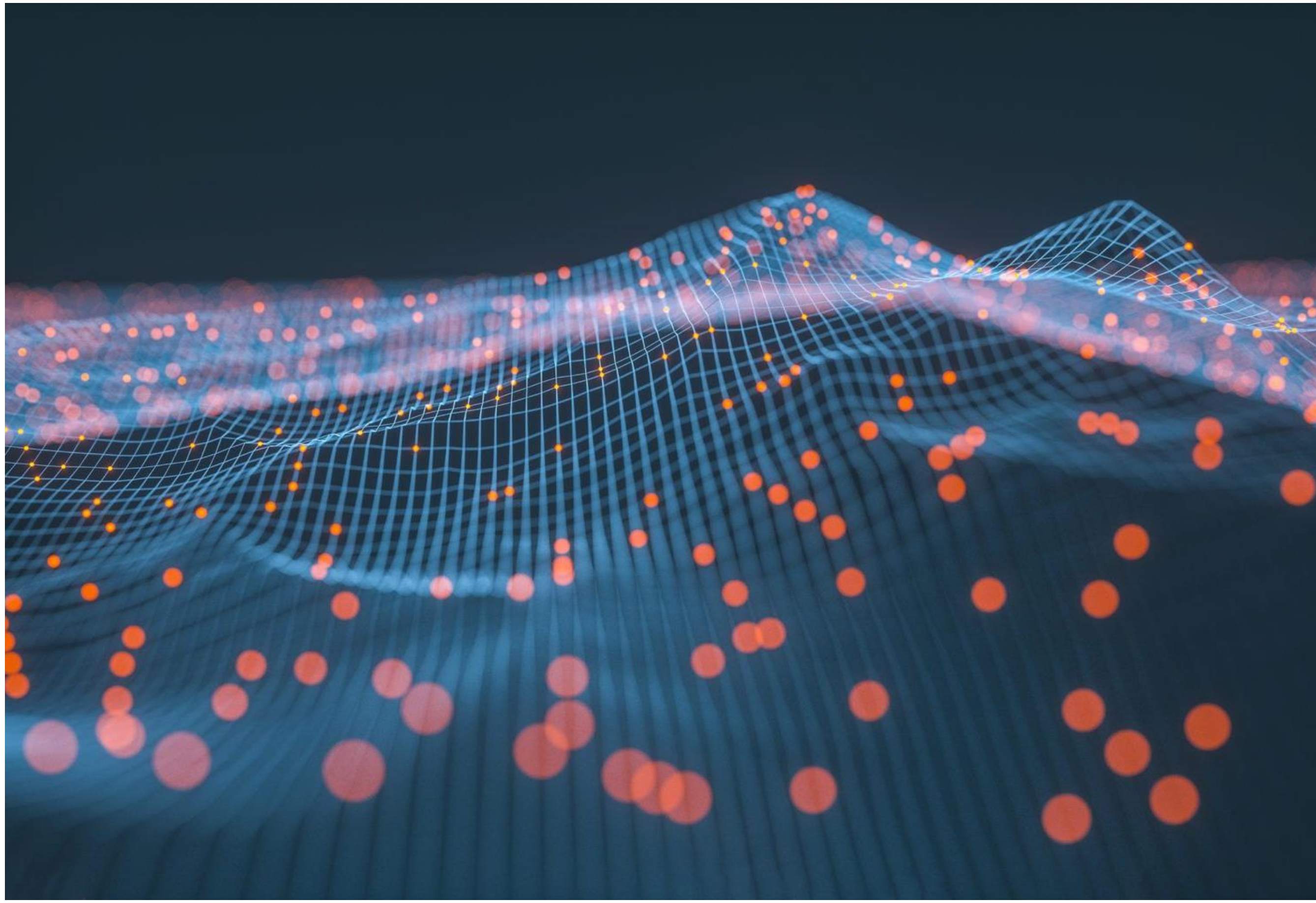
AI for Optimized SRF Performance of CEBAF Operations

Chris Tennant

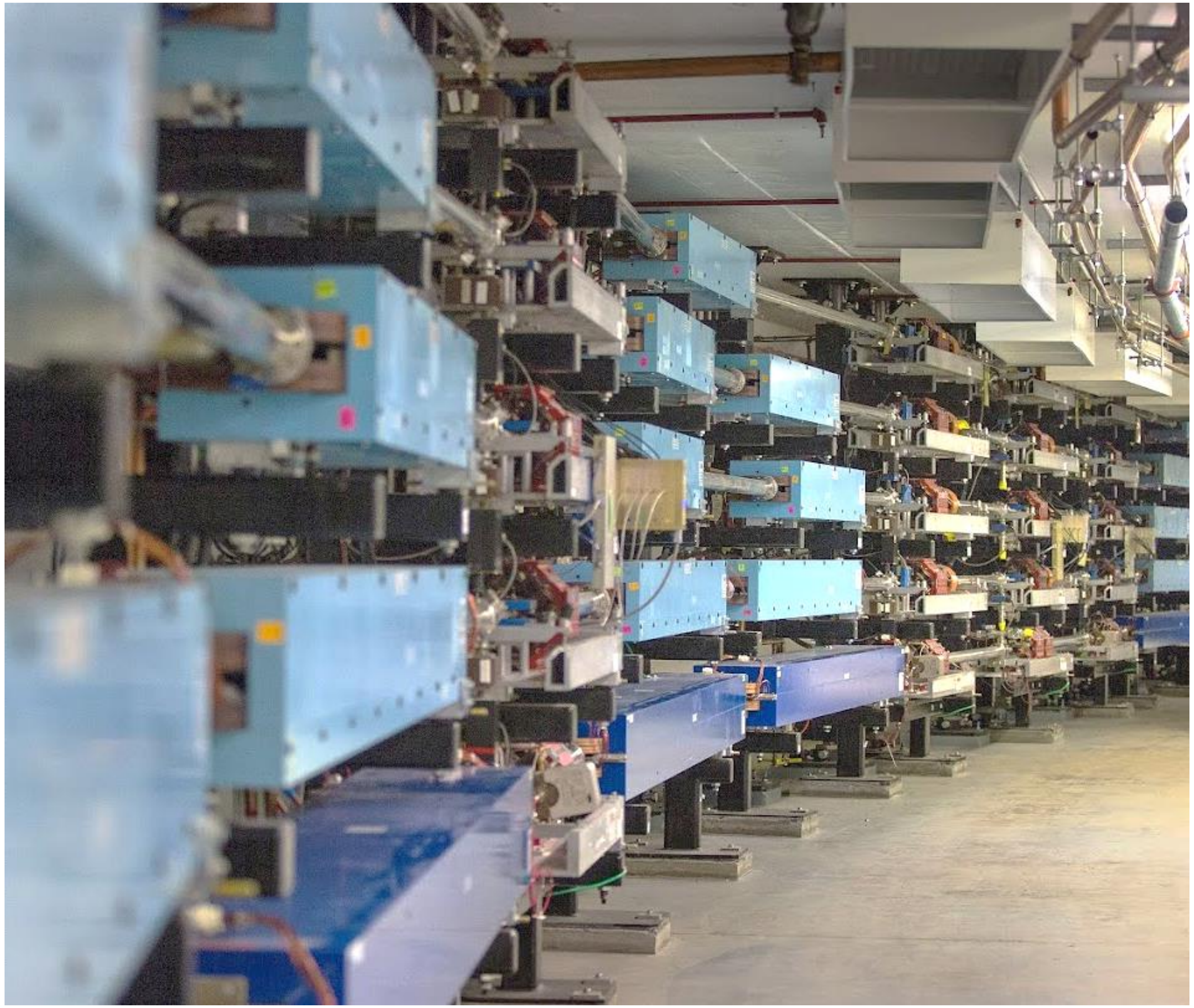
Kawser Ahammed, Adam Carpenter, Hal Ferguson, Steven Goldenberg, Khan Iftekharuddin, Jiang Li, Md. M. Monibor Rahman, Riad Suleiman, Dillon Thomas, Dennis Turner

DOE PI AI/ML Exchange Meeting | December 5, 2024



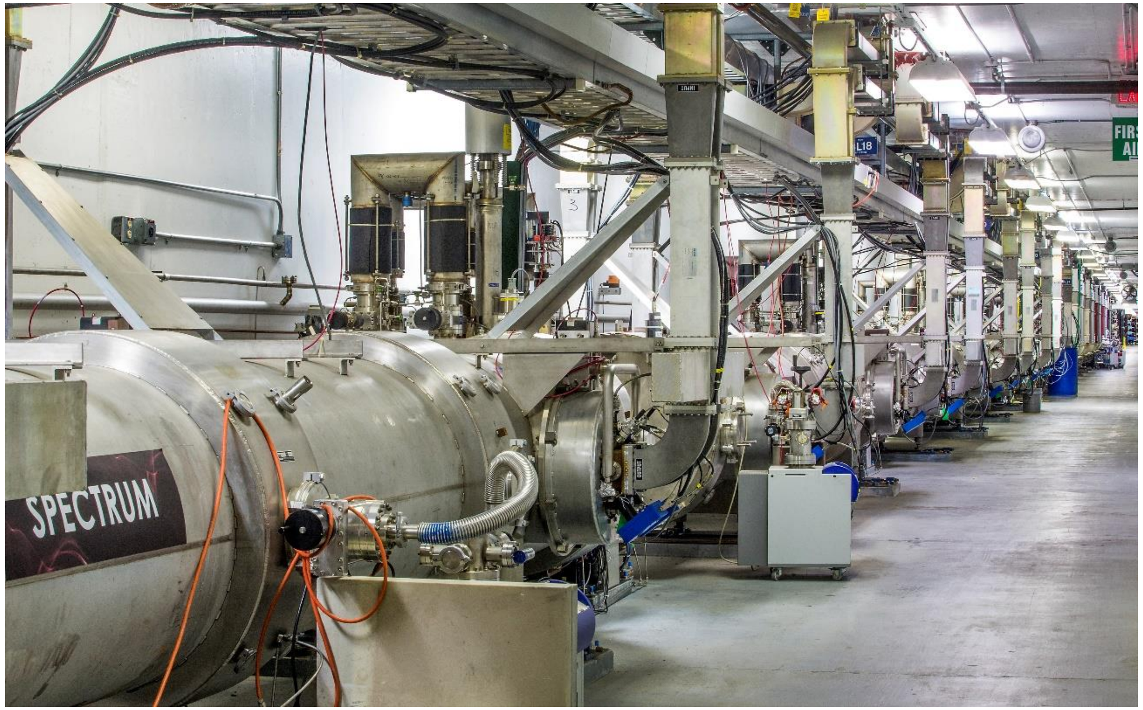
Outline

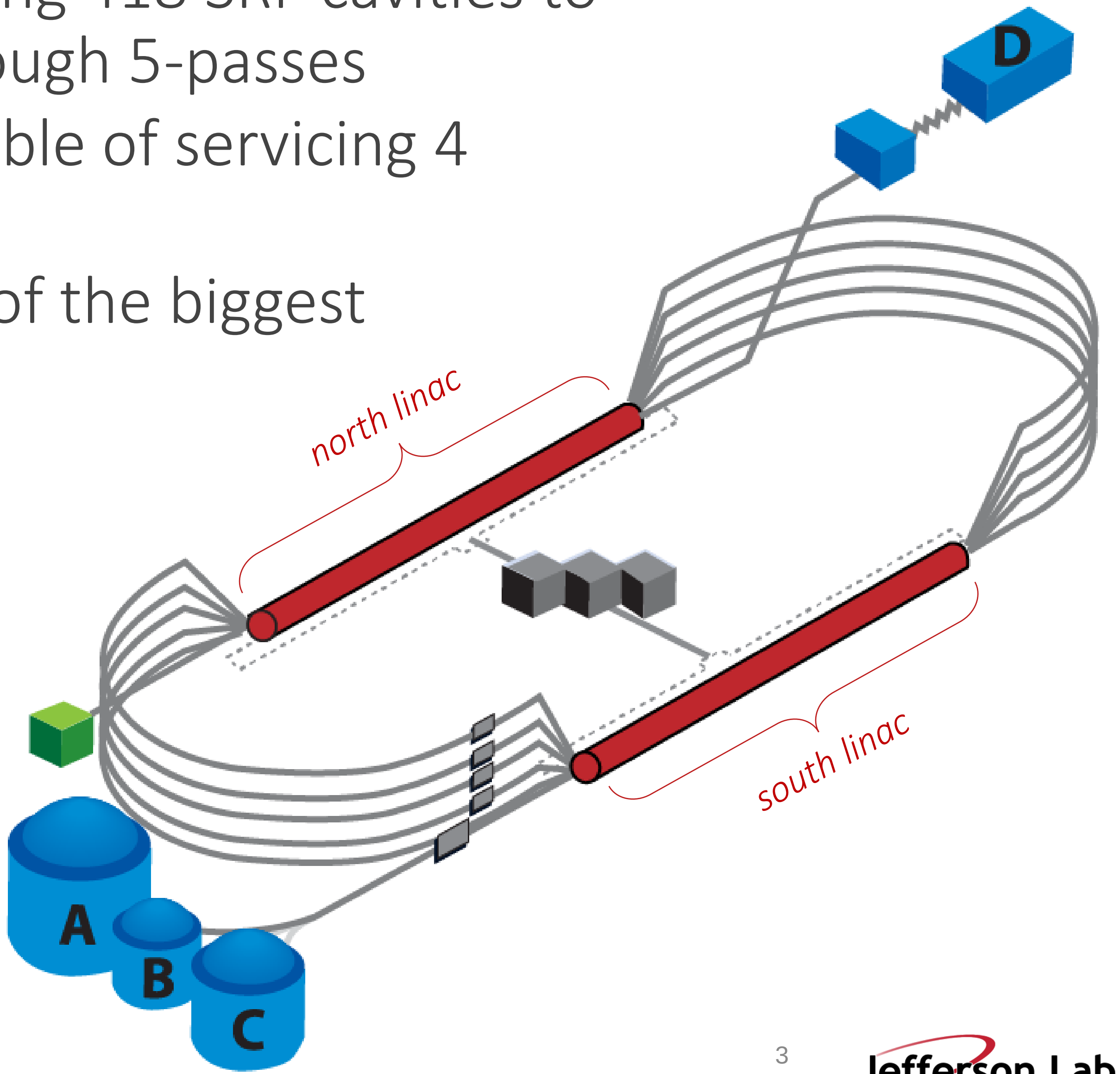
- Jefferson Laboratory
- FOA LAB 20-2261: Year 4 Status
  - ✓ *Anomalous Cavity Detection*
  - ✓ *Cavity Fault Prediction*
  - ✓ *Field Emission Management*
- Project Summary



