
СНИЖЕНИЕ РИСКОВ И ЛИКВИДАЦИЯ ПОСЛЕДСТВИЙ ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЙ. ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ ПРИ ЧС

Научная статья

УДК 556.06, 556.043, 556.535.2; DOI: 10.61260/1998-8990-2026-1-8-27

ПРОГНОЗИРОВАНИЕ СУТОЧНЫХ УРОВНЕЙ ВОДЫ РЕКИ ПУР ЗА НАВИГАЦИОННЫЙ ПЕРИОД НА ОСНОВЕ ТРАНСФОРМЕРНЫХ НЕЙРОННЫХ СЕТЕЙ

✉ Волкова Надежда Александровна.

Российский государственный гидрометеорологический университет;

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

✉ navolkova@aari.ru

Аннотация. В условиях наблюдаемых климатических изменений Арктические регионы становятся все более уязвимыми к экстремальным гидрологическим явлениям, таким как наводнения. Традиционные методы прогнозирования зачастую не способны уловить сложные нелинейные и долгосрочные зависимости в гидрометеорологических данных, что создает значительные риски для управления техносферной безопасностью прибрежных территорий. Цель исследования: разработка и апробация модели прогнозирования суточных уровней воды арктической реки Пур в замыкающем створе Самбург в навигационный период на основе архитектуры трансформерных нейронных сетей для повышения точности и заблаговременности предупреждения о паводковых ситуациях. В исследовании использована оригинальная архитектура нейронной сети типа Transformer, адаптированная для задач прогнозирования временных рядов. Модель включает механизм многоголового самовнимания, позиционного кодирования и полносвязные слои. Обучение и валидация модели проведены на массиве многолетних суточных данных, включающих целевые показатели уровня воды и 12 сопутствующих гидрометеорологических признаков. Для оценки точности применялись метрики MAE, RMSE и MAPE. Построенная модель продемонстрировала высокую точность в прогнозировании суточной динамики уровня воды. Вычисленные метрики качества (MAE = 10,76, RMSE = 14,00, MAPE = 2,69 %) подтверждают ее эффективность. Модель успешно выявляет как сезонные тенденции, так и аномальные паводковые пики, что важно для систем раннего предупреждения. Трансформерная модель представляет собой перспективный инструмент для решения задач оперативного прогнозирования уровня режима рек в сложных условиях Арктики. Внедрение разработанного подхода в практику управления рисками позволяет перейти от реактивных к проактивным стратегиям, обеспечивая своевременное принятие управленческих решений и снижение техногенных и природных рисков.

Ключевые слова: уровень режим, прогнозирование, река Пур, трансформерные нейронные сети, навигационная безопасность

Для цитирования: Волкова Н.А. Прогнозирование суточных уровней воды реки Пур за навигационный период на основе трансформерных нейронных сетей // Проблемы управления рисками в техносфере. 2026. № 1 (77). С. 8–27. DOI: 10.61260/1998-8990-2026-1-8-27

© Санкт-Петербургский университет ГПС МЧС России, 2026

Scientific article

FORECASTING DAILY WATER LEVELS OF THE PUR RIVER FOR THE NAVIGATION PERIOD BASED ON TRANSFORMER NEURAL NETWORKS

✉Volkova Nadezhda A.

Russian state hydrometeorological university;

Arctic and antarctic research institute, Saint-Petersburg, Russia

✉navolkova@aari.ru

Abstract. In the face of ongoing climate change, Arctic regions are becoming increasingly vulnerable to extreme hydrological events such as floods. Traditional forecasting methods are often unable to capture complex nonlinear and long-term dependencies in hydrometeorological data, posing significant risks to the management of coastal technosphere safety. The objective of this study was to develop and test a model for forecasting daily water levels in the Arctic Pur River at the Samburg outlet during the navigation season, based on a transformer neural network architecture. This model aims to improve the accuracy and lead time of flood warnings. The study utilized a unique Transformer-type neural network architecture adapted for time series forecasting. The model incorporates a multi-headed self-attention mechanism, positional coding, and fully connected layers. Training and validation were conducted on a multi-year daily data set, including target water level indicators and 12 associated hydrometeorological features. The MAE, RMSE, and MAPE metrics were used to assess accuracy. The constructed model demonstrated high accuracy in forecasting diurnal water level dynamics. The calculated quality metrics (MAE = 10,76, RMSE = 14,00, MAPE = 2,69 %) confirm its effectiveness. The model successfully identifies both seasonal trends and abnormal flood peaks, which is important for early warning systems. The transformer model is a promising tool for solving problems of operational river level forecasting in challenging Arctic conditions. Implementation of the developed approach in risk management practices allows for a transition from reactive to proactive strategies, ensuring timely management decisions and mitigating man-made and natural risks.

Keywords: water level regime, forecasting, the Pur River, transformer neural networks, navigation safety

For citation: Volkova N.A. Forecasting daily water levels of the Pur river for the navigation period based on transformer neural networks // Problemy upravleniya riskami v tekhnosfere = Problems of risk management in the technosphere. 2026. № 1 (77). P. 8–27. DOI: 10.61260/1998-8990-2026-1-8-27

Введение

Навигационный период в Арктическом регионе представляет собой важный временной интервал для транспортной логистики, снабжения удалённых территорий и промышленного развития. Однако его кратковременность и изменчивость гидрологических условий создают комплекс серьёзных сложностей для обеспечения безопасной и устойчивой хозяйственной деятельности. Традиционно для прогнозирования уровней воды в арктических реках применяются физико-математические модели, основанные на уравнениях гидродинамики и балансовых соотношениях [1]. Такие модели, однако, требуют детальных данных о рельефе, грунтах, характеристиках русла и метеорологических параметрах, которые в условиях Арктики часто являются недостаточными или недоступными [2]. Кроме того, эти модели плохо адаптируются к быстро меняющимся климатическим условиям и часто не способны адекватно учитывать нелинейные эффекты, связанные с таянием многолетней мерзлоты, изменением режима осадков и частотой экстремальных явлений.

Широкое распространение в оперативной гидрологической практике получили статистические методы, в частности, модели авторегрессии (ARIMA, SARIMA), которые

хорошо работают со стационарными рядами данных [3]. Однако их применение в Арктике ограничено ввиду выраженной нестационарности гидрологических процессов, наличия резких сезонных переходов и экстремальных выбросов, связанных с паводками. Статистические модели также не позволяют эффективно интегрировать множество взаимосвязанных гидрометеорологических факторов, что снижает их прогностическую силу в условиях сложной мультифакторной среды.

С развитием технологий машинного обучения появились новые возможности для прогнозирования временных рядов. Рекуррентные нейронные сети (RNN), в частности их модификации с долгой краткосрочной памятью (LSTM) и управляемыми рекуррентными блоками (GRU), продемонстрировали высокую эффективность в моделировании последовательностей за счёт способности улавливать долговременные зависимости [4, 5]. Ряд исследований подтвердил применимость LSTM-сетей для прогнозирования уровней воды и стока рек, в том числе в условиях Севера [6]. Тем не менее данным архитектурам присущи ограничения, связанные с последовательной природой обработки данных, что приводит к высоким вычислительным затратам при работе с длинными рядами и сложностями в моделировании глобальных зависимостей между далеко отстоящими друг от друга временными точками.

Перспективным направлением в обработке последовательностей стало появление архитектуры Transformer, основанной исключительно на механизмах внимания (attention) и не содержащей рекуррентных связей [7]. Трансформеры позволяют параллельно обрабатывать всю входную последовательность, что значительно ускоряет обучение, а механизм самовнимания (self-attention) даёт возможность напрямую моделировать зависимости между любыми элементами ряда, независимо от расстояния между ними. Эта архитектура стала де-факто стандартом в задачах обработки естественного языка и компьютерного зрения, а в последние годы активно исследуется для анализа временных рядов [8, 9].

Несмотря на растущий интерес к применению трансформеров в геонауках, их адаптация для прогнозирования гидрологических процессов, особенно в условиях Арктики, изучена недостаточно. Большинство существующих работ сосредоточено на прогнозировании стока или уровня воды в умеренных широтах с использованием относительно коротких исторических рядов. Отсутствуют исследования, комплексно оценивающие эффективность трансформерных архитектур для моделирования суточной динамики уровня воды в арктических реках с учётом широкого набора гидрометеорологических предикторов в течение продолжительного навигационного периода.

