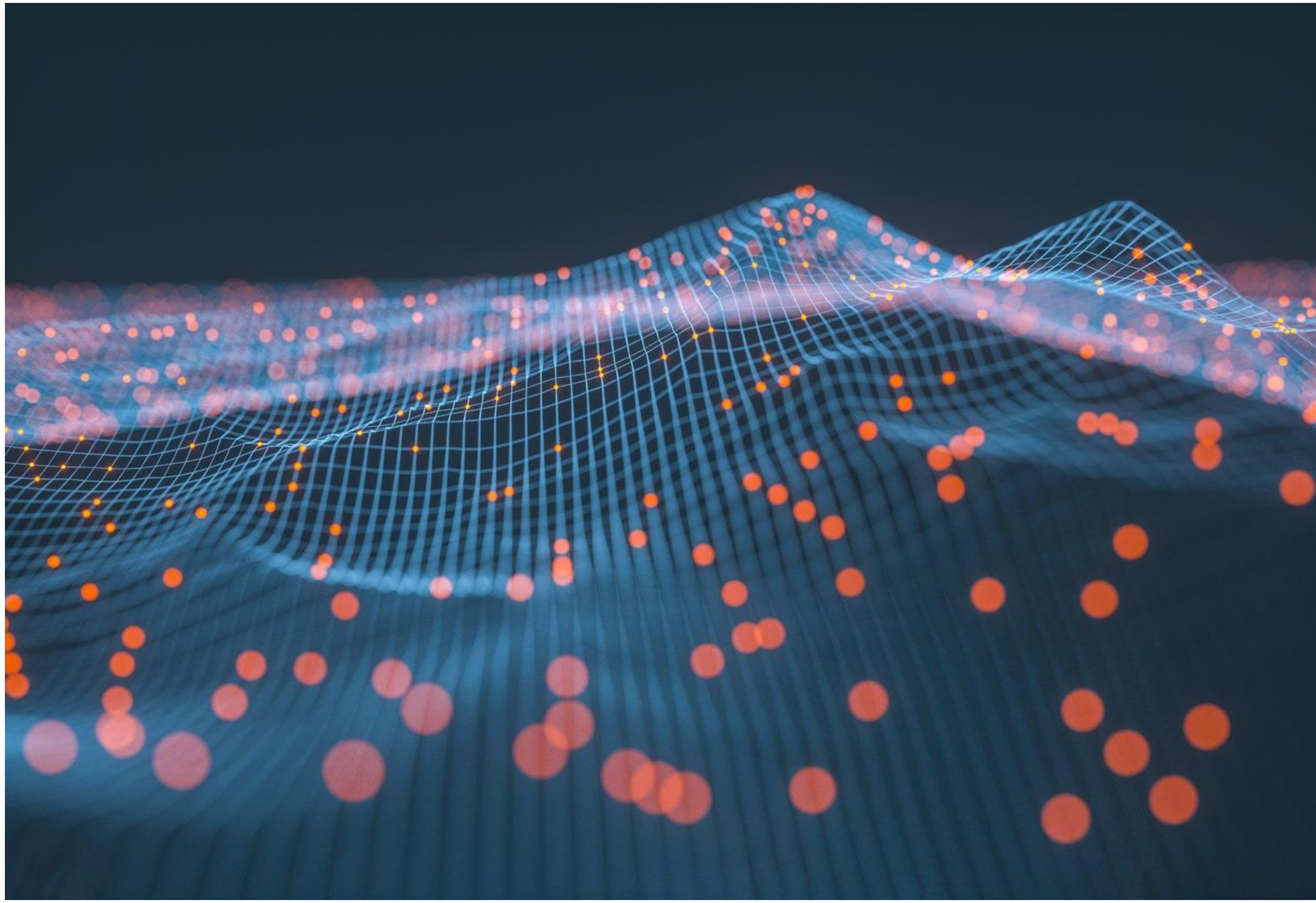
# AI for Optimized SRF Performance of CEBAF Operations

Chris Tennant

*for the Jefferson Laboratory Team*

DOE PI AI/ML Exchange Meeting | November 30, 2022



Jefferson Lab

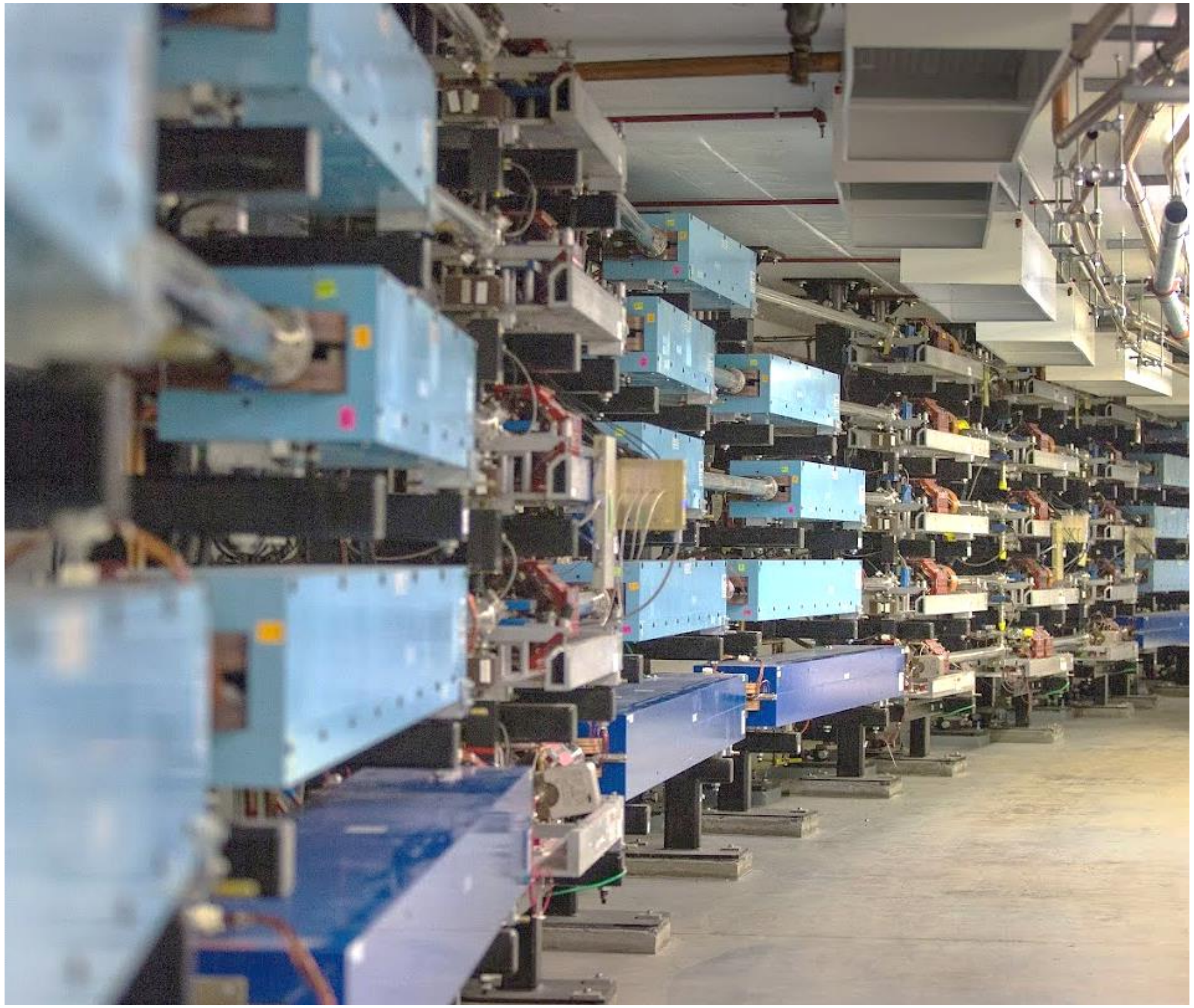
OLD DOMINION UNIVERSITY

U.S. DEPARTMENT OF ENERGY | Office of Science

JSA

# Outline

- **Jefferson Laboratory**
- **FOA LAB 20-2261: Year 2 Status**
  - ✓ *Cavity Instability Detection*
  - ✓ *C100 Fault Prediction*
  - ✓ *Field Emission Management*
- **Project Summary**
  - ✓ *Deliverables and Schedule*
  - ✓ *Budget*

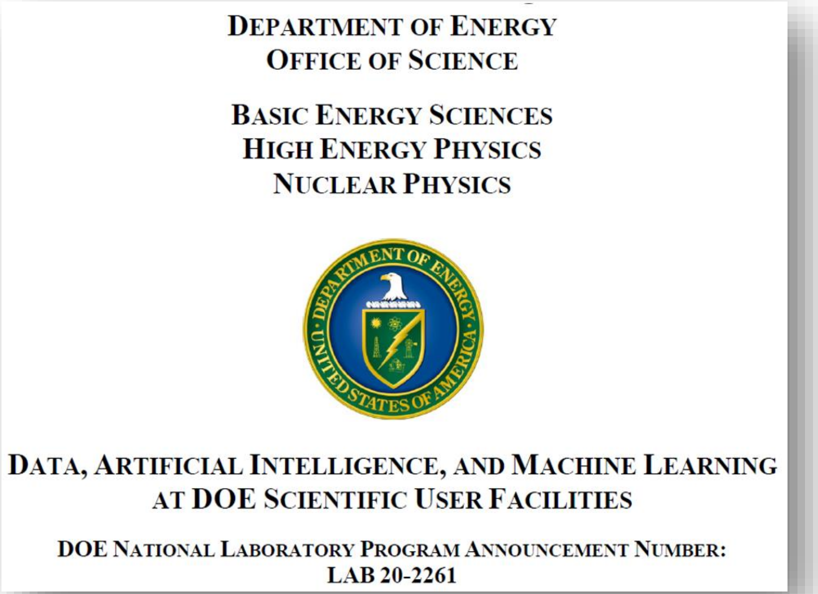


# “AI for Optimized SRF Performance of CEBAF Operations”

The project builds on a recent successful effort at Jefferson Lab to implement AI at CEBAF and seeks to extend the work for optimizing SRF operations. Specifically, the proposal presents a multi-faceted approach to:

A. develop tools to automate cavity instability detection
B. provide real-time fault prediction for C100 cavities
C. 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

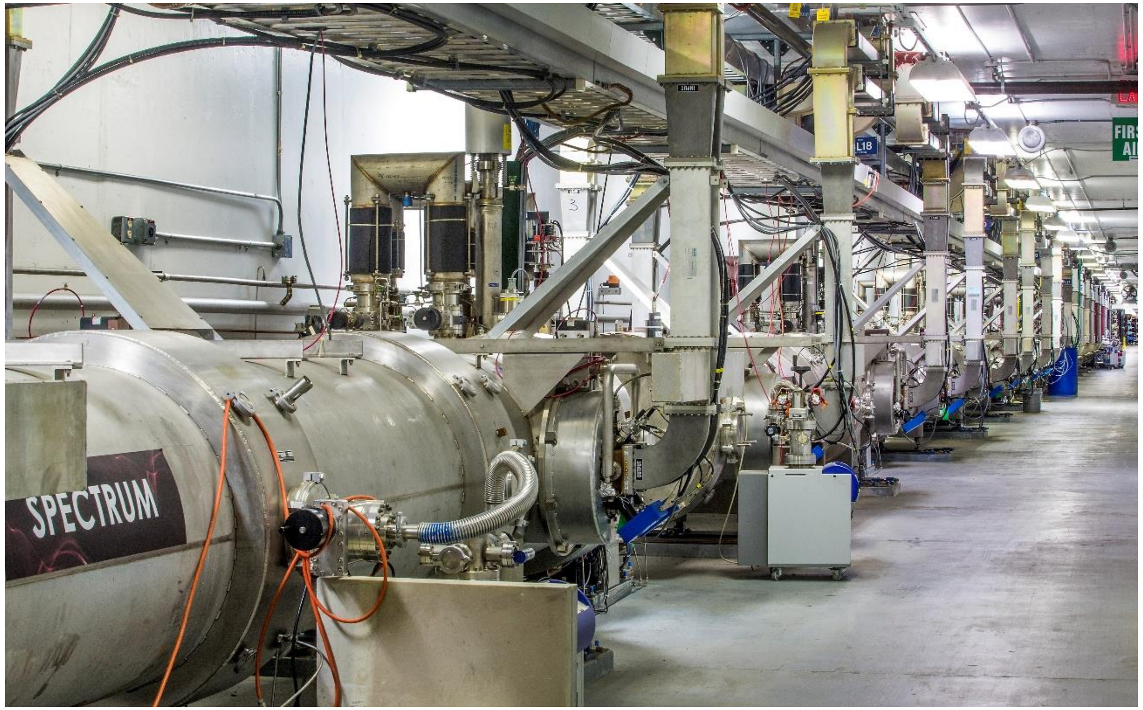
BASIC ENERGY SCIENCES
HIGH ENERGY PHYSICS
NUCLEAR PHYSICS

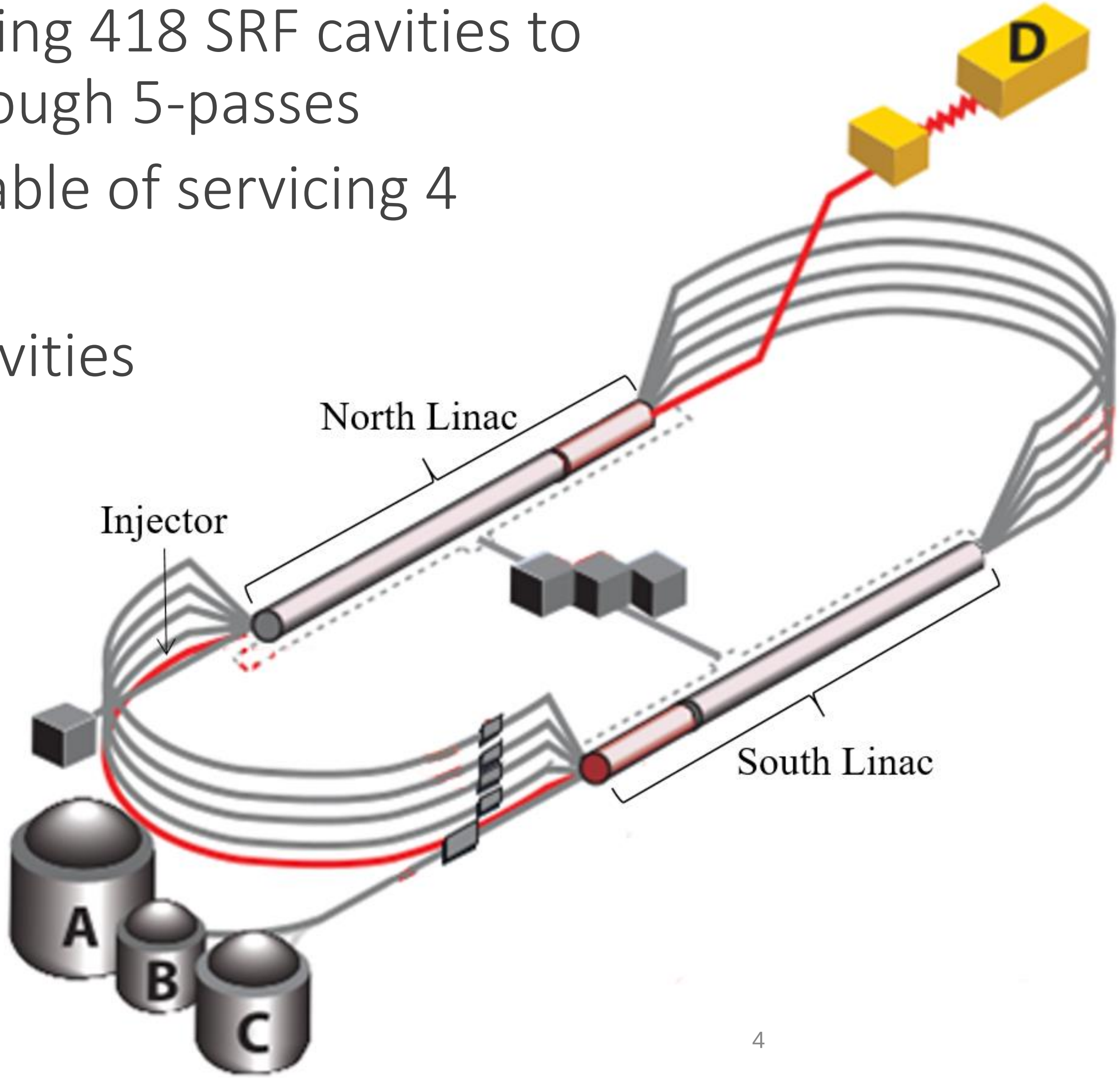
DATA, ARTIFICIAL INTELLIGENCE, AND MACHINE LEARNING AT DOE SCIENTIFIC USER FACILITIES

DOE NATIONAL LABORATORY PROGRAM ANNOUNCEMENT NUMBER: LAB 20-2261

# Continuous Electron Beam Accelerator Facility

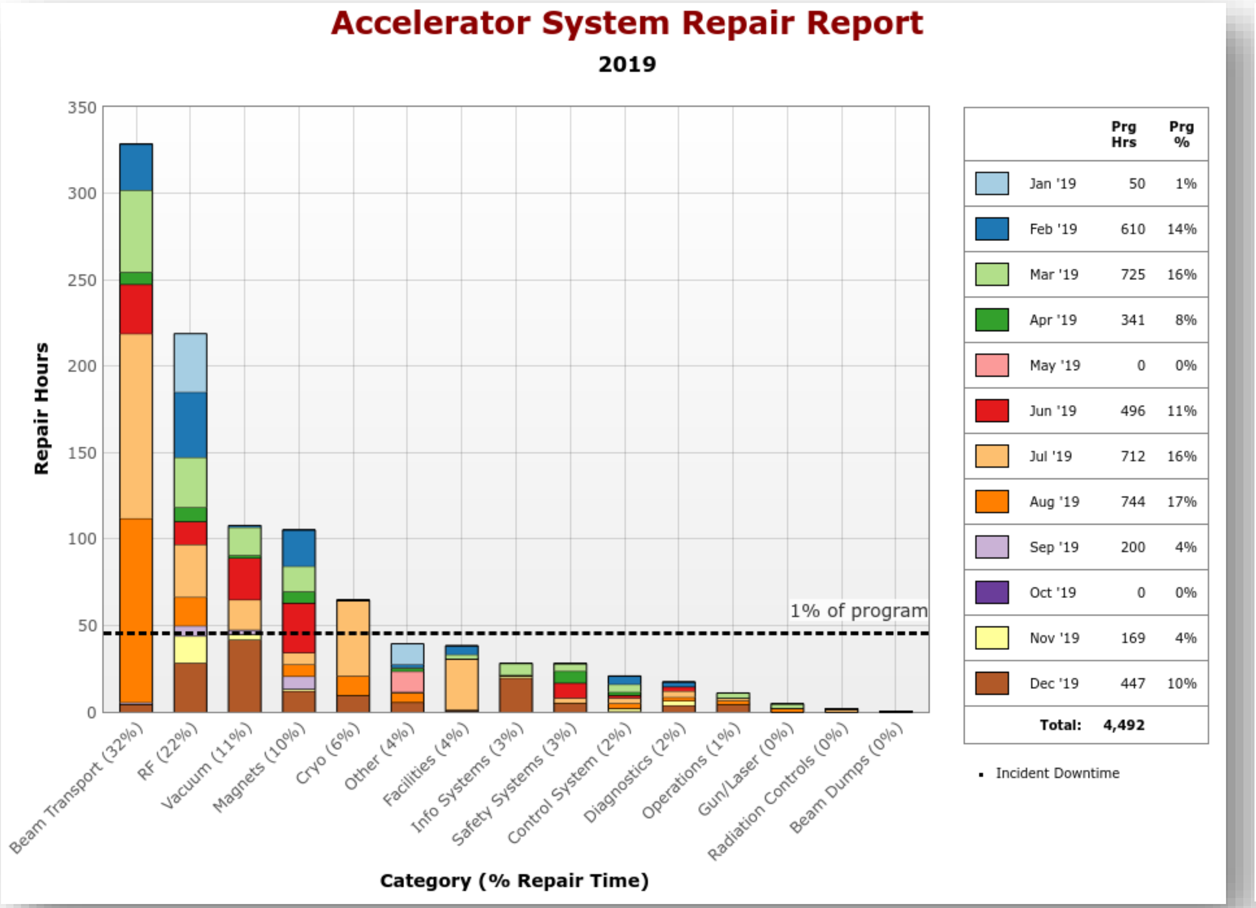
- CEBAF is a CW recirculating linac utilizing 418 SRF cavities to accelerate electrons up to 12 GeV through 5-passes
- it is a nuclear physics user-facility capable of servicing 4 experimental halls simultaneously
- the heart of the machine is the SRF cavities

