

## Оценка эффективности мультисезонных моделей машинного обучения для оперативного распознавания озимых культур на больших территориях

Д. Е. Плотников<sup>1</sup>, Ю. Ш. Бойматов<sup>2</sup>, Е. С. Ёлкина<sup>1</sup>,  
Е. В. Щербенко<sup>3</sup>, А. С. Плотникова<sup>4</sup>

<sup>1</sup> *Институт космических исследований РАН, Москва, 117997, Россия*  
*E-mail: dmitplot@d902.iki.rssi.ru*

<sup>2</sup> *Московский физико-технический институт (НИУ)*  
*Долгопрудный, Московская обл., 141700, Россия*

<sup>3</sup> *ООО «Институт космических исследований Земли», Москва, 121205, Россия*

<sup>4</sup> *Центр по проблемам экологии и продуктивности лесов РАН*  
*Москва, 117997, Россия*

Оценивались возможности ансамблевых методов машинного обучения Random Forest и XGBoost (*англ.* eXtreme Gradient Boosting) для распознавания озимых культур на больших территориях при обучении на мультисезонных наборах данных за пять последовательных лет. В рамках исследования были проведены два эксперимента: по сравнению эффективности мультисезонных моделей указанных архитектур и по сравнению мультисезонной модели и серии односезонных моделей оптимальной архитектуры. Оценка результатов экспериментов проводилась для региона исследования, охватывающего более 94 % посевных площадей озимых Российской Федерации, на основе сравнения с данными Росстата и опорными картами, полученными методом экспансии обучающей выборки, с использованием метрики RMSE (*англ.* root mean squared error) и коэффициента детерминации  $R^2$ , а также метрики  $F$ -score по классу «озимые». В первом эксперименте была продемонстрирована перспективность мультисезонных моделей, причём мультисезонная модель Random Forest показала себя эффективнее и стабильнее, чем XGBoost, улучшение RMSE по сравнению с опорными картами составило в среднем 28 тыс. га и наблюдалось для любого из пяти сезонов; относительное улучшение составило 15 % в среднем, достигая 35 %. Результаты второго эксперимента по сравнению пяти сезонных моделей и мультисезонной модели Random Forest также указали на эффективность мультисезонной модели: суммарное за пять сезонов значение RMSE мультисезонной модели в 1,32–2,02 раз меньше, чем для односезонных моделей, а среднее значение  $F$ -score для любой односезонной модели находится в диапазоне 0,75–0,79, в то время как для мультисезонной модели достигает значения 0,84.

**Ключевые слова:** мультисезонная модель, Random Forest, XGBoost, спутниковое картографирование, озимые культуры, LOWESS, временные ряды, сеточный поиск, машинное обучение

Одобрена к печати: 30.09.2024

DOI: 10.21046/2070-7401-2024-21-5-116-129

### Введение

Спутниковый мониторинг является источником объективной и оперативной информации о размещении и состоянии сельскохозяйственных культур, способствуя обеспечению продовольственной безопасности на различных уровнях пространственного охвата (Bégué et al., 2018; Weiss et al., 2020; Wu et al., 2023). Одна из востребованных задач в области дистанционного мониторинга сельскохозяйственной растительности — раннее распознавание посевов культур и групп культур на национальном уровне. Результаты оперативного распознавания используются для заблаговременной оценки состояния зерновых и прогнозирования урожайности, улучшая информационное обеспечение сельскохозяйственного сектора экономики.

Методы дистанционного распознавания и оценки культур текущего сезона вегетации обычно основаны на использовании пространственно и тематически представленной обучающей выборки — репрезентативной информации об их расположении (Blickensdörfer et al.,

2022; Valero et al., 2016; Waldner et al., 2019). Ряд подходов используют информацию о сроках сева и этапах развития (Foerster et al., 2012; Skakun et al., 2019; Zhong et al., 2016), другие используют устойчивые различия сезонной динамики дистанционно измеряемых характеристик растений (Massey et al., 2017, Palchoudhuri et al., 2018), а также перенос обученных моделей на другие сезоны вегетации (Wang et al., 2019). Часто используются временные серии предварительно обработанных спутниковых изображений, которые доказали свою эффективность для решения этих задач (Плотников и др., 2020а; Blickensdörfer et al., 2022; Massey et al., 2017). В частности, для решения различных задач оперативного и исторического мониторинга используются безоблачные композитные изображения (Лупян и др. 2021, Трошко и др., 2022), содержащие ежедневные измерения коэффициентов спектральной яркости земной поверхности и производные вегетационные индексы, восстановленные на основе весовой интерполяции LOWESS (*англ.* LOcally WEighted Scatterplot Smoothing) (Plotnikov et al., 2022). Использованию спутниковых данных для распознавания и оценки состояния сельскохозяйственных культур на локальном уровне в различных регионах мира и России посвящено множество работ (Плотников и др., 2011; Повх и др., 2013; Савин и др., 2011; Терехин, 2017; Blickensdörfer et al., 2022; Kussul et al., 2015; Lebourgeois et al., 2017; Skakun et al., 2019; Zhong et al., 2016), однако публикаций на тему методов для национального и глобального охвата немного. Одним из методов, работающих в оперативном режиме на больших территориях, выступает разработанный в ИКИ РАН автоматический метод на базе экспансии обучающей выборки для раннего распознавания озимых культур России на национальном уровне пространственного охвата с региональной точностью от 70 до 98 % (Лупян и др. 2021; Плотников и др., 2017; Трошко и др., 2022). Указанный метод для работы использует спутниковые данные только текущего сезона вегетации и априорные предположения о фенологии культур, являясь одним из немногих решений, предоставляющих репрезентативную информацию о размещении озимых культур на больших территориях за произвольный сезон вегетации. Разработка автоматизированных методов сезонного картографирования и оценки культур на национальном и глобальном охвате на основе данных высокого пространственного разрешения ведётся также коммерческими организациями (например, OneSoil), однако эти результаты недоступны для бесплатного использования или для научных исследований.

Мультисезонная универсальность и устойчивость методов оперативного распознавания культур может быть повышена усилением разносезонной информацией. Такого рода подходы нацелены на создание более адаптивной и осведомлённой модели на базе методов машинного обучения, которая обеспечит картографирование культур и групп культур произвольного сезона вегетации без использования сезон-специфичных данных, включая обучающую выборку или фенологическую информацию целевого сезона вегетации. Для решения такого рода задач используются методы машинного обучения, включая нейросетевые архитектуры LSTM (*англ.* Long-Short-Term Memory) и MLP (*англ.* Multi-Layer Perception), а также архитектуры на основе ансамблей деревьев решений Random Forest и их градиентного бустинга (Breiman 2001; Xu et al., 2020; Zhang et al., 2022). Кроме этого, была показана эффективность использования имитационного моделирования для этих целей, если на этапе параметризации имитационных моделей используются разносезонные спутниковые и климатические данные (Плотников и др., 2018). Тем не менее, указанные методы ранее применялись только на локальном уровне пространственного охвата.

В настоящей работе исследуются возможности методов машинного обучения ансамблевых архитектур на базе деревьев решений Random Forest и XGBoost (*англ.* eXtreme Gradient Boosting) для распознавания озимых культур на больших неоднородных территориях, при этом для построения моделей одновременно используются разносезонные предикторы и автоматически созданные разметки. Цель работы заключается в создании и оценке универсальных, т.е. устойчиво применимых для произвольного сезона вегетации, моделей для распознавания озимых культур на больших территориях, их сравнении с соответствующими односезонными моделями, обученными на одинаковых исходных наборах входных данных.

