Interagency Technical Nuclear Forensics Technical Review

SAND2018-7346PE

Statistical Support for Experimental Design/Planning for Plutonium Signatures Project

John Lewis, Adah Zhang, Daniel Ries (SNL)

Christine Anderson-Cook (LANL)

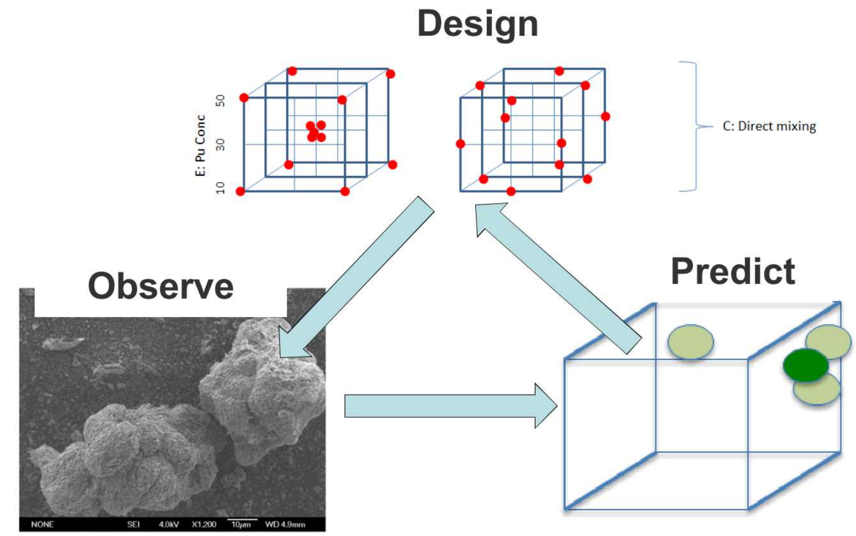
08/02/2018

SNL.16.002/17.002, LANL.16.009/17.010
Funding Agency: DHS CWMD

Statistical Support for Design/Analysis of Experiments

Objective: 1) Facilitate efficient data collection through designed experiments to accelerate learning. 2) Characterize relationships to understand origin of material using statistically sound analysis strategies that maximize understanding of experimental results.

Relevance: Maximizing what can be learned from experimentation given the high costs of collecting data.



Approach: Collaboration with Pu Expert Panel and other SMEs to make best use of resources during experimentation. Using innovative leading-edge statistical methods to model relationships and predict new data.

Personnel Support: 4 statisticians (1 @ LANL, 3 @ SNL)

Accomplishments/Results:

- Analyzed/presented initial set bench-scale Pu(III) data
- Designing Pu(IV) Studies: Identified factors, constructed initial sequential design options.
  Inverse Prediction methodology Development
- Two accepted peer-reviewed manuscripts
- One submitted
- Four conference talks
  - Joint Research Conference, June 2018.
  - INMM Proc. & Sig. Disc. Workshop, May 2018
  - Actinide Separations Conference, May 2017
  - Joint Statistical Meetings, Aug 2016

Overview

SNL Statistical Sciences Group

Statistical Sciences, Org 9136

SANDIA STATISTICAL SCIENCES

Plan. Analyze. Inform.

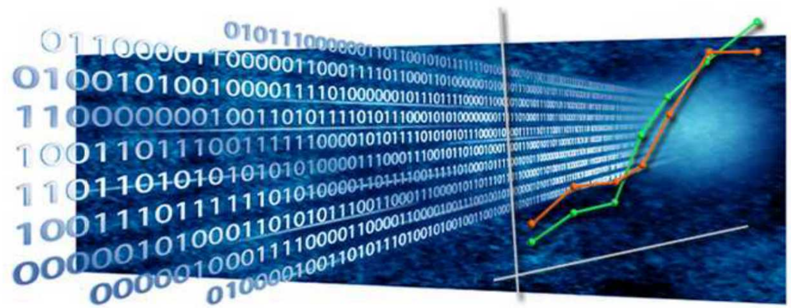
8 PhD Statisticians
7 MS Statisticians

Sandia's Statistical Sciences department provides innovative statistical solutions for complex national security problems through research, collaboration, and education. We develop and deliver tools and methods to better inform the collection, analysis, and interpretation of data.

