

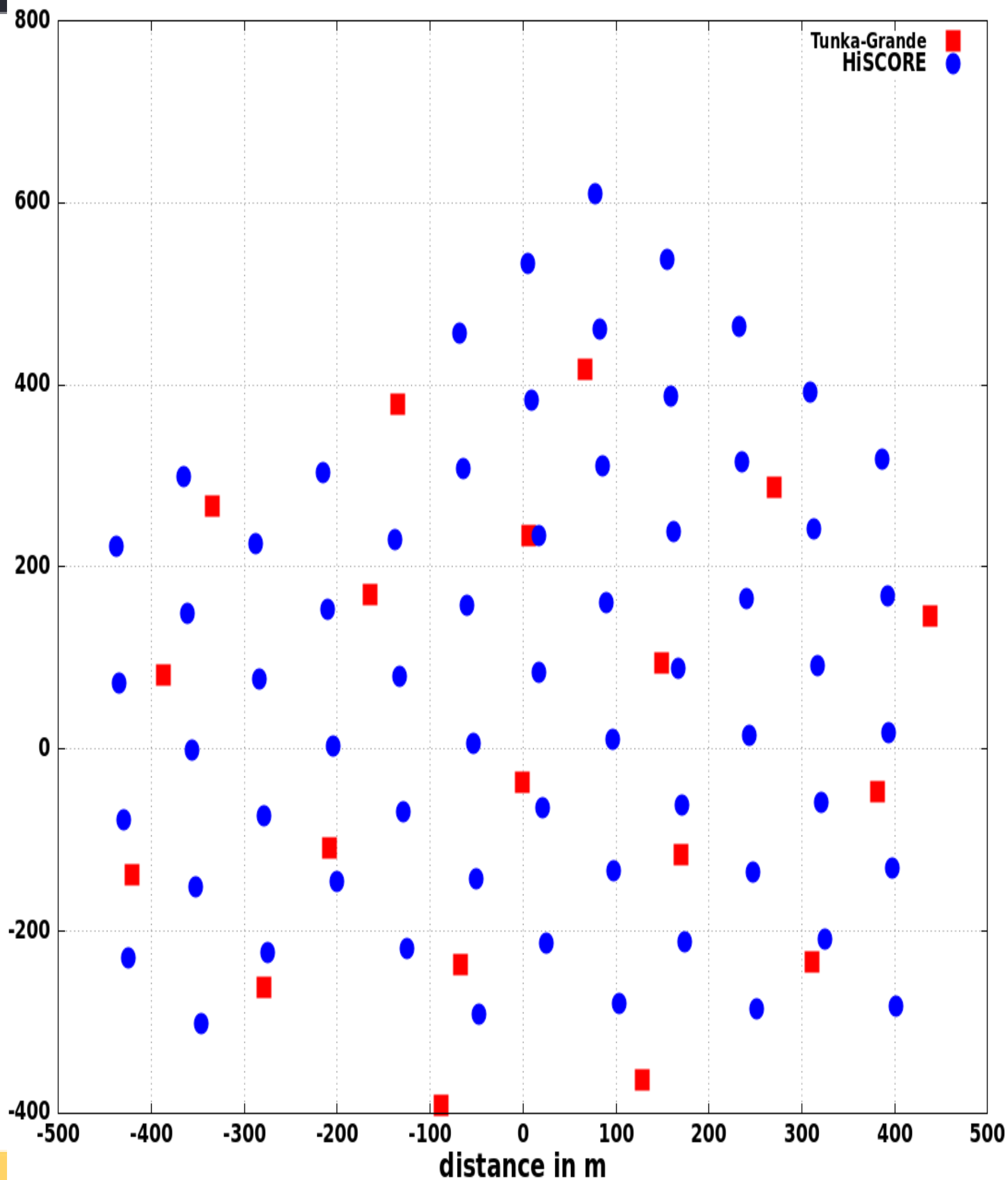
# Метод исследования массового состава космических лучей по данным гамма-обсерватории TAIGA на основе искусственной нейросети

