IceTop-CNN

Ethan Dorr & Dr. Frank McNally
Mercer University

Outline

- Project Overview
  - Science Goals
  - Educational Objective
  - Project design
- Baseline Results
- Getting Started
- Planned Projects

Science Goals

Goal: a fast framework for reconstructing properties of CR primaries

- **Quantity over (peak) quality**
  - Maximize event statistics and wide field-of-view for **anisotropy-related analyses**
- **Maximize input information**
  - Use **low-level inputs** of charge and time for each tank
  - Potentially sensitive to energy (charge deposited), core position and direction (hit timing), composition (curvature and smoothness of shower profile)…

Potential use cases:

- Seed finder for alternative reconstruction tools
- Spectral anisotropy
- Rigidity dependence of anisotropy

Educational Objective

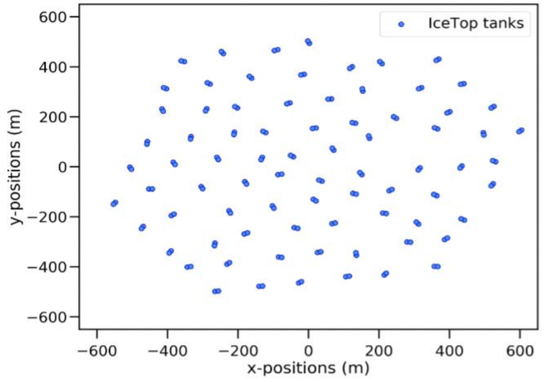
Goal: an introduction to machine learning, created by and for undergraduate students

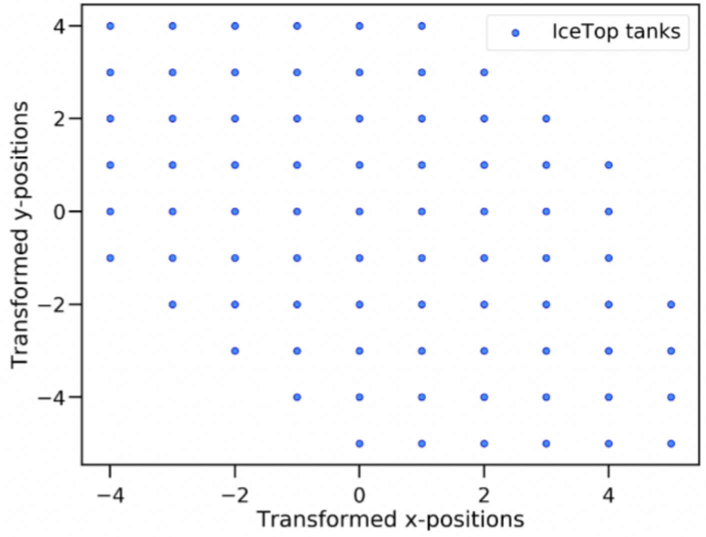
- **Start from zero**
  - Assume **no prior programming knowledge**
  - Run on IceCube computing resources
- **Modular**
  - **Easy to adjust** model complexity, reconstructed quantity, input information…
  - Framework for creating and tracking multiple student projects
- **Assessment ready**
  - Ability to compare performance between students

Project Design

Method

- Treat each event as a 10px x 10px padded “image”
- Use **charge and time** from both tanks at each station
- Convolutional Neural Network (CNN) encodes spatial information about the detector (not perfectly)
- Trained on simulation datasets from 2012
  - Minimal snow accumulation



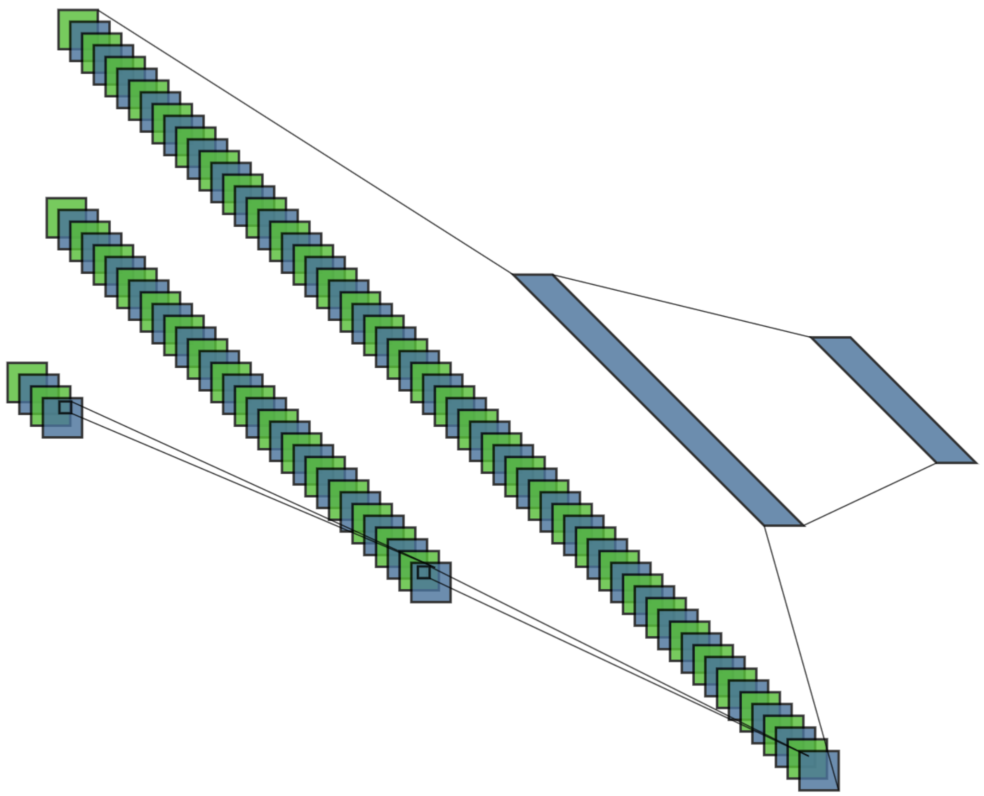


Source: Bauwens, I. (2020)

Project Design

Baseline Architecture

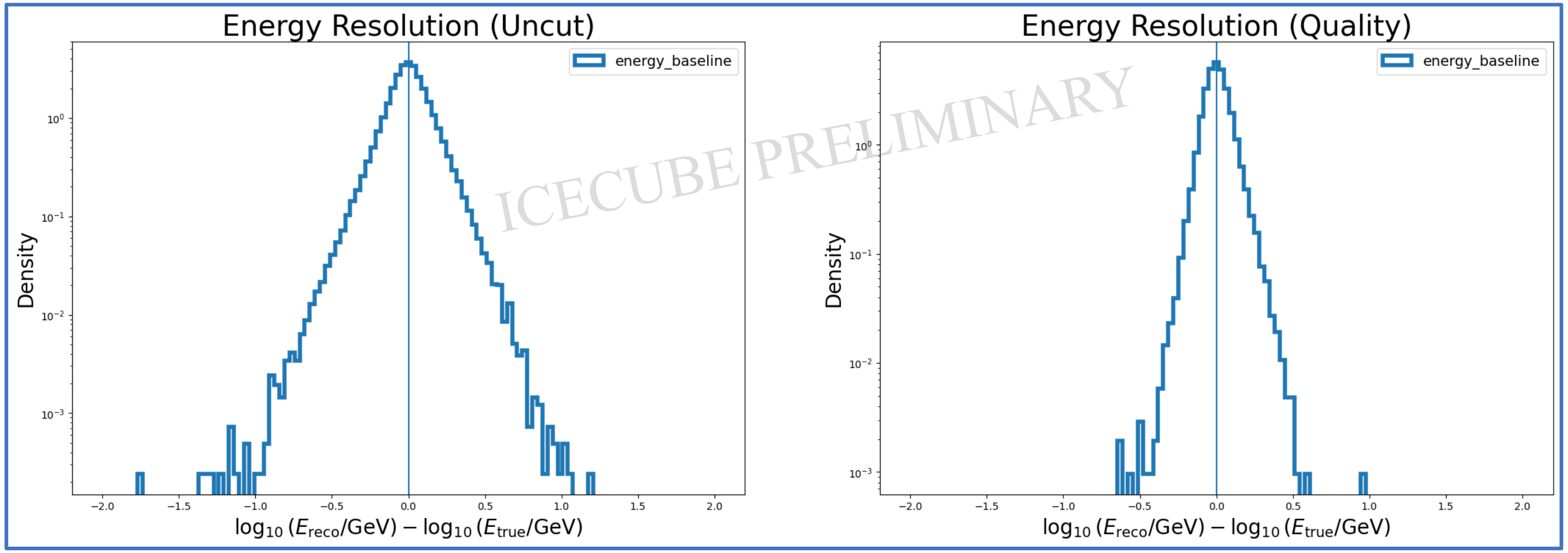
- Double convolutional layer
  - 32 and 64 filters, kernel size of 3
- Array flattened
  - Optional addition of **high-level parameters**
- Two dense layers
  - 64 and 32 nodes
- Output



