# Analyzing TAIGA data with machine learning

Based on materials obtained during the implementation of the RSF grant 24-11-00136.

**Alexander Kryukov* on behalf of the RNF team**

*Lomonosov Moscow State University, Skobeltsyn Institute of Nuclear Physics (SINP MSU, Moscow)*

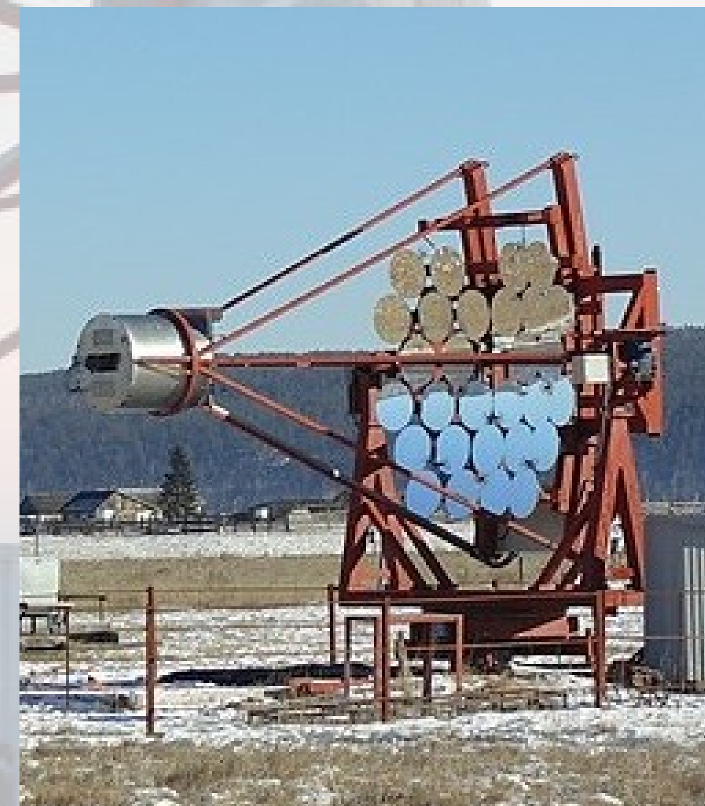
** Email: kryukov@theory.sinp.msu.ru*

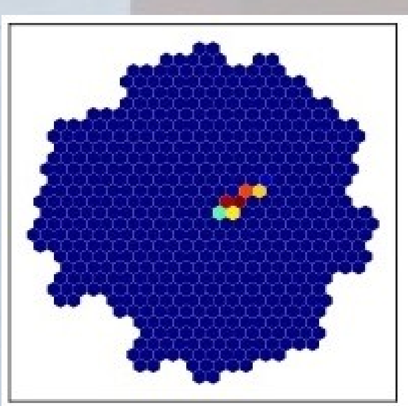
# Outline

- Ground-based experimental TAIGA
  - IACT – Imaging Atmosphere Cherenkov Telescope
  - HiSCORE - Wide-angle Cherenkov detector array
- Main tasks of experimental facilities
- Traditional methods in gamma astronomy (Hillas parameters)
- HiSCORE: EAS parameters
- IACT+HiSCORE: Energy of primary particles
- Conclusion

# TAIGA complex

Charged cosmic rays and high energy gamma rays interact with the nuclei of the atmosphere. The result is extensive air showers (EAS) of secondary particles emitting Cherenkov light. Imagine Atmospheric Cherenkov Telescopes (IACT) register the light.

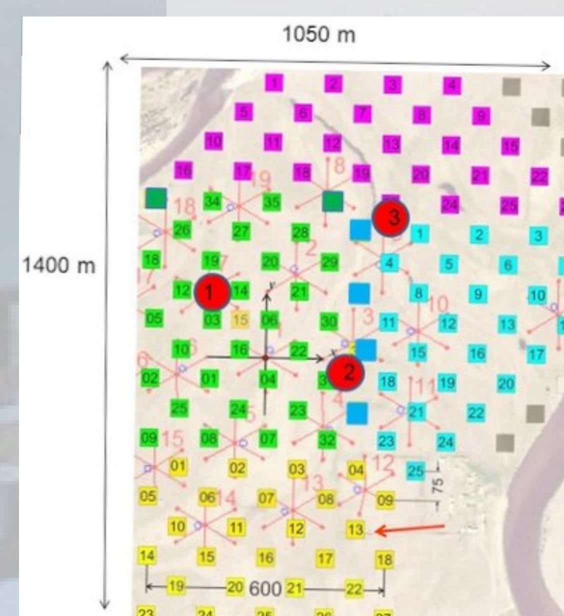




Detected data form "images" of the air shower

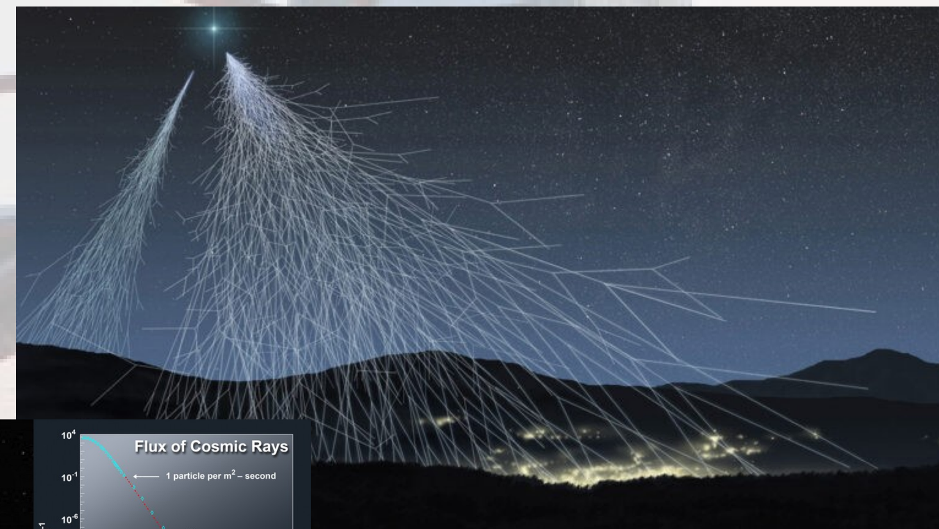
# HiSCORE

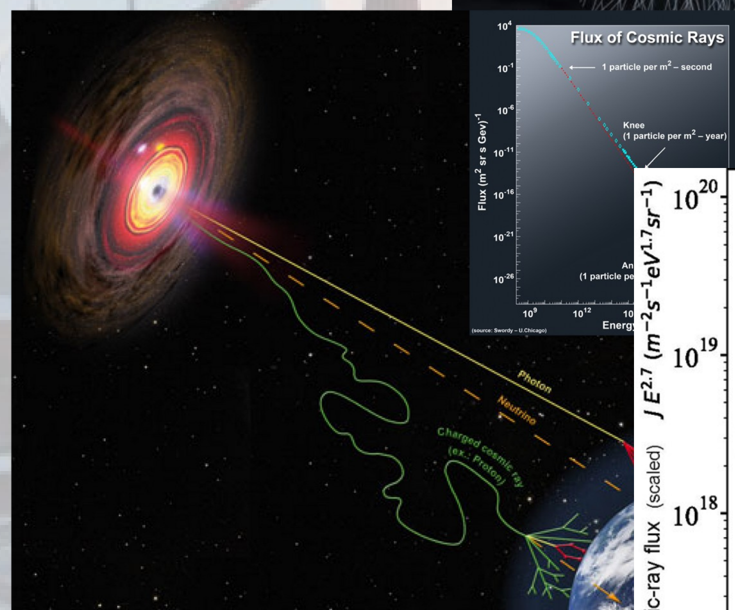
- TAIGA-HiSCORE is an array of integrating Cherenkov detector stations with a wide Field of View (FoV ~0.6 sr). Each detector station consists of 4 large area PMTs (20 cm and 25 cm in diameter), located next to each other and equipped with a Winston cone to increase the effective light collection area by 4 times.
- The detector stations are placed at distances of 150-200m from each other. The ~100 stations of the array cover an area of ~1 km² (which, at a later phase of the experiment, can be extended to ~10 km²).

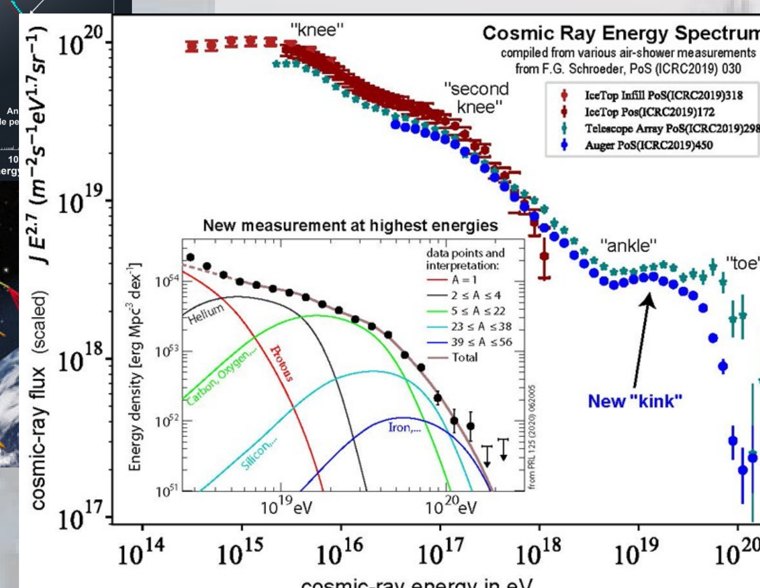


# Main tasks

- Identification of primary particle
  - Gammas, hadrons, nuclei (He, CNO, Fe, ...)
- Mass spectra
- Energy spectra
  - Gammas, hadrons, nuclei

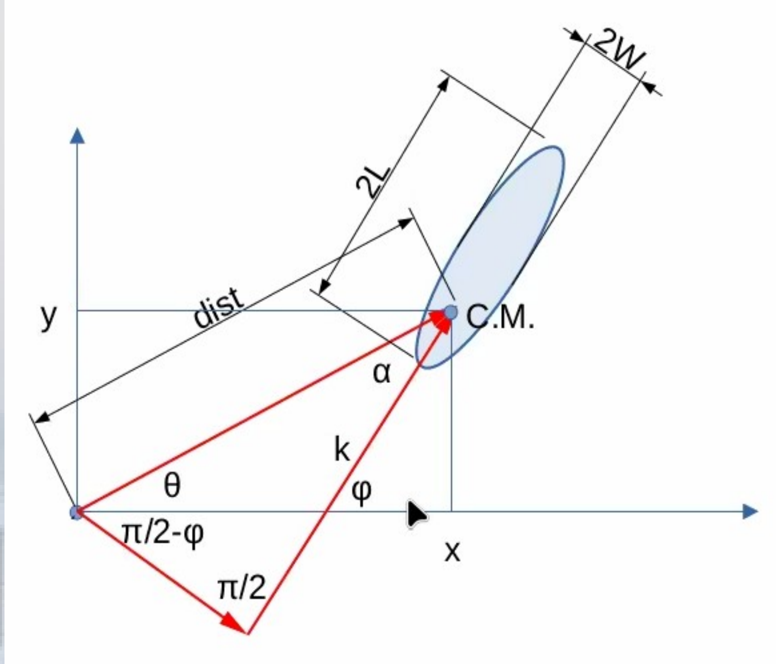




# IACT image and Hillas parameters

For each event image we can calculate the so-called Hillas parameters, which form a set of geometric features of the image




These parameters are widely used in gamma-ray astronomy for gamma/hadron separation

