Los Alamos
NATIONAL LABORATORY

Opportunities for Machine Learning and Artificial Intelligence

Scott Vander Wiel

WANDA 2025
February 10, 2025

LA-UR-25-21135

Summary

ML methods can assist evaluators and experimentalists to

- root out potential biases in evaluated ND
- reduce impact of systematic errors
- avoid understated uncertainties
- integrate analyses across diverse datasets
- optimize experiments (e.g., materials, geometry, measurements)

AI methods are exploding.

It’s bleeding-edge, but disruptive computing science is impressive and could dramatically accelerate analysis of diverse and complex datasets.

Projects and Teams

^{252}Cf PFNS
Small

AIACHNE **AI / ML Informed Californium Chi Nuclear Data Experiment**

D Brown[1], A. Carlson[2], M. Grosskopf[3], R. Haight[3], K. Kelly[3],
D. Neudecker[3], B. Pritychenko[1], S. Vander Wiel[3], N. Walton[3,4]

DOE Office of Science

[1]BNL, [2]NIST, [3]LANL, [4]UTK

^{239}Pu fast
Medium

EUCLID **Experiments Underpinned by Computational Learning for Improvements in Nuclear Data**

J. Alwin, B. Bell, A. Clark, T. Cutler, M. Grosskopf, W. Haeck, M. Herman,
J. Hutchinson, N. Kleedtke, J. Lamproe, R.C. Little, I. Michaud, D. Neudecker,
M. Rising, T. Smith, N. Thompson, S. Vander Wiel, N. Wynne

LANL

Pu,U,Cu,Al,B,Be,Cr…
Large & Diverse

PARADIGM **Shifting the Nuclear Data Evaluation Paradigm: Parallel Approach of Differential and Integral Measurements**

K. Amundson, B. Bell, P. Brain, T. Cutler, F. Diaby, M. Devlin,
N. Gibson, M. Grosskopf, J. Hutchinson, T. Kawano, F. Kazuki, A. Khatiwada,
N. Kleedtke, E. Leal Cidoncha, B. Little, A.E. Lovell, A. McHugh,
D. Neudecker, A. Stamatopoulos, C. Thompson, S. Vander Wiel, N. Walton

LANL

Next-Generation Computational Methods

Machine Learning (ML) methods can target these goals:

1. Integrate diverse datasets to produce ND with higher-quality uncertainties
2. Optimize new experiments to reduce ND uncertainties & application bounds
3. **Speed up the pipeline from experiment to evaluation**

Outline

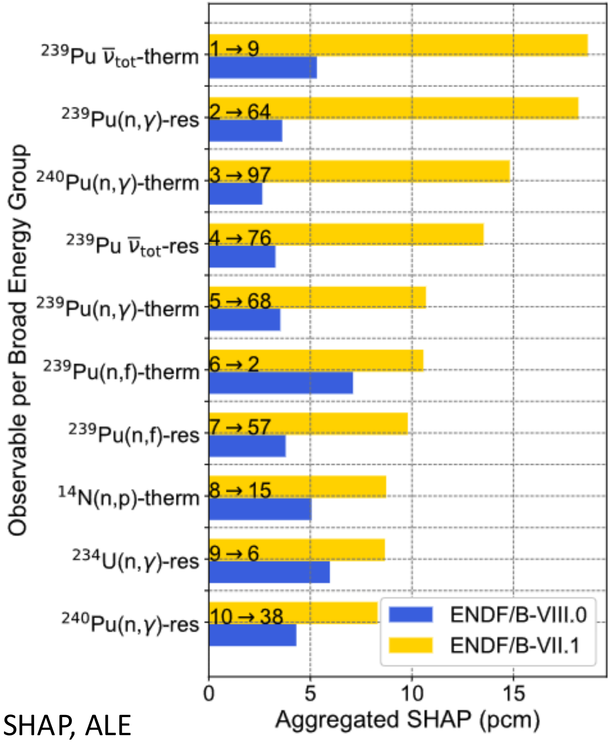
- Illustrate ML work applied to these goals
- Suggest how Artificial Intelligence (AI) could **further accelerate the pipeline**

A) ML Applied to Root Out Biases That Are Difficult for Experts to Find

- **ML regression methods** explain biases over a suite of criticality benchmarks:
  - mismatch b/w simulated and observed k_{eff}
  - fit to thousands of ND sensitivities
- **ML interpretability methods** quantify which ND quantities best explain the biases
  - Found some known issues & other new ones

ML methods assist ND experts to identify signal in large messy datasets. Keep experts in the loop!

