.S. and Kiss, A.A. (2014): Generalised Computational Methods in Thermodynamics, in: Computer Aided Chemical Engineering. Elsevier B.V., Amsterdam, the Netherlands, pp. 157–200. https://doi.org/10.1016/B978-0-444-62700-1.00005-X

Engineering Toolbox. (2008), "Diffusion Coefficients of Gases in Water", available at: https://www.engineeringtoolbox.com/diffusion-coefficients-d_1404.html (accessed 16 July 2020).

Gaucher, E.C., Tournassat, C., Pearson, F.J., Blanc, P., Crouzet, C., Lerouge, C. and Altmann, S. (2009), "A robust model for pore-water chemistry of clayrock", *Geochimica et Cosmochimica Acta*, Elsevier Ltd, Vol. 73 No. 21, pp. 6470–6487, https://doi.org/10.1016/j.gca.2009.07.021

Gaus, I., Azaroual, M. and Czernichowski-Lauriol, I. (2005), "Reactive transport modelling of the impact of CO2 injection on the clayey cap rock at Sleipner (North Sea)", *Chemical Geology*, available at:https://doi.org/10.1016/j.chemgeo.2004.12.016.

Gimmi, T., Leupin, O.X., Eikenberg, J., Glaus, M.A., Van Loon, L.R., Waber, H.N., Wersin, P., et al. (2014), "Anisotropic diffusion at the field scale in a 4-year multi-tracer diffusion and retention experiment – I: Insights from the experimental data", *Geochimica et Cosmochimica Acta*, Vol. 125, pp. 373–393, , https://doi.org/10.1016/j.gca.2013.10.014

Grunau, H.R. (1987), "A worldwide look at the cap-rock problem", *Journal of Petroleum Geology*, Vol. 10 No.

3, pp. 245–265, https://doi.org/10.1111/j.1747-5457.1987.tb00945.x

IAEA. (2011), *Disposal of Radioactive Waste.*, IAEA safety standards series, ISSN 1020–525X ; no. SSR-5, *IAEA Safety Standards Series, SSR-5*, International Atomic Energy Agency, Vienna, available at: http://www-pub.iaea.org/MTCD/Publications/PDF/Pub1449_web.pdf (accessed 21 June 2017).

Kluyver, T., Ragan-kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., et al. (2016), "Jupyter Notebooks—a publishing format for reproducible computational workflows", *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pp. 87–90.

Kulik, D.A. (2010): Geochemical thermodynamic modelling of ion partitioning, in: Ion Partitioning in Ambient-Temperature Aqueous Systems. Mineralogical Society of Great Britain & Ireland, London, pp. 65–138. https://doi.org/10.1180/EMU-notes.10.3

Kulik, D.A., Wagner, T., Dmytrieva, S. V., Kosakowski, G., Hingerl, F.F., Chudnenko, K. V. and Berner, U.R. (2012), "GEM-Selektor geochemical modeling package: revised algorithm and GEMS3K numerical kernel for coupled simulation codes", *Computational Geosciences*, Vol. 17 No. 1, pp. 1–24, https://doi.org/10.1007/s10596-012-9310-6

Leupin, O.X., Bernier-Latmani, R., Bagnoud, A., Moors, H., Leys, N., Wouters, K. and Stroes-Gascoyne, S. (2017), "Fifteen years of microbiological investigation in Opalinus Clay at the Mont Terri rock laboratory (Switzerland)", *Swiss Journal of Geosciences*, Springer International Publishing, Vol. 110 No. 1, pp. 343–354, https://doi.org/10.1007/s00015-016-0255-y

OECD. (2000), *Porewater Extraction from Argillaceous Rocks for Geochemical Characterisation Methods and Interpretation*, Radioactive waste management, ISSN 1990-0325, edited by "Organisation for Economic Co-operation and Development", Nuclear Energy Agency, Organisation for Economic Co-operation and Development, available at: https://doi.org/10.1787/19900325.

Parkhurst, D.L. and Appelo, C.A.J. (2013), "Description of Input and Examples for PHREEQC Version 3 — A Computer Program for Speciation , Batch-Reaction , One-Dimensional Transport , and Inverse Geochemical Calculations. U.S. Geological Survey Techniques and Methods, book 6, chapter A43, 497 p.", *U.S. Geological Survey Techniques and Methods, Book 6, Chapter A43*, pp. 6-43A.

Pearson, F., Tournassat, C. and Gaucher, E.C. (2011), "Biogeochemical processes in a clay formation in situ experiment: Part E – Equilibrium controls on chemistry of pore water from the Opalinus Clay, Mont Terri Underground Research Laboratory, Switzerland", *Applied Geochemistry*, Vol. 26 No. 6, pp. 990–1008 , https://doi.org/10.1016/j.apgeochem.2011.03.008

Rumble, J.R. (Ed.). (2019), *CRC Handbook of Chemistry and Physics*, 100th Edit., CRC Press/Taylor & Francis, Boca Raton, FL.

Selley, R.C. and Sonnenberg, S.A. (2014), *Elements of Petroleum Geology: Third Edition*, *Elements of Petroleum Geology: Third Edition*, Elsevier Inc., available at: https://doi.org/10.1016/C2010-0-67090-8.

Van Loon, L.R., Soler, J.M., Müller, W. and Bradbury, M.H. (2004), "Anisotropic Diffusion in Layered Argillaceous Rocks: A Case Study with Opalinus Clay", *Environmental Science & Technology*, Vol. 38 No. 21, pp. 5721–5728, https://doi.org/10.1021/es049937g

Van Loon, L. R., Müller, W. and Iijima, K. (2005). Activation energies of the self-diffusion of HTO, $^{22}Na^{+}$ and $^{36}Cl^{-}$ in a highly compacted argillaceous rock (Opalinus Clay). *Applied Geochemistry*, 20 (5), pp. 961–972, https://doi.org/10.1016/j.apgeochem.2004.10.007

Van Loon, L.R., Wersin, P., Soler, J.M., Eikenberg, J., Gimmi, T., Hernán, P., Dewonck, S., et al. (2004), "In-situ diffusion of HTO, $22Na^{+}$, Cs^{+} and I^{-} in Opalinus Clay at the Mont Terri underground rock laboratory", *Radiochimica Acta*, Vol. 92 No. 9–11, pp. 757–763, https://doi.org/10.1524/ract.92.9.757.54988

Vinsot, A., Appelo, C.A.J., Lundy, M., Wechner, S., Lettry, Y., Lerouge, C., Fernández, A.M., et al. (2014), "In situ diffusion test of hydrogen gas in the Opalinus Clay", *Geological Society, London, Special Publications*, Vol. 400 No. 1, pp. 563–578, https://doi.org/10.1144/SP400.12

Vinsot, A., Appelo, C.A.J., Lundy, M., Wechner, S., Cailteau-Fischbach, C., de Donato, P., Pironon, J., et al. (2017), "Natural gas extraction and artificial gas injection experiments in Opalinus Clay, Mont Terri rock laboratory (Switzerland)", *Swiss Journal of Geosciences*, Vol. 110 No. 1, pp. 375–390, https://doi.org/10.1007/s00015-016-0244-1

Wagner, T., Kulik, D.A., Hingerl, F.F. and Dmytrieva, S. V. (2012), "GEM-SELEKTOR geochemical modeling package: TSolMod library and data interface for multicomponent phase models", *The Canadian Mineralogist*, Mineralogical Association of Canada, Vol. 50 No. 5, pp. 1173–1195 , https://doi.org/10.3749/canmin.50.5.1173

Wersin, P., Gaucher, E.C., Gimmi, T., Leupin, O.X., Mäder, U.K., Pearson, F.J., Thoenen, T., et al. (2009), *Geochemistry of Pore Waters in Opalinus Clay at Mont Terri: Experimental Data and Modelling*, Mont Terri Project, St. Ursanne, Switzerland. Mont Terri Technical Report 2008-06. See https://www.mont-terri.ch/content/mont-terri-internet/de/documentation/free-reports.download/mont-terri-internet/de/documents/technical-reports/TR2008-06.pdf (accessed 20/09/2021)

xGEMS (2021). See https://bitbucket.org/gems4/xgems.git (accessed 08/09/2021).

Tables

Table 1: Gas diffusion coefficients at 20 °C and 101.325 kPa (Rumble, 2019) *and diffusion coefficients in water at 25 °C* (Appelo and Vinsot, 2012*; Engineering Toolbox, 2008)*

Species	**Gas diffusion coefficient ($\times 10^{-4}$ m^2/s)**	**Binary gas mixture for gas diffusion coefficient**	**Diffusion coefficient in water ($\times 10^{-9}$ m^2/s)**
H_2O	0.242	Large excess of air	2.24
H_2 (Hit_2)	0.794	Equimolar mixture with Ar	5.13
He	0.726	Equimolar mixture with Ar	7.29
N_2	0.190	Equimolar mixture with Ar	1.96
Ne	0.313	Equimolar mixture with Ar	4.04
Ar	0.189	Large excess of air	2.45
CO_2	0.148	Equimolar mixture with Ar	1.92
CH_4 ($CitH_4$)	0.208	Equimolar mixture with N_2	1.85
O_2	0.187	Equimolar mixture with Ar	2.42
“default”	0.189	Identical to Ar	2.24

Table 2: Equilibrated Opalinus Clay system in terms of the liquid phase, gas phase, cation exchanger phase (OpaExchanger), and mineral phases. The total volume of the Opalinus Clay system is normalised to 1m^3. The numbers were rounded for better readability.

		Opalinus Clay		**Borehole**	
Phase name		**Phase volume: cm^3**	**Phase mass: g**	**Phase volume: cm^3**	**Phase mass: g**
Liquid phase		160	162	160	162
Gas phase		0.0	0.0	9402.3	39.087
	Mole fraction			Ar: 0.9985 CO_2: 3.0×10^{-4} $CitH_4$: 2.6×10^{-5} H_2: 9.6×10^{-4} He: 1.0×10^{-8} Hit_2: 5.3×10^{-9} N_2: 0.00021 Ne: 1.0×10^{-8} O_2: 0.0	
Cation exchange phase for Opalinus Clay: "Opa"		0.044	7.34	2.7×10^{-7}	5.1×10^{-5}
	Amount of Opa: meq	0.272		2×10^{-6}	
	Mole fraction	Opa_2Ca: 0.051 Opa_2Fe: 0.00051		Opa_2Ca: 0.046 Opa_2Fe: 4.5×10^{-4}	

		OpaK: 0.344 Opa_2Mg: 0.043 OpaNa: 0.560 Opa_2Sr: 0.0015		OpaK: 0.25 Opa_2Mg: 0.038 OpaNa: 0.663 Opa_2Sr: 2.3×10^{-5}	
Chlorite CCca-2		3.2	9.37	0.08	0.22
Illite(Mg)		268.0	750.5	2.7	7.7
Kaolinite		213.0	553.6	49.7	129.1
Biomass		1.00×10^{-8}	1.00×10^{-08}	1×10^{-10}	1×10^{-10}
Dolomite		0.42	1.2	0.041	0.12
Calcite		167.6	454.2	1.7	4.5
Siderite		10.5	41.5	32.3	127.3
Pyrite		0.076	0.38	0*	0*
Mackinawite (FeS)		0	0	0*	0*
Celestite		6.027	23.9		
Quartz(alpha)		171.6	454.5	11.1	29.5

* The precipitation/occurrence of pyrite and mackinawite in the borehole were suppressed. The amount of mackinawite in the borehole is kinetically controlled during injection of hydrogen.

Table 3: Comparison of pore-water properties and composition from Vinsot et al. (2014) and this work.

Water analysis from 5. September 2012 (Table 2 of Vinsot et al. (2014))		**Modeled values for Opalinus Clay (this work, 0.22 MPa)**		**Initial setup for borehole system (0.25 MPa)**
Lab pH	6.9	pH	7.6	7.8
Lab temperature: °C	20	Temperature: °C	15.6	15.6
log(pCO_2)	-2.0	log(pCO_2)	-2.99	-3.12
Species	Concentration: mmol/l	Species/element	Concentration: mmol/kg	Concentration: mmol/kg
Na	255	Total Na	254.0	247.1
K	1.47	Total K	5.8	3.5
NH_4	0.524		-	
Mg	18.6	Total Mg	15.5	8.8
Ca	16.3	Total Ca	17.1	9.5
Sr	0.369	Total Sr	6.1	6×10^{-3}
Cl	295	Total Cl	291.4	286.1
TIC	2.5	Total C	1.22	1.25

SO_4	14.4	Total S	17.0	16.3
		Total N	4.4	8.1×10^{-4}
		Non-reactive CH_4	 0.3	1.1×10^{-4}

Figures

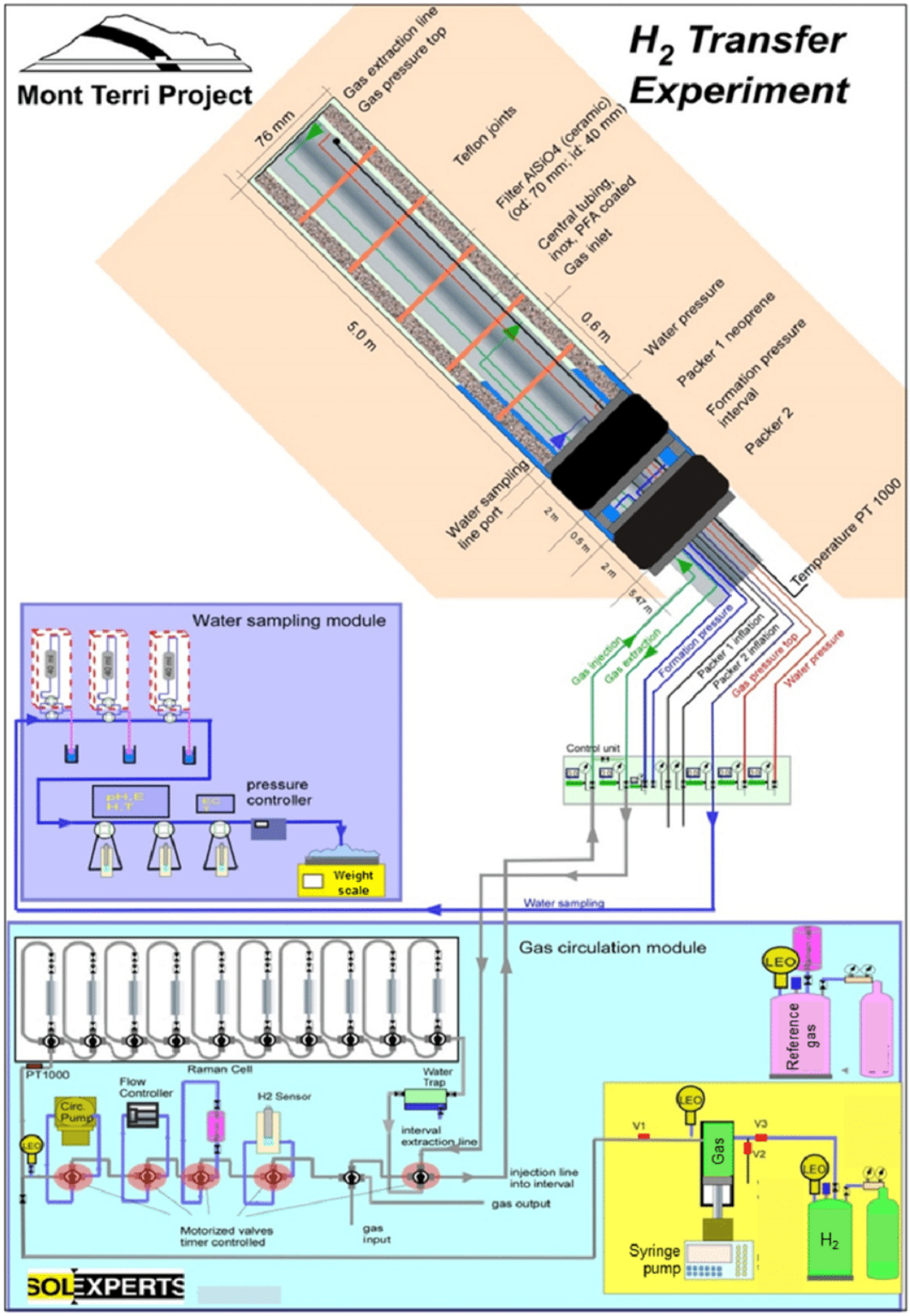


Figure 1: Experimental layout of the Mont Terri HT experiment. (from Vinsot et al., 2014, Fig.1).

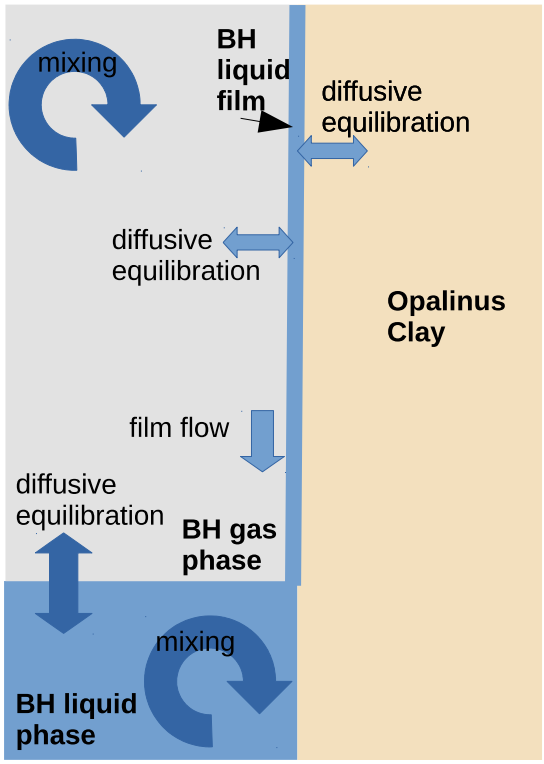


Figure 2: Conceptual description of processes controlling the gas and water equilibria in the borehole.

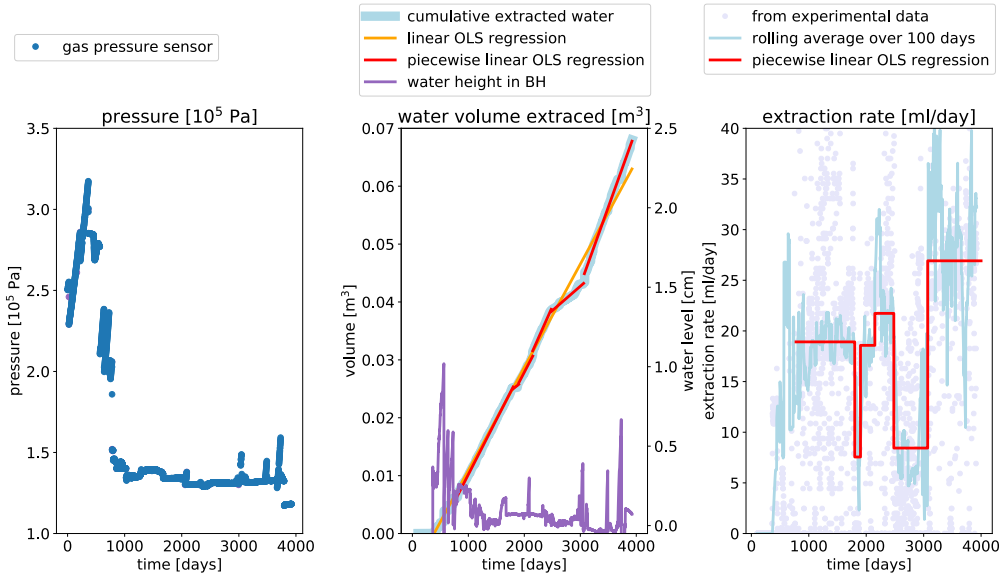


Figure 3: Gas pressure evolution (left), cumulated extracted water and water height (middle) in the borehole, and approximation of water inflow rates by fitting cumulated water extraction (middle and right) after start of the hydrogen injection phase (day 778).

