

Использование методов  
машинного обучения для решения  
задач мониторинга пахотных  
земель в ХФИЦ ДВО РАН

Докладчик: Дубровин Константин Николаевич,  
Вычислительный центр ДВО РАН – обособленное подразделение  
ХФИЦ ДВО РАН

# ЦИФРОВЫЕ ТЕХНОЛОГИИ В ЗЕМЛЕДЕЛИИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Принято Постановление Правительства РФ от **14.05.2021** № 731 «О Государственной программе *эффективного вовлечения в оборот земель сельскохозяйственного назначения*

- вовлечение в оборот 13 млн га ранее использовавшихся и не использовавшихся сельхозугодий.

Создание единой карты-схемы земель сельхозназначения РФ (для каждого из субъектов РФ): границы ЗСН (с установлением границ сельскохозяйственных угодий). Хабаровский край, Приморский край, Амурская область - 2023 г.

*Указ Президента РФ от 07.05.2024 N 309 "О национальных целях и стратегических задачах развития Российской Федерации на период до 2030 года и перспективу до 2036".*

достижение к 2030 году "**цифровой зрелости**" ...ключевых отраслей экономики предполагающей **автоматизацию** ...на основе данных с учетом ускоренного внедрения технологий обработки **больших объемов данных, машинного обучения** и искусственного интеллекта



# Реализация государственной программы

Россия / Дальневосточный федеральный округ / Хабаровский край / ...

2024

Яровые и прочие

Озимые

Фильтры:

Сводная информация



По данным МСХ сев составляет (План), га 59 870

По данным РОУ АПК сев составляет, га 56 600

Количество полей, шт 1 608

Внесено в ЕФИС ЗСН, га 38 340,68

Процент наполнения: 67,74

Дата и время расчета отчета 24.10.2025 13:21:05

Россия / Дальневосточный федеральный округ / Хабаровский край / ...

2025

Яровые и прочие

Озимые

Фильтры:

Сводная информация



По данным МСХ сев составляет (План), га 56 670

По данным РОУ АПК сев составляет, га 60 363

Количество полей, шт 1 394

Внесено в ЕФИС ЗСН, га 37 437,32

Процент наполнения: 62,02

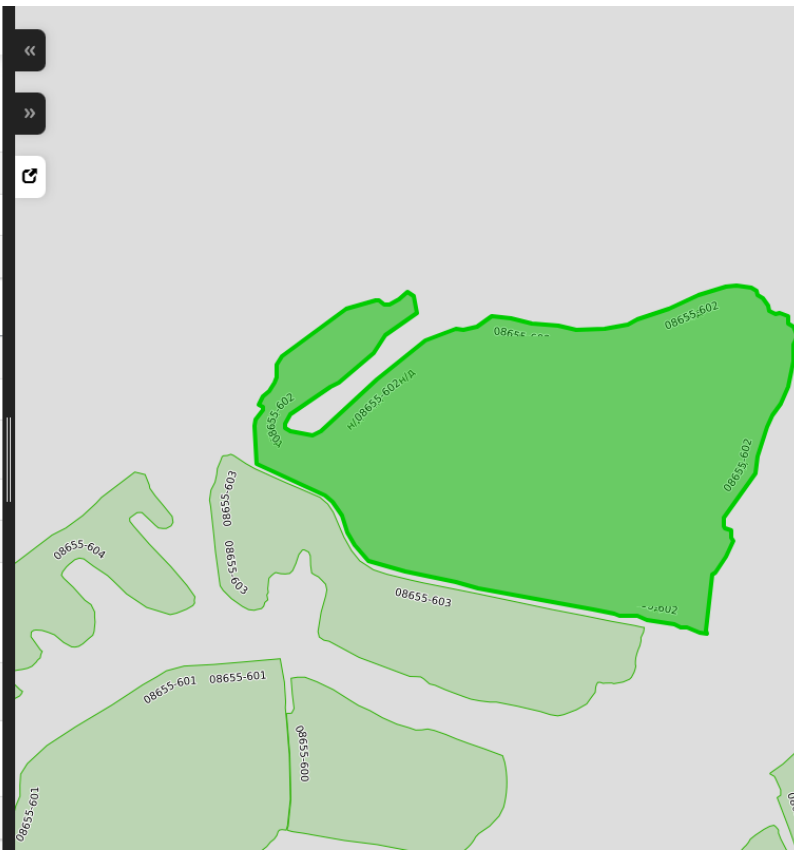
Дата и время расчета отчета 26.10.2025 12:23:53

Текущий сезон

Регион [Хабаровский край](#)  
 Муниципальное образование [Хабаровский муниципальный район](#)  
 Номер поля ЕФИС ЗСН 08655-602

Поле [Права зарегистрированные](#) [Права без регистрации](#)

Площадь, га 112,55  
 Географические координаты 48°19'26.424"N 134°55'55.552"E  
 Идентификатор поля в региональной системе нет данных  
 Внутрихозяйственный номер поля нет данных  
 Вид угодий [Пашня](#)  
 Кадастровый номер 27:17:0625001:2440-2443,2445,2447,2449,2450,2459,2463,2466,2563,2680,2682,2684,2762,2764,2766,2819,2865,3141,3155,3541,3848,4651,4652,4653,4654,4655,4658,4659,4660,4661-4675,6075,6079  
 Факт использования/неиспользования земель [Используется](#)  
 Об отнесении земельного участка к особо ценным продуктивным сельскохозяйственным угодьям нет данных  
 Примечания нет данных



Действия	Номер севооборота	Сезон	Культура	ОКПД 2	Сорт семян	Площадь сева, га	Цель севооборота	Род (культура)
	<a href="#">08655-602-2020-1</a>	2020	Нет данных	нет данных	нет данных	112,54535	Посев/ посадка	нет дан
	<a href="#">08655-602-2019-1</a>	2019	Нет данных	нет данных	нет данных	112,54535	Посев/ посадка	нет дан
	<a href="#">08655-602-2018-1</a>	2018	Нет данных	нет данных	нет данных	112,54535	Посев/ посадка	нет дан
	<a href="#">08655-602-2018-2</a>	2018	Нет данных	нет данных	нет данных	112,54535	Посев/ посадка	нет дан

Россия / [Дальневосточный федеральный округ](#) / [Хабаровский край](#)

/ [Хабаровский муниципальный район](#) / 86550-1524

Текущий сезон

Регион [Хабаровский край](#)

Муниципальное образование [Хабаровский муниципальный район](#)

Номер поля ЕФИС ЗСН 86550-1524

Поле [Права зарегистрированные](#) [Права без регистрации](#)

Площадь, га 68,71

Географические координаты 48°35'4.812"N 135°6'48.367"E

Идентификатор поля в региональной системе нет данных

Внутрихозяйственный номер поля нет данных

Вид угодий Нет данных

Кадастровый номер 27:17:0329001:4554, 27:17:0329001:4555, 27:17:0329001:4559

Факт использования/ неиспользования земель Нет данных

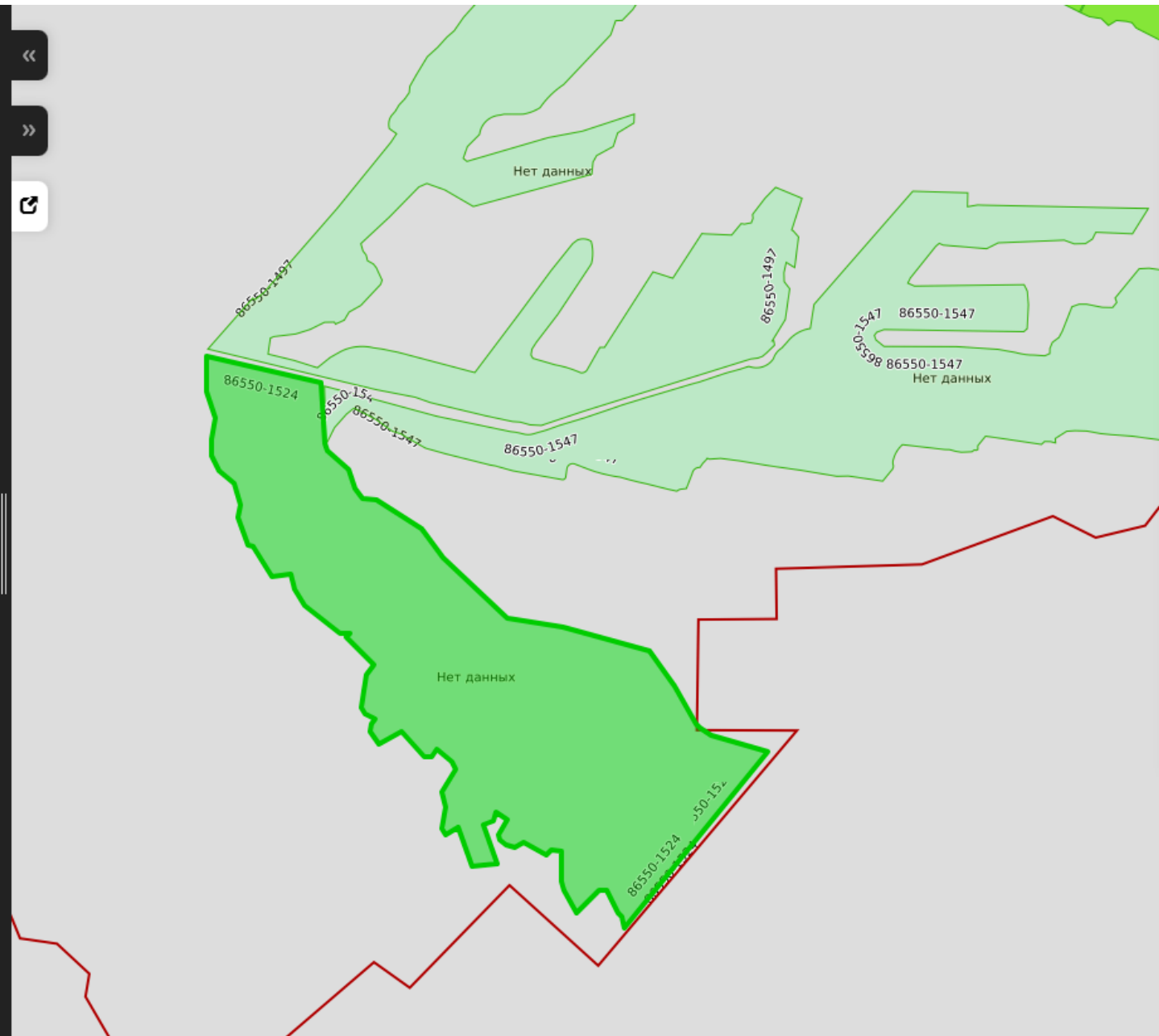
Об отнесении земельного участка к особо ценным продуктивным сельскохозяйственным угодьям нет

Примечания нет данных

[Показать на карте](#)

[Редактировать](#)

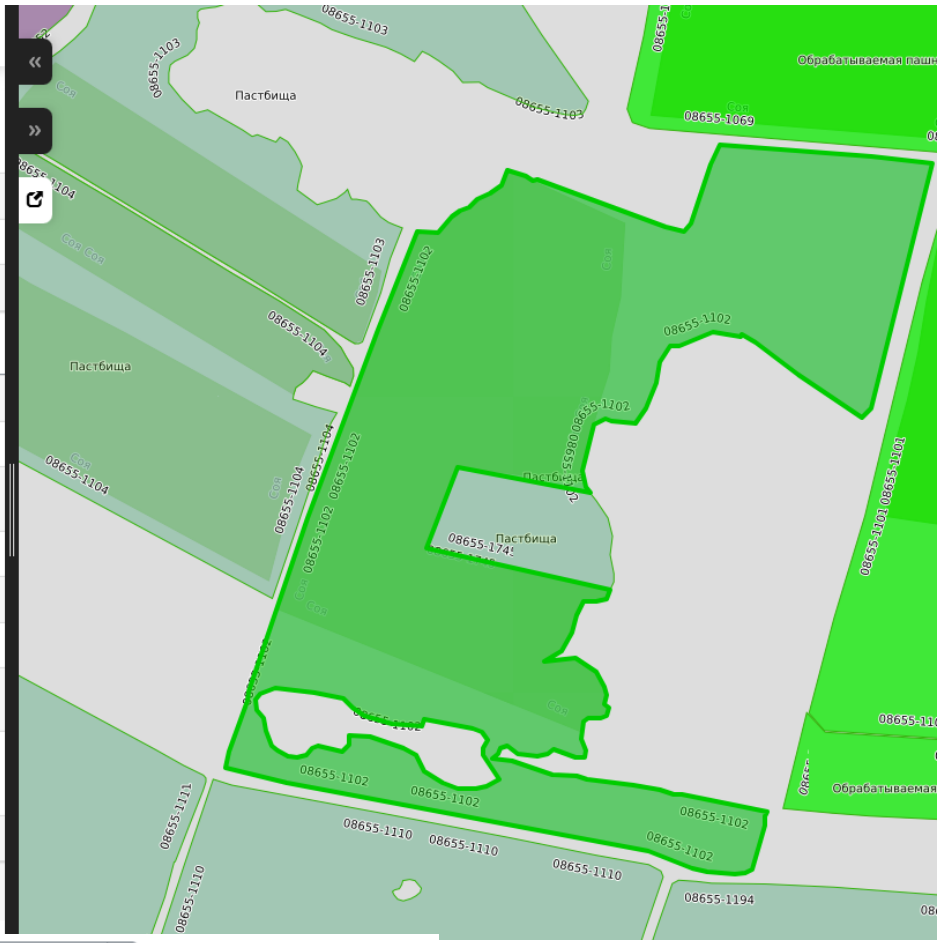
[Удалить](#)







Текущий сезон

Регион	Хабаровский край
Муниципальное образование	Хабаровский муниципальный район
Номер поля ЕФИС ЗСН	08655-1102
Поле	Права зарегистрированные    Права без регистрации
Площадь, га	40,3
Географические координаты	48°30'30.497"N 135°16'11.700"E
Идентификатор поля в региональной системе	нет данных
Внутрихозяйственный номер поля	нет данных
Вид угодий	Пастбища
Кадастровый номер	0329201:100
Факт использования/ неиспользования земель	Используется
Об отнесении земельного участка к особо ценным продуктивным сельскохозяйственным угодьям	нет данных
Примечания	нет данных

Показать на карте    Редактировать    Удалить



Действия	Номер севооборота	Сезон	Культура	ОКПД 2	Сорт семян	Площадь сева, га	Цель севооборота
 	08655-1102-2025-1	2025	Соя	01.11.81.110 - Бобы соевые для посева	ГРЭЙ	20,29738	Посев/ посадка
 	08655-1102-2024-1	2024	Соя	01.11.81.110 - Бобы соевые для посева	ВНИИС 18	40,29519	Посев/ посадка

# Неверно указана культура

ЭФИС ЭСН

Госмониторинг

Российская Федерация | Хабаровский край | Хабаровский муниципальный район | 08655-1025

Поле

Общая информация  
История вводов  
Журнал изменений  
История организаций  
Другие акты

СЕВООБОРОТ

План / Факт

МОНИТОРИНГ

Вегетационный индекс  
Сравнение инд. вегетационных индексов  
Сравнение NDT со средними значениями  
Сравнение значений NDT с эталоном  
Погодные условия

