Machine Learning for Improving Accelerator and Target Performance

HVCM Anomaly Detection

Yasir Alanazi (Jefferson Lab)

Collaborators:
M. Schram, K. Rajput, S. Goldenberg (Jefferson Lab)
C. Pappas, M. Radaideh, D. Lu, X. Zhao, P. Ramuhalli, S. Cousineau (ORNL)

ORNL is managed by UT-Battelle, LLC for the US Department of Energy
JLab is managed by Jefferson Sc. Assoc., LLC for the US Department of Energy

Agenda

1 Introduction

HVCM-background & Motivation
Project Overview

2 Data Description

Data extraction from SNS
- Normal waveforms
- Abnormal waveforms

Data preparation
Abnormal data

3 Methodology

VAE background
Anomaly detection approach
Multi-module CVAE
Results

4 Loss Landscape

Motivation
Method background
Results

5 Conclusion

1. Introduction

HVCM - Background & Motivation

- The High Voltage Converter Modulators (HVCMs) are used to power the linear accelerator (linac) klystrons at the Spallation Neutron Source (SNS)
- The HVCMs consist of multiple modules working cooperatively to produce high quality neutron beams at SNS facility
- There are 15 modules that slightly differ in their designs to accommodate the different voltage values and types of klystrons
- The HVCMs occasionally experience failures which can result in a day or more of lost operation time

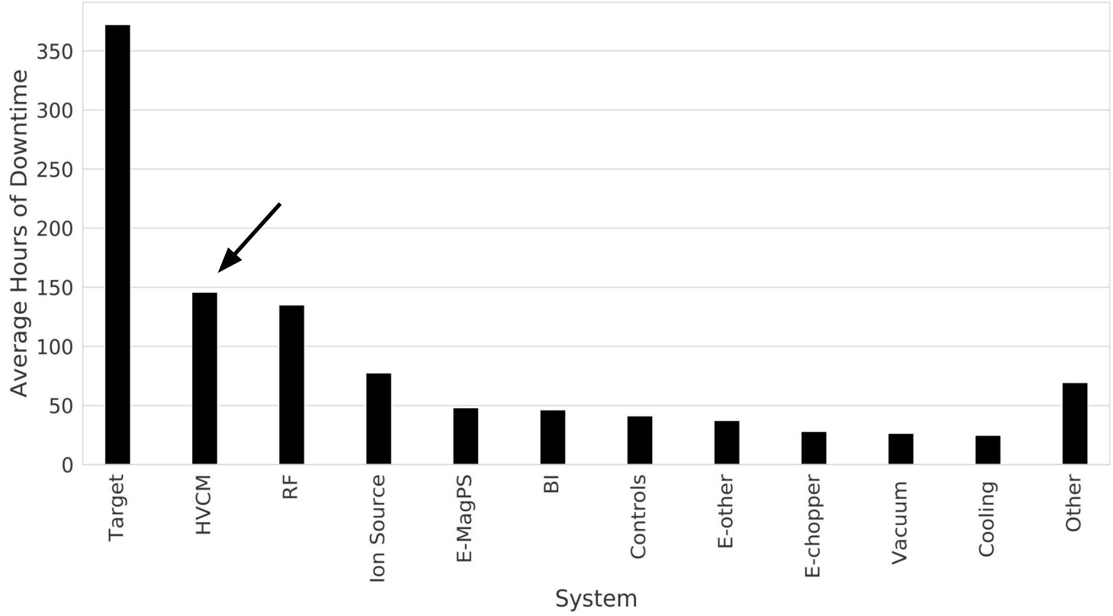


Figure 1: SNS Unscheduled Downtime by System. On average, HVCM is the second leading source of downtime after Target. The average is calculated from fiscal year 2007 to 2021.

1. Introduction

HVCM Anomaly Detection Overview

- **<u>Goal</u>**: Predict an upcoming machine failure before it occurs to improve the reliability of the HVCMs and reduce the down time for the SNS facility
- **<u>How:</u>** We use pulses leading to failure because we believe there is a sign about upcoming anomaly event before it happens
- **<u>Methodology:</u>**
  - Multi-module Conditional VAE (CVAE)
    - Train a CVAE that combines all 15 modules
    - Compare the results with a Single-module VAE
- **<u>Evaluation:</u>**
  - Use experimental data extracted from SNS
  - Evaluate the accuracy of distinguishing normal from abnormal
  - Evaluate the model loss landscape performance

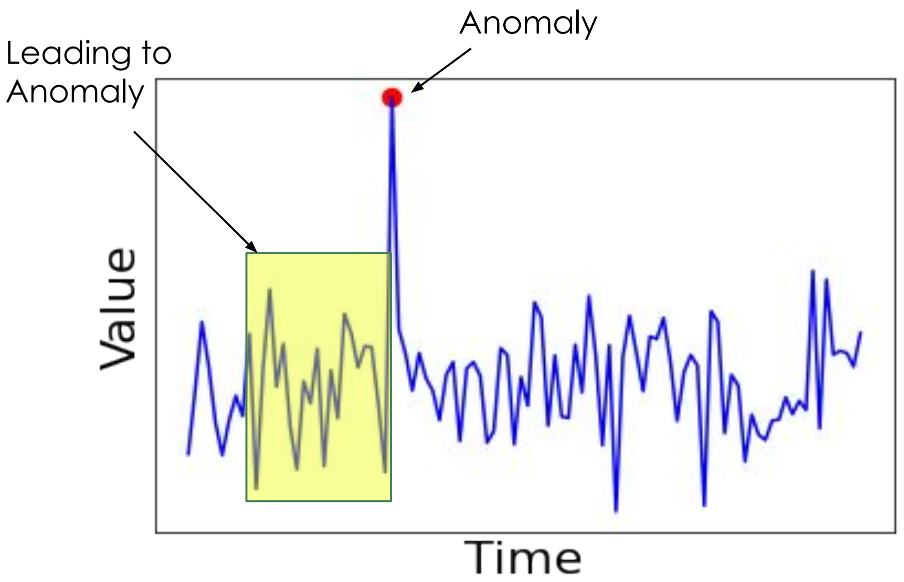


Figure 2: Toy example shows a visual representation of anomaly (red dot)) and data leading to anomaly (yellow shaded area).

Agenda

1 Introduction

HVCM background & motivation
Project Overview

2 Data Description

Data extraction from SNS
- Normal waveforms
- Abnormal waveforms

Data preparation
Abnormal data

3 Methodology

VAE background
Anomaly detection approach
Multi-module CVAE
Results

4 Loss Landscape

Motivation
Method background
Results

5 Conclusion

2. Data Description

Data Extraction

- We train and test our methodology on experimental data extracted from SNS
- Normal data: we extract all three macro-pulses and label it as "***Normal***"
- Abnormal data: we extract the first macro-pulse (pre-fault) and label it as "***Abnormal***"

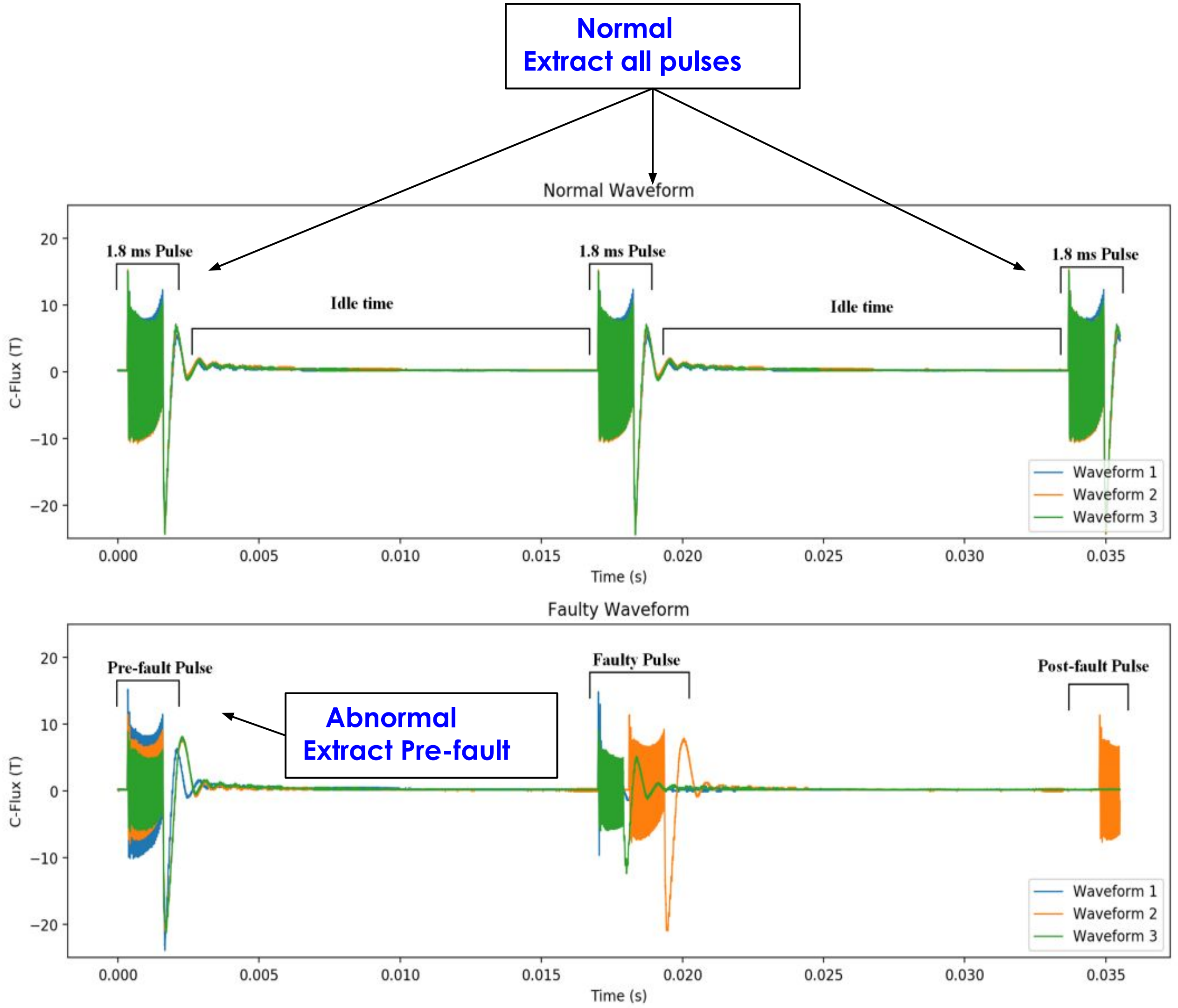


Figure 4: Top figure shows three normal macro-pulses. Lower figure shows pre-fault, fault and post fault pulses respectively.

