

## Нейронная сеть, основанная на физике

### Гравитационная модель III поколения

Джон Мартин<sup>1\*</sup>

и Ханспетер Шaub<sup>2</sup>

<sup>1\*</sup> Факультет аэрокосмической инженерии, Мэрилендский университет,

4298 кампус Dr, Колледж-Парк, 20742, Мэриленд, США.

<sup>2</sup> Энн и Х. Дж. Смид, Факультет аэрокосмической техники

наук, Университет Колорадо в Боулдере, 431 UCS, Боулдер,

80309, Колорадо, США.

\* Автор (ы)-корреспондент. E-mail (ы): [jmartin@umd.edu](mailto:jmartin@umd.edu);

Участвующие авторы: [hanspeter.schaub@colorado.edu](mailto:hanspeter.schaub@colorado.edu);

#### Абстрактный

Научное машинное обучение и появление нейронных сетей, основанных на физике (PINN), демонстрируют значительный потенциал в их способности находить решения сложных дифференциальных уравнений. За последние два года была проделана большая работа по разработке PINNs, способных решать задачу моделирования гравитационного поля, то есть изучать дифференцируемую форму гравитационного потенциала на основе оценок положения и ускорения. В то время как предыдущие гравитационные модели ПИННА (PINN-GMS) продемонстрировали преимущества в компактности модели, устойчивости к шуму и эффективности выборки; остаются ключевые проблемы моделирования, которые призван решить этот документ. В частности, в этом документе представлено третье поколение основанной на физике модели гравитации нейронной сети (PINN-GM-III), которая решает проблемы ошибки экстраполяции, смещения в сторону маловысотных выборок, числовой нестабильности на больших высотах, и соответствующих граничных условий посредством многочисленных модификаций конструкции модели. PINN-GM-III тестируется путем моделирования известного астероида с неоднородной плотностью, и его характеристики оцениваются с использованием семи основных показателей, которые демонстрируют его сильные стороны по сравнению с его предшественниками и другими аналитическими и численными гравитационными моделями.

Научное машинное обучение, Нейронные сети, Основанные на физике, Ключевые слова: Астродинамика, Гравитационное моделирование

## 1 Введение

Почти все задачи астродинамики связаны с силой тяжести. Будь то в небесной механике, оптимизации траектории, сближении космических аппаратов, осведомленности о космической обстановке или других проблемах, гравитация часто играет значительную - если не доминирующую - роль в динамике системы. Простая повсеместность этой силы является свидетельством ее значимости, но, несмотря на это, существует на удивление мало способов представить эту силу с высокой точностью. Построение высокоточных гравитационных моделей отныне называется проблемой гравитационного моделирования, и в настоящее время универсального решения не существует. Некоторые гравитационные модели хорошо работают для крупных планетных тел, но не работают при моделировании объектов экзотической формы, таких как астероиды и кометы. Другие модели могут работать с неправильной геометрией, но требуют допущений и сопряжены с высокими вычислительными затратами. По своей сути, все современные модели гравитации таят в себе уникальные подводные камни, которые препятствуют стандартизации во всем сообществе. Следовательно, исследователи изучают альтернативные стратегии решения проблемы гравитационного моделирования в надежде найти универсальную модель. Одно из обнадеживающих направлений исследований касается использования нейронных сетей и научного машинного обучения для обхода многих из этих прошлых трудностей.

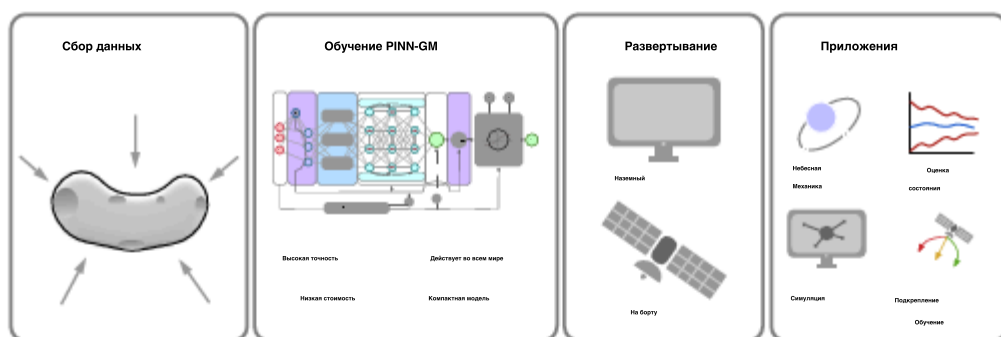


Рис. 1. PINN-GMs предлагают высокоточные и недорогие гравитационные решения, которые могут быть развернуты в различных вариантах использования и приложениях.

Научное машинное обучение предлагает использовать нейронные сети для изучения высокоточных моделей сложных дифференциальных уравнений и их решений [1, 2]. Нейронные сети, основанные на физике (PINN), являются одним из классов моделей, доступных для решения таких задач. PINN - это нейронные сети, специально обученные представлять решение дифференциального уравнения способом, учитывающим соответствующие дифференциальные и физические ограничения. Это соответствие, или "физическая информирование", достигается путем дополнения функции сетевой стоимости этими дифференциальными ограничениями, так что изученная модель наказывается за нарушение любой из лежащих в ее основе физических правил. Было показано, что благодаря этому изменению конструкции PINN-коды достигают более высокой точности при меньшем объеме данных, чем их традиционные аналоги из нейронных сетей [3].

Недавно исследователи начали изучать использование PINNs для решения проблемы моделирования гравитации [4, 5]. Эти модели, называемые моделями гравитации ПИННА (PINN-GMS), продемонстрировали значительные перспективы в своей способности извлекать высокоточные модели гравитации из разреженных и зашумленных данных в условиях как больших, так и малых тел. Несмотря на их предварительный успех, эти модели также продемонстрировали ряд недостатков, которые этот документ призван устранить. В частности, в этом документе представлено третье поколение PINN-GM (PINN-GM-III), которое решает проблемы ошибки экстраполяции, числовой нестабильности, смещения в сторону образцов с малой высотой и соответствия граничным условиям посредством многочисленных модификаций конструкции модели.

## 2 Общие сведения

Существуют два семейства гравитационных моделей: аналитические и численные.

Аналитические модели выводятся из первых принципов, таких как уравнение Лапласа, в то время как численные модели строятся на основе данных..... Каждая модель имеет свой собственный набор преимуществ и недостатков, и выбор того, какую модель использовать, часто диктуется областью применения. В следующем разделе подробно описаны доступные гравитационные модели и их соответствующие плюсы и минусы.

### 2.1 Модель сферических гармоник.

В начале 1900-х годов было предложено, чтобы сферические гармонические базисные функции могли быть наложены друг на друга для получения высокоточной оценки гравитационного потенциала [6]:

$$U(r) = \frac{\mu}{r} + \sum_{l=0}^{\infty} \sum_{m=0}^l \frac{R^l}{r^{l+1}} P_{l,m}(\sin \phi) C_{l,m} \cos(m\lambda) + S_{l,m} \sin(m\lambda) \quad (1)$$

Уравнение (1) называется сферической гармонической гравитационной моделью, где  $r$  - это расстояние до точки поля,  $\mu$  - гравитационный параметр тела,  $R$  - радиус, описывающий окружность тела,  $l$  - степень модели,  $m$  - это порядок модели,  $C_{l,m}$  - коэффициенты Стокса,  $\lambda$  - долгота,  $\phi$  - широта и  $P_{l,m}$  - связанные полиномы Лежандра [7].

Сферическая гармоническая гравитационная модель является основной моделью для выбора для представления гравитационных полей крупных планетарных тел, таких как Земля [8], Луна [9], и Марс [10]. Одним из наиболее убедительных преимуществ сферической модели гармонической гравитации является то, что она может компактно фиксировать одно из крупнейших гравитационных возмущений: сплюснутость планеты. Поскольку Земля и эти другие крупные тела вращаются вокруг своих осей, возникает центробежное ускорение, создающее сплющивание тела из сферы в сплюснутый эллипсоид [11]. Это уплощение или избыток массы вблизи экватора называется планетарной сплюснутостью, и его наличие оказывает значительное влияние на траектории космических аппаратов. Поэтому важно, чтобы это возмущение было точно зафиксировано в гравитационной модели. Сферическая гармоническая гравитационная модель упрощает это, требуя

только одна гармоника в расширении отражает это сжатие (C

2.0<sup>1</sup> или его

альтернативная форма  $J_2 = -C_{2,0}$ ).

В то время как сферические гармоники эффективны для представления сплюснутости планеты

, им трудно моделировать оставшиеся и более дискретные гравитационные

возмущения, такие как горные хребты, границы тектонических плит и кратеры.

Неоднородности, как известно, трудно представить с использованием функций периодического базиса,

часто требуется наложение сотен тысяч гармоник

для преодоления трехмерного эквивалента явления Гиббса [12]. Эти

гармоники должны не только накладываться друг на друга, они должны регрессировать, и

высокочастотные сигналы становится все труднее обнаруживать из-за их слабой заметности с (R/r) точки.

. ]].].].]. член в уравнении (1). Взяты вместе, эти условия

приводят к созданию больших моделей с неэффективной памятью, которые становятся сложно

содержать на борту космического корабля. Кроме того, эти высокочастотные модели могут потребовать

значительных вычислительных затрат, поскольку все связанные с ними полиномы Лежандра должны

быть повторно оценены в каждой точке поля. Эти вычисления являются рекурсивными и поэтому не

могут быть распараллелены, что приводит к неизбежному

, которое может серьезно ограничить как наземное моделирование, так и алгоритмы, основанные на сложности

борту [13].

Помимо вычислительной неэффективности, сферические гармоники также имеют

операционные ограничения. Вывод этой модели требует, чтобы все элементы массы

существовали внутри сферы фиксированного радиуса (радиус Бриллюэна). Если специалистам по

динамике потребуется оценка потенциала или ускорения в этой сфере, модель рискует

отклониться от (R/r)<sup>1</sup> срок в расширении. Хотя такие эффекты незначительны для планет или лун,

имеющих форму, близкую к сферической, они могут стать проблематичными в условиях малых

тел. Для астероидов или комет, которые обладают сильно несферической геометрией,

таких как Эрос или Итокава, эти эффекты могут привести к большим ошибкам в прогнозируемых

ускорениях и потенциально поставить под угрозу безопасность космических аппаратов [14].

## 2.2 Многогранная Гравитационная модель

Многогранную тяжести модели обеспечивает альтернативу сферической Гар-

Моник модель в этих условиях, предлагая решение, которое сохраняет

достоверность вниз на поверхность любого тела независимо от его формы. Такая

стабильность делает полиэдральную модель особенно популярной для исследований

малых тел, таких как астероиды и кометы. Если доступна модель формы тела - набор

треугольных граней и вершин, который отражает геометрию объекта, -

то связанный с этой геометрией гравитационный потенциал может быть вычислен

в предположении о постоянной плотности посредством:

$$\nabla U = -G\sigma \sum_{e \in \text{ребра}} E_e \cdot r_e + L_e + G\sigma \sum_{f \in \text{грани}} F_f \cdot r_f + \omega_f \quad (2)$$

где G - гравитационная постоянная,  $\sigma$  - плотность тела,  $E_e$  является

реберной диадой,  $r_e$  является вектором положения между центром ребра и

точкой поля,  $L_e$  является аналогом потенциального вклада ребра,  $F_f$  is

диада face normal,  $r_i$  - расстояние между нормалью грани и точкой поля, и  $\omega_i$  является аналогом потенциального вклада лица [14].

Хотя многогранная гравитационная модель позволяет обойти численное расхождение внутри сферы Бриллюэна, у нее есть свои уникальные недостатки. Прежде всего, эта 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, чтобы вычислить ускорение в одной точке поля. При вычислении многих ускорений или распространения траекторий, этот computational нагрузки может привести к чрезмерно длительным временем автономной работы. Хотя эта модель может легче использовать преимущества распараллеливания, чем сферические гармоники, такие вычислительные возможности недоступны на борту космических аппаратов, что по сути ограничивает эту модель наземным моделированием.

Другим недостатком многогранной гравитационной модели является ее допущения. 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 [15, 16]. Кроме того, многогранная гравитационная модель предполагает, что модель формы тела уже существует. Хотя эти модели могут быть получены на месте, этот процесс на практике является нетривиальным и отнимает много времени [17-19].

## 2.3 Другие Аналитические модели

Были предложены дополнительные аналитические модели, которые пытаются обойти некоторые проблемы, с которыми сталкиваются популярные сферические гармонические и многогранные модели. Модель эллипсоидной гармоники придерживается подхода, аналогичного подходу к сферическим гармоникам, но вместо этого использует базисные функции эллипсоидной гармоники [20]. Этот выбор позволяет получить более плотный эллипсоид, ограничивающий тело, чем сферические гармоники, сводя к минимуму область, в которой модель потенциально может расходиться. Несмотря на это, расхождение остается возможным внутри ограничивающего эллипсоида, и модель по-прежнему сталкивается с теми же проблемами при представлении разрыва с периодическими базисными функциями.

Альтернативно, модель внутренней сферической гармоники инвертирует классическую формулировку сферической гармоники и может моделировать локальную область, граница которой пересекает только одну точку на поверхности тела. [21]. Эта модель поддерживает стабильность вплоть до этой единственной точки, что делает ее ценной для точной посадки операций; однако решение становится недействительным в любой другой точке поверхности и за пределами соответствующей локальной сферы. Внутренняя сферическая гравитационная модель Бесселя расширяет эту тему, используя функции Бесселя, а не сферические гармоники, и способна обеспечить более широкую область применимости. Однако, это происходит за счет громоздкой аналитики и сохраняет некоторые из проблем, связанных с использованием гармонических базисных функций для улавливания неоднородностей [22].

В отличие от гармонических моделей, гравитационные модели маскона распределяют набор точечных элементов массы внутри тела, сумма которых может сформировать глобальное представление гравитационного поля [23]. К сожалению, ускорения могут становиться неточными в точках поля вблизи отдельных масконов [24]. Гибридные модели mascon предлагают несколько более надежную альтернативу чисто масконовскому подходу, представляя

каждый маскон имеет сферическую гармоническую модель с низкой точностью воспроизведения, но это включает в себя дополнительную сложность регрессии и остается подверженным как проблемам, связанным с масконными, так и сферическими гармоническими моделями [25].

2.4 Традиционные модели машинного обучения

В качестве альтернативы аналитическим моделям недавние попытки исследуют использование машинного обучения для изучения представлений сложных гравитационных полей на основе данных, управляемых способом. В принципе, модели машинного обучения могут изучать высокоточные гравитационные поля без каких-либо внутренних предположений о рассматриваемом теле, сохраняя при этом низкие вычислительные затраты.

Модель	Параметры	Данные для обучения	Средняя ошибка [%]	Действителен глобально
Общий результат [26]	12,960,000	3,600	1.5%	<i>x</i>
NNs [27]	1,575,936	800,000	0.35%	<i>x</i>
ВЯЗы[28]	100,000	768,000	1-10%	<i>x</i>
Геодезическая сеть [29]	91,125	500,000	0.36%	<i>x</i>
ПИНН-GM-III	2,211	4,096	0.30%	✓

Таблица 1: Статистика гравитационной модели машинного обучения - См. Приложение C

Гауссовские процессы

Первая предложенная попытка использовать машинное обучение для решения задачи гравитационного моделирования была представлена в 2019 году и использовала гауссовские процессы для изучения сопоставления положения и ускорения из набора обучающих данных [26]. Гауссовские процессы подходят для определения предварительного распределения по функциям и обновления этого предыдущего на основе наблюдаемых данных. Для этого требуется, чтобы пользователь указал 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.

Использование гауссовых процессов для моделирования силы тяжести выгодно, поскольку оно обеспечивает вероятностную оценку неопределенности в предсказании модели; однако модель плохо масштабируется для больших наборов данных. В частности, процесс Гауса-сиана характеризуется ковариационной матрицей, построенной на основе обучающих данных. Эта ковариационная матрица масштабируется как  $O(p^2)$  где  $p$  - размер обучающих данных. Такое масштабирование усложняет гауссовские процессы с вычислительной точки зрения для использования больших объемов данных, поскольку размер ковариационной матрицы и сложность соответствующей инверсии быстро растут при небольших наборах обучающих данных. В качестве примера, в справочной 26, гауссовский процесс регрессирует всего по 3600 точкам данных, в результате чего получается ковариационная матрица модели из более чем 12 960 000 параметров. Поскольку размер модели быстро растет, исследователи должны тщательно выбирать, какие обучающие данные использовать, и неизбежно оказываются неспособными

моделирует всю область тестирования, подчеркивая внутреннее ограничение этих моделей. Помимо этого факта, исследование демонстрирует, что эти модели, после установки, могут обеспечить быстрое время оценки и относительно высокую точность прогнозов, близких к данным обучения.

## Машины с экстремальным обучением

В качестве альтернативы, машины экстремального обучения (ELM) представляют собой однослойные нейронные сети, которые работают путем случайной инициализации весов из входных данных в скрытый слой, а затем вычисления весов для выходного слоя с использованием метода наименьших квадратов, чтобы минимизировать некоторую квадратичную функцию затрат, такую как:

$$L(\theta) = \frac{1}{N} \sum_{i=0}^N \|y_i - \hat{y}_i(\theta)\|^2 \quad (3)$$

где  $\hat{y}_i(\theta)$  является ли модель машинного обучения прогнозом на входе  $x_i$  с параметрами обучаемой модели  $\theta$  [30]. ELM имеет преимущество перед другими методами, поскольку его можно обучить за одну итерацию с использованием метода наименьших квадратов, в отличие от нейронных сетей, которые требуют множества итераций стохастического постепенного спуска и обратного распространения. ELMs были использованы для изучения сопоставления положения и ускорения астероида Итокава и снова демонстрируют высокую скорость оценки с относительно высокой точностью [28]. Тем не менее, для моделей требуются большие обучающие наборы данных (> 750 000 данных), длительное время обучения (приблизительно 40 часов), а их производительность не оценивается на больших высотах.

