

A.I. Assisted Experiment Control and Calibration

Thomas Britton^ξ, Michael Goodrich^ξ,
Naomi Jarvis*, Torri Jeske^ξ, Nikhil Kalra^ξ,
David Lawrence^ξ, Diana McSpadden^ξ,
Kishansingh Rajput^ξ

^ξThomas Jefferson National Accelerator Facility
*Carnegie Mellon University

Summary of expenditures by fiscal year (FY)*:

	FY20 (\$k)	FY21 (\$k)	FY22 (\$k)	Totals (\$k)
a.) Funds allocated	270	270	270	810
b.) Actual Costs to date	270	270	270	810

total award for 3 years: \$810k

- Mostly labor cost

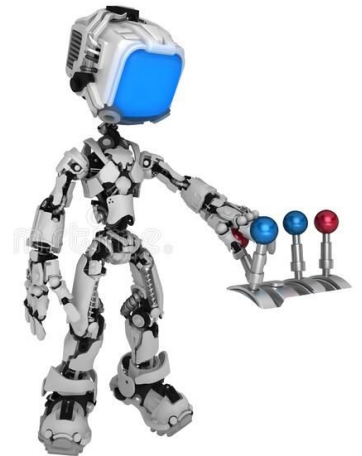
**n.b. funds come at end of fiscal year so are actually spent during the following fiscal year*

Motivation

- Sensitive detectors need to be calibrated to obtain optimal resolution
- Calibrations cause a delay between data collection and analysis (weeks-months)
 - Multiple iterations are needed to converge to final set of constants

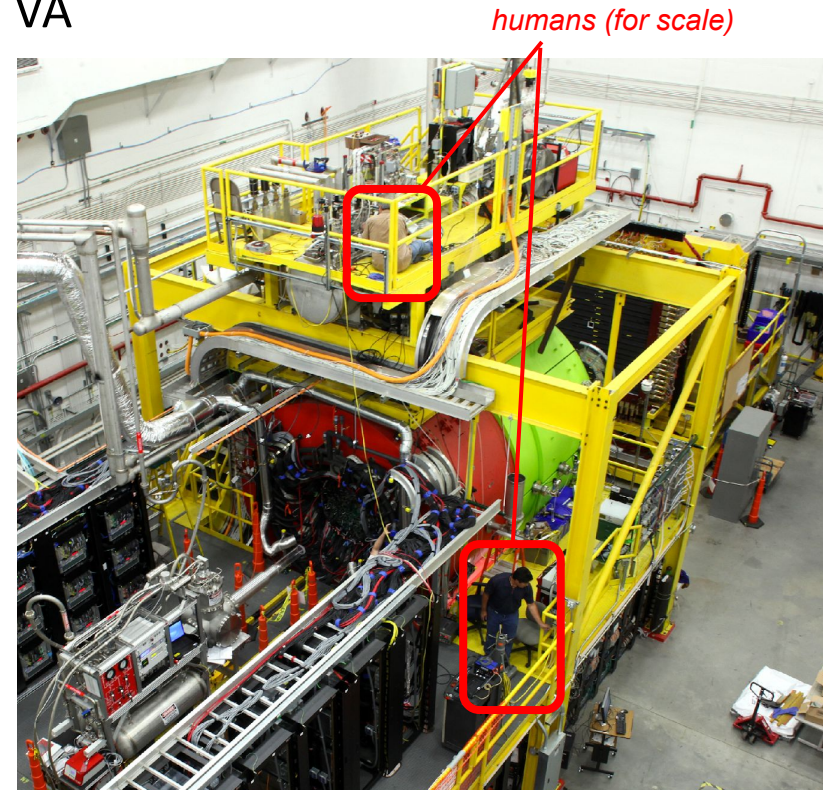
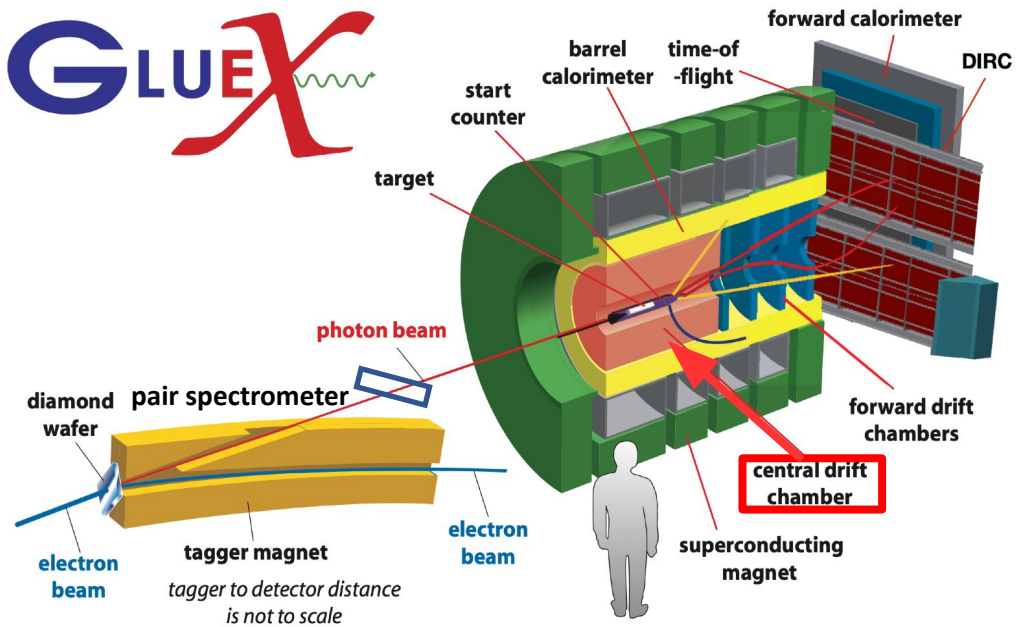
Main Goal:

Dynamically adjust the controls of a sensitive detector to reduce or eliminate the need for calibration



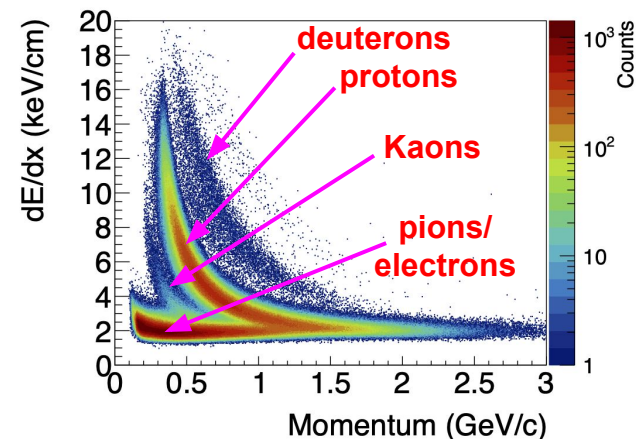
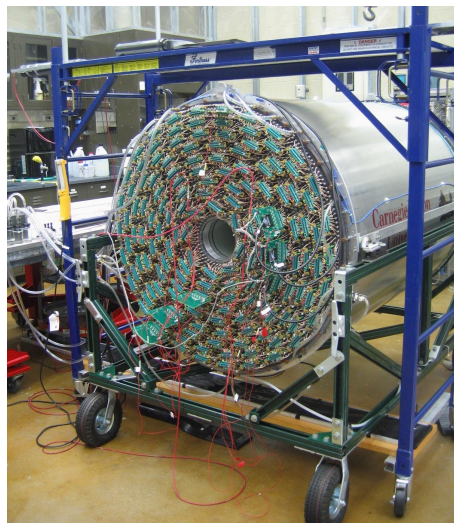
The GlueX Detector

GlueX detector located in Hall D at Jefferson Lab, VA



The CDC (= Central Drift Chamber)

- 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

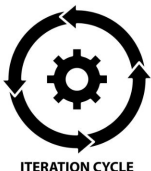


Conventional Calibration and Motivation for ML

Motivation: Conventional vs. Online, ML Calibration Paradigms

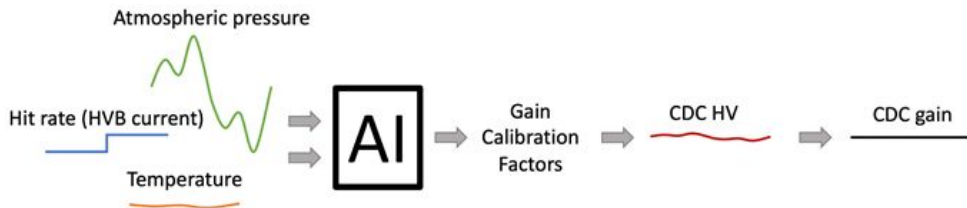
Conventional

- **Calibrate:** calibration values **iteratively**, produced after the experiment
 - ~2 hour runs
- **Control:** CDC operating voltage is **fixed** at 2125 V

