

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ РОССИЙСКОЙ АКАДЕМИИ НАУК

На правах рукописи

Куринов Кирилл Олегович

**Изучение энергетического спектра космических лучей в области
энергий 10-100 ПэВ с использованием нейтронной компоненты
ШАЛ**

1.3.15 — Физика атомного ядра и элементарных частиц

Диссертация на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
к.ф.-м.н. О. Б. Щеголев

Москва — 2025

Оглавление

	Стр.
Введение	4
Глава 1. Описание установки	13
1.1. ENDA-LHAASO	13
1.2. ENDA-INR	14
1.3. Эн-детектор и методика регистрации	16
Глава 2. Выделение сигналов от тепловых нейтронов	22
2.1. Базисный метод	22
2.2. Создание обучающей выборки	23
2.3. Архитектура нейронной сети	26
2.4. Результаты сравнения	28
Глава 3. Моделирование эксперимента	30
3.1. Моделирование ШАЛ	30
3.2. Описание программы моделирования в Geant4	31
3.2.1. Моделирование отклика эн-детектора на различные частицы	32
3.2.2. Моделирование числа зарегистрированных нейтронов	35
3.2.3. Моделирование отклика установки на прохождение ШАЛ	37
3.3. Программа быстрого моделирования отклика установки	38
3.3.1. Сравнение результатов с Geant4	39
3.3.2. Моделирование выборок по отклику установки	41
Глава 4. Реконструкция событий	43
4.1. Эффективность регистрации ШАЛ	43
4.2. Базисный метод восстановления параметров ШАЛ	45
4.2.1. Восстановление направления прихода ливня	45
4.2.2. Восстановление положения оси и мощности ливня	47

4.2.3. Оценка точности восстановленных параметров	56
4.3. Реконструкция с помощью методов машинного обучения	58
4.3.1. Восстановление направления прихода ливня	59
4.3.2. Восстановление положения оси ливня	62
4.3.3. Реконструкция мощности и полного числа адронов ливня	64
4.4. Восстановление энергии первичной частицы	69
4.5. Оценка неопределенности реконструкции	80
Глава 5. Обработка экспериментальных данных	83
5.1. Анализ работы установки	83
5.2. Амплитудная калибровка эн-детекторов	84
5.3. Временное разрешение и временная калибровка установки	86
5.4. Программа визуализации событий	88
5.5. Результаты измерительной серии	90
5.5.1. Отбор событий	90
5.5.2. Реконструированные параметры ШАЛ	91
5.5.3. Временное распределение тепловых нейтронов в ШАЛ	92
5.5.4. Зависимость среднего числа зарегистрированных нейтронов от суммарного энерговыделения и мощности ливня	95
5.5.5. ФПР тепловых нейтронов в ШАЛ	98
5.5.6. Распределение ШАЛ по числу нейтронов	100
Заключение	106
Список основных обозначений	108
Список иллюстраций	109
Список таблиц	116
Список литературы	117

Введение

Актуальность темы. С момента открытия в 1912 году Виктором Гессом космических лучей, прошло уже больше ста лет, но фундаментальные вопросы, относительно источников, механизмов ускорения и распространения космических лучей остаются пока до конца не ясны. Исследование энергетического спектра, а также массового состава космических лучей, даст ключ к ответу на вышеописанные вопросы. К настоящему времени было сделано большое количество сообщений об энергетическом спектре, а также о массовом составе космических лучей, однако между ними существуют серьезные разногласия, особенно относительно последнего.

Энергетический спектр и массовый состав космических лучей измеряется прямыми методами на высотных аэростатах и спутниках (JACEE [1], Dapre [2], Calet [3]) до энергий ~ 50 ТэВ (были опубликованы данные с энергией выше 100 ТэВ, но они имеют большие статистические ошибки [4]). При больших энергиях используются не прямые измерения с применением установок по регистрации широких атмосферных ливней (ШАЛ). Полученный по данным различных установок энергетический спектр первичного космического излучения, подчиняется степенному закону с несколькими особенностями: излом в области нескольких ПэВ (“колени”), второй излом в области $\sim 10^{17}$ эВ (“второе колено”), обратный изгиб в области $10^{18} - 10^{19}$ эВ (“лодыжка”) и обрезание при энергиях выше $5 \cdot 10^{19}$ эВ (“ГЗК-эффект”).

Причина появления колена в энергетическом спектре до сих пор неизвестна. Существуют различные модели, которые относятся примерно к двум различным группам.

- Первая группа – космологическая и утверждает, что колено может являться результатом соответствующего механизма генерации, ускорения или свойства распространения космических лучей.
- Вторая группа связана с моделями, использующими новые процессы фи-

зики элементарных частиц или еще не подтвержденные предположения о свойствах задействованных частиц, а также с феноменологией ШАЛ.

Важной особенностью космологической модели является то, что многие механизмы ускорения и распространения космических лучей предсказывают зависимость наблюдаемых энергетических спектров от заряда первичной частицы, что должно приводить к последовательным изломам в спектре для каждой отдельной компоненты, то есть с увеличением энергии излом сначала возникает в спектре протонов, потом в гелии, углероде и т.д. [5, 6, 7, 8].

Совершенно другой подход основан на использовании новой физики элементарных частиц в атмосфере. В этих моделях часть энергии первичного адронного взаимодействия рассеивается в нерегистрируемых каналах ШАЛ. Эти каналы представлены образованием пока не обнаруженных тяжелых частиц [9] или мюонов и нейтрино высоких и сверхвысоких энергий (10-100 ТэВ и выше) [10]. Поскольку такие частицы не могут быть зарегистрированы обычными детекторами, первичная энергия недооценивается, что приводит к наблюдаемому излому восстановленного энергетического спектра. Последние результаты экспериментов ЛНС на встречных pp пучках с энергией в системе центра масс $\sqrt{s} = 14$ ТэВ, что эквивалентно энергии космических лучей $\sim 10^{17}$ эВ не подтверждают эту гипотезу [11]. Но стоит учесть, что на данный момент эффекты при протон-ионных столкновениях, которые являются доминирующими при развитии ШАЛ, не могут быть точно предсказаны из данных по pp и $p - Pb$ пучкам. Поэтому в настоящее время (2025 год) на ЛНС проводится эксперимент на $p - O$ пучках, который поможет уменьшить неопределенности ядерных эффектов при развитии каскада, а также скорректировать модели адронных взаимодействий.

Другая теория, которая объясняет также и появление второго колена, предложена [12] и утверждает, что изменение показателя наклона спектра космических лучей на энергии $\sim 4 \cdot 10^{15}$ эВ может быть вызвано началом реакции обратного бета-распада на реликтовых нейтрино $p + \bar{\nu} \rightarrow n + e^+$, что стано-

вится энергетически возможным при такой энергии, при условии, что масса нейтрино равна 0.4 эВ. Кроме того, если гипотеза верна, то колено должно быть исключительной особенностью протонной компоненты космических лучей. Второе колено соответствует пороговой энергии для той же реакции для ядер гелия и ожидается на уровне $\sim 3 \cdot 10^{17}$ эВ.

Феноменологический подход к проблеме излома состоит в том, что при энергиях ниже ~ 100 ТэВ/нуклон (т.е. ~ 100 ТэВ для протона и ~ 5 ПэВ для ядра железа) адроны не доходят до уровня наблюдения и их появление с ростом первичной энергии может являться причиной наблюдаемого в электромагнитной компоненте излома по числу частиц [13].

Характерное “колени” наблюдается во всех компонентах ливня, однако ожидаемая величина этого излома не согласуется с результатами экспериментальных измерений. Так при одной и той же величине излома первичного спектра, излом в мюонной и адронной компоненте ливня должен быть больше, чем в электромагнитной, однако наблюдается обратная картина. Без детального измерения спектров отдельных массовых компонент фундаментальный механизм “колени” остается неизвестным, для решения этой проблемы необходимо проводить мультикомпонентные исследования ШАЛ.

В подавляющем большинстве для исследования ШАЛ использовались методы, основанные на регистрации электромагнитной и мюонной компонент с использованием сцинтилляционных, черенковских и флуоресцентных детекторов. В очень редких случаях использовались калориметры для регистрации адронной компоненты. Их применение не получило широкого распространения и чаще всего они имеют весьма ограниченную площадь несопоставимую с размерами изучаемых ливней из-за экономических и технических причин.

Хотя адронная компонента ливня составляет всего около 1% от общего числа частиц в ШАЛ, она оказывает самое большое влияние на развитие каскада, поскольку является источником всех других компонент ливня. Адронная компонента чувствительна к первичному составу космических лучей и инфор-

мация о ней позволяет восстанавливать энергию первичной частицы.

В 2003 году Ю. В. Стенькин из ИЯИ РАН предложил регистрировать адроны вблизи оси ливня путем регистрации задержанных тепловых нейтронов, которые рождаются в результате ядерной реакции в веществе окружающей среды (например, в почве, зданиях, материалах детектора и воздухе) [14]. Из результатов моделирования было получено, что число тепловых нейтронов прямо пропорционально количеству адронов ШАЛ, попавших на площадь установки [15]. Из чего следует, что метод позволяет регистрировать адронную компоненту ливня по всей площади установки. Созданный для этой задачи детектор позволяет одновременно регистрировать также и электромагнитную компоненту ливня, из-за чего и получил свое название электронно-нейтронный детектор (эн-детектор).

В настоящее время эн-детекторы применяются для регистрации ШАЛ в установке ENDA (и её прототипе ENDA-INR) в эксперименте LHAASO [16], в установках PRISMA-36 и УРАН, входящих в экспериментальный научный комплекс НЕБОД [17]. Вышеперечисленные установки позволяют восстановить энергетический спектр первичных космических лучей, используя информацию как по нейтронной, так и по электромагнитной компоненте, что поможет для более точного исследования области “колена”.

Помимо установок для исследования ШАЛ, эн-детекторы применяются для анализа вариаций нейтронного фона в геофизических исследованиях [18, 19].

Целью данной работы являлась разработка методов анализа и реконструкции энергии первичных космических лучей в области энергий от 10-100 ПэВ по данным установки ENDA-INR.

Для достижения поставленной цели необходимо было решить следующие задачи:

1. Создать программу быстрого моделирования отклика установки ENDA-

INR.

2. Создать алгоритм выделения сигналов от тепловых нейтронов в эн-детекторах.
3. Создать метод для определения энергии первичной частицы с использованием информации о нейтронной компоненте ливня.
4. Обработать и проанализировать экспериментальные данные установки ENDA-INR, сравнить с результатами моделирования и данными других установок.

Методы исследования. Для достижения поставленных целей используются методы машинного обучения (нейронные сети, градиентный бустинг над решающими деревьями, случайный лес), методы Монте-Карло и методы статистической обработки данных.

Основные положения, выносимые на защиту.

1. Разработан метод выделения сигналов от тепловых нейтронов с использованием сверточных нейронных сетей в эн-детекторах.
2. Разработан метод реконструкции энергии первичной частицы с использованием данных о нейтронной компоненте ШАЛ и с оценкой неопределенности такого восстановления с помощью методов машинного обучения.
3. Функция пространственного распределения тепловых нейтронов в ливне и интегральный спектр по числу нейтронов в ШАЛ, полученный в ходе моделирования эксперимента ENDA-INR, проведенного с использованием программ CORSIKA и разработанной в ходе выполнения работы программы быстрого Монте-Карло моделирования.
4. Точность восстановления параметров ШАЛ (направление прихода и положение оси ливня, полное число адронов и заряженных частиц), полученные при использовании разработанной методики обработки событий.

5. Полученная по данным эксперимента функция пространственного распределения тепловых нейтронов в ливне для установки ENDA-INR.
6. Полученный по данным эксперимента интегральный спектр по числу нейтронов в ШАЛ для установки ENDA-INR

Научная новизна. Впервые был разработан и применен метод восстановления энергии космических лучей с использованием данных о нейтронной компоненте ШАЛ с помощью методов машинного обучения. Впервые был разработан и применен метод выделения сигналов от тепловых нейтронов в эндетекторах с использованием сверточных нейронных сетей. Впервые был проведен анализ экспериментальных данных установки ENDA-INR и получены следующие результаты: временное и пространственное распределение тепловых нейтронов в ШАЛ, зависимости мощности ливня и полного энергосодержания в событии от зарегистрированного числа нейтронов и распределение ШАЛ по числу нейтронов.

Теоретическая и практическая значимость. Независимое измерение энергетического спектра космических лучей в области энергий 10 - 100 ПэВ необходимо для проверки теоретических моделей, объясняющих их происхождение и механизмы ускорения. Использование методов машинного обучения при обработке данных позволяет улучшить точность результатов, за счет более эффективного использования информации.

Степень достоверности и апробация работы. Достоверность полученных результатов обеспечивается с помощью различных проверок, проведенных в данной работе, сравнением с результатами других экспериментов, а также публикацией результатов данной работы в рецензируемых научных журналах и апробацией на научных конференциях.

Основные результаты диссертации представлены на следующих конференциях:

1. 64-ая Всероссийская научная конференция МФТИ (2021)

2. 37 и 38 Всероссийская конференция по космическим лучам (2022, 2024)
3. 4-ый и 5-ый Международный симпозиум по космическим лучам и астрофизике (ISCRA – 2023, 2025)
4. 2-ая конференция LHAASO (2024)

Результаты также демонстрировались и обсуждались на семинарах НИЯУ МИФИ и рабочих встречах группы ENDA.

Работа поддержана грантом Российского научного фонда №23-72-01086, "Моделирование эксперимента ENDA-LHAASO". В составе коллектива авторов получено свидетельство о регистрации двух программ ЭВМ: "Программа моделирования установки ENDA", № 2024667840 от 30.07.2024 и "Программа быстрого моделирования установки ENDA-INR", №2025667745 от 09.07.2025.

Публикации по теме диссертации. Основные результаты по теме диссертации изложены в 5 печатных изданиях в журналах, рекомендованных ВАК.

1. Куринов К.О., Кулешов Д.А., Лагуткина А.А., Стенькин Ю.В., Щеголев О.Б. Выделение сигналов от тепловых нейтронов в электронно-нейтронных детекторах с использованием сверточных нейронных сетей в эксперименте ENDA // Журнал Экспериментальной и Теоретической Физики — 2023. — Т. 163. — С. 524–530.
2. Kurinov K.O., Kuleshov D.A., Stenkin Yu.V., Shchegolev O.B. Techniques for data analysis and primary mass reconstruction in the ENDA experiment // Physics of Atomic Nuclei — 2023. — Т. 86. — С. 1063–1068.
3. Куринов К.О., Кулешов Д.А., Малый И.О., Стенькин Ю.В., Щеголев О.Б. Первые экспериментальные результаты, полученные на установке ENDA-INR // Журнал Экспериментальной и Теоретической Физики — 2025 — Т. 168. — С. 343–349.
4. Куринов К.О., Кулешов Д.А., Малый И.О., Стенькин Ю.В., Щеголев О.Б. Быстрое Монте-Карло моделирование установки ENDA // Краткие

сообщения по физике ФИАН — 2025 — Т. 9. — С. 15–26.

5. *Куринов К.О., Кулешов Д.А., Малий И.О., Стенькин Ю.В., Степанов В.И., Щеголев О.Б.* Восстановление направления прихода ШАЛ в эксперименте ENDA-INR // Ядерная физика и инжиниринг — 2026 — Т. 17.— С. 236–240.

Личный вклад. Личный вклад автора в получение научных результатов, лежащих в основе диссертации, является определяющим. Автором были выполнены этапы по предобработке данных, выборе архитектуры и обучению сверточной нейронной сети для задачи выделения сигналов от тепловых нейтронов в эн-детекторах. Также автором была создана программа быстрого Монте-Карло моделирования отклика установки ENDA-INR, которая была основана на зависимостях и распределениях, полученных из программы моделирования на Geant4. Автор разработал методику реконструкции параметров ШАЛ с использованием методов машинного обучения, которая включает в себя оценку неопределенности восстановленных параметров, а также провел сравнение с традиционными методами реконструкции для установки ENDA-INR. Автором был разработан метод реконструкции энергии первичной частицы с использованием информации о нейтронной компоненте с использованием методов машинного обучения и проведено сравнение качества реконструкции для трёх алгоритмов (линейная регрессия с полиномиальными признаками, случайный лес и градиентный бустинг над решающими деревьями), а также с результатами двух базисных подходов. Автором была проведена предобработка экспериментальных данных установки ENDA-INR, работа по определению временных задержек между детекторами и временного разрешения, получена функция пространственного распределения тепловых нейтронов в ШАЛ, а также интегральный спектр по числу нейтронов в ливне, проведено сравнение с результатами моделирования и с экспериментальными результатами установок PRISMA-32, PRISMA-YBJ и KASCADE.

Структура и объем работы. Диссертация состоит из оглавления, введения, пяти разделов, заключения, списка иллюстраций, списка таблиц, словаря терминов и списка литературы из 84 наименований. Основной текст занимает 126 страниц.

Глава 1

Описание установки

1.1. ENDA-LHAASO

Установка ENDA, создаваемая в рамках уникального эксперимента нового поколения LHAASO, позволит провести мультикомпонентное исследование ШАЛ с регистрацией всех компонент ливня. Эксперимент LHAASO предназначен для измерения энергетического спектра, массового состава и анизотропии космических лучей в диапазоне энергий от 10^{12} до 10^{18} эВ. Основной частью LHAASO является установка Km2A [20], созданная для регистрации ШАЛ и состоящая из 5216 электромагнитных сцинтилляционных детекторов площадью по 1 м^2 и 1188 водных черенковских мюонных детекторов с площадью по 36 м^2 (рис. 1.1). Установка имеет большую эффективную площадь около 1 км^2 и очень плотную решетку детекторов, что приводит к угловому разрешению в несколько десятых долей градуса. Также у установки очень выгодное расположение, на высоте 4410 м над уровнем моря, на уровне максимального развития каскада, в области первого “колена”, что сводит к минимуму флуктуации при развитии ливня, которые являются самой большой проблемой при изучении ШАЛ. Одним из главных результатов эксперимента является открытие 43 источников гамма-излучения сверхвысоких энергий, максимальная энергия фотонов в которых достигает 2.5 ПэВ, что является свидетельством существования “Пэ-Ватронов” в нашей галактике, которые ускоряют частицы космических лучей до энергий, превышающих область колена, до десятков ПэВ [21].

Измерение спектров отдельных частиц, особенно легкой компоненты в области колена, является ключевым моментом в установлении связи между небесными источниками и потоками, регистрируемыми на Земле. Именно в этом и состоит главная задача установки ENDA-64, которая на данный момент состоит из 4 кластеров по 16 эн-детекторов в каждом, общая площадь 10^3 м^2 . В

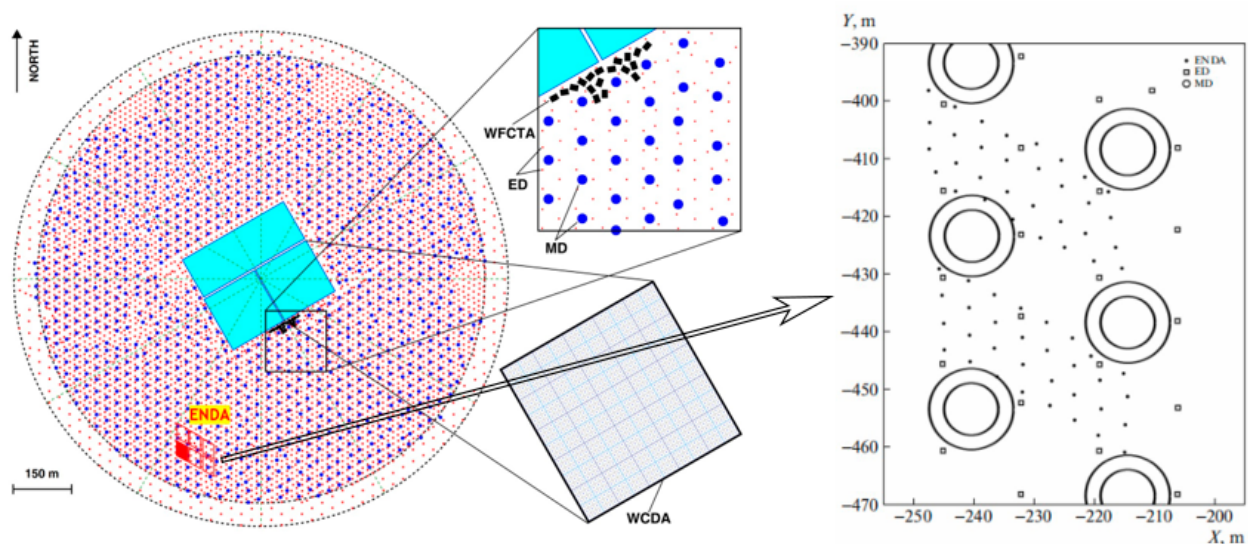


Рис. 1.1: Расположение детекторов эксперимента LHAASO

дальнейшем планируется расширение установки до 400 эн-детекторов с площадью 10^4 м^2 для изучения энергетического спектра тяжелой компоненты (Fe).

Прототипы установки ENDA были созданы на территории ИЯИ РАН в Москве (ENDA-INR [22]) и в Китае в деревне Янгбаджинг (PRISMA-YBJ [23]) на высоте 4300 м над уровнем моря, они состоят из 16 эн-детекторов и были созданы для изучения ШАЛ с энергией выше 1 ПэВ, тестирования методики регистрации и обработки сигналов.

1.2. ENDA-INR

Эн-детекторы установки ENDA-INR, располагаются по квадратной сетке 4×4 с шагом в 5 м на грунте (рис. 1.2). Из-за того, что прототип расположен под открытым небом, для защиты от дождя и снега эн-детекторы помещаются в дополнительный стальной внешний корпус.

Сигналы с эн-детекторов снимаются одновременно с 8-го и 5-го промежуточного динода, 4-х дюймового фотоэлектронного умножителя (ФЭУ) модели CR-165, производства Beijing Hamamatsu. Промежуточный динод используется для увеличения динамического диапазона и используется в тех случаях, когда

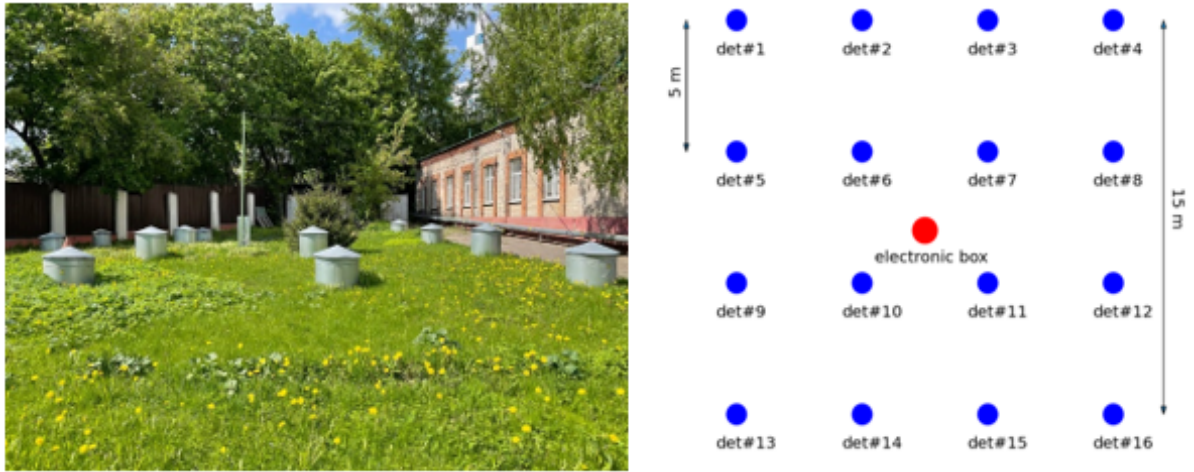


Рис. 1.2: Фото установки ENDA-INR (слева), конфигурация детекторов (справа)

сигнал по восьмому диноду находится в насыщении, его коэффициент усиления в ~ 30 раз меньше, чем у восьмого. Сигналы с восьмого динода ФЭУ поступают на дискриминатор-интегратор-усилитель (ДИУ), который расположен непосредственно в детекторе, где интегрируются в течение 1 мкс. Сигналы с промежуточного динода поступают на интегральный усилитель (ИУ), где также интегрируются в течение 1 мкс. В работах по изучению вариаций фонового потока тепловых нейтронов используется время интегрирования, равное 5-10 мкс. После усиления сигнал поступает по кабелю на вход АЦП. В качестве АЦП используется 32-х канальный DT5740D, производства CAEN (цена деления 0.5 мВ/канал), шаг дискретизации 32 нс. АЦП расположен в металлическом блоке в центре установки, вместе с высоковольтным источником питания и персональным компьютером, на который записываются данные с установки с помощью онлайн программы.

У прототипа есть два режима работы: режим совпадений (триггер формируется при срабатывании двух и более детекторов) и режим единичных срабатываний (срабатывание только одного детектора), который используется для калибровки детекторов. Триггер в детекторе формируется, когда амплитуда сигнала превышает значение в 3 мВ. Нормировка каналов АЦП в число частиц проводилась из сравнения показаний эн-детекторов установки ENDA и пласти-

ческих детекторов KM2A-ED эксперимента LHAASO в совместных событиях. Для установки ENDA-INR одному каналу АЦП соответствует ~ 1.5 частицы, а динамический диапазон по заряженной компоненте составляет $9 \div 3.2 \cdot 10^4$ частиц на детектор или $24 \div 8.5 \cdot 10^4$ частиц на м^2 .

1.3. Эн-детектор и методика регистрации

Конструкция эн-детектора (рис. 1.3 слева) является стандартной для сцинтилляционных детекторов ШАЛ, однако вместо толстого пластического сцинтиллятора используется специализированный тонкий (толщина 50 мг/см^2) сцинтиллятор на основе неорганического сцинтилляционного компаунда $\text{ZnS}(\text{Ag}) + \text{B}_2\text{O}_3$ с необогащенным бором. Сцинтиллятор изготовлен путём спекания кристаллов ZnS и B_2O_3 , которые затем помещаются в слой оптически прозрачного силикона. Образец сцинтиллятора представлен на рис. 1.3 (справа).

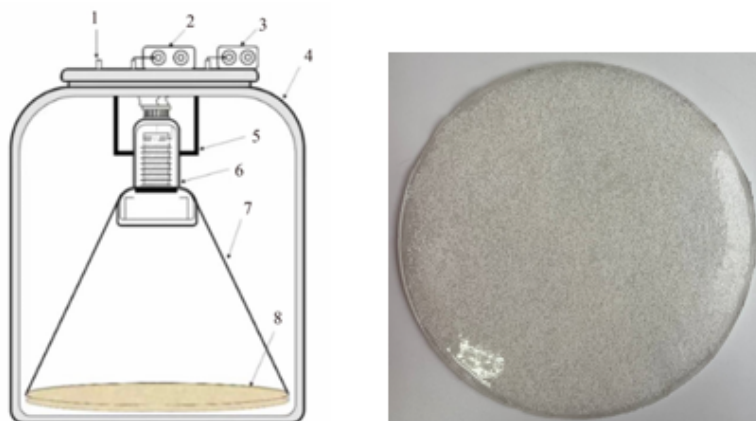
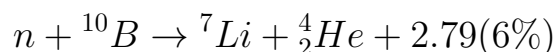
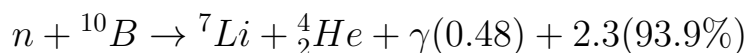


Рис. 1.3: Схема эн-детектора (1 – высоковольтный входной порт, 2 – ДИУ 8-го динода, 3 – выход 5-го динода, 4 – пластиковый корпус, 5 – держатель ФЭУ, 6 – ФЭУ, 7 – светособирающий конус, 8 – сцинтиллятор) (слева), образец сцинтиллятора (справа).

Из-за малой толщины сцинтиллятора сигнал от прохождения одиночных заряженных частиц лежит ниже порога регистрации, но в то же время сцин-

тиллятор обладает большим световыходом при регистрации тепловых нейтронов. Эффективность захвата теплового нейтрона в сцинтилляторе составляет $\sim 20\%$. Нейтроны регистрируются с помощью следующих реакций:



Легкие релятивистские частицы (β , γ , μ) возбуждают, в основном, быструю компоненту (~ 40 нс) сцинтиллятора $\text{ZnS}(\text{Ag})$, в то время как, тяжелые частицы (α , Li^7) возбуждают также и медленные компоненты сцинтиллятора (несколько микросекунд и более).

Для заряженной компоненты фронт интегрированного импульса круто нарастающий, но из-за того, что исходный импульс не является прямоугольным, а имеет экспоненциальный хвост, то время нарастания фронта равно не (~ 40 нс), а (~ 200 нс). Для нейтронной компоненты из-за увеличенного времени собирания заряда, увеличивается и фронт нарастания интегрированного импульса. На рисунке 1.4 представлены нормированные усредненные фронты импульса для заряженной и нейтронной компоненты.

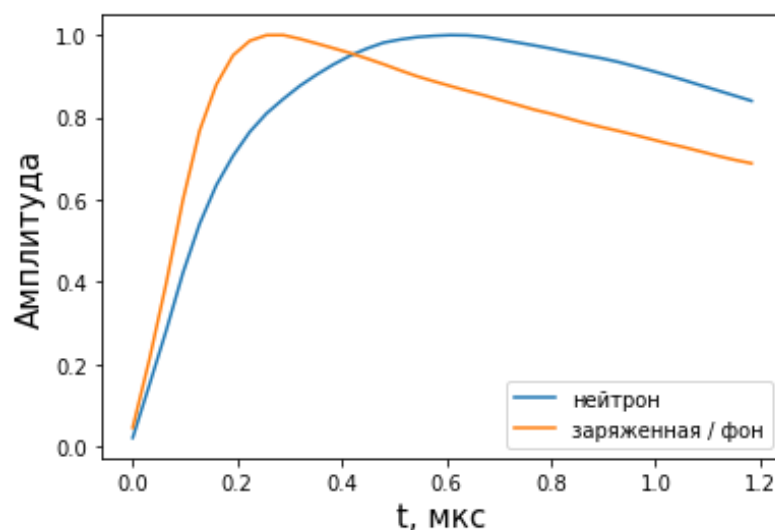


Рис. 1.4: Фронт импульса заряженной и нейтронной компоненты

Совпадателные импульсы от фронта ливня регистрируются детектора-

ми в течение первой мкс, нейтроны же около 1 мс живут в грунте и более 30 мс в воздухе, и им необходимо время для замедления в грунте (~ 500 мкс). Чтобы избежать просчетов, связанных с прохождением мощных ШАЛ, вызывающих насыщение электроники и связанным с этим мертвым временем, вводится задержка начала счета нейтронов, и они регистрируются во временном интервале от 100 мкс до 5 мс после фронта ливня. Порог по амплитуде для задержанных импульсов составляет 10 каналов АЦП или 5 мВ. При этом часть задержанных импульсов создается одновременным прохождением нескольких заряженных частиц через сцинтиллятор, а также шумами ФЭУ. Для выделения нейтронных импульсов используются специальные алгоритмы (см. Главу 2).

Осциллограмма события состоит из 164500 точек, где первый импульс соответствует энерговыделению от прохождения фронта ШАЛ, а затем следуют импульсы от задержанных нейтронов. На рисунке 1.5 показана осциллограмма, зарегистрированного события ШАЛ со всех сработавших детекторов (отмечены разным цветом) с установки ENDA-INR.

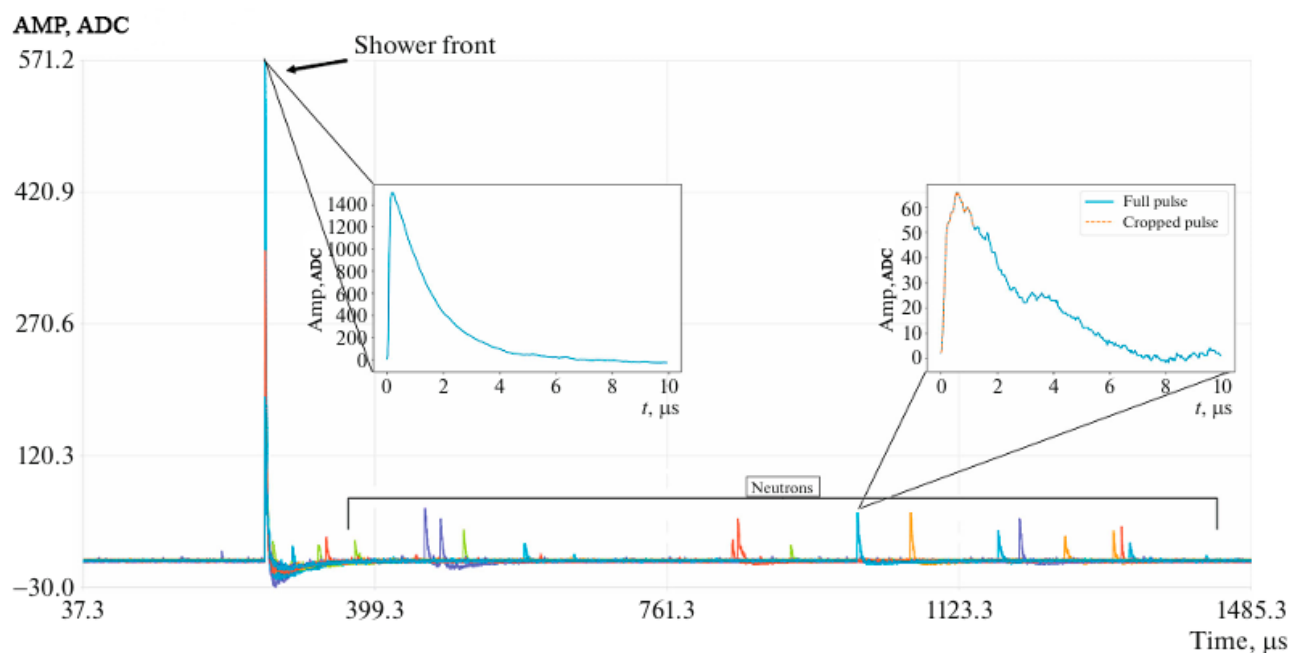


Рис. 1.5: Осциллограмма события ШАЛ

На рисунке 1.6 представлена визуализация события с самым большим числом зарегистрированных нейтронов за выбранный период обработки данных установки ENDA-INR (см. Главу 5).

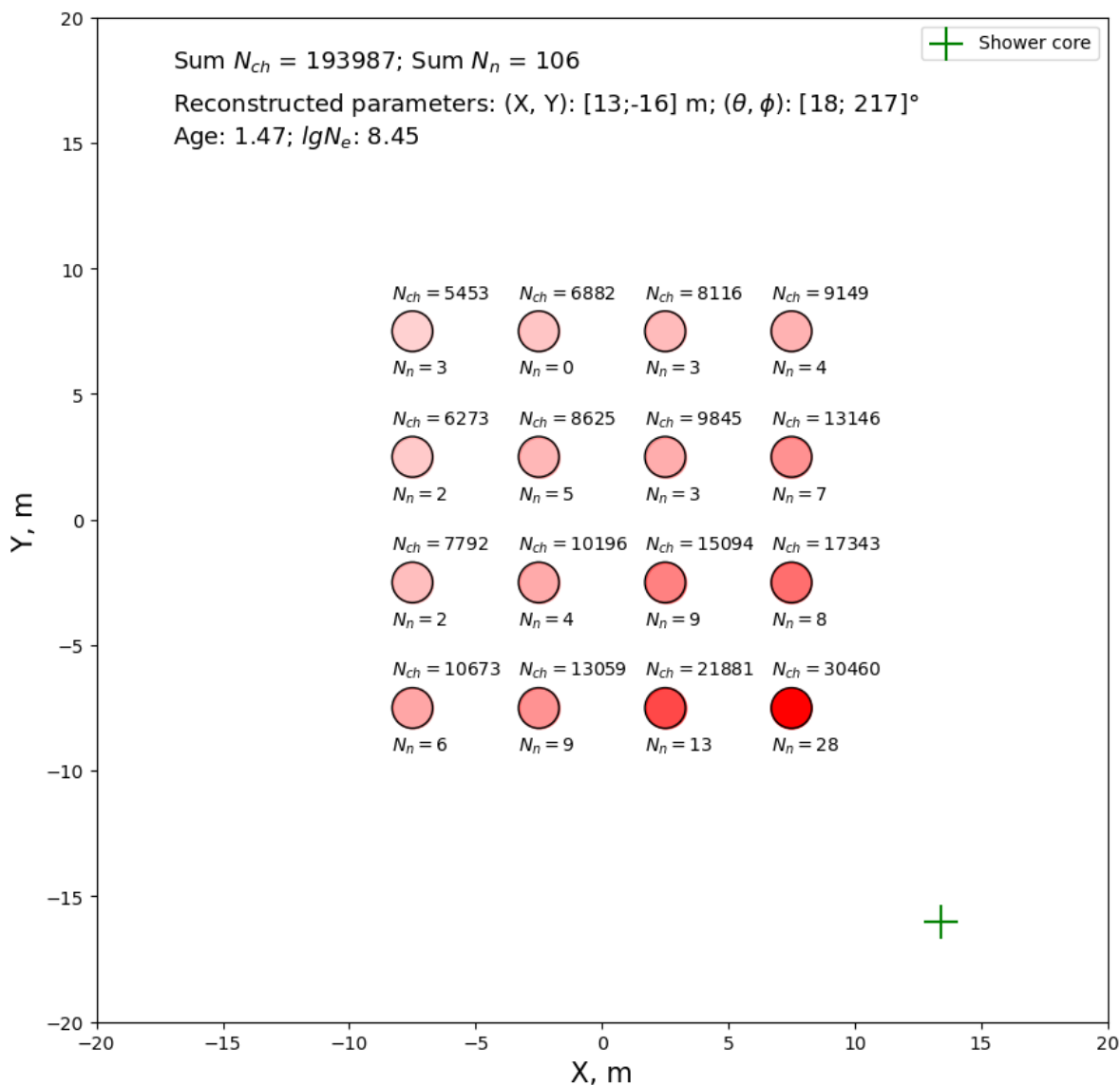


Рис. 1.6: Визуализация события ШАЛ с наибольшим числом зарегистрированных нейтронов

Над детекторами указано зарегистрированное число заряженных частиц в эн-детекторе, а под детектором указано зарегистрированное число нейтронов. Зеленым крестом обозначено реконструированное положение оси ливня (см. Главу 4), а цвет эн-детекторов обозначает количество зарегистрированных заряженных частиц. На рисунке 1.7 представлена визуализация того же собы-

тия, но с осциллограммой каждого эн-детектора отдельно. Штриховой линией отмечен порог по амплитуде для задержанного импульса.

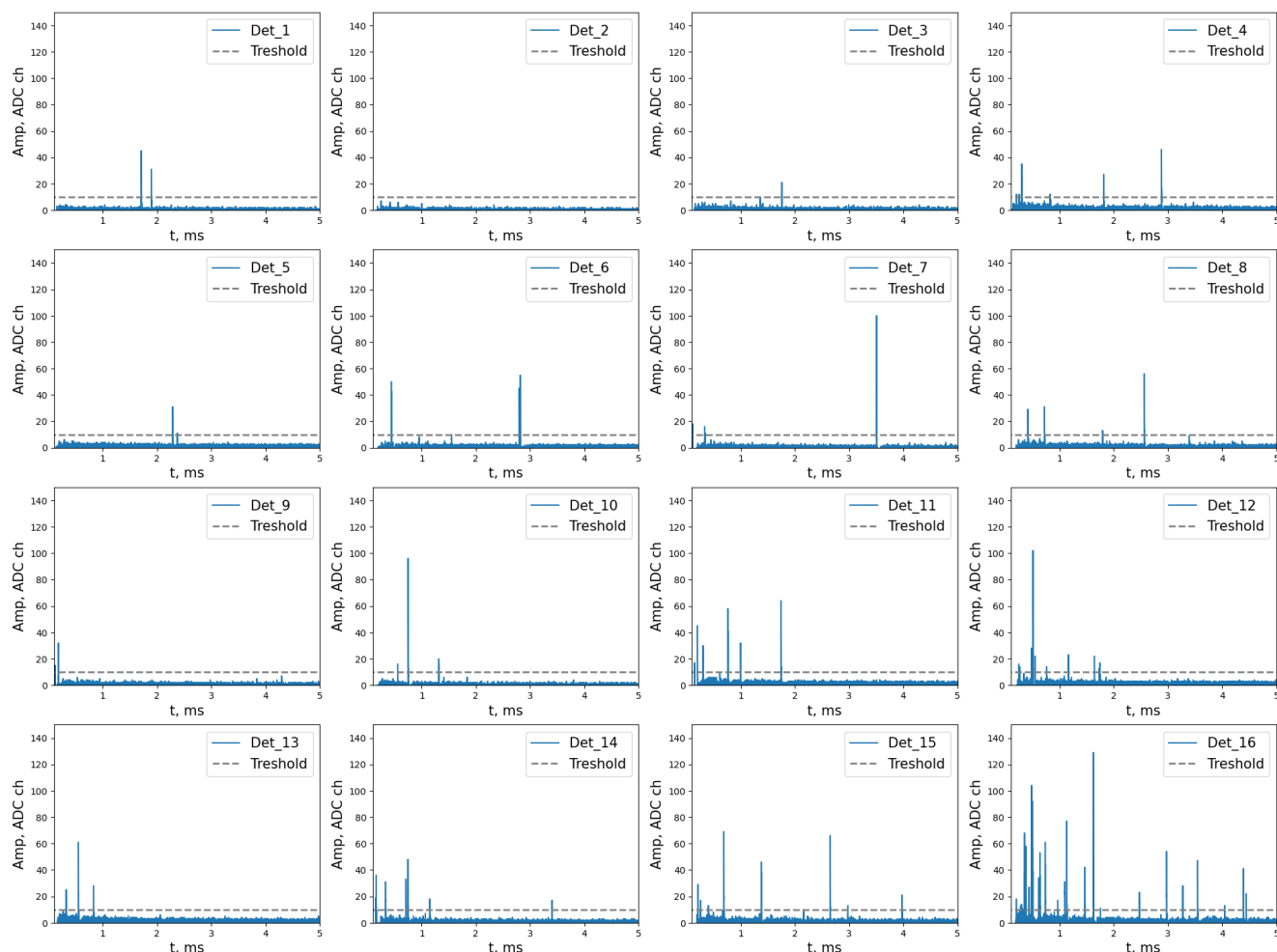


Рис. 1.7: Визуализация события ШАЛ с наибольшим число зарегистрированных нейтронов с осциллограммой каждого отдельного эн-детектора

Онлайн программа, анализирующая входной сигнал, присваивает каждому событию метку триггера второго уровня, которая называется мастером (М), он имеет битовое двоичное представление, где значение каждого бита определяется следующими условиями:

1. сработали два и более детектора (с порогом 6 каналов АЦП).
2. суммарное энерговыделение ≥ 300 каналов АЦП.
3. суммарное число нейтронов ≥ 5 .

