Applications of Neural Networks for Anomaly Detection in Particle Accelerators

Jonathan Edelen, Kevin Bruhwiler, Nathan Cook, Christopher Hall, and Stephen Webb

Kevin Brown, and Vincent Schoefer

*jedelen@radiasoft.net

November 2nd 2021

AI for Particle Accelerators, X-ray Beamlines, and Electron Microscopy Workshop @ ANL

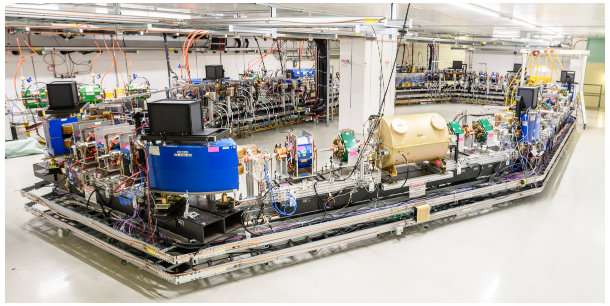
This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics Award Number DE-SC0019682 and Office of High Energy Physics Award Number DE-SC0019979

Overview of accelerator operations

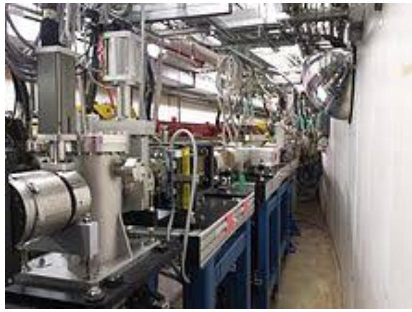
Accelerator R+D

Machine Development Time

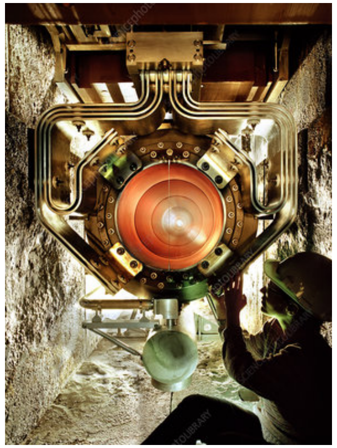
Specialized R+D Facilities



Beam for Experimentalists



Small single user end stations

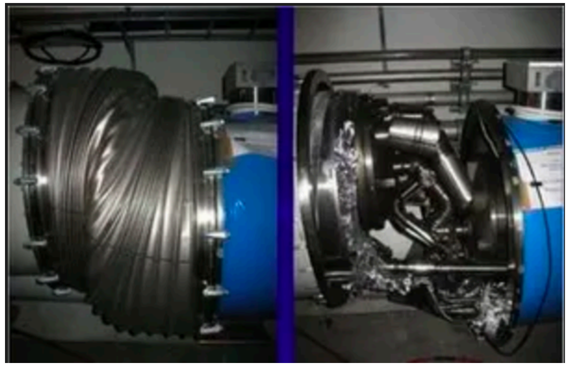


Large experimental collaborations

Down Time

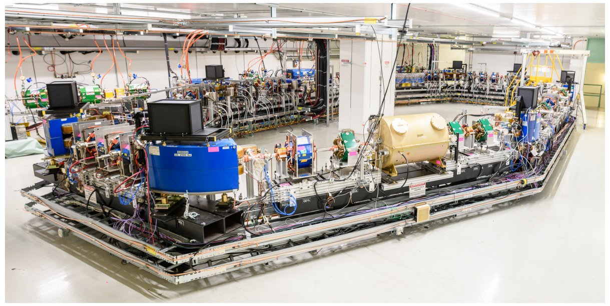
Scheduled Maintenance

Unscheduled Maintenance

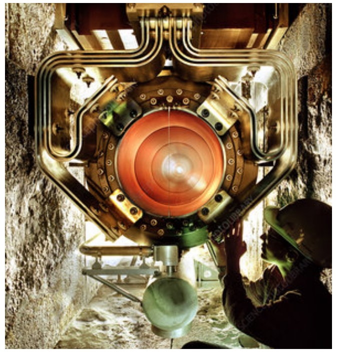


Machine learning applications for anomaly detection

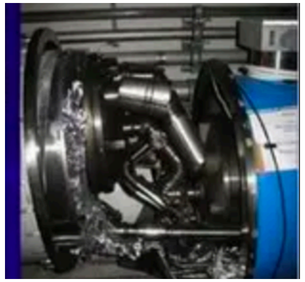
Accelerator R+D



Beam for Experimentalists



Down Time



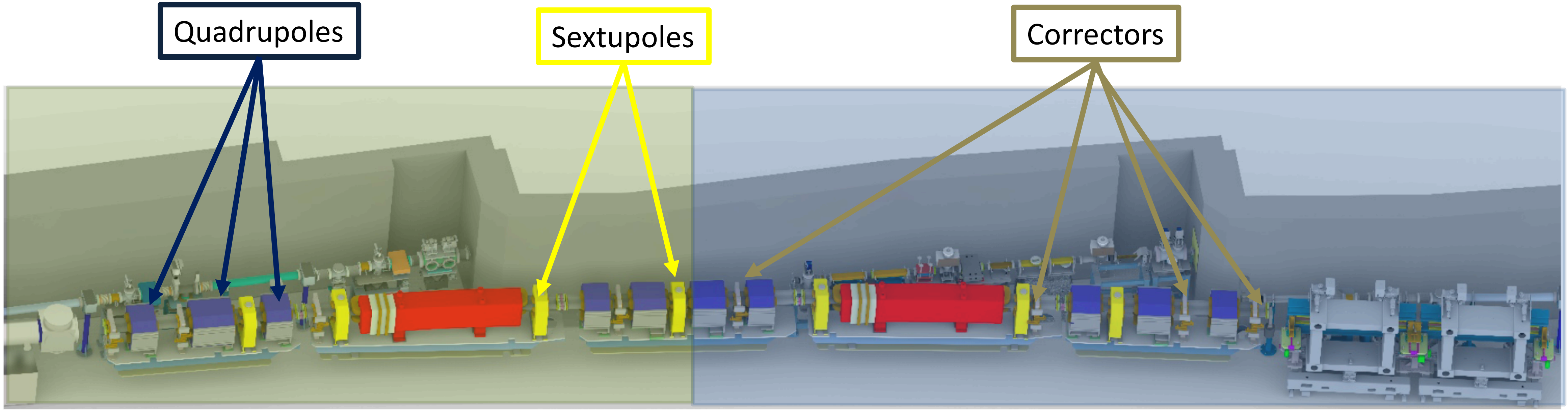
Inverse Models
(unsupervised and semi-supervised detection of beamline errors)

Autoencoders
(unsupervised and semi-supervised detection of fault precursors)

Detecting faulty magnet power supplies in the APS

- Can we predict if a fault will occur?
  - If yes, can we predict which magnet will fault

- Components of interest
  - 1320 magnet power supplies / 40 sectors (each has A (green) and B (blue) sections)

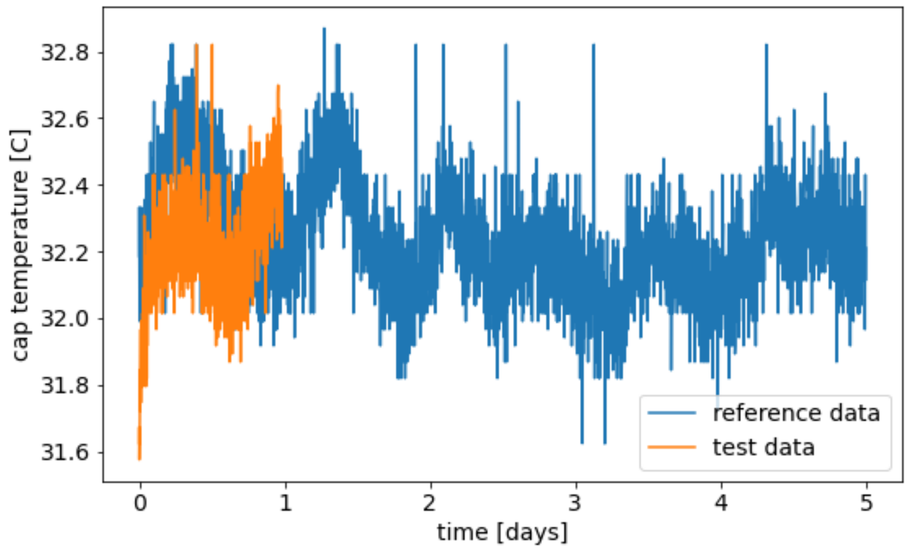


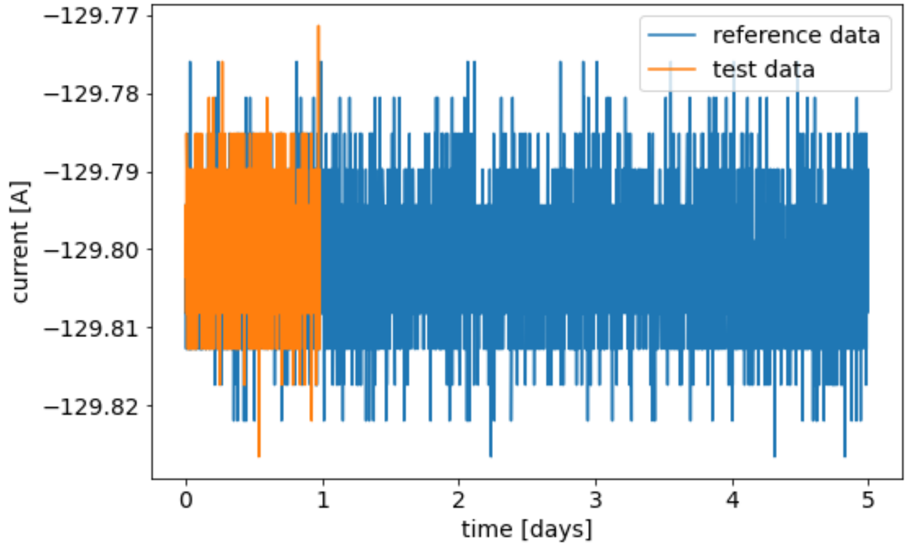
https://www.energy.gov/sites/prod/files/2019/04/f62/Advanced-Photon-Source-Upgrade-Project.pdf

Detecting faulty magnet power supplies in the APS

- Time series data for 1320 magnets
  - Power supply cap temperature
  - Current
  - Magnet temperature

- Reference data (blue)
  - No fault occurs in vicinity, normal operations
- Test data (orange)
  - Magnet failure occurs
  - Data is clipped and does not include final minutes prior to magnet fault

