Updated Vertex Selection BDT for Atmospheric Neutrinos

Andy Chappell

14/02/2022 FD Sim/Reco Meeting

DEEP UNDERGROUND NEUTRINO EXPERIMENT

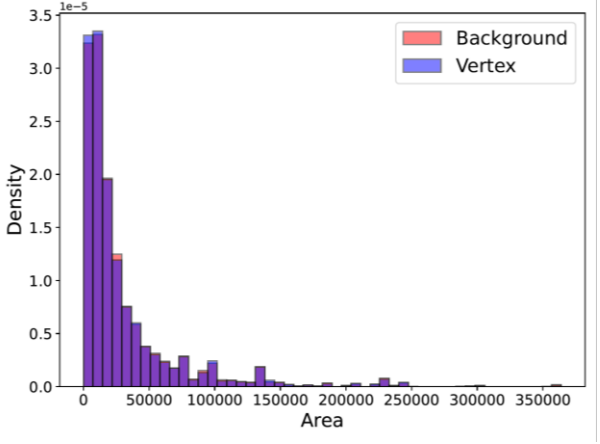
WARWICK

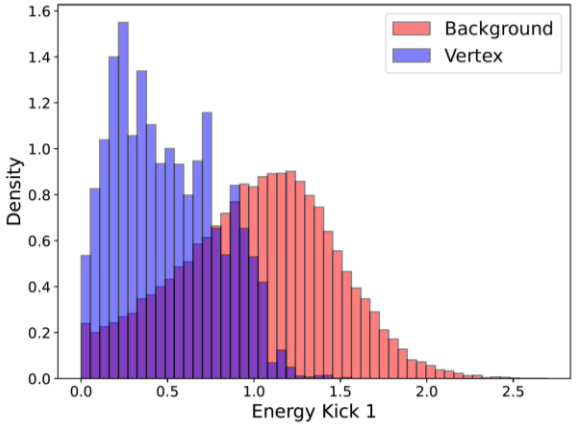
Beam assumption

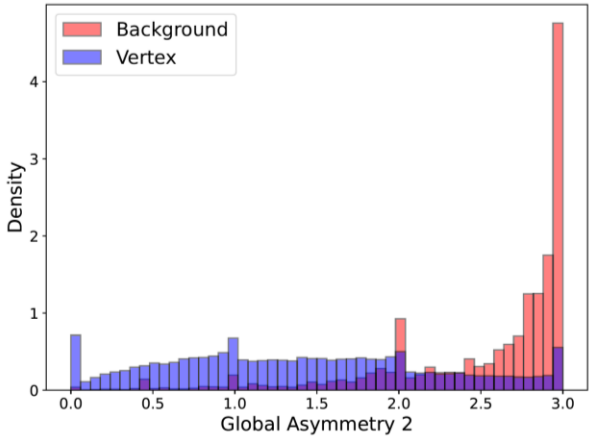
WARWICK

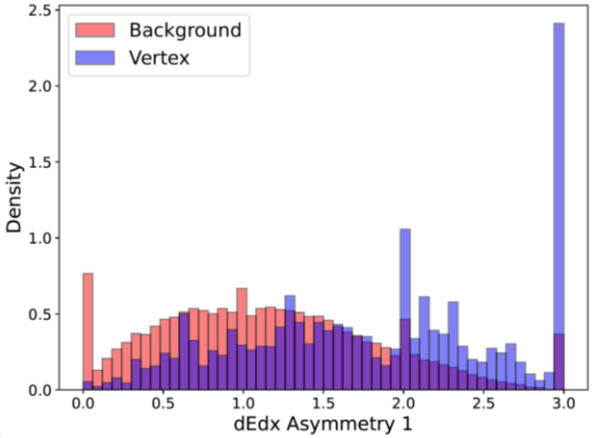
- v03_26_03 and earlier of LArContent assumes a beam
  - Negatively impacts atmospheric reconstruction for “backward going” interactions
- Updated LArContent to allow beam deweight to be optional
  - For technical reasons, you can’t just drop the variable, but update avoids failure when running in non-beam mode
- Running over atmospherics sample NNBarAtm_hA_BR_dune10kt_1x2x6_RITM1251140_reco
- Region and Vertex BDTs have been retrained