SNL supports both pre/post detonation nuclear forensics including Pu/U Signatures projects, multivariate algorithms enabling group inclusion/exclusion classification, and analyses of NUDET signals.

LANL Statistical Sciences Group

CCS-6



26 Ph.D. Statisticians
3 MS Statisticians

The LANL Statistical Sciences group provides statistical reasoning and rigor to multidisciplinary scientific investigations and the development, application, and communication of cutting-edge statistical sciences research.

LANL supports capability development for both pre and post detonation nuclear forensics including design of experiments, modeling and analysis through Pu and U focused projects.

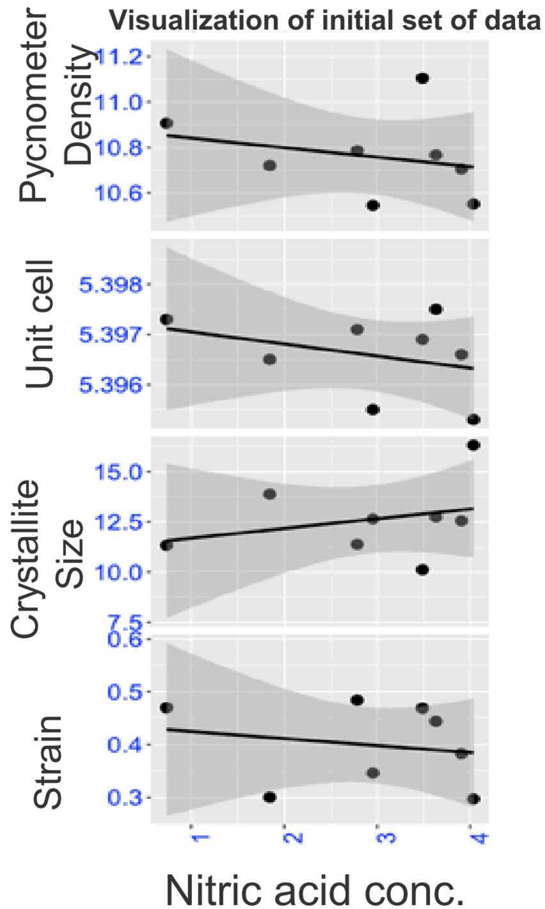
Research/Project Description

Goal: Identify signatures of nuclear forensic value in plutonium (Pu) materials that can be related to the processing conditions used to produce them.
Approach: Consulting with experts, design experiments to collect relevant data.
Ongoing/Potential Studies: Pu(III), Pu(IV), metal – each requires a staged approach

Pu (III)

Pu Expert Panel (PEP) recommended bench-scale experiments to study signatures derived from PuO_2 produced via Pu(III) oxalate precipitation process.

Designed Experiments currently ongoing at PNNL



Planned: 76 bench-scale (10g) runs to inform lab-scale (200g) runs.

