УДК 504.453

М. А. Смирнов, С. С. Кладова, Ю. В. Королёва

ИСПОЛЬЗОВАНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ В АНАЛИЗЕ И ПРОГНОЗИРОВАНИИ ПОКАЗАТЕЛЕЙ АЗОТНО-ФОСФОРНОГО ЦИКЛА В БАССЕЙНЕ РЕКИ ПРЕГОЛИ

Балтийский федеральный университет им. И. Канта, Калининград, Россия
Поступила в редакцию 08.09.2025 г.
Принята к публикации 25.11.2025 г.
doi: 10.5922/vestniknat-2026-1-3

Для цитирования: *Смирнов М. А., Кладова С. С., Королёва Ю. В.* Использование математических методов в анализе и прогнозировании показателей азотно-фосфорного цикла в бассейне реки Преголи // Вестник Балтийского федерального университета им. И. Канта. Сер.: Естественные науки. 2026. №1. С. 41—53. doi: 10.5922/vestniknat-2026-1-3.

Для анализа многомерных необработанных данных состояния водных экосистем зачастую требуется применение современных аналитических инструментов, способных выявлять скрытые закономерности. Цель настоящего исследования — анализ многолетних данных мониторинга качества воды р. Преголи (г. Калининград) для выявления основных факторов, определяющих изменчивость гидрохимического режима, и оценки их временной динамики. Для анализа данных по показателям азотного и фосфорного циклов были применены методы многомерной статистики: иерархический кластерный анализ (метод Уорда), факторный анализ (метод минимальных остатков с вращением «varimax») и анализ временных рядов (LOESS-сглаживание). В результате кластерного анализа было выделено три группы наблюдений, отражающих различные периоды поступления биогенных элементов. Факторный анализ позволил идентифицировать два латентных фактора: азотную нагрузку от органических источников и соотношение органических и минеральных форм азота. Анализ временных рядов факторных оценок выявил нелинейную динамику и долгосрочные тренды в изменении качества воды. Продемонстрировано, что комплексное применение хемометрических методов является эффективным инструментом для диагностики состояния городских водных объектов, идентификации источников загрязнения и обеспечивает научную основу для разработки управленческих решений в области охраны водных ресурсов.

Ключевые слова: многомерный статистический анализ, факторный анализ, кластерный анализ, качество воды, соединения азота и фосфора

© Смирнов М. А., Кладова С. С., Королёва Ю. В., 2026

Введение

В условиях интенсивного антропогенного воздействия на окружающую среду водные экосистемы подвергаются значительной нагрузке, что приводит к ухудшению их экологического состояния. Вода является основным ресурсом, необходимым для существования всех живых организмов. На качество природных вод влияют различные факторы, включая атмосферные условия, геологические особенности и растительность, но в городской среде доминирующим становится антропогенный фактор. В связи с этим особую актуальность приобретает разработка эффективных методов мониторинга, анализа и прогнозирования состояния водных объектов.

Традиционно для оценки качества природных вод используются классификации, основанные на нормативно-правовых документах. Водный кодекс Российской Федерации устанавливает критерии качества водных объектов по физико-химическим и биологическим показателям, определяя соответствие различным категориям водопользования. Классический подход к оценке качества воды также включает использование комплексных индексов, таких как индекс загрязненности воды (далее — ИЗВ), который позволяет получить интегральную оценку состояния водного объекта на основе нескольких ключевых показателей. Данные подходы, хотя и обеспечивают стандартизированную оценку, имеют ограничения в выявлении скрытых взаимосвязей между параметрами качества воды [1].

Альтернативным подходом является применение хемометрических методов, которые позволяют извлекать скрытую информацию из многомерных данных мониторинга. В мировой практике успешно применяются исследования с использованием кластерного анализа и анализа главных компонент для характеризации подземных вод различных регионов, где выявляются группы наблюдений с разной степенью минерализации и источниками формирования химического состава. Анализ поверхностных вод урбанизированных территорий с применением факторного анализа позволяет идентифицировать основные параметры, ответственные за вариабельность качества воды, и сократить количество показателей, необходимых для эффективного мониторинга.

Оценка качества воды необходима не только для удовлетворения текущих потребностей, но и для прогнозирования состояния водных объектов в будущем. Такой прогноз позволяет предотвращать превышение ассимиляционной емкости водных экосистем, а регулярный мониторинг способствует выявлению закономерностей в изменениях их состояния. Традиционные подходы, основанные на анализе отдельных параметров, зачастую не позволяют выявить скрытые взаимосвязи между переменными, характеризующими качество воды. Следовательно, применение современных методов многомерного статистического анализа становится необходимым этапом в комплексной оценке состояния водных объектов [2].

