Extraction of the muon signal using LSTM in the TA experiment

Kalashev Oleg, Suzdalov Gleb
INR RAS
QUARKS-2026

TA experiment

We want to extract the muon signal from the full SD detector signal

In the future, with such data, it will be possible to test various interaction models for the compatibility of LHC data and astrophysical experiments.

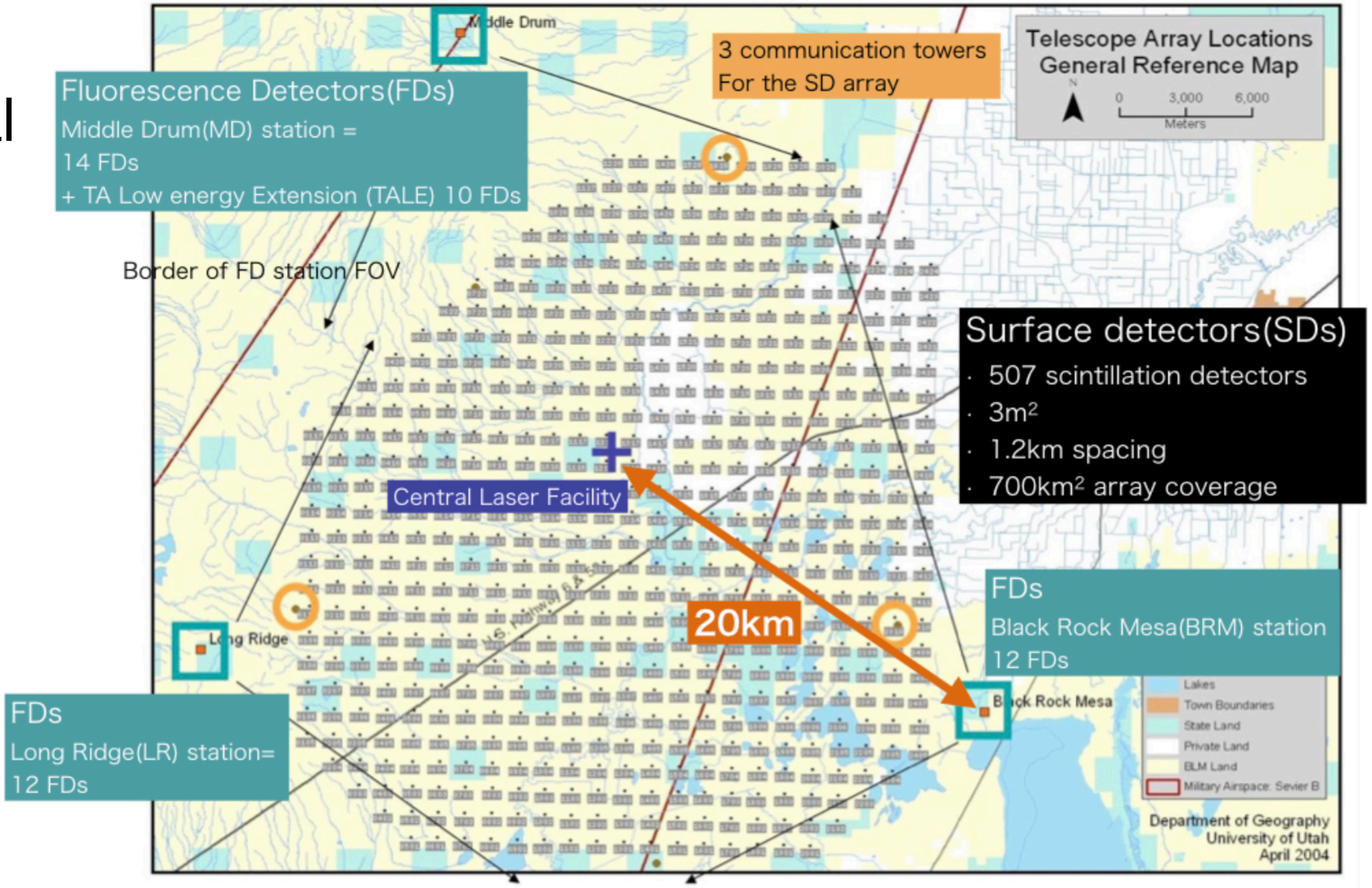


Fig.1: Telescope array experiment scheme

Data production

Corsika simulation → Simulation of the detector response → Reconstruction of the zenith angle and the core distance

Choosing an interaction model

For details, see: 1403.0644

As a result, we get a dataset containing about 2M events.