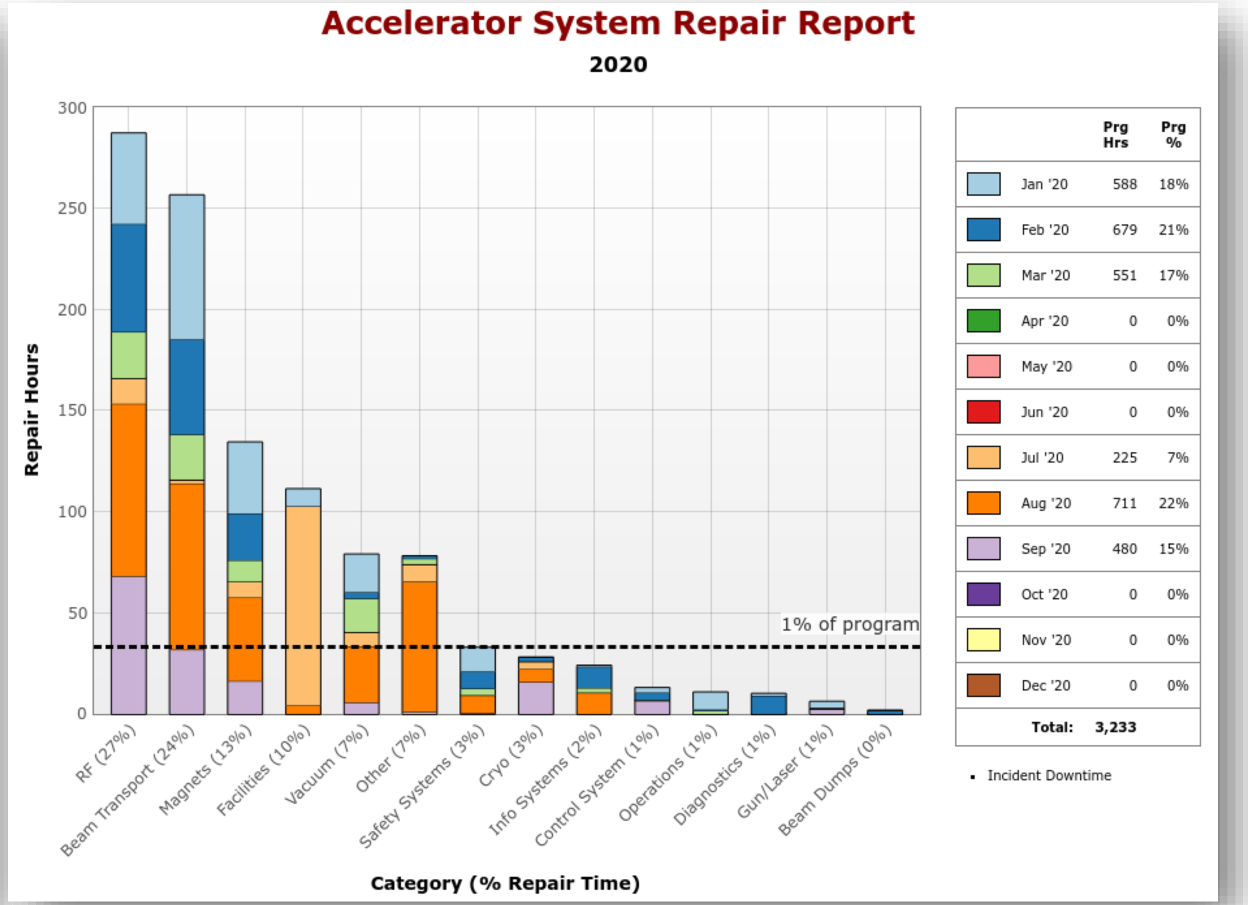


# CEBAF Down Time Manager

- RF related issues are consistently one of the biggest contributors to downtime



# PROJECT A

PI: Dennis Turner

Graduate Student: Hal Ferguson (ODU)

# Project A: Cavity Instability Detection

- **Goal:**

  automate the process of identifying unstable SRF cavities

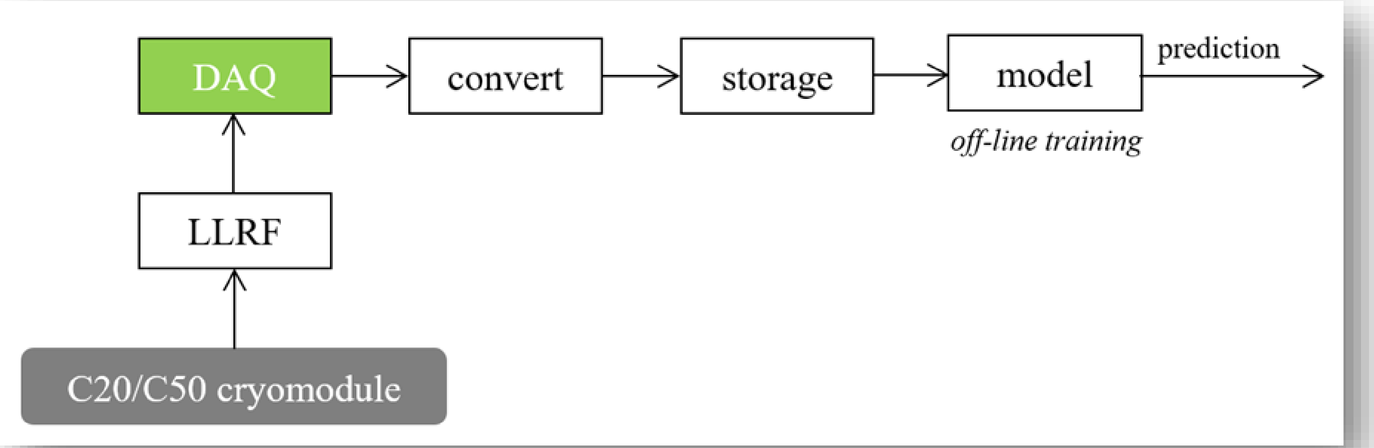
- **Description:**

  SRF cavities can become unstable without presenting faults, identifying these unstable cavities with present diagnostics is difficult and time-consuming

- **Solution**:

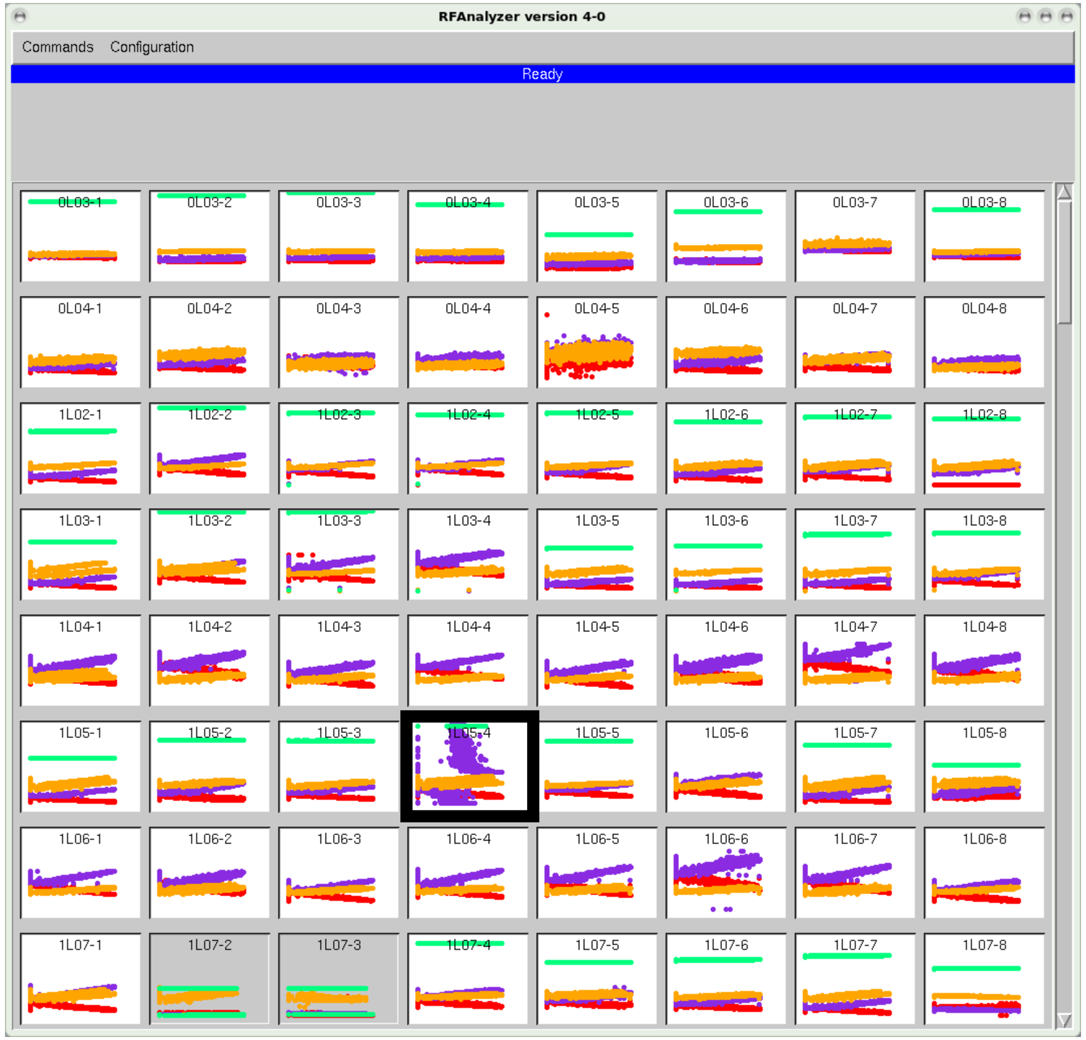
  (1) develop and install a new fast DAQ system for the legacy SRF cavities

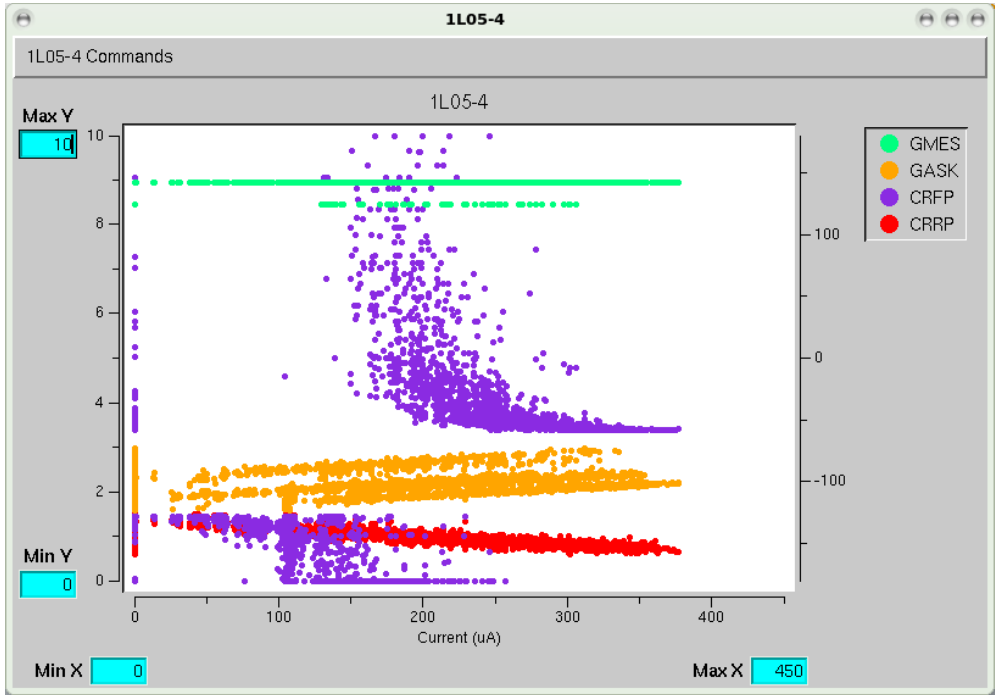
  (2) apply ML to identify unstable cavities



# Cavity Instability Detection: Current Approach

## RF Analyzer Tool



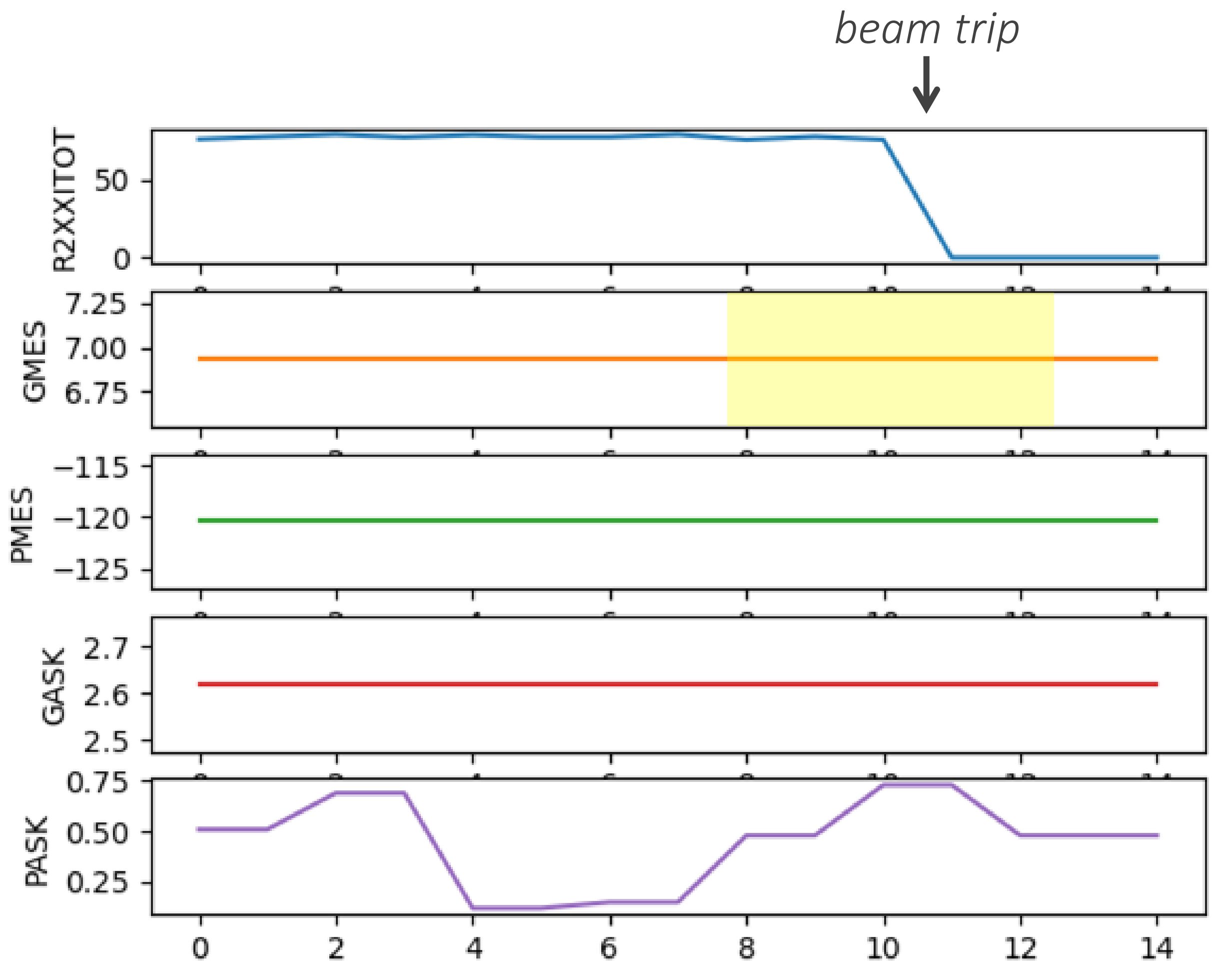

- note, this represents an obvious example
- not all instances are so easily detectable by an operator

