

<https://doi.org/10.55764/2957-9856/2025-1-31-40.4>

МРНТИ 52.01.11

УДК 551.579

Н. Т. Кемербаев<sup>1</sup>, К. Б. Самарханов<sup>\*2</sup>, Г. А. Мусина<sup>3</sup>, М. Х. Шкиева<sup>4</sup>

<sup>1</sup> К. т. н., генеральный директор (ТОО «GeoID», Астана, Казахстан; [n.kemberbaeyev@geo-id.kz](mailto:n.kemberbaeyev@geo-id.kz))

<sup>2\*</sup> К. г. н., заместитель директора по R&D, главный научный сотрудник (ТОО «GeoID», Международный научный комплекс «Астана», Астана, Казахстан; [kanat.baurzhanuly@gmail.com](mailto:kanat.baurzhanuly@gmail.com))

<sup>3</sup> Начальник отдела камеральной обработки (ТОО «GeoID», Астана, Казахстан; [g.mussina@geo-id.kz](mailto:g.mussina@geo-id.kz))

<sup>4</sup> Заместитель руководителя отдела обработки пространственных данных и ДЗЗ (ТОО «GeoID», Астана, Казахстан; [maral.shkiyeva@gmail.com](mailto:maral.shkiyeva@gmail.com))

## ОПЫТ УЛУЧШЕНИЯ ПРОСТРАНСТВЕННОГО РАЗРЕШЕНИЯ ДАННЫХ СПУТНИКОВОЙ ГРАВИМЕТРИИ GRACE ДЛЯ БАССЕЙНА РЕКИ ЖАЙЫК (УРАЛ) С ПРИМЕНЕНИЕМ МАШИННОГО ОБУЧЕНИЯ

**Аннотация.** Представлено исследование по использованию машинного обучения для улучшения пространственного разрешения спутниковых гравиметрических данных GRACE для бассейна реки Жайык (Урал). Работа направлена на решение проблемы низкого пространственного разрешения данных GRACE, что затрудняет детальный анализ водных ресурсов на региональном уровне. Методология исследования использует статистический метод уменьшения масштаба с использованием модели Random Forest. Для улучшения пространственного разрешения использовались дополнительные наборы данных, включающие цифровую модель рельефа, индексы растительности (NDVI) и индексы водной поверхности (NDWI). Результаты исследования продемонстрировали эффективность предлагаемой методологии: улучшенные данные обеспечивают более точное распределение эквивалентной толщины водного слоя при сохранении общих статистических свойств. Коэффициент корреляции Пирсона составил 0,9424, что означает существенную степень надежности результатов. Практическая значимость этого исследования заключается в перспективном применении разработанного подхода для улучшения оценки водных ресурсов, прогнозирования гидрологических событий и разработки планов по адаптации к изменению климата в бассейне реки Жайык.

**Ключевые слова:** гравиметрия, GRACE, Казахстан, машинное обучение, гидрология, цифровая модель рельефа, пространственное разрешение.

**Введение.** Спутниковая миссия (GRACE) преобразила дистанционное зондирование гравитационного поля Земли, предложив исключительные возможности для мониторинга изменений в мировых водных ресурсах [1]. Однако ограниченное пространственное разрешение данных GRACE (55-111 км) уменьшает их полезность для региональных и локальных исследований. Это особенно актуально для бассейна реки Жайык (Урал), имеющего важное экологическое и экономическое значения для Республики Казахстан и Российской Федерации [2]. Жайык является жизненно важным источником питьевой воды, поддерживающим сельскохозяйственный и промышленный потенциал и баланс экосистемы на прилегающих территориях. В последние десятилетия наблюдаются изменение климата, увеличение антропогенной нагрузки, недостаток водных ресурсов, что в свою очередь требует развития современных методов мониторинга и управления водными ресурсами. Наблюдаемые колебания климата и неравномерность его изменения затрудняют прогнозирование водности реки и управление водными последствиями [3, 4].

Глобальные климатические изменения вносят коррективы в проведение научно-исследовательской деятельности и постоянного контроля за получением актуальной информации в сфере экологии, состояния водных ресурсов и влияния климата на сельскохозяйственное производство в стране. Эксплуатация водных ресурсов, изменение ландшафтов – все это требует использования современных методов исследований с применением спутниковых данных [5].

В связи с участившимися паводками в Республике Казахстан в весенний период возникает необходимость в проведении исследований в области прогнозирования таких ситуаций [6]. Для них необходима актуальная гидрологическая информация, которую можно получить с помощью

спутниковых платформ GRACE-FO [7]. Данные GRACE показывают результаты измерений гравитации в связи с месячными изменениями массы. В свою очередь, измеряемая масса – это изменения гравитации, вызванные изменениями объема воды в водохранилищах, морях и океанах, атмосферной влаги и ледников, состояние обмена между ними. Это облегчает возможность получения и обработки информации. Вертикальные пределы водной массы измеряются в сантиметрах, которые называют эквивалентной толщиной водного слоя.

Машинное обучение (Machine learning, ML) является перспективным инструментом для повышения пространственного разрешения данных дистанционного зондирования через алгоритмы, которые могут анализировать большие объемы данных и выявлять скрытые связи между различными переменными. Для этого используется метод улучшения пространственного разрешения (downscaling), который преобразует выходную информацию глобальной модели с низким разрешением в модель с высоким пространственным разрешением. К основным методам уменьшения пространственного масштабирования относятся метод динамического уменьшения масштаба и метод статистического уменьшения масштаба [8]. В отличие от метода динамического масштабирования метод статистического масштабирования позволяет гибко строить модели и вводить коэффициенты моделирования с высоким разрешением, что значительно улучшает пространственное разрешение гидрологических переменных [9]. В этой статье применен нелинейный статистический метод уменьшения масштаба – уменьшение масштаба на основе модели Random Forest. Применение машинного обучения для исследований на территории Казахстана описано в работе [10], метода улучшения пространственного разрешения – в [11, 12], модели Random Forest – в [13]. Применение алгоритмов улучшения пространственного разрешения предполагает использование различных независимых переменных, включая топографию [14].