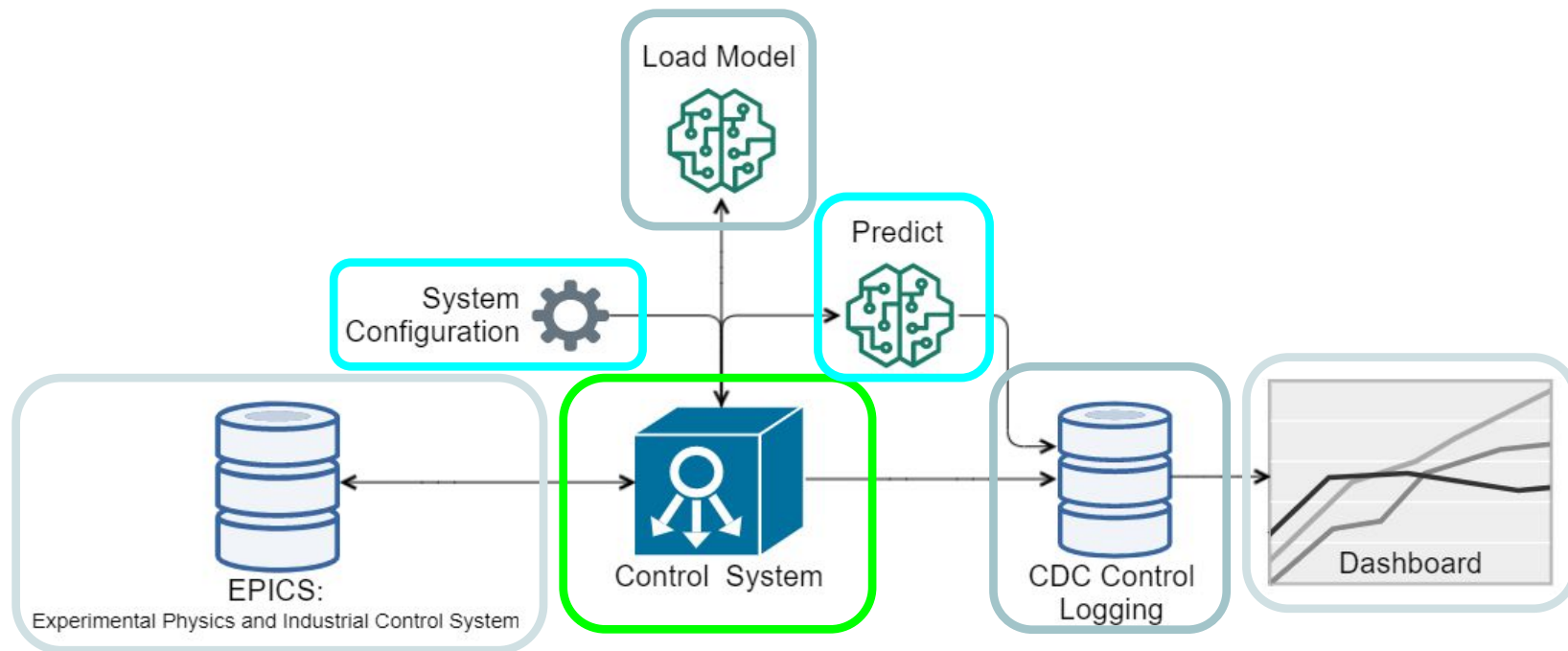




Online and ML

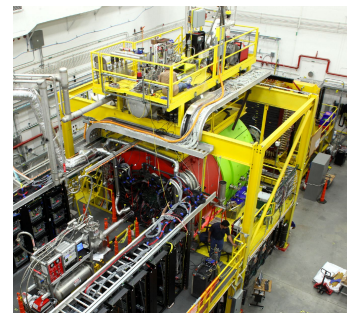
- **Control:** Stabilize detector response to changing environmental/experimental conditions by **adjusting** CDC HV
- **Calibrate:** **online** calibration values produced during the experiment



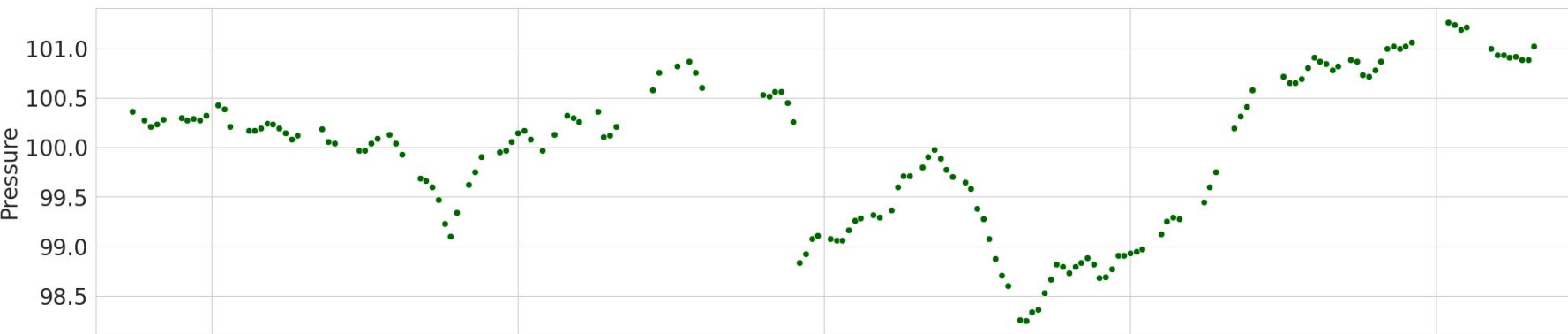
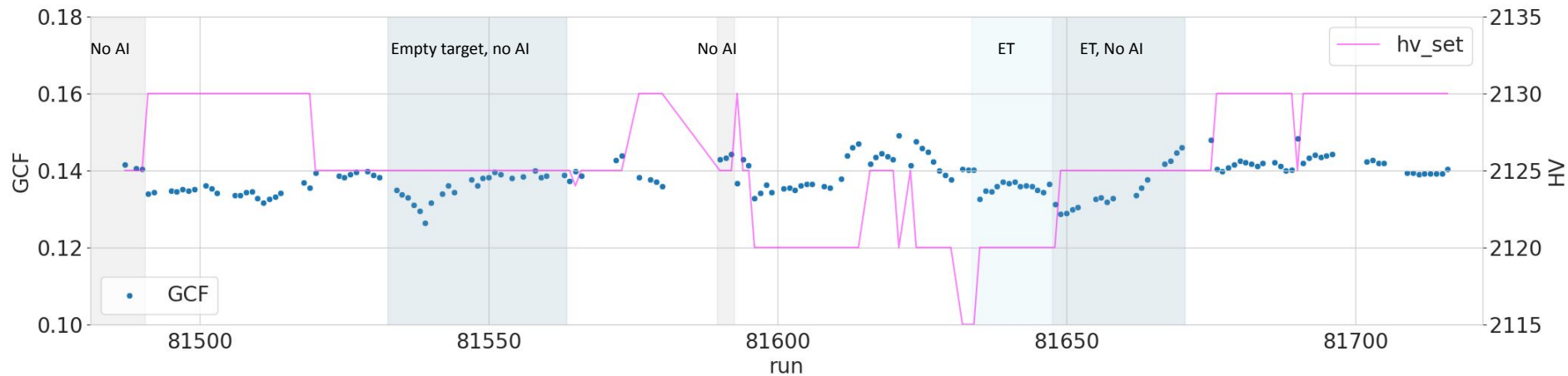
Integrating AI/ML into Control System



- Mid-October to early November 2021
- PrimEx- η running with GlueX Detector in Hall-D
 - Run plan was to have small amount of data with Solenoid on but most with it off
- Planned to test AI system over 2 days when solenoid was on
- Background levels were improved significantly with solenoid on
 - PI's changed plan and ran with it on for ~2weeks 
 - Atmospheric pressure did not change as much as we wanted 

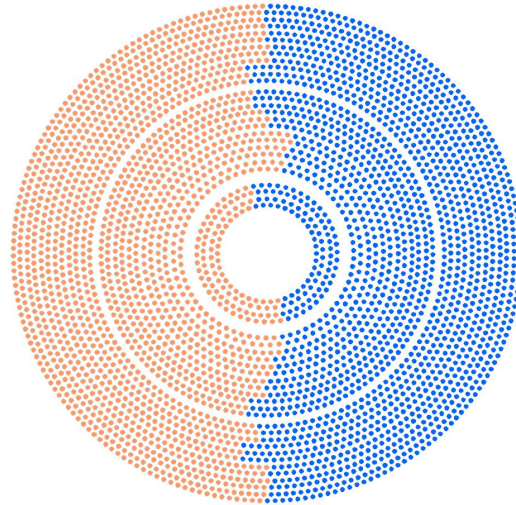
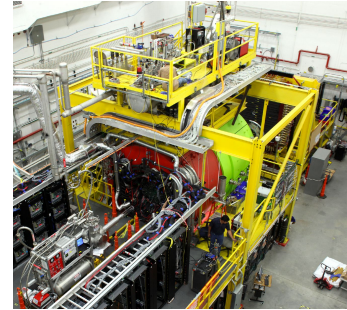


Gain correction factors from conventional calibrations

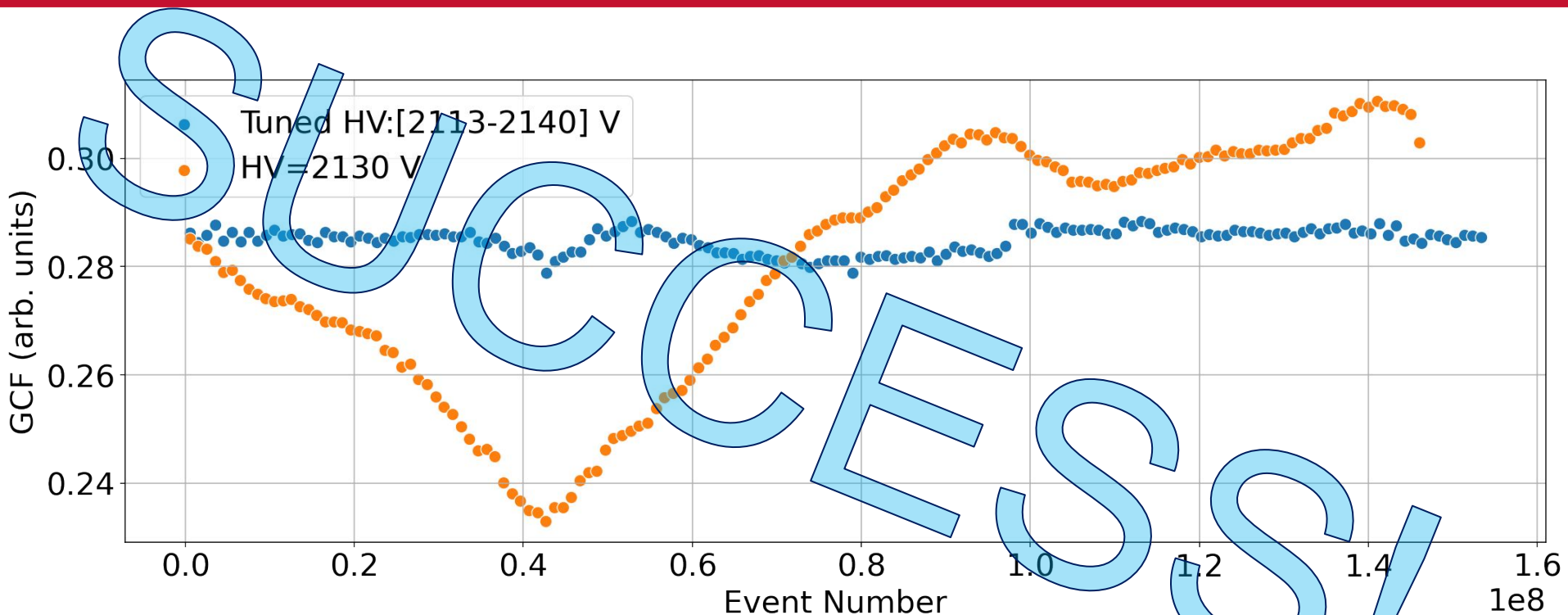


Automation Test with Cosmic Rays

- Two weeks in March 2022
- Half of sense wires controlled by AI/ML. Other half used fixed HV
- Fully automated with AI/ML adjustments every 5 minutes
- No beam. Cosmics only.



