

ОРИГИНАЛЬНАЯ СТАТЬЯ

УДК 631.1:004.048

<https://doi.org/10.26897/2687-1149-2025-1-41-52>**Машинное обучение при прогнозировании продуктивности севооборотов*****В.К. Каличкин¹, Д.С. Федоров², К.Ю. Максимович³, В.С. Риксен⁴***^{1,2,3,4} Сибирский федеральный научный центр агробиотехнологий Российской академии наук; р.п. Краснообск, Россия¹ vk.kalichkin@gmail.com; <https://orcid.org/0000-0002-7765-3451>² dima.fedorov99@mail.ru; <https://orcid.org/0000-0001-8678-400X>³ kiri-maksimovi@mail.ru; <https://orcid.org/0000-0002-9563-4641>⁴ riclog@mail.ru; <https://orcid.org/0000-0001-9898-4754>

Аннотация. Севооборот способствует поддержанию устойчивых систем земледелия. Применение машинного обучения позволит более эффективно проектировать и прогнозировать продуктивность севооборотов. Традиционные методы обработки данных не отвечают требованиям интеллектуального земледелия. С целью оценки применения машинного обучения выполнено построение моделей прогнозирования продуктивности севооборотов на основе применения 6 алгоритмов: дерево решений (CART); случайный лес (RF); бутстрэп-агрегирование (Bagging); градиентный бустинг (Gradient Boosting); экстремальный градиентный бустинг (Baging XGBoost); искусственная нейронная сеть (ANN). В исследованиях использованы временные ряды данных по продуктивности 9 типов севооборотов на трех уровнях применения техногенных средств, полученные в лесостепи Приобья Новосибирской области Сибирским НИИ земледелия и химизации сельского хозяйства СФНЦА РАН в течение 1999-2019 гг. В качестве дополнительного предиктора в модели был включен показатель атмосферного увлажнения в виде стандартизованного индекса осадков (Standardized Precipitation Index – SPI), рассчитанный как средний показатель атмосферного увлажнения для мая-июля за ротацию каждого из анализируемых севооборотов. Установили, что модели, описывающие продуктивность севооборотов на основе алгоритмов ANN, Gradient Boosting и XGBoost, характеризовались наиболее высокими прогностическими способностями в зависимости от складывающихся условий атмосферного увлажнения и уровня интенсификации технологии возделывания ($R^2 = 0,90..0,93$). Сравнительный анализ показал, что модель на основе экстремального градиентного бустинга демонстрирует наилучшие показатели с коэффициентом детерминации (R^2) 0,93, среднеквадратичной ошибкой (RMSE) 2,34 и средней абсолютной ошибкой (MAE) 1,81. Продемонстрирована возможность применения методов машинного обучения в качестве эффективного инструментария для прогнозирования продуктивности севооборотов.

Ключевые слова: севооборот, продуктивность севооборотов, прогнозирование, машинное обучение, искусственный интеллект

Благодарности. Авторы выражают признательность Н.В. Васильевой (в.н.с. лаборатории плодородия СФНЦА РАН, канд. с.-х. наук) и Г.М. Захарову (зав. лабораторией севооборотов СФНЦА РАН, канд. с.-х. наук) за предоставленную информацию по полевым опытам с севооборотами.

Для цитирования: Каличкин В.К., Федоров Д.С., Максимович К.Ю., Риксен В.С. Машинное обучение при прогнозировании продуктивности севооборотов // Агроинженерия. 2025. Т. 27. № 1. С. 41-52. <https://doi.org/10.26897/2687-1149-2025-1-41-52>

ORIGINAL ARTICLE

Machine learning in predicting crop rotation productivity***V.K. Kalichkin¹, D.S. Fedorov², K. Yu. Maksimovich³, V.S. Riksen⁴***^{1,2,3,4} Siberian Federal Scientific Centre of AgroBioTechnologies of the Russian Academy of Sciences; Krasnoobsk, Russia¹ vk.kalichkin@gmail.com; <https://orcid.org/0000-0002-7765-3451>² dima.fedorov99@mail.ru; <https://orcid.org/0000-0001-8678-400X>³ kiri-maksimovi@mail.ru; <https://orcid.org/0000-0002-9563-4641>⁴ riclog@mail.ru; <https://orcid.org/0000-0001-9898-4754>

Abstract. Crop rotation contributes to maintaining sustainable farming systems. The application of machine learning will enable more efficient design and prediction of crop rotation productivity. Traditional data processing methods do not meet the requirements of intelligent farming. To evaluate the application of machine learning, models

for predicting crop rotation productivity were built based on six algorithms: decision tree (CART), random forest (RF), bootstrap aggregating (Bagging), Gradient Boosting, extreme gradient boosting (XGBoost), and artificial neural network (ANN). The study used time series data on the productivity of nine types of crop rotations at three levels of technogenic inputs, obtained in the forest-steppe of the Ob region in Novosibirsk Oblast by the Siberian Research Institute of Agriculture and Chemicalization of Agriculture of the SFSCA RAS during 1999-2019. As an additional predictor, the model included an atmospheric moisture indicator in the form of the Standardized Precipitation Index (SPI), calculated as the average atmospheric moisture indicator for May-July over the rotation of each analyzed crop rotation. Models describing crop rotation productivity based on ANN, Gradient Boosting, and XGBoost algorithms were characterized by the highest predictive abilities depending on the prevailing atmospheric moisture conditions and the level of cultivation technology intensification ($R^2 = 0.90 \dots 0.93$). Comparative analysis showed that the model based on extreme gradient boosting demonstrates the best performance with a determination coefficient (R^2) of 0.93, root mean square error (RMSE) of 2.34, and mean absolute error (MAE) of 1.81. The possibility of applying machine learning methods as an effective tool for predicting crop rotation productivity has been demonstrated.

Keywords: crop rotation, crop rotation productivity, forecast, forecasting, machine learning, artificial intelligence

Acknowledgments. The authors express their gratitude to N.V. Vasilyeva (Lead Researcher of the Fertility Laboratory of SFSCA RAS, CSc (Ag) and G.M. Zakharov (Head of the Crop Rotation Laboratory of SFSCA RAS, CSc (Ag)) for providing information on field experiments with crop rotations.

For citation: Kalichkin V.K., Fedorov D.S., Maksimovich K.Y., Riksen V.S. Machine learning in predicting crop rotation productivity. *Agricultural Engineering (Moscow)*. 2025;27(1):41-52 (In Russ.). <https://doi.org/10.26897/2687-1149-2025-1-41-52>

Введение

Севооборот играет центральную роль в поддержании устойчивых систем земледелия. Это связано с тем, что урожайность каждой культуры, а также количество и тип ресурсов, необходимых для достижения планируемой урожайности, зависят не только от технологии возделывания каждой культуры, но и от долгосрочного воздействия культурооборота на физические, химические и биологические факторы окружающей среды, плодородие и фитосанитарное состояние почв. Эти эффекты в первую очередь определяются сочетанием видов сельскохозяйственных культур, частотой их выращивания и последовательностью, а также остаточными эффектами в межвегетационные периоды. Севооборот или сочетание севооборотов в хозяйстве обуславливают достаточно фиксированные потребности в производственных ресурсах: рабочей силе, технике, складских помещениях и т.д., а также в движении денежных средств. Поэтому выбор возделываемых культур и их размещение по территории землепользования в контексте природных и социально-экономических условий хозяйствования на заданном горизонте планирования лежат в основе управления системой земледелия [1, 2]. В этих решениях сконцентрирована вся сложность, связанная с рациональной организацией территории землепользования и формированием системы земледелия на уровне конкретного хозяйства. Правильно спроектированный и последовательно осуществленный севооборот позволяет не только предотвратить деградацию почв, но и мобилизовать питательные вещества в почве, прервать циклы болезней

и ограничить распространение сорняков и вредителей. Как правило, это отражается на урожайности культур и качестве готовой продукции [3-6].

В отечественном научном дискурсе исследования по проблемам севооборотов связаны в основном с оптимизацией структуры землепользования и территориальной организацией севооборотов с помощью линейного программирования [7-11]. Проектирование севооборотов рассматривалось как статичная концепция, в которой решение принимается только один раз при планировании севооборотов, и чаще всего было основано на процедуре оптимизации нескольких или одного (например, экономического) критериев. Учет неопределенности информации определялся в качестве стохастических факторов или вероятности возникновения, но эта вероятность оставалась статической независимо от динамичной эволюции различных ограничений.

