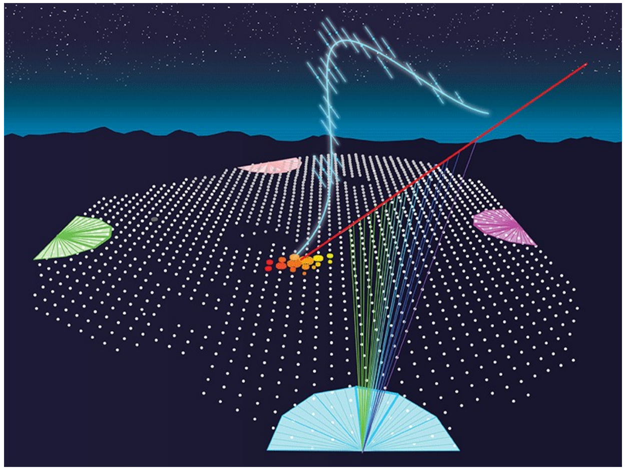
X_{max} Reconstruction with AugerPrime using Deep Learning

PIERRE AUGER OBSERVATORY

Sonja Schröder
Bergische Universität Wuppertal

Big Data Science in Astroparticle Research - Workshop
17-19 February 2020
-
18.02.20

Pierre Auger Observatory



Largest Observatory measuring Extended Air Showers (EAS) initiated by cosmic rays

Surface Detector (SD)

- 1660 Water Cherenkov Detectors (WCDs)
- Determine primary particle energy based on amount of Cherenkov light produced in each station
- Determine trajectory of incoming cosmic ray with differences in detection time of stations

Fluorescence Detector

- Fluorescence light traces longitudinal shower development

Radio Detector

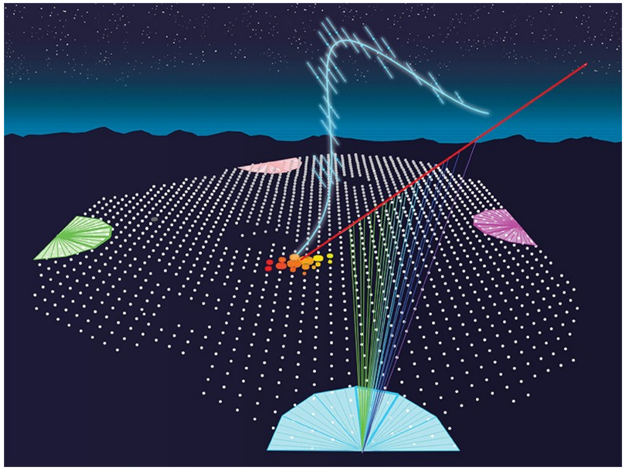
- Measures electromagnetic component of EAS in form of short radio pulses

Upgrade AugerPrime

Plastic scintillator detector on top of SD stations as part of upgrade 'AugerPrime'

Two detectors (WCD and SSD) with different responses to muonic and electromagnetic component of air showers

Improved event wise measurement of mass composition of primary particle using the 24/7 on-time of the SD



Goal of this analysis:

- Reconstruction of X_{max} with AugerPrime simulations
- Using only
  - WCD and SSD traces
  - Arrival time in stations
  - SD reconstructed Energy and Shower Axis (Zenith)

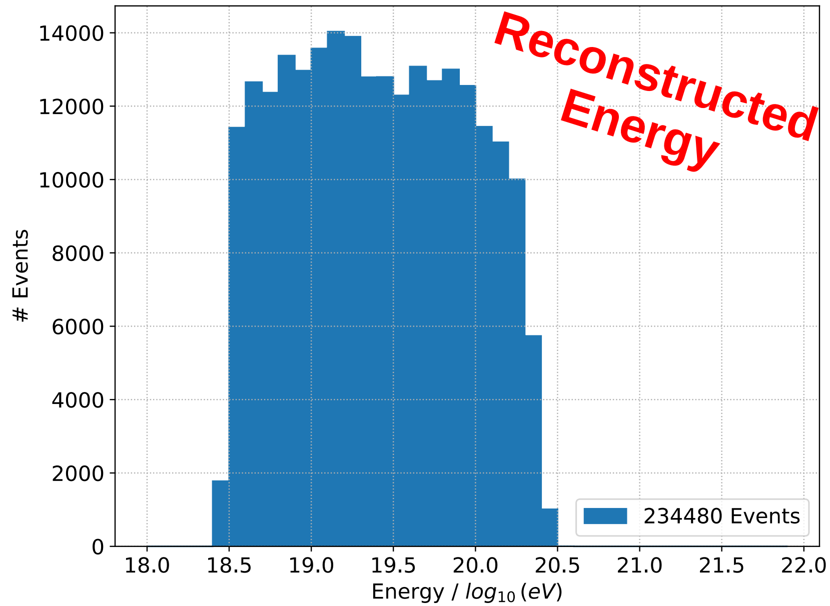
CORSIKA Simulations & Setup

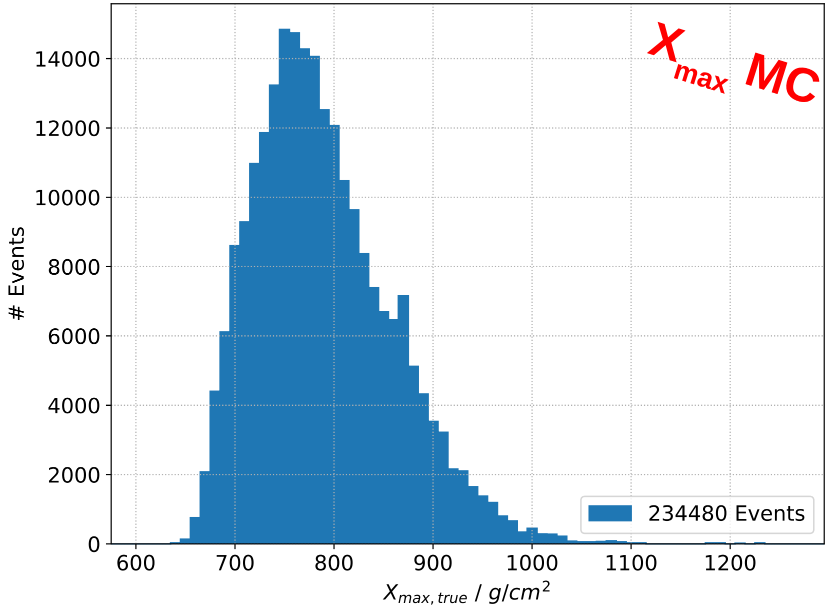
CORSIKA simulations

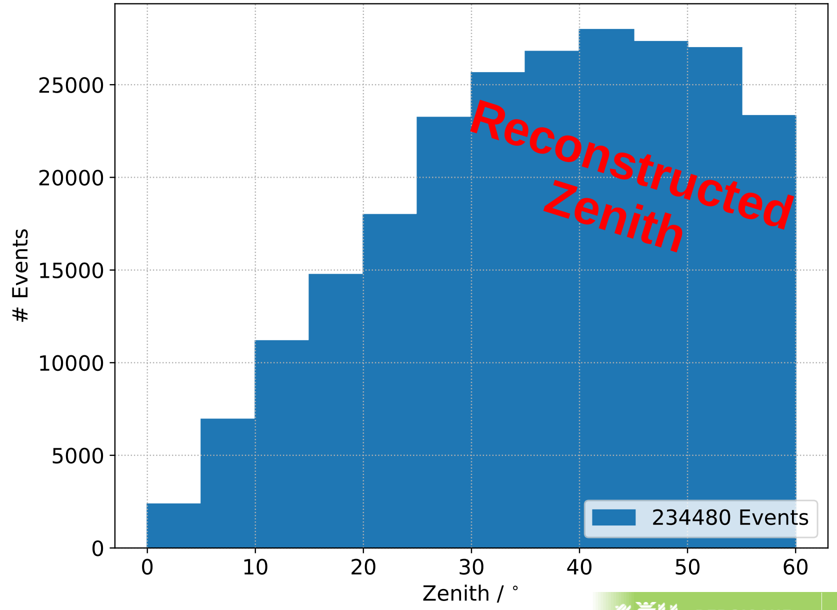
- 5000 proton, 5000 iron
- Energy range: 18.4 - 20.5 log(E) → flat in energy
- Zenith range: 0° - 60°

Offline Framework

- SD and SSD detector simulations
- Throwing each event onto different position in the array 30 times





