

https://doi.org/10.51885/3134-8009_IJS_2026_2_15

MPNТИ 36.16.43, 37.31.23, 28.23.25

ПОВЫШЕНИЕ ТОЧНОСТИ ЛОКАЛЬНОЙ МОДЕЛИ ГЕОИДА ГОРОДА АСТАНА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

АСТАНА ҚАЛАСЫНЫҢ ЖЕРГІЛІКТІ ГЕОИД МОДЕЛІНІҢ ДӘЛДІГІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ПАЙДАЛАҢУ АРҚЫЛЫ АРТТЫРУ

IMPROVING THE ACCURACY OF THE LOCAL GEOID MODEL OF ASTANA CITY USING MACHINE LEARNING METHODS

А.С. Уразалиев ¹, Д.А. Шоганбекова ^{1,2*}, М.С. Қожахметов ^{1,3}, Н.Н. Жақсығұл ^{1,2}

¹Институт ионосферы, г. Алматы, Казахстан

²Международная образовательная корпорация, г. Алматы, Казахстан

³Казахский национальный исследовательский технический университет им. К.И. Сатпаева,
г. Алматы, Казахстан

*Автор-корреспондент: Шоганбекова Дания Асыгатовна, e-mail: d.shoganbekova@ionos.kz

Ключевые слова:

моделирование, модель
геоида, LSMSA,
преобразование
Гельмерта, регрессоры
машинного обучения,
гауссовская процессная
регрессия, регрессия
опорных векторов.

АННОТАЦИЯ

Разработка высокоточной региональной модели геоида является важным этапом модернизации высотной основы Казахстана и повышения эффективности определения высот по данным ГНСС. В работе предложен гибридный подход к уточнению локальной модели геоида для района г. Астана, сочетающий базовое гравиметрическое решение LSMSA и статистическое моделирование остаточных поправок. Базовая поверхность геоида рассчитывается методом модификации формулы Стокса по методу наименьших квадратов с аддитивными поправками и проверяется контрольными пунктами ГНСС/Нивелирования. Для уменьшения расхождений с наблюдениями формируется корректирующая поверхность по остаткам с использованием 7-параметрического преобразования Гельмерта и регрессионных моделей машинного обучения (GPR, SVR, LSBoost). На независимой тестовой подвыборке для SVR получено снижение RMSE с 0,062 до 0,047 м, что соответствует уменьшению ошибки на 24%. Результаты показывают, что предложенная схема уменьшает разброс остаточных ошибок и улучшает согласование модели геоида с данными ГНСС/Нивелирования.

Түйінді сөздер:

модельдеу, геоид моделі,
LSMSA, Гельмерт
түрлендіруі, машиналық
оқыту регрессорлары,
гаустық процестік

ТҮЙІНДЕМЕ

Жоғары дәлдікті өңірлік геоид моделін әзірлеу Қазақстанның биіктік негізін жаңғыртудың және ЖНСЖ деректері бойынша биіктікті анықтау тиімділігін арттырудың маңызды кезеңі болып табылады. Жұмыста Астана қаласы ауданы үшін жергілікті геоид моделін нақтылауға арналған, LSMSA гравиметриялық базалық



регрессия, тірек
векторлар регрессиясы.

шешімін және қалдық түзетулерді статистикалық модельдеуді біріктіретін гибриді тәсіл ұсынылған. Геоидтың базалық беті аддитивті түзетулері бар Стокс формуласын ең кіші квадраттар әдісімен модификациялау арқылы есептеліп, ЖНСЖ /Нивелирлеу бақылау пункттері бойынша тексеріледі. Бақылаулармен айырмашылықтарды азайту үшін қалдықтар негізінде 7-параметрлі Гельмерт түрлендіруі және машиналық оқыту регрессиялық модельдері (GPR, SVR, LSBoost) қолданылып, түзету беті құрылады. Тәуелсіз тестік таңдамада SVR үшін RMSE 0.062-ден 0.047 м-ге дейін төмендеді, бұл қателіктің 24%-ға азаюына сәйкес келеді. Нәтижелер ұсынылған схеманың қалдық қателердің шашырауын азайтып, геоид моделінің ЖНСЖ /Нивелирлеу деректерімен сәйкестігін жақсартатынын көрсетеді.

Keywords:

modeling, geoid model,
LSMSA, Helmert
transformation, machine
learning regressors,
Gaussian process
regression, support vector
regression.

ABSTRACT

The development of a high-precision regional geoid model is an important step in the modernization of Kazakhstan's vertical reference framework and in improving the efficiency of height determination from GNSS data. This paper proposes a hybrid approach to refining the local geoid model for the Astana area by combining the LSMSA gravimetric base solution with statistical modeling of residual corrections. The base geoid surface is computed using the Least Squares Modification of Stokes' Formula with additive corrections and is validated against GNSS/levelling control points. To reduce discrepancies with observations, a correction surface is formed from residuals using a 7-parameter Helmert transformation and machine learning regression models (GPR, SVR, LSBoost). On the independent test subset, SVR reduced RMSE from 0.062 to 0.047 m, corresponding to a 24% decrease in error. The results show that the proposed scheme reduces the spread of residual errors and improves the agreement of the geoid model with GNSS/levelling data.

ВВЕДЕНИЕ

Создание точной модели геоида является одной из ключевых задач современной геодезии, поскольку она лежит в основе единой национальной высотной системы и согласованного определения высот по данным ГНСС. Для Казахстана это особенно актуально из-за больших размеров территории и контрастного рельефа — от Каспийской низменности до горных систем Тянь-Шаня и Алтая. В таких условиях глобальные модели геопотенциала (ГГМ), включая EGM2008 и XGM2019e, как правило, дают лишь усредненное описание гравитационного поля и не всегда обеспечивают требуемую точность для инженерно-геодезических приложений, что обуславливает необходимость разработки регионально адаптированных моделей (Shoganbekova et al., 2025). Традиционно при построении геоида используются данные нивелирования и ГНСС, однако такие наблюдения затратны и ограничены по плотности, а также могут содержать пропуски и неоднородности (Elshewy et al., 2024; Chen et al., 2022).

Высокоточные региональные решения рассматриваются как важная часть формирования государственной системы координат и высот и создают основу для перехода от трудоемкого геометрического нивелирования к ГНСС-технологиям, где ортометрическая высота определяется как разность эллипсоидальной высоты и высоты геоида.

Среди современных гравиметрических подходов широкое распространение получил метод модификации формулы Стокса по методу наименьших квадратов с аддитивными поправками (LSMSA, «метод КТН») (Abbak & Ustun, 2015; Işık et al., 2021), обеспечивающий

согласованное объединение аномалий силы тяжести, топографических поправок по ЦМР и данных ГТМ с учетом усечения и спектральных ошибок. Однако даже при физически корректной постановке нередко сохраняются остаточные расхождения с данными ГНСС/Нивелирования, обусловленные несогласованностями датумов, шумом измерений и локальными неоднородностями (Abdalla & Mogren, 2015). Для их уменьшения применяют параметрическое преобразование (например, по Гельмерту), но глобальные параметрические модели не всегда способны описать пространственно сложные систематические эффекты.

