

Нейронные сети на основе физики для моделирования гравитационного поля

Автор:

Дж. Р. Мартин

Бакалавр физики и астрономии, Университет Северной Каролины в Чапел-Хилл, 2018 г.

Бакалавр музыки, Университет Северной Каролины в Чапел-Хилл, 2018 г.

Магистр аэрокосмических инженерных наук, Университет Колорадо в Боулдер, 2021 г.

Диссертация, представленная на рассмотрение

Факультета высшей школы

Университета Колорадо в частичном соответствии

требованиям, предъявляемым к степени

Доктор философии

Кафедра аэрокосмических инженерных наук

2023

Члены Комитета:

Hanspeter Schaub, Chair

Джей Макмахон

Дэниел Ширес

Захари Санберг

Роберто Фурфаро

Мартин Дж. Р. (доктор философии, аэрокосмические инженерные науки)

Нейронные сети на основе физики для моделирования гравитационного поля

Диссертация под руководством проф. Hanspeter Schaub

Гравитация является одним из самых вездесущих сил в астродинамики. Движение каждого планеты, астероид, космический корабль и неразрывно под влиянием гравитационных сил на объекты ближнего и Дальнего. Несмотря на это, универсальной модели этой силы не существует. Скорее, динамикам приходится выбирать между множеством различных моделей силы тяжести, каждая из которых имеет свои уникальные преимущества и недостатки.

Например, некоторые модели быстры в вычислениях, но им не хватает аналитической строгости; некоторые достигают высокой точности

но сопряжены с вычислительными ограничениями, а другие по-прежнему построены на внутренних допущениях или имеют ограниченную эксплуатационную достоверность. Для решения этих проблем в этом тезисе предлагается основанная на физике

Нейросетевая гравитационная модель, или PINN-GM, которая отвлекает внимание от аналитических подходов

в сторону моделей, основанных на данных. В частности, PINN-GM использует последние достижения в области

научного машинного обучения, чтобы объединить возможности нейронных сетей с динамическими системами

теория для создания высокоточных моделей сложных динамических систем без ущерба для аналитических

строгость. Благодаря многочисленным итерациям разработки PINN-GM теперь обеспечивает высокую точность и быстроту.

время выполнения, эффективность обработки данных, глобальную достоверность и точную дифференцируемость. В совокупности эти

характеристики делают PINN-GM хорошо подходящим для оказания помощи астродинамикам в различных областях применения

включая обучение с подкреплением, периодическое обнаружение орбиты и определение орбиты.

Посвящаю себя

Моим учителям.

Благодарности

Говорят, для этого нужна целая деревня...

Что бы я ни говорил, здесь просто слишком много людей, которых нужно поблагодарить, и одной страницы недостаточно.

этого недостаточно. Спасибо доктору шоб за предоставленную мне возможность, чтобы перевернуть его в новое поле. Спасибо

мой комитета по многих разговорах, как профессиональный, так и Личный на мое путешествие

научные круги. Спасибо моему брату, который научил меня лазать (мы уже добрались до вершины?).

Спасибо всем сотрудникам AVS, новым и старым, за воспоминания, советы и пончики. Спасибо семье

Берроуз и Мартин за празднование взлетов и поддержку падений. Небольшая благодарность

моей кошке Кэлли, чье неприличное мяуканье побуждает меня ходить в лабораторию каждое утро. Большинство

главное, благодаря Мелисса, - за все.

Содержание

Глава

1 Введение	1
1.1 Значение	6
1.2 История гравитационных моделей	8
1.3 Гравитационные Миссии	11
1.4 Цель диссертации	11
2 Обзор задачи гравитационного моделирования	14
2.1 Обзор	14
2.2 Аналитические модели	14
2.2.1 Сферические Гармоники	14
2.2.2 Эллипсоидальные Гармоники	16
2.2.3 Внутренние Сферические Гармоники / Модели Бесселя	17
2.2.4 Масконы	17
2.2.5 Многогранная модель	17
2.2.6 Гетерогенный Многогранник	19
2.3 Модели машинного обучения	19
2.3.1 Нейронные сети	19
2.3.2 Экстремальные Обучающие Машины	21
2.3.3 Гауссовские процессы	22

2.3.4	Геодезическая сеть	23
2.3.5	Нейронные сети, основанные на физике	23
	3 Основанные на физике Нейросетевые модели Гравитации	25
3.1	Поколение I	25
3.1.1	Сетевая архитектура и гиперпараметры	28
3.1.2	Сферические Гармонические характеристики	30
3.1.3	Репрезентативная Компактность	35
3.1.4	Обобщение	40
3.1.5	Скорость вычислений	45
3.1.6	Производительность сети Применительно к гравитационному полю Луны	46
3.2	Поколение II	54
3.2.1	Нормализация	55
3.2.2	Разработка характеристик	55
3.2.3	Дополнительные ограничения PINN	56
3.2.4	Измененные Сетевые Архитектуры	57
3.2.5	Производительность	59
3.3	Поколение III	72
3.3.1	Усилия по предварительной обработке	72
3.3.2	Модифицированная функция потерь для учета выборок с большой высотой	75
3.3.3	Улучшайте числовые показатели, изучая Косвенный показатель Потенциального	76
3.3.4	Выполнение граничных условий для избежания ошибки экстраполяции	80
3.3.5	Использование ранее существовавшей информации о силе тяжести в решении PINN-GM	83
3.4	Производительность PINN-GM-III: Сравнительное исследование	84
3.4.1	Поколение I Против поколения III	85
3.4.2	Поколение II против поколения III	88
3.5	Характеристики PINN-GM-III: Астероид неоднородной плотности	88

4 Приложение I: обучение с подкреплением	94
4.1 Формулировка Марковского процесса Принятия решений	97
4.2 Критик мягких Актеров	98
4.3 Окружающая среда	100
4.4 Эксперимент	102
4.5 Результаты	103
5 Приложение II: Периодическое обнаружение орбиты	108
5.1 Предыстория	109
5.2 Методология	111
5.2.1 Автоматическая Дифференциация	114
5.2.2 Характеристика PINN-GM	115
5.3 Метод декартовой съемки ПИННА-Г.М.	118
5.4 Способ съемки элемента орбиты PINN-GM	121
5.5 Метод съемки элемента с ограниченной орбитой PINN-GM	123
5.6 Дополнительные Начальные условия	126
6 Приложение III: Оценка и фильтрация гравитационного поля	128
6.1 Обзор	128
6.2 Автономная оценка	130
6.3 Онлайн-оценка	136
6.3.1 Фильтр Калмана	137
6.3.2 Компенсация Динамической модели	139
6.3.3 Фильтр Калмана PINN-GM	139
6.4 Установка проблемы	142
6.5 Показатели	145
6.6 Эксперименты	147
6.6.1 Поиск по гиперпараметру	147

6.6.2 Чувствительность к геометрии орбиты	153
7 Выводов	156
Библиография	160
Приложение	
Подробности обучения PINN-GM-III	172
В Скорость обучения Алгоритму отжига	174
С Предотвращение спектрального смещения с помощью отображения признаков Фурье	178
D Комментарии о прошлых результатах машинного обучения	181
Алгоритм E Pines	183
E.1 Введение	183
Состав E.2 Pines	186
E.2.1 Гравитационный потенциал	186
E.2.2 Гравитационное ускорение	187
193 Алгоритм	
F GPU Программное обеспечение F.1 GPGPU	193
F.2 Аппаратное обеспечение GPGPU	194
F.3 Общие стратегии оптимизации GPGPU	194
Основные Процедуры составления рецептур F.4 Pines	196
F.4.1 Основная процедура 1: Сокращение объема данных	196
F.4.2 Основная процедура 2: Матрица Лежандра	197
F.4.3 Ограничение памяти	199

F4.4 Граница вычисления	200
F5 Масштабируемость до созвездий	201
F6 Контрольные показатели	202
F7 Заключение	205

Таблицы

Таблица

3.1	Общие гиперпараметры для традиционных и основанных на физике нейронных сетей	
	обучен в этой работе	30
3.2	Уникальные гиперпараметры для традиционных и основанных на физике нейронных сетей	
	обучен в этой работе	31
3.3	Номинальные Гиперпараметры	62
3.4	Статистика гравитационной модели машинного обучения - См. Приложение D	72
5.1	Начальное распределение элементов орбиты	116
5.2	Начальные условия и решения	125
6.1	Начальные элементы орбиты	145
6.2	Начальное состояние, ковариация, матрица шума процесса и матрица шума измерения	146
6.3	Орбиты для определения метрики траектории	146
6.4	Таблица показателей для стандартных гравитационных моделей.	147
6.5	Гиперпараметры	148
A.1	Гиперпараметры по умолчанию для PINN-GM-III	173
C.1	Гиперпараметрический поиск для отображения объектов Фурье.	179
F1	Буферы, передаваемые графическому процессору на протяжении всего срока службы алгоритма	200

Рисунки

Рисунок

1.1	Гравитационные возмущения Земли и Луны	3
1.2	Использование научных моделей машинного обучения для решения задачи гравитационного моделирования. . .	4
1.3	Астероиды часто имеют сложную геометрию, приводящую к нерегулярным гравитационным полям.	7
 2.1 Вычислите время и конечную ошибку, связанную с моделированием космического аппарата на расстоянии 600 км высота орбиты в течение четырех часов в реальном времени, используя сosen выработки как функции сферических гармоник степени.		16
3.1	PINN-GM Поколения I	26
3.2	Карта да на земной сфере Бриллюэна	31
3.3	Карта да Земли на высоте LEO (приблизительно 420 км)	32
3.4	График MRSE(), MRSE(Δ) и MRSE () как функции суммарных параметров, p , в использовалась сферическая гармоническая гравитационная модель, где $p = l(l+1)$	33
3.5	График MRSE как функции общих параметров модели, стр. Сплошные линии представляют представление сферической гармоники. Пунктирные линии представляют традиционную нейронную сети. Линии с круговыми маркерами представляют нейронные сети, основанные на физике. 36	36

3.6	Подмножество промежуточных преобразований и результирующая базисная функция для $N = 40$	
традиционная нейронная сеть. Каждая строка соответствует одному слою сети		
в порядке ввода (верхний) и вывода (нижний). Отдельные графики представляют собой		
нормализованный и безразмерный вывод функции активации конкретного узла, когда		
оценивается по всей земной сфере Бриллюэна 38		
3.7	Увеличенный последний слой рисунка 3.6, представляющий предсказанные декартовы компоненты	
векторов ускорения, нанесенных на сферу Бриллюэна 38		
3.8	Гравитационная модель Земли с использованием (а) представления нейронной сети и (б) представления сферической гармоники	
при примерно одинаковом количестве свободных па-		
раметров ($r \approx 3,000$) 39		
3.9	MRSE из	A_m (сверху) и F_m (нижняя) для традиционных (штриховая линия) и физической-
прошлого года (сплошная) нейронные сети преобразуются в эквивалентный сферический гармонический		
степень как функция высоты. Синяя гистограмма представляет данные обучения		
распределение 42		
3.10	обучение распределения данных и эквивалентных сферических гармоник степени на различных АЛТИ-	
установки для $\beta = 3$ и $\beta = 10$ наборов данных. Сплошные линии представляют PINNs и накаты		
линии представляют традиционные нейронные сети 43		
3.11	Общее время оценки для оценки 10 000 случайных данных с использованием различных гравитационных моделей	45
3.12	Контрастное распределение гравитационного поля и ускорений Земли и Луны 47	
3.13	График MRSE как функции общих параметров модели для Луны. Пунктирные линии	
представляют традиционные нейронные сети. Линии с круговыми маркерами представляют собой		
нейронные сети, основанные на физике 49		
3.14	Гравитационная модель Луны с использованием (а) полной сферической гармонической модели $l = 1000$, (б)	
представление PINNA с $p = 3,040$ и (с) низкая точность воспроизведения $l = 55 \Rightarrow p = 3080$		
сферическое гармоническое представление 50		

3.15 MRSE из A_m (сверху) и F_m (нижняя) для традиционных (штриховая линия) и физической- прошлого года (сплошная) нейронные сети преобразуется в эквивалентный сферический гармонический степень как функция высоты для Луны. Синяя гистограмма представляет	
распределение обучающих данных	53
3.16 PINN-GM Поколения II (1)	54
3.17 Измененная Структура Сети	58
3.18 Тренировочные распределения Астероид 433-Эрос	60
3.19 Остатки ускорения, $ a - a^* / a \times 100$, их скользящие средние и обучающие данные распределение в зависимости от радиуса. Синие точечные диаграммы соответствуют ошибке PINN обучен на идеальных измерениях, а зеленый соответствует PINN, обученному на зашумленных измерениях. Пунктирные линии представляют погрешность сферической гармоники. Модели соответствуют тем же данным. Серая гистограмма представляет радиальное распределение. из обучающих данных.	63
3.20 Распределение данных Eros	65
3.21 ускорение остатков функция распределения обучение с $J_{\text{одо}}$	показать
как дополнительные ограничения физики помогают уменьшить чувствительность модели к шуму в обучении сведения	66
3.22 Ошибка модели за пределами сферы Бриллюэна (снаружи) как функция количества обучающих данных- и шума, добавленного к обучающим данным.	68
3.23 Ошибка модели внутри сферы Бриллюэна (внутренняя часть) в зависимости от объема обучения данные и шум добавлены к обучающим данным.	69
3.24 Ошибка модели на поверхности астероида (surface) в зависимости от количества обучающих данных и шума, добавленного к обучающим данным.	70
3.25 PINN-GM поколения III с новыми модификациями, помещенными в темно-серые коробки ..	73

3.26 РАЗЛИЧНЫЕ ФУНКЦИИ ПОТЕРЬ ИЗМЕНЯЮТ ПРОИЗВОДИТЕЛЬНОСТЬ СЕТИ НА БОЛЬШИХ И МАЛЫХ ВЫСОТАХ. Синий	
точки - это отдельные ошибки тестовых данных, синяя гистограмма - распределение	
данных обучения, а серая линия представляет собой среднюю ошибку тестирования в пределах скользящего окна	
в 100 баллов.....	77
3.27 ИСТИННЫЙ ПОТЕНЦИАЛ U (слева) БЫСТРО УМЕНЬШАЕТСЯ ДО ЧИСЛЕННО НЕБЛАГОПРИЯТНЫХ ЗНАЧЕНИЙ, ТОГДА КАК	
по сравнению с косвенным потенциалом $U_{N,N}$ (справа) численно остается хорошо обусловленным между	
[1, 1].....	79
3.28 ВЛИЯНИЕ ОБУЧЕНИЯ ПОТЕНЦИАЛЬНОГО ПРОКСИ-СЕРВЕРА НА ОШИБКУ СЕТЕВОГО ВЫВОДА.	79
3.29 верхний ряд: процент погрешности PINN-GM внутри (0-3P) и снаружи (3P-10p) из	
домен обучения. Серая вертикальная линия радиусе Бриллюэна. Зеленая вертикальная линия - максимальная.-	
минимальные границы обучающих данных. Нижняя строка: % ошибка PINN-GM в плоскости XY... .	82
3.30 Средняя ошибка ускорения в процентах, определенная в уравнении (3.33) PINN-GM-I (вверху)	
и PINN-GM-III (внизу), прошедшее обучение на Земле.....	85
3.31 Средняя ошибка ускорения в процентах для обученных PINN-II (вверху) и PINN-III (внизу)	
об Эросе.....	87
3.32 Астероид с неоднородностью массы (вверху) и соответствующим полем тяжести на поверхности	
когда предполагается постоянная плотность в сравнении с истинным полем (внизу слева и внизу справа	
соответственно). ..	89
3.33 Ошибка многогранного ускорения постоянной плотности, указанная в виде $\mu \pm \sigma$ (макс.).	90
3.34 Ошибка ускорения PINN-GM-III указана как $\mu \pm \sigma$ (макс.).	90
3.35 Поверхностная погрешность полиэдральной гравитационной модели постоянной плотности и PINN-GM-III.	
3.36 Скорость распространения и погрешность многогранной модели постоянной плотности и PINN-	
GM-III.	92
4.1 Траектории, снятые без включенного агента расширенного безопасного режима. Более темные оттенки	
соответствуют более ранним участкам траектории.....	101
4.2 Средняя доходность в зависимости от времени на часах	104

4.3	Траектории, выбранные с включенным агентом расширенного безопасного режима.	105
4.4	Показатели успешности агентов, прошедших обучение в различных средах	107
5.1	Схема прежнего декартова метода съемки (вверху) и нового орбитального элемента метод съемки (внизу).	110
5.2	Процентная погрешность вектора ускорения для каждой гравитационной модели	116
5.3	Метрики, сравнивающие орбиты, генерируемые PINN-GM и многогранной гравитацией Модели.	117
5.4	Результаты декартового метода съемки. Верхняя строка: ошибка решения в размерности- упорядоченные координаты (тильда) и безразмерные координаты (без тильды). Середина. ряд: начиная орбите (серый) и найденного решения (цвет) распространяется на одного орбиты. Нижний ряд: решения распространяется на 10 орбит.	120
5.5	Результаты метода съемки орбитальных элементов. Верхний ряд: ошибка решения после одна орбита в размерных координатах (тильда) и безразмерных координатах (без тильды). Средний ряд: начальная орбита (серый) и обнаруженные решения (цветной) распространяется на одну орбиту. Нижний ряд: найденные решения распространялись на 10 орбит.	124
5.6	Решения орбиты, найденные методом декартовой съемки (синий), неограниченная съемка ОЕ- метод редактирования (красный) и метод съемки с ограниченным удалением (зеленый) по сравнению с исходной орбитой (серый / черный).	125
5.7	Средняя процентная ошибка решений с использованием декартова LPE и ОЕ LPE	126
6.1	Случайный выбор науки, БЛИЗКОЙ к Сапожной, вращается вокруг 433-Го Эроса	131
6.2	Ошибка регрессированной гравитационной модели Эроса ПИННА-ГМ без выброса	133
6.3	Ошибка регрессированной гравитационной модели Эроса ПИННА-ГМ с измерениями выбросов (Пурпурный)	134
6.4	Фильтр Калмана PINN-GM	140
6.5	Неоднородности астероида	143

6.6	Верхний ряд: оценены истинные ускорения гравитационного поля астероида неоднородной плотности вдоль декартовых плоскостей из $[2R, 2R]$. Нижние ряды: процентная погрешность ускорения для различные гравитационные модели ограничены погрешностью 10%.....	144
6.7	Траектория движения космического аппарата вокруг астероида	145
6.8	Гиперпараметрический поиск PINN-GM, инициализированный с постоянной плотностью poly- модель хедрала.....	151
6.10	Гиперпараметрический поиск PINN-GM, инициализированный моделью точечной массы.....	152
B.1	% Погрешность модели в зависимости от высоты с использованием различных комбинаций предварительной подготовки и адаптивные алгоритмы отжига со скоростью обучения.....	176
E. 2	Усовершенствованные гравитационные алгоритмы E.1 GPGPU для программного обеспечения для высокоточной астрономии обеспечивают- поддержку высокоточного распространения по орбите с меньшим временем выполнения.....	184
E. 2	вычислить время и окончательной ошибки, связанная с имитации ка при 600 км высота орбиты в течение четырех часов в реальном времени, используя сосен выработки как функции сферических гармоник степени.....	185
F1	Метод сокращения объема данных	197
F3	Потребляемая видеопамять в зависимости от I_{\max} . и количества космических аппаратов, Nc/c	202
F4	Поток данных при вызове ядра графического процессора	206

Глава 1

Введение

Астродинамика - это изучение движения объектов в пространстве. Будь то космический корабль или небесные тела

цель состоит в том, чтобы лучше понять лежащие в основе силы, действующие на эти системы и их характер.-

упорядочить или усилить результирующее движение. Несмотря на простоту всеобъемлющей темы, астродинамика

он полон разнообразных академических тем, охватывающих оптимизацию траектории, полет в формации, оценку состояния

, теорию управления, динамику ориентации и даже дистанционное зондирование. Среди широкого спектра

проблемы, одна вещь связывает их все вместе: сила тяжести.

Учитите тот факт, что почти все проблемы астродинамики начинаются с уравнений движения

вида:

$$\ddot{r} = -\nabla U(r) + \ddot{a} \quad (1.1)$$

где U - модель гравитационного потенциала, \ddot{a} - остальные ускорения. Значение

презентация этого дифференциального уравнения, только подчеркивает значимость гравитационных сил.

Почти во всех случаях, сила тяжести является доминирующей силой, а остальные возмущения

среднее. Однако, несмотря на важную роль, что гравитация играет во всех астродинамики проблемы, не универсальный

модель сила существует. Вместо этого, dynamicists осталось выбрать между множеством аналитических

или численные приближения, называемые гравитационными моделями, каждая из которых имеет свои соответствующие

преимущества и недостатки.

Простейшая гравитационная модель предполагает, что все объекты могут быть представлены в виде идеальной точки

массы, гравитационный потенциал которых принимает форму

$$U(r) = -\frac{GM}{r} \quad (1.2)$$

где μ - гравитационный параметр тела, а r - расстояние от тела. В то время как этот

выбор может быть достаточно для первого порядка анализа и доказательства правильности концепции работы, оно в значительной мере становится

менее надежны при переходе в сценарии реальной миссии. В этих условиях, выше точность модели

требуются, и специалисты по динамике вынуждены выбирать между множеством различных моделей, включая

сферические гармоники, но не ограничиваясь ими (2), эллипсоидальные гармоники (3), масконы (4), многогранные

модели (5), экстремальные обучающие машины (6), геодезические сети (7), и так далее.

У каждой из этих моделей есть компромиссы, но все они признают одно: гравитационные поля

намного сложнее, чем приближение точечной массы. Каждое небесное тело обладает ассилиацией.-

бозонами и гетерогенной плотности, или у поверхности усеяна кратерами, горы, минеральные

депозиты и прочие геологические особенности, что каждый вносит свою уникальную возмущение в

общее гравитационное поле. Достаточно взглянуть на печально текстурированные поверхности Земли и Луны

показано на рисунке 1.1 в качестве доказательства. Рассматривать эти тела как идеальные сферы или даже сплюснутые эллипсоиды,

это грубое упрощение истинной системы. Если dynamicists и космических аппаратов пренебрегать этими высшее-

заказ на гравитационные характеристики, последствия могут варьироваться от постоянного отклонения от исходных

траекторий до отсутствия входных коридоров или, в худших случаях, неожиданного перевода орбитальных аппаратов в режим

зондов-перехватчиков. Чтобы избежать этих сценариев, крайне важно, чтобы у специалистов по динамике были эффективные и

точные модели для представления этих гравитационных возмущений с высокой точностью.

Существует долгая история и богатый объем литературы, посвященные построению высоких-

точные модели гравитации. Большая часть этой литературы исследует способы представления гравитационных полей.

аналитически, используя хитроумные математические приемы и базовые функции для обеспечения сходимости.

Хотя эти модели предлагают убедительные физические интерпретации и желаемую математическую поддержку.-

во-первых, многие из них имеют эксплуатационные ограничения, допущения или накладные вычислительные расходы. Хотя

эти аналитические модели обеспечили значительный прогресс сообществу специалистов по моделированию гравитации, их

сохраняющиеся недостатки выявляют фундаментальный вопрос, стоящий перед сообществом разработчиков динамики: существуют ли

способы построения универсальной модели сложных динамических сил, свободного от этих недостатков?

Чтобы ответить на этот вопрос, нужно сначала признать, что модели динамики, как известно,

трудно построить. В идеале эти модели должны быть точными, компактными и эффективными с точки зрения вычислений,

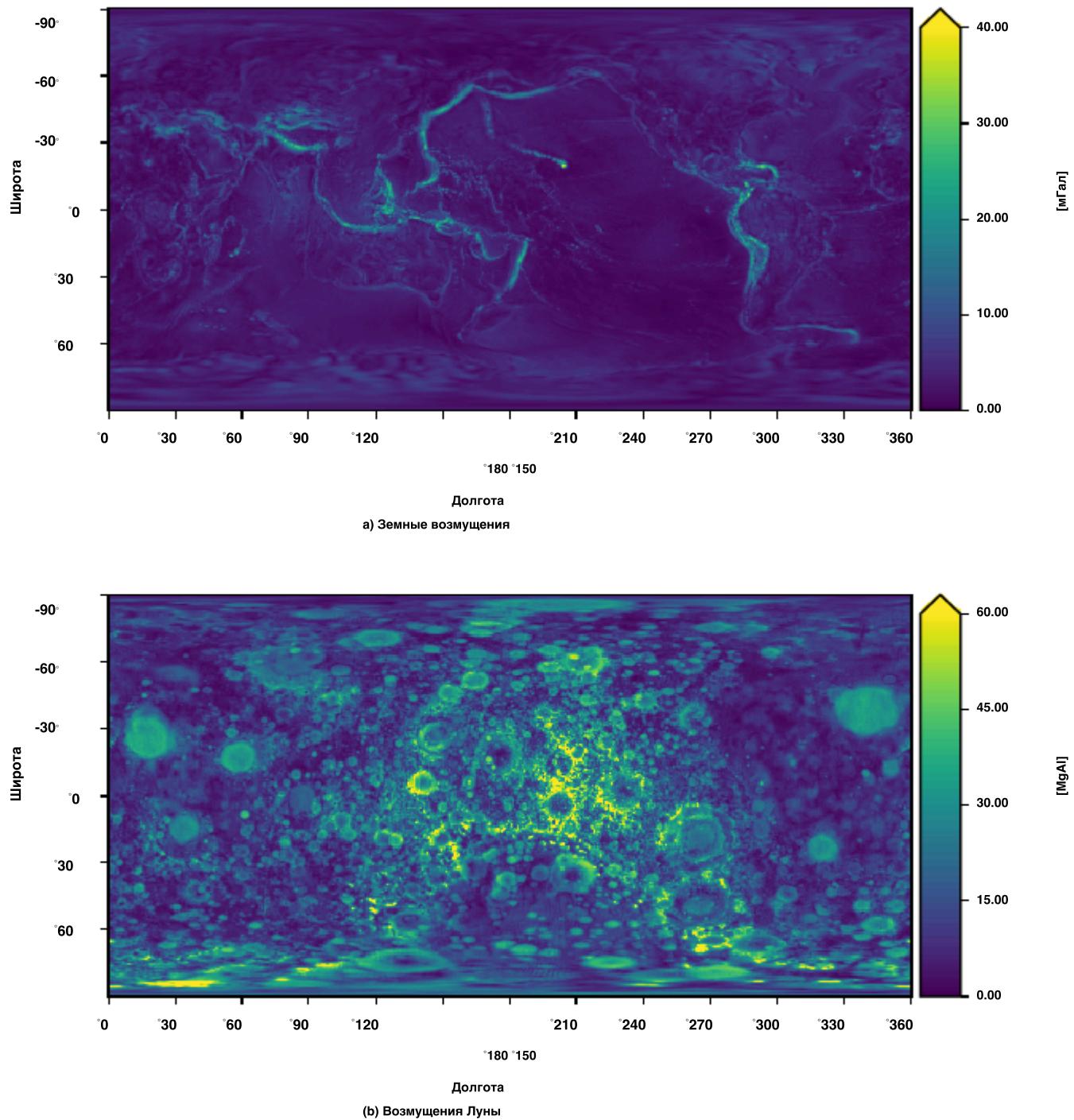


Рисунок 1.1: Гравитационные возмущения Земли и Луны

и физически совместимые; однако эти цели часто находятся в естественном соперничестве с одним-

Другое. Во многих случаях высокоточные модели отличаются множеством параметров, большим объемом памяти,

и высокими вычислительными затратами. Компактные модели требуют меньших накладных расходов, но, как правило, теряют точность.

или срезают углы в своей аналитической строгости. Остается открытым вопрос, существуют ли вообще модели, для

которых не нужно идти на эти компромиссы. Для случая моделирования гравитационного поля еще не создана

модель, которая была бы достаточно гибкой, чтобы представлять наиболее доминирующие возмущения с высокой точностью-

при сохранении небольшого объема памяти и быстрого времени выполнения. Если такая модель

если бы они существовали, их применение было бы далеко идущим - расширение бортовых возможностей космических аппаратов

наведение, навигация и управление, повышение точности наземного моделирования при проектировании траектории.

и планирование миссий, а также неизученные усовершенствования для планетологов и геофизиков

которые используют эти модели для других целей.

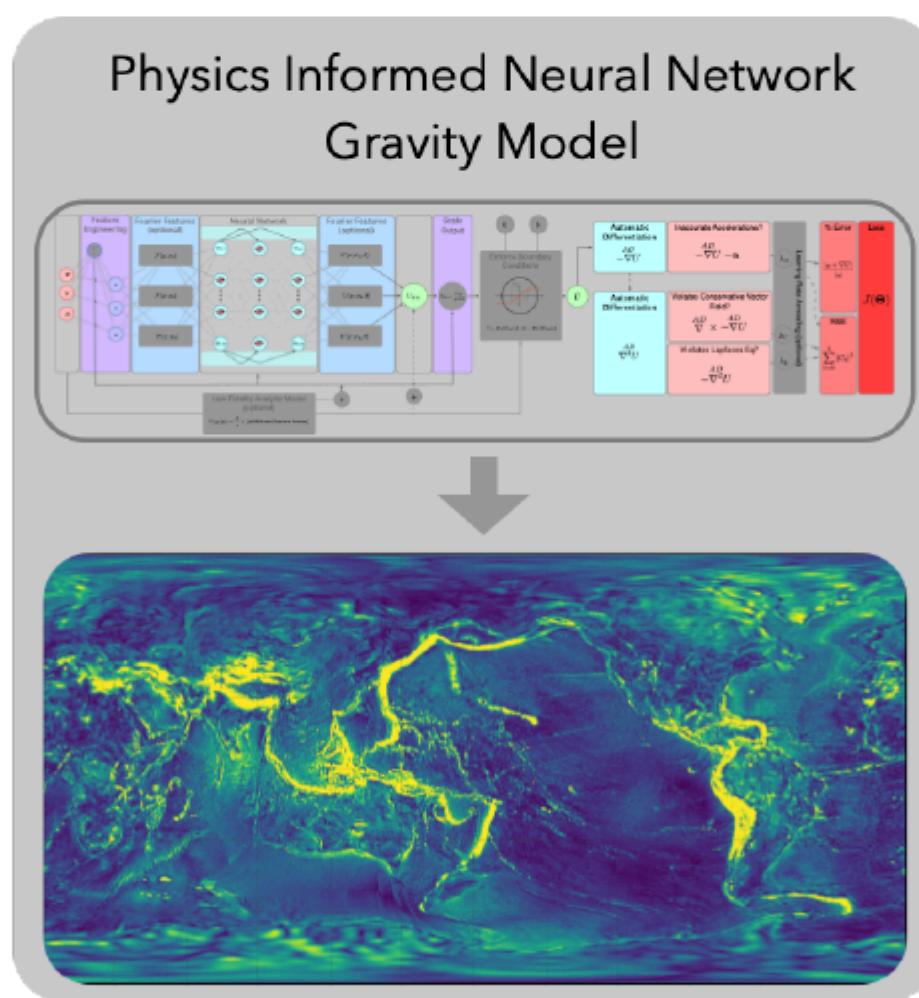


Рисунок 1.2: Использование научных моделей машинного обучения для решения задачи гравитационного моделирования.

Этот тезис направлен на то, чтобы сделать шаг к созданию универсального, высокоточного гравитационного устройства.

Модель. Путем смещения фокуса от чисто аналитических подходов, это работа, а не вводит

класс гравитационная модель, которая использует новые методы в новой области научных машина

Обучения, или SciML. SciML - это подмножество искусственного интеллекта, которое стремится создавать высокоточные модели сложных динамических систем на основе данных.

(8). Существует большое разнообразие

инструменты-кандидаты и аппроксиматоры функций в этой области машинного обучения, но один из самых популярных

Нейронная сеть на основе физики (PINN).

PINN - это нейронные сети, специально разработанные для изучения решений дифференциальных уравнений (9).

Хотя технически это может быть достигнуто с помощью традиционных нейронных сетей - выборка значений

динамической системы в различных положениях и моментах времени и регрессия нейронной сети

для этого привязки к данным обеспечивают более удовлетворяющую и мощную структуру. В частности, привязки к данным

вводят расширенную функцию потерь, которая одновременно наказывает за неправильное моделирование обучения

данных, а также включая дополнительные условия, которые ограничивают модели, когда он нарушает дифференциала

уравнения истинной системы. Это простое изменение дает удивительную сумму коммунальных,

представление dynamicists мощный и гибкий способ представления сложных динамических явлений при

убедившись, что узнал решение по сути жалобы с основной физикой. Не только

помогает ли это устраниТЬ опасения по поводу обоснованности изученного решения, но эти изменения также приводят к

значительным улучшениям точности модели, обобщаемости и эффективности обработки данных.

В этом тезисе исследуется, как нейронные сети, основанные на физике, могут быть использованы для решения задачи

моделирования силы тяжести. Вместо того, чтобы предписывать аналитические базисные функции, которые поставляются со своими

собственные уникальные проблемы, в последующих главах демонстрируется PINNs можете узнать более компактным

представления гравитационного потенциала при сохранении желательных физических свойств

их аналитические предшественников. После нескольких итераций разработки PINN-GM теперь производит-

гравитационные модели, которые обеспечивают высокую точность, быстрое время выполнения, эффективность обработки данных, глобальную достоверность,

и точную дифференцируемость. Взятые вместе, эти характеристики делают PINN-GM хорошо подходящим для

оказания помощи астродинамикам в различных приложениях, включая обучение с подкреплением, периодическую орбиту

обнаружение и определение орбиты.

1.1 Значение

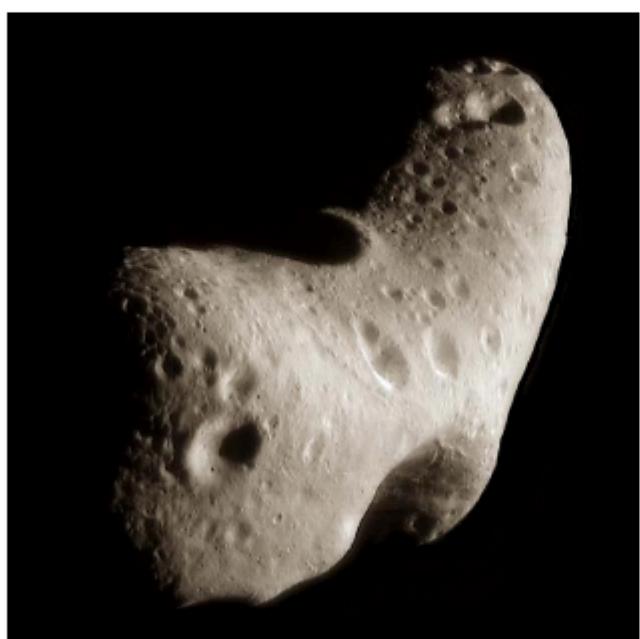
Как говорилось выше, гравитационное поле, моделирование является одним из наиболее фундаментальных проблем в astrodynamic. Динамика движения. Почти во всех космических динамических систем, гравитация играет ключевую роль в руководстве, навигации и управление рассматриваемым объектом. Рассмотрим наиболее распространенные проблемы в астрономии-динамика: задача о двух телах, задача о трех телах, полет строем, сближение, сближение операции и стыковка, расчет траектории и т.д. Для каждой из этих тем гравитационная модель может полностью изменить характер проектного пространства и найденных решений. Несмотря на это, динамики-специалисты часто полагаются на модель силы тяжести точечной массы или альтернативы с низкой точностью, поскольку переход к модели с более высокой точностью рискует привести к тому, что проблема достигнет неприемлемого уровня сложности.

