Los Alamos NATIONAL LABORATORY

AI/ML Surrogate-Accelerated Learning and Inference in Nuclear Data Problems

Ajeeta Khatiwada

Workshop for Applied Nuclear Data Activities (WANDA) 2026

Feb 11, 2026

LA-UR-26-20774

Motivation

Nuclear data problems are often:

- High-dimensional
- Computationally expensive

Many coupled reaction channels, nuclei, energy regime etc.

Optimization and inference can require 1000s to millions repeated forward model calls

Traditional workflows:

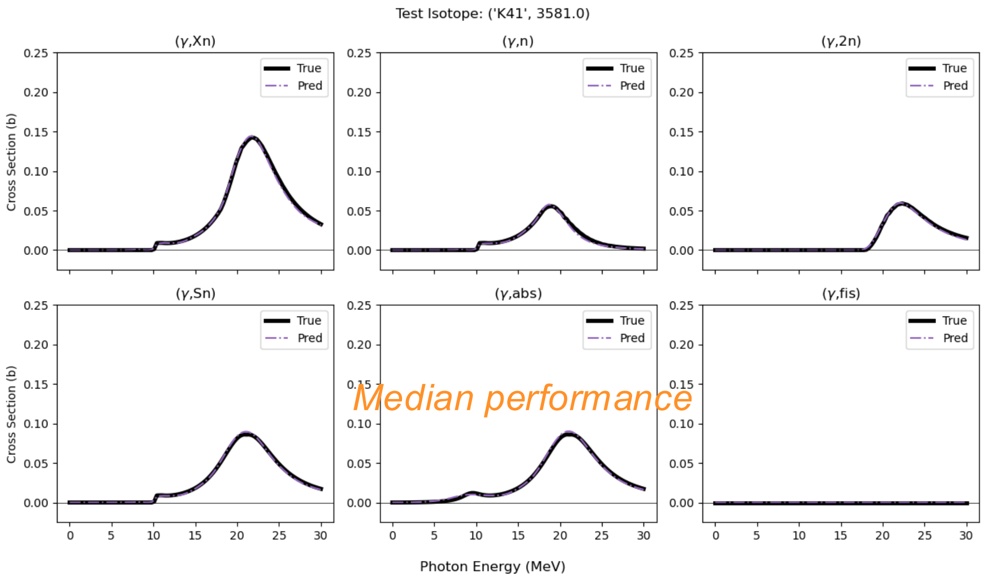
- Introduce biases and assumptions
- Limit optimization
- Limit full uncertainty exploration

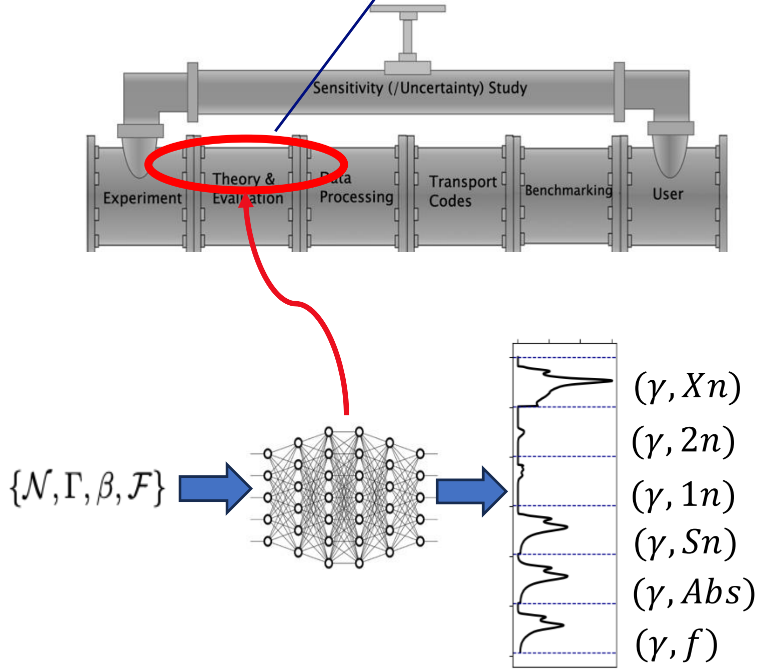
Surrogate models can enable otherwise intractable learning and inference with reduced computational load

