Optimizing a Cosmic-ray Energy Estimator with Machine learning for the HAWC observatory

Workshop on Machine Learning for Analysis of High-Energy Cosmic Particles

University of Delaware

January 30, 2025

HAWC
High Altitude Water Cherenkov
Gamma-Ray Observatory

J. Jaime (UIS),

T. Capistrán* (UNITO),

I. Torres (INAOE),

for the HAWC Collaboration

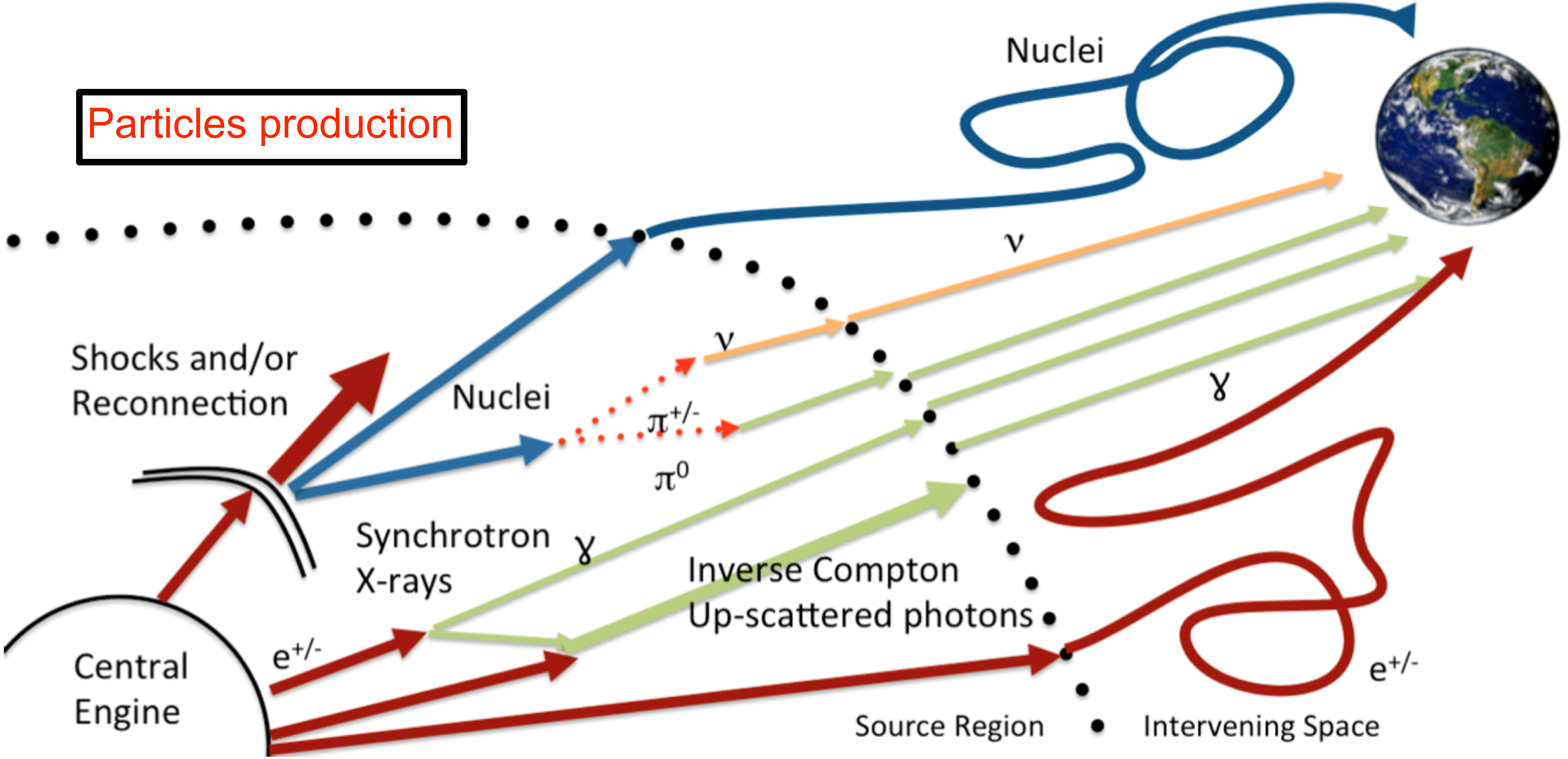
*speaker

Universidad Industrial de Santander

UNIVERSITÀ DI TORINO

INAOE

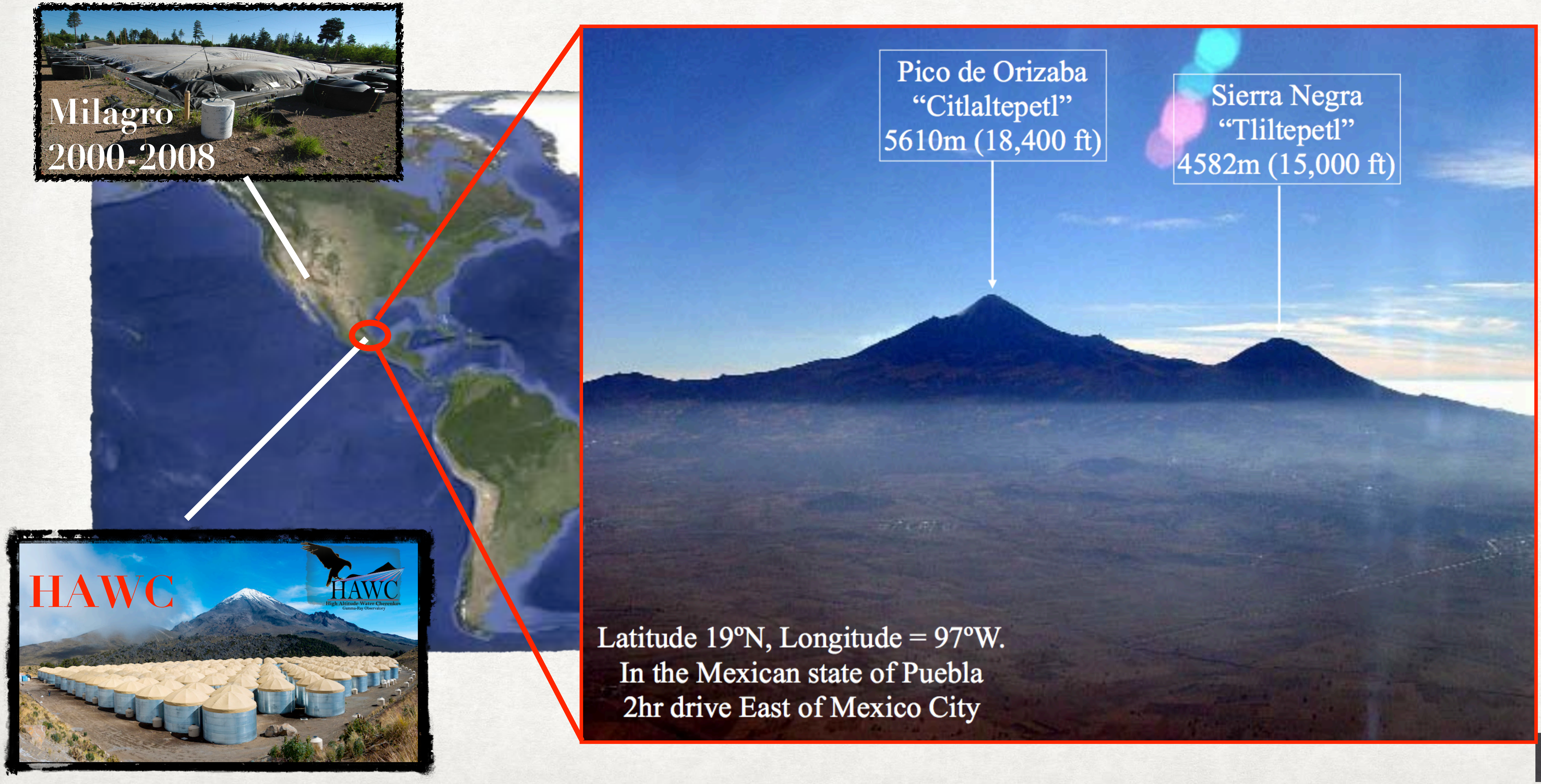
Particles production



Pretz, J. (2016), https://doi.org/10.22323/1.236.0025

Gamma-ray observatories

High-Altitude Water Cherenkov (HAWC)

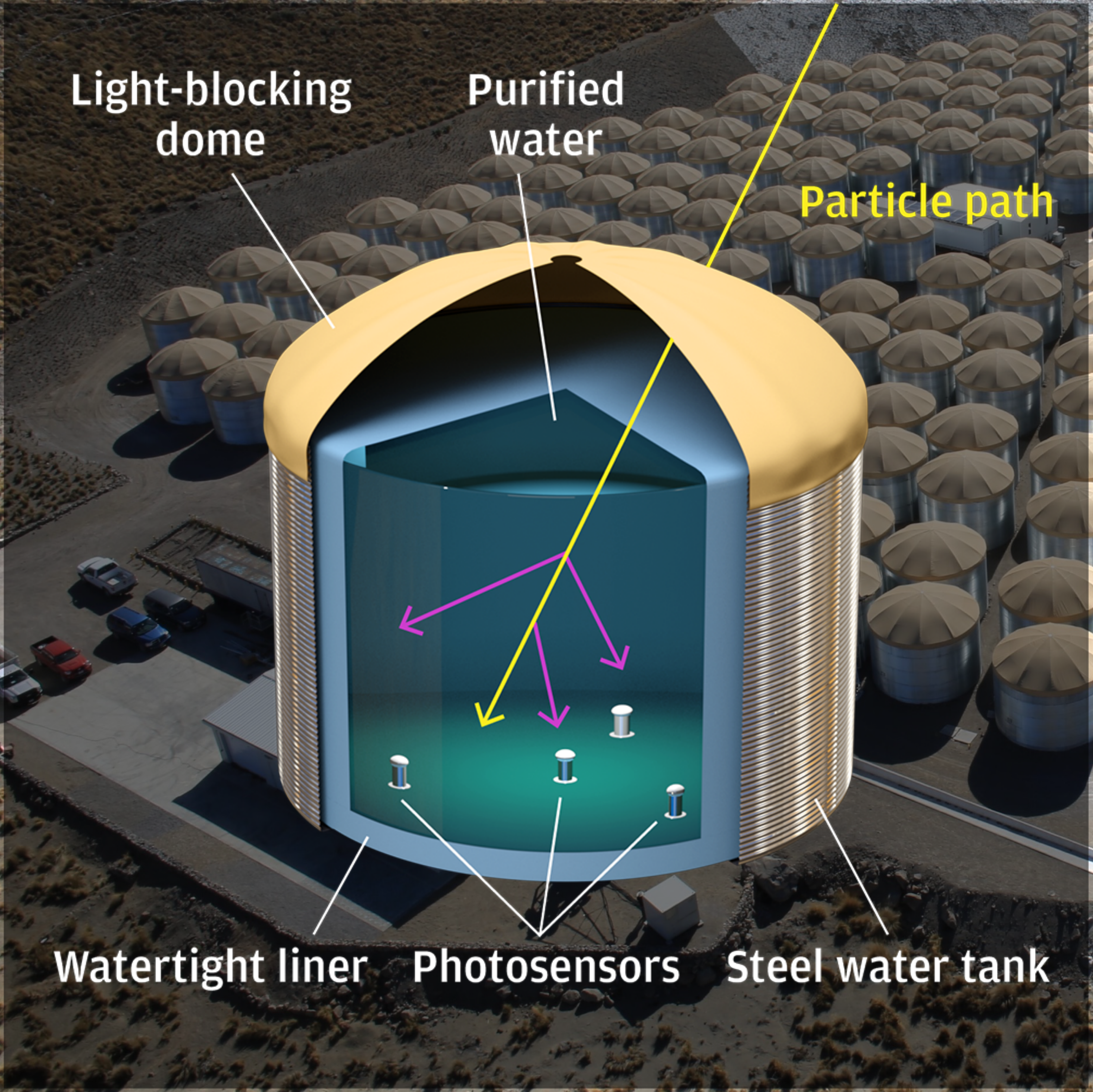


HAWC: High-Altitude Water Cherenkov

Some characteristics:

- DC > 95%
- Wide F.O.V.
- 300 WCD ->1200 PMT
- At 4,100 m.a.s.l.
- Cosmic ray energy: 1 TeV to 1 PeV