Cosmics Test Results



Conventional in orange
ML-tuned in blue

The Gaussian process model

ML Technique

Gaussian Process (GP)

- 3 features:
 - **atmospheric pressure** within the hall
 - **Gas temperature** within CDC
 - **CDC high voltage board current** a measure of luminosity
- 601 runs from 2020 and 2021 run periods
 - Pressure balanced for low, medium and high pressure
 - 80 / 20 train test split
- **1 target:** the traditional Gain Correction Factor (**GCF**)
- GP calculates PDF over admissible functions that fit data
- GP provides the standard deviation
 - we can exploit for uncertainty quantification (UQ)
- GP kernel:
 - Radial Basis Function + White noise
 - Compared isotropic (1 length scale) and anisotropic (length scale per input variable) kernels

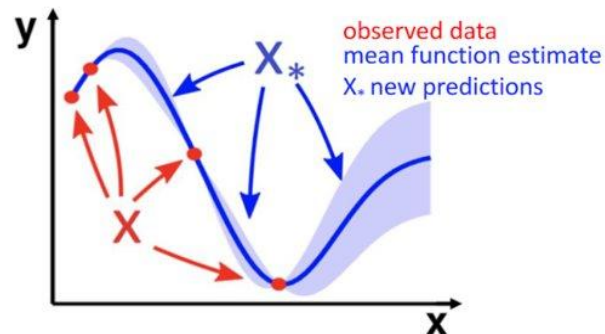


Illustration training a Gaussian process

Our goal was better than a 5% error

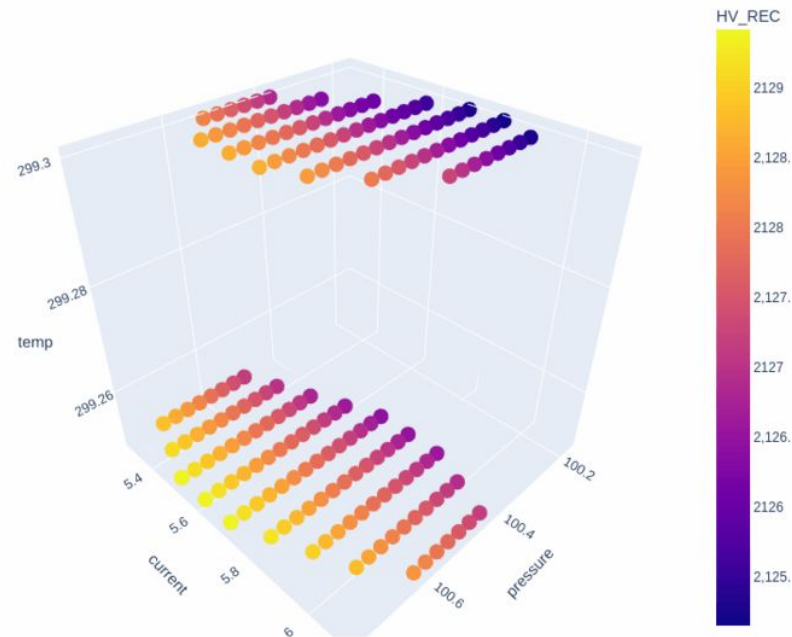
RBF kernel (length scale(s))		RMSE	Mean % err
Isotropic (1.412)	0.97	0.002	0.8%
Anisotropic (1.4,1.17,.171)	0.97	0.002	0.8%

Uncertainty Quantification

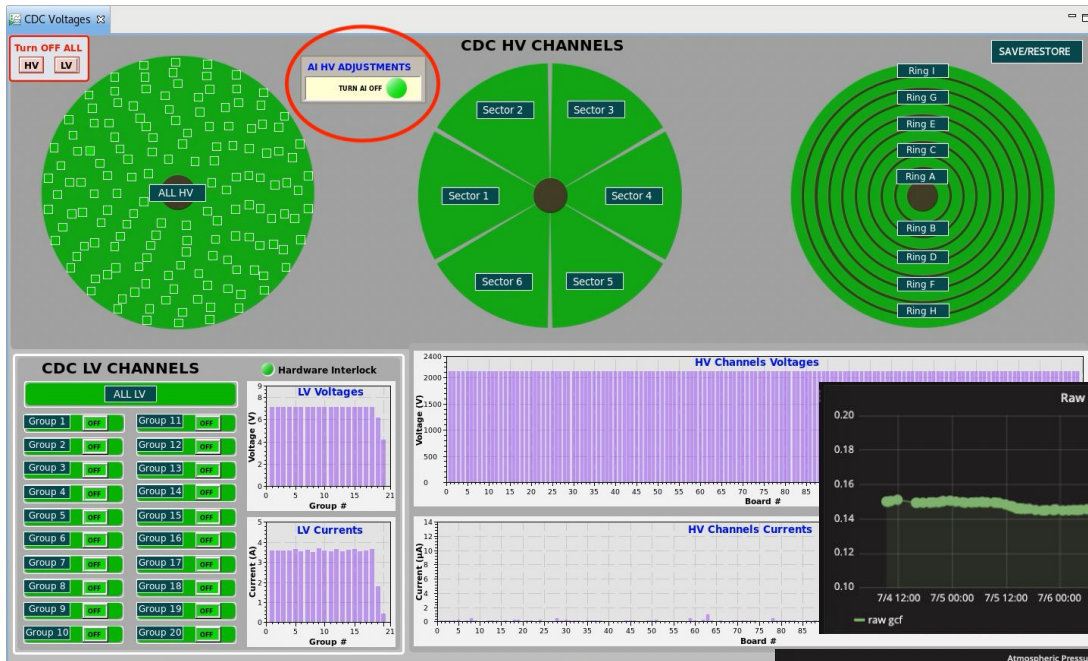
We created a system to automate the learning process as environmental and experimental conditions change:

1. A system that knows when it is **certain** and **controls** the experiment
2. Says “I don’t know” when **uncertain**, and **collects** more data and “**learns**”
3. Online retraining, evaluation of retrained model... (*future*)
4. Implement the retrained model that should be certain for more conditions

Threshold $\geq .7\%$

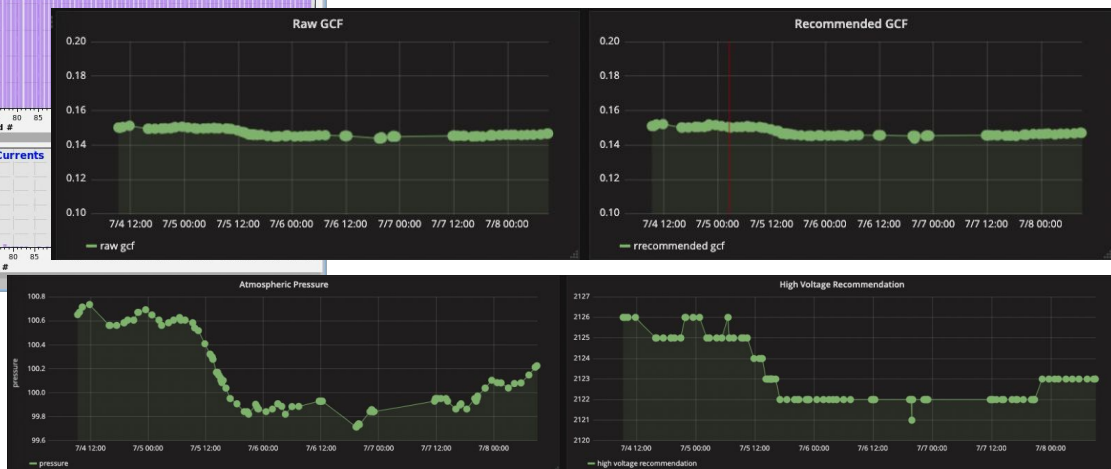


Integrating AI/ML into Standard Operations



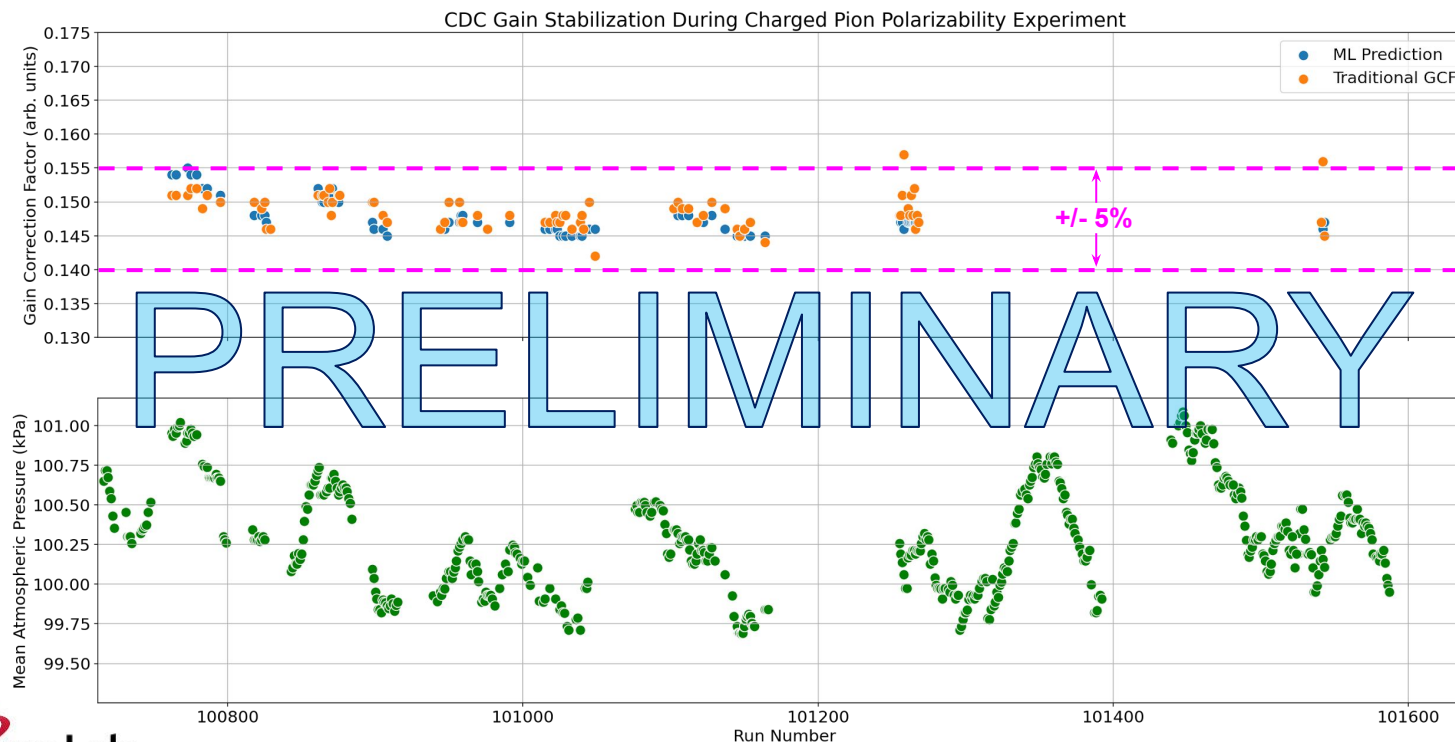
A switch was added to CDC Control GUI to allow shift takers to disable the AI/ML control completely.