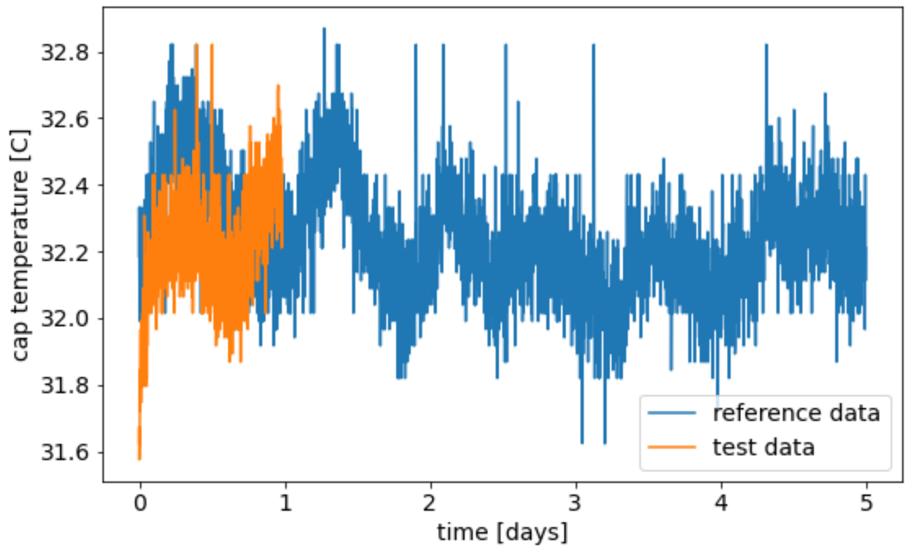


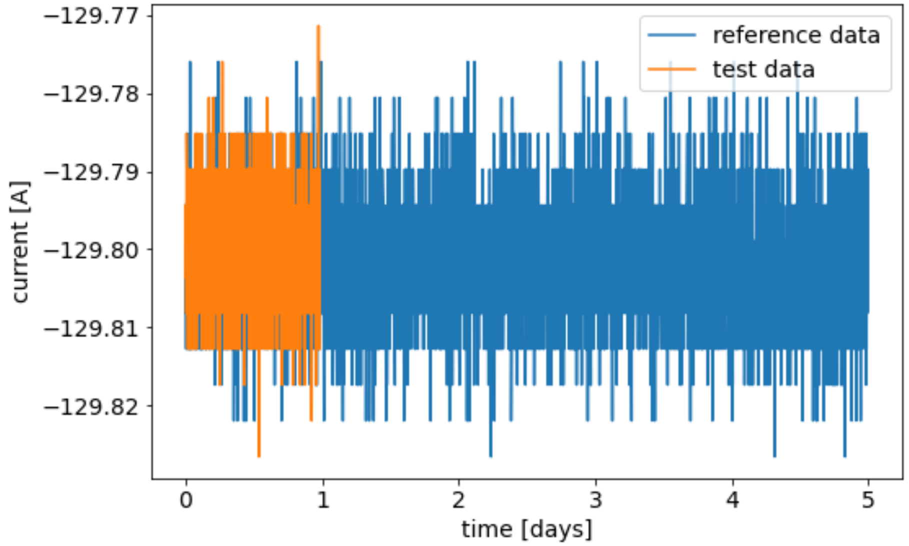


Detecting faulty magnet power supplies in the APS

- Time series data for 1320 magnets
  - Power supply cap temperature
  - Current
  - Magnet temperature

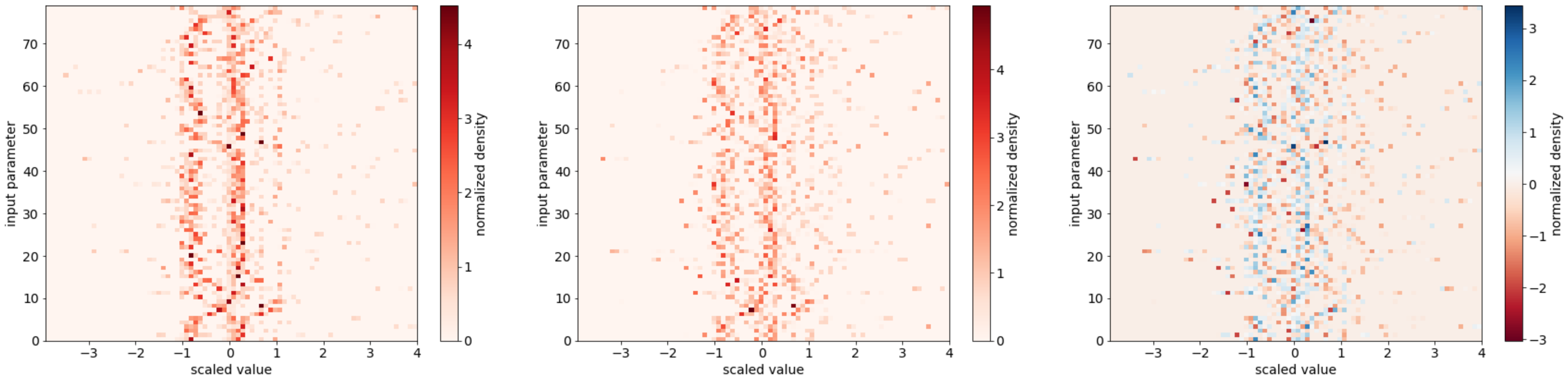
- Simplifications
  - Aggregate by each section in a sector: sum current across magnets in a sector (80 inputs/outputs)
  - Aggregate by magnet type in sectors: sum current across each magnet type in a sector (320 inputs/outputs)
  - Consider magnet current or temperature





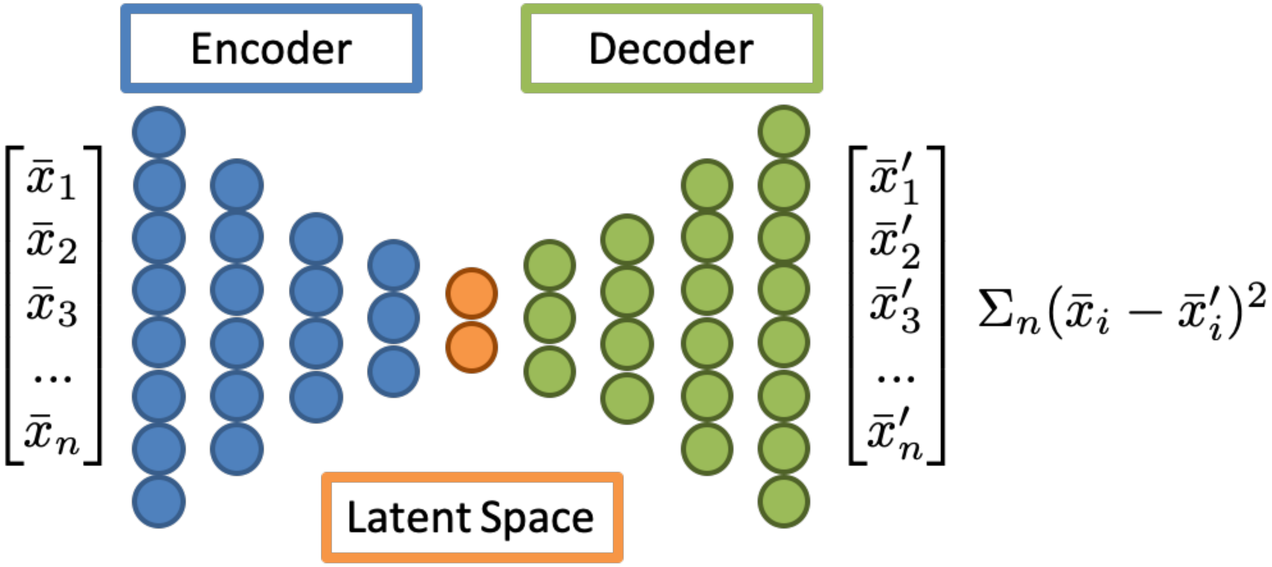
Aggregating by sector

- Overview of dataset
  - Reference data (left) used for training and validation
  - Test data (middle) with known anomalies
  - Histogram difference (right)
- Clear visual differences but datasets are qualitatively similar
- Differences between sectors are apparent



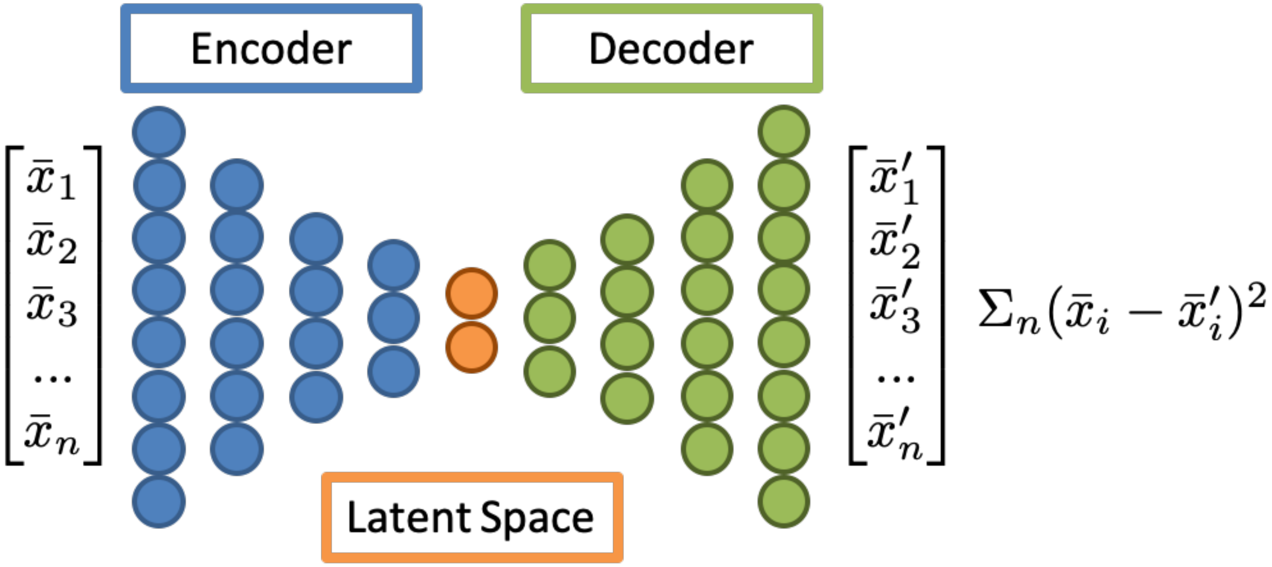
Machine learning for anomaly detection

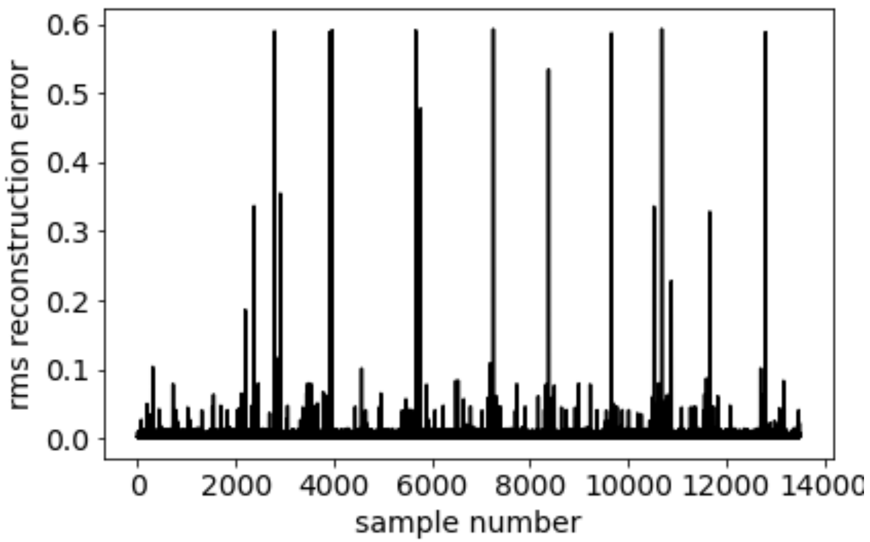
- Reconstruct unknown data using an autoencoder
  - Train and validate the autoencoder on known good datasets
  - Test on unknown data (may be good or bad)
  - Measure the degree to which the autoencoder successfully reconstructs the unknown data



