Large-scale reactive transport simulations of uranium migration in Opalinus Clay accelerated by means of surrogate models

Marco De Lucia[*,1], Max Lübke[2], Theresa Hennig[1], Bettina Schnor[2]

[1] GFZ German Research Centre for Geosciences, Fluid Systems Modelling, Potsdam, Germany

[2] University of Potsdam, Institute of Computer Science, Potsdam, Germany

* delucia@gfz-potsdam.de

Long-term safety assessment of potential radioactive waste disposal sites relies on coupled reactive transport simulations on the host rock scale and beyond. They need to be able to capture both geological heterogeneities and complex geochemical processes of different radionuclide species. This represents a challenge since such coupled simulations require considerable computing resources, especially for geochemistry, which is usually the computational bottleneck. One possible solution which is increasingly explored by the scientific community is to leverage methods from AI/ML (Artificial Intelligence, Machine Learning) to approximate the expensive numerical solutions of geochemistry.

The migration of uranium, the main component of spent fuel, in the Opalinus Clay is used here as an example. The Opalinus Clay is the chosen host rock in Switzerland and also considered in the German site selection process. Radionuclide transport in argillaceous formations like the Opalinus Clay occurs via molecular diffusion and is retarded by sorption processes on the inherent clay minerals. Both processes depend on the pore water geochemistry and mineralogy. In the geochemical system of the formation, uranium forms aqueous ternary complexes with calcium and carbonate, and thus the speciation depends on the pore water composition [1]. However, the geochemical conditions can vary across a formation and within geological time scales what, in turn, affects sorption processes and ultimately the migration lengths [2]. At the Mont Terri system, for instance, the pore water geochemistry profiles are characterised by gradients towards the surrounding aquifers that represent the current hydroegological boundaries. Previous geochemical simulations showed that uranium migration lengths resulting from reactive transport simulations can vary by several tens of meters depending on the underlying geochemical conditions, mineralogy and thermodynamic data (Fig. 1).

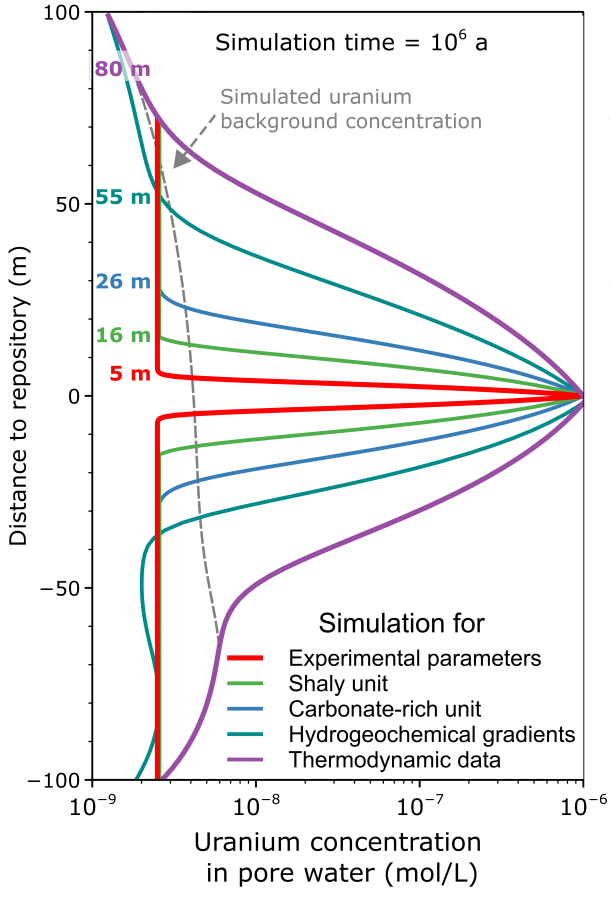

Figure 1: Simulated migration lengths of uranium for the hydrogeological system at Mont Terri (Switzerland) as a function of clay mineralogy (green and blue line), considering the hydrogeological system (cyan line) and different thermodynamic data sets (purple line).

In this contribution we showcase an hybrid physics-data driven regression approach called DecTree [3], which consists in a recursive partitioning of the parameter space of geochemical models based on knowledge about the ongoing processes. Within each partition, a model reduction is performed by identifying functional dependencies between outputs and a minimum number of inputs of the system. In each subspace, either a simple regression of low dimensionality or further constraints such as mass and charge balance can be used to perform the approximated prediction of all the output variables. The advantage of such hybrid approach is an increased robustness, since physical constraints are embedded into the approach, and the negligible training time, which usually burdens classical pure data-driven approaches such as artificial neural networks or Gaussian processes, which however are still faster to predict especially when run on GPU. We evaluate the performance and the accuracy of our approach based on benchmarks and datasets for uranium migration in Opalinus Clay subject to surface complexation, cation exchange, and mineral precipitation as defined within the ML Benchmark initiative in the DONUT/EURAD project.

We compare the performance gain and the accuracy loss caused by the DecTree-approximated solutions in coupled diffusive reactive transport models using the HPC simulator POET [4] on uniform grids of 10^5-10^6 elements with heterogeneous properties. Furthermore, POET implements storage of exact numerical solutions of geochemistry in highly efficient Distributed Hash Tables (DHT) and provides the ability to interpolate new predictions from appropriate DHT entries, retrieved through automatic clustering of the multivariate data [5]. Furthermore, POET can be coupled with AI libraries through a high-level interface in the R language, and can hence incrementally train surrogate models such as neural networks during coupled simulations and use their predictions.

We draw lessons learned by testing and evaluating these different approaches and discuss their caveats. These approximations can lead to dramatic speedups of coupled simulations, hence enabling models of unprecedented complexity or comprehensive uncertainty analyses without recurring to oversimplifications due to the computational burden of classical numerical simulations, both requirements for proper Safety Case assessments.

References

[1] T. Hennig et al., “Simulation of diffusive uranium transport and sorption processes in the Opalinus Clay”, Appl. Geochemistry, 123, 104777, 2020, doi: 10.1016/j.apgeochem.2020.104777.

[2] T. Hennig and M. Kühn, “Reactive transport simulations of uranium migration in the Opalinus Clay depend on ion speciation governed by underlying thermodynamic data”, Adv. Geosci., vol. 58, pp. 11-18, 2022, doi: 10.5194/adgeo-58-11-2022

[3] M. De Lucia and M. Kühn, “DecTree v1.0 - Chemistry speedup in reactive transport simulations: Purely data-driven and physics-based surrogates”, Geosci. Mod. Devel., n. 14, pp. 4713–473, 2021, doi: 10.5194/gmd-14-4713-2021.

[4] M. De Lucia et al., “POET (v0.1): speedup of many-core parallel reactive transport simulations with fast DHT lookups”, Geosci. Mod. Devel, n. 14, pp. 7391–7409, 2021, doi: 10.5194/gmd-14-7391-2021.

[5] M. Lübke, “Design, implementation and validation of an interpolation-based surrogate model for the POET simulator”. Master thesis in Computer Science, University of Potsdam, 18 July 2023.