В настоящее время в целях научного обеспечения интеллектуального земледелия возникает необходимость усилить исследования по применению машинного обучения, которое является подмножеством искусственного интеллекта. С помощью традиционных методов обработки данных невозможно удовлетворить постоянно растущие требования интеллектуального земледелия, что является важным препятствием для извлечения ценной информации из полевых опытов научных и учебных учреждений. Исследований в этой области недостаточно, а имеющиеся в отечественной научной литературе касаются в основном прогнозирования урожайности сельскохозяйственных

культур с использованием методов глубокого машинного обучения. В качестве примера можно привести работы, в которых используется искусственная нейронная сеть (ИНС) с алгоритмом обратного распространения ошибки [12], ИНС в форме многослойного персептрона MLP [13, 14] и ИНС глубокого обучения со сверточными слоями (СНС) [15]. Статей по использованию алгоритмов машинного обучения для прогнозирования продуктивности севооборотов на основе данных длительных временных рядов полевых опытов в отечественных библиографических базах нами не обнаружено.

В англоязычной литературе исследования по данному направлению представлены довольно широко. В публикациях подчеркивается, что существенные достижения в обработке данных были получены при использовании методов машинного обучения, которые отличаются их способностью обрабатывать различные типы входных данных и решать нелинейные задачи [16, 17]. Кроме того, машинное обучение в отличие от глубокого обучения позволяет проводить его на небольших наборах данных и опирается на вмешательство человека для классификации данных и выделения атрибутов. Машинное обучение функционирует как послушный робот. Закономерности в данных анализируются для дальнейшего их прогноза [18]. Для формирования прогнозных (предиктивных) моделей различных объектов сельского хозяйства используются следующие алгоритмы: дерево решений (DT) [19]; случайный лес (RF) [20]; метод опорных векторов (SVM) [21]; регрессия гауссовского процесса (GPR) [22]; метод К-ближайших соседей (KNN) [23]; ансамблевые методы, включающие в себя различные варианты бустинга [24], и др.

Цель исследований: создание прогнозных моделей продуктивности севооборотов на основе различных алгоритмов машинного обучения и оценки их предиктивной способности.

Материалы и методы

Машинное обучение проводили с использованием данных временных рядов длительных полевых опытов, выполненных в лесостепи Приобья Новосибирской области Сибирским НИИ земледелия и химизации сельского хозяйства СФНЦА РАН в течение 1999-2019 гг. (табл. 1). В опытах изучалось 9 типов севооборотов, ориентированных на производство зерна: А – бессменная пшеница; В – зерновой; С – зерновой с бобовыми (вико-овес на зерно); D – зерновой с масличными (рапс); Е – зернопаровой с яровыми культурами; F – зернопаровой с озимой рожью; G – зерноотравной с озимой рожью и донником; H – зерноотравной (клевер); I – зерноотравной (вико-овес на з/м). Все севообороты реализованы во времени и пространстве в 3-кратной повторности на делянках площадью 475 м², всего экспериментальных полей севооборотов – 104. В расчетах использованы временные ряды продуктивности севооборотов, полученные на трех уровнях применения техногенных средств: 1 – без удобрений и пестицидов; 2 – удобрения + гербициды; 3 – удобрения + фунгицид + гербицид + инсектицид. Для оценки продуктивности севооборотов был произведен перерасчет фактической урожайности возделываемых культур в зерновые единицы (з.ед.) с помощью коэффициентов перевода продукции растениеводства. Атмосферное увлажнение в годы проведения опытов учитывали с помощью стандартизованного индекса осадков (Standardized Precipitation Index – SPI) [25], описываемого формулой:

$$SPI = \frac{p_i - \bar{p}}{p_i}, \quad (1)$$

где p_i – фактическое количество осадков, зафиксированное в определенный период времени i ; \bar{p} – общее среднее значение осадков за весь анализируемый период; p_i – среднее значение осадков за период времени i [25].

Фрагмент таблицы исходных данных, используемых для прогнозирования

Таблица 1

Fragment of the initial data table used for forecasting

Table 1

Уровни интенсификации <i>Intensification levels</i>	SPI	Тип севооборота <i>Crop rotation type</i>	Выход ц з.ед/га площади севооборота <i>Grain output, hwt/crop rotation area, ha</i>
1	-0,96	Бессменная пшеница / <i>Continuous wheat</i>	4,5
1	-0,29	Зернопаровой с яровыми культурами / <i>Grain fallow with spring crops</i>	10,7
2	0,35	Зерновой / <i>Grain</i>	14,0
2	0,32	Зерновой с масличными (рапс) / <i>Grain with oilseeds (rape)</i>	15,0
3	0,66	Зерновой с бобовыми (вико-овес на зерно) / <i>Grain with legumes (vetch-oats for grain)</i>	19,2
3	0,26	Зернопаровой с озимой рожью / <i>Grain fallow with winter rye</i>	25,6

За основной SPI, участвующий в математических расчетах, принят средний показатель атмосферного увлажнения мая-июля за 4 года ротации каждого из проанализированных севооборотов. Для расчетов SPI использованы временные ряды по данным поста метеонаблюдений г. Новосибирска (источником данных был web-ресурс «<http://www.pogodaiklimat.ru>»). Расчеты SPI выполнены в ПО Drought Indices Calculator (DrinC) с открытым исходным кодом (<https://drought-software.com>).

Выбор методов статистического анализа данных был обусловлен особенностями структуры и характером распределения исходных данных: несоответствие модели нормальному закону распределения; сравнительно небольшой объем выборки; наличие дискретных и континуальных факторов, сложные нелинейные корреляционные связи. Исходный набор данных включал в себя 462 наблюдения, которые были разделены на обучающую и тестовую выборки в соотношении 80 и 20% соответственно.

Для предварительной обработки данных и выявления наиболее значимых признаков применяли методы непараметрической статистики. В наборе данных присутствуют категориальные (тип севооборота, уровень интенсификации) и количественные (стандартизированный индекс осадков SPI, продуктивность севооборотов) характеристики. Связь типа севооборота с его продуктивностью оценивалась с помощью критерия Краскела-Уоллиса (H-test). Для оценки связи продуктивности севооборота с количественными характеристиками применяли ранговый коэффициент корреляции Спирмена и дисперсионный анализ (ANOVA). Уровень значимости (p-значение) оценивали на уровне 0,05. В ходе исследований построены модели, включающие в себя дерево решений (алгоритм Classification and Regression Tree – CART), случайный лес (алгоритм Random Forest), бутстрэп-агрегирование (алгоритм Bagging), градиентный бустинг (алгоритм Gradient boosting) и его вариацию – экстремальный градиентный бустинг (алгоритм EXtreme Gradient Boosting – XGBoost), а также искусственная нейронная сеть (ANN).

Дерево решений – это алгоритм машинного обучения, который строит модель классификации или регрессии в виде древовидной структуры, где каждый внутренний узел представляет собой условие на значения признаков, ветви отражают результат проверки, а листья содержат выходные значения [26]. Общий вид модели дерева решений можно представить следующей формулой:

$$f(x) = \sum_{m=1}^M c_m \mathbb{I}(x \in R_m), \quad (2)$$

где x – входной вектор; R_m – m -й регион (лист дерева); c_m – константа, соответствующая региону R_m ; \mathbb{I} – индикаторная функция; M – общее число листьев [26].

Случайный лес – это ансамблевый метод машинного обучения, который строит множество деревьев решений на случайных подвыборках обучающего набора данных и усредняет результаты для повышения точности прогнозирования и контроля переобучения [27]. Общая формула модели случайного леса –

$$f(x) = \frac{1}{N} \sum_{n=1}^N f_n(x), \quad (3)$$

где $f_n(x)$ – n -е дерево решений; N – общее число деревьев в ансамбле [27].

Бутстрэп-агрегирование – это ансамблевый метод машинного обучения, который комбинирует множество моделей, обученных на случайных подвыборках исходного набора данных с помощью бутстрэпа, и усредняет их прогнозы для повышения точности и стабильности [28]. Общий вид модели бутстрэп-агрегирования:

$$f(x) = \frac{1}{M} \sum_{m=1}^M f_m(x), \quad (4)$$

где $f_m(x)$ – модель, обученная на n -й бутстрэп-выборке; M – общее число моделей в ансамбле [28].