Monitoring of the entire system was put onto a Grafana server.



Fully Automated System Deployed - *RoboCDC*

- Charged Pion Polarizability (CPP) – Spring 2022
 - Used at the start of each run in the experiment



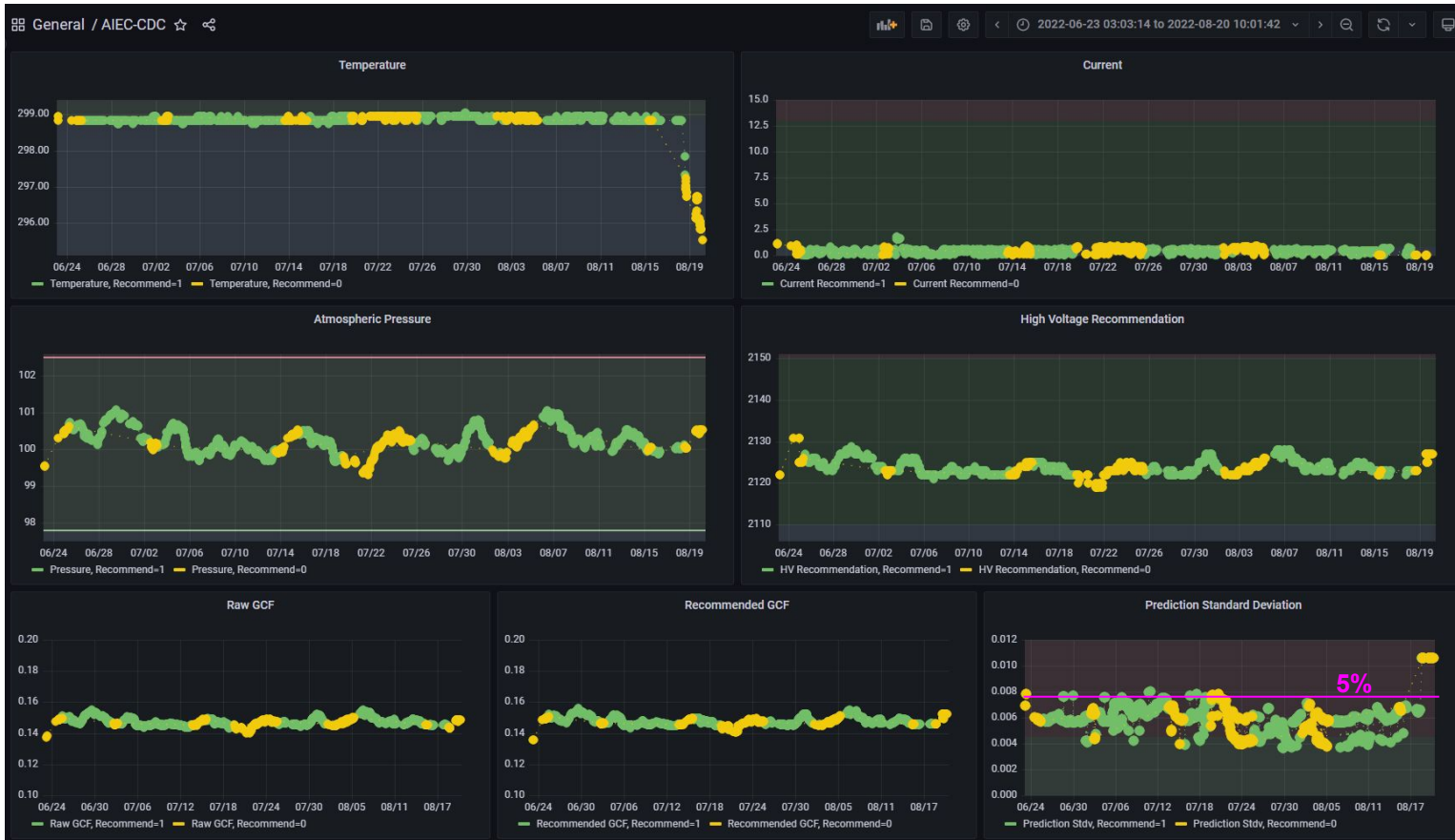
Fully deployed system during CPP Experiment in Spring 2022

green = AI controlled

yellow = fixed HV

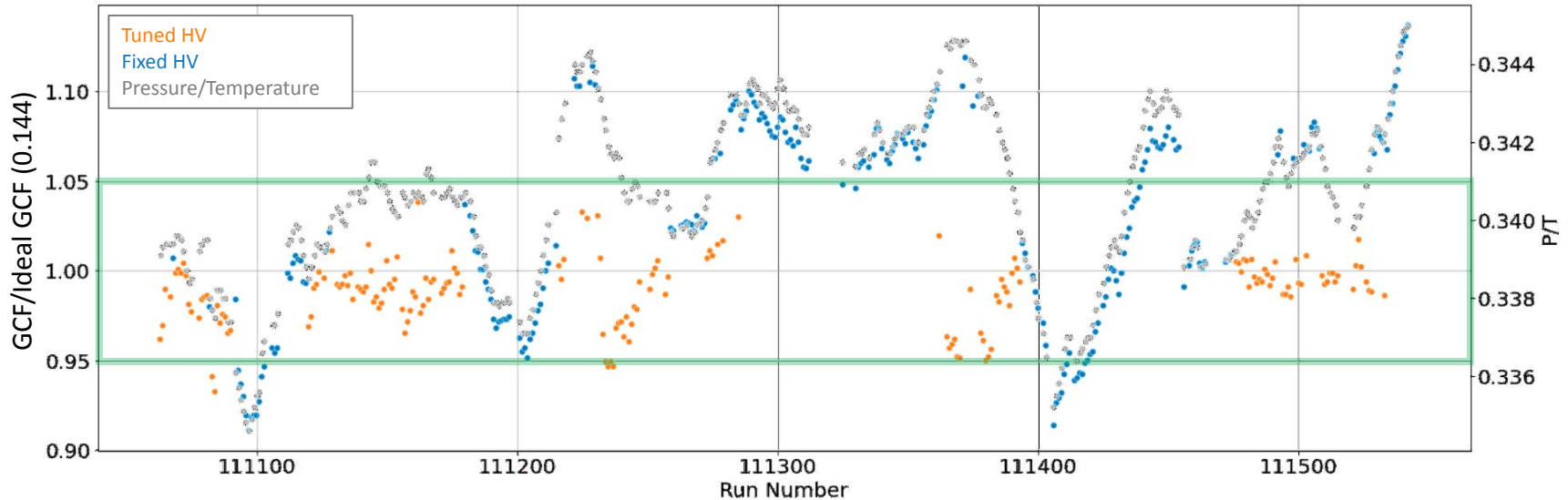
The CPP experiment setup was different from that for our training data set.

This required the system to drop into observation mode a number of times

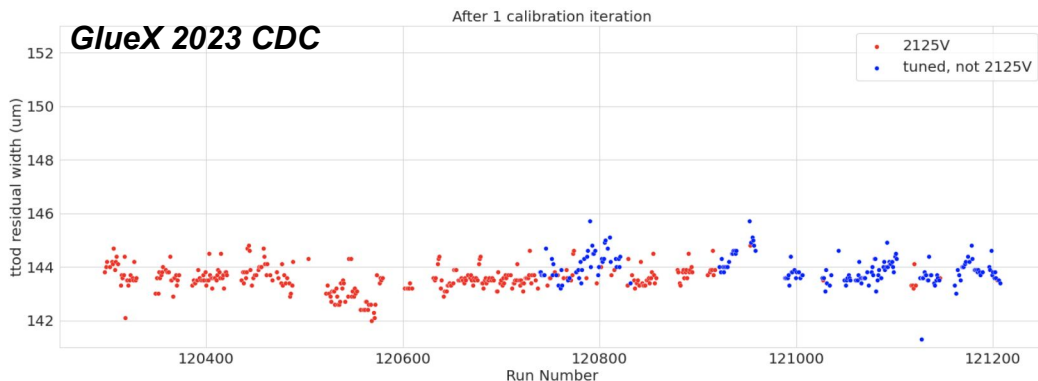
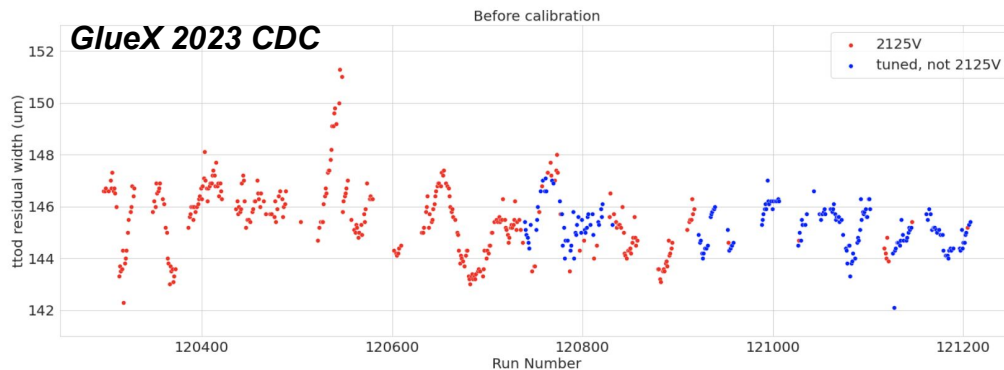


Deployment 3 – PrimEx- η June-Dec 2022

- GCF obtained from dE/dx after the run
- Preliminary results show GCF predominantly within 5% of ideal value for runs with tuned HV
- Plot of GCF/ideal for **tuned HV** and **fixed HV** also shows pressure/temperature



GlueX CDC Time to Distance Calibration



Tuning HV to stabilize gains results in comparable or better time-to-distance resolution prior to calibrating*.

**addresses concern noted in the proposal*

Bonus: Calibration technique for CDC Time-to-Distance developed as result of AIEC effort led to single iteration procedure. This new procedure replaced one entailing a series of 3 to 4 iterations of track reconstruction and re-fitting, and gave better and more consistent resolution.

Part II:

With the successful deployment of the automated AI control of the CDC, attention was turned to another detector system as outlined in the proposal.

The first candidate was the CLAS12 drift chambers.