ENDF/B-VII.1 top 10 bias candidates vs. VIII.0

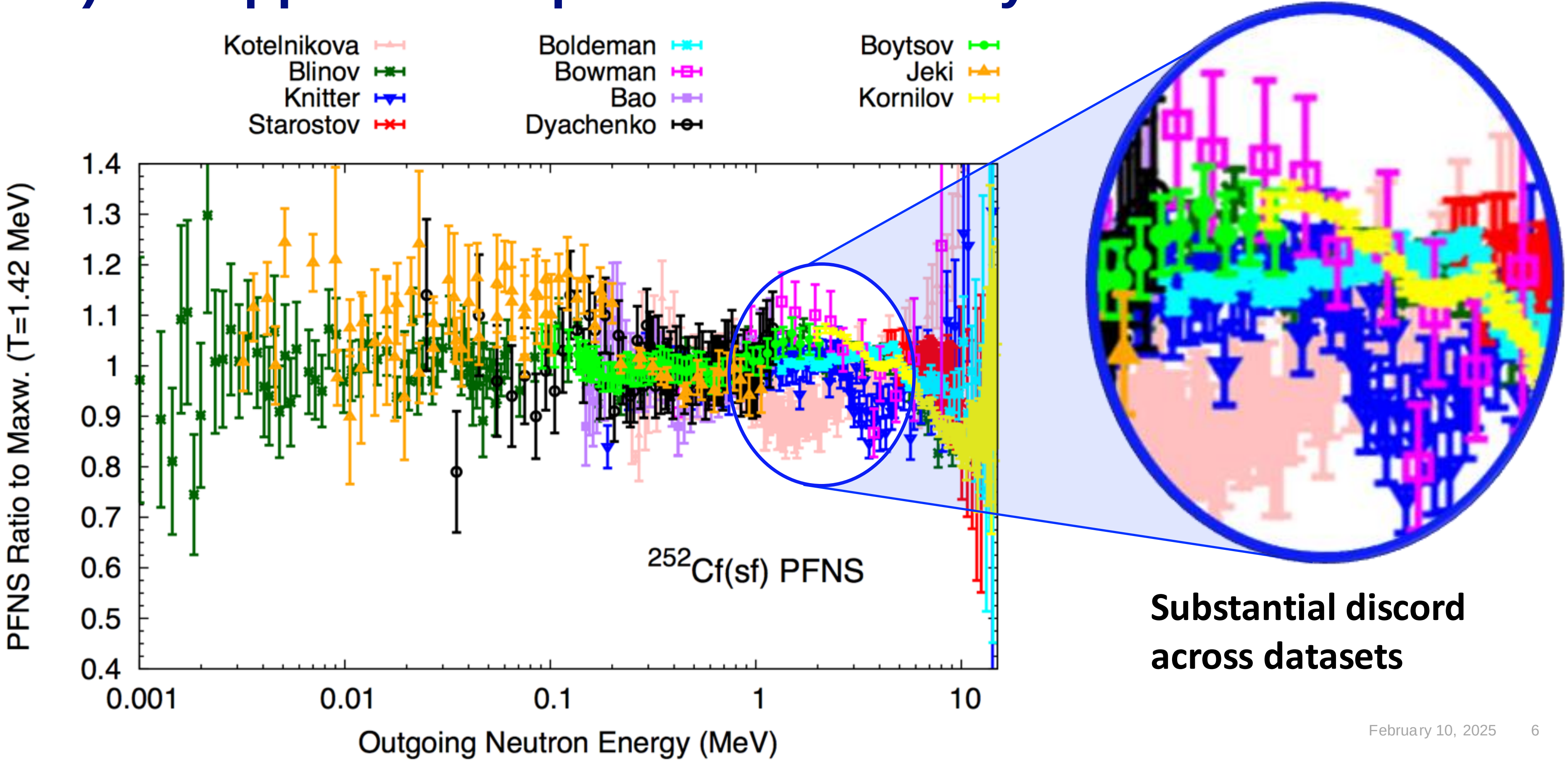


ML Methods: random forest, neural net, elastic net, support vector machine, SHAP, ALE

Neudecker, Denise, et al. "Enhancing nuclear data validation analysis by using machine learning." *Nuclear Data Sheets* 167 (2020): 36-60.