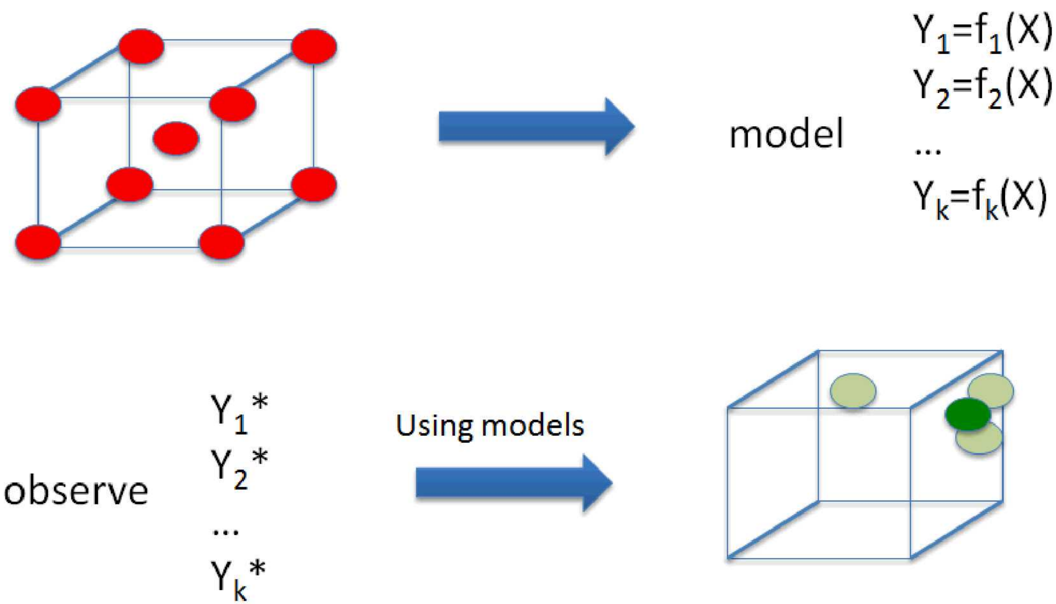
Pu(IV) and Metal

Similar approach ongoing. Consulting with PEP and other experts to:

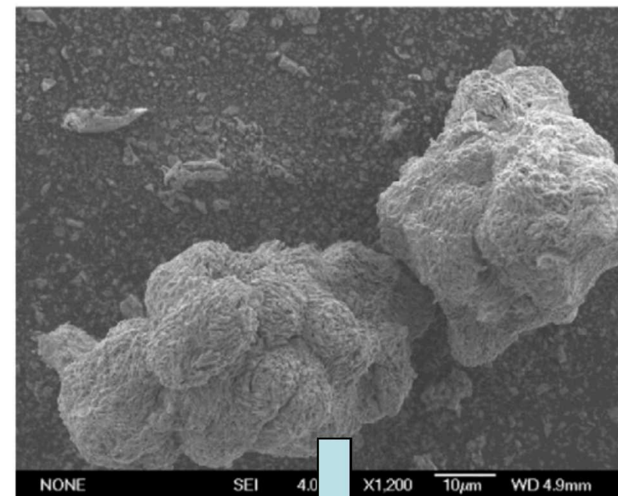
- Understand currently available data
- Identify areas needing study
    - E.g., Scaling effects
- Identify important processing parameters to study

Research/Project Description (cont'd)

- The long term goal is to infer the set of experimental conditions used to produce material based on the material's observed properties using an **inverse modeling approach**.



Success: good ability to identify location in X-space that produced Y*



Some relevant questions:

How was this material produced?

Who produced it?