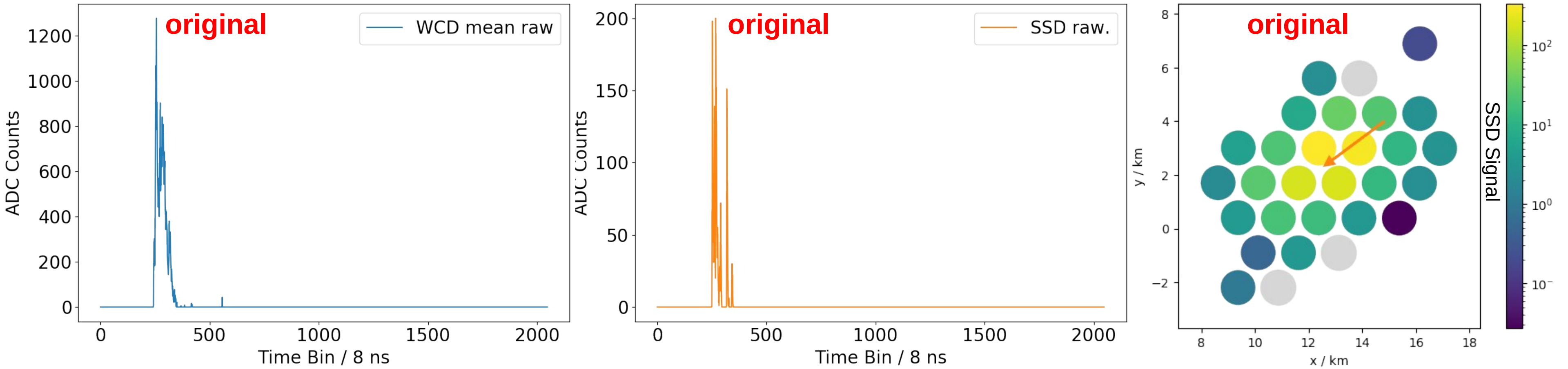


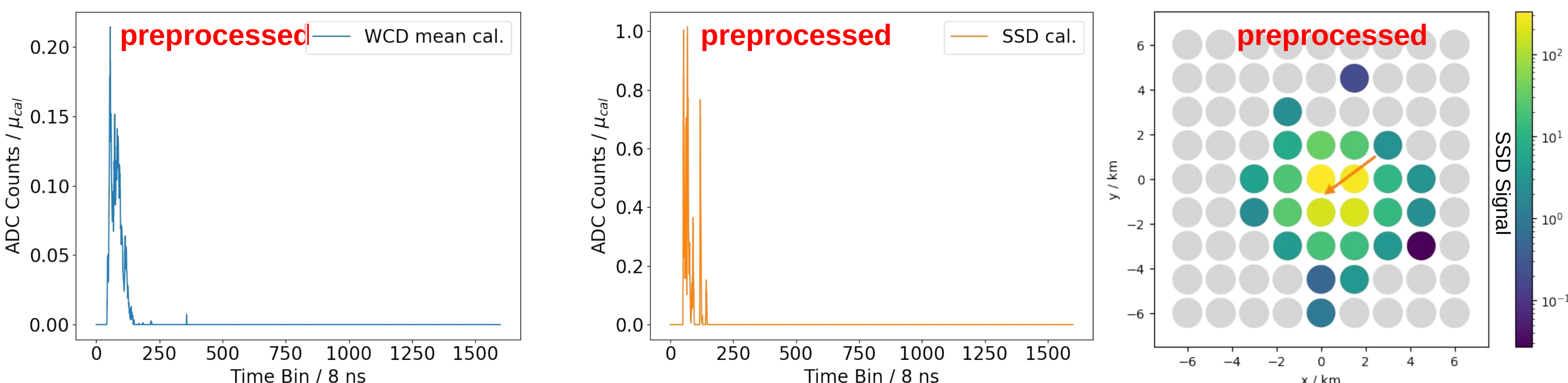
Preprocessing of Full Detector Simulations

Using raw SSD and WCD traces, remove empty bins from start before trigger

Norm traces by dividing by calibration constant (single muon charge)

Hexagonal to Cartesian grid, zero-pad with “empty” stations, centered around hottest station

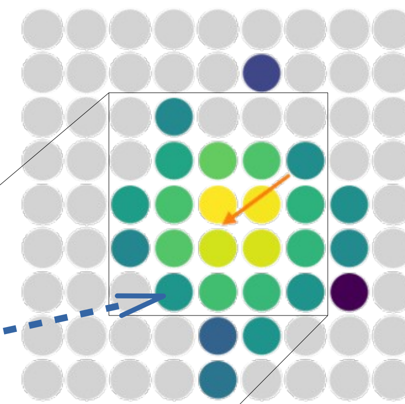




Network Architecture

Information stored in 9x9 maps:

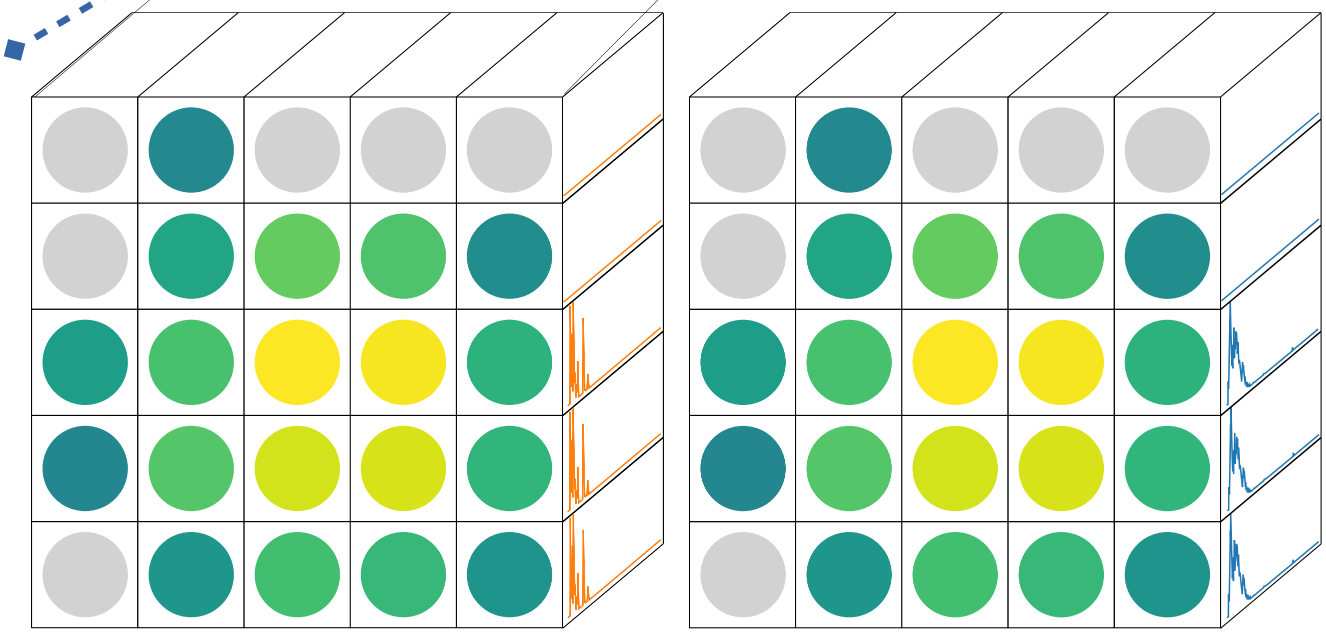
- A 3D map of station positions and WCD traces
- A 3D map of station positions and SSD traces
- A 2D map of station positions and vectors containing relative station arrival time Δt_i and event *E and θ*



inspired by
network architecture
developed in Aachen
https://arxiv.org/pdf/1708.00647.pdf

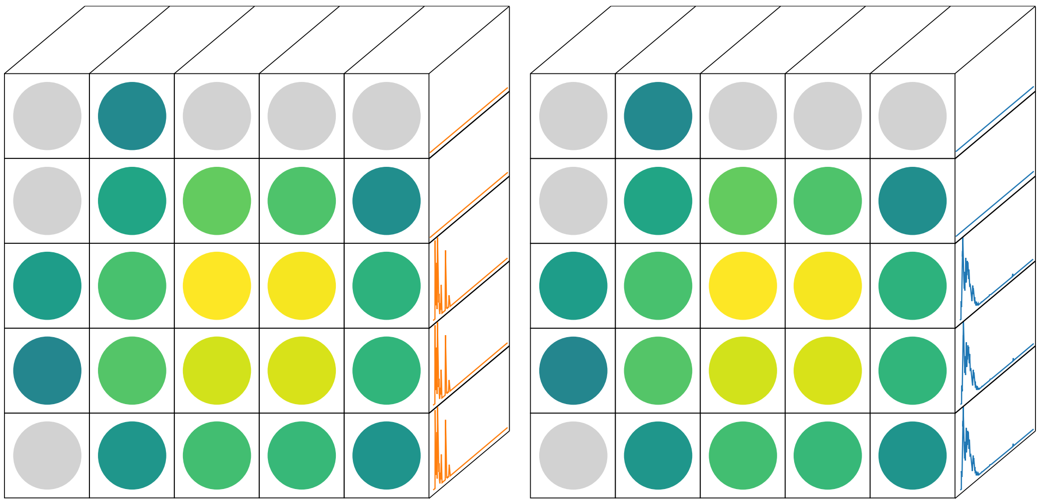
2D map with additional info in vector:

$$\begin{pmatrix} \Delta t_i / ns \\ E / \log 10 (eV) \\ \theta / rad \end{pmatrix}$$