Цель этого исследования – разработка метода масштабирования GRACE с использованием машинного обучения и вспомогательных данных дистанционного зондирования Земли. Основные задачи включают этапы формирования обучающего набора на основе данных Sentinel-2, FABDEM и GRACE, а также обучения модели Random Forest для прогнозирования эквивалентной толщины слоя воды (LWE Thickness). Предлагаемая методика позволит увеличить детализацию данных GRACE для более точного мониторинга изменений водных ресурсов в бассейне р. Жайык. Будущая валидация позволит сравнить результаты с наземными измерениями РГП «Казгидромет» и полевыми исследованиями.

**Материалы и методы.** Ниже приведены данные, использованные в этом исследовании (таблица 1). Набор включает географические данные о характеристиках поверхности, обзор географии бассейна, гидрологических ресурсов и растительного покрова. Набор данных WWF HydroSHEDS используется для определения границ бассейна.

Таблица 1 – Исходные данные, использованные для улучшения пространственного разрешения эквивалентной толщины водного слоя реки Жайык (Урал)

Table 1 – Input data used to improve the spatial resolution of the equivalent water layer thickness of the Zhaiyk River (Ural)

Тип данных	Источник	Год	Разрешение, м	Параметры	Применение
Границы бассейна	WWF HydroSHEDS v1 [15]	2000	Уровень 4	HYBAS_ID 2040067740	Определение области интереса
Данные космической гравиметрии	NASA GRACE MASS GRIDS V03 [16]	2024 (март)	55 660	Толщина эквивалентного слоя воды (LWE)	Основной параметр для улучшения разрешения
Цифровая модель рельефа	FABDEM [17]	2000	30	Высота над уровнем моря	Входной параметр для модели
Космические снимки	Copernicus Sentinel-2 [18]	2023-2024 (март-октябрь)	10	Мультиспектральные данные	Расчет индексов, входные данные для модели
Уклон местности	Производный от FABDEM	–	30	Градус уклона	Входной параметр для модели
Индекс растительности	NDVI, рассчитанный из Sentinel-2 [19]	2023-2024 (март-октябрь)	10	От -1 до 1	Входной параметр для модели
Водные поверхности	NDWI, рассчитанный из Sentinel-2 [20]	2023-2024 (март-октябрь)	10	От -1 до 1	Создание водной маски

Набор данных WWF HydroSHEDS используется для определения границ бассейна. Данные GRACE служат основным источником для получения результата с высоким пространственным разрешением, тогда как дополнительные наборы данных предлагают дополнительные входные данные для моделирования и анализа. Они обеспечивают эквивалентные измерения толщины воды с пространственным разрешением 55 км за март 2024 года.

Извлечение информации об уклоне проводится с использованием цифровых данных высоты высокого разрешения (30 м) из модели FABDEM. Индексы растительности (NDVI) и воды (NDWI) вычислялись по оптическим спутниковым изображениям Sentinel-2 с марта по октябрь 2023–2024 годов с пространственным разрешением 10 м.

Такая интеграция источников данных позволяет проводить комплексный анализ параметров окружающей среды и водных ресурсов в выбранном речном бассейне.

#### Алгоритм улучшения пространственного разрешения данных, полученных с GRACE-FO.

Представлен подход, использующий методы машинного обучения для улучшения пространственного разрешения данных GRACE для бассейна реки Жайык (Урал). Он направлен на устранение недостатков текущих продуктов GRACE и на лучшее понимание пространственно-временной динамики водных ресурсов в исследуемой области.

Улучшение географического разрешения данных GRACE имеет три основных преимущества:

повышение точности оценки водных ресурсов на уровне суббассейна, способность обнаруживать тенденции и региональные аномалии;

повышение эффективности объединения данных GRACE с другими источниками географических данных;

использование методов машинного обучения для решения этой проблемы открывает новые возможности для обработки и интерпретации спутниковых гравиметрических данных.

Для улучшения пространственного разрешения данных, полученных с GRACE-FO, построен алгоритм, включивший предварительную обработку и применение модели Random Forest (рисунок 1).

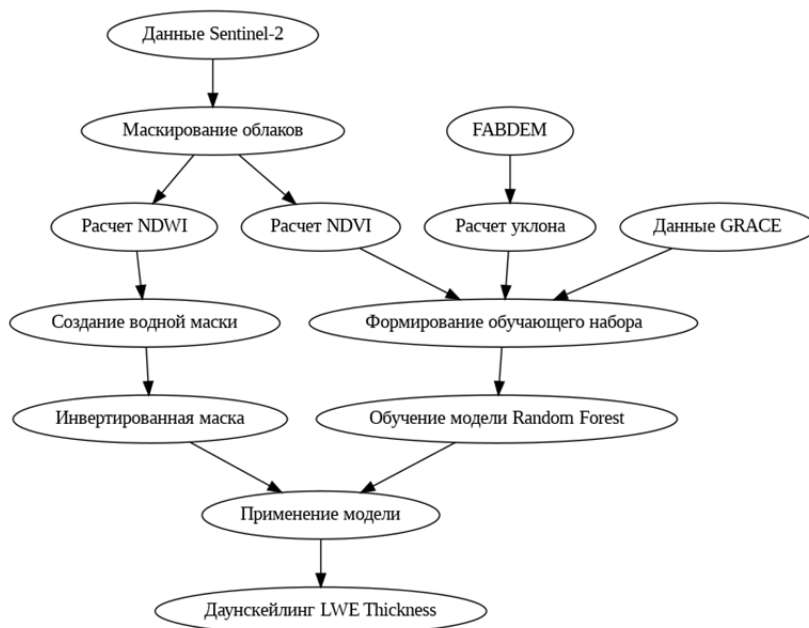


Рисунок 1 – Алгоритм улучшения пространственного разрешения данных, полученных с GRACE-FO

Figure 1 – Algorithm for improving the spatial resolution of GRACE-FO data

Для реализации алгоритма подготовлен и применен код на языке JavaScript на платформе Google Earth Engine [21]. Ниже описано содержание алгоритма.