Continuous Electron Beam Accelerator Facility

- CEBAF is a CW recirculating linac utilizing 418 SRF cavities to accelerate electrons up to 12 GeV through 5-passes
- it is a nuclear physics user-facility capable of servicing 4 experimental halls simultaneously
- RF related issues are consistently one of the biggest contributors to downtime





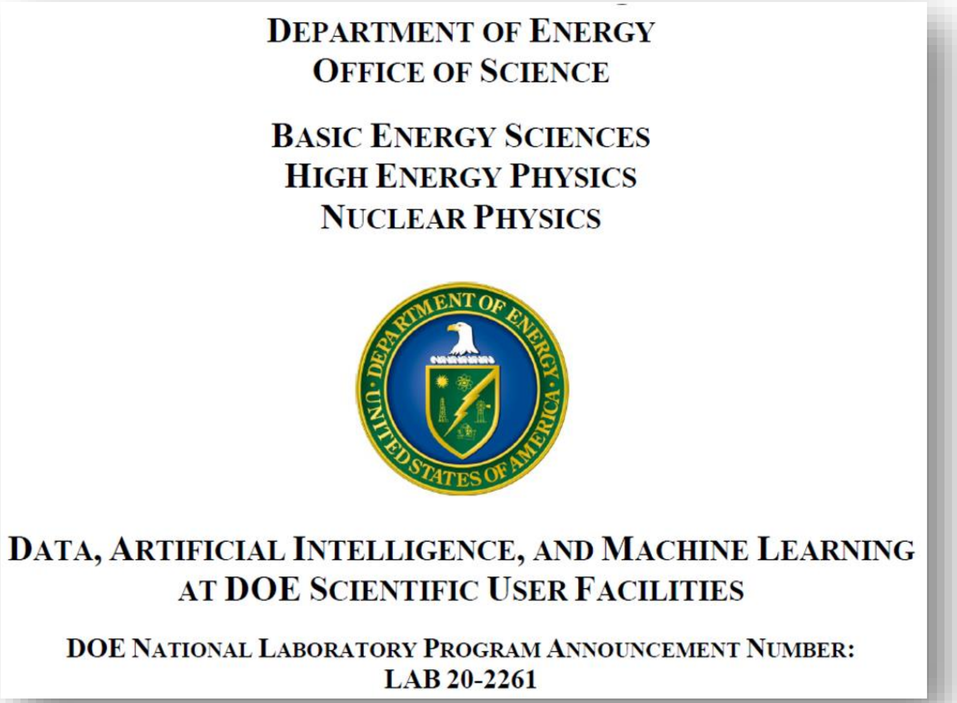
Jefferson Lab

"AI for Optimized SRF Performance of CEBAF Operations"

The proposal presents a multi-faceted approach to:

1. develop tools to automate cavity instability detection
2. provide real-time fault prediction for C100 cavities
3. minimize radiation levels due to field emission in the linacs

Improving SRF performance in these ways would translate to increased beam availability and reliability of CEBAF, increased beam-on-target for nuclear physics users, and meet DOE's mission to maximize scientific output per operating dollar.



DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

BASIC ENERGY SCIENCES
HIGH ENERGY PHYSICS
NUCLEAR PHYSICS

DATA, ARTIFICIAL INTELLIGENCE, AND MACHINE LEARNING
AT DOE SCIENTIFIC USER FACILITIES

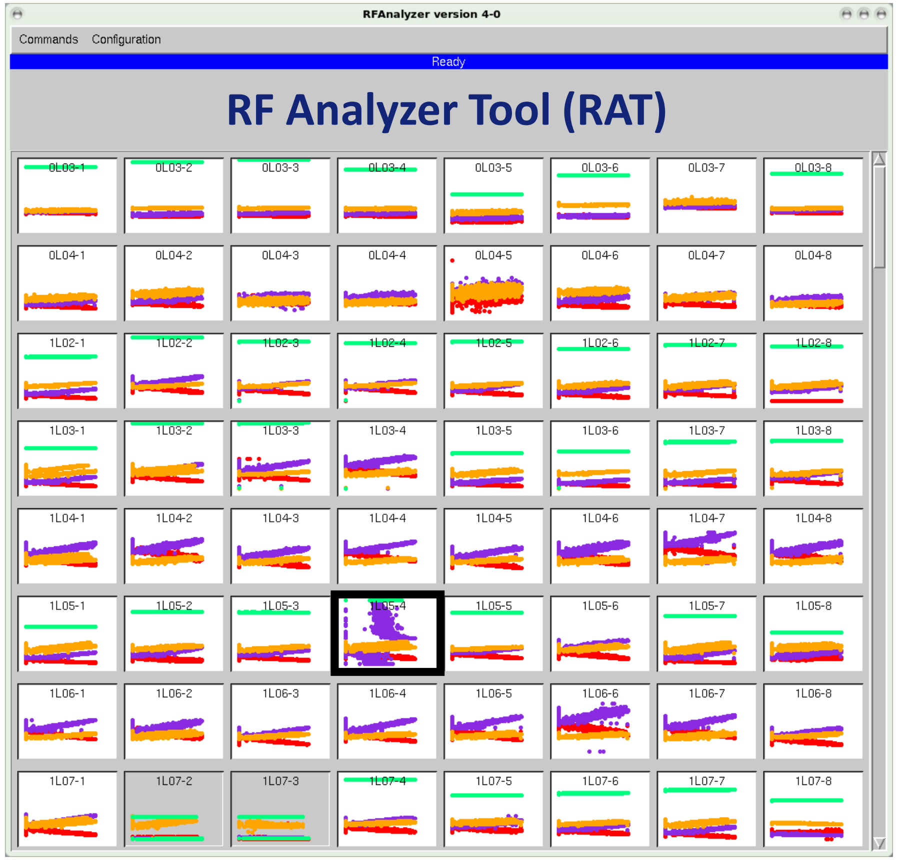
DOE NATIONAL LABORATORY PROGRAM ANNOUNCEMENT NUMBER:
LAB 20-2261

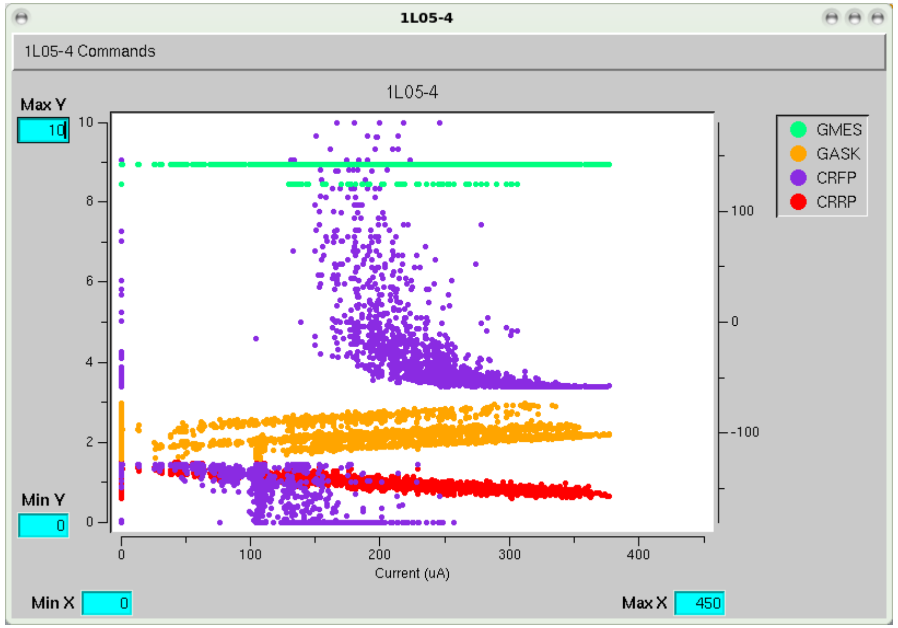
Anomalous Cavity Detection

PI: Dennis Turner
Graduate Student: Hal Ferguson (ODU)

Anomalous Cavity Detection

- **Goal:** automate the process of identifying RF cavities that exhibit anomalous behavior, but do *not* present as a fault
- **Previous Implementation:** manual inspection of hundreds of plots looking for an outlier
- **New Implementation:** use unsupervised learning to identify anomalous cavity behavior

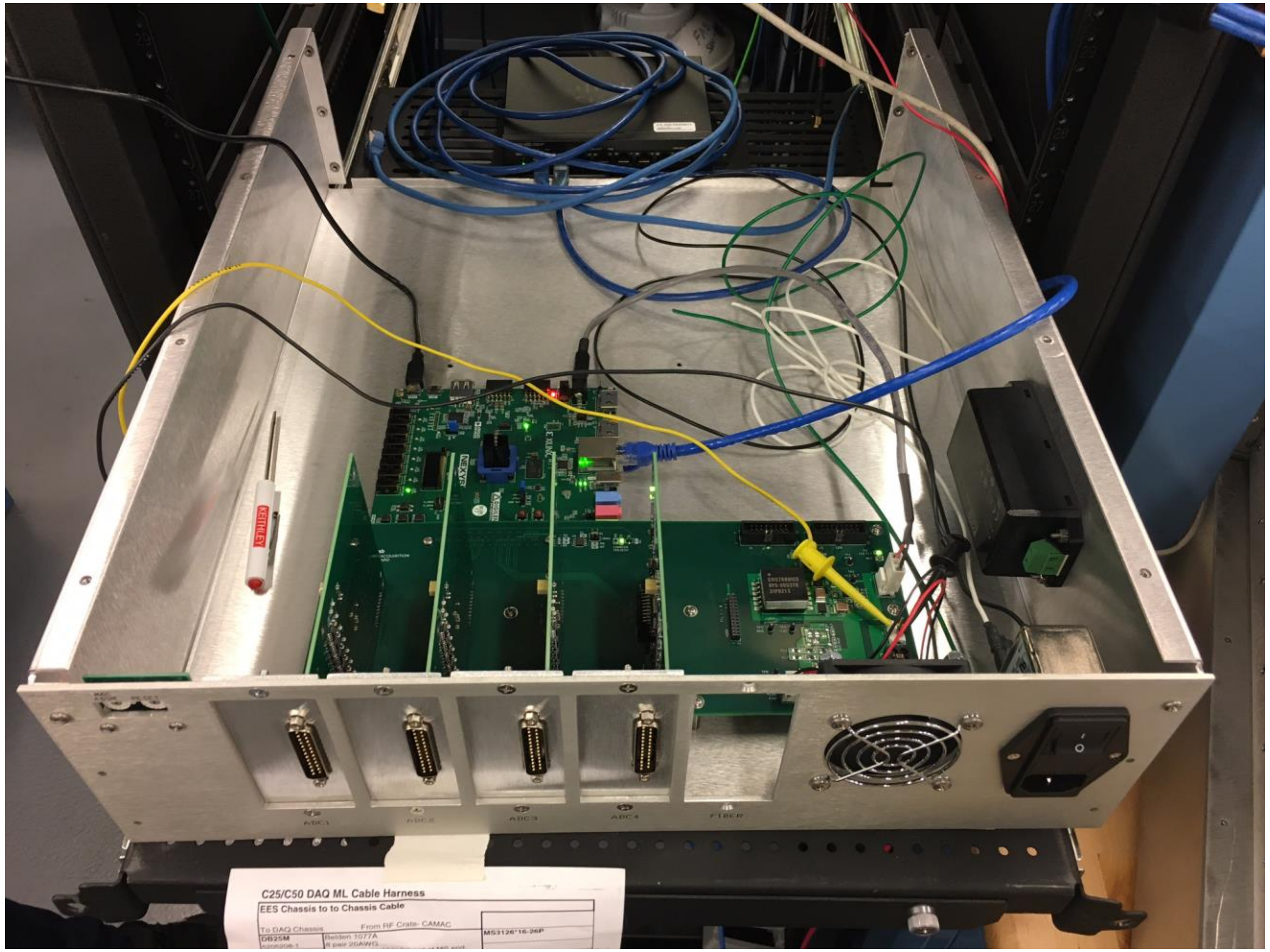




- this represents an obvious example
- not all instances are so easily detectable

