

**Долгих В. П., Кусайко Н. П.
Донбасский государственный технический университет
E-mail: vidoscience@mail.ru

ПРИНЦИПЫ ПОСТРОЕНИЯ МЕТОДИКИ НЕЙРОСЕТЕВОГО РАСЧЕТА ХИМИЧЕСКИХ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ВОДЫ (НА ПРИМЕРЕ р. БЕЛАЯ, ЛНР)

Предложена методика нейросетевого расчета химических показателей качества воды. Приведены этапы расчета, которые учитывают мониторинговую информацию о рассчитываемом компоненте, наличие регламентированных расчетных методик, внешние факторы, оказывающие влияние на загрязнение водосборной территории, самообучение и проверку построенной модели на адекватность путем выполнения серии симуляций на данных, не входящих в обучающую выборку нейросети. На примере данных по сухому остатку апробирован алгоритм градиентного бустинга (CatBoost), погрешность которого в сравнении с экспериментальными данными составила 4,83 %. Алгоритм расчета учитывает влияние внешних факторов (метеоданных) и внутренних связей между другими компонентами.

Ключевые слова: речной сток, алгоритм градиентного бустинга (CatBoost), методика расчета химических показателей качества воды, математическое моделирование, сухой остаток.

Финансирование: исследования выполнены за счет средств федерального бюджета (код темы: FRRU-2024-0004 в ЕГИСУ НИОКТР).

Постановка проблемы, обоснование ее актуальности. Трансформация процессов создания, передачи, хранения и обработки данных, полученных в результате экологического мониторинга, является ключевым направлением в распоряжении Правительства Российской Федерации от 15 декабря 2023 г. № 3664-р [1]. Это направление также актуально для водосборной территории реки Белая, которая является правым притоком реки Лугань и входит в бассейн Северского Донца. Река Белая подвержена влиянию затопленных угольных шахт и потребительскому интересу со стороны промышленных предприятий и населения.

Учитывая множественные и многокомпонентные природные и антропогенные факторы, сосредоточение населения, проживающего вблизи водного объекта, оценка практической применимости цифровой трансформации в контексте локального экологического мониторинга водных объектов усложняется, но не теряет своей практической значимости. При разработке подобных систем необходимо учитывать методическое, информационное и приборное обеспе-

чение, с помощью которого достигается заданный уровень детализации получения и применения мониторинговой информации.

Гидрологические приложения, такие как расчетные программы, модели и методы, которые используются для планирования и управления водными ресурсами, опираются на репрезентативность гидрологических данных. Эти данные, представленные в виде временных рядов, являются основой для принятия решений, а также служат базой для прогнозирования и изучения гидрологических явлений [2]. Решение практических задач водопользования в промышленности и населенных пунктах требует комплексного подхода к анализу водных ресурсов. Необходимо учитывать не только гидрологические и гидрометеорологические характеристики водных объектов и их изменения во времени, но и химический состав речных и подземных вод, включая виды и концентрации загрязняющих веществ.

Вследствие повышенной техногенной нагрузки, которая сформировалась в результате закрытия угольных шахт Донбас-

са [3, 4], водосборная территория реки Белая характеризуется своими особенностями по виду и концентрации загрязняющих веществ. Данные обстоятельства имеют критически важное значение как для промышленных предприятий ООО «Южный горно-металлургический комплекс», так и для населения.

Проблемы создания и функционирования систем локального мониторинга загрязнения водных объектов, в которых уделяется внимание разработке методик расчета и оценки химических показателей качества воды, описаны в многочисленных публикациях, например, Слеймана Алаа [5], Е. А. Белозерова [6]. Авторы подчеркивают, что, несмотря на различия в оборудовании и методах анализа, сбор точных и детализированных данных о состоянии воды (с учетом пространственных, временных, количественных и качественных характеристик) представляет собой значительную трудность.

Закономерности распространения химических компонентов, попадающих в воду с подземными и поверхностными водами, осадками, отмечены в работах С. Е. Гулько [7], Ю. А. Рыбниковой [8], И. В. Смирновой [9] и др. Разработанные модели, отражающие теоретические или полуэмпирические подходы к описанию процессов появления того или иного компонента, характеризуются своей детерминированностью. Это в свою очередь приводит к рассмотрению ограниченного числа значимых параметров и локальному применению научных разработок для конкретного участка исследования.