Использование гравитационных моделей низкой точности, однако, становится неустойчивым. Поскольку околоземные, окололунные и дальние космические миссии продолжают усложняться, растет одновременный потребность в более мощных динамических моделях, помогающих космическим аппаратам в достижении их целей. Последние десятилетия, потраченные на изучение астероидов и комет, или малых тел, являются убедительным примером этого явления. Для самых ранних миссий 1980 - х годов , таких как Giotto (10) и Галилей (11), цель заключалась в простом облете небольшого объекта исследования (кометы Галлея и астероида Гаспра соответственно). Для этих миссий не требовалось чрезвычайно точные гравитационные модели, только достаточно точные чтобы сделать космический корабль в непосредственной близости от объекта для захвата изображения перед переходом в глубокий космос. Позже, эти миссии стали, требующих большей точности, так как дизайнеры стремились место космический аппарат на орбите вокруг этих тел (ВБЛИЗИ Шумейкера, 1996 г. (12)) или для проведения нескольких полетов (Deep Space 1, 1998 (13)). Как только было продемонстрировано, что космический корабль может выходить на орбиту, приоритет был смешен затем на сбор материала (Stardust, 1999 (14)), целенаправленно воздействующий на объект (Deep impact, 2005 (15)), даже приземляясь на поверхность тела (Розетта, 2004 (16); Hayabusa, 2003 (17)). Эти все более амбициозные цели продолжаются и по сей день в таких миссиях, как Hayabusa2(18) и ОСИРИС-Рекс (19) попытка вернуть образцы для астероидов Рюгу и Бенну соответственно. Нетрудно экстраполировать эти тенденции и признать, что миссии будут продолжать усложняться в обозримом будущем, и специалистам по динамике потребуется все более точная

динамические модели соответствуют этим амбициям. Из этих сил особенно важно точно отразить эффекты гравитационного поля астероида. Эти тела могут иметь экзотическую геометрию и профили плотности (см. Рисунок 1.3), который, если его не учитывать, может вывести космический корабль на траектории, которые могут столкнуться с телом. При избежании катастрофического разрушения наивысший приоритетные высокоточные гравитационные модели также могут быть полезны для точного приземления. Рассмотрим недавнюю миссию OSIRIS-REx, флагманскую фаза которой изначально предназначалась для направьте космический корабль в 25-метровую зону посадки. После обнаружения того, что астероид был покрыт валунами, самая большая из доступных посадочных площадок простиралась всего

(20). В

эти моменты, точное знание гравитационного поля имеет важное значение для обеспечения безопасности миссии в тачдаун.



(a) - 433-Эрос (21)



(b) 25413-Итокава (22)

Рисунок 1.3: Астероиды часто имеют сложную геометрию, приводящую к нерегулярным гравитационным полям.

Помимо применения гравитационных моделей к миссиям с малыми телами, существуют и другие академические сообщества, использующие высокоточные представления гравитационных полей. В частности, гравитационная модель-
eling сыграла ключевую роль в решении задач геофизики, океанологии и климатологии, а также планетологии-.
. Гравитационные модели обычно используются для лучшего понимания водных ресурсов по всему миру (23),

колебания уровня моря (24), и таяние ледников (25). Эти исследования играют решающую роль в прогнозировании

засух (26) и наводнения (27). Аналогичным образом, модели гравитационного поля используются для лучшего понимания

недр Земли (28) и движение тектонических плит (29). Последнее имеет решающее значение для

приложения в GNSS, которые, в свою очередь, помогают определять высоты, системы отсчета и координаты

системы, используемые в других дисциплинах (30).

В этой работе предлагается новая гравитационная модель, которая направлена на обеспечение полезности каждой из этих моделей.

области; однако уроки, извлеченные в результате разработки этой модели, выходят за рамки космоса

научное сообщество. В частности, эта работа фундаментально исследует, как машинное обучение может быть

используется для построения высокоточных динамических моделей; исследует, как лучше системно обнаруживать

базисные функции и координировать описания сложных динамических явлений. В принципе,

уроки, собранные здесь, могут принести пользу сообществам, которые стремятся изучить альтернативы, основанные на данных.

В противном случае громоздкие аналитические модели. Акцент на гравитационном моделировании предлагает продуктивный тест

стенд для научного сообщества машинного обучения в целом, поскольку эти системы часто инвариантны во времени-

и имеют хорошо изученные ранее существовавшие модели. Исходя из этого, практики SciML могут использовать это

проблема тестирования новых методов машинного обучения без введения дополнительных накладных расходов

для более сложных динамических систем.

1.2 История гравитационных моделей

Поиск высокоточных гравитационных моделей начался еще до появления космических аппаратов. В то время как забава-

законы тяготения были впервые охарактеризованы Ньютоном в конце 17 века, и только в

начале 1900-х годов ученые начали создавать модели более высокой точности. Рапп предоставляет подробную

история этих разработок для моделей земного притяжения в его обзорной статье (31), который является суммой-

маринован здесь для удобства. За период с 1901 по 1958 год ученые впервые признали, что

Поле тяготения Земли отличается от поля приближения точечной массы. Чтобы учесть это, они

сначала разработали гравитационные модели, которые представляли собой суперпозицию изменений силы тяжести по широте (зональных на-

турбулентности), а позже расширил эти модели, включив в них изменения долготы - предшественника

используемой сегодня сферической гармонической гравитационной модели. Для оценки параметров этих моделей,

геодезисты усреднили наземные измерения силы тяжести в ячейках фиксированного размера (30x30, 10x10, 5x5,

и 1x1) и регрессировали соответствующие коэффициенты.

С запуском спутника и других ранних спутников геодезисты, такие как Каула, впервые использовали спутниковые данные для оценки коэффициента select в 1960-х годах. Аналогично, такие учреждения, как Смитсоновский институт

Астрофизическая обсерватория, Лаборатория прикладной физики и Лаборатория военно-морского вооружения

рассчитали свои собственные гравитационные модели на основе спутниковых данных, называемых моделями SE, APL и NWL

соответственно. На протяжении 1960-х и в 1970-е годы были предложены дополнительные методы для объединения

наземных и спутниковых данных, позволяющие создавать модели более высокого уровня с большей точностью. В 1970-х годах,

Центр космических полетов имени Годдарда представил свою собственную гравитационную модель GEM, которая началась с GEM-I

в 1972 году (32), но продолжал обновляться до GEM-10C, который достиг 180 градусов в 1978 году. В

в 1980-х годах внедрение спутникового лазерного слежения обеспечило более высокое качество и больший объем

спутниковых данных для обновления этих моделей. Данные таких миссий, как Lageos и TOPEX / POSEIDON, были

используются с более поздними поколениями GEM для формирования таких моделей, как GEM-L2 (33) и GEM-T1 (34)

соответственно. Аналогичным образом, исследователи из Центра космических исследований Техасского университета в

Остин начал разрабатывать свои модели TEG, в которых использовались данные отслеживания с нескольких спутников.

В 1990-х годах GSFC и Центр космических исследований Калифорнийского университета в Остине объединили усилия и создали

Модели JGM 1-3 (35; 36). Совсем недавно Модели земного притяжения (EGM) от Национального агентства геопространственной разведки (National Geospatial Intelligence Agency)

представляют одни из самых точных моделей. Среди наиболее известных-

EGM96 (37), в котором достигнуто сферическое гармоническое решение, охватывающее 360 градусов. В 2008 году,

NGA представила модель EGM2008, которая достигла степени 2190 и в настоящее время является самой точной.

общедоступная сферическая гармоническая модель Земли (38).

Изучение гравитационного поля Луны шло по аналогичной тенденции, хотя и более медленными темпами из-за ограниченных данных. Одна из первых гравитационных моделей лунного поля появилась в результате программы "Аполлон"

. Изучая движение космического корабля, который вращался вокруг Луны, исследователи поняли

что орбитальное движение космического корабля было нарушено над определенной областью лунной поверхности. Эти

возмущения были позже охарактеризованы как концентрации массы на поверхности Луны, получившие название масконов,

которые можно эффективно смоделировать как суперпозицию нескольких точечных элементов массы, встроенных

на поверхности Луны или чуть ниже Нью (39). Анализ этих масконов продолжался до конца 20-го века с такими миссиями, как миссия Клементина (40), Лунный Разведывательный орбитальный аппарат (41), а также Лаборатория гравитационного восстановления и внутренних работ (GRAIL) (42) продолжает проводить измерения топологии Луны и гравитационного поля, чтобы помочь построить как более точные модели маскона, так и сферические гармонические модели.

Помимо изучения крупных планетных тел и спутников, исследование малых тел также привело к значительной работе по моделированию гравитационного поля. Как обсуждалось, гравитационные поля малых тел могут быть общеизвестно, что их сложно представить, особенно учитывая их неправильную геометрию и неизвестную плотность профиля. Первоначально исследователи продолжали моделировать эти поля со сферическими гармониками; однако, как отметил Бриллюэн в 1933 году, сферическая гармоническая модель начинает расходиться, когда космический корабль входит в сферу, которая ограничивает все элементы массы (43). Это расхождение делает сферическую гармонику модель, особенно рискованная при сближении или посадке.

В 1997 г., Вернер и Ширс представили многогранную гравитационную модель, которая предлагает нерасходящееся решение для представления гравитационных полей этих тел. (5). Если модель формы многогранника если рассматриваемое малое тело существует, то существует способ вычислить потенциал этого тела аналитически в предположении постоянной плотности. Такая модель предлагает решение, которое остается стабильна на поверхности, что делает ее популярным и выгодным выбором для многих миссий с небольшими объектами. Примечательно, что эта модель отличается относительно высокими вычислительными затратами и постоянной плотностью предположение оказалось неверным в некоторых случаях, как показано для астероида Бенну из последних данных OSIRIS-REx (44).

Недавно специалисты по динамике начали изучать альтернативные способы представления гравитационных полей этих экзотических тел, в меньшей степени полагаясь на аналитические приближения системы, а вместо этого на большее количество данных-модели, основанные на данных. Эти усилия включали использование гауссовых процессов (45), экстремальное обучение машины (6), и нейронные сети (46). Каждая из этих моделей обладает вычислительными преимуществами по сравнению с некоторыми из их аналитических аналогов, но им также не хватает аналитических тонкостей прошлых подходов.

На сегодняшний день ни одна из этих моделей не демонстрирует, что они удовлетворяют важным дифференциальным свойствам, и относительно небольшая работа была затрачена на тщательную проверку этих моделей. Во многих случаях изученные

решение тестируется только в локальных регионах или в пределах обучающих данных. Эти модели также может прийти с большой след памяти через десятки или сотни тысяч параметров, и многие требуют большого количества обучающих данных к регрессии, который может быть трудно получить в практика.

1.3 Гравитационные Миссии

Было осуществлено несколько миссий с целью уточнения моделей гравитационного поля для Земли и Луны. Для Земли первой из этих миссий была сложная полезная нагрузка миниспутника (CHAMP), запущенный в 2000 году компанией GFZ Potsdam. CHAMP предназначался для проведения различных атмосферных и ионосферных исследований в дополнение к улучшению характеристик силы тяжести и магнитного поля Земли. месторождения в течение пяти лет (47). Вслед за CHAMP в 2002 году был запущен эксперимент по восстановлению гравитации и климата (GRACE), в котором использовались два спутника-близнеца и их относительное расстояние. возможность производить более точные измерения гравитационного поля (48). Используя это расстояние, миссия смогла восстановить подробные измерения распределения воды (49), таяние ледникового покрова (25), и вносить свой вклад в распространение самых современных знаний об изменении климата (23). За GRACE последовал проект Gravity Field and Steady-State Ocean Circulation Explorer (GOCE), в котором использовался высокочувствительный градиентометр для оценки временных вариаций гравитационного поля (50). После GOCE и конца GRACE был запущен GRACE Follow On (GRACE-FO), который продолжил усилия оригинальной миссии. заключительная миссия GRACE по получению оценок гравитационного поля, изменяющихся во времени (51). Между этими миссиями, в 2011 году также была запущена Лаборатория восстановления гравитации и внутренних помещений (GRAIL), целью которой было измерить гравитационное поле Луны, используя концепцию миссии, аналогичную GRACE (52).

1.4 Цель диссертации

Этот тезис представляет собой первый всеобъемлющий документ, в котором обсуждается, как проектировать, обучать, и проверять основанные на физике модели гравитации нейросетей (PINN-GM). Очевидно, что этот тезис отвечает на вопросы, в том числе: Как дизайн PINN-GM влияет на его производительность? Как можно сконструировать входные функции, максимизирующие параметрическую емкость модели, избегая при этом

численная нестабильность? Есть ли преимущества в включении ранее существовавших моделей в систему,

и если да, то как этого можно достичь? Какую роль качество и распределение обучения

данные играют в производительности модели? Эти вопросы, среди прочих, исследуются с помощью

нескольких поколений разработок, описанных в следующих главах.

Более того, в этом тезисе также предлагается ряд новых способов характеристики гравитационной модели.

Производительность. Вместо случайной выборки контрольных точек и вычисления соответствующей ошибки ускорения-

В этом тезисе вводятся дополнительные меры для обеспечения более полного описания

характеристик модели. Эти меры включают оценку общей ошибки интегрирования во время траектории

распространение, тестирование точности модели в диапазоне высот в рамках обучения и за его пределами

данные, измерение компактности модели как зависимости точности от общих параметров и тестирование

модели с различным распределением данных и качеством. Эти показатели становятся полезными не только для

оценки производительности PINN-GM по сравнению с прошлыми моделями, но также обеспечивают основу

для руководства разработкой будущих поколений PINN-GM.

Благодаря этим исследованиям было показано, что PINN-GM является первой моделью, управляемой данными, которая

изучает новые мощные базовые функции сложных гравитационных систем, применяя дифференциальные

ограничения к изученному решению. Эти ограничения обеспечивают PINN-GM большую точность моделирования

и эффективность обработки данных, чем их предшественники, работающие с аналитикой и данными, при сохранении

небольшого объема памяти и быстрого времени выполнения. В совокупности эти атрибуты обеспечивают PINN-GM

большой полезности для сообщества астродинамика в большой, как показано через три кейса приме-

кации. Первый из этих приложений демонстрирует, как PINN-GM высокой точностью и быстро

оценка скорости может улучшить качество автономные агенты космических аппаратов, подготовленных через ре-

злоупотребление обучения. Второе приложение использует дифференцируемость PINN-GM знакомства

возможные периодические орбиты вокруг сложных гравитационных тел в произвольных наборах элементов. Третье

и последнее приложение исследует, как PINN-GMS может быть напрямую интегрирована в определение орбиты.-

конвейеры и фильтры для оценки сложных гравитационных полей как онлайн, так и офлайн-способом за

меньшее время, чем существующие подходы.

Подводя итог, нейросетевая гравитационная модель на основе физики представляет собой новый способ построения

высокоточные модели гравитационных полей как для больших, так и для малых небесных тел таким образом, чтобы не было многих подводных камней прежних подходов. В этом тезисе представлен обзор усовершенствований generational этой модели и освещены возможные приложения в рамках более широкой астродинамики.

Сообщество.

Глава 2

Обзор проблемы гравитационного моделирования

2.1 Обзор

Существуют два семейства гравитационных моделей: аналитические и численные. Аналитические модели принимают различные формы, но по своей сути их объединяет тот факт, что они основаны на первых принципах, предлагая решения фундаментальных дифференциальных уравнений, таких как уравнение Лапласа. В отличие от этого, численные модели строятся на основе данных - регрессионные функции, способные к интерполяции между измеренными данными. Каждое семейство моделей имеет свои соответствующие преимущества и недостатки, и выбор модели для использования часто определяется приложением. В следующей главе подробно описаны доступные гравитационные модели и их соответствующие плюсы и минусы.

2.2 Аналитические модели

2.2.1 Сферические гармоники

В начале 1900-х годов было предложено наложить базисные функции сферических гармоник для получения высокоточной оценки гравитационного потенциала (43).

$$U(r) = \frac{\mu}{r^l} \sum_{m=0}^l P_{l,m} (\sin \phi) C_{l,m} \cos(m\lambda) + S_{l,m} \sin(m\lambda) \quad (2.1)$$

Уравнение (2.1) называется сферической гармонической гравитационной моделью, где r - радиус до точки поля

, μ - гравитационный параметр тела, R - радиус, описывающий окружность тела,

l - степень модели, m - порядок модели, $C_{l,m}$

и $S_{l,m}$ - коэффициенты Стокса,

λ - долгота, ϕ - широта, и $P_{l,m}$

являются связанными полиномами Лежандра (2).

Сферическая гармоническая гравитационная модель является основной моделью для представления гравитации поля крупных планетарных тел, таких как Земля (53), Луна (54), и Марс (55). Одно из наиболее убедительных преимуществ сферической гармонической гравитационной модели заключается в том, что она позволяет компактно захватывать большое гравитационное возмущение сплюснутости планеты. Поскольку Земля и эти другие большие тела врачаются вокруг своих осей, возникает центробежное ускорение, которое сплющивает тело. из сферы в сплюснутый эллипсоид (56). Это уплощение или избыток массы вблизи экватора является называют планетарной сжатие, и ее присутствие имеет значительное влияние на траектории космических аппаратов.

Поэтому важно, что это шевеления записывается точно в гравитационной модели. В сферическую гармоническую гравитационную модель упрощает задачу, требуя только одной гармоники в разложении чтобы зафиксировать эту сплюснутость ($C_{2,0}$ или его альтернативная форма $J_2 = -C_{2,0}$).

В то время как сферические гармоники эффективны для представления сплющенности планеты и других возмущений глобального масштаба, они с трудом моделируют оставшиеся и более дискретные гравитационные возмущения. такие объекты, как горные хребты, границы тектонических плит и кратеры. Известно, что неоднородности сложно представить с помощью периодических базисных функций, часто для этого требуются сотни тысяч переменных.

Гармоники накладываются друг на друга, чтобы преодолеть трехмерный эквивалент феномена Гиббса (57). Необходимо не только наложить эти гармоники, но и уменьшить их, и высокочастотные сигналы будут обнаруживать все труднее из-за $(R/r)^l$ член в уравнении (2.1). Взятые вместе, эти результаты в условиях неэффективных моделей памяти, которые становятся трудно держать внутреннее пространство-ремесло. Кроме того, эти высококачественные модели поставляются с высоким вычислительным затратам. Высокая степень сферических гармоник модели требуют пересмотра всех соответствующих полиномов Лежандра на каждом точке поля. Данный расчет является рекурсивным и, следовательно, не банальный способ распараллелить эти computations существует. Это приводит к неизбежному $O(n^2)$ вычислительная сложность, которая может серьезно ограничить как наземное моделирование, так и бортовые алгоритмы (58) - см . Приложение E и F для деталей реализации.

Помимо вычислительной неэффективности, spherical harmonics также имеет эксплуатационные ограничения. Вывод этой модели требует, чтобы все элементы массы находились внутри сферы фиксированного радиуса (Радиус Бриллюэна). Если динамикам требуется оценить потенциал или ускорение в пределах этой сферы, то

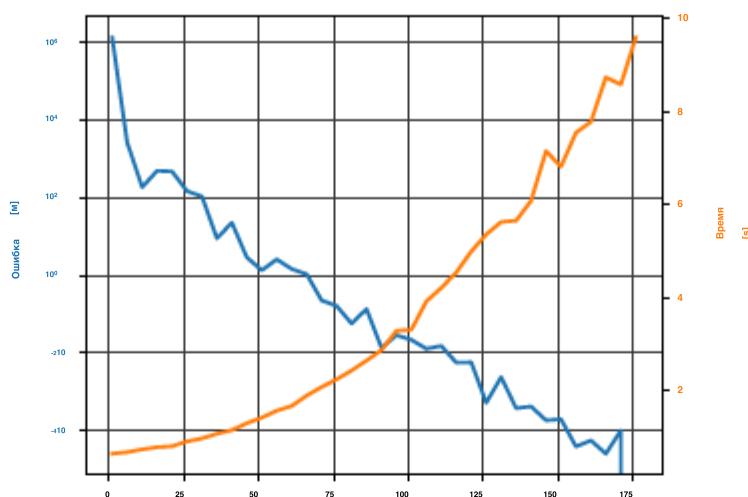


Рисунок 2.1: Вычислите время и конечную ошибку, связанную с моделированием космического аппарата на высоте 600 км, находящегося на орбите в течение четырех часов в режиме реального времени, используя формулировку Пайнса как функцию степени сферической гармоники.

модель может отличаться благодаря $(R/r)^l$ термин в расширении. Хотя такие эффекты незначительны для

планет или лун, близких к сферической форме, они могут стать проблематичными в условиях малых тел. Для астероидов

или комет, которые имеют сильно несферическую геометрию, таких как Эрос или Итокава, эти эффекты могут привести

к большим ошибкам в прогнозируемых ускорениях и потенциально поставить под угрозу безопасность космического аппарата (5).

2.2.2 Эллипсоидальные гармоники

Модель эллипсоидальных гармоник придерживается подхода, аналогичного подходу к сферическим гармоникам, но

вместо этого использует базисные функции эллипсоидальных гармоник (3). На первый взгляд потенциал можно представить

как

$$V = \sum_{n=0}^{\infty} \sum_{p=1}^{2n+1} a_{np} E_{np}(\lambda_1) E_n(\lambda_2) E_n(\lambda_3) \quad (2.2)$$

где a_{np} - константы, E - функции Ламе первого рода, λ_i являются эллипсоидальными координатами, n

- степень, p - собственные значения. Этот выбор позволяет получить более плотный эллипсоид, ограничивающий

тела, чем сферические гармоники, минимизируя область, в которой модель потенциально может расходиться.

Несмотря на это, расходжение остается возможным внутри ограничивающего эллипсоида, и модель по-прежнему

страдает от тех же проблем при представлении разрыва с помощью периодических базисных функций.

2.2.3 Внутренние сферические гармоники / модели Бесселя

Модель внутренних сферических гармоник инвертирует классическую формулировку сферических гармоник

и может моделировать локальную область, граница которой пересекает только одну точку на поверхности

тело (59). Явно, вместо вычисления потенциала с использованием закона косинуса и разложения на множители

от $1/r$ до

$${}^eU(r, \phi, \lambda) = G \int_M \frac{1}{r} \frac{1}{\sqrt{1 - 2(\rho_r \cos S) + \frac{r^2}{\rho^2}}} dm \quad (2.3)$$

вместо этого внутренняя сферическая гармоническая модель учитывает ρ

$${}^eU(r, \phi, \lambda) = G \int_M \frac{1}{\rho} \frac{1}{\sqrt{1 - 2(\rho_r \cos S) + \frac{r^2}{\rho^2}}} dm \quad (2.4)$$

Эта модель сохраняет устойчивость вплоть до этой единственной точки, что делает ее полезной для точной посадки

однако решение становится недействительным в любой другой точке поверхности и за пределами

соответствующей локальной сферы. Модель внутренней сферической гравитации Бесселя расширяет эту тему,

используя функции Бесселя, а не сферические гармоники, и способна обеспечить более широкую область действия

. Тем не менее, это происходит за счет громоздких аналитика и сохраняет некоторые

проблем с использованием гармонических базисных функций для захвата разрыва (60).

2.2.4 Масконы

Гравитационные модели Mascon распределяют набор точечных элементов массы внутри тела, сумма которых

может формировать глобальное представление гравитационного поля (61). К сожалению, ускорения могут

становиться неточными в точках поля рядом с отдельными масконами (4). Гибридные модели mascon предлагают

несколько более надежную альтернативу чисто масконовскому подходу, представляя каждый маскон с

сферическая гармоническая модель с низкой точностью воспроизведения, но она включает дополнительную сложность регрессии

и остается подверженной как вызовам моделей маскона, так и сферической гармоники (62).

2.2.5 Многогранная модель

Многогранная гравитационная модель обеспечивает альтернативу сферической гармонической модели в этих параметрах

предлагая решение, которое сохраняет актуальность вплоть до поверхности любого тела независимо от

форма. Такая стабильность делает многогранную модель основной гравитационной моделью для исследования

малых тел, таких как астероиды и кометы. Если доступна модель формы тела (коллекция

треугольных граней и вершин, которые отражают геометрию объекта), то связанный с этим

гравитационный потенциал этой геометрии может быть вычислен в предположении постоянной плотности

через:

$$\nabla U = \sum_{e \in \text{ребра}} -G\sigma_e E_e \cdot r_e + \sum_{f \in \text{грани}} L_f + G\sigma_f F_f \cdot r_f + \omega_f \quad (2.5)$$

где G - гравитационная постоянная, σ - плотность тела, E_e является граничной диадой, r_e является ли

вектор положения между центром ребра и точкой поля, L_e является аналогом потенциального

вклада ребра, F_f является ли лицо нормальной диадой, r_f является расстоянием между гранью по нормали

и точкой поля, и ω_f является аналогом потенциального вклада лица (5).

Многогранная гравитационная модель позволяет обойти численное расхождение в системе Бриллюэна сферы, что делает ее чрезвычайно популярным выбором для операций с малыми телами. Во многих работах эта модель используется для характеристики динамической среды вокруг малых тел и проектирования соответствующих траекторий (63; 64; 65; 66). Более того, эта модель находится в стадии активной разработки, с помощью геопоисковики разрабатывают модели, учитывающие неопределенность формы (67) и более эффективные стратегии оценки (68). Однако у этой модели есть некоторые недостатки. Прежде всего, это гравитационная модель может быть дорогостоящей с точки зрения вычислений. Модели формы с высоким разрешением содержат сотни тысяч граней и вершин, которые необходимо перебирать по отдельности для вычисления ускорения в одной точке поля. При вычислении большого количества ускорений или распространяющихся траекторий это вычислительная нагрузка может привести к чрезмерно длительному времени выполнения. Хотя эта модель может использовать преимущества параллелизации проще, чем сферические гармоники, такие вычислительные возможности не являются доступными на борту космического аппарата - по сути, ограничивающий эту модель наземным моделированием.

Дополнительно многогранная гравитационная модель предполагает, что исследователям известен профиль плотности для рассматриваемого тела. Чаще всего предполагается, что плотность постоянна, но литература показывает, что такое предположение не обязательно верно (69; 44). Кроме того, многогранная гравитация требует, чтобы модель формы тела уже существовала. Хотя эти модели можно приобрести

на месте этот процесс является нетривиальным и отнимает много времени на практике (70; 71; 72).

2.2.6 Неоднородный многогранник

Альтернативная формулировка многогранной модели предлагает рассматривать каждую грань формы

модель как тетраэдр, соединенный с центром формы. Тетраэдры можно разрезать на

множество частей, и каждой части присваиваются разные значения плотности на основе некоторого возможного распределения плотности-

распределение (73). Это позволяет провести аналитическую аппроксимацию тела с неоднородной плотностью, но

продолжает страдать от допущений, сделанных относительно возможного распределения плотности.

2.3 Модели машинного обучения

В качестве альтернативы аналитическим моделям недавние попытки исследуют использование машинного обучения для

изучения представлений сложных гравитационных полей на основе данных. В принципе, эти машинные

обучающие модели позволяют изучать гравитационные поля с высокой точностью, не делая предположений о рассматриваемом теле

при сохранении низких вычислительных затрат.

2.3.1 Нейронные сети

Нейронных сетей серия узнал, нелинейные преобразования картографических данных из входного

пространство для желаемого вывода пространства путем минимизации заданной функции потерь, например, означает в квадрате

ошибка:

$$J(\Theta) = \frac{1}{N} \sum_{i=1}^N \|y_i - \hat{y}(x_i | \Theta)\|^2 \quad (2.6)$$

где y_i является истинным результатом, $\hat{y}(x_i | \Theta)$ является прогнозируемым результатом работы искусственной нейронной сети с учетом

вектор обучаемых параметров Θ , который включает веса, w , и смещения, b , сети, и

N_f - общее количество точек, используемых для обучения сети.

Обычно эти сети строятся как серия плотно соединенных скрытых слоев с

N узлами на слой, выраженными как:

$$h_{i(k)}^{(k)} = \sigma \sum_{j=1}^{N_{k-1}} w_{ij} h_j^{(k-1)} + b_k \quad k \in \{1, \dots, N\}_{\text{макс}} \quad (2.7)$$

где $h^{(k)}$ является k -м скрытым слоем, i - узел в слое, w_{ij} веса, соединяющие

скрытые слои, b_i являются смещениями, прикрепленными к узлам в слое, а σ - нелинейным.

преобразование (обычно сигмоидальное, гиперболическое касательное или выпрямленная линейная единица). Обратите внимание, что $h^{(0)} = x$,

$$\text{и } h^{(k_{\max})} = \hat{y}.$$

Нейронная сеть обучается путем итеративного обновления весов и отклонений модели, чтобы

минимизировать уравнение (2.6) такой, что:

$$\begin{aligned} {}^*w &= \arg \min_{\in \Theta_{Wt}} (\text{Дж}(wt)); \\ {}^*b &= \arg \min_{\in \Theta_b} (\text{Дж}(b)) \end{aligned} \quad (2.8)$$

который может быть решен с помощью алгоритма градиентного спуска, такого как Adam или SGD (74; 75)

$$\Theta^{m+1} = \Theta^m - \nabla_{\Theta^m} \text{Джерой}^m(\Theta), \quad (2.9)$$

где η -скорость обучения и M является подготовка итерации, и градиент, сделанных с использованием автоматического дифференцирования.

Нейронные сети демонстрируют значительный потенциал в своей способности к регрессии высокоточный модели сложных явлений в самых различных научных областях. Несмотря на это гораздо критицизм существует в отношении этих моделей, а их overparameterization может привести к переобучению ведущий к ненадежным моделям за пределами первоначального обучения. Кроме того, эти модели часто описываются как "черного ящика" регрессоры, которые трудно интерпретировать или предоставить какие-либо аналитические гарантии. Несмотря на их недостатки, в качестве кандидата были предложены нейронные сети.

Решение проблемы гравитационного моделирования. Путем обучения данным о местоположении и ускорении в литературе было показано, что нейронные сети способны представлять сложные гравитационные явления (46).

Автоматическое дифференцирование

Автоматическое дифференцирование - это метод вычисления точной производной алгоритма по отношению к любым входным данным. Это делается путем построения вычислительного графика и использования либо прямого

или обратная форма цепного правила, такая как:

$$\begin{aligned}\frac{\partial y}{\partial x} &= \frac{\partial y}{\partial x^n} \frac{\partial x^n}{\partial w_1} \\ &= \frac{\partial y_n \partial w}{\partial w} \frac{\partial w_1}{\partial w} \frac{\partial w_2}{\partial w} \\ &= \dots\end{aligned}$$

где x - входные данные для некоторого произвольного алгоритма, w_n являются последовательностью промежуточных вычислений

, выполняемых для получения конечного результата, y .

Поскольку все алгоритмы построены на основе элементарных функций с известными производными,

части каждого промежуточного выражения всегда могут быть вычислены вместе с исходным вычислением.

Это свойство гарантирует, что часть выходного y по отношению к любому входному x может быть вычислена

автоматически. Автоматическое дифференцирование наиболее известно своим применением в рамках глубокого обучения,

где оно обычно применяется в алгоритмах стохастического градиентного спуска, используемых для обучения нейронных сетей

(76). В частности, уравнение 2.9 именно здесь применяется автоматическое дифференцирование. Градиент

заданной функции потерь относительно весов и отклонений сети может быть вычислен

автоматически и применен при снижении для перемещения параметров сети в направлении, которое позволит

минимизировать эту функцию потерь.

2.3.2 Машины экстремального обучения

Поскольку нейронные сети часто подвергаются обоснованной критике из-за их чрезмерной параметризации-

были разработаны альтернативные модели, чтобы смягчить некоторые из этих проблем. В частности, был предложен

альтернативный класс модели машинного обучения, называемый машинами экстремального обучения,

или ELMs. ELMs представляют собой однослойные нейронные сети, которые подбираются случайным образом путем инициализации весов

от входов к скрытому слою, а затем решить для Весов выходного слоя с помощью

наименьших квадратов подход к минимизации некоторой квадратичной функции затрат, таких как:

$$L(\theta) = \frac{1}{N} \sum_{i=0}^N \|y_i - \hat{y}_i(\theta)\|^2 \quad (2.10)$$

где $Y^*(x; \theta)$ является ли модель машинного обучения прогнозом на входе x с обучаемыми параметрами модели θ

6 (77). ELM имеет преимущество перед другими методами, потому что его можно обучить за одну итерацию

использование метода наименьших квадратов в отличие от нейронных сетей, которые требуют множества итераций стохастического метода.

градиентный спуск и обратное распространение. ELM достаточно большой ширины также были доказаны как

универсальные аппроксиматоры функций **(78)**. Недавно было показано, что ELMs успешно моделирует гравитационное поле

астероида Итокава в для сценария приземления и продемонстрировали высокую скорость оценки

с относительно высокой точностью **(6)**. Тем не менее, несмотря на простоту модели, ELMs создают

различные проблемы. А именно, ELMs требуют выполнения больших матричных инверсий для выполнения наименьших

квадратов соответствуют друг другу, что существенно ограничивает объем данных, которые могут быть использованы для обновления модели в любой момент времени

в данный момент времени.

2.3.3 Гауссовские процессы

В дополнение к ELMs, гауссовские процессы также были предложены в качестве возможных решений для

задачи гравитационного моделирования **(45)**. Гауссовские процессы подгоняются путем указания предварительного распределения по

функции и обновление предыдущей на основе наблюдаемых данных. Для этого пользователю необходимо указать некоторую

функцию ядра, которая обеспечивает показатель сходства между данными, и вычисление ковариационной матрицы

между всеми точками данных с использованием этой функции. Как только ковариационная матрица вычислена, ее

можно инвертировать и использовать для оценки среднего значения и ковариации изученной функции в контрольной точке.

С помощью гауссовых процессов для гравитационного моделирования является полезным, поскольку он обеспечивает возможную оценку неопределенности в модели прогнозирования; в то же время модель не масштабируется

хорошо для больших наборов данных. В частности, гауссовский процесс характеризуется ковариационной матрицей

, построенной на основе обучающих данных. Эта ковариационная матрица масштабируется как $O(n^2)$ где n -

размер обучающих данных. Этот скейлинг делает его вычислительно сложной задачей для гауссовых процессов на

рычагах - возраст больших объемов данных, так как размер ковариационной матрицы и сложность соответствующей инверсии,

растет с небольших наборов данных для обучения. В качестве примера, в справочной **(45)** процесс Гаусса-сиана

был регрессирован всего по 3600 точкам данных, в результате чего была получена модельная ковариационная матрица

из более чем 12 960 000 параметров. Поскольку размер этой модели быстро растет, исследователи должны тщательно

выберите, какие обучающие данные использовать, поскольку эти модели неизбежно не могут моделировать всю область тестирования

. Помимо этого факта, исследование демонстрирует, что эти модели, однажды подходящие, позволяет достичь

быстрого времени оценки и относительно высокой точности прогнозов, близких к данным обучения.

2.3.4 Геодезическая сеть

В 2021 году, Ссылка (7) предложил использовать поля нейронной плотности для изучения карт плотности для

малых тел. Это исследование вдохновлено недавними разработками машинного обучения в области нейронной

Поля сияния (NeRFs), которые используются для построения трехмерных моделей формы на основе относительно разреженных изображений

данные (79). После обучения прогнозы плотности могут быть интегрированы численно для получения соответствующих данных.-

соответствующие гравитационные потенциалы и ускорения. Эта работа продемонстрировала многообещающие результаты. -

достижение конкурентоспособной точности ускорения и способности учитывать неоднородные плотности.;

однако, в моделях довольно большие (свыше 90,000 параметров), численные Интегра-

цию, необходимую для вычисления ускорений является вычислительно дорого, а модели продолжают

требует большой подготовки наборов данных.

2.3.5 Нейронные сети, основанные на физике

Одним из недостатков традиционных моделей машинного обучения является то, что им не хватает аналитической основы

. Нет гарантий, что модель по своей сути удовлетворяет некоторым дифференциальному

уравнения и точно проанализировать по всей проблемной области. Когда нейронных сетей, экстрем

обучения машин, или GeodesyNets обучены представляют собой гравитационное поле, исследователи не могут обеспечить

что эти модели по своей сути соответствуют лежащей в основе физике, такой как уравнение Лапласа

или свойствам консервативного векторного поля.

В 2019 году Раисси и др. признали эту проблему и предположили, что модели нейронных сетей

не нужно быть агностиком в отношении физики, которая управляет функцией, которую они пытаются воспроизвести-

возмущаться (9). Вместо этого Раисси утверждает, что сети могут быть специально обучены для обеспечения того, чтобы обученные

представления подчиняются базовому дифференциальному уравнению с введением оригинальной физики-

Информированная нейронная сеть (PINN). PINNs вводят дифференциальные уравнения и граничные условия

в функцию затрат традиционных нейронных сетей, использующую автоматическое дифференцирование для обеспечения того, чтобы полученные решения естественным образом удовлетворяют этим ограничениям.

