Data-Driven Detection, Isolation, Identification, and Prediction of Accelerator Fault Events

Chris Tennant | *Jefferson Lab*

AI/ML for Particle Accelerator, X-Ray Beamlines and Electron Microscopy

November 2, 2021

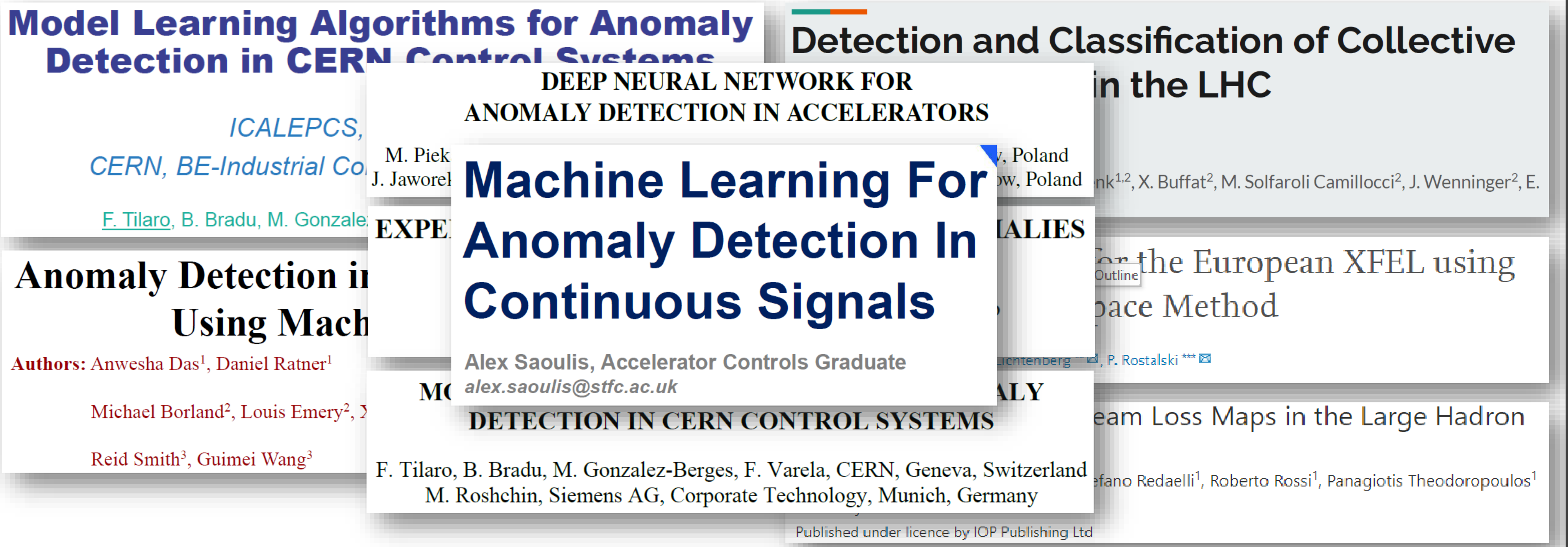
special thanks to Adam Carpenter and Lasitha Vidyaratne

Outline

- ***Some Definitions***
- ***Case Study: SRF Cavities***
  - ✓ isolation and identification
    - *post-fault and post-run*
  - ✓ prediction
- ***Data***
- ***Summary***

AI and Particle Accelerators

- particle accelerators represent the most complex scientific instruments designed, built, and operated
- there is clear motivation to maximize scientific output per operating dollar



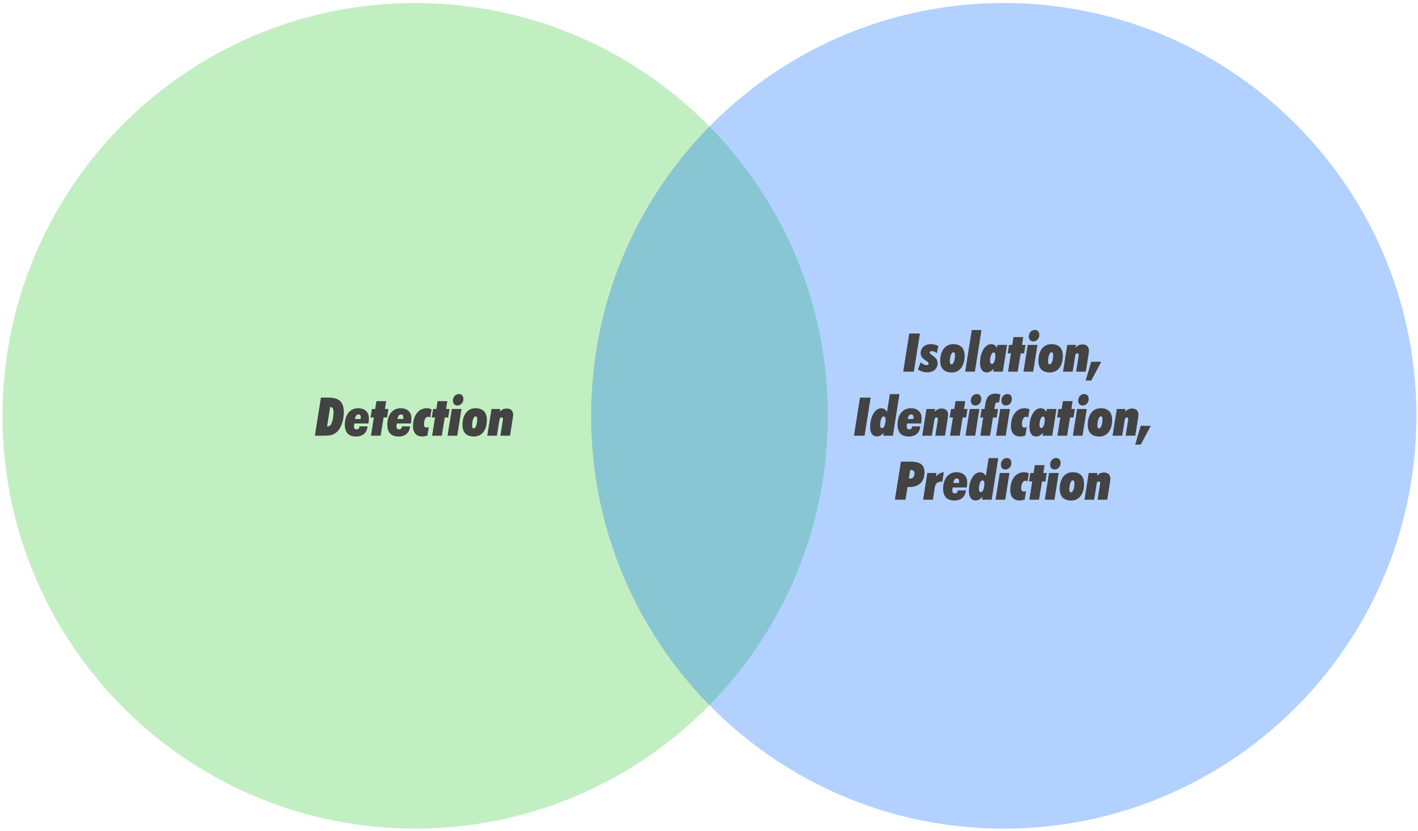
Definitions

- ***Fault***: an unpermitted deviation of at least one characteristic property or parameter of the system from acceptable, usual or standard conditions
- ***Fault Detection***: monitoring measured variables to determine if a fault has occurred (if a fault has occurred, it may be important to determine the time at which the fault occurred)
- ***Fault Isolation***: determining the location of a fault once it is known that a fault has occurred
- ***Fault Identification***: determining the type of fault
- ***Fault Prediction***: providing advanced warning of an impeding fault

(R. Isermann, "*Fault Diagnosis Systems*")