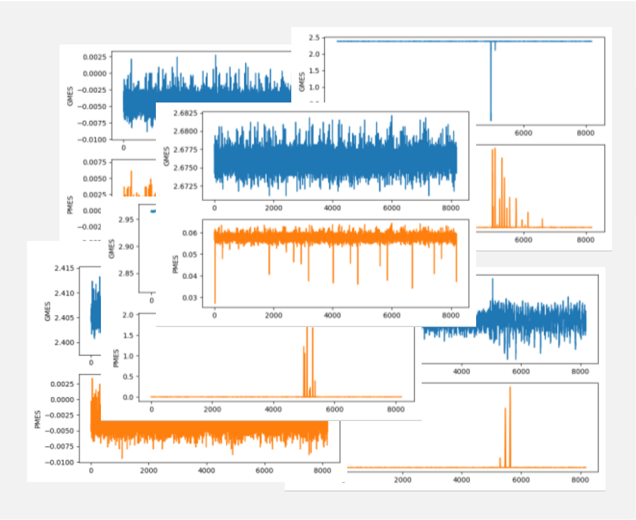
Anomalous Cavity Detection: Data Acquisition System (DAQ)

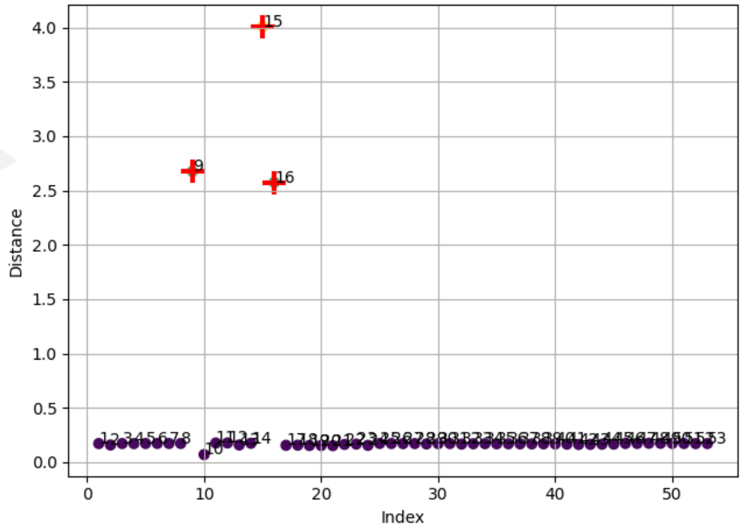
- 16 DAQs for NL (reduced scope due to rising costs)
- all legacy cryomodules in NL are outfitted with DAQs



Anomalous Cavity Detection: Workflow

- when a machine trip occurs, we collect RF signals from across all DAQs if the fault involves a BLM, ion chamber, or BLA trip but <u>*not*</u> a cavity trip
  - ✓ this represents data that potentially exhibits anomalous cavity behavior
  - ✓ data from 1 machine trip = 16 cryomodules x 8 cavities/cryomodule x 2 signals/cavity = 256 signals
- our initial workflow consisted of
  - ✓ extract *n* features from each cavity's pair of signals
  - ✓ use PCA to reduce dimensionality from 2*n* to 2 for visualization
  - ✓ compute centroid of data points
  - ✓ compute distance of every data point from centroid and plot

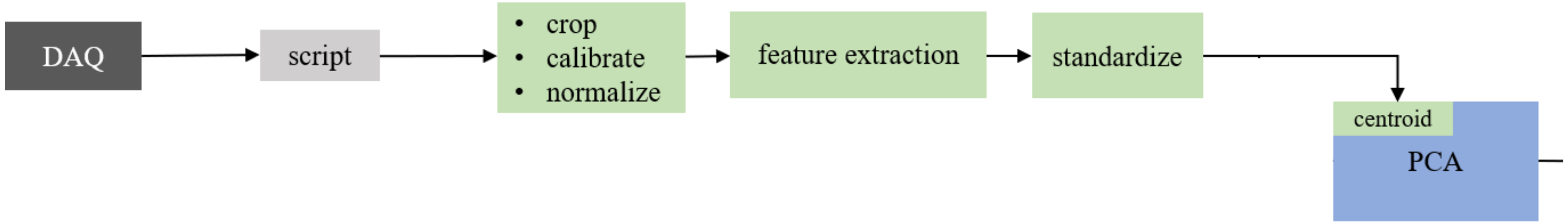




- however, it became evident that each cavity has its own distinctive characteristics
- many of the outliers were cavities that were consistently more noisy – *but stable* – compared to other cavities

Anomalous Cavity Detection: Workflow

- rather than compare many cavities across a single timestamp, build a workflow that compares each cavity across time



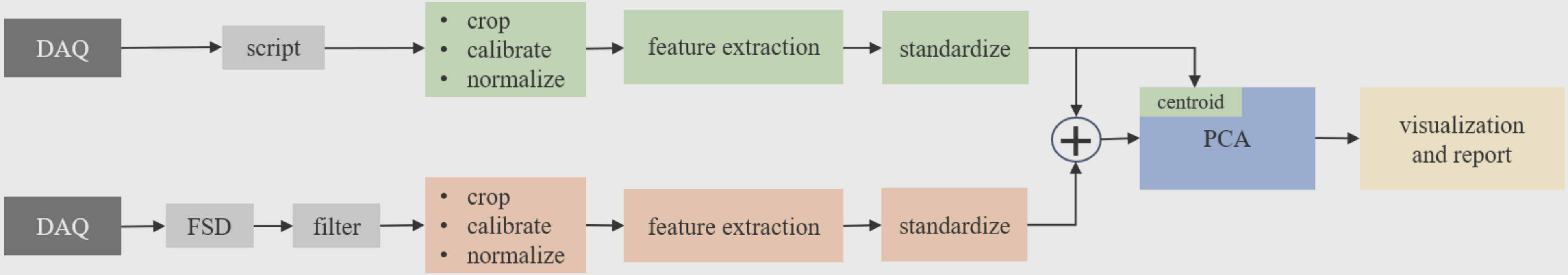
- new workflow
  - ✓collect RF signals from each cavity at various times during normal operation → *normal*
  - ✓train PCA model on normal data from previous 7 days
  - ✓collect RF signals associated with machine faults from previous 24 hours → *potential*
  - ✓every day, <u>for each cavity</u>, generate a PCA plot showing the *normal* and *potential* data
  - ✓compile report of cavities with the largest distances from centroid of their respective normal data

```
Top 25 Cavities with the Highest Distance from Centroid for all timestamps:
============================
Timestamp: 2024-05-19_23_53_08, Cavity Name: R1C6, Distance to Centroid: 55.48696304072229
Timestamp: 2024-05-19_23_24_48, Cavity Name: R126, Distance to Centroid: 36.67617775495357
Timestamp: 2024-05-19_23_40_46, Cavity Name: R126, Distance to Centroid: 34.817115260624156
Timestamp: 2024-05-19_23_30_17, Cavity Name: R126, Distance to Centroid: 32.624079589243046
Timestamp: 2024-05-19_23_28_13, Cavity Name: R126, Distance to Centroid: 32.55733745821345
```


Jefferson Lab

Anomalous Cavity Detection: Workflow

- rather than compare different cavities across a single timestamp, build a workflow that compares a each cavity across time



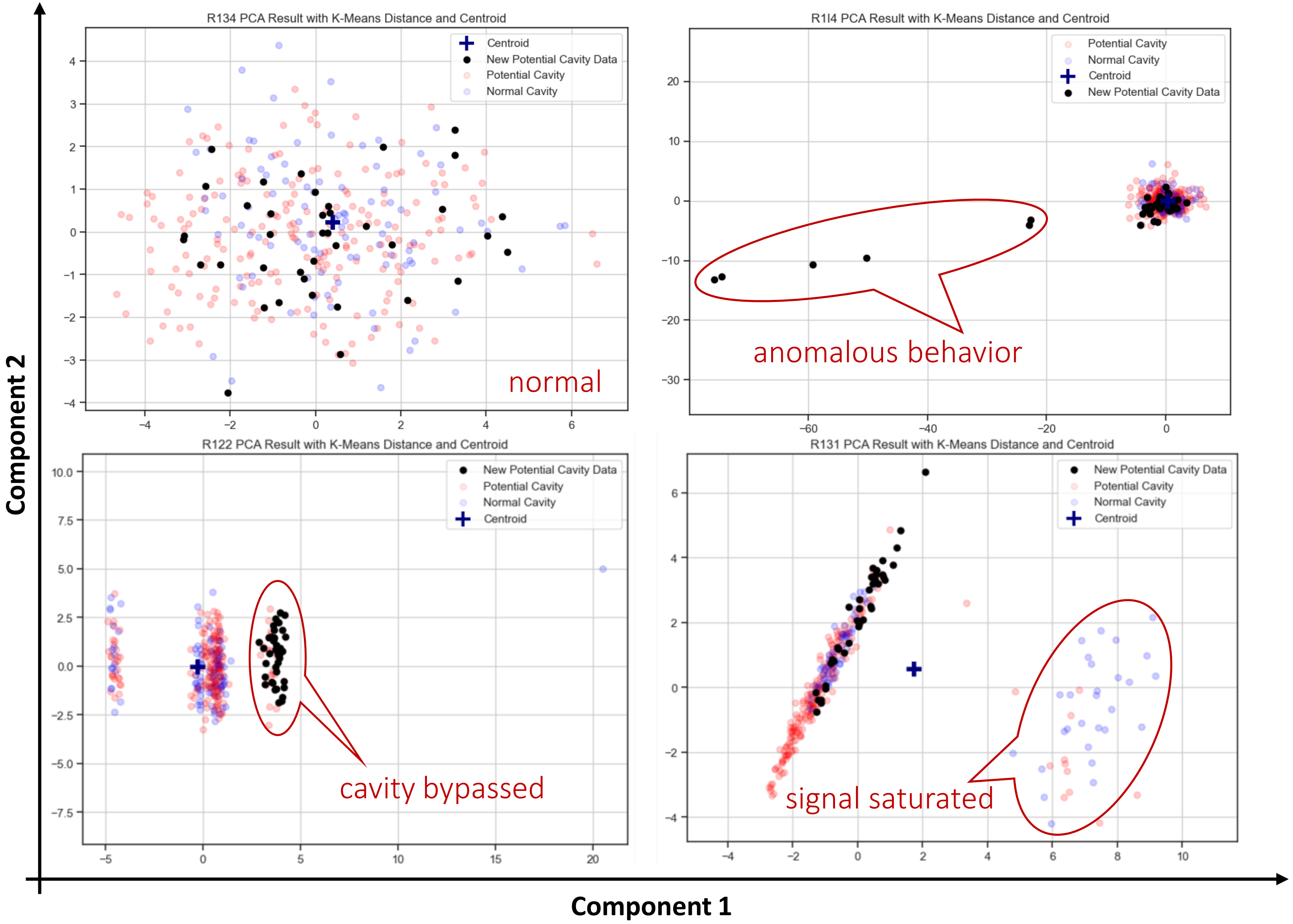
- new workflow

The intuition is simple. Look at data that is potentially anomalous and highlight instances where it differs from normal operation over the previous week.

```
Top 25 Cavities with the Highest Distance from Centroid for all timestamps:
============================
Timestamp: 2024-05-19_23_53_08, Cavity Name: R1C6, Distance to Centroid: 55.48696304072229
Timestamp: 2024-05-19_23_24_48, Cavity Name: R126, Distance to Centroid: 36.67617775495357
Timestamp: 2024-05-19_23_40_46, Cavity Name: R126, Distance to Centroid: 34.817115260624156
Timestamp: 2024-05-19_23_30_17, Cavity Name: R126, Distance to Centroid: 32.624079589243046
Timestamp: 2024-05-19_23_28_13, Cavity Name: R126, Distance to Centroid: 32.55733745821345
```