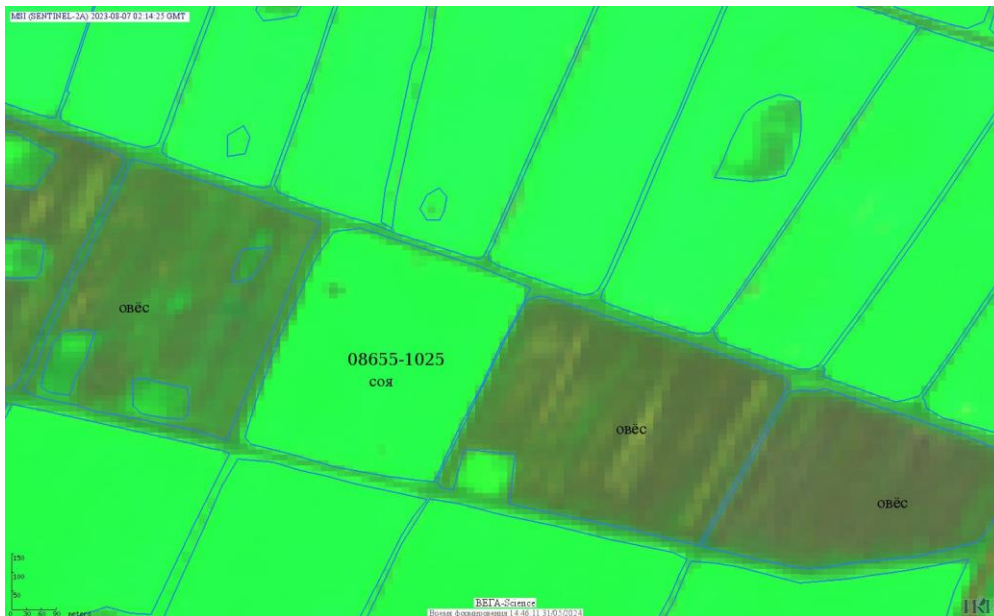
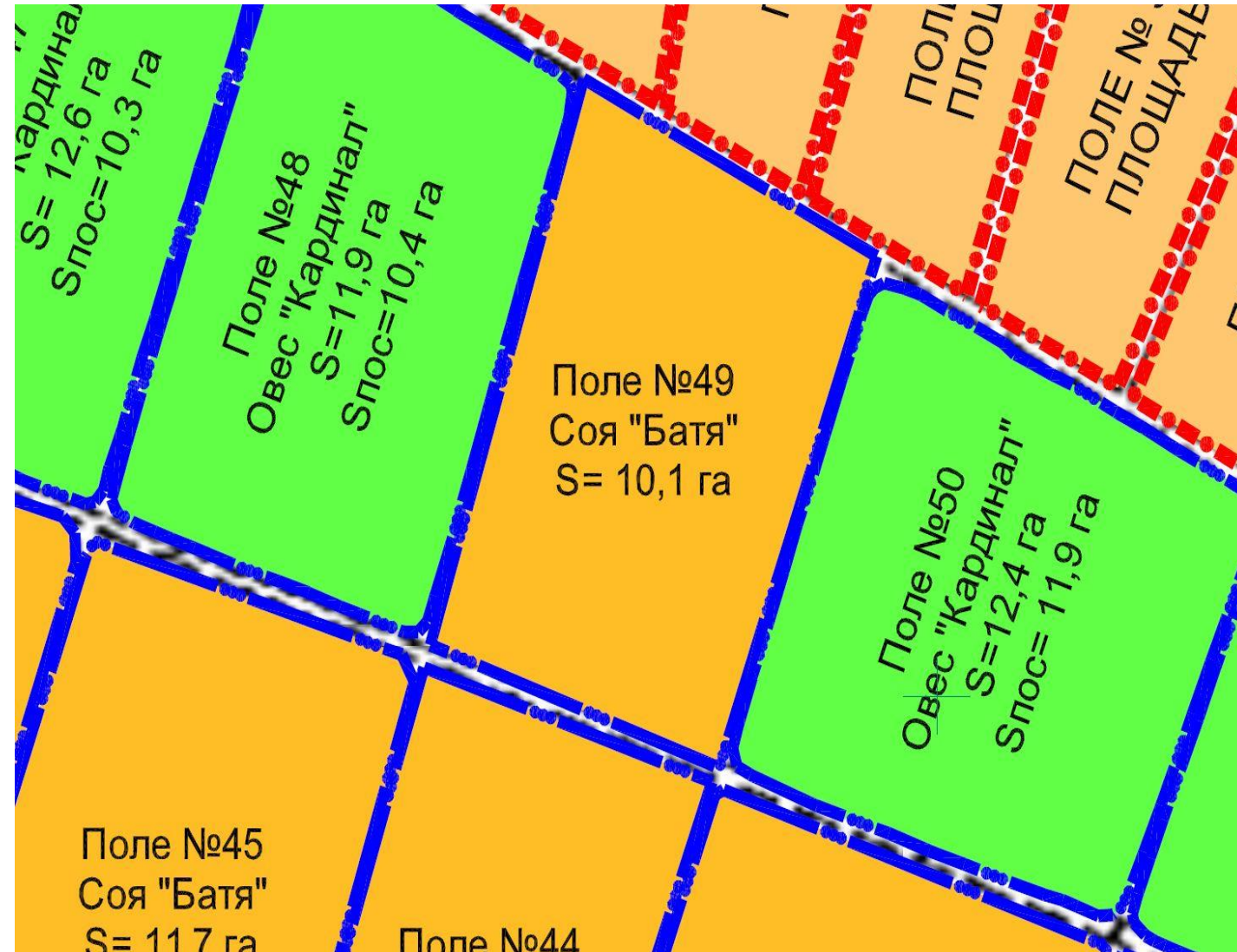
АКС

Обследования

Период	Культура, план	Культура, факт	Площадь, сева, га	Площадь, га	Урожайность, ц/га	Валовый сбор, ц	МНТ, млн
2024	нет данных	овес	10,1	10,44	нет данных	нет данных	27200
2023	овес	овес	10,44	10,44	0	0	0
2020	нет данных	соя	10,48	10,44	13,1	137,28	0
2019	нет данных	нет данных	10,48	10,44	нет данных	нет данных	0
2018	нет данных	соя	10,43	10,44	14,4	150,24	нет д.
2018	нет данных	соя	10,43	10,44	14,4	150,24	нет д.

Фильтр Экспорт Добавить культуру

Ср. темп 18,6 °C



# Неактуальные контуры полей

ГОСМОНИТОРИНГ

Риски: [Дальневосточный федеральный округ] [Сибирский федеральный округ] [Хабаровский краевой округ]

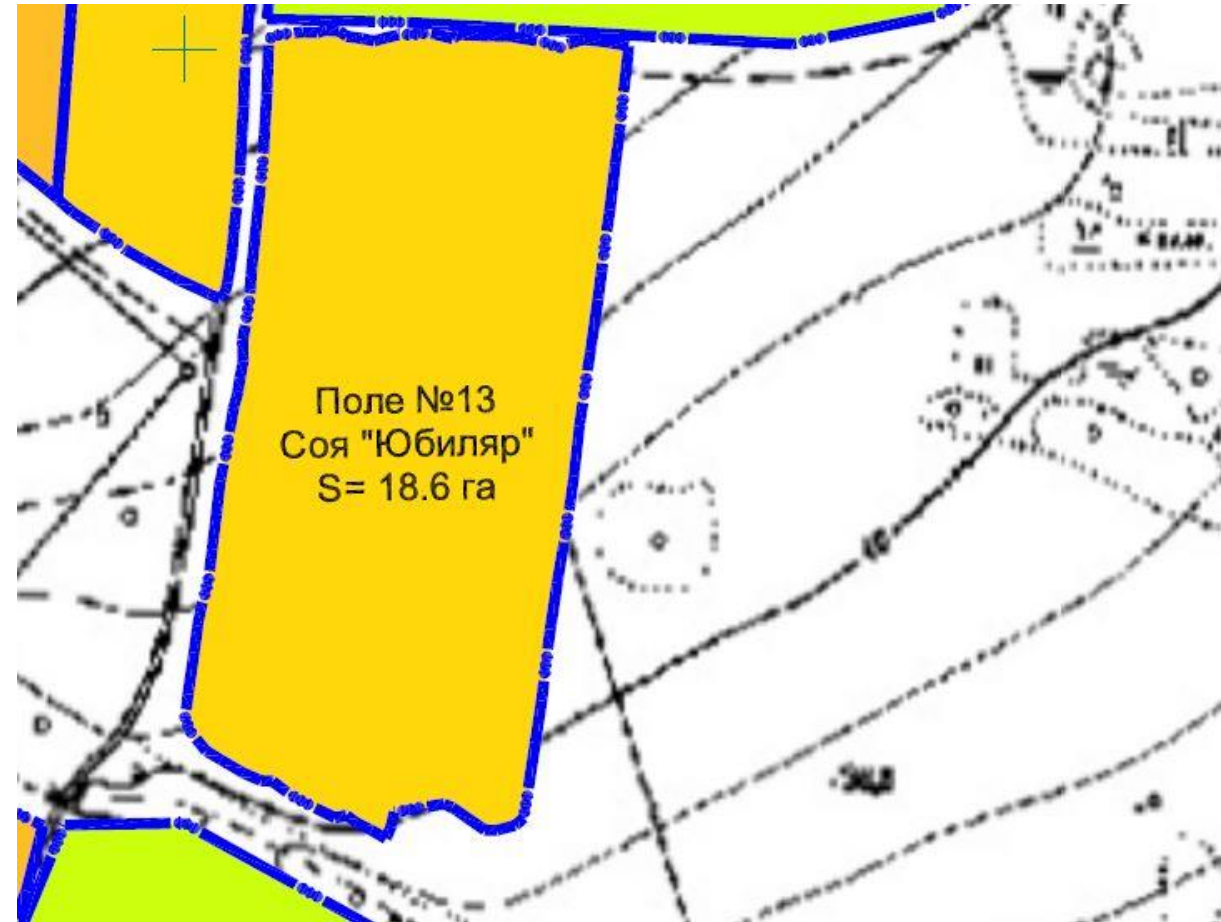
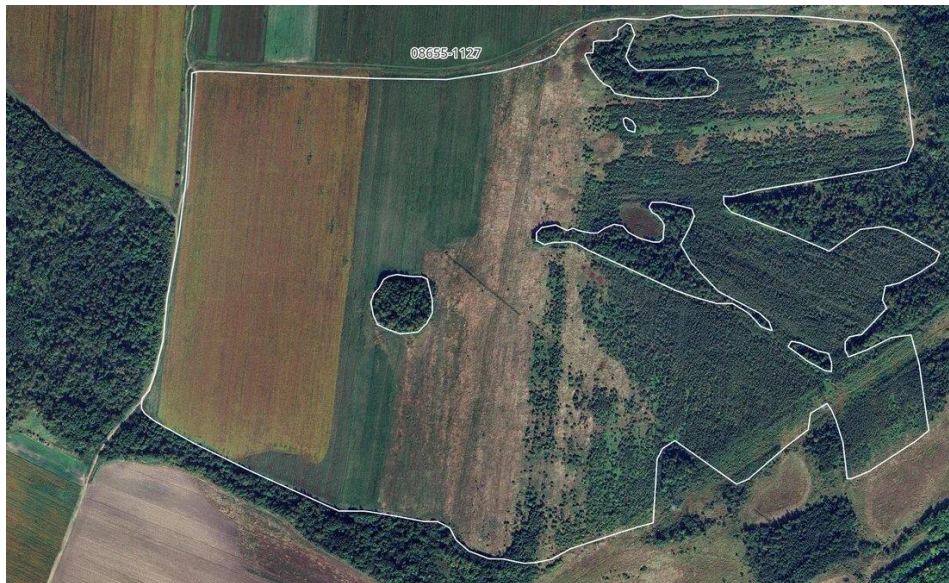
08655-1127

Период сбора	Культура, факт	Культура, план	Площадь, факт, га	Площадь, план, га	Урожайность, ц/га	Валовый сбор, ц
2023	соя	соя	84.12	84.12	0	0
2020	нет данных	овес яровой	84.42	84.12	19.1	1 612.30
2019	нет данных	зерноовощные (злаковые + бобовые культуры)	84.42	84.12	нет данных	нет данных
2018	нет данных	овес яровой	84.05	84.12	19.9	1 672.66
2018	нет данных	овес яровой	84.05	84.12	19.9	1 672.66
2017	нет данных	пшеница яровая	84.05	84.12	18.6	1 563.20
2016	нет данных	пшеница яровая	84.05	84.12	нет данных	нет данных

Поле №13



Соя "Юбиляр"

S = 18.6 га



# В 2025 году внесено сотрудниками ДВ НИИСХ



Действия	Номер севооборота	Сезон	Культура	ОКПД 2	Сорт семян	Площадь сева, га	Цель севообор
 	08655-1127-2025-19	2025	Овес яровой	01.11.33.111 - Семена овса	КАРДИНАЛ	5,30657	Посев/ посадка
 	08655-1127-2025-6	2025	Пшеница мягкая яровая	01.11.12.122 - Семена яровой мягкой пшеницы	АНФЕЯ	3,90278	Посев/ посадка
 	08655-1127-2025-5	2025	Пшеница мягкая яровая	01.11.12.120 - Пшеница яровая мягкая	АНФЕЯ	0,1994	Посев/ посадка
 	08655-1127-2025-8	2025	Пшеница мягкая яровая	01.11.12.122 - Семена яровой мягкой пшеницы	АНФЕЯ	20,1138	Посев/ посадка

# Мониторинг пахотных земель: основные задачи исследований



- **Распознавание и уточнение границ пахотных земель. Классификация земных покровов. Автоматизированная подготовка shp-файлов.**
- Создание и ведение архивов спутниковых данных, необходимых для управления системами цифрового земледелия. Накопление, хранение, обработка информации о севообороте, урожайности культур, границах отдельных полей
- Изучение особенностей вегетационного развития сельхозкультур ДФО. Моделирование сезонного хода индексов вегетации с использованием данных ДЗЗ. Решение задач точного земледелия для Дальнего Востока (оценка межполевой и внутрислолевой однородности, засоренности посевов).
- **Классификация сельскохозяйственных земель южной части Дальнего Востока (мониторинг и идентификация посевов, контроль севооборотов)**
- Развитие методов прогнозирования урожайности сельскохозяйственных культур на основе данных дистанционного зондирования Земли
- Мониторинг и математическое моделирование агроэкосистем: мониторинг растительных покровов и пахотных земель Хабаровского края (оценка ресурсного потенциала) с использованием данных ДЗЗ. Построение эффективных моделей управления почвенными ресурсами

# Данные ДЗЗ

	<b>Sentinel-2</b>	NDVI, EVI, NDWI, LSWI, 10 м
	<b>Sentinel-1</b> <b>ALOS-2 (PALSAR-2)</b>	RVI, DpRVI, VH\VV, 12-20 м
	<b>MODIS Terra</b> (ЦКП ИКИ-МОНИТОРИНГ)	NDVI, LAI - 250 м, 7 days, VEGA- Science
	<b>DJI Mavic3</b> <b>Multispectral</b>	NDVI, EVI, 1 см
	<b>Landsat-8,9</b>	NDVI, EVI, 30 м
	<b>Метеор-М №2</b> (ЦКП ИКИ-МОНИТОРИНГ)	NDVI, EVI, 60 м

# Схема мониторинга пахотных земель с использованием данных ДЗЗ и методов машинного обучения



# Машинное обучение

ФЕДЕРАЛЬНОЕ АГЕНТСТВО  
ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ



НАЦИОНАЛЬНЫЙ  
СТАНДАРТ  
РОССИЙСКОЙ  
ФЕДЕРАЦИИ

ГОСТ Р  
71476—  
2024  
(ИСО/МЭК 22989:  
2022)

Искусственный интеллект

КОНЦЕПЦИИ И ТЕРМИНОЛОГИЯ  
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

(ISO/IEC 22989:2022, Information technology — Artificial intelligence —  
Artificial intelligence concepts and terminology, MOD)

Издание официальное

**3.3.5 машинное обучение; MO (machine learning, ML):** Процесс оптимизации параметров модели с помощью вычислительных методов таким образом, чтобы поведение модели отражало данные и/или опыт.

**3.3.6 алгоритм машинного обучения (machine learning algorithm):** Алгоритм определения параметров модели машинного обучения в соответствии с заданными критериями на основе данных.

*Пример — Рассмотрим задачу определения параметров линейной функции с одной переменной  $y(x) = \theta_0 + \theta_1 x$ , где  $y$  — значение функции,  $x$  — независимая переменная,  $\theta_0$  — свободный член (значение функции при  $x = 0$ ) и  $\theta_1$  — коэффициент. В машинном обучении процесс определения свободного члена и коэффициентов линейной функции известен как линейная регрессия.*

**3.3.7 модель машинного обучения (machine learning model):** Математическая конструкция, генерирующая логический вывод или прогноз на основе входных данных и/или информации.

*Пример — По результатам обучения модели, представленной в виде линейной функции с одной переменной  $y(x) = \theta_0 + \theta_1 x$ , с использованием линейной регрессии, итоговая модель могла бы выглядеть как  $y(x) = 3 + 7x$ .*

# Машинное обучение в разных отраслях науки

## ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ПОСЕЩАЕМОСТИ ФУТБОЛЬНЫХ МАТЧЕЙ

*Сергей Викторович Галицын, доктор педагогических наук, профессор, Олег Зуфарович Зиганишин, доцент, председатель федерации хоккея с мячом Хабаровска, Дальневосточная государственная академия физической культуры, г. Хабаровск; Константин Николаевич Дубровин, инженер, Вычислительный центр Дальневосточного отделения Российской академии наук, Хабаровск; Николай Егорович Еришов, старший научный сотрудник, кандидат физико-математических наук, Дальневосточный научно-исследовательский институт сельского хозяйства, Хабаровск; Павел Анатольевич Ткаченко, кандидат педагогических наук, доцент, Дальневосточная государственная академия*

76

*Ученые записки университета имени П.Ф. Лесгафта. – 2019. – № 10 (176).*

*физической культуры, Хабаровск*

### Аннотация

Для прогнозирования посещаемости матчей Российской Премьер-лиги была создана рекуррентная двухслойная нейронная сеть Элмана. В рамках исследования определены 6 основных факторов, влияющих на посещаемость матчей. Для обучения сети использовались значения факторов и посещаемость 448 матчей РПЛ, а в качестве тестового множества использовались значения факторов и данные о посещаемости с 24 матчей того же турнира в мае 2019 года. Для оценки работоспособности модели было вычислено значение метрики RMSE (корень из среднеквадратической ошибки). По результатам исследования, можно говорить, что модели на базе нейронных сетей могут эффективно использоваться для прогнозирования посещаемости футбольных матчей.