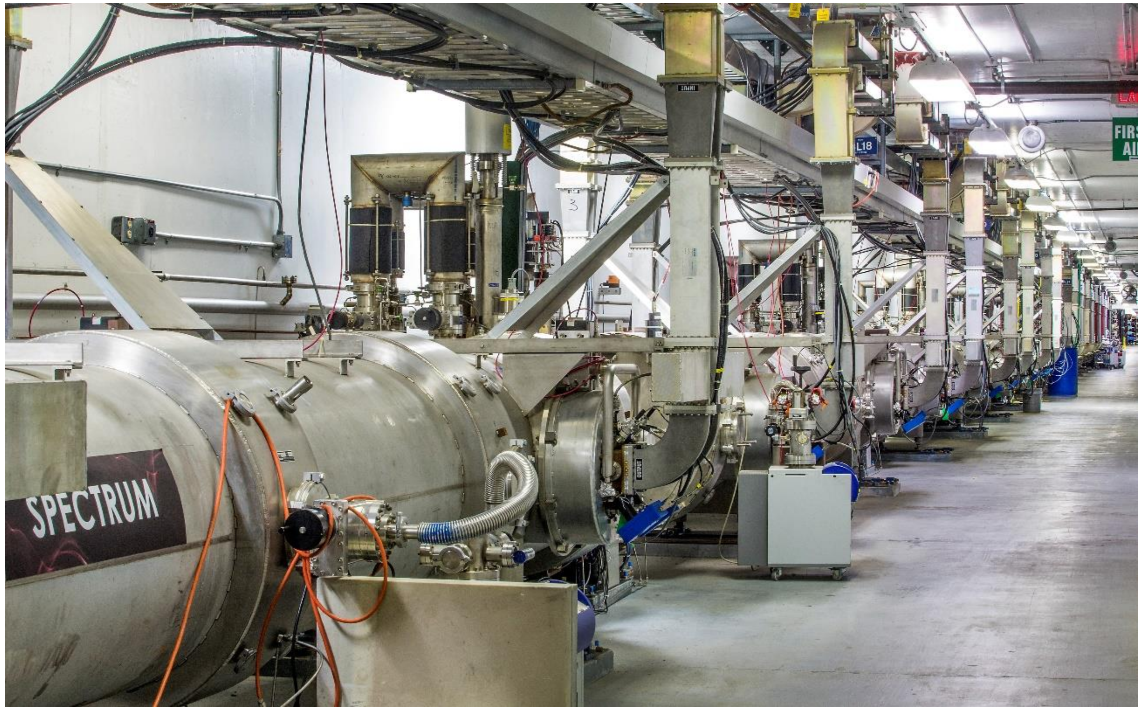
Detection vs (Isolation, Identification, Prediction)

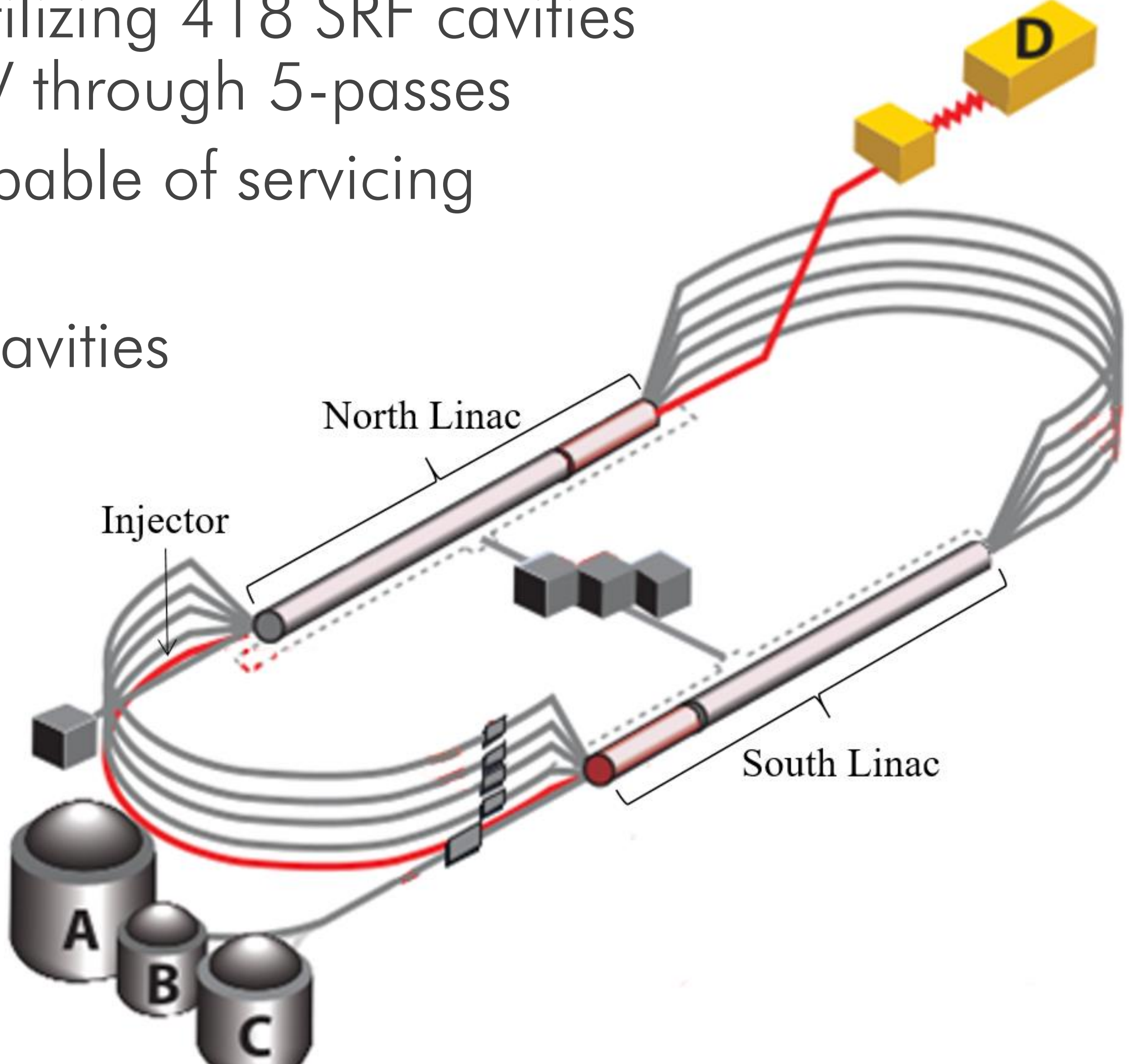
- machine protection systems, personal safety systems, alarms, and other engineered systems are able to detect many types of faults
- in this talk the focus is on faults that have already been detected



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

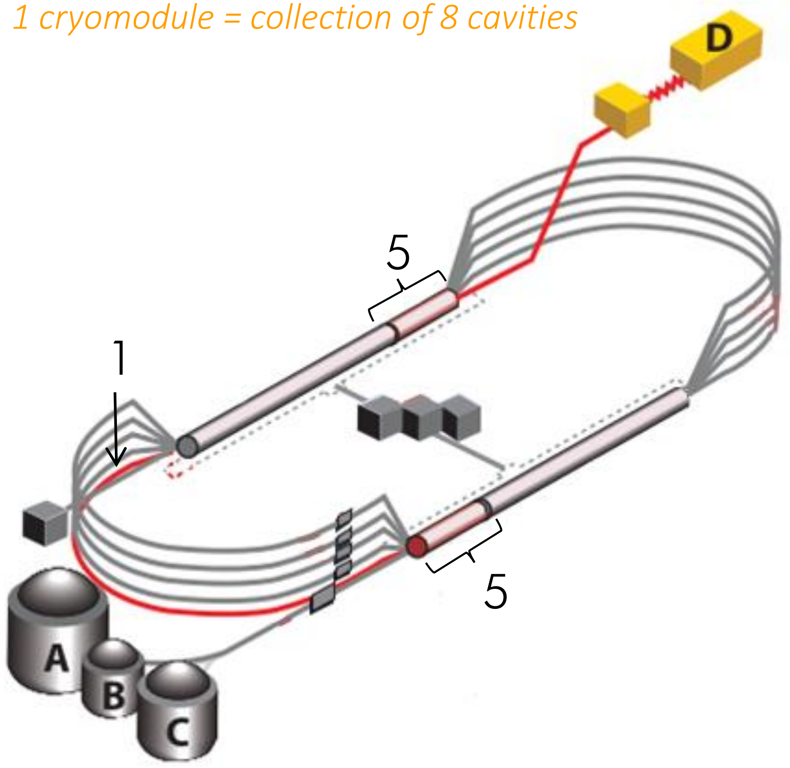




Case Study: SRF Cavities

we have the ability to record high-fidelity data from 12 cryomodules

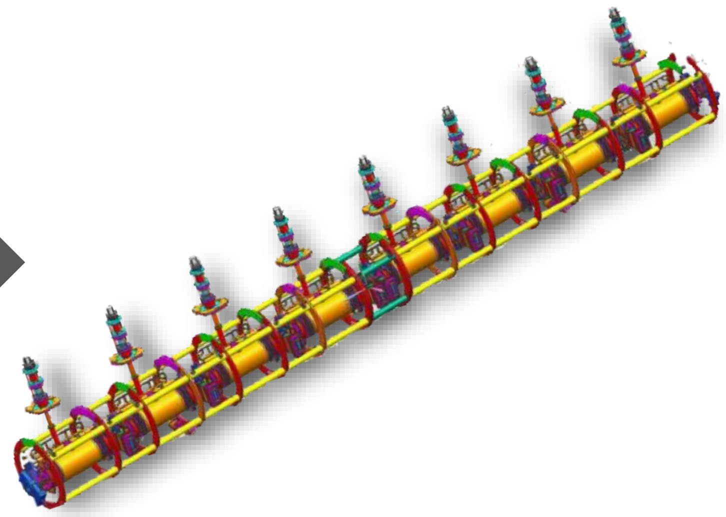
1 cryomodule = collection of 8 cavities



FAULT ISOLATION

Which of the 8 cavities faulted first?

