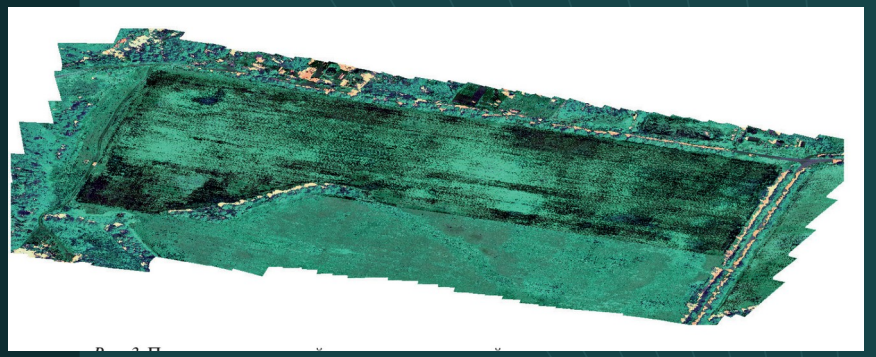
НАУЧНАЯ КОНФЕРЕНЦИЯ • 2026

Предсказание обеспеченности посевов пшеницы макро- и микроэлементами

по мульти- и гиперспектральным данным дистанционного зондирования с применением методов машинного обучения и пространственной кросс-валидации

А. А. Струев, Е. П. Митрофанов, В. П. Якушев,
В. М. Буре, О. А. Митрофанова



12 нутриентов · 200 наблюдений

Актуальность и цель исследования

Проблемы лабораторной диагностики и дистанционного мониторинга

01

Лабораторная диагностика

Точность в отдельных точках, но высокая стоимость, трудоёмкость и отсутствие непрерывной карты.

02

Косвенный спектральный сигнал

Для Cu, Zn, Co и ряда других элементов нет собственных полос поглощения в VNIR.

03

Завышенная оценка моделей

Соседние пробы похожи; случайное разбиение смешивает пространственно зависимые наблюдения.

Цель: выявить воспроизводимый сигнал для 12 нутриентов и честно оценить точность его прогноза

Исходные данные и целевые показатели

Одно поле в Ленинградской области • один вегетационный сезон

200
наблюдений

12
целевых показателей

1
поле

2
фазы развития

НАБОР ЦЕЛЕВЫХ ПЕРЕМЕННЫХ

Макроэлементы и формы азота

N · P · K · Ca · Mg · S · NO_3

Микроэлементы

Fe · Mn · Cu · Zn · Co

Предсказывается концентрация в надземной биомассе, а не содержание элементов в почве

Мульти- и гиперспектральные данные

Спектральное и пространственное разрешение используемых сенсоров

МУЛЬТИСПЕКТРАЛЬНАЯ СЪЁМКА

5 каналов

Синий · зелёный · красный
красный край · ближний ИК

ГИПЕРСПЕКТРАЛЬНАЯ СЪЁМКА

300 каналов

390–1031 нм · ≈ 2 нм на канал
push-broom: последовательные полосы

5 см/пиксель Мультиспектральная ортомозаика

≈ 80 % поля Гиперспектральное покрытие

≈ 115 точек Доступно для сравнения сенсоров

Формирование признакового пространства

Три группы признаков для мультиспектральной модели

44

Вегетационных индекса

NDVI, NDRE, GNDVI, CCCI, индексы пигментов

7 × 7

Оконные статистики

Среднее, медиана, разброс, процентили, вариация

3

Семейства текстур

GLCM, фильтры Габора, локальные бинарные шаблоны (LBP)