## Методы

В рамках исследования была проведена серия экспериментов по сравнению эффективности двух групп моделей, первая из которых (мультисезонные или универсальные модели) использует для обучения одновременно весь многолетний разносезонный набор признаков и выборок, а вторая группа моделей обучается на наборе признаков и выборке одного сезона вегетации. В качестве исследуемых методов машинного обучения выступали два ансамблевых алгоритма, основанных на деревьях решений, — так называемый «случайный лес» или Random Forest (Breiman, 2001) и экстремальный градиентный бустинг деревьев решений в реализации XGBoost (Chen, Guestrin, 2016).

Основные различия методов XGBoost и Random Forest следующие:

1. Градиентный бустинг последовательно добавляет слабые модели, состоящие из взаимосвязанных и последовательно построенных деревьев, в то время как случайный лес использует ансамбль параллельных и независимых деревьев решений (рис. 1).

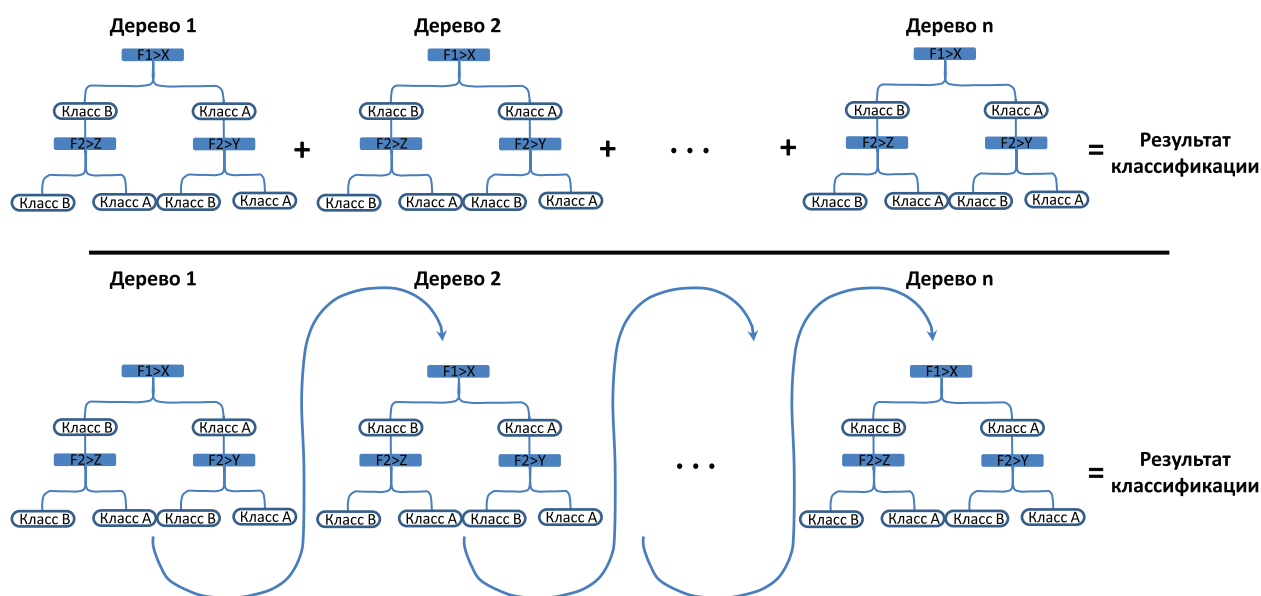


Рис. 1. Схемы работы случайного леса (вверху) и экстремального градиентного бустинга (внизу) на основе деревьев решений

2. В связи с большей склонностью к переобучению градиентный бустинг в реализации XGBoost использует регуляризации L1 и L2 для борьбы с переобучением, в то время как случайный лес обычно не использует регуляризацию.
3. XGBoost использует первую и вторую производную функции потерь для ускорения обучения и повышения точности предсказаний, в то время как случайный лес не использует градиентные методы.
4. Случайный лес, благодаря своей архитектуре, легче поддерживает параллельную обработку и распределённое обучение на нескольких ядрах (процессорах, машинах).

Кроме этого, имеющиеся в открытом доступе программные реализации этих методов поддерживают возможность работы с большими объёмами данных и устойчивы к переобучению, что важно при использовании автоматически создаваемой и содержащей возможные ошибки разметки. Наконец, реализации этих методов имеют достаточное быстродействие, чтобы обеспечить обработку больших объёмов одновременно подаваемых разносезонных входных данных за приемлемое время.

Алгоритмы реализованы в среде python3 с использованием библиотек sklearn, RAPID, CuMI, pandas, dask. В работе задействованы параллельные вычисления на графическом процессоре для ускорения процесса разработки и исследования. Для работы с графическим про-

цессором использовалась библиотека CUDA (*англ.* Compute Unified Device Architecture), а также инструменты для параллельных вычислений, написанные с её использованием (Garland et al., 2008). Для подбора гиперпараметров моделей использовался поиск по сетке (*англ.* grid search) (Probst et al., 2019). В работе использовалось разделение данных на обучающее и тестовое подмножества в соотношении 80 на 20 % соответственно. В качестве функции потерь использовалась метрика Binary Cross-Entropy Loss, при параметризации — метрика  $F$ -score.

Оптимизация работы с использованием библиотеки dask позволила существенно сократить использование памяти в представлении данных. Реализация алгоритма с использованием библиотеки CuMI, работающей на технологии CUDA при использовании графического процессора NVIDIA GeForce RTX 4060, позволила ускорить обучение и классификацию более чем в 44 раза по сравнению с работой на обычном 4-ядерном процессоре.

Оценка эффективности моделей проводилась на основе среднеквадратического отклонения RMSE (*англ.* root mean squared error) и коэффициента детерминации  $R^2$  при сравнении с данными официальной статистики Росстата по площади озимых культур. Кроме этого, в эксперименте по сравнению универсальной (мультисезонной) и односезонных моделей вычислялась метрика  $F$ -score по классу «озимые» на основе опорных карт озимых культур.

## Данные

Получаемые на основе весовой интерполяции LOWESS мультиспектральные ежедневные композитные изображения формируются с использованием ресурсов Центра коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений (ЦКП) «ИКИ-Мониторинг» (Лупян и др., 2019) и доступны пользователям платформ семейства «Вега» (<http://sozvezdie-vega.ru/>) в качестве продуктов под названием TSE (*англ.* Time Series Enhancement). Они создаются по данным различных спутниковых систем высокого, среднего и низкого пространственного разрешения — Sentinel-2 (MSI, *англ.* Multispectral Instrument), «Метеор-М» № 2-2.2 (Комплекс многозональной спутниковой съёмки, KMCC-М и KMCC-2), Terra/Aqua (MODIS, *англ.* Moderate Resolution Imaging Spectroradiometer) и SNPP (*англ.* Suomi NPP, ранее NPOESS Preparatory Project, NPP) (VIIRS, *англ.* Visible Infrared Imaging Radiometer Suite) — в историческом (продукт HiTSE, *англ.* Historical TSE) и оперативном (продукт OpTSE, *англ.* Operative TSE) режимах и используются для оперативного картографирования и оценки состояния сельскохозяйственной растительности, используемых земель и решения ряда других задач (Денисов и др., 2024; Кашницкий и др., 2023; Плотников и др., 2020б, 2022; Kolbudaev et al., 2021; Plotnikov et al., 2023). В работе использовались временные серии ежедневных композитных изображений HiTSE NDVI (*англ.* Normalized Difference Vegetation Index) по данным прибора MODIS с пространственным разрешением 250 м (рис. 2, см. с. 120).

В качестве опорных карт использовались созданные методом экспансии обучающей выборки карты озимых культур на территорию России за весенне-летний период вегетации (Плотников и др., 2017) (см. рис. 2).

