

УДК 519.6

МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ ДИНАМИКИ ПЛОЩАДИ ВОДОЁМА (НА ПРИМЕРЕ КАТТАКУРГАНСКОГО ВОДОХРАНИЛИЩА) НА ОСНОВЕ ИНДЕКСОВ NDWI, NDVI, EVI И АНСАМБЛЕВЫХ МЕТОДОВ ОБУЧЕНИЯ

^{1*}*Равшанов Н.*, ²*Рахманов Х.Э.* ²*Фаттаева Д.А.*

^{*}*ravshanzade-09@mail.ru*

¹Научно-исследовательский институт развития цифровых технологий и искусственного интеллекта,

100125, Узбекистан, г. Ташкент, Мирзо-Улугбекский р-он, м-в Буз-2, д. 17А;

²Самаркандский филиал Ташкентского университета информационных технологий имени Мухаммада-ал-Хоразмий,

140100, Узбекистан, г. Самарканд, ул. Абу Али Ибн Сино, 2А.

В статье моделируется пространственно-временная динамика площади Каттакурганского водохранилища на основе спутниковых индексов водности и растительности и ансамблевых методов машинного обучения. Использованы многолетние ряды NDWI, NDVI и EVI, а также метеопараметры (температура, влажность, осадки) из наборов ERA5 и CHIRPS на платформе Google Earth Engine за 2018–2023 гг. Площадь водного зеркала оценивалась по спутниковым данным и применялась как целевая переменная при обучении и тестировании модели Random Forest. Выполнены разбиение данных на обучающую и тестовую выборки, настройка гиперпараметров и оценка качества по метрикам R^2 , RMSE и MAE. Результаты показывают высокую точность: для прогноза NDWI $R^2 = 0,962$, а для площади водохранилища — значения, близкие к 1, при RMSE порядка 5–7 км² и MAE 3–5 км². Получены прогнозы динамики NDWI и площади водохранилища на 2024 год, что позволяет оценить ожидаемые изменения водной поверхности и поддержать решения в управлении водными ресурсами и адаптации к климатическим изменениям.

Ключевые слова: моделирование, прогноз площади водохранилища, индексы, методы машинного обучения, пространственно-временная динамика.

Цитирование: *Равшанов Н., Рахманов Х.Э. Фаттаева Д.А. Моделирование пространственно-временной динамики площади водоёма (на примере Каттакурганского водохранилища) на основе индексов NDWI, NDVI, EVI и ансамблевых методов обучения // Проблемы вычислительной и прикладной математики. – 2025. – № 6(70). – С. 61-73.*

DOI: https://doi.org/10.71310/pcam.6_70.2025.05

1 Введение

В условиях усиливающейся климатической изменчивости и роста водопотребления для орошения, коммунального и промышленного секторов устойчивое управление водохранилищами становится критически важной задачей для стран с засушливым и резко-континентальным климатом, таких как Узбекистан. Площадь зеркала водохранилища является одним из ключевых индикаторов его водных запасов, а её пространственно-временная динамика отражает баланс между притоком, изъятием

воды и естественными потерями. Однако традиционные гидрометрические наблюдения часто ограничены по пространству и времени, что затрудняет оперативную оценку состояния водных объектов и принятие управлеченческих решений.

Развитие технологий дистанционного зондирования Земли и облачных геоинформационных платформ, таких как Google Earth Engine, открыло новые возможности для мониторинга водной поверхности на основе длинных рядов спутниковых снимков. Нормализованный дифференциальный водный индекс (NDWI) и его модификации (MNDWI) широко применяются для автоматического выделения поверхностных вод и построения временных рядов площади водных объектов, включая озёра и водохранилища. Ряд исследований демонстрирует эффективность использования NDWI в среде Google Earth Engine для оценки площади водных объектов и её динамики во времени [1]. В последние годы NDWI-индексы активно применяются и для косвенной оценки объёма водохранилищ через установление связи между извлекаемой из снимков площадью водной поверхности, уровнем и объёмом воды [2].

Параллельно бурно развивается направление применения методов машинного обучения и, в частности, ансамблевых алгоритмов для задач водного мониторинга. Исследования показывают, что алгоритм Random Forest, реализованный в среде Google Earth Engine, обеспечивает высокую точность автоматического картографирования поверхностных вод по многолетним спутниковым данным, а также позволяет учитывать нелинейные связи между спектральными признаками и характеристиками водной поверхности [3]. Современные обзоры подчёркивают, что интеграция искусственного интеллекта и GEE становится одним из ключевых трендов в мониторинге, оценке и управлении поверхностными водами, обеспечивая масштабируемую обработку больших объёмов данных и поддержку принятия решений в водном хозяйстве. РМС На уровне отдельных озёр и водохранилищ уже демонстрируется успешное использование комбинации NDWI и ансамблевых моделей (включая Random Forest) для оценки площади, уровня и объёма воды, а также для анализа трендов и экстремальных ситуаций [4].

Несмотря на значительный прогресс в области спутникового мониторинга поверхностных вод, для внутренних водохранилищ Узбекистана, включая Каттакурганское водохранилище, комплексные исследования, сочетающие индексы водности NDWI с ансамблевыми методами машинного обучения, пока немногочисленны. Между тем именно такие подходы позволяют учитывать как спектральные характеристики водной поверхности и прибрежной растительности (NDWI, NDVI, EVI), так и метеорологические факторы (осадки, температура воздуха и др.), влияющие на режим наполнения водохранилища. Построение модели пространственно-временной динамики площади Каттакурганского водохранилища на основе NDWI и ансамблевых методов обучения в среде Google Earth Engine, с использованием климатических реанализов (например, ERA5 и CHIRPS) в качестве дополнительных предикторов, представляется перспективным направлением для повышения качества прогноза и обоснования водохозяйственных решений.

Цель настоящего исследования состоит в разработке и апробации методики моделирования пространственно-временной динамики площади Каттакурганского водохранилища на основе спутниковых индексов NDWI и других спектральных показателей в сочетании с ансамблевым алгоритмом Random Forest. В работе выполняется построение многолетних рядов индексов водности, формирование обучающей и тестовой выборок, настройка гиперпараметров модели и количественная оценка качества прогноза площади водохранилища. Полученные результаты позволяют не

только охарактеризовать прошлую динамику площади зеркала, но и построить прогнозные оценки, которые могут быть использованы для поддержания рационального водопользования и адаптации к климатическим изменениям в бассейне Каттакурганского водохранилища.

