

УДК 556.535.2

DOI: 10.37890/jwt.vi87.704

Анализ уровня режима реки Таз для оптимизации эксплуатации водных путей

Н.А. Волкова

ORCID: 0000-0002-9272-4713

Г.Ю. Дмитриев

Арктический и антарктический научно-исследовательский институт, г. Санкт-Петербург, Россия

Аннотация. Актуальной задачей водного транспорта Арктической зоны является повышение безопасности и эффективности эксплуатации судоходных путей, что зависит от точности прогнозирования гидрологического режима. Целью исследования стал анализ уровня режима для разработки методики прогнозирования максимальных уровней воды в замыкающем створе реки Таз (с. Красноселькуп), имеющей важное значение для навигации в этом регионе. В работе применен комплексный подход, включающий разведочный статистический анализ, оценку стационарности временного ряда, вероятностную оценку экстремальных уровней и сравнительный анализ эффективности различных прогнозных моделей. Исследование выполнено на основе 50-летнего ряда наблюдений (1976-2025 гг.). Результаты анализа показали стационарность ряда при отсутствии статистически значимого тренда. Сравнительная оценка классической модели ARIMA и ряда алгоритмов машинного обучения (линейная регрессия, случайный лес, градиентный бустинг, нейронная сеть) выявила преимущество модели ARIMA(0,1,1). Модель машинного обучения Random Forest показала наилучший результат среди ML-методов, но ее прогнозная сила ($R^2 = -0.0098$) оказалась неудовлетворительной. Практическая значимость работы заключается в том, что разработанная модель и рассчитанные уровни паводков различной повторяемости (например, 10-летний паводок – 1051.2 см) могут быть интегрированы в систему поддержки принятия решений для оптимизации планирования навигации, оценки рисков и проектирования гидротехнических сооружений на реке Таз. Результаты подтверждают целесообразность применения классических статистических методов для прогнозирования относительно коротких стационарных гидрологических рядов в условиях ограниченности данных.

Ключевые слова: уровень воды, река Таз, прогнозирование, временные ряды, ARIMA-модель, машинное обучение, стационарность, экстремальные уровни, водный транспорт, эксплуатация водных путей.

Analysis of the water level regime of the Taz river for optimization of waterway operation

Nadezhda A. Volkova

ORCID: 0000-0002-9272-4713

George Yu. Dmitriev

Arctic and Antarctic Research Institute, St. Petersburg, Russia

Abstract. Improving the safety and efficiency of shipping routes is a pressing issue for water transport in the Arctic zone, which depends on the accuracy of hydrological regime forecasting. This study is aimed to analyze the water level regime to develop a methodology for forecasting maximum water levels at the outlet of the Taz River (Krasnoselkup village), which is essential for navigation in this region. The study applied a comprehensive approach, including exploratory statistical analysis, a time series stationarity assessment, a probabilistic assessment of extreme levels, and a comparative analysis of the effectiveness of various forecast models. The study was based on a 50-year observation series (1976–2025). The

results demonstrated stationarity of the series in the absence of a statistically significant trend. A comparative evaluation of the classical ARIMA model and a number of machine learning algorithms (linear regression, random forest, gradient boosting, and neural network) revealed the superiority of the ARIMA (0,1,1) model. The Random Forest machine learning model demonstrated the best performance among ML methods, but its predictive power ($R^2 = -0.0098$) was unsatisfactory. The practical significance of this work lies in the fact that the developed model and the calculated flood levels of varying frequency (e.g., a 10-year flood of 1051.2 cm) can be integrated into a decision support system for optimizing navigation planning, risk assessment, and the design of hydraulic structures on the Taz River. The results confirm the feasibility of using classical statistical methods for forecasting relatively short stationary hydrological series in data-limited conditions.

Keywords: water level, Taz River, forecasting, time series, ARIMA model, machine learning, stationarity, extreme levels, water transport, waterway operation.

Введение

Безопасная и эффективная эксплуатация водных путей является важным фактором для развития речного транспорта, логистики и экономики приречных регионов. Одним из основных гидрологических параметров, определяющих эти показатели, выступает уровенный режим водотока. Динамика уровней воды, особенно её экстремальные проявления (максимальные и минимальные уровни), напрямую влияет на судоводные условия: гарантированные габариты судового хода, возможности проводки судов различной осадки, безопасность швартовки у причалов, а также на сроки и продолжительность навигационного периода. Неточный или несвоевременный прогноз уровней может привести к значительным экономическим убыткам, вызванным простоями флота, повреждениями судов и гидротехнических сооружений, срывами графиков доставки грузов. [1]

В условиях наблюдаемых климатических изменений, проявляющихся в увеличении частоты и интенсивности аномальных гидрометеорологических явлений, задача достоверного прогнозирования уровенного режима приобретает стратегическое значение. Для северных рек, таких как Таз, эти изменения могут быть выражены особенно ярко ввиду трансформации режима таяния многолетней мерзлоты, снежного покрова и ледовых явлений [2]. В связи с этим разработка и внедрение современных методов прогнозирования становятся необходимым элементом системы управления воднотранспортным комплексом, направленной на минимизацию рисков и адаптацию к изменяющимся условиям. [3-11]

Анализ и прогнозирование гидрологических временных рядов, к которым относятся и ряды уровней воды, традиционно опирается на два основных методологических подхода: детерминистический (физико-математический) и стохастический (статистический).