Visualization credit: https://alexlenail.me/NN-SVG/LeNet.html

Results: Reading the Plots

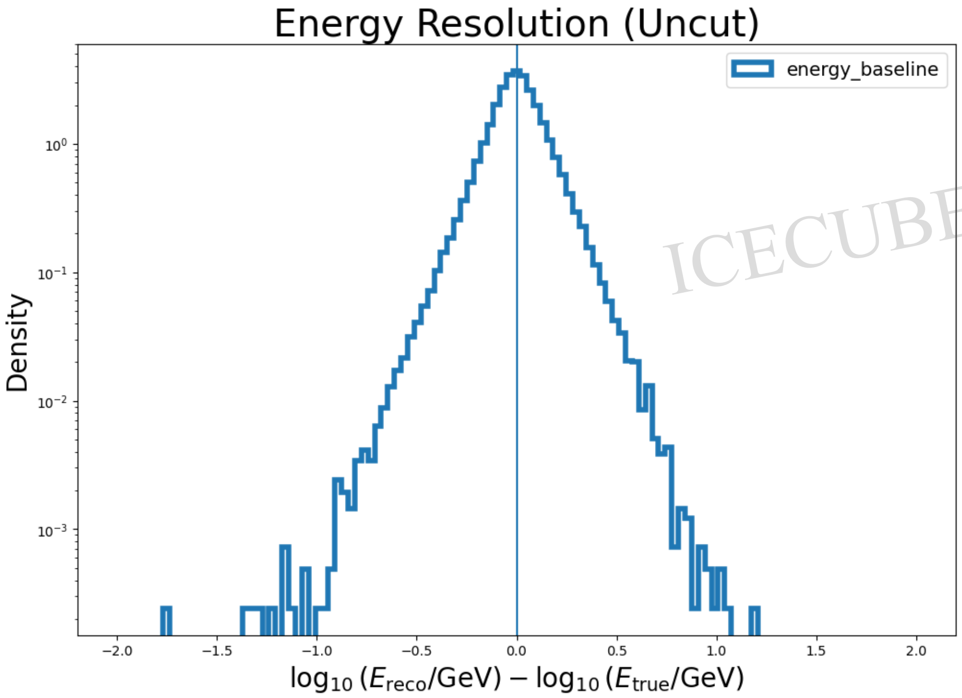
- Images shown for assessment on **all events** (*left*) and **quality cut** events (*right*)
- Models always trained on all events
- Energies in $\log_{10}(E/\mathrm{GeV})$
- 10% of simulation set aside for assessment

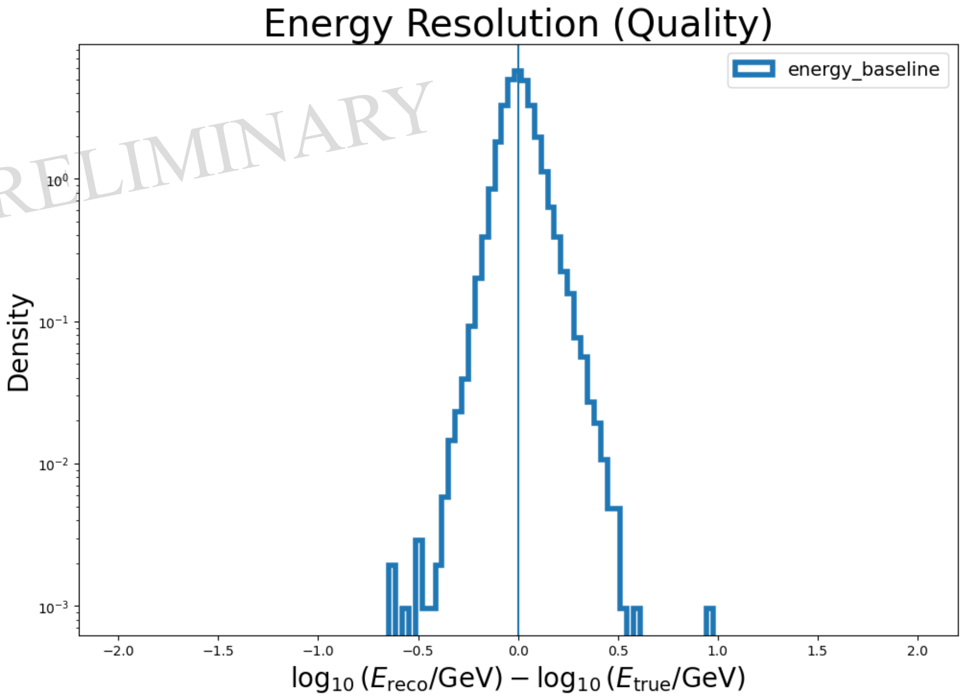


Quality-cut events must pass standard IceCube quality filters

- Maximum charge, station density, containment, etc.
- Zenith angle must be less than 40 degrees