В настоящей работе в качестве исходных данных были использованы пять наборов, состоящих из признаков и разметки. В качестве признаков использовалась сезонная временная серия ежедневных композитных изображений HiTSE NDVI с 1 января по 15 июля, охватывающая период весенней ревегетации, созревания и уборки озимых, а также карта агроклиматических зон GAEZ (*англ.* Global Agro-Ecological Zones) (Fischer et al., 2012), в качестве разметки — соответствующие карты размещения озимых культур на территорию Европейской части РФ. Каждый набор соответствовал одному сезону из интервала с 2015 по 2019 г. включительно. Кроме этого, для оценки результатов распознавания в работе использовались статистические данные Росстата по площади озимых культур на уровне субъектов Российской Федерации ([rosstat.gov.ru](http://rosstat.gov.ru)) за соответствующий интервал лет.

Регион исследования включал территорию 36 субъектов Российской Федерации европейской России, целиком попадающих внутрь прямоугольника со стороной примерно 2500 км,

задаваемого четырьмя гранулами MODIS, включая h19v04, h20v04, h21v03 и h21v04. Помимо субконтинентальных размеров, исследуемая территория имеет ярко выраженный широтный градиент почвенно-климатических условий и агропрактик: вариабельность среднегодового количества осадков превышает 400 мм (от 400 до 800 мм), вариабельность сумм накопленной температуры превышает 1000 °С (от 2500 до 3500 °С), даты начала сева и весенней ревегетации озимых различаются на два месяца, территория охватывает более 50 различных зон агростратификации GAEZ. Указанные 36 субъектов федерации являются ключевыми регионами озимого клина в РФ: общая среднемноголетняя площадь озимых в этих субъектах за пятилетний интервал анализа достигает 15,6 млн га, что составляет 94,5 % от средней посевной площади РФ за этот же интервал времени (16,5 млн га).

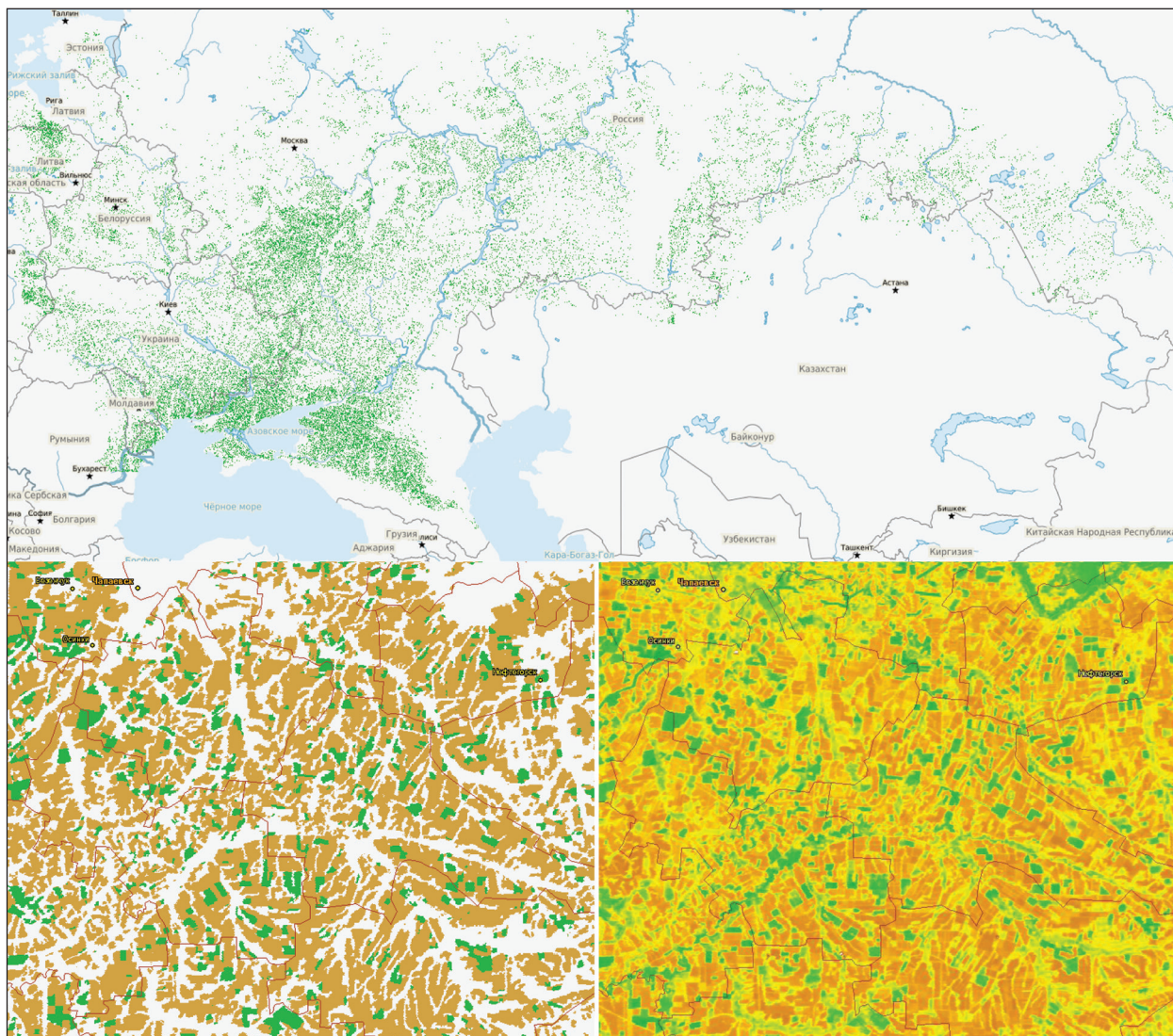


Рис. 2. Карта озимых культур сезона уборки 2020 г. на территорию озимого клина РФ и сопредельных стран (зелёный цвет, *вверху*), а также фрагмент карты озимых (зелёный цвет) на фоне карты пахотных земель (коричневый цвет) на часть территории Самарской области (*слева внизу*) совместно с ежедневным композитным изображением NDVI OpTSE по данным MODIS за 15 мая 2020 г. (*справа внизу*)

## Дизайн экспериментов

На первом этапе экспериментов на основе одинаковых наборов разносезонных признаков и опорных карт были обучены и оптимизированы две универсальные модели для распознавания озимых на больших территориях: на базе Random Forest и на базе XGBoost. После этого

они были развёрнуты на каждом из исследуемых сезонов 2015–2019 гг. с получением оценок площади озимых на уровне субъектов федерации для полученных карт и для опорных карт, выступающих в роли контроля, с вычислением RMSE и  $R^2$  при сопоставлении с данными статистики Росстата по площади озимых культур (рис. 3, слева).

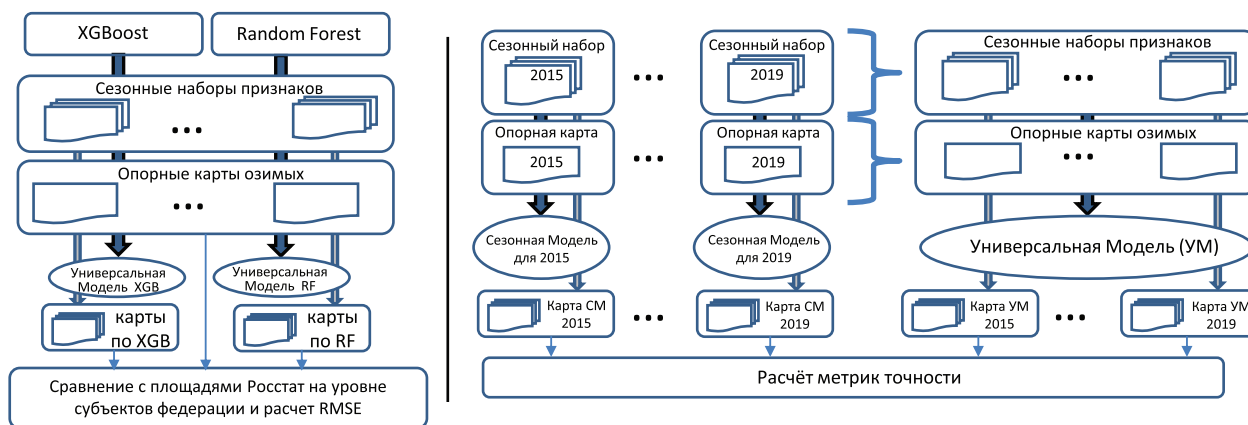


Рис. 3. Схема анализа данных при сравнении эффективности универсальных моделей XGBoost и Random Forest, а также оценке результатов и опорных карт на основе статистических данных Росстата (слева), а также схема анализа данных при сравнении эффективности сезон-специфичных моделей с универсальной моделью с расчётом метрик точности (справа); широкими стрелками обозначены процессы обучения моделей, узкими — классификация, тонкими — анализ полученных результатов