Соответственно, если было выполнено условие (1), то в десятичной записи $M = 1$, если (1) и (2), то $M = 3$, если выполнены все условия, то $M = 7$. Для

контроля вариаций фона тепловых нейтронов (а также контроля работоспособности онлайн программы) используется программный триггер, названный нулевым мастером ($M = 0$), который запускается каждые 10 секунд. Суммарное по всем детекторам для установки ENDA-INR число фоновых нейтронов подчиняется распределению Пуассона, и его среднее составляет 0.47 нейтронов на событие, для заряженной/шумовой компоненты среднее значение составляет 0.63 импульса на событие.

Глава 2

Выделение сигналов от тепловых нейтронов

Выделение импульсов от нейтронов основано на различии форм импульсов, получаемых при захвате в сцинтилляторе тепловых нейтронов, легких релятивистских заряженных частиц и шумов. Существуют различные способы идентификации нейтронных импульсов, включающие использование аналитических методов [24, 25, 26], а также методов машинного обучения, таких как метод опорных векторов [27], использование полносвязных нейронных сетей и метод k-средних [28], а также применение сверточных нейронных сетей (Convolutional Neural Networks) [29]. В статье [29] приведено сравнение, в котором алгоритм разделения сигналов от заряженной и нейтронной компоненты для сцинтиллятора на основе ${}^6\text{LiF:ZnS(Ag)}$ с примесью поливинилтолуола показал лучшее качество по сравнению с другими аналитическими методами разделения сигналов.

В работе предлагается метод выделения нейтронных импульсов для эндетекторов с применением сверточных нейронных сетей и проводится сравнение с базисным методом, применяемым в настоящее время на установке.

2.1. Базисный метод

Метод интегрирования заряда [26], основан на оценке отношения заряда, собранного от "быстрой" компоненты сцинтиллятора к заряду, собранному за время достижения максимальной амплитуды интегрированного сигнала. Для нейтронных импульсов это отношение всегда меньше 1, для легких релятивистских заряженных частиц и фона отношение близко или равно 1. Для получения хорошего качества классификации сигналов по форме импульса данным методом необходимо время интегрирования порядка 10 мкс. Ради улучшения временного разрешения, необходимого для счета нейтронных импульсов, сопро-

вождающих ливень, в установке используется компромиссное время интегрирования, равное всего лишь 1 мкс. Однако в этом случае алгоритм разделения на основе интегрирования заряда дает низкое качество классификации.

На данный момент в установке применяется алгоритм [22], который в дальнейшем мы будем называть базисным, согласно которому на основе анализа распределения по фронту нарастания интегрированного импульса, выбирается пороговое значение, и все сигналы, чей фронт нарастания импульса лежит слева от этого значения, считаются сигналами от заряженной компоненты и фоном, а справа - сигналами от нейтронной компоненты.

2.2. Создание обучающей выборки

Для обучения сверточной нейронной сети была сформирована обучающая выборка на основе импульсов, полученных из набора осциллограмм, записанных в течение 11 месяцев работы установки ENDA-INR. Для формирования выборки по заряженной компоненте, были взяты осциллограммы с установки, работающей в режиме единичных срабатываний. В данном режиме импульсы являются одиночными и могут быть сигналами как от заряженной компоненты, так и от нейтронной. Чтобы оставить импульсы только от заряженной компоненты, импульсы, классифицированные базисным методом как нейтронные, были удалены из выборки. Выборка по нейтронной компоненте формировалась из запаздывающих импульсов больших событий ШАЛ. Под большими событиями подразумеваются события, в которых число зарегистрированных нейтронов лежит выше пуассоновского фона (верхнее распределение на рисунке 2.1). Также был установлен порог на суммарное энерговыделение в детекторах в 500 мВ. В таких событиях доля заряженной компоненты подавлена относительно нейтронной компоненты и можно все запаздывающие импульсы считать нейтронными.

Количество задержанных импульсов на ливень распределено по Пуас-

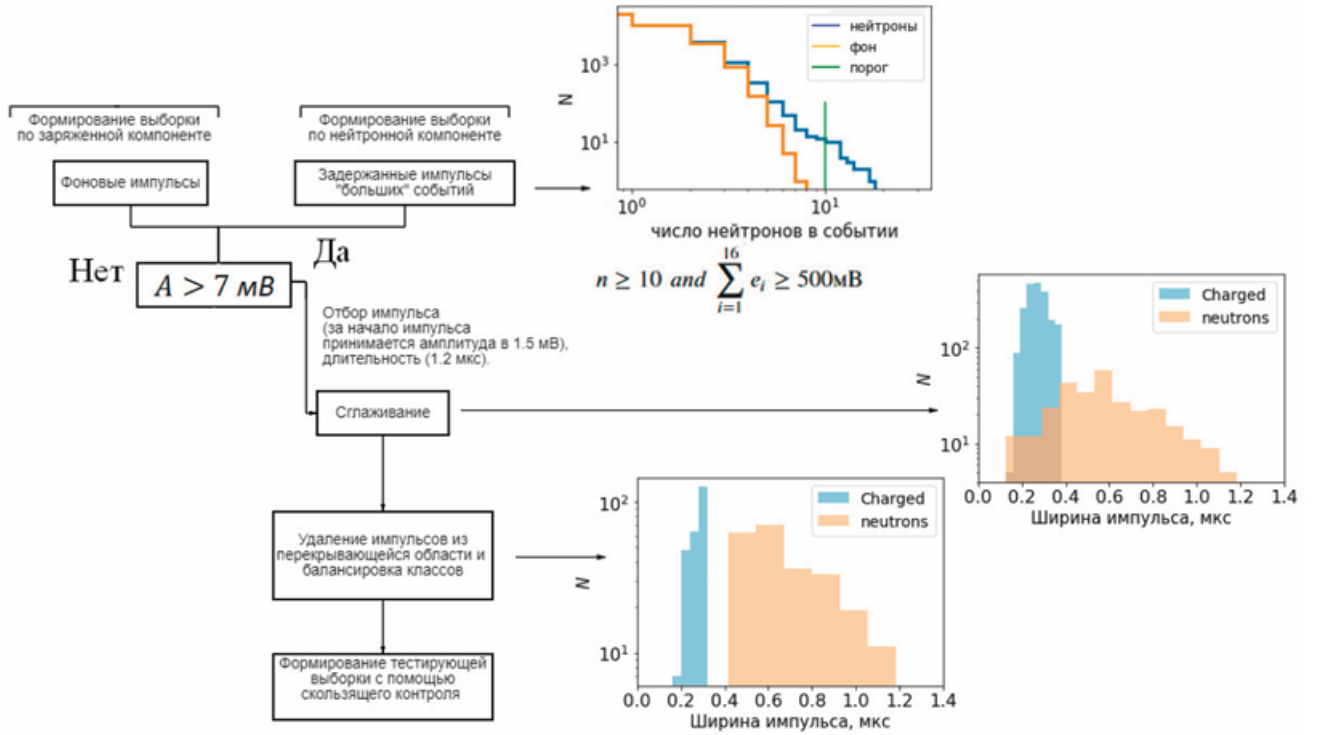


Рис. 2.1: Алгоритм формирования обучающей выборки.

сону с математическим ожиданием, равным 1.1. Среднее значение количества задержанных нейтронных импульсов на одно событие по полученной выборке равно 18.8, и из-за того, что вклад в задержанные импульсы вносят и фоновые нейтроны, и заряженная компонента, то можно утверждать, что заряженная компонента в больших событиях подавлена относительно нейтронной более чем в 17 раз.

Импульсы отбирались из условия превышения амплитуды импульса порога в 7 мВ. Все импульсы из выборки имели длительность в 1.2 мкс или в 38 точек (при шаге дискретизации в 32 нс). За начало импульса был принят порог в 0.5 мВ. После этого, все сигналы сглаживались с помощью скользящей регрессии [30]. В результате вышеописанных действий было отобрано 2210 импульсов заряженной и 348 импульсов нейтронной компоненты.

На рисунках 2.2, 2.3 представлены распределения по амплитуде и длительности фронта импульса для обеих компонент, где под длительностью фронта понимается интервал времени между началом импульса и его максимумом.

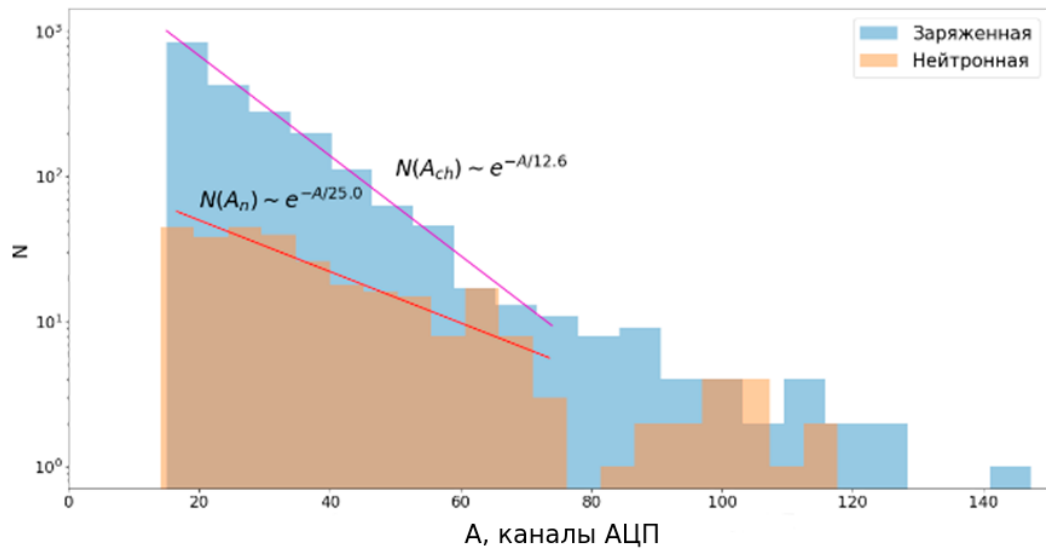


Рис. 2.2: Распределение по амплитуде сигнала

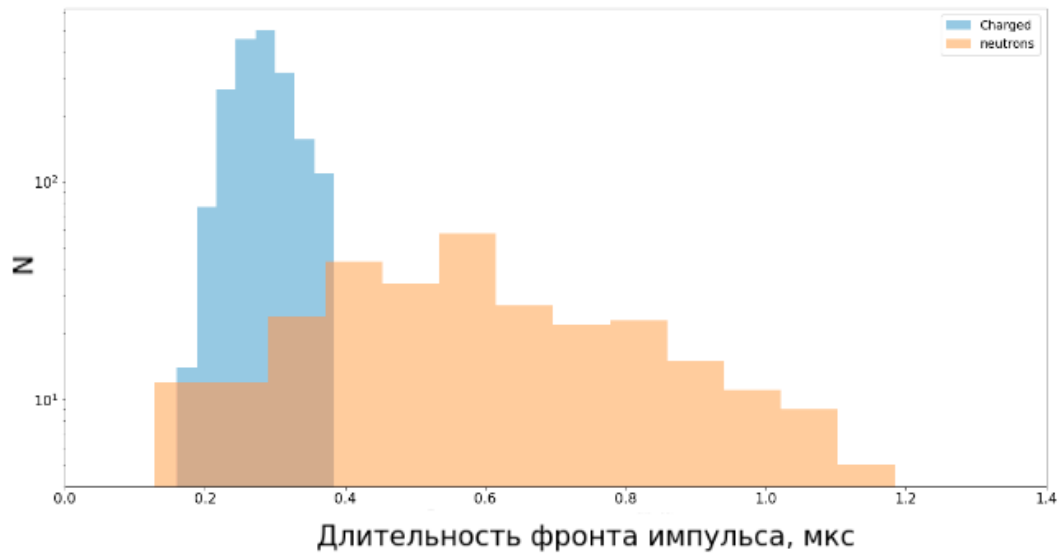


Рис. 2.3: Распределение по длительности фронта импульса

После построения распределения по времени нарастания фронта импульса, для использования аппроксимирующих свойств нейронной сети из выборки были удалены все импульсы, чьи времена нарастания фронта находились в пересекающейся области. Затем выборки для каждого класса были отбалансированы, чтобы число импульсов в обоих классах было равно. Конечное распределение по длительности фронта импульса для обоих классов представлено на самом нижнем распределении на рисунке 2.1. По полученной выборке были

построены фронты импульса обеих компонент. Максимальная амплитуда усредненного фронта импульса нейтронной компоненты оказалась в 1.4 раза больше, чем у заряженной компоненты.

Эффективность модели оценивалась с помощью k-fold кросс-валидации [31], где вся выборка разбивалась на 3 равные подвыборки и точность модели оценивалась по среднему значению на всех подвыборках.

2.3. Архитектура нейронной сети

Для создания нейросети использовался фреймворк PyTorch [32], архитектура сети представлена на рисунке 2.4.

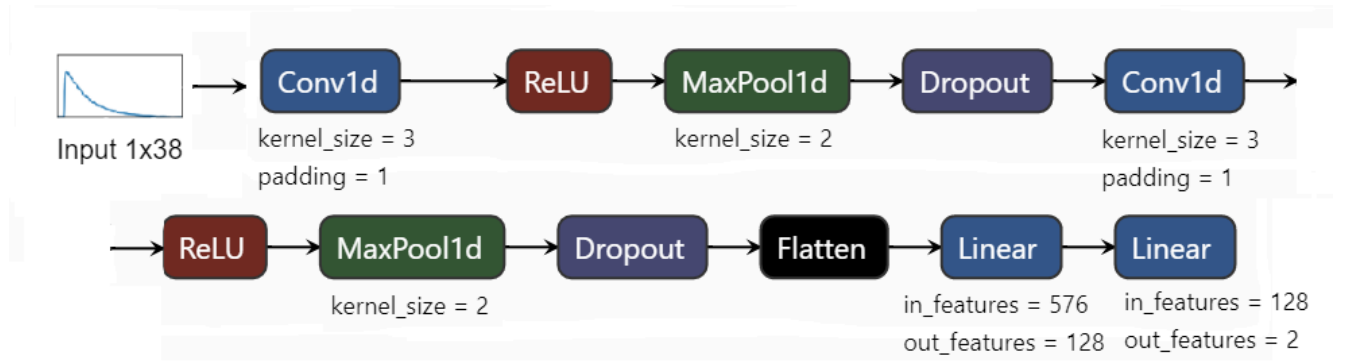


Рис. 2.4: Архитектура нейронной сети

В работе применялась нейронная сеть, состоящая из двух частей. Первая часть представляет из себя сверточную сеть, входом которой является вектор размерности $(N, 38)$, где N — это размер батча, а выходом вектор (эмбеддинг), размерности $(N, 128)$. Этот слой обучался с помощью функции потерь Sub-center ArcFace [33], которая имеет следующий вид:

$$\mathcal{L}_{ArcFace_{subcenter}} = -\log \frac{e^{\cos(\theta_{i,y_i}+m)}}{e^{\cos(\theta_{i,y_i}+m)} + \sum_{j=1, j \neq y_i}^N e^{\cos(\theta_{i,j})}}, \quad (2.1)$$

где $\theta_{i,j} = \arccos(\max_k(W_{j_k}^T x_i))$, $k \in \{1, \dots, K\}$ - это косинусное расстояние между элементом выборки и одним из центров класса.

Отличительная идея функции потерь (2.1) состоит в введении гиперпараметра m , отвечающего за некоторую пустую область (decision margin) между элементами разных классов в гиперпространстве размерности, соответствующей размеру эмбединга. Благодаря этому, элементы, лежащие на границе классов и в пересекающихся областях, классифицируются точнее, чем при использовании стандартной перекрестной энтропии.

На рисунке 2.5 представлена визуализация эмбедингов с помощью метода t-SNE [34]. Слева изображено представление выборки входных данных, справа результат после обучения.

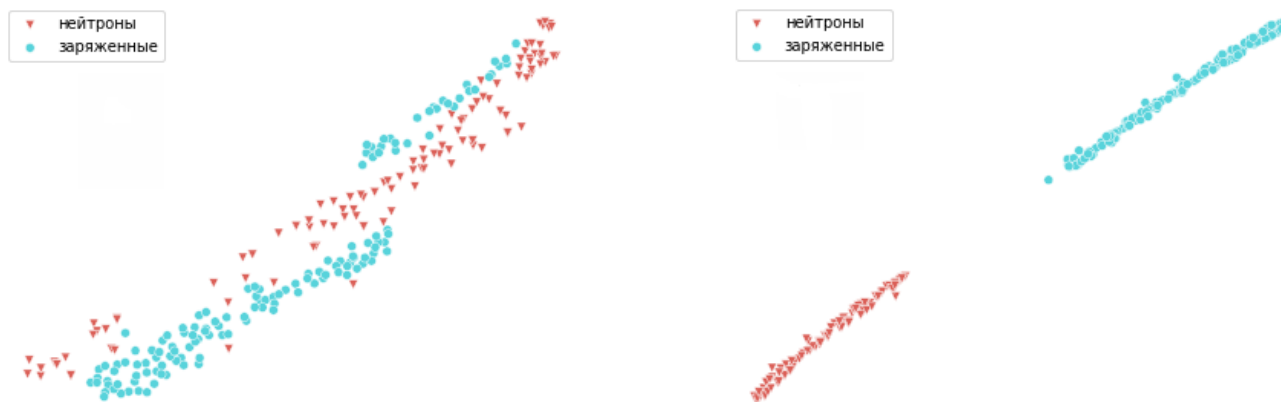


Рис. 2.5: t-SNE представление эмбедингов. Исходные данные (слева), после обучения (справа).

Вторая часть являлась логистической регрессией, которая переводила эмбединг в конечный вектор из двух значений, которые соответствовали вероятностям принадлежности к каждому из классов (заряженной компоненте или нейтронной).

Модель обучалась в течение 100 эпох при размере батча в 32 элемента. В качестве оптимизатора использовался Adam [35] с коэффициентом скорости обучения $\text{lr} = 10^{-3}$ и весовым коэффициентом L2 регуляризации, равным 10^{-5} . На рисунке 2.6 представлен график потерь и точности в зависимости от эпохи обучения.

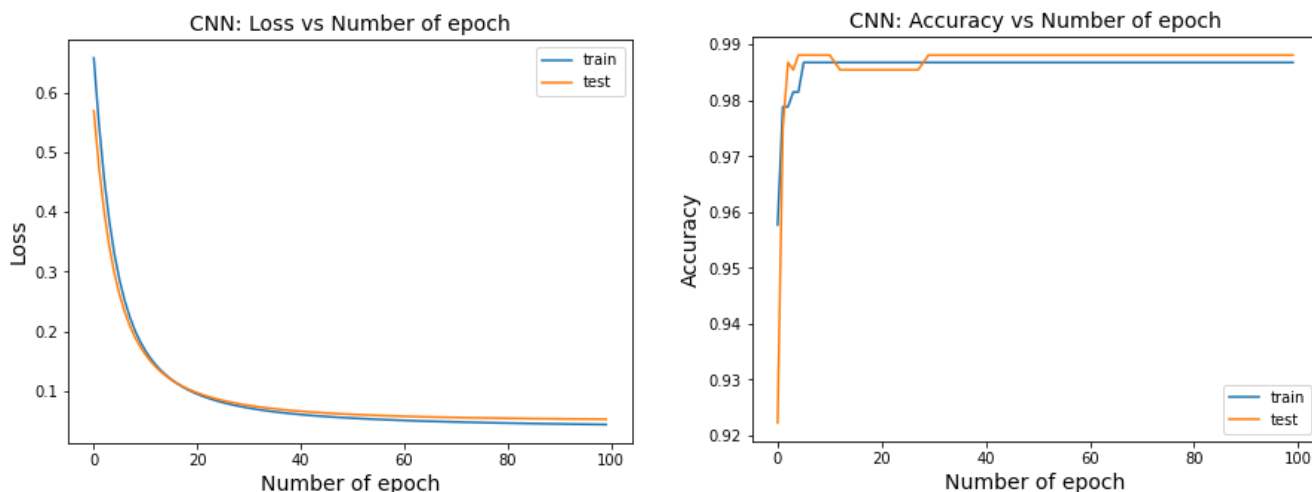


Рис. 2.6: Графики обучения нейронной сети. График зависимости потерь от эпохи (слева), доли верных ответов от эпохи (справа)

2.4. Результаты сравнения

Эффективность методов разделения определялась путем оценки основных метрик классификации на основе матрицы ошибок, таких как precision, recall и F1 мера на тестирующей выборке. Значение метрик оценивалось как средневзвешенное значение по обоим классам. На рисунке 2.7 представлены матрицы ошибок для метода, использующегося в настоящее время в установке, и метода, основанного на машинном обучении.

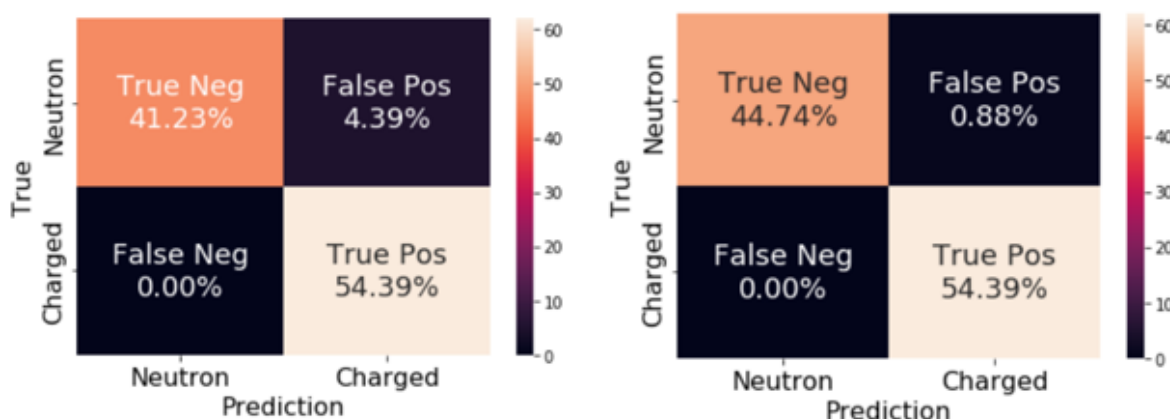


Рис. 2.7: Матрица ошибок. Базисный метод (слева), CNN (справа)

Значения, расположенные в левом нижнем и правом верхнем углу мат-

рицы, показывают процент неверно классифицированных импульсов. В таблице 2.1 представлены значения основных метрик для сравнения для тестирующей выборки.

Таблица 2.1: Сравнение результатов для тестирующей выборки

Метрика	Базисный метод	CNN
Precision	95%	99%
Recall	96%	99%
F1-score	96%	99%

Алгоритм с использованием сверточных нейронных сетей показал лучшее качество классификации на тестирующей выборке по сравнению с базисным методом, который применяется на данное время в установке. В дальнейшем разработанный метод будет применен в эксперименте ENDA.

Глава 3

Моделирование эксперимента

Моделирование эксперимента состоит из двух частей. Первая часть связана с моделированием ШАЛ, вторая - с моделированием отклика детекторов на его прохождение. Для последнего использовалась программа моделирования, написанная с помощью программного пакета Geant4 [36]. Недостатком данной программы является её низкая скорость вычислений. Одна из целей данной работы заключалась в создании программы быстрого Монте-Карло моделирования и максимального приближения её к результатам программы, написанной в Geant4.

3.1. Моделирование ШАЛ

Для моделирования ШАЛ в работе использовалась программа CORSIKA [37] версии 7.7402. Для моделирования электромагнитной компоненты ливня используется программа EGS4, для низких энергий адронных взаимодействий используется модель GHEISHA 2002d [38], а для высоких QGSJET-II-04 [39], EPOS-LHC v.3400 [40] и SIBYLL-2.3d [41]. В настоящий момент нет однозначного ответа на вопрос о том, какая модель лучше, к тому же в различных моделях предсказывается разное содержание частиц в ливне. Уровень наблюдения был равен 170 м над уровнем моря, геомагнитное поле равно $B_x = 16.5 \mu T$, $B_y = 49.6 \mu T$. Для адронов (кроме π^0) нижний энергетический порог был равен 50 МэВ; для $e^\pm$, γ и π^0 он равен 60 кэВ, для $\mu^\pm$ 100 МэВ.

Ливни разыгрывались без процедуры прореживания в энергетическом диапазоне от 10^{15} до 10^{17} эВ с показателем дифференциального спектра $(\gamma + 1) = 2$, который затем в ходе анализа перевзвешивался к 2.7. Зенитный угол был в диапазоне от 0 до 45 градусов в предположении изотропности распределения направления прихода по небесной сфере. В качестве первичных частиц были

взяты p, He, N, Si, Fe .

3.2. Описание программы моделирования в Geant4

Программа не включает в себя процесс моделирования оптических фотонов, и соответственно, ФЭУ и электроники, т.к. на данном этапе существуют большие сложности, связанные с определением оптических свойств используемого в эксперименте сцинтиллятора. Поэтому в программе моделирования рассчитывается только энерговыделение в сцинтилляторе (интегрированное в течение первой мкс после прихода фронта ливня) при прохождении через него заряженных частиц и число зарегистрированных нейтронов, которое определяется условием образования α -частицы и ${}^7\text{Li}$ в аналогичном эксперименту временном интервале.

В модель детектора была заложена упрощенная модель сцинтиллятора, представляющая собой сплошной слой компаунда $\text{ZnS}(\text{Ag}) + \text{B}_2\text{O}_3$, (массовое соотношение которых 65% на 35%) диспергированный в слое силикона, пластиковый корпус детектора, грунт и воздух. Сцинтиллятор задан диском с диаметром 69 см и толщиной 0.15 мм, который помещается в слой силикона толщиной 3 мм. Корпус детектора представляет собой полый цилиндр с высотой 70 см и толщиной стенок 8 мм. За материал грунта был взят бетон с плотностью 1.8 г/см^3 , который был задан прямоугольным параллелепипедом с толщиной в 20 метров. В качестве списка физических моделей взаимодействия частиц использовался набор QGSP_BIC_HP, который является комбинацией модели "QGS" (Quark Gluon String), используемой для описания адронных взаимодействий высоких энергий ($> \sim 15 \text{ ГэВ}$), "P" (Precompound), для описания процессов фрагментации и снятия возбуждения ядер, "BIC" (Binary Cascade), для описания адронных взаимодействий низких энергий ($< \sim 10 \text{ ГэВ}$) и "HP" (High Precision neutron model), которая используется для описания нейтронных взаимодействий повышенной точности для энергий ($< 20 \text{ МэВ}$).

В выходной файл программы записывается энергосодержание и число зарегистрированных нейтронов в каждом из детекторов, а также дополнительная информация о событии ШАЛ, аналогичная экспериментальному формату данных, за исключением дополнительной информации о направлении прихода ливня, типе и энергии первичной частицы.

3.2.1. Моделирование отклика эн-детектора на различные частицы

Для получения необходимых зависимостей и распределений, которые в дальнейшем будут использоваться в программе быстрого Монте-Карло, с помощью Geant4 было проведено моделирование отклика эн-детектора на все типы вторичных частиц ШАЛ при различных фиксированных энергиях (таблица 3.1).

Таблица 3.1: Тип и энергия разыгранных частиц

Тип частицы	Фиксированная энергия, МэВ
γ	$5 \cdot 10^{-2}, 10^{-1}, 3 \cdot 10^{-1}, 5 \cdot 10^{-1}, 1, 3, 5, 7, 10, 20,$ $30, 50, 70, 10^2, 3 \cdot 10^2, 10^3, 10^4$
$e^{\pm}$	$1, 3, 5, 7, 10, 20, 30, 50, 70, 10^2, 3 \cdot 10^2, 10^3, 10^4$
$\mu^{\pm}$	$10, 30, 10^2, 3 \cdot 10^2, 10^3, 10^4, 10^5$
h кроме π^0	$50, 10^2, 3 \cdot 10^2, 5 \cdot 10^2, 10^3, 3 \cdot 10^3, 10^4, 10^5$

По полученным выборкам были найдены зависимости эффективности одиночной частицы дать отклик детектора от её энергии (рисунок 3.1). Под откликом детектора в данном случае понимается любое ненулевое энергосодержание в сцинтилляторе.

Из-за того, что сцинтиллятор тонкий, одиночные гамма-кванты почти не взаимодействуют с сцинтиллятором и их эффективность дать энергосодержание в детекторе находится на уровне нескольких процентов. Электроны с энергией

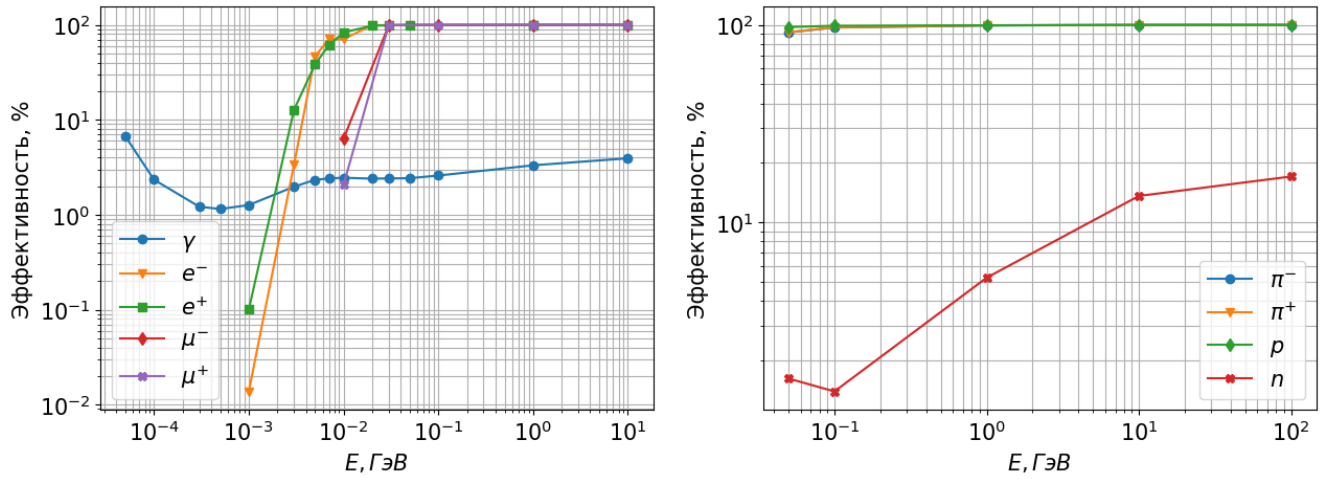


Рис. 3.1: Эффективность γ , $e^\pm$, $\mu^\pm$ (слева) и p , n , $\pi^\pm$ (справа) дать отклик в эн-детекторе в зависимости от энергии частицы

ниже 3 МэВ в подавляющем большинстве не долетают до сцинтиллятора и теряют всю свою энергию в пластиковой бочке. Заряженные адроны и мюоны при энергии 30 МэВ и выше почти всегда дают отклик детектора. Эн-детекторы установки ENDA-INR имеют внешний стальной корпус толщины 0.5 мм (рис. 3.2), из-за чего вероятность регистрации гамма-квантов возрастает за счет их конверсии в электрон-позитронные пары. Для численной оценки этого влияния было проведено сравнение эффективности гамма-квантов дать отклик эн-детектора со стальным корпусом и без.

Как видно из графика эффективность для эн-детектора со стальным корпусом ниже при низких энергиях, что связано с тем, что гамма-кванты, взаимодействуя со стальным корпусом, рождают электрон-позитронную пару, а их энергии не хватает пройти пластиковый корпус, при больших же энергиях эффективность дать отклик детектора выше $\sim 1\%$. В п. 3.2.3 будет показано, что вклад гамма-квантов в суммарное энерговыделение в событиях ШАЛ находится на уровне $\sim 10\%$, и из-за чего влиянием внешнего стального корпуса было решено пренебречь и при моделировании он не закладывался в модель установки.

Частица, не попавшая в детектор, тоже может дать свой вклад в энерго-

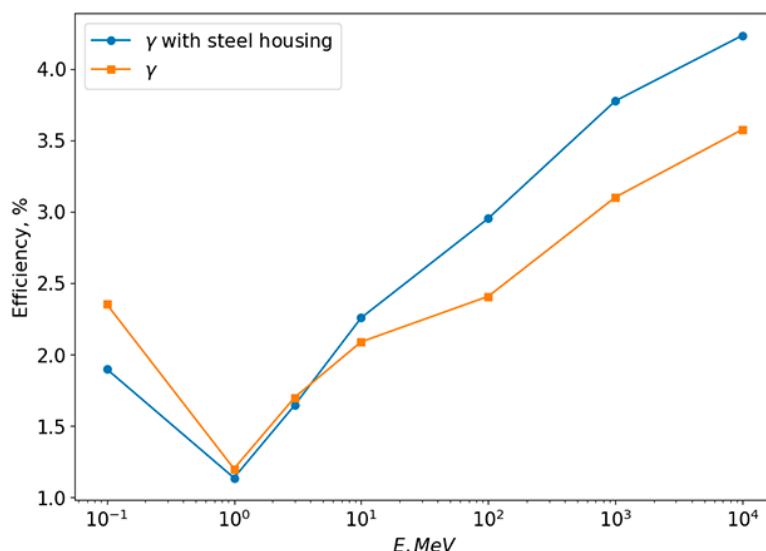
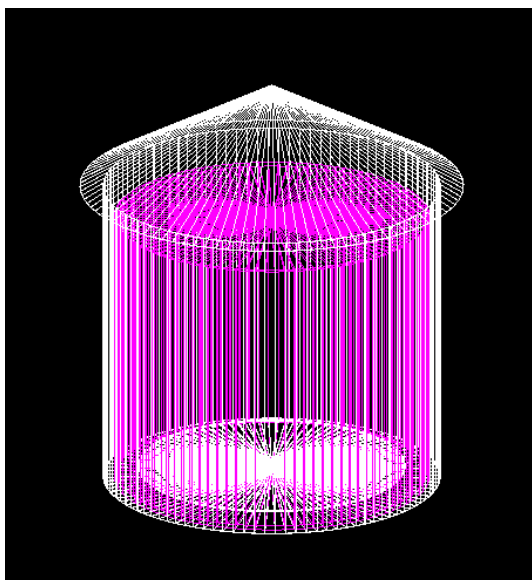


Рис. 3.2: Визуализация модели эн-детектора со стальным корпусом (слева), эффективность гамма-кванта дать отклик эн-детектора (справа) со стальным корпусом (синяя кривая), без стального корпуса (оранжевая кривая)

выделение, заряженные частицы за счет тормозного излучения, а гамма-кванты за счет рождения электрон-позитронных пар. Для учета таких событий, были получены зависимости эффективности дать отклик детектора в зависимости от расстояния от центра детектора для гамма-квантов и электронов (рисунок 3.3).

Из рисунка видно, что уже на расстоянии 1 м от центра детектора эффективность дать ненулевое энергосодержание в детекторе падает на два порядка для гамма-квантов и на три порядка для электронов. С целью упрощения вычислений принималось, что позитроны имеют аналогичную электронам зависимость эффективности и такие же распределения по энергосодержанию. Также, так как вклад от тормозных потерь может дать любая заряженная частица, но электронно-фотонная компонента составляет около 98% всех частиц ливня (в центральной области ШАЛ), то вклад в энергосодержание от других частиц, не попавших в детектор, не учитывался.

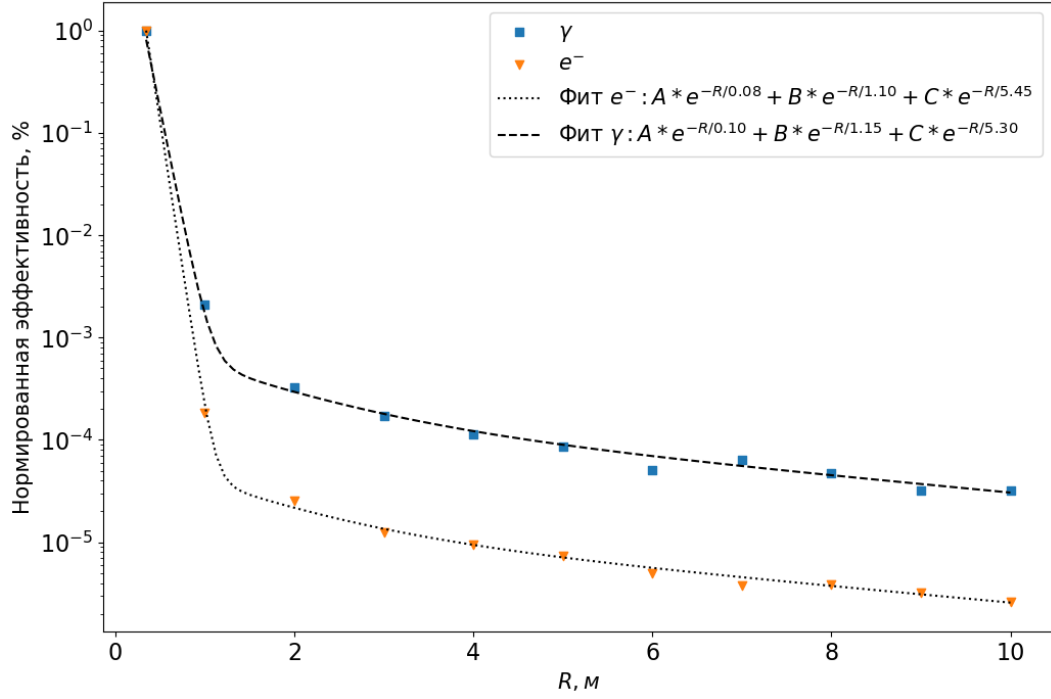


Рис. 3.3: Зависимость эффективности дать отклик детектора в зависимости от расстояния от центра детектора для гамма-квантов и электронов

3.2.2. Моделирование числа зарегистрированных нейтронов

Нейтроны, регистрируемые эн-детекторами, рождаются в адрон-ядерных и фотоядерных взаимодействиях. Их число зависит от энергии налетающего адрона или гамма-кванта, а также от расстояния от места реакции до центра детектора. Для получения первой зависимости были смоделированы выборки, в которых адроны и гамма-кванты различных энергий бросались в центр эн-детектора, затем были оценены средние значения числа зарегистрированных нейтронов на 1 брошенную частицу и получена степенная аппроксимация вида $\bar{n} \sim E^\alpha$ (рисунок 3.4), где $\alpha = 0.53 \pm 0.04$ для протонов и $\alpha = 0.54 \pm 0.02$ для гамма-квантов. Полученные зависимости хорошо согласуются с результатами, полученными на нейтронном мониторе, где средняя кратность регистрируемых испарительных нейтронов от протонов космических лучей была $\sim E^{0.56}$ для $5 < E \leq 1000$ ГэВ [42].