Прогресс в машинном обучении расширил возможности геодезического моделирования за счет гибких, основанных на данных, методов корректировки остаточных расхождений при вычислении геоида. Показано, что подходы на основе опорных векторов и нейросетевые архитектуры эффективно аппроксимируют высоты геоида по данным ГНСС/Нивелирования, а более поздние работы развили эти идеи, предложив регрессионные схемы повышенной сложности (включая ансамблевые и эвристические модели), которые нередко превосходят классические интерполяторы в локальных тестовых районах (Zaletnyik et al., 2008; Ghaffari-Razin & Hooshangi, 2024). Существенный интерес представляют ядровые методы, прежде всего гауссовская процессная регрессия (GPR), а также гибридные решения, где глобальные модели геопотенциала дополняются ML-моделированием остатков с учетом рельефа и характеристик окрестности (Konakoglu & Akar, 2021; Pham et al., 2025). В целом ML-подходы ценны способностью описывать нелинейные зависимости без жесткой параметризации, работать при нерегулярной пространственной выборке и часто снижать среднеквадратическую ошибку и разброс остатков по сравнению с IDW, кригингом и решениями, основанными только на ГТМ, особенно при ограниченной плотности опорных данных (Konakoglu & Akar, 2022; Liu et al., 2024). На практике чаще всего применяются SVR, GPR и усиление методом наименьших квадратов (LSBoost), при этом гиперпараметры обычно подбирают перебором по сетке или байесовской оптимизацией с k-кратной кросс-валидацией (Butt et al., 2021).

В настоящей работе методы машинного обучения рассматриваются не как замена гравиметрической модели, а как компактный корректирующий слой, дополняющий базовое LSMSA-решение. В качестве объекта исследования выбран район г. Астана, для которого построено базовое гравиметрическое решение LSMSA на основе доступных гравиметрических данных, глобальной модели XGM2019e_2159 и ЦМР FABDEM с разрешением 30 м. Далее остаточные расхождения с данными ГНСС/Нивелирования моделируются двумя классами подходов: 7-параметрическим преобразованием Гельмерта и регрессорами машинного обучения (GPR, SVR, LSBoost), что позволяет сравнить их по точности и устойчивости на независимой тестовой выборке. Такой подход обеспечивает целостную схему построения локальной модели геоида для Астаны, в которой физически обоснованная компонента задается LSMSA, а локальные систематические эффекты учитываются через моделирование остаточных поправок.

Территория исследования. На севере Казахстана, в пределах окрестностей города Астана, расположена исследуемая территория, ограниченная координатами 49,767°–52,5° с. ш. и 70,05°–72,783° в. д. (рисунок 1). В геоморфологическом отношении она приурочена к бассейну реки Ишим. Для данной местности характерны волнистые равнины, развитая речная сеть, незначительные высотные различия и повсеместное распространение четвертичных отложений. При таких природных условиях пространственная изменчивость гравитационного поля определяется главным образом неоднородностью плотности отложений, изменением мощности рыхлых пород, а также локальным развитием процессов эрозии и аккумуляции. По сравнению с горными территориями роль топографического фактора здесь существенно менее выражена (Reguzzoni et al., 2021).

Несмотря на то, что рассматриваемая территория относится к районам с низкой сейсмической активностью, для нее характерны медленные деформации земной коры, а также воздействие сезонной гидрологической нагрузки. Последняя связана с таянием снежного покрова, изменением влажности грунтов, колебанием уровней воды в бассейне реки Ишим и регулирующим влиянием водохранилищ. Совместное действие указанных факторов приводит к возникновению пространственно согласованных колебаний очень малой амплитуды как в гравитационном поле, так и в вертикальных смещениях земной поверхности (Hu et al., 2021).

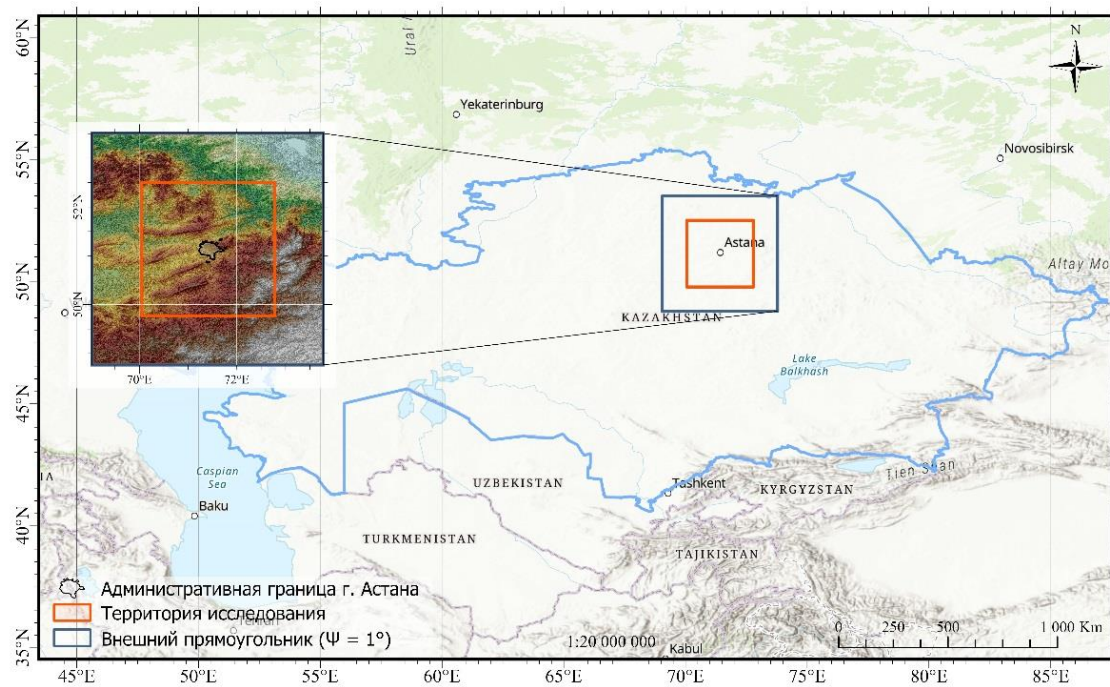


Рисунок 1. Расположение района исследования

Примечание – составлено авторами

С целью ослабления краевых эффектов вычисления проводились в пределах внутренней прямоугольной области, тогда как этапы подготовки и предварительной обработки исходных данных выполнялись в расширенной зоне, включавшей буфер шириной 1° . Применение такой схемы способствует повышению устойчивости вычислительных процедур и обеспечивает корректный учет топографических воздействий наряду с длинноволновой составляющей гравитационного поля (El Shouny et al., 2021).

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Гибридная модель геоида в данном исследовании построена на основе четырех взаимодополняющих наборов данных: наземных гравиметрических измерений, глобальной модели геопотенциала (ГТМ), цифровой модели рельефа (ЦМР) высокого разрешения и пунктов ГНСС/Нивелирования. В настоящем исследовании длинноволновая составляющая гравитационного поля представлена глобальной гравитационной моделью XGM2019e_2159, поскольку это одна из наиболее актуальных комбинированных моделей, интегрирующих спутниковые данные (GRACE, GOCE) и наземные гравиметрические данные [Liang et al., 2020; Zingerle et al., 2020]. Топографическая информация и поправки за рельеф получены на основе набора FABDEM с разрешением 30 м, обеспечивающего улучшенное представление высот поверхности земли (Li et al., 2022). Однако основными

источниками информации являются гравиметрические данные – сформированные на основе оцифрованных исторических съемок, проверенные современными наземными измерениями и согласованные с глобальными моделями – а также данные ГНСС/Нивелирования, которые используются как для AI-поддерживаемой корректировки геоида, так и для независимой валидации модели геоида (Shoganbekova et al., 2025a; Shoganbekova et al., 2025b).

