Los Alamos
NATIONAL LABORATORY

Machine Learning for Hyperspectral X-ray Imaging Analysis

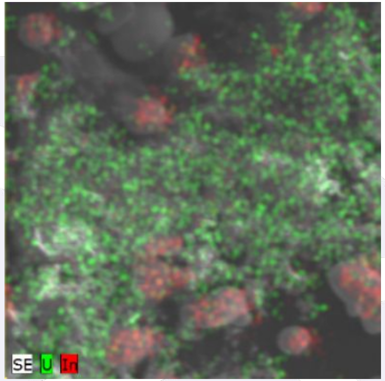
Matthew Carpenter, NEN-1 (PI)
Emily Stark Teti, CCS-3
Sebastian Salazar, NEN-1

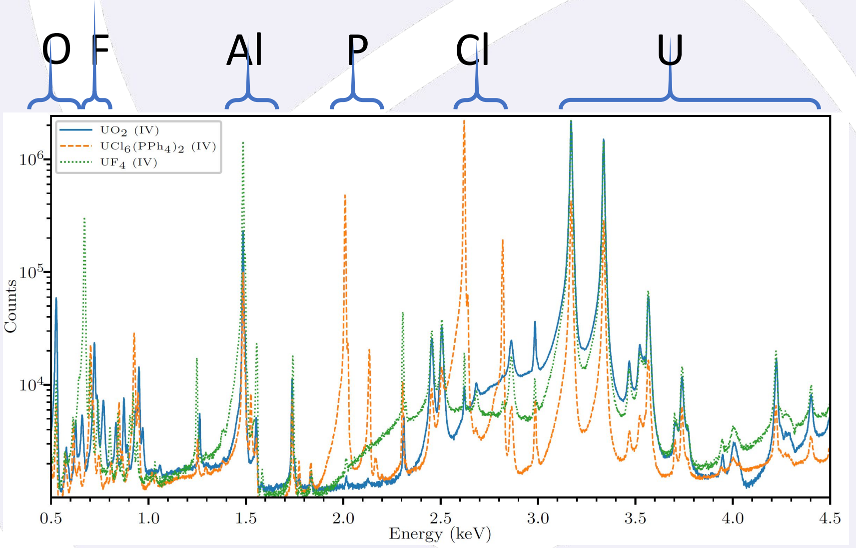
September 13, 2021

LA-UR-21-28927

Background and Goals: Signatures in XES

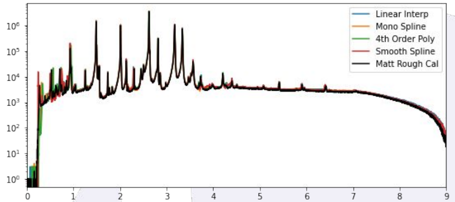
- New instrument: Hyperspectral X-ray Imaging detector takes high-resolution x-ray emission spectra of samples in electron microscope
- Problem: How can we *quickly* identify relevant sample parameters (like actinide chemical species or oxidation state)?
  □*Rapid feedback is key*
- This project: study machine learning methods of spectral analysis for x-ray emission, build tools for application to future instrument