Задача анализа и прогнозирования качества природных вод часто осложняется спецификой данных мониторинга, которые характеризуются наличием пропусков, высокой размерностью и сложными нелиней-

ными взаимосвязями между гидрохимическими, гидрологическими и другими показателями. Оценка качества воды требует мониторинга широкого спектра физических, химических и биологических параметров, что приводит к формированию сложных матриц данных для анализа качества воды [3; 4]. Долгосрочные программы мониторинга качества воды генерируют большие массивы данных, которые часто трудно интерпретировать. Для автоматизации этого процесса и поиска скрытых взаимосвязей между показателями успешно применяются современные статистические подходы и методы. Среди них выделяются кластерный анализ (далее — CA), анализ главных компонент (далее — PCA), факторный анализ (далее — FA) и метод множественной импутации (далее — MI) [5—7].

CA представляет собой метод классификации без предварительных гипотез о распределении данных, позволяющий группировать данные мониторинга на основе их сходства. Хемометрические методы идентифицируют естественные кластерные структуры и группируют переменные на основе сходства между образцами [8; 9]. Существуют два основных подхода к CA: иерархический и неиерархический. В иерархическом подходе анализ начинается с того, что каждый объект представляется отдельным кластером, затем объекты последовательно объединяются на основе их сходства, пока все они не будут объединены в один кластер. Алгоритм кластеризации базируется на вычислении матрицы расстояний между объектами с использованием различных метрик расстояния [10].

PCA является одним из наиболее эффективных методов снижения размерности данных и считается методом без каких-либо предположений, что делает его более надежным по сравнению с FA при отклонениях от нормального распределения. Математически PCA основан на сингулярном разложении матрицы данных или на собственном разложении ковариационной матрицы. Метод основан на ортогональном преобразовании исходных, возможно, коррелирующих, переменных в набор линейно некоррелирующих переменных, называемых главными компонентами. Число главных компонент всегда меньше или равно числу исходных переменных.

Такой подход способствует оперативному и точному диагностированию экологической ситуации [11], что особенно важно в условиях быстрого изменения среды под влиянием как антропогенных, так и природных факторов.

Настоящее исследование посвящено анализу многолетних данных, собранных в рамках мониторинга состояния воды на р. Преголе, расположенной в черте г. Калининграда. Пробоотбор проводился по 21 гидрохимическому показателю, включая параметры азотного и фосфорного цикла, концентрации тяжелых металлов и общие физико-химические характеристики.

Материалы и методы

Отбор проб проводился в районе морского торгового порта ежемесячно с марта 2013 г. по декабрь 2024 г. (рис. 1) по 21 показателю: азотного и фосфорного циклов, металлам, физическим параметрам: температу-

ра (°С), pH, ион аммония (мг/дм^3), нитрит-ионы (мг/дм^3), нитрат-ионы (мг/дм^3), фосфат-ионы, железо общее (мг/дм^3), марганец (мг/дм^3), общий фосфор (мг/дм^3), нефтепродукты (мг/дм^3), кислород растворенный, взвешенные вещества, хлориды (мг/дм^3), БПК 5 (мг/дм^3), азот общий (мг/дм^3), медь (мг/дм^3), цинк (мг/дм^3), никель (мг/дм^3), фенолы (мг/дм^3), сульфат-ионы (мг/дм^3), формальдегид (мг/дм^3).