## Нейронные сети

В качестве альтернативы ELMs также были предложены нейронные сети для изучения гравитационной модели на основе данных о положении и ускорении [31]. Эти модели также обучены минимизировать функцию затрат, но делают это с использованием некоторой разновидности стохастического градиентного спуска - итеративного обновления весов сети с помощью оптимизационных алгоритмов, таких как ADAM [32], NADAM [33], LBFGS [34] и т.д. Преимущества использования нейронной сети заключаются в их способности представлять сильно нелинейные функции, но они традиционно чрезмерно параметризованы, что приводит к созданию больших моделей, которыми было бы сложно управлять непосредственно на борту. Подобно ELM, модели, обученные в эталонных [31] требовался большой набор обучающих данных (>1 000 000 данных), и их эффективность не оценивалась на больших высотах.

## Геодезическая сеть

В 2021 году, Ссылка 29 предложил использовать нейронные поля плотности для изучения карт плотности малых тел. Это исследование основано на недавних разработках машинного обучения в области нейронных полей излучения (NeRFs), которые используются для построения 3D-сцен на основе относительно разреженных данных изображения [35]. После обучения прогнозы плотности могут быть интегрированы численно для получения соответствующих гравитационных потенциалов и ускорений. В этой работе проявились promising результаты - достижение конкурентных ускорение точность и умение

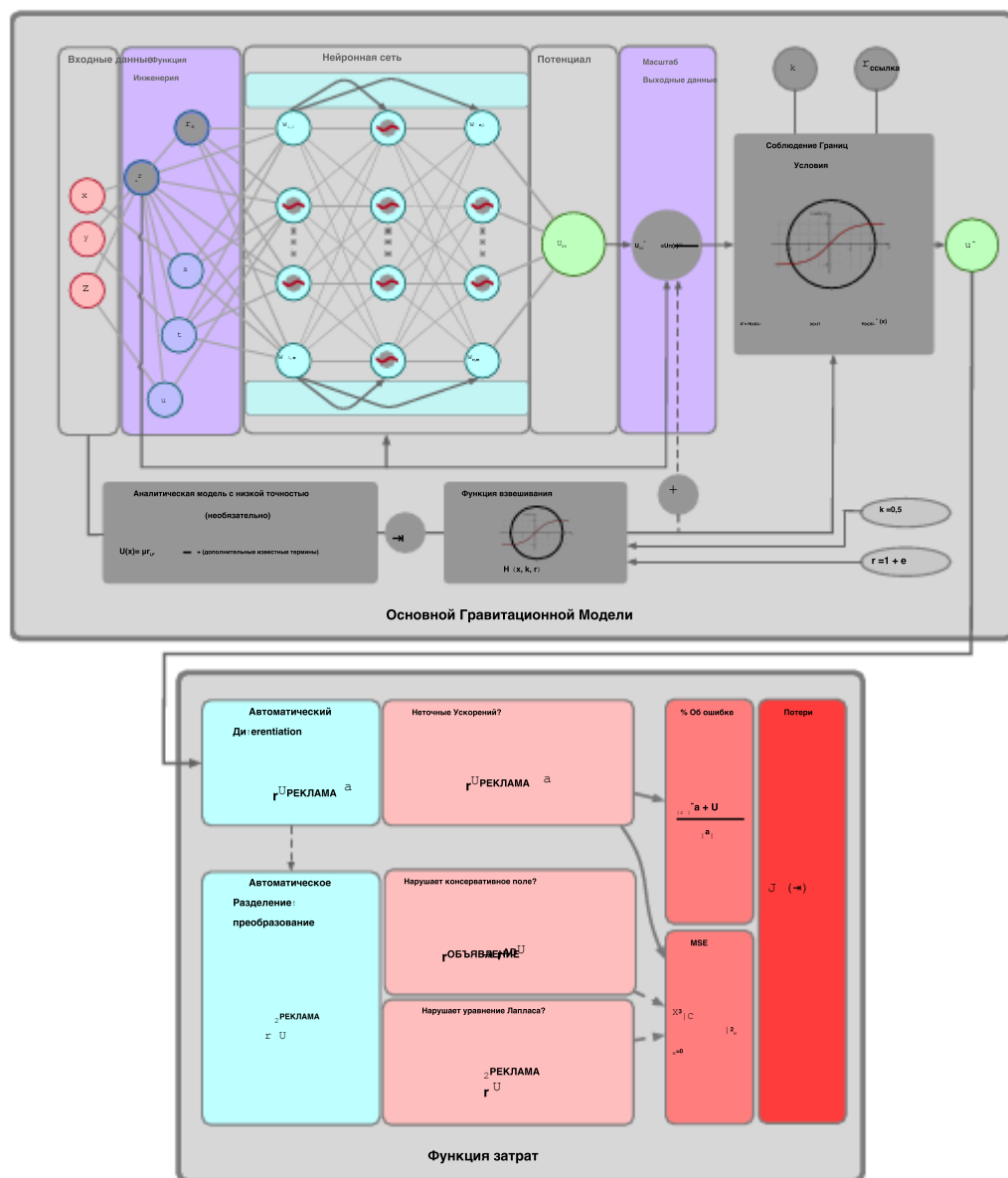


Рис. 2: PINN-GM поколения III с новыми модификациями, помещенными в темно-серые коробки

учитывают неоднородные плотности; однако модели относительно большие (более 90 000 параметров), численное интегрирование, необходимое для расчета ускорений, требует больших вычислительных затрат, и модели продолжали требовать больших обучающих наборов данных.

## 2.5 Модель гравитации нейросети на основе физики.

Как обсуждалось ранее, нейронные сети, основанные на физике (PINN), представляют собой принципиально новый класс алгоритмов машинного обучения, которые явно включают



известные дифференциальные ограничения на их функцию стоимости [3]. Наказывая за нарушение этих ограничений во время обучения, сети вынуждены изучать решения, которые естественным образом соответствуют лежащей в их основе физике.

В 2022 году Ссылка 4 предложил основанную на физике нейросетевую модель гравитации (PINN-GM) в качестве возможного решения проблемы гравитационного моделирования . Вместо подготовки традиционных нейронных сетей со стоимостью Функции, как то, что, показано в уравнении (3), PINN-GM использовал функцию затрат от которой использует известное дифференциальное уравнение а 
$$-\nabla U = \text{через}$$

$$L(\theta) = \frac{1}{N} \sum_{i=0}^N \left\| \nabla U(x_i) - \nabla U(x_i, \theta) \right\|^2_a \tag{4}$$

Вывод осуществляется с помощью автоматической дифференцировки нейронной сети [36]. где Ссылка 4 исследуется полезность этого подхода к моделированию в качестве альтернативы сферическим гармоникам для случаев Земли и Луны.. Результаты демонстрируют, что в некоторых высотных режимах PINN-GM может обеспечивать ускорение, сравнимое по точности с высокоточными моделями сферических гармоник, используя почти на порядок меньше параметров. Позже эта работа была расширена до настроек для малых тел, и были внесены дополнительные модификации дизайна, включая дополнительные физические ограничения, модифицированную архитектуру сети и разработку функциональных возможностей для получения высокоточных гравитационных моделей с использованием на порядок меньших наборов данных, чем в предыдущих подходах [5].

Несмотря на многообещающие разработки для гравитационной модели ПИННА, эта работа выявляет ранее незамеченные подводные камни этих моделей и предлагает конструктивные изменения для их устранения. В явном виде в этом документе представлена гравитационная модель ПИННА третьего поколения (PINN-GM-III), которая включает в себя множество конструктивных изменений, направленных на решение проблем численной нестабильности, ухудшения характеристик на больших высотах, ошибки экстраполяции и многого другого.

### 3 PINN-GM-III

Гравитационная модель PINN третьего поколения (PINN-GM-III) представляет собой коллекцию новых конструктивных решений для устранения различных недостатков прошлых поколений, одновременно повышая точность и надежность модели. Архитектурный обзор этих изменений показан темно-серым цветом на рисунке 2. Все модификации дизайна подробно описаны в подразделах ниже.

#### 3.1 Предварительная обработка и разработка функций.

Продуманная предварительная обработка обучающих данных для нейронных сетей может улучшить результаты обучения [38]. Одной из наиболее распространенных практик является нормализация данных для существования в пределах [-1, 1], поскольку многие функции активации их наибольшей нелинейности в этом режиме. Нелинейности - это то, что обеспечивает демонстрируют нейронные сети свои мощные возможности аппроксимации, и если входные данные насыщают функции активации, то возможности сети по моделированию заключаются в

## Алгоритм 1 Алгоритм PINN-GM-III

- 1: Сбор обучающих данных (x, a) из: (a) ранее существовавшей модели (b) оперативных оценок состояния [37]
- 2: Безразмерность обучающих данных ▷ Сек. 3.1
- 3: Преобразовать в неособые сферические координаты ( $r_n$ ,  $r_e$ , s, t, u) ▷ Сек. 3.1
- 4: Распространяются по нейронной сети
- 5: Выводят прокси-потенциал  $U_{NN}$  ▷ Сек. 3.3
- 6: Преобразуйте прокси-потенциал в истинный потенциал  $U_{NN}$  ▷ Сек. 3.3
- 7: (Необязательно) Предохранитель с взвешенным потенциалом низкой точности  $U^*$
- 8: Наложите граничные условия на потенциал  $U^*$  ▷ Сек. 3.4
- 9: сети U Автодифференцируйте потенциал (AD)
- 10: для получения ускорения a если обучение, то
- 11: Суммируйте процент ошибки ускорения и ▷ Сек. 3.2
- 12: среднеквадратичное значение для формирования потерь  $\hat{L}$  через РЕКЛАМУ
- 13: (Необязательно) Вычислите якобиан для Вычислите градиенты функции
- 14: потерь Обновите параметры сети
- 15: завершите, если

уменьшено. Кроме того, нормализация также улучшает числовую стабильность во время обучения и снижает риск некорректных числовых операций, если тренировочные значения слишком велики или малы.

Традиционно обучающие данные нормализуются независимо друг от друга, например, если входными данными являются x, а выходными - y, то нормализация приводит к  $[x \rightarrow [-1, 1]]$ . Однако для PINN данные  $x \in [x_{\min}, x_{\max}]$ ,  $y \in [y_{\min}, y_{\max}]$  могут разделять единицы измерения, и эта несвязанная нормализация может привести к несоответствию физике. Поэтому для задачи моделирования гравитационного поля важно нормализовать входные и выходные данные с учетом размеров. В PINN-GM-III это достигается путем нормализации положения и потенциала на характерную длину  $x^*$  равно радиусу планеты R и максимальному потенциальному значению в обучающих данных U соответственно. Используя эти характеристики скаляры, постоянная времени могут быть вычислены и использованы в сочетании с  $x^*$  для безразмерности ускорений. Явно это проявляется через:

$$x = \frac{x}{x^*}, \quad U = \frac{U}{U^*}, \quad a = \frac{a}{a^*} \quad (5)$$

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

$$x^* = R, \quad U^* = \max(U_n) \quad \text{или} \quad U^* = U \text{ ЕСЛИ, я),} \quad a^* = \frac{x^*}{\tau_t} \quad (6)$$

где R - максимальный радиус небесного тела,  $U_n$  является ли истинный гравитационный потенциал на тренировочном уровне на уровне:  $U_n$ , x и  $U_{LF}$  есть ли какие-либо малоточные устройства потенциальные вклады, уже учтенные в PINN-GM (обсуждаются)

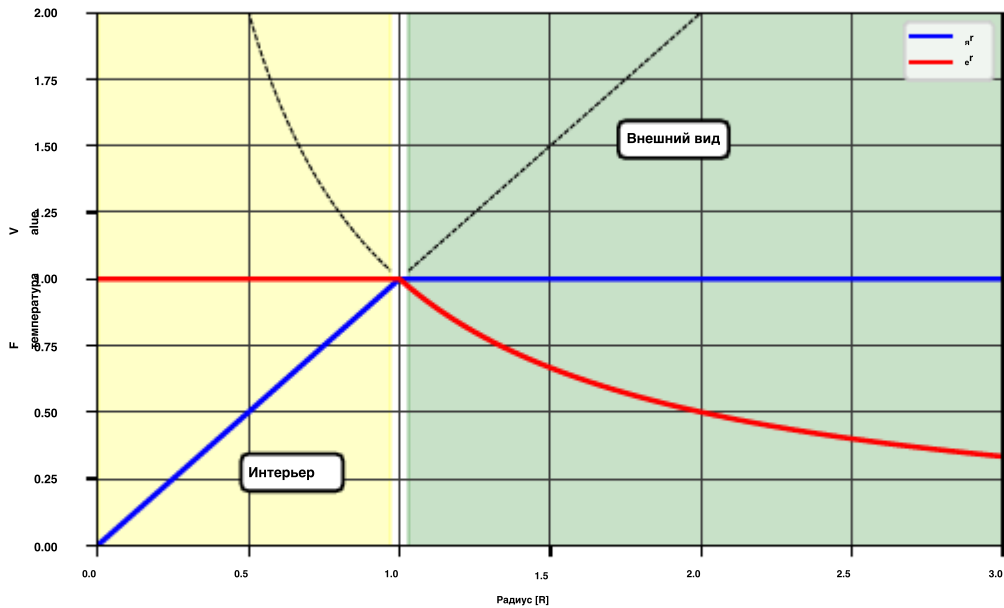


Рис. 3: Визуализация  $r_n$  и  $r_e$  возможности.

в Разделе 3.5), и

$$t' = \frac{\sqrt{x}}{U} \quad (7)$$

Помимо предварительной обработки данных, тщательная разработка характеристик входных данных сети может улучшить численную стабильность и сходимость сети. Для PINN- GM-III безразмерные декартовы координаты положения преобразуются в 5-мерное сферическое координатное описание  $(r_n, r_e, s, t, u)$  где  $r_e$  и  $r_n$  являются двумя прокси радиуса точки поля  $r$ , определенными как

$$r_n = \begin{cases} r & r \in [0, R] \\ 1 & r \in [R, \infty) \end{cases} \quad \text{и} \quad r_e = \begin{cases} 1 & r \in [0, R] \\ \frac{1}{r} & r \in [R, \infty) \end{cases} \quad (8)$$

Как показано на рисунке 3, эти прокси-переменные для расстояния гарантируют, что сетевые входные данные со значениями  $r \rightarrow 0$  или  $r \rightarrow \infty$  никогда не поставят под угрозу числовую стабильность сети. Аналогично,  $s, t$  и  $u$  представляют собой синус угла между точкой поля и каждой декартовой осью, определенной как  $x/r, y/r$  и  $z/r$  [39]. Эти преобразования гарантируют, что все входные объекты существуют в численно желаемом диапазоне  $[1, 1]$  независимо от местоположения точки поля.

### 3.2 Модифицированная функция потерь для учета большой высоты

#### Образцы

Функция потерь для исходного PINN-GM представляет собой

среднеквадратичное значение (RMS) показатель ошибки:

$$L_{\text{RMS}}(\theta) = \frac{1}{N} \sum_{i=0}^N \left( \hat{\nabla U}(\theta) - \nabla U(\mathbf{x}_i) \right)^2 \quad (9)$$

Эта функция потерь часто используется в задачах регрессии машинного обучения, поскольку она минимизирует наиболее явные расхождения между истинными и прогнозируемыми значениями.

Для PINN-GM это означает минимизацию разницы между разностным потенциалом сети,  $U = a$ .

, чтобы удовлетворить  $\nabla U(\mathbf{x}) = \nabla U$  уравнение.  $\theta$ , и истинное ускорение,  $a$ .

Несмотря на свою популярность, эта функция потерь имеет неожиданный недостаток при применении к задаче гравитационного моделирования.

Гравитационные ускорения, возникающие ближе к небесному телу, имеют гораздо большую величину, чем ускорения, возникающие на больших высотах. Как следствие, даже небольшие относительные ошибки в прогнозах для малых высот будут казаться непропорционально большими по сравнению с любыми ошибками для больших высот. Это означает, что гравитационные модели, обученные с использованием функции среднеквадратичной стоимости, всегда будут отдавать приоритет точному моделированию скорости в низковисотных режимах в ущерб высотным режимам, даже если высотные прогнозы более ошибочны в относительном смысле.

Чтобы решить эту проблему, PINN-GM-III дополняет исходную функцию среднеквадратичных потерь дополнительным членом ошибки в среднем проценте за счет:

$$L_{\text{PMC}\%}(\theta) = \frac{1}{N} \sum_{i=0}^N \left( \hat{\nabla U}(\theta) - \nabla U(\mathbf{x}_i) \right)^2 + \frac{1}{N} \sum_{i=0}^N \left| \frac{\hat{\nabla U}(\theta) - \nabla U(\mathbf{x}_i)}{\nabla U(\mathbf{x}_i)} \right| \quad (10)$$

Благодаря включению процентной потери погрешности, PINN-GM-III больше не зависит исключительно от абсолютных величин векторов ускорения, но вместо этого также минимизирует относительные ошибки. Это изменение конструкции гарантирует, что все выборки, независимо от высоты, вносят существенный вклад в функцию потерь и распределяются сетью по приоритетам во время обучения.