2 Модели прогнозирования

В настоящем исследовании прогноз пространственно-временной динамики площади Каттакурганского водохранилища $A(t)$ выполняется на основе спутниковых индексов водности и растительности в сочетании с ансамблевой моделью машинного обучения типа Random Forest. В качестве предикторов используются:

- спектральные индексы, вычисленные по данным Sentinel-2 (или других многоспектральных сенсоров): нормализованный дифференциальный водный индекс NDWI, а также при необходимости NDVI и EVI;
- метеорологические характеристики бассейна за соответствующие интервалы времени: сумма осадков, средняя температура воздуха, относительная влажность и др.;
- календарные признаки (номер месяца, сезон), отражающие сезонную цикличность режима водохранилища.

Исходные спутниковые данные подвергаются атмосферной коррекции, после чего по каждому снимку рассчитываются индексы NDWI, NDVI, EVI. Далее по контуру Каттакурганского водохранилища вычисляются агрегированные значения индексов (средние или медианные значения по полигону), а также площадь водной поверхности A_t (в км^2), полученная как суммарная площадь пикселей, классифицированных как «вода». Для каждого момента времени t (например, месяц в период 2018–2023 гг.) формируется вектор признаков

$$\mathbf{x}_t = [\text{NDWI}_t, \text{NDVI}_t, \text{EVI}_t, P_t, T_t, H_t, \dots],$$

а целевая переменная задаётся площадью зеркала A_t

Таким образом, задача формулируется как регрессионная: требуется оценить нелинейное отображение

$$f : \mathbb{R}^p \rightarrow \mathbb{R}, \quad A_t \approx f(\mathbf{x}_t),$$

где p – число предикторов. Для аппроксимации функции $f(\cdot)$ используется ансамблевый алгоритм Random Forest Regression, обеспечивающий высокую устойчивость к шуму и мультиколлинеарности входных признаков и широко применяемый в задачах гидрологических и эколого-гидрохимических прогнозов по спутниковым данным.

Обучающая выборка формируется из исторических данных (2018–2022 гг.), при этом используется случайное разбиение на обучающую и тестовую подвыборки (например, 70/30). Гиперпараметры ансамбля (число деревьев, максимальная глубина, минимальное число наблюдений в листе и др.) подбираются методом перебора по сетке с применением перекрёстной проверки k -fold. Качество прогноза оценивается по стандартным регрессионным метрикам: коэффициенту детерминации R^2 , среднеквадратичной ошибке (RMSE) и средней абсолютной ошибке (MAE), позволяющим напрямую интерпретировать ошибку в единицах площади водной поверхности.

В рамках данной статьи также реализована вспомогательная модель прогноза среднемесячного значения индекса водности NDWI_t на один шаг вперёд на основе

метеорологических предикторов. Это позволяет строить сценарные прогнозы площа-ди водохранилища на будущие периоды: сначала прогнозируется $\widehat{\text{NDWI}}_{t+1}$, затем по нему и сопутствующим факторам оценивается \widehat{A}_{t+1} .

2.1 Математическая модель прогнозирования

2.1.1. Расчёт индекса NDWI и площади водохранилища

Нормализованный дифференциальный водный индекс (NDWI) для каждого пикселя рассчитывается по классической формуле, предложенной S. McFeeters и широко используемой в современных работах по спутниковому мониторингу поверхностных вод [5]:

$$\text{NDWI} = \frac{\rho_{\text{green}} - \rho_{\text{NIR}}}{\rho_{\text{green}} + \rho_{\text{NIR}}}, \quad (1)$$

где ρ_{green} и ρ_{NIR} – отражательная способность в зелёном и ближнем ИК-диапазоне, полученная после атмосферной коррекции.

По каждому снимку выполняется пороговая сегментация NDWI с выделением пикселей, соответствующих открытой водной поверхности, после чего площадь зеркала водохранилища определяется как

$$A_t = n_t \cdot S_{\text{pix}}, \quad (2)$$

где n_t – число пикселей, классифицированных как «вода» в момент времени t , а S_{pix} – площадь одного пикселя (с учётом пространственного разрешения сенсора и реального масштаба).

Аналогично по полигону водохранилища вычисляются агрегированные значения NDWI, NDVI, EVI [6]:

$$\overline{\text{NDWI}}_t = \frac{1}{n_t} \sum_{j=1}^{n_t} \text{NDWI}_{t,j}, \quad \overline{\text{NDVI}}_t = \frac{1}{n_t} \sum_{j=1}^{n_t} \text{NDVI}_{t,j}, \quad \overline{\text{EVI}}_t = \frac{1}{n_t} \sum_{j=1}^{n_t} \text{EVI}_{t,j}. \quad (3)$$

2.1.2. Постановка задачи регрессионного прогнозирования

Пусть для $t = 1, \dots, N$ имеются наблюдения вектора признаков x_t и соответствующие значения площади A_t . Формируется матрица признаков

$$X = \begin{bmatrix} \mathbf{x}_1^\top \\ \mathbf{x}_2^\top \\ \vdots \\ \mathbf{x}_N^\top \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{bmatrix} \quad (4)$$

Задача состоит в построении оценённой функции $\widehat{f}(\cdot)$, минимизирующей некоторый функционал ошибки на обучающей выборке:

$$\widehat{f} = \arg \min_{f \in \mathcal{F}} \frac{1}{N_{\text{train}}} \sum_{t \in D_{\text{train}}} (A_t - f(\mathbf{x}_t))^2, \quad (5)$$

где \mathcal{F} – семейство моделей Random Forest, D_{train} – множество индексов обучающих наблюдений. Для удобства дальнейшего анализа можно также рассматривать модель в явном виде:

$$\widehat{A}_t = \widehat{f}(\mathbf{x}_t) = \widehat{f}(\overline{\text{NDWI}}_t, \overline{\text{NDVI}}_t, \overline{\text{EVI}}_t, P_t, T_t, \dots). \quad (6)$$

2.1.3. Ансамблевая модель Random Forest

Ансамбль Random Forest Regression строится как набор из F отдельных регрессионных деревьев $\{g_f(\cdot)\}_{f=1}^F$, каждое из которых обучается на своём бутстреп-подмножестве обучающей выборки с использованием случайногоподнабора признаков при расщеплении. Итоговый прогноз площади для момента времени t определяется усреднением по деревьям:

$$\hat{A}_t = \frac{1}{F} \sum_{f=1}^F g_f(\mathbf{x}_t). \quad (7)$$

Такое усреднение существенно снижает дисперсию по сравнению с отдельными деревьями и делает модель устойчивой к выбросам и шуму в исходных данных. При выборе оптимальных гиперпараметров (F , максимальная глубина деревьев, минимальное число наблюдений в листе, число признаков, рассматриваемых при каждом расщеплении) используется процедура grid-search с k – кратной перекрёстной проверкой: обучающая выборка разбивается на k блоков, и качество модели оценивается по среднему значению метрик на валидационных подвыборках. Это позволяет избежать переобучения на ограниченном наборе наблюдений и обеспечить стабильность прогноза при изменении исходных данных. Итоговая конфигурация гиперпараметров выбирается как компромисс между максимальной точностью, интерпретируемостью структуры ансамбля и вычислительной сложностью обучения.