The most important Hillas parameters are:

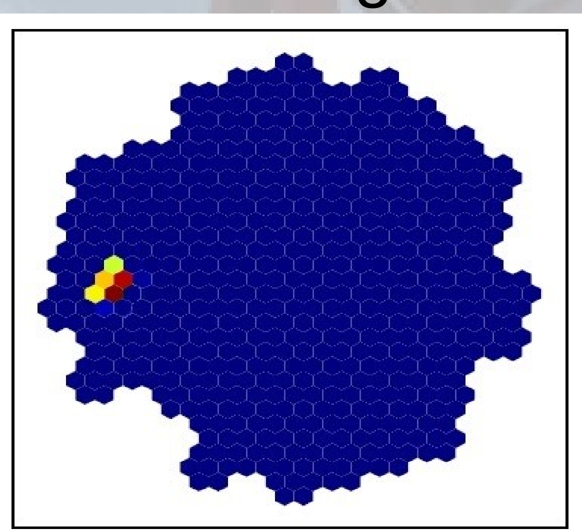
- Image brightness (called image size)
- Width and length of the ellipse
- Number of triggered pixels
- Distance
- Angles: alpha, phi, theta

The key parameter is the energy of the primary particle (can not be directly calculated but mainly correlates with the image size and distance)

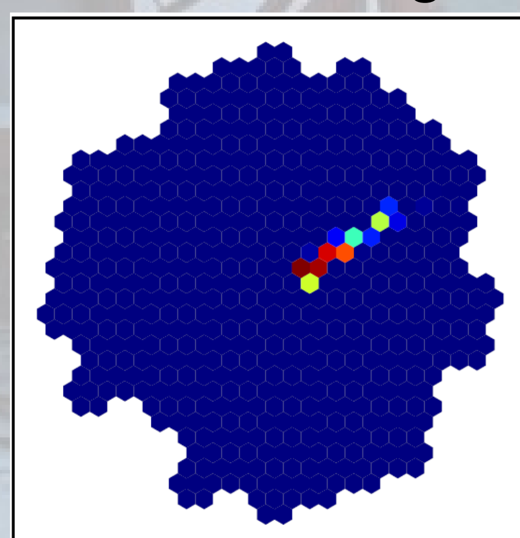
# IACT: Gamma vs. hadron images

- High energy gamma rays
  - the particles of interest (0.01% of all particles)
- Hadrons background (mostly protons)

Proton image

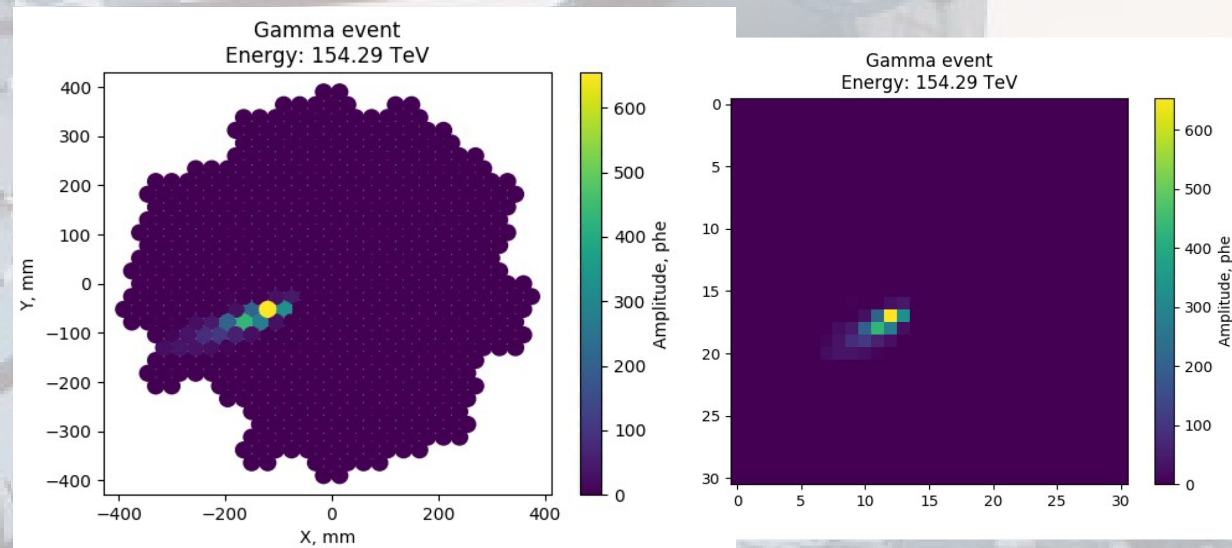


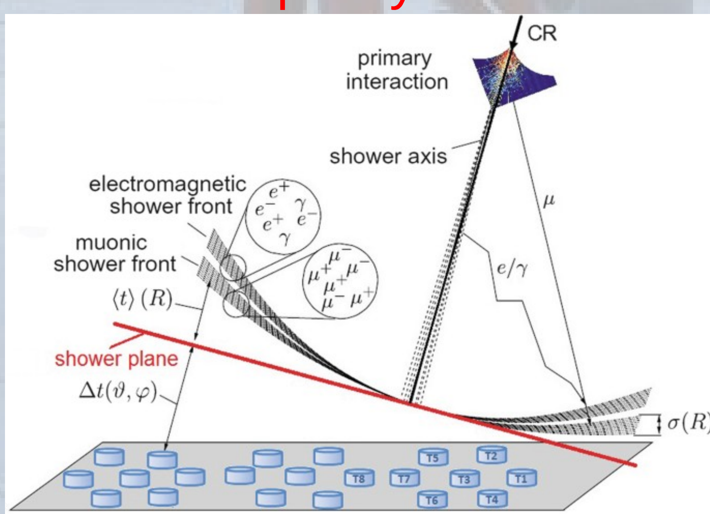
Gamma image

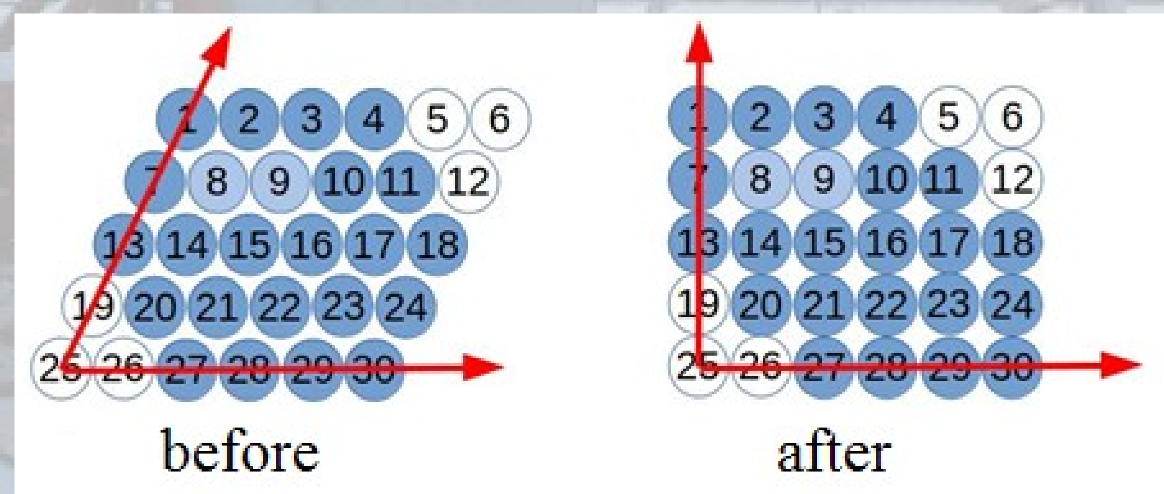


# Python frameworks: TensorFlow and PyTourch

- TensorFlow and PyTorch work with squared matrix
- There are technics of transformation:
  - approximation
  - re-bining
  - oblique system of axes