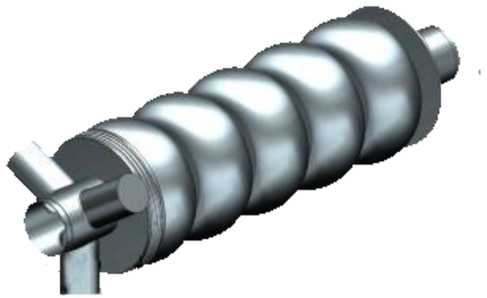
17 signals/cavity × 8 cavities = 136 signals



FAULT IDENTIFICATION

What kind of trip was it?

17 signals

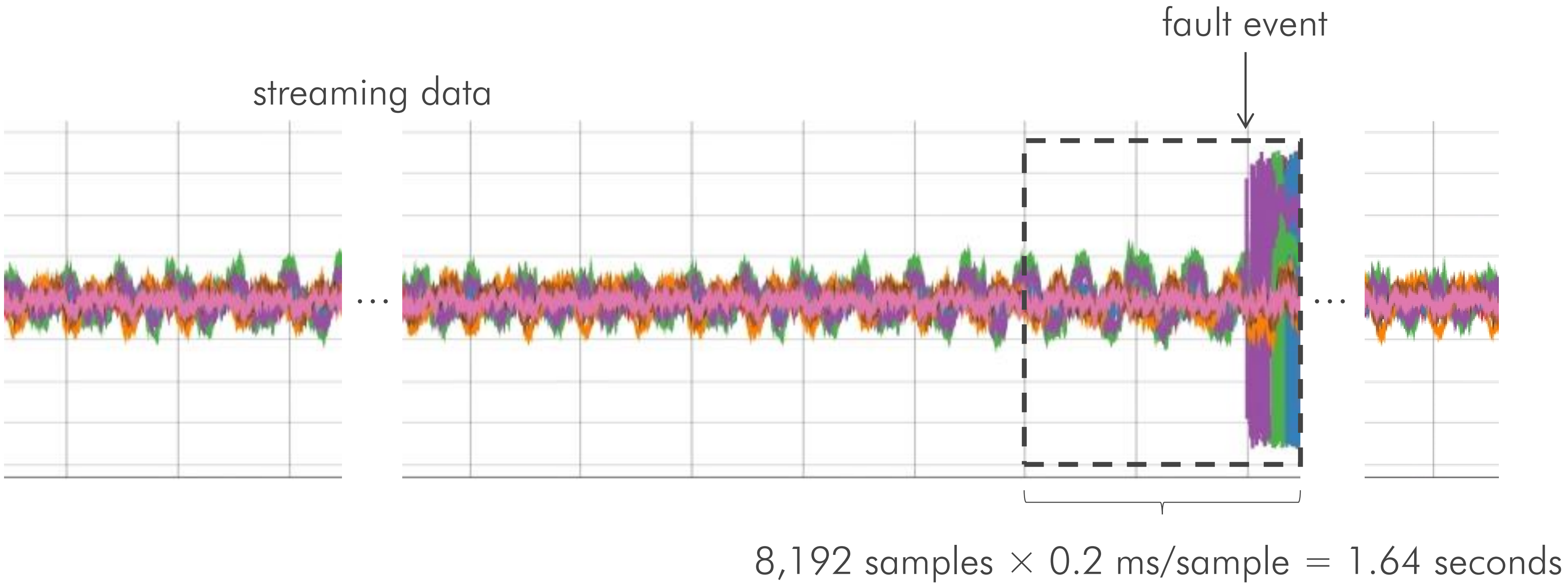


train a model to correctly classify the cavity and type of RF fault given waveform data

$\underbrace{\text{train a model}}_{\text{machine learning}}$ to correctly $\underbrace{\text{classify the cavity and type of RF fault}}_{\text{multi-class classification}}$ given $\underbrace{\text{waveform data}}_{\text{time-series data}}$

Data Acquisition System

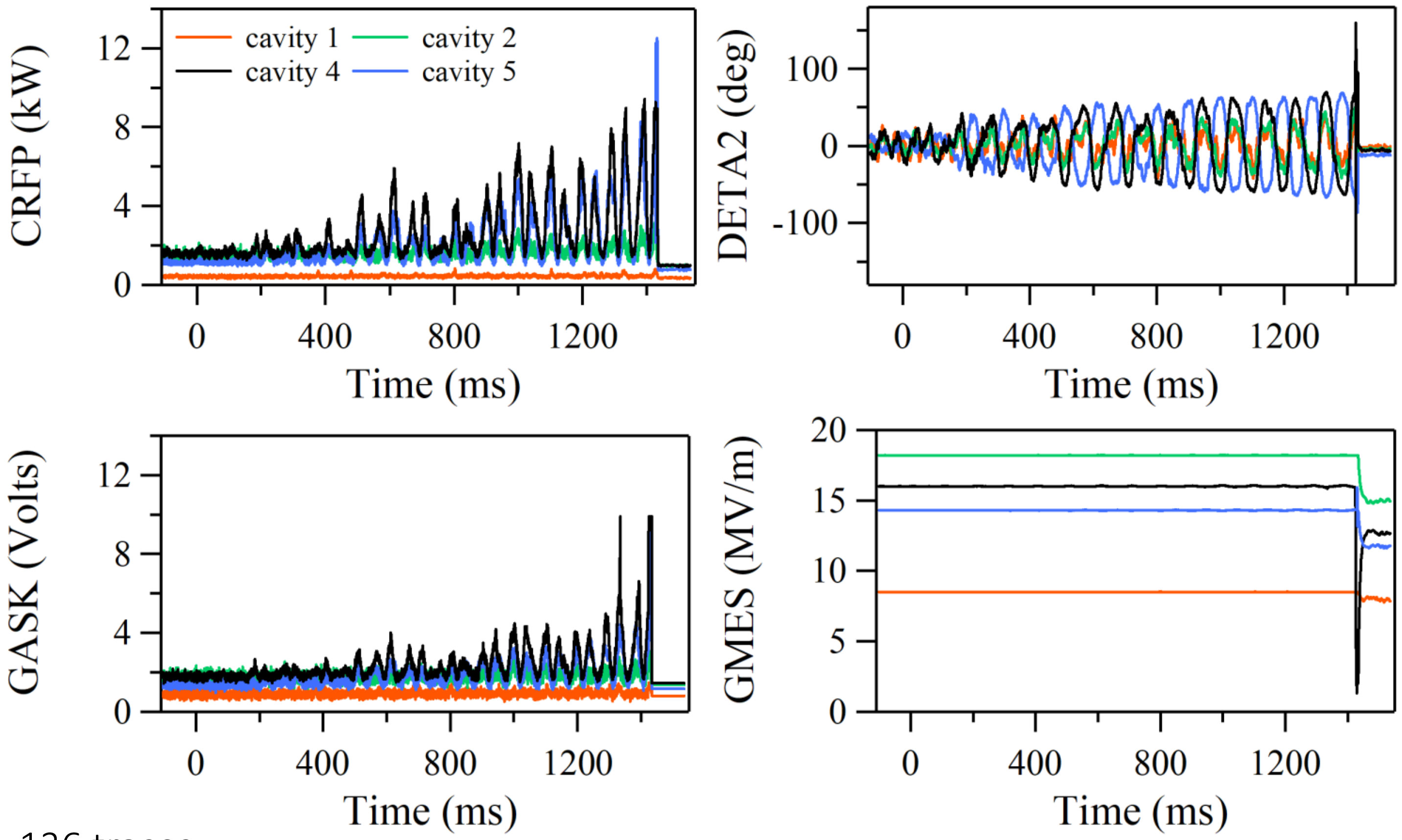
- waveform harvester was developed to capture RF time-series signals after a fault and write them to file for later analysis
  - ✓ each of the 17 harvested waveform signals is 8,192 points long
  - ✓ trigger set such that 94% of the recorded data precedes the fault and 6% after
  - ✓ pre-fault data provides valuable information about the root cause of the trip