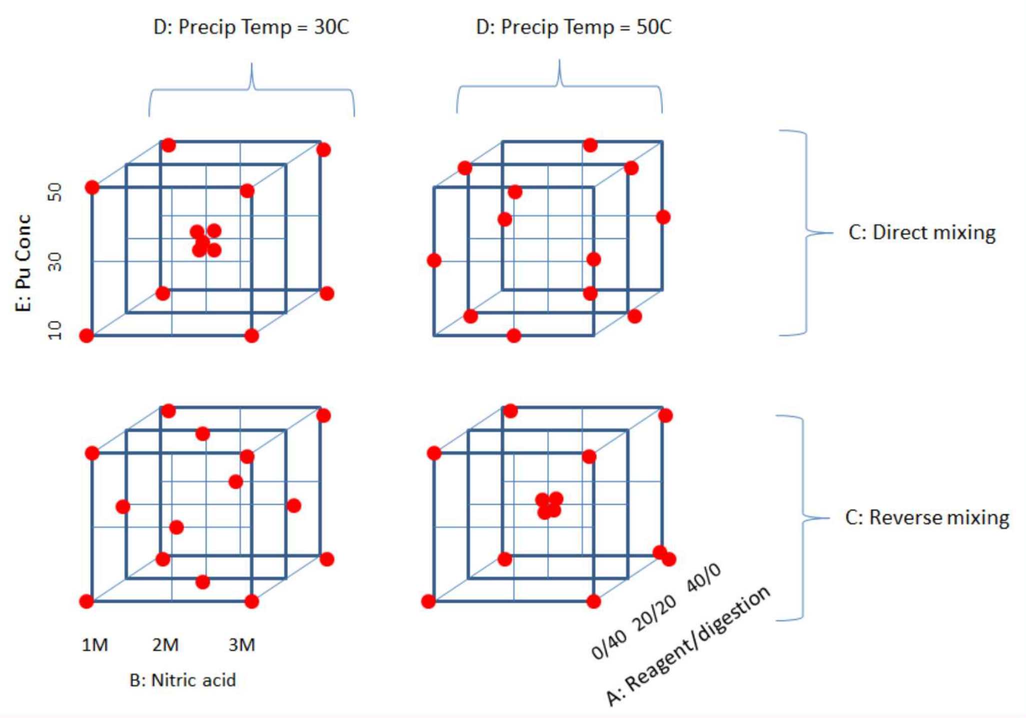
Who does it belong to?

Technical Approach: Designed Experiment

Subset of I-optimal design for the Pu (III) Oxalate Bench Scale Study:

- Sample sizes chosen to make prediction performance balanced across the design space
- I-optimality: minimizes average prediction uncertainty
    - Runs spread across design space
    - Some replication
    - Precise predictions is a good criterion for inverse prediction

Oxalate feed in solution at 0.9M



Well designed experiments help to:

- Characterize the input/output relationship
- Increase precision of inverse predictions
- Understand multi-scale issues – where should we run more expensive 200g runs?

Technical Approach: Sequential Design

Planning for possible Pu (IV) experimentation is ongoing

- **Suggested Approach:** Sequential Experimental Design in two stages geared towards precise estimation of the **input/output relationship**.
- True relationship is often approximated with a '**Second Order Model**':

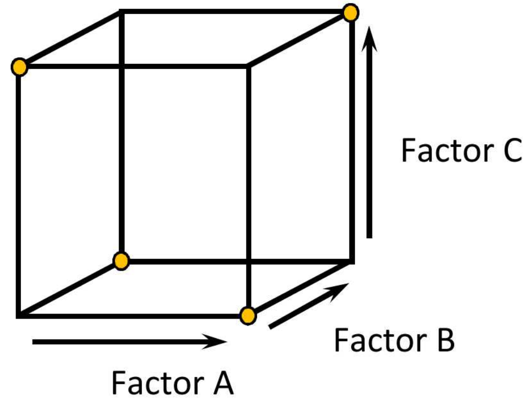
$$Y \approx \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_1 X_2$$

Intercept	Main Effects (at least 2 levels)	Quadratic Effects (at least 3 levels)	Interaction Effect (at least 2 levels)
β_0	$\beta_1 X_1 + \beta_2 X_2$	$\beta_{11} X_1^2 + \beta_{22} X_2^2$	$\beta_{12} X_1 X_2$

Technical Approach: Sequential Design

Suggested Approach: Sequential Experimental Design in two stages

Stage One – 'Screening Experiment'