Example Features







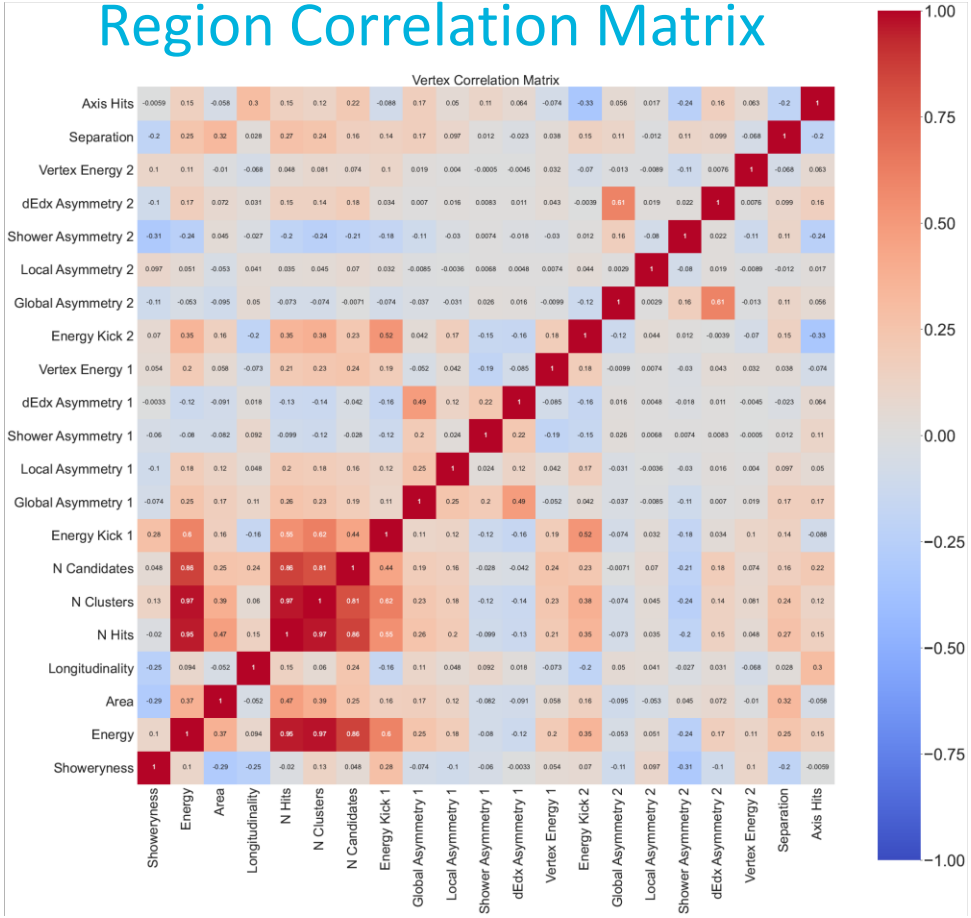


Region finding BDT

WARWICK

- Pandora’s vertex selection BDT takes a two-tiered approach
- Vertices are selected based on their scores in coarse-grained regional comparisons and by fine-grained comparisons within local regions
- We’ll look first at the region finding BDT and then the fine tuning BDT

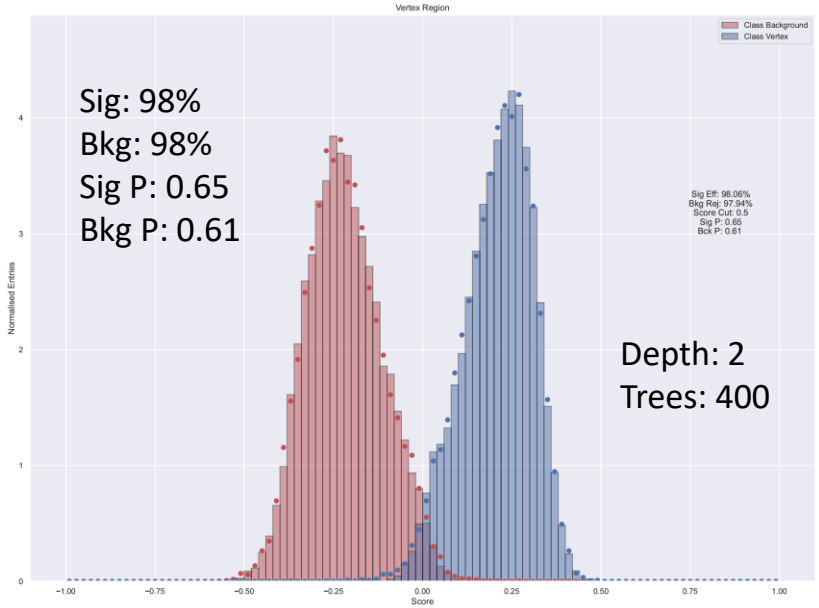
Region Correlation Matrix



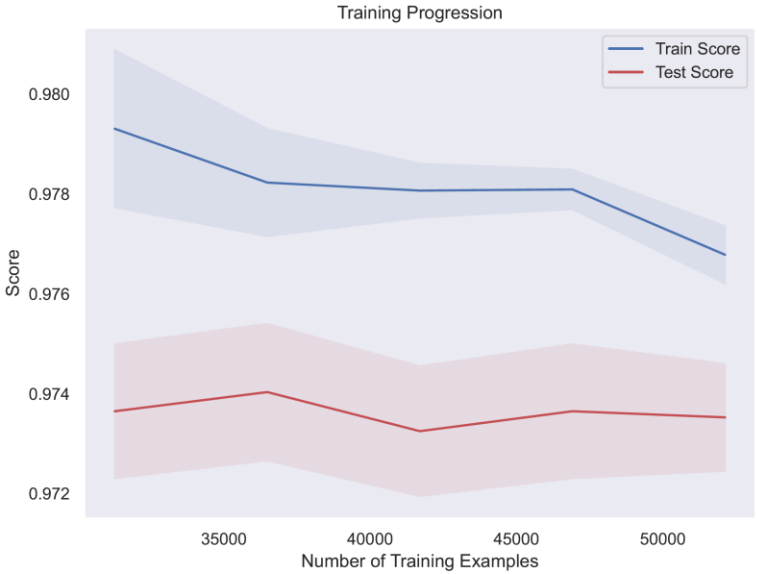
- Removing highly correlated features (> 0.8)