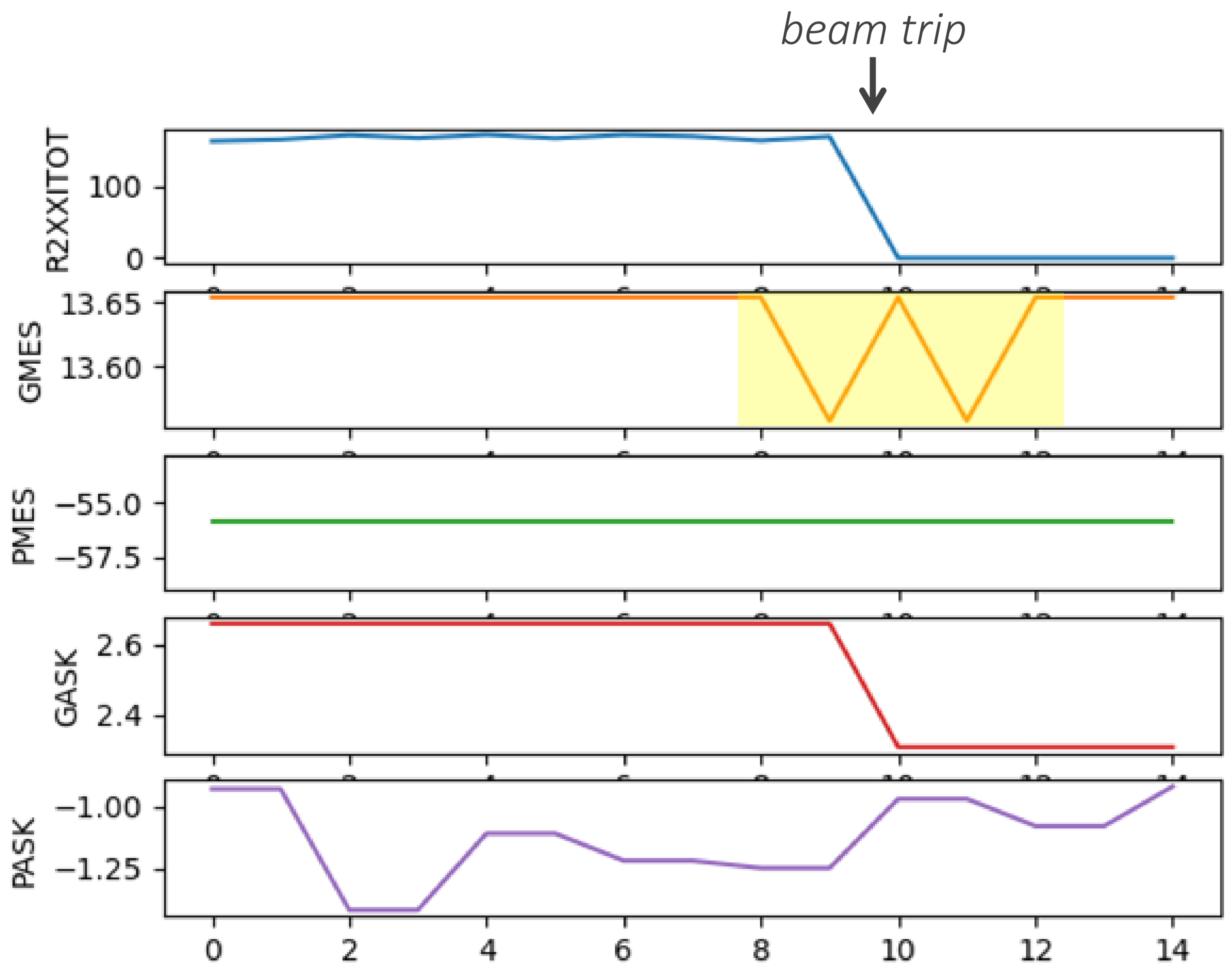


Jefferson Lab

# Cavity Instability Detection: Slow Data

- collect and label "slow" data from the archiver



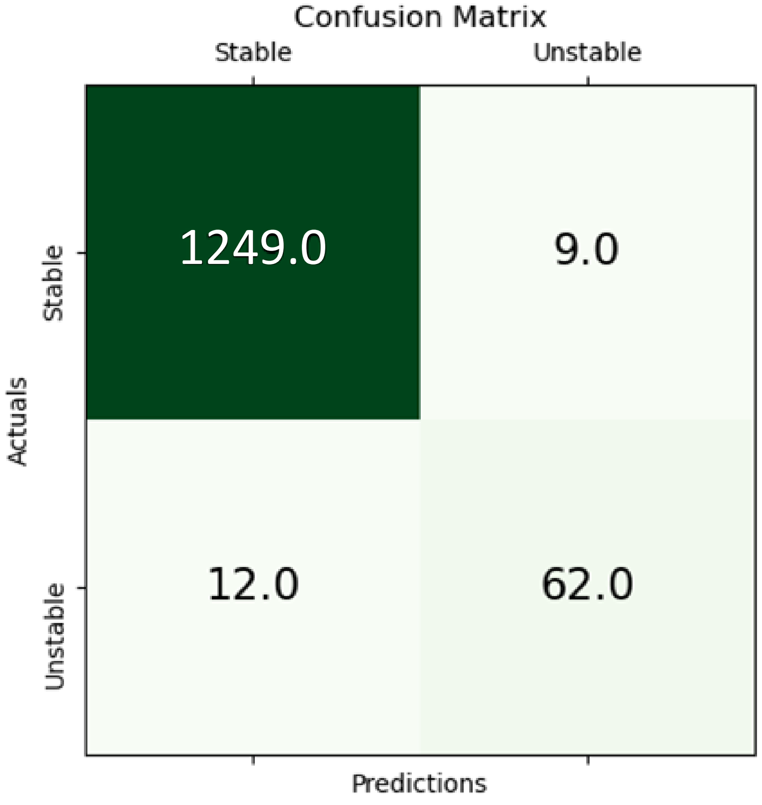



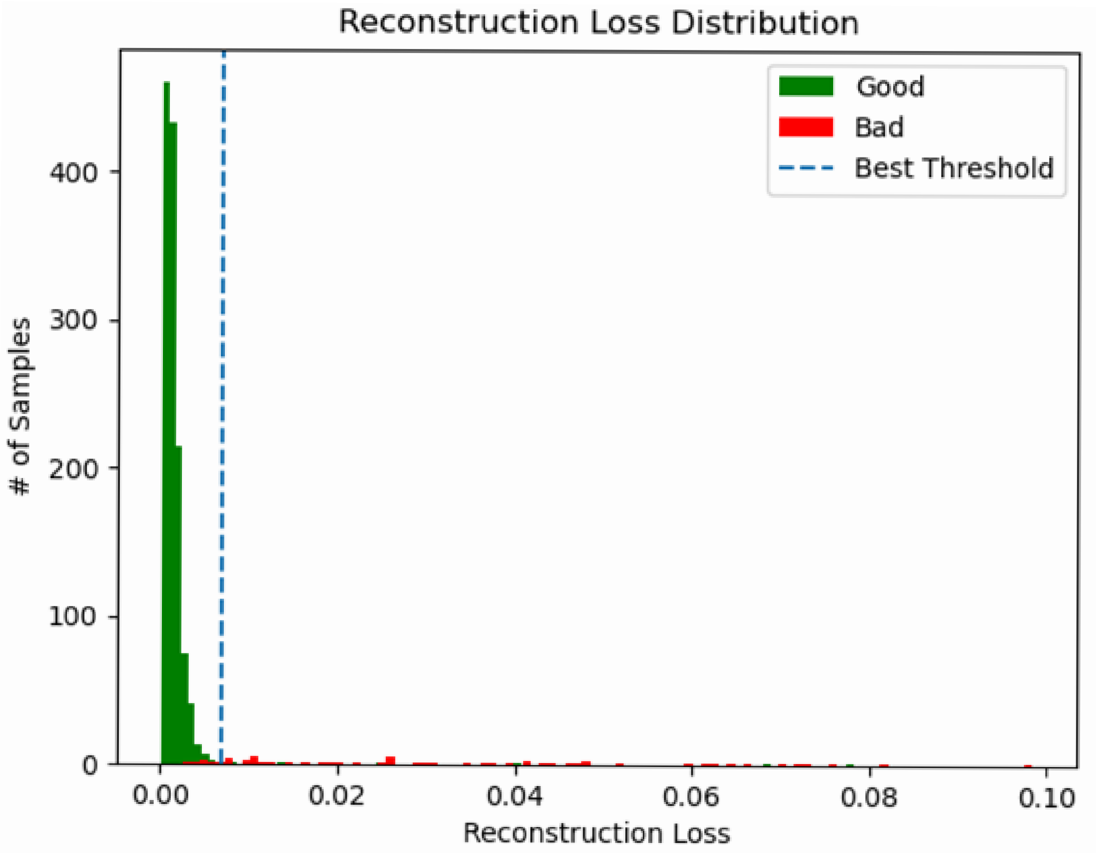
# Cavity Instability Detection: Slow Data

- autoencoder architecture
  - ✓train on “normal” samples, anomalous conditions revealed in poor reconstruction errors

at the chosen threshold:

accuracy = 0.9842
precision = 0.9928
recall = 0.9904
F1-score = 0.9916



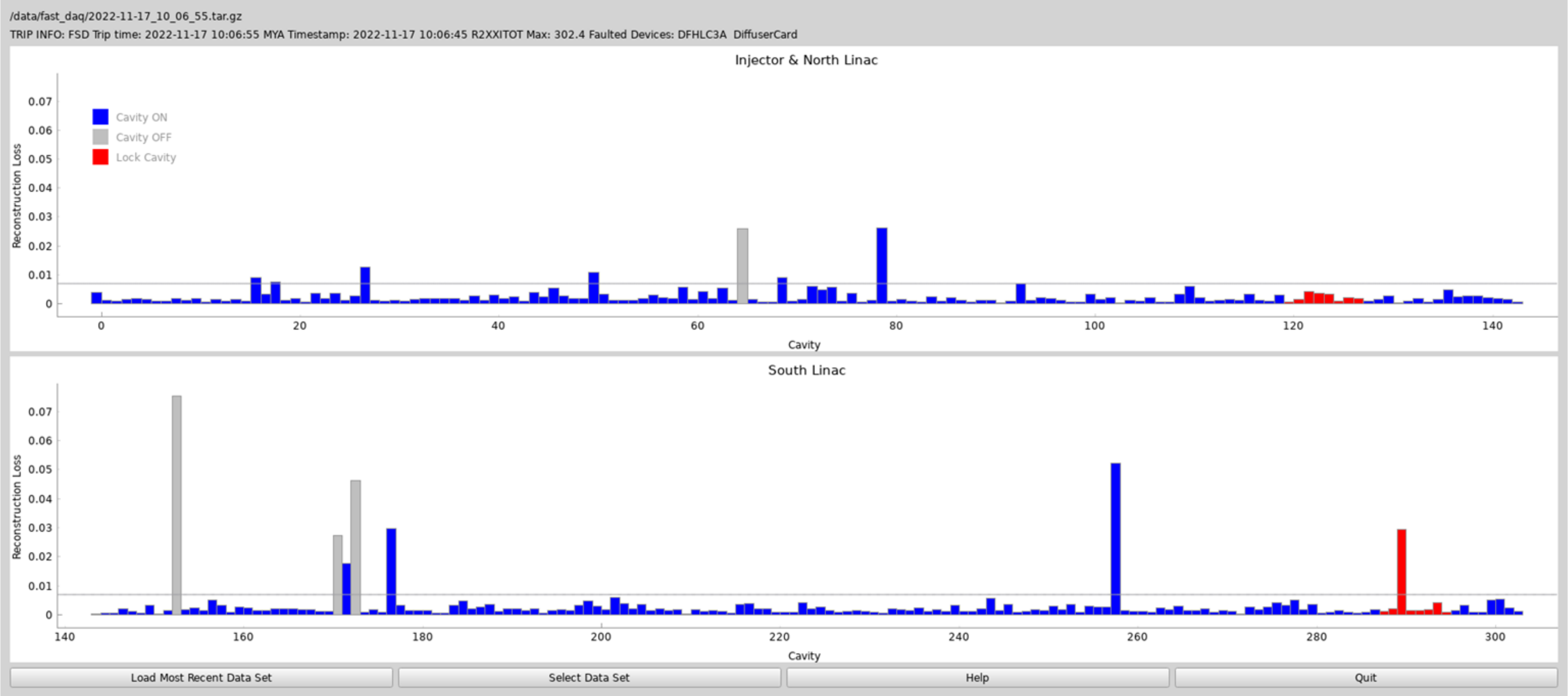





Jefferson Lab

# Cavity Instability Detection: User Interface

- displays reconstruction loss per cavity using archived data for one trip
- higher reconstruction loss → higher likelihood that the cavity presented an instability
- clicking on a bar for a particular cavity opens a plot of the raw archived data



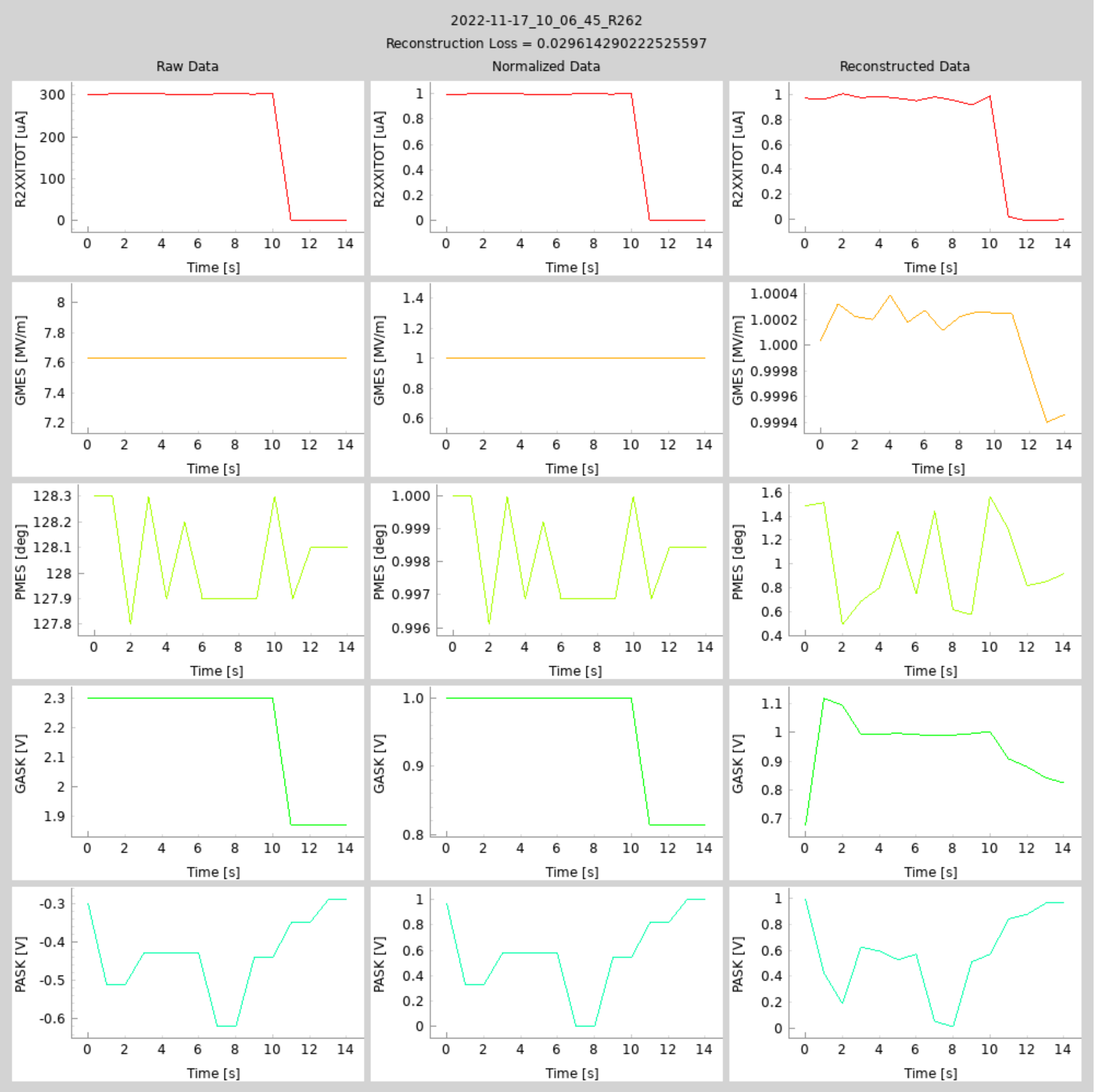


Jefferson Lab

# Cavity Instability Detection: User Interface

- operator interface is nearly ready for deployment
- documentation and installation remain
- an identical interface will be created to display results of the fast DAQ autoencoder model

*note PMES instability around the time of trip →*

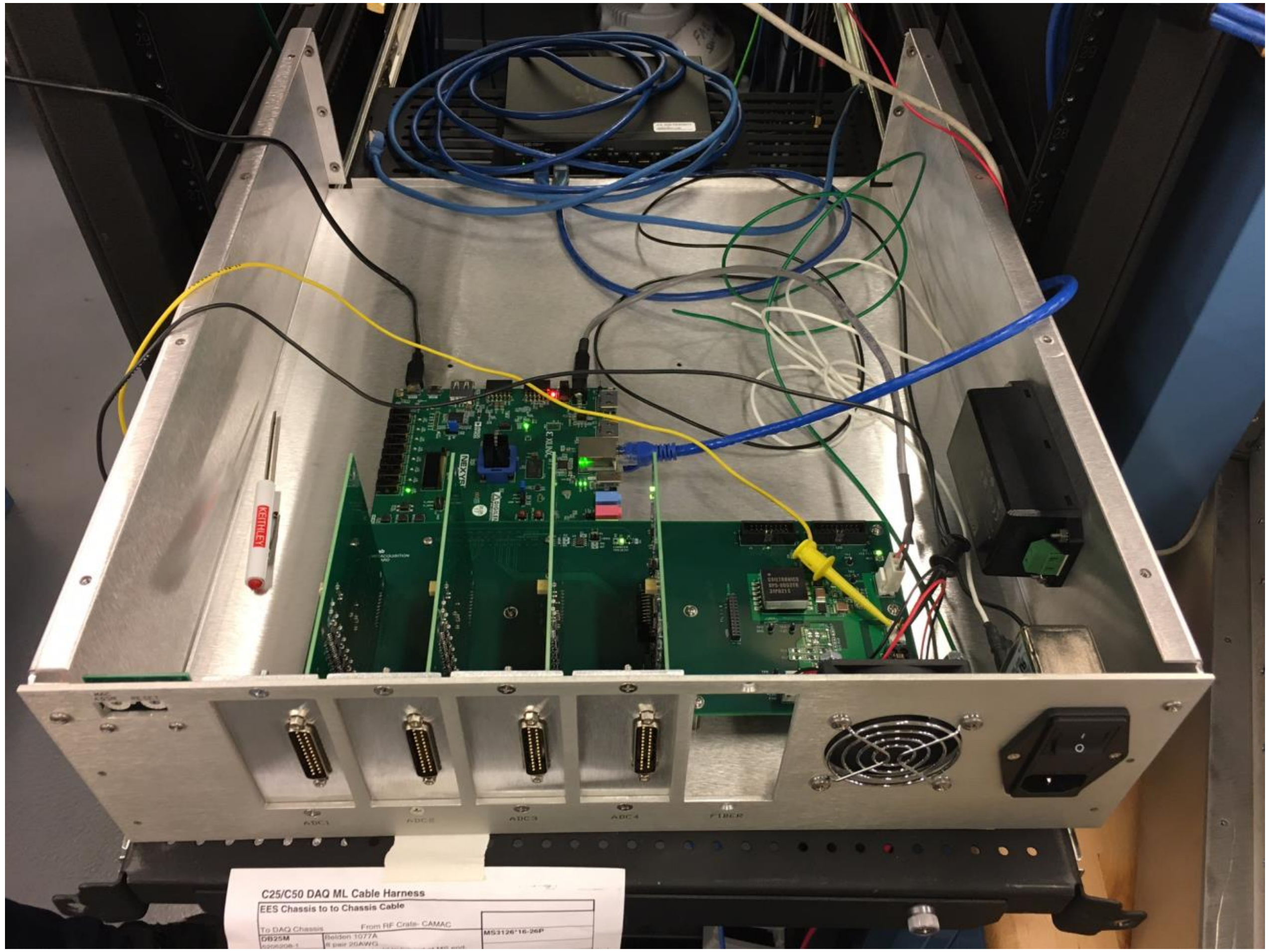




Jefferson Lab

# Cavity Instability Detection: Data Acquisition System (DAQ)

- all parts procured, fabrication and testing in progress
- 20 DAQs for NL (reduced scope due to rising costs)