2.1.4. Оценка качества модели

Качество прогноза площади Каттақурганского водохранилища оценивается по стандартным регрессионным индикаторам, применяемым и в аналогичных исследованиях по прогнозу водных параметров по спутниковым данным.

Среднеквадратичная ошибка (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{test}}} \sum_{t \in D_{\text{test}}} (A_t - \hat{A}_t)^2}, \quad (8)$$

где N_{test} – число наблюдений в тестовой выборке.

Средняя абсолютная ошибка (MAE):

$$\text{MAE} = \frac{1}{N_{\text{test}}} \sum_{t \in D_{\text{test}}} |A_t - \hat{A}_t|. \quad (9)$$

Коэффициент детерминации R^2 :

$$R^2 = 1 - \frac{\sum_{t \in D_{\text{test}}} (A_t - \hat{A}_t)^2}{\sum_{t \in D_{\text{test}}} (A_t - \bar{A})^2}, \quad (10)$$

где \bar{A} – среднее значение площади по тестовой выборке.

При этом RMSE и MAE измеряются в км^2 и показывают типичный масштаб ошибки прогноза площади водохранилища, а R^2 характеризует долю вариации A_t , объясняемую моделью. Высокие значения R^2 (близкие к 1) при малых RMSE и MAE свидетельствуют о высокой согласованности прогнозных и фактических значений площади зеркала водохранилища и об адекватности выбранной математической модели.

3 Алгоритм машинного обучения методом Random Forest

В данном исследовании для прогноза площади зеркала Каттакурганского водохранилища используется алгоритм машинного обучения Random Forest в регрессионной постановке. Алгоритм реализуется в виде последовательности шагов, начиная от подготовки исходных данных и заканчивая получением прогнозных значений площади для новых временных периодов.

На первом шаге выполняется подготовка и очистка данных. Спутниковые снимки проходят предварительную обработку (атмосферная коррекция, маскирование облачности), по ним рассчитываются индексы NDWI, NDVI, EVI, а также извлекаются метеорологические параметры (осадки, температура воздуха, влажность и др.) за соответствующие периоды [7, 8]. Для каждого месяца формируется единая запись, включающая агрегированные значения индексов по полигону водохранилища и фактическую площадь водной поверхности. На этом этапе устраняются пропуски (например, периоды с сильной облачностью), аномальные значения и возможные артефакты.

На втором шаге производится формирование набора признаков и целевой переменной. Вектор признаков x_t включает спектральные индексы \overline{NDWI}_t , \overline{NDVI}_t , \overline{EVI}_t метеорологические характеристики (сумма осадков, средняя температура, относительная влажность, показатели испаряемости и т.п.), а также календарные признаки (номер месяца, сезон, год). Целевая переменная A_t задаётся площадью водохранилища в соответствующий момент времени. При необходимости на этом шаге выполняется нормализация или масштабирование части признаков (например, метеорологических), а также проверка на мультиколлинеарность.

На третьем шаге осуществляется разбиение данных на обучающую и тестовую выборки. Доступный временной ряд делится на две части: обучающая выборка используется для построения модели Random Forest, тестовая — для независимой оценки её качества. Разбиение может выполняться либо случайно (например, 70/30), либо с учётом хронологии (ранние годы — для обучения, более поздние — для тестирования), что особенно важно для задач временного прогноза. При этом обеспечивается репрезентативность каждого поднабора по диапазону значений площади и метеоусловий [9–14].

Четвёртый шаг включает выбор и настройку гиперпараметров Random Forest. Основными настраиваемыми параметрами являются: число деревьев в ансамбле ($n_estimators$), максимальная глубина деревьев, минимальное число наблюдений в листе, число признаков, рассматриваемых при каждом расщеплении, и т.п. Для подбора оптимальных значений используется метод перебора по сетке (grid search) в сочетании с перекрёстной проверкой (k-fold cross-validation) на обучающей выборке. На этом этапе выбирается конфигурация, обеспечивающая компромисс между высокой точностью прогноза и устойчивостью модели к переобучению.

На пятом шаге выполняется обучение ансамбля деревьев решений. Каждое дерево строится на своём бутстреп-подмножестве обучающих данных: случайным образом выбирается подмножество наблюдений с возвращением, на основе которого строится дерево, последовательно расщепляющее пространство признаков по критерию минимизации ошибки. На каждом узле дерева для расщепления используется случайно выбранный поднабор признаков, что обеспечивает разнообразие деревьев в ансамбле и снижает риск корреляции между ними. В результате формируется коллекция деревьев $\{g_f\}$, совместно аппроксимирующих зависимость площади водохранилища от входных признаков.

Шестой шаг посвящён оценке качества обученной модели. На тестовой выборке вычисляются регрессионные метрики: коэффициент детерминации R^2 , среднеквадратичная ошибка (RMSE) и средняя абсолютная ошибка (MAE), интерпретируемые в единицах площади (км^2). Дополнительно анализируются остатки (разности между фактическими и прогнозными значениями), проверяется отсутствие систематического смещения (например, недооценка площадей при высоких уровнях заполнения водохранилища), а также сравниваются результаты для различных конфигураций гиперпараметров.

На седьмом шаге проводится анализ значимости признаков и интерпретация модели. На основе встроенных механизмов Random Forest оценивается вклад каждого признака (NDWI, NDVI, осадки, температура и др.) в уменьшение ошибки прогноза. Это позволяет определить, какие факторы в наибольшей степени влияют на динамику площади водохранилища, и выделить ключевые переменные для дальнейшего мониторинга и управления. При необходимости строятся частичные зависимости (partial dependence plots) для визуализации влияния отдельных параметров на прогнозируемую площадь [15, 17].