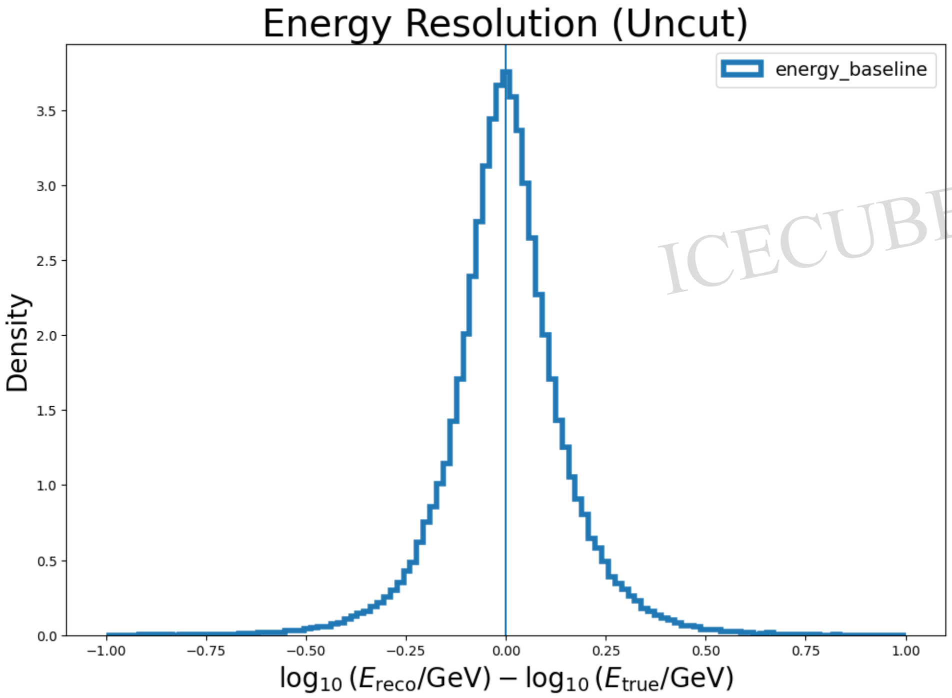
Results: Energy Resolution

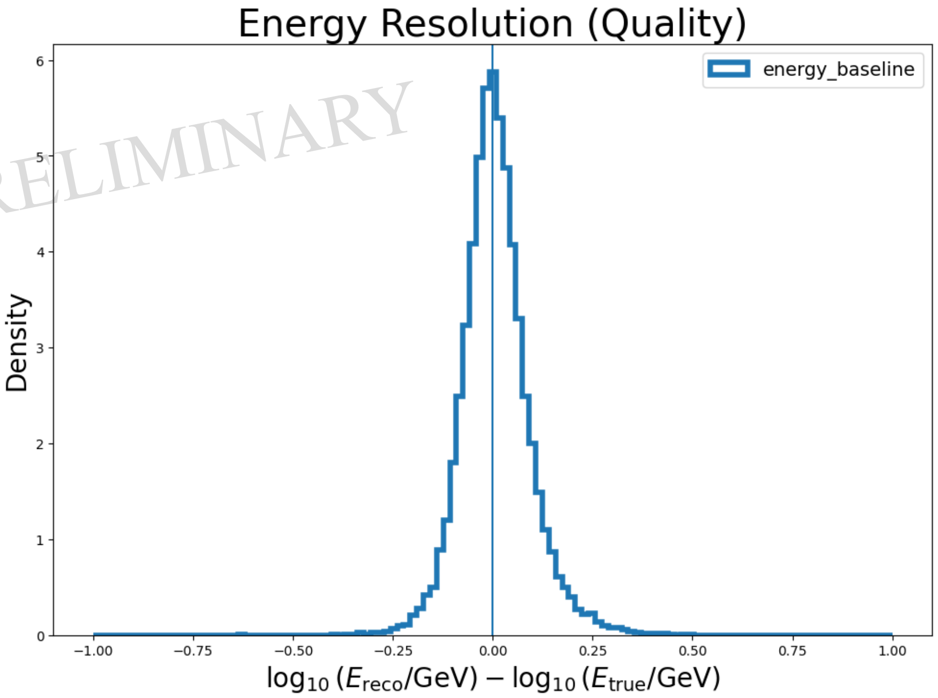




Event Sample	Uncut	Quality Cut
Energy Resolution (μ ± 1σ)	0.00 +0.13 -0.12	0.00 +0.08 -0.07
$\Delta\log_{10}(E) \leq 0.1$	60.0%	80.6%

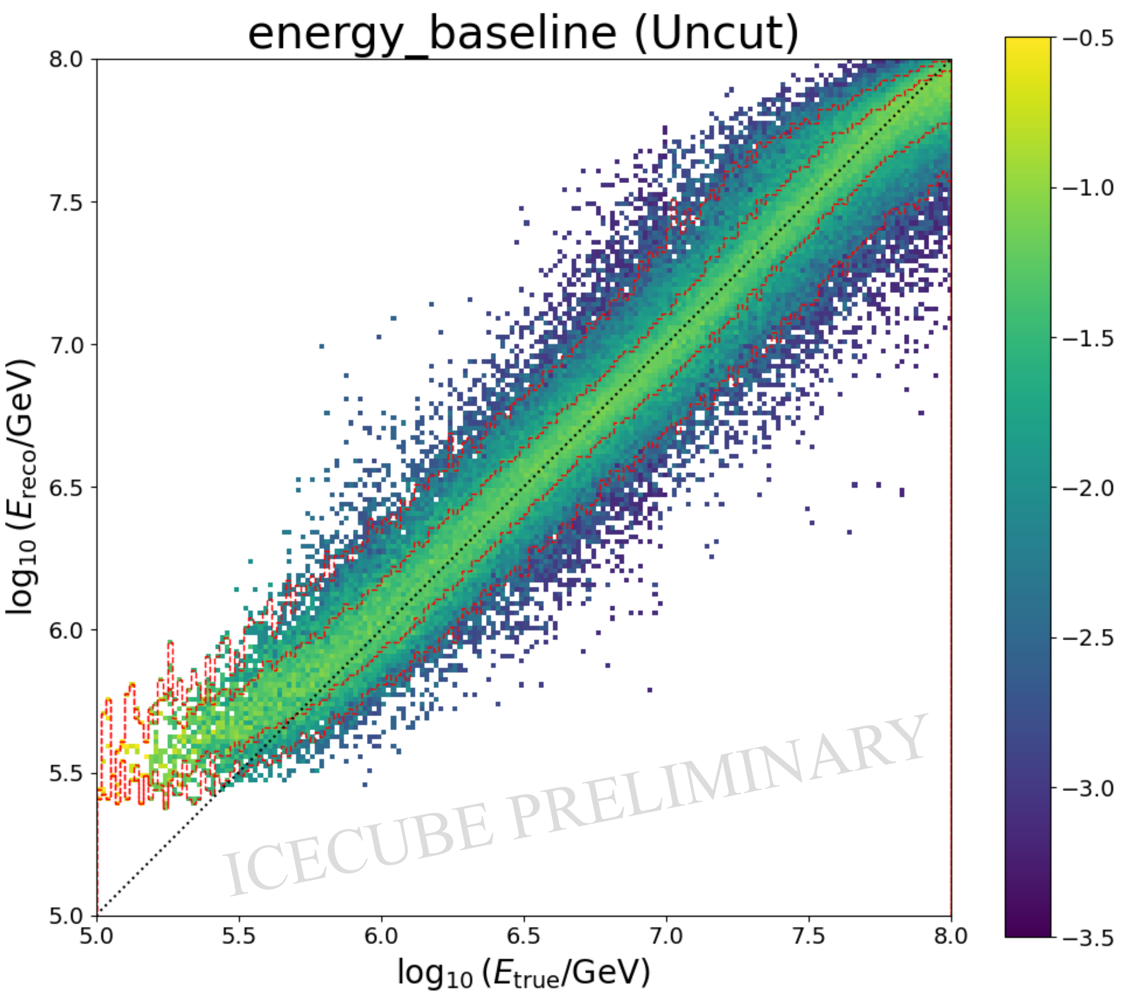
Results: Energy Resolution (Un-logged)