3D map combining station position and traces
SSD and WCD

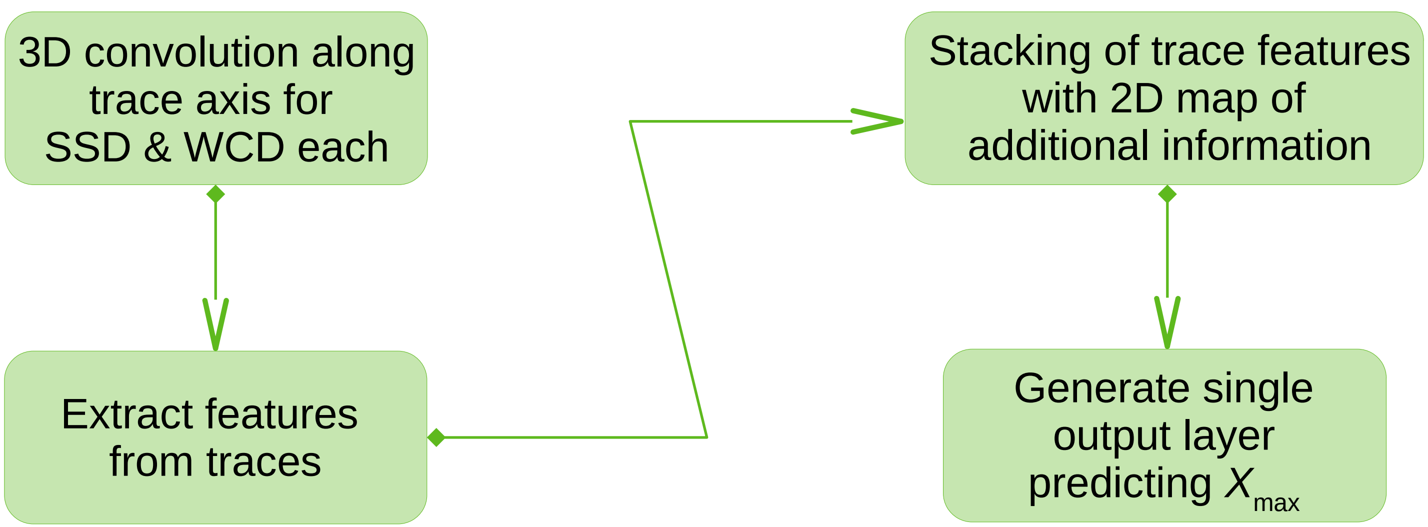
Network Architecture



3D map combining station position and traces
SSD and WCD

2D map with additional info in vector:

$$\begin{pmatrix} \Delta t_i / ns \\ E / \log 10 (eV) \\ \theta / rad \end{pmatrix}$$



X_{max} Reconstruction Setup

Input data fed to model via `fit_generator/predict_generator` Keras functions

→ Streaming input files from directory with `generator` function

Evaluate trained network based on

- Correlation coefficient of $X_{max,true}$ and $X_{max,\ DNN}$ distributions
- Check reconstruction bias and resolution
  - Split $X_{max,true}$ and $X_{max,DNN}$ values by energy bins (size 0.1 $\log_{10}$ eV)
  - Fit $\Delta X_{max} = X_{max,\ true} - X_{max,\ DNN}$

    distribution for each energy bin with superposition of 2 Gaussian distributions
  - μ from fit represents the **reconstruction bias**
  - σ from fit represents the **reconstruction resolution**

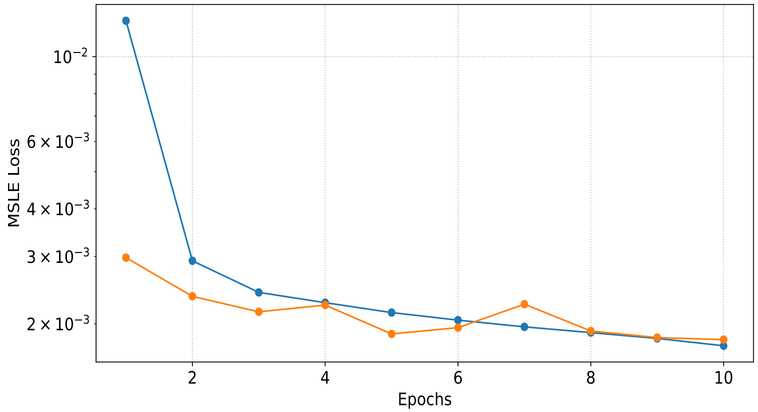
	# Events
Total	234480
Train	140704
Validate	46896
Test	46880
	Settings
Batch size	16
Epochs	10
Learning rate	0.001
Training time (on GPU cluster)	~ 1.5 days

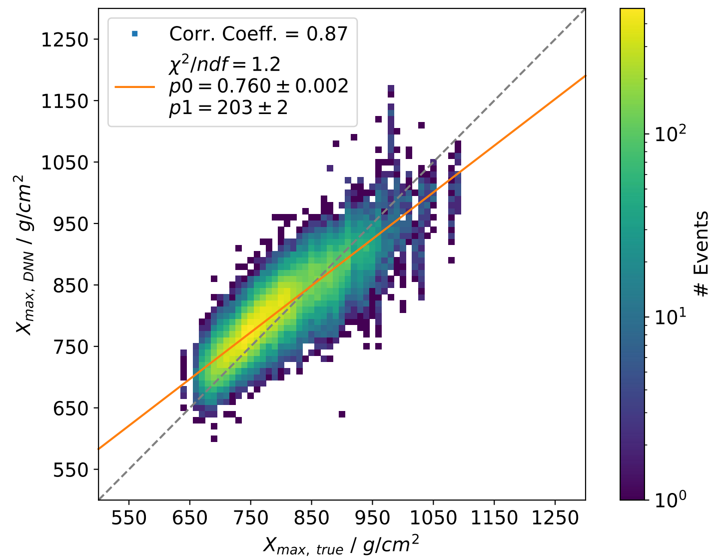
GPU Cluster
CPU: Intel® Xeon® E5-2660
RAM: 128 GB
GPU: 5 × NVIDIA Tesla K20m
Using: Tensorflow + Keras

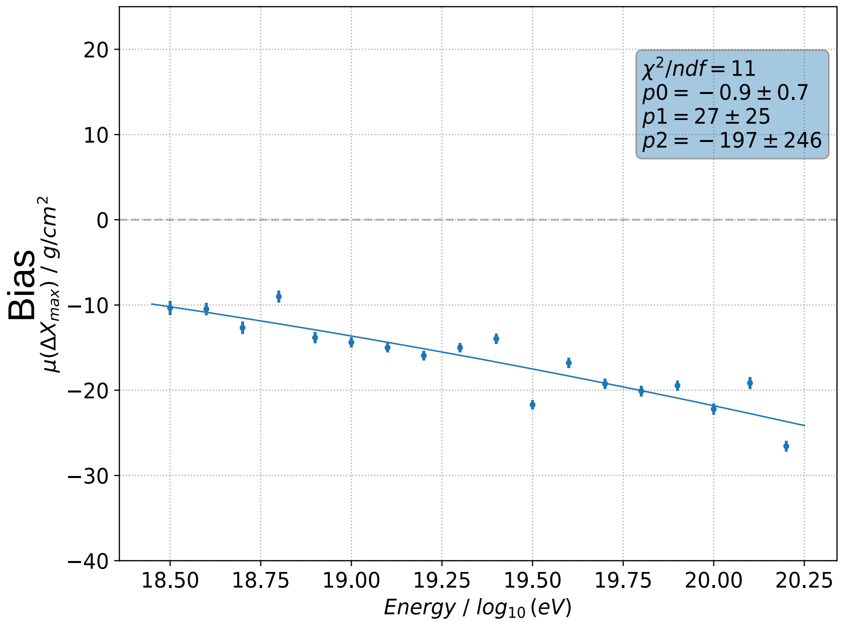
X_{max} Reconstruction Results (combined)

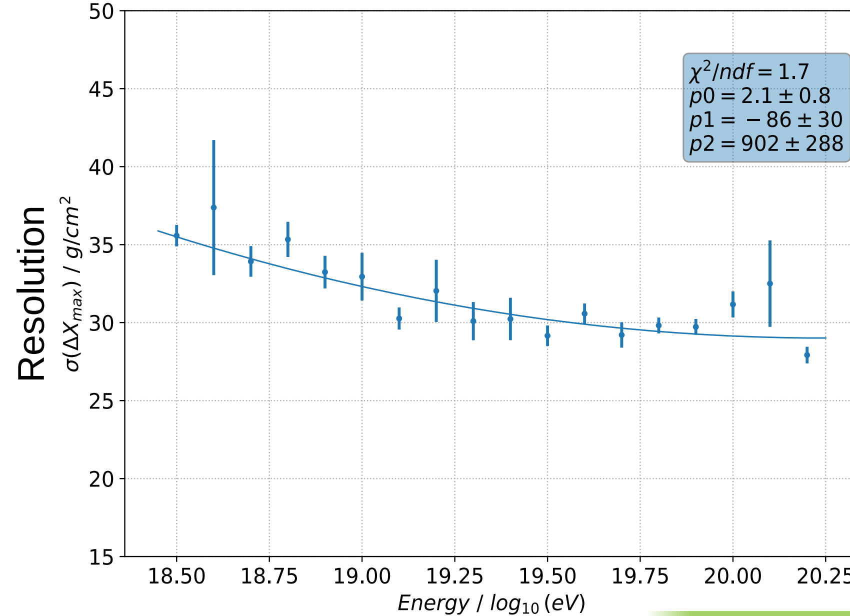
Correlation of 87% between MC truth and prediction