Как показал краткий обзор литературных источников, общим методологическим недостатком подходов прошлых лет к работе систем локального мониторинга загрязнения водных объектов выступает их слабая цифровизация и автоматизация. В условиях растущих возможностей машинного обучения и сквозных цифровых технологий перед исследователями открываются перспективы использования нейротехнологий и искусственного интеллекта, а

также построения новой методологической базы для оценки химических показателей качества воды и их прогнозирования.

В связи с этим, *целью* настоящей работы является установление принципов построения методики нейросетевого расчета химических показателей качества воды (на примере р. Белая, ЛНР).

Объект исследования — процесс изменения концентрации химических компонентов в речной воде.

Предмет исследования — оценка химических показателей качества воды (на примере сухого остатка).

Задачи исследования:

- анализ существующих подходов к описанию методик для расчета и оценки химических показателей качества воды;
- установление принципов построения методики нейросетевого расчета химических показателей качества воды, включающей описание каждого этапа;
- проверка адекватности методики на основании экспериментальных данных.

Методы исследования. Предлагается общая методика нейросетевого расчета химических показателей качества воды, которая выполнена с использованием алгоритма градиентного бустинга (CatBoost) на данных с водосборной территории р. Белая. Причиной выбора данного алгоритма стала возможность применения ансамблевого обучения, которое в отличие от классических регрессионных моделей типа ARIMA или SARIMAX объединяет ряд слабых моделей прогнозирования. Алгоритм настроен таким образом, чтобы добавлять новые модели в ансамбль. При этом каждая из новых обучается исправлять ошибки предыдущих моделей, что в итоге позволяет достичь более точных прогнозов [10]. Нейросети обладают двойной функциональностью: они могут служить как самостоятельной основой для расчетов, так и дополнять существующие модели. Их основная роль заключается в адаптации уже применяемых методик к фактическим экспериментальным данным. При этом при проектировании каждой

нейросетевой модели ее топология, архитектура и метод обучения выбираются с учетом специфики исследуемого компонента и имеющихся априорных сведений о закономерностях его распространения.

Применение нейросетей для расчета концентраций химических показателей качества воды с возможностью адаптации к водосборной территории реки Белая схематически изображено на рисунке 1.



Рисунок 1 — Методика применения нейросетей для расчета концентраций компоненты химического состава в воде

Методика предусматривает выполнение следующих основных этапов:

1. На **первом этапе** происходит сбор мониторинговой информации о рассчитываемом компоненте. Сбор данных для мониторинга должен быть как можно более исчерпывающим, включая наибольшее количество точек наблюдения и диапазонов возможных концентраций. Информация о замеренных концентрациях применяется как для обучения моделей нейронных сетей, так и для проверки их адекватности.

2. На **втором этапе** определяются наличие регламентированных расчетных методик. Валидация применимых методик осуществляется путем сравнения их расчетных результатов с экспериментальными данными, полученными на начальном этапе. Способ валидации определяется спецификой методики и анализируемого компонента в водной пробе. На основании результатов валидации принимается решение о целесообразности доработки регламентированной методики с использованием дополнительной нейросетевой модели. В случае положительного решения формируется гибридная модель, интегрирующая регламентированную и нейросетевую методики. При этом результаты расчетов по утвержденным методикам выступают в качестве одного из входных параметров для дополнительного нейросетевого блока гибридной модели.

3. На **третьем этапе** выполняется оценка внешних факторов, которые воздействуют на уровень загрязнения водосборной территории. Если на предыдущем этапе было установлено наличие стандартизированной методики для проведения расчетов, то входные параметры этой методики требуют особого внимания.

В зависимости от характера процессов, протекающих на территории и определяющих степень загрязнения, эти внешние факторы могут быть учтены в гибридной модели двумя путями: либо как часть расчетов, выполняемых по регламентированной методике, либо как отдельные, независимые входные данные для нейросетевого

блока модели. В случае отсутствия регламентированной методики, необходимо провести самостоятельный анализ существующих факторов. Все идентифицированные факторы должны быть представлены в числовой, измеримой форме и будут использоваться в качестве входных данных для нейросетевой модели, предназначенной для расчета концентраций определенного компонента.

