ML Applications and Experiences at BNL hadron injector complex

Weijian (Lucy) Lin

wlin1@bnl.gov

Dedicated AI/ML Accelerator Test Facility

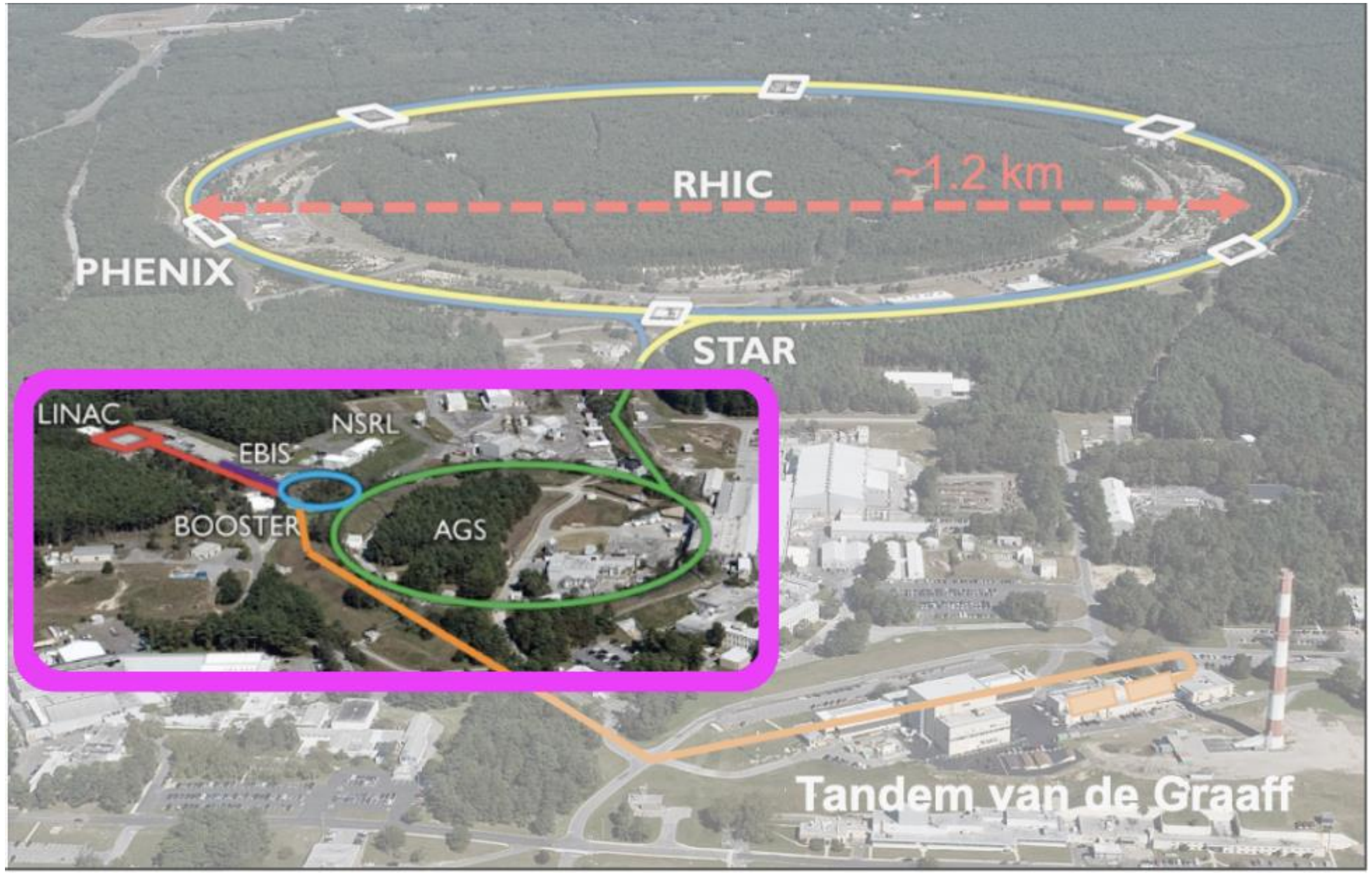
July 15-17, 2026

Hadron Injector Complex at BNL

Accelerator Rings	Circumference [m]
Booster	201
AGS	807
RHIC	3833

Typical Top Energies [Total, GeV/N]