Соответственно, научная новизна настоящего исследования заключается в целенаправленной адаптации и валидации архитектуры Transformer для решения задачи прогнозирования суточных уровней воды средней арктической реки (на примере р. Пур) в условиях выраженной нестационарности, сезонности и мультифакторности процессов. В отличие от предыдущих исследований, в данной работе:

- предложена модифицированная архитектура Transformer, оптимизированная для работы с многомерными гидрометеорологическими временными рядами;
- осуществлено комплексное тестирование модели на протяжённом многолетнем массиве данных, характерном для Арктического региона;
- проведена детальная оценка точности прогноза не только в целом, но и в ключевых для навигации диапазонах уровней воды (высоких, средних, низких).

Цель работы – разработать, реализовать и апробировать модель прогнозирования суточных уровней воды р. Пур в створе с. Самбург на навигационный период на основе архитектуры Transformer, оценив её точность и потенциал для повышения заблаговременности предупреждений о паводковых ситуациях и оптимизации управления судоходством.

В качестве объекта исследований выбрана средняя арктическая р. Пур в замыкающем створе с. Самбург. Река образуется от слияния рек Пяку-Пур и Айваседа-Пур.

Площадь бассейна в месте слияния этих рек равна 57 500 км² и возрастает к устью р. Пур до 112 000 км². От истоков р. Пяку-Пур общая длина р. Пур составляет 1 024 км [2]. Русло реки на всем протяжении песчаное, подверженное значительным деформациям при изменениях скоростного режима потока. В паводковый период на различных участках реки наблюдается довольно сильный размыв и намыв. В русле реки отлагается масса песка, образуя множество песчаных осередков, кос, островов. Половодье сопровождается значительными подъемами воды, превышающими межень на 3,6–4,8 м. Продолжительность разлива, когда вода выходит на пойму, составляет 6–8 дн. [10–12].

Научная новизна исследования

Научная новизна данного исследования определяется совокупностью теоретических и методических решений, направленных на преодоление ограничений существующих подходов к гидрологическому прогнозированию в условиях Арктики.

Впервые предложена и реализована модифицированная архитектура нейронной сети типа Transformer, специально спроектированная для прогнозирования суточных уровней воды в условиях средней арктической реки. Новизна заключается в синтезе ключевых компонентов трансформера (механизм многоголового самовнимания, позиционное кодирование) с процедурами предобработки и структурой данных, учитывающими специфику арктических гидрометеорологических рядов: высокую нестационарность, наличие экстремальных паводковых пиков, выраженную сезонность и мультифакторность.

Работа предлагает формальную постановку задачи прогнозирования уровня воды как задачи регрессии, где целевая переменная (уровень воды в день $t+1$) моделируется на основе многомерной входной последовательности фиксированной длины, включающей как историю целевого показателя, так и синхронные временные ряды 12 гидрометеорологических признаков. Такой подход позволяет модели выявлять скрытые кросскорреляционные и причинно-следственные связи между разнородными факторами (температура, осадки, индексы циркуляции и т.д.) и откликом речной системы.

В отличие от многих исследований, использующих данные одной станции, в работе предложен метод агрегации данных с привязкой к бассейну реки. Данные метеостанции Тарко-Сале, расположенной в центральной части бассейна Пура, используются в качестве прокси-показателей для всего водосбора, что является инженерным упрощением, обоснованным для средних арктических рек с недостаточной сетью наблюдений.

Впервые для арктических условий проведена детализированная оценка точности модели не только интегральными метриками (MAE, RMSE, MAPE), но и в разрезе трёх критически важных для навигации и безопасности диапазонов уровней воды: высокого (паводкоопасного), среднего (нормальной навигации) и низкого (судоходно-проблемного). Полученные результаты (например, систематическое запаздывание в прогнозе пиков) имеют не только прикладное, но и научное значение, указывая на направления дальнейшего совершенствования архитектуры.

Разработан и описан комплексный конвейер предобработки, решающий специфические проблемы реальных арктических данных: парсинг нестандартных временных меток, обработка пропусков и аномалий в условиях малого объёма данных, согласованная нормализация многомерных рядов. Методическая ценность данного конвейера выходит за рамки конкретной модели и может быть использована в других исследованиях по арктической гидрологии.

Научная новизна работы носит комплексный характер и лежит на стыке машинного обучения, гидрологии и методов обработки данных, внося вклад в развитие инструментария для прогнозного моделирования в условиях изменяющегося климата Арктики.

Математическая постановка задачи

Пусть процесс формирования суточного уровня воды в замыкающем створе реки описывается многомерным временным рядом. Обозначим:

– T – дискретное время (день);

– $y_t \in \mathbb{R}$ – целевая переменная, представляющая собой среднесуточный уровень воды (в см) в день t , где $t = 1, \dots, N$, N – общая длина ряда наблюдений;

– $x_t \in \mathbb{R}^m$ – вектор входных признаков (предикторов) в день t , где $m = 12$. Этот вектор включает гидрометеорологические параметры, измеренные в тот же день t , но оказывающие влияние на уровень воды в последующие дни:

$$x_t = \left(x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(m)} \right)^T.$$

Конкретно, x_t включает: температуру воздуха, осадки, скорость ветра, высоту снежного покрова, температуру почвы, атмосферное давление, дефицит насыщения, относительную влажность, температуру точки росы, индексы АО и НАО, уровень воды в верхнем створе (п. Уренгой).

Задача прогнозирования формулируется следующим образом: на основе известной истории наблюдений за L предыдущих дней (длина *окна контекста* или *look-back period*) спрогнозировать значение уровня воды на один шаг вперёд (прогноз на завтра).

Формально, для любого момента времени $t \geq L$ задаётся *входная последовательность* (контекстное окно) X_t , представляющая собой матрицу размерности $L \times (m + 1)$, которая объединяет историю целевой переменной и всех предикторов:

$$X_t = \begin{bmatrix} y_{t-L} & x_{t-L}^T \\ y_{t-L+1} & x_{t-L+1}^T \\ \dots & \dots \\ y_{t-1} & x_{t-1}^T \end{bmatrix}.$$

Цель – найти такую функцию (модель) $f_\theta: \mathbb{R}^{L \times (m+1)} \rightarrow \mathbb{R}$, параметризованную вектором θ , которая минимизирует ошибку прогноза на одношаговом горизонте:

$$\hat{y}_t = f_\theta(X_t),$$

где \hat{y}_t – прогнозируемое моделью значение уровня воды в день t .

Параметры модели θ находятся в процессе обучения путём минимизации эмпирического риска – среднеквадратичной ошибки (MSE) на множестве всех доступных последовательностей обучающей выборки \mathcal{D}_{train} :

$$\theta^* = \arg \min_{\theta} L(\theta) = \arg \min_{\theta} \frac{1}{|\mathcal{D}_{train}|} \sum_{(X_t, y_t) \in \mathcal{D}_{train}} (y_t - f_\theta(X_t))^2.$$

В качестве f_θ предлагается использовать нейронную сеть с архитектурой Transformer, адаптированную для регрессии временных рядов. Модель состоит из:

1) Блока проекции и позиционного кодирования. Входная матрица X_t линейно проецируется в пространство размерности d_{model} . К полученным векторам поэлементно

добавляется детерминированное синусоидальное позиционное кодирование $P \in \mathbb{R}^{L \times d_{model}}$, чтобы передать модели информацию о временном порядке: $H^{(0)} = X_t W_{proj} + P$.

2) Стека энкодеров. Состояние $H^{(0)}$ последовательно обрабатывается N идентичными слоями энкодера. Каждый слой l состоит из:

– блока многоголового самовнимания (Multi-Head Self-Attention, MHSA):

$$head_i = \text{Attention}(H^{(l-1)}W_i^Q, H^{(l-1)}W_i^K, H^{(l-1)}W_i^V), i = 1, \dots, h$$

$$\text{MHSA}(H^{(l-1)}) = \text{Concat}(head_1, \dots, head_h)W^O,$$

где W_i^Q , W_i^K , W_i^V , W^O – обучаемые матрицы весов, а функция Attention определена стандартным образом [19].