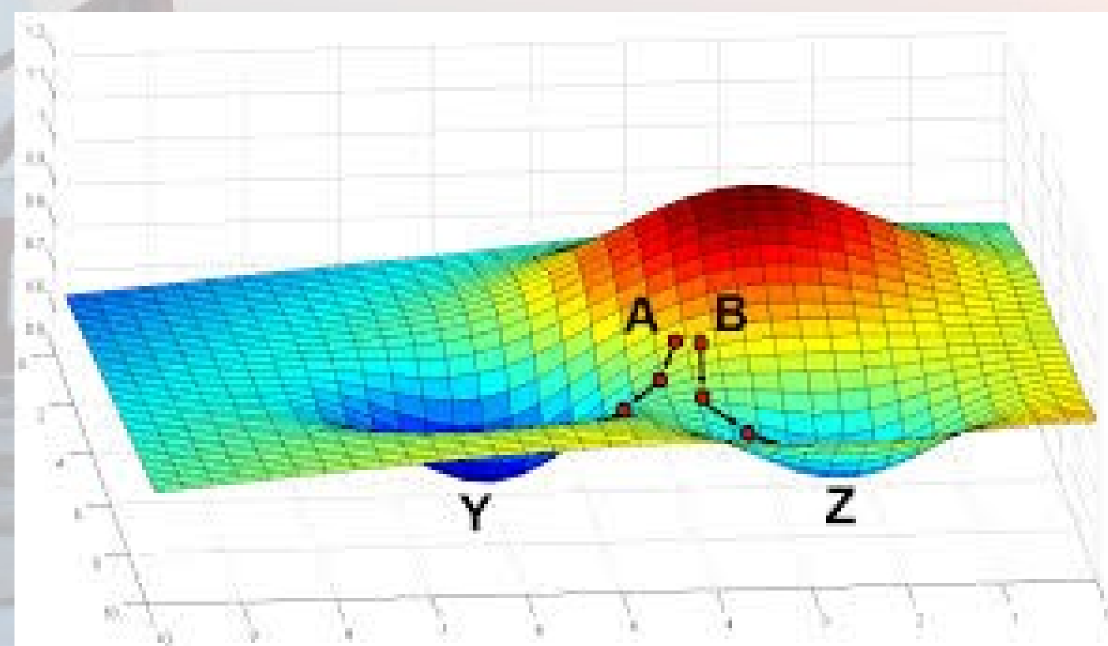


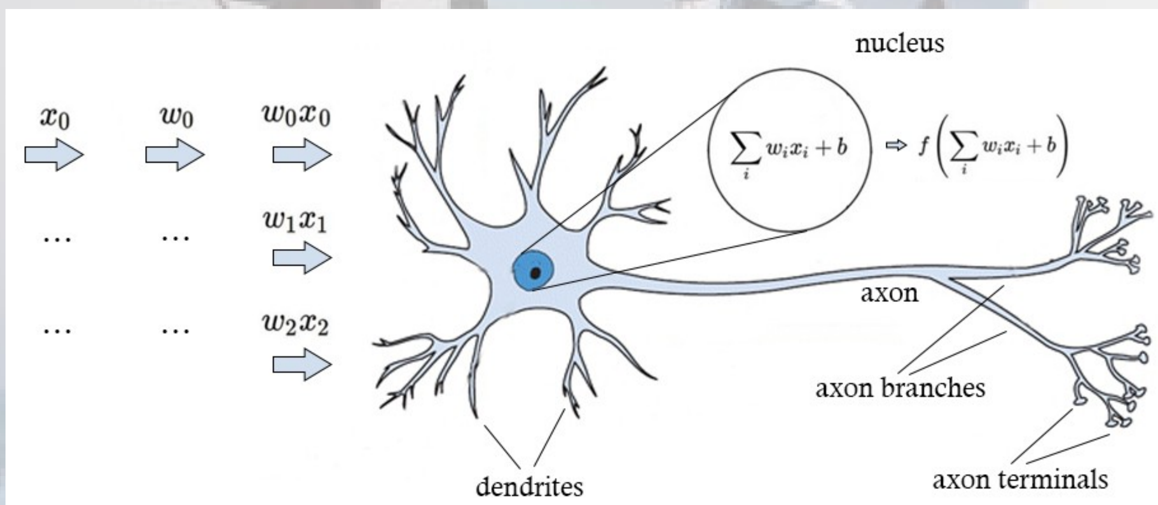
# General approach

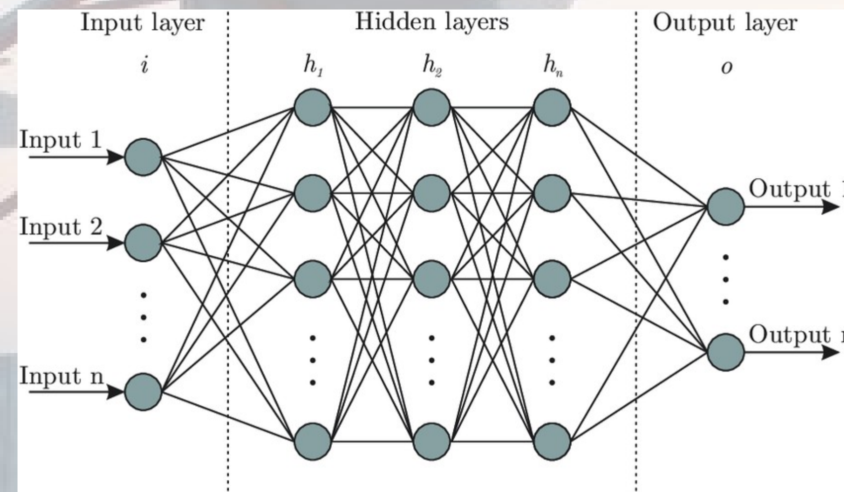
- $F(x)=0 \Rightarrow x=x_0$
- $F(x) \Rightarrow f=F^2(x)$: $\min_{x=x0} f(x)$
- Gradient descent: $x_{n+1}=x_n+\eta\nabla f(x_n)$
- f(x) → f(x,w), for example $f\in P_n(x)$
- f(x) → ANN(x,w)

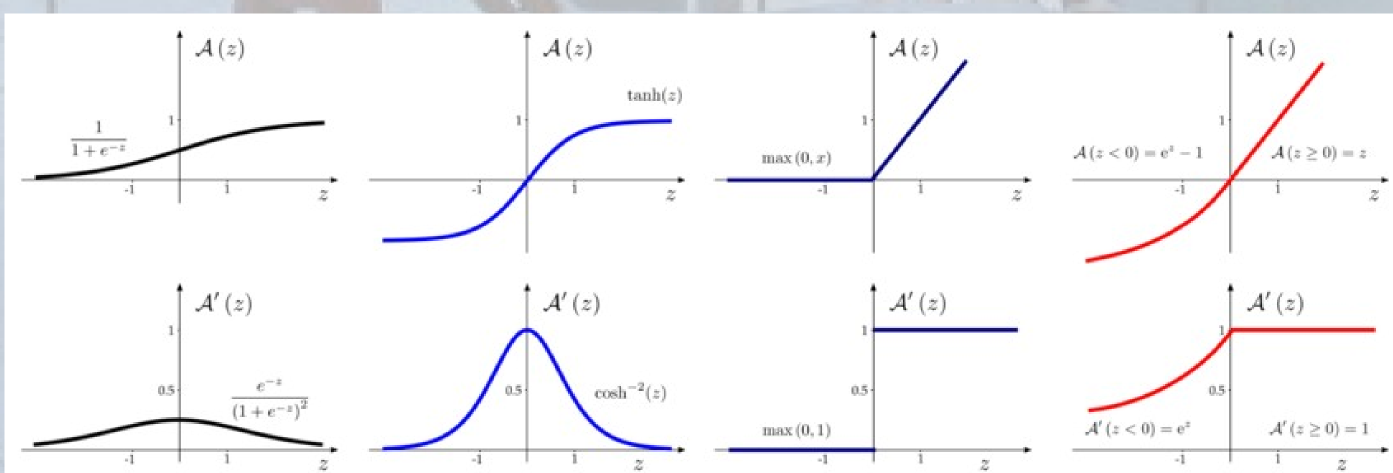


# ANN for beginners

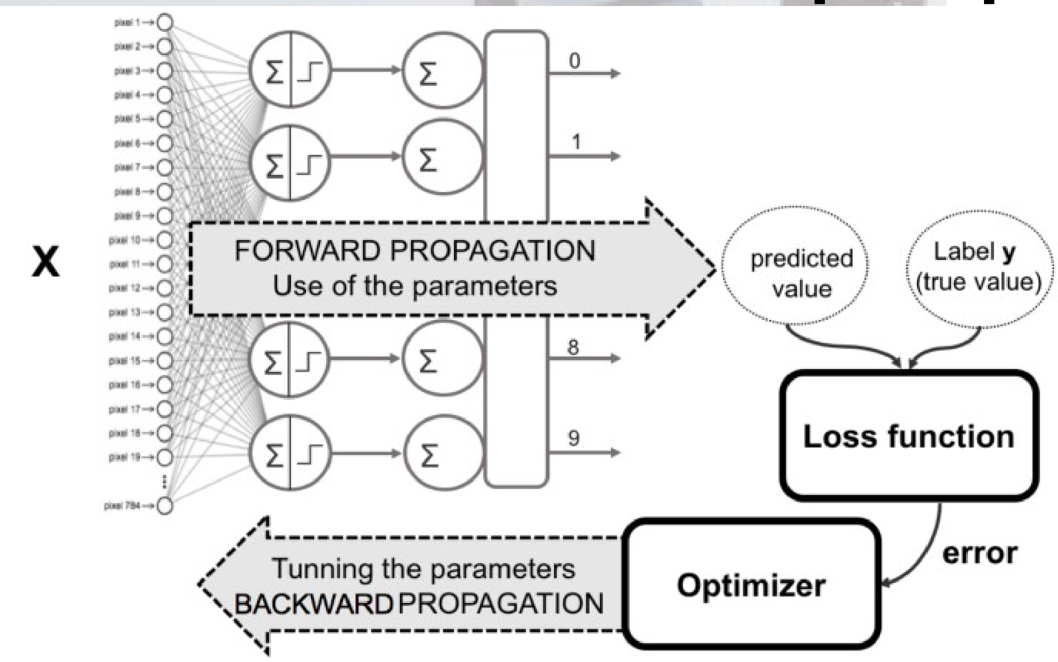
- The structure of the artificial neuron was inspired by the natural neuron in the brain.