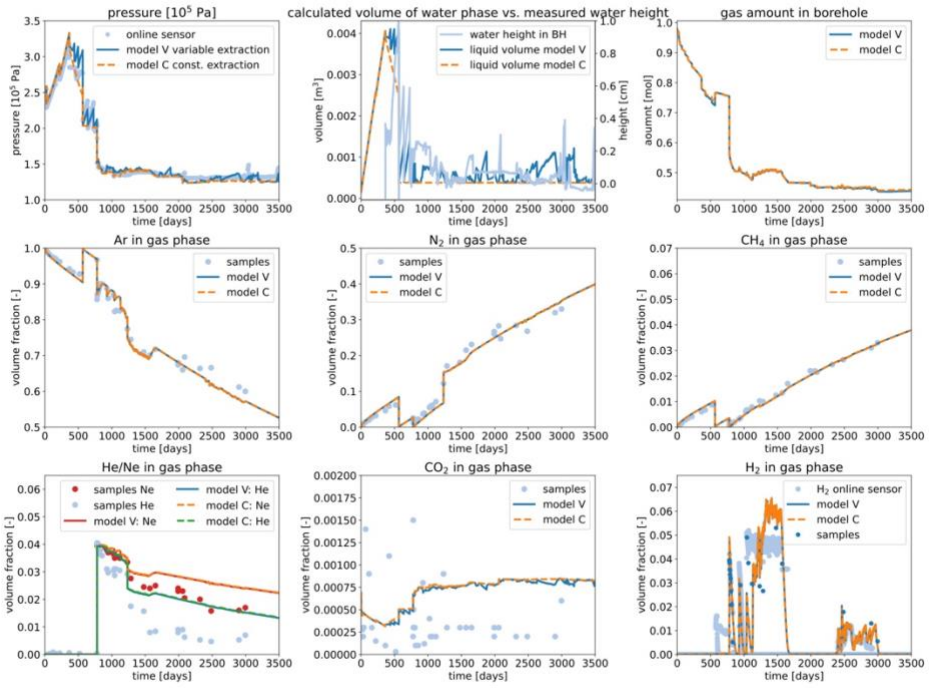


Figure 4: Comparison of the calibrated reference model V with measured data and an alternative model C for constant water inflow and extraction. The maximum biomass in the borehole available for hydrogen consumption is 1.2×10^{-4} mol.

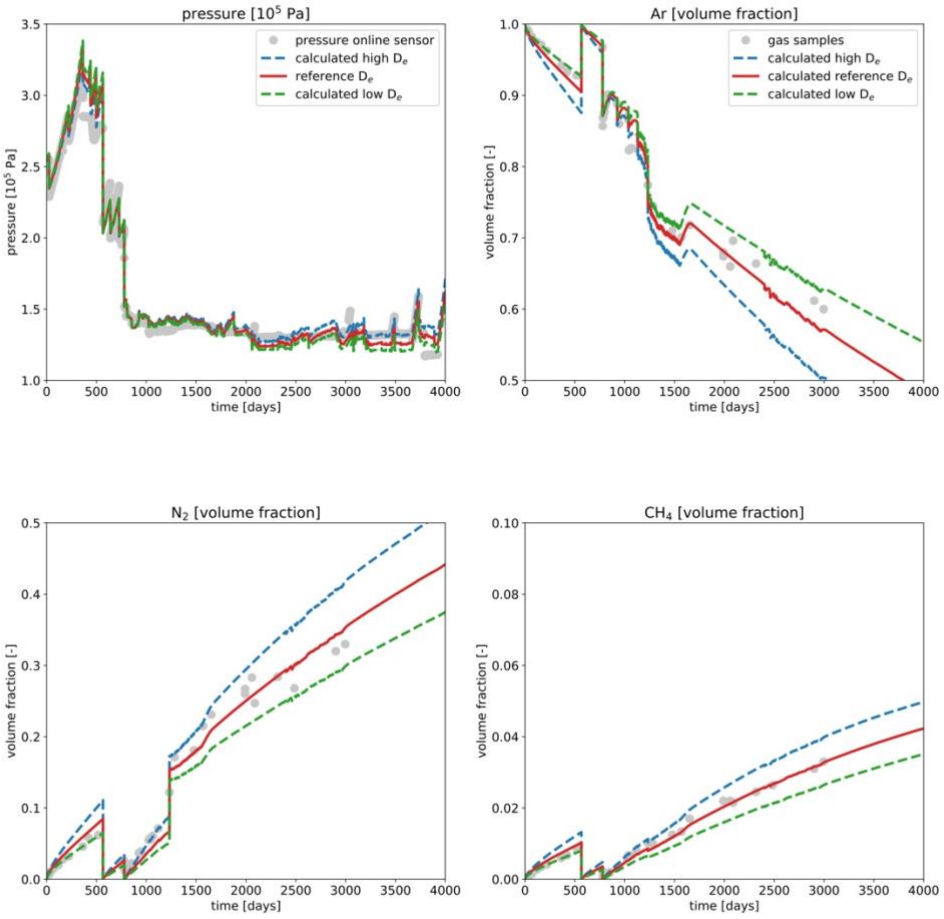


Figure 5: Influence of effective gas diffusion coefficients on the predicted pressure and gas phase composition in borehole. The variable water extraction model Vwas used for the simulations.

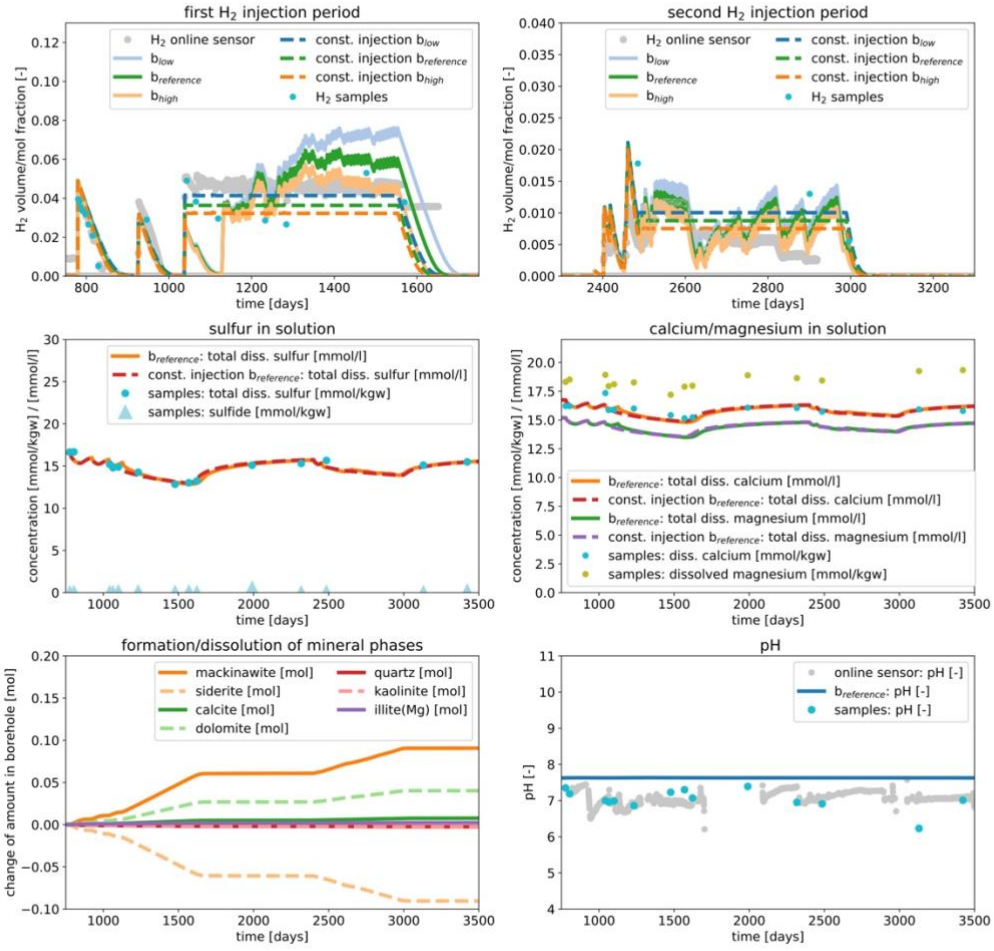


Figure 6: Evolution of hydrogen, sulphur, calcium concentrations, biomass, selected mineral phases and pH during injection of hydrogen. The reference model b_{ref} contains maximally 1.2×10^{-4} mol biomass in the borehole, while the low biomass model b_{low} contains maximally 1.1×10^{-4} and the high biomass model b_{high} maximally 1.3×10^{-4} mol biomass.

Supplement to

Hydrogen gas transfer between a borehole and claystone: Experiment and geochemical model

Leonardo Hax Damiani[1,3], Georg Kosakowski[1*], Agnès Vinsot[2], Sergey V. Churakov[1,3]

[1] *Laboratory for Waste Management, Paul Scherrer Institut, Villigen Switzerland*

[2] *Andra, CMHM, Bure, France*

[3] *Institute of Geological Sciences, University of Bern, Bern, Switzerland*

[*] *corresponding author: georg.kosakowski@psi.ch*

Chronology of tests

The test interval was closed 1.2 days after the end of the drilling operation and filled with argon at a pressure of 1.3 ($\pm$0.1) bar.
On day 23, the external gas circulation module was connected, which added 0.00093 m^3 (0.93 litres) for gas circulation. The gas circulation module also allows online measurements of the borehole's gas atmosphere composition and includes cylinders for sampling the gas.
The **"calibration phase"** starts with installing the gas circulation module, in which the composition of the gas phase and the pressure in the borehole were monitored. From day 23 on, periodically, gas samples were taken and analysed. An increase in the gas pressure was observed. On day 220, the water line valve was opened, 100ml of water was removed, and a gas leak occurred at the end of such removal. On day 364, 160ml of water and gas were sampled. In addition, an automatic water removal system was installed to keep the water amount in the borehole constant. Based on the long-term automatically extracted water amount, the rock water inflow was approximated to be about 18 ml/day. Inverse modelling of water inflow rates indicated that water inflow had to be smaller before installing the automatic water extraction system (Vinsot et al., 2017). Between days 452 and 565, the automatic water extraction was interrupted. On day 566, the test interval was flushed with argon, and on day 575, the gas circulation was interrupted. This first experimental phase is essentially a diffusion test that includes the out-diffusion of argon and the in-diffusion of gases dissolved in the Opalinus Clay pore water, specifically nitrogen, CO_2, methane, and other light alkanes. Back fitting the gas phase evolution in the borehole allows estimating effective diffusion coefficients for dissolved gases in the Opalinus Clay
On day 779, the second experimental phase started. The **"first hydrogen injection phase"** consists of a diffusion experiment with inert gases (helium, neon, argon) and hydrogen, which is known to potentially participate in various chemical and biochemical reactions (Carden and Paterson, 1979). First, the borehole was filled with a new gas mixture (85 vol% Ar, 5 vol% H_2, 5 vol% Ne, and 5 vol% He) at about 1.5 bar. It was observed that hydrogen concentration dropped much faster than inert tracer gases.

On day 926, a single hydrogen injection (14.3 mmol) was performed. On day 1038, a semi-continuous hydrogen injection started. The hydrogen content in the borehole gas phase was kept constant at a value of about 4.5%. On day 1233, a gas leak in the circulation system occurred, making it necessary to consider that 68 mmol nitrogen entered the gas circuit simultaneously as the borehole pressure kept stable at about 1.35 bar. The semi-continuous injection was stopped on day 1553. After that, gas circulation and water extraction systems were still running, which corresponds to a long-term diffusion test.
In November 2015, on day 2402, the "**second hydrogen injection phase**" started. Initially, several hydrogen pulse injections were conducted over about two months. Moreover, on day 2485, a semi-continuous injection was performed, which set the hydrogen fraction in the gas phase to about 2%. After about 1.5 years, on day 3029, the injection was stopped. Similarly to the period after the first hydrogen injection phase, the gas circulation and water extraction proceeded further. As the borehole's gas phase was not artificially replaced since the first hydrogen injection phase, this is a very long-term equilibration of the borehole and adjacent clay rock, interrupted only by the hydrogen injection experiments.

FE solution of the diffusion equation

The model presented in this work is based on the reactive transport framework developed and described in Damiani et al. (2020). Python-based interface and the UFL language to the FEniCS framework (Alnæs et al., 2015; Hoffman et al., 2012) are used to describe the transport of dissolved species in a porous medium.

The partial differential equation (PDE) for the diffusion in a water-saturated porous medium is

$$\frac{\partial \sigma C}{\partial t} = \nabla\left(\phi D_p \nabla C\right) + f, \qquad (1)$$

where ϕ denotes the porosity [-], C denotes volumetric concentration [mol m^{-3}], D_p is the species-dependent pore diffusion coefficient [m^2 s^{-1}] which is related to the effective diffusion coefficient $D_e = \phi D_p$, and f is a source/sink term [mol m^{-3} s^{-1}]. For a system with constant porosity, the equation can be simplified dividing by ϕ and get

$$\frac{\partial C}{\partial t} = \nabla D_p \nabla C + \frac{f}{\phi}. \qquad (2)$$

Forward Euler approach was adopted for time integration of Equation 2 for which the time derivative of the unknown concentrations is replaced by $\left(C^{i+1} - C^i\right)/\Delta t$ with time step size Δt and known concentrations C^i at the current time step (i) and unknown concentrations C^{i+1} at next time step (i+1):

$$\frac{C^{i+1} - C^i}{\Delta t} = \nabla D_p \nabla C^{i+1} + \frac{f^i}{\phi}. \qquad (3)$$

This transforms into a sequence of spatial and temporal problems starting with a known C_0 at t=0

$$\begin{aligned} C^0 &= C_0;\ f^0 = f_0 \\ C^{i+1} - \Delta t \nabla D_p \nabla C^{i+1} &= C^i + \Delta t \frac{f^i}{\phi}. \end{aligned} \qquad (4)$$

The term on the left side includes the unknown concentrations and, on the right side, all known terms, representing the discretised bi-linear formulation.
To solve the equation system with FEniCS, the PDEs (one per transported species) are transformed into a variational problem by multiplying it with a continuous and differentiable function *v*, the so-called test function, and integrating it over the domain *Ω*. Second-order terms $\nabla D_p \nabla C$ are integrated by parts. The unknown concentrations C^{n+1} are referred to as *C* (also called trial function) and ordered according to standard notation $a(C,v) = L_{n+1}(v)$ (with *a* the left side and *L* the right side)

$$a(C,v) = \int_\Omega \left(Cv + \Delta t D_p \cdot \nabla C \cdot \nabla v\right) dx$$
$$L_{i+1}(v) = \int_\Omega \left(C^i + \Delta t \frac{f^{i+1}}{\phi}\right) v dx \quad (5)$$

Neumann boundaries (flux boundaries) can be added to the right side of Equation (5)

$$L_{i+1}(v) = \int_\Omega \left(C^i + \Delta t \frac{f^{i+1}}{\phi}\right) v dx + \int_\Gamma \frac{j}{\phi} \Delta t v ds \quad (6)$$

where *j* is the flux [mol m^{-2} s^{-1}] across boundary Γ.

Dirichlet boundary conditions, i.e., setting fixed concentrations at a boundary, are set explicitly as part of the trial space.

Identical pore diffusion coefficients for mass transport in the Opalinus Clay formation was applied for all aqueous species, except for dissolved gases. This ensures charge conservation during the transport of dissolved species, without the need to solve computationally expensive Nernst-Planck equations for charge-coupled transport of ionic species. Dissolved gases are uncharged neutral species and can be modelled independently with diffusion coefficients taken from Vinsot et al. (2017). In FEniCS, a setup was implemented where the underlying matrices for solving a single transport equation are assembled once and then re-used to calculate transport for each species separately and sequentially. This reduces computing time, as the assembly of matrices is a relatively time-consuming task.

Model implementation

The implementation for coupling transport and chemical equilibrium calculations is based on the reactive transport framework described in Damiani et al. (2020) and uses a sequential non-iterative approach (SNIA).
At each time step, the diffusion equations for all dissolved species are solved and then new concentrations are updated (optionally) by calculating new chemical equilibria between pore-water and mineral phases with the xGEMS thermodynamic solver. xGEMS (https://bitbucket.org/gems4/xgems/) is a C++ and python API (application programming interface) for accessing the GEMS3K thermodynamic solver.

Based on the symmetry of the experimental setup, a 2D triangular mesh in the form of an annulus sector with 8421 nodes and 16000 elements represents the 3D rock surrounding the borehole (Figure S1). The mesh is oriented perpendicular to the borehole and parallel to the bedding of the Opalinus Clay rock. The inner circle corresponds to the borehole wall, while the outer circle has a radius of 2 m. The mesh is refined towards the inner circle, as spatial and temporal concentration changes at the borehole walls are expected to be highest. Besides, for an accurate

calculation of the flux, an adequate spatial resolution of concentration gradients is required at the boundary.

A simplified flow chart for the simulations is shown in Figure S2. At the beginning of the simulations, concentrations for dissolved species, including dissolved gases, are set based on the equilibrated Opalinus Clay setup.

The straight left and right boundaries of the modelling domain are treated as "no-flux" boundaries. The outer circle of the annulus sector is a Dirichlet boundary with constant concentrations, identical to initial concentrations in the whole domain. The inner circle, which corresponds to the borehole wall, also forms a Dirichlet boundary. Concentrations at the inner circle are assumed to be in equilibrium with the borehole system and are updated each time step according to the borehole equilibration results.

The model's workflow represents a typical sequential non-iterative approach (SNIA) for reactive transport calculations. After reading input files, initialising the transport solver, and the geochemical solver instances, the time loop starts by calculating transport equations. If desired, at each FE node, the chemical system is equilibrated based on the newly calculated concentrations from the transport solution. This is done via multi-threaded instances of the GEMS3K solver (Kulik et al., 2012), which are executed concurrently to speed up the calculations. After updating the concentrations, the fluxes across the inner boundary (borehole wall) are calculated and, after being converted to species amounts, are added/removed from the borehole system during the current time step. Besides, the influx of water into the borehole is done by adding a defined amount of Opalinus Clay pore water with a composition equivalent to the water at the borehole wall.

The species fluxes change species concentrations in the liquid phase of the borehole. Since the borehole's liquid film at its surface and the bulk water accumulated at its bottom are treated as one homogeneous liquid phase, it is not possible to resolve any differences in terms of composition between water films and bulk water.
The difference between solute concentrations in the water film and the accumulated water at the bottom is likely to be negligible, as the residual water amount in the borehole did not exceed ~570 ml during the automatic water extraction period and is most of the time much less. For an inflow of rock pore water with the rate of 18 ml/day, at least 1/30 of the water would be replaced per day.

Finally, other events like the gas phase sampling are implemented by adjusting the bulk composition in the borehole and possibly changing the reference volume of the gas circulation system.

After a first equilibration of the borehole composition, a fraction of the aqueous phase in the borehole is removed to keep the amount of water in the borehole constant. This mimics the automatic water extraction, conducted only if the pressure and the amount of water in the borehole exceed pre-defined thresholds.

After the water extraction routine, the borehole's system is equilibrated again, and the consumption of hydrogen is calculated with the kinetic equations. After adjusting the bulk composition in the borehole by adding "reactive hydrogen" and removing "non-reactive hydrogen", and dissolving the required amount of siderite, the borehole system needs to be equilibrated again.

Finally, the simulation time step is incremented, and the calculation loop starts again by updating the transport boundary condition at the borehole wall based on the last equilibrated borehole composition.

A variable time-stepping scheme is implemented for the time loop. The initial minimum time step size is increased successively by multiplication with 1.1 until a maximum time step size is reached. If an event changes the composition of the borehole and the boundary condition for the transport solver significantly, the time step size is set to the minimum value. It was found that for the investigated scenarios, a value of 1 000 s for the minimum time step and 10 000 s for the maximum time step gives a good compromise between the need to resolve temporal changes in fluxes across the borehole wall and the need to minimize computational costs of the simulation runs.

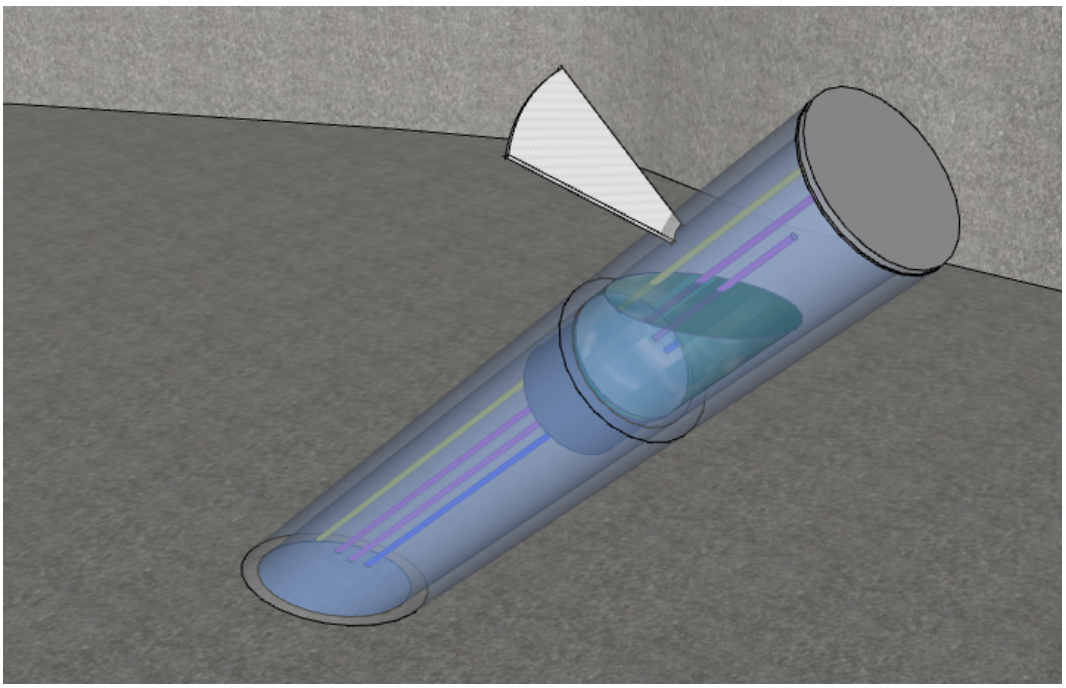


Figure S1: Conceptual sketch of the borehole and the FE-mesh used for modelling diffusive transport in Opalinus Clay. Please note that the size of borehole is not in correct relation to the size of the FE-mesh used for modelling. The radius of the borehole should be 0.038 m with a length of 5 m for the injection interval; the FE-mesh should extend up to 2 m from borehole centre.