	Au	Pol. Protons
Linac (H^-)	--	1.1
Booster	1	2.3
AGS	10	23.8
RHIC	100	255



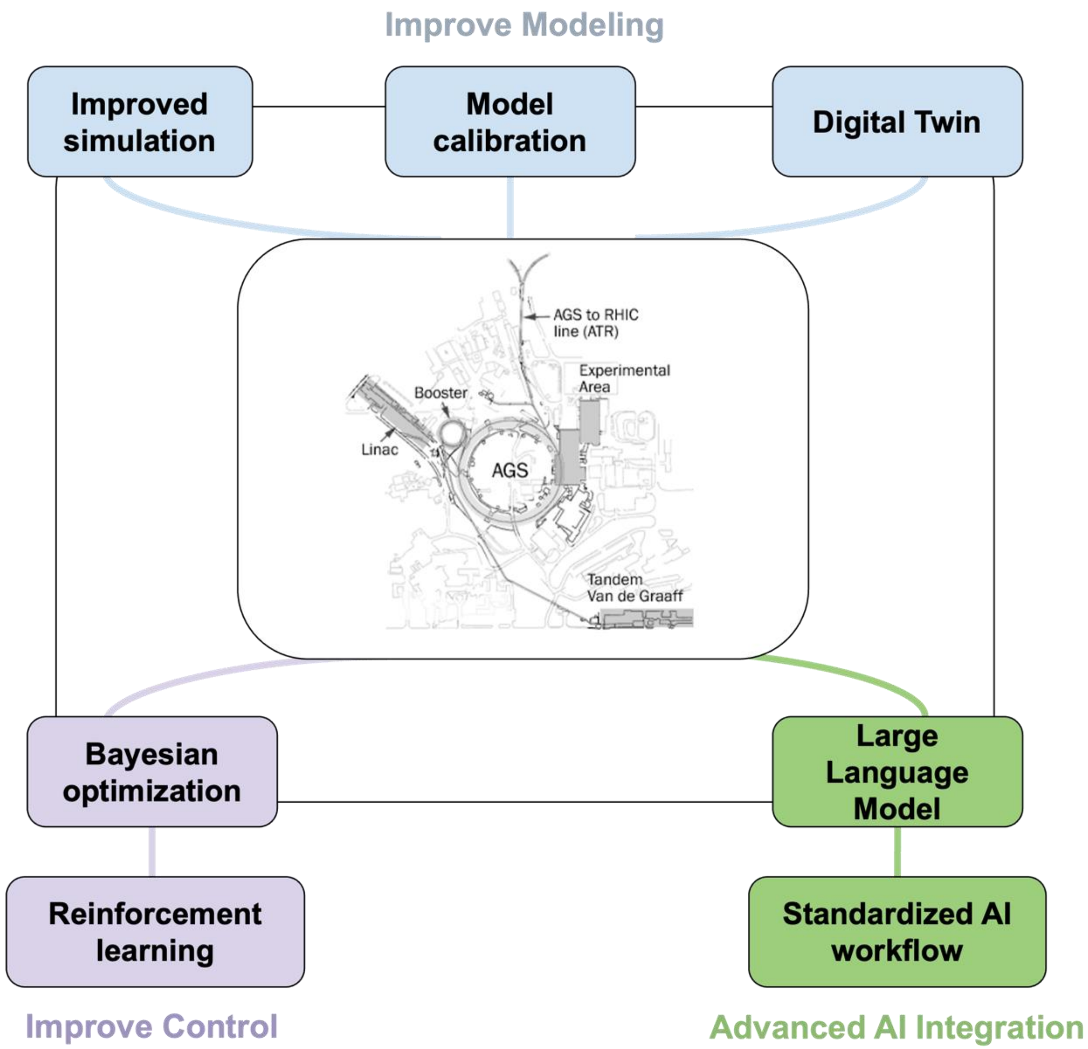
Machine Learning at BNL hadron injectors

Goal: Improve beam quality and operational efficiency.

This talk

- Automated beam tuning
- Operational Digital Twin
- Data-driven model calibration
- Fast adaptive control
- AI-ready data standard development

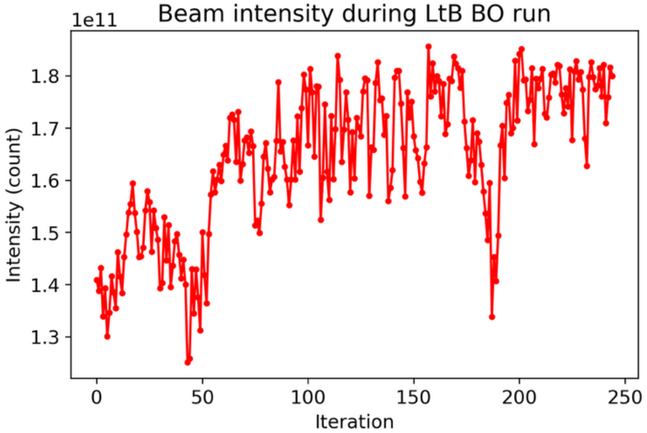
Better models + better data + better optimization = more autonomous accelerator operation