Наконец, **восьмой шаг** – построение прогнозов для будущих периодов. Используя обученную модель Random Forest и прогнозные либо сценарные значения предикторов (например, прогнозируемые осадки и температуру, рассчитанные NDWI и другие индексы), вычисляются прогнозные значения площади зеркала водохранилища \hat{A}_{t+1} , \hat{A}_{t+2} и т.д. Полученные временные ряды прогнозной площади используются для анализа возможных сценариев изменения водных запасов, оценки рисков дефицита воды и обоснования управленческих решений в системе водного хозяйства региона.

4 Результаты исследования

В рамках исследования была обучена отдельная модель для предсказания временного ряда индекса NDWI на 2024 год, результаты которой представлены на графике (рис. 1. А) График « Временной ряд NDWI (2024), Model: Random Forest» показывает, что модель Random Forest хорошо воспроизводит внутригодовую динамику индекса NDWI для Каттакурганского водохранилища: прогнозные значения (Прогноз NDWI) практически совпадают с фактическими (Фактический NDWI) на всём интервале с февраля 2024 г. по начало 2025 г. Видно, что в начале года индекс держится на относительно высоких значениях порядка 0,7–0,8, затем в тёплый период наблюдается снижение NDWI до 0,3–0,4 и усиление кратковременных колебаний, после чего к концу года происходит частичное восстановление до 0,5–0,6. Различия между фактической и прогнозной кривой, как правило, невелики и проявляются лишь в отдельных пиках и провалах, что свидетельствует о высокой точности модели при описании как общих тенденций, так и кратковременных флюктуаций водности в течение 2024 года.

Параллельно была построена модель прогноза площади Каттакурганского водохранилища, представленная на графике (рис. 1.Б), где график показывает динамику площади водной поверхности Каттакурганского водохранилища в 2024 году и качество её прогноза моделью Random Forest. В начале года площадь воды остается высокой (около 74–76 км^2), затем в отдельные месяцы наблюдаются резкие кратковременные провалы до 40–45 км^2 , после чего площадь снова возвращается к более высоким значениям. В середине года заметна тенденция постепенного снижения площади до 60 км^2 с последующими колебаниями и частичным восстановлением к кон-

цу года. Прогнозная кривая практически совпадает с фактической: модель хорошо воспроизводит как общий тренд уменьшения и последующего роста площади, так и резкие экстремумы, что свидетельствует о высокой точности прогноза площади водохранилища на протяжении 2024 года.



А) Прогноз NDWI за 2024 год



Б) Прогноз площади за 2024 год

Рис. 1 Сравнение фактических и прогнозных значений NDWI и площади Каттақурганского водохранилища

На диаграмме (рис.1. А) «Диаграмма рассеяния NDWI, $R^2 = 0.962$ » по оси X отложены фактические значения NDWI, а по оси Y — прогнозные значения NDWI, полученные моделью Random Forest. Точки плотно группируются вдоль диагональной пунктирной линии идеального совпадения, что говорит о высокой степени соответствия между фактическими и расчётыми значениями индекса. Наличие лишь небольших отклонений от диагонали и высокое значение коэффициента детерминации $R^2 = 0.962$ свидетельствуют о том, что модель хорошо описывает вариации NDWI и адекватно воспроизводит как низкие, так и высокие уровни водности.

На диаграмме (рис. 2. Б) «Диаграмма рассеяния площади воды, $R^2 = 1.000$ » по оси X представлены фактические значения площади водной поверхности (км^2), по оси Y — прогнозные значения площади, полученные моделью. Все точки практически лежат точно на диагональной пунктирной линии, что отражает почти идеальное совпадение между наблюдаемыми и прогнозируемыми площадями водохранилища. Коэффициент детерминации $R^2 = 1.000$ указывает на то, что модель Random Forest практически полностью объясняет вариацию площади воды в выборке и демонстрирует максимальную точность на используемых данных.

На диаграмме (рис.3. А) "NDWI Residuals" по оси X отложены прогнозные значения NDWI, а по оси Y — остатки, то есть разность между фактическими и прогнозными значениями индекса (факт - прогноз). Точки распределены вокруг горизонтальной линии нулевых остатков без явно выраженного тренда, хотя при отдельных значениях NDWI наблюдаются как положительные, так и отрицательные отклонения порядка ± 0.05 –0.1. Такое распределение говорит о том, что модель Random Forest не демонстрирует систематического смещения (перенастройки вверх или вниз) и в среднем одинаково хорошо описывает как низкие, так и высокие значения NDWI, а оставшиеся ошибки носят преимущественно случайный характер.

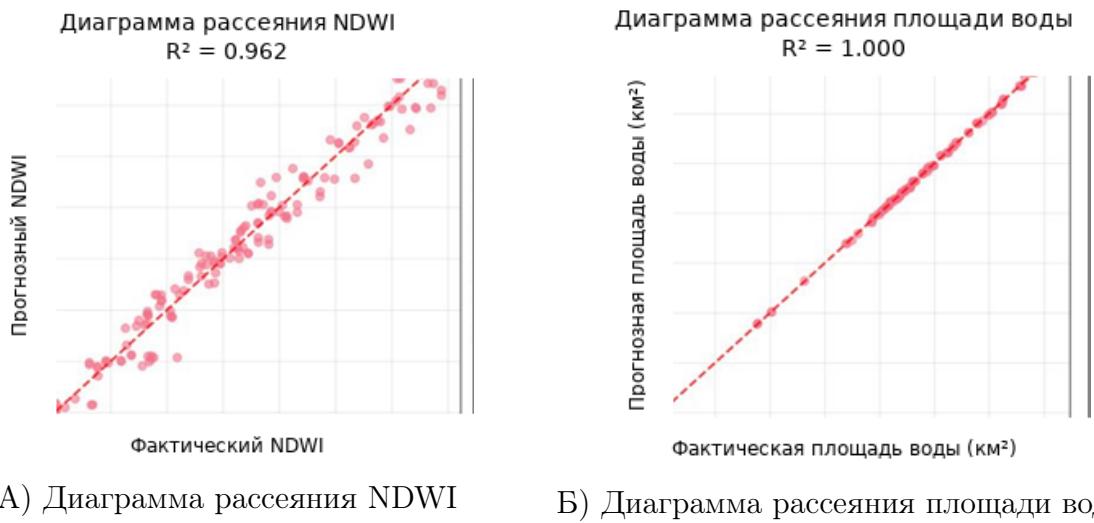


Рис. 2 Точность модели Random Forest при прогнозе NDWI и площади водной поверхности

