

Использование многомерного анализа для выделения гидрохимических маркеров антропогенного воздействия

Л.В. Новикова, Н.Ю. Степанова*

Казанский (Приволжский) федеральный университет,
РФ, 420008, РТ, г. Казань, ул. Кремлевская, д.18

* Адрес для переписки: Ludmila.Novikova@kpfu.ru

Реферат. Нефтегазодобывающая деятельность остается приоритетным источником загрязнения водных экосистем. Однако отсутствие специфических маркеров затрудняет идентификацию виновного в причинении накопленного экологического ущерба. Целью работы было выявление комплекса индикаторных показателей характерных для нефтедобывающей деятельности, с использованием многомерного статистического анализа. Исследование проведено на основе 232 проб поверхностных вод, отобранных на реке Шешме, в водосборном бассейне которой ведется нефтедобыча, в сравнении с рекой Мёшей, в бассейне которой осуществляется преимущество аграрное производство. Анализировались 30 параметров, включая стандартные показатели качества воды, элементный анализ, органические загрязняющие вещества (по ХПК, БПК₅). Применены методы описательной статистики (тест Шапиро-Уилка для проверки нормальности), непараметрического сравнения групп (Манна-Уитни, U-критерий), анализ главных компонент (PCA), иерархическая кластеризация (метод Ward.D2) и метод K-средних (K-means) с валидацией, пермутационный многомерный анализ дисперсии (PERMANOVA).

По результатам анализа гидрохимических данных реки достоверно различаются по 65.6% показателей. Выявлена триада специфических маркеров нефтедобывающей деятельности: хлориды (в исследуемой реке медиана в 6.2 раза выше), молибден и стронций (оба маркера пластовых вод), а также кадмий и свинец. Анализ главных компонент показал четкое разделение облаков точек гидрохимических данных двух рек (перекрытие менее 10%). Основной вклад в PC1 вносит барий (9.63%), в PC2 – БПК₅ (10.73%). Кластеризация выделила три уровня загрязнения: условно-чистая вода ($n \approx 71$), среднее загрязнение ($n \approx 140$) и высокое загрязнение ($n \approx 16$ все пробы из р. Шешмы). PERMANOVA подтвердила статистическую значимость различий ($F_{1,220} = 19.039$, $p = 0.001$).

На основе многомерного анализа идентифицированы пять специфических индикаторов нефтедобывающей деятельности, рекомендованные к внедрению в системы производственного и экологического мониторинга. Работа имеет важное практическое применение для идентификации источников загрязнения и установления ответственности за накопленный экологический ущерб в нефтедобывающих регионах, а также теоретическое значение для развития многомерных методов экологической диагностики.

Ключевые слова. Нефтяное загрязнение, поверхностные воды, гидрохимический мониторинг, PCA.

Using multidimensional analysis to identify water pollution markers

L.V. Novikova, N.Yu. Stepanova*

Kazan Federal University,
18, Kremlyovskaya str., 420008, Kazan, Russian Federation

*Correspondence address: *Ludmila.Novikova@kpfu.ru*

Abstract. Oil and gas production remain a primary source of pollution for aquatic ecosystems. However, the lack of specific markers complicates the identification of parties responsible for cumulative environmental damage. The aim of this study was to identify a set of indicators specific to oil production activities through multivariate statistical analysis. The study was conducted using 232 surface water samples collected from the Sheshma River, with a catchment area dominated by oil extraction, compared to the Myosha River, where the catchment is predominantly agricultural. Thirty parameters were analyzed, including standard water quality indicators, elemental analysis, and organic pollutants (COD, BOD₅). The following methods were used: descriptive statistics (Shapiro-Wilk test for normality), nonparametric comparison of groups (the Mann-Whitney U test), principal component analysis (PCA), hierarchical clustering (Ward's D2 method), the K-means method with validation, and permutational multivariate analysis of variance (PERMANOVA).

According to the data analysis, the rivers differ reliably in 65.6% of the parameters. A triad of specific markers of oil production activity was identified: chlorides (6.2-fold excess), molybdenum and strontium (both markers of formation waters), as well as cadmium and lead. Principal component analysis showed a clear separation of the point clouds of hydrochemical data of the two rivers (overlapping less than 10%). The main contribution to PC1 is made by barium (9.63%), and to PC2 – by BOD₅ (10.73%). Clustering identified three levels of pollution: conventionally clean water ($n \approx 71$), moderate pollution ($n \approx 140$), and high pollution ($n \approx 16$, all samples from the Sheshma River). PERMANOVA confirmed the statistical significance of the differences ($F_{1,220}=19.039$, $p=0.001$).

Based on a multivariate analysis, five specific indicators of oil production activities were identified and are recommended for implementation in industrial and environmental monitoring systems. This work has important practical applications for identifying pollution sources and establishing liability for environmental damage in oil-producing regions, as well as theoretical significance for the development of multivariate methods in environmental diagnostics.

Keywords. Oil pollution, surface water, hydrochemical monitoring, PCA.

Введение

Нефтегазодобывающая деятельность по-прежнему остается приоритетным загрязнителем водных объектов. Причиной являются не только аварийные случаи разлива нефтепродуктов (Гарелина, 2025; Кленкин, Агапов, 2011), но и результаты штатной работы производственных объектов¹⁾ (Рахматуллин и др., 2023; Титов, Ходов, 2015). В воду могут попадать сырая нефть, нефтепродукты, шлам, буровой раствор, пластовые воды, содержащие различные химические соединения естественного происхождения (хлориды, сульфаты, тяжелые металлы, органические загрязняющие вещества, включая ПАУ, ВТЕХ), а также используемые для поддержания нефтеотдачи пластов (загустители, коагулянты, флокулянты, ингибиторы коррозии, регуляторы, стабилизаторы, эмульгаторы) (Лешкович и др., 2020; Bakke et al., 2013; Yusta-García et al., 2017).

Развитая нефтегазовая промышленность обуславливает высокие техногенные нагрузки на водные объекты и соответственно требуется наблюдение за их состоянием. Традиционные показатели, такие как нефтепродукты, уже не являются специфическими и не могут в полной мере охарактеризовать влияние нефтяной промышленности на состояние поверхностных и подземных вод в силу масштабного загрязнения окружающей среды (ОС) и их поступления в водные объекты с поверхностным стоком.

Целью нашей работы было выявление индикаторов нефтегазодобывающей деятельности с применением многомерного статистического анализа для последующего использования их в рамках производственного экологического мониторинга и при идентификации виновного в причинении накопленного экологического вреда.

Методы и материалы

Для выделения показателей, характерных для нефтедобывающей деятельности, были выбраны реки, сходные по географическим и гидрологическим условиям, но отличающиеся отсутствием или наличием нефтегазодобывающей деятельности в бассейне реки. Для соблюдения данного условия использовались такие критерии как географическая близость объектов; равенство площади и зонирования водосборного бассейна; схожесть преобладающих типов хозяйственной деятельности, объема стока; а также отсутствие крупных населенных пунктов, промышленных предприятий и искусственного зарегулирования стока.

По результатам предварительного анализа была выбрана р. Шешма, на водосборе которой ведется нефтедобыча, в качестве референтного водного объекта – р. Мёша. Обе реки являются притоками 1 порядка р. Камы.

¹⁾ Влияние добычи и разработки нефти и газа на экологию водно-болотных угодий района "Нигер-дельта" URL: <https://www.elibrary.ru/item.asp?id=39948240>.

Площадь водосбора р. Мёши составляет 4.18 тыс. км², длина реки 186.4 км, густота речной сети 0.35 км·км⁻². По гидрохимическому режиму вода относится к гидрокарбонатно-кальциевому типу с малой минерализацией в межень. Площадь водосбора р. Шешмы составляет 6.2 тыс. км², длина – 234.1 км, густота речной сети – 0.34 км·км⁻². Вода в реке гидрокарбонатно-хлоридно-кальциевая, средней минерализации (до 700 мг·л⁻¹) (Водные объекты..., 2018; Ермолаев и др., 2007). Питание у обеих рек смешанное, а бассейны имеют сходное распределение площадей под разными классами поверхностей земли (табл. 1).

Таблица 1. Классы поверхностей земли: ESA WorldCover 10m v200, %

Table 1. Land cover classes: ESA WorldCover 10m v200, %

Классы поверхностей земли: ESA WorldCover 10m v200	Мёша	Шешма
Земли под древесно-кустарниковой растительностью	0.0%	0.0%
Земли, занятые травянистой растительностью	21.7%	28.6%
Пахотные земли:	49.0%	45.3%
Земли населенных пунктов:	1.6%	0.6%
Голая или разреженная растительность (пустыни, скалистые территории):	0.1%	0.1%
Земли водного фонда	2.6%	0.4%
Травянистые водно-болотные угодья:	0.3%	0.2%
Древесный покров (леса и лесостепи):	24.7%	24.8%

В бассейне р. Шешмы ведутся как сельскохозяйственная, так и нефтедобывающая деятельности, на водосборе р. Мёши только сельскохозяйственная. В бассейнах обеих рек отсутствуют крупные промышленные предприятия; загрязнение поверхностных вод происходит в результате эрозионных процессов, несоблюдения агротехнических мероприятий при обработке почв, диффузного стока с территории водосбора. На водосборе р. Шешмы возможно антропогенное воздействие при разработке и эксплуатации кустов скважин, аварийных порывах нефтепроводов.