Deep Learning Surrogate enabled MCMC evaluation of photonuclear cross sections

2.5M X speedup

Emulates the theory models (CoH_3) accurately

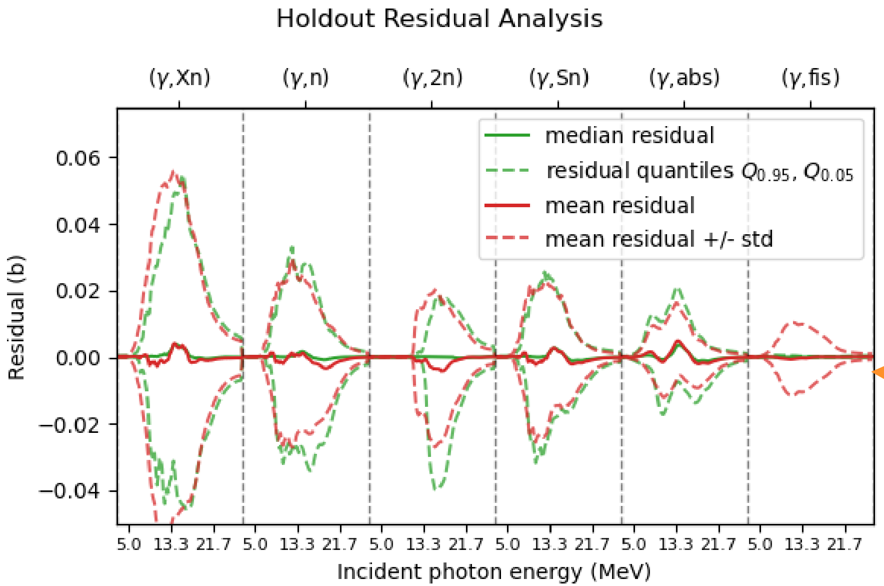


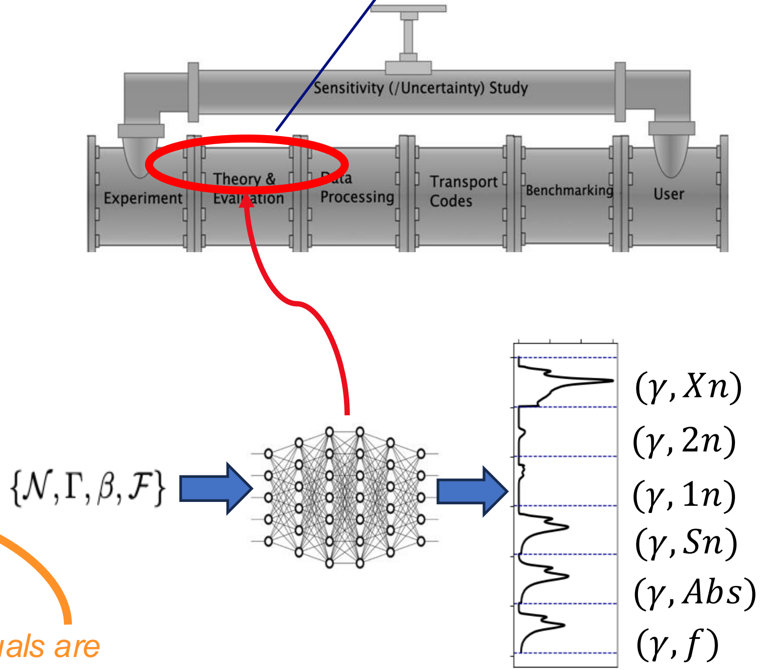


Deep Learning Surrogate enabled MCMC evaluation of photonuclear cross sections

2.5M X speedup

Accurate predictions for held-out data outside of the training phase space

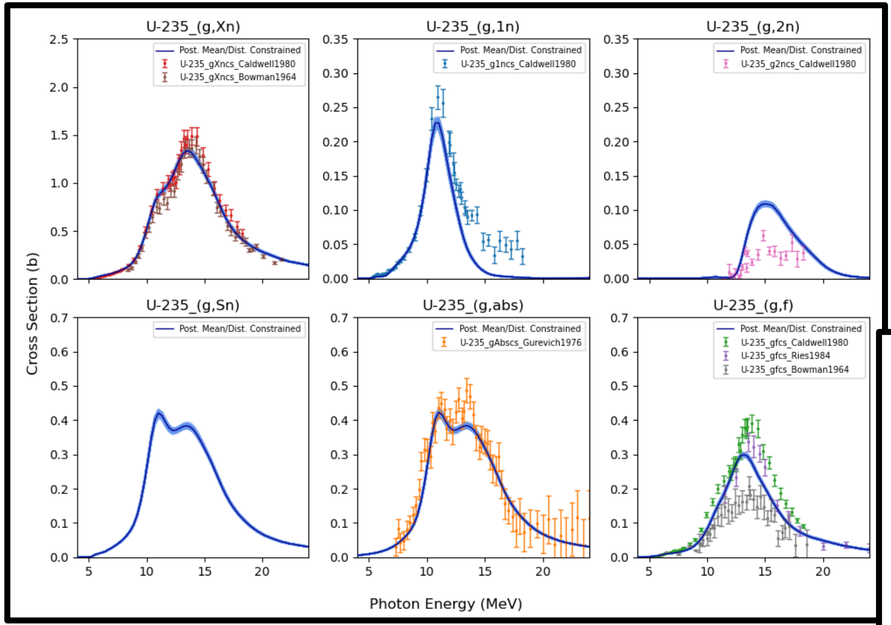


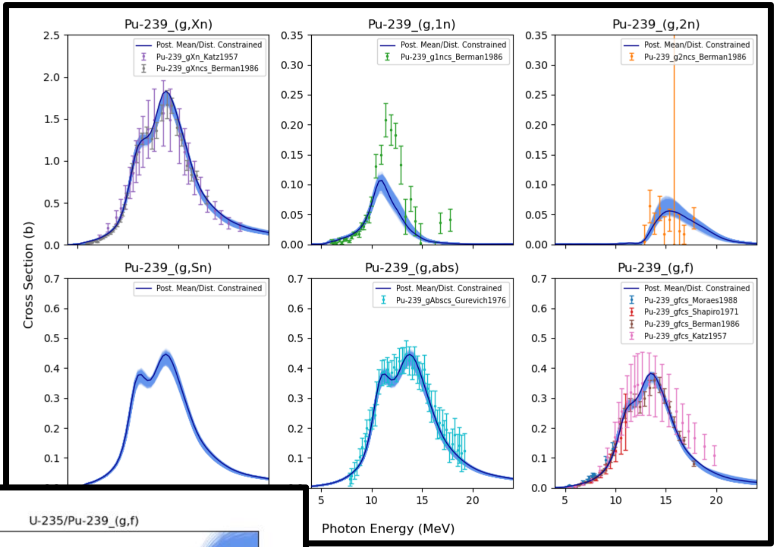


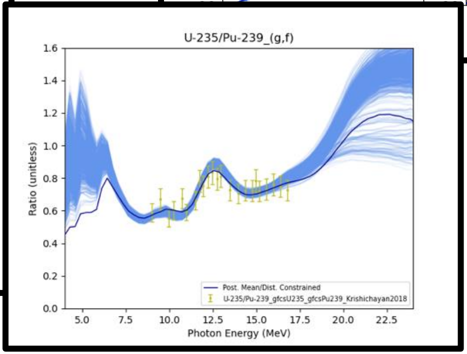