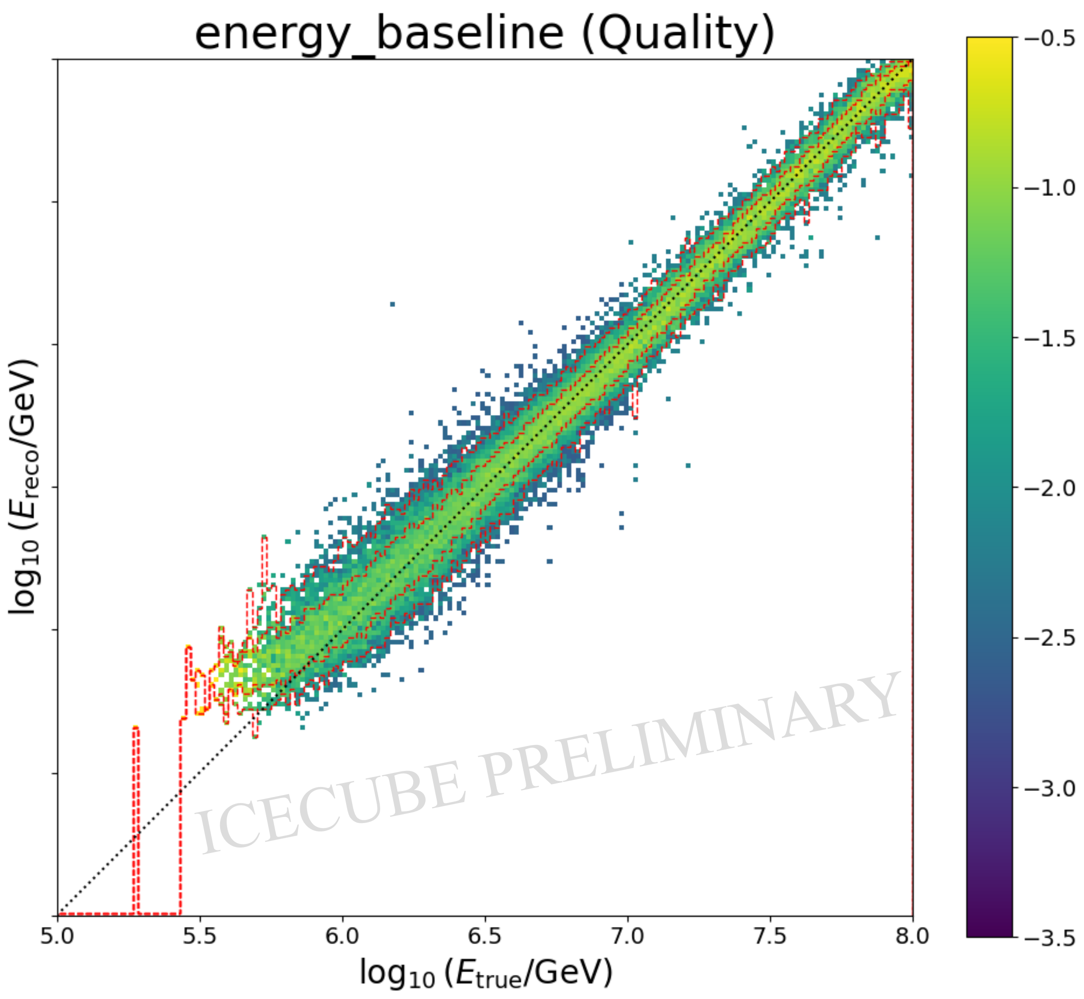




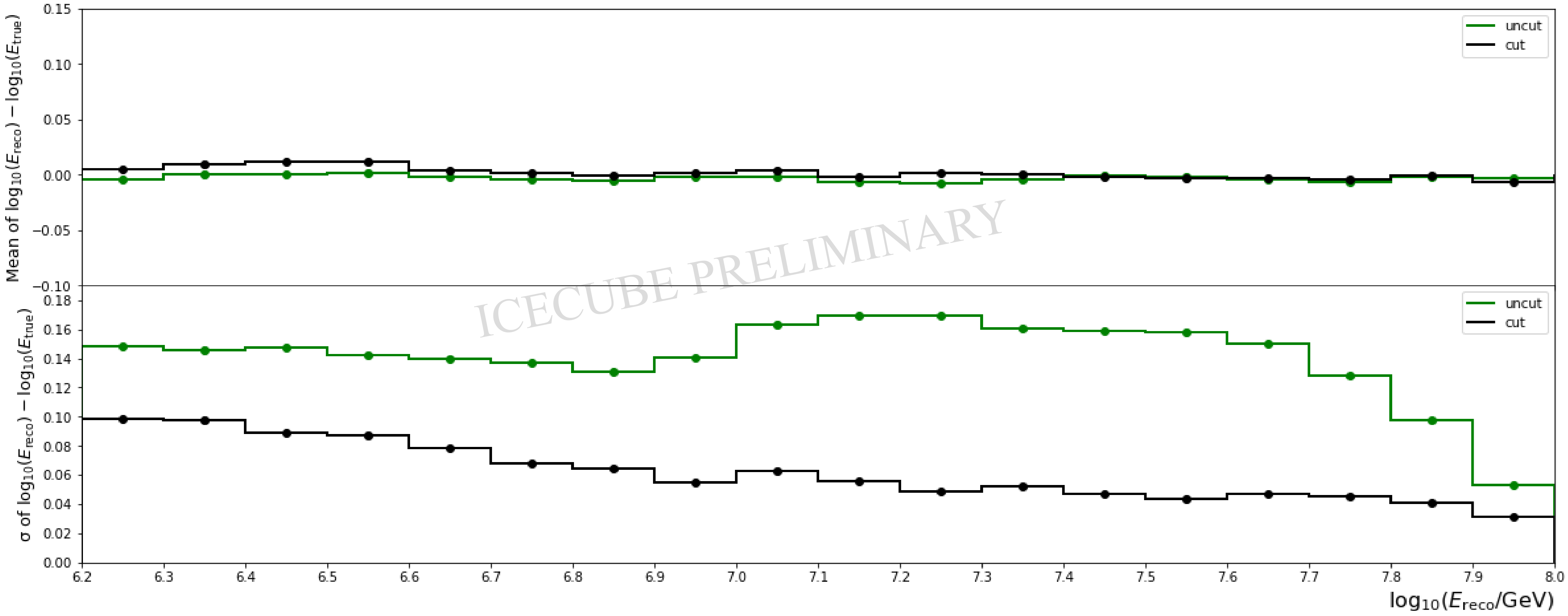
Event Sample	Uncut	Quality Cut
Energy Resolution (μ ± 1σ)	0.00 +0.13 -0.12	0.00 +0.08 -0.07
$\Delta\log_{10}(E) \leq 0.1$	60.0%	80.6%