Jefferson Lab

Motivation

- labeling is hard
  - ✓ have a subject matter expert with 30+ years SRF experience to label fault events
  - ✓ closer to annotating medical images than distinguishing between cats and dogs



17 signals/cavity × 8 cavities = 136 traces

Benefit of Fault Isolation and Identification

Post-Run Analysis

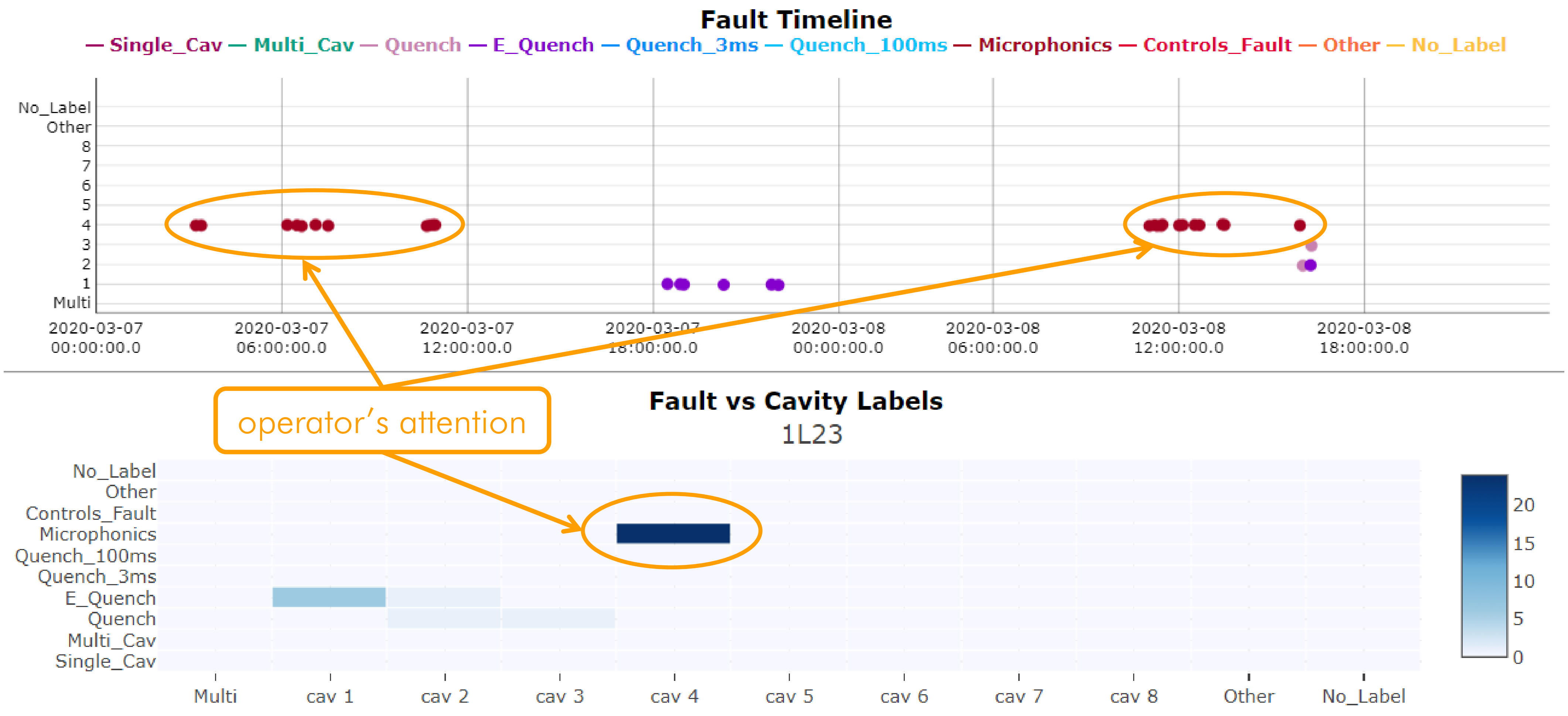
- use aggregate statistics for data-driven guidance for maintenance and/or upgrade activities
  - ✓analysis of fall 2018 data indicated three cryomodules in the South Linac were prone to microphonic-based faults → provided justification to perform microphonics hardening (installing tuner dampers) → reduced microphonics-based trip rates → gradients could be increased in those cryomodules

Post-Fault Analysis

- provides critical feedback to control room operators
- fault types get mapped to actions for the operators
  - ✓"if Fault *A* happens *X* times within *Y* minutes, drop gradient in the cavity by *Z* MV/m"
  - ✓"if Fault *B* happens *X* times within *Y* minutes, contact a SME"

Visualization and Communication

- for ML models to be effective, information must be communicated clearly and concisely
- visualize spatial and temporal nature of model predictions

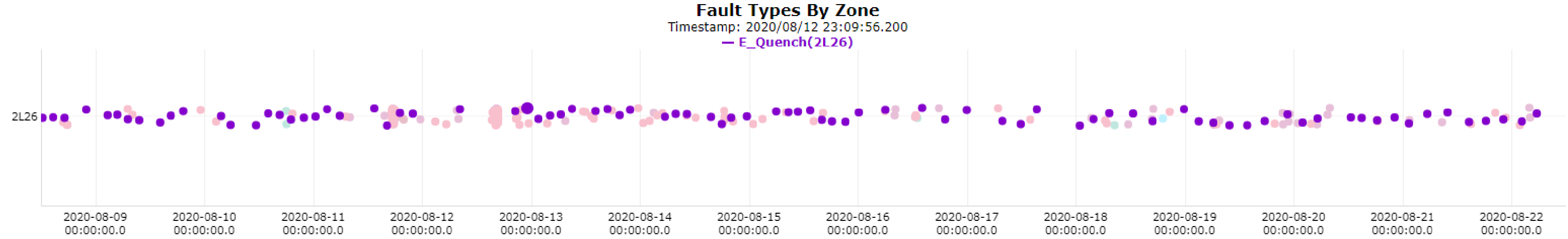


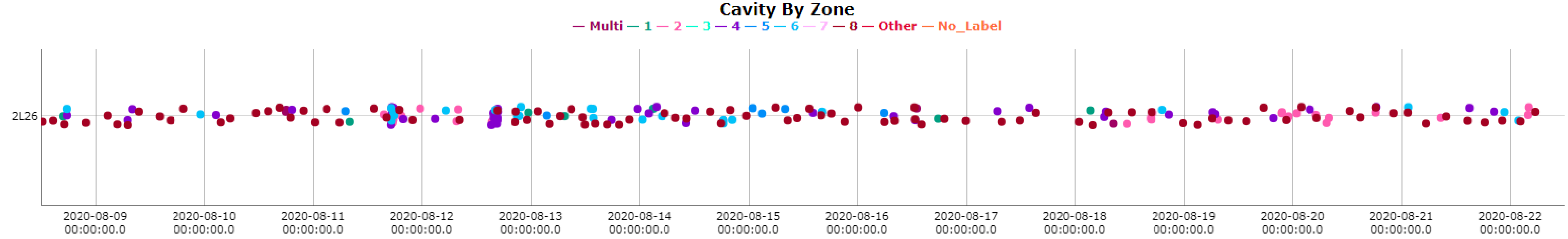
(C. Tennant, PRAB **23**, 114601 (2020))