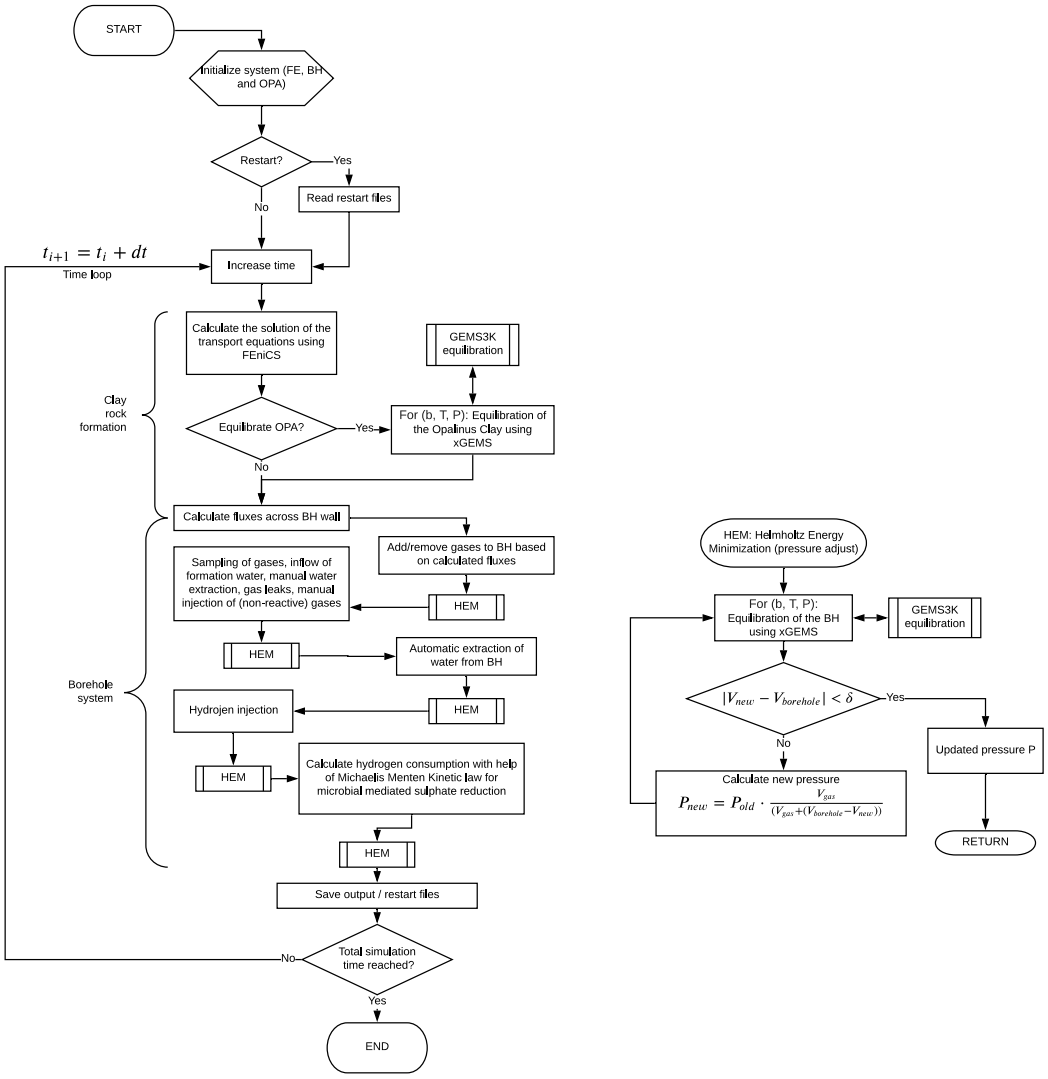
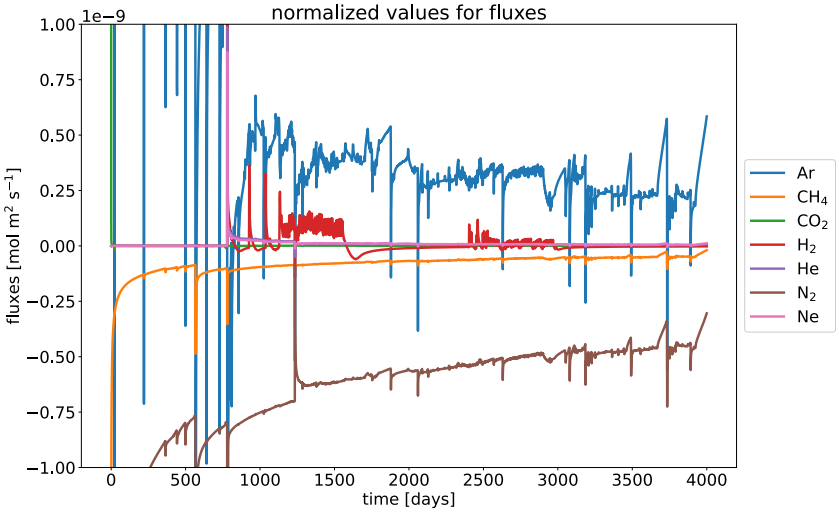


Figure S2: Flow chart diagram of the modelling algorithm implemented using python scripting to simulate experimental data from the HT experiment.

a)



b)

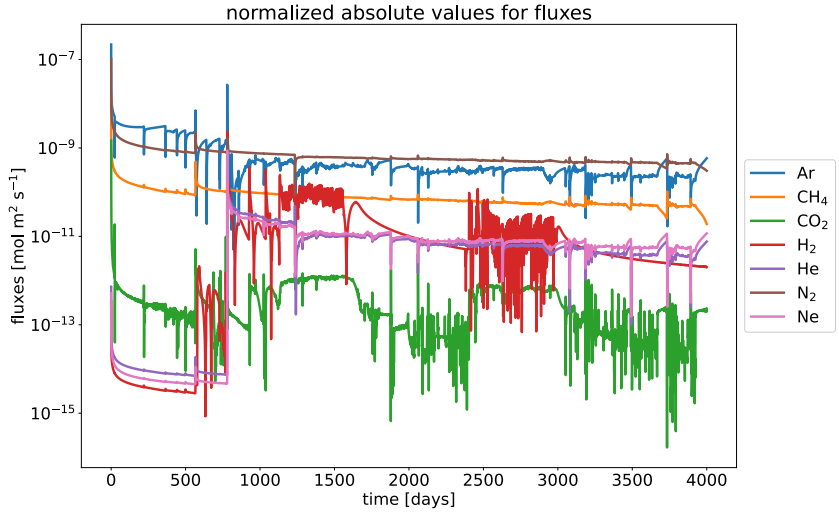


Figure S3: Fluxes of dissolved gases across the borehole wall calculated from the FE transport model in claystone. Subfigure a) uses linear axis for fluxes. Negative fluxes indicate flux direction towards the borehole, while positive fluxes are towards the claystone. Subfigure b) uses logarithmic axis to plot absolute values for fluxes.

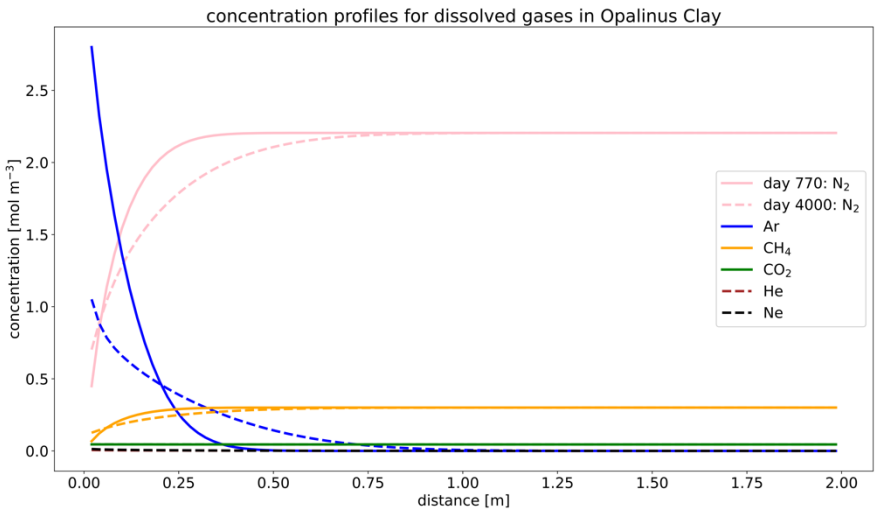


Figure S4: Concentration profiles in radial direction for selected dissolved gases in clayrock water for two simulation times. Solid lines denote day 770, about 9 days before the first hydrogen injection. Dashed lines denote the concentration at day 4000.

Thermodynamic setup

A unified setup in terms of phase and species definition based on the Thermoddem DB (Blanc et al., 2012) has been used for all the thermodynamic calculations. Some of the gases of interest, specifically N_2, CH_4, and H_2, are considered non-reactive species in the gas and water phases.

Opalinus Clay setup

The thermodynamic setup for Opalinus Clay is based on the approach of modelling Opalinus Clay pore water described in Pearson et al. (2011). The pore water composition is described by equilibria with mineral phases and a cation exchanger phase, which is implemented as an ideal solid solution phase (Kulik, 2010).
Fixing pore water composition by mineral equilibria makes it challenging to accurately reproduce the average Opalinus Clay pore water composition given in Vinsot et al. (2014). The pore water compositions can be manipulated by changing the cation exchanger's composition, adjusting exchanger constants or choosing other minerals (clay) phases for equilibration. Pearson et al. (2011) compared Opalinus Clay water models that use different combinations of clay minerals for equilibration. Scoping calculations suggest that the combination of kaolinite, illite, celestite, dolomite, and some minor amount of daphnite gives a pore-water composition in Opalinus Clay, which is close to the desired composition in terms of major cations and sulphate concentrations. A comparison of pore-water properties between the model and those used by Vinsot et al. (2014) is provided in the paper. The concentration of strontium deviates considerably from measured values because, in the model, celestite is used to fix both sulphate and strontium concentrations. Potassium concentrations are lower than measured, but they are in line with those of Pearson et al. (2011). The partial pressure of CO_2 at equilibrium with water is one order of magnitude lower than the value deduced from the observed water composition. Finally, magnesium concentrations are slightly lower than measured and modelled by Pearson et al. (2011) due to the influence of changed cation exchanger composition. The agreement with the desired pore-water composition can be further improved by adjusting the mineral phase stabilities,

cation exchanger constants, and by adaption of cation exchanger capacity and occupancy.
The modelled pH is set by equilibria with alumino-silicate minerals, specifically illite, kaolinite, and chlorite. The calculated value of 7.6 is comparable with the values calculated by Pearson et al. (2011) but is also significantly higher than the measured values.

Borehole setup

The same thermodynamic setup used for the Opalinus Clay chemical system definition is used to equilibrate the borehole system with a different bulk composition. An initial setup in terms of phase composition included is given in Table 2 of the paper and the major properties of the aqueous water phase in Table 3. The system was created by equilibrating the equivalent of 10 ml of Opalinus Clay with 0.16 litres of Opalinus Clay pore water and 6.762 moles of Ar. The system's total initial volume is about 9.6 litres, of which about 9.4 litres is occupied by the gas phase, 0.16 litre by the liquid phase, and 0.04 litres by several mineral phases. In addition, siderite was added to be the source of iron for microbial mediated sulphate reduction during hydrogen injection.
This setup gives an equilibrated borehole water composition different to that of the Opalinus Clay pore water, as the liquid to solid ratio is much higher in the borehole. Specifically, the much smaller cation exchanger capacity and a gas phase's existence enforces different equilibrium concentrations for most dissolved species. Also, the celestite phase was removed from the borehole's setup to avoid a buffering of sulphate concentrations during microbial mediated consumption of hydrogen.

It is assumed that hydrogen consumption in the borehole is mainly caused by H_2-fueled, kinetically controlled microbial mediated sulphate reduction (Appelo and Vinsot, 2012; Bagnoud, Chourey, et al., 2016; Bagnoud, Leupin, et al., 2016; Boylan et al., 2019; Leupin et al., 2017; Vinsot et al., 2014, 2017).

How long does it take to equilibrate water at the borehole wall?

Tokunaga and Wan (1997) estimated the water layer thickness to vary between 2 and 70 µm for a water film flowing along fracture surfaces of porous rock. For a very rough estimation of the time needed to equilibrate such a water film with the adjacent rock, it was assumed that water is entering homogeneously over the total borehole surface and forms a film with a thickness of 10 µm. The amount of water in the film layer covering the entire borehole surface corresponds to a water volume of ~1.2e-5 m^3 (12 ml). The influx of Opalinus Clay pore water was estimated to be about 18 ml/day (Vinsot et al., 2017). Therefore, it would take about less than one day to completely replenish such a layer with the rock pore water solution. According to the square root relationship for diffusion in an infinite medium x ~$\sqrt{2Dt}$ (Crank, 1975) with a gas diffusion coefficient of D ~2e-9 m/s in the water, any concentration differences across the water film are expected to level out in much less than one day.

References

Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., et al. (2015), "The FEniCS project version 1.5", *Archive of Numerical Software*, Vol. 3 No. 100, available at:https://doi.org/10.11588/ANS.2015.100.20553.

Appelo, C.A.J. and Vinsot, A. (2012), *HT(Hydrogen Transfer) Experiment: Model for Transport of Gases Ar, H_2, He, Ne, N_2, CO_2 and Alkanes with Reactions for Hydrogen in Opalinus Clay*, Mont Terri Technical Note TN 2012-58, Federal Office of Topography (swisstopo), Wabern.

Bagnoud, A., Chourey, K., Hettich, R.L., de Bruijn, I., Andersson, A.F., Leupin, O.X., Schwyn, B., et al. (2016), "Reconstructing a hydrogen-driven microbial metabolic network in Opalinus Clay rock", *Nature Communications*, Nature Publishing Group, Vol. 7 No. 1, p. 12770.

Bagnoud, A., Leupin, O.X., Schwyn, B. and Bernier-Latmani, R. (2016), "Rates of microbial hydrogen oxidation and sulfate reduction in Opalinus Clay rock", *Applied Geochemistry*, Pergamon, Vol. 72, pp. 42–50.

Blanc, P., Lassin, A., Piantone, P., Azaroual, M., Jacquemet, N., Fabbri, A. and Gaucher, E.C. (2012), "Thermoddem: A geochemical database focused on low temperature water/rock interactions and waste materials", *Applied Geochemistry*, Vol. 27 No. 10, pp. 2107–2116.

Boylan, A.A., Perez-mon, C., Guillard, L., Burzan, N., Loreggian, L., Maisch, M., Kappler, A., et al. (2019), "H_2-fuelled microbial metabolism in Opalinus Clay", *Applied Clay Science*, Elsevier, Vol. 174 No. March, pp. 69–76.

Carden, P.O. and Paterson, L. (1979), "Physical, chemical and energy aspects of underground hydrogen storage", *International Journal of Hydrogen Energy*, Pergamon, Vol. 4 No. 6, pp. 559–569.

Crank, J. (1975), *The Mathematics of Diffusion*, Second Edition, Clarendon Press, Oxford.

Damiani, L.H., Kosakowski, G., Glaus, M.A. and Churakov, S. V. (2020), "A framework for reactive transport modeling using FEniCS–Reaktoro: governing equations and benchmarking results", *Computational Geosciences*, Vol. 24 No. 3, pp. 1071-1085.

Hoffman, B.J., Jansson, J., Degirmenci, C., Jansson, N. and Nazarov, M. (2012), *Automated Solution of Differential Equations by the Finite Element Method*, edited by Logg, A., Mardal, K.-A. and Wells, G., Vol. 84, Springer Berlin Heidelberg, Berlin, Heidelberg, available at:https://doi.org/10.1007/978-3-642-23099-8.

Kulik, D.A. (2010), "Geochemical thermodynamic modelling of ion partitioning", *EMU Notes in Mineralogy*, Vol. 10, pp. 65–138.

Kulik, D.A., Wagner, T., Dmytrieva, S. V., Kosakowski, G., Hingerl, F.F., Chudnenko, K. V. and Berner, U.R. (2012), "GEM-Selektor geochemical modeling package: revised algorithm and GEMS3K numerical kernel for coupled simulation codes", *Computational Geosciences*, Vol. 17 No. 1, pp. 1–24.

Leupin, O.X., Bernier-Latmani, R., Bagnoud, A., Moors, H., Leys, N., Wouters, K. and Stroes-Gascoyne, S. (2017), "Fifteen years of microbiological investigation in Opalinus Clay at the Mont Terri rock laboratory (Switzerland)", *Swiss Journal of Geosciences*, Springer International Publishing, Vol. 110 No. 1, pp. 343–354.

Pearson, F., Tournassat, C. and Gaucher, E.C. (2011), "Biogeochemical processes in a clay

formation in situ experiment: Part E – Equilibrium controls on chemistry of pore water from the Opalinus Clay, Mont Terri Underground Research Laboratory, Switzerland", *Applied Geochemistry*, Vol. 26 No. 6, pp. 990–1008.

Thiel, J., Byrne, J.M., Kappler, A., Schink, B. and Pester, M. (2019), "Pyrite formation from FeS and H_2S is mediated through microbial redox activity", *Proceedings of the National Academy of Sciences*, National Academy of Sciences, Vol. 116 No. 14, pp. 6897–6902.

Tokunaga, T.K. and Wan, J. (1997), "Water film flow along fracture surfaces of porous rock", *Water Resources Research*, Vol. 33 No. 6, pp. 1287–1295.

Vinsot, A., Appelo, C.A.J., Lundy, M., Wechner, S., Lettry, Y., Lerouge, C., Fernández, A.M., et al. (2014), "In situ diffusion test of hydrogen gas in the Opalinus Clay", *Geological Society, London, Special Publications*, Vol. 400 No. 1, pp. 563–578.

Vinsot, A., Appelo, C.A.J., Lundy, M., Wechner, S., Cailteau-Fischbach, C., de Donato, P., Pironon, J., et al. (2017), "Natural gas extraction and artificial gas injection experiments in Opalinus Clay, Mont Terri rock laboratory (Switzerland)", *Swiss Journal of Geosciences*, Vol. 110 No. 1, pp. 375–390.

Chapter 4

Machine learning assisted geochemical reactive transport modeling: a unified modular computational framework

Machine learning assisted geochemical reactive transport modeling: a unified modular computational framework

LEONARDO H. DAMIANI[1,2], ROBIN L. I. HALLER[2], GEORG KOSAKOWSKI.[1], GEORGE D. MIRON.[1], DMITRII A. KULIK.[1], SERGEY V. CHURAKOV.[1,2], NIKOLAOS PRASIANAKIS.[1]

[1]Laboratory for Waste Management, Paul Scherrer Institut, 5232 Villigen, Switzerland
[2]University of Bern, Institute of Geological Sciences, 3012 Bern, Switzerland

Corresponding author: Nikolaos Prasianakis (e-mail: nikolaos.prasianakis@psi.ch).

ABSTRACT Reactive transport modeling is a standard computational tool used to analyze and predict physical and chemical phenomena. The objective of reactive transport codes is to carry out accurate simulations of natural systems; hence, the more complex and detailed the phenomena analyzed by the reactive transport model, the higher the computational costs. Machine learning and other new computing techniques allow modern reactive transport codes to tackle complex problems in full detail by reducing computational costs, for example, by using high-performance computing and GPU architecture. This article presents a geochemical machine learning framework that bundles together several tools in order to carry out reactive transport simulations using deep neural networks (DNN) instead of the traditional chemical solver. A framework of this kind has enormous potential, since it makes it possible to first create the training dataset with a traditional geochemical solver, then to use machine learning to identify patterns and train the neural network (or any other major machine learning technique) to predict correct outputs, and, finally, to harness the deep neural network to perform fast geochemical equilibrium calculations needed by the reactive transport simulations. In order to validate the framework presented here, we consider a typical geochemical application in which a $1.0\,\mathrm{mol}\ MgCl_2$ is injected into a calcite column, causing calcite dissolution and dolomite precipitation. The results show a very high level of agreement between the traditional geochemical solver and the accelerated machine learning based approach. Additionally, we present a second application, in which a second DNN is used with the supporting role of providing thermodynamic data for the reactive transport simulation, making possible a more comprehensive emulation of the geochemical solver.