Neudecker, Denise, et al. "Informing nuclear physics via machine learning methods with differential and integral experiments." *Physical Review C* 104.3 (2021): 034611.

B) ML Applied to Capture Uncertain Systematic Errors

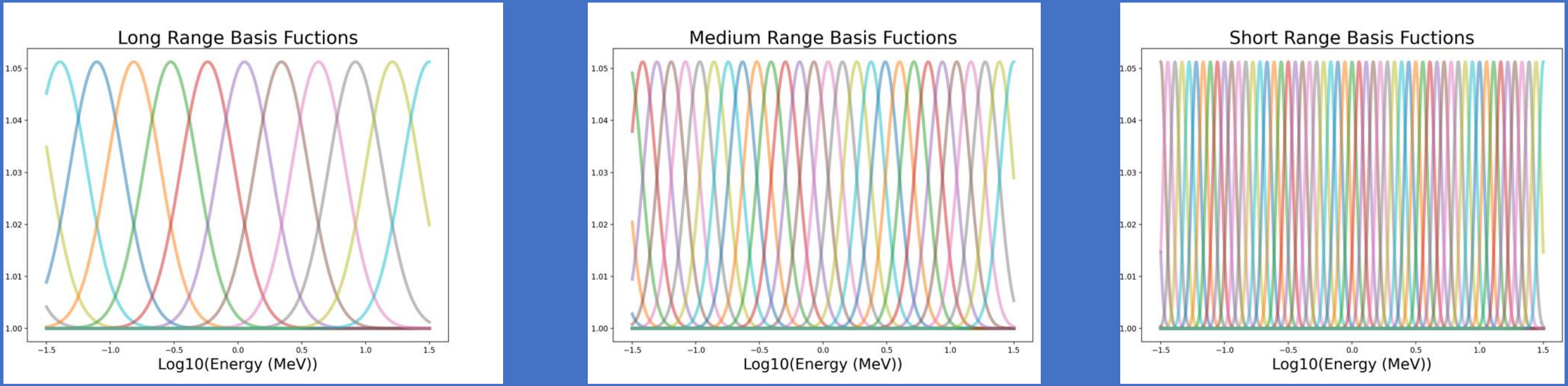


Substantial discord across datasets

B) ML Applied to Capture Uncertain Systematic Errors

ML sparsity methods can pick out localized biases

Large Library of Potential Biases



ML Methods: sparsity, global/local shrinkage, horseshoe priors

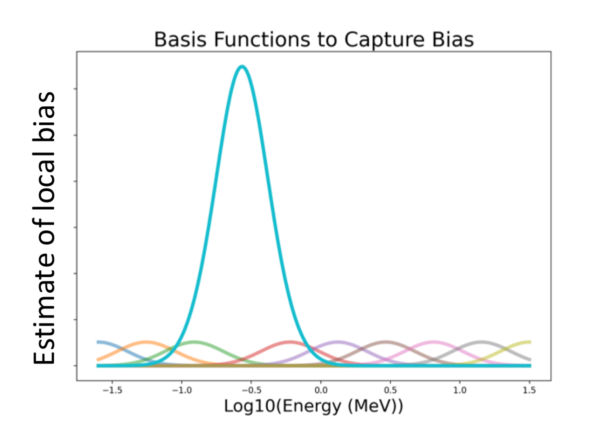
B) ML Applied to Capture Uncertain Systematic Errors

ML sparsity methods can pick out localized biases

Can identify *potential* root causes from experiment features

- E.g. experiments with ^{6}Li detectors are biased high at 100 to 300 keV (a known issue)