Окно 7 × 7 при 5 см/пиксель → 35 × 35 см: локальная структура посевов

Построение регрессионных моделей

Отбор и настройка выполняются строго внутри обучающих подвыборок

01 Спектральные данные

02 ВИ, окна, текстуры

03 mRMR топ-50

04 Регрессоры + Optuna

05 Buffered-LOO

06 12 карт нутриентов

Cu: |r| = 0,034

dCor = 0,366 → возможна нелинейная связь

Предотвращение утечек

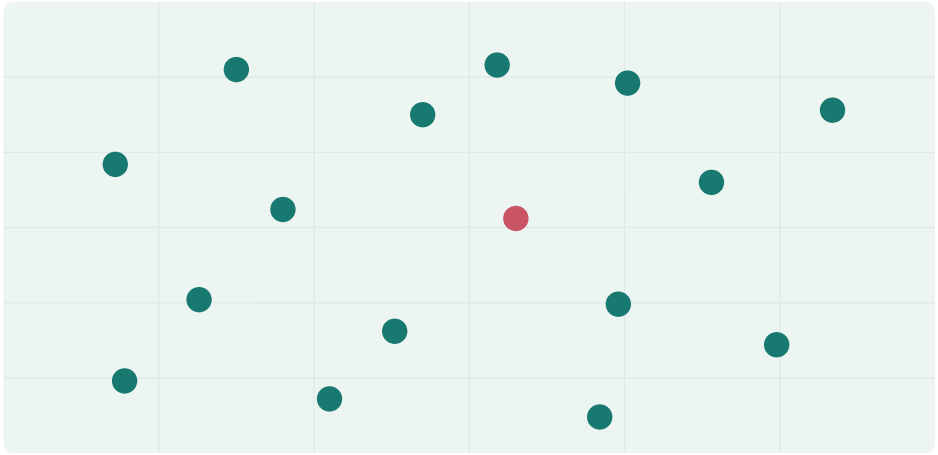
Нормализация · импутация · mRMR · C-Mixup — только на train

В статье исследованы Bayesian Ridge, ElasticNet, RF, ExtraTrees, XGBoost, LightGBM и SVR

Пространственная кросс-валидация

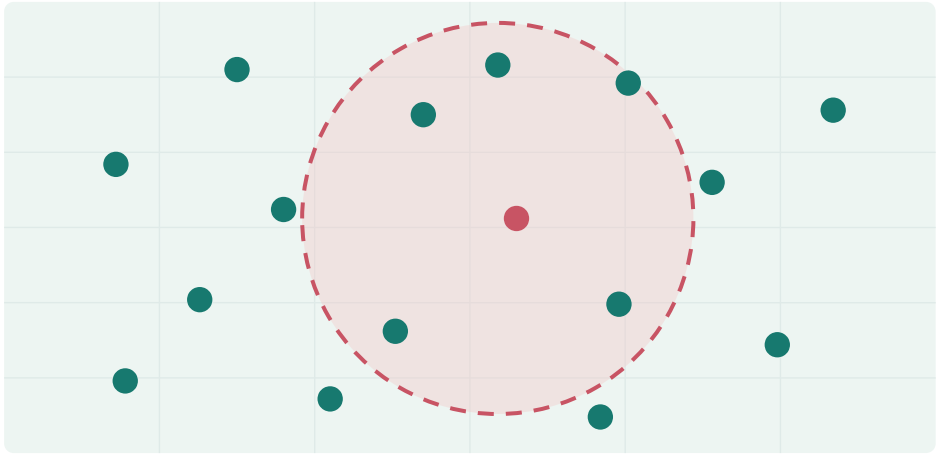
Соседние наблюдения пространственно зависимы — тестовые точки нужно изолировать

