AI Experimental Calibration and Control

Thomas Britton

David Lawrence, Michael Goodrich, Torri Jeske, Nikhil Kalra, Diana McSpadden, Kishansingh Rajput

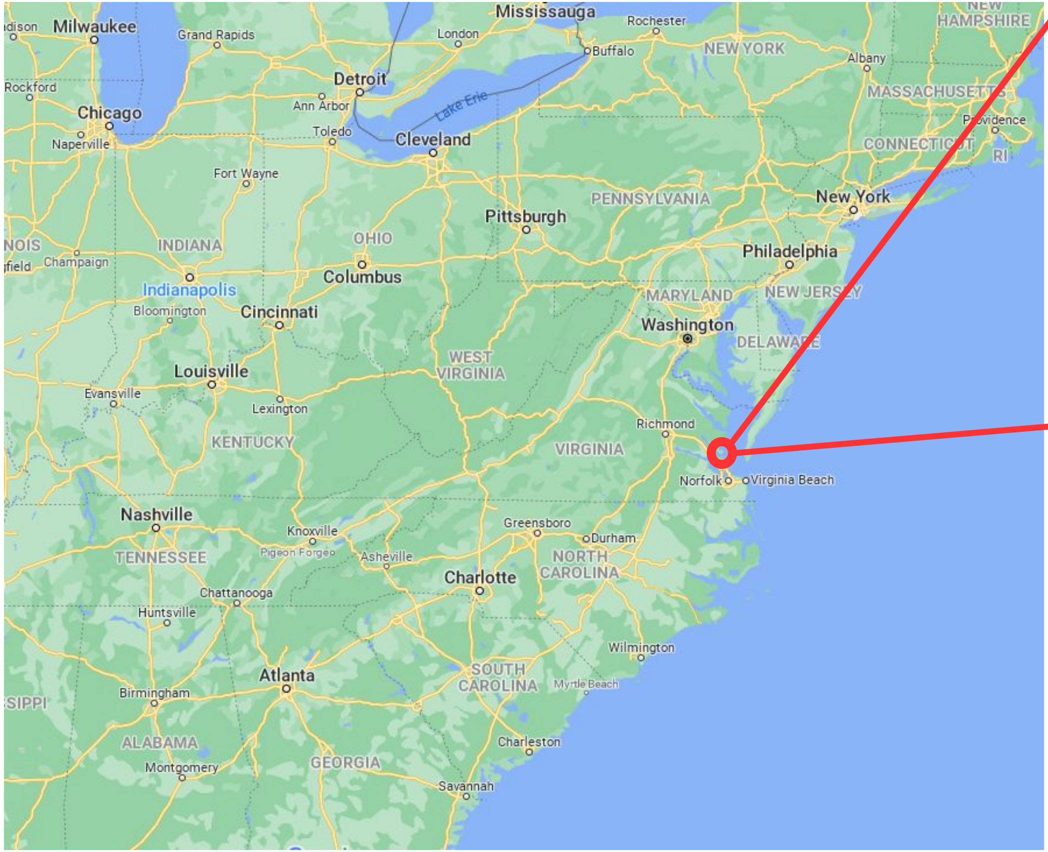
EPSCI

Naomi Jarvis

Carnegie Mellon University

U.S. DEPARTMENT OF ENERGY
JSA
GlueX
Jefferson Lab
Thomas Jefferson National Accelerator Facility