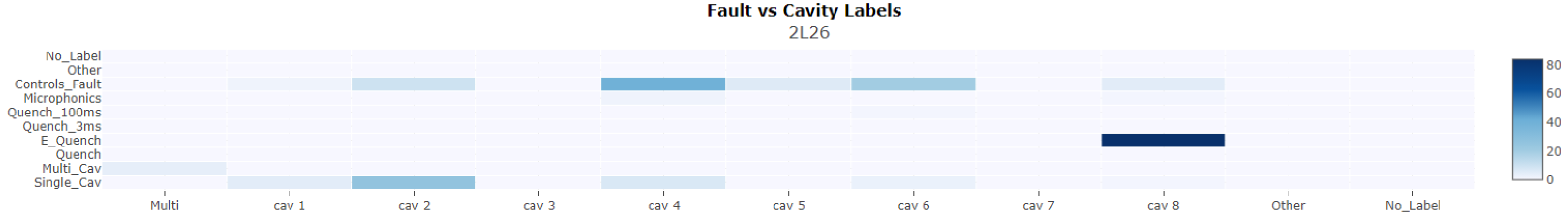
(courtesy A. Carpenter)

Post-Fault: Actionable Information

- cavity 8 in cryomodule 2L26 plagued by electronic quenches



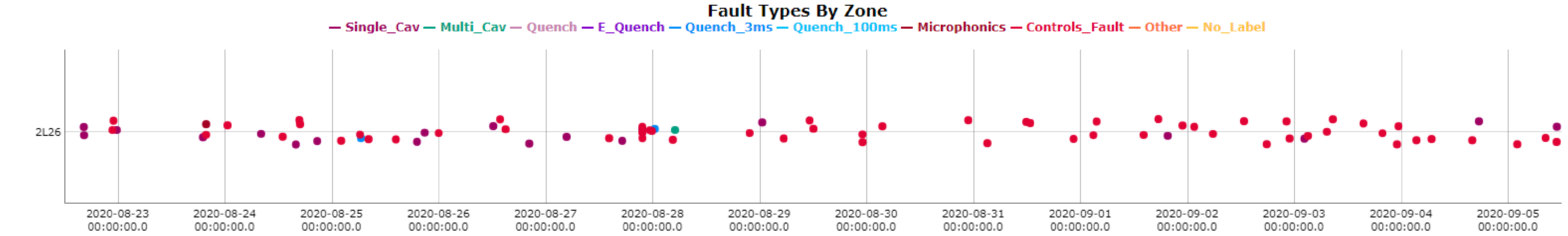


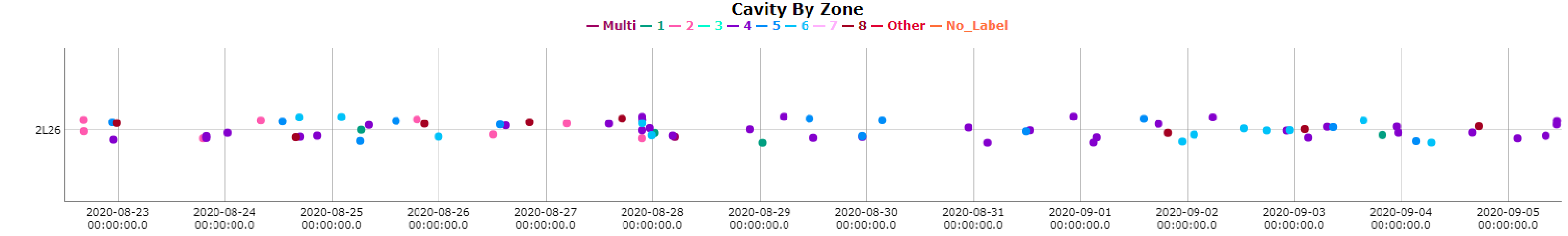


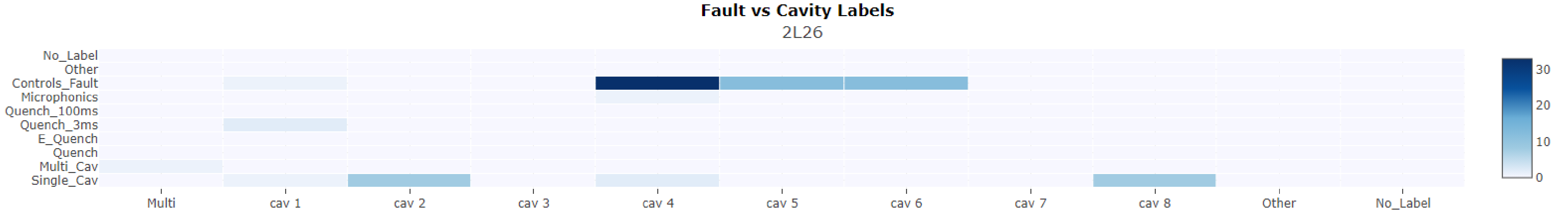
Jefferson Lab

Post-Fault: Actionable Information

- turn down gradient September 5, 2020 and faults went away completely



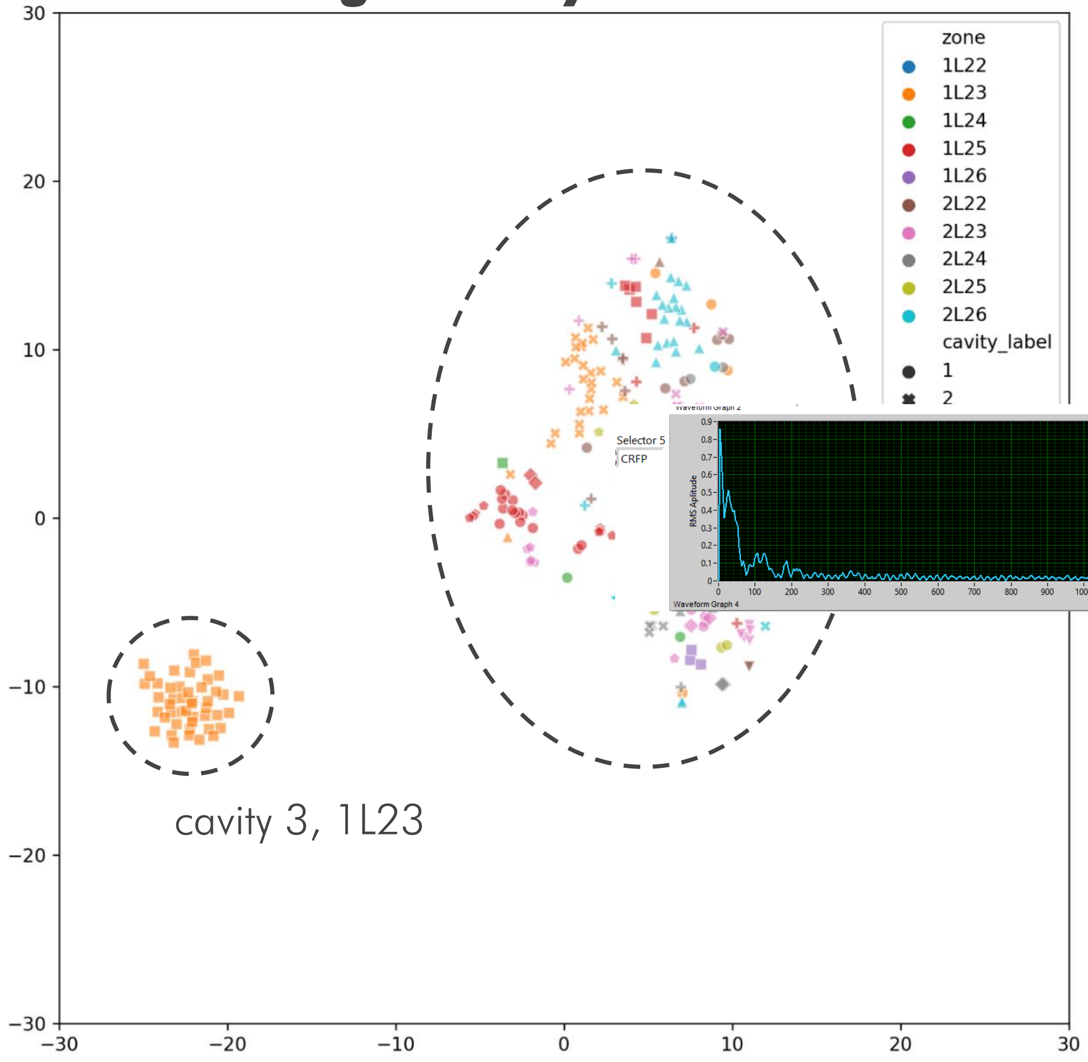