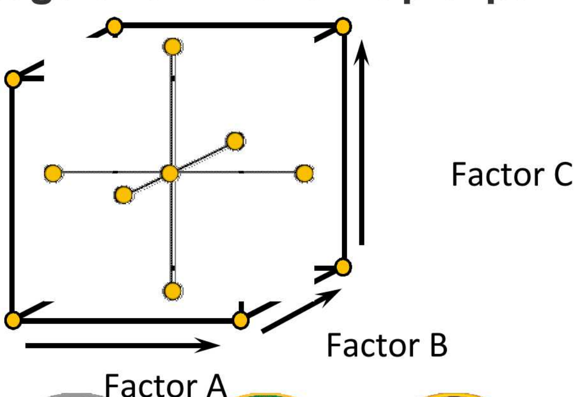


Many factors - identify the most important ones with a relatively small number of runs

Often 'Supersaturated' - i.e. cannot estimate **all** effects separately from each other

Example – Definitive Screening Designs (DSD): Can estimate quadratic effects when 3 or fewer factors are 'active'/important

Stage two – 'Follow-up experiment'

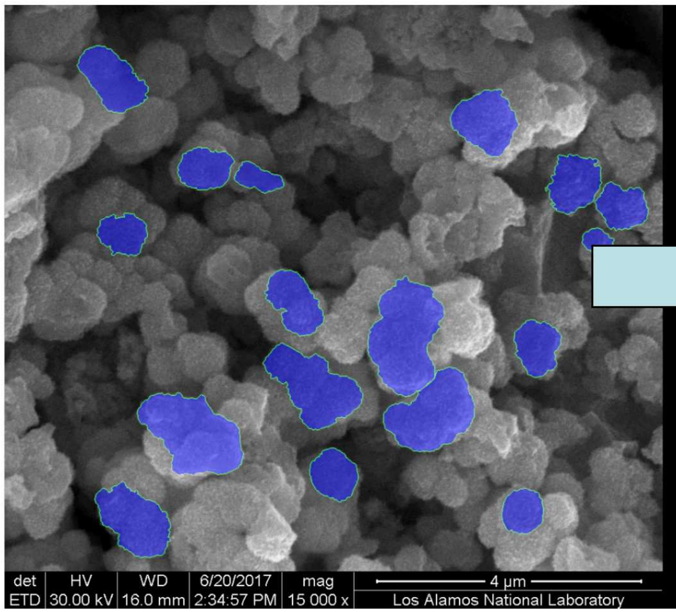


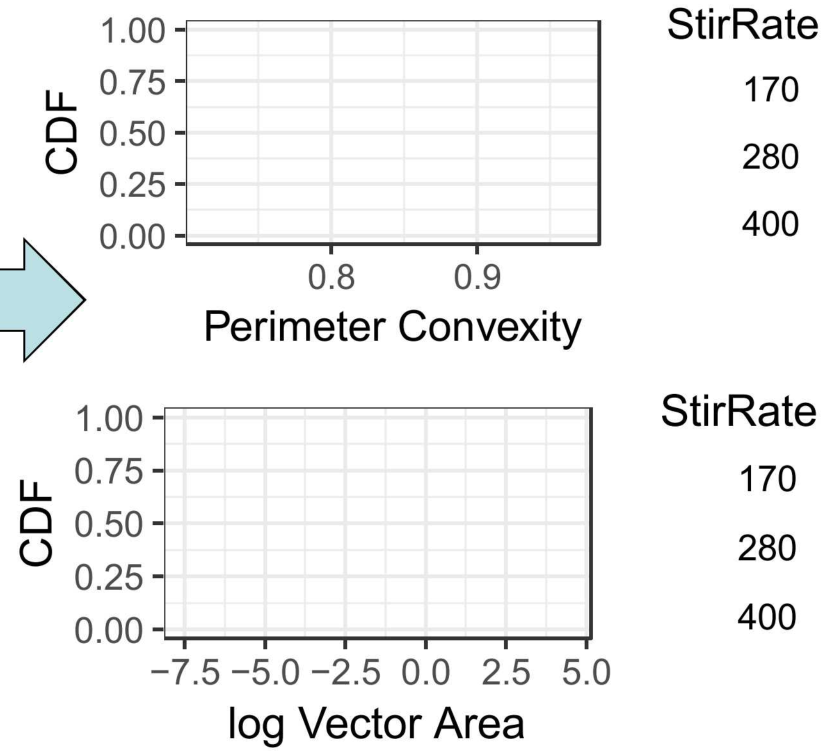
Sparsity-of-Effects Principle: there are only a few important factors which are identified in stage one

More substantial coverage of the smaller dimensional space can be achieved.