Results: Energy Resolution vs True Energy

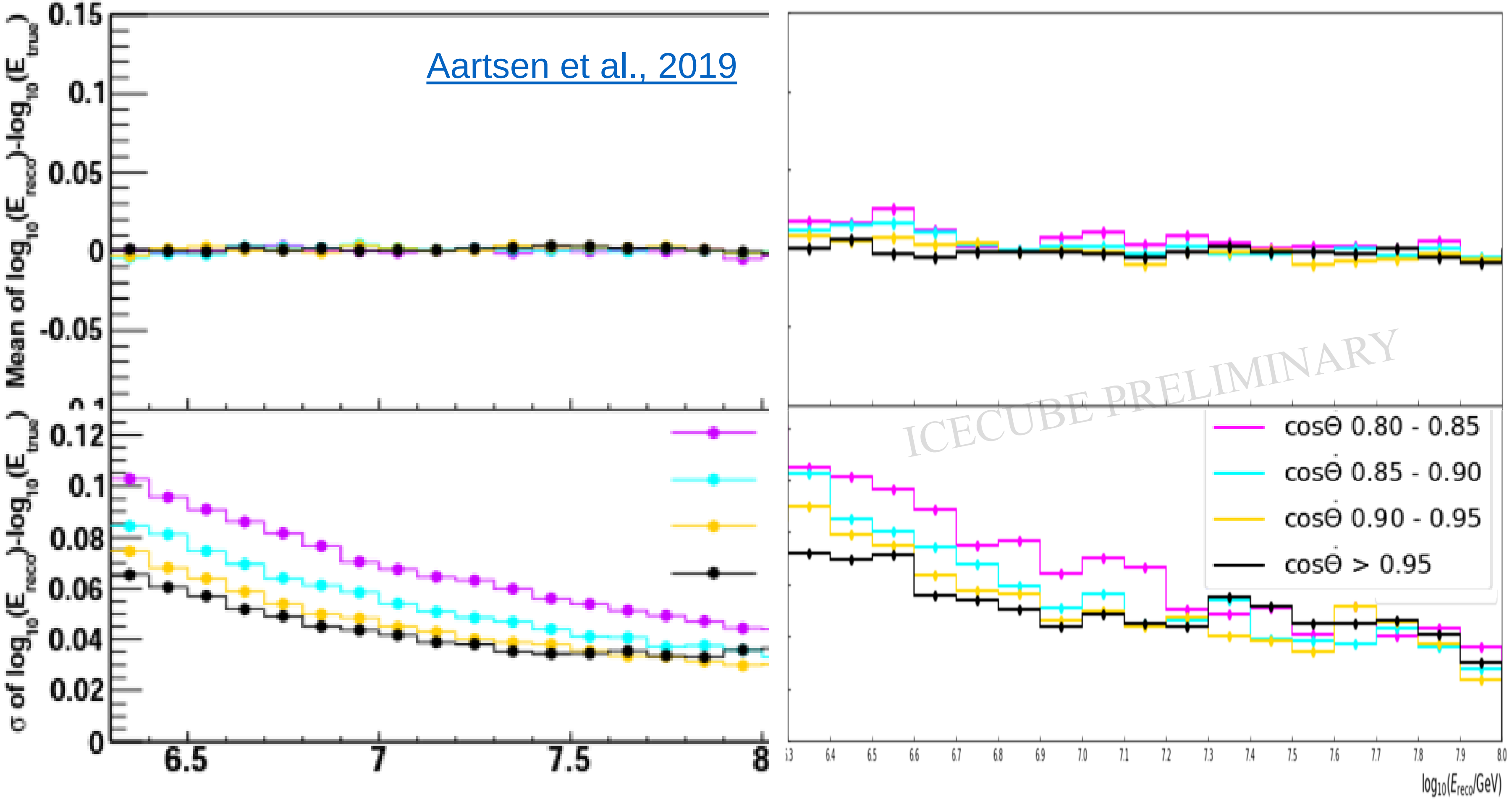




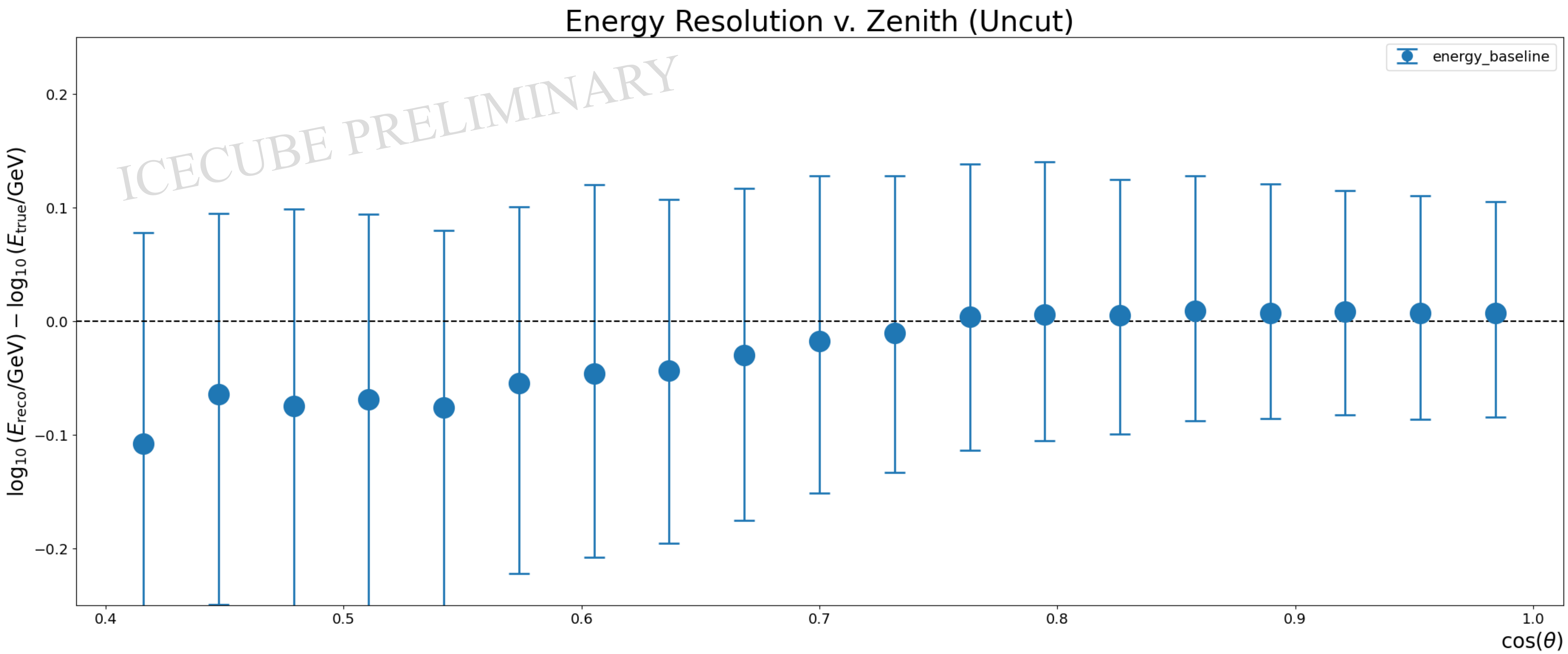
Results: Energy Resolution



Results: Energy Resolution vs. 3-Year



Results: Energy Resolution vs Zenith



Installation & Usage

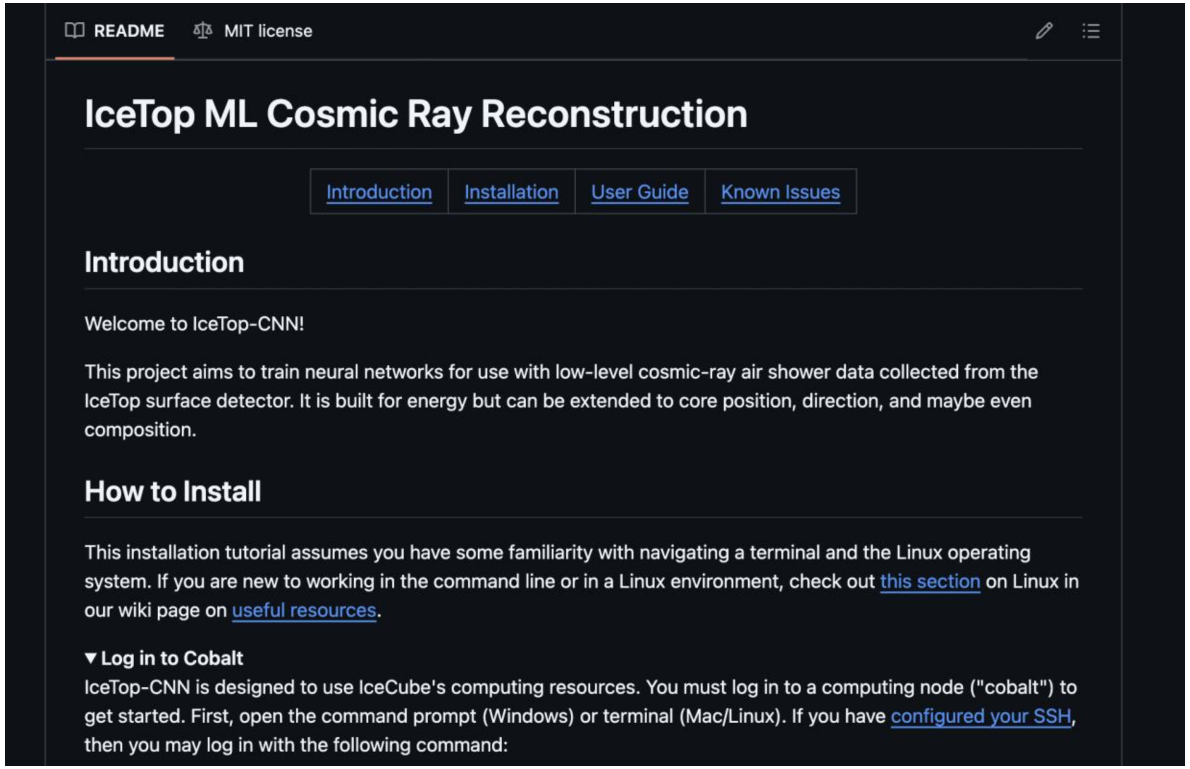
- Main resource: GitHub
  - Installation instructions
  - **Wiki**
- Designed to run on IceCube servers
  - Utilizes HTCondor
  - Avoids workstation-specific errors

Installation Steps

- Clone GitHub repository
- Run setup script
- Activate virtual environment

GitHub Repository

- https://github.com/fmcnallyi3/icetop-cnn


README MIT license

IceTop ML Cosmic Ray Reconstruction

Introduction | Installation | User Guide | Known Issues

Introduction

Welcome to IceTop-CNN!