KS Tests

- Tested various depth/estimator count combinations for BDT
- Checked KS test scores compatible with training and validation sets drawn from same distribution

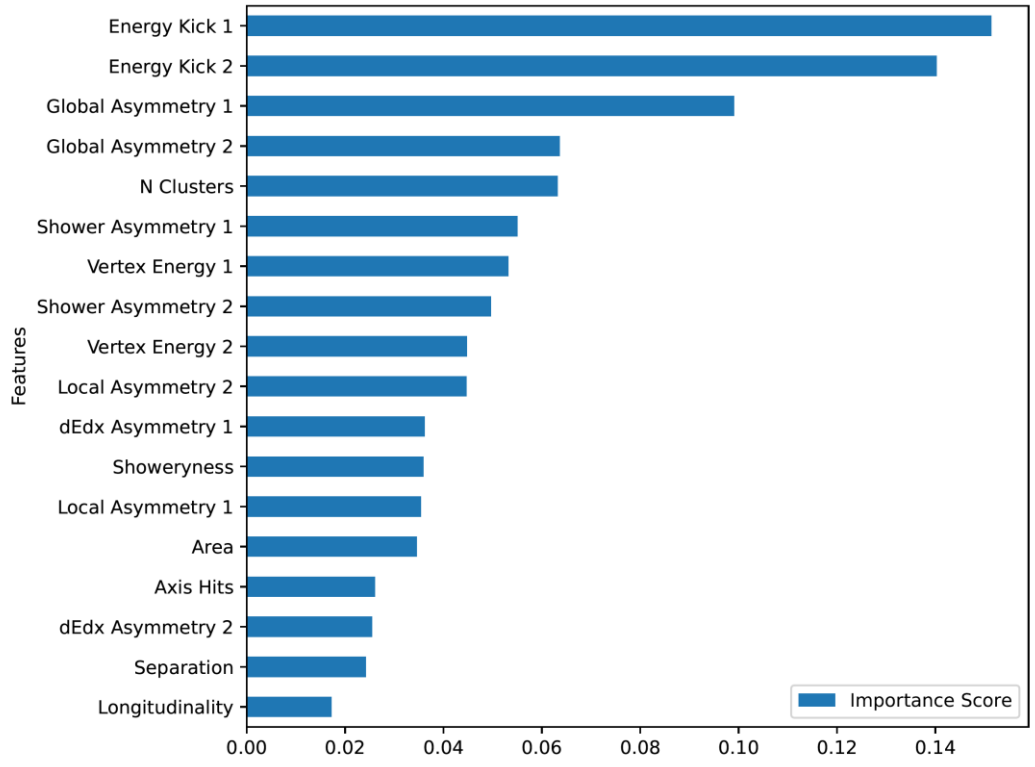


Training Set Size

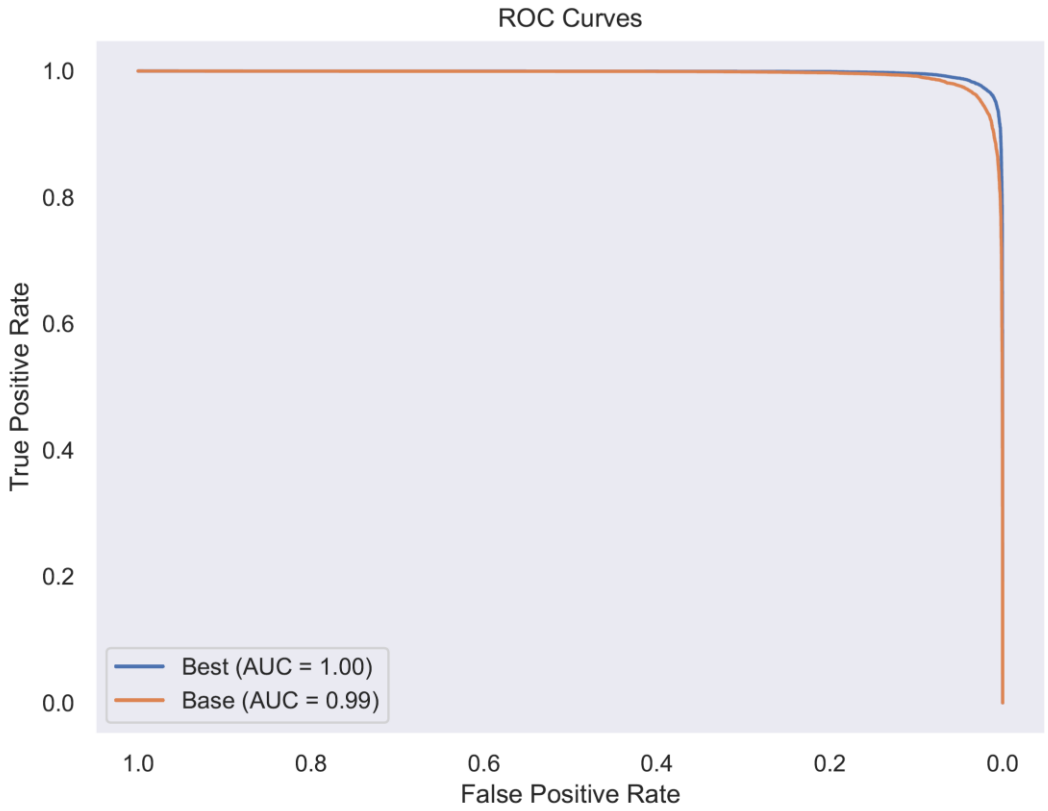


- Test scores stable across a broad range of sample sizes