INDEX TERMS Artificial neural network, finite element method, geochemical modeling, reactive transport in porous media, machine learning

I. INTRODUCTION

REACTIVE transport modeling is a valuable tool for analyzing coupled physical and chemical processes in the natural sciences [1]. These models consider geochemical properties and reactions (e.g., the precipitation and dissolution of minerals, swelling and sorption mechanisms) combined with mass transport phenomena (e.g., diffusion, advection). They have a variety of applications, for example, in CO_2 sequestration [2], deep geological repositories for radioactive waste [3], and geothermal energy [4]. Application-related geochemical characteristics have led to different numerical algorithms and specific functionalities, producing several reactive transport packages over the years [5], [6].

Geochemical equilibration is known to be the bottleneck in reactive transport modeling (RTM) [7]. Moreover, depending on the complexity of the geochemical system described (e.g., the number of components, the presence of ideal and non-ideal mineral and fluid phases, chemical reactions of interest), such simulations can be highly time-consuming and computationally expensive.

Consequently, alternative solutions and efforts to reduce computational costs and overcome this bottleneck have emerged. For example, one straightforward approach is to adopt a simplified chemical system with fewer phases and components that focuses on specific aspects of interest, while ignoring others [8], [9]. Look-up tables have also been used to approximate the chemical reactions and achieve very similar results with a numerically efficient and truthfully equivalent approach [10]–[14].

Recently, numerous studies have adopted machine learn-

ing (ML) and artificial intelligence (AI) techniques. They both perform well and can deal with large amounts of data relatively accurately, for example, drug discovery and description [15], hand-gesture recognition [16], image analysis and reconstruction [17], protein structure prediction [18], and water treatment [19]. More specifically, neural networks (NN) have gained popularity in, for example, combustion [20] and geochemical modeling [21]–[25]. These innovative techniques have enormous potential to decrease the computational costs incurred by algorithms of this kind.

Some criticism has been directed towards the claim that ML represent actual *learning* [26]: One of the arguments stresses that these algorithms do not learn as humans do, but instead associate statistical analysis and correlations to produce correct answers that are perhaps not always grounded in valid reasons. Nevertheless, the neural networks are supposed to mimic human neurons. In most applications, the algorithms are data-driven and can encode the structures and underlying physics presented in the form of a multidimensional dataset. Even if they do not outperform the accuracy of human experience in relation to several problems, they certainly exceed human throughput capacity.

In NN mathematical descriptions, some hyperparameters (e.g., the number of neurons, layers, activation function, kernel initialization) control the NN settings and are responsible for the behavior of the NN model, whose aim is to produce accurate outputs from the inputs provided. Currently, there is no predefined method for selecting the best set of hyperparameters [27]. However, while finding the appropriate hyperparameters is highly relevant, it is not a straightforward task. Moreover, the training routine can take a long time and be very computationally intensive.

Finding the right combination of an efficient training algorithm with a good set of hyperparameters is fundamental when it comes to creating a NN model that correctly identifies the patterns from the training dataset and produces accurate outputs. For example, a NN with more neurons can solve complex problems, but, at the same time, it can also result in slow performance and overfitting (i.e., when the data outliers are mistakenly considered part of the training data) [28].

In this paper, we present a geochemical reactive transport modeling framework that enables the efficient, flexible, and accurate replacement of the traditional geochemical solver with an NN model. Additionally, we provide the complete geochemical neural network pipeline in a single bundled and modular framework: from the dataset generation based on a choice of a traditional chemical solver, through searching for and tuning the proper hyperparameters, to the creation and training of the NN model, and, finally, the deployment of the NN model in a wholly integrated reactive transport modeling framework in order to perform geochemical simulations. This pipeline logic is presented in Fig. 1.

Using a NN to emulate the traditional geochemical solver can dramatically decrease simulation time, while maintaining good accuracy. It is an exciting approach to tackling complex reactive transport modeling problems using a powerful technique with state-of-the-art AI computational tools. The framework presented here allows us to explore the hyperparameters of the NN, in order to find the most suitable ones in a self-contained, easy and interactive way. Moreover, it makes it possible to train the models—a very computationally demanding task—in high-performance computing (HPC) facilities without any code change. Finally, one can compare the results of different NN models and examine how they perform against the fully coupled reactive transport results. Once the neural network has been produced and benchmarked in, e.g., a 1D reactive transport test case, its efficiency is assessed and may be used for the demanding 3D reactive transport simulations. A typical example of high-performance computing of a 3D reactive transport simulation (e.g., using the lattice Boltzmann method) is usually described in a 1000×10^3 computational nodes. While mass transport operations are generally fast, the slow geochemical calculations must be performed at every computational node and at every time step. With time steps being of the order of 100K–1M, the resulting number of geochemical calculations is more than one quadrillion (10×10^{15}) per simulation, making it impractical to use detailed geochemistry in the traditional sense. Therefore, the use of neural networks opens new horizons, since it enables the acceleration of geochemical calculations by a factor of approximately 10×10^3 to 10×10^4, essentially making the calculation of chemistry computationally less expensive than the transport.

This paper describes the details of the transport and chemical solvers, as well as how the sequential non-iterative approach (SNIA) is used to couple both. Then, we present the artificial neural network definitions and properties. After that, we validate the methodology using a benchmark in which a calcite column is flushed with $MgCl_2$, causing calcite dissolution and dolomite precipitation. Additionally, we propose a reactive transport setup using an auxiliary neural network that provides specific thermodynamic data in order to emulate the traditional geochemical solver in a more comprehensive way. To conclude, we examine the temporal and computational costs, summarize our work, and discuss future possibilities for the geochemical deep learning neural network infrastructure presented here.

II. METHODOLOGY

For this paper, we use a variation of the reactive transport framework defined in [6]. It performs reactive transport simulations in fully saturated multiphase-multicomponent heterogeneous systems with chemical equilibrium and kinetics, using the Gibbs Energy Minimization (GEM) algorithm.

For the chemical equilibrium calculations, we use Reaktoro [29], a robust geochemical framework for modeling chemically reactive systems that calculates the equilibrium condition for a given chemical system using the GEM algorithm. It supports non-ideal multicomponent solid solutions, a variety of activity models, and several equations of states for multicomponent fluid systems.

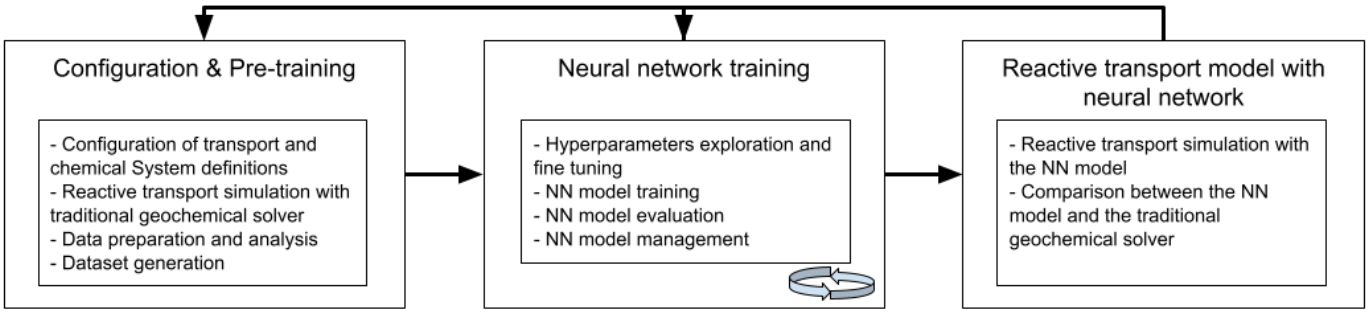


FIGURE 1. Pipeline for the geochemical machine learning framework showing the fully integrated bundled infrastructure

For the reactive transport simulation module, the partial differential equation (PDE) solver FEniCS [30] is used, which describes the physical mass transport of solutes using its weak formulation via the Unified Form Language (UFL) [31]. Moreover, it offers several Finite Element (FE)–based numerical methods, which enables great flexibility by giving full access to the PDE level and compatibility with distributed memory HPC architecture PETSc-based parallelization.

A. TRANSPORT OF SOLUTES

Diffusion and advection are mass transport processes that depend on concentration and pressure gradients [32]. In a general multicomponent form description [33], the diffusive and advective fluxes in a porous medium are:

$$\theta \frac{\partial c_i}{\partial t} = -\nabla \cdot \theta \big(V c_i - D_e (\nabla c_i) \big) \tag{1}$$

where c_i is the volumetric concentration ($\mathrm{mol} \cdot \mathrm{m}^{-3}$) of the i-th transported species, t is time (s), θ is the porosity (dimensionless), V is the pore velocity ($\mathrm{m} \cdot \mathrm{s}^{-1}$), and D_e is the effective diffusion coefficient ($\mathrm{m}^2 \cdot \mathrm{s}^{-1}$). An implicit time discretization scheme is used due to its stability (respecting the Von Neumann criteria [34], [35]) when coupling transport and chemical solvers using operator-splitting [36] in combination with a short time step length. The implicit time step scheme is:

$$\frac{\partial c_i}{\partial t} \approx \frac{c^n - c^{n-1}}{\Delta t} \tag{2}$$

where c^n is the current unknown concentration, c^{n-1} is the known concentration from the previous time step and Δt is the time step size. The corresponding weak formulation of (1), using the Lagrange finite element family and combining (2), is:

$$\int_\Omega \theta c_i^n v_i dx = \int_\Omega \big(\theta c_i^{n-1} - \theta \Delta t (V c_i^n - D_e \nabla c_i^n) \big) \cdot \nabla v_i dx \tag{3}$$

where v_i is the test function, a smooth, continuous, and differentiable function inside the domain [37].

B. CHEMICAL EQUILIBRIUM CALCULATION

Two popular methods for performing equilibrium calculations are Gibbs Energy Minimization (GEM) [38]and the Law of Mass Action (LMA) [39], which use Gibbs energy (G) and the equilibrium constant (K) respectively. Both have been extensively studied and used in RTM over the years [24], [40]–[45] and they are mathematically equivalent under certain conditions [46]. As stated in [5], LMA is used in popular RTM codes (e.g., EQ3/6 [47], Phreeqc [48], GWB [49]) and it is a stoichiometric-based approach that benefits from the fact that multiple thermodynamic databases are available that contain the equilibrium constants of reactions [50]. GEM is the preferred method used in this paper for modeling the multiphase-multicomponent geochemical systems, thanks to the advantages it confers when modeling multiphase systems, including non-ideal solid solutions, and the fact that it is based on the standard chemical potentials. It is used via the chemical solver Reaktoro [29]. We refer to [29] and [38] for the governing equations and details of the GEM chemical equilibrium algorithm.

This paper aims to present an alternative approach in which a deep neural network replaces the chemical solver (e.g., GEMS or Reaktoro) during the reactive transport simulation to reduce the overall simulation time. However, the traditional solver is still very relevant, since it provides the training data, in addition to being used for cross-validation and benchmarking.

C. COUPLING OF TRANSPORT AND CHEMICAL SOLVER

In order to preserve the modular nature of the framework and to quickly provide the possibility of coupling the different modules together, we adopt SNIA [51], [52]. Using SNIA, we couple the transport solver, which is responsible for the physical mass transfer processes, and the chemical solver or NN, which is responsible for the chemical reactions and equilibration. For each time step, the partial differential equations that describe the transport of aqueous components are solved first. Then, the result is forwarded to the chemical solver or NN, which calculates each node's new equilibrium condition. After that, the new amounts of transported species and minerals are updated based on the equilibration results. The loop is then repeated until the total simulation time is reached.

III. GEOCHEMICAL MACHINE LEARNING APPROACH

The traditional deterministic programming algorithms require the input data and a set of instructions for producing the output. When using ML, this approach is replaced with a less constrained, self-learning, output-oriented concept: The input and output are provided beforehand, while a choice among several ML training algorithms is responsible for calibrating and adjusting the internal parameters of the NN model in order to reproduce the output. One of the challenges of replacing the chemical solver with an NN is the number of inputs and outputs, because the inputs and outputs directly impact its accuracy and ease of training. Essentially, numerous inputs mean that the data lies in multidimensional surfaces, and many neurons and layers might be needed for such a description. The larger the network, the more challenging it is to train it to achieve the desired accuracy. Moreover, as stated in [21], the simple feedforward NN architectures are not fully mass conservative, and if the mass balance error is significant and biased, it may grow cumulatively with each time step.

A. NEURAL NETWORK CONCEPTS

Neural networks (NN) are traditionally applied to classification tasks, in which the output is a class or category, as well as to data regression problems, in which the output is a number. NN is composed of interconnected neuron nodes responsible for processing the inputs to produce outputs [53]. The "neurons," or processing units, generate outputs based on a non-linear function of the sum of its inputs. Therefore, the more information and guidelines are provided to the neurons during the training period, the higher the chances of the NN's internal algorithm making appropriate decisions and producing accurate outputs.

NN with multiple layers were initially presented in [54]–[56] and the term "deep learning" was coined a few years later in [57]. Nevertheless, only after the dissemination of the effective techniques presented in [58] did the deep neural networks (DNN) start to gain popularity due to the new efficient training routines and their ability to deal with a large amount of data. Fig. 2 shows a feed-forward DNN architecture, for illustration purposes, in which the multiple neurons from each layer are fully connected.

Fig. 3 shows how the neuron weights the inputs and produces the output using a non-linear activation function. Each neural network property (e.g., number of layers, neurons, activation function) has its advantages and limitations, and often one needs to find the best combination empirically. For a detailed description, we refer to [59].

The training process of a NN optimizes each neuron's weight based on the samples from the training dataset through an interactive process in which, for a given set of inputs, the correct parameters are given at the output [60]. In order to achieve this, it is possible to use, for example, gradient descent algorithms. The adjustment of the iterative weights is known as the back-propagation phase. It is a very popular procedure and a crucial phase in neural network

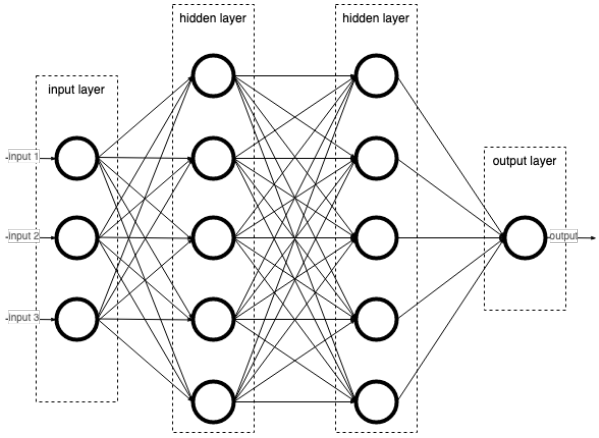


FIGURE 2. Deep neural network architecture with three inputs, two hidden layers, and one output

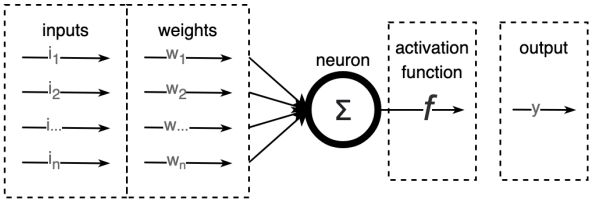


FIGURE 3. Each neuron produces a nonlinear output based on the weighted summation of its inputs. Nonlinearity is achieved using an activation function

training, as it is responsible for tuning the weights in order to ensure lower error rates on later iterations [61]–[64].

There are many techniques, algorithms, and approaches to training a DNN [65], [66], but there are still many unresolved challenges, such as how to efficiently and accurately train a high dimension neural network [67], [68]. One popular technique used to assist training routines is early stopping callback [69], which is commonly used to monitor and halt training when a particular metric of interest (e.g., accuracy) does not improve after a certain threshold. Another important technique is the dropout, which randomly ignores certain neurons during the training routine, causing an overall improvement in the results and thus helping mitigate the effects of overfitting [70]. The objective is to keep the network as simple as possible (with a reduced number of layers and neurons) and sufficiently robust to learn and predict accurately. This ensures a rapidly trainable and high-performing neural network.

NN hyperparameters play an important role in its behavior, as well as in the optimization process [59], [71]. A few of these hyperparameters are summarized as follows:

- Layers: a container of neurons that receives an input, performs a calculation, and produces outputs. In Fig. 2, the layers are indicated with dashed rectangles.
- Neurons: the processing units, which perform mathematical functions on their inputs in order to produce outputs. The neurons are placed inside layers, as indicated in Fig. 2, and their behavior is detailed in Fig. 3.

- Batch size: the number of samples used in each training iteration before updating the weight of each neuron. Usually, the use of small batch sizes is desirable in order to stimulate smaller, but consistent updates to the learning algorithm.
- Epoch: the number of iterations (or cycles) for which the learning algorithm will use the dataset. Typically, a high number of epochs is used in order to allow the learning algorithm to evaluate the data multiple times and to decrease prediction errors by adjusting the weights.
- Learning rate (LR): the size of the change in weight at each step based on the error metric. A small LR results in a slow training routine, while an excessively large LR results in unstable training.
- Momentum: increases or reduces the impact of the LR. It can accelerate the learning algorithm or decrease the changes, once a minimum gradient of descent has been found.
- Activation function: is the mathematical function mechanism that produces an output based on the neuron's weights. It significantly impacts the NN's performance, and different layers can have different activation functions. Popular activation functions are, for example, the sigmoids, the tanh and the rectified linear activation function (RELU), with the latter returning the value provided as input or zero if the input is equal or smaller than zero.
- Kernel initialization: the initial weights attributed to the neurons. Usually, a random distribution is preferred to avoid the influence of an initial seed.
- Optimizers: the training algorithms that adjust the weights of the neurons and learning rate during training, in order to reduce the model's losses and to produce the most accurate output possible. Several training algorithms exist (each with its own advantages and disadvantages): e.g., Stochastic Gradient Descent (SGD) [72], RMSprop [73], Adam [72], and Adadelta [74]. New algorithms are constantly being developed in this rapidly evolving field.

B. GEOCHEMICAL SOLVER EMULATION INFRASTRUCTURE

In this paper, a regression NN model is used. It takes as its input the amount of each independent component (in moles) and produces as an output an equilibrium condition in amounts of aqueous species and minerals (in moles). Alternatively, the output could be expanded to include other properties of the traditional geochemical solver, such as pH, Eh, volume, and pressure. The trained NN model is coupled within the SNIA algorithm, as shown in Fig. 4, and provides a seamless tool in the reactive transport framework.

The size and quality of the existing samples in the training dataset are essential to producing an accurate neural network, and there are multiple ways to create the dataset. For demonstration purposes, we use the output of the fully coupled reactive transport model to constrain the dataset and to ensure that the dataset represents the whole reactive transport simulation domain.

The geochemical ML infrastructure presented in this paper is based on TensorFlow, a free and open-source machine learning application, and Keras, a high-level Python interface that facilitates the usage of TensorFlow. In addition, Keras provides methods for building and training neural networks, enabling fast experimentation and the exploration of machine learning algorithms in a few lines of code. For the dataset generation, this work uses the chemical solver Reaktoro [29], and ThermoFun [75]. ThermoFun is an open-source provider of thermodynamic properties of substances and reactions at the temperature and pressure of interest [75]. The geochemical ML infrastructure, presented in this work, has been bundled together using Python, Conda packages, and Jupyter notebooks, allowing for the quick deployment of the framework, as well as for a seamless user experience [76].

C. HYPERPARAMETER SEARCH

Finding the appropriate set of hyperparameters that effectively produces a performant NN is a challenging task (encapsulated in the middle component of Fig. 1). It involves: defining a complete set of parameters; performing a full training; evaluating the NN model; conducting the reactive transport simulation with the NN and analyzing the results; fine-tuning it and tracking its evolution; and, if necessary, restarting the iteration. The geochemical ML infrastructure presented in this paper includes an integrated hyperparameter search tool, in order to assist with the iterative task of finding suitable parameters. It allows the user to fix a set of hyperparameters while varying others and compares them side by side, making it possible to graphically visualize which ones are likely to be the most suitable combination. For demonstration purposes, this section presents the evaluation of the hyperparameter search tool, while exploring the most suitable activation function, epoch, and batch size for the calcite-dolomite application presented in the next section. When attempting to establish the most suitable value for a given hyperparameter, all the others are fixed. The other neural network training routine parameters are defined as follows: 6 inputs and 5 outputs; dataset size with $45,000$ entries; 3 hidden layers with 12 neurons each; and Adam as the training algorithm optimizer and relu as the activation function. The evaluation is based on Tensorflow's metrics: loss (mean squared error or MSE), accuracy (ACC), and root mean squared error (RMSE). Table 1 presents the (a) accuracy, (b) RMSE, and (c) loss (MSE) for the different NN models evaluated during the hyperparameter search routine, when changing the number of epochs, batch size, and activation function.