*Courtesy to Majdi Radaideh

2. Data Description

How hard is it to identify abnormal waveforms?

- Some abnormal waveforms can be easily identified using clustering algorithms, or visualization techniques, such as, histograms, box plots, .. etc

- However, many other examples fall within the statistics of normal data and cannot be easily separated

- We need a better technique!

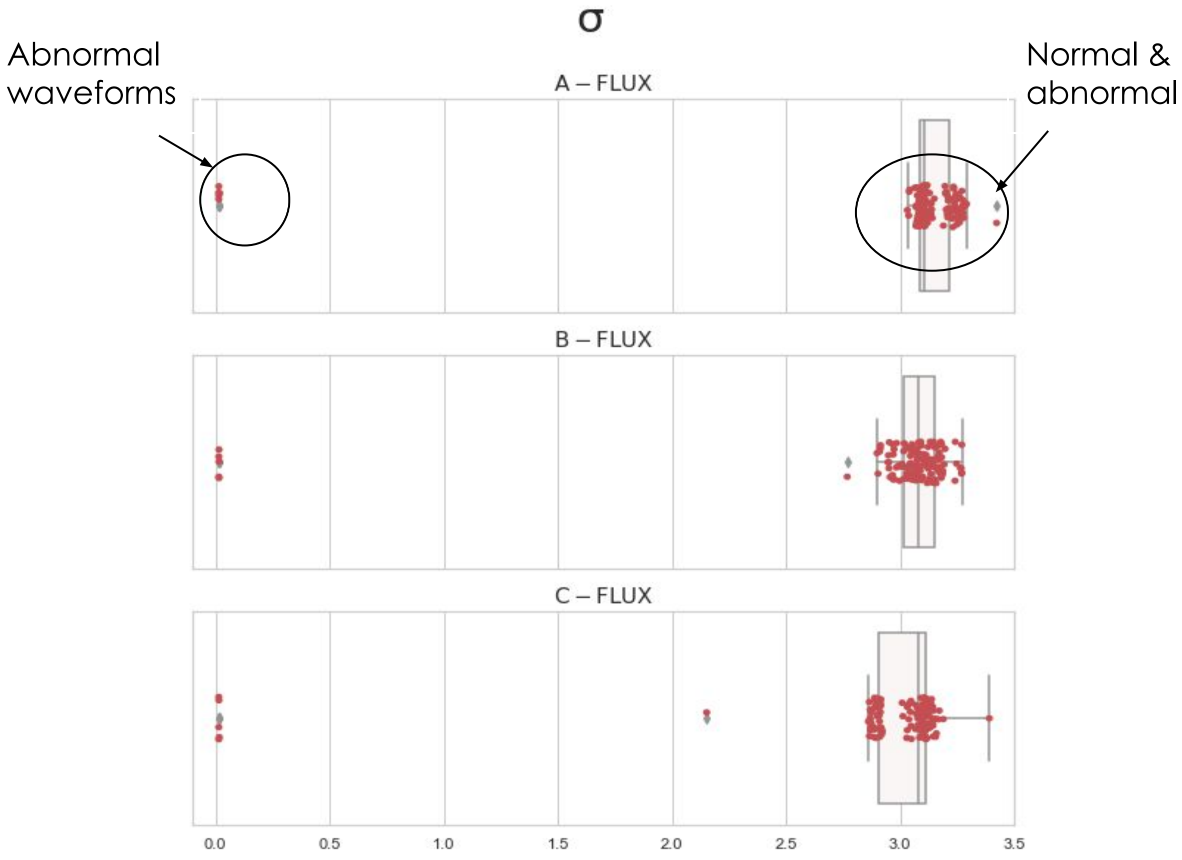


Figure 8. The std distributions of normal & abnormal waveforms

OAK RIDGE National Laboratory | SPALLATION NEUTRON SOURCE

Jefferson Lab

Agenda

1 Introduction

HVCM background & motivation
Project Overview

2 Data Description

Data extraction from SNS
- Normal waveforms
- Abnormal waveforms

Data preparation
Abnormal data

3 Methodology

VAE background
Anomaly detection approach
Multi-module CVAE
Results

4 Loss Landscape

Motivation
Method background
Results

5 Conclusion

3. Methodology

What is a Variational Autoencoder (VAE)?

- Variational Autoencoder (VAE) is class of Machine Learning that provides a probabilistic manner for describing input data in latent space

- VAE consists of two Neural Networks:
    - <u>Encoder</u>: projects the input data into a probability distribution by estimating a mean and a standard deviation parameters of that distribution

    - <u>Decoder</u>: learns how to reconstruct the data from the learned distributions

- The model loss function consists of:
    - Kullback–Leibler divergence:
        - Ensure the prior distribution to be as close as possible to the estimated one
    - Reconstruction error:
        - Minimize the difference between input and output data (e.g. Mean Squared Error)

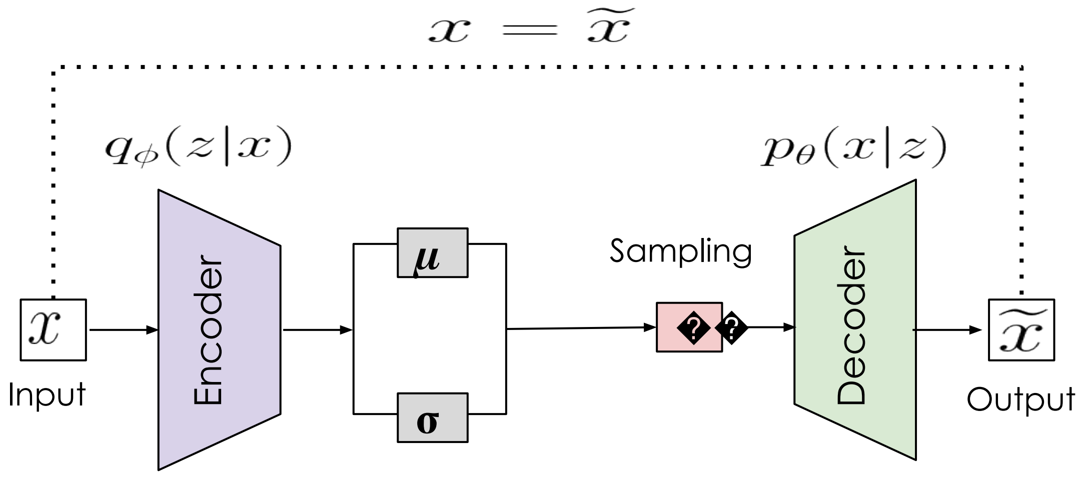
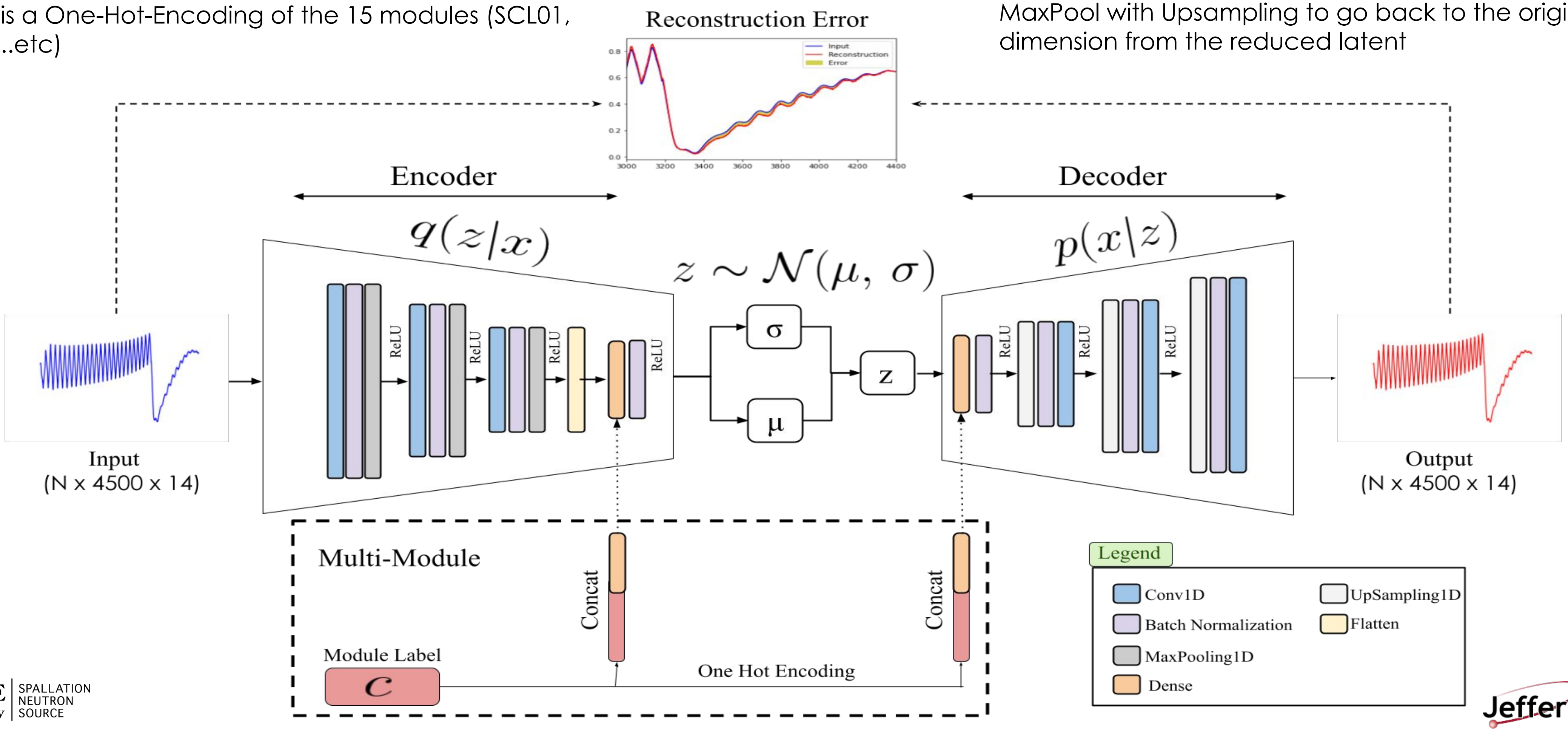