Machine learning for anomaly detection

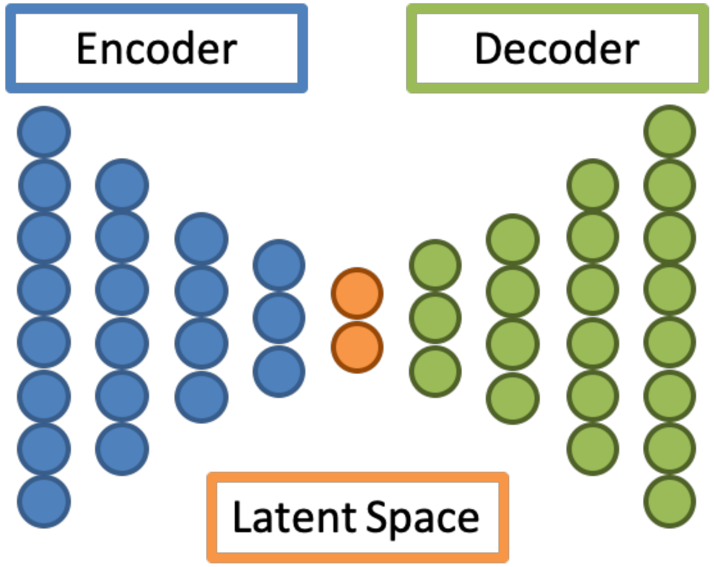
- Reconstruct unknown data using an autoencoder
  - Train and validate the autoencoder on known good datasets
  - Test on unknown data (may be good or bad)
  - Measure the degree to which the autoencoder successfully reconstructs the unknown data



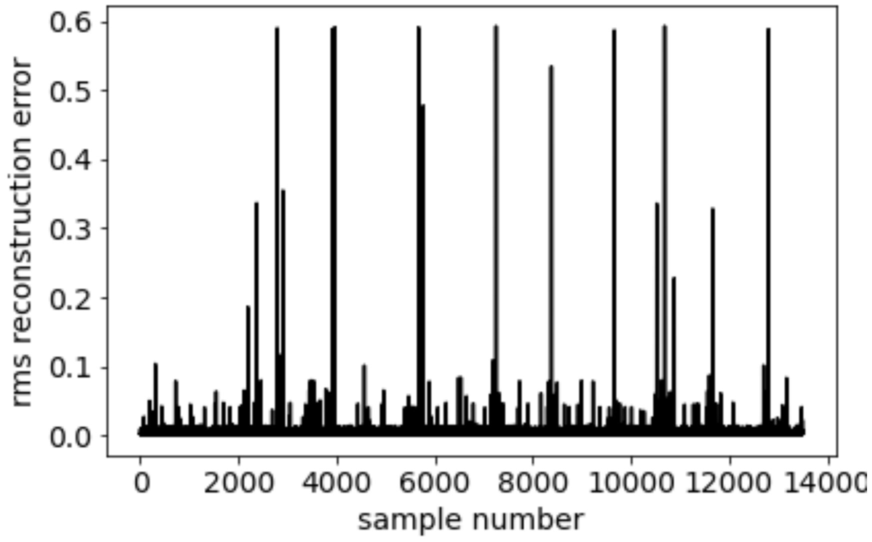


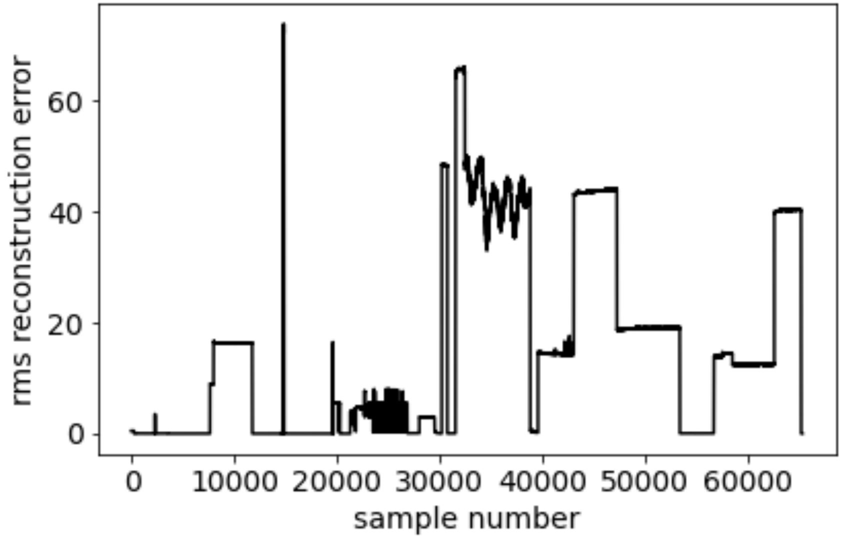
Machine learning for anomaly detection

- Reconstruct unknown data using an autoencoder
  - Train and validate the autoencoder on known good datasets
  - Test on unknown data (may be good or bad)
  - Measure the degree to which the autoencoder successfully reconstructs the unknown data



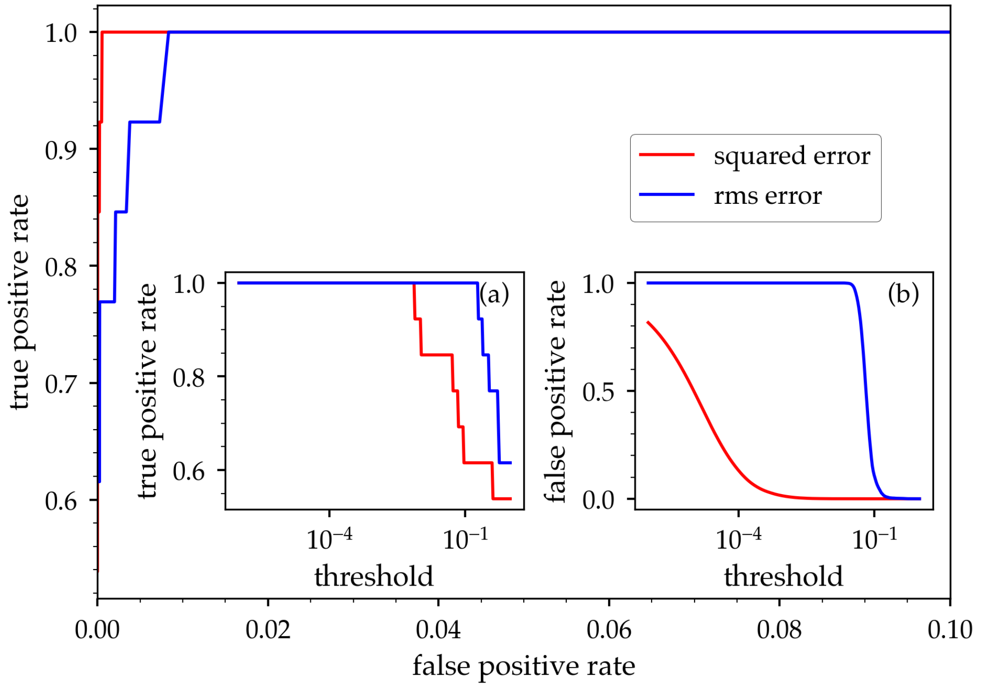
$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \cdots \\ \bar{x}_n \end{bmatrix} \qquad \begin{bmatrix} \bar{x}'_1 \\ \bar{x}'_2 \\ \bar{x}'_3 \\ \cdots \\ \bar{x}'_n \end{bmatrix} \qquad \Sigma_n(\bar{x}_i - \bar{x}'_i)^2$$



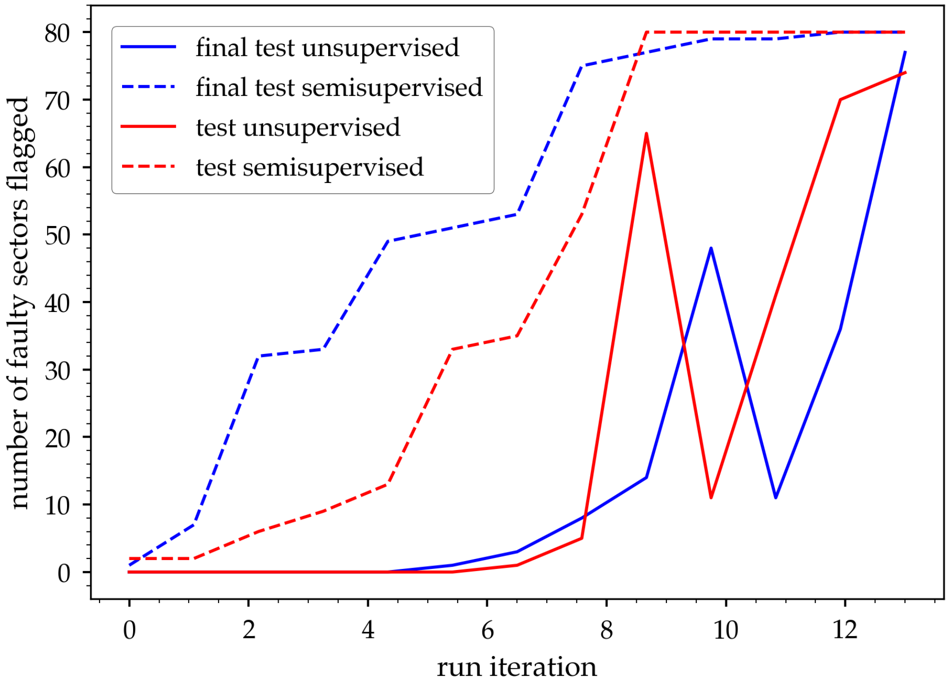


Reconstruction of reference data and test data (by sector)

Region of convergence plot for the RMS error and squared error evaluation metrics. The main plot shows the true positive rate vs the false positive rate as a function of anomaly threshold. Inset a) shows the true positive rate as a function of the error threshold and inset b) shows the false positive rate as function of the error threshold. Note that the threshold is normalized to the peak value of the reconstruction error computed on the reference data.