Е.А. Кравченко, А.Вайдянатан

От имени коллаборации TAIGA

Новосибирской Государственный Университет/ ИЯФ СО РАН

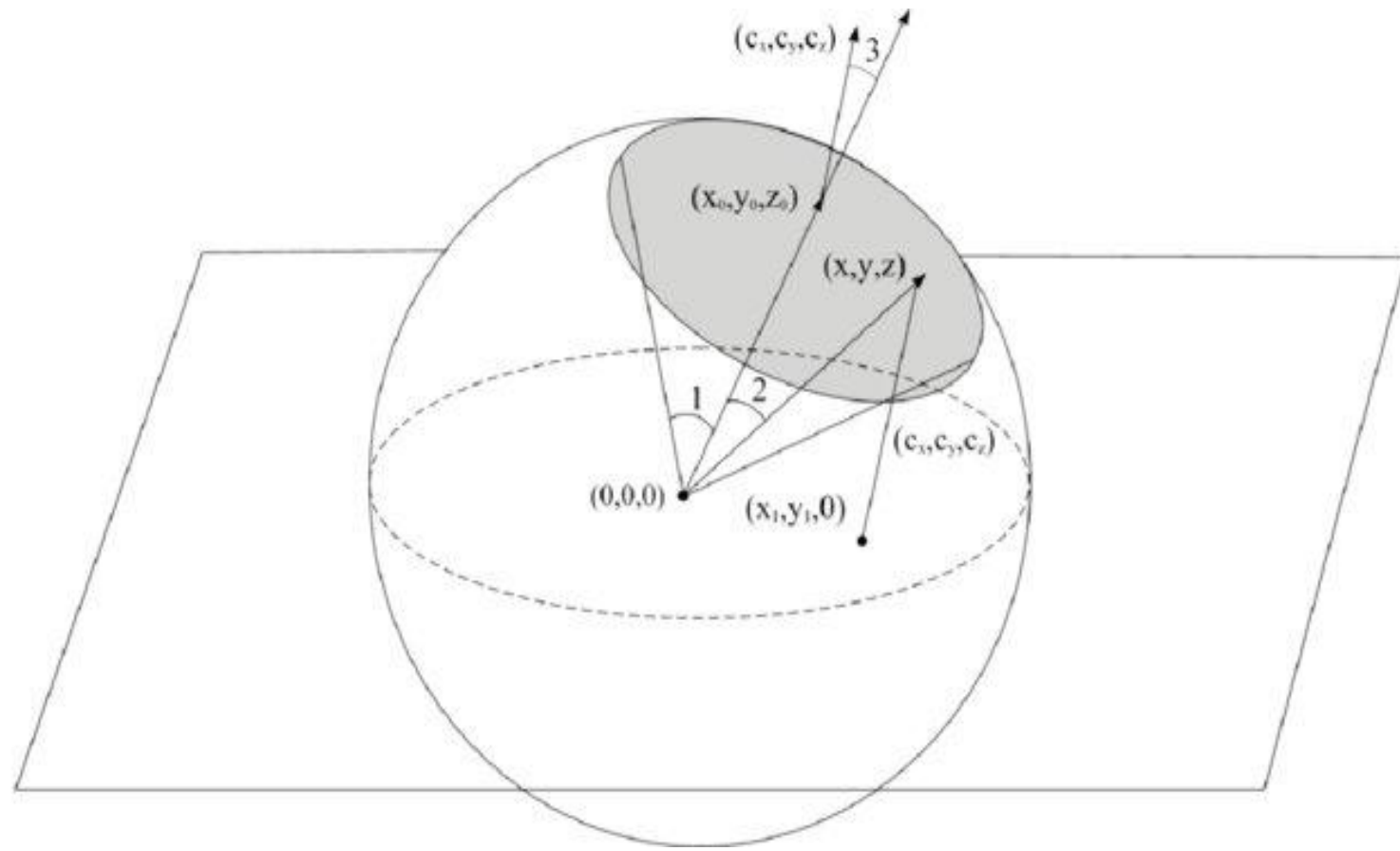
- Основная цель – разработать метод определения элементного состава ПКЛ по совместным данным установок HiSCORE и Tunka-GRANDE в диапазоне энергий 1 – 10 ПэВ.
- Совместное использование данных черенковских и сцинтилляционных детекторов позволит улучшить качество идентификации.



Процедура моделирования включает в себя следующие шаги:

1. Моделирование ШАЛ с использованием пакета CORSIKA
2. Расчет отклика в детекторах HiSCORE с использованием CORSIKA
3. Отбор вторичных частиц на уровне земли с помощью пакета COAST
4. Моделирование взаимодействие вторичных частиц с наземными и подземными сцинтилляционными детекторами установки Tunka-GRANDE с использованием пакета GEANT4

Моделировались 19 станций Tunka-GRANDE и близлежащие 62 станции HiSCORE.



Основные параметры HiSCORE в моделировании:

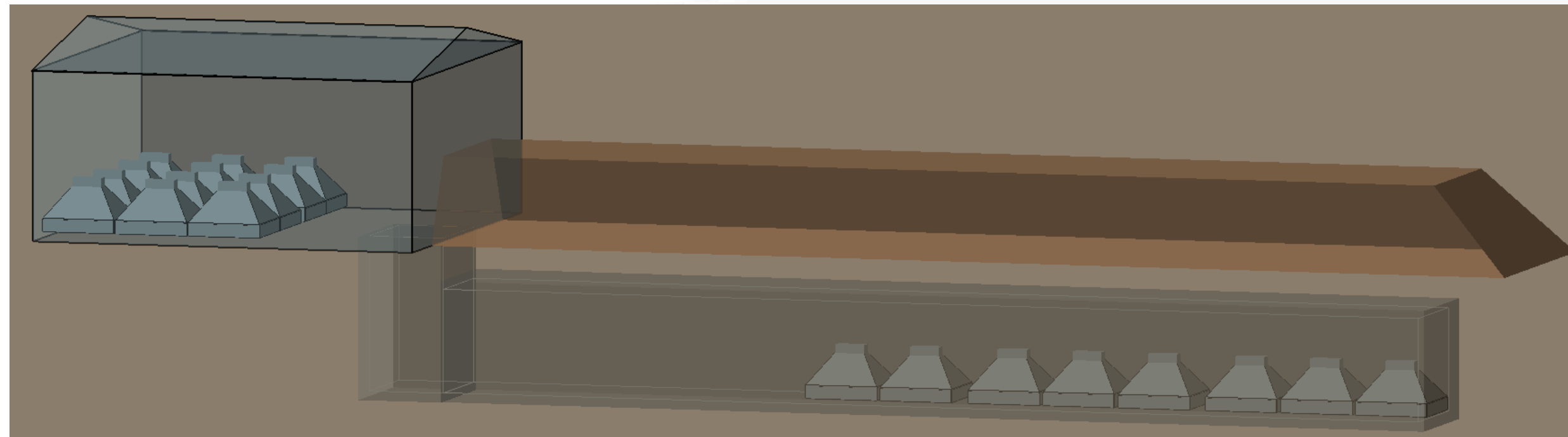
Угол наклона -  $25^\circ$

Угол обзора -  $60^\circ$

Площадь регистрации -  $0.5 \text{ m}^2$

Модель разработана в ИГУ:

M Ternovoy, I Kotovschikov, N. M. Budnev et.al, Simulation of the Tunka-Grande, TAIGA-Muon and TAIGA-HiSCORE arrays for a search of astrophysical gamma quanta with energy above 100 TeV, Journal of Physics Conference Series 1847(1), 2021