В качестве примера рассмотрим следующее произвольное дифференциальное уравнение:

$$f'(x) + f(x) + f(x) = 0, \quad (2.11)$$

Предположим, что существуют измерения x , а также соответствующие значения $f(x)$. Традиционный нейронная сеть может использовать эти наблюдения в качестве обучающих данных для изучения отображения из $x \rightarrow f(x|\Theta)$ путем минимизации функции затрат $J(\Theta) = \|f(x) - f(x|\Theta)\|^2$. Риск обучения сети с помощью эта конкретная функция потерь заключается в том, что сеть не знает, что отображение $f(x|\Theta)$ должно также удовлетворять уравнению. (2.11). PINNs меняет эту парадигму, вставляя исходное дифференциальное уравнение в функцию стоимости:

$$J(\Theta) = \frac{1}{N} \sum_{i=1}^N \|f(x_i) - f(x_i|\Theta)\|^2 + \|f'(x_i|\Theta) + f(x_i|\Theta) + f(x_i|\Theta)\|^2, \quad (2.12)$$

где производные сети $f'(x_i|\Theta)$ берутся с автоматическим дифференцированием. Эта стоимость функция, хотя и аналогична эквалайзеру. (2.6), не только наказывает за ошибочные значения $f(x|\Theta)$, но и наказывается, когда изученная функция нарушает дифференциальную форму задачи. Этот дополнительный термин служит формой регуляризации в процессе обучения, которая может привести к улучшенным решениям, которые удобно также удовлетворять важным физическим свойствам. Проведена большая работа по изучению и полезность этих PINNs. Недавняя работа показала, что эти модели нередко обладают преимуществами данных эффективность и точность моделирования на их традиционные аналоги.

Глава 3

Модели гравитации нейронной сети, основанные на физике

В этой диссертации исследуется использование нейронной сети, основанной на физике (PINN), для регрессии

формы гравитационного потенциала. Гравитационная модель PINN претерпела несколько поколений

этап разработки - на каждом этапе вносятся дополнительные усовершенствования для повышения точности модели,

надежность и эффективность обработки данных. В этой главе освещаются хронологические изменения каждого поколения.

и соответствующие характеристики на каждом этапе.

3.1 Поколение I

Прошлые попытки моделирования гравитации продемонстрировали, что машинное обучение может быть использовано для регрессии

модели гравитационного поля (6; 45). Это может быть достигнуто путем сбора оценок космического аппарата

положение, x и ускорение, и, используя этот набор данных для обучения, чтобы соответствовать нейронной сети или

экстремальные машинного обучения путем минимизации следующей функции затрат

$$J_{00}(\Theta) = \frac{1}{N} \sum_{i=1}^N \| \hat{a}_i - f(x_i, \Theta) \|^2 \quad (3.1)$$

где $f(x, \Theta)$ - изученное отображение вектора ускорения, параметризованное весами и

смещения сети Θ (т.е. $f(x, \Theta) = \hat{a}$).

Эти предыдущие попытки, хотя и являются важным вкладом в литературу, не пользуются особым успехом-

поскольку они не зависят от физики системы, которую они пытаются представить. В отличие от

сферические гармонические или многогранные гравитационные модели, решения, полученные с помощью нейронных сетей, не имели

не знали, что гравитационные ускорения, которые они представляют, являются побочным продуктом более

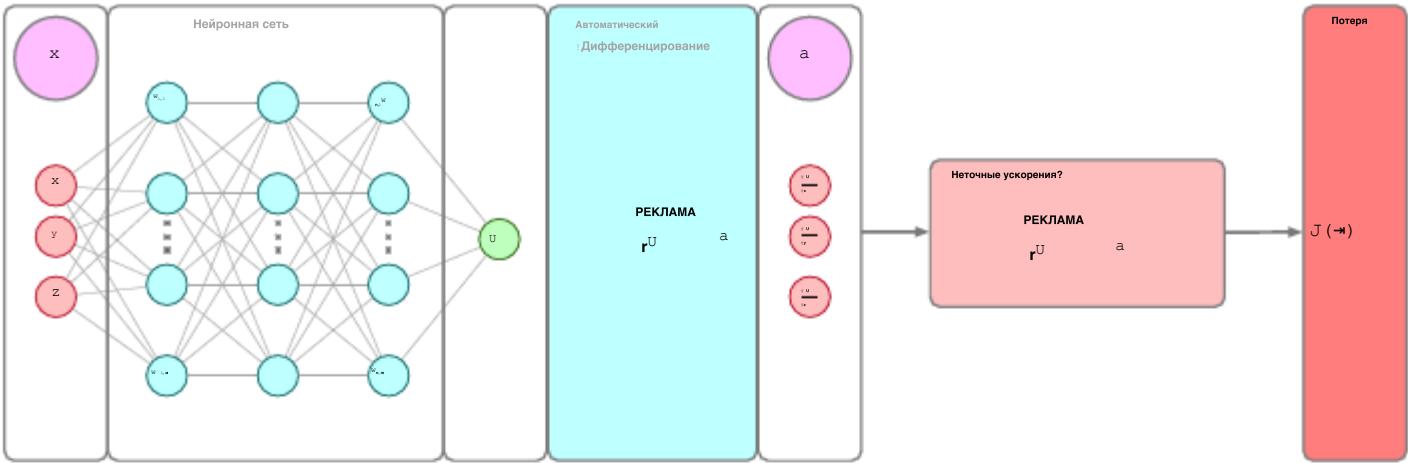


Рисунок 3.1: Поколение I ПИННА-GM

фундаментальная скалярная потенциальная функция. Более того, в этих решениях не учитывается, что гравитационная-

экономическая сила консервативна, равно как и то, что поведение на больших высотах должно стремиться к нулю.

До недавнего времени у исследователей не было способа включить эти знания в построение этих моделей.

модели машинного обучения. Однако в 2019 году Раисси и др. представили нейронную сеть, основанную на физике

(PINN), которая дала первый взгляд на то, как этого можно достичь (9).

Как обсуждалось в главе 2, PINN - это модели машинного обучения, которые используют расширенную функцию потерь

. Эти функции потерь не только наказывают за неточности моделирования между прогнозируемым

и истинным значением, как и в случае с традиционным машинным обучением, но они также наказывают за нарушения

лежащих в основе дифференциальных уравнений системы, которые они пытаются смоделировать. В случае

что касается проблемы моделирования гравитационного поля, физика подразумевает, что гравитационные ускорения на самом деле являются

побочными продуктами более фундаментального скалярного потенциала через:

$$\mathbf{a} = -\nabla U \quad (3.2)$$

ПИНс может использовать это дифференциальное уравнение непосредственно в функции затрат:

$$_A\mathbf{J}(\Theta) = \frac{1}{N} \sum_{n=1}^N \mathbf{a}_n + \nabla \hat{\mathbf{u}}(\mathbf{r}_n \| \Theta), \quad , \quad (3.3)$$

где a_r - расчетное ускорение в положении r , U^* - изученная потенциальная функция, а ∇U^* -

оператор градиента, применяемый посредством автоматического дифференцирования

Уравнение 3.3 это основанная на физике функция потерь. Вместо того, чтобы просто минимизировать разницу- между прогнозируемым ускорением и истинным ускорением, уравнение. 3.3 обеспечивает, чтобы предсказанное ускорение было связано с гравитационным потенциалом. Таким образом, нейронная сеть не представляет вектор ускорения напрямую, а вместо этого изучает скалярную потенциальную функцию, u^* и обеспечивает соответствие ее отрицательного градиента измеренному ускорению.

Одним из преимуществ этого подхода является повышенная эффективность обработки данных. При обучении традиционная нейронная сеть непосредственно изучает вектор ускорения, три компонента вектора эффективно рассматриваются как отдельные функции для изучения сетью. Фактически, эти компоненты по конструкции ортогональны, так что между ними нет наблюдаемой взаимосвязи. Таким образом, традиционная нейронная сеть должна изучать базовый набор, который может решать три разные задачи одновременно. Гравитационная модель PINN, напротив, должен только учиться скаляр по- циала функции; использование всех трех составляющих ускорения к регрессии один объект.

Как следствие, PINNs можете более эффективно использовать то же количество данных для обучения, как сокращает время тренировки и помогает облегчить dynamicists от необходимости собирать большие объемы данных.

Следует отметить, что к функции потерь могут быть добавлены дополнительные ограничения, основанные на физических данных. Например, функция затрат может также включать штраф за:

(1) Мис-моделирование потенциальной функции на себя, $U(p) - U^*(P||\Theta)$

= 0 (2) нарушении граничных условия $U(P) = 0$, так как $\|P\| \rightarrow \infty$ (3) не удовлетворяют скромным векторного поля, свойства, как $\nabla^2 u^* = 0$ и

$$\nabla x \cdot u^* = 0.$$

Это первое поколение PINN-GM, или PINN-GM-I, намеренно исключает эти дополнительные ограничения, основанные на физике по двум причинам. Необязательное ограничение (1) опущено, поскольку область потенциала на порядок больше области соответствующих ускорений. Эти отличающиеся весы представляют численных задач для обучения, и могут ввести слишком большие градиенты в метод обратного распространения ошибки. Кроме того, этот выбор отражает более общие обстоятельства где высокоточный потенциал рассматриваемого корпуса не существует и, следовательно, не может быть использован.

Ограничения (2) и (3) опущены, поскольку, на ранних стадиях этой работы они имели тенденцию

приближать решение PINN для U к нулю или просто к медленной сходимости. Это объясняется

многоцелевой оптимизацией, которую алгоритмы градиентного спуска с трудом приспособливают (80).

Очевидно, что ограничения консервативного векторного поля могут быть выполнены, если сеть идентифицирует

$U = 0$ как оптимальное решение. Это может вынудить сеть отказаться от ненулевого решения, ориентированного

по ограничению $a + \nabla U = 0$. Теоретически, эти конкурирующие цели могут быть лучше сбалансированы

за счет динамической корректировки темпов обучения с учетом каждой из объективных затрат(81). Эта стратегия заключается в

изучить в последующих поколениях PINN-ГМ.

Наконец, следует подчеркнуть, что представленные PINN тяжестью формулировка модель не

требуется действующая гравитационная модель для получения данных для обучения. PINNs можно обучать на месте, полагаясь

на оценки ускорений либо с помощью конечной разности относительных скоростей, либо с помощью более продвинутых методов

фильтрации. В этой главе основное внимание уделяется проблеме гравитационного моделирования, а не

проблема оценки силы тяжести. Таким образом, существующие высокоточные модели используются в качестве обучающих данных для

этой главы, но глава 6 мы проведем более подробное обсуждение того, как можно обучать пиннов.

на месте без ранее существовавшей гравитационной модели.

3.1.1 Архитектура сети и используются для отбора признаков

Все сети, обученные в предстоящих экспериментах, проведение поделиться архитектура восемь густо

связанные скрытыми слоями с N вершинами каждого слоя. Выбор из восьми скрытых слоев предложили Дезир-

в состоянии баланса между пропускной способностью сети и разумные сроки обучения. Как уже говорилось, сетей

подготовки данных состоит из установки и пары ускорение создается на основе существующих высококачественный

гравитационные модели. Для всех сетей данные о местоположении предварительно обрабатываются с помощью преобразования min-max

соответствуют каждому компоненту из Γ в пределах [-1, 1]. Эта нормализация гарантирует, что входные данные

существуют в благоприятной, нелинейной части функции активации сети, а также способствуют

числовой стабильности во время обучения. Кроме того, выходные данные ускорения также нормализованы

с помощью преобразования min-max. Обратите внимание, что для традиционных нейронных сетей преобразование min-max-

формирование может применяться для каждого отдельного компонента вектора ускорения. Для PINNs,

однако все компоненты должны масштабироваться одновременно, чтобы размеры соответствовали физическим требованиям

для функции изученного потенциала.

Для PINN-GM-I гиперпараметрами, которые оказывают наибольшее влияние на производительность, являются скорость обучения и размер мини-пакета. Если размер пакета слишком мал, алгоритм градиентного спуска может перемещаться в направлениях, отличных от локального минимума, что приводит к увеличению времени обучения. К сожалению, небольшие пакеты часто неизбежны, когда в доступном графическом процессоре недостаточно видеопамяти для хранения всего набора данных на устройстве. В этом случае меньшая и более осторожная скорость обучения должна использоваться для обеспечения того, чтобы оптимизатор быстро не привел сеть к неоптимальному минимуму, которого позже будет сложно избежать. И наоборот, если можно использовать пакет большого размера, более высокие темпы обучения поощряются, поскольку они обеспечивают меньшую продолжительность обучения, которая с большей вероятностью будет снижаться в направлении истинного градиента функции затрат.

Независимо от начальной скорости обучения, может быть выгодно медленно снижать скорость обучения ближе к концу обучения, чтобы предотвратить колебания весов выше стоимости локального минимума функции (82). Для достижения этого к процессу обучения применяется экспоненциальный спад.

Оценить:

$$\eta_i = \eta_0 * \text{вейлонопленный } \alpha, -\frac{\alpha}{\alpha - \eta_0} \quad , \quad i \geq i_0 \quad (3.4)$$

где η_i скорость обучения в эпоху i , i_0 является эталонной эпохой, после которой начинается распад, α равно

масштабный коэффициент, α - скорость затухания, а η_0 - начальная скорость обучения.

Дополнительным критическим гиперпараметром является функция активации. При обучении PINN она важно, чтобы выбранная функция активации имела достаточно высокий порядок непрерывности. Если функция активации не имеет гладких производных высокого порядка, и если градиенты сети

взяты с использованием автоматического дифференцирования для обеспечения соблюдения физических ограничений, функция стоимости

больше не будет хорошо себя вести при градиентном спуске. Таким образом, рекомендуется избегать использования

популярные устраниены единица измерения (Релуи) или дырявой Релу и вместо этого выбрать функции с бесконечным

заказы преемственности, как гиперболический тангенс или гауссовой экспоненциальная линейная единица (Джелу). В

остальные гиперпараметры по умолчанию для PINN-GM-I показаны в таблице 3.1 и любой обычай

гиперпараметры приведены в таблице 3.2.

Каждая сеть обучается в течение 100 000 эпох с использованием оптимизатора Adam в Tensorflow 2.4

¹ на

видеокарте NVIDIA RTX 2060. Такая продолжительность обучения гарантировала постоянную потерю проверки

плавный переход, указывающий на завершение обучения. Все сетевые веса инициализируются с использованием

Схемы равномерной инициализации Xavier, подробно описанной в Ссылке (83).

Таблица 3.1: Общие гиперпараметры для традиционных и основанных на физике нейронных сетей, обученных в этой работе

Скорость обучения по	Значение	Гиперпараметр	Значение
гиперпараметру , η_0	0.005	Инициализатор	Glorot Uniform
Начинается эпоха Терпения, η_0	25000	Эпохи	100000
Скорость распада LR-планировщика, α	0,5	преобразование x	MinMax
Масштабный коэффициент, σ	25000	Активация	GELU
Оптимизатор	Адам	Количество слоев	8

3.1.2 Характеристики сферических гармоник

Сферические гармоники - это фактическая гравитационная модель для крупных планетарных тел. Наиболее точный сферический гармонический гравитационной модели Земли гравитационная модель Земли 2008

(СГЭ-2008), который содержит сферических гармоник, которые распространяются на степень и порядок 2,160 - до-

подсчет более 4 000 000 параметров (53). Хотя сферические гармоники особенно удобны при

представлении гравитационных характеристик глобального масштаба с гармониками низкой степени, их первоначальная эффективность

не сохраняется при более высоких степенях.

Чтобы продемонстрировать, рассмотрим характерные гравитационные характеристики, которые остаются после удаления точечная масса, сплюснутость планеты и ускорения при наклоне планеты (т.е. ускорения

, создаваемые гармониками выше степени и порядка 2). Для просмотра этих характеристик вводится переменная δ -

введена:

$$\delta a(r) = -\nabla U_{SH} \text{ Истина}(r) - -\nabla U_{2SH}(r), \quad (3.5)$$

¹ <https://www.tensorflow.org/>

Таблица 3.2: Уникальные гиперпараметры для традиционных и основанных на физике

нейронных сетей, обученных в этой работе

Тип сети	Узлы на уровне (N)	Параметры модели	Размер пакета	трансформация
Традиционный	10	843	262144	Минмакс
Традиционный	20	3083	262144	минмакс
Traditional	40	11763	262144	минмакс
Traditional	80	45923	262144	минмакс
PINN	10	820	262144	Униформа
PINN	20	3040	262144	Униформа
PINN	40	11680	262144	Униформа униформа
PINN	80	45760	131072	

где $U_{\text{Истина SH}}$ построена с использованием EGM-2008, расширенного до степени $l = 1000$, $U_{2\text{SH}}$ это то же самое

модель расширена только до степени и порядка 2, а градиент потенциала рассчитан с использованием алгоритма Пайнса

, чтобы избежать особенностей на полюсах (84). Рисунок 3.2 показывает эти возмущения, вычисленные на

сфера Земли Бриллюэна. Этот рисунок подтверждает, что возмущения являются прерывистыми элементами

в земной коре, например, горные хребты, такие как Гималаи и Анды, зоны тектонической субдукции, такие как

лучше всего видны в Тихом океане, и горячие точки, разбросанные по всему земному шару. Эти выводы согласуются с интуицией.

поскольку ускорения прямо пропорциональны градиенту потенциала ($a = -\nabla U$), а большие

смещения в массиве суши приводят к большим изменениям потенциала.

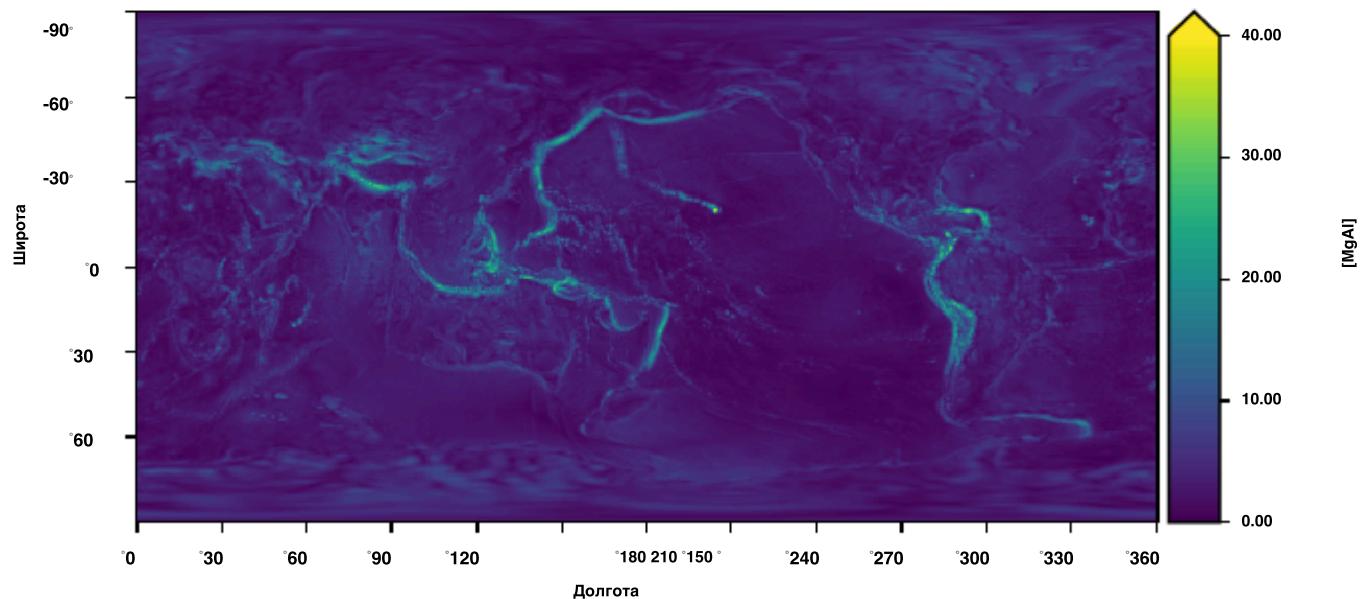


Рисунок 3.2: Карта да на земной сфере Бриллюэна

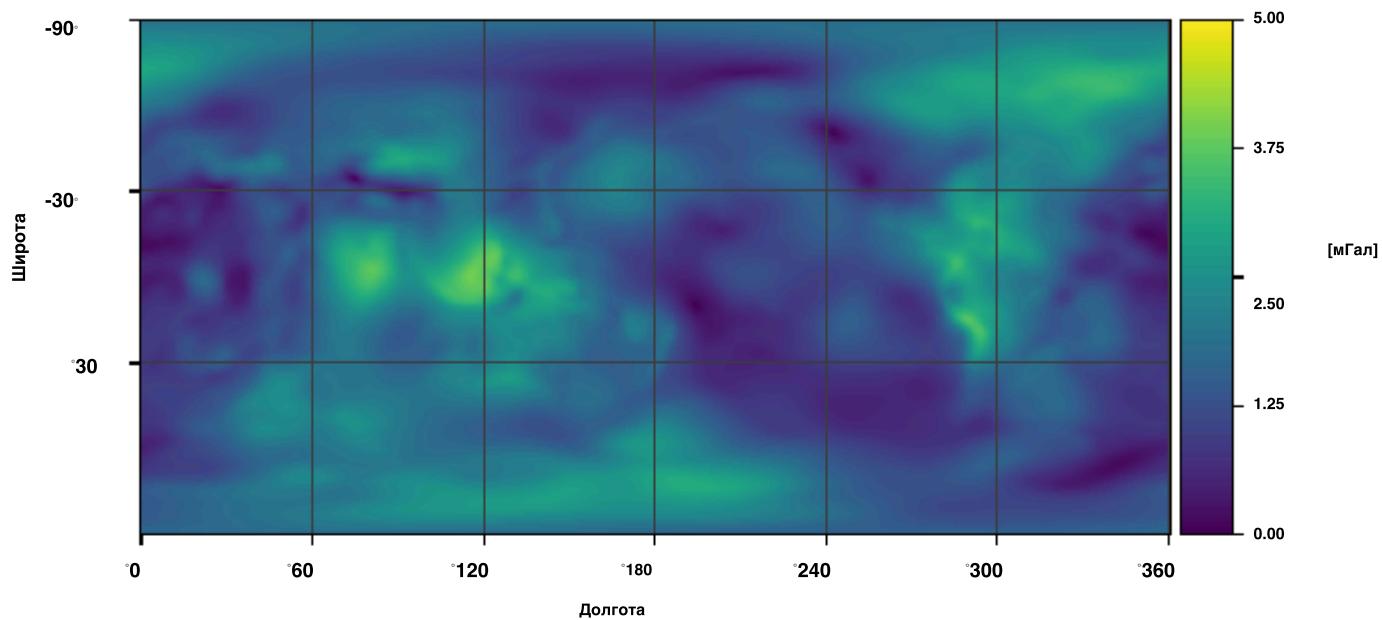


Рисунок 3.3: Карта ёа Земли на высоте LEO (приблизительно 420 км)

Эти возмущения являются особенно важными характеристиками для учета с помощью гравитационной модели, во-вторых

только к сплющенности планеты. Поэтому разумно задаться вопросом, насколько эффективно сферические

гармонические модели отражают эти возмущения. В частности, когда сферическая гармоника низкой степени

используются модели, насколько велика погрешность в этих важных частях гравитационного поля? Для исследования,

вводится показатель среднеквадратичной ошибки (MRSE):

$$\text{MRSE}(A) = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta a(r_i)}, \quad r_i \in A, \quad (3.6)$$

A - набор положений, для которых оценивается гравитационное поле, т.е. точек поля, N ,

где общее количество точек поля в наборе

δa и a_p является обобщением уравнения. (3.5) используется для измерения

соответствующей погрешности ускорения низкоточкой сферической гармонической модели посредством:

$$\delta a(r_i) = \| -\nabla U_{\text{Истина}}(r_i) + \nabla U_{\text{model}}(r_i) \|_p, \quad (3.7)$$

где P представляет максимальное число показателей / коэффициентов, используемых в гравитационном поле модель

проводится оценка.

Для характеристики погрешности низкой точности сферических гармонических моделей, MRSE метрика АП-

курсировали до трех наборов данных: A , F и $C. B$. Набор данных A включает $N_f = 250\,000$ полевых точек

в сетке Фибоначчи на земной сфере Бриллюэна. Сетка Фибоначчи выбрана для обеспечения близкого

изотропное распределение данных о сфере Бриллюэна, что позволяет избежать кластеризации данных на

полюсы были установлены для равномерного распределения по широте и долготе (85). Второй набор,

F_A , является подмножеством из A , которое включает только точки данных в пределах доминирующих гравитационных возмущений.

В частности,, F_A генерируется путем выбора точек поля в пределах

A ускорение которого превышает 2 ступени-

максимальное отклонение среднего ускорения, такое, что

$$F_A : \delta_a(r) - \delta_a > 2\sigma_a(r). \text{ Третий } A$$

окончательный набор C полняет set

$F_C : A F_A /$, представляющий фон гравитационного поля.

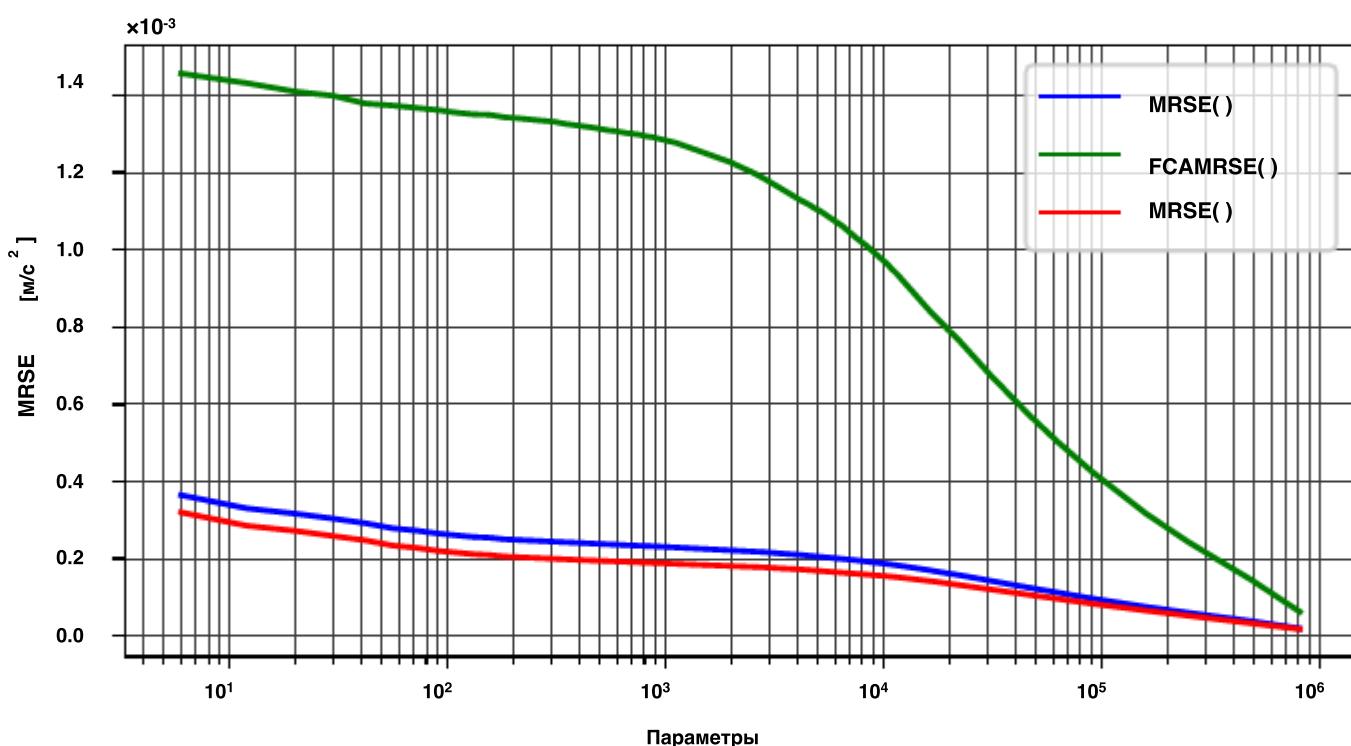


Рисунок 3.4: График MRSE (A), MRSE(F) и MRSE(C) как функции суммарных параметров, p , в

использовалась сферическая гармоническая гравитационная модель, где $p = l(l+1)$

С и представлен на рисунке 3.4 как функция A, F и метрика MRSE применяются к наборам

полных параметров, используемых в представлении сферической гармоники. Этот рисунок количественно определяет отношение-

зависимость между размером модели сферической гармоники и точностью модели. Синяя кривая MRSE для

A рассказывает обманчиво простую историю: чем больше параметров используется в сферической гармонической

модели, тем выше средняя точность по сфере Бриллюэна. Однако, если данные разделены между

динамики значимые функции

F по сравнению с менее значимыми функциями

C, представляющий гораздо больший интерес-

появляется результат редактирования. В частности, ошибка в пределах доминирующих гравитационных характеристик,

F , устранена

в 10 раз больше, чем его фоновый аналог в режиме с низким параметром. Фактически, для этого требуется

около 10 000 параметров, или сферическая гармоническая модель со степенью 110 или выше, до этого разрыва.

производительность начинает снижаться заметными темпами.

Несоответствие между ошибкой моделирования

A и F демонстрирует, как сферический-

мониторное представление изо всех сил пытается уловить возмущения в порядке динамической значимости. Хотя

первоначальная эффективность представления сплюснутости Земли неоспорима, удобство не исключает-

приводят к следующим по важности возмущениям. Это связано с тем, что сферические гармоники имеют приоритетное значение.

подгонка предписанной геометрии к системе, в которой эта геометрия отсутствует естественным образом.

Как следствие, сферическая гармоническая модель должна накладывать множество частот высокого порядка /

гармоники, прежде чем фиксировать эти возмущения.

Для специалиста по динамике это означает, что следующие по значимости возмущения за пределами J_2 требуется,

как минимум, модель сферической гармоники, превышающая степень $l = 110$, если они стремятся включить-

перенесите динамические эффекты возмущений Земли высокого порядка в их применение. В некоторых случаях

это может не представлять проблемы. Когда существует достаточно высокоточная модель и

исследователь не ограничен в вычислениях, сферические гармоники в конечном итоге будут сходиться даже по

прерывистым элементам. Однако на борту космического аппарата вычислительные ресурсы могут быть ограничены

или для рассматриваемого тела может не существовать высокоточной сферической гармонической модели. В этих условиях-

это может отрицательно сказаться на работе в течение короткого периода времени и вблизи крупных элементов поверхности.

сферическими гармониками, неспособными эффективно представить эти возмущения.

Более того, результаты, показанные на рисунке 3.4 мотивируют, почему это исследование обращается к изученным нейронным сетям

представления о гравитации в сети как альтернатива сферическим гармоникам. Сферические гармоники

Базис по своей сути ограничен в разрешении прерывистых возмущений при сохранении компактной модели

размер. Возмущения, присутствующие на Земле, требуют маловолновых гармоник, которые являются только

присутствует в расширениях высокой степени. Нейронные сети, напротив, не имеют предписанных базовых

функций и соответствующих характеристических длин волн. Не существует присущего минимума в 10 000

параметры, необходимые для представления конкретного горного хребта или другого разрыва. Скорее,

нейронные сети изучают удобную основу, которая представляет наиболее важные возмущения поля

независимо от их геометрии или масштаба. В принципе, нейросетевые модели гравитации могут

следовательно, давать более компактные представления, достигающие сопоставимой, если не большей, точности

, чем традиционные сферические гармоники.

3.1.3 Репрезентативная Компактность

Следующий эксперимент направлен на то, чтобы определить, может ли PINN-GM-I создавать столь же точные модели силы тяжести.

гравитационные модели используют меньше параметров, чем их сферический гармонический аналог. Для этого эксперимента,

параметры модели, p , относятся к общему количеству коэффициентов, используемых в сферической гармонической модели

или к количеству обучаемых весов и отклонений сетевой модели.

Эксперимент начинается с обучения как традиционных, так и основанных на физике нейронных сетей на

5 000 000 парах векторов положения / ускорения, которые выбираются случайным образом из равномерного распределения

по высоте (0-420 км), широте и долготе. Подготовлено в общей сложности восемь сети, каждый с

модели разной мощности, представленные в таблице 3.2. После обучения показатель MRSE (Eq. (3.6)) является

используется для оценки производительности каждой сети с использованием тех же данных сетки Фибоначчи, которые использовались ранее

для генерации рисунка 3.4 - $\gamma^r \in \{A, F, C\}$. Образцы теста Фибоначчи обеспечивают полностью несвязанный

набор данных из обучающих данных для обеспечения справедливой оценки производительности сети. MRSE

для сетей сопоставлен с MRSE сферической гармонической модели и представлен на

Рисунке 3.5.

Рисунок 3.5 демонстрирует, что существует широкий диапазон, в котором нейронные сети генерируют

более точные модели гравитационного поля, использующие меньшее количество параметров по сравнению с их сферическими аналогами.-

монические аналоги. Этот диапазон охватывает от 1000 до 50 000 параметров или от сферической

степени гармоники $l = 30$ до $l = 225$. Лучше всего это видно на наборе данных F , где PINNs

может создавать модели равной точности, используя на порядок меньше параметров, чем

сферические гармоники. Традиционные нейронные сети также создают более компактные

представления, хотя и меньшие , чем модель PINN аналогичного размера.

Тем не менее, существуют условия, при которых преимущество PINN в компактности становится менее очевидным-

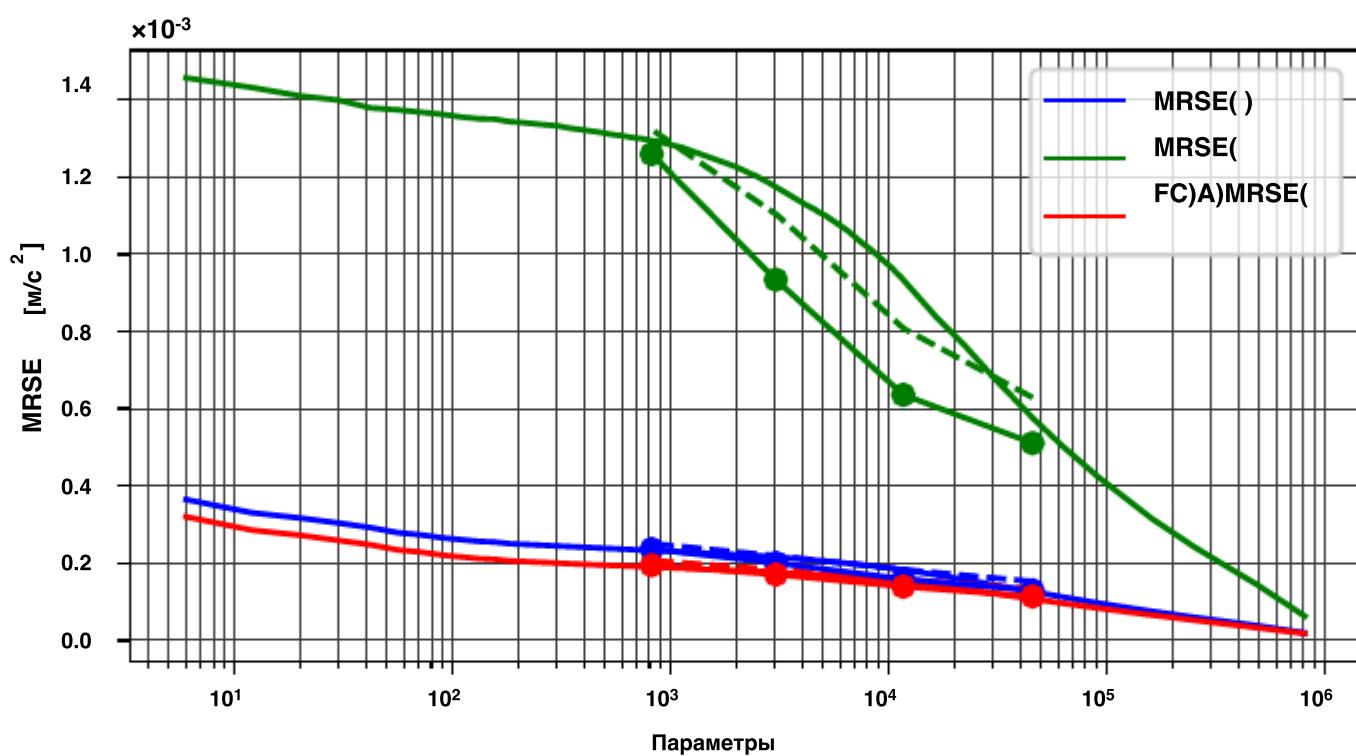


Рисунок 3.5: График MRSE как функции общих параметров модели, стр. Сплошные линии представляют сферическое гармоническое представление. Пунктирные линии представляют традиционные нейронные сети. Линии с круговыми маркерами представляют собой нейронные сети, основанные на физике

ent. В частности, обратите внимание на нейронные сети с наименьшей и наибольшей пропускной способностью ($N = 10, 80$).