Technical Approach: Inverse Modeling

Image Segmentation of material leads to many **distributional responses** that potentially contain a wealth of nuclear forensic signature information





Sample SEM Image with Segmentation in MAMA Software

Data from a bench-scale ammonium diuranate experiment led by Marianne Wilkerson at LANL

Technical Approach: Inverse Modeling

<u>One Approach:</u>
Aggregation to summary statistics (e.g. a mean) to get scalar responses.

Regress on the means and perform inverse prediction.

$$y_j = x'\beta_j + \epsilon$$

$$x^* = \underset{x}{\operatorname{argmin}} \sum_{j=1}^{q} \left(\hat{y}_j - y_j^*\right)^2$$

$$\hat{y}_j = x'\hat{\beta}_j$$

y_j^* - new observations

<u>Example of an inverse prediction solution</u>

From Initial Experiment

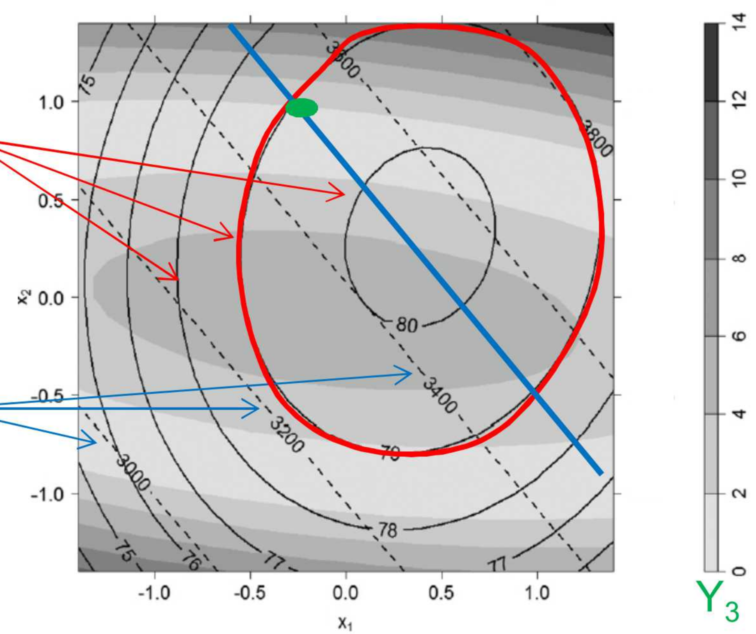
$Y_1 = f_1(X_1,\ldots,X_k) + \varepsilon$

