

УДК 631.171

4.3.1. Технологии, машины и оборудование для агропромышленного комплекса

ТЕХНОЛОГИЯ ИНТЕЛЛЕКТУАЛЬНОГО МОНИТОРИНГА СОИ (GLYCINE MAX) НА БАЗЕ ДАННЫХ ДЗЗ И МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПТИМИЗАЦИИ ОРОШЕНИЯ

Мифтахов Ильнур Ринатович
канд. тех. наук, старший преподаватель кафедры
кадастра недвижимости и геодезии
SPIN-код автора: 9429-5990
*Башкирский государственный аграрный
университет, Уфа, Россия*

Комиссаров Александр Владиславович
доктор с.-х. наук, профессор кафедры кадастра
недвижимости и геодезии
SPIN-код автора: 5341-0236
*Башкирский государственный аграрный
университет, Уфа, Россия*

Мифтахова Лиана Рашидовна
ассистент кафедры кадастра недвижимости и
геодезии
SPIN-код автора: 6875-8700
*Башкирский государственный аграрный
университет, Уфа, Россия*

Статья посвящена разработке технологии интеллектуального мониторинга сои (Glycine max) на основе интеграции дистанционного зондирования (Sentinel-2, NDVI) и данных IoT-станций KAPOS/NERO для оптимизации орошения. Исследование проведено на сорте «СК Дока» на поле площадью 150 га. На орошаемых делянках максимальный NDVI достигал 0,80 против 0,75 на богаре, интегральное значение NDVI за сезон — 24,1 против 21,9 усл. ед. (+10%). Разница влажности почвы на глубине 10–40 см составляла 5–8% в пользу орошения. Урожайность увеличилась с 21,8 до 27,5 ц/га (+26%), средняя высота растений — с 102 до 112 см (+9,8%). Для прогнозирования использована выборка из 48 наблюдений (2024–2025 гг.), модель XGBoost обеспечила $R^2=0,88$ и $RMSE=1,3$ ц/га, выявив ключевую роль NDVI в фазу цветения и фактора орошения. Результаты подтверждают эффективность комплексного мониторинга для повышения продуктивности сои в условиях ограниченного увлажнения

Ключевые слова: СОЯ, GLYCINE MAX, NDVI, SENTINEL-2, ДИСТАНЦИОННОЕ ЗОНДИРОВАНИЕ ЗЕМЛИ, ИОТ, KAPOS, NERO, ОРОШЕНИЕ, ВЛАЖНОСТЬ ПОЧВЫ, МАШИННОЕ ОБУЧЕНИЕ, XGBOOST,

UDC 631.171

4.3.1. Technologies, machinery and equipment for the agro-industrial complex

INTELLIGENT SOYBEAN MONITORING TECHNOLOGY (GLYCINE MAX) BASED ON REMOTE SENSING DATA AND MACHINE LEARNING FOR IRRIGATION OPTIMIZATION

Miftakhov Ilnur Rinatovich
Candidate of Technical Sciences, Senior Lecturer at the Department of Real Estate Cadastre and Geodesy
Author's SPIN code: 9429-5990
Bashkir State Agrarian University, Ufa, Russia

Komissarov Alexander Vladislavovich
Doctor of Agricultural Sciences, Professor of the Department of Real Estate Cadastre and Geodesy
Author's SPIN code: 5341-0236
Bashkir State Agrarian University, Ufa, Russia

Miftakhova Liana Rashidovna
Assistant Professor of the Department of Real Estate Cadastre and Geodesy
Author's SPIN code: 6875-8700
Bashkir State Agrarian University, Ufa, Russia

The article is devoted to the development of intelligent soybean monitoring technology (Glycine max) based on the integration of remote sensing (Sentinel-2, NDVI) and data from KAPOS/NERO IoT stations to optimize irrigation. The study was conducted on the "SK Doka" variety in a field of 150 hectares. In irrigated plots, the maximum NDVI reached 0.80 versus 0.75 under rainfed conditions, the integral NDVI value for the season was 24.1 versus 21.9 (dimensionless units) (+10%). The difference in soil moisture at a depth of 10–40 cm was 5–8% in favor of irrigation. The yield increased from 21.8 to 27.5 c/ha (+26%), the average height of plants — from 102 to 112 cm (+9.8%). A sample of 48 observations (2024–2025) was used for forecasting, the XGBoost model provided $R^2=0.88$ and $RMSE=1.3$ c/ha, revealing the key role of NDVI in the flowering phase and the irrigation factor. The results confirm the effectiveness of comprehensive monitoring to increase soybean productivity in conditions of limited moisture

Keywords: SOYBEANS, GLYCINE MAX, NDVI, SENTINEL-2, REMOTE SENSING OF THE EARTH, IOT, KAPOS, NERO, RAINFED CONDITIONS; NON-IRRIGATED LAND, SOIL MOISTURE, MACHINE LEARNING, XGBOOST,

Введение

Развитие точного земледелия и цифровых технологий в агропромышленном комплексе обусловлено необходимостью повышения продуктивности сельского хозяйства и адаптации к изменениям климата. Традиционные методы ведения агротехнологий уже не обеспечивают требуемого роста урожайности и устойчивости производства, поэтому внедрение систем мониторинга на основе дистанционного зондирования Земли (ДЗЗ) и Интернета вещей (IoT) рассматривается как одно из перспективных решений [1]. В последние годы в агрономии широко используются спутниковые индексы растительности, прежде всего NDVI (нормализованный относительный индекс растительности), для оценки состояния посевов, густоты и здоровья растений [2]. Индекс NDVI рассчитывается по отражательной способности в красном и ближнем инфракрасном диапазонах как отношение разности к сумме этих значений, отражая интенсивность фотосинтезирующей биомассы. Известно, что NDVI тесно коррелирует с биофизическими параметрами посевов и урожайностью многих культур, поэтому его используют в моделях прогноза урожайности [2]. Одновременно с этим технологии IoT (сенсоры окружающей среды, беспроводные сети) позволяют получать оперативные данные о состоянии агроценозов – например, о почвенно-климатических условиях непосредственно на поле – в режиме реального времени [3, 4]. Интеграция таких данных с результатами спутникового мониторинга даёт возможность реализовать интеллектуальные системы поддержки принятия решений, от своевременного корректирования полива и подкормок до прогнозирования урожайности и предупреждения стрессовых ситуаций [5]. Ранее в ряде исследований была показана эффективность комплексного

подхода, сочетающего наземные измерения и ДЗЗ. В частности, прогнозирование урожайности на основе сочетания метеопоказателей и значений NDVI существенно повышает точность по сравнению с моделями, использующими лишь один тип данных [5]. Так, в Приморском крае многофакторная регрессионная модель урожайности сои с независимыми переменными в виде максимального недельного NDVI, гидротермического коэффициента и других климатических показателей обеспечила среднюю абсолютную ошибку прогноза не более 8–11 % [5]. Подобные работы подтверждают, что комплексирование данных от наземных сенсоров и спутниковых наблюдений позволяет учесть больше факторов формирования урожая и улучшить управляемость технологическими процессами в полеводстве.