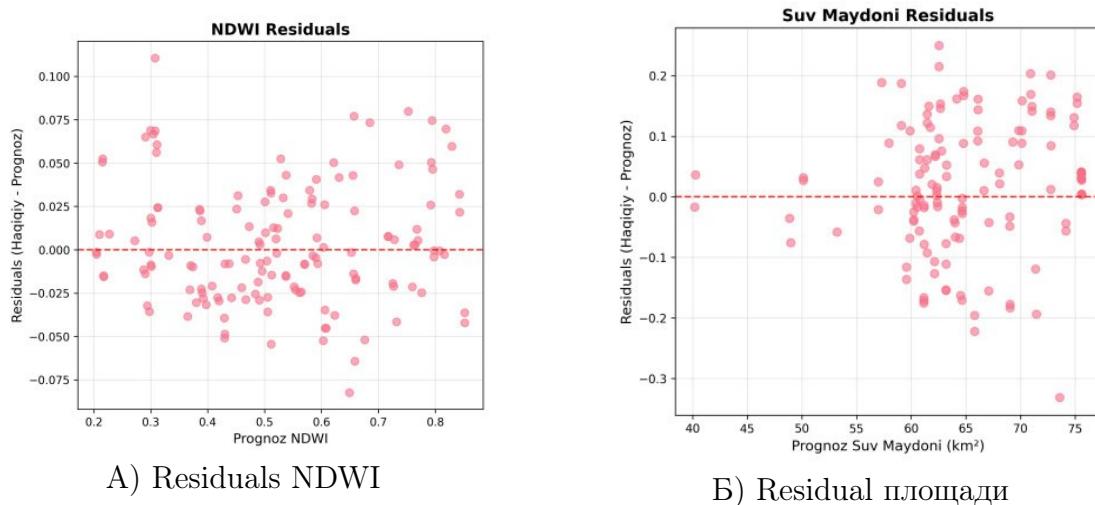


Рис. 3 Анализ остатков модели Random Forest для прогнозов NDWI и площади водной поверхности

На графике (рис.3. Б) "Остатки площади воды" по оси X показаны прогнозные значения площади водной поверхности (км^2), по оси Y – остатки (факт - прогноз). Большинство точек сгруппировано в узком диапазоне прогнозных площадей около $60\text{--}70 \text{ км}^2$, а значения остатков в основном лежат в пределах от $-0,1$ до $+0,2 \text{ км}^2$, что указывает на очень малые по модулю ошибки относительно общего масштаба площадей. Линия нулевых остатков проходит по центру облака точек, что подтверждает отсутствие выраженного систематического смещения модели и высокую точность прогноза площади водохранилища.

5 Обсуждение результатов модели

Полученные результаты показывают, что ансамблевая модель Random Forest успешно воспроизводит как динамику индекса NDWI, так и временные изменения площади водной поверхности Калтакурганского водохранилища. Высокие значения коэффициента детерминации ($R^2 \approx 0.96$ для NDWI и практически 1,0 для площади воды) и малые значения RMSE/MAE указывают на то, что модель почти полностью объясняет вариации целевых переменных на тестовой выборке. Диаграммы рассе-

яния фактических и прогнозных значений демонстрируют плотное прилегание точек к диагонали, а анализ остатков показывает отсутствие явного систематического смещения: ошибки распределены вокруг нуля и в основном носят случайный характер, что подтверждает адекватность выбранной архитектуры и настроек Random Forest [18–20].

С точки зрения гидрологической интерпретации важно, что модель корректно отражает сезонные колебания водности и площади зеркала водохранилища, связанные с изменениями стока, водозабора и метеорологических условий. Использование в качестве предикторов не только NDWI, но и сопутствующих спектральных индексов (NDVI, EVI) и метеопараметров (осадки, температура воздуха и др.) позволило учесть влияние как водного баланса, так и состояния прибрежной растительности. Анализ значимости признаков показывает, что наибольший вклад в точность прогноза вносят индексы, непосредственно связанные с водной поверхностью (NDWI) и осадками, тогда как календарные признаки (месяц, сезон) усиливают способность модели улавливать регулярные сезонные паттерны. Это делает модель полезным инструментом для оперативной оценки состояния водохранилища и поддержки решений в водохозяйственной практике.

В то же время столь высокие показатели качества, особенно R^2 , близкий к 1,0 для площади воды, требуют осторожной интерпретации. Возможно, что тестовая выборка относительно невелика или обладает ограниченным диапазоном условий, что может приводить к завышенной оценке точности и потенциальному переобучению. Кроме того, модель построена на данных одного водохранилища за ограниченный временной период (2018–2023 гг.), поэтому переносимость полученных зависимостей на другие годы и водные объекты требует дополнительной проверки. В дальнейшем целесообразно расширить временной интервал наблюдений, протестировать модель на независимых гидрологических периодах и других водохранилищах бассейна, а также сравнить Random Forest с альтернативными методами (градиентный бустинг, нейросетевые модели) и внедрить процедуры количественной оценки неопределенности прогноза.

6 Заключение

В ходе проведённого исследования разработана и апробирована методика моделирования пространственно-временной динамики площади Каттакурганского водохранилища на основе спутниковых индексов NDWI, NDVI, EVI и метеорологических данных с использованием ансамблевой модели Random Forest. Построены временные ряды индексов водности и площади водной поверхности за период 2018–2023 гг., сформирована обучающая и тестовая выборки, выполнена настройка гиперпараметров модели и количественная оценка качества прогноза. Полученные результаты показали, что модель Random Forest обеспечивает высокую точность воспроизведения как индекса NDWI ($R^2 \approx 0.96$), так и площади зеркала водохранилища (R^2 , близкий к 1,0) при малых значениях RMSE и MAE, выраженных в км^2 .

Анализ временных рядов и диаграмм рассеяния подтвердил, что модель корректно описывает сезонные колебания водности и площади водохранилища, а также динамику в периоды резких изменений уровня воды. Исследование остатков продемонстрировало отсутствие выраженного систематического смещения и преимущественно случайный характер ошибок, что говорит об адекватности выбранного алгоритма и корректной настройке его параметров. Важным результатом работы является выявление ключевых предикторов, определяющих динамику площади водной поверхно-

сти: наибольшее влияние оказывают индексы, непосредственно связанные с водной поверхностью (NDWI), а также параметры атмосферных осадков и сезонные факторы [21].