Данные ГНСС/Нивелирования. Для независимой валидации локальной модели геоида использована сеть из 70 пунктов ГНСС/Нивелирования в пределах территории исследования. Их пространственное распределение приведено на рисунке 2. Набор сформирован в рамках проекта по развитию Национальной инфраструктуры пространственных данных («Геопортал открытых пространственных данных Республики Казахстан», 2026) и включен в официальную государственную систему координат QazTRF-23 («Государственное геодезическое обеспечение», 2026). Для каждого пункта доступны нормальные высоты в системе БСВ-77 и эллипсоидальные высоты, полученные по результатам ГНСС-наблюдений.

По классу точности в выборке представлены 5 пунктов 1-го класса, 22 пункта 2-го класса и 43 пункта 3-го класса. Нормальные высоты в сети имеют следующие статистические параметры: min 220,5 м, mean 392,8 м, max 609,4 м.

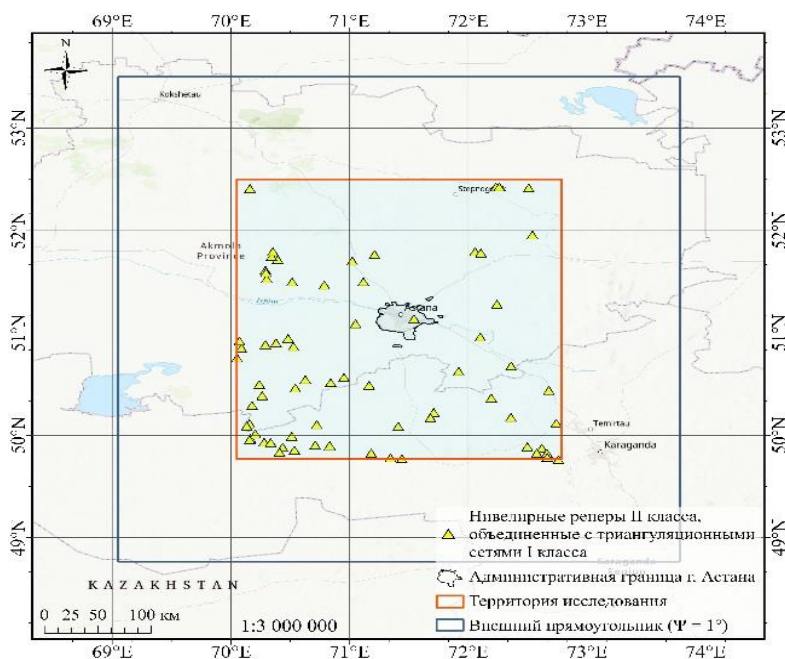


Рисунок 2. Пространственное распределение пунктов ГНСС/Нивелирования национальной нивелирной сети

Примечание – составлено авторами

Методы локального моделирования геоида. В работе используется комбинированная методика, в которой метод модификации формулы Стокса по методу наименьших квадратов с аддитивными поправками (LSMSA, «метод КТН») дополняется инструментами машинного обучения для повышения точности локальной модели геоида. Общий алгоритм построения разделен на три ключевых этапа (рисунок 3).

На *первом этапе* формируется базовая гравиметрическая поверхность геоида по схеме LSMSA (Sjöberg, 2003). Данный подход обеспечивает строгую постановку задачи и позволяет согласованно объединять разнородные источники информации – наземные

аномалии силы тяжести, глобальные модели геопотенциала (ГГМ), цифровые модели рельефа и данные ГНСС/Нивелирования – в единую региональную модель.

В рамках LSMSA параметры модификации ядра Стокса оптимизируются методом наименьших квадратов, что направлено на минимизацию среднеквадратической ошибки вычисленных значений высоты геоида (Sjöberg & Bagherbandi, 2017). При этом учитываются:

- ошибки усечения и спектральные составляющие погрешностей аномалий силы тяжести;
- регуляризация плохо обусловленной системы для гармоник высоких степеней;
- аддитивные поправки, связанные с влиянием рельефа, атмосферы, эллипсоидальных эффектов и операций продолжения вниз.

Итогом первого этапа является гравиметрическая модель N_{LSMSA} , используемая в дальнейшем как базовое решение для последующего уточнения.

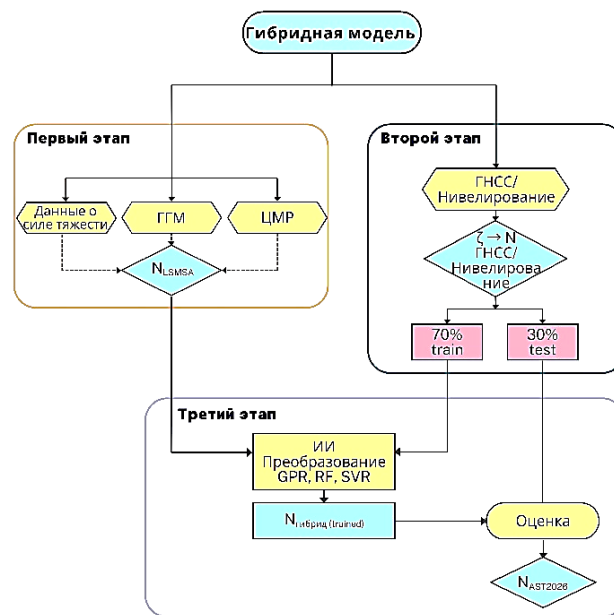


Рисунок 3. Процесс построения гибридной модели геоида

Примечание – адаптировано по (Urazaliyev et al., 2025)

На *втором этапе* значения $N_{GNSS/lev}$ разбиваются на обучающую и тестовую подвыборки в пропорции 70 % и 30 % соответственно для последующего сопоставления 7-параметрического преобразования Гельмерта и AI-регрессоров. Разбиение 70/30 в формате hold-out является распространенной практикой для малых и умеренных наборов данных ГНСС/Нивелирования и соответствует подходам, применяемым в ряде современных работ по моделированию геоида (Elshevy et al., 2024; Konakoglu & Akar, 2021; Pham et al., 2025).

Третий этап методики включает построение корректирующей поверхности по точечным остаткам между значениями высоты геоида, полученными по данным ГНСС/Нивелирования, и гравиметрическим базовым решением, после чего данная поправка накладывается на поверхность LSMSA для формирования итоговой гибридной модели. В традиционной постановке согласование выполняют с использованием преобразования Гельмерта (3, 5 или 7-параметрического), позволяющего описать систематические расхождения через сдвиги, вращения и масштабные искажения. Однако в условиях различий высотных датумов или наличия топографических несогласованностей преобразование по Гельмерту может приводить к сохранению систематических ошибок.

Поэтому для более гибкого и точного описания остаточного поля предлагается применять AI-регрессию, аппроксимирующую разность между значениями ГНСС/Нивелирования и моделью LSMSA (Urazaliyev et al., 2025).

Введем вектор признаков $\chi = [E, N, N_{LSMSA}]$, включающий локальные координаты восток–север (в километрах) и базовое значение высоты геоида LSMSA, интерполированное в положение пункта (в метрах). Для каждого контрольного пункта вычисляются точечные остатки в виде

$$y(\chi) = \Delta N(\chi) = N_{GNSS/Lev}(\chi) - N_{LSMSA}(\chi), \quad (1)$$

далее обучается регрессор $\hat{f}(\chi)$, построенный по данным, так что $\hat{f}(\chi) \approx \Delta N(\chi)$. После этого итоговую гибридную модель геоида формируют путем наложения соответствующего корректирующего поля на базовую поверхность,

$$N_{hyb}(\chi) = N_{LSMSA}(\chi) + f(\chi). \quad (2)$$

Для аппроксимации ΔN в пределах района исследования применяются три взаимодополняющих регрессионных подхода: гауссовская процессная регрессия (GPR), усиление методом наименьших квадратов (LSBoost) и регрессия опорных векторов (SVR). Модели обучаются на 70% набора данных ГНСС/Нивелирования и оцениваются на оставшихся 30% в рамках hold-out проверки.