This project aims to train neural networks for use with low-level cosmic-ray air shower data collected from the IceTop surface detector. It is built for energy but can be extended to core position, direction, and maybe even composition.

How to Install

This installation tutorial assumes you have some familiarity with navigating a terminal and the Linux operating system. If you are new to working in the command line or in a Linux environment, check out this section on Linux in our wiki page on useful resources.

▼ **Log in to Cobalt**

IceTop-CNN is designed to use IceCube's computing resources. You must log in to a computing node ("cobalt") to get started. First, open the command prompt (Windows) or terminal (Mac/Linux). If you have configured your SSH, then you may log in with the following command:

Settings/Features

- Training/validation split proportion
- Batch size
- Toggle infill detectors (additional detectors clustered around center)
- Toggle soft local coincidences
- Preprocess charge (min, max, average, etc.)
- Preprocess time
- Toggle directional data
- Maximum epochs
- Early stopping
- Plateau reduction
- Automatic loss graph creation
- …and more

Fundamental Conceptions/Challenges

- Models train best with as much data as can be provided
  - Includes “lower-quality” events
- Introducing additional model complexity rapidly results in overfitting
- Minimally-preprocessed data leads to best model performance
  - Exception: normalization
- Hyperparameter tuning has minimal impact on results

Future Work

- Behavior over time (snow depth)
- Inclusion of high-level parameters
- Multiparameter reconstructions
- Spectral anisotropy
- Composition

Acknowledgements

- **Project Supervisor**
  - Dr. Frank McNally

- **Early group**
  - Brandon Pascali (‘22)
  - Peter Richardson (‘22)
  - Tyler Sledge (‘20)
  - Roy Wood (‘20)

- **Special thanks**
  - Dr. Anthony Choi

Grant Clark (‘26)

Caden Hamrick (‘23)

Kennedy Mays (‘24)

Kriti Mittal (‘24)

Marlon Oliver (‘24)

Charbel Youhanna (‘26)

Thank You

Questions?

Backup Slides

Bonus: Summary of Tests

TEST	VARIATIONS	RESULT
Pulse Series	-OfflineIceTopTankPulses -LaputopSeededTankPulses -RTSeededTankPulses -SnowUnattenuatedPulses	Best resolution achieved with OfflineIceTopTankPulses
Local Coincidences	-HLC only -SLC included	Best resolution achieved with both HLC and SLC pulses
STA5 Filter	-Events must pass STA3 -Events must pass STA5	Best resolution achieved with events that only must pass STA3
Layer Merging	-Sum charge -Mean charge -Mean time	Best resolution achieved with unmerged layers
Time Shifting	-No shift -All time values shifted by minimum nonzero time	Resolution improves with minimum relative time at 0

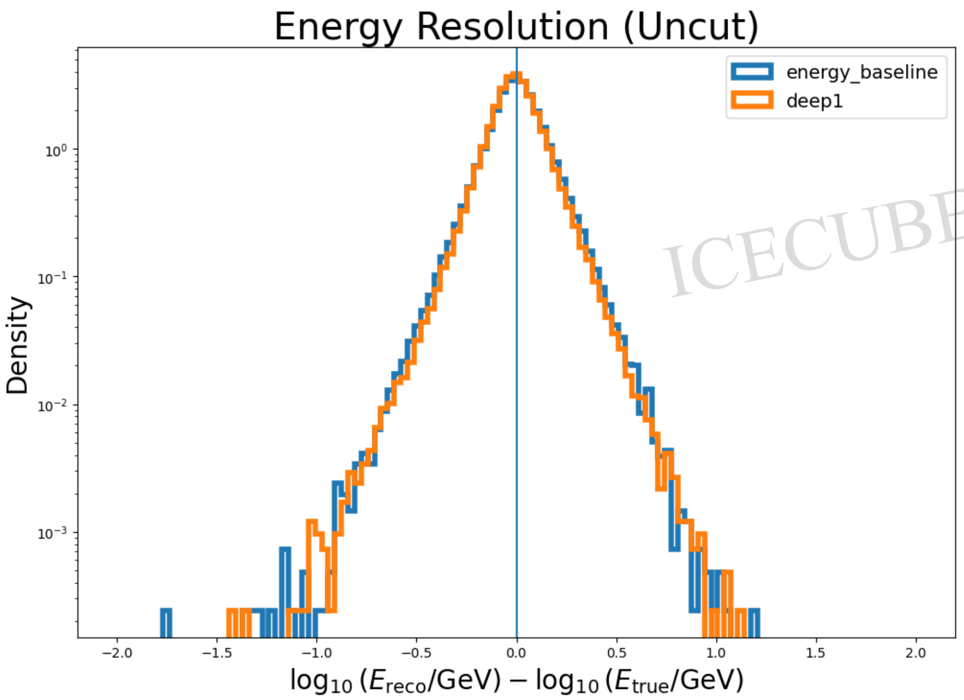
Summary of Tests contd.

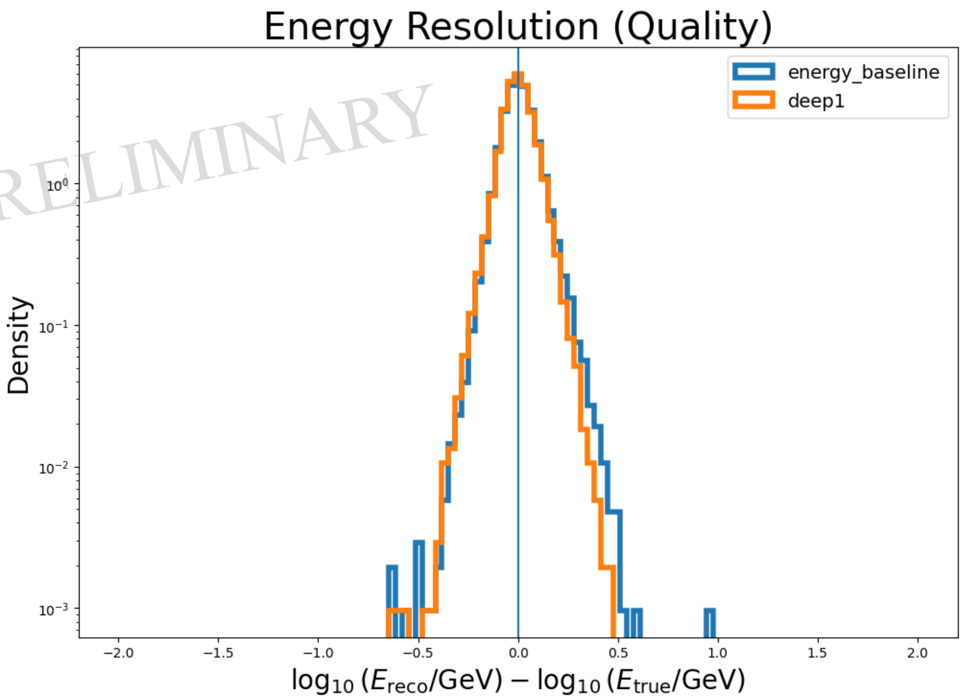
TEST	VARIATIONS	RESULT
Event Reflecting	-Horizontal -Vertical -Diagonal	Best resolution achieved with no reflected events in training dataset
Event Rotating	- 90° orthogonal rotations - 60° hexagonal rotations	Best resolution achieved with no rotated events in training dataset
Dropout	Many	Best resolution achieved with 20% dropout layers between dense layers
Pooling	-Max pooling -Average pooling	Best resolution achieved with max pooling between every two convolutional layers
Optimizer	-SGD -Adam -AdamW	Best resolution achieved with weighted AdamW optimizer

Summary of Tests contd.