Horseshoe Induces Sparsity

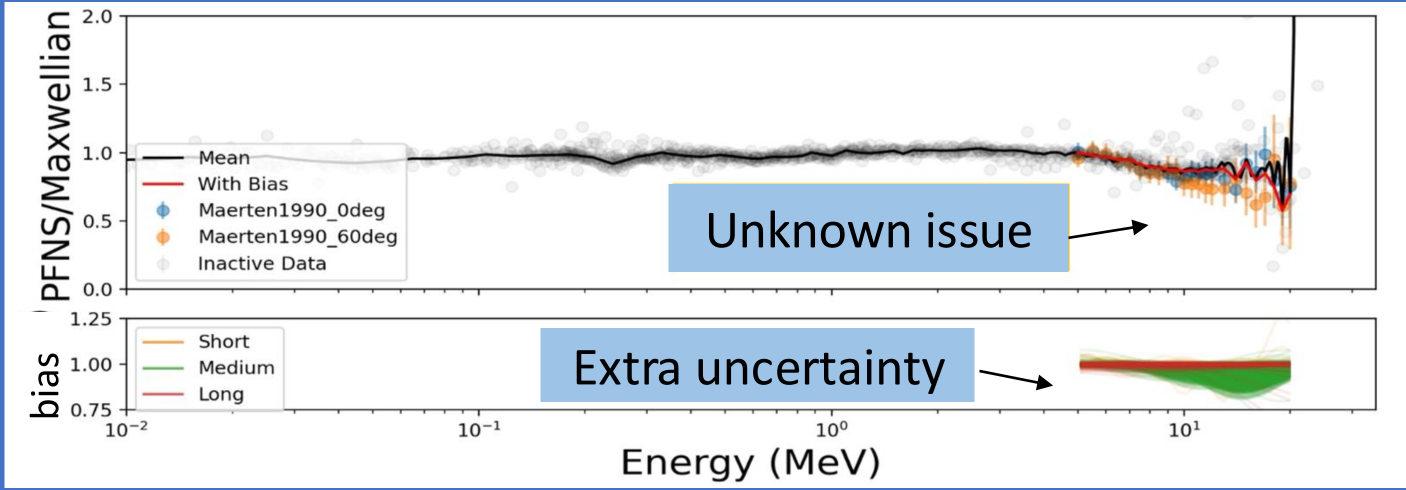


B) ML Applied to Capture Uncertain Systematic Errors

ML sparsity methods in action

- Identifies disagreement above 10 MeV and increases uncertainty there
- Can scale to much larger cases (e.g. *PARADIGM*)

^{252}Cf(sf) PFNS



ML methods capture additional uncertainty in the face of unknown systematic errors

C) Challenges with Integrating Diverse Datasets

- 20 candidate differential experiments

Candidate experiment selection

LANSCE Los Alamos Neutron Science Center

NCERC

- 6 candidate integral experiments
- **Dim: 6 x ~12,200**

Nuclear Theory

- Sampled model curves for ^{63,65}Cu, ^{239,240}Pu
- **Dim: ~2,200 x 1,000 ND**

Historic Differential Experiments

- 122 differential data sets
- **Dim: 8,400 data points**

Historic Integral Experiments

- 46 k_{eff} benchmarks with sensitivity profiles
- **Dim: 46 x ~12,200**

Evaluation

- ENDF/B-VIII.0 mean values and covariance from ^{1}H, ^{9}Be, ^{10,11}B, ^{12}C, ^{16}O, ^{27}Al, ^{52}Cr, ^{56}Fe, ^{208}Pb, ^{235,238}U
- **Dim: 10,000 ND**