– покомпонентный полносвязный слой (Position-wise Feed-Forward Network, FFN):

$$\text{FFN}(z) = \max(0, zW_1 + b_1)W_2 + b_2;$$

– оба подблока используют остаточные соединения (residual connections) и слоевую нормализацию (layer norm).

3) Регрессионной зависимости. После последнего слоя энкодера получается обогащённое контекстом представление всей последовательности $H^{(N)} \in \mathbb{R}^{L \times d_{model}}$. Для прогноза используется только вектор последнего временного шага $h_L^{(N)}$ (или агрегация всех шагов). Он пропускается через полносвязный слой для получения итогового скалярного прогноза:

$$\hat{y}_t = h_L^{(N)} w_{out} + b_{out}.$$

Следовательно, задача сводится к обучению параметров:

$$\theta = \{W_{proj}, W_i^Q, W_i^K, W_i^V, W^O, W_1, b_1, W_2, b_2, w_{out}, b_{out}, \dots\}$$

трансформерной сети f_θ на исторических данных для минимизации ошибки предсказания \hat{y}_t относительно фактического значения y_t .

Методология и алгоритмы

Для прогнозирования уровней воды в створе р. Пур – с. Самбург написана программа на языке Python, которая реализует трансформерную нейронную сеть (Transformer Neural Network), адаптированную для задач прогнозирования временных рядов. Предложенная архитектура сохраняет ключевые принципы оригинальной модели трансформера, одновременно учитывая специфику временных данных [13–18].

Основные компоненты архитектуры включают:

1. Энкодер трансформера (Transformer Encoder), оснащённый многоголовым механизмом внимания (multi-head self-attention), что позволяет модели одновременно улавливать различные аспекты зависимости между элементами входной последовательности.

2. Позиционное кодирование (positional encoding), интегрированное во входные представления для явного учёта временного порядка наблюдений, поскольку исходная архитектура трансформера не содержит рекуррентных или свёрточных компонентов, способных самостоятельно учитывать последовательность.

3. Полносвязные (dense) слои на выходе энкодера, предназначенные для преобразования скрытых представлений в конечный прогнозируемый вектор.

Структура обеспечивает эффективное моделирование как локальных, так и глобальных временных зависимостей в данных, что важно при прогнозировании многомерных или долгосрочных временных рядов.

Энкодер – это основная часть архитектуры трансформера, отвечающая за анализ входной последовательности (в данном случае – временного ряда) и извлечение из неё значимых зависимостей. В отличие от рекуррентных сетей (RNN, LSTM), энкодер трансформера обрабатывает все элементы последовательности параллельно, что делает обучение значительно быстрее.

Стандартный энкодер состоит из нескольких одинаковых слоёв (блоков), каждый из которых включает:

- механизм многоголового самовнимания (multi-head self-attention), позволяющий каждому элементу «взаимодействовать» с другими;
- полносвязный (feed-forward) нейросетевой блок, применяющий нелинейные преобразования к каждому элементу независимо.

Эти компоненты снабжены механизмами остаточного соединения (residual connection) и нормализации по слоям (layer normalization), что улучшает стабильность и сходимость обучения.

В задачах прогнозирования временных рядов энкодер принимает на вход последовательность прошлых значений и формирует для каждого шага «богатое» скрытое представление, учитывающее контекст всей последовательности.

Архитектура трансформера не содержит рекуррентных или свёрточных структур, поэтому по умолчанию она не знает порядка элементов во входной последовательности – для неё все токены (временные точки) «равноправны». Однако во временных рядах порядок критически важен. Чтобы передать модели информацию о позиции каждого элемента в последовательности, к входным векторам добавляется позиционное кодирование – специальный вектор, зависящий только от индекса временного шага.

Центральным элементом предложенной модели является механизм самовнимания (self-attention), позволяющий каждому элементу входной последовательности взаимодействовать с другими элементами с учётом их относительной значимости. Формально механизм самовнимания определяется следующим выражением:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V,$$

где $Q \in \mathbb{R}^{n \times d_k}$ – матрица запросов (queries); $K \in \mathbb{R}^{n \times d_k}$ – матрица ключей (keys); $V \in \mathbb{R}^{n \times d_v}$ – матрица значений (values); d_k – размерность векторов ключей и запросов, n – длина входной последовательности.

Деление на d_k в аргументе функции softmax применяется для стабилизации градиентов при обучении, предотвращая чрезмерно большие значения скалярных произведений при увеличении размерности. Полученные веса внимания определяют, насколько сильно каждый элемент последовательности должен «внимать» другим элементам при формировании своего скрытого представления.

Самовнимание – это механизм, позволяющий каждому элементу последовательности взвешенно «внимать» другим элементам, включая самого себя. Механизм вычисляет степень релевантности между всеми парами элементов и формирует новое представление каждого элемента как взвешенную сумму всех остальных.

В рамках механизма самовнимания каждый элемент входной последовательности генерирует вектор запроса Q , который формализует информационные потребности данного элемента относительно контекста всей последовательности.

Математически, для входной последовательности длиной n и размерности модели d матрица запросов формируется как:

$$Q = X \cdot W^Q,$$

где $X \in \mathbb{R}^{n \times d}$ – матрица входных эмбеддингов; $W^Q \in \mathbb{R}^{d \times d_k}$ – обучаемая матрица весов проекции.

Семантически, каждый вектор запроса q_i кодирует репрезентацию, определяющую, какие аспекты других элементов последовательности наиболее релевантны для текущей позиции i . В контексте временных рядов, данный механизм позволяет идентифицировать структурные аналогии и паттерны, распределенные во временной области.

Параллельно с формированием запросов, каждый элемент последовательности проецируется в пространство ключей посредством матрицы K :

$$K = X \cdot W^K,$$

где $W^K \in \mathbb{R}^{d \times d_k}$ – обучаемая матрица параметров.

Ключевые векторы k_j выполняют функцию дескрипторов содержательной информации, доступной в каждой позиции j . В процессе вычисления осуществляется попарное сравнение запроса q_i со всеми ключами k_j через операцию скалярного произведения:

$$e_{ij} = \frac{q_i \cdot k_j^\top}{\sqrt{d_k}}.$$

Нормализация на $\sqrt{d_k}$ стабилизирует градиенты в процессе обучения, компенсируя рост величины скалярного произведения с увеличением размерности.

Третьей компонентой механизма внимания выступает матрица значений:

$$V = X \cdot W^V,$$

где $W^V \in \mathbb{R}^{d \times d_v}$ – параметры проекции.

Векторы значений v_j включают непосредственно информационное содержание, которое будет агрегироваться в выходное представление. После вычисления матрицы коэффициентов внимания $A = \text{softmax}(E)$ осуществляется взвешенное суммирование:

$$z_i = \sum_j a_{ij} \cdot v_j,$$

где a_{ij} – элементы матрицы внимания, отражающие степень релевантности элемента j для позиции i .

Механизм внимания реализует адаптивную процедуру контекстуализации, где каждое выходное представление формируется как взвешенная комбинация всех элементов входной последовательности, с весами, определяемыми совместимостью соответствующих запросов и ключей.

В рамках многоголового внимания (multi-head attention) вычисления выполняются параллельно в нескольких «головах», каждая из которых обучается на различных линейных проекциях входных данных. Результаты объединяются и дополнительно преобразуются линейным слоем, что значительно расширяет выразительную способность модели и позволяет ей улавливать разнородные зависимости в данных.

Структура данных и предобработка

В рамках исследования использованы два типа данных – целевые и входные временные ряды.

Целевые данные представлены в файле `target_data.csv`. Данный файл содержит одномерный временной ряд, отражающий среднесуточные уровни воды за навигационный период в исследуемом створе по годам. Данные получены из гидрологических ежегодников [19] и Автоматизированной информационной системы государственного мониторинга водных объектов (АИС ГМВО)¹.