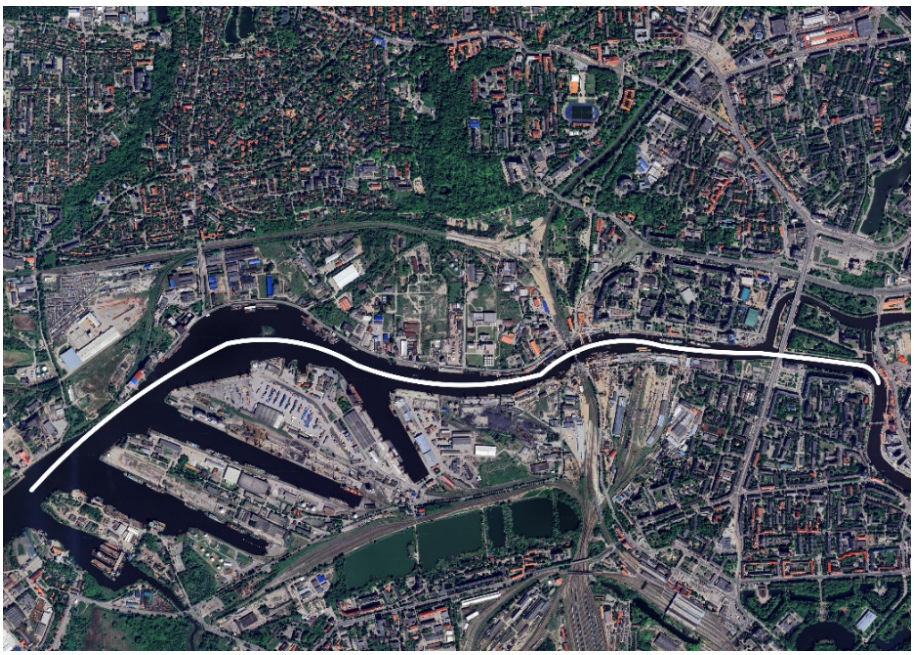

Рис. 1. Картосхема района исследования

Из общего массива данных были выделены показатели азотного и фосфорного циклов как наиболее информативные для характеристики экологического состояния водной среды. Анализ проводился по следующим показателям: ион аммония, нитрит-ионы, нитрат-ионы, фосфат-ионы, общий фосфор и азот общий. Выбор данных показателей обусловлен их ключевой ролью в оценке эвтрофикации водных объектов и определении основных источников загрязнения. Показатели азотного и фосфорного циклов являются лимитирующими биогенными элементами в пресноводных экосистемах и служат индикаторами антропогенного воздействия на водные объекты.

Для прогнозирования данных выбран промежуток с января по декабрь 2023 г. с целью использования в модели как прошлых, так и будущих данных и сравнения прогнозируемых данных с реальными.

Применены следующие методы статистического анализа:

— иерархический CA методом Уорда с евклидовым расстоянием для выявления групп наблюдений со схожими характеристиками;

— FA методом минимальных остатков с вращением «varimax» для снижения размерности данных и выявления латентных факторов;

— анализ временных рядов с применением LOESS-сглаживания для исследования динамики факторов во времени;

— MI для прогнозирования значений гидрохимических показателей.

Статистический анализ выполнен в программной среде R версии 4.3.0 с использованием пакетов cluster, psych, forecast, mice, ggplot2.

FA тесно связан с PCA, но имеет свои концептуальные особенности и математические преимущества в контексте обнаружения латентных структур. Он направлен на выявление скрытых (латентных) переменных, которые объясняют взаимосвязи между наблюдаемыми параметрами и имеют определенный физический или химический смысл. Модель FA предполагает, что наблюдаемые переменные являются линейными комбинациями небольшого числа общих и специфических факторов [6; 7; 12]. Общие факторы объясняют корреляции между переменными, а специфические — отвечают за остаточную дисперсию, включая ошибку измерения и уникальную вариабельность каждой переменной. В математическом выражении модель FA представляется как $X = LFм + \varepsilon$, где X — матрица наблюдаемых переменных (исходные гидрохимические показатели), L — матрица факторных нагрузок, F — матрица общих факторов, ε — вектор остатков, включающий специфические факторы и ошибки измерения [13].

Метод MI представляет собой современный статистический подход к определению пропущенных данных и прогнозированию будущих значений, основанный на байесовской статистике. Хотя классические методы PCA и FA широко используются для анализа качества воды, они имеют существенные ограничения при работе с неполными данными, которые часто встречаются в долгосрочных исследованиях мониторинга. Процесс MI состоит из трех основных этапов: импутация (создание заполненных наборов данных), анализ (применение статистических методов к каждому набору) и объединение (комбинирование результатов). На этапе импутации создается несколько (обычно от 5 до 20) заполненных наборов данных, где пропущенные значения заменяются различными правдоподобными значениями, полученными из наблюдаемых данных. Затем каждый заполненный набор данных анализируется отдельно с использованием стандартных статистических методов. Наконец, результаты всех анализов объединяются для получения итоговых оценок параметров и их стандартных ошибок [14; 15].

Результаты

1. Дендрограмма и иерархический кластерный анализ CA.

Для выявления групп наблюдений со схожими характеристиками применен иерархический кластерный анализ методом Уорда. Результаты кластеризации визуализированы в виде дендрограммы, где по горизонтальной оси расположены даты наблюдений, а по вертикальной — высота кластеризации (рис. 2). Для определения оптимального числа кластеров проведен анализ структуры дендрограммы с выделением трех основных групп наблюдений.