*Предварительная обработка данных.* Для уменьшения влияния облаков на изучаемые данные на снимках Sentinel-2 использовалась процедура маскировки облаков. Она выполнялась с применением канала QA (оценки качества) снимков Sentinel-2, который предоставляет данные о качестве пикселей, включая наличие облаков и перистых облаков.

Для построения модели использовались следующие предикторы:

1. Топографические особенности. Топографические характеристики (ЦМР FABDEM) включены для учета влияния рельефа на распределение водных ресурсов. Включена высота над уровнем моря, полученная из данных FABDEM. Затем рассчитывается градиент рельефа для каждого пикселя ЦМР. Используемый алгоритм анализирует четыре соседних пикселя для определения локального градиента поверхности.

2. Индекс растительности. Для оценки состояния растительности рассчитан нормализованный разностный индекс растительности (NDVI). Данные Sentinel-2 с марта по октябрь 2023-2024 годов использовались для оценки сезонных закономерностей фенологических фаз растительности. Нами использованы каналы 8 (842 нм) и 4 (665 нм) спутника Sentinel-2. Значения NDVI варьируются от -1 до +1, причем повышенные значения указывают на большую плотность и активность растительности.

3. Маска воды. Нормализованный разностный индекс воды (NDWI) находится для установления маски поверхности воды. Спутник Sentinel-2 использовал каналы 8 (842 нм) и 3 (560 нм). NDWI позволяет идентифицировать водоемы и оценивать содержание влаги. Значения NDWI варьируются от -1 до +1, причем значения, превышающие 0,5, указывают на наличие водоемов. Пиксели, показывающие значения NDWI выше определенного порога, классифицируются как вода и используются для исключения водоемов из процедуры улучшения пространственных данных GRACE.

*Улучшение пространственного разрешения.* Показан результат использования алгоритма Random Forest – метода машинного обучения для улучшения пространственного разрешения данных GRACE. Этот подход был выбран из-за его способности умело управлять нелинейными зависимостями и его устойчивости к переобучению [22]. По определению автора метода, Random Forest – это классификатор, состоящий из набора древовидных классификаторов  $\{h(x, \theta_k), k = 1, \dots\}$ , где  $\{\theta_k\}$  являются независимыми одинаково распределенными случайными векторами, и каждое дерево отдает единичный голос за самый популярный класс на входе  $x$ .

Эквивалентное значение толщины воды из данных GRACE выступало в качестве зависимой переменной, тогда как абсолютная высота, уклон и NDVI служили независимыми переменными. Обученная модель использовалась на наборе данных высокого разрешения для получения улучшенной оценки толщины LWE. В приложении модели применялась инвертированная водная маска, что облегчало исключение водоемов из процедуры уменьшения масштаба.

*Методика оценки надежности полученного результата.* При оценке надежности полученного результата применены метод расчета корреляции Пирсона, коэффициент детерминации, коэффициент статистической значимости корреляции и среднеквадратическое отклонение значений. Для этого на район исследований создан векторный полигональный слой с разбиением на прямоугольные сегменты. На каждый сегмент извлечена статистическая информация по исходному и производному (улучшенному) набору данных GRACE.

**Результаты и их обсуждение.** *Статистическое описание предикторов.* В рамках исследования были проанализированы основные параметры, характеризующие рельеф и растительный покров бассейна реки Жайык. Статистическое описание этих параметров представлено в таблице 2.

Анализ данных цифровой модели рельефа (FABDEM) показывает значительное колебание высотных отметок в исследуемом регионе (рисунок 2, а). Минимальная высота равна -27,9 м, что соответствует прибрежной зоне Каспийского моря. Максимальная высота составляет 1061,2 м в возвышенностях горной цепи Мугалжар (Урал) в верхней части бассейна.

Таблица 2 – Параметры высот, уклона и NDVI

Table 2 – Elevation, slope and NDVI parameters

Параметры	FABDEM, м	Уклон, град.	NDVI
Минимум	-27,982	0,000	-0,129
Максимум	1061,216	18,630	0,580
Среднее	249,850	0,867	0,174
Медиана	243,872	0,561	0,154
Стандартное отклонение	157,187	1,021	0,078

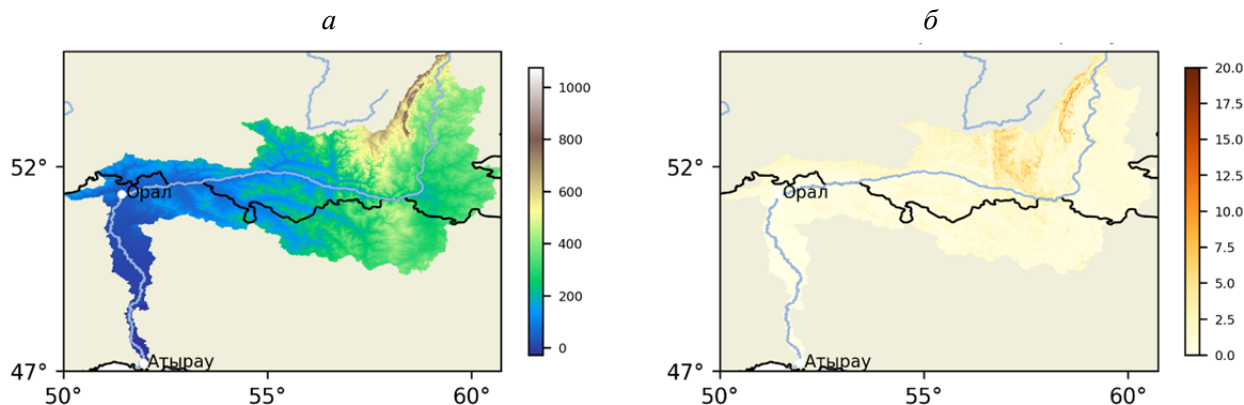


Рисунок 2 – Топографические данные бассейна реки Жайык (Урал):  
*a* – абсолютная высота (FabDEM), м; *б* – уклон поверхности, град.