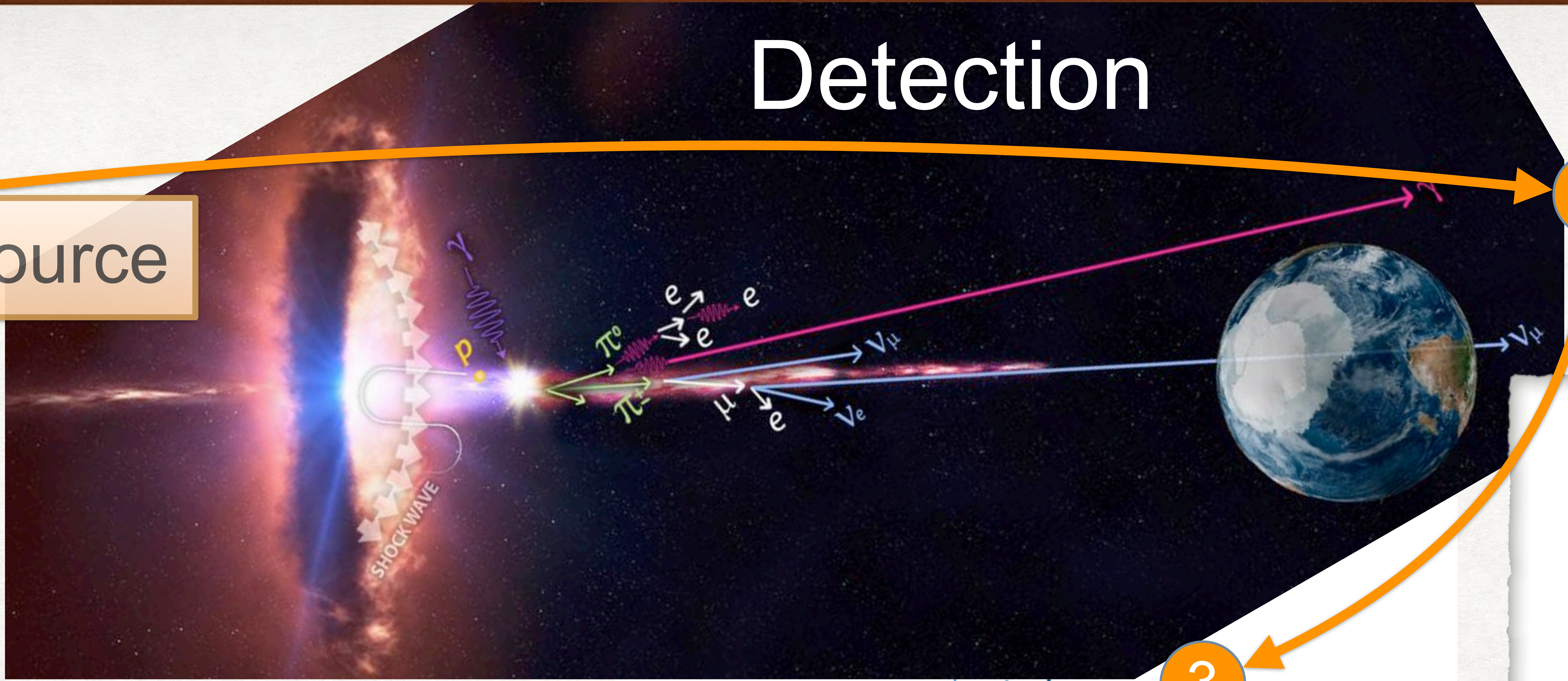
Water Cherenkov Detector (WCD)



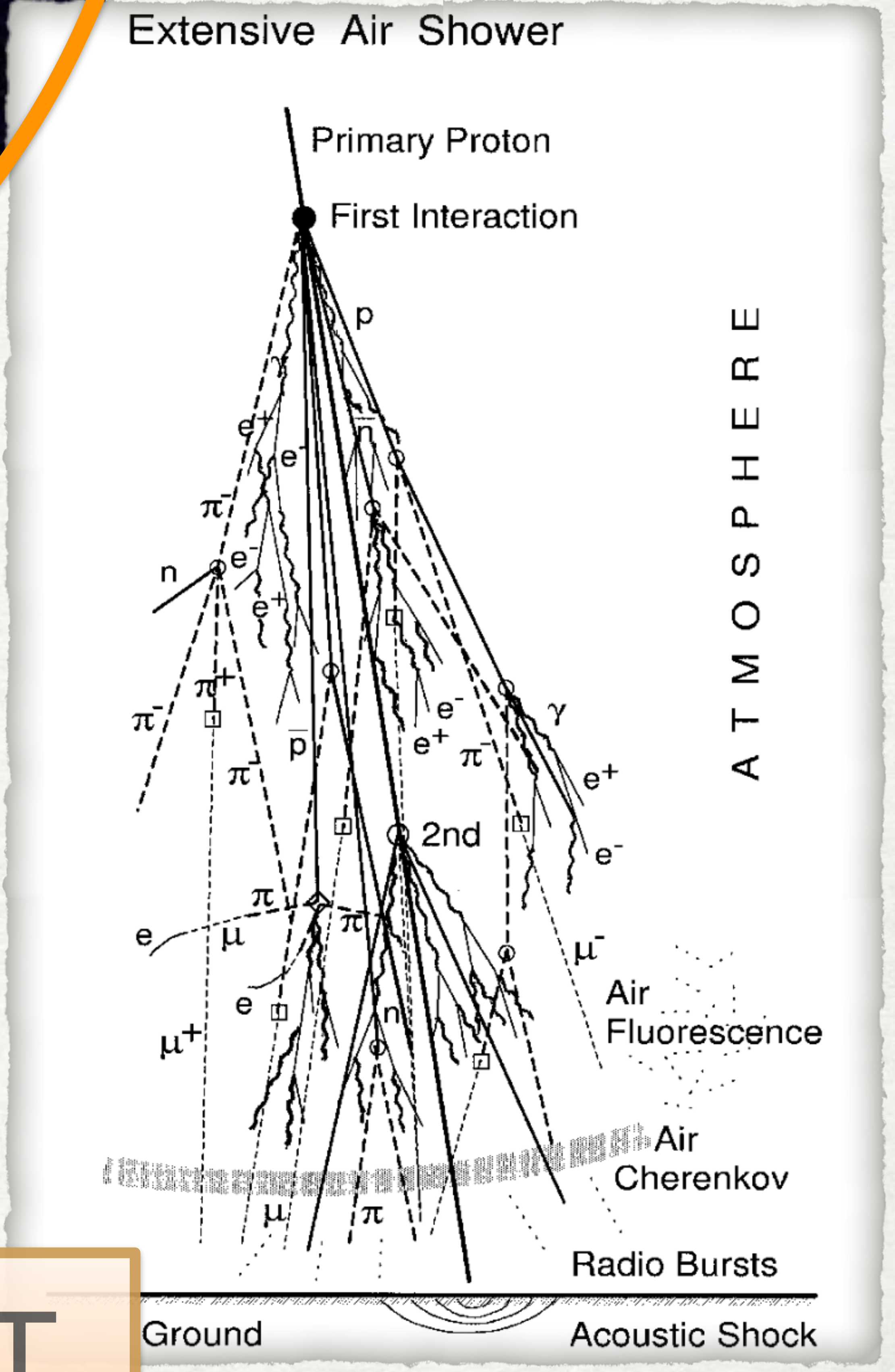
Photomultiplier tube (PMT)