X_{max} reconstruction with DNN possible

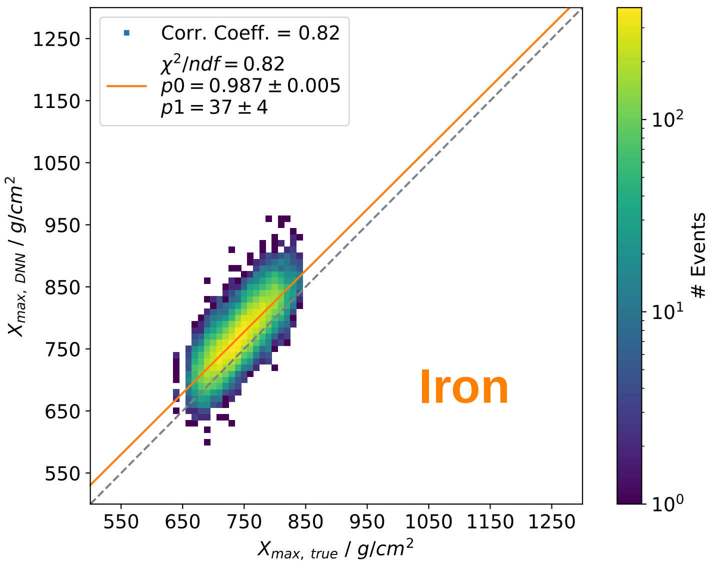


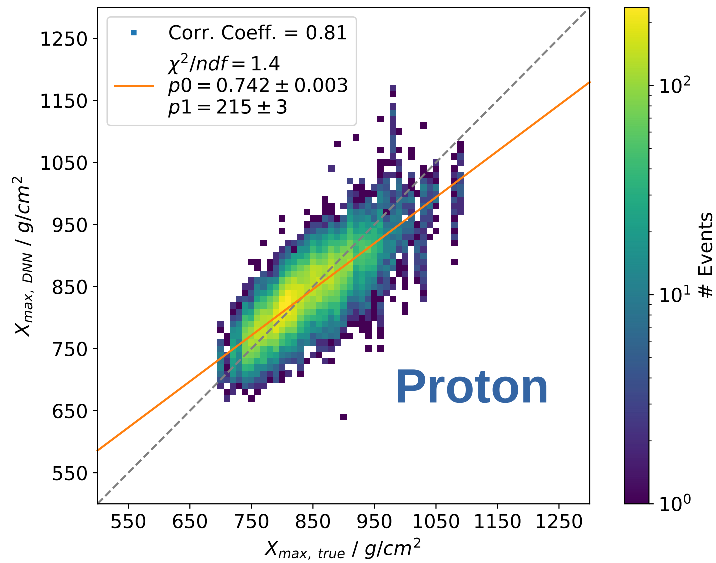


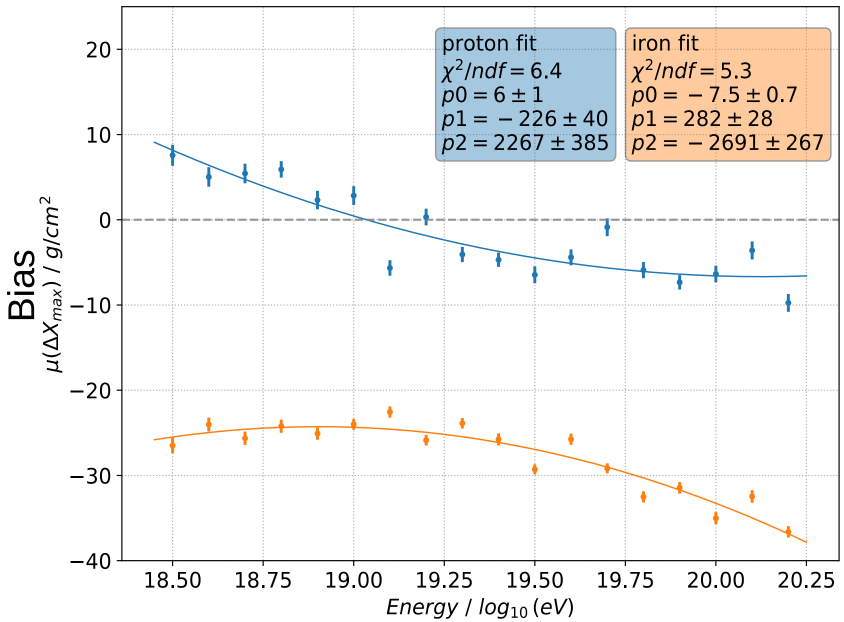


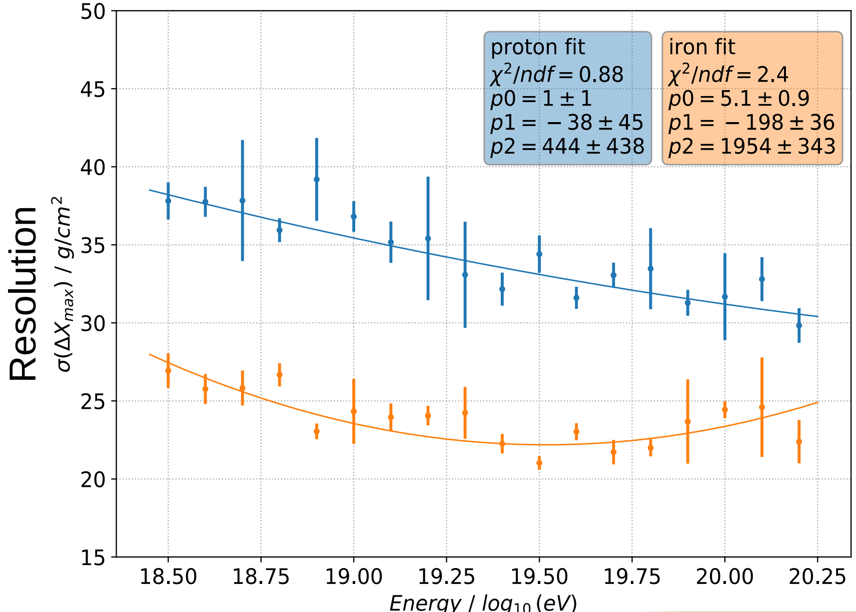


X_{max} Reconstruction Results by Primary Particle









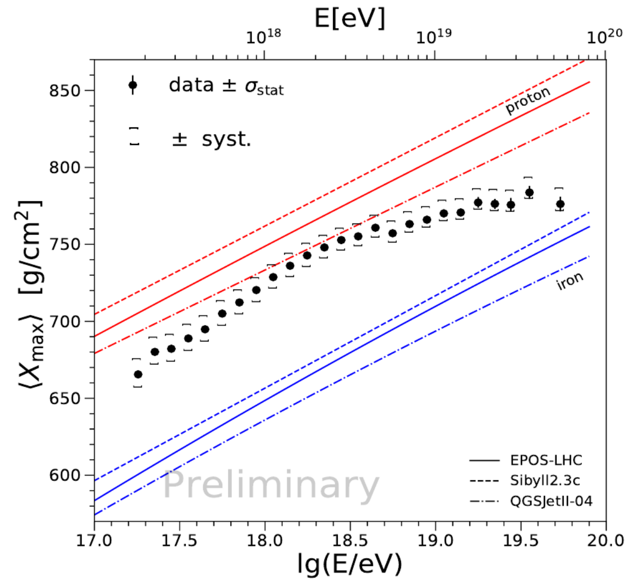
Elongation Rate (ER)

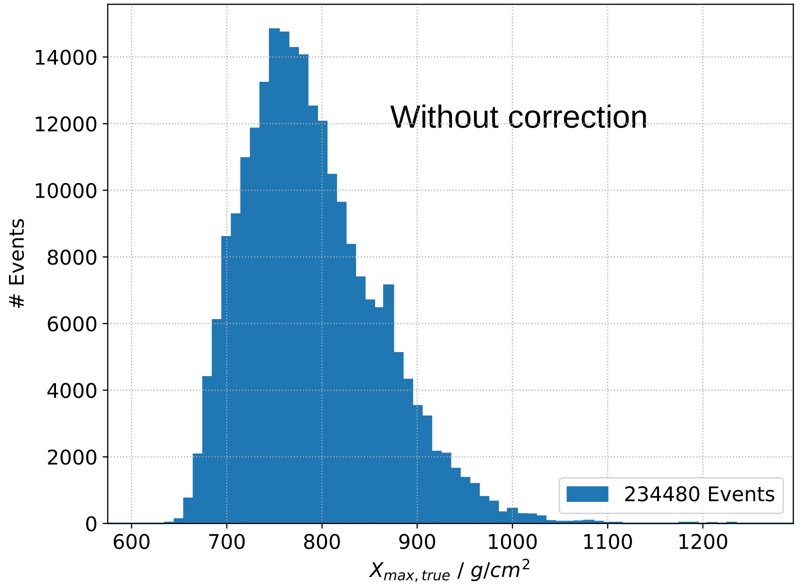
Elongation rate:

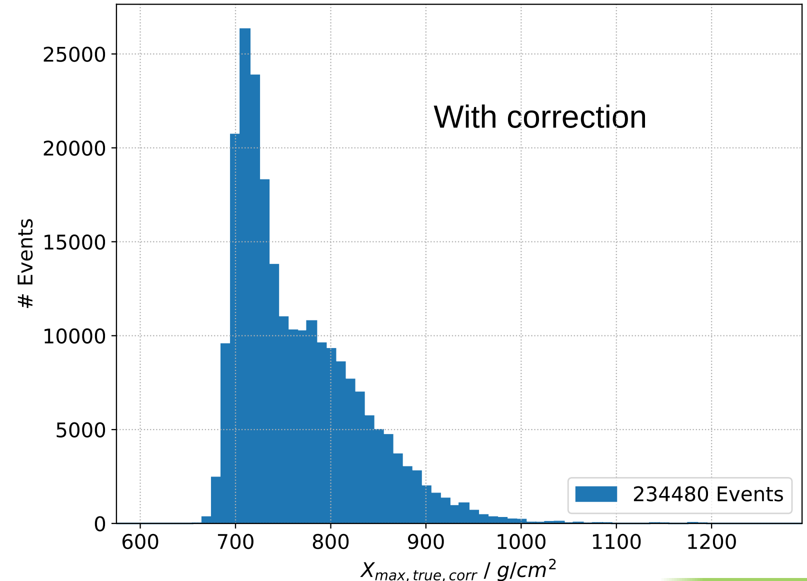
- Rate of change of the shower maximum with logarithm of energy
- $X_{max,true,corr}$ =