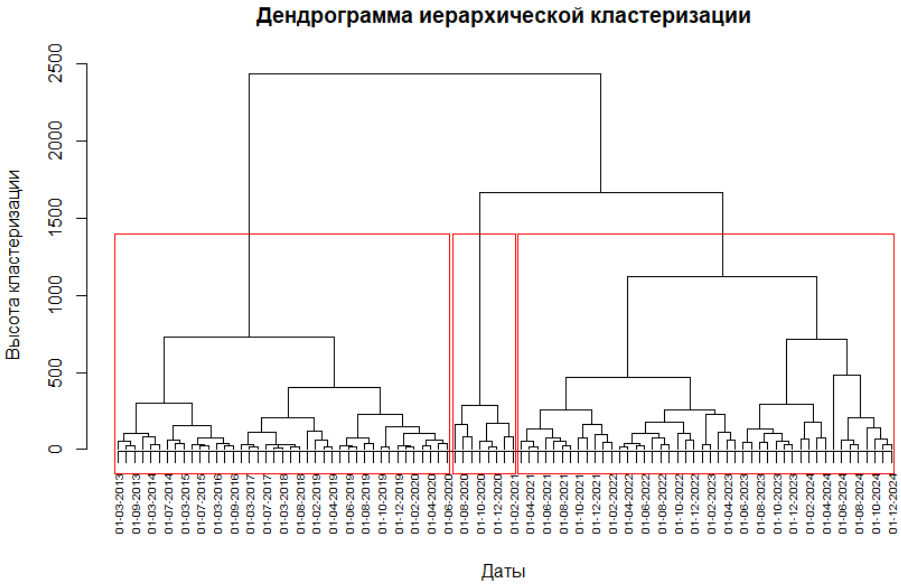

Рис. 2. Дендрограмма иерархической кластеризации

По оси абсцисс расположены даты отбора проб, по оси ординат — высота кластеризации, отражающая степень различия между кластерами. Выделены три основных кластера, отражающие периоды с различными гидрохимическими режимами.

Высота кластеризации варьируется от 0 до 2500 единиц. Выявлены три основных кластера с высотой слияния, равной 1400. Первый кластер включает наблюдения с 1 по 42, характеризующиеся относительно низкими значениями азотных показателей. Третий кластер охватывает период с 50 по 96 наблюдения и соответствует максимальным концентрациям биогенных элементов. Второй кластер содержит данные, сильно отличающиеся от остальных. Продемонстрировано четкое разделение данных на временные группы с различными гидрохимическими характеристиками. Структура дендрограммы указывает на наличие периодов с контрастными режимами поступления биогенных элементов в водную среду.

2. FA.

Для снижения размерности данных и выявления латентных факторов применен FA методом минимальных остатков с вращением «varimax». Предварительно проведена стандартизация данных и проверка пригодности выборки для факторного анализа. Метод минимальных остатков обеспечивает устойчивые результаты при отклонениях от нормального распределения. Выделено два фактора, объясняющих основную долю дисперсии в данных по показателям азота и фосфора (рис. 3).

В математическом выражении модель FA можно представить как $X = LF + \varepsilon$, где X — матрица наблюдаемых переменных (исходные гидрохимические показатели), L — матрица факторных нагрузок, F — матрица факторов, ε — вектор остатков (уникальных компонент).

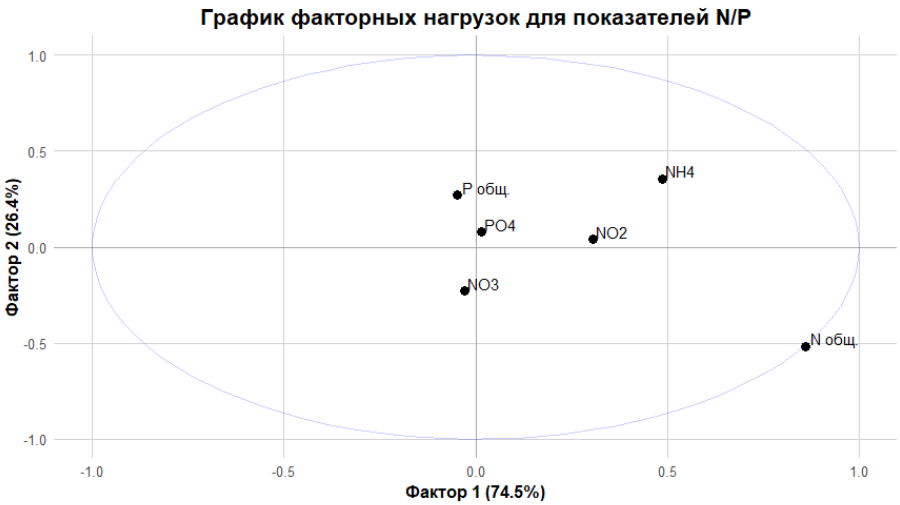

Рис. 3. График факторных нагрузок

Расчет факторных оценок для каждого наблюдения позволил перейти от исходных шести переменных к двум интегральным показателям. Факторные оценки нормализованы со средним значением 0 и стандартным отклонением 1. Это обеспечивает сопоставимость результатов анализа временных рядов.