По результатам этого эксперимента была идентифицирована наиболее эффективная универсальная модель и для неё был проведён второй этап экспериментов. Он заключался в сравнении показателей точности универсальной модели с показателями сезон-специфичной модели, обученной по данным только целевого сезона вегетации из диапазона 2015–2019 гг. и созданной на базе того же метода машинного обучения (см. рис. 3, справа). Для оценки эффективности моделей использовались метрики  $F$ -score и RMSE, сравнение проводилось с опорными картами размещения озимых и со статистическими данными Росстата.

## Результаты

Изначально в процессе оптимизации параметров каждая сезонная и универсальная модель настраивалась независимо. Однако было обнаружено, что использование для всех моделей единых параметров, полученных в процессе оптимизации универсальных моделей, качественно не влияет на результаты и на взаимное положение моделей в рейтинге. Количественные различия также находились в пределах погрешности оценок, поэтому для удобства описания были приняты единые оптимальные параметры.

Параметризация методом поиска по сетке дала следующие параметры моделей: параметр  $n\_estimators$  (число деревьев) равен 500, параметр  $max\_depth$  (максимальная глубина дерева) для XGBoost равен 6, коэффициент прореживания равен 4.

### Оценка эффективности мультисезонных моделей Random Forest и XGBoost

Обученные универсальные модели Random Forest и XGBoost были развёрнуты на каждом из сезонов в интервале с 2015 по 2019 г., после чего полученные на их основе карты, а также контрольные опорные карты озимых были использованы для расчёта площади обнаруженных озимых в границах каждого из 36 субъектов РФ региона исследования. Оценки площади озимых сравнивались с соответствующими данными Росстата с вычислением метрики ошибки RMSE для каждого сезона. На рис. 4 (см. с. 122) приведены оценки ошибок RMSE

для каждого из трёх источников, для информации приведены суммарные посевные площади озимых культур региона исследования по данным Росстата.

Улучшение оценок площади озимых наиболее эффективной модели по отношению к опорной разметке в терминах RMSE составляет в среднем 15 % и может достигать 35 % (для 2016 г.), что указывает на практическую значимость использования мультисезонных моделей. При сравнении с контрольными картами более точными и устойчивыми во времени выглядят результаты универсальной модели с архитектурой Random Forest. В частности, Random Forest демонстрирует минимальное значение ошибки в сравнении с XGBoost и опорными картами в любой из исследуемых сезонов. При этом можно увидеть, что универсальная модель XGBoost в ряде случаев имеет даже более высокую ошибку, чем опорные карты.

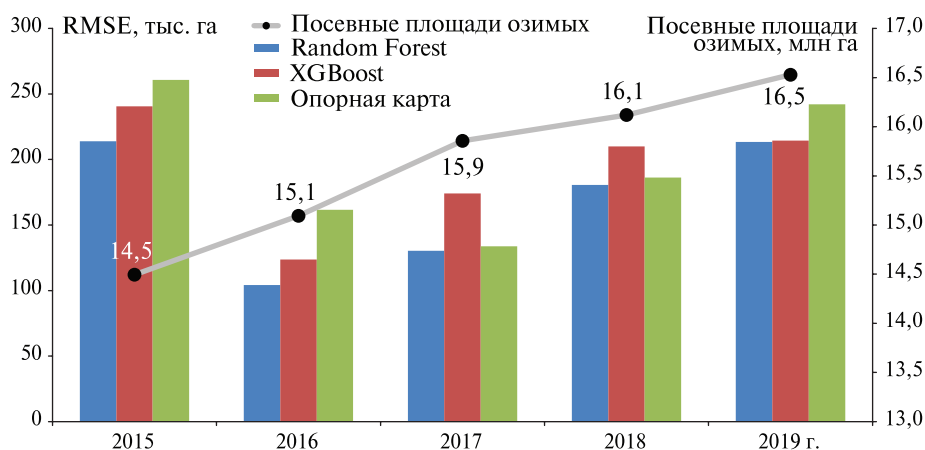


Рис. 4. Представленные для трёх источников данных (по модели Random Forest — синий цвет, по модели XGBoost — красный, по опорным картам — зелёный) значения метрики ошибок RMSE при сравнении с данными Росстата (левая вертикальная ось), а также динамика суммы посевных площадей озимых субъектов региона исследования с 2015 по 2019 г. (правая вертикальная ось)

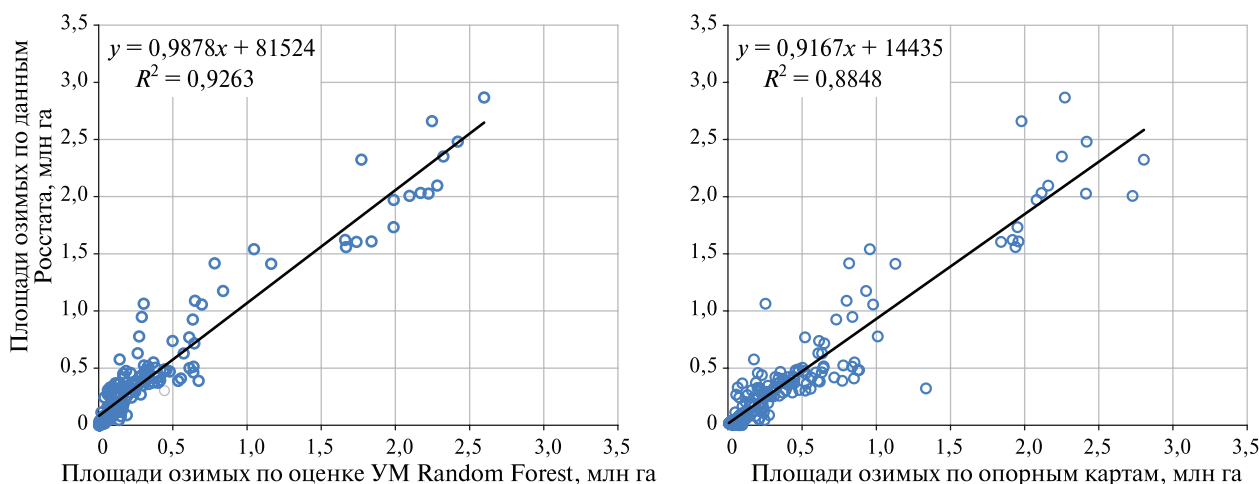


Рис. 5. Диаграммы рассеяния при сравнении оценок площади озимых на основе универсальной модели Random Forest (слева) и опорных карт (справа) с данными Росстата на уровне субъектов федерации за 2015–2019 гг.

