УДК 628.31

ПРОГНОЗИРОВАНИЕ КАЧЕСТВА СТОЧНЫХ ВОД И ОЦЕНКА ЭКОЛОГИЧЕСКИХ РИСКОВ НА СТАНЦИИ АЭРАЦИИ ВОДОКАНАЛА

Розенталь[1] О. М., Самбурский[2] Г. А.

[1]Институт водных проблем РАН, 119333, г. Москва, ул. Губкина, д. 3, omro3@yandex.ru
[2]ФГБОУ ВО Российский Химико-технологический университет им. Д.И. Менделеева, 125047, г. Москва, Миусская пл., 9

Аннотация. В работе представлены результаты многолетнего (2019–2023 гг.) анализа качества сточных вод, поступающих на станцию аэрации водоканала, с применением гибридного подхода, объединяющего статистические методы, генеративные нейросети, рекуррентные архитектуры (LSTM, Transformer) и методы объяснимого искусственного интеллекта (SHAP). Выявлены устойчивые тренды снижения концентраций взвешенных веществ (в 4,5 раза), аммонийного азота (в 4,5 раза) и БПК (в 2,6 раза) при сохранении высоких уровней нитратов (≈90 мг/л) и фосфатов. Разработана прогнозная модель с MAPE 11,0% и R^2 0,91, позволяющая осуществлять раннее предупреждение о превышениях нормативов. С помощью SHAP-анализа идентифицированы ключевые факторы, влияющие на качество очистки: расход сточных вод (вклад 25%), температура (18%) и предыдущие значения нитратов (32%). Предложены технологические и организационные меры по снижению экологических рисков. Результаты демонстрируют потенциал ИИ для перехода от реактивного к предиктивному управлению очистными сооружениями.
Предмет исследования: возможность оптимизации работы канализационных очистных сооружений
Материалы и методы: методы математической статистики и моделирования, прогностическая модель работы очистных сооружений
Результаты: получены результаты математического моделирования работы очистных сооружений, согласующиеся с фактическими данными за период, сделаны рекомендации по модернизации действующих сооружений
Выводы: предложен комплекс технологических и организационных мер, направленных на снижение экологических рисков и переход к предиктивному управлению очистными сооружениями. Сделанные предложения целесообразно учесть при формировании программы модернизации очистных сооружений. Полученные результаты демонстрируют высокий потенциал гибридных ИИ-решений для повышения эффективности и экологической безопасности станций аэрации. Дальнейшие исследования целесообразно направить на создание цифрового двойника очистных сооружений с интеграцией прогнозных моделей в реальном времени

Ключевые слова: очистка сточных вод, искусственный интеллект, LSTM, Transformer, SHAP, прогнозирование качества воды, оценка рисков

ВВЕДЕНИЕ

Станции аэрации являются ключевыми элементами городской инфраструктуры, обеспечивающими биологическую очистку сточных вод перед сбросом в водные объекты. Современные требования к качеству очистки основываются на позициях технологического нормирования (НДТ) и в целом системно ужесточаются. При этом для веществ, не относящихся к технологически нормируемым, требования практически всегда основываются, исходя из ПДК водоемов рыбохозяйственного назначения. Традиционные подходы к контролю качества основаны на периодическом отборе проб и лабораторном анализе, что не позволяет оперативно реагировать на изменения состава стоков и предотвращать аварийные сбросы.

В последние годы искусственный интеллект активно внедряется в практику управления очистными сооружениями. Обзорные работы показывают, что ИИ прошёл путь от базового прогнозирования качества до интегрированной оптимизации процессов и стратегий сокращения выбросов. Особый интерес представляют рекуррентные нейросетевые архитектуры (LSTM, GRU, Transformer), способные эффективно моделировать нелинейные и инерционные процессы, характерные для биологической очистки. Наряду с прогнозированием всё большее внимание уделяется интерпретируемости моделей: методы объяснимого ИИ (XAI), такие как SHAP, позволяют выявлять ключевые факторы, влияющие на качество очистки, и принимать обоснованные управленческие решения-.

Настоящая работа посвящена комплексному анализу многолетних данных о качестве сточных вод, сбрасываемых через выпуск №2 станции аэрации водоканала, с применением гибридного подхода на основе ИИ. Цель исследования - выявление трендов, сезонных закономерностей и аномалий, построение прогнозных моделей, оценка рисков для водопользователей и разработка практических рекомендаций по оптимизации режимов очистки.

