

УДК 628.35

**МОДЕЛИРОВАНИЕ ПРОЦЕССА БИОЛОГИЧЕСКОЙ ОЧИСТКИ
СТОЧНЫХ ВОД НЕФТЕХИМИЧЕСКОГО ПРОИЗВОДСТВА**

© 2025 Е.С. Балымова, А.Р. Новикова, Р.К. Закиров, Ф.Ю. Ахмадуллина

Казанский национальный исследовательский технологический университет, г. Казань, Россия

Статья поступила в редакцию 06.02.2025

В статье рассматриваются вопросы цифровой трансформации и внедрения прорывных технологий в сфере водоотведения и биологической очистки сточных вод нефтехимического производства. В условиях нестабильности по составу и расходу сточных вод, высоких концентраций токсичных веществ в стоках производств моделирование процесса биоочистки и реализация алгоритмов для поддержки принятия решений по сбросу сточных вод на биологические очистные сооружения становится ключевым направлением для обеспечения устойчивого развития урбанизированных территорий и защиты окружающей среды. В статье изложены результаты прогнозирования оптимальных параметров биологической очистки сточных вод на действующей биостанции нефтехимического предприятия в Республике Татарстан, который поможет принимать обоснованные решения операторами очистных сооружений относительно возможности приема сточных вод на очистку, условий их обработки и прогнозирования эффективности процессов биологической очистки. Методология исследования включает в себя лабораторное моделирование процессов биологической очистки сточных вод, используя пробы активного ила и сточных вод с действующей биостанции нефтехимического предприятия. На основании экспериментальных данных проведен анализ влияния токсикантов (бензол, толуол, стирол, фенол, этилбензол, формальдегид) на эффективность очистки и физико-химические и биохимические свойства активного ила, а также показатели общей загрязненности и токсичности сточных вод. Результаты исследования позволяют не только повысить точность прогнозирования, но и оптимизировать рабочие процессы на очистных сооружениях, что, несомненно, повысит эффективность очистки высоконцентрированных токсичных сточных вод. Предложенное решение – оптимальная математическая модель «случайный лес», адекватно описывающая процесс биологической очистки сточных вод нефтехимического комплекса, способствует устойчивому развитию и безопасной эксплуатации нефтехимических производств. Проведенные исследования направлены на экологическое оздоровление региона и муниципальных образований, снижение техносферной нагрузки, вызванной негативным воздействием сброса недостаточно очищенных сточных вод в природные водоприемники.

Ключевые слова: активный ил, сточная вода, биологическая очистка, машинное обучение, токсиканты, гидробиологические индексы, математическая модель, случайный лес.

DOI: 10.37313/1990-5378-2025-27-1-166-175

EDN: WBGMYX

ВВЕДЕНИЕ

Развитие и цифровая трансформация различных отраслей экономики приводит к внедрению на предприятиях прорывных технологий, в том числе цифровых. В Российской Федерации принят ряд стратегических документов, направленных на цифровизацию экономики, таких как «Стратегия развития информационного общества в Российской Федерации на 2017-2030 годы» и программа «Цифровая экономика Российской Федерации» [1].

С учетом возрастающего интереса к цифровизации и государственной политике в данной области, обеспечение разработки алгоритмов и программного обеспечения, которые оптимизируют рабочие процессы, особенно связанные с принятием решений, оказывающих влияние устойчивое развитие урбанизированных территорий и состояние окружающей (природной) среды, становится актуальной и важной задачей [2]. К сфере обеспечения безопасности и экологической устойчивости

Балымова Елена Сергеевна, кандидат технических наук, доцент кафедры промышленной биотехнологии.

E-mail: ilc2013@inbox.ru

Новикова Айгуль Рафаэлевна, аспирант кафедры промышленной биотехнологии. E-mail: aignovikova@yandex.ru

Закиров Рустем Каюмович, кандидат технических наук, доцент кафедры промышленной биотехнологии.

E-mail: zakrus@mail.ru

Ахмадуллина Фарида Юнусовна, старший преподаватель кафедры промышленной биотехнологии.

E-mail: akhmadullina1951@mail.ru

производств можно отнести проблемы повышения эффективности очистки на биостанциях химической и нефтехимической отраслей, которые обрабатывают сложные по составу высококонцентрированные токсичные сточные воды [3].

Целью исследования являлась разработка алгоритма принятия решения персоналом очистных сооружений о возможности приема на очистку, условий проведения процесса и прогнозирования его эффективности на биологических очистных сооружениях.

Задачи, решаемые в работе:

1. Оценить влияние приоритетных токсикантов на процесс биологической очистки сточных вод нефтехимического производства;
2. Провести подготовку экспериментальных данных к анализу данных для повышения точности моделей прогноза;
3. Исследовать данные путем построения математических моделей с выявлением оптимальной модели, адекватно описывающей процесс биологической очистки;
4. Оценить возможность применения алгоритмических средств машинного обучения для принятия решений оператором очистных сооружений по очистке и сбросу сточных вод.

МЕТОДИКА ИССЛЕДОВАНИЙ

Биологическая очистка сточных вод – эффективный способ удаления органических веществ микроорганизмами, с помощью которых находящиеся в сточных водах в растворенном или коллоидном состоянии органические вещества разлагаются и удаляются из очищаемой воды. Объектом исследования являлся процесс биологической очистки сточных вод нефтехимического производства, содержащих различные токсиканты.