Как видно из рисунка 3.4 число зарегистрированных нейтронов на 1 бро-

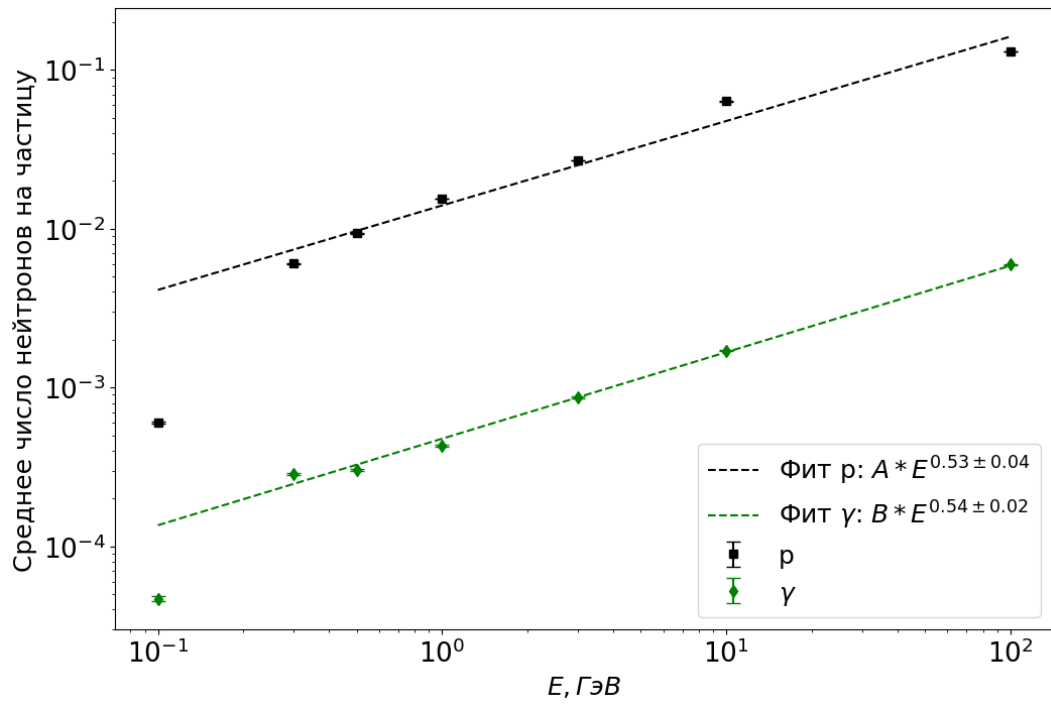


Рис. 3.4: Зависимость среднего числа зарегистрированных нейтронов в зависимости от энергии налетающей частицы для протона и гамма-кванта

шенную частицу почти на два порядка меньше для гамма-квантов, чем для протонов, а энергии в 100 МэВ не хватает для рождения нейтронов в грунте. Отсюда следует, что массовое рождение нейтронов на уровне наблюдения начинается тогда, когда высокоэнергичные адроны ШАЛ достигают этого уровня.

Для получения функции пространственного распределения (ФПР) зарегистрированных нейтронов относительно “родительских” адронов были смоделированы выборки, в которых высокоэнергичные протоны бросались на разных расстояниях от центра детектора с шагом в 1 метр. Для оценки влияния погодных условий на число зарегистрированных нейтронов была взята модель сухого грунта и грунта с 10% влажностью (рисунок 3.5).

Вид ФПР из рисунка 3.5 можно описать двумя экспонентами следующего вида (3.1):

$$f(R) = A \times e^{-R/r_1} + B \times e^{-R/r_2}, \quad (3.1)$$

где $r_1 = 0.36$ м и $r_2 = 2.65$ м.

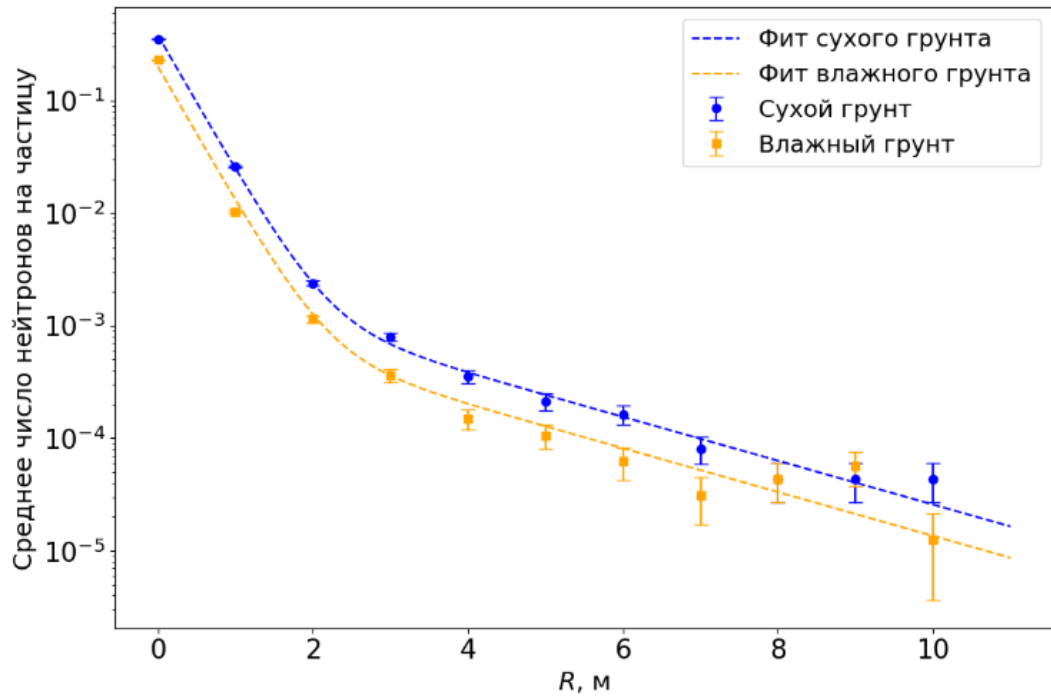


Рис. 3.5: Зависимость среднего числа зарегистрированных нейтронов на 1 протон в зависимости от расстояния от места ядерной реакции до детектора

Первый параметр (3.1) связан с площадью детектора, а второй связан с нейтронами, которые родились в грунте и пришли с больших расстояний через диффузию в воздухе. В рамках упрощения считалось, что параметры экспонент не зависят от влажности грунта, а меняется только положение ФПР относительно положения при сухом грунте. Как видно из рисунка, такой фит, умноженный на коэффициент, связанный с уменьшением числа нейтронов, неплохо описывает данные для 10% влажного грунта. Среднее же число зарегистрированных нейтронов оказалось в 1.67 раз больше для сухого грунта.

3.2.3. Моделирование отклика установки на прохождение ШАЛ

На данном этапе было проведено моделирование отклика установки на прохождение ШАЛ. Выборка состояла из 6000 событий, где в качестве первичной частицы был выбран протон, с энергией от 1 до 100 ПэВ и $\frac{dN}{dE_0} \sim E_0^{-2.7}$, а ось разыгрывалась равномерно в радиусе от 0 до 15 м от центра установки. На

рисунке 3.6 представлены распределения по энергосвечению в эн-детекторах для каждого отдельного типа вторичных частиц ШАЛ.

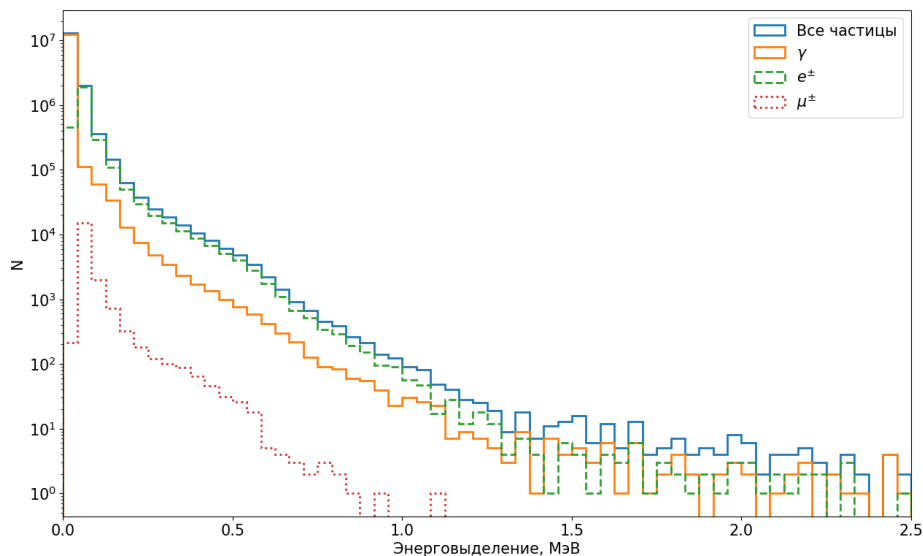


Рис. 3.6: Распределения энергосвечения в эн-детекторе от прохождения ШАЛ

Основной вклад в энергосвечение вносят электроны и позитроны, их доля относительно всех других частиц составляет $\sim 88.7\%$, доля гамма-квантов $\sim 10.3\%$. Средняя по всем событиям ШАЛ вероятность дать отклик детектора для $e^\pm$ равна 85.4% , для гамма-квантов 2.3% . Среднее энергосвечение для $e^\pm$ равно 70.9 кэВ, для мюонов 78.7 кэВ, а для гамма-квантов 2.4 кэВ.

За коэффициент пересчета от суммарного энергосвечения к числу заряженных частиц было выбрано значение среднего энергосвечения от $e^\pm$, так как их вклад в суммарное энергосвечение подавляющий.

3.3. Программа быстрого моделирования отклика установки

Программа быстрого Монте-Карло была написана на языке C++ с использованием технологий параллельного программирования. Энергосвечение в программе разыгрывается на основе распределений по энергосвечению и функций плотности вероятности, которые были получены в п. 3.2.1, а число зарегистрированных нейтронов в эн-детекторе на основе функций, полученных

в п. 3.2.2.

Для сэмплирования полученных распределений использовался метод Неймана (выборка с отклонением) и взвешенная линейная интерполяция для распределений [43] (3.2).

$$\hat{p}_{i,t} = \frac{t - t_0}{t_1 - t_0} \cdot (p_{i,t_1} - p_{i,t_0}) + p_{i,t_0}; t \in \{t_0, t_1\}; \sum_{i=0}^N p_{i,t} = 1, \quad (3.2)$$

где N — количество бинов, t — значение промежуточной энергии, t_0, t_1 — левая и правая граница интервала энергии, $p_{i,t}$ — функция массы вероятности.

На рисунке 3.7 представлен пример полученных распределений по энергосодержанию для μ^- с наложенным распределением, полученным из программы быстрого моделирования.

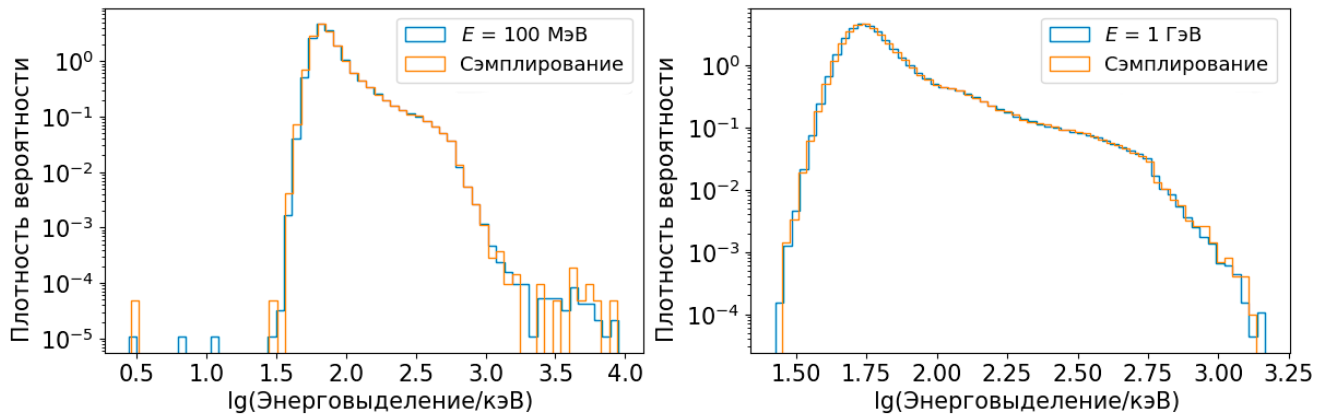


Рис. 3.7: Распределения энергосодержаний в эн-детекторе для мюона с энергией в 100 МэВ (слева) и 1 ГэВ (справа)

3.3.1. Сравнение результатов с Geant4

Сравнение программ моделирования проводилось на основе рассчитанных откликов установки на прохождения ШАЛ. На первом этапе с помощью программы CORSIKA были смоделированы выборки при фиксированных энергиях первичной частицы (1, 3 и 10 ПэВ), в качестве первичной частицы был взят протон, зенитный угол был равен 0° . Уровень наблюдения был равен 4400

м над уровнем моря. Ось ливня бросалась в одно и то же место около одного из центральных детекторов. Затем было проведено сравнение числа зарегистрированных частиц в обеих программах моделирования. На рисунке 3.8 представлены распределения по числу частиц в четырех детекторах, расположенных на разном расстоянии от оси ШАЛ для энергии первичной частицы в 1 ПэВ, вертикальными линиями обозначены средние значения соответствующих распределений. Распределения энерговыделений для всех детекторов приведено в приложении.

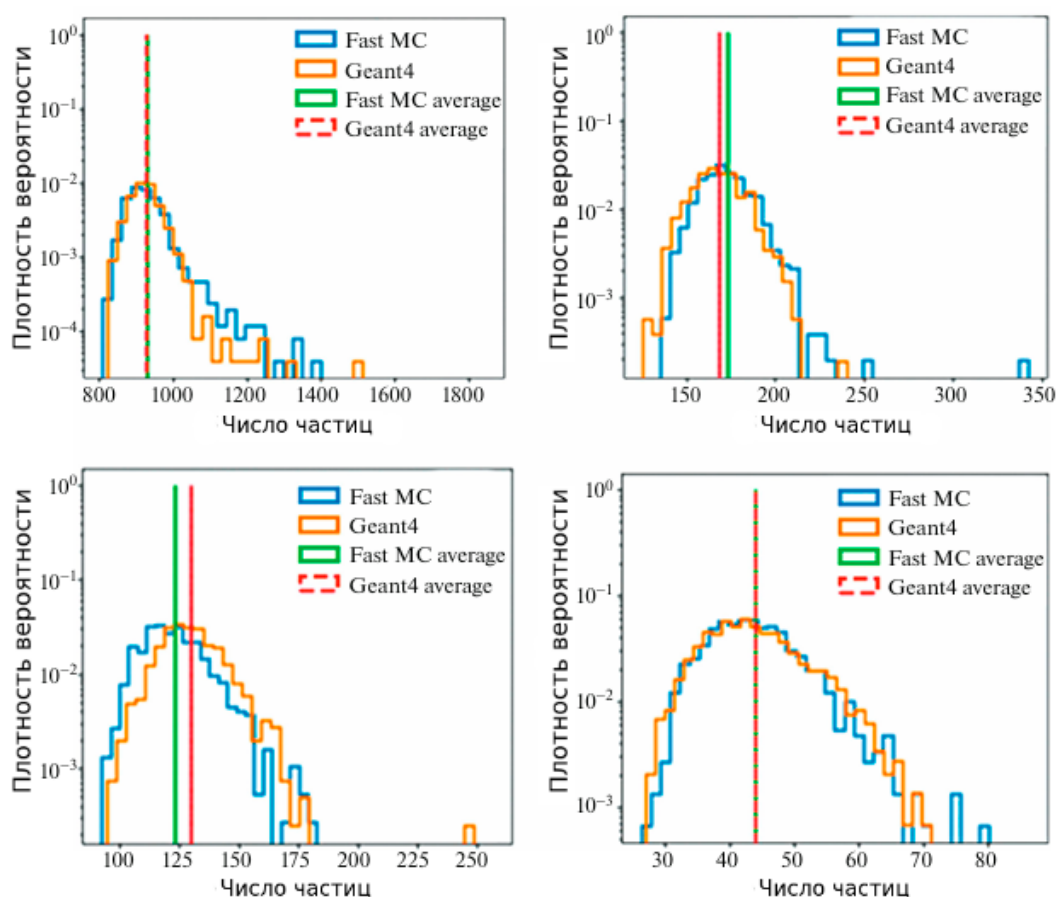


Рис. 3.8: Распределения по числу частиц в четырёх детекторах для программы быстрого моделирования (синий цвет) и Geant4 (оранжевый цвет)

Как видно из рисунка 3.8, обе программы дали очень близкие результаты, а относительная разница среднего числа частиц не превышала 5%. Для проверки корректности розыгрыша числа зарегистрированных нейтронов сравнивались распределения суммарного числа зарегистрированных нейтронов в

события в обеих программах моделирования (рис. 3.9).

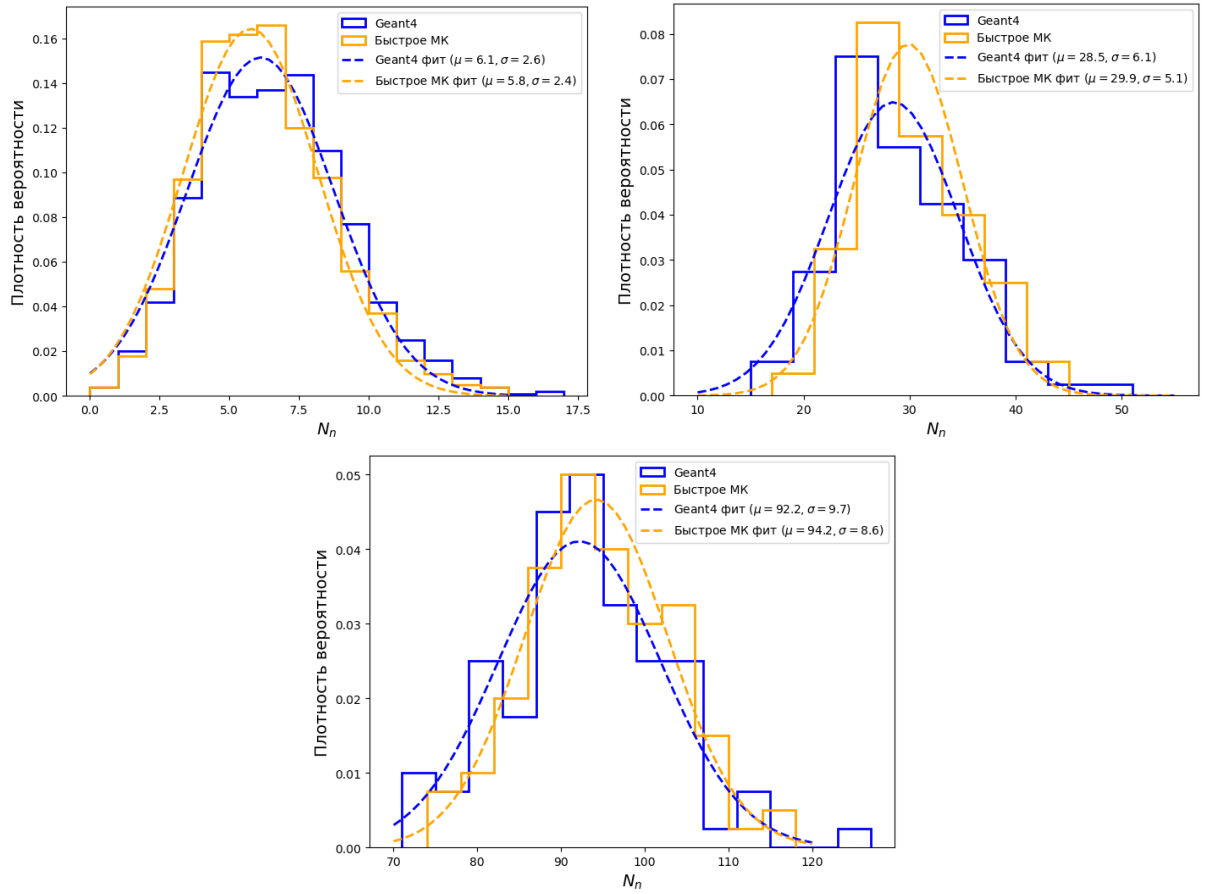


Рис. 3.9: Распределения по числу зарегистрированных нейтронов для программы быстрого моделирования (оранжевый цвет) и Geant4 (синий цвет) для $E_0 = 1$ ПэВ (слева сверху), $E_0 = 3$ ПэВ (справа сверху) и для $E_0 = 10$ ПэВ (снизу)

Распределения из рисунков выше аппроксимировались нормальным распределением, в результате относительная ошибка среднего значения оказалась на 5.2% ниже при энергии 1 ПэВ, на 4.7% при 3 ПэВ и на 2.1% выше при 10 ПэВ для программы быстрого Монте-Карло, а время вычисления оказалось меньше в ~ 2000 раз.

3.3.2. Моделирование выборок по отклику установки

На данном этапе с помощью программы быстрого моделирования и событий ШАЛ из пункта 3.1 были получены выборки по отклику установки

ENDA-INR. Положение оси каждого ливня разыгрывалось равномерно в радиусе 30 метров от центра установки (рис. 3.10), каждый ливень был использован несколько раз для увеличения итоговой статистики событий.

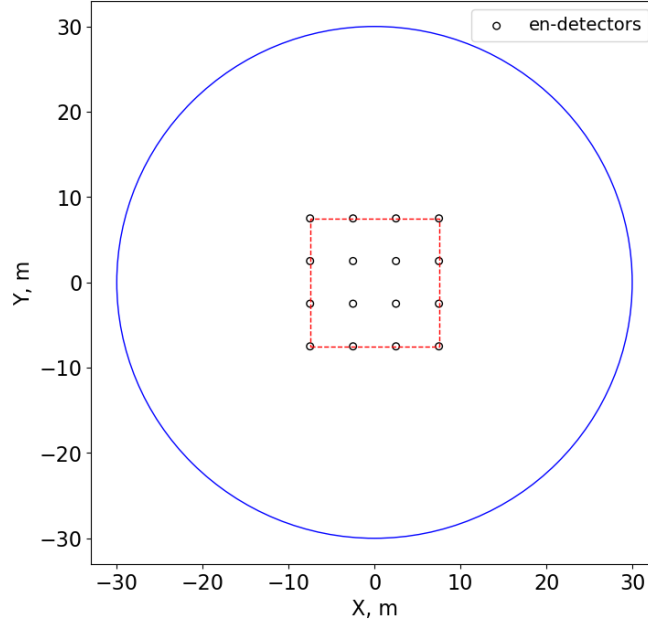


Рис. 3.10: Красная штриховая область – границы установки, синий круг – зона розыгрыша оси ШАЛ

Для обучения моделей машинного обучения (см. Главу 4) использовалась выборка, полученная с помощью модели QGSJET-II-04. В качестве валидационных выборок, целью которых служила проверка алгоритмов реконструкции, были выбраны независимые выборки, созданные с использованием моделей EPOS-LHC, SIBYLL и QGSJET-II-04.

Для модели QGSJET-II-04 размер обучающей выборки составил $2.2 \cdot 10^5$ событий, а валидационной - $8.1 \cdot 10^4$ событий, для модели EPOS-LHC размер выборки составил $7.9 \cdot 10^4$ событий, а для модели SIBYLL-2.3d - $8.1 \cdot 10^4$ событий. Из-за того, что 100%-ая эффективность регистрации ШАЛ для легкой компоненты наступает по энергии первичной частицы раньше, чем у тяжелой компоненты (см. 4.1), соотношение массовых групп в исходной выборке неравномерное.

Глава 4

Реконструкция событий

4.1. Эффективность регистрации ШАЛ

Эффективность регистрации события ШАЛ (ε_{trig}) определяется как отношение числа событий, давших срабатывание триггера (N_{trig}), к их полному числу (N_{tot}) (4.1):

$$\varepsilon_{eff} = \frac{N_{trig}}{N_{tot}}. \quad (4.1)$$

На рисунке 4.1 представлены зависимости эффективности регистрации для триггеров второго уровня от энергии первичной частицы для протона и ядер железа.

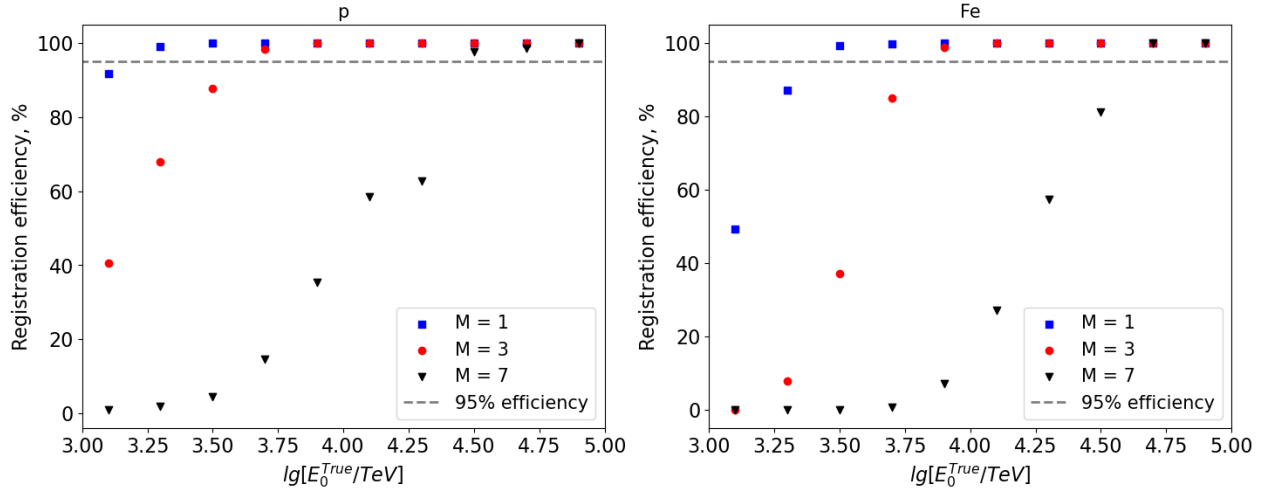


Рис. 4.1: Зависимость эффективности регистрации от энергии первичной частицы для протона (слева) и железа (справа)

Эффективность регистрации для $M=1$ превышает 95% с энергии ~ 1.5 ПэВ, для протона и с ~ 2.5 ПэВ для железа, для $M=3$ с энергии ~ 5 ПэВ и ~ 7.9 ПэВ, а для $M=7$ с ~ 25 ПэВ и с ~ 35 ПэВ соответственно.

На рисунке 4.2 представлены зависимости эффективности регистрации для двух различных критериев отбора в зависимости от энергии первичной частицы.

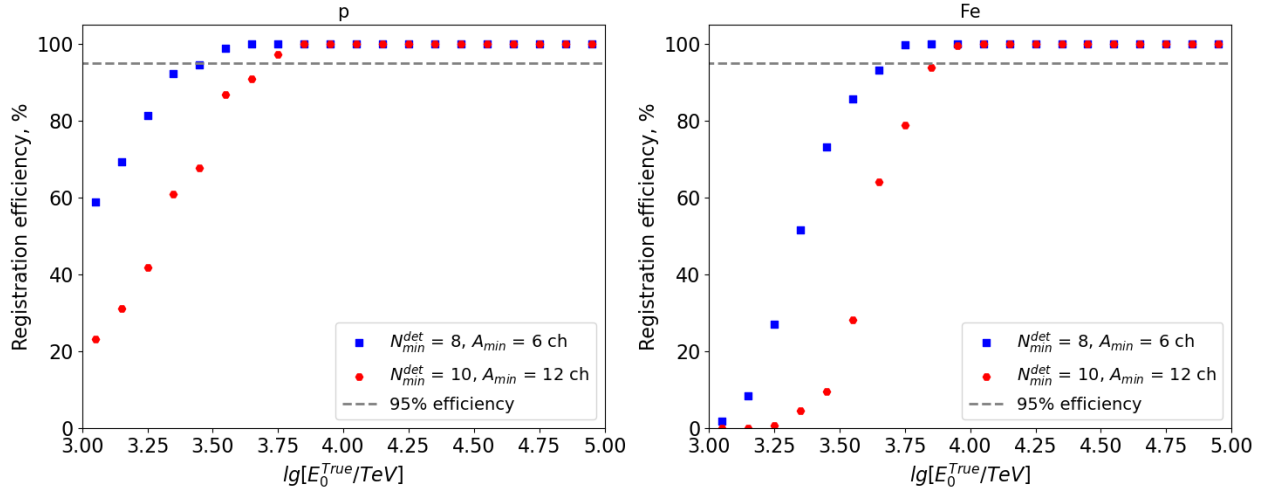


Рис. 4.2: Зависимость эффективности регистрации от энергии первичной частицы для протона (слева) и железа (справа). N_{min}^{det} , A_{min} — минимальное количество и амплитуда сработавших эн-детекторов

Эффективность регистрации для критерия с $N_{min}^{det}=8$ и $A_{min}=6$ каналов АЦП превышает 95% с энергии ~ 3.2 ПэВ, для протона и с ~ 5 ПэВ для железа, для критерия с $N_{min}^{det}=10$ и $A_{min}=12$ каналов АЦП с энергии ~ 5 ПэВ, и с ~ 7.9 ПэВ соответственно. 100%-ая эффективность регистрации достигается с энергии ~ 5.0 ПэВ и ~ 7.1 ПэВ для двух критериев отбора для протона и с ~ 6.3 ПэВ и ~ 8.9 ПэВ для железа. Более жесткий критерий отбора, соответствующий $N_{min}^{det}=10$ и $A_{min}=12$ каналам АЦП, будет в дальнейшем использоваться при обработке экспериментальных данных (см. Главу 5).

На рисунке 4.3 представлен разыгранный энергетический спектр первичных частиц, после отбора событий с критерием $N_{min}^{det}=8$ и $A_{min}=6$ каналов АЦП.

Помимо оценки эффективности регистрации, также важно оценить и эффективность реконструкции (ε_{rec}), так как реконструкция событий, прошедших триггер установки, не всегда дает положительный результат, в результате чего утерянные события искажают искомые реконструируемые спектры. Эффективность реконструкции определяется качеством реконструкции параметров ШАЛ (подробнее см. 4.2.2 и 4.3.1).

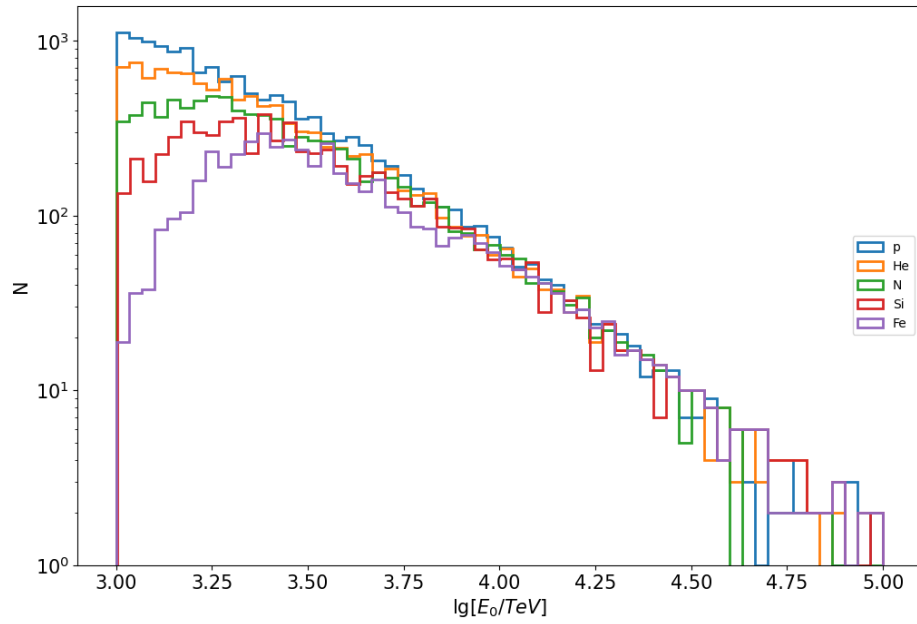


Рис. 4.3: Разыгранный энергетический спектр после первичного критерия отбора событий

4.2. Базисный метод восстановления параметров ШАЛ

Под восстановлением параметров ШАЛ понимается нахождение координат положения оси ливня в плоскости установки X и Y , направление его прихода (θ, φ) , полного числа заряженных частиц N_e , параметра возраста ливня s . Возраст ливня связан с наклоном функции пространственного распределения (ФПР) частиц ливня и является параметром, чувствительным к массовому составу космических лучей.

Традиционным подходом для нахождения параметров ливня считается метод максимального правдоподобия (ММП), в котором используют различные аналитические аппроксимации ФПР частиц ливня, исходя из теоретических, экспериментальных или модельных соображений.

4.2.1. Восстановление направления прихода ливня

Направление прихода ливня, задаваемое с помощью зенитного и азимутального угла (θ, φ) , определяется по относительным временам срабатывания

детекторов с помощью аппроксимации фронта ШАЛ. Из-за малых размеров установки использовалось приближение плоского фронта ливня.

Вторичные частицы ливня с большими временными задержками относительно времени прихода ливня могут привести к большим ошибкам при восстановлении направления прихода, поэтому в качестве аналитического алгоритма оптимизации был выбран метод наименьших квадратов с итеративным перевзвешиванием (IRLS, Iteratively Reweighted Least Squares) [44], который является более устойчивым к выбросам в данных алгоритмом. Для сравнения результатов в качестве базисного алгоритма для реконструкции был использован метод наименьших квадратов (МНК).

На каждом шаге в методе IRLS происходит решение взвешенной линейной задачи наименьших квадратов (4.2):

$$\beta^{(k+1)} = \arg \min_{\beta} \sum_{i=1}^{N_{det}} \omega_i^{(k)} \left| y_i - X_i \beta \right|^2 = (X^T W^{(k)} X)^{-1} X^T W^{(k)} y, \quad (4.2)$$

где β – матрица оптимизируемых коэффициентов $[\sin\theta\cos\theta \quad \sin\theta\sin\varphi \quad \cos\theta]$, k – шаг алгоритма (в МНК $k = 0$), N_{det} – число сработавших детекторов, X – матрица с координатами детекторов, y – матрица с временными задержками ($c \cdot t_i$), c – скорость света (0.2998 м/нс), $W^{(k)}$ – диагональная матрица весов (единичная матрица в МНК) с элементами $\omega_i^{(k)}$, задаваемых по (4.3):

$$\omega_i^{(k)} = \frac{1}{\max \left\{ \delta, |y_i - X_i \beta^{(k)}| \right\}}, \quad (4.3)$$

где $\delta = 10^{-5}$.

Окончание итерационного процесса достигается при выполнении условия (4.4):

$$\left\| \beta^{(k+1)} - \beta^{(k)} \right\|_1 \leq \varepsilon, \quad (4.4)$$

где за ε было принято значение равное 10^{-5} .

Оценка качества реконструкции будет проведена в п. 4.3.1.

4.2.2. Восстановление положения оси и мощности ливня

В традиционном подходе ключевым этапом реконструкции параметров ШАЛ является выбор ФПР. Наиболее часто используемой функцией, описывающей плотность пространственного распределения частиц ливня, является функция Нишимуры-Каматы-Грейзена (НКГ) (4.5) [45]. При выводе этой функции ливень рассматривался как сугубо электромагнитный без учета ядерно-каскадного процесса, а также для нулевого энергетического порога электронов. Впоследствии стало ясно, что такая аппроксимация недостаточно хорошо описывает ФПР реального ливня [46, 47].

$$\rho_e^{NKG}(r, s) = \frac{1}{r_m^2} \frac{\Gamma(4.5 - s)}{2\pi\Gamma(s)\Gamma(4.5 - 2s)} \left(\frac{r}{r_m}\right)^{s-2} \left(1 + \frac{r}{r_m}\right)^{s-4.5}, \quad (4.5)$$

где $0.5 < s < 1.5$, r_m — радиус Мольера (76 м для уровня Москвы).

В связи с этим в литературе существуют различные варианты ФПР, а также различные модификации функции НКГ. Модификации связаны с различными вариациями показателей степени, нормировочных коэффициентов, а также мольеровского радиуса и введением различных параметров возраста. Например, введение поперечного параметра возраста $s_\perp$ (модификация Учайкина или Лагутина) (4.6) [48].

$$\rho_e(r, s_\perp) = m^{-2} \cdot \rho_e^{NKG}\left(\frac{r}{m}, s_\perp\right), \quad (4.6)$$

где $m = 0.78 - 0.21 \cdot s_\perp$, и $0.6 \leq s_\perp \leq 1.8$.

Или же введением локального параметра возраста $s_{loc}(r)$, зависящего от расстояния от оси ливня (4.7).

$$\rho_e(r, s_{loc}) = C(s_{loc}) \cdot \left(\frac{r}{r_m}\right)^{s_{loc}+\alpha(r)-2} \left(1 + \frac{r}{r_m}\right)^{s_{loc}+\alpha(r)-4.5}, \quad (4.7)$$

где параметризация поправки $\alpha(r)$ выбирается экспериментально или на основе моделирования (в работе использовалась параметризация, полученная в эксперименте ШАЛ-МГУ [49] (рис. 4.4), а $C(s_{loc})$ — нормировочная константа, которая вычисляется численно, $0.5 < s_{loc} < 1.5$.

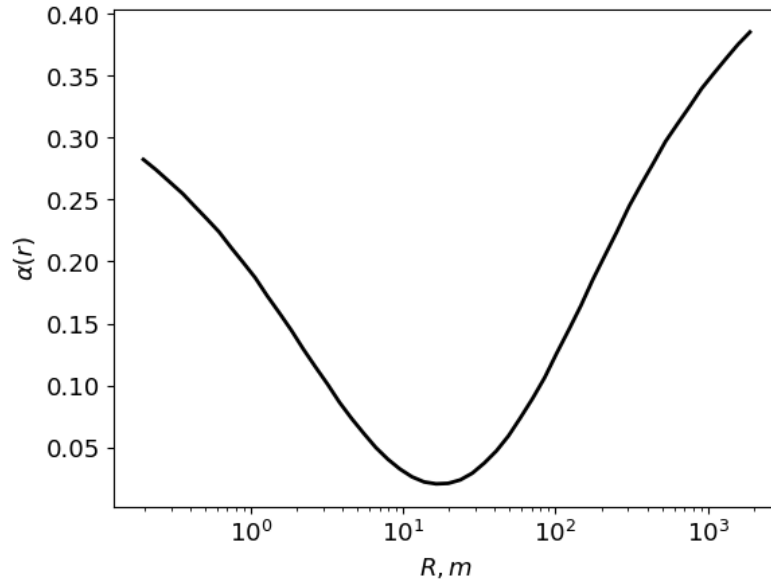


Рис. 4.4: Параметризация поправки $\alpha(r)$

На первом этапе была проведена качественная оценка аппроксимации плотности электронной компоненты, которая была получена с помощью программы CORSIKA для трёх вышеописанных ФПР. Для этого была получена средняя плотность электронной компоненты по 100 вертикальным ливням от протона с энергией $1 \div 3$ ПэВ и проведен фит с помощью функций (4.5-4.7) с использованием МНК (рис. 4.5).

На рисунке 4.6 представлен график отношения истинной плотности к оцененной с помощью вышеописанных ФПР.

Из приведенных выше рисунков видно, что НКГ даёт завышенную плотность электронной компоненты ШАЛ в области близкой к оси ливня, и заниженную на больших расстояниях от оси. Наиболее близкие значения к истинной плотности дают ФПР (4.6) и (4.7). ФПР (4.6) имеет заниженную оценку плотности на больших расстояниях от оси ливня (от 100 м), а ФПР (4.7) - в области от 20 до 50 метров, обе функции имеют заниженную оценку в области близкой к оси ливня.