FA выделил два основных фактора, объясняющих 26,7 % общей дисперсии данных. Первый фактор (MR1) объясняет 17,9 % дисперсии, второй фактор (MR2) — 8,8 % дисперсии.

Фактор 1 характеризуется высокими положительными нагрузками по азоту общему (0,861), иону аммония (0,487) и нитрит-ионам (0,307). Данная структура нагрузок отражает процессы поступления и трансформации органического азота. Высокая корреляция общего азота с аммонием указывает на деструкцию органических соединений с образованием аммиака. Положительная связь с нитритами свидетельствует о протекании первой стадии нитрификации. Фактор интерпретируется как азотная нагрузка от органических источников.

Фактор 2 демонстрирует контрастную структуру нагрузок: положительную по иону аммония (0,357) и отрицательную по общему азоту (–0,517). Такое соотношение указывает на различные механизмы поступления этих форм азота. Отрицательная корреляция между аммонием и общим азотом может отражать поступление неорганических форм азота на фоне снижения органической составляющей. Фактор интерпретируется как соотношение органических и минеральных форм азота.

Нитрат-ионы, фосфат-ионы и общий фосфор не продемонстрировали значимых факторных нагрузок: фосфат-ионы — 0,156 на первый фактор, общий фосфор — 0,198 на первый фактор. Это указывает на независимость динамики фосфорных соединений от азотного цикла в данных условиях.

3. Анализ временных рядов.

Для исследования динамики выделенных факторов во времени построены временные ряды факторных оценок (рис. 4, 5). Применение LOESS-сглаживания проводилось с параметром сглаживания 0,3, что обеспечило оптимальный баланс между выделением долгосрочного тренда и сохранением значимых краткосрочных колебаний. Статистическая значимость трендов оценивалась через доверительные интервалы сглаженных кривых.

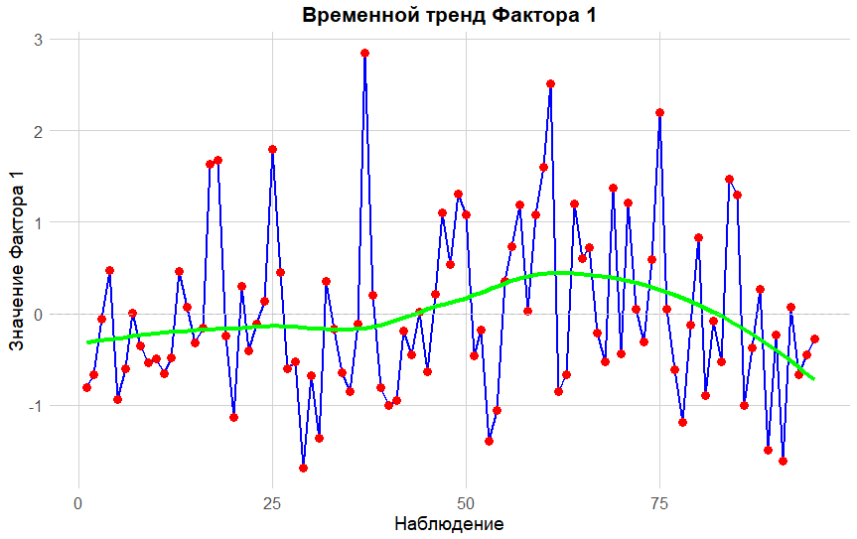

Рис. 4. Временной тренд фактора 1

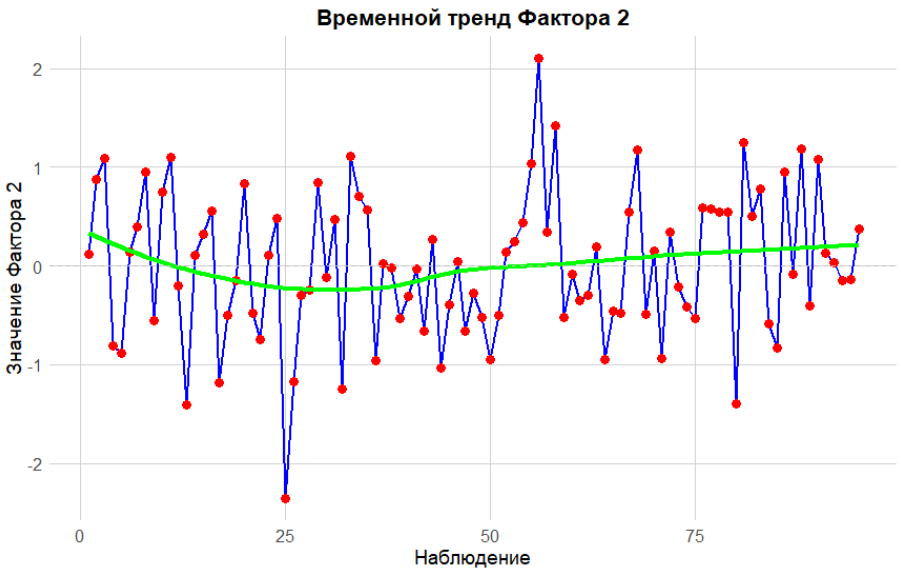

Рис. 5. Временной тренд фактора 2

4. Прогностическая модель.