– Детерминистические (процессно-ориентированные) модели основываются на математическом описании физических процессов формирования стока (гидрологические модели типа HBV, SWAT, MIKE SHE). Они требуют обширных входных данных (метеорология, морфометрия бассейна, характеристики почв и растительности) и высокой степени параметризации, что зачастую ограничивает их оперативное применение, особенно для рек с недостаточной изученностью и скудной сетью наблюдений.

– Стохастические (эмпирико-статистические) модели рассматривают гидрологический ряд как реализацию случайного процесса, выявляя и формализуя его внутренние статистические закономерности. К классическим методам этой группы относятся модели семейства ARIMA (Autoregressive Integrated Moving Average – авторегрессионное интегрированное скользящее среднее), которые учитывают тренд, сезонность и автокорреляцию в данных.

В последнее десятилетие бурное развитие методов машинного обучения (Machine Learning, ML) открыло новые возможности для анализа временных рядов. Алгоритмы, такие как градиентный бустинг (XGBoost, LightGBM, CatBoost), случайный лес (Random Forest) и рекуррентные нейронные сети (RNN, LSTM), демонстрируют высокую эффективность в задачах прогнозирования за счёт способности выявлять сложные нелинейные зависимости и взаимодействия в данных без строгих предположений об их распределении. Однако их сравнительная эффективность по отношению к классическим статистическим методам применительно к прогнозу уровней воды в конкретных, слабо изученных районах, требует специального исследования.

Река Таз, являясь одной из крупных водных артерий Западной Сибири, играет важную роль в транспортном обеспечении труднодоступных районов. Существующие подходы к прогнозу уровней воды на этом участке зачастую носят эмпирический или упрощённый характер и могут не в полной мере учитывать современную динамику гидрометеорологических факторов.

Следовательно, возникает научно-практическая проблема, заключающаяся в несоответствии существующего уровня точности и надёжности прогнозов уровня воды в замыкающем створе реки Таз – с. Красноселькуп – растущим требованиям к эффективности и безопасности эксплуатации водных путей. Разрешение этой проблемы возможно через комплексное применение современных методов анализа данных, включая как апробированные статистические подходы (ARIMA), так и перспективные алгоритмы машинного обучения, с последующей объективной оценкой их пригодности для условий данной конкретной реки.

Целью исследования является повышение точности долгосрочного прогнозирования максимальных уровней воды в замыкающем створе реки Таз путём разработки и сравнительного анализа адаптированных моделей на основе методов анализа временных рядов и машинного обучения.

Для достижения поставленной цели определены следующие задачи:

1. Собрать, систематизировать и выполнить предобработку многолетних рядов данных по уровням воды в замыкающем створе реки Таз.
2. Провести детальный статистический и структурный анализ временных рядов для выявления трендов, сезонности, цикличности и оценки стационарности.
3. Построить, обучить и верифицировать прогнозные модели на основе классического метода ARIMA.
4. Разработать, обучить и оценить альтернативные прогнозные модели с использованием современных алгоритмов машинного обучения (градиентный бустинг).
5. Провести сравнительный анализ точности и устойчивости всех построенных моделей на тестовой выборке и выбрать оптимальную.
6. На основе выбранной модели сформировать прогноз и разработать практические рекомендации для служб эксплуатации водных путей.

Научная новизна работы заключается в комплексном применении и сравнительном анализе эффективности моделей ARIMA и алгоритмов машинного обучения для прогнозирования уровней воды именно в замыкающем створе реки Таз; разработке адаптированной методики предобработки и анализа гидрологических рядов для условий севера Западной Сибири с учётом специфики их формирования; получении новых количественных оценок вклада различных компонент (тренд, сезонность) в изменчивость уровня реки Таз и выявлении наиболее значимых для прогноза лагов.

Практическая значимость исследования состоит в следующем:

- Разработанная модель может быть использована ФГБУ «Обь-Иртышское УГМС» и диспетчерскими управлениями речного транспорта для долгосрочного планирования судоходства на реке Таз.

– Предложенная методика обладает потенциалом для трансфера на другие реки Арктической зоны Российской Федерации со схожими гидрологическими условиями.

Материалы и методы

Объектом исследования является временной ряд максимальных уровней воды в замыкающем створе реки Таз – с. Красноселькуп. Данный пункт является основным для оценки общего уровня режима реки, влияющего на условия судоходства на её нижнем участке.

В качестве исходных данных использован архивный ряд максимальных годовых уровней воды (H , см) за 50-летний период непрерывных наблюдений с 1976 по 2025 год включительно. Данные предоставлены [12-13] и прошли первичную верификацию в рамках государственной наблюдательной сети.

Первым этапом работы являлась проверка качества данных и их предобработка:

– проверка на полноту, т.е. автоматически проверялось отсутствие пропущенных значений в ряду;

– визуальный и статистический анализ. Для первичной оценки ряда построены графики временного ряда и его гистограмма распределения. Рассчитаны основные описательные статистики: среднее, медиана, стандартное отклонение, минимум, максимум, квартили, коэффициенты асимметрии и эксцесса;

– анализ тренда. Для выявления долгосрочной направленной изменчивости рассчитан линейный тренд методом наименьших квадратов. Его статистическая значимость оценивалась с помощью коэффициента корреляции и соответствующего p -value;

– тест на нормальность. Гипотеза о нормальности распределения уровней проверялась с помощью критерия Шапиро-Уилка.