Модель GPR реализована с использованием ядра с автоматическим определением релевантности (ARD) на основе квадратично-экспоненциальной ковариационной функции (GPR, ARD-ядро). В этой постановке на скрытое поле $f(\cdot)$ задается априорное распределение гауссовского процесса при независимом и одинаково распределенном гауссовом шуме наблюдений,

$$f(\chi) \sim \mathcal{GP}(0, k_{\theta}(\chi, \chi')), \quad y_i = f(\chi_i) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_n^2). \quad (3)$$

Для повышения адаптивности GPR применяется схема автоматического определения релевантности (ARD), при которой для признаков E, N и N_{LSMSA} оцениваются отдельные масштабы длины. В вычислительной реализации в качестве основной ковариационной функции используется ядро Матерна $\nu = 3/2$; при необходимости допускается переход к ядру Матерна $\nu = 5/2$ или к квадратично-экспоненциальному ядру. Далее запишем:

$$r(\chi, \chi') = \sqrt{\sum_j \frac{(x_j - x'_j)^2}{l_j^2}} \quad (4)$$

Ядро имеет вид:

$$k_{Matèrn-3/2} = \sigma_f^2 (1 + \sqrt{3}r) e^{-\sqrt{3}r}, \quad k_{SE} = \sigma_f^2 \exp(-\frac{1}{2} \sum_j \frac{(x_j - x'_j)^2}{l_j^2}). \quad (5)$$

При заданных обучающих входах X и целевых значениях y апостериорный прогноз в точке x^* характеризуется математическим ожиданием и дисперсией,

$$\mu_*(x_*) = k_*^T (K + \sigma_n^2 I)^{-1} y, \quad \sigma_*^2(x_*) = k_{**} - k_*^T (K + \sigma_n^2 I)^{-1} k_*, \quad (6)$$

при этом гиперпараметры $\theta = \{\sigma_f, \{l_j\}, \sigma_n\}$ определяются посредством максимизации логарифма предельного правдоподобия. Разнородность единиц измерения входных переменных (E, N в км; N_{LSMSA} в м) учитывается за счет z-стандартизации и механизма ARD.

Регрессия опорных векторов (SVR, RBF-ядро). В модели SVR подбирается функция с малой нормой в пространстве RKHS при условии, что отклонения (остатки) остаются внутри ε -нечувствительной зоны:

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad s. t. \quad \begin{cases} y_i - f(\chi_i) \leq \varepsilon + \xi_i, \\ f(\chi_i) - y_i \leq \xi_i^*, \\ \xi_i, \xi_i^* \geq 0. \end{cases} \quad (7)$$

Согласно теореме о представлении, искомая функция в SVR может быть записана в виде линейной комбинации ядерных функций по обучающим точкам:

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) k(\chi_i, \chi) + b, \quad (8)$$

где в данной работе используется RBF-ядро

$$k(\chi, \chi') = \exp\left(-\frac{\|\chi - \chi'\|^2}{2\sigma^2}\right) \quad (9)$$

Параметры C , ε и σ (KernelScale) задают компромисс между гладкостью аппроксимации и устойчивостью к шуму. Аналогично постановке для GPR входные переменные (E , N , N_{LSMSA}) используются в исходных единицах (координаты – в км, N_{LSMSA} – в м) и перед обучением приводятся к стандартизованному виду.

Усиление методом наименьших квадратов (LSBoost). Метод LSBoost строит аддитивную модель как ансамбль регрессионных деревьев, последовательно уменьшая квадратичную функцию потерь. В качестве начального приближения берется константная модель, минимизирующая сумму квадратов отклонений:

$$f_0(\chi) = \arg \min_c \sum_i (y_i - c)^2 = \bar{y} \quad (10)$$

Далее для $m=1, \dots, M$, $r_{im} = y_i - f_{m-1}(\chi_i)$, по которым обучается очередное дерево h_m , после чего выполняется обновление

$$h_m \leftarrow \text{tree fit to } \{(\chi_i), r_{im}\}, f_m(\chi) = f_{m-1}(\chi) + v h_m(\chi), v \in (0, 1] \quad (11)$$

Регуляризация задается числом итераций M , коэффициентом обучения v и параметрами сложности деревьев (например, минимальным размером листа). В результате формируется гибкая кусочно-гладкая аппроксимация $\Delta N(E, N, N_{LSMSA})$, дополняющая более гладкие ядровые модели GPR и SVR.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

На рисунке 4 представлены этапы формирования высоты геоида методом LSMSA, который состоит из вычисления приближенных высот, аддитивных поправок и итоговой высоты геоида. Показанное распределение характеризует региональную структуру поверхности геоида, восстановленную по данным наземной гравиметрии с учетом длинноволновой составляющей, заданной моделью XGM2019e_2159.

В таблице 1 приведены статистические характеристики аддитивных поправок: топографической, поправки за аналитическое продолжение вниз, эллипсоидальной и атмосферной.

Таблица 1. Статистика примененных аддитивных поправок

Поправки	Min, м	Max, м	Mean, м	STD, м
Ntopo	-0,0500	-0,0029	-0,0169	0,0064
Ndws	-0,0060	0,0243	0,0064	0,0041
Nell	-0,0006	0,0000	-0,0003	0,0001
Natm	0,0002	0,0007	0,0004	0,0001
All	-0,0330	0,0051	-0,0104	0,0062
Примечание – составлено авторами				