Для прогнозирования значений гидрохимических показателей разработана модель на основе множественной импутации. Метод позволяет учесть неопределенность в данных и получить более робастные прогнозы по сравнению с детерминистическими подходами. Модель использует информацию о временной структуре данных и взаимосвязях между показателями для восстановления пропущенных значений и прогнозирования будущих наблюдений. Качество модели оценивалось с помощью кросс-валидации и сравнения прогнозных значений с фактическими данными. Однако, поскольку на данный момент не существует «универсальной» модели прогнозирования, результаты множественной импутации носят рекомендательный характер и позволяют лишь предположить состояние водного объекта в рамках исследуемого отрезка времени.

Временной тренд фактора 1 демонстрирует нелинейную динамику с выраженными пиковыми значениями. Максимальное значение фактора 1 достигает 2,84 единиц на 42-м наблюдении, что соответствует периоду максимальной азотной нагрузки. Минимальные значения (–1,67) зафиксированы в начальный период наблюдений. Амплитуда колебаний составляет 4,51 единиц, что превышает утроенное стандартное отклонение. LOESS-сглаживание выявляет куполообразную тенденцию с подъемом в средний период и снижением к концу наблюдений. Высокочастотные колебания указывают на эпизодический характер поступления азотных соединений. Сглаженный тренд фактора 1 показывает подъем с 15-е по 60-е наблюдение с последующим спадом до конца периода. Максимум тренда (0,52) приходится на 58-е наблюдение. Коэффициент автокорреляции первого порядка составляет 0,23, указывая на слабую временную зависимость.

Фактор 2 демонстрирует меньшую изменчивость с максимумом 2,11 на 56-м наблюдении и минимумом –2,31 на 25-м наблюдении. Тренд характеризуется начальным снижением до 25-го наблюдения с последующей стабилизацией около нулевого уровня. Стандартное отклонение составляет 0,89 единиц против 1,12 для первого фактора. Синхронность отдельных пиков с фактором 1 свидетельствует о комплексном воздействии источников загрязнения на различные формы азота.

Среднеквадратичная ошибка (далее — RMSE) представляет собой меру точности модели сглаживания. Для фактора 1 RMSE составила 0,978, для фактора 2 — 0,812. Меньшее значение RMSE для второго фактора указывает на более регулярный характер процессов, им описываемых.

Коэффициент детерминации (далее — R^2) отражает долю дисперсии, объясняемую трендовой составляющей. Значения R^2 равны 0,216 для фактора 1 и 0,123 для фактора 2. Относительно низкие значения соответствуют высокой изменчивости гидрохимических показателей в речных системах.

Тест Дики — Фуллера является статистическим критерием проверки стационарности временных рядов. Результаты теста (p-value < 0,05 для

обоих факторов) подтверждают стационарность рядов после удаления трендовой составляющей. Это позволяет применять стандартные методы анализа временных рядов для дальнейшего исследования.

5. Источники загрязнения.

Структура факторных нагрузок позволяет идентифицировать основные источники поступления азотных соединений. Фактор 1 отражает влияние коммунально-бытовых сточных вод, содержащих органические соединения азота. Пиковые значения могут соответствовать периодам залповых сбросов или снижения эффективности очистки.

Фактор 2 может отражать воздействие атмосферных осадков и поверхностного стока, приносящих минеральные формы азота. Отрицательная корреляция с общим азотом указывает на разбавление органических загрязнителей при поступлении поверхностных вод.

Сезонная динамика факторов обусловлена температурной зависимостью биохимических процессов трансформации азота, изменениями водности реки и вариациями антропогенной нагрузки. Интенсификация процессов аммонификации и нитрификации в теплый период года объясняет сезонные колебания соотношения различных форм азота.

Заключение

Применение статистических методов анализа обработанных гидрохимических данных реки Преголя за период 2013—2024 гг. позволяет сформулировать следующие выводы.

Иерархический CA методом Уорда выявил три основных кластера наблюдений, отражающих различные режимы поступления биогенных элементов в водную среду. Структура дендрограммы демонстрирует четкое временное разделение данных с высотой кластеризации от 0 до 2500 единиц, указывая на наличие периодов с контрастными гидрохимическими характеристиками.