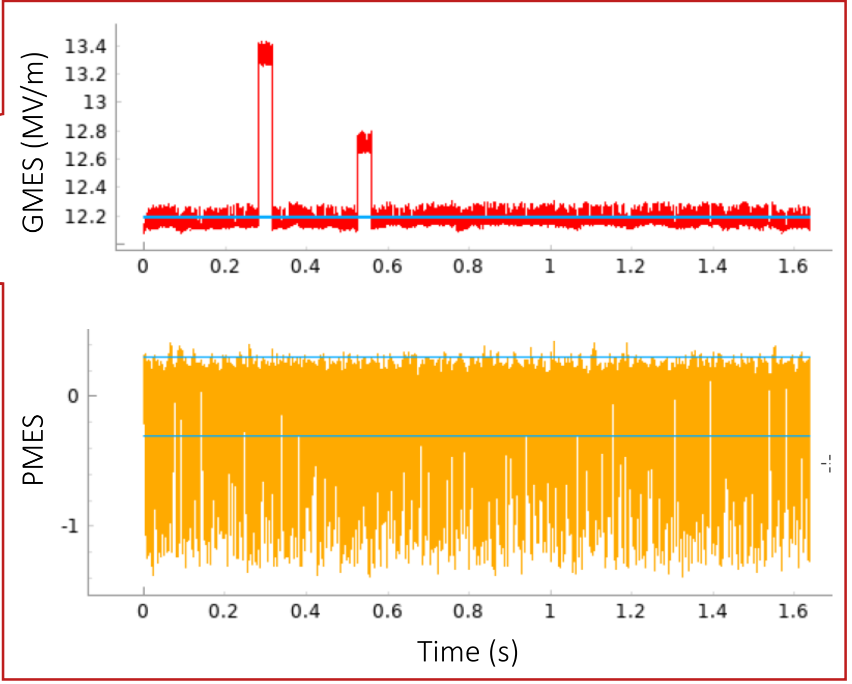

Jefferson Lab

Anomalous Cavity Detection: PCA Plots



Anomalous Cavity Detection: Results

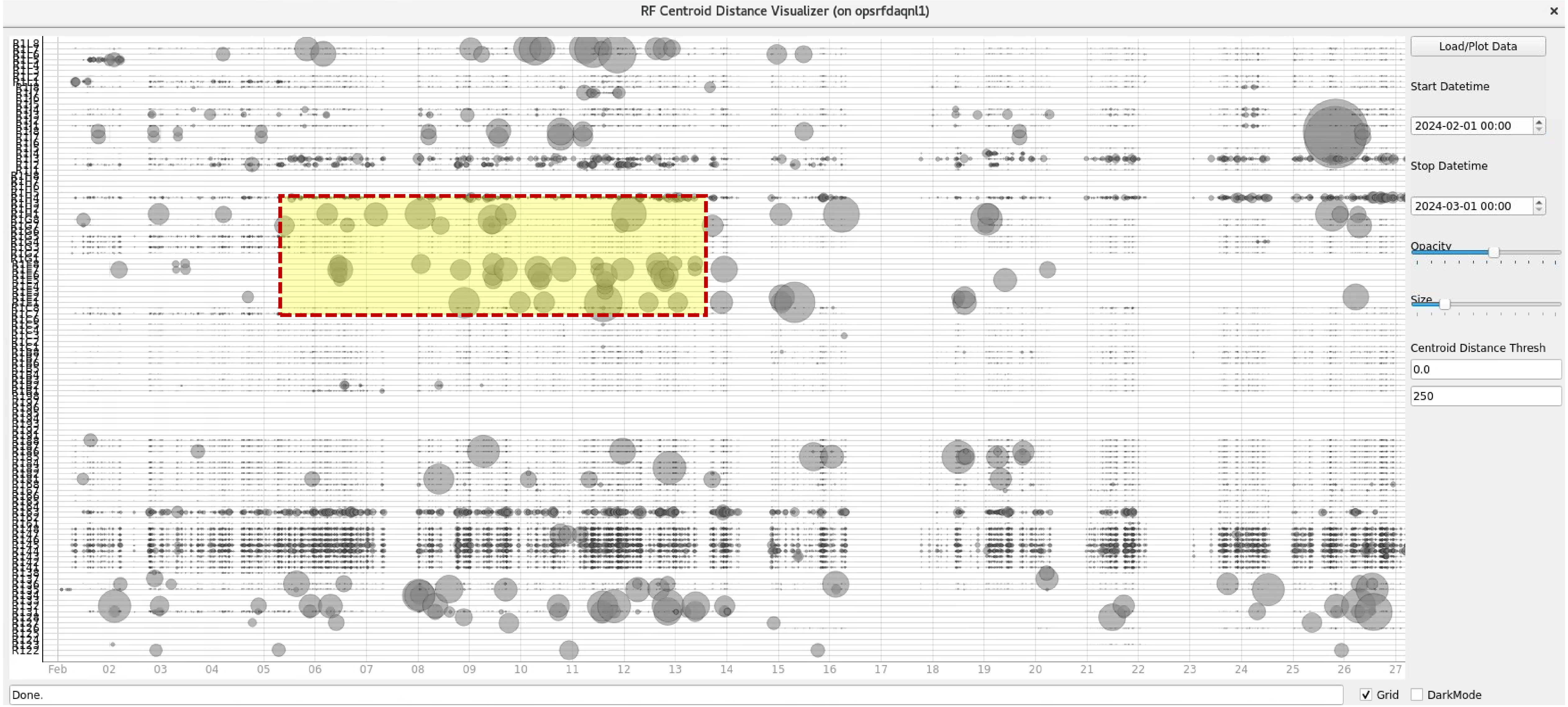
- the framework was deployed and operational in CEBAF during the Spring 2024 run and proved to be very effective at identifying anomalous signals
- due to lingering issues with the DAQ system, most of the anomalies are traced to DAQ system “features” and not to cavity behavior
  - ✓ e.g. single-valued (saturated) signals, noise in the system
- several successes are worth noting:
  - ✓ RF control module replacement
  - ✓ microphonics and power supply failure on a cryomodule
  - ✓ *note: we analyzed data for less than 4 months and effectively from only 8 of the 38 legacy cryomodules*



- for the next CEBAF operational run:
  - ✓ attenuators have been installed to avoid saturated signals
  - ✓ waveforms from C100 cavities – which already have a system to output fast sampled signals – added to the framework
    - *NL coverage: 23/25 cryomodules*
    - *SL coverage: 5/25 cryomodules*

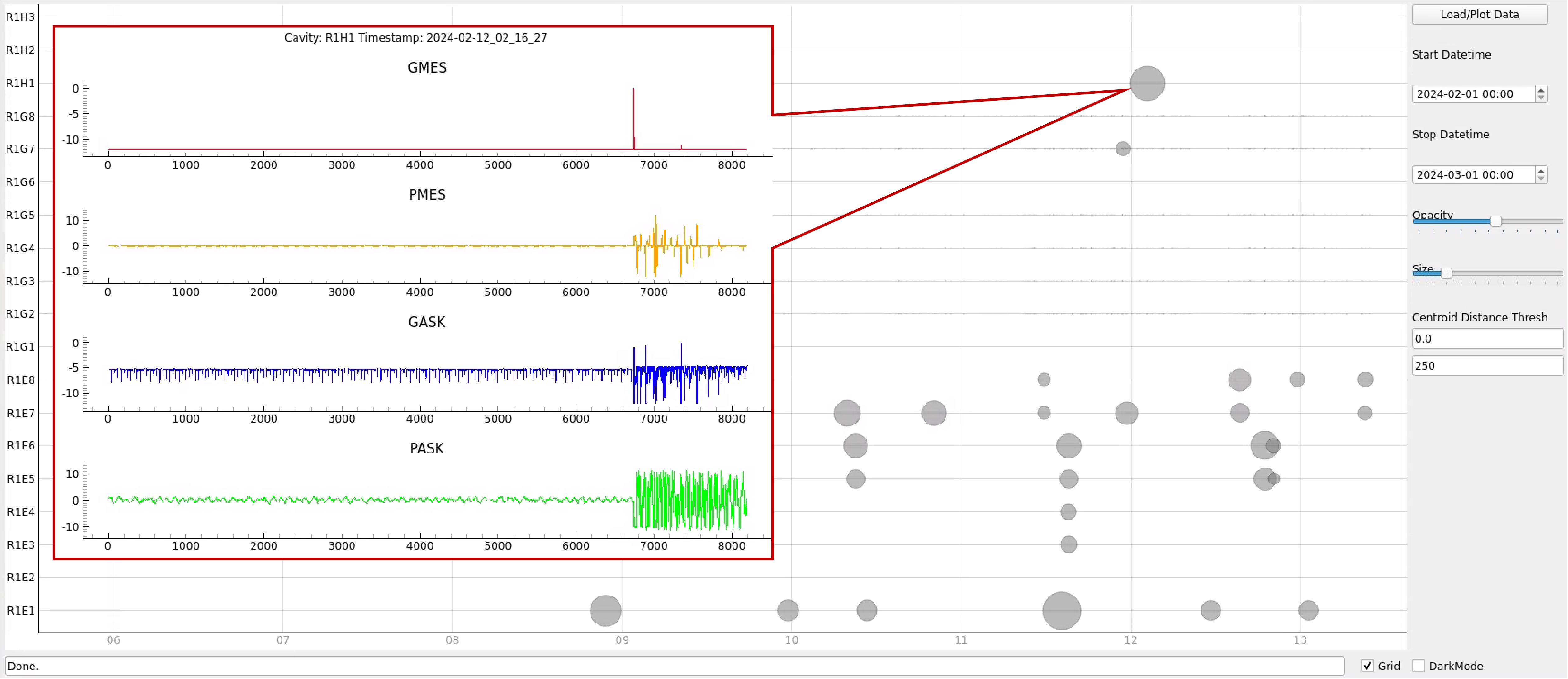
Jefferson Lab

Interactive Timeline View



Jefferson Lab

Interactive Timeline View

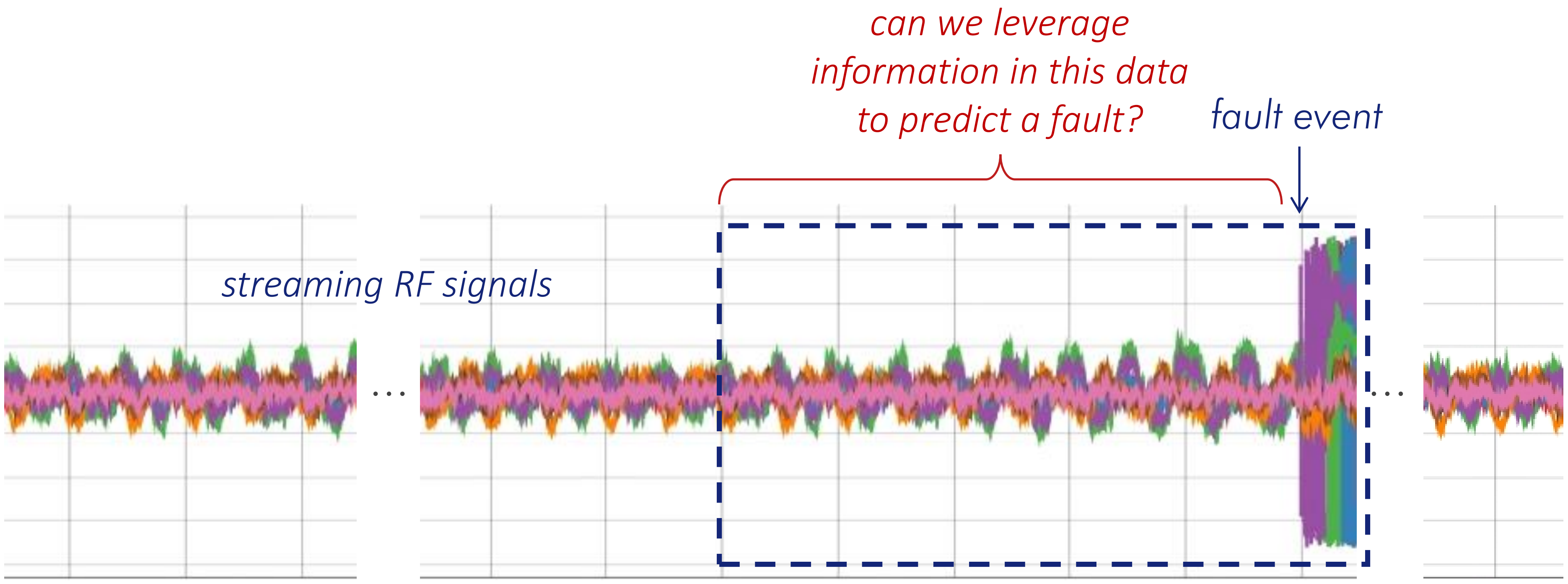


Cavity Fault Prediction

PI: Chris Tennant
Graduate Student: Md. Monibor Rahman (ODU)

Cavity Fault Prediction

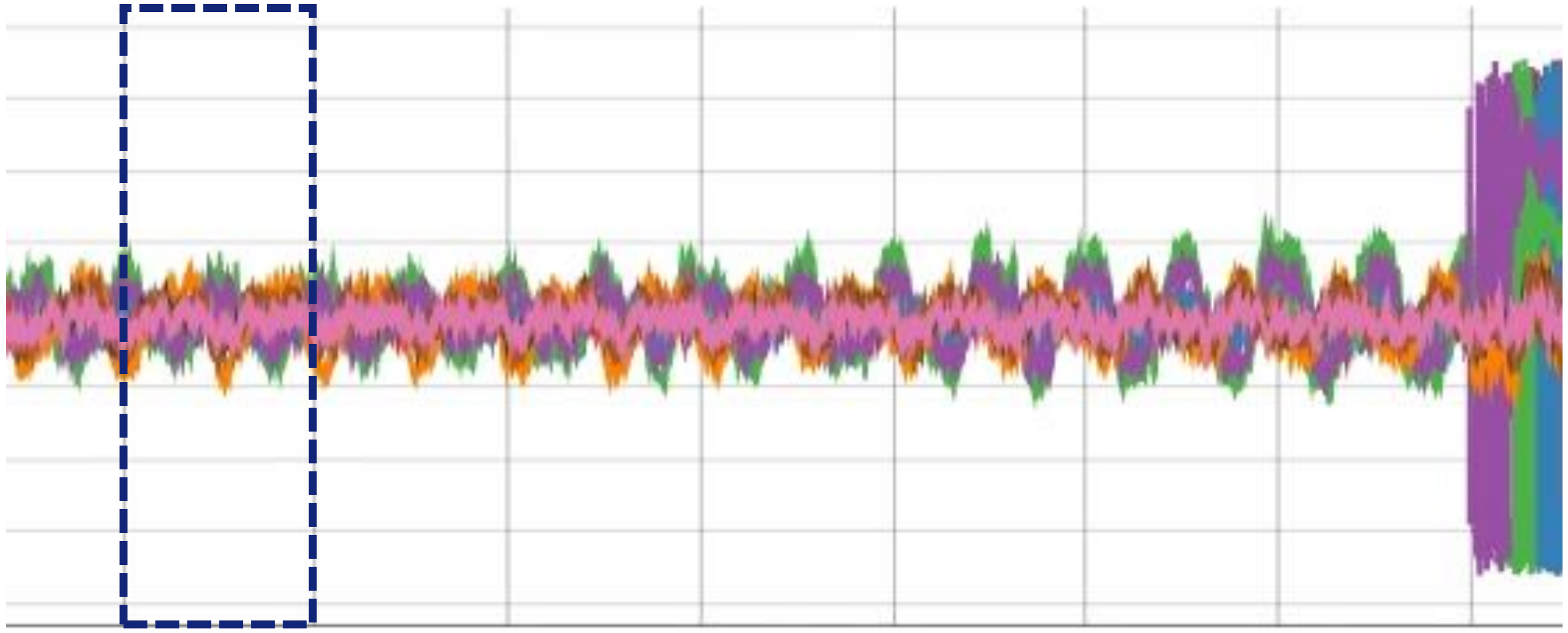
- **Goal:** predict if an RF cavity fault will occur
- **Previous Implementation:** *N/A*
- **New Implementation:** use deep learning to identify features in pre-fault data to predict slow developing cavity faults