Пробы поверхностных вод отбирались по течению обеих рек в соответствии с требованиями нормативной документации (РД 52.24.353-12, 2012) с глубины 0.5 м от поверхности (рис. 1). На р. Мёше на 21 станции было отобрано 122 пробы, на р. Шешме на 18 станциях – 110 проб.

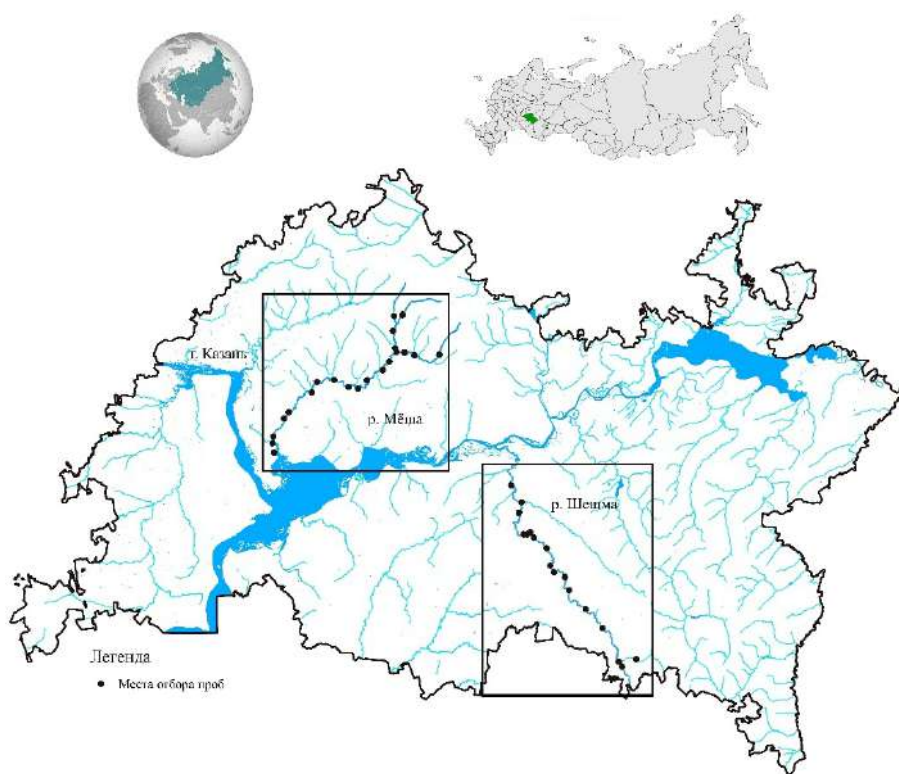


Рисунок 1. Схема отбора проб поверхностных вод

Figure 1. Surface water sampling map

Для пробоподготовки и анализа использовались утвержденные методики. В пробах воды определяли pH и электропроводность (РД 52.24.495-2017)²⁾, растворенный кислород (РД 52.24.419-2019)³⁾, БПК₅ (ПНД Ф 14.1:2:3:4.123-97)⁴⁾. Хлориды определяли аргентометрическим методом (ПНД Ф 14.1:2.96-97)⁵⁾, сульфаты – турбидиметрическим (ПНД Ф 14.1:2.159-2000)⁶⁾, гидрокарбонаты – титриметрическим (ПНД Ф 14.2.99-97, 2017)⁷⁾,

²⁾ РД 52.24.495-2017 Водородный показатель и удельная электрическая проводимость вод методика выполнения измерений электрометрическим методом.

³⁾ РД 52.24.419-2019 Массовая концентрация растворенного кислорода в водах: методика выполнения измерений йодометрическим методом.

⁴⁾ ПНД Ф 14.1:2:3:4.123-97 Методика выполнения измерений биохимического потребления кислорода после n-дней инкубации (БПК_{полн.}) в поверхностных пресных, подземных (грунтовых), питьевых, сточных и очищенных сточных водах.

⁵⁾ ПНД Ф 14.1:2.96-97 Методика выполнения измерений массовой концентрации хлоридов в пробах природных и очищенных сточных вод аргентометрическим методом.

⁶⁾ ПНД Ф 14.1:2.159-2000 Методика выполнения измерений массовой концентрации сульфат-ионов в пробах природных и сточных вод турбидиметрическим методом.

⁷⁾ ПНД Ф 14.2.99-97 Методика выполнения измерений массовой концентрации гидрокарбонатов в пробах природных вод титриметрическим методом.

ХПК – фотометрическим (ПНД Ф 14.1:2:4.210-2005)⁸⁾, азотную группу – фотометрическим: NH_4^+ с реактивом Несслера (ПНД Ф 14.1:2:4.262-2010)⁹⁾, NO_2^- – с реактивом Грисса (ПНД Ф 14.1:2:4.3-95)¹⁰⁾, NO_3^- – с салициловой кислотой (ПНД Ф 14.1:2:4.4-95)¹¹⁾, фосфаты – фотометрическим с молибдатом аммония (ПНД Ф 14.1:2:3:4.112-2023)¹²⁾, нефтепродукты (НП) – ИК-спектрофотометрии (ПНД Ф 14.1:2:4.168-2000)¹³⁾, СПАВ (ГОСТ 31857-2012, 2014)¹⁴⁾, гравиметрическим – взвешенные вещества (ВВ) (ПНД Ф 14.1:2.110-97)¹⁵⁾ и сухой остаток (ПНД Ф 14.1:2.114-97)¹⁶⁾, атомно-эмиссионной спектрометрии – Al, As, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Zn, Mo, Ba, Sr, Se, V, P (ПНД Ф 14.1:2:4.135-98)¹⁷⁾.

Статистическая обработка результатов осуществлялась в программе Statistica 8.0 и в RStudio с использованием языка R. Проверка на нормальность проводилась с использованием тестов нормальности Шапиро-Уилка, сравнение групп с использованием t-теста Манна-Уитни для данных с распределением отличным от нормального. Также были проведены анализ главных компонент (PCA), кластеризация данных (HCA и k-means) с последующей валидацией результатов с использованием Silhouette Score, Davies-Bouldin index, Calinski-Harabasz index. Статистическую значимость различий между реками оценивали с использованием пермутационного многомерного анализа дисперсии (PERMANOVA). Картографический материал оформлен в QGIS 3.4.15.

Для разработки количественной системы оценки риска загрязнения вод применили логистическую регрессию и дерево решений (Decision Tree). Отбор переменных проводился методом последовательного включения с кросс-валидацией ($k=5$). Кадмий исключен из-за переобучения на обучающей

⁸⁾ ПНД Ф 14.1:2:4.210-2005 Методика измерений химического потребления кислорода (ХПК) в пробах питьевых, природных и сточных вод фотометрическим методом.

⁹⁾ ПНД Ф 14.1:2:4.262-2010 Методика измерений массовой концентрации ионов аммония в питьевых, поверхностных (в том числе морских) и сточных водах фотометрическим методом с реактивом Несслера.

¹⁰⁾ ПНД Ф 14.1:2:4.3-95 Методика измерений массовой концентрации нитрит-ионов в питьевых, поверхностных и сточных водах фотометрическим методом с реактивом Грисса.

¹¹⁾ ПНД Ф 14.1:2:4.4-95 Методика измерений массовой концентрации нитрат-ионов в питьевых, поверхностных и сточных водах фотометрическим методом с салициловой кислотой.

¹²⁾ ПНД Ф 14.1:2:3:4.112-2023 Методика измерений массовой концентрации фосфат-ионов (в том числе с пересчетом на массовую концентрацию фосфора фосфатов) в пробах питьевых, природных (поверхностных и подземных) и сточных вод фотометрическим методом с молибдатом аммония.

¹³⁾ ПНД Ф 14.1:2:4.168-2000 Методика измерений массовой концентрации нефтепродуктов в пробах питьевых, природных и очищенных сточных водах методом ИК-спектрофотометрии с применением концентратометров серии КН.

¹⁴⁾ ГОСТ 31857-2012 Методы определения содержания поверхностно-активных веществ.

¹⁵⁾ ПНД Ф 14.1:2.110-97 Методика выполнения измерений содержаний взвешенных веществ и общего содержания примесей в пробах природных и очищенных сточных вод гравиметрическим методом.

¹⁶⁾ ПНД Ф 14.1:2.114-97 Методика выполнения измерений массовой концентрации сухого остатка в пробах природных сточных вод.

выборке ($AUC_{обуч}=0.98$ против $AUC_{провер}=0.72$), свинец – из-за недообучения (низкие β и $p>0.5$). В качестве независимых переменных выбраны три гидрохимических показателя: концентрация хлоридов (Cl), молибдена (Mo) и стронция (Sr), которые были выделены в качестве маркеров по результатам исследования. Зависимая переменная кодировалась бинарным образом: 1 = воды подвергнутые воздействию нефтедобычи (р. Шешма), 0 = «фоновые воды» (р. Мёша).