FA идентифицировал два латентных фактора, объясняющих 26,7 % общей дисперсии показателей азотного и фосфорного цикла. Первый фактор (17,9 % дисперсии) интерпретируется как азотная нагрузка от органических источников с доминированием общего азота, аммония и нитритов. Второй фактор (8,8 % дисперсии) отражает соотношение органических и минеральных форм азота, характеризуясь контрастными нагрузками по аммонию и общему азоту.

Анализ временных рядов факторных оценок с применением LOESS-сглаживания выявил нелинейную динамику с выраженными пиковыми значениями и куполообразными трендами. Статистические характеристики временных рядов (RMSE 0,978 и 0,812; R^2 0,216 и 0,123) соответствуют высокой природной изменчивости гидрохимических показателей речных систем.

Фосфорные соединения продемонстрировали независимость от динамики азотного цикла, не показав значимых факторных нагрузок. Данная особенность указывает на различные механизмы и источники поступления соединений фосфора в водную среду.

Структура факторных нагрузок позволяет идентифицировать основные источники загрязнения: коммунально-бытовые сточные воды (органические соединения азота), атмосферные осадки и поверхностный сток (минеральные формы), сезонные изменения.

Интегрированный подход к анализу качества природных вод, сочетающий иерархический CA, FA, исследование временных рядов и прогностическое моделирование, представляет собой эффективный инструментарий для комплексной оценки состояния водных объектов в условиях антропогенного воздействия.

Применение метода Уорда для кластеризации обеспечивает выявление временных групп с различными гидрохимическими характеристиками, позволяя идентифицировать периоды аномального поступления загрязняющих веществ. Факторный анализ с вращением «varimax» эффективно редуцирует размерность данных и выявляет латентные факторы, отражающие основные процессы формирования качества воды.

MI демонстрирует преимущества перед детерминистическими подходами при работе с неполными данными мониторинга, обеспечивая корректную оценку неопределенности результатов. Метод позволяет учитывать временную структуру данных и взаимосвязи между показателями, однако не дает гарантированной точности.

Используемый подход обеспечивает выявление скрытых взаимосвязей между показателями, идентификацию источников загрязнения и разработку научно обоснованных рекомендаций по управлению качеством водных ресурсов.

Список литературы

1. *РД 52.24.643-2002.* Методические указания. Метод комплексной оценки степени загрязненности поверхностных вод по гидрохимическим показателям : руководящий документ. Утв. и введ. в действие Федеральной службой России по гидрометеорологии и мониторингу окружающей среды 4 декабря 2002 г. М., 2002.

2. *Ngwenya B., Paepae T., Bokoro P. N.* Monitoring ambient water quality using machine learning and IoT: A review and recommendations for advancing SDG indicator 6.3.2 // Journal of Water Process Engineering. 2025. Vol. 73. P. 107664. doi: 10.1016/j.jwpe.2025.107664.

3. *Шелутко В. А.* Методы обработки и анализа геоэкологической информации. СПб., 2020.

4. *Шилина М. В., Мусатова О. В., Ивановский В. В.* Биометрия : учеб.-метод. комплекс. Витебск, 2011.

5. *Kannel P. R., Lee S., Kanel S. R., Khan S. P.* Chemometric application in classification and assessment of monitoring locations of an urban river system // Analytica Chimica Acta. 2007. Vol. 582, №2. doi: 10.1016/j.aca.2006.09.006.

6. *Belkhiri L., Boudoukha A., Mouni L., Baouz T.* Application of multivariate statistical methods and inverse geochemical modeling for characterization of groundwater — A case study: Ain Azel plain (Algeria) // Geoderma. 2010. Vol. 159, №3-4. doi: 10.1016/j.geoderma.2010.08.016.

7. *Singh Y., Kumar M.* Interpretation of water quality parameters for Villages of Sanganer Tehsil, by using Multivariate Statistical analysis // Journal of Water Resource and Protection. 2010. Vol. 02, №10. doi: 10.4236/jwarp.2010.210102.

8. *Kausher R., Sinha A. Kr., Singh R.* Chemometric appraisal of groundwater and surface water quality for domestic, irrigation and industrial purposes in the coal mining province of Mahan River catchment area // Desalination and Water Treatment. 2023. Vol. 311. doi: 10.5004/dwt.2023.29971.

9. *Sharma S.* Use of Multivariate Analytical Methods in Assessment of River Water Quality // International Journal of Earth Sciences and Engineering. 2017. Vol. 10, №02. doi: 10.21276/ijee.2017.10.0236.