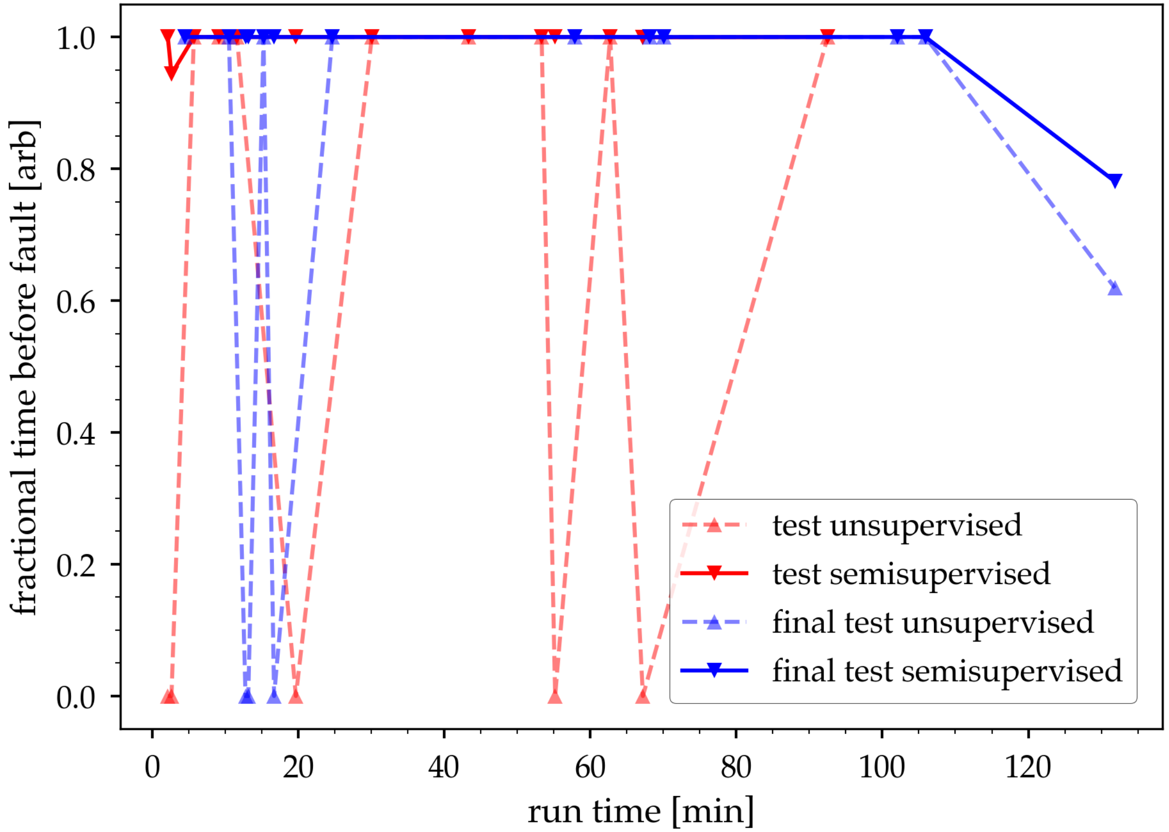


Number of faulty sectors for a given fault run. The data are sorted by the number of faulty sectors identified in the semisupervised case.

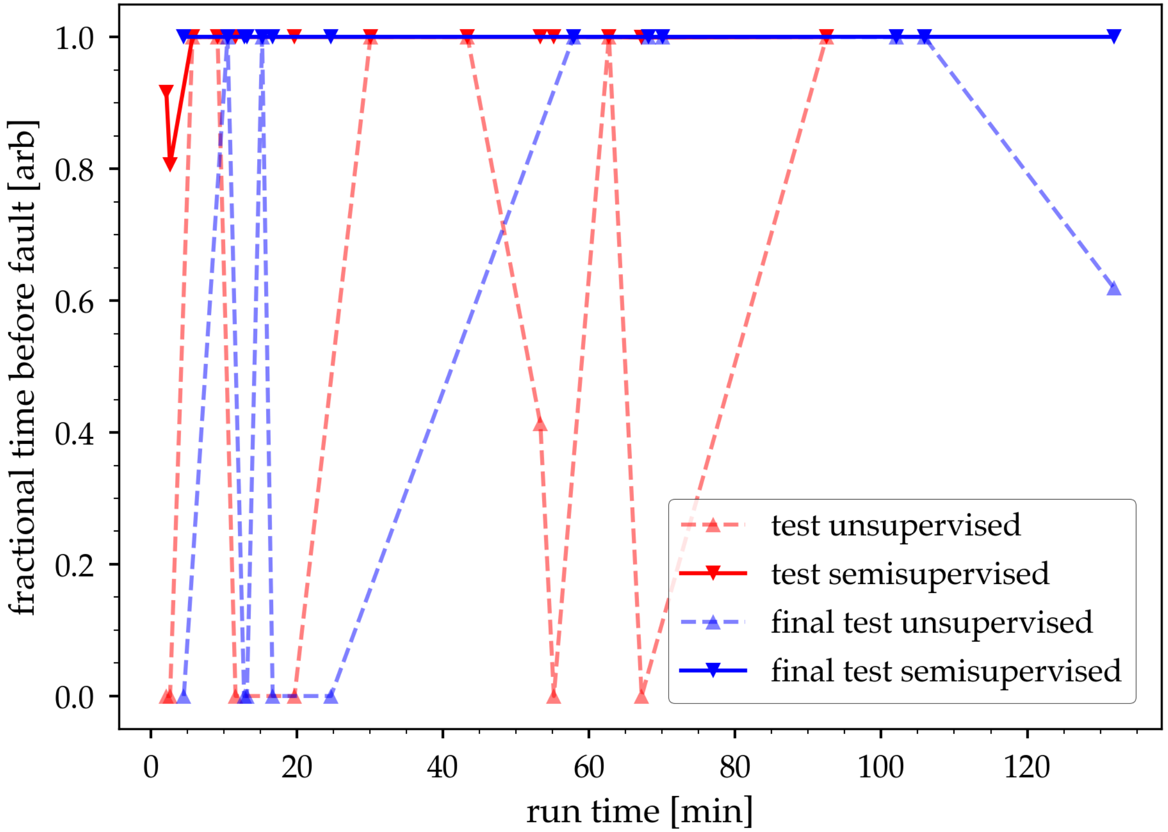


Forecasting faults using unsupervised and semi-supervised learning

First indication of an anomaly as a function of the run time for the fault data using the RMS error metric. Red is the data used to tune the detection threshold while blue is the final test data that is not used in any of the training or parameter tuning. The dashed lines represent the unsupervised case while the solid line is the semisupervised case.

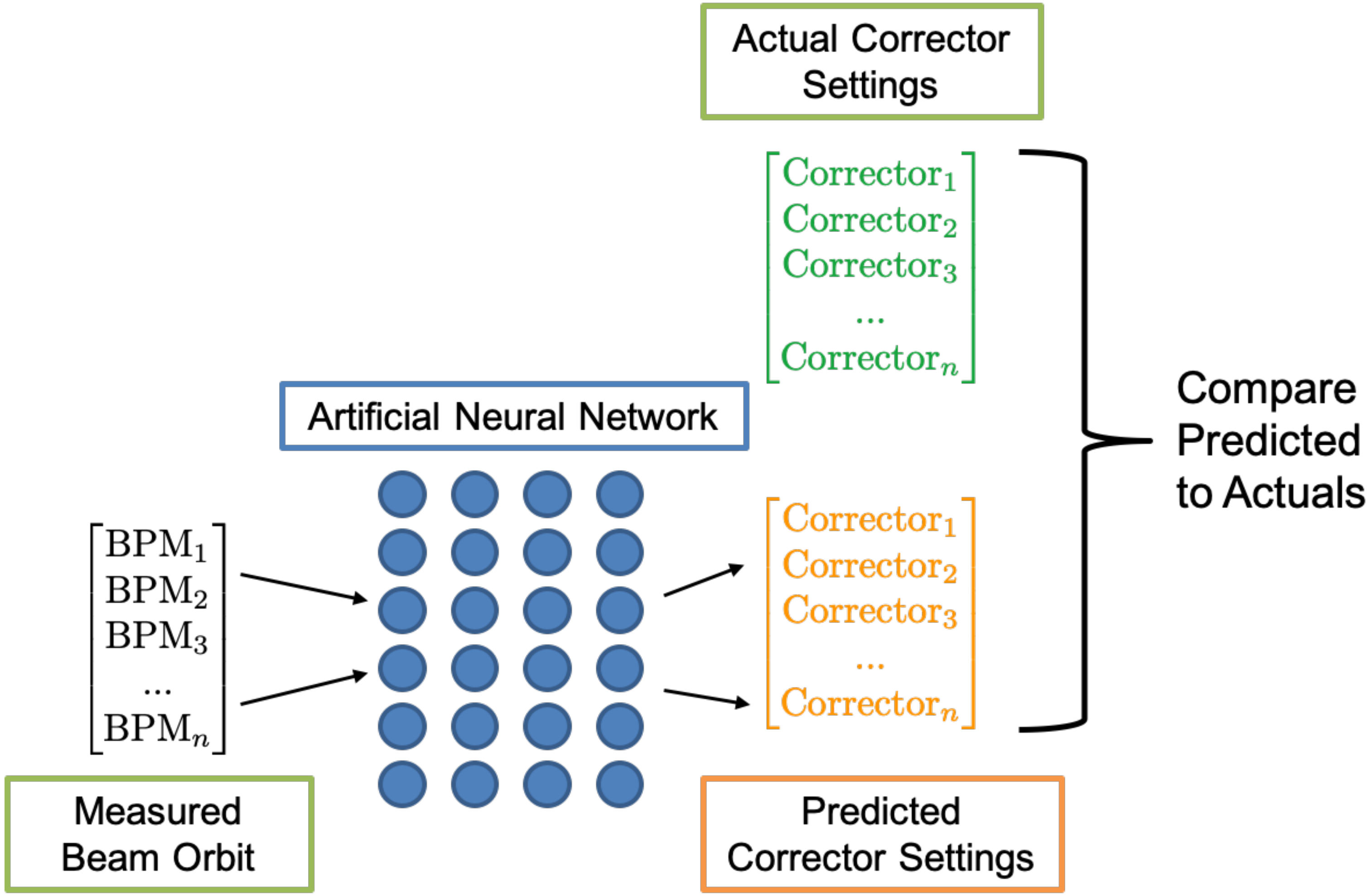


First indication of an anomaly as a function of the run time for the fault data using the squared error metric. Red is the data used to tune the detection threshold while blue is the final test data that is not used in any of the training or parameter tuning. The dashed lines represent the unsupervised case while the solid line is the semisupervised case