Jefferson Lab

Cavity Fault Prediction: Sliding Window

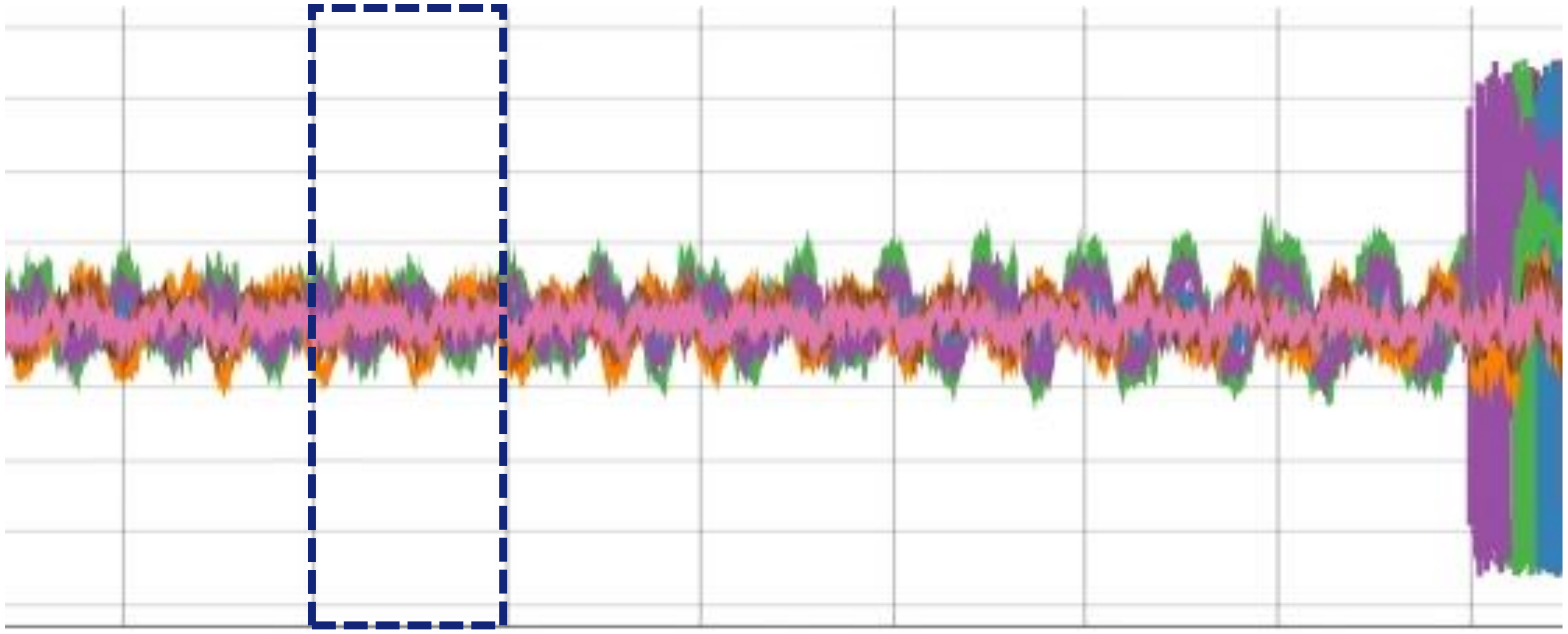
- what should the duration of the time window be?
- how many consecutive windows should be used to make a prediction?
- for the choice of those parameters, is the predictive power sufficient?



Jefferson Lab

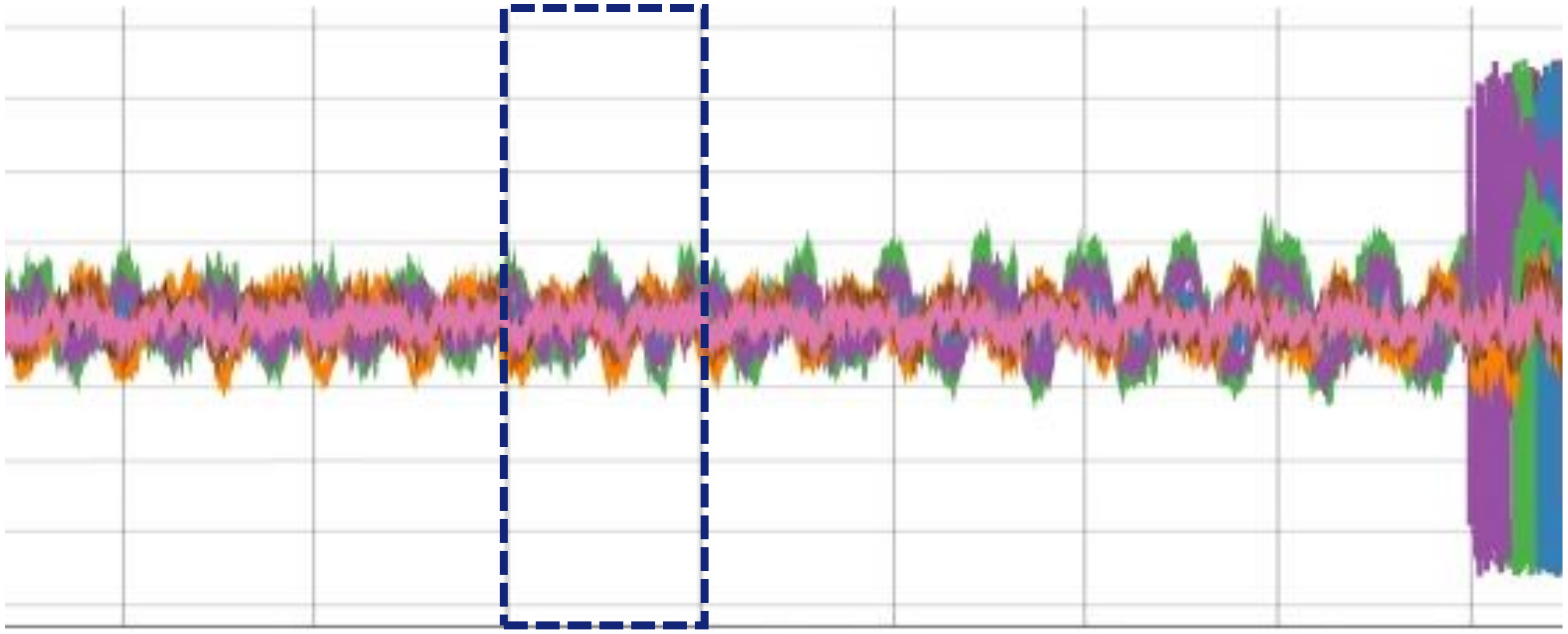
Cavity Fault Prediction: Sliding Window

- what should the duration of the time window be?
- how many consecutive windows should be used to make a prediction?
- for the choice of those parameters, is the predictive power sufficient?



Cavity Fault Prediction: Sliding Window

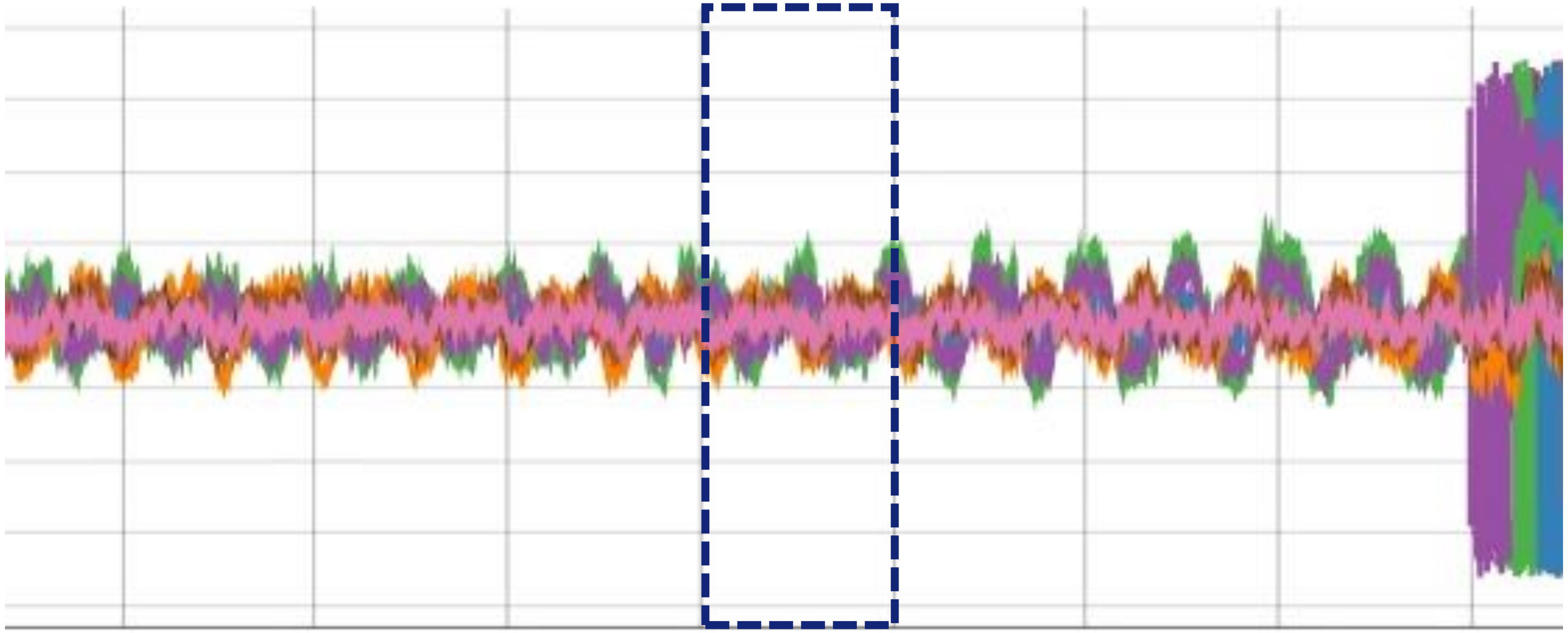
- what should the duration of the time window be?
- how many consecutive windows should be used to make a prediction?
- for the choice of those parameters, is the predictive power sufficient?



Jefferson Lab

Cavity Fault Prediction: Sliding Window

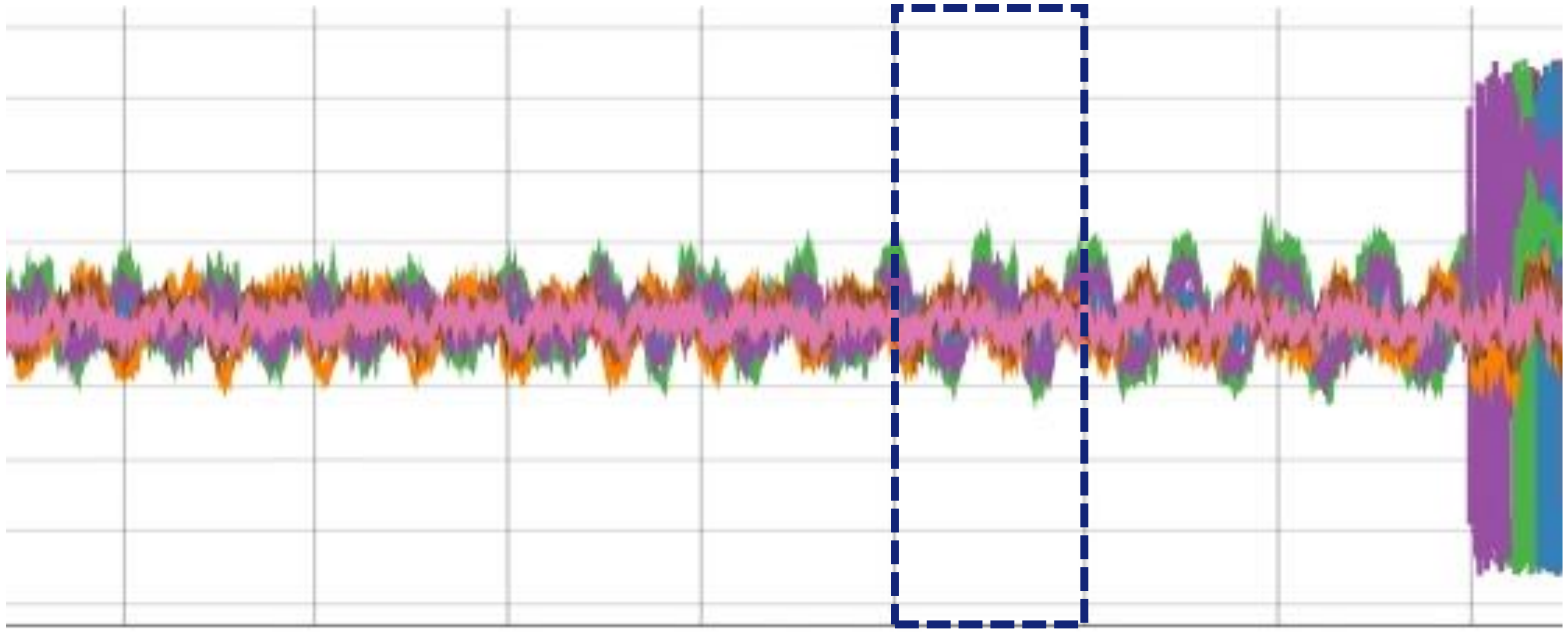
- what should the duration of the time window be?
- how many consecutive windows should be used to make a prediction?
- for the choice of those parameters, is the predictive power sufficient?