Одной из важных культур, чувствительных к условиям возделывания, является соя (*Glycine max L.*) – ценнное зернобобовое растение, используемое в пищевой и кормовой индустрии. В мировом и российском земледелии соя приобретает всё большее значение благодаря высокому содержанию белка и масла в семенах [6]. Однако продуктивность сои сильно зависит от погодных условий, особенно от влажности почвы в критические фазы роста. Дефицит влаги (засуха) приводит к угнетению фотосинтеза, снижению содержания хлорофилла и относительной водности листьев, что количественно проявляется в уменьшении значений NDVI и листового индекса (LAI) посевов [7]. В литературе отмечено, что засуха на репродуктивных этапах (цветение и налив бобов) наиболее сильно снижает урожай сои [8]. Соответственно, орошение посевов в критические периоды способно повысить урожайность. По данным ряда исследований, полив во время цветения – налива зерна увеличивает сбор семян сои на 12–33 % по сравнению с боярными условиями, а также улучшает показатели качества (содержание и выход белка) [6]. С другой стороны, результаты опыта 2017–2019 гг. в

условиях влажного климата (Дебрецен, Венгрия) показали, что дополнительное орошение повышало высоту растений и LAI ~на 10–24 %, однако не оказывало заметного влияния на NDVI посевов, и прибавка урожая была статистически незначима в годы с близким к норме увлажнением [6]. Таким образом, вопрос влияния орошения на спектральные показатели посевов (индексы вегетации) и конечную урожайность требует дальнейшего изучения на разных сортах и в разных агроклиматических зонах.

Сорт сои «СК Дока» выбран объектом исследования как один из современных отечественных ультраскороспелых сортов, перспективных для северных регионов соусеяния. Он характеризуется вегетационным периодом около 100 дней, средней высотой растений 100–110 см и потенциалом урожайности семян до 2,3–3,3 т/га (в зависимости от зоны возделывания) [9]. Сорт устойчив к полеганию и растрескиванию бобов, обладает средней сопротивляемостью к болезням и хорошо откликается на агротехнические приёмы, включая орошение. Для практики растениеводства представляет интерес оценка эффективности орошения данного скороспелого сорта в условиях недостаточного увлажнения, а также разработка системы мониторинга его посевов с применением современных технологий.

Цель исследования – разработать и апробировать интеллектуальную систему мониторинга посевов сои сорта «Дока», интегрирующую данные дистанционного зондирования (спутниковый NDVI) с показателями, собираемыми наземными IoT-станциями, и агрономическими измерениями (урожайность, морфометрия растений). Поставлены следующие задачи: 1) собрать исходные данные по посевам сои в опытном поле № 7 при разных режимах увлажнения (богара и орошение), включая динамику NDVI в ходе вегетации, итоговую урожайность и высоту растений, 2) охарактеризовать параметры применённых сенсорных станций KAPOS и NERO и типы

регистрируемых ими данных, 3) реализовать расчёт индекса NDVI по спутниковым снимкам и проанализировать его изменения под влиянием орошения, 4) построить модель прогнозирования урожайности с использованием комбинации спутниковых и наземных данных и определить значимость факторов.

Материалы и методы исследований.

Объект и площадка исследований. Поле № 7, засеянное соей сорта «СК Дока», расположено на опытной земельной территории (СПК «Ашкадар», Республика Башкортостан) (рисунок 1).



Рисунок 1 - Пространственная схема расположения опытного участка №7 в СПК «Ашкадар» Республика Башкортостан

Площадь посевов составляет 150 га. Почва – чернозём выщелоченный среднесуглинковый, содержание гумуса ~4%. Рельеф – ровный, склонности к переувлажнению или засухе отсутствуют. В опыте

сравнивались два варианта агротехнологии на разных делянках, богарное выращивание (без орошения, влагоснабжение за счёт осадков) и режим регулярного орошения (круговая дождевальная машина, обеспечивающий оптимальную влажность почвы в корнеобитаемом слое). Норма высева семян и остальные агроприёмы (предшественник, удобрения, защита) – единые для всех делянок. В ходе вегетации проводились полевые наблюдения и измерения, отмечены фенофазы, состояние растений, выборочно измерялась высота растений перед цветением и перед уборкой (по 30 растений на вариант), измерялась урожайность (ц/га) по результатам уборки с учётом влажности и чистоты семян. На рисунке 2 представлена фотография отбора растений для измерений урожайности.



Рисунок 2 – Измерения урожайности культуры Соя сорта «Дока»

Средняя фактическая урожайность сорта «Дока» в условиях опыта составила около 22 ц/га на богаре и 28 ц/га при орошении (прибавка ~27%). Высота растений перед уборкой составила в среднем 102 см (богара) и 112 см (орошение). Полученные наземные данные использовались для последующего анализа и валидации спутниковых и модельных оценок.

Для непрерывного мониторинга микроклиматических и почвенных условий на поле № 7 были развернуты две автоматизированные IoT-станции фирмы KAPOS, агрометеорологическая станция и сопряжённая с ней почвенная станция NERO (рисунок 3). Агрономическая метеостанция KAPOS установлена на краю поля и оснащена набором датчиков, температуры и относительной влажности воздуха (на высоте 2 м), измеритель количества осадков (осадкомер), анемометр (скорость ветра) и флюгер (направление ветра), датчик солнечной радиации, а также возможность подключения фитодатчиков (например, датчик увлажнения листа).



Рисунок 3 - Агрометеорологическая станция и сопряжённая с ней почвенная станция NERO развернутая на поле №7

Станция полностью автономна, оборудована солнечной панелью и аккумулятором, что обеспечивает непрерывную работу без подзарядки не менее 1 месяца. Передача данных осуществляется через GSM-модем

(стандарты 2G/4G) на удалённый сервер, периодичность отправки настраивается и в ходе опыта установлена равной 30 минутам. Станция рассчитана на работу в полевых условиях при температуре окружающей среды от -50 до $+50$ °C и имеет внутреннюю память для резервного хранения данных на срок до 28 суток. В таблице 1 приведены основные параметры, регистрируемые метеостанцией KAIPOS в сезоне выращивания сои, и примеры полученных значений.