Для получения экспериментальных данных были отобраны пробы активного ила и сточных вод на действующих биологических очистных сооружениях нефтехимического предприятия Республики Татарстан. Процесс биологической очистки моделировался в лабораторных биореакторах, в которых активный ил и сточные воды смешивались в соотношении 1:2 на основании производственных данных. Доза активного ила варьировалась в диапазоне 2,5-3,5 г/дм³, продолжительность процесса биологической очистки – 16 ч (согласно технологическим расчетам). В процессе биоочистки контролировали концентрацию растворенного кислорода с использованием анализатора растворенного кислорода МАРК-303М (ООО «ВЗОР», Россия), а также температуру и активную реакцию среды (pH) с использованием pH-метра pH-150МИ (ООО «Измерительная техника», Россия). На рис. 1 представлена схема экспериментальной установки, использованной для проведения исследований.

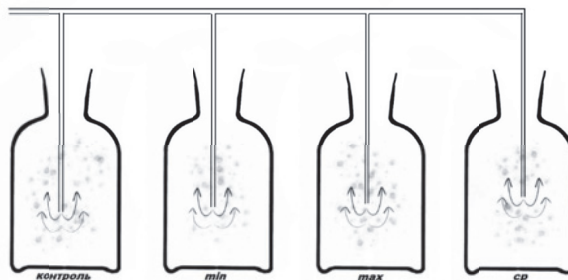


Рис. 1. Экспериментальная установка:

- контроль – эксперимент без добавления токсиканта;
- min – эксперимент с минимальным содержанием токсиканта;
- max – эксперимент с максимальным содержанием токсиканта;
- ср – эксперимент со средним содержанием токсиканта

В сточных водах изучаемого предприятия присутствуют различные экотоксиканты в широком диапазоне концентраций (табл. 1). Для оценки влияния токсикантов на процесс биологической очистки в промаркированные биореакторы добавляли расчетное количество исследуемых токсикантов с обязательным учетом их исходной концентрации в стоках. Изучаемый диапазон концентраций экотоксикантов выбран согласно данным производственного экологического мониторинга поступающих сточных вод действующих очистных сооружений нефтехимического предприятия. Остальные нормируемые показатели сточных вод находились в пределах нормы согласно технологическому регламенту изучаемых очистных сооружений. Экспериментальные исследования предполагали оценку индивидуального влияния экотоксиканта на процесс биологической очистки.

В ходе проведенных исследований осуществляли обязательное тестирование контрольной пробы, представляющей собой смесь активного ила со сточными водами без добавления токсикантов. Пробы иловой смеси для анализа ключевых контролируемых показателей активного ила и сточной воды отбирались:

Таблица 1. Диапазон концентраций изучаемых экотоксикантов

Наименование вещества	Концентрация в сточных водах, мг/дм ³			Норма в химстоке по технологическому регламенту, мг/дм ³
	мин	макс	сред	
формальдегид	0,032±0,002	133,0±6,7	9,3±0,5	15
фенол	0,0050±0,0003	3,3±0,2	0,50±0,03	5
бензол	0,050±0,003	21,5±1,1	1,1±0,1	10
толуол	0,0050±0,0003	22,0±1,1	0,50±0,03	2
этилбензол	0,024±0,001	232,0±11,6	13,6±0,7	2
стирол	0,050±0,003	4,2±0,2	0,40±0,02	2

- для сточных вод: химическое потребление кислорода (ХПК) в соответствии с ПНД Ф 14.1:2:3.100-97, мутность по ПНД Ф 14.1:2:3.4.213-05, токсичность по ПНД Ф Т 14.1:2:3.13-06 и концентрации токсикантов.

- для активного ила: иловый индекс и доза ила согласно ФР 1.31.2008.04398, дегидрогеназная активность (ДАИ) [4], каталазная активность (КА) [5], гидробиологический анализ с расчетом индекса биоразнообразия Шеннона [6];

Для предобработки полученных данных использовались методы корреляционного анализа, позволяющие установить взаимосвязи между контролируемыми параметрами, а также оценивали минимальные, максимальные, средние и медианные значения показателей.

Для моделирования процесса биологической очистки сточных вод параметрами в данной работе выбраны:

X_1 – токсикант (табл.1);

X_2 – исходная концентрация токсиканта в сточной воде, мг/дм³;

X_3 – время аэрации активного ила со сточной водой, ч;

X_4 – гидробиологический индекс Шеннона активного ила до смешения со сточной водой (исходный ил);

X_5 – ДАИ исходного ила по отношению к водопроводной воде ($ДАИ_{ВВисх}$), мг/г;

X_6 – ДАИ исходного ила по отношению к сточной воде ($ДАИ_{СВисх}$), мг/г;

X_7 – ДАИ исходного ила по отношению к надиловой жидкости ($ДАИ_{НЖисх}$), мг/г;

X_8 – КА исходного ила ($КА_{исх}$), мг H_2O_2 /г·мин;

X_9 – ХПК/доза ила, мг/г;

X_{10} – иловый индекс, см³/г;

X_{11} – токсичность сточной воды, %;

Y_1 – гидробиологический индекс Шеннона активного ила в процессе биологической очистки (измерение разнообразия родов, учитывая их однородность);

Y_2 – ДАИ по отношению к надиловой жидкости в процессе биологической очистки ($ДАИ_{НЖ}$), мг/г;

Y_3 – ДАИ по отношению к сточной воде в процессе биологической очистки ($ДАИ_{СВ}$), мг/г;

Y_4 – КА в процессе биологической очистки (активность фермента, катализирующего реакцию разложения перекиси водорода), мг H_2O_2 /г·мин.