Таким образом, из двух архитектур, исследованных с целью построения универсальной мультисезонной модели для картографирования озимых на больших территориях, более эффективной оказалась архитектура Random Forest. Диаграммы рассеяния, характеризующие улучшение точности мультисезонной модели по сравнению с контрольными опорными картами, представлены на рис. 5. Диаграммы характеризуют динамику согласованности оценок площади озимых культур с данными Росстата при оценке с помощью универсальной модели

Random Forest (см. рис. 5, слева) и опорных карт (см. рис. 5, справа), интегрально за все исследуемые сезоны. Универсальная модель характеризуется более высоким коэффициентом детерминации, угловой коэффициент регрессионной прямой ближе к единице; кучность новых оценок возле линии регрессии повысилась, а их отклонение от линии 1:1 снизилось по сравнению с опорными картами.

### Оценка сезонной точности универсальной модели Random Forest

На втором этапе экспериментов универсальная модель сравнивалась с серией моделей, обученных на основе данных только одного сезона вегетации; в качестве рабочей архитектуры использовался метод Random Forest. В рамках этого исследования каждая из пяти сезонных моделей, а также универсальная модель, были развёрнуты на каждом из пяти исследуемых сезонов вегетации. Для всех моделей были использованы одинаковые значения параметров, полученные в результате параметризации мультисезонных моделей для первой серии экспериментов: число деревьев, равное 500, и коэффициент прореживания, равный 4. При оценке эффективности моделей использовались статические данные Росстата на уровне субъектов РФ и метрика RMSE, а также метрика  $F$ -score и опорные карты озимых культур.

Результаты сравнения пяти сезонных моделей и универсальной модели с данными Росстата по площади озимых на уровне субъектов РФ для всей территории исследования представлены в табл. 1. Нижняя строка таблицы соответствует сумме ошибок каждой модели за пять исследуемых сезонов.

Таблица 1. Посезонные и интегральные значения ошибки RMSE для пяти сезонных моделей и универсальной модели Random Forest для распознавания озимых культур, млн га

Сезон	Модель RF 2015	RF 2016	RF 2017	RF 2018	RF 2019	Универсальная модель RF
2015 г.	0,456	0,526	0,579	0,395	0,375	<b>0,318</b>
2016 г.	0,217	0,272	0,441	0,344	0,312	<b>0,083</b>
2017 г.	0,438	0,536	0,312	0,490	0,245	<b>0,160</b>
2018 г.	0,340	0,307	0,310	0,263	<b>0,230</b>	0,258
2019 г.	0,557	0,522	0,514	0,503	0,260	<b>0,253</b>
<b>Сумма RMSE</b>	2,008	2,165	2,156	1,995	1,422	<b>1,072</b>

Примечание: полужирным начертанием выделены минимальные значения RMSE в каждой строке.

Таблица 2. Посезонные значения метрики  $F$ -score для пяти сезонных моделей и универсальной модели Random Forest для распознавания озимых культур

Сезон	Модель RF 2015	RF 2016	RF 2017	RF 2018	RF 2019	Универсальная модель RF
2015 г.	<b>0,832</b>	0,765	0,742	0,771	0,716	0,814
2016 г.	0,767	0,792	0,733	0,757	0,812	<b>0,827</b>
2017 г.	0,737	0,826	0,783	0,757	0,791	<b>0,882</b>
2018 г.	0,759	0,752	0,796	<b>0,814</b>	0,808	0,810
2019 г.	0,707	0,797	0,743	0,740	0,842	<b>0,883</b>

Примечание: полужирным начертанием выделены максимальные значения  $F$ -score в каждой строке.

Кроме этого, сравнительный анализ карт, полученных на основе этих моделей, был выполнен с помощью метрики  $F$ -score с использованием опорной разметки в качестве эталона (табл. 2).



## Обсуждение

По результатам первой серии экспериментов (см. рис. 4) можно сделать выводы, что универсальные непараметрические модели для распознавания озимых на больших территориях на базе Random Forest и XGBoost в среднем оказываются точнее, чем опорные карты озимых, на которых эти модели были обучены; улучшение среднемноголетнего за пять рассматриваемых лет значения RMSE составило 28 и 4 тыс. га соответственно. При этом архитектура модели имеет значение: метод Random Forest показал себя эффективнее и устойчивее, чем метод XGBoost, продемонстрировав более высокие показатели в каждый из сезонов.

Ранее было отмечено, что автоматически создаваемые методом экспансии обучающей выборки опорные карты характеризуются общей точностью детектирования озимых культур в диапазоне 0,85–0,95, а ошибки носят в целом случайный характер в пространстве и во времени (Плотников и др., 2017). Непараметрические методы достаточно легко подстраиваются под разметку и поэтому во многом воспроизводят её ошибки, что заметно по скоррелированной динамике RMSE обеих универсальных моделей между сезонами (см. рис. 4), однако метод Random Forest обладает более высокой способностью к генерализации, в то время как XGBoost более подвержен переобучению, поэтому его сложнее использовать для улучшения содержащих ошибки опорных данных.

Интересно отметить, что во второй серии экспериментов при сравнении с данными Росстата универсальная модель Random Forest почти всегда оказывается лучше, чем любая сезон-специфичная модель (табл. 1). Более того, во многих случаях сезон-специфичная модель целевого сезона вегетации в среднем оказывалась хуже, чем сезон-специфичная модель иного сезона вегетации: такое произошло для всех сезонов, кроме сезона 2019 г. При сравнении же с опорной разметкой, на которой и происходило обучение (табл. 2), лучшей по метрике  $F$ -score всегда оказывалась либо сезон-специфичная модель, либо универсальная модель. При этом разница  $F$ -score между лучшей и второй по точности моделью, если это была нецелевая сезон-специфичная модель, почти всегда превышает существенное значение (обычно 0,05 и более). Такого рода созависимость с разметкой, при отсутствии аналогичных трендов в табл. 1, где сравнение производилось с полностью независимым источником, также указывает на восприимчивость Random Forest к возможным недостаткам обучающих данных, подтверждая, тем не менее, способность к генерализации разносезонной информации. Последнее качество позволяет использовать универсальную модель этой архитектуры для улучшения результатов картографирования озимых культур на больших территориях.

Дальнейший анализ универсальных моделей требует использования проверенной экспертами точной геопространственной информации о размещении озимых культур на рассматриваемой территории выращивания озимых за пять исследуемых сезонов, что выходит за рамки настоящего исследования. Тем не менее, можно отметить, что уменьшение среднеквадратической региональной ошибки оценки на 28 тыс. га является заметным улучшением, поскольку для более чем 20 % субъектов из зоны интереса среднегодовая посевная площадь озимых культур не превышает и половины этого значения.

## Заключение

В рамках настоящей работы впервые были получены мультисезонные непараметрические модели, обученные на основе одновременного использования разносезонной информации в целях распознавания озимых культур на больших территориях озимого клина Российской Федерации. Кроме этого, в рамках работы было показано, что использование мультисезонных моделей может повысить точность распознавания озимых по сравнению с автоматическими методами картографирования культур, использующих спутниковые данные только целевого сезона вегетации. Таким образом, продемонстрирована перспективность таких моделей при решении задач оперативного сельскохозяйственного спутникового мониторинга различных сезонов вегетации на больших неоднородных территориях.

Исследование оценки эффективности мультисезонных моделей для распознавания культур выполнено за счёт гранта Российского научного фонда № 23-27-00412, <https://rscf.ru/project/23-27-00412/>.