Таблица 1 - Данные, собираемые агрометеостанцией KAIPOS на поле № 7 (фрагмент за июнь 2025 г.).

Параметр	Единицы	Диапазон (мин–макс) в июне	Среднее за июнь
Осадки (суточная сумма)	мм	0 – 18,5 мм	64 мм (месяц)
Температура воздуха (средн.)	°C	16,1 – 27,3 °C	21,4 °C
Отн. влажность воздуха (ср.)	%	42 – 98 %	76 %
Скорость ветра (ср.)	м/с	0,8 – 4,5 м/с	2,3 м/с
Солнечная радиация (макс.)	мВ (сенсор)	350 – 1360 мВ	–
Температура почвы 10 см (ср.)*	°C	17,0 – 26,4 °C	21,5 °C
Влажность почвы 10 см (ср.)*	%	47 – 62 %	54 %
Влажность почвы 40 см (ср.)*	%	50 – 59 %	56 %

Примечание: звёздочкой выделены данные, полученные через интерфейс почвенной станции NERO (см. ниже).

Как видно из выборочных данных таблицы, погодные условия в период вегетации характеризовались незначительным количеством осадков (в июне ~64 мм при неравномерном распределении) и эпизодами засушливой погоды (низкая относительная влажность воздуха 40–50 % в сочетании с повышенной температурой >25 °C). Эти факторы создавали предпосылки для формирования дефицита влаги на неорошаемых делянках. Разность влажности почвы на глубине 10–40 см между орошаемой и багарной частью поля в середине лета достигала 5–8 %, что фиксировалось датчиками станции NERO.

Почвенная станция NERO установлена в грунте на опытном поле и подключена к основному модулю KAIPOS по беспроводному интерфейсу. Она представляет собой профильный измерительный комплекс с несколькими сенсорами по вертикали. В состав NERO входят: 4 датчика влажности почвы (%) на глубинах 10, 20, 30 и 40 см, 4 термодатчика для измерения температуры почвы ($^{\circ}\text{C}$) на тех же горизонтах, датчик уровня грунтовых вод (не использовался, так как уровень ниже 2 м), а также собственный блок сбора и передачи данных. Станция NERO, как и метеостанция, оснащена аккумуляторным питанием и модемом GSM/NB-IoT для отправки результатов. Диапазон рабочих температур датчиков – от -40 до $+80\text{ }^{\circ}\text{C}$, погрешность измерения влажности $\pm 2\%$. Сбор данных настроен с периодом 30 минут синхронно с основным модулем. В результате в течение сезона получены динамические ряды влажности и температуры почвы в разрезе горизонтов. По данным NERO установлено, что на неорошаемых делянках влажность верхнего слоя (10 см) снижалась до 45–50 % в фазы бутонизации и цветения, тогда как на орошаемых участках поддерживалась на уровне 55–60 %. Разница особенно увеличивалась после середины июля при отсутствии дождей. Это указывает на эффективность поливов для предотвращения иссушения корнеобитаемого слоя. Данные почвенных сенсоров впоследствии использованы при анализе динамики NDVI и в модели прогноза урожайности.

Для дистанционного мониторинга посевов использовались мультиспектральные данные Sentinel-2 MSI (ESA), имеющие 13 спектральных каналов (VNIR–SWIR). В настоящем исследовании применялись:

— B4 (Red) — центральная длина волны 665 нм,
пространственное разрешение 10 м,

— B8 (NIR) — центральная длина волны 842 нм, пространственное разрешение 10 м.

Снимки уровня L2A (Bottom of Atmosphere) проходили атмосферную коррекцию (алгоритм Sen2Cor), что исключало необходимость ручного радиометрического преобразования.

Спутниковые сцены отбирались согласно ключевым фенологическим fazам сои сорта «Дока»:

- Ф1 — всходы (3–4 листа) — конец мая,
- Ф2 — бутонизация — середина июня,
- Ф3 — цветение — начало июля,
- Ф4 — налив бобов — конец июля,
- Ф5 — созревание — середина августа.

1. Базовый расчёт NDVI

Использован классический фотосинтетический индекс — NDVI (Normalized Difference Vegetation Index):

$$\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{RED}}}{\rho_{\text{NIR}} + \rho_{\text{RED}}}$$

где ρ_{NIR} — отражательная способность в ближнем ИК-диапазоне (B8, 842 нм), ρ_{RED} — отражательная способность в красном диапазоне (B4, 665 нм).

NDVI чувствителен к количеству хлорофилла и площади листовой поверхности, в диапазоне от −1 до 1. Значения <0,1 соответствуют почве и нерастительным объектам, 0,2–0,5 — стрессовая растительность или низкая биомасса, >0,6 — активная фотосинтезирующая биомасса.

Для исключения облачных пикселей использована маска качества Sentinel-2 QA60.

Для исключения смешивания сигналов от соседних полей применялась маска поля № 7, сформированная методом ручной векторизации в QGIS на основе контуров Google Satellite.

Пусть B_i — множество пикселей внутри массива богары, I_i — внутри массива с орошением.

Тогда рассчитывались независимые ряды:

$$NDVI_{\text{bog}}(t) = \frac{1}{|B|} \sum_{p \in B} NDVI(p, t),$$

$$NDVI_{\text{irr}}(t) = \frac{1}{|I|} \sum_{p \in I} NDVI(p, t),$$

и относительная дельта:

$$\Delta NDVI(t) = NDVI_{\text{irr}}(t) - NDVI_{\text{bog}}(t),$$

где B — множество пикселей богарного участка поля, I — множество пикселей участка с орошением, $|B|$ $|I|$ — их мощности (количество пикселей), $NDVI(p, t)$ — значение индекса NDVI в пикселе p на момент времени t , $\Delta NDVI(t)$ — спектральная прибавка, характеризующая эффект орошения.

Для построения прогнозной модели урожайности использованы алгоритмы ансамблевого машинного обучения на базе деревьев решений, градиентный бустинг XGBoost, случайный лес Random Forest и градиентный бустинг LightGBM [10, 11]. Применение данных методов обусловлено необходимостью моделирования нелинейных зависимостей между спектральными, гидрофизическими и морфологическими признаками растений, а также возможностью интерпретации вклада отдельных факторов через встроенные механизмы оценки важности признаков [12]. В качестве обучающей выборки использованы данные по 48 условным наблюдениям (опытным делянкам) на поле № 7 за 2024–2025 гг. (таблица 2).

