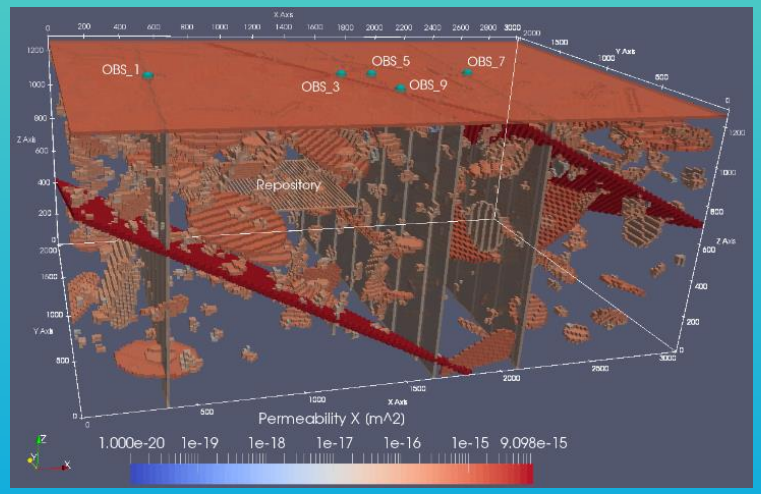


Community Database Development and Application of Surface Complexation and Hybrid ML Approaches to Reactive Transport Modeling and Performance Assessment

U.S. Nuclear Waste Technical Review Board
Fact Finding Meeting
July 19, 2022

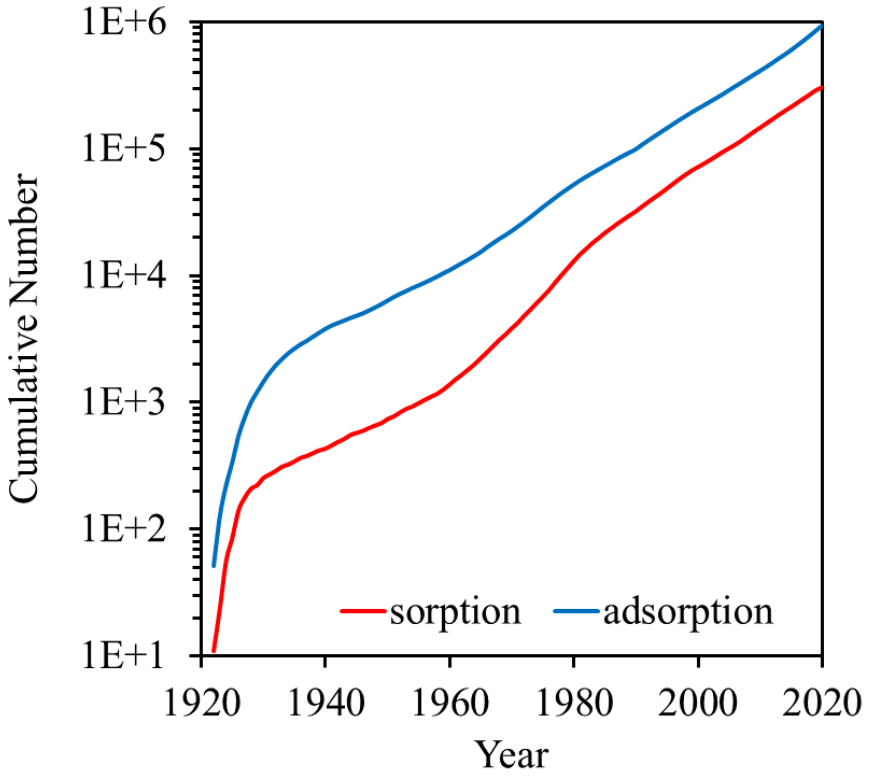
Dr. Mavrik Zavarin, Dr. Elliot Chang
Lawrence Livermore National Laboratory

LLNL-PRES-837417
This work was performed under the auspices of the U.S. Department of Energy By Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

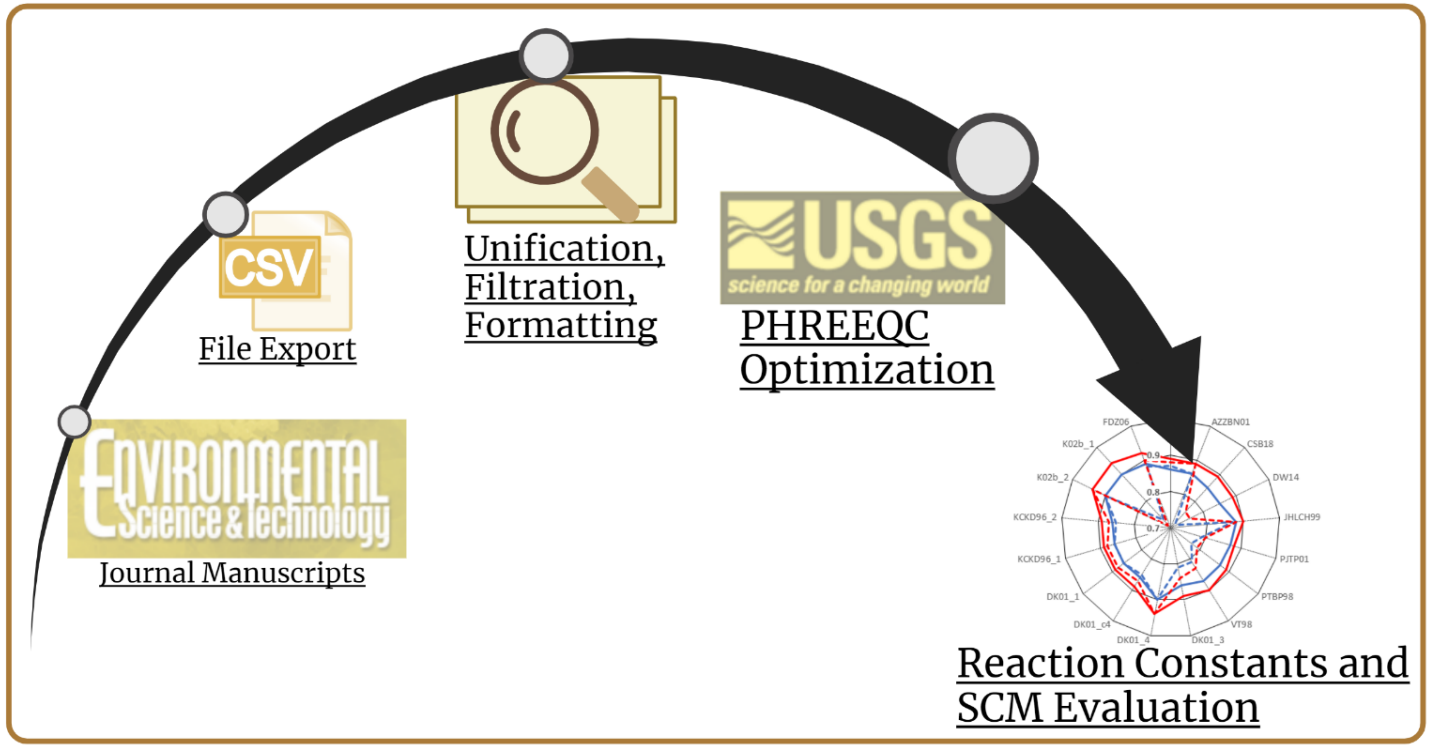
FAIR data is critical to developing defensible models