Figure 10: A typical VAE consists of an encoder that projects the input data into a smaller representation (z), and a decoder that takes z as inputs to reconstruct the input data.

3. Methodology

Multi-module CVAE

- We implement a Multi-module Conditional VAE (CVAE), motivated by the architecture or CVAE*
- The model is conditioned by the component c (red box), which is a One-Hot-Encoding of the 15 modules (SCL01, SCl05, ..etc)
- The encoder uses three 1D CNN blocks (Conv1D, BN, MaxPool)
- The decoder uses also three 1D CNN block, but replaces MaxPool with Upsampling to go back to the original dimension from the reduced latent



3. Methodology

Results

- Three faults (DV/DT, FLUX, and IGBT)
- Kernel density estimate (KDE) plot to show the distributions of the reconstruction error using, Mean Squared Error (MSE)
- The Receiver Operating Characteristic (ROC) curve shows the accuracy performance at various threshold settings
- The other faults show reasonable separation with AUC values ranging from 0.83 to 0.93 (they can be found in the paper*)

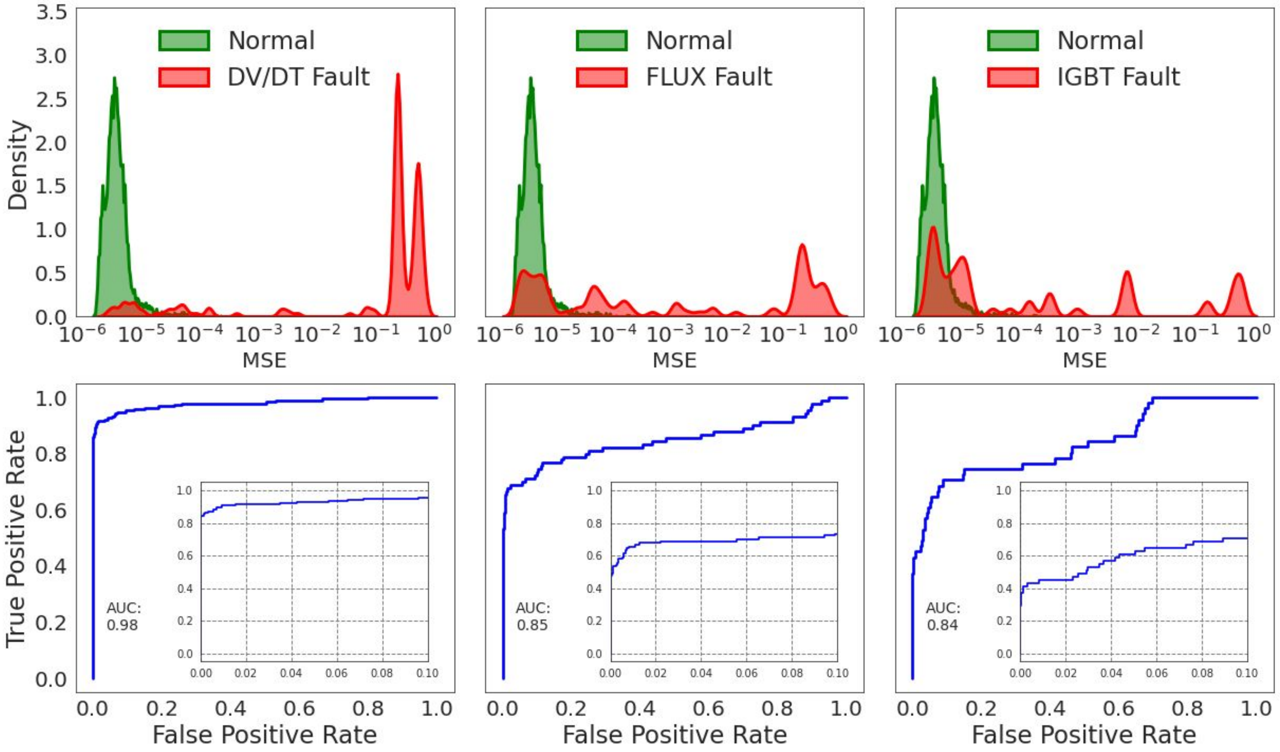


Figure 13: KDE distributions of the MSE from reconstructing normal (grey color) and faulty waveforms (blue color) for six fault types, with the corresponding ROC curve for each fault.

*https://arxiv.org/pdf/2304.10639.pdf

3. Methodology

Results: Comparison

- Compare the results with a Single-module based VAE
- Single-module is a VAE that is trained individually for each module
- The Multi-module (top row) shows <u>smaller reconstruction error,</u>
- for all SCL modules (SCL01, SCL05, ..etc), using 3-FLUXs features
- This allows to increase the separation between normal and abnormal waveforms

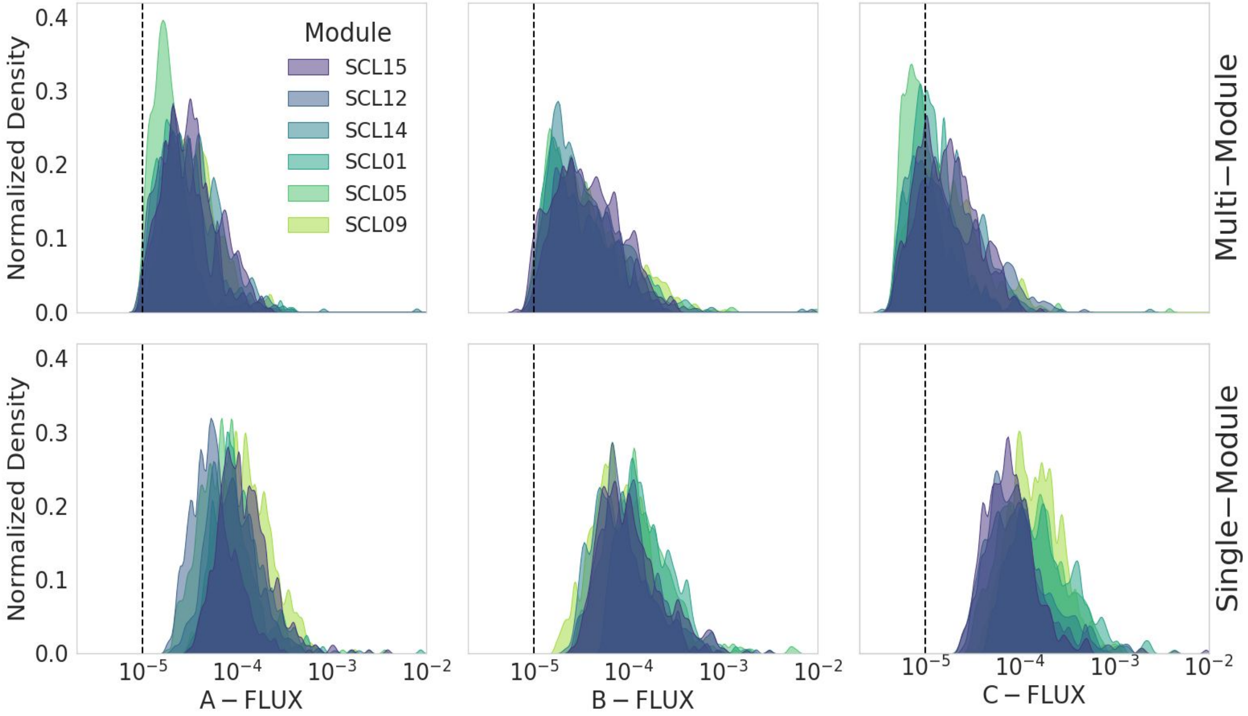


Figure 14: Normalized density estimation plot shows the reconstruction error of normal waveforms using multi-module and single-module. The multi-module model shows smaller reconstruction error, where the distributions of most of the individual systems are more shifted to the dashed black line at MSE= 10^{-5}

3. Methodology

Results: Comparison with errorbar

- Compare AUC values between two methods
- Multi-module has higher AUC values for almost all faults and modules
- The error bar is generated by using the probabilistic encoder model
- Sampling from the estimated parameters (mean & standard deviation) at inference time
- 100 replicas are generated from the model