На следующем этапе была произведена реконструкция параметров ливня с помощью ММП и с использованием всех трёх ФПР, для выборки, описанной

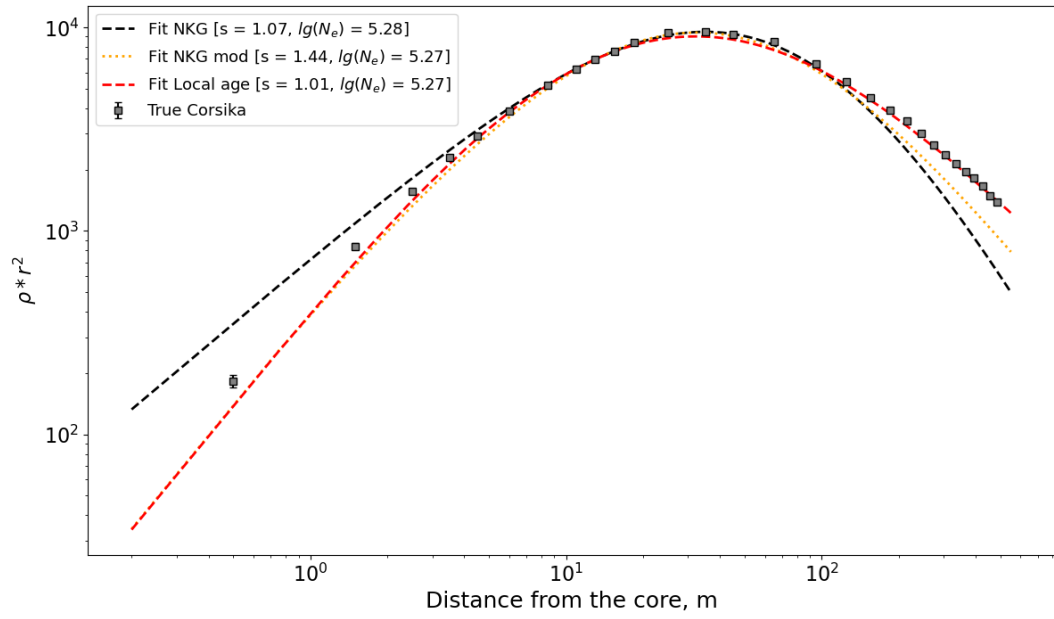


Рис. 4.5: Истинная плотность электронной компоненты и наложенные фиты различных ФПР. Функция (4.5) - чёрный цвет, (4.6) - жёлтый цвет, (4.7) - красный цвет

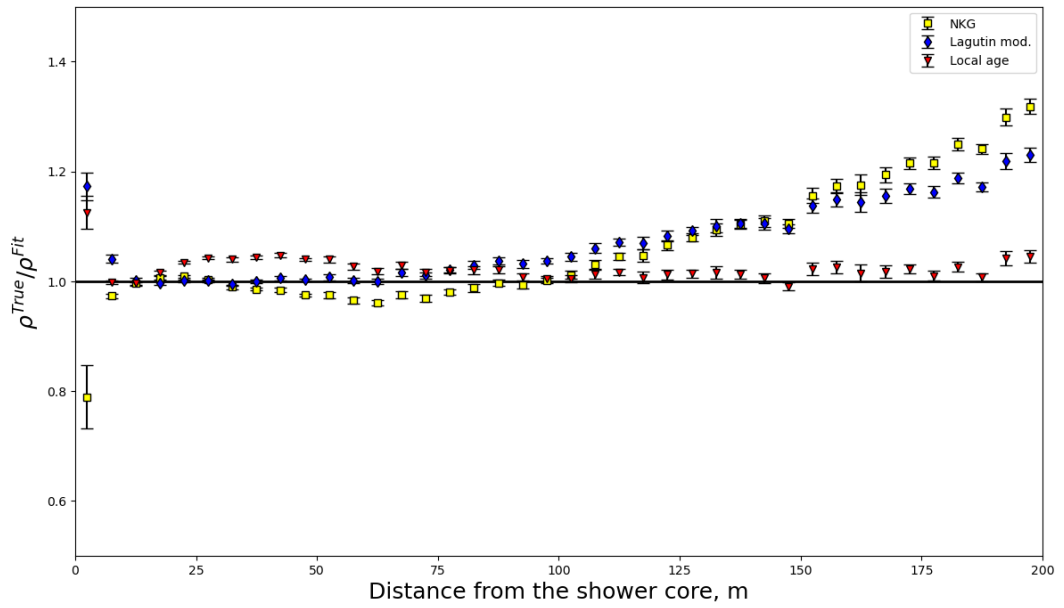


Рис. 4.6: Отношение истинной плотности к полученной в результате аппроксимации. Функция (4.5) - жёлтый цвет, (4.6) - синий цвет, (4.7) - красный цвет

в п. 3.3.2.

Логарифмический вид функции правдоподобия имел следующий вид

(4.8):

$$\mathcal{L} = \sum_i (n_i \ln N_i - N_i - \ln(n_i!)) + \sum_j \left[\ln \frac{1}{\sigma_j \sqrt{2\pi}} - \left(\frac{n_j - N_j}{2\sigma_j} \right)^2 \right], \quad (4.8)$$

где n_i - число зарегистрированных частиц в детекторе, $N_i = N_e \cdot \rho_e \cdot S$ - ожидаемое число частиц в плоскости перпендикулярной оси, которые могут попасть в детектор, площадью S на расстоянии r от оси ливня, N_e - мощность ливня, $\sigma_j = \sqrt{N_j}$.

При поиске максимума функции правдоподобия член $\ln(n_i!)$ можно отбросить, так как при переборе различных положений оси число попавших в детектор частиц является постоянной величиной. Первая сумма отвечает Пуассоновскому распределению и работает для малого числа частиц ($n < 25$), а вторая соответствует нормальному распределению и работает при ($n \geq 25$).

Для нахождения максимума функции правдоподобия применялся алгоритм Нелдера-Мида (симплекс-метод) [50]. Также были протестированы и другие алгоритмы оптимизации, такие как: метод Метрополиса-Гастингса [51] и алгоритм Пауэлла [52]. Качество восстановления параметров ливня у представленных алгоритмов очень близкие и находятся в границах 1%, лучшее же качество реконструкции показал симплекс-метод. Нулевые приближения для параметров определялись по следующим формулам (4.9):

$$x_0 = \sum_{i=1}^{N_{det}} (n_i \cdot X_i) / N_{ch}; \quad y_0 = \sum_{i=1}^{N_{det}} (n_i \cdot Y_i) / N_{ch}; \quad N_{e0} = \frac{N_{ch}}{S \cdot \sum_{i=1}^{N_{det}} \rho_e(r_i, s_0)}, \quad (4.9)$$

где N_{det} - общее число детекторов, (X_i, Y_i) - координаты i -го детектора, N_{ch} - суммарное число зарегистрированных частиц.

Нулевые приближения для возраста ливня разыгрывались по нормальному закону со средним значением равным 1.1, и среднеквадратичным отклонением, равным 0.1, $s_0 \sim \mathcal{N}(1.1, 0.1)$.

Для выбора диапазона отбираемых значений возраста ШАЛ, для событий, в которых положение оси ливня было определено в границах установки,

были построены двумерные распределения, где по оси абсцисс отложен логарифм отношения реконструированной и истинной мощности ливня, а по оси ординат - реконструированное значение возраста ливня (рис. 4.7), после чего была выбрана область, соответствующая $|\lg(N_e^{Rec.}/N_e^{True})| \leq 0.5$. За истинную мощность ливня принималось полное число электронов и позитронов в ливне на уровне наблюдения, влиянием гамма-квантов было пренебрежено из-за низкой чувствительности сцинтиллятора к ним (см п. 3.2.3).

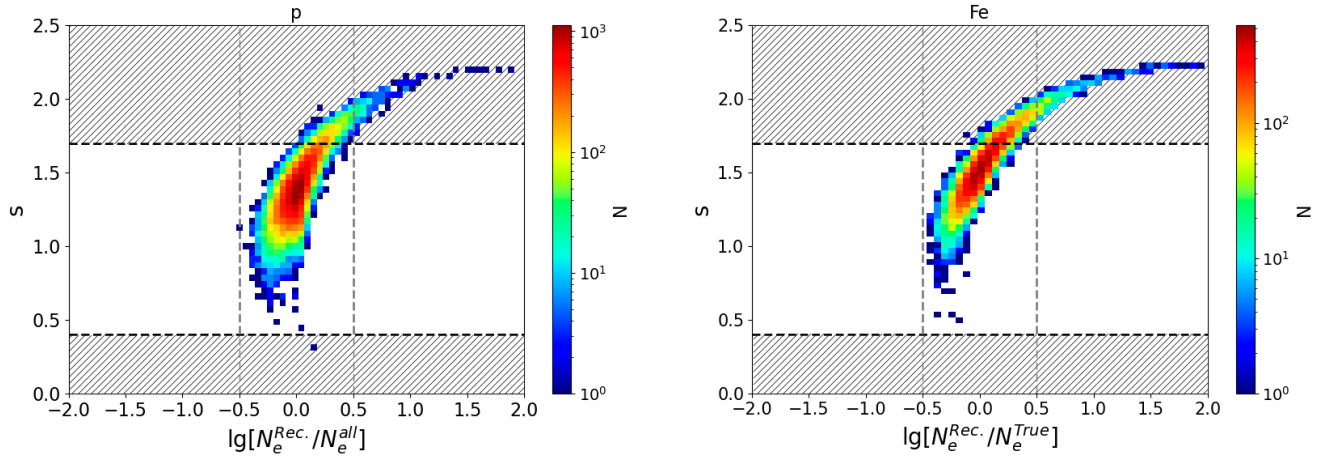


Рис. 4.7: Двумерные распределения логарифма отношения реконструированной и истинной мощности ШАЛ к реконструированному возрасту ливня для протона (слева) и железа (справа). Штриховая линия обозначает отбрасываемую область

Как видно из рисунка, ошибка реконструкции мощности ливня растет с увеличением значения возраста ливня, события с малым значением возраста ($\sim 0.4 \div 0.6$) имеют малую ошибку реконструкции. Исходя из анализа полученных распределений, для каждой ФПР были установлены следующие ограничения на значения возраста ливня:

1. $0.4 \leq s \leq 1.4$ для функции (4.5);
2. $0.4 \leq s_{\perp} \leq 1.7$ для функции (4.6);
3. $0.4 \leq s_{loc} \leq 1.3$ для функции (4.7).

После этого были отобраны события, в которых реконструированная ось ливня лежала в границах установки, а затем оценивались метрики классификации на основе её истинного положения, точность реконструкции оси, за которую принималась квантиль 68% распределения погрешности определения оси ливня, и абсолютная погрешность восстановления мощности ливня (err_{N_e}), определяемая как $\sigma(\lg(N_e^{\text{Rec.}}/N_e^{\text{True}}))$. В процессе классификации реконструированные события делились на события, в которых ось ливня была определена в границах установки и за её пределами. Метрика recall (efficiency) показывает, сколько событий с реконструированной осью ливня в границах установки удалось восстановить от всех событий, находящихся в её пределах, а метрика precision (purity) соответствует точности такой классификации. За оценку эффективности реконструкции (ε_{rec}) принималась метрика среднего гармонического recall и precision. В качестве итогового значения метрики принималось её среднее значение для обоих классов, так как классы не сбалансированы и событий в границах установки намного меньше. Результаты сравнения представлены в таблице 4.1.

Таблица 4.1: Сравнение результатов восстановления различных ФПР

ФПР	err_{N_e}	ΔR_{68} , м	Recall, %	Precision, %	$\varepsilon_{\text{rec}}^{\text{core}}$, %
(4.5)	0.17	1.73	84.3	81.6	82.9
(4.6)	0.17	1.71	87.0	83.2	84.9
(4.7)	0.19	2.11	87.3	81.9	84.2

Из результата сравнения видно, что наилучшее качество отбора осей показала модификация Лагутина (4.6), в дальнейшем она использовалась при обработке данных моделирования и эксперимента. События, в которых положение оси реконструировалось в границах установки, но истинное положение которой было за её пределами, присутствовали в выборках при использовании всех трёх ФПР. Для устранения таких событий, была использована информация о нейтронной компоненте. На рисунке 4.8 представлено распределение по-

грешности определения оси ливня во всех отобранных событиях и в событиях с нейтронами.

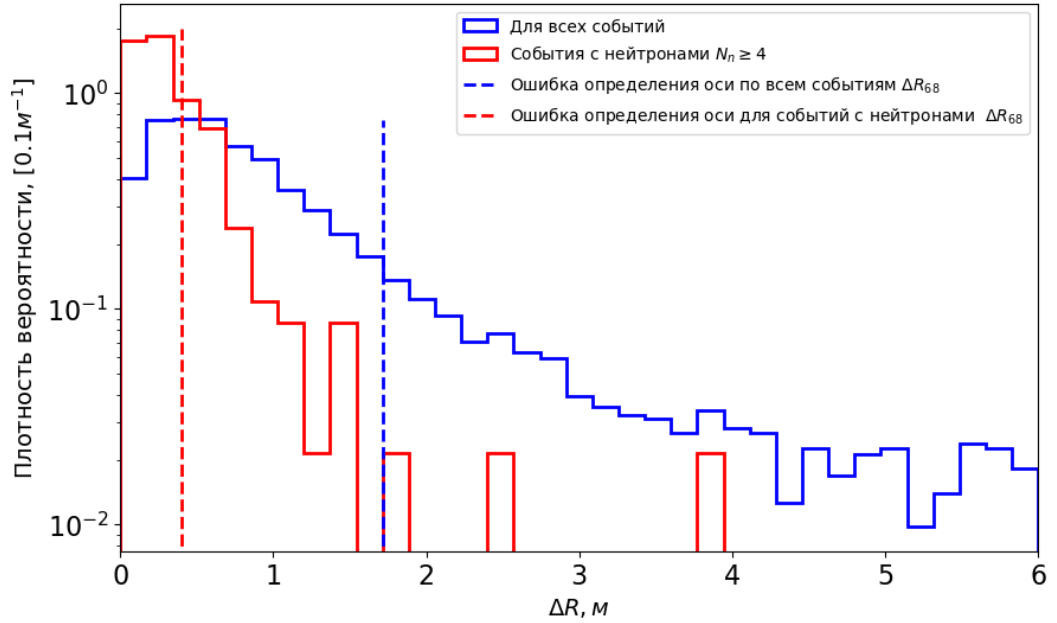


Рис. 4.8: Распределение погрешности определения оси ливня во всех реконструированных событиях (синяя кривая) и в событиях с $N_n \geq 4$ (красная кривая)

При использовании информации о нейтронной компоненте ошибка определения оси ливня уменьшается в ~ 4 раза и также приводит к минимизации больших ошибок, т.к. высокоэнергетические адроны ливня, которые рожают испарительные нейтроны, находятся вблизи оси ШАЛ.

При выборе области отбора восстановленных осей ливня для событий с $N_n \geq 4$, было проведено сравнение качества реконструкции для случая с границами, равными границам установки, и с границами с отступом в 1 м от крайних детекторов. На рисунке 4.9 показаны двумерные распределения положений осей ШАЛ, где левый рисунок соответствует реконструированным осям ливня, по середине их истинное положение, а справа их разница для варианта отбора осей ливня в границах установки.

Как видно из рис. 4.9, при использовании области, равной границам установки, возникает избыток событий в углах установки. Возникает это из-за большего телесного угла и эффекта “притяжения осей” к установке, т.е. события в

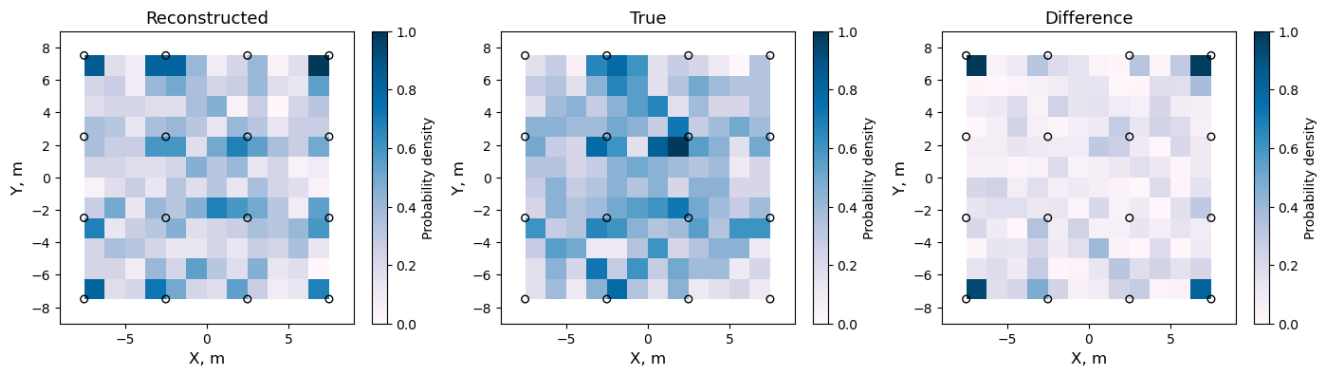


Рис. 4.9: Тепловая карта распределения осей ШАЛ с отбором в границах установки. Реконструированное положение (слева), истинное (по центру), разница (справа)

которых положение оси находится за границами установки, нередко восстанавливается в её границах.

При использовании же границ с отступом от границ на 1 метр реконструированные оси ливня относительно равномерно распределены по всей площади (рис. 4.10).

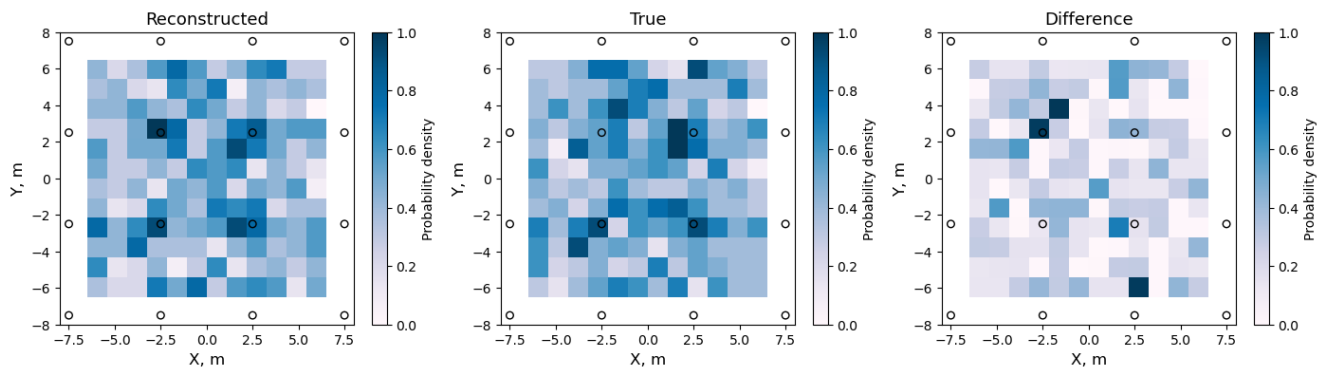


Рис. 4.10: Тепловая карта распределения осей ШАЛ с отбором с отступом в 1 метр от границ установки. Реконструированное положение (слева), истинное (по центру), разница (справа)

Затем также была проведена оценка качества реконструкции для отбора осей ливня в границах установки (I), для границ с отступом (II) и для отбора с отступом и требованием наибольшего энерговыделения в центральных четырех

детекторах (III). Результаты сравнения представлены в таблице 4.2.

Таблица 4.2: Сравнение результатов восстановления для различных критериев отбора событий

Критерий отбора	err_{N_e}	ΔR_{68} , м	Recall, %	Precision, %	ε_{rec}^{core} , %
(I)	0.07	0.46	89.8	90.4	90.0
(II)	0.07	0.40	95.8	96.0	95.9
(III)	0.07	0.40	82.1	89.0	82.9

Из результата сравнения видно, что наилучшее качество отбора осей показал критерий отбора с отступом от границ установки, в дальнейшем он использовался при обработке данных. Для этого критерия отбора 95% выборки лежит в пределах одного метра ошибки, а менее 1% событий с реконструированным положением оси ливня в выбранной области находилось за границами установки.

На рисунке 4.11 представлены корреляционные зависимости истинной мощности ливня от её реконструированного значения.

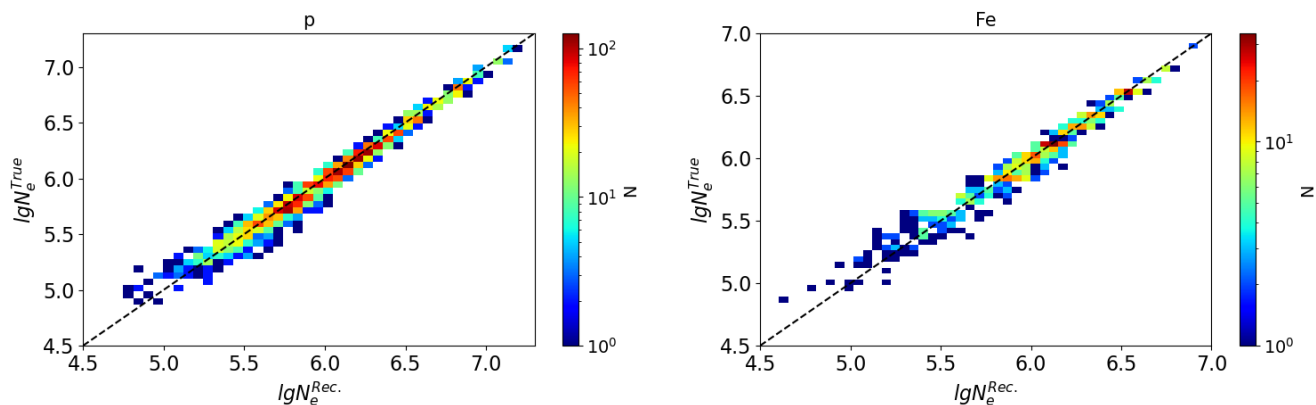


Рис. 4.11: Корреляционные зависимости истинной мощности ливня от реконструированного значения для протона (слева) и ядра железа (справа), штриховая линия соответствует прямой пропорциональности

Абсолютная погрешность восстановления мощности для протона соста-

вила 0.05, а для ядер железа 0.07, средние значения разности логарифмов составили $1.9 \cdot 10^{-2}$ и $4.0 \cdot 10^{-3}$ соответственно.

4.2.3. Оценка точности восстановленных параметров

Для получения информации о точности реконструированных параметров ШАЛ использовалась информация Фишера, которая основана на оценке второй частной производной функции правдоподобия в точке экстремума. Благодаря неравенству Рао-Крамера (4.10) нижняя граница для оценки дисперсии неизвестного параметра:

$$\begin{aligned} D(\hat{\theta}) &\geq |j(\theta)|^{-1} \\ j(\theta) &= -\mathbb{E}(H), \end{aligned} \tag{4.10}$$

где $j(\theta)$ - информационная матрица Фишера, H - матрица Гессе.

Таким образом, чем более выпукла функция правдоподобия в точке экстремума, тем меньшей оценки дисперсии восстановленного параметра она соответствует. События с большими ошибками реконструируемых параметров свидетельствуют о неадекватности априорных предположений о виде распределения, задаваемого в функции правдоподобия. Такие события могут быть связаны с большими флуктуациями при развитии ливня или энерговыделения в детекторе. Исходя из полученных оценок на погрешность восстановленных параметров, можно накладывать дополнительные критерии для отбора событий.

На рисунке 4.12 представлен пример события с большой ошибкой реконструкции положения оси ливня. Зеленой штриховой линией обозначено реконструированное положение оси ШАЛ, а синим крестом её истинное положение. Градиентом желтого цвета обозначено число частиц в каждом детекторе, а градиентом от синего до красного цвета – логарифм функции правдоподобия, которое было посчитано в каждой точке с шагом в 0.1 метр в границах от -30 до 30 метров по оси X и Y. Значение функции правдоподобия было рассчитано при фиксированных значениях мощности и возраста ливня.

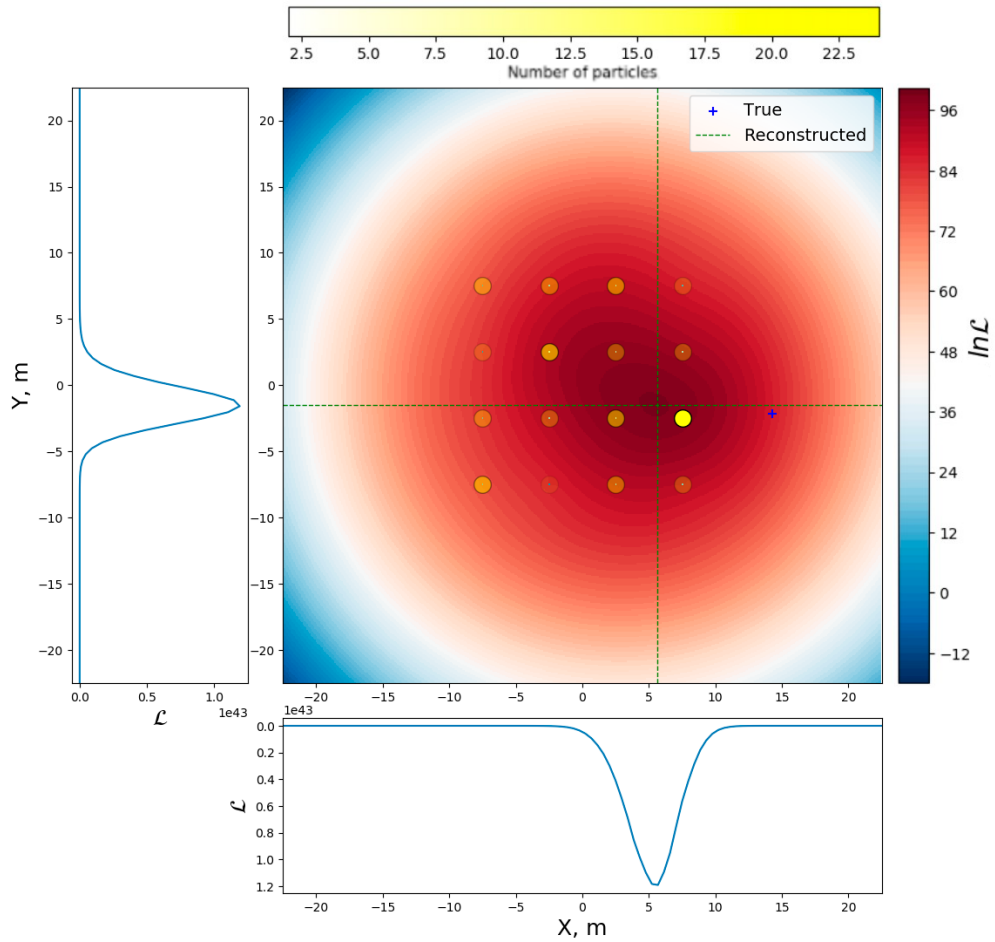


Рис. 4.12: Пример события с большой ошибкой реконструкции положения оси ШАЛ

Как видно из рисунка, в событии наблюдаются большие флуктуации энерговыведения (в некоторых детекторах на большем расстоянии от оси ливня энерговыведение выше, чем в близлежащих), что приводит к тому, что функция правдоподобия в точке экстремума становится менее острой. На рисунке 4.13 представлен пример события с малой ошибкой реконструкции.

Флуктуации энерговыведения в данном событии меньше, и детекторы с наибольшим энерговыведением лежат вблизи оси ливня, что приводит к более выпуклой функции правдоподобия в точке своего максимума и меньшей оценке погрешности восстановленных параметров ШАЛ.

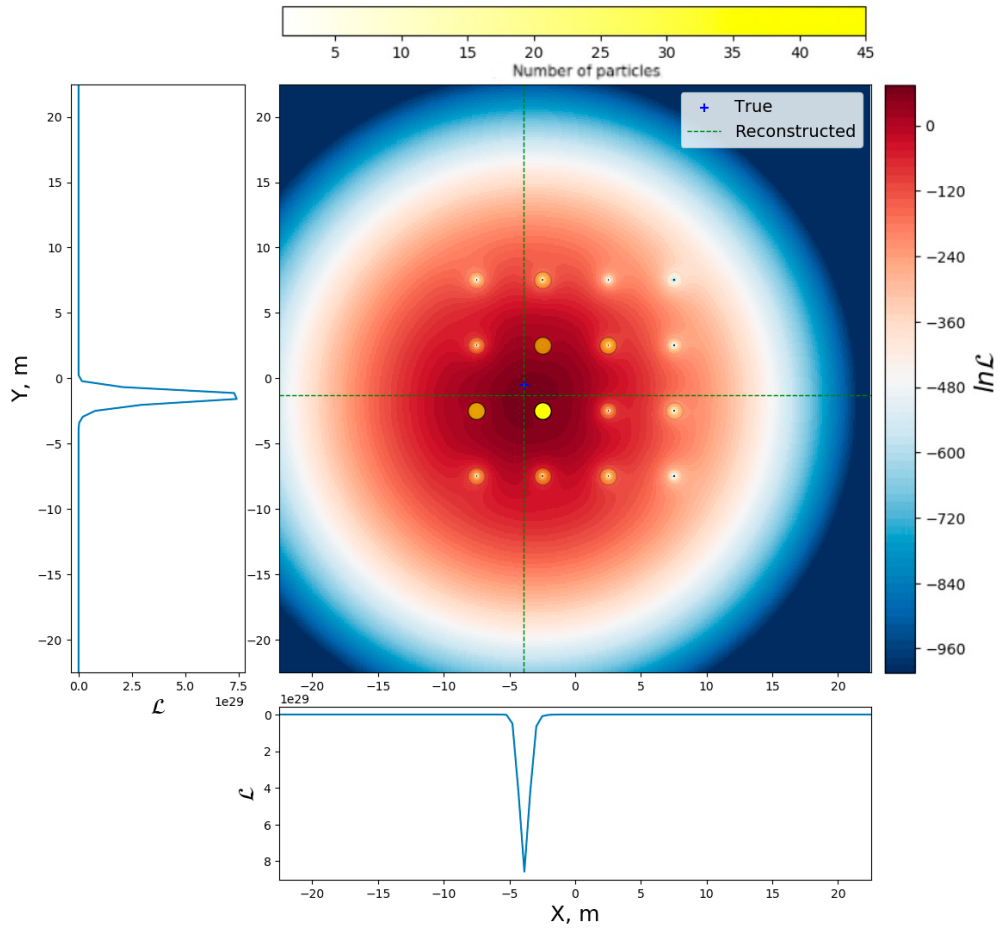


Рис. 4.13: Пример события с малой ошибкой реконструкции положения оси ШАЛ

4.3. Реконструкция с помощью методов машинного обучения

Для реконструкции параметров ШАЛ также можно применять различные методы машинного обучения, в этом случае модель учится выявлять общие закономерности по эмпирическим данным и затем используются для предсказания целевых переменных. Такой подход применяется во многих экспериментах [53, 54, 55]. Первая работа с использованием методов машинного обучения, в задаче реконструкции параметров ШАЛ, которую смог найти автор, датируется 1992 годом [56], и в ней для эксперимента KASCADE с помощью нейронной сети определялось положение оси ливня.

У данного подхода по сравнению с аналитическим есть преимущество в том, что модель в процессе обучения может построить решающую функцию,

которая лучше описывает данные. Возникает это из-за хороших аппроксимирующих свойств (в частности, для моделей глубокого обучения), а также из-за того, что модель может извлечь дополнительную информацию из обучающей выборки, которая не учтена в аналитическом подходе.

Модели обучаются с учителем на выборке искусственных размеченных данных с совокупностью прецедентов - пар “вход-целевая переменная”. Параметризация целевой переменной так же, как и архитектура модели, влияет на итоговое качество реконструкции. В ходе исследования были проведены сравнительные анализы различных параметризаций целевой переменной и алгоритмов машинного обучения с целью выявления оптимальной модели реконструкции каждого отдельного параметра ШАЛ.

4.3.1. Восстановление направления прихода ливня

Для восстановления направления прихода ШАЛ а качестве алгоритмов реконструкции были выбраны полносвязные нейронные сети (ANN), случайный лес (RF) [57] и градиентный бустинг над решающими деревьями (GBDT) [58].

Выбор полносвязных нейронных сетей обусловлен их универсальными аппроксимирующими свойствами [59], которые возрастают с увеличением числа слоев сети и числа нейронов в них. Для её создания использовался фреймворк Pytorch, модель состояла из 6 слоев, в качестве функции активации использовалась ReLu ($f(x) = \max(0, x)$), а в качестве функции потерь MSE (Minimum Square Error).

Алгоритм случайного леса основан на использовании ансамбля решающих деревьев, к одним из главных достоинств алгоритма относят его устойчивость к переобучению и широкую область применения. Для создания модели использовался фреймворк Scikit-learn [60].

Алгоритм градиентного бустинга над решающими деревьями считается одним из самых универсальных и сильных методов машинного обучения.

Для создания модели GBDT был использован фреймворк Catboost [61]. Модели GBDT и RF включали в себя ансамбли из 1000 деревьев, максимальная глубина которых была равна 5. Признаками служили времена срабатывания всех 16 эн-детекторов.

Лучшие результаты реконструкции были достигнуты при использовании параметризации целевых переменных с помощью отображения на единичную окружность: $(\theta, \varphi) \rightarrow \nu$, где $y \rightarrow (\cos(\nu), \sin(\nu))$. Переход к искомым углам (θ, φ) осуществляется через $\text{atan2}(\cos(\nu), \sin(\nu))$.

Для поиска лучшего алгоритма реконструкции было проведено сравнение распределений по углу отклонения ψ восстановленного вектора прихода от истинного для. За угловое разрешение установки была принята квантиль 68% распределения ψ (4.11).

$$\psi = \arccos \left(\frac{A_i B_i}{|\vec{A}| |\vec{B}|} \right), \quad (4.11)$$

где $\vec{A}$, $\vec{B}$ - реконструированный и истинный вектор прихода ШАЛ.

Лучшее угловое разрешение среди методов машинного обучения было получено в методе GBDT (3.28°) против (3.38°) для RF и (3.95°) для ANN (по данным протонной выборки). Для базисного МНК угловое разрешение составило (5.20°), а для алгоритма IRLS (3.04°). На рисунке 4.15 представлены распределения по ψ для МНК, IRLS и GBDT для выборки по p и Fe.

Для дальнейшей реконструкции параметров ШАЛ отбирались события с реконструированным зенитным углом $\theta^{Rec} \leq 30^\circ$. Также по аналогии с п. 4.2.1 была оценена эффективность реконструкции ε_{rec} . Значения метрик для выборки по протону и железу оказались очень близкими в пределах 1%, в таблице 4.3 приведены усредненные значения метрик и угловых разрешений для обеих выборок для алгоритмов МНК, IRLS и GBDT.

Лучшего углового разрешения среди всех алгоритмов реконструкции удалось добиться с помощью алгоритма IRLS, по сравнению с базисным МНК, удалось улучшить угловое разрешение в ~ 1.74 раза. Результаты GBDT по угло-

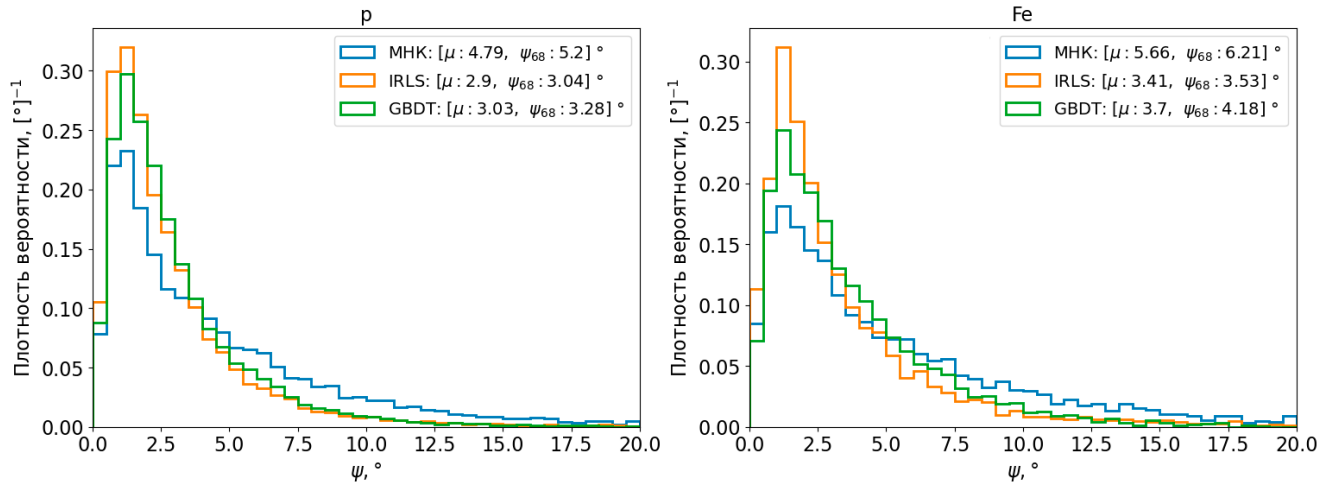


Рис. 4.14: Распределение по ψ для МНК (синяя кривая), IRLS (оранжевая кривая), GBDT (зеленая кривая) для протона (слева) и железа (справа)

Таблица 4.3: Сравнение метрик классификации и углового разрешения

Алгоритм	$\psi_{68}, ^\circ$	Recall, %	Precision, %	$\varepsilon_{rec}^{angle}, \%$
МНК	5.71	85.4	83.9	84.6
IRLS	3.28	89.9	91.1	90.4
GBDT	3.73	88.6	91.4	89.8

вому разрешению оказались немного хуже результатов IRLS, на менее чем 0.5° . С помощью метода IRLS удалось реконструировать около 90% всех событий. Итоговая оценка углового разрешения установки составляет ~ 3.3 градуса.

На рис. 4.16 представлена зависимость эффективности реконструкции направления прихода ШАЛ (слева) и углового разрешения (справа) от числа зарегистрированных нейтронов в событии.

Эффективность реконструкции, как и угловое разрешение, улучшается при увеличении числа нейтронов в событии, при $N_n \geq 4$ угловое разрешение составляет для протона составляет 1.23° , а для железа 1.05° . 95%-я эффективность достигается при $N_n \geq 2$, или при $lg(N_{ch}) \geq 2.6$.

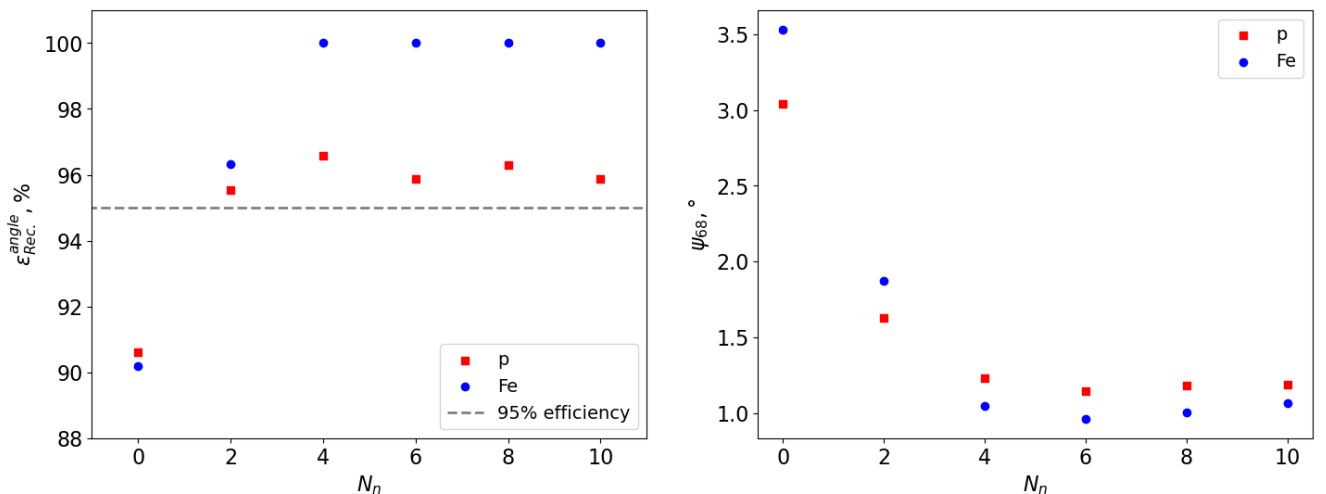


Рис. 4.15: Зависимость эффективности реконструкции направления прихода ШАЛ (слева) и углового разрешения (справа) от числа зарегистрированных нейтронов в событии

4.3.2. Восстановление положения оси ливня

При использовании методов машинного обучения при восстановлении положения оси ливня возможны два подхода: решение задачи регрессии через прямое предсказание координат оси (I) или через решение задачи классификации с предсказанием метки класса (“ось в границах установки”, “ось не в границах установки”) (II). При использовании второго подхода, возможно применить гибридный метод (III), восстанавливая положение оси с помощью ММП для событий, которые были отобраны моделью машинного обучения.

В качестве признаков использовались значения числа зарегистрированных заряженных частиц и нейтронов в каждом эн-детекторе и реконструированный зенитный угол. В таблице 4.4 приведены результаты сравнения качества реконструкции в событиях без дополнительного отбора по числу нейтронов.

Лучшие результаты восстановления показал гибридный метод, по сравнению с традиционным подходом (таблица 4.1) удалось улучшить эффективность реконструкции положения оси на 4.6%, а при использовании (II) подхода на 2.5%.

Таблица 4.4: Сравнение качества реконструкции положения оси ШАЛ для различных параметризаций целевой переменной при использовании машинного обучения.

Метод	err_{N_e}	ΔR_{68} , м	Recall, %	Precision, %	ε_{rec}^{core} , %
(I)	-	1.50	77.9	93.2	83.3
(II)	-	-	83.3	93.5	87.4
(III)	0.16	1.13	87.3	92.2	89.5

В событиях с $N_n \geq 4$ с использованием гибридного метода (I) абсолютная ошибка восстановления мощности составила 0.06, погрешность определения оси ливня 0.38 м, $\text{recall} = 95.3\%$, $\text{precision} = 95.9\%$, $\varepsilon_{rec}^{core} = 95.6\%$, что оказалось очень близким к результатам базисного подхода (см. таблицу 4.2).

На рисунке 4.17 представлена зависимость разрешения определения положения оси ливня (ΔR_{68}) от энергии первичной частицы.