4. На **четвертом этапе** на основании проведенного анализа второго и третьего этапов разрабатывается и обучается нейросетевой алгоритм для вычисления концентрации химического компонента в водной среде. Экспериментальные данные и результаты расчета нормативов концентрации указанного вещества формируют базу для модели.

5. На **пятом этапе** проводится оценка качества построенной модели на независимых данных, не участвовавших в процессе обучения. Если результаты проверки показывают достаточную точность, модель признается пригодной для расчетов концентрации компонентов. В противном случае необходимо вернуться к этапу анализа влияющих факторов и перестроить модель, учитывая новые данные.

Практическая реализация данной методики требует адаптации алгоритмов построения и обучения нейронных сетей под специфику каждой отдельной задачи.

Для улучшения сходимости разработанные нейронные сети были обучены на данных химического анализа проб воды на реке Белая, полученных сотрудниками Молодежной научно-исследовательской лаборатории геоэкологии и прикладной химии в 2024 году (с июня по ноябрь 2024 года). В качестве основного компонента выбран сухой остаток, данные по которому взяты вблизи поссовета пгт. Бугаевка. Место отбора проб выбрано не случайно, потому что данный створ является равномерным и замыкающим перед Исаковским водохранилищем, т. е. аккумулирующим изменение всего ряда химических показателей.

С помощью действующей нормативной методики (для сухого остатка — РД 52.24.468-2019) были установлены соответствия между значениями концентраций компонента и наборами ключевых метеорологических показателей, зафиксированных специальными службами (гидромет, погода.ру и т. д.). В качестве задаваемого массива были выбраны параметры, фрагмент которых приведен в таблице 1.

Фактические замеры проводились не реже одного раза в неделю, но для более полной картины изменения сухого остатка использовался метод линейной интерполяции.

В качестве примера структуры нейросети, используемой для определения концентраций компонентов, были выбраны следующие характеристики (рис. 2):

1. Алгоритм расчета — градиентный бустинг (CatBoost).

2. Количество скрытых слоев — 1.

3. Количество входных нейронов — 8. Данный показатель характеризует наличие внешних факторов и внутренних взаимосвязей.

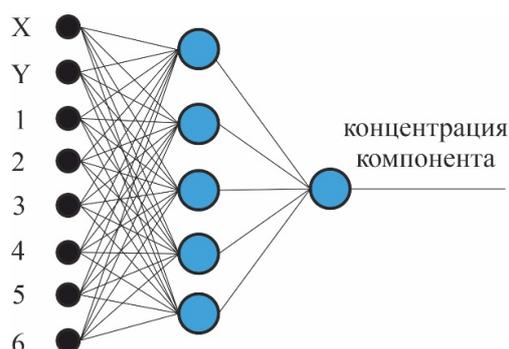
На начальном этапе разработки нейронной сети ключевой задачей был отбор наиболее важных входных данных. Для улучшения точности прогнозирования концентрации компонента, по сравнению с существующими методами, в качестве предикторов были выбраны расчетные значения концентраций примеси и основные метеорологические параметры. Изучение взаимосвязей между концентрациями примеси, зафиксированными в точке отбора проб на реке Белая, и метеорологическими условиями на ее водосборной территории выявило значительное влияние таких факторов, как места пересечения водоносных горизонтов, ранее существовавшие и частично или полностью ликвидированные шахтные выработки, а также температура и атмосферное давление. Эти метеорологические параметры были выбраны в качестве входных данных для нейронной сети.

Точность нейросетевой модели расчета концентраций сухого остатка наглядно продемонстрирована на рисунке 3. Средняя относительная погрешность, полученная с применением данной методики и нормативных коэффициентов, составила 4,83 %.

Таблица 1

Фрагмент массива исходных данных

№ п/п (i)	Дата замера	Местоположение	Метеоданные					Химический анализ
			Средняя температура воздуха [https://global-weather.ru], °C	Среднее атмосферное давление, [https://www.meteoblue.com], (мм рт. ст.)	Средняя скорость ветра [https://www.meteoblue.com], м/с	Направление ветра, [https://www.meteoblue.com], (°)	Среднее количество осадков [https://www.meteoblue.com], мм	
		Координаты, °						Сухой остаток, мг/дм ³
1	01.06.24	48,420454 38,859154	22,5	742	3,1	СЗ	0	1253
2	02.06.24		24	743	3,6	З	0	1261
3	03.06.24		24	742	5,0	ЮЗ	0,1	1268
...
181	28.11.24		1	747	2,8	ССЗ	0	2470
182	29.11.24		1	751	1,7	В	0	2473
183	30.11.24		0,5	755	2,2	ССВ	0	2475