"Collecting and generating data is outpacing its assimilation, interpretation, and understanding"
-B. Helland, Associate Director of the Office of Science's Advanced Scientific Computing Research (ASCR) program

Data Mining Potential



Data Mining Approach



Overview of sorption modeling workflow

Data mining from literatures

- Data importing
- Original image / Metadata
- Data (Exp. conditions, K_d, etc.)

Database construction

- Data imports to MS ACCESS
- Reference organizing – dataset table (experimental conditions)
- Produce data table (sorption data)

Data formatting

- Data unification (+ error estimation / propagation)
- Data filtration for modeling
- Data export as CSV format for modeling

L-SCIE
(LLNL Surface Complexation/Ion Exchange)

OPENSOURCE CODE AVAILABLE AT: https://ipo.llnl.gov/technologies/software/llnl-surface-complexation-database-converter-scdc

Surface titration modeling

- Coded in Python environment
- Automized PHREEQC/PEST simulation for surface titration of minerals
- Produce optimized surface acidity constants

L-ASTM
(LLNL Automized Surface Titration Model)

input

Traditional SC modeling

- Coded in Python environment
- Automized PHREEQC/PEST simulation (*various combinations* of surface reactions)
- Produce optimized surface complexation constants

L-ASCM
(LLNL Automized Surface Complexation Model)

Intercomparison

ML based SC modeling

- Coded in Python environment
- Coupled with PHREEQC for aqueous speciation
- Predict partition of adsorbate based on machine learning: random forest algorithm

L-SURF
(LLNL Speciation Updated Random Forest)

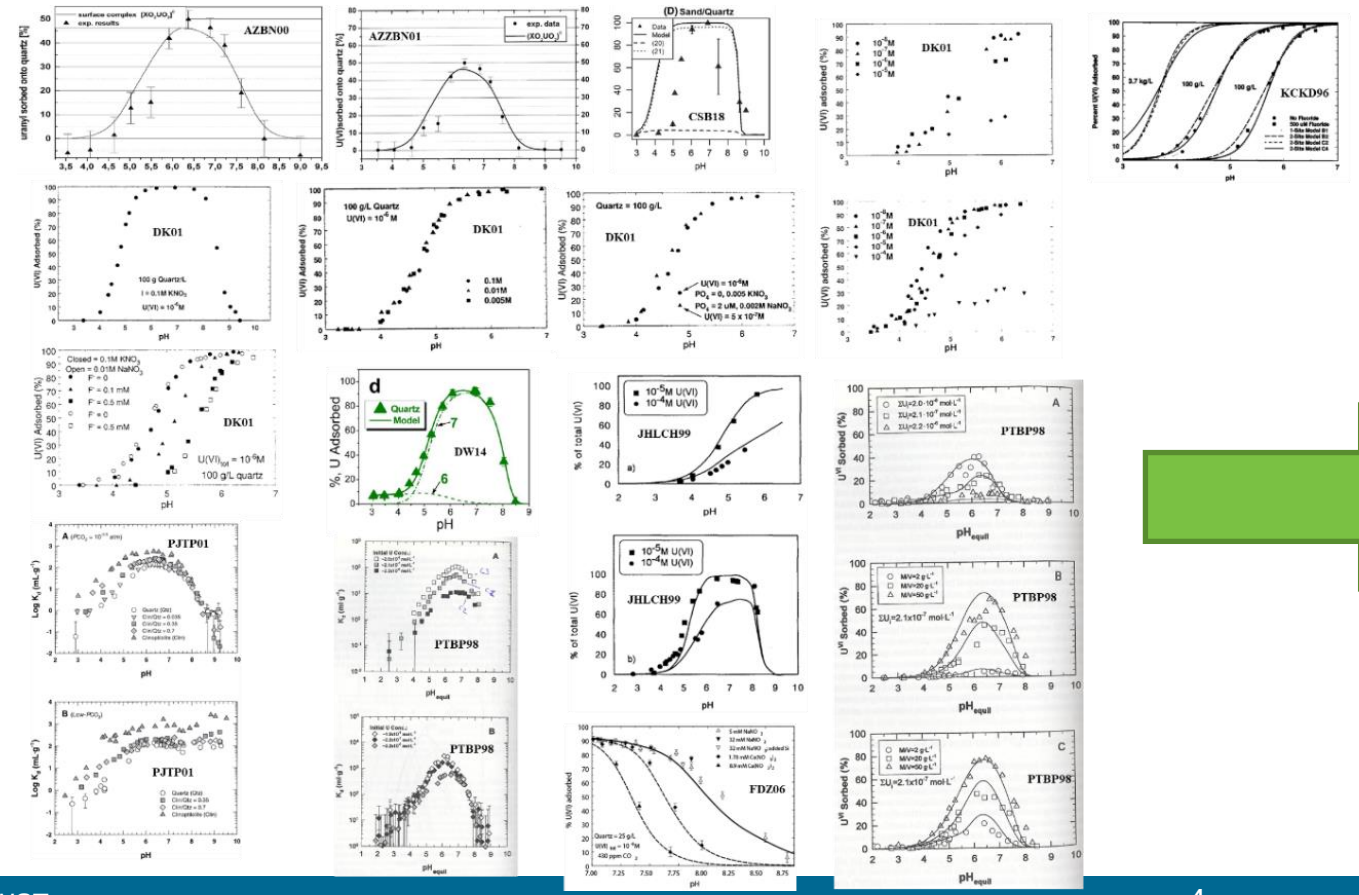
What the L-SCIE database and workflow does