МАТЕРИАЛЫ И МЕТОДЫ

Характеристика объекта исследования - Южная аэрационная станция (ЮАС) г. Екатеринбург осуществляет приём и очистку сточных вод от потребителей города. ЮАС Проектная мощность ЮАС 550 000 куб м. в сутки (крупнейшее очистное сооружения по постановлению Правительства РФ №1430). В

настоящий момент станция находится в процессе модернизации. Состав очистных сооружений соответствует классической компоновке для станций аэрации. Основные задачи, помимо окисления органической компоненты стоков, связаны с достижением технологических показателей НДТ (ТП НДТ) по азотной группе и фосфору. Так как водоканал Екатеринбурга принимает в свои сети стоки промышленных абонентов, то на ЮАС приходят сточные воды, содержащие металлы, ПАВ, нефтепродукты и ряд других неспецифических для хозяйственно-бытовых сточных вод загрязнителей.

Данные - исходные данные представляют собой результаты анализов проб сточных вод за период 2019–2023 гг. (около 150 записей). Перечень контролируемых показателей включает: взвешенные вещества, аммоний-ион, нитрит-ион, нитрат-ион, фосфаты (по фосфору), сульфаты, нефтепродукты, сухой остаток, хлориды, цинк, фенол, железо общее, марганец, хром (3+ и 6+), никель, алюминий, медь, ХПК, БПК, АПАВ, а также расход сточных вод.

Предобработка данных - Выполнена очистка от явных ошибок. Пропуски (≈12% от общего объёма) восполнены с помощью генеративно-состязательных сетей (Generative Adversarial Networks, GAN). GAN обеспечивает интеллектуальное восполнение пропусков в данных. В классической постановке GAN [1] генератор G и дискриминатор D играют минимаксную игру (1):

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}\left[\log\left(1 - D\big(G(z)\big)\right)\right], \quad (1)$$

Здесь:

- $E_{x\sim p\ data(x)}$ — математическое ожидание по распределению реальных данных.
- $E_{z\sim p\ z(z)}$ — математическое ожидание по распределению входного шума.
- D(x) — вероятность того, что образец x является реальным.
- D(G(z)) — вероятность того, что сгенерированный образец принадлежит к реальным данным.
- G(z) — сгенерированные данные (выход генератора);
- x – реальные данные, z – шумовой вектор.

Функции потерь:

$$L_D = - E_{x\sim p_data}[\log D(x)] - E_{z\sim pz}[\log(1 - D(G(z)))] \quad (2)$$

$$L_G = - E_{z\sim pz}[\log(1 - D(G(z)))] \quad (3)$$

Для каждого пропущенного показателя синтезировано по 10 вариантов, взято медианное значение.

Анализ временных рядов - для выделения трендов использован фильтр Ходрика–Прескотта (HP) [2,3], решающий задачу минимизации:

$$\min_{\{x_t\}}\left[\sum_{t=1}^{T}(y_t - x_t)^2 + \lambda\sum_{t=2}^{T-1}\big((T_{t+1} - T_t) - (T_t - T_{t-1})\big)^2\right] \quad (4)$$

$У_t$ – исходный ряд, T_t – тренд, λ = 14400 для месячных данных

STL декомпозиция (Seasonal and Trend decomposition using LOESS) раскладывает ряд на основные компоненты:

$$У_t = T_t + S_t + R_t \quad (5)$$

T_t – тренд, S_t – сезонная компонента (период 12 месяцев), R_t - остаток

Детектирование аномалий - Использованы два метода: *Isolation Forest и автоэнкодер* для обнаружения аномальных сбросов (залпов)

Изолирующий лес (Isolation Forest) [4,5]. Аномальный балл:

$$s(x,n) = 2^{-\frac{E[h(x)]}{c(n)}}, \quad (6)$$

где:

- $E[h(x)]$ — средняя длина пути (число разбиений) от корня до листа, в котором оказывается точка x, усреднённая по всем деревьям леса;
- $c(n)$ — средняя длина пути в случайном бинарном дереве поиска для nнаблюдений:

$$c(n) = \begin{cases} 0, & n \le 1, \\ 2H_{n-1} - \frac{2(n-1)}{n}, & n > 1, \end{cases} \quad (7)$$

где $H_{n-1} = \ln(n-1) + \gamma$

$\gamma \simeq 0.5772$ – (постоянная Эйлера-Маскерони).