- Region BDT performant even for quite small samples
  - ~36K training examples equivalent to about 4K events

Feature Importance



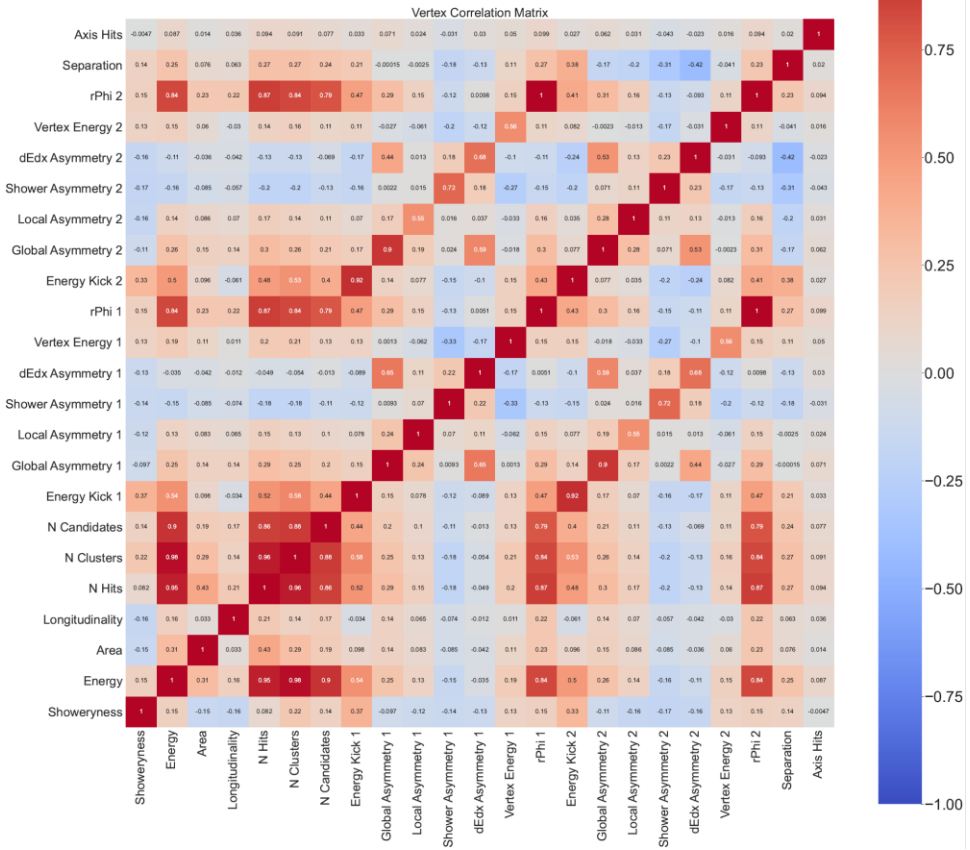
ROC Curves


ROC Curves
True Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
Best (AUC = 1.00)
Base (AUC = 0.99)
1.0
0.8
0.6
0.4
0.2
0.0
False Positive Rate