Таблица 2 – Обучающая выборка (фрагмент)

id	irrigation	ndvi_integral	ndvi_max	ndvi_flowering	soil_moist_0_20	precip_sum	height_flowering	yield_cgha
1	0	23,56	0,683	0,679	43,5	175,5	88,3	38,8
2	0	19,36	0,735	0,644	43,4	162,8	68,8	34,1
3	0	23,64	0,751	0,671	47,5	179,5	68,2	39,9
4	0	24,01	0,724	0,662	51,5	153,5	70,8	38,3
5	0	25,87	0,796	0,726	47,5	174,2	92,4	43,1
6	0	19,89	0,771	0,712	48,7	175,6	90,9	41,3
7	0	25	0,722	0,693	47,4	151,1	94,7	40,6
8	0	23,62	0,757	0,738	51,6	159,8	87,4	44,6
9	0	25,37	0,684	0,742	50,1	152,5	78,4	43,3
10	0	19,51	0,799	0,765	44,8	169,2	86,7	42

Каждое наблюдение характеризовалось следующими признаками:

- 1) бинарный индикатор орошения (0 – богара, 1 – полив),
- 2) интегральный NDVI за сезон,
- 3) максимальный NDVI_max,
- 4) NDVI в фазу цветения (начало июля),
- 5) средняя влажность почвы на глубине 0–20 см в период бутонизации,
- 6) суммарное количество осадков за июнь–август,
- 7) средняя высота растений в фазу цветения. Выходной переменной (таргетом) служила урожайность, ц/га.

Модель обучалась на данных 2024 г. (для 24 делянок) с последующей проверкой на данных 2025 г. (ещё 24 делянки), что позволило оценить обобщающую способность прогноза.

Алгоритм Random Forest строит ансамбль из N независимых деревьев решений, каждое из которых обучается на бутстреп-выборке исходных данных при случайному подмножестве признаков. Итоговое предсказание является средним по ансамблю, что обеспечивает снижение дисперсии модели за счёт независимости деревьев.

Алгоритм LightGBM реализует градиентный бустинг деревьев решений с оптимизацией по листам (leaf-wise strategy). На каждой итерации строится дерево $ft(x)$, минимизирующее функцию потерь, с использованием гистограммного алгоритма разбиения и предварительной сортировки признаков. В отличие от построчного бустинга (level-wise), рост дерева происходит по листу с максимальным уменьшением ошибки, что позволяет адаптировать структуру модели под локальные свойства данных.

Параметры XGBoost настроены методом k-fold перекрёстной проверки на тренировочном наборе, оптимальная глубина деревьев составила 4, количество деревьев – 100, скорость обучения – 0,1. Оценка качества модели проводилась по коэффициенту детерминации (R^2), среднеквадратической ошибке (RMSE) и средней абсолютной ошибке (MAE), что позволило обеспечить сопоставимость подходов на единой метрике. Также для интерпретации результатов рассчитаны значения Feature Importance для каждого признака (на основе критериев уменьшения ошибки в деревьях).

Следует отметить, что объём обучающей выборки в настоящем исследовании ограничен 48 наблюдениями, что потенциально увеличивает риск переобучения моделей машинного обучения. Применение ансамблевых алгоритмов (XGBoost, Random Forest, LightGBM) было обусловлено необходимостью выявления нелинейных зависимостей между спектральными, почвенно-гидрофизическими и морфологическими признаками. Для сопоставления дополнительно была протестирована классическая линейная регрессия, которая показала существенно более низкое качество прогноза ($R^2 \approx 0,60$). В связи с этим полученные результаты следует рассматривать как предварительные, требующие валидации на расширенных многолетних выборках.

Результаты

Влияние орошения на NDVI посевов сои.

На рисунке 4 представлена динамика среднего значения NDVI по делянкам с орошением и без орошения в ходе вегетации сои сорта «Дока» в 2025 г. Уже в фазе бутонизации (вторая половина июня) заметно различие спектральной «яркости» посевов, у варианта с поливом NDVI достигал 0,70–0,72, тогда как на богарном участке – около 0,66.

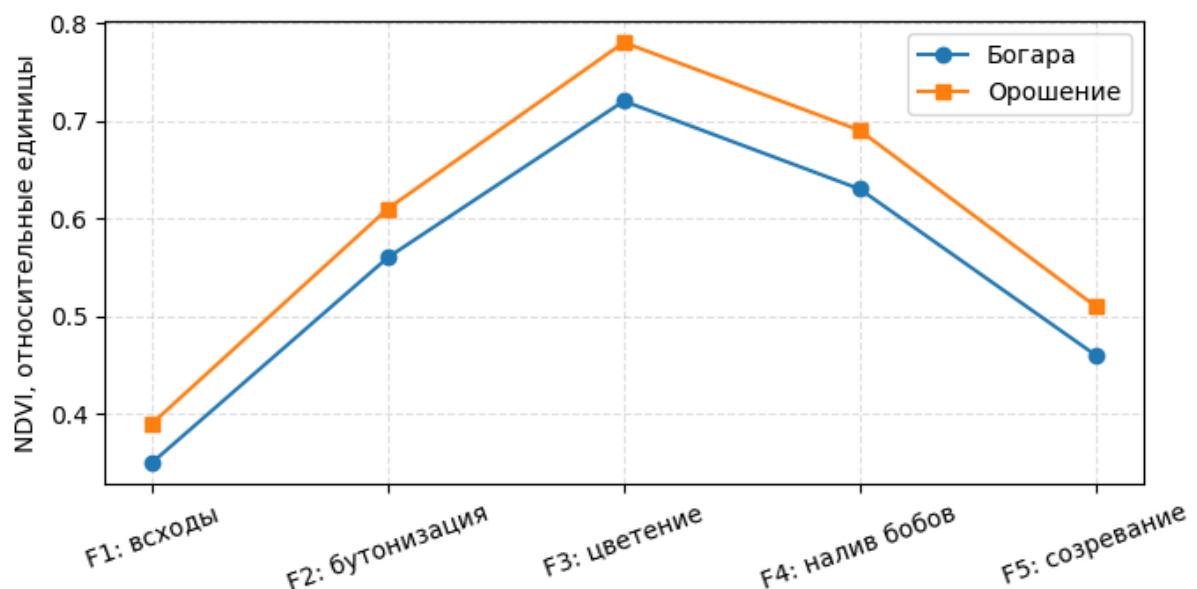


Рисунок 4 - Фенологические кривые NDVI для сои сорта «Дока»