Значение s>0.6 считается аномальным

Автоэнкодер [5]. Ошибка реконструкции: использовалась среднеквадратичная ошибка (MSE), как стандарт для непрерывных данных, где d — размерность вектора данных.

$$L_{\text{MSE}}(x,\hat{x}) = \frac{1}{d}\sum_{i=1}^{d}(x_i - \hat{x}_i)^2. \quad (8)$$

Значения s>0.6s>0.6 считаются аномальными.

Порог аномалии - 95-й процентиль распределения ошибок.

Прогнозирование - Построены модели LSTM и Transformer. *LSTM (Long Short-Term Memory)* – это нейросеть для работы с

долгосрочными зависимостями (табл. 1) во временных рядах с тремя управляющими "вентилями" (забывающий, входной, выходной), которые регулируют поток информации [6].

Таблица 1. Уравнения LSTM-ячейки
Table 1. Equations of an LSTM cell

Парамерт	Уравнение	Комментарий
Забывающий вентиль (forget gate)	$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right)$	решает, какую часть старого состояния ячейки c_{t-1}сохранить:
Входной вентиль (input gate)	$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$	определяет, сколько новой информации добавлять
Кандидат на состояние ячейки (candidate cell state)	$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$	предлагает новые значения для добавления
Обновление состояния ячейки (cell state)	$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$	Здесь $f_t \odot c_{t-1}$«забывает» часть старой памяти, а $i_t \odot \tilde{c}_t$добавляет новую.
Выходной вентиль (output gate)	$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$	решает, какая часть состояния ячейки пойдёт на выход:
Скрытое состояние (hidden state)	$h_t = o_t \odot \tanh(c_t)$	фактический выход ячейки

Используется: f_t, i_t, o_t - входной, забывающий и выходной вентили; σ- сигмоида, ⊙-поэлементное произведение. Архитектура 2 слоя по 64 нейрона, Dropout 0.2

Transformer [7-10]. Attention Mechanism:

Для набора запросов Q, ключей Kи значений Vвыход механизма внимания:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V \quad (9)$$

Многоголовое внимание:

$$\text{MultiHead}(X) = \text{concat} \cdot W_O. \quad (10)$$

$$\text{concat} = [\text{head}_1; \text{head}_2; \ldots; \text{head}_H] \quad (11)$$

$$\text{head}_i = \text{Attention}\left(QW_Q^i, KW_K^i, VW_V^i\right) \quad (12)$$

Принято: архитектура: 4 головы внимания, 2 слоя, размерность 128

Метрики оценки: *регрессия*

- *MSE (Mean Squared Error)*: $\frac{1}{N}\sum(y_i - \hat{y}_i)^2$. Чувствителен к выбросам.
- *RMSE*: $\sqrt{\text{MSE}}$ — в тех же единицах, что и целевая переменная.
- *MAE (Mean Absolute Error)*: $\frac{1}{N}\sum|y_i - \hat{y}_i|$. Более устойчив к выбросам, легче интерпретировать.
- *MAPE (Mean Absolute Percentage Error)*: средняя ошибка в процентах. Удобна для сравнения моделей на разных масштабах, но ломается, если $y_i \approx 0$.
- R^2: доля объяснённой дисперсии. Показывает, насколько модель лучше простого среднего.

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \quad (13)$$

где:

- $SS_{\text{res}} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ — остаточная сумма квадратов (необъяснённая дисперсия),
- $SS_{\text{tot}} = \sum_{i=1}^{n}(y_i - \bar{y})^2$ — общая сумма квадратов (полная дисперсия),
- y_i — фактическое значение,
- $\hat{y}_i$ — предсказанное моделью,
- $\bar{y}$ — среднее по y.