Differences in Collaboration Culture?

```
Terminal - mysql -h clasdb.jlab.org -u clas12reader - 136x66
mysql> SELECT assignments.created AS created,count(typeTables.id) AS Nentries,typeTables.name,directories.name AS directory from assignm
ents,constanSets,typeTables,directories WHERE assignments.constanSetId=constanSets.id AND constanSets.constanTypeId=typeTables.id A
ND directories.directoryId=typeTables.id GROUP BY typeTables.id ORDER BY Nentries DESC LIMIT 100;
```

created	Nentries	name	directory
2018-01-05 12:11:18	4073	offset	
2018-07-07 13:22:16	1219	time2dist	time_to_distance
2018-07-05 15:46:31	848	fcup	runcontrol
2018-04-23 13:54:44	439	alignment	dc
2018-07-14 12:27:01	300	time_offsets	FTOF
2018-07-14 12:27:02	248	tres	FTOF
2018-08-15 19:33:20	222	hwp	runcontrol
2017-08-30 12:34:30	225	time	FTOF
2017-08-30 12:34:30	222	time	htcc
2016-04-08 00:19:29	208	region	dc
2019-04-25 12:38:22	208	superlayer	dc
2018-07-03 04:09:40	193	TimeOffsets_Layer	cnd
2018-07-01 10:41:59	184	time_offsets	ctof
2016-02-08 19:20:24	181	attenuation	ftof
2016-02-08 17:42:54	179	gain_balance	ftof
2016-02-08 19:40:55	176	effective_velocity	ftof
2016-02-08 19:34:20	173	status	FTOF
2019-04-22 09:09:42	169	time_walk_pos	ftof
2018-06-26 04:31:40	160	TurnLoss	dc
2018-07-03 04:10:48	159	TimeOffsets_LR	cnd
2018-07-04 05:33:29	159	EFV	cnd
2017-07-13 06:16:59	137	Attenuation	cnd
2016-03-07 11:09:13	136	attenuation	ctof
2017-07-13 11:20:56	133	Energy	ctof
2016-03-07 11:09:18	133	effective_velocity	ctof
2016-03-07 11:09:26	132	status	ctof
2016-03-07 11:09:22	132	gain_balance	ctof
2019-12-27 09:49:58	122	time_walk_exp	ftof
2020-01-14 11:09:10	121	tres	ctof
2017-08-30 05:55:36	118	time_offsets	ftcal
2020-07-06 18:36:00	101	time_parms	rtpc
2019-04-23 16:19:21	99	hps	ctof
2021-09-10 16:41:41	98	position	beam
2018-06-26 17:26:06	92	TBCorrections	l_time_corrections
2016-12-07 04:34:52	90	charge_to_energy	l_fthodo
2016-04-18 16:26:44	90	gain	hcc
2019-04-16 09:16:30	88	time_walk	ftcal
2016-11-23 08:17:03	86	charge_to_energy	ftcal
2019-05-21 15:13:02	77	misalignments	rich
2018-06-11 19:11:20	73	photon_sf	eb
2016-03-17 11:06:34	71	time_walk	ftof
2016-11-25 12:28:56	69	status	l_fthodo
2020-07-07 11:09:21	60	recon_parms	rtpc
2018-02-27 10:57:23	49	spe	l_tccc
2016-03-31 23:03:54	46	dc_resolution	l_signal_generation
2019-06-11 19:12:20	41	electron_sf	sb
2019-05-16 11:40:26	37	time_offset	rich
2019-04-19 10:43:47	36	ministagger	dc
2018-06-12 11:17:00	35	fadc_offset	ftof
2018-06-03 08:47:36	35	timing	ec
2020-10-01 20:07:36	34	slr	runcontrol
2019-05-14 11:30:41	34	time_walk	rich
2016-11-25 12:31:28	33	time_offsets	l_fthodo
2020-02-05 09:06:12	28	gain_balance	rtpc
2018-06-23 16:47:37	23	fadc_offset	ctof
2018-01-05 12:05:26	20	config	rf
2018-02-03 19:25:00	20	drift_fullfield	bmt_mv
2019-07-26 16:14:37	20	endplatesbow	dc
2017-02-18 18:18:53	19	attenuation	ec
2017-02-19 15:12:16	18	gain	ec
2018-05-07 16:50:45	14	wire_status	tracking

CLAS12

Total CCDB assignments (All years):
13,740

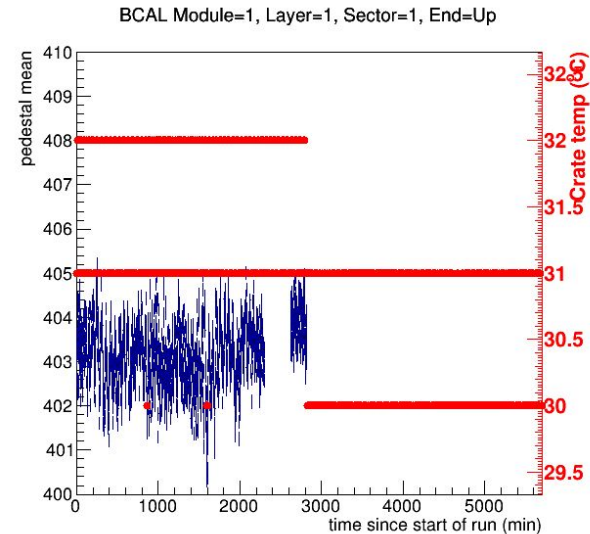
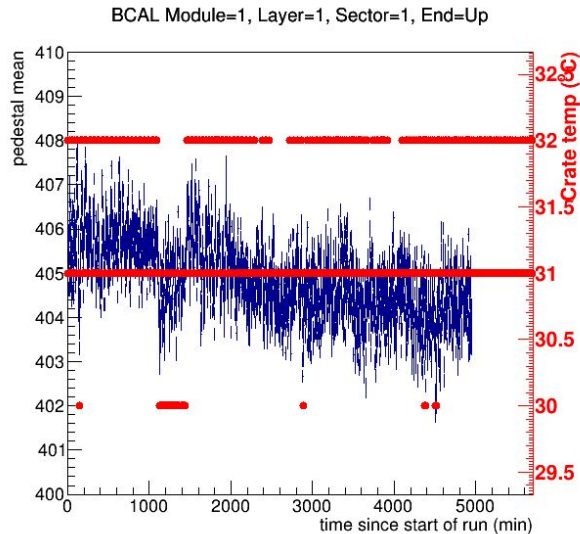
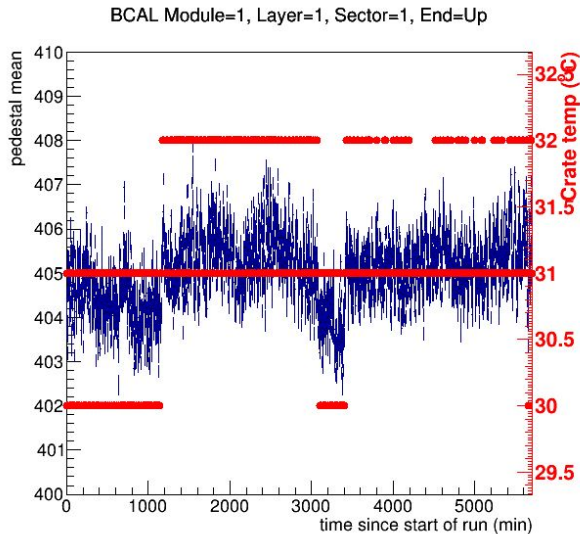
```
Terminal - mysql -h halldb.jlab.org -u ccdb_user ccdb - 136x66
mysql> SELECT assignments.created AS created,count(typeTables.id) AS Nentries,typeTables.name,directories.name AS directory from assignm
ents,constanSets,typeTables,directories WHERE assignments.constanSetId=constanSets.id AND constanSets.constanTypeId=typeTables.id A
ND directories.directoryId=typeTables.id GROUP BY typeTables.id ORDER BY Nentries DESC LIMIT 100;
```

created	Nentries	name	directory
2014-08-04 18:44:44	24994	tdc_time_offsets	microscope
2014-08-04 18:04:29	24192	tdc_time_offsets	microscope
2014-08-04 18:02:19	19156	fadc_time_offsets	microscope
2015-11-12 18:16:10	15961	tdc_timewalk	hodoscope
2014-09-17 22:23:45	15927	base_time_offset	START_COUNTER
2014-08-04 18:42:46	14780	base_time_offset	hodoscope
2014-09-17 22:23:52	14278	base_time_offset	hodoscope
2014-08-15 16:22:49	13662	digit_scales	DCD
2014-09-25 21:48:58	13068	base_time_offset	FCAL
2014-09-17 22:23:59	13023	base_time_offset	microscope
2014-11-07 23:40:24	12847	base_time_offset	psip_spectrom
2014-09-17 22:23:31	12438	base_time_offset	FDC
2014-09-17 22:23:24	10929	base_time_offset	BCAL
2014-09-17 22:23:17	10193	base_time_offset	DCD
2014-09-17 22:23:38	10016	base_time_offset	TOF
2015-10-15 14:41:28	9875	drift_parameters	CCD
2014-11-07 23:40:12	9622	adc_timing_offsets	fine
2015-12-12 22:04:31	9554	tdc_timewalk_corrections	microscope
2019-10-22 23:25:48	8283	accidental_scaling_factor	ANALYSIS
2015-05-14 21:42:13	6726	time_offset	RF
2014-11-07 23:39:54	6649	adc_timing_offsets	CCD
2017-02-25 20:09:00	6651	current_map_epics	ELECTRON_BEAM
2014-11-07 23:39:57	6649	adc_timing_offsets	coarse
2014-07-01 06:32:09	6206	endpoint_energy	PHOTON_BEAM
2014-05-16 16:04:33	5971	tdc_timing_offsets	START_COUNTER
2014-05-16 16:04:36	5956	adc_timing_offsets	START_COUNTER
2019-10-21 16:41:29	5867	base_time_offset	TOF2
2015-05-14 21:42:17	5821	time_offset_var	RF
2017-05-22 12:53:05	4582	ADC_offsets	FCAL
2016-07-13 16:40:43	3706	block_mc_efficiency	FCAL
2014-05-22 17:39:26	3574	adc_timing_offsets	BCAL
2014-06-24 13:49:42	3518	wire_timing_offsets	package2
2014-06-24 13:49:48	3501	wire_timing_offsets	package3
2014-06-24 13:49:36	3479	wire_timing_offsets	package1
2014-06-24 13:49:54	3469	wire_timing_offsets	package4
2017-02-25 20:09:00	3315	time_stamp_to_units	ELECTRON_BEAM
2018-08-02 09:48:28	3118	gain_dcca_correction	CCD
2014-06-03 17:07:56	2979	adc_timing_offsets	TOF
2018-10-23 11:05:50	2782	tagged	tigh
2018-10-23 11:05:51	2782	untagged	tigh
2018-10-23 11:05:48	2782	tagged	tigh
2018-10-23 11:05:49	2782	untagged	tigh
2018-10-23 11:05:52	2648	trig_live	lumn
2019-09-10 14:02:27	2507	adc_timing_offsets	TOF2
2020-09-10 14:02:33	2494	propagation_speed	TOF2
2020-02-10 09:42:33	2490	timewalk_parms_SPAR	TOF2
2020-02-11 15:31:52	2489	walkcorr_type	TOF2
2020-02-11 19:35:06	2485	timing_offsets_SPAR	TOF2
2021-04-26 06:36:26	2468	ps_counters	ELECTRON_BEAM
2014-01-17 15:27:11	2118	block_quality	FCAL
2014-02-04 18:09:15	2072	pedestals	TOF
2014-01-17 15:27:09	1868	gains	FCAL
2014-05-22 17:39:25	1640	TDG_offsets	BCAL
2017-06-13 16:15:30	1369	timewalk_parms_NERAMP	TOF
2017-07-06 07:36:24	1349	timing_offsets_NERAMP	TOF
2017-07-06 07:36:25	1348	propagation_speed_NERAMP	TOF
2019-01-10 13:09:41	1309	base_time_offset	DIRC
2014-02-04 18:09:12	1244	timing_offsets	TOF
2018-11-25 15:44:32	1223	gdcs	CCAL
2014-12-07 20:42:12	1040	tdc_shift	TOF