X — координаты точки расчета (широта); Y — координаты точки расчета (долгота);
 1 — средняя температура воздуха (°C); 2 — среднее атмосферное давление (мм рт. ст.);
 3 — средняя скорость ветра (м/с); 4 — направление ветра (°);
 5 — среднее количество осадков (мм); 6 — расчетное значение концентрации примеси, полученное экспериментальным путем на основании методик (мг/дм³)

Рисунок 2 — Структура нейросети для расчета концентрации компонентов химического состава воды

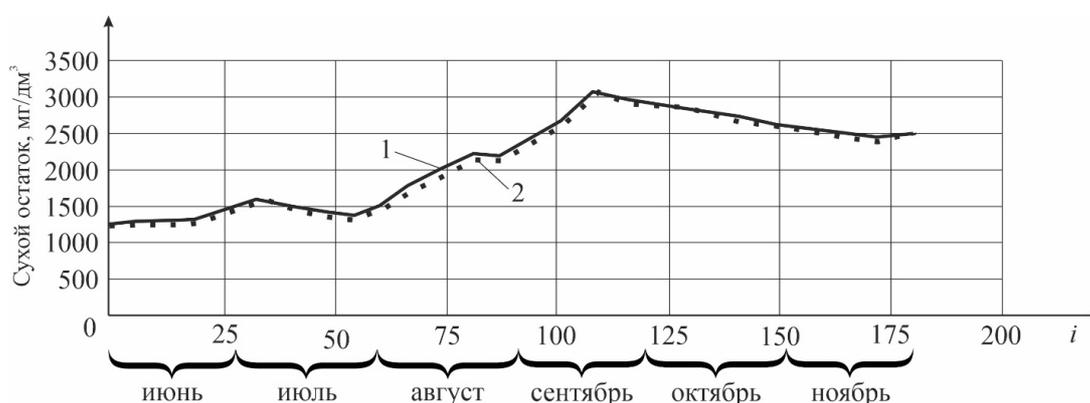


Рисунок 3 — Сравнение экспериментально измеренных (1) и скорректированных нейросетью (2) концентраций сухого остатка

Выводы и направление дальнейших исследований. Несмотря на ограниченность исследуемого ряда наблюдений вследствие слабой экспериментальной сети экологического мониторинга, на примере сухого остатка были получены значения концентрации химического показателя качества воды. Погрешность измерений, выполненных с помощью нейросетевой методики, в основе которой применен алгоритм градиентного бустинга (CatBoost), составила 4,83 %. Разработанная методика, использующая нейросетевой подход, обеспе-

чивает повышение адекватности результатов за счет обучения модели на массиве экспериментальных наблюдений. Это позволяет эффективно компенсировать систематические и случайные ошибки, присутствующие в данных.

Рассмотренная методика будет включена в программно-технический комплекс, одной из задач которого будет выявление закономерностей многофакторного процесса формирования стока поверхностных водных объектов с учетом влияния шахтных вод.

Список источников

1. Стратегическое направление в области цифровой трансформации отрасли экологии и природопользования, относящейся к сфере деятельности Министерства природных ресурсов и экологии Российской Федерации : распоряжение Правительства Российской Федерации от 15 декабря 2023 г. № 3664-р. URL: <http://static.government.ru/media/files/Mqq6qYnбисSm7CbAqQXy0GHESYTQ9K4.pdf>.
2. Красногорская Н. Н., Нафикова Э. В. Геоэкологическая оценка и прогнозирование опасных природно-техногенных процессов на водосборе реки. М. : Инновационное машиностроение, 2015. 242 с.
3. Крамаренко А. А., Коптева А. К., Лысенко И. Л. О наличии гидрогеологической связи между горными выработками закрытых шахт в горнопромышленных районах на примере участка зоны водосбора р. Белая Перевальского района Луганской Народной Республики // Экологический вестник Донбасса. 2023. № 9. С. 27–38. EDN: QVWFSJ
4. Мохов А. В. Трансформация гидродинамических характеристик горного массива на участках освоения каменноугольных залежей подземным способом : автореф. дис. ... д-ра техн. наук. Ростов н/Д, 2015. 22 с.
5. Слейман А., Козлов Д. В. Моделирование речного стока в условиях недостаточного информационного обеспечения // Гидротехническое строительство, 2024. № 7. С. 7–16. DOI: 10.34831/EP.2024.97.56.002. EDN: XWDHTT
6. Белозерова Е. А., Красногорская Н. Н. Разработка геоинформационной системы поддержки принятия решений при управлении водными ресурсами // Гидрометеорология и экология. 2021. № 65. С. 702–725. DOI: 10.33933/2713-3001-2021-65-702-725. EDN: RZGGLF