Модели обучены на 83 пробах воды (42 – из р. Мёша, 41 – из р. Шешма). Для нормализации распределения и стабилизации дисперсии концентрации Cl, Mo и Sr в логистической регрессии подвергли логарифмическому преобразованию ($\log_{10}$) с последующей стандартизацией методом z-score. Логистическая регрессия реализована с использованием функции связи logit в обобщенной линейной модели (GLM). Параметры модели оценены методом максимального правдоподобия.

Дерево решений построено алгоритмом CART (Classification and Regression Trees) с параметром сложности $cp = 0.01$ и минимальным числом наблюдений в узле $minsplit = 5$. Качество моделей оценено стандартными метриками классификации: точность (Accuracy), чувствительность/Полнота (Sensitivity), специфичность (Specificity), точность прогнозов (Precision), F1-Score, площадь под ROC-кривой (AUC)

Статистическую значимость коэффициентов моделей оценили с использованием z-тестов при уровне значимости $\alpha = 0.05$. Расчеты выполнены в программной среде R (версия 4.5) с использованием пакетов tidyverse, gpart, pROC и ggplot2.

Результаты

Описательная статистика

Анализ данных (табл. 2) показал, что реки достоверно различаются по 65.6% показателей, качество воды в р. Шешме отличается от контрольной реки в худшую сторону.

Наибольшая кратность превышения ПДК отмечается по содержанию в воде р. Шешмы растворенных форм Cu и Mn – от 3 до 8 ПДК. При этом среднее содержание Mn в р. Мёше превышает содержание в Шешме в 1.8 раза (рис. 2). Кратность превышения ПДК по содержанию Cu в р. Шешме в среднем в 1.5 раза выше, чем в р. Мёше. Содержание Se превышает ПДК в среднем в 5 раз в обеих реках.

¹⁷⁾ ПНД Ф 14.1:2:4.135-98 МВИ массовых концентраций металлов методом атомно-эмиссионной спектроскопии с индуктивно связанной плазмой в питьевой, природной, сточной водах и атмосферных осадках.

Таблица 2. Химический состав природных вод

Table 2. Chemical composition of natural waters

Показатель	Мёша	Шешма	Показатель	Мёша	Шешма
pH*	$\frac{8.11 \pm 0.04}{8.05}$	$\frac{8.25 \pm 0.04^1}{8.34}$	Al ²	$\frac{26.72 \pm 1.89}{23.40}$	$\frac{52.92 \pm 11.44}{29.10}$
Гидро-карбонаты ^{3*}	$\frac{277.85 \pm 10.61}{300.84}$	$\frac{364.00 \pm 12.35}{370.00}$	As ²	$\frac{11.40 \pm 1.79}{6.82}$	$\frac{9.63 \pm 1.33}{4.34}$
Нитрит-ион ^{3*}	$\frac{0.08 \pm 0.008}{0.06}$	$\frac{0.06 \pm 0.009}{0.04}$	Ba ^{2*}	$\frac{54.58 \pm 0.75}{54.94}$	$\frac{68.16 \pm 0.47}{67.79}$
Нитрат-ион ^{3*}	$\frac{5.64 \pm 0.32}{5.58}$	$\frac{8.77 \pm 0.69}{9.10}$	Cd ^{2*}	$\frac{0.017 \pm 0.008}{0}$	$\frac{1.29 \pm 0.30}{0}$
Ион аммония ^{3*}	$\frac{0.14 \pm 0.02}{0.09}$	$\frac{0.15 \pm 0.03}{0.03}$	Co ²	$\frac{12.36 \pm 1.45}{17.05}$	$\frac{8.33 \pm 1.20}{0.05}$
Сульфаты ³	$\frac{144.15 \pm 6.84}{131.15}$	$\frac{134.76 \pm 9.01}{129.33}$	Cr*	$\frac{0.70 \pm 0.22}{0}$	$\frac{0.80 \pm 0.15}{0.45}$
Фосфаты ³	$\frac{0.05 \pm 0.007}{0.03}$	$\frac{0.093 \pm 0.01}{0.10}$	Cu ^{2*}	$\frac{5.56 \pm 0.91}{1.50}$	$\frac{6.83 \pm 0.90}{3.20}$
Хлориды ^{3*}	$\frac{11.26 \pm 0.36}{11.03}$	$\frac{63.45 \pm 2.51}{68.77}$	Fe ²	$\frac{94.06 \pm 16.57}{41.80}$	$\frac{77.87 \pm 5.19}{75.75}$
Сухой остаток ^{3*}	$\frac{538.14 \pm 21.73}{534.00}$	$\frac{605.27 \pm 16.79}{622.00}$	Ni ²	$\frac{5.92 \pm 1.04}{0}$	$\frac{4.89 \pm 0.78}{0.30}$
ВВ ^{3*}	$\frac{20.90 \pm 1.97}{12.05}$	$\frac{120.82 \pm 18.24}{164.00}$	Mn ^{2*}	$\frac{58.34 \pm 3.66}{54.04}$	$\frac{33.08 \pm 2.29}{26.88}$
Минерализация ^{3*}	$\frac{616.41 \pm 24.69}{670.30}$	$\frac{758.32 \pm 23.40}{772.00}$	Mo ^{2*}	$\frac{3.87 \pm 0.15}{3.76}$	$\frac{4.84 \pm 0.09}{4.91}$
БПК ³	$\frac{1.37 \pm 0.07}{1.22}$	$\frac{1.38 \pm 0.18}{1.39}$	Pb ^{2*}	$\frac{0.34 \pm 0.11}{0}$	$\frac{6.11 \pm 1.19}{0}$
СПАВ ^{3*}	$\frac{0.003 \pm 0.0005}{0}$	$\frac{0.01 \pm 0.002}{0.01}$	Se ²	$\frac{11.30 \pm 0.27}{11.30}$	$\frac{11.18 \pm 0.25}{11.36}$
НП ³	$\frac{0.03 \pm 0.003}{0.02}$	$\frac{0.04 \pm 0.005}{0.02}$	Sr ^{2*}	$\frac{1098.63 \pm 69.79}{1032.16}$	$\frac{1272.96 \pm 27.46}{1267.01}$
ХПК ^{3*}	$\frac{13.24 \pm 0.67}{11.80}$	$\frac{19.24 \pm 1.91}{19.20}$	V ^{2*}	$\frac{0.995 \pm 0.05}{0.98}$	$\frac{1.15 \pm 0.04}{1.15}$
p ²	$\frac{15.09 \pm 4.24}{2.55}$	$\frac{11.31 \pm 4.23}{0}$	Zn ^{2*}	$\frac{11.47 \pm 1.54}{1.42}$	$\frac{14.10 \pm 1.47}{10.30}$

Примечания: ¹ Среднее ± ошибка / медиана;

² мкг л⁻¹;

³ мг л⁻¹;

* выделены достоверно различающиеся показатели.