{ }

Эти конкретные сети имеют менее выраженное преимущество в компактности, которое предполагается как

результат действия двух факторов. В случае $N = 10$ сети не имеют достаточно высокого

способность моделирования представлять нелинейные возмущения гравитационного поля. С помощью простого

$p \approx 1000$ обучаемых параметров, эти небольшие сети не обладают необходимой параметрической гибкостью-

сари для захвата прерывистых возмущений высокого порядка. Случай $N = 80$, напротив, обладает такой

большой пропускной способностью сети, что ему удается достаточно хорошо моделировать эти особенности - фактически настолько хорошо, что

сеть начинает перестраиваться под обучающие данные и страдает при тестировании на новых данных. Это можно

исправить с помощью дополнительных обучающих данных или более продуманной конструкции модели, как будет показано в последующих разделах

.

Компактность преимущества, представленные здесь, предполагают, что модели машинного обучения могут

чтобы узнать набор базисных функций, что значительно более эффективно, представляющий высокого порядка

возмущения гравитационного поля Земли. В случае традиционных нейронных сетей эти базовые функции

и промежуточные нелинейные преобразования, используемые для их генерации, могут быть нанесены непосредственно

на график, как показано на рисунке 3.6.

Рисунок 3.6 может быть интерпретировано следующим образом: верхняя строка представляет первый слой в сети

- т.е. входные данные нормализованных декартовых векторов положения. Поскольку эти изображения генерируются на

сфера Бриллюэна, компоненты x и y векторов положения уменьшаются на полюсах и

увеличиваются на экваторе, тогда как z -компонент линейно увеличивается от южного к северному полюсу. В

вторая строка представляет взвешенные линейные комбинации первых трех входных данных, которые затем передаются

через функцию активации гауссовой экспоненциальной линейной единицы (GELU). Эта строка соответствует

к выходам первого скрытого уровня сети. Это нелинейное преобразование повторяется

для каждого промежуточного скрытого слоя до 9-го ряда рисунка (8-го скрытого слоя), который

представляет изученные базовые функции. Затем эти предпоследние функции объединяются линейным способом

без преобразования ГЕЛУ для получения трех предсказанных составляющих ускорения в земной сфере Бриллюэна

(рисунок 3.7).

Примечательно, что по мере того, как входные сигналы распространяются глубже в сеть, соответствующие выходные сигналы

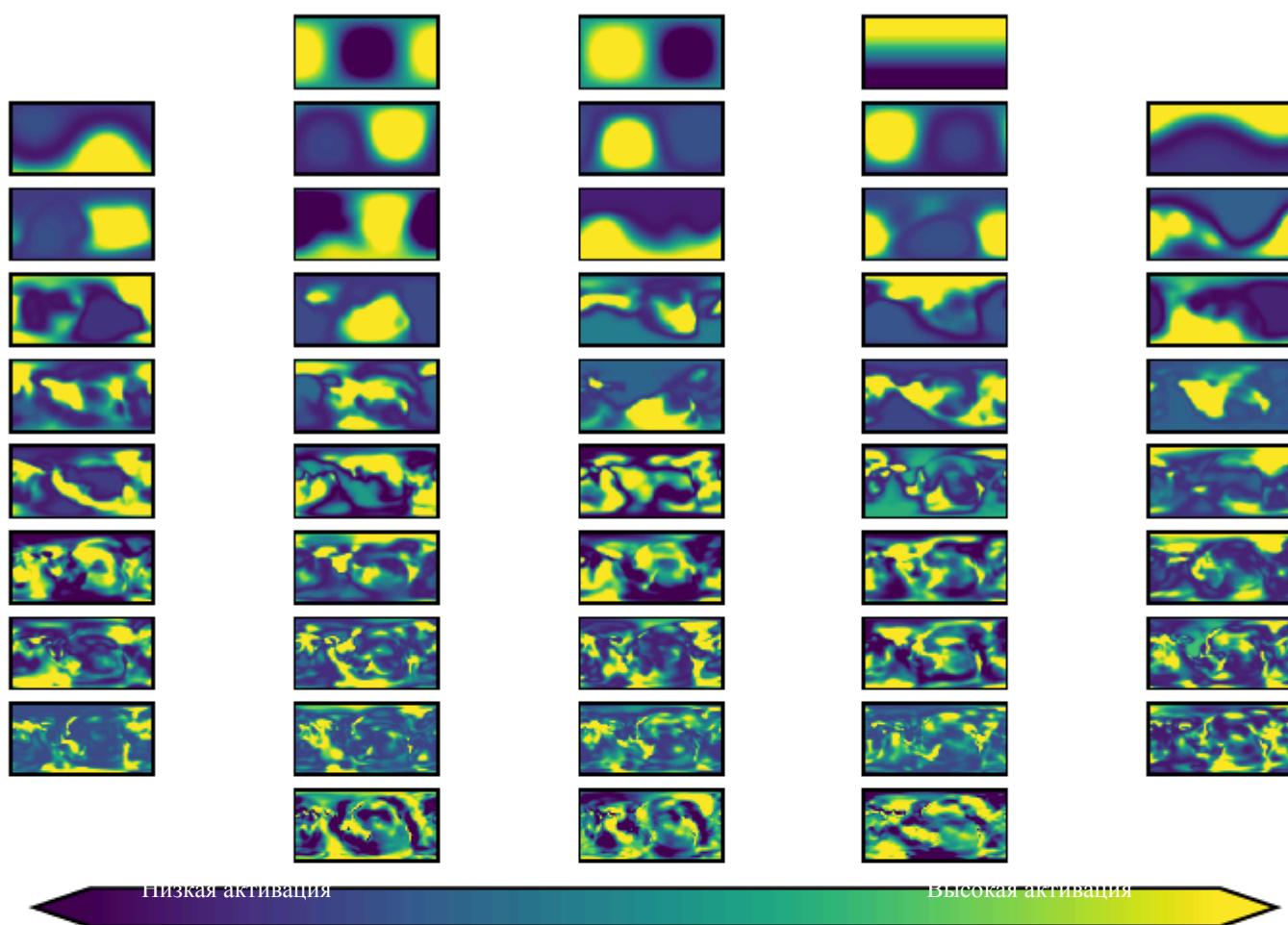


Рисунок 3.6: Подмножество промежуточных преобразований и результирующая базисная функция для традиционной нейронной сети $N = 40$. Каждая строка соответствует одному уровню сети в порядке от ввода (верхний) к выводу (нижний). Отдельные графики представляют собой нормализованный и безразмерный результат функции активации конкретного узла при оценке в земной сфере Бриллюэна

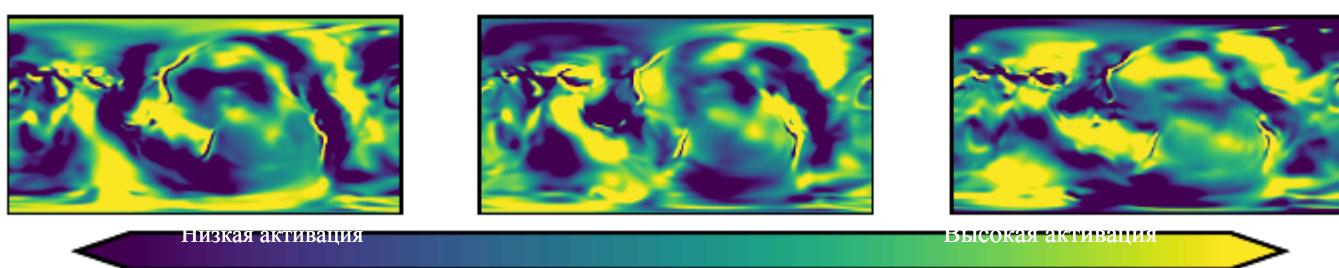


Рисунок 3.7: Увеличенный последний слой рисунка 3.6 представляющие предсказанные декартовы компоненты векторов ускорения, нанесенных на сферу Бриллюэна

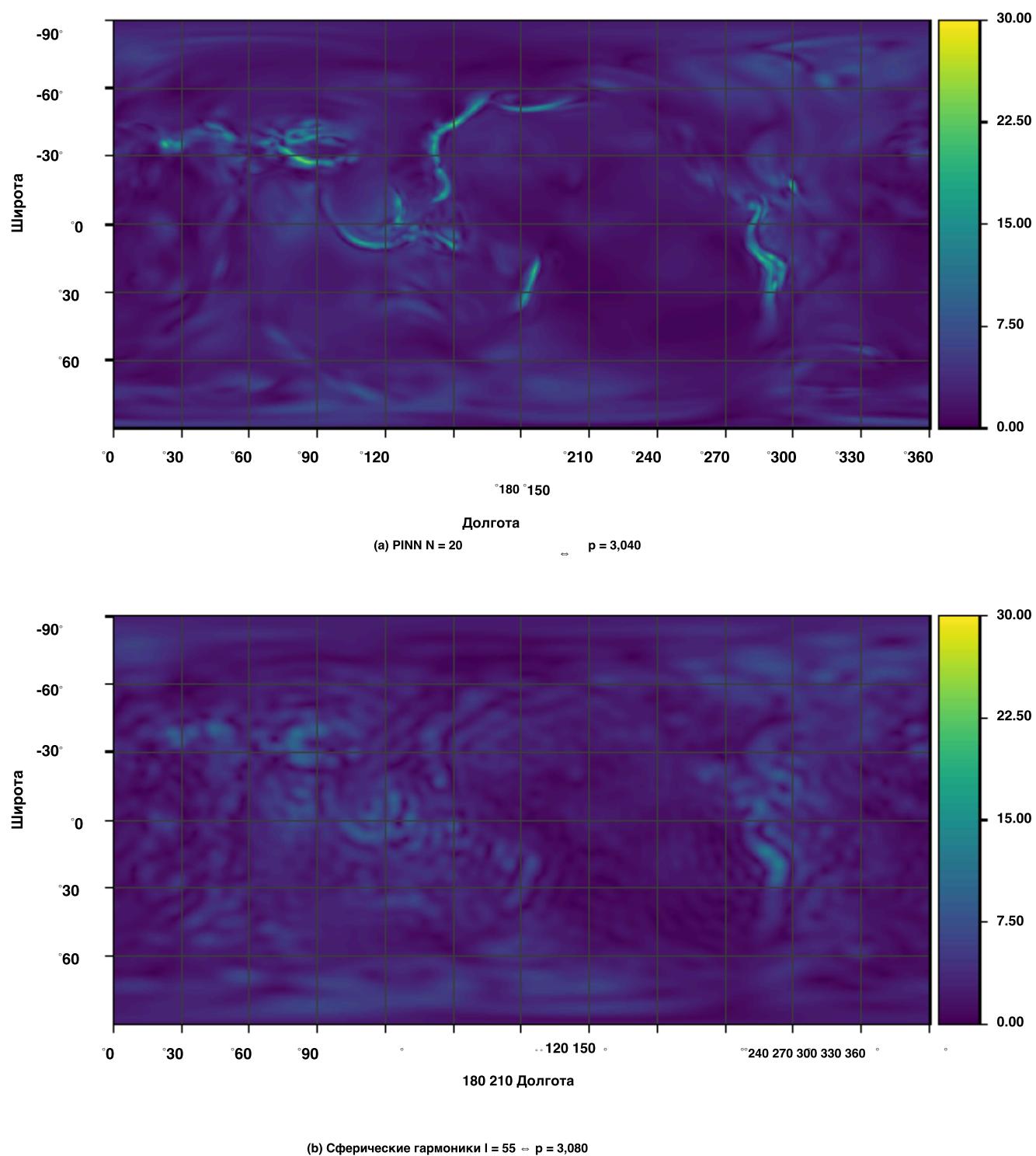


Рисунок 3.8: Гравитационная модель Земли с использованием (а) представления нейронной сети и (б) сферического гармонического представления при примерно одинаковом количестве свободных параметров ($p \approx 3000$)

скрытые слои становятся все более сложными. Более ранние слои активируются в широких областях по всей площади всей сферы Бриллюэна, в то время как более глубокие слои активируются в более локализованных объектах. Это показывает, как более мелкие слои в сетях имеют тенденцию разрешать высокоуровневые абстрагированные пространства объектов в то время как более глубокие слои начинают соответствовать конкретным возмущениям рассматриваемого тела.

Рисунок 3.8 наглядно демонстрирует разницу между моделированием прерывистых возмущений Земли-с использованием сферической гармонической модели с низким параметром и более гибкой нейронной сетью представление. Рисунок 3.8а показывает, как сетевое представление способно генерировать чувствительную-простой и точный базисный набор, способный представлять наиболее заметные возмущения. И наоборот, сферические гармоники предписывают колебательные базисные функции, которые не поддаются моделированию.- сплошные горные хребты и зоны субдукции, как показано на рисунке 3.8б. Фактически, сферический базис гармоник может оставлять непреднамеренные волновые структуры, которые скрывают важные возмущения.

Рисунок 3.8 тем самым подчеркивается один из важных выводов этого исследования: астродинамикам не нужно применять универсальную основу для каждого гравитационного поля (т. Е. Сферические гармоники); вместо этого нейронная сеть предоставляет астродинамикам возможность генерировать уникальные базовые функции, которые максимально эффективны для их задачи моделирования удельного веса.

3.1.4 Обобщение

Предыдущий эксперимент был сосредоточен на точности PINN-GM-I в сфере Бриллюэна, где гравитационные возмущения наиболее заметны. В то время как поверхность тела обеспечивает наиболее сложная динамика из-за незначительных возмущений большинство космических аппаратов работают на больших высотах. динамика имеет тенденцию к упрощению. Как показано на рисунке 3.3, возмущения имеют тенденцию к уменьшаются по величине и охватывают большие пространственные масштабы по мере увеличения высоты. Это зависит от $(R / r)^l$ срок в пределах Eq. 2.1 который быстро уменьшает вклад высокой степени / малой длины волны гармоники при больших радиусах. Предлагается второй эксперимент, в котором исследуется, насколько хорошо PINN-GM-I обобщает данные для этих больших высот, а также изучает роль распределения обучения данные о производительности сети.

Равномерное Распределение

Первая часть этого эксперимента обучает традиционные гравитационные модели и модели PINN с использованием

5 000 000 пар положения и ускорения, равномерно распределенных между 0-420 километрами. Это

представляет наиболее полный и оптимистичный набор обучающих данных для определения верхней границы

производительности. Производительность модели измеряется с использованием той же широты и долготы, что и

сетка Фибоначчи с 250 000 данными, подробно описанная в разделе. 3.1.3, однако сетка оценивается на высотах

варьируется от 0 до 500 км с шагом 10 км. Эти сетки остаются неизменными

A_m и F_m множеств,

просто оценивается на разных высотах, которые отныне будут именоваться

A_m и F_m , где

$m \in [0, 500]$ км. Затем метрика MRSE применяется к этим наборам высоты, преобразуется в

эквивалентный градус модели сферической гармоники и отображается на рисунке 3.9.

Рисунок 3.9 демонстрирует, что производительность сетей снижается в зависимости от высоты.

В среднем сети превосходят свой соответствующий сферический гармонический эквивалент для

первые 100 км, но за пределами 100 км сферические гармоники становятся более компактным представлением.

На всех тестируемых высотах сети остаются более точными, чем сферические градуса и порядка 25

гармоническая модель, предполагающая, что сети хорошо работают в области обучающих данных

и остаются жизнеспособным вариантом для моделирования гравитационного поля вплоть до высоты над уровнем моря.

Неравномерные распределения

Во второй части эксперимента исследуются характеристики PINN-GM-I при обучении.-

данные распределены неравномерно по высоте. В более реалистичных условиях полета космический аппарат

не сможет легко пролетать близко к поверхности, и образцы должны быть взяты в первую очередь из ог-

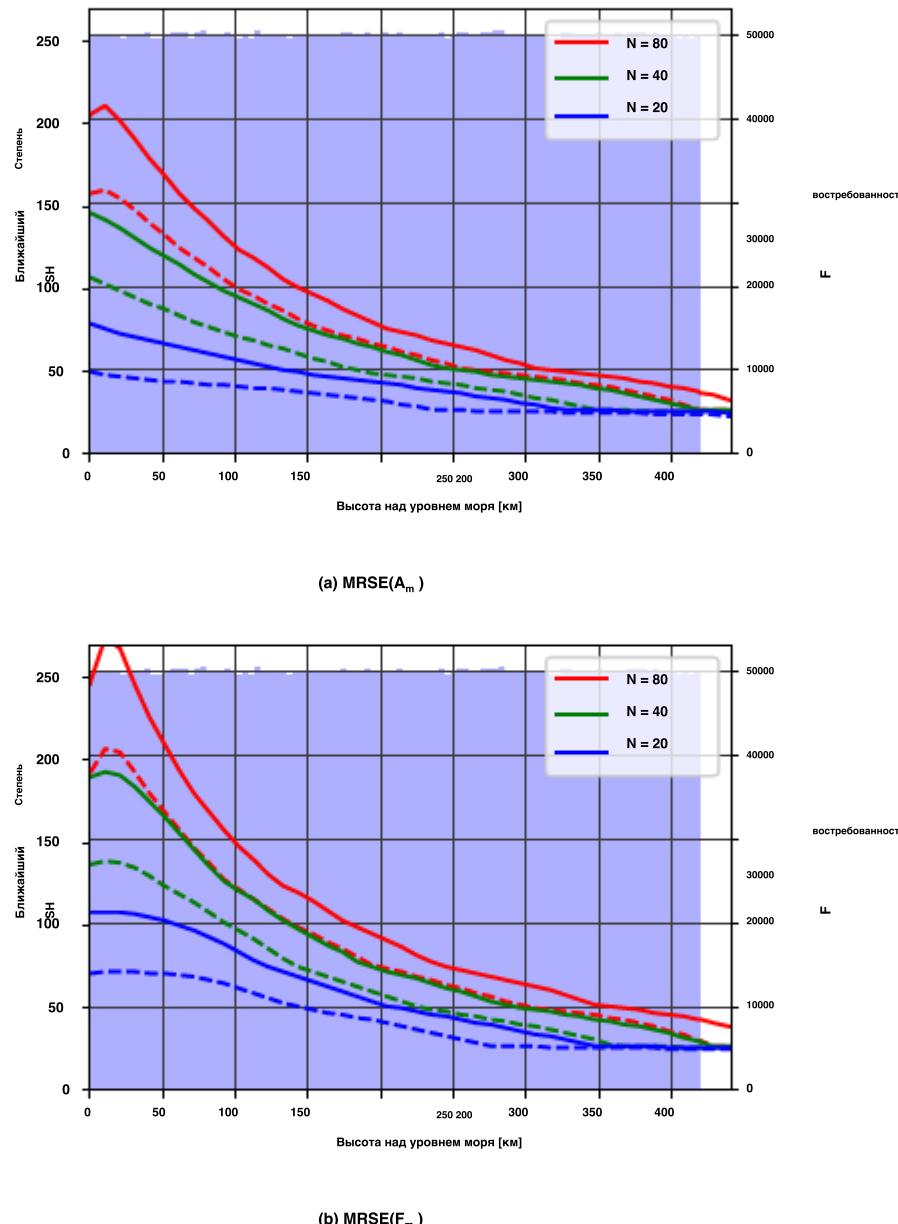
бит. Чтобы отразить эти обстоятельства, два новых набора данных из 1 000 000 пар положение / ускорение

взяты из следующего экспоненциального распределения:

$$E(x, x_0, \beta) = \text{опыт} \frac{-\beta x}{-\|x - x_0\|}, \quad (3.8)$$

где x_0 - базовая высота 420 км, а β - параметр масштабирования. Первый набор данных

устанавливает $\beta = 10$ км и моделирует распределение данных , собранных космическим аппаратом, которое начинается с высокой



A_m F_m (внизу) для традиционного (пунктирного) и основанного на физике рисунка

3.9: MRSE (вверху) и (сплошная) нейронные сети преобразуются в эквивалентную степень сферической гармоники в зависимости от высоты. Синяя гистограмма представляет распределение обучающих данных

высотной орбиты перед постепенным сходом с орбиты. Таким образом, большая часть данных будет поступать с

режим рабочей орбиты с более редкими измерениями ближе к поверхности тела. Второе распределение

устанавливает $\beta = 3$ км и представляет спутник на эксцентрической орбите, который остается на высоте рабочей орбиты и

собирает данные с нее. Вблизи поверхности могут проводиться нечастые измерения.

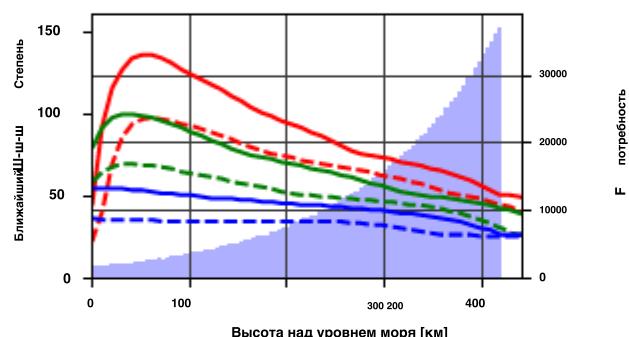
но практически все данные собираются на высотах, превышающих 200 км. Обратите внимание, что для

чтобы предотвратить выборку изнутри сферы Бриллюэна, распределение для обоих наборов данных ограничено

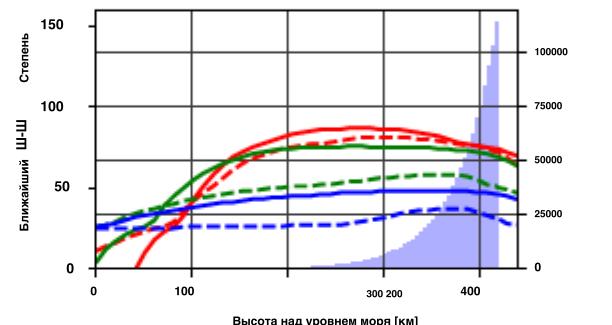
до $x \in [0, 420]$ км. Результаты для $N = 20, 40, 80$ традиционных сетей и PINN показаны

}

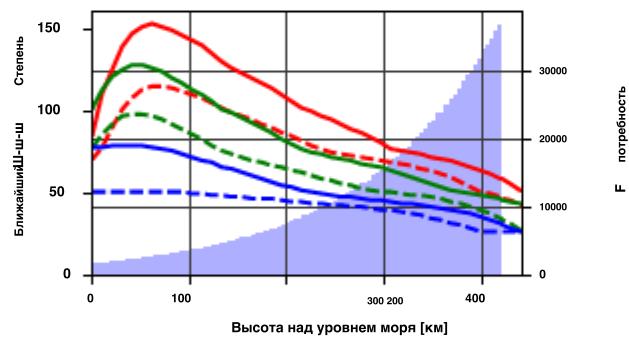
на рисунке 3.10.



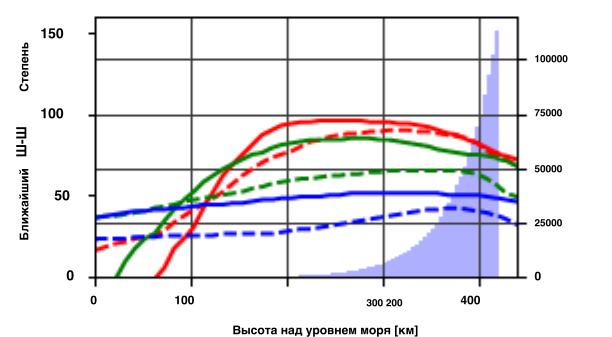
A(a) MRSE() $\beta = 10$



A(b) MRSE() $\beta = 3$



F(c) MRSE() $\beta = 10$



F(d) MRSE() $\beta = 3$

Рисунок 3.10: тренировка распределения данных и эквивалентных сферических гармоник

степени на различных АЛТИ- установки для $\beta = 3$ и $\beta = 10$ наборов данных. Сплошными

линиями обозначены точки, а пунктирными - традиционные нейронные сети

Обсуждение

Цифры 3.9 и 3.10 продемонстрируйте, что как традиционная нейронная сеть, так и PINN gravity модели наиболее продуктивны при работе с образцами на малой высоте, даже если они используются нечасто. Рисунки 3.10b и 3.10d показывают, что сетям сложно точно предсказать ускорения на Сфера Бриллюэна не предоставила никаких данных, но располагала всего лишь 2000 данными, собранными вблизи поверхности, сети достигают существенно лучшей производительности на меньших высотах, как показано на рисунках 3.10a и 3.10c.

Такая чувствительность к выборке на малых высотах подчеркивает присущую первым поколениям PINN-GM. В частности, эти модели нейронных сетей обучены минимизировать функцию затрат, которая сравнивает квадрат разницы между истинным и прогнозируемым ускорениями. Это непреднамеренно сначала смещает сеть для точного моделирования образцов на малой высоте, потому что именно на малых высотах векторы ускорений являются наибольшими. На этих малых высотах даже небольшое относительное моделирование ошибки будут вносить больший вклад в функцию затрат, чем большие относительные ошибки на больших высотах. Автор: как следствие, PINN-GM раннего поколения всегда будут отдавать приоритет этим образцам, даже если они относительно небольшая часть от общего набора данных обучения.

Несмотря на этот перекос, стоит отметить, что физика-сообщил сетей последовательно отрицательную форму их традиционные аналоги сетях в очередной раз подчеркивая их высокой эффективностью. Есть однако существуют незначительные исключения в низковысотном режиме распределения $\beta = 3$. В этих режимах сети полностью экстраполируют за пределы своих исходных обучающих данных, поэтому как правило, не рекомендуется использовать ни традиционную, ни гравитационную модели PINNA.

Эти результаты обобщения предполагают, что PINN-GM-I наиболее предпочтительна при использовании вблизи поверхности тела. Для полетов космических аппаратов наземного базирования это не является обычным режимом эксплуатации поскольку на этих высотах одно только атмосферное сопротивление создало бы большую динамическую неопределенность чем гравитационные возмущения высокого порядка. Для тел с очень тонкими или несуществующий атмосферой однако, (например, Луны или малых органов), снижение высоты орбиты являются реальными и представляют собой реальную использовать чехол для сети представительства. Последующие поколения PINN-GM исследуют способы устранения этой проблемы зависимости от высоты, как будет показано в главе 3.3.

3.1.5 Скорость вычислений

Последний эксперимент для PINN-GM-I исследует, насколько быстро обученные сетевые модели

могут быть выполнены по сравнению с другими популярными представлениями о гравитации. Гравитация нейронной сети

модели записываются в Tensorflow 2.4 и выполненных на NVIDIA на RTX 2060 GPU и на

Ryzen 3400G для GPU и процессора случаев соответственно. Рисунок 3.11 показывает время выполнения повторно-

требуется оценить ускорения 10 000 случайно распределенных данных о местоположении для гравитационных моделей

с увеличением числа параметров. Сначала сравниваются характеристики PINN и spheri-

каль-гармоническое представление, а затем к многогранной гравитационной модели. Сферические гармоники

Каждое представление генерируется на основе модели EGM-2008 с разной степенью усечения.

В полиздрической модели протестированы использовать все более и более ухудшаться модели форму 433-Эрос, созданный с помощью

Блендер². Два аналитических представления написаны на Python, скомпилированы точно в срок с использованием

Numba³, и выполняется на процессоре Ryzen 3400G.

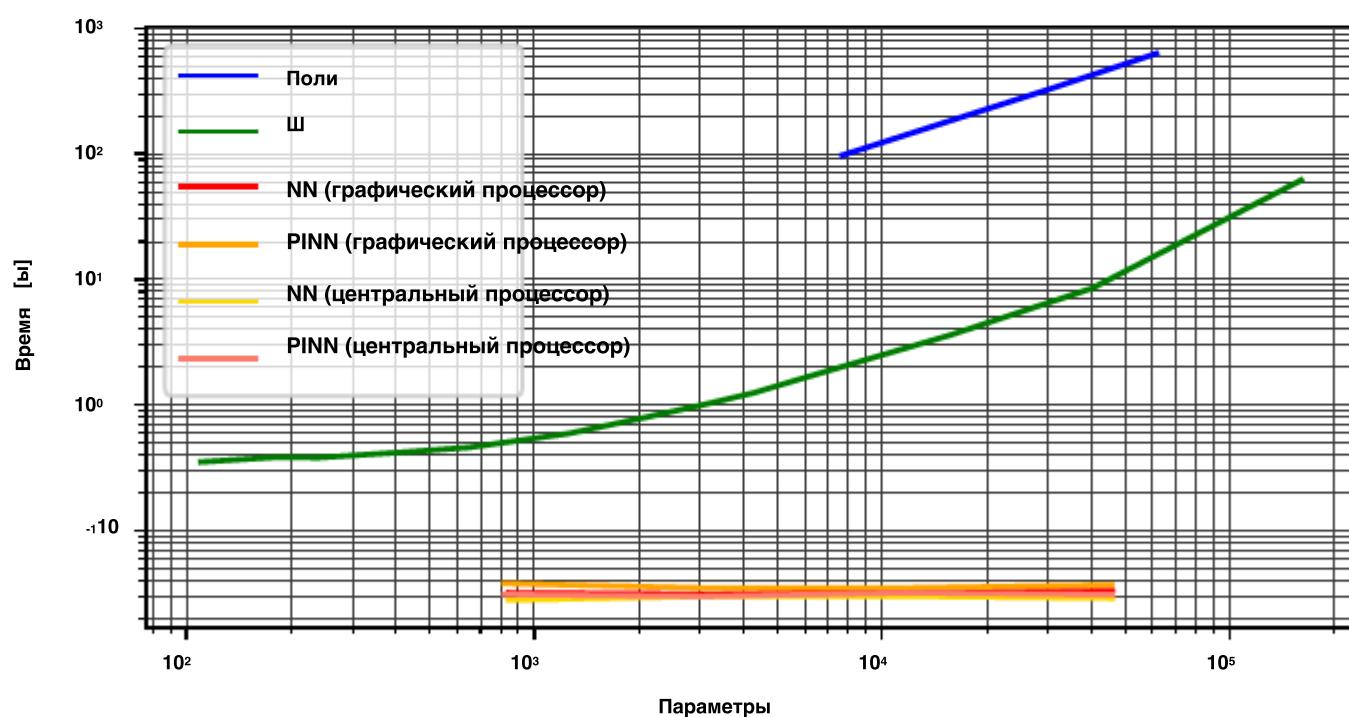


Рисунок 3.11: Общее время оценки 10 000 случайных данных с использованием различных гравитационных моделей

² <https://www.blender.org>

³ <https://numba.pydata.org/>

Рисунок 3.11 демонстрирует, что представление сферических гармоник относительно эффективно

для оценки с низкой степенью, но она также имеет самый крутой градиент, поскольку в нее включено больше параметров

модель. Это проверяет сферические гармоники.

$O(n^2)$ сложность. Многогранное представление - это

безусловно, оценка занимает больше всего времени - почти на два порядка больше, чем у модели сферических гармоник, оснащенной тем же количеством параметров.

Нейронная сеть использует тот же набор параметров, что и нейронная сеть. Нейронная сеть

сетевое представление, выполняемое либо на графическом процессоре, либо на центральном процессоре, значительно эффективнее обоих из них

эти представления - с производительностью, которая почти на порядок выше, чем

протестированная модель с наименьшей сферической гармоникой ($l = 10$), не зависящая от количества параметров

с помощью этих сетей. Эти результаты обнадеживают как для использования в моделировании, так и на борту

космический аппарат, производящий высокоточные оценки гравитационного поля, которые могут быть использованы для определения траектории движения

проектирование или бортовой контроль.

3.1.6 Производительность сети Применительно к гравитационному полю Луны

Представленные на данный момент эксперименты сосредоточены конкретно на Земле, но выводы

о компактности модели и ее обобщении не обязательно универсальны для любого небесного тела, чье

гравитационное поле традиционно представлено сферическими гармониками. Чтобы продемонстрировать это, предыдущие эксперименты

повторяются для тела с характерно отличающимися возмущениями: Луны.

Луна предлагает интересную точку сравнения с Землей, поскольку гравитационный коэффициент-

колебания Луны происходят значительно чаще и большей величины. Земные

возмущения обычно генерируются крупными, нечастыми и локализованными геологическими структурами (горами-

отложения, границы тектонических плит и т.д.). Возмущения Луны, напротив, генерируются

кратерами и связанными с ними масконами, которые покрывают большую часть ее поверхности. Как будет показано в следующем примере.-

возможно, сложность гравитационного поля Луны затрудняет эффективное моделирование таких возмущений

более сложной задачей для PINN-GM-I.

Рисунок 3.12 цель - сопоставить различия между двумя телами и их гравитационными полями.

Качественно обратите внимание, насколько проще гравитационное поле Земли по сравнению с полем Луны.

Кратеры на Луне не только образуют почти случайную топологию поверхности, но и переплетаются

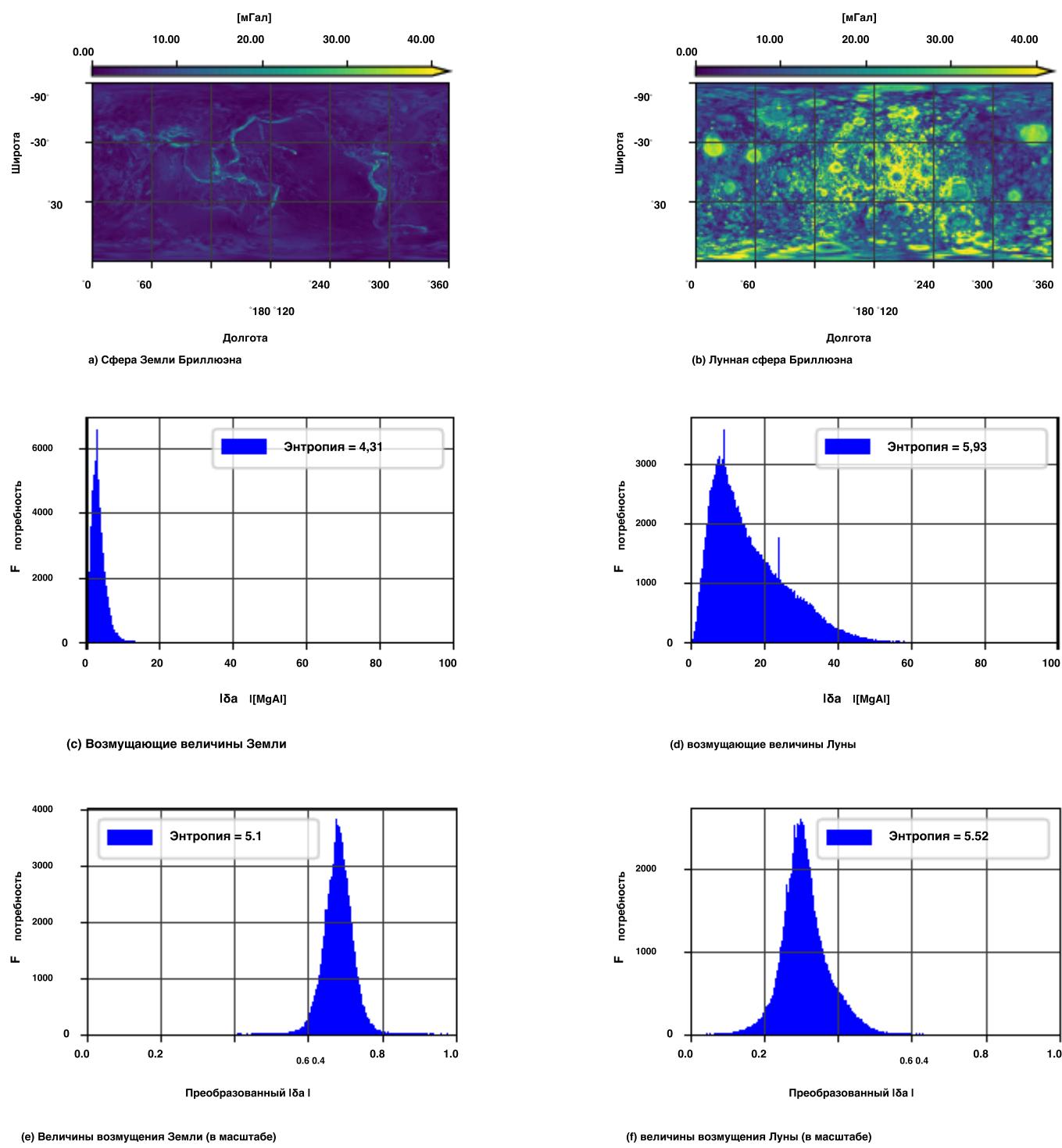


Рисунок 3.12: Контрастное распределение гравитационного поля и ускорений Земли и Луны

некоторые из самых сильных возмущений по сравнению с самыми слабыми. Возмущения Земли в сравнении, гораздо более исправны. Сложность двух полей определяется количественно путем оценки Энтропии Шеннона распределений ускорений, показанных на рисунках 3.12c и 3.12d. Для Земли, энтропия поля составляет 4,31 натс (единица измерения информации, выраженная в базе e), тогда как энтропия для Луны составляет 5,93 натс. Это говорит о том, что гравитационное поле Луны содержит примерно в пять раз больше информации в среднем больше информации о гравитационном поле Земли - значительно более сложная задача задача моделирования для нейронных сетей. Эти распределения предварительно обрабатываются, прежде чем использовать их в качестве обучающие данные для сетей, поэтому рисунки 3.12e и 3.12f показывают, как изменяются распределения после того, как к векторам ускорения было применено равномерное преобразование min-max, как подробно описано в разделе. 3.1.1.