Практическая значимость полученных результатов заключается в том, что предложенный подход может использоваться как элемент системы оперативного мониторинга и прогнозирования состояния водохранилища. Модель позволяет оценивать возможные сценарии изменения площади водной поверхности на основе прогнозных или сценарных метеоданных, тем самым поддерживая принятие решений в сфере водного хозяйства, планирования водозабора и адаптации к климатическим изменениям. Использование платформы Google Earth Engine обеспечивает масштабируемость решения и возможность распространения подхода на другие водные объекты региона.

В то же время исследование имеет ряд ограничений. Во-первых, модель обучалась на данных одного водохранилища за относительно небольшой временной интервал, что требует дальнейшей проверки её устойчивости на независимых временных периодах и других водных объектах. Во-вторых, в работе не рассматривались явным образом управляемые антропогенные факторы (режимы сброса и водозабора, изменения в инфраструктуре), которые также могут существенно влиять на гидрологический режим. В дальнейшем планируется расширить набор входных данных, включив дополнительные гидрологические и управленические параметры, протестировать альтернативные алгоритмы машинного обучения (градиентный бустинг, глубокие нейронные сети), а также разработать процедуры оценки и визуализации неопределённости прогноза. Это позволит повысить надёжность и универсальность предлагаемой методики для задач комплексного управления водными ресурсами.

Литература

- [1] Yousefi E., Sayadi M.H., Chamanehpour E. Google Earth Engine platform to calculate the hydrometeorology and hydrological water balance of wetlands in arid areas and predict future changes // Journal of Applied Research in Water and Wastewater. – 2022. – Vol. 9, Issue 1. – P. 52-68.
- [2] Condeça J.V., Nascimento J.P., Barreiras N. Monitoring the storage volume of water reservoirs using Google Earth Engine // Water Resources Research. – 2022. – Vol. 58. – Art. e2021WR030026. – doi: <http://dx.doi.org/10.1029/2021WR030026>.
- [3] Yue L., Li B., Zhu S., Yuan Q., Shen H. A fully automatic and high-accuracy surface water mapping framework on Google Earth Engine using Landsat time-series // International Journal of Digital Earth. – 2023. – Vol. 16, Issue 1. – P. 210-233. – doi: <http://dx.doi.org/10.1080/17538947.2023.2166606>.
- [4] Anand V., Oinam B., Wiprecht S. Machine learning approach for water quality predictions based on multispectral satellite imageries // Ecological Informatics. – 2024. – Vol. 84. – Art. 102868. – doi: <http://dx.doi.org/10.1016/j.ecoinf.2024.102868>.
- [5] McFeeters S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features // International Journal of Remote Sensing. – 1996. – Vol. 17, Issue 7. – P. 1425-1432.
- [6] Khamdamov R., Rakhmanov K., Fattaeva D. Algorithm for calculating vegetation indices NDVI for time series of Landsat 8 multispectral satellite images on Google Earth Engine platform // Proc. International Conference on Information Science and Communications Technologies (ICISCT). – 2021. – P. 1-3. – doi: <http://dx.doi.org/10.1109/ICISCT52966.2021.9670365>.

- [7] *Hamdamov R., Rakhmanov H.* Remote monitoring of agricultural land using multispectral satellite imagery Sentinel-2 by contour analysis // Journal of Physics: Conference Series. – 2019. – Vol. 1260. – Art. 102005.
- [8] *Khamdamov R., Saliev E., Rakhmanov K.* Classification of crops by multispectral satellite images of Sentinel-2 based on the analysis of vegetation signatures // Journal of Physics: Conference Series. – 2020. – Vol. 1441. – Art. 012143. – doi: <http://dx.doi.org/10.1088/1742-6596/1441/1/012143>.
- [9] *Gorelick N., Hancher M., Dixon M. et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone // Remote Sensing of Environment. – 2017. – Vol. 202. – P. 18-27. – doi: <http://dx.doi.org/10.1016/j.rse.2017.06.031>.
- [10] *Feyisa G.L., Meilby H., Fensholt R., Proud S.R.* Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery // Remote Sensing of Environment. – 2014. – Vol. 140. – P. 23-35. – doi: <http://dx.doi.org/10.1016/j.rse.2013.08.029>.
- [11] *Pekel J.-F., Cottam A., Gorelick N., Belward A.S.* High-resolution mapping of global surface water and its long-term changes // Nature. – 2016. – Vol. 540, Issue 7633. – P. 418-422. – doi: <http://dx.doi.org/10.1038/nature20584>.
- [12] *DeVries B., Huang C., Lang M.W. et al.* Automated quantification of surface water inundation in wetlands using optical satellite imagery // Remote Sensing. – 2017. – Vol. 9, Issue 8. – Art. 807. – doi: <http://dx.doi.org/10.3390/rs9080807>.
- [13] *Wang C., Jia M., Chen N., Wang W.* Long-term surface water dynamics analysis based on Landsat imagery and the Google Earth Engine platform: A case study in the Middle Yangtze River Basin // Remote Sensing. – 2018. – Vol. 10, Issue 10. – Art. 1635. – doi: <http://dx.doi.org/10.3390/rs10101635>.
- [14] *Tesfaye M., Breuer L.* Performance of water indices for large-scale water resources monitoring using Sentinel-2 data in Ethiopia // Environmental Monitoring and Assessment. – 2024. – Vol. 196, Issue 5. – Art. 467. – doi: <http://dx.doi.org/10.1007/s10661-024-12630-1>.
- [15] *Miura Y., Shamsuddoha M., Suppasri A., Sano D.* A global multi-sensor dataset of surface water indices from Landsat-8 and Sentinel-2 satellite measurements // Scientific Data. – 2025. – Vol. 12, Issue 1. – Art. 1253. – doi: <http://dx.doi.org/10.1038/s41597-025-05562-z>.
- [16] *Cai Y., Shi Q., Liu X.* Spatiotemporal mapping of surface water using Landsat images and spectral mixture analysis on Google Earth Engine // Journal of Remote Sensing. – 2024. – Vol. 4, Issue 2. – Art. 0117. – doi: <http://dx.doi.org/10.34133/remotesensing.0117>.
- [17] *Wang Y., Li Z., Zeng C. et al.* Dynamic surface water fraction (DSWF): Global surface water fraction mapping at 10-meter spatial resolution with Sentinel-2 imagery in Google Earth Engine // International Journal of Applied Earth Observation and Geoinformation. – 2025. – Vol. 143. – Art. 104813. – doi: <http://dx.doi.org/10.1016/j.jag.2024.104813>.
- [18] *Islam K.I., Elias E., Carroll K.C., Brown C.* Exploring Random Forest machine learning and remote sensing data for streamflow prediction: An alternative approach to a process-based hydrologic modeling in a snowmelt-driven watershed // Remote Sensing. – 2023. – Vol. 15, Issue 16. – Art. 3999. – doi: <http://dx.doi.org/10.3390/rs15163999>.
- [19] *Jin H., Fang S., Chen C.* Mapping of the spatial scope and water quality of surface water based on the Google Earth Engine cloud platform and Landsat time series // Remote Sensing. – 2023. – Vol. 15, Issue 20. – Art. 4986. – doi: <http://dx.doi.org/10.3390/rs15204986>.