Объяснимый ИИ ((SHapley Additive exPlanations, или аддитивные объяснения Шепли, SHAP)

Формально вклад признака $\boldsymbol{j}$для объекта $\boldsymbol{x}$ — это значение Шепли:

$$\phi_j = \sum_{S \subseteq \{1,\ldots,d\}\setminus\{j\}} \frac{|S|!(d-|S|-1)!}{d!}[f(S \cup \{j\}) - f(S)], \quad (14)$$

Аддитивность. Для каждого объекта выполняется:

$$f(x) = E[f(X)] + \sum_{j=1}^{d} \phi_j(x), \quad (15)$$

То есть прогноз равен среднему прогнозу по выборке плюс сумма вкладов всех признаков.

Оценка рисков - Вероятность превышения ПДК:

$$P(\text{превышение}) = \frac{1}{1+e^{(\beta(0)+\Sigma\beta(i)x(i)}} \quad (16)$$

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Таблица 2. Среднегодовые концентрации ключевых ингредиентов (мг/дм³)
Table 2. Average annual concentrations of key ingredients (mg/dm³)

Показатель	**2019**	**2020**	**2021**	**2022**	**2023**	**Изменение 2023/2019, раз**
Взвешенные вещества	57,7	34,7	20,7	11,6	13,1	↓ **в 4,4**
Аммоний-ион	3,6	3,5	1,1	0,73	0,79	↓ **в 4,6**
Нитрит-ион	0,12	0,15	0,26	0,18	0,42	↑ в 3,5
Нитрат-ион	7,2	8,8	62,9*	99,5	86,3	↑ в 12,0
Фосфаты (по Р)	3,0	2,1	2,6	1,25	1,26	↓ **в 2,4**
БПК	31,5	~30	~25	11,2	12,1	↓ **в 2,6**
Нефтепродукты	0,08	~0,15	~0,20	0,18	0,16	↑ в 2,0
Железо общее	0,27	0,28	0,07	0,09	0,10	↓ в 2,7
Цинк	0,089	0,096	0,055	0,045	0,052	↓ в 1,7
Марганец	0,129	0,125	0,076	0,077	0,092	↓ в 1,4
Хлориды	~68	~72	~72	78	82	↑ в 1,2
ХПК	-	-	~80	67	109	↑ в 1,4

* аномально высокие нитраты в 2021 г., вероятно, связаны с изменением технологического режима или ошибкой измерений.

Тренды основных показателей

Наблюдается устойчивое снижение взвешенных веществ, аммония, фосфатов и БПК, что свидетельствует о повышении эффективности очистных сооружений и реализации предприятием плана снижения сбросов. В то же время концентрации нитратов остаются на уровне 86–100 мг/л, что значительно превышает нормативы для водоёмов рыбохозяйственного назначения (40 мг/л) и указывает на недостаточную денитрификацию. Рост нитритов в 2023 г. (0,42 мг/л против 0,18 в 2022) может сигнализировать о нестабильности процессов нитрификации.

Сезонные колебания - STL-декомпозиция выявила следующие закономерности:

- Весна (март–май): взвешенные вещества на 25–30% выше среднегодовых (паводковый смыв).
- Лето (июнь–август): аммоний и фосфаты на 15–20% выше (активизация биологических процессов, снижение растворённого кислорода в приёмнике).
- Осень (сентябрь–ноябрь): рост нитратов (увеличение нагрузки на аэротенки).
- Зима (декабрь–февраль): минимальные значения большинства показателей.

Аномалии

Автоэнкодер (порог 95-й процентиль) идентифицировал следующие значимые события:

1. 10.01.2019 – множественное превышение по взвешенным, БПК, нефтепродуктам, металлам (залповый сброс).
2. 18.02.2019 – взвешенные 581,5 мг/л (сбой в работе отстойников).
3. Апрель–май 2021 – нитраты> 60 мг/л.
4. 07.02.2023 – цинк 0,156 мг/л, железо 0,219 мг/л (локальный промышленный выброс).
5. 04.10.2022 – железо 0,313 мг/л (вероятно, коррозионный выброс).

Прогнозирование

Transformer показал наилучшие результаты, что согласуется с данными других исследований, подтверждающих эффективность архитектур на основе внимания для прогнозирования качества воды-.

Прогноз на январь–февраль 2024 г. (средние значения): взвешенные - 11–13 мг/л; аммоний - 0,70–0,75 мг/л; нитраты - 85–95 мг/л; фосфаты - 1,20–1,30 мг/л; нефтепродукты - 0,14–0,18 мг/л.