$$X_{max,true} - 60\ \mathrm{g/cm^2} \times (\log_{10}(E/\mathrm{eV}) - 19)$$

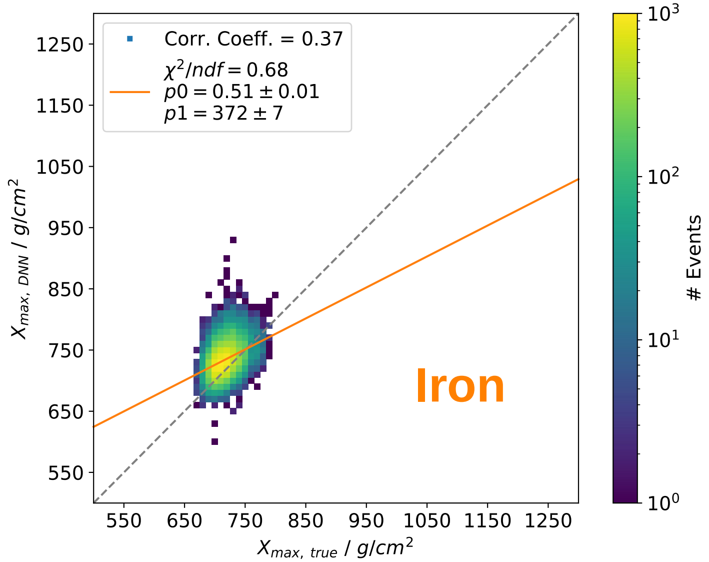
$\log_{10}$ of reference energy of
10 EeV i.e. 10^{19}eV

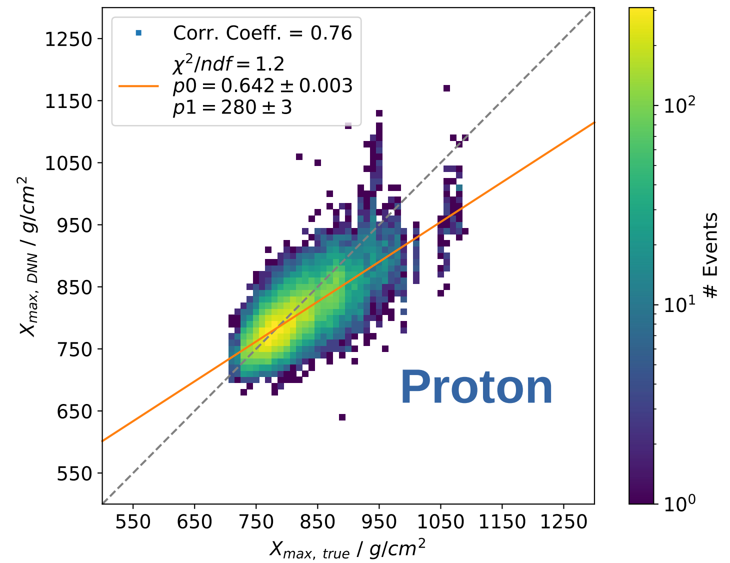


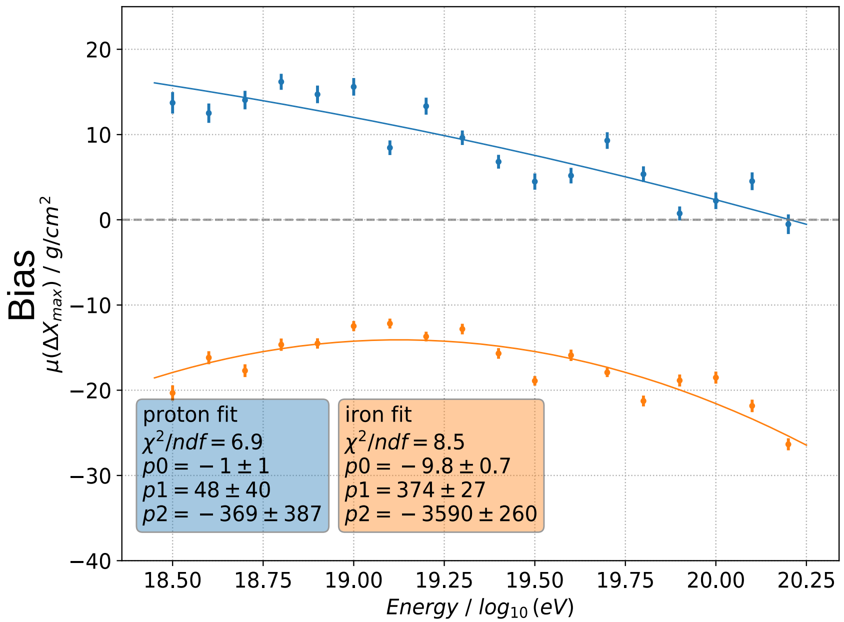


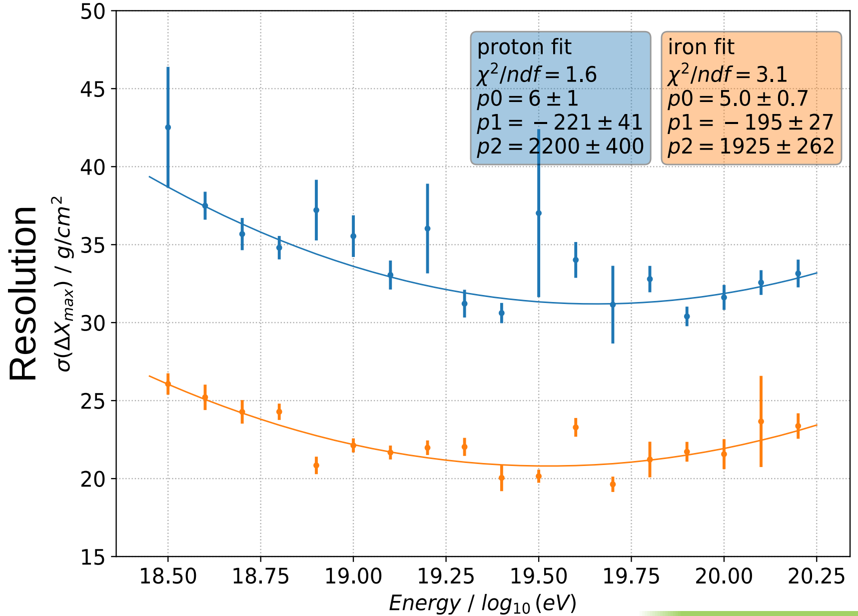


X_{max} Reconstruction Results with ER









Event Weighting (EW)

X_{max} distribution for training/validation is not flat

Events with X_{max} of ~720 g/cm^2 seen by model more often

How to prevent model from drawing prediction from ~ <X_{max}> ?

→ Event weighting:

e.g. to get weights between 1 and 5:

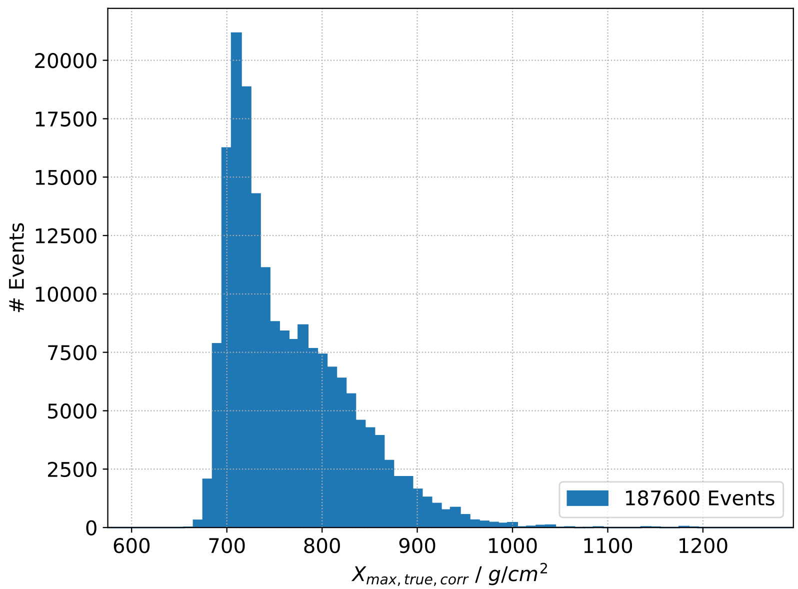
$$w = \frac{1}{\left(\frac{hist_{counts}}{max\left(hist_{counts}\right)}\right) \times 0.8 + 0.2}$$

- Re-weighting X_{max} distribution by assigning higher weight to Xmax values with less statistics

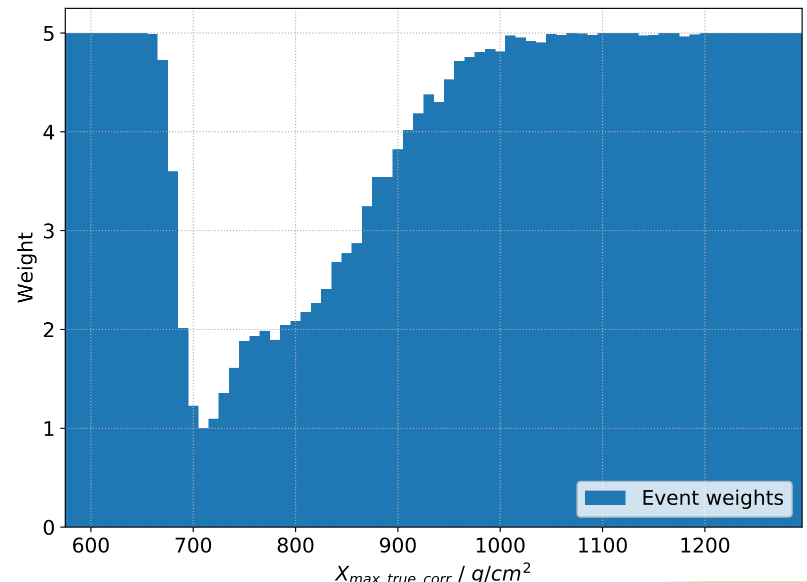
  → Network pays more attention to events with higher event, i.e. mimicking flat distribution, pulling network away from the mean