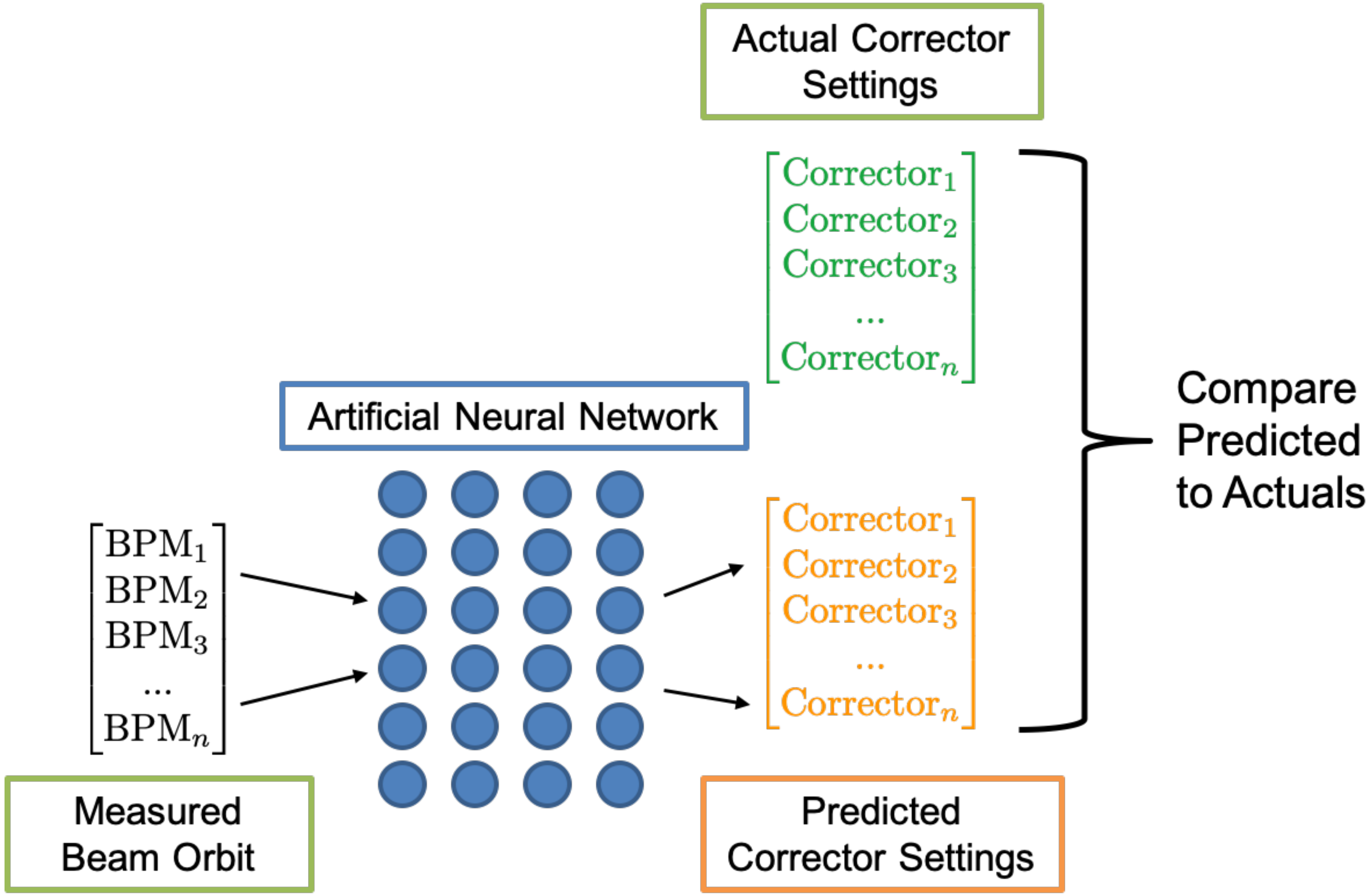
Inverse models for diagnostics

- Inverse models as a diagnostic in a supervised fashion
  - Direct comparison between predicted settings and actual settings informs operations of a potential anomaly with that magnet



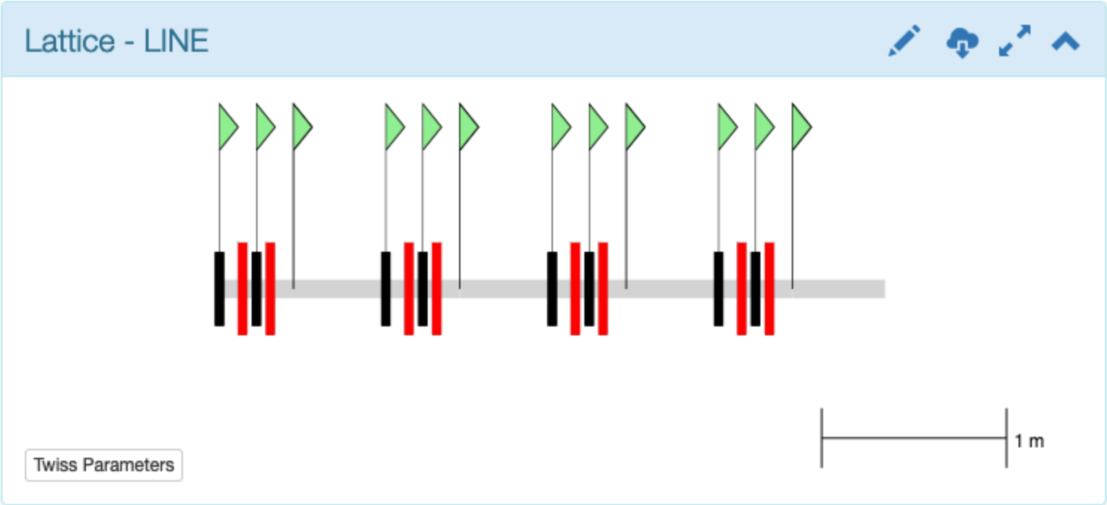
Inverse models for diagnostics

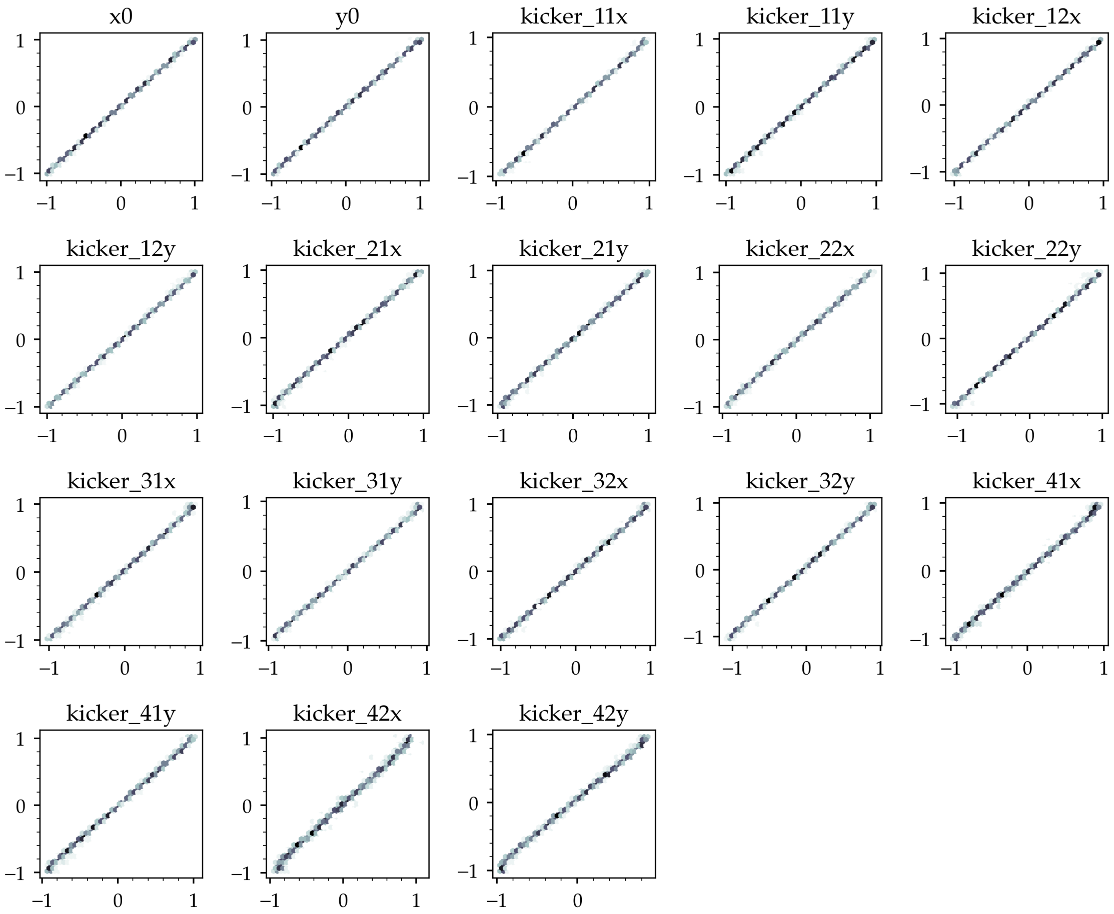
- Inverse models as a diagnostic in a supervised fashion
  - Direct comparison between predicted settings and actual settings informs operations of a potential anomaly with that magnet
- Inverse models as a diagnostic in an unsupervised fashion
  - Assumptions
    - model errors are caused by other beamline elements
    - each beam-line element will have a unique error signature