Detection

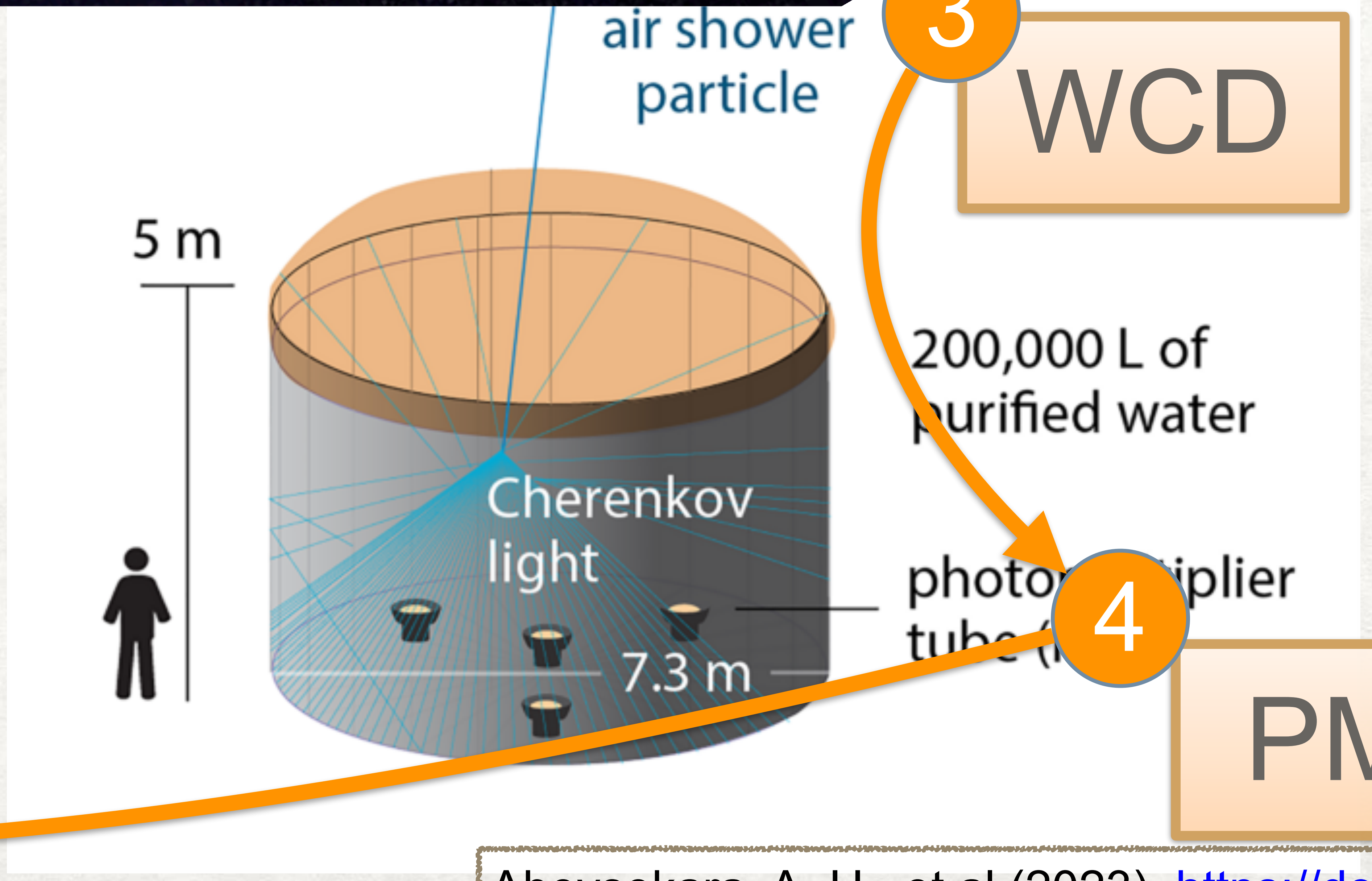
1 Source



2 Atmosphere



3 WCD



4 PMT

5 ICN/UMD cluster

Abeysekara, A. U., et al (2023), https://doi.org/10.1016/j.nima.2023.168253

Detection

Reconstruction

1 Source

2 Atmosphere

3 WCD

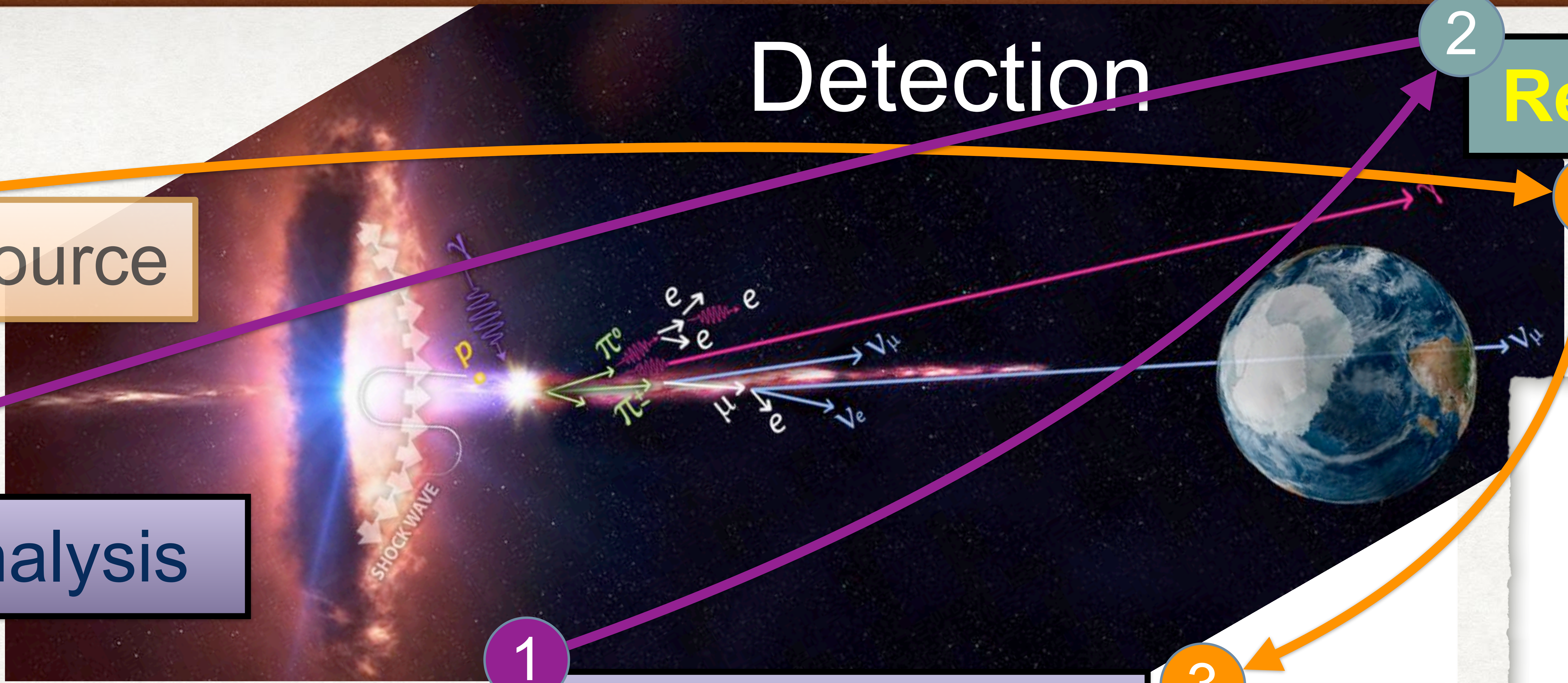
4 PMT

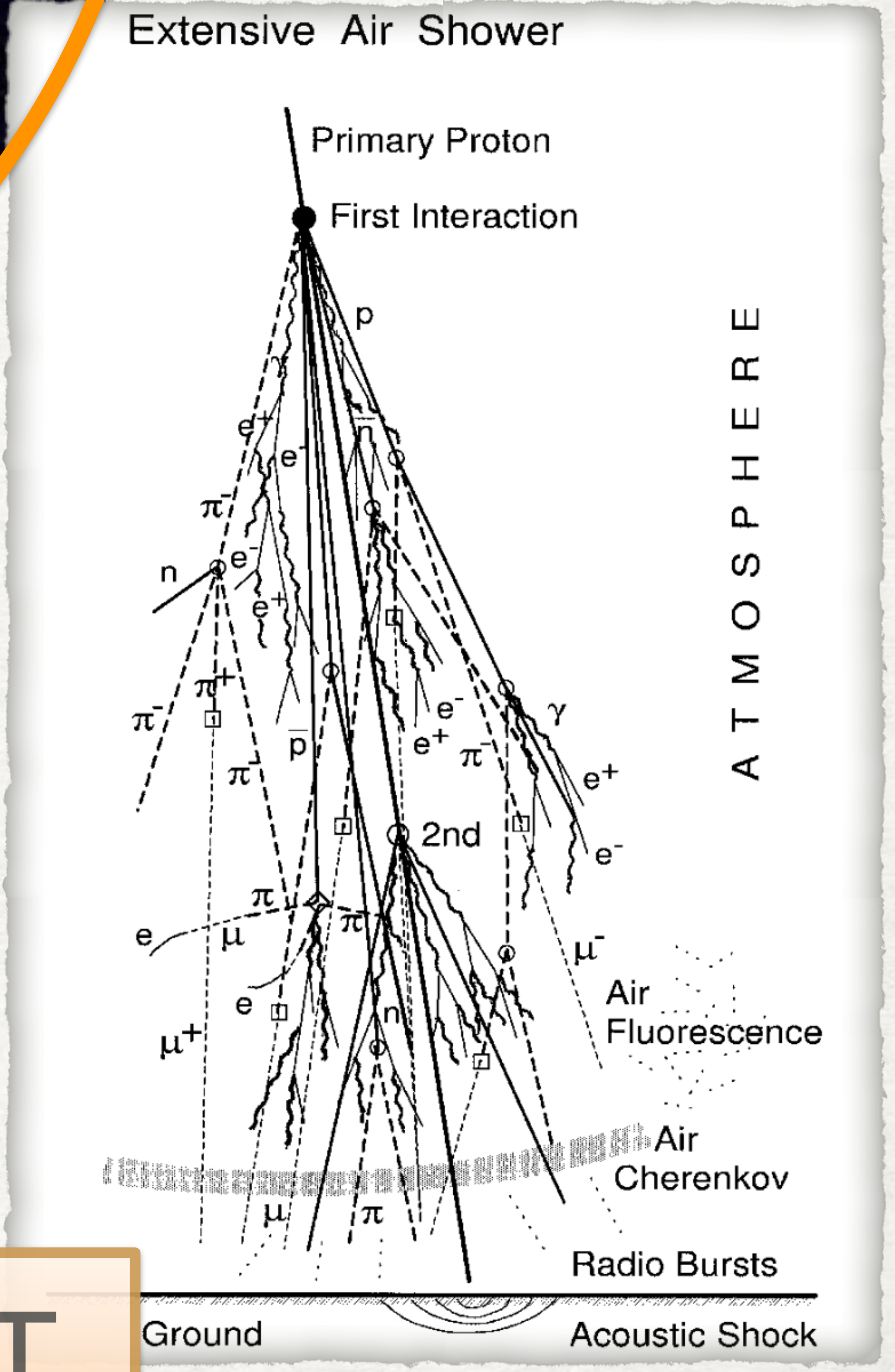
5 ICN/UMD cluster

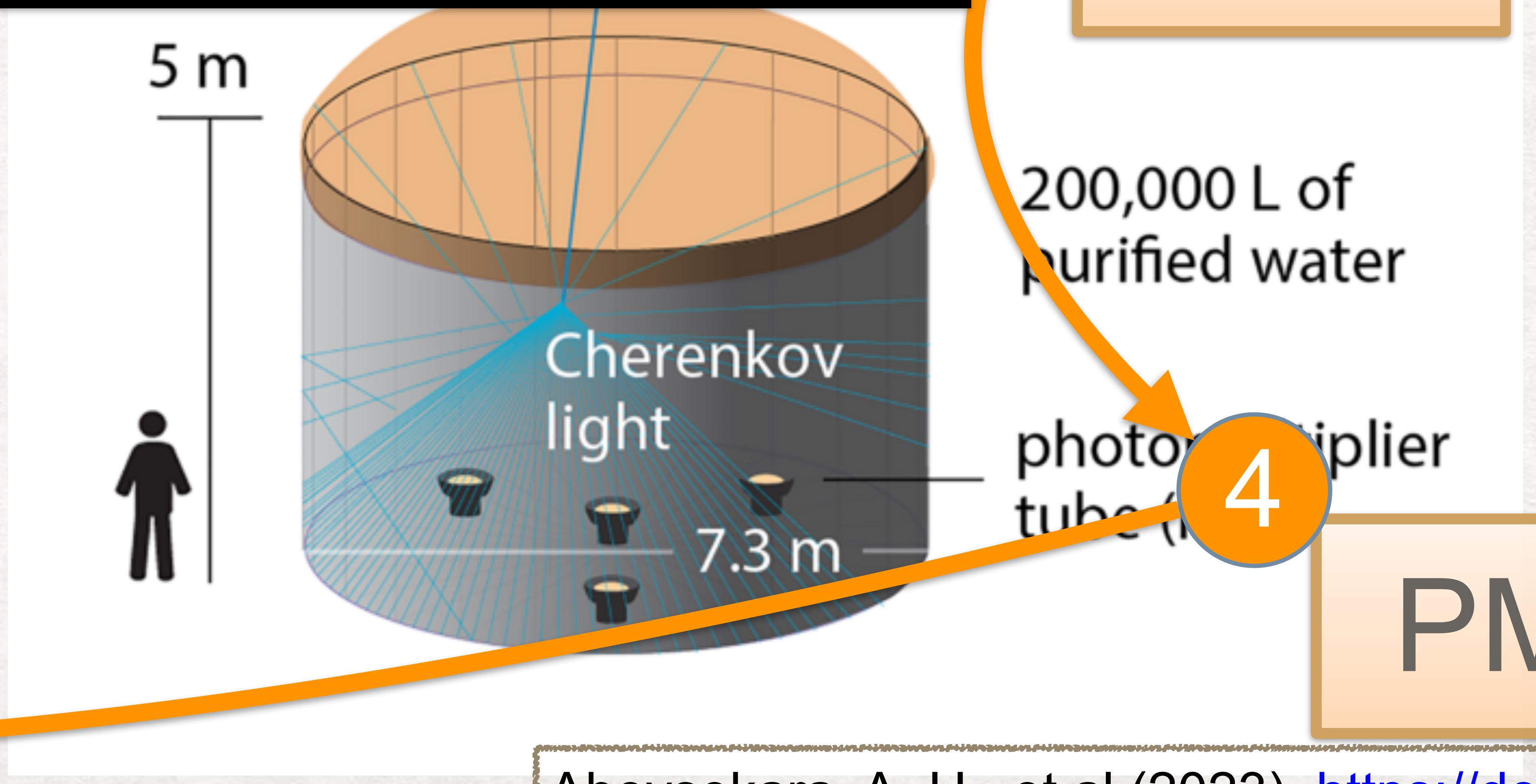
1 Calibration

2 Reconstruction

3 Analysis

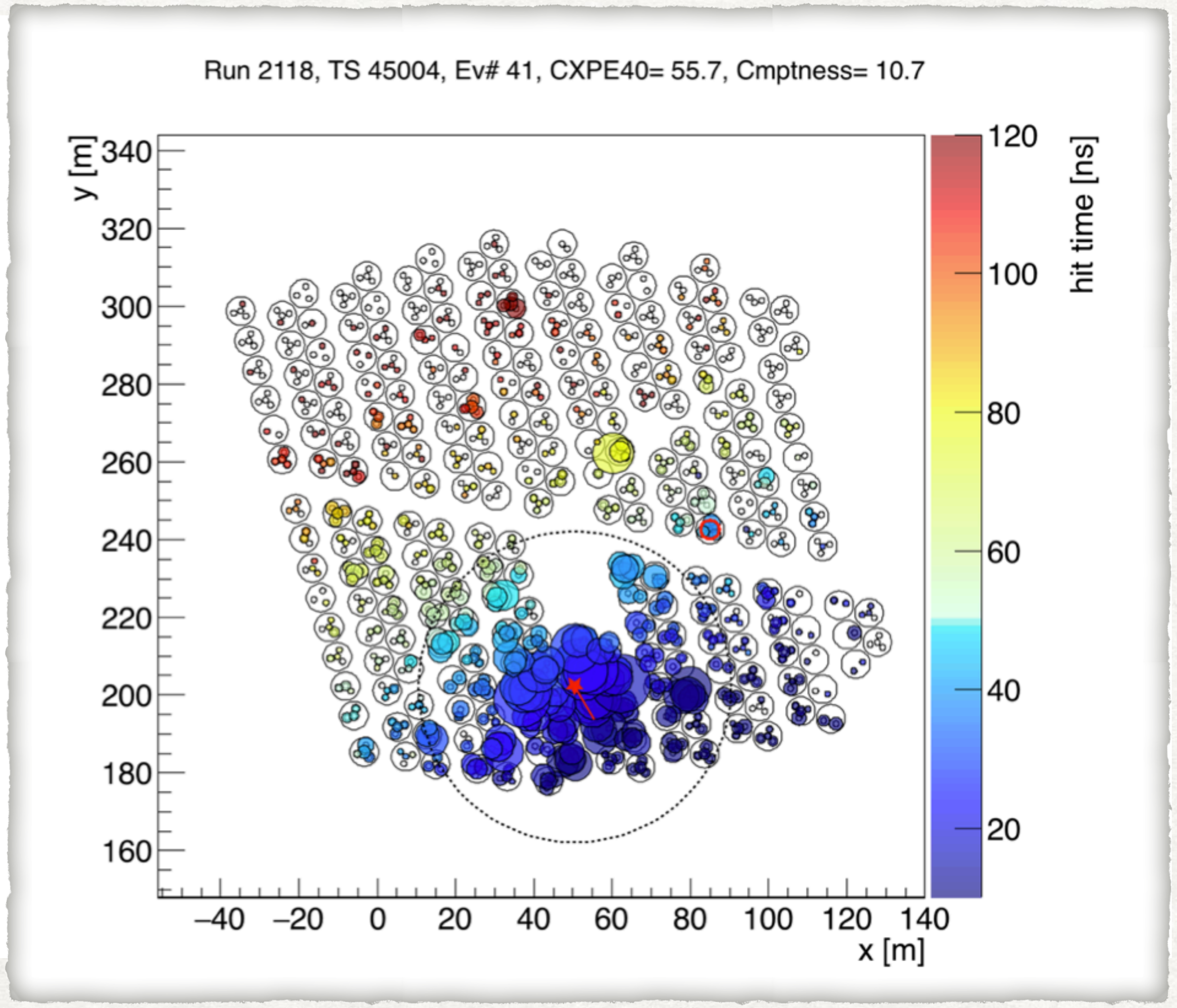






Abeysekara, A. U., et al (2023), https://doi.org/10.1016/j.nima.2023.168253

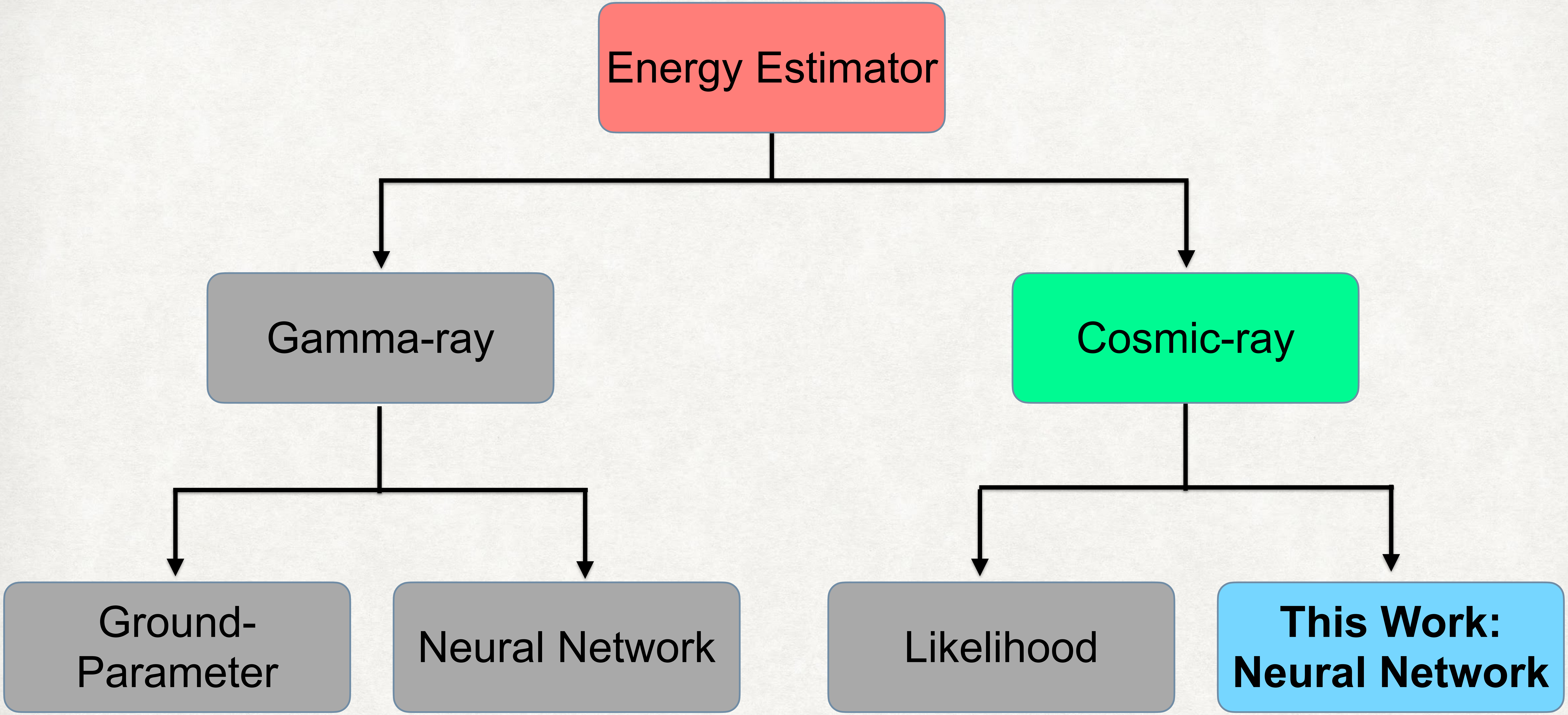
Event simulation detected by HAWC



A. Timing information allows us determine where the particle comes.
B. Energy deposition in each PMT:
    - The shower core.
    - Gamma or Hadron?
    - **Primary particle energy**.

Pretz, J. (2016), https://doi.org/10.22323/1.236.0025

Abeysekara, A. U., et al (2023), https://doi.org/10.1016/j.nima.2023.168253



Abeysekara, A. U., et al (2019), https://doi.org/10.3847/1538-4357/ab2f7d