После предварительной обработки гравитационное поле Луны по-прежнему содержит больше информации по сравнению с Землей хотя и всего на 50%. Это предполагает, что как в необработанном, так и в предварительно обработанном виде гравитационное поле Луны гравитационное поле - более сложная задача моделирования.

Компактность

Точность модели PINNA-GM-I оценивается как функция параметров и сравнивается с сферической гармонической гравитационной моделью на рисунке 3.13. Рисунок воспроизводит тот же эксперимент из раздела. 3.1.3, оценивая ошибку только по доминирующему гравитационным возмущениям, F , а также по средним и фоновым данным (A и C). Вместо обучения сетей по полевым точкам, нарисованным равномерно на высоте 0-420 км, обучающие данные охватывают диапазон высот 0-50 км. Эта более узкая область обучения пытается отразить тот факт, что космические аппараты могут вращаться на меньших высотах вокруг тел, подобных Луне, учитывая ее очень тонкую атмосферу. Кроме того, истинный потенциал, Y_{TruthSH} создан с использованием модели гравитационного поля Луны GRGM1200A (54).

Сферическая гармоническая гравитационная модель значительно более эффективна при представлении Гравитационное поле Луны больше, чем у Земли, как показано на рисунке 3.14. Для Земли, почти 10 000 коэффициенты сферических гармоник требуются, чтобы начать сходиться к набору характеристик F , тогда как для Луны требуется всего 300 коэффициентов сферических гармоник. Это говорит о том, что длинноволновые гармоники играют большую роль в представлении гравитационных особенностей Луны, что является удивительным фактом.-

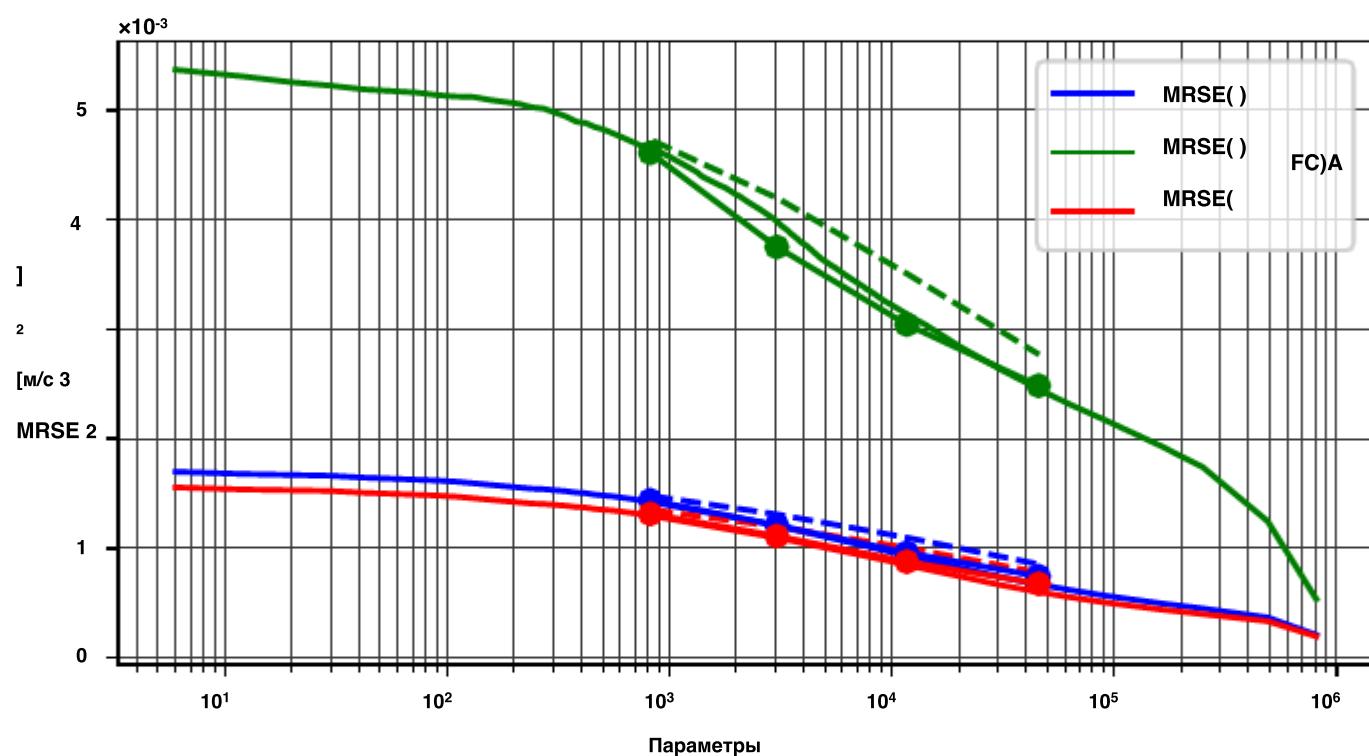


Рисунок 3.13: График MRSE как функции общих параметров модели для Луны. Пунктирные линии представляют традиционные нейронные сети. Линии с круговыми маркерами представляют собой нейронные сети, основанные на физике

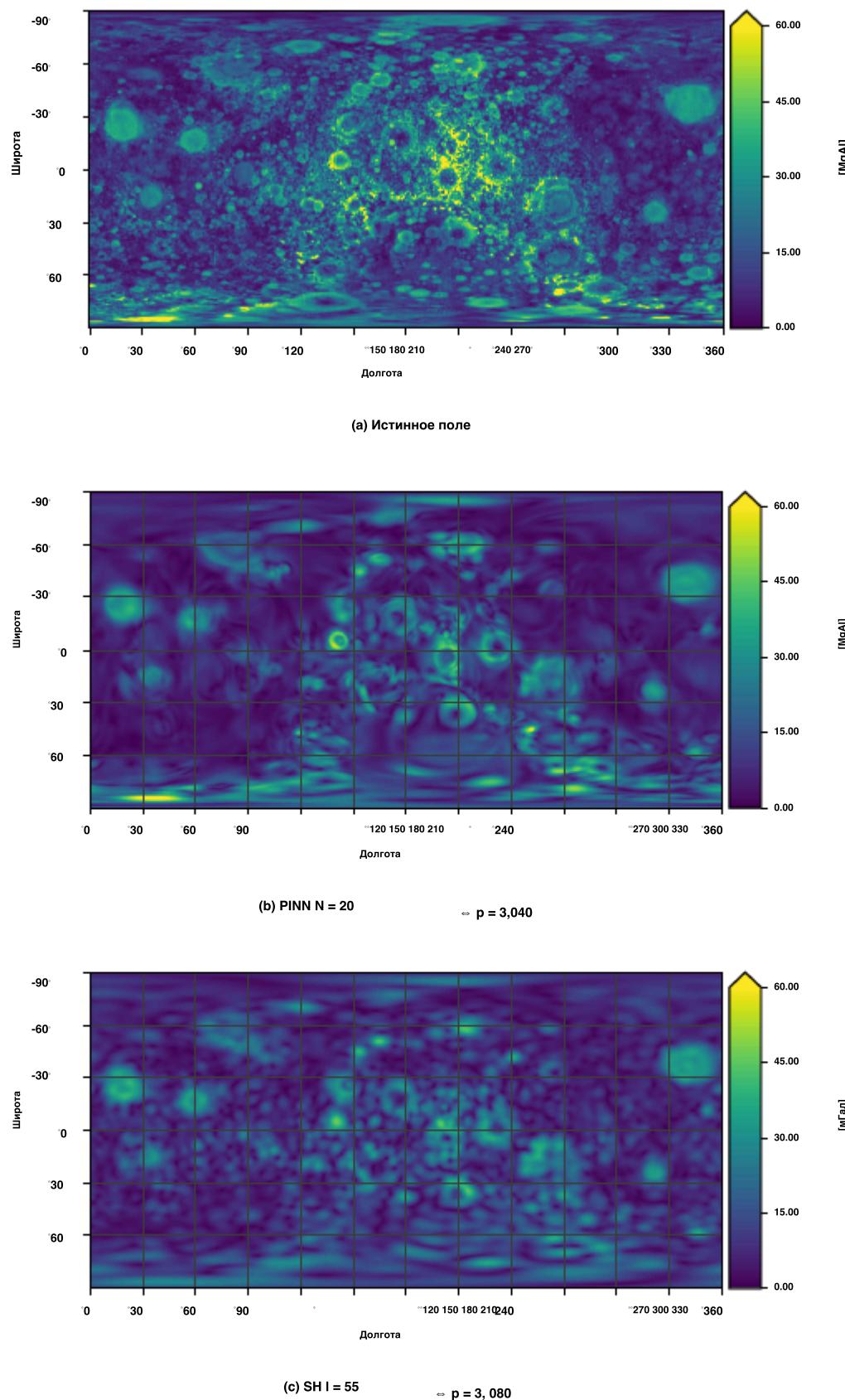


Рисунок 3.14: Гравитационная модель Луны с использованием (а) полной сферической гармонической модели $I =$

1000, (б) представления ПИННА с $p = 3\,040$ и (в) представления сферической гармоники с низкой точностью $I = 55 \Leftrightarrow p = 3080$

учитывая резкие прерывистые особенности, присутствующие в поле Луны. Эффективность этих гармоник низкой степени объясняется почти изотропной природой возмущений. Даже если гармоники низкой степени не могут эффективно уменьшить погрешность отдельного прерывистого объекта, такого как кратер. Они вносят небольшой вклад во все кратеры, покрывающие всю поверхность. Это коллективный вклад приводит к ощутимому снижению погрешности, наблюдаемой в наборе функций,

F

Тем не менее, PINN-GM-I по-прежнему имеет небольшое преимущество перед сферическими гармониками с точки зрения компактность модели, позволяющая получить сопоставимые оценки гравитации с использованием примерно на 30% меньшего количества параметров в лучшем случае, чем сферические гармоники. Относительно небольшой прирост производительности является интересным результат объясняется двумя факторами: во-первых, вышеупомянутой эффективностью длинноволнового диапазона в сферическом гармоническом представлении. Аналитическая модель просто более эффективна для Луны чем он был для Земли, так сетях конкурируют с более продуктивные представления.

Второй вклад в расхождение показателей является большей сложностью / высшее энтропия гравитационного поля Луны. Сетям поручено генерировать дистрибутив который содержит больше информации, чем присутствовало в поле притяжения Земли, т.е. Сетям приходится работать усерднее. Остается возможным, что существует более высокая производительность сети, но наблюдать такая производительность сетей, вероятно, потребует дополнительных обучающих данных, дополнительных функций проектирования и более длительного времени обучения, чтобы компенсировать эту более сложную проблему.

Требуется дальнейшая работа, чтобы определить, существует ли более фундаментальная взаимосвязь между энтропия гравитационного поля и способность к успешной нейросетевой модели гравитации. Потому что среды с высокой гравитационной энтропией будут генерировать ускорения, которые кажутся неструктурированными или случайная, вполне возможно, что все базисных наборов (научен или аналитический) будет равной борьбе моделирования типы полей.

Обобщение

Следуя той же процедуре, представленной в разд. 3.1.4, производительность PINN-GM-I тестируется на различных высотах, чтобы определить, как эти сети распространяются на разные орбиты. В Тестируемые высоты варьируются в пределах 0-55 км и MRSE для каждого набора данных A_m , F_m и C_m показано

на Рисунке 3.15.

Общие характеристики PINN-GM-I для Луны в точности повторяют характеристики земных сетей

. Как традиционные, так и основанные на физике нейронные сети, как правило, обладают наибольшей точностью.

на малых высотах, что еще раз подчеркивает внутреннюю склонность этих моделей к выборкам на малых высотах. Это

говорили, существует менее выраженной пиковой производительности PINN-гроссмейстеры, представляющие Луне, чем

те земли. Эта разность отнесена к высокой энтропии Луны гравитационное поле. Поля

большую сложность, в конечном счете, выступает как тип регуляризации, которая помогает предотвратить сети

переобучение к обучающим данным. Возмущения настолько разнообразны и постоянно присутствующим, что в сети

не развивайте чрезмерную уверенность в их изученном решении. Аналогичные эффекты наблюдаются при использовании

небольших мини-пакетов для обучения нейронных сетей. Эти небольшие пакеты могут создавать более шумный градиент.

снижение, которое требует больше времени для сближения, но помогает оптимизатору в изучении структуры затрат

и, в конечном итоге, может создавать более надежные решения(86). Разница здесь заключается не в уменьшении, а в

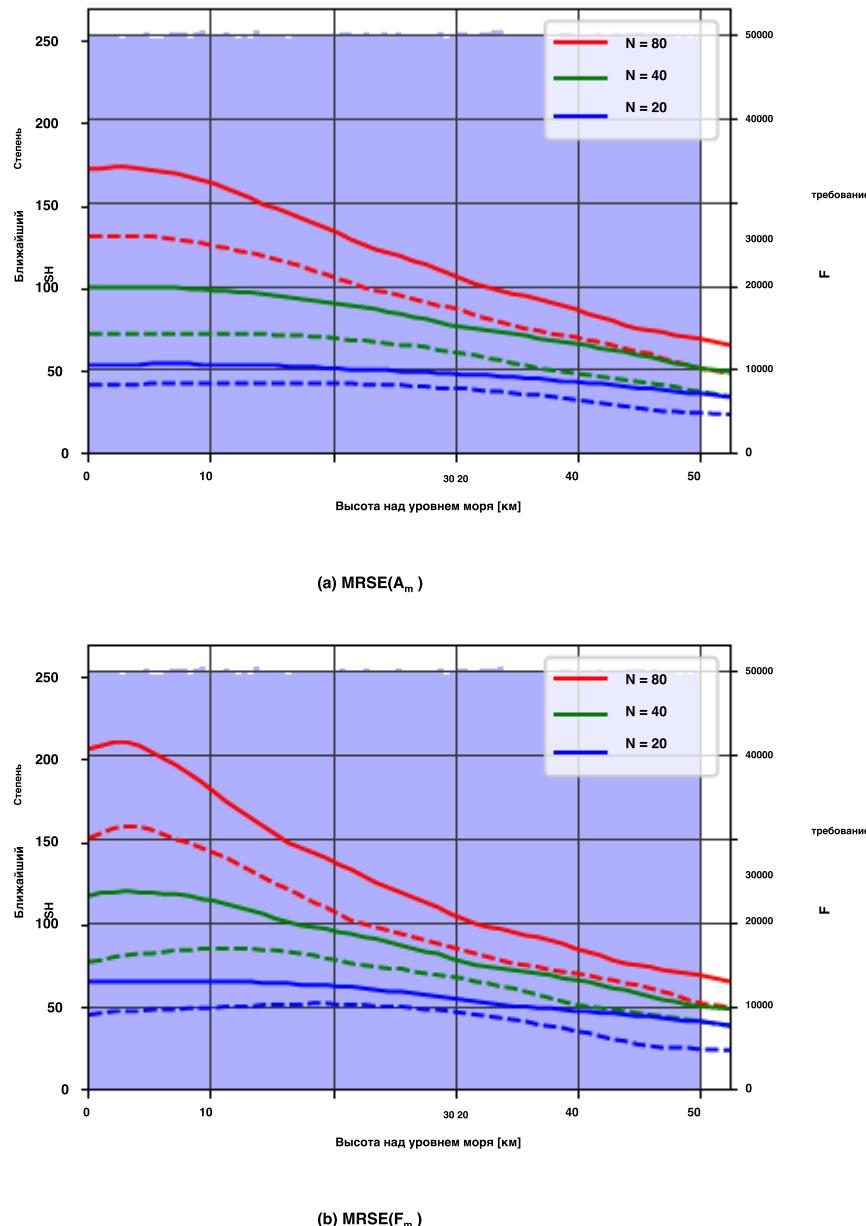
размер партии, сложность структуры затрат увеличиваются для достижения аналогичного эффекта.

Подводя итог, PINN-GM-I демонстрирует некоторые преимущества перед сферическими гармониками, когда

представляющий гравитационные поля крупных небесных тел, таких как Земля и Луна. Среди

Наиболее убедительных из этих преимуществ является быстрая выполнимость этих моделей по сравнению с их аналитическими

аналоги, а также их модели меньшего размера и памяти.



3.15: MRSE (вверху) и (сплошная) нейронные сети преобразуются в эквивалентную степень сферической гармоники в зависимости от высоты Луны. Синяя гистограмма представляет распределение обучающих данных

3.2 Поколение II

Хотя результаты PINN-GM-I дают обнадеживающий старт использованию PINN для представления поля гравитации крупных небесных тел, остаются открытые вопросы относительно квантов и ограничений гравитационной модели PINN особенно в контексте малого-тело разведка.

Например, являются ли эти сетевые представления одинаково надежными при моделировании гравитационного поля несферических небесных тел? Продолжают ли сети генерировать точные потенциалы и ускорения внутри сферы Бриллюэна? Какие условия для данных обучения необходимы для создания надежной модели и насколько чувствительны эти сети к шуму в данных обучения?

Кроме того, было проделано мало работы по обсуждению того, как эти сети могут быть спроектированы и настроены для повышения производительности. В справочнике (87), только одно физическое ограничение используется в функции стоимости ($a = -\nabla U$), однако могут быть введены другие ограничения. Существуют ли такие ограничения способствовать обобщению или многоцелевой характер дополнительных ограничений препятствует процессу обучения? Исключает ли плотно связанная архитектура сети возможность сети для изучения абстрактных симметрий и законов сохранения, как предложено в Ссылке (81)?

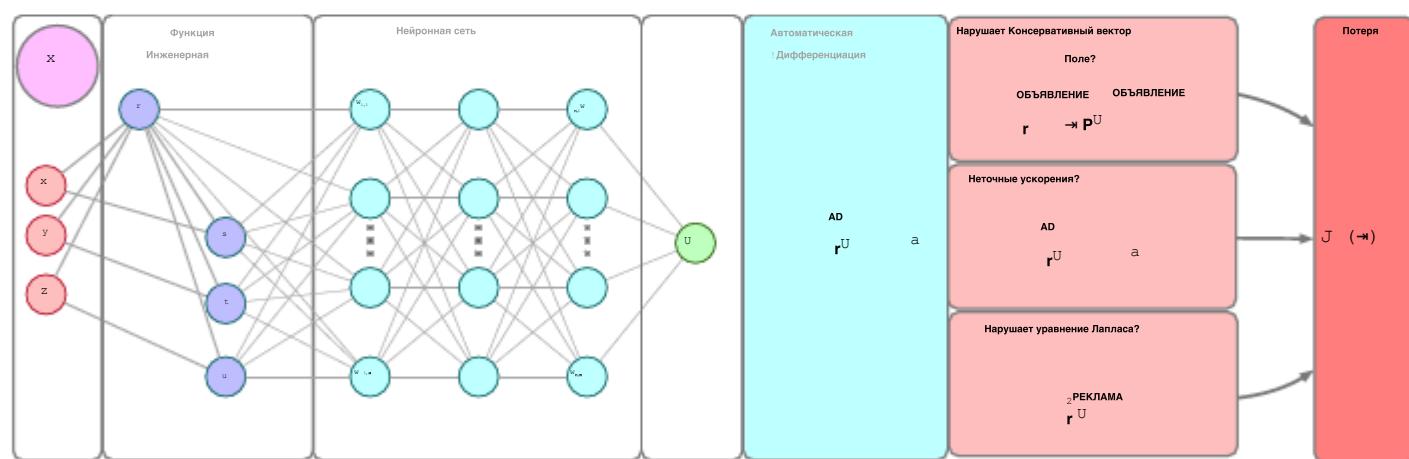


Рисунок 3.16: PINN-GM поколения II (1)

В этом разделе представлен PINN-GM поколения II, или PINN-GM-II, который предлагает обзор новых модификаций конструкции оригинального PINN-GM, направленных на решение этих вопросов. В частности, этот раздел посвящен моделированию гравитационного поля малых тел, исследуя, как PINN-GM может быть спроектирован для создания высокоточных моделей гравитационного поля астероида 433-Эрос. Модель

подробно описаны конструктивные изменения PINN-GM-II и их соответствующее влияние на производительность ниже.

3.2.1 Нормализация

Во многих приложениях машинного обучения, данные, используемые для обучения нейронных сетей, нормализованы чтобы гарантировать, что веса и отклонения, полученные сетью, не имеют чрезмерно большой величины, с целью избежать числовых проблем (88). В PINN-GM первого поколения обучающие данные нормализуются с помощью преобразования min-max для каждого компонента ускорения таким образом, чтобы все обучающие данные существовали в диапазоне от -1 до 1. Хотя этот подход эффективен, на самом деле он не гарантирует числовую стабильность выходных данных сети. Нормализация положения и ускорения до [-1,1] не обеспечивает гарантирует диапазон потенциальных, фактических выходных данных нейронной сети. Чтобы гарантировать что числовая стабильность выходных данных сети является приоритетной, PINN-GM-II вводит другую безразмерность посредством:

$$x^* = \frac{x - x_0}{x_s} \quad (3.9)$$

$$U = \frac{U - U_0}{U_s} \quad (3.10)$$

$$a^* = \frac{\dot{x} dU}{U dx} \quad (3.11)$$

(3.12)

где x_0 соответствует минимальному значению всех компонентов положения и x_s соответствует диапазону вектора положения. U_0 и U_s являются минимумом и диапазоном возможного.

3.2.2 Проектирование объектов

В PINN-GM-I входными данными сетей являются данные о местоположении, представленные в декартовых координатах. Хотя этот подход по своей сути не является проблематичным, область каждого немасштабированного объекта (x, y, z) простирается от $(-\infty, \infty)$, заставляя сеть размещать существенно более широкое пространство объектов, чем если бы оно было представлено в сферических координатах. PINN-GM-II изменяет эту конструкцию, вместо этого переводя-формирование входных данных PINN в 4D сферическом наборе координат r, s, t и u перед их передачей

через сеть. Эти переменные определяются как:

$$\begin{aligned} x^2 + y^2 + z^2 &= r^2 \\ x &= r \cos \theta \sin \phi \\ s_r &= \frac{r}{\sqrt{x^2 + y^2}} \\ t &= \frac{y}{\sqrt{x^2 + y^2}} \\ u &= \frac{z}{\sqrt{x^2 + y^2}} \end{aligned}$$

где s , t и u представляют касательные углы между точками поля и каждой из декартовых точек

единичные векторы x^{\wedge} , y^{\wedge} и z^{\wedge} (84).

Этот 4D-сферический набор координат выбран вместо более традиционного 3D-набора (r , θ и ϕ) по

нескольким причинам. Во-первых, значения s , t и u , естественно, находятся между желаемыми границами

[−1, 1] принимая во внимание, что координаты θ и ϕ трехмерного набора имеют разрыв из-за переноса угла.

Во-вторых, набор координат 4D позволяет избежать сингулярности в градиенте потенциала на полюсах.

Это можно увидеть при оценке трехмерного сферического градиента при $\phi =$

-90° и 90° как показано на рисунке:

$$\nabla U = \frac{\partial U}{\partial r} + \frac{1}{r} \frac{\partial U}{\partial \phi} + \frac{\partial U}{\partial \theta} \quad \text{так как } \phi \text{ от 0 до } \pi \quad (3.13)$$

Это инженерная функция, которая преобразует декартово положение в сферическую координату 4D.-

nates выполняется в рамках официального графика Tensorflow в качестве дополнительного слоя к нейронной сети

Модель. Интегрируя это преобразование в модель, это означает, что потенциал все еще может быть увеличен.-

математически дифференцированный по отношению к декартовым входным данным. Эта конструкция обеспечивает лучшее из

обоих наборов координат: нейронная сеть, которая обучается на значительно уменьшенном пространстве объектов (с увеличением-

(для определения сходимости и точности сети) при сохранении производных в удобном декартовом виде

базис.

3.2.3 Дополнительные ограничения PINN

В дополнение к изменению входных функций, PINN-GM-II вводит модифицированную функцию потерь

которые включают дополнительные ограничения, основанные на физических данных, для повышения точности моделирования и надежности-

несс. Очевидно, что функция потерь, предложенная для PINN-GM-I, наказывала только за ошибки в прогнозируемом

ускорении (обозначенное нижним индексом \dot{A} в уравнении 3.3). Хотя эта функция потерь эффективна, она не учитывает

другие динамические свойства, которые должны соблюдаться гравитационной моделью. Например,,

сила тяжести является консервативной силой - следовательно, скалярный потенциал, изучаемый сетью

должен также подчиняться этим дополнительным физическим свойствам:

$$\nabla^2 U = 0 \quad \text{Лапласиан (L)} \quad (3.14)$$

$$\nabla \times a = 0 \quad \text{Завиток (C)} \quad (3.15)$$

Эти консервативные свойства векторного поля могут быть встроены в исходную функцию потери связи-

посредством:

$$J_{ALC}(\Theta) = \frac{1}{N} \sum_{i=1}^N \|a_i - \nabla f(x_i | \Theta)\|^2 + \frac{\lambda_{AD}}{2} \|\nabla f(x_i | \Theta)\|^2 + \frac{\lambda_{PROMO}}{2} \|\nabla_{PROMO} \times \nabla_{OBYAVLENIE} f(x_i | \Theta)\|^2 \quad (3.16)$$

Кроме того, если существует модель гравитационного потенциала, функция потерь может быть расширена до

оштрафовать эти ошибки моделирования (нижний индекс P):

$$J_{APLC}(\Theta) = J_{ALC}(\Theta) + \frac{1}{N} \|U_p - f(|\Theta|)\|_2^2 \quad (3.17)$$

Эти дополнения к функции затрат должны действовать как форма регуляризации во время обучения, которая

может помочь повысить надежность, особенно при наличии шума.

3.2.4 Модифицированные сетевые архитектуры

Помимо включения дополнительных физических ограничений в функцию сетевых потерь, PINN-

GM-II также представляет альтернативную сетевую архитектуру. Для справки (81), замечено, что

успех многих современных моделей машинного обучения основан на встраивании групп симметрии или

инвариантности в сетевых архитектурах (будь то трансляционные инвариантности в сверточных сетях

или временные инвариантности в рекуррентных нейронных сетях). Внедрение этих свойств в сеть

архитектуры для физических процессов могут быть сложными, поскольку они предполагают наличие ранее известных знаний об этих группах

группы симметрии, которые могут отсутствовать. По этой причине обычно полагаются только на полностью

подключенные сети при моделировании физических процессов в отличие от более продвинутых архитектур.

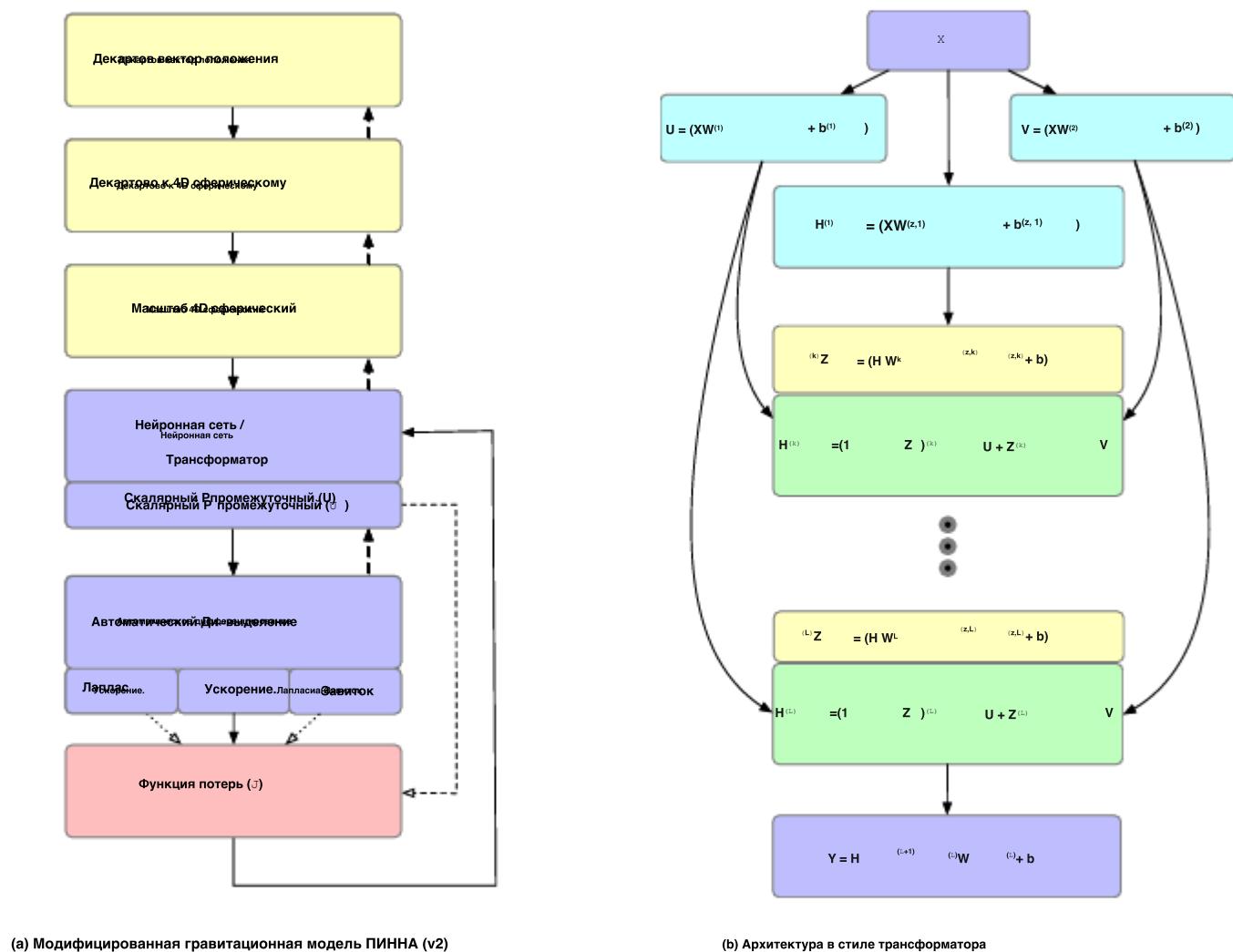


Рисунок 3.17: Модифицированная структура сети

Ссылка (81) подразумевает, что этот выбор может быть недостаточным, поскольку пока не существует доказательства сходимости полностью связанных нейронных сетей, основанных на физике. Как следствие, исследователям остается только надеяться, что полностью связанные сети обеспечат достаточные возможности моделирования и гибкость представления решения соответствующего дифференциального уравнения. Ссылка (81) поддерживает дополните это утверждение, представив, как относительно простое изменение сетевой архитектуры может сигнализировать- значительно повысить точность изученной модели в наборе архетипических задач PINN.

В этой работе используется предложенная архитектура за счет расширения полностью подключенной сети двумя кодерами, проект которого входит в высоком мерном пространстве. Кроме того, скрытые состояния усиливаются при остаточных соединений с предыдущих слоев путем мультипликативной, поэлементное возведение в- teractions как показано на рисунке 3.17b. Эти модификации вносят относительно незначительные изменения в объем памяти и вычислительных ресурсов полностью подключенной сети, но обеспечивают заметный прирост в точности модели. Это изменение архитектуры будет называться преобразованием с учетом физики. (PIT) для оставшейся части этого раздела.

3.2.5 Производительность

Чтобы исследовать влияние этих конструктивных изменений на производительность, предлагается серия экспериментов . Эти эксперименты выполняются путем последовательного внесения различных изменений в конструкцию , представленных выше, и характеризующих производительность модели на каждом промежуточном этапе.

Для начала устанавливаются базовые показатели для оценки того, насколько хорошо работает гравитационная модель ПИННА. изучить гравитационное поле астероида 433-Эрос на основе оценок положения и ускорения. Эти эксперименты разработаны в предположении, что существует уже существующая высокоточная гравитационная модель (будь то сферические гармоники, многогранники, масконы или другие), из которых могут быть получены точные или зашумленные измерения. могут быть получены и использованы в качестве обучающих данных для создания представления ПИННА. Для этих экспериментов, предполагается, что многогранное представление является основной истиной и использует Glaskell⁴

модель формы. Предлагаются три уникальных набора данных для характеристики влияния количества данных и распределения на производительность модели.

⁴ <https://arcnav.psi.edu/urn:nasa:pds:gaskell.ast-eros.shape-model>

Равномерное распределение 0 $-3R \leq r \leq R$

При первом обучающем распределении данные отбираются равномерно от поверхности астероида до $\pm 3R$, в три раза превышающего максимальный радиус астероида, R (рисунок 3.19a). Эта выборка представляет наилучший сценарий, при котором сеть обучается на данных, охватывающих всю проблемную область (внутри сферы Бриллюэна, на уровне сферы Бриллюэна и выше нее).