Отношение ХПК/доза ила использована в работе для оценки нагрузки на активный ил, выражаемой в количестве поступающих со сточной водой загрязнений по ХПК, приходящееся на единицу массы активного ила. Показатель ХПК использован с учетом меньшей по сравнению с биологическим потреблением кислорода продолжительностью его лабораторного определения для повышения экспрессности контроля и моделирования процесса биологической очистки на нефтехимических предприятиях.

В процессе эксперимента производили отборы проб для измерения контрольных показателей через 15 мин, 1 ч, 2 ч, 3 ч и 16 ч после смешения активного ила со сточными водами. Экспериментальные исследования предполагали полнофакторный эксперимент для всех изучаемых токсикантов с учетом диапазона изменения их концентраций в трех повторностях для исключения случайных погрешностей. Общий объем датасета составил 4560 строк.

ПОДГОТОВКА ДАННЫХ ДЛЯ ИССЛЕДОВАНИЯ

Для достижения поставленных целей в рамках настоящего исследования был использован язык программирования Python в среде Jupyter, который представляет собой высокоуровневый язык программирования общего назначения. Данный язык обеспечивает высокую производительность разработчиков и читаемость кода, а также обеспечивает легкую переносимость созданных про-

граммных приложений [7]. Python также является открытым фреймворком, позволяющим проводить анализ данных, статистическое моделирование и машинное обучение.

Для моделирования процесса биологической очистки сточных вод в Jupyter создан новый лист (Notebook). В первую очередь организовали импорт используемых библиотек для работы с данными, используя функцию `import`.

Для проверки типа данных в столбцах, а также выявления пропусков в ячейках, используя функцию выведения информации по таблице (`info`), была проанализирована информация в таблице исходных данных. По выведенной информации были выявлены пропуски в столбцах X_4 , X_5 , X_{10} , X_{11} , Y_1 , Y_2 , Y_3 , Y_4 . По результатам проверки данных ошибок по типам данных нет. На следующем этапе проверили данные на наличие явных повторов, используя функцию выведения явных повторений (`uplicated`). Повторов в данных не обнаружено, приступили к следующему этапу – построение корреляции между числами (рис. 2), для чего использовали функцию подсчета корреляции (`corr`).

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Y1	Y2	Y3	Y4
1.000000	0.140913	0.115794	-0.003711	-0.269714	-0.159633	0.107971	0.010862	0.029783	0.091429	0.128943	0.033742	-0.139388	-0.335967	0.008045
0.140913	1.000000	0.010008	-0.094307	-0.120932	-0.057522	-0.117648	-0.100811	-0.177839	-0.052302	0.209482	0.088740	0.037001	-0.133860	-0.038260
0.115794	0.010008	1.000000	0.150878	0.094439	0.019345	-0.022605	-0.011917	-0.007350	0.208073	-0.144977	0.140508	-0.073277	0.002077	0.150584
-0.003711	-0.094307	0.150878	1.000000	0.341472	0.015166	-0.438892	0.776683	0.107514	0.077839	0.080312	0.141910	0.192154	0.267040	0.568233
-0.269714	-0.120932	0.094439	0.341472	1.000000	0.036618	0.014663	0.192271	0.019219	-0.050538	-0.007000	0.107333	0.474712	0.640228	0.295296
-0.159633	-0.057522	0.019345	0.015166	0.036618	1.000000	0.115986	-0.064885	-0.320186	-0.037340	0.113556	0.165017	0.031600	0.152579	-0.086734
0.107971	-0.117648	-0.022605	-0.438892	0.014663	0.115986	1.000000	-0.554492	-0.194354	-0.004421	-0.070743	0.242472	0.099432	0.017953	-0.498552
0.010862	-0.100811	-0.011917	0.776683	0.192271	-0.064885	-0.554492	1.000000	0.269958	-0.113864	0.031377	0.057724	0.039629	0.108035	0.400637
0.029783	-0.177839	-0.007350	0.107514	0.019219	-0.320186	-0.194354	0.269958	1.000000	0.063086	-0.277327	-0.446699	-0.003163	0.076643	0.418344
0.091429	-0.052302	0.208073	0.077839	-0.050538	-0.037340	-0.004421	-0.113864	0.063086	1.000000	-0.149910	-0.080068	-0.060388	-0.065696	0.138569
0.128943	0.209482	-0.144977	0.080312	-0.087860	0.113556	-0.070743	0.031377	-0.277327	-0.149910	1.000000	0.276974	0.105405	-0.007489	-0.126149
0.033742	0.088740	0.140508	0.141910	0.187333	0.165017	0.242472	0.057724	-0.446699	-0.080068	0.276974	1.000000	0.147188	0.070671	-0.314071
-0.139388	0.037001	-0.073277	0.192154	0.474712	0.031600	0.099432	0.039629	-0.003163	-0.060388	0.105405	0.147188	1.000000	0.521566	0.208783
-0.335967	-0.133860	0.002077	0.267040	0.640228	0.152579	0.017953	0.108035	0.076643	-0.065696	-0.007489	0.070671	0.521566	1.000000	0.246160
0.008045	-0.038260	0.150584	0.568233	0.295296	-0.086734	-0.498552	0.400637	0.418344	0.138569	-0.126149	-0.314071	0.209783	0.246160	1.000000

Рис. 2. Корреляция между значениями переменных

Наиболее значительная корреляция отмечена между индексом Шеннона (X_4) и каталазной активностью (X_8), а также между исходной ДАИ по отношению к очищенной воде и ДАИ по отношению к сточной воде в процессе очистки (соответственно X_5 и Y_3). В первом случае это связано с откликом биомассы на негативное воздействие, то есть снижение биоразнообразия, оцениваемое по индексу Шеннона, коррелирует с ростом каталазной активности, как показателя стресса биомассы.