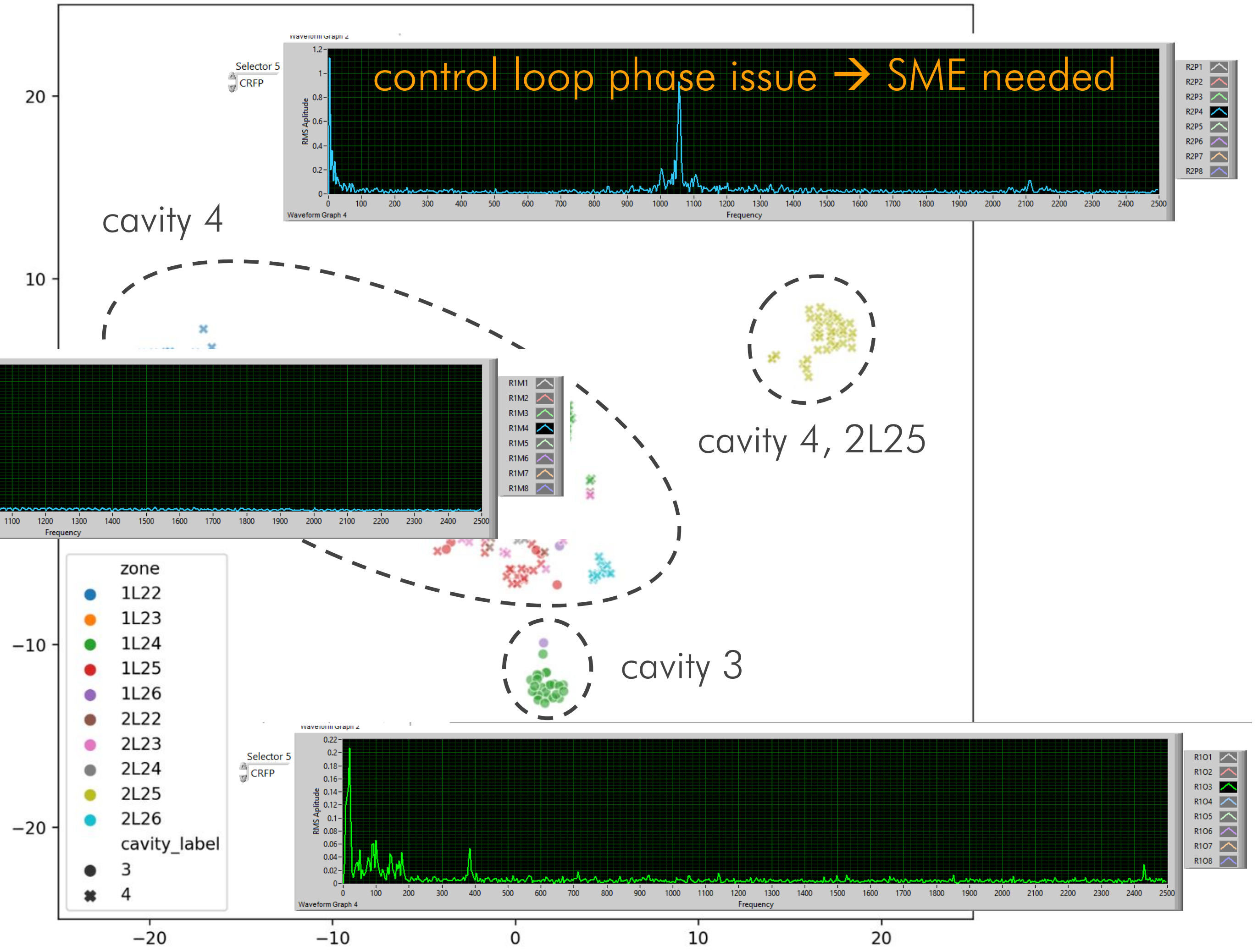
Jefferson Lab

Post-Run: Dimensionality Reduction

Single Cavity Turn-off



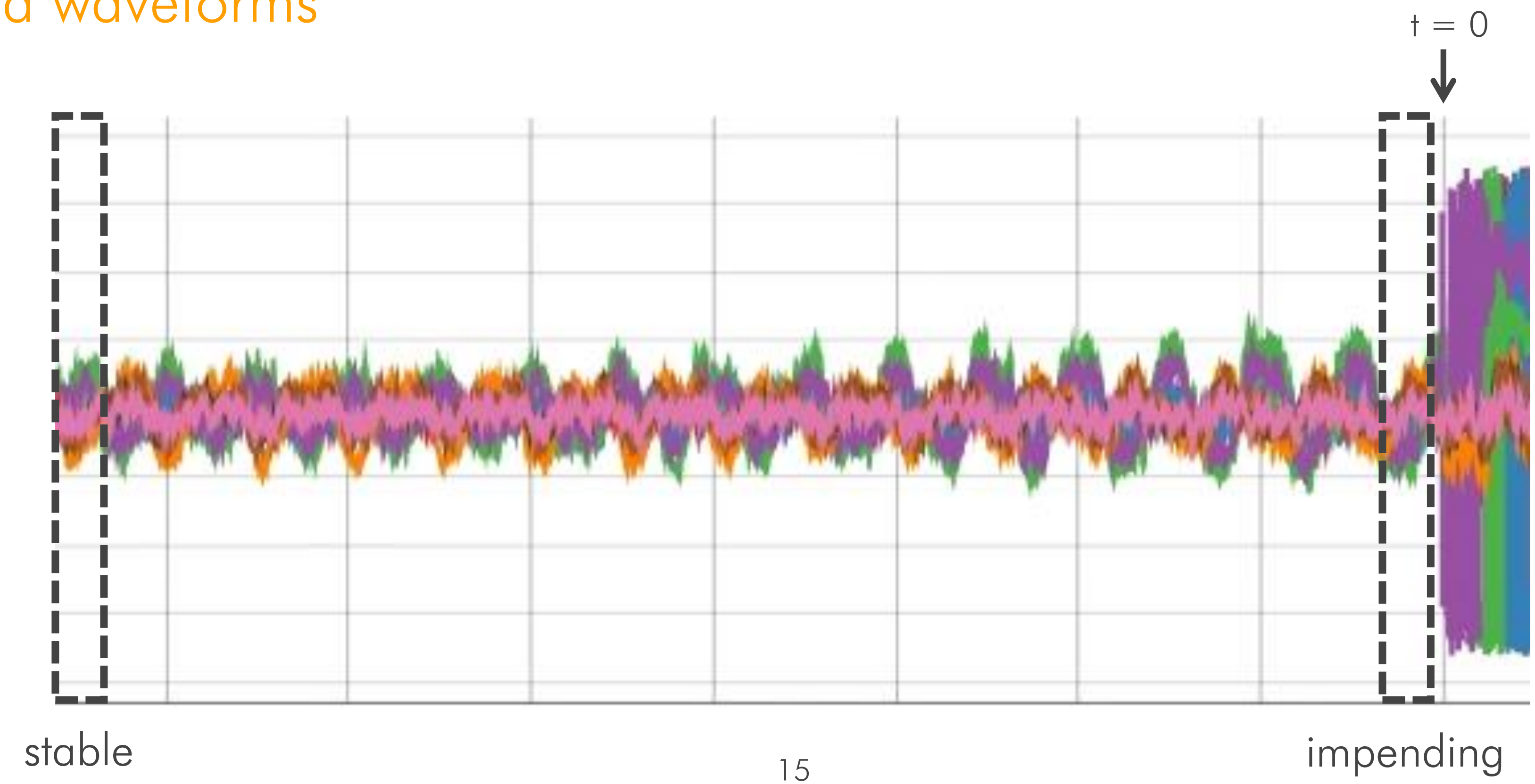
Controls Faults in Cavities 3 and 4



Jefferson Lab

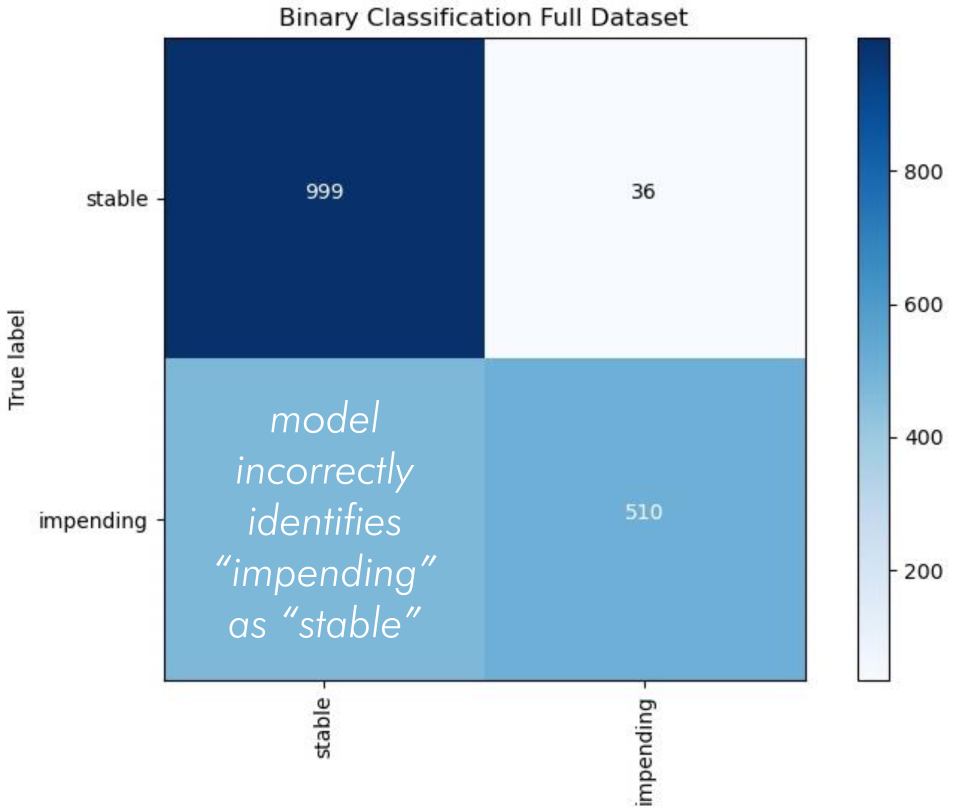
From Isolation and Identification to Prediction

- fault prediction
  - ✓ near-term: fault avoidance
  - ✓ longer-term: predictive maintenance/prognostics