GEANT4 модель станции Tunka-Grande

## Основные параметры моделирования

- 
- CORSIKA-77410, с QGSJET-II-04 и GHEISHA I

порог по энергии  $e^\pm$  и  $\gamma$  - 0.5 MeV

порог по кин. энергии адронов - 50 MeV

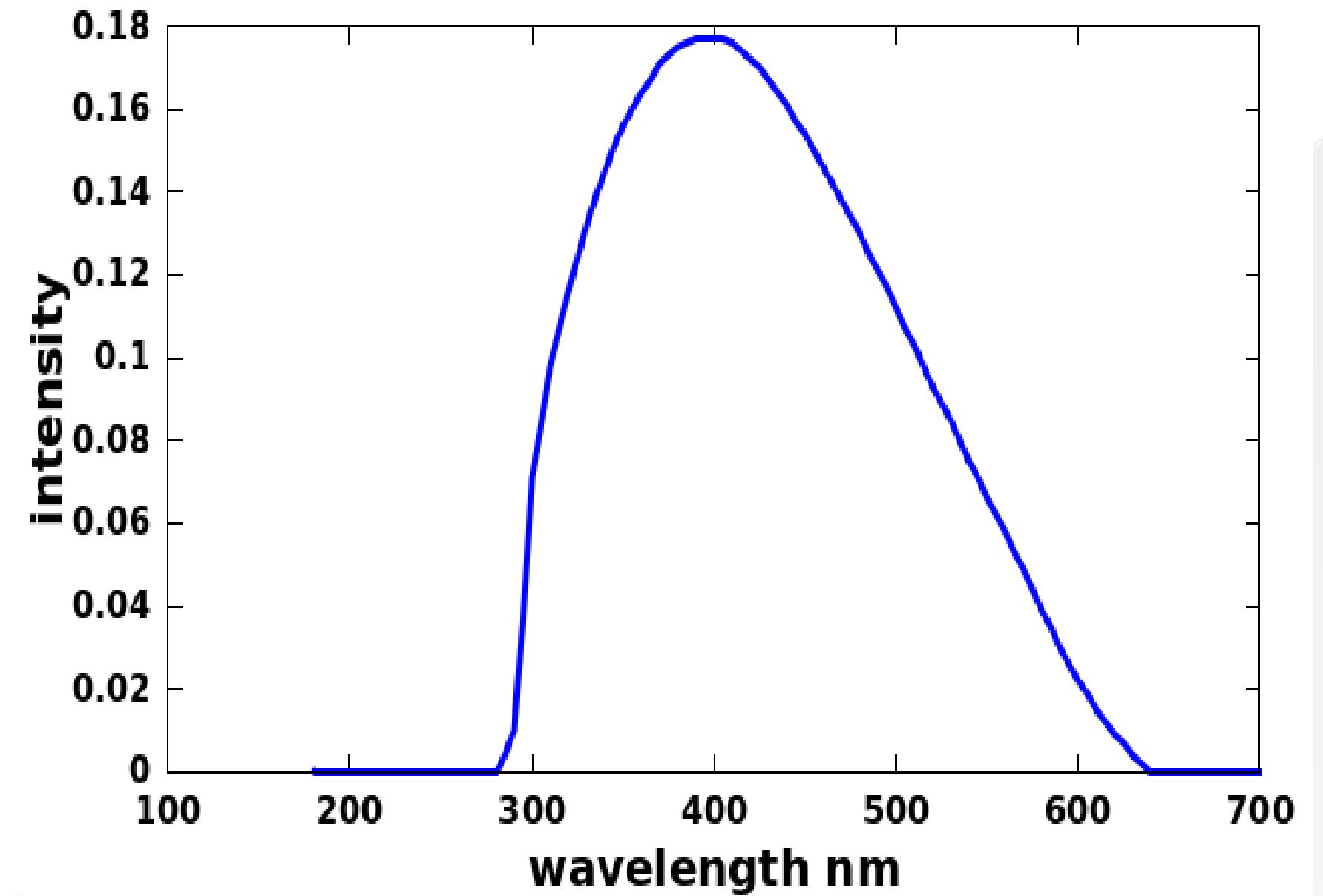
порог по кин. энергии мюонов - 10 MeV

диапазон длин волн для Черенковских фотонов:  