A FODO cell toy problem

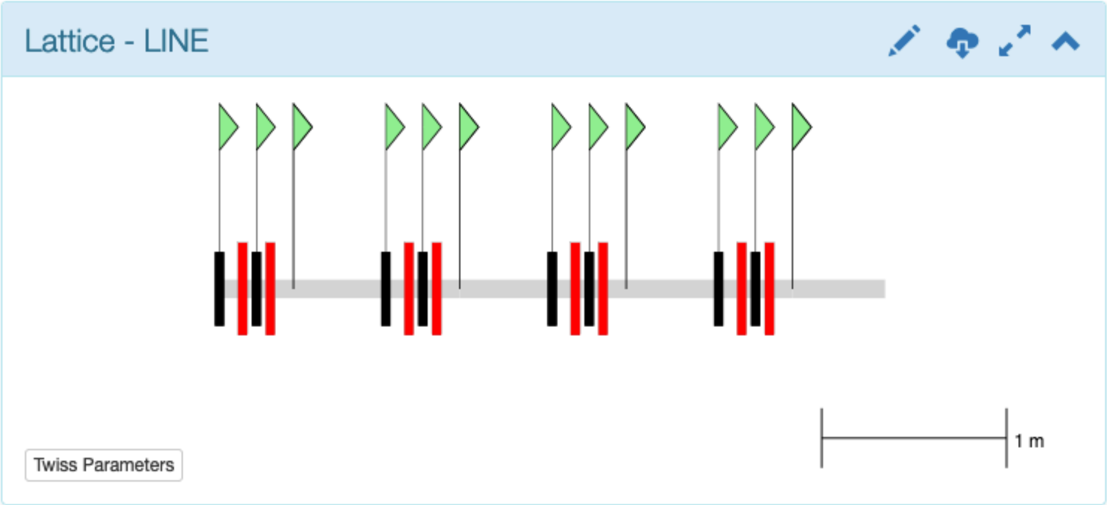
- Prototype the algorithm on a FODO cell
  - Neural network trained to predict corrector settings from BPM measurements

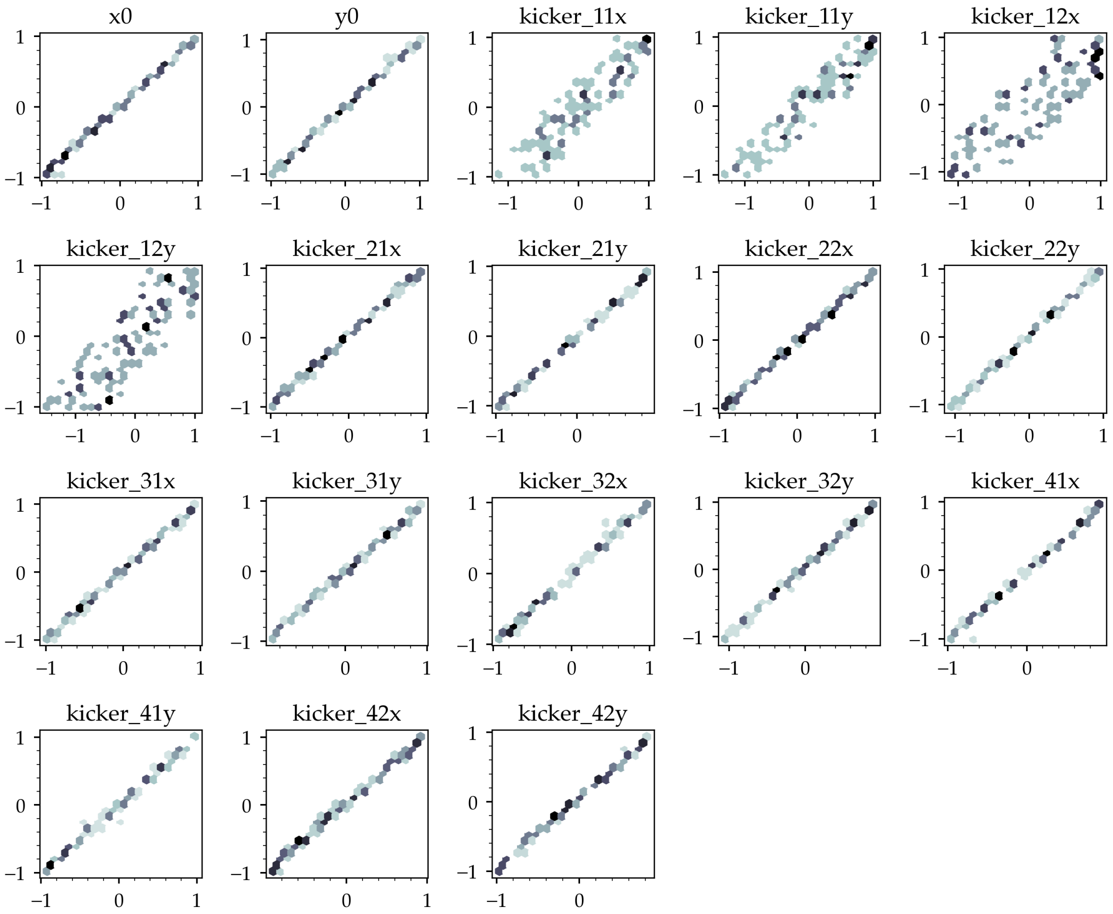




A FODO cell toy problem

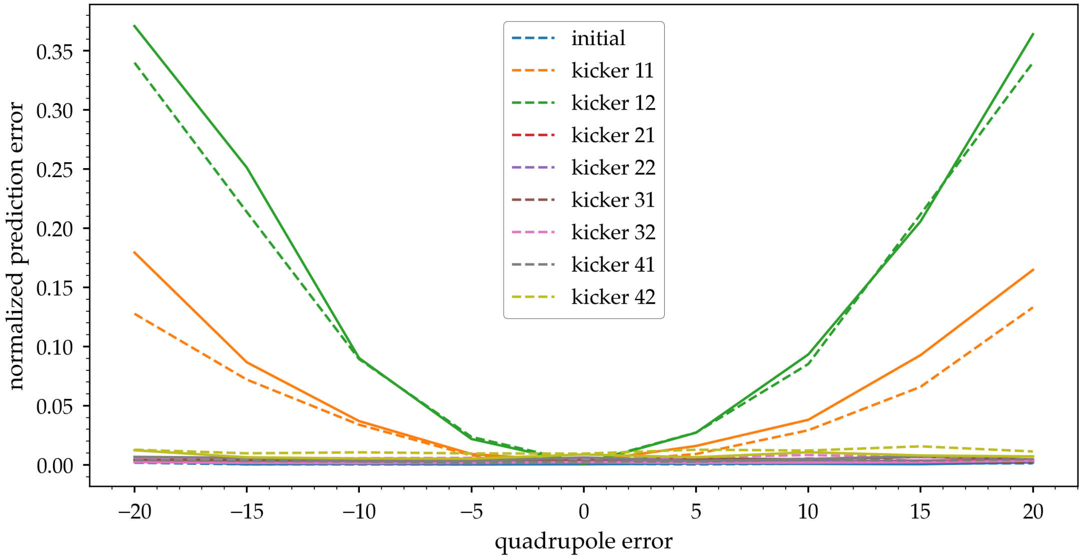
- Prototype the algorithm on a FODO cell
  - Neural network trained to predict corrector settings from BPM measurements
  - Test neural network on data with single quadrupole error