Таблица 3. Сравнение точности моделей (усреднено по всем показателям)
Table 3. Comparison of model accuracy (averaged across all indicators)

Модель	**MAE, мг/л**	**RMSE, мг/л**	**MAPE, %**	**R^2**
LSTM	4,2	7,8	12,3	0,89
Transformer	3,9	7,1	11,0	0,91

Таблица 4. Вероятность превышения нормативов в ближайшие 30 суток
Table 4. Probability of exceeding the standards within the next 30 days

Показатель	ПДК (рыбхоз) / ТП НДТ, мг/л	Фактический диапазон	Вероятность превышения, %	Уровень риска
Взвешенные	+0,25 к фону /10	2–18	22	Средний
$БПК_5$	2,1/8	6–14	98	Критический
Аммоний-ион	0,4/1,0	0,25–1,61	68	Высокий
Нитрат-ион	40/40	58–155	85	Критический
Фосфаты	0,2/0,5	0,53–1,39	100	Критический
Нефтепродукты*	0,05–0,1	0,02–0,44	55	Средний
Цинк*	0,01–0,03	0,019–0,156	70	Высокий
ХПК	15/40	40–215	100	Критический

*не технологические показатели для ЮАС

Оценка рисков

БПК, ХПК, нитраты и фосфаты систематически превышают нормативы, создавая угрозу эвтрофикации водного объекта. Экологические риски связаны с потенциальной эвтрофикацией водоисточника за счет сброса биогенных компонентов (нитраты и фосфаты). Риск для питьевого водоснабжения связан с нитратами (выше 40 мг/л).

Интерпретация с помощью SHAP - SHAP-анализ модели прогноза нитратов выявил следующие вклады факторов:

1. Предыдущее значение нитратов (лаг 1) - 32% (инерционность системы).
2. Расход сточных вод - 25% (отрицательная корреляция: увеличение расхода снижает концентрацию).
3. Температура воды (сезон) - 18% (летом нитрификация активнее, денитрификация ограничена).
4. Концентрация аммония - 12% (прямая связь: больше аммония → больше нитратов).
5. Фосфаты - 8% (косвенное влияние на активность микроорганизмов).
6. Металлы (цинк, медь) - 5% (токсическое ингибирование).

Эти результаты согласуются с литературными данными, где аэрация, расход и эффективность удаления взвешенных веществ называются ключевыми предикторами.

РЕКОМЕНДАЦИИ ПО УПРАВЛЕНИЮ КАЧЕСТВОМ СТОЧНЫХ ВОД

На основе проведённого анализа и SHAP-интерпретации предлагаются следующие мероприятия.

Технологические решения

1. Внедрение стадии денитрификации. Для снижения нитратов с 90 до < 40 мг/л необходимо создать аноксидную зону в аэротенках или выделить отдельный денитрификатор с подачей органического субстрата. Денитрификация позволяет восстанавливать нитраты до газообразного азота.
2. Оптимизация аэрации. Установка частотных регуляторов и датчиков растворённого кислорода для поддержания оптимального режима (2–3 мг O_2/л в аэробной зоне).
3. Увеличение времени пребывания в первичных отстойниках для снижения взвешенных веществ.
4. Химическое осаждение фосфора. В периоды повышенных концентраций фосфатов (летние месяцы) рекомендуется реагентная обработка (соли алюминия или железа).
5. Дополнительная очистка от нефтепродуктов**.** Флотационная установка или фильтры-адсорберы.

Организационные меры

1. Ужесточение мониторинга промышленных предприятий-абонентов, особенно в отношении металлов и нефтепродуктов.
2. Разработка регламента действий при аномалиях - переключение потоков в резервные ёмкости при обнаружении залповых сбросов.
3. Внедрение автоматической системы раннего предупреждения на основе прогнозной модели ИИ (интеграция в SCADA) с оповещением операторов при вероятности превышения> 80%. Современные исследования показывают, что такие системы позволяют перейти от реактивного управления к предиктивному-.

Долгосрочное планирование

1. Провести ревизию трубопроводов и оборудования для выявления источников повышенного содержания металлов.
2. Рассмотреть строительство дополнительных полей фильтрации или биоплато для доочистки перед сбросом.
3. Разработать программу снижения удельных сбросов на 20% к 2026 г. с промежуточным контролем.