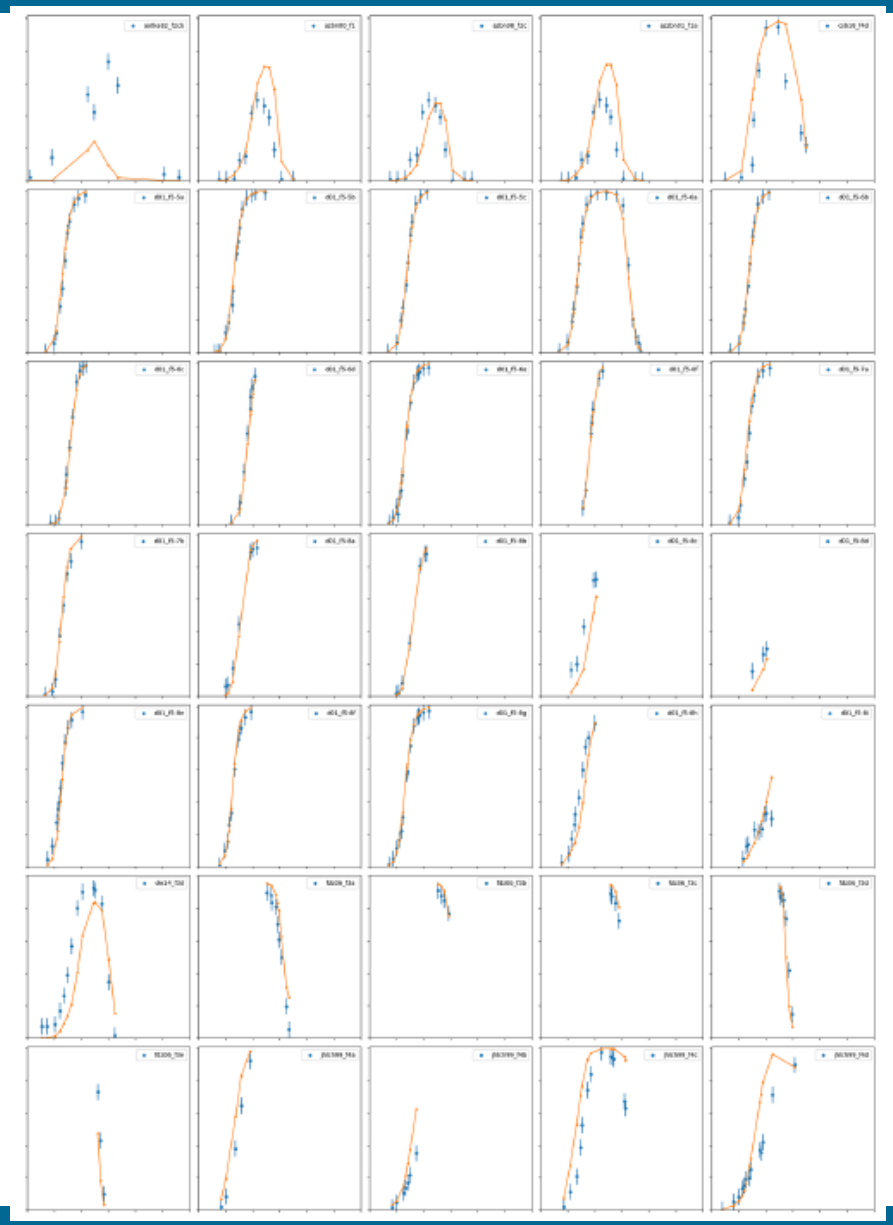
Environmental Science & Technology
pubs.acs.org/est
Article

Community Data Mining Approach for Surface Complexation Database Development

Mavrik Zavarin,*,# Elliot Chang,*,# Haruko Wainwright, Nicholas Parham, Rahul Kaukuntla, Jadallah Zouabe, Amanda Deinhart, Victoria Genetti, Sam Shipman, Frank Bok, and Vinzenz Brendler

Cite This: https://doi.org/10.1021/acs.est.1c07109
Read Online



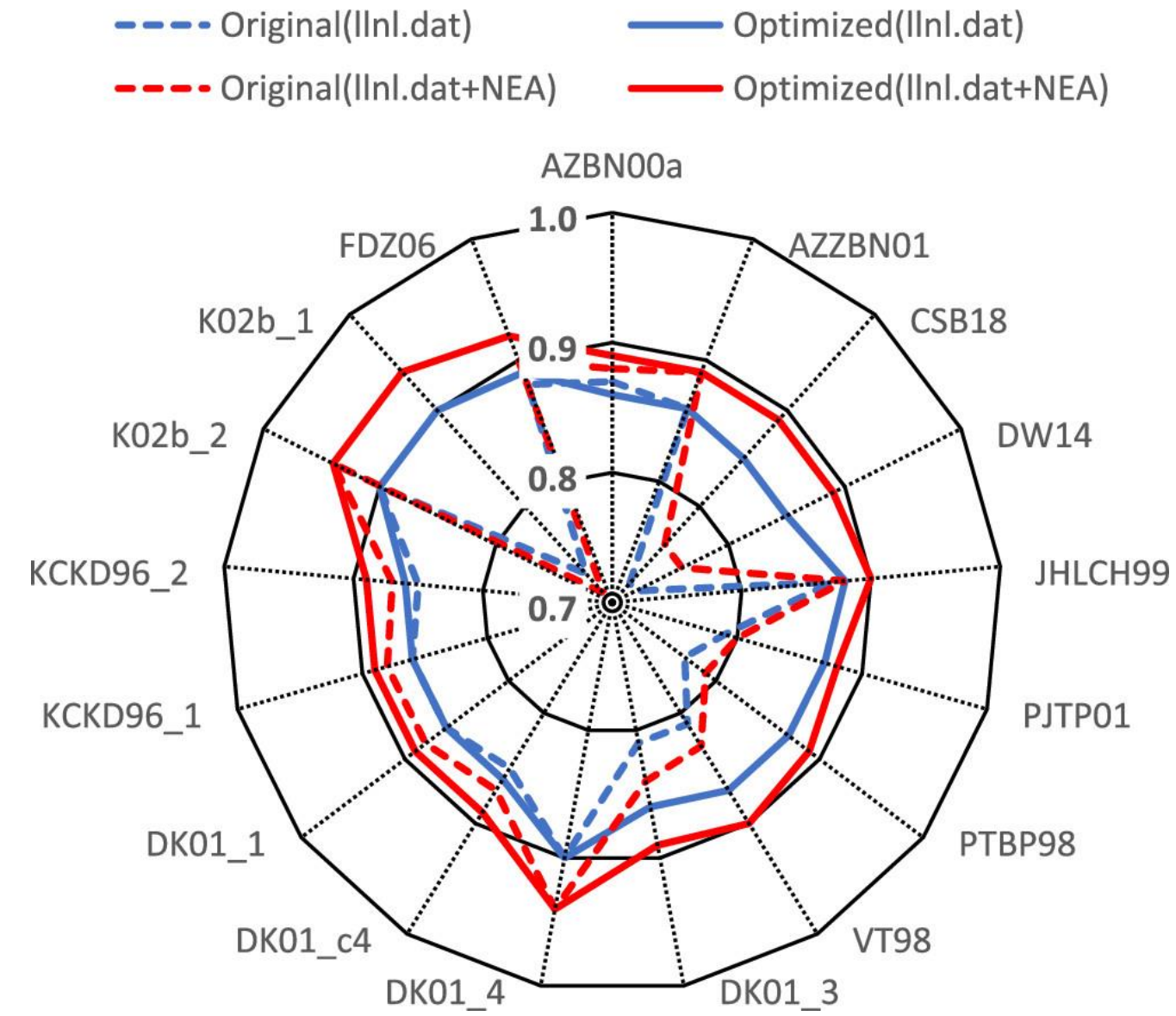


Model testing using published SCMs

Example: U(VI) sorption to quartz and intercomparison of published SCMs (from RES³T*)