The purpose of a metric is to indicate a NN's property, which is normally intended to be minimized or maximized through the iterations and learning process. Additionally, [53] is referred to for details of each activation function, which are omitted here because they fall outside the scope of this work. Moreover, the framework presented here also provides

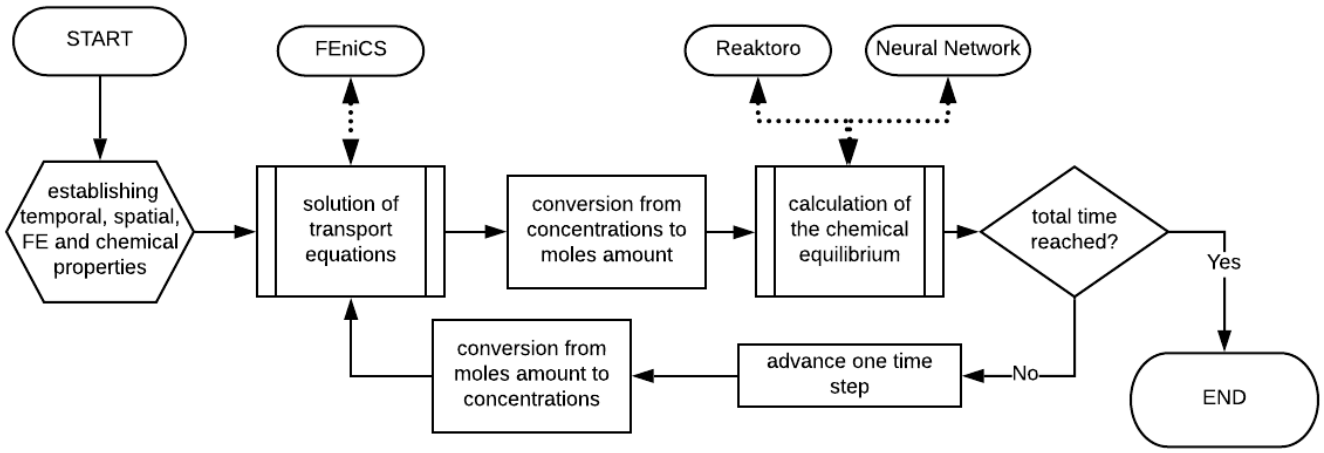


FIGURE 4. SNIA with the two possibilities for calculation of the chemical equilibrium condition: chemical solver or neural network

TABLE 1. Compilation of the NN's metrics (loss, accuracy, and RMSE) generated during the hyperparameter's search for number of epoch, batch size, and activation function.

epoch	batch	metrics		
#	size	loss	accuracy	RMSE
50	16	2.8370×10^{-5}	9.9980×10^{-1}	5.3263×10^{-3}
50	32	3.3975×10^{-5}	9.9980×10^{-1}	5.8288×10^{-3}
50	64	4.4702×10^{-5}	9.9980×10^{-1}	6.6860×10^{-3}
500	16	2.3201×10^{-5}	9.9980×10^{-1}	4.8167×10^{-3}
500	32	2.2338×10^{-5}	9.9980×10^{-1}	4.7263×10^{-3}
500	64	2.7731×10^{-5}	9.9980×10^{-1}	5.2660×10^{-3}
1000	16	2.7046×10^{-5}	9.9928×10^{-1}	5.2006×10^{-3}
1000	32	2.3455×10^{-5}	9.9980×10^{-1}	4.8430×10^{-3}
1000	64	4.9052×10^{-3}	9.6088×10^{-1}	7.0037×10^{-2}

activation	metrics		
function	loss	accuracy	RMSE
relu	2.2732×10^{-5}	9.9980×10^{-1}	4.7678×10^{-3}
softmax	1.8440×10^{-4}	9.9960×10^{-1}	1.3579×10^{-2}
tanh	3.0838×10^{-5}	9.9933×10^{-1}	5.5532×10^{-3}
sigmoid	2.3661×10^{-5}	9.9980×10^{-1}	4.8643×10^{-3}
hard sigmoid	4.9281×10^{-3}	9.6088×10^{-1}	7.0200×10^{-2}
softplus	2.4107×10^{-5}	9.9980×10^{-1}	4.9099×10^{-3}
softsign	2.4948×10^{-5}	9.9946×10^{-1}	4.9948×10^{-3}

a graphical comparison between the trained NN models to facilitate the comprehension of the impact of different hyperparameters on the overall quality of the NN model and on the simulation. Fig. 5 and Fig. 6 present the NN's metrics evaluation for the first 5 epochs.

In Fig. 7, which shows a comparison between the predicted and true values for different activation functions, the perfect identity diagonal plot represents an accurate model, in which the NN's predictions match the true values from the dataset. It is possible to verify that for this specific trial of dataset and combination of hyperparameters, the hard sigmoid activation function does not perform well (indicated by the horizontal lines). However, relu and tanh are suitable candidates (indicated by their overlap with the identity diagonal).

TABLE 2. Boundary condition (BC) and initial condition (IC) definition based on the bulk composition (independent components) of the chemical system.

Element	BC [mol]	IC [mol]
C	$2.707\,53 \times 10^{-2}$	$4.878\,00 \times 10^{-1}$
Ca	$4.878\,00 \times 10^{-1}$	$4.878\,00 \times 10^{-1}$
Cl	$6.149\,09 \times 10^{-1}$	$3.158\,36 \times 10^{-21}$
H	$3.413\,26 \times 10^{1}$	$3.506\,43 \times 10^{1}$
Mg	$3.074\,55 \times 10^{-1}$	$3.158\,36 \times 10^{-21}$
O	$7.700\,30 \times 10^{1}$	$7.735\,12 \times 10^{1}$
Si	29.9277	29.1784

IV. APPLICATIONS

A. DOLOMITIZATION APPLICATION

The benchmark presented here is inspired by [24], [38], [77], [78]. A 1M $MgCl_2$ solution flows, injected from the left boundary, into a 1D $0.5\,\mathrm{m}$ long calcite column with a fixed effective porosity of 0.32 and 50 equidistant nodes $0.01\,\mathrm{m}$ apart from each other. The simulation runs for a total time of $10,000s$ with a time step $\Delta t = 200\,\mathrm{s}$. The advective velocity is $9.375 \times 10^{-5}\,\mathrm{m\,s^{-1}}$ and the diffusion coefficient is $6.281\,25 \times 10^{-7}\,\mathrm{m^2\,s^{-1}}$.

The temperature and pressure of the chemical system are, respectively, $60°C$ (333.15 K) and $100\,\mathrm{bar}$.The independent components of the chemical system are presented in table 2. The boundary condition (BC) amounts are strong Dirichlet BC imposed at the left boundary of the domain ($x = 0.0$). The initial conditions (IC) are the initial amounts of the species throughout the domain ($0.0 < x <= 0.5$). Table 3 presents the chemical system phases and speciation.

For this benchmark, a reduced chemical system with only the main species was adopted in order to slightly reduce the chemical system's dimensionality. The transported species are Mg^{+2}, Cl^{-} and Ca^{+2}. The reactive mineral phases are calcite ($CaCO_3$) and dolomite ($CaMg(CO_3)_2$).

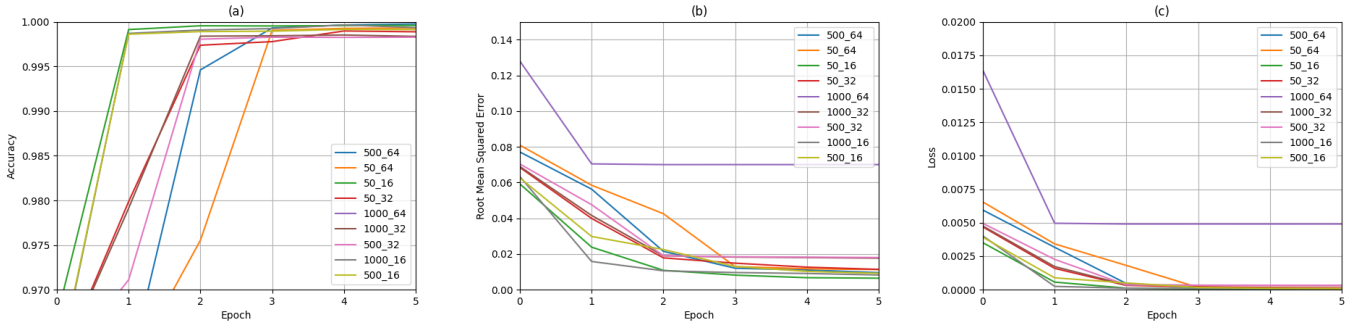


FIGURE 5. Plot evaluation of the metrics using the hyperparameter search tool with regard to number of epochs and batch size (subtitle format: epoch_batch)

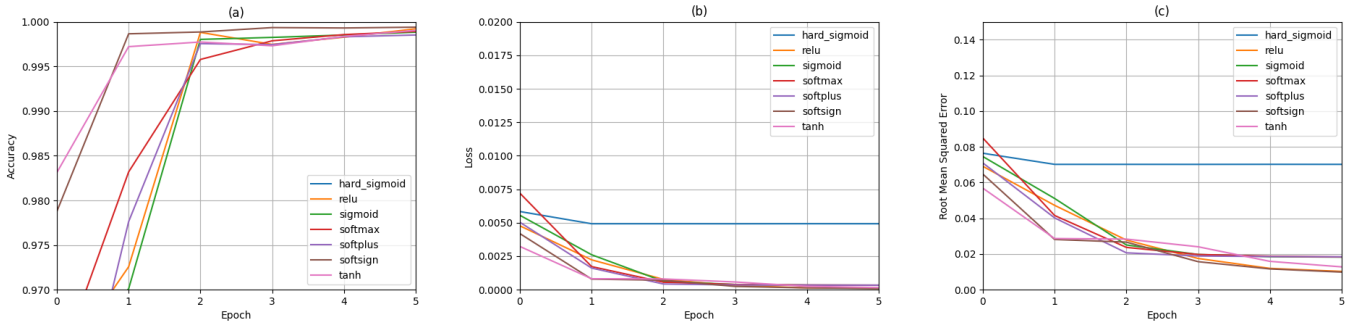


FIGURE 6. Plot evaluation of metrics using the hyperparameter search tool with regard to the optimizer

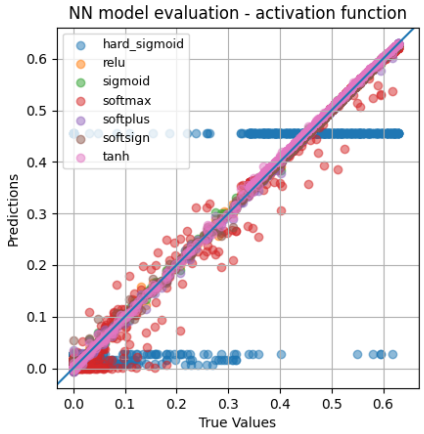


FIGURE 7. Evaluation of models with different activation functions and comparison with the identity diagonal

TABLE 3. Phases and species in the chemical composition definition.

Phase									
Aqueous							Quartz	Calcite	Dolomite
$H_2O(l)$	H^+	OH^-	Ca_2^+	Cl^-	HCO_3^-	Mg_2^+	SiO_2	$CaCO_3$	$CaMg(CO_3)_2$

The benchmark workflow is divided into three connected, sequentially ordered parts: the definition of the chemical system and the generation of the training dataset; the generation of the NN model after the learning routines and hyperparameters tuning; and, finally, the reactive transport simulation using the neural network as an emulation of the chemical solver, as well as the comparison between the results from both reactive transport approaches and the computational costs. All three parts are bundled together into a unique geochemical machine learning infrastructure presented in this paper.

The training dataset is composed of $915,000$ samples and was generated using a refined spatial discretization mesh, in order to ensure that the entire simulation domain is represented in the dataset. Each node exports its bulk composition and equilibrium results at every iteration during the simulation time, plus a variation of $\pm 5\%$ around the exact values. Each sample is composed of independent components (inputs) and equilibrated composition of transported species and minerals (outputs) for a specific time step, as presented in equation 4.

$$input = \begin{bmatrix} C \\ Ca \\ Cl \\ H \\ O \\ Mg \end{bmatrix} \quad output = \begin{bmatrix} \mathrm{Mg}^{+2} \\ \mathrm{Cl}^{-} \\ \mathrm{Ca}^{+2} \\ calcite \\ dolomite \end{bmatrix} \tag{4}$$

Quartz is present in the chemical system as an inert (non-reactive) phase solely for the purpose of porosity adjustment. For this reason, it was removed from the training dataset. Table 4 presents the hyperparameters of the DNN used in

TABLE 4. Deep neural network hyperparameter details.

Hyperparameter						
Layers	Neurons	Initializer	Optimizer	Activation function	Epoch	Batch size
3	7	uniform	adam	relu	1000	32

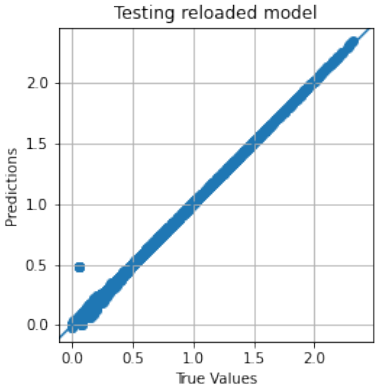


FIGURE 8. Evaluation of the DNN through a comparison between the true values and the NN predictions

the emulation of the geochemical solver for the benchmark problem.

Fig. 8 presents the model's comparison of predictions with true values. The predictions of the NN model match with a high degree of agreement the identity diagonal. An outsider point is present on the evaluation plot, but does not strongly impact the reactive transport model.

Fig. 9 presents the comparison after $10,000s$ between the reactive transport model using Reaktoro and the alternative approach using NN. On the left side, the transported species are presented, on the right side, the minerals. The inflow of $MgCl_2$, starting from the left-side boundary, causes dissolution of the existing calcite, followed by dolomite precipitation. It is possible to verify that there is a high level of agreement between the NN predictions and the full equilibrium calculation using Reaktoro. It is also interesting to observe the behavior of the aqueous Cl^-, which is a non-reactive aqueous tracer. It shows the advective-diffusive transport front without interacting with the minerals, but it follows the precipitation-dissolution front.

1) Computational cost analysis

A larger dataset, with $1.178\,212 \times 10^7$ entries, is used to analyze the computational and temporal costs between NN and Reaktoro. Fig. 10 shows that, on average, each NN prediction or Reaktoro equilibration takes, respectively, 4.063×10^{-8} s and 1.8914×10^{-4} s.

Depending on the number of nodes, the computational cost is most likely to derive from the chemical equilibration part of the model. For temporal analysis purposes, the calcite-dolomite benchmark presented above was performed again, but this time with $10,000$ nodes (instead of 50). For this variation, the time consumed in the chemical equilibrium

TABLE 5. Hyperparameter details of the ThermoFun thermodynamic neural network used for better geochemical solver emulation.

Hyperparameter						
Layers	Neurons	Initializer	Optimizer	Activation function	Epoch	Batch size
5	50	lecun uniform	adamax	relu	100	32

corresponds to approximately 95% of the total time and 5% of the transport solver. With this NN, one could achieve an increase of speed of around 4650, that is, three to four orders of magnitude, which is in agreement with [23]. Such an increase in speed would make it possible to tackle even more complex problems.

B. THERMODYNAMIC NEURAL NETWORK

The use of a DNN as an emulated chemical solver, as presented in section IV-A, brings with it certain limitations. For example, the reaction details and substance's thermodynamic properties are no longer available, unless they are explicit outputs of the DNN that had been previously defined when preparing the dataset for the training routine. In order to provide a more comprehensive geochemical ML reactive transport infrastructure, this paper also presents an approach to better emulating the chemical solver by providing thermodynamic data with an auxiliary DNN. ThermoFun [75] is used to generate a species-related thermodynamic database, which, based on the temperature and pressure, provides the Gibbs energy, enthalpy, entropy, heat capacity, and volume. The details of the neural network are presented in the equation 5. The NN details are presented in 5.

$$input = \begin{bmatrix} temperature \\ pressure \end{bmatrix} output = \begin{bmatrix} Gibbs\ energy \\ enthalpy \\ entropy \\ heatcapacity \\ volume \end{bmatrix} \tag{5}$$

Fig. 11 shows on the left and the right, respectively, the comparison between different optimizers and the evaluation of the finished, trained model.

With an auxiliary neural network focused on thermodynamic properties, the reactive transport simulation can access thermodynamic information from the substances and reactions while acquiring the necessary performance and flexibility, as a result of its not depending on the traditional chemical solver. These thermodynamic properties can be used, for example, to detect problems with the evolution of the reactive transport simulation, because the approach presented in this work is not fully mass conservative. An extension of the developed framework presented in this work would be to perform smart adjustments on the main neural network chemical equilibration predictions based on the auxiliary thermodynamic data neural network: a thermodynamically consistent neural network architecture.

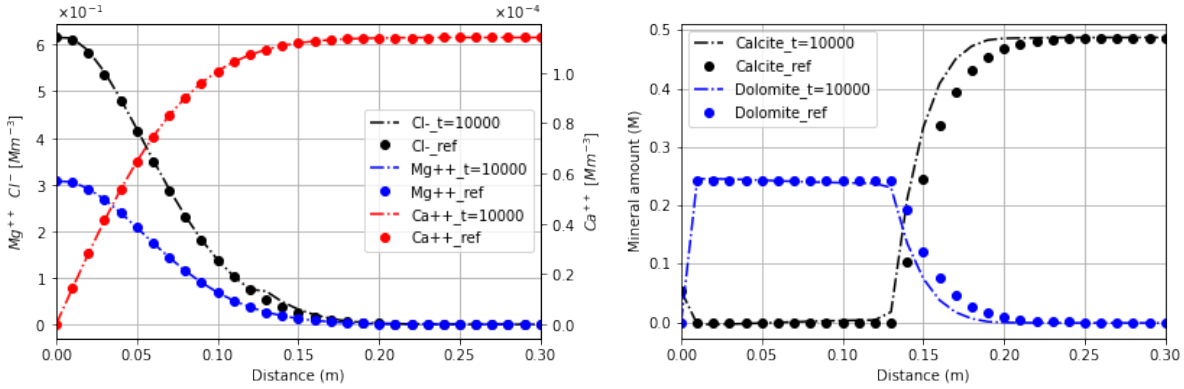


FIGURE 9. Aqueous transported species comparison between DNN and Reaktoro

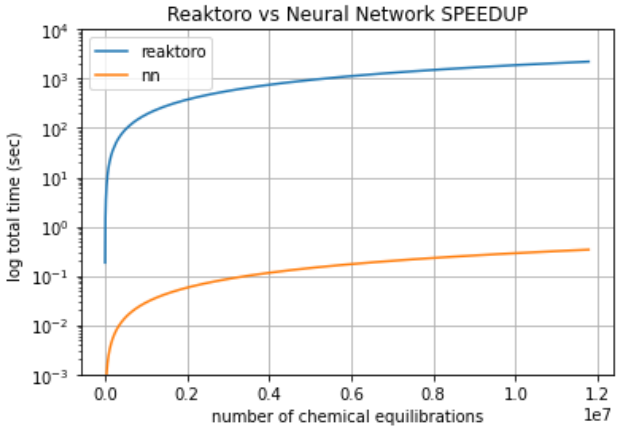


FIGURE 10. Comparison of total time needed to perform $1.178\,212 \times 10^7$ predictions (DNN) and equilibrations (Reaktoro)

V. SUMMARY AND CONCLUSION

Machine learning and neural network algorithms are innovative techniques that are being widely applied. They help us deal with and understand the enormous amount of data generated nowadays. This paper presents a fully integrated state-of-the-art machine learning infrastructure that enables the exploration and design of reactive transport models using neural networks instead of direct couplings with the traditional chemical solver. This represents an innovative engine that brings machine learning possibilities to reactive transport modeling. The machine learning infrastructure presented in this paper makes it possible to: (i) generate training data sets using a geochemical solver (e.g., GEMS or Reaktoro), also as a result of a simple full reactive transport simulation; (ii) train and test the neural networks and explore the multiple hyperparameters, in order to verify which are the most appropriate; (iii) validate the resulting neural network against the reactive transport benchmark and compare its performance with that of the traditional geochemical solver. This infrastructure allows us to overcome two severe and inherent limitations of chemical solvers: the non-parallelizable nature of the chemical solvers and their non-transferability to various computer architectures. At the same time, the neural network is, per se, highly parallelizable and may be deployed in all existing computer architectures both in software and recently also in hardware mode (e.g., neural engines of M1 Apple CPU, Intel, GPU tensor cores). Moreover, it takes advantage of high-performance computing, while maintaining accuracy and a flexible description of the physical and chemical phenomena. Among other challenges, there is no predefined recipe for finding the most suitable hyperparameter for neural networks. For this reason, a single framework that makes it possible to define, train, and test hyperparameters, as well as to compare the results of a neural network with those of a chemical solver has immense potential, when it comes to designing and solving complicated reactive transport simulations. We stress that it is necessary for there to be an initial phase in which the dataset is generated and the neural network trained. However, the results for the benchmark presented here display an extremely high level of agreement and, even though the increase in speed is not generalizable to all applications (since it depends heavily on the chemical system's definition and training routine), it demonstrates the possibility of taking reactive transport modeling to the next level of high-performance computing and of addressing more complex problems in full detail. For complex or high-performance computing simulations, the training cost is a small fraction of the overall computational time, providing a one to three orders of magnitude (depending on the mass transport solver) overall acceleration to find the solution. In future research, the current machine learning infrastructure could generate auxiliary neural networks that emulate other processes from the chemical solvers. For example, in a similar way to the proof-of-concept, in which a DNN was used to provide thermodynamic properties based on ThermoFun, one could integrate an additional neural network, in order to describe non-ideal solid solutions. Additionally, by taking advantage of the Jupyter notebook approach, it would be possible to make the geochemical machine learning infrastructure presented in this work available to many users via an online

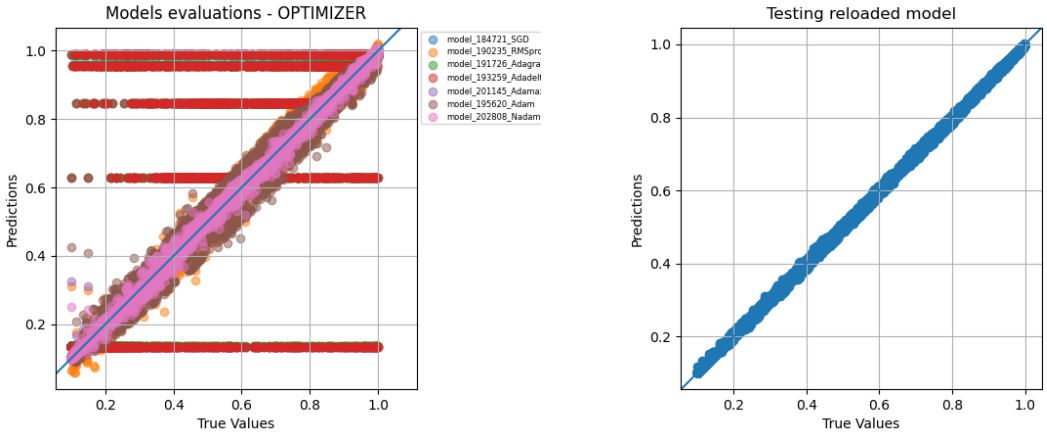


FIGURE 11. Comparison with different optimizers and final evaluation of the ThermoFun DNN model

web app without code changes. Finally, the modular nature of the framework allows different testing combinations of chemical solvers, neural network architectures, and machine learning techniques, as well as of different mass transport solvers.