Alfaro, R., et al (2017), https://doi.org/10.1103/PhysRevD.96.122001

Data used

Eight species simulated: Carbon, Helium, Iron, Magnesium, Neon, Oxygen, Proton, and Silicon

Hadronic models employed in **CORSIKA: FLUKA & QGJET-II-04**

Spectrum: **Power law** with a spectral index of **-2**, covering an energy range from **5 GeV to 2 PeV**.

Albert, A., et al (2024), https://doi.org/10.3847/1538-4357/ad5f2d

Events were **selected** based with the following criteria:

- **Successful reconstruction**.
- Zenith angle between **0º and 35º**.
- At least **20%** of the HAWC array was involved.

Building the model

The model was built using:

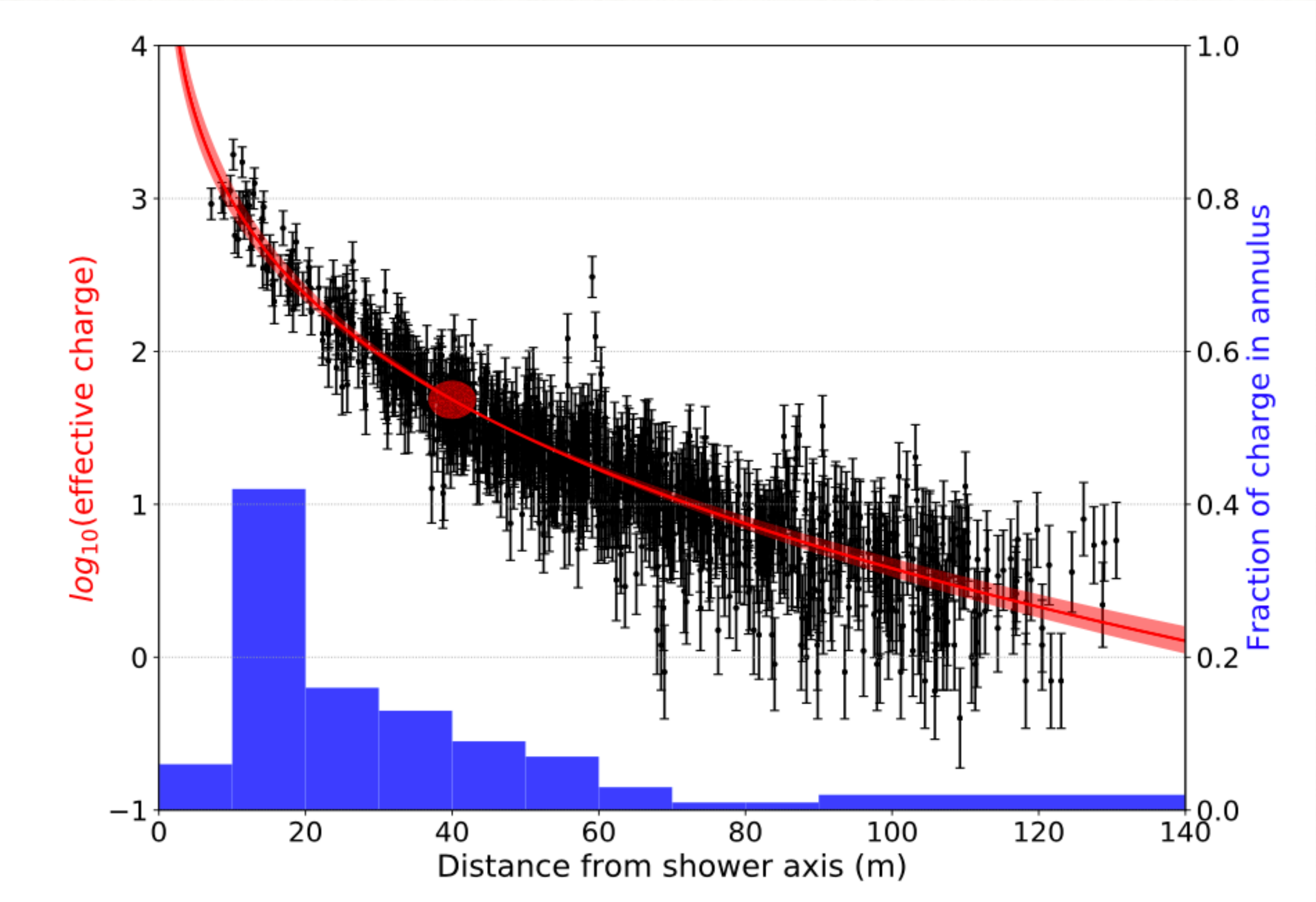
- Learning model: **Supervised**
- Training data: **proton**,
- Test data: **eight** particles
- Package: **TMVA of ROOT**
- Architecture: **14:10:10:1**
- Three models operated: low, medium and high energy

Alvarado, D. A., et al (2023), https://doi.org/10.22323/1.444.0402

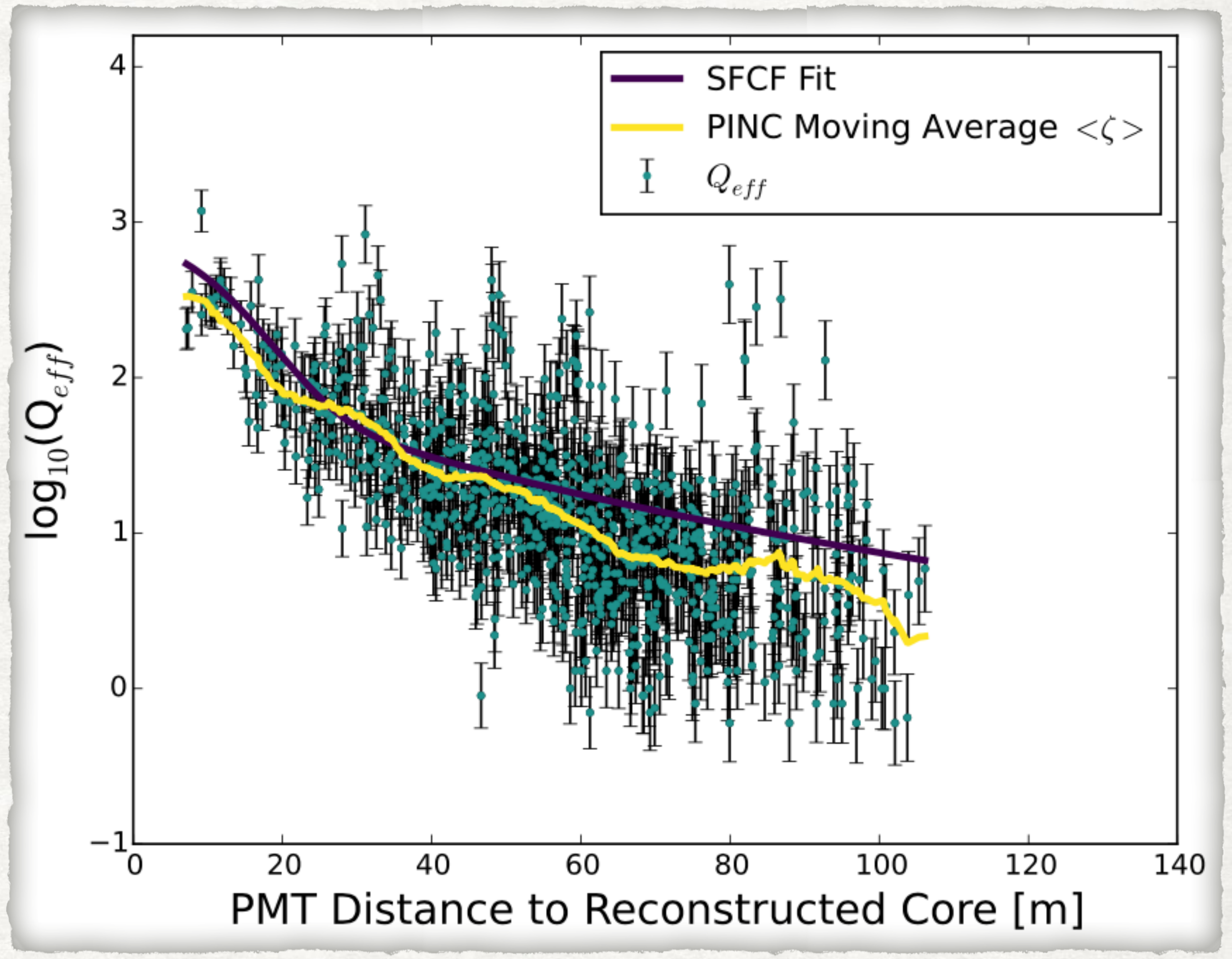
The model was built using:

