AI to Guide Sorption Data Acquisition and Assimilation into Uncertainty Quantifications for the Nuclear Waste Disposal Performance Assessment

PI: Haruko Wainwright (Massachusetts Institute of Technology)

Collaborators: Brian Powell (Co-PI), (Clemson University), **Senior Personnel**: Mavrik Zavarin, (Lawrence Livermore National Laboratory), Liange Zheng, Marcus Noack (Lawrence Berkeley National Laboratory)

Program: Spent Fuel, Waste Science & Technology and Integrated Waste Management

ABSTRACT:

The objective of this project is to develop machine learning (ML) and AI toolsets to effectively expand the global sorption database—the datasets collected by multiple institutions around the world—and to assimilate these datasets into the uncertainty quantification (UQ) of radionuclide transport for the performance assessment (PA) of nuclear waste repositories. The AI toolsets will guide the geochemical data acquisitions to further improve the UQ. Our particular focus is on surface complexation and ion exchange sorption processes, which are the major retardation mechanisms for radionuclides. The innovative aspects of this project include (1) physically-informed machine learning under probabilistic programming languages to estimate the posterior distributions of surface complexation model (SCM) parameters, based on the global datasets that can be incorporated into the thermal-hydrological-chemical (THC) and reactive transport models as a part of PA; (2) Bayesian model averaging to integrate the uncertainty associated with multiple SCMs; and (3) autonomous experimental design to optimize the subsequent experimental data acquisition, to reduce the uncertainty in the overall PA.

We will take advantage of recent advances, including (a) the Lawrence Livermore National Laboratory-Surface Complexation/Ion Exchange (L-SCIE) database to unify community adsorption experiments and metadata in a findable, accessible, interoperable, and reusable (FAIR) format, (b) probabilistic programming language, which is a class of programming languages that are specifically designed for probabilistic models and statistical inference, (c) emulators for complex reactive transport modeling, (d) the gpCAM framework, which is a mathematical framework and Python software for Gaussian-Process (GPs) function approximation, uncertainty quantification, Bayesian optimization and autonomous experimentation. We will demonstrate and validate our approach using laboratory sorption and transport experiments. This study will demonstrate the full cycle of UQ from the global sorption database to reactive transport models of uranium through an engineered barrier system (bentonite clay buffer) and clay host rock, as well as the identification of experimental conditions that can reduce the uncertainty and the actual experiments for performance confirmation.

We aim to establish a direct link between the global geochemical database and radionuclide transport models, including the parameter and model uncertainty. Such a comprehensive UQ can transfer the uncertainty from all the existing experimental data to simulation results in a seamless manner. In addition, this direct link enables feedback from model results to data acquisition, optimizing subsequent experimental conditions in such a way as to reduce the uncertainty in the model results. Although the database development and THC models are currently funded under the DOE-NE, there is no framework to bridge these two. We propose to fill this critical gap based on innovative AI/ML applications.