## Литература

1. Денисов П. В., Трошко К. А., Луян Е. А. и др. Спутниковое дистанционное зондирование Земли в задаче установления границ земель сельскохозяйственного назначения // *Землеустройство, кадастр и мониторинг земель*. 2024. Т. 19. № 4(231). С. 227–232. DOI: 10.33920/sel-4-2404-04.
2. Кашиницкий А. В., Луян Е. А., Плотников Д. Е., Толпин В. А. Анализ возможности использования данных различного пространственного разрешения при проведении мониторинга объектов // *Современные проблемы дистанционного зондирования Земли из космоса*. 2023. Т. 20. № 2. С. 60–74. DOI: 10.21046/2070-7401-2023-20-2-60-74.
3. Луян Е. А., Прошин А. А., Бурцев М. А. и др. Опыт эксплуатации и развития центра коллективного пользования системами архивации, обработки и анализа спутниковых данных (ЦКП «ИКИ-Мониторинг») // *Современные проблемы дистанционного зондирования Земли из космоса*. 2019. Т. 16. № 3. С. 151–170. DOI: 10.21046/2070-7401-2019-16-3-151-170.
4. Луян Е. А., Середа И. И., Денисов П. В. и др. Дистанционный мониторинг состояния озимых культур зимой 2020–2021 гг. на Европейской территории России // *Современные проблемы дистанционного зондирования Земли из космоса*. 2021. Т. 18. № 1. С. 165–172. DOI: 10.21046/2070-7401-2021-18-1-165-172.
5. Плотников Д. Е., Барталев С. А., Жарко В. О., Михайлов В. В., Просяникова О. И. Экспериментальная оценка распознаваемости агрокультур по данным сезонных спутниковых измерений спектральной яркости // *Современные проблемы дистанционного зондирования Земли из космоса*. 2011. Т. 8. № 1. С. 199–208.
6. Плотников Д. Е., Барталев С. А., Луян Е. А., Толпин В. А. Оценка точности выявления посевов озимых культур в весенне-летний период вегетации по данным прибора MODIS // *Современные проблемы дистанционного зондирования Земли из космоса*. 2017. Т. 14. № 4. С. 132–145. DOI: 10.21046/2070-7401-2017-14-4-132-145.
7. Плотников Д. Е., Хвостиков С. А., Барталев С. А. Метод автоматического распознавания сельскохозяйственных культур на основе спутниковых данных и имитационной модели развития растений // *Современные проблемы дистанционного зондирования Земли из космоса*. 2018. Т. 15. № 4. С. 131–141. DOI: 10.21046/2070-7401-2018-15-4-131-141.
8. Плотников Д. Е., Ёлкина Е. С., Дунаева Е. А. и др. (2020а) Развитие метода автоматического распознавания озимых культур на основе спутниковых данных для оценки их состояния на территории Республики Крым // *Таврический вестн. аграрной науки*. 2020. № 1(21). С. 64–82. DOI: 10.33952/2542-0720-2020-1-21-64-83.
9. Плотников Д. Е., Колбудаев П. А., Жуков Б. С. и др. (2020б) Публикация коллекции мультиспектральных измерений прибором КМСС-М (КА «Метеор-М» № 2) для количественной оценки характеристик земной поверхности // *Современные проблемы дистанционного зондирования Земли из космоса*. 2020. Т. 17. № 7. С. 276–282. DOI: 10.21046/2070-7401-2020-17-7-276-282.
10. Плотников Д. Е., Колбудаев П. А., Луян Е. А. Автоматический метод субпиксельной географической привязки спутниковых изображений КМСС-М на основе актуализируемого эталона низкого пространственного разрешения // *Компьютерная оптика*. 2022. Т. 46. № 5. С. 818–827. DOI: 10.18287/2412-6179-СО-1098.
11. Повх В. И., Шляхова Л. А., Боева И. Н. Оценка структуры посевных площадей сельскохозяйственных культур по спутниковой информации высокого разрешения // *Современные проблемы дистанционного зондирования Земли из космоса*. 2013. Т. 10. № 4. С. 224–228.
12. Савин И. Ю., Барталев С. А., Луян Е. А., Толпин В. А., Медведева М. А., Плотников Д. Е. Спутниковый мониторинг воздействия засухи на растительность (на примере засухи 2010 года в России) // *Современные проблемы дистанционного зондирования Земли из космоса*. 2011. Т. 8. № 1. С. 150–162.
13. Терехин Э. А. Применение дискриминантного анализа для распознавания посевов сельскохозяйственных культур // *Современные проблемы дистанционного зондирования Земли из космоса*. 2017. Т. 14. № 7. С. 89–99. DOI: 10.21046/2070-7401-2017-14-7-89-99.
14. Трошко К. А., Денисов П. В., Дунаева Е. А. и др. Особенности развития зерновых культур в России в 2022 году по данным дистанционного мониторинга // *Современные проблемы дистанционного зондирования Земли из космоса*. 2022. Т. 19. № 6. С. 301–307. DOI: 10.21046/2070-7401-2022-19-6-301-307.

15. *Bégué A., Arvor D., Bellon B. et al.* Remote sensing and cropping practices: A Review // *Remote Sensing*. 2018. V. 10. No. 1. Article 99. <https://doi.org/10.3390/rs10010099>.
16. *Blickensdörfer L., Schwieder M., Pflugmacher D. et al.* Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat-8 data for Germany // *Remote Sensing of Environment*. 2022. V. 269. Article 112831. <https://doi.org/10.1016/j.rse.2021.112831>.
17. *Breiman L.* Random forests. *Machine learning*. 2001. V. 45. P. 5–32. DOI: 10.1023/A:1010933404324.
18. *Chen T., Guestrin C.* XGBoost: A scalable tree boosting system // *KDD '16: Proc. 22<sup>nd</sup> ACM SIGKDD Intern. Conf. Knowledge Discovery and Data Mining*. 2016. P. 785–794. DOI: 10.1145/2939672.2939785.
19. *Fischer G., Nachtergaele F., Prieler S., Teixeira E., Tóth G., Velthuizen H., Verelst L., Wiberg D.* Global AgroEcological Zones (GAEZ ver3.0). Model documentation. IIASA/FAO. 2012. 196 p.
20. *Foerster S., Kaden K., Förster M., Itzerott S.* Crop type mapping using spectral-temporal profiles and phenological information // *Computers and Electronics in Agriculture*. 2012. V. 89. P. 30–40. DOI: 10.1016/j.compag.2012.07.015.
21. *Garland M., Le Grand S., Nickolls J. et al.* Parallel computing experiences with CUDA // *IEEE Micro*. 2008. V. 28(4). P. 13–27. <https://doi.org/10.1109/mm.2008.57>.
22. *Kolbudaev P.A., Plotnikov D.E., Loupian E.A. et al.* The methods and automatic technology aimed at imagery georeferencing, cloud screening, atmospheric and radiometric correction of KMSS-M satellite data // *E3S Web of Conf.* 2021. V. 333. Article 01006. DOI: [doi.org/10.1051/e3sconf/202133301006](https://doi.org/10.1051/e3sconf/202133301006).
23. *Kussul N., Skakun S., Shelestov A. et al.* Regional scale crop mapping using multi-temporal satellite imagery // *Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2015. V. XL-7/W3. P. 45–52. <https://doi.org/10.5194/isprsarchives-XL-7-W3-45-2015>.
24. *Lebourgeois V., Dupuy S., Vintrou E. et al.* A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM) // *Remote Sensing*. 2017. V. 9. Article 259. <https://doi.org/10.3390/rs9030259>.
25. *Massey R., Sankey T., Congalton R. et al.* MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types // *Remote Sensing of Environment*. 2017. V. 198. P. 490–503. DOI: 10.1016/j.rse.2017.06.033.
26. *Palchoudhuri Y., Valcarce-Diñeiro R., King P., Sanabria-Soto M.* Classification of multi-temporal spectral indices for crop type mapping: a case study in Coalville, UK // *J. Agricultural Science*. 2018. P. 1–13. DOI: 10.1017/S0021859617000879.
27. *Plotnikov D.E., Loupian E.A., Kolbudaev P.A. et al.* Daily surface reflectance reconstruction using LOWESS on the example of various satellite systems // *IEEE Xplore. 8<sup>th</sup> Intern. Conf. Information Technology and Nanotechnology (ITNT)*. 2022. DOI: 10.1109/ITNT55410.2022.9848630.
28. *Plotnikov D.E., Kolbudaev P.A., Matveev A.M. et al.* Accuracy assessment of atmospheric correction of KMSS-2 Meteor-M No. 2.2 data over Northern Eurasia // *Remote Sensing*. 2023. V. 15. Iss. 18. Article 4395. DOI: [doi.org/10.3390/rs15184395](https://doi.org/10.3390/rs15184395).
29. *Probst P., Wright M.N., Boulesteix A.L.* Hyperparameters and tuning strategies for random forest // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2019. V. 9. Iss. 3. Article e1301. <https://doi.org/10.1002/widm.1301>.
30. *Skakun S., Vermote E., Franch B. et al.* Winter wheat yield assessment from Landsat-8 and Sentinel-2 data: incorporating surface reflectance, through Phenological fitting, into regression yield models // *Remote Sens.* 2019. V. 11. Iss. 15. Article 1768. <https://doi.org/10.3390/rs11151768>.
31. *Valero S., Morin D., Inglada J. et al.* Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions // *Remote Sensing*. 2016. V. 8. Iss. 1. Article 55. <https://doi.org/10.3390/rs8010055>.
32. *Waldner F., Schucknecht A., Lesiv M. et al.* Conflation of expert and crowd reference data to validate global binary thematic maps // *Remote Sensing of Environment*. 2019. Vol. 221. P. 235–246. DOI: 10.1016/j.rse.2018.10.039.
33. *Wang S., Azzari G., Lobell D.B.* Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques // *Remote Sensing of Environment*. 2019. V. 222. P. 303–317. <https://doi.org/10.1016/j.rse.2018.12.026>.
34. *Weiss M., Jacob F., Duveiller G.* Remote sensing for agricultural applications: A meta-review // *Remote Sensing of Environment*. 2020. V. 236. Article 111402. <https://doi.org/10.1016/j.rse.2019.111402>.
35. *Wu B., Zhang M., Zeng H. et al.* Challenges and opportunities in remote sensing-based crop monitoring: A review // *National Science Review*. 2023. Article nwc290. <https://doi.org/10.1093/nsr/nwc290>.
36. *Xu J., Zhu Y., Zhon R. et al.* DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping // *Remote Sensing of Environment*. 2020. V. 247. Article 111946. <https://doi.org/10.1016/j.rse.2020.111946>.
37. *Zhang C., Di L., Lin L. et al.* Towards automation of in-season crop type mapping using spatiotemporal crop information and remote sensing data // *Agricultural Systems*. 2022. V. 201. Article 103462. <https://doi.org/10.1016/j.agry.2022.103462>.