Чтобы проиллюстрировать влияние новой функции потерь на производительность модели, предлагается тест, который обучает две PINN-GM на 1000 парах данных о местоположении / ускорении, распределенных на расстоянии 0-50 радиусов над поверхностью астероида Эрос. Один PINN-GM обучается с использованием исходной среднеквадратичной функции потерь, а другой обучается с использованием расширенной функции потерь. После обучения каждая сетевая работа оценивается на наборе из 10 000 случайно распределенных тестовых точек от 0 до 50R, а процентная погрешность их векторов ускорения представлена как функция высоты на рисунке 4. Рисунок 4a подтверждает, что сети, обученные с помощью функции среднеквадратичных потерь, непропорционально отдают предпочтение низковисотным полевым точкам (< 5R) в ущерб высотным полевым точкам (> 5R), несмотря на то, что они были обучены во всей области 0-50R. Напротив, модель ПИННА-GM, обученная с помощью расширенной функции потерь (рисунок 4b), отдает приоритет точному моделированию на всех высотах.

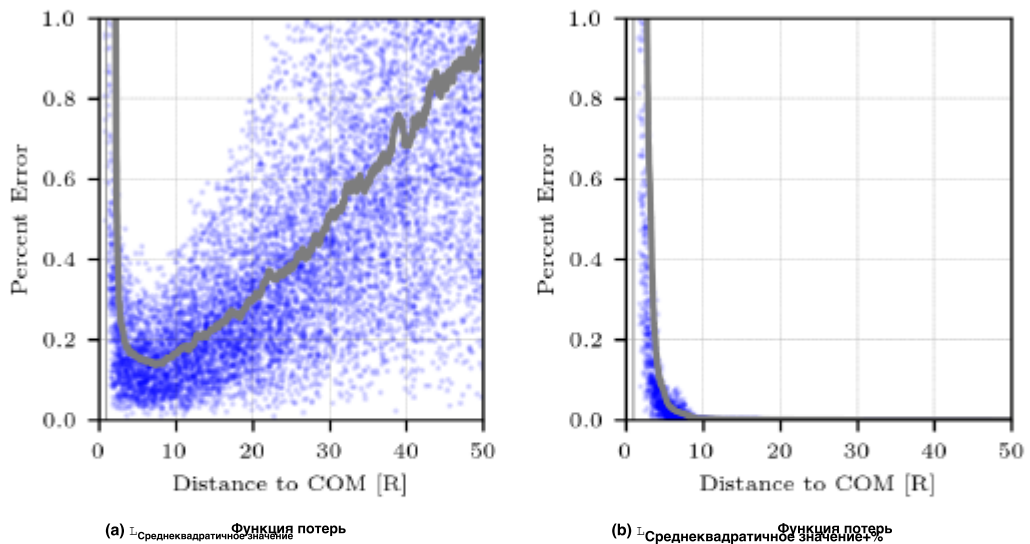


Рис. 4: Различные функции потерь изменяют производительность сети на больших и малых высотах. Синие точки обозначают индивидуальные ошибки в тестовых данных, а серая линия - среднюю ошибку теста в пределах скользящего окна в 250 пунктов.

### 3.3 Улучшайте числовые показатели, изучая прокси для Потенциального

Все гравитационные потенциалы затухают на больших высотах в соответствии с обратным степенным законом, равным  $1/r$ . Это затухание создает численные проблемы при эксплуатации моделей PINN-GM на большой высоте. Учтите, что наибольший гравитационный потенциал, представленный сетью, не имеет размерности, равной единице из уравнений (5) и (6). Для полевых точек на достаточно большой высоте нейронной сети потребуется вычислить потенциалы, которые затухают до значений, меньших или равных точности станка (т.е.  $U(r) \leq \epsilon_{\text{машинна}}$ ). Представляя их в значительной степени критичные числовые масштабы для одной и той же нейронной сети нежелательны и могут привести к числовой нестабильности во время обучения и вывода, а также к ненужному ограничению максимальной высоты, для которой жизнеспособна модель.

PINN-GM-III решает это явление, изучая более благоприятный с числовой точки зрения показатель потенциала, который позже может быть преобразован в правильный порядок величины. Явно, PINN-GM-III узнает прокси для потенциала, обозначаемого как  $U_{NN}$  и определяемого как

$$U_{NN} = U * n(r); \quad n(r) = \begin{cases} 1 & r < R \\ r & r > R \end{cases} \quad (11)$$

где  $U$  - истинный потенциал, а  $n(r)$  - масштабирующая функция. При введении этой функции масштабирования прямой выходной сигнал нейронной сети,  $U_{NN}$  всегда остается ограниченным и центрированным относительно безразмерного значения  $\mu$ . Такое центрирование снижает риск ошибки в расчетах и позволяет избежать преждевременной укупорки

точность модели на больших высотах. После того, как сеть выдает свое значение прокси-потенциала, оно затем явно преобразуется в истинное потенциальное распределение / правильный порядок величины путем деления на предписанную функцию масштабирования через

$$\hat{U}_{NN} = \frac{U_{NN}}{n(r)} \quad (12)$$

Эффект этого изменения показан на рисунке 5 который показывает истинный потенциал и косвенный потенциал для астероида Эрос в диапазоне 0-10R. Вместо того, чтобы изучать распределение, которое непрерывно затухает на больших высотах, сеть вместо этого изучает распределение, центрированное и ограниченное относительно истинного, безразмерного значения  $\mu$ . По мере увеличения высоты приближение массы точки становится все более репрезентативным для истинной системы, позволяя распределению сходиться к истинному значению  $\mu$ .

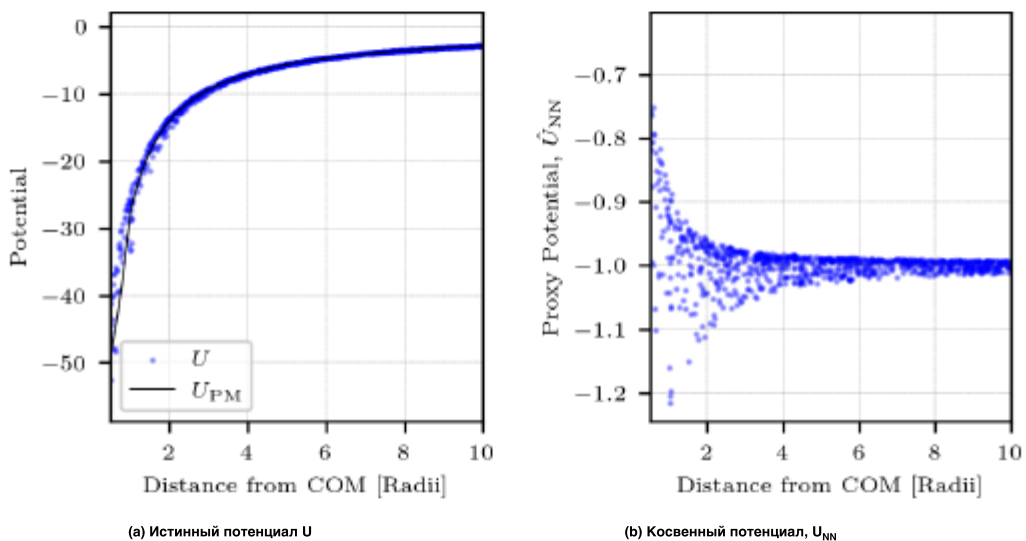


Рис. 5: Истинный потенциал  $U$  (слева) быстро уменьшается до численно неблагоприятных значений, тогда как по сравнению с косвенным потенциалом  $U_{NN}$  (справа) остается центрированным и ограниченным относительно безразмерного значения  $\mu$ .

### 3.4 Выполнение граничных условий с помощью модели

#### Проектирование во избежание ошибки экстраполяции

Хотя PIN-коды чаще всего разрабатываются для удовлетворения физических требований с помощью их функции затрат, существуют и другие способы обеспечения соответствия с помощью разработки самой модели машинного обучения. Ссылка 40 показано, как можно спроектировать модели машинного обучения для плавного сочетания сетевого решения и известных аналитических граничных условий за счет использования функций, подобных Хевисайду .

Для достижения аналогичного эффекта PINN-GM-III предлагает следующую конструкцию модификацию для обеспечения соблюдения соответствующих граничных ограничений:

$$\hat{U}(r) = w_{NN}(r) \hat{U}_{NN}(r) + w_{BC}(r) U_{BC}(r) \quad (13)$$

- предсказанный гравитационный потенциал,  $U_{BC}$  где  $\hat{U}_{NN}$  это потенциал на в граничное условие, и  $w_{NN}$  и  $w_{BC}$  это веса, зависящие от высоты для каждый определяется как:

$$w_{До н.} (r, k, r_{ссылка}) = H(r, k, r_{ссылка}) \quad (14)$$

$$w_{NN} (r, k, r_{ссылка}) = 1 - H(r, k, r_{ссылка}) \quad (15)$$

где  $H(r)$  - функция, основанная на Хевисайде, определяемая как

$$H(r, k, r_{ссылка}) = \frac{1 + \tanh(k(r - r_{ссылка}))}{2} \quad (16)$$

где  $r$  - радиус точки поля,  $r_{ref}$  - опорный радиус, а  $k$  - параметр сглаживания для управления более непрерывным или дискретным переходом.  $r_{ссылка}$  обычно выбирается максимальная высота границ обучения, а рекомендуемое значение  $k$  равно 2. Этот эвристический выбор уравнивает уверенность в том, что модель быстро перейдет к аналитической модели вне обучающих данных без слишком быстрого применения перехода, что изменит градиент потенциала и внесет ошибку в прогнозы ускорения. Визуализация этих изменяющихся весов показана на рисунке 6.

Уравнение (16) обеспечивает плавный переход PINN-GM в известное граничное условие при превышении контрольной высоты  $r_{ссылка}$ . Таким образом, модель машинного обучения может использовать гибкость моделирования нейронной сети для представления сложных областей гравитационного поля вблизи тела, для которых граничное условие не имеет значения, но затем плавно уменьшать ответственность сети в пределе по мере приближения модели к границе.

Для задачи гравитационного моделирования существует множество способов, которыми может проявляться этот выбор конструкции. В пределе  $r \rightarrow \infty$  потенциал

уменьшается до нуля, как обсуждалось в разделе 3.3. Однако, установив  $U_{BC} = 0$  и  $r_{ссылка} = \infty$

Уравнение (13) непрактичен, поскольку требует, чтобы нейронная сеть должна

изучить модель потенциала для всей области  $r \in [0, \infty)$ . Более полезным выбором чтобы использовать информацию, полученную из сферической гармонической гравитационной модели, и распознать является то, что высокочастотные компоненты гравитационного потенциала затухают

до нуля быстрее, чем вклад точечной массы на больших высотах. т.е.

$$U_{BC}(r) = U_{LF} = \frac{1}{r} + \sum_{l=0}^n \sum_{m=0}^l \frac{r_l}{r^{l+1}} - \dots \quad (17)$$

как  $r \rightarrow \infty$ .

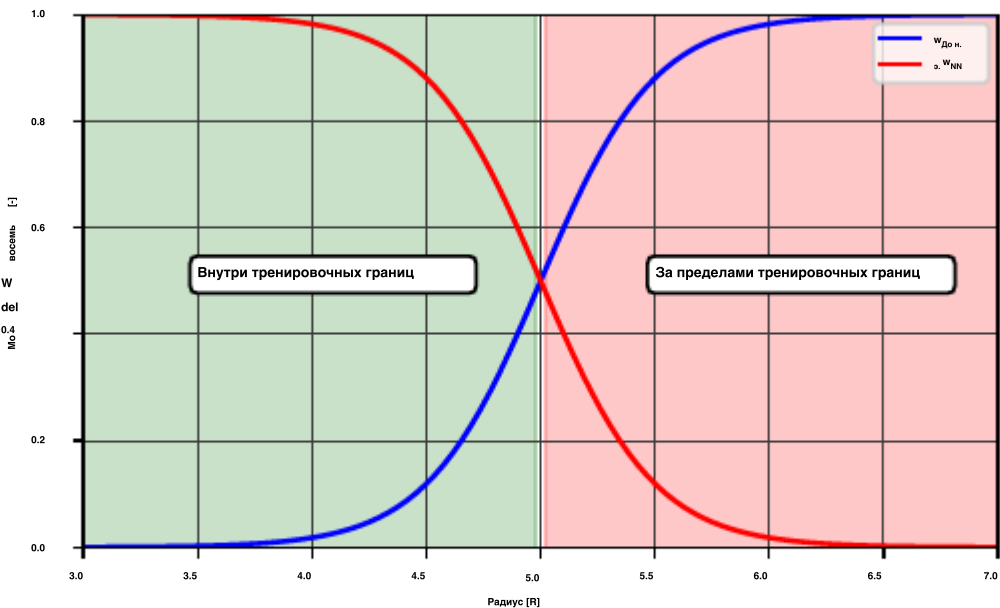


Рис. 6. Визуализация относительных весов моделей аналитики и нейросети, определенных с помощью  $H(r, k, r_{\text{ссылка}})$ .

Это наблюдение подразумевает, что  $U$  вставим образом, PINN-GM-III устанавливает  $U_{\text{BC}} \approx$   $U_{\text{ан.}} + f(r)$  в уравнении (13), где  $f(r)$  -  $r$  любые члены более высокого порядка в модели сферической гармонической силы тяжести, которые пользователь знает априори и желает оставить как часть граничного условия.

### 3.5 Использование ранее существовавшей информации о силе тяжести в Решение ПИННА-ГМ

В дополнение к обеспечению соблюдения граничных условий посредством проектирования модели, еще одним конструктивным решением, предоставляемым PINN-GM-III, является возможность объединения предыдущих гравитационных моделей с нейросетевым решением. Например, большинство крупных небесных тел демонстрируют планетарную сплюснутость, которая кратко описывается с помощью символа  $C_{2,0}$  коэффициент сферической гармоник, как описано в разделе 2. Вместо того, чтобы требовать от сети повторного изучения этого заметного и легко наблюдаемого возмущения, эта информация может быть непосредственно объединена с сетевой моделью. Таким образом, PINN-GM-III может предсказывать ускорения, используя низко-точную аналитическую модель первого порядка с сетью, ответственной за улавливание возмущений высокого порядка посредством:

$$\hat{U}(r) = w_{\text{NN}} U(r) + \hat{u}_{\text{LF}}(r) + w_{\text{BC}} U_{\text{LF}}(r) \tag{18}$$

где  $U_{\text{LF}}$  относится к известному, аналитическая модель с низкой точностью, такая как  $U_{\text{LF}} = -\frac{GM}{r^2} + U$ .



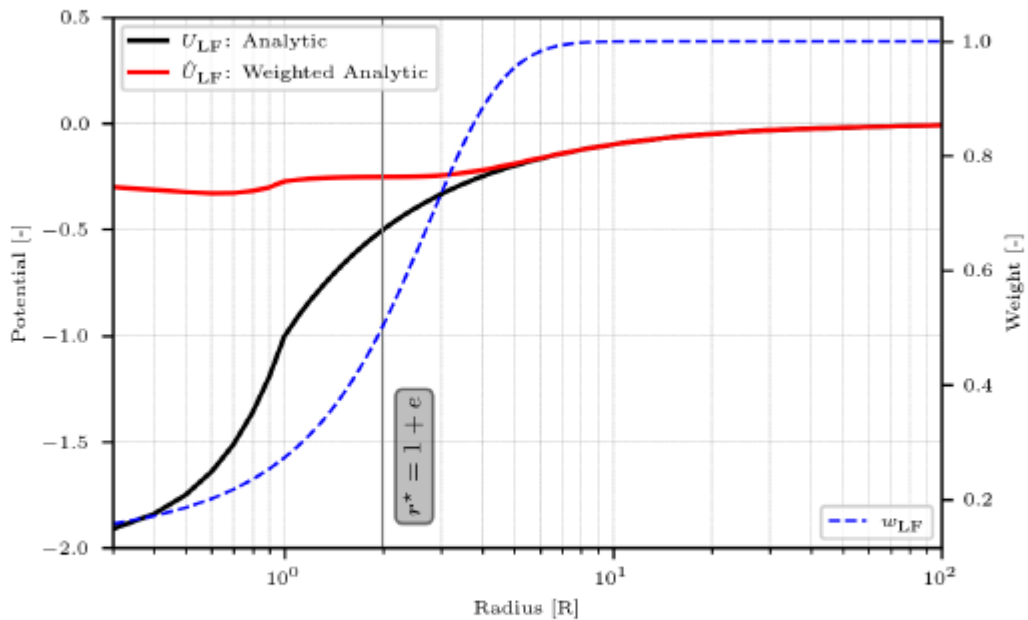


Рис. 7: Потенциал с низкой точностью U

становится  $U_{LF}^*$  перед слиянием с U

имеет пониженный вес

на малых высотах до  $U_{LF}^{NN*}$ 

В случае моделирования силы тяжести малых тел, где геометрия может значительно отличаться от точечной массы или низкочастотного сферического гармонического приближения, включение прошлых аналитических моделей должно выполняться осторожно, поскольку аналитические модели часто действительны только на больших высотах. Чтобы гарантировать, что эти аналитические модели используются только там, где это уместно, повторно используется гиперболическая функция слияния касательных  $H(x, r, k)$ . В частности, для небольших кузовов PINN-GM-III использует

$$u_{NN}^*(r) = w_{LF} U_{LF}(r) \quad (19)$$

как аналитический вклад в модель с низкой точностью, где  $w_{LF}$ 

=

$$H(r, R^*, k), R^* = 1 + e \text{ и } k = 0.5, \text{ где } e - \text{эксцентриситет тела, вычисляемый через } e = 1 -$$

$$b^2/a^2.$$
 Это гарантирует, что на большой высоте, где аналитическая аппроксимация наиболее точна, она в полной мере вносит вклад в

окончательное изученное решение, но на меньших высотах аналитический вклад имеет меньший вес

для окончательного решения. Визуализация этой взвешенной аналитической модели с низкой точностью

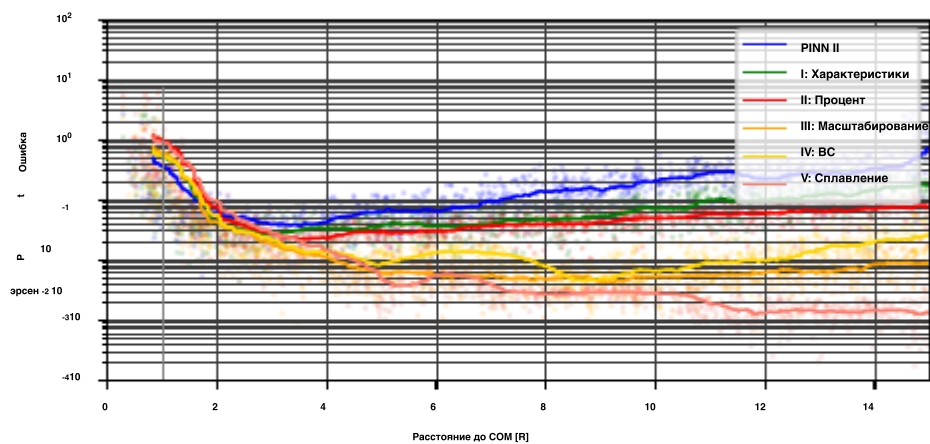
показана на рисунке 7, и показывает, как приближение точечной массы с низкой точностью

имеет пониженный вес на малых высотах, но полностью способствует на больших высотах.