- **Optimization is non-unique but some models outperform others**
- **Updates to thermodynamic database have greater impact on SCM performance than the SCMs.**

* RES³T Rossendorf Expert System for Surface and Sorption Thermodynamics



International engagements in L-SCIE development

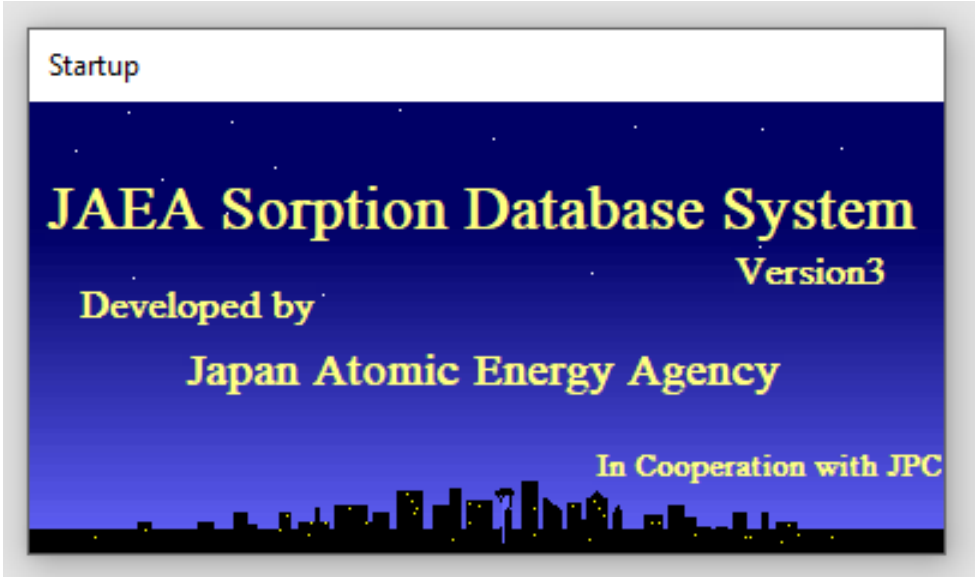
The L-SCIE database status:

- **Total data points: 27,000**
- **References: 243**

Database engagements:

- **JAEA (Yukio Tachi)**
  https://migrationdb.jaea.go.jp/sdb_e2/sdb_pre_e.html
  - **17,000 (of 70,000 available) JAEA data added to L-SCIE database (new total: 44,000!)**

- **HZDR (Vinzenz Brendler)**
  https://www.hzdr.de/db/res3t.login
  - **Database with 7550 surface complexation reaction constants mined from the literature (3398 references)**
  - **NEW: SOrption REference DAtabase (SOREDA)**