Случайное разбиение



Тестовая точка может соседствовать с train

Buffered leave-one-out



Тестовая точка + исключение соседей

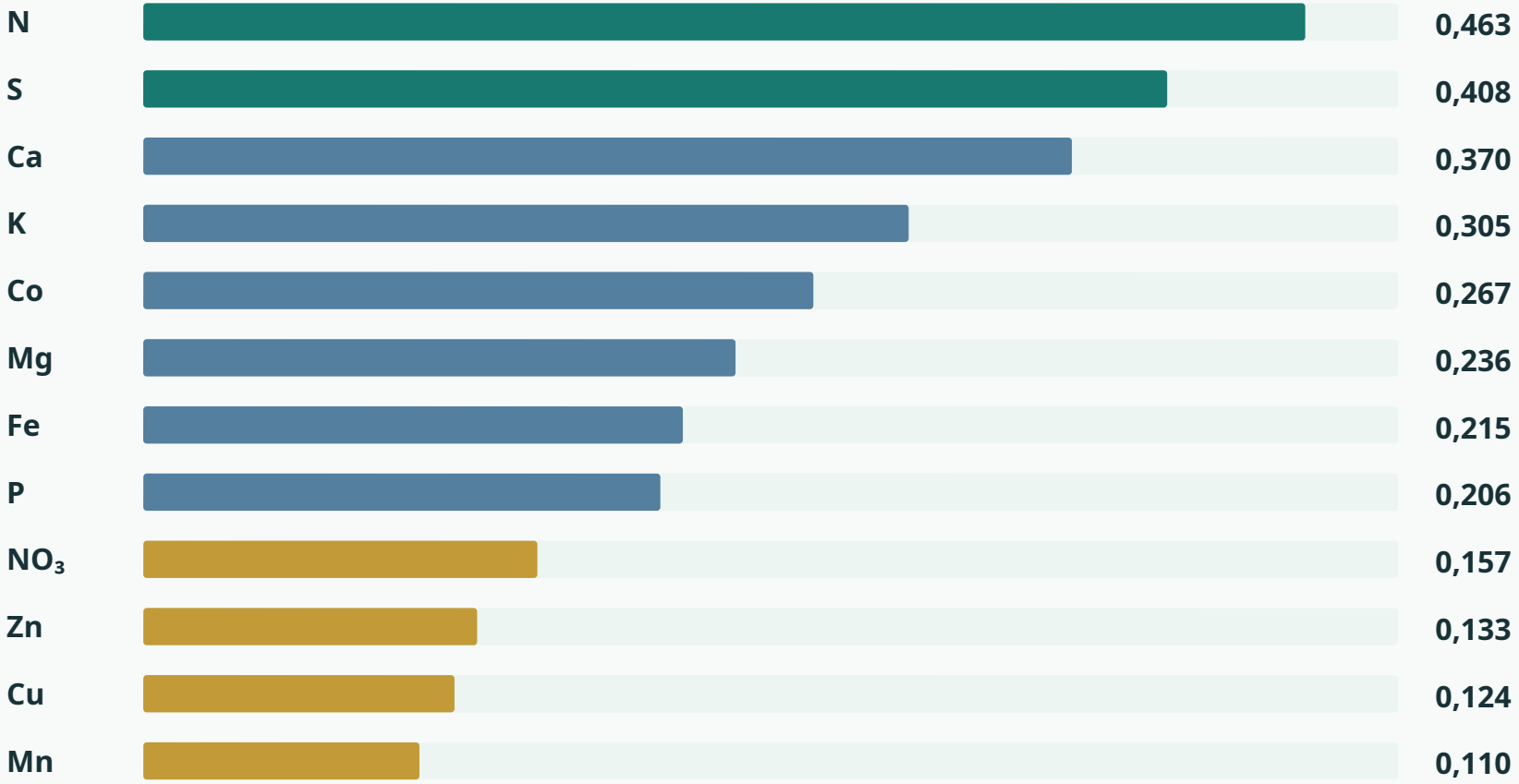
10/12 значимый индекс Морана

30–100 м радиусы буферов

+0,13 среднее завышение R^2

Результаты прогнозирования 12 нутриентов

Основной мультиспектральный конвейер • Pollux • n = 200 • buffered-LOO



0,463

лучший результат для азота

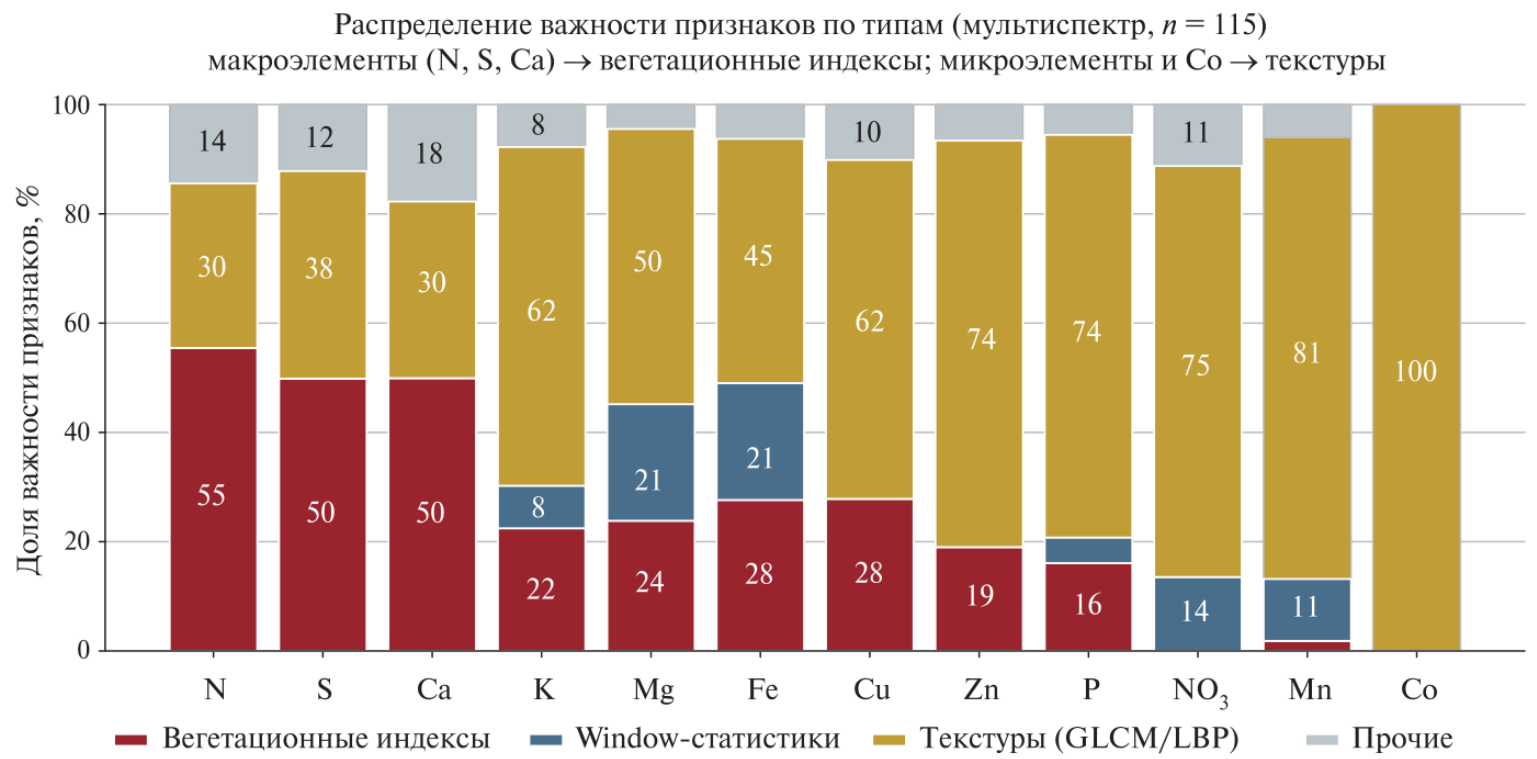
95 % ДИ

включает 0 для NO_3, Cu и Mn

R^2: 0,11–0,46 • Две модели ≥ 0,4; шесть — 0,2–0,4; четыре — ниже 0,2

Важность спектральных и текстурных признаков

Рисунок 1 из статьи: важность групп признаков в лучших мультиспектральных моделях



50–55 %

вегетационные индексы для N, S и Ca

62–100 %

текстуры для многих микроэлементов и кобальта

Важность модели ≠ доказательство причинной биохимической связи

Сравнение мультиспектральных камер

Даже камеры с одинаковым числом каналов могут давать разные признаки

10 / 12

нутриентов: Pollux показал более высокую максимальную корреляцию

r = 0,08–0,26

Корреляция спектральных измерений двух камер в одинаковых точках

Предполагаемая причина — различие центральных длин волн и ширины каналов, особенно «красного края»