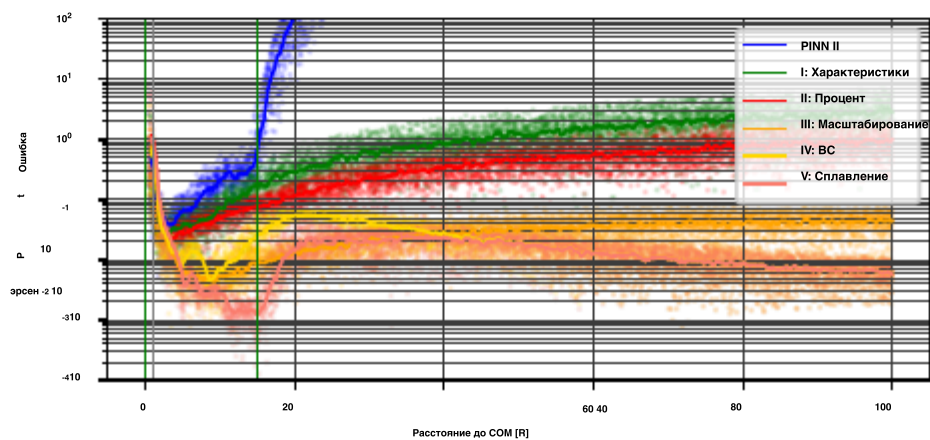
### 3.6 Анализ модификаций PINN-GM-III в модели

#### Точность

Чтобы визуализировать, как предлагаемые изменения влияют на точность модели, обучаются пять PINN-GM, последовательно добавляя каждую из перечисленных выше модификаций. Каждый PINN-GM обучается с использованием одного и того же набора обучающих данных из 5000 точек, равномерно распределенных между поверхностью астероида Эрос и высотой над уровнем моря



(а) Ошибка внутри границы обучения



(b) Ошибка за пределами тренировочных границ

Рис. 8: Процентная погрешность ускорения в зависимости от высоты после последовательного применения каждой из предложенных модификаций PINN III.

из 15R. Соответствующая ошибка ускорения указана в обучающем распределении на рисунке 8а, и вне обучающего распределения на рисунке 8б.

Рисунок 8б подчеркивает, что прежний PINN-GM-II не способен точно моделировать образцы на большой высоте и быстро отклоняется при тестировании за пределы обучающих данных. Применение новых радиальных характеристик (I) значительно уменьшает эту ошибку экстраполяции; однако ошибка на больших высотах все еще монотонно увеличивается и достигает максимума, в среднем, в 3% при 100R. Включение процентной ошибки в функцию потерь (II) дополнительно снижает эти ошибки - с 3% примерно до 1% - за счет снижения чувствительности к полевым точкам, расположенным на малой высоте. Прокси-потенциальная функция и (III) выпадает ошибка упали почти двух порядков - от 1% до 0,03%, хотя ошибка плато после

примерно 40R. Когда добавляется граничное условие (IV), происходит переход нейросетевого решения в аналитическое приближение точечной массы, и погрешность в пределе уменьшается до 0,007%. Включение граничного условия действительно приводит к небольшому снижению производительности в точке перехода, близкой к 15R, поскольку функция перехода вносит нежелательное изменение в градиент потенциала. Наконец, наложение аналитического потенциала на потенциал нейронной сети дополнительно уменьшает ошибку в пределах обучающих данных и возвращает к минимальной ошибке в 0,007% на пределе большой высоты. Обратите внимание, что существует числовой предел процентной ошибки, равный приблизительно 0,001%, когда различия между прогнозом модели и достоверностью на практике функционально идентичны.

## 4 Примерное исследование: Астероид неоднородной плотности.

Чтобы дополнительно проиллюстрировать возможности моделирования PINN-GM-III, проводится тематическое исследование, в котором PINN-GM-III обучается представлять гравитационное поле астероида неоднородной плотности, смоделированного по образцу 433-Eros. Гетерогенные эрогенные астероиды представляют собой особенно сложный сценарий для гравитационных моделей, поскольку распределение их внутренней плотности непосредственно не наблюдается. Часто делаются предположения о постоянной плотности, особенно в случае многогранной модели, но недавние результаты показывают, что это предположение не всегда справедливо [16]. Некоторые астероиды могут содержать над и под собой плотные области внутри своих недр или, возможно, были образованы в результате слияния двух астероидов, каждый из которых имеет разную плотность. Для захвата этих неоднородностей, dynamicists в настоящее время принимают обоснованные предположения об этих дистрибутивах, основанных на астероид гравитационного поля и фигуры и сравнивать измеренные сферические Гар-Моник коэффициентов в отношении тех, которые будут генерироваться по их предполагали профиля [41]. В лучшем случае этот процесс оставляет исследователям эвристические оценки профиля плотности тела, которые могут быть использованы в гетерогенных формах многогранной модели. Однако чаще всего обычной практикой остается просто исходить из предположения о постоянной плотности. В этом разделе рассматриваются разветвления этого выбора и подчеркивается, как PINN-GM может обойти многие из последствий этой практики.

### 4.1 Настройка проблемы.

Для этого эксперимента в астероид 433-Эрос были введены две небольшие неоднородности 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, и они симметрично смещены вдоль оси  $x$  на 0,5 R (см. Рисунок 9). Гравитационный вклад этих элементов массы накладывается на гравитационное поле многогранной модели с постоянной плотностью. Это имитирует гравитационное поле единого тела, образованного двумя слившимися астероидами с разной характерной плотностью. Выбор в пользу того, чтобы масса каждого элемента составляла  $\pm 10\%$  от общей массы, обоснован на основании литературы с аналогичными распределениями плотности кандидатов [42].

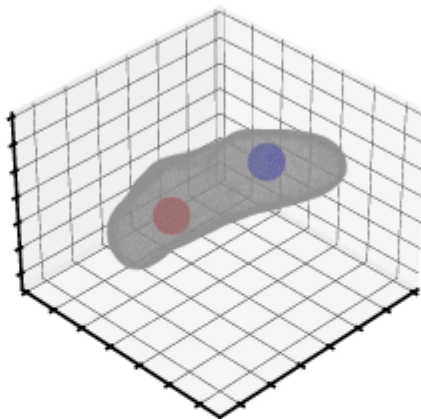
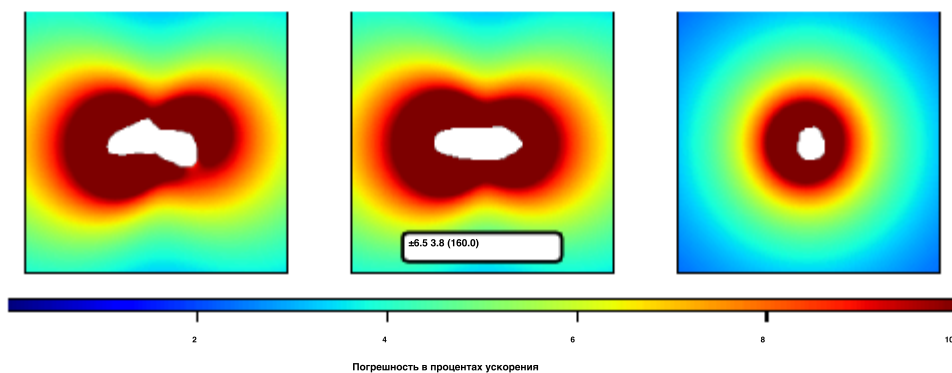


Рис. 9: Астероид с наложенными неоднородностями массы

Рис. 10: Многогранная погрешность постоянной плотности, выраженная в виде  $e_a$ 

$$\pm e_a(e_{\text{Maxc}}).$$

Эти небольшие неоднородности массы могут иметь серьезные последствия для соответствующего гравитационного поля астероида. Если специалисты по динамике продолжают моделировать тело в предположении постоянной плотности, траектории космических аппаратов будут быстро отклоняться от исходных и могут поставить под угрозу их безопасность на более длительных промежутках времени или во время критически важных операций. Эта ошибка лучше всего проиллюстрирована на рисунке 10 который показывает относительную ошибку ускорения многогранной модели с постоянной плотностью в плоскостях XY, XZ и YZ с точностью до 3R и сообщает среднее значение, стандартное отклонение и максимальные ошибки. На больших высотах предположение о постоянной плотности приводит к относительно небольшим ошибкам, составляющим в среднем приблизительно 4% погрешности. Однако на меньших высотах предположение о постоянной плотности становится все более ненадежным. В пределах 2 радиусов от центра масс средняя ошибка превышает 10% и достигает 160% вблизи поверхности.

Эти ошибки показывают, почему необходимы альтернативные гравитационные модели. Несмотря на то, что многогранная модель с постоянной плотностью предлагает наиболее убедительную основу для учета влияния неправильной геометрии тела на гравитационное поле, ее недостаточно для моделирования неоднородностей плотности. Кроме того, требуется модель формы

более 200 700 граней для представления полной геометрии с высоким уровнем точности. Даже если бы многогранная модель лучше подходила для учета неоднородностей, она все равно оставалась бы дорогостоящим вариантом из-за большого объема памяти и вычислительных ресурсов. PINN-GM предоставляет возможное решение этой проблемы с возможностью регрессировать поле независимо от геометрии тела и профиля плотности, используя значительно меньшее количество параметров.

Чтобы проиллюстрировать это, PINN-GM-II и PINN-GM-III обучаются на основе набора данных о положении и ускорении для регрессии модели этого сложного поля. Это достигается с использованием сетевой архитектуры из шести скрытых уровней с 32 узлами на каждом уровне (приблизительно 6500 параметров) и регрессируется при идеальных условиях передачи данных. Обучающий набор данных включает 90 000 точек данных, равномерно распределенных в диапазоне от 0 до 10R, а также 200 000 образцов, распределенных непосредственно на поверхности астероида. Эти оптимистичные условия данных используются для установления приблизительной верхней границы производительности PINN-GMS при данной параметрической мощности. Условия неоптимальных и разреженных данных исследуются далее в разделе 5.

## 4.2 Показатели

Производительность PINN-GM-II и III сравнивается с производительностью полиэдральной модели с постоянной плотностью с помощью комплексного набора показателей, описанных в следующем подразделе. Эти показатели исследуют производительность модели в пределах обучающих данных и за их пределами, а также характеризуют дополнительные представляющие интерес свойства, такие как время выполнения и накопленная ошибка траектории.

### Плоскости Метрические

Первый показатель оценивает среднюю процентную ошибку ускорения изученной гравитационной модели вдоль трех декартовых плоскостей (XY, XZ, YZ), расположенных между  $[-5R, 5R]$ , где R - радиус астероида. Поле оценивается на сетке точек размером 200x200 вдоль каждой плоскости, и средняя процентная ошибка вычисляется как

$$P = \frac{1}{N} \sum_{i=1}^N \frac{|a_{\text{верно}} - a_{\text{PINN}}|}{|a_{\text{истинный}}|} \times 100 \quad (20)$$

Этот показатель предназначен для обеспечения высокоуровневой оценки производительности модели в широком диапазоне рабочих режимов.

### Показатели обобщения.

Вторая, третья и четвертая метрики исследуют обобщение модели в диапазоне высот как в пределах границ обучения, так и за их пределами. В явном виде средняя ошибка ускорения оценивается как функция высоты, и разделяется на три режима тестирования: внутренний, внешний и экстраполяционный. Внутренняя метрика оценивает погрешность в пределах ограничивающей сферы радиуса R. Внешняя метрика исследует погрешность между R и максимальной высотой над уровнем моря

обучающего набора данных на уровне 10R. Наконец, показатель экстраполяции измеряет ошибку для высот за пределами обучающего набора данных в диапазоне от 10R до 100R. Обратите внимание, что для каждой единицы радиуса 500 образцов равномерно распределены по высоте для получения тестового набора.

### Метрика поверхности.

Пятый показатель оценивает среднюю ошибку ускорения по всем граням на поверхности модели формы с высоким разрешением. Для гетерогенной плотности модели 433-Eros, используемой в этом эксперименте, для вычисления среднего значения используется 200 700 фасетов. Этот показатель используется для характеристики производительности модели в наиболее сложной и динамически значимой области месторождения.

### Показатели траектории.

Шестой показатель оценивает накопленную ошибку распространения траектории за один день моделирования для космического аппарата, находящегося на низковысотной полярной орбите вокруг вращающегося астероида:

$$S = \int_{t=0}^T \Delta X_n(t) dt \approx \sum_{j=0}^N |\Delta_n X| I(\tau_j) \quad (21)$$

используемая орбита определяется как  $\{a, e, i, \omega, \Omega, M\} = \{32 \text{ км},$

$0,1, 90^\circ, 0, 0^\circ, 0^\circ\}$ , и астероид вращается со скоростью  $\omega_0$  градусов в секунду по оси z. Накопленная ошибка определения местоположения служит практическим показателем для определения того, какую величину ошибки могут ожидать проектировщики траекторий при использовании этих моделей. Наконец, седьмая метрика оценивает общее время, необходимое для распространения по орбите.

$= 0.00073$

## 4.3 Характеристики PINN-GM-III

Эти семь показателей используются для оценки эффективности трех моделей гравитации : полиэдральной гравитационной модели с постоянной плотностью высокого разрешения, PINN-GM-II и PINN-GM-III. Многогранная модель содержит 200 700 граней, а модели ПИННА-ГМ содержат шесть скрытых слоев по 32 узла в каждом, что соответствует размеру модели примерно в 6500 параметров. Поли- хедральная модель предварительно вычисляется независимо от обучающих данных, и PIN-коды обучаются с использованием распределения, описанного выше, и с использованием гиперпараметров по умолчанию, определенных в Приложении B. Соответствующие результаты показаны на рисунке 11.