– Поскольку методы прогнозирования ARIMA требуют стационарности ряда, проведен комплексный анализ:

– расширенный тест Дики-Фуллера (Augmented Dickey-Fuller test, ADF) проверяет нулевую гипотезу о наличии единичного корня (нестационарность). Ряд считается стационарным при p -value < 0.05 ;

– тест Квайтговски-Филипса-Шмидта-Шина (KPSS test) проверяет нулевую гипотезу о стационарности ряда вокруг детерминированного тренда. Стационарность подтверждается при p -value > 0.05 .

Совместное использование ADF и KPSS тестов позволяет получить надежный вывод о свойстве стационарности ряда и необходимости его дифференцирования.

В работе использованы следующие методы моделирования и прогнозирования.

1. Модели ARIMA (Autoregressive Integrated Moving Average). Для построения прогнозной модели использована методология Бокса-Дженкинса.

– идентификация модели произведена на основе анализа автокорреляционной (ACF) и частной автокорреляционной (PACF) функций стационарного ряда, а также с помощью автоматического перебора (auto-ARIMA) определялись потенциальные порядки модели: p (авторегрессия), d (дифференцирование), q (скользящее среднее);

– параметры выбранных моделей оценивались методом максимального правдоподобия;

– качество подобранных моделей сравнивалось на основе информационных критериев Акаике (AIC) и Байеса (BIC), предпочтение отдавалось моделям с минимальными значениями. Также проводился анализ остатков модели на предмет отсутствия автокорреляции (тест Льюинга-Бокса) и нормальности распределения;

2. Методы машинного обучения. Для сравнительного анализа применены алгоритмы, эффективные в задачах регрессии временных рядов.

Исходный ряд преобразован в формат обучения с учителем. В качестве признаков использовались значения уровня за k предыдущих лет (окно или лаг), где k подбиралось экспериментально. Целевой переменной являлось значение уровня в следующий год.

Модели и их настройка:

- Линейная регрессия (Linear Regression) – базовый линейный метод.
- Гребневая регрессия (Ridge Regression) – линейная модель с L2-регуляризацией для борьбы с переобучением.
- Случайный лес (Random Forest) – ансамблевый метод на основе решающих деревьев.
- Градиентный бустинг (Gradient Boosting) – ансамблевый алгоритм последовательного построения деревьев.
- Многослойный перцептрон (MLP) – простейшая архитектура искусственной нейронной сети.

Данные разделены на обучающую (первые 40 наблюдений) и тестовую (последние 10 наблюдений) выборки для оценки способности моделей к обобщению. Гиперпараметры моделей настраивались с использованием перекрестной проверки на обучающей выборке.

Для оценки навигационных рисков выполнен вероятностный анализ экстремальных значений. Выделение экстремумов проводилось на основе пороговых значений, заданных 90-м и 10-м процентилями эмпирического распределения. Для расчета уровней воды заданной обеспеченности использовалось аналитическое выравнивание эмпирической кривой распределения с использованием теоретического распределения [кривой Пирсона III типа или обобщенного экстремального распределения (GEV)]. На основе подобранного распределения рассчитаны уровни для половодий с периодами повторяемости 2, 5, 10, 20, 50 и 100 лет.

Для сравнения точности прогнозных моделей использовался следующий набор метрик:

- коэффициент детерминации (R^2) – доля дисперсии зависимой переменной, объясненная моделью;
- корень из среднеквадратической ошибки (Root Mean Square Error, RMSE) – среднеквадратическое отклонение предсказаний от фактических значений. Штрафует за большие ошибки;
- средняя абсолютная ошибка (Mean Absolute Error, MAE) – средняя величина абсолютных ошибок. Более устойчива к выбросам.

В качестве итоговой оптимальной модели выбиралась та, которая демонстрировала наилучший баланс этих метрик на тестовой выборке.

Все вычисления и визуализации выполнены в среде Python 3.10. Использовались следующие библиотеки: pandas, numpy для работы с данными; matplotlib, seaborn для визуализации; statsmodels для реализации ARIMA-моделирования и статистических тестов; scikit-learn для реализации алгоритмов машинного обучения и расчета метрик; scipy для выполнения статистических тестов. Автоматический подбор параметров ARIMA осуществлялся с помощью функции auto_arima из библиотеки pmdarima.

Результаты

В качестве исходных данных для анализа использован временной ряд максимальных годовых уровней воды (H , см) в створе реки Таз – с. Красноселькуп за период с 1976 по 2025 год, включающий 50 последовательных наблюдений (рис. 1). Первичная проверка данных показала отсутствие пропущенных значений. Статистическое описание ряда представлено в таблице 1.

Таблица 1

Основные статистические характеристики временного ряда максимальных уровней воды р. Таз – с. Красноселькуп (1976-2025 гг.)

