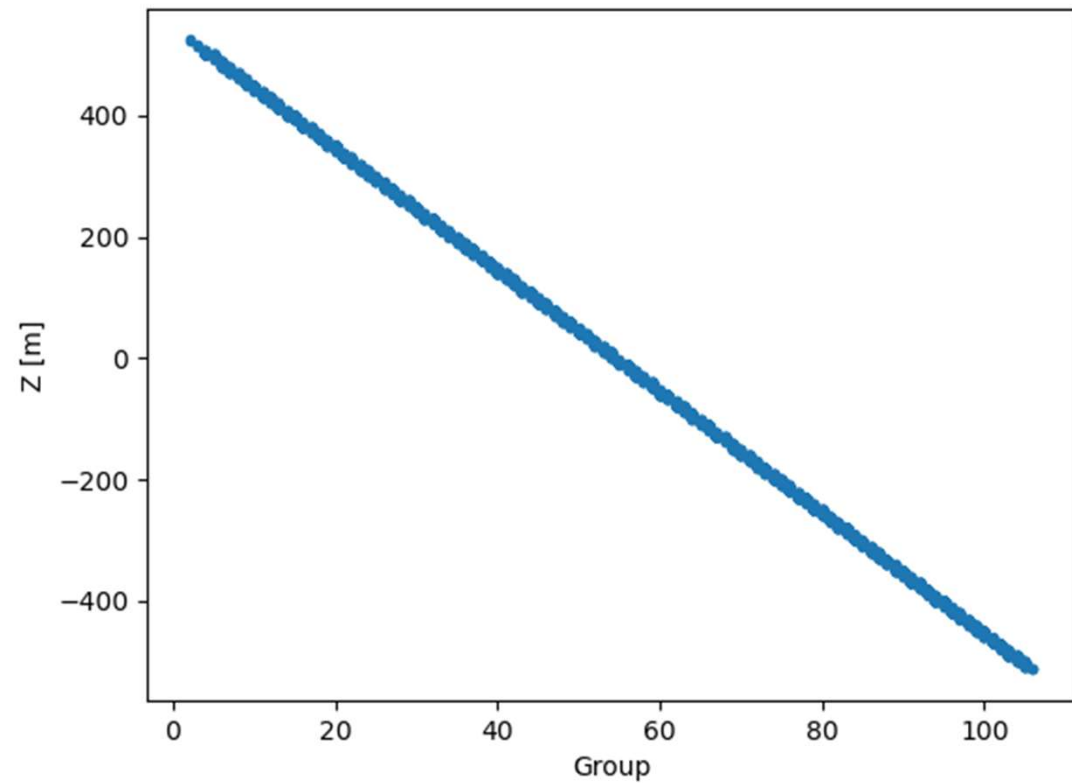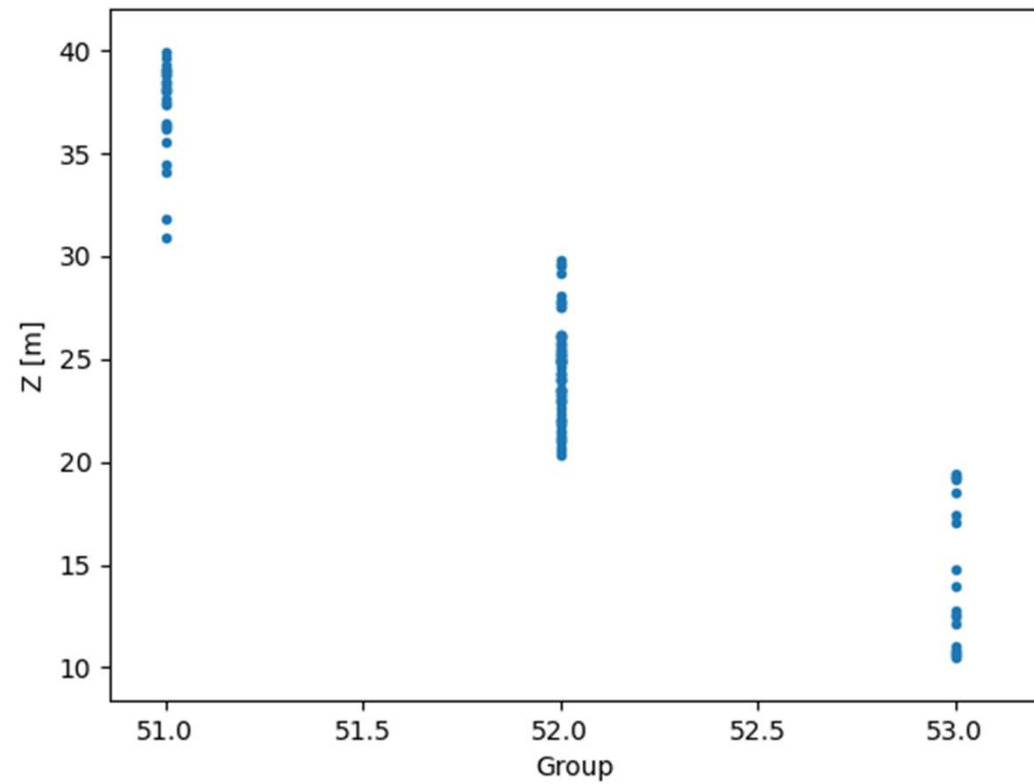# RIDE Update

Sofus Stray

Recap

- We expect DOM charge responses to be similar in the same depth level

- Group DOMs into 106 groups based on z-position

Recap

- We expect DOM charge responses to be similar in the same depth level

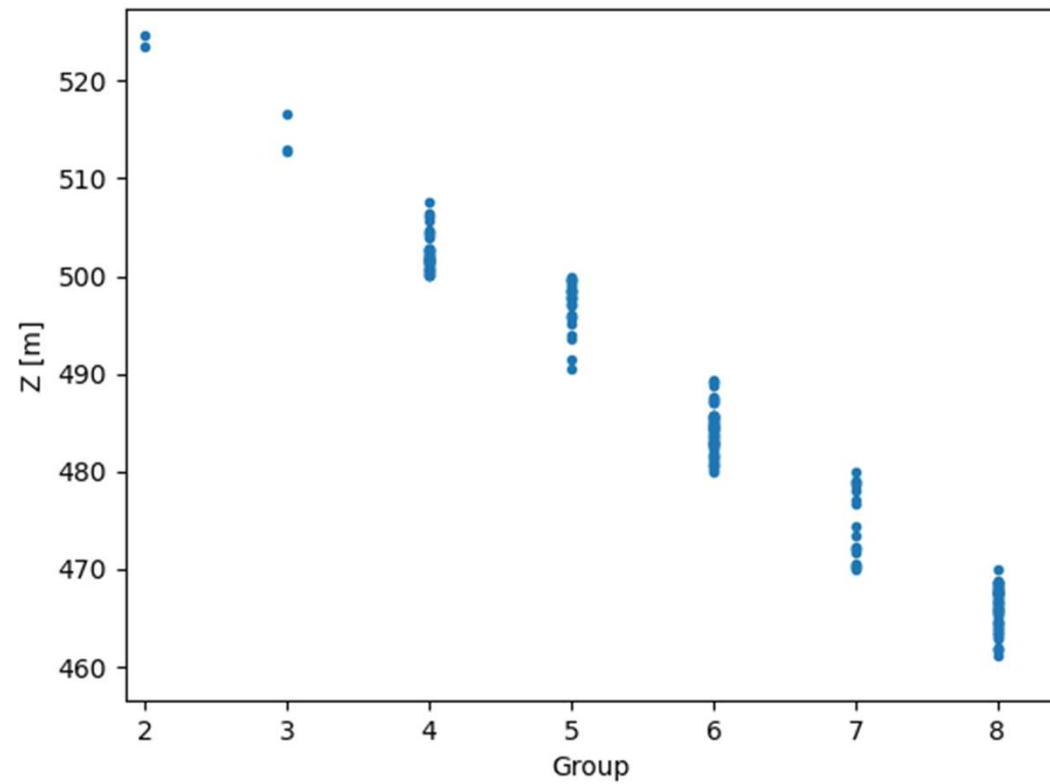- Group DOMs into 106 groups based on z-position