ACKNOWLEDGMENT

This work was supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 15.0186-2. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Swiss Government. Authors receive partial financial support from Nagra. The research leading to these results has received funding from the European Union's Horizon 2020 Research and Training Programme of the European Atomic Energy Community (EURATOM) (H2020 - NFRP - 2014 / 2015) under grant agreement n^o662147 (CEBAMA).

REFERENCES

[1] C. I. Steefel, D. J. DePaolo, and P. C. Lichtner, "Reactive transport modeling: An essential tool and a new research approach for the Earth sciences," Earth and Planetary Science Letters, vol. 240, pp. 539–558, 2005.

[2] T. Xu, E. Sonnenthal, N. Spycher, and K. Pruess, "TOUGHREACT - A simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and CO2geological sequestration," Computers and Geosciences, vol. 32, pp. 145–165, 2006.

[3] G. Kosakowski and N. Watanabe, "OpenGeoSys-Gem: A numerical tool for calculating geochemical and porosity changes in saturated and partially saturated media," Physics and Chemistry of the Earth, vol. 70-71, pp. 138–149, 2014.

[4] N. Garapati, B. M. Adams, J. M. Bielicki, P. Schaedle, J. B. Randolph, T. H. Kuehn, and M. O. Saar, "A Hybrid Geothermal Energy Conversion Technology - A Potential Solution for Production of Electricity from Shallow Geothermal Resources," in Energy Procedia, vol. 114, pp. 7107–7117, Elsevier Ltd, 2017.

[5] C. I. Steefel, C. A. J. Appelo, B. Arora, D. Jacques, T. Kalbacher, O. Kolditz, V. Lagneau, P. C. Lichtner, K. U. Mayer, J. C. L. Meeussen, S. Molins, D. Moulton, H. Shao, J. Šimůnek, N. Spycher, S. B. Yabusaki, and G. T. Yeh, "Reactive transport codes for subsurface environmental simulation," Computational Geosciences, vol. 19, no. 3, pp. 445–478, 2015.

[6] L. H. Damiani, G. Kosakowski, M. A. Glaus, and S. V. Churakov, "A framework for reactive transport modeling using FEniCS–Reaktoro: governing equations and benchmarking results," Computational Geosciences, vol. 24, no. 3, pp. 1071–1085, 2020.

[7] A. M. M. Leal, D. A. Kulik, and M. O. Saar, "Ultra-Fast Reactive Transport Simulations When Chemical Reactions Meet Machine Learning: Chemical Equilibrium," arXiv, 2017.

[8] S. Bejaoui and B. Bary, "Modeling of the link between microstructure and effective diffusivity of cement pastes using a simplified composite model," Cement and Concrete Research, vol. 37, no. 3, pp. 469–480, 2007.

[9] Y. Li, M. Niewiadomski, E. Trujillo, and S. P. Sunkavalli, "Tougher: A user-friendly graphical interface for TOUGHREACT," Computers & Geosciences, vol. 37, pp. 775–782, 2011.

[10] S. W. Wang, H. Levy, G. Li, and H. Rabitz, "Fully equivalent operational models for atmospheric chemical kinetics within global chemistry-transport models," Journal of Geophysical Research: Atmospheres, vol. 104, pp. 30417–30426, 1999.

[11] K. Fukumoto and Y. Ogami, "Combustion simulation technique for reducing chemical mechanisms using look-up table of chemical equilibrium calculations: Application to CO-H 2-air turbulent non-premixed flame," Computers and Fluids, vol. 66, pp. 98–106, 2012.

[12] B. Fiorina, D. Veynante, and S. Candel, "Modeling combustion chemistry in large eddy simulation of turbulent flames," Flow, Turbulence and Combustion, vol. 94, pp. 3–42, 2015.

[13] Y. Huang, H. Shao, E. Wieland, O. Kolditz, and G. Kosakowski, "A new approach to coupled two-phase reactive transport simulation for long-term degradation of concrete," Construction and Building Materials, vol. 190, pp. 805–829, 2018.

[14] Y. Huang, H. Shao, E. Wieland, O. Kolditz, and G. Kosakowski, "Two-phase transport in a cemented waste package considering spatio-temporal evolution of chemical conditions," npj Materials Degradation, vol. 5, pp. 1–14, 2021.

[15] D. A. Winkler, "Neural networks as robust tools in drug lead discovery and development," Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology, vol. 27, no. 2, pp. 139–167, 2004.

[16] G. Li, H. Tang, Y. Sun, J. Kong, G. Jiang, D. Jiang, B. Tao, S. Xu, and H. Liu, "Hand gesture recognition based on convolution neural network," Cluster Computing, vol. 22, pp. 2719–2729, 2019.

[17] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep Convolutional Neural Network for Image Deconvolution," in Advances in Neural Information Processing Systems (Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and Q. K. Weinberger, eds.), Curran Associates, Inc., 2014.

[18] M. Torrisi, G. Pollastri, and Q. Le, "Deep learning methods in protein structure prediction," 2020.

[19] M. R. Gadekar and M. M. Ahammed, "Modelling dye removal by adsorption onto water treatment residuals using combined response surface methodology-artificial neural network approach," Journal of Environmental Management, vol. 231, pp. 241–248, 2019.

[20] F. C. Christo, A. R. Masri, and E. M. Nebot, "Artificial neural network

implementation of chemistry with pdf simulation of H2/CO2 flames," Combustion and Flame, vol. 106, pp. 406–427, 1996.

[21] D. Guérillot and J. Guérillot, "Geochemical Equilibrium determination using an artificial neural network in compositional reservoir flow simulation," 16th European Conference on the Mathematics of Oil Recovery, ECMOR 2018, pp. 697–707, 2018.

[22] E. Laloy and D. Jacques, "Emulation of CPU-demanding reactive transport models: a comparison of Gaussian processes, polynomial chaos expansion, and deep neural networks," Computational Geosciences, vol. 23, pp. 1193–1215, 2019.

[23] N. I. Prasianakis, R. Haller, M. Mahrous, J. Poonoosamy, W. Pfingsten, and S. V. Churakov, "Neural network based process coupling and parameter upscaling in reactive transport simulations," Geochimica et Cosmochimica Acta, vol. 291, pp. 126–143, 2020.

[24] A. M. Leal, S. Kyas, D. A. Kulik, and M. O. Saar, "Accelerating Reactive Transport Modeling: On-Demand Machine Learning Algorithm for Chemical Equilibrium Calculations," Transport in Porous Media, vol. 133, pp. 161–204, 2020.

[25] J. Lapeyre, T. Han, B. Wiles, H. Ma, J. Huang, G. Sant, and A. Kumar, "Machine learning enables prompt prediction of hydration kinetics of multicomponent cementitious systems," Scientific Reports, vol. 11, p. 3922, 2021.

[26] M. Mitchell, "Why AI is Harder Than We Think," in GECCO '21: Proceedings of the Genetic and Evolutionary Computation Conference, p. 3, 2021.

[27] E. Haber and L. Ruthotto, "Stable architectures for deep neural networks," Inverse Problems, vol. 34, p. 014004, 2018.

[28] X. Ying, "An Overview of Overfitting and its Solutions," Journal of Physics: Conference Series, vol. 1168, 2019.

[29] A. M. Leal, M. J. Blunt, and T. C. LaForce, "Efficient chemical equilibrium calculations for geochemical speciation and reactive transport modelling," Geochimica et Cosmochimica Acta, vol. 131, pp. 301–322, 2014.

[30] M. S. Alnaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells, "The FEniCS Project Version 1.5," Archive of Numerical Software, vol. 3, pp. 9–23, 2015.

[31] M. S. Alnæs, A. Logg, K. B. Ølgaard, M. E. Rognes, and G. N. Wells, "Unified form language: A domain-specific language for weak formulations of partial differential equations," ACM Transactions on Mathematical Software, vol. 40, pp. 1–37, 2014.

[32] C. I. Steefel and K. Maher, "Fluid-rock interaction: A reactive transport approach," Reviews in mineralogy and geochemistry, vol. 70, no. 1, pp. 485–532, 2009.

[33] J. Bear and Y. Bachmat, Introduction to Modeling of Transport Phenomena in Porous Media. Dordrecht: Springer Netherlands, 1990.

[34] J. G. Charney, R. Fjortoft, and J. V. Neumann, "Numerical Integration of the Barotropic Vorticity Equation," Tellus, vol. 2, pp. 237–254, 1950.

[35] P. Wesseling, "Von Neumann stability conditions for the convection-diffusion equation," IMA Journal of Numerical Analysis, vol. 16, no. 4, pp. 583–598, 1996.

[36] A. D. Daus and E. O. Frind, "An alternating direction galerkin technique for simulation of contaminant transport in complex groundwater systems," Water Resources Research, vol. 21, pp. 653–664, 1985.

[37] M. Renardy and R. C. Rogers, An Introduction to Partial Differential Equations, vol. 13 of Texts in Applied Mathematics. Springer-Verlag New York, 2 ed., 2004.

[38] D. Kulik, T. Wagner, S. V. Dmytrieva, G. Kosakowski, F. F. Hingerl, K. V. Chudnenko, and U. R. Berner, "GEM-Selektor geochemical modeling package: Revised algorithm and GEMS3K numerical kernel for coupled simulation codes," Computational Geosciences, vol. 17, no. 1, pp. 1–24, 2013.

[39] F. Morel and J. Morgan, "A Numerical Method for Computing Equilibria in Aqueous Chemical Systems," Environmental Science and Technology, vol. 6, pp. 58–67, 1972.

[40] T. J. Wolery and L. J. Walters, "Calculation of equilibrium distributions of chemical species in aqueous solutions by means of monotone sequences," Journal of the International Association for Mathematical Geology, vol. 7, pp. 99–115, 1975.

[41] W. Smith and R. Missen, Chemical Reaction Equilibrium Analysis: Theory and Algorithms. Wiley, 1982.

[42] M. H. Reed, "Calculation of multicomponent chemical equilibria and reaction processes in systems involving minerals, gases and an aqueous phase," Geochimica et Cosmochimica Acta, vol. 46, pp. 513–528, 1982.

[43] P. Waage and C. M. Gulberg, "Studies concerning affinity," Journal of Chemical Education, vol. 63, no. 12, pp. 1044–1047, 1986.

[44] J. Borge, "Reviewing some crucial concepts of gibbs energy in chemical equilibrium using a computer-assisted, guided-problem-solving approach," Journal of Chemical Education, vol. 92, pp. 296–304, 2015.

[45] R. E. Ferner and J. K. Aronson, "Cato Guldberg and Peter Waage, the history of the Law of Mass Action, and its relevance to clinical pharmacology," 2016.

[46] A. M. Leal, D. A. Kulik, W. R. Smith, and M. O. Saar, "An overview of computational methods for chemical equilibrium and kinetic calculations for geochemical and reactive transport modeling," Pure and Applied Chemistry, vol. 89, pp. 597–643, 2017.

[47] T. Wolery, "EQ3/6, a software package for geochemical modeling of aqueous systems: Package overview and installation guide (Version 7.0)," tech. rep., Lawrence Livermore National Laboratory (LLNL), Livermore, CA (United States), 1992.

[48] D. Parkhurst and C. Appelo, "Description of Input and Examples for PHREEQC Version 3—A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations," U.S. Geological Survey Techniques and Methods, p. 497 p, 2013.

[49] C. M. Bethke, Geochemical and biogeochemical reaction modeling: Second edition, vol. 9780521875547. Cambridge University Press, 2007.

[50] M. Blasco, M. J. Gimeno, and L. F. Auqué, "Comparison of Different Thermodynamic Databases used in a Geothermometrical Modelling Calculation," Procedia Earth and Planetary Science, vol. 17, pp. 120–123, 2017.

[51] A. L. Walter, E. O. Frind, D. W. Blowes, C. J. Ptacek, and J. W. Molson, "Modeling of multicomponent reactive transport in groundwater: 1. Model development and evaluation," Water Resources Research, vol. 30, no. 11, pp. 3137–3148, 1994.

[52] J. Carrayrou, R. Mosé, and P. Behra, "Operator-splitting procedures for reactive transport and comparison of mass balance errors," Journal of Contaminant Hydrology, vol. 68, pp. 239–268, 2004.

[53] J. Heaton, Applications of Deep Neural Networks. arxiv 2009.05673v2, 2021.

[54] A. G. Ivakhnenko and V. G. Lapa, Cybernetics and forecasting techniques, vol. 8. New York, American Elsevier Pub. Co., 1967.

[55] A. G. Ivakhnenko, "Polynomial Theory of Complex Systems," IEEE Transactions on Systems, Man and Cybernetics, vol. 1, no. 4, pp. 364–378, 1971.

[56] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biological Cybernetics, vol. 36, pp. 193–202, 1980.

[57] R. Dechter, "Learning while searching in constraint-satisfaction-problems," in AAAI'86: Proceedings of the Fifth AAAI National Conference on Artificial Intelligence, pp. 178–183, 1986.

[58] G. E. Hinton, S. Osindero, and Y. Teh, "A Fast Learning Algorithm for Deep Belief Nets Yee-Whye Teh," Neural Computation, vol. 18, pp. 1527–1554, 2006.

[59] A. Koutsoukas, K. J. Monaghan, X. Li, and J. Huan, "Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data," Journal of Cheminformatics, vol. 9, p. 42, 2017.

[60] Q. Li and S. Hao, "An optimal control approach to deep learning and applications to discrete-weight neural networks," in 35th International Conference on Machine Learning, ICML 2018, vol. 7, pp. 4642–4661, International Machine Learning Society (IMLS), 2018.

[61] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, no. 6088, pp. 533–536, 1986.

[62] S. M. Bohte, J. N. Kok, and H. La Poutré, "Error-backpropagation in temporally encoded networks of spiking neurons," Neurocomputing, vol. 48, pp. 17–37, 2002.

[63] S. M. Bohte, "Error-backpropagation in networks of fractionally predictive spiking neurons," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 6791 LNCS, pp. 60–68, Springer, Berlin, Heidelberg, 2011.

[64] M. Mirsadeghi, M. Shalchian, S. R. Kheradpisheh, and T. Masquelier, "STiDi-BP: Spike time displacement based error backpropagation in multilayer spiking neural networks," Neurocomputing, vol. 427, pp. 131–140, 2021.

[65] H. Larochelle, Y. Bengio, J. Louradour, and P. Lamblin, "Exploring Strategies for Training Deep Neural Networks," Journal of Machine Learning Research, vol. 10, pp. 1–40, 2009.
[66] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Training Very Deep Networks," Advances in Neural Information Processing Systems, vol. 2015-January, pp. 2377–2385, 2015.
[67] I. A. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," Journal of Microbiological Methods, vol. 43, pp. 3–31, 2000.
[68] I. Arel, D. Rose, and T. Karnowski, "Deep machine learning-A new frontier in artificial intelligence research," IEEE Computational Intelligence Magazine, vol. 5, pp. 13–18, 2010.
[69] A. Effland, E. Kobler, K. Kunisch, and T. Pock, "Variational Networks: An Optimal Control Approach to Early Stopping Variational Methods for Image Restoration," Journal of Mathematical Imaging and Vision, vol. 62, pp. 396–416, 2020.
[70] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014.
[71] L. N. Smith, "A disciplined approach to neural network hyper-parameters: part 1 – learning rate, batch size, momentum, and weight decay," 2018.
[72] H. Zhong, Z. Chen, C. Qin, Z. Huang, V. W. Zheng, T. Xu, and E. Chen, "Adam revisited: a weighted past gradients perspective," Frontiers of Computer Science, vol. 14, p. 145309, 2020.
[73] G. Hinton, N. Srivastava, and K. Swersky, "Neural Networks for Machine Learning - Lecture 6 RMSprop," tech. rep., cs.toronto.edu, 2012.
[74] M. D. Zeiler, "ADADELTA: An Adaptive Learning Rate Method," arxiv 1212.5701, 2012.
[75] G. D. Miron, A. M. Leal, and A. Yapparova, "Thermodynamic Properties of Aqueous Species Calculated Using the HKF Model: How Do Different Thermodynamic and Electrostatic Models for Solvent Water Affect Calculated Aqueous Properties?," Geofluids, vol. 2019, p. 24, 2019.
[76] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, and C. Willing, "Jupyter Notebooks—a publishing format for reproducible computational workflows," in Positioning and Power in Academic Publishing: Players, Agents and Agendas - Proceedings of the 20th International Conference on Electronic Publishing, ELPUB 2016, pp. 87–90, IOS Press BV, 2016.
[77] P. Engesgaard and K. L. Kipp, "A geochemical transport model for redox-controlled movement of mineral fronts in groundwater flow systems: A case of nitrate removal by oxidation of pyrite," Water Resources Research, vol. 28, pp. 2829–2843, 1992.
[78] H. Shao, S. V. Dmytrieva, O. Kolditz, D. Kulik, W. Pfingsten, and G. Kosakowski, "Modeling reactive transport in non-ideal aqueous–solid solution system," Applied Geochemistry, vol. 24, no. 7, pp. 1287–1300, 2009.