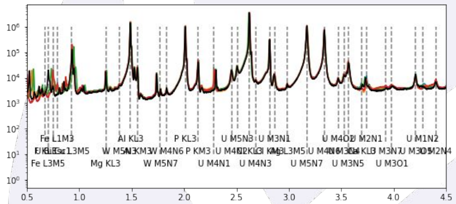


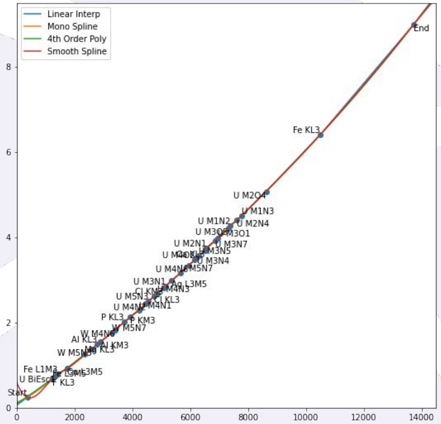


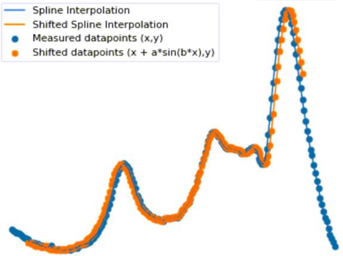
Data Handling and Pre-Processing

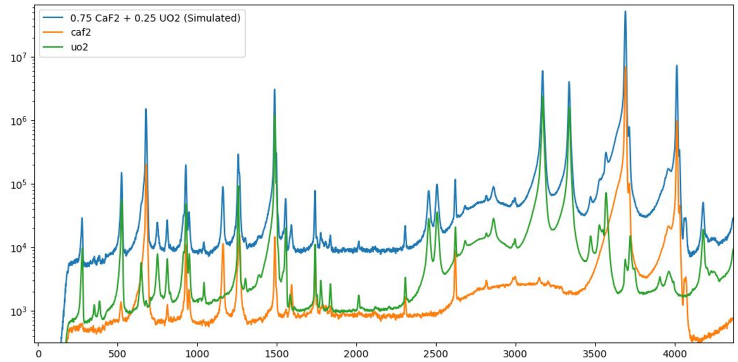
- Spectra are difficult to calibrate: non-linear, many pixels
  □Create guided-learning calibration routine to identify peaks/features and apply different calibration functions
- Available training data is sparse (1-3 spectra per sample)
  □Create Monte-Carlo-based spectrum generator that resamples and combines spectra + contaminants from pool, and can alter energy calibration







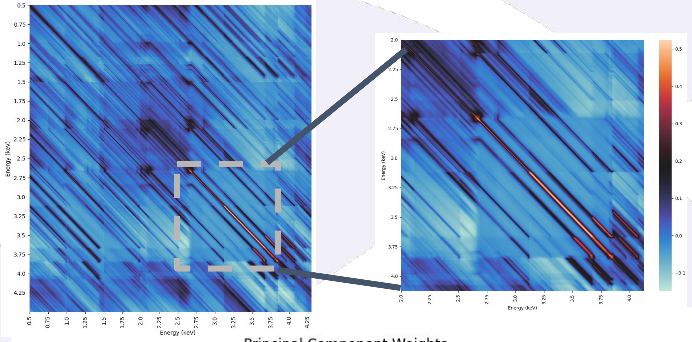


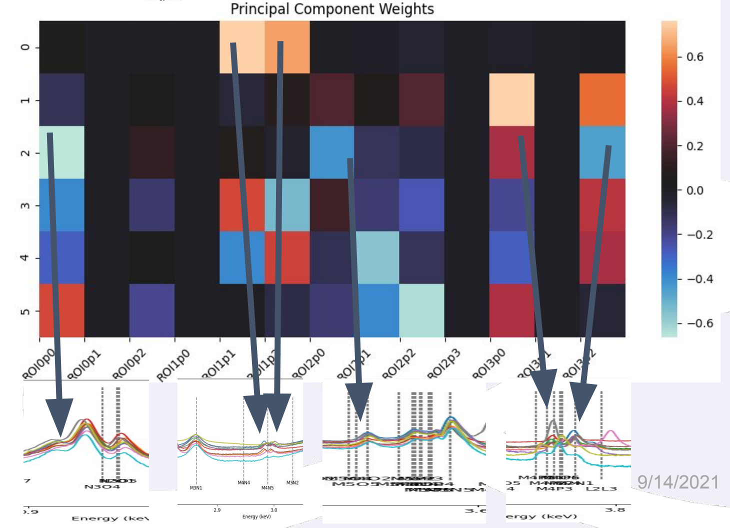


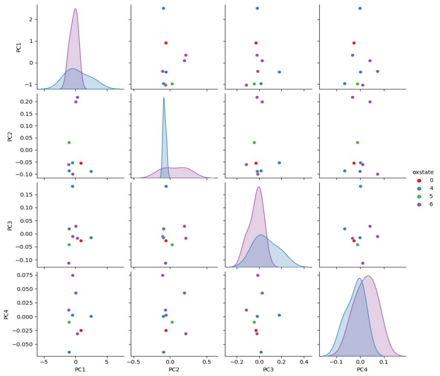
Traditional Machine Learning and Data Discovery

- Data discovery: identify regions of interest (ROIs) that have high information content
  - Develop relevant statistics to detect correlation in spectral features
- Analyze core data set with principal component analysis (PCA)
  - Unsupervised and supervised learning