Recap

- We expect DOM charge responses to be similar in the same depth level

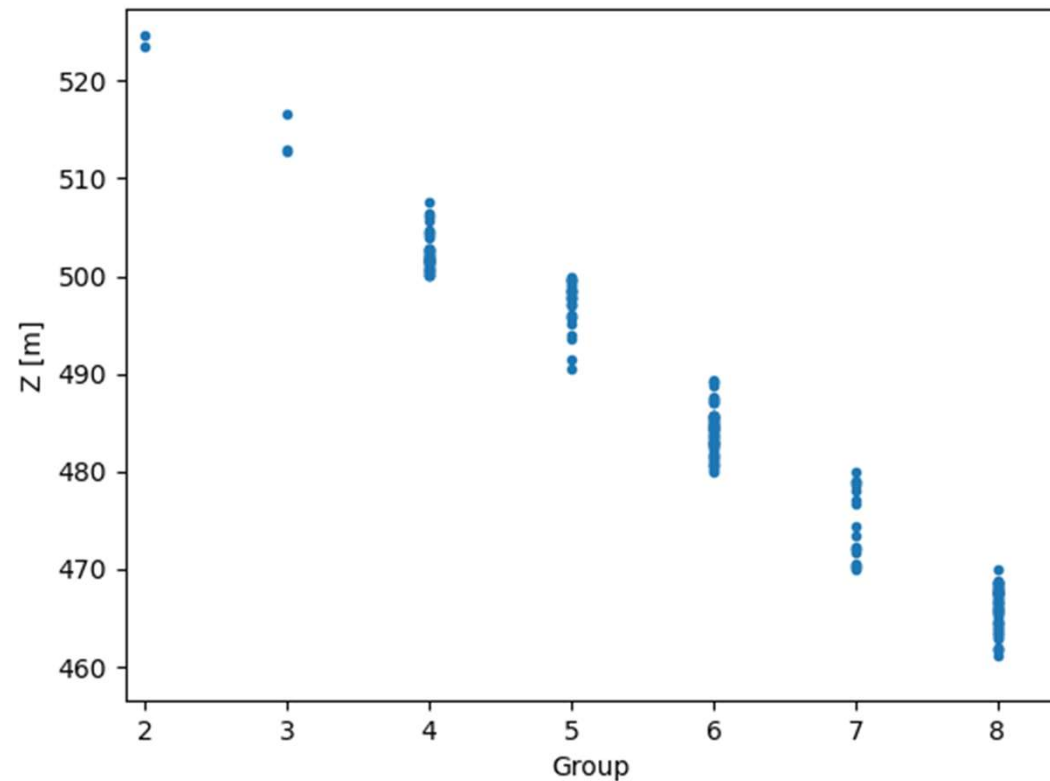- Group DOMs into 106 groups based on z-position

Recap

- We expect DOM charge responses to be similar in the same depth level

- Group DOMs into 106 groups based on z-position

- Calculate mean charge of each DOM

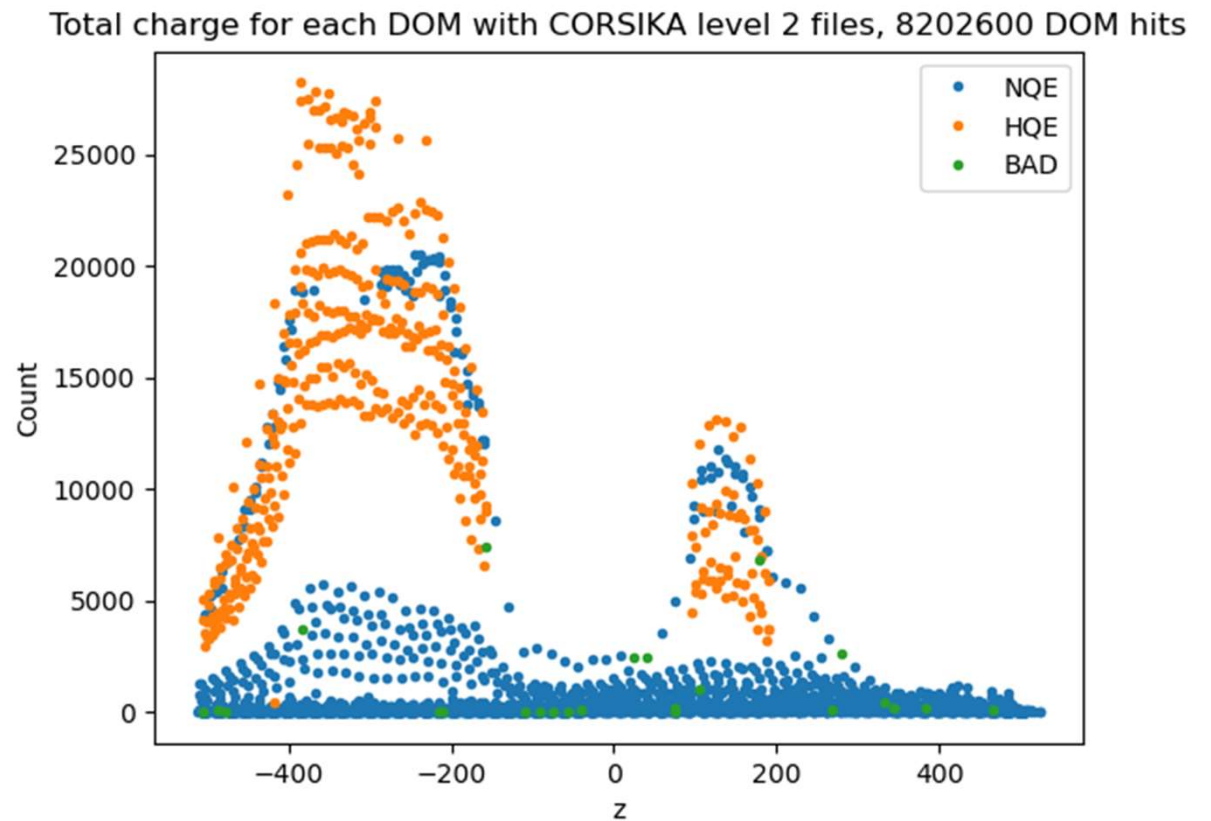- Divide each charge by the group's median

- $RIDE_i = \dfrac{\left(\frac{\Sigma_{events}\, q}{\Sigma_{eve}\quad hit}\right)_i}{\left(\frac{\Sigma_{event}\quad q}{\Sigma_{even}\quad hit}\right)_{monitor}}$

- **Expectations: NQE DOMs have ride value of 1, HQE of 1.35**

Recent results

- Ran through 10,000 CORSIKA L2 simulation files

- Used TCN (neural network approach) to predict stopped muons

- Calculate total charge and RIDE-value for each DOM

- Only select DOMs within 75m of muon track

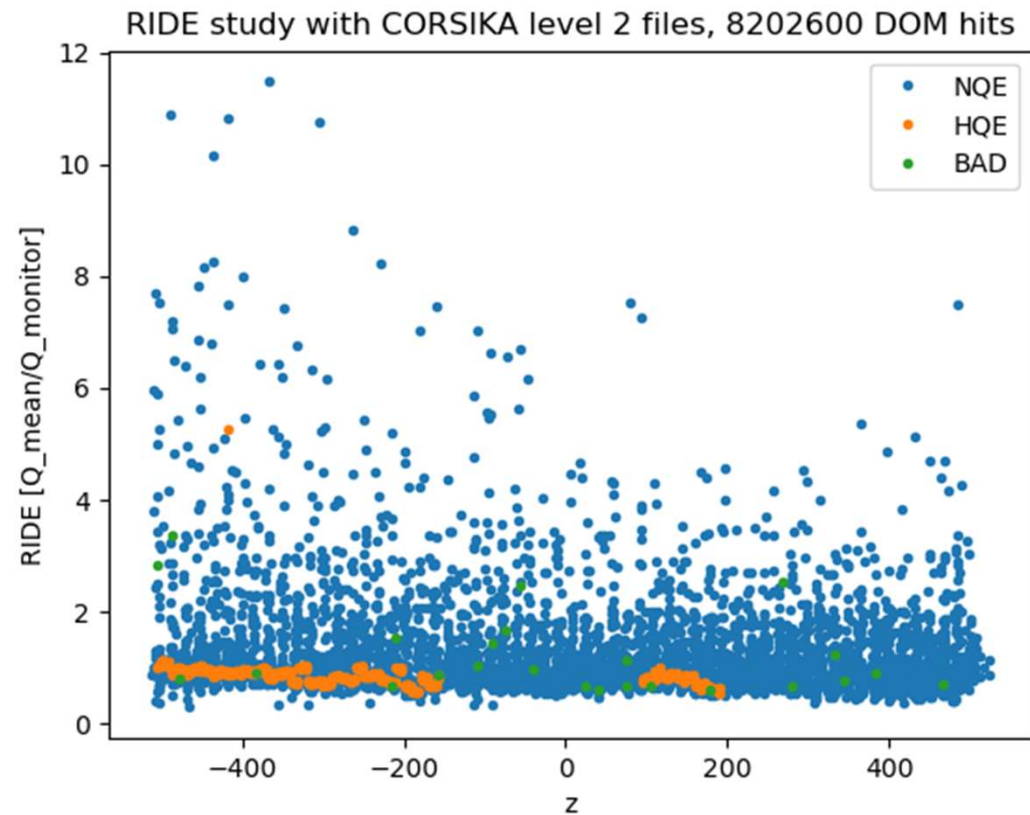Total charge for each DOM with CORSIKA level 2 files, 8202600 DOM hits