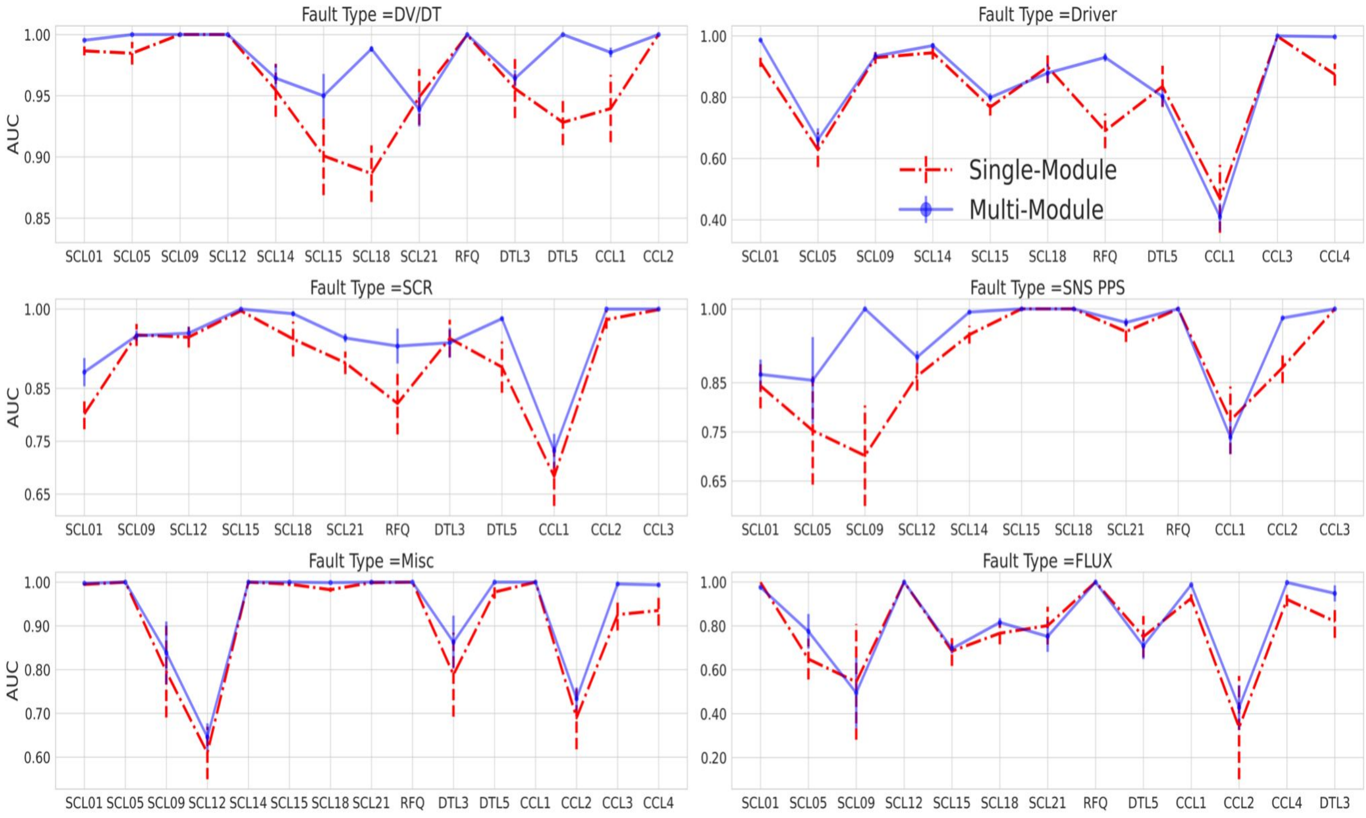


Figure 15: Compare the AUC values between single-module and multi-module using six types of faults across several modules. The error bar is plus/minus 1 Standard Deviation (SD) error generated by sampling the latent Z of each method

Agenda

1 Introduction

HVCM background & motivation
Project Overview

2 Data Description

Data extraction from SNS
- Normal waveforms
- Abnormal waveforms

Data preparation
Abnormal data

3 Methodology

VAE background
Anomaly detection approach
Multi-module CVAE
Results

4 Loss Landscape

Motivation
Method background
Results

5 Conclusion

4. Loss Landscape

Motivation

- The performance of NNs can be impacted by several factors such as variable initialization, optimizers, network architectures, batch sizes, ..etc
- Studying the effects of various hyper-parameters is challenging because their loss values live in a high-dimensional space
- Looking at 1D loss curve does not tell us the whole story!
- We need more information about the internal behaviour of the model

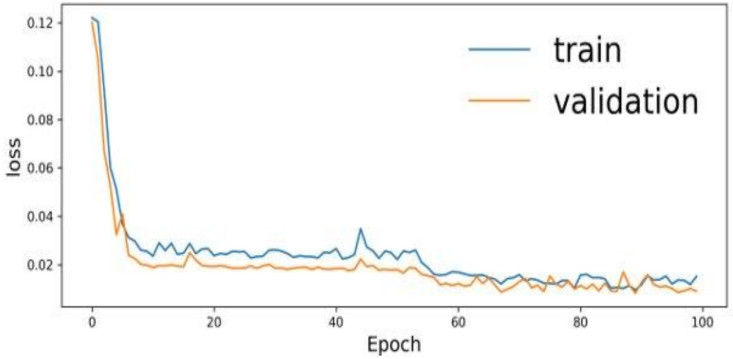


Figure 16: Loss values of VAE trained on CCL module

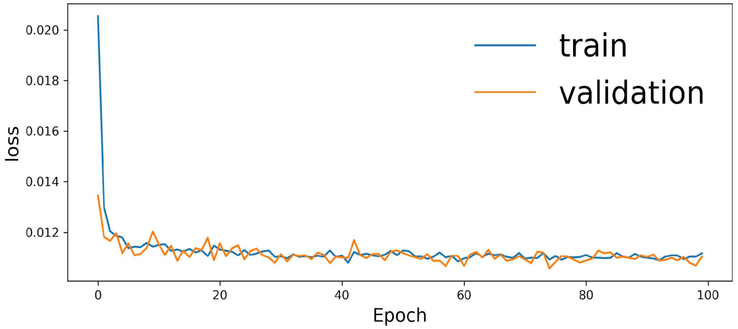
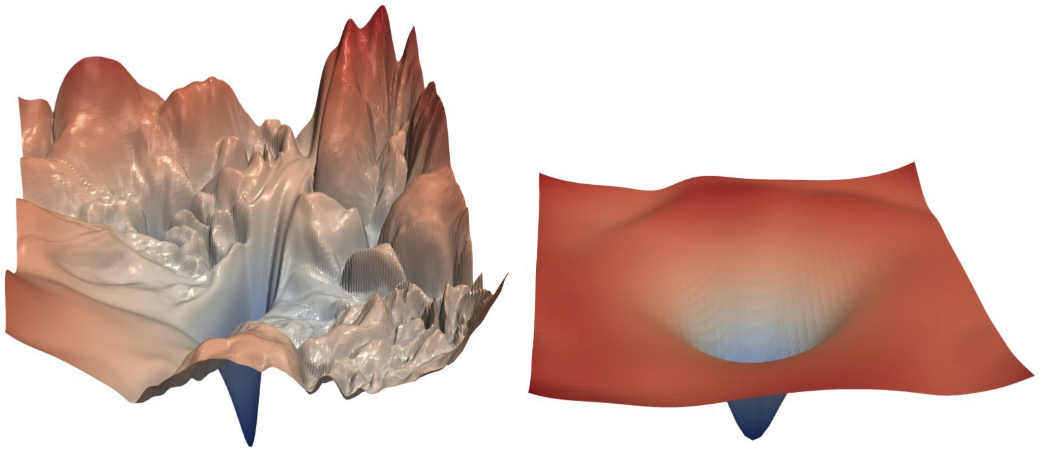


Figure 17: Loss values of VAE trained on all modules

4. Loss Landscape

Filter Normalization background

- It has been proposed to visualize the loss landscape
- The loss landscape can show the convexity/non-convexity of the trained models
- Can explain why certain choice of NNs architectures are easier to train than others. (i.e. skip connections)
- Smooth loss landscapes (right plot) tend to generalize better



*Hao Li, and et al, Visualizing the Loss Landscape of Neural Nets. NIPS, 2019.

OAK RIDGE National Laboratory | SPALLATION NEUTRON SOURCE

Jefferson Lab

4. Loss Landscape

Filter Normalization: Results

- Visualize the loss landscape of Single- and Multi-module
- We show the results for CCL4 module
- Single-module has chaotic loss surface, while Multi-module has smooth, convex-like loss surface
- Is this due to the random weights initialization?

