Speeding up reactive transport simulations: statistical surrogates and caching of simulation results in lookup tables

Marco De Lucia\textsuperscript{1}, Michael Kühn\textsuperscript{1,2}, Alexander Lindemann\textsuperscript{3}, Max Lübke\textsuperscript{3}, Bettina Schnor\textsuperscript{3}

\textsuperscript{1} GFZ German Research Centre for Geosciences, Fluid Systems Modelling, Germany
\textsuperscript{2} Potsdam University, Institute of Geosciences - Hydrogeology, Germany
\textsuperscript{3} Potsdam University, Institute of Computer Science - Operating Systems and Distributed Systems, Germany

EGU2020-17719 - ERE 6.1
Chemistry is the computational bottleneck in coupled reactive transport simulations: use pre-trained surrogates instead. **Real chemistry called only if their inaccuracy is too large**

- Operator splitting, sequential coupling, only advection
- Multiple multivariate regression (one regressor per output)
- Model hierarchy: mass balance as accuracy control
- R/PHREEQC 1D benchmark implementation: https://gitext.gfz-potsdam.de/delucia/RedModRphree

Using xgboost as regressor, speedup is achieved already for 50-elements grid and it increases for larger grids. **Trade-off**: accuracy vs speedup

---

delucia@gfz-potsdam.de

Surrogates, DHT

EGU2020-17719
Caching chemistry in Distributed Hash Tables (DHT) for further reuse during simulations produces large speedup in common scenarios: making good use of available RAM

- Low overhead: MPI One-Sided Communication
- Simple to implement in existing code
- No accuracy loss (fixed significant digits)
- Retrieves exact input combination

Mpi: master-slave with variable package size
Ketzin grid, 648420 elements, 200 iterations
Homogeneous initial state

Next steps
Combine use of surrogates and caching (DHT) to make large scale reactive transport possible on desktop PC!

delucia@gfz-potsdam.de  Surrogates, DHT  EGU2020-17719