Recent results

- Ran through 10,000 CORSIKA L2 simulation files

- Used TCN (neural network approach) to predict stopped muons

- Calculate total charge and RIDE-value for each DOM

- Only select DOMs within 75m of last 200 meters of muon track

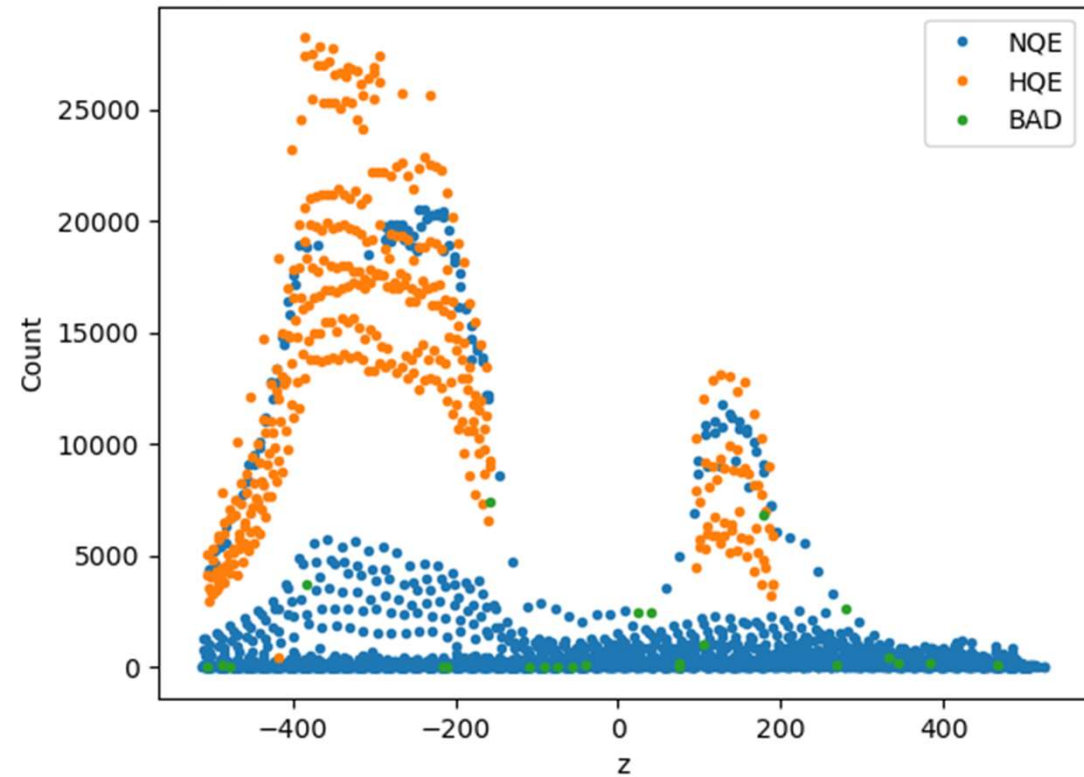RIDE study with CORSIKA level 2 files, 8202600 DOM hits

Problems

Problem 1: Some NQE DOMs have a much higher total charge than expected
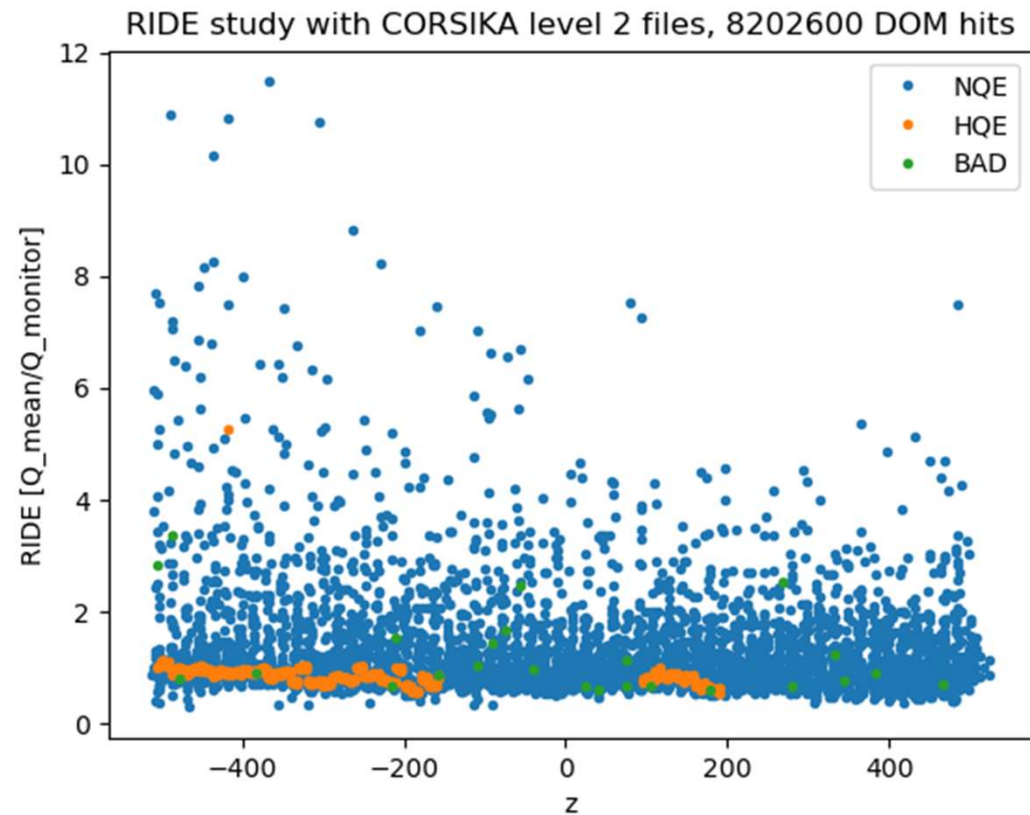
- More on that later

Total charge for each DOM with CORSIKA level 2 files, 8202600 DOM hits

Problems

Problem 2: NQE DOMs have a higher RIDE value despite HQE DOMs generally having more total charge

- Reason: Mean charge is calculated from every DOM *with* a charge response instead of every could that could have a charge response

Problems

Problem 2: NQE DOMs have a higher RIDE value despite HQE DOMs generally having more total charge

- Reason: Mean charge is calculated from every DOM *with* a charge response instead of every could that could have a charge response

- Solution: Calculate mean charge from every DOM within radius of muon track
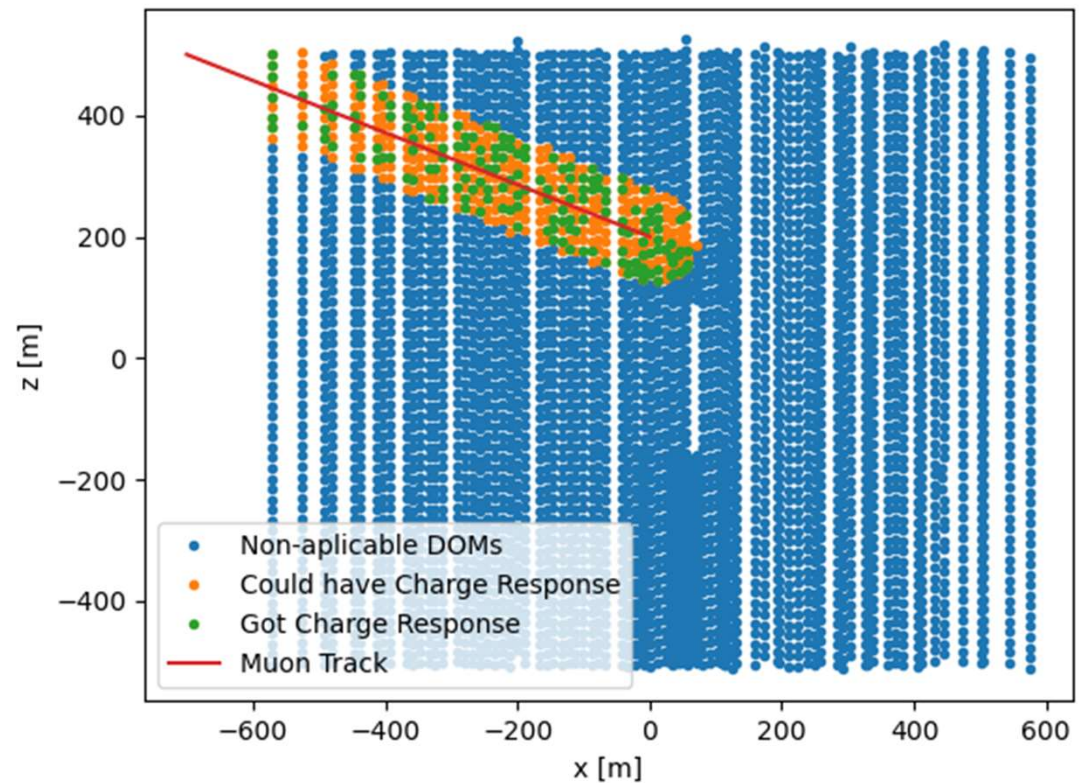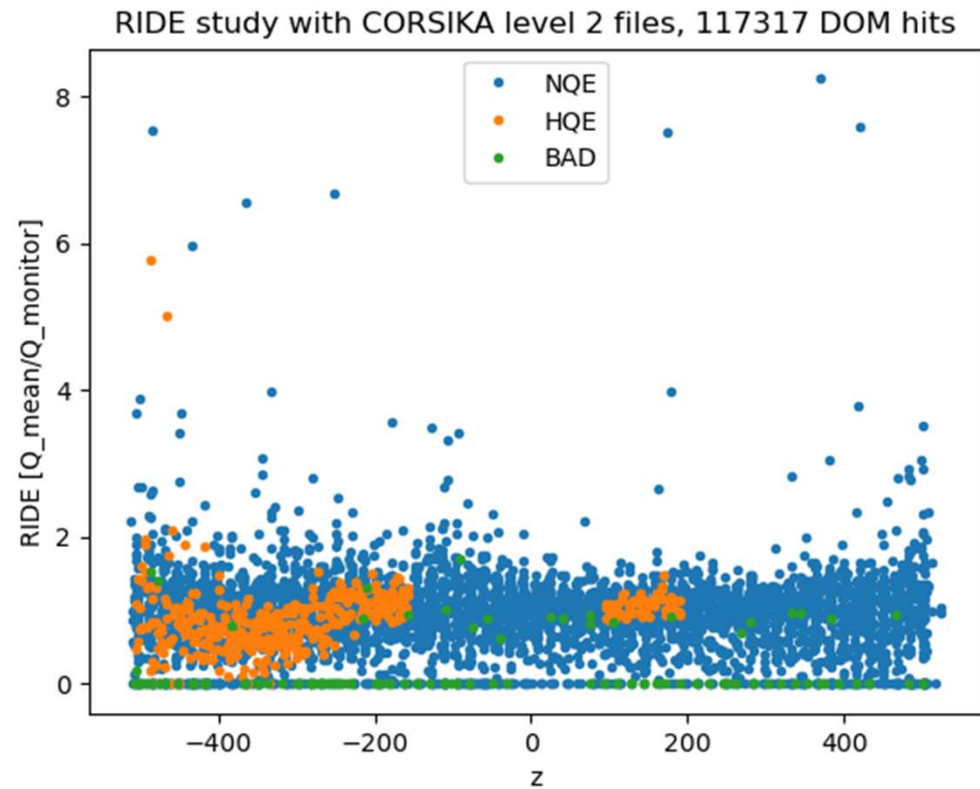
New Results