- Learning model: **Supervised**
- Training data: **eight** particles,
- Test data: **eight** particles
- Test data: **eight** particles Package: **TensorFlow**
- Architecture: **36:256:128:64:32:1**
- One model was trained

Information on the input

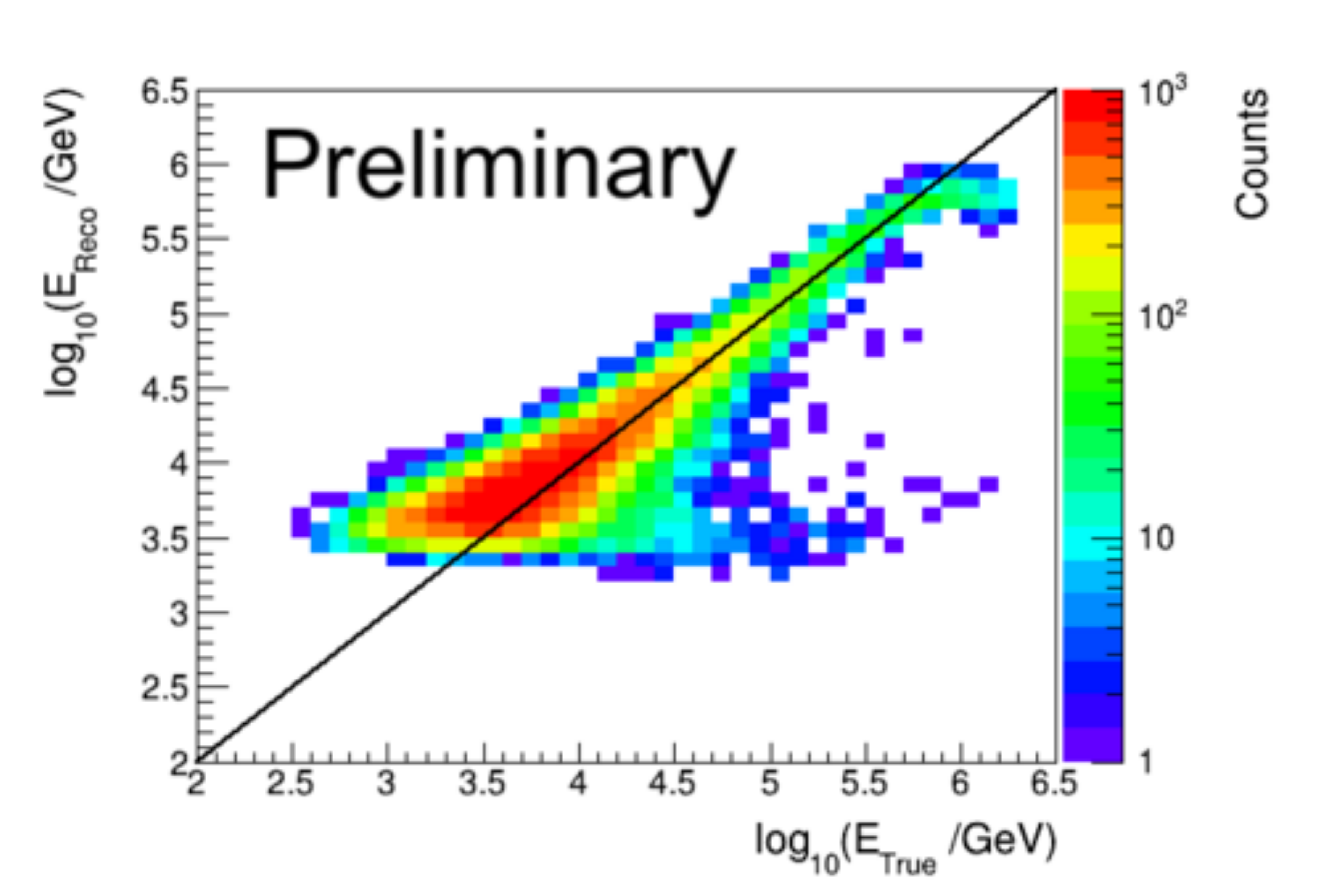
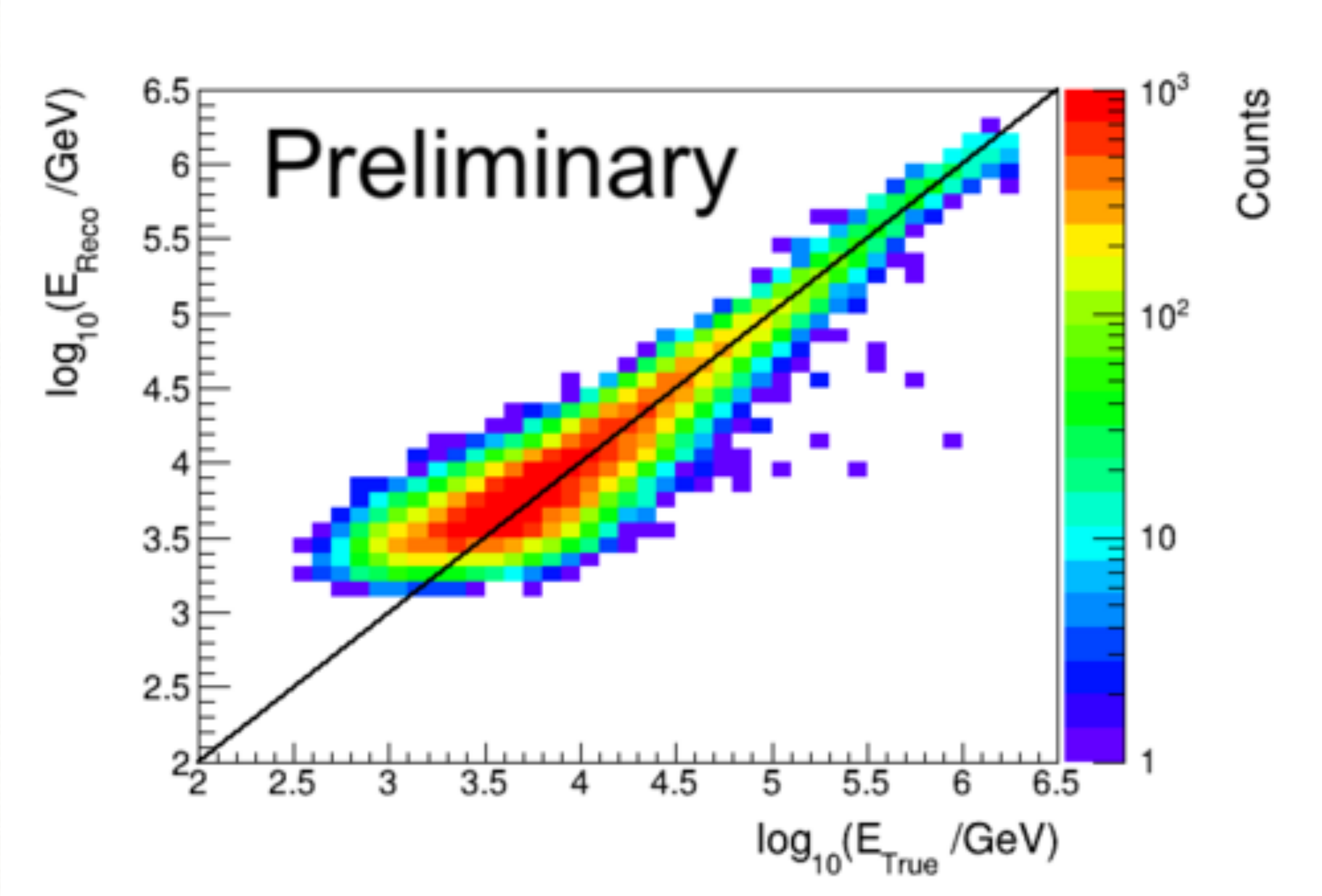
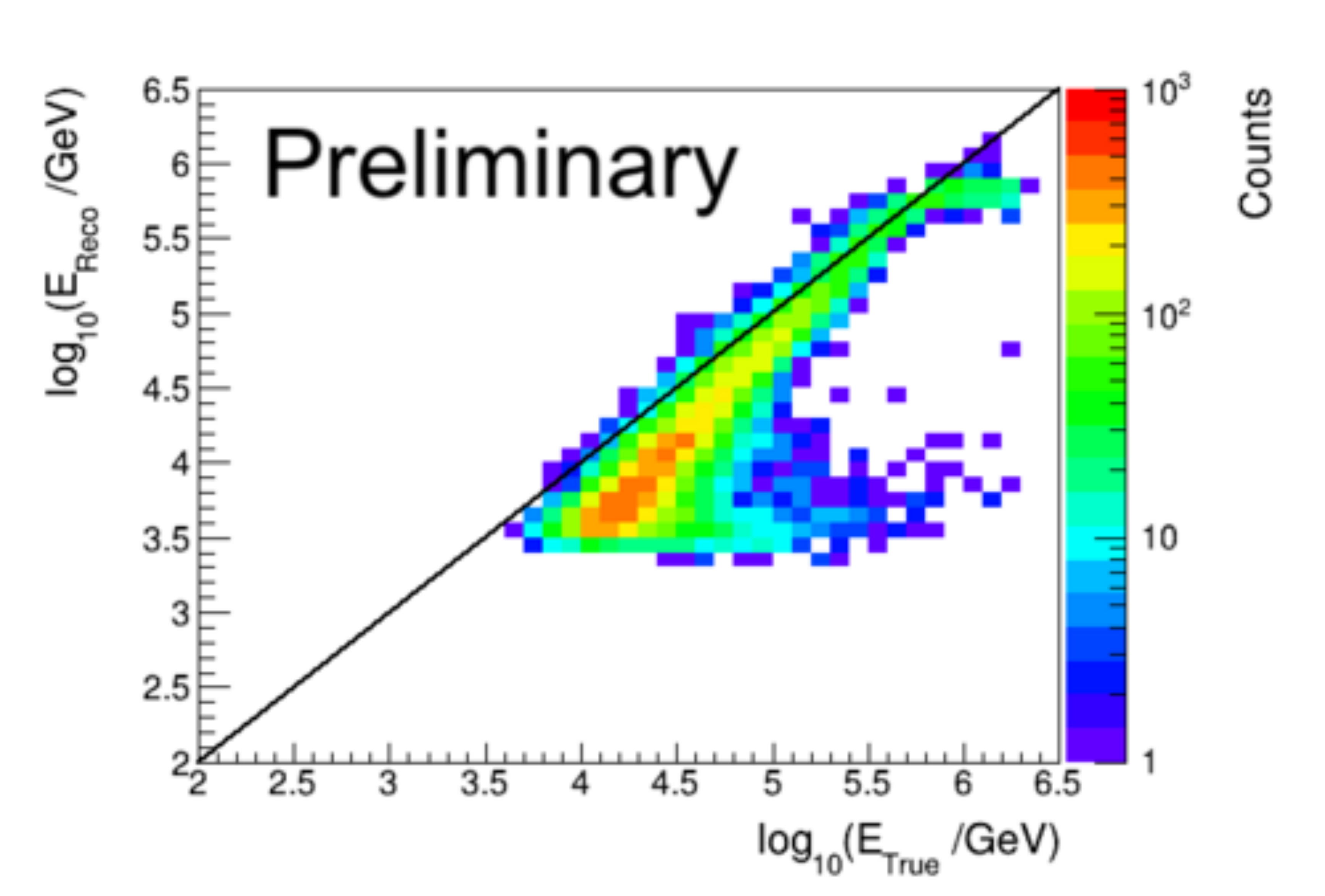
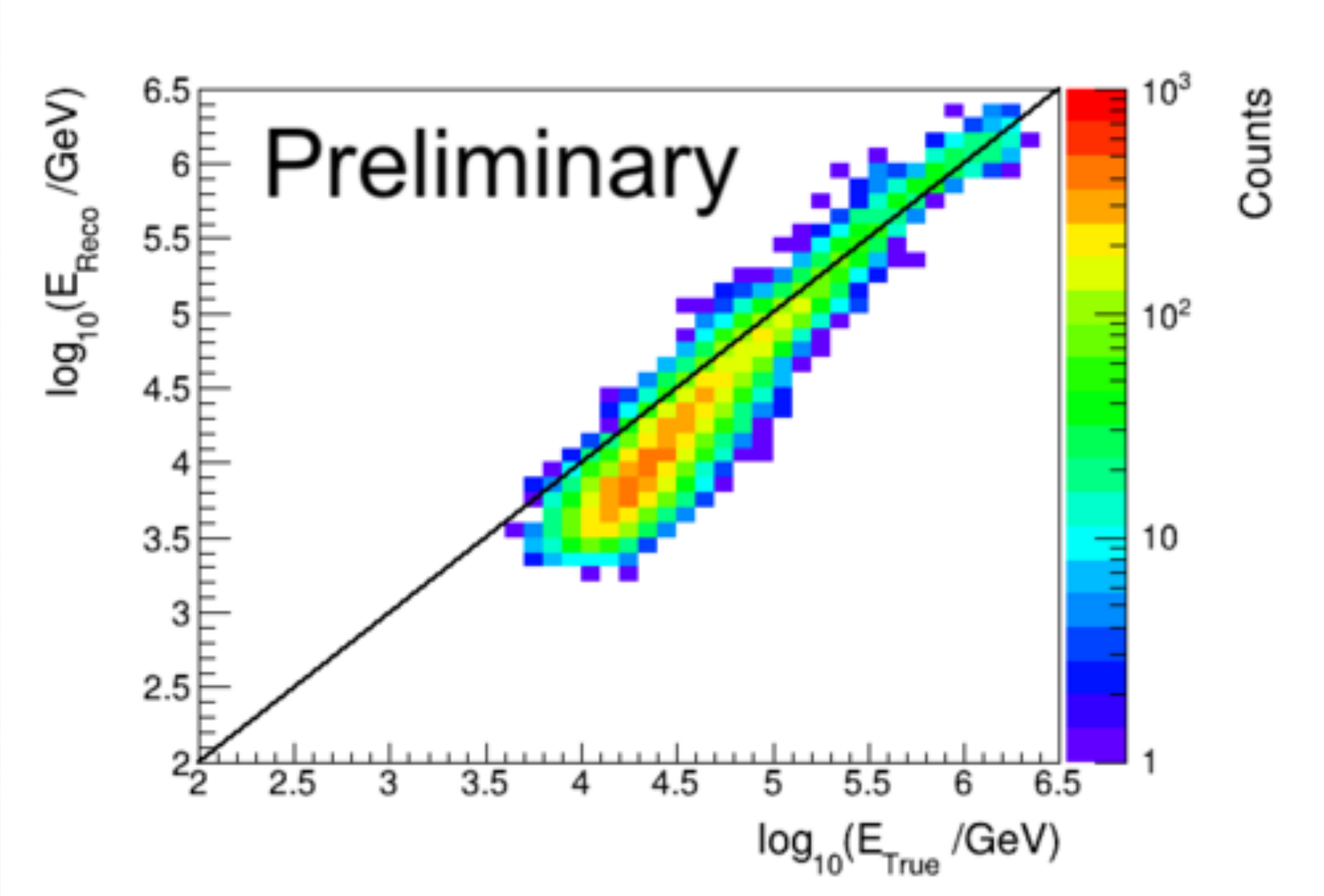