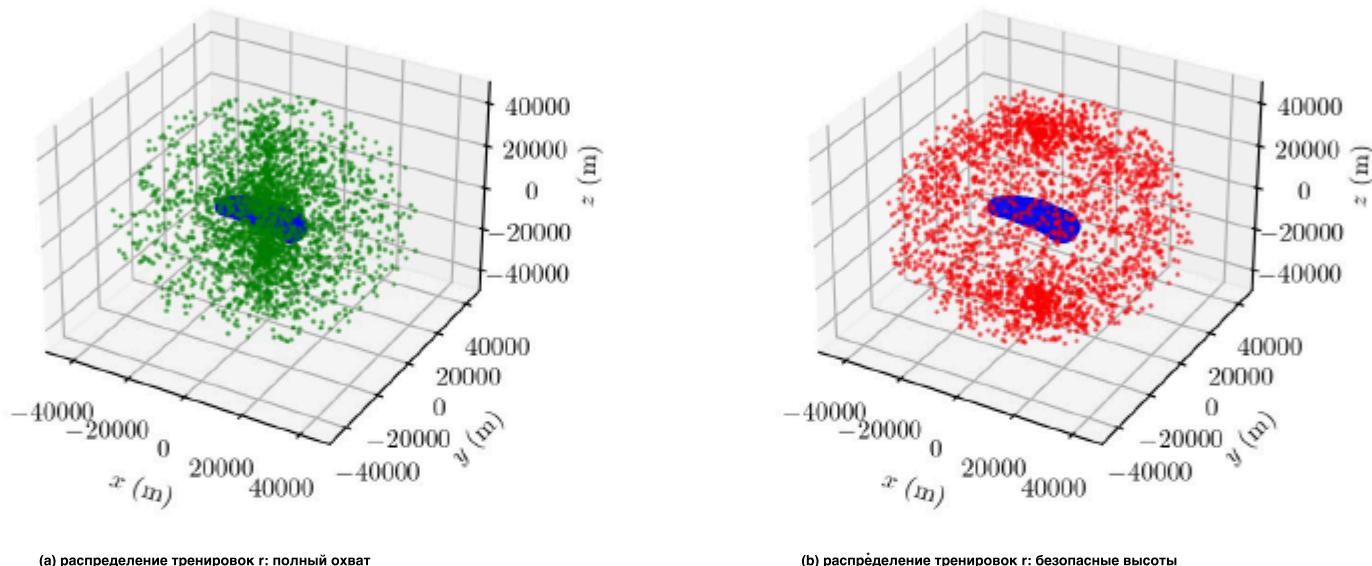


Рисунок 3.18: Тренировочные распределения Астероид 433-Эрос

Равномерное распределение $2R$ $-3R \leq r \leq R$

Второе распределение обучения отражает более реалистичные обстоятельства, когда выборки собраны равномерно, но только с диапазоном безопасных высот над астероидом (в 2-3 раза превышающих максимальный радиус). Такое распределение показано на рисунке 3.19b. Эти высоты, как правило, подходят для оценки сферических гармонических моделей низкой степени или сбора изображений, которые могут быть использованы для разрешения модели формы для использования в многогранной гравитационной модели.

Гибридное распределение 2R**3R плюс разреженные выборки на малой высоте (r)**

Окончательное обучающее распределение формирует гибридный набор из двух предыдущих распределений. В

Большинство выборок собирается из 2R

-3R (r^*) распределение, но 10% приходится на промежуток между

поверхностью и 2R. Эти более редкие образцы могут представлять два потенциальных явления: активный астероид

выброс или передовые концепции миссий, такие как гравитационные попперы. Как недавно было обнаружено в радиопередаче OSIRIS-REx

научный эксперимент, некоторые астероиды выбрасывают частицы, которые можно отследить с помощью оптических измерений

и их траектории, по оценкам, используются для восстановления гравитационного поля (89; 90). Эти частицы сильно

улучшенное разрешение исходного радио наука возвращает Осириса-Рекса и, следовательно,

мотивированное новая технология миссия под названием Гравити попперс (91). Гравитационные попперы - это технология

демонстрация, которая запускает искусственные, неконтролируемые зонды с корабля-носителя на поверхность

небольшого тела. Затем эти зонды отрываются от поверхности астероида, и результирующая динамика

используются для помощи в оценке гравитационного поля тела.

Экспериментальная установка

Эти эксперименты начинаются с обучения PINN-GM-Is только новым правилам нормирования и fea-

применены инженерные модификации. Из каждого из вышеупомянутых выборок отобрано 5000 обучающих данных.-

распределенные распределения используются для обучения соответствующей гравитационной модели PINNA за 7500 эпох. Это

затем процесс повторяется с добавлением искусственной ошибки к измерениям ускорения для создания

более реалистичного, зашумленного набора данных, который может быть получен в процессе определения орбиты с использованием

компенсации динамической модели. Эти зашумленные оценки генерируются путем добавления вектора ошибки с

случайная ориентация и величина, равные 20% от истинной величины вектора ускорения. Следовательно, ошибка

негауссова, где чем больше истинное ускорение, тем больше соответствующая ошибка

. Остальные гиперпараметры, используемые для обучения этих сетей, перечислены в таблице 3.3.

После подготовки каждой модели, 20,000 тестовых образцов отбираются случайным образом с поверхности

тела на высоту, соответствующую радиусе 3R. Остатки между истинным и прогнозируемым значениями

ускорения в местах тестирования нанесены на график рядом с их скользящим средним. Кроме того, два

Таблица 3.3: Номинальные гиперпараметры

Параметр	Значение	Параметр Инициализации	Значение
Активация	GELU	веса	8
Оптимизатора	5000	скрытых слоев	20
Размера пакета	Adam	узлов на слой	Норма Славы
Эпохи	7,500	Скорость обучения	0.002

сферические гармонические модели (одна степени и порядка 4, другая степени и порядка 8) соответствуют

тем же 5000 обучающим данным с использованием регрессии методом наименьших квадратов и тестируются на тех же 20 000 тестовых выборках

. Сферическая гармоническая модель остатки и средняя ошибка также построены, чтобы служить

точка сравнения. Эти кривые представлены на соответствующее обучение распределения данных

гистограммы на рис. 3.19.

Исходные результаты

В лучшем случае (г без шума) модели PINN обеспечивают ускорение с

средней погрешностью менее 3% на всем пути до поверхности - обнадеживающий результат, учитывая, что

сети было предоставлено всего 5000 точек данных. Когда в оценки вводится ошибка ускорения в 20%.

ошибка PINN также увеличивается, но остается ниже среднего значения в 10%.

регресс сферических гармоник моделей, напротив, никогда не достигают в среднем более чем на 5% об ошибке -

независимо от количества ошибок в оценках - и видно расширяются по мере приближения к

более низких высотах.

В случае тренировочных пинов на r^* , изученные представления достигают средней ошибки <0,1%

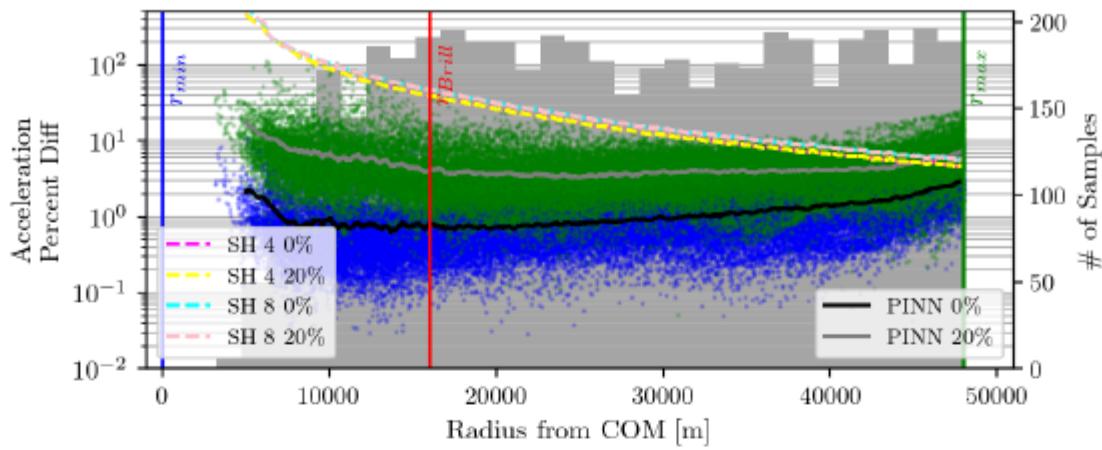
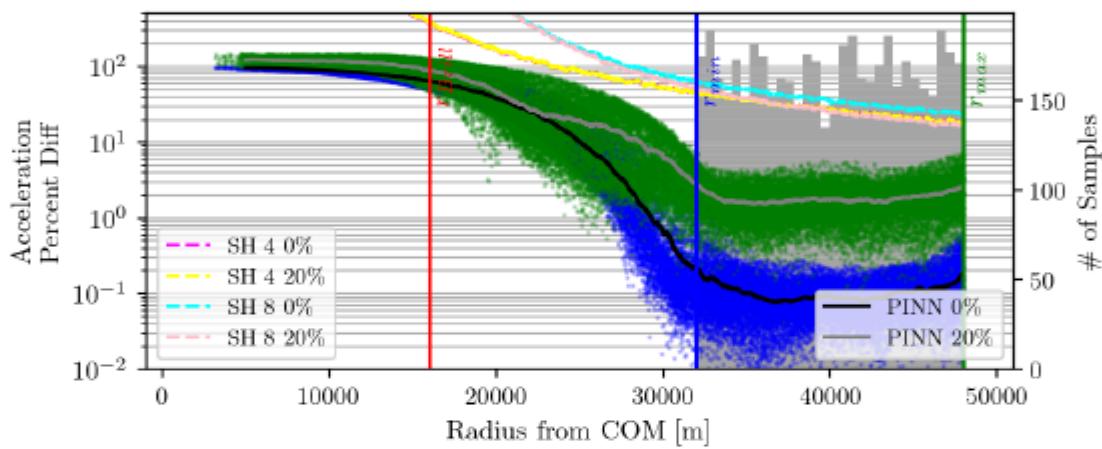
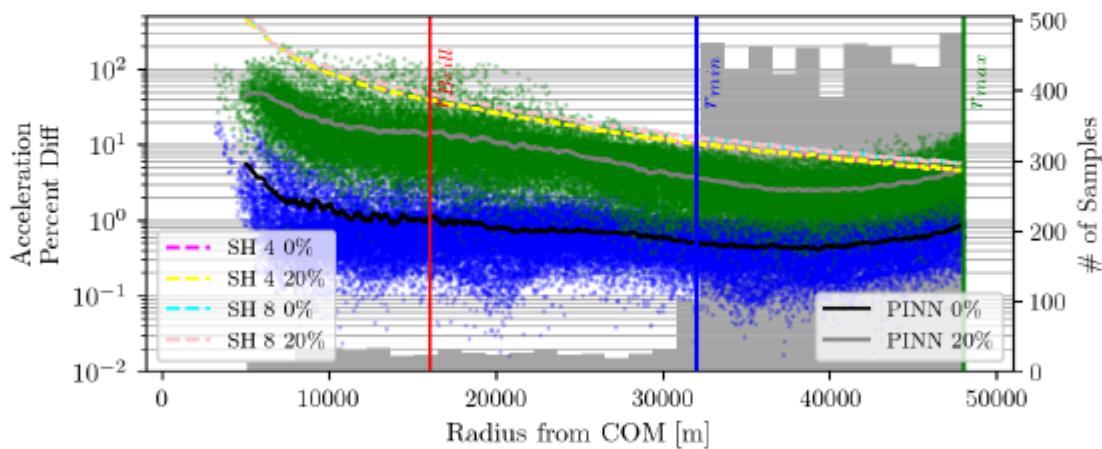
в области обучения при идеальных измерениях и < 3%, когда измерения зашумлены.

Однако PINN страдают от ошибки экстраполяции за пределы этих границ. Хотя такое поведение

нежелательно, оно характерно не только для моделей PINN. Регрессирующие сферические гармонические модели

также страдают от большей ошибки экстраполяции. Более того, даже в области обучения сферические модели

модели гармоник генерируют более высокую ошибку, чем PINN. Как для PINN, так и для сферических гармоник,

(a) Обучающее распределение g и соответствующие ошибки модели.(b) Обучающее распределение g и соответствующие ошибки модели.(c) обучающее распределение g и соответствующие ошибки модели.

$$\|a - a^\dagger\| / \|a\| \approx$$

100, их скользящие средние и данные обучения Рисунок 3.19:

Остаточные значения ускорения, распределение в зависимости от радиуса. Синие диаграммы рассеяния соответствуют

ошибке PINN, обученной на идеальных измерениях, а зеленый соответствует PINN, обученному на зашумленных измерениях.

Пунктирные линии представляют погрешность подгонки сферических гармонических моделей к одним и тем же данным.

Серая гистограмма представляет радиальное распределение обучающих данных.

эти результаты свидетельствуют о том, что обобщать модели требуют обучающих выборок из меньших высотах, если они используются в непосредственной близости к телу.

Чтобы выяснить, как много образцов с низкой высоте необходимы, распределение P добавок
распределение r с 500 дополнительными образцами, собранными между поверхностью астероида и радиусом $2R$.

Благодаря включению этого относительно небольшого набора измерений точность моделирования PINN значительно стабилизируется
. Если ошибка не включена в обучающие данные, средняя ошибка PINN

остается $< 1\%$ для большинства высот. Модель PINN выполняется только на поверхности тела.

достигните средней ошибки моделирования в 5%. Когда ошибка добавляется к данным обучения ускорению, точность моделирования снижается, но общая ошибка остается ниже 10% для большинства высот

и достигает максимальной средней ошибки моделирования в 40% только при оценке на поверхности - и то, и другое
значительное повышение точности по сравнению с сетями, обученными только на r . Этот эксперимент
особенно обнадеживает, поскольку он подразумевает, что PINNs довольно эффективно извлекают информацию
о гравитационном поле, когда им предоставляется доступ к разреженным образцам на малой высоте. Такие результаты согласуются с
выводы PINN-GM-I и предоставляют убедительные доказательства в пользу таких концепций, как гравитационные попперы.

Учитывая, как мало данных требуется для создания обобщаемых моделей, возможно, нет необходимости во многих
гравитационные попперы или даже попперы, которые могут создавать различные траектории.

Эксперимент с архитектурой и потерями

Оценить, насколько улучшаются дополнительные физические ограничения и модифицированная архитектура сети.
точности моделирования, проводится еще один эксперимент, в ходе которого штыри и ямки тренируются с различными
функции затрат на все более разреженных наборах данных ($N = \{2500, 1250, 625\}$), состоящих из пар
положение / ускорение, которые выбираются случайным образом от 0 до $3R$. После обучения ошибка моделирования
этих представлений PINN и PIT затем оценивается на отдельных тестовых наборах данных, состоящих
из 20 000 случайно распределенных пар положение / ускорение. Первый тестовый набор содержит 20
000 пар положений / ускорений, отобранных за пределами сферы Бриллюэна и в пределах $3R$ от тела
(снаружи), второй тестовый набор отбирается между поверхностью астероида и сферой Бриллюэна
(внутри), и, наконец, третий набор выборок положений / ускорений из 20 000 выборок случайным образом

выбранные грани модели многогранной формы (поверхности). Эти три области тестового набора данных показаны

на рисунке 3.20. Каждая сетевая модель обучается с использованием только одной из функций затрат (J_A , J_{AP} , J_{ALC})

$J_{AP LC}$ и функция затрат, не основанная на физических данных, обозначаемая J_{00} ИЛИ

$$J_{00} = \frac{1}{N} \sum_{i=0}^{N-1} \| \mathbf{a}_i - \hat{\mathbf{a}}_i \|^2 \quad (3.18)$$

Кроме того, каждая из моделей обучается на наборах данных с повышением уровня ошибок, добавляемых к

данные об ускорении. Ошибочные оценки ускорения генерируются таким же образом, как

представлены в разделе. 3.2.5. Процент тестируемых ошибок составляет 0%, 10% и 20% от истинного ускорения.

величина. Три модели сферической гармонической силы тяжести (соответствующие степени $l = \{4, 8,$

$16\}$) также регрессируются в каждом обучающем наборе данных с использованием метода

наименьших квадратов и оцениваются с использованием тех же тестовых наборов данных.

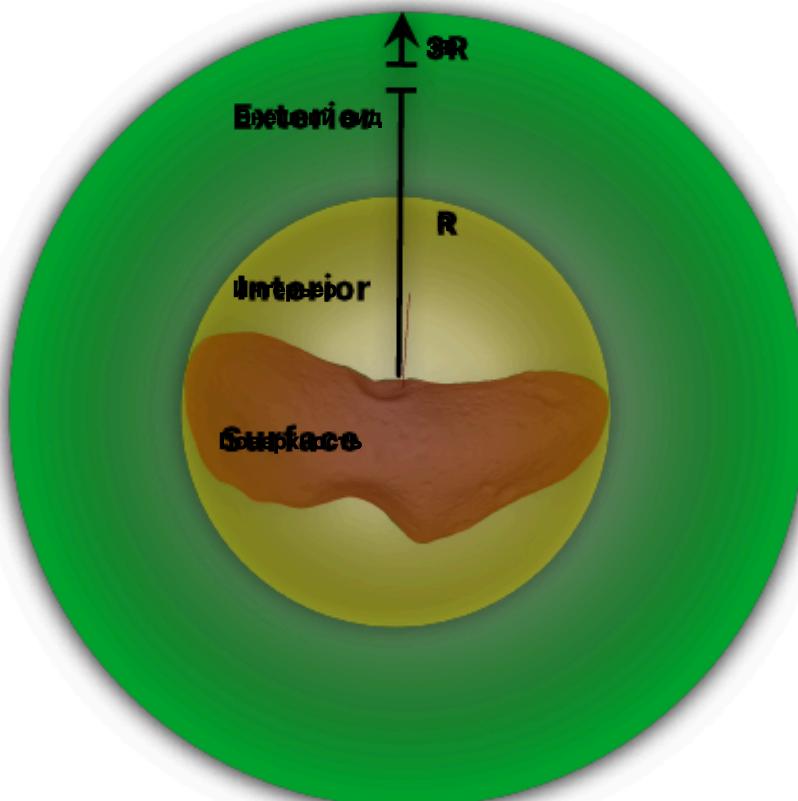


Рисунок 3.20: Распределение данных Eros

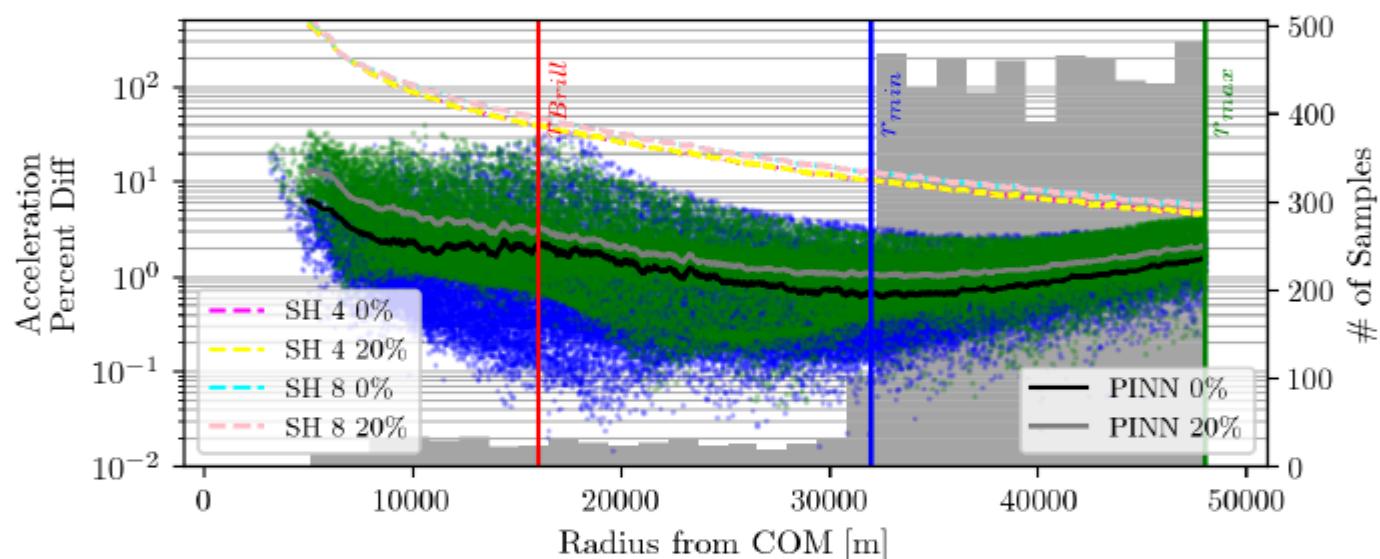


Рисунок 3.21: Остатки ускорения как функция распределения обучения с помощью J_{ALC}
демонстрируют
как дополнительные физические ограничения помогают снизить чувствительность модели к шуму в обучающих данных

Характеристики гравитационных моделей PINNs, PITs и spherical harmonics отображаются следующим образом

функция объема данных для обучения, H , и ошибки в обучающих данных (0%, 10% и 20%). В

соответствующие результаты для каждого тестового набора данных (внешний, внутренний и поверхностный) показаны на рисунке 3.22,

Рисунок 3.23, и Рисунок 3.24 соответственно. Обратите внимание, что ошибки модели определены количественно как ускорение

процентное распределение ошибок, и они представлены в виде прямоугольных графиков. В рамке представлены

1-й и 3-й квартили (от 25% до 75%) процентного распределения ошибок, т.е. 50% данных, находятся внутри

рамки. Стока внутри рамки соответствует медиане или 2-му квартилю. Линии, расширяющиеся

за рамкой соответствует 1.5

\times IQR, где IQR - межквартильный диапазон, а точки

за пределами этих линий соответствуют выбросам. Кроме того, рисунок 3.22, Рисунок 3.23, и Рисунок 3.24 сгруппируйте

модели PINN, spherical harmonics (SH) и PIT в разные цветовые группы. Таким образом, строки

внутри каждого кластера условные обозначения соответствуют моделям на рисунке, нанесенным слева направо.

Рисунок 3.22, 3.23, и 3.24 продемонстрируйте ряд интересных особенностей поведения PINN

и модели гравитации ЯМы. Во-первых, дополнительные ограничения функции затрат действуют как форма

регуляризации регрессии. При наличии достаточного количества данных ($N = 2500$) с практически нулевым

шум, PIN-коды с несколькими ограничениями (ALC, APLC), как правило, работают хуже, чем их более простые аналоги

(00, A, AP). Однако, когда шум в наборе данных увеличивается или количество

выборки уменьшаются, сети с несколькими ограничениями становятся более надежными. Эти сети подтверждают это.

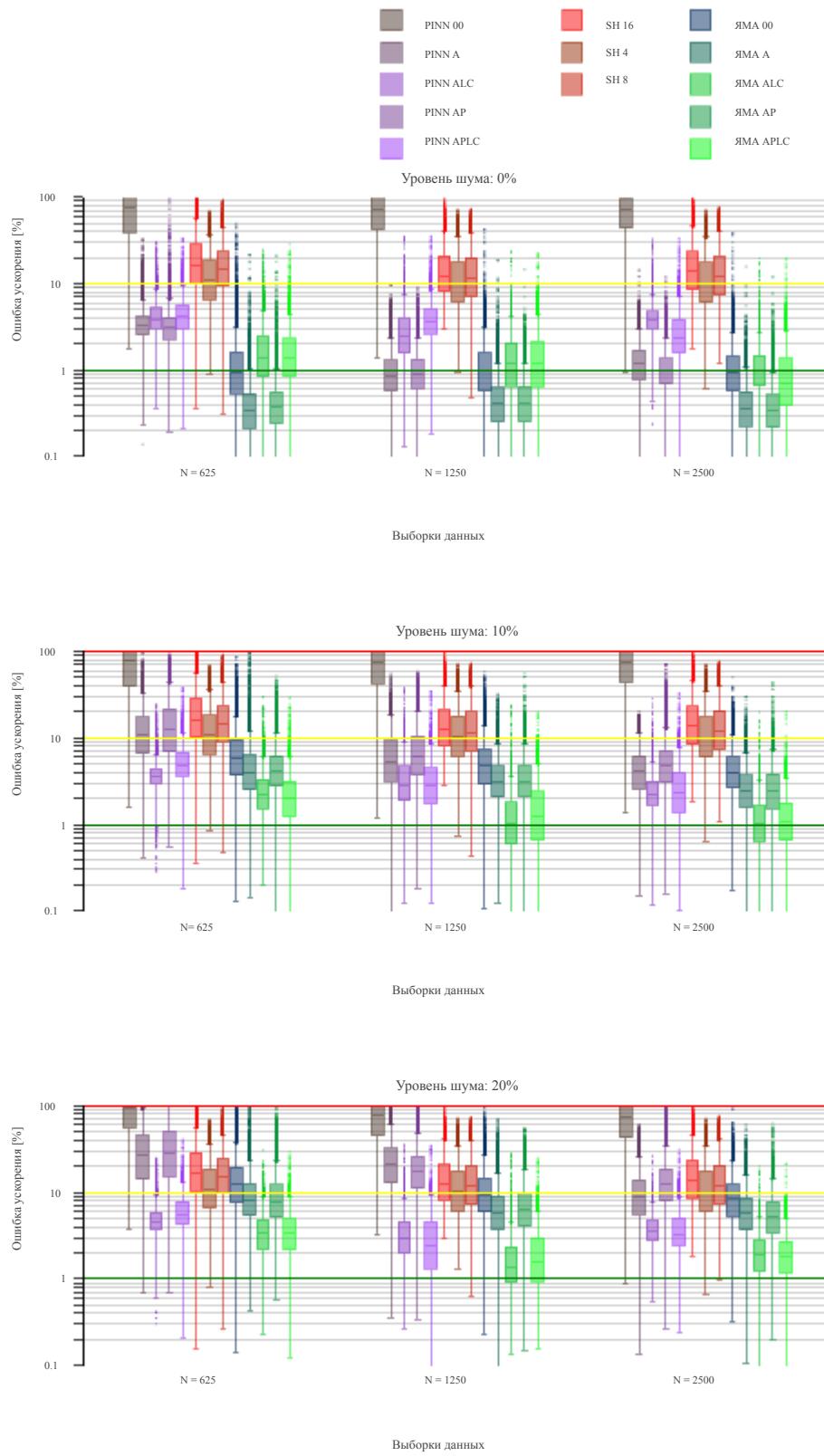
дополнительные ограничения сводят к минимуму риск чрезмерной адаптации к шуму или разреженным данным и являются более

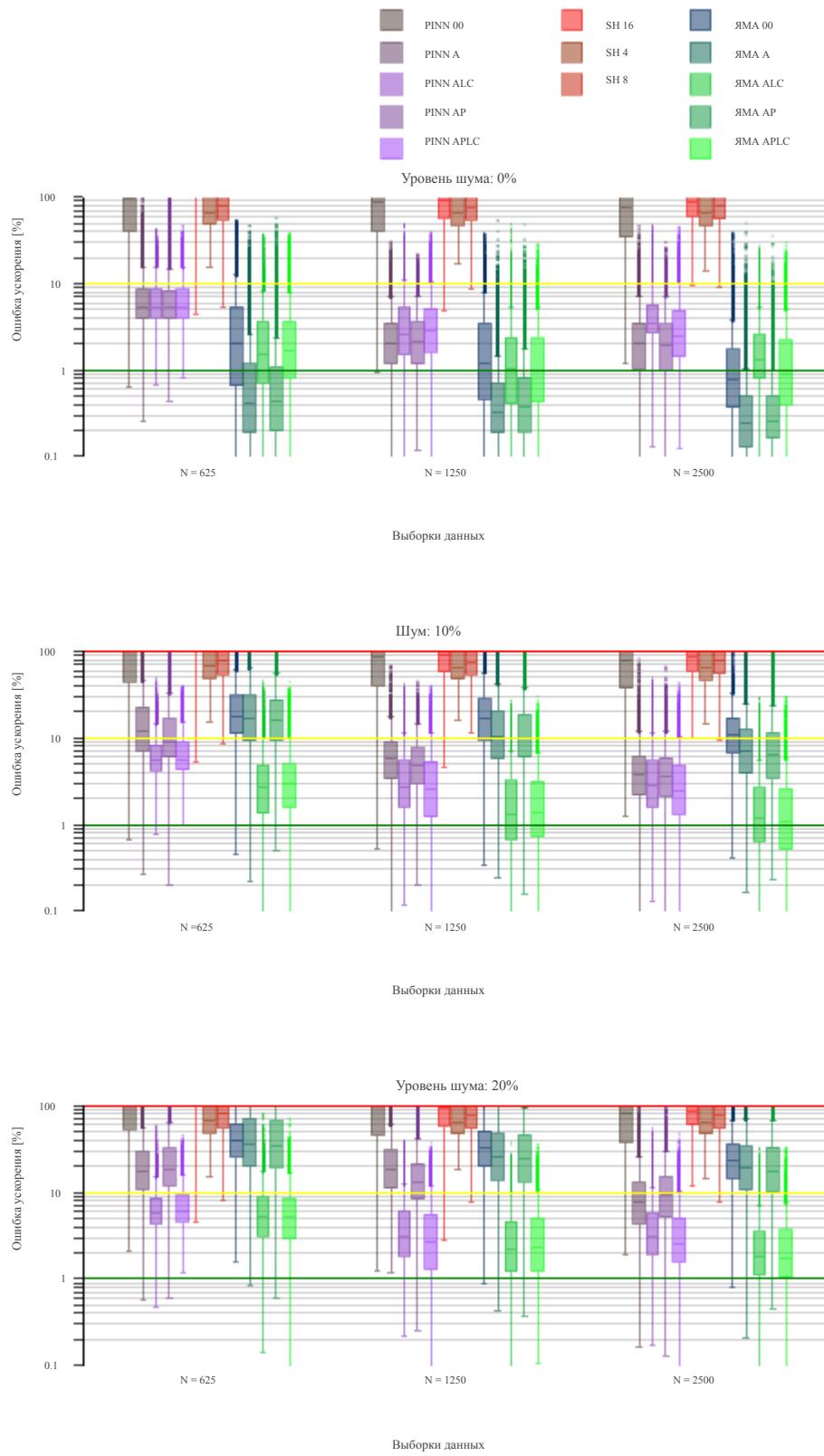
надежной формулировкой в неопределенных или новых средах. В этом также можно убедиться, сравнив

Рисунки 3.19 и 3.21 и отмечая, как сеть PINN ALC, обученная на зашумленном г-распределении,

функционально нечувствительна к шуму.

Вторая интересная находка из Рисунков 3.22, 3.23, и 3.24 является последовательным улучшением производительности-улучшение качества ям по сравнению с шипами. Как в режимах с низкой, так и с высокой дискретизацией архитектура, основанная на преобразователе, может обеспечить снижение погрешности до порядка-несмотря на то, что обучение проводится с использованием тех же данных, что лучше всего проиллюстрировано на примерах с низким уровнем шума на рисунке 3.22 и 3.23. Однако стоит отметить, что трансформаторы также более подвержены переоснащению, чем штыри. из-за их большей способности к моделированию, что видно при использовании более простых физических решений.-





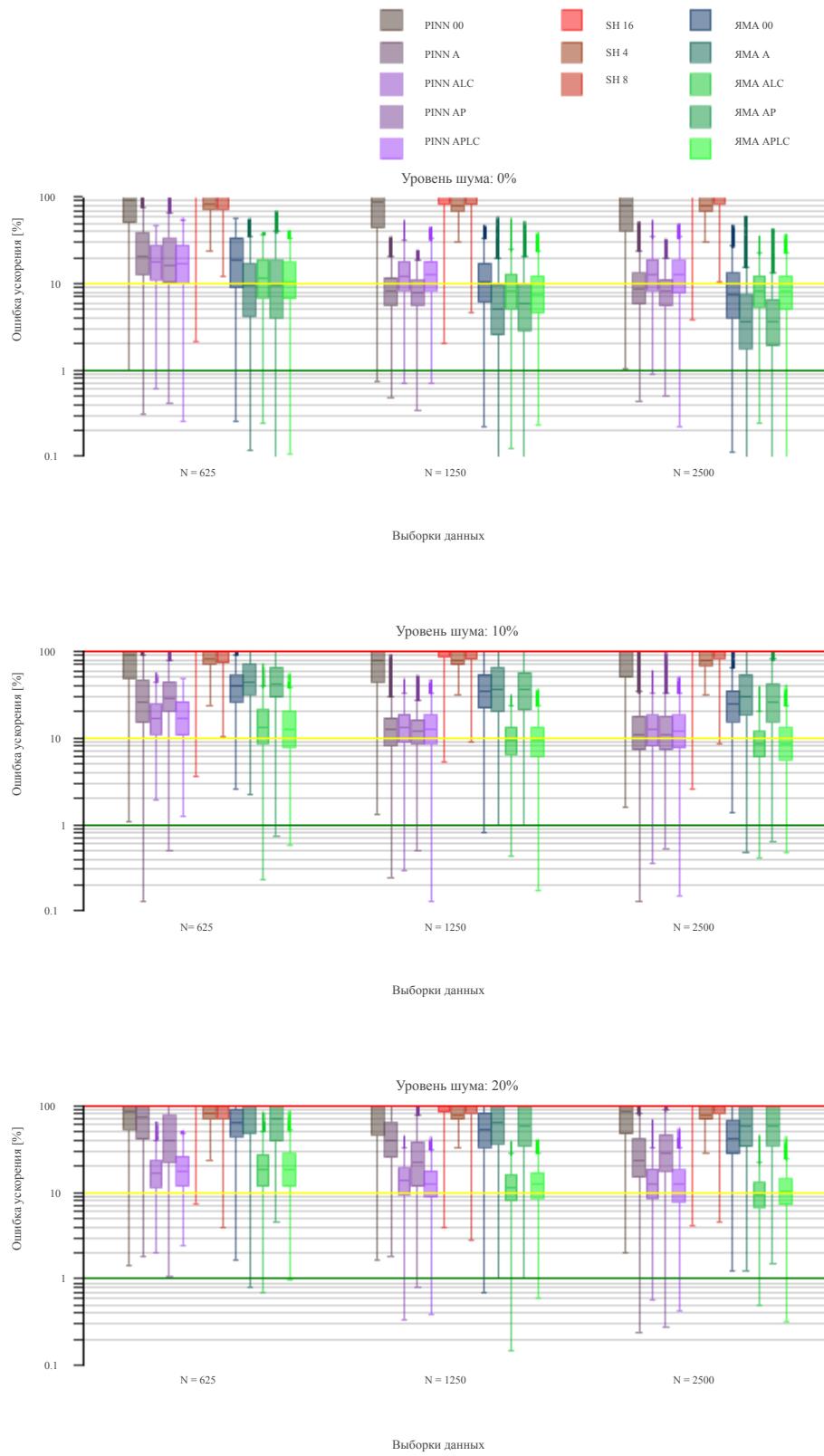


Рисунок 3.24: Ошибка модели на поверхности астероида (surface) в зависимости от количества обучающих данных и шума, добавленного к обучающим данным.

ограничения (00, A, AP) для 10%-ных и 20%-ных случаев шума на рисунках 3.23 и 3.24. В случае 0% шума,

такую гибкость моделирования можно считать сильной стороной, поскольку меньшее количество физических ограничений позволяет

модели сходиться быстрее, чем их более упорядоченные варианты. Однако в более практических

в приложениях, например, когда поле неизвестно априори, рекомендуется использовать дополнительные

ограничения, чтобы свести к минимуму риск переобучения.

Наконец, как штыри, так и углубления обеспечивают значительный прирост производительности по сравнению со своим сферическим гармоническим предшественником-

не только во внутреннем и поверхностном распределении, для которого сферические гармоники

будут расходиться, но и внешнего распространения, где сферических гармоник часто имели Аналит-

скую возвышенность. Это неизменно хуже производительность сферических гармоник представительства

функция небольшого размера выборки, случайного распределения этих выборок и малой высоты над уровнем моря

данные - каждый из которых способен создать числовые проблемы при оценке сферических коэффициентов гармоник методом наименьших квадратов

. В совокупности PINN-GM-II демонстрирует отличные эксплуатационные характеристики при применении в условиях небольшого корпуса

. Благодаря улучшенной разработке характеристик, дополнительным условиям потери и улучшенной сети

архитектура PINN-GM-II обеспечивает значительно лучшую эффективность выборки, чем его предшественник.

Более того, PINN-GM-II демонстрирует устойчивость к зашумленным данным и успешно восстанавливает

низковысотные режимы с относительно небольшим количеством выборок.

3.3 Поколение III

Третье и самое последнее поколение гравитационной модели PINN (PINN-GM-III) представляет собой дополнительных модификаций для решения различных болевые точки остаются без внимания по прошлых поколений, а также повышает точность и надежность модели. Наглядное изображение этих изменения показаны темно-серым цветом на рис. 3.25 и может быть сопоставлен с предыдущим PINN-GM-II на рисунке 3.16. Все модификации подробно описаны в разделах ниже.