Jefferson Lab

Cavity Fault Prediction: Sliding Window

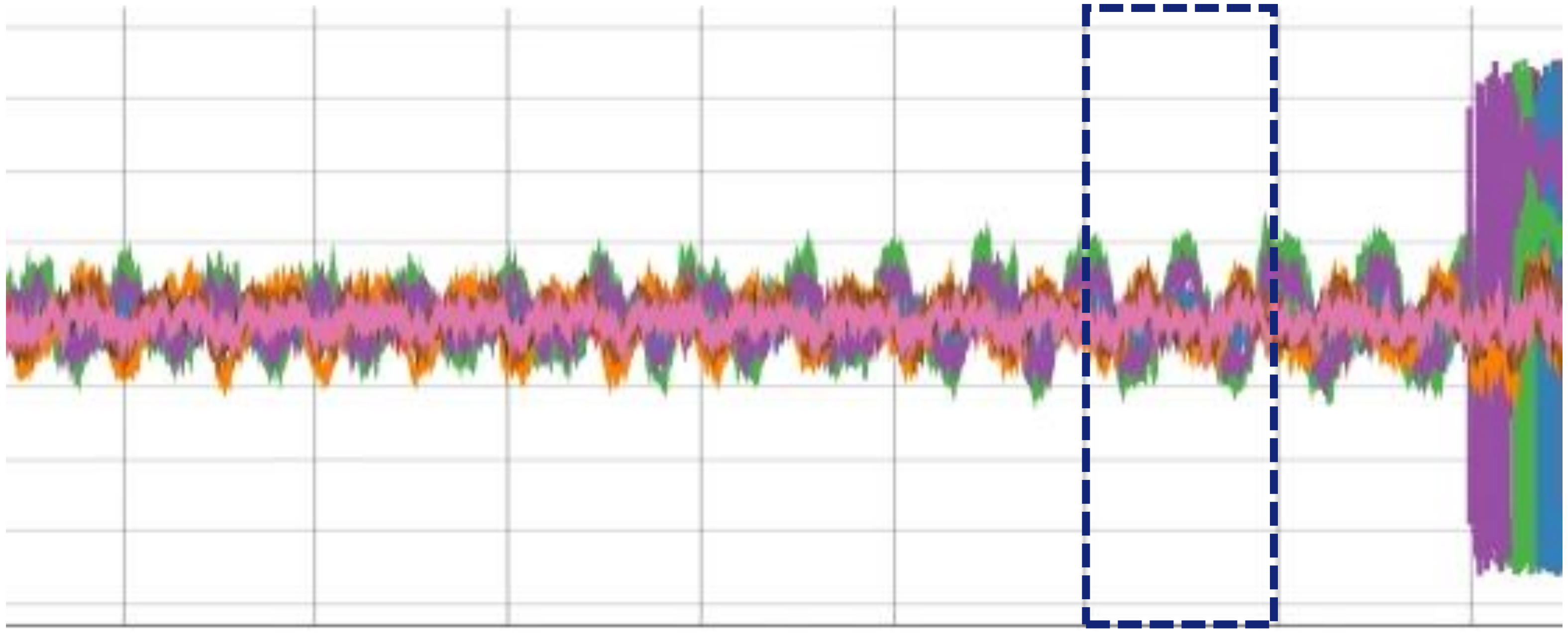
- what should the duration of the time window be?
- how many consecutive windows should be used to make a prediction?
- for the choice of those parameters, is the predictive power sufficient?



Jefferson Lab

Cavity Fault Prediction: Sliding Window

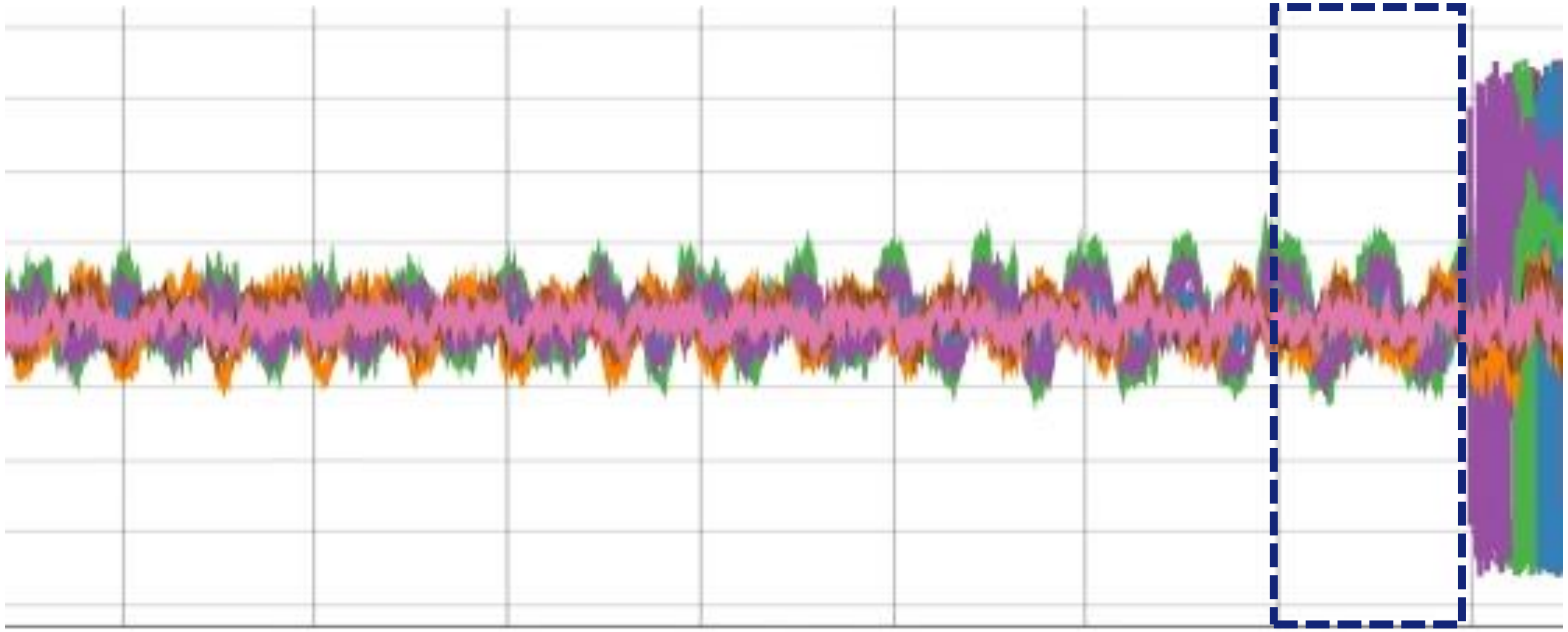
- what should the duration of the time window be?
- how many consecutive windows should be used to make a prediction?
- for the choice of those parameters, is the predictive power sufficient?



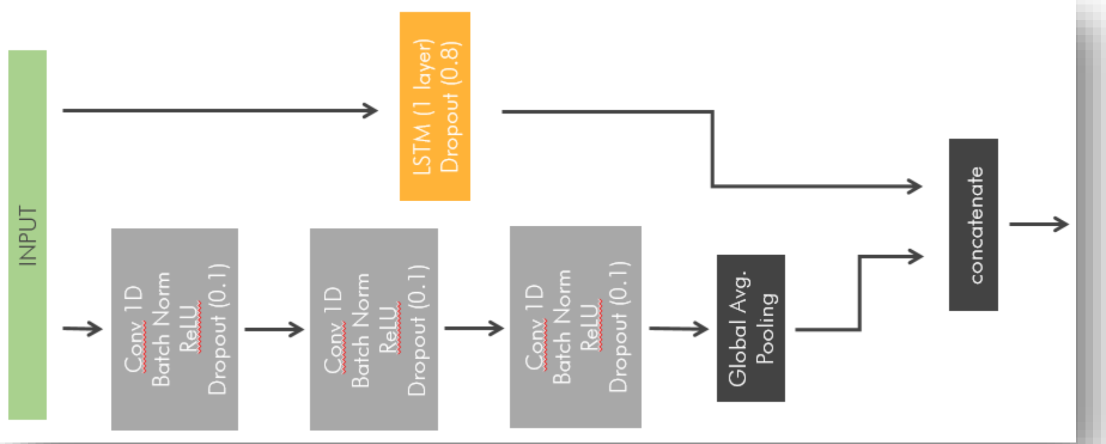
Jefferson Lab

Cavity Fault Prediction: Sliding Window

- what should the duration of the time window be? 100 ms
- how many consecutive windows should be used to make a prediction? 3
- for the choice of those parameters, is the predictive power sufficient? yes



Cavity Fault Prediction: Model Development



- model architecture: CNN + LSTM
  - ✓ the spatial features produced by the CNN branch and the temporal features generated by the LSTM layers are concatenated and passed through a fully connected layer
- model training
  - ✓ collect examples of *slow* faults from the last several years
  - ✓ collect examples from normal operation
  - ✓ train the model to distinguish between normal and pre-faulty signals by using 100 ms random samples from each type
    - *use the faulty samples within 500 ms of the fault onset – further from the fault and the signal is difficult to distinguish from a normal signal*
- model optimization: minimize false positives
  - ✓ number of consecutive windows to use for a fault prediction (3)
  - ✓ fine-tuning confidence threshold (cavity dependent)

Jefferson Lab

Cavity Fault Prediction: Simulated Deployment

- collected semi-continuous *normal* data from March 3-6, 2023 and tested model

	Cavity 1	Cavity 2	Cavity 3	Cavity 4	Cavity 5	Cavity 6	Cavity 7	Cavity 8
# of Examples*	6856	6588	6944	6874	6770	6816	6932	6949
Normal	6856	6588	6937	6874	6770	6816	6932	6949
Faulty	0	0	7	0	0	0	0	0
Accuracy	100	100	99.90	100	100	100	100	100

**1 example = 16 inferences*

- viewed collectively:
  - ✓ a total of 875,643 out of 875,664 inferences correctly predicted normal data
- collected 33 labeled *faults* from March 7-20, 2023 and tested model

Fault Type	Predicted	Not Predicted	Total	Accuracy
Slow	4	1	5	80.0%
Fast	1	27	28	3.6%

model was not trained on pre-fault data from fast faults

✓ accurately identifies normal data (with minimal false positives)

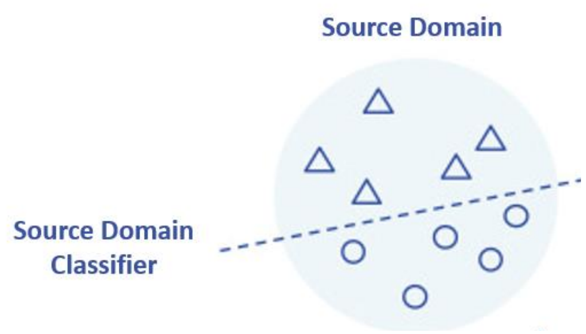
✓ is able to predict slow faults

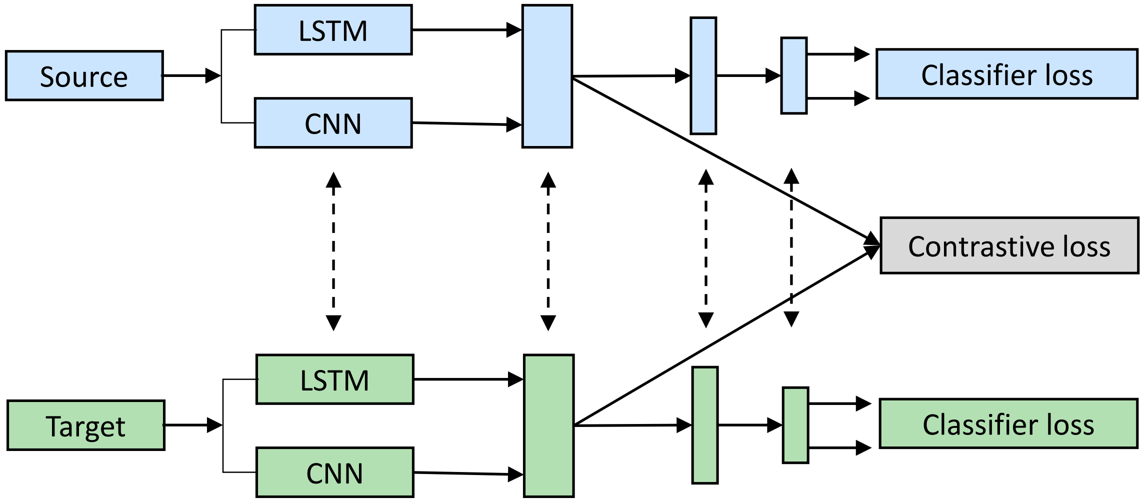
- *with hundreds of milliseconds prior to the fault*