Abeysekara, A. U., et al (2019), https://doi.org/10.3847/1538-4357/ab2f7d



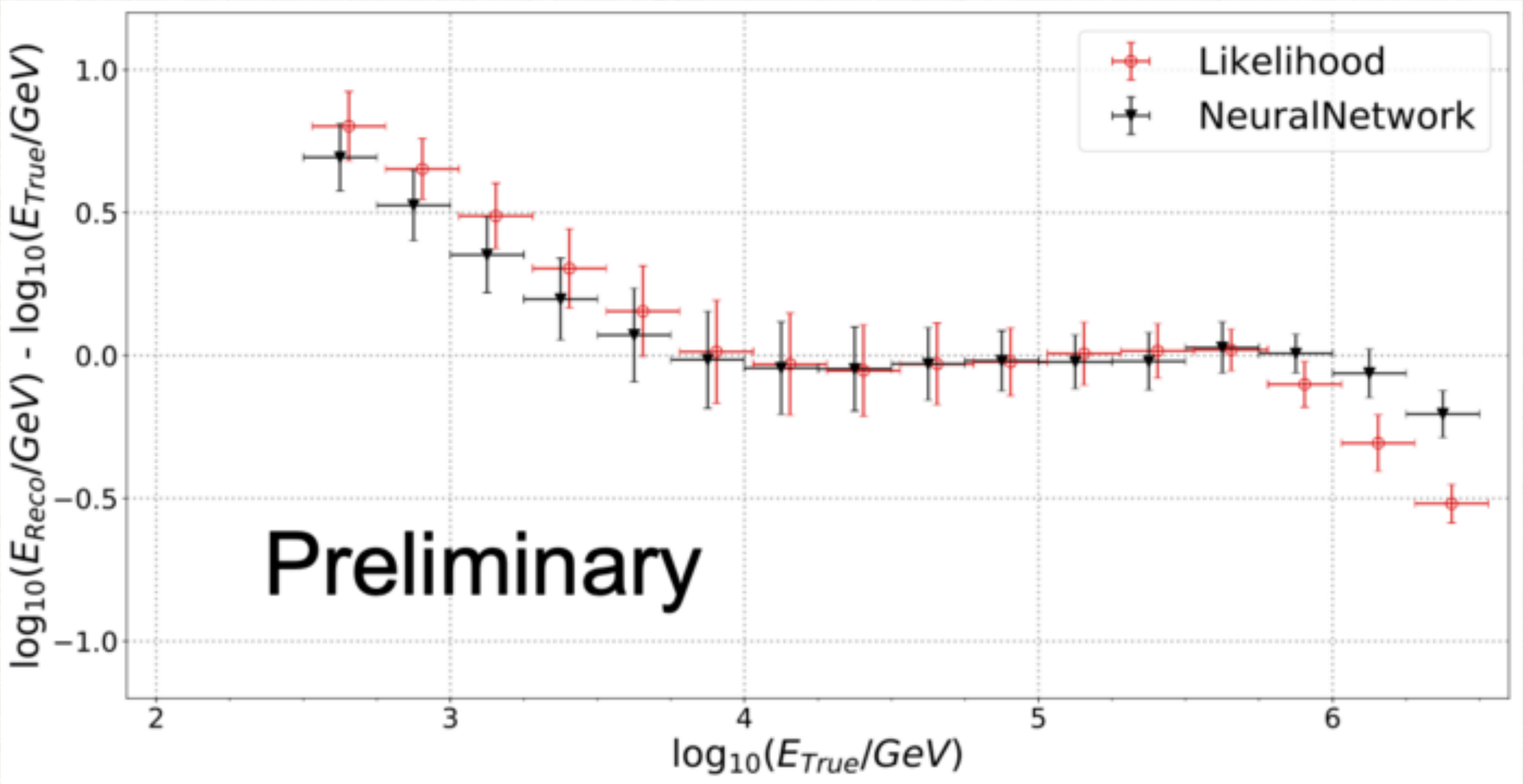
Cosmic Ray

Abeysekara, A. U., et al (2017), https://doi.org/10.3847/1538-4357/aa7555

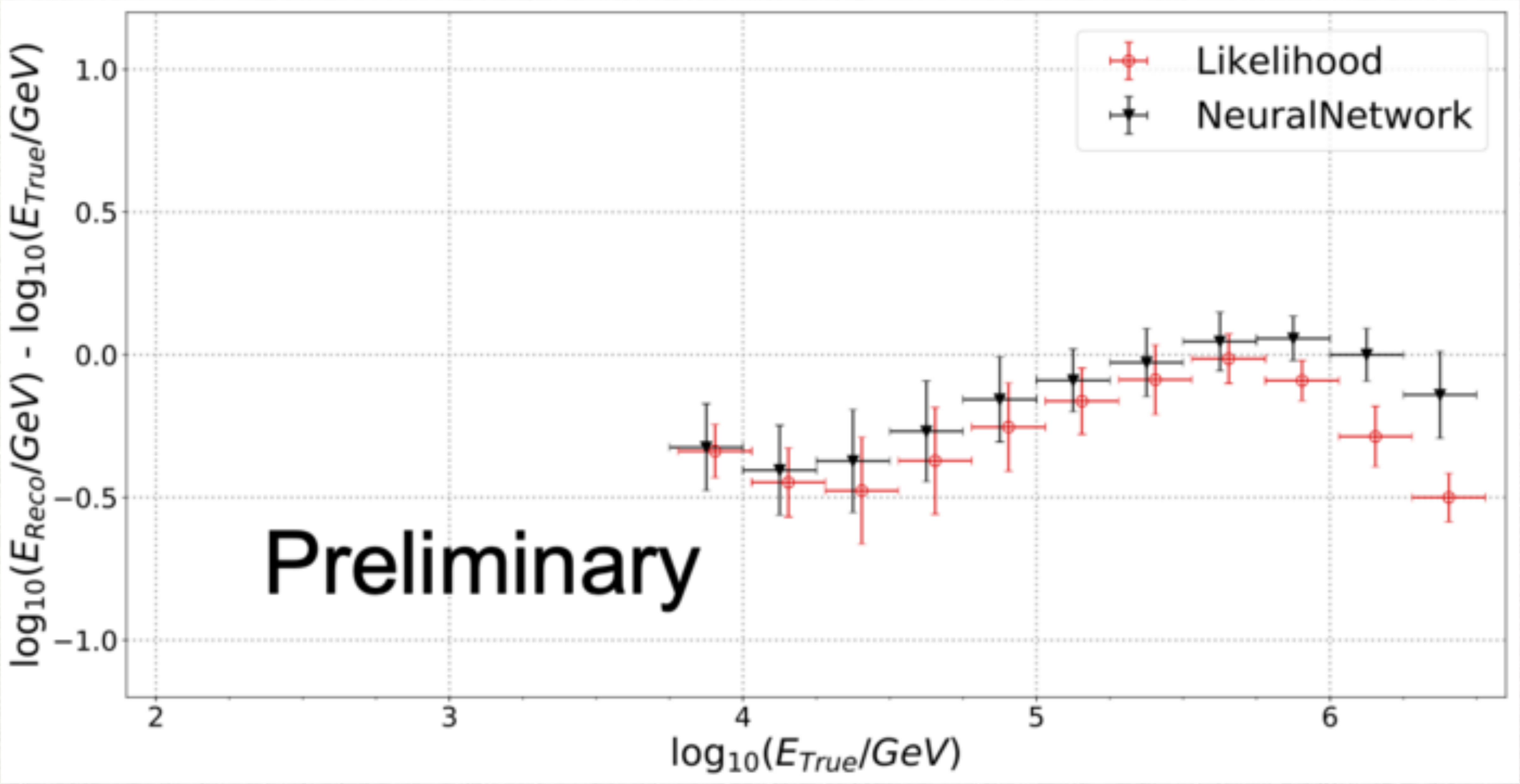
Likelihood	Neural Network trained in TMVA	
 	 	Proton
 	 	Iron