Градиентный бустинг – это ансамблевый метод машинного обучения, который строит модель в виде ансамбля слабых предикторов, обычно деревьев решений, последовательно добавляя их к ансамблю и корректируя ошибки предыдущих моделей с помощью градиентного спуска [29].

Экстремальный градиентный бустинг – это усовершенствованный алгоритм машинного обучения, представляющий собой оптимизированную распределенную реализацию градиентного бустинга [29]. XGBoost эффективно применяет принципы параллельного построения деревьев решений (известного также как GBDT или GBM) в рамках парадигмы градиентного бустинга. Общая формула модели градиентного бустинга:

$$f(x) = \sum_{m=1}^M \gamma_m h_m(x), \quad (5)$$

где $h_m(x)$ – базовый предиктор (дерево решений); γ_m – коэффициент усадки, регулирующий вклад каждого дерева; M – число итераций бустинга [29].

Искусственная нейронная сеть – это модель машинного обучения, состоящая из взаимосвязанных узлов (нейронов), организованных в слои, которая может обучаться на данных для решения задач регрессии и классификации [30]. Общий вид модели искусственной нейронной сети –

$$y = f\left(\sum_{n=1}^N w_n x_n + b\right), \quad (6)$$

где x_n – входные значения; w_n – веса; b – смещение; f – функция активации; N – число входов [30].

Оценку производительности моделей проводили с использованием трех статистических параметров: MAE, RMSE и R^2 соответственно, используя в качестве известных параметров: y – истинное значение; \hat{y} – прогнозируемое значение; \bar{y} – среднее значение истинных значений; N – размер выборки.

MAE (Mean Absolute Error) – средняя абсолютная ошибка. Эта метрика измеряет среднее абсолютное отклонение прогнозируемых значений от истинных:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (7)$$

RMSE (Root Mean Squared Error) – корень из среднеквадратичной ошибки. Это метрика, которая измеряет стандартное отклонение ошибок между прогнозируемыми и истинными значениями:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}. \quad (8)$$

R^2 (R Squared) – коэффициент детерминации. Эта метрика измеряет долю дисперсии зависимой переменной, которая объясняется независимыми переменными. Он может принимать значения от 0 до 1, где 1 означает идеальное совпадение прогнозов и истинных значений:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}. \quad (9)$$

Статистические расчеты и графические построения выполнены средствами языка программирования Python в интерактивном блокноте Jupyter, а также средствами языка программирования для статистической обработки данных и работы с графикой R в интегрированной среде разработки R-Studio. Модели продуктивности севооборотов реализованы в авторской программе Crop Yield Analysis & Forecast (CYAF) [31, 32].

Результаты и их обсуждение

Концептуализация подхода «Прогнозирование продуктивности севооборотов», представляющая детальную информацию из 8 основных шагов от этапа подготовки данных – формирования архитектуры, обучения и тестирования моделей – к прогнозированию продуктивности севооборотов, представлена на рисунке 1.

Ключевыми этапами исследований являются:

1. Сбор и предобработка данных включая формирование временных рядов продуктивности севооборотов и атмосферного увлажнения (SPI).

2. Анализ данных и инжиниринг признаков для выявления значимых предикторов и формирования репрезентативного набора данных.

3. Выбор и оптимизация моделей машинного обучения включая линейные модели, ансамблевые методы и нейронные сети.

4. Оценка и интерпретация результатов с использованием метрик эффективности (RMSE, MAE, R^2) и методов интерпретации моделей.

Особое внимание уделяется этапу обучения и оптимизации моделей, когда для повышения точности прогнозирования применяются методы кросс-валидации и подбора гиперпараметров. Итеративный характер процесса обеспечивается обратными связями между этапами, позволяя корректировать выбор признаков и настройки моделей на основе полученных результатов. Заключительные этапы включают в себя развертывание модели и ее непрерывное улучшение, что обеспечивает практическую применимость и адаптивность подхода к изменяющимся агроэкологическим условиям.

На основе эвристического подхода были выделены признаки, определяющие продуктивность севооборотов. Предварительный анализ данных выявил статистически значимую связь ($p < 0,05$) между продуктивностью севооборотов и тремя факторами: атмосферным увлажнением, выраженным через SPI; уровнем интенсификации технологии возделывания; типом севооборота. Графики зависимости демонстрируют характер этих взаимосвязей (рис. 2).

Анализ графиков зависимости продуктивности севооборотов от ключевых факторов в лесостепи Приобья за период 1999-2019 гг. выявил следующие закономерности:

1. Влияние атмосферного увлажнения (рис. 2а). Наблюдается линейная зависимость между стандартизированным индексом осадков (SPI) и продуктивностью севооборотов ($p < 0,05$), что указывает на умеренную положительную связь двух переменных. С увеличением SPI наблюдается пропорциональный рост продуктивности севооборотов, что подчеркивает лимитирующую роль влагообеспеченности для условий лесостепной зоны юга Западной Сибири.

2. Влияние уровня интенсификации (рис. 2б). График демонстрирует устойчивый рост продуктивности при переходе от низкого к высокому уровню использования техногенных ресурсов (интенсификация земледелия). Это свидетельствует о значительном потенциале повышения урожайности культур в севообороте за счет интенсификации агротехнологий.

3. Влияние типа севооборота (рис. 2в). Наблюдаются существенные различия в продуктивности между различными типами севооборотов. Наибольшая продуктивность получена при реализации зерноотравных севооборотов, особенно с включением клевера (тип H) и комбинации озимой ржи с донником (тип G). За ними следуют зерноотравной с вико-овсом на з/м (I), зерновой с бобовыми (C), зернопаровой с озимой рожью (F), зерновой (B),