- Looks only at last 200 meters

- Calculates mean charge for every DOM within track radius

- Less statistics due to ongoing bug

- Completely different shape
  - Almost a reversal of previous shape
  - Highly unlikely to be due to statistics

Total charge for each DOM with CORSIKA level 2 files, 117317 DOM hits
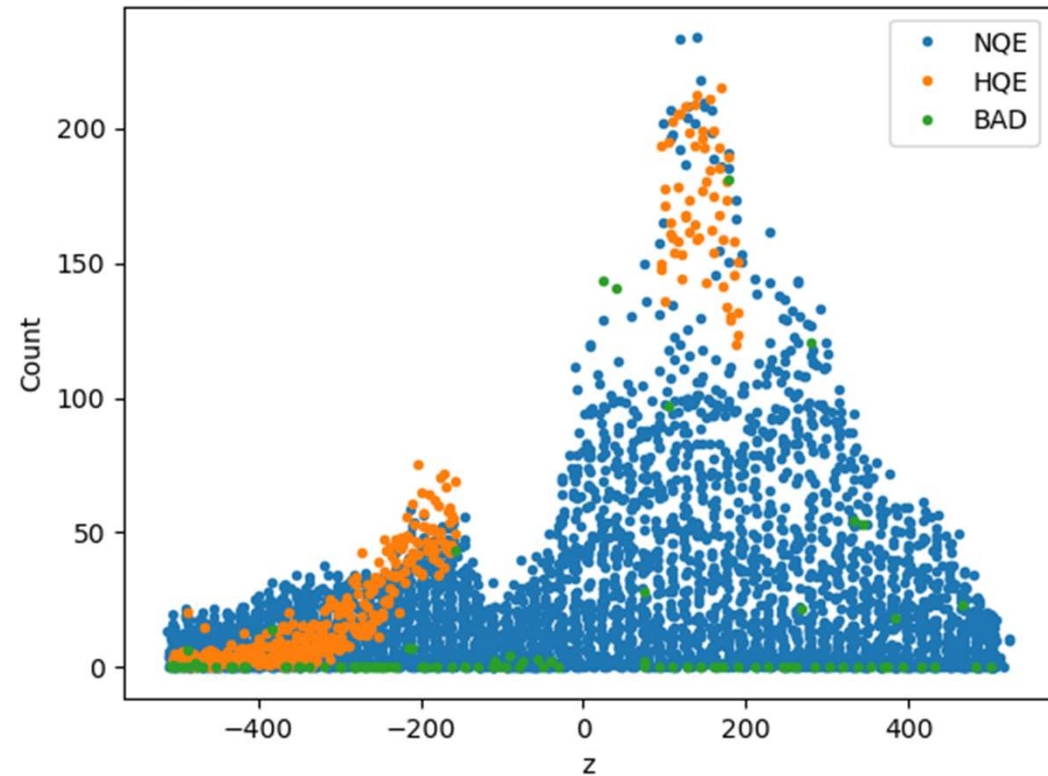
New Results