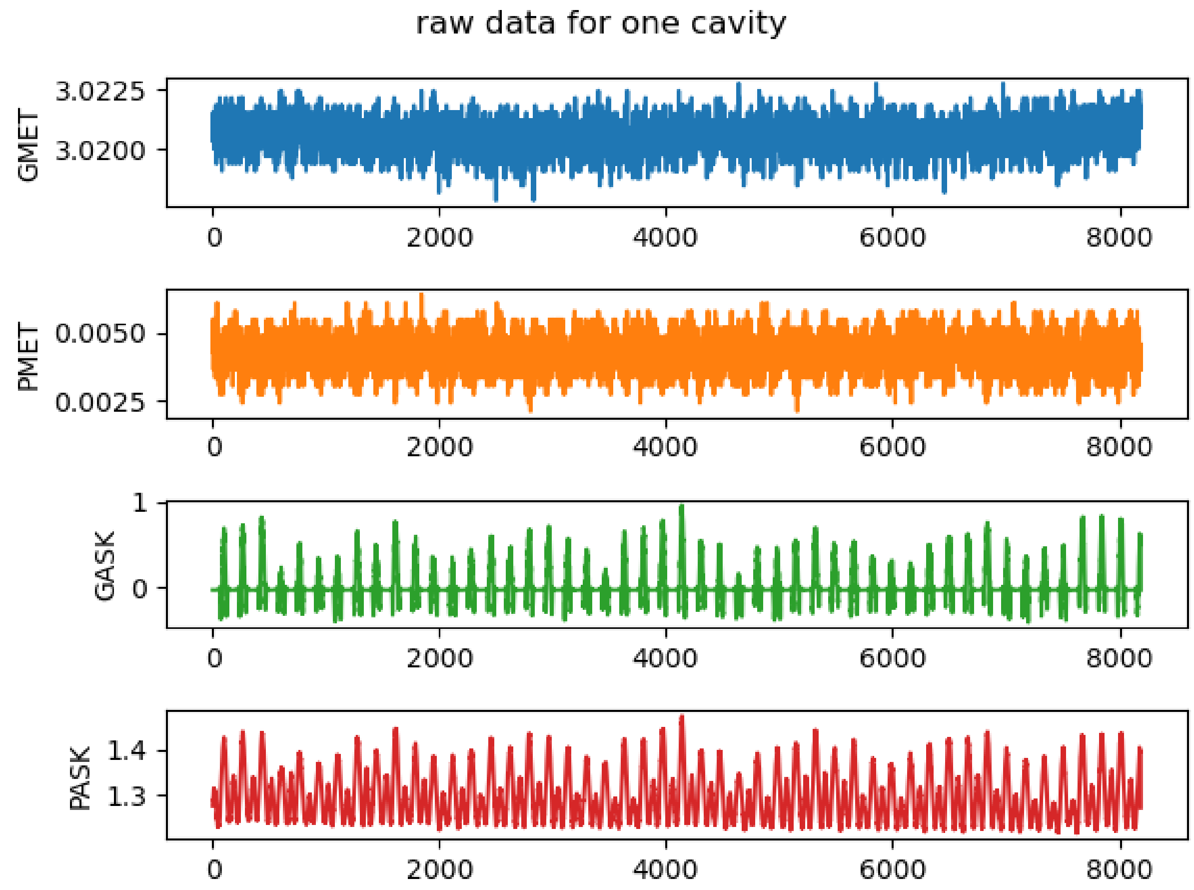


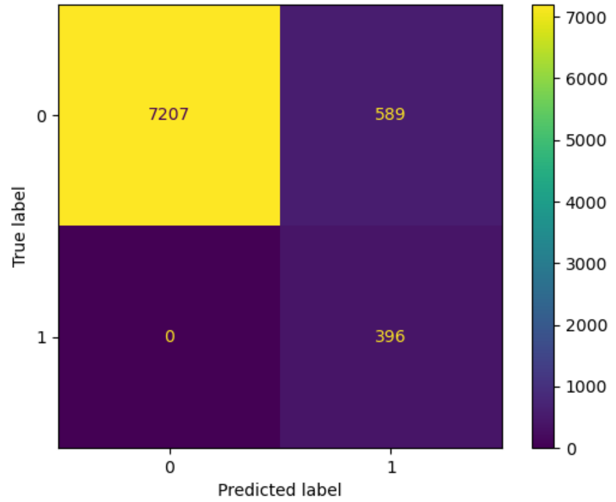


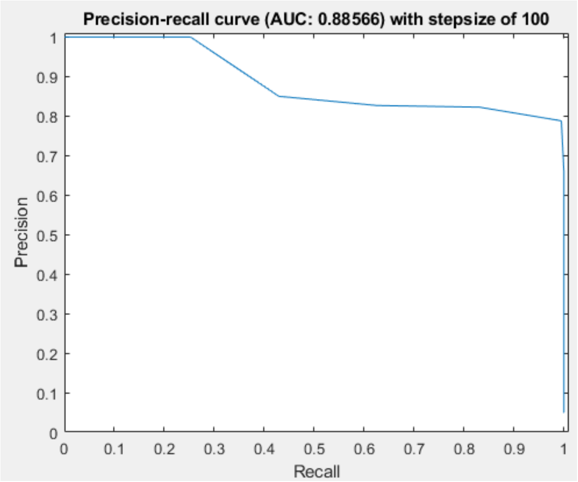

Jefferson Lab

# Cavity Instability Detection: Data Acquisition System (DAQ)

- prototype chassis installed and running software collecting data
- working on autoencoder using this fast data




```
Precision: 0.402
Recall: 1.000
Accuracy: 0.928
F1 Score: 0.573
```


Jefferson Lab

# PROJECT B

PI: Chris Tennant

Graduate Student: Md. Monibor Rahman (ODU)

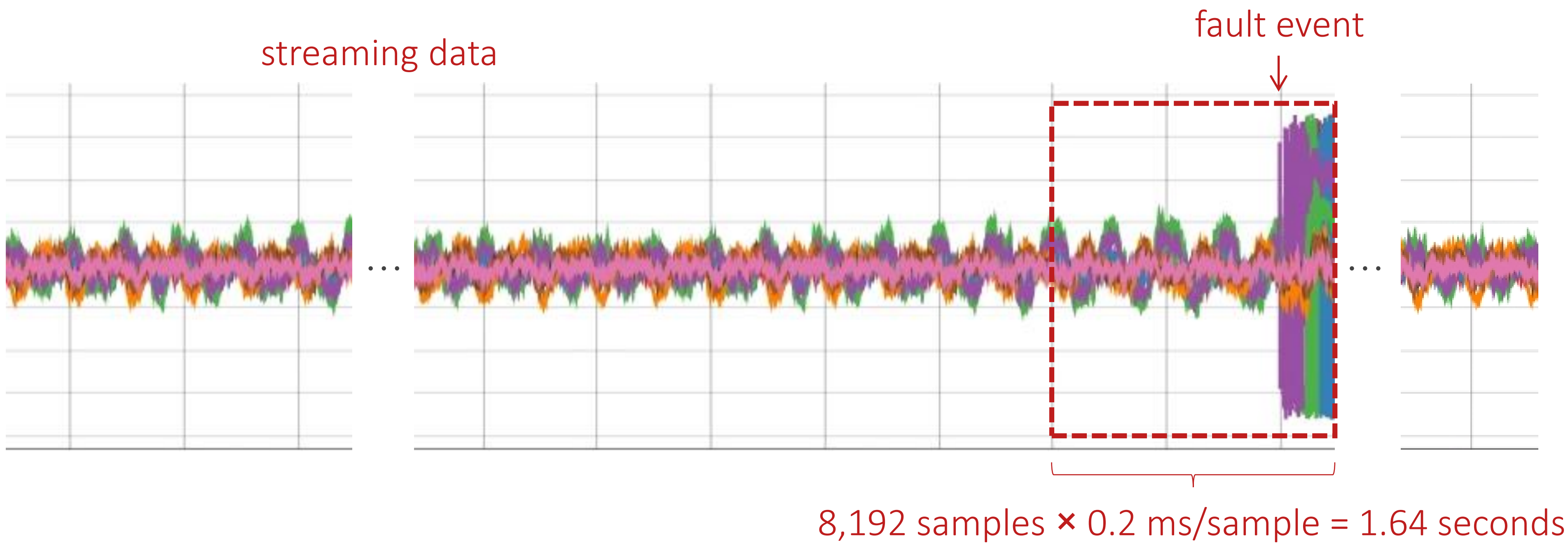
# Project B: C100 Fault Prediction

- **Goal:**

  Proactively *predict* if a C100 cavity fault will occur

- **Description:**

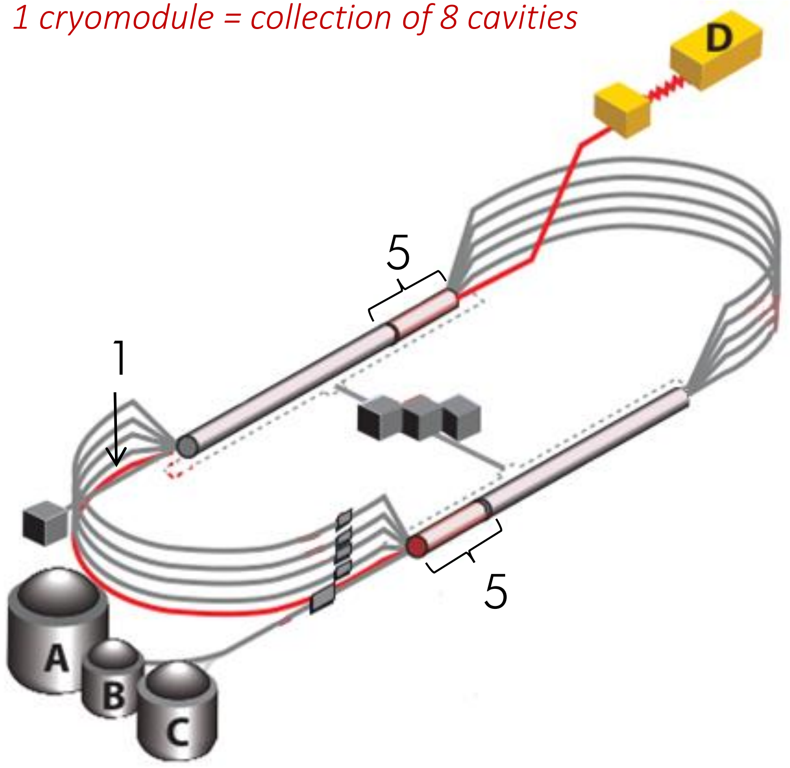
  Currently deployed ML models analyze data *after* a fault has occurred. Investigate the use of machine learning to predict if a fault will occur.



# C100 Fault Isolation and Identification: Present

have the ability to record high-fidelity data from 11 cryomodules

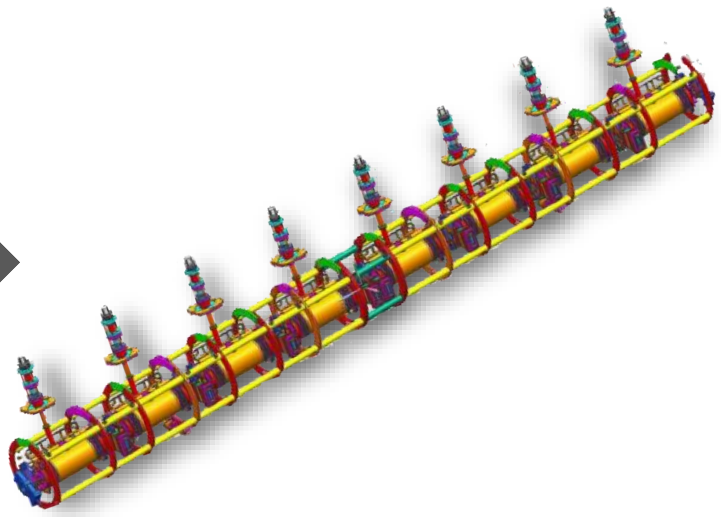
*1 cryomodule = collection of 8 cavities*




**FAULT ISOLATION**

Which of the 8 cavities faulted first?

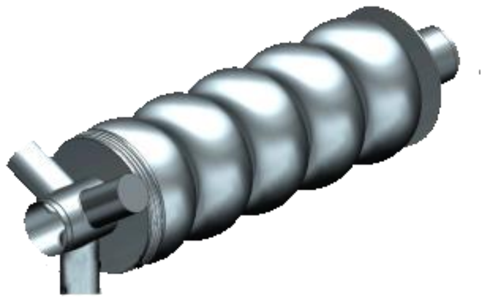
*17 signals/cavity × 8 cavities = 136 signals*



**FAULT IDENTIFICATION**

What kind of trip was it?

*17 signals*



train a model to correctly classify the <u>cavity</u> and <u>type</u> of RF fault given waveform data

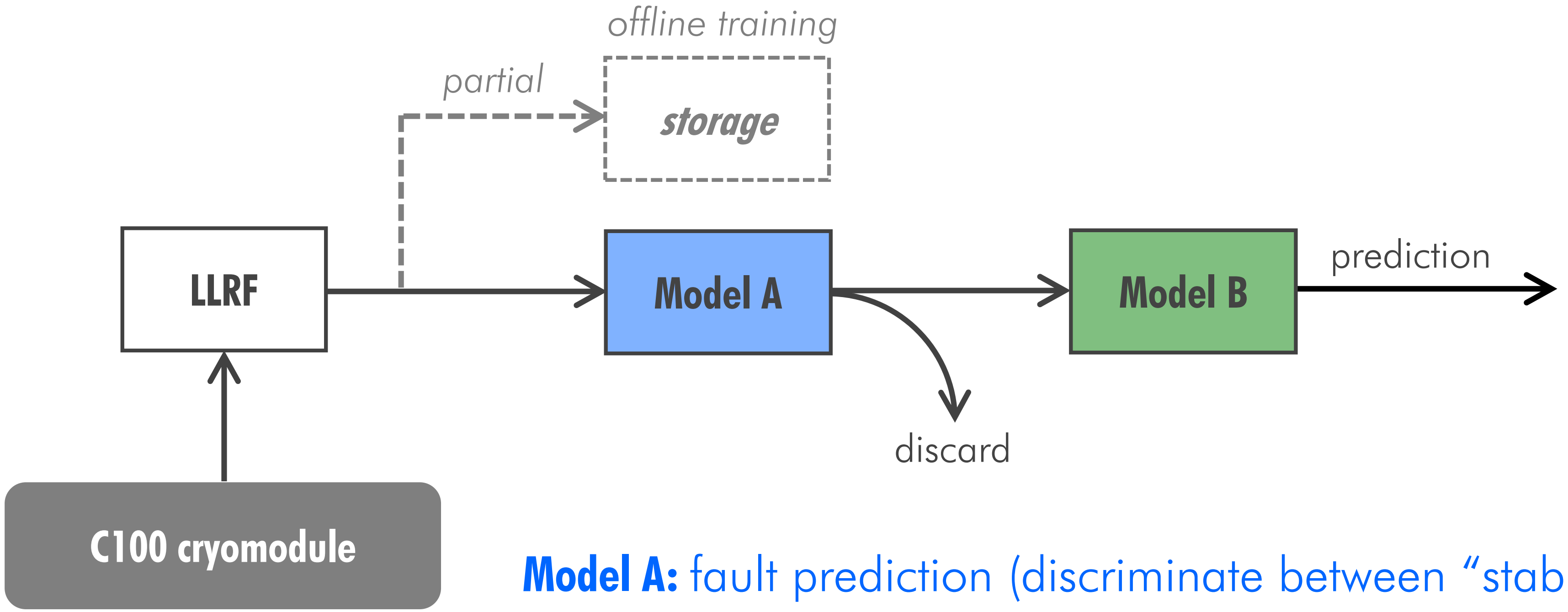
machine learning

multi-class classification

time-series data

# Fault Classification → Fault Prediction

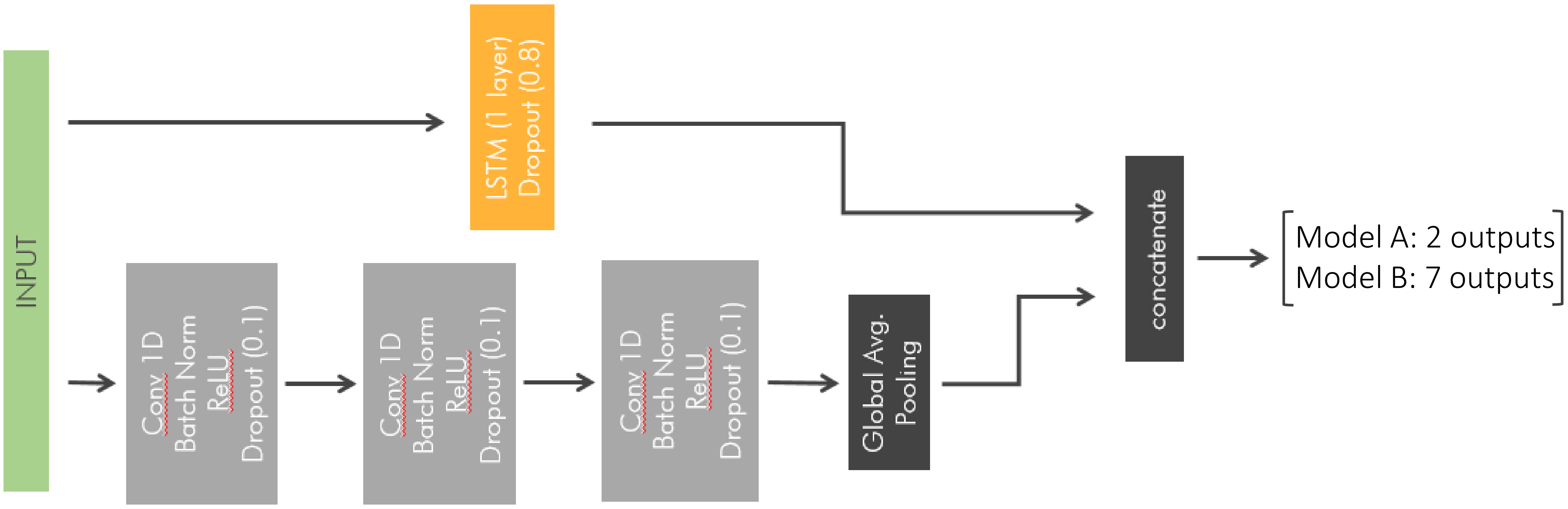
- small portion of waveforms around fault event are used for training classifiers
  - ✓uses *static* datasets
- modifications to LLRF system will allow us to continuously stream data
- investigate if data prior to fault contains enough information to predict event