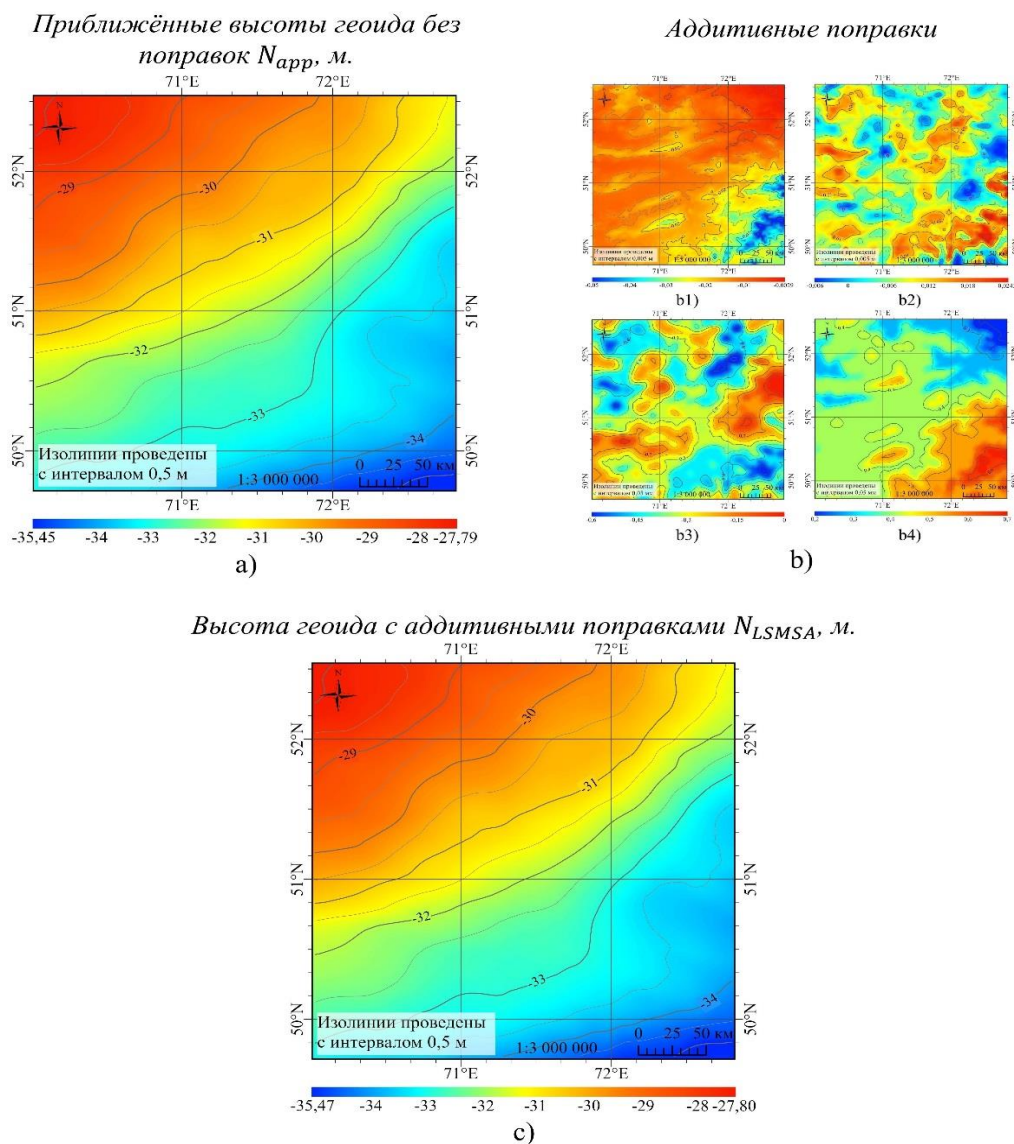


Рисунок 4. Формирование высоты геоида методом LSMSA:

а – Приближенные высоты геоида без поправок N_{app} (м); б – Аддитивные поправки: b1) Топографические поправки (м), b2) Поправки за аналитическое продолжение вниз (м), b3) Эллипсоидальные поправки (мм), b4) Атмосферные поправки (мм). с) Высота геоида с аддитивными поправками N_{LSMSA} (м).

Примечание – составлено авторами

Результаты расчета по методу LSMSA показывают, что высота геоида в пределах исследуемой территории изменяется от -35,477 м до -27,803 м; среднее значение составляет -31,723 м при стандартном отклонении 1,659 м.

Корректирующая поверхность с использованием ИИ. Были сформированы остатки между наблюдаемыми значениями высоты геоида $N_{GNSS/Lev}$ и значениями N_{LSMSA} , интерполированными в положения 70 контрольных пунктов (обучающая выборка – 50, тестовая – 20).

В таблице 2 приведены статистические характеристики распределения ошибок для исходного варианта (До) и для набора моделей корректировки (После) с отдельной оценкой по выборкам ALL/TRAIN/TEST.

Таблица 2. Статистика остаточных ошибок при корректировке геоида
для выборок ALL/TRAIN/TEST, м

Метод	Разделение	Кол-во	Mean	Median	STD	RMSE	MAE	IQR	Min	Max
До										
-	All	70	-0,015	-0,005	0,054	0,056	0,042	0,071	-0,179	0,116
	Train	50	-0,011	-0,003	0,052	0,053	0,041	0,072	-0,170	0,116
	Test	20	-0,024	-0,010	0,059	0,062	0,043	0,065	-0,179	0,073
После										
Гель-мерт	All	70	0,001	0,004	0,048	0,048	0,034	0,046	-0,161	0,107
	Train	50	0,000	0,003	0,043	0,043	0,030	0,039	-0,161	0,107
	Test	20	0,002	0,009	0,060	0,058	0,046	0,087	-0,148	0,089
SVR (RBF)	All	70	0,000	0,002	0,041	0,041	0,033	0,068	-0,114	0,142
	Train	50	0,002	0,003	0,038	0,038	0,031	0,078	-0,051	0,142
	Test	20	-0,006	-0,001	0,048	0,047	0,037	0,057	-0,114	0,076
GPR (ARD)	All	70	-0,001	0,003	0,043	0,043	0,031	0,039	-0,140	0,100
	Train	50	0,000	0,002	0,041	0,040	0,029	0,039	-0,133	0,100
	Test	20	-0,004	0,006	0,050	0,049	0,035	0,047	-0,140	0,059
LSBoost	All	70	-0,001	0,000	0,028	0,028	0,014	0,006	-0,133	0,062
	Train	50	0,000	0,000	0,004	0,004	0,003	0,003	-0,011	0,013
	Test	20	-0,002	0,019	0,054	0,052	0,043	0,069	-0,133	0,062
Примечание – составлено авторами										

В таблице 2 приведены сводные характеристики остатков по контрольным пунктам ГНСС/Нивелирования для базового решения LSMSA (До/-) и вариантов корректировки (После) с отдельной оценкой по ALL/TRAIN/TEST. Для исходного варианта сохраняется отрицательное смещение, а уровень ошибок на тестовой подвыборке остается выше, чем по всей выборке (RMSE: 0,062 м для TEST при 0,056 м для ALL), что указывает на необходимость дополнительного согласования гравиметрической поверхности с наземными наблюдениями.

7-Параметрическое преобразование Гельмерта ожидаемо приводит среднюю ошибку к значениям, близким к нулю, и уменьшает RMSE по всей выборке до 0,048 м. Однако на независимой тестовой выборке улучшение менее выражено (RMSE = 0,058 м), а сохраняющийся разброс ошибок показывает, что одной глобальной параметрической модели недостаточно для описания остаточного поля.

Более заметный эффект на TEST обеспечивают регрессоры машинного обучения. Для SVR (RBF) и GPR (ARD) RMSE на тестовых точках составляет 0,047–0,049 м при малом смещении и более компактном распределении ошибок, что отражает более гибкое описание пространственной структуры остатков по сравнению с преобразованием Гельмерта. Однако модель LSBoost показала признаки переобучения. При практически нулевой ошибке на обучающей (TRAIN) выборке (RMSE $\approx$ 0,004 м) на тестовых данных (TEST) качество ухудшается (RMSE $\approx$ 0,052 м). Это указывает на необходимость дополнительного контроля сложности модели, если рассматривать данный метод как основной способ корректировки.

Гибридная модель геоида (LSMSA + ИИ). На данном этапе была построена корректирующая поверхность по остаткам с применением регрессионных моделей машинного обучения. Это позволило добиться повышения согласованности между базовым гравиметрическим решением LSMSA и наземными наблюдениями ГНСС/Нивелирования. Как видно из таблицы 2, SVR с RBF-ядром показывает наиболее устойчивое качество на

независимой тестовой подвыборке. Это, в свою очередь, послужило обоснованием для выбора данного метода при формировании итоговой гибридной модели геоида.

Гибридная поверхность строится как суперпозиция базовой модели и корректирующего поля, вычисленного регрессором:

$$N_{AST2026} = N_{LSMSA} + f_{SVR}(\Delta X) \quad (12)$$

где $f_{SVR}(\Delta X)$ аппроксимирует остаточную поправку ΔN , а χ – вектор признаков, включающий локальные координаты восток–север и значение N_{LSMSA} , интерполированное в соответствующей точке. Далее модель f_{SVR} применяется ко всем узлам расчетной сетки, после чего полученное поле поправок добавляется к N_{LSMSA} . Тем самым сохраняется физически согласованная длинноволновая компонента, определяемая LSMSA.