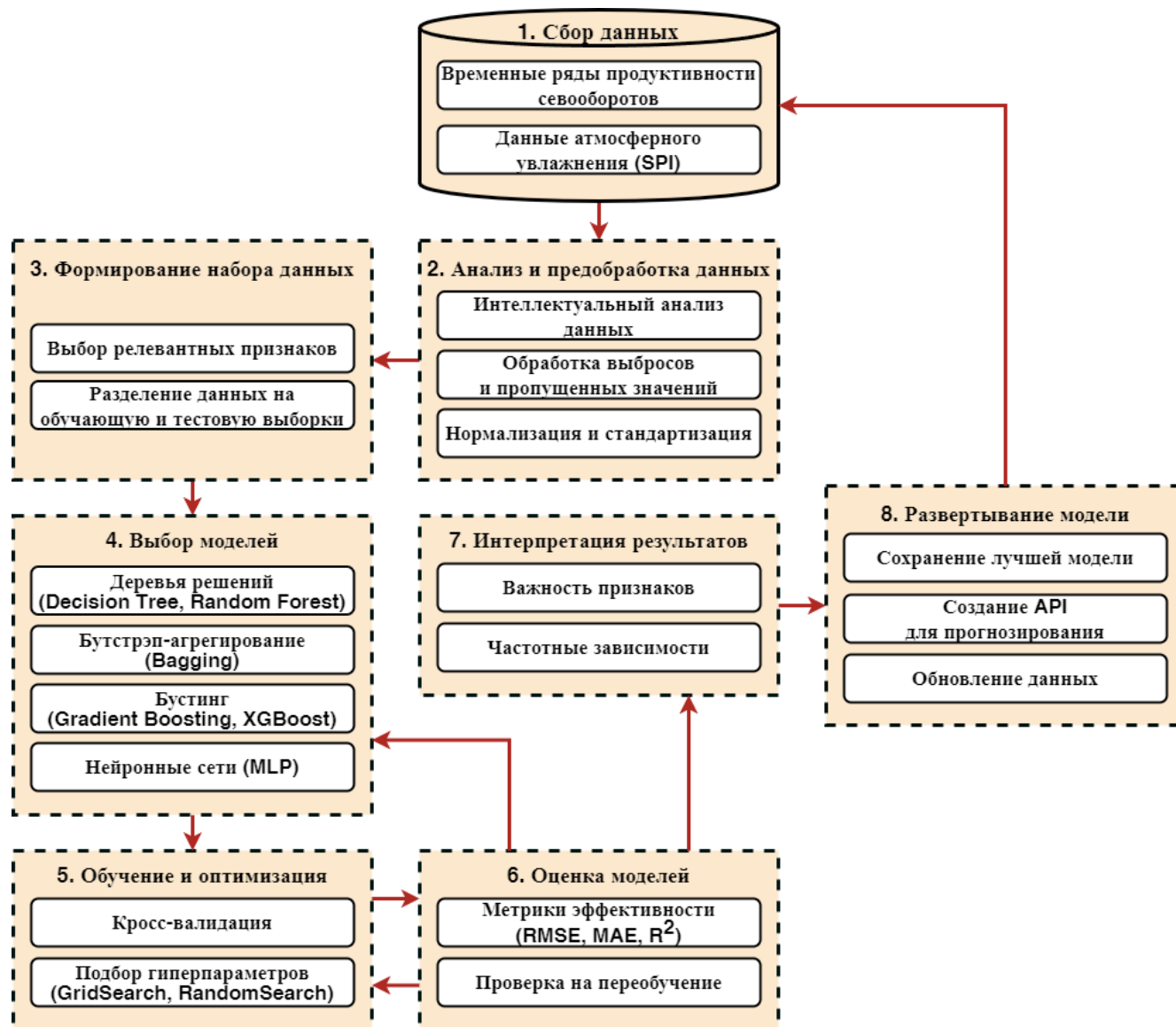


Рис. 1. Концептуальная основа исследования, объясняющая методологию проектирования предиктивных моделей для прогнозирования продуктивности севооборотов с использованием временных рядов данных многолетних стационаров СФНЦА РАН и методов машинного обучения

Fig. 1. Conceptual framework of the study, explaining the methodology for designing predictive models to forecast crop rotation yield using time series data from long-term stationary experiments of SFSCA RAS and machine learning methods

зернопаровой с яровыми культурами (Е), зерновой с масличными (D), и наименее продуктивным оказалось бессменное возделывание пшеницы (А).

Высокая продуктивность зернотравяных севооборотов может быть объяснена благоприятным воздействием многолетних трав на структуру и плодородие почвы, а также обогащением почвы азотом за счет симбиотической азотфиксации на бобовых. Низкая эффективность бессменного возделывания пшеницы подтверждает необходимость научно обоснованного чередования культур.

Метрика важности переменных модели случайного леса, измеряющая среднее снижение точности при перестановке входных признаков [27], показала значимые результаты для всех выбранных предикторов (рис. 3).

Результаты анализа важности отдельных признаков с помощью алгоритма RF показали, что атмосферное увлажнение, выраженное через «SPI», обладает наибольшим значением чистоты узла. Это свидетельствует о его ключевой роли при прогнозировании продуктивности севооборотов. Переменные «уровень интенсификации» (характеризующий комплексность использования техногенных ресурсов) и «тип севооборота» (зерновые, зернопаровые, зернотравяные и др.) также являлись статистически значимыми (во всех случаях $p < 0,05$), но их вклад в увеличение чистоты узлов несколько меньший, чем у «SPI». Результаты предварительной работы были использованы в дальнейшем при построении предиктивных моделей продуктивности севооборотов.

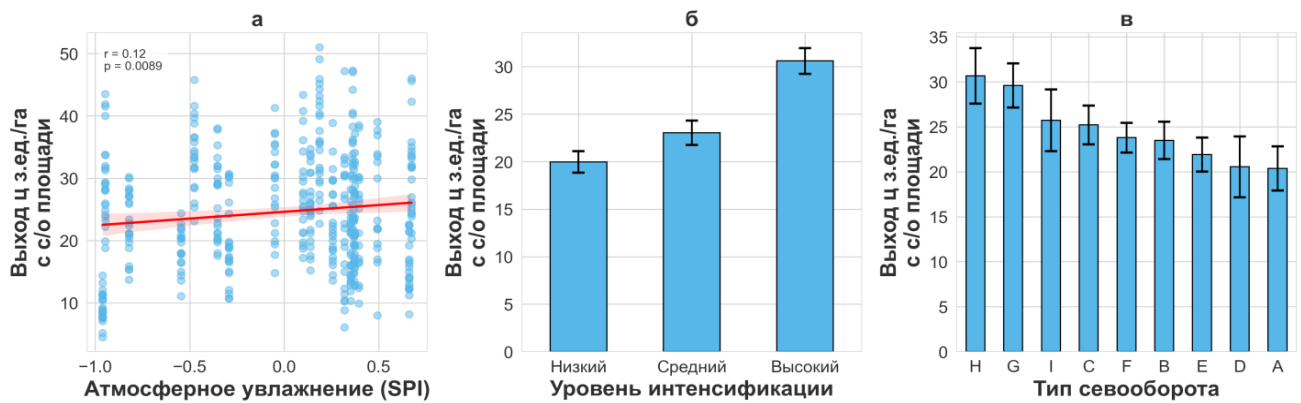


Рис. 2. Зависимость показателя продуктивности севооборотов от атмосферного увлажнения (а), уровня интенсификации (б), типа севооборота (в):

А – бесменная пшеница; В – зерновой; С – зерновой с бобовыми (вико-овес на зерно); D – зерновой с масличными (рапс); E – зернопаровой с яровыми культурами; F – зернопаровой с озимой рожью; G – зернотравяной с озимой рожью и донником; H – зернотравяной (клевер); I – зернотравяной (вико-овес на 3/м)

Fig. 2. Graphical relationship between crop rotation productivity indicators and selected predictors – atmospheric moisture (a), intensification level (b), type of crop rotation (c):

A – continuous wheat, B – grain, C – grain with legumes (vetch-oats for grain), D – grain with oilseeds (rapeseed), E – grain-fallow with spring crops, F – grain-fallow with winter rye, G – grain-grass with winter rye and melilot, H – grain-grass (clover), I – grain-grass (vetch-oats for green mass)

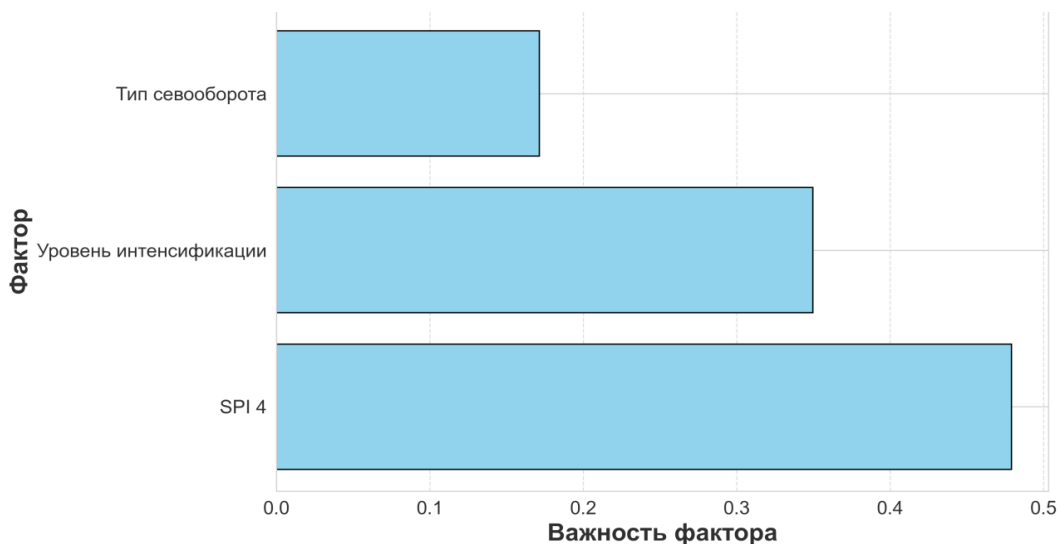


Рис. 3. Результаты анализа важности отдельных признаков с помощью алгоритма RF

Fig. 3. Analysis results showing the importance of individual factors (based on the RF algorithm)

Оценка точности и предиктивной способности моделей осуществлена с помощью основных статистических метрик (табл. 2).

Установлено, что модель XGBoost превзошла другие алгоритмы прогнозирования. Для нее характерны наименьшие значения MAE и RMSE, а также наибольший показатель R², что демонстрирует ее высокую эффективность в прогнозировании продуктивности севооборотов. Значения MAE и RMSE, полученные в моделях CART, RF и Bagging, не имели статистически значимых различий (H-test > 0,05). Значения R² для этих моделей находились в диапазоне 0,80...0,85.