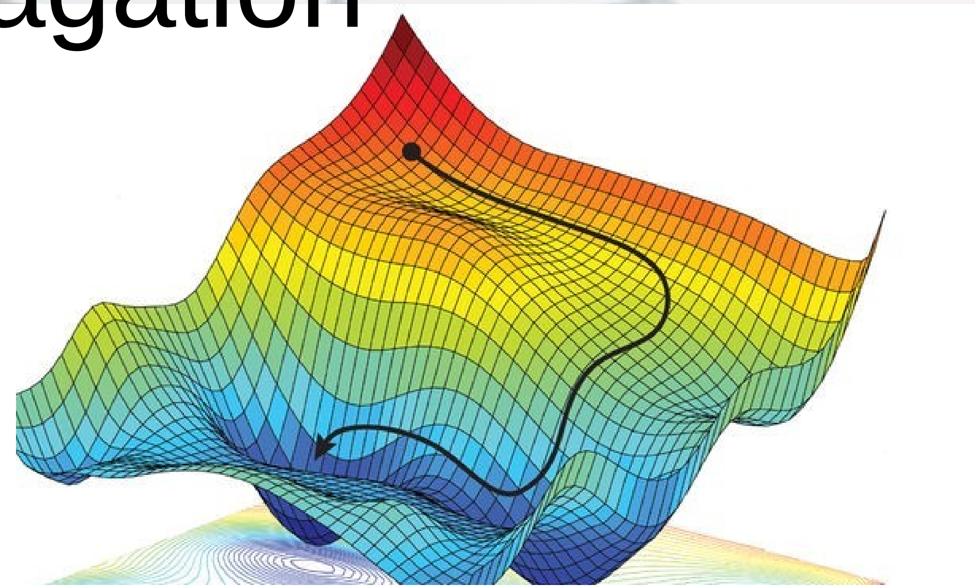


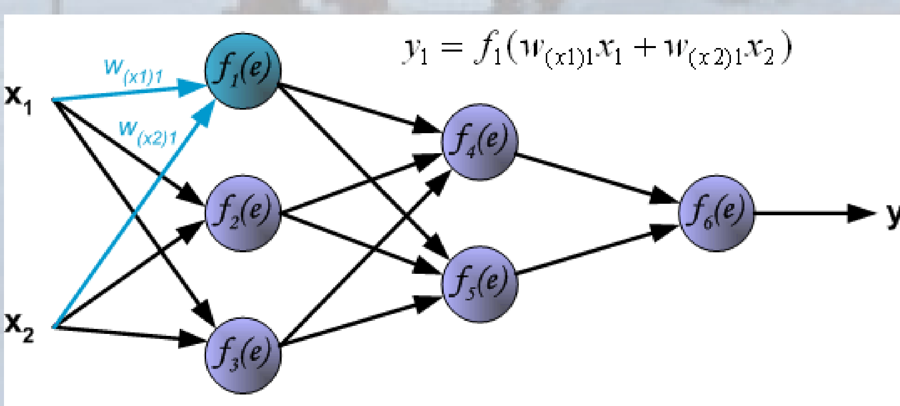



# Forward and back propagation



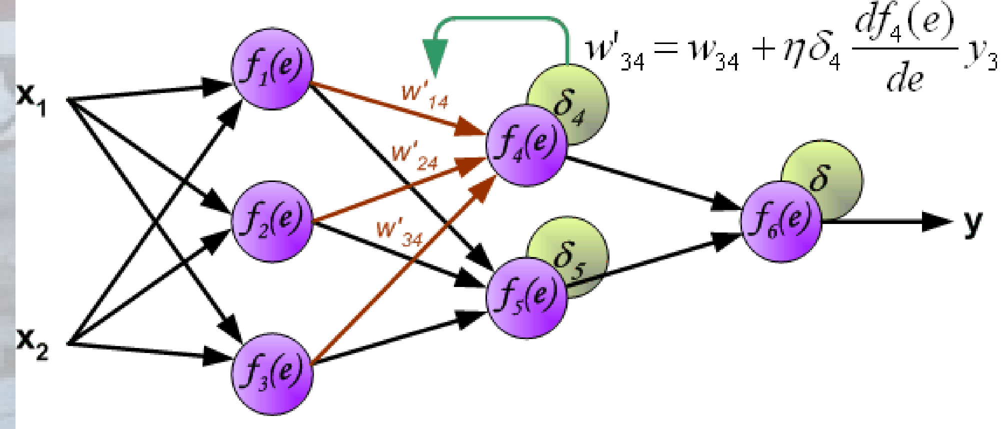






$$w'_{14} = w_{14} + \eta\delta_4 \frac{df_4(e)}{de} y_1$$

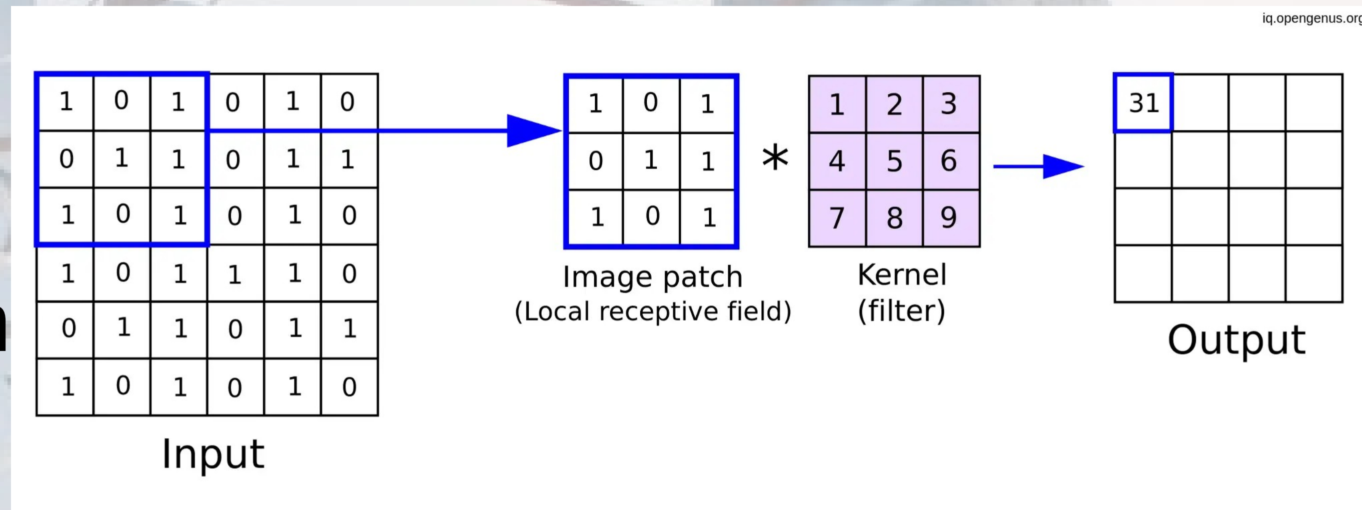
$$w'_{24} = w_{24} + \eta\delta_4 \frac{df_4(e)}{de} y_2$$

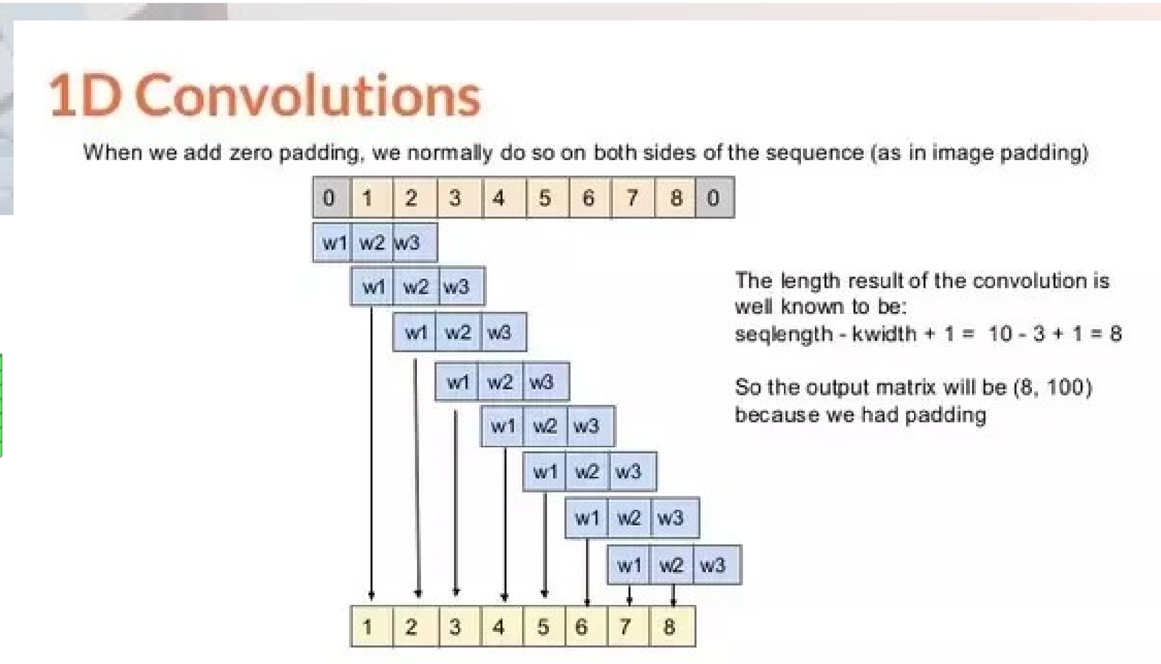
$$w'_{34} = w_{34} + \eta\delta_4 \frac{df_4(e)}{de} y_3$$

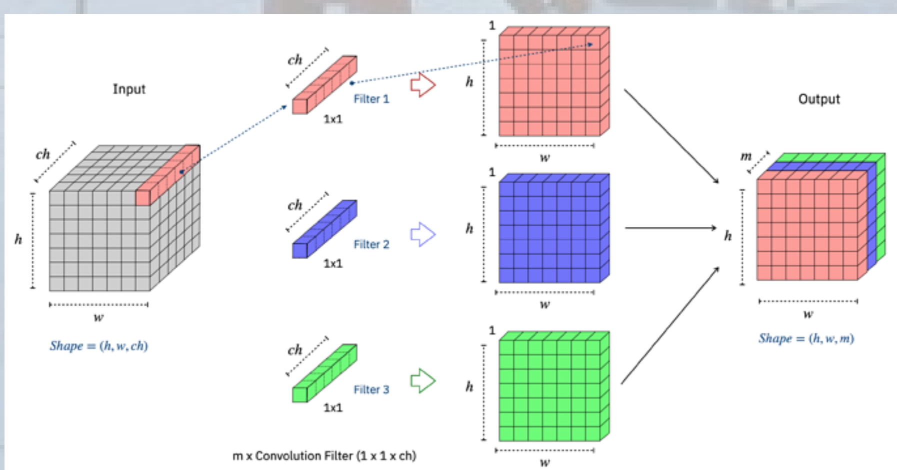


# Convolutional ANN

- Convolution
- 1D convolution
- 1x1 convolution

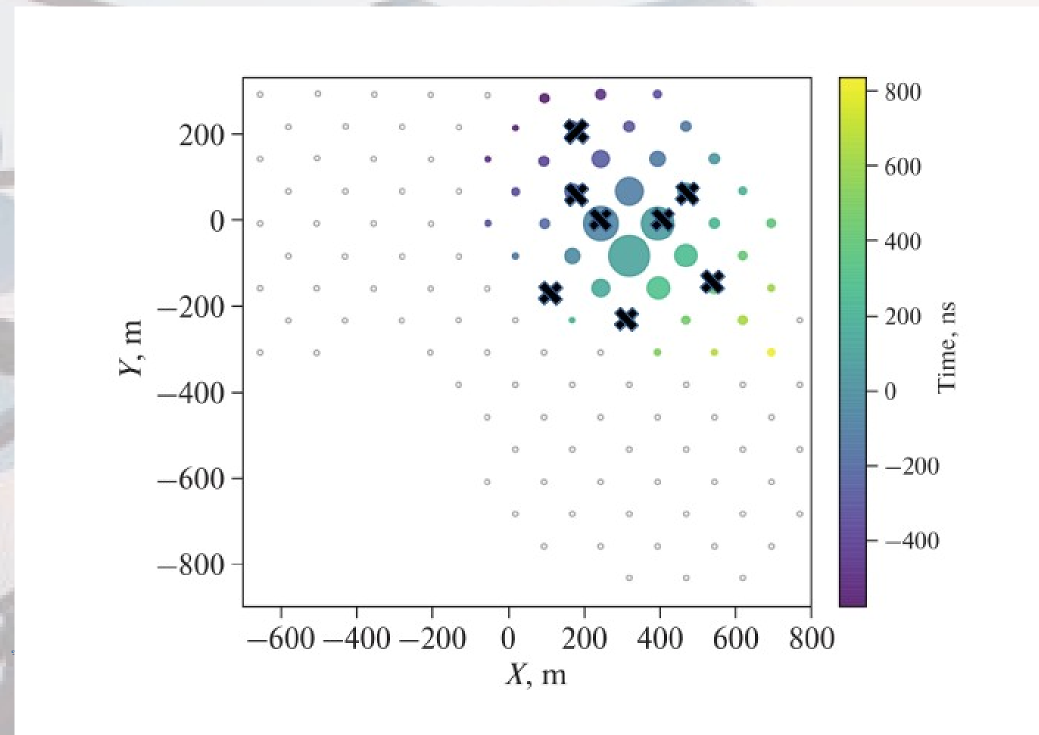


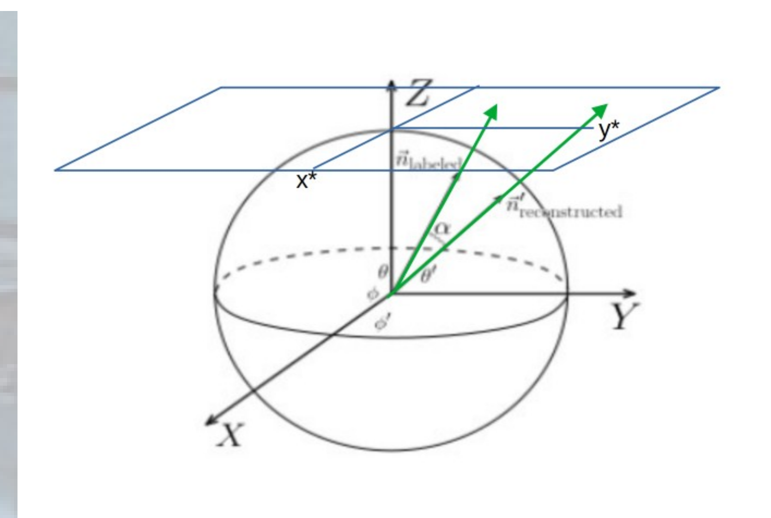



# Training set

- Для обучения используются 74419 протонных событий, с минимумом 10 сработавших станций и 300 ф.э. на станцию. Среднее число сработавших станций на событие 16,9.
- Из данных формируются входные векторы для сети, содержащие K=8 наборов по 7 значений для выбранных станций: T (нс), √Q, log10 Q, (для значений Q в ф.э.), s.d. T (нс), x, y, z (м).
- Станции всегда отсортированы по возрастанию T и сдвинуты так, чтобы у первой станции T, x, y, z были нулевыми.
- В качестве меток используются координаты x*, y* точек пересечения вектора направления оси ШАЛ (с началом в начале координат) с горизонтальной плоскостью z=100.