JLab

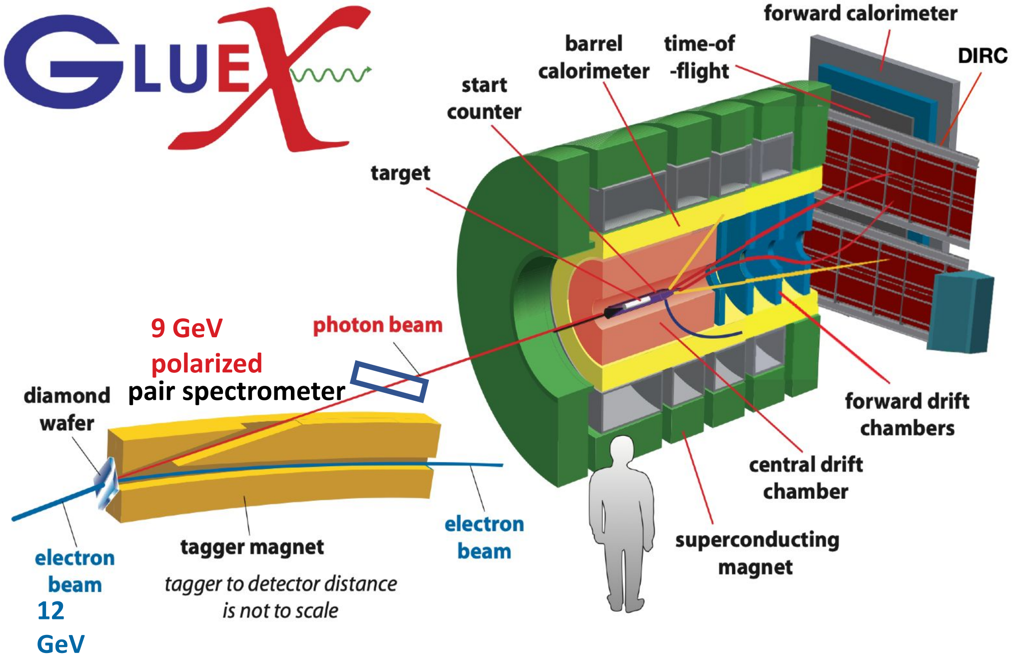


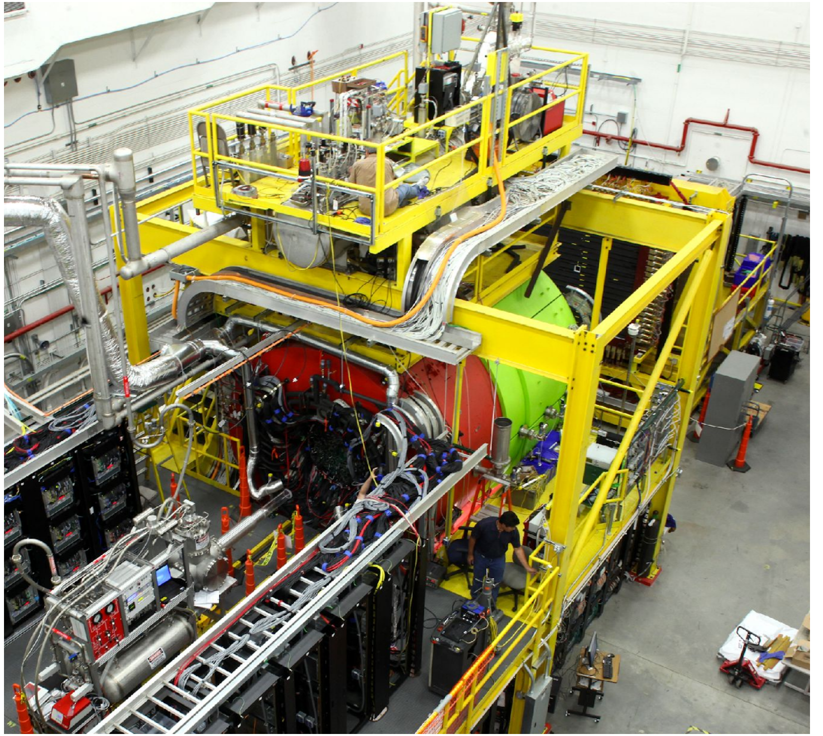
4 experimental halls
Up to 12 GeV electron beam

Focusing (initially) on **HallD**

GlueX

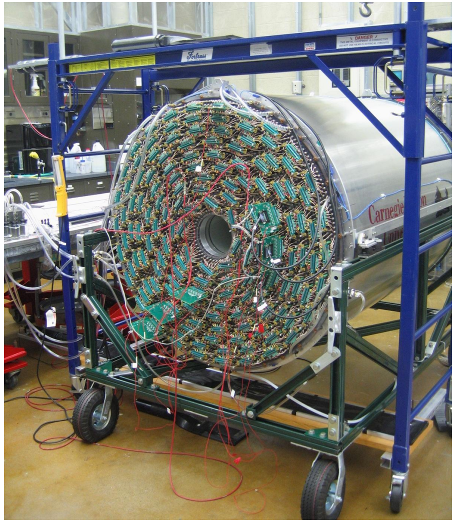
GlueX detector located in Hall D at Jefferson Lab, VA