На основе анализа значений MAE, RMSE и R² модель CART характеризовалась наиболее низкой

способностью к прогнозированию продуктивности севооборотов. В сравнении с высокой эффективностью ансамблевых и нейросетевых методов (Gradient Boosting, XGBoost, ANN) модель CART продемонстрировала относительно высокие значения MAE и RMSE. Это указывает на практическую склонность данного алгоритма к переобучению, а также на ограниченную способность к генерализации результатов в контексте изменчивости данных. Модель искусственной нейронной сети характеризовалась сопоставимыми значениями с ансамблевыми методами по показателям MAE и RMSE, но не достигла ожидаемых показателей прогностической точности, что обусловлено, видимо, ограниченностью объема обучающих данных.

Значения продуктивности севооборотов для различных предиктивных моделей, полученные на основе применения алгоритмов RF, CART и Bagging, имеют наибольшую вариативность и отражают менее стабильное прогнозирование (рис. 4). Модели, построенные с помощью алгоритмов Gradient Boosting,

XGBoost и нейронная сеть ANN, продемонстрировали максимально близкое совпадение значений с линией идеального соответствия, что указывает на их высокую прогностическую способность.

По результатам исследования различных алгоритмов машинного обучения установлено явное

Таблица 2

Результаты оценки точности предиктивных моделей

Table 2

The results of evaluating the accuracy of predictive models

Модель / Model	Метрики оценки точности моделей Metrics for assessing model accuracy		
	MAE	RMSE	R ²
Дерево решений (CART) / Decision Tree (CART)	3,15	3,99	0,80
Случайный лес (RF) / Random Forest (RF)	2,91	3,86	0,81
Бутстрэп-агрегирование (Bagging) / Bootstrap Aggregation (Bagging)	2,70	3,47	0,85
Градиентный бустинг (Gradient Boosting) / Gradient Boosting	1,87	2,43	0,92
Экстремальный градиентный бустинг (XGBoost) / Extreme Gradient Boosting (XGBoost)	1,81	2,34	0,93
Искусственная нейронная сеть (ANN) / Artificial Neural Network (ANN)	1,98	2,88	0,90

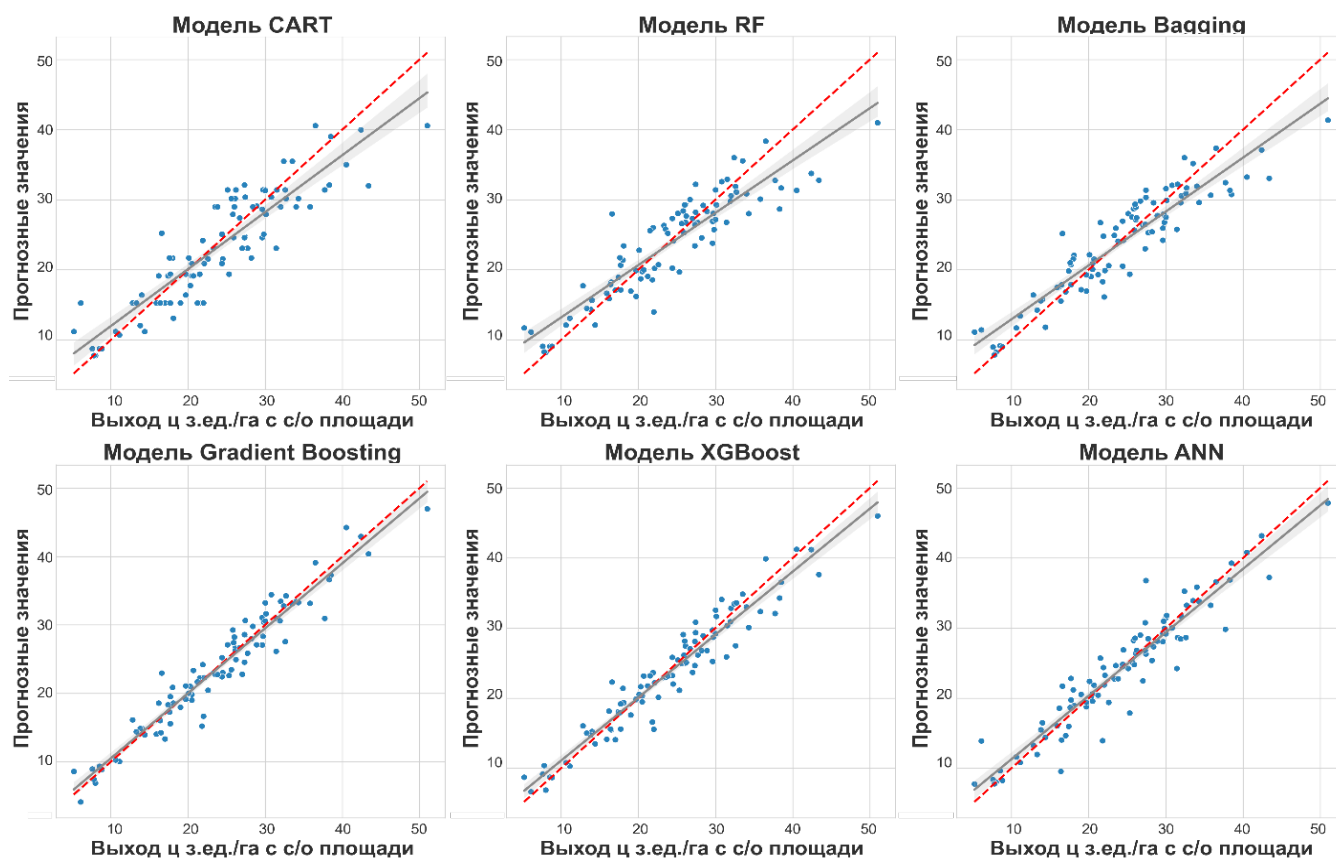


Рис. 4. Диаграммы рассеяния прогнозных и фактических значений продуктивности, полученные с помощью различных алгоритмов машинного обучения:

точки – фактические (ось X) и прогнозные (ось Y) значения продуктивности севооборотов; диагональная линия (y = x) – случаи, когда прогнозные значения в точности совпадают с фактическими; серая область вокруг линии тренда – 95%-ный доверительный интервал

Fig. 4. Scatter plots of predicted and actual productivity values obtained using various machine learning algorithms: points – actual (X-axis) and predicted (Y-axis) values of crop rotation productivity; diagonal line on the graph (y = x line) – cases where predicted values exactly match actual values; gray area around the trend line – 95% confidence interval

преимущество ансамблевых (Gradient Boosting, XGBoost) и нейросетевой (ANN) моделей в общей точности прогнозирования продуктивности севооборотов. Преимущество данных моделей отмечалось и в других исследованиях – например, при прогнозировании урожайности кукурузы и ее гибридов [33, 34], а в работе [35] особо подчеркнута эффективность модели регрессии XGBoost с точностью прогноза 93,8%. В наших исследованиях относительно высокие прогностические показатели моделей ANN, Gradient Boosting и XGBoost были достигнуты на основе использования наблюдений многолетних полевых стационарных опытов, однако данные модели обычно классифицируются как «черный ящик», поскольку способы получения прогнозов не являются непосредственно очевидными и им обычно не хватает прозрачности и интерпретируемости [36]. Поэтому для решения проблемы интерпретируемости моделей машинного обучения в последнее время развиваются искусственный интеллект и различные инструменты, а именно интерпретируемое машинное обучение [37].

В мировой науке по исследованию применения машинного обучения (в последнее время – преимущественно глубокого) для прогнозирования различных предметных областей сельского хозяйства осуществляется практика включения широкого комплекса информации о состоянии растительного покрова и свойствах почвы, фенологических наблюдениях, агроклиматических условиях местности и др., что позволяет повысить точность предиктивных моделей до относительно высоких значений [38, 39]. Модели с включением детализированной информации о рабочем участке, культурах и окружающей среде с использованием данных временных рядов полевых опытов также позволяют повысить предсказательную способность алгоритмов машинного обучения. Поэтому многоуровневая параметризация предиктивных моделей, повышение точности предсказаний и их локальная

адаптация требуют большего количества агроэкологических данных о возделываемых культурах и окружающей среде, которые можно получить с помощью различных методов мониторинга на конкретном пространственном объекте [40]. При этом нативный, проксимальный и дистанционный мониторинг становятся безальтернативным подходом в контексте использования методов искусственного интеллекта для формирования цифровых систем поддержки принятия решений в земледелии и масштабирования пространства.