# Structure ANN

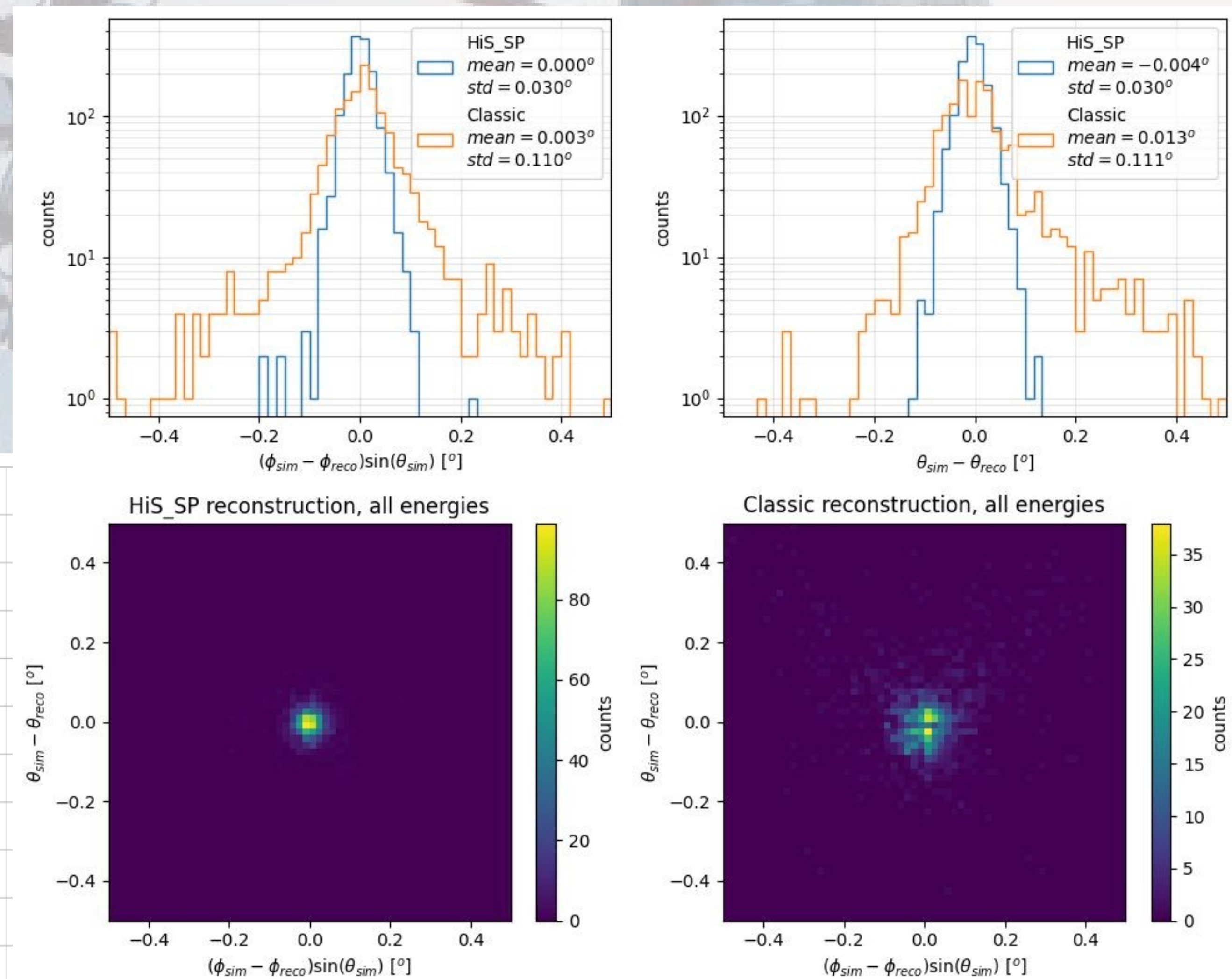
- The station set size is $K = 8$, the data is organized into 7 channels, for a total of 56 input values.
- A 1D convolutional network with a filter size of 2 is used.
  - The first layer is separate, followed by 16 residual learning blocks, each with 2 layers of 512 channels, followed by a convolutional layer with a filter of 1 to reduce the number of channels to 32 (also residual), and an output fully connected layer with 2 values.
  - Within the res blocks, channel multiplication is performed: if we have 2c channels, then for all $i \le c$, we find the product of channel values i and i + c and concatenate these c products with the main 2c channels. BatchNorm and activation are then applied to all 3c channels.
- SiLU (Sigmoid Linear Unit: $x/(1+e–x)$) activation is used, except for output neurons, which have linear activation.
- A total of ~21 million trainable parameters are used.

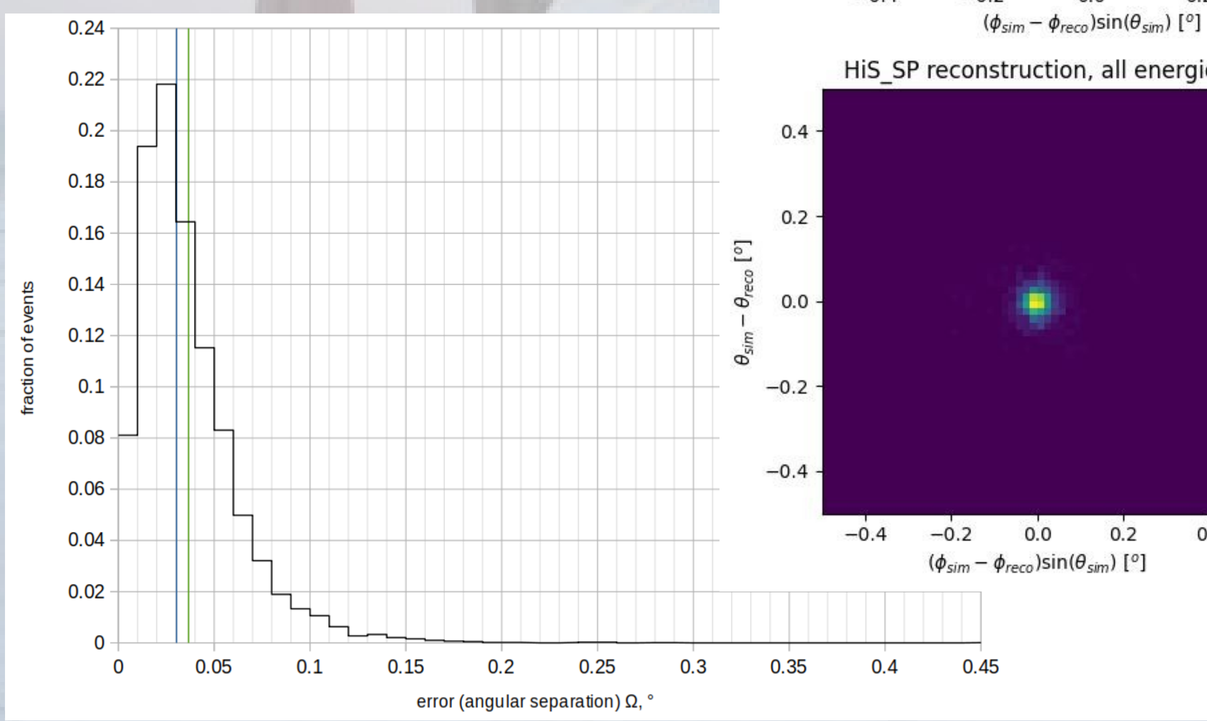
# Training a neural network

- 124,416 training examples are used per epoch. Every two epochs (at the end of each training epoch), random disjoint subsets of size K from the triggered stations are reselected for each event.
- On odd epochs (except at the end of training), a reflection augmentation is performed: the sign of the y-coordinate for all stations and the azimuth angle φ are flipped.
- 1000+20 epochs were trained, with a learning rate of 0.001 except for the last 20; in the last 20 epochs, it decreases exponentially by a factor of 1024. The loss function is MSE.

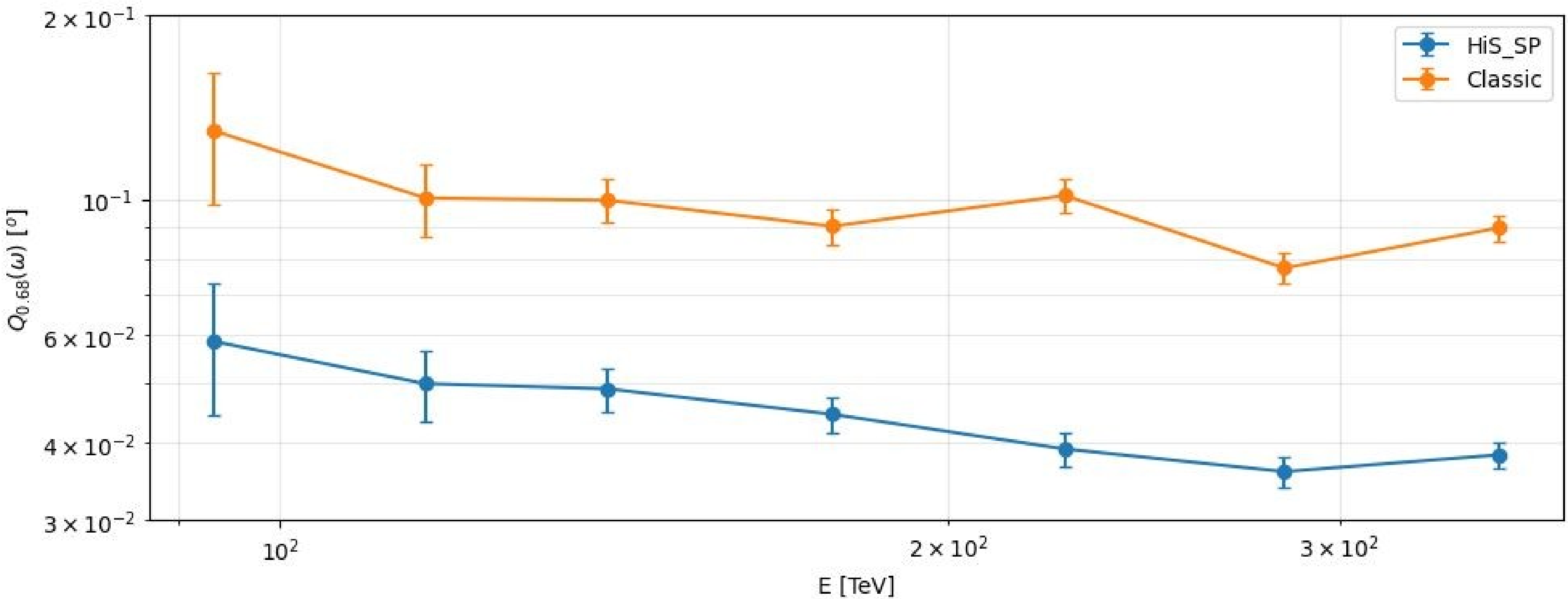
# Preliminary results

- The average error ω of the trained network is 0.057°.
- For M=10, the value of Ω10 is 0.0392°.