Figure 2 – Topographic data of the Zhaiyk River basin (Ural):  
*a* – absolute height (FabDEM), m; *b* – surface slope

Средняя высота составляет 249,85 м, хотя медиана равна 243,87 м, что указывает на преобладание ровных и умеренно волнистых ландшафтов. Значительное стандартное отклонение (157,187 м) подразумевает большое разнообразие высот по всему бассейну.

Анализ уклонов рельефа выявляет преимущественно ровный характер рельефа (см. рисунок 2, *б*). Медиана составляет 0,561°, средний уклон – 0,867°. При максимальном уклоне 18,630° ограниченные области с большей высотой показывают свое присутствие либо в верховьях реки, либо на крутых склонах. Стандартное отклонение уклона 1,021° указывает на незначительное изменение этого значения по бассейну.

Нормализованный индекс разности растительности (NDVI) помогает оценить состояние растительного покрова (рисунок 3, *а*).

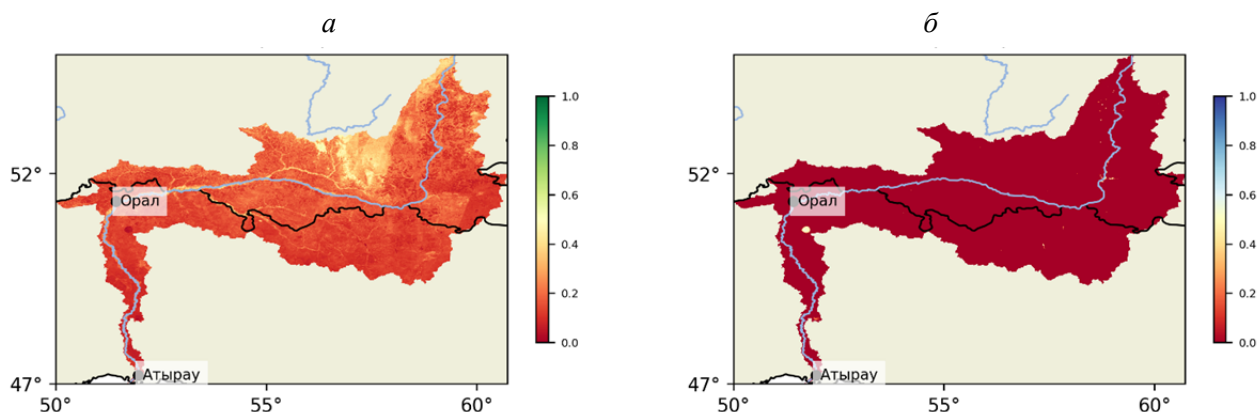


Рисунок 3 – Спектральные индексы на бассейн реки Жайык (Урал): *a* – NDVI; *б* – NDWI

Figure 3 – Spectral indices for the Zhaiyk River basin (Ural): *a* – NDVI; *b* – NDWI

Среднее NDVI составляет 0,174, тогда как медиана – 0,154, что предполагает довольно ограниченный рост растительности, типичный для степных и полупустынных сред. Области с более густой растительностью, вероятно в пойменных или орошаемых зонах, четко видны с максимальным показателем NDVI 0,580.

При минимальном показателе NDVI -0,129 водные поверхности или области, отмеченные редкой растительностью, наверняка себя проявят. Среднеквадратическое отклонение NDVI, составляющее 0,078, указывает на небольшую степень вариации растительного покрова в бассейне. Распределение водных ресурсов зависит от особенностей рельефа и растительности, поэтому они могут существенно влиять на результаты масштабирования данных GRACE. В то время как

изменения высоты и уклона могут влиять на локальные гидрологические процессы, изменения NDVI отражают изменения в эвапотранспирации и потреблении воды растительностью по всему бассейну.

Рассчитан нормализованный индекс водной поверхности NDWI (см. рисунок 3, б) и создана маска водной поверхности для анализа.

**Исходные и улучшенные данные GRACE.** Представлены результаты улучшения пространственного разрешения данных толщины LWE GRACE в бассейне реки Жайык посредством применения метода уменьшения масштаба. В таблице 3 приведены сравнительные результаты между исходными и улучшенными данными.

Таблица 3 – Сравнение исходных и улучшенных данных толщины эквивалентного слоя воды (LWE, см)

Table 3 – Comparison of original and improved water equivalent layer thickness (LWE, cm) data

Параметры	Исходные данные	Улучшенные данные	Изменение
Среднее значение	9,735	9,660	-0,075
Медиана	11,591	10,374	-1,217
Стандартное отклонение	7,445	5,980	-1,465
Минимальное значение	-16,119	-16,119	0
Максимальное значение	19,657	19,657	0

Визуализация исходных (рисунок 4, а) и улучшенных данных толщины эквивалентного водного слоя (см. рисунок 4, б) наглядно показывает, каким образом распределена вода на выбранную дату по бассейну реки Жайык (Урал).