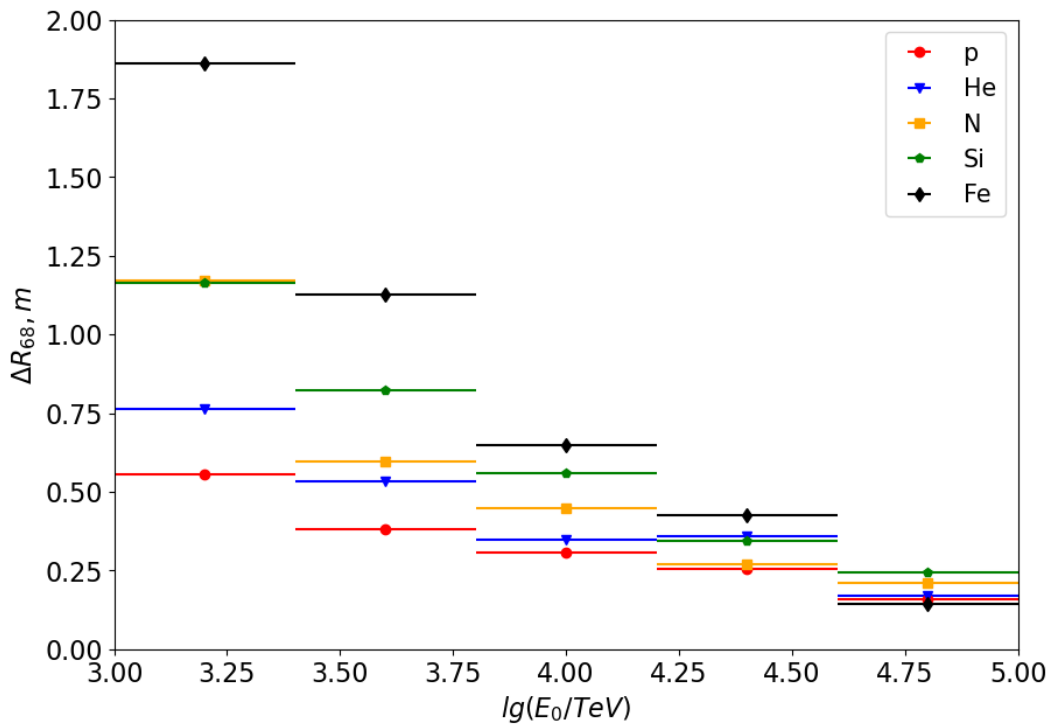


Рис. 4.16: Зависимость разрешения определения положения оси ливня от энергии первичной частицы

При энергии в 10 ПэВ, разрешение определения положения оси ШАЛ составляет $0.3 \div 0.7$ м.

На рис. 4.18 представлена зависимость эффективности реконструкции оси ливня для различных порогов по суммарному числу заряженных частиц и нейтронов в событии.

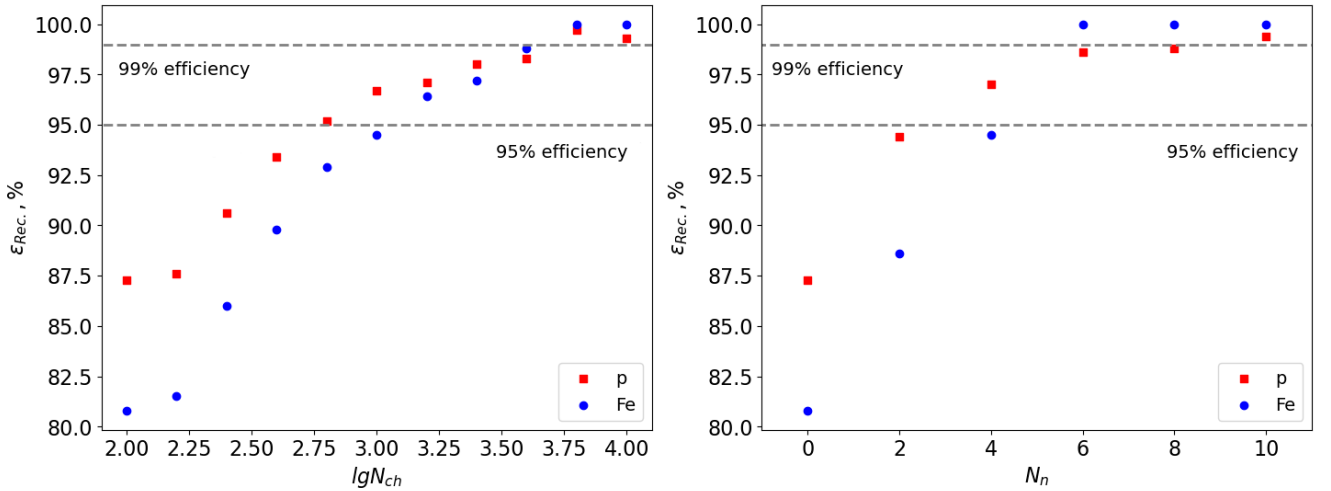


Рис. 4.17: Зависимость эффективности реконструкции положения оси ливня для различных порогов по числу заряженных частиц (слева) и числу нейтронов (справа)

При увеличении суммарного числа заряженных частиц и нейтронов в событии возрастает эффективность реконструкции положения оси ливня, 95%-я эффективность достигается при $lgN_{ch} > 3$ и $N_n > 4$.

4.3.3. Реконструкция мощности и полного числа адронов ливня

С применением методов машинного обучения можно обучить модель предсказывать полное число частиц ливня или полное число частиц определенной компоненты или же восстанавливать энергию первичной частицы без промежуточных стадий, используя в качестве признаков лишь показания детекторов (подробнее в п. 4.4).

На данном этапе была создана модель с помощью алгоритма GBDT для

предсказания мощности и полного числа адронов ШАЛ. В качестве признаков использовались показания эн-детекторов и реконструированные параметры, полученные в базисном подходе. На рисунке 4.18 представлено распределение относительной ошибки реконструкции мощности ливня для базисного подхода и гибридного с использованием машинного обучения.

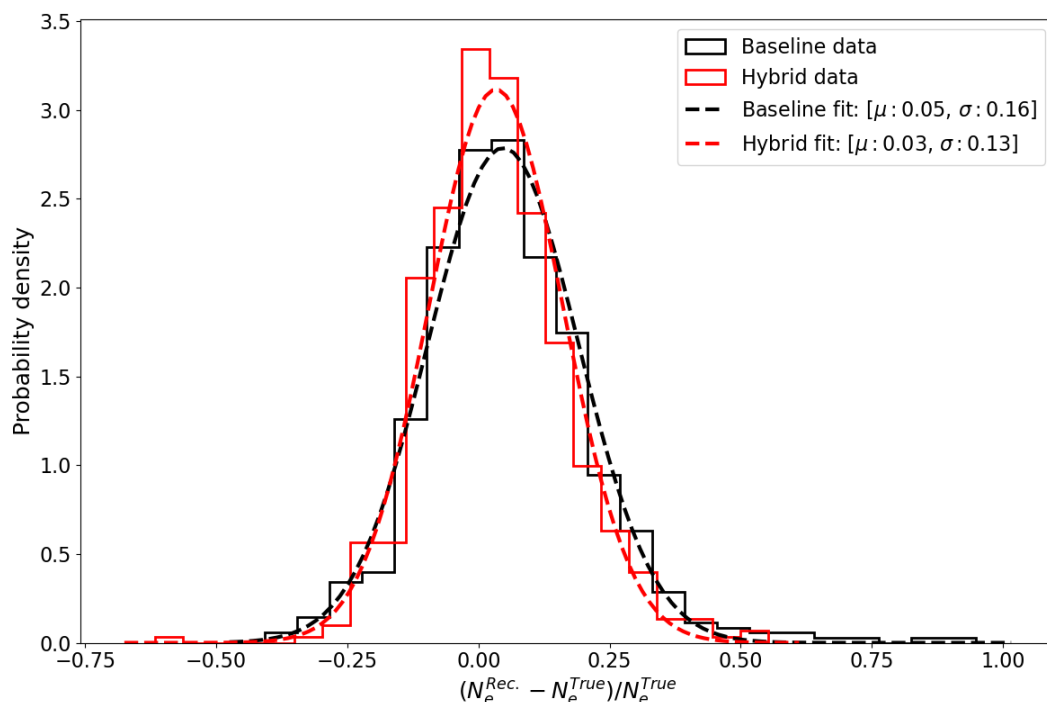


Рис. 4.18: Распределение относительной ошибки реконструкции мощности ливня

Распределения были профитированы распределением Гаусса, значение среднего оказалось ближе к нулю для гибридного метода, по сравнению с базисным подходом удалось улучшить качество реконструкции на $\sim 17.6\%$.

На рисунке 4.19 представлена зависимость относительной ошибки реконструкции мощности ливня от логарифма истинной мощности для базисного и гибридного подхода.

С увеличением истинной мощности ливня относительная ошибка реконструкции уменьшается с 16.5% до 8.3% для гибридного подхода и с 17.7% до 7.2% для базисного подхода.

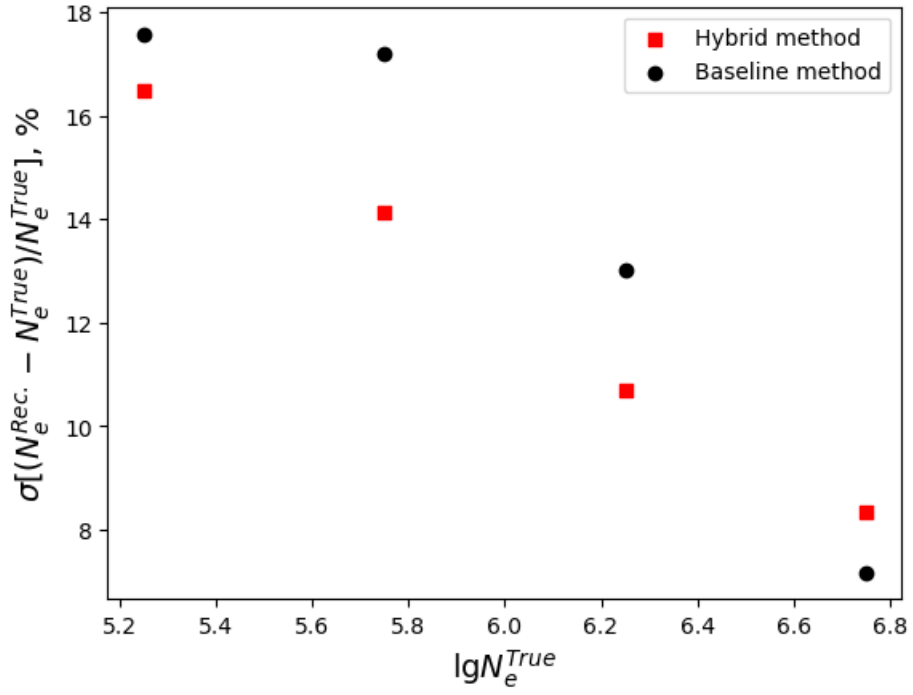


Рис. 4.19: Зависимость относительной ошибки реконструкции мощности ливня от логарифма истинной мощности

Перед реконструкцией полного числа адронов ливня, была проанализирована зависимость N_h от N_n . На (рис. 4.20) представлены результаты для первичных протонов, железа и смешанной выборки, которая состояла из всех разыгранных первичных частиц в равной пропорции.

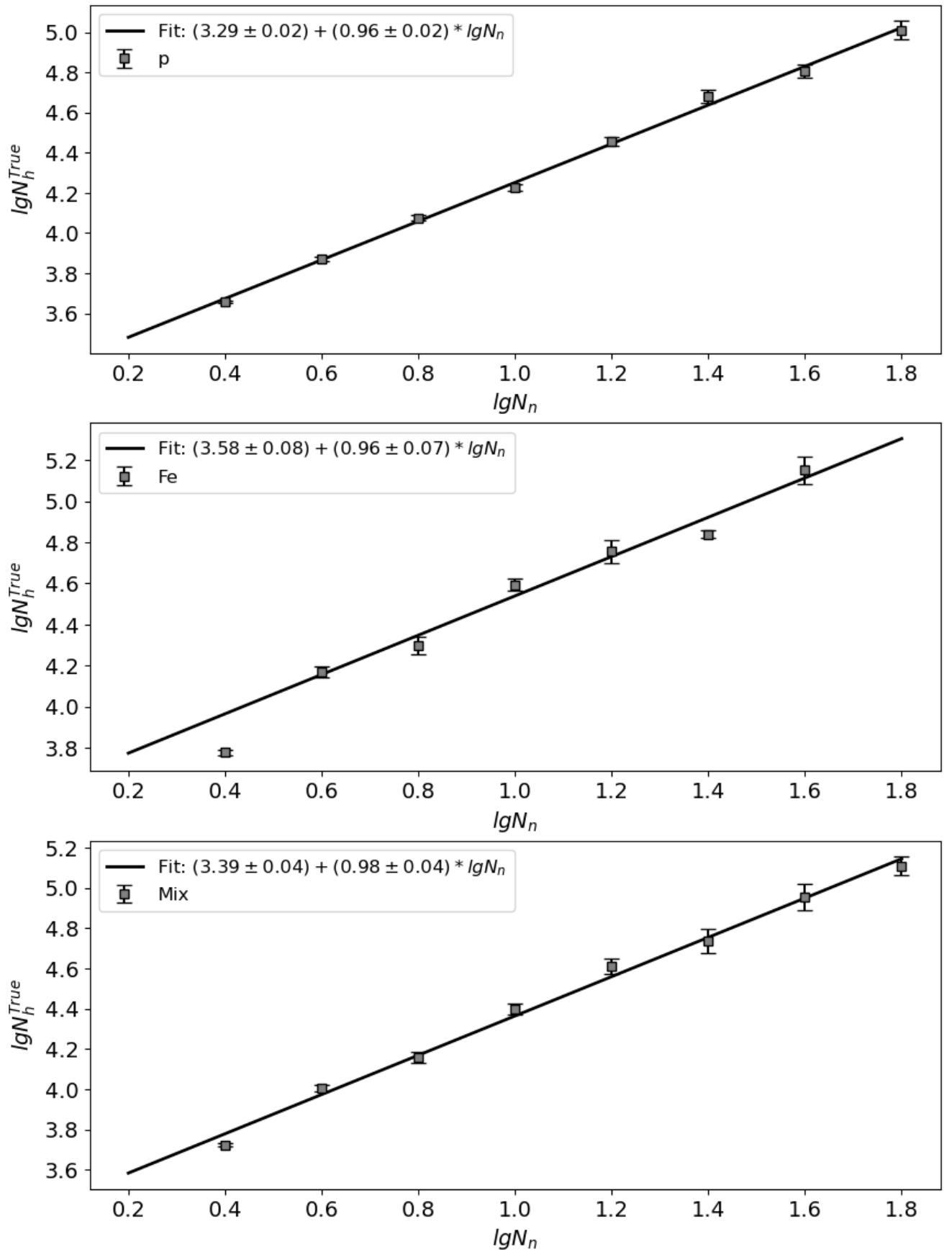


Рис. 4.20: Зависимость полного числа адронов от числа зарегистрированных нейтронов для ливней от протона (сверху), от железа (по центру) и для смешанного состава (снизу)

Полученная зависимость была профитирована линейной функцией в двойном логарифмическом масштабе (4.12):

$$\lg(N_h) = a \times \lg(N_n) + b. \quad (4.12)$$

Фит проводился в области $0.6 \leq \lg(N_n) \leq 1.6$, где результаты для протона: $\alpha_p = 0.96 \pm 0.02$, для железа $\alpha_{Fe} = 0.96 \pm 0.07$, а для смешанной выборки $\alpha_{mix} = 0.98 \pm 0.04$.

Подход с использованием формулы (4.12) был принят в качестве базисного метода реконструкции полного числа адронов в ШАЛ. Подход с использованием машинного обучения был основан на применении модели GBDT, где в качестве признаков модели использовались значения зарегистрированного числа нейтронов в каждом из детекторов и реконструированные параметры ШАЛ.

Лучшего качества восстановления удалось добиться с помощью гибридного метода, где абсолютная ошибка реконструкции ($\lg(N_h^{Rec.}/N_h^{True})$) составила 0.14, а для базисного метода 0.24. Относительная же ошибка реконструкции для гибридного метода составляет 42%.

На рисунке 4.21 представлено распределение абсолютной ошибки реконструкции полного числа адронов ливня (слева) и корреляция реконструированного и истинного значения N_h для гибридного метода (справа).

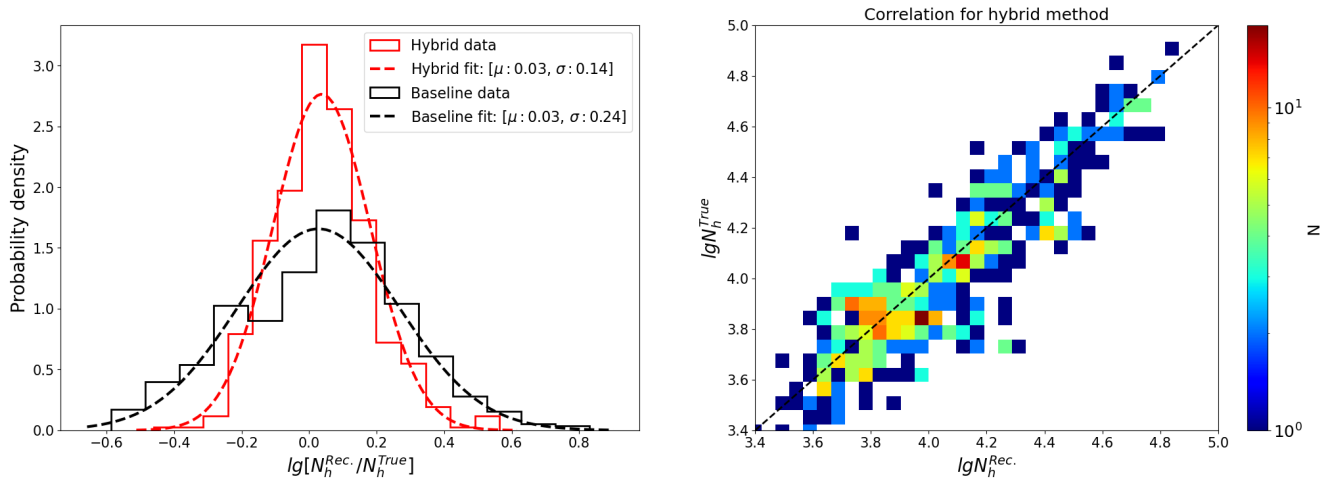


Рис. 4.21: Распределение абсолютной ошибки реконструкции полного числа адронов ливня (слева) для гибридного (красная кривая) и базисного подхода (черная кривая), и корреляция реконструированного и истинного значения N_h для гибридного метода (справа)

4.4. Восстановление энергии первичной частицы

Реконструкция энергии во многих экспериментах основана на использовании степенной зависимости первичной энергии E_0 от полного числа частиц определенной компоненты ливня $N_x(N_e, N_\mu, N_n)$ [62, 63, 64] с некоторыми коэффициентами a, b (4.13):

$$\lg(E_0) = a \times \lg(N_x) + b. \quad (4.13)$$

Гораздо более качественным параметром для перехода к энергии первичной частицы является оценка плотности частиц на определенном расстоянии от оси ливня [65, 66], так как этот параметр является менее флуктуирующим. Однако для установки ENDA-INR этот подход не применим из-за её небольших размеров.

Коэффициенты для (4.13) были получены с использованием выборки для модели QGSJET-II-04, эта же выборка использовалась для обучения моделей машинного обучения. Валидация качества реконструкции происходила на

выборках для модели ЕРОС-LHC, Sibyll-2.3d и n независимых событий модели QGSJET-II-04. Реконструкция производилась для событий с $N_n \geq 4$ и прошедших критерии отбора (таблица 4.5).

Таблица 4.5: Критерии отбора и процент отброшенных событий

Критерий	Доля событий [%]
$lgN_e^{Rec.} > 5$	99.3
$\theta^{Rec.} < 30^\circ$	87.7
$ X^{Rec.} \& Y^{Rec.} \leq 6.5$ м	32.4
$0.4 \leq s \leq 1.7$	32.2

Зависимость $lgE_0(lgN_n)$ представлена на (рис. 4.22), верхняя панель соответствует первичным протонам, по центру - железу, а нижняя - смешанной выборке по всем разыгранным первичным частицам в равной пропорции.

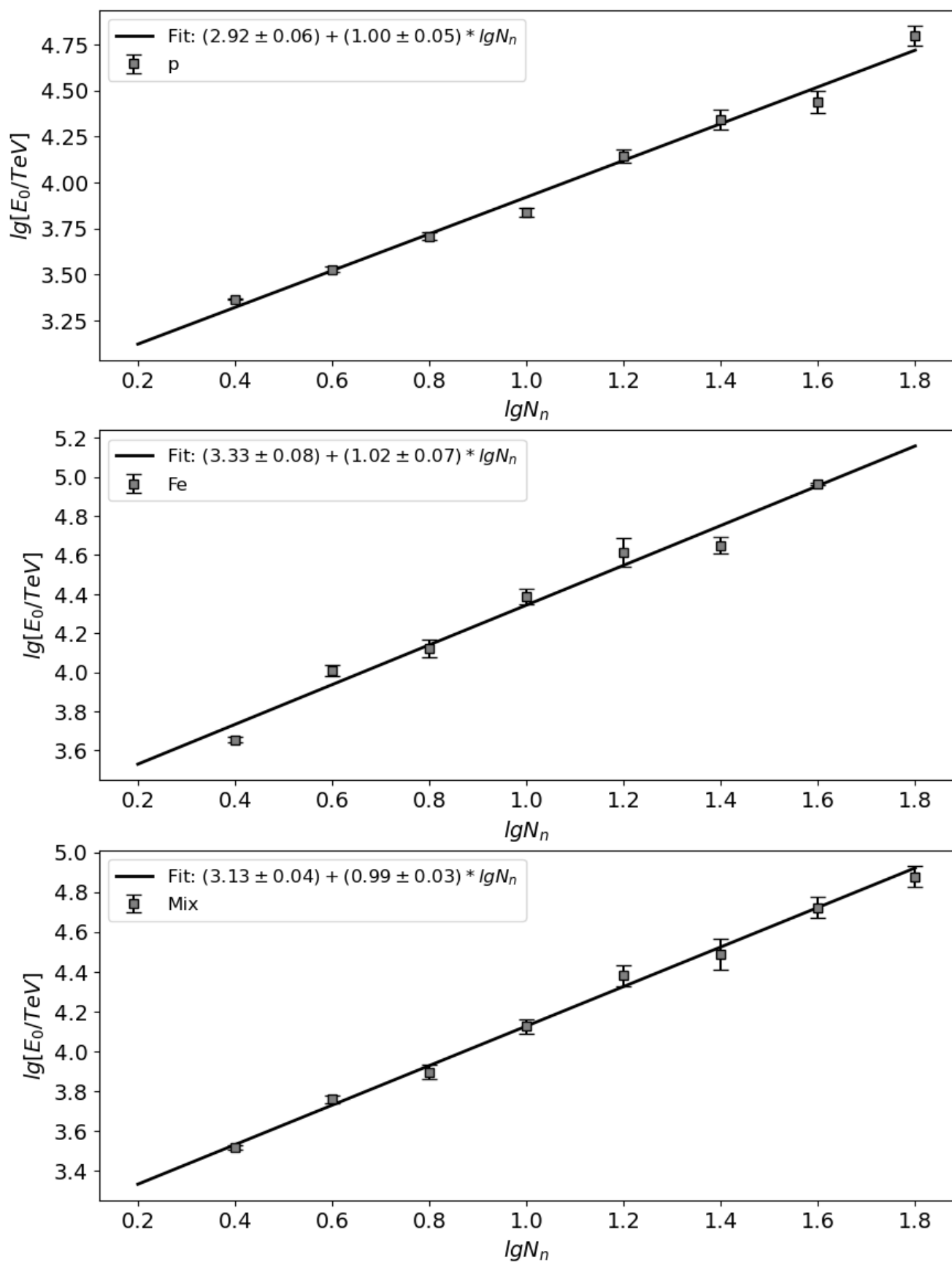


Рис. 4.22: Зависимость первичной энергии от числа зарегистрированных нейтронов в событии для первичных протонов (сверху), железа (центр), смешанный состав (снизу)

Фит проводился в области $0.6 \leq \lg N_n \leq 1.6$, где результаты для смешанной выборки (4.14):

$$a_{mix} = 0.99 \pm 0.03, \quad b_{mix} = 3.13 \pm 0.04. \quad (4.14)$$

С целью получения оценки зависимости (4.13) от химического состава были получены соответствующие коэффициенты для протона и железа (4.15):

$$\begin{aligned} a_p &= 1.00 \pm 0.05, \quad b_p = 2.92 \pm 0.06, \\ a_{Fe} &= 1.02 \pm 0.07, \quad b_{Fe} = 3.33 \pm 0.08. \end{aligned} \quad (4.15)$$

На рисунке 4.23 представлена зависимость $\lg E_0(\lg N_e^{Rec})$. Фит проводился в области $5.5 \leq \lg N_e^{Rec} \leq 6.75$, где результаты для смешанной выборки (4.16):

$$a_{mix} = 0.83 \pm 0.02, \quad b_{mix} = -1.08 \pm 0.14. \quad (4.16)$$

Результаты для протона и железа приведены в (4.17).

$$\begin{aligned} a_p &= 0.90 \pm 0.05, \quad b_p = -1.78 \pm 0.30, \\ a_{Fe} &= 0.83 \pm 0.02, \quad b_{Fe} = -0.95 \pm 0.18. \end{aligned} \quad (4.17)$$

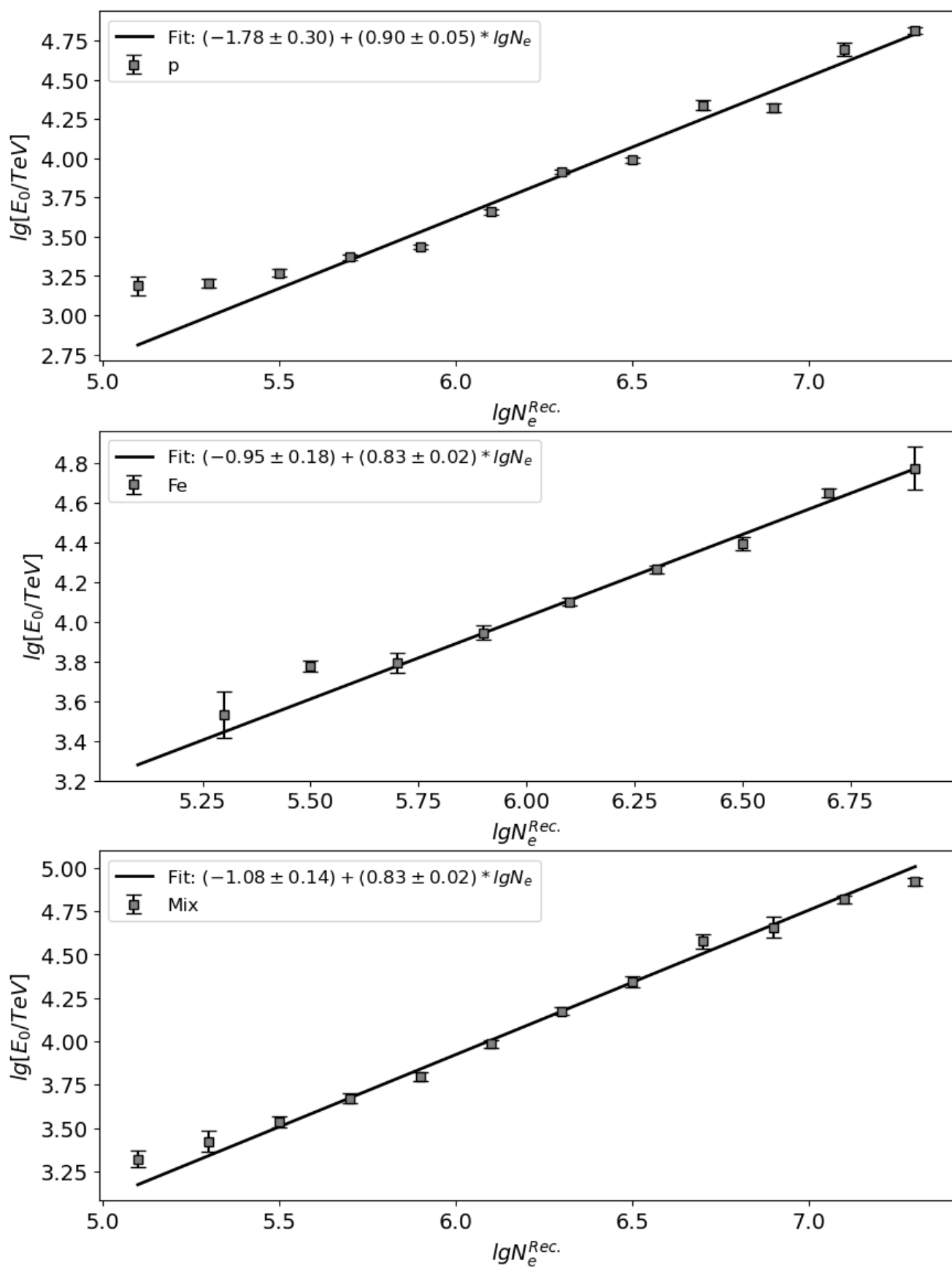


Рис. 4.23: Зависимость первичной энергии от реконструированной мощности ШАЛ для первичных протонов (сверху), железа (центр), смешанная выборка (снизу)

Как видно из (4.15) и (4.17), смещение между компонентами при использовании мощности ливня для реконструкции первичной энергии больше (на 47%), чем при использовании числа зарегистрированных нейтронов (10%).

Реконструкция первичной энергии происходила без знания о типе первичной частицы, в качестве коэффициентов (4.13) были взяты значения, полученные для смешанной выборки (4.14) и (4.16). Полученные распределения были профитированы распределением Гаусса, лучшее качество при использовании базисного подхода показала реконструкция с использованием мощности ливня $\sigma = 0.27$, при использовании числа нейтронов $\sigma = 0.35$.

В подходе с использованием машинного обучения можно создать модель, которая помимо N_e и N_n будет обрабатывать и другие признаки, чувствительные к энергии первичной частицы. В работе использовались следующие признаки: зенитный угол (θ), десятичный логарифм числа нейтронов $\lg(N_n)$, десятичный логарифм мощности $\lg(N_e)$, возраст (s) и положение оси (X, Y) ШАЛ.

В качестве алгоритмов были выбраны: линейная регрессия с полиномиальными признаками (LRPF), случайный лес (RF) и градиентный бустинг над решающими деревьями (GBDT). Лучшее качество реконструкции было получено в методе GBDT $\sigma = 0.24$, для случайного леса и линейной регрессии с полиномиальными признаками $\sigma = 0.25$ (рис. 4.24).

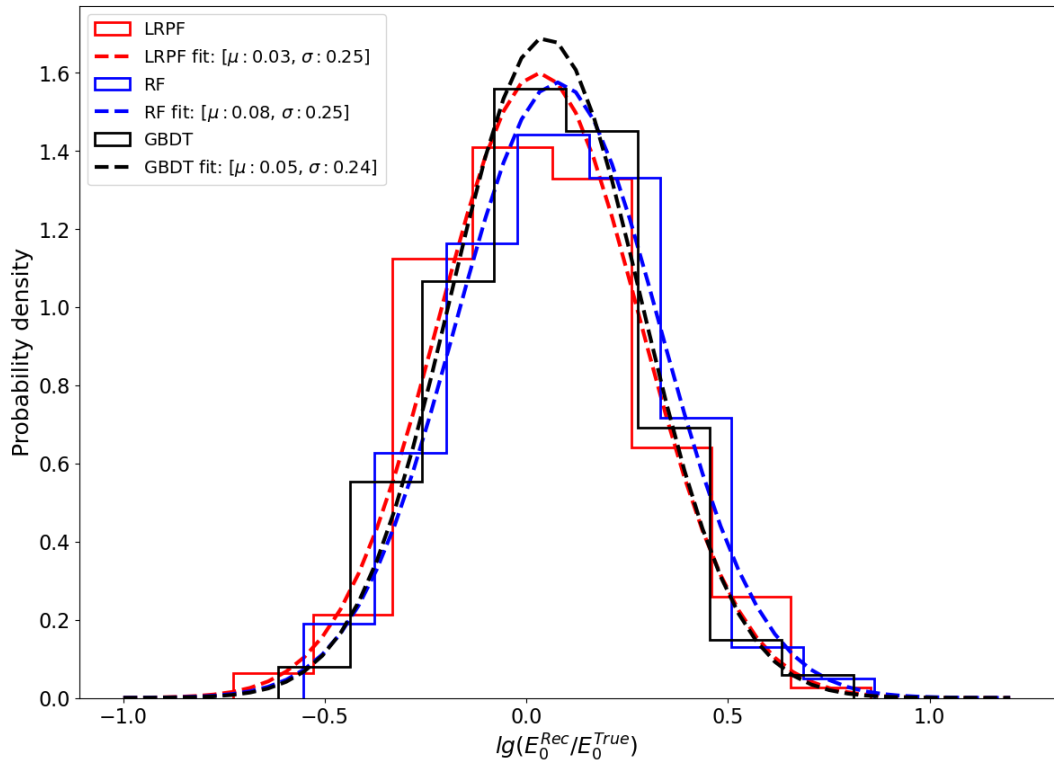


Рис. 4.24: Распределения отношения реконструированной и истинной энергии первичной частицы для LRPF, RF и GBDT

На рисунке 4.25 приведены сравнения качества реконструкции энергии первичной частицы для метода GBDT с базисным подходом с использованием мощности ливня и числа зарегистрированных нейтронов для модели QGSJET-II-04.

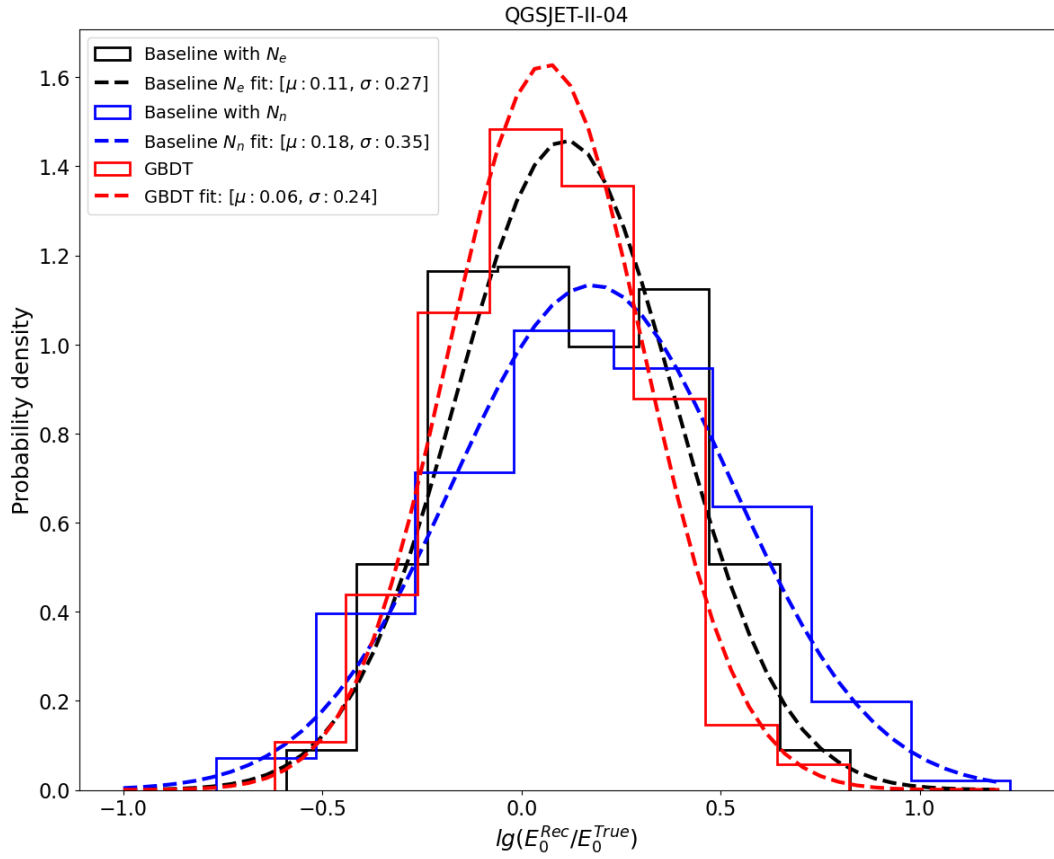


Рис. 4.25: Распределения отношения реконструированной и истинной энергии первичной частицы для алгоритма GBDT и базисного подхода для модели QGSJET-II-04

На рисунке 4.26 приведены сравнения качества реконструкции первичной энергии для модели EPOS-LHC, а на рисунке 4.27 для модели Sibyll-2.3d.

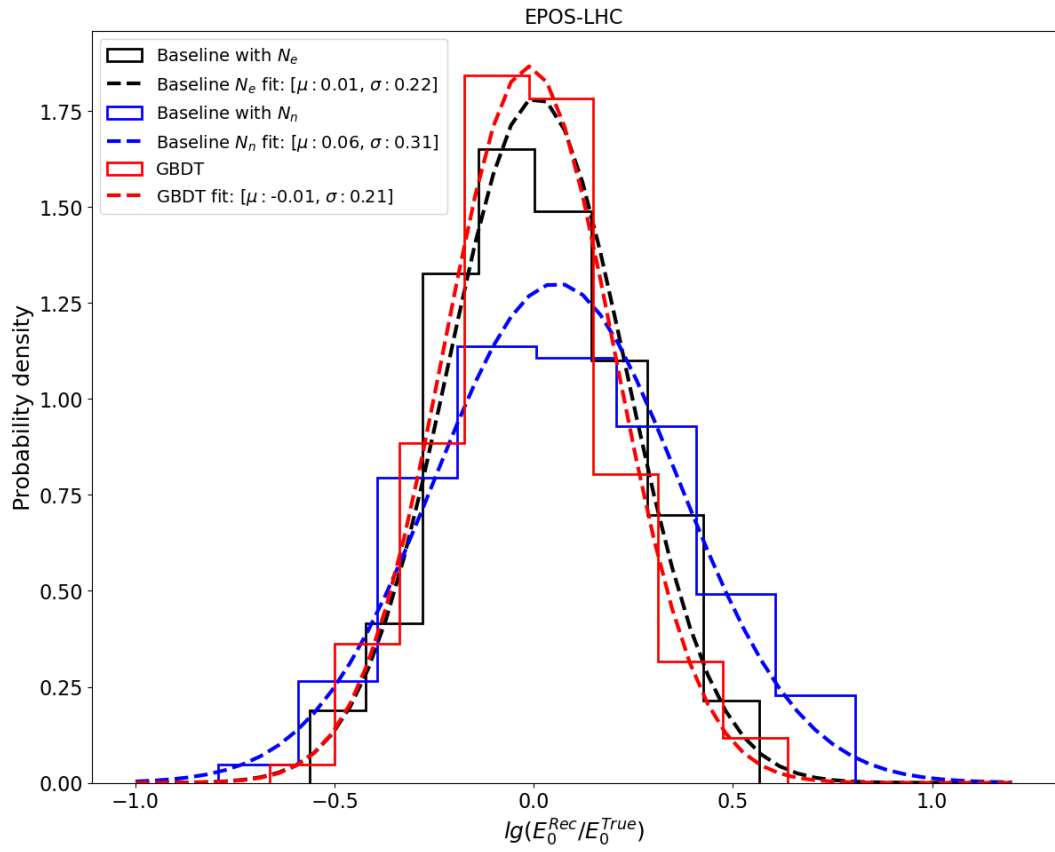


Рис. 4.26: Распределения отношения реконструированной и истинной энергии первичной частицы для алгоритма GBDT и базисного подхода для модели EPOS-LHC

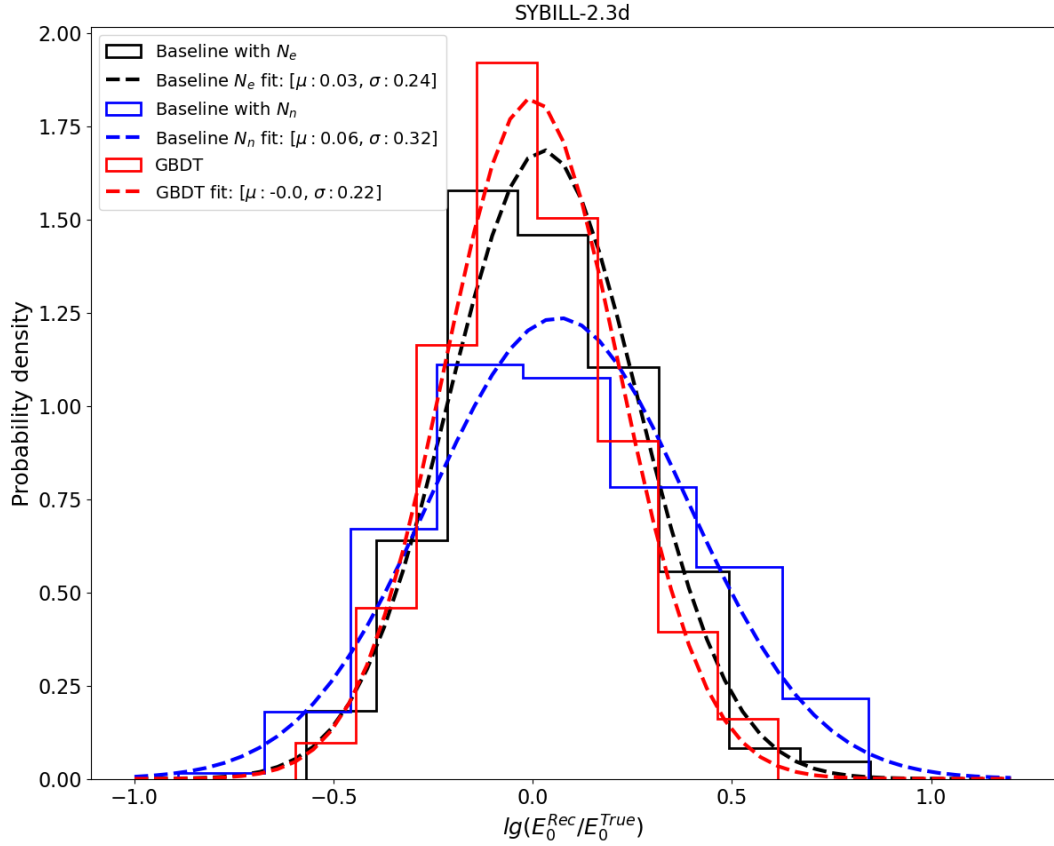


Рис. 4.27: Распределения отношения реконструированной и истинной энергии первичной частицы для алгоритма GBDT и базисного подхода для модели Sibyll-2.3d

Лучшие результаты были получены с помощью алгоритма GBDT, по сравнению с базисным методом реконструкции с помощью мощности ливня качество реконструкции удалось улучшить на 10.5% для модели QGSJET-II-04, на 4.3% для модели EPOS-LHC и на 7.6% для модели Sibyll-2.3d.