280 - 640 nm

Квантовая эффективность ФЭУ станций HiSCORE

ШАЛ разыгрывались в круге с радиусом: 200 м

Число черенковских фотонов в «банче»: 30



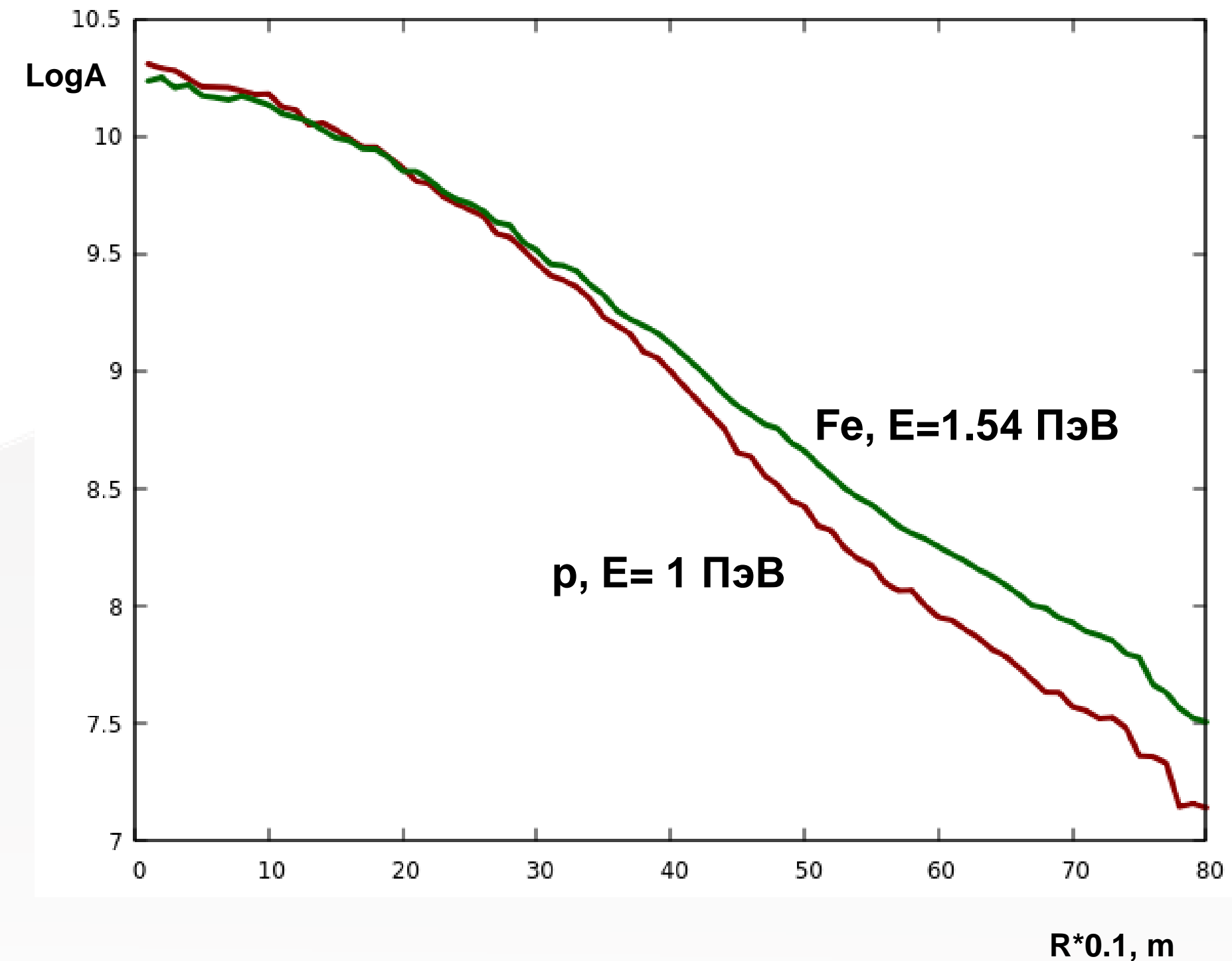
# Подготовка данных

В установке HiSCORE энергия определяется по амплитуде сигнала на расстоянии 200 м от оси ливня.