Двухслойная сеть хорошо подходит для задач прогнозирования. Она состоит из входного слоя (значений факторов), промежуточного и выходного слоя. На скрытый слой включает в себя нейроны, преобразующие входные данные в вид, с которым может работать нейрон выходного слоя. На выходном слое находится нейрон, выдающий итоговый результат. Рекуррентная двухслойная сеть Элмана получается из простой двухслойной сети введением обратных связей от выходов внутренних нейронов, что позволяет учесть предысторию наблюдаемых процессов и накопить информацию для выработки правильной стратегии управления. На рисунке 1 приведена схема нейронной сети Элмана.

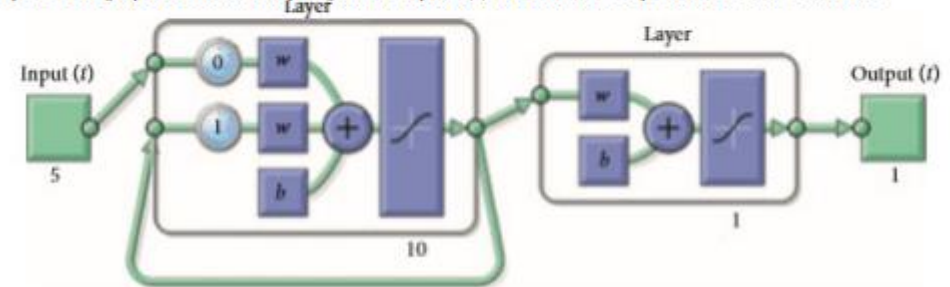


Рисунок 1 – Схема нейронной сети

Таблица 2 – Параметры нейронной сети Элмана

Количество слоёв	2
Количество нейронов на скрытом слое	10
Функция активации на скрытом слое	Гиперболический тангенс
Алгоритм обучения	Обратное распространение ошибки
Оптимизационный метод	Стохастический градиентный спуск с шагом 0,001
Регуляризация	Байесовская регуляризация
Количество эпох обучения	562
Функция потерь	Среднеквадратическое отклонение

# Машинное обучение в разных отраслях науки

Вычислительные технологии

Том 24, № 6, 2019

## Об алгоритмах компьютерного зрения для поиска порывов в сетчатых ограждающих конструкциях\*

А. С. СМАГИН<sup>1,2</sup>, К. Н. ДУБРОВИН<sup>1,†</sup>

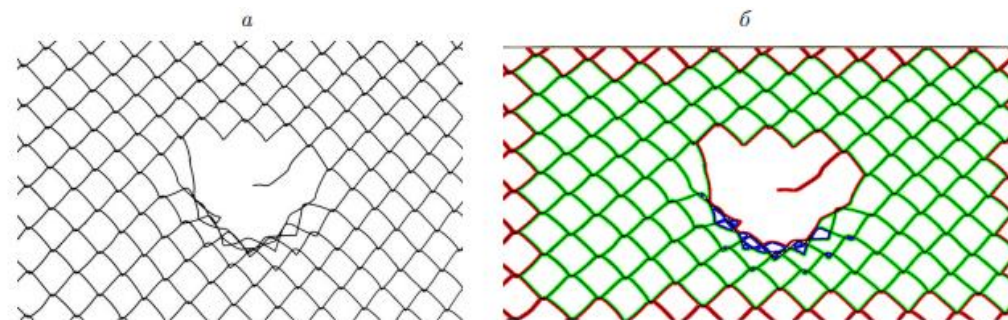
<sup>1</sup>Вычислительный центр ДВО РАН, Хабаровск, Россия

<sup>2</sup>Институт горного дела ДВО РАН, Хабаровск, Россия

<sup>†</sup>Контактный e-mail: nobforward@gmail.com

Рассмотрены вопросы автоматизированного мониторинга состояния сетчатых ограждающих конструкций, которые используются на предприятиях, производящих морские биоресурсы. Предложен алгоритм выделения порывов сетчатых ограждений в подводных условиях с применением методов компьютерного зрения, который реализован в виде комплекса программ на языке Python. Приведены результаты работы программного комплекса. Показано, что методы компьютерного зрения эффективно справляются с определением целостности ячеек сети на слабо- и среднешумленных изображениях. Для работы в более сложных оптических условиях в состав программного комплекса предложено включить нейросетевой модуль.

Проверка целостности сетчатого ограждения осуществляется после завершения фильтрации и получения его не зашумлённого изображения. На этом этапе производится поиск всех контуров на изображении при помощи функции FindContours библиотеки OpenCV. Далее проводится их классификация. Если контур является замкнутым многоугольником и его периметр находится в диапазоне от  $0.5P$  до  $2P$ , где  $P$  – средний периметр, то этот контур классифицируется как целая ячейка сети и помечается зелёным цветом. Незамкнутый контур или замкнутый контур с периметром больше, чем  $2P$ , помечается красным цветом и классифицируется, как порыв. Замкнутый контур с малым периметром (меньше  $0.5P$ ) воспринимается системой как сильно деформированный (для таких ячеек велика вероятность порыва в будущем) и помечается синим цветом. На рисунках 3а и 3б представлен результат работы программы, определяющей целостность сети.



# Прогнозирование урожайности: выбор факторов

$x_1$  – максимальное значение NDVI по маске пахотных земель муниципалитета (Bege-Science).

По данным реанализа (Bege-Science):

$x_2$  – гидротермический коэффициент Селянинова;

$x_3$  – количество дней с активной вегетацией (с среднесуточной температурой выше 10°C) ;

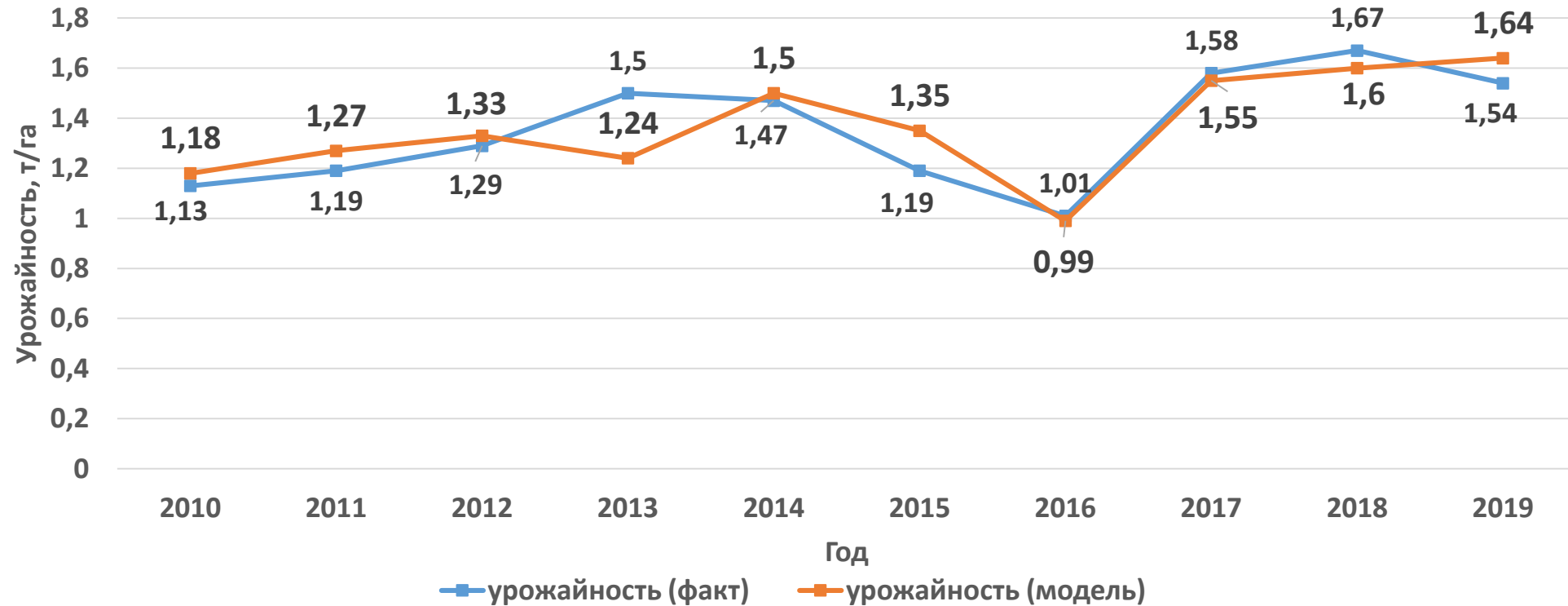
$x_4$  – сумма среднесуточных температур верхнего слоя почвы (глубина до 10 см) ,°C;

$x_5$  – средняя влажность верхнего слоя почвы (глубина до 10 см), %;

$x_6$  – показатель фотосинтетически активной радиации (ГДж\*м<sup>2</sup>).

Predicting soybean yield at the regional scale using remote sensing and climatic data /  
A. Stepanov, T. Aseeva, K. Dubrovin, A. Sorokin // Remote Sensing. – 2020. – Vol. 12,  
No. 12. – P. 1936. – DOI 10.3390/rs12121936.

# Реальная и расчётная урожайность сои в Хабаровском районе

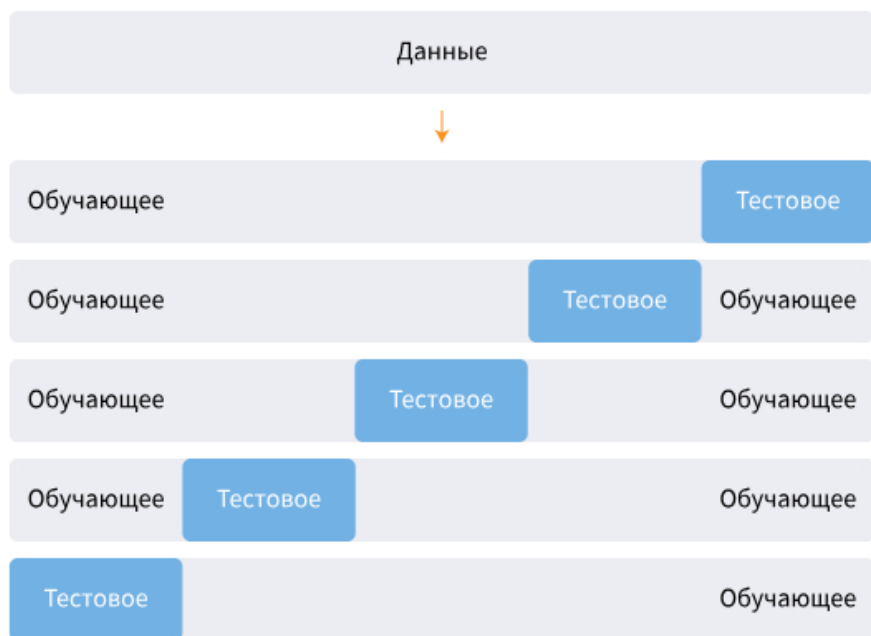


$$y = -8,24 + 9,39x_1 + 0,036x_3$$

$$R^2 = 0,72$$

Ошибка **кросс-валидации** 7,8 %. Ошибка прогнозирования на 2019 год – 6,3 %.

# Кросс-валидация



	$R^2_{cv}$	RMSE <sub>cv</sub> (t/ha)	MAPE <sub>cv</sub> , %	APE <sub>cv</sub> , %
2010	0.70	0.11	6.1	3.5
2011	0.73	0.11	5.8	7.9
2012	0.73	0.12	6.5	4.9
2013	0.81	0.07	4.0	22.7
2014	0.72	0.12	6.5	1.2
2015	0.79	0.10	5.1	18.3
2016	0.62	0.12	6.4	2.2
2017	0.70	0.11	6.1	2.9
2018	0.68	0.11	5.7	8.2
2019	0.72	0.11	6.2	5.8
$\bar{x}$		0.11	5.9	7.8
$\Delta\bar{x}$		0.01	0.6	5.1

**Оценка производительности:** Позволяет более точно оценить, как модель будет работать на новых, ранее невидимых данных, и избежать переобучения.

**Сравнение моделей:** Помогает выбрать лучшую модель для конкретной задачи, сравнивая их производительность на одних и тех же данных.

**Эффективное использование данных:** Позволяет максимально полно использовать имеющиеся данные, так как каждый объект данных участвует и в обучении, и в тестировании.

# Раннее прогнозирование урожайности сои

Stepanov, A. Function fitting for modeling seasonal normalized difference vegetation index time series and early forecasting of soybean yield / A. Stepanov, K. Dubrovin, A. Sorokin // The Crop Journal. – 2022. – Vol. 10, No. 5. – P. 1452-1459. – DOI 10.1016/j.cj.2021.12.013.

## Регрессионные модели для прогнозирования урожайности на районном уровне

$$y = -3,41 + 3,53NDVI_{max} + 0,015D$$

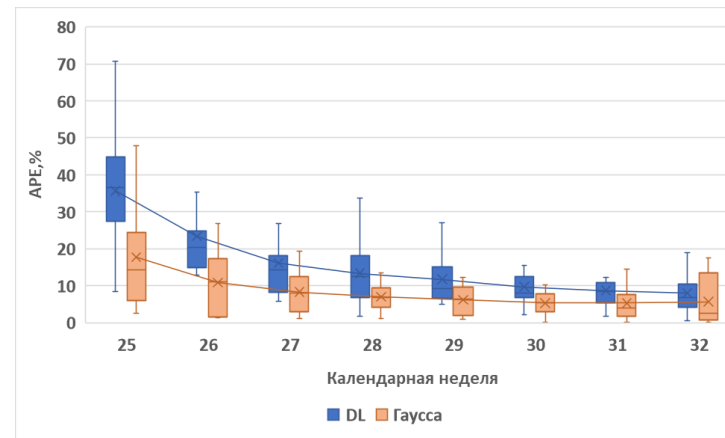
(Ленинский р-н ЕАО)

$$y = -2,23 + 1,94NDVI_{max} + 0,017D$$

(Октябрьский р-н ЕАО)

- $y$  – урожайность сои, т/га;
- $NDVI_{max}$  – значения максимума  $NDVI$  пахотных земель;
- $D$  – количество дней с активными температурами от начала вегетации до недели с  $NDVI_{max}$ .