Dataset used

- We use QGSJET II-03 model for particle interactions

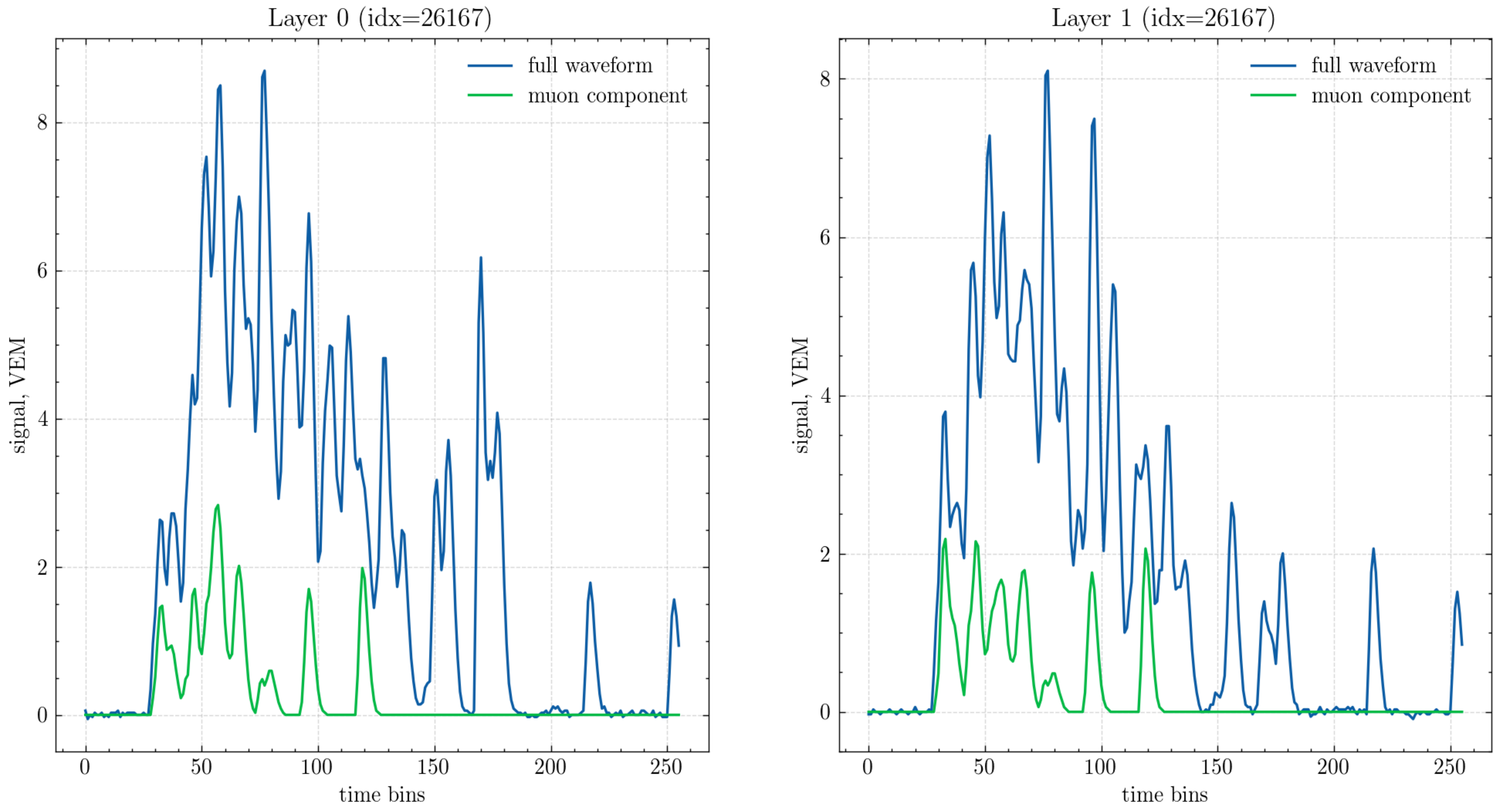


Fig.2: Input data example

LSTM

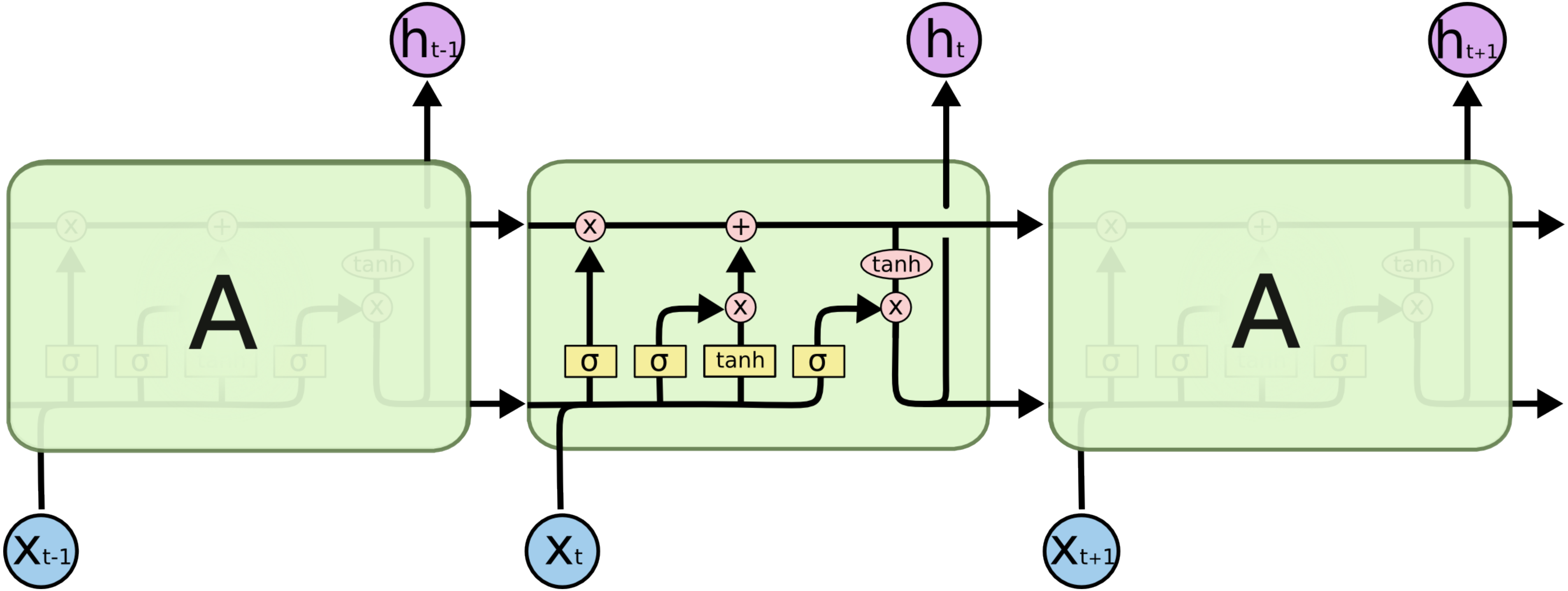


Fig.3: LSTM block diagram

Neural network architecture

We took the architecture used for a similar problem in the Pierre Auger experiment, see 2103.11983