Например, если взять ШАЛ от p и Fe одинаковой энергии, то амплитуда сигнала на 200 метров будет разной. Это приводит к систематике в определении энергии.

В дальнейшем мы моделировали ШАЛ от различных элементов подбирая энергию так, чтобы амплитуда сигналов на 200 метрах была одинаковой. Рассматривались только вертикальные события на начальном этапе (0-15°)

Proton	helium	carbon	nitrogen	oxygen	silicon	calcium	iron
1000	1130	1270	1285	1310	1390	1480	1540

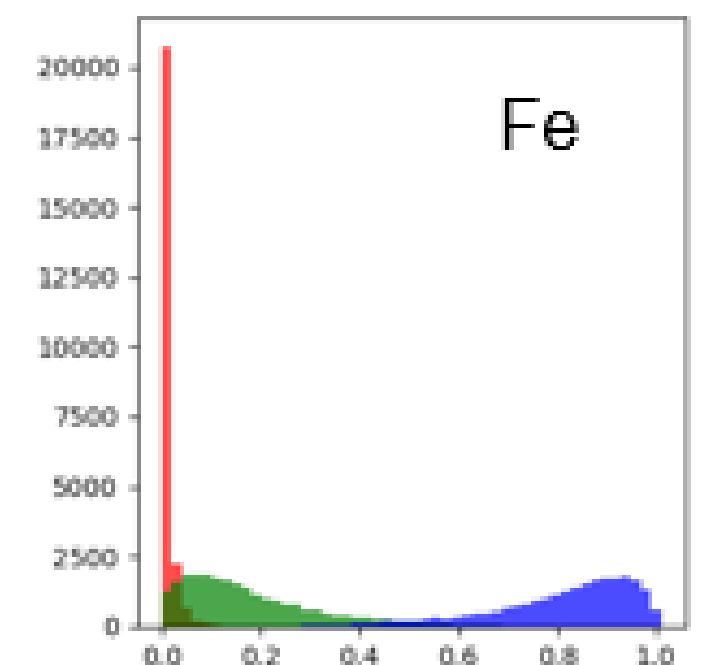
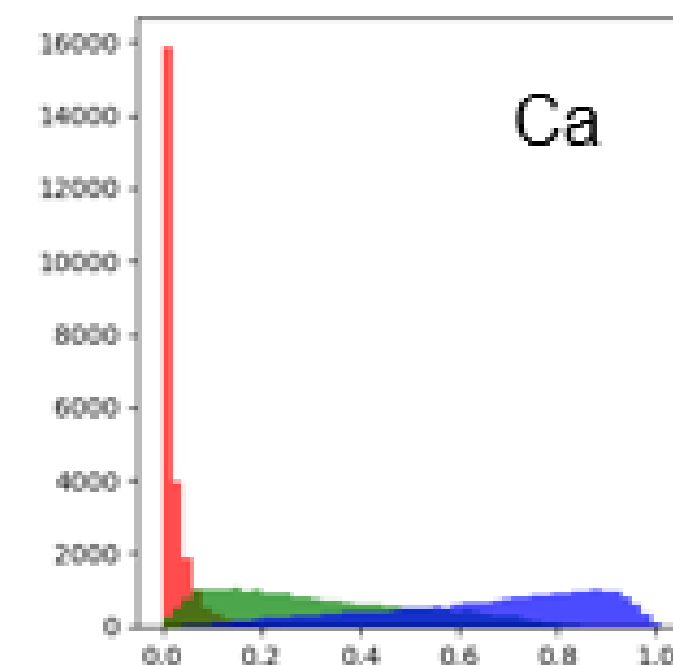
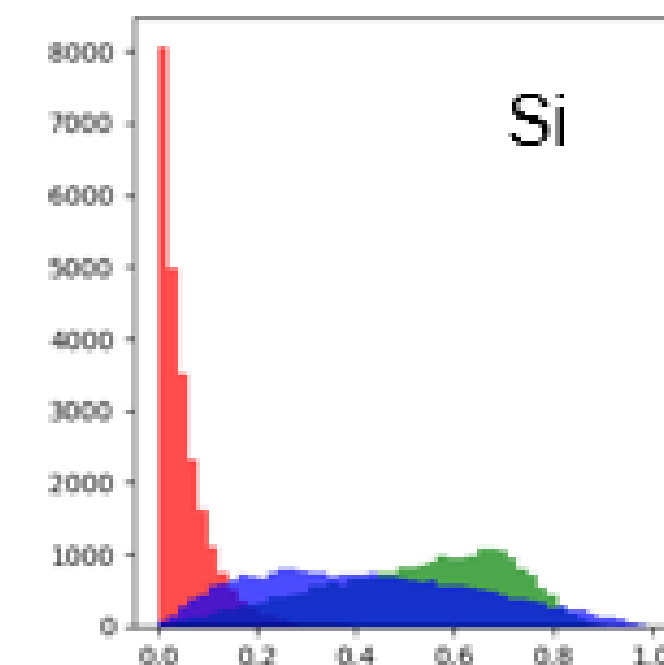
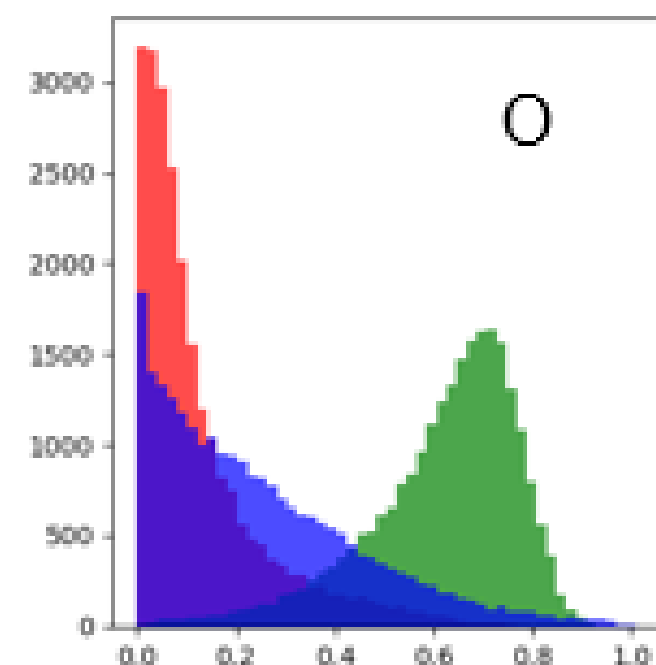
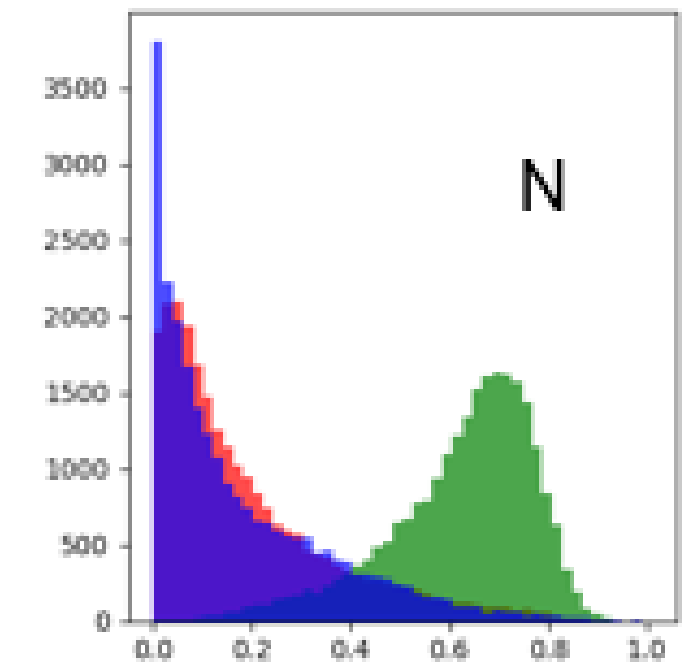
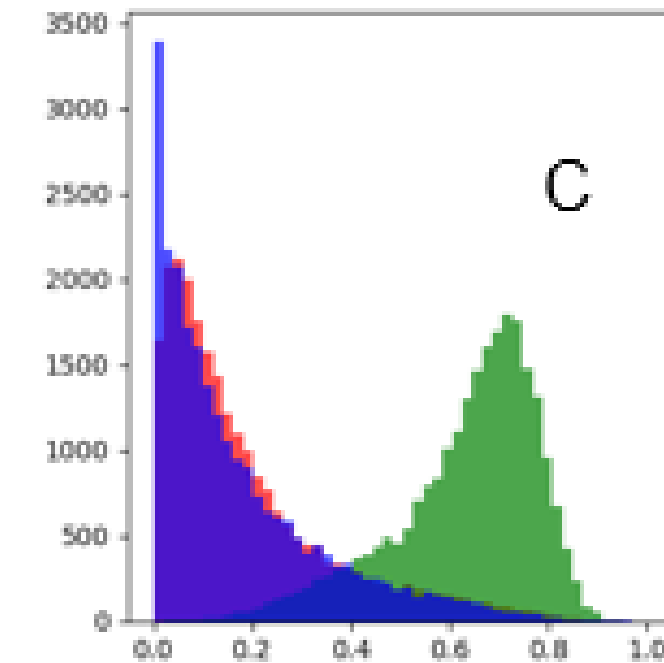
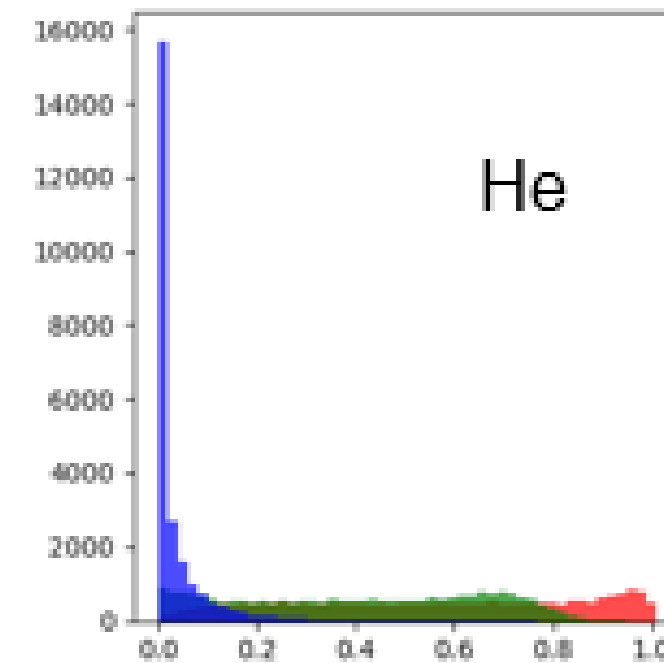
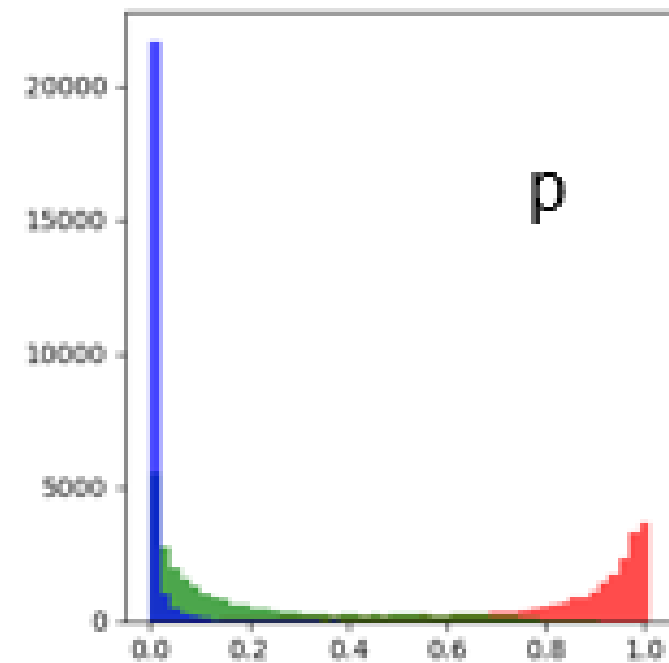