Выводы

В результате реализации 6 моделей машинного обучения для прогнозирования продуктивности севооборотов оценена точность их предиктивной способности. Модели продуктивности севооборотов на основе алгоритмов ANN, Gradient Boosting, XGBoost обладали наиболее высокими прогностическими способностями в зависимости от типа севооборота, атмосферного увлажнения и уровня интенсификации технологии возделывания ($R^2 = 0,90 \dots 0,93$). Высокая эффективность применения ансамблевых и нейросетевых методов объясняется их способностью решать нелинейные задачи взаимодействия, включенные в модели признаков. Модель на основе экстремального градиентного бустинга демонстрирует наилучшие показатели с коэффициентом детерминации (R^2) 0,93, среднеквадратичной ошибкой (RMSE) 2,34 и средней абсолютной ошибкой (MAE) 1,81.

Продолжением исследований могут стать разработка и адаптация гибридных прогнозных моделей продуктивности севооборотов, которые будут базироваться на интеграции многоуровневой параметризации экспертных знаний в земледелии и методов машинного обучения. Мониторинговые исследования агрофитоценозов могут стать безальтернативным решением проблемы нехватки данных в области сельскохозяйственной науки и практики.

Список источников

1. Кирюшин В.И. Управление плодородием почв и продуктивностью агроценозов в адаптивно-ландшафтных системах земледелия // Почвоведение. 2019. № 9. С. 1130-1139. <https://doi.org/10.1134/S0032180X19070062>
2. Степных Н.В., Нестерова Е.В., Заргарян А.М. и др. Стратегическое значение диверсификации растениеводства // Земледелие. 2022. № 2. С. 7-13. EDN: DNVBGB
3. Лошаков В.Г. Развитие учения о севообороте в РГАУ-МСХА им. К.А. Тимирязева // Земледелие. 2017. № 2. С. 3-9. EDN: YLMWMD
4. Завьялова Н.Е., Фомин Д.С., Тетерлев И.С. Влияние севооборотов и бессменных посевов на агрохимические свойства и азотный режим дерново-подзолистой почвы Предуралья // Агрохимия. 2019. № 1. С. 5-10. <https://doi.org/10.1134/S0002188119010162>

References

1. Kiryushin V.I. The management of soil fertility and productivity of agrocenoses in landscape adaptive farming systems. *Pochvovedenie*. 2019;9:1137-1139. (In Russ.) <https://doi.org/10.1134/S0032180X19070062>
2. Stepanykh N.V., Nesterova E.V., Zargaryan A.M., Kopylova S.A. Strategic importance of crop production diversification. *Zemledelie*. 2022;2:7-13. (In Russ.)
3. Лошаков В.Г. Development of the doctrine about crop rotation in RSAU-MTAA. *Zemledelie*. 2017;2:3-9. (In Russ.)
4. Zavyalova N.E., Fomin D.S., Teterlev I.S. Effect of crop rotations and monoculture on agrochemical properties and nitrogen regime of sod-podzolic soil of the Cis-Ural region. *Agrohimiya*. 2019;1:5-10. (In Russ.) <https://doi.org/10.1134/S0002188119010162>
5. Kurdyukova O.N. Weed infestation of crops and short rotating crops productivity in the steppe zone.

5. Курдюкова О.Н. Засоренность посевов и продуктивность короткороотационных севооборотов степной зоны // Вестник Красноярского государственного аграрного университета. 2022. № 7 (184). С. 69-76. <https://doi.org/10.36718/1819-4036-2022-7-69-76>
6. Акименко А.С., Свиридов В.И., Дудкина Т.А. и др. Обоснование направления интенсификации производства зерна в севооборотах лесостепи Центрального Черноземья // Земледелие. 2023. № 1. С. 3-7. <https://doi.org/10.24412/0044-3913-2023-1-3-7>
7. Макарова Е.И. Экономико-математическое моделирование оптимизации структуры посевных площадей аграрных предприятий // Вестник Воронежского государственного аграрного университета. 2015. № 4-2. С. 174-178. EDN: VNTSYJ
8. Нуретдинов И.Г., Нуретдинова Ю.В., Хасянов О.Р. Оптимизация структуры посевных площадей сельскохозяйственных культур методом экономико-математического моделирования: практический аспект // Вестник Казанского государственного аграрного университета. 2015. Т. 10, № 4. С. 24-27. EDN: VLQMFT
9. Свиридов В.И. Методологические и методические аспекты проектирования оптимальной структуры посевных площадей в условиях перехода к адаптивно-ландшафтному земледелию // Вестник Курской государственной сельскохозяйственной академии. 2018. № 2. С. 4-10. EDN: YVOILT
10. Никитина К.Р. Оптимизация структуры севооборота с применением метода квадратичного программирования // Экономика. Управление. Инновации. 2019. № 2 (6). С. 92-96. EDN: KMMBUK
11. Апажев А.К., Шекихачев Ю.А. Оптимизация функционирования сельскохозяйственных производственных систем // Известия Кабардино-Балкарского государственного аграрного университета им. В.М. Кокова. 2022. № 1 (35). С. 81-89. <https://doi.org/10.55196/2411-3492-2022-1-35-81-89>
12. Дудко Ю.В., Маринкин Е.Б., Владимирова Т.М. Применение нейросетевых технологий в предсказании плодородия почв с зернобобовыми культурами и прогнозировании их урожайности // Вопросы науки: инноватика, техника и технологии. 2019. № 1. С. 67-72. EDN: YXPXCH
13. Рогачев А.Ф. Системный анализ и прогнозирование временных рядов урожайности на основе автокорреляционных функций и нейросетевых технологий // Известия Нижневолжского агроуниверситетского комплекса: Наука и высшее профессиональное образование. 2018. № 3 (51). С. 309-316. EDN: VOQSWI
14. Рогачев А.Ф., Мелихова Е.В. Обоснование алгоритмов и инструментария для нейросетевого прогнозирования урожайности агрокультур с использованием ретроспективных данных // Известия Нижневолжского агроуниверситетского комплекса: наука и высшее профессиональное образование. 2020. № 1 (57). С. 290-302. <https://doi.org/10.32786/2071-9485-2020-01-29>
15. Лебедь Н.И., Токарев К.Е., Нехорошев Д.Д. и др. Разработка методов управления продукционными процессами фитоагроценозов в условиях точного земледелия с использованием гибридных автоматизированных систем и алгоритмов искусственного интеллекта // Международный сельскохозяйственный журнал. 2024. № 2. С. 201-204. EDN: WMQCNC
16. Benos L., Tagarakis A.C., Dolias G. et al. Machine learning in agriculture: A comprehensive updated review. *Sensors*. 2021;21(11):3758. <https://doi.org/10.3390/s21113758>
17. Attri I., Awasthi L.K., Sharma T.P. Machine learning in agriculture: a review of crop management applications. *Multimedia Tools and Applications*. 2024;83(5):12875-12915. <https://doi.org/10.1007/s11042-023-16105-2>
18. Taye M.M. Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers*. 2023;12(5):91. <https://doi.org/10.3390/computers12050091>
19. Kalichkin V.K., Alsova O.K., Maksimovich K.Y. Application of the decision tree method for predicting the yield of spring wheat. *IOP Conference Series: Earth and Environmental Science*. IOP Publishing. 2021;839(3):032042. <https://doi.org/10.1088/1755-1315/839/3/032042>
20. Elavarasan D., Vincent P.M. D.R. A reinforced random forest model for enhanced crop yield prediction by integrating agrarian *Bulletin of KSAU*. 2022;7;184:69-76. (In Russ.). <https://doi.org/10.36718/1819-4036-2022-7-69-76>
6. Akimenko A.S., Sviridov V.I., Dudkina T.A., Vavin V.G., Dolgoplova N.V. Substantiation of the direction of grain production intensification in crop rotations of the forest-steppe of the Central Chernozem region. *Zemledelie*. 2023;1:3-7. <https://doi.org/10.24412/0044-3913-2023-1-3-7>
7. Makarova E.I. Economic and mathematical modeling of structural optimization of cropping system of agricultural enterprises. *Vestnik of Voronezh State Agrarian University*. 2015;4-2:174-178. (In Russ.).
8. Nuretdinov I.G., Nuretdinova Yu.V., Khasyanov O.R. Optimization of sown area structure of crops by economic and mathematical modeling: practical aspect. *Vestnik of the Kazan State Agrarian University*. 2015;10(4):24-27. (In Russ.).
9. Sviridov V.I. Methodological and methodological aspects of designing the optimal structure of acreage in the transition to adaptive landscape farming. *Vestnik Kurskoy Gosudarstvennoy Selskhozaystvennoy Akademii*. 2018;2:4-10. (In Russ.).
10. Nikitina K.R. Crop rotation optimization using the quadratic programming approach. *Economics. Management. Innovations*. 2019;2:92-96. (In Russ.).
11. Apazhev A.K., Shekikhachev Yu.A. Optimization of the functioning of agricultural production systems. *Izvestiya Kabardino-Balkarskogo gosudarstvennogo agrarnogo universiteta im. V.M. Kokova = Izvestiya of the Kabardino-Balkarian State Agrarian University named after V.M. Kokov*. 2022;1:81-89. <https://doi.org/10.55196/2411-3492-2022-1-35-81-89>
12. Dudko Yu.V., Marinkin E.B., Vladimirova T.M. Use of neural network technologies in predicting soil fertility when growing leguminous crops and predicting their yield. *Voprosy Nauki: Innovatika, Tekhnika i Tekhnologii*. 2019;1:67-72. (In Russ.).
13. Rogachev A.F. System analysis and forecasting of yield time series based on autocorrelation functions and neural network technologies. *Proceedings of Lower Volga Agro-University Complex: Science and Higher Education*. 2018;3:309-316. (In Russ.).
14. Rogachev A.F., Melikhova, E.V. Justification of algorithms and tools for neural network forecasting of agricultural productivity using retrospective data. *Proc. of the Lower Volga AgroUniversity Comp*. 2020;1:290-302. (In Russ.). <https://doi.org/10.32786/2071-9485-2020-01-29>
15. Lebed N.I., Tokarev K.E., Nekhoroshev D.D., Storozhakov S.Yu., Popov A.Yu. Development of methods for managing the production processes of phyto-agrocenoses in precision farming using hybrid automated systems and artificial intelligence algorithms. *Mezhdunarodnyi Sel'skhozaystvennyi Zhurnal*. 2024;2:201-204. (In Russ.).
16. Benos L., Tagarakis A.C., Dolias G. et al. Machine learning in agriculture: A comprehensive updated review. *Sensors*. 2021;21(11):3758. <https://doi.org/10.3390/s21113758>
17. Attri I., Awasthi L.K., Sharma T.P. Machine learning in agriculture: a review of crop management applications. *Multimedia Tools and Applications*. 2024;83(5):12875-12915. <https://doi.org/10.1007/s11042-023-16105-2>
18. Taye M.M. Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers*. 2023;12(5):91. <https://doi.org/10.3390/computers12050091>
19. Kalichkin V.K., Alsova O.K., Maksimovich K.Y. Application of the decision tree method for predicting the yield of spring wheat. *IOP Conference Series: Earth and Environmental Science*. IOP Publishing. 2021;839(3):032042. <https://doi.org/10.1088/1755-1315/839/3/032042>
20. Elavarasan D., Vincent P.M. D.R. A reinforced random forest model for enhanced crop yield prediction by integrating agrarian