Результаты пространственной реализации показаны на рисунке 5. На рисунке 5, а представлено корректирующее поле ΔN_{SVR} , отражающее пространственную структуру поправок, необходимых для согласования LSMSA с наблюдениями. На рисунке 5, б приведена итоговая поверхность высоты геоида $N_{LSMSA+SVR}$, рассчитанная на сетке с разрешением $2' \times 2'$. Визуально видно, что корректирующее поле имеет локальный характер и изменяется в пределах исследуемой области, тогда как общая форма поверхности геоида определяется базовым гравиметрическим решением.

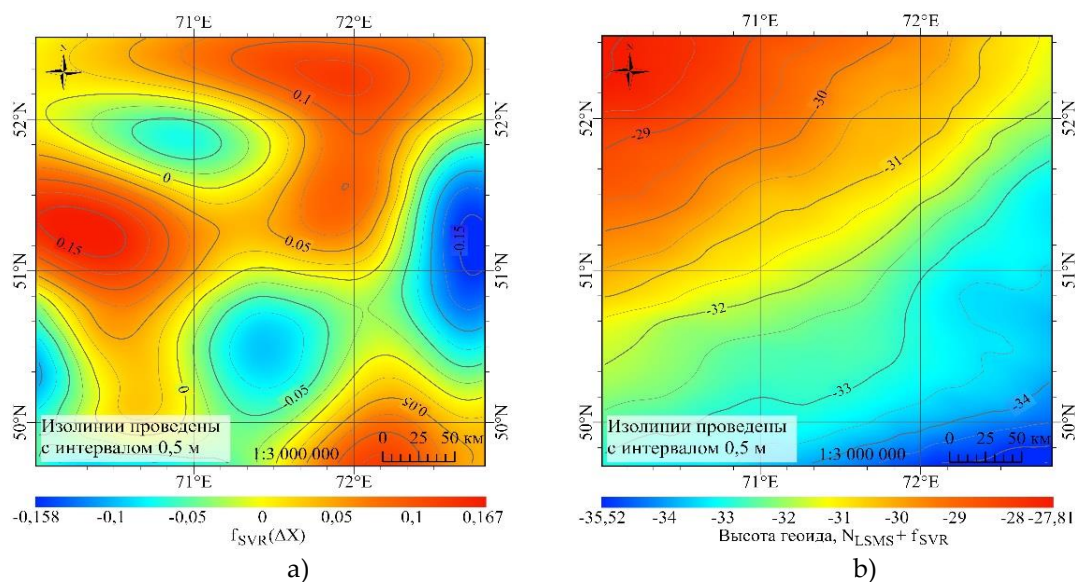


Рисунок 5. а – Корректирующее поле ΔN_{SVR} (м);
б – Итоговая поверхность высоты геоида $N_{LSMSA+SVR}$ (м).

Примечание – составлено авторами

На рисунке 6 представлено распределение ошибок до и после введения корректирующей поправки. Для исходного варианта (рисунок 6, а) характерно более выраженное смещение, тогда как после применения SVR-коррекции (рисунок 6, б) распределение становится ближе к нулевому центру. Это подтверждается снижением RMSE на независимой тестовой подвыборке по сравнению как с базовым решением LSMSA, так и с 7-параметрическим преобразованием Гельмерта (таблица 2). Полученные результаты показывают эффективность выбранной гибридной схемы при уточнении поверхности геоида на исследуемой территории.

Результаты расчетов показывают, что базовое гравиметрическое решение LSMSA обеспечивает физически согласованную поверхность геоида для района Астаны, однако без

дополнительного согласования с данными ГНСС/Нивелирования сохраняются остаточные систематические и случайные составляющие. При оценке до преобразования (Before, таблица 2) наблюдается устойчивое отрицательное смещение (Mean на TEST = -0,024 м), а уровень ошибок на независимых точках остается заметным (RMSE на TEST = 0,062 м). Это означает, что даже в условиях относительно спокойного рельефа локальные несогласованности между гравиметрическим решением и контрольными наблюдениями проявляются статистически значимо и требуют корректировки остаточного поля.

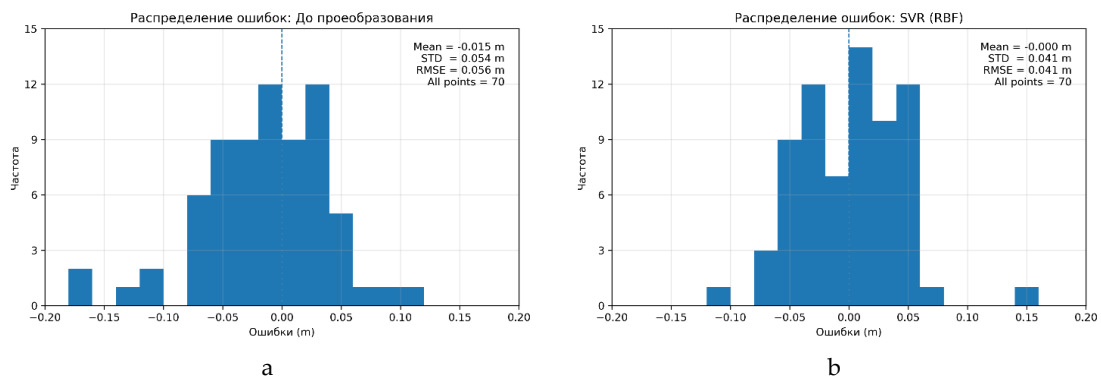


Рисунок 6. Гистограммы распределения ошибок модели геоида до и после коррекции на основе SVR: а – до коррекции; б – после коррекции, м

Примечание – составлено авторами

7-параметрическое преобразование Гельмерта ожидаемо решает задачу глобального выравнивания уровней. Систематический сдвиг практически устраняется (Mean на TEST $\approx$ 0,002 м), а общая ошибка уменьшается (RMSE на ALL снижается с 0,056 до 0,048 м, то есть примерно на 14%). Однако на тестовой подвыборке эффект оказывается существенно слабее: RMSE снижается лишь с 0,062 до 0,058 м (около 6 %). Такая разница между ALL и TEST указывает, что параметрическая модель, эффективно снимая общий сдвиг, не полностью воспроизводит пространственно неоднородную структуру остатков, которая проявляется именно при проверке на независимых точках. Дополнительно это подтверждается тем, что на TEST разброс по STD практически не уменьшается (0,059 $\rightarrow$ 0,060 м).