- Providing weights for each event in `fit_generator` function via `sample_weights` parameter

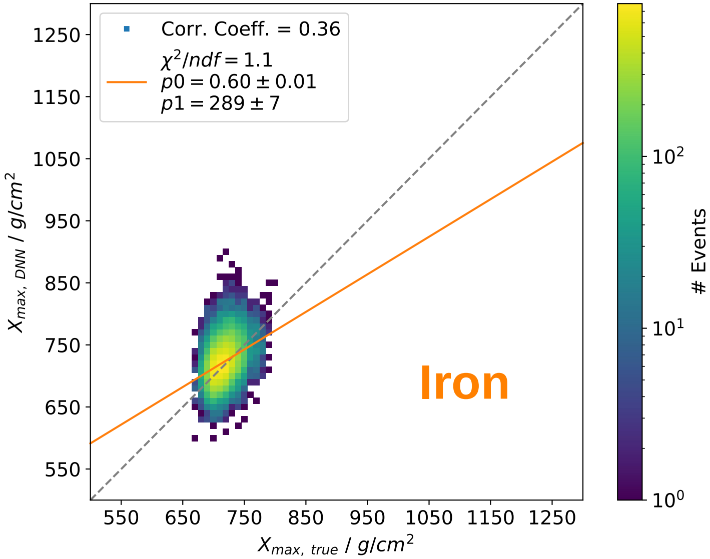
Input X_{max} distribution

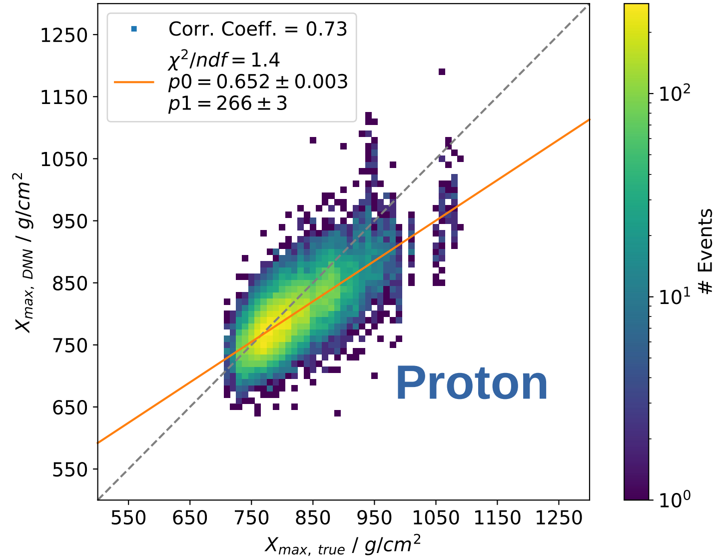


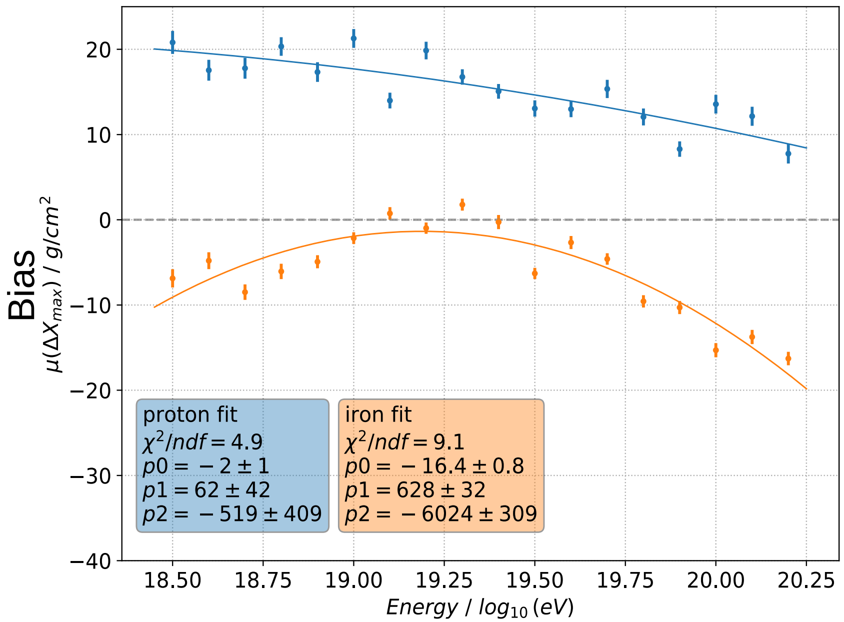
Calculated corresponding weights

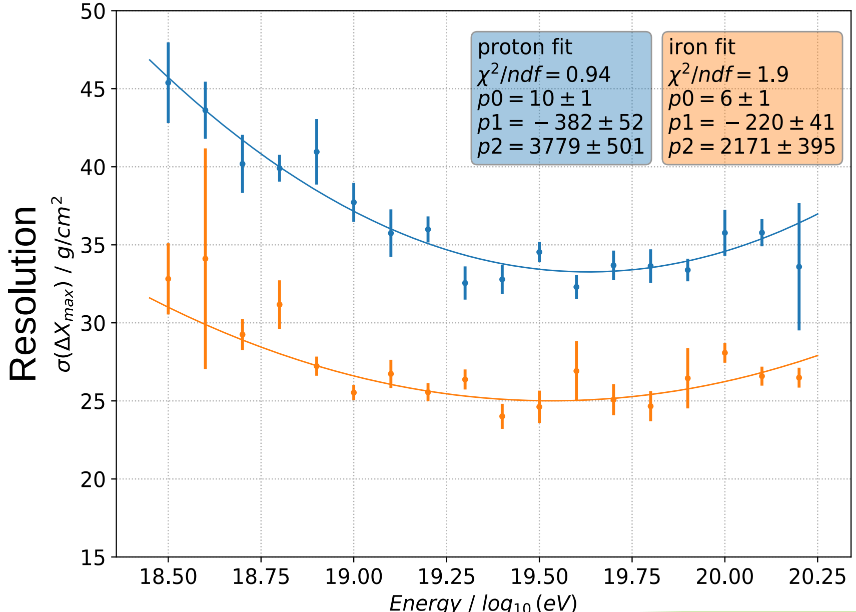


X_{max} Reconstruction Results with ER & with EW









Outlook

Identified points for improvement:

- 9 × 9 station grid possibly cuts out stations with signal especially at higher energies

→ Increase grid size to not cut off outer stations

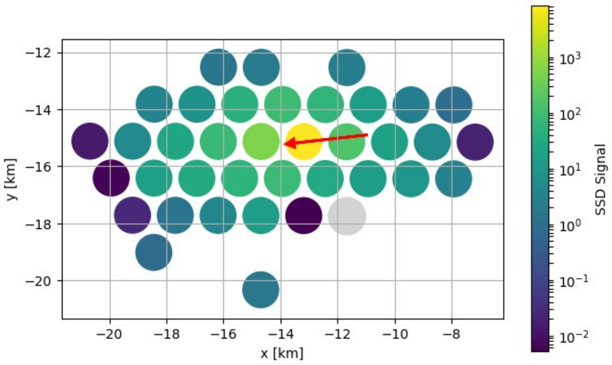
- Bias still challenging, Iron always heavier than MC / proton always lighter

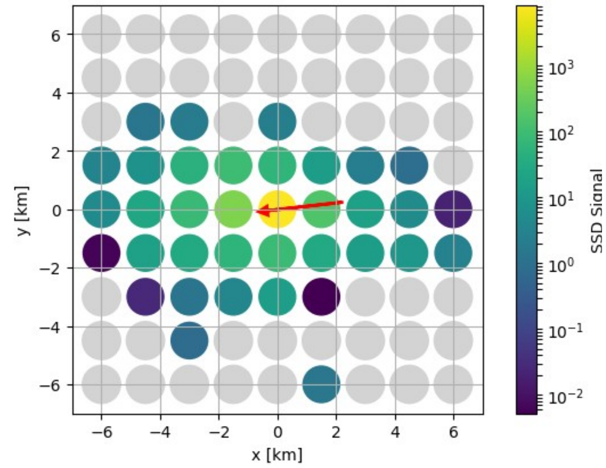
→ Implement different type of model structure:

  - Network that is similar to speech recognition/comparison/identifying fake speech,...