Fine-tuning BDT

WARWICK

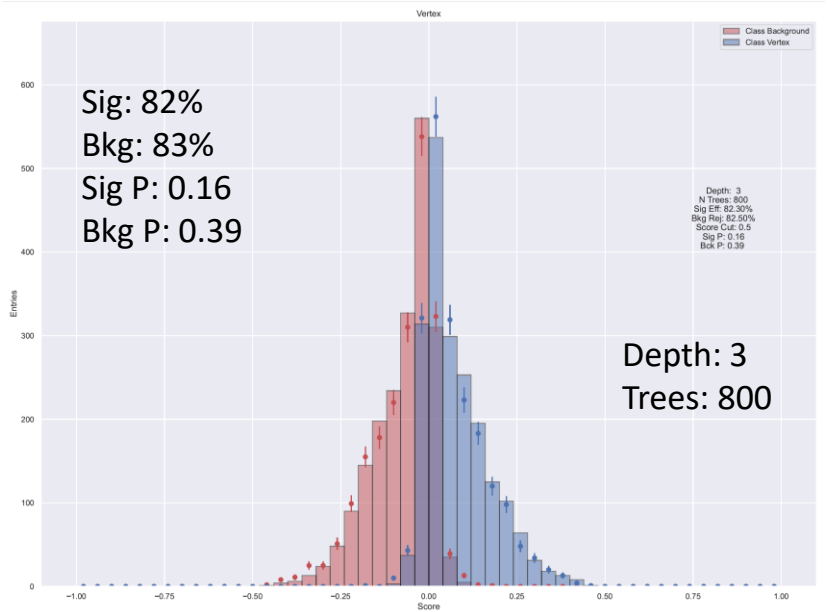
Vertex Correlation Matrix



- Removing highly correlated features (> 0.8)

KS Tests

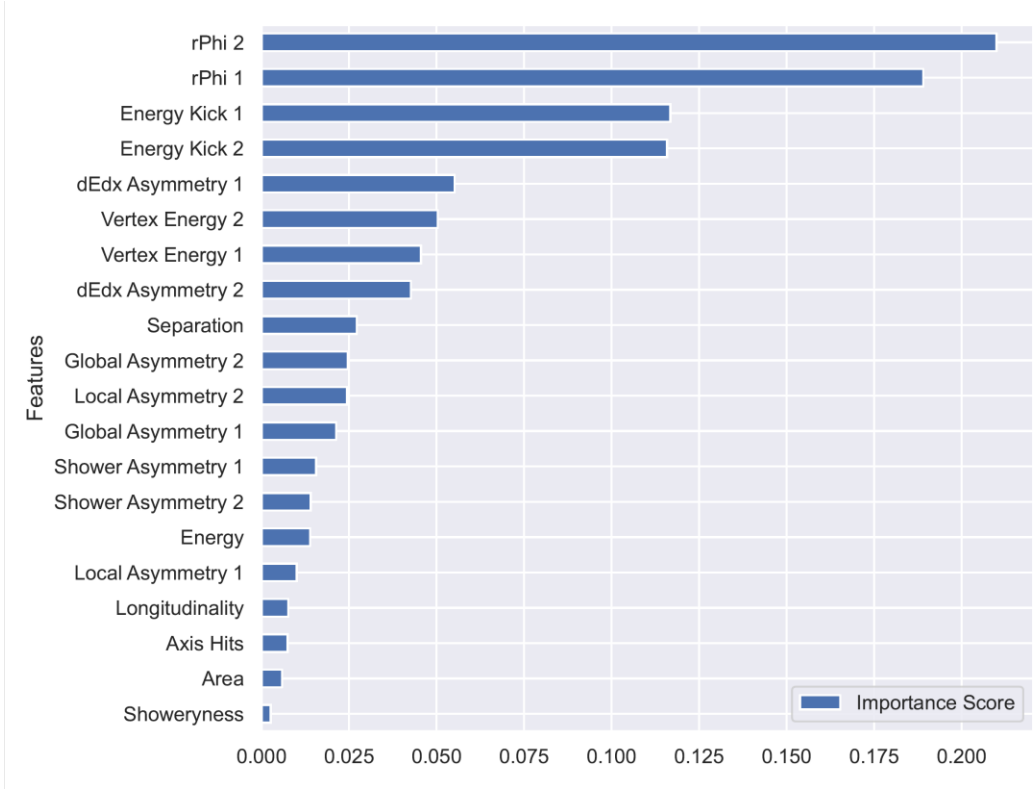
- Tested various depth/estimator count combinations for BDT
- Checked KS test scores compatible with training and validation sets drawn from same distribution