Chapter 5

Summary and outlook

5.1 Summary

Reactive transport modelling is the basic tool used in modern computational geochemistry. In practice, its predictive accuracy in relation to complex natural phenomena and its ability to account for the relevant process couplings is often limited by the performance and flexibility of the available numerical tools. In this research project, an open-source simulation framework was developed, which provides a versatile modelling environment, while simultaneously ensuring full control over the numerical methods. Taken as a whole, it represents a goal-oriented and easy-to-learn framework for reactive transport modelling, one that combines various computing engines optimized for state-of-the-art hardware platforms that are designed to describe physical and chemical processes. Basic reactive transport equations that are predefined for typical geochemical problems are easily scalable by means of the coupling of additional physical and chemical phenomena using a simple, intuitive, high-level programming language. To cite just one example, the framework has been benchmarked for electrochemically coupled multicomponent diffusion across charged or uncharged membranes, as well as in combination with complex geochemical reactions (Chapter 2). This flexible and easily customizable framework makes possible the further development of the model, as well as geochemical applications that are not easily implemented in existing reactive transport codes. Accordingly, the framework was used to investigate and model the data collected from the Hydrogen Transfer (HT) experiment (Chapter 3) conducted at the Mont Terri underground rock laboratory in Switzerland. Modelling the HT experiment requires a coupled geochemical setup, in which the evolution of a borehole gas and water phase sets boundary conditions for the transport of dissolved gases in the surrounding Opalinus Clay formation. In turn, the evolution of the borehole gas phase is back-coupled to the diffusive transport of dissolved gases in the Opalinus Clay, the flow of formation water into the borehole, the extraction of accumulated water from the borehole, the injection of hydrogen and other gases into the borehole, and the microbially mediated hydrogen reaction.

The customized model accurately reproduces both the observed temporal evolution of gas composition in the borehole and the composition of the extracted water. The modelling study demonstrated that fluxes of non-reactive gases across the borehole wall are primarily constrained by the diffusive transport of dissolved gases in Opalinus Clay. The estimated effective diffusion coefficients displayed a high level of agreement with those measured by previous laboratory and

field studies on Opalinus Clay conducted at Mont Terri. The experimentally observed reduction in sulphate concentration is controlled by the interplay between the inflow of Opalinus Clay water, water extraction via an automatic extraction system, and the microbially driven consummation of hydrogen via sulphate reduction.

For many geochemical applications, chemical equilibria calculations are one of the key processes controlling the evolution of the system. Unfortunately, the chemical solvers often represent a computational bottleneck for reactive transport modelling. The reactive transport framework developed here provides a straightforward, modular method for coupling external chemical solvers. Machine learning (ML) and artificial intelligence (AI) are innovative technologies that are being widely applied in both research and industry. They facilitate the processing and understanding of the enormous amounts of data being generated nowadays. For example, the use of ML makes it possible to improve the performance of geochemical calculations by several orders of magnitude, without compromising modelling accuracy. Chapter 4 of the thesis describes an innovative modelling engine that applies ML techniques to reactive transport modelling in a novel way.

The developd machine learning toolbox presented and tested in chapter 4 makes it possible: (i) to generate training data sets using a geochemical solver (e.g., GEMS or Reaktoro), as the result of a simple full reactive transport simulation; (ii) to train and test the neural networks and explore the multiple hyperparameters, in order to verify which are the most appropriate; (iii) to validate the resulting neural network against the reactive transport benchmark and to compare its performance with that of the traditional geochemical solver. Applications of the ML-based approach enables us to overcome two severe limitations inherent in chemical solvers: (i) their non-parallelizable nature and (ii) their non-transferability to various computer architectures. By contrast, the neural network is highly parallelizable and can be deployed in all existing computer architectures, both in software and, more recently, in hardware mode (e.g., the neural engines of M1 Apple CPU, Intel, GPU tensor cores). These advantages were demonstrated in order to compare reactive transport models using neuronal networks with traditional reactive models that use direct coupling with the traditional chemical solvers.

Moreover, neural network–based chemical-solver surrogate models make it possible to take advantage of high-performance computing, while maintaining accuracy and preserving a flexible description of the physical and chemical phenomena. Finally, the modular nature of the framework allows for different testing combinations of chemical solvers, neural network architectures, and machine learning techniques, as well as different mass transport solvers.

5.2 Outlook

The developer experience and numerous numerical benchmarks presented in this thesis suggest that the future development of reactive transport codes will be driven by the need to be more realistic regarding the description and coupling of the processes and to consider the different spatial and temporal scales of these processes. Furthermore, the multiscale aspect, in particular, anticipates the need for the development of faster and more flexible computer code, capable of exploiting the rapidly evolving modern computer architecture.

The framework developed here was specifically designed to be easily extendable, to allow

the coupling of other physical processes, and to create a thermo-hydro-chemical-mechanical (THMC) framework. The insertion of thermal, hydraulic, and continuum mechanical processes (e.g., deformation) is straightforward. They can be implemented in the form of additional coupled Partial Differential Equations (PDE) in the FEnICS software. Other mechanical processes (e.g., crack formation) will require the additional coupling of specialized codes into the framework.

The chemical description and equilibrium calculation of the natural systems must be reliable and based on accurate (and accessible) thermodynamic data. It could be that, depending on the application, a specific choice of chemical solver is more appropriate — for example, law-of-mass action (LMA) or Gibbs Energy Minimization (GEM). Therefore, it is recommendable to analyze the availability of thermodynamic data needed to describe the system.

Pore and molecular-scale phenomena are of fundamental importance for the understanding mechanism of reactive transport processes in porous media. However, consideration of pore-scale processes in large continuum-scale models requires some level of abstraction and approximation. Therefore, in practice, pore-scale phenomena are considered in large-scale models in the form of effective parametric relationships.

The spatial heterogeneities and characteristic time scale of coupled dynamic processes have an enormous impact on the overall stability, robustness, and applicability of the reactive transport simulations. Small-scale phenomena usually impose limitations on the integration time scale and model resolution, as well as heavily impacting the overall computational performance.

Considering the recent evolution of software and hardware, the reactive transport codes must take advantage of the available computational power by using parallelization and HPC techniques in order to improve performance and enable a more detailed description of the physical and chemical processes. A general problem in software development is the rapid evolution of new hardware. Therefore, modularity and the possibility of quickly exchanging existing software is needed, if it cannot be effectively used on new computing platforms.

Moreover, given that the chemical equilibrium calculation is usually the most time-consuming task for the models, it is necessary to consider feasibility, by taking into account the computational costs of the reactive transport simulations and, if necessary, possible simplifications in order to enable performance gains. For example, for large-scale simulations as chemical engineering applications with different geometries of the modelling domain, but with a similar chemical setup, one should also consider using surrogate chemical solvers, such as neural networks or lookup tables.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X., 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. Technical report, Google Research.

Al Reda, S. M., Yu, C., Berthe, G., and Matray, J. M., 2020. Study of the permeability in the Opalinus clay series (Mont Terri - Switzerland) using the steady state method in Hassler cell. Journal of Petroleum Science and Engineering, 184.

Alexander, W. R., Reijonen, H. M., and McKinley, I. G., 2015. Natural analogues: studies of geological processes relevant to radioactive waste disposal in deep geological repositories. Swiss Journal of Geosciences, 108(1):75–100.

Alnaes, M. S., Logg, A., Oelgaard, K. B., Rognes, M. E., and Wells, G. N., 2012. Unified Form Language: A domain-specific language for weak formulations of partial differential equations. ACM Transactions on Mathematical Software, 40(2).

Alnaes, M. S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M. E., and Wells, G. N., 2015. The FEniCS Project Version 1.5. Archive of Numerical Software, 3(100):9–23.

Appelo, C. A. J., Verweij, E., and Schäfer, H., 1998. A hydrogeochemical transport model for an oxidation experiment with pyrite/calcite/exchangers/organic matter containing sand. Applied Geochemistry, 13(2):257–268.

Appelo, C. A. J., Van Loon, L. R., and Wersin, P., 2010. Multicomponent diffusion of a suite of tracers (HTO, Cl, Br, I, Na, Sr, Cs) in a single sample of Opalinus Clay. Geochimica et Cosmochimica Acta, 74(4):1201–1219.

Archie, G. E., 1942. The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics. Transactions of the AIME, 146(1):54–62.

Balay, S., Abhyankar, S., Adams, M. F., Brown, J., Brune, P., Buschelman, K., Dalcin, L., Dener, A., Eijkhout, V., Gropp, W. D., Kaushik, D., Knepley, M. G., May, D. A., McInnes, L. C.,

Mills, R. T., Munson, T., Rupp, K., Sanan, P., Smith, B. F., Zampini, S., Zhang, H., and Zhang, H., 2021a. PETSc Users Manual. Technical Report ANL-95/11 - Revision 3.15, Argonne National Laboratory.

Balay, S., Abhyankar, S., Adams, M. F., Brown, J., Brune, P., Buschelman, K., Dalcin, L., Dener, A., Eijkhout, V., Gropp, W. D., Kaushik, D., Knepley, M. G., May, D. A., McInnes, L. C., Mills, R. T., Munson, T., Rupp, K., Sanan, P., Smith, B. F., Zampini, S., Zhang, H., and Zhang, H. PETSc Web page, 2021b.

Barry, D. A., Prommer, H., Miller, C. T., Engesgaard, P., Brun, A., and Zheng, C., 2002. Modelling the fate of oxidisable organic contaminants in groundwater. Advances in Water Resources, 25(8-12):945–983.

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

Bear, J. and Bachmat, Y., 1990. Introduction to Modeling of Transport Phenomena in Porous Media. Springer Netherlands, Dordrecht.

Bear, J. and Verruijt, A., 1987. Modeling Groundwater Flow and Pollution. Springer Netherlands, Dordrecht.

Berkowitz, B., Dror, I., Hansen, S. K., and Scher, H., 2016. Measurements and models of reactive transport in geological media. Reviews of Geophysics, 54(4):930–986.

Bilke, L., Flemisch, B., Kalbacher, T., Kolditz, O., Helmig, R., and Nagel, T., 2019. Development of Open-Source Porous Media Simulators: Principles and Experiences. Transport in Porous Media, 130(1):337–361.

Blechschmidt, I. and Vomvoris, S. 5 - Relevance of underground rock laboratories for deep geological repository programs. In Apted, M. J. and Ahn, J., editors, Geological Repository Systems for Safe Disposal of Spent Nuclear Fuels and Radioactive Waste (Second Edition), Woodhead Publishing Series in Energy, pages 113–142. Woodhead Publishing, second edi edition, 2017.

Bos, M. and Meershoek, H. Q., 1972. A computer program for the calculation of equilibrium concentrations in complex systems. Analytica Chimica Acta, 61(2):185–199.

Bossart, P., Meier, P. M., Moeri, A., Trick, T., and Mayor, J. C., 2002. Geological and hydraulic characterisation of the excavation disturbed zone in the Opalinus Clay of the Mont Terri Rock Laboratory. Engineering Geology, 66(1-2):19–38.

Bossart, P., Bernier, F., Birkholzer, J., Bruggeman, C., Connolly, P., Dewonck, S., Fukaya, M., Herfort, M., Jensen, M., Matray, J. M., Mayor, J. C., Moeri, A., Oyama, T., Schuster, K., Shigeta, N., Vietor, T., and Wieczorek, K., 2017. Mont Terri rock laboratory, 20 years of research: introduction, site characteristics and overview of experiments. Swiss Journal of Geosciences, 110(1):3–22.

Bradbury, M. and Baeyens, B., 2009. Sorption modelling on illite. Part II: Actinide sorption and linear free energy relationships. Geochimica et Cosmochimica Acta, 73(4):1004–1013.

Carrayrou, J., Mosé, R., and Behra, P., 2004. Operator-splitting procedures for reactive transport and comparison of mass balance errors. Journal of Contaminant Hydrology, 68(3):239–268.

Charney, J. G., Fjörtoft, R., and Neumann, J. V., 1950. Numerical Integration of the Barotropic Vorticity Equation. Tellus, 2(4):237–254.

Christo, F. C., Masri, A. R., and Nebot, E. M., 1996. Artificial neural network implementation of chemistry with pdf simulation of H_2/CO_2 flames. Combustion and Flame, 106(4):406–427.

Crerar, D. A., 1975. A method for computing multicomponent chemical equilibria based on equilibrium constants. Geochim. Cosmochim. Acta, 39:1375–1384.

Dalcin, L. D., Paz, R. R., Kler, P. A., and Cosimo, A., 2011. Parallel distributed computing using Python. Advances in Water Resources, 34(9):1124–1139.

Damiani, L. H., Kosakowski, G., Glaus, M. A., and Churakov, S. V., 2020. A framework for reactive transport modeling using FEniCS–Reaktoro: governing equations and benchmarking results. Computational Geosciences, 24(3):1071–1085.

De Windt, L., Pellegrini, D., and Van Der Lee, J., 2004. Coupled modeling of cement/claystone interactions and radionuclide migration. Journal of Contaminant Hydrology, 68(3-4):165–182.

Epstein, N., 1989. On tortuosity and the tortuosity factor in flow and diffusion through porous media. Chemical Engineering Science, 44(3):777–779.

Flukiger, F. and Bernard, D., 2009. A new numerical model for pore scale dissolution of calcite due to CO_2 saturated water flow in 3D realistic geometry: Principles and first results. Chemical Geology, 265(1-2):171–180.

Gaucher, E. C. and Blanc, P., 2006. Cement/clay interactions - A review: Experiments, natural analogues, and modeling. Waste Management, 26(7):776–788.

Ge, X., Schaap, M., Kranenburg, R., Segers, A., Jan Reinds, G., Kros, H., and De Vries, W., 2020. Modeling atmospheric ammonia using agricultural emissions with improved spatial variability and temporal dynamics. Atmospheric Chemistry and Physics, 20(24):16055–16087.

Geuzaine, C. and Remacle, J.-F., 2009. Gmsh: a three-dimensional finite element mesh generator with built-in pre-and post-processing facilities. International Journal for Numerical Methods in Engineering, 79(11):1309–1331.

Gharasoo, M., Centler, F., Regnier, P., Harms, H., and Thullner, M., 2012. A reactive transport modeling approach to simulate biogeochemical processes in pore structures with pore-scale heterogeneities. Environmental Modelling and Software, 30:102–114.

Gimmi, T. and Kosakowski, G., 2011. How mobile are sorbed cations in clays and clay rocks? Environmental Science and Technology, 45(4):1443–1449.

Glarborg, P., Miller, J. A., Ruscic, B., and Klippenstein, S. J., 2018. Modeling nitrogen chemistry in combustion. Progress in Energy and Combustion Science, 67:31–68.

Glaus, M., Aertsens, M., Appelo, C., Kupcik, T., Maes, N., Van Laer, L., and Van Loon, L., 2015. Cation diffusion in the electrical double layer enhances the mass transfer rates for Sr_2^+, CO_2^+ and Zn_2^+ in compacted illite. Geochimica et Cosmochimica Acta, 165:376–388.

Glaus, M. A., Frick, S., Rossé, R., and Loon, L. R. V., 2010. Comparative study of tracer diffusion of HTO, $^{22}Na^+$ and $^{36}Cl^-$ in compacted kaolinite, illite and montmorillonite. Geochimica et Cosmochimica Acta, 74(7):1999–2010.

Godinho, J. R. and Withers, P. J., 2018. Time-lapse 3D imaging of calcite precipitation in a microporous column. Geochimica et Cosmochimica Acta, 222:156–170.

Guérillot, D. and Bruyelle, J., 2020. Geochemical equilibrium determination using an artificial neural network in compositional reservoir flow simulation. Computational Geosciences, 24(2): 697–707.

Guerillot, D. R. and Bruyelle, J., 2016. History matching methodology using an optimal neural network proxy and a global optimization method. In 3rd EAGE Integrated Reservoir Modelling Conference, volume 2016, pages 1–5. European Association of Geoscientists and Engineers, EAGE.

Helgeson, H. C. and Kirkham, D. H., 1974a. Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures; II, Debye-Huckel parameters for activity coefficients and relative partial molal properties. American Journal of Science, 274 (10):1199–1261.

Helgeson, H. C. and Kirkham, D. H., 1974b. Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures; I, Summary of the thermodynamic/electrostatic properties of the solvent. American Journal of Science, 274(10): 1089–1198.

Helgeson, H. C. and Kirkham, D. H., 1976. Theoretical Prediction of the Thermodynamic Properties of Aqueous Electrolytes at High Pressures and Temperatures - 3. Equation of State for Aqueous Species at Infinite Dilution. American Journal of Science, 276(2):97–240.

Hensen, C., Landenberger, H., Zabel, M., Gundersen, J. K., Glud, R. N., and Schulz, H. D., 1997. Simulation of early diagenetic processes in continental slope sediments off southwest Africa: The computer model CoTAM tested. Marine Geology, 144(1-3):191–210.

Hinton, G. E., Osindero, S., and Teh, Y., 2006. A fast learning algorithm for deep belief nets. Neural Computation, 18:1527–1554.

Hoch, A. R., Baston, G. M. N., Glasser, F. P., Hunter, F. M. I., and Smith, V., 2012. Modelling evolution in the near field of a cementitious repository. Mineralogical Magazine, 76(8): 3055–3069.

Hooyman, G. J., 1961. On Thermodynamic Coupling of Chemical Reactions. Proceedings of the National Academy of Sciences, 47(8):1169–1173.

Hundsdorfer, W. and Verwer, J. G., 1995. A note on splitting errors for advection-reaction equations. Applied Numerical Mathematics, 18(1):191–199.

Hunt, A. G. and Sahimi, M. Flow, Transport, and Reaction in Porous Media: Percolation Scaling, Critical-Path Analysis, and Effective Medium Approximation, 2017.

Idiart, A., Laviña, M., Kosakowski, G., Cochepin, B., Meeussen, J. C., Samper, J., Mon, A., Montoya, V., Munier, I., Poonoosamy, J., Montenegro, L., Deissmann, G., Rohmen, S., Damiani, L. H., Coene, E., and Nieves, A., 2020. Reactive transport modelling of a low-pH concrete / clay interface. Applied Geochemistry, 115:104562.

Ingri, N., Kakolowicz, W., Sillén, L. G., and Warnqvist, B., 1967. High-speed computers as a supplement to graphical methods-V1 1 Part IV-Arkiv Kemi, 1964, 23, 97; for commsnt on spelling "program" see Talanta, 1967, 14, 833. Haltafall, a general program for calculating the composition of equilibrium mixtures. Talanta, 14(11):1261–1286.

Jatnieks, J., De Lucia, M., Dransch, D., and Sips, M., 2016. Data-driven Surrogate Model Approach for Improving the Performance of Reactive Transport Simulations. In Energy Procedia, volume 97, pages 447–453. Elsevier Ltd.

Jenni, A., Mäder, U., Lerouge, C., Gaboreau, S., and Schwyn, B., 2014. In situ interaction between different concretes and Opalinus Clay. Physics and Chemistry of the Earth, 70-71: 71–83.

Jenni, A., Gimmi, T., Alt-Epping, P., Mäder, U., and Cloet, V., 2017. Interaction of ordinary Portland cement and Opalinus Clay: Dual porosity modelling compared to experimental data. Physics and Chemistry of the Earth, 99:22–37.

Johnson, J. W., Oelkers, E. H., and Helgeson, H. C., 1992. SUPCRT92: A software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000°C. Computers and Geosciences, 18(7):899–947.

Kanit, T., Forest, S., Galliet, I., Mounoury, V., and Jeulin, D., 2003. Determination of the size of the representative volume element for random composites: statistical and numerical approach. International Journal of Solids and Structures, 40(13):3647–3679.

Keyes, D. E., McInnes, L. C., Woodward, C., Gropp, W., Myra, E., Pernice, M., Bell, J., Brown, J., Clo, A., Connors, J., Constantinescu, E., Estep, D., Evans, K., Farhat, C., Hakim, A., Hammond, G., Hansen, G., Hill, J., Isaac, T., Jiao, X., Jordan, K., Kaushik, D., Kaxiras, E., Koniges, A., Lee, K., Lott, A., Lu, Q., Magerlein, J., Maxwell, R., McCourt, M., Mehl, M., Pawlowski, R., Randles, A. P., Reynolds, D., Rivière, B., Rüde, U., Scheibe, T., Shadid, J., Sheehan, B., Shephard, M., Siegel, A., Smith, B., Tang, X., Wilson, C., and Wohlmuth, B., 2013. Multiphysics simulations: Challenges and opportunities. International Journal of High Performance Computing Applications, 27(1):4–83.

Kharaka, Y. K., Gunter, W. D., Aggarwal, P. K., Perkins, E. H., and Debraal, J. D., 1988. SOLMINEQ. 88: A computer program for geochemical modeling of water-rock interactions. U.S. Geological Survey.

Kickmaier, W., 2002. The role of rock laboratories. Technical report, Nagra, Bulletin No. 34, Wettingen, Switzerland.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., and Willing, C., 2016. Jupyter Notebooks—a publishing format for reproducible computational workflows. In Positioning and Power in Academic Publishing: Players, Agents and Agendas - Proceedings of the 20th International Conference on Electronic Publishing, ELPUB, pages 87–90. IOS Press BV.

Koroleva, M., Alt-Epping, P., and Mazurek, M., 2011. Large-scale tracer profiles in a deep claystone formation (Opalinus Clay at Mont Russelin, Switzerland): Implications for solute transport processes and transport properties of the rock. Chemical Geology, 280(3-4):284–296.

Kosakowski, G. and Berner, U., 2013. The evolution of clay rock/cement interfaces in a cementitious repository for low- and intermediate level radioactive waste. Physics and Chemistry of the Earth, 64:65–86.