Начиная с эксперимента по обобщению, верхняя часть рисунка 11 демонстрирует , что многогранная модель постоянной плотности дает наибольшую ошибку в тестовой области. PINN-GM-II значительно лучше работает вблизи поверхности, но точность ухудшается на больших высотах из-за неисправной функции среднеквадратичных потерь. Напротив, PINN-GM-III выдает погрешности от на один до трех порядков величины ниже, чем полигональная модель с постоянной плотностью, обеспечивая превосходные характеристики как на малых, так и на больших высотах. Модель

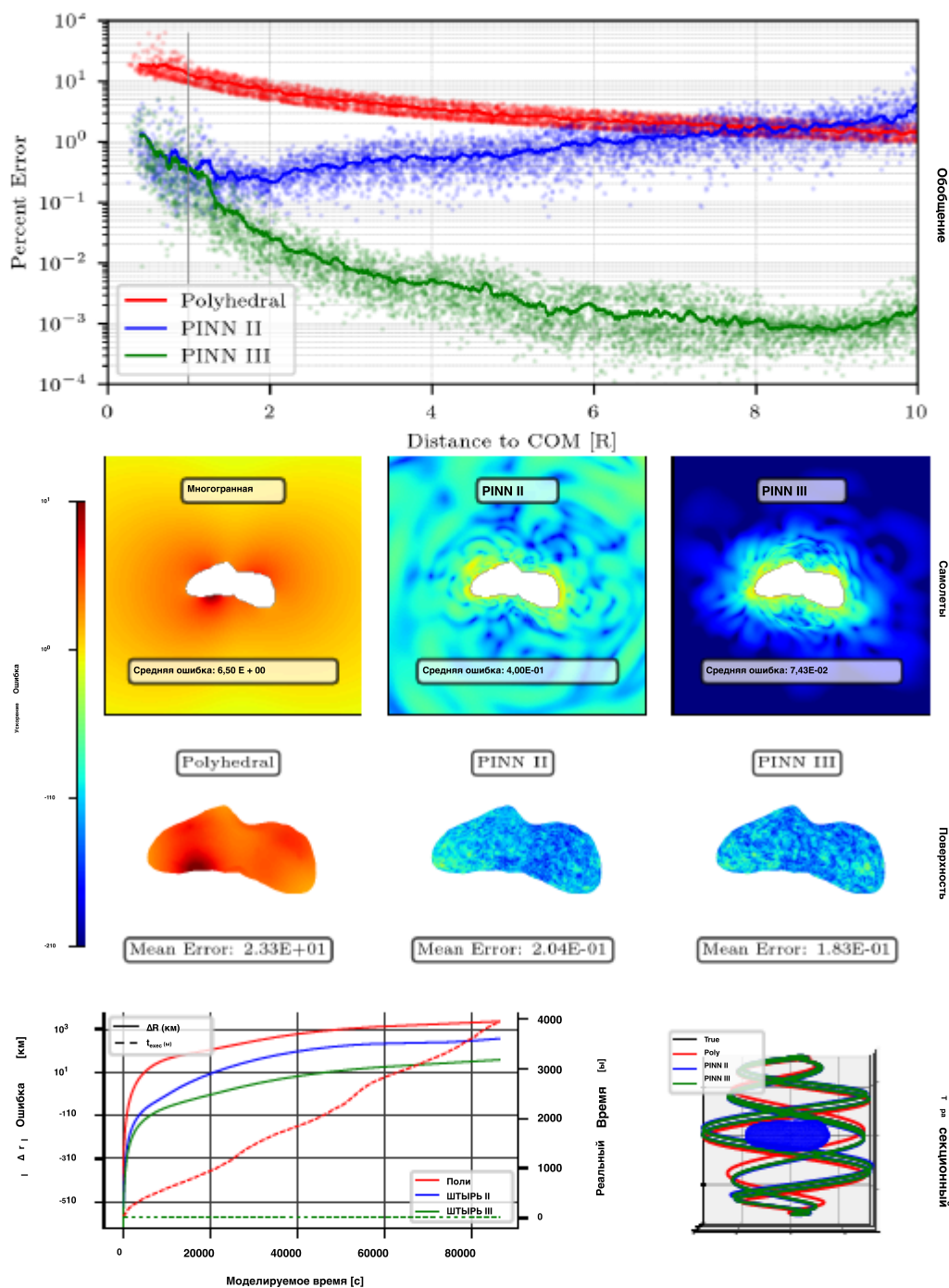


Рис. 11: Все предложенные показатели, оцененные для многогранной модели с постоянной плотностью, PINN-GM-II и PINN-GM-III.

плоскости и метрики поверхности дополняют это описание, как показано в средних строках рисунка 11. Здесь полиэдральная гравитационная модель дала среднюю погрешность ускорения в 6,5% по метрике planes, а PINN-III снижает ее на

почти на два порядка величины. На первый взгляд, аналогичный эффект показан с полиэдральной моделью, имеющей ошибку в среднем 23%, и PINN-III, снижающей эту ошибку до 0,18%.

Нижняя часть строки рисунка 11 показан эксперимент с траекторным распространением и продемонстрирована значительно более высокая скорость распространения PINN-II и меньшая ошибка распространения. Выведение космического аппарата на полярную орбиту вокруг вращающегося астероида в течение одного моделируемого дня полета требует более одного часа расчетного времени с использованием многогранной модели и накапливает более 2000 километров ошибки. Напротив, PINN-III требует всего 8 секунд для распространения и накапливает ошибку в 38 км. Обратите внимание, что более интуитивная мгновенная ошибка определения местоположения в конце орбиты составляет 4,6 км для многогранной модели и 0,1 км для PINN-III. Накопленная ошибка выбирается для построения графика, поскольку она обеспечивает монотонно возрастающую меру ошибки распространения траектории.

В совокупности PINN-III обеспечивает неизменно более высокую производительность, чем как его аналитические, так и численные предшественники, предлагая значительные улучшения в точности моделирования по широкому диапазону показателей при использовании значительно меньшего объема параметрических данных. Примечательно, что производительность PINN-GM может улучшиться благодаря моделям с большей производительностью и дополнительным данным для обучения. Хотя эти показатели производительности обнадеживают, они являются репрезентативными только для идеальных условий моделирования, для которых существуют идеальные данные о положении и ускорении, которые могут быть использованы для обучения. В следующем разделе рассматривается, что происходит при менее идеальных условиях обработки данных, и расширяется исследование для изучения эффективности PINN-III по сравнению с дополнительными аналитическими и численными моделями.

## 5 Сравнительное исследование

В этом разделе проводится сравнительное исследование для оценки точности PINN-GM-III по сравнению с другими популярными гравитационными моделями. Очевидно, что точечная масса (PM), сферическая гармоника (SH), маскон, многогранник, машина для экстремального обучения (ELM), традиционная нейронная сеть (TNN), модели PINN-GM-I и PINN-GM-II соответствуют различным условиям данных и затем оцениваются с использованием вышеупомянутых показателей.

Каждая модель подгоняется четыре раза, изменяя количество данных и параметрической емкости каждой модели. Очевидно, что каждая модель подходит один раз для 500 точек данных и один раз для 50 000 точек данных, чтобы проиллюстрировать чувствительность модели к количеству данных. В обоих случаях данные равномерно распределены в диапазоне 0-10R с использованием астероида одинаковой неоднородной плотности из раздела 4. Кроме того, каждая модель тестируется при "малой" и "большой" параметрической емкости. Эти метки соответствуют общему количеству параметров  $p$ , содержащихся в модели (например, коэффициентам Стокса для сферической гармонической модели, граням и вершинам в модели формы или общим весам и смещениям в нейронной сети). Каждая маленькая модель содержит примерно 250 параметров, тогда как большие модели имеют примерно 30 000 параметров. В совокупности эти различные данные и параметры  $g$  позволяют понять, какие модели способны поддерживать конкурентоспособную производительность в условиях нехватки данных и памяти. Примечание



	Плоскости		Обобщение		Поверхность		Пространственный		Вспомогательный	
	Ошибка [%]	Экстраполировать Ошибка [%] <sup>пр.</sup>	Внешняя	Ошибка [%] Ошибка [%] в внутренняя	Ошибка [%]	р значение	Ошибка [км] Время [с] Распространение	р значение	Регрессия Время [с]	
PM	52.8	53.4	53.5	44.6	69.6	1.2e5	0.0	1.0	0.7	
SH	D	0.3	5.0	D	D	4.4e3	0.1 1.6	240.0 204.0	4.3 NA	
Poly	6.2	0.3	3.3	15.9	23.3	1.9e3	0.6	220.0	4.9	
Масконы	3.1	0.3	2.8	4.4	41.4	8.2e3	6.8	280.0	2.2	
ВЯЗ	D	2.7	D	D	D	2.0e9	149.9	257.0	726.1	
TNN	14.4	D D	D	8.3	31.1	5.8e5	114.5	256.0	800.2	
ШТЫРЬ I	0.6	D	37.6	1.7	12.7	2.5e4	22.4	200.0	971.6	
ШТЫРЬ II	0.4	0.1	0.6	2.6	17.7	273.1	22.4	227.0	862.0	
ШТЫРЬ III			0.1			56.5				
N = 50000 образцов;                      Размер модели = "Маленький"										
PM	NA	NA	NA	NA	NA	NA	NA	NA	NA	
SH	D	0.3	5.0	D	D	4.4e3 2.3e3	0.4	3.1e4 3.0e4	8.2e3 NA	
Poly	6.5 0.0	0.4	3.5	16.5	D	2.7e3	132.3 6.2	3.0e4	471.7	
Масконы	6.5 0.0	0.0	0.0	0.1	23.3 32.8	5.1e7	33.5	3.6e4	262.3	
ВЯЗ	D D	D	D	D	D 8.8	6.4e5	62.2	2.9e4	973.3	
TNN	10.4	D	D	D	2.8	1.1e4	35.1	2.9e4	995.3	
ШТЫРЬ I	0.0	D	8.3	4.2	A не	9.6	15.3	3.0e4	1.7e3	
ШТЫРЬ II	0.0	1.1	0.0	0.1...0.1	3.3	8.4	18.5	3.0e4	1.6e3	
ШТЫРЬ III			0.0							
N = 50000 образцов;                      Размер модели = "Большой"										
PM	49.9	50.3	50.5	43.2	69.6	1.2e5	0.0	1.0	0.6	
SH	25.7	0.3	5.0	D	D	4.4e3	0.1 1.6	240.0 204.0	2.7 NA	
Poly	6.2	0.3	3.3	15.9	23.3	1.9e3	0.5	220.0	2.1	
Масконы	2.7	1.8	1.9	12.3	D	6.4e3	0.9	280.0	0.0	
ВЯЗ	74.2	D	D	98.0	99.3	6.6e4	75.7	257.0	154.9	
TNN	D	D	D	D	D	4.5e5	84.0	256.0	129.0	
ШТЫРЬ I	38.3	D	75.2	23.8	42.0	4.7e4	23.9	200.0	133.2	
ШТЫРЬ II	2.5	0.3	2.9	8.2	30.4	3.0e3	14.0	227.0	128.7	
ШТЫРЬ III	1.5		0.4	8.6	31.3	1.1e3				
N = 500 образцов;                      Размер модели = "Маленький"										
PM	NA	NA	NA	NA	NA	NA	NA	NA	NA	
SH	D	0.3	5.0	D	D	4.4e3 2.3e3	0.3	3.1e4 3.0e4	97.7	
Poly	6.5 0.7	0.4	3.5	16.5 10.9	23.3 D	2.7e3	132.3 6.9	3.0e4	NA 138.6	
Масконы	75.8	0.0	0.0	97.3	99.1	6.5e4	1.1	3.6e4	4.3	
ВЯЗ	D	D	D	D	D	7.8e5	162.7	2.9e4	145.2	
TNN	49.3	D	D	37.8	37.7	5.5e4	51.0	2.9e4	150.3	
ШТЫРЬ I	0.7	D	D	5.5	25.9	196.5	15.4	3.0e4	177.6	
ШТЫРЬ II	1.6	0.4	0.2...0.2	17.3	51.0	406.6	20.8	3.0e4	173.9	
ШТЫРЬ III										
N = 500 образцов;                      Размер модели = "Большой"										

Таблица 2: Показатели силы тяжести, оцененные после регрессии по данным бесшумного обучения.

что каждая модель имеет несколько отличающиеся процедуры регрессии и размер модели расчеты, которые подробно описаны в Приложении D.

Регрессия с идеальными данными

Эксперимент начинается с получения идеальных тренировочных данных о положении и ускорении без ошибок. Показатели, связанные с каждой регрессированной моделью, представлены в таблице 2. Как в условиях большого объема, так и в условиях нехватки данных, а также при использовании больших и малых параметрических моделей производительность PINN-GM-III по семи показателям остается конкурентоспособной, если не превосходит, все предыдущие аналитические и численные модели. Очевидно, что в условиях большого объема данных, когда N = 50 000, модель PINN-III обычно является наиболее эффективной моделью как в больших, так и в малых параметрических конфигурациях. Как показано в первых двух подтаблицах в таблице 2,

PINN-III превосходит аналитическую модель только по общему

времени распространения траектории; однако меньшее время распространения этих моделей достигается за счет более высокой общей накопленной ошибки траектории.

По проценту погрешностей PINN-III уступает PINN-II только на поверхности и

с небольшим отрывом. Более того, есть один случай, когда модель маскона

действительно имеет меньшую ошибку в областях за пределами обучающих данных;

однако этого следует ожидать, учитывая, что PINN-III переходит в приближение

точной массы с низкой точностью воспроизведения за пределами обучающих данных,

как описано в разделе 3.4.

В условиях нехватки данных, когда существует только  $N = 500$  данных в диапазоне от 0

до 10R, производительность всех моделей, как и ожидалось, ухудшается; однако PINN-III

остается наиболее надежным вариантом. Как показано в двух нижних подтаблицах в

Таблице 2, PINN-III - единственная модель, помимо многогранного решения, которая

никогда не расходится - где расхождение определяется как ошибки  $> 100\%$  и

обозначается буквой "D" в таблице. В то время как PINN-II имеет несколько меньшую

погрешность вблизи поверхности, а модель маскона имеет меньшую погрешность на

больших высотах, обе эти модели расходятся в других регионах, в то время как PINN-III - нет. Более

того, PINN-III остается относительно устойчивым к переобучению, даже когда

параметрическая емкость увеличена до  $p$

$$\approx 30,000.$$

Примечательно, что, несмотря на то, что производительность PINN-III является наиболее точным

вариантом в условиях большого объема данных и наиболее надежным в условиях нехватки

данных, следует отметить, что эти улучшения достигаются за счет увеличения времени регрессии. Все

результаты, сгенерированные в таблице 2 обучаются на 2-ядерном процессоре с доступом к

графическому процессору A100. В этой конфигурации процесс обучения PINN-GM неизменно

занимает больше всего времени - от 120 секунд для небольших моделей с небольшим объемом

данных до более чем 30 минут для самых больших моделей с наибольшим объемом данных.

Хотя время регрессии, представленное для аналитических моделей, неизменно

быстрее, следует отметить, что это может быть не всегда так. Все аналитические

модели, протестированные в этом эксперименте, регрессируют последовательно небольшими партиями, чтобы

избежать больших объемов памяти и времени вычислений, которые были бы необходимы для выполнения

полного обращения матрицы наименьших квадратов. Этот выбор приносит в жертву некоторую

производительность аналитических моделей, но в противном случае регресс этих моделей был бы

невозможен, учитывая доступные вычислительные ресурсы. И наоборот., время подготовки PINN-GMs

может быть значительно сокращено, если использовать мини-партии большего размера. Несмотря

на то, что используемые графические процессоры A100 обладают достаточным объемом памяти для

моделирования таких больших объемов партий, это достигается за счет использования несколько менее

точных моделей, как показано в исследованиях абляции Приложения В. Результаты, представленные в

таблице 2 являются попыткой сбалансировать эти факторы, но следует

подчеркнуть, что пользователи могут эффективно и точно распределять время регрессии как для

аналитических, так и для числовых моделей.

## Регрессия с зашумленными данными.

Прошлый эксперимент продемонстрировал, как работают аналитические и численные

модели гравитации в предположении о идеально измеренных тренировочных данных.

	Плоскости		Обобщение		Поверхность		Траектория		Вспомогательная	
	Ошибка [%]	Екстр Ошибка [%] <sup>90</sup>	Внешний	Ошибка [%] в terior	Ошибка [%]	Р размещение	Ошибка [с] Время [с]	Р аппа	Регрессия Время [с]	
PM	52.8	53.3	53.4	44.6	69.6	1.2e5	0.0 0.1	1.0	0.7	
SH	D	0.3	5.0	D	D	4.4e3	NA	240.0 NA	5.5 NA	
Poly.	NA	NA	NA	NA	NA	NA	0.5	220.0	5.6	
Масконы	3.3	3.0	3.0	4.6	41.6	9.0e3	7.8	280.0	2.5	
ВЯЗ	D	D	D	D	D	1.9e9	91.6	257.0	232.5	
TNN	16.3	D D	D	10.1	33.3	7.6e5	132.9	256.0	683.3	
ШТЫРЬ I	2.1	D-	38.4	3.2	15.1	4.8e3	29.7	200.0	141.4	
ШТЫРЬ II	3.9	0.3	2.9	8.3	35.0	1.4e3	15.8	227.0	37.8	
ШТЫРЬ III			2.0			3.0e3				
N = 50000 образцов;					Размер модели = "Маленький"					
PM	NA	NA	NA	NA	N	NA	NA	NA	NA	
SH	D	0.3	5.0	D	D	4.4e3 NA	0.4	3.1e4 NA	6.0e3 NA	
Poly.	D	NA	5.0	NA	D N	2.4e3	NA 6.2	3.0e4	471.9	
Масконы	NA	0.1	NA 0.2	9.3	D D D	4.7e7	22.4	3.6e4	255.7	
ВЯЗ	1.3	D	D	D	D	8.0e5	29.0	2.9e4	59.8	
TNN	13.8	D	29.9	D	17.0	1.3e4	28.8	2.9e4	77.8	
ШТЫРЬ I		D	2.2	6.1	8.9	799.9	15.3	3.0e4	81.3	
ШТЫРЬ II	1.4...3.4	D	2.0	2.7	12.7	6.3e3	20.2	3.0e4	66.9	
ШТЫРЬ III		0.3		3.7						
N = 50000 образцов;					Размер модели = "Большой"					
PM	50.1	50.6	50.7	43.3	69.6	1.2e5	0.0	1.0	0.7 1.5	
SH	25.1	0.3	5.0	D	D	4.4e3	0.1 NA	240.0 NA	NA	
Poly.	NA	NA	NA	NA	NA	NA	0.3	220.0	2.5	
Масконы	34.1	28.4	28.7	98.4	D	5.6e4	0.7	280.0	0.0	
ВЯЗ	75.6	D	D	97.9	99.3	6.6e4	19.8	257.0	27.2	
TNN	D	D D	D	D	D	6.6e5	86.0	256.0	127.0	
ШТЫРЬ I	32.8	D-	70.3	19.4	36.6	3.6e4	22.3	200.0	18.0	
ШТЫРЬ II	5.8	0.3	6.6	12.0	31.9	2.5e3	14.6	227.0	12.9	
ШТЫРЬ III	10.7		3.4	58.0	82.4	2.2e4				
N = 500 образцов;					Размер модели = "Маленький"					
PM	NA	NA	NA	NA	NA	NA	NA	NA	NA	
SH	D	NA	5.0 NA	D	D	4.4e3 NA	0.3 NA	3.1e4 NA	91.9	
Poly.	NA D			NA D	NA D	NA	NA	3.0e4	NA 157.5	
Масконы	77.6	0.3 NA	D 48.8	97.2	99.0	6.5e4	1.0	3.6e4	4.4	
ВЯЗ	30.5	0.6 D-D-D-D-D-D-	5.0	D	D	5.0e5	20.8	2.9e4	18.0	
TNN				17.4	33.2	3.2e4	40.8	2.9e4	17.9	
ШТЫРЬ I	5.9		A не	13.4	32.3	7.1e3	16.0	3.0e4	22.4	
ШТЫРЬ II	6.1	0.4	3.0	27.1	49.5	6.4e3	20.8	3.0e4	19.8	
ШТЫРЬ III										
N = 500 образцов;					Размер модели = "Большой"					

Таблица 3: Показатели силы тяжести, оцененные после регрессии по зашумленным (10%) обучающим данным.