ЗАКЛЮЧЕНИЕ

В работе выполнен комплексный анализ многолетних данных о качестве сточных вод станции аэрации с применением современного арсенала методов ИИ. Основные выводы:

1. За период 2019–2023 гг. достигнуто значительное снижение концентраций взвешенных веществ (в 4,4 раза), аммонийного азота (в 4,6 раза), фосфатов (в 2,4 раза) и БПК (в 2,6 раза), что

свидетельствует об успешной модернизации очистных сооружений.

2. Сохраняются критические риски по нитратам (превышение ПДК в 85% случаев), фосфатам (100%), БПК (98%) и ХПК (100%), что создаёт угрозу эвтрофикации водного объекта-приёмника.

3. Разработанная прогнозная модель на основе Transformer обеспечивает MAPE 11,0% и R^2 0,91, что позволяет осуществлять раннее предупреждение о превышениях нормативов.

4. С помощью SHAP-анализа идентифицированы ключевые управляемые факторы: расход сточных вод (вклад 25%) и интенсивность аэрации (косвенно через температуру и аммоний), воздействуя на которые можно целенаправленно снижать концентрации нитратов.

5. Предложен комплекс технологических и организационных мер, направленных на снижение экологических рисков и переход к предиктивному управлению очистными сооружениями. Сделанные предложения целесообразно учесть при формировании программы модернизации ЮАС (третий этап). Полученные результаты демонстрируют высокий потенциал гибридных ИИ-решений для повышения эффективности и экологической безопасности станций аэрации. Дальнейшие исследования целесообразно направить на создание цифрового двойника очистных сооружений с интеграцией прогнозных моделей в реальном времени.

СПИСОК ЛИТЕРАТУРЫ

1. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, 27.
2. Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29(1), 1–16.
3. Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, 6(1), 3–73.
4. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422.
5. Sakurada, M., & Yairi, T. (2014). Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction. *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, 4–11.
6. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.
7. 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.
8. Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30.
9. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts.
10. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). John Wiley & Sons.

FORECASTING WASTEWATER QUALITY AND ASSESSING ENVIRONMENTAL RISKS AT THE WATER UTILITY AERATION STATION

Rosental O.M.[1], Samburskii G.A.[2]

[1]Water problem institute Russian Academy of Science, 119330, Gubkina str., 3
[2]D.I. Mendeleev Russian University of Chemistry and Technology, 125047, Russia, Moscow, 9 Miusskaia Sq.

Abstract. The paper presents the results of a multi-year (2019-2023) analysis of the quality of wastewater entering the aeration plant using a hybrid approach combining statistical methods, generative neural networks, recurrent architectures (LSTM, Transformer) and methods of explicable artificial intelligence (SHAP). stable trends of decreasing concentrations of suspended solids (4.5 times), Ammonium nitrogen (4.5 times) and BOD (2.6 times) were revealed while maintaining high levels of nitrates (≈90 mg/l) and phosphates. A predictive model with a mape of 11.0% and R2 of 0.91 has been developed, which allows for early warning of regulatory exceedances. using SHAP analysis, the key factors influencing the quality of treatment were identified: wastewater consumption (contribution of 25%), temperature (18%) and previous nitrate values (32%). Technological and organizational measures to reduce environmental risks are proposed. The results demonstrate the potential of ai for the transition from reactive to predictive management of wastewater treatment plants.
Subject of the study: The possibility of optimizing the operation of wastewater treatment facilities.
Materials and methods: methods of mathematical statistics and modeling, a predictive model of wastewater treatment facility operation.
Results: The results of mathematical modeling of wastewater treatment facility operation were obtained; they are consistent with the actual data for the period; recommendations for modernizing existing facilities were made.
Conclusions: A set of technological and organizational measures aimed at reducing environmental risks and transitioning to predictive management of wastewater treatment facilities has been proposed. It is advisable to take the proposed measures into account when developing a wastewater treatment facility modernization program. The results obtained demonstrate the high potential of hybrid ai solutions to improve the efficiency and environmental safety of aeration stations. Further research should be focused on creating a digital twin of wastewater treatment facilities with the integration of predictive models in real time.

Key words: wastewater treatment, artificial intelligence, LSTM, Transformer, SHAP, water quality forecasting, risk assessment