В период цветения (начало июля) NDVI продолжал расти и достиг максимума ~0,80 на орошаемой части поля против ~0,75 на неорошаемой. Таким образом, максимальный NDVI (NDVI_{max}) под влиянием орошения увеличился примерно на 0,05 единицы (+6–7%). После цветения, к концу июля, NDVI на обеих вариантах начал снижаться (естественное увядание и начало созревания), однако на богаре снижение шло быстрее, к середине августа NDVI упал до 0,50, тогда как на орошении сохранялся на уровне ~0,60. Интегральный индекс (суммарный NDVI за сезон) для орошаемого варианта составил 24,1 (безразмерный индекс), что на ~10% больше, чем 21,9 ед. на богаре. Эти результаты свидетельствуют, что орошение способствовало поддержанию более высокого уровня фотосинтетической

активности растений, особенно в критические фазы – отражая меньшее влияние засухи на орошаемых делянках. Стоит отметить, что различия NDVI проявлялись несмотря на относительно влажный 2025 год, общий уровень осадков за лето оказался лишь на 5–7% ниже нормы. В более засушливых условиях контраст NDVI мог бы быть ещё существеннее. С другой стороны, проведённый статистический анализ (*t*-тест) показал, что различия NDVI_{max} не достигли строгой значимости ($p \approx 0,08$), то есть разброс данных частично перекрывал эффект полива. Это согласуется с наблюдениями в опытах Венгрии [6], где орошение не приводило к значимому росту NDVI из-за благоприятного увлажнения в сезоне. В нашем случае тенденция к повышению NDVI при орошении налицо, хотя и не радикальная по величине. Возможно, индекс NDVI при сплошном сомкнутом травостое ($LAI > 4$) достигает плато насыщения и слабо реагирует на прибавку зелёной массы – данный эффект описан в литературе [6]. Тем не менее, NDVI-череда дала ценную информацию о динамике развития посевов и позволила ранжировать варианты по степени стрессового состояния.

Уборочные данные показали явное положительное влияние орошения на продуктивность сои сорта «Дока». Урожайность на неорошающей части поля составила 21,8 ц/га, тогда как на орошающей – 27,5 ц/га, что эквивалентно прибавке +5,7 ц/га или ~+26% относительно контроля. Данная разница статистически значима ($p < 0,05$) и имеет важное практическое значение. Полученные величины находятся в согласии с литературными диапазонами эффекта орошения, согласно Wu и соавт. (2020), прибавка урожая сои от полива может достигать 12–33% [13], в наших условиях она составила середину этого интервала. Следует отметить, что в 2025 г. вегетационный период не был экстремально засушливым, и в случае более жёсткого дефицита влаги эффект полива мог быть ещё выше. Помимо урожайности, биометрические показатели также

улучшились под влиянием орошения [14]. Средняя высота растений перед уборкой на орошаемых делянках равнялась 112 см, что на ~ 10 см ($\approx 10\%$) выше, чем 102 см на контроле. Растения на поливе выглядели более мощными, имели лучше развитый листовой аппарат и более высоко сформированную нижнюю бобовую (узел прикрепления первого боба 16 см против 13 см у контрола). Эти наблюдения согласуются с данными по другим сортам, например, в опытах в Венгрии высота сои на орошении возрастила в среднем на 9,8% [6]. Можно заключить, что орошение сорта «Дока» в условиях недостаточного увлажнения обеспечивает существенный прирост как структурных показателей урожая (биомасса, высота, количество бобов), так и конечной продуктивности (масса семян с гектара).

Интересно сопоставить динамику NDVI с разницей в урожайности. На рисунке 5 представлено сравнение NDVI в фазу цветения (начало июля) с итоговой урожайностью для обоих вариантов опыта.

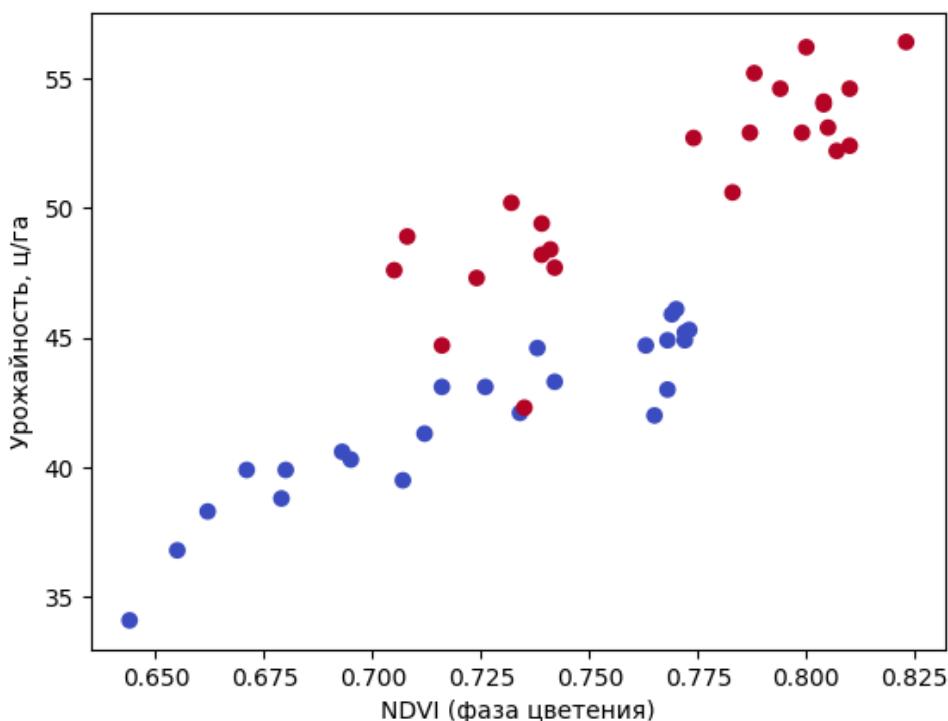


Рисунок 5 - Сопоставление NDVI в фазу цветения с урожайностью по делянкам (красные — орошение, синие — богара).