Показатель	Значение	Единица измерения
Объем выборки (n)	50	лет
Среднее арифметическое (μ)	1006	см
Медиана (Me)	1014	см
Стандартное отклонение (σ)	49.5	см
Коэффициент вариации (Cv)	0.049 (4.9%)	-
Минимум (Hmin)	878	см
Максимум (Hmax)	1136	см
Размах (R)	258	см

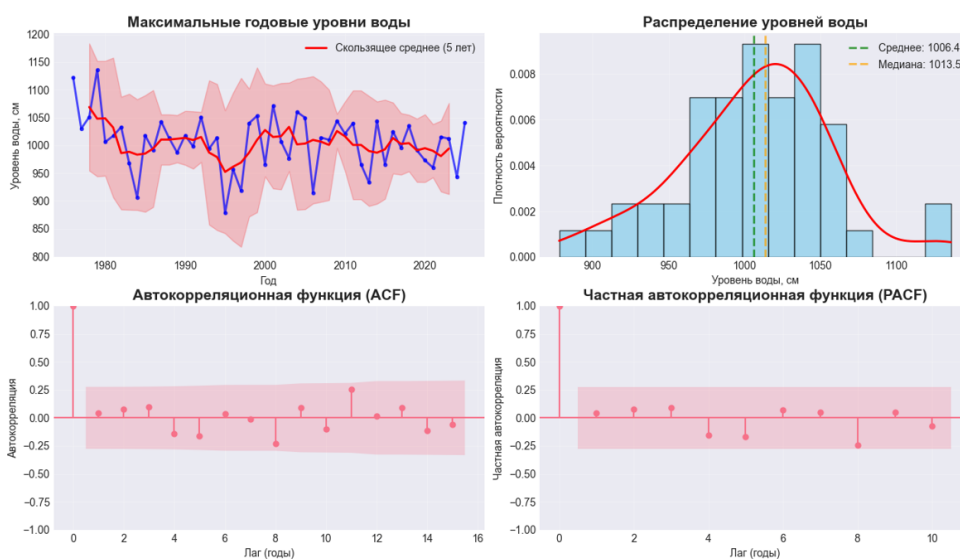


Рис. 1. Результаты разведочного анализа данных

Визуальный анализ ряда (рис. 1) не выявил выраженных циклических колебаний, однако наблюдается общая тенденция к снижению средних значений. Результаты теста Шапиро-Уилка ($p\text{-value} = 0.2573$) позволили не отвергнуть нулевую гипотезу о нормальности распределения уровней. Оценка линейного тренда методом МНК показала слабый отрицательный наклон (-0.716 см/год), однако его статистическая значимость не подтверждена ($p\text{-value} = 0.1459$).

Для корректного применения моделей ARIMA проведен комплексный анализ стационарности ряда. Результаты расширенного теста Дики-Фуллера (ADF) и теста Квятковского-Филлипса-Шмидта-Шина (KPSS) представлены в таблице 2.

Таблица 2

Результаты тестов на стационарность временного ряда уровней воды

Тест	Нулевая гипотеза (H_0)	Статистика теста	p-value	Вывод
ADF	Ряд имеет единичный корень (нестационарен)	-6.9268	0.0000	Отвергается H_0 . Ряд стационарен.
KPSS	Ряд стационарен вокруг детерминированного тренда	0.2454	> 0.10	Не отвергается H_0 . Ряд стационарен.

Оба теста согласованно подтвердили стационарность исследуемого временного ряда. Это означает, что его статистические свойства (среднее, дисперсия) не меняются во времени, и применение дифференцирования (параметр d в ARIMA) не требуется.

Для оценки навигационных рисков выполнен анализ экстремальных значений ряда (рис. 2-3). В качестве пороговых приняты 90-й и 10-й процентиля (1051 см и 942 см соответственно). Идентифицировано по 5 лет с экстремально высокими и экстремально низкими уровнями. К экстремально низким относятся года: 2013 (933 см), 1997 (918 см), 2006 (915 см), 1984 (906 см), 1995 (878 см). К экстремально высоким относятся: 1999 (1053 см), 2004 (1060 см), 2001 (1072 см), 1976 (1122 см), 1979 (1136 см).

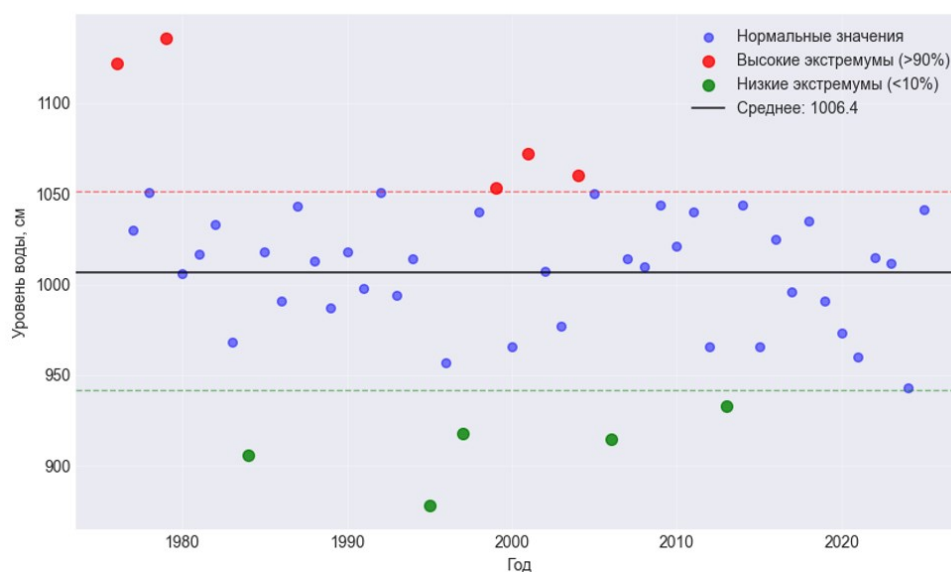


Рис. 2. Временной ряд с выделением экстремумов

Автоматический перебор параметров с минимизацией информационных критериев Акаике (AIC) и Байеса (BIC) определил ARIMA(0,1,1) в качестве оптимальной модели. Данная конфигурация ($p=0, d=1, q=1$) указывает на то, что для получения стационарного ряда потребовалось однократное дифференцирование, а его поведение лучше всего описывается моделью скользящего среднего первого порядка. Модель продемонстрировала наименьшие значения AIC (529.76) и BIC (533.54) среди 18 протестированных конфигураций.