В гибридной схеме методы машинного обучения используются не как замена LSMSA, а для моделирования остаточной поправки ΔN в зависимости от локального положения точки и значения N_{LSMSA} . При таком подходе физически согласованная основа сохраняется за базовым решением, тогда как регрессионная модель описывает локальные систематические особенности и пространственное распределение ошибок. По результатам оценки на независимой тестовой выборке для SVR (RBF) получено $RMSE = 0,047$ м, а для GPR (ARD) – 0,049 м. По отношению к исходному решению LSMSA наблюдается уменьшение RMSE примерно на 24% для SVR (0,062 $\rightarrow$ 0,047 м) и на 21% для GPR (0,062 $\rightarrow$ 0,049 м). По сравнению с 7-параметрическим преобразованием Гельмерта улучшение также сохраняется: около 19% для SVR (0,058 $\rightarrow$ 0,047 м) и 16% для GPR (0,058 $\rightarrow$ 0,049 м). Изменения затрагивают не только RMSE. Для SVR и GPR на тестовой подвыборке уменьшается и стандартное отклонение ошибок (STD), на 19% и 15% соответственно. Кроме того, значение среднего смещения снижается до сантиметрового уровня. Для GPR величина (MEAN) по сравнению с базовым решением уменьшается приблизительно на 83% (0,024 $\rightarrow$ 0,004 м), а для SVR – на 75% (0,024 $\rightarrow$ 0,006 м). Это указывает на то, что ML-коррекция лучше описывает локальную структуру остаточного поля, чем глобальная 7-параметрическая модель. Однако для LSBoost получены менее однозначные результаты. С одной стороны, по

совокупной выборке значение метрики остается низким ($RMSE = 0,028$ м), а на обучающей подвыборке ошибки практически отсутствуют ($RMSE \approx 0,004$ м). С другой стороны, на тестовой подвыборке составляет около 0,052 м, что хуже результатов SVR и GPR, при заметном расхождении между TRAIN и TEST. Такое соотношение обычно рассматривается как признак переобучения, при котором модель слишком точно воспроизводит обучающие данные, но хуже работает на независимых точках. Поэтому при данном объеме обучающей выборки для LSBoost необходим более детальный контроль сложности модели. Тогда как SVR и GPR показывают более устойчивые результаты при переносе на тестовые данные.

Таким образом, при сохранении физически обоснованной базы LSMSA в рамках выполненного эксперимента наилучшие результаты при сопоставлении с данными ГНСС/Нивелирования показали методы SVR (RBF) и GPR (ARD). Из них минимальное значение RMSE на независимой тестовой подвыборке получено для SVR при близких значениях смещения и разброса.

ЗАКЛЮЧЕНИЕ

В работе для района г. Астана была построена локальная модель геоида на основе гравиметрического решения LSMSA и последующей корректировки остаточного поля по контрольным пунктам ГНСС/Нивелирования. Базовое решение LSMSA обеспечивает физически обоснованное описание поверхности геоида, однако проверка на независимой тестовой подвыборке показала, что без дополнительной коррекции сохраняются как систематическое смещение, так и разброс ошибок. Для варианта Before/None на тестовой выборке получены $RMSE = 0,062$ м и $Mean = -0,024$ м, что указывает на необходимость локального уточнения модели.

Использование 7-параметрического преобразования Гельмерта позволило существенно уменьшить среднее смещение, однако снижение ошибки на независимых точках оказалось сравнительно небольшим. RMSE на TEST уменьшился лишь до 0,058 м, то есть примерно на 6%. Более высокий эффект был достигнут при моделировании остаточной поправки с помощью методов машинного обучения. Для моделей SVR (RBF) и GPR (ARD) значение RMSE на TEST составило 0,047–0,049 м, что соответствует уменьшению ошибки по сравнению с базовым решением на 21–24%. Одновременно сократились разброс ошибок и величина смещения, достигнув сантиметрового уровня.

Для LSBoost, напротив, полученные результаты оказались менее устойчивыми. При очень высокой точности на обучающей подвыборке качество на тестовых данных заметно снижается, что свидетельствует о переобучении в рамках рассматриваемой конфигурации и требует дополнительного ограничения сложности модели. В целом выполненные расчеты показывают, что гибридный подход, объединяющий LSMSA и коррекцию остаточного поля методами машинного обучения, позволяет сохранить физически обоснованную основу гравиметрического решения и одновременно повысить его соответствие данным ГНСС/Нивелирования. Среди рассмотренных вариантов наиболее устойчивый результат на независимой тестовой подвыборке показала модель SVR (RBF).

КОНФЛИКТ ИНТЕРЕСОВ: Авторы заявляют об отсутствии конфликта интересов.

ФИНАНСИРОВАНИЕ: Данное исследование финансировалось Комитетом науки Министерства науки и высшего образования Республики Казахстан, грант № AP26197867.

ДОСТУПНОСТЬ ДАННЫХ: Оригинальные материалы, представленные в этом исследовании, включены в статью. С дополнительными вопросами можно обратиться к автору-корреспонденту.

БЛАГОДАРНОСТИ: Мы выражаем благодарность всем авторам статей, упомянутых в данном обзоре.

УВЕДОМЛЕНИЕ ОБ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: Авторы не применяли инструменты искусственного интеллекта (ИИ) при подготовке данной научной статьи, включая помощь в написании и редактировании текста, проверку фактов и анализ данных.

СПИСОК ЛИТЕРАТУРЫ

- Abbak, R. A., & Ustun, A. (2015). A software package for computing a regional gravimetric geoid model by the KTH method. *Earth science informatics*, 8(1), 255-265. <https://doi.org/10.1007/s12145-014-0149-3>
- Abdalla, A., & Mogren, S. (2015). Implementation of a rigorous least-squares modification of Stokes' formula to compute a gravimetric geoid model over Saudi Arabia (SAGEO13). *Canadian journal of earth sciences*, 52(10), 823-832. <https://doi.org/10.1139/cjes-2014-0192>
- Butt, J., Wieser, A., Gojic, Z., & Zhou, C. (2021). Machine learning and geodesy: A survey. *Journal of Applied Geodesy*, 15(2), 117-133. <https://doi.org/10.1515/jag-2020-0043>
- Chen, F., Zhang, X., Guo, F., Zheng, J., Nan, Y., & Freeshah, M. (2022). TDS-1 GNSS reflectometry wind geophysical model function response to GPS block types. *Geo-Spatial Information Science*, 25(2), 312-324. <https://doi.org/10.1080/10095020.2021.1997076>
- El Shouny, A., Al-Karagy, E. M., Mohamed, H. F., & Dawod, G. M. (2018). Gis-based accuracy assessment of global geopotential models: A case study of Egypt. *American Journal of Geographic Information System*, 7(4), 118-124. <https://doi.org/10.5923/j.ajgis.20180704.03>
- Elsheawy, M. A., Thanh, P. T., Elsheshtawy, A. M., Refaat, M., & Freeshah, M. (2024). A novel approach for optimizing regional geoid modeling over rugged terrains based on global geopotential models and artificial intelligence algorithms. *The Egyptian Journal of Remote Sensing and Space Sciences*, 27(4), 656-668. <https://doi.org/10.1016/j.ejrs.2024.09.002>
- Ghaffari-Razin, S. R., & Hooshangi, N. (2024). Efficiency of machine learning models in estimation of local geoid height with GPS/Leveling measurements. *Scientific-Research Quarterly of Geographical Data (SEPEHR)*, 33(130), 99-117.
- Hu, Z., Zhang, Z., Sang, Y. F., Qian, J., Feng, W., Chen, X., & Zhou, Q. (2021). Temporal and spatial variations in the terrestrial water storage across Central Asia based on multiple satellite datasets and global hydrological models. *Journal of Hydrology*, 596, 126013. <https://doi.org/10.1016/j.jhydrol.2021.126013>
- Information system "State geodetic support" (2025). [cited March 5, 2026]. <https://ggo.gov.kz/geoportal> (In Russ.).
- Işık, M. S., Erol, B., Erol, S., & Sakil, F. F. (2021). High-resolution geoid modeling using least squares modification of Stokes and Hotine formulas in Colorado. *Journal of geodesy*, 95(5), 49. <https://doi.org/10.1007/s00190-021-01501-z>
- Konakoglu, B., & Akar, A. (2021). GEOID UNDULATION PREDICTION USING GAUSSIAN PROCESSES REGRESSION: A CASE STUDY IN A LOCAL REGION IN TURKEY. *Acta Geodynamica et Geomaterialia*, 18(1). <https://doi.org/10.13168/AGG.2021.0001>
- Konakoglu, B., & Akar, A. (2022). Prediction of geoid undulation using approaches based on GMDH, M5 model tree, MARS, GPR, and IDP. *Acta Geodaetica et Geophysica*, 57(2), 293-315. <https://doi.org/10.1007/s40328-022-00378-4>
- Li, H., Zhao, J., Yan, B., Yue, L., & Wang, L. (2022). Global DEMs vary from one to another: an evaluation of newly released Copernicus, NASA and AW3D30 DEM on selected terrains of China using ICESat-2 altimetry data. *International Journal of Digital Earth*, 15(1), 1149-1168. <https://doi.org/10.1080/17538947.2022.2094002>
- Liang, W., Li, J., Xu, X., Zhang, S., & Zhao, Y. (2020). A high-resolution Earth's gravity field model SGG-UGM-2 from GOCE, GRACE, satellite altimetry, and EGM2008. *Engineering*, 6(8), 860-878. <https://doi.org/10.1016/j.eng.2020.05.008>