    → Trace by trace comparison to extract features based on differences in trace types

  - Hexagonal convolutions (6-fold rotational symmetry), especially suitable for tasks with already hexagonal layout





Additional plans/ideas:

- Further increase event statistics and add Helium and Oxygen as primaries
- Primary-specific event weighting
- Try on real data from pre-production array

Sonja Schröder | soschroeder@uni-wuppertal.de

Conclusions

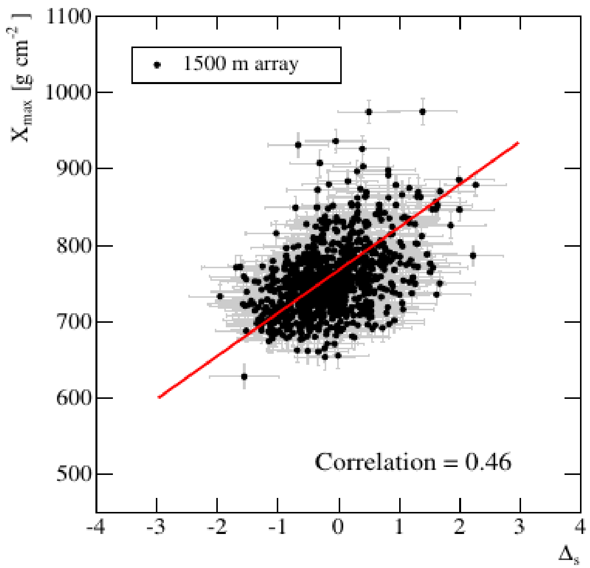
X_{max} reconstruction with DNN is possible

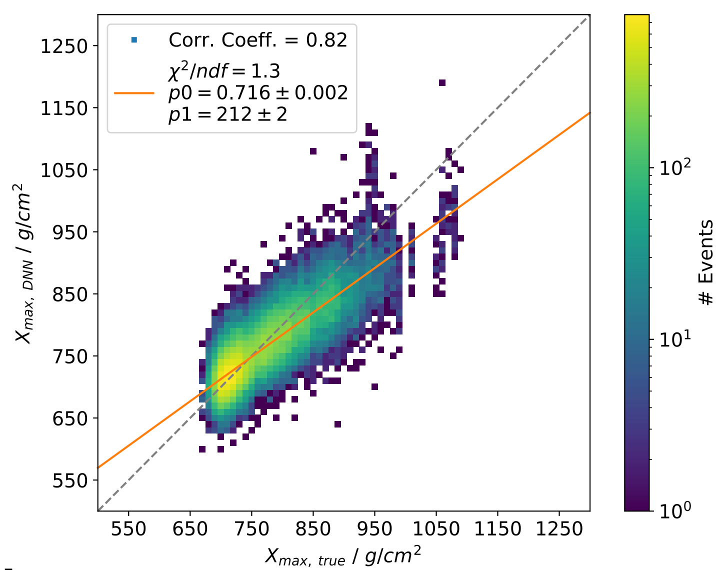
- Currently correlation of 82%
- Resolution of ~ 30 g/cm^2

Successfully implemented

- X_{max} correction via elongation rate → removing energy dependent bias
- X_{max} event weighting to mimic flat distribution → reducing primary dependent bias

Current model already performs better than classic methods



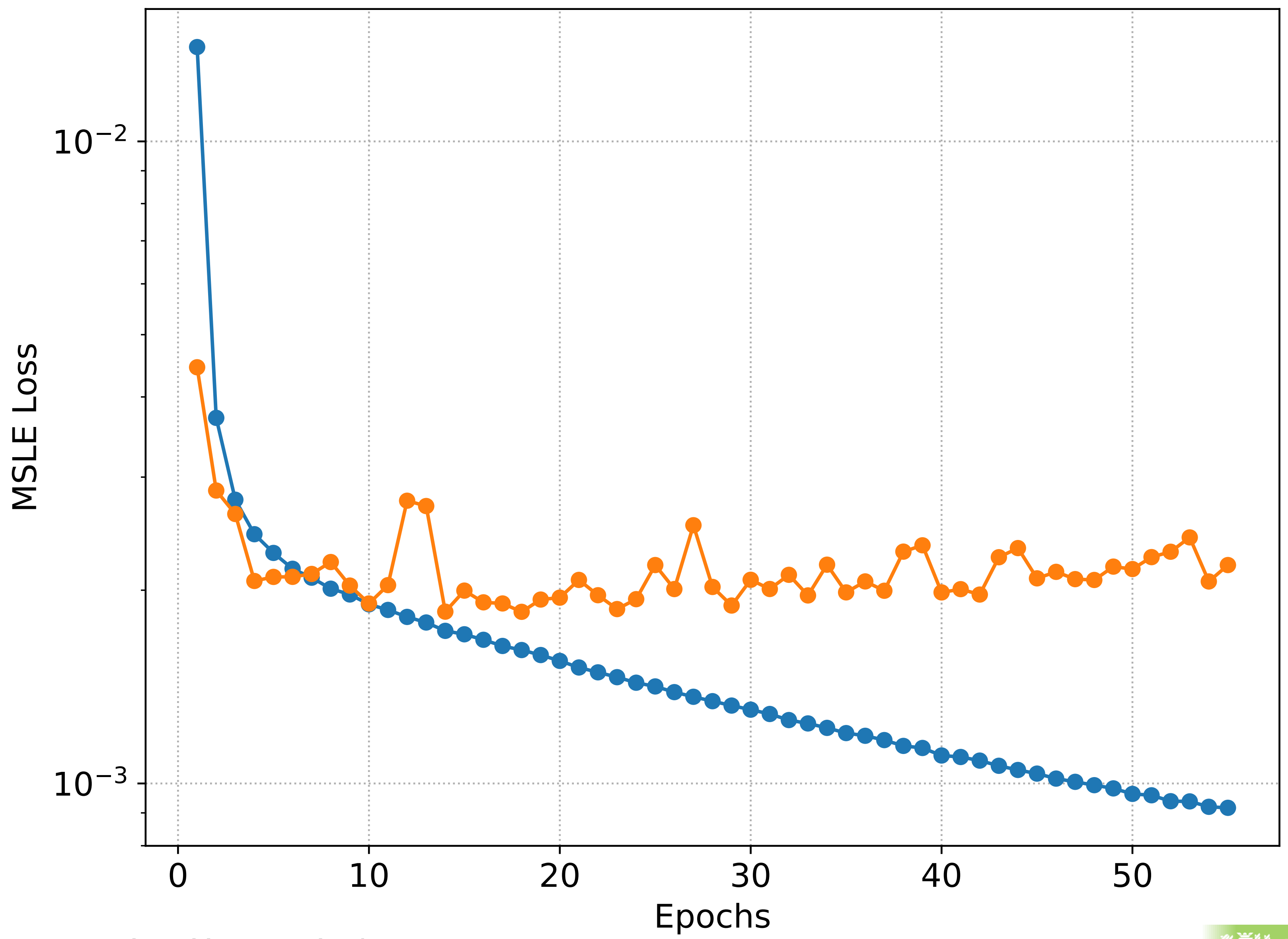


Thank you!

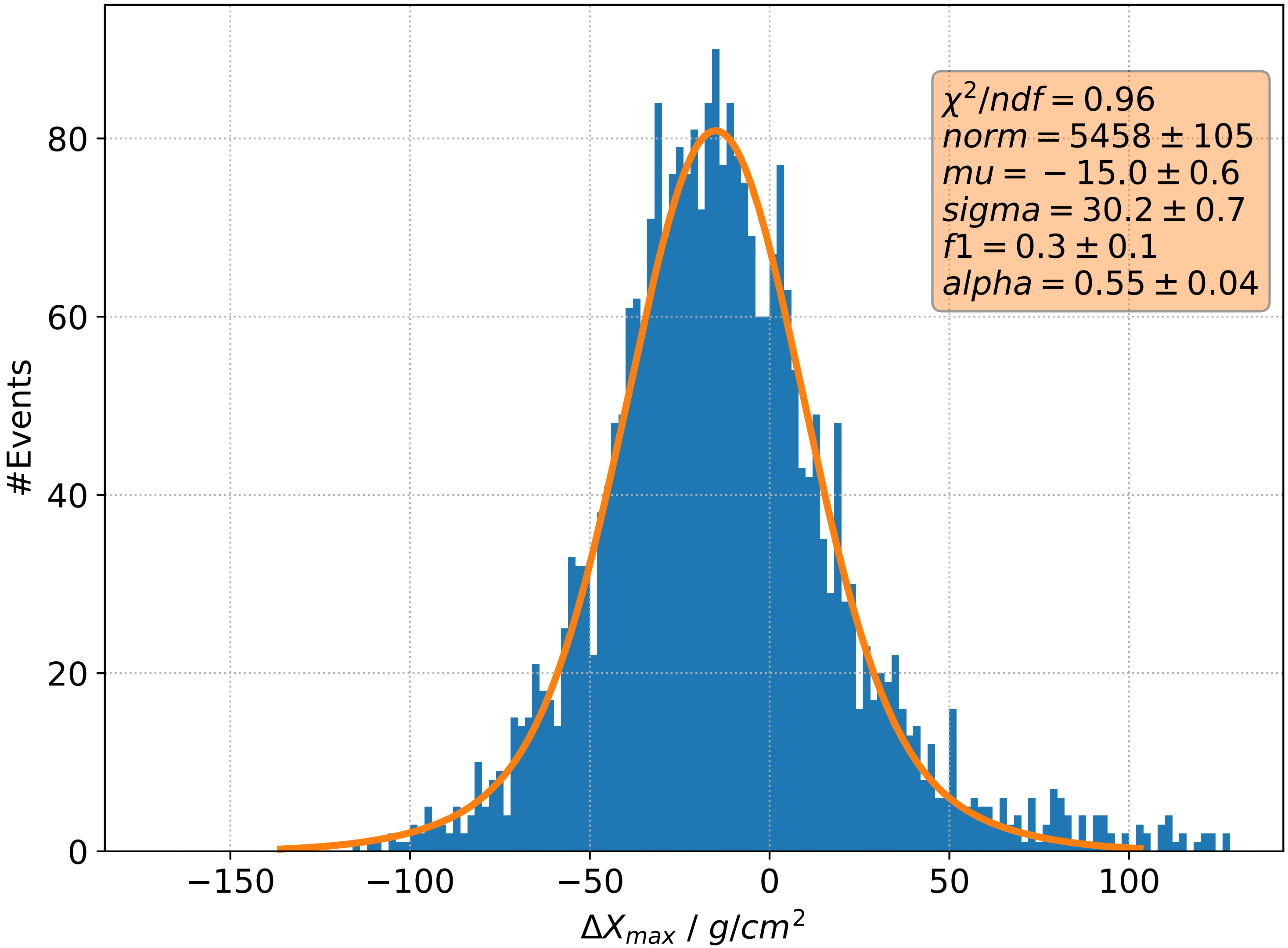
Any questions or feedback?