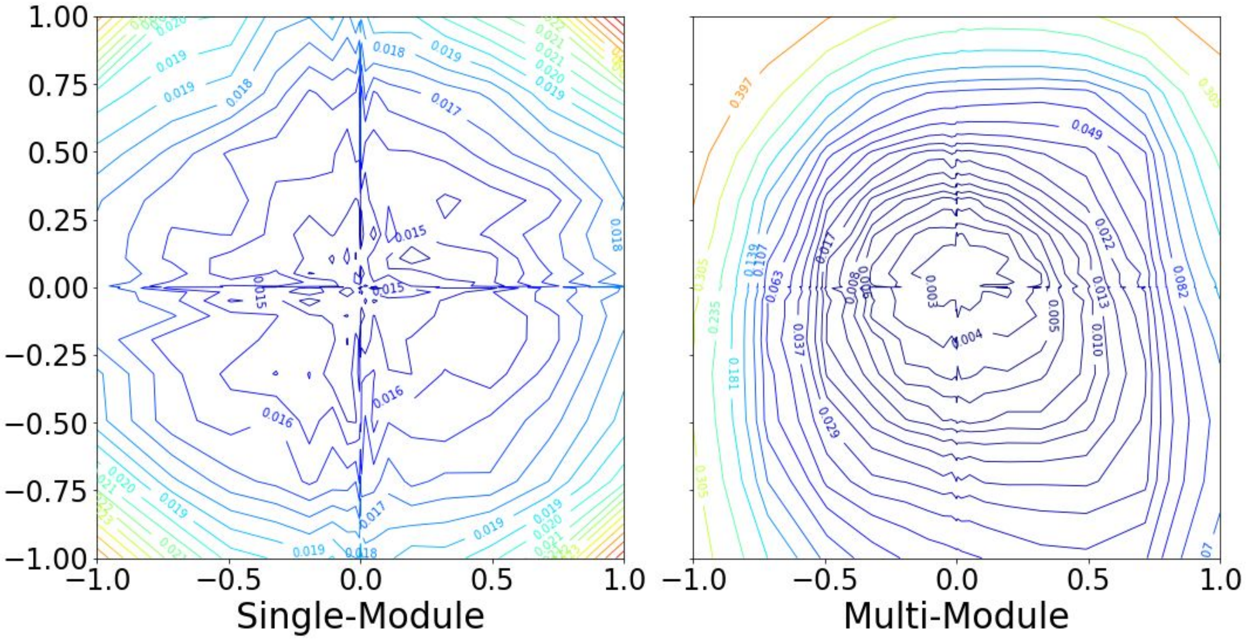


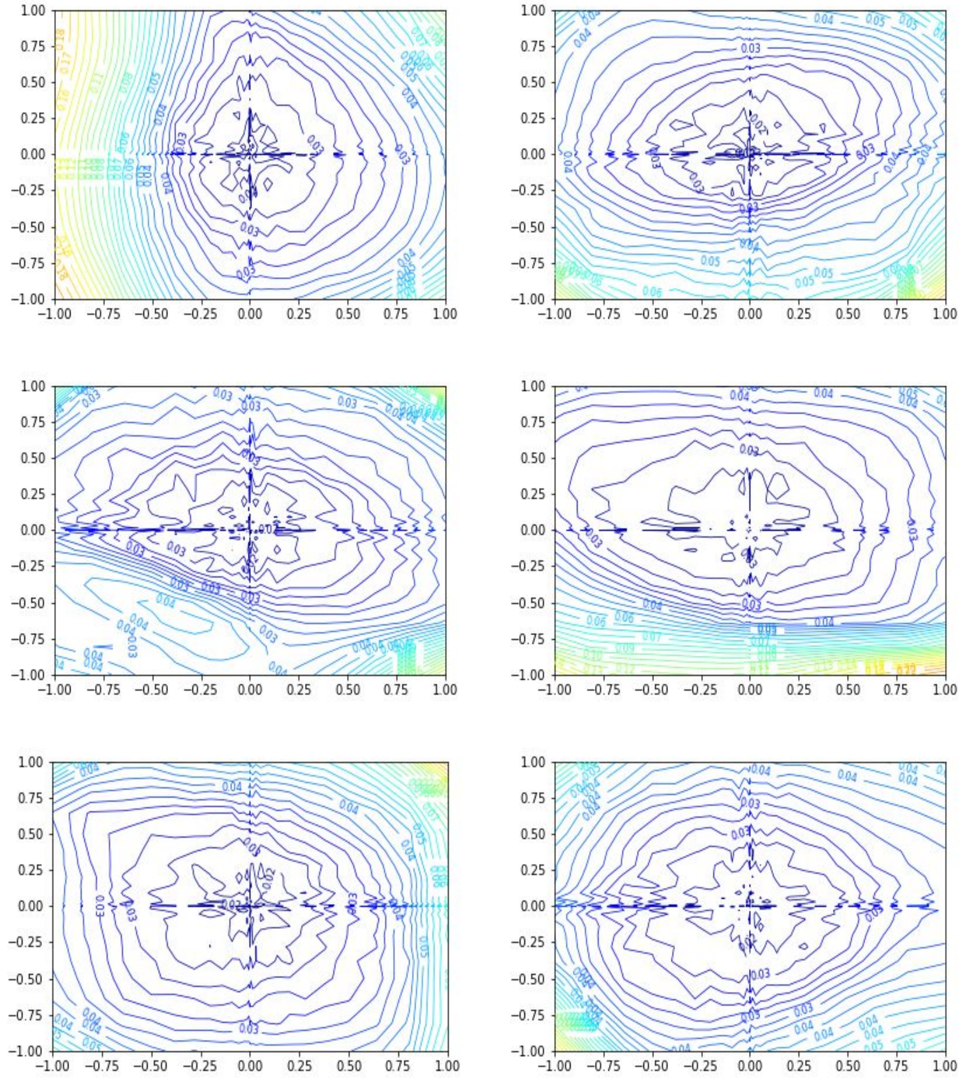
Figure 18: Loss surface, where x- and y-axis are two random directions in weights space generated using filter normalization method.

4. Loss Landscape

Filter Normalization: Results

- Train Single-module using CCL4 file multiple times
- For each replica, we start with different weights initialization
- The results show that for multiple trials, Single-module produces chaotic loss surface
- This suggests that Hyperparameter optimization (HPO) and Neural Network Search (NSA) is needed for each module when train Single-module
- But only needed once for Multi-module model!

Single-Module



OAK RIDGE National Laboratory | SPALLATION NEUTRON SOURCE

Jefferson Lab

Agenda

1 Introduction

HVCM background & motivation

Project Overview

2 Data Description

Data extraction from SNS

- Normal waveforms
- Abnormal waveforms

Data preparation

Abnormal data

3 Methodology

VAE background

Anomaly detection approach

Multi-module CVAE

Results

4 Loss Landscape

Motivation

Method background

Results

5 Conclusion

5. Conclusion

- Implementing a Conditional Variational Autoencoder (CVAE) to detect anomalies in the HCVMs at SNS

- Multi-module can laren from different files and generalize better than Single-module approach

- Using the probabilistic encoder model, Multi-module produces higher AUC values for almost all faults with smaller uncertainty band

- Using loss landscape analysis, Multi-module shows convex-like loss surface, while Single-module has chaotic behaviour

References & Acknowledgment

References:

[1] Yasir Alanazi, Malachi Schram, Kishansingh Rajput, Steven Goldenberg, Lasitha Vidyaratne, Chris Pappas, Majdi I. Radaideh, Dan Lu, Pradeep Ramuhalli, and Sarah Cousineau: Multi-module based CVAE to predict HVCM faults in the SNS accelerator. arXiv:2304.10639 [cs], 2023.

[2] Majdi I. Radaideh; Chris Pappas; Jared Walden; Dan Lu; Lasitha Vidyaratne; Thomas Britton; Kishansingh Rajput; Malachi Schram; Sarah Cousineau. Time Series Anomaly Detection in Power Electronics Signals with Recurrent and Convlstm Autoencoders. DOI: 10.2139/ssrn.4069225, 2022

[3] Sohn, Kihyuk and Lee, Honglak and Yan, Xinchen. Learning Structured Output Representation using Deep Conditional Generative Models. Advances in Neural Information Processing Systems, 2015.

[4] Hao Li and Zheng Xu and Gavin Taylor and Christoph Studer and Tom Goldstein. Visualizing the Loss Landscape of Neural Nets. arXiv:1712.09913 [cs], 2018

Acknowledgment:

- This research has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The Jefferson Science Associates (JSA) operates the Thomas Jefferson National Accelerator Facility for the DOE under Contract No. DE-AC05-06OR23177. This research used resources at the Spallation Neutron Source, a DOE Office of Science User Facility operated by the Oak Ridge National Laboratory.

Thank You!

Yasir Alanazi
alanazi@jlab.org

Backup

2. Data Description

Data Preparation

- After data extraction, we save 14 waveforms into a 3D tensor of shape (samples, timestep, features)
  - where timestep is 4500 and 14 features represent different waveforms

- The 14 waveforms are:
  - Six IGBT current waveforms
  - Three magnetic flux density in the phases A, B, and C of the resonant circuit (Figure 5 shows an example)
  - Two waveforms represent the cap bank voltage and the cab bank current
  - Two waveforms represent the modulator output voltage and the modulator current
  - One waveform represents the time change of the modulator output voltage

- The total number of normal samples is 7246 for all modules combined, where the number of abnormal samples is 1080 waveforms

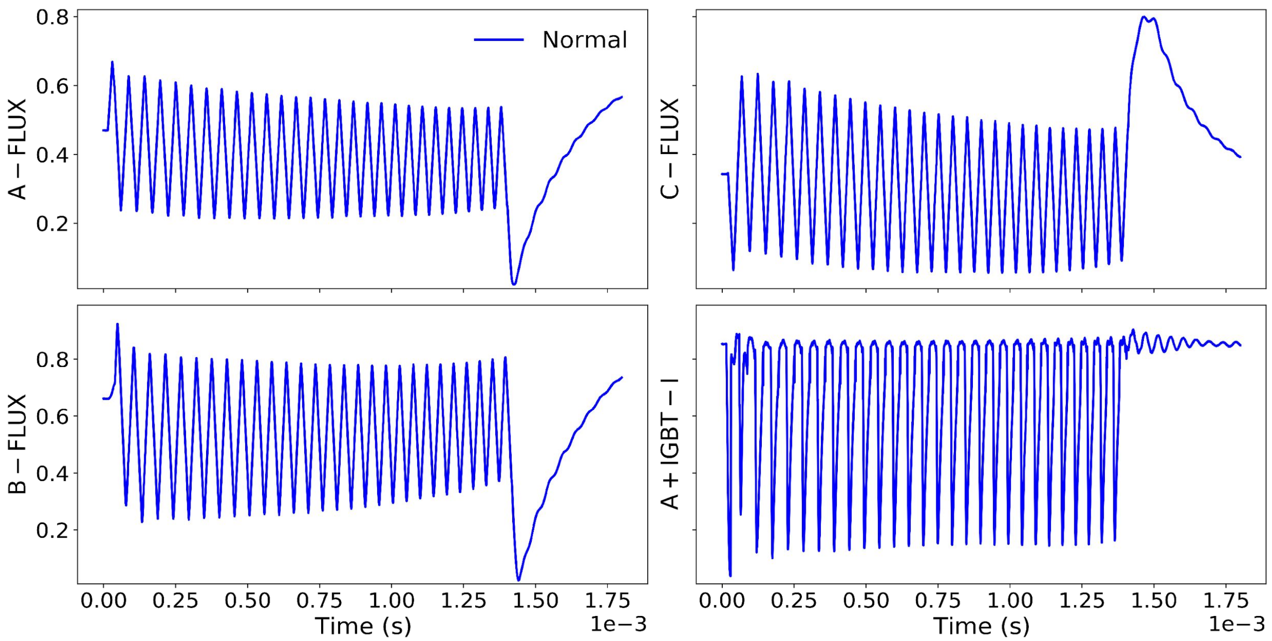


Figure 5: Different representation of a normal example (A, B, and C FLUX, and A+IGBT-I). The waveforms are normalized between 0 and 1.