Смещение среднего значения μ относительно нуля зависит от химического состава выборки (см. п. 3.3.2), поэтому необходимо было провести оценку ошибки реконструкции для каждого отдельного типа первичной частицы.

Оценкой энергетического разрешения служило среднеквадратичное отклонение параметра ξ (4.18):

$$\xi = \frac{\lg E_0^{Rec.} - \lg E_0^{True}}{\lg E_0^{True}} \quad (4.18)$$

В таблице 4.6 представлены результаты среднего и среднеквадратичного

отклонения параметра ξ для трёх валидационных выборок для реконструкции с использованием метода GBDT.

Таблица 4.6: Результаты оценки точности реконструкции первичной энергии при восстановлении с помощью метода GBDT

Первичная частица	QGSJET-II-04		EPOS-LHC		Sibyll-2.3d	
	$\mu[\xi], \%$	$\sigma[\xi], \%$	$\mu[\xi], \%$	$\sigma[\xi], \%$	$\mu[\xi], \%$	$\sigma[\xi], \%$
p	4.17	6.64	2.77	5.46	3.53	5.87
He	-0.73	5.66	0.62	5.62	0.54	5.76
N	-2.61	4.52	-2.44	4.63	-2.13	3.91
Si	-3.67	4.49	-3.56	4.11	-3.00	3.91
Fe	-4.33	3.30	-4.51	3.61	-4.78	4.17

В таблице 4.7 представлены результаты среднего и среднеквадратичного отклонения параметра ξ для трёх валидационных выборок для реконструкции с использованием мощности ливня.

Таблица 4.7: Результаты оценки точности реконструкции первичной энергии при восстановлении с помощью мощности ливня

Первичная частица	QGSJET-II-04		EPOS-LHC		Sibyll-2.3d	
	$\mu[\xi], \%$	$\sigma[\xi], \%$	$\mu[\xi], \%$	$\sigma[\xi], \%$	$\mu[\xi], \%$	$\sigma[\xi], \%$
p	6.45	7.17	4.47	5.16	5.45	6.31
He	-0.13	4.25	1.66	5.45	1.50	5.18
N	-2.49	2.74	-2.54	4.13	-1.98	3.43
Si	-4.93	3.71	-4.08	3.26	-2.93	3.92
Fe	-5.63	3.14	-6.12	2.91	-5.34	3.63

В таблице 4.8 представлены результаты среднего и среднеквадратичного отклонения параметра ξ для трёх валидационных выборок для реконструкции

с использованием суммарного числа нейтронов в событии.

Таблица 4.8: Результаты оценки точности реконструкции первичной энергии при восстановлении по суммарному числу нейтронов

Первичная частица	QGSJET-II-04		EPOS-LHC		Sibyll-2.3d	
	$\mu[\xi]$, %	$\sigma[\xi]$, %	$\mu[\xi]$, %	$\sigma[\xi]$, %	$\mu[\xi]$, %	$\sigma[\xi]$, %
p	8.95	9.45	6.02	8.31	6.58	9.23
He	2.22	8.50	4.23	7.73	3.41	7.98
N	-1.24	6.92	-0.08	6.99	-0.60	6.56
Si	-4.02	5.14	-3.78	4.96	-1.74	7.03
Fe	-4.07	6.62	-6.20	4.30	-4.32	5.90

При использовании алгоритма GBDT значение среднего отклонения ξ не превышает 4.5%, а энергетическое разрешение менее 7%.

4.5. Оценка неопределенности реконструкции

Модели машинного обучения в задачах регрессии обычно дают только точечные предсказания. В связи с чем, они представляют из себя “черный ящик”, в котором решающая функция задана неявно. Из-за этого у модели, особенно это свойственно глубоким нейросетям может возникнуть ситуация чрезмерной уверенности, когда, например с большой вероятностью предсказывается определенная метка класса на данных другого класса, которого могло и не быть в обучающей выборке [67]. Чтобы получить больше информации и оценить степень уверенности о предсказании существуют различные вероятностные и байесовские методы.

Неопределенности в машинном обучении могут быть поделены на два класса, неопределенность в данных (aleatoric uncertainty) и неопределенность в знаниях (epistemic uncertainty). Неопределенность в данных возникает из-за

присущего данным перекрытия классов (рисунок 4.28 (слева)) или шума и не может быть уменьшена с увеличением обучающей выборки. Неопределенность в знаниях возникает из-за присущего модели недостатка знаний о входных данных из регионов, либо удаленных от обучающих данных (рисунок 4.28 (справа)), либо слабо охваченных ими (рисунок 4.28 (в центре)) [68, 69]. Неопределенность в знаниях же может быть уменьшена с увеличением размера обучающей выборки.

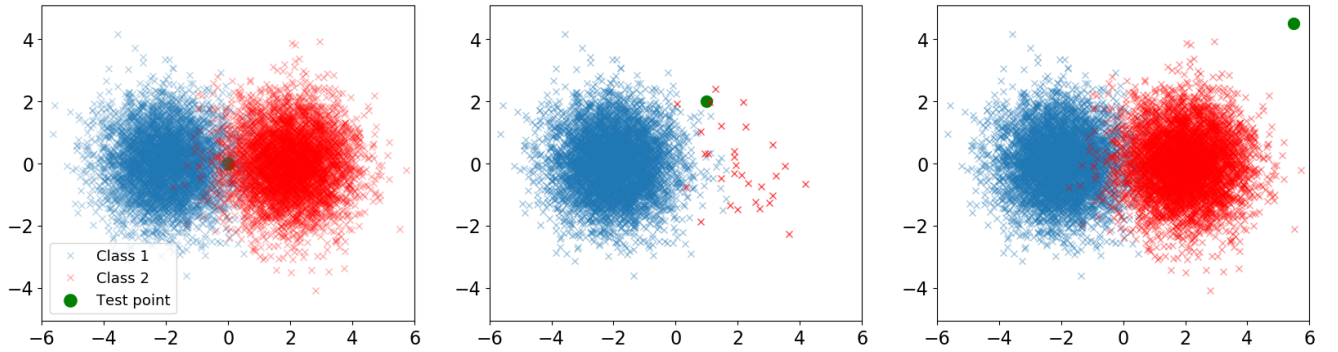


Рис. 4.28: Неопределенность в данных (слева), неопределенность в знаниях (центр, справа). Зеленая точка – объект из выборки, красные точки – обучающая выборка

Неопределенность в данных можно оценивать с помощью вероятностной регрессии. Из экспериментов по космическим лучам данный метод применяется в обсерватории IceCube [70] для оценки неопределенности восстановленной энергии и направления прихода нейтрино.

Основан метод на определении функции потерь в виде Гауссовской функции правдоподобия (4.19):

$$\mathcal{L} = \ln(\sqrt{2\pi\sigma^2}) + \frac{(y - \mu)^2}{\sigma^2}, \quad (4.19)$$

где μ, σ - реконструируемые параметры, y - целевая переменная.

Результаты по оценке неопределенности в данных для энергии первичной частицы, направления прихода и положения оси ливня для установки ENDA-INR представлены на рисунке 4.29.

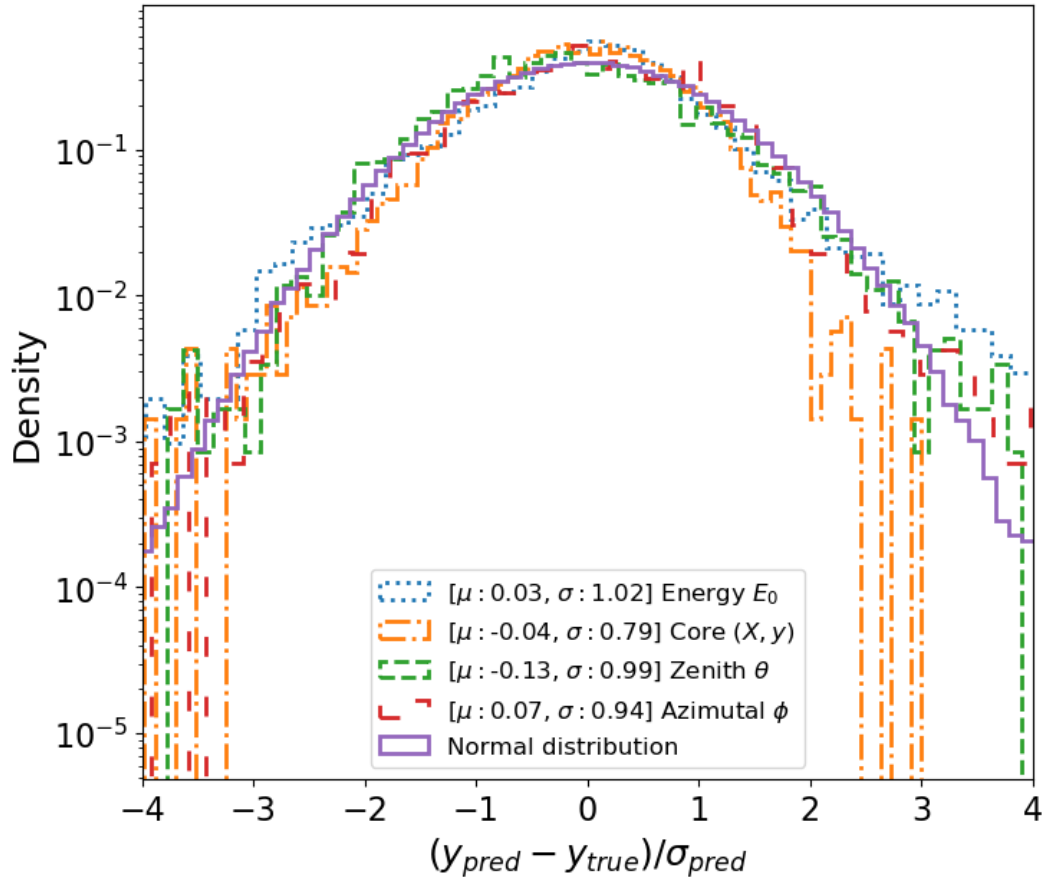


Рис. 4.29: Распределение разностей между предсказанным и истинным значением целевой переменной, деленное на предсказанное среднеквадратичное отклонение в сравнении с нормальным распределением, что указывает на хорошую оценку неопределенности

В работе для оценки неопределенности в данных и в знаниях использовался метод ансамблей [71, 72]. Подход основан на создании независимого ансамбля моделей, где каждая модель предсказывает свое среднее значение и дисперсию. Неопределенность в прогнозах, обусловленная неопределенностью знаний, выражается как уровень разброса, или “несогласованности” моделей в ансамбле.

Глава 5

Обработка экспериментальных данных

5.1. Анализ работы установки

При обработке экспериментальных данных была взята информация за промежуток времени с 01.09.23 по 05.02.25, в течение которого конфигурация установки была неизменна. На рисунке 5.1 представлен итоговый рабочий цикл установки, который был получен с помощью нулевого мастера.

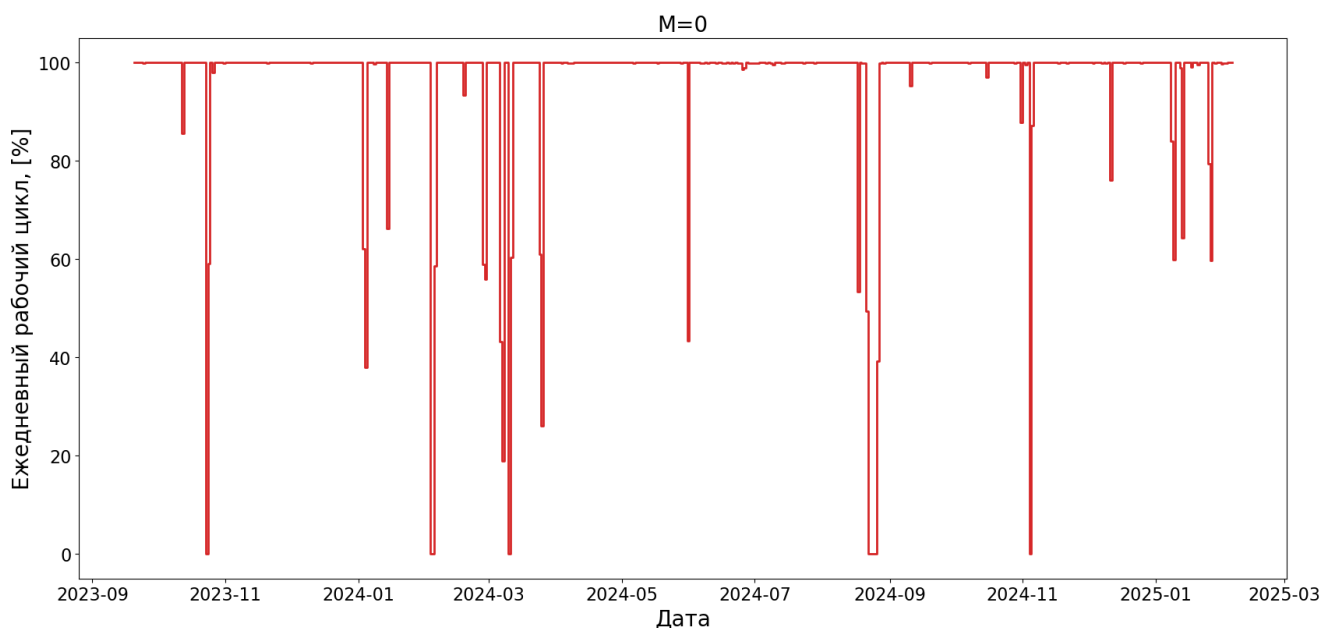


Рис. 5.1: Ежедневный рабочий цикл установки

Провалы на графике связаны с техническими неполадками (отключение электроэнергии или зависание программы). Также программа останавливалась, когда на установке проводились ремонтные работы. Итоговый рабочий цикл установки составил 96.1%.

Для контроля работоспособности установки оценивается количество триггеров по первому и третьему мастеру, среднее энерговыделение и среднее число зарегистрированных нейтронов в детекторе за промежуток времени, равный 6 часам. В результате анализа этих параметров, события в которых

наблюдались отклонения от стабильного режима работы установки, были удалены из дальнейшей обработки. На рисунке 5.2 представлены темпы счета для $M = 3$ и $M = 7$.

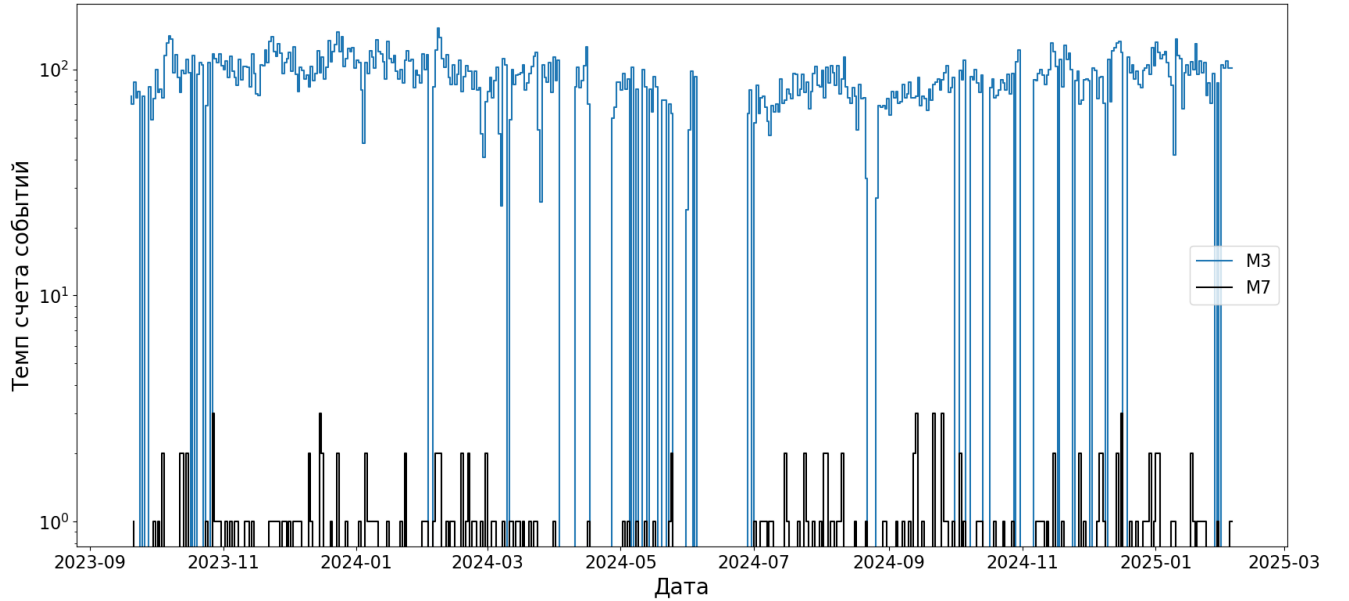


Рис. 5.2: Темп счета триггеров второго уровня

5.2. Амплитудная калибровка эн-детекторов

Для объединения показаний динодов необходимо было определить коэффициенты сшивки. Для этого была построена диаграмма рассеяния амплитуд по всем отобранным событиям, где по оси абсцисс отложена амплитуда по пятому, а по оси ординат по восьмому диноду, затем была выбрана область до насыщения восьмого динода и проведена линейная аппроксимация, при этом за начало прямой было принято начало координат. На рисунке 5.3 представлены итоговые диаграммы рассеяния для первых трёх эн-детекторов и приведены коэффициенты сшивки между динодами.

Итоговая амплитуда отдельного эн-детектора (A) определялась согласно следующим условиям (5.1):

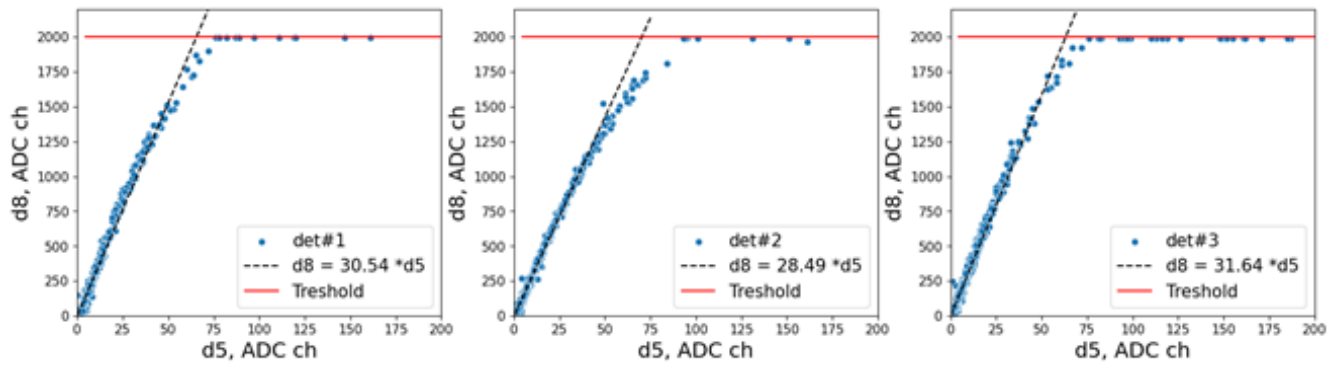


Рис. 5.3: Диаграммы рассеяния (черным пунктиром обозначена линейная аппроксимация), красной линией обозначено максимальное значение для 8 динода

$$A = \begin{cases} K \cdot A_{d5} & \text{если } A_{d8} > 650, A_{d5} > 10 \text{ и } A_{d5} > A_{d8} \\ A_{d8} & \text{иначе} \end{cases} \quad (5.1)$$

где K - коэффициент сшивки, A_{d8} и A_{d5} - значения амплитуд по 8 и 5 диноду в каналах АЦП.

Амплитудная калибровка детекторов проводилась с использованием режима единичных срабатываний, в котором записывается энерговыделение каждого отдельно сработавшего эн-детектора. Исходя из полученных амплитудных распределений подбирался коэффициент усиления отдельных детекторов для приведения их к усредненному уровню по всей установке. Подобную калибровку можно проводить и по событиям с первым и третьим мастером в режиме совпадений, ввиду осевой симметрии установки и предположения об изотропности азимутального угла прихода ШАЛ. На рисунке 5.4 представлены итоговые распределения энерговыделений для всех эн-детекторов.

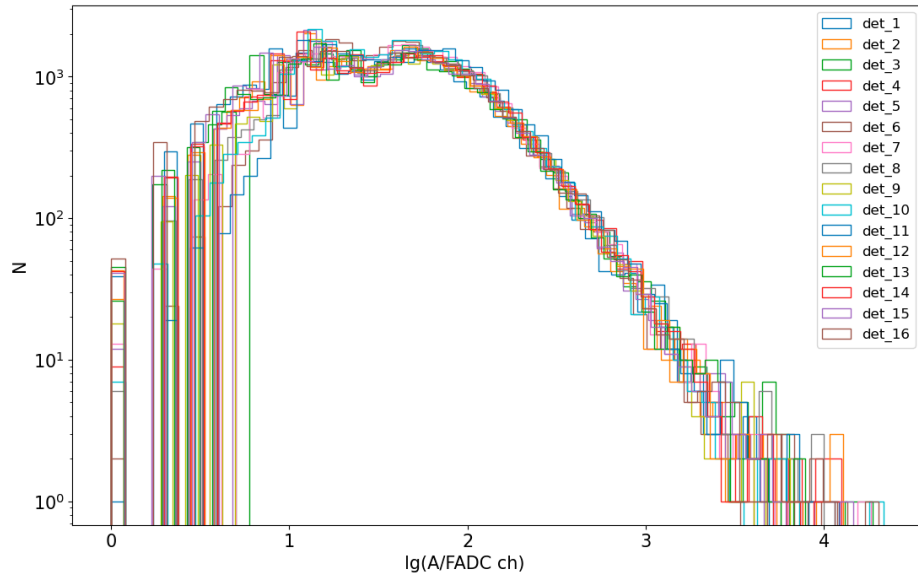


Рис. 5.4: Распределение по энерговыделению каждого отдельного детектора

5.3. Временное разрешение и временная калибровка установки

В качестве меры временного разрешения установки использовалась величина среднеквадратичного отклонения распределения “разности разностей”, $\Delta T = (t_1 - t_2) - (t_3 - t_4)$, что представляет из себя линейную комбинацию относительных времен срабатывания 4-х детекторов, находящихся в углах установки. Временем срабатывания эн-детектора считается момент пересечения сигналом уровня, равного половине его амплитуды.

Из-за осевой симметрии установки и изотропного направления прихода ШАЛ по азимутальному углу распределение по ΔT представляет из себя симметричное распределение, не зависящее от угла прихода и может быть использовано для оценки временного разрешения детекторов [73]. На рисунке 5.5 представлены распределения по ΔT , полученное из результатов моделирования (для протона и ядер железа) и эксперимента.

Полученные распределения фитировались нормальным распределением, после чего были оценены значения среднего (μ) и среднеквадратичного отклонения (σ). По данным моделирования $\sigma_p = 29.3$ нс для протона и $\sigma_{Fe} = 28.1$

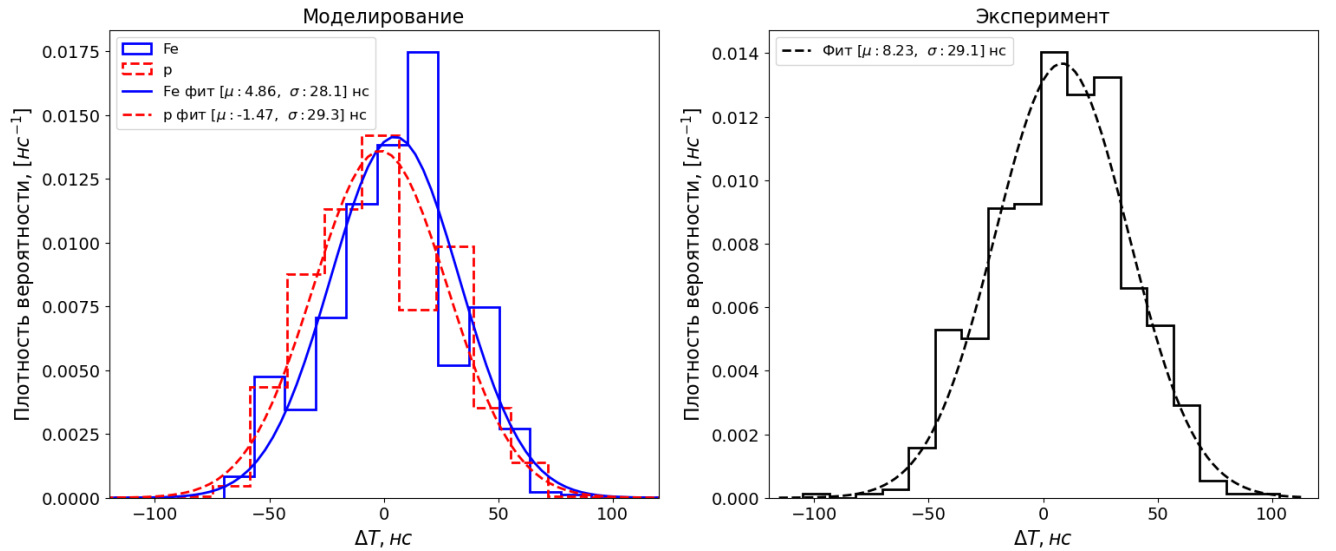


Рис. 5.5: Распределение по ΔT в моделировании (слева) и эксперименте (справа)

нс для ядер железа, по данным эксперимента $\sigma_{exp} = 29.1$ нс. Исходя из того, что измеряемая величина ΔT складывается из четырех независимых случайных величин, то оценка временного разрешения одного детектора составляет ~ 14.4 нс.

Из-за того, что временное разрешение тесно связано с угловым, а экспериментальные и расчетные оценки временного разрешения оказались очень близкими, то можно сделать вывод о близости реального значения углового разрешения установки к оценке, полученной в п. 4.3.1.

Значительное отклонение среднего значения от нуля для экспериментального распределения, объясняется наличием неучтенных задержек, связанных с задержками срабатывания ФЭУ, временем распространения сигнала по кабелям и временем обработки сигналов измерительными каналами регистрирующей электроники. Учет данных задержек необходим для корректной реконструкции направления прихода ливня.

Для их определения были вычислены и построены распределения по разности времен срабатывания между каждой парой детекторов. Исходя из уже упомянутого предположения об изотропии азимутального угла прихода ШАЛ, полученные распределения должны быть симметричны. Смещения их средних

значений относительно нуля имеют систематический характер, определяя искомые временные задержки между детекторами.

На рисунке 5.6 приведены распределения разности времен срабатывания первого эн-детектора относительно второго и третьего.

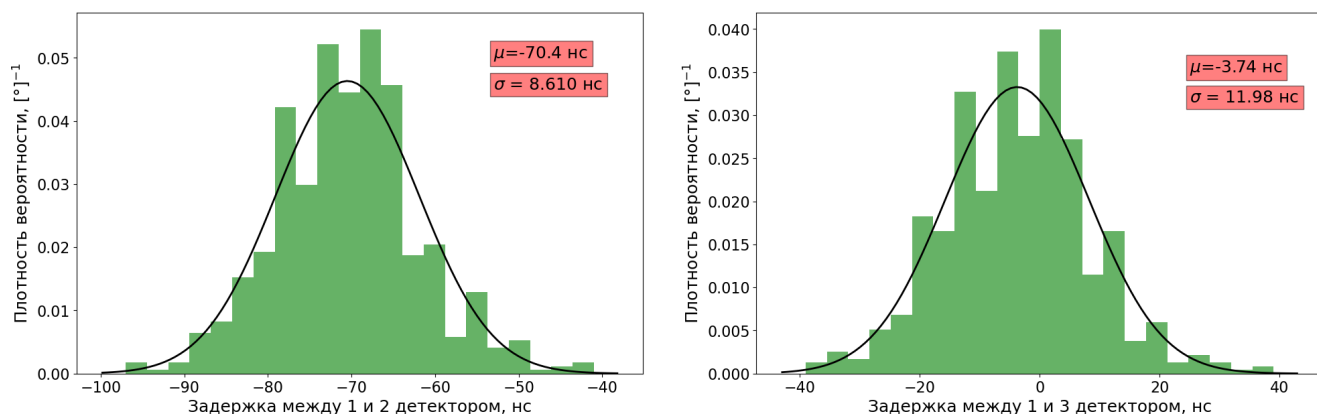


Рис. 5.6: Распределения разности времен срабатывания 1 эн-детектора относительно 2 и 3

5.4. Программа визуализации событий

Для удобного просмотра событий на языке C++ была написана программа визуализации событий и их осциллограмм для установок ENDA и ENDA-INR. На рисунке 5.7 представлена двумерная визуализация события ШАЛ с наибольшим значением числа зарегистрированных нейтронов (за отобранный период времени) в котором реконструированная ось ливня находилась в границах установки. На рисунке эн-детекторы изображены кругами, и их размер определяет количество зарегистрированных нейтронов. Над детекторами выводится число зарегистрированных заряженных частиц, а под детектором число зарегистрированных нейтронов. Красным крестом обозначено реконструированное положение оси ливня, а оценка первичной энергии составляет $\sim 1.8 \cdot 10^{17}$ эВ, если первичная частица была протоном, или $\sim 3.3 \cdot 10^{17}$ эВ, если ядром железа.

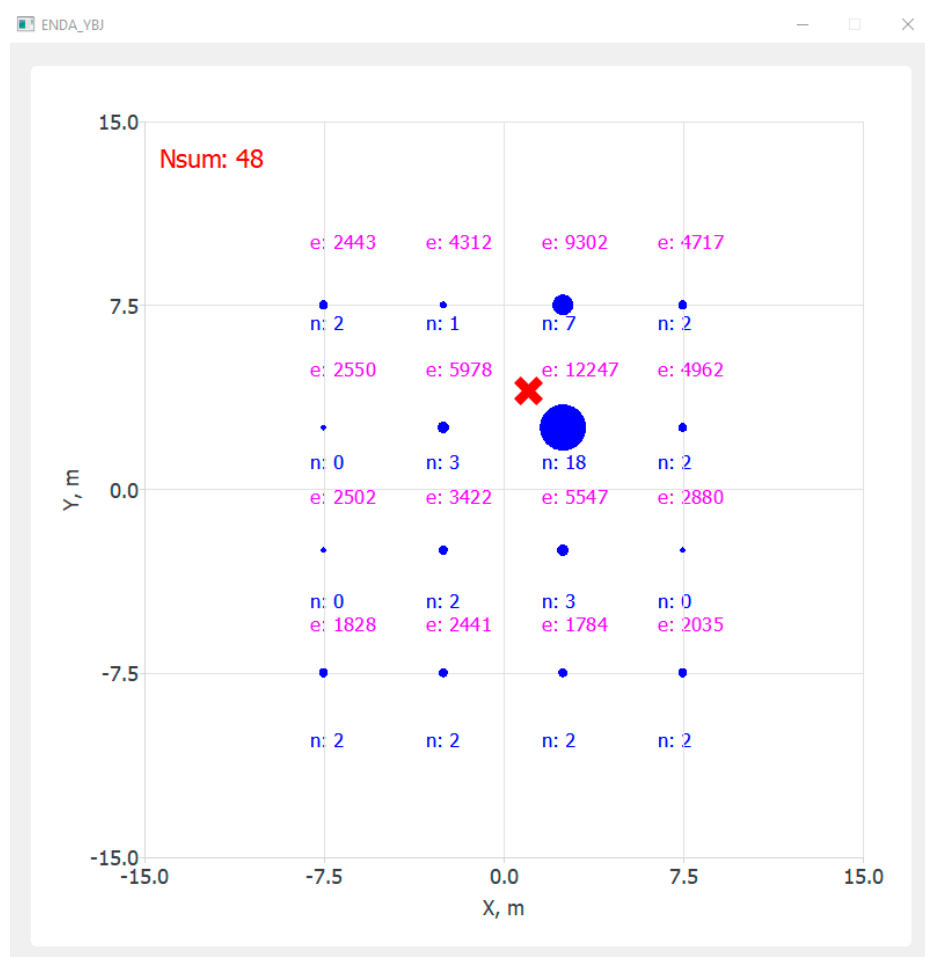


Рис. 5.7: Визуализация события ШАЛ с наибольшим числом зарегистрированных нейтронов в границах установки

На рисунке 5.8 представлена визуализация осциллограммы представленного выше события. На панели отрисовки можно увеличить и рассмотреть желаемую область, а также отрисовать линии только определенных детекторов.

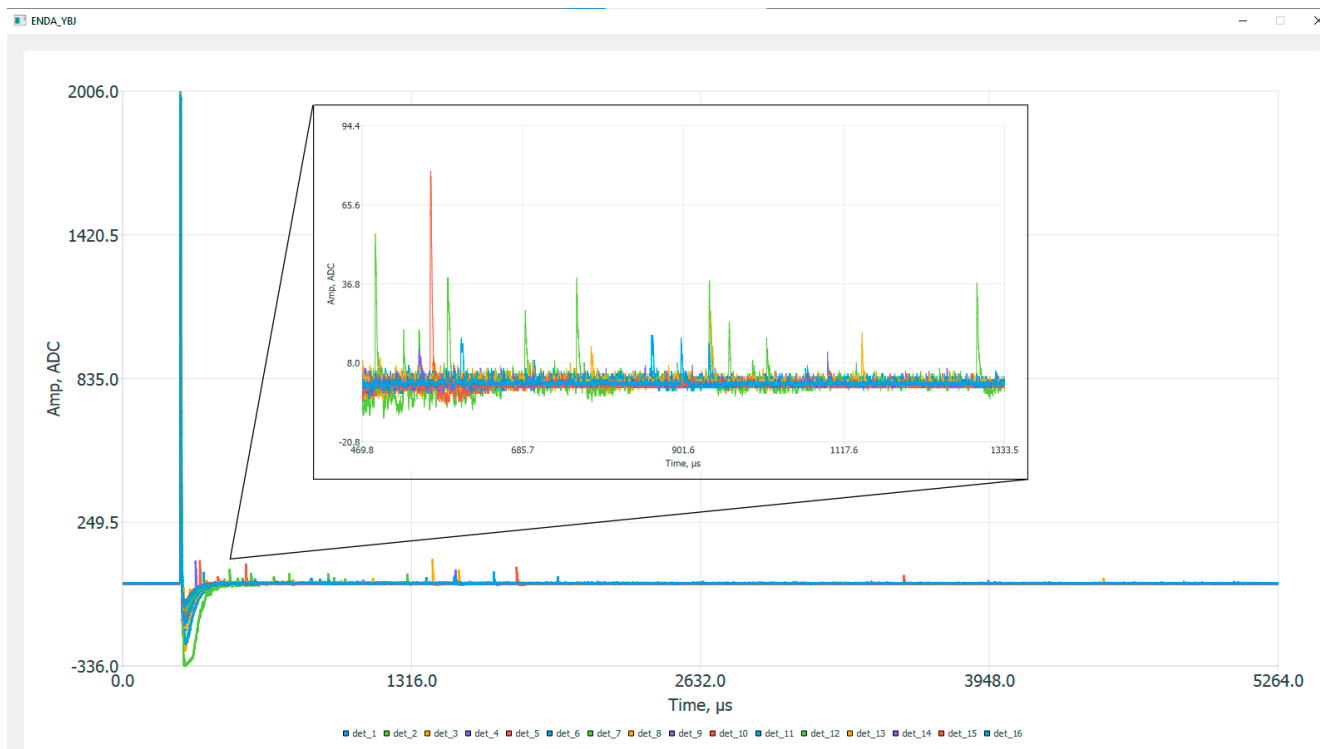


Рис. 5.8: Визуализация осциллограммы события ШАЛ

5.5. Результаты измерительной серии

5.5.1. Отбор событий

Расположение установки ENDA-INR в относительном центре Москвы, рядом с ТЭЦ, приводит к тому, что даже в отфильтрованных данных присутствуют шумовые события. Для их устранения были установлены более жесткие, чем в моделировании, критерии отбора. Требовалось, чтобы минимальное суммарное число заряженных частиц в событии было не меньше 225, чтобы энерговыделение в эн-детекторе было больше чем $k \cdot n_n$, где n_n - зарегистрированное число нейтронов в детекторе, а коэффициент $k = 100$ и был получен эмпирически. Также отбирались события, в которых сработали минимум 10 детекторов с порогом в 12 каналов АЦП (~ 18 частиц). Помимо этого, по данным осциллограмм проводилась проверка отсутствия наложения двух ШАЛ в одном событии.

Всего было отобрано около 77 000 событий. Событий с осциллограммой, которая записывается в случае большого суммарного энергосвечения ($N_{ch} > 6000$) или большого числа зарегистрированных нейтронов ($N_n > 4$), было 715.

5.5.2. Реконструированные параметры ШАЛ

В экспериментальных данных направление прихода ШАЛ можно определить только из данных осциллограмм событий. На рисунке 5.9 представлено реконструированное распределение косинуса зенитного и азимутального угла.

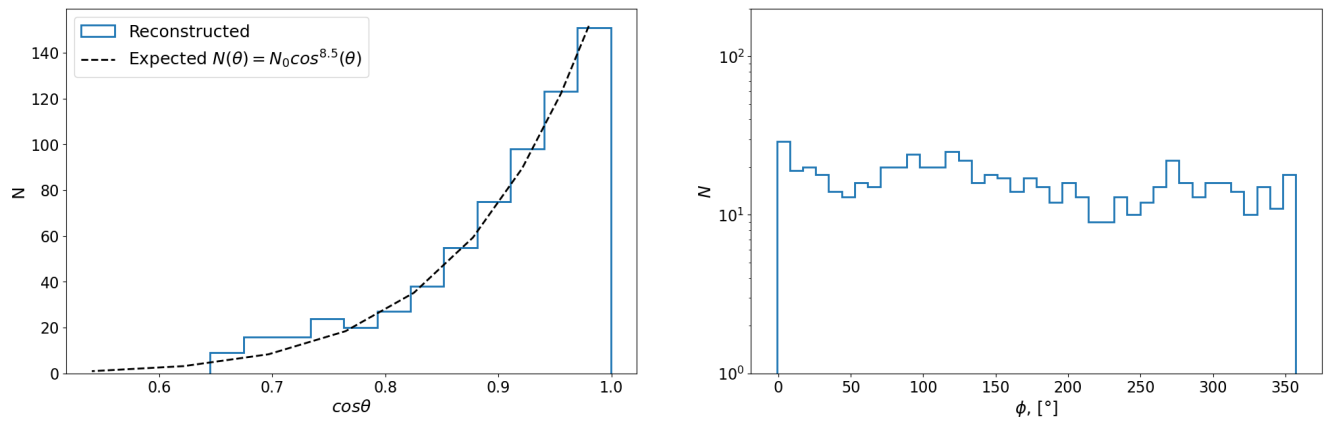


Рис. 5.9: Распределения по косинусу реконструированного зенитного угла (слева) и азимутальному углу (справа)

Распределение по косинусу зенитного угла описывается степенным законом $\sim \cos^\alpha \theta$, где $\alpha = 8.60 \pm 0.54$, что находится в хорошем согласии с ожидаемым значением $\sim 8.5 \div 9$ для уровня моря [74]. Распределение по азимутальному же углу близко к равномерному.

На рисунке 5.10 (слева) изображено распределение реконструированного возраста ливня по данным эксперимента и моделирования. Среднее значение возраста, полученное по экспериментальным данным, составляет 1.35, со среднеквадратичным отклонением 0.38, а по данным моделирования $\mu = 1.39$ с $\sigma = 0.27$. Так как возраст ливня является массово зависимым параметром, а определенного химического состава на данном этапе в выборку не закладыва-

лось (см. п. 3.3.2), то сравнение с моделированием носило качественный характер.

Реконструированное положение осей ливня, представлено на рисунке 5.10 (справа).