Сравнение мульти- и гиперспектральных подходов

n ≈ 115 • различающиеся наборы признаков при одинаковой выборке



11 / 12

нутриентов лучше у мультиспектра

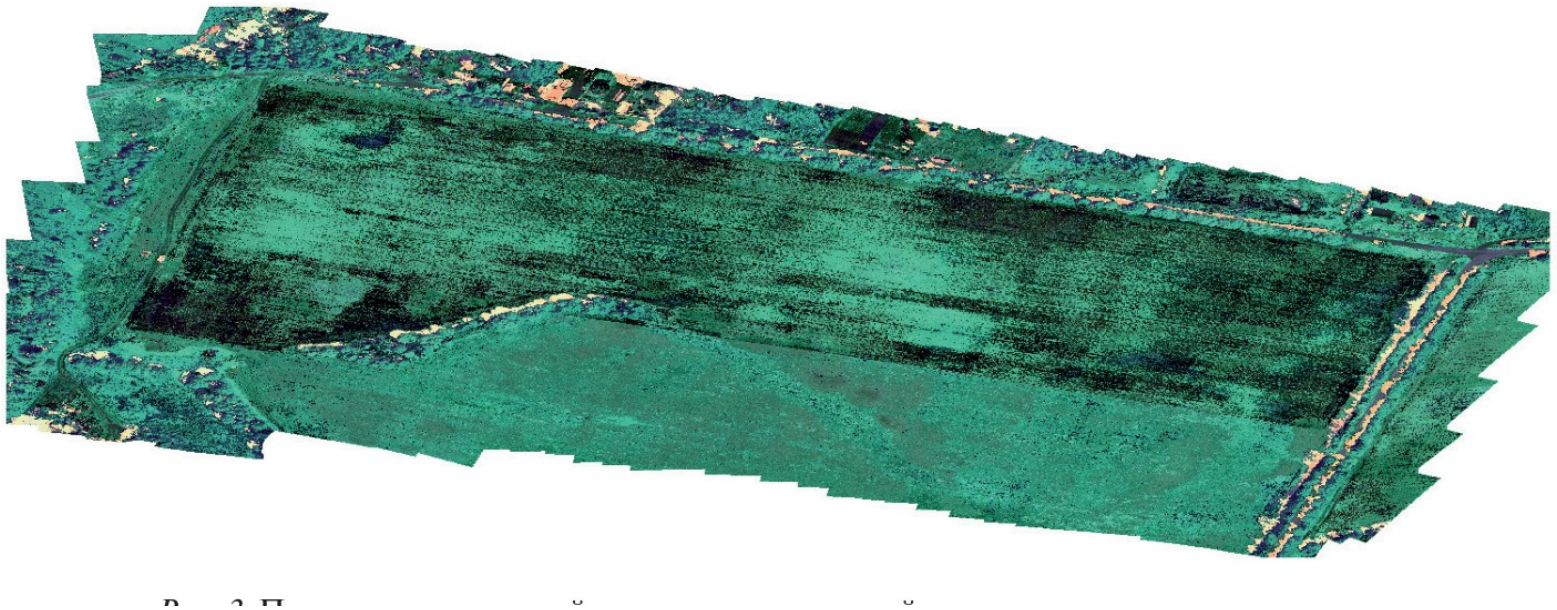
53–80 %

информации гиперспектра вне типичного диапазона

Данные из рис. 2 статьи; сравнение конфигураций, а не спектральной информативности самих сенсоров

Построение карт концентраций нутриентов

Прогнозирование в каждом пикселе с геопривязкой и оценкой неопределённости



Пример из статьи: ложное RGB-композитное отображение трёх прогнозных каналов

12

каналов GeoTIFF

512 × 512

тайлы с перекрытием

σ

карта неопределённости

Дальнейшее применение: ГИС-анализ и дифференцированное внесение удобрений после агрономической проверки

Струев и соавт. • Современные проблемы ДЗЗ из космоса, 2026 • DOI: 10.21046/2070-7401-2026-23-4-282-292

Результаты исследования и направления дальнейшей работы

01

Текстуры

Сильный информационный канал для микроэлементов и Co

02

Пространственная проверка

Случайная CV завышает R^2 в среднем на 0,13

03

Прикладной результат

12-канальные карты концентраций с геопривязкой

Ограничения: одно поле и один сезон • следующий этап: кросс-полевая проверка, гиперспектральные ортомозаики, SWIR

БЛАГОДАРЮ ЗА ВНИМАНИЕ!

Благодарности

Финансирование исследования и сведения о проекте

Российский научный фонд (РНФ)

Грант № 24-16-00203

НАИМЕНОВАНИЕ ПРОЕКТА

Методология обработки данных дистанционного зондирования получаемых при мульти- и гиперспектральной съёмке посевов зерновых культур и сопряженных полевых измерений

Исследование выполнено за счёт гранта Российского научного фонда № 24-16-00203.

https://rscf.ru/project/24-16-00203/