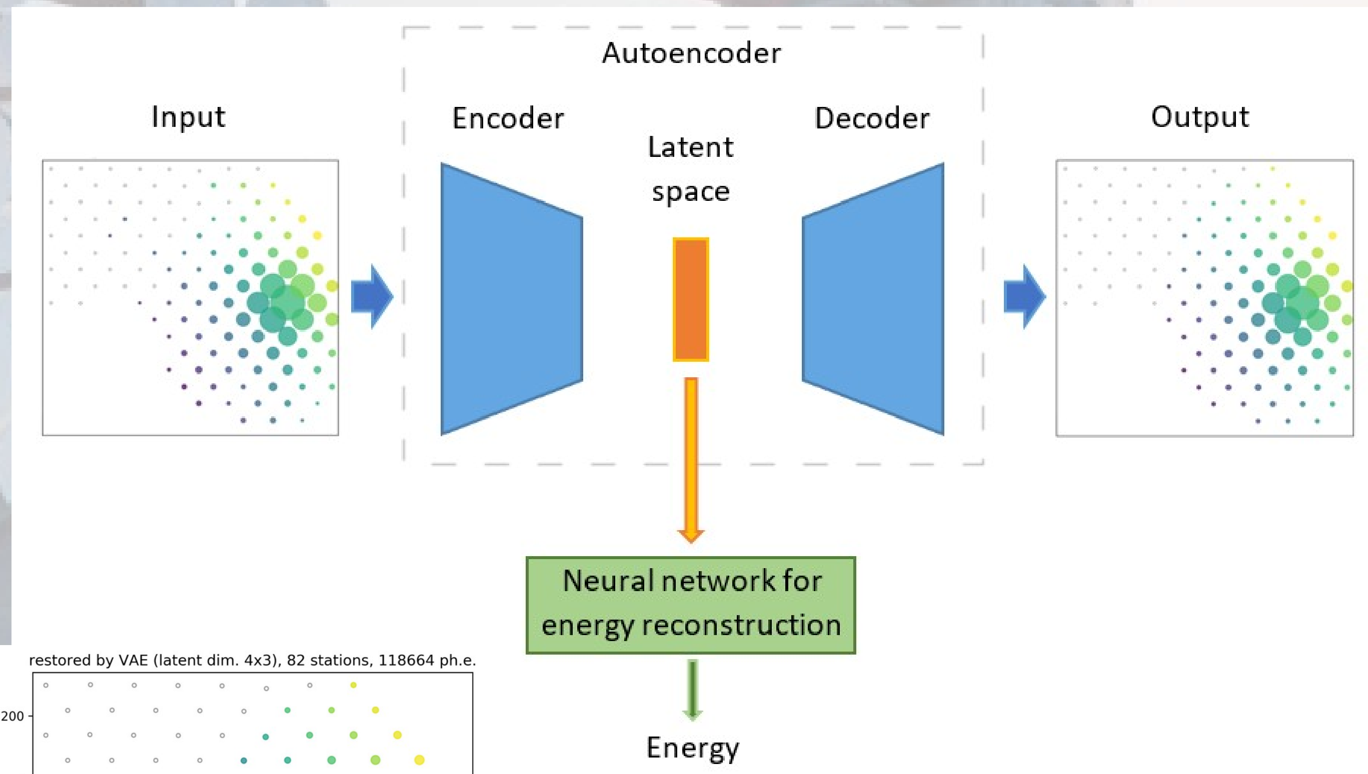


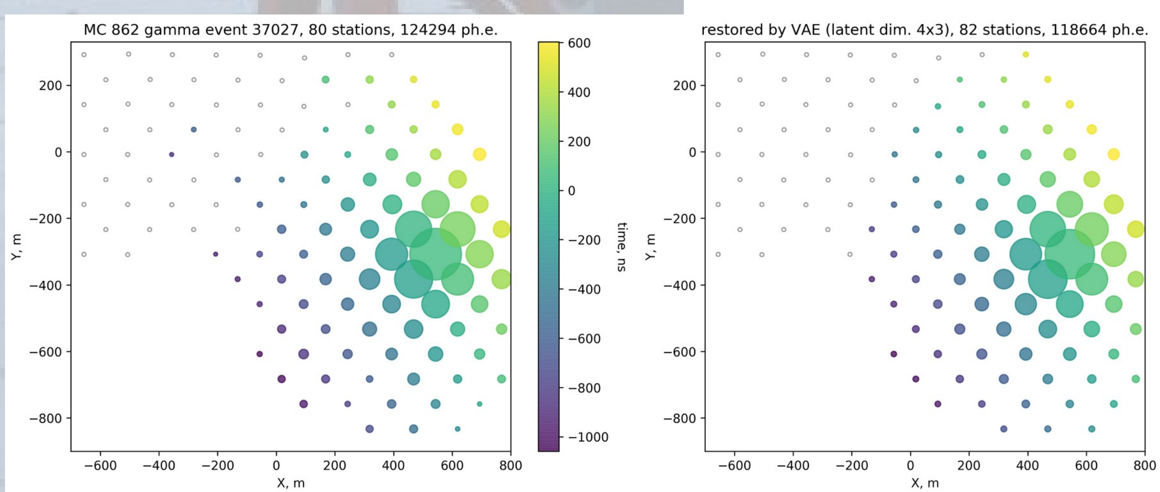



# Preliminary results



# AE as the tool for extracting essential features

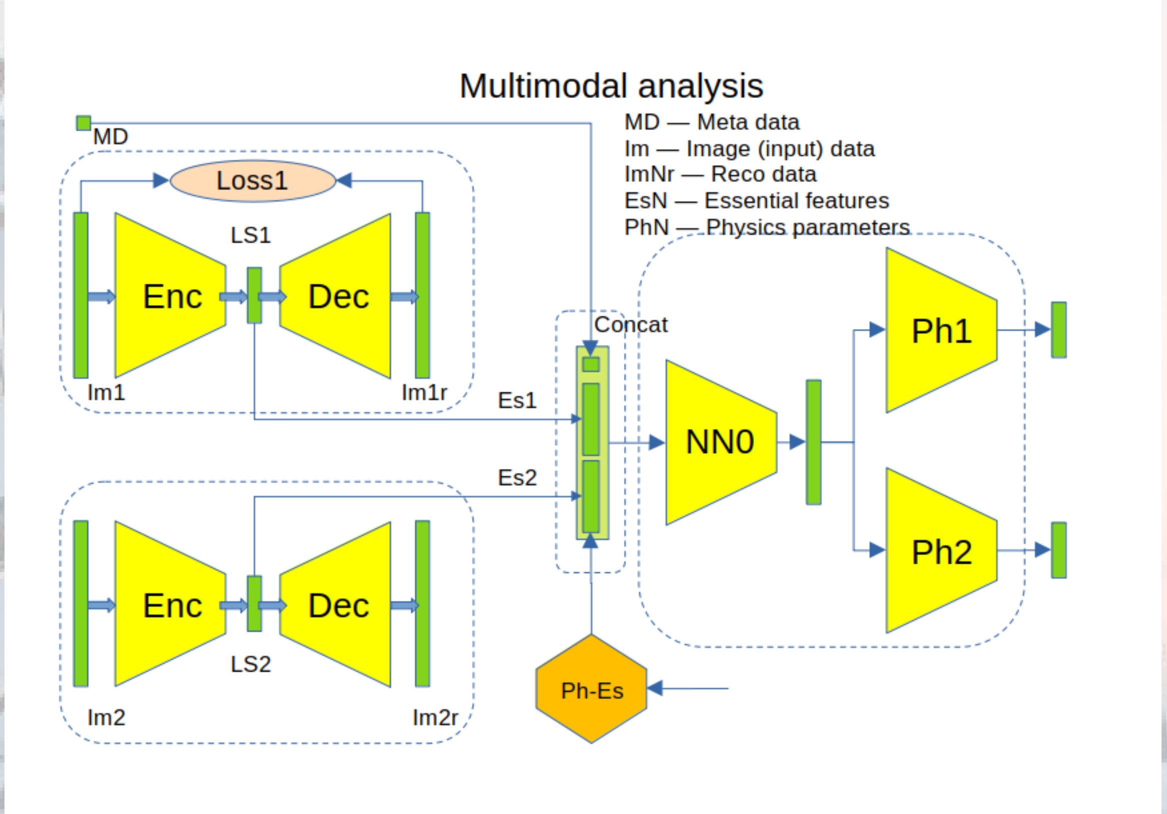


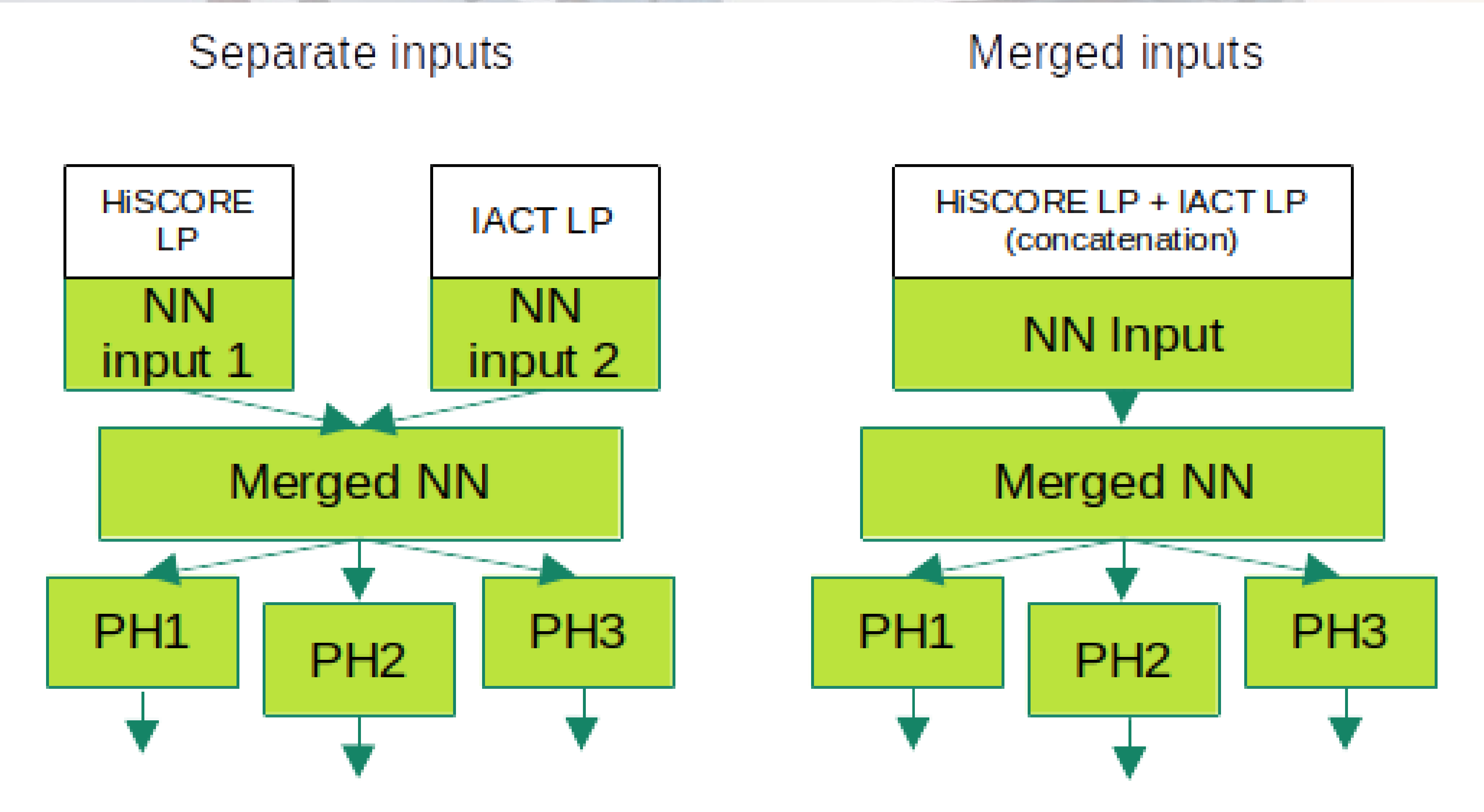
# Multimodal analysis

- AE as extractor of essensial features
  - Alternative to Hillas parameters
- Multimodal analisys:

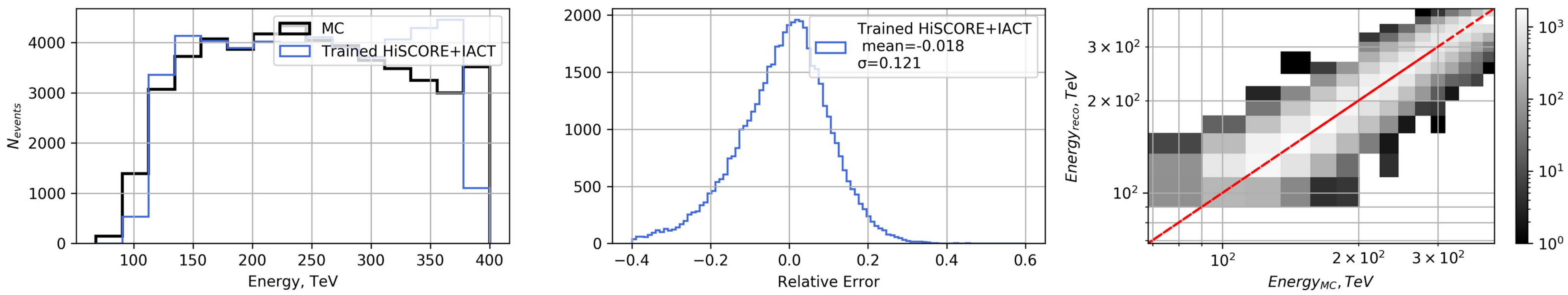
Joint analysis of data from several types, in particular from installations of different types (for example, IACT+HiS)

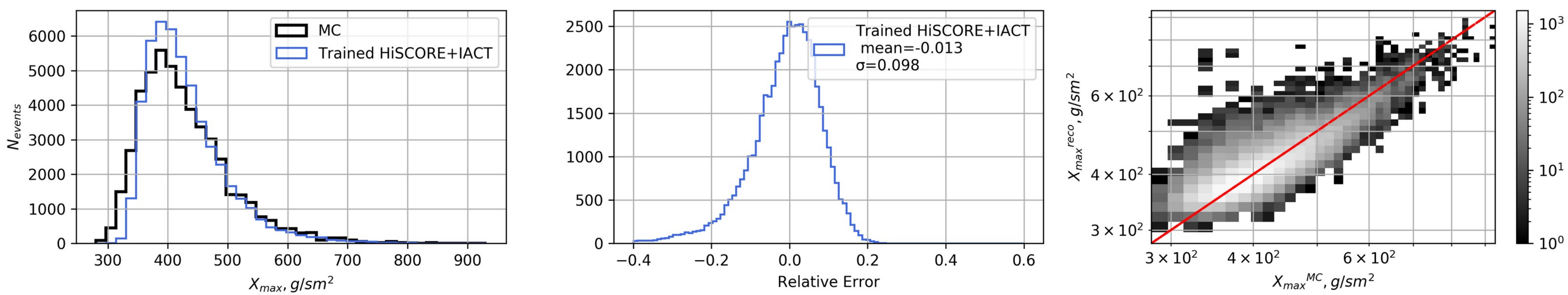


# Real structure of ANN



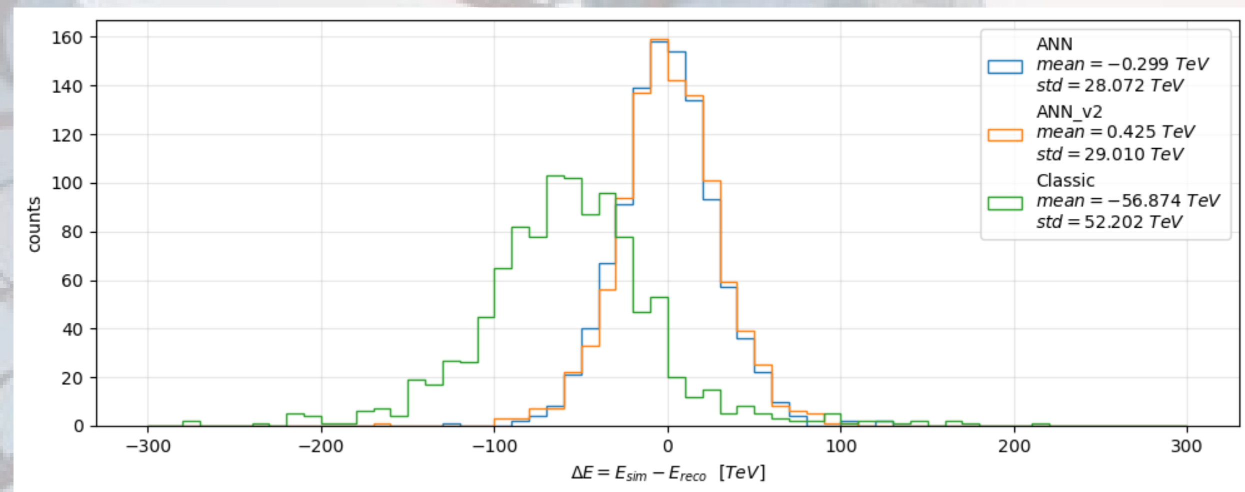
# Energy and Xmax (preliminary)

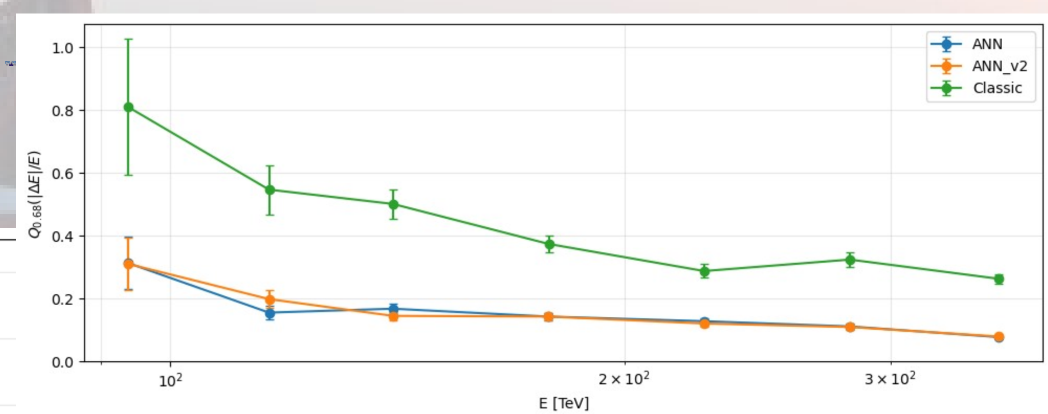


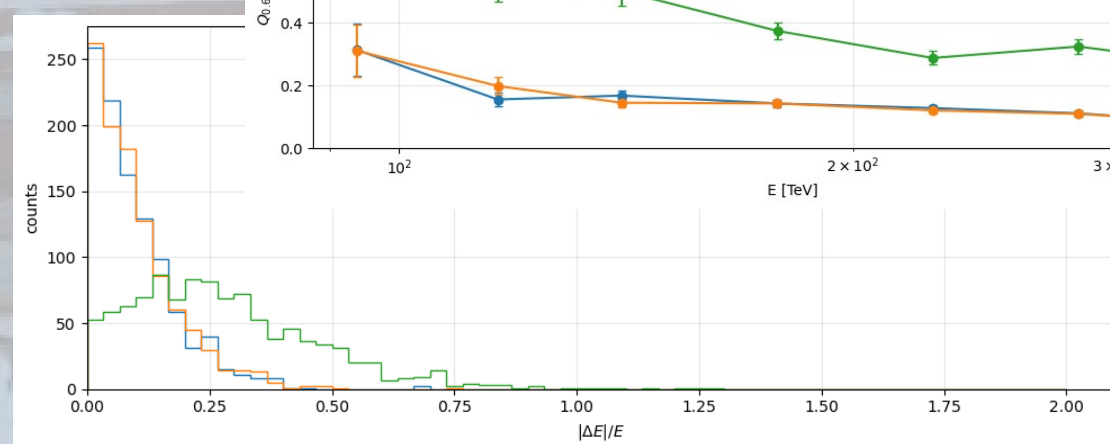



# Comparison with traditional method

- The shift of mean is virtually zero.
- The standard deviation of the error is 2 times smaller.
- The energy determination accuracy in the 40-400 TeV range is two or more times higher.