Kosakowski, G. and Watanabe, N., 2014. OpenGeoSys-Gem: A numerical tool for calculating geochemical and porosity changes in saturated and partially saturated media. Physics and Chemistry of the Earth, 70-71:138–149.

Kosakowski, G., Berner, U., Wieland, E., Glaus, M., and Degueldre, C., 2014. Nagra Technical Report 14-11: Geochemical Evolution of the L/ILW Near-field. Technical report, Paul Scherrer Institut.

Kulik, D. A. and Kersten, M., 2001. Aqueous Solubility Diagrams for Cementitious Waste Stabilization Systems: II, End-Member Stoichiometries of Ideal Calcium Silicate Hydrate Solid Solutions. Journal of the American Ceramic Society, 84(12):3017–3026.

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

Lai, P., Moulton, K., and Krevor, S., 2015. Pore-scale heterogeneity in the mineral distribution and reactive surface area of porous rocks. Chemical Geology, 411:260–273.

Laloy, E. and Jacques, D., 2019. Emulation of CPU-demanding reactive transport models: a comparison of Gaussian processes, polynomial chaos expansion, and deep neural networks. Computational Geosciences, 23(5):1193–1215.

Lapeyre, J., Han, T., Wiles, B., Ma, H., Huang, J., Sant, G., and Kumar, A., 2021. Machine learning enables prompt prediction of hydration kinetics of multicomponent cementitious systems. Scientific Reports, 11(1):3922.

Leal, A., Kulik, D. A., Kosakowski, G., and Saar, M. O., 2016. Computational methods for reactive transport modeling: An extended law of mass-action, xLMA, method for multiphase equilibrium calculations. Advances in Water Resources, 96:405–422.

Leal, A., Kulik, D. A., Smith, W. R., and Saar, M. O., 2017. An overview of computational methods for chemical equilibrium and kinetic calculations for geochemical and reactive transport modeling. Pure and Applied Chemistry, 89(5):597–643.

Leal, A., Kyas, S., Kulik, D. A., and Saar, M. O., 2020. Accelerating Reactive Transport Modeling: On-Demand Machine Learning Algorithm for Chemical Equilibrium Calculations. Transport in Porous Media, 133:161–204.

Li, L., Maher, K., Navarre-Sitchler, A., Druhan, J., Meile, C., Lawrence, C., Moore, J., Perdrial, J., Sullivan, P., Thompson, A., Jin, L., Bolton, E. W., Brantley, S. L., Dietrich, W. E., Mayer, K. U., Steefel, C. I., Valocchi, A., Zachara, J., Kocar, B., Mcintosh, J., Tutolo, B. M., Kumar, M., Sonnenthal, E., Bao, C., and Beisman, J., 2017. Expanding the role of reactive transport models in critical zone processes. Earth-Science Reviews, 165:280–301.

Lichtner, P. C., 1985. Continuum model for simultaneous chemical reactions and mass transport in hydrothermal systems. Geochimica et Cosmochimica Acta, 49(3):779–800.

Liu, C., Shang, J., and Zachara, J. M., 2011. Multispecies diffusion models: A study of uranyl species diffusion. Water Resources Research, 47(12).

Liu, J., Pereira, G. G., Liu, Q., and Regenauer-Lieb, K., 2016. Computational challenges in the analyses of petrophysics using microtomography and upscaling: A review. Computers and Geosciences, 89:107–117.

Lothenbach, B., Kulik, D. A., Matschei, T., Balonis, M., Baquerizo, L., Dilnesa, B., Miron, G. D., and Myers, R. J., 2019. Cemdata18: A chemical thermodynamic database for hydrated Portland cements and alkali-activated materials. Cement and Concrete Research, 115:472–506.

Lovley, D. R. and Chapelle, F. H., 1995. Deep subsurface microbial processes. Reviews of Geophysics, 33(3):365.

Mäder, U., Jenni, A., Lerouge, C., Gaboreau, S., Miyoshi, S., Kimura, Y., Cloet, V., Fukaya, M., Claret, F., Otake, T., Shibata, M., and Lothenbach, B., 2017. 5-year chemico-physical evolution of concrete–claystone interfaces, Mont Terri rock laboratory (Switzerland). Swiss Journal of Geosciences, 110(1):307–327.

Maes, N., Wang, L., Delecaut, G., Van Geet, M., Put, M., Weetjens, E., Marivoet, J., Lee, J., Warwick, P., Hall, A., Walker, G., Maes, A., Bruggeman, C., Bennett, D., Hicks, T., Higgo, J., and Galson, D., 2004. Migration Case Study: Transport of radionuclides in a reducing clay

sediment (TRANCOM-II). Final Scientific and Technical Report of the EC TRANCOM-II project. Belgian Nuclear Research Centre.

Mancini, A., Wieland, E., Geng, G., Dähn, R., Skibsted, J., Wehrli, B., and Lothenbach, B., 2020. Fe(III) uptake by calcium silicate hydrates. Applied Geochemistry, 113.

Marty, N. C., Tournassat, C., Burnol, A., Giffaut, E., and Gaucher, E. C., 2009. Influence of reaction kinetics and mesh refinement on the numerical modelling of concrete/clay interactions. Journal of Hydrology, 364(1-2):58–72.

Marty, N. C., Bildstein, O., Blanc, P., Claret, F., Cochepin, B., Gaucher, E. C., Jacques, D., Lartigue, J. E., Liu, S., Mayer, K. U., Meeussen, J. C., Munier, I., Pointeau, I., Su, D., and Steefel, C. I., 2015. Benchmarks for multicomponent reactive transport across a cement/clay interface. Computational Geosciences, 19(3):635–653.

Mazurek, M. and de Haller, A., 2017. Pore-water evolution and solute-transport mechanisms in Opalinus Clay at Mont Terri and Mont Russelin (Canton Jura, Switzerland). Swiss Journal of Geosciences, 110(1):129–149.

Mazurek, M., Gautschi, A., Marschall, P., Vigneron, G., Lebon, P., and Delay, J., 2008. Transferability of geoscientific information from various sources (study sites, underground rock laboratories, natural analogues) to support safety cases for radioactive waste repositories in argillaceous formations. Physics and Chemistry of the Earth, Parts A/B/C, 33:S95–S105. Clays in Natural & Engineered Barriers for Radioactive Waste Confinement.

Meeussen, J. C. L., 2003. ORCHESTRA: An Object-Oriented Framework for Implementing Chemical Equilibrium Models. Environ. Sci. Technol., 37(6):1175–1182.

Meissner, H. P., Kusik, C. L., and Dalzell, W. H., 1969. Equilibrium composition with multiple reactions. Industrial and Engineering Chemistry Fundamentals, 8(4):659–665.

Millington, R. J., 1959. Gas diffusion in porous media. Science, 130(3367):100–102.

NAGRA, 2002. TECHNICAL REPORT 02-05 National Cooperative for the Disposal of Radioactive Waste. Technical report, NAGRA.

NAGRA, 2008. Effects of post-disposal gas generation in a repository for low- and intermediate-level waste sited in the Opalinus Clay of Northern Switzerland (Technical Report 08-07). Technical report, NAGRA, Wettingen, Switzerland.

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

Narsilio, G. A., Buzzi, O., Fityus, S., Yun, T. S., and Smith, D. W., 2009. Upscaling of Navier–Stokes equations in porous media: Theoretical, numerical and experimental approach. Computers and Geotechnics, 36(7):1200–1206.

NEA. The environmental and ethical basis of geological disposal of long-lived radioactive wastes, 1995.

Necib, S., Diomidis, N., Keech, P., and Nakayama, M., 2017. Corrosion of carbon steel in clay environments relevant to radioactive waste geological disposals, Mont Terri rock laboratory (Switzerland). Swiss Journal of Geosciences, 110(1):329–342.

Nordstrom, D. K., Plummer, L. N., Wigley, T. M. L., Wolery, T. J., Ball, J. W., Jenne, E. A., Bassett, R. L., Crerar, D. A., Florence, T. M., Fritz, B., Hoffman, M., Holdren, G. R., Lafon, G. M., Mattigod, S. V., Mcduff, R. E., Morel, F., Reddy, M. M., Sposito, G., and Thrailkill, J. A Comparison of Computerized Chemical Models for Equilibrium Calculations in Aqueous Systems, volume 93, chapter 38, pages 857–892. ACS Symposium Series, 1979.

Nussbaum, C., Bossart, P., Amann, F., and Aubourg, C., 2011. Analysis of tectonic structures and excavation induced fractures in the Opalinus Clay, Mont Terri underground rock laboratory (Switzerland). Swiss Journal of Geosciences, 104(2):187–210.

Pape, H., Clauser, C., and Iffland, J., 1999. Permeability prediction based on fractal pore-space geometry. Geophysics, 64(5):1447–1460.

Parkhurst, D. and Appelo, C., 2013. Description of Input and Examples for PHREEQC Version 3—A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations. U.S. Geological Survey Techniques and Methods, page 497 p.

Pearson, F. J., Arcos, D., Bath, A., Boisson, J.-Y., Fernandez, A. M., Gäbler, H.-E., Gautschi, E., Griffault, L., Hernán, P., and Waber, H. N., 2003. Mont Terri project: geochemistry of water in the Opalinus Clay Formation at the Mont Terry rock laboratory. Bundesamt für Wasser und Geologie, BWG, Bern, 5 edition.

Poonoosamy, J., Kosakowski, G., Van Loon, L. R., and Mäder, U., 2015. Dissolution-precipitation processes in tank experiments for testing numerical models for reactive transport calculations: Experiments and modelling. Journal of Contaminant Hydrology, 177-178:1–17.

Poonoosamy, J., Klinkenberg, M., Deissmann, G., Brandt, F., Bosbach, D., Mäder, U., and Kosakowski, G., 2020. Effects of solution supersaturation on barite precipitation in porous media and consequences on permeability: Experiments and modelling. Geochimica et Cosmochimica Acta, 270:43–60.

Prasianakis, N. I., Curti, E., Kosakowski, G., Poonoosamy, J., and Churakov, S. V., 2017. Deciphering pore-level precipitation mechanisms. Scientific Reports, 7(1):1–9.

Prasianakis, N. I., Haller, R., Mahrous, M., Poonoosamy, J., Pfingsten, W., and Churakov, S. V., 2020. Neural network based process coupling and parameter upscaling in reactive transport simulations. Geochimica et Cosmochimica Acta, 291:126–143.

Reed, M. H., 1982. Calculation of multicomponent chemical equilibria and reaction processes in systems involving minerals, gases and an aqueous phase. Geochimica et Cosmochimica Acta, 46(4):513–528.

Reid, R. C., Sherwood, T. K., Prausnitz, J. M., and Poling, B. E., 1977. The Properties of Gases and Liquids. McGraw-Hill Education, 4 edition.

Saheb, M., Descostes, M., Neff, D., Matthiesen, H., Michelin, A., and Dillmann, P., 2010. Iron corrosion in an anoxic soil: Comparison between thermodynamic modelling and ferrous archaeological artefacts characterised along with the local in situ geochemical conditions. Applied Geochemistry, 25(12):1937–1948.

Samson, E. and Marchand, J., 1999. Numerical Solution of the Extended Nernst–Planck Model. Journal of Colloid and Interface Science, 215(1):1–8.

Schmidhuber, J., 2015. Deep Learning in neural networks: An overview. Neural Networks, 61: 85–117.

Seigneur, N., Mayer, K. U., and Steefel, C. I., 2019. Reactive Transport in Evolving Porous Media. Reviews in Mineralogy and Geochemistry, 85(1):197–238.

Sheikholeslami, M., Haq, R. u., Shafee, A., Li, Z., Elaraki, Y. G., and Tlili, I., 2019. Heat transfer simulation of heat storage unit with nanoparticles and fins through a heat exchanger. International Journal of Heat and Mass Transfer, 135:470–478.

Skipper, N. T., Lock, P. A., Titiloye, J. O., Swenson, J., Mirza, Z. A., Howells, W. S., and Fernandez-Alonso, F., 2006. The structure and dynamics of 2-dimensional fluids in swelling clays. Chemical Geology, 230(3-4):182–196.

Sohrabi, R., Omlin, S., and Miller, S. A., 2019. GEYSER: 3D thermo-hydrodynamic reactive transport numerical simulator including porosity and permeability evolution using GPU clusters. Computational Geosciences, 23(6):1317–1330.

Steefel, C. I. and Lasaga, A. C., 1994. A coupled model for transport of multiple chemical species and kinetic precipitation/dissolution reactions with application to reactive flow in single phase hydrothermal systems. American Journal of Science, 294(5):529–592.

Steefel, C. I. and Maher, K., 2009. Fluid-rock interaction: A reactive transport approach. Reviews in mineralogy and geochemistry, 70(1):485–532.

Steefel, C. I., DePaolo, D. J., and Lichtner, P. C., 2005. Reactive transport modeling: An essential tool and a new research approach for the Earth sciences. Earth and Planetary Science Letters, 240(3-4):539–558.

Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., Lagneau, V., Lichtner, P. C., Mayer, K. U., Meeussen, J. C. L., Molins, S., Moulton, D., Shao, H., Šimůnek, J., Spycher, N., Yabusaki, S. B., and Yeh, G. T., 2015. Reactive transport codes for subsurface environmental simulation. Computational Geosciences, 19(3):445–478.

Stroes-Gascoyne, S., Schippers, A., Schwyn, B., Poulain, S., Sergeant, C., Simonoff, M., Le Marrec, C., Altmann, S., Criepi, T. N., Mauclaire, J. L., Mckenzie, J., Daumas, S., Vinsot, A., Beaucaire, C., and Matray, J.-M., 2007a. Microbial Community Analysis of Opalinus Clay Drill Core Samples from the Mont Terri Underground Research Laboratory, Switzerland. Geomicrobiology Journal, 24:1–17.

Stroes-Gascoyne, S., Schippers, A., Schwyn, B., Poulain, S., Sergeant, C., Simonoff, M., Marrec, C. L., Altmann, S., Nagaoka, T., Mauclaire, L., McKenzie, J., Daumas, S., Vinsot, A., Beaucaire, C., and Matray, J. M., 2007b. Microbial Community Analysis of Opalinus Clay Drill Core Samples from the Mont Terri Underground Research Laboratory, Switzerland. Geomicrobiology Journal, 24(1):1–17.

Su, D., Ulrich Mayer, K., and MacQuarrie, K. T., 2017. Parallelization of MIN3P-THCm: A high performance computational framework for subsurface flow and reactive transport simulation. Environmental Modelling & Software, 95:271–289.

Su, D., Mayer, K. U., and MacQuarrie, K. T., 2020. MIN3P-HPC: A High-Performance Unstructured Grid Code for Subsurface Flow and Reactive Transport Simulation. Mathematical Geosciences, pages 1–34.

Thoenen, T., Hummel, W., Berner, U., and Curti, E., 2014. The PSI/Nagra Chemical Thermodynamic Database 12/07. Technical report, Nuclear Energy and Safety Research Department Laboratory for Waste Management (LES).

Thury, M. and Bossart, P., 1999. The Mont Terri rock laboratory, a new international research project in a Mesozoic shale formation, in Switzerland. Engineering Geology, 52:347–359.

Tournassat, C. and Appelo, C., 2011. Modelling approaches for anion-exclusion in compacted Na-bentonite. Geochimica et Cosmochimica Acta, 75(13):3698–3710.

Trangenstein, J. A., 1987. Customized minimization techniques for phase equilibrium computations in reservoir simulation. Chemical Engineering Science, 42(12):2847–2863.

Trotignon, L., Devallois, V., Peycelon, H., Tiffreau, C., and Bourbon, X., 2007. Predicting the long term durability of concrete engineered barriers in a geological repository for radioactive waste. Physics and Chemistry of the Earth, Parts A/B/C, 32(1):259–274.

Tsang, C. F., Barnichon, J. D., Birkholzer, J., Li, X. L., Liu, H. H., and Sillen, X., 2012. Coupled thermo-hydro-mechanical processes in the near field of a high-level radioactive waste repository in clay formations. International Journal of Rock Mechanics and Mining Sciences, 49:31–44.

Tu, J. V., 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. Journal of Clinical Epidemiology, 49(11):1225–1231.

Van Loon, L. R., Soler, J. M., Müller, W., and Bradbury, M. H., 2004. Anisotropic Diffusion in Layered Argillaceous Rocks: A Case Study with Opalinus Clay. Environ. Sci. Technol., 38 (21):5721–5728.

Van Loon, L. R., Glaus, M. A., and Müller, W., 2007. Anion exclusion effects in compacted bentonites: Towards a better understanding of anion diffusion. Applied Geochemistry, 22(11): 2536–2552.

Van Loon, L. R., Leupin, O. X., and Cloet, V., 2018. The diffusion of SO_4^{2-} in Opalinus Clay: Measurements of effective diffusion coefficients and evaluation of their importance in view of microbial mediated reactions in the near field of radioactive waste repositories. Applied Geochemistry, 95:19–24.

Vinsot, A., Appelo, C. A. J., Lundy, M., Wechner, S., Cailteau-Fischbach, C., de Donato, P., Pironon, J., Lettry, Y., Lerouge, C., and De Cannière, P., 2017. Natural gas extraction and artificial gas injection experiments in Opalinus Clay, Mont Terri rock laboratory (Switzerland). Swiss Journal of Geosciences, 110(1):375–390.

W. E. Dibble, W. A. T., 1981. Non-equilibrium water/rock interactions. I. Model for interface-controlled reactions. Geochim. Cosmochim. Acta, 45:79–92.

Walter, A. L., Frind, E. O., Blowes, D. W., Ptacek, C. J., and Molson, J. W., 1994. Modeling of multicomponent reactive transport in groundwater: 1. Model development and evaluation. Water Resources Research, 30(11):3137–3148.

Wei, H., Bao, H., and Ruan, X., 2020. Machine learning prediction of thermal transport in porous media with physics-based descriptors. International Journal of Heat and Mass Transfer, 160.

Wigger, C. and Van Loon, L. R., 2018. Effect of the pore water composition on the diffusive anion transport in argillaceous, low permeability sedimentary rocks. Journal of Contaminant Hydrology, 213:40–48.

Wolery, T. J., 1992. EQ3/6, A Software Package for Geochemical Modeling of Aqueous Systems: Package Overview and Installation Guide (Version 7.0). Technical report, Lawrence Livermore National Laboratory.

Xu, T., Sonnenthal, E., Spycher, N., and Pruess, K., 2006. TOUGHREACT - A simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and CO_2 geological sequestration. Computers and Geosciences, 32(2):145–165.

Xu, T., Senger, R., and Finsterle, S., 2008. Corrosion-induced gas generation in a nuclear waste repository: Reactive geochemistry and multiphase flow effects. Applied Geochemistry, 23(12): 3423–3433.

Xu, T., Zheng, L., and Tian, H., 2011. Reactive transport modeling for CO_2 geological sequestration. Journal of Petroleum Science and Engineering, 78(3-4):765–777.

Yeh, G.-T. and Tripathi, V. S., 1991. A Model for Simulating Transport of Reactive Multispecies Components: Model Development and Demonstration. Water Resources Research, 27(12): 3075–3094.

Yekta, A., Salinas, P., Hajirezaie, S., Amooie, M. A., Pain, C. C., Jackson, M. D., Jacquemyn, C., and Soltanian, M. R., 2021. Reactive transport modeling in heterogeneous porous media with dynamic mesh optimization. Computational Geosciences, 25(1):357–372.

Zeggeren, F. v., 1970. The computation of chemical equilibria. Cambridge University Press.

Zeleznik, F. J. and Gordon, S., 1968. Calculation of Complex Chemical Equilibria. Industrial & Engineering Chemistry, 60(6):27–57.

Declaration of consent

Declaration of consent

on the basis of Article 18 of the PromR Phil.-nat. 19

Name/First Name: Hax Damiani / Leonardo

Registration Number: 16-129-587

Study program: Doctoral studies in Earth Sciences (Geology)

Bachelor ☐ Master ☐ Dissertation ☑

Title of the thesis: A novel reactive transport framework for fluid-rock interaction analysis: computational approach, applications and benchmarks

Supervisor: Prof. Dr. Sergey V. Churakov
PD Dr. Georg Kosakowski

I declare herewith that this thesis is my own work and that I have not used any sources other than those stated. I have indicated the adoption of quotations as well as thoughts taken from other authors as such in the thesis. I am aware that the Senate pursuant to Article 36 paragraph 1 litera r of the University Act of September 5th, 1996 and Article 69 of the University Statute of June 7th, 2011 is authorized to revoke the doctoral degree awarded on the basis of this thesis.
For the purposes of evaluation and verification of compliance with the declaration of originality and the regulations governing plagiarism, I hereby grant the University of Bern the right to process my personal data and to perform the acts of use this requires, in particular, to reproduce the written thesis and to store it permanently in a database, and to use said database, or to make said database available, to enable comparison with theses submitted by others.

Place/Date

Signature