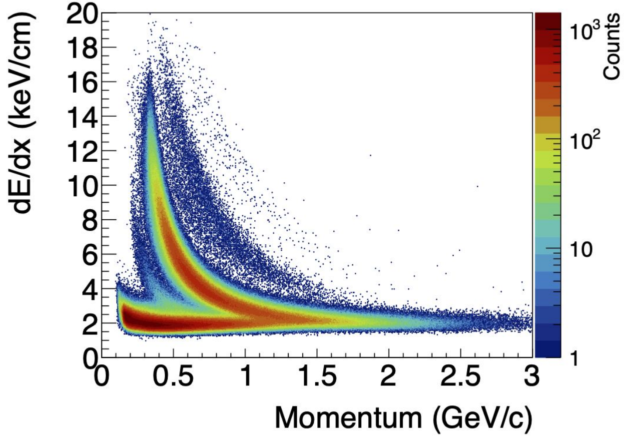




The CDC

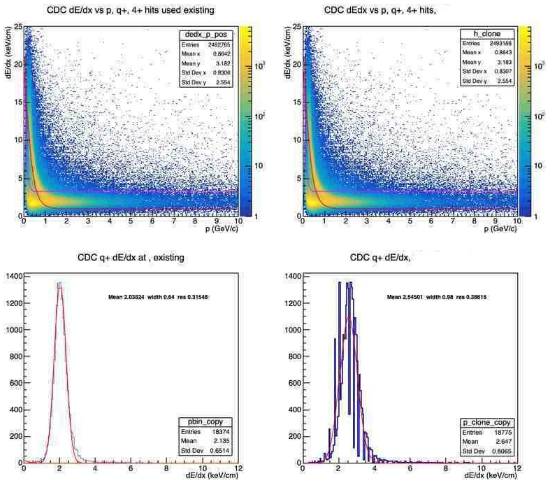
- 1.5m long x 1.2m diameter cylinder; central hole for beam, target and start counter scintillators
- 3522 anode wires at 2125V inside 1.6cm diameter straw
- Ar/CO2 gas mix, approx. 30 Pa above atmospheric pressure
- Measures drift time and deposited charge





Motivation

- Calibrations cause a delay between data collection and analysis
- At present several calibration rounds are used, due to interplay between subdetector calibrations
- Calibration could be made more efficient using AI (less iterations)
  - Less cpu time
  - Less personal attention from experts
  - * We expect to fine-tune the calibrations in the usual way
- **CDC gain calibrations** have the most variation **+/- 33%**

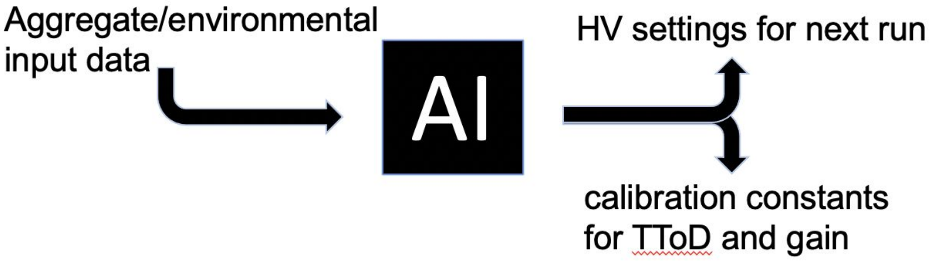


Artificially modified gain 33%

- If we know what gain to expect before taking data, we can **adjust the HV to maintain constant gain**
  - Perhaps eliminating the need to perform gain calibrations at all…

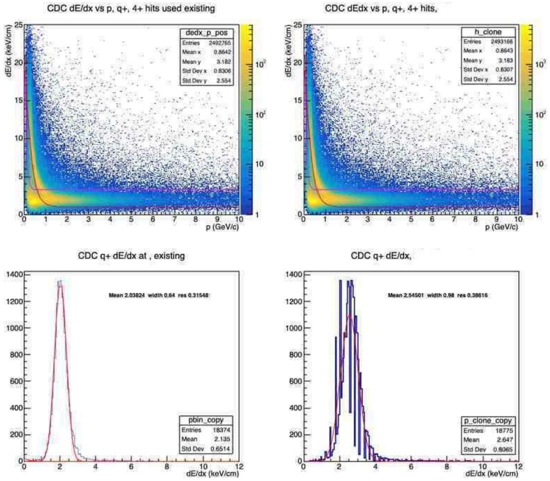
Goals

- AI-recommended HV settings to maintain GlueX Central Drift Chamber gain
  - E.g. Chamber gain is sensitive to atmospheric pressure
- **Have neural network determine calibration constants as quickly as possible**
  - **Reduce time for offline calibration**
- Apply tech to other detector systems such as CLAS12 spectrometer