Корреляция между ДАИ_{ВВСК} и ДАИ_{СВ} также закономерна. Физический смысл ДАИ_{ВВ} – это отношение ферментов дегидрогеназ активного ила к биологически очищенной воде, в которой биологически окисляемый субстрат отсутствует, а ДАИ_{СВ} – тоже отношение к сточной воде, в которой достаточно субстрата и может присутствовать токсикант. Таким образом, чем меньше ДАИ_{ВВ} в начале процесса, а значит активный ил полностью прошел регенерацию, тем меньше будет и ДАИ_{СВ}, значит выше будет ДАИ_{СВ}, как реакция на поступление субстрата.

Далее изучили основную информацию о распределении данных, содержащихся в столбцах для чего использовали функцию числового распределения (`describe`). Для того чтобы не потерять точность при предсказании моделями значений Y_1 – Y_4 , значения этих столбцов нужно проверить на наличие выбросов значений, для этого были использованы функция построения графиков широкого назначения `boxplot` и выбран график распределения данных по квартилям, представленный рис. 3.

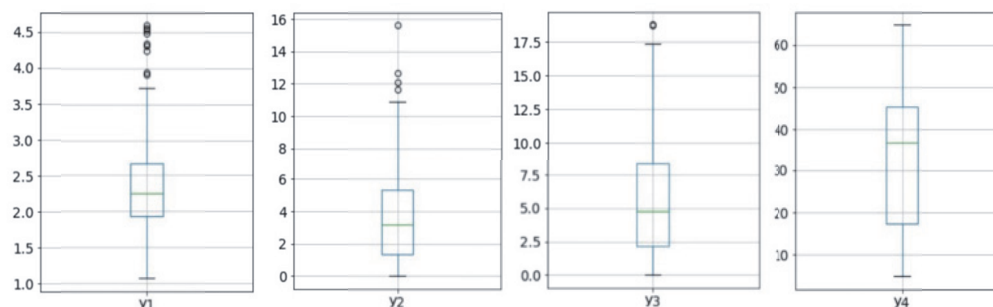


Рис. 3. Гистограмма распределения данных

По гистограмме распределения выбросами в данных будут являться числа ниже первого квартиля и выше третьего. В столбце функции Y_1 такими выбросами являются значения ниже 1,0 и выше 3,5. Выбросами столбца Y_2 являются данные ниже нуля и выше 11. Выбросами столбца Y_3 являются данные выше 17,5. В столбце Y_4 выбросов не обнаружено.

Использование машинного обучения для анализа экспериментальных данных показало, что 25 строк в представленной таблице содержат ошибочные записи. Отсутствие предварительных этапов по подготовке и очистке данных могло бы привести к обучению прогностических моделей биологической очистки сточных вод на неверных данных, что существенно ограничивает их точность и надежность на реальных объектах негативного воздействия на окружающую среду. Пробелы в строках данных препятствуют построению эффективных предиктивных моделей. Поэтому на следующем этапе исследования моделирование процессов биологической очистки сточных вод нефтехимического предприятия осуществлялось на основе подготовленного и очищенного датасета.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Исследование данных путем построения моделей

В данной работе были использованы четыре модели: линейная регрессия (Linear Regression), полиномиальная регрессия второго порядка (Polynomial Features), «дерево решений» (Decision Tree Regressor) и «случайный лес» (Random Forest Regressor).

Линейная регрессия – одна из важнейших и широко используемых техник регрессии, позволяющий установить взаимосвязь между функцией и независимыми переменными. Одним из его достоинств является легкость интерпретации результатов [8].

Полиномиальная регрессия описывает нелинейную взаимосвязь между значениями X и соответствующими условным средним Y , обозначаемым как $Y = f(X)$. Преимуществом полиномиальной регрессии является её способность наиболее точно аппроксимировать отношения между зависимой и независимыми переменными благодаря использованию многочленов, что позволяет эффективно учитывать широкий диапазон кривизны данных. Однако следует отметить, что полиномиальная зависимость очень чувствительна к выбросам, которые мы исключили на первом этапе работы с выборкой экспериментальных данных [9].

Модели «дерево решений» и «случайный лес» являются быстрыми по сравнению с другими алгоритмами и наиболее часто применяемыми алгоритмами контролируемого машинного обучения, выполняющими одновременно задачи регрессии и классификации [10–12].

Модель «случайный лес» (Random Forest) представляет собой ансамблевый метод, использующий множество деревьев решений, каждый из которых строится на случайно сформированном подмножестве исходных данных. Набор таких классификаторов формирует «лес», в котором каждое дерево генерируется на основе метрик выбора признаков, таких как прирост информации, соотношение прироста и индекс Джини [13]. Интерпретация модели «Случайный лес» более сложная по сравнению с «деревом решений», где можно легко определить результат, следуя по пути в «дереве» и даже формулируя правила по ходу движения по «дереву».