- Looks only at last 200 meters

- Calculates mean charge for every DOM within track radius

- Less statistics due to ongoing bug

- More noisy HQE RIDE
  - Almost certainly due to low statistics

- "Right side" of HQE cluster generally stable

- Less NQE fraction with higher RIDE than HQE clusters

RIDE study with CORSIKA level 2 files, 117317 DOM hits

Quick comparison

Old results

Total charge for each DOM with CORSIKA level 2 files, 8202600 DOM hits

RIDE study with CORSIKA level 2 files, 8202600 DOM hits

New Results

Total charge for each DOM with CORSIKA level 2 files, 117317 DOM hits
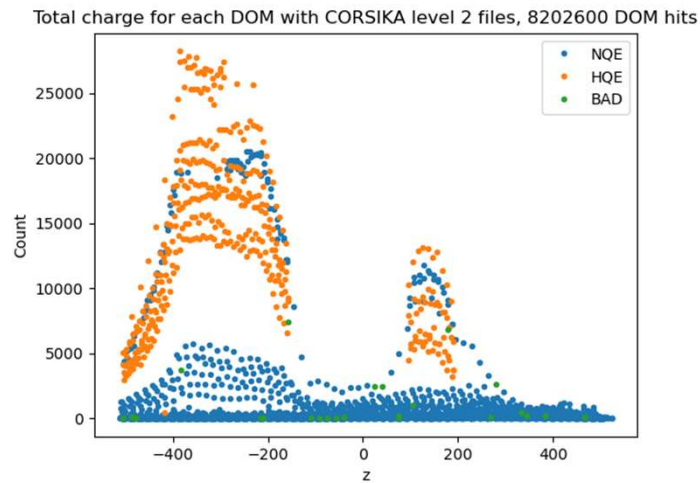
RIDE study with CORSIKA level 2 files, 117317 DOM hits
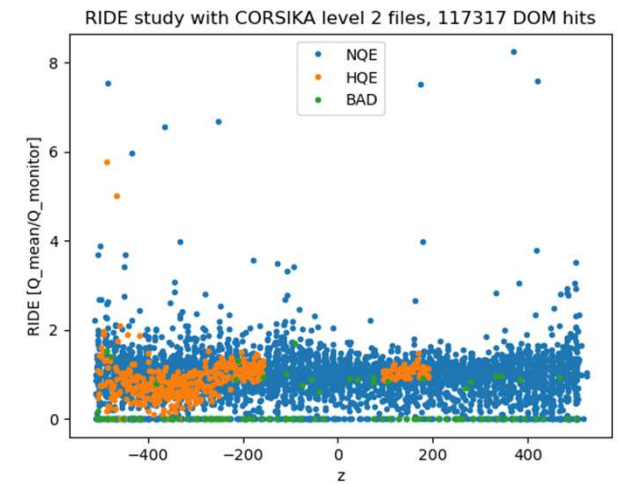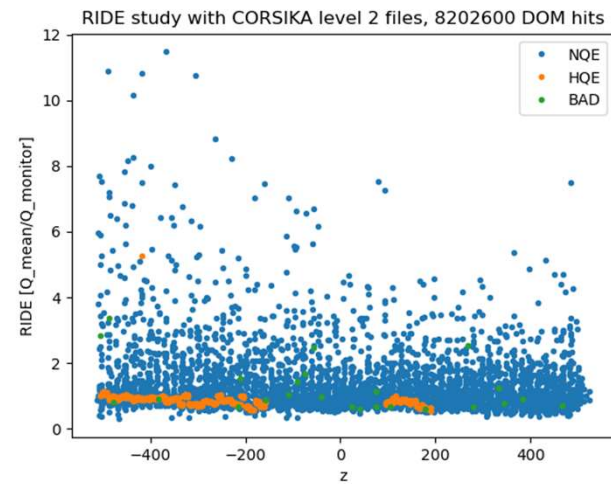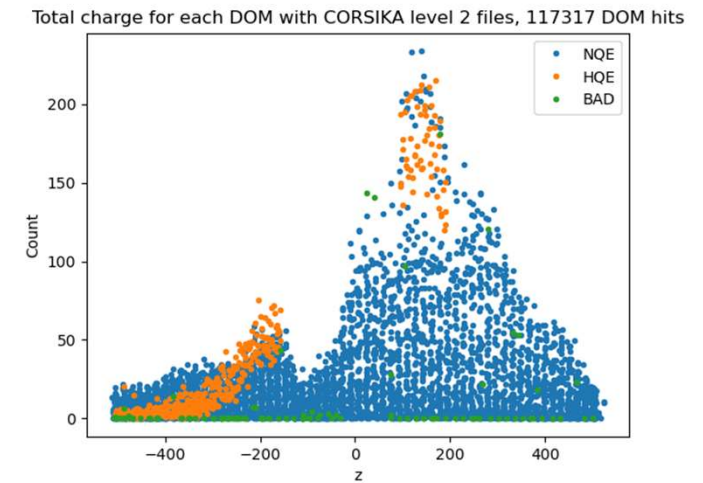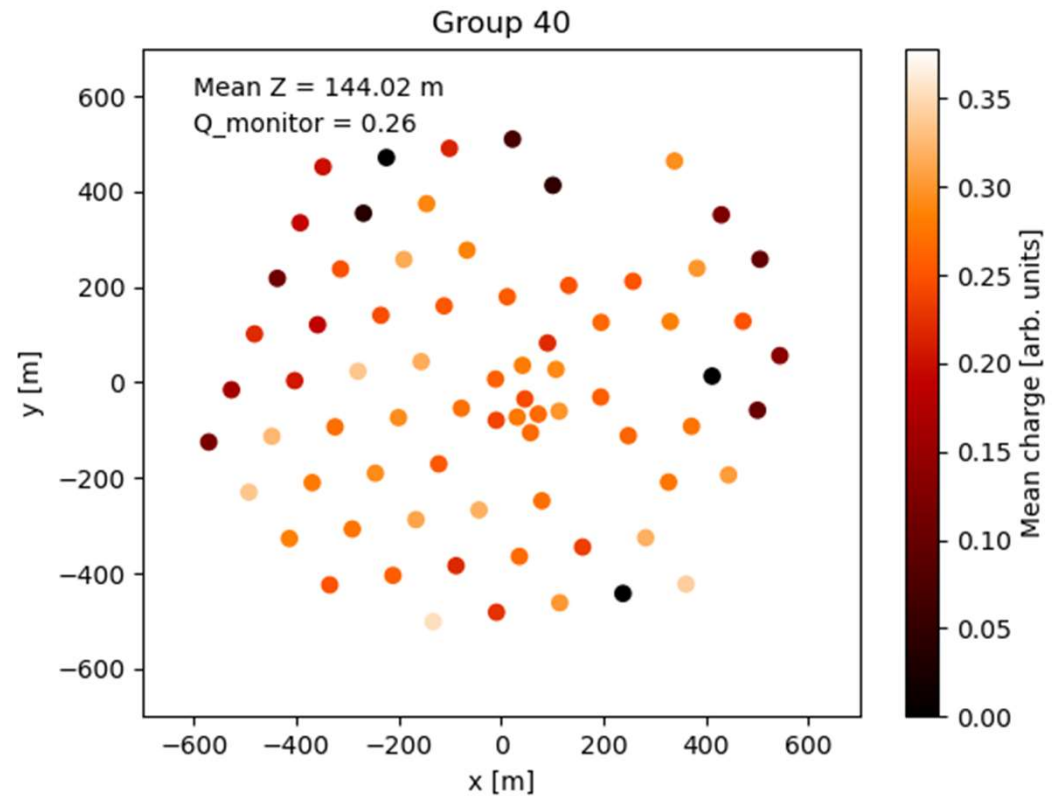
Discussion

- Possible reasons
  - Bugs in the code
  - Wrong implementation of new mean
  - TCN prediction issues
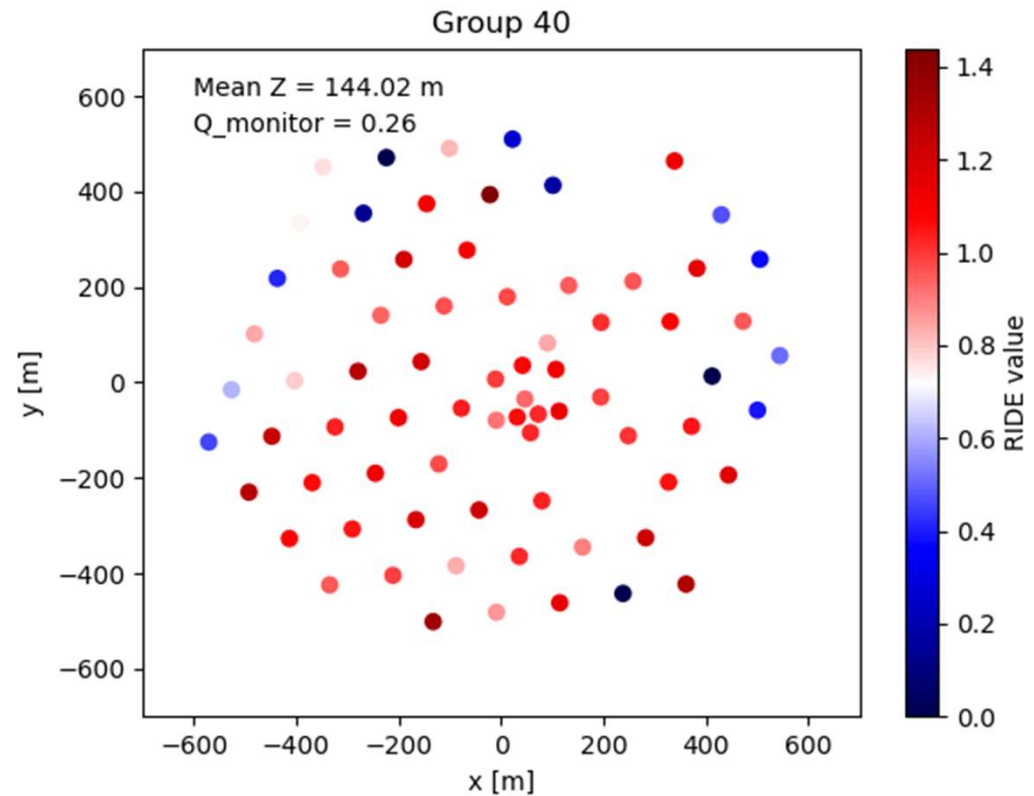- Quick detour before going further

Analysis of single group

- Group 40 picked arbitrarily (just needs to be a group with many DOMs)

- Fairly unstable results
  - Likely from lack of statistics
 - Possibly for reasons discussed later

Analysis of single group

- Group 40 picked arbitrarily (just needs to be a group with many DOMs)

- We should expect all DOMs to have a RIDE value of 1

Analysis of single group

- Histogram of group 40, showing monitor

- For all DOMs to have a RIDE value around the monitor, we would expect a much tighter distribution

- Some values, like the mean charge of 0, is entirely explained by lack of statistics
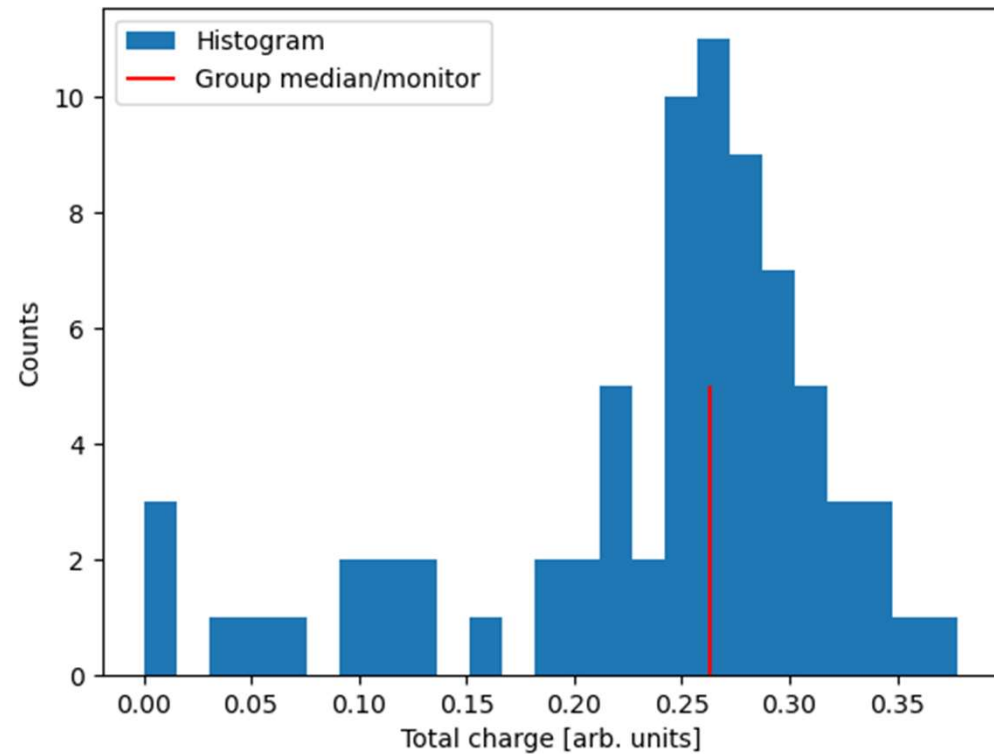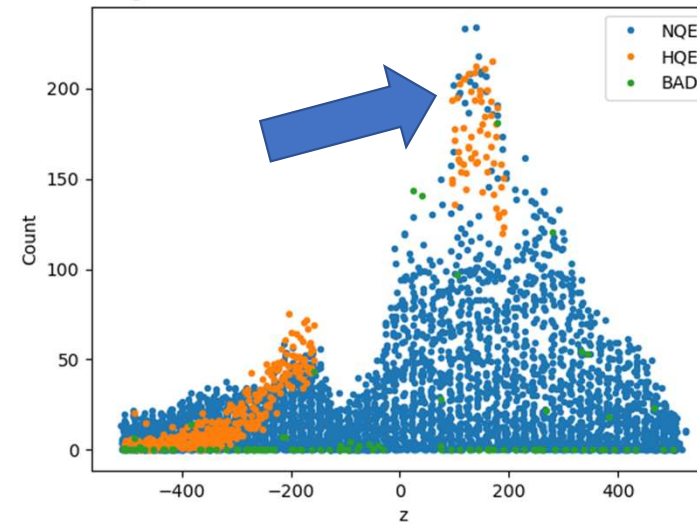
Unexpected high HQE charge analysis