Входные данные организованы в виде серии файлов `data_YYYY.csv`, где YYYY обозначает год наблюдений (например, `data_2020.csv`, `data_2021.csv` и т.д.). Каждый файл содержит многомерный временной ряд, состоящий из 12 признаков, характеризующих гидрометеорологические факторы, потенциально влияющие на уровень воды. Предикторный горизонт охватывает данные за предшествующие сезоны, в частности среднесуточные значения за период с 1 сентября предшествующего года до 28 февраля года, для которого создается прогноз. В качестве входных данных используются следующие параметры по метеорологической станции Тарко-Сале: температура воздуха; атмосферные осадки; средняя скорость ветра; высота снежного покрова; температура поверхности почвы; атмосферное давление на уровне станции; дефицит насыщения водяного пара; относительная влажность; температура точки росы [20]. Дополнительно используются индекс арктической осцилляции (АО), индекс Северо-Атлантического колебания (NAO) и ежедневные данные по уровню воды р. Пур в створе Уренгой за тот же промежуток времени. Индексы североатлантического колебания (NAO) и арктического колебания (АО) получены с сайтов https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/ao.shtml и <https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml>.

Все временные ряды имеют одинаковую частоту дискретизации (суточные измерения), что обеспечивает согласованность при построении обучающих последовательностей.

Для обеспечения корректного функционирования нейросетевой модели и повышения качества прогноза реализован комплексный конвейер предобработки данных, включающий следующие этапы:

А. Парсинг временных меток. В исходных данных даты представлены в нестандартном формате «01.сен», «15.окт» и т.п. с использованием сокращённых названий месяцев на русском языке. Для корректного преобразования таких строк в объекты даты реализована процедура парсинга с явным указанием локали (`locale='ru_RU.UTF-8'`) и соответствующего шаблона («%d.%b»). Это позволило точно сопоставить каждое наблюдение с календарной датой и обеспечить корректную временную упорядоченность данных.

Б. Обработка пропущенных значений и артефактов. Анализ выявил наличие пропущенных значений (NaN), а также отдельных аномальных измерений (например, физически невозможные резкие скачки, не соответствующие динамике процесса). Для устранения этих проблем применён следующий подход:

¹ Автоматизированная информационная система государственного мониторинга водных объектов. URL: <https://gmvo.skniivh.ru> (дата обращения: 01.09.2025).

- пропущенные значения в признаковых рядах интерполированы линейно с учётом временной структуры;
- аномалии в целевом ряде идентифицированы с помощью метода межквартильного размаха (IQR) и заменены на значения, полученные с использованием скользящего окна (rolling window) с медианной фильтрацией;
- в случае длительных пропусков (более 5 последовательных дней) соответствующие участки исключены из обучающей выборки.

В. Нормализация признаков. Для стабилизации процесса обучения и ускорения сходимости модели все входные признаки (12 временных рядов) нормализованы с использованием StandardScaler из библиотеки scikit-learn. Данный метод преобразует каждый признак к нулевому среднему и единичной дисперсии:

$$x' = \frac{x - \mu}{\sigma},$$

где μ и σ – среднее и стандартное отклонение, вычисленные только по обучающей выборке, чтобы избежать утечки информации из валидационной и тестовой частей.

Г. Формирование обучающих последовательностей. Для адаптации данных к архитектуре трансформера, работающей с фиксированной длиной контекста, исходные временные ряды были преобразованы в набор оконных последовательностей. Каждая последовательность включает: 60 последовательных временных шагов входных признаков (матрица размером 60×12); соответствующее целевое значение – уровень воды на следующем шаге.

Представленный подход позволяет модели извлекать как краткосрочные, так и среднесрочные зависимости в данных. Общее количество сформированных последовательностей зависит от длины исходного ряда и шага сдвига окна (в работе использован шаг, равный 1, для максимального использования данных).

В результате описанной предобработки получен структурированный, нормализованный и очищенный набор данных, пригодный для обучения и оценки нейросетевой модели прогнозирования уровней воды.

Реализация модели

Для решения задачи прогнозирования уровней воды разработана специализированная нейросетевая модель на основе архитектуры трансформера, реализованная в виде класса TransformerTimeSeries, унаследованного от базового класса nn.Module библиотеки PyTorch. Архитектура модели адаптирована под специфику многомерных временных рядов и включает следующие компоненты.

1. Проекция входных признаков. Входные данные представляют собой последовательность длиной 60 временных шагов с 12 признаками на каждом шаге (тензор размерности 60×12). Для приведения входного пространства к размерности, требуемой трансформером (d_{model}), применяется линейный слой (полносвязный), осуществляющий проекцию каждого временного шага из 12-мерного пространства в d_{model} -мерное и позволяющий эффективно обрабатывать многомерные входы в рамках стандартной архитектуры трансформера.

2. Позиционное кодирование. Поскольку архитектура трансформера не содержит рекуррентных или свёрточных механизмов, явно учитывающих порядок элементов, для передачи информации о временной позиции каждого шага используется синусоидальное позиционное кодирование, предложенное в работе [17]. Векторы позиционного кодирования предварительно вычисляются для максимальной длины последовательности и добавляются к проектированным входным представлениям до подачи в энкодер.

3. Энкодер трансформера. Основу модели составляет стек из N идентичных слоёв энкодера трансформера. Каждый слой включает:

- многоголовый механизм самовнимания (multi-head self-attention), позволяющий модели одновременно улавливать различные типы временных зависимостей (локальные, сезонные, долгосрочные);

- полносвязную промежуточную сеть (position-wise feed-forward network) с нелинейной активацией;

- остаточные соединения (residual connections) и нормализацию по слоям (layer normalization) вокруг каждого подблока для стабилизации градиентов и ускорения сходимости.

4. Декодер для регрессионного прогноза. В отличие от классического трансформера, используемого в задачах машинного перевода, в данной работе декодер в традиционном смысле отсутствует. Вместо него применяется упрощённый регрессионный «головной» модуль, который преобразует скрытые представления, полученные на выходе энкодера, в конечный прогноз. Конкретно, используется агрегация (в текущей реализации – взятие последнего временного шага выхода энкодера) с последующим применением одного или нескольких полносвязных слоёв, завершающихся линейным выходом, соответствующим прогнозируемому значению (уровню воды).

Модель полностью дифференцируема и обучается с использованием стандартных методов оптимизации (Adam) и функции потерь (среднеквадратичная ошибка, MSE).

Архитектура и гиперпараметры модели подобраны на основе компромисса между выразительной способностью (expressive capacity), вычислительной сложностью и объёмом доступных данных. Использованы значения, представленные в табл. 1.

Таблица 1

Параметры конфигурации модели

Параметр	Значение
Длина входной последовательности	60 временных шагов
Размерность модели (d_{model})	128
Количество голов внимания	8
Количество слоёв энкодера	4
Вероятность dropout	0.2

Длина последовательности (60) выбрана с учётом характерного периода гидрологических процессов (например, двухмесячного окна для учёта осадков и таяния снега).

Размерность модели (128) обеспечивает достаточную ёмкость для кодирования 12 входных признаков и их взаимодействий, не приводя к чрезмерному переобучению на имеющемся объёме данных.

8 голов внимания позволяют модели одновременно фокусироваться на различных аспектах временной структуры (например, краткосрочные колебания и долгосрочные тренды).

4 слоя энкодера обеспечивают иерархическое извлечение признаков: от локальных зависимостей на нижних уровнях до глобальных паттернов на верхних.

Dropout с вероятностью 0,2 применяется после каждого подблока внимания и после полносвязных слоёв для регуляризации и повышения обобщающей способности модели.

Все параметры зафиксированы в конфигурационном файле и могут быть легко модифицированы для проведения экспериментов с масштабированием архитектуры или адаптации к другим временным рядам.

Обучение и валидация модели

Процесс обучения модели TransformerTimeSeries организован с использованием современных практик глубокого обучения, направленных на обеспечение стабильной сходимости, предотвращение переобучения и достижение высокого качества прогноза.