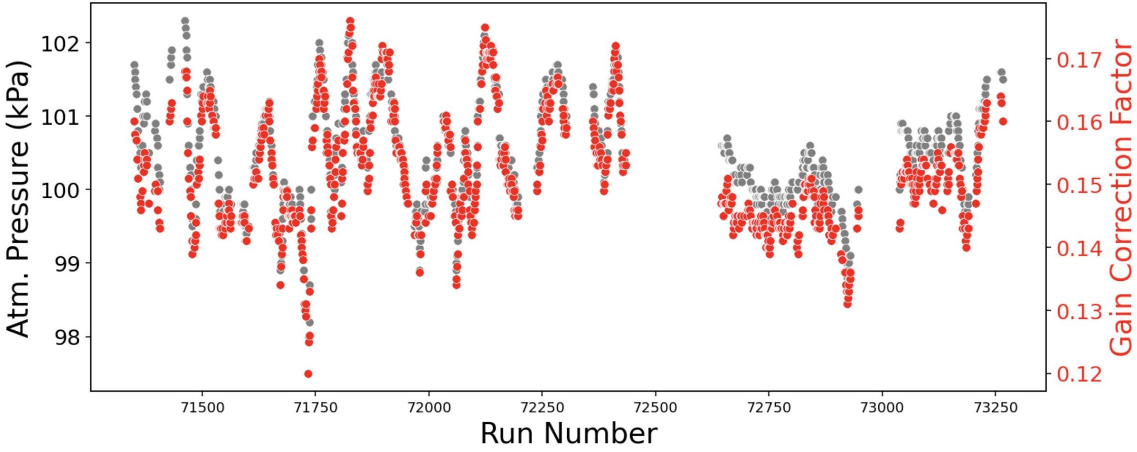
Plan of attack

- Start with the gains
  - Use **traditional methods as “ground truth”**
- Develop a voltage recommender
  - Stabilize gains
- Time-to-distance
  - First with traditional methods
  - Then with physics based methods
- Application to other experiments

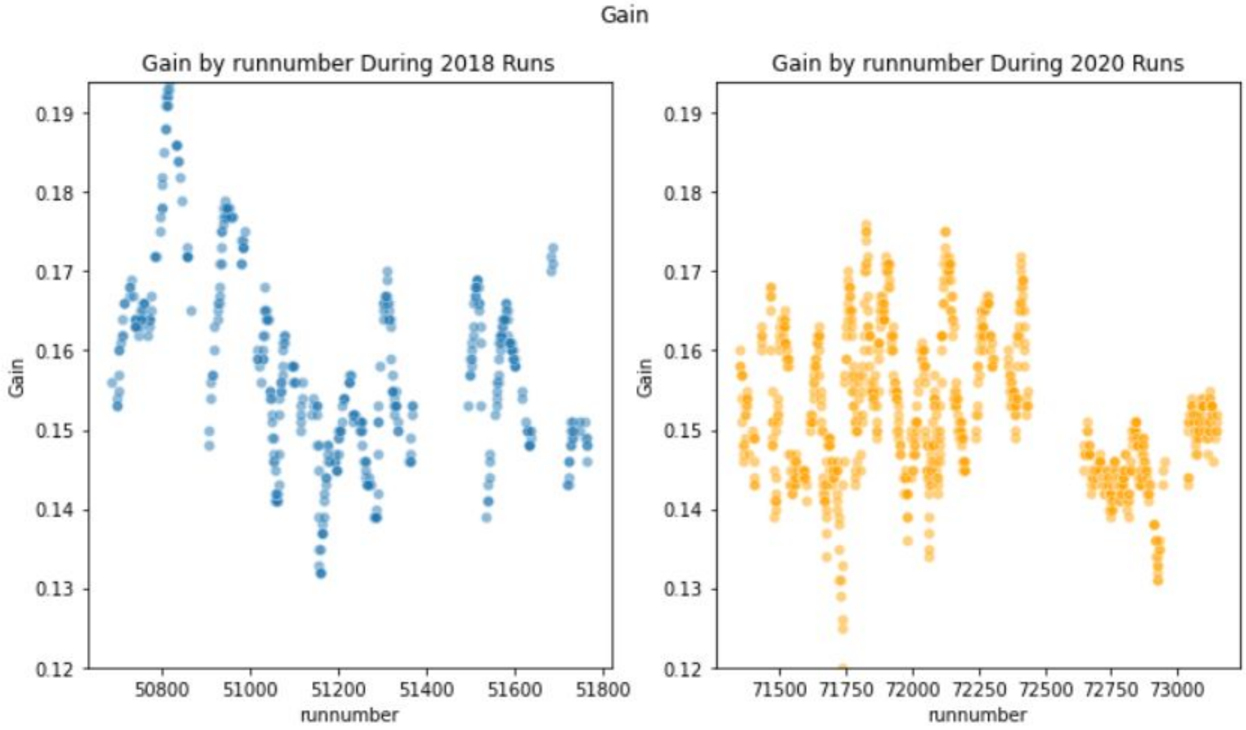


CDC gains

- Time to distance -> track-fitting, vertex resolution and dE/dx resolution
- Gain -> stable dE/dx throughout the run, affects PID selections in analysis. Environmental conditions, eg atmospheric pressure, affect the chamber gain
- Data-taking divided into *runs* (up to 2h), each session of data taking spans several months