Residuals are comparable to regions well represented in training set

Surrogate enables fast full-MC evaluation

Consistent posterior inference across coupled reaction channels, nuclei, and observables







Especially valuable for photonuclear data with inconsistencies and data–model tension

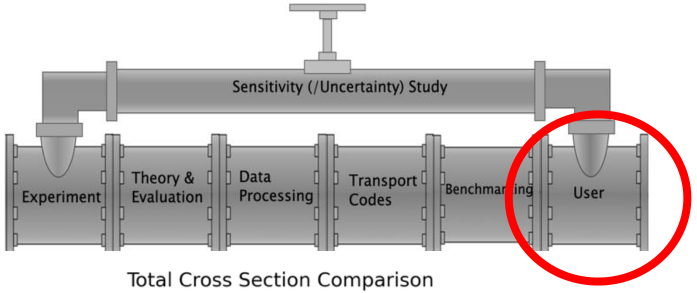
Surrogate-enabled Reinforcement Learning for Energy Group Optimization

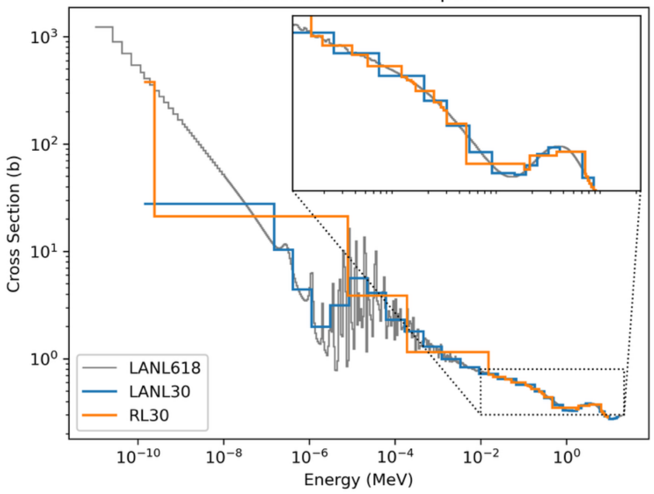
- Energy group structure used in nuclear data and transport calculations affects accuracy vs. runtime