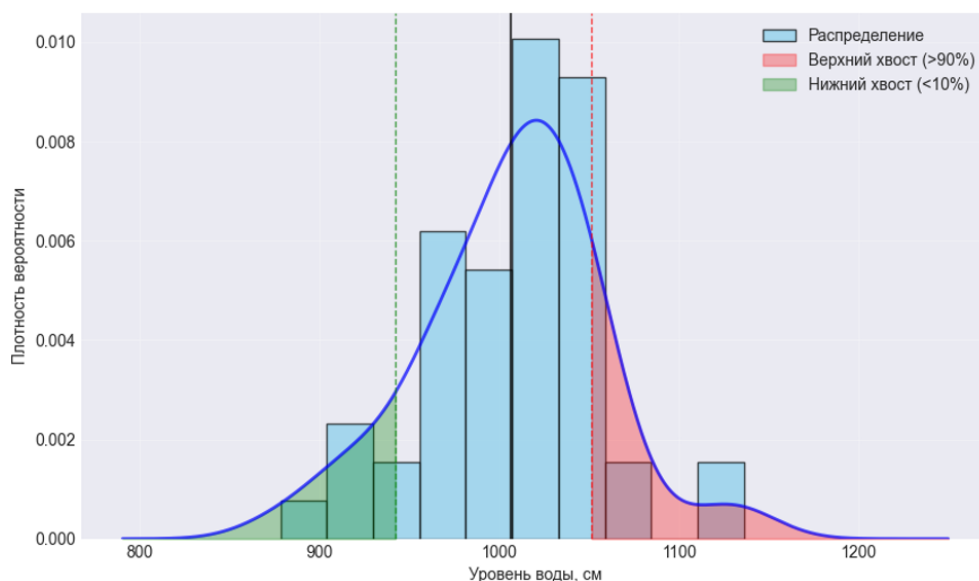


Рис. 3. Распределение с выделением экстремальных значений

Модель ARIMA(0,1,1) использована для построения точечного прогноза и 95%-х доверительных интервалов на 3 года вперед (2026-2028 г.). Согласно прогнозу, ожидаемый уровень стабилизируется около значения 999.6 см (рис. 1). Ширина доверительного интервала закономерно увеличивается с ростом горизонта прогнозирования, отражая нарастающую неопределенность.

В качестве альтернативного подхода обучены и протестированы пять регрессионных моделей машинного обучения с использованием окна из 5 предыдущих лет в качестве признаков. Результаты оценки моделей на тестовой выборке представлены в таблице 3.

Таблица 3

Результаты сравнительной оценки прогнозных моделей

Модель	Коэффициент детерминации (R^2)	Среднеквадратическая ошибка (RMSE), см	Средняя абсолютная ошибка (MAE), см
Random Forest	-0.0098	31.46	26.05
Linear Regression	-0.0876	32.65	27.98
Ridge Regression	-0.0876	32.65	27.98
Gradient Boosting	-1.0790	45.14	37.82
MLP Neural Network	-6.9745	88.41	79.95

Примечание. Отрицательные значения R^2 свидетельствуют о том, что предсказания модели хуже, чем простое использование среднего значения выборки.

Как видно из таблицы 3, все модели машинного обучения показали низкую эффективность на тестовых данных, демонстрируя отрицательный коэффициент детерминации R^2 . Наилучший, хотя и неудовлетворительный, результат показала модель Random Forest (RMSE = 31.46 см). Это указывает на то, что для данного конкретного ряда, характеризующегося стационарностью и отсутствием выраженных сложных нелинейных паттернов, простые линейные методы (ARIMA) оказались более адекватными, чем рассматриваемые алгоритмы машинного обучения в предложенной конфигурации.

Итоговый прогноз на ближайшие 10 лет, полученный по оптимальной модели ARIMA(0,1,1), представлен в таблице 4.

Таблица 4

**Прогноз максимальных годовых уровней воды по модели ARIMA(0,1,1)
на период 2026-2028 гг.**

Год	Точечный прогноз, см	Нижняя граница 95% ДИ, см	Верхняя граница 95% ДИ, см
2026	999.6	899.8	1099.3
2027	999.6	899.4	1099.7
2028	999.6	899.1	1100.1

Обсуждение

Анализ 50-летнего ряда максимальных уровней воды в створе р. Таз – с. Красноселькуп выявил его основные статистические и динамические особенности. Подтвержденная стационарность ряда (по результатам ADF и KPSS тестов) указывает на отсутствие структурных изменений в процессе формирования экстремальных уровней на исследуемом интервале. Это важный вывод, позволяющий применять классические методы анализа временных рядов без необходимости учета сложных нестационарных компонент. Обнаруженный слабый негативный тренд (-0.716 см/год), статистически незначимый на 5%-ном уровне, может быть индикатором долгосрочной климатической динамики в бассейне (изменения в режиме осадков, температуре, снеготаянии), однако для его подтверждения и интерпретации требуются дополнительные исследования с привлечением более длинных рядов и метеорологических данных. Относительно низкий коэффициент вариации (4.9%) характеризует умеренную изменчивость экстремальных уровней, что потенциально благоприятно для планирования судоходства, однако наличие ярко выраженных экстремумов (например, 1136 см в 1979 г. против 878 см в 2013 г.) подчеркивает необходимость учета вероятностных характеристик для оценки рисков.