CDC gains



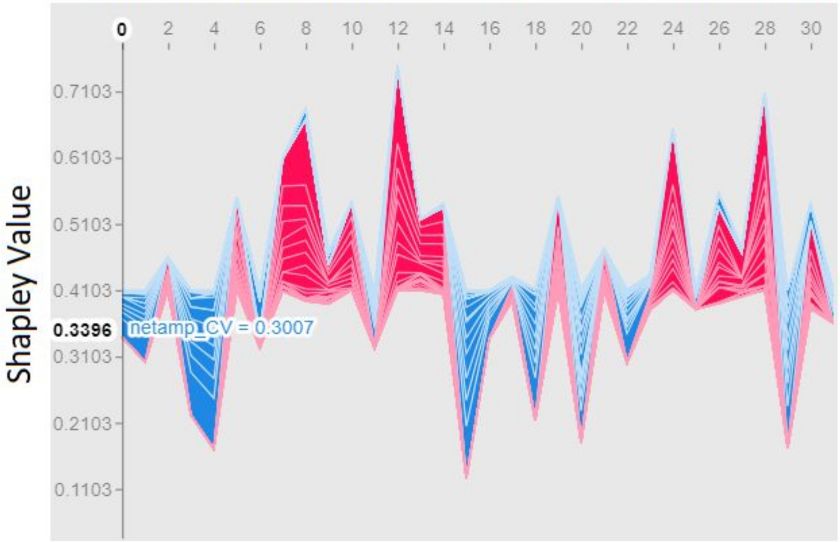
Can we use AI to predict existing gain constants to within ~1%?

CDC gains

For regression problems, there are a number of available evaluation methods.

We implemented Shapley values.

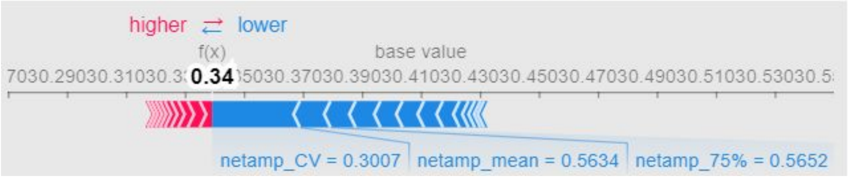
Non-linear relationship of 39 features on Gain constant



Test data set

"The Shapley value is a framework originally proposed in the context of game theory to determine individual contributions of a set of cooperating players" - Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications | IEEE Journals & Magazine | IEEE Xplore

The first test data set run:
effect of features on Gain constant for this run



U.S. DEPARTMENT OF ENERGY JSA Jefferson Lab

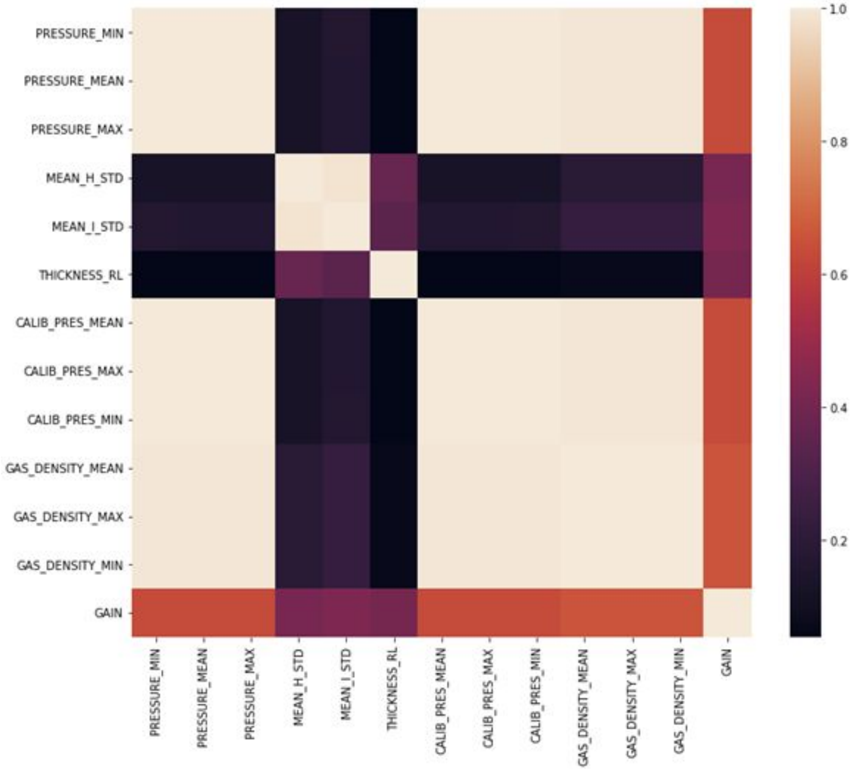
Feature importance

Input Features:

- Aggregate features per run from experimental data and EPICS system:
    - Netamp = pulse height - pedestal, momentum, track angle, drift time

- Split data into train and test sets:
    - 438 runs from 2018
        - 350 train
        - 88 test
    - 897 runs from 2020
        - 717 train
        - 180 test

- **Iterate feature importance** to help with feature engineering and minimize needed data/model size