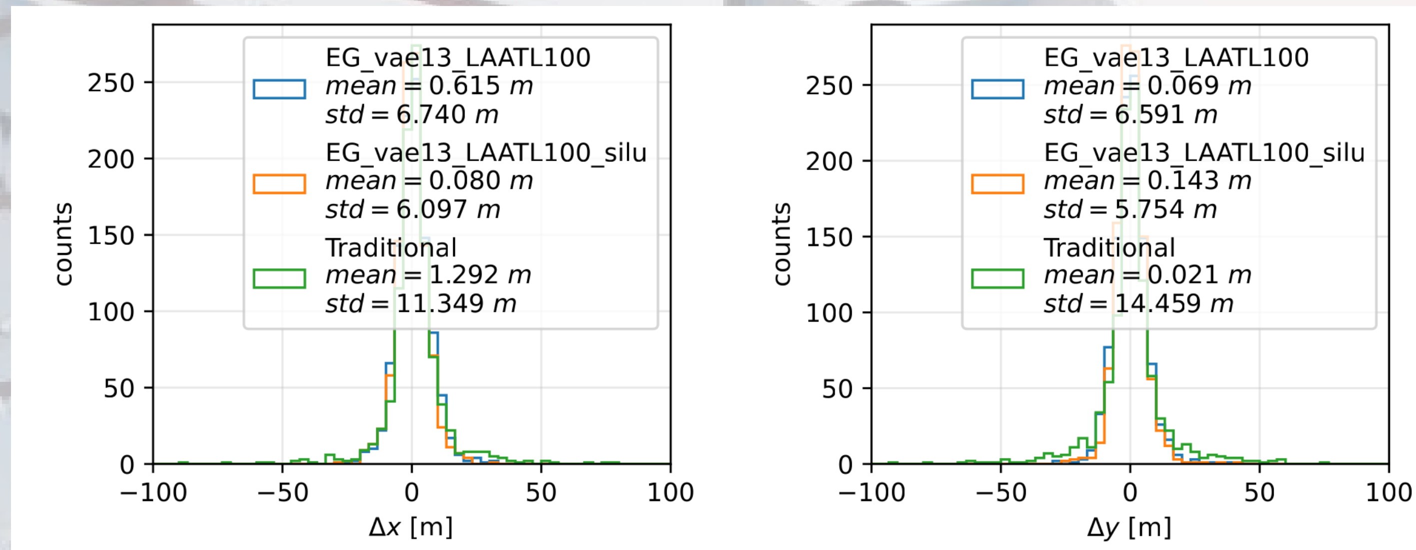



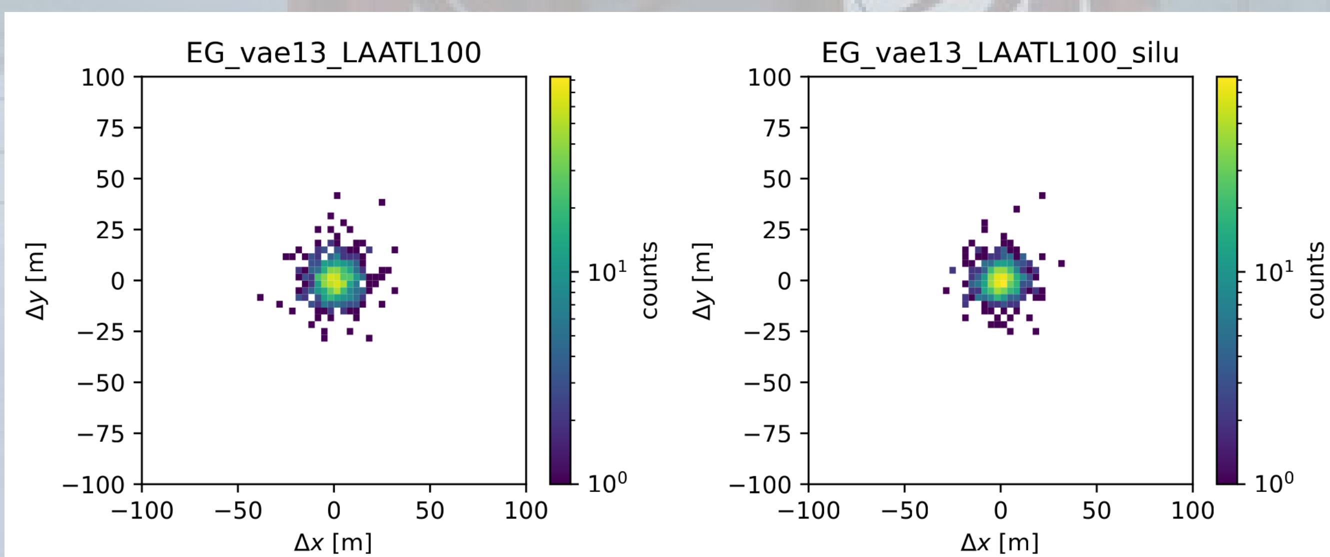



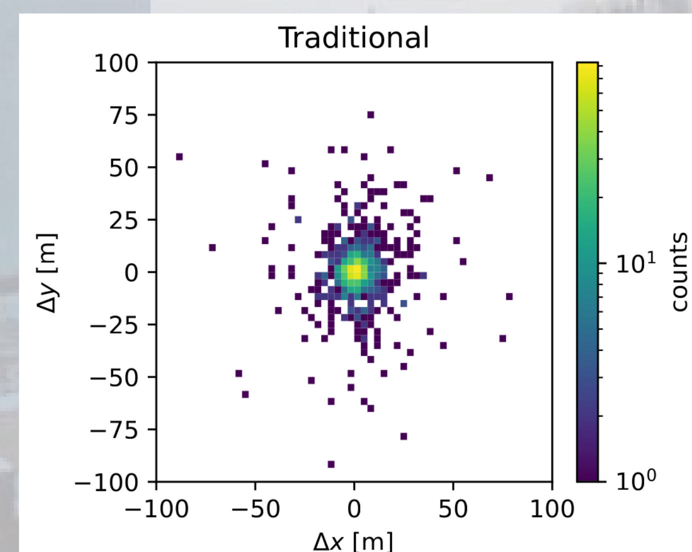
# EAS arrival point

HiS+1*IACT

- sig_Δx=6.7m
- sig_Δy=6.6m
- 2 times better then traditional method.

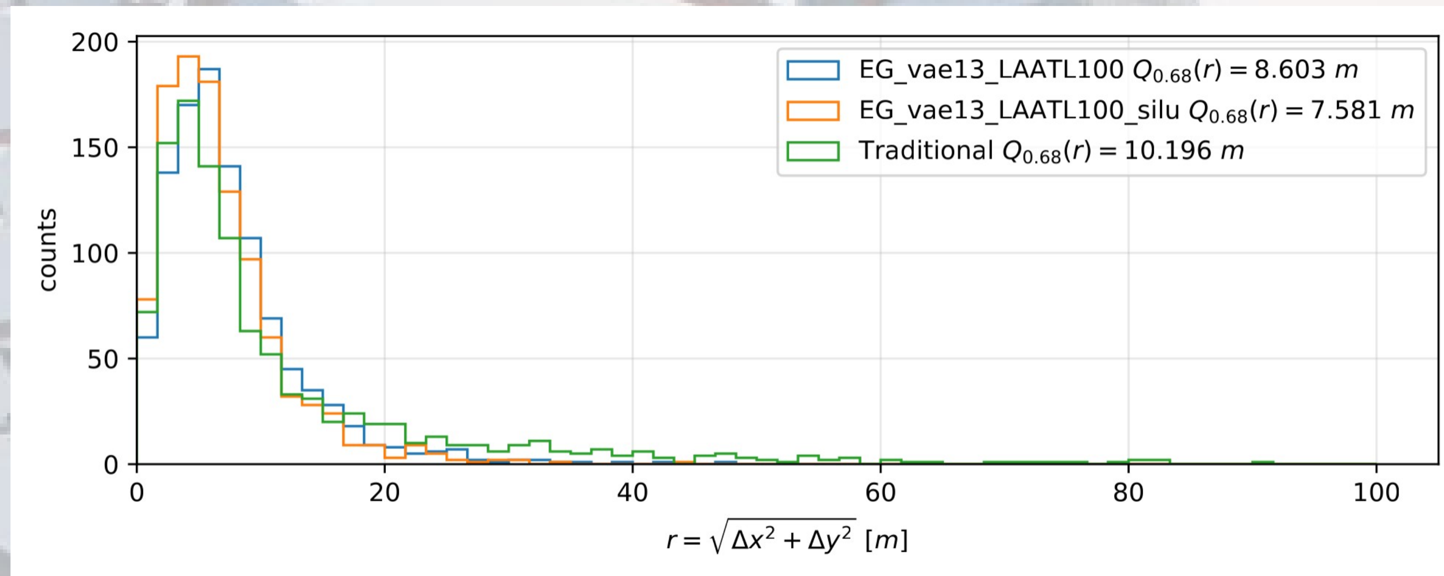


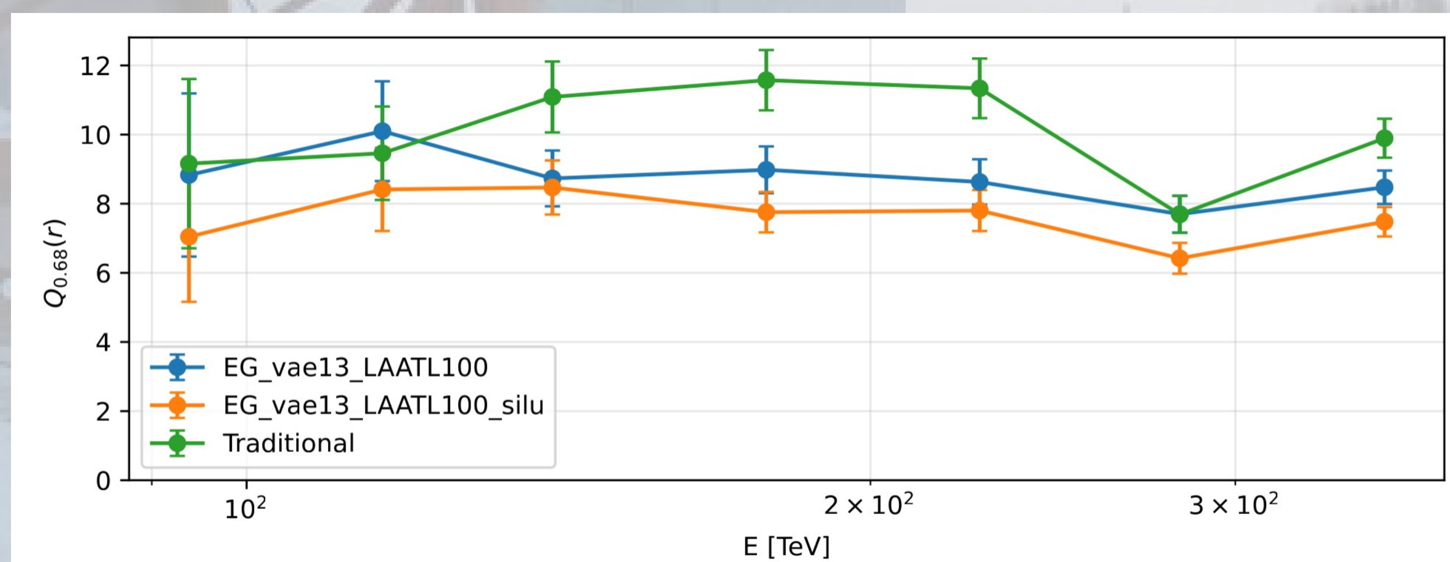



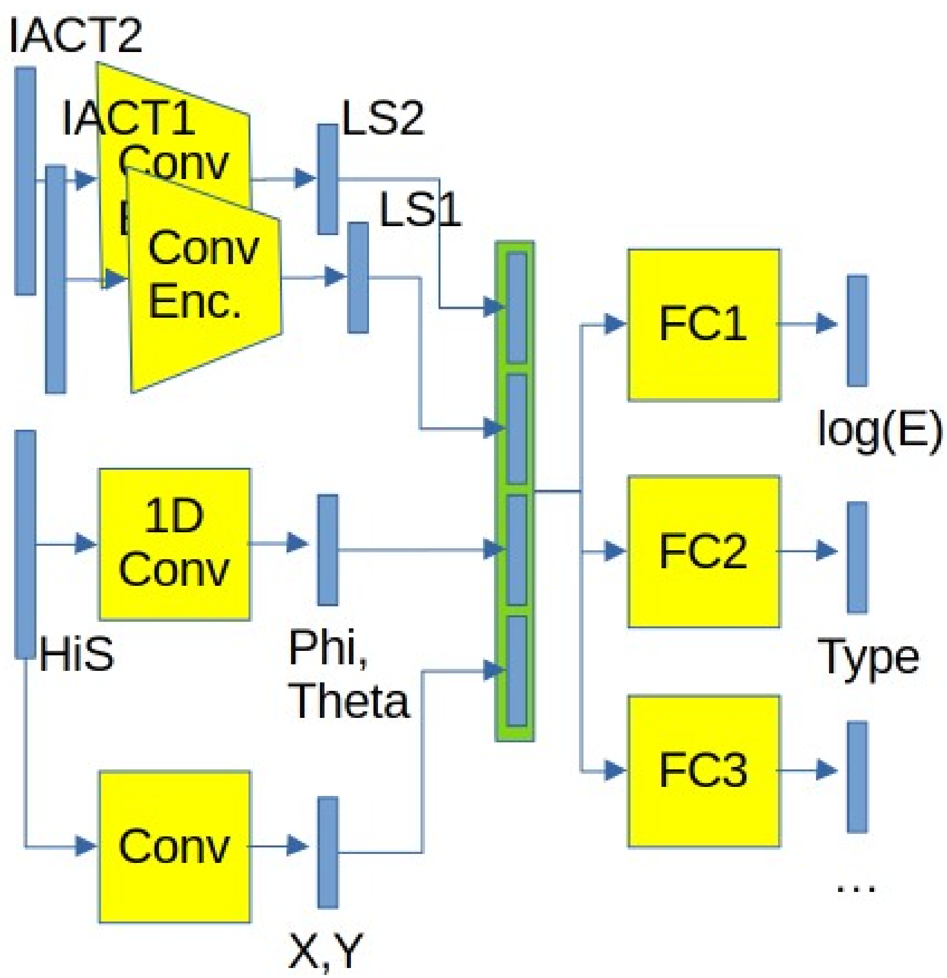
# EAS arrival point

- Distance error (68%)
  - ANN = 8m
  - Trad = 10m



# General structure of neural network




- We combine several IACTs.
- We add the HiS data.
- The output is the energy and particle type.
  - Including heavy elements.
- Other parameters.

# Conclusion

- Machine learning is a powerful tool for analyzing cosmic ray data obtained by ground-based facilities such as HiSCORE.
- A one-dimensional convolutional neural network model is very promising for reconstructing the axis direction of extensive air showers.
- Machine learning allows for more accurate reconstruction of extensive air shower parameters than traditional methods.

# RNF Grant 24-11-00136

- A.Kryukov (head of team)
- J.Dubenskaya
- E.Gres
- E.Postnikov
- A.Razumov
- P.Volchugov
- D.Zhurov

Assisiated members:

- A.Demichev
- S.Polyakov


The author expresses gratitude to all colleagues on the RSF grant, as well as the TAIGA collaboration, for their support and assistance in the work.

The TAIGA IACT located in Tunka valley, Russia