## 3-х классовая классификация на основе машинного обучения

- В работе использовалась Scikit-learn библиотека ([scikit-learn.org](http://scikit-learn.org)).
- В качестве входных параметров использовались 38 амплитуд с поверхностных и подземных счетчиков станций Tunka-Grande + 62 амплитуды со станций HiSCORE. Полное число входных параметров 100
- Мы использовали классификатор на основе алгоритма «Random forest»
- Число особенностей (features) было ограничено 10 (`max_features = 10`) согласно правилу, что их число не должно превышать корень из числа входных параметров
- Число деревьев фиксировано 100 (`n_estimators`).
- Глубина деревьев не фиксирована (`max_depth = None`)
- Классификатор был натренирован на 3 типах событий: p, N, Fe
- Статистика – 96000 событий каждого типа для тренировки, 24000 – для проверки
- Проверка точности проводилась для 8 типов событий: p, N, Fe + He, C, O, Si, Ca

# Как это работает.

- Классификатор выдает 3 значения «вероятностей»: **p-like**, **N-like**, **Fe-like**
- Сравнивая вероятности можно вынести решения по типу частицы.





# Предварительные результаты. E – 1 ПэВ

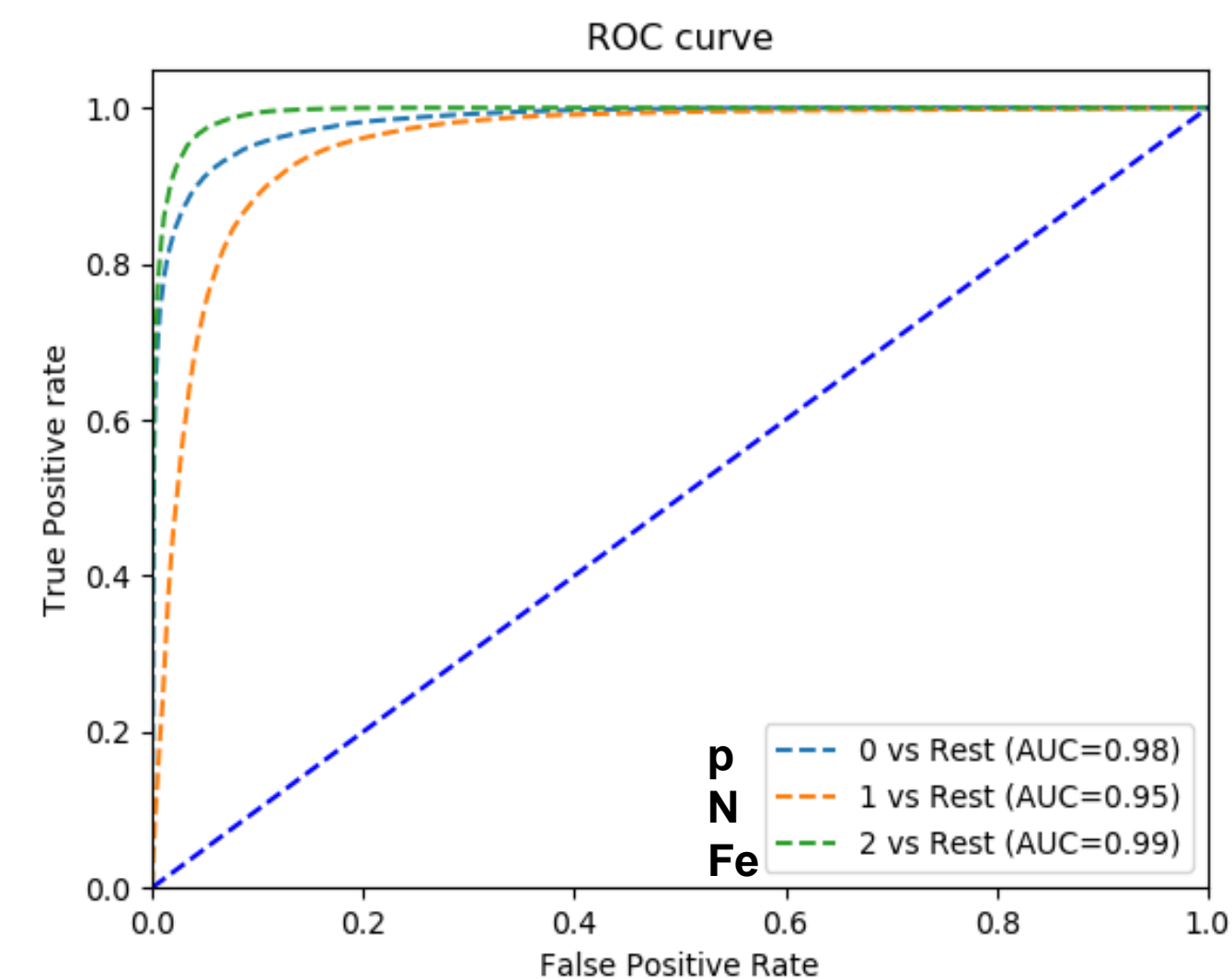
Классификация по максимальному значению вероятности

Тип тестового набора	p-like ID, %	N-like ID, %	Fe-like ID, %
<b>p</b>	<b>87</b>	<b>13</b>	<b>0</b>
He	56	43	1
C	8	85	7
<b>N</b>	<b>8</b>	<b>84</b>	<b>8</b>
O	3	84	13
Si	0	61	39
Ca	0	23	77
<b>Fe</b>	<b>0</b>	<b>4</b>	<b>96</b>

Классификация с ослабленным требованием на p-like ID (вероятность > 0,2)

Тип тестового набора	p-like ID, %	N-like ID, %	Fe-like ID, %
<b>p</b>	<b>97</b>	<b>3</b>	<b>0</b>
He	86	13	1
C	36	57	7
<b>N</b>	<b>37</b>	<b>55</b>	<b>8</b>
O	21	65	14
Si	3	59	38
Ca	0	23	77
<b>Fe</b>	<b>0</b>	<b>4</b>	<b>96</b>

Кривая ошибок



## Заключение

Предложен метод определения состава ПКЛ по совместным данным установок HiSCORE и Tunka-GRANDE.

Достоверность идентификации типа ядер составляет ~85% для CNO и 90% для протонов и 95% для Fe. Возможна модернизация метода с повышением достоверности идентификации протонов до 97% за счет снижения достоверности идентификации для CNO

Проверка параметров идентификации при энергиях до 10 ПэВ будет проведена в ближайшее время.

При дальнейшем развитии установки TAIGA-Muon достоверность идентификации будет улучшена.

**Спасибо**