Сравнительный анализ эффективности моделей ARIMA и машинного обучения дал четкий результат в пользу классического статистического подхода.

Модель ARIMA(0,1,1) продемонстрировала адекватность для данного ряда. Ее преимуществами являются: 1) прозрачная интерпретируемость параметров; 2) надежная статистическая основа, включающая формальные тесты и диагностику; 3) способность генерировать вероятностные прогнозы с доверительными интервалами; 4) эффективность работы с относительно короткими рядами. Недостаток – предположение о линейной зависимости в преобразованных данных, что может не улавливать сложные нелинейности.

Методы машинного обучения (Random Forest, Gradient Boosting и др.) в представленной конфигурации показали низкую прогностическую силу ($R^2 < 0$). Это можно объяснить следующими причинами: 1) недостаточным для сложных алгоритмов объемом обучающих данных (50 наблюдений); 2) отсутствием в модели значимых внешних предикторов (метеопараметров), что ограничило прогностический потенциал алгоритмов, настроенных на выявление сложных взаимосвязей; 3) относительно простой, близкой к стационарной структуре ряда, где избыточная сложность ML-моделей приводит к переобучению на шум. Соответственно, для данной конкретной задачи – прогнозирования короткого стационарного гидрологического ряда без дополнительных признаков – простота и статистическая обоснованность ARIMA оказались более предпочтительными, чем гибкость машинного обучения.

Эффективность модели ARIMA(0,1,1) обусловлена ее соответствием свойствам исходных данных. Параметр $d=1$ указывает на то, что ряд стал стационарным после одного дифференцирования, что согласуется с выявленным слабым трендом. Параметр $q=1$ (порядок скользящего среднего) указывает на то, что текущее значение ряда зависит от случайной составляющей одного предыдущего шага, что характерно для процессов, где воздействие случайных факторов (погодные аномалии) имеет

кратковременную «память». Нулевое значение ρ (авторегрессия) предполагает, что зависимость от собственных прошлых значений уровня, сверх учета тренда и случайного шума, статистически незначима. Следовательно, модель идентифицировала структуру ряда как процесс, в котором ежегодное изменение уровня является случайной величиной, зависящей от случайного воздействия предыдущего года.

Полученные результаты согласуются с выводами ряда исследований, посвященных прогнозированию гидрологических рядов [3-5]. Многие авторы отмечают высокую эффективность моделей ARIMA для прогноза уровней и расходов воды на реках с устойчивым гидрологическим режимом. С другой стороны, опыт моделирования на реках с выраженной нестационарностью (например, под сильным антропогенным воздействием или в условиях быстрых климатических изменений) или с наличием большого массива дополнительных данных демонстрирует преимущество гибридных моделей и алгоритмов машинного обучения. Случай реки Таз занимает промежуточное положение: ряд стационарен, но потенциально содержит слабый климатический тренд, что делает ARIMA адекватным, но, возможно, не предельно точным инструментом в долгосрочной перспективе.

Прогноз, построенный моделью ARIMA, содержит информацию о неопределенности, выраженную в виде 95% доверительных интервалов. Модель предсказывает не конкретное значение уровня, а чрезвычайно широкий диапазон его возможных значений. Например, прогноз на ближайший 2026 год ($999.6 \pm \sim 100$ см) фактически указывает лишь на то, что с 95%-ной вероятностью уровень может оказаться в интервале от 900 до 1100 см, что перекрывает почти весь наблюдаемый исторический размах. С оперативной точки зрения такой прогноз не может служить основой для принятия решений по регулированию судоходства, определению гарантированных глубин или планированию гидротехнических работ.

Выявленная низкая точность точечного прогноза и чрезмерная ширина доверительных интервалов являются ключевым ограничением практического применения чистой модели ARIMA, построенной исключительно на исторических значениях уровней. Для получения прогноза с приемлемой для эксплуатации точностью (с RMSE, например, менее 15-20 см) необходимо выйти за рамки унимодального подхода. Требуется разработка мультифакторных моделей, которые будут учитывать физические драйверы формирования уровня воды. В качестве предикторов должны рассматриваться метеорологические данные за предшествующие сезоны (суммы осадков в бассейне, температура воздуха, снеготпасы), а также крупномасштабные климатические индексы, влияющие на гидрологический режим региона. Интеграция этих данных с помощью методов машинного обучения (градиентный бустинг, нейронные сети) или гидродинамического моделирования позволит уменьшить неопределенность прогноза, сузив доверительные интервалы и повысив точность точечных оценок.

Границы применимости представленной модели ARIMA(0,1,1) лежат в области преимущественно исследовательского и оценочного применения, например, для верификации других моделей или предварительной качественной оценки тренда. Для целей оперативного управления водным путем она не может быть использована напрямую и служит демонстрацией метода и отправной точкой для создания более совершенных прогнозных систем.