TEST	VARIATIONS	RESULT
Loss Function	-Mean absolute error -Mean squared error -Huber	Best resolution achieved with Huber loss function
Reduce Learning Rate on Plateau	-No change in learning rate -Learning rate reduced upon plateau in validation loss	Best resolution achieved with a learning rate reduction by a factor of 0.4 after 10 epochs with no decrease in validation loss

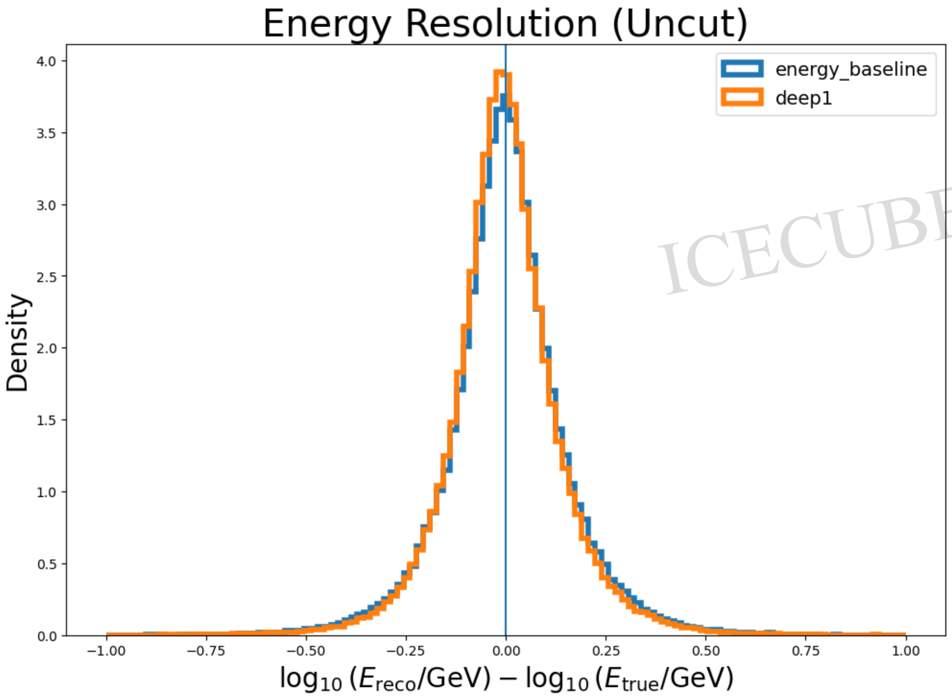
Double Bonus: Model Comparison

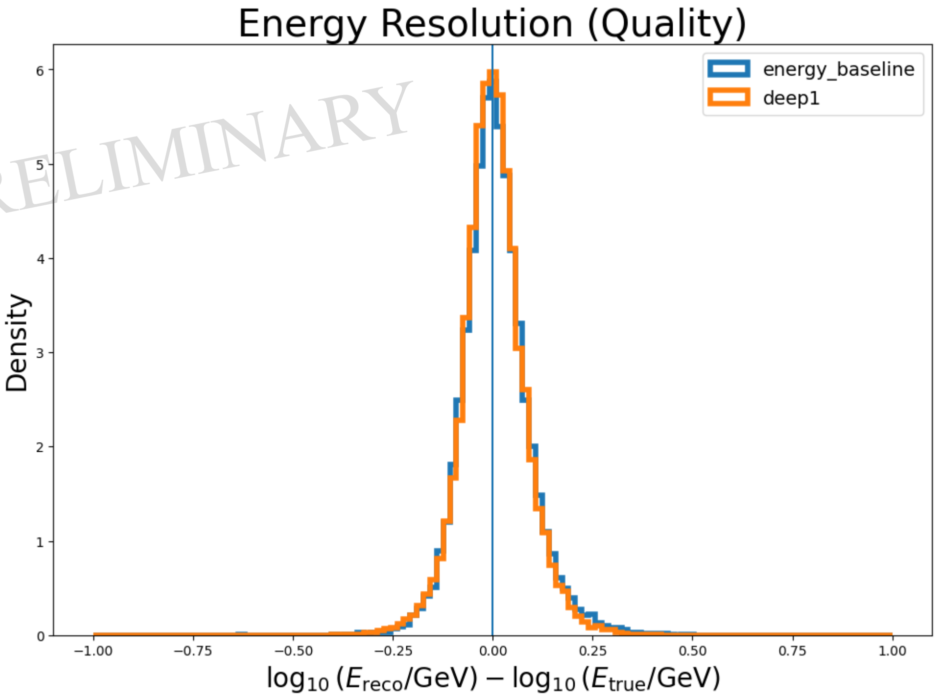




Event Sample	Uncut	Quality Cut
Energy Resolution ($\mu \pm 1\sigma$)	0.00 +0.13 -0.12	0.00 +0.08 -0.07
$\Delta\log_{10}(E) \leq 0.1$	60.0%	80.6%

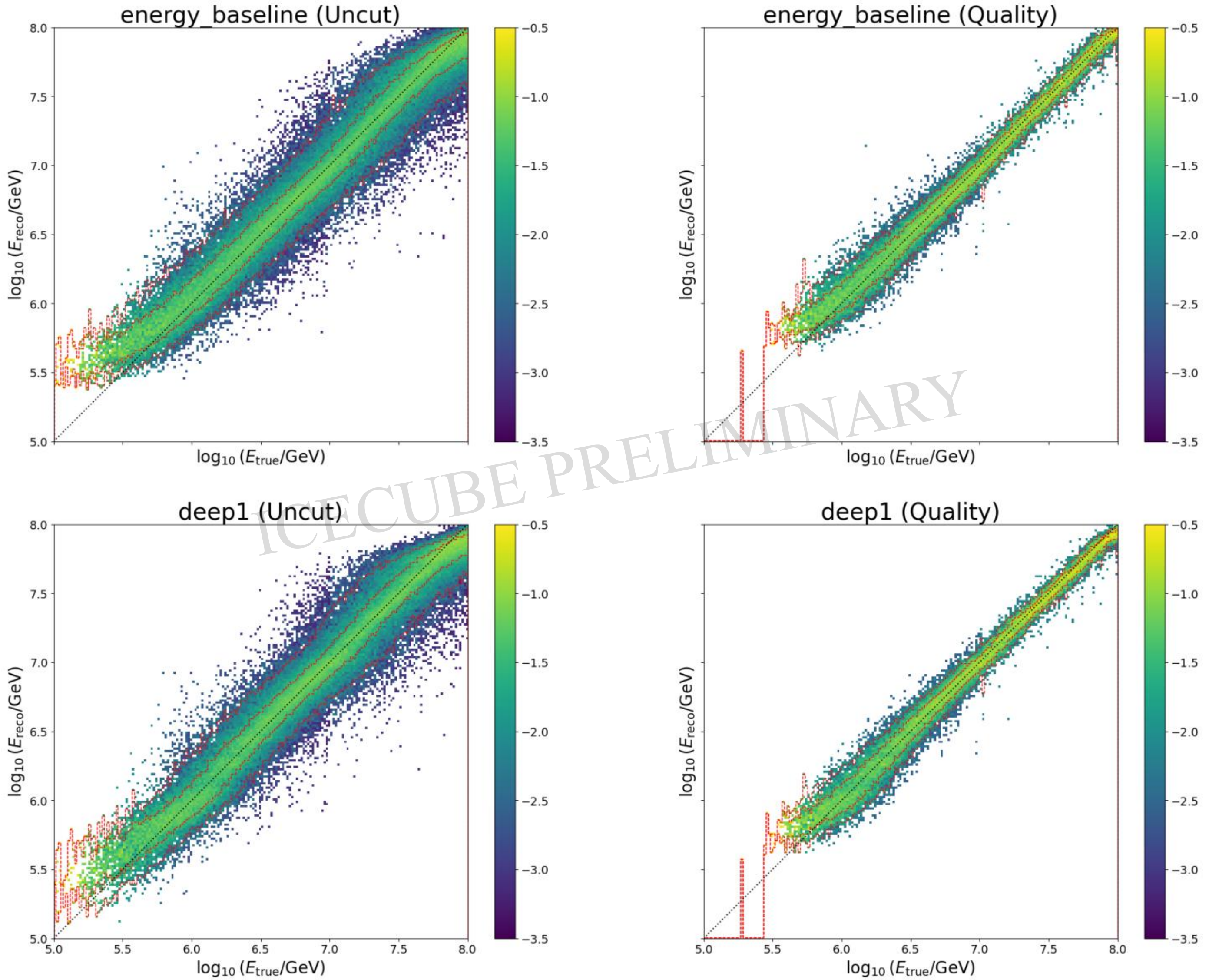
Model Comparison contd.