- Figure out what causes the HQE anomalies in both results

- Are abnormal HQEs from the same string?

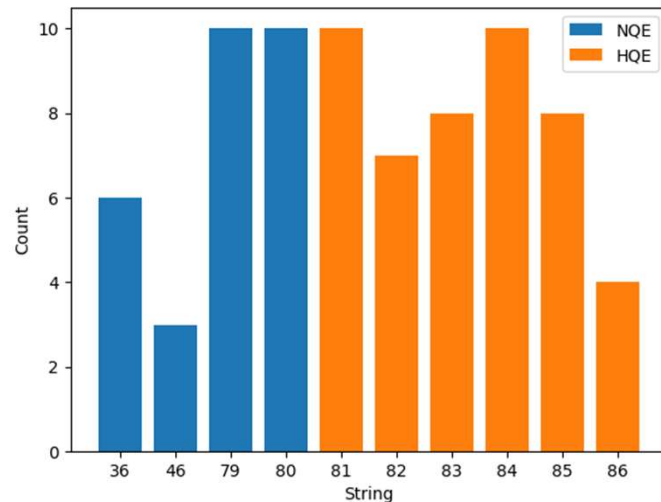Total charge for each DOM with CORSIKA level 2 files, 8202600 DOM hits
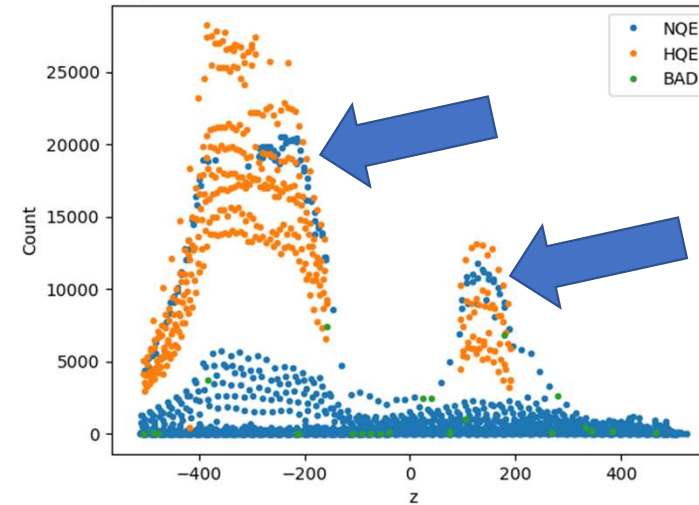
Total charge for each DOM with CORSIKA level 2 files, 117317 DOM hits
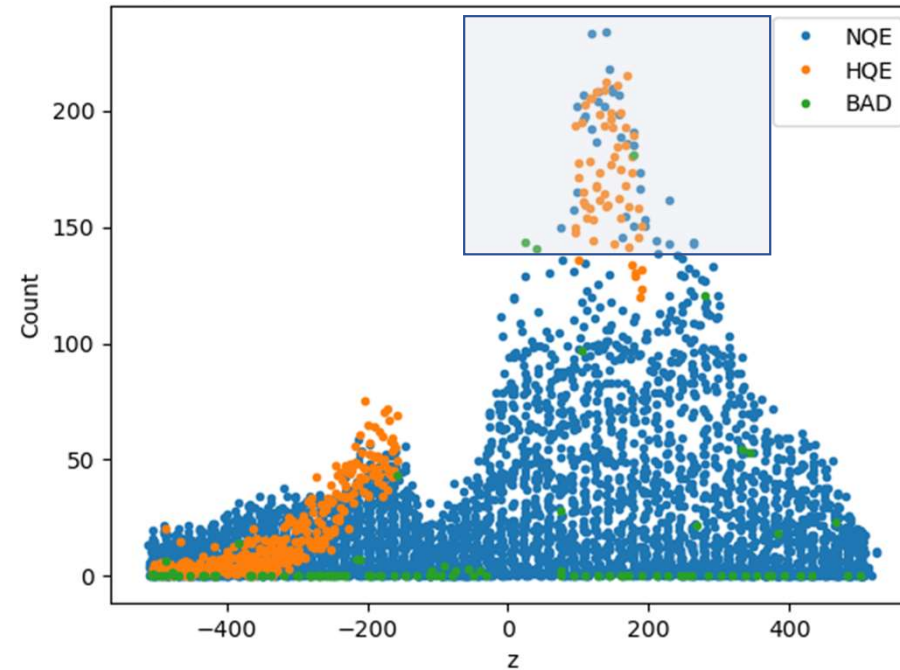
Unexpected high HQE charge analysis

- Figure out what causes the HQE anomalies in both results

- Are abnormal HQEs from the same string?

- Possibly: Analysis of DOMs with >150 total charge come from following strings:

Total charge for each DOM with CORSIKA level 2 files, 117317 DOM hits

Take note of low statistics: These strings *might* not be responsible with more data
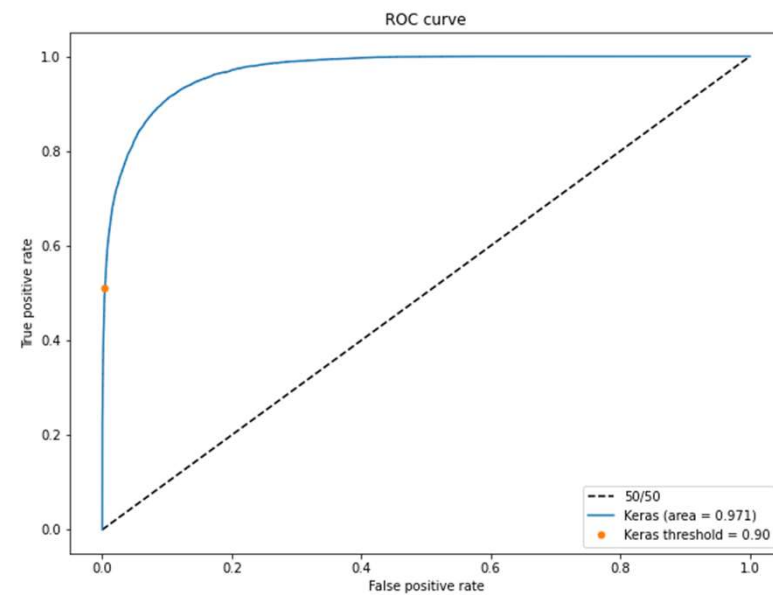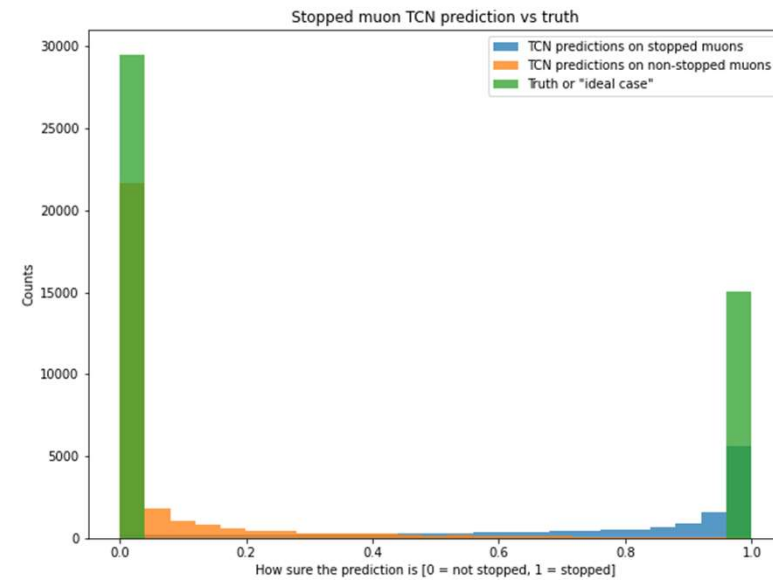
Back to old/new data discrepancy