- Liu, Y., Zhang, Y., Pang, Q., Liu, S., Li, S., Shi, X., ... & Wu, Y. (2024). Gravity predictions in data-missing areas using machine learning methods. *Remote Sensing*, 16(22), 4173. <https://doi.org/10.3390/rs16224173>
- Open geospatial data of Republic of Kazakhstan (2025). [cited March 5, 2026]. <https://map.gov.kz/> (In Russ.).
- Pham, H. T., Awange, J., Claessens, S., & Kuhn, M. (2025). Local hybrid geoid/quasigeoid development using machine learning with consideration of neighbour spatial characteristics. *Survey Review*, 1-33. <https://doi.org/10.1080/00396265.2025.2543647>
- Reguzzoni, M., Carrion, D., De Gaetani, C. I., Albertella, A., Rossi, L., Sona, G., ... & Sansó, F. (2021). Open access to regional geoid models: the International Service for the Geoid. *Earth system science data*, 13(4), 1653-1666. <https://doi.org/10.5194/essd-13-1653-2021>
- Shoganbekova, D., Urazaliyev, A., Nurakynov, S., Kozhakhmetov, M., Zhaksygul, N., & Sermiagin, R. (2025a). Combination of Soviet-Era Surface Gravity and Modern Satellite Data for Geoid Model Computation: A Case Study for Kazakhstan. *Computation*, 13(11), 260. <https://doi.org/10.3390/computation13110260>
- Shoganbekova, D., Urazaliyev, A., Sermiagin, R., Nurakynov, S., Kozhakhmetov, M., Zhaksygul, N., & Islyamova, A. (2025b). Evaluation of a Soviet-Era Gravimetric Survey Using Absolute Gravity Measurements and Global Gravity Models: Toward the First National Geoid of Kazakhstan. *Geosciences*, 15(10), 404. <https://doi.org/10.3390/geosciences15100404>
- Sjöberg, L. E. (2003). A general model for modifying Stokes' formula and its least-squares solution. *Journal of geodesy*, 77(7), 459-464. <https://doi.org/10.1007/s00190-003-0346-1>
- Sjöberg, L. E., & Bagherbandi, M. (2017). *Gravity inversion and integration*. Cham, Switzerland: Springer International Publishing.
- Urazaliyev, A., Shoganbekova, D., Nurakynov, S., Kozhakhmetov, M., Zhaksygul, N., & Sermiagin, R. (2025). Hybrid Geoid Modelling with AI Enhancements: A Case Study for Almaty, Kazakhstan. *Algorithms*, 18(12), 737. <https://doi.org/10.3390/a18120737>
- Zaletnyik, P., Völgyesi, L., & Paláncz, B. (2008). Modelling local GPS/levelling geoid undulations using Support Vector Machines. *Periodica polytechnica civil engineering*, 52(1), 39-43. <https://doi.org/10.3311/pp.ci.2008-1.06>
- Zingerle, P., Pail, R., Gruber, T., & Oikonomidou, X. (2020). The combined global gravity field model XGM2019e: P. Zingerle et al. *Journal of geodesy*, 94(7), 66. <https://doi.org/10.1007/s00190-020-01398-0>

Авторлар туралы мәліметтер

Информация об авторах

Information about authors



Уразалиев Асет Сейсенбекович – «Геодезия» мамандығы бойынша философия докторы (PhD), «Ионосфера институты» ЖШС-нің Спутниктік және геодинамикалық зерттеулер зертханасының меңгерушісі

Уразалиев Асет Сейсенбекович – доктор философии (PhD) по специальности «Геодезия», заведующий лабораторией спутниковых и геодинамических исследований ТОО «Институт ионосферы»

Urazaliyev Asset Seisenbekovich – Doctor of Philosophy (PhD) in the specialty "Geodesy", Head of the Laboratory of Satellite and Geodynamic Research at LLP "Institute of Ionosphere",

e-mail: a.urazaliyev@ionos.kz,

ORCID: <https://orcid.org/0000-0001-7444-2897>



Шоганбекова Дания Асыгатовна – «Геодезия» мамандығы бойынша философия докторы (PhD), «Ионосфера институты» ЖШС-нің бас ғылыми қызметкері, Халықаралық білім беру корпорациясының қауымдастырылған профессоры

Шоганбекова Дания Асыгатовна – доктор философии (PhD) по специальности «Геодезия», главный научный сотрудник ТОО «Институт ионосферы», ассоциированный профессор Международной образовательной корпорации

Shoganbekova Daniya Assygatovna – Doctor of Philosophy (PhD) in Geodesy, Chief Researcher at LLP “Institute of Ionosphere, Associate Professor at the International Educational Corporation,

e-mail: d.shoganbekova@ionos.kz,

ORCID: <https://orcid.org/0000-0002-6825-4774>



Қожахметов Мағжан Серікжанұлы – «Ионосфера институты» ЖШС-нің ғылыми қызметкері, Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университетінің «Геокеңістіктік цифрлық инженерия» мамандығы бойынша PhD докторанты

Қожахметов Мағжан Серікжанұлы – научный сотрудник ТОО «Институт ионосферы», PhD-докторант по специальности «Геопространственная цифровая инженерия» Казахского Национального Исследовательского Технического университета им. К.И. Сатпаева

Kozhakhmetov Magzhan Serikzhanuly – researcher at LLP “Institute of Ionosphere”, PhD doctoral student in Geospatial Digital Engineering at the Kazakh National Research Technical University named after K.I. Satpayev,

e-mail: m.kozhakhmetov@ionos.kz,

ORCID: <https://orcid.org/0009-0004-9433-8674>



Жақсығұл Найля Нұрланқызы – «Ионосфера институты» ЖШС-нің ғылыми қызметкері, Халықаралық білім беру корпорациясының «Геодезия» мамандығы бойынша PhD докторанты

Жақсығұл Найля Нұрланқызы – научный сотрудник ТОО «Институт ионосферы», PhD-докторант по специальности «Геодезия» Международной образовательной корпорации

Zhaksygyul Nailya Nurlankyzy – researcher at LLP “Institute of Ionosphere”, PhD doctoral student in Geodesy at the International Educational Corporation,

e-mail: n.zhaksygyul@ionos.kz,

ORCID: <https://orcid.org/0009-0005-3597-5056>