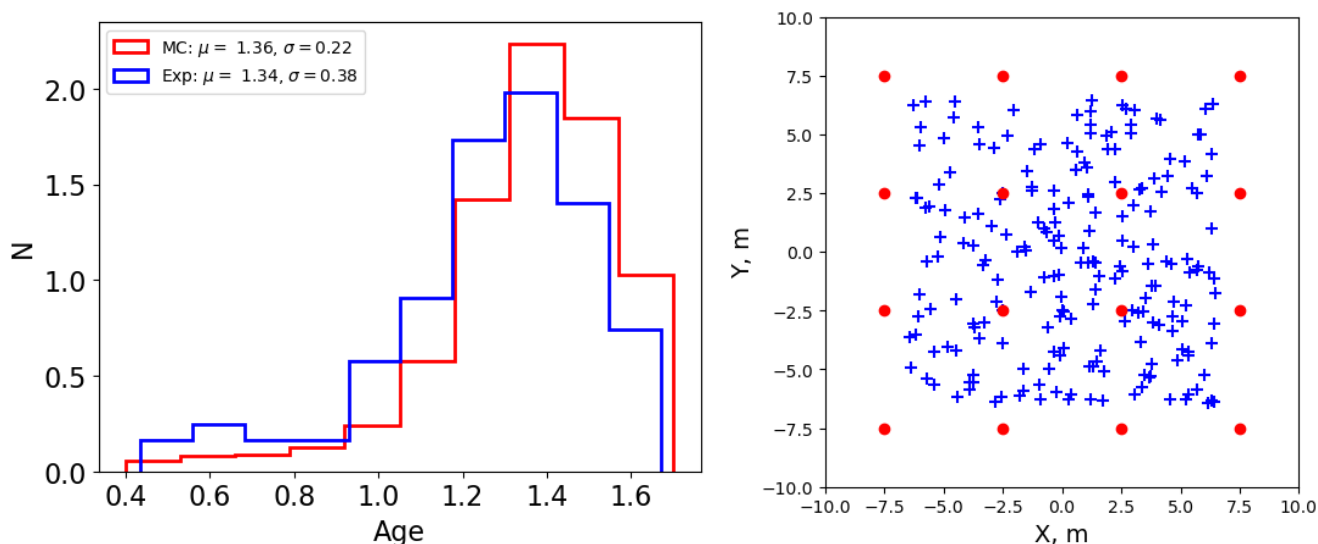


Рис. 5.10: Распределение реконструированного возраста ливня (слева) и реконструированное положение осей ливня, после критериев отбора (справа)

Из рисунка 5.10 (справа) видно, что реконструированное положение осей ливня относительно равномерно заполняют отобранную область.

5.5.3. Временное распределение тепловых нейтронов в ШАЛ

Временное распределение тепловых нейтронов (временные задержки нейтронов относительно фронта ливня) зависит от параметров внешней среды, в частности от химического состава и влажности грунта. Для численной оценки этого влияния было проведено моделирование для сухого грунта, грунта с 10% и 20% влажностью. На рисунке 5.11 представлены полученные временные распределения без нормировки с дифференциальным распределением, полученным из эксперимента (рис. 5.12). В моделировании выборка для сухого и влажного грунта формировалась на одном и том же наборе событий ШАЛ

для протона при фиксированной энергии первичной частицы. К временному распределению, полученному в моделировании, был добавлен экспериментально измеренный фон, разыгрываемый по Пуассону со средним 0.47 нейтронов на событие, равномерно в интервале от 100 мкс до 5 мс.

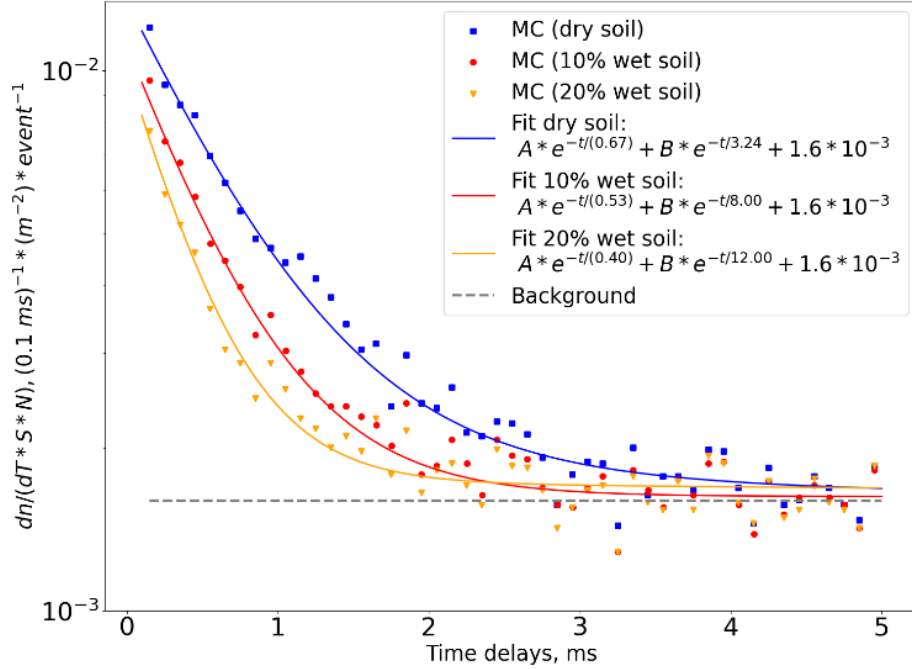


Рис. 5.11: Временное распределение тепловых нейтронов (данные моделирования)

Полученные временные распределения можно фитировать функцией вида (5.2):

$$f(t) = A \times e^{-t/\tau_1} + B \times e^{-t/\tau_2} + C. \quad (5.2)$$

Первая экспонента (5.2) описывает компоненту временного распределения, связанную с локально рожденными нейтронами, где параметр τ_1 связан с временем жизни нейтрона в грунте. Для сухого грунта $\tau_1 = 0.67$ мс, для грунта с 10% влажностью $\tau_1 = 0.53$ мс, с 20% влажностью $\tau_1 = 0.40$ мс, а в эксперименте $\tau_1 = 0.54$ мс. Из полученных результатов видно, что экспериментальные результаты не противоречат модели с 10% влажным грунтом. Вторая компонента связана с нейтронами, рожденными в атмосфере (либо рожденными в грунте, но затем попавшими в атмосферу), и приходящие с больших рассто-

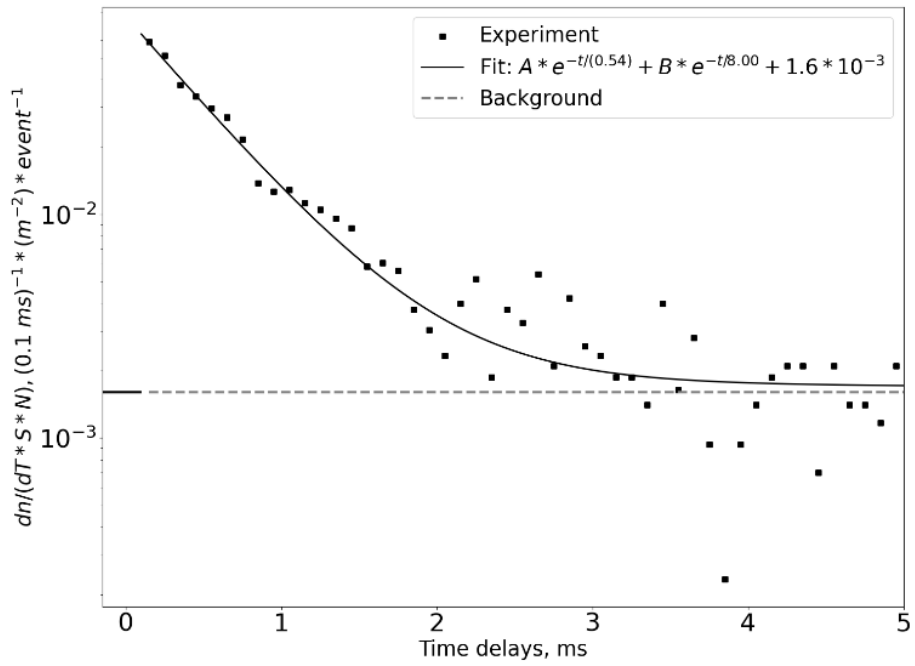


Рис. 5.12: Временное распределение тепловых нейтронов (данные эксперимента)

аний с большой временной задержкой [75], следует учесть, что фон тепловых нейтронов в этой области оказывает преобладающее влияние. На рисунке 5.12 горизонтальной прямой показан фон, оцененный по нулевому мастеру и равный $1.60 \cdot 10^{-3} \text{ см}^{-2} \text{ с}^{-1}$, он близок к уровню, на которое выходит временное распределение, и находится в хорошем согласии с фоном тепловых нейтронов вблизи поверхности земли, который зависит от места проведения измерений $10^{-3} \text{ см}^{-2} \text{ с}^{-1}$ [18].

В моделировании эффективность регистрации нейтронов определяется вероятностью захвата теплового нейтрона в сцинтилляторе ($\varepsilon_{\text{сцин.}}$) и оценивается величиной $\approx 19\%$, что подтверждается экспериментальными измерениями [76]. В нашем же эксперименте эффективность регистрации нейтронов меньше из-за установленного порога по амплитуде импульса ($\varepsilon_{\text{порог}}$) (5 мВ или 10 каналов АЦП) и из-за алгоритма выделения сигналов от тепловых нейтронов ($\varepsilon_{\text{отбор}}$) (из-за чего теряется $1 \div 2\%$ нейтронных импульсов) (см. Главу 2). Эффективность с учетом порога по амплитуде можно оценить с помощью амплитудного распределения нейтронных импульсов (рисунок 5.13), предполагая, что

распределение описывается экспонентой вплоть до нуля, по формуле (5.3):

$$\varepsilon_{\text{порог}} = e^{-\xi/\beta} = e^{-10/26.1} \approx 0.68, \quad (5.3)$$

где ξ - порог по амплитуде, β - показатель экспоненты, полученный в результате фита.

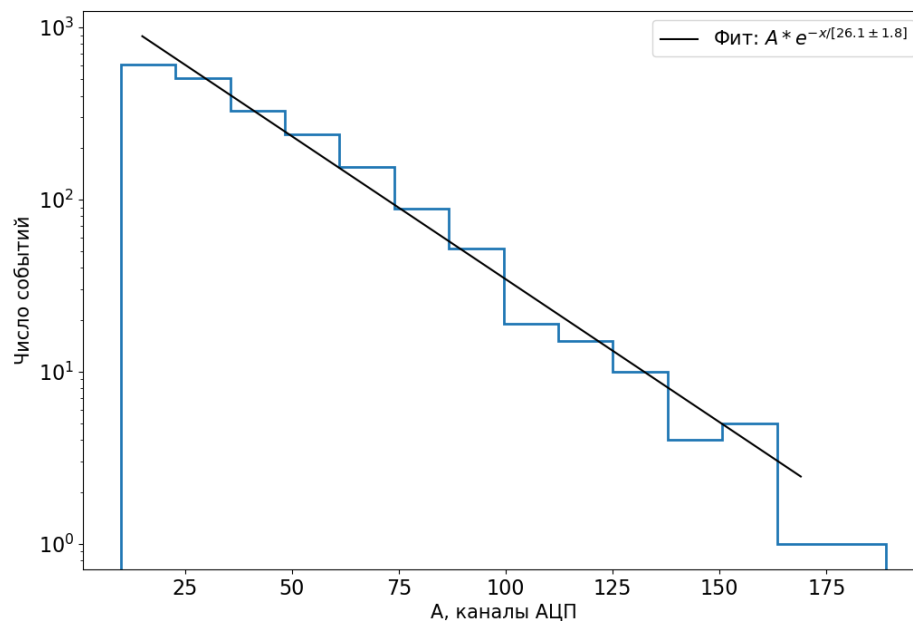


Рис. 5.13: Амплитудное распределение нейтронных импульсов

Из-за чего итоговая эффективность регистрации нейтронов ($\varepsilon_{\text{рег.}}$) равна $\varepsilon_{\text{рег.}} = \varepsilon_{\text{сцин.}} \times \varepsilon_{\text{порог}} \times \varepsilon_{\text{отбор}} = 0.19 \times 0.68 \times 0.98 \approx 12.7\%$. В дальнейшем при сравнении экспериментальных результатов с результатами моделирования на это будет сделана поправка. Также к данным моделирования будет применена поправка связанная с влажностью грунта (см. п. 3.2.2) так как моделирование проводилось для сухого грунта.

5.5.4. Зависимость среднего числа зарегистрированных нейтронов от суммарного энерговыделения и мощности ливня

На рисунке 5.14 показана зависимость среднего числа зарегистрированных нейтронов от суммарного числа зарегистрированных частиц в событии

ШАЛ. При построении данной зависимости не проводилось селекции событий, в которых ось ливня лежала в области отбора (см. п. 4.2.2), так как энергосодержание в детекторе и зарегистрированное число нейтронов уже тесно связано с расстоянием от детектора до оси ШАЛ. Из данных эксперимента был вычтен фон тепловых нейтронов.

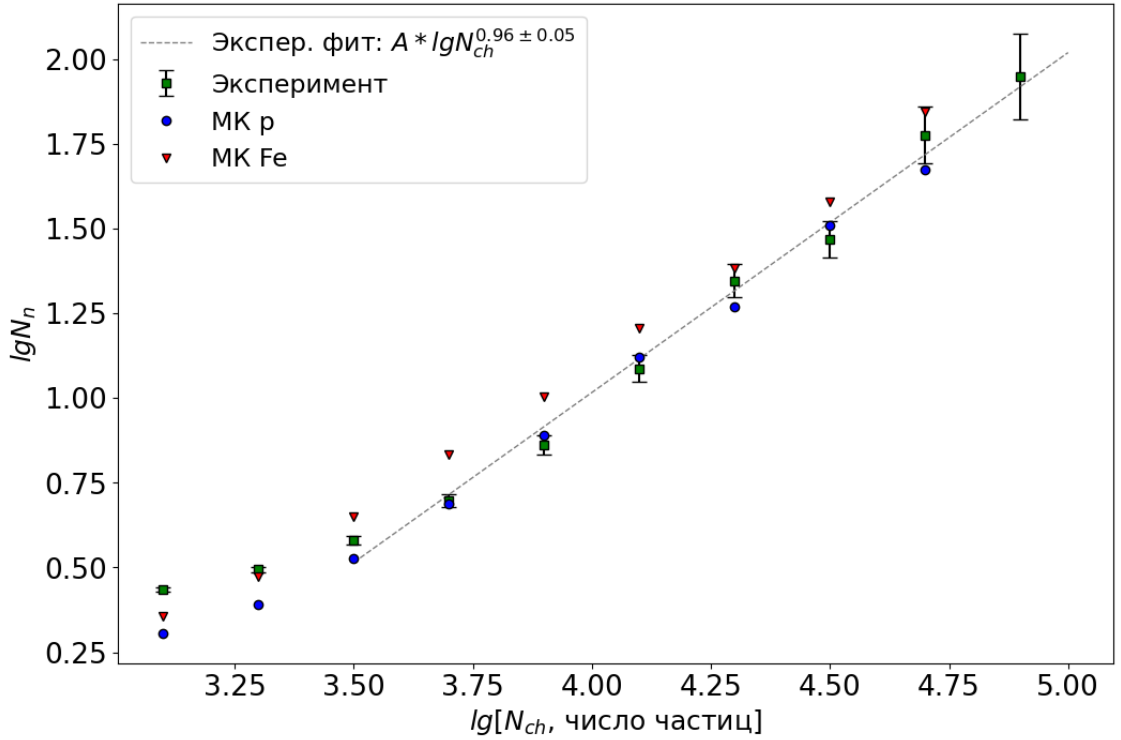


Рис. 5.14: Зависимость среднего числа зарегистрированных нейтронов от суммарного числа зарегистрированных частиц в событии

Полученная зависимость имеет нелинейную и линейную область, где первая связана с отсутствием высокоэнергичных адронов. Линейная область описывается степенным законом $\lg N_n \sim \lg N_{ch}^\alpha$ с $\alpha = 0.96 \pm 0.05$, где фит проводился от 3.5 до 5 по логарифму суммарного числа частиц. По числу нейтронов линейная область соответствует $N_n \geq 4$. Полученный показатель степени находится в хорошем согласии с данными моделирования, где для протонов $\alpha_p = 0.90 \pm 0.05$, а для ядер железа $\alpha_{Fe} = 0.94 \pm 0.07$, а также с экспериментальным результатом, свидетельствующим о пропорциональности плотности адронной компоненты и плотности потока заряженных частиц [77].

Затем были отобраны события, в которых реконструированное положение оси ливня лежало в границах установки, и построена зависимость среднего числа зарегистрированных нейтронов от восстановленной мощности ливня (рис. 5.15). На рисунке представлены экспериментальные точки в области $N_n \geq 4$, жирной линией показан фит в диапазоне, доступном экспериментальным данным, тонкой линией – экстраполяция на меньшую мощность.

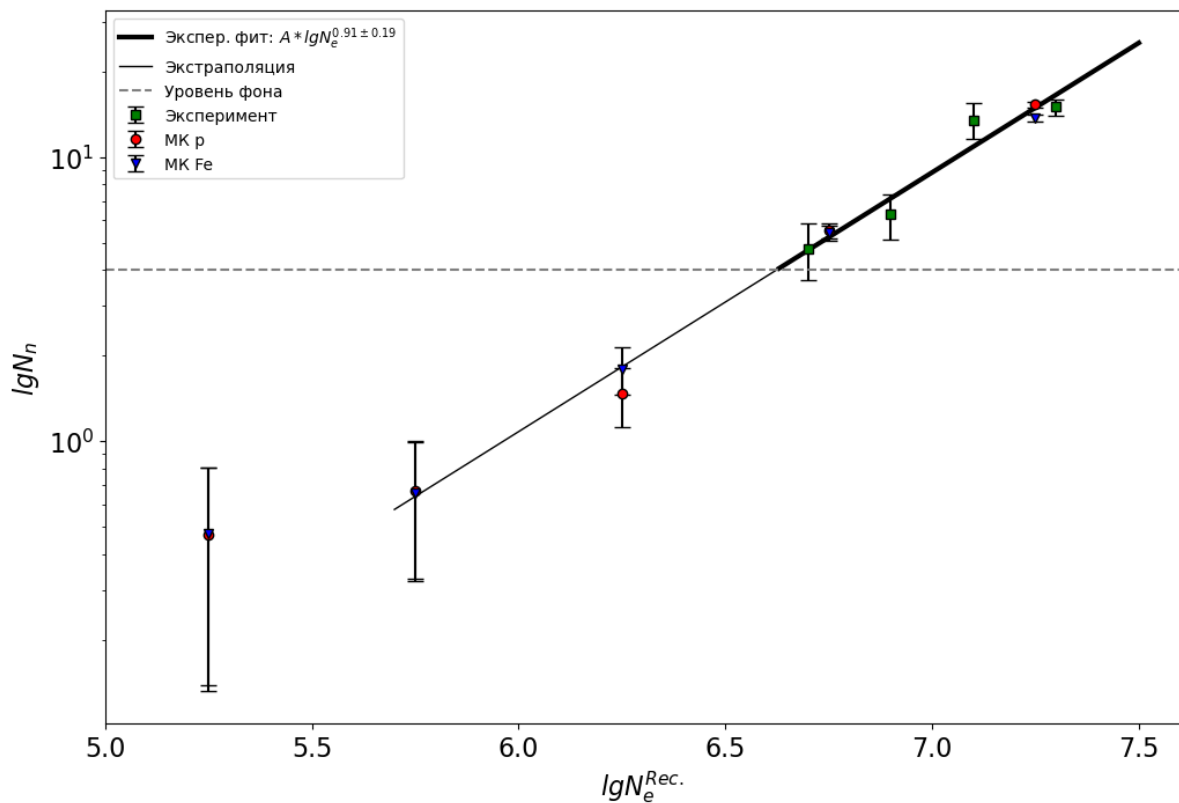


Рис. 5.15: Зависимость числа зарегистрированных нейтронов от реконструированной мощности ливня

Зависимость описывается степенным законом $\sim \lg N_e^\alpha$, где по данным эксперимента $\alpha_{exp} = 0.91 \pm 0.19$. По данным моделирования для протона $\alpha_p = 0.93 \pm 0.06$, а для железа $\alpha_{Fe} = 0.89 \pm 0.02$, где фит проводился в диапазоне от 5.75 до 7.25 по логарифму мощности.

5.5.5. ФПР тепловых нейтронов в ШАЛ

Была построена ФПР нейтронной компоненты для событий, в которых $N_n \geq 4$, в эксперименте таких событий было 39. На рисунке 5.16 представлены результаты моделирования с кривой, соответствующей функции Кемпа [78], которая является аппроксимацией пространственного распределения адронов в ШАЛ. Из рисунка 6 видно, что такая аппроксимация хорошо описывает ФПР тепловых нейтронов.

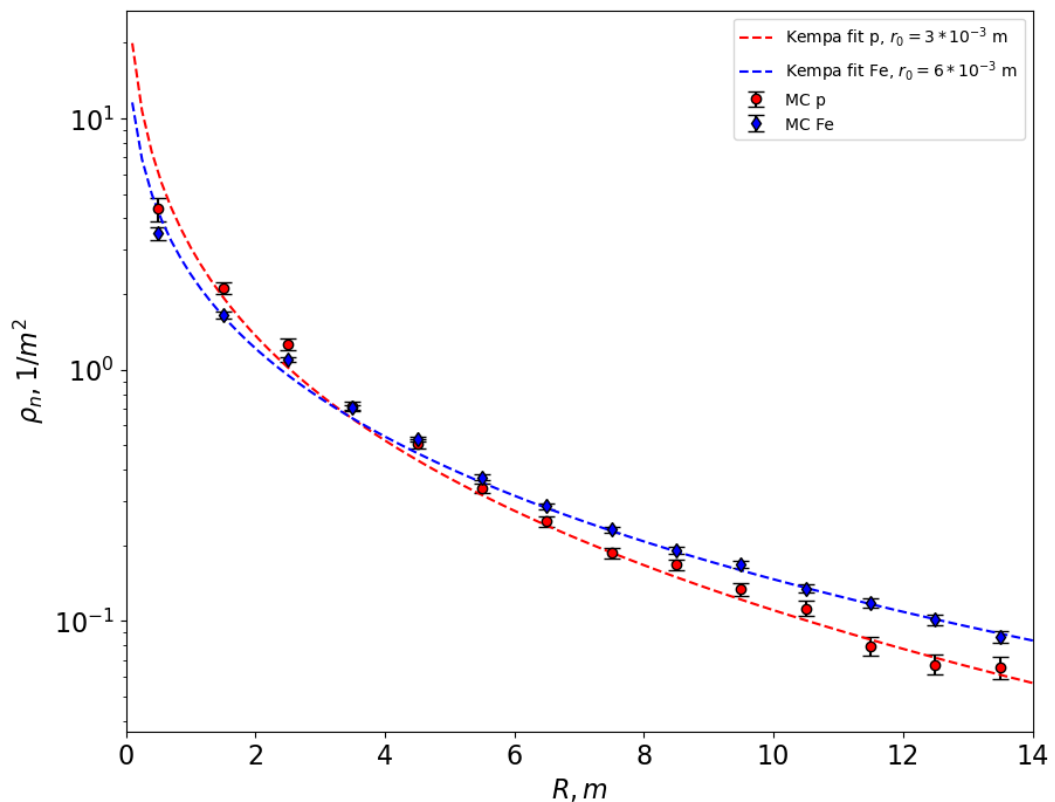


Рис. 5.16: ФПР тепловых нейтронов (результаты моделирования)

Затем на данные моделирования был наложен экспериментально оцененный фон тепловых нейтронов и проведено сравнение с ФПР, полученной из данных эксперимента (рисунок 5.17).

Данные можно описать функцией следующего вида (5.4):

$$\rho_n(R) = A \times e^{-R/r_1} + B \times e^{-R/r_2} + C, \quad (5.4)$$

где $r_1 = 1.24$ м, $r_2 = 6.01$ м, а $C = 3.19$ м⁻².

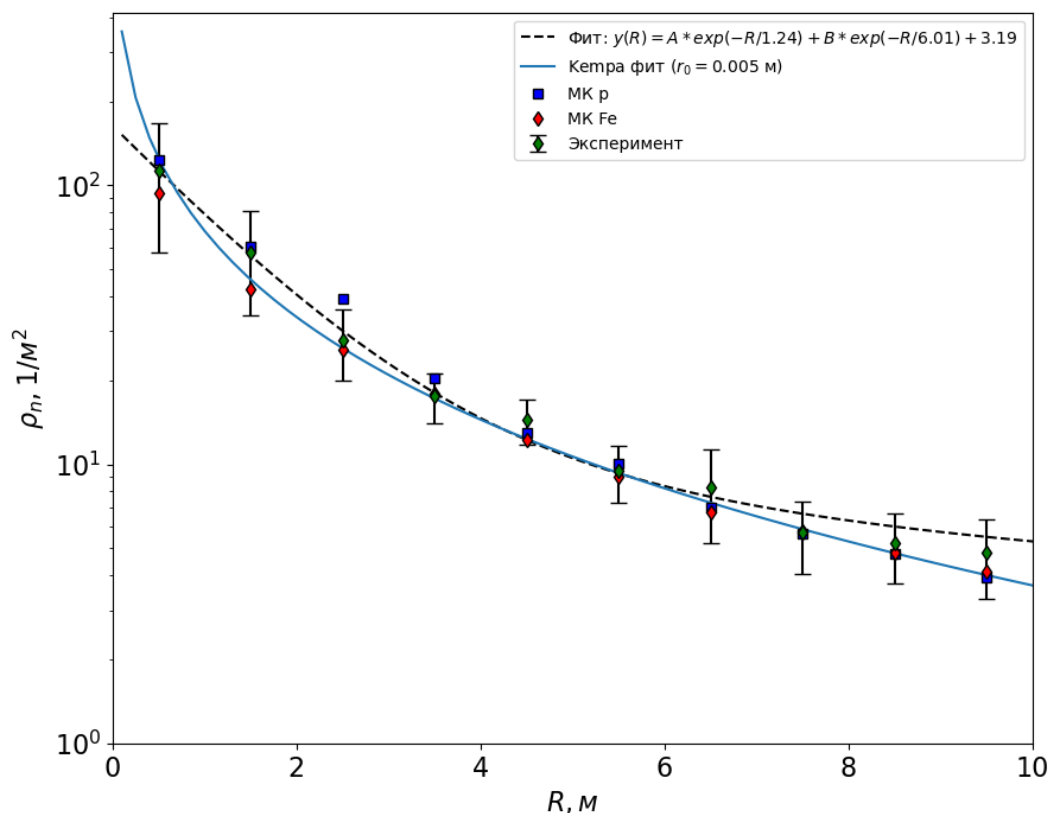


Рис. 5.17: ФПР тепловых нейтронов

ФПР тепловых нейтронов на уровне моря впервые была получена на установке PRISMA-32. Её вид также представляет из себя сумму двух экспонент [79, 80], с параметрами $r_1 = 1.4$ м, $r_2 = 8.2$ м.

Вторая экспонента также, как и вторая компонента во временном распределении, возникает из-за нейтронов, пришедших из атмосферы. Из-за того, что установки находятся в разных внешних условиях (PRISMA-32 расположена в здании), параметр r_2 будет отличаться. Параметра же r_1 обоих экспериментов находится в хорошем согласии друг с другом и связан с характерным расстоянием родительских адронов относительно оси ливня.

На малых расстояниях от оси ШАЛ регистрируются в основном нейтроны, рожденные высокоэнергичными адронами, возникает это из-за того, что нейтроны имеют малую длину диффузии в плотном веществе, а адроны имеют крутую ФПР в этой области, что приводит к тому, что ФПР тепловых нейтронов близка к ФПР родительских адронов.

5.5.6. Распределение ШАЛ по числу нейтронов

Так как тепловые нейтроны ШАЛ рождаются преимущественно высокоэнергетичными адронами ливня, то показатель наклона спектра по числу нейтронов должен совпадать с показателем наклона спектра по числу адронов (см. п. 4.3.3).

В моделировании интегральный спектр по числу нейтронов был получен для трёх различных выборок. В первой выборке энергетический спектр разыгрывался по степенному закону $(\gamma + 1) = 2.7$, во второй был заложен излом в области 3 ПэВ, при котором показатель спектра изменялся на $(\gamma + 1) = 3.1$ (рис. 5.18), в третьей выборке энергетический спектр разыгрывался по модели Gaisser H3a [81] (рис. 5.19). В качестве первичной частицы для первой и третьей выборки были взяты все разыгранные типы первичных частиц, а для второй выборки только протоны с целью оценки величины показателя наклона спектра нейтронов в случае возможного излома энергетического спектра.

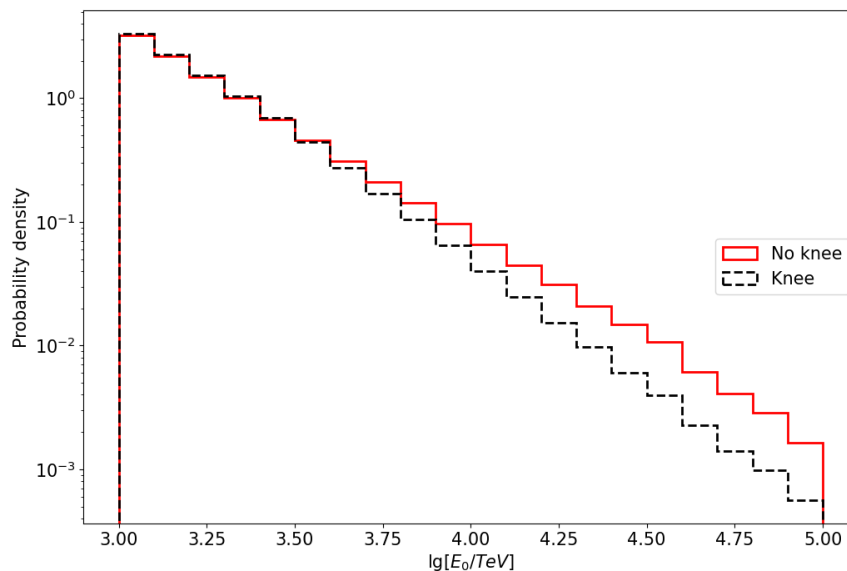


Рис. 5.18: Первичный энергетический спектр с изломом и без излома

При построении интегрального спектра по числу нейтронов, использовались события, в которых реконструированное положение оси ливня находилось в границах установки. К зарегистрированному числу нейтронов был наложен

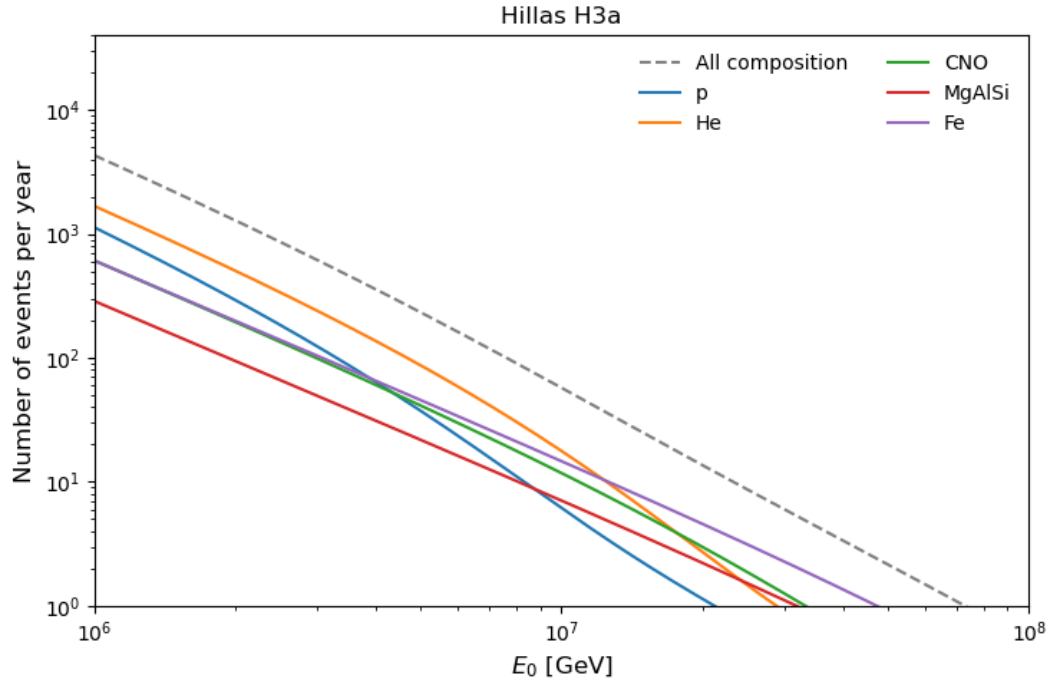


Рис. 5.19: Первичный энергетический спектр нормированный к модели Gaisser H3a

экспериментально измеренный фон. На рисунке 5.20 представлены интегральные спектры по числу нейтронов для протонных ливней в случае излома первичного спектра и без излома.

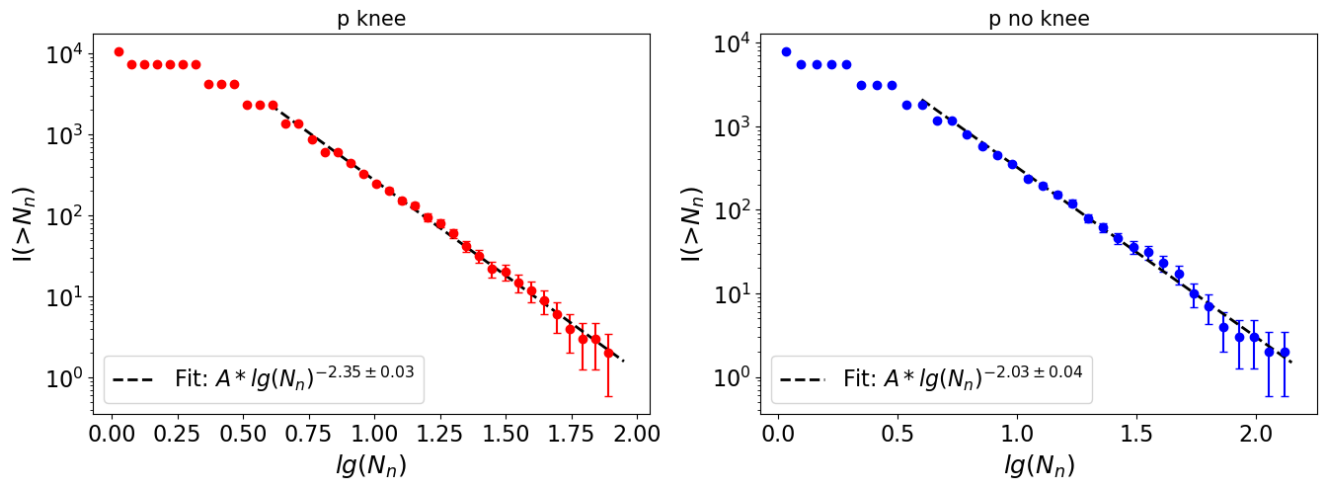


Рис. 5.20: Интегральный спектр по числу нейтронов для протонов с изломом в первичном спектре (слева) и без излома (справа)

Полученный спектр соответствует степенному закону с показателем

$\beta_{knee} = -2.35 \pm 0.03$ для выборки с изломом первичного спектра и с $\beta = -2.03 \pm 0.04$ для выборки без излома. На рисунке 5.21 представлены интегральные спектры по числу нейтронов в ШАЛ для первичных протонов, ядер гелия и железа, для выборки без излома первичного спектра. Спектры были домножены на указанные в легенде коэффициенты с целью устранения сливания их друг с другом.

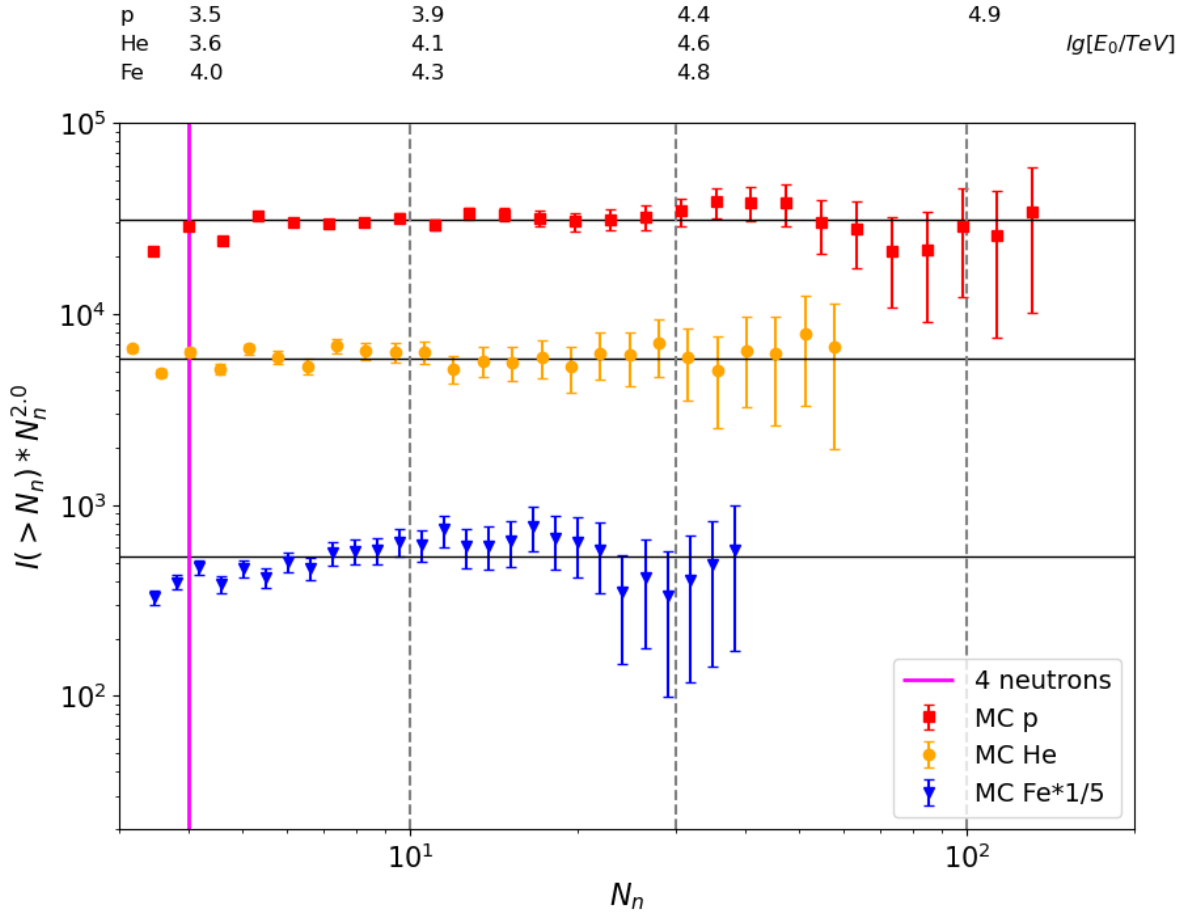


Рис. 5.21: Интегральный спектр по числу нейтронов для энергетического спектра без излома для p , He и Fe

Полученные показатели спектра: $\beta_p = -2.03 \pm 0.04$ для протонной выборки, $\beta_{He} = -1.99 \pm 0.04$ для выборки гелия и $\beta_{Fe} = -2.01 \pm 0.13$ для выборки с железом.

На рисунке 5.22 представлен интегральный спектр по числу нейтронов для выборки, в которой энергетический спектр и химический состав был задан

согласно модели Gassier H3a.

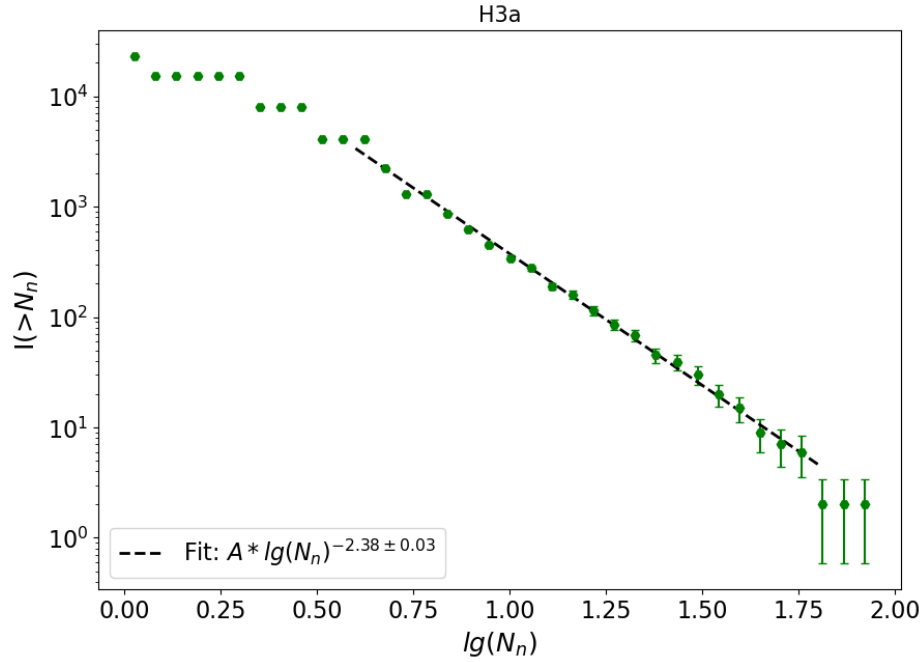


Рис. 5.22: Интегральный спектр по числу нейтронов для энергетического спектра модели Gassier H3a

Полученный спектр соответствует степенному закону с показателем $\beta = -2.38 \pm 0.03$.