BACKUP SLIDES

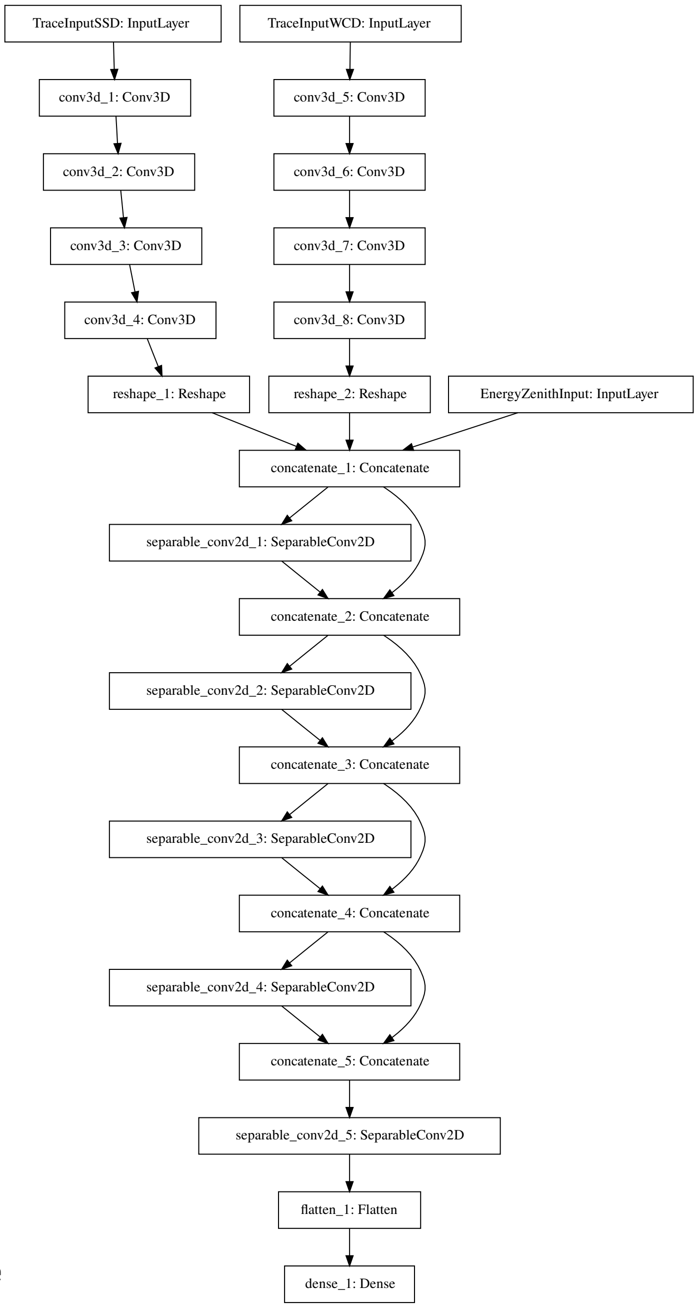
Loss Evolution over 55 Epochs



Example Fit Double Gaussian



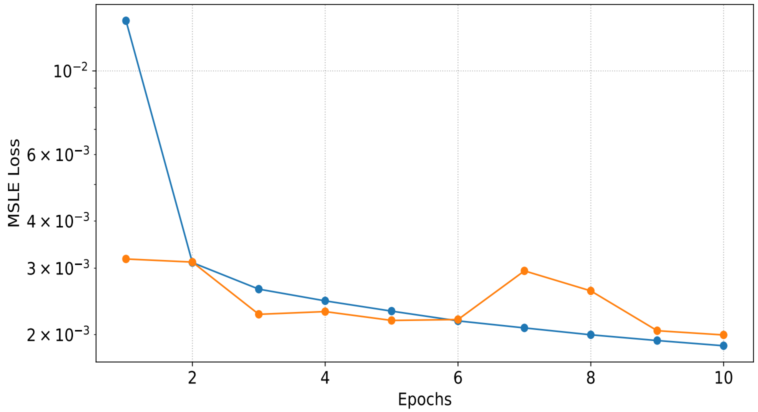
Full Network Structure

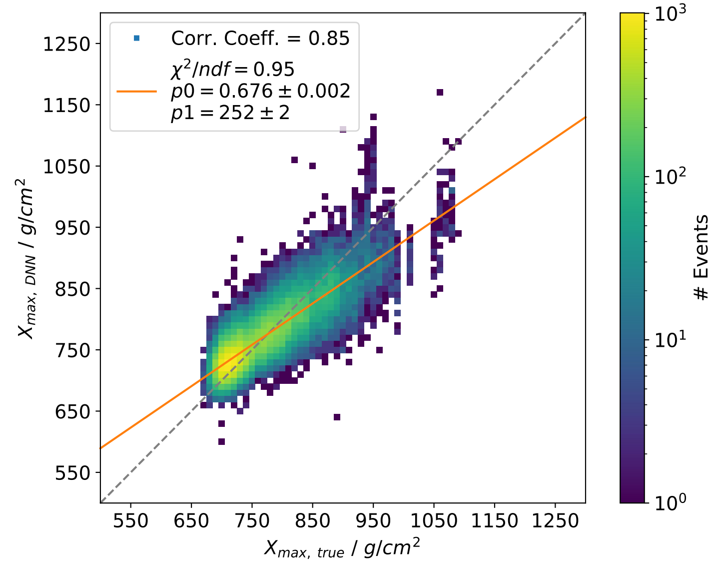


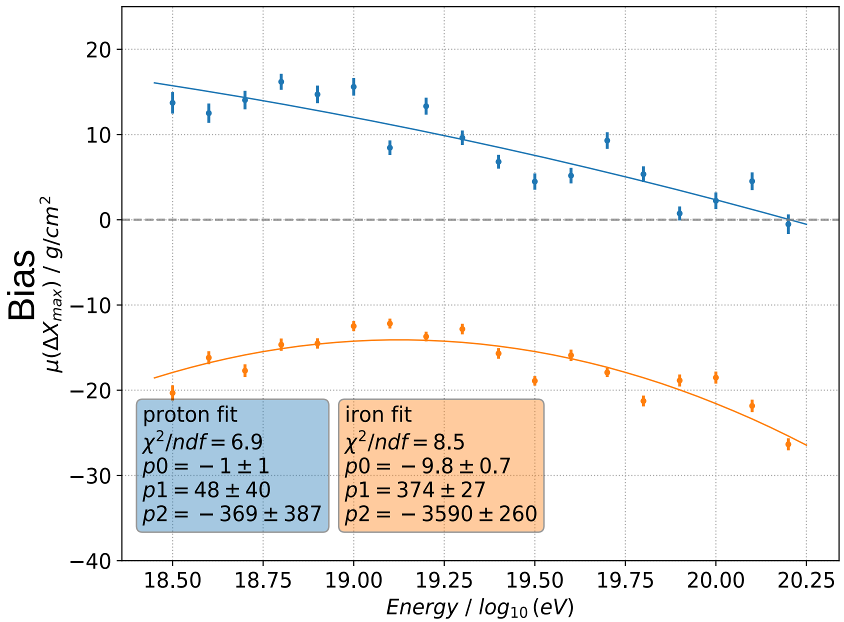
X_{max} Reconstruction Results with ER

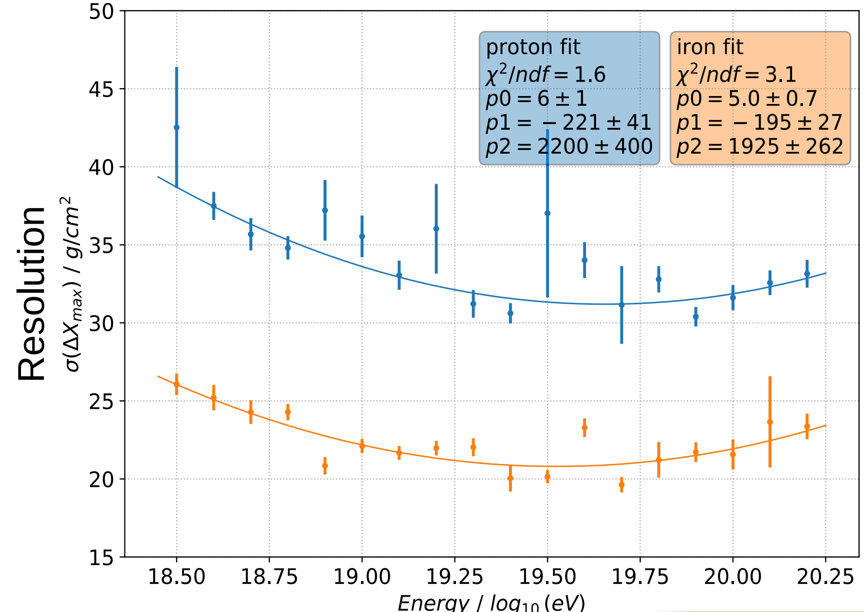
Overall correlation decreases

Improves bias and resolution of Iron









X_{max} Reconstruction Results with ER & with EW

Overall correlation decreases further

Bias on Iron improves, while resolution is worse