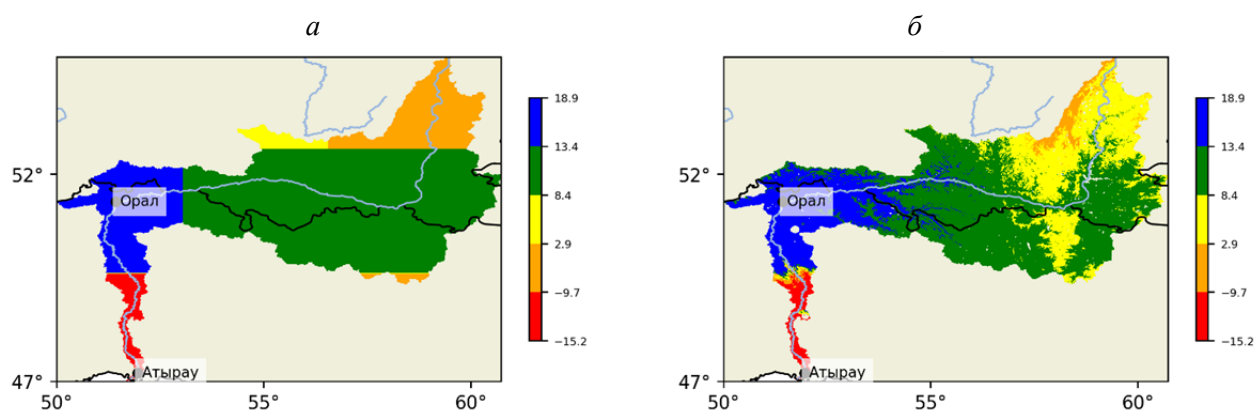


Рисунок 4 – Толщина эквивалентного слоя воды (LWE) в бассейне реки Жайык (Урал):  
а – исходные данные; б – данные с улучшенным пространственным разрешением

Figure 4 – LWE thickness in the Zhaiyk River basin (Ural): a – raw data; b – data with improved spatial resolution

Применение метода улучшения масштаба позволило извлечь статистические свойства данных толщины эквивалентного водного слоя (LWE):

- 1) средняя толщина эквивалентного слоя воды (LWE) уменьшилась с 9,735 до 9,665 см, что указывает на снижение на -0,075 см;
- 2) медиана толщины эквивалентного слоя воды (LWE) снизилась на 1,217 см – с 11,591 до 10,374 см;
- 3) стандартное отклонение показывает наиболее четкий сдвиг; оно уменьшилось с 7,445 до 5,0 см, разница составляет -1,465 см;
- 4) минимальные и максимальные значения толщины LWE остаются прежними (-16,119 и 19,656 см соответственно), тем самым сохраняя общий диапазон данных.

Важно, что уменьшение масштаба выполнялось с использованием дополнительных предикторов, включая рельеф (средняя высота 249,85 м, средний уклон 0,87°) и растительность (средний NDVI 0,174). Эти параметры позволили учесть особенности местного ландшафта, одновременно улучшив пространственное разрешение данных GRACE.

Анализ распределения значений GRACE выявил четкую пространственную дифференциацию исследуемой территории. В северо-западной части региона, в Западно-Казахстанской области, наблюдаются максимальные показатели, варьирующиеся от 11,6 до 19,7 см. Эта зона характеризуется наибольшими положительными аномалиями.

Центральная часть исследуемой территории, включая Оренбургскую область РФ и прилегающие районы, демонстрирует умеренно положительные значения – 8-11,6 см. Примечательно, что в направлении с запада на восток наблюдается постепенное снижение показателей, формируя плавный переход к зоне со значениями 2,4-8 см в Башкортостане (РФ).

В южной части карты, в Атырауской области, показатели достигают минимума -16,12 см в прибрежной части Каспийского моря.

*Оценка надежности полученного результата.* Для оценки полученного результата на район исследований создан векторный полигональный слой с разбиением на прямоугольные сегменты. На каждый сегмент извлечена статистическая информация по исходному и производному (улучшенному) набору данных GRACE. Таким образом, получен ряд данных из 181 значения средней толщины эквивалента водного слоя на оба набора данных, после чего оценены статистическая значимость и коэффициент корреляции (рисунок 5).

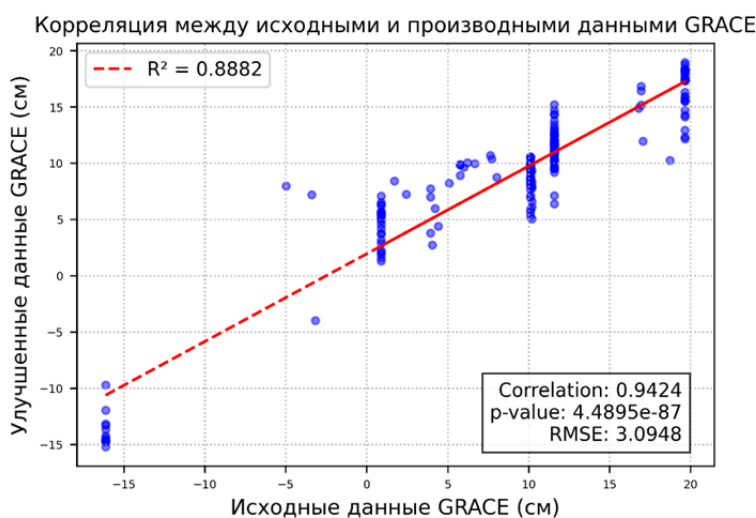


Рисунок 5 – Корреляция между исходными и производными данными GRACE

Figure 5 – Correlation between GRACE raw and derived data

Коэффициент корреляции Пирсона демонстрирует сильную положительную корреляцию и равняется 0,9424, что является высоким показателем. Коэффициент детерминации  $R^2$  указывает, что модель объясняет 88,82% вариации в данных.