## Оценка точности прогнозирования урожайности



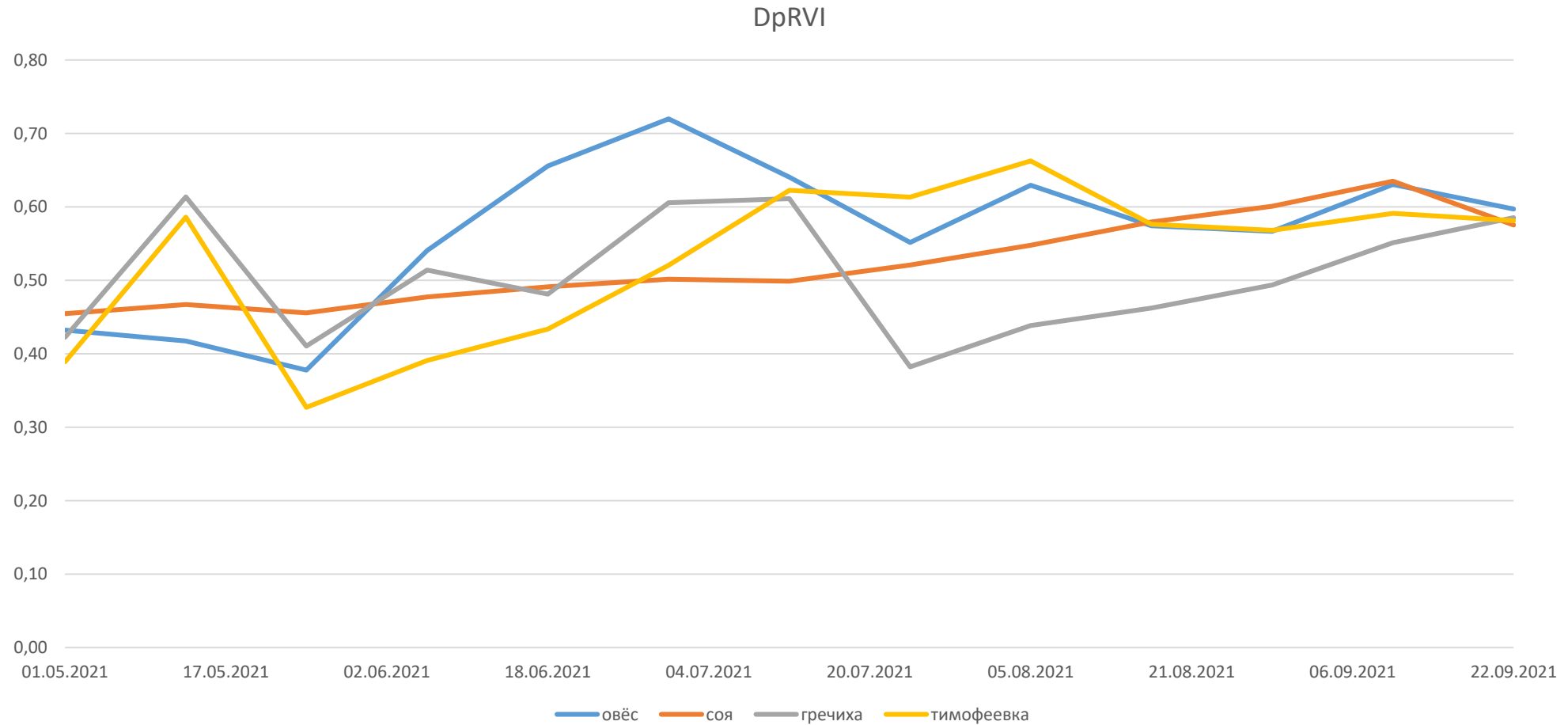
Неделя	P-value	MAPE, %	
		Гаусса	DL
25	P<0,05	14,9	32,8
26	P<0,05	12,9	22,5
27	P<0,05	9,3	16,3
28	P<0,05	7,3	12,7
29	P<0,05	6,2	11,7
30	P<0,05	5,1	9,9

Уже на 27 неделе (первая декада июля), за 5–6 недель до достижения максимума возможно прогнозирование урожайности с точностью свыше 90 %

# РАЗРАБОТКА ПОДХОДОВ К КЛАССИФИКАЦИИ ПАХОТНЫХ ЗЕМЕЛЬ С ИСПОЛЬЗОВАНИЕМ ИНДЕКСА DpRVI (2021)

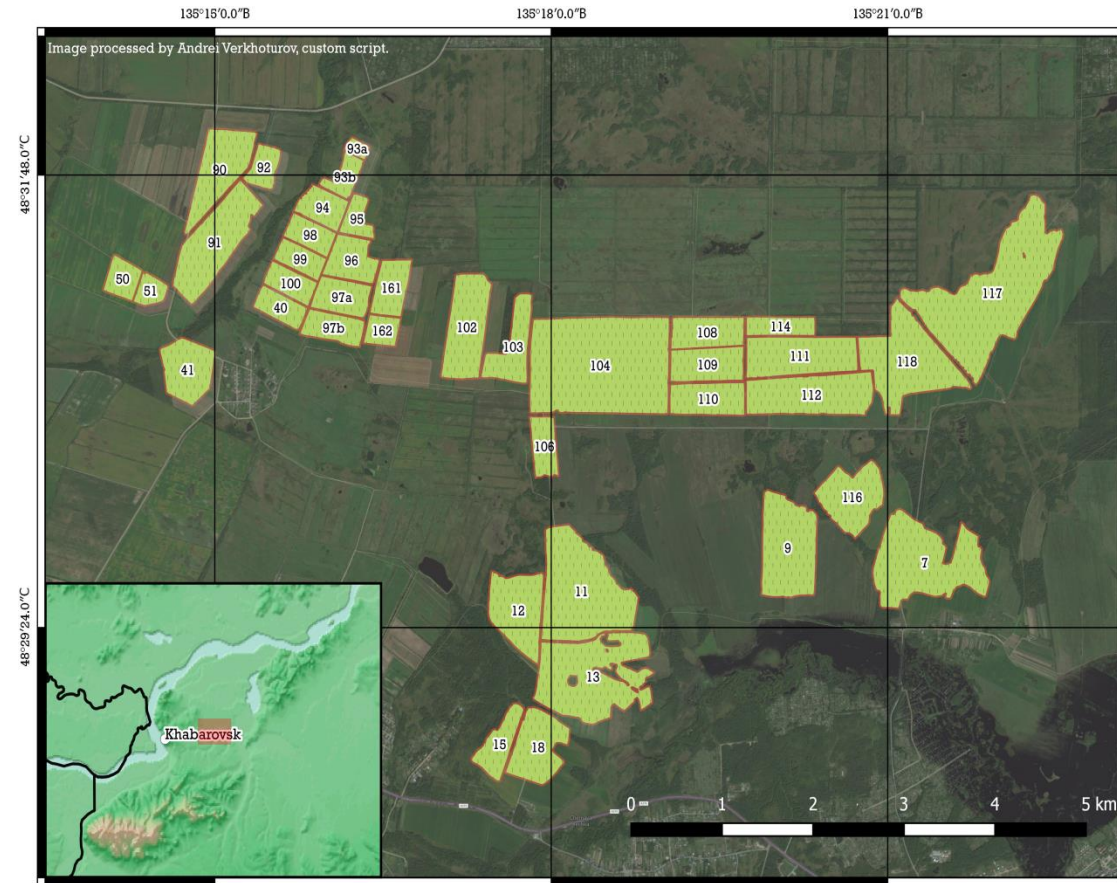
- Построение кривых сезонного хода вегетационных индексов для различных с/х культур, фильтрация и аппроксимация данных, изучение возможности применения вегетационных индексов для идентификации культур
- Проведение классификации пахотных земель разными методами машинного обучения, оценка точности, выбор оптимального метода

# Сезонный ход DpRVI для разных культур (отдельные поля, 2021)



# Предобработка исходных данных

- Исходные данные – ряды сезонного хода DpRVI для каждого пикселя 36 опытных полей ДВ НИИСХ (общая площадь 990 га). Срок исследования – с 2017 по 2021 год.
- Фильтрация по датам – **в каждом ряду по 13 значений**: с начала мая по конец сентября.
- **Фильтрация нерелевантных пикселей** – для каждого поля на каждую дату рассчитывается средний DpRVI и СКО. Значения, не попадающие в  $2\sigma$ -интервал, удаляются. **Пиксели с пропусками в данных не попадают в обучающий датасет.**

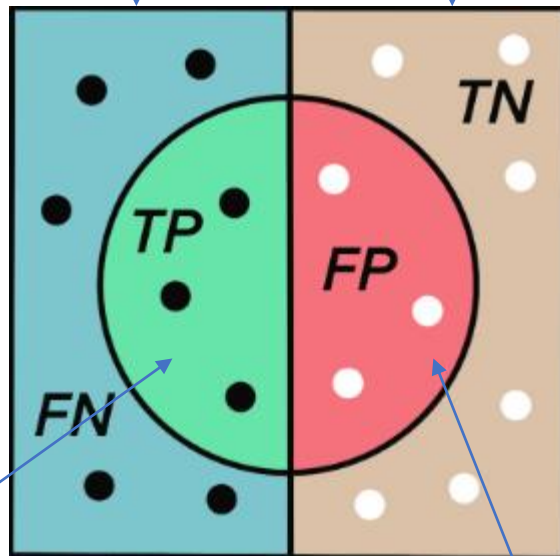


# Данные для классификации

- Обучающая выборка – ряды значений DpRVI для **пикселей 20 опытных полей**. Период – 2018-2021 год. Всего 39528 рядов.
- Тестовая выборка – значения DpRVI для **пикселей 16 опытных полей**. Период – 2017-2021 год. Всего 14048 рядов.
- Все пиксели разбиты на 4 класса: 1 – овёс, 2 – соя, 3 – гречиха, 4 – тимофеевка луговая.

# Оценка точности классификации

Объекты класса 1    Объекты класса 2



Объекты, отнесённые  
классификатором  
к классу 1

Объекты, отнесённые  
классификатором  
к классу 2

Общая точность:

$$OA = \frac{TP + TN}{TP + TN + FN + FP} * 100\%$$

Рассчитывается для каждого класса отдельно:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * Recall * Precision}{Recall + Precision}$$

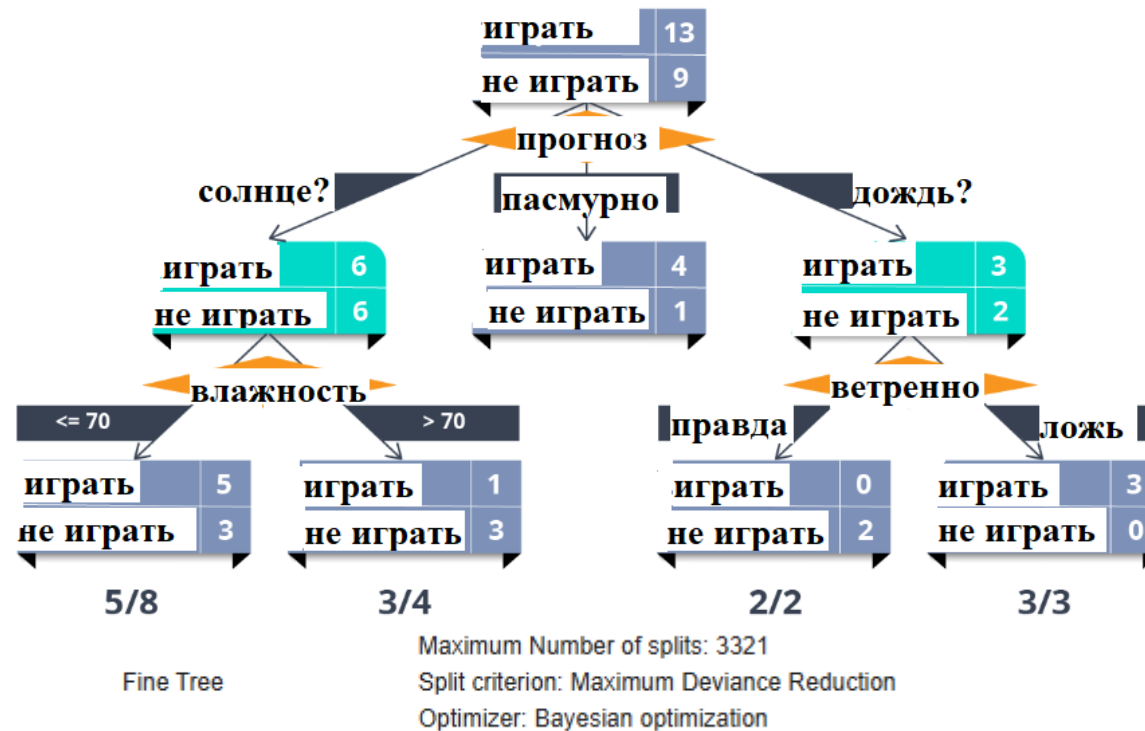
$$\kappa = \frac{N \sum_{i=1}^r TP_i - \sum_{i=1}^r X_i + X_{+i}}{N^2 - \sum_{i=1}^r X_i + X_{+i}} \text{ (каппа Коэна)}$$

$N$  – общее число пикселей,  $r$  – количество классов,  $TP_i$  – количество верно классифицированных пикселей класса  $i$ ,  $X_{i+}$  – Количество пикселей класса  $i$  в исходном наборе,  $X_{+i}$  – Количество пикселей класса  $i$  по результатам классификации,  $FP_i$  – количество пикселей, неверно отнесённых к классу  $i$ ,  $FN_i$  – количество пикселей класса  $i$ , неверно отнесённых к другим классам.

# Выбор метода классификации

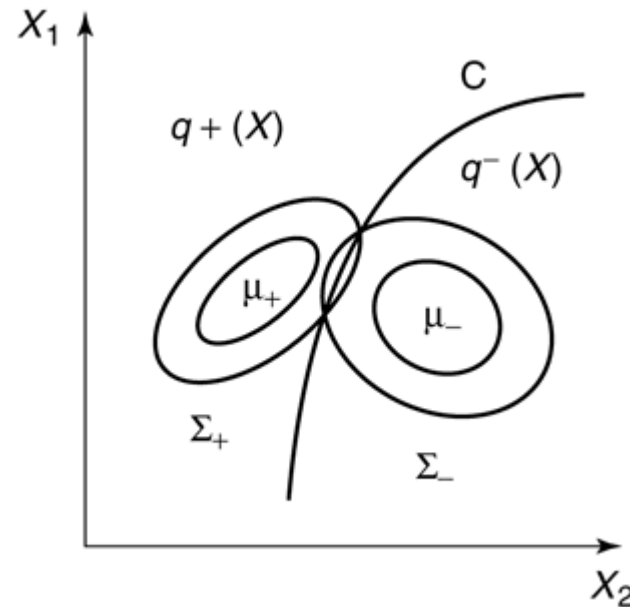
Метод	Время обучения, с.	Точность на обучающем множестве, %	Точность на тестовом множестве, %
Деревья решений	62	90,8	71,1
Квадратичный дискриминантный анализ	30	90,9	81,9
Наивный байесовский классификатор	3612	89,0	73,4
к ближайших соседей	526	99,8	72,4
RUSBoost	1774	95,9	74,4

# Дерево решений



Chen, Y.; Lu, D.; Moran, E.; Batistella, M.; Dutra, L.V.; Sanches, I.D.; da Silva, R.F.B.; Huang, J.; Luiz, A.J.B. Mapping croplands, cropping patterns, and crop types using MODIS time-series data. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, 69, 133–147.

# Квадратичный дискриминантный анализ



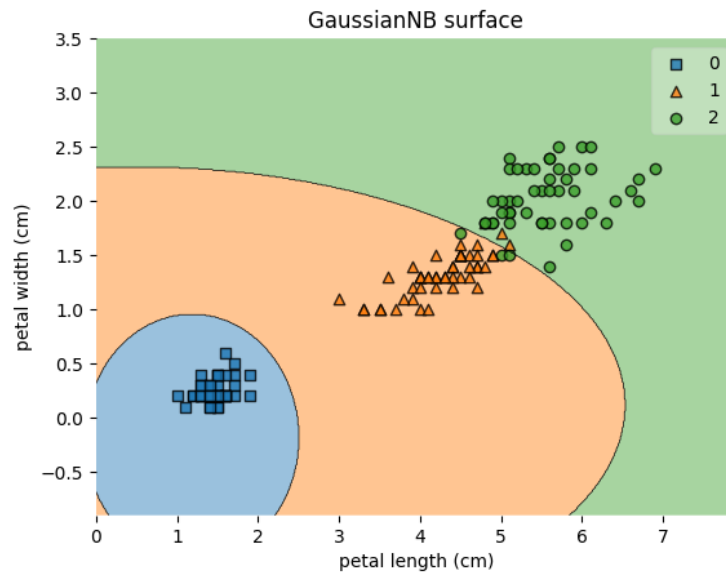
Quadratic Discriminant Analysis

Covariance Structure: Full

Ghojogh, B.; Crowley, M. Linear and Quadratic Discriminant Analysis: Tutorial. *arXiv* **2019**, arXiv:1906.02590.

Choubin, B.; Moradi, E.; Golshan, M.; Adamowski, J.; Sajedi-Hosseini, F.; Mosavi, A. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci. Total Environ.* **2019**, 651, 2087–2096.

# Наивный байесовский метод



$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B)}$$

Gaussian Naive Bayes

Kernel type: Triangle  
Support: Unbounded

Zhang, H. The Optimality of Naive Bayes. In *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference*; FLAIRS: Miami Beach, FL, USA, 2004.

Castro, W.; De-la-Torre, M.; Avila-George, H.; Torres-Jimenez, J.; Guivin, A.; Acevedo-Juárez, B. Amazonian cacao-clone nibs discrimination using NIR spectroscopy coupled to naïve Bayes classifier and a new waveband selection approach. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *270*, 120815.