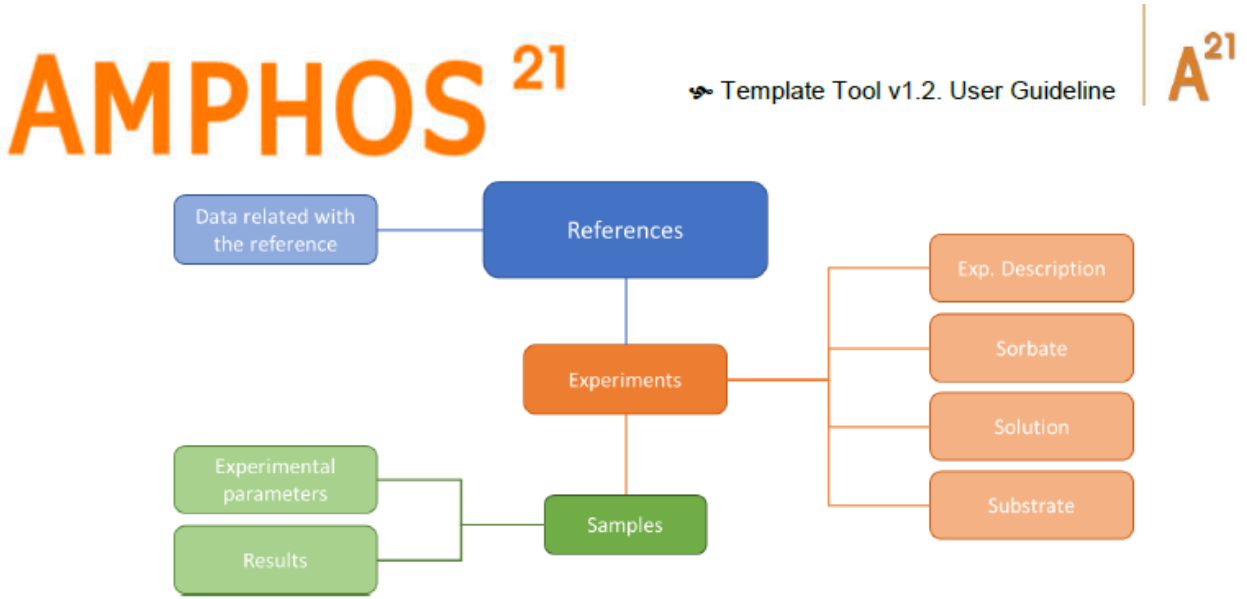


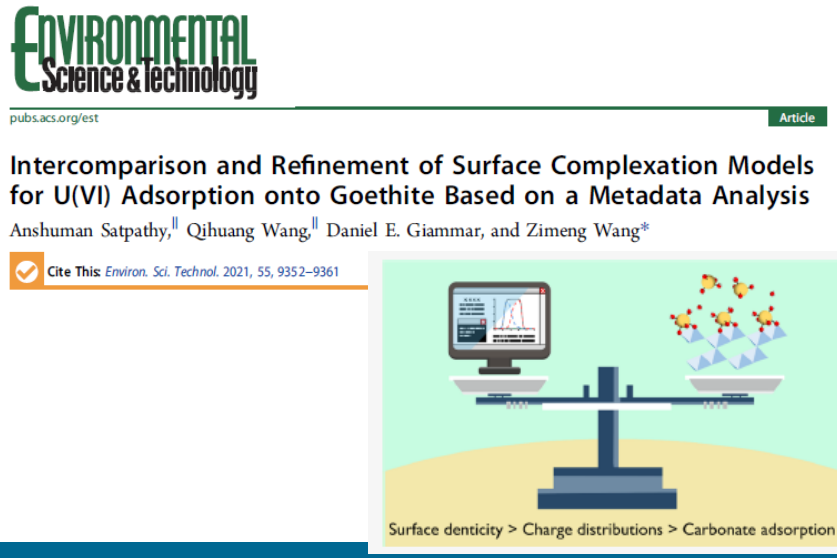
RES³T

Rossendorf Expert System for
Surface and Sorption Thermodynamics

International engagements in L-SCIE development

- Amphos21 (David Garcia) in collaboration with **Belgian Agency for Radioactive Waste Management, Ondraf-Niras (Stéphane Brassinnes)**
- PSI (Maria Marquez) effort to develop database for clay sorption data
- Fudan U. (Zimeng Wang)





A mechanistic description of Ni and Zn sorption on Na-montmorillonite
Part I: Titration and sorption measurements

Bart Baeyens *, Michael H. Bradbury

Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland

Received 23 October 1996; revised 11 December 1996; accepted 11 December 1996

Sorption modelling on illite Part I: Titration measurements and the sorption of Ni, Co, Eu and Sn

M.H. Bradbury, B. Baeyens *

Laboratory for Waste Management, Paul Scherrer Institut, 5232 Villigen PSI, Switzerland

Received 14 July 2008; accepted in revised form 10 November 2008

Research topic in Frontiers in Nuclear Engineering

Research Topic

Sorption Processes in Nuclear Waste Management: Data knowledge Management and New Methodologies for Data Acquisition/Prediction

Submit your abstract | Submit your manuscript | Participate

frontiers in Nuclear Engineering

Radioactive Waste Management

Topic Editors

David García
Amphos 21 (Spain)
Barcelona, Spain

Follow

10 publications

Romain V.H. Dagnelie
Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA)
Paris, France

Follow

20 publications

Mavrik Zavarin
Lawrence Livermore National Laboratory (DOE)
Livermore, United States

Follow

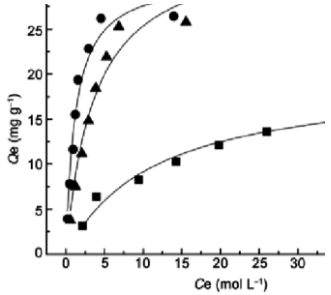
96 publications

Quantifying mineral-based radionuclide retardation

Approach

- Langmuir Isotherm

Visualization



Xiang et al. (2013)

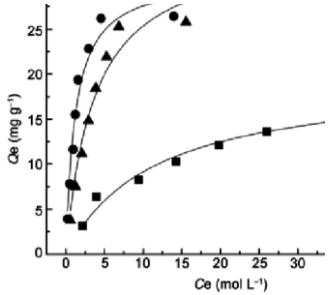
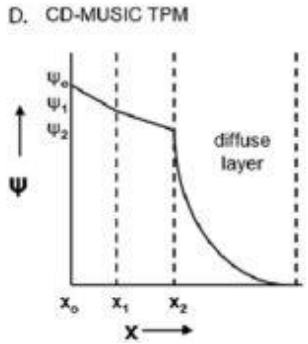
Method

- Empirical fitting.

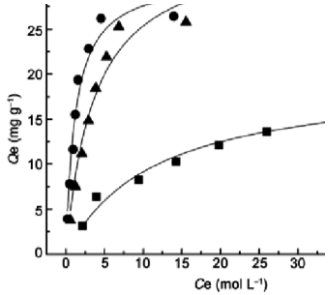
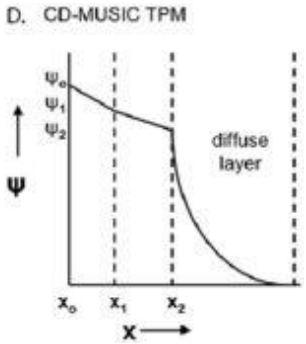
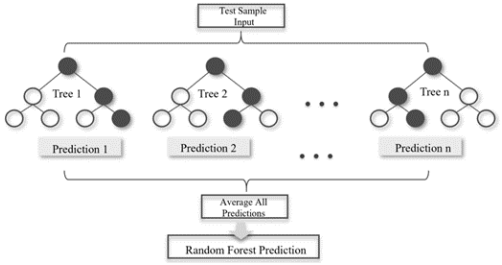
Limitation

- No mechanism inferred.

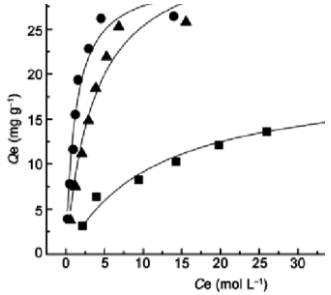
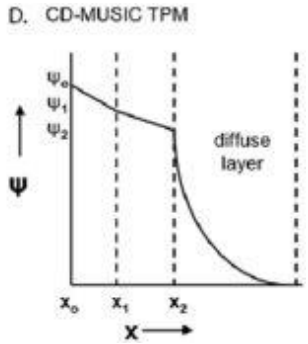
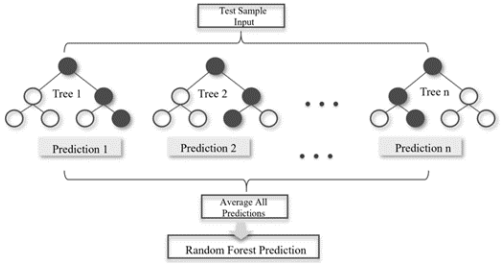
Quantifying mineral-based radionuclide retardation

Approach	Visualization	Method	Limitation
▪ Langmuir Isotherm	 Xiang et al. (2013)	▪ Empirical fitting.	▪ No mechanism inferred.
▪ Surface Complexation Model	 Goldberg et al. (2007)	▪ Fit with mechanistic descriptions and simplifications.	▪ Restrictive in assumptions.