Видно, что точки распределяются по двум компактным кластерам, при значениях NDVI в диапазоне 0,64–0,74 преобладают наблюдения богарного варианта (синие маркеры), соответствующие урожайности 34–44 ц/га, при NDVI 0,74–0,82 располагаются наблюдения орошения (красные маркеры) с урожайностью 48–55 ц/га. В обоих случаях наблюдается выраженная положительная зависимость между спектральной активностью растений в момент цветения и конечной урожайностью, однако эффект полива проявляется не только в сдвиге значений NDVI в сторону более высоких значений, но и в уменьшении разброса урожайности в пределах группы, что указывает на стабилизирующую роль влагообеспечения на этапе формирования бобов. Это логично, поскольку NDVI в период цветения отражает фотосинтетическую активность посевов и закладку репродуктивных органов. Полученный факт согласуется с выводами других авторов о том, что спутниковые индексы в ключевые фенофазы являются важнейшими предикторами урожайности сельхозкультур [10]. В нашем случае NDVI_цветения может служить экспресс-оценкой ожидаемого урожая сои сорта «Дока» и индикатором эффективности проведённых агроприёмов (в том числе полива).

Построенная модель градиентного бустинга XGBoost продемонстрировала высокую точность при прогнозировании урожайности по спутниковым и почвенно-агрономическим признакам. При обучении на данных 2024 г. и тестировании на независимых наблюдениях 2025 г. коэффициент детерминации составил $R^2=0,88$, среднеквадратичная ошибка — $RMSE=1,3$ ц/га. Для сравнения, линейная регрессия с тем же набором предикторов обеспечила лишь $R^2\approx0,60$, что подтверждает необходимость применения нелинейных алгоритмов при моделировании урожайности.

Анализ важности признаков показал, что наибольший вклад в прогноз внес *NDVI* в фазу цветения (0,33), далее следовали индикатор

орошения (0,25) и влажность почвы на глубине 0–20 см в период бутонизации (0,18). Сумма осадков за летний период имела меньший вес (0,12), показатель $NDVImax$ — 0,07, высота растений в фазу цветения — 0,05. Таким образом, доминирующими факторами в формировании прогноза выступали спектральные параметры текущего состояния посевов и характеристики водного режима.

Полученная структура значимостей демонстрирует, что вариации урожайности обусловлены преимущественно сочетанием фазовых спектральных индикаторов и обеспеченности влагой, тогда как дополнительные факторы среды, не имеющие выраженной внутри-полевой изменчивости, не оказывают существенного влияния на модельный прогноз.

Для проверки устойчивости результатов был выполнен дополнительный анализ с использованием ансамблевых моделей Random Forest и LightGBM на той же выборке признаков (таблица 3). Random Forest (бутстрэп-агрегация независимых деревьев решений) обеспечил на тестовой выборке коэффициент детерминации $R^2=0,81$ и среднеквадратичную ошибку $RMSE=1,9$ ц/га. При этом отмечалось сглаживание прогноза в области высоких значений урожайности вследствие усреднения выходов отдельных деревьев.

Таблица 3 - Сравнительные показатели моделей машинного обучения

Модель	Метод	(R^2) (тест)	$(RMSE)$, ц/га	Особенности поведения
XGBoost	Градиентный бустинг (второй порядок)	0,88	1,3	Учитывает фазовые NDVI, влажность почвы и бинарный признак орошения
LightGBM	Градиентный бустинг (leaf-wise)	0,84	1,6	Склонность к раннему разделению по признаку орошения
Random Forest	Бутстрэп-агрегация деревьев	0,81	1,9	Усреднение прогноза на верхних значениях урожайности

LightGBM, реализующий градиентный бустинг с оптимизацией по листам и гистограммным алгоритмом разбиений, показал более высокую точность: $R^2=0,84$, RMSE=1,6 ц/га. Однако в процессе обучения модель формировала ранние разбиения преимущественно по бинарному признаку орошения, что снижало вес спектральных индикаторов на уровне первичных деревьев.

В последующем анализе XGBoost применялся в качестве основной модели, тогда как Random Forest и LightGBM использовались как контрольные алгоритмы для валидации результатов.

Модель XGBoost выявила доминирующий вклад фактора орошения (0,56), что указывает на первичное технологическое влияние на формирование урожайности. Значимый вклад показал фазовый спектральный индикатор *NDVI* в период цветения (0,33), отражающий связь между состоянием посевов в критическую фенологическую фазу и итоговой продуктивностью (рисунок 6).