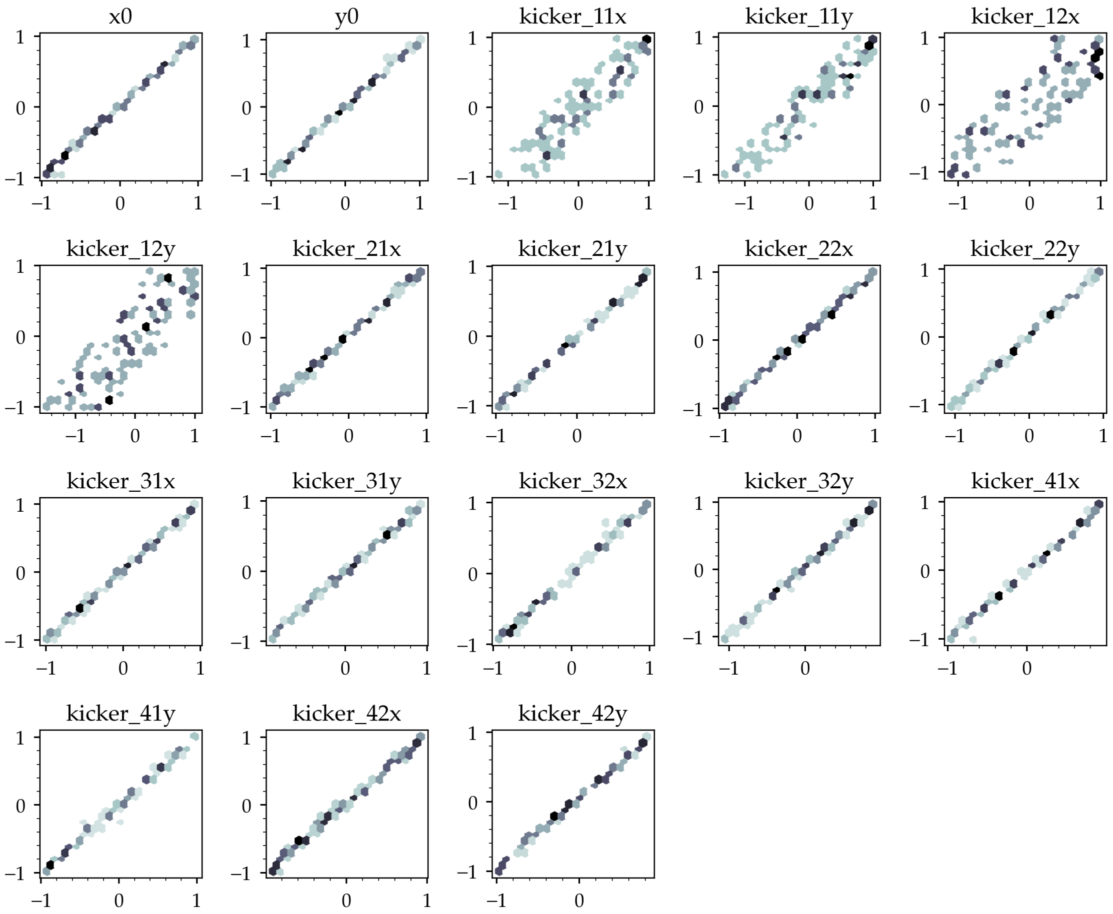




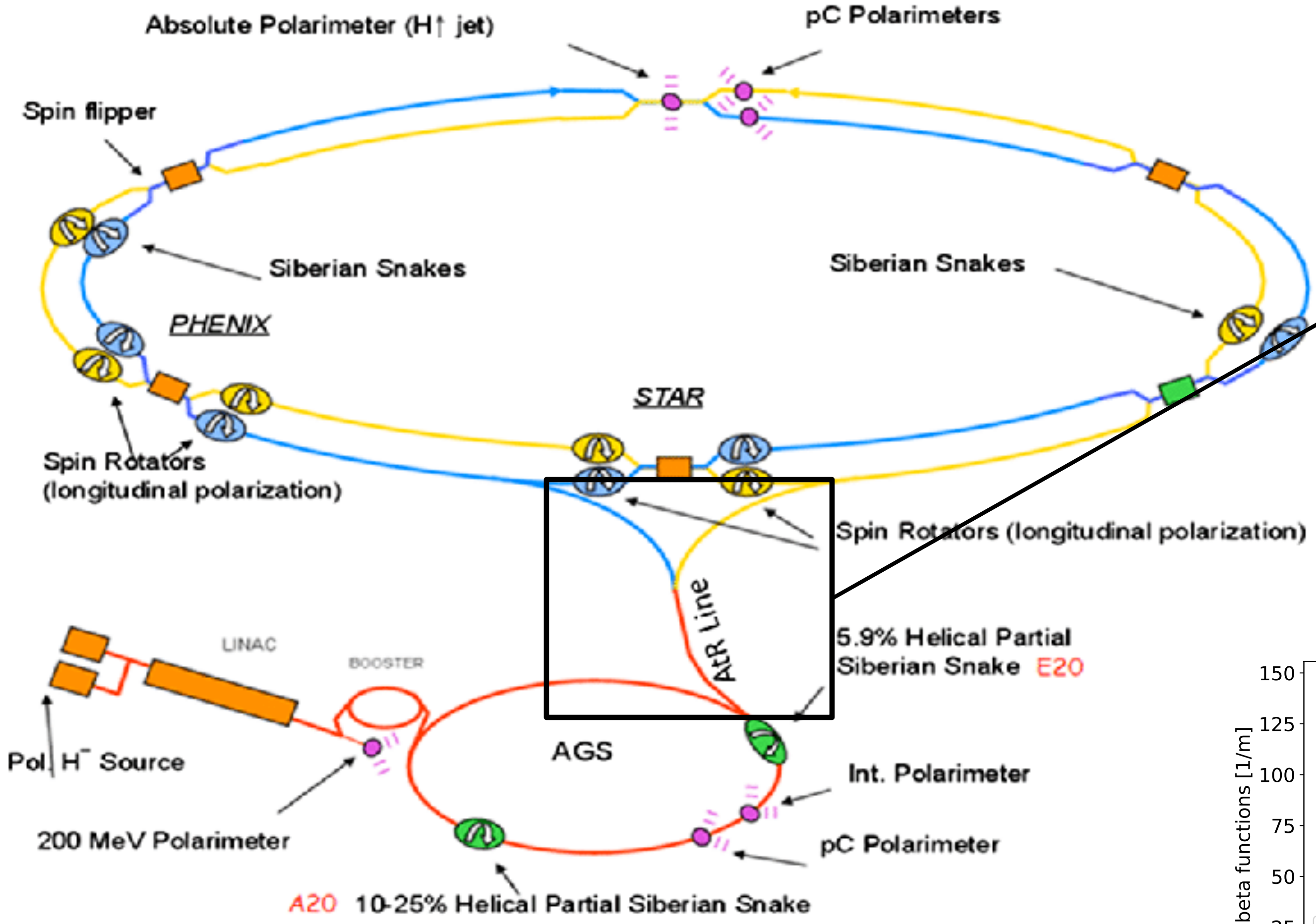
A FODO cell toy problem

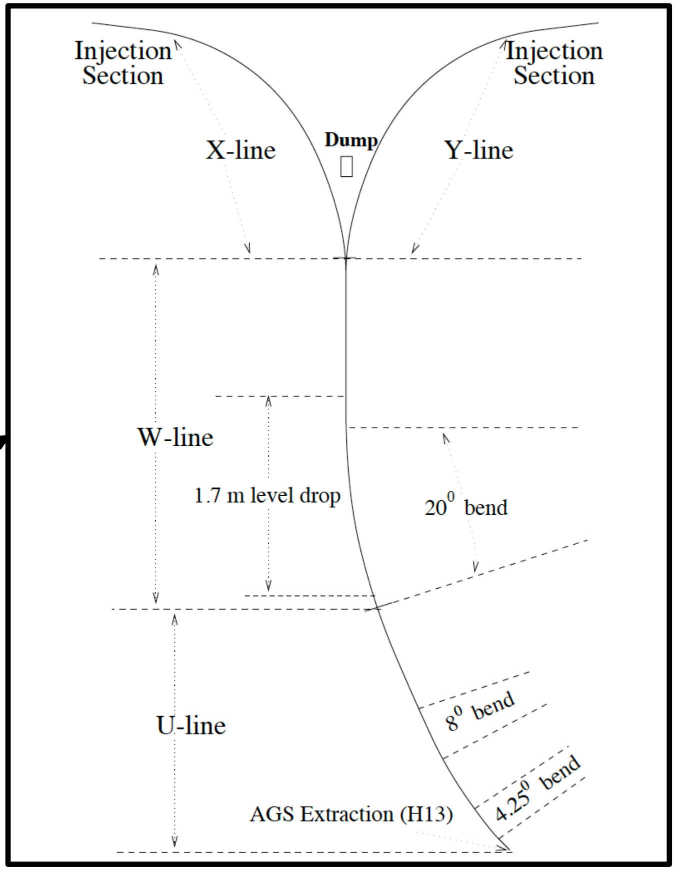
- Prototype the algorithm on a FODO cell
  - Neural network trained to predict corrector settings from BPM measurements
  - Test neural network on data with single quadrupole error
  - Study correlation between quadrupole error and model prediction error