Quantifying mineral-based radionuclide retardation

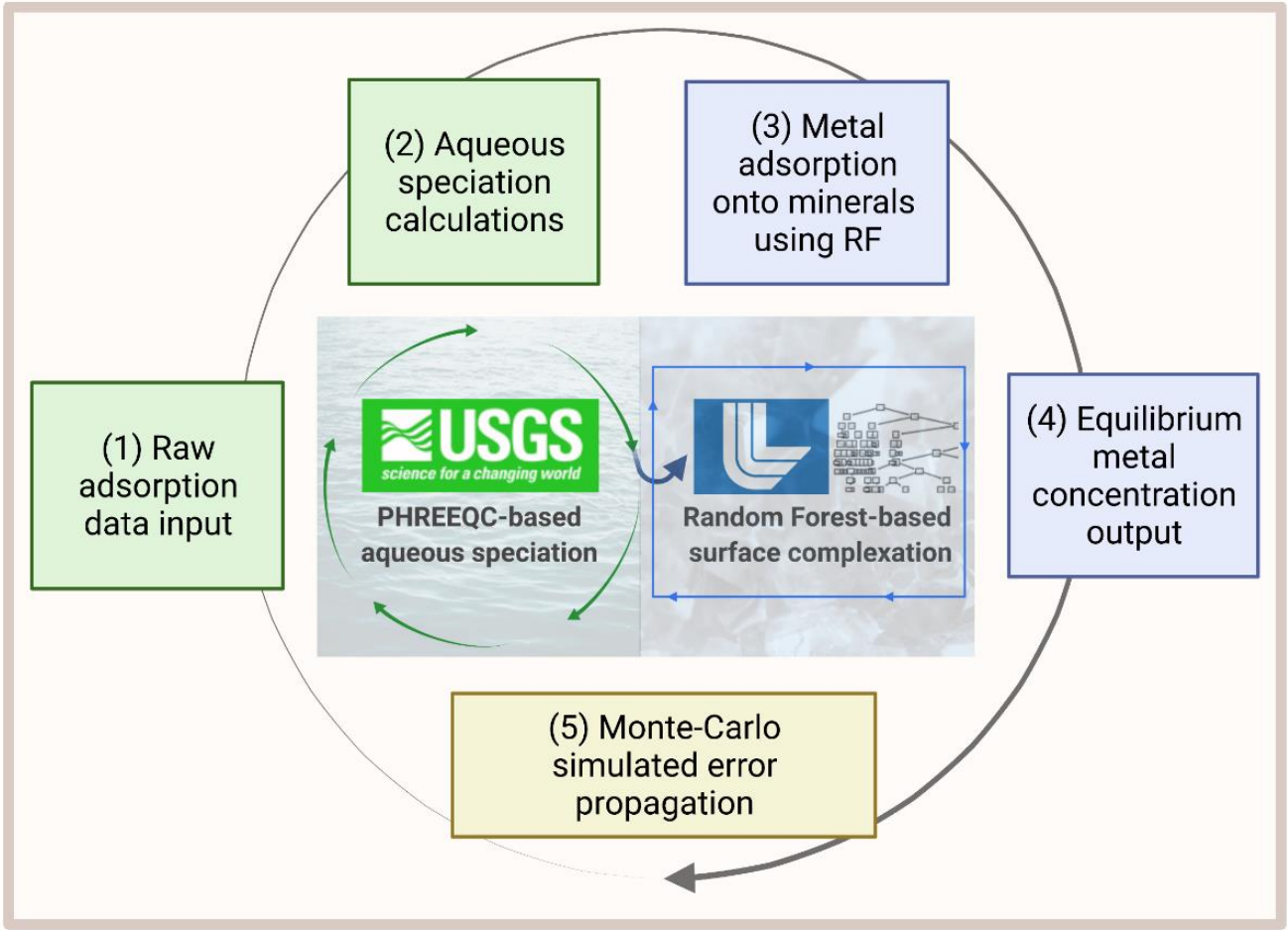
Approach	Visualization	Method	Limitation
▪ Langmuir Isotherm	 Xiang et al. (2013)	▪ Empirical fitting.	▪ No mechanism inferred.
▪ Surface Complexation Model	 Goldberg et al. (2007)	▪ Fit with mechanistic descriptions and simplifications.	▪ Restrictive in assumptions.
▪ Machine Learning Algorithm	 Beigzadeh et al. (2020).	▪ Data-driven regression development.	▪ Pure ‘black-box’ approach.

Quantifying mineral-based radionuclide retardation

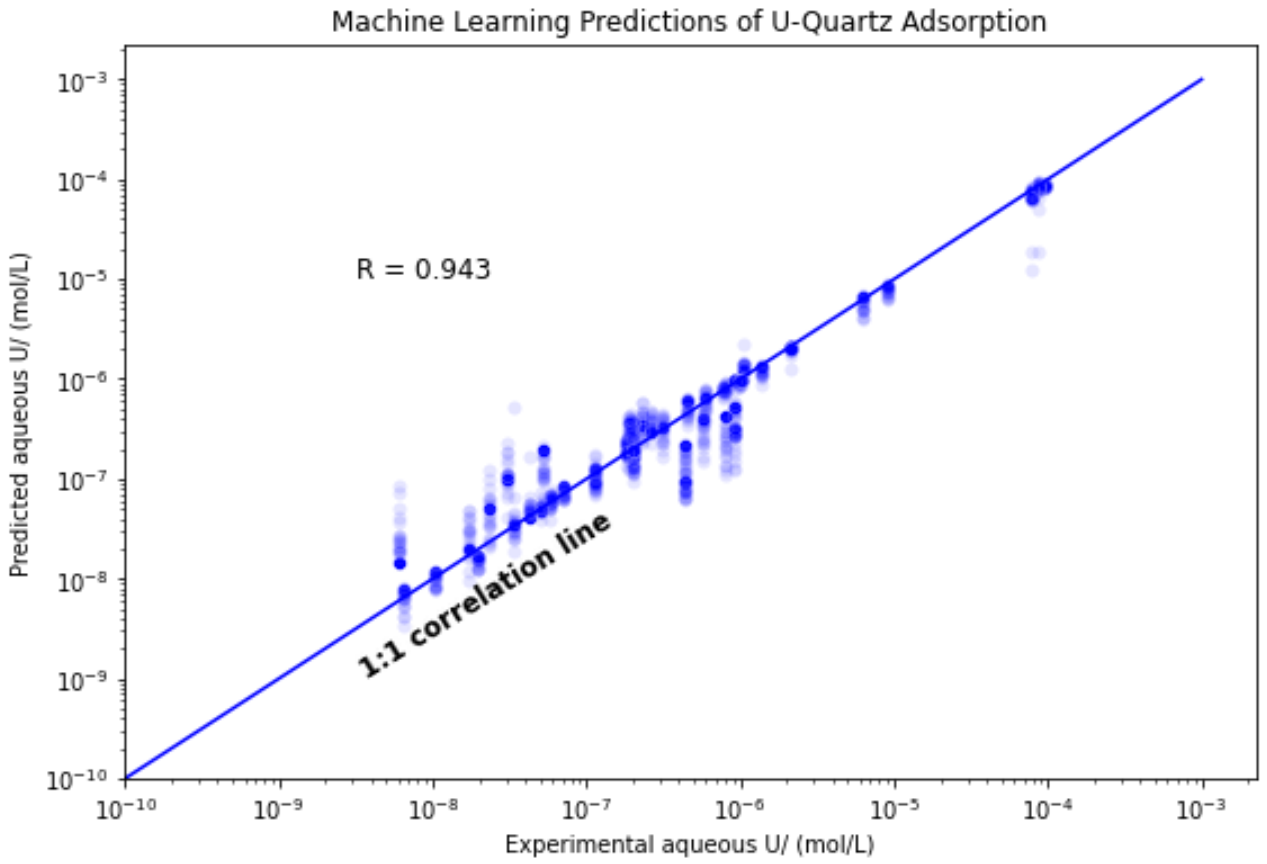
Approach	Visualization	Method	Limitation
▪ Langmuir Isotherm	 Xiang et al. (2013)	▪ Empirical fitting.	▪ No mechanism inferred.
▪ Surface Complexation Model	 Goldberg et al. (2007)	▪ Fit with mechanistic descriptions and simplifications.	▪ Restrictive in assumptions.
▪ Machine Learning Algorithm	 Beigzadeh et al. (2020).	▪ Data-driven regression development.	▪ Pure ‘black-box’ approach.