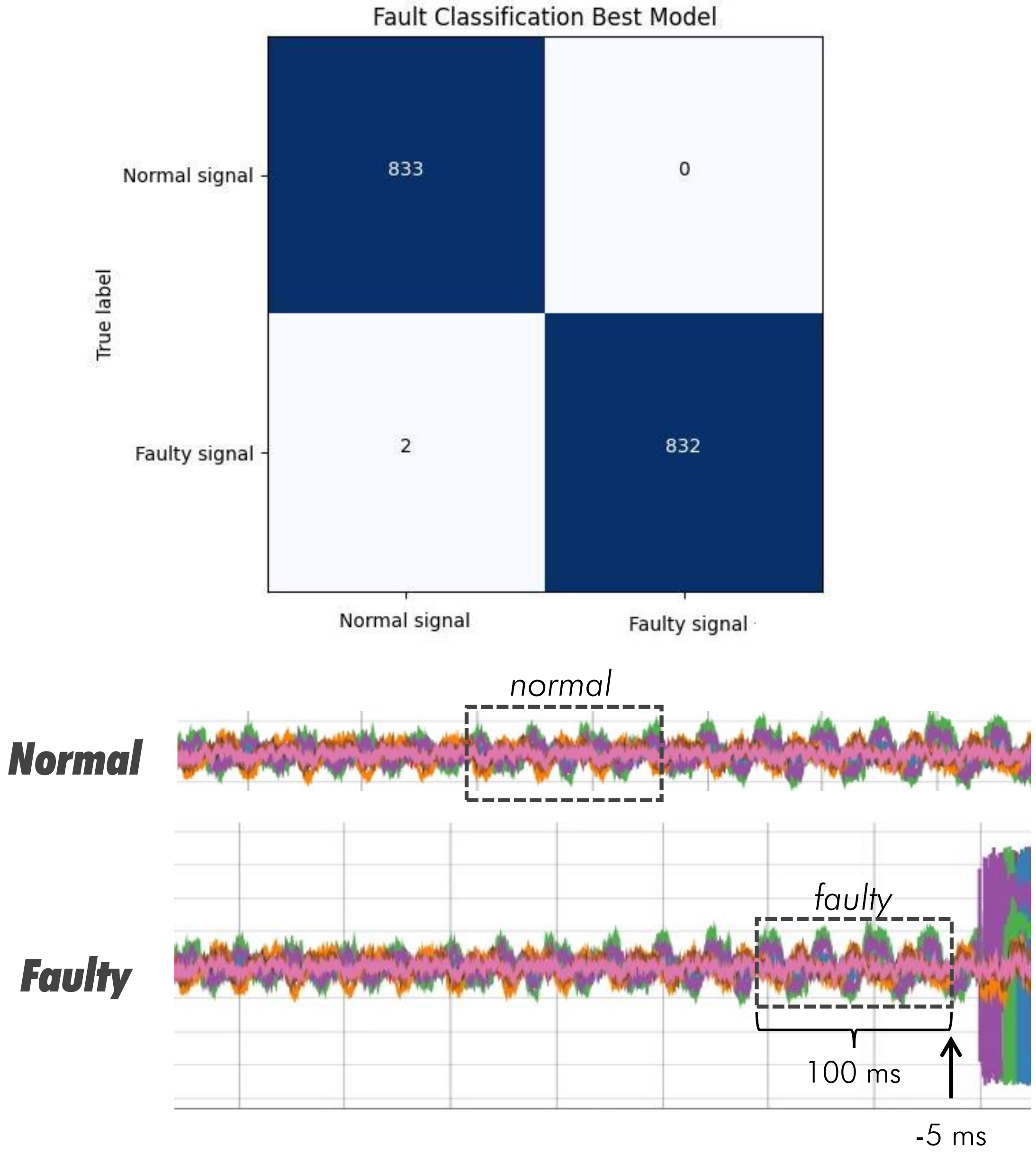

**Model A:** fault prediction (discriminate between "stable" and "impending")
**Model B:** fault-type prediction (classify fault)

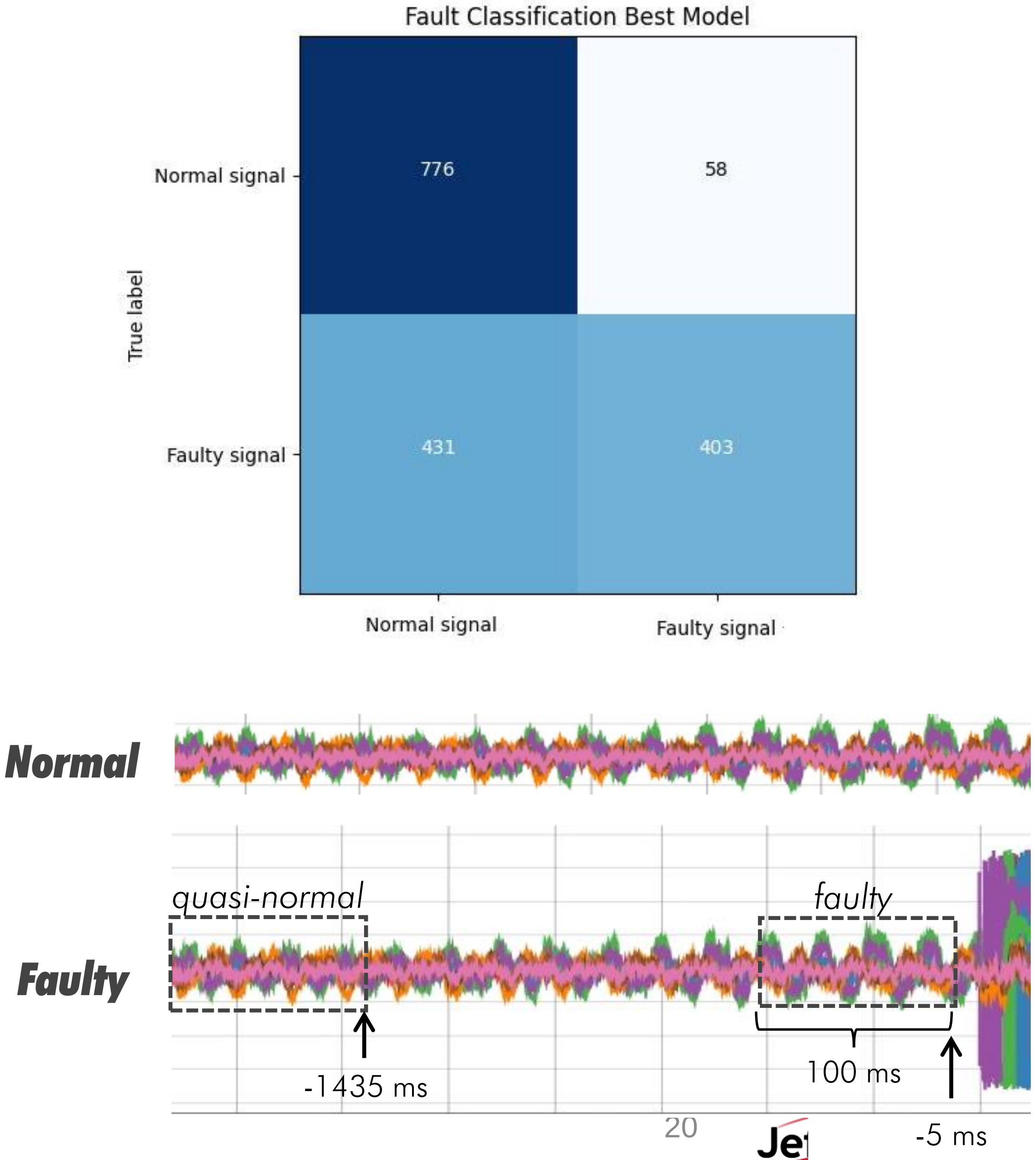
# Hybrid Deep Learning Model for Fault-type Prediction

- 1D CNN – LSTM model architecture for *both* model A and B



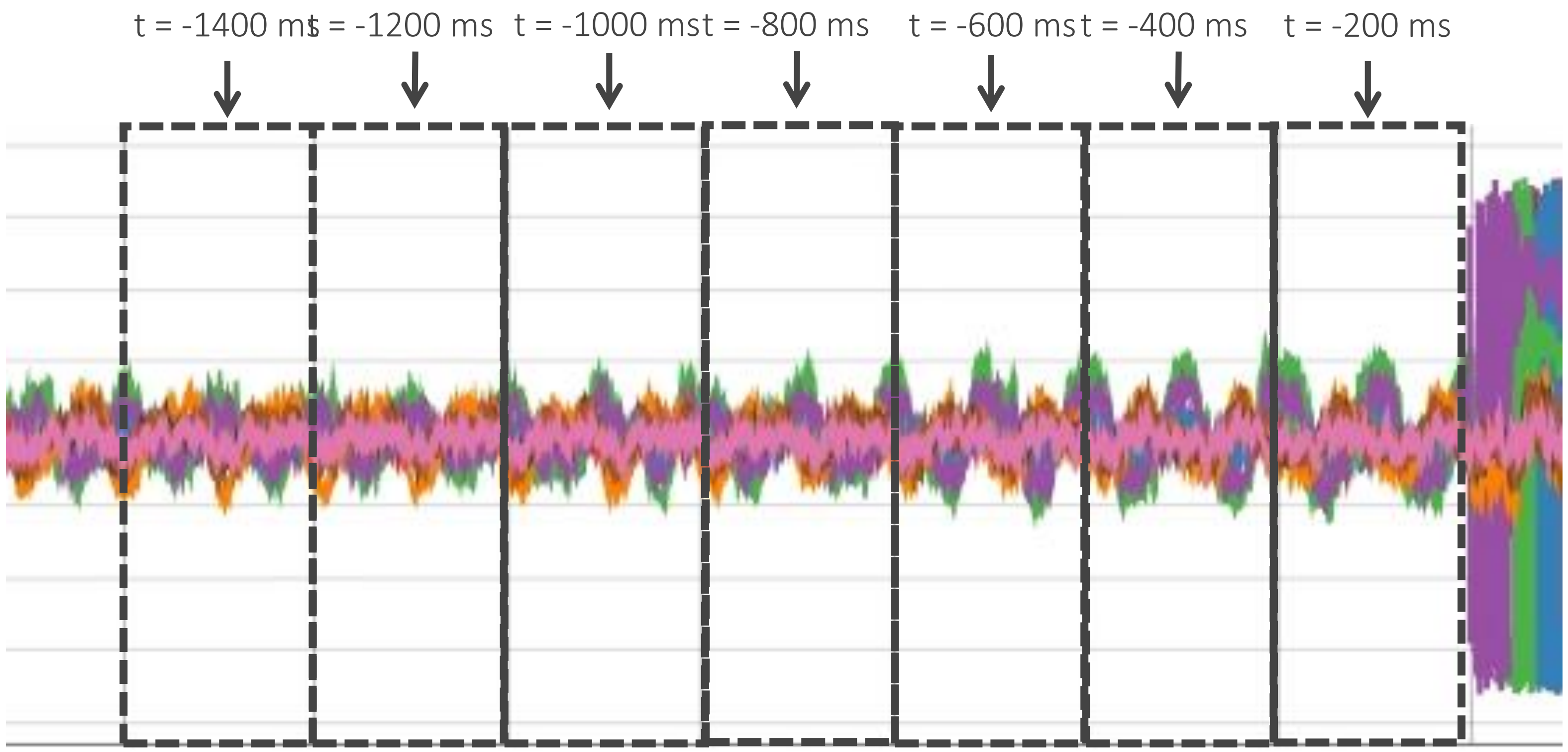
# Model A: Binary Classifier



# Model B: Fault Classifier

- can data prior to event accurately predict the fault type?
  - ✓ use saved waveforms