Коэффициент статистической значимости p-value указан на графике и равен 4,4895e-87, что является чрезвычайно малым значением (практически нулевым). Это свидетельствует о статистически значимой корреляции между исходными и улучшенными данными GRACE. Такое малое значение p-value говорит о том, что вероятность получить наблюдаемую корреляцию случайно практически равна нулю, что подтверждает достоверность выявленной взаимосвязи между двумя наборами данных,

**Заключение.** Данные, полученные в результате применения методики исследования, показали:

1. Средняя толщина эквивалентного слоя воды (LWE) уменьшилась незначительно, и это указывает на то, что водные ресурсы исследуемой области в целом сохраняются.
2. Медиана толщины эквивалентного слоя воды (LWE) сократилась на 1,217 см, что указывает на сдвиг в сторону более низких значений.
3. Стандартное отклонение показывает наиболее четкий сдвиг, оно уменьшилось на -1,465 см. Это говорит о том, что после снижения масштаба толщина эквивалентного слоя воды (LWE) распределена более равномерно, а экстремальные значения более сглажены.

Минимальные и максимальные значения толщины LWE схожие, тем самым сохраняется общий диапазон данных. Эти данные показывают, что, несмотря на сохранение общего объема и экстремальных значений, процедура уменьшения масштаба дала более детальное и плавное распределение толщины эквивалентного слоя воды (LWE).

В пространственном распределении значений GRACE наблюдается четкая широтная зональность с тенденцией к снижению показателей в южном направлении. Особый интерес представляет область с крайне низкими значениями в южной части карты, в Атырауской области, где показатели достигают минимума -16,119 см. Эта аномалия пространственно совпадает с Прикаспийской низменностью и может быть обусловлена специфическими геологическими и гидрогеологическими условиями региона.

Результаты демонстрируют потенциал метода улучшения пространственного разрешения с применением машинного обучения для детализации данных о водных ресурсах, что может быть особенно полезно для местного управления водными ресурсами и мониторинга окружающей среды в бассейне реки Жайык (Урал) и других территорий. Результаты представляют практический интерес, но требуют дополнительной проверки на локальных данных.

Валидация результатов по наземным данным гидропостов на р. Жайык от РГП «Казгидромет» или полевых исследований не проводилась, так как изначально целью этой работы было тестирование алгоритма по улучшению пространственного разрешения данных. Однако для дальнейшего практического применения такого метода планируется сравнение полученных данных с измерениями на гидрологических постах.

В дополнение нами получено свидетельство о внесении в государственный реестр прав на объекты, охраняемые авторским правом №54285 от 06.02.2025 г. на программное обеспечение «GDA (Алгоритм улучшения пространственного разрешения данных спутниковой гравиметрии GRACE)».

**Финансирование.** Исследование выполнено при финансовой поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан (грант № BR21882366).

## ЛИТЕРАТУРА

- [1] Chen J., et al. Global Ocean Mass Change From GRACE and GRACE Follow-On and Altimeter and Argo Measurements // *Geophys Res Lett*, Blackwell Publishing Ltd. – 2020. – Vol. 47, No. 22.
- [2] Плеханов П. А. Гидрологические риски природного характера и их предупреждение в Казахстане // *Центрально-Азиатский журнал исследований водных ресурсов*. – 2017. – Вып. 3, № 1. – С. 19-25.
- [3] Tulemisova G., et al. Ecological state of the river Ural // *Chemical Bulletin of Kazakh National University*. – 2017. – No. 2. – P. 18-24.
- [4] Plekhanov P. A., Medeu N. N., Skufin P. Hydrological risks and their prevention in Kazakhstan // *International Journal of Hydrology*. – 2019. – Vol. 3, No. 1.
- [5] Gyawali B., et al. Filling Temporal Gaps within and between GRACE and GRACE-FO Terrestrial Water Storage Records: An Innovative Approach // *Remote Sens (Basel)*. – 2022. – Vol. 14, No 7. – P. 1565.
- [6] Revilla-Romero B., et al, On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions // *Remote Sens (Basel)*, Multidisciplinary Digital Publishing Institute. – 2015. – Vol. 7, No. 11. – P. 15702-15728.
- [7] Chang L., Sun W., Consistency analysis of GRACE and GRACE-FO data in the study of global mean sea level change // *Geod Geodyn, KeAi Communications Co.* – 2022. – Vol. 13, No. 4. – P. 321-326.
- [8] Liu Y., G. W., F. J., Z. K. A Summary of Methods for Statistical Downscaling of Meteorological Data // *Advances in Earth Science*. – 2011. – Vol. 26, No. 8. – P. 837.
- [9] Chen L. et al. Downscaling of GRACE-derived groundwater storage based on the random forest model // *Remote Sens (Basel)*, MDPI AG. – 2019. – Vol. 11, No. 24.
- [10] Hu Z, et al, Temporal and spatial variations in the terrestrial water storage across Central Asia based on multiple satellite datasets and global hydrological models // *J. Hydrol. (Amst)*. – 2021. – Vol. 596. – P. 126013.
- [11] Peng Y. et al. Future challenges of terrestrial water storage over the arid regions of Central Asia // *International Journal of Applied Earth Observation and Geoinformation*. – 2024. – Vol. 132. – P. 104026.
- [12] Mannig B. et al. Dynamical downscaling of climate change in Central Asia // *Glob Planet Change*. – 2013. – Vol. 110. – P. 26-39.
- [13] Yin W. et al. Improving the resolution of GRACE-based water storage estimates based on machine learning downscaling schemes // *J. Hydrol. (Amst)*. – 2022. – Vol. 613. – P. 128447.
- [14] Jyolsna P.J., Kambhammettu B.V.N.P., Gorugantula S. Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes // *Hydrological Sciences Journal*. – 2021. – Vol. 66, No. 5.