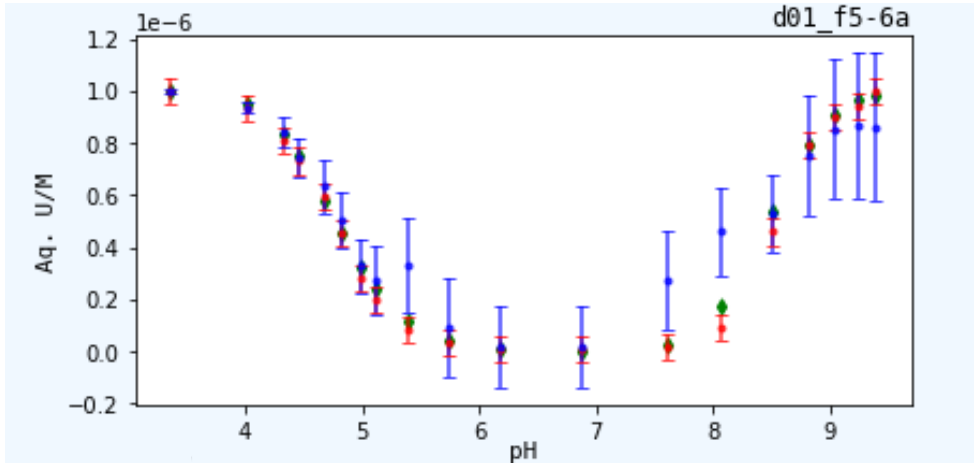
Hybrid ML approach to quantifying mineral-fluid partitioning

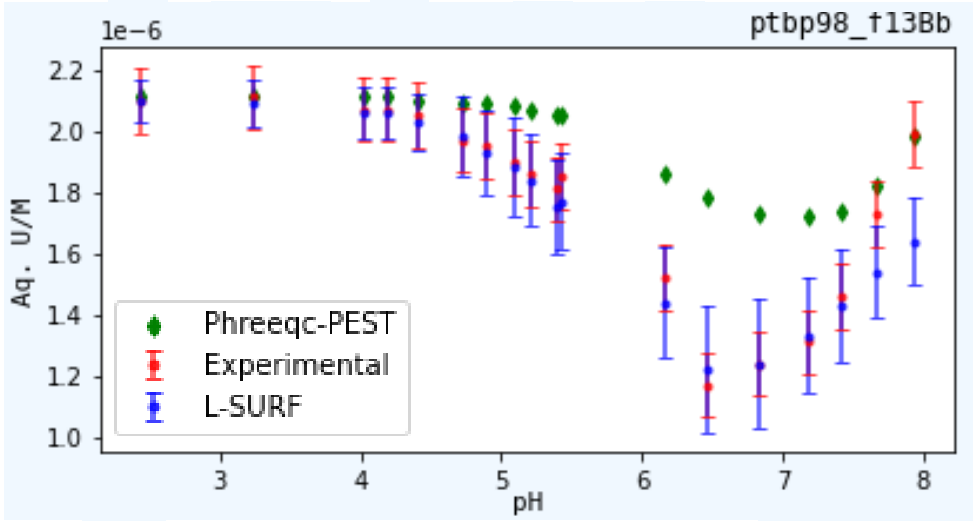
- (1) Adsorption data, thermodynamic databases are **imported**.
- (2) Aqueous speciation calculations are conducted, and important geochemical features are **stored**.
- (3) Features are inputted to **train and test** a random-forest model describing mineral-metal interactions.
- (4) Equilibrium metal concentration **outputted**.
- (5) Monte-Carlo iterations run to **propagate uncertainty**.