В качестве алгоритма оптимизации выбран адаптивный метод Adam (Adaptive Moment Estimation), сочетающий преимущества методов AdaGrad и RMSProp. Adam эффективно работает с разреженными градиентами и обеспечивает быструю сходимость даже при нестационарных целевых функциях. Начальная скорость обучения (learning rate) установлена равной 0,0005, что позволило избежать неустойчивости на начальных этапах обучения при сохранении достаточной скорости продвижения к оптимуму.

Поскольку задача формулируется как регрессионное прогнозирование уровня воды, в качестве целевой функции использована среднеквадратичная ошибка (Mean Squared Error, MSE):

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2,$$

где y_i – истинное значение уровня воды; \hat{y}_i – прогноз модели; N – размер батча. MSE чувствительна к крупным ошибкам, что способствует более точному моделированию экстремальных гидрологических событий.

Для снижения риска переобучения применена L2-регуляризация (также известная как weight decay), встроенная непосредственно в оптимизатор Adam. Коэффициент регуляризации установлен равным $1e-4$, что обеспечивает мягкое штрафование больших весов без чрезмерного подавления выразительной способности модели.

Для адаптации скорости обучения в процессе тренировки использован ReduceLROnPlateau – стратегия, которая снижает learning rate, если метрика на валидационной выборке перестает улучшаться в течение заданного числа эпох («плато»). Конкретные параметры: фактор уменьшения – 0,5, patience (число эпох ожидания) – 5, минимальный learning rate – $1e-7$. Такой подход позволяет модели «замедляться» вблизи оптимума и избегать пропуска узких минимумов функции потерь.

Обучение проводилось с использованием мини-батчей фиксированного размера (32), что обеспечивало стабильность градиентных оценок и эффективное использование памяти GPU.

Для объективной оценки обобщающей способности модели и предотвращения переобучения реализован строгий протокол валидации.

Исходный временной ряд разделён на обучающую и валидационную выборки в соотношении 80 %: 20 % с сохранением хронологического порядка (то есть без случайного перемешивания). Такой подход соответствует реалистичному сценарию прогнозирования: модель обучается на прошлых данных и тестируется на будущих, что исключает утечку информации из будущего в прошлое.

В процессе обучения после каждой эпохи вычислялась метрика качества на валидационной выборке (MSE). Модель, продемонстрировавшая наилучшее значение валидационной ошибки, автоматически сохранялась на диск в формате состояния PyTorch (state_dict).

Для мониторинга сходимости и диагностики проблем (переобучение, недообучение, неустойчивость) реализована визуализация динамики обучения. После завершения тренировки строятся графики:

- значения функции потерь (MSE) на обучающей и валидационной выборках в зависимости от номера эпохи;
- изменение скорости обучения в ходе работы планировщика.

Эти графики включены в отчёт как часть анализа поведения модели и служат основой для принятия решений о необходимости модификации архитектуры или гиперпараметров.

Метрики качества

Для объективной количественной оценки точности прогнозирования уровня воды программой вычисляется набор стандартных метрик регрессионного анализа, каждая из которых характеризует ошибку модели с различных точек зрения. Использование нескольких метрик позволяет получить комплексное представление о качестве прогноза, учитывая как абсолютную величину отклонений, так и их относительную значимость в контексте масштаба целевой переменной.

1. Mean Absolute Error (MAE) – средняя абсолютная ошибка. MAE представляет собой среднее арифметическое абсолютных разностей между прогнозируемыми и фактическими значениями:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

где y_i – истинное значение уровня воды; \hat{y}_i – прогноз модели; N – количество наблюдений.

MAE выражается в тех же единицах, что и целевая переменная (например, сантиметрах уровня воды), и устойчива к выбросам по сравнению с MSE/RMSE, поскольку не возводит ошибки в квадрат, и, следовательно, лучше отражает типичную ошибку прогноза.

2. Root Mean Square Error (RMSE) – корень из среднеквадратичной ошибки. RMSE определяется как квадратный корень из среднего значения квадратов ошибок:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

Как и MAE, RMSE измеряется в единицах целевой переменной, однако она сильнее штрафует крупные ошибки из-за квадратичной зависимости. RMSE служит индикатором рискованных прогнозов и чувствительности модели к экстремальным событиям.

3. Mean Absolute Percentage Error (MAPE) – средняя абсолютная процентная ошибка. MAPE выражает среднюю абсолютную ошибку в процентах от истинного значения:

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

MAPE позволяет оценить относительную точность прогноза, независимо от абсолютного масштаба уровня воды. Однако следует учитывать, что MAPE не определена при $y_i=0$ и может быть нестабильной при малых значениях y_i .

Использование всех трёх метрик в совокупности позволяет оценить среднюю ошибку в физических единицах (MAE); выявить влияние крупных отклонений и экстремальных событий (RMSE); проанализировать относительную точность прогноза, независимо от масштаба (MAPE).

Все метрики автоматически сохраняются в лог-файл и визуализируются в виде таблиц и графиков совместно с прогнозами, что упрощает интерпретацию результатов и принятие решений на основе модели.

Результаты прогнозирования

По завершении процесса обучения и оценки программа автоматически формирует следующий комплекс выходных артефактов, обеспечивающих воспроизводимость, интерпретируемость и практическую применимость результатов.

Обученная модель записывается в файл `best_transformer_model.pth`, содержащий сохранённые веса нейросетевой архитектуры `TransformerTimeSeries`, соответствующие наилучшей эпохе по метрике валидации. Модель может быть повторно загружена для последующего инференса, дообучения или интеграции в операционные системы мониторинга водных ресурсов.

Прогнозные значения записываются в файл `water_level_predictions.csv` в формате CSV с разделителем «точка с запятой», включающий временные метки и соответствующие прогнозируемые уровни воды. Файл структурирован таким образом, чтобы обеспечить совместимость с существующими гидрологическими базами данных и системами визуализации.

Метрики качества по результатам работы программы на тестовой выборке составили: MAE 10,76 см за навигационный период, RMSE 14,00 см, MAPE 2,69 %.

Результаты работы программы представлены на рис. 1, 2. На рис. 3 приведен пример прогнозирования уровней воды р. Пур в створе с. Самбург за навигационный период 2021 г.

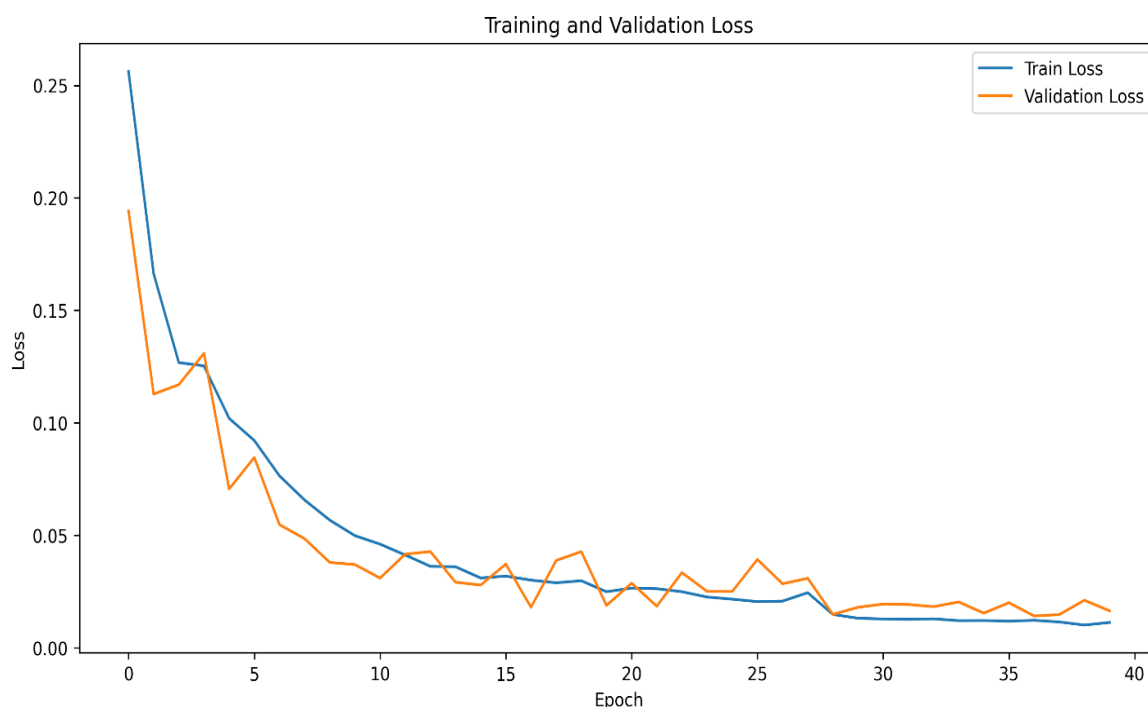


Рис. 1. Кривые функции потерь (MSE) на обучающей и валидационной выборках по эпохам