See Peter Brain's PARADIGM poster tomorrow

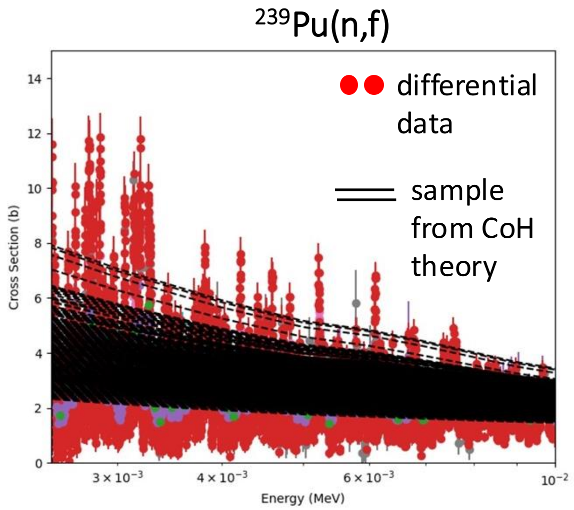
C) Challenges with Integrating Diverse Datasets

Processing challenges to wrangle the data

- Diverse energy specifications
- Large-scale computation of sensitivities

Analysis challenges for plausible uncertainties

- Discord (again)
  - No evaluation adequately fits all data
  - USU: unrecognized sources of uncertainty
- Missing physics
  - e.g., unresolved resonances in theory models
  - extensions underway
- The ML formulation must reconcile data tensions
  - to avoid understating combined uncertainties.



Physics Needs: Advance theory and model development

ML Needs: Scaling workflows to larger diverse collections

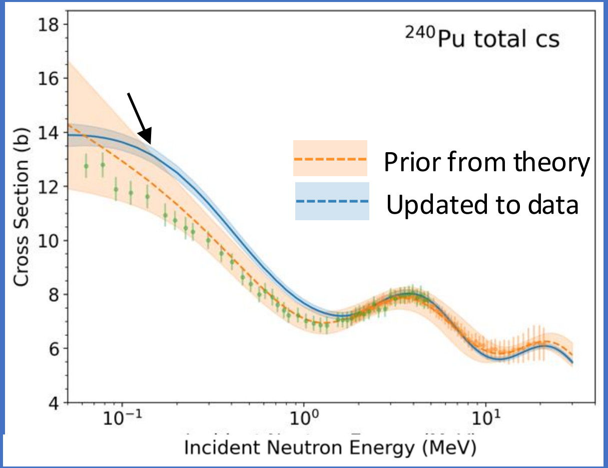
D) Statistical ML Applied to Produce Better Uncertainties

ML methods can relieve systematic tensions within and between datasets

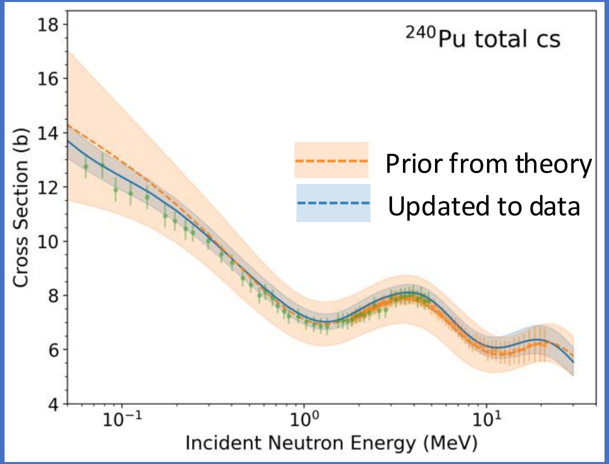
- E.g. shows benefit of adding Gaussian Processes uncertainties
- Combined uncertainties should be validated with statistical diagnostics

Arrow shows a wild side-effect of overly-tight correlations claimed in a ^{240}Pu(n,f) dataset

Standard Analysis (GLS)