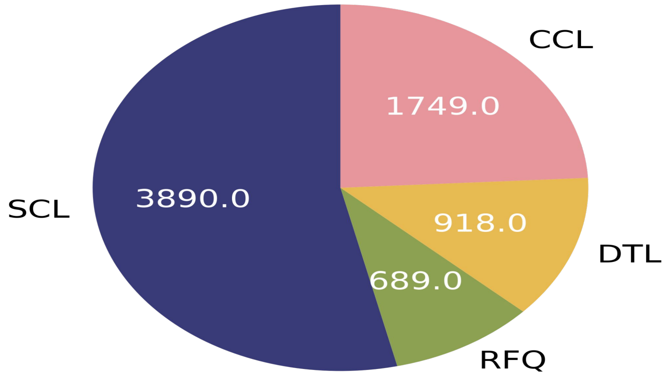


Figure 6: Number of normal samples for the 4 main modules.

2. Data Description

Faults Grouping

- There are several abnormal waveform types exist for different modules (Figure 7 - outer bar chart)
- In this work, we group the the abnormal waveforms into 9 related categories (Figure 7 - inner bar chart)
- This increases the number of statistics and allows for more meaningful results evaluation
- The grouped abnormal categories are:
  - DV/DT, Driver, SCR, SNS PPS, Misc, FLUX, IGBT, CB, and TPS
- In our analysis we focus on detecting the grouped 9 faults (Figure 7 - inner bar chart)

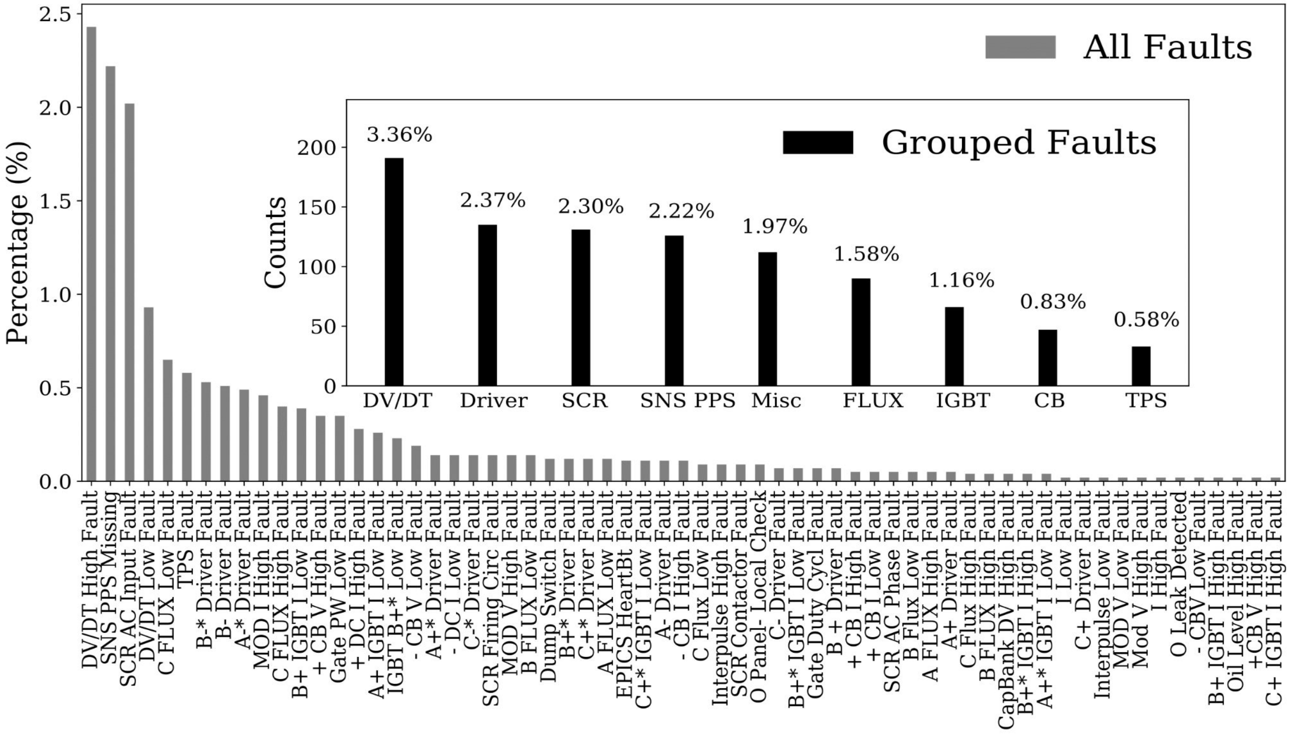


Figure 7: The outer figure (grey bars) shows percentage of abnormal waveform types with respect to all data including normal. The inner figure (black bars) shows the counts of abnormal data after regrouping

3. Methodology

Results

- Box plot shows the reconstruction error, Mean Squared Error (MSE) distribution of Normal (left), and Abnormal (right) for each module using different waveform features
- Overall, the normal examples show smaller MSE than abnormal, allowing us to set a threshold to classify them
- There are 1080 samples for normal and abnormal samples, where abnormal includes all fault types combined

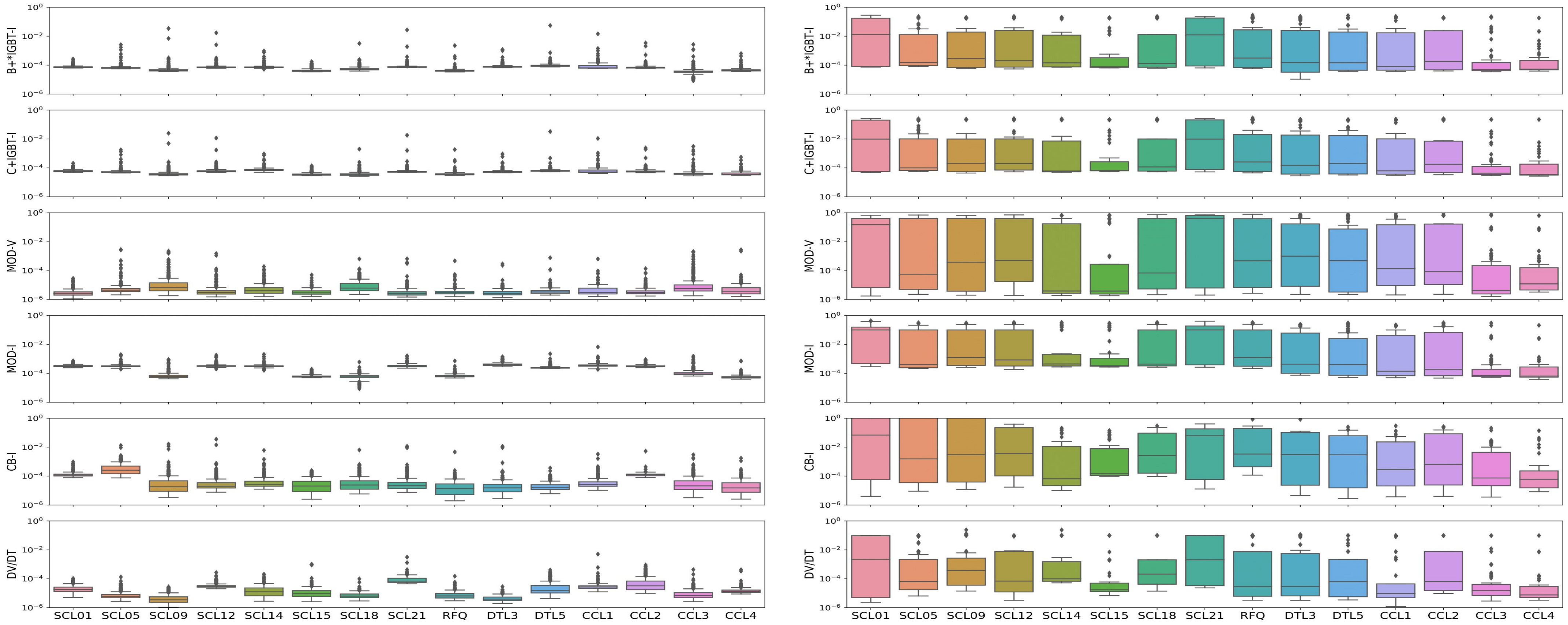


Figure 12: Box plot shows the MSE reconstruction error distributions of normal (left plot) and abnormal (right plot) using multi-module. The x-axis shows all trained modules and the y-axis shows six source of waveforms.