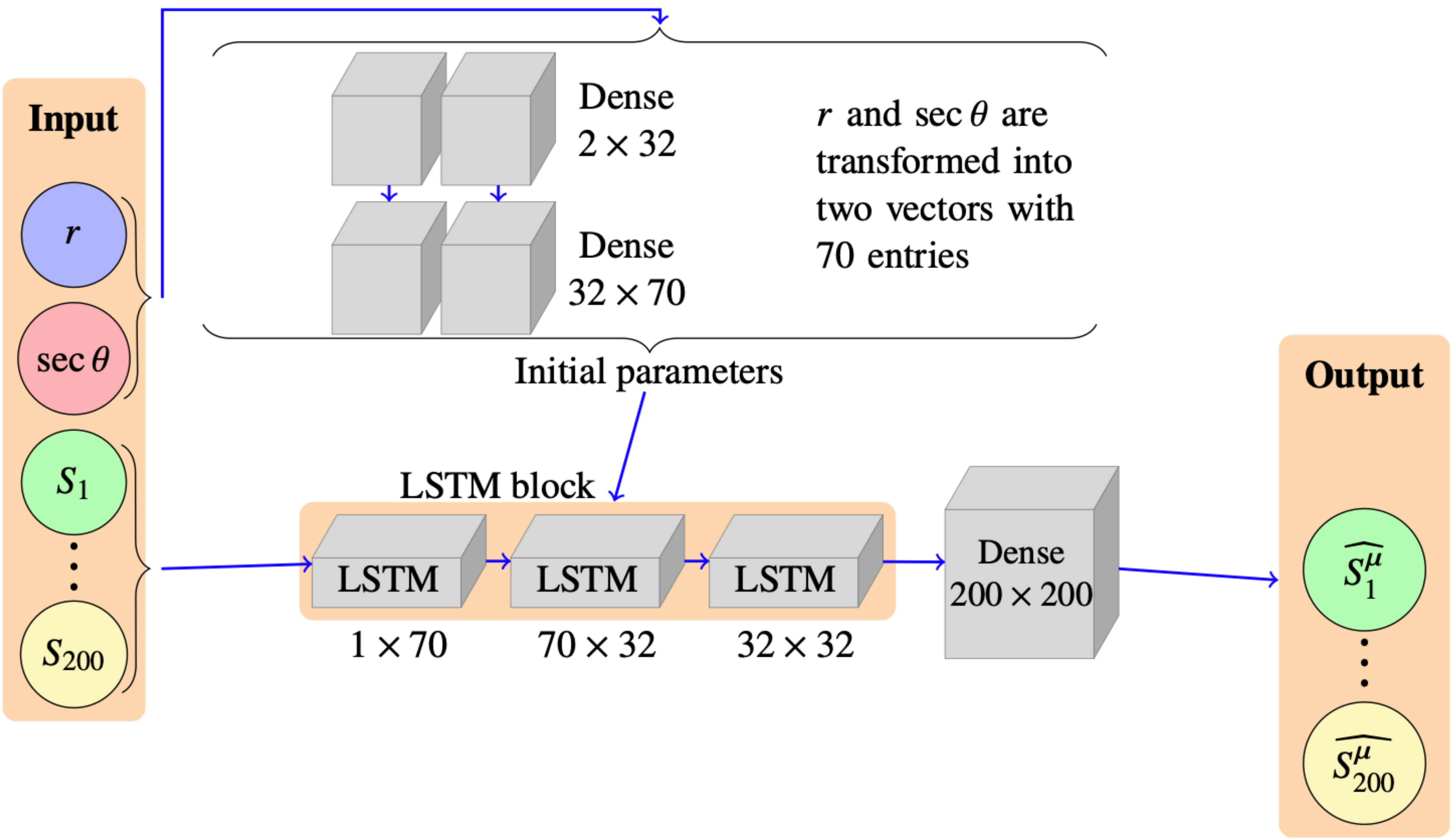


Fig.4: Original architecture for Pierre Auger experiment

Changes made:

- We take the signals from both layers as input, the first two frames (256 bins), so the input sequence is 256*2, and the first LSTM takes two sequences.
- As output, the neural network produces a muon signal on both layers with a size of 256*2

Loss function: $L = \frac{1}{256 \cdot 2} \sum_{j=0}^{1} \sum_{i=1}^{256} (S_{ij,pred} - S_{ij,true})^2$

Data preprocessing

- We only keep events where two layers are triggered
- We use reconstruction to restore zenith angle and distance from the core
- Zenith angle range from 0 to 45
- The number of stations that registered an event is greater than 5, and then we only keep the event with the largest integral signal
- Distance from the core and sec of zenith angle are scaled to be between 0 and 1
- The full signals are individually scaled to be between 0 and 1, and the muon signal is scaled by the same factor