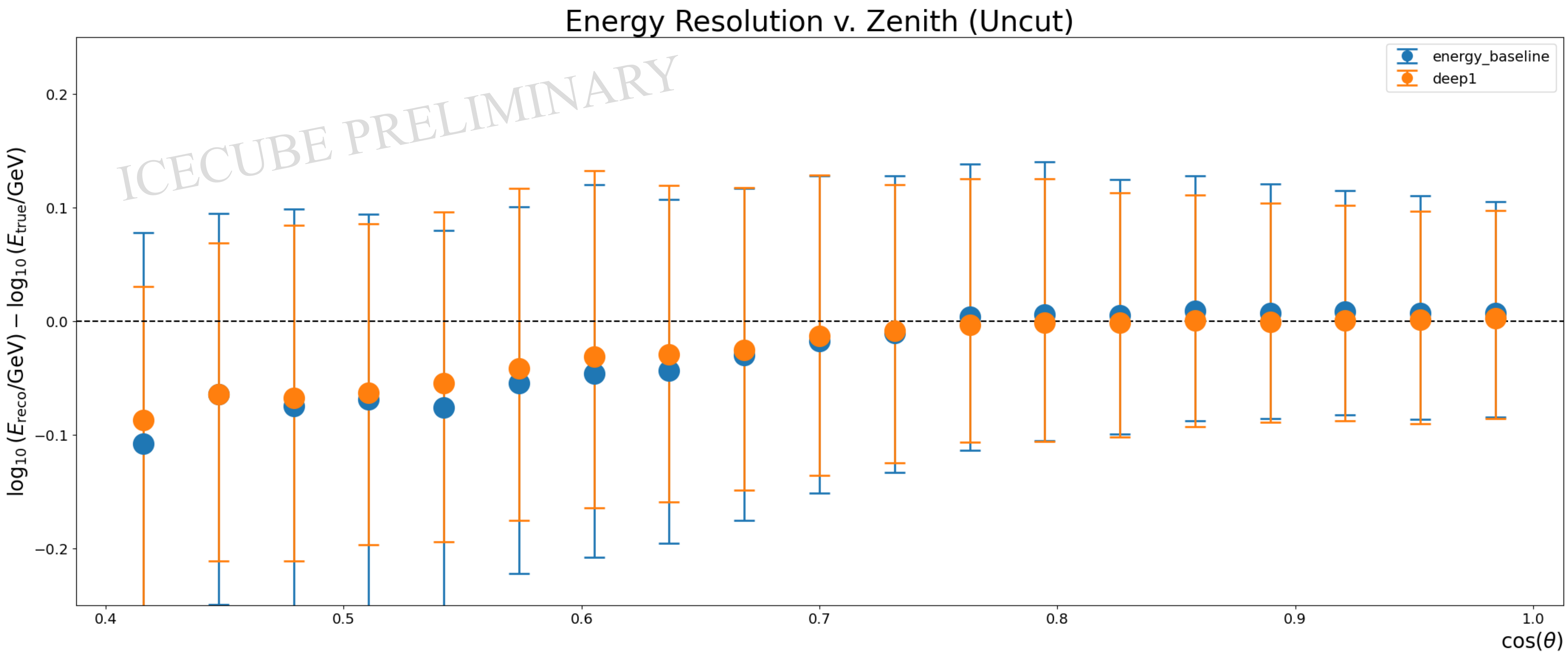


Event Sample	Uncut	Quality Cut
Energy Resolution (μ ± 1σ)	0.00 +0.13 -0.12	0.00 +0.08 -0.07
$\Delta\log_{10}(E) \leq 0.1$	60.0%	80.6%

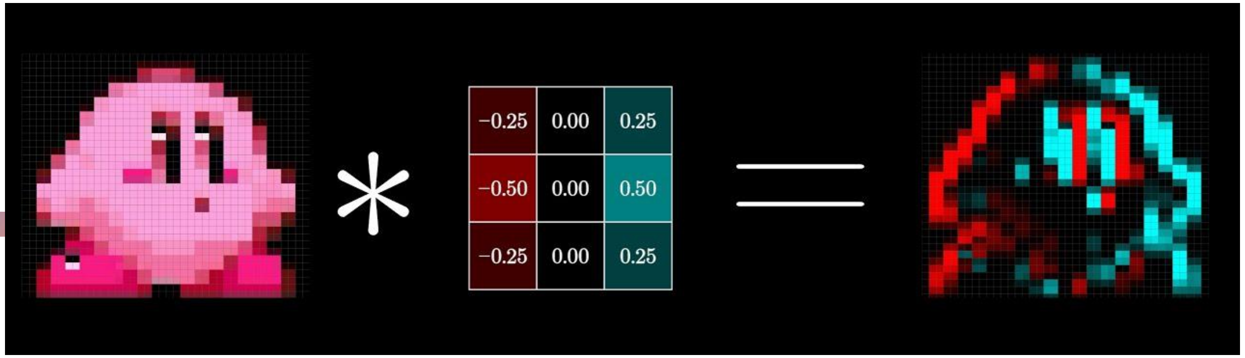
Model Comparison contd.



Model Comparison contd.



Introduction



Convolutional Neural Network (CNN): Neural network that specializes in extracting spatial information from an image.

- Complex image features built from simpler features
- Utilizes convolutions with sliding weighted **kernels** to extract spatial features

Idea is that neural network will detect lower-level features difficult for humans to recreate in higher-level parameters

Utilize **low-level charge and time** information to **minimize** potential information loss

Motivation is to provide energy estimation tool for potential use in anisotropy studies

Statistics for Data

NUCLEI	NO CUT	QUALITY CUT
Proton (12360)	328,012	85,212
Helium (12630)	319,051	81,297
Oxygen (12631)	308,358	76,978
Iron (12362)	294,281	72,411
TOTAL	1,249,702	315,898

Results: Energy Resolution Mean & Bias