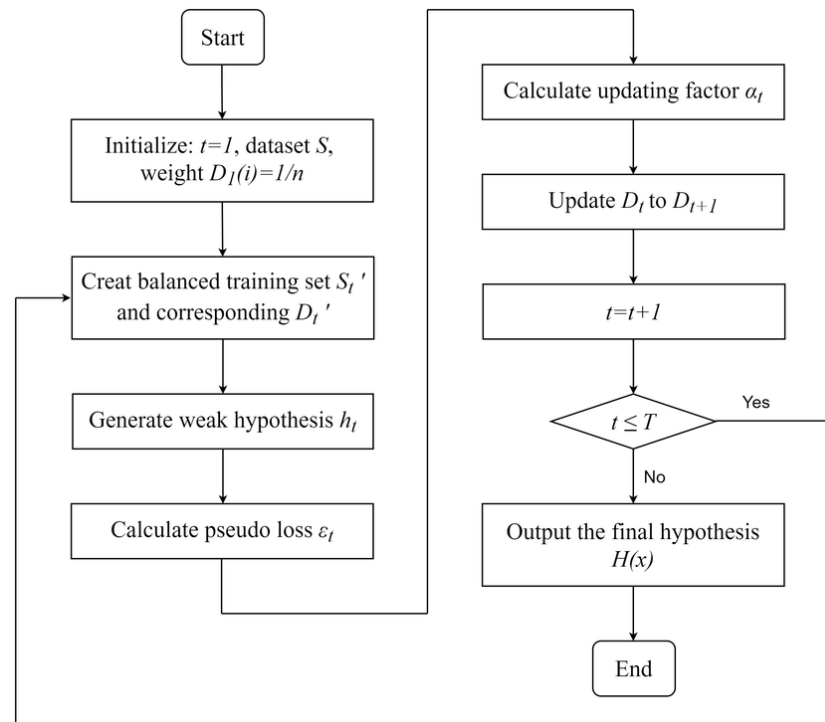
# Метод k ближайших соседей



Hechenbichler, K.; Schliep, K. Weighted k-Nearest-Neighbor Techniques and Ordinal Classification. Master's Thesis, Massey University, Palmerston North, New Zealand, 2004.

Wilson, B.T.; Lister, A.J.; Riemann, R.I. A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *For. Ecol. Manag.* **2012**, 271, 182–198.

# Метод RUSBoost



RUSBoost

Maximum number of splits: 5455  
Number of learners: 72  
Learning rate: 0.1

Seiffert, C.; Khoshgoftaar, T.; Van Hulse, J.; Napolitano, A. RUSBoost: Improving Classification Performance when Training Data is Skewed. In Proceedings of the 19th IEEE International Conference on Pattern Recognition, Tampa, FL, USA, 8–11 December 2008.

# Результаты классификации (Quadratic Discriminant)

- ОА (обучающая выборка, **10-кратная кросс-валидация**) = 90,9 %  
(попиксельно)
- ОА (тестовая выборка) = 81,9%
- $UA_1 = 61,3\%$ ;  $UA_2 = 95,7\%$ ;  $UA_3 = 56,0\%$ ;  $UA_4 = 61,9\%$
- $PA_1 = 74,1\%$ ;  $PA_2 = 90,0\%$ ;  $PA_3 = 68,9\%$ ;  $PA_4 = 57,6\%$
- $K = 0,67$  (существенное соответствие)
- Правильно было классифицировано 15 полей из 16

*1 – овёс, 2 – соя, 3 – гречиха, 4 – тимофеевка луговая*





# Результаты классификации (Quadratic Discriminant)

овёс	1316	41	105	313	74.1%	25.9%
соя	69	8226	600	246	90.0%	10.0%
гречиха	3	314	899	89	68.9%	31.1%
тимоф	759	14	2	1052	57.6%	42.4%
	овёс	соя	гречиха	тимоф		

**Матрица несоответствий**

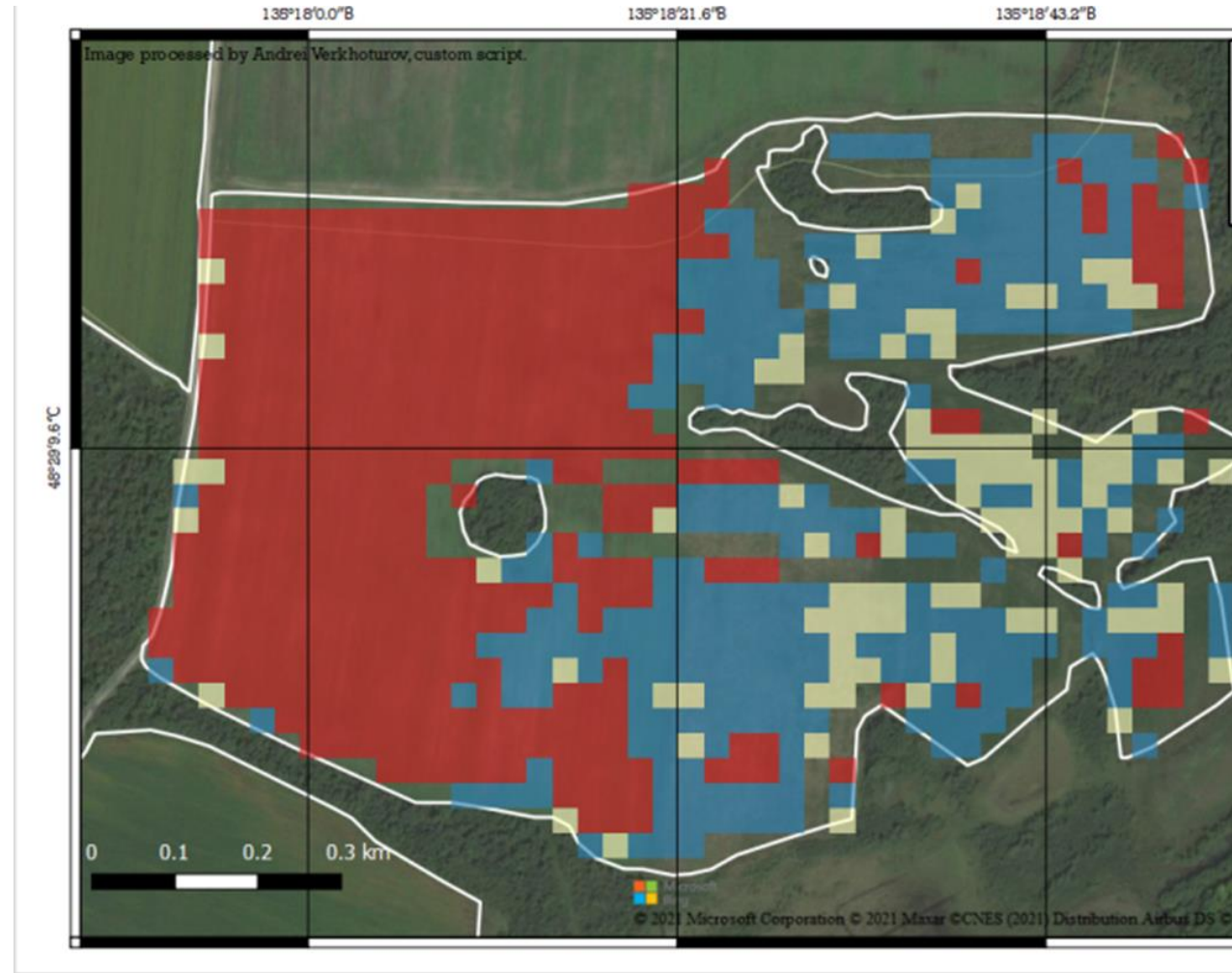


Опытное поле № 90: овёс  
с подсевом тимофеевки, 2021 год

	соя		овёс
	гречиха		тимофеевка

# Выявление ошибок при определении реальных границ поля

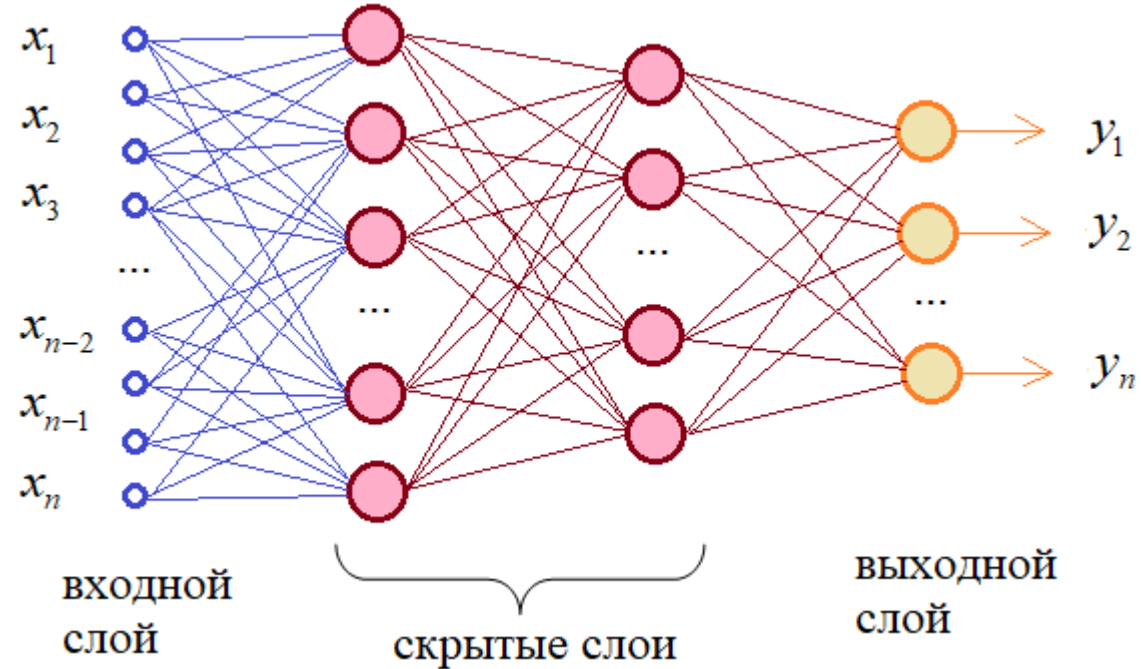
- Не всегда границы сева совпадают с кадастровыми границами поля



# Нейронные сети

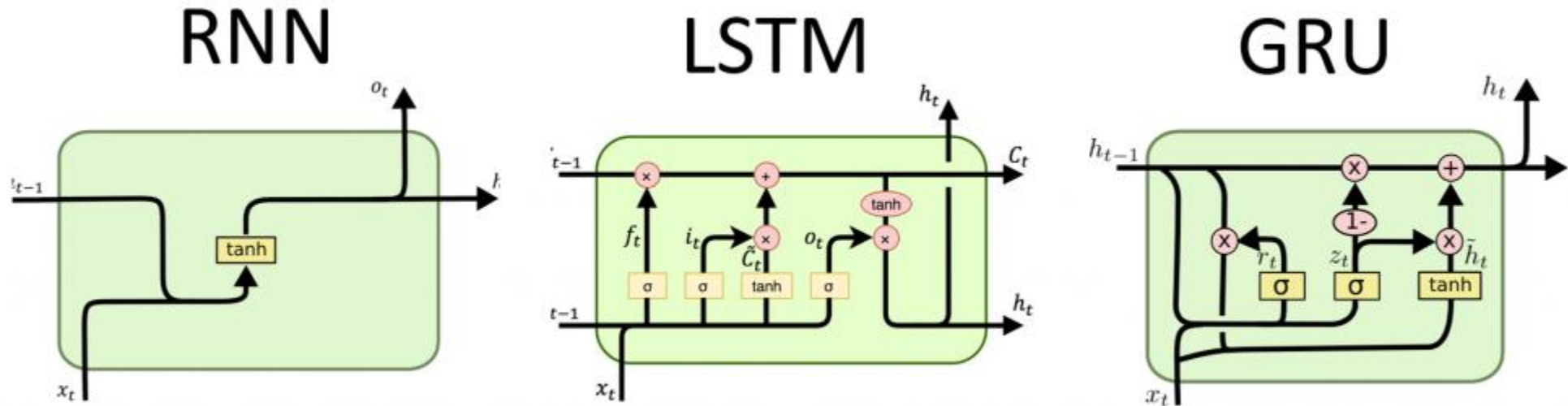
- Нейронная сеть; искусственная нейронная сеть (neural network, NN, neural net, artificial neural network): Сеть из двух или более слоев, состоящих из нейронов, соединенных взвешенными связями с регулируемыми весовыми коэффициентами; при этом каждый нейрон получает входные данные и вырабатывает результат.
- Нейрон (neuron): Базовый элемент, получающий одно или несколько входных значений и вырабатывающий выходное значение посредством комбинирования входных значений и применения функции активации к результату комбинирования.

# Полносвязная нейронная сеть



Николенко С., Кадурын А., Архангельская Е. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.: ил. — (Серия «Библиотека программиста»). ISBN 978-5-496-02536-2.

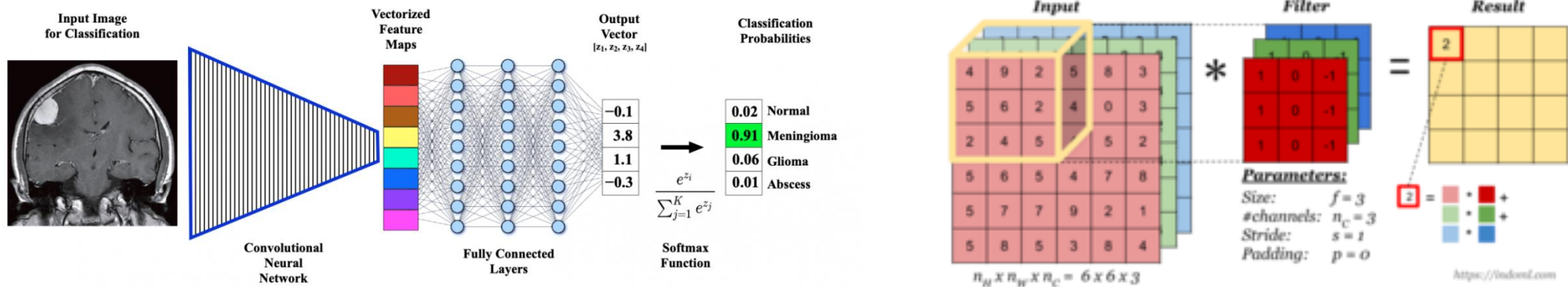
# Рекуррентная нейронная сеть



Рекуррентная нейронная сеть (recurrent neural network, RNN): Нейронная сеть, в которой как выходные данные предыдущего слоя, так и результаты предыдущего шага вычислений подаются на вход текущему слою.

S. Ge, O. Antropov, W. Su, H. Gu and J. Praks, "Deep Recurrent Neural Networks for Land-Cover Classification Using Sentinel-1 INSAR Time Series," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, pp. 473-476, doi: 10.1109/IGARSS.2019.8900088.

# Сверточная нейронная сеть

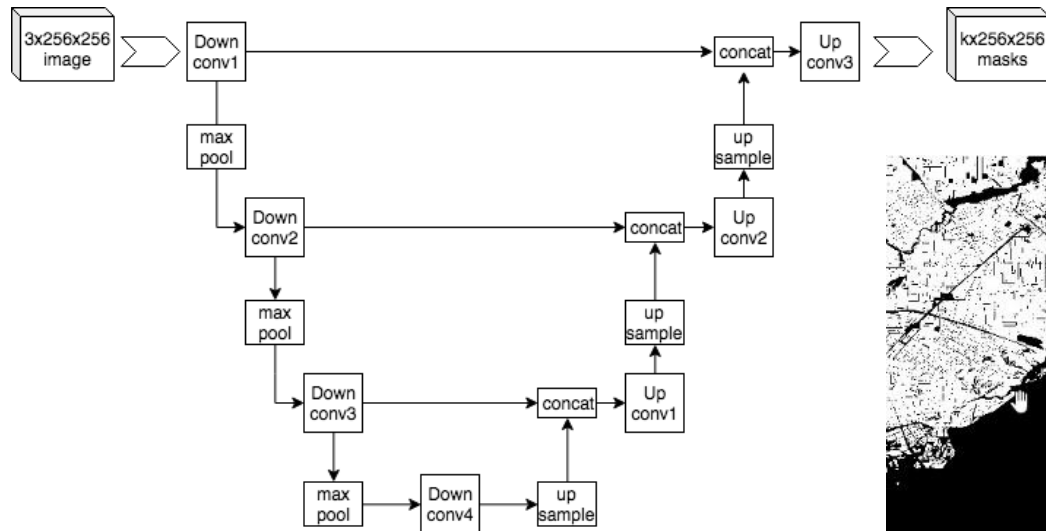


Yan Y., Ryu Y. Exploring Google Street View with deep learning for crop type mapping // ISPRS Journal of Photogrammetry and Remote Sensing. 2021. Vol. 171. P. 278-296. Doi: 10.1016/j.isprsjprs.2020.11.022.