Results #1: Capacity to generate high quality predictions

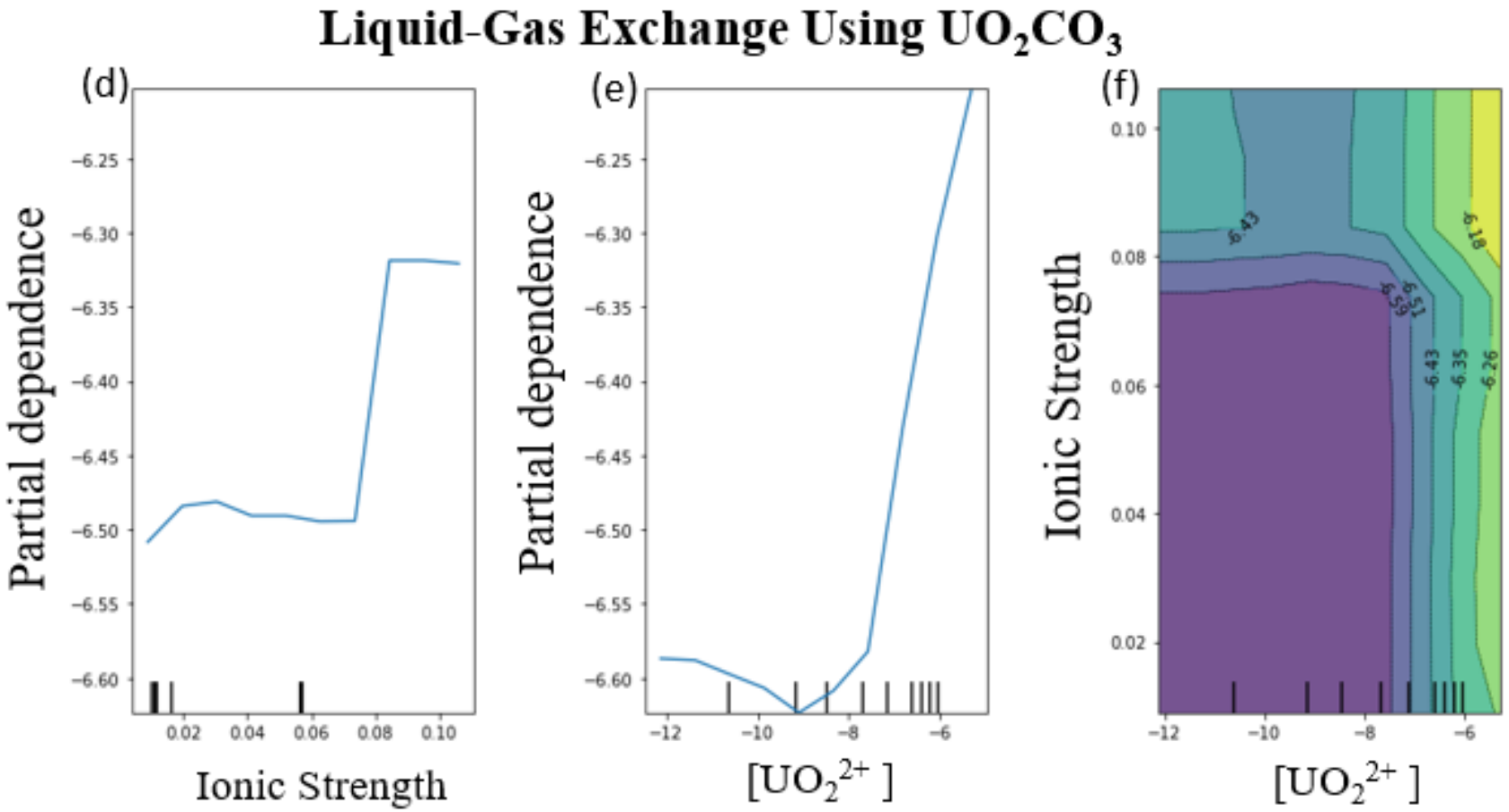






Results #2: Data-driven method to conduct sensitivity analysis

- Highlight parameter spaces that most readily impact sorption



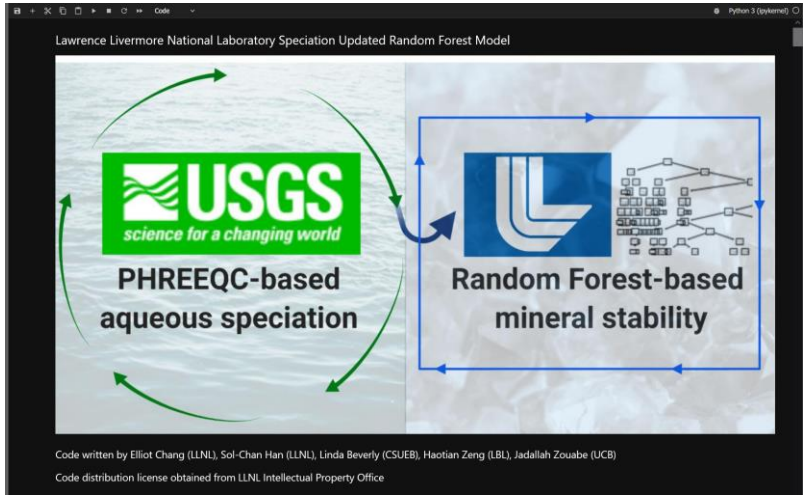
Concluding remarks

- LLNL's new database assimilates FAIR community data,
- *Enabling the new possibility for…*
- LLNL's hybrid ML modeling:
  - High throughput, high quality predictions
- LLNL's automated surface complexation modeling:
  - Ability to quickly update reaction constants

- Increased power of modeling through international engagements with nuclear waste community.

- "Collecting and generating data is outpacing its assimilation, interpretation, and understanding" -B. Helland @ AI4ESP Introduction, October 25, 2021

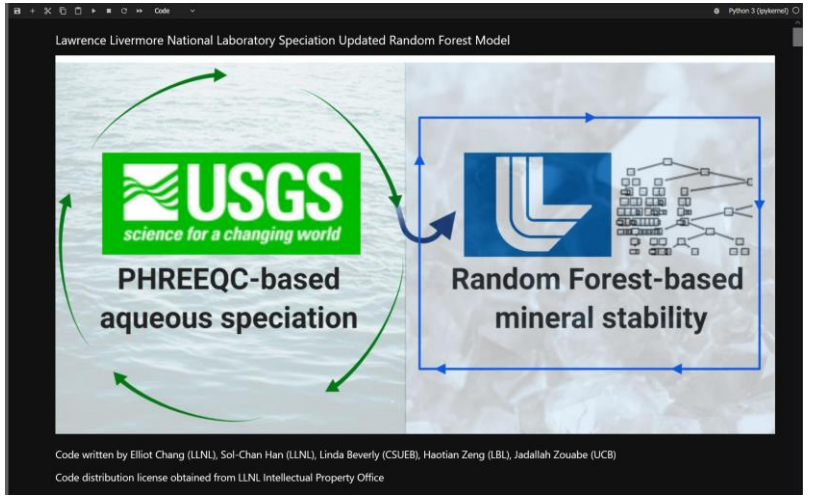
Publicly Available License to LLNL Work Package (July 2022)



Concluding remarks

- LLNL's new database assimilates FAIR community data,
- *Enabling the new possibility for…*
- LLNL's hybrid ML modeling:
  - High throughput, high quality predictions
- LLNL's automated surface complexation modeling:
  - Ability to quickly update reaction constants
- Increased power of modeling through international engagements with nuclear waste community.
- "Collecting and generating data is outpacing its assimilation, interpretation, and understanding" -B. Helland @ AI4ESP Introduction, October 25, 2021

Publicly Available License to LLNL Work Package (July 2022)



Concluding remarks

- LLNL's new database assimilates FAIR community data,
- *Enabling the new possibility for…*
- LLNL's hybrid ML modeling:
  - High throughput, high quality predictions
- LLNL's automated surface complexation modeling:
  - Ability to quickly update reaction constants
- Increased power of modeling through international engagements with nuclear waste community.
- "Collecting and generating data is outpacing its assimilation, interpretation, and understanding" -B. Helland @ AI4ESP Introduction, October 25, 2021

Publicly Available License to LLNL Work Package (July 2022)