Added GP Uncertainty



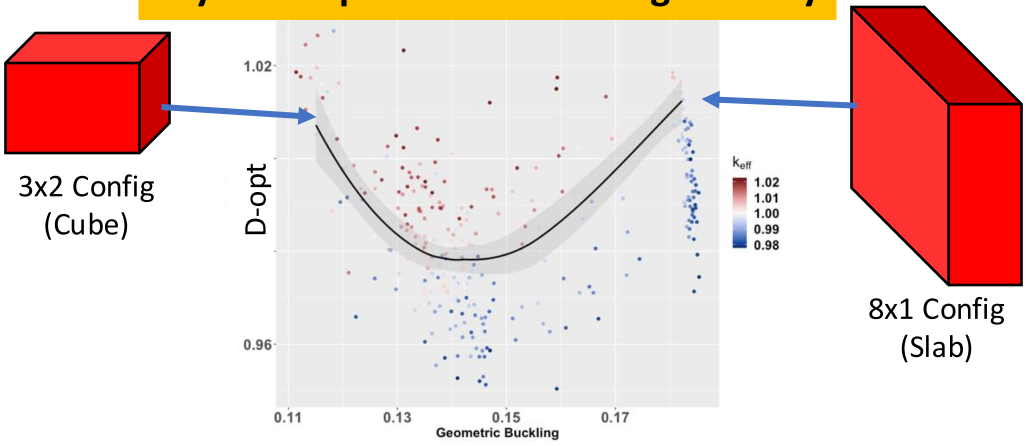
Need: automated methods to find and loosen overly-constrained data uncertainties where needed

E) ML Applied to Optimize Experiments

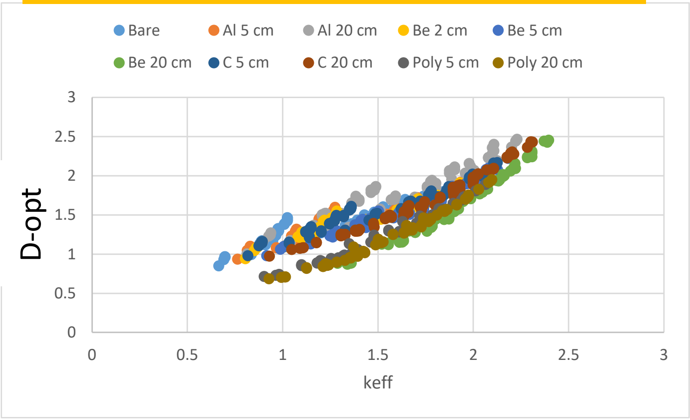
ML optimization methods search and **optimize proposed experiments** to best reduce uncertainties and compensating errors

D-opt measures uncertainty volume for correlated ND. It is targetable to subsets of ND (e.g. fast energy range) and to **specific applications suites**

Bayesian optimization of Pu geometry



D-opt down-select to Al reflector



Michaud, Isaac, et al. "Expert-in-the-loop design of integral nuclear data experiments." Statistical Analysis and Data Mining: The ASA Data Science Journal 17.2 (2024): e11677.

ML Methods: Bayesian optimization, genetic algorithms, D-opt metric

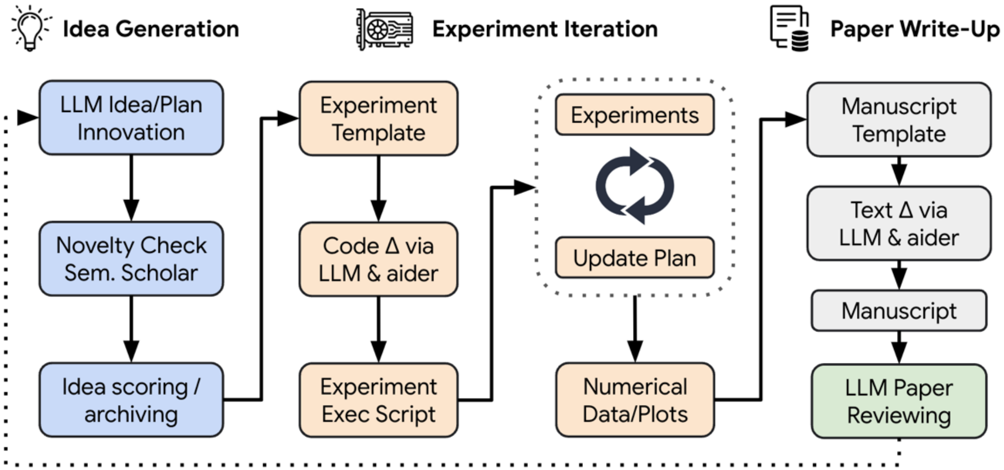
New *AI Scientists* automate model building and comparison, vastly accelerating scientific throughput

Large Language Model are leveraged to do compute-based scientific research:

- synthesize literature,
- propose experiments,
- write code and execute
- refine, and
- write a paper

The field is screaming fast!