Этот эксперимент рассказывает о том, как модели проезда по более сложной кон-  
ловиях, где обучающих данных может быть шумно. Для достижения этого выполняются те  
же самые регрессии, но каждый вектор ускорения в обучающих данных  
возмущается в случайном направлении на 10% от исходной величины ускорения.  
Эта ошибка выбрана в качестве преувеличенного тестового примера, чтобы определить,  
какие из этих моделей остаются надежными в более стрессовых условиях обработки  
данных. Многогранная модель намеренно исключена из этого эксперимента, поскольку не  
существует прямого способа включить неопределенность векторов ускорения в результирующую  
модель формы. Результаты этого эксперимента представлены в таблице 3.

Из зашумленной регрессии можно сделать такие же выводы, как и из бесшумной регрессии.  
PINN-III остается единственным надежным вариантом, который никогда не выходит из строя  
как в условиях большого объема данных, так и при их недостатке, а также для больших размеров

и небольшие размеры параметрической модели. Примечательно, что гравитационная модель *mascon* предлагает убедительные, а иногда и превосходные характеристики в условиях большого объема данных, но она склонна к отклонениям на малых высотах, в отличие от PINN-III. Важно отметить, что ни одна из гравитационных моделей не работает особенно хорошо на малых высотах в условиях нехватки данных. Хотя PINN-III официально никогда не расходится, он может давать ошибки, превышающие 50% внутри ограничивающей сферы и >80% на поверхности. Хотя это лучше, чем другие доступные аналитические модели, рекомендуется избегать работы в этих маловысотных режимах до тех пор, пока не будут собраны дополнительные данные, при которых производительность может достигать <10% погрешности во внутреннем режиме, как показано в условиях с большим количеством данных. На больших высотах PINN-III обеспечивает гораздо более удовлетворительную производительность с погрешностями менее 3% на всех высотах выше 1R.

В дополнение к основным показателям производительности, общее время регрессии, необходимое для соответствия моделям PINN, гораздо более конкурентоспособно в условиях зашумленных данных, чем в условиях идеально измеренных данных. Это результат критериев ранней остановки, используемых во время тренировки. Очевидно, что процесс обучения преждевременно приостанавливается, если потери при проверке не продолжают уменьшаться в течение конечного набора эпох. Это помогает предотвратить чрезмерную подгонку под ошибочные данные и, следовательно, приводит к гораздо более быстрому времени регрессии. При идеальных условиях данных этот критерий остановки никогда не срабатывает, следовательно, значительно увеличивается время обучения.

## 6 Выводов

Научное машинное обучение и нейронные сети, основанные на физике (PINNs), предлагают новый и привлекательный способ решения проблемы моделирования гравитационного поля. Вместо того чтобы использовать предписывающие аналитические модели силы тяжести, которые сопряжены с различными ограничениями и подводными камнями, ПИННы могут изучить удобные представления гравитационного потенциала, сохраняя при этом желаемые физические свойства и небольшие размеры модели. Несмотря на то, что предыдущие поколения PINN-GM дали ранее представление о потенциальных преимуществах этого класса гравитационных моделей машинного обучения, оставалось решить множество проблем и получить дополнительные возможности для развития. В этом документе представлена гравитационная модель ПИННА третьего поколения, PINN-GM-III, которая специально разработана для преодоления этих прошлых проблем.

Очевидно, что прошлые поколения PINN-GM изо всех сил пытались включить данные о большой высоте в свою регрессию, непреднамеренно отдавая приоритет выборкам для малой высоты в результате смещенной функции затрат. Эти предыдущие модели также страдают от численной нестабильности на больших высотах, склонны к ошибкам экстраполяции и не имеют способа включить предыдущие модели в свою регрессию. PINN-GM-III предлагает ключевые конструктивные изменения, такие как использование более надежной функции потерь, изучение косвенных потенциалов, включение известных граничных условий и использование многопозиционных точных моделей для решения этих и других проблем.

Вместе эти конструктивные изменения позволяют создавать более точные и надежные модели гравитации с машинным обучением, чем предыдущие поколения. При моделировании

несмотря на неоднородную плотность астероидов, PINN-GM-III неизменно достигает ничтожных, если не превосходящих, характеристик по сравнению с предыдущими поколениями и конкурирующими аналитическими моделями по целому ряду показателей. В лучшем случае PINN-GM-III демонстрирует улучшение точности на порядки по сравнению с многогранными моделями с постоянной плотностью при доле размера модели, обеспечивая более быстрое распространение в разы с меньшей накопленной ошибкой. Кроме того, эти модели поддерживать конкурентоспособные моделирования точность на больших и малых форм, и в обоих сведения богатые и скудные данные условия. В будущей работе будет продолжено изучение модификаций конструкции, которые могут улучшить характеристики модели, особенно возвращение к крупным небесным телам и изучение способов, с помощью которых высокочастотные компоненты могут быть изучены и представлены более эффективно с учетом разреженных данных.

## 7 Заявлений и деклараций

От имени всех авторов соответствующий автор заявляет, что нет конфликта интересов.

## 8 Благодарностей

В этой работе использовался высокопроизводительный вычислительный ресурс Alpine в Университете Колорадо в Боулдере. Alpine совместно финансируется Университетом Колорадо в Боулдере, Университетом Колорадо Анштутц, Государственным университетом Колорадо и Национальным научным фондом (премия 2201538).

## Список литературы

- [1] Куомо С., ди Кола В.С., Джампаоло Ф., Розза Г., Раисси М., Пикчиалли, Ф.: Научное машинное обучение с помощью нейронных сетей, основанных на физике: где мы находимся и что дальше. arXiv (2022)
- [2] Карниадакис, Г.Е., Кеврекидис, И.Г., Лу, Л., Пердикарис, П., Ван, С., Янг, Л.: Машинное обучение на основе физики. Обзоры Nature Physics 3(6), 422-440 (2021). <https://doi.org/10.1038/s42254-021-00314-5>
- [3] Раисси, М., Пердикарис, П., Карниадакис, Г.Э.: Нейронная система, основанная на физике. сети: платформа глубокого обучения для решения прямых и обратных задач, связанных с нелинейными дифференциальными уравнениями в частных производных. Журнал вычислительной физики 378, 686-707 (2019). <https://doi.org/10.1016/j.jcp.2018.10.045>
- [4] Мартин Дж., Шауб Х.: Нейронные сети на основе физики для гравитационного поля моделирование Земли и Луны. Небесная механика и динамика астрономия 134(2) (2022). <https://doi.org/10.1007/s10569-022-10069-5>

- [5] Мартин Дж., Шауб Х.: Основанные на физике нейронные сети для гравитации моделирование поля малых тел. Небесная механика и динамическая астрономия, 28 (2022) [6] Бриллюэн, М.: уравнения с производными частями второго порядка. Domaines множественная связь. Функции сфер без антиподов 4, 173-206 (1933)
- [7] Каула, В.М. Теория спутниковой геодезии: применение спутников для Геодезия. Издательство Блейсделл, Уолтем, Массачусетс. (1966)
- [8] Павлис, Н.К., Холмс, С.А., Кеньон, С.К., Шмидт, Д., Триммер, Р.: А Предварительная гравитационная модель до степени 2160. В: Гравитация, геоид и космические миссии, том 129, стр. 18-23. Springer-Verlag, Berlin/Heidelberg (2005). [https://doi.org/10.1007/3-540-26932-0\\_4](https://doi.org/10.1007/3-540-26932-0_4)
- [9] Гуссенс С., Лемуан Ф., Сабака Т., Николас Дж., Мазарико Э., Роуландс, Д., Лумис, Б., Чинн, Д., Нейман, Г., Смит, Д., Зубер, М.: Глобальная модель гравитационного поля Луны 1200 степени и порядка с использованием данных миссии GRAIL. В: 47-я ежегодная лунная и планетарная научная конференция. Конференция по лунным и планетарным наукам, стр. 1484- 1484 (2016)
- [10] Дженова А., Гуссенс С., Лемуан Ф.Г., Мазарико Э., Нейман Г.А., Смит, Д.Э.Н., Зубер, М.Т.: Сезонное и статическое гравитационное поле Марса от MGS, Mars Odyssey и MRO radio science. Икар 272, 228-245 (2016). <https://doi.org/10.1016/j.icarus.2016.02.050>
- [11] Гравитация, Дж.: Гравитационное поле Земли 2.1, том 1, стр. 64 (2016). <https://doi.org/10.1002/14651858.CD000031.pub4.www.cochranelibrary.com>
- [12] Хьюитт Э., Хьюитт Р.Э.: Феномен Гиббса-Уилбрэхема: эпизод в анализе Фурье. Архив истории точных наук 21 (2), 129-160 (1979). <https://doi.org/10.1007/BF00330404>
- [13] Мартин, Дж.Р., Шауб, Х.: Реализация GPGPU сферической модели Пайнса Гармоническая гравитационная модель. В: Уилсон, Р.С., Шан, Дж., Хауэлл, К.К., Хутс, Ф.Р. (ред.) Конференция специалистов по астеродинамике AAS / AIAA. Univelt Inc., Виртуальное событие (2020)
- [14] Вернер Р., Ширс Д.: Внешняя гравитация многогранника, полученного и по сравнению с представлениями о гармонической и масконной гравитации астероида 4769 Касталия. Небесная механика и динамическая астрономия 65(3), 313-344 (1997). <https://doi.org/10.1007/BF00053511>
- [15] Ширс Д.Дж., Хушалани Б., Вернер Р.А.: оценка плотности астероидов распределения по форме и гравитационной информации. Планетарный и космический

Наука 48(10), 965-971 (2000). [https://doi.org/10.1016/s0032-0633\(00\)](https://doi.org/10.1016/s0032-0633(00)00064-7)

00064-7

- [16] Ширс, Д.Дж., Френч, А.С., Трикарико, П., Чесли, С.Р., Такахаси, Ю., Фарноккья, Д., Макмахон, Дж.У., Брэк, Д.Н., Дэвис, А.Б., Баллуз, Р.Л., Джавин, Э.Р., Розитис, Б., Эмери, Дж.П., Райан, А.Дж., Парк, Р.С., Раш, Б.П., Мastroдемос, Н., Кеннеди, Б.М., Беллероуз, Дж., Люби, Д.П., Велес, Д., Вон, А.Т., Леонард, Дж.М., Гирарт, Дж., Пейдж, Б., Антреазян, П., Мазарико, Э., Гетцанданнер, К., Роулэндс, Д., Моро, М.К., Смолл, Дж., Хайсмит, Д.Э.Н., Гуссенс, С., Палмер, Э.Э., Вейрих, Дж.Р., Гаскелл, Р.В., Барнуин, О.С., Дейли, М.Г., Сибрук, Дж.А., Аль-Асад, М.М., Филпотт, Л.К., Джонсон, К.Л., Харцелл, К.М., Гамильтон, В.Е., Мишель, П., Уолш, К. Дж., Нолан, М.К., Лоретта, Д.С.: Неоднородное распределение массы груды обломков астероида (101955) Бенну. Достижения науки 6(41) (2020). <https://doi.org/10.1126/sciadv.abc3350>
- [17] Зубер, М.Т., Смит, Д.Э.Н., Ченг, А.Ф., Гарвин, Дж.Б., Ааронсон, О., Коул, Т.Д., Данн, П.Дж., Го, Ю., Лемуан, Ф.Г., Нейман, Г.А., Роулэндс, Д.Д., Торренс, М.Х.: Форма 433 Эроса с помощью лазерного дальномера БЛИЖНЕГО действия Шумейкера. Наука 289(5487), 2097-2101 (2000). <https://doi.org/10.1126/science.289.5487.2097>
- [18] Шепард, М.К., Тимерсон, Б., Ширс, Д.Дж., Беннер, Л.А.М., Джорджини, Дж.Д., Хауэлл Э.С., Магри К., Нолан М.К., Спрингманн А., Тейлор П.А., Виркки А.: Пересмотренная модель формы астероида (216) Клеопатра. Икар 311, 197-209 (2018). <https://doi.org/10.1016/j.icarus.2018.04.002>
- [19] Миллер Дж.К., Коноплив А.С., Антреазян П.Г., Борди Дж.Дж., Чесли С., Хелфрич, К.Э.Н., Оуэн, У.М., Ван, Т.К., Уильямс, Б.Г., Йоменс, Д.К., Ширс, Д.Дж.: Определение формы, силы тяжести и состояния вращения астероида 433 Эрос. Икар 155(1), 3-17 (2002). <https://doi.org/10.1006/икар.2001.6753>
- [20] Ромен Г., Жан-Пьер Б.: Эллипсоидальные гармонические разложения гравитационный потенциал: теория и применение. Небесная механика и динамическая астрономия 79(4), 235-275 (2001). <https://doi.org/10.1023/A:1017555515763>
- [21] Такахаси Ю., Ширс Д.Дж., Вернер Р.А.: Поверхностные гравитационные поля для астероидов и комет. Журнал руководства, контроля и динамики 36 (2), 362-374 (2013). <https://doi.org/10.2514/1.59144>
- [22] Такахаси Ю., Ширс Д.Дж.: Поля тяготения на поверхности малых тел посредством сферических гармонических разложений. Небесная механика и динамическая астрономия 119(2), 169-206 (2014). <https://doi.org/10.1007/s10569-014-9552-9>
- [23] Мюллер, А.П.М., Шегрен, В.Л.: Масконы: концентрация лунной массы

161(3842), 680-684 (1968)

[24] Тардивел, С.: Пределы приближения Маскона к однородному многограннику-  
neous. В: Конференция специалистов по аэродинамике AIAA / AAS,  
стр. 1-13. Американский институт аэронавтики и астронавтики, Рестон,  
Вирджиния (2016). <https://doi.org/10.2514/6.2016-5261>

[25] Виттик П.Т., Рассел Р.П.: Представления силы тяжести в смешанной модели для  
малых небесных тел с использованием масконов и сферических гармоник. Космическая  
механика и динамическая астрономия 131(7), 31-31 (2019). <https://doi.org/10.1007/s10569-019-9904-6>

[26] Гао А., Ляо В.: Эффективный метод моделирования гравитационного поля для малых тел  
основанный на регрессии гауссовского процесса. Acta Astronautica 157 (декабрь  
2018), 73-91 (2019). <https://doi.org/10.1016/j.actaastro.2018.12.020>

[27] Ченг Л., Ван З., Цзян Ф.: Контроль в режиме реального времени для оптимальной подачи топлива  
Посадка на Луну на основе интерактивного алгоритма глубокого обучения с  
подкреплением. Аэродинамика 3(4), 375-386 (2019). <https://doi.org/10.1007/s42064-018-0052-2>

[28] Фурфаро Р., Барокко Р., Линарес Р., Топпуто Ф., Редди В., Симо,  
Дж., Ле Корре, Л.: Моделирование гравитационного поля нерегулярных малых тел с помощью  
машин экстремального обучения и байесовской оптимизации. Достижения в космических  
исследованиях (июнь) (2020). <https://doi.org/10.1016/j.asr.2020.06.021>

[29] Иззо Д., Гомес П.: Геодезия нерегулярных малых тел с помощью нейронной плотности  
поля. Инженерия связи 1(1), 48 (2022). <https://doi.org/10.1038/s44172-022-00050-3>

[30] Хуан, Г.-Б., Чжу, К.-Ю., Сю, К.-К.: Машина экстремального обучения: теория  
и приложения. Нейрокомпьютинг 70(1-3), 489-501 (2006). <https://doi.org/10.1016/j.neucom.2005.12.126>

[31] Ченг Л., Ван З., Сонг Ю., Цзян Ф.: Оптимальное управление в реальном времени для  
нерегулярных посадок на астероиды с использованием глубоких нейронных  
сетей. Acta Astronautica 170 (январь 2019), 66-79 (2020). <https://doi.org/10.1016/j.actaastro.2019.11.039>

[32] Кингма, Д.П., Б.А., Дж.: Адам: метод стохастической оптимизации.  
3-я Международная конференция по учебным представлениям, ICLR 2015 -  
Материалы конференции, 1-15 (2014)

[33] Дозат, Т.: Включение импульса Нестерова в Адама

[34] Ночедал, Дж., Райт, С.Дж.: Численная оптимизация, 2-е изд. изд. Спрингер  
Серия исследований операций. Спрингер, Нью-Йорк (2006)



- [35] Милденхолл Б., Шринивасан П.П., Танчик М., Бэррон Дж. Т., Рамамурти, R., Ng, R.: NeRF: Представление сцен в виде полей нейронного излучения для синтеза зрения. Коммуникации ACM 65(1), 99-106 (2022). <https://doi.org/10.1145/3503250>
- [36] Байдин А.Г., Перлмуттер Б.А., Сискинд Дж.М.: Автоматическое дифференцирование в машинном обучении: обзор. Журнал исследований машинного обучения 18, 1-43 (2018) [37] Мартин Дж.Р., Шауб Х.: предварительный анализ силы тяжести малого тела Оценка поля с использованием нейронных сетей, основанных на физике, и фильтров Калмана . международный астронавтический конгресс, 10 (2022)
- [38] Гуденору И., Бенджио Ю., Курвиль А.: Глубокое обучение. MIT Press, ???
- [39] Пайнс, С.: Единообразное представление гравитационного потенциала и его Производные. Журнал AIAA 11(11), 1508-1511 (1973). <https://doi.org/10.2514/3.50619>
- [40] Чжу К., Лю З., Ян Дж.: Машинное обучение для производства присадок к металлам факторинг: прогнозирование температуры и динамики жидкости в бассейне расплава с использованием нейронных сетей, основанных на физике. arXiv (2020) [41] Такахаси, Ю., Ширс, Д.Дж.: Распределение плотности, основанное на морфологии оценка для малых тел. Icarus 233, 179-193 (2014). <https://doi.org/10.1016/j.icarus.2014.02.004>
- [42] КАНАМАРУ М., САСАКИ С.: Оценка распределения плотности внутри- Японского общества авиационных и космических наук, аэрокосмических технологий Япония 17(3), 270-275 (2019). <https://doi.org/10.2322/tastj.17.270>
- [43] Хендрикс Д., Гимпел К.: Линейные единицы погрешности по Гауссу (GELUs). arXiv (2020)

## Подробности обучения PINN-GM-III

Как и PINN-GM-II, PINN-GM-III построен на основе базовой сети, которая имеет сходство с архитектурой трансформатора. Сама сеть состоит из многослойного восприятия с прямой связью, которому предшествует слой встраивания. Пропускные соединения подключаются между уровнем встраивания и каждым из скрытых слоев, а последний уровень использует функции линейной активации для получения сетевого прогноза потенциала прокси.