Для полученных и обработанных экспериментальных данных на примере токсиканта бензола и переменной Y_1 были оценены все исследуемые предиктивные модели с целью выявления наиболее адекватной модели, описывающей процесс биологической очистки.

Результаты анализа показали, что модель линейной регрессии демонстрирует предсказание значений Y_1 для токсиканта бензола с коэффициентом детерминации $R^2 = 0.82$, что соответствует 82%. Полиномиальная регрессия обеспечивает более высокий коэффициент детерминации ($R^2 = 0.97$ или 97%). Однако следует отметить, что данная модель имеет ограничение: не способна одновременно строить предсказания для нескольких зависимых переменных в рамках одного набора данных, при попытке предсказания других прогнозируемых переменных Y_2 , Y_3 и Y_4 она возвращает значения для прогноза лишь первой функции (Y_1), что ограничивает её применение для дальнейших расчетов и прогноза различных показателей процесса биологической очистки при воздействии токсикантов.

Для моделирования процесса биологической очистки в широком диапазоне концентраций бензола при использовании модели «дерево решений» (рис. 4) метрика точности модели составила 0.93 (или 93%).

Предсказание гидробиологического индекса Шеннона активного ила в процессе биологической очистки по модели «случайный лес» (рис. 5) при количестве деревьев 8 и глубине 10 достигается точность 0.97 (или 97%) по метрике Accuracy.

На основании проведенного анализа, модель «случайный лес» продемонстрировала наилучшие результаты по метрикам качества среди всех исследованных моделей, поэтому дальнейшие предсказания для всех токсикантов осуществлялись с применением этой модели [14]. Результаты расчетов показателя качества модели «случайный лес» Accuracy для прогноза в процессе биологической очистки индекс Шеннона (Y_1), дегидрогеназной (Y_2 и Y_3) и каталазной (Y_4) активности ила для каждого исследованного токсиканта приведены в табл. 2.

```

#модель дерева решений
best_model = None
best_result = 0
for depth in range(1, 5):
    model = DecisionTreeRegressor(random_state=12345, max_depth=depth) # обучение модели с заданной глубиной дерева
    model.fit(features_benzol_y1, target_benzol_y1) # обучение модели
    result = model.score(features_benzol_y1, target_benzol_y1) # качество модели
    print('max_depth =', depth, ': ', end='')
    print(result)
    if result > best_result:
        best_model = model
        best_result = result
        best_depth = depth

print("Accuracy лучшей модели:", best_result, "Глубина дерева:", best_depth)

max_depth = 1 : 0.611714466761513
max_depth = 2 : 0.8076859910469963
max_depth = 3 : 0.906498434682722
max_depth = 4 : 0.9334100474873684
Accuracy лучшей модели: 0.9334100474873684 Глубина дерева: 4

```

Рис. 4. Обучение модели «дерево решений» и расчет метрики Accuracy

```

#модель случайного леса
best_model = None
best_result = 0
best_est = 0
best_depth = 0
for est in range(1, 11):
    for depth in range(1, 11):
        model = RandomForestRegressor(random_state=12345, max_depth = depth, n_estimators=est) # модель с заданным количеством деревьев
        model.fit(features_benzol_y1, target_benzol_y1) # обучение модели на тренировочной выборке
        result = model.score(features_benzol_y1, target_benzol_y1) # качество модели на валидационной выборке
        if result > best_result:
            best_model = model # наилучшую модель
            best_result = result # наилучшее значение метрики accuracy на валидационных данных
            best_est = est
            best_depth = depth

print("Accuracy наилучшей модели на валидационной выборке:", best_result, "Количество деревьев:", best_est, "Максимальная глубина:", best_depth)

Accuracy наилучшей модели на валидационной выборке: 0.9747468990960966 Количество деревьев: 8 Максимальная глубина: 1
0

```

Рис. 5. Обучение модели «случайный лес» и расчет метрики Accuracy

Таблица 2. Показатель качества модели Accuracy для модели «случайный лес»

Токсикант	Y ₁	Y ₂	Y ₃	Y ₄
Бензол	0,97	0,98	0,96	0,96
Толуол	0,95	0,93	0,97	0,96
Стирол	0,96	0,92	0,91	0,97
Фенол	0,94	0,95	0,96	0,97
Этилбензол	0,96	0,91	0,93	0,93
Формальдегид	0,97	0,93	0,98	0,96

С точностью, представленной в табл. 2, построены модели предсказания для токсикантов (бензол, толуол, стирол, фенол, этилбензол, формальдегид) в изученном диапазоне изменения параметров (X) для принятия решения оператором о возможности приема на узел биологической очистки сточных вод и их очистки.

Реализация алгоритма для процедуры принятия решений по сбросу сточных вод на биологические очистные сооружения

В рамках исследования проведена апробация на биологических очистных сооружениях нефтехимического предприятия разработанного алгоритма для моделирования состояния активного ила в процессе биоочистки.

Пример использования разработанного алгоритма на основе модели «случайный лес» для токсиканта бензол демонстрирует, что оператор, вводя текущие показатели качества активного ила (X_1 - X_{11}) и данные о сточных водах, поступающих на очистку (в качестве примера, рис. 6), может получить результаты прогнозирования показателей состояния активного ила.

Моделирование позволяет оценить условия функционирования активного ила под воздействием определенных токсикантов, реакцию активного ила на поступление субстрата, т.е. возможность оценки бактериальной трансформации загрязняющих веществ сточных вод, а также сукцессии биоценоза.



Рис. 6. Моделирование показателей биологической очистки по модели «случайный лес»

Индекс биоразнообразия Шеннона зависит от двух факторов: от числа родов в биоценозе и распределения численности особей между отдельными родами. Чем выше индекс биоразнообразия, тем лучше будет обеспечено качество биологической очистки сточных вод [6,14].

При заданных в примере условиях модель прогнозирует, что по окончании процесса биологической очистки сточных вод с содержанием 1,0 мг/дм³ бензола индекс Шеннона снизится относительно исходного значения и составит 1,96, что хорошо коррелирует с экспериментальными и производственными данными, полученными на этапе адаптации и внедрения разработанного алгоритма на действующие очистные сооружения. Также модель предсказывает ДАИ_{СВ}, равное 2,26 мг/г, ДАИ_{НЖ}, равное 1,19 мг/г. Фермент дегидрогеназа относится к классу оксидоредуктаз, катализирующих реакцию дегидрогенизации, т.е. отщепления водорода у органического соединения. Фермент дегидрогеназа относится к конститутивным ферментам, так как дегидрирование занимает основную долю процессов при биологической очистке [4].

Такие высокие значения дегидрогеназной активности по отношению к надильовой жидкости и ее снижение по отношению к сточной воде свидетельствуют о повышенном остаточном содержании субстрата и промежуточных продуктов его метаболизма в клетках активной биомассы. Это делает невозможным организацию эффективного нового цикла процесса биологической очистки без глубокой регенерации активного ила. Продолжительность регенерации в данном случае может быть рассчитана оператором путем изменения продолжительности аэрации (X_3) в разработанном алгоритме.

На основании данных моделирования оператор очистных сооружений может принять обоснованное решение о возможности приема сточных вод для очистки на основе текущих технологических показателей процесса, необходимости их разбавления, используя для этого различные режимы работы насосного оборудования или регулирования запорной арматуры, или отказа в приеме стоков и перевода их в запасные емкости в случае их несоответствия установленным требованиям. В рассматриваемой моделируемой технологической ситуации прогнозируется рост каталазной активности до 45,78 мг H₂O₂/г·мин, то есть рост показателя стресса биомассы. Это хорошо согласуется с экспериментальными данными по оценке токсичности бензола и продуктов его биотрансформации [15].

Фермент каталаза находится почти во всех аэробно дышащих клетках и у некоторых факультативных

тивных анаэробов. Учитывая, что наибольшую роль в процессе биологической очистки сточных вод занимают окислительные процессы, перспективным является оценка и прогноз каталазной активности ила. Однако активность этого фермента значительно различается в зависимости от вида, возраста, стадии развития, физиологии и состояния организма [5].

Таким образом, разработанный алгоритм будет использован для оперативного управления, принятия технологических решений и информационной поддержки персонала очистных сооружений. Это несомненно повысит оперативность принятия решений и качество очистки сточных вод, особенно в условиях стажировки и работы нового персонала, что может существенно снизить процент недостаточно очищенных сточных вод за счет более точного регулирования процесса биологической очистки.

ВЫВОДЫ И ЗАКЛЮЧЕНИЕ

В результате проведенного исследования были получены экспериментальные данные о состоянии активного ила под действием широкого диапазона концентраций токсикантов (бензол, толуол, стирол, фенол, этилбензол, формальдегид) в процессе биологической очистки сточных вод, формирующихся на урбанизированных территориях на примере крупного промышленного нефтехимического комплекса Республики Татарстан. Данные подготовлены к анализу путем удаления повторов, выбросов и проверки пропусков с последующим удалением или заменой на медиану по признакам корреляции.

Получена оптимальная математическая модель, адекватно описывающая процесс биологической очистки сточных вод нефтехимического комплекса, с высоким значением коэффициента детерминации для всех изученных экотоксикантов (87-98%) – модель «случайный лес». Проведенные исследования направлены на экологическое оздоровление региона и муниципальных образований, снижение техносферной нагрузки, вызванной негативным воздействием сброса недостаточно очищенных сточных вод в природные водоприемники.

Реализация управленческих решений на основе разработанного алгоритма предоставляют персоналу возможность обоснованно принимать на очистку и распределять по очередям сточные воды при текущих технологических показателях, что позволит предупредить потенциально возможные техногенные чрезвычайные ситуации на объекте негативного воздействия на окружающую природную среду.

СПИСОК ЛИТЕРАТУРЫ

1. Указ Президента РФ от 9 мая 2017 г. №203 «О Стратегии развития информационного общества в Российской Федерации на 2017 - 2030 годы» // URL: <http://www.kremlin.ru/acts/bank/41919> (дата обращения 20.12.2024).
2. Яговкин, Н.Г. Совершенствование процесса принятия человеком управленческих решений при эксплуатации технологического оборудования / Яговкин Н.Г. // Известия Самарского научного центра РАН. – 2024. – Т. 26. – № 1. – С. 141-145.
3. Никифорова, Т.И. Цифровизация ЖКХ как основа развития отрасли / Т.И. Никифорова, Н.И. Нижальская // Индустриальная экономика. – 2022. – Т. 4. – № 2. – С. 125-129.
4. Михлин, Д.М. Биологическое окисление / Д.М. Михлин. – М.: Изд-во АН СССР, 1956. – 443 с.
5. Негода, Л.Л. Оценка состояния биоценоза биологической пленки по активности окислительно-восстановительного фермента - каталазы / Л.Л. Негода // Химия и технология воды. – 1986. – Т. 8. – № 6. – С. 82-86.
6. Achmadulina, F.Y. Comparison of bioindicator eukaryotes of activated sludge biocenoses on two water-treatment plants: A case study / F.Y. Achmadulina, R.K. Zakirov, E.S. Balymova, V. Denisova, T. Brovdová, J. Trögl, M. Neruda // Nova Biotechnologica et Chimica. 2017. Vol. 16. № 1. P. 54-60.
7. Якимчик, А.И. Jupyter Notebook: система интерактивных научных вычислений / А.И. Якимчик // Геофизический журнал. – 2019. – Т. 41. – № 2. – С. 112-121.
8. Коросов, А.В. Нейронные сети для экологии: введение / А.В. Коросов // Принципы экологии. – 2023. – № 3. – С. 76-96.
9. Iskandaryan D., Ramos F., Trilles S. Air quality prediction in smart cities using machine learning technologies based on sensor data: a review // Applied Sciences. 2020. Vol. 10. № 7. P. 2401.
10. Плескачев, Ю.А. Применение методов машинного обучения в агентном моделировании / Ю.А. Плескачев, Д.М. Радченко, Д.Ю. Евдокимов, М.В. Юлусов // Новое в экономической кибернетике. – 2024. – № 2. – С. 11-27.
11. Мифтахова, А.А. Применения метода дерева решений для решения задач классификации и прогнозирования / А.А. Мифтахова // Инфокоммуникационные технологии. – 2016. – Т. 14. № 1. – С. 64-70.
12. Зинченко, Т.Д. Донные сообщества и абиотические факторы: анализ статистической связи с использованием индекса нестабильности и метода виртуальных видов / Т.Д. Зинченко, В.К. Шитиков, Л.В. Головатюк // Журн. Сиб. федер. ун-та. Биология. – 2021. – Т.14. – № 2. – С. 119-132.
13. Шитиков, В.К. Классификация, регрессия и другие алгоритмы Data Mining с использованием R / В.К. Шитиков, С.Э. Мастицкий // Электронная книга. – 2017. – URL: <https://ranalytics.github.io/data-mining/> (дата обращения 10.12.2024).

14. Новикова, А.Р. Моделирование процесса биологической очистки сточных вод нефтехимического производства / А.Р. Новикова, Е.С. Балымова, Р.К. Закиров, Ф.Ю. Ахмадуллина // Экология и безопасность жизнедеятельности промышленно-транспортных комплексов ELPIT 2023. Сборник трудов девятого международного экологического конгресса (одиннадцатой международной научно-технической конференции). – Тольятти, 2023. – С. 42-48.
15. Михайлова, И.В. Влияние бензола на иммунную систему и некоторые механизмы его действия / И.В. Михайлова, А.И. Смолягин, С.И. Красиков, А.В. Караулов // Иммунология. – 2014. – №1. – С. 51-55.

FORECASTING OPTIMAL WORK PROCESS PARAMETERS BIOLOGICAL WASTEWATER TREATMENT BASED ON NUMERICAL MODELING

© 2025 E.S. Balymova, A.R. Novikova, R.K. Zakirov, F.Yu. Akhmadullina

Kazan National Research Technological University, Kazan, Russia

The article discusses the issues of digital transformation and implementation of breakthrough technologies in the field of water disposal and biological treatment of wastewater in petrochemical production. In conditions of instability in the composition and flow rate of wastewater, high concentrations of toxic substances in industrial effluents, modeling the biotreatment process and implementing algorithms to support decision-making on wastewater discharge to biological treatment facilities is becoming a key area for ensuring the sustainable development of urbanized areas and environmental protection. The article presents the results of forecasting the optimal parameters of biological wastewater treatment at an operating biostation of a petrochemical enterprise in the Republic of Tatarstan, which will help make informed decisions by treatment facility operators regarding the possibility of accepting wastewater for treatment, the conditions for their treatment and forecasting the effectiveness of biological treatment processes. The research methodology includes laboratory modeling of biological wastewater treatment processes using samples of activated sludge and wastewater from an operating biostation of a petrochemical enterprise. Based on the experimental data, an analysis of the influence of toxicants (benzene, toluene, styrene, phenol, ethylbenzene, formaldehyde) on the efficiency of purification and the physicochemical and biochemical properties of activated sludge, as well as indicators of general pollution and toxicity of wastewater, was carried out. The results of the study allow not only to increase the accuracy of forecasting, but also to optimize work processes at treatment facilities, which will undoubtedly increase the efficiency of purification of highly concentrated toxic wastewater. The proposed solution - the optimal mathematical model «random forest», adequately describing the process of biological purification of wastewater of the petrochemical complex, contributes to the sustainable development and safe operation of petrochemical industries. The conducted studies are aimed at environmental improvement of the region and municipalities, reducing the technosphere load caused by the negative impact of the discharge of insufficiently purified wastewater into natural water receivers.

Keywords: activated sludge, wastewater, biological treatment, machine learning, toxicants, hydrobiological indices, mathematical model, random forest.

DOI: 10.37313/1990-5378-2025-27-1-166-175

EDN: WBGMYX

REFERENCES

1. Ukaz Prezidenta RF ot 9 maya 2017 g. №203 «O Strategii razvitiya informacionnogo obshchestva v Rossijskoj Federacii na 2017 - 2030 gody» // URL: <http://www.kremlin.ru/acts/bank/41919> (data obrashcheniya 20.12.2024).
2. Yagovkin, N.G. Sovershenstvovanie processa prinyatiya chelovekom upravlencheskih reshenij pri ekspluatatsii tekhnologicheskogo oborudovaniya / Yagovkin N.G. // Izvestiya Samarskogo nauchnogo centra RAN. – 2024. – T. 26. – № 1. – S. 141-145.
3. Nikiforova, T.I. Cifrovizaciya ZhKH kak osnova razvitiya otrasli / T.I. Nikiforova, N.I. Nizhal'skaya // Industrial'naya ekonomika. – 2022. – T. 4. – № 2. – S. 125-129.
4. Mihlin, D.M. Biologicheskoe okislenie / D.M. Mihlin. – M.: Izd-vo AN SSSR, 1956. – 443 s.
5. Negoda, L.L. Ocenka sostoyaniya biocenoza biologicheskoy plenki po aktivnosti okislitel'no-vosstanovitel'nogo fermenta - katalazy / L.L. Negoda // Himiya i tekhnologiya vody. – 1986. – T. 8. – № 6. – S. 82-86.
6. Achmadulina, F.Y. Comparison of bioindicator eukaryotes of activated sludge biocenoses on two water-treatment plants: A case study / F.Y. Achmadulina, R.K. Zakirov, E.S. Balymova, V. Denisova, T. Brovdiová, J. Trögl, M. Neruda // Nova Biotechnologica et Chimica. 2017. Vol. 16. № 1. P. 54-60.
7. Yakimchik, A.I. Jupyter Notebook: sistema interaktivnyh nauchnyh vychislenij / A.I. Yakimchik // Geofizicheskij zhurnal. – 2019. – T. 41. – № 2. – S. 112-121.
8. Korosov, A.V. Nejronnye seti dlya ekologii: vvedenie / A.V. Korosov // Principy ekologii. – 2023. – № 3. – S. 76-96.
9. Iskandaryan D., Ramos F., Trilles S. Air quality prediction in smart cities using machine learning technologies based on sensor data: a review // Applied Sciences. 2020. Vol. 10. № 7. P. 2401.

10. *Pleskachev, Yu.A.* Primenenie metodov mashinnogo obucheniya v agentnom modelirovanii / Yu.A. Pleskachev, D.M. Radchenko, D.Yu. Evdokimov, M.V. Yulusov // *Novoe v ekonomicheskoy kibernetike*. – 2024. – № 2. – S. 11-27.
11. *Miftahova, A.A.* Primeneniya metoda dereva reshenij dlya resheniya zadach klassifikacii i prognozirovaniya / A.A. Miftahova // *Infokommunikacionnye tekhnologii*. – 2016. – T. 14. № 1. – С. 64-70.
12. *Zinchenko, T.D.* Donnye soobshchestva i abioticheskie faktory: analiz statisticheskoy svyazi s ispol'zovaniem indeksa nestabil'nosti i metoda virtual'nyh vidov / T.D. Zinchenko, V.K. Shitikov, L.V. Golovatyuk // *Zhurn. Sib. feder. un-ta. Biologiya*. – 2021. – T.14. – № 2. – S. 119-132.
13. *Shitikov, V.K.* Klassifikaciya, regressiya i drugie algoritmy Data Mining s ispol'zovaniem R / V.K. Shitikov, S.E. Mastickij // *Elektronnaya kniga*. – 2017. – URL: <https://ranalytics.github.io/data-mining/> (data obrashcheniya 10.12.2024).
14. *Novikova, A.R.* Modelirovanie processa biologicheskoy ochistki stochnyh vod neftekhimicheskogo proizvodstva / A.R. Novikova, E.S. Balymova, R.K. Zakirov, F.Yu. Ahmadullina // *Ekologiya i bezopasnost' zhiznedeyatel'nosti promyshlenno-transportnyh kompleksov ELPIT 2023. Sbornik trudov devyatogo mezhdunarodnogo ekologicheskogo kongressa (odinnadcatoy mezhdunarodnoj nauchno-tekhnicheskoy konferencii)*. – Tol'yatti, 2023. – S. 42-48.
15. *Mihajlova, I.V.* Vliyanie benzola na immunnuyu sistemu i nekotorye mekhanizmy ego dejstviya / I.V. Mihajlova, A.I. Smolyagin, S.I. Krasikov, A.V. Karaulov // *Immunologiya*. – 2014. – № 1. – S. 51-55.

Elena Balymova, Candidate of Technics, Associate Professor at the Industrial Biotechnology Department.

Email: ilc2013@inbox.ru

Aigul Novikova, Graduate Student at the Industrial Biotechnology Department. E-mail: aignovikova@yandex.ru

Rustem Zakirov, Candidate of Technics, Associate Professor at the Industrial Biotechnology Department. E-mail: zakrus@mail.ru

Farida Akhmadullina, Senior Lecturer at the Department of Industrial Biotechnology. E-mail: akhmadullina1951@mail.ru