GlueX

Total CCDB assignments (All years):
410,367

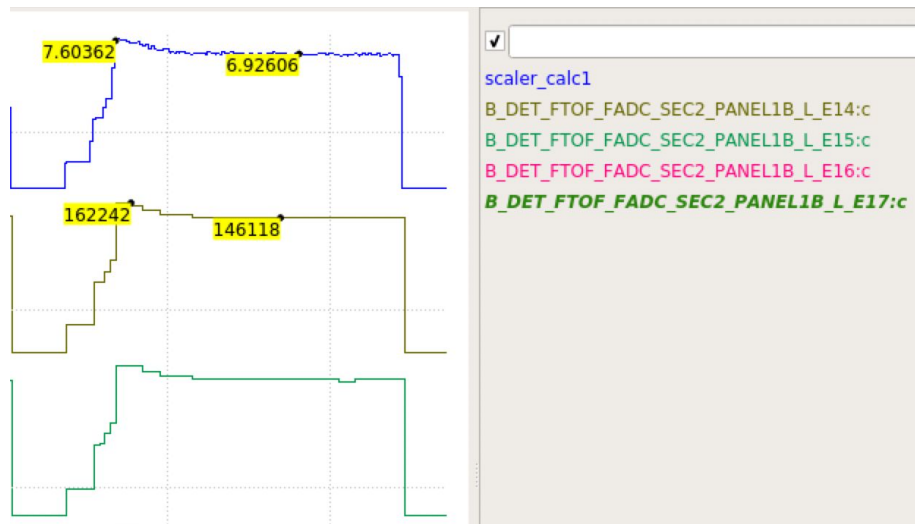
GlueX Barrel Calorimeter Pedestals



A colleague suggested to look at controlling fan speed to stabilize pedestals read by a flash ADC for a calorimeter.

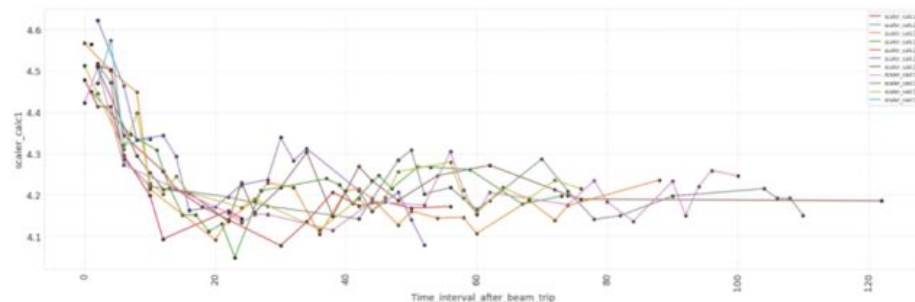
Rejected due to crate temp. measurements being too coarse.

CLAS12 Calorimeter and TOF detector gains after beam trip



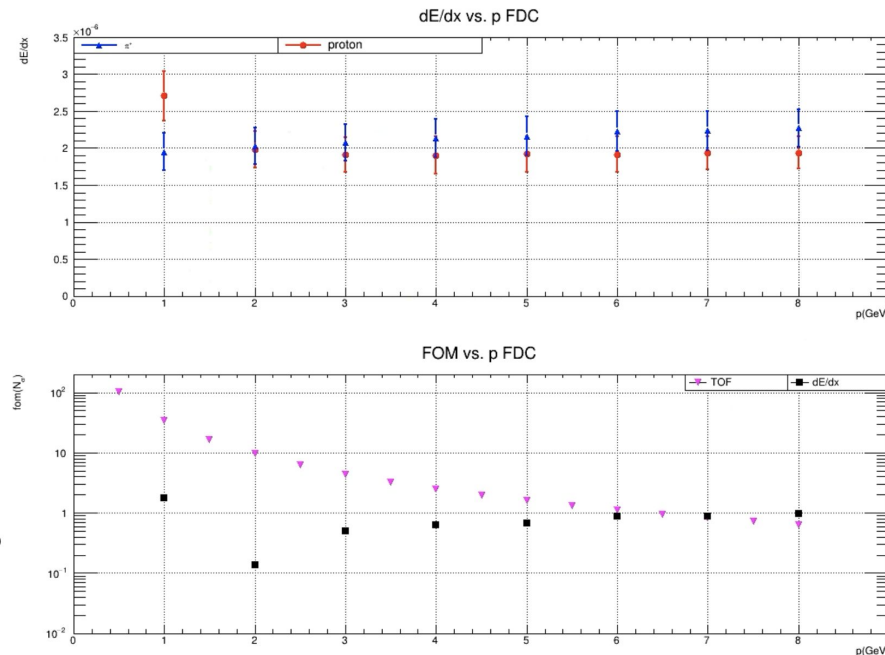
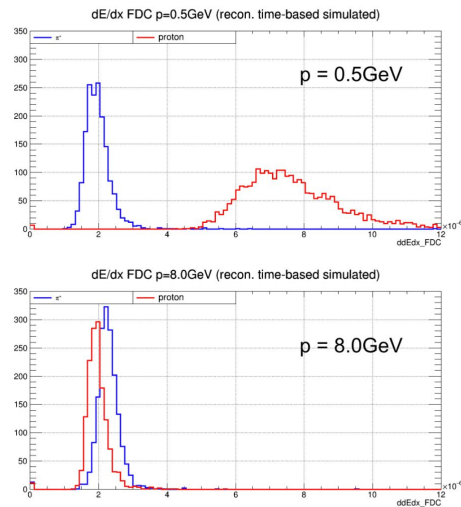
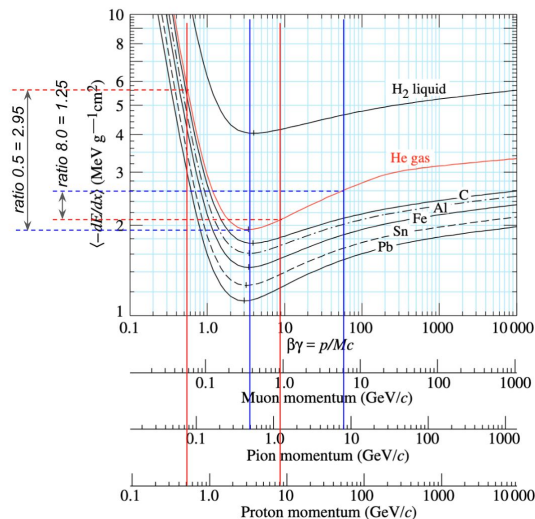
A colleague suggested to look at stability of gains in EM Calorimeter to just after beam trips. This could possibly benefit TOF PMTs as well.

Rejected due to observed effect being almost completely due to beam current overshoot upon recovery.



GlueX FDC PID from dE/dx

34. Passage of Particles Through Matter

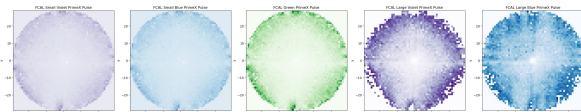


Unable to identify an approved experiment requiring good PID for particles > 5GeV

Traditional Calibration:

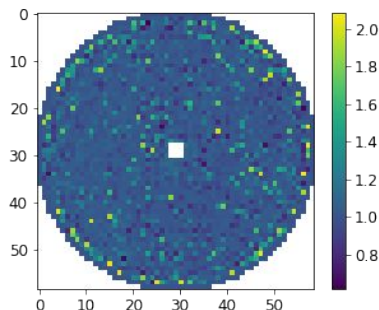
- iterative over π^0 s
- Requires particle reconstruction
- Statistics sometimes difficult

Can we use the LED monitoring system and Machine Learning?

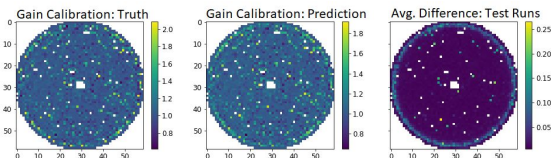


=

Gain calibration values

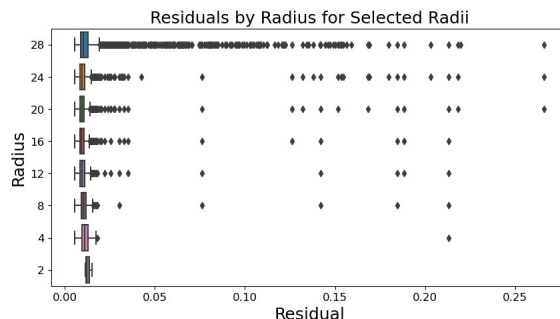


Can ML learn traditional calibrations?