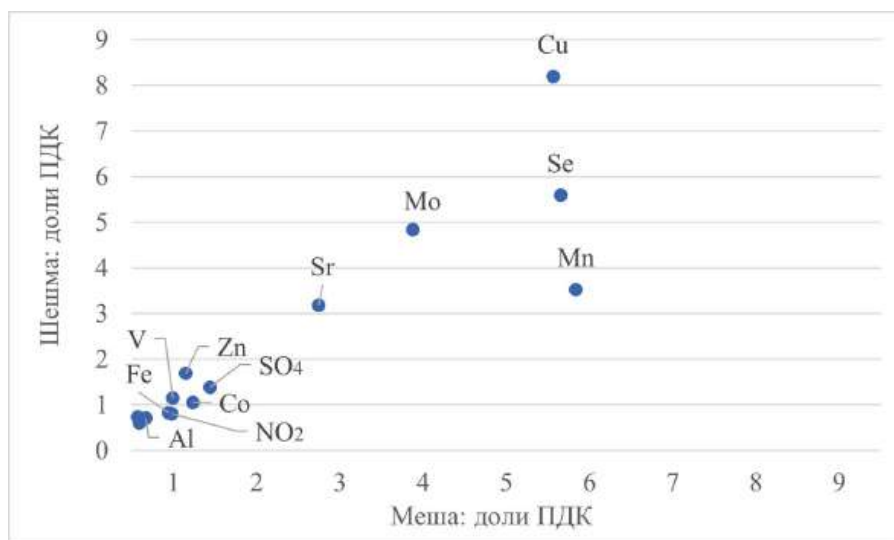


Рисунок 2. Кратность превышения ПДК

Figure 2. The rate of excess of the MPC

Выделяются показатели, содержание которых в воде не превышает ПДК, но сильно различается в реках. Так среднее содержание Cd в 75 раз выше в р. Шешме – $1.29 \text{ мкг} \cdot \text{л}^{-1}$ против $0.017 \text{ мкг} \cdot \text{л}^{-1}$ в р. Мёше. Сильная правосторонняя асимметрия в распределении Cd (1.89 в Шешме) подтверждает наличие регулярного антропогенного загрязнения вод, а не диффузного фонового содержания как в р. Мёше, где асимметрия составляет 4.2. Также высокая асимметрия Cd в р. Шешме говорит о том, что среднее значение не репрезентативно и необходимо использовать медиану, являющуюся более адекватным показателем «центра» распределения. Аналогичная ситуация наблюдается и в отношении Pb. Среднее содержание хлоридов в 6 раз выше в воде р. Шешме, что может быть обусловлено попаданием высокоминерализованных промышленных вод, применяемых на нефтепромыслах для поддержания пластового давления. Асимметрия данных по Sr близка к 0, что говорит о регулярном поступлении элемента в воду. Несмотря на то, что среднее содержание Zn в воде между реками отличается всего в 1.2 раза, медианы различаются в 7 раз (достоверность различия по тесту Манна-Уитни $p=0.004$). Содержание ВВ в р. Шешме в 6 раз выше по сравнению с р. Мёшей. В р. Шешме выше содержание нитратов, но ниже содержание нитритов и ионов аммония по сравнению с р. Мёшей, что свидетельствует об активно протекающих процессах нитрификации. Содержание НП не превышает установленные нормативы в воде обеих рек.

PERMANOVA

Пермутационный многомерный анализ дисперсии (PERMANOVA) выявил высокосignимые различия в гидрохимическом составе между исследуемыми реками ($F_{1,220} = 19.039$; $p = 0.001$). Фактор 1 "тип хозяйственной

деятельности в водосборе", объясняет 5.5% общей вариации (табл. 3), в то время как сезонность (фактор 2) влияет на 11.0% вариации. Данные факторы объясняют небольшую часть вариации, что обусловлено исключительной сложностью изучения природных вод, подверженных как влиянию множества локальных источников загрязнения, так и неучтенных глобальных факторов (поверхностный сток, изменение климата, аэрогенный перенос загрязняющих веществ (ЗВ) и т.д.). Ординационные графики (NMDS, PCoA), позволяющие по совпадению паттернов утверждать о реальности и устойчивости структуры данных, показывают ясное разделение облаков точек с минимальным перекрытием (рис. 3).

Таблица 3. Результаты анализа PERMANOVA

Table 3. Results of the PERMANOVA analysis

Источник вариации	R ²	Статистическая значимость	Интерпретация
Сезон	11.0%	$**p < 0.001$	Основной фактор – летние/зимние колебания
Река (Мёша / Шешма)	5.5%	$***p < 0.001$	Вторичный фактор – различие типа хозяйственной деятельности в водосборе
Остаточная вариация	81.3%	-	Неучтенные факторы

Две реки достоверно различаются по химическому составу на уровне значимости $p < 0.001$. Критерий F-статистика равный 15.34 указывает на сильное влияние на качество воды типа хозяйственной деятельности в бассейне реки. Значение $R^2 = 0.05$ (5.5%) указывает на то, что различие между Шешмой и Мёшей объясняют менее 6% общей вариабельности данных, но это не противоречит выводу о существенных различиях – в многомерных экологических данных высокая статистическая значимость может сопровождаться низким эффект-размером из-за большой внутригрупповой дисперсии.

Данные по р. Мёше на NMDS-графиках (рис. 3) представлены преимущественно компактным, плотным облаком точек, сгруппированным в центральной части ординационного пространства, что позволяет говорить о низкой внутригрупповой вариабельности, т.е. относительно постоянстве состава; низкой вероятности экстремальных значений показателей; ввиду низкой вариабельности и повторяемости параметров в разных пробах – о прогнозируемости качества воды. Все вышеперечисленное подтверждает обоснованность выбора р. Мёши в качестве реки сравнения в нашем исследовании.

Данные по р. Шешме представлены широким, рассеянным облаком точек с множественными выбросами, что указывает на высокую внутригрупповую вариабельность, нестабильность состава воды вследствие повышенного антропогенного стресса.

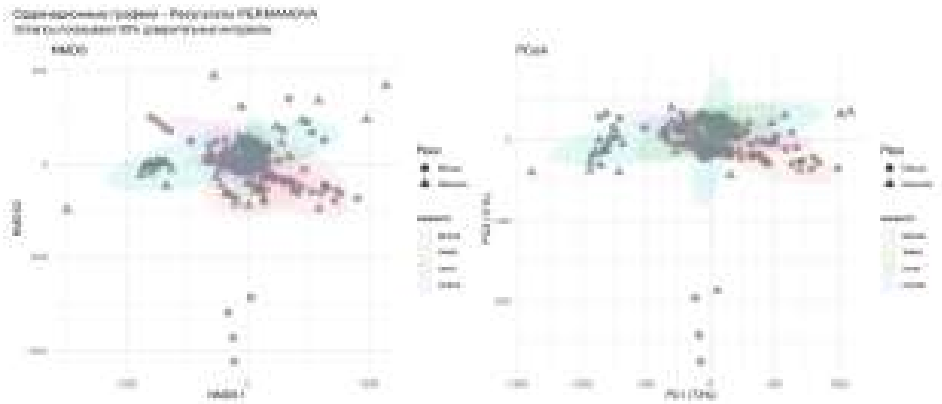


Рисунок 3. Ординационные графики

Рисунок включает в себя два метода ординации nMDS (неметрическое многомерное шкалирование) и PCoA (анализ главных координат)

Figure 3. Ordination plots

The figure includes two ordination methods: nMDS (non-metric multidimensional scaling) and PCoA (principal coordinates analysis)

Сезонность имеет более выраженный эффект и является основным фактором, структурирующим данные, объясняя 11.02% вариации (величина *F*-статистики – 10.25, $p = 0.001$). В р. Шешме по сравнению с р. Мёшей более отчетливо видна дифференциация по сезонам. Весенние и осенние пробы расположены ближе к центру; летние смещены в положительном направлении PC1/NMDS1 и характеризуются более высокой вариабельностью и высокими значениями показателей. Зимой наблюдаются минимальные значения показателей с низкой вариабельностью для обеих рек.

Для корректного сравнения качества воды между реками следует отбирать пробы в один и тот же сезон, поскольку сезонные вариации (11%) сопоставимы или превосходят различия по типу хозяйственной деятельности в бассейне реки (5.5%), а статистическая значимость различия двух рек высокая ($R^2 = 0.0765$ (7.65%), *F*-значение 19.04 и $p = 0.001$) и исключение сезонного фактора позволит более точно учитывать влияние типа хозяйственной деятельности в водосборе.

PCA (Анализ главных компонент)

На рис. 4, иллюстрирующем результаты PCA, показано, что данные концентрируются в группы с перекрытием менее 10%. Представленные компоненты *Dim 1* и *Dim 2* объясняют 54.7% различий, носят не случайный характер и указывают на наличие нескольких независимых источников загрязнения.

Первая компонента (*Dim 1*) объясняет главное различие между реками – общий уровень загрязнения. Существенный вклад в первую компоненту вносят такие показатели как Ba, (9.63%), NO₂ (7.34%), Zn (6.36%), Pb (6.04%), Cu (5.60%), Mo (5.22%). PC1 выступает как ось "антропогенной нагрузки". Вторая компонента (*Dim 2*) отражает вторичное различие и основной вклад вно-

сят такие показатели как Sr (11.27%), БПК₅ (10.73%), Cd (6.54%), As (5.48%), Se (6.04%), ХПК (5.54%).

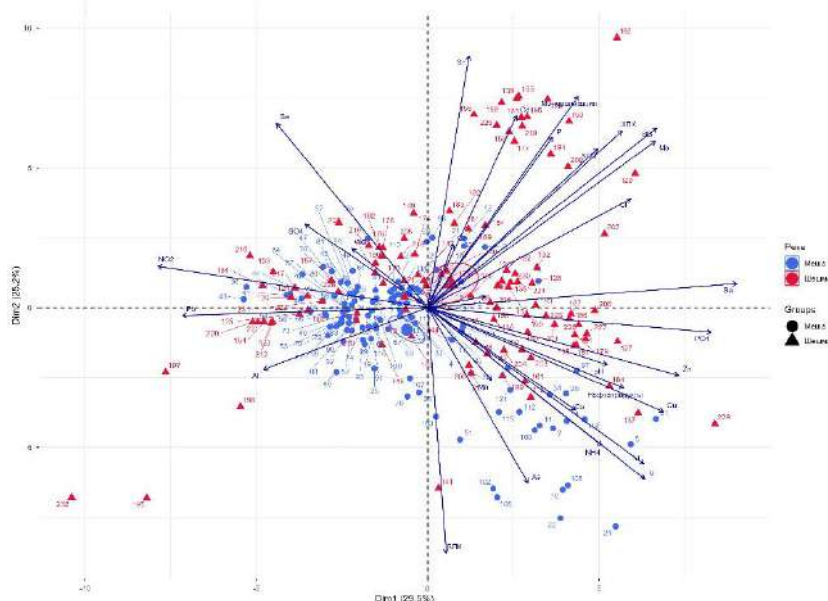


Рисунок 4. Анализ главных компонент (PCA)