- initial step: discriminate between "stable" and "impending" fault conditions
  - ✓ use saved waveforms

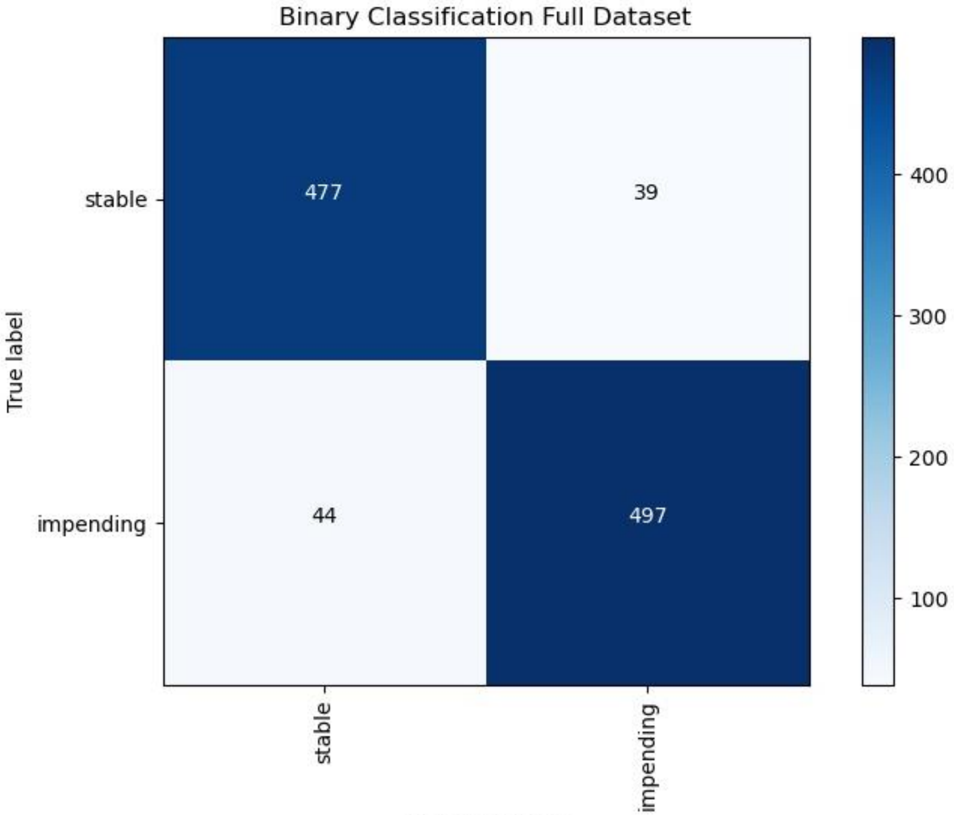


Jefferson Lab

Initial Step: Binary Classifier



accuracy = 74.74%



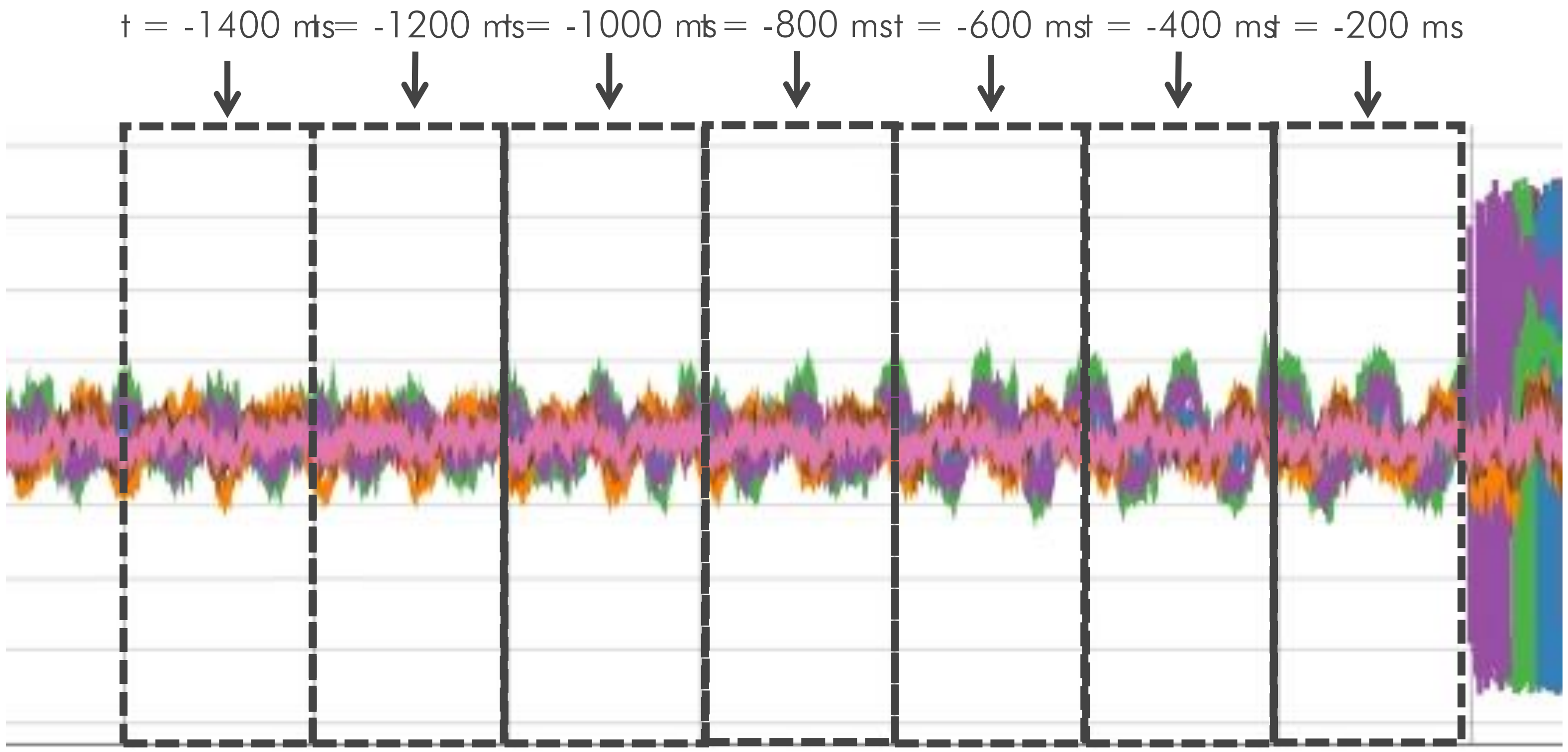
accuracy = 92.13%

- remove fault types which do not show any precursors

	Precision	Recall	f1-score	Support
Stable	0.9155	0.9244	0.9199	516
Impending	0.9272	0.9186	0.9229	541
Accuracy	0.9213			