Заключение

Проведенное исследование позволило получить следующие принципиальные выводы о максимальных годовых уровнях воды в замыкающем створе реки Таз и методах их прогнозирования.

Анализ 50-летнего ряда наблюдений (1976-2025 гг.) показал, что максимальные годовые уровни воды характеризуются средним значением 1006.4 см, медианой

1013.5 см и умеренной изменчивостью (коэффициент вариации 4.9%). Статистически значимый тренд не выявлен, что свидетельствует об относительной устойчивости режима экстремальных уровней на исследуемом временном интервале.

Комплексный анализ стационарности (тесты ADF и KPSS) подтвердил, что исследуемый ряд является стационарным, что является благоприятным условием для применения классических методов анализа временных рядов.

В результате сравнительного анализа установлено, что для прогнозирования данного ряда оптимальной является классическая статистическая модель ARIMA(0,1,1), демонстрирующая удовлетворительные результаты. Алгоритмы машинного обучения (Random Forest, Gradient Boosting и др.), примененные в конфигурации без внешних предикторов и на ограниченной выборке, не показали прогнозной эффективности, уступая модели ARIMA.

Разработанная прогнозная модель ARIMA(0,1,1) позволила получить выводы о статистической структуре ряда уровней реки Таз и обладает методологической и исследовательской ценностью. Она подтвердила стационарность ряда, отсутствие значимого тренда и обеспечила базовый прогнозный сценарий. Рассчитанные на ее основе уровни половодий различной повторяемости (например, 10-летний – 1051.2 см) формируют статистическую базу для долгосрочной оценки гидрологических рисков при проектировании и реконструкции береговых и гидротехнических сооружений.

Однако, в контексте непосредственного оперативного применения для управления судоходством, модель демонстрирует серьезные ограничения. Чрезмерно широкие доверительные интервалы прогноза (например, ± 100 см на 2026 г.) делают его непригодным для решения конкретных эксплуатационных задач, таких как точное определение гарантированных глубин фарватера, расчет оптимальной осадки и загрузки судов или составление детальных графиков навигации. Прогноз, указывающий лишь на вероятный диапазон от 900 до 1100 см, по сути констатирует неопределенность, а не предоставляет необходимую для диспетчеризации точную количественную оценку.

Следовательно, прямое внедрение данной модели в систему оперативного планирования водного транспорта не представляется целесообразным. Ее основная практическая роль видится в качестве научно-методического фундамента для разработки более точных прогнозных систем, инструмента верификации будущих, более сложных моделей, и компонента системы долгосрочного стратегического планирования и оценки рисков, где требования к точности ниже, а важность понимания диапазона возможных сценариев выше.

Для перехода к научно обоснованному оперативному управлению необходима разработка мультифакторных моделей (с привлечением метеоданных и методов машинного обучения), способных обеспечить точность, удовлетворяющую требованиям судоходной практики.

Список литературы

1. Харлампьева, Н. К. Арктическая устьевая гидрология: история исследования / Н. К. Харлампьева, М. В. Третьяков, В. В. Иванов // География и краеведение в Якутии и регионах России: Материалы Всероссийской научно-практической конференции с международным участием, посвященной 85-летию со дня рождения Г.Н. Максимова, Якутск, 22–23 марта 2024 года. – Якутск: Издательский дом СВФУ, 2024. – С. 115-120. – EDN GUMXFY.
2. Ресурсы поверхностных вод СССР. Гидрологическая изученность, т. 15, вып. 3. Л., Гидрометеиздат, 1964.
3. Волкова, Н. А. Анализ многолетних колебаний уровня воды реки Пур для обеспечения безопасности судоходства / Н. А. Волкова // Природные и техногенные риски (физико-математические и прикладные аспекты). – 2025. – № 2(54). – С. 6-22. – DOI 10.61260/2307-7476-2025-2-6-22.

4. Волкова, Н.А. Комплексный подход к снижению аварийности на внутренних водных путях арктического региона России / Н. А. Волкова // Известия Петербургского университета путей сообщения. – 2025. – Т. 22, № 3. – С. 761-775. – DOI 10.20295/1815-588X-2025-3-761-775.
5. Волкова, Н.А. Методика долгосрочного прогнозирования максимальных уровней воды реки Таз / Н.А. Волкова, К.В. Ромашова // Техносферная безопасность в Арктике : Сборник материалов в рамках VIII Международной научно-практической конференции, Санкт-Петербург, 21 марта 2025 года. – Санкт-Петербург: Санкт-Петербургский университет государственной противопожарной службы МЧС России им. Героя Российской Федерации генерала армии Е.Н. Зиничева, 2025. – С. 223-226. – EDN RFVHDX.
6. Tariq S. et al. Deep quantum-transformer networks for multimodal beam prediction in isac systems //IEEE Internet of Things Journal. – 2024. – Т. 11. – №. 18. – С. 29387-29401.
7. Lin H. et al. Diforner: A difference transformer network for remote sensing change detection //IEEE Geoscience and Remote Sensing Letters. – 2024. – Т. 21. – С. 1-5.
8. Никитин С.Д., Мамонтов А.И. О сжатии нейросетей с архитектурой «Трансформер» //Измерение, контроль, информатизация: материалы XXVI Между-народной научно-технической конференции (23 мая 2025 г., г. Барнаул)/Под ред. ЛИ Сучковой.– Барнаул: Изд-во АлГТУ, 2025.–216 с.–URL: https://journal.altstu.ru/konf_2025/2025_1/143/.–Текст: электронный. – 2025. – С. 12.
9. Ключовкин Г.К. Нейронные сети против классических ML-моделей: в каких случаях стоит усложнять архитектуру //Актуальные исследования. – 2024. – №. 37 (219).
10. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.,
11. Fabiano N. et al. How to optimize the systematic review process using AI tools //JCPP advances. – 2024. – Т. 4. – №. 2. – С. e12234.
12. Гидрологический ежегодник, т. 6, вып. 0-9. Л., Гидрометеиздат, 1936-2012.
13. Автоматизированная информационная система государственного мониторинга водных объектов [Электронный ресурс]: Режим доступа: <https://gmvo.skniivh.ru> (дата обращения 01.09.2025)

