

# **New Approaches to Surface Complexation Modeling and Application of Machine Learning to Radionuclide-Mineral Interface Reactions**

**MAVRIK ZAVARIN<sup>1</sup>, HARUKO WAINWRIGHT<sup>2</sup>,  
JADALLAH ZOUABE<sup>3</sup> AND ELLIOT CHANG<sup>2</sup>**

<sup>1</sup>LLNL

<sup>2</sup>Lawrence Berkeley National Laboratory

<sup>3</sup>UC Berkeley

Presenting Author: [zavarin1@llnl.gov](mailto:zavarin1@llnl.gov)

A fundamental approach to nuclear waste repository research involves the collection of experimental data, development of numerical models, and their application in performance assessment models to inform society of impacts and risks associated with repository scenarios. Machine Learning (ML) is poised to fundamentally change how these predictive tools quantify impacts and risks associated with siting nuclear waste repositories.

In our research, we focus on the experimental data that can be categorized as sorption data. A recent effort at Lawrence Livermore National Laboratory, in coordination with the Helmholtz Zentrum Dresden Rossendorf (HZDR) partners (RES<sup>3</sup>T database, <https://www.hzdr.de/db/RES3T.login>), has been developing a data digitization pipeline for application of ML as well as traditional surface complexation (SC) modeling. To date, the manual digitization of data has yielded a LLNL SCIE database that includes 211 references and a total of 22,732 individual digitized data and associated metadata. With this existing database, we have begun developing mechanistic SC and ML models to simulate radionuclide reactions at the mineral-water interface. Here, we provide a case study on U(VI) sorption to quartz to illustrate the modeling pipelines. A Random Forest ML model was trained and validated to produce "smart K<sub>d</sub>" (L/g) results given various geochemical features, ultimately yielding a validation R<sup>2</sup> score of 94%. A SC modeling approach was developed by coupling the PHREEQC code with the PEST parameter estimation code to optimize surface stability constants. Importantly, both the ML and the SC modeling approaches incorporate uncertainty quantification that can be propagated through to performance assessment calculations. In addition, ML allows us to compare the performance of different SC models reported in RES<sup>3</sup>T against the global datasets in the LLNL SCIE database through automation of PHREEQC-PEST optimization routines. While a number of challenges are still present (e.g. automation of various component of the data assimilation and interpretation in the data stream), these ML approaches will provide a robust and nimble framework for evaluating nuclear waste repository performance that is adaptable to the unique timescales that repository performance, oversight, and monitoring will require.