Из данных эксперимента интегральный спектр по числу нейтронов в ШАЛ был получен с использованием различных алгоритмов выделения нейтронных импульсов из осциллограмм и различных порогов на амплитуду импульса. Были использованы описанные в главе 2 алгоритмы, а также вариант без отбора по форме импульса. Результаты представлены на рисунке 5.23 и в таблице 5.1.

Измеренные показатели степени β согласуются с данными установки PRISMA-32 [82] и с результатом PRISMA-YBJ [83], а также с данными адронного калориметра KASCADE, в котором показатель спектра наклона по числу адронов был в диапазоне $1.9 \div 2.0$ в зависимости от выбора пороговой энергии [84].

На рисунке 5.24 представлен экспериментально полученный интеграль-

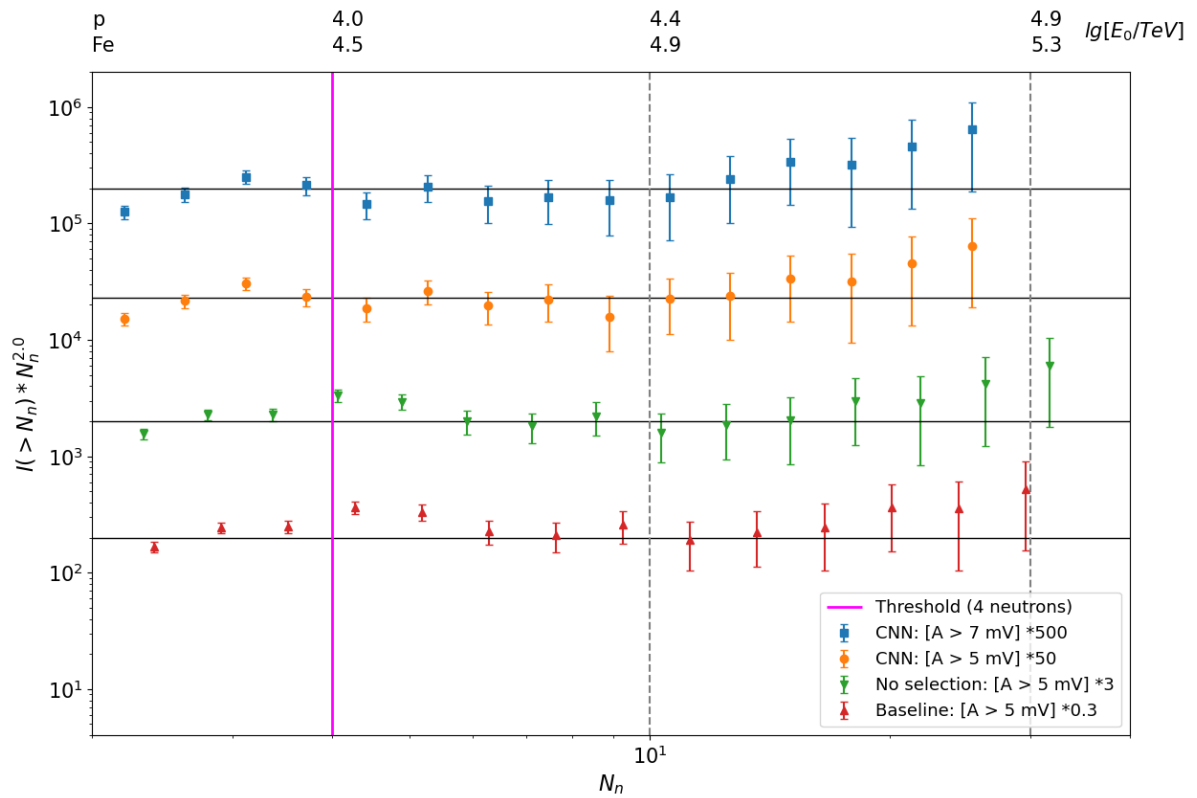


Рис. 5.23: Экспериментально полученный интегральный спектр по числу нейтронов

Таблица 5.1: Значение показателя спектра по числу нейтронов для различных вариантов отбора

Метод отбора	Порог по амплитуде	β
CNN	5 мВ	-1.96 ± 0.20
CNN	7 мВ	-2.01 ± 0.18
Базисный алгоритм	5 мВ	-2.02 ± 0.22
Без отбора	5 мВ	-2.16 ± 0.18

ный спектр по числу нейтронов для метода отбора с помощью CNN и порогом по амплитуде 7 мВ совместно с результатом моделирования для спектра, заданного по модели Gassier H3a, с учетом поправки на эффективность регистрации и влажности грунта.

Большие статистические ошибки, связанные с малой статистикой, на

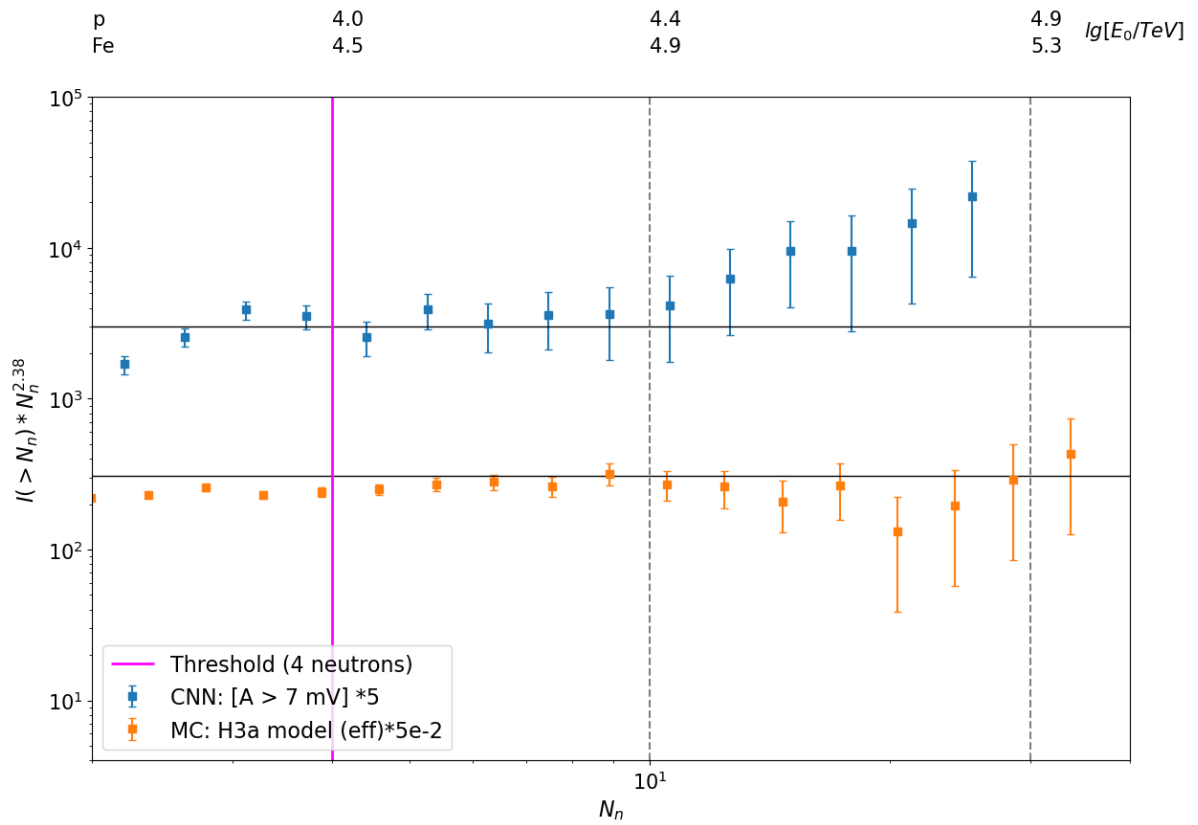


Рис. 5.24: Интегральный спектр по числу нейтронов полученный по данным эксперимента для отбора с помощью CNN (с порогом 7 мВ) и по результатам моделирования для модели энергетического спектра Gassier H3a

данный момент не позволяют однозначно исключить какую-либо рассматриваемую модель энергетического спектра из рассмотрения.

Заключение

Основные результаты диссертационной работы заключаются в следующем.

1. Был разработан метод выделения сигналов от тепловых нейтронов с использованием сверточных нейронных сетей в эн-детекторах. Метод показал лучшее качество по сравнению с используемым на данный момент в установке и в дальнейшем будет применен для обработки данных эксперимента ENDA.
2. Была создана программа быстрого Монте-Карло моделирования отклика установки ENDA на прохождение ШАЛ. Программа показала результаты, близкие к результатам программы моделирования на Geant4, а её время вычисления оказалось меньше в ~ 2000 раз. По данным моделирования была получена функция пространственного распределения тепловых нейтронов в ШАЛ, а также интегральный спектр по числу нейтронов в ливне.
3. Была получена оценка точности реконструкции параметров ливня (направление прихода и положение оси ливня, полное число адронов и заряженных частиц) для базисного подхода и подхода с использованием методов машинного обучения.
4. Был разработан метод реконструкции энергии первичной частицы с использованием данных о нейтронной компоненте ШАЛ и с оценкой неопределенности такого восстановления с помощью методов машинного обучения. С помощью методов машинного обучения удалось улучшить качество реконструкции первичной энергии на $\sim 10.5\%$.
5. По данным эксперимента ENDA-INR была получена функция пространственного распределения тепловых нейтронов в ливне, которая показала хорошее согласие с результатами моделирования и с результатом установки PRISMA-32.

6. По данным эксперимента ENDA-INR был получен интегральный спектр по числу нейтронов в ШАЛ, имеющий степенной вид с показателем ~ 2.0 , который показал хорошее согласие с результатом установок PRISMA-32, PRISMA-YBJ и с результатом установки KASCADE для спектра по числу адронов.

Автор выражает искреннюю признательность своему научному руководителю Щеголеву О. Б. за предоставленную тему диссертации, ценные рекомендации, постоянное внимание и поддержку на всех этапах работы. Также автор благодарит Стенькина Ю. В. за активные процессы обсуждения результатов, полезные советы и важные замечания. Большую благодарность автор выражает Кулешову Д. А. за переданные знания, связанные с экспериментом и процессом моделирования.

Автор благодарит коллектив НОЦ НЕВОД за приглашения на семинары, которые оказали плодотворное влияние на степень проработки диссертации и отдельно выражает благодарность Громушкину Д. М. и Богданову А. Г.

Также автор искренне благодарен профессору Пекинского института физики высоких энергий Ма Синьхуа и всем китайским коллегам, участвующим в проекте ENDA, за ценные обсуждения и поддержку.

Словарь терминов

Кросс-валидация – метод оценки модели и её поведения на независимых данных.

Батч – партия данных.

Функция потерь – функция, которая характеризует потери при неправильном принятии решений на основе наблюдаемых данных.

Гиперпараметр – параметр алгоритма, значение которого устанавливается перед запуском процесса обучения.

Метрика – мера качества работы алгоритма на выборке.

Переобучение - явление, когда построенная модель хорошо объясняет примеры из обучающей выборки, но относительно плохо работает на примерах, не участвовавших в обучении.

Эпоха – одна итерация в процессе обучения.

Признаки – переменные, описывающие характеристики объекта.

Объект – некоторый элемент из всего множества данных, для которого производятся предсказания целевой переменной (или переменных).

Список иллюстраций

1.1	Расположение детекторов эксперимента LHAASO	14
1.2	Фото установки ENDA-INR (слева), конфигурация детекторов (справа)	15
1.3	Схема эн-детектора (1 – высоковольтный входной порт, 2 – ДИУ 8-го динода, 3 – выход 5-го динода, 4 – пластиковый корпус, 5 – держатель ФЭУ, 6 – ФЭУ, 7 – светособирающий конус, 8 – сцинтиллятор) (слева), образец сцинтиллятора (справа).	16
1.4	Фронт импульса заряженной и нейтронной компоненты	17
1.5	Осциллограмма события ШАЛ	18
1.6	Визуализация события ШАЛ с наибольшим числом зарегистрированных нейтронов	19
1.7	Визуализация события ШАЛ с наибольшим числом зарегистрированных нейтронов с осциллограммой каждого отдельного эн-детектора	20
2.1	Алгоритм формирования обучающей выборки.	24
2.2	Распределение по амплитуде сигнала	25
2.3	Распределение по длительности фронта импульса	25
2.4	Архитектура нейронной сети	26
2.5	t-SNE представление эмбедингов. Исходные данные (слева), после обучения (справа).	27
2.6	Графики обучения нейронной сети. График зависимости потерь от эпохи (слева), доли верных ответов от эпохи (справа)	28
2.7	Матрица ошибок. Базисный метод (слева), CNN (справа)	28
3.1	Эффективность γ , $e^\pm$, $\mu^\pm$ (слева) и p , n , $\pi^\pm$ (справа) дать отклик в эн-детекторе в зависимости от энергии частицы	33

3.2	Визуализация модели эн-детектора со стальным корпусом (слева), эффективность гамма-кванта дать отклик эн-детектора (справа) со стальным корпусом (синяя кривая), без стального корпуса (оранжевая кривая)	34
3.3	Зависимость эффективности дать отклик детектора в зависимости от расстояния от центра детектора для гамма-квантов и электронов	35
3.4	Зависимость среднего числа зарегистрированных нейтронов в зависимости от энергии налетающей частицы для протона и гамма-кванта	36
3.5	Зависимость среднего числа зарегистрированных нейтронов на 1 протон в зависимости от расстояния от места ядерной реакции до детектора	37
3.6	Распределения энергосвечения в эн-детекторе от прохождения ШАЛ	38
3.7	Распределения энергосвечиваний в эн-детекторе для мюона с энергией в 100 МэВ (слева) и 1 ГэВ (справа)	39
3.8	Распределения по числу частиц в четырёх детекторах для программы быстрого моделирования (синий цвет) и Geant4 (оранжевый цвет)	40
3.9	Распределения по числу зарегистрированных нейтронов для программы быстрого моделирования (оранжевый цвет) и Geant4 (синий цвет) для $E_0 = 1$ ПэВ (слева сверху), $E_0 = 3$ ПэВ (справа сверху) и для $E_0 = 10$ ПэВ (снизу)	41
3.10	Красная штриховая область – границы установки, синий круг – зона розыгрыша оси ШАЛ	42
4.1	Зависимость эффективности регистрации от энергии первичной частицы для протона (слева) и железа (справа)	43

4.2	Зависимость эффективности регистрации от энергии первичной частицы для протона (слева) и железа (справа). N_{min}^{det} , A_{min} — минимальное количество и амплитуда сработавших эн-детекторов	44
4.3	Разыгранный энергетический спектр после первичного критерия отбора событий	45
4.4	Параметризация поправки $\alpha(r)$	48
4.5	Истинная плотность электронной компоненты и наложенные фиты различных ФПР. Функция (4.5) - чёрный цвет, (4.6) - жёлтый цвет, (4.7) - красный цвет	49
4.6	Отношение истинной плотности к полученной в результате аппроксимации. Функция (4.5) - жёлтый цвет, (4.6) - синий цвет, (4.7) - красный цвет	49
4.7	Двумерные распределения логарифма отношения реконструированной и истинной мощности ШАЛ к реконструированному возрасту ливня для протона (слева) и железа (справа). Штриховая линия обозначает отбрасываемую область	51
4.8	Распределение погрешности определения оси ливня во всех реконструированных событиях (синяя кривая) и в событиях с $N_n \geq 4$ (красная кривая)	53
4.9	Тепловая карта распределения осей ШАЛ с отбором в границах установки. Реконструированное положение (слева), истинное (по центру), разница (справа)	54
4.10	Тепловая карта распределения осей ШАЛ с отбором с отступом в 1 метр от границ установки. Реконструированное положение (слева), истинное (по центру), разница (справа)	54
4.11	Корреляционные зависимости истинной мощности ливня от реконструированного значения для протона (слева) и ядра железа (справа), штриховая линия соответствует прямой пропорциональности	55

4.12	Пример события с большой ошибкой реконструкции положения оси ШАЛ	57
4.13	Пример события с малой ошибкой реконструкции положения оси ШАЛ	58
4.14	Распределение по ψ для МНК (синяя кривая), IRLS (оранжевая кривая), GBDT (зеленая кривая) для протона (слева) и железа (справа)	61
4.15	Зависимость эффективности реконструкции направления прихо- да ШАЛ (слева) и углового разрешения (справа) от числа заре- гистрированных нейтронов в событии	62
4.16	Зависимость разрешения определения положения оси ливня от энергии первичной частицы	63
4.17	Зависимость эффективности реконструкции положения оси лив- ня для различных порогов по числу заряженных частиц (слева) и числу нейтронов (справа)	64
4.18	Распределение относительной ошибки реконструкции мощности ливня	65
4.19	Зависимость относительной ошибки реконструкции мощности ливня от логарифма истинной мощности	66
4.20	Зависимость полного числа адронов от числа зарегистрирован- ных нейтронов для ливней от протона (сверху), от железа (по центру) и для смешанного состава (снизу)	67
4.21	Распределение абсолютной ошибки реконструкции полного числа адронов ливня (слева) для гибридного (красная кривая) и базис- ного подхода (черная кривая), и корреляция реконструированно- го и истинного значения N_h для гибридного метода (справа) . . .	69
4.22	Зависимость первичной энергии от числа зарегистрированных нейтронов в событии для первичных протонов (сверху), железа (центр), смешанный состав (снизу)	71

4.23	Зависимость первичной энергии от реконструированной мощности ШАЛ для первичных протонов (сверху), железа (центр), смешанная выборка (снизу)	73
4.24	Распределения отношения реконструированной и истинной энергии первичной частицы для LRPF, RF и GBDT	75
4.25	Распределения отношения реконструированной и истинной энергии первичной частицы для алгоритма GBDT и базисного подхода для модели QGSJET-II-04	76
4.26	Распределения отношения реконструированной и истинной энергии первичной частицы для алгоритма GBDT и базисного подхода для модели EPOS-LHC	77
4.27	Распределения отношения реконструированной и истинной энергии первичной частицы для алгоритма GBDT и базисного подхода для модели Sibyll-2.3d	78
4.28	Неопределенность в данных (слева), неопределенность в знаниях (центр, справа). Зеленая точка – объект из выборки, красные точки – обучающая выборка	81
4.29	Распределение разностей между предсказанным и истинным значением целевой переменной, деленное на предсказанное среднеквадратичное отклонение в сравнении с нормальным распределением, что указывает на хорошую оценку неопределенности	82
5.1	Ежедневный рабочий цикл установки	83
5.2	Темп счета триггеров второго уровня	84
5.3	Диаграммы рассеяния (черным пунктиром обозначена линейная аппроксимация), красной линией обозначено максимальное значение для 8 динода	85
5.4	Распределение по энерговыделению каждого отдельного детектора	86

5.5	Распределение по ΔT в моделировании (слева) и эксперименте (справа)	87
5.6	Распределения разности времен срабатывания 1 эн-детектора относительно 2 и 3	88
5.7	Визуализация события ШАЛ с наибольшим числом зарегистрированных нейтронов в границах установки	89
5.8	Визуализация осциллограммы события ШАЛ	90
5.9	Распределения по косинусу реконструированного зенитного угла (слева) и азимутальному углу (справа)	91
5.10	Распределение реконструированного возраста ливня (слева) и реконструированное положение осей ливня, после критериев отбора (справа)	92
5.11	Временное распределение тепловых нейтронов (данные моделирования)	93
5.12	Временное распределение тепловых нейтронов (данные эксперимента)	94
5.13	Амплитудное распределение нейтронных импульсов	95
5.14	Зависимость среднего числа зарегистрированных нейтронов от суммарного числа зарегистрированных частиц в событии	96
5.15	Зависимость числа зарегистрированных нейтронов от реконструированной мощности ливня	97
5.16	ФПР тепловых нейтронов (результаты моделирования)	98
5.17	ФПР тепловых нейтронов	99
5.18	Первичный энергетический спектр с изломом и без излома	100
5.19	Первичный энергетический спектр нормированный к модели Gaisser H3a	101
5.20	Интегральный спектр по числу нейтронов для протонов с изломом в первичном спектре (слева) и без излома (справа)	101

5.21	Интегральный спектр по числу нейтронов для энергетического спектра без излома для p , He и Fe	102
5.22	Интегральный спектр по числу нейтронов для энергетического спектра модели Gassier H3a	103
5.23	Экспериментально полученный интегральный спектр по числу нейтронов	104
5.24	Интегральный спектр по числу нейтронов полученный по данным эксперимента для отбора с помощью CNN (с порогом 7 мВ) и по результатам моделирования для модели энергетического спектра Gassier H3a	105

Список таблиц

2.1	Сравнение результатов для тестирующей выборки	29
3.1	Тип и энергия разыгранных частиц	32
4.1	Сравнение результатов восстановления различных ФПР	52
4.2	Сравнение результатов восстановления для различных критериев отбора событий	55
4.3	Сравнение метрик классификации и углового разрешения	61
4.4	Сравнение качества реконструкции положения оси ШАЛ для раз- личных параметризаций целевой переменной при использовании машинного обучения.	63
4.5	Критерии отбора и процент отброшенных событий	70
4.6	Результаты оценки точности реконструкции первичной энергии при восстановлении с помощью метода GBDT	79
4.7	Результаты оценки точности реконструкции первичной энергии при восстановлении с помощью мощности ливня	79
4.8	Результаты оценки точности реконструкции первичной энергии при восстановлении по суммарному числу нейтронов	80
5.1	Значение показателя спектра по числу нейтронов для различных вариантов отбора	104

Список литературы

1. Cosmic-ray proton and helium spectra: results from the JACEE experiment / K Asakimori, TH Burnett, ML Cherry et al. // *The Astrophysical Journal*. — 1998. — Vol. 502, no. 1. — P. 278.
2. Measurement of the cosmic ray proton spectrum from 40 GeV to 100 TeV with the DAMPE satellite / DAMPE collaboration, Q An, R Asfandiyarov et al. // *Science advances*. — 2019. — Vol. 5, no. 9. — P. eaax3793.
3. Observation of spectral structures in the flux of cosmic-ray protons from 50 GeV to 60 TeV with the calorimetric electron telescope on the international space station / O Adriani, Y Akaike, K Asano et al. // *Physical Review Letters*. — 2022. — Vol. 129, no. 10. — P. 101102.
4. Григоров Н. Л. и др. Изучение энергетического спектра и состава первичных космических лучей в области высоких и сверхвысоких энергий на ИСЗ "Протон-1" и "Протон-2" // *Космические исследования*. — 1967. — Vol. 5, no. 3. — P. 395.
5. Peters B. Primary cosmic radiation and extensive air showers // *Il Nuovo Cimento (1955-1965)*. — 1961. — Vol. 22, no. 4. — Pp. 800–819.
6. Berezhko EG, Ksenofontov LT. Composition of cosmic rays accelerated in supernova remnants // *Journal of Experimental and Theoretical Physics*. — 1999. — Vol. 89, no. 3. — Pp. 391–403.
7. Kobayakawa K, Honda YS, Samura T. Acceleration by oblique shocks at supernova remnants and cosmic ray spectra around the knee region // *Physical Review D*. — 2002. — Vol. 66, no. 8. — P. 083004.
8. Candia Julian, Roulet Esteban, Epele Luis N. Turbulent diffusion and drift in galactic magnetic fields and the explanation of the knee in the cosmic ray spectrum // *Journal of High Energy Physics*. — 2003. — Vol. 2002, no. 12. — P. 033.

9. *Kazanas D, Nicolaidis A.* Cosmic Ray "Knee": A Herald of New Physics? // *arXiv preprint astro-ph/0103147*. — 2001.
10. *Petrukhin AA.* Problem of the knee and very high energy muons // *Proceedings of the 27th International Cosmic Ray Conference*. 07-15 August, 2001. Hamburg, Germany. Under the auspices of the International Union of Pure and Applied Physics (IUPAP)., p. 1768. — Vol. 5. — 2001. — P. 1768.
11. *Pierog Tanguy.* Air shower simulation with a new generation of post-LHC hadronic interaction models in CORSIKA // *Proc. Sci.* — 2017. — Vol. 301. — P. 100.
12. *Wigmans Richard.* On big bang relics, the neutrino mass and the spectrum of cosmic rays // *Nuclear Physics B-Proceedings Supplements*. — 2000. — Vol. 85, no. 1-3. — Pp. 305–310.
13. *Stenkin Yuri V.* Does the "knee" in primary cosmic ray spectrum exist? // *Modern Physics Letters A*. — 2003. — Vol. 18, no. 18. — Pp. 1225–1234.
14. *Stenkin Yuri V.* On the PRISMA project // *Nucl. Phys. B Proc. Suppl.* — 2009. — Vol. 196, no. 293.
15. Thermal neutron flux produced by EAS at various altitudes / Yu V Stenkin, VV Alekseenko, DM Gromushkin et al. // *Chinese physics C*. — 2013. — Vol. 37, no. 1. — P. 015001.
16. *Collaboration Lhaaso et al.* The LHAASO experiment: from gamma-ray astronomy to cosmic rays, Nucl // *Part. Phys. Proc.* — 2016. — Vol. 279, no. 281. — P. 166.
17. NEVOD—An experimental complex for multi-component investigations of cosmic rays and their interactions in the energy range 1–10¹⁰ GeV / II Yashin, MB Amelchakov, II Astapov et al. // *Journal of Instrumentation*. — 2021. — Vol. 16, no. 08. — P. T08014.
18. Регистрация потока тепловых нейтронов вблизи поверхности Земли / ДМ Громушкин, ВВ Алексеенко, АА Петрухин et al. // *Известия*

- Российской академии наук. Серия физическая.* — 2009. — Vol. 73, no. 3. — Pp. 425–427.
19. Долговременные вариации природного потока тепловых нейтронов на высоте 4300 м над уровнем моря / ЮВ Стенькин, ВВ Алексеенко, АС Багрова et al. // *Известия Российской академии наук. Серия физическая.* — 2017. — Vol. 81, no. 2. — Pp. 179–180.
 20. He Huihai, Collaboration Lhaaso. Design of the LHAASO detectors // *Radiation Detection Technology and Methods.* — 2018. — Vol. 2, no. 1. — P. 7.
 21. Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ -ray Galactic sources / Zhen Cao, FA Aharonian, Q An et al. // *Nature.* — 2021. — Vol. 594, no. 7861. — Pp. 33–36.
 22. Performances of ENDA-INR prototype array / O Shchegolev, V Alekseenko, D Kuleshov et al. // *Journal of Physics: Conference Series* / IOP Publishing. — Vol. 1690. — 2020. — P. 012011.
 23. Detection of thermal neutrons with the PRISMA-YBJ array in extensive air showers selected by the ARGO-YBJ experiment / Bruno Bartoli, Paolo Bernardini, XJ Bi et al. // *Astroparticle physics.* — 2016. — Vol. 81. — Pp. 49–60.
 24. Ranucci Gioacchino. An analytical approach to the evaluation of the pulse shape discrimination properties of scintillators // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment.* — 1995. — Vol. 354, no. 2-3. — Pp. 389–399.
 25. Study of the thermal neutron detector ZnS (Ag)/LiF response using digital pulse processing / F Pino, Luca Stevanato, Davide Cester et al. // *Journal of Instrumentation.* — 2015. — Vol. 10, no. 08. — P. T08005.
 26. An algorithm for charge-integration, pulse-shape discrimination and estimation of neutron/photon misclassification in organic scintillators / JK Polack, M Flaska, A Enqvist et al. // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment.* — 2015. — Vol. 795. — Pp. 253–267.

27. Machine learning n/γ discrimination in CLYC scintillators / E Doucet, T Brown, P Chowdhury et al. // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. — 2020. — Vol. 954. — P. 161201.
28. Machine learning for digital pulse shape discrimination / TS Sanderson, CD Scott, M Flaska et al. // 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC) / IEEE. — 2012. — Pp. 199–202.
29. Pulse shape discrimination and exploration of scintillation signals using convolutional neural networks / Jack Griffiths, Steven Kleinegesse, D Saunders et al. // *Machine Learning: Science and Technology*. — 2020. — Vol. 1, no. 4. — P. 045022.
30. *Cleveland William S, Devlin Susan J*. Locally weighted regression: an approach to regression analysis by local fitting // *Journal of the American statistical association*. — 1988. — Vol. 83, no. 403. — Pp. 596–610.
31. *Refaeilzadeh Payam, Tang Lei, Liu Huan*. Cross-validation // *Encyclopedia of database systems*. — Springer, 2009. — Pp. 532–538.
32. Pytorch: An imperative style, high-performance deep learning library / Adam Paszke, Sam Gross, Francisco Massa et al. // *Advances in neural information processing systems*. — 2019. — Vol. 32.
33. Sub-center arcface: Boosting face recognition by large-scale noisy web faces / Jiankang Deng, Jia Guo, Tongliang Liu et al. // *European Conference on Computer Vision* / Springer. — 2020. — Pp. 741–757.
34. *Maaten Laurens van der, Hinton Geoffrey*. Visualizing data using t-SNE // *Journal of machine learning research*. — 2008. — Vol. 9, no. Nov. — Pp. 2579–2605.
35. *Kinga Diederik, Adam Jimmy Ba et al*. A method for stochastic optimization // *International conference on learning representations (ICLR)* / California;. — Vol. 5. — 2015.

36. Geant4—a simulation toolkit / Sea Agostinelli, John Allison, K al Amako et al. // *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. — 2003. — Vol. 506, no. 3. — Pp. 250–303.
37. CORSIKA: A Monte Carlo code to simulate extensive air showers / Dieter Heck, Johannes Knapp, JN Capdevielle et al. // *Report fzka*. — 1998. — Vol. 6019, no. 11.
38. *Fesefeldt HS*. GHEISHA program // *Pitha-85-02, Aachen*. — 1985.
39. *Kalmykov NN, Ostapchenko SS, Pavlov AI*. Quark-gluon-string model and EAS simulation problems at ultra-high energies // *Nuclear Physics B-Proceedings Supplements*. — 1997. — Vol. 52, no. 3. — Pp. 17–28.
40. EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider / Tanguy Pierog, Iu Karpenko, Judith Maria Katzy et al. // *Physical Review C*. — 2015. — Vol. 92, no. 3. — P. 034906.
41. sibyll: An event generator for simulation of high energy cosmic ray cascades / RS Fletcher, TK Gaisser, Paolo Lipari, Todor Stanev // *Physical Review D*. — 1994. — Vol. 50, no. 9. — P. 5710.
42. *Nieminen M, Torsti JJ, Valtonen E*. Analysis of neutron-multiplicity frequencies observed in a double monitor // *Physical Review D*. — 1982. — Vol. 26, no. 5. — P. 1036.
43. *Leśniak Damian, Sieradzki Igor, Podolak Igor*. Distribution-interpolation trade off in generative models // International Conference on Learning Representations. — 2019.
44. *Holland Paul W, Welsch Roy E*. Robust regression using iteratively reweighted least-squares // *Communications in Statistics-theory and Methods*. — 1977. — Vol. 6, no. 9. — Pp. 813–827.
45. *Kamata Koichi, Nishimura Jun*. The lateral and the angular structure functions of electron showers // *Progress of Theoretical Physics Supplement*. — 1958. — Vol. 6. — Pp. 93–155.

46. Lateral distribution of charged particles in giant air showers above 1 EeV observed by AGASA / S Yoshida, N Hayashida, K Honda et al. // *Journal of Physics G: Nuclear and Particle Physics*. — 1994. — Vol. 20, no. 4. — P. 651.
47. The lateral distribution of extensive air showers produced by cosmic rays above 1019 eV as measured by water-Čerenkov detectors / RN Coy, G Cunningham, CL Pryke, AA Watson // *Astroparticle Physics*. — 1997. — Vol. 6, no. 3-4. — Pp. 263–270.
48. *Lagutin AA, Pljasheshnikov AV, Uchaikin VV*. The radial distribution of electromagnetic cascade particles in the air // 16th International Cosmic Ray Conference, Vol. 7. Published by the Institute for Cosmic Ray Research, University of Tokyo, 3-2-1, Midori-cho Tanashi, Tokyo 18 JAPAN Edited by Saburo Miyake. International Union of Pure and Applied Physics.; Nihon Gakujutsu Kaigi.; Nihon Butsuri Gakkai. LCCN: 80-502341, p. 18. — Vol. 7. — 1979. — P. 18.
49. Full Monte-Carlo description of the Moscow State University extensive air shower experiment / Yu A Fomin, NN Kalmykov, IS Karpikov et al. // *Journal of Instrumentation*. — 2016. — Vol. 11, no. 08. — P. T08005.
50. *Nelder John A, Mead Roger*. A simplex method for function minimization // *The computer journal*. — 1965. — Vol. 7, no. 4. — Pp. 308–313.
51. *Chib Siddhartha, Greenberg Edward*. Understanding the metropolis-hastings algorithm // *The american statistician*. — 1995. — Vol. 49, no. 4. — Pp. 327–335.
52. *Powell Michael JD*. A fast algorithm for nonlinearly constrained optimization calculations // Numerical Analysis: Proceedings of the Biennial Conference Held at Dundee, June 28–July 1, 1977 / Springer. — 2006. — Pp. 144–157.
53. *Erdmann Martin, Glombitza Jonas, Walz David*. A deep learning-based reconstruction of cosmic ray-induced air showers // *Astroparticle Physics*. — 2018. — Vol. 97. — Pp. 46–53.
54. Using deep learning to enhance event geometry reconstruction for the telescope array surface detector / Dmitri Ivanov, Oleg E Kalashev, M Yu Kuznetsov

- et al. // *Machine Learning: Science and Technology*. — 2020. — Vol. 2, no. 1. — P. 015006.
55. *Shchegolev O.* Deep Learning Techniques Applications for the ENDA Experiment Data Analysis // *Physics of Atomic Nuclei*. — 2021. — Vol. 84, no. 6. — Pp. 915–918.
 56. *Mayer H.J.* A neural network algorithm for core location analysis at large extended air shower arrays // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. — 1992. — Vol. 317, no. 1-2. — Pp. 339–345.
 57. *Rigatti Steven J.* Random forest // *Journal of insurance medicine*. — 2017. — Vol. 47, no. 1. — Pp. 31–39.
 58. *Friedman Jerome H.* Greedy function approximation: a gradient boosting machine // *Annals of statistics*. — 2001. — Pp. 1189–1232.
 59. *Колмогоров Андрей Николаевич.* О представлении непрерывных функций нескольких переменных в виде суперпозиций непрерывных функций одного переменного и сложения // Доклады Академии наук / Российская академия наук. — Vol. 114. — 1957. — Pp. 953–956.
 60. *Kramer Oliver.* Scikit-learn // Machine learning for evolution strategies. — Springer, 2016. — Pp. 45–53.
 61. CatBoost: unbiased boosting with categorical features / Liudmila Prokhorenko, Gleb Gusev, Aleksandr Vorobev et al. // *Advances in neural information processing systems*. — 2018. — Vol. 31.
 62. Energy spectrum of primary cosmic rays between 1014.5 and 1018 eV / M Nagano, T Hara, Y Hatano et al. // *Journal of Physics G: Nuclear Physics*. — 1984. — Vol. 10, no. 9. — P. 1295.
 63. Энергетический спектр космических лучей из анализа электронной, мюонной и черенковской компонент широких атмосферных ливней / ЕА Вишневская, ВН Калмыков, НН Калмыков et al. // *Известия*

- Российской академии наук. Серия физическая.* — 2002. — Vol. 66, no. 11. — Pp. 1567–1570.
64. Primary cosmic ray energy spectrum above 1 PeV as measured by the PRISMA-YBJ array / Y Stenkin, VV Alekseenko, OB Shchegolev et al. // 35th International Cosmic Ray Conference (ICRC2017). — Vol. 301. — 2017. — P. 488.
 65. Measurement of Primary Energy of Air Showers in the Presence of Fluctuations. / AM Hillas, DJ Marsden, JD Hollows, HW Hunter // pp 1001-6 of 12th International Conference on Cosmic Rays. Vol. 3. Hobart, Australia International Union of Pure and Applied Physics (1971). / Univ. of Leeds, Eng. — 1971.
 66. Energy spectrum and mass composition of cosmic rays, by the data of the tunka-133 array / SF Berezhnev, NM Budnev, OA Gress et al. // *Bulletin of the Russian Academy of Sciences: Physics.* — 2015. — Vol. 79, no. 3. — Pp. 344–347.
 67. *Nguyen Anh, Yosinski Jason, Clune Jeff.* Deep neural networks are easily fooled: High confidence predictions for unrecognizable images // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2015. — Pp. 427–436.
 68. *Gal Yarin.* Uncertainty in deep learning: Ph.D. thesis. — 2016.
 69. *Malinin Andrey.* Uncertainty estimation in deep learning with application to spoken language assessment: Ph.D. thesis. — 2019.
 70. A convolutional neural network based cascade reconstruction for the IceCube Neutrino Observatory / R Abbasi, Markus Ackermann, Jenni Adams et al. // *Journal of Instrumentation.* — 2021. — Vol. 16, no. 07. — P. P07041.
 71. *Díaz Esteban, Spagnoli Giovanni.* Natural gradient boosting for probabilistic prediction of soaked CBR values using an explainable artificial intelligence approach // *Buildings.* — 2024. — Vol. 14, no. 2. — P. 352.
 72. *Malinin Andrey, Prokhorenkova Liudmila, Ustimenko Aleksei.* Uncertainty in gradient boosting via ensembles // *arXiv preprint arXiv:2006.10562.* — 2020.

73. *Alekseenko V. V. Bakatanov V. N. Djappuev D. D.* “Carpet” array: estimation of angular resolution using a Cherenkov radiation detector // *preprint/Inst. for Nucl. Res. Russ. Acad. Sci.* — 2003.
74. *Grieder Peter KF.* Extensive air showers: high energy phenomena and astrophysical aspects—a tutorial, reference manual and data book. — 2010.
75. Изучение временной структуры нейтронной компоненты ШАЛ / ДМ Громушкин, АА Петрухин, ЮВ Стенькин, ИИ Яшин // *Известия Российской академии наук. Серия физическая.* — 2011. — Vol. 75, no. 3. — Pp. 403–405.
76. *Stenkin Yuri V.* Large scintillator detector for thermal neutron recording // *Nuclear Track Detectors: Design, Methods and Applications ISBN.* — 2010. — Pp. 978–1.
77. Широкие атмосферные ливни космических лучей / Николай Алексеевич Добротин, Георгий Тимофеевич Зацепин, Иосиф Леонидович Розенталь et al. // *Успехи физических наук.* — 1953. — Vol. 49, no. 2. — Pp. 185–242.
78. *Kempa J.* Hadrons in extensive air showers. I: Theoretical analysis // *Il Nuovo Cimento A (1965-1970).* — 1976. — Vol. 31, no. 4. — Pp. 568–580.
79. Novel method for detecting the hadronic component of extensive air showers / ДМ Громушкин, VI Volchenko, АА Petrukhin et al. // *Physics of Atomic Nuclei.* — 2015. — Vol. 78, no. 3. — Pp. 349–352.
80. *Щеголев ОБ, Стенькин ЮВ.* Функция пространственного распределения электронов и тепловых нейтронов в широких атмосферных ливнях на уровне моря // *Краткие сообщения по физике Физического института им. ПН Лебедева Российской Академии Наук.* — 2016. — Vol. 43, no. 7. — Pp. 24–31.
81. *Gaisser Thomas K, Stanev Todor, Tilav Serap.* Cosmic ray energy spectrum from measurements of air showers // *Frontiers of Physics.* — 2013. — Vol. 8, no. 6. — Pp. 748–758.

82. Исследование нейтронной компоненты ШАЛ на установке PRISMA-32 / ДМ Громушкин, ФА Богданов, ЮВ Стенькин et al. // *Физика элементарных частиц и атомного ядра*. — 2018. — Vol. 49, no. 4. — Pp. 1056–1063.
83. Study of the EAS size spectrum in thermal neutrons measured by the PRISMA-YBJ array / Oleg Shchegolev, Victor Alekseenko, Anastasia Bagrova et al. // EPJ Web of Conferences / EDP Sciences. — Vol. 145. — 2017. — P. 19021.
84. *Hörandel JR, Collaboration KASCADE et al.* Measurement of the Primary Cosmic Ray Energy Spectrum using the Hadronic Air Shower Component // Proceedings of the 27th International Cosmic Ray Conference. 07-15 August, 2001. Hamburg, Germany. Under the auspices of the International Union of Pure and Applied Physics (IUPAP)., p. 137. — Vol. 1. — 2001. — P. 137.