Train/Test/Validation split

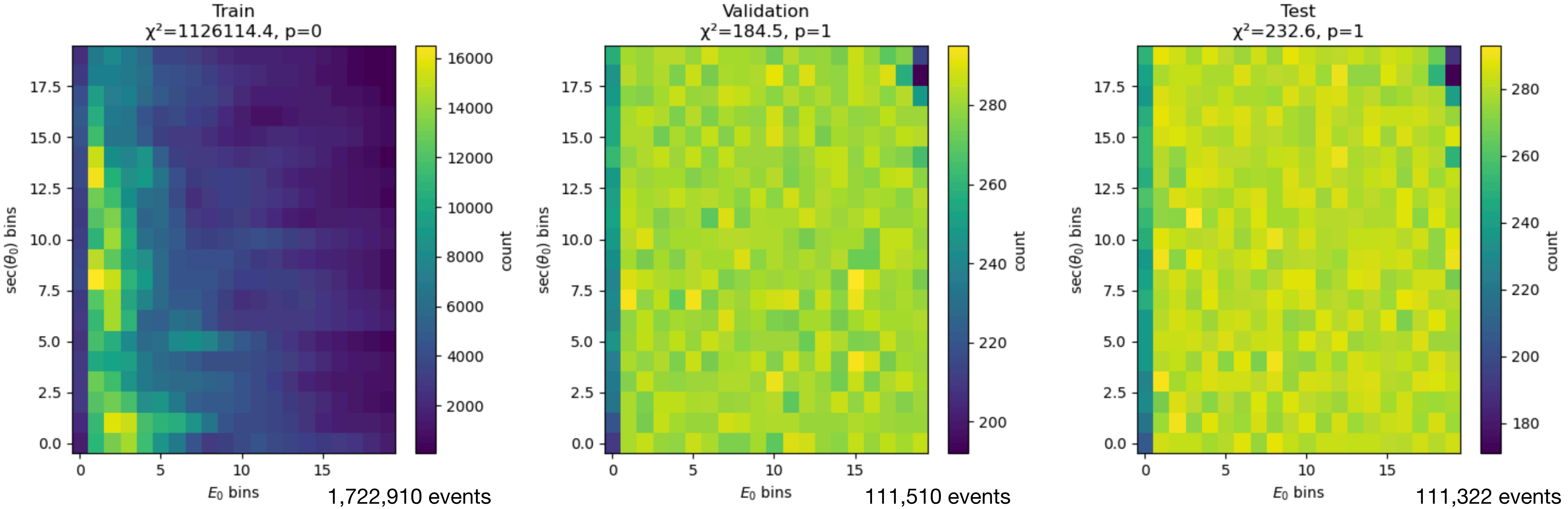


Fig.5: Train/Test/Validation split

The data is split in such a way that the test and validation datasets have a uniform distribution of energy and $\sec(\theta)$. The remaining data goes to the training dataset. The graphs show the 2-d χ^2 test for compatibility with a uniform distribution