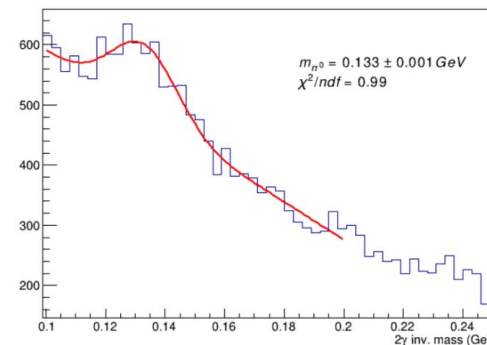
Average results over 5-fold cross validation

dataset	fold idx	average residual ↓	mape ↓	mse ↓
unmasked	average	0.258	23.848	5.183
masked	average	0.027	2.370	0.004



Initial Physics Comparison

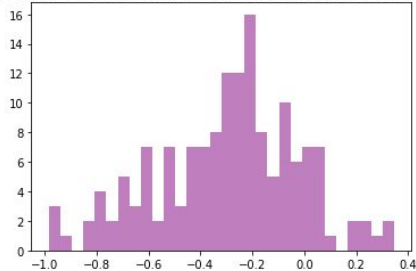
- Does prediction accuracy result in good physics results?
- We have an initial π^0 analysis
 - Single run, entire FCAL



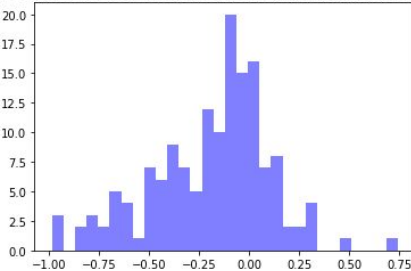
- π^0_{PDG} mass: **134.98 MeV**
- Using our calibrations: **133.31 MeV**

Correlation Plots

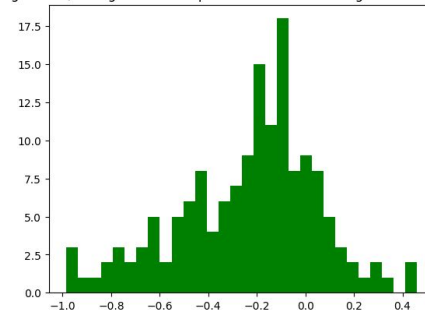
Q0 (Rings 10-17) histogram of stats.pearsonr correlation for gain and SV pulse: 30 Bins



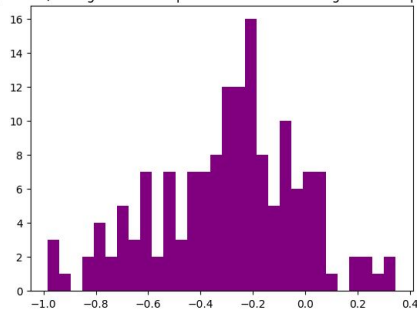
Q0 (Rings 10-17) histogram of stats.pearsonr correlation for gain and SB pulse: 30 Bins



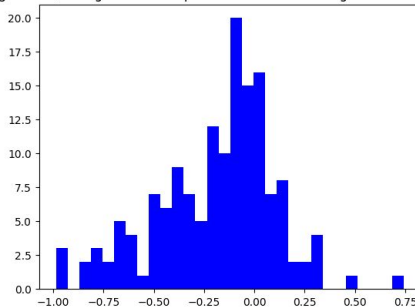
Q0 (Rings 10-17) histogram of stats.pearsonr correlation for gain and G pulse: 30 Bins



Q0 (Rings 10-17) histogram of stats.pearsonr correlation for gain and LV pulse: 30 Bins

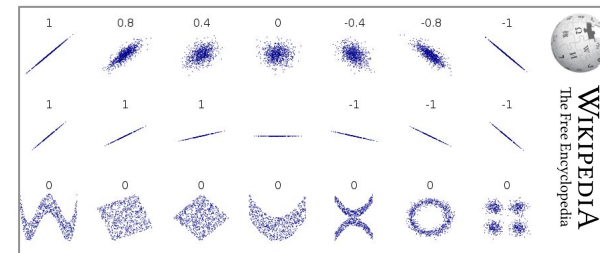


Q0 (Rings 10-17) histogram of stats.pearsonr correlation for gain and LB pulse: 30 Bins



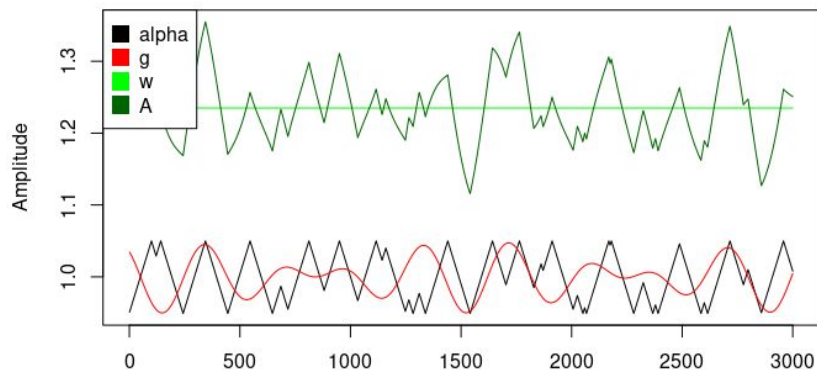
Pearson correlation coefficient for LED peaks and calibration gains for different colors

Expect anti-correlation (values <0) but scale is smaller than expected

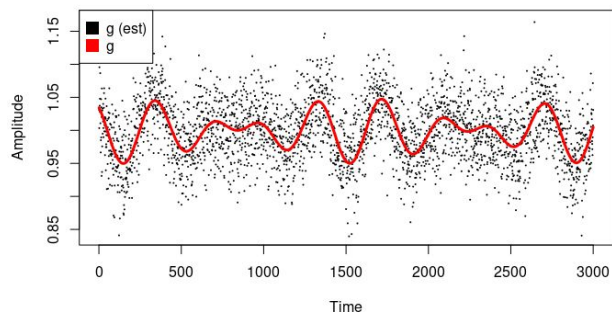


Simulation of the LED gain monitoring system

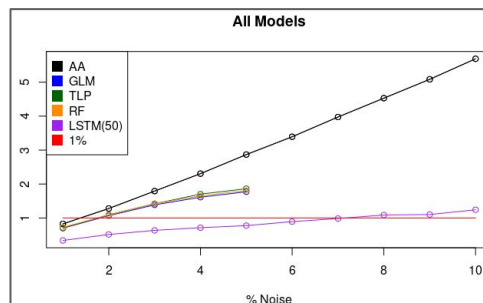
Gain Factors (0 pct noise)



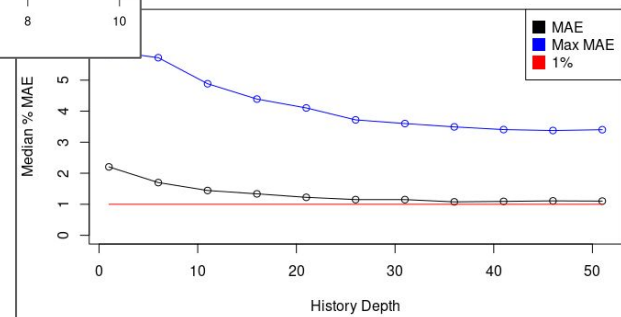
AA fit at 5 pct Noise



- A = ADC readout amplitude for each block and discrete time index.
- g = PMT gain for each block and discrete time index.
- ω = Optical Coupling constant for each block.
- α = Amplitude of LED pulsar for each discrete time index.
- R = Radiation Damage for each block and discrete time index.



LSTM: 9 percent Noise

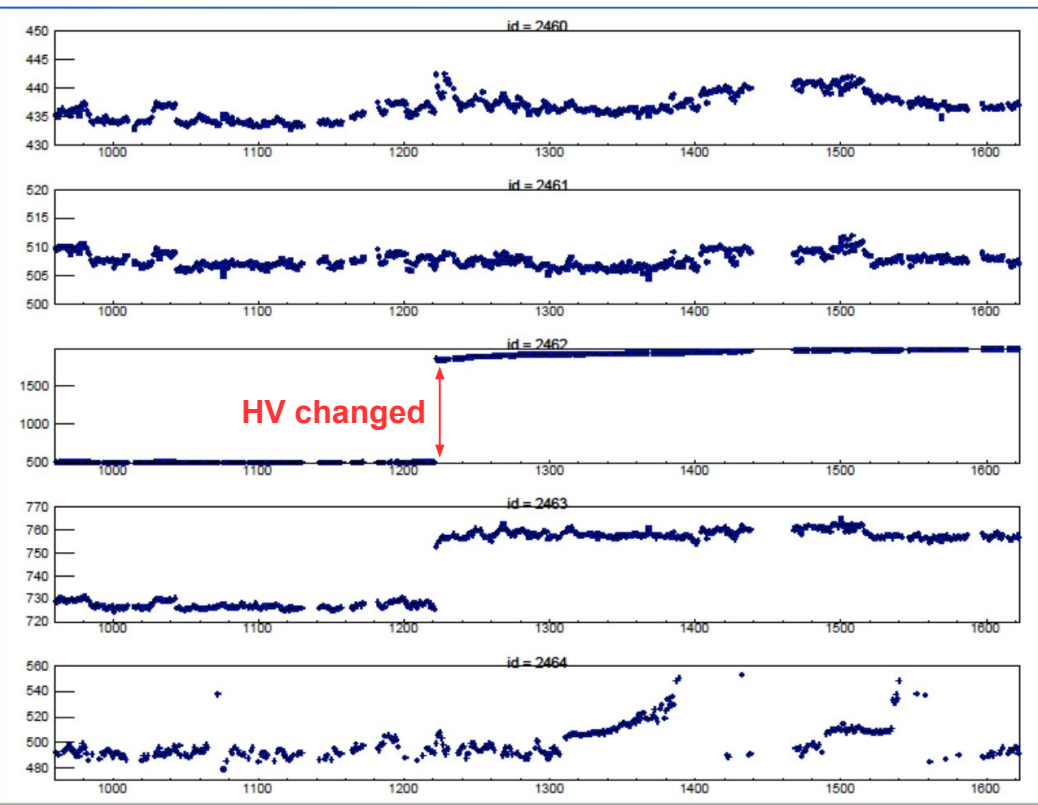


conclusion: AI can extract gains from this simulation model at ~1% level with significant measurement noise

Independent Analysis of LED gain monitoring system

“The clear correlation between led gain change or sometimes even jumps and π^0 energy gain has not been seen so far, which is puzzling”

Plots are for 5 different PMTs. LED amplitude as seen by PMT normalized to LED pulser amplitude as function of time.



Summary

- **Reproduced** calib. **constants** for GlueX CDC using **AI model** with same inputs as classic method
- Successfully **predicted** GCF calibrations using environmental data from GlueX 2018 and 2020 runs
- Successful **deployment** of AI detector control system (Gaussian Process model)
- Successful **deployment** of UQ aware system for CPP experiment in summer 2022
 - **Now part of standard operations!**
- Investigated GlueX FCAL LED monitoring system
 - UVA Capstone project for team of 4 DS students on automating Calorimeter
 - Created simulation. Extracted accurate calibrations using LSTM model
 - Unable to identify strong correlations in real data

Publications

Jeske, T., McSpadden, D., Kalra, N., Britton, T., Jarvis, N., & Lawrence, D. (2023, February). Using AI to predict calibration constants for the central drift chamber in GlueX at Jefferson Lab. In Journal of Physics: Conference Series (Vol. 2438, No. 1, p. 012132). IOP Publishing.