Пробы, отобранные в р. Шешме, обозначены красными треугольниками, в р. Мёше – синими кругами; стрелками обозначены нагрузки переменных, показывающие как исходные параметры связаны с главными компонентами; обведенные группы – наиболее часто встречающиеся значения показателей

Figure 4. Principal component analysis (PCA)

Samples collected from the Sheshme River are marked with red triangles, from the Myosha River with blue circles; the arrows indicate the loadings of the variables, showing how the original parameters are related to the principal components; the circled groups represent the most frequently occurring values of the indicators

Доминирующим показателем с вкладом в PC1 в 9.63% и коэффициентом качества представления ($\cos^2$) равным 0.853, является Ba – классический маркер техногенного воздействия в регионах добычи углеводородов, активно используемый в буровых растворах (Ананьев и др., 1978; Wang et al., 2024), высвобождается в процессе окисления углеводородов микроорганизмами (Jones et al., 2023). Вклад нитритов составляет 7.34% в PC1 с отрицательной координатой, что является показателем сельскохозяйственной деятельности, характерной для водосбора обеих рек. Zn и Pb вносят в PC1 6.36% и 6.04% соответственно с высокими коэффициентами качества представления ($\cos^2 = 0.563$ и 0.535).

Sr и Cd вносят высокий вклад в PC2 (11.27% и 6.54% соответственно) с высоким коэффициентом качеством представления ($\cos^2$ составляет 0.851 и 0.494 соответственно), означая очень хорошую представленность этих показателей.

Профили загрязнения рек принципиально отличаются. Р. Мёша характеризуется более низкими и стабильными значениями анализируемых показателей. Преобладающий массив данных химического состава воды р. Мёши на рис. 2 расположен в 1 четверти с низкими отрицательными значениями Dim1

и Dim 2, что свидетельствует о более низком загрязнении проб в том числе Ba, Cu, Zn, Pb по сравнению с р. Шешмой.

В пробах воды обеих рек содержание НП не превышало ПДК (0.04 ± 0.005 и 0.03 ± 0.003 в Шешме и Мёше соответственно), что снижает эффективность использования их в качестве маркера нефтедобывающей деятельности.

Несмотря на сопоставимые значения БПК₅ между реками (разница не более 1%), в р. Шешме медианальное значение ХПК в 6 раз выше, что может свидетельствовать о наличии органических веществ, устойчивых к биохимическому окислению. Показатель ХПК вносит 5.54% в PC2 и 3.80% в PC1 с хорошим представлением на обеих осях ($\cos^2 = 0.418$ и 0.336). Вклад в PC2 БПК₅ составляет 10.73% (коэффициент представления $\cos^2 = 0.810$).

Кластерный анализ

Для углубленного анализа многомерных данных качества воды применялись два независимых метода кластеризации: иерархическая кластеризация (НС) методом Ward.D2 (рис. 5) и методом К-средних (K-means) (рис. 6). Графики (рис. 5 и 6) показывают распределение точек данных по двум главным компонентам PC1 и PC2, которые объясняют 29.5% и 5.2% общей дисперсии соответственно.

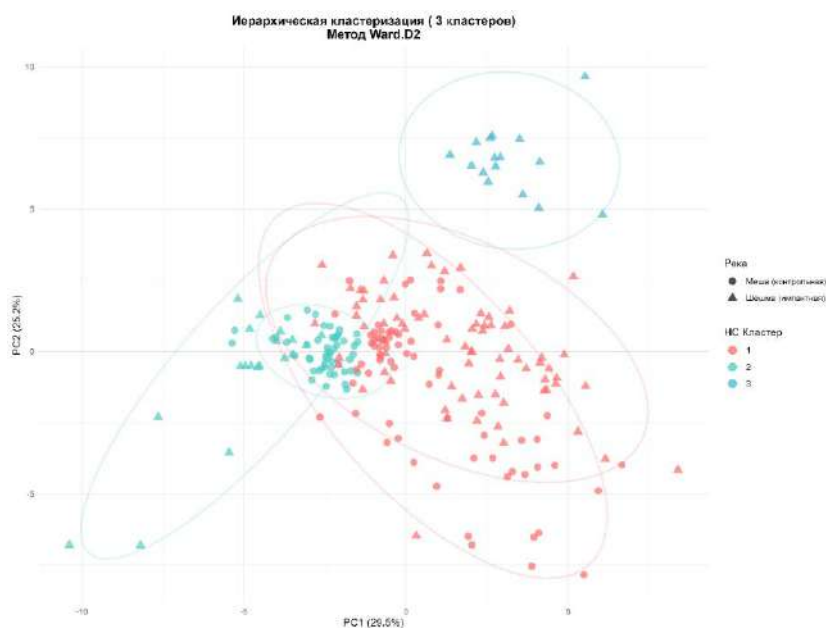


Рисунок 5. Иерархическая кластеризация методом Ward.D2 (НС)

Точки данных обозначены разными символами и цветами в зависимости от типа реки и кластера, к которому они принадлежат; контурные линии вокруг точек данных показывают границы кластеров

Figure 5. Hierarchical clustering (Ward.D2)

Data points are labeled with different symbols and colors depending on the river type and the cluster to which they belong; the contour lines around the data points indicate the cluster boundaries

Использование двух методов позволяет валидировать результаты и выявить прочные закономерности в структуре данных, которые воспроизводятся независимо от метода. На основании метода Elbow plot было определено, что оптимальное число кластеров составляет 3.

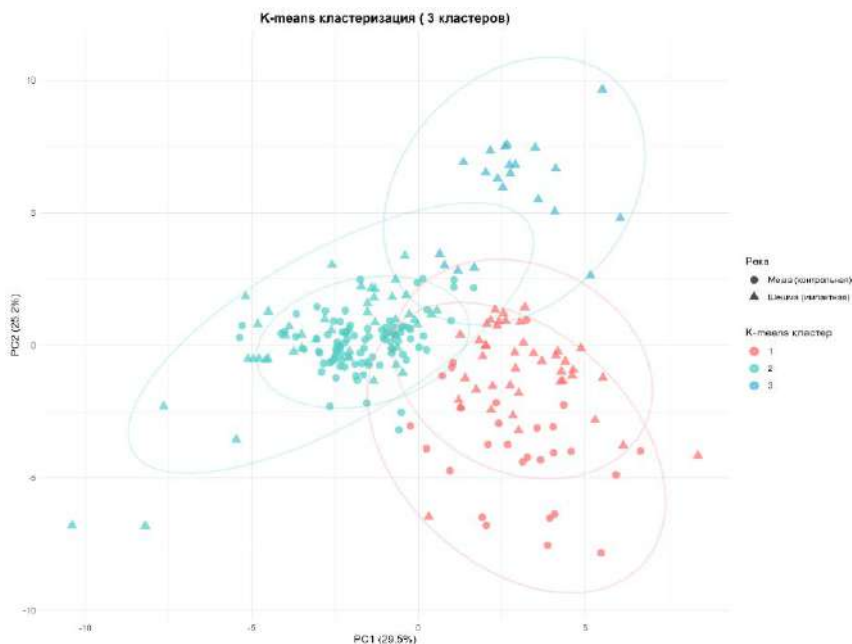


Рисунок 6. Кластеризация методом К-means с тремя кластерами

Точки данных обозначены разными символами и цветами в зависимости от типа реки и кластера, к которому они принадлежат; контурные линии вокруг точек данных показывают границы кластеров

Figure 6. K-means clustering with three clusters

Data points are identified by different symbols and colors depending on the river type and the cluster to which they belong; the contour lines around the data points indicate the cluster boundaries

Оба метода показали похожие результаты, выделив 3 кластера (рис. 7) учитывающие степень антропогенного воздействия. Первый метод (НС) показал себя более успешно, так как он выделил группу отдельно стоящих проб в новый кластер. Данные пробы по результатам PCA были наиболее загрязненными.

В первый кластер вошли пробы со средним уровнем загрязнения. В данном кластере преобладают данные р. Шешмы. Второй кластер представлен преимущественно пробами из р. Мёши с небольшой долей проб из менее загрязненных участков р. Шешмы. 3 кластер представлен исключительно пробами из р. Шешмы, характеризуются высоким уровнем загрязнения.

Для оценки согласованности методов была построена матрица ошибок (confusion matrix), которая показала, что оба метода достигают хорошей согласованности в определении кластера 3. Тем не менее, появление в 3 кластере дополнительных точек может быть обусловлено тем, что метод К-средних «не видит» структуру «экстремальных» загрязнений, что делает его использование при анализе экологических данных затруднительным.