- **Sakana AI (see figure at right):** github.com/SakanaAI/AI-Scientist/tree/main
- **Agent Laboratory:** github.com/SamuelSchmidgall/AgentLaboratory/tree/main
- **Aviary:** arxiv.org/abs/2412.21154

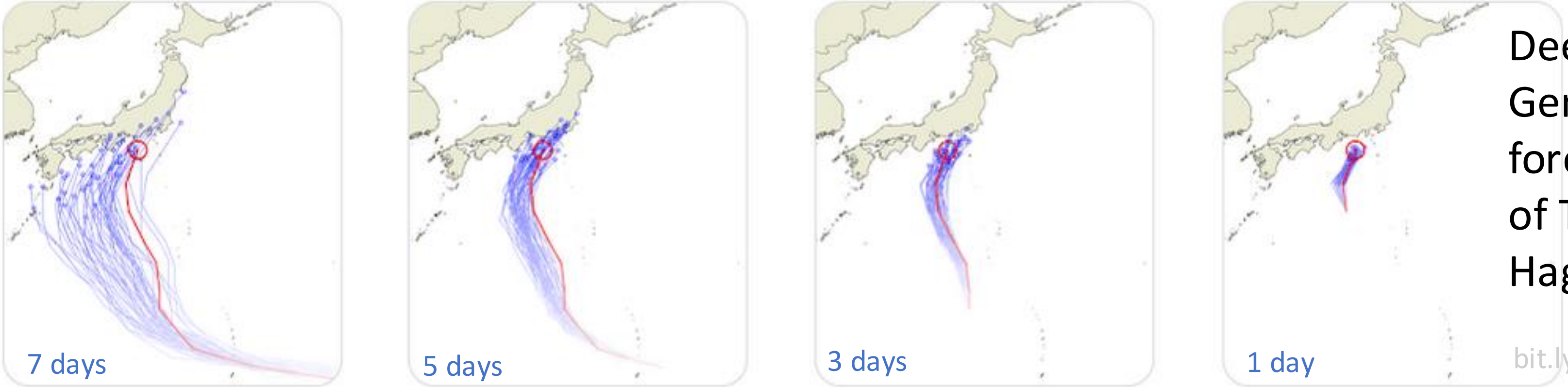


Autonomous agents could assist ND evaluation by implementing multiple ML methods to fuse diverse datasets and recommend what works best to deliver vetted uncertainties.

Los Alamos NATIONAL LABORATORY

Needs: demonstration project; integration with AI researchers

AI methods underlying popular image & video generators (like DALL·E) are turning to hard science problems



DeepMind GenCast forecasts of Typhoon Hagibis

bit.ly/42HGz61

Generative diffusion methods could be trained on a sample of ND curves from theory and asked to create similar curves that match given differential data. These tools can ingest vast datasets and give fast answers.

Caution: current approaches do not explicitly handle physical constraints and correlated uncertainties. ND applications could push the field!

Diverse expertise is needed for development and validation

Summary

ML methods can assist evaluators and experimentalists to

- root out potential biases in evaluated ND
- translate discrepancies between datasets into appropriate uncertainties
- integrate analyses across diverse datasets
- optimize experiments (e.g., materials, geometry, measurements)

AI methods are exploding.

It’s bleeding-edge, but disruptive computing science is impressive and could dramatically accelerate analysis of diverse and complex datasets.

ML and AI are well-poised to continue to speed up and improve the ND pipeline.