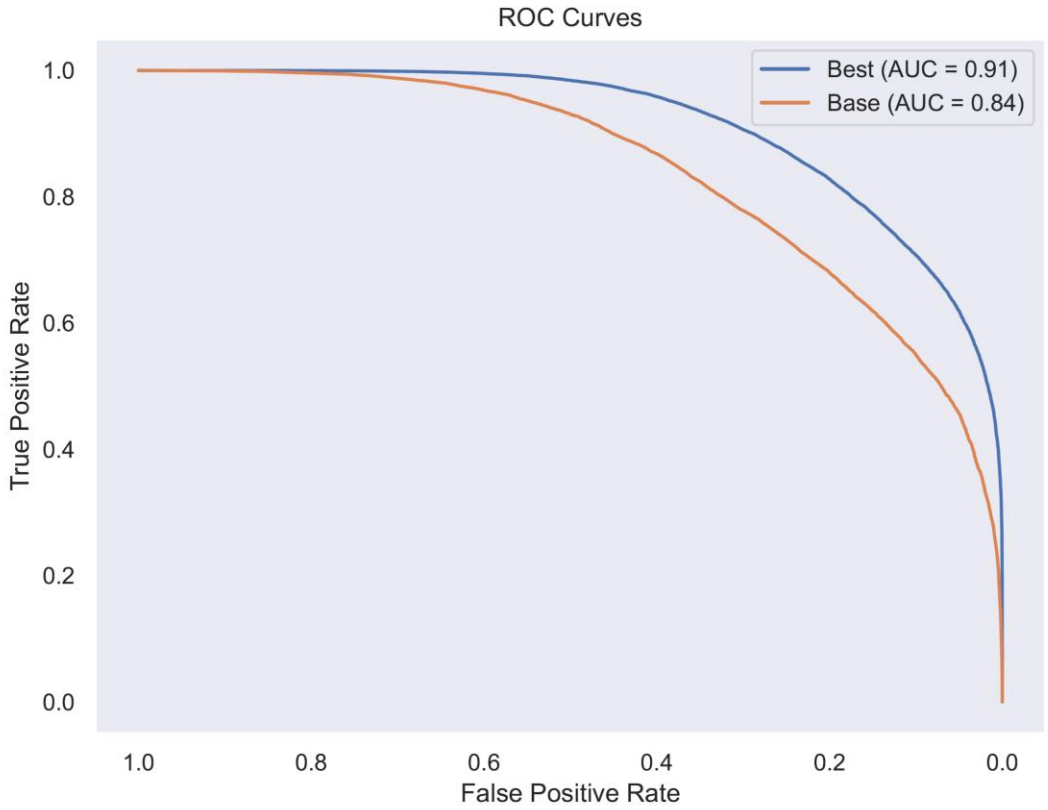
Training Set Size

- Need many more training examples for fine tuning
- Performance stabilizes 80K-120K examples
  - About 21-32K events

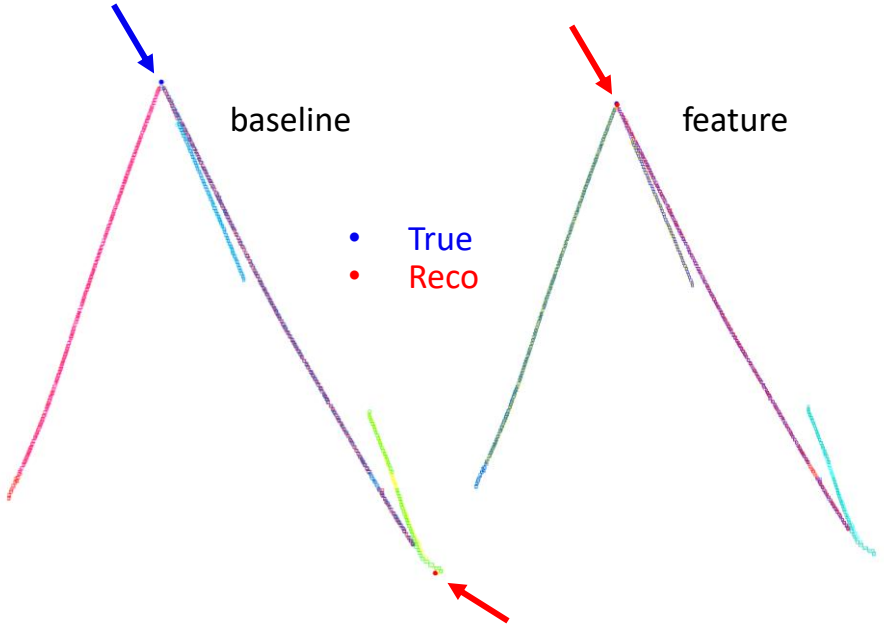
Feature Importance



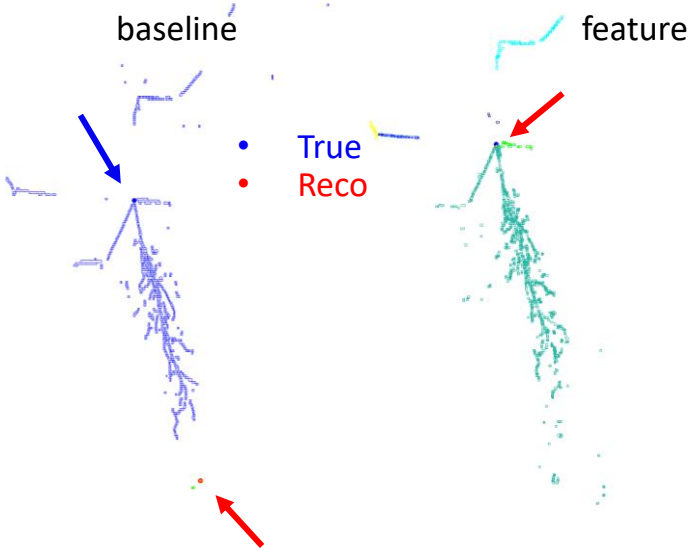
ROC Curves


ROC Curves
Best (AUC = 0.91)
Base (AUC = 0.84)
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
1.0
0.8
0.6
0.4
0.2
0.0