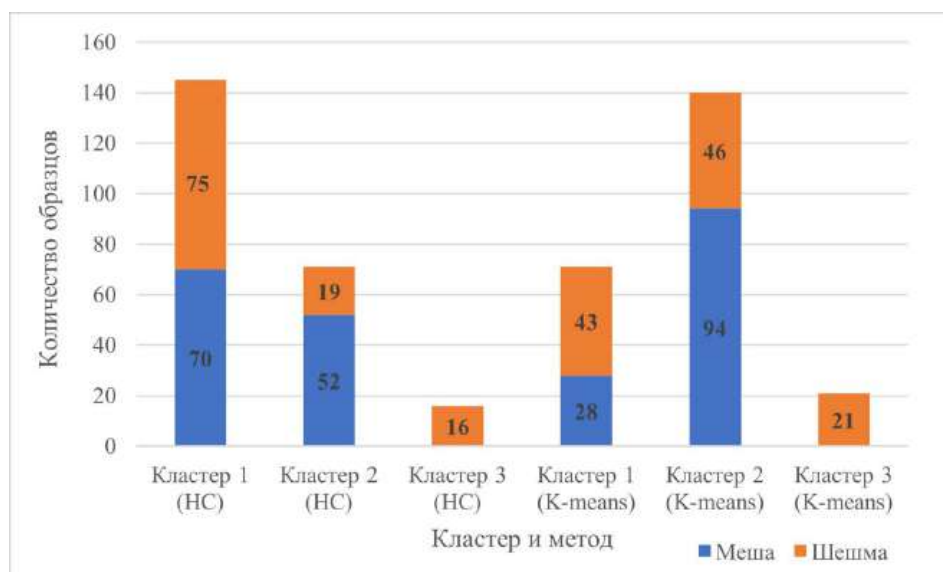


Рисунок 7. Сравнение результатов методами иерархической кластеризации (HC) и с использованием К-средних

Figure 7. Comparison of hierarchical clustering (HC) and K-means clustering results

Таким образом проведенный кластерный анализ позволяет выделить в реках 3 уровня загрязнения: условно-чистая вода (≈ 71 проба), со средним уровнем загрязнения (≈ 140 проб) и с высоким уровнем загрязнения (≈ 16 проб). Отсутствие в 3 кластере проб из р. Мёши позволяет сделать вывод, что он выделен не на основе природной вариабельности воды, а обусловлен исключительно антропогенным воздействием. Это подтверждается тем, что он состоит из проб, разошедшихся по осям РСА далеко от основного облака. Расхождение по матрице ошибок в 69 пробах между HC-2 и K-means-1 (в кластерах 1 и 2) и их смещение обусловлено размытостью границ между условно-чистой и загрязненной водой.

Река Мёша, несмотря на наличие диффузного загрязнения, может быть отнесена к условно-чистой и ее можно использовать в качестве реки-сравнения в проведенном исследовании. К основным комплексным показателям, характеризующим воду как условно чистую, могут быть отнесены такие показатели как ХПК, СПАВ, ВВ, а также содержание элементов – Cd, As и Sr. Также в проведенном РСА данные сконцентрированы более плотно, демонстрируя стабильное, репрезентативное состояние с низкой внутривидовой вариабельностью.

В р. Шешме только 26-30% проб можно отнести к условно чистым. Река демонстрирует высокую вариабельность, что обусловлено антропогенным воздействием. Основным фактором для выделения 3 кластера послужили концентрации таких элементов как Ba ($\cos^2 = 0.853$) и Sr ($\cos^2 = 0.851$), уровень БПК₅ ($\cos^2 = 0.810$), а также Zn ($\cos^2 = 0.563$), Pb ($\cos^2 = 0.535$) и минерализация ($\cos^2 = 0.600$).

Проведена валидация результатов кластеризации гидрохимических данных рек Мёша (контрольная) и Шешма (исследуемая) с использованием четырех независимых метрик качества: Silhouette Score, Davies-Bouldin Index, Calinski-Harabasz Index и двумерного дисперсионный анализ ANOVA с post-hoc тестом Tukey HSD. Результаты показывают высокую статистическую значимость и хорошую качественную валидацию найденной структуры данных. Двумерный ANOVA показал статистически значимые различия между кластерами по PC1 ($F_{2,229} = 78.04, p < 0.001$) и PC2 ($F_{2,229} = 117.00, p < 0.001$). Post-hoc тест Tukey HSD подтвердил, что все три кластера попарно различаются ($p < 0.01$), при этом границы между условно-чистой и загрязненной водой остаются размытыми ($p = 0.06$).

Выделенная трехуровневая структура загрязнения (низкое → среднее → высокое) является статистически значимой, структурно обоснованной и методологически надежной.

Выделение маркеров

Загрязнение воды в результате нефтедобывающей деятельности объясняет только 5.5% общей многомерной вариации, это указывает на высокое содержание отдельных компонентов и их уровень превышает содержание аналогичных показателей в контрольной реке.

Гидрохимические данные по р. Мёше статистически более однородны (узкое облако NMDS), а вариабельность обусловлена, главным образом, сезонностью и пространственными различиями. К условно чистым относится 43% проб, к доминирующим факторам вариации относятся естественные геохимические процессы.

Гидрохимические данные р. Шешма, напротив, характеризуется высокой вариабельностью с выраженным "хвостом" экстремальных значений. К сильно загрязненным относится 14.5% всех проб, демонстрируя наличие точечных источников загрязнения. Вариабельность обусловлена взаимодействием естественных процессов формирования качества воды и антропогенным загрязнением.

На основе комплексного анализа данных были идентифицированы пять специфических маркеров нефтедобывающей деятельности (табл. 4), которые могут быть рекомендованы для улучшения действующей системы экологического мониторинга.

В качестве дополнительных маркеров можно использовать ВВ, медианальное значение которых примерно в 14 раз выше в р. Шешме, а также Ва ($\cos^2=0.853$), одного из наиболее доминирующих показателей в PC1. Комбинированное применение маркеров позволяет уверенно идентифицировать влияние нефтедобывающей деятельности на качество поверхностных вод и дифференцировать антропогенное загрязнения от геохимического фона.

При организации мониторинга следует не только отслеживать такой параметр как превышение ПДК, но и анализировать превышение медианальных значений контрольных рек в 2-3 раза.

Таблица 4. Индикаторы нефтедобывающей деятельности

Table 4. Indicators of oil production activity

Показатель	Содержание в воде		Отношение содержания в воде	Интерпретация
	Мёша	Шешма		
Хлориды мг·л ⁻¹	11.03	68.77	6.2	Входит в состав пластовых вод (Шипаева, 2025)
Кадмий мкг·л ⁻¹	0.017	1.29	76.0	Входит в состав сырой нефти (Osuji, Onojake, 2004)
Свинец мкг·л ⁻¹	0.35	6.11	17.5	Входит в состав сырой нефти (Osuji, Onojake, 2004)
Молибден мкг·л ⁻¹	3.76	4.91	1.3	Присутствует в пластовых водах; вклад в PC1 = 5.22% (Мударисова, Пунанова, 2024; Скибицкая и др., 2019)
Стронций мкг·л ⁻¹	1032.16	1267.01	1.2	Входит в состав нефтей и пластовых вод; $\cos^2 = 0.851$; асимметрия ≈ 0 (Ибрагимов, 2011; Шипаева, 2025)

Модели прогноза

Разработанная логистическая регрессионная модель для бинарной классификации («чистая вода» / «загрязненная вода»), описываемая уравнением (1), показала высокую дискриминирующую способность в классификации данных по степени техногенного воздействия ($AUC = 0.9779$, доверительный интервал 95%: 0.958–0.99).

$$P(\text{загрязнение}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 Cl + \beta_2 Mo + \beta_3 Sr)}} \quad , \quad (1)$$

где:

P – вероятность загрязнения (от 0 до 1). $P > 0.5$ = "загрязнено"

$\beta_0 = -10.299$, $\beta_1 = 0.1359$, $\beta_2 = 2.0947$, $\beta_3 = -0.0028$ – коэффициенты регрессии (оцениваются методом максимального правдоподобия); $\text{logit}(P) = \ln(P/(1-P))$ – логарифм отношения шансов (от $-\infty$ до $+\infty$). Преобразует вероятность в линейную шкалу.

Модель достигла общей точности классификации 94.98% (78 из 83 проб), чувствительности 95.0% (38 из 40 загрязненных образцов) и специфичности 93.02% (40 из 42 чистых проб). Статистически значимыми предикторами загрязнения являются хлориды и Мо. Увеличение концентрации хлоридов на 1 мг/л ассоциировалось с увеличением логит-отношения шансов

загрязнения на 0.136 единиц ($z = 3.59, p < 0.001$), что соответствует увеличению шансов загрязнения на 14.6% ($OR = 1.15$; 95% CI : 1.07–1.22). Мо показал значительно более сильный эффект: увеличение концентрации на 1 $\text{мкг}\cdot\text{л}^{-1}$ увеличивало шансы загрязнения в 8.12 раз ($\beta = 2.095, OR = 8.12$; 95% CI : 1.82–36.32, $z = 2.73, p = 0.006$). Стронций не показал статистически значимого влияния на риск загрязнения ($\beta = -0.003, p = 0.085$).