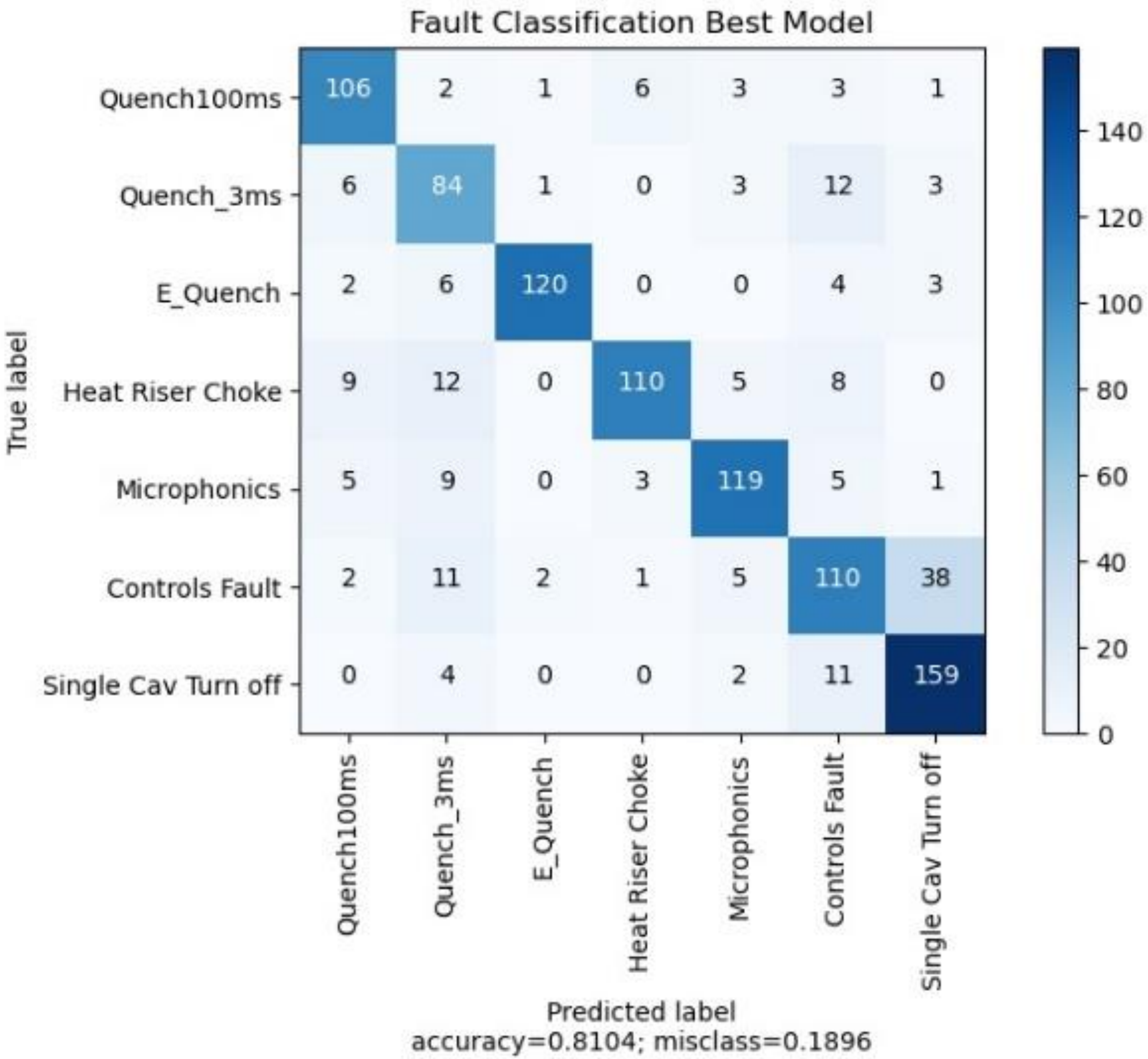




Jefferson Lab

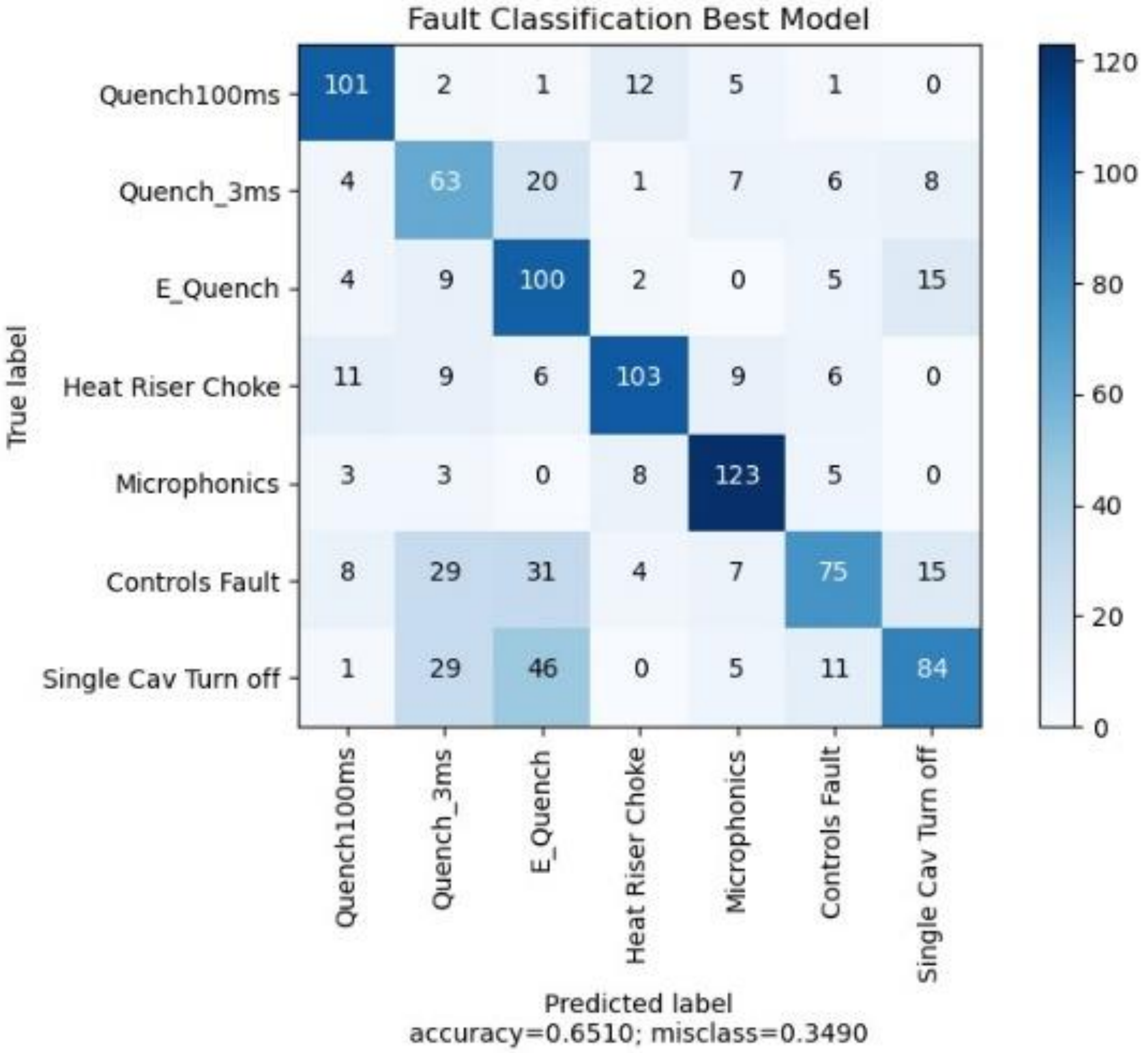
# Model B: Fault Classifier

## 0 ms prior to fault



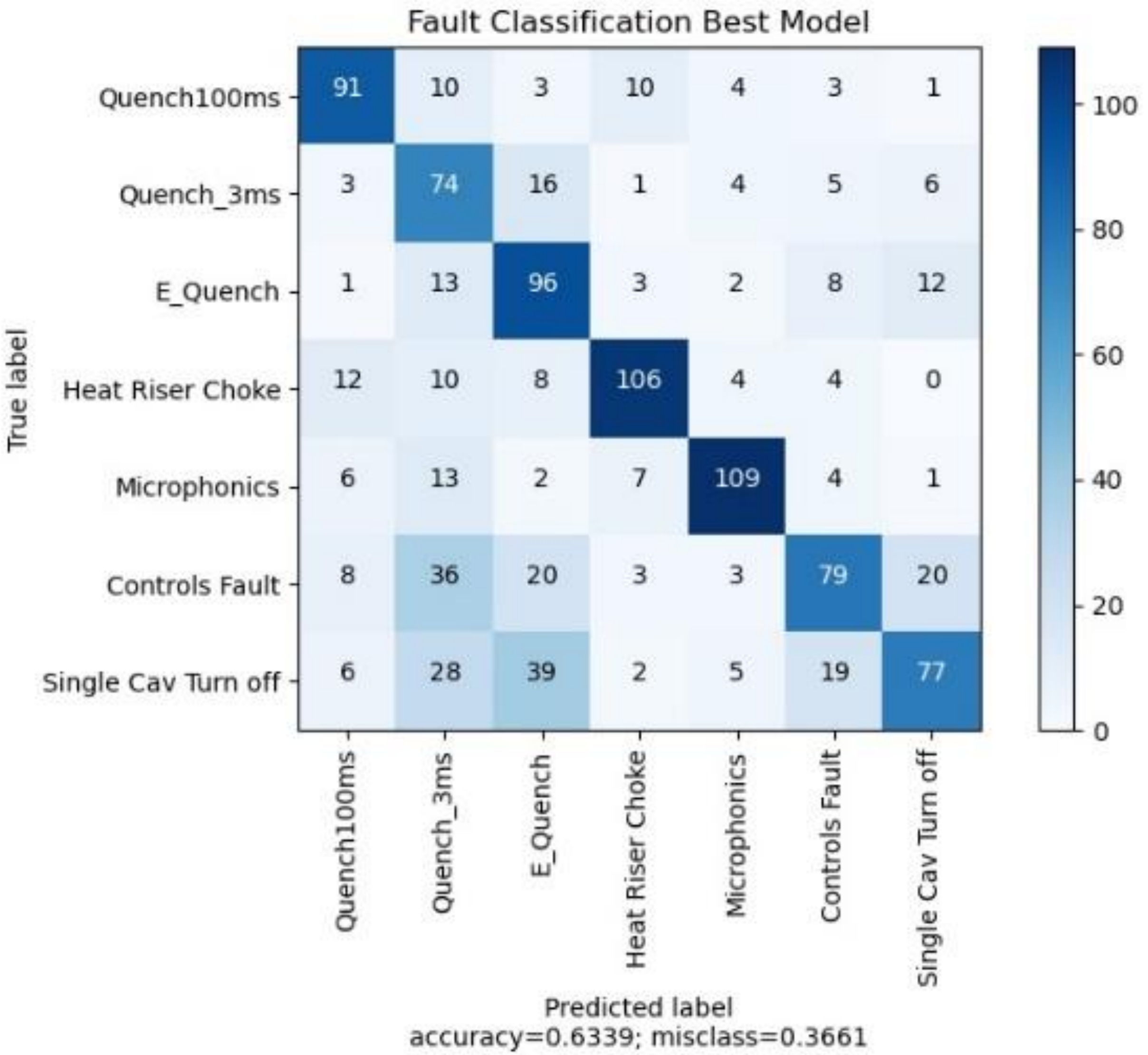
# Model B: Fault Classifier

## 20 ms prior to fault



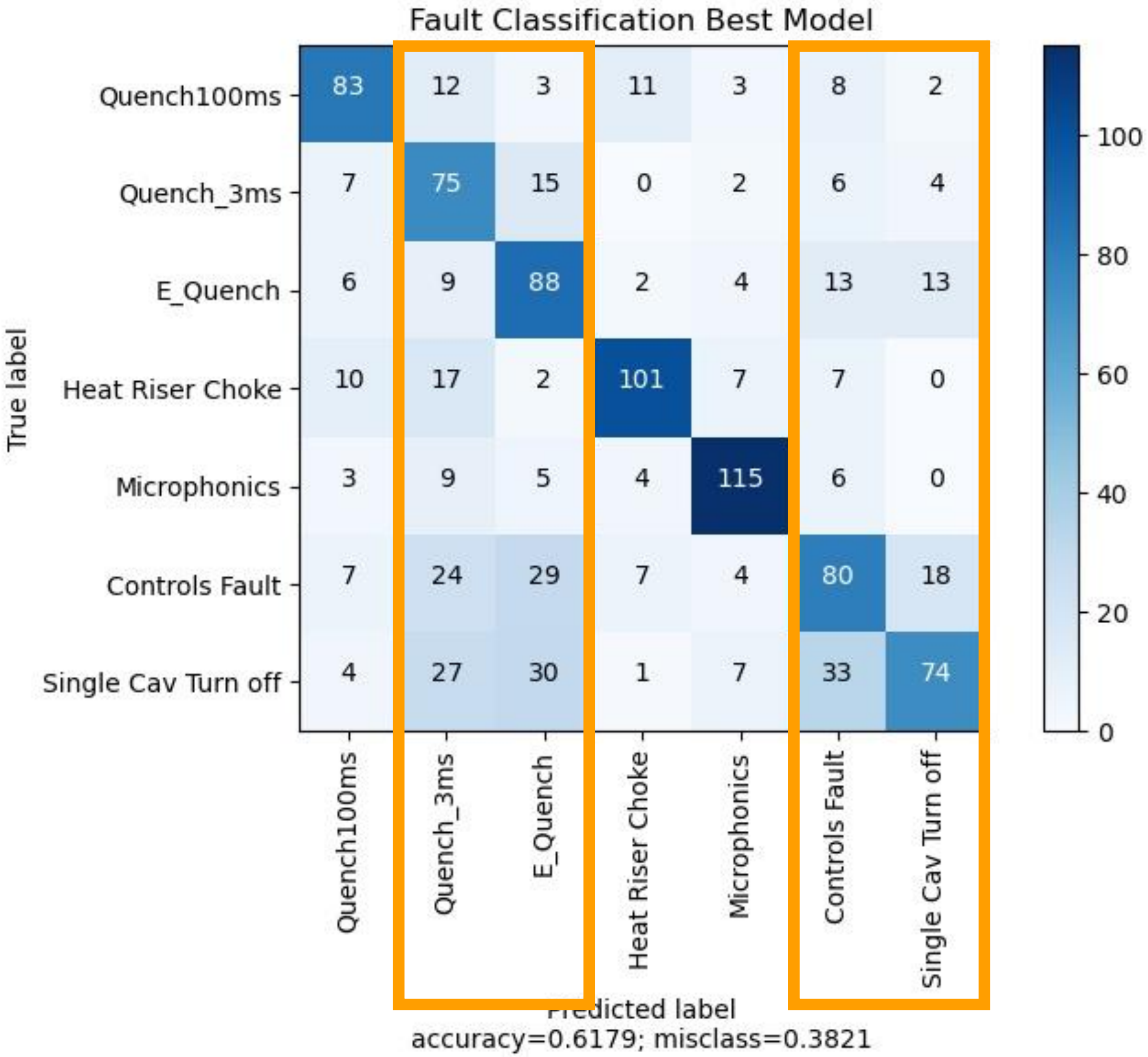
# Model B: Fault Classifier

## 50 ms prior to fault



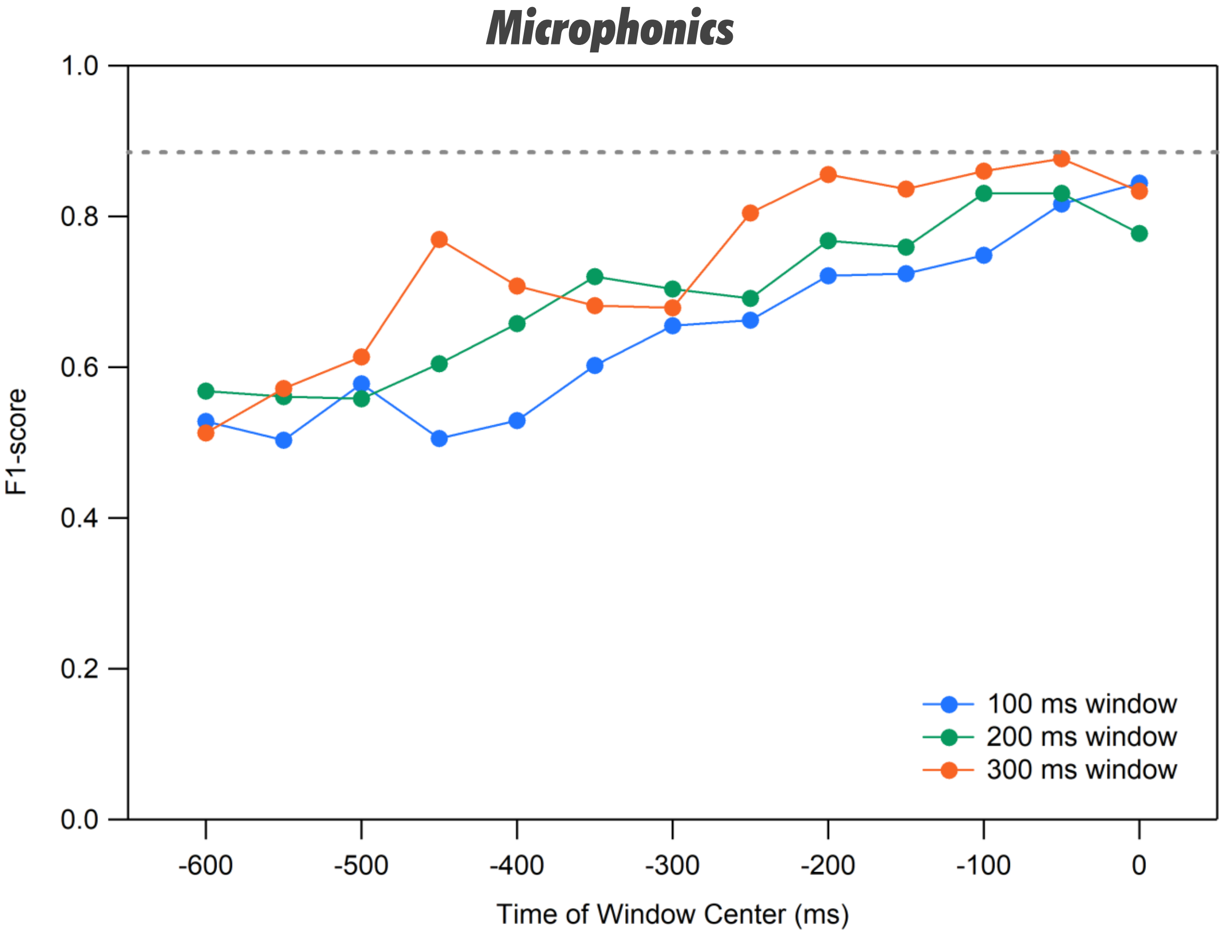
# Model B: Fault Classifier

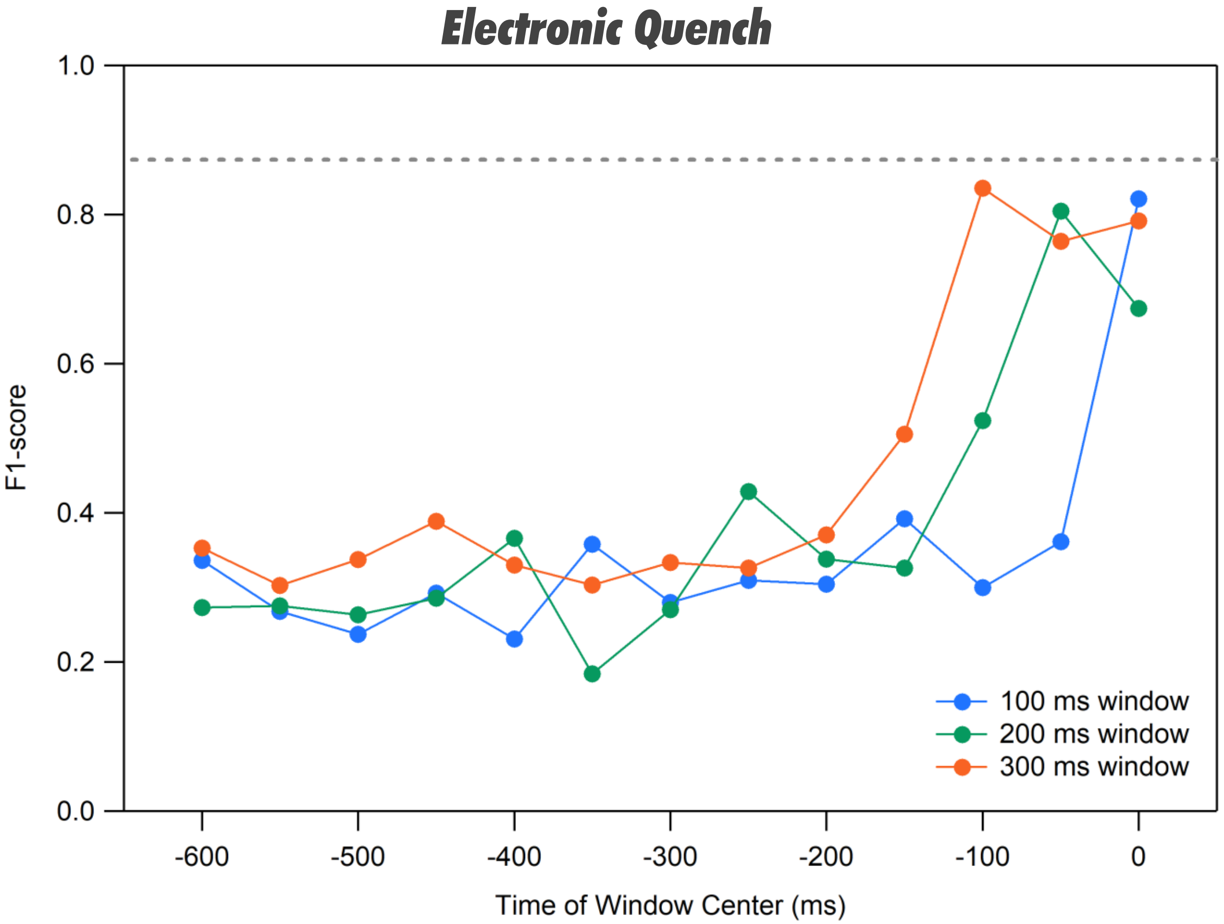
## 100 ms prior to fault



# Model B: Fault Classifier

- initial results suggests that for some fault types, prediction is possible





Jefferson Lab

# PROJECT C

PI: Adam Carpenter and Riad Suleiman

Graduate Student: Kawser Ahammed (ODU)

# Project C: Field Emission Management

- **Goal:**

  maintain low levels of field emitted (FE) radiation without invasive interruptions to physics

- **Description:**

  use ML to model radiation levels and allow for off-line optimization of gradient distribution, identify cavities where FE onsets have changed

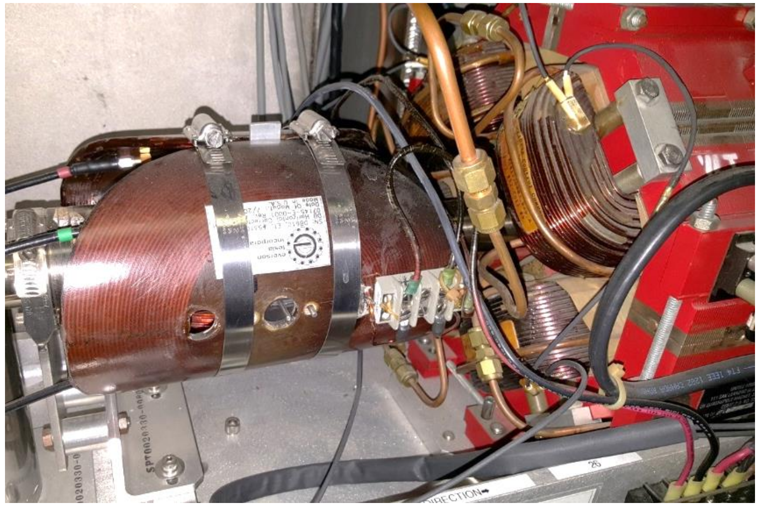
- **Solution:**

  optimize surrogate model to minimize radiation via gradient reduction

*radiation area*

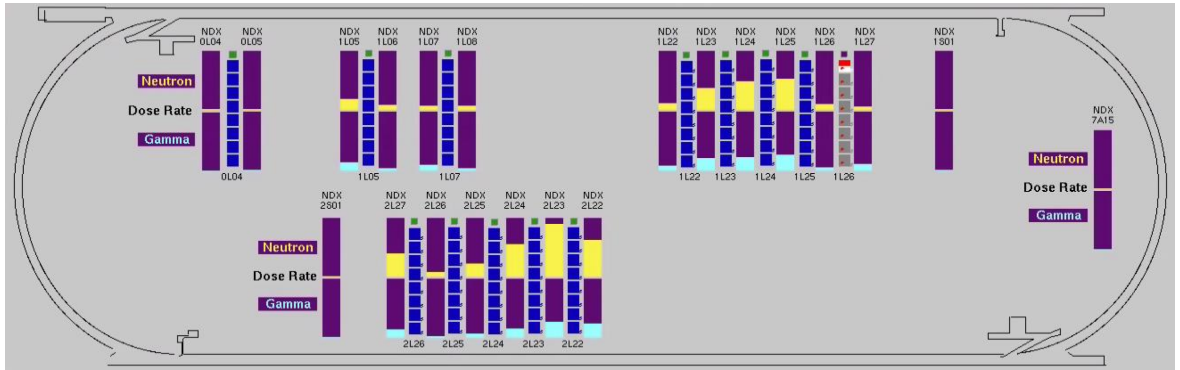

*damaged beamline valve*


*damaged magnet and cables*



# Field Emission Management: Data Requirements

- Jefferson Lab designed, installed, and commissioned a new neutron and gamma radiation detection system* focused on FE radiation
  - ✓operational August 2021
  - ✓measure neutron dose rates correctly in the presence of photon radiation
  - ✓detectors are “blind” to low energy photons and electrons
  - ✓integrated into EPICS with signals for gamma and neutron dose rates
  - ✓wide dynamic range
  - ✓currently have 22 detectors installed