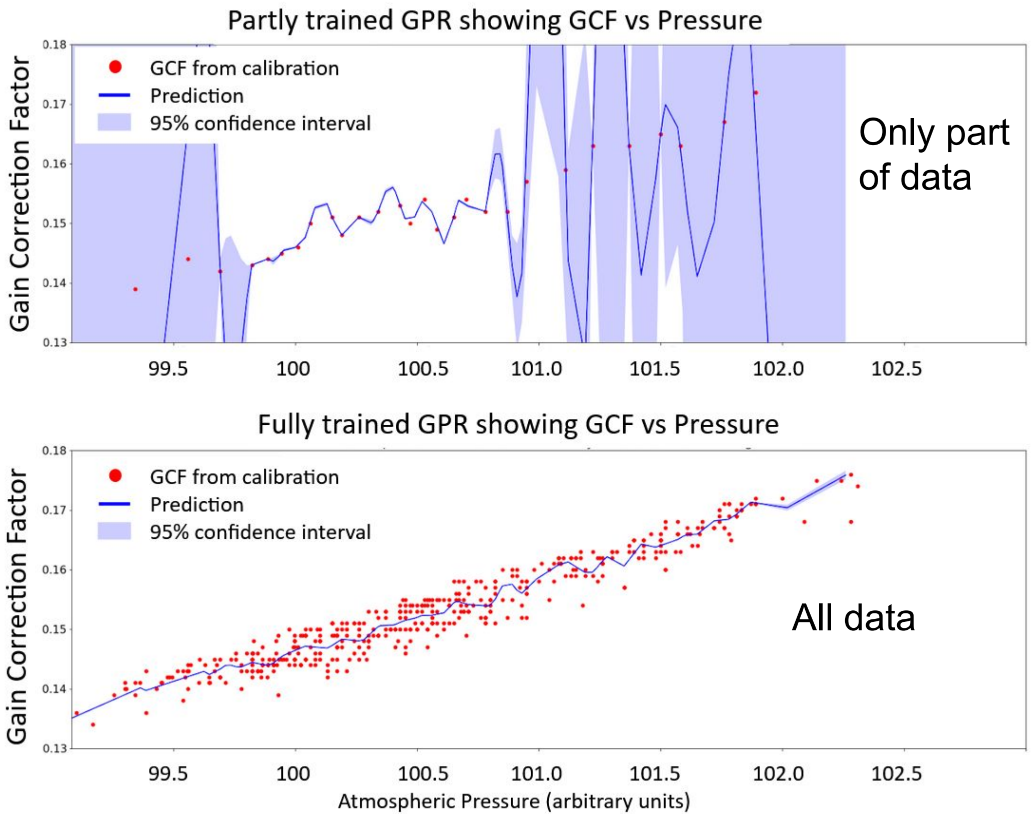


Gaussian Process Regression (GPR)

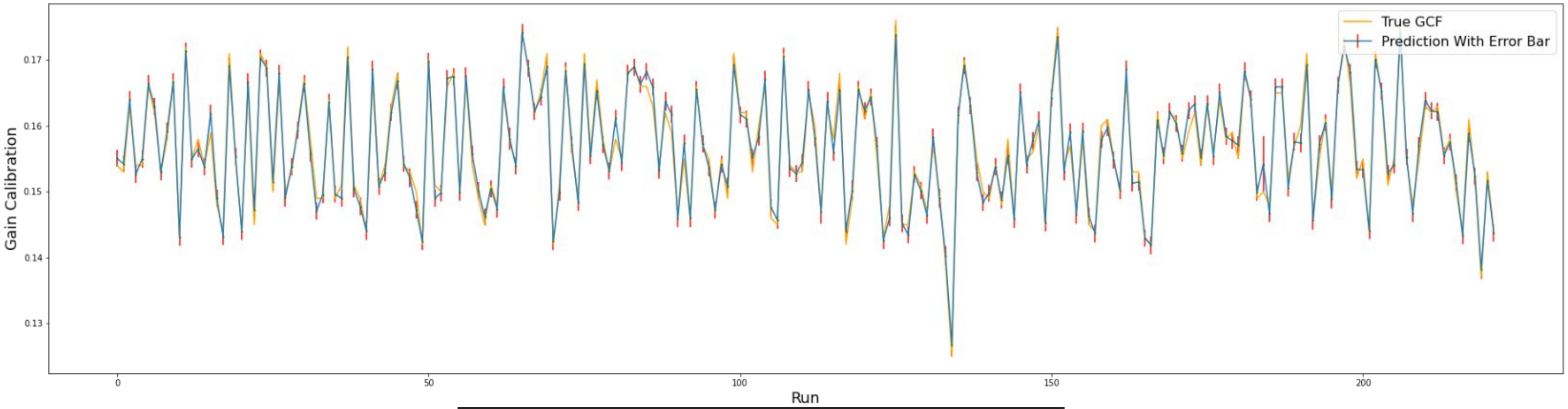
2020 run data (filtered)
- 430 training observations
- 106 testing observations

AI/ML methods applied:
NN,
Random Forest,
XGBoost,
GPR

Gaussian Process Regression
- Suited to small data set
- Provides uncertainty quantification