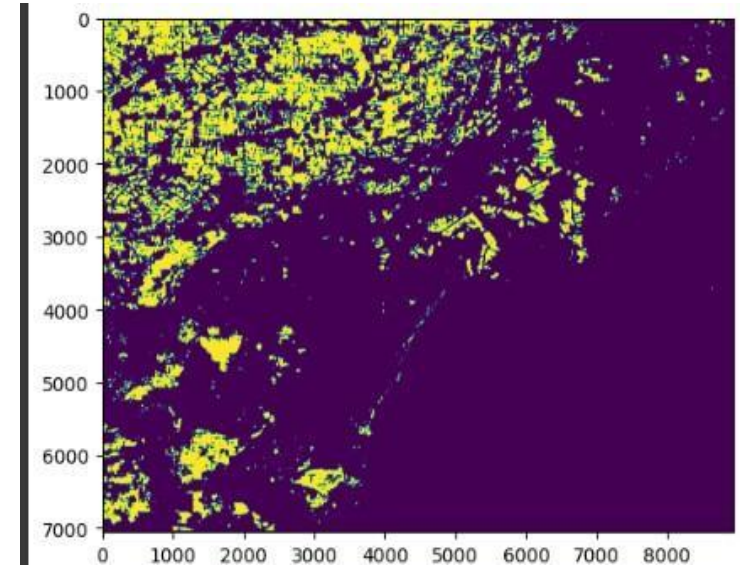
Shi, Z.; Fan, J.; Du, Y.; Zhou, Y.; Zhang, Y. LULC-SegNet: Enhancing Land Use and Land Cover Semantic Segmentation with Denoising Diffusion Feature Fusion. Remote Sens. 2024, 16, 4573.

<https://doi.org/10.3390/rs16234573>

# Архитектура U-Net



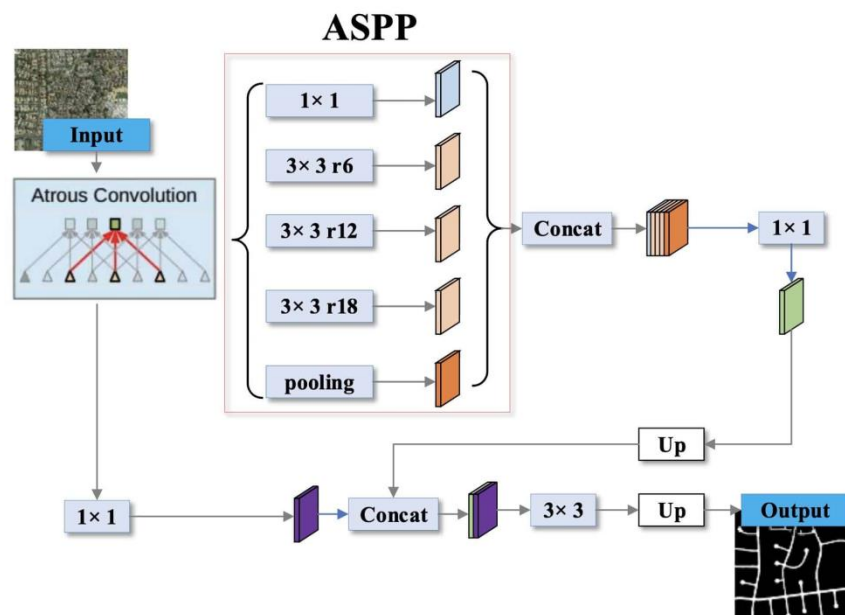
Результаты работы построенного алгоритма для Вяземского района (1 - исходная маска, 2 - предсказанная маска)



Deep one-class crop classification via positive and unlabeled learning for multi-modal satellite imagery / Lei L., Wang X., Zhong Y., et al. // International Journal of Applied Earth Observation and Geoinformation. 2021. Vol. 105. 102598.

Утляков, А. А. Использование данных ДЗЗ для автоматической сегментации сельскохозяйственных полей с применением U-Net / А. А. Утляков // Информационные технологии и высокопроизводительные вычисления : Материалы VII Международной научно- практической конференции, Хабаровск, 11–13 сентября 2023 года. – Хабаровск: Федеральное государственное бюджетное учреждение науки Хабаровский Федеральный исследовательский центр Дальневосточного отделения Российской академии наук, 2023. – С. 223-226.

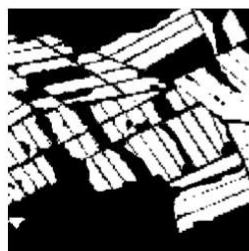
# Архитектура DeepLab



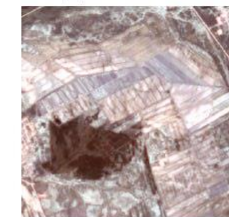
Загруженный тестовый снимок  
2024 года:



Полученная маска 2024 года:



Загруженный тестовый снимок  
2023 года:



Полученная маска 2023 года:



<https://github.com/leimao/DeepLab-V3> - репозиторий Google с кодом DeepLab

Manzanarez, S.; Manian, V.; Santos, M. Land Use Land Cover Labeling of GLOBE Images Using a Deep Learning Fusion Model. Sensors 2022, 22, 6895.

<https://doi.org/10.3390/s22186895>

# Архитектура YOLO

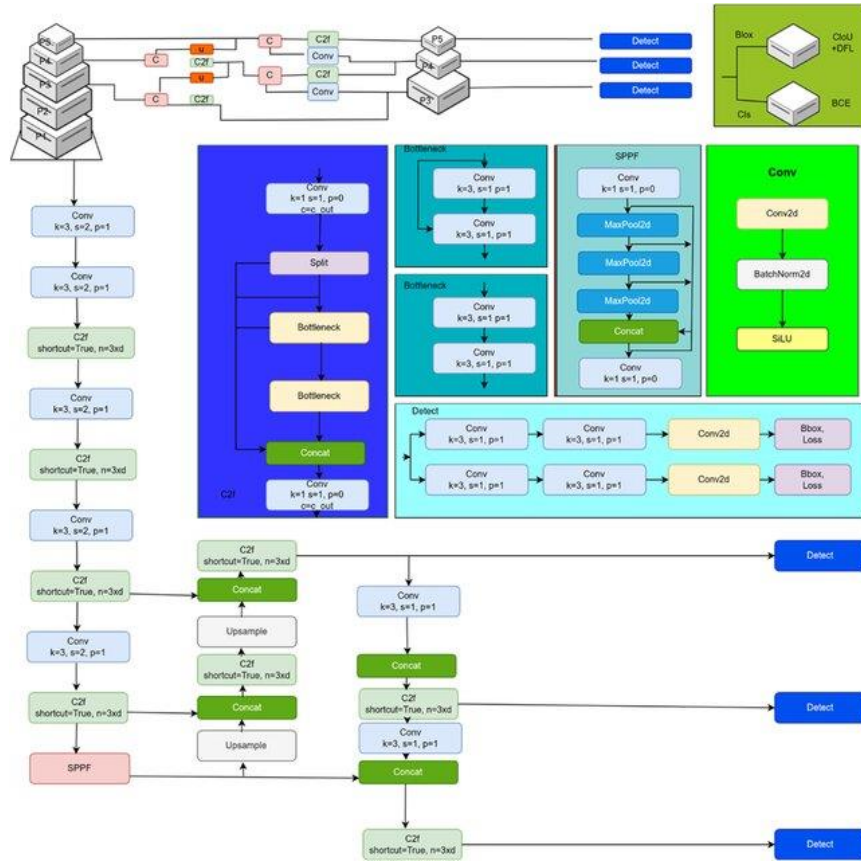


Рисунок 1 – Ручная разметка



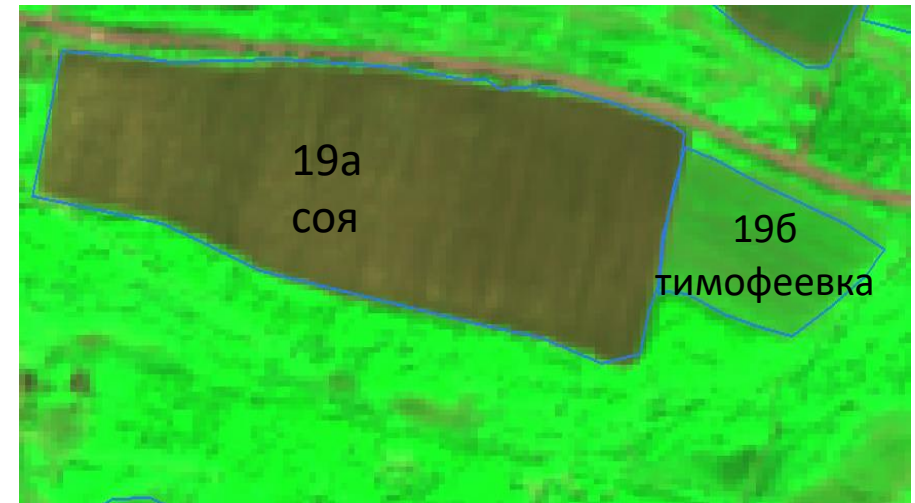
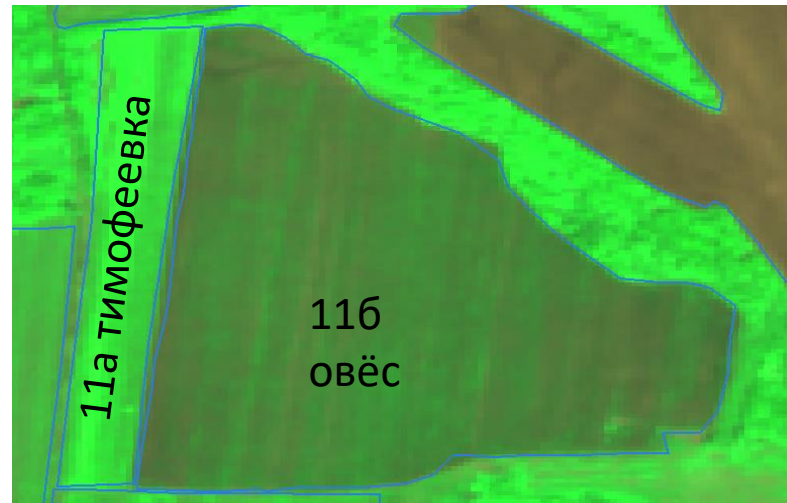
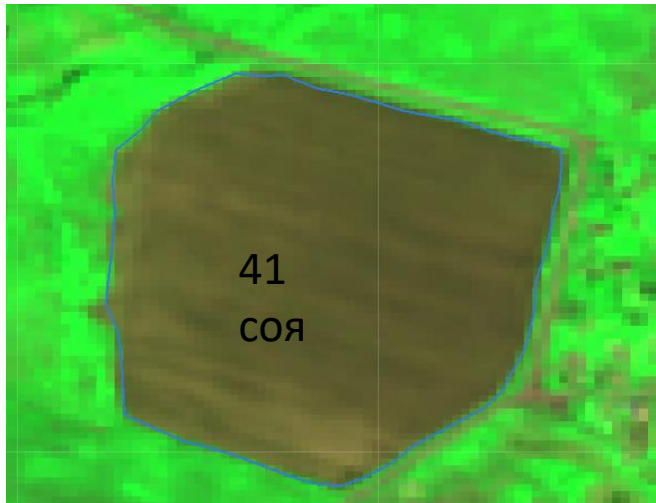
Рисунок 2 – Автоматическая разметка

Поляков, А. Н. Использование модели YOLO для сегментации пахотных полей / А. Н. Поляков, К. Э. Полякова // Вычислительные технологии и прикладная математика : Материалы III научной конференции с международным участием, Комсомольск-на-Амуре, 07–11 октября 2024 года. – Комсомольск-на-Амуре: Комсомольский-на-Амуре государственный университет, 2024. – С. 283-286.

Guo, X., et al. (2022). YOLOv5: Improving Object Detection and Segmentation Performance with Cross-Stage Feature Pyramid Networks. Computer Vision and Image Understanding, 217, 103374.

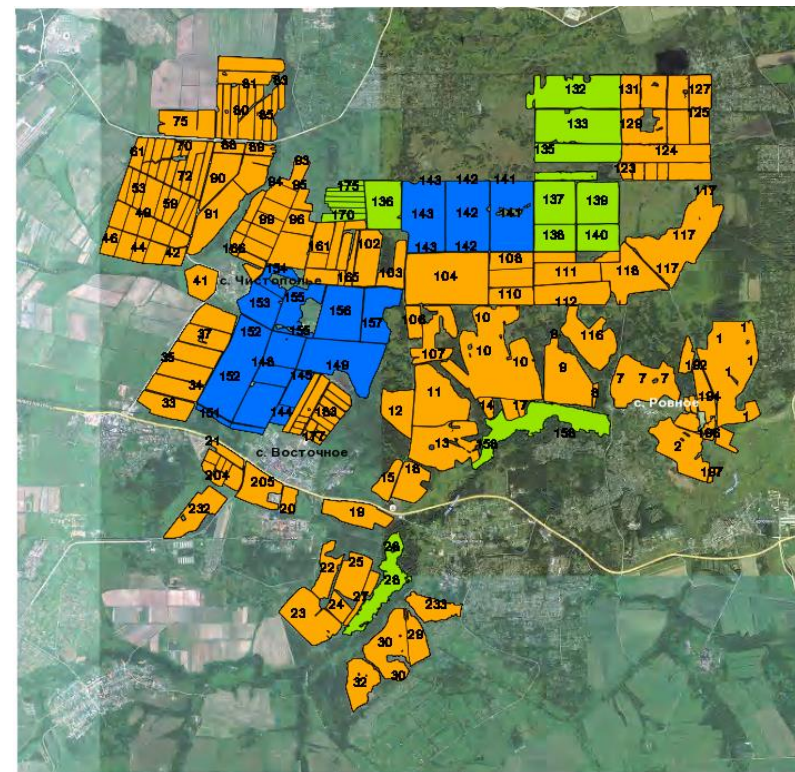
# Реальные контуры полей

- Контуры полей формировались по результатам анализа спутниковых снимков Sentinel-2 и наземной экспертизы



# КЛАССИФИКАЦИЯ ЗЕМЕЛЬ СЕЛЬСКОХОЗЯЙСТВЕННОГО НАЗНАЧЕНИЯ ДАЛЬНЕГО ВОСТОКА С ИСПОЛЬЗОВАНИЕМ ВРЕМЕННЫХ РЯДОВ ВЕГЕТАЦИОННЫХ ИНДЕКСОВ (2022)

- Поля Дальневосточного научно-исследовательского института сельского хозяйства (Хабаровский район, Хабаровский край), 2021 год



	Залежь	Пар	Соя	Гречиха	Тимофеевка	Овёс	Всего
Кол-во полей	60	45	46	14	2	3	170
Площадь, га	1577	411	1294	293	60	27	3662

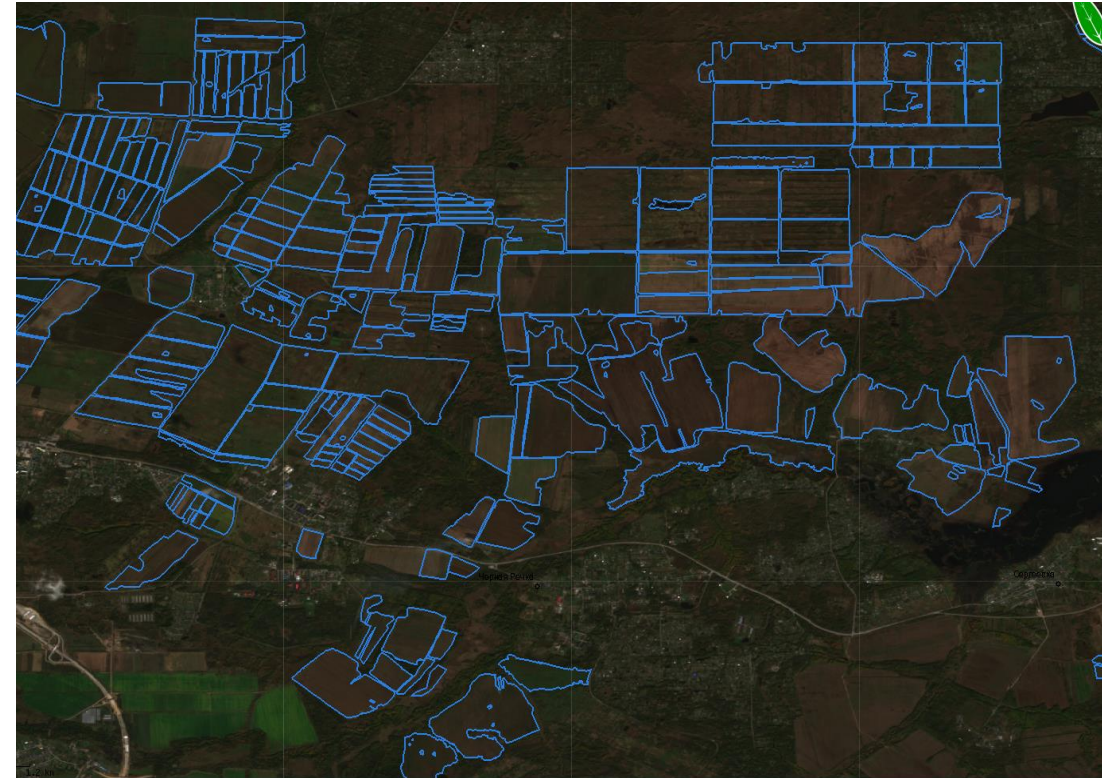
# Предобработка исходных данных

Исходные данные:

170 полей

- А) ряды сезонного хода **NDVI** для каждого пикселя (23 снимка Sentinel-2 с 1.05 по 31.10.2021 г.)
- Б) ряды сезонного хода **DpRVI** для каждого пикселя (16 снимков Sentinel-1 с 1.05 по 28.10.2021 г.)