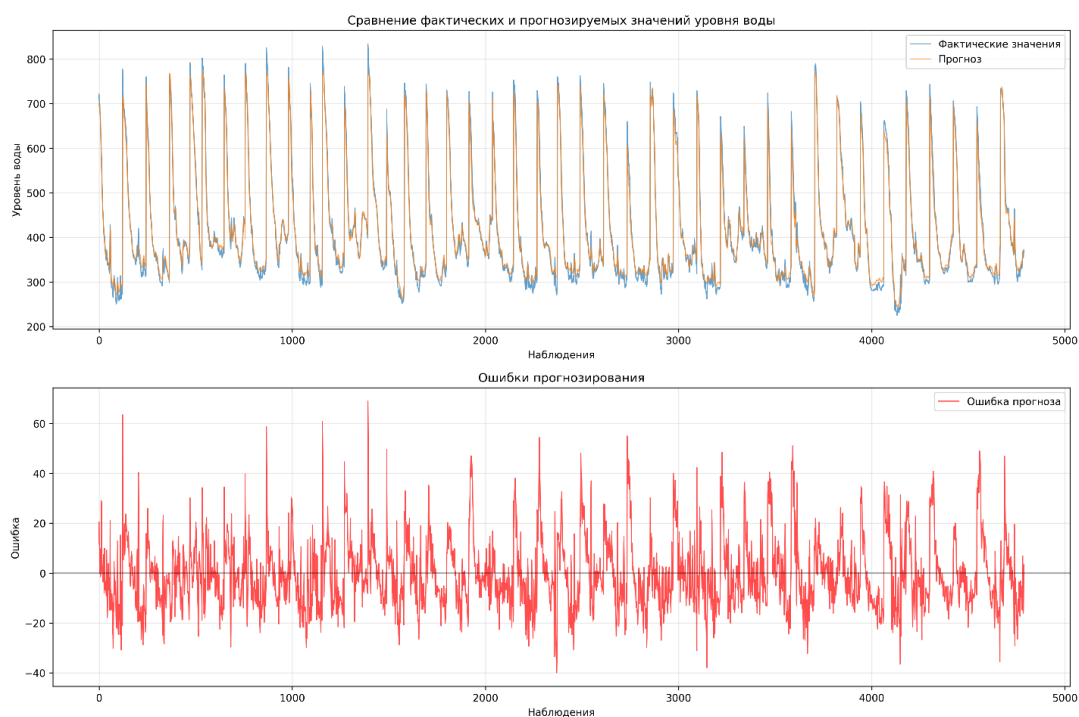


Рис. 2. Визуальное сопоставление истинных и прогнозируемых значений уровня воды и ошибки прогнозирования по р. Пур – с. Самбург

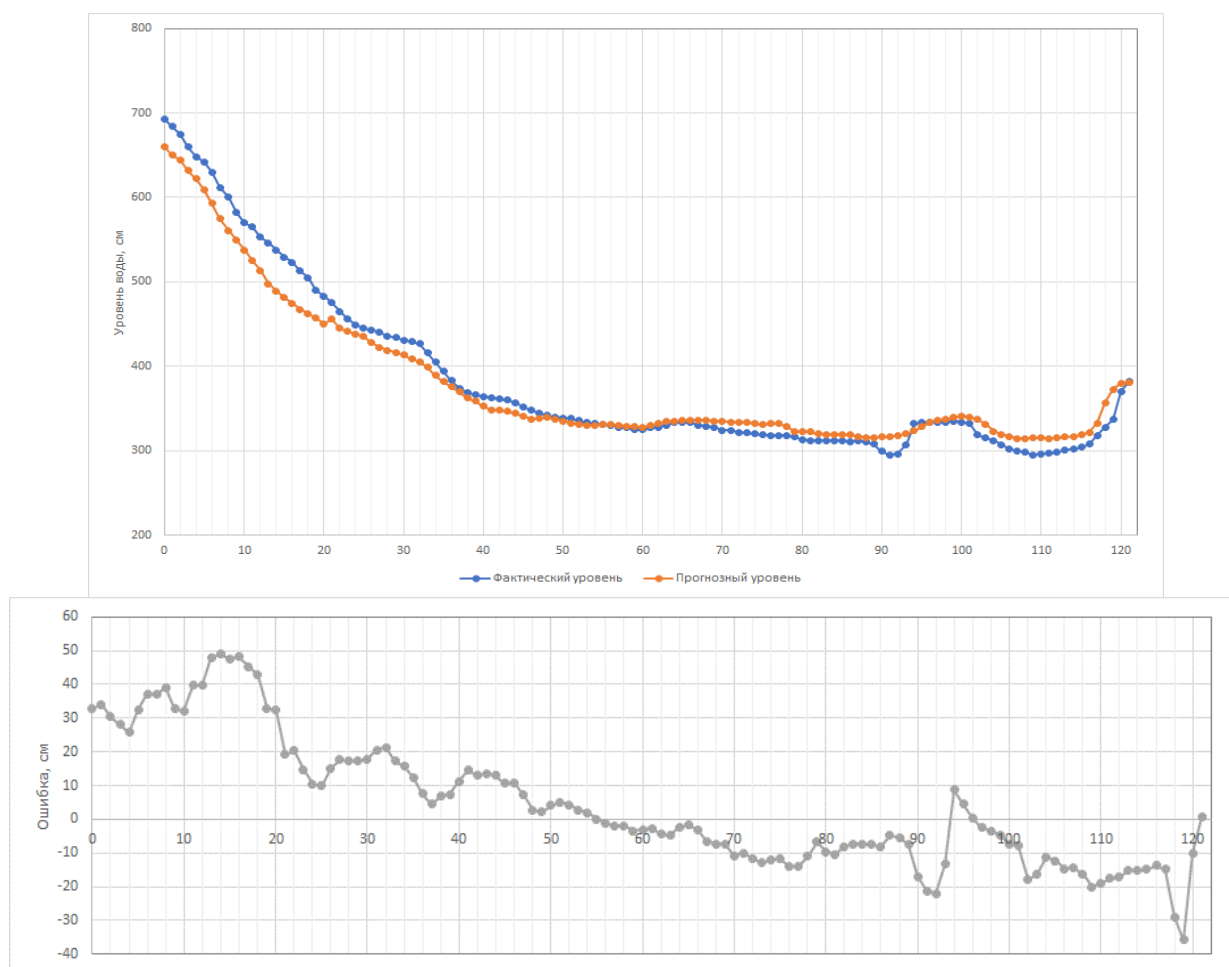


Рис. 3. Пример прогноза уровней воды за навигационный период за 2021 г. по р. Пур – с. Самбург. MAE 15,22 см, RMSE 19,46 см, MAPE 3,70 %

По результатам прогнозирования на примере 2021 г. проведён анализ точности гидрологической модели с разбивкой по трём функциональным диапазонам: высоким, средним и низким уровням воды, имеющим различное значение для судоходства и управления гидрологическими рисками.

Высокие уровни воды (опасный период, 370–693 см) характеризуются хорошей точностью прогноза, однако наблюдается систематическое запаздывание в определении пиковых значений. Самое значительное расхождение зафиксировано при максимальном уровне 693 см, который был недооценён моделью на 33 см (прогноз – 660 см). В то же время при уровне 382 см достигается высокая точность: прогноз составил 381 см, что свидетельствует об адекватной работе модели в пределах нижней границы опасного диапазона.