Results: Waveform predictions

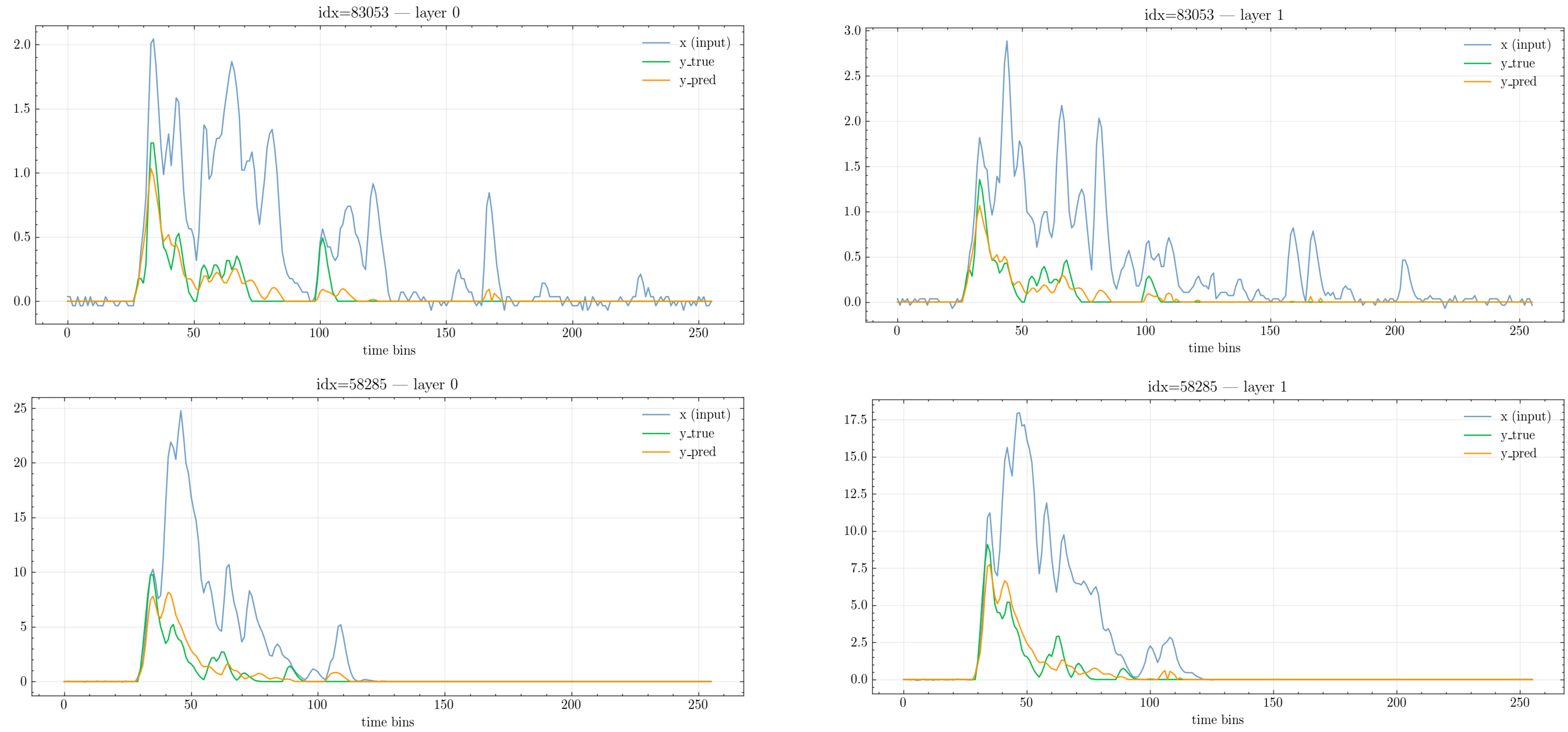


Fig.6: Waveform predictions examples