Correlation Matrix

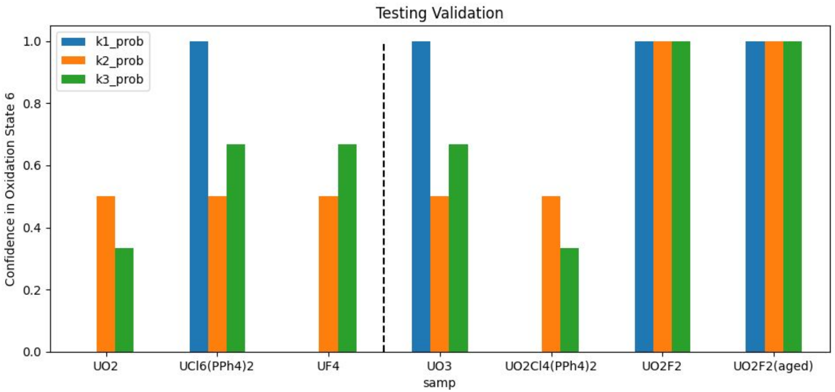






PCA Outcomes

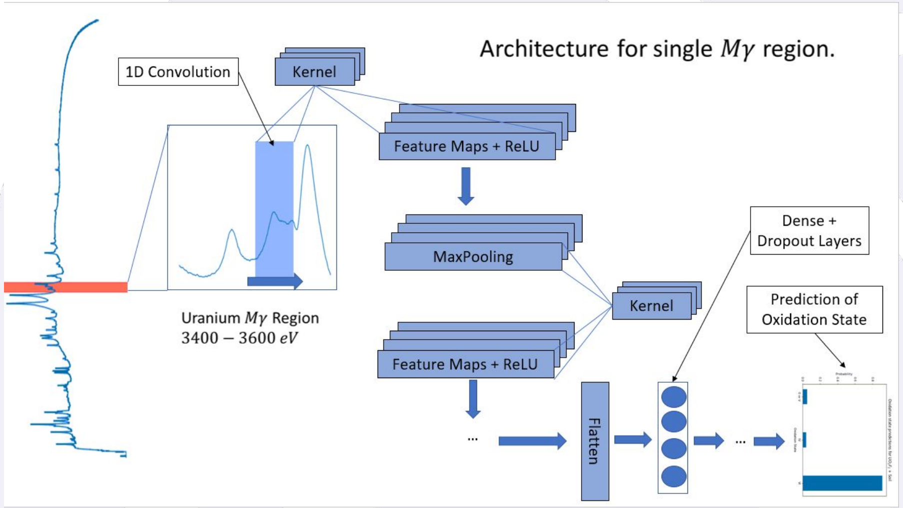
Both unsupervised and supervised models fail to predict oxidation state of test data



Model		Test Accuracy
Logistic Regression		3/7
K-Nearest Neighbors	K = 1	3/7
	K = 2	3/7
	K = 3	4/7
Decision Tree		3/7

Convolutional Neural Network (CNN)

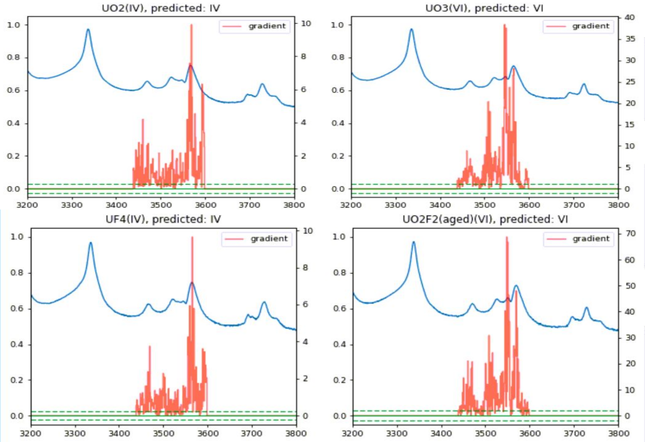
- Train and verify on generated data sets (40k to 100k+ spectra) with varying parameters
- Approaches:
  - Full spectrum
  - Select ROIs (variable)
  - Multiple training instances for statistical variation
- Output:
  - Chemical species (e.g. UF4, UO2)
  - Oxidation state (e.g. IV or VI)