Модель	Параметры	Обучающие данные	Средняя ошибка [%]	Действителен во всем мире
GP (45)	12,960,000	3,600	1.5%	✗
NNs (92)	1,575,936	800,000	0.35%	✗
ВЯЗЫ (6)	100,000	768,000	1-10%	✗
Геодезическая сеть (7)	91,125	500,000	0.36%	✗
PINN-GM-III	3,048	4,096	0.20%	✓

Таблица 3.4: Статистика гравитационной модели машинного обучения - См. Приложение D

3.3.1 Усилия по предварительной обработке

PINN-GM-III продолжает поиск наиболее продуктивной нормировки входных данных PINN-GM. Как обсуждалось ранее, одной из наиболее распространенных практик является нормализация данных таковы, что они существуют в пределах $[-1, 1]$, поскольку многие функции активации демонстрируют свою большую нелинейность в этом режиме. Нелинейности - это то, что обеспечивает нейронным сетям их мощную возможность аппроксимации, и если входов насытить функции активации, тогда Сети снижается способность моделирования. Кроме того, нормализация также улучшает устойчивость во время обучения, и снижает риск возникновения плохой обусловленности матрицы операции, если обучение слишком большие или маленькие.

Традиционно, входы и выходы нейронной сети нормированы таким образом агностик друг друга, например, если входы равны x , а выходы равны y , то нормализация дает $[x_{\min}, x_{\max}] \rightarrow [-1, 1]$. Однако в PINNs входные и выходные данные могут иметь общие единицы измерения, и эта несвязанная нормализация может привести к несоответствующим физическим условиям. Для моделирования гравитационного поля

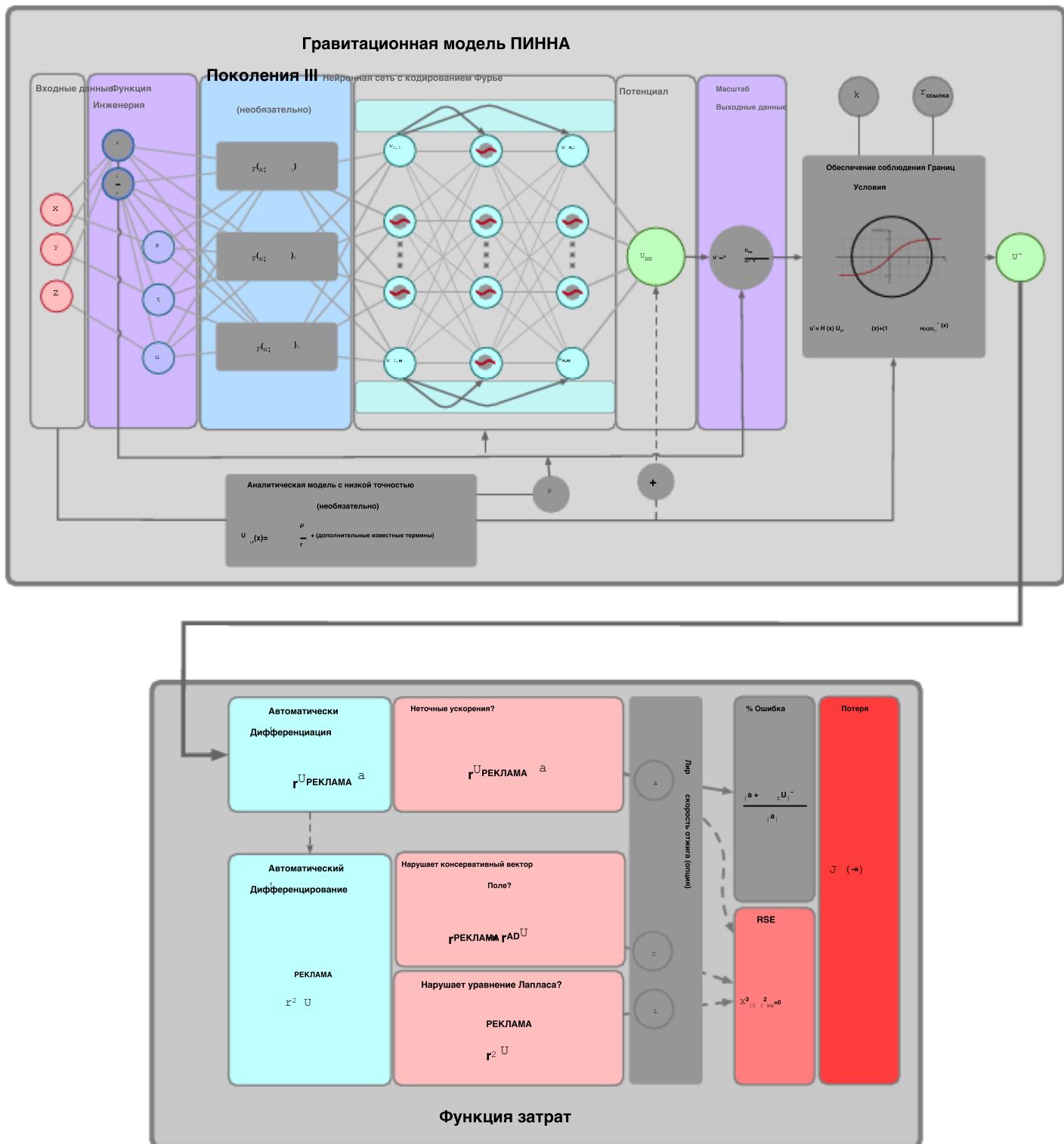


Рисунок 3.25: PINN-GM поколения III с новыми модификациями, помещенными в темно-серые коробки

проблема, поэтому важно нормализовать входные и выходные данные с учетом размеров

. В PINN-GM-III это достигается путем нормализации положения и потенциала на величину

характерная длина x^* равен радиусу планеты R и максимальному потенциальному значению в тренировке

данные U соответственно. Используя эти характеристические скаляры, можно вычислить и использовать постоянную времени

в сочетании с x^* для безразмерности ускорений. Явно это проявляется через:

$$x = \frac{x}{x^*}, \quad U = \frac{U}{U^*}, \quad a = \frac{a}{a^*} \quad (3.19)$$

где x^* , U^* , и a^* еще определяются ли константы без размерности как:

$$x^* = R \quad (3.20)$$

$$U^* = \max(U_r - U_{\text{ЕСЛИ}}, r) \quad (3.21)$$

$$t^* = \frac{\sqrt{2}x}{U} \quad (3.22)$$

$$a^* = \frac{x^*}{t^{*2}} \quad (3.23)$$

где R - максимальный радиус небесного тела, U_r есть истинный гравитационный потенциал в

обучение датум на r , x и U L_F любое низкое качество потенциальных взносы уже учтены

в пределах PINN-GM (обсуждается в разделе 3.3.5).

Поскольку сетевые входы и выходы не имеют размеров, вводится один дополнительный этап предварительной обработки

для обеспечения того, чтобы все входы оставались ограниченными. Как и в случае с PINN-GM-II, декартовы координаты положения без измерения-

преобразуются в их сферическую координату 4D

описание (r, s, t, u) , где r - радиус, а s, t и u - синус углов между

точка поля и каждая декартова ось, определенные как $x/r, y/r$ и z/r (84). Это преобразование гарантирует, что значения $s,$

t и u остаются в численно желаемом диапазоне $[-1, 1]$. Радиальная координата, однако, не гарантирует

тех же удобств. Для точек поля, которые существуют на бесконечных радиусах от тела,

координата r может масштабироваться от $[0, \infty)$, которые могут создавать численные проблемы для сети. Для

чтобы обойти эти трудности, вводятся два прокси-значения радиуса, r_{clip}

и $(1/r)_{clip}^{-r}$

$clip$

является радиальной составляющей, которая обрезается при значениях, превышающих единицу, а $(1/r)_{clip}$

является обратной величиной

от r , которая также обрезается, когда ее значение превышает единицу. Это гарантирует, что информация о радиальном

координаты всегда фиксируются, но убедитесь, что они никогда не выходят за пределы [0, 1] и остаются численно стабильными для обучения.

3.3.2 Модифицированная функция потерь для учета выборок с большой высотой

Функция потерь для исходного PINN-GM представляет собой показатель среднеквадратичной ошибки (RMS):

$$L_{\text{RMS}}(\theta) = \sqrt{\frac{1}{N} \sum_{i=0}^N \| \hat{\nabla U}(x_i | \theta) - a_i \|^2} \quad (3.24)$$

Эта функция потерь особенно распространена в регрессионных задачах машинного обучения с контролем, и она поощряет модели к минимизации наиболее волнистых расхождений между истинными и прогнозируемыми значениями.

Для PINN-GM это означает минимизацию разницы между дифференцированным потенциалом сети-

тиал, $\hat{\nabla U}(x_i | \theta)$, и истинное ускорение, a_i , удовлетворяющее дифференциальному уравнению $-\nabla U = a$. Несмотря

при своей популярности эта функция потерь сопряжена с неожиданными недостатками для гравитационного моделирования проблема, как показано в предыдущих разделах.

В частности, гравитационные ускорения, создаваемые ближе к небесному телу, имеют значительно большую величину, чем ускорения, создаваемые на больших высотах. Как следствие, любые небольшие относительные погрешности в малых высотах прогнозы будут выглядеть непропорционально большой по сравнению с любым высокая ошибки высоте. Поэтому гравитационные модели, обученные с помощью этой функции стоимости всегда приоритет низкий-образцы высоте над высокой пробы высота, даже если высотные выборки более ошибочны в относительном смысле. Это было продемонстрировано как PINN-GM-I, так и PINN-GM-II, для которых обе модели приоритетными для точного моделирования малых высотах образцы, даже когда большинство данных существует на больших высотах.

Чтобы устранить этот конструктивный недостаток, PINN-GM-III использует новую функцию потери среднего процента ошибок:

$$L_{\%}(\theta) = \frac{1}{N} \sum_{i=0}^N \frac{| \hat{\nabla U}(x_i | \theta) - a_i |}{| \hat{\nabla U}(x_i | \theta) |} \quad (3.25)$$

Используя вместо этого процент потери ошибки, PINN-GM-III больше не зависит от абсолютной величины-значения векторов ускорения, но вместо этого стремится минимизировать относительные ошибки. Этот выбор гарантирует, что все образцы, независимо от высоты, в равной степени влияют на функцию потерь и

точно моделируются сетью. Если у пользователя есть приложения, которые особенно чувствительны к

ускорения на малой высоте, такие как посадка или управление касанием - Уравнение (3.25) может

при необходимости быть дополнен уравнением (3.24).

Чтобы проиллюстрировать, как функция потерь в сети влияет на производительность модели, предлагается тест

который обучает две PINN-GM на 5000 парах обучающих данных о местоположении / ускорении, распределенных с расстояния

0-15 радиусов Земли над поверхностью Земли. Один PINN-GM обучается с использованием исходной функции среднеквадратичных потерь

, а другой обучается с использованием функции потери процента ошибок. После обучения каждая сеть

оценивается на наборе из 10 000 случайно распределенных тестовых точек в области 0-15R и их

результатирующих среднеквадратичных значениях и процентах погрешности⁵ представлены в зависимости от высоты над уровнем моря на рисунке 3.26.

Рисунок 3.26а подтверждает, что сети, обученные с использованием функции среднеквадратичных потерь, непропорционально отличаются

предпочтение отдается низковысотным полевым точкам (0-2R) в ущерб высокогорным полевым точкам (5-15R), несмотря на

обучен по всей области 0-15R. Напротив, модель PINN-GM обучена с использованием функции потери процента ошибок

(рисунок 3.26c) отдает приоритет точному моделированию на всех высотах и удобному

выдает более низкие среднеквадратичные значения для образцов с большой высотой, как показано на рисунке 3.26d.

3.3.3 Улучшайте числовые показатели, изучая Косвенный показатель Потенциального

Все гравитационные потенциалы затухают на больших высотах в соответствии со степенным законом вида $1/r^p$

Как обсуждалось, это затухание создает проблемы при использовании функции среднеквадратичных потерь, но оно также вводит

другие численные трудности для нейронных сетей. Учтите, что наибольший гравитационный потенциал

представленный сетью, не имеет размерности до величины 1 из уравнений (3.19) и

(3.21). Для полевых точек, расположенных на достаточно большой высоте, нейронной сети необходимо будет вычислить

потенциалы, которые распадаются до значений, меньших или равных точности станка (т.е. $U(r)$

$$\rightarrow U \leq \Gamma_{\text{крит}}$$

$\epsilon_{\text{машина}}$). Представление этих совершенно разных числовых масштабов с помощью одних и тех же матричных операций

(нейронная сеть) нежелательно и может привести к числовой нестабильности во время обучения и вывода.

Как следствие, эта нестабильность излишне ограничивает максимальную высоту, на которую рассчитана модель

⁵ Обратите внимание, что сообщаемые ошибки касаются только вклада ускорения выше степени сферической гармоники 2-го порядка модели - т. е. Того, насколько точно сеть фиксирует все оставшиеся гравитационные возмущения за пределами $C_{2,0}$

и С

2.2*

Это позволяет избежать этих вкладов, скрывающих ошибку моделирования более прерывистых объектов, и следовать процедуре, приведенной в разделе 3.1.2.

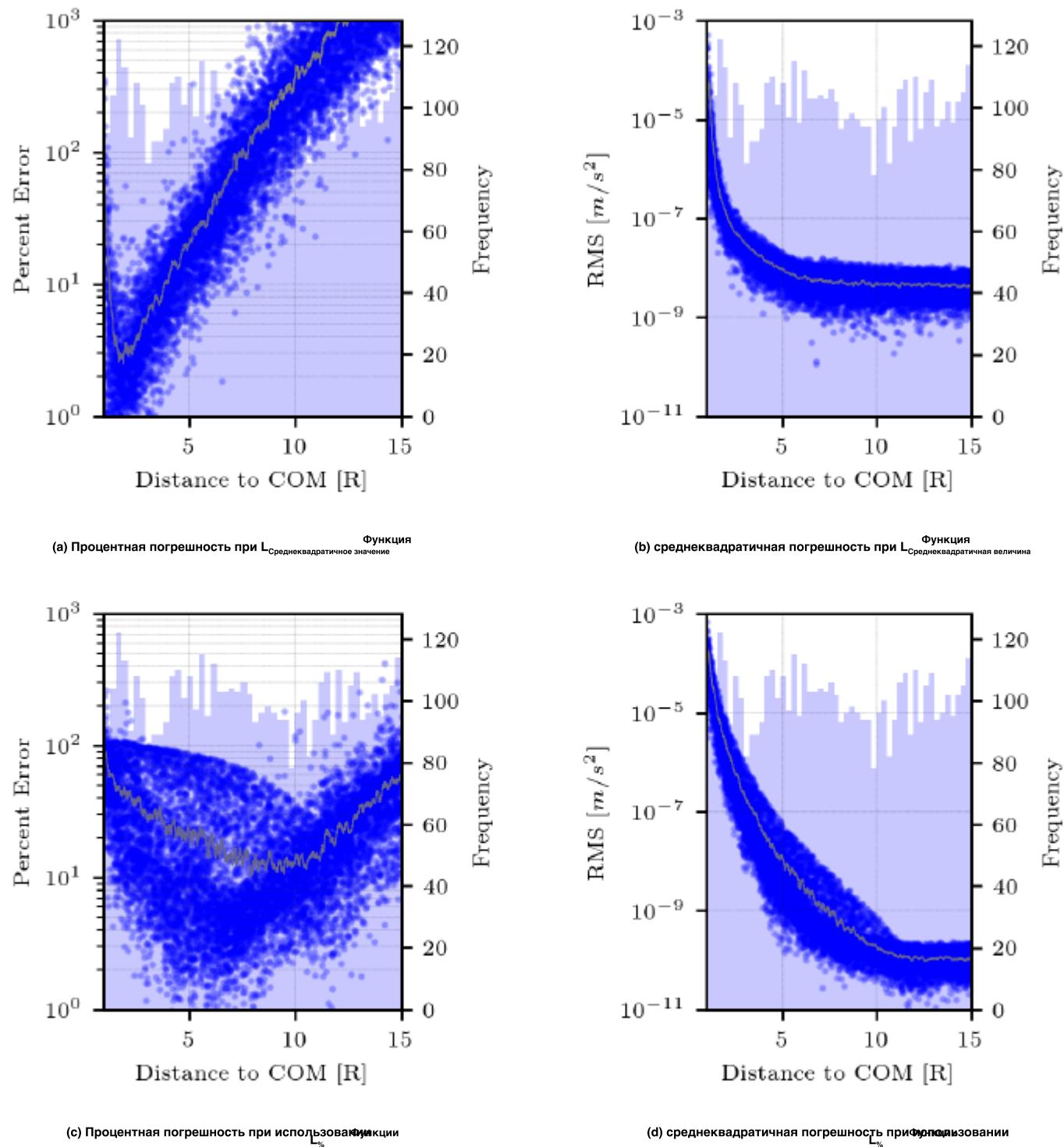


Рисунок 3.26: Различные функции потерь изменяют производительность сети на больших и малых высотах.

Синие точки обозначают индивидуальные ошибки в тестовых данных, синяя гистограмма - распределение обучающих данных, а серая линия - средняя ошибка теста в пределах скользящего окна в 100 пунктов.

жизнесспособный.

PINN-GM-III решает это явление, изучая более благоприятное с числовой точки зрения представление- потенциала, который позже может быть преобразован в надлежащий порядок величины. Явно,

PINN-GM-III предлагает изучить не зависящий от высоты прокси для потенциала, U_{NN} определяется как

$$U_{NN} = Ur^p \quad (3.26)$$

где U - истинная безразмерная потенциальная функция, r - безразмерный радиус

точки поля, а p определяется через

$$p = \frac{l_r + 1}{l_r + 2} \quad l_r \geq 0 \quad r^{-1} \quad (3.27)$$

где l_r включена ли максимальная степень сферической гармоники в необязательную аналитическую часть

модель, U_{LF} (Раздел 3.3.5).

U_{NN} разработан таким образом, чтобы последовательно охватывать область [-1,1] независимо от высоты. В отличие от

true potential, U_{NN} значительно меньший риск введения численных ошибок и преждевременно

точность покрыва модели на больших высотах. После того, как сеть выдает свои прогнозы на

косвенный потенциал затем преобразуется в истинное распределение потенциала позже в модели

посредством

$$\hat{U}_{NN} = \frac{U_{NN}}{r^p} \quad (3.28)$$

Это гарантирует, что цифры сетевого вывода остаются хорошо обусловленными, в то же время позволяя

модели выдавать правильное значение / порядок величины потенциала.

Рисунок 3.27 визуализирует разницу между истинным и косвенным потенциалом как функцию высоты и фигуры 3.28 показано соответствующее влияние на погрешность модели. В частности,

Рисунок 3.27а показывает 10 000 значений потенциала Земли⁶ распределен между 0-5R (радиусами) от

поверхности Земли. Видно, что возмущения затухают со скоростью $1/r^4$ соответствует стандарту Equa-

⁶ Обратите внимание, что эти значения исключают вклад сферической гармонической модели степени и порядка 2 (т.е. $\delta U = U_{LF} - U_{NN}$ где U_{LF} является сферической гармонической моделью степени $= l_2$).

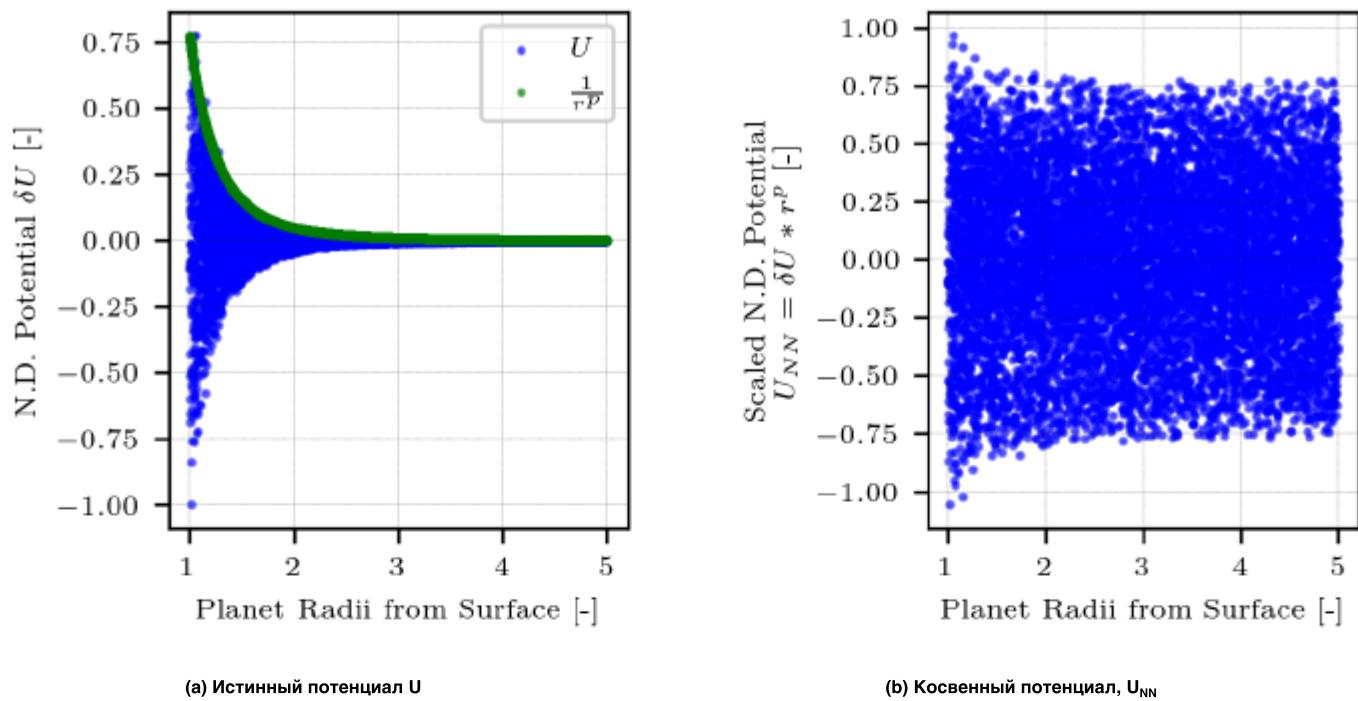


Рисунок 3.27: Истинный потенциал U (слева) быстро уменьшается до численно неблагоприятных значений, тогда как по сравнению с прокси-потенциалом U_{NN} (справа) остается численно хорошо обусловленным между [1, 1].

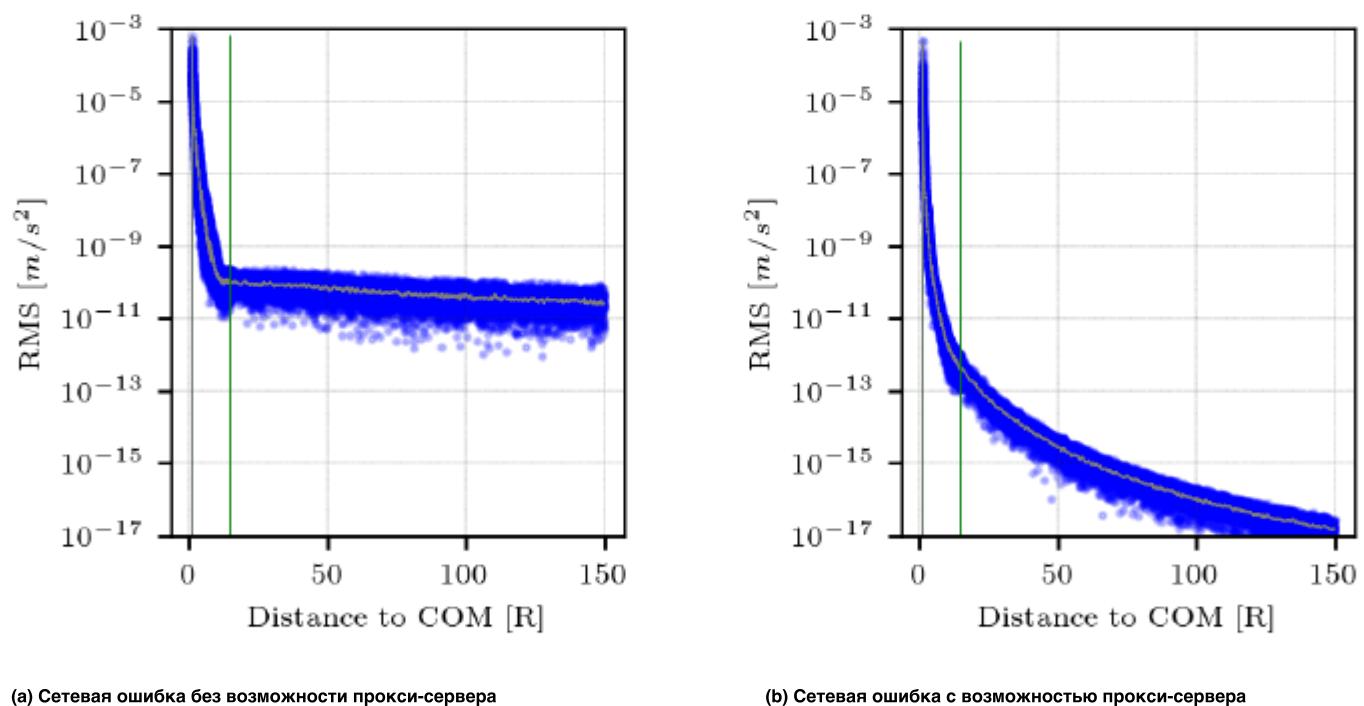


Рисунок 3.28: Влияние обучения потенциального прокси-сервера на ошибку сетевого вывода.

определение (3.27). В отличие от этого, распределение косвенного потенциала показано на рисунке 3.27b. Обратите внимание, как

значения стабильно существуют в более благоприятной с численной точки зрения области [1, 1] независимо от высоты.

Полезность такого масштабирования в рамках PINN-GM-III показана в эксперименте, в котором обучаются две модели

PINN-GM. Обе модели обучаются на 100 выборках, случайным образом распределенных от 0-

15R от поверхности Земли. Одна сеть обучена изучать потенциал напрямую, без использования

масштабирования, предложенного в уравнении (3.26) (Рисунок 3.28а). Другой PINN-GM обучен изучать

потенциальный прокси, который позже делится на r чтобы преобразовать его в правильный порядок величины

(Рисунок 3.28б). Затем среднеквадратичная ошибка каждой модели оценивается на 10 000 контрольных точках, которые охватывают диапазон

от 0 до 150R.

Рисунок 3.28а ясно показывает, что PINN-GMS, обученные распознаванию потенциала, непосредственно попадают в

числовой барьер примерно при $r = 12R$. За пределами этой критической высоты модельные цифры

становятся плохо подготовленными и не могут представлять эти очень маленькие значения - прежде всего

ограничение диапазона высот, на которых может использоваться эта модель. Напротив, когда PINN-

GM обучается распознаванию косвенного потенциала, который позже масштабируется до правильного порядка величины

с помощью уравнения (3.28) больше не существует критической высоты или числового минимума, который нельзя было бы превысить

3.3.4 Выполнение граничных условий для избежания ошибки экстраполяции

Хотя PIN-коды чаще всего разрабатываются для удовлетворения физических требований с помощью их функции затрат,

существуют и другие способы обеспечения соответствия посредством разработки самой модели машинного обучения

. Ссылка (93) заметки как модели машинного обучения может быть разработана, чтобы плавно перейти

в известных граничных условий с помощью Хевисайд функции.

В PINN-GM-III предложены следующие модификации конструкции для обеспечения соблюдения соответствующих граничных

ограничения:

$$u^*(r) = 1 - H(r)U_{NN} \quad \text{где } U_{NN} \text{ - предсказанный потенциал, } H(r) \text{ - Хевисайд функция, } U_{BC}(r) \text{ - потенциал при граничном условии.} \quad (3.29)$$

где U_{NN} -

является потенциалом при граничном условии,

а $H(r)$ - функция, основанная на Хевисайде, определяемая как

$$H(r) = \frac{1 + \tanh(k(r - r_{\text{ ссылка}}))}{2} \quad (3.30)$$

где r - радиус точки поля, r_{ref}

- опорный радиус, а k - параметр сглаживания

для управления более непрерывным или дискретным переходом. Обратите внимание, что как r_{ref} , так и k могут быть заданными пользователем-

или изученными параметрами.

Уравнение (3.30) обеспечивает плавный переход PINN-GM в известное граничное условие

после контрольной высоты $r_{\text{ ссылка}}$. Таким образом, модель машинного обучения может использовать

гибкость моделирования нейронной сети для представления сложных областей гравитационного поля вблизи

тела, для которого граничное условие не имеет значения, но затем плавно уменьшить

ответственности в пределе по мере приближения модели к границе.

Существует множество способов принудительного выполнения граничного условия. В пределе

$r \rightarrow \infty$, потенциал спадает до нуля, как описано в разделе 3.3.3. Однако, установив U_{BC} = 0 и

$r_{\text{ ссылка}} = \infty$ в уравнении (3.29) непрактичен, поскольку требует, чтобы нейронная сеть

изучила модель потенциала для всей области r

$\in [0, \infty)$. Более полезным выбором является использование модели с более низкой

точностью воспроизведения в пределе, где более высокая точность не требуется. PINN-GM-III достигает этого, принимая

выводы из модели сферической гармонической гравитации и признание того, что высокочастотные компоненты

гравитационного потенциала убывают до нуля быстрее, чем вклад точечной массы на больших высотах

. Т.е.

$$U_{\text{до нашей эры}}(r) = U_{LF} = \frac{n}{r} + \sum_{l=0}^{\infty} \sum_{m=0}^{l} \frac{R}{r^l} \frac{1}{r^m} (\dots) \quad (3.31)$$

как $r \rightarrow \infty$.

Это наблюдение подразумевает, что $U_{BC}(r)$ может быть установлен на $\frac{1}{r}$ предполагая, что $r \gg R$. Таким образом, PINN-

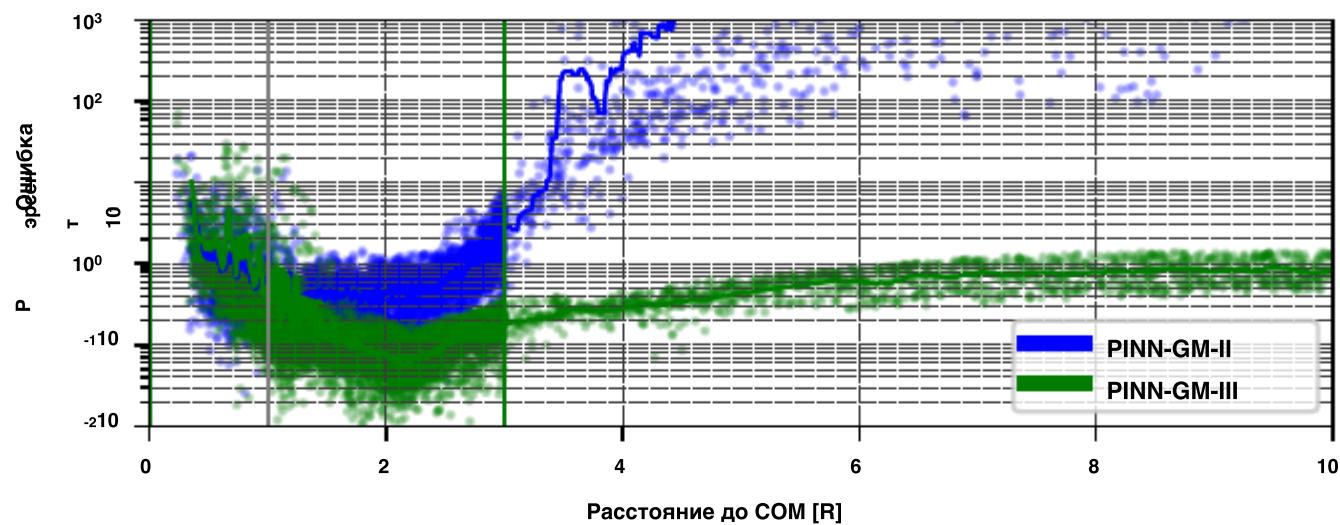
$+ f(r)$ в уравнении (3.29), где $f(r)^\mu$ - любые члены более высокого порядка в наборах GM-III U_{BC}

r сферическая гармоника = модель силы тяжести, которую пользователь знает априори и желает оставить как часть

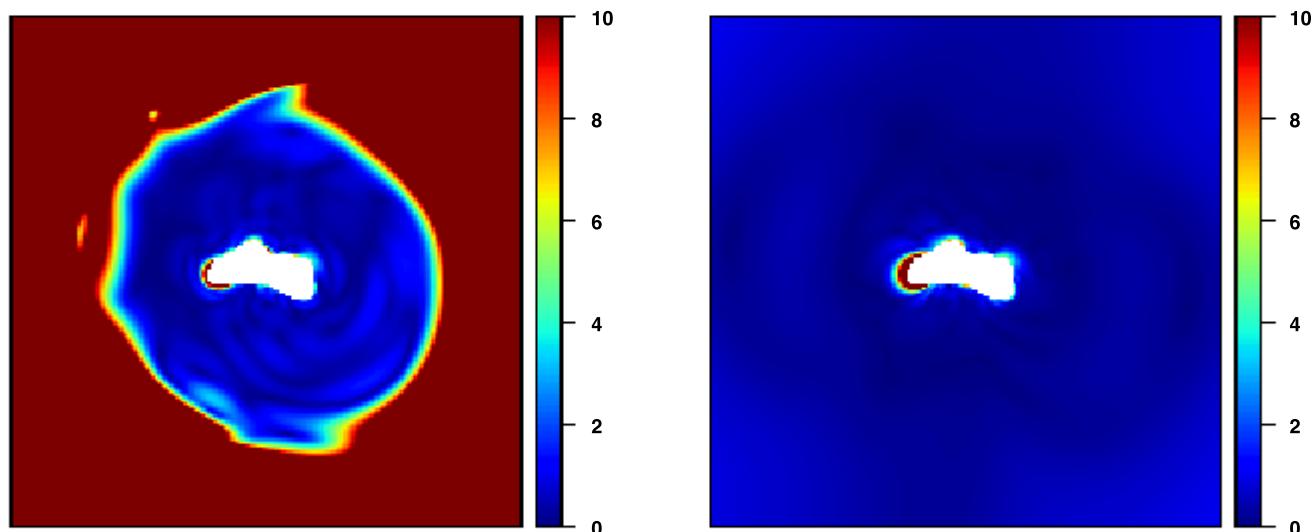
граничное условие.

Чтобы проиллюстрировать эффект от выполнения этих граничных условий, проводится эксперимент

где проходят обучение PINN-GM-II и PINN-GM-III. PINN-GM-II не содержит никакой информации о-



(а) показатели PINN-GM-II и III внутри ($0R-3R$) и за пределами ($3R-10R$) тренировочной области.



(б) Ошибка PINN-GM-II%

(с) Ошибка PINN-GM-III%

Рисунок 3.29: Верхняя строка: Процент ошибки PINN-GM внутри ($0-3R$) и за пределами ($3R-10R$) области обучения. Серая вертикальная линия - радиус Бриллюэна. Зеленая вертикальная линия - максимальные границы обучающих данных. Нижняя строка: % ошибки PINN-GM в плоскости XY.

размышление о граничном условии при разработке модели, тогда как PINN-GM-III включает в себя

предлагаемый переход, отраженный уравнением (3.29). Каждая сеть содержит восемь скрытых слоев с

10 узлов, и обе модели обучаются на ту же 1000 точек данных, охватывающих 0-3Р о

астероид Эрос. Средние ошибки в обеих моделях отражаются как функция высоты от 0-10 рублей

на рисунке 3.29а, а поперечные сечения в декартовой плоскости XY показаны на рисунках 3.29б и 3.29с.

Рисунок 3.29 показывает, что при уравнении (3.29) не включен в конструкцию модели, PINN-

GM расходится при тестировании за пределами своих обучающих данных. Напротив, при уравнении-

(3.29) входит в комплект, PINN-GM-III способен поддерживать низкие погрешности даже на больших высотах, для которых

отсутствуют обучающие данные. Этот разительный контраст в поведении модели демонстрирует преимущество

обеспечения соблюдения граничных условий при проектировании модели. PINN-GM-III может использовать нейронную сеть

для изучения расширенного представления гравитационной модели в рамках обучающих данных, но как только

модель выходит за эти пределы и гарантированно работает не хуже, чем точечное приближение массы-

или сферическая гармоническая модель с низкой точностью. Это мощная разработка для этих машин.