✓ can do so in the context of a highly imbalanced data set

Cavity Fault Prediction: Domain Adaptation

- **domain shift**: refers to a situation where the distribution of the data used to train a machine learning model differs from the distribution of the data the model is applied to during inference or deployment.
- **domain adaptation**: aligning the source and target data distributions through techniques like adversarial training or feature alignment





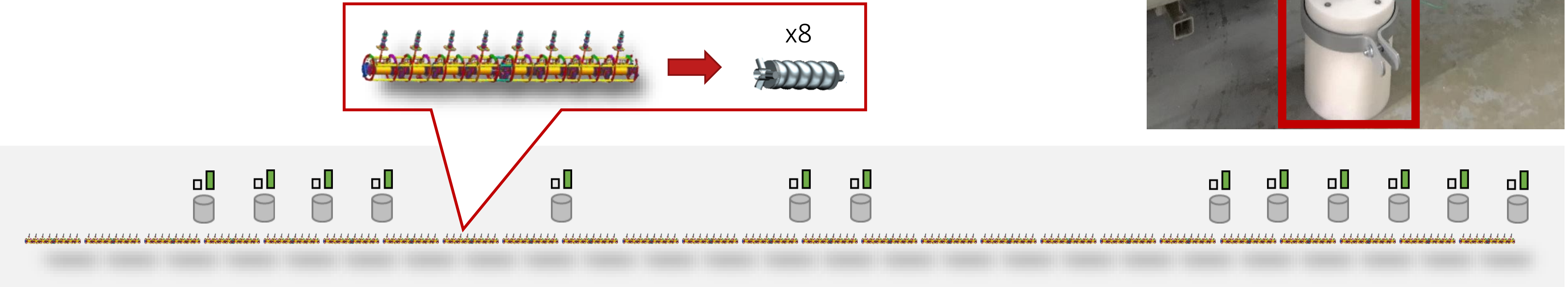
- current efforts are aimed at implementing and evaluating a variety of techniques for domain adaptation
  - ✓i.e. incorporating contrastive loss to make the same (differing) labels from different domains more alike (different)

Field Emission Management

PI: Adam Carpenter and Riad Suleiman
Postdoc: Steven Goldenberg
Graduate Student: Kawser Ahammed (ODU)

Field Emission Management

- **Goal:** reduce radiation due to SRF cavities through gradient redistribution through an automated process
- **Previous Implementation:** manual, trial and error tuning of cavity gradients during a dedicated beam study (i.e. no beam to users)
- **New Implementation:** pair an optimization algorithm with a surrogate model of radiation readings



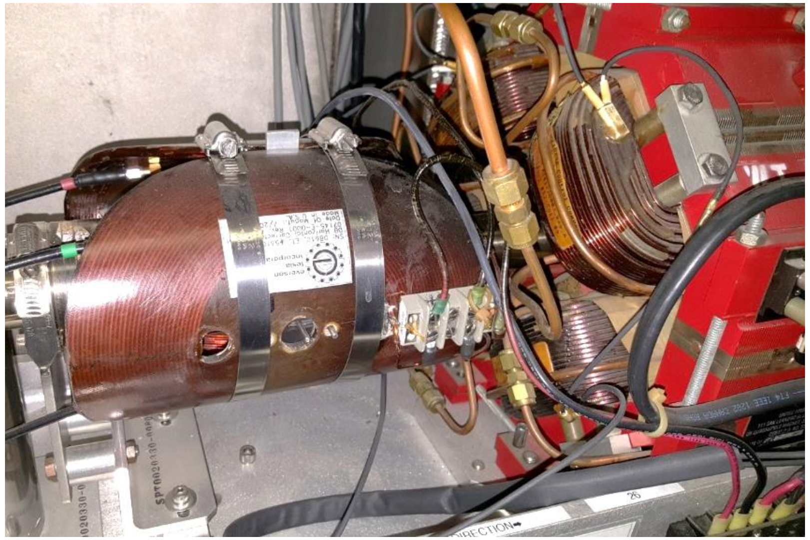
Field Emission

- field emission (FE) is process where electrons are emitted from a cavity wall when experiencing high electric field
- FE electrons can be captured by CEBAF cavities and accelerated over long distances to produce radiation, excess heat load, and vacuum spikes
- radiation can damage and activate beamline components

radiation area

damaged beamline valve

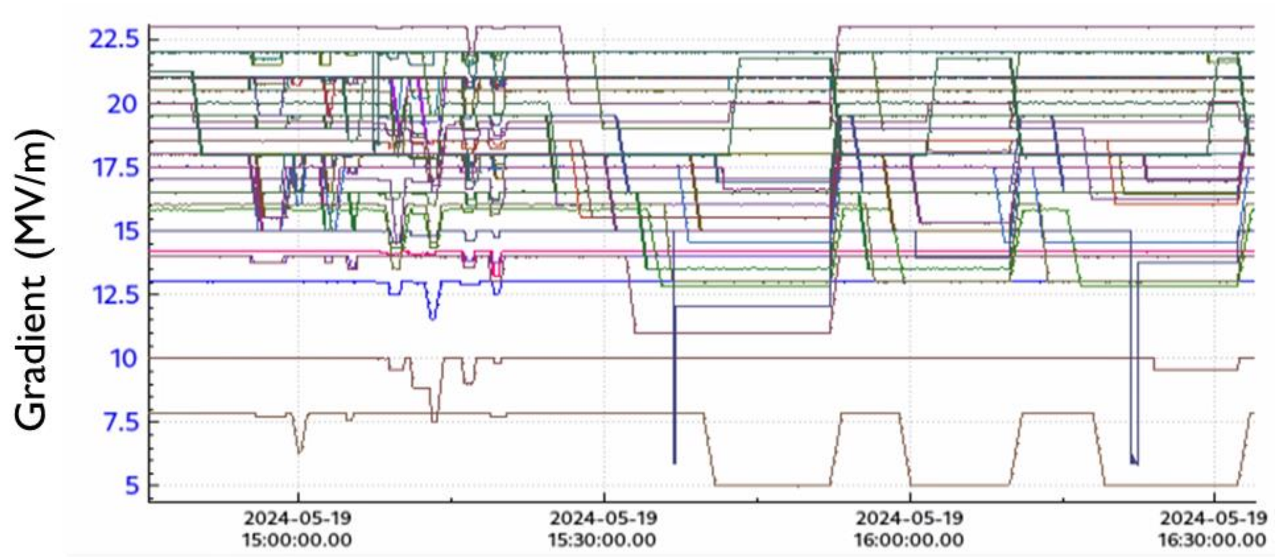
damaged magnet and cables

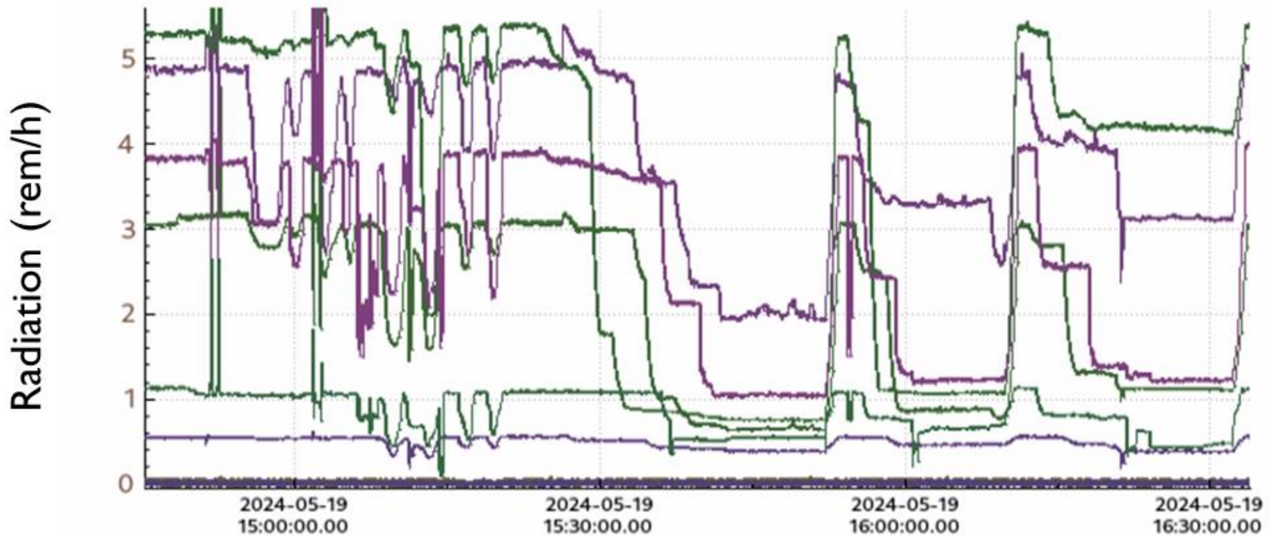


Jefferson Lab

Field Emission Management: Data Collection

- invasive data collection software
  - ✓changes gradients to effect radiation response
  - ✓change random combinations of cavities
    - *sample from all cavities or every n^{th} cryomodule*
  - ✓change individual cavities
  - ✓typical range is [-4, +0.5] MV/m from starting point
  - ✓monitors linac systems and pauses operation if anything goes wrong
- typically takes 2-4 hours, as little as 30 minutes is useful if no problems
  - ✓RF and/or cryo problems can slow progress



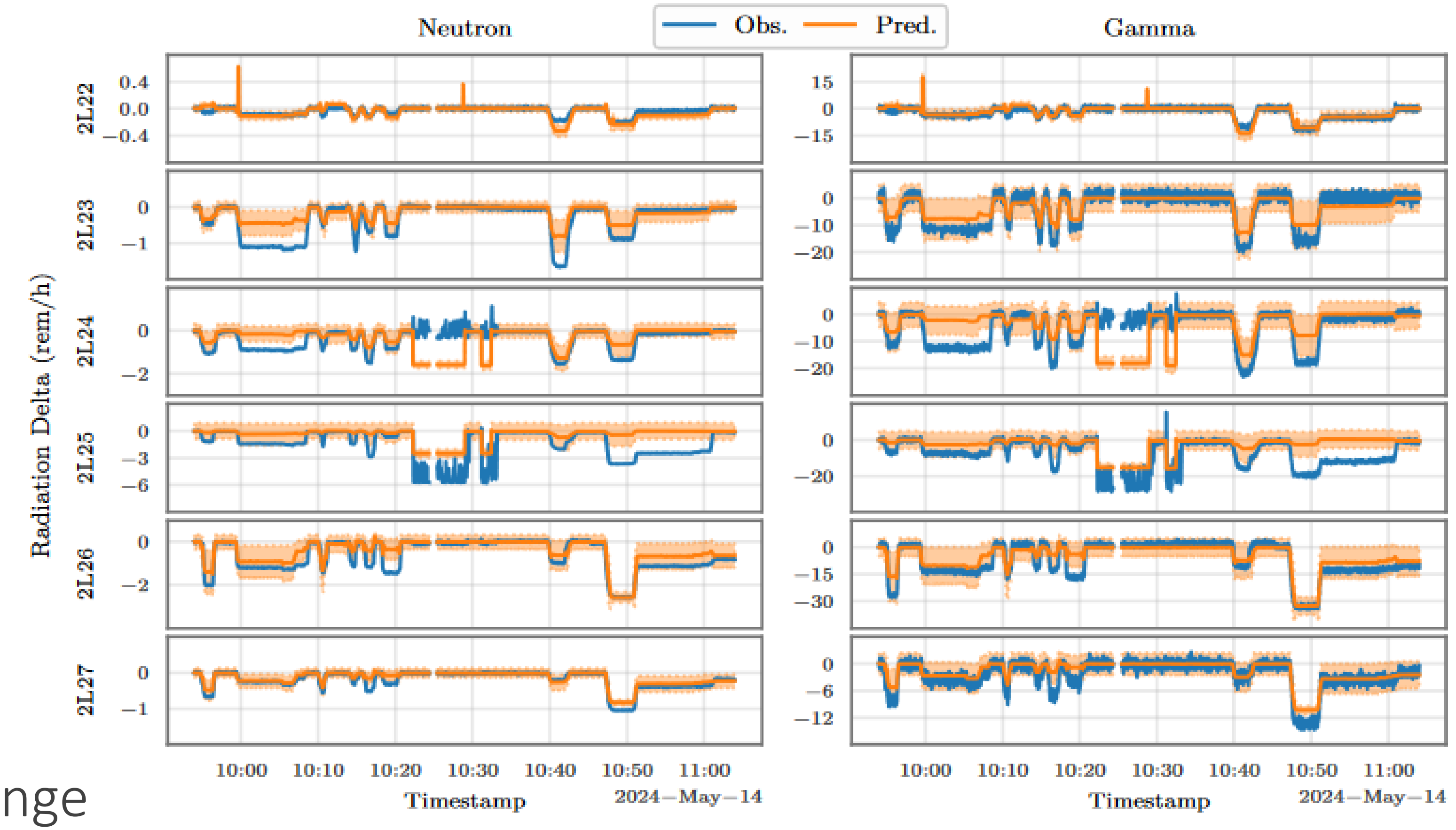


Jefferson Lab

Field Emission Management: Models

- train one model per NDX detector
  - ✓inputs are cavity gradients from cryomodules neighboring the detector
  - ✓simultaneous quantile regression neural network model
  - ✓prediction error estimated by difference between 84th and 16th quantile (one sigma for Gaussian)
  - ✓overall predictions are reasonable, but sometimes misses individual contributions
- collecting sufficient data is a challenge
  - ✓huge sample space (2^{200} = 1.6e60)
  - ✓using neighboring cryomodules is a trade off
  - ✓training, validation, or test sets cover small portion of possible gradient distributions