10. *Gajbhiye S., Sharma S.K., Awasthi M.K.* Application of Principal Components Analysis for Interpretation and Grouping of Water Quality Parameters // International Journal of Hybrid Information Technology. 2015. Vol. 8, №4. doi: 10.14257/ijhit.2015.8.4.11.

11. *Ling T.Y., Soo C.L., Liew J.J. et al.* Application of Multivariate Statistical Analysis in Evaluation of Surface River Water Quality of a Tropical River // Journal of Chemistry. 2017. Vol. 2017. doi: 10.1155/2017/5737452.

12. *Михальчук А.А.* Многомерный статистический анализ эколого-геохимических измерений. Ч. II : Ординация и классификация. СПб., 2012.

13. *Ouyang Y., Nkedi-Kizza P., Wu Q.T. et al.* Assessment of seasonal variations in surface water quality // Water research. 2006. Vol. 40, №20. P. 3800—3810. doi: 10.1016/j.watres.2006.08.030.

14. *Subba Rao N.* Application of Principal Components Analysis for Interpretation and Grouping of Water Quality Parameters // Oriental Journal of Chemistry. 2012. Vol. 28. P. 1749—1751. doi: 10.14257/ijhit.2015.8.4.11.

15. *ГОСТ* Р 58556-2019. Оценка качества воды водных объектов с экологических позиций : национальный стандарт Российской Федерации. Утв. и введ. в действие Приказом Федерального агентства по техническому регулированию и метрологии от 17 октября 2019 г. №965-ст. М., 2019.

Об авторах

Матвей Александрович Смирнов — магистрант, Балтийский федеральный университет им. И. Канта, Россия.

E-mail: maasmirnov@stud.kantiana.ru

Софья Сергеевна Кладова — асп., Балтийский федеральный университет им. И. Канта, Россия.

E-mail: sonickm12@gmail.com

Юлия Владимировна Королева — канд. геогр. наук, доц., Балтийский федеральный университет им. И. Канта, Россия.

ORCID: 0000-0001-7612-4454

E-mail: yu.koroleff@yandex.ru

SPIN-код: 6276-7017

M. A. Smirnov, S. S. Kladova, Y. V. Koroleva

USE OF MATHEMATICAL METHODS IN THE ANALYSIS AND FORECASTING OF NITROGEN-PHOSPHORUS CYCLE INDICATORS IN THE PREGOLYA RIVER BASIN

Immanuel Kant Baltic Federal University, Kaliningrad, Russia
Received 08 September 2025
Accepted 25 November 2025
doi: 10.5922/vestniknat-2026-1-3

To cite this article: Smirnov M. A., Kladova S. S., Koroleva Y. V., 2026, Use of mathematical methods in the analysis and forecasting of nitrogen-phosphorus cycle indicators in the Pregolya river basin, *Vestnik of Immanuel Kant Baltic Federal University. Series: Natural Sciences*, №1. P. 41—53. doi: 10.5922/vestniknat-2026-1-4.

The analysis of multidimensional raw data on the state of aquatic ecosystems often requires the use of modern analytical tools capable of identifying hidden patterns. The purpose of this study is to analyze long-term water quality monitoring data for the Pregolya River (Kaliningrad) in order to identify the main factors determining variability in the hydrochemical regime and to assess their temporal dynamics. To analyze data on indicators of the nitrogen and phosphorus cycles, methods of multivariate statistics were applied, including hierarchical cluster analysis (Ward's method), factor analysis (minimum residual method with varimax rotation), and time-series analysis (LOESS smoothing). As a result of the cluster analysis, three groups of observations were identified, reflecting different periods of nutrient input. Factor analysis made it possible to identify two latent factors: "nitrogen load from organic sources" and "the ratio of organic and mineral forms of nitrogen." Time-series analysis of factor scores revealed nonlinear dynamics and long-term trends in changes in water quality. The study demonstrates that the integrated application of chemometric methods serves as an effective tool for diagnosing the condition of urban water bodies, identifying pollution sources, and providing a scientific basis for the development of management decisions in the field of water resource protection.

Keywords: multivariate statistical analysis, factor analysis, cluster analysis, water quality, nitrogen and phosphorus compounds

The authors

Matvey A. Smirnov, master's student, Immanuel Kant Baltic Federal University, Russia.
E-mail: maasmirnov@stud.kantiana.ru

Sofia S. Kladova, PhD student, Immanuel Kant Baltic Federal University, Russia.
E-mail: sonickm12@gmail.com

Yulia V. Koroleva, PhD in Geography, Associate Professor, Immanuel Kant Baltic Federal University, Russia.
ORCID: 0000-0001-7612-4454
E-mail: yu.koroleff@yandex.ru
SPIN code: 6276-7017