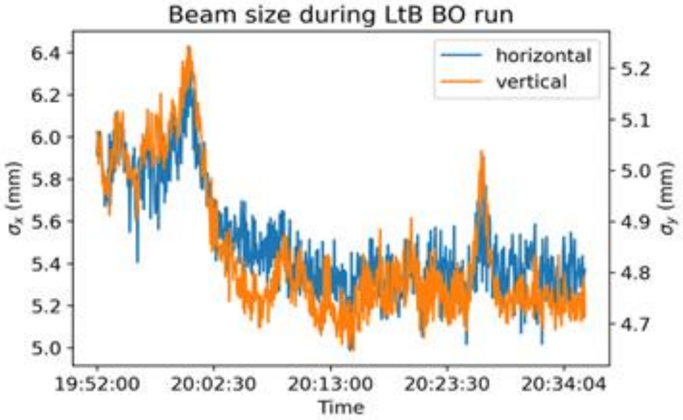


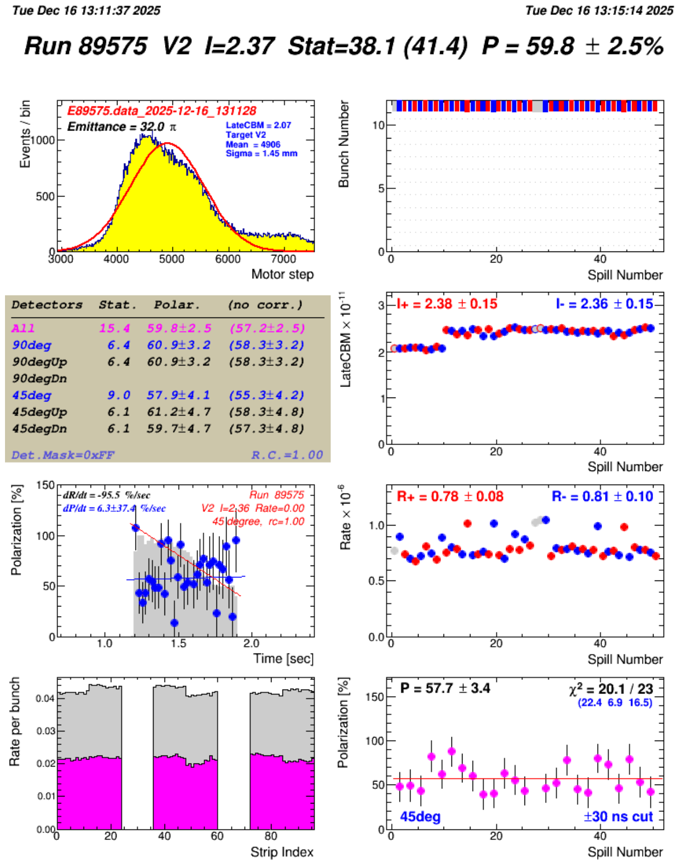
Injection Optimization with BO

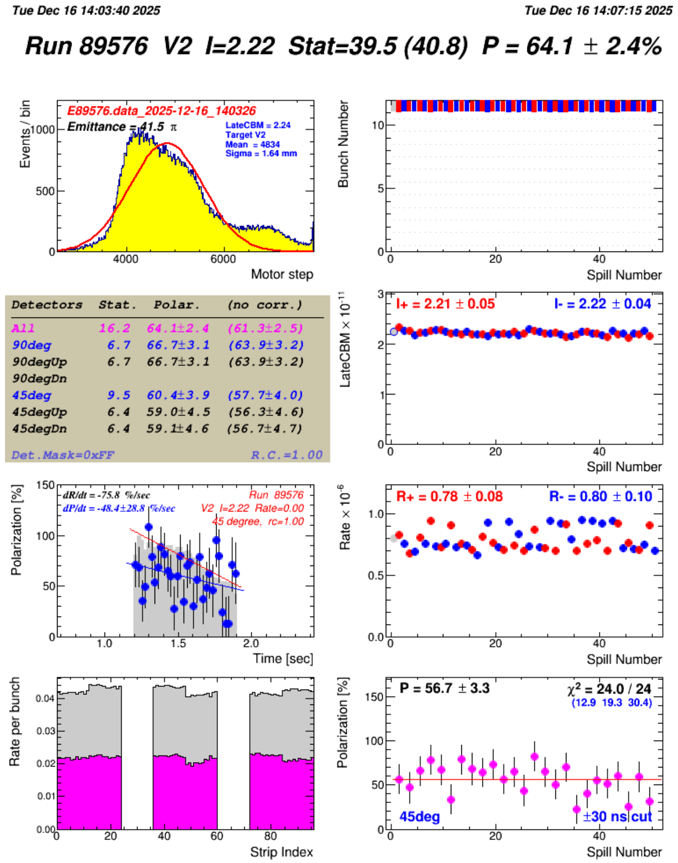
LtB Injection Optimization

- BO maximizes Booster beam intensity using LtB optics.
- Beam size decrease in both planes during LtB optimization.
- Polarization increase observed, but instrumentation limitation is significant.