- [20] *Santecchia V., Bovenga F., Refice A., Nutricato R., Lovergine F.P., Nitti D.O.* Assessment of Landsat-8 and Sentinel-2 water indices for surface water extent monitoring in a Mediterranean reservoir // *Journal of Imaging*. – 2023. – Vol. 9, Issue 5. – Art. 99. – doi: <http://dx.doi.org/10.3390/jimaging9050099>.
- [21] *Wang Y., Li Z., Zeng C. et al.* An urban water extraction method combining deep learning and Google Earth Engine // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2020. – Vol. 13. – P. 769-782. – doi: <http://dx.doi.org/10.1109/JSTARS.2020.2971783>.

UDC 519.6

MODELING THE SPATIO-TEMPORAL DYNAMICS OF A RESERVOIR AREA (USING THE KATTAKURGAN RESERVOIR AS AN EXAMPLE) BASED ON NDWI, NDVI, EVI INDICES AND ENSEMBLE LEARNING METHODS

^{1*}*Ravshanov N.,* ²*Rakhmanov Kh.E.* ²*Fattaeva D.A.*

^{*}*ravshanzade-09@mail.ru*

¹Digital Technologies and Artificial Intelligence Development Research Institute,
17A, Buz-2, Tashkent, 100125 Uzbekistan;

²Samarkand branch of Tashkent University of Information Technologies,
2A, Abu Ali Ibn Sino Str., Samarkand, 140100 Uzbekistan.

The paper models the spatiotemporal dynamics of the Kattakurgan Reservoir area using satellite-based water and vegetation indices combined with ensemble machine-learning methods. Multi-year time series of NDWI, NDVI, and EVI are used together with meteorological variables (air temperature, humidity, precipitation) from the ERA5 and CHIRPS datasets on the Google Earth Engine platform for 2018–2023. The reservoir water-surface area, derived from satellite imagery, serves as the target variable for training and validating a Random Forest model. The study includes a train–test split, hyperparameter tuning, and performance evaluation using R^2 , RMSE, and MAE. The results indicate high accuracy: $R^2 = 0.962$ for NDWI prediction, and values close to 1 for reservoir area, with RMSE about 5–7 km^2 and MAE 3–5 km^2 . Forecasts of NDWI and reservoir area for 2024 are produced to assess expected changes in water extent and support water-management decisions and climate adaptation planning.

Keywords: modeling, reservoir area forecast, indices, machine learning methods, spatiotemporal dynamics.

Citation: Ravshanov N., Rakhmanov Kh.E. Fattaeva D.A. 2025. Modeling the spatio-temporal dynamics of a reservoir area (using the Kattakurgan Reservoir as an example) based on NDWI, NDVI, EVI indices and ensemble learning methods. *Problems of Computational and Applied Mathematics*. 6(70): 61-73.

DOI: https://doi.org/10.71310/pcam.6_70.2025.05

ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОЙ И ПРИКЛАДНОЙ МАТЕМАТИКИ

№ 6(70) 2025

Журнал основан в 2015 году.

Издается 6 раз в год.

Учредитель:

Научно-исследовательский институт развития цифровых технологий и
искусственного интеллекта.

Главный редактор:

Равшанов Н.

Заместители главного редактора:

Арипов М.М., Шадиметов Х.М., Ахмедов Д.Д.

Ответственный секретарь:

Убайдуллаев М.Ш.

Редакционный совет:

Азамов А.А., Алоев Р.Д., Амиргалиев Е.Н. (Казахстан), Арушанов М.Л.,
Бурнашев В.Ф., Джумаёзов У.З., Загребина С.А. (Россия), Задорин А.И. (Россия),
Игнатьев Н.А., Ильин В.П. (Россия), Иманкулов Т.С. (Казахстан),
Исмагилов И.И. (Россия), Кабанихин С.И. (Россия), Карабчик В.В. (Россия),
Курбонов Н.М., Маматов Н.С., Мирзаев Н.М., Мухамадиев А.Ш., Назирова Э.Ш.,
Нормуродов Ч.Б., Нуралиев Ф.М., Опанасенко В.Н. (Украина),
Расулмухамедов М.М., Расулов А.С., Садуллаева Ш.А.,
Старовойтов В.В. (Беларусь), Хаётов А.Р., Халдзигитов А., Хамдамов Р.Х.,
Хужаев И.К., Хужаеров Б.Х., Чье Ен Ун (Россия), Шабозов М.Ш. (Таджикистан),
Dimov I. (Болгария), Li Y. (США), Mascagni M. (США), Min A. (Германия),
Singh D. (Южная Корея), Singh M. (Южная Корея).

Журнал зарегистрирован в Агентстве информации и массовых коммуникаций при
Администрации Президента Республики Узбекистан.

Регистрационное свидетельство №0856 от 5 августа 2015 года.

ISSN 2181-8460, eISSN 2181-046X

При перепечатке материалов ссылка на журнал обязательна.

За точность фактов и достоверность информации ответственность несут авторы.

Адрес редакции:

100125, г. Ташкент, м-в. Буз-2, 17А.

Тел.: +(998) 712-319-253, 712-319-249.

Э-почта: journals@airi.uz.

Веб-сайт: <https://journals.airi.uz>.

Дизайн и вёрстка:

Шарипов Х.Д.

Отпечатано в типографии НИИ РЦТИИ.

Подписано в печать 25.12.2025 г.

Формат 60x84 1/8. Заказ №8. Тираж 100 экз.

PROBLEMS OF COMPUTATIONAL AND APPLIED MATHEMATICS

No. 6(70) 2025

The journal was established in 2015.
6 issues are published per year.

Founder:

Digital Technologies and Artificial Intelligence Development Research Institute.

Editor-in-Chief:

Ravshanov N.

Deputy Editors:

Aripov M.M., Shadimetov Kh.M., Akhmedov D.D.

Executive Secretary:

Ubaydullaev M.Sh.