- Possible reasons
  - Bugs in the code
  - Wrong implementation of new mean
  - TCN prediction issues

Back to old/new data discrepancy

- Possible reasons
  - Bugs in the code
  - Wrong implementation of new mean
 - TCN prediction issues

TCN Recap

- Neural Network

- Predicts whether muon is stopped or not

- Trained on muon gun data

- Performs well on said data

Test performance on CORSIKA data

- Should've been done earlier

- Completely terrible performance

- Essentially 50/50

- Possible reasons:
  - **Bug in definition of Truth**
  - **Features don't share same distribution**
  - Event ID conflicts
  - Muongun data can't predict CORSIKA data (would be very weird)
  - Possible overtraining (unlikely)

Stopped muon TCN prediction vs truth

ROC curve
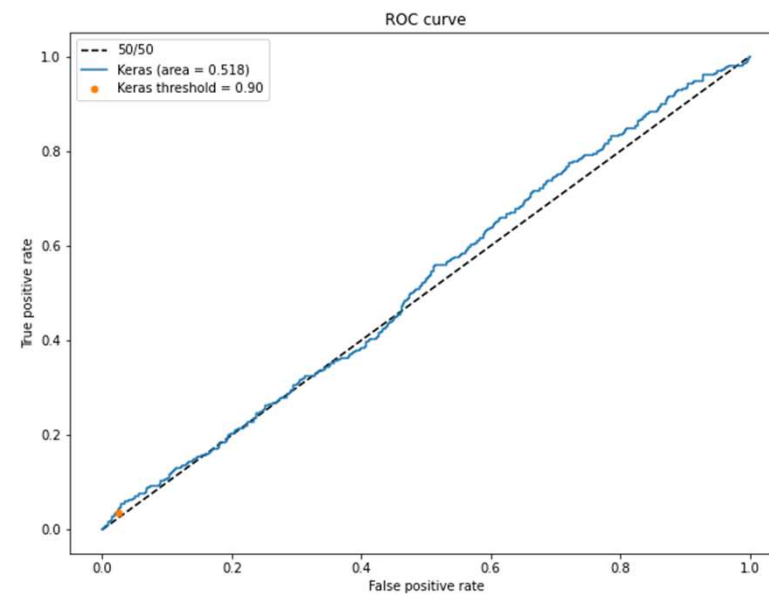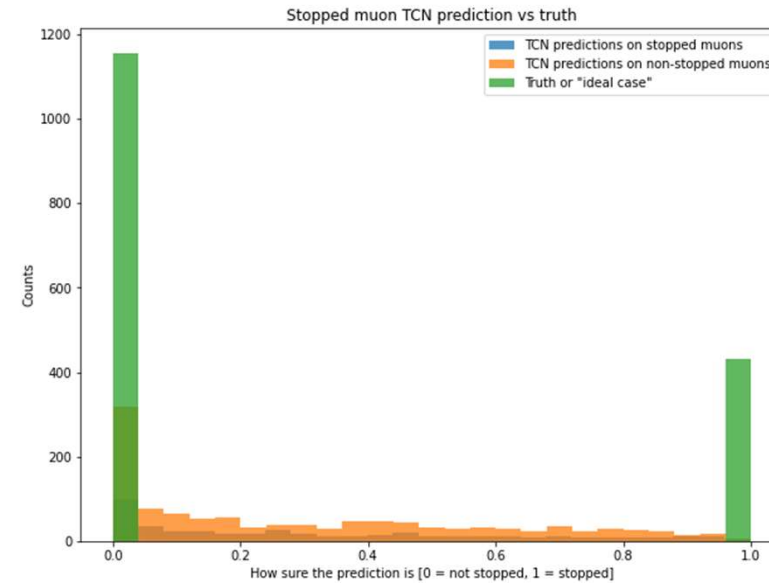
Bugs in Truth definition

- Most likely reason

- Similar culprit (though not a bug) could be a much lesser % of stopped muons in CORSIKA data
  - This doesn't explain why the separation is so unclean

Stopped muon TCN prediction vs truth

CORSIKA

Stopped muon TCN prediction vs truth
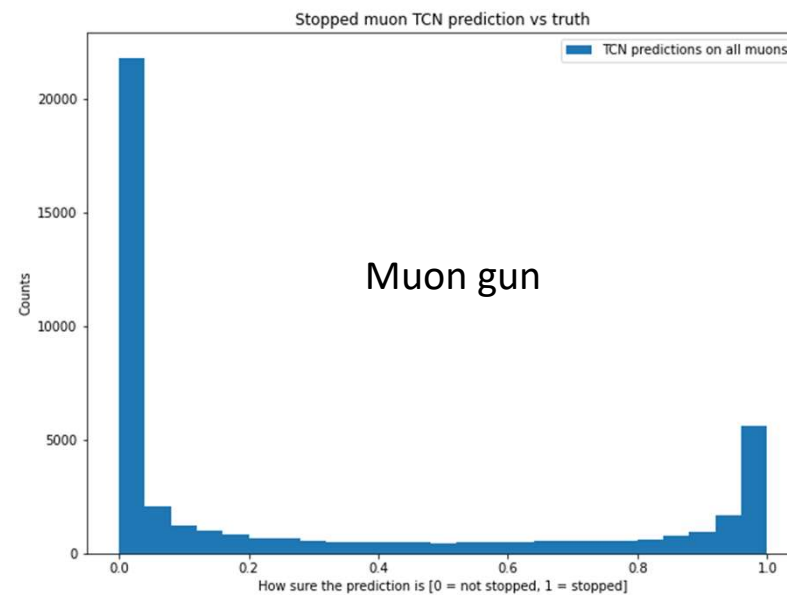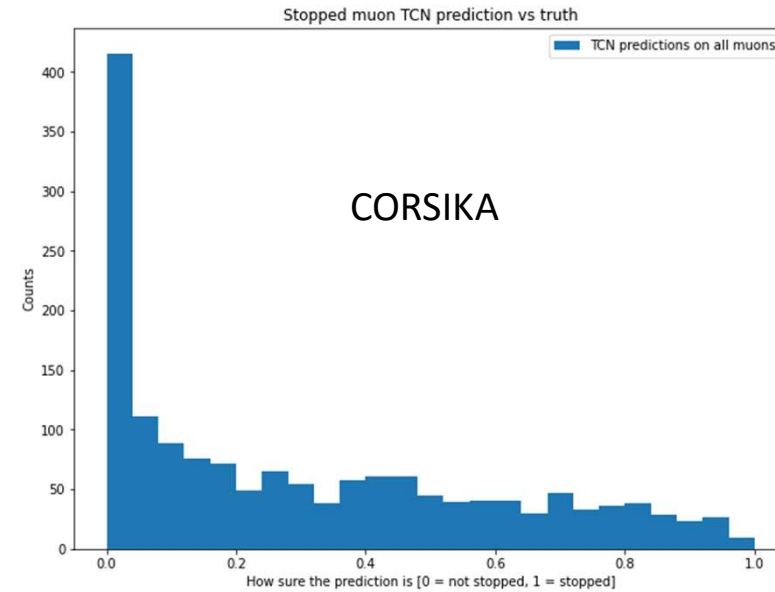
Muon gun