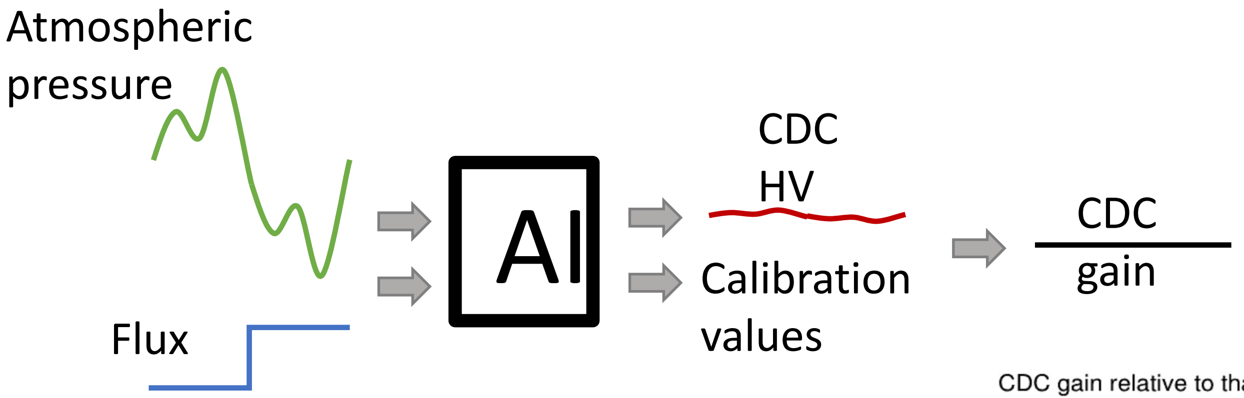
CDC gains results




```
Predictions on 222 Training and Test Data Sets
_______________________________________________
Minimum Different Between Truth and Pred: 0.000005
Maximum Different Between Truth and Pred: 0.005996
Mean Different Between Truth and Pred: 0.001026
Minimum Perc Dif: 0.000029
Maximum Perc Dif: 0.039976
Mean Perc Dif: 0.006556
Mean GCF: 0.15697727272727272
------------------------------
IF WE JUST USED THE MEAN GAIN INSTEAD OF PREDICTION?:
Minimum Perc Dif Between Truth and Mean Truth: 0.000145
Maximum Perc Dif Between Truth and Mean Truth: 0.145820
Mean Perc Dif Between Truth and Mean Truth: 0.044877
```

AI solution better than just using the mean gain

HV controls + gains

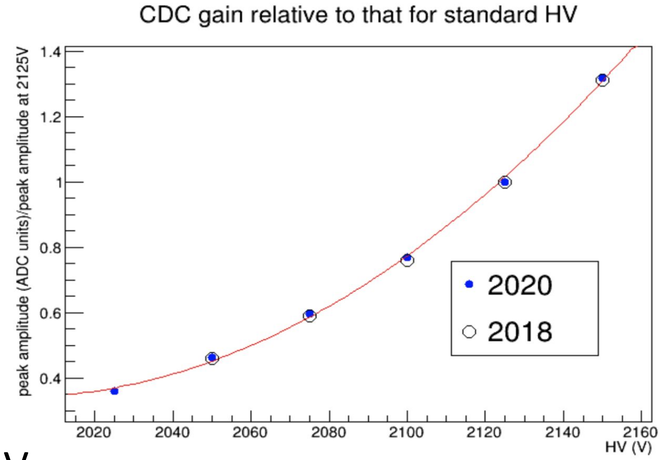


- AI predicts Gain Correction Factor (GCF) for 2125V

- Ask AI for ideal GCF, at std pressure (101.3 kPa)
- Ask AI for expected GCF at pressure right now

- Calculate relative change in gain needed

- Use known behaviour of relative gain vs HV to find desired HV



time-to-distance

- Current calibration method produces 6 unique calibration constants from fit to data