19. Kalichkin V.K., Alsova O.K., Maksimovich K.Y. Application of the decision tree method for predicting the yield of spring wheat. *IOP Conference Series: Earth and Environmental Science*. IOP Publishing, 2021;839(3):032042. <https://doi.org/10.1088/1755-1315/839/3/032042>
20. Elavarasan D., Vincent P.M. D.R. A reinforced random forest model for enhanced crop yield prediction by integrating agrarian parameters. *Journal of Ambient Intelligence and Humanized Computing*, 2021;12:10009-10022. <https://doi.org/10.1007/s12652-020-02752-y>
21. Kok Z.H., Shariff A.R.M., Alfatni M.S.M., Khairunni-za-Bejo S. Support vector machine in precision agriculture: a review. *Computers and Electronics in Agriculture*. 2021;191:106546. <https://doi.org/10.1016/j.compag.2021.106546>
22. Hegedus P.B., Maxwell B.D., Mieno T. Assessing performance of empirical models for forecasting crop responses to variable fertilizer rates using on-farm precision experimentation. *Precision Agriculture*. 2023;24:677-704. <https://doi.org/10.1007/s11119-022-09968-2>
23. Gopal M.P.S., Bhargavi R. Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms. *Applied Artificial Intelligence*. 2019;33(7):621-642. <https://doi.org/10.1080/08839514.2019.1592343>
24. Han J., Zhang Z., Cao J. et al. Prediction of winter wheat yield based on multi-source data and machine learning in China. *Remote Sensing*. 2020;12(2):236. <https://doi.org/10.3390/rs12020236>
25. Tigkas D., Vangelis H., Tsakiris G. DrinC: a software for drought analysis based on drought indices. *Earth Science Informatics*. 2015;8:697-709. <https://doi.org/10.1007/s12145-014-0178-y>
26. Breiman L., Friedman J., Olshen R.A., Stone C.J. Classification and Regression Trees (1st ed.). Chapman and Hall/CRC. 1984;368. <https://doi.org/10.1201/9781315139470>
27. Breiman L. Random forests. *Machine learning*. 2001;45:5-32. <https://doi.org/10.1023/A:1010933404324>
28. Breiman L. Bagging predictors. *Machine learning*. 1996;24:123-140. <https://doi.org/10.1007/BF00058655>
29. Friedman J.H. Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*. 2001;29(5):1189-1232. <https://doi.org/10.1214/aos/1013203451>
30. Haykin S. Neural networks: A comprehensive foundation. Prentice Hall PTR. Upper Saddle River, NJ. United States. 1994;768.
31. Каличкин В.К., Федоров Д.С., Альсова О.К., Максимович К.Ю. Разработка программы анализа и прогнозирования урожайности сельскохозяйственных культур // Достижения науки и техники АПК. 2022. Т. 36, № 1. С. 51-56. https://doi.org/10.53859/02352451_2022_36_1_51
32. Программа анализа и прогнозирования урожайности сельскохозяйственных культур: Свидетельство о государственной регистрации программы для ЭВМ (CYAF – Crop Yield Analysis & Forecast) № 2021777894 / Д.С. Федоров, О.К. Альсова, В.К. Каличкин, К.Ю. Максимович; 08.11.2021.
33. Du Z., Yang L., Zhang D. et al. Corn variable-rate seeding decision based on gradient boosting decision tree model. *Computers and Electronics in Agriculture*. 2022;198:107025. <https://doi.org/10.1016/j.compag.2022.107025>
34. Sarijaloo F.B., Porta M., Taslimi B., Pardalos P.M. Yield performance estimation of corn hybrids using machine learning algorithms. *Artificial Intelligence in Agriculture*. 2021;5:82-89. <https://doi.org/10.1016/j.aiia.2021.05.001>
35. Awasthi P., Mishra S., Gupta N. Performance Assessment of Machine Learning Techniques for Corn Yield Prediction. *International Conference on Advanced Network Technologies and Intelligent Computing*. Cham: Springer Nature Switzerland. 2022;320-335. https://doi.org/10.1007/978-3-031-28183-9_23
36. Wolanin A., Mateo-García G., Camps-Valls G. et al. Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. *Environmental Research Letters*. 2020;15(2):024019. <https://doi.org/10.1088/1748-9326/ab68ac>
37. Hassija V., Chamola V., Mahapatra A. et al. Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*. 2023;16:45-74. <https://doi.org/10.1007/s12559-023-10179-8>
38. Wang X., Huang J., Feng Q., Yin D. Winter wheat yield prediction at county level and uncertainty analysis parameters. *Journal of Ambient Intelligence and Humanized Computing*. 2021;12:10009-10022. <https://doi.org/10.1007/s12652-020-02752-y>
21. Kok Z.H., Shariff A.R.M., Alfatni M.S.M., Khairunni-za-Bejo S. Support vector machine in precision agriculture: a review. *Computers and Electronics in Agriculture*. 2021;191:106546. <https://doi.org/10.1016/j.compag.2021.106546>
22. Hegedus P.B., Maxwell B.D., Mieno T. Assessing performance of empirical models for forecasting crop responses to variable fertilizer rates using on-farm precision experimentation. *Precision Agriculture*. 2023;24:677-704. <https://doi.org/10.1007/s11119-022-09968-2>
23. Gopal M.P.S., Bhargavi R. Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms. *Applied Artificial Intelligence*. 2019;33(7):621-642. <https://doi.org/10.1080/08839514.2019.1592343>
24. Han J., Zhang Z., Cao J. et al. Prediction of winter wheat yield based on multi-source data and machine learning in China. *Remote Sensing*. 2020;12(2):236. <https://doi.org/10.3390/rs12020236>
25. Tigkas D., Vangelis H., Tsakiris G. DrinC: a software for drought analysis based on drought indices. *Earth Science Informatics*. 2015;8:697-709. <https://doi.org/10.1007/s12145-014-0178-y>
26. Breiman L., Friedman J., Olshen R.A., Stone C.J. Classification and Regression Trees (1st ed.). Chapman and Hall/CRC. 1984;368. <https://doi.org/10.1201/9781315139470>
27. Breiman L. Random forests. *Machine learning*. 2001;45:5-32. <https://doi.org/10.1023/A:1010933404324>
28. Breiman L. Bagging predictors. *Machine learning*. 1996;24:123-140. <https://doi.org/10.1007/BF00058655>
29. Friedman J.H. Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*. 2001;29(5):1189-1232. <https://doi.org/10.1214/aos/1013203451>
30. Haykin S. Neural networks: A comprehensive foundation. Prentice Hall PTR. Upper Saddle River, NJ. United States. 1994;768.
31. Kalichkin V.K., Fedorov D.S., Alsova O.K., Maksimovich K.Yu. Development of software for analyzing and forecasting crop yields. *Achievements of Science and Technology in Agro-Industrial Complex*. 2022;36(1):51-56. (In Russ.). https://doi.org/10.53859/02352451_2022_36_1_51
32. Crop Yield Analysis & Forecast (CYAF) software. Fedorov D.S., Alsova O.K., Kalichkin V.K., Maksimovich K.Yu. Certificate of state registration of software No. 2021777894, 08.11.2021. (In Russ.).
33. Du Z., Yang L., Zhang D. et al. Corn variable-rate seeding decision based on gradient boosting decision tree model. *Computers and Electronics in Agriculture*. 2022;198:107025. <https://doi.org/10.1016/j.compag.2022.107025>
34. Sarijaloo F.B., Porta M., Taslimi B., Pardalos P.M. Yield performance estimation of corn hybrids using machine learning algorithms. *Artificial Intelligence in Agriculture*. 2021;5:82-89. <https://doi.org/10.1016/j.aiia.2021.05.001>
35. Awasthi P., Mishra S., Gupta N. Performance Assessment of Machine Learning Techniques for Corn Yield Prediction. *International Conference on Advanced Network Technologies and Intelligent Computing*. Cham: Springer Nature Switzerland. 2022;320-335. https://doi.org/10.1007/978-3-031-28183-9_23
36. Wolanin A., Mateo-García G., Camps-Valls G. et al. Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. *Environmental Research Letters*. 2020;15(2):024019. <https://doi.org/10.1088/1748-9326/ab68ac>
37. Hassija V., Chamola V., Mahapatra A. et al. Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*. 2023;16:45-74. <https://doi.org/10.1007/s12559-023-10179-8>
38. Wang X., Huang J., Feng Q., Yin D. Winter wheat yield prediction at county level and uncertainty analysis