38. Zhong L., Hu L., Yu L., Gong P., Biging G.S. Automated mapping of soybean and corn using phenology // J. Photogrammetry and Remote Sensing. 2016. V. 119. P. 151–164. <https://doi.org/10.1016/j.isprsjprs.2016.05.014>.

## Performance assessment of multi-season machine learning models for large-scale in-season winter crops mapping

D. E. Plotnikov<sup>1</sup>, Yu. Sh. Boimatov<sup>2</sup>, E. S. Elkina<sup>1</sup>,  
E. V. Shcherbenko<sup>3</sup>, A. S. Plotnikova<sup>4</sup>

<sup>1</sup> Space Research Institute RAS, Moscow 117997, Russia  
E-mail: dmitplot@d902.iki.rssi.ru

<sup>2</sup> Moscow Institute of Physics and Technology  
Dolgoprudny, Moscow Region 141700, Russia

<sup>3</sup> OOO Space Research Institute for the Earth, Moscow 121205, Russia

<sup>4</sup> Center for Forest Ecology and Productivity RAS, Moscow 117997, Russia

In this study, we have evaluated the potential of ensemble machine learning methods Random Forest and XGBoost for winter crops mapping over large areas when training on multi-season datasets for five consecutive years. Two experiments were conducted: accuracy intercomparison of multi-season models of two architectures and intercomparison of multi-season and a series of single-season models of optimal architecture. The results of the experiments were provided for the study region covering more than 94 % of winter crops areas sown in Russia, based on comparison with the Rosstat (Federal State Statistics Service) data and the results of the automatic benchmark method using RMSE (root mean squared error) and coefficient of determination  $R^2$ , as well as the  $F$ -score metric for the winter crop class. In the first experiment, the potential of multi-season models was demonstrated, and the Random Forest multi-season model proved to be more efficient and more stable than XGBoost, the RMSE improvement compared to the reference maps averaged 28 th. ha and was observed for all of the five seasons; the relative improvement was 15 % on average, reaching 35 %. The results of the second experiment which compared the five single-season models and the multi-season model also indicated the effectiveness of the multi-season model: the cumulative RMSE of the multi-season model for 5 seasons is 1.32–2.02 times less than of single-season models, and the average  $F$ -score for any multi-season model falls in the range of 0.75–0.79, while for the multi-season model, it is as high as 0.84.

**Keywords:** multi-season model, Random Forest, XGBoost, satellite mapping, winter crops, LOWESS, time series, grid search, machine learning

Accepted: 30.09.2024

DOI: 10.21046/2070-7401-2024-21-5-116-129

## References

1. Denisov P. V., Troshko K. A., Loupian E. A. et al., Satellite remote sensing for boundaries determination of agricultural lands, *Land management, cadastre and land monitoring*, 2024, Vol. 19, No. 4(231), pp. 227–232 (in Russian), DOI: 10.33920/sel-4-2404-04.
2. Kashnitskii A. V., Loupian E. A., Plotnikov D. E., Tolpin V. A., Analysis of the possibility of using different spatial resolution data for objects monitoring, *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2023, Vol. 20, No. 2, pp. 60–74 (in Russian), DOI: 10.21046/2070-7401-2023-20-2-60-74.
3. Loupian E. A., Proshin A. A., Bourtsev M. A. et al., Experience of development and operation of the IKI-Monitoring center for collective use of systems for archiving, processing and analyzing satellite data, *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2019, Vol. 16, No. 3, pp. 151–170 (in Russian), DOI: 10.21046/2070-7401-2019-16-3-151-170.
4. Loupian E. A., Sereda I. I., Denisov P. V. et al., Remote monitoring of winter crops condition in the winter of 2020/2021 in the European territory of Russia, *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2021, Vol. 18, No. 1, pp. 165–172 (in Russian), DOI: 10.21046/2070-7401-2021-18-1-165-172.