Jeske, T., McSpadden, D., Kalra, N., Britton, T., Jarvis, N., & Lawrence, D. (2022). AI for Experimental Controls at Jefferson Lab. Journal of Instrumentation, 17(03), C03043.

Key Presentations

AI Driven Experiment Calibration and Control [Slides](#)

Britton, T., Lawrence, D., Jeske, T., McSpadden, D., & Jarvis, N. (2023, May 8–12). AI Driven Experiment Calibration and Control [Conference presentation]. Conference on Computing in High Energy & Nuclear Physics, Norfolk, VA, United States. <https://indico.jlab.org/event/459/contributions/11374/>

Using Machine Learning to control the GlueX Central Drift Chamber [Slides](#)

Jarvis, N., Britton, T., Jeske, T., Lawrence, D., & McSpadden, D. (2023, January 9–13). Using Machine Learning to control the GlueX Central Drift Chamber [Conference presentation]. Conference on High Energy Physics in LHC Era, Valparaíso, Chile. <https://indico.cern.ch/event/1158681/contributions/5192980/>

Control and Calibration of GlueX Central Drift Chamber Using Gaussian Process Regression [Poster Paper](#)

McSpadden, D., Jeske, T., Jarvis, N., Britton, T., Lawrence, D., & Kalra, N. (2022, December 3). Control and Calibration of GlueX Central Drift Chamber Using Gaussian Process Regression [Poster Presentation]. Machine Learning and the Physical Sciences workshop at NeurIPS, New Orleans, LA, United States. <https://ml4physicalsciences.github.io/2022/>

Gaussian process for calibration and control of GlueX Central Drift Chamber [Slides](#)

Jeske, T., Britton, T., Kalra, N., Jarvis, N., McSpadden, D. & Lawrence, D. (2022, October 23-28). Gaussian process for calibration and control of GlueX Central Drift Chamber [Conference Presentation]. Advanced Computing and Analysis Techniques in Physics Research, Bari, Italy. <https://indico.cern.ch/event/1106990/contributions/4998092/>

Acknowledgements

This work was supported by the US DOE as LAB 20-2261.

The Carnegie Mellon Group is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, DOE Grant No. DE-FG02-87ER40315.

GlueX acknowledges the support of several funding agencies and computing facilities: www.gluex.org/thanks



Backup slides

HV Channel Segmentation (Prepping for Cosmics Test)

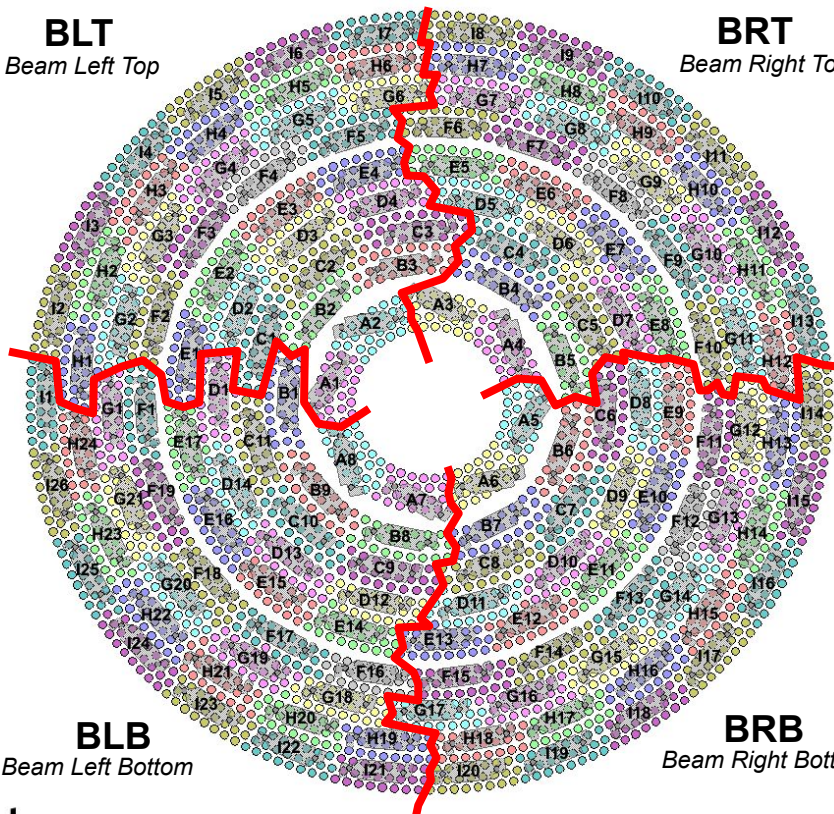
GlueX CDC

View from upstream

- A: 1-2
- B: 2-3
- C: 1-3
- D: 2-4
- E: 1-4
- F: 2-5
- G: 2-6
- H: 1-6
- I: 2-7

BLT
Beam Left Top

BRT
Beam Right Top



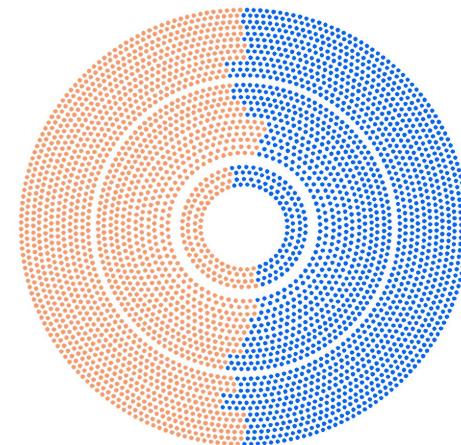
- A: 7-8
- B: 1,8-9
- C: 9-11
- D: 1, 12-14
- E: 14-17
- F: 1,16-19
- G: 1,18-21
- H: 19-24
- I: 1,21-26

BLB
Beam Left Bottom

BRB
Beam Right Bottom

- A: 3-4
- B: 4-5
- C: 4-5
- D: 5-7
- E: 5-8
- F: 6-10
- G: 7-11
- H: 7-12
- I: 8-13

- A: 5-6
- B: 6-7
- C: 6-8
- D: 8-11
- E: 9-13
- F: 11-15
- G: 12-17
- H: 13-18
- I: 14-20



Split the CDC into 2 halves

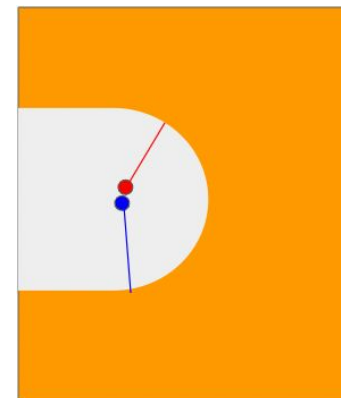
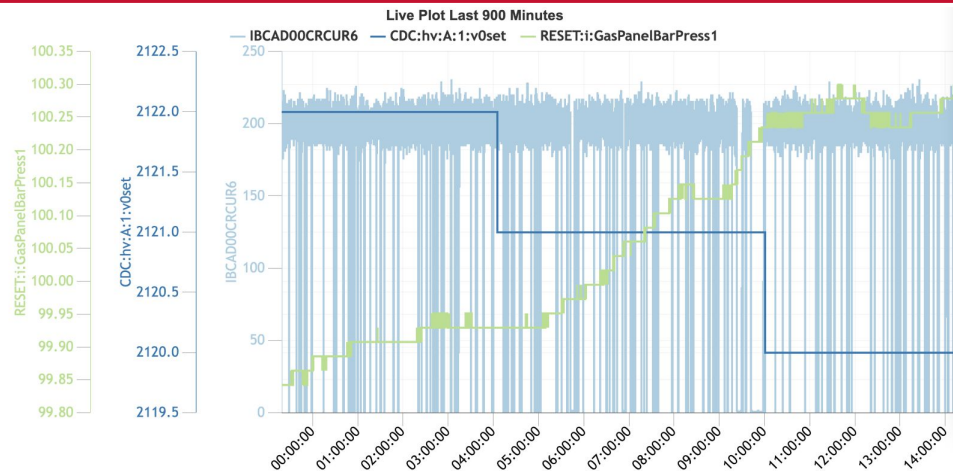
- Leave one side at a fixed HV (conventional)
- Let the ML control the other
- **Autonomously** adjust HV every 5 min

Observed Behavior that was Unexpected

Plot to the right shows HV setting was dropping while atmospheric pressure was rising during period of constant beam current. This is the opposite of what is expected.

Issue turned out to be due to using point on surface of minimum acceptable uncertainty with the minimal Euclidean distance to actual point in feature space.

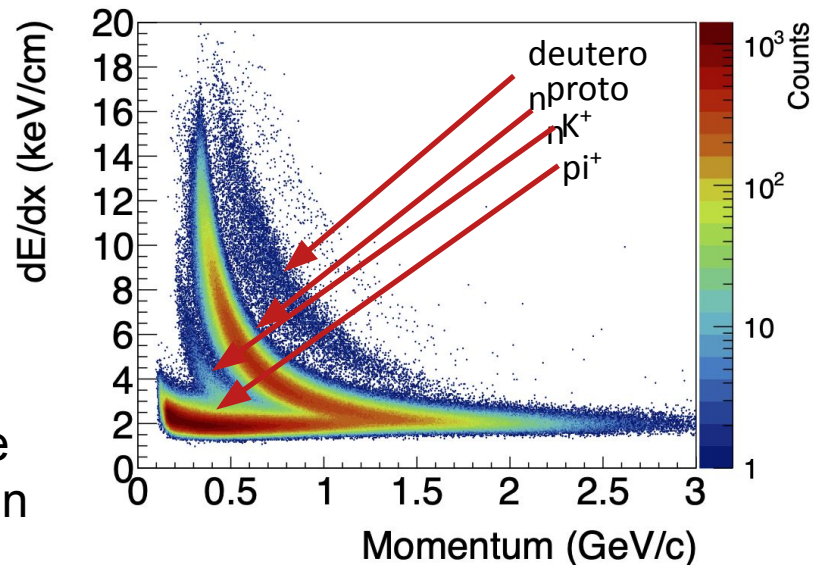
A small change in location in feature space could result in a large change in the projected location on the surface of uncertainty.



n.b. the GCF value was actually still within the few percent tolerance for operations.

CDC Calibrations

- Gain affects PID selections in analysis
 - Sensitive to environmental conditions
 - Atmospheric pressure
 - Temperature
 - Sensitive to experimental conditions
 - Beam conditions change with the experiment
- **Traditionally:**
 - GCF obtained from Landau fit to amplitude
 - Calibration constants are generated per run
 - Approximately 2 hours of beam time



FCAL Radiation damage