Results: integrals of muon trace

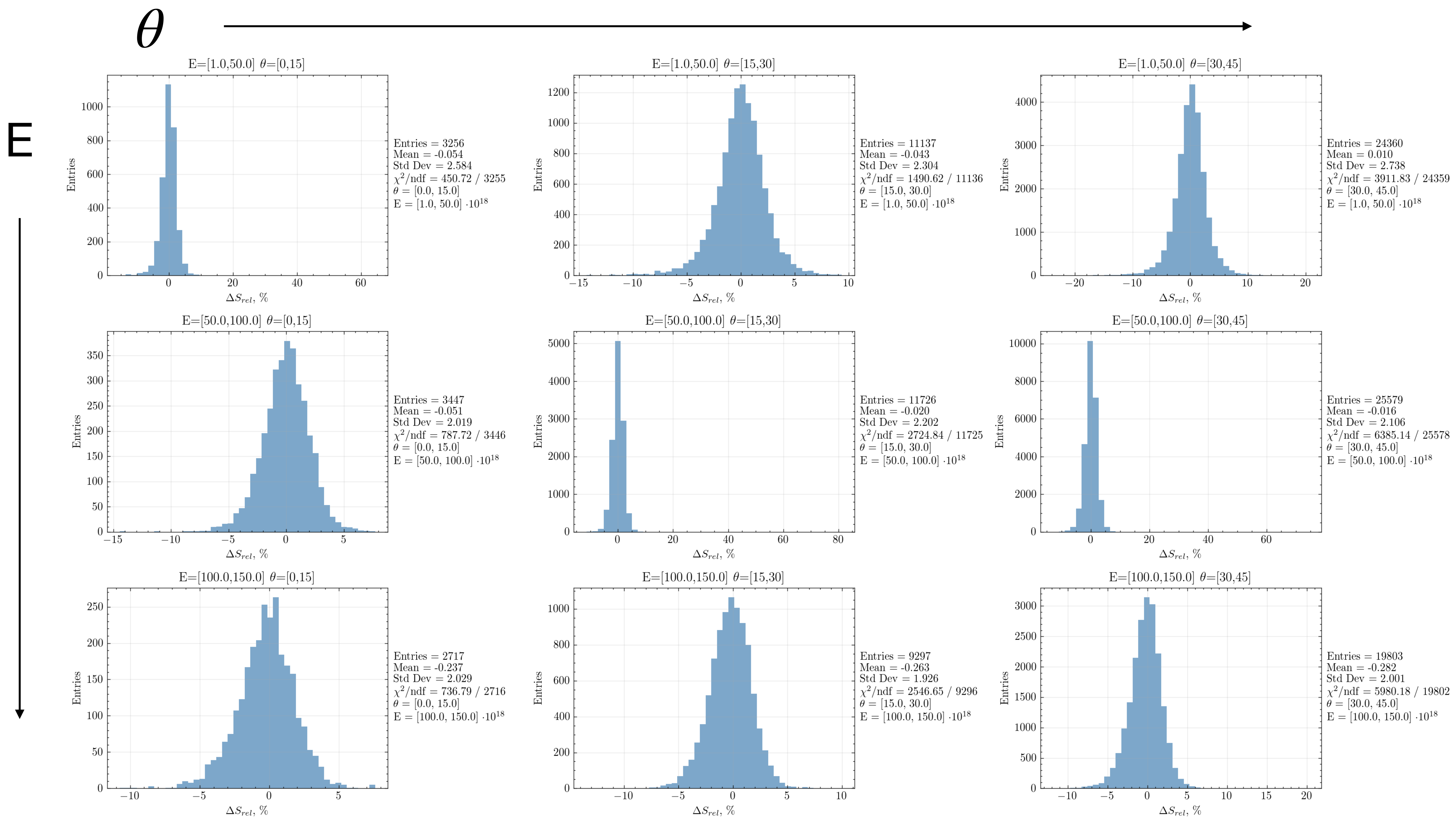
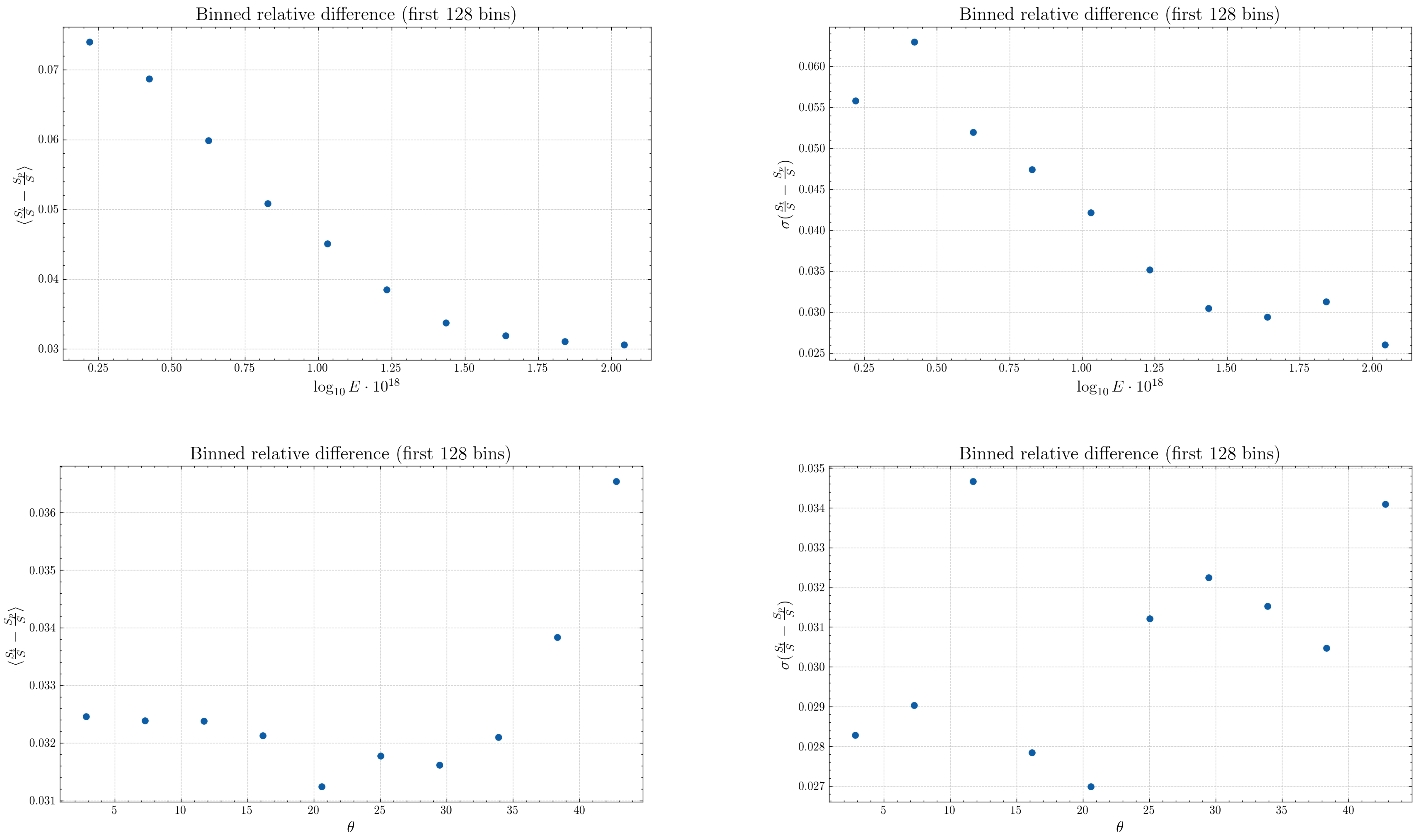