Radiation Changes From Baseline



Jefferson Lab

Optimization Problem

- objectives:
  - ✓minimize total neutron radiation
  - ✓minimize total gamma radiation
  - ✓minimize total neutron uncertainty
  - ✓minimize total gamma uncertainty
  - ✓maximize linac energy

$$\begin{aligned} \underset{G}{\text{minimize}} \quad & R_n(G), R_\gamma(G), U_n(G), U_\gamma(G), -E(G) \\ \text{subject to} \quad & |E(G) - E^*| \le \epsilon, \\ & g_i^l \le g_i \le g_i^u, \quad \forall g_i \in G. \end{aligned}$$

- constraints:
  - ✓require linac energy within tolerance (±0.5 MeV)
  - ✓bound cavity changes based on real operational constraints
  - ✓limit gradients to within [-3, +0.5] MV/m of initial setting to mitigate OOD effects
- NSGA2 handles all of above
  - ✓greedy algorithm only focuses minimizing total neutron radiation, but respects constraints
- minimizing RF trip rate is an obvious new objective
- uncertainty objectives have had mixed effectiveness depending on the UQ technique

Field Emission Management: Optimization Software

- loads a historical CEBAF configuration
- ML-model provides radiation and uncertainty estimates for objectives
- supports multiple optimizers
- runs optimization for user-specified number of iterations
- lets user investigate family of different solutions
- provides shell script to apply a given solution

Jefferson Lab

Field Emission Management: Demonstrations

- six gradient distributions generated and applied
  - ✓ three models trained on different ranges of historical data
  - ✓ two different optimization strategies
- March 29 - May 19 model demonstration built on solution of May 9 - NSGA2 run
- gradient distributions only consider field emission and energy gain
  - ✓ no trip rates, klystron power limits, energy lock cavities, etc.

Model	Optimizer	E_{gain} Delta (MeV)	Neutron Start (rem/h)	Neutron Change (rem/h)	Gamma Start (rem/h)	Gamma Change (rem/h)	Arc Trip Rate (Trips/Hour)***
March29-May14	Greedy	-23.59*	19.01	**-13.61 (-72%)**	288.81	-163.75	6.93
March29-May14	Greedy	-3.63*	19.01	**-11.27 (-59%)**	288.81	-117.67	6.93
May9	NSGA2	-4.08*	19.01	**-8.53 (-45%)**	288.81	-112.55	6.92
May9	Greedy	-3.70*	19.01	**-8.05 (-42%)**	288.81	-107.01	6.92
March29-May19	NSGA2	-0.49	10.93	**-5.45 (-50%)**	183.30	-57.24	11.57
March29-May19**	NSGA2	-4.47	19.01	**-13.98 (-74%)**	288.81	-169.80	11.57

**Bypassed cavity resulted in ~3 MeV of unexpected energy loss*

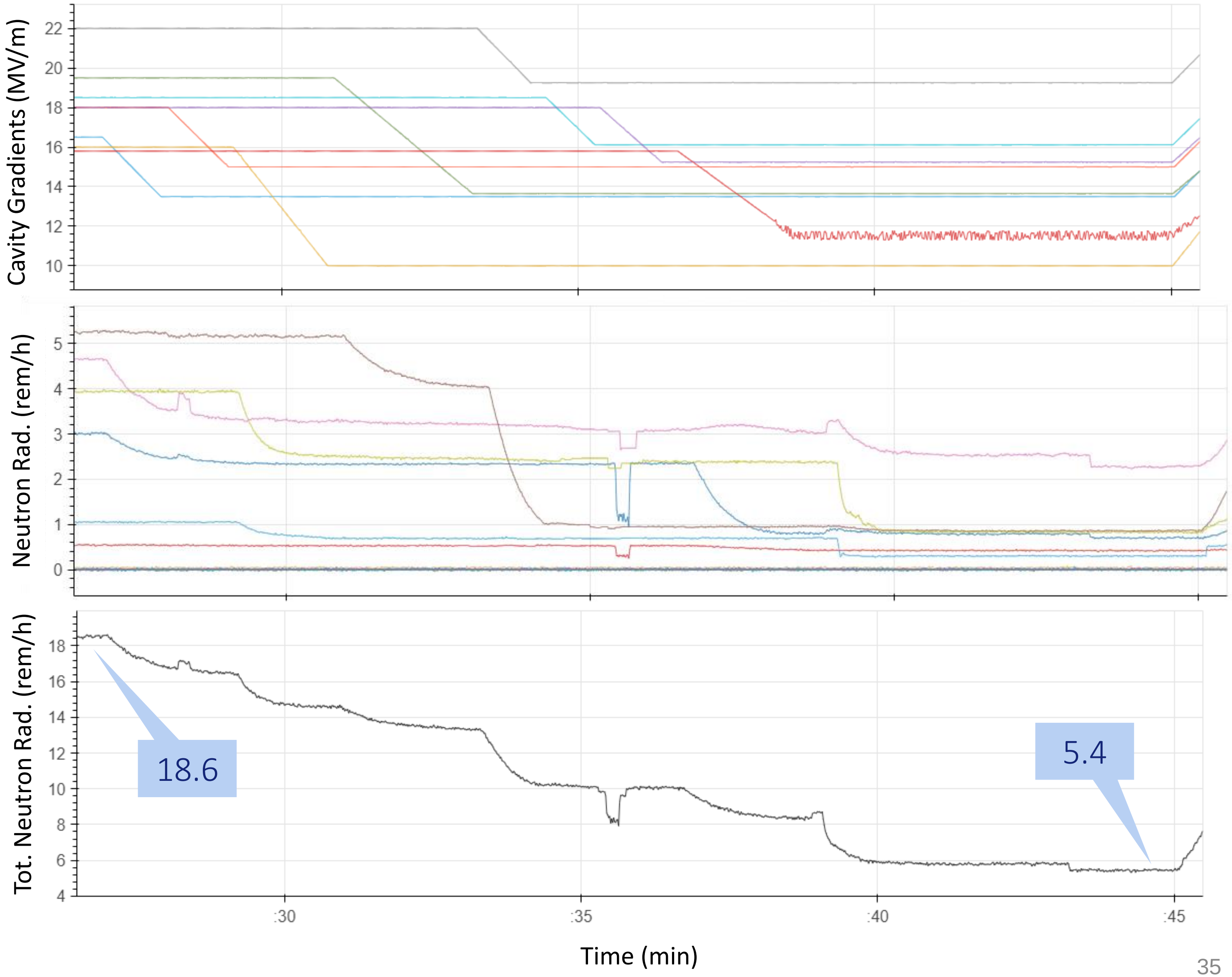
***Includes effects of May9-NSGA2 and March29-May19 combined*

****Baseline trip rate 3.90 trips/hour*

Field Emission Management: Demonstration

optimized solution found *offline* – no dedicated beam studies required

13 rem/h decrease in neutron radiation



not shown: cavity gradients that were increased to maintain linac energy

not shown: 145 rem/h reduction in gamma radiation

Jefferson Lab

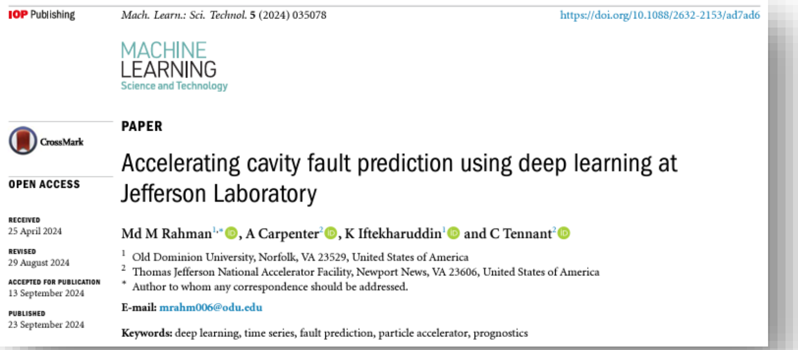
SUMMARY

Anomalous Cavity Detection

- deployed in CEBAF and demonstrated to be a robust solution moving forward
- because we train new models every day using the latest 7 days of data, our framework automatically adapts to domain drift
- proposal submitted to management to fund additional DAQs to provide coverage of legacy cryomodules in SL
  - ✓ would provide fast-sampled RF signals from 94% of the North and South Linacs
- prepared manuscript *"Detecting Anomalous SRF Cavity Behavior with Unsupervised Learning"*
  - ✓ submitted to journal, under review
- H. Ferguson to defend his dissertation in Spring 2025

Fault Prediction

- achieved a proof-of-principle demonstration that has important implications for future RF system design
- prepared manuscript *"Accelerating Cavity Fault Prediction Using Deep Learning at Jefferson Lab"*
  - ✓ published in the **Journal of Machine Learning: Science and Technology**


IOP Publishing
Mach. Learn.: Sci. Technol. **5** (2024) 035078
https://doi.org/10.1088/2632-2153/ad7ad6
MACHINE LEARNING Science and Technology

CrossMark

OPEN ACCESS

RECEIVED
25 April 2024
REVISED
29 August 2024
ACCEPTED FOR PUBLICATION
13 September 2024
PUBLISHED
23 September 2024

PAPER

Accelerating cavity fault prediction using deep learning at Jefferson Laboratory

Md M Rahman[1,*], A Carpenter[2], K Iftekharuddin[1] and C Tennant[2]
1 Old Dominion University, Norfolk, VA 23529, United States of America
2 Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, United States of America
* Author to whom any correspondence should be addressed.
E-mail: mrahm006@odu.edu

Keywords: deep learning, time series, fault prediction, particle accelerator, prognostics

- Md. M. Rahman to defend his dissertation spring 2025

Jefferson Lab

Field Emission Management

- demonstrated the ability to significantly reduce FE radiation in an automated manner
- presented to Operations as a viable solution moving forward
  - ✓ note, as with most of these applications, there is significant overhead in implementing and maintaining this systems to operational standards beyond the initial proof-of-principle demonstration
- prepared manuscript *"Data-Driven Gradient Optimization for Field Emission Management in a Superconducting Radio-Frequency Linac"*
  - ✓ submitted to journal and under review

arXiv > physics > arXiv:2411.07018 Search... Help

Physics > Accelerator Physics

[Submitted on 11 Nov 2024]

Data-Driven Gradient Optimization for Field Emission Management in a Superconducting Radio-Frequency Linac

Steven Goldenberg, Kawser Ahammed, Adam Carpenter, Jiang Li, Riad Suleiman, Chris Tennant

Field emission can cause significant problems in superconducting radio-frequency linear accelerators (linacs). When cavity gradients are pushed higher, radiation levels within the linacs may rise exponentially, causing degradation of many nearby systems. This research aims to utilize machine learning with uncertainty quantification to predict radiation levels at multiple locations throughout the linacs and ultimately optimize cavity gradients to reduce field emission induced radiation while maintaining the total linac energy gain necessary for the experimental physics program. The optimized solutions show over 40% reductions for both neutron and gamma radiation from the standard operational settings.

https://doi.org/10.48550/arXiv.2411.07018

- K. Ahammed to defend his dissertation ~2026

Jefferson Lab

Project Summary: Major Deliverables and Schedule

This proposal has five primary objectives:

- Develop an online AI model to predict an impending cavity fault in C100 cryomodules from streaming data
2. Perform data mining on saved (minutes-long) continuous RF data streams from C100 cavities to extract useful, but as yet unidentified, information for CEBAF operations
- Develop AI models to minimize radiation levels due to field emission in the Injector and each of CEBAF's SRF linacs by optimizing cavity gradient distribution, subject to operational constraints (e.g., beam energy and trip rate)
- Develop a general purpose DAQ for streaming continuous RF signals from the 40 legacy C20/C50 cryomodules
- Develop an AI model to automatically detect transient SRF cavity instabilities in C20/C50 cavities

Project	Deliverable	Date
Cavity Instability Detection	Publish manuscript, prepare framework for next operational run	01/2025
C100 Fault Prediction	Demonstrate domain-adaptation	01/2025
Field Emission Management	Publish manuscript	01/2025

Project Summary: Annual Budget

	FY 2020	FY2021	FY2022	Total
a) Funds allocated	$450,000	$450,000	$450,000	$1,350,000
b) Actual costs to date	$450,000	$450,000	$437,564	$1,337,564
c) Uncosted commitments	$0	$0	$0	$0
d) Uncommitted funds (d=a-b-c)	$0	$0	$12,436	$12,436

Thank you.

This work is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Contract No. DE-AC05-06OR23177.