- [15] Lehner B., Verdin K., Jarvis A., HydroSHEDS Technical Documentation. World Wildlife Fund US. Washington. HydroSHEDS Technical Documentation. World Wildlife Fund US. – Washington, 2006.
- [16] Jet Propulsion Laboratory (JPL). GRACE/GRACE-FO Monthly Mass Grids – JPL Global Mascons.
- [17] Saberi A. et al. Accuracy assessment and improvement of SRTM, ASTER, FABDEM, and MERIT DEMs by polynomial and optimization algorithm: A case study (Khuzestan Province, Iran) // *Open Geosciences*. – 2023. – Vol. 15, No. 1.
- [18] Davis E., Wang C., Dow K. Comparing Sentinel-2 MSI and Landsat 8 OLI in soil salinity detection: a case study of agricultural lands in coastal North Carolina // *Int J. Remote Sens.* Taylor & Francis. – 2019. – Vol. 40, No 16. – P. 6134-6153.
- [19] Maselli F. et al. Evaluation of Terra/Aqua MODIS and Sentinel-2 MSI NDVI data for predicting actual evapotranspiration in Mediterranean regions // *Int J. Remote Sens.* Taylor & Francis. – 2020. – Vol. 41, No. 14. – P. 5186-5205.
- [20] Yang X. et al. Mapping of urban surface water bodies from sentinel-2 MSI imagery at 10 m resolution via NDWI-based image sharpening // *Remote Sens (Basel)*. – 2017. – Vol. 9, No. 6. – P. 1-19.
- [21] Kumar L., Mutanga O. *Google Earth Engine Applications* // *Google Earth Engine Applications*. – 2019.
- [22] Breiman L. *Random forests* // *Mach Learn. Springer*. – 2001. – Vol. 45, No. 1. – P. 5-32.

## REFERENCES

- [1] Chen J. et al. Global Ocean Mass Change From GRACE and GRACE Follow-On and Altimeter and Argo Measurements // *Geophys Res Lett.* Blackwell Publishing Ltd. 2020. Vol. 47, No. 22.
- [2] Plekhanov P. A. Hydrological risks of natural character and their prevention in Kazakhstan // *Central Asian Journal of Water Resources Research*. 2017. Vol. 3, No. 1. P. 19-25 (in Russ.).
- [3] Tulemisova G. et al. Ecological state of the river Ural // *Chemical Bulletin of Kazakh National University*. 2017. No. 2. P. 18-24.
- [4] Plekhanov P. A., Medeu N. N., Skufin P. Hydrological risks and their prevention in Kazakhstan // *International Journal of Hydrology*. 2019. Vol. 3, No. 1.
- [5] Gyawali B. et al. Filling Temporal Gaps within and between GRACE and GRACE-FO Terrestrial Water Storage Records: An Innovative Approach // *Remote Sens (Basel)*. 2022. Vol. 14, No. 7. P. 1565.
- [6] Revilla-Romero B. et al. On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions // *Remote Sens (Basel)*. Multidisciplinary Digital Publishing Institute. 2015. Vol. 7, No. 11. P. 15702-15728.
- [7] Chang L., Sun W. Consistency analysis of GRACE and GRACE-FO data in the study of global mean sea level change // *Geod Geodyn. KeAi Communications Co*. 2022. Vol. 13, No. 4. P. 321-326.
- [8] Liu Y., G.W., F.J., Z.K. A Summary of Methods for Statistical Downscaling of Meteorological Data // *Advances in Earth Science*. 2011. Vol. 26, No. 8. P. 837.
- [9] Chen L. et al. Downscaling of GRACE-derived groundwater storage based on the random forest model // *Remote Sens (Basel)*. MDPI AG. 2019. Vol. 11, No. 24.
- [10] Hu Z. et al. Temporal and spatial variations in the terrestrial water storage across Central Asia based on multiple satellite datasets and global hydrological models // *J. Hydrol. (Amst)*. 2021. Vol. 596. P. 126013.
- [11] Peng Y. et al. Future challenges of terrestrial water storage over the arid regions of Central Asia // *International Journal of Applied Earth Observation and Geoinformation*. 2024. Vol. 132. P. 104026.
- [12] Mannig B. et al. Dynamical downscaling of climate change in Central Asia // *Glob Planet Change*. 2013. Vol. 110. P. 26-39.
- [13] Yin W. et al. Improving the resolution of GRACE-based water storage estimates based on machine learning downscaling schemes // *J. Hydrol. (Amst)*. 2022. Vol. 613. P. 128447.
- [14] Jyolsna P.J., Kambhammettu B.V.N.P., Gorugantula S. Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes // *Hydrological Sciences Journal*. 2021. Vol. 66, No. 5.
- [15] Lehner B., Verdin K., Jarvis A. HydroSHEDS Technical Documentation. World Wildlife Fund US. Washington. HydroSHEDS Technical Documentation. World Wildlife Fund US. Washington, 2006.
- [16] Jet Propulsion Laboratory (JPL). GRACE/GRACE-FO Monthly Mass Grids – JPL Global Mascons.
- [17] Saberi A. et al. Accuracy assessment and improvement of SRTM, ASTER, FABDEM, and MERIT DEMs by polynomial and optimization algorithm: A case study (Khuzestan Province, Iran) // *Open Geosciences*. 2023. Vol. 15, No. 1.
- [18] Davis E., Wang C., Dow K. Comparing Sentinel-2 MSI and Landsat 8 OLI in soil salinity detection: a case study of agricultural lands in coastal North Carolina // *Int J. Remote Sens.* Taylor & Francis. 2019. Vol. 40, No. 16. P. 6134-6153.
- [19] Maselli F. et al. Evaluation of Terra/Aqua MODIS and Sentinel-2 MSI NDVI data for predicting actual evapotranspiration in Mediterranean regions // *Int J Remote Sens.* Taylor & Francis. 2020. Vol. 41, No. 14. P. 5186-5205.
- [20] Yang X. et al. Mapping of urban surface water bodies from sentinel-2 MSI imagery at 10 m resolution via NDWI-based image sharpening // *Remote Sens (Basel)*. 2017. Vol. 9, No. 6. P. 1-19.
- [21] Kumar L., Mutanga O. *Google Earth Engine Applications* // *Google Earth Engine Applications*. 2019.
- [22] Breiman L. *Random forests* // *Mach Learn. Springer*. 2001. Vol. 45, No 1. P. 5-32.