7. Гулько С. Е. *Научные основы экологически безопасных технологий при использовании шахтных вод* : дис. ... д-ра тех. наук. Донецк, 2019. 328 с.

8. Рыбникова Л. С. *Процессы формирования подземных вод в горнодобывающих районах Среднего Урала на постэксплуатационном этапе* : автореф. дис. ... д-ра геол.-минерал. наук. Москва, 2019. 45 с.

9. Смирнова И. В., Вознюк Ю. С. *Расчёт подземного стока в реку Белая гидрохимическими методами* // Экологический вестник Донбасса. 2023. № 9. С. 41–46. EDN: MTJOQA

10. Hancock J., Khoshgoftaar T. *CatBoost for Big Data: an Interdisciplinary Review* // *Journal of Big Data*. 2020. Vol. 7 (1). URL: https://www.researchgate.net/publication/346496325_CatBoost_for_big_data_an_interdisciplinary_review (date of treatment: 24.06.2025). DOI: 10.1186/s40537-020-00369-8

© Долгих В. П., Кусайко Н. П., 2025

Рекомендована к печати к.т.н., доц. каф. ЭБЖД ДонГТУ Павловым В. И., заместителем начальника отдела по водо- и кислородоснабжению отдела главного энергетика ООО «ЮГМК» Стрельченко Н. В.

Статья поступила в редакцию 06.08.2025.

СВЕДЕНИЯ ОБ АВТОРАХ

Долгих Виталий Павлович, канд. техн. наук, научный руководитель Молодежной научно-исследовательской лаборатории геоэкологии и прикладной химии
Донбасский государственный технический университет,
г. Алчевск, Россия, e-mail: vidoscience@mail.ru

Кусайко Наталья Петровна, директор Научного центра мониторинга окружающей среды, старший научный сотрудник Молодежной научно-исследовательской лаборатории геоэкологии и прикладной химии
Донбасский государственный технический университет
г. Алчевск, Россия

***Dolgikh V. P., Kusayko N. P.** (Donbass State Technical University, Alchevsk, Russia, *e-mail: vidoscience@mail.ru)

PRINCIPLES OF DEVELOPING A NEURAL NETWORK METHOD FOR CALCULATING CHEMICAL WATER QUALITY INDICATORS (THROUGH THE EXAMPLE OF THE BELAYA RIVER, LPR)

A method for neural network calculation of chemical water quality indicators has been proposed. The outlined stages of calculation consider monitoring information about the calculated component, the availability of regulated calculation methods, external factors that influence pollution of the water-collecting area, self-learning and validation of the constructed model for adequacy by performing a series of simulations on data not included in the neural network's training set. Using the example of data on dry residue, a gradient boosting algorithm (CatBoost) has been battle-tested, whose error in comparison with experimental data was 4,83 %. The calculation algorithm takes into account the influence of external factors (meteorological data) and the internal connections between other components.

Key words: river run-off, gradient boosting algorithm (CatBoost), method for calculating chemical water quality indicators, mathematic simulation, dry residue.

Funding: the studies were carried out with funding from the federal budget (theme code: FRRU-2024-0004 in the Unified state information system for accounting the research, experimental and technical works).

INFORMATION ABOUT THE AUTHORS

Dolgikh Vitaliy Pavlovich, PhD in Engineering, Head of Youth Research Laboratory of Geoecology and Applied Chemistry

*Donbass State Technical University,
Alchevsk, Russia, e-mail: vidoscience@mail.ru*

Kusayko Natalia Petrovna, Director of the Scientific Center for Environmental Monitoring, Senior Researcher of the Youth Research Laboratory of Geoecology and Applied Chemistry

*Donbass State Technical University,
Alchevsk, Russia*