  618 energy group solve: 14 min vs. 30 group solve: 2 s

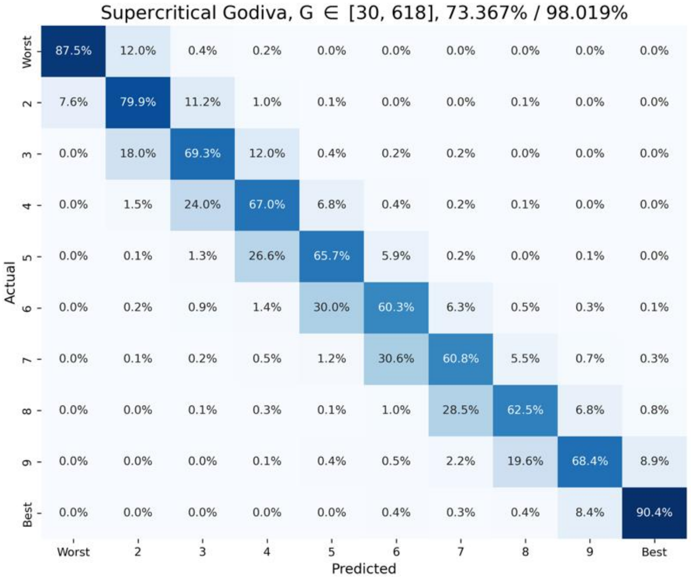
- Brute-force/group optimization are infeasible

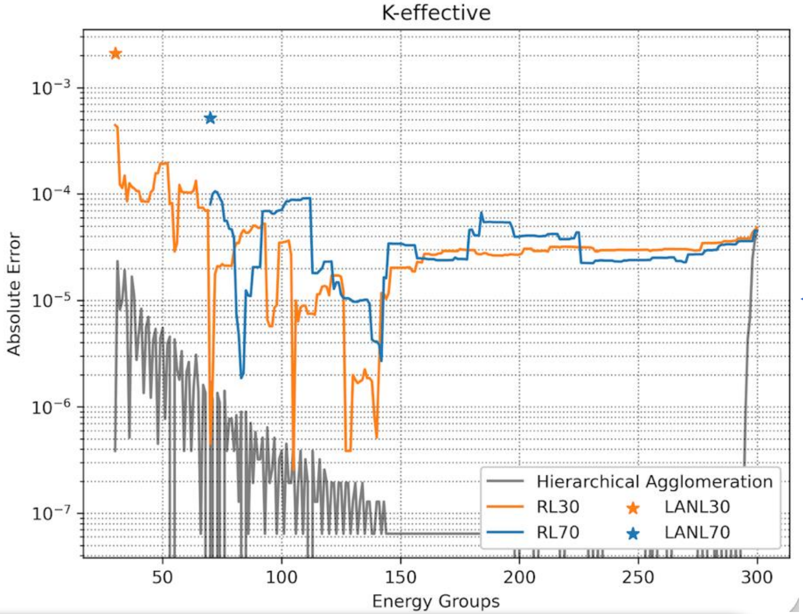
- AI/ML surrogate modeling of the transport problem using CNN + LSTM and embedded inside a Reinforcement Learning framework for group optimization





RL optimized group structure results





- Reusable CNN + LSTM based surrogate model for neutron transport.
- 1000+ days of RL training reduced to < 1 day training + 16 day data collection with surrogate
- RL optimized group structure outperforms commonly used ones

Key Finding: From speedup to capability

- Speed is means to an end; the goal is increased capability.
  - Optimization, UQ, inference previously infeasible become possible.
- Lessons:
  - Surrogates must be uncertainty aware
  - Hybrid workflows are often a good middle ground for preserving physics and making tiny steps
  - Physics still leads
  - Compatible with the ND culture of verification, validation and traceability.

ajeeta@lanl.gov

Future direction: From one-off success to sustained capability

- How to go from one-off models to reusable infrastructure
  - For e.g., deploy CoH_3 surrogate model as a shared tool or framework that can be used across projects.
- From just speed metrics to decision-quality metrics
  - Metric for success cannot just be one thing. Accuracy, precision, defensibility etc. need to be factored.
- Ad-Hoc data to managed data generation
  - Training surrogates requires intentional, funded use of HPC or computational resources to generate high-value simulation data. This needs to be planned, not incidental.
- Novelty to Trust
  - UQ, validation protocols, and reproducibility build confidence, and investment need to be in these areas

The need

- Even small targeted investments in surrogate-embedded workflows have been shown to be effective in unlocking disproportionate gains in capability
- Methods fail when ML and nuclear physics are treated as siloed. Surrogates are not generic tools – their validity depends on physics assumptions, parameterization and uncertainty interpretation
- In practice, the people building and using these must be trained in both the nuclear data problem and the ML methods well enough to understand their failure modes