Fig.7: Distribution $\frac{S_{pred} - S_{true}}{S_{full}}$ at different energies and zenith angles

- When calculating all subsequent metrics, we take only the first frame (128 bins) of the signal

Dependence of the relative error on angles and energy



- The error decreases with increasing energy

- At large angles, the error increases slightly

Fig.8: Dependence of the relative error on angles and energy

Integrals of the trace distribution

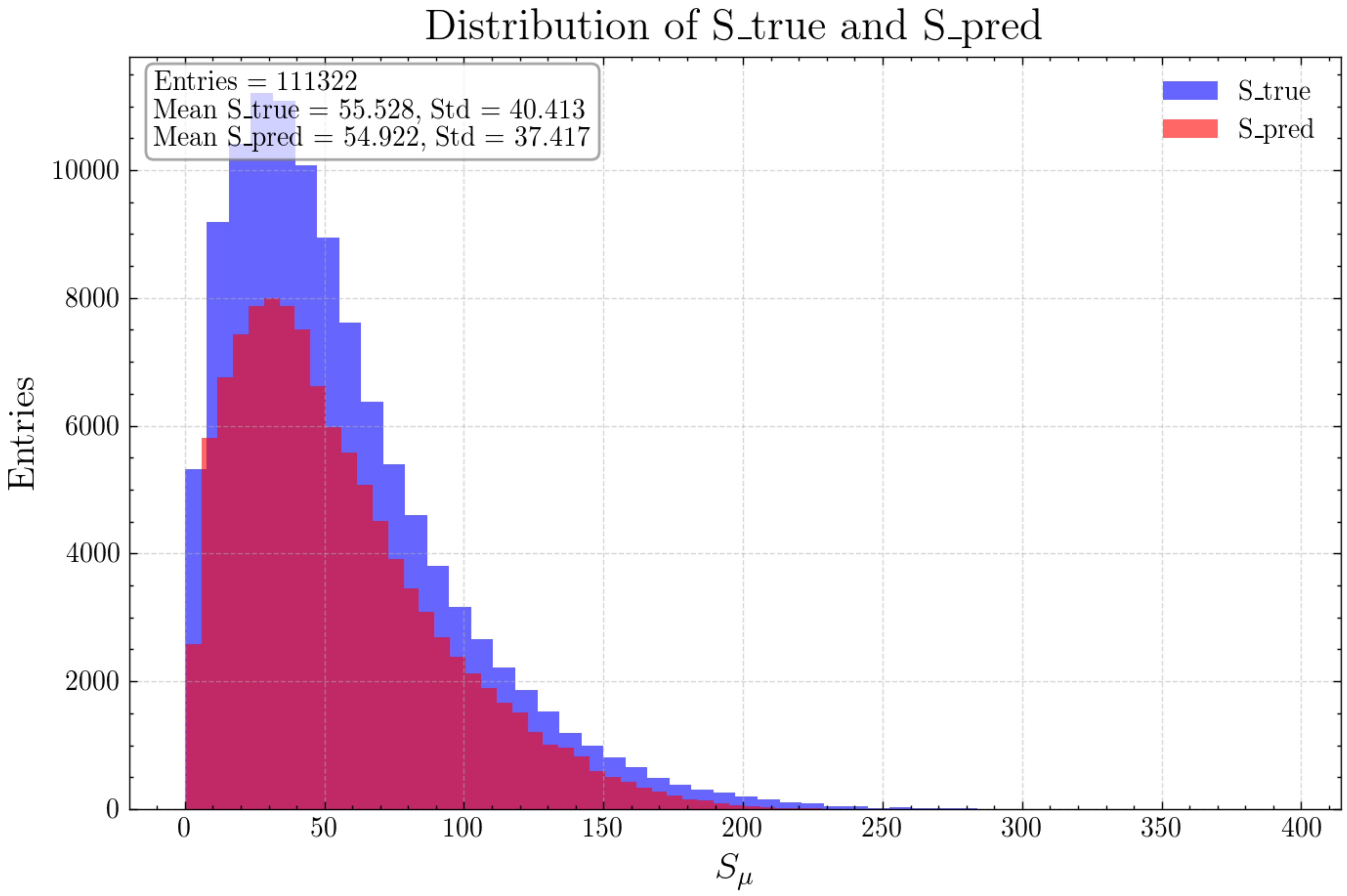


Fig.9: Distribution of predicted and true integrals of the trace

- You can see that the model underestimates the predictions
- Possible ways to avoid this problem: modifying the loss function with penalties for underestimating peaks

Evaluation of peak prediction quality

- To identify a peak, we use the scipy find peaks function with default parameters
- The maximum allowed peak offset value is 3 bins
- To evaluate quality, we use the analog of precision/recall, defined as follows:

  TP — the peak is present in both true and pred, value: $\min\left(\frac{f_p}{f_t}, 1\right)$

  FP — the peak is present only in pred, value: $\frac{f_p}{\max(f_t)}$

  FN — the peak is present only in true, value: $\frac{f_t}{\max(f_t)}$

  Precision = $\frac{TP}{TP + FP}$ Recall = $\frac{TP}{TP + FN}$

- Results:
  Layer 0: Precision=0.810, Recall=0.842
  Layer 1: Precision=0.795, Recall=0.843

- Problems: the metric value strongly depends on the peak search parameters and the allowed shifts

Conclusions

- We have a neural network that can extract the muon signal in TA experiment
- The average relative error in predicting the muon trace is about 0.05%
- We proposed a metric for checking the quality of peak prediction

Further work

- Verify that the neural network's predictions are model independent
- Test the transformer architecture to solve such problems

Thank you for your attention