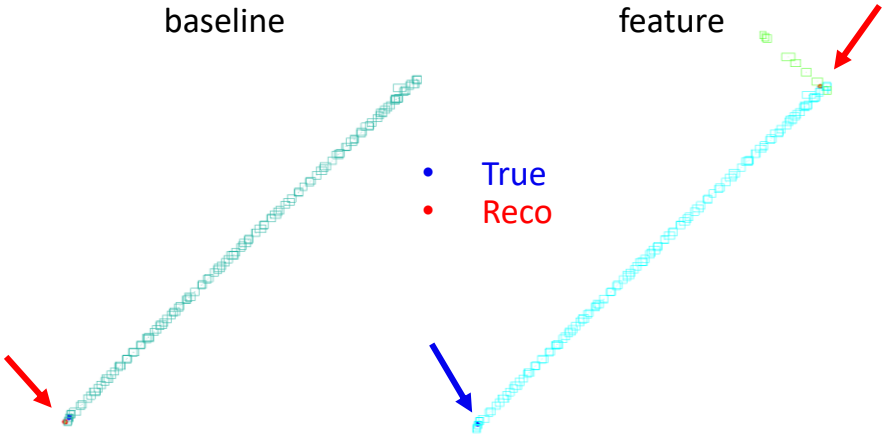
Event comparisons

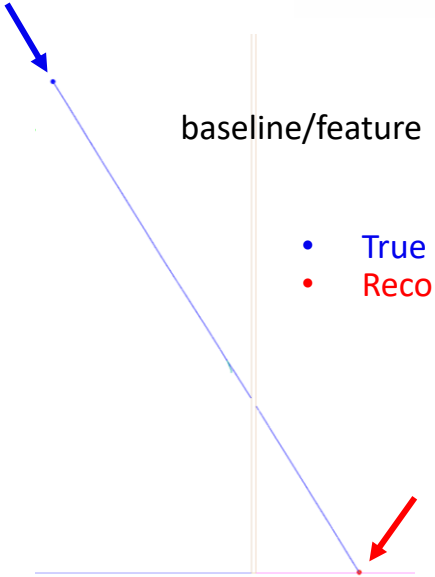


- Better performance for backward going events



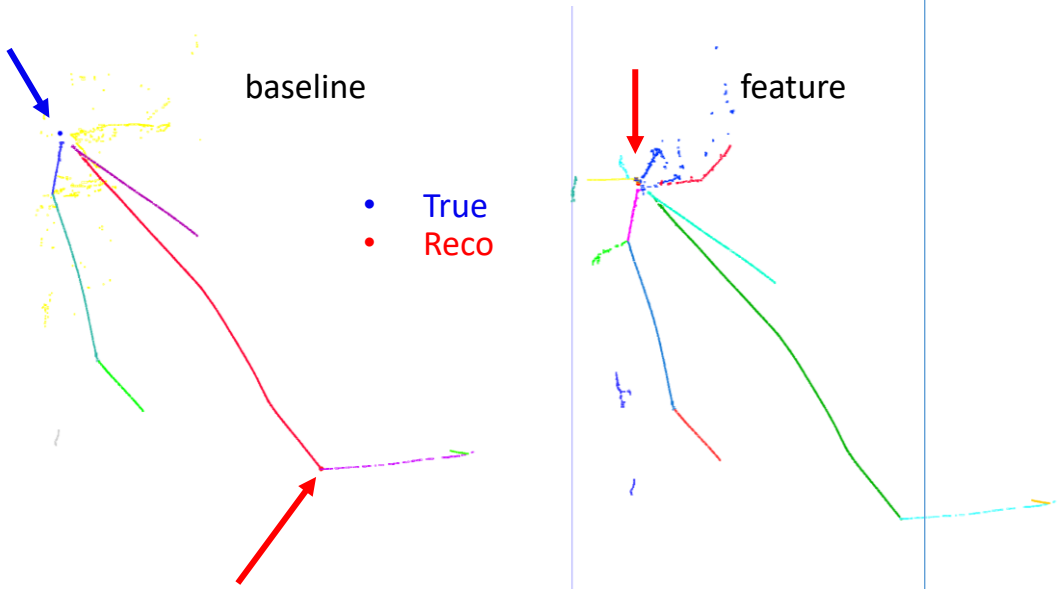
Event comparisons

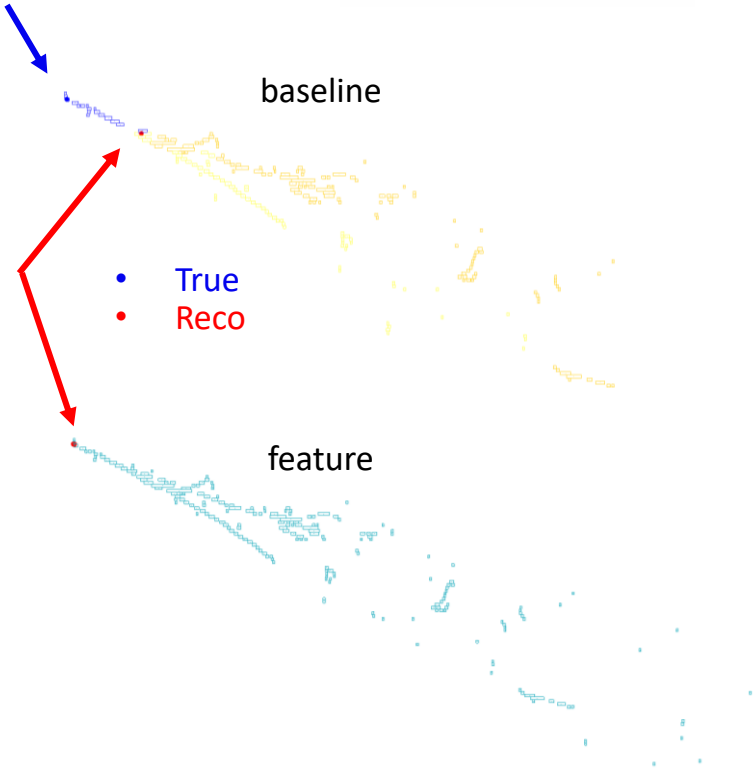




- Remaining/new errors are often ambiguous cases

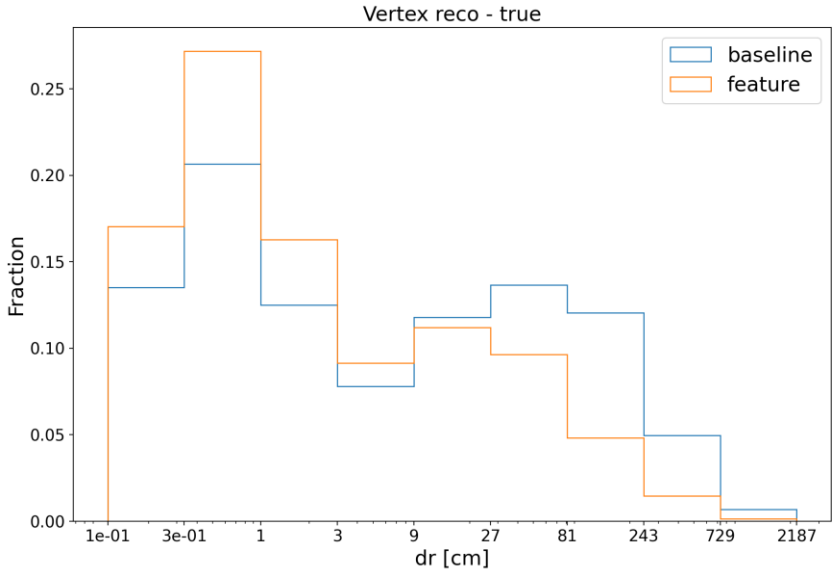
Event comparisons

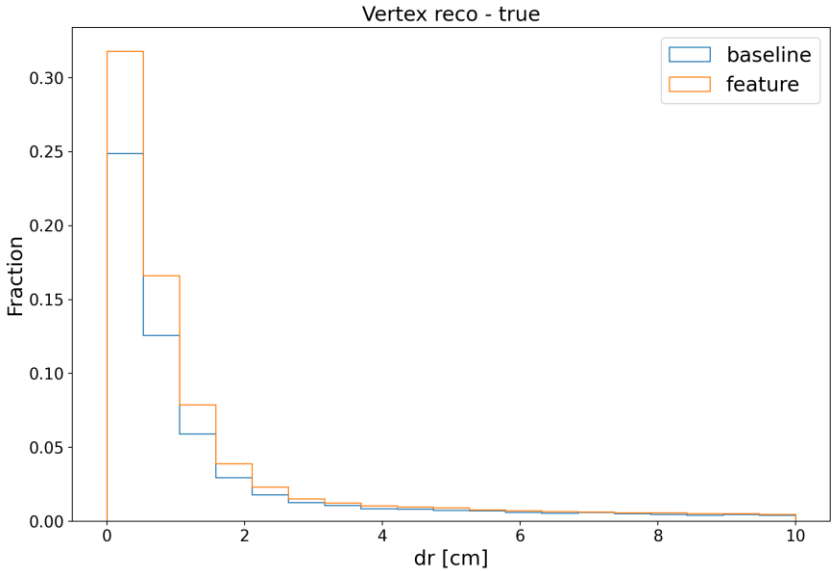




- Better performance for backward going events

Performance comparisons





- Accelerator (baseline) vertexing dr68: 26.1 cm
- Atmospheric (feature) vertexing dr68: 5.3 cm

Summary

WARWICK

- Removing the beam assumption notably improves vertexing performance
- Though suppressed, some large errors are still present, so further gains likely possible
- Updated algorithm and XML to be available in the near-term Pandora release
- Ongoing development and next steps
  - Look for further algorithmic improvements to vertexing
  - Neural network-based vertexing in development