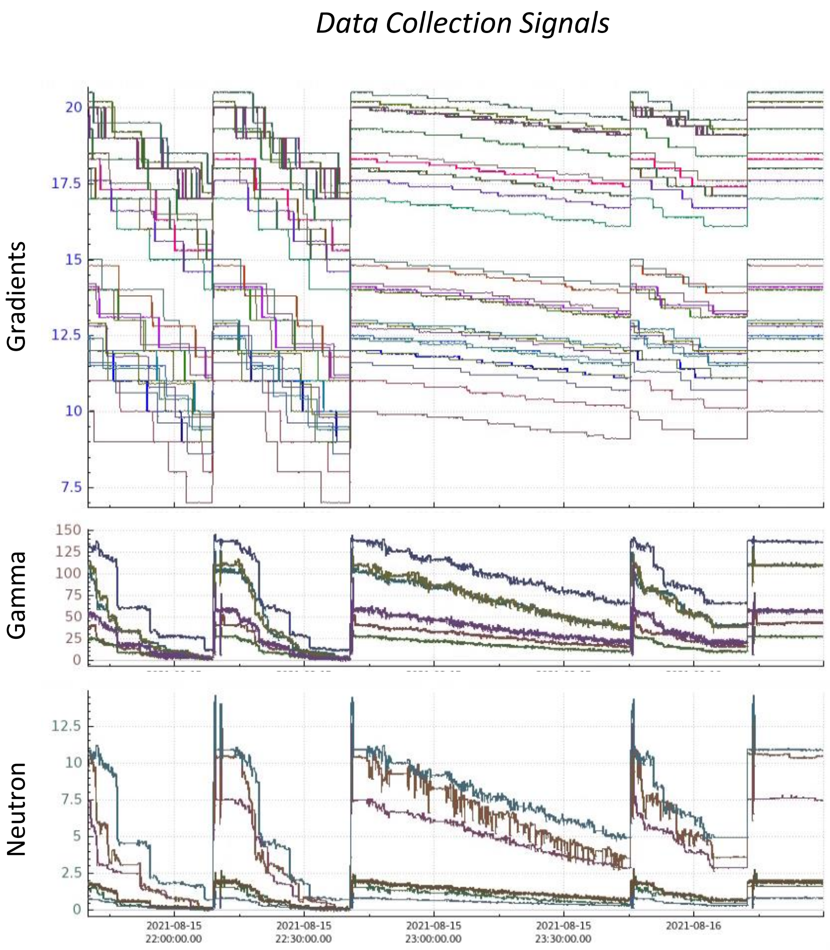


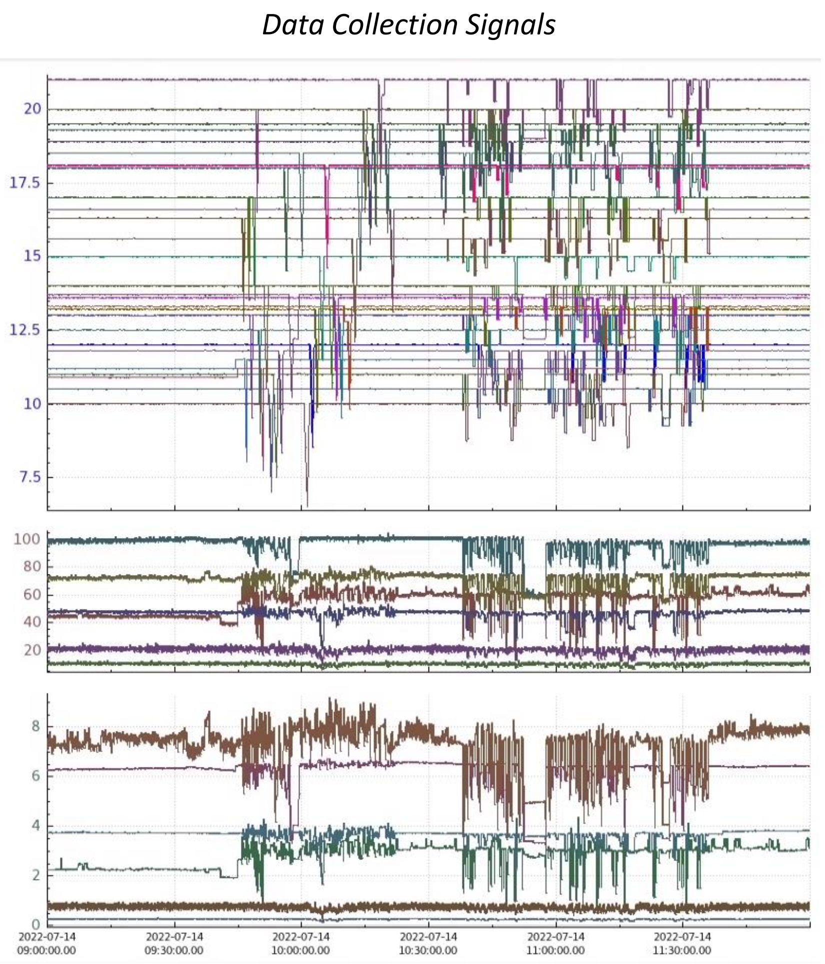

*P. Degtiarenko, US Patent 10,281,600

# Data Collection: New and Improved

- pros of new approach
  - ✓ less correlated inputs
  - ✓ better sampling of input space around operational values
  - ✓ clearly indicates major field emitters
- cons of new approach
  - ✓ slower, so fewer samples
  - ✓ smaller range of radiation values observed
- streamlined process, able to be run by operators

| Old Method | New Method |
| --- | --- |





Jefferson Lab

# Data Collection: Models

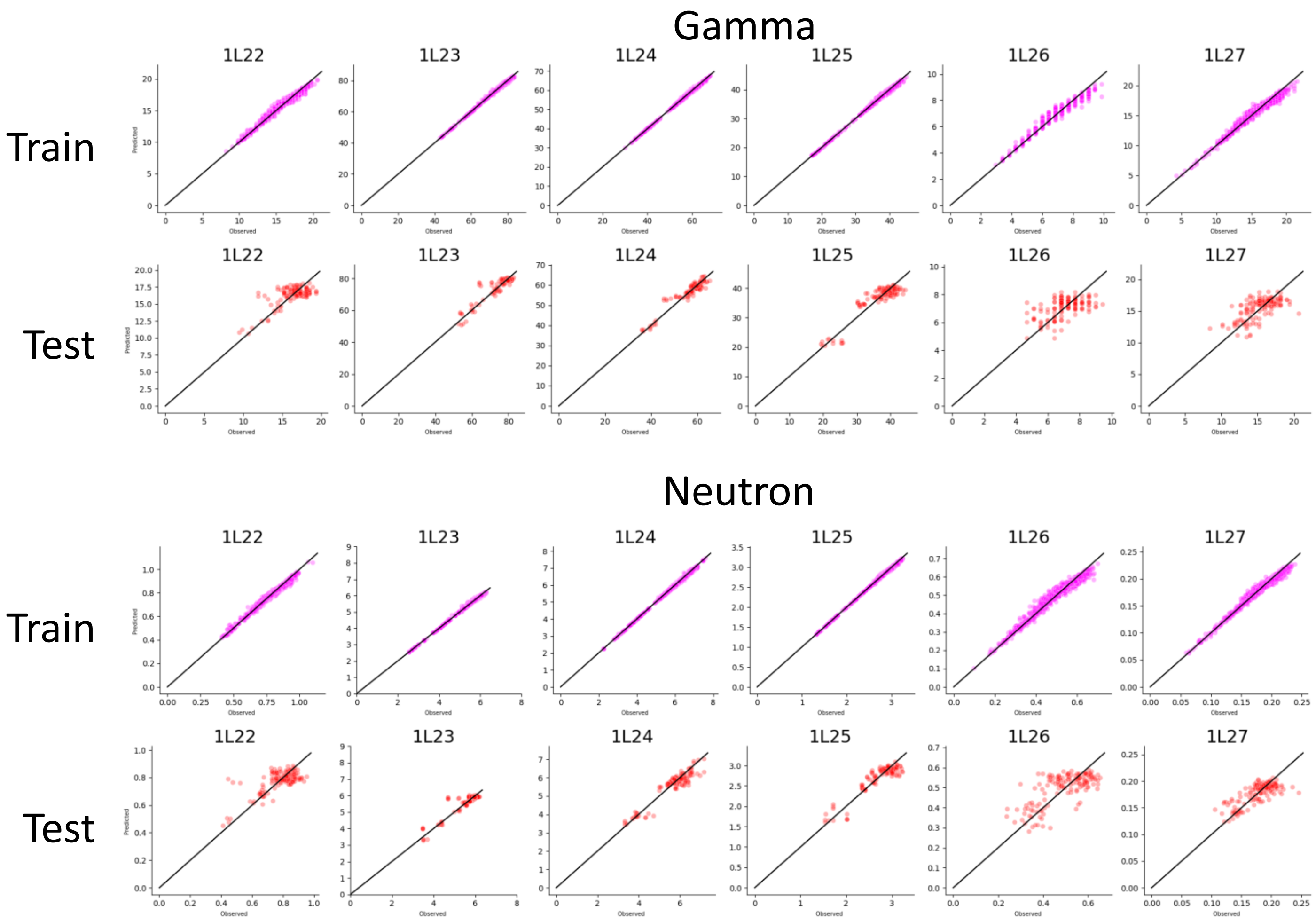
- no FE onset required as input
- no feature engineering
- currently training MLP and XGBoost models
- XGBoost performs better, likely due to limited data
    - ✓ no extrapolation – likely pushes us to MLP

**XGBoost Results on September 7 Data**

| | R-Squared | MSE | MAE | MAPE |
|---|---|---|---|---|
| Train | 0.981 | 0.062 | 0.133 | 0.012 |
| Test | 0.652 | 1.815 | 0.701 | 0.062 |

**XGBoost**

*Observed vs Predicted Plots*





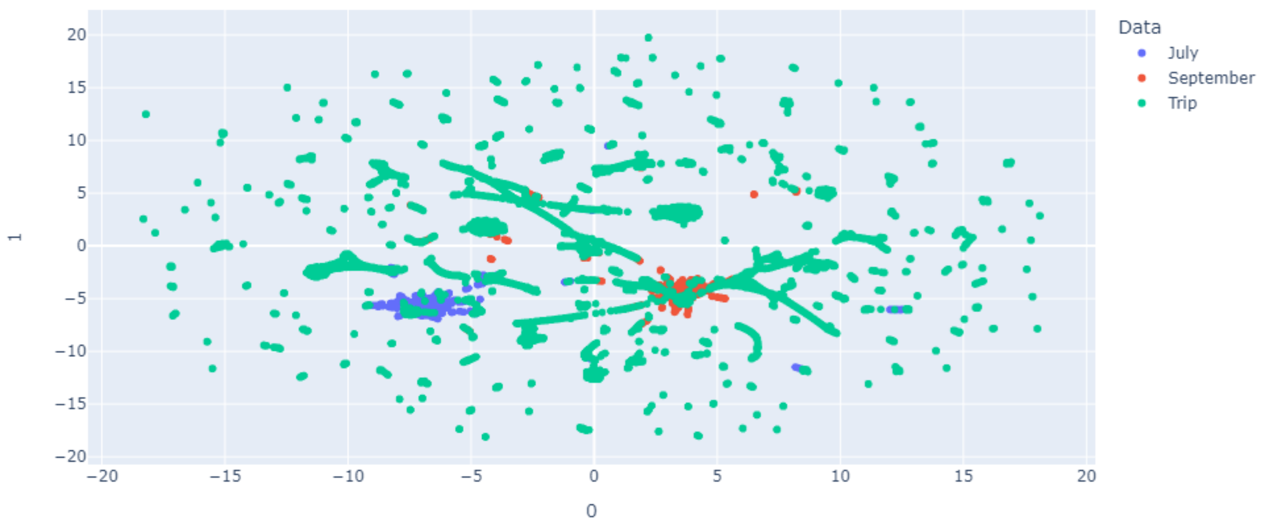

Jefferson Lab

# Dimensionality Reduction

- use dimensionality reduction techniques to visualize input data
  - ✓ reduce 32-dimensional gradient inputs to 2-dimensions
- assess:
  - ✓ how similar or dissimilar data sets are
  - ✓ how does (parasitic) “Trip” data compare to data collected invasively (i.e., “July”)

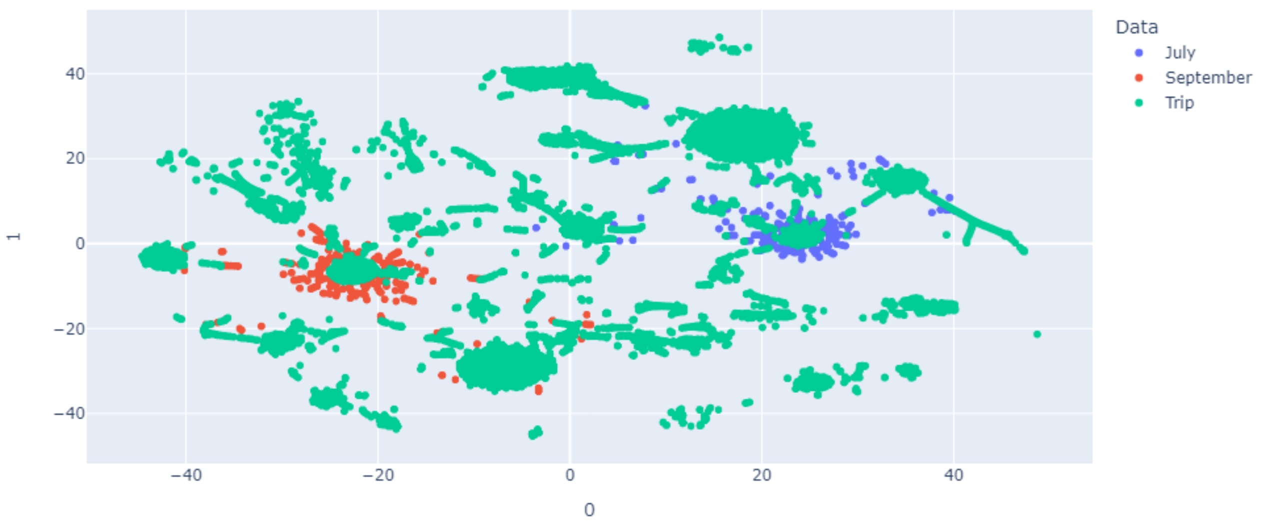
**UMAP**

(n_neighbors=100, min_dist=.1




**t-SNE**

Perplexity=400

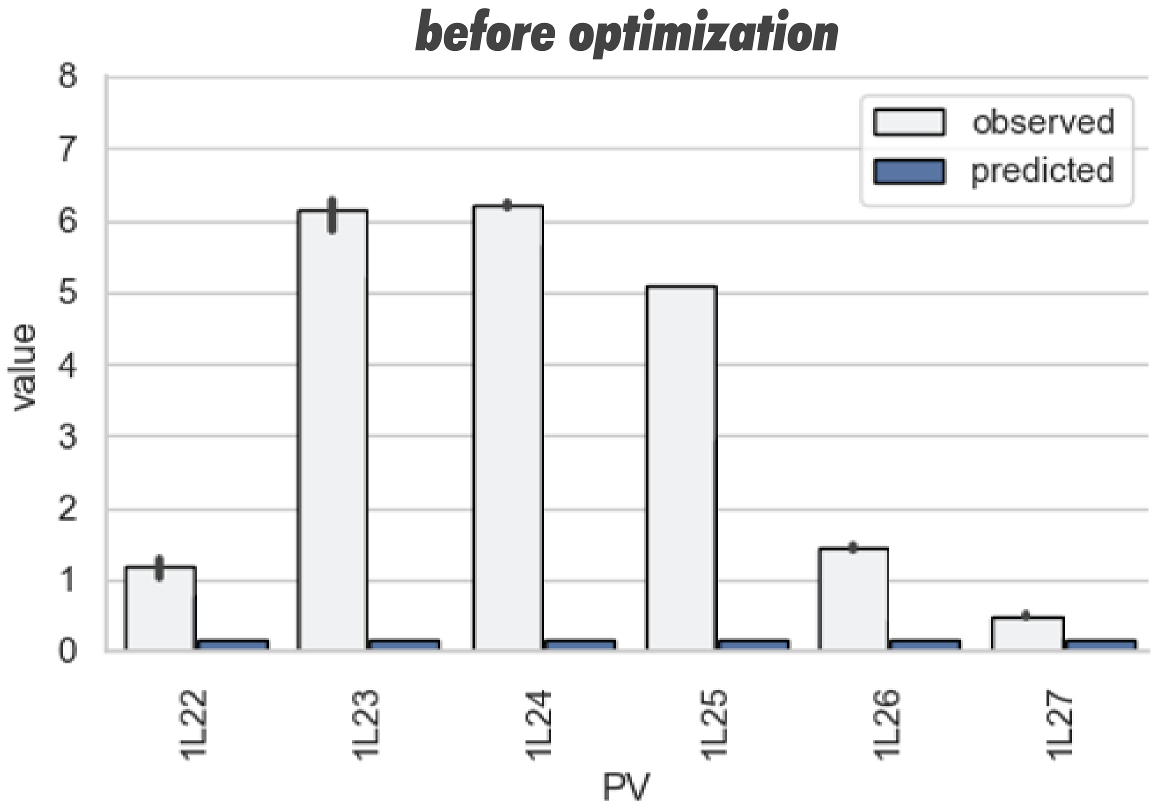





Jefferson Lab

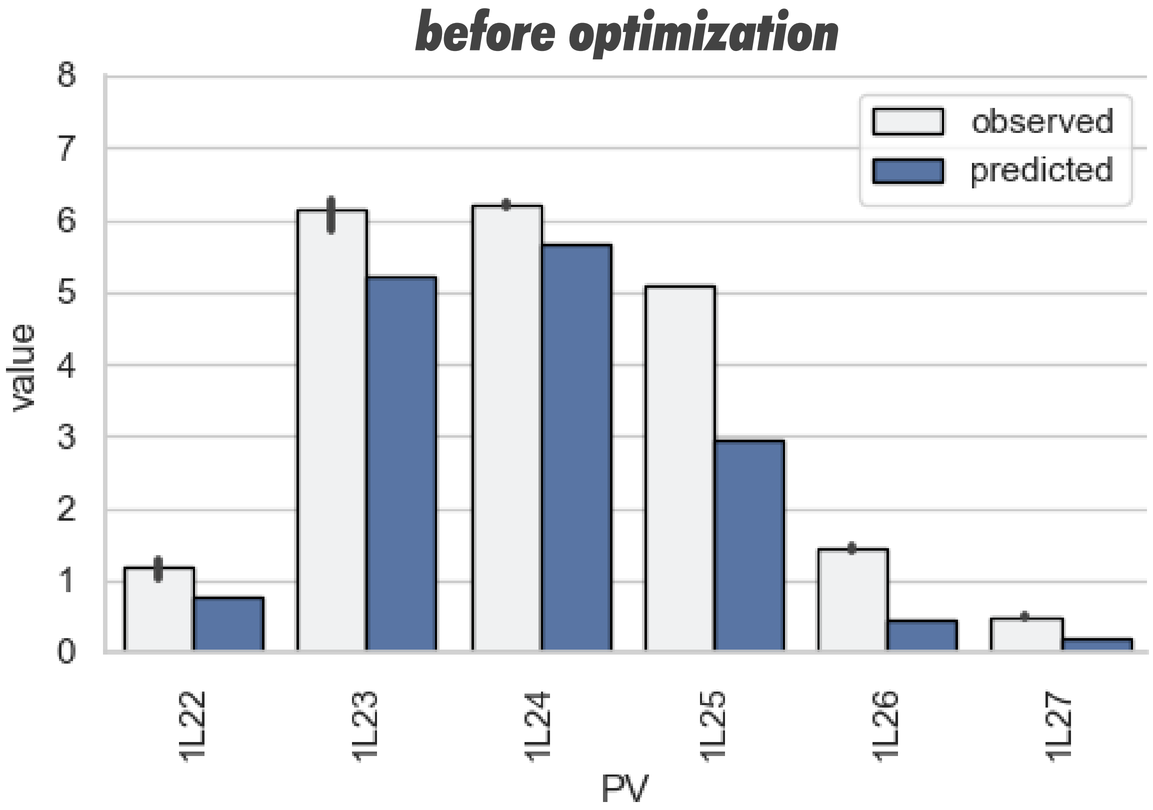
# Field Emission Management: Proof-of-Concept Demonstration