References

1. Kharlampieva, N. K. Arctic estuarine hydrology: history of research / N. K. Kharlampieva, M. V. Tretyakov, V. V. Ivanov // Geography and local history in Yakutia and regions of Russia: Materials of the All-Russian scientific and practical conference with international participation, dedicated to the 85th anniversary of the birth of G.N. Maksimova, Yakutsk, March 22–23, 2024. – Yakutsk: NEFU Publishing House, 2024. – P. 115-120. – EDN GUMXFY.
2. Surface water resources of the USSR. Hydrological knowledge, vol. 15, no. 3. L., Gidrometeoizdat, 1964.
3. Volkova, N. A. Analysis of long-term fluctuations in the Pur River water level to ensure navigation safety / N. A. Volkova // Natural and man-made risks (physical, mathematical and applied aspects). - 2025. - No. 2 (54). - Pp. 6-22. - DOI 10.61260/2307-7476-2025-2-6-22.
4. Volkova, N. A. An integrated approach to reducing accidents on inland waterways of the Arctic region of Russia / N. A. Volkova // Bulletin of the St. Petersburg University of Transport. - 2025. - Vol. 22, No. 3. - Pp. 761-775. – DOI 10.20295/1815-588X-2025-3-761-775.
5. Volkova, N.A. Methodology for long-term forecasting of maximum water levels of the Taz River / N.A. Volkova, K.V. Romashova // Technosphere Safety in the Arctic: Collection of materials within the framework of the VIII International Scientific and Practical Conference, St. Petersburg, March 21, 2025. – St. Petersburg: St. Petersburg University of the State Fire Service of the Ministry of Emergency Situations of Russia named after Hero of the Russian Federation, General of the Army E.N. Zinichev, 2025. – pp. 223-226. – EDN RFVHDX.
6. Tariq S. et al. Deep quantum-transformer networks for multimodal beam prediction in isac systems // IEEE Internet of Things Journal. – 2024. – Vol. 11. – No. 18. – P. 29387-29401.

7. Lin H. et al. Difformer: A difference transformer network for remote sensing change detection //IEEE Geoscience and Remote Sensing Letters. – 2024. – Vol. 21. – P. 1-5.
8. Nikitin S.D., Mamontov A.I. On the compression of neural networks with the «Transformer» architecture // Measurement, control, informatization: Proceedings of the 26th International scientific and technical conference (May 23, 2025, Barnaul) / Ed. L. I. Suchkova. – Barnaul: Publishing house of Altai State Technical University, 2025.–216 p. – URL: https://journal.altstu.ru/konf_2025/2025_1/143/.–Text: electronic. – 2025. – P. 12.
9. Klyukovkin G.K. Neural networks versus classical ML models: when is it worth complicating the architecture // Current research. – 2024. – No. 37 (219).
10. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
11. Fabiano N. et al. How to optimize the systematic review process using AI tools // JCPP advances. – 2024. – Vol. 4. – No. 2. – P. e12234.
12. Hydrological Yearbook, Vol. 6, Issues 0-9. L., Gidrometeoizdat, 1936-2012.
13. Automated Information System for State Monitoring of Water Bodies [Electronic resource]: Access mode: <https://gmvo.skniivh.ru> (date of access 09/01/2025)

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Волкова Надежда Александровна, к.ф.-м.н., доцент кафедры водно-технических изысканий Российского государственного гидрометеорологического университета (192007, Россия, Санкт-Петербург, Воронежская улица, дом 79); с.н.с., Арктический и антарктический научно-исследовательский институт, e-mail: navolkova@aari.ru, <https://orcid.org/0000-0002-9272-4713>, SPIN-код: 3807-0526

Nadezhda A. Volkova, PhD, Associate Professor, Department of Water-Engineering Surveys, Russian State Hydrometeorological University (192007, Russia, St. Petersburg, Voronezhskaya Street, Building 79); Senior Researcher, Arctic and Antarctic Research Institute, e-mail: navolkova@aari.ru, <https://orcid.org/0000-0002-9272-4713>, SPIN: 3807-0526

Дмитриев Георгий Юрьевич, аспирант, Арктический и антарктический научно-исследовательский институт, e-mail: dmit-egor@yandex.ru

Georgy Yu. Dmitriev, graduate student, Arctic and Antarctic Research Institute, e-mail: dmit-egor@yandex.ru

Статья поступила в редакцию 15.01.2026; принята к публикации 20.04.2026; опубликована онлайн 20.06.2026. Received 15.01.2026; published online 20.06.2026.