5. Plotnikov D. E., Bartalev S. A., Zharko V. O. Mihailov V. V., Prosyannikova O. I., An experimental assessment of crop types recognisability using time-series of intra-seasonal spectral reflectance measurements by satellite sensor, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 199–208 (in Russian).
6. Plotnikov D. E., Bartalev S. A., Loupian E. A., Tolpin V. A., Accuracy assessment for winter crops mapping in spring-summer growing season with MODIS data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2017, Vol. 14, No. 4, pp. 132–145 (in Russian), DOI: 10.21046/2070-7401-2017-14-4-132-145.
7. Plotnikov D. E., Khvostikov S. A., Bartalev S. A., Method for automated crop types mapping using remote sensing data and a plant growth simulation model, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2018, Vol. 15, No. 4, pp. 131–141 (in Russian), DOI: 10.21046/2070-7401-2018-15-4-131-141.
8. Plotnikov D. E., Elkina E. S., Dunaieva Ie. A. et al. (2020a), Development of the method for automatic winter crops mapping by means of remote sensing aimed at crops state assessment over the Republic of Crimea, *Taurida Herald of the Agrarian Sciences*, 2020, No. 1(21), pp. 64–82 (in Russian), DOI: 10.33952/2542-0720-2020-1-21-64-83.
9. Plotnikov D. E., Kolbudaev P. A., Zhukov B. S. et al. (2020b), The collection of multispectral KMSS-M (Meteor-M No. 2) satellite data aimed at quantitative assessment of the Earth surface, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2020, Vol. 17, No. 7, pp. 276–282 (in Russian), DOI: 10.21046/2070-7401-2020-17-7-276-282.
10. Plotnikov D. E., Kolbudaev P. A., Loupian E. A., An automatic method for subpixel registration of KMSS-M imagery based on coarse-resolution actualized reference, *Computer Optics*, 2022, Vol. 46, No. 5, pp. 818–827 (in Russian), DOI: 10.18287/2412-6179-CO-1098.
11. Povkh V. I., Shlyakhova L. A., Boeva I. N., Assessment of the structure of agricultural crop areas using high resolution satellite data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2013, Vol. 10, No. 4, pp. 224–228 (in Russian).
12. Savin I. Yu., Bartalev S. A., Loupian E. A., Tolpin V. A., Medvedeva M. A., Plotnikov D. E., Satellite monitoring of vegetation affected by drought (using drought 2010 in Russia as an example), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 150–162 (in Russian).
13. Terekhin E. A., Application of discriminant analysis for recognition of agricultural crops, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2017, Vol. 14, No. 7, pp. 89–99 (in Russian), DOI: 10.21046/2070-7401-2017-14-7-89-99.
14. Troshko K. A., Denisov P. V., Dunaeva E. A. et al., Features of grain crops development in Russia in 2022 based on remote sensing data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2022, Vol. 19, No. 6, pp. 301–307, DOI: 10.21046/2070-7401-2022-19-6-301-307 (in Russian).
15. Bégué A., Arvor D., Bellon B. et al., Remote sensing and cropping practices: A review, *Remote Sensing*, 2018, Vol. 10, No. 1, Article 99, <https://doi.org/10.3390/rs10010099>.
16. Blickensdörfer L., Schwieder M., Pflugmacher D. et al., Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat-8 data for Germany, *Remote Sensing of Environment*, Vol. 269, 2022, Article 112831, <https://doi.org/10.1016/j.rse.2021.112831>.
17. Breiman L., Random Forests, *Machine Learning*, 2001, Vol. 45, No. 1, pp. 5–32, DOI: 10.1023/A:1010933404324.
18. Chen T., Guestrin C., XGBoost: A scalable tree boosting system, *KDD '16: Proc. 22<sup>nd</sup> ACM SIGKDD Intern. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794, DOI: 10.1145/2939672.2939785.
19. Fischer G., Nachtergaele F., Prieler S., Teixeira E., Tóth G., Velthuizen H., Verelst L., Wiberg D., *Global AgroEcological Zones (GAEZ ver3.0). Model documentation*, IIASA/FAO, 2012, 196 p.
20. Foerster S., Kaden K., Förster M., Itzerott S., Crop type mapping using spectral-temporal profiles and phenological information, *Computers and Electronics in Agriculture*, 2012, Vol. 89, pp. 30–40, DOI: 10.1016/j.compag.2012.07.015.
21. Garland M., Le Grand S., Nickolls J., et al., Parallel Computing Experiences with CUDA, *IEEE Micro*, 2008, V. 28(4), pp. 13–27, <https://doi.org/10.1109/mm.2008.57>.
22. Kolbudaev P. A., Plotnikov D. E., Loupian E. A. et al., The methods and automatic technology aimed at imagery georeferencing, cloud screening, atmospheric and radiometric correction of KMSS-M satellite data, *E3S Web of Conf.*, 2021, Vol. 333, Article 01006, DOI: [doi.org/10.1051/e3sconf/202133301006](https://doi.org/10.1051/e3sconf/202133301006).
23. Kussul N., Skakun S., Shelestov A. et al., Regional scale crop mapping using multi-temporal satellite imagery, *Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2015, Vol. XL-7/W3, pp. 45–52, <https://doi.org/10.5194/isprsarchives-XL-7-W3-45-2015>.
24. Lebourgeois V., Dupuy S., Vintrou E. et al., A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM), *Remote Sensing*, 2017, Vol. 9, Article 259, <https://doi.org/10.3390/rs9030259>.

25. Massey R., Sankey T., Congalton R. et al., MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types, *Remote Sensing of Environment*, 2017, Vol. 198, pp. 490–503, DOI: 10.1016/j.rse.2017.06.033.
26. Palchoudhuri Y., Valcarce-Diñeiro R., King P., Sanabria-Soto M., Classification of multi-temporal spectral indices for crop type mapping: a case study in Coalville, UK, *J. Agricultural Science*, 2018, pp. 1–13, DOI: 10.1017/S0021859617000879.
27. Plotnikov D.E., Loupian E.A., Kolbudaev P.A. et al., Daily surface reflectance reconstruction using LOWESS on the example of various satellite systems, *IEEE Xplore. 8<sup>th</sup> Intern. Conf. Information Technology and Nanotechnology (ITNT)*, 2022, DOI: 10.1109/ITNT55410.2022.9848630.
28. Plotnikov D.E., Kolbudaev P.A., Matveev A.M. et al., Accuracy assessment of atmospheric correction of KMSS-2 Meteor-M No. 2.2 data over Northern Eurasia, *Remote Sensing*, 2023, Vol. 15, Issue 18, Article 4395, DOI: doi.org/10.3390/rs15184395.
29. Probst P., Wright M.N., Boulesteix A.L., Hyperparameters and tuning strategies for random forest, *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 2019, Vol. 9, Issue 3, Article e1301, <https://doi.org/10.1002/widm.1301>.
30. Skakun S., Vermote E., Franch B. et al., Winter wheat yield assessment from Landsat-8 and Sentinel-2 data: incorporating surface reflectance, through Phenological fitting, into regression yield models, *Remote Sens.*, 2019, Vol. 11, Issue 15, Article 1768, <https://doi.org/10.3390/rs11151768>.
31. Valero S., Morin D., Inglada J. et al., Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions, *Remote Sensing*, 2016, Vol. 8, Issue 1, Article 55, <https://doi.org/10.3390/rs8010055>.
32. Waldner F., Schucknecht A., Lesiv M. et al., Conflation of expert and crowd reference data to validate global binary thematic maps, *Remote Sensing of Environment*, 2019, Vol. 221, pp. 235–246, DOI: 10.1016/j.rse.2018.10.039.
33. Wang S., Azzari G., Lobell D.B., Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques, *Remote Sensing of Environment*, 2019, Vol. 222, pp. 303–317, <https://doi.org/10.1016/j.rse.2018.12.026>.
34. Weiss M., Jacob F., Duveiller G., Remote sensing for agricultural applications: A meta-review, *Remote Sensing of Environment*, 2020, Vol. 236, Article 111402, <https://doi.org/10.1016/j.rse.2019.111402>.
35. Wu B., Zhang M., Zeng H. et al., Challenges and opportunities in remote sensing-based crop monitoring: A review, *National Science Review*, 2023, Article nwac290, <https://doi.org/10.1093/nsr/nwac290>.
36. Xu J., Zhu Y., Zhong R. et al., DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping, *Remote Sensing of Environment*, 2020, Vol. 247, Article 111946, <https://doi.org/10.1016/j.rse.2020.111946>.
37. Zhang C., Di L., Lin L. et al., Towards automation of in-season crop type mapping using spatiotemporal crop information and remote sensing data, *Agricultural Systems*, 2022, Vol. 201, Article 103462, <https://doi.org/10.1016/j.agsy.2022.103462>.
38. Zhong L., Hu L., Yu L., Gong P., Biging G.S., Automated mapping of soybean and corn using phenology, *J. Photogrammetry and Remote Sensing*, 2016, Vol. 119, pp. 151–164, <https://doi.org/10.1016/j.isprsjprs.2016.05.014>.