Editorial Council:

Azamov A.A., Aloev R.D., Amirgaliev E.N. (Kazakhstan), Arushanov M.L.,
Burnashev V.F., Djumayozov U.Z., Zagrebina S.A. (Russia), Zadorin A.I. (Russia),
Ignatiev N.A., Ilyin V.P. (Russia), Imankulov T.S. (Kazakhstan), Ismagilov I.I. (Russia),
Kabanikhin S.I. (Russia), Karachik V.V. (Russia), Kurbonov N.M., Mamatov N.S.,
Mirzaev N.M., Mukhamadiev A.Sh., Nazirova E.Sh., Normurodov Ch.B., Nuraliev F.M.,
Opanasenko V.N. (Ukraine), Rasulov A.S., Sadullaeva Sh.A., Starovoitov V.V. (Belarus),
Khayotov A.R., Khaldjigitov A., Khamdamov R.Kh., Khujaev I.K., Khujayorov B.Kh.,
Chye En Un (Russia), Shabozov M.Sh. (Tajikistan), Dimov I. (Bulgaria), Li Y. (USA),
Mascagni M. (USA), Min A. (Germany), Singh D. (South Korea), Singh M. (South
Korea).

The journal is registered by Agency of Information and Mass Communications under the
Administration of the President of the Republic of Uzbekistan.

The registration certificate No. 0856 of 5 August 2015.

ISSN 2181-8460, eISSN 2181-046X

At a reprint of materials the reference to the journal is obligatory.

Authors are responsible for the accuracy of the facts and reliability of the information.

Address:

100125, Tashkent, Buz-2, 17A.
Tel.: +(998) 712-319-253, 712-319-249.
E-mail: journals@airi.uz.
Web-site: <https://journals.airi.uz>.

Layout design:

Sharipov Kh.D.

DTAIDRI printing office.

Signed for print 25.12.2025

Format 60x84 1/8. Order No. 8. Print run of 100 copies.

Содержание

<i>Алимова Н.Б., Паровик Р.И.</i>	
Программный комплекс FrOsFHN для количественного и качественного анализа дробного осциллятора ФитцХью-Нагумо с переменной памятью	5
<i>Эшкулов М.У., Хамдамов Р.Х.</i>	
Проектирование и анализ системы солнечного водоснабжения для многоэтажных жилых зданий на основе булева программирования	18
<i>Равшанов Н., Усмонов Л.С.</i>	
Трёхмерная математическая модель и алгоритм численного решения для мониторинга и прогнозирования процессов подземного выщелачивания в пористой среде	26
<i>Каландаров А.А.</i>	
Численное моделирование связанной динамической задачи термоупругости в напряжениях	48
<i>Равшанов Н., Рахманов Х.Э. Фаттаева Д.А.</i>	
Моделирование пространственно-временной динамики площади водоёма (на примере Каттакурганского водохранилища) на основе индексов NDWI, NDVI, EVI и ансамблевых методов обучения	61
<i>Хажиев И.О., Шобдаров Э.Б.</i>	
Регуляризация начально-краевой задачи для неоднородного параболического уравнения с меняющимся направлением времени	74
<i>Равшанов Н., Боборахимов Б.И., Бердиев М.И.</i>	
Модель и алгоритмы классификации аномальных явлений на основе сходимости акустико-визуальных сигналов	88
<i>Рустамов Н., Мухамеджанов Н.Б.</i>	
Конструкция и принцип работы когенеративного фрактального солнечного коллектора	103
<i>Холиев Э.Ч., Тураев Д.Ш.</i>	
Численное решение плоскорадиальной граничной обратной задачи для уравнения нестационарной релаксационной фильтрации жидкости в пористой среде	112
<i>Ахмедов Д.М., Маматова Н.Х.</i>	
Оптимальный метод приближённого решения гиперсингулярных интегральных уравнений	124
<i>Шадиметов Х.М., Элмуратов Г.Ч.</i>	
Оптимизация приближенного вычисления интегралов от быстроосцилирующих функций в пространстве Соболева комплекснозначных функций	132
<i>Зиякулова Ш.А.</i>	
Об оптимальных итерационных и прямых методах решения задачи Дирихле для уравнения Пуассона	143

Contents

<i>Alimova N.B., Parovik R.I.</i>	
FrOsFHN software package for quantitative and qualitative analysis of the FitzHugh-Nagumo fractional oscillator with variable memory	5
<i>Eshkulov M.U., Khamdamov R.Kh.</i>	
Design and analysis of solar water supply system for multi-story residential buildings based on Boolean programming	18
<i>Ravshanov N., Usmonov L.S.</i>	
Three-dimensional mathematical model and numerical solution algorithm for monitoring and predicting in-situ leaching processes in porous medium	26
<i>Kalandarov A.A.</i>	
Numerical simulation of the coupled dynamic problem of thermoelasticity in stresses	48
<i>Ravshanov N., Rakhmanov Kh.E. Fattaeva D.A.</i>	
Modeling the spatio-temporal dynamics of a reservoir area (using the Kattakurgan Reservoir as an example) based on NDWI, NDVI, EVI indices and ensemble learning methods	61
<i>Khajiev I.O., Shobdarov E.B.</i>	
Regularization of the initial-boundary value problem for a inhomogeneous parabolic equation with changing time direction	74
<i>Ravshanov N., Boborakhimov B.I., Berdiev M.I.</i>	
Model and algorithms for classifying anomalous phenomena based on the convergence of acoustic-visual signals	88
<i>Rustamov N., Mukhamejanov N.B.</i>	
Design and operating principle of a cogenerative fractal solar collector	103
<i>Kholiyarov E.Ch., Turaev D.Sh.</i>	
Numerical solution of plane-radial boundary value inverse problem for the equation of non-stationary relaxation filtration of fluid in a porous medium	112
<i>Akhmedov D.M., Mamatova N.H.</i>	
An optimal method for the approximate solution of the hypersingular integral equations	124
<i>Shadimetov Kh.M., Elmuratov G.Ch.</i>	
Optimization of approximate computation of integrals of rapidly oscillating functions in the Sobolev space of complex-valued functions	132
<i>Ziyakulova Sh.A.</i>	
On optimal iterative and direct methods for solving the Dirichlet problem for the Poisson equation	143

HI SOBLASH VA AMALIY МАТЕМАТИКА MUAMMOLARI



ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОЙ
И ПРИКЛАДНОЙ МАТЕМАТИКИ
PROBLEMS OF COMPUTATIONAL
AND APPLIED MATHEMATICS