OAK RIDGE National Laboratory | SPALLATION NEUTRON SOURCE

Jefferson Lab

3. Methodology

Results: Multi-module vs Single-module

- We also compare our results with a single-module based VAE
- Single-module is a VAE that is trained individually for each module, where in this case it has been trained 12 times for the SCLs, RFQ, CCLs, and DTLs
- Recap, multi-module is only trained once by combining all modules together
- We show normalized KDE between the two methods. As expected, the multi-module is learning from the multiple modules and produce lower MSE values than single-module when reconstructing normal waveforms
- The multi-module model shows smaller reconstruction error, where the distributions of most of the systems are more shifted to the dashed black line at MSE=10^{-5}
- The results show that when combining all models the model can learn more and produce smaller error, which allows us to have higher separation between normal and abnormal behavior

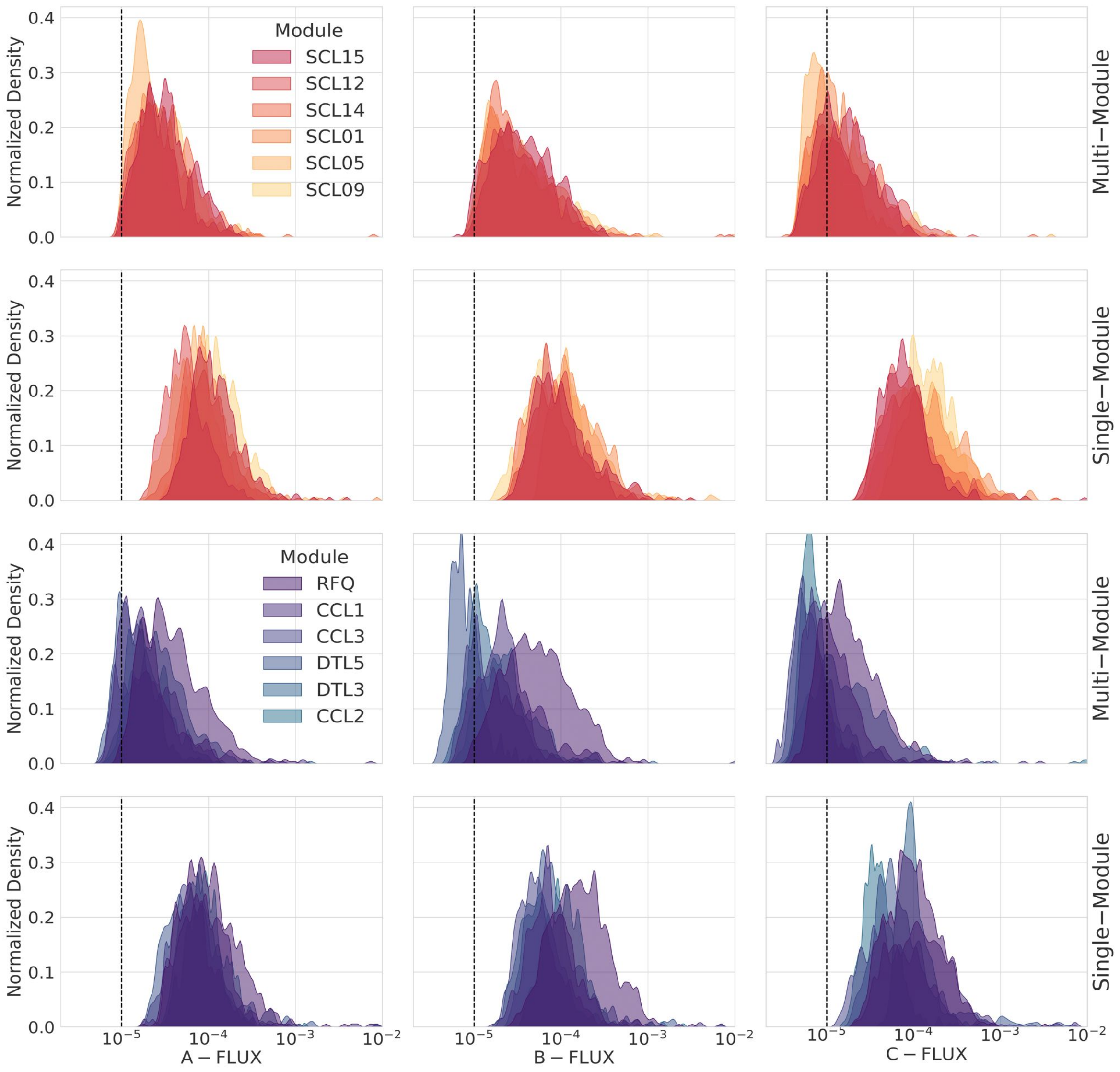


Figure 14: Normalized density estimation plot shows the reconstruction error of normal waveforms using multi-module and single-module. The multi-module model shows smaller reconstruction error, where the distributions of most of the individual systems are more shifted to the dashed black line at MSE= 10^{-5}

OAK RIDGE National Laboratory | SPALLATION NEUTRON SOURCE

Jefferson Lab

3. Methodology

Results

- Three faults (DV/DT, SCR, and SNS PPS)
- Kernel density estimate (KDE) plot to show the distributions of the reconstruction error using,
- Mean Squared Error (MSE)
- The Receiver Operating Characteristic (ROC) curve shows the accuracy performance at various threshold settings
- The other faults show reasonable separation with AUC values ranging from 0.83 to 0.93 (they can be found in the paper*)

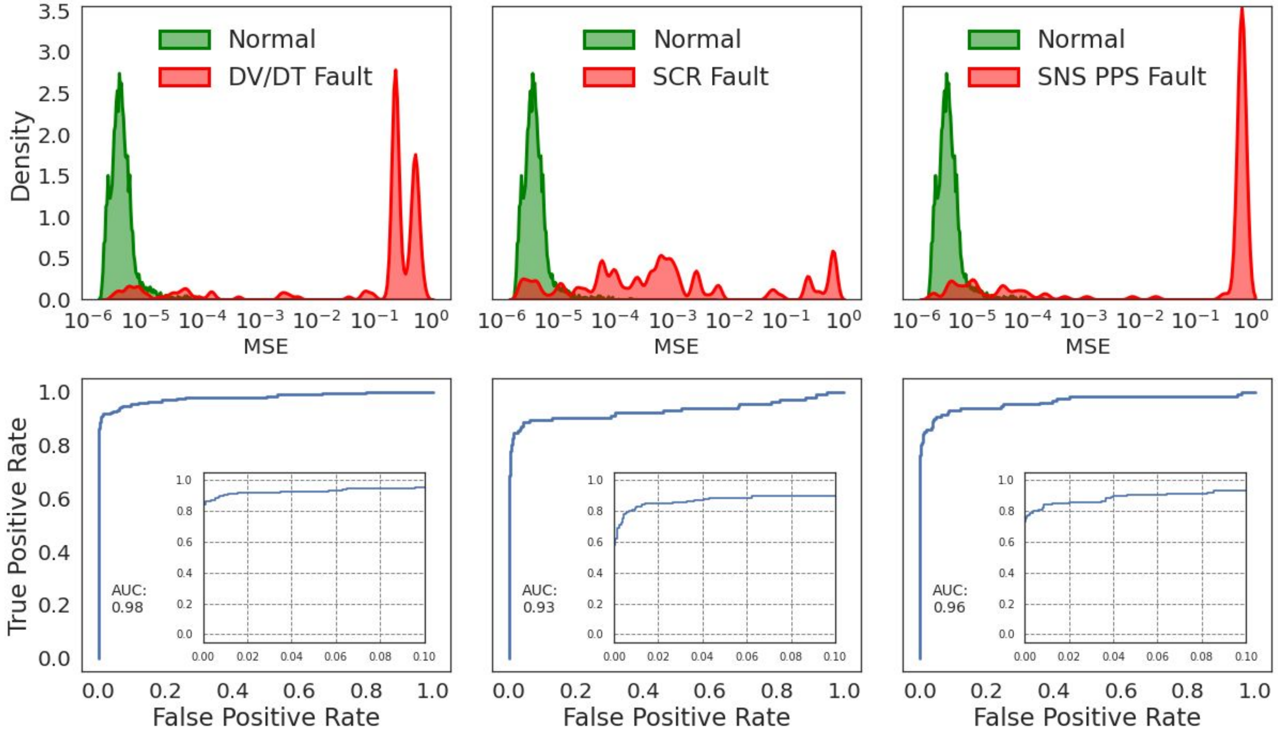


Figure 13: KDE distributions of the MSE from reconstructing normal (grey color) and faulty waveforms (blue color) for six fault types, with the corresponding ROC curve for each fault.

*https://arxiv.org/pdf/2304.10639.pdf

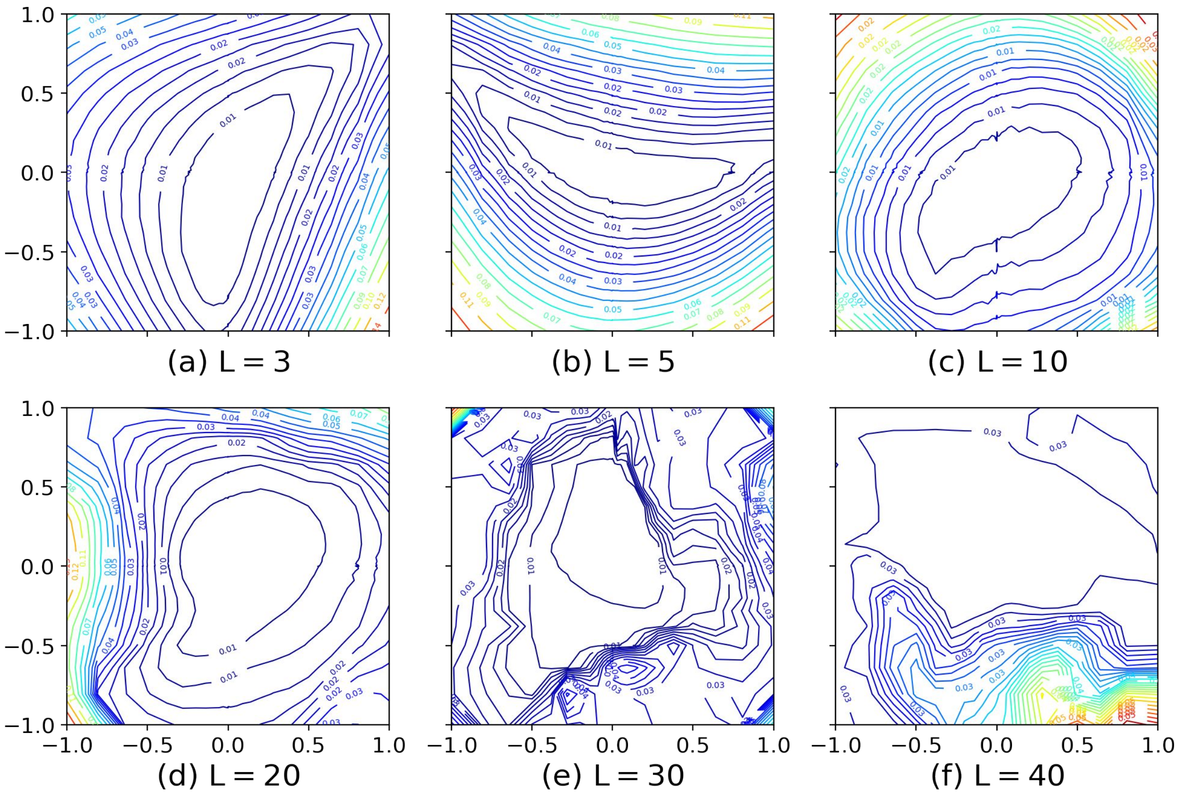
OAK RIDGE National Laboratory | SPALLATION NEUTRON SOURCE