В отличие от PINN-GM-II, все эксперименты, протестированные в этой статье, не используют функцию потерь с несколькими ограничениями. В частности, в PINN-

При внедрении GM-II функция потерь включала штрафы за

$$\nabla^2 U = 0 \text{ и}$$

$\nabla \times \nabla \nabla U = 0$  в дополнение к  $-\nabla U$  - = а 0. Хотя функция многоцелевых потерь в прошлом приносила пользу при обучении сетей по зашумленным данным, она показано, что в исследованиях абляции, показанных в Приложении, больше нет необходимости. [В](#). Гиперпараметры по умолчанию, используемые для обучения PINN-GM-III, перечислены в таблице [4](#).

Гиперпараметр	Значение	Гиперпараметр	Значение
скорость обучения	$\sim 2$	lr планировщик	плато
размер пакета	$2^{11}$	lr терпение скорость	1,500
количество эпох	$2^{13}$	затухания	0.5
оптимизатор	Адам [32]	минимальная	0.001
функция активации	GELU [43]	дельта мин lr	$1e-6$

Таблица 4: Набор гиперпараметров по умолчанию для обучения нейронной сети

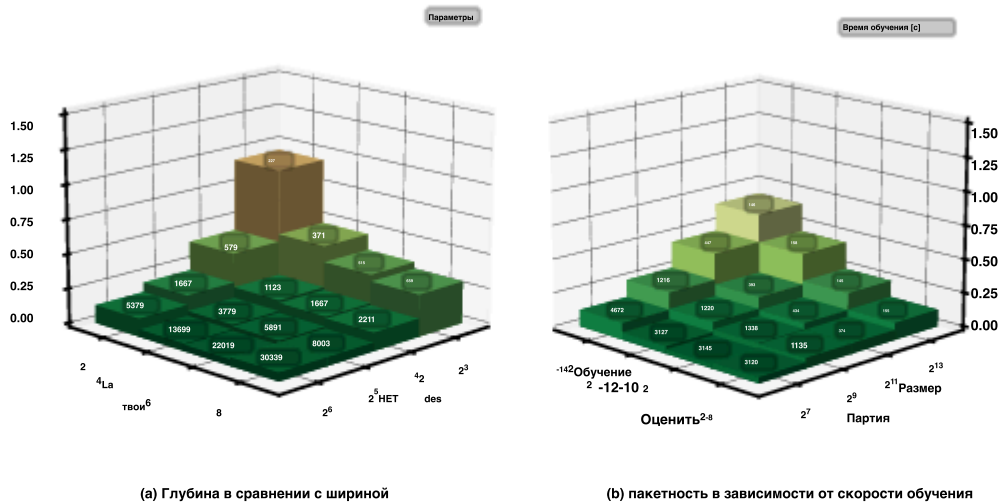
Исследования В-абляции

Нейронные сети могут быть чувствительны к различным гиперпараметрам, поэтому этот раздел призван охарактеризовать чувствительность PINN-GM к набору основных гиперпараметров , а также определить влияние различных размеров сети и количества / качества обучающих данных. Очевидно, что глубина сети, ширина, размер пакета, скорость обучения, эпохи и функция потерь различаются, так же как общий объем данных и их качество. Как и прежде, неоднородная плотность астероида Эрос подробно описана в разделе [5](#) используется для предоставления обучающих данных. 32 768 обучающих данных создаются равномерно в диапазоне от 0 до 3R, а производительность сети оценивается с использованием среднепроцентной ошибки, усредненной по 4096 отдельным проверочным выборкам.

В.1 Размер сети

В первом тестировании исследуется чувствительность PINN-GM к размеру сети, варьирующемуся как по глубине, так и по ширине сети. Глубина сети варьируется между 2, 4, 6, 8 скрытыми слоями, а ширина варьируется между 8, 16, 32, 64 узлами на скрытый слой. Остальные гиперпараметры остаются неизменными по умолчанию, приведенным в таблице [4](#). Средняя процентная погрешность моделей приведена на рисунке [12a](#) с соответствующим размером модели или общими параметрами, пригодными для обучения, наложенными на столбики.

Рисунок [12a](#) демонстрирует, что PINN-GM-III хорошо работает в сети разного размера. Самые маленькие сети будут работать хуже, чем самые крупные сети, учитывая их значительно меньшую параметрическую емкость, однако, несмотря на это, погрешность всех моделей остается значительно ниже 1,5%. Самая маленькая модель из 227 обучаемых параметров имеет в среднем ошибку 0,9%, тогда как самая большая модель с 30 339 параметрами достигает ошибки 0,3%. Этот рисунок иллюстрирует, что для оптимальных моделей требуется минимум 16 узлов с четырьмя скрытыми слоями. Дополнительные уровни и узлы продолжают незначительно повышать производительность, но увеличат размер модели. Таким образом, наступает момент, когда



Что касается скорости обучения, то чем больше, тем лучше. Вероятно, это побочный продукт выбранного планировщика скорости обучения. Как правило, высокая скорость обучения может препятствовать тому, чтобы сети сходились к своему локальному минимуму; однако планировщики скорости обучения снижают этот риск, постоянно снижая скорость обучения, чтобы решение могло достичь своего локального минимума. Для этих экспериментов скорость обучения снижается в 0,5 раза каждые 1500 эпох, что не приводит к снижению достоверности . Учитывая эту конфигурацию планировщика скорости обучения., в скорости обучения не менее  $2^{-8}$  рекомендуется для большинства применений.

### Б. 3 количество данных и эпох

Из этих двух начальных экспериментов выбирается окончательный набор оптимизированных гиперпараметров и архитектуры моделей для дальнейших экспериментов. В частности, скорость обучения  $2^{-8}$  и размер пакета  $2^{11}$  назначаются для двух размеров ушек- GM-III: ушек малого и УШЕК большого размера. PINN-Small использует 2-уровневую конфигурацию сети с 8 узлами на каждый уровень (227 параметров), тогда как PINN-Large использует 8-уровневую конфигурацию с 64 узлами (30 339 параметров). Используя эти две сети с фиксированными гиперпараметрами, проводится дополнительный эксперимент, в ходе которого исследуется чувствительность этих моделей к количеству обучающих данных и продолжительности обучения. Очевидно, что как PINN-Small, так и PINN-Large обучаются с увеличением объемов обучающих данных в диапазоне от  $\{2^9, 2^{11}, 2^{13}, 2^{15}\}$  выборки и увеличения общего количества периодов обучения  $\{2^9, 2^{11}, 2^{13}, 2^{15}\}$ . Как и прежде, средняя процентная ошибка оценивается и представлена на основе продолжительности обучения на рисунках 13а и 13б.

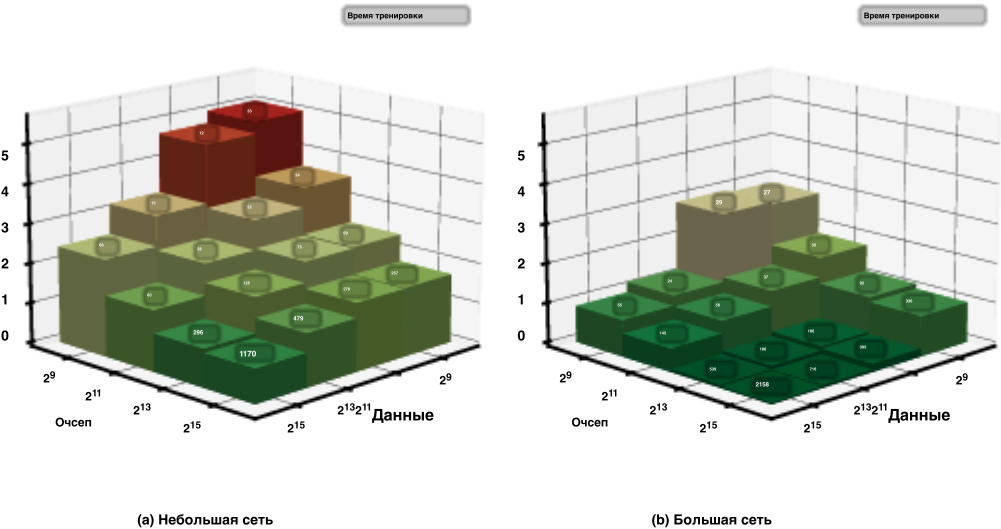


Рис. 13: Данные в сравнении с эпохами

Рисунки 13а демонстрирует, что, несмотря на небольшие размеры, PINN- Small способен достигать ошибок всего в 1% при наличии достаточных данных для обучения

и время обучения. Даже в условиях наибольшего дефицита данных модель остается способной сходиться к решениям с погрешностью более 5% и улучшаться до значений менее 1% при наличии достаточного времени обучения и данных. Аналогичным образом, рисунок 13b демонстрирует, что PINN-Large неизменно выигрывает от дополнительных обучающих данных и эпох, при этом погрешность наиболее точных моделей достигает 0,04%. В отличие от моделей PINN-Small, модели PINN-Large всегда оставались с погрешностью ниже 3% даже в режимах с низким объемом данных и эпохой. Взяты вместе, результаты больших и малых сетей предполагают, что для оптимальной производительности PINN-GM-III следует обучать как минимум 8192 эпохи, но все равно можно использовать больше, если позволяют вычислительные ресурсы.

## V.4 Качество данных и физические ограничения

В заключительном эксперименте исследуется, как качество обучающих данных и физические ограничения влияют на производительность. Этот эксперимент начинается с того, что берется каждый вектор ускорения и добавляется 10% от их величины в некотором случайном направлении к вектору истины. Это негауссово распределение предназначено для того, чтобы подчеркнуть эти модели и определить, насколько хорошо они сходятся к истинному решению, несмотря на ошибку в данных. Эксперимент продолжается путем обучения как PINN-Small, так и PINN-Large на увеличивающихся объемах обучающих данных в диапазоне  $(1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536, 131072, 262144, 524288, 1048576, 2097152, 4194304, 8388608, 16777216, 33554432, 67108864, 134217728, 268435456, 536870912, 1073741824, 2147483648, 4294967296, 8589934592, 17179869184, 34359738368, 68719476736, 137438953472, 274877906944, 549755813888, 1099511627776, 2199023255552, 4398046511104, 8796093022208, 17592186044416, 35184372088832, 70368744177664, 140737488355328, 281474976710656, 562949953421312, 1125899906842624, 2251799813685248, 4503599627370496, 9007199254740992, 18014398509481984, 36028797018963968, 72057594037927936, 144115188075855872, 288230376151711744, 576460752303423488, 1152921504606846976, 2305843009213693952, 4611686018427387904, 9223372036854775808, 18446744073709551616, 36893488147419103232, 73786976294838206464, 147573952589676412928, 295147905179352825856, 590295810358705651712, 1180591620717411303424, 2361183241434822606848, 4722366482869645213696, 9444732965739290427392, 18889465931478580854784, 37778931862957161709568, 75557863725914323419136, 151115727451828646838272, 302231454903657293676544, 604462909807314587353088, 1208925819614629174706176, 2417851639229258349412352, 4835703278458516698824704, 9671406556917033397649408, 19342813113834066795298816, 38685626227668133590597632, 77371252455336267181195264, 154742504910672534362390528, 309485009821345068724781056, 618970019642690137449562112, 1237940039285380274899124224, 2475880078570760549798248448, 4951760157141521099596496896, 9903520314283042199192993792, 19807040628566084398385987584, 39614081257132168796771975168, 79228162514264337593543950336, 158456325028528675187087900672, 316912650057057350374175801344, 633825300114114700748351602688, 1267650600228229401496703205376, 2535301200456458802993406410752, 5070602400912917605986812821504, 10141204801825835211973625643008, 20282409603651670423947251286016, 40564819207303340847894502572032, 81129638414606681695789005144064, 162259276829213363391578010288128, 324518553658426726783156020576256, 649037107316853453566312041152512, 1298074214633706907132624082305024, 2596148429267413814265248164610048, 5192296858534827628530496329220096, 10384593717069655257060992658440192, 20769187434139310514121985316880384, 41538374868278621028243970633760768, 83076749736557242056487941267521536, 166153499473114484112975882535043072, 332306998946228968225951765070086144, 664613997892457936451903530140172288, 1329227995784915872903807060280344576, 2658455991569831745807614120560689152, 5316911983139663491615228241121378304, 10633823966279326983230456482242756608, 21267647932558653966460912964485513216, 42535295865117307932921825928971026432, 85070591730234615865843651857942052864, 170141183460469231731687303715884105728, 340282366920938463463374607431768211456, 680564733841876926926749214863536422912, 1361129467683753853853498429727072845824, 2722258935367507707706996859454145691648, 5444517870735015415413993718908291383296, 10889035741470030830827987437816582766592, 21778071482940061661655974875633165533184, 43556142965880123323311949751266331066368, 87112285931760246646623899502532662132736, 174224571863520493293247799005065324265472, 348449143727040986586495598010130648530944, 696898287454081973172991196020261297061888, 1393796574908163946345982392040522594123776, 2787593149816327892691964784081045188247552, 5575186299632655785383929568162090376495104, 11150372599265311570767859136324180752990208, 22300745198530623141535718272648361505980416, 44601490397061246283071436545296723011960832, 89202980794122492566142873090593446023921664, 178405961588244985132285746181186892047843328, 356811923176489970264571492362373784095686656, 713623846352979940529142984724747568191373312, 1427247692705959881058285969449495136382746624, 2854495385411919762116571938898990272765493248, 5708990770823839524233143877797980545530986496, 11417981541647679048466287755595961091061972992, 22835963083295358096932575511191922182123945984, 45671926166590716193865151022383844364247891968, 91343852333181432387730302044767688728495783936, 182687704666362864775460604089535377456991567872, 365375409332725729550921208179070754913983135744, 730750818665451459101842416358141509827966271488, 1461501637330902918203684832716283019655932542976, 2923003274661805836407369665432566039311865085952, 5846006549323611672814739330865132078623730171904, 11692013098647223345629478661730264157247460343808, 23384026197294446691258957323460528314494920687616, 46768052394588893382517914646921056628989841375232, 93536104789177786765035829293842113257979682750464, 187072209578355573530071658587684226515959365500928, 374144419156711147060143317175368453031918731001856, 748288838313422294120286634350736906063837462003712, 1496577676626844588240573268701473812127674924007424, 2993155353253689176481146537402947624255349848014848, 5986310706507378352962293074805895248510699696029696, 11972621413014756705924586149611790497021399392059392, 23945242826029513411849172299223580994042798784118784, 47890485652059026823698344598447161988085597568237568, 95780971304118053647396689196894323976171195136475136, 191561942608236107294793378393788647952342390272950272, 383123885216472214589586756787577295904684780545900544, 766247770432944429179173513575154591809369561091801088, 1532495540865888858358347027150309183618739122183602176, 3064991081731777716716694054300618367237478244367204352, 6129982163463555433433388108601236734474956488734408704, 12259964326927110866866776217202473468949912977468817408, 24519928653854221733733552434404946937899825954937634816, 49039857307708443467467104868809893875799651909875269632, 98079714615416886934934209737619787751599303819750539264, 196159429230833773869868419475239575503198607639501078528, 392318858461667547739736838950479151006397215279002157056, 784637716923335095479473677900958302012794430558004314112, 1569275433846670190958947355801916604025588861116008628224, 3138550867693340381917894711603833208051177722232017256448, 6277101735386680763835789423207666416102355444464034512896, 12554203470773361527671578846415332832204710888928069025792, 25108406941546723055343157692830665664409421777856138051584, 50216813883093446110686315385661331328818843555712276103168, 100433627766186892221372630771322662657637687111424552206336, 200867255532373784442745261542645325315275374222849104412672, 401734511064747568885490523085290650630550748445698208825344, 803469022129495137770981046170581301261101496891396417650688, 1606938044258990275541962092341162602522202993782792835301376, 3213876088517980551083924184682325205044405987565585670602752, 6427752177035961102167848369364650410088811975131171341205504, 12855504354071922204335696738729300820177623950262342682411008, 25711008708143844408671393477458601640355247900524685364822016, 51422017416287688817342786954917203280710495801049370729644032, 102844034832575377634685573909834406561420991602098741459288064, 205688069665150755269371147819668813122841983204197482918576128, 411376139330301510538742295639337626245683966408394965837152256, 822752278660603021077484591278675252491367932816789931674304512, 1645504557321206042154969182557350504982735865633579863348609024, 3291009114642412084309938365114701009965471731267159726697218048, 6582018229284824168619876730229402019930943462534319453394436096, 13164036458569648337239753460458804039861886925068638906788872192, 26328072917139296674479506920917608079723773850137277813577744384, 52656145834278593348959013841835216159447547700274555627155488768, 105312291668557186697918027683670432318895095400549111254310977536, 210624583337114373395836055367340864637790190801098222508621955072, 421249166674228746791672110734681729275580381602196445017243910144, 842498333348457493583344221469363458551160763204392890034487820288, 1684996666696914987166688442938726917102321526408785780068975640576, 3369993333393829974333376885877453834204643052817571560137951281152, 6739986666787659948666753771754907668409286105635143120275902562304, 13479973333575319897333507543509815336818572211270286240551805124608, 26959946667150639794667015087019630673637144422540572481103610249216, 53919893334301279589334030174039261347274288845081144962207220498432, 107839786668602559178668060348078522694548577690162289924414440996864, 215679573337205118357336120696157045389097155380324579848828881993728, 431359146674410236714672241392314090778194310760649159697657763987456, 862718293348820473429344482784628181556388621521298319395315527974912, 1725436586697640946858688965569256363112777243042596638790631055949824, 3450873173395281893717377931138512726225554486085193277581262111899648, 6901746346790563787434755862277025452451108972170386555162524223799296, 13803492693581127574869511724554050904902217944340773110325048447598592, 27606985387162255149739023449108101809804435888681546220650096895197184, 55213970774324510299478046898216203619608871777363092441300193790394368, 110427941548649020598956093796432407239217743554726184882600387580788736, 220855883097298041197912187592864814478435487109452369765200775161577472, 441711766194596082395824375185729628956870974218904739530401550323154944, 883423532389192164791648750371459257913741948437809479060803100646309888, 1766847064778384329583297500742918515827483896875618958121606201292619776, 3533694129556768659166595001485837031654967793751237916243212402585239552, 7067388259113537318333190002971674063309935587502475832486424805170479104, 14134776518227074636666380005943348126619871175004951664972849610340958208, 28269553036454149273332760011886696253239742350009903329945699220681916416, 56539106072908298546665520023773392506479484700019806659891398441363832832, 113078212145816597093331040047546785012958969400039613319782796882727665664, 226156424291633194186662080095093570025917938800079226639565593765455331328, 452312848583266388373324160190187140051835877600158453279131187530910662656, 904625697166532776746648320380374280103671755200316906558262375061821325312, 1809251394333065553493296640760748560207343510400633813116524750123642650624, 3618502788666131106986593281521497120414687020801267626233049500247285301248, 7237005577332262213973186563042994240829374041602535252466099000494570602496, 14474011154664524427946373126085988481658748083205070504932198000989141204992, 28948022309329048855892746252171976963317496166410141009864396001978282409984, 57896044618658097711785492504343953926634992332820282019728792003956564819968, 115792089237316195423570985008687907853269984665640564039457584007913129639936, 231584178474632390847141970017375815706539969331281128078915168015826259279872, 463168356949264781694283940034751631413079938662562256157830336031652518559744, 926336713898529563388567880069503262826159877325124512315660672063305037119488, 1852673427797059126777135760139006525652319754650249024631321344126610074238976, 3705346855594118253554271520278013051304639509300498049262642688253220148477952, 7410693711188236507108543040556026102609279018600996098525285376506440296955904, 14821387422376473014217086081112052205218558037201992197050570753012880593911808, 29642774844752946028434172162224104410437116074403984394101141506025761187823616, 59285549689505892056868344324448208820874232148807968788202283012051522375647232, 118571099379011784113736688648896417641748464297615937576404566024103044751294464, 237142198758023568227473377297792835283496928595231875152809132048206089502588928, 474284397516047136454946754595585670566993857190463750305618264096412179005177856, 948568795032094272909893509191171341133987714380927500611236528192824358010355712, 1897137590064188545819787018382342682267975428761855001222473056385648716020711424, 3794275180128377091639574036764685364535950857523710002444946112771297432041422848, 758855036$