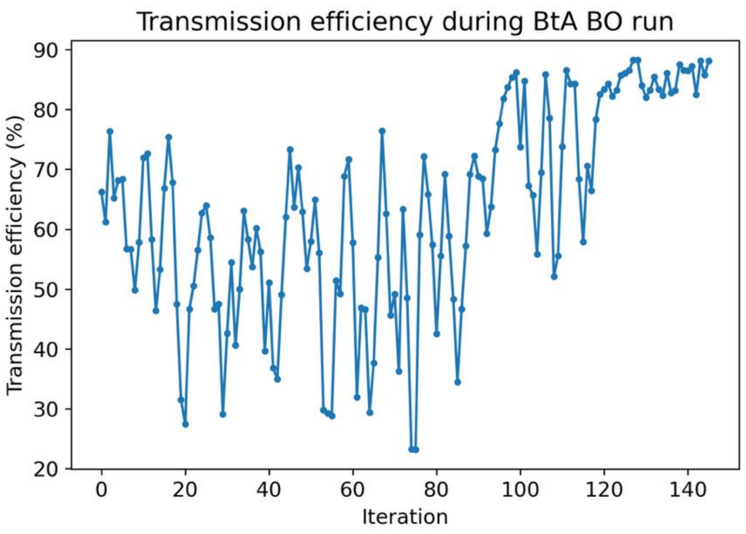


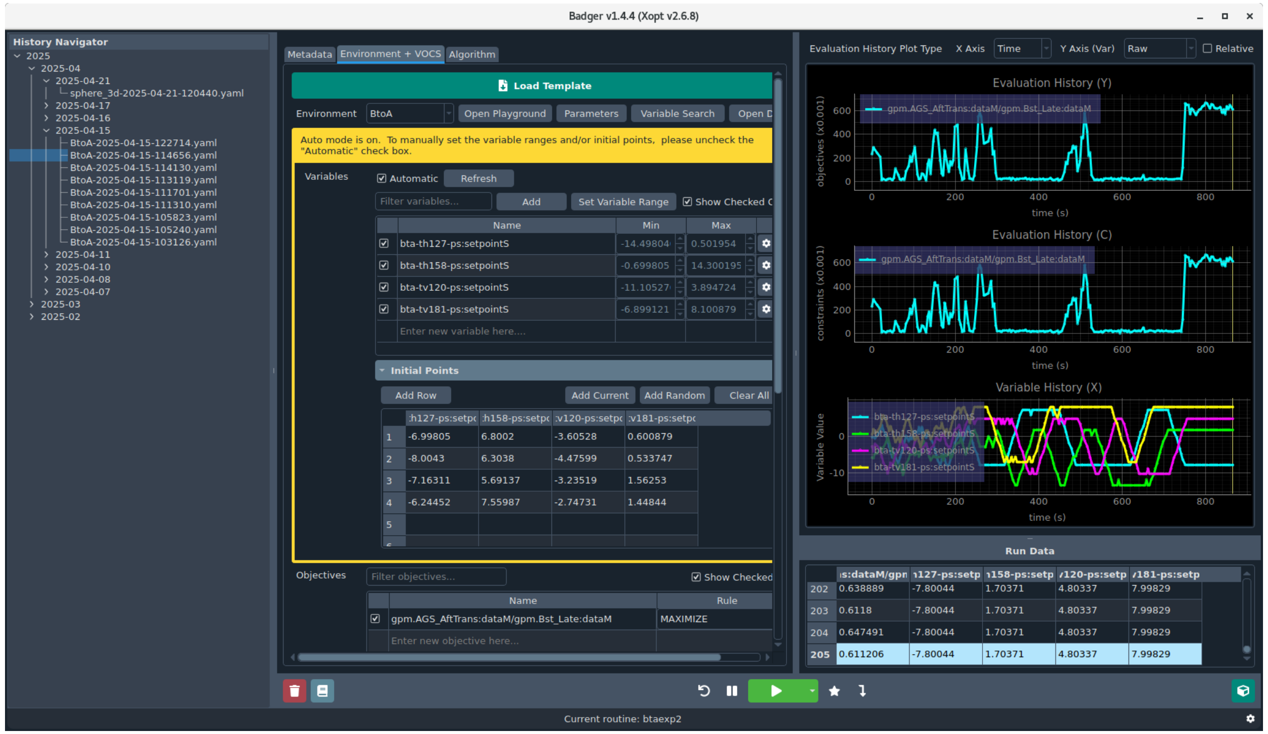


Injection Optimization with BO

BtA Injection Optimization

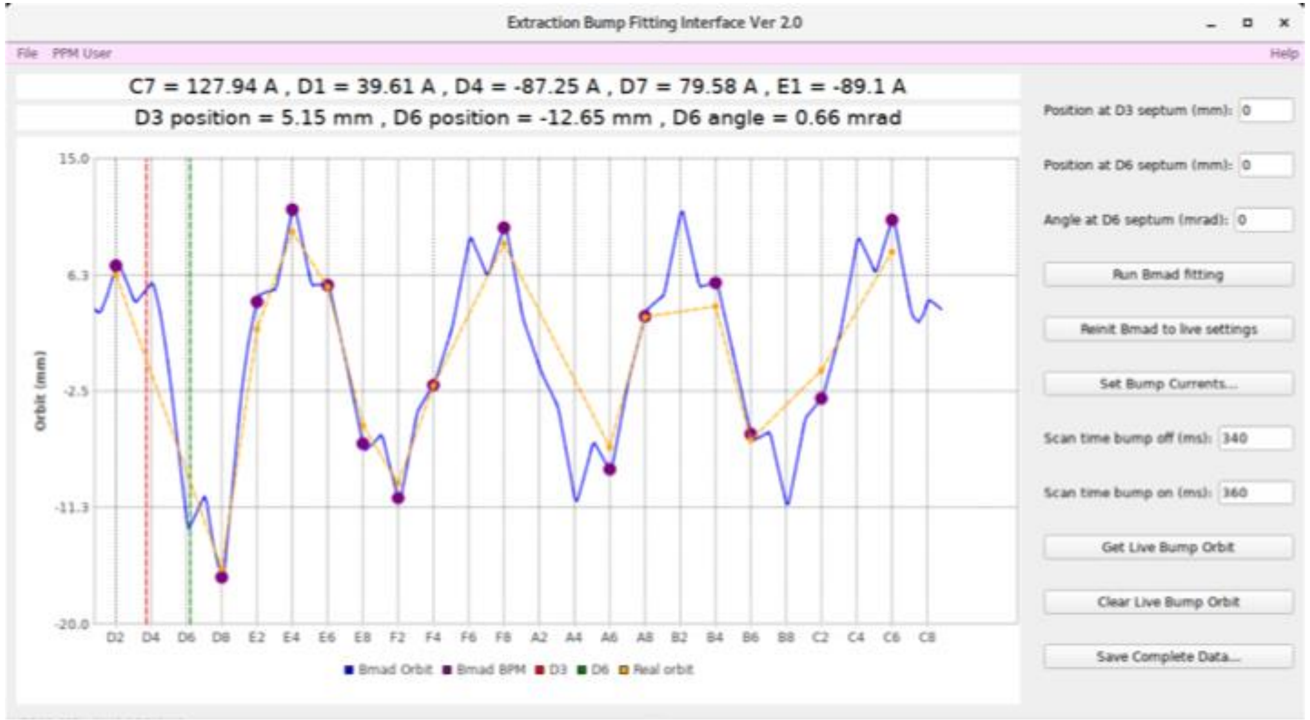
- BO maximizes beam transmission efficiency using BtA optics.
- 20% improvement from operational settings.
- Badger successfully deployed in the control room after customization for the BNL control system.
- Emittance measurement from IPM is not reliable, no working BPM in BtA line.