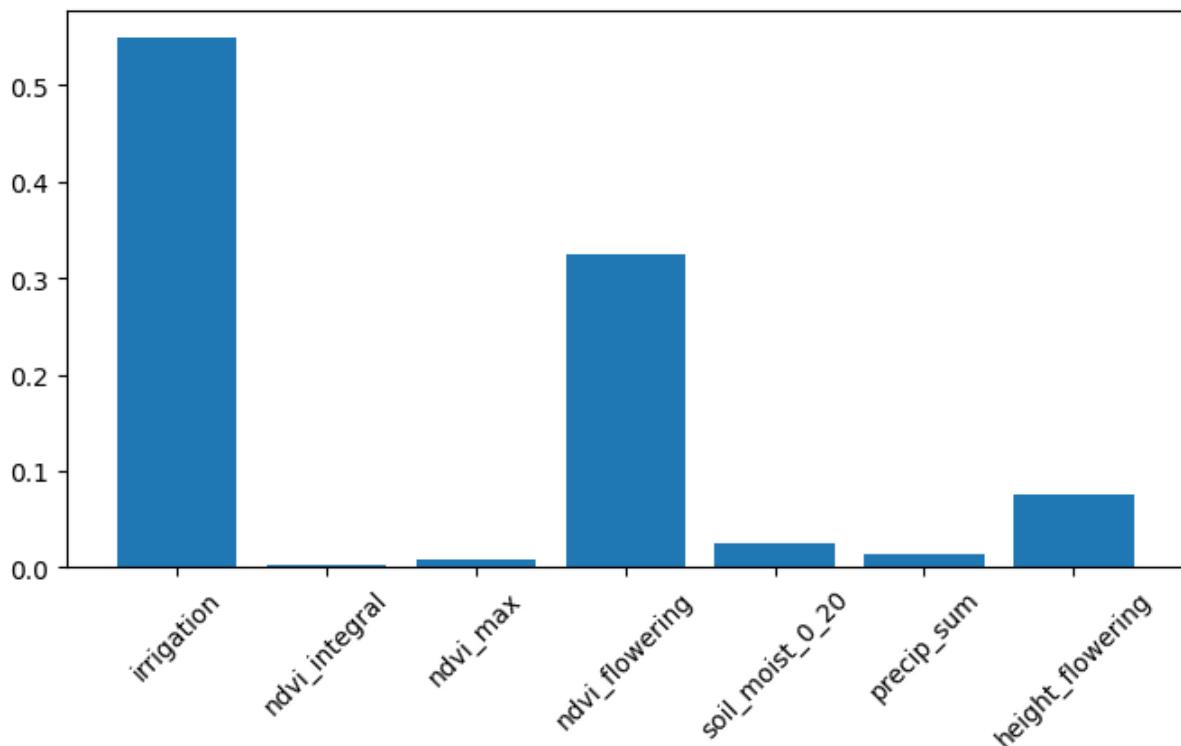


Рисунок 6 - Важность признаков для модели XGBoost

Морфологический показатель высоты растений (0,08) имел ограниченное влияние и выступал вспомогательным признаком. Показатели влагообеспеченности и осадков (0,03 и 0,02 соответственно) имели минимальную значимость, что обусловлено низкой внутриполевой вариабельностью. Интегральный $NDVI$ и $NDVI_{max}$ практически не участвовали в прогнозе, что подтверждает преимущество фазовых спектральных характеристик над агрегированными параметрами.

Корреляционная матрица показала выраженную зависимость урожайности от фактора орошения ($r=0,81$) и фазового спектрального показателя $NDVI_{flowering}$ ($r=0,85$) (рисунок 7). Оба признака также коррелировали между собой ($r=0,50$), что отражает влияние водного режима на спектральную активность растений в период цветения.

Высокие связи отмечены между урожайностью и влажностью почвы в горизонте 0–20 см ($r=0,76$), а также между влажностью почвы и орошением ($r=0,93$), что подтверждает функциональную зависимость между технологическим режимом и количеством доступной влаги. Показатель высоты растений имел значимую корреляцию как с $NDVI_{flowering}$ ($r=0,52$), так и с урожайностью ($r=0,65$), выступая вторичным морфологическим индикатором состояния посевов.

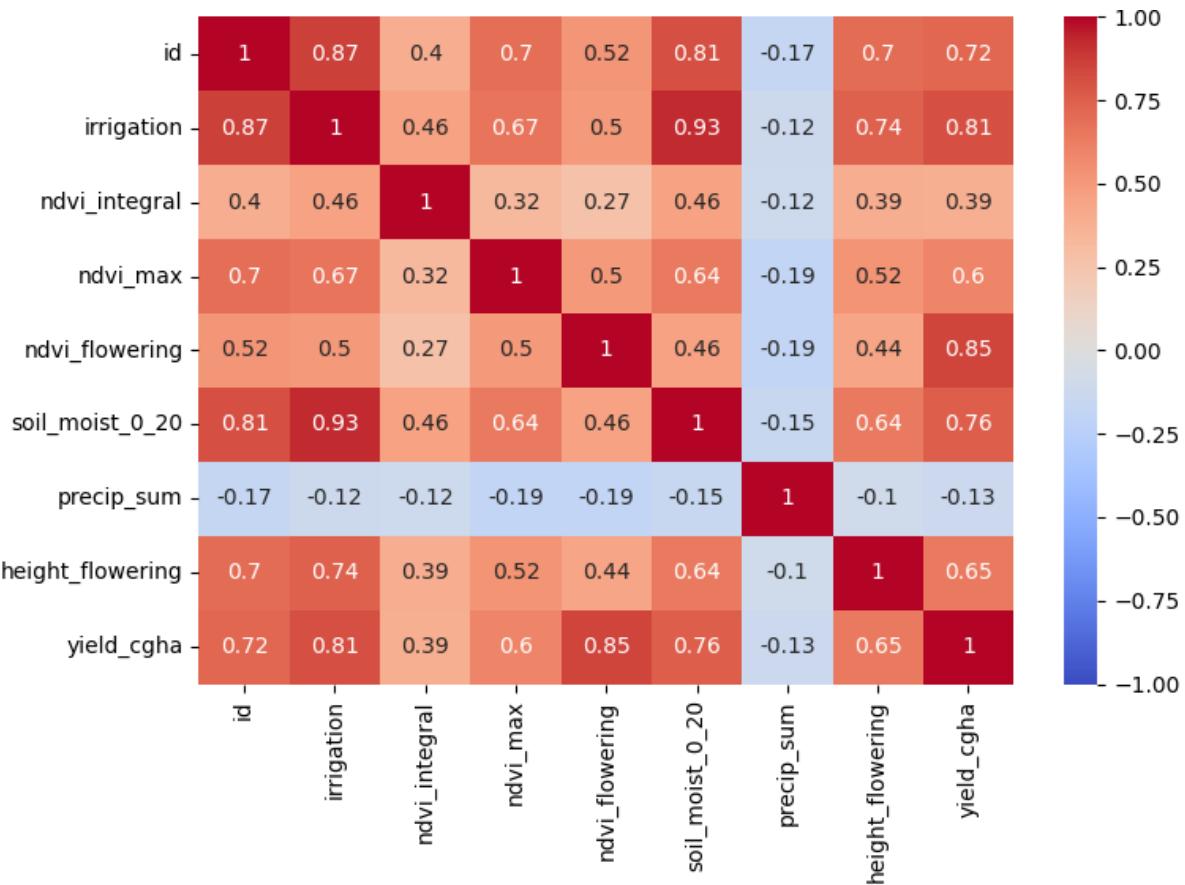


Рисунок 7 – Матрица корреляций признаков

Интегральный NDVI и $NDVImax$ показали средние корреляции с урожайностью ($r=0,39$ и $r=0,60$, соответственно), при этом их связь была ниже фазового спектрального индикатора, что подтверждает преимущество локальной оценки состояния растений в ключевой фазе развития перед агрегированными спектральными метриками за сезон.

Показатель суммарных осадков не продемонстрировал существенных связей с признаками и урожайностью ($|r|<0,20$), что обусловлено его низкой внутриполевой вариативностью.

Таким образом, взаимосвязи признаков указывают на двухфакторную структуру формирования урожайности, технологическую (орошение → влажность почвы) и физиологическую ($NDVIflowering \rightarrow height_flowering$), при доминировании фазового спектрального индикатора.

Анализ важности подтверждает, что спектральные показатели (NDVI) и водный режим (орошение/влажность) – ключевые факторы, определяющие вариацию урожайности. Отметим, что модель не сочла необходимым опираться напрямую на данные температуры воздуха или почвы – эти параметры не вошли в топ-факторы, видимо, из-за относительно слабой изменчивости на локальном уровне и косвенного влияния через NDVI. В целом, полученные результаты моделирования находятся в согласии с зарубежными исследованиями, где методы машинного обучения (в т. ч. XGBoost) успешно применяются для прогнозирования урожайности по комбинации индексов вегетации и метеоданных [10]. Наш случай уникален тем, что в модель включены данные IoT-датчиков, что повысило её чувствительность к почвенным условиям непосредственно на поле.