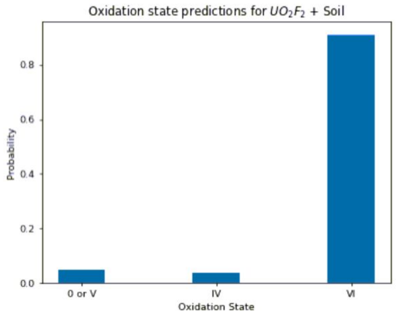


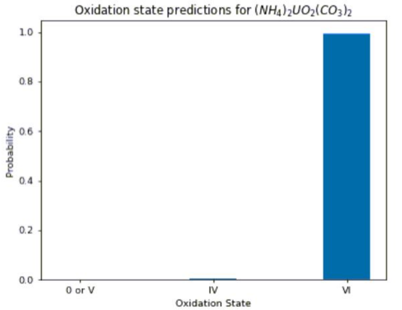
CNN Results

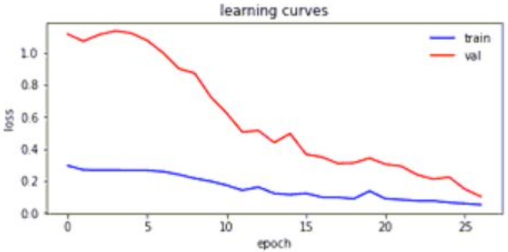
- CNNs are able to reliably determine oxidation state using 1 or more ROIs
- Chemical species is less reliable, but there is good promise for progress with further development
- Essential to “salt” training data with noise (contaminants + energy calibration shift) for reliable training

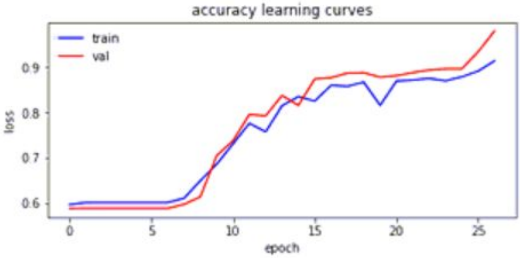
Saliency Maps





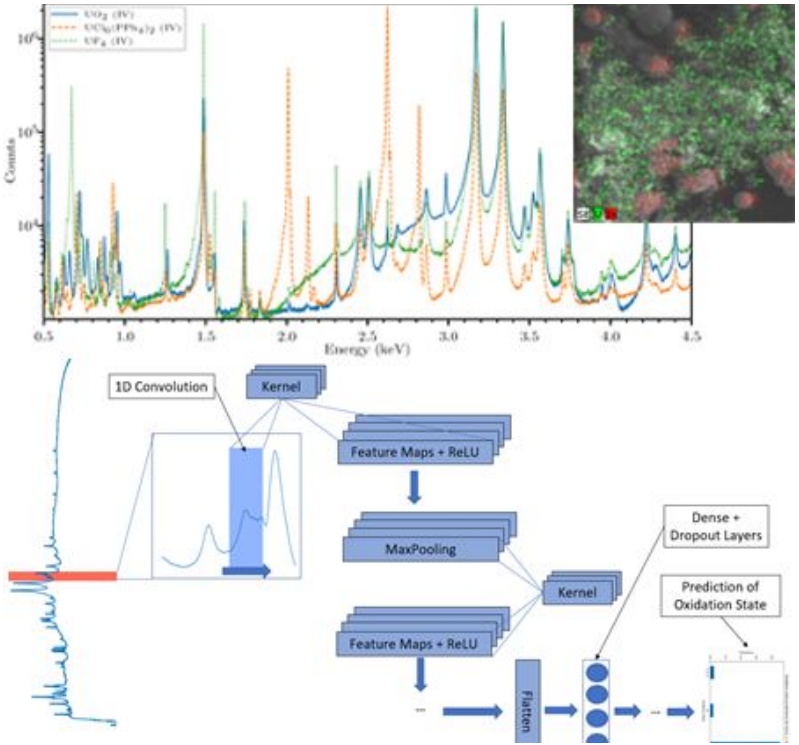






Los Alamos
NATIONAL LABORATORY

Machine Learning for Hyperspectral X-ray Imaging Analysis



The Machine Learning for Hyperspectral X-ray Imaging Analysis project investigates multiple machine learning methods to determine actinide chemical state from measured x-ray emission spectra. This will allow spectra measured in an electron microscope (top left) to be analyzed in real-time to create chemical state maps (top right), using, for example, convolutional neural networks (bottom).

Project Description

Research machine learning approaches to chemical speciation analysis in high-resolution x-ray emission spectroscopy

Project Outcomes

- -Created tools for data handling/preprocessing: DDAC and MCS
- -Studied feature-based ML models (PCA), conclusion: insufficient for purpose
- -Studied CNN models (oxidation state, species), conclusion: promising results, useful for analysis as-is
- -Follow-on work: expand CNN models with new compounds, integrate into future HXI instrument processing stream
- -Student applied to postdoc position
- -Two students visited LANL for first time (worked remotely)
- -Intend to present and submit proceedings paper at CoDA 2022 conference

PI: Matthew Carpenter

Total Project Budget: $67,000

ISTI Focus Area: Data Science and Artificial Intelligence Infrastructure

END