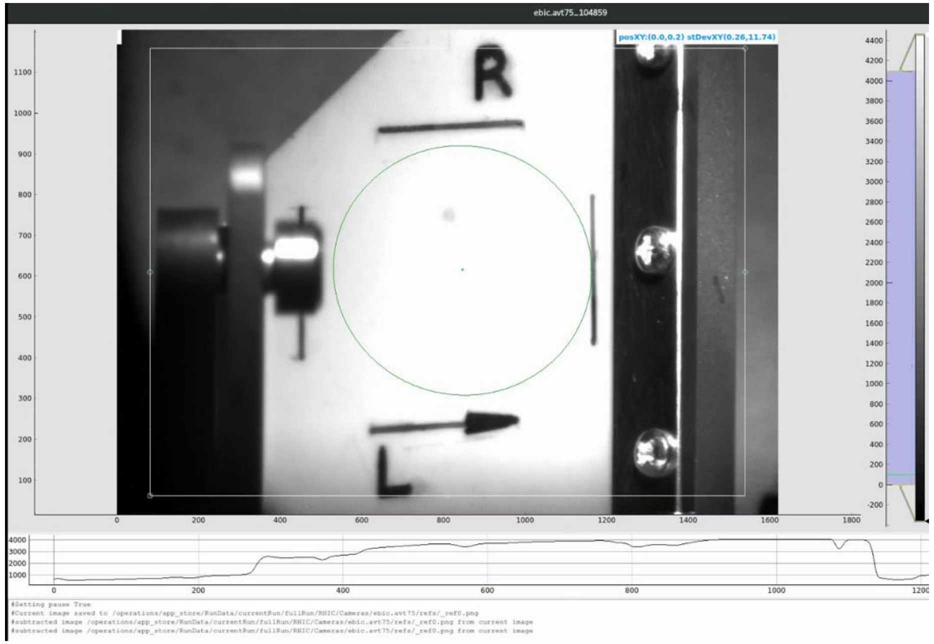




Digital Twin for Booster slow extraction bump

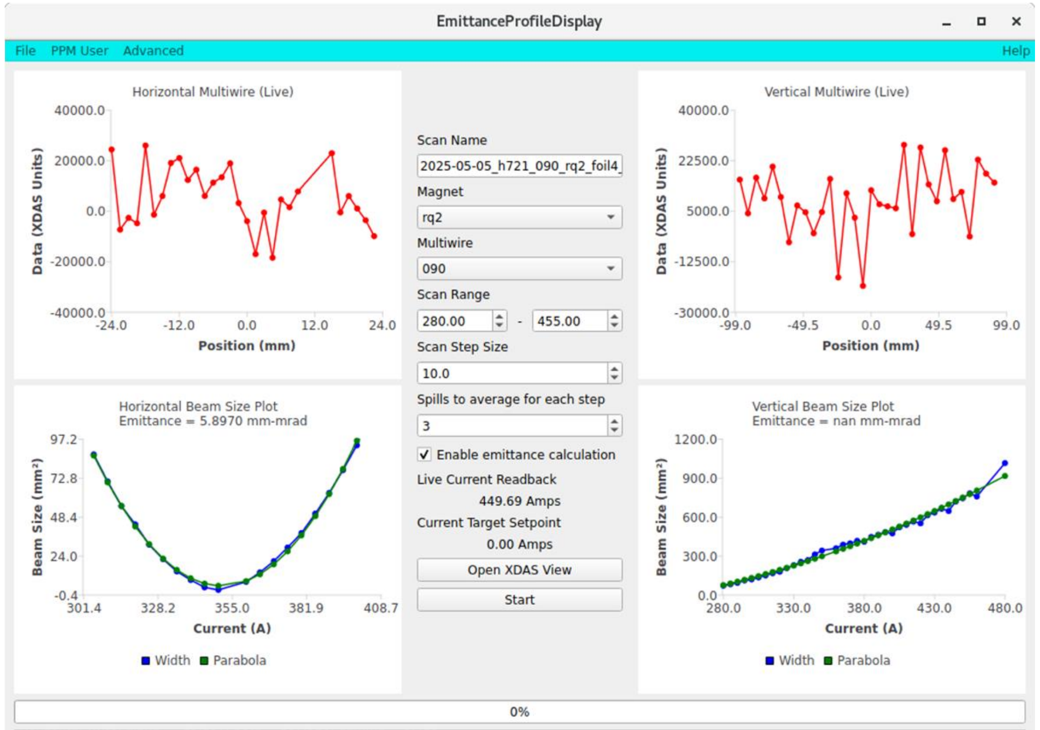
- Operational DT GUI with **bidirectional** interaction between the physical and virtual machines.
- Can expand to other scenarios (e.g. fast extraction bumps) with easier model-machine validation. BPM measurement unavailable during normal slow extraction process (RF off).
- Diagnostics between thin (D3) and thick (D6) septa are sparse, difficult to reconstruct design bump. Camera was installed at thick septum for additional info, but no image processing in place.