Jefferson Lab

4.Loss Landscape

- Train 6 VAE models using different number of CNN layers (L):
    - L = 3, 5, 10, 20, 30, and 40.

- Transition from smooth loss surface to chaotic behavior as the number of layers increased

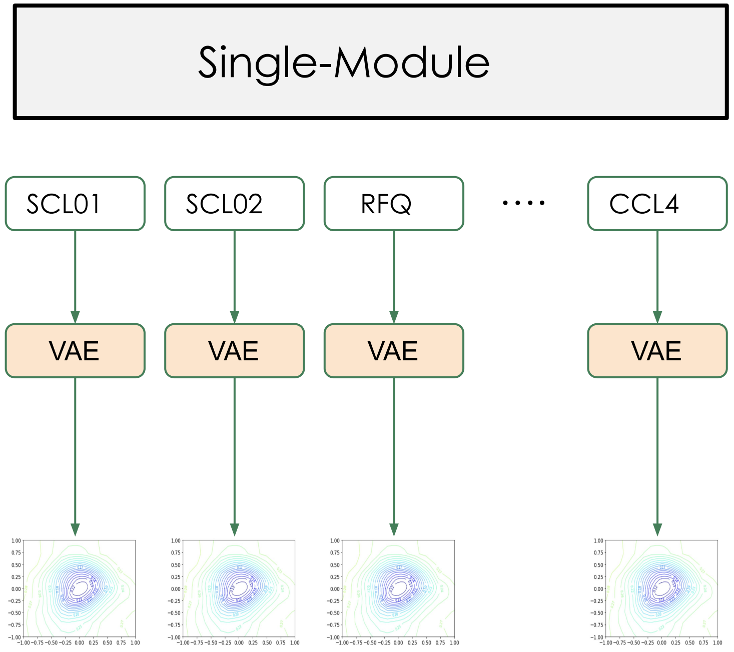


2D visualization of the loss surface of the Single Module-based trained using different number of Conv1D layers, where L is the number of layers in the encoder and decoder.

4. Loss Landscape

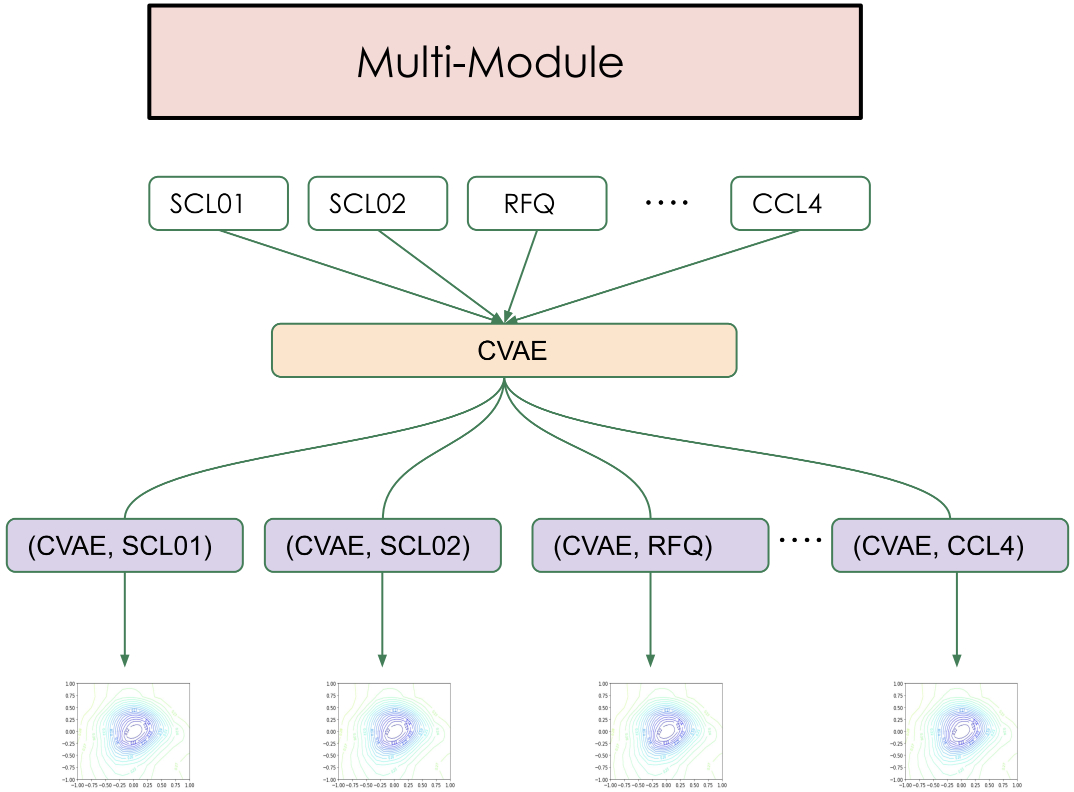
- <u>Single-Module</u>: Train an individual Variational Autoencoder (VAE) for each subsystem (e.g. SCL01).

- Visualize the loss for each model.



- <u>Multi-Module</u>: Train a conditional Variational Autoencoder (CVAE) combining all subsystems together.

- Visualize the loss of the model for each file.



CVAE Model Calibration

- The majority of the Miscalibration Area is less than 5%

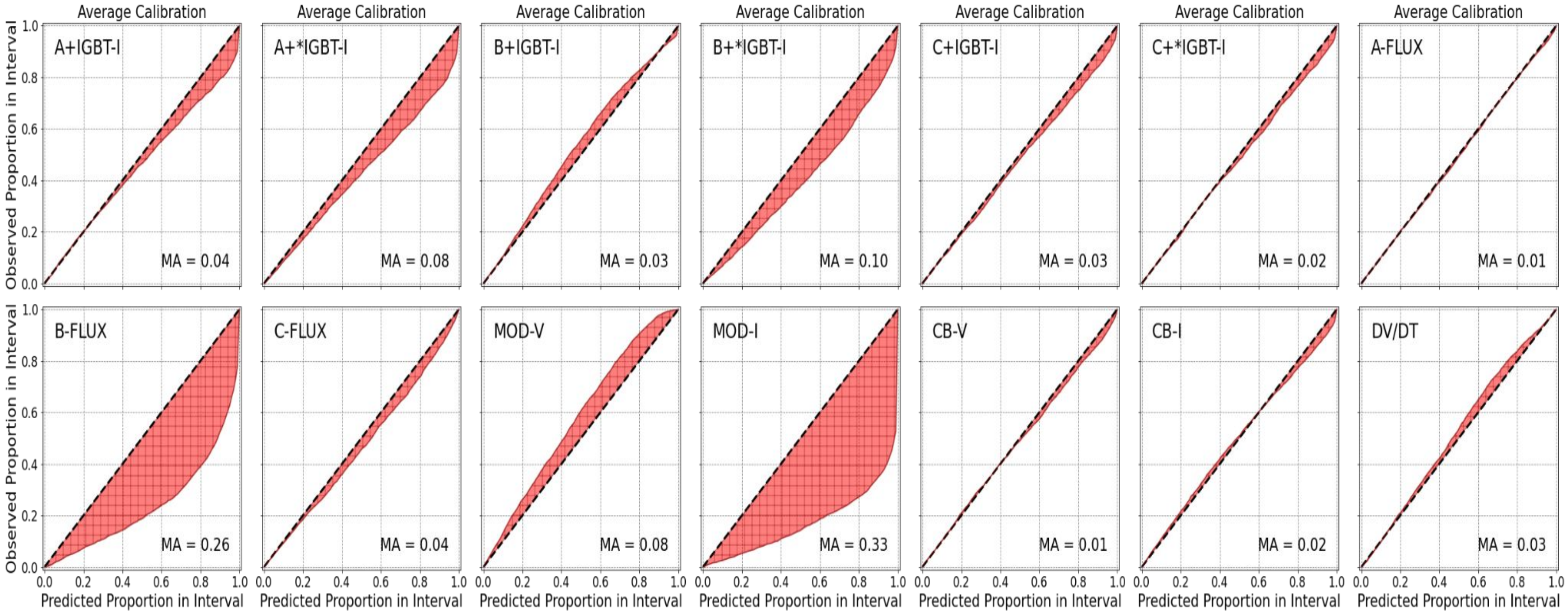


Figure2: SCL01 Miscalibration Area (MA) for all waveforms