Фильтрация нерелевантных пикселей – для каждого поля на каждую дату рассчитывается средние значения и СКО для NDVI и DpRVI. Значения, не попадающие в  $2\sigma$ -интервал, удаляются. Пиксели **с менее чем половиной значений** не попадают в датасет.



# Каскадная схема классификации

- Определение залежи и парующих земель



- Определение полей с соей (самая распространённая культура, 1294 га)

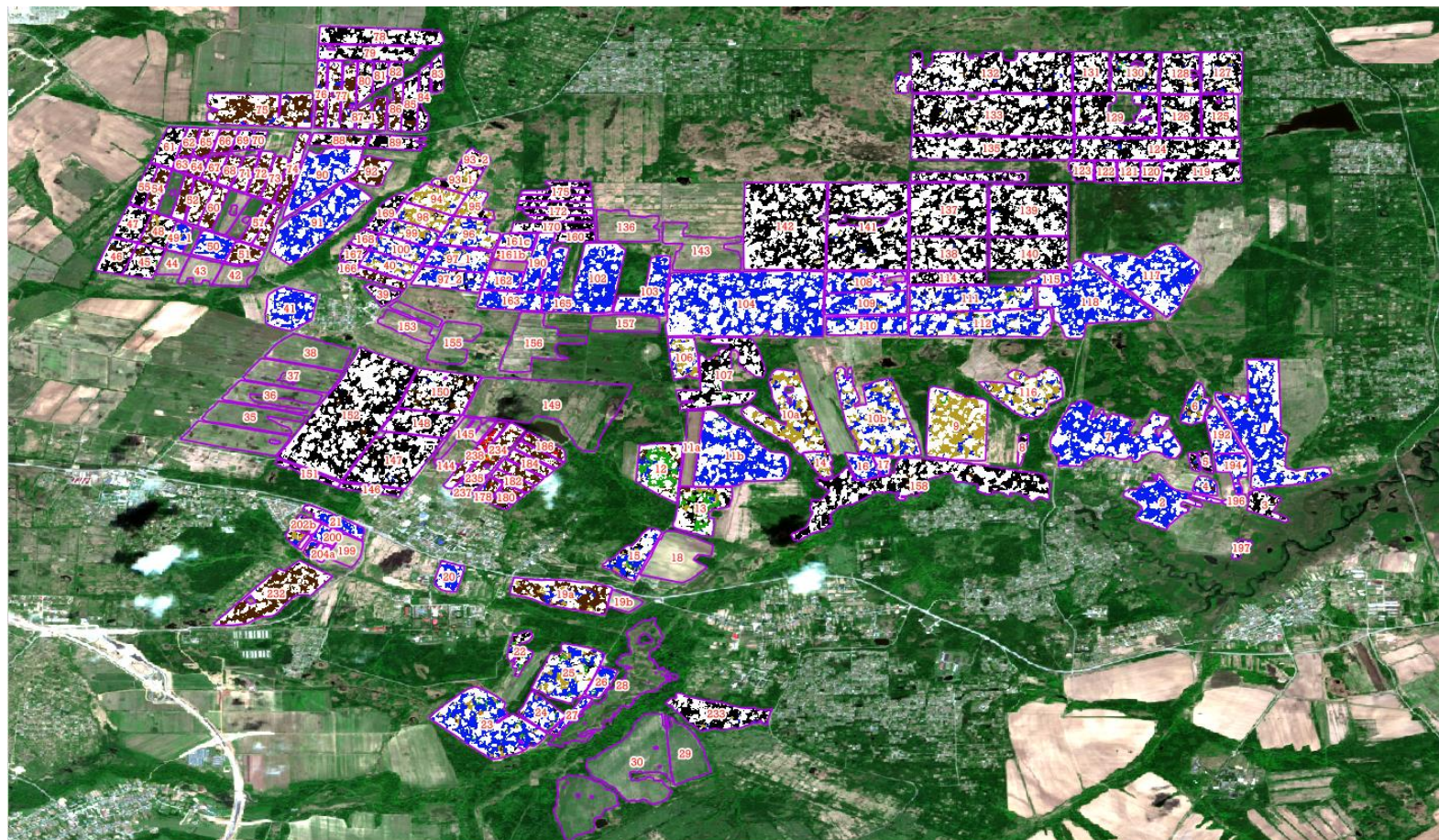


- Определение полей с гречихой (293 га)



- Классификация оставшихся пикселей: тимофеевка и овёс

# Карта-схема пахотных земель ДВ НИИСХ



- соя 
- гречиха 
- тимофеевка 
- овёс 
- залежь 
- пар 

# Результаты классификации

- $OA_{NDVI} = 94,0 \%$

- $F1_{\text{пар}} = 0,85$

- $F1_{\text{залежь}} = 0,96$

- $F1_{\text{соя}} = 0,97$

- $F1_{\text{гречиха}} = 0,93$

- $F1_{\text{тимофеевка}} = 0,71$

- $F1_{\text{овёс}} = 0,17$

- $OA_{DpRVI} = 86,0 \%$

- $F1_{\text{пар}} = 0,76$

- $F1_{\text{залежь}} = 0,93$

- $F1_{\text{соя}} = 0,90$

- $F1_{\text{гречиха}} = 0,62$

- $F1_{\text{тимофеевка}} = 0,50$

- $F1_{\text{овёс}} = 0,07$

# Идентификация культур на полях

## **NDVI**

162 поля из 170 идентифицировано верно (**95,3 %**):

Залежь 58/60

Соя 46/46

Пар 43/45

Гречиха 13/14

Тимофеевка 2/2

Овёс 0/3

## **DpRVI**

155 полей из 170 идентифицировано верно (**91,2 %**):

Залежь 60/60

Соя 46/46

Пар 37/45

Гречиха 10/14

Тимофеевка 2/2

Овёс 0/3

# Классификация пахотных земель Хабаровского края с использованием методов машинного обучения(ИТНРС-2023)

- Используются снимки с показателем облачности менее 20 % и маски облачности “scl\_cloud\_high\_proba”, “scl\_cloud\_shadow”
- Проведена 2σ-фильтрация данных на каждом поле
- Ряды с количеством значений менее 6 (кол-во параметров при **аппроксимации**) удалялись

2021

23 снимка

Месяц	Количество снимков
Май	5
Июнь	3
Июль	2
Август	3
Сентябрь	5
Октябрь	5

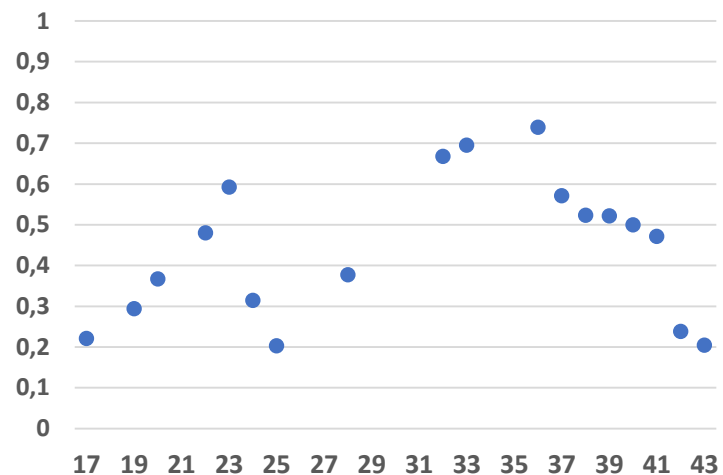
2022

16 снимков

Месяц	Количество снимков
Май	2
Июнь	3
Июль	1
Август	3
Сентябрь	4
Октябрь	3

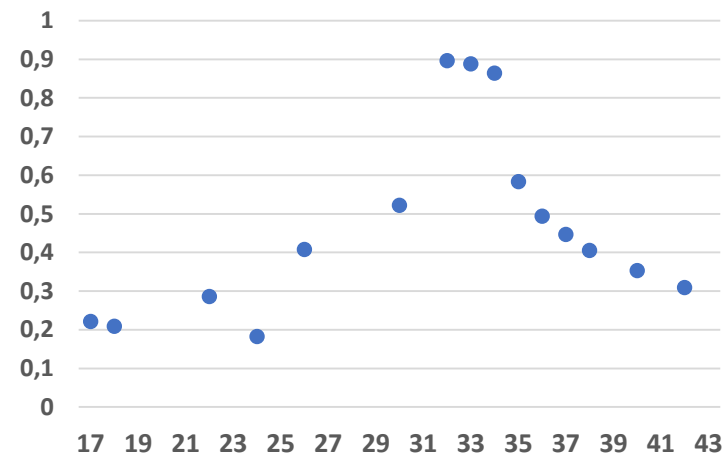
# Недельные композиты NDVI

NDVI 2021



Необходимо восстановить 9 значений из 27!

NDVI 2022



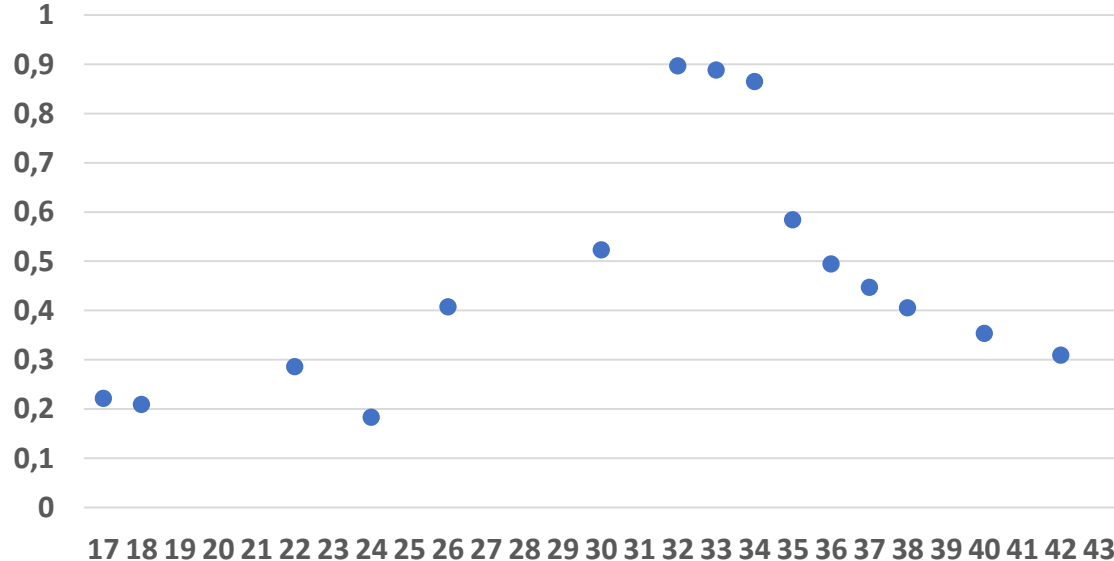
Необходимо восстановить 12 значений из 27!!

# Аппроксимация

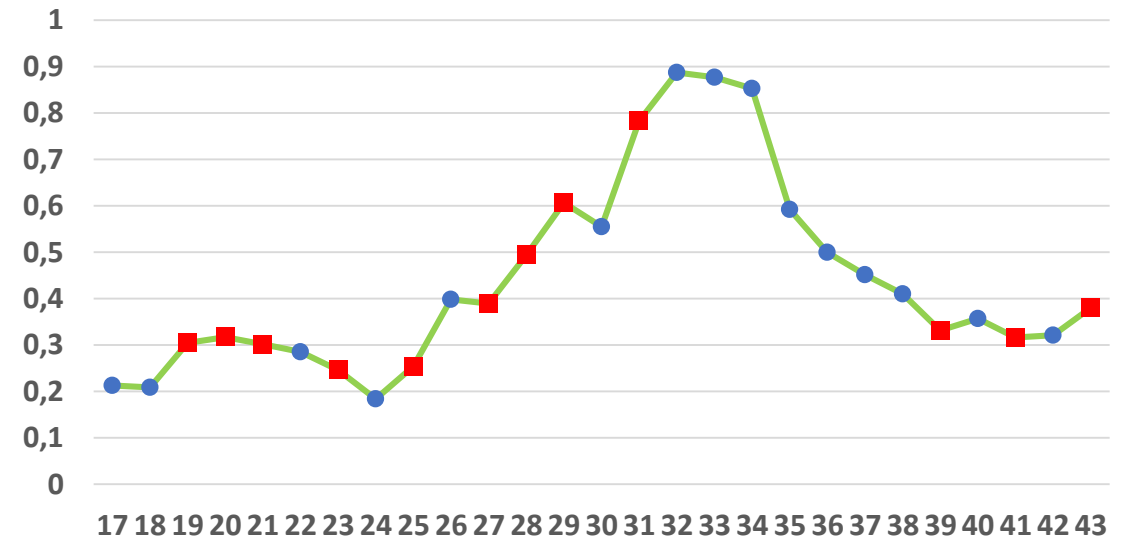
Для восстановления данных использовалась аппроксимация при помощи тригонометрической функции (двойной синусоидальной функции):

$$f(x) = a_1 \sin(b_1 x + c_1) + a_2 \sin(b_2 x + c_2)$$

NDVI-1-2022-недельные композиты



NDVI-1-2022 восстановленные композиты



# Сбалансированные выборки

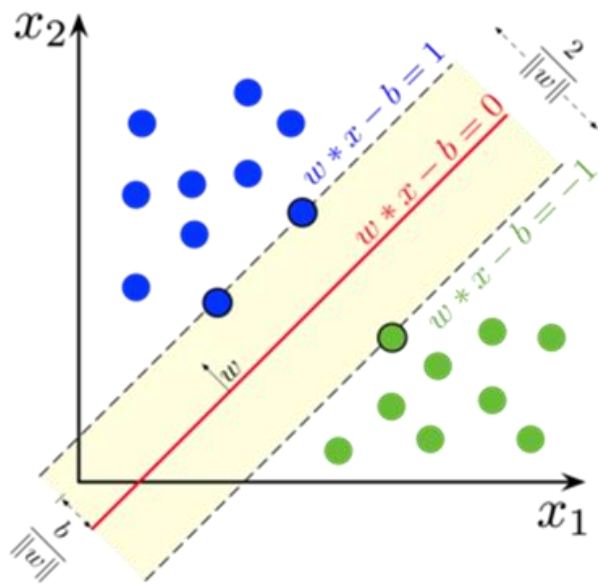
## Обучающая

	Залежь	Соя	Гречиха	Многолетние травы	Овёс	Всего
Кол-во полей	20	20	10	22	14	113
Площадь, га	501	499	438	467	499	2906
Пиксели NDVI	48500	42956	40945	32695	40968	245887

## Тестовая

	Залежь	Соя	Гречиха	Многолетние травы	Овёс	Всего
Кол-во полей	10	9	11	14	11	55
Площадь, га	200	204	168	201	189	862
Пиксели NDVI	18664	19650	15500	18480	17370	89664