Digital Twin for NASA Space Radiation Laboratory

- Operational DT GUI running in digital shadow mode.
- Initial conditions after slow extraction hard to predict without dedicated measurements. Full slow extraction simulation is computationally prohibitive for online applications. Created automated quad scan GUI for initial parameter back-propagation, but horizontal is difficult due to dispersion.



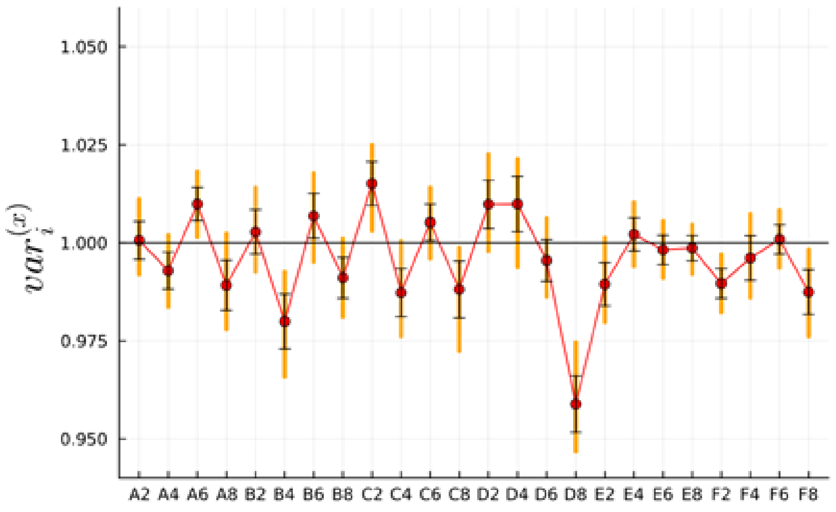
Uncertainty Quantification to Improve Booster Model

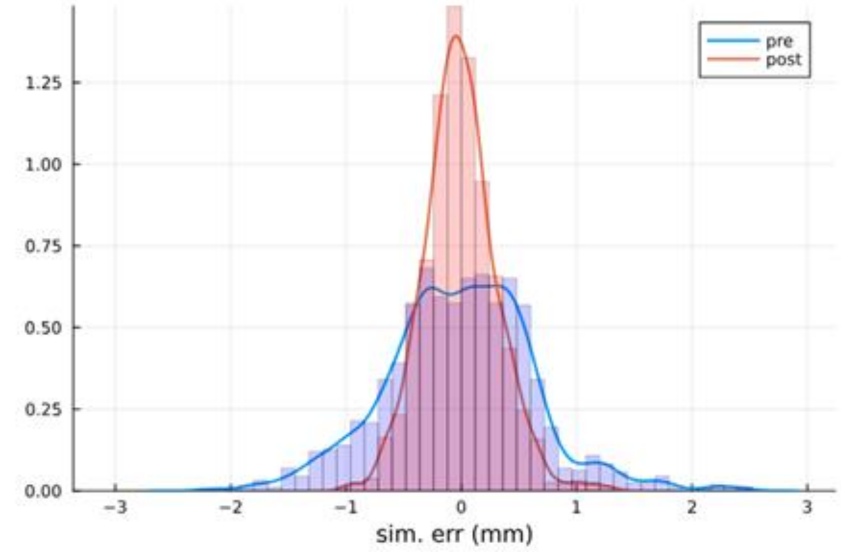
- Use Bayesian Uncertainty Quantification (UQ) to characterize and reduce model-data mismatches, assisting Digital Twin development.

- Bayesian UQ analyzes orbit response discrepancies.
- Neural network surrogate model accelerates UQ process.

- Bayesian UQ quantifies quadrupole field errors contributing to model discrepancies.
- Incorporating UQ results improves agreement between simulation and measurement.

- Anticipated turn-by-turn BPM upgrade will significantly improve the available data for UQ.