$Y_2 = f_2(X_1,\ldots,X_k) + \varepsilon$

$Y_3 = f_q(X_1,\ldots,X_k) + \varepsilon$

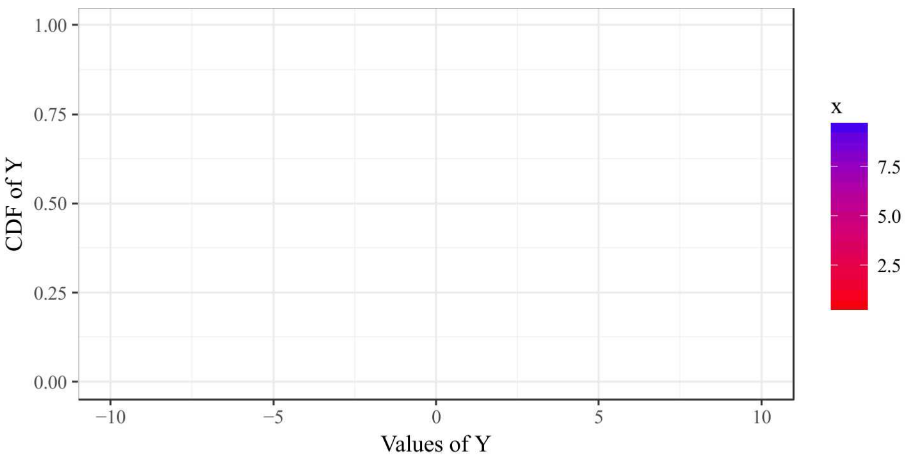
New observation
$(Y^*_1, Y^*_2, Y^*_3) = (79, 3500, 2)$

Suggests $(X_1, X_2) \approx (-0.2, 1.0)$



Technical Approach: Distributional Responses

Aggregation results in a potential loss of information



Example: input X affects the spread of the distributions, but the means remain the same.

Regression on the means will yield weak regression fits.

Better Approaches – utilize the functional nature of the responses:

- Functional Regression using segmentation results
- Image/shape analysis (potential future work) directly on the images

$$W_{ij}(t) = \boldsymbol{X}_i' \boldsymbol{\beta}_j(t) + \epsilon_{ij}(t),$$

$$\widehat{\boldsymbol{X}_i} = \arg\min_x \sum_{j=1}^{q} \int \left| \widehat{W_{ij}}(t) - W_{ij}(t)^* \right|^2 dt,$$

$$\widehat{W_{ij}}(t) = \boldsymbol{X}_i' \widehat{\boldsymbol{\beta}}_j(t).$$

This approach is more likely to capture relationships related to changes in the distributional response as a whole.

Results / Accomplishments

Pu (III) Study

- Pu (III) Oxalate Bench Scale Experiments on going
- 76-runs planned – analyzed initial set of data and presented to stakeholders and PEP

Pu (IV) Study

- Identified factors to study
- Presented *sequential design* approach to stakeholders and PEP
- Compared 6-7 ‘stage one’ designed experiments – current suggestion is roughly 30 runs for each stage of experimentation to characterize second-order statistical model.

Inverse Modeling Methodological Development

- **Two accepted peer reviewed journal articles and one submitted:**
  - Lewis, J., Zhang, A., & Anderson-Cook, C. “Leveraging Multiple Statistical Methods for Inverse Prediction in Nuclear Forensic Applications” Chemometrics and Intelligent Laboratory Systems.
  - Thomas et al. “Selecting an Informative/Discriminating Multivariate Response for Inverse Prediction” Journal of Quality Technology
  - Ries et al. “Utilizing Distributional Measurements of Material Characteristics from SEM Images for Inverse Prediction” Submitted to Journal of Nuclear Materials Management
- **Presentations:**
  - Edward Thomas presented at the Joint Research Conference on selecting an informative response set.
  - Daniel Ries presented on functional inverse regression methods at the Institute for Nuclear Materials Management Workshop.
  - Adah Zhang presented an analysis of historical Pu(IV) production data at the 41st annual Actinides Separations conference
  - John Lewis presented at the Joint Statistical Meetings on selecting an informative response set.

Future Plans

Pu (III) Study

- Continue to analyze bench-scale data as it becomes available
- Use analyses to inform larger lab-scale runs and develop inverse prediction models.

Pu (IV) Study

- Analyze currently available data on Pu(IV) and consult with PEP and stakeholders to
  - Help inform the needed experimental studies
  - Understand differences between types of production techniques (continuous, batch)
  - Understand factors that should vary and the levels required in the experiments
  - Understand scaling effects due to differences in production scale.

Inverse Modeling Methodological Development

- Account for correlations in the multivariate distributional / functional responses
- Use the SEM images directly (before segmentation) to predict production parameters.