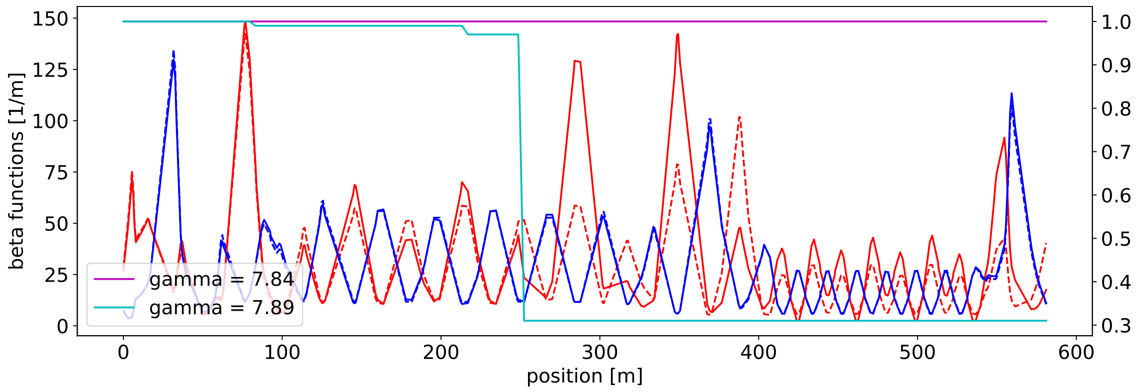




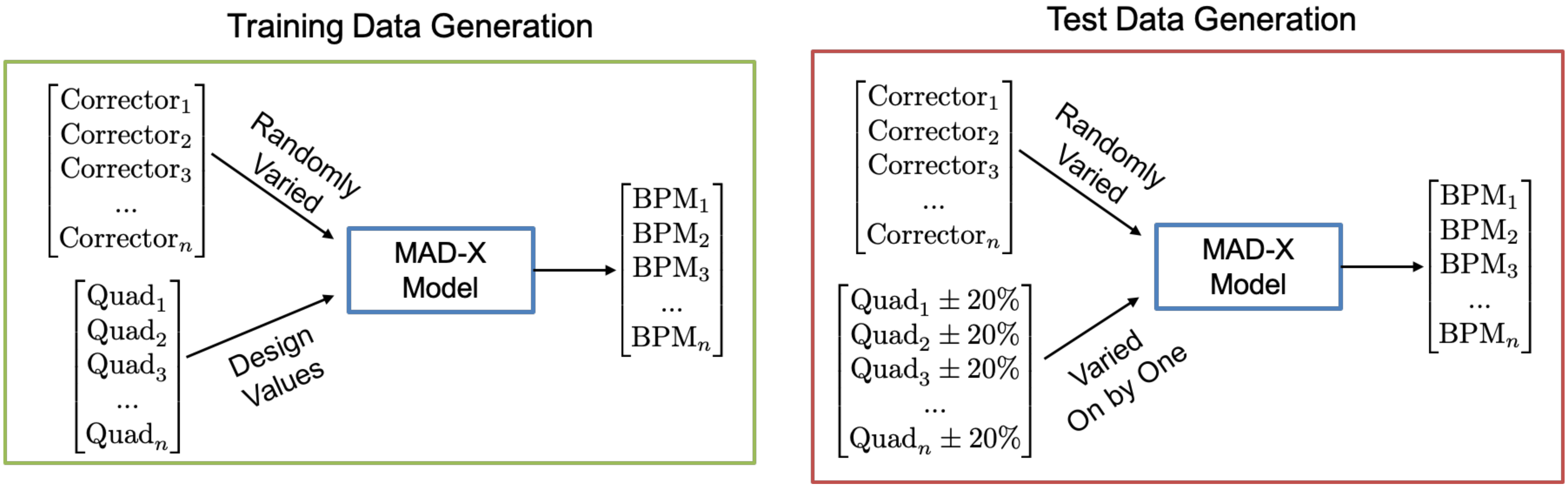
AGS to RHIC transfer line





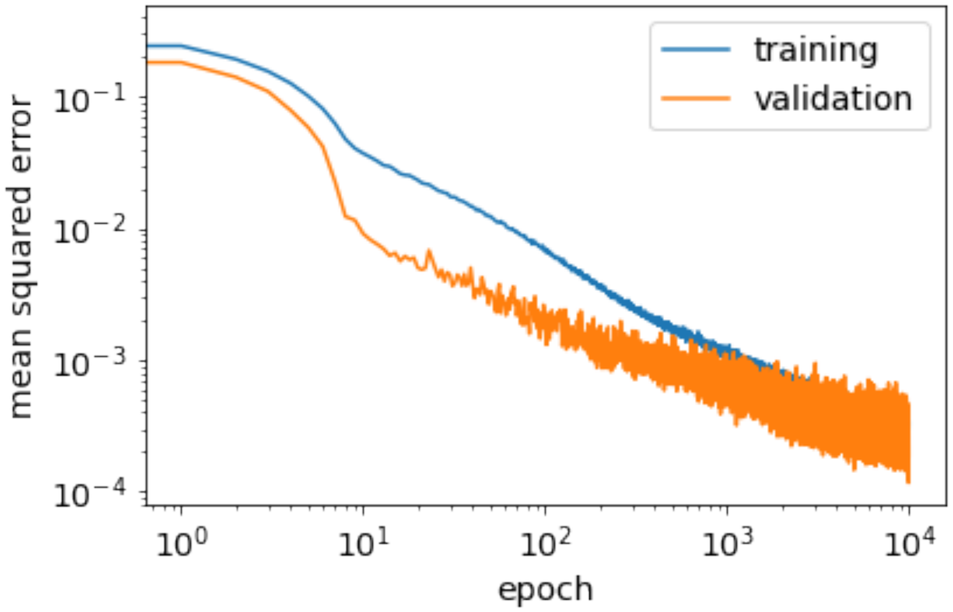


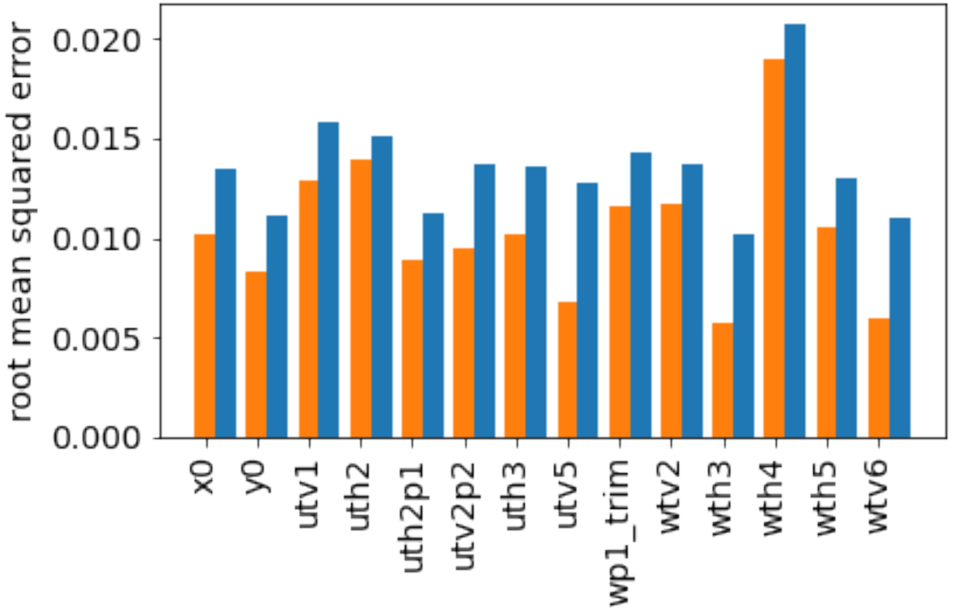
AGS to RHIC transfer line study



AGS to RHIC transfer line study

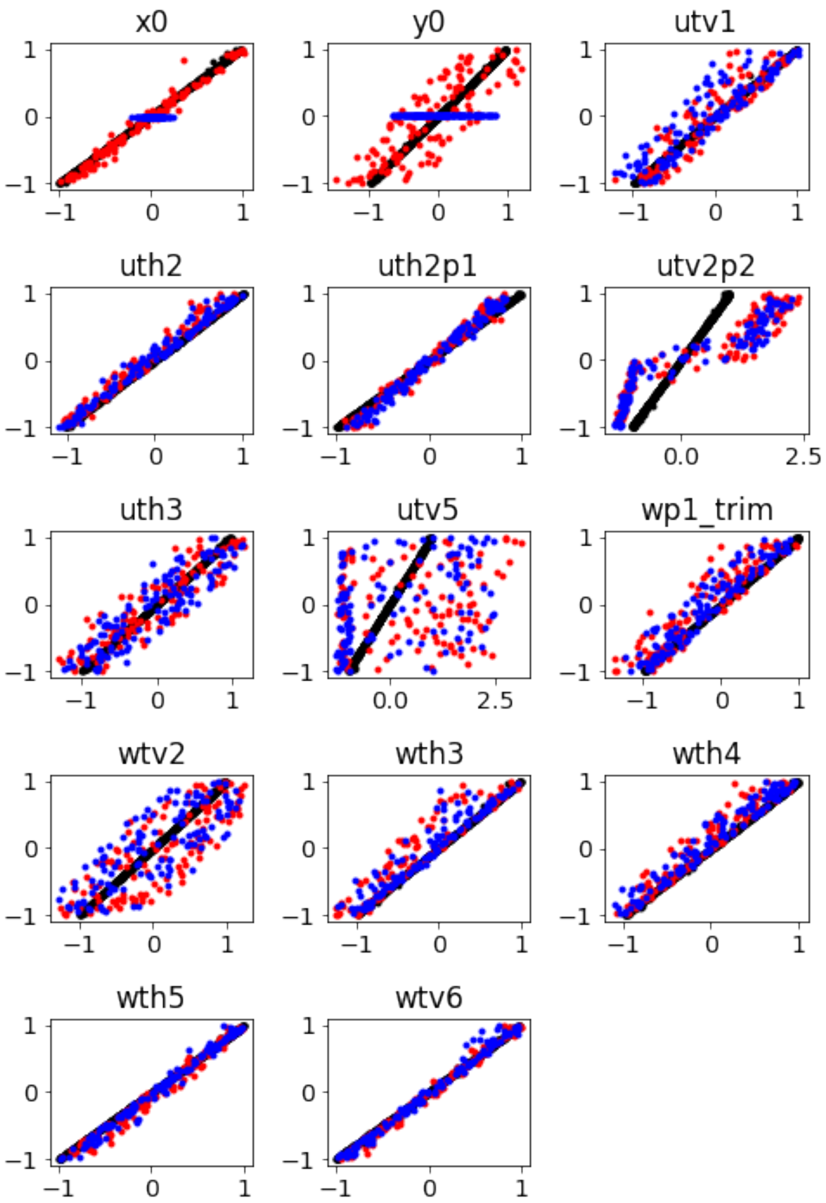
- Inverse model trained using 5000 samples, randomly varying the corrector strengths and beam initial positions.
- Removed four correctors (utv4, uth6, utv7, and wth1) from the inverse model due to degeneracy issues.
  - In future work we will address this issue
- Model / Training Parameters:
  - For this study the data were split into 80% training and 20% validation
  - 5 dense layers with 45 nodes each
  - Gaussian noise for regularization
  - Rectified linear units for the activation functions





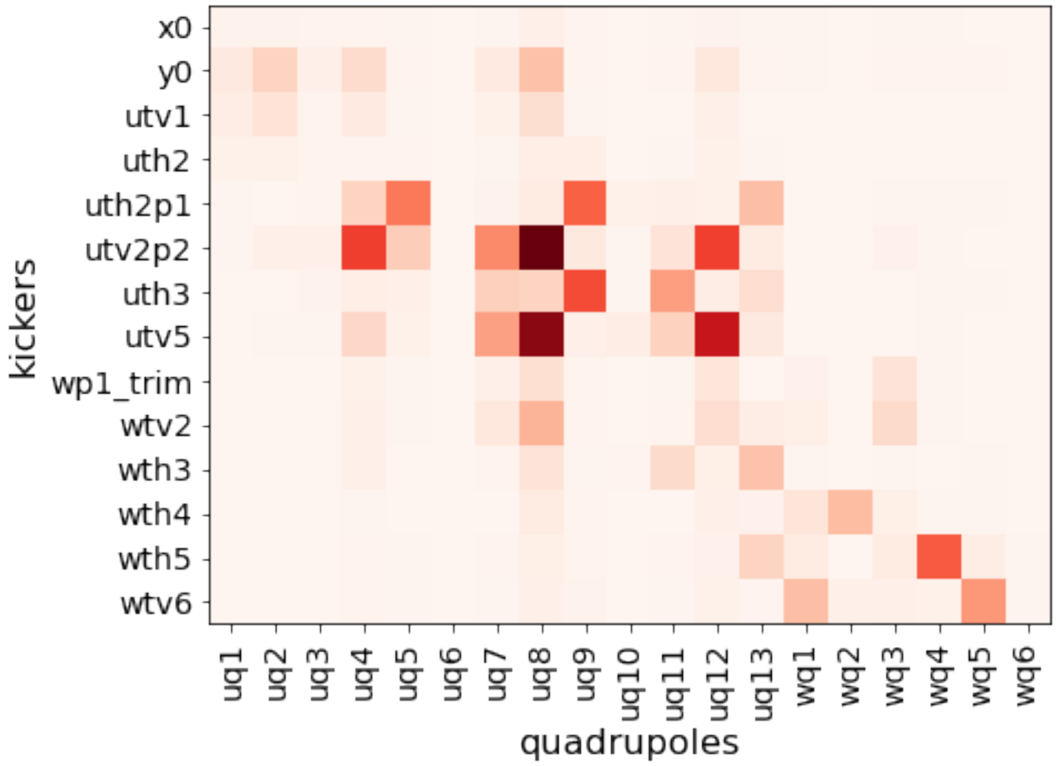
AGS to RHIC transfer line study

- Two configurations were used: one where the initial positions were also varied randomly and one where the initial positions were not varied.
- Right: Predicted corrector settings vs the ground truth for the validation set
  - Black: without quadrupole errors
  - Red: a single quadrupole error and random initial position errors
  - Blue: a single quadrupole error without initial position errors



AGS to RHIC transfer line study

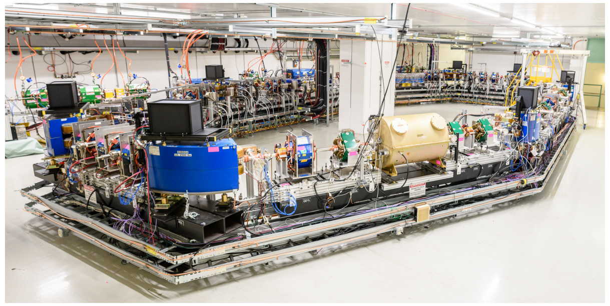
- Sensitivity of each corrector prediction to a particular quadrupole
  - Unique signatures for each quadrupole
  - The model clearly identifies errors in these magnets without any explicit knowledge of their existence



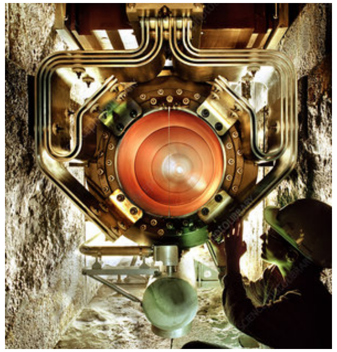
- Future work
  - Use signatures to predict unknown quadrupole errors
  - Use model errors to tune out quadrupole errors

Conclusions

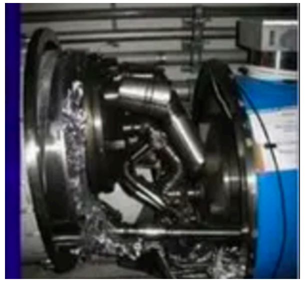
Accelerator R+D



Beam for Experimentalists



Down Time



Inverse Models
(unsupervised and semi-supervised detection of beamline errors)

Promising studies on FODO example and ATR line. Working towards deployment

Autoencoders
(unsupervised and semi-supervised detection of fault precursors)

Promising studies on the APS storge ring. Working towards new applications.

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.