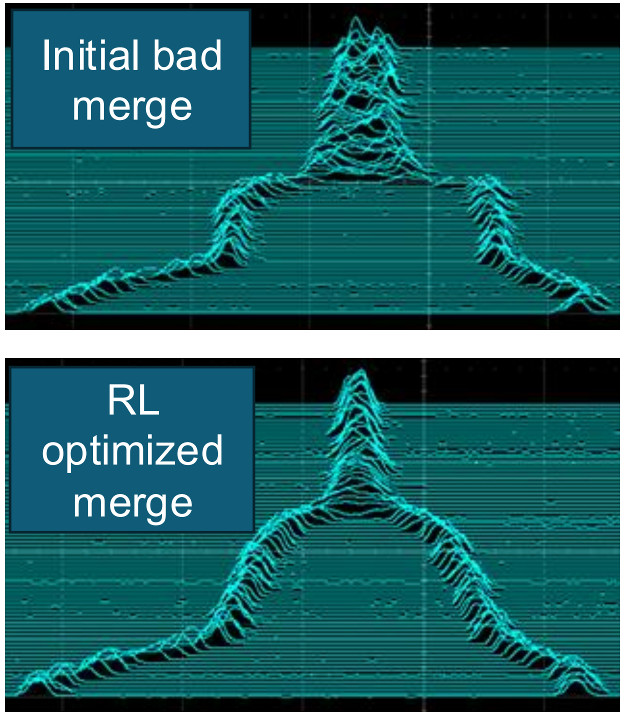




AGS Bunch Merging Optimization with RL

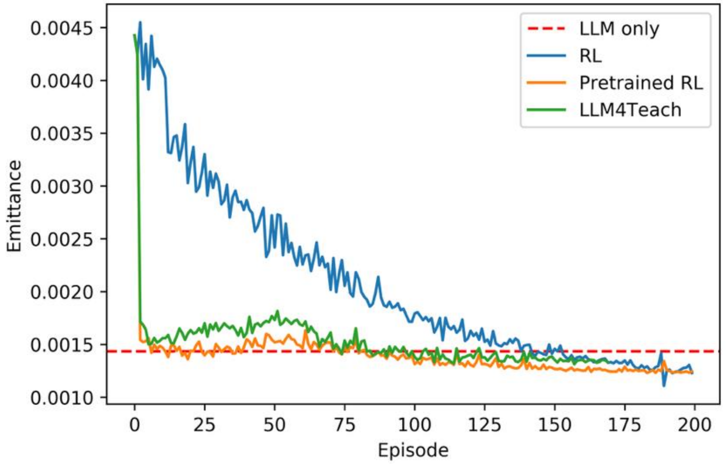
- Obtain a good merged bunch profile – low emittance growth, no loss, centered and stable.

- Longitudinal merge data had to be digitized and logged; operators tune by looking at remote scope signals.



RL Optimization

- RL agent trained to minimize emittance growth.
- Begins from poor merge conditions; rapidly converges to good RF settings.
- Comparable performance to operational settings.
- LLM-based policy teacher for RL agent in development, leads to faster convergence and better performance.