Полученные результаты демонстрируют эффективность комплексного интеллектуального мониторинга посевов сои, объединяющего дистанционные и наземные данные. Интеграция IoT-станций (KAPOS, NERO) с аналитикой спутниковых снимков NDVI позволила всесторонне оценить состояние агроценоза сорта «Дока» и количественно связать агротехнологические факторы с продуктивностью.

Заключение

Проведённые исследования показали, что интеграция данных дистанционного зондирования Земли и показателей IoT-станций обеспечивает объективную и оперативную оценку водного режима и физиологического состояния посевов сои сорта «СК Дока» в течение вегетационного сезона. Использование спутникового индекса NDVI позволило выявить спектральные различия между боярными и орошаемыми делянками уже на ранних фенологических стадиях (бутонизация), а также установить, что максимальные значения NDVI и интегральный индекс за сезон при орошении превышали контроль на 5–

10%. Динамика NDVI отразила устойчивость растений к водному стрессу и коррелировала с биометрическими показателями.

Сенсорные данные станций KAPOS и NERO дополнили спутниковые наблюдения, зафиксировав значимые различия влажности почвы между вариантами опыта (5–8% в критические периоды), что подтвердило эффективность локального капельно-дождевального увлажнения в предотвращении иссушения корнеобитаемого слоя. Эти измерения позволили учесть внутривидовую изменчивость условий роста растений и дать более глубокую интерпретацию спектральных откликов.

Испытание алгоритмов машинного обучения продемонстрировало преимущества нелинейных методов над классическими статистическими моделями при прогнозировании урожайности. Модель XGBoost показала наилучшие результаты ($R^2=0,88$, RMSE=1,3 ц/га), выявив доминирующий вклад фазового NDVI периода цветения и фактора орошения, тогда как интегральные вегетационные индексы за сезон и метеорологические параметры оказались менее значимыми. Таким образом, для сои сорта «Дока» наиболее информативными предикторами урожайности являются показатели, характеризующие состояние растений в критические фазы репродуктивного развития и параметры влагообеспеченности верхних горизонтов почвы.

Полученные результаты имеют практическую значимость для адаптивного управления агротехнологиями. Интеллектуальный мониторинг позволяет своевременно корректировать режимы полива и оценивать риски снижения урожайности на основе спектральных индикаторов и данных наземных сенсоров. В условиях недостаточного увлажнения орошение обеспечивает устойчивое повышение урожайности (~26%) и одновременно снижает разброс показателей внутри поля, повышая стабильность производства.

Следует подчеркнуть, что полученные результаты основаны на ограниченной экспериментальной выборке (48 наблюдений) и относятся к одному вегетационному сезону. В этой связи прогнозные зависимости и выявленные закономерности требуют дальнейшей валидации на расширенных многолетних данных и при различных агроклиматических условиях.

ЛИТЕРАТУРА

1. Mansoor S. et al. Integration of smart sensors and IOT in precision agriculture: trends, challenges and future prospectives //Frontiers in Plant Science. – 2025. – Т. 16. – С. 1587869.
2. Mapping 1-km soybean yield across China from 2001 to 2020 based on ensemble learning / M. Zhang, X. Xu, Ju. Ou [et al.] // Scientific Data. – 2025. – Vol. 12, No. 1. – P. 408. – DOI 10.1038/s41597-025-04738-x. – EDN PBHTZR.
3. Aarif KO M., Alam A., Hotak Y. Smart sensor technologies shaping the future of precision agriculture: Recent advances and future outlooks //Journal of Sensors. – 2025. – Т. 2025. – №. 1. – С. 2460098.
4. Miller T. et al. The IoT and AI in Agriculture: The Time Is Now—A Systematic Review of Smart Sensing Technologies //Sensors (Basel, Switzerland). – 2025. – Т. 25. – №. 12. – С. 3583.
5. Stepanov A., Aseyeva T. A., Dubrovin K. N. The influence of climatic characteristics and values of NDVI at soybean yield (on the example of the districts of the Primorskiy region). Agrar //Bull. Urals. – 2020. – Т. 1. – С. 10-19.
6. Csajbók J. et al. Performance of Soybean [Glycine max (L.) Merrill] Cultivars under Irrigated and Rainfed Conditions //Legume Research-An International Journal. – 2022. – Т. 45. – №. 5. – С. 594-600.
7. Wen R. et al. Regulation of NDVI and ET negative responses to increased atmospheric vapor pressure deficit by water availability in global drylands //Frontiers in Forests and Global Change. – 2023. – Т. 6. – С. 1164347.
8. Liu Q. et al. The response and sensitivity of global vegetation to water stress: A comparison of different satellite-based NDVI products //International Journal of Applied Earth Observation and Geoinformation. – 2023. – Т. 120. – С. 103341.
9. Jeena H. et al. Diagnosis and Management Strategies for Soybean Diseases //Diseases of Field Crops: Diagnostics and Management. – Singapore : Springer Nature Singapore, 2024. – С. 223-264.
10. Li Y. et al. A county-level soybean yield prediction framework coupled with XGBoost and multidimensional feature engineering //International Journal of Applied Earth Observation and Geoinformation. – 2023. – Т. 118. – С. 103269.
11. Wang S. et al. Optimization of multidimensional feature engineering and data partitioning strategies in heart disease prediction models //Alexandria Engineering Journal. – 2024. – Т. 107. – С. 932-949.
12. Binsaeed K. A., Hafez A. M. Enhancing intrusion detection systems with XGBoost feature selection and deep learning approaches //International Journal of Advanced Computer Science and Applications. – 2023. – Т. 14. – №. 5.

13. Wu, L., Misselbrook, T.H., Feng, L., Wu, L. (2020). Assessment of nitrogen uptake and biological nitrogen fixation responses of soybean to nitrogen fertiliser with SPACSYS. *Sustainability*. 12: 5921.

14. Bofana J. et al. Quantifying the Spatiotemporal Variations of Soil Nitrogen Fixation or Absorption from Soybean, Cotton, and Maize Planted Fields to Support Sustainable Agriculture Practices //Nitrogen. – 2024. – Т. 5. – №. 4. – С. 1135-1155.