Средние уровни воды (нормальная навигация, 300–370 см) демонстрируют стабильно высокую точность прогнозирования. Средняя абсолютная ошибка в данном диапазоне составляет 5–15 см, что соответствует требованиям к оперативному планированию судоходства.

Низкие уровни воды (проблемы для судоходства, 295–300 см) характеризуются удовлетворительной точностью. Отмечена тенденция к небольшой переоценке: минимальный зафиксированный уровень (295 см) был спрогнозирован как 315 см, что может приводить к избыточной осторожности при планировании рейсов, но не создаёт критических рисков.

Модель проявляет ряд сильных сторон. В том числе она корректно отслеживает общую тенденцию спада уровней, точно прогнозирует стабильные периоды (например, дни 30–90) и адекватно реагирует на резкие изменения гидрологической обстановки (дни 119–121). Вместе с тем выявлены области, требующие улучшения: запаздывание реакции на резкие подъёмы уровней (дни 93–96 и 117–120); систематическая недооценка экстремальных паводковых пиков; незначительная переоценка уровней в периоды низкой воды.

Поскольку при использовании модели обеспечивается заблаговременность предупреждения за 3–4 мес., то точность определения пиковых значений (± 30 – 40 см) достаточна для принятия превентивных мер по защите инфраструктуры и населения. В условиях низких уровней воды точность прогноза в критическом диапазоне (± 20 см) позволяет оптимизировать грузоподъёмность судов и существенно снизить риски посадки на мель.

Заключение

Исследование посвящено решению актуальной научно-прикладной задачи – повышению точности и заблаговременности прогнозирования уровня режима арктических рек на примере р. Пур. В результате выполненной работы сформулированы следующие основные выводы, непосредственно аргументированные элементами научной новизны исследования:

1. Адаптированная трансформерная архитектура является эффективным инструментом для моделирования нелинейной динамики уровней воды в условиях Арктики. Разработанная модель на основе энкодера Transformer с механизмом многоголового самовнимания (8 голов, 4 слоя, $d_{model} = 128$) продемонстрировала способность улавливать сложные многомасштабные зависимости в данных. Высокие значения метрик качества ($MAE = 10,76$ см, $RMSE = 14,00$ см, $MAPE = 2,69$ %) на независимой тестовой выборке подтверждают, что предложенная архитектура успешно преодолевает ограничения традиционных методов (ARIMA, LSTM) при работе с нестационарными, зашумлёнными и мультифакторными рядами, характерными для арктического региона.

2. Формализация задачи как многомерной регрессии на длинных последовательностях с использованием контекстного окна в 60 сут позволяет модели выявлять скрытые кросс-корреляции между гидрометеорологическими факторами. Использование в качестве входных данных объединённой матрицы, включающей историю целевого уровня и 12 синхронных

признаков за двухмесячный период, обеспечило модели возможность обучаться не на изолированном ряде уровней, а на комплексном «слежке» состояния водосбора. Это позволило выявить и количественно оценить вклад таких факторов, как индексы АО/НАО и температура почвы, в формирование уровня режима, что ранее в оперативных моделях для данной реки не учитывалось в полной мере.

3. Разработанный метод интеграции прокси-данных с метеостанции Тарко-Сале обеспечил достаточную информативность модели при ограниченной сети наблюдений. Несмотря на отсутствие плотной сети гидрометеорологических постов в бассейне р. Пур, использование данных одной репрезентативной станции в качестве агрегированного сигнала для всего водосбора доказало свою эффективность. Модель, обученная на таких данных, успешно прогнозирует как фоновую сезонную динамику, так и реакцию реки на экстремальные метеособытия, что подтверждает корректность выбранного инженерного упрощения для средних арктических рек.

4. Детализированный анализ точности по функциональным диапазонам уровней выявил сильные стороны модели и направления для её дальнейшего совершенствования. Проведённая стратификация ошибок показала, что модель наиболее точна в диапазоне средних уровней (ошибка 5–15 см), важном для планирования навигации. Систематическая недооценка экстремальных пиков (запаздывание на 1–2 дня, ошибка до 33 см) указывает на необходимость модификации механизма внимания для лучшего захвата резких, кратковременных сигналов. При этом точность прогноза в опасном и судоходно-проблемном диапазонах (± 20 –40 см при заблаговременности 3–4 мес.) является операционно достаточной для принятия превентивных мер.

5. Созданный воспроизводимый конвейер предобработки данных (парсинг, интерполяция, нормализация, формирование окон) является существенным методическим вкладом и обеспечивает устойчивость модели к артефактам реальных наблюдений. Эффективность сложной нейросетевой архитектуры обеспечена за счёт тщательной подготовки данных. Реализованные методы обработки пропусков, фильтрации аномалий и хронологического разделения выборки позволили избежать переобучения и «утечки» будущего в прошлое, гарантировав объективность оценки. Данный конвейер может быть использован как типовое решение для других малоизученных арктических бассейнов.

Теоретическая значимость заключается в успешной трансфере передовой архитектуры машинного обучения (Transformer) в предметную область арктической гидрологии, что расширяет методологический аппарат дисциплины и открывает пути для создания более адекватных моделей в условиях меняющегося климата.

Практическая значимость состоит в создании работоспособного инструмента для поддержки принятия решений. Внедрение модели позволит управлениям по гидрометслужбе повысить заблаговременность и надёжность предупреждений о паводках, судоходным компаниям оптимизировать логистику и минимизировать риски аварий, органам управления реализовать проактивные стратегии защиты прибрежной инфраструктуры и населения, сократить потенциальный социально-экономический ущерб.

Перспективы исследований лежат в области интеграции спутниковых данных, разработки гибридных физико-статистических моделей и применения методов Explainable AI (ХАИ) для интерпретации внутренних механизмов принятия решений трансформерной сетью.

Список источников

1. Харлампыева Н.К., Третьяков М.В., Иванов В.В. Арктическая устьевая гидрология: история исследования // География и краеведение в Якутии и регионах России: материалы Всероссийской науч.-практ. конф. с междунар. участием, посвященной 85-летию со дня рождения Г.Н. Максимова. Якутск: Издательский дом СВФУ, 2024. С. 115–120. EDN GUMXFY.

2. Ресурсы поверхностных вод СССР: Гидрологическая изученность. Л.: Гидрометеиздат, 1964. Т. 15. Вып. 3.