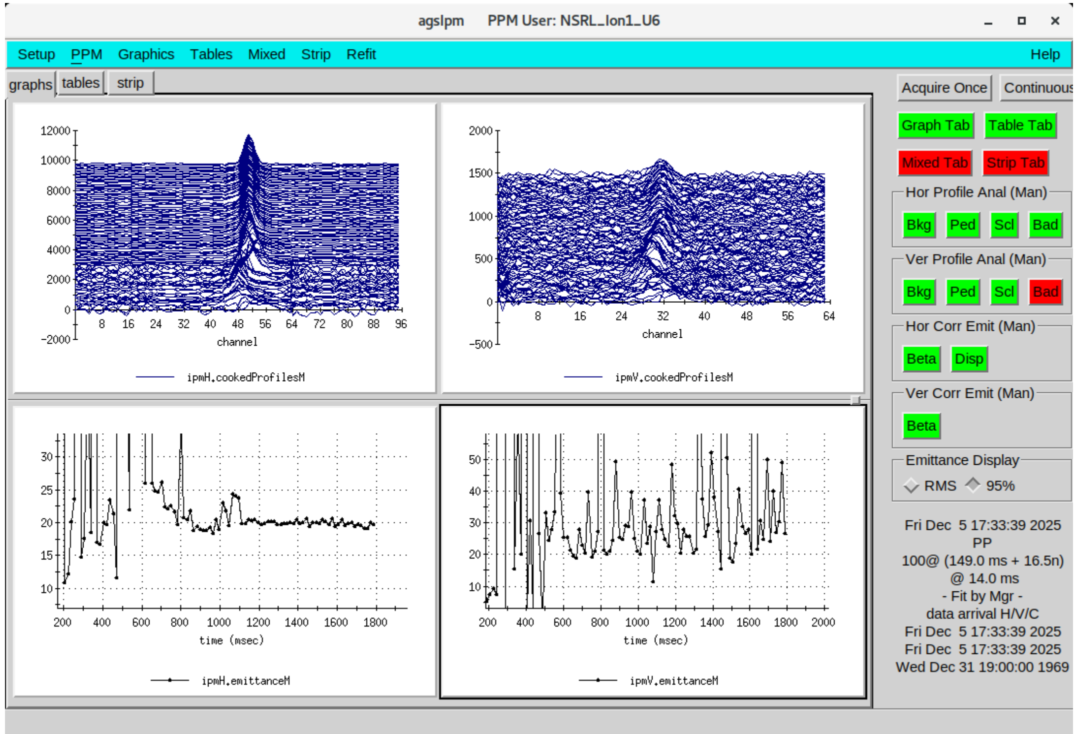


Improving AGS IPM Emittance Measurement with ML

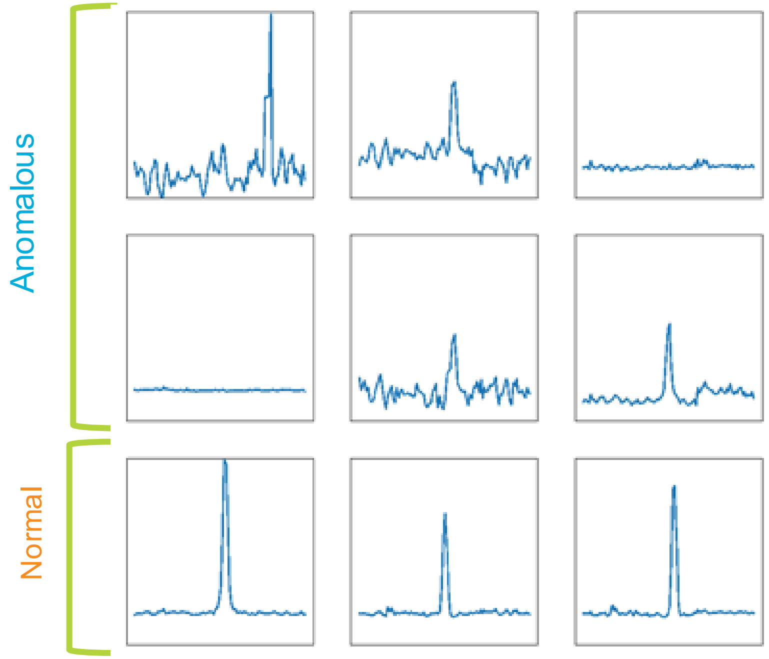
Goal: Improve emittance reconstruction accuracy and robustness of AGS Ion Profile Monitor.

Approach: Gaussian process based surrogate model for fast IPM profile inference; unsupervised anomaly classifier to detect atypical IPM profiles.

Challenge: Limited availability of high-quality training data from both IPM and eIPM.



Classifier predictions on archive data



NP AI-Ready Accelerator Data (NARAD)

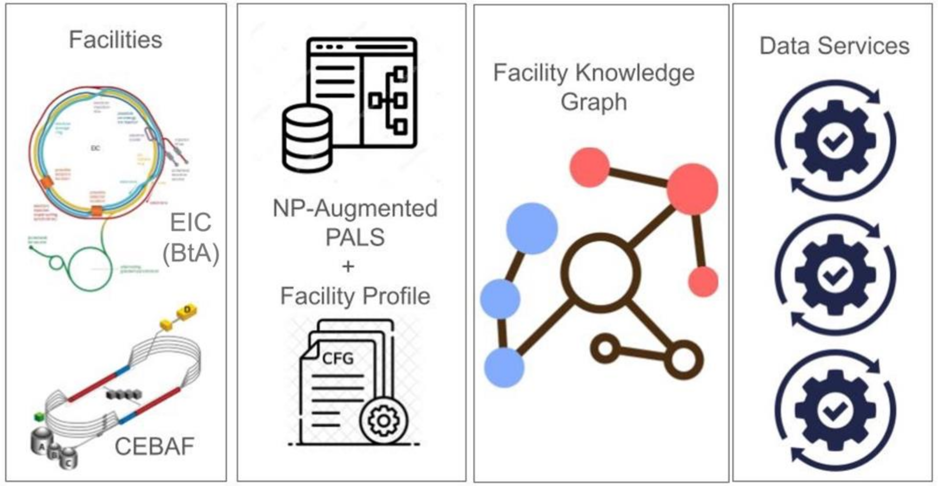
Goal: Develop FAIR, context-rich, AI-ready accelerator data and common ontology for cross-facility AI applications (BNL, JLab, LBNL, PNNL).

BNL BtA Experience

- Legacy control systems are difficult to generalize: ADO vs. EPICS
- Metadata / Context can be sparse or implicit: lattice configuration, device hierarchy, calibration, units, machine state etc.
- Data correlation is nontrivial: asynchronous data streams (magnet settings, beam profiles, timestamps)

Standard Development Effort

- PALS + LinkML + knowledge graphs: machine-readable accelerator semantics
- Genesis Mission data card: toward community standards for AI-ready data curation, preprocessing, and quality assessment



Agentic AI for Accelerator Operations

Current capabilities

- Domain-specific RAG using GPT-OSS-120B over 300-400k documents (HTML, PDF, Wiki, etc.)
- AI assistants for control system queries, database search, eLog analysis, and troubleshooting

Remaining challenges

- Natural language query for ~2.7M ADO parameters, structured context from scattered documents
- Secure agent interaction with control systems behind the firewall, local vs. commercial LLM
- Building operator trust and encouraging routine use

ML Experience at BNL: Takeaways

- **AI works:** Bayesian Optimization, Reinforcement Learning, surrogate models, and Digital Twins already provide measurable operational benefits.
- **Diagnostics are the bottleneck:** Reliable measurements, metadata, and standardized data infrastructure are essential for scalable AI.
- **Physics + AI + Data Infrastructure:** Future accelerator control will integrate first-principles models, machine learning, interoperable data standards, and intelligent agents.

Looking forward

- Invest in next-generation diagnostics
- Develop trustworthy Digital Twins with uncertainty quantification
- Build standardized AI-ready data infrastructure through NARAD
- Expand online Bayesian Optimization and Reinforcement Learning
- Human-in-the-loop accelerator automation

Acknowledgements: A Community Effort

Brookhaven National Laboratory

CLASSE Cornell Laboratory for Accelerator-based ScienceS & Education

Jefferson Lab Thomas Jefferson National Accelerator Facility

Fermilab

SLAC National Accelerator Laboratory

Rensselaer

radiasoft

Berkeley Lab

Pacific Northwest National Laboratory

EIC-Beam AI Working Group now organizes monthly meetings: https://indico.global/category/1612/