Модель дерева решений показала еще более высокие показатели классификации, достигнув точности 98.80% (82 из 83 проб), чувствительности 100% (41 из 41 загрязненных проб) и специфичности 97.67% (41 из 42 чистых проб). Модель идеально определила все загрязненные образцы без ложных отрицательных результатов, допустив только 1 ложное положительное значение (рис. 8).

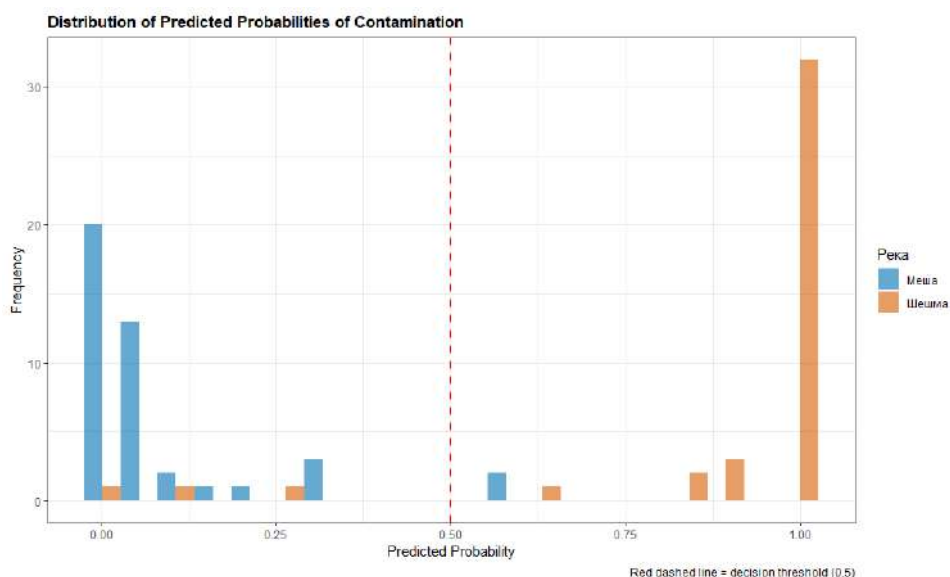


Рисунок 8. Распределение предсказанных вероятностей загрязнения
Ось X показывает предсказанную вероятность загрязнения P ($0 =$ чистая, $1 =$ загрязненная), ось Y – частоту (количество проб). Красная пунктирная линия отмечает порог равный 0.5, который разделяет пробы на «чистые» и «загрязненные»

Figure 8. Distribution of predicted contamination probabilities
The x -axis shows the predicted probability of contamination P ($0 =$ clean, $1 =$ contaminated), and the y -axis shows the frequency (number of samples). The red dotted line marks the threshold of 0.5, which separates the samples into “clean” and “contaminated”

Ранговый анализ важности переменных показал доминирующую роль хлоридов в классификации загрязнения (важность = 39.54, что составляет 47.6% от общей информационной ценности), за которыми следуют Мо (14.78, 17.8%) и Sr (12.94, 15.6%). Основное правило разделения в корневом узле дерева: концентрация хлоридов $< 30.1 \text{ мг л}^{-1}$ классифицирует воду как чистую (класс 0, р. Мёша), концентрация $\geq 30.1 \text{ мг л}^{-1}$ – как загрязненную (класс 1, р. Шешма). Это пороговое значение хорошо согласуется с данными по р. Мёше, где максимальное значение содержания хлоридов составляет 26.9 мг л^{-1} .

Доминирующая роль хлоридов как индикатора может быть объяснена несколькими факторам: 1. хлориды ассоциированы с нефтедобывающей деятельностью; 2. хлориды отличаются консервативным поведением в речных экосистемах (слабо участвуют в химических и биологических процессах); 3. концентрация хлоридов в р. Мёше (11.26 ± 0.36 мг л⁻¹) существенно отличается от значения в р. Шеше (63.45 ± 2.51 мг л⁻¹), что обеспечивает четкое разделение классов в моделях.

Mo, демонстрирующий второе место по важности (14.78) и наибольшее отношение шансов (*OR*) равное 8.12, также является значимым индикатором антропогенного воздействия со стороны нефтедобывающей деятельности. *Sr* не показал статистически значимого влияния ($p=0.085$), но его присутствие в модели не ухудшает её производительность и может быть полезным для контроля геохимических факторов.

Выводы

Комплексный подход, объединяющий анализ главных компонент (PCA), пермутационный многомерный анализ дисперсии (PERMANOVA) и классификационные модели (логистическая регрессия и дерево решений), продемонстрировал высокую эффективность для идентификации гидрохимических маркеров нефтяного загрязнения водных объектов.

К основным преимуществам используемой методологии можно отнести статистическую робастность (независимые методы взаимно подтверждают различие между исследуемым объектом и объектом сравнения), выделение специфических маркеров и дифференциацию источников нагрузки с использованием кластеризации (НС, K-means) и ANOVA, позволивших разделить пробы по уровню загрязнения, исключив природную вариабельность. Предлагаемая методология может быть адаптирована для других нефтедобывающих регионов РФ. Результаты исследования могут быть использованы в процедурах установления ответственности за накопленный экологический ущерб.

По результатам исследования рекомендовано включить в систему производственного мониторинга нефтегазодобывающих управлений специфические индикаторы: триаду маркеров (хлориды-Mo-Sr), используемую в регрессионной модели, а также Cd и Pb. При принятии управленческих решений предлагается руководствоваться не только ПДК, но и превышением медианных значений контрольных рек более 2 раз.

Разработанные классификационные модели (логистическая регрессия и дерево решений) на основе трех гидрохимических показателей (хлориды, Mo, Sr) обеспечивают эффективную диагностику техногенного загрязнения вод рек региона с точностью 94-99%.

Работа выполнена за счет средств субсидии, выделенной Казанскому федеральному университету для выполнения государственного задания в сфере научной деятельности, проект № FZSM-2024-0004.

Список литературы

Ананьев, А.Н., Комяков, Ю.А., Литвишко В.Г. (1978) *Утяжелитель для буровых растворов на нефтяной основе*, Авторское свидетельство № 601299 А1 СССР, МПК C09K 8/32, № 2121277.

Водные объекты Республики Татарстан. Гидрографический справочник (2018) Казань, изд-во «Фолиант», 512 с.

Гарелина, С.А. (2025) Комплексное решение для защиты береговой полосы Черного моря от разливов нефти и нефтепродуктов (мазута), *Научные и образовательные проблемы гражданской защиты*, № 3(66), с. 22-31.

Ермолаев, О.П., Игонин, М.Е., Бубнов, А.Ю., Павлова, С.В. (2007) *Ландшафты Республики Татарстан. Региональный ландшафтно-экологический анализ*, Казань, Слово, 441 с.

Ибрагимов, Р.Л. (2011) Оценка возможности использования подземных вод нефтяных месторождений Республики Татарстан в качестве гидроминерального сырья, *Георесурсы, геознергетика, геополитика*, № 2(4), URL: https://oilgasjournal.ru/vol_4/ibragimov.pdf (дата обращения: 22.02.2026).

Кленкин, А.А., Агапов, С.А. (2011) Динамика распределения нефтепродуктов в воде и донных отложениях Азовского и Черного морей после аварии судов в Керченском проливе, *Водные ресурсы*, т. 38, № 2, с. 214-222.

Лешкович, Н.М., Арнбрехт, А.Э., Викулов, Г.Е. (2020) Воздействие нефтегазового комплекса на окружающую среду, *Булатовские чтения*, т. 5. с. 138-141.

Мударисова, Р.А., Пунанова, С.А. (2024) Особенности образования залежи и состава сверхвязкой нефти Горского месторождения Республики Татарстан, *Экспозиция Нефть Газ*, № 8, с. 28-33, doi: 10.24412/2076-6785-2024-8-28-33.

Рахматуллин, Н.Р., Сулейманов, Р.А., Валеев, Т.К. (2023) Пути совершенствования системы производственного и природоохранного контроля при разработке и эксплуатации нефтегазовых месторождений морского шельфа, *Тенденции развития науки и образования*, № 98-7, с. 129-133, doi: 10.18411/tmio-06-2023-390.

РД 52.24.353-12 (2012) *Отбор проб поверхностных вод суши и очищенных сточных вод*, Ростов на Дону, 40 с.

Скибицкая, Н.А., Бурханова, И.О., Большаков, М.Н., Навроцкий, О.К. (2019) Перспективность нефтегазоконденсатных и газоконденсатных месторождений как источников полиметаллического сырья, *Актуальные проблемы нефти и газа*, 1(24), URL: <https://doi.org/10.29222/ipng.2078-5712.2019-24.art6> (дата обращения: 22.02.2026).

Титов, В.Н., Ходов, Д.А. (2015) Основные экологические проблемы нефтяного комплекса Саратовской области, *Защита окружающей среды в нефтегазовом комплексе*, № 3, с. 35-40.