36. Wolanin A., Mateo-García G., Camps-Valls G. et al. Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. *Environmental Research Letters*. 2020;15(2):024019. <https://doi.org/10.1088/1748-9326/ab68ac>

37. Hassija V., Chamola V., Mahapatra A. et al. Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*. 2023;16:45-74. <https://doi.org/10.1007/s12559-023-10179-8>

38. Wang X., Huang J., Feng Q., Yin D. Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. *Remote Sensing*. 2020;12(11):1744. <https://doi.org/10.3390/rs12111744>

39. Reichstein M., Camps-Valls G., Stevens B. et al. Deep learning and process understanding for data-driven Earth system science. *Nature*. 2019;566:195-204. <https://doi.org/10.1038/s41586-019-0912-1>

40. Каличкин В.К., Максимович К.Ю. Методология формирования цифровой системы управления земледелием // Сибирский вестник сельскохозяйственной науки. 2024. Т. 54, № 3. С. 5-20. EDN: OJMVRJ

Информация об авторах

¹ **Каличкин Владимир Климентьевич**, д-р с.-х. наук, руководитель научного направления по земледелию и агрохимии СФНЦА РАН; SPIN-код: 1587-4615, AuthorID: 247907; vk.kalichkin@gmail.com; <https://orcid.org/0000-0002-7765-3451>

² **Федоров Дмитрий Сергеевич**, младший научный сотрудник лаборатории агроклиматических исследований СФНЦА РАН; SPIN-код: 9530-8327, AuthorID: 1133884; dima.fedorov99@mail.ru; <https://orcid.org/0000-0001-8678-400X>

³ **Максимович Кирилл Юрьевич**, канд. биол. наук, научный сотрудник лаборатории агроклиматических исследований СФНЦА РАН; SPIN-код: 4700-3900, AuthorID: 1025027; kiri-maksimovi@mail.ru; <https://orcid.org/0000-0002-9563-4641>

⁴ **Риксен Вера Сергеевна**, заведующий лабораторией агроклиматических исследований СФНЦА РАН; SPIN-код: 7858-1352, AuthorID: 1068243; riclog@mail.ru; <https://orcid.org/0000-0001-9898-4754>

^{1,2,3,4} Сибирский федеральный научный центр агроботехнологий Российской академии наук; 630501, Российская Федерация, Новосибирская область, р.п. Краснообск, ул. Центральная, 2б

Вклад авторов

В.К. Каличкин – концептуализация, методология, администрирование проекта, проведение исследования, создание окончательной версии (доработка) рукописи и ее редактирование; Д.С. Федоров – программное обеспечение, верификация данных, формальный анализ, проведение исследования, визуализация; К.Ю. Максимович – концептуализация, формальный анализ, проведение исследования, администрирование проекта, создание окончательной версии (доработка) рукописи и ее редактирование;

В.С. Риксен – методология, верификация данных, администрирование данных, формальный анализ, проведение исследования.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов и несут ответственность за плагиат.

Статья поступила 18.08.2024, после рецензирования и доработки 07.10.2024; принята к публикации 06.12.2024

in main wheat-producing regions of China with deep learning approaches. *Remote Sensing*. 2020;12(11):1744. <https://doi.org/10.3390/rs12111744>

39. Reichstein M., Camps-Valls G., Stevens B. et al. Deep learning and process understanding for data-driven Earth system science. *Nature*. 2019;566:195-204. <https://doi.org/10.1038/s41586-019-0912-1>

40. Kalichkin V.K., Maksimovich K.Yu. Methodology for forming a digital farming management system. *Siberian Herald of Agricultural Science*. 2024;54(3):5-20. (In Russ.)

Author Information

¹ **Vladimir K. Kalichkin**, DSc (Ag), Head of Agriculture and Agrochemistry Research Area, SFSCA RAS; vk.kalichkin@gmail.com; <https://orcid.org/0000-0002-7765-3451>.

² **Dmitry S. Fedorov**, Junior Researcher, Agro-Climatic Research Laboratory, SFSCA RAS; dima.fedorov99@mail.ru; <https://orcid.org/0000-0001-8678-400X>.

³ **Kirill Yu. Maksimovich**, PhD (Biology), Research Associate, Agro-Climatic Research Laboratory, SFSCA RAS; kiri-maksimovi@mail.ru; <https://orcid.org/0000-0002-9563-4641>.

⁴ **Vera S. Riksen**, Head of Agro-Climatic Research Laboratory, SFSCA RAS; riclog@mail.ru; <https://orcid.org/0000-0001-9898-4754>.

^{1,2,3,4} Siberian Federal Scientific Center of Agrobiotechnologies at the Russian Academy of Sciences; 630501, Tsentralnaya Str., 2b, Krasnoobsk, Novosibirsk Oblast, Russia

Author Contributions

V.K. Kalichkin – conceptualization, methodology, project administration, conducting research, writing – review and editing of the manuscript;

D.S. Fedorov – software, data verification, formal analysis, investigation, visualization;

K.Yu. Maksimovich – conceptualization, formal analysis, investigation, project administration, writing – review and editing of the manuscript;

V.S. Riksen – methodology, data verification, data administration, formal analysis, investigation.

Conflict of interests

The authors declare no conflict of interests regarding the publication of this article and bear equal responsibility for plagiarism.

Received 18.08.2024; Revised 07.10.2024; Accepted 06.12.2024