Some needs to take advantage of ML/AI:

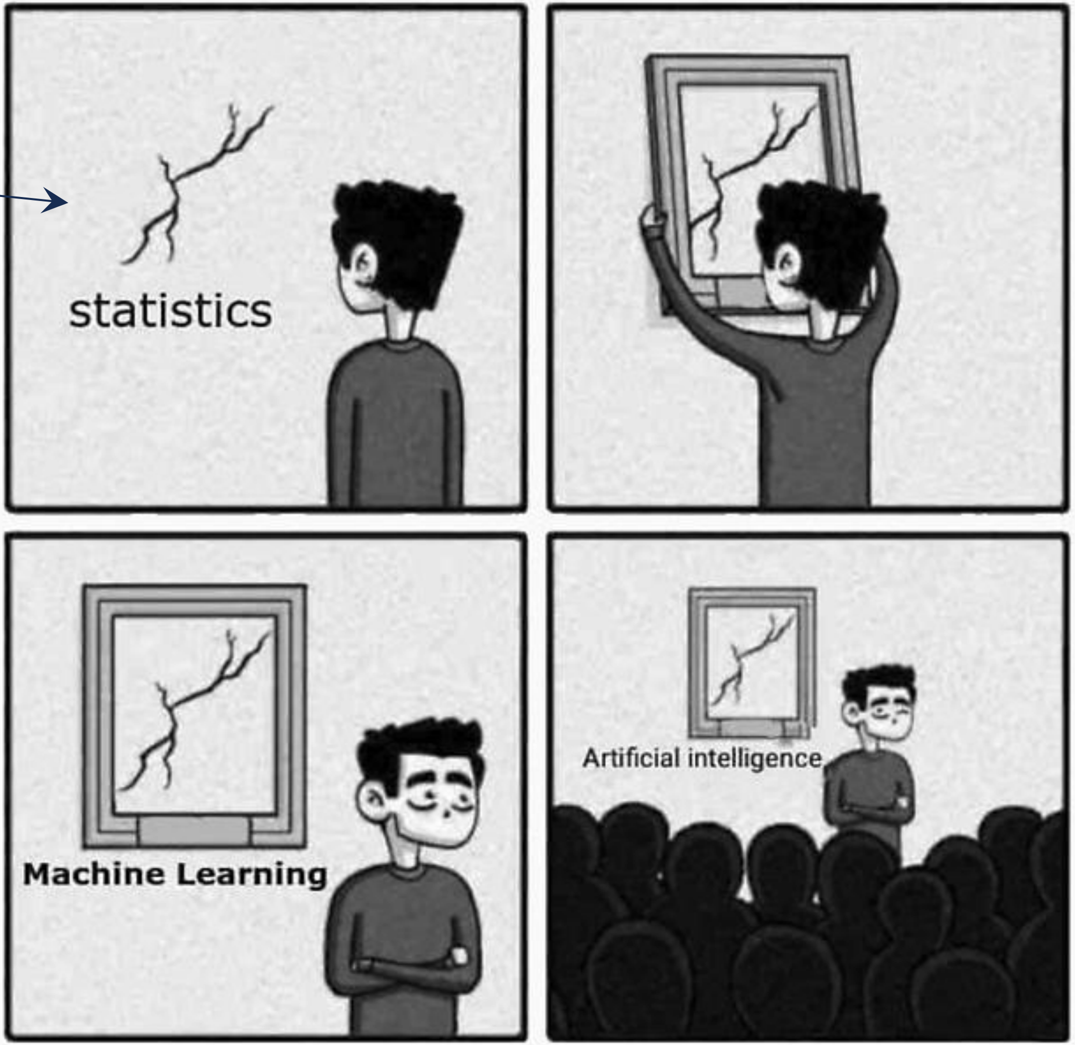
- Further collaborations with ML scientists
  - Tailor ML methods to ND analysis workflows
  - Require validation of uncertainties in large-scale analyses
- Infrastructure
  - Accessible data, codes, etc.
  - Accessible worked examples
  - Shared development
- For AI (FAST moving)
  - Integration with the AI researchers
  - Capable computing resources

The needs are cross-disciplinary.

Teams require
- **experimentalists**
- **evaluators**
- **physicists**
- **data science researchers**

Confession

I mostly live here



Organizations, facilities, libraries, codes

Acknowledgements

- Research reported in this publication was supported by the U.S. Department of Energy
  - LDRD program at Los Alamos National Laboratory and
  - Office of Nuclear Physics, under the Nuclear Data InterAgency Working Group Research Program.

- NCERC is supported by the DOE Nuclear Criticality Safety Program, funded and managed by the National Nuclear Security Administration for the Department of Energy.