3. Time Series Analysis: Forecasting and Control / G.E.P. Box [et al.] // John Wiley & Sons, 2015. 720 p.
4. Hochreiter S., Schmidhuber J. Long short-term memory // *Neural computation*. 1997. Т. 9. № 8. С. 1735–1780.
5. Learning phrase representations using RNN encoder-decoder for statistical machine translation / K. Cho [et al.] // *arXiv preprint arXiv:1406.1078*. 2014.
6. Rainfall–runoff modelling using long short-term memory (LSTM) networks / F. Kratzert [et al.] // *Hydrology and Earth System Sciences*. 2018. Т. 22. № 11. С. 6005–6022.
7. Attention is all you need / A. Vaswani [et al.] // *Advances in neural information processing systems*. 2017. Т. 30. № 1. С. 5998–6008.
8. Deep transformer models for time series forecasting: The influenza prevalence case / N. Wu [et al.] // *arXiv preprint arXiv:2001.08317*. 2020.
9. Informer: Beyond efficient transformer for long sequence time-series forecasting / H. Zhou [et al.] // *Proceedings of the AAAI conference on artificial intelligence*. 2021. Т. 35. № 12. С. 11106–11115.
10. Волкова Н.А. Анализ многолетних колебаний уровня воды реки Пур для обеспечения безопасности судоходства // *Природные и техногенные риски (физико-математические и прикладные аспекты)*. 2025. № 2 (54). С. 6–22. DOI: 10.61260/2307-7476-2025-2-6-22
11. Волкова Н.А. Комплексный подход к снижению аварийности на внутренних водных путях арктического региона России // *Известия Петербургского университета путей сообщения*. 2025. Т. 22. № 3. С. 761–775. DOI: 10.20295/1815-588X-2025-3-761-775
12. Волкова Н.А., Ромашова К.В. Методика долгосрочного прогнозирования максимальных уровней воды реки Таз // *Техносферная безопасность в Арктике: материалы в рамках VIII Междунар. науч.-практ. конф.* СПб.: С.-Петерб. ун-т ГПС МЧС России, 2025. С. 223–226. EDN RFVHDX.
13. Deep quantum-transformer networks for multimodal beam prediction in isac systems / S. Tariq [et al.] // *IEEE Internet of Things Journal*. 2024. Т. 11. № 18. С. 29387–29401.
14. Diformer: A difference transformer network for remote sensing change detection / H. Lin [et al.] // *IEEE Geoscience and Remote Sensing Letters*. 2024. Т. 21. С. 1–5.
15. Никитин С.Д., Мамонтов А.И. О сжатии нейросетей с архитектурой «Трансформер» // *Измерение, контроль, информатизация: материалы XXVI Междунар. науч.-техн. конф.* / под ред. Л.И. Сучковой. Барнаул: Изд-во АлтГТУ, 2025. 216 с.
16. Ключовкин Г.К. Нейронные сети против классических ML-моделей: в каких случаях стоит усложнять архитектуру // *Актуальные исследования*. 2024. № 37 (219).
17. Attention is all you need / A. Vaswani [et al.] // *Advances in Neural Information Processing Systems*. 2017. № 30. P. 5998–6008.
18. How to optimize the systematic review process using AI tools / N. Fabiano [et al.] // *JCPP advances*. 2024. Т. 4. № 2. С. e12234.
19. Гидрологический ежегодник. Л., Гидрометеиздат, 1936–2012. Т. 6. Вып. 0-9.
20. Свидетельство о государственной регистрации базы данных № 2020621470 Рос. Федерация. «Научно-прикладной справочник «Климат России»: № 2020620899: заявл. 09.06.2020; опубли. 18.08.2020 / В.Н. Разуваев, О.Н. Булыгина, Н.Н. Коршунова [и др.]; заявитель федеральное государственное бюджетное учреждение «Всероссийский научно-исследовательский институт гидрометеорологической информации – Мировой центр данных». EDN MXYBEY.

References

1. Harlamp'eva N.K., Tret'yakov M.V., Ivanov V.V. Arkticheskaya ust'evaya gidrologiya: istoriya issledovaniya // *Geografiya i kraevedenie v Yakutii i regionah Rossii: materialy Vserossijskoj nauch.-prakt.konf.s mezhdunar. uchastiem, posvyashchennoj 85-letiyu so dnya rozhdeniya G.N. Maksimova*. Yakutsk: Izdatel'skij dom SVFU, 2024. S. 115–120. EDN GUMXFY.

2. Resursy poverhnostnyh vod SSSR: Gidrologicheskaya izuchennost'. L.: Gidrometeoizdat, 1964. T. 15. Vyp. 3.
3. Time Series Analysis: Forecasting and Control / G.E.P. Box [et al.] // John Wiley & Sons, 2015. 720 p.
4. Hochreiter S., Schmidhuber J. Long short-term memory // Neural computation. 1997. T. 9. № 8. S. 1735–1780.
5. Learning phrase representations using RNN encoder-decoder for statistical machine translation / K. Cho [et al.] // arXiv preprint arXiv:1406.1078. 2014.
6. Rainfall–runoff modelling using long short-term memory (LSTM) networks / F. Kratzert [et al.] // Hydrology and Earth System Sciences. 2018. T. 22. № 11. S. 6005–6022.
7. Attention is all you need / A. Vaswani [et al.] // Advances in neural information processing systems. 2017. T. 30. № 1. S. 5998–6008.
8. Deep transformer models for time series forecasting: The influenza prevalence case / N. Wu [et al.] // arXiv preprint arXiv:2001.08317. 2020.
9. Informer: Beyond efficient transformer for long sequence time-series forecasting / H. Zhou [et al.] // Proceedings of the AAAI conference on artificial intelligence. 2021. T. 35. № 12. S. 11106–11115.
10. Volkova N.A. Analiz mnogoletnih kolebanij urovnya vody reki Pur dlya obespecheniya bezopasnosti sudohodstva // Prirodnye i tekhnogennye riski (fiziko-matematicheskie i prikladnye aspekty). 2025. № 2 (54). S. 6–22. DOI 10.61260/2307-7476-2025-2-6-22
11. Volkova N.A. Kompleksnyj podhod k snizheniyu avarijnosti na vnutrennih vodnyh putyah arkticheskogo regiona Rossii // Izvestiya Peterburgskogo universiteta putej soobshcheniya. 2025. T. 22. № 3. S. 761–775. DOI: 10.20295/1815-588X-2025-3-761-775
12. Volkova N.A., Romashova K.V. Metodika dolgosrochnogo prognozirovaniya maksimal'nyh urovnej vody reki Taz // Tekhnosfermaya bezopasnost' v Arktike: materialy v ramkah VIII Mezhdunar. nauch.-prakt. konf. SPb.: S.-Peterb. un-t GPS MCHS Rossii, 2025. S. 223–226. EDN RFVHDX.
13. Deep quantum-transformer networks for multimodal beam prediction in isac systems / S. Tariq [et al.] // IEEE Internet of Things Journal. 2024. T. 11. № 18. S. 29387–29401.
14. Difformer: A difference transformer network for remote sensing change detection / H. Lin [et al.] // IEEE Geoscience and Remote Sensing Letters. 2024. T. 21. S. 1–5.
15. Nikitin S.D., Mamontov A.I. O szhatii nejrosetej s arhitekturoj «Transformer» // Izmerenie, kontrol', informatizaciya: materialy XHVI Mezhdunar. nauch.-tekhn. konf. / pod red. L.I. Suchkovej. Barnaul: Izd-vo AltGTU, 2025. 216 c.
16. Klyukovkin G.K. Nejronnye seti protiv klassicheskikh ML-modelej: v kakih sluchayah stoit uslozhnyat' arhitekturu // Aktual'nye issledovaniya. 2024. № 37 (219).
17. Attention is all you need / A. Vaswani [et al.] // Advances in Neural Information Processing Systems. 2017. № 30. P. 5998–6008.
18. How to optimize the systematic review process using AI tools / N. Fabiano [et al.] // JCPP advances. 2024. T. 4. № 2. S. e12234.
19. Gidrologicheskij ezhegodnik. L., Gidrometeoizdat, 1936–2012. T. 6. Vyp. 0-9.
20. Svidetel'stvo o gosudarstvennoj registracii bazy dannyh № 2020621470 Ros. Federaciya. «Nauchno-prikladnoj spravochnik «Klimat Rossii»: № 2020620899: zayavl. 09.06.2020; opubl. 18.08.2020 / V.N. Razuvaev, O.N. Bulygina, N.N. Korshunova [i dr.]; zayavitel' federal'noe gosudarstvennoe byudzhethoe uchrezhdenie «Vserossijskij nauchno-issledovatel'skij institut gidrometeorologicheskoy informacii – Mirovoj centr dannyh». EDN MXYBEY.

Информация о статье:

Статья поступила в редакцию: 10.10.2025; одобрена после рецензирования: 10.12.2025;
принята к публикации: 21.01.2026

The information about article:

The article was submitted to the editorial office: 10.10.2025; approved after review: 10.12.2025;
accepted for publication: 21.01.2026

Информация об авторах:

Волкова Надежда Александровна, доцент кафедры водно-технических изысканий Российского государственного гидрометеорологического университета (192007, Санкт-Петербург, Воронежская ул., д. 79); старший научный сотрудник Арктического и антарктический научно-исследовательского института (199397, Санкт-Петербург, ул. Беринга, д. 38Г), кандидат физико-математических наук, e-mail: navolkova@aari.ru, <https://orcid.org/0000-0002-9272-4713>, SPIN-код: 3807-0526

Information about the authors:

Volkova Nadezhda A., associate professor of department of water-engineering surveys of Russian state hydrometeorological university (192007, Saint-Petersburg, Voronezhskaya st., 79); senior researcher of Arctic and Antarctic research institute (199397, Saint-Petersburg, Bering st., 38G), candidate of physico-mathematical sciences, e-mail: navolkova@aari.ru, <https://orcid.org/0000-0002-9272-4713>, SPIN: 3807-0526