Н. Т. Кемербаев<sup>1</sup>, Қ. Б. Самарханов\*<sup>2</sup>, Г. А. Мусина<sup>3</sup>, М. Х. Шкиева<sup>4</sup>

<sup>1</sup> Т. ғ. к., бас директоры («GeoID» LLC, Астана, Қазақстан; *n.kemerbaeyev@geo-id.kz*)

<sup>2\*</sup> География ғылымдарының кандидаты, директордың ҒЗТҚЖ жөніндегі орынбасары, бас ғылыми қызметкер («GeoID» ЖШС, «Астана» халықаралық ғылыми кешені, Астана, Қазақстан; *kanat.baurzhanuly@gmail.com*)

<sup>3</sup> Кеңселерді өңдеу бөлімінің басшысы («GeoID» ЖШС, Астана, Қазақстан; *g.mussina@geo-id.kz*)

<sup>4</sup> Кеңістіктік деректерді өңдеу және қашықтықтан зондтау бөлімі бастығының орынбасары («GeoID» LLC, Астана, Қазақстан; *maral.shkiyeva@gmail.com*)

### МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНА ОТЫРЫП, ЖАЙЫҚ (ОРАЛ) ӨЗЕНІ БАССЕЙНІ ҮШІН GRACE СПУТНИКТИК ГРАВИМЕТРИЯ ДЕРЕКТЕРІНІҢ КЕҢІСТІКТИК АЖЫРАТЫМДЫЛЫҒЫН ЖАҚСARTУ ТӘЖІРИБЕСІ

**Аннотация.** Мақалада Жайық (Орал) өзені бассейні үшін grace спутниктік гравиметриялық деректерінің кеңістіктік ажыратымдылығын жақсарту үшін машиналық оқытуды қолдану бойынша зерттеу ұсынылған. Жұмыс аймақтық деңгейде су ресурстарын егжей-тегжейлі талдауды қиындататын grace деректерінің төмен кеңістіктік ажыратымдылығы мәселесін шешуге бағытталған. Зерттеу әдістемесі Random Forest моделін қолдана отырып, масштабты азайтудың статистикалық әдісін қолданады. Кеңістіктік ажыратымдылықты жақсарту үшін сандық рельеф моделін, өсімдік жамылғысының индекстерін (NDVI) және су бетінің индекстерін (ndwi) қамтитын қосымша мәліметтер жиынтығы қолданылды. Зерттеу нәтижелері ұсынылған Әдістеменің тиімділігін көрсетті: жақсартылған деректер жалпы статистикалық қасиеттерді сақтай отырып, су қабатының баламалы қалыңдығының дәл таралуын қамтамасыз етеді. Пирсонның корреляция коэффициенті 0.9424 болды. бұл нәтижелердің сенімділігінің айтарлықтай дәрежесін білдіреді. Бұл зерттеудің практикалық маңыздылығы Су ресурстарын бағалауды жақсарту, гидрологиялық оқиғаларды болжау және Жайық өзені бассейніндегі климаттың өзгеруіне бейімделу жоспарларын әзірлеу үшін әзірленген тәсілді перспективалық қолдану болып табылады.

**Түйін сөздер:** гравиметрия, GRACE, Қазақстан, машиналық оқыту, гидрология, рельефтің цифрлық моделі, кеңістіктік рұқсат.

N. T. Kemerbayev<sup>1</sup>, K. B. Samarkhanov\*<sup>2</sup>, G. A. Mussina<sup>3</sup>, M. Kh. Shkiyeva<sup>4</sup>

<sup>1</sup> Candidate of Technical Sciences, General Director («GeoID» LLP, Astana, Kazakhstan; *n.kemerbaeyev@geo-id.kz*)

<sup>2\*</sup> Candidate of Geographical Sciences, Deputy Director for R&D, Chief Scientific Associate («GeoID» LLP, International Scientific Complex «Astana», Astana, Kazakhstan; *kanat.baurzhanuly@gmail.com*)

<sup>3</sup> Head of desk processing department («GeoID» LLP, Astana, Kazakhstan; *g.mussina@geo-id.kz*)

<sup>4</sup> Deputy Head of Spatial Data Processing and Remote Sensing of the Earth Department («GeoID» LLP, Astana, Kazakhstan; *maral.shkiyeva@gmail.com*)

### EXPERIENCE IN IMPROVING THE SPATIAL RESOLUTION OF GRACE SATELLITE GRAVIMETRY DATA FOR THE ZHAIYK RIVER BASIN (URAL) USING MACHINE LEARNING

**Abstract.** The article presents a study on using machine learning to improve the spatial resolution of GRACE satellite gravimetric data for the Zhaiyk River basin (Ural). The work aims to solve the problem of low spatial resolution of GRACE data, making it difficult to analyze water resources in detail at the regional level. The research methodology uses a statistical method of scaling down using the Random Forest model. Additional datasets were used to improve spatial resolution, including a digital terrain model, vegetation indices (NDVI), and water surface indices (NDWI). The results demonstrated the effectiveness of the proposed methodology: improved data provides a more accurate distribution of the equivalent thickness of the water layer while maintaining general statistical properties. The Pearson correlation coefficient was 0.9424, which means a significant degree of reliability of the results. The practical significance of this study lies in the prospective application of the developed approach to improve the assessment of water resources, forecasting hydrological events, and developing plans for adaptation to climate change in the Zhaiyk River basin.

**Keywords:** gravimetry, GRACE, Kazakhstan, machine learning, Hydrology, digital terrain model, spatial resolution.