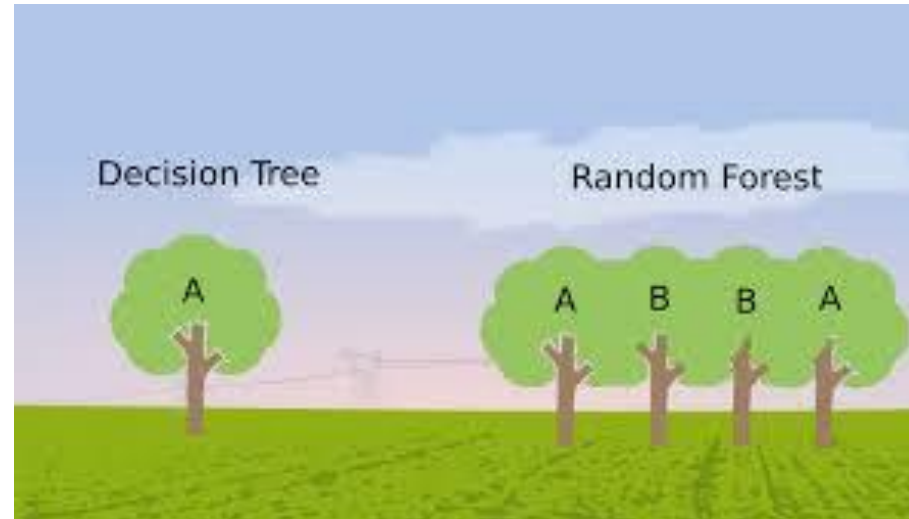
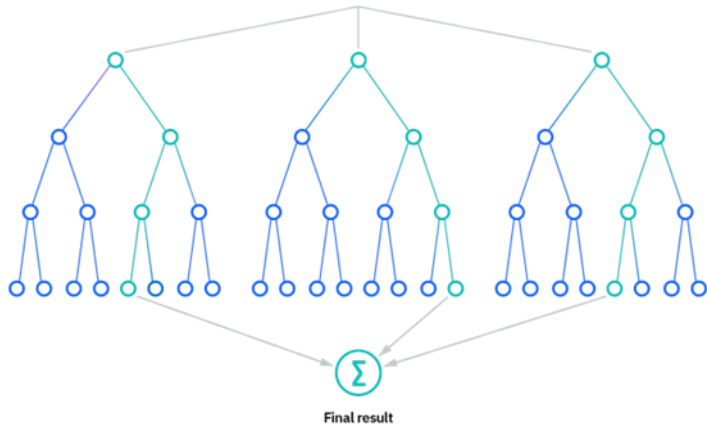
# Метод опорных векторов



Parameter	Value
Penalization norm	L2
Loss function	Squared hinge
Multi-class strategy	ovr
Regularization parameter C	100

Zhang, H.; Kang, J.; Xu, X.; Zhang, L. Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China. *Comput. Electron. Agric.* 2020, 176, 105618.

# Метод Random Forest

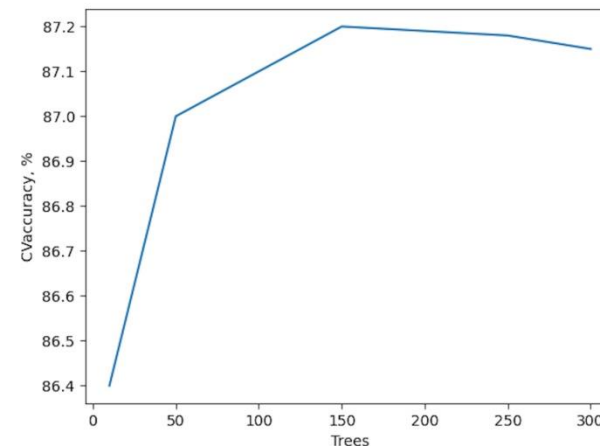


`sklearn.ensemble.RandomForestClassifier`

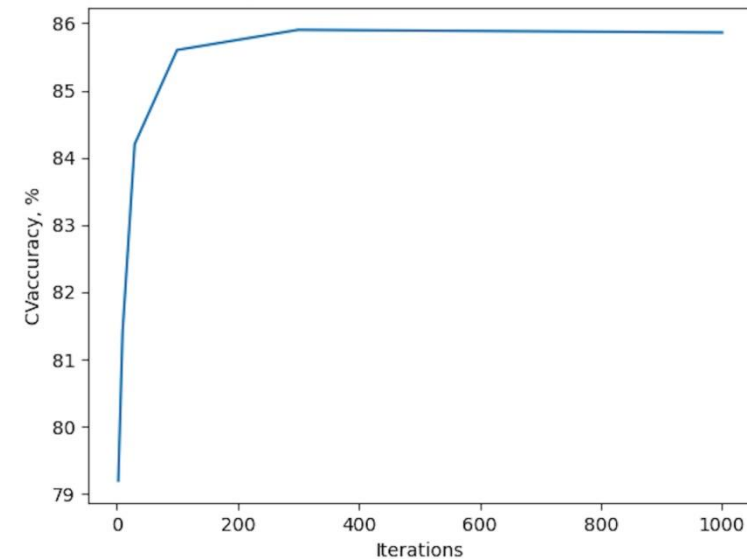
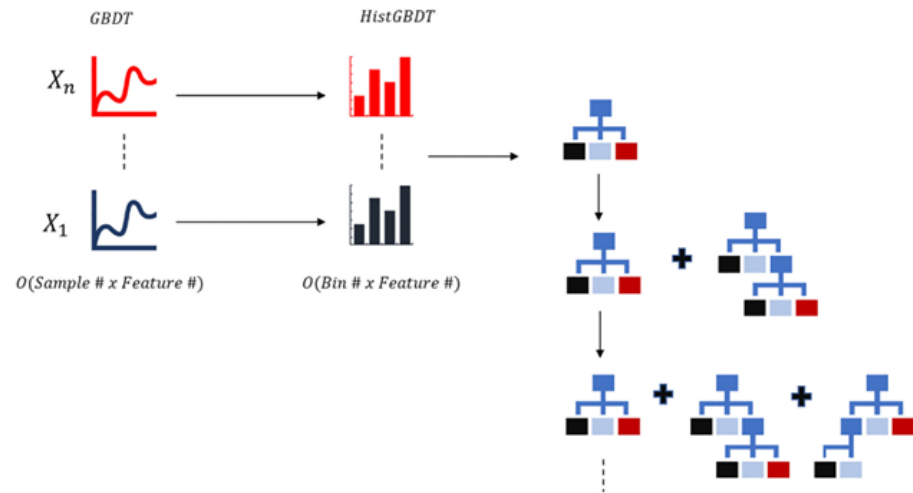
Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32.

Tran, K.H.; Zhang, H.K.; McMaine, J.T.; Zhang, X.; Luo, D. 10 m crop type mapping using Sentinel-2 reflectance and 30 m cropland data layer product. Int. J. Appl. Earth Obs. Geoinf. 2022, 107, 102692.

Количество деревьев 150



# Гистограммный градиентный бустинг



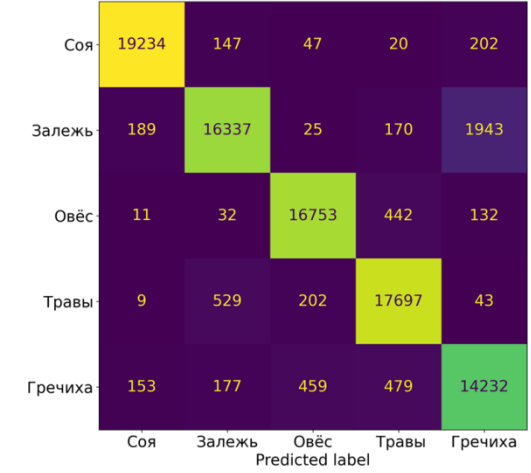
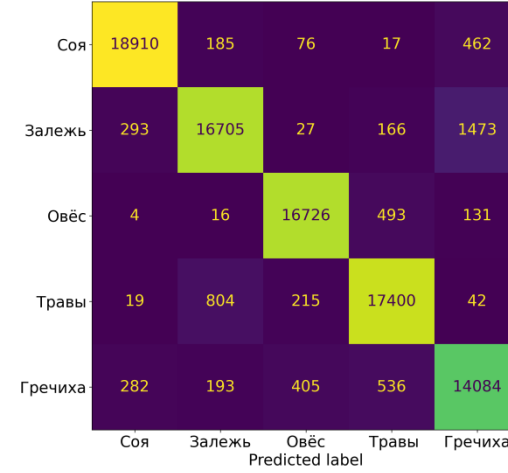
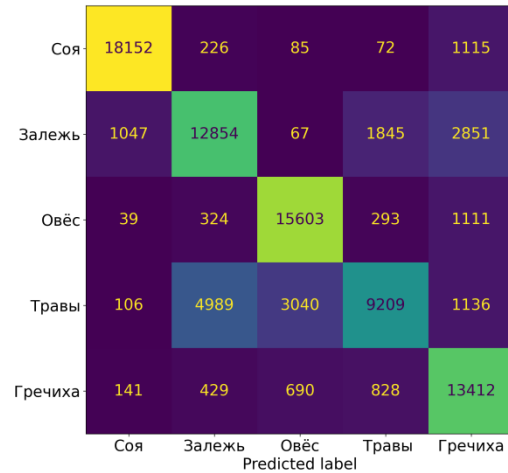
Количество итераций 300

**`sklearn.ensemble.HistGradientBoostingClassifier`**

Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16), San Francisco, CA, USA, 13–17 August 2016.

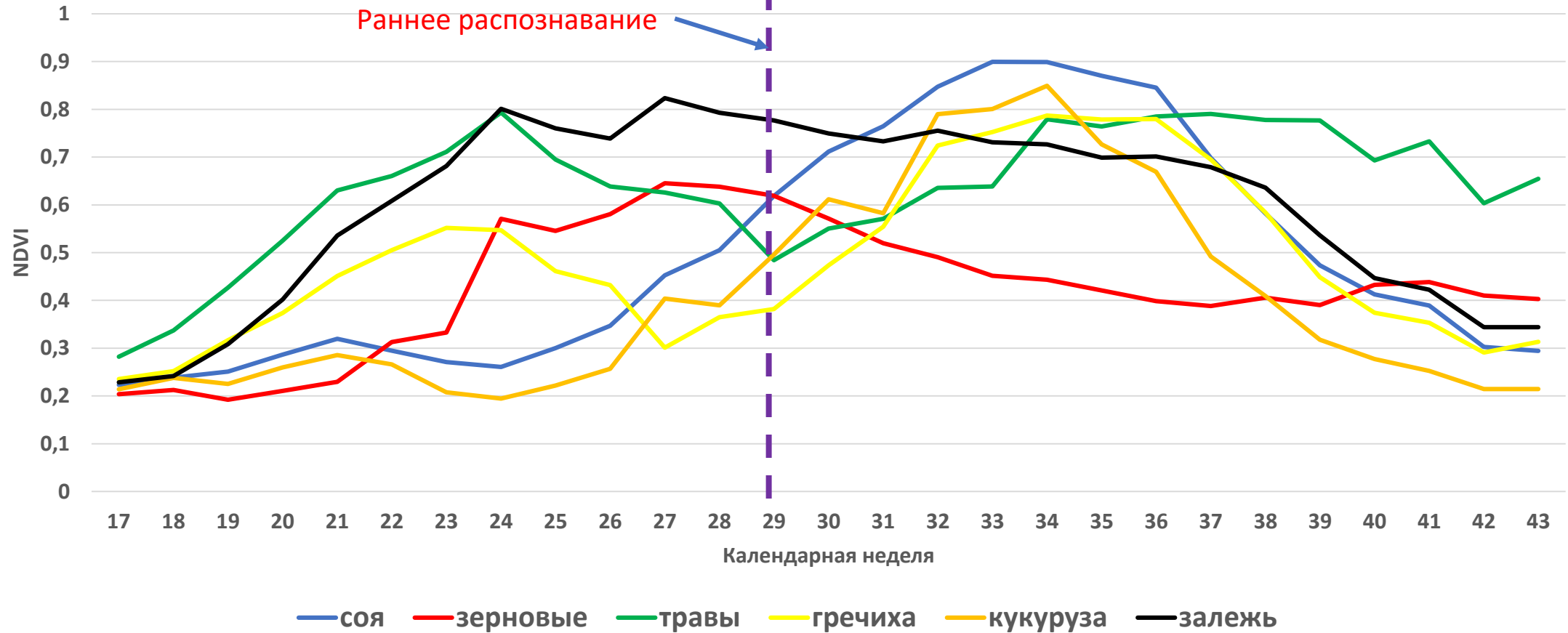
Candido, C.; Blanco, A.C.; Medina, J.; Gubatanga, E.; Santos, A.; Sta Ana, R.; Reyes, R.B. Improving the consistency of multi-temporal land cover mapping of Laguna Lake watershed using light gradient boosting machine (LightGBM) approach, change detection analysis, and Markov chain. Remote Sens. Appl. Soc. Environ. 2021, 23, 100565.

# Результаты классификации



Метод	SVM	RF	HGB
$OA, \%$	77.2	93.5	94.0
$F1_{\text{со́я}}$	0.928	0.966	0.980
$F1_{\text{за́лежь}}$	0.686	0.914	0.910
$F1_{\text{овёс}}$	0.847	0.961	0.961
$F1_{\text{тра́вы}}$	0.599	0.938	0.949
$F1_{\text{гречи́ха}}$	0.764	0.989	0.888
Время обучения, с.	98.66	146.16	10.85

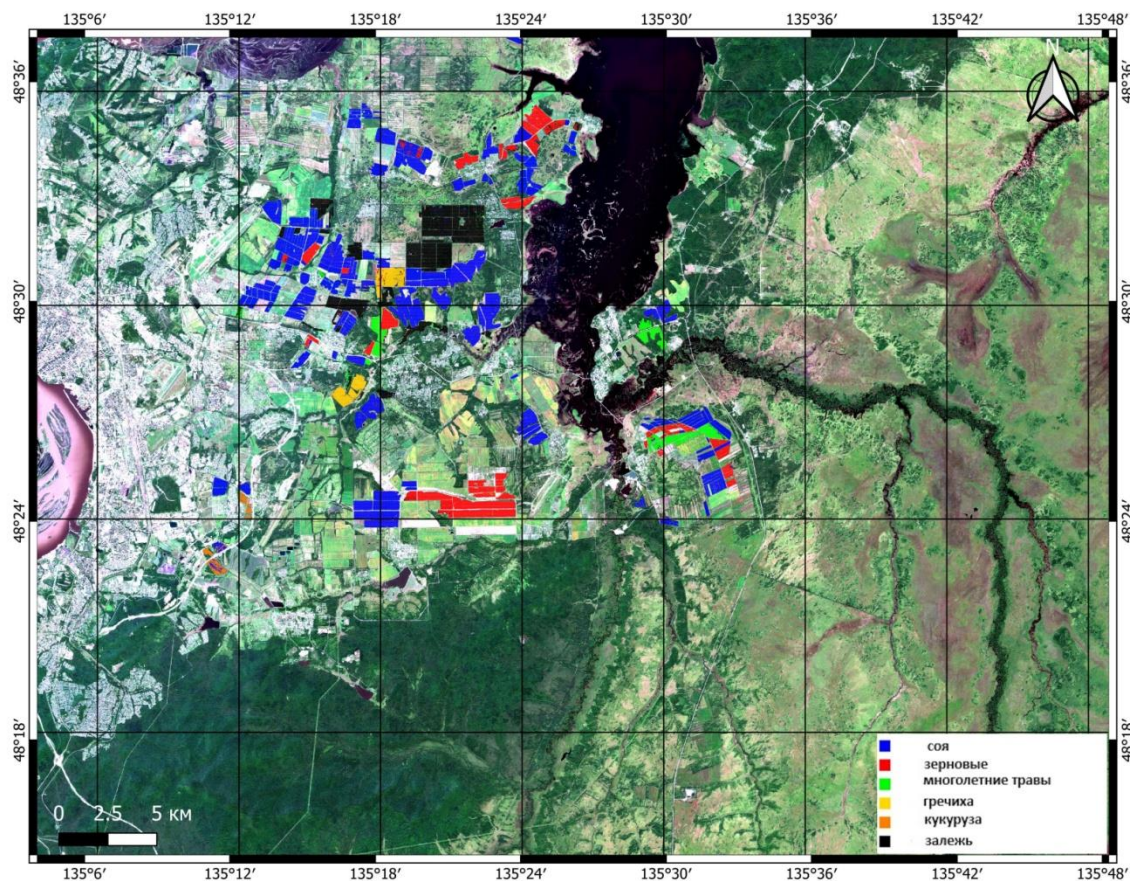
# ИСПОЛЬЗОВАНИЕ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ ДЛЯ МОНИТОРИНГА ПАХОТНЫХ ЗЕМЕЛЬ ХАБАРОВСКОГО КРАЯ (2024)