Alvarado, D. A., et al (2023), https://doi.org/10.22323/1.444.0402

Proton

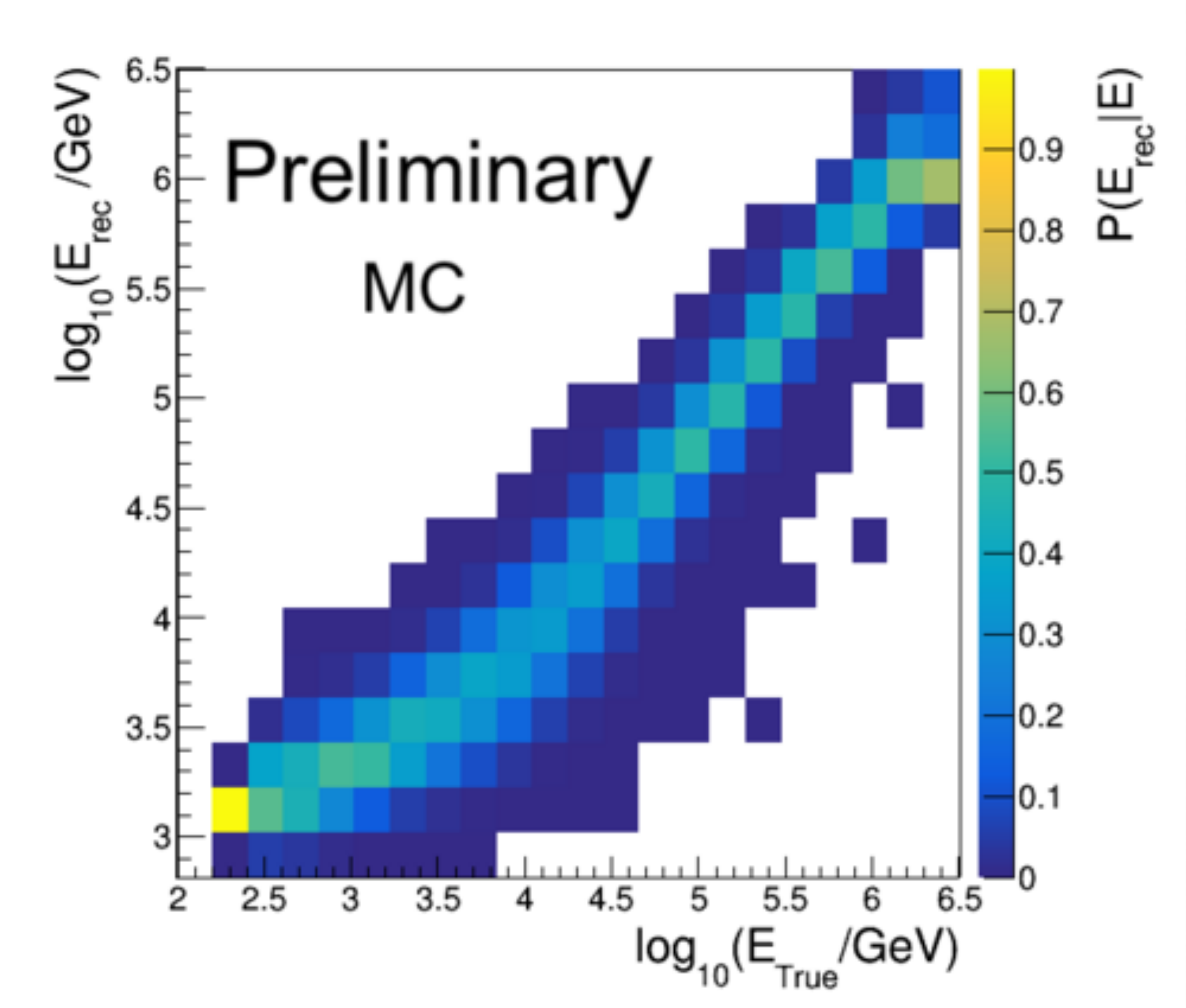


Iron

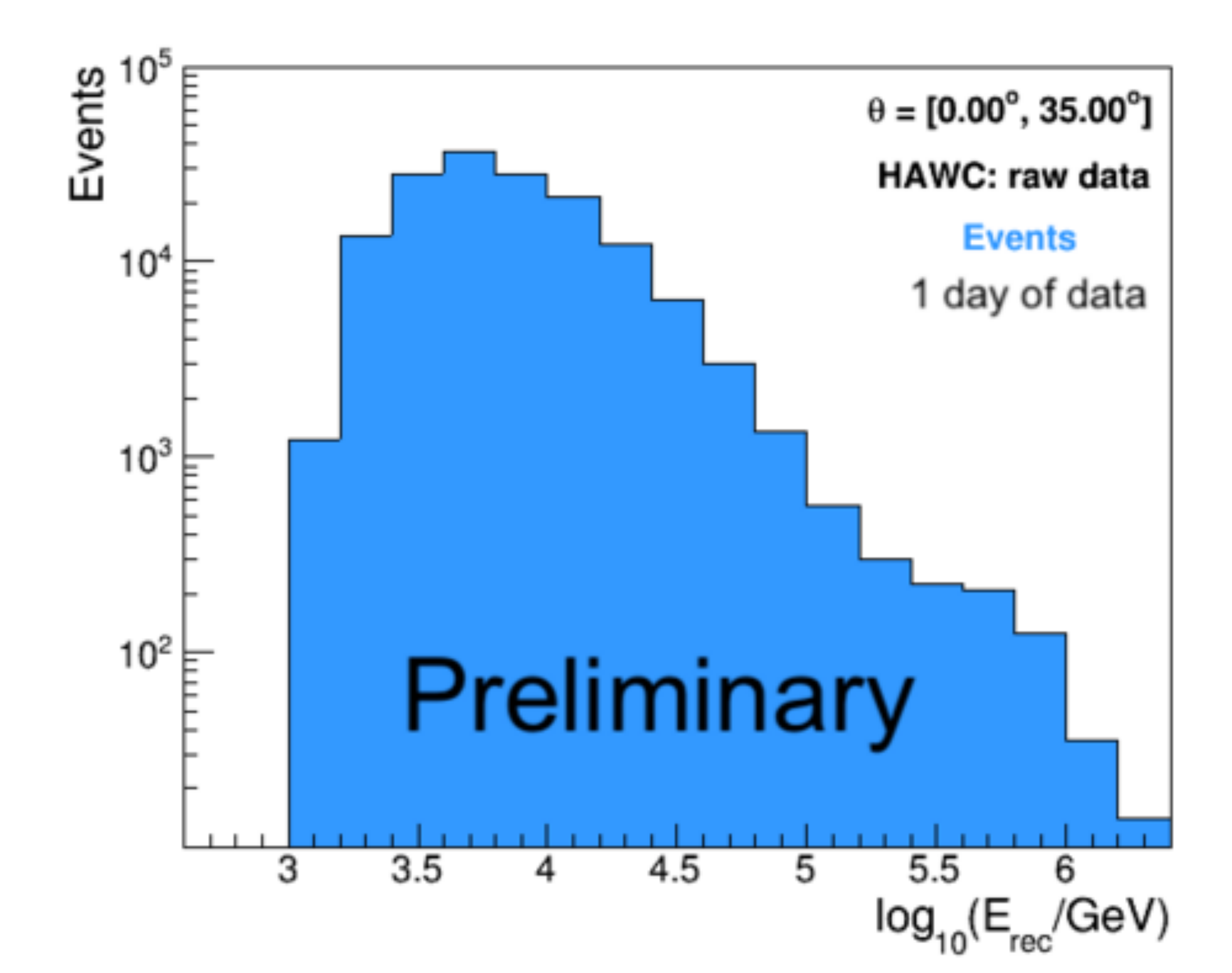


Alvarado, D. A., et al (2023), https://doi.org/10.22323/1.444.0402

Neural Network trained in TensorFlow

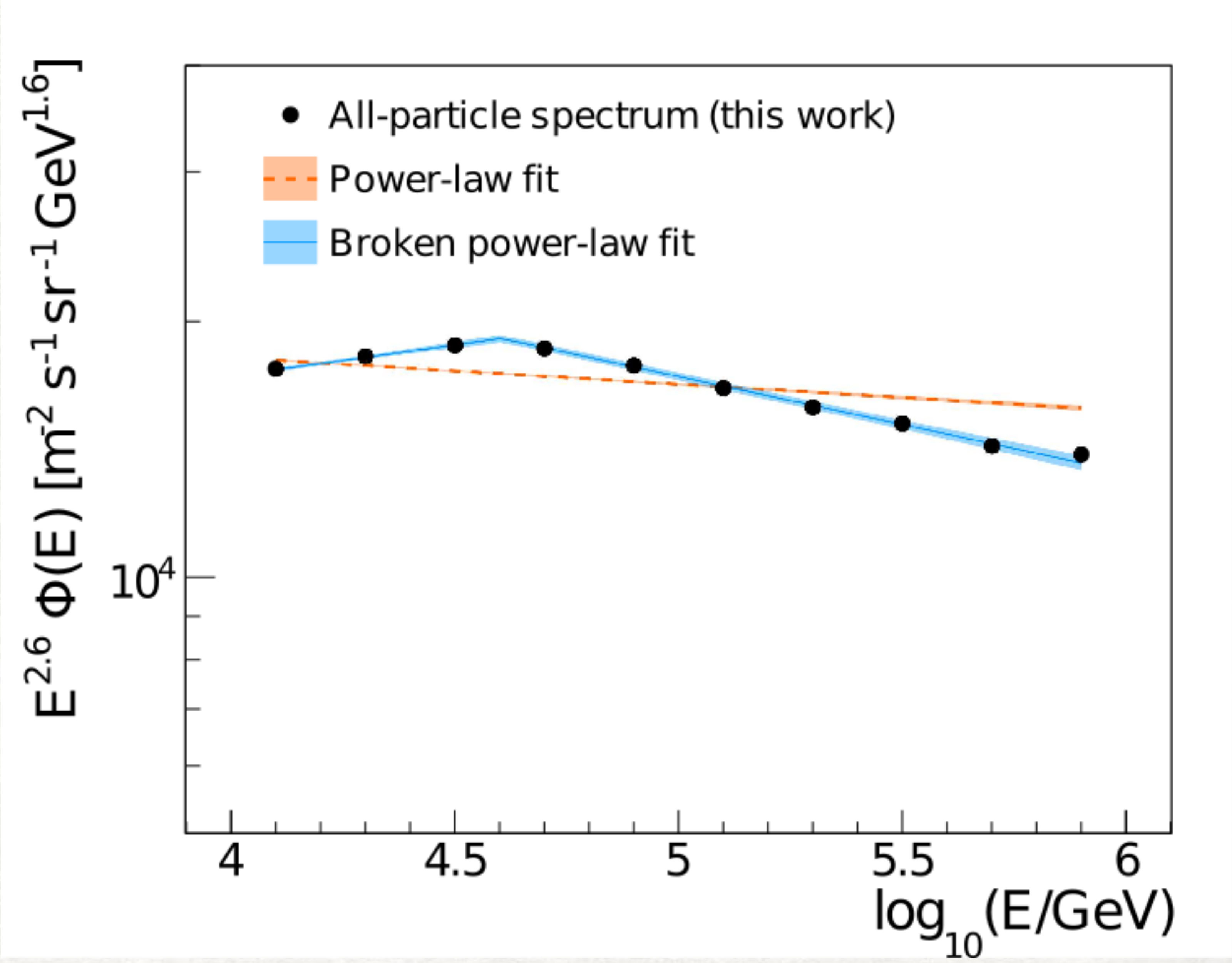


Response Matrix



Output Distribution using one day of HAWC data

The next step is to obtain the spectrum and compare it with the latest publication.



Alfaro, R., et al (2024), https://doi.org/10.1016/j.astropartphys.2024.103077

Summary

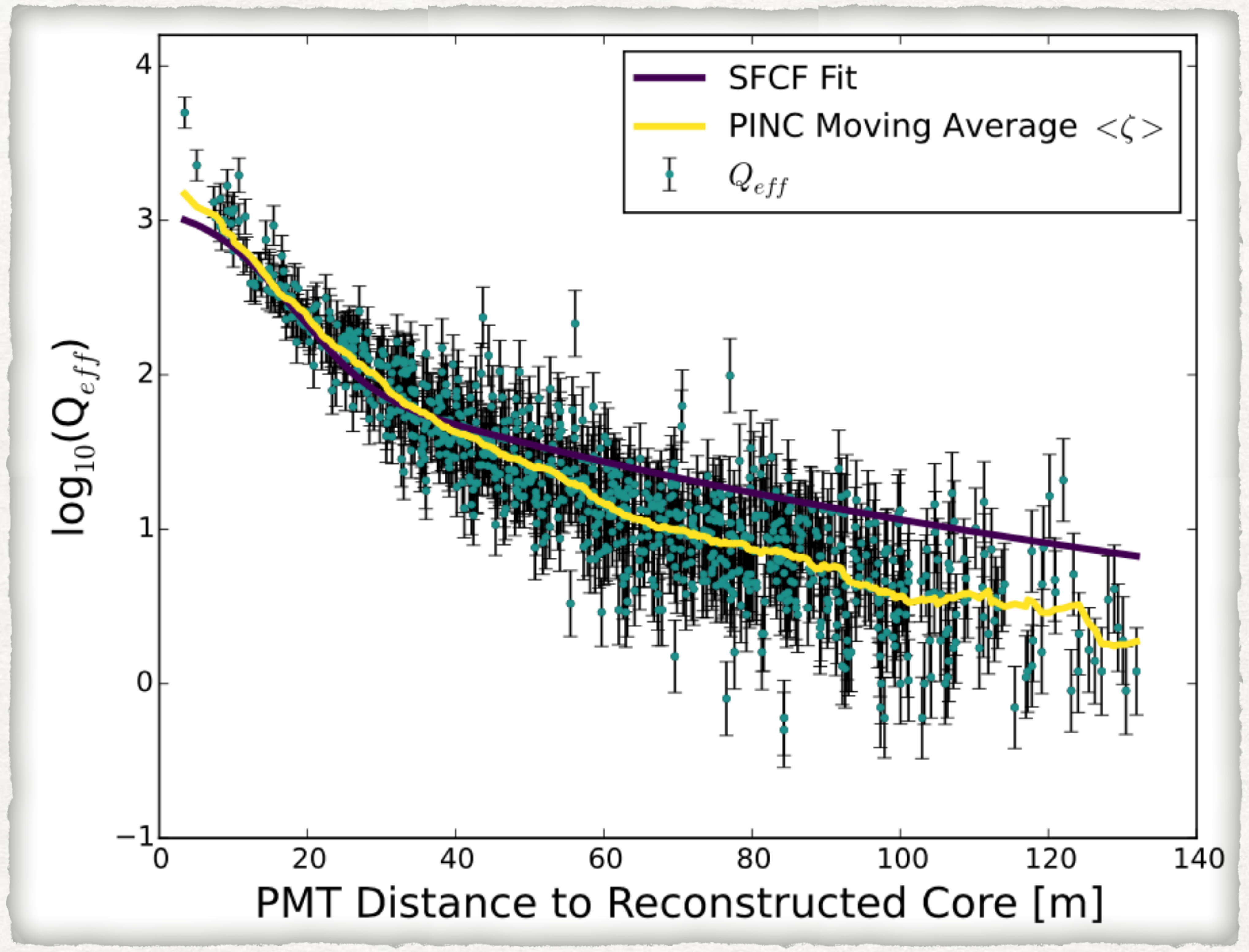
- The HAWC experiment is a state-of-the-art cosmic ray detector operating in the energy range of 1 TeV to 1 PeV.
- A machine learning model was developed to predict the energy of cosmic rays, improving the reconstruction according with MC simulation.
- This model was applied to one day of HAWC data to evaluate its performance with real data. We will be obtain the spectrum to validate it against the recently published spectrum.