Features don't share same distribution

- x and y strongly share distributions

- CORSIKA x/y slightly more strongly centered around 0

- Overrepresentation of DOM hits in lower/higher depth levels for CORSIKA/muon gun respectively
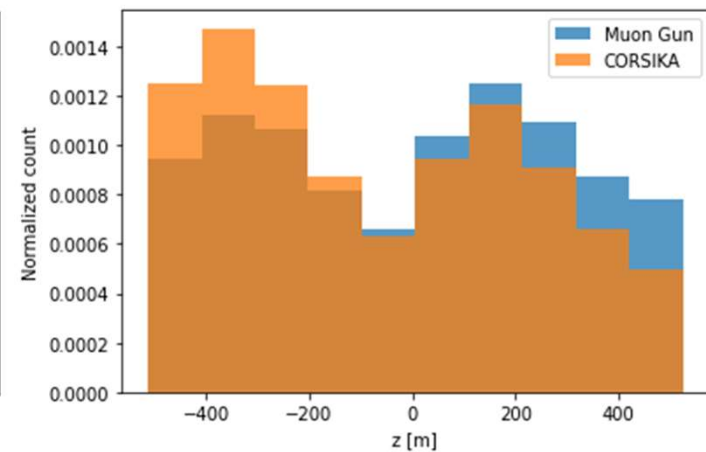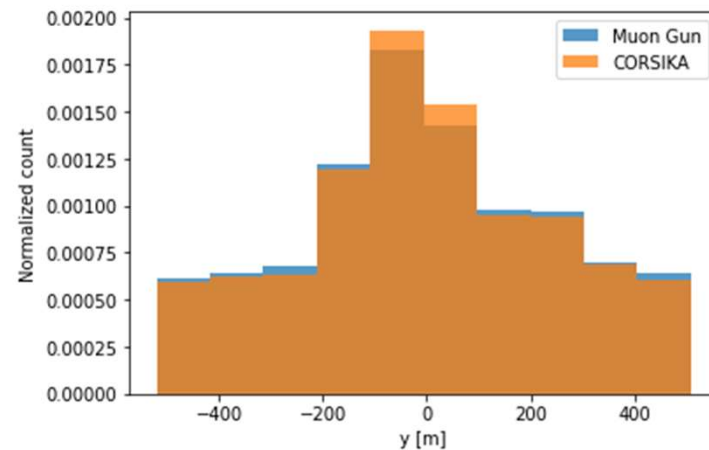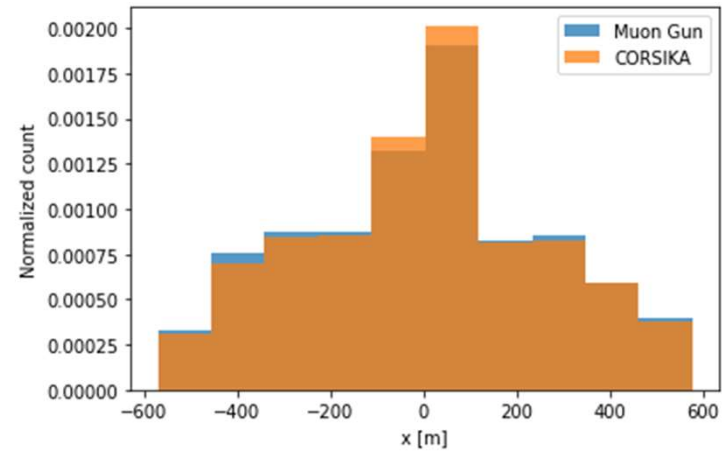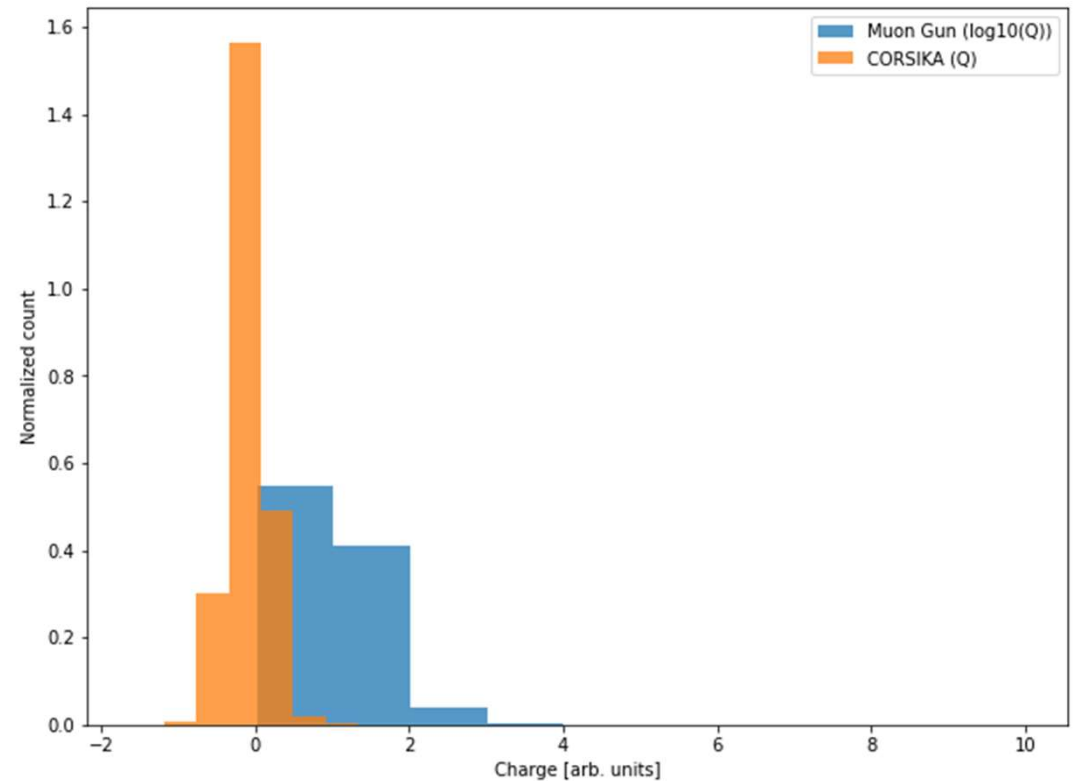
Features don't share same distribution

- [25 bins instead of 10 for a bit more detail]

- Time and charge has significant non-overlapping

- No strong structures appearing

- Muon gun has a much larger time tail
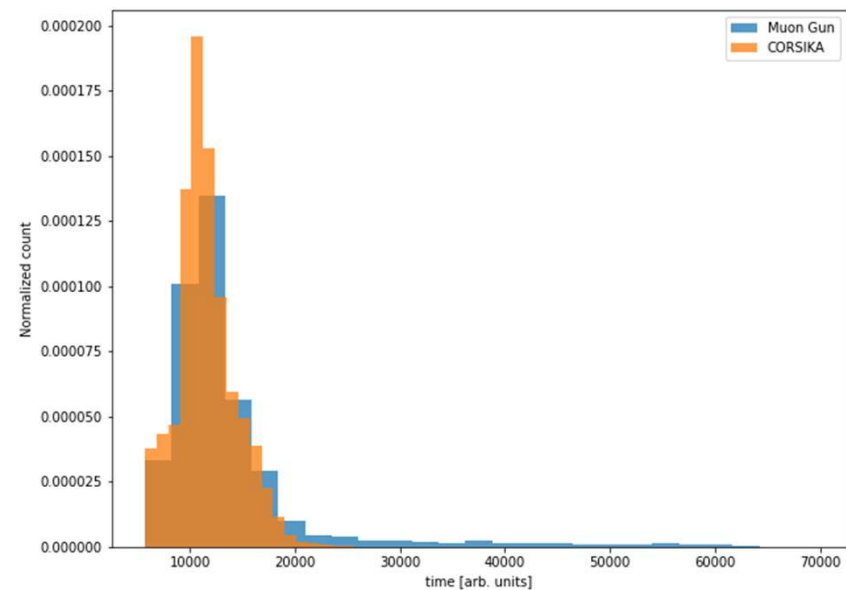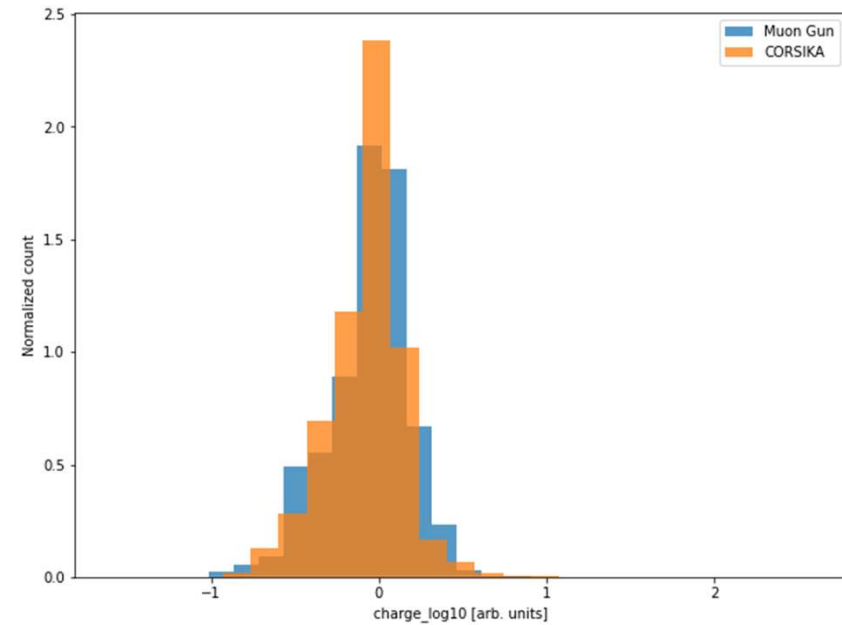
Features don't share same distribution

- Bug was also found: Muon gun data uses log10 of charge, CORSIKA just uses charge

- Performance test was with bug fixed

- RIDE calculation did not have the bug fixed

- Prediction is still worthless post-bug-fix

- Does not by itself explain the issue

Features don't share same distribution

- Uneven distributions may be rectified by better statistics
- Overall distributions alone don't account for distribution in each event
  - Event ID bugs could still be the culprit

Going forward

- Scour trough code and fix any and all bugs

- Improve statistics of current data

- Look through the TCN code

- Analyse more modern CORSIKA files (current is 2012)

- After above is fixed:
  - Get true stopping variables
  - Run on actual data
  - Compare performance to FiniteReco + MPE