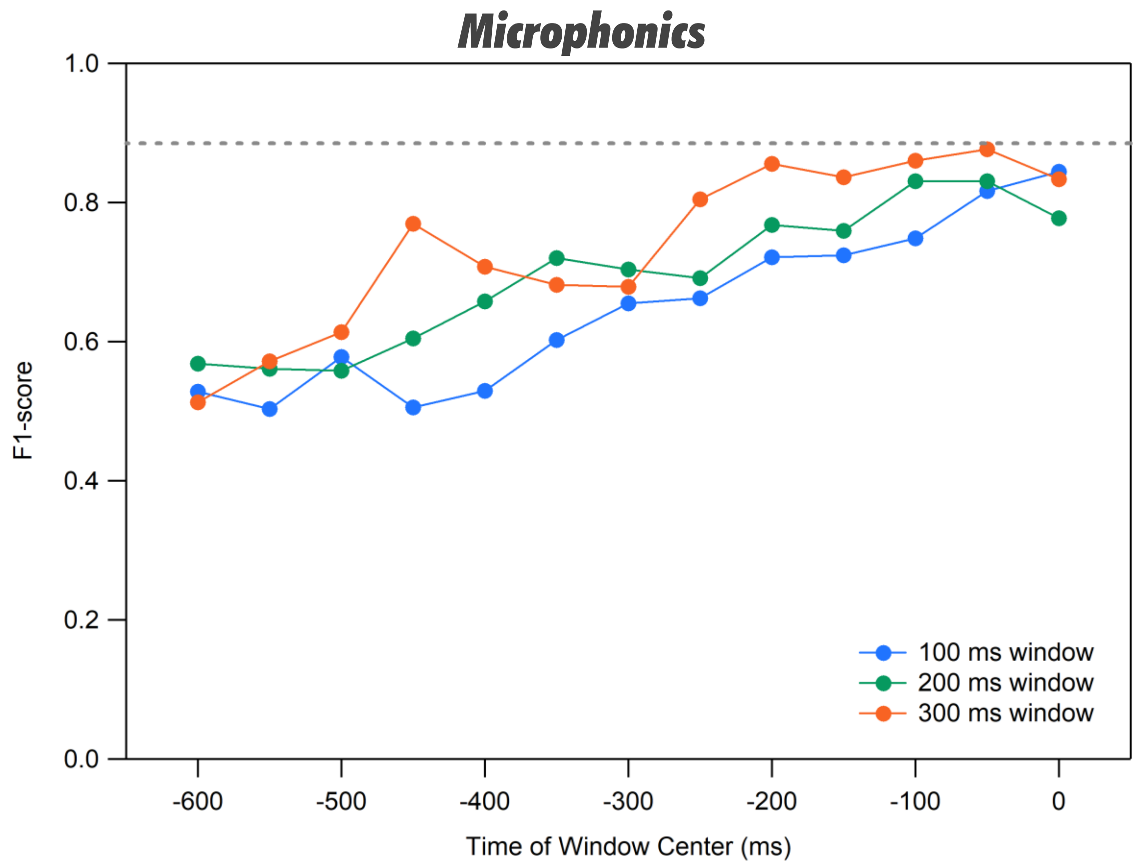
Intermediate Step: Sliding Window

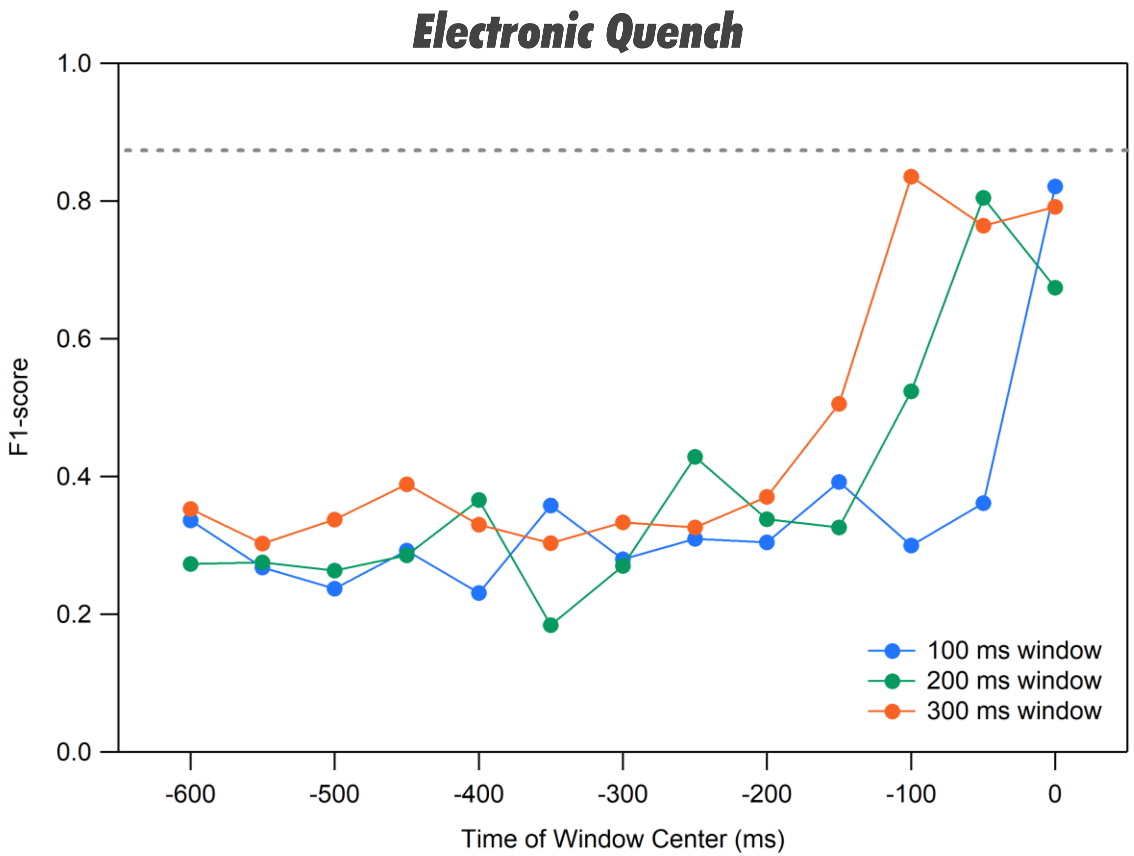
- can data prior to event accurately predict the fault type?
  - ✓ use saved waveforms



Intermediate Step: Sliding Window

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





- motivates continued study
  - ✓what kind of targeted mitigations could be implemented in those time-scales?

Data: Fueling AI

Detection → *Isolation* → *Identification* → *Prediction*

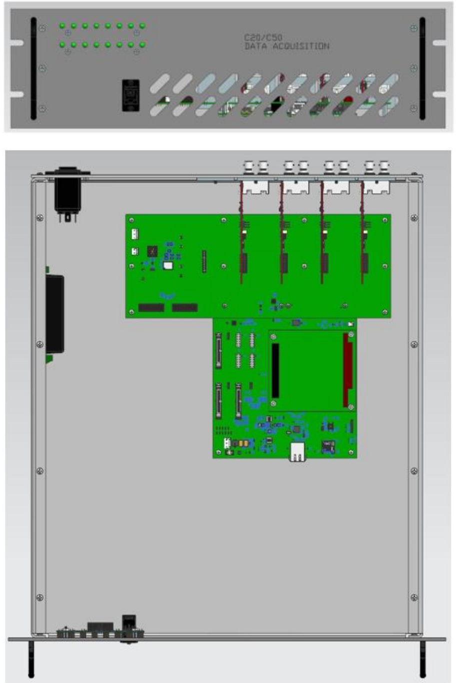
beam off

high sample frequency "snapshots"

high sample frequency streaming data

- commensurate increase in data quality required

SRF cavity instability detection in legacy cryomodules



prototype DAQ for legacy CEBAF cryomodules

field emission management

JLab designed radiation detector

Summary

- detecting, localizing (isolation) and classifying (identification) faults represent areas ripe for ML application
- the transition to fault prediction often represents an ultimate goal
- higher quality data is needed as you move along the spectrum from detection to isolation to identification to prediction
  - ✓ access to information-rich data is critical
- to achieve good performance, in addition to better data, may also need additional and/or different data
  - ✓ growth in ML must necessarily be accompanied by more and/or better data

Thank You.

tennant@jlab.org