Gracias! Thank you! Grazie!

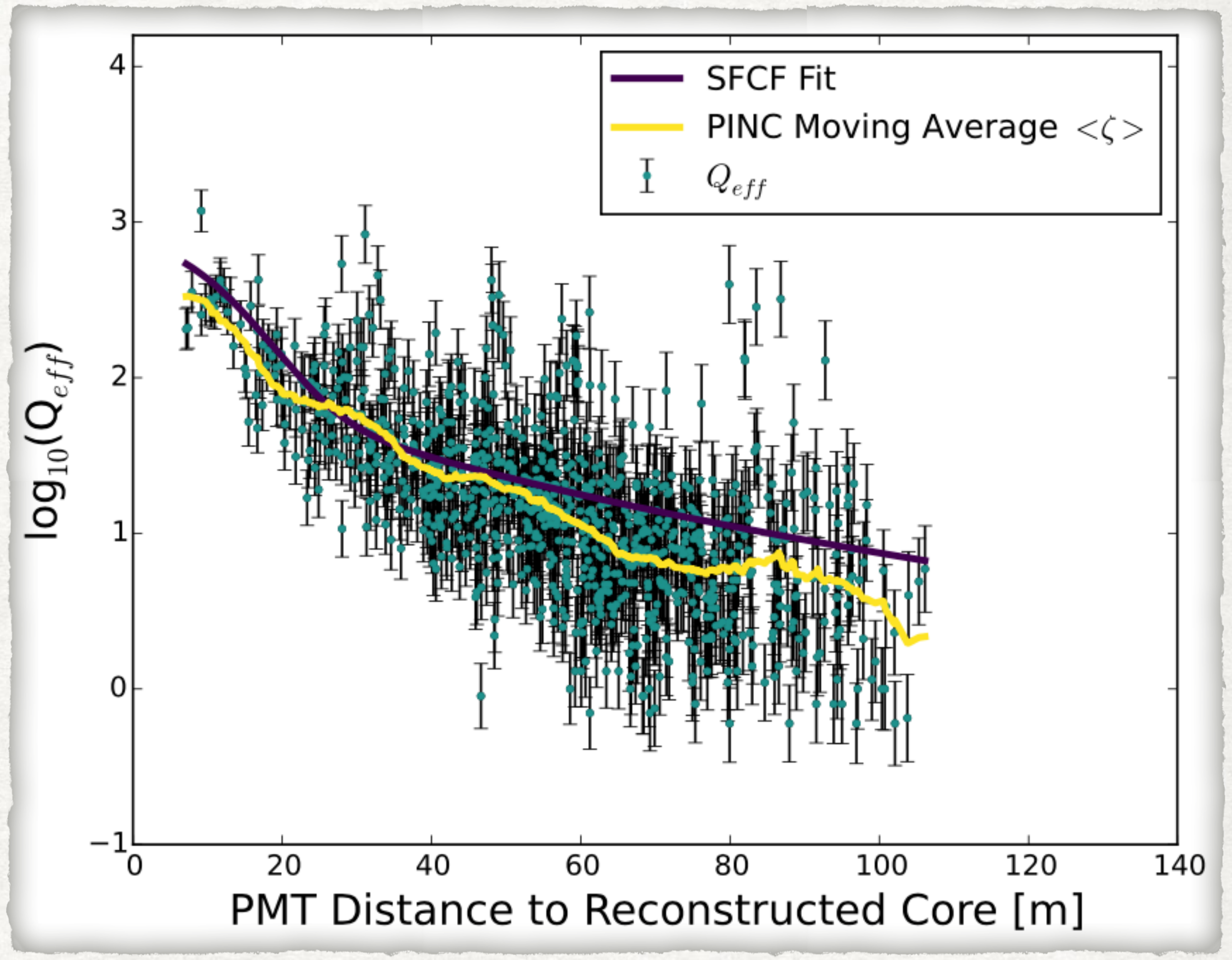
Backslides

Input parameters

- **LIC**
- disMax
- LDFChi2
- LDFAmp
- **PINC**



Photon from Crab Nebula



Cosmic Ray

Abeysekara, A. U., et al (2017), https://doi.org/10.3847/1538-4357/aa7555

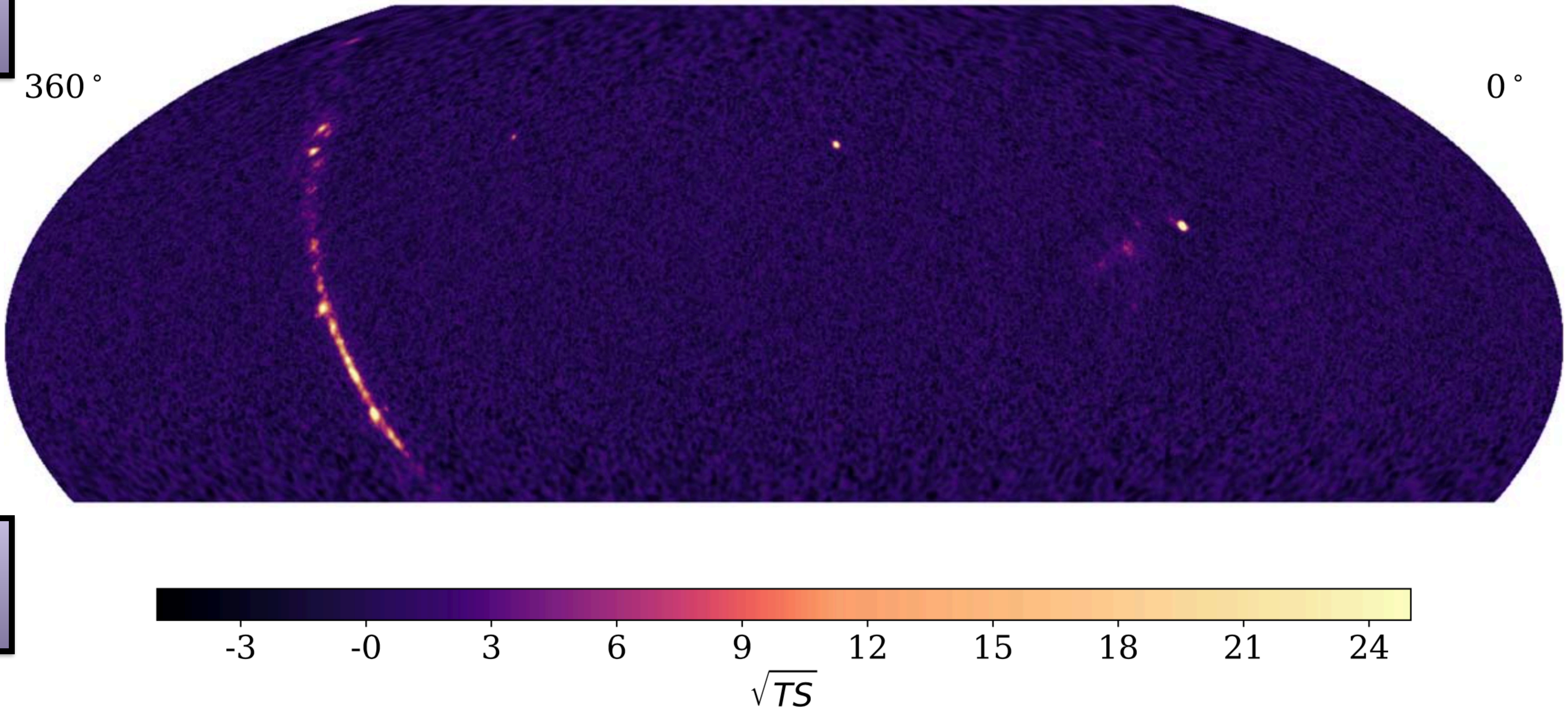
$$LIC = \log_{10} \frac{1}{compactness} = \log_{10} \frac{CxPE_{40}}{nHit}$$

$$NKG = A\ \rho^{s-3}\ (1+\rho)^{s-4.5}$$

$$PINC = \frac{1}{N} \sum_{i=0}^{N} \frac{[\log_{10}(q_i) - \langle \log_{10}(q_i) \rangle]^2}{\sigma^2}$$

The main reconstruction parameters

Energy



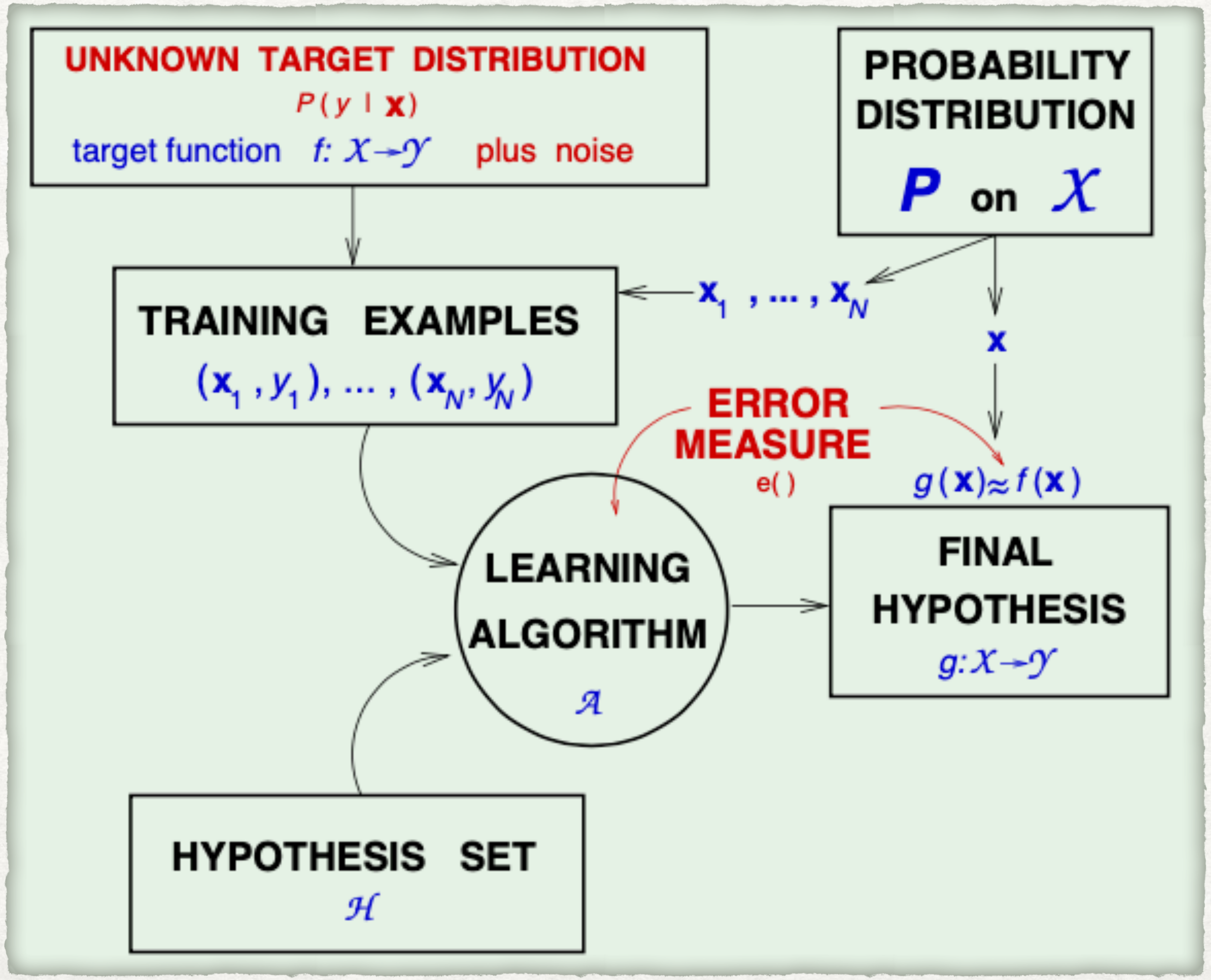
Position

Figure 1. All-sky significance map in celestial coordinates, assuming a point-source hypothesis. The bright band on the left is part of the Galactic plane (see Figures 4–7), and the bright region on the right is the Galactic anticenter region containing the Crab Nebula and the Geminga halo (see Figure 3). The two off-plane hotspots are the two TeV-bright blazars Mrk 421 (right) and Mrk 501 (left).

A. A. U. Abeysekara (2020) DOI:10.3847/1538-4357/abc2d8

Type of specie

Machine Learning Techniques (MLT)



Abu-Mostafa, Y. S., Magdon-Ismail, M.,, Lin, H.-T. (2012). *Learning From Data*. AMLBook