изучение гравитационных моделей как демонстрация того, что эти модели могут быть надежно использованы за пределами их обучающих данных

и всегда будут работать так же хорошо или даже лучше, чем их аналоги с низкой точностью.

3.3.5 Использование существующей информации о силе тяжести в решении PINN-GM

В дополнение к обеспечению соблюдения граничных условий с помощью модели, другой дизайн

выбор включен PINN-GM-III-это способность предохранителей априори гравитационная модель с нейронной

сетевое решение. Например, большинство крупных небесных тел демонстрируют планетарного сжатия, который является

лаконично запечатленный с помощью $\omega_{2,0}$ коэффициент сферической гармоники, обсуждаемый в главе 2.

Вместо того, чтобы требовать от сети повторного изучения этого заметного и легко наблюдаемого возмущения, эта информация для-

моделирование может быть непосредственно встроено в сетевую модель. Таким образом, PINN-GM-III может прогнозировать

ускорения за счет использования аналитической модели первого порядка с низкой точностью и сети, ответственной за

улавливание возмущений высокого порядка с помощью:

$$u^*(r) = 1 - H(r) U_{\text{H}}(r) + U_{\text{NN}}(r) + H(r) U_{\text{LF}}(r) \quad (3.32)$$

где U_{LF} относится к известной аналитической модели с низкой точностью, такой как $U_{LF} = \frac{\mu}{r} + U$.

Алгоритм 1: алгоритм PINN-GM-III

```

1 Сбор обучающих данных ( $x_a$ ,  $y_a$ ) из:
  (a) ранее существовавшая модель
  (b) оперативные оценки состояния (94) Безразмерность обучающих данных // Сек . 3.3.1
2 Преобразовать в сферические координаты ( $r_{clip}$ ,  $1/r$ ,  $s$ ,  $t$ ,  $u$ ) // сек . 3.3.1
3 (Необязательно) Кодирование входных данных в пространство Фурье // Приложение . C
4 Распространять через выходной прокси-потенциал обученной нейронной сети // Сек . 3.3.3
5 Преобразуйте потенциал посредника в истинный потенциал // Сек . 3.3.3
6 Применение граничных условий к потенциальному сетевому // Сек . 3.3.4

// Потенциал автоматической
дифференциации (AD) для обеспечения ускорения если
Тренировка , то 7
  Прибавьте к потере процент ошибки ускорения // Сек . 3.3.2
9 (Необязательно) Вычислить якобиан для  $a$  через AD (необязательно) Выполните отжиг скорости обучения
  // Приложение . B
10 (Необязательно) Добавьте отожженный  $\nabla^2 u^*$  отношение к функции потерь
  функция Обновления сетевых параметров
  Вычислите градиенты потерь

```

3.4 Характеристики PINN-GM-III: сравнительное исследование

В этом разделе сравниваются характеристики PINN-GM-III с характеристиками двух его моделей

предшественники, PINN-GM-I и PINN-GM-II. Модели обучены в условиях, описанных в их оригинальных работах.

PINN-GM-I обучен представлять гравитационное поле Земли, и PINN-GM-II

обучены представлять астероид 433-Эрос. В обоих случаях оцениваются характеристики моделей

для различных ключевых конфигураций гиперпараметров, включая количество обучающих данных, количество

периодов обучения и общую емкость модели. Для удобства чтения, PINN-GM-I, PINN-

GMI и PINN-GM-III будут обозначаться как PINN-I, PINN-II и PINN-III соответственно.

В каждом эксперименте сравниваются мощности трех моделей. Емкость модели определяется

количеством узлов на скрытый слой, $N_{узлов} = \{20, 40, 80\}$, для которого существует в общей сложности 8

скрытых слоев. Каждая модель обучается 25 раз, для каждой используется разное количество обучающих данных и

периоды обучения выбираются из $N_{данных} = \{2^{10}, 2^{11}, 2^{12}, 2^{13}, 2^{14}\}$ и $N_{эпох} = \{2^{12}, 2^{13}, 2^{14}, 2^{15}, 2^{16}\}$

соответственно. Остальные подробности, касающиеся обучения PINN-III, приведены в Приложении A.

3.4.1 Поколение I против поколения III

В первом эксперименте исследуются характеристики PINN-III по сравнению с оригиналом

PINN-I. Каждая сеть обучается на данных, полученных из высокоточного статического гравитационного поля Земли

модель: Earth Gravity Model 2008 или EGM 2008 (95). Эта модель распространяется на степень и порядок

2160 но этот эксперимент ограничен степенью и порядком 1000 (

~ 1 миллион параметров), чтобы гарантировать

точность воспроизведения достаточно высока для улавливания высокочастотных возмущений.

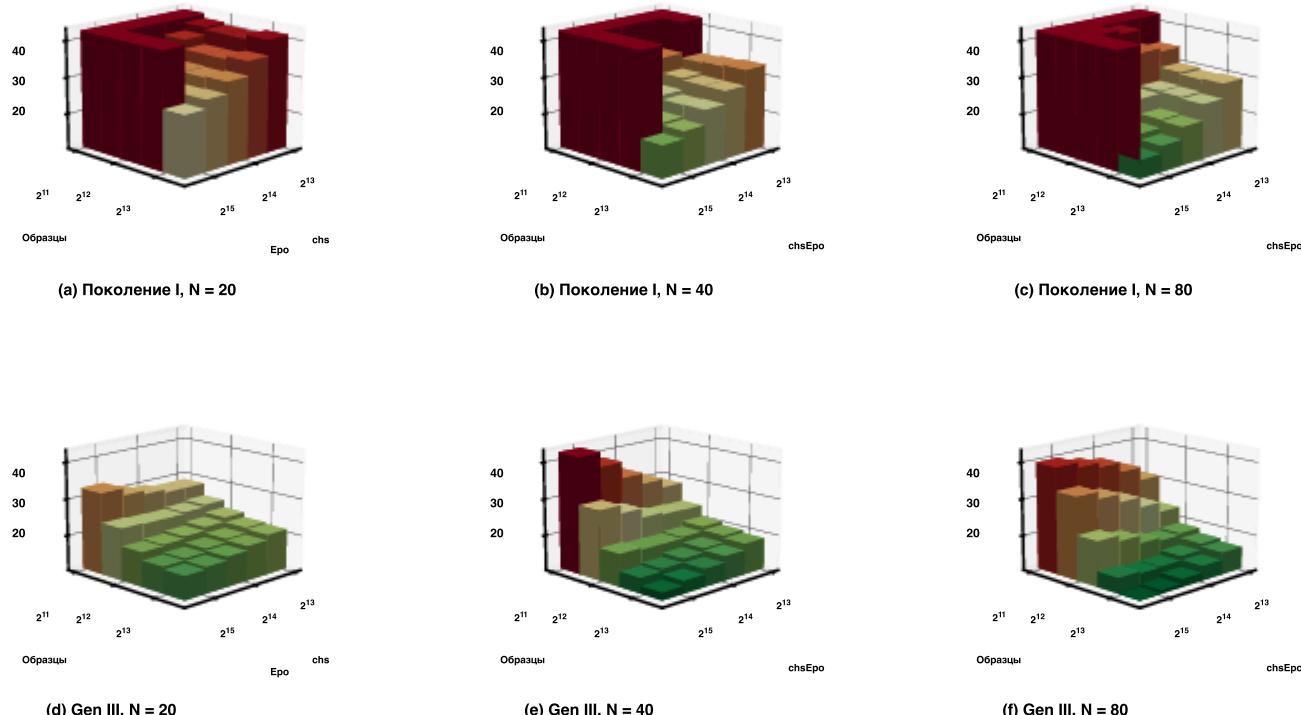


Рисунок 3.30: Средняя процентная погрешность ускорения, определенная в уравнении (3.33)

PINN-GM-I (вверху) и PINN-GM-III (внизу), обученных на Земле.

Обучающие данные для каждой сети распределяются случайным образом от радиуса Земли до

приблизительно высота над уровнем моря 420 км. Как только точки поля выбраны, их соответствующие значения

ускорения вычисляются с использованием EGM2008 и вместе формируют тренировочные данные. Остальные

гиперпараметры для сети PINN-III включены в таблицу A.1 а гиперпараметры по умолчанию-

для сети PINN-I приведены в таблице 3.1.

Производительность каждой сети оценивается с использованием показателя средней процентной ошибки, вычисленного на основе

отдельный тестовый набор из 50 000 случайно распределенных полевых точек в тех же пределах, что и обучающие данные

. Обратите внимание, что показатель процентной ошибки целенаправленно исключает вклад ускорения из

точечной массы Земли и сплющенности планеты через:

$$P = \frac{1}{N} \sum_{i=0}^N \frac{\| \hat{a}_i - a_i \|_2 }{\| a_i \|_2} * 100 \quad (3.33)$$

где \hat{a} - сетевые ускорения, произведенные с помощью $-\nabla U$ и a - произведенные ускорения

точки массового и сроки планетарного скатия сферической гармонической тяжести модели. В

бездействие a_2 от процента ошибки делает его легче определить, насколько хорошо PINN-GM захвата

остальные возмущения высокого порядка, такие как горные хребты и тектонические плиты.

Рисунок 3.30 показывает, что между PINN-I и

PINN-III наблюдается значительное повышение производительности. Для случаев с низкой пропускной способностью, когда количество узлов на уровне равно 20 (размер модели

≈ 3000 параметров), PINN-I последовательно выдает прогнозы с низкой точностью (ошибка $\geq 35\%$). В

случаях с низким объемом данных некоторые выводы даже расходятся. Напротив, PINN-III последовательно

сходится к решению с ошибками менее 30% даже в случаях с низким объемом данных. Аналогичные результаты

проявляются для случаев с более высокой производительностью при $N = \{40, 80\}$, при этом PINN-III

стабильно превосходит своего предшественника по всем объемам данных и продолжительности обучения.

Примечательно, что эти результаты не только подразумевают, что PINN-III способен обеспечить значительно меньшую ошибку

, чем PINN-I, но также и то, что этого можно достичь с меньшим объемом данных, меньшим временем обучения и меньшими

Модели. Учитите тот факт, что 20-узловая модель PINN-III обучалась с $N_{\text{данными}} = 1024$ и

$N_{\text{эпох}} = 4096$ (рисунок 3.30d самый маленький образец, наименьшее эпохи) имеет практически идентичные ошибки до 80

узел PINN-я модель с $H_{\text{сведения}} = 32, 768$ и $H_{\text{эпох}} = 8192$ (рис. 3.30a большинство выборок, второе место

наименьшее количество эпох). Это говорит о том, что PINN-III, который на 94% меньше, обучался с 97%

меньше данных для обучения, и на 50% меньше эпох способен достичь такую же точность, как

PINN-И.

Один дополнительный отметить результат рисунок 3.30 это что, модели PINN-III большой емкости

может возникнуть переобучение. При недостатке данных ($N_{\text{данных}} = 1024$) ошибка начинает увеличиваться

при более длительных тренировках, что лучше всего показано на рисунках 3.30e и 3.30f. В то время как

переоснащение PINN-III остается менее проблематичным, чем то, что наблюдается с PINN-I, это может быть

тривально исправлено путем добавления небольшого количества отсева к каждому слою, включения физических ограничений Лапласа

и curl в функцию затрат или просто путем добавления дополнительных обучающих данных. The

отсев и дополнительные ограничения намеренно не включены в этот эксперимент, поскольку они действуют как

регулирующие условия, которые могут затруднить максимальную производительность этих сетей в условиях нехватки данных

.

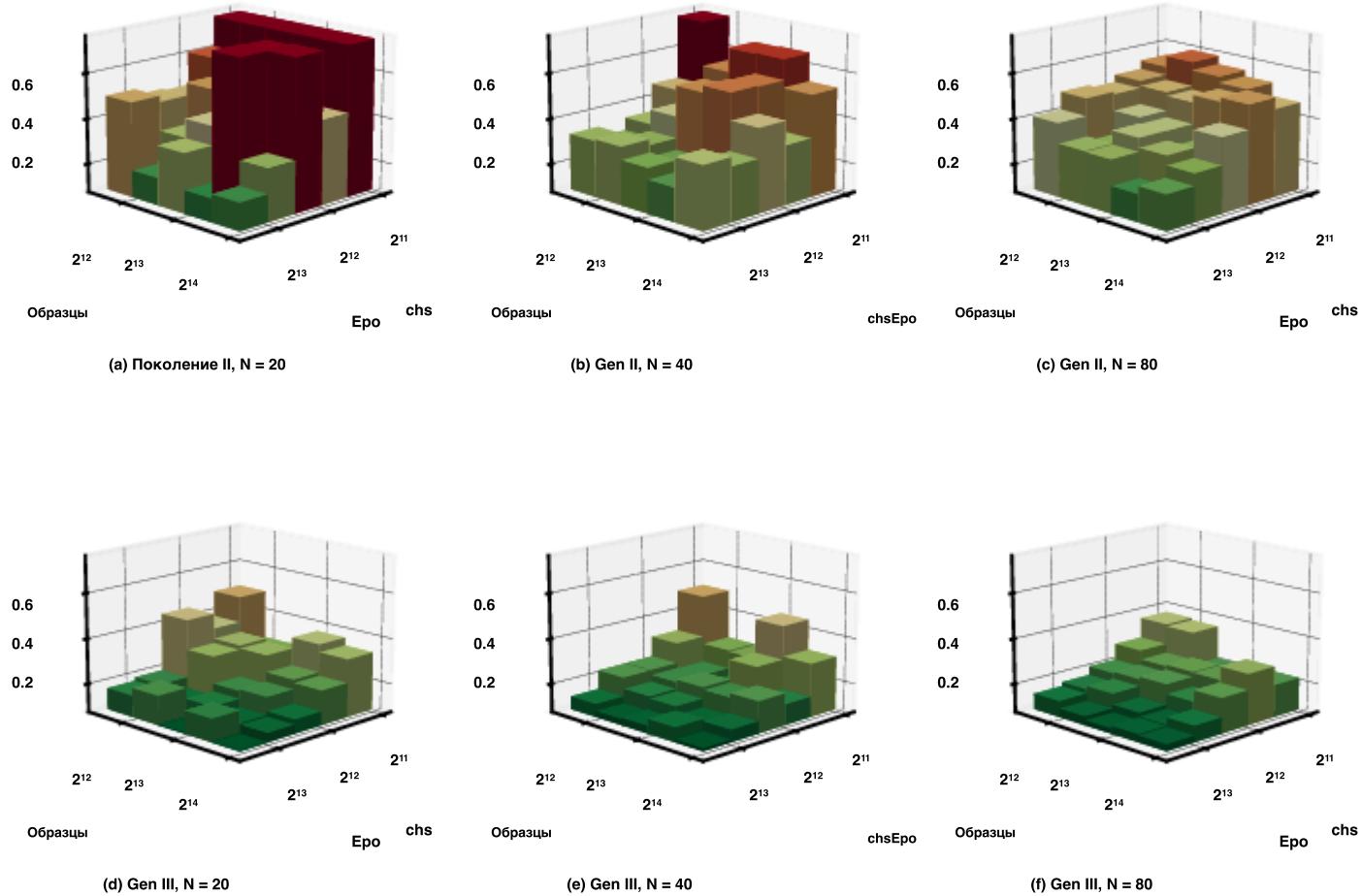


Рисунок 3.31: Средняя ошибка ускорения в процентах для PINN-II (вверху) и PINN-III (внизу), обученных на Eros.

3.4.2 Поколение II против поколения III

Во втором эксперименте исследуются усовершенствования PINN-III по сравнению с его более поздним предшественником

PINN-II. В этом эксперименте среда тестирования переходит на астероид 433-

Eros, для которого первоначально был разработан PINN-II. Вместо использования сферической гармоники

гравитационная модель, истинные ускорения вычисляются с использованием многогранной модели, основанной на низком-

модель формы Eros с высоким разрешением, состоящая из приблизительно 11 000 общих граней и вершин. В

Обучающие данные равномерно распределены между поверхностью астероида и максимальным радиусом

, равным трем радиусам от центра масс.

Рисунок 3.31 иллюстрирует ошибку ускорения для PINN-II и PINN-III. Обратите внимание, что в процентах

указанные ошибки относятся к полному вектору ускорения ($\|\mathbf{a} - \hat{\mathbf{a}}\| / \|\mathbf{a}\|$), а не с точечной массой и

термины сплюснутости удалены. Хотя PINN-II более надежен, чем PINN-I, он приближается к разумному

решение во всех протестированных случаях PINN-III продолжает превосходить его. В то время как PINN-II и III

поддерживают погрешность < 1,5%, PINN-III неизменно остается ниже < 0,5%, улучшая производительность

в 3 раза. PINN-III также способен повысить производительность при меньшем объеме данных, времени обучения,

и меньших размерах модели. Самые маленькие модели PINN-III приводили к ошибкам, сопоставимым с

большинство гиперпараметров протестировано с 80-узловым PINN-II.

3.5 Характеристики PINN-GM-III: астероид неоднородной плотности

Одной из проблем моделирования гравитационных полей малых тел является неопределенность, связанная с

распределение внутренней плотности астероида или кометы. Недавняя работа показала, что плотность постоянна.

предположение, часто используемое в полиэдральной гравитационной модели, не всегда справедливо (44). Некоторые астероиды

может содержать над и под плотным области в пределах их интерьер, или, возможно, были сформированы два

астероиды сливаюсь воедино, каждый с различной плотностью. К сожалению внутренней структуры

эти маленькие тела чрезвычайно сложно измерить, поскольку для этого требуется непосредственное зондирование внутренней части

объекта. В качестве менее навязчивой альтернативы, специалисты по динамике в настоящее время делают обоснованные предположения

об этих распределениях на основе гравитационного поля и формы астероида и сравнивают измеренные

коэффициенты сферических гармоник по сравнению с коэффициентами , которые были бы сгенерированы их предполагаемым профилем (73).

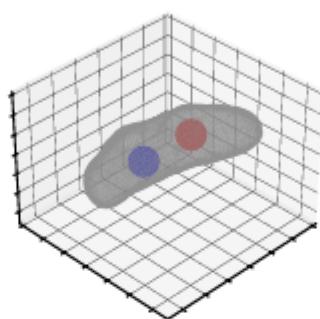
В лучшем случае этот процесс оставляет исследователям эвристические оценки профиля плотности тела

которые могут быть использованы в гетерогенных формах полиэдральной модели (73). Тем не менее, это остается

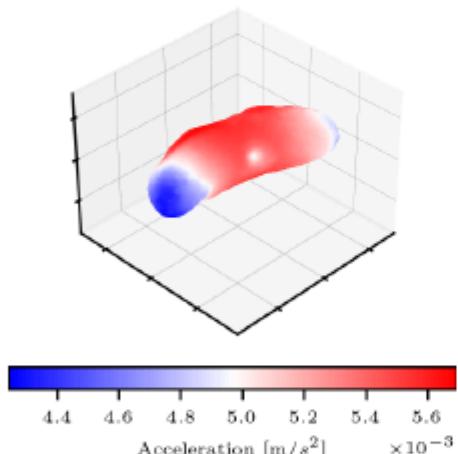
обычная практика - просто исходить из предположения о постоянной плотности. В этом разделе рассматриваются

последствия этого выбора и подчеркивается, как PINN-GM может обойти многие последствия

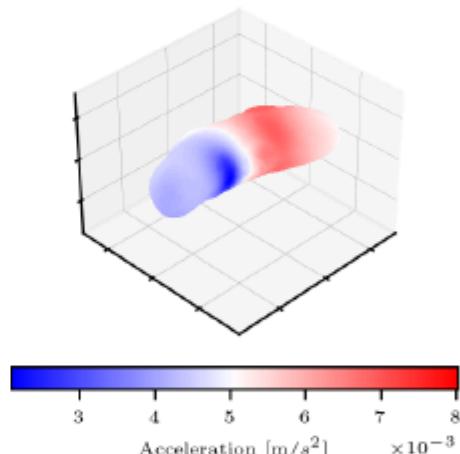
этой практики.



(a) вызванная неоднородность



(b) гравитация (предполагаемая постоянная плотность)



(c) гравитация (с неоднородностью)

Рисунок 3.32: Астероид с неоднородностью массы (вверху) и соответствующим полем притяжения на поверхности , когда предполагается постоянная плотность в сравнении с истинным полем (внизу слева и внизу справа соответственно).

Для этого эксперимента астероиду 433-Эрос вводится небольшая неоднородность массы.

В одном полушарии элемент массы добавляется к внутренней части тела, а в другом полушарии,

элемент массы удаляется. Каждый элемент массы содержит 10% от общей массы астероида,

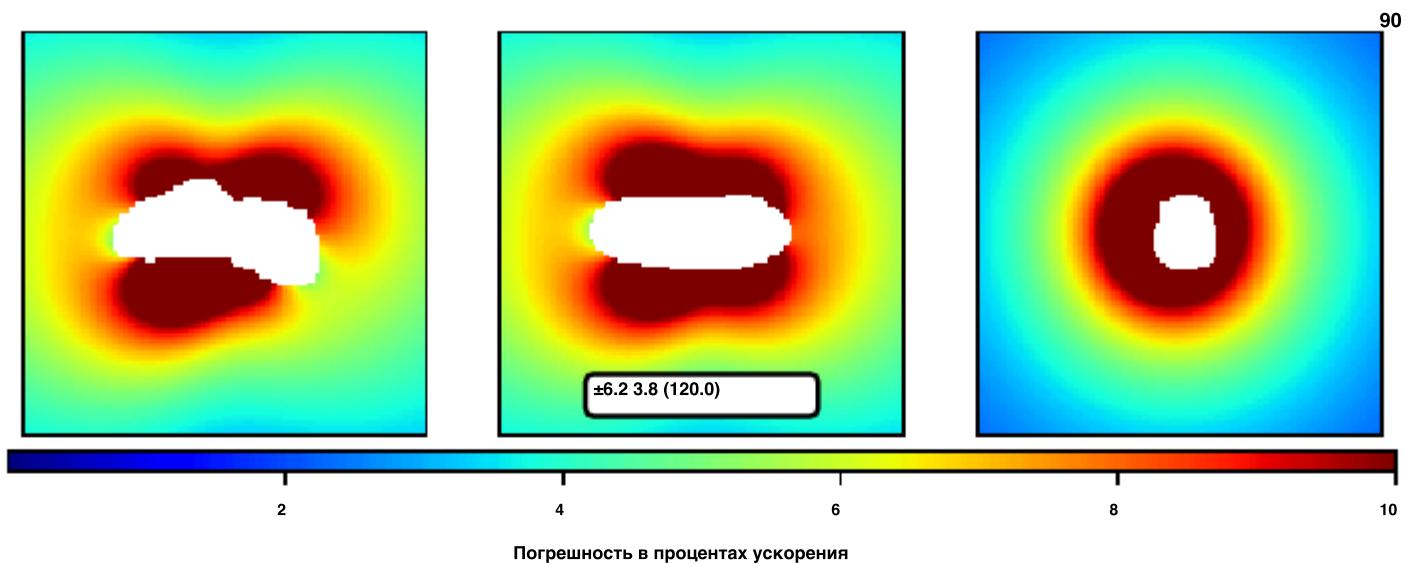


Рисунок 3.33: Ошибка многогранного ускорения постоянной плотности, указанная в виде μ

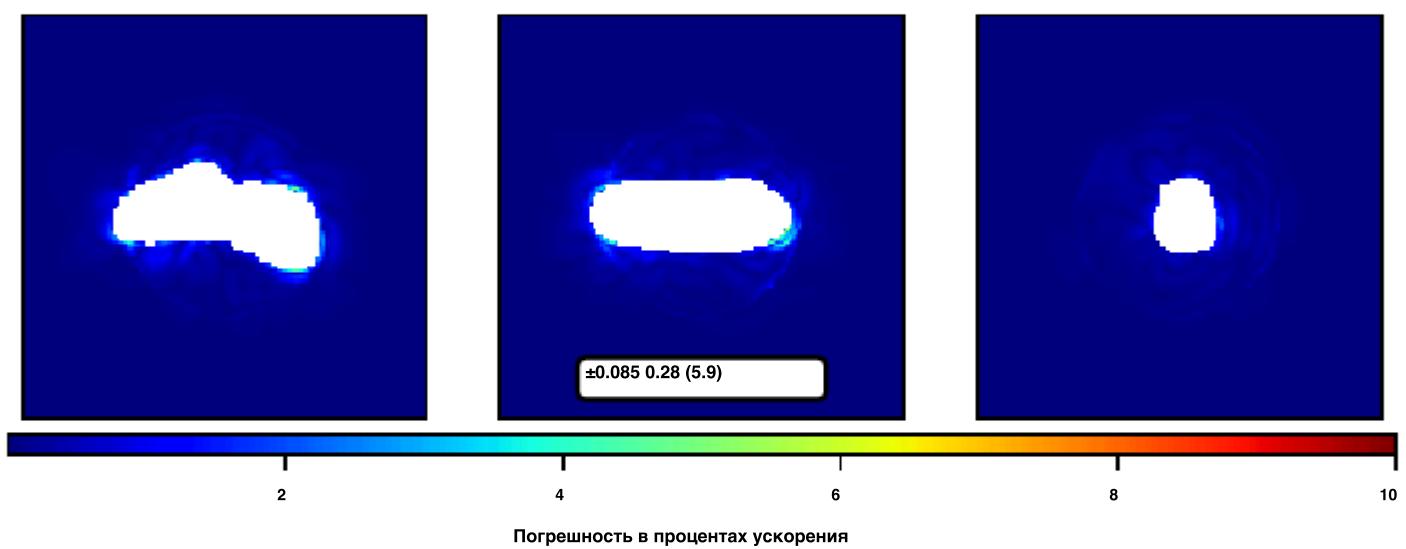


Рисунок 3.34: Ошибка ускорения PINN-GM-III указана в виде μ

и они симметрично смешены вдоль оси x на $0,33R$ (см. Рисунок 3.32a). Этот выбор

предназначен для имитации сценария, в котором два астероида с разной плотностью могли слиться в

один. Выбор сделать каждый элемент массы 10% от общей массы обоснован на основе литературы

с аналогичными распределениями плотности кандидатов (96).

Эта небольшая неоднородность может иметь большое значение для гравитационного поля астероида. Рис-

уре 3.32б изображает поле в предположении постоянной плотности, тогда как рисунок 3.32с Шоу

истинное гравитационное поле неоднородной плотности. Из этих цифр очевидно, что постоянная

предположение о плотности может привести к совершенно разным гравитационным характеристикам и, соответственно, к пространственному-

траектории космических аппаратов.

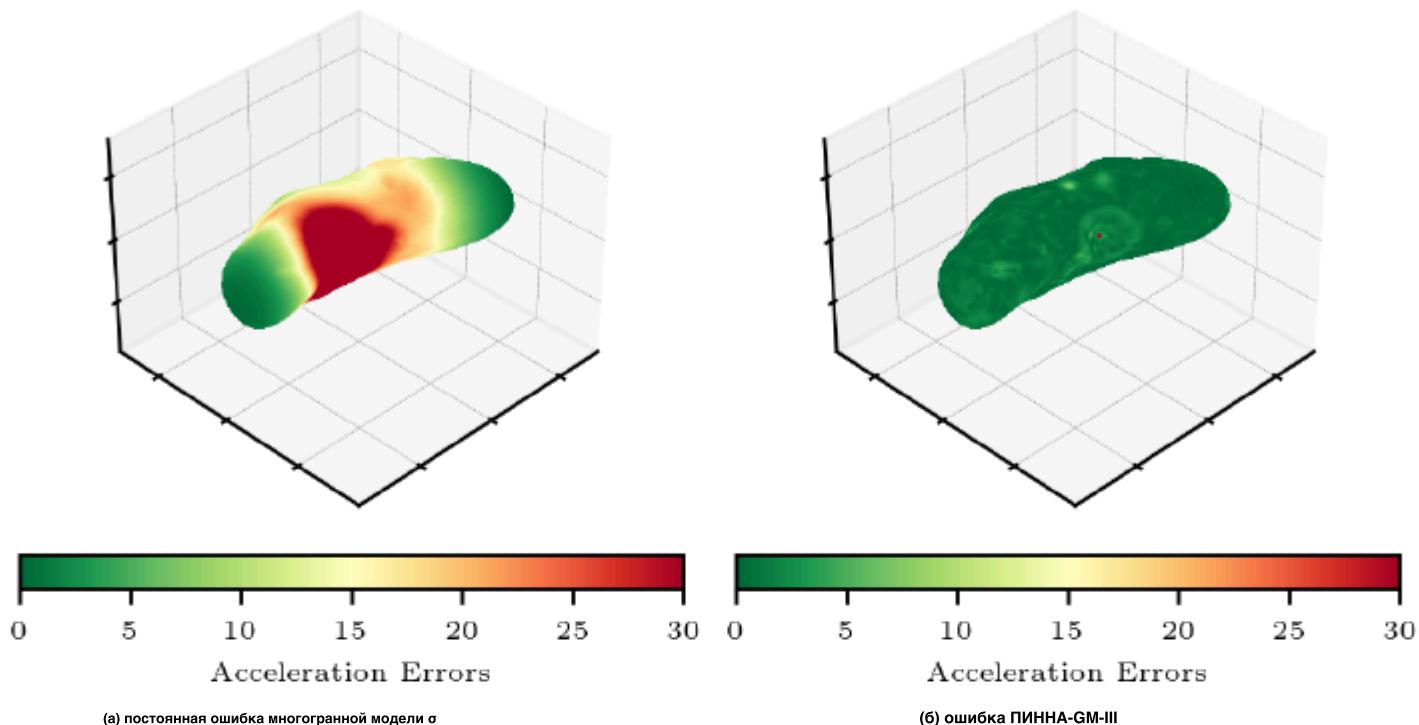


Рисунок 3.35: Поверхностная погрешность многогранной гравитационной модели постоянной плотности и PINN-GM-III.

Рисунок 3.33 иллюстрирует ошибку ускорения многогранной модели с постоянной плотностью на плоскости XY, XZ и YZ изменяются на 3R и сообщают о среднем значении, стандартном отклонении и максимуме ошибки ускорения. Предположение о постоянной плотности становится все более неверным вблизи поверхности, получения предельной ошибки, превышающие 100%. Ошибки этих величин внес значительный риск для миссий, которые стремятся к земле или работать в непосредственной близости от астероида.

В качестве альтернативы, PINN-GM-III обучается представлять гравитационное поле астероида неоднородной плотности. Очевидно, что обучающие данные для PINN-GM представляют собой надмножество двух наборов данных: первый набор данных содержит сведения исключительно выборку по периметру модели формы, где каждый положение точки данных соотносится с одним из аспектов модели полиэдрической формы (200,700 граней). Второй набор данных одного миллиона точек данных, охватывающих период от поверхности астероида на 10 радиусы. При экспериментах в разделе 3.4.2 изученные данные-низкая производительность PINN-GM-III, этот эксперимент целенаправленно изучает производительность в среде, богатой данными, при достаточном обучении

образцы существуют.

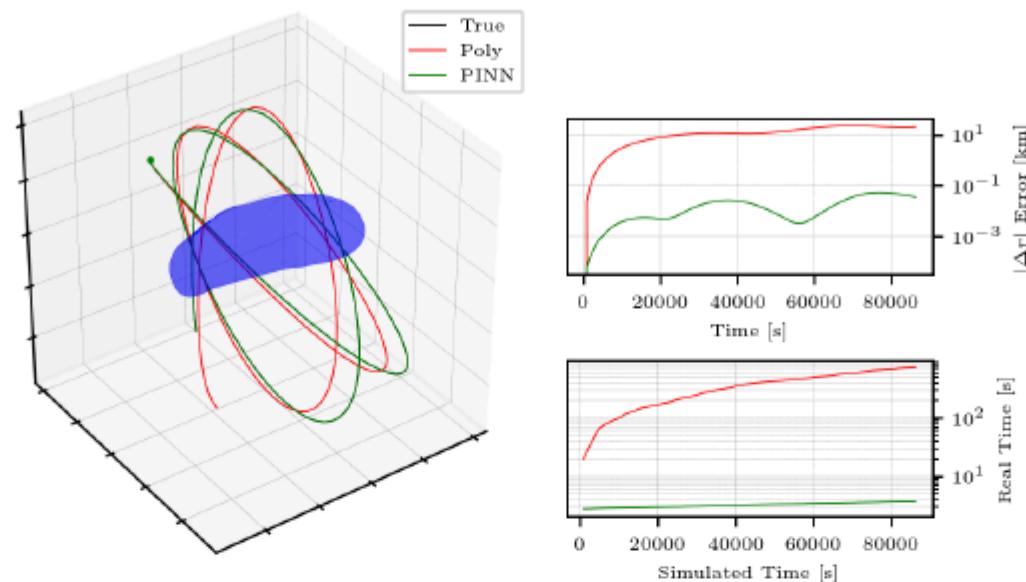


Рисунок 3.36: Скорость распространения и погрешность многогранной модели постоянной плотности и PINN-GM- III.

Производительность обученного PINN-GM визуализирована на рисунках 3.34 и 3.35. Средняя и максимальная погрешность PINN-GM почти на два порядка величины ниже, чем у постоянной многогранной модели плотности как на больших высотах, так и на поверхности. Обратите внимание, что PINN-GM-III содержит всего около 3000 параметров в отличие от многогранной модели, которая включает более 200 000 граней и 100 000 вершин. Эти результаты демонстрируют, что PINN-GM-III находится в состоянии для достижения порядка уменьшения величины ошибки с помощью двух порядков меньше параметров в модели.

PINN-GM-III не только более точен и компактен, но и более эффективен в вычислительном отношении. Это показано на рисунке 3.36 в котором распространяется набор начальных условий один раз с использованием многогранной модели постоянной плотности и один раз с обученным PINN-GM-III. На рисунке показано, что для распространения космического аппарата в течение одного моделируемого дня требуется примерно 30 минут в режиме реального времени, тогда как для PINN-GM-III требуется всего три секунды - ускорение в 600 раз. Более того, рисунок 3.36 также подчеркивается риск принятия предположения о постоянной плотности, когда вращающиеся по орбите небольшие тела экзотической формы. После одного дня распространения состояния при постоянной плотности полиэдральная гравитационная модель траектории эволюционирует так, что нормы ошибок определения положения превышают

20 км от истинной траектории. PINN-GM, напротив, не делает никаких предположений и выдает нормы ошибок < 0,1 км.

Глава 4

Приложение I: обучение с подкреплением.

Обучение с подкреплением - это структура, используемая для обучения агентов, принимающих решения, выполнять некоторые действия.

сложное поведение. Это может проявляться в обучении автомобилей самостоятельному вождению (97), здание

рекомендательные системы, которые подсказывают, какое шоу netflix посмотреть следующим (98), или даже обучение шахматам

боты как превзойти международных гроссмейстеров (99). По своей сути обучение с подкреплением - это

выполняется путем неоднократного выполнения агентом действий в некоторой среде и получения

сигнала вознаграждения, указывающего, было ли действие хорошим или плохим. Например, если шахматный бот выполняет

ход, который ставит противника под контроль, получит положительную награду. Напротив,,

если бот выполнит ход, который поставит под угрозу безопасность его собственного короля, он получит отрицательную награду.

награда. Постоянно пробуя разные приемы и оценивая связанное с ними долгосрочное вознаграждение,

обучающий агент с подкреплением создает историю опыта, из которой он может извлечь больше уроков

сложное и желательное поведение.

Обучение с подкреплением носит чрезвычайно общий характер, и существует широкий спектр возможных применений-

это привлекло внимание многих научных и инженерных сообществ (100; 101).

Во многих случаях эти сообщества стремятся обучить агентов более сложному поведению или решениям-

которые превосходят современные методы. Область астрономики не является исключением.

За последние несколько лет многие исследователи начали изучать, как можно использовать обучение с подкреплением

для создания более гибких и надежных автономных космических аппаратов для исследования дальнего космоса.

Эти усилия включают изучение того, как космические аппараты могут разрабатывать более надежные алгоритмы наведения

способен справляться с неопределенностью окружающей среды (102; 103), определение траектории переноса кандидатов-