1. set CEBAF to same gradient distribution as September 7 baseline
2. apply model-based optimized gradients to CEBAF



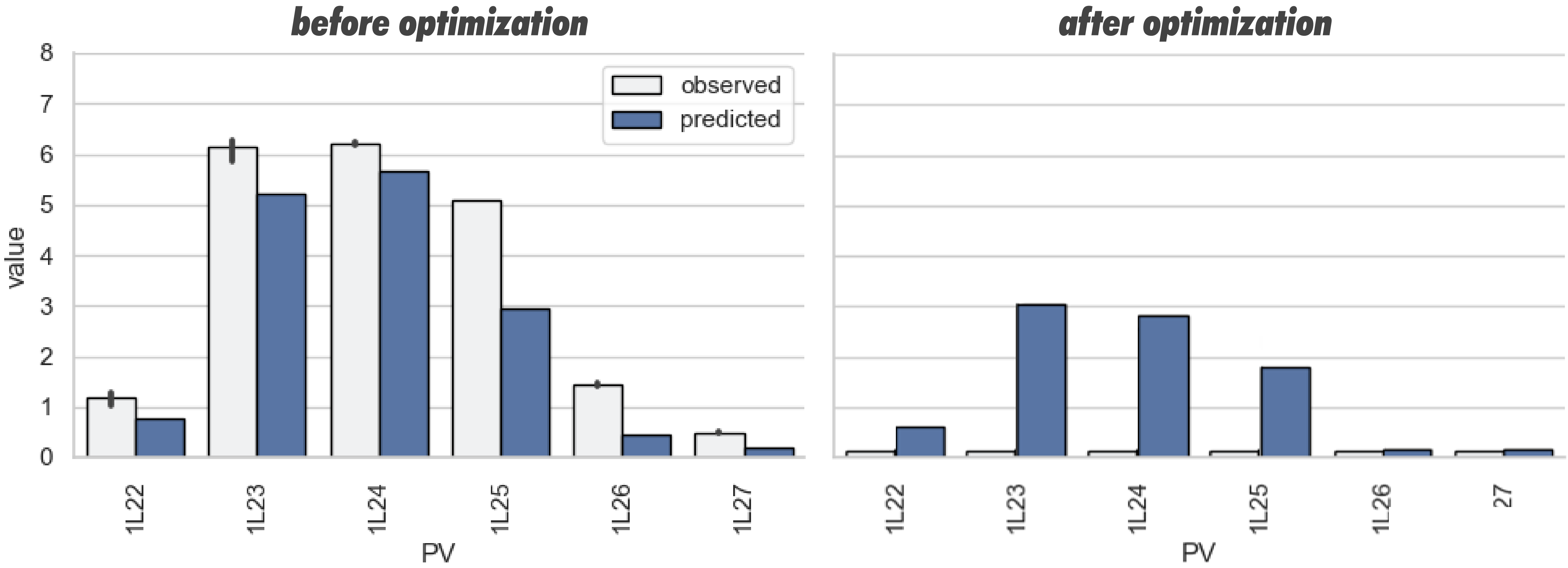
# Field Emission Management: Proof-of-Concept Demonstration

1. set CEBAF to same gradient distribution as September 7 baseline
2. apply model-based optimized gradients to CEBAF



# Field Emission Management: Proof-of-Concept Demonstration

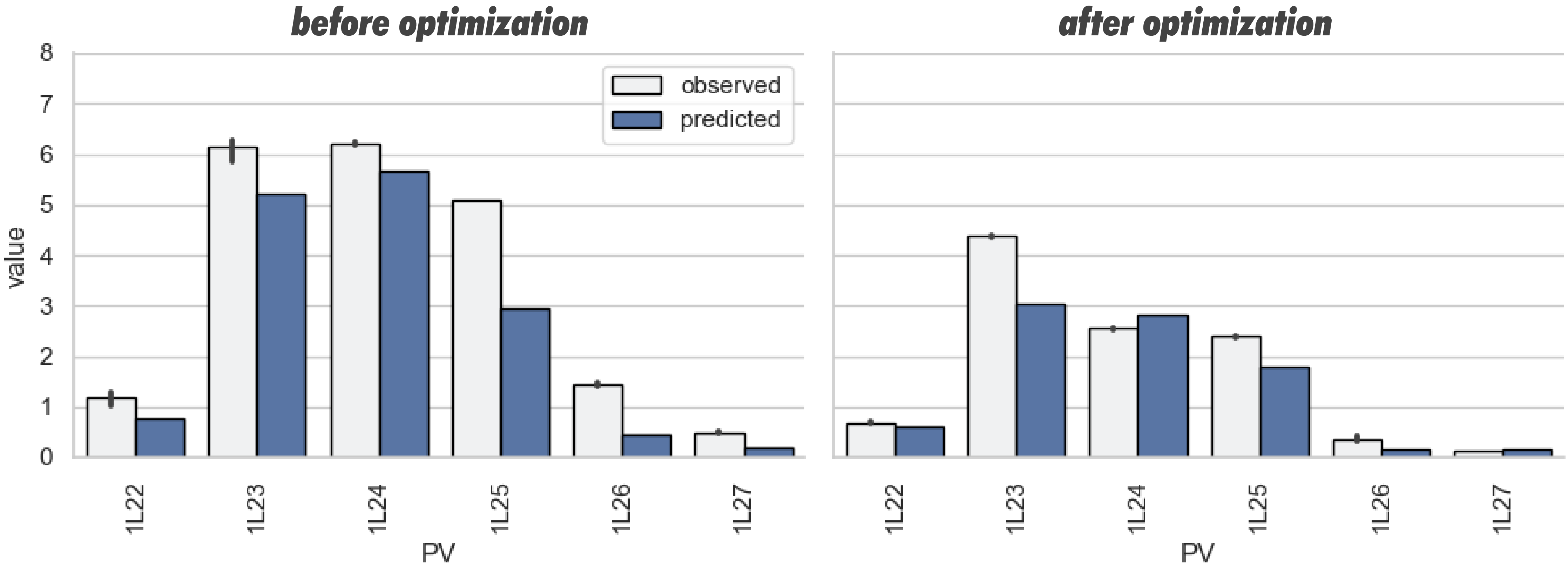
1. set CEBAF to same gradient distribution as September 7 baseline
2. apply model-based optimized gradients to a portion of the NL in CEBAF



# Field Emission Management: Proof-of-Concept Demonstration

1. set CEBAF to same gradient distribution as September 7 baseline
2. apply model-based optimized gradients to CEBAF

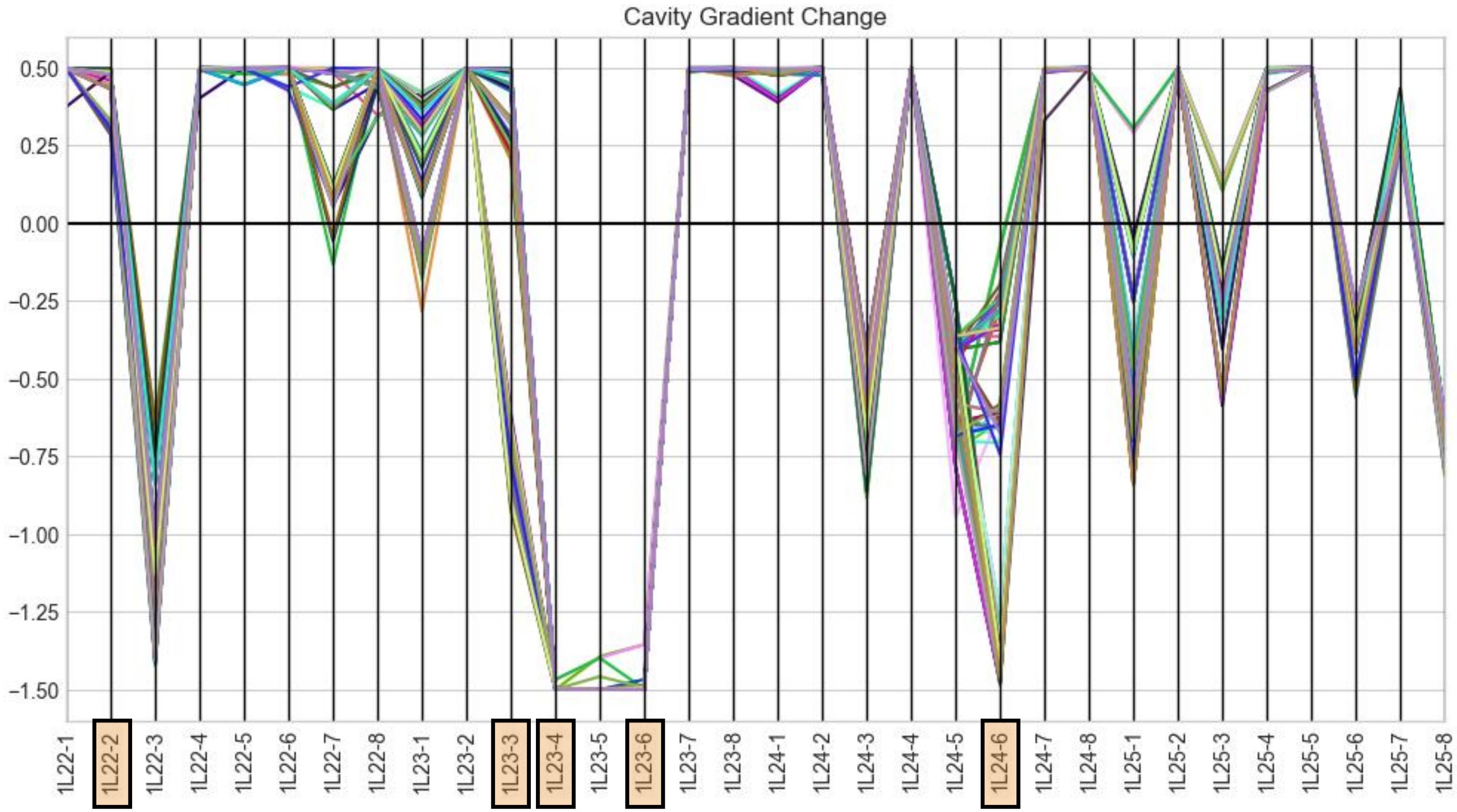
*12 rem/hour decrease for 5 MV/m reduction in gradient*

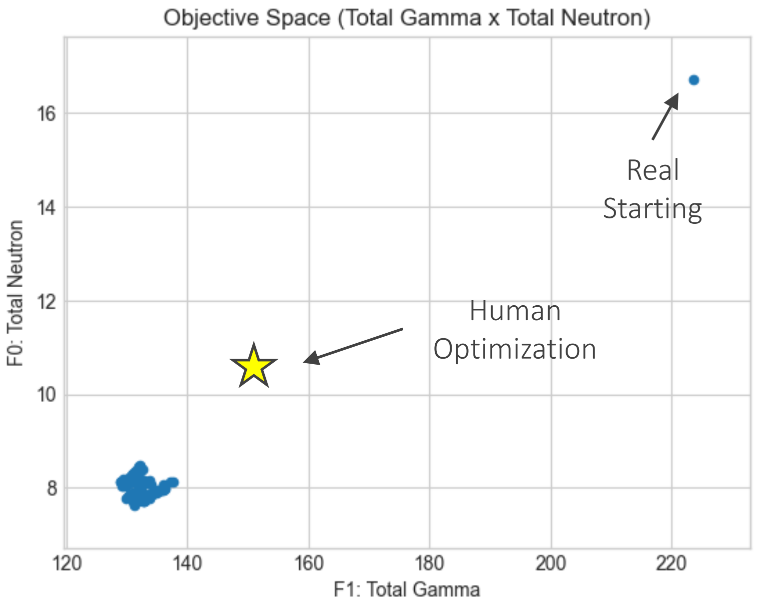


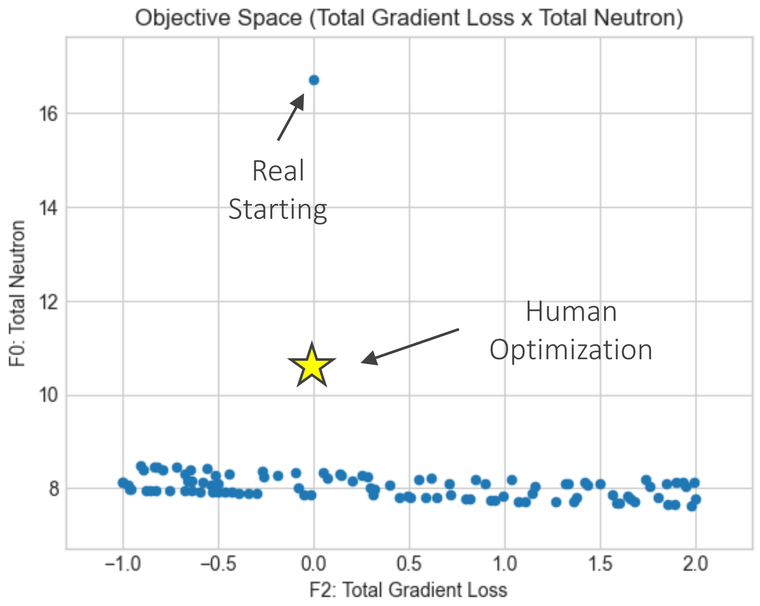
# Field Emission Management: Optimization

- transition to a genetic algorithm (GA) to optimize gradient distributions

Sep Model



# SUMMARY

# Data: The Fuel for Machine Learning

- accelerators produce a lot of data
  - ✓ CEBAF continuously archives 300,000+ signals
- however, it's not all useful for ML applications
- ML projects at JLab only possible because of newly available data

***C100 cavity fault classification# → digital LLRF + waveform harvester***

***C100 cavity fault prediction → digital LLRF + streaming data***

***field emission management → NDX detectors****

***cavity instability detection → fast DAQ***

- data reliability is critical!

#work supported by JLab LDRD

*P. Degtiarenko, US Patent 10,281,600

# Data: The Fuel for Machine Learning

- focus in Year 2 was developing models using alternate available data
- significant delays have plagued our ability to use fast and/or streaming data
  A. cavity instability detection: DAQ system
     - 🚩 *supply chain and other pandemic related issues caused delays (2+ years)*
  B. C100 fault prediction: dual-buffer firmware upgrade
     - 🚩 *bench tests ongoing, however experiencing 2-year delay from expected deployment*
  C. field emission management: NDX detectors
     - *built, tested, commissioned, installed, and operational*
- focus of Year 3 will be to continue making progress in getting systems in place to collect data required for developing machine learning models
- have – or will have – sources of high quality data to enable continued work in this area for the foreseeable future (beyond the life of the FOA)

# Year 2 Progress

***Project A: Cavity Instability Detection***

- DAQs in the process of being fabricated, tested and installed
- good progress on ML model for slow and fast data, user interface developed

***Project B: C100 Fault Prediction***

- binary classifier shows excellent performance and will be deployed shortly
- fault-type classifier shows good performance as well

***Project C: Field Emission Management***

- proof-of-principle demonstration showing the utility of the surrogate model
- work to better understand data and how to best maintain model performance over time

- three posters presented at 2022 NAPAC conference
  - ✓ "Initial Studies of SRF Cavity Fault Prediction at Jefferson Laboratory"
  - ✓ "Using AI for Management of Field Emission in SRF Linacs"
  - ✓ "SRF Cavity Instability Detection with Machine Learning at CEBAF"

# Project Summary: Major Deliverables and Schedule

*CEBAF Scheduled Accelerator Down: March – July, 2023*

| Project | Deliverable | Date |
|---|---|---|
| *Cavity Instability Detection* | Installation of 20 production DAQs | 03/2023 |
| | Deployment of user interface | 03/2023 |
| | Training and testing of ML model using *fast* data | 03/2023 |
| | Incorporate transient energy signals from BPM data | 07/2023 |
| | Deploy ML model in CEBAF | 07/2023 |
| *C100 Fault Prediction* | Deploy binary classifier in CEBAF for testing and evaluation | 12/2022 |
| | Train and test ML regression model (deploy if performance is acceptable) | 01/2023 |
| | Deploy fault type classifier to work with binary classifier | 02/2023 |
| | Implement streaming data capability and use with deployed models | 07/2023 |
| *Field Emission Management* | Use dimensionality reduction to visualize and understand data sets | 02/2023 |
| | Develop optimization software for use with surrogate model to optimize gradients | 02/2023 |
| | Develop whole (NL) linac surrogate model | 07/2023 |

# Project Summary: Annual Budget

| | **FY2020** ($k) | **FY2021** ($k) | **Total** ($k) |
|---|---|---|---|
| a) Funds allocated | 450,000 | 450,000 | $900,000 |
| b) Actual costs to date | 450,000 | 214,287 | $664,287 |

- *took awhile to find second graduate student*
- *took even longer to find third graduate student*
- *have not been able to replace PostDoc*

# Acknowledgements

*Kawser Ahammed*

*Rama Bachimanchi*

*Adam Carpenter*

*Hal Ferguson*

*Khan Iftekharuddin*

*James Latshaw*

*Jiang Li*

*Theo McGuckin*

*Md. Monibor Rahman*

*Riad Suleiman*

*Dennis Turner*

*And others!*

# Thank You