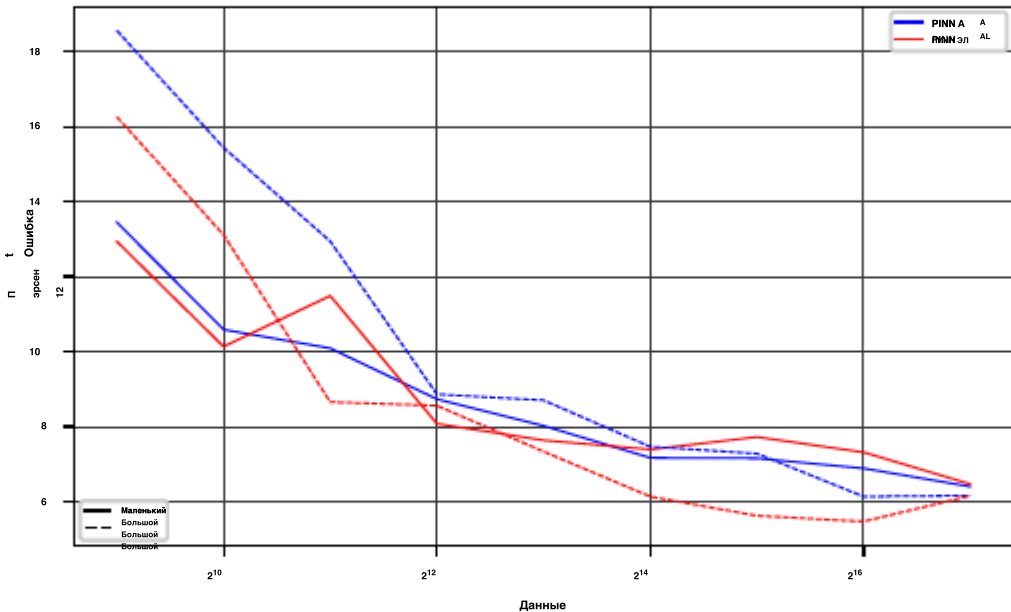


Рис. 14: Ошибка как функция обучающих данных и функции потерь

что касается системы, то модели могут лучше игнорировать части обучающих данных, которые принципиально несовместимы с физикой.

PINN-Large более подвержен ошибкам в режимах с низким уровнем данных, достигая 19% ошибки в случае с 512 точками данных. Вероятно, это проявление переоснащения ограниченных обучающих данных, учитывая значительно более высокую параметрическую емкость модели, чем у модели ПИННА-Смолла. В модели ПИННА-Большого размера дополнительный компонент Лапласа к функции потерь (PINN AL) действительно немного улучшает производительность, помогая упорядочить модель и минимизировать тенденции к ее переобучению; однако в модели малого размера преимущество значительно менее очевидно. Учитывая, что дополнительное ограничение Лапласа-  $\text{c} \cdot \text{lap}$  намного дороже в вычислительном отношении, чем простое ограничение PINN A, обычно рекомендуется просто придерживаться ограничения A, если в настройках с ограничением вычислений.

## Комментарии С о прошлом Машинном обучении

### Производительность

Таблица 1 освещает общую производительность прошлых и нынешних моделей гравитации, основанных на машинном обучении..... . Все значения взяты из соответствующего справочника, но необходим дальнейший контекст, поскольку каждая модель оценивала точность по-разному и на разных астероидах. Цель этого раздела - предоставить соответствующие сведения относительно этих показателей для полноты картины.

Для гравитационной модели гауссовского процесса, представленной в ссылке 26, количество параметров модели явно не указано, но может быть выведено.

Гауссовские процессы определяются их ковариационной матрицей и функцией ядра.

Масштабы ковариационной матрицы в качестве  $O(N)^2$  и максимальное количество использованных точек данных составили  $N = 3\,600$ . Это говорит о том, что минимальное количество параметров, используемых в модели, составляет 12 960 000. Точность для этих моделей также указывается при фиксированных радиусах от центра масс для каждого астероида, а не во всей области, поэтому значения можно считать верхними границами. Более того, показано, что модель расходится на больших высотах, следовательно, недействительна в глобальном масштабе.

Для гравитационной модели экстремальной обучающей машины, представленной в Ссылке 28, размер модели определяется тем фактом, что в ELM имеется 50 000 скрытых узлов. Первые 50 000 параметров представляют собой случайные веса, соединяющие входные данные со скрытым слоем, а вторые 50 000 параметров - это полученные веса от скрытого слоя к выходному слою, что составляет в общей сложности 100 000 параметров модели. Смоделированный астероид - 25143 Итокава. Ошибка в ссылке 28 указывается в абсолютных, а не относительных величинах, однако, используя рисунок 9, можно приблизить, что относительная погрешность варьируется от 1% до 10%.

Для справки 31, сообщается, что гравитационная модель нейронной сети использует 512 узлов на скрытый слой для 6 скрытых слоев ( $512^6 = 1,572\,864$  параметра). В документе сообщается, что было сгенерировано 1 000 000 обучающих данных, которые были разделены в соотношении 8: 2 между данными обучения и тестирования, в общей сложности 800 000 обучающих данных. Исследованный астероид также называется 433-Эрос, и средняя относительная ошибка набора тестов указана в таблице 3 как 0,35%.

Для геодезических сетей [29], используются сирены из 9 скрытых слоев по 100 узлов в каждом ( $9 \cdot 100^2 = 90,000$ ). Изучены четыре астероида: Бенну, Чурюмова-Герасименко, Эрос и Итокава. В их дополнительных материалах (таблица S4), сообщается об относительной погрешности в отношении Эроса на трех характерных высотах. На самой низкой высоте средняя ошибка составляет 0,571%, а на самой высокой - 0,146%. В попытках количественно оценить погрешность по всему режиму больших и малых высот эти значения усредняются для заявленного значения 0,359%.

Наконец, для PINN-GM-III указана средняя ошибка для сети, обученной с 8 скрытыми слоями по 16 узлов на слой. Модель была обучена на 4096 точках данных, распределенных между 0-10R, с использованием тех же гиперпараметров, что указаны в Приложении B, и достигнута средняя ошибка в наборе проверки 0,3%.

## Подробности регрессии D

В разделе 5, несколько моделей гравитации регрессируют и тестируются на различных условиях данных. Процедура регрессии для нейронных сетей согласуется с предыдущими объяснениями; однако остальные модели требуют дополнительного контекста. Цель этого раздела - представить этот контекст, объяснив, как каждая из оставшихся гравитационных моделей регрессирует по данным, обсуждаемым в разделе 5.

## Сферические гармоники D.1

Сферическая гармоническая гравитационная модель регрессируется путем решения линейной системы для их коэффициентов Стокса. Эта система имеет следующий вид:

$$\mathbf{a} = \mathbf{H} \mathbf{c} \quad (24)$$

$$\mathbf{C}^T \mathbf{a} = \mathbf{H}^T \mathbf{H} \mathbf{c} \quad (25)$$

$$(\mathbf{C}^T \mathbf{C}) \mathbf{C}^T \mathbf{a} = \mathbf{c} \quad (26)$$

где  $\mathbf{c}$  - вектор коэффициентов сферической гармоники  $C_{lm}$

are и  $S_{lm}$

вектор ускорений,  $\mathbf{a}$  - якобиан

$\frac{\partial \mathbf{a}}{\partial \mathbf{c}}$  и

$\mathbf{r}$  - это позиция

вектор для каждой точки поля. Прямые решения методом наименьших квадратов могут быть подходящими для небольших систем, но по мере роста размера сферической гармонической модели и увеличения объема данных этот процесс быстро становится неосуществимым с вычислительной точки зрения. Более того, модели сферических гармоник высокой степени создают численные нестабильности, поскольку наблюдаемость гармоник быстро уменьшается при  $(R/r)^l$ , часто делая  $\mathbf{H}^T \mathbf{H}$  плохо обусловленные.

Существует множество стратегий для смягчения этих проблем. Одним из наиболее популярных является использование правила Каулы, формы гребневой регрессии, которая может использоваться для регуляризации коэффициентов сферической гармоники и оказания помощи в плохо обусловленном характере регрессии высокой степени. Явно, вместо того, чтобы стремиться минимизировать среднеквадратичное значение ошибки решения

$$L(\mathbf{c}) = \|\mathbf{a} - \mathbf{H} \mathbf{c}\|_2^2 \quad (27)$$

вместо этого добавьте штрафной член, чтобы обеспечить уменьшение коэффициентов по величине для гармоник более высокой степени через

$$L(\mathbf{c}) = \|\mathbf{a} - \mathbf{H} \mathbf{c}\|_2^2 + \lambda \|\mathbf{c}\|_2^2 \quad (28)$$

где  $\Gamma$  - матрица регуляризации, определенная с помощью правила Каулы

$$\Gamma_{ll} = \frac{\alpha}{l(l+1)} \quad (29)$$

где  $\alpha$  - указанная пользователем константа, обычно выбираемая с помощью перекрестной проверки. Соответствующее решение для гребневой регрессии тогда становится

$$(\mathbf{H}^T \mathbf{H} + \mathbf{\Gamma})^{-1} \mathbf{H}^T \mathbf{a} = \mathbf{c} \quad (30)$$

что дает решение со все более малыми коэффициентами сферической гармоники при высокой степени. Это одновременно устраняет некорректный характер исходной регрессии, а также обеспечивает основу для поддержки регрессии с низким объемом данных. Хотя использование правила Каулы смягчает некорректность исходной регрессии, инвертирование остается дорогостоящим с точки зрения вычислений  $(\mathbf{H}^T \mathbf{H} + \mathbf{\Gamma})$  для больших



наборы данных. Это решается с помощью двух стратегий: 1) рекурсивного метода наименьших квадратов и 2) итеративной регрессии коэффициентов. Рекурсивный метод наименьших квадратов последовательно передает небольшие партии данных для поддержания удобства вычислений, в то время как итеративная регрессия коэффициентов регрессирует гармоники низкой степени отдельно от гармоник высокой степени.

Рекурсивный метод наименьших квадратов выполняется партиями по 100 пар положение / ускорение с использованием следующих рекурсивных соотношений

$$K_{k+1}^{-1} = K_k^{-1} - \frac{K_k^{-1} H_{k+1}^T (I + H_k K_k^{-1} H_k^T)^{-1} H_{k+1} K_k^{-1}}{1 + H_{k+1} K_k^{-1} H_{k+1}^T} \quad (31)$$

$$c_{k+1} = c_k + K_{k+1}^{-1} (H_{k+1}^T c_k - H_{k+1} K_{k+1}^{-1} H_{k+1}^T c_k) \quad (32)$$

где  $K$  выполняется только один раз, и итеративная регрессия коэффициента равна 5000 коэффициентов за раз.

Последнее замечание, связанное с регрессией модели сферической гармоники для сравнительного исследования, заключается в том, что образцы под поверхностью Бриллюэна  $r < R$  намеренно исключены из регрессии из-за  $(R/r)^l$  масштабирование в гармонической модели. Для моделей с очень высокой степенью этот член расходится при оценке на выборках ниже Бриллюэна, что нарушает регрессию.

## D.2 Масконы

Как и сферические гармоники, регрессия Маскона также является формой регрессии методом наименьших квадратов. В частности, набор из  $N$  масконов равномерно распределен по объему модели формы, и их соответствующие массы регрессируют. В идеале регрессия маскона также должна включать нелинейное ограничение, применяемое ко всем элементам массы, которое гарантирует, что они всегда будут иметь значение больше нуля для обеспечения соответствия физике - т. Е. Не существует таких вещей, как отрицательные массы. Однако такое физическое соответствие часто приводит к значительной потере точности моделирования. Поэтому раздел 5 позволяет регрессировать отрицательные массы, чтобы дать модели наиболее сильные аргументы против PINN-GM.

В дополнение к допуску отрицательных масс, регрессия маскона также требует итеративной подгонки, когда размеры модели велики и имеется много данных. Для этого масконы распределяются партиями по 500 штук с использованием всех доступных данных. После подгонки каждой партии их вклад в векторы ускорения удаляется, следующая партия случайным образом распределяется по объему, и модель подгоняется к остаткам ускорения. Этот процесс повторяется до тех пор, пока не будут достигнуты элементы общей массы.

## D.3 Машины для экстремального обучения.

Регрессия экстремальных обучающих машин также очень похожа на регрессию сферической гармоники в том смысле, что она также использует гребневую регрессию и рекурсивную оптимизацию по методу наименьших квадратов. .... Очевидно, что модели ELM регрессируют путем применения матрицы регрессии с фиксированным гребнем  $\Gamma_{ELM}$   $= \alpha_2$  и разбиение набора данных на пакеты, которые будут применяться рекурсивно с использованием эквалайзеров. 32. В частности, оно также опирается на РАН-дом нелинейной проекции в пространстве большей размерности до линейного

## 42 Основанная на физике модель гравитации нейросети поколения III

регрессия, подробно описанная в [28]. Для сравнительного исследования в этой статье, все входные и выходные данные предварительно обрабатываются с использованием преобразования min-max в  $[0, 1]$  и применения сигмоидной функции активации в скрытых узлах.