$$d(t) = f_\delta \left(\frac{d_0(t)}{f_0} P + 1 - P \right)$$

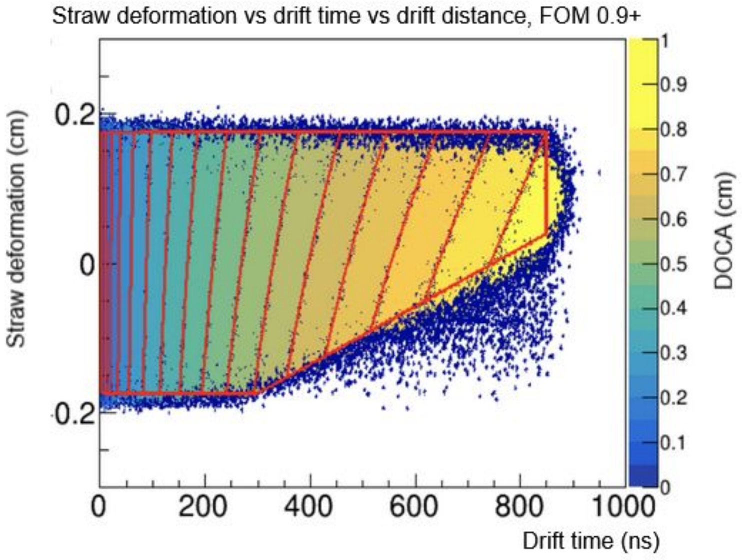
$$f_\delta = a\sqrt{t} + bt + ct^3$$

$$f_0 = a_1\sqrt{t} + b_1 t + c_1 t^3$$

$$a = a_1 + a_2|\delta|$$

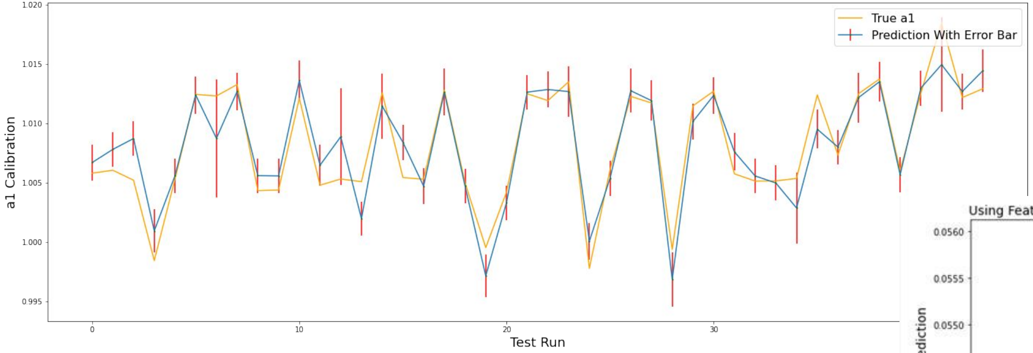
$$b = b_1 + b_2|\delta|$$

$$c = c_1 + c_2|\delta|$$



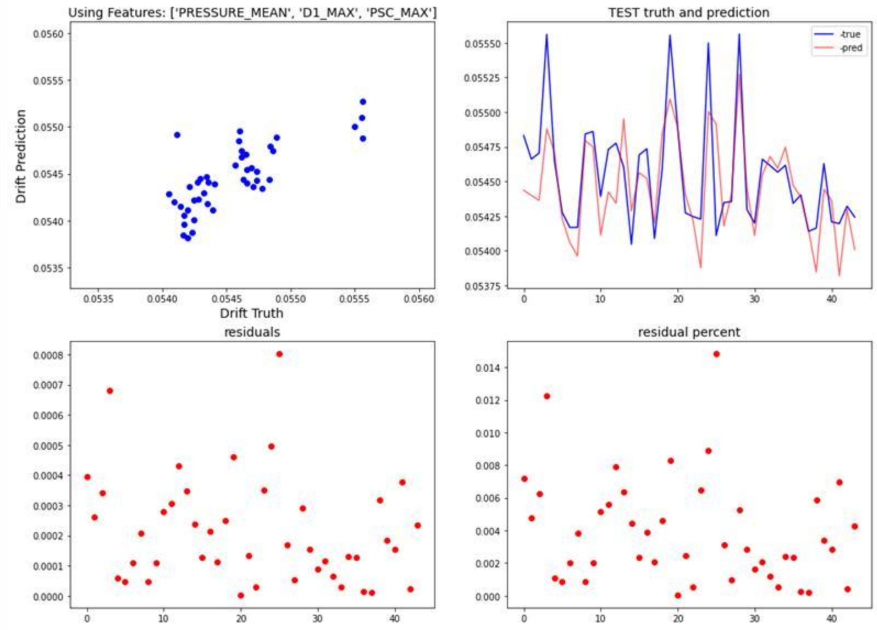
Model development for calibration constants is in very early stages

Early results



A1 has the biggest effect so concentrating efforts here

First results look promising



Future work

- Calibrate 2018 and 2020 runs with both predicted gain correction factor and time-to-distance calibrations
  - evaluate IF and "how many" iterations of traditional calibration are needed to equate to the AI’s calibration.
- Evaluate the data collected in 2021 (where AI was setting the voltage)
  - determine if stabilization of gain (through HV control) is improving the stability of dE/dx and thus improving things like PID
  - Integrate the AI-recommended voltage into existing control software for ease of expert evaluation and use.
- Apply Gain and Time-To-Distance AI to other detectors, i.e. CLAS12

Conclusion

- Promising early results:
  - Very preliminary results show we can decrease time to calibrate
  - Have taken parasitic data in which the AI “controls” the CDC HV.
    - Will analyze for gain stability

- There is a need to explore physics based metrics for success
  - Have seen the result of beam trips affecting the “ground truth” and thus the model