Шипаева, М.С. (2025) *Геохимический мониторинг подземных вод нефтяных месторождений для повышения эффективности их разработки (на примере Республики Татарстан)*, дисс. ... канд. геол.-мин. наук, Казань, 147 с.

Bakke, T., Klungsoyr, J., Sanni, S. (2013) Environmental impacts of produced water and drilling waste discharges from the Norwegian offshore petroleum industry, *Marine Environmental Research*, vol. 92, pp. 154-169.

Jones, K.L., Ziegler, B.A., Davis, A.M., Cozzarelli, I.M. (2023) Attenuation of Barium, Strontium, Cobalt, and Nickel Plumes Formed during Microbial Iron Reduction in a Crude-Oil-Contaminated Aquifer ACS, *Earth and Space Chemistry*, vol. 7(7), pp. 1322-1336, doi: 10.1021/acsearthspacechem.2c00387.

Osuji, LC, Onojake, CM. (2004) Trace heavy metals associated with crude oil: a case study of Ebocha-8 oil-spill-polluted site in Niger Delta, Nigeria, *Chem Biodivers*, vol. 1(11), pp. 1708-1715, doi: 10.1002/cbdv.200490129. PMID: 17191811.

Wang, F., Zhao, T., Yang, J.Z., et al. (2024) Research on the identification method of barium toxic substance content of polysulfonate cuttings and reducing soils, *Journal of Environmental Engineering Technology*, vol. 14(6), pp. 1969-1976, doi: 10.12153/j.issn.1674-991X.20240266.

Yusta-García, R., Orta-Martínez, M., Mayor, P., González-Crespo, C., Rosell-Melé A. (2017) Water contamination from oil extraction activities in Northern Peruvian Amazonian rivers, *Environmental Pollution*, vol. 225, pp. 370-380.

References

Ananyev, A.N., Komyakov, Yu.A., Litvishko, V.G. (1978) *Utyazhelitel' dlya burovyykh rastvorov na neftyanoy osnove* [Weighting agent for oil-based drilling fluids], USSR, Pat. 601299 A1, IPC C09K 8/32, no. 2121277.

Vodnyye ob'yekty Respubliki Tatarstan: Gidrogeograficheskiy spravochnik [Water bodies of the Republic of Tatarstan: Hydrographic reference book] (2018) Foliant, Kazan, Russia, 512 p.

Garelina, S.A. (2025) *Kompleksnoye resheniye dlya zashchity beregovoy polosy Chernogo morya ot razlivov nefi i nefteproduktov (mazuta)* [Comprehensive solution for coastal protection the Black sea from oil spills and petroleum products (fuel oil)], *Nauchnyye i obrazovatel'nyye problemy grazhdanskoy zashchity*, no. 3(66), pp. 22-31.

Yermolaev, O.P., Igonin, M.E., Bubnov, A.Yu., Pavlova, S.V. (2007) *Landshafty Respubliki Tatarstan. Regional'nyy landshaftno-ekologicheskii analiz* [Landscapes of the Republic of Tatarstan. Regional landscape-ecological analysis], Slovo, Kazan, Russia, 441 p.

Ibragimov, R.L. (2011) *Otsenka vozmozhnosti ispol'zovaniya podzemnykh vod neftyanykh mestorozhdeniy Respubliki Tatarstan v kachestve gidromineral'nogo*

syr'ya [Assessment of the possibility of using groundwater from oil fields of the Republic of Tatarstan as hydromineral raw materials], *Actual Problems of Oil and Gas*, no. 2(4), available at: https://oilgasjournal.ru/vol_4/ibragimov.pdf (accessed 22 Feb 2026).

Klenkin, A.A., Agapov, S.A. (2011) Dinamika raspredeleniya nefteproduktov v vode i donnykh otlozheniyakh Azovskogo i Chernogo morey posle avarii sudov v Kerchenskom prolive [Dynamics of oil product distributions in water and bottom sediments of the Sea of Azov and the Black Sea after ship accidents in the Kerch Strait], *Water Resources*, vol. 38, no. 2, pp. 214-222.

Leshkovich, N.M., Arnbreht, A.E., Vikulov, G.E. (2020) Vozdeystviye neftegazovogo kompleksa na okruzhayushchuyu sredu [Environmental impact of the oil and gas complex], *Bulatovskiy chteniya*, vol. 5, pp. 138-141.

Mudarisova, R.A., Punanova, S.A. (2024) Osobennosti obrazovaniya zalezhi i sostava sverkhvyazkoy nefti Gorskogo mestorozhdeniya Respubliki Tatarstan [Peculiarities of the formation of the deposit and composition of ultraviscous oil Gorsky field of the Republic of Tatarstan], *Exposition Oil Gas*, issue 8, pp. 28-33, doi: 10.24412/2076-6785-2024-8-28-33.

Rakhmatullin, N.R., Suleymanov, R.A., Valeyev, T.K. (2023) Puti sovershenstvovaniya sistemy proizvodstvennogo i prirodookhrannogo kontrolya pri razrabotke i ekspluatatsii neftegazovykh mestorozhdeniy morskogo shel'fa [Ways to improve the system of industrial and environmental control in the development and operation of offshore oil and gas fields], *Tendentsii razvitiya nauki i obrazovaniya*, vol. 98, issue 7, pp. 129-133, doi: 10.18411/trnio-06-2023-390.

RD 52.24.353-12 (2012) *Otbor prob poverkhnostnykh vod sushi i ochishchennykh stochnykh vod* [Sampling of inland surface waters and treated wastewater], Rostov-on-Don, Russia, 40 p.

Skibitskaya, N.A., Burkhanova, I.O., Bol'shakov, M.N., Navrotskiy, O.K. (2019) Perspektivnost' neftegazokondensatnykh i gazokondensatnykh mestorozhdeniy kak istochnikov polimetallicheskogo syr'ya [Prospectivity of oil-gas-condensate and gas-condensate fields as polymetallic feedstock sources], *Actual Problems of Oil and Gas*, issue. 1(24), available at: <https://doi.org/10.29222/ipng.2078-5712.2019-24.art6> (accessed 22 Feb 2026).

Titov, V.N., Khodov, D.A. (2015) Osnovnyye ekologicheskiye problemy neftyanogo kompleksa Saratovskoy oblasti [The main environmental problems of the oil industry in Saratov region], *Zashchita okruzhayushchey sredy v neftegazovom komplekse*, issue 3, pp. 35-40.

Shipaeva, M.S. (2025) *Geokhimicheskii monitoring podzemnykh vod neftnykh mestorozhdeniy dlya povysheniya effektivnosti ikh razrabotki (na primere Respubliki Tatarstan)* [Geochemical monitoring of groundwater in oil fields to improve the efficiency of their development (case study of the Republic of Tatarstan)], Candidate's thesis, Kazan Federal University, Kazan, Russia, 147 p.

Bakke, T., Klungsøyr, J., Sanni, S. (2013) Environmental impacts of produced water and drilling waste discharges from the Norwegian offshore petroleum industry, *Marine Environmental Research*, vol. 92, pp. 154-169.

Jones, K.L., Ziegler, B.A., Davis, A.M., Cozzarelli, I.M. (2023) Attenuation of Barium, Strontium, Cobalt, and Nickel Plumes Formed during Microbial Iron Reduction in a Crude-Oil-Contaminated Aquifer ACS, *Earth and Space Chemistry*, vol. 7(7), pp. 1322-1336, doi: 10.1021/acsearthspacechem.2c00387.

Osuji, LC, Onojake, CM. (2004) Trace heavy metals associated with crude oil: a case study of Ebocha-8 oil-spill-polluted site in Niger Delta, Nigeria, *Chem Biodivers*, vol. 1(11), pp. 1708-1715, doi: 10.1002/cbdv.200490129. PMID: 17191811.

Wang, F., Zhao, T., Yang, J.Z. (2024) Research on the identification method of barium toxic substance content of polysulfonate cuttings and reducing soils, *Journal of Environmental Engineering Technology*, vol. 14(6), pp. 1969-1976, doi: 10.12153/j.issn.1674-991X.20240266.

Yusta-García, R., Orta-Martínez, M., Mayor, P., González-Crespo, C., Rosell-Melé A. (2017) Water contamination from oil extraction activities in Northern Peruvian Amazonian rivers, *Environmental Pollution*, vol. 225, pp. 370-380.

Статья поступила в редакцию (Received): 02.03.2026.

Статья доработана после рецензирования (Revised): 18.03.2026.

Для цитирования / For citation:

Новикова, Л.В., Степанова, Н.Ю. (2026) Использование многомерного анализа для выделения гидрохимических маркеров антропогенного воздействия, *Экологический мониторинг и моделирование экосистем*, т. XXXVII, № 1-2, с. 133-155, doi: 10.24412/2782-3237-2026-1-2-133-155.

Novikova, L.V., Stepanova N.Yu. (2026) Using multidimensional analysis to identify water pollution markers, *Environmental Monitoring and Ecosystem Modelling*, vol. XXXVII, no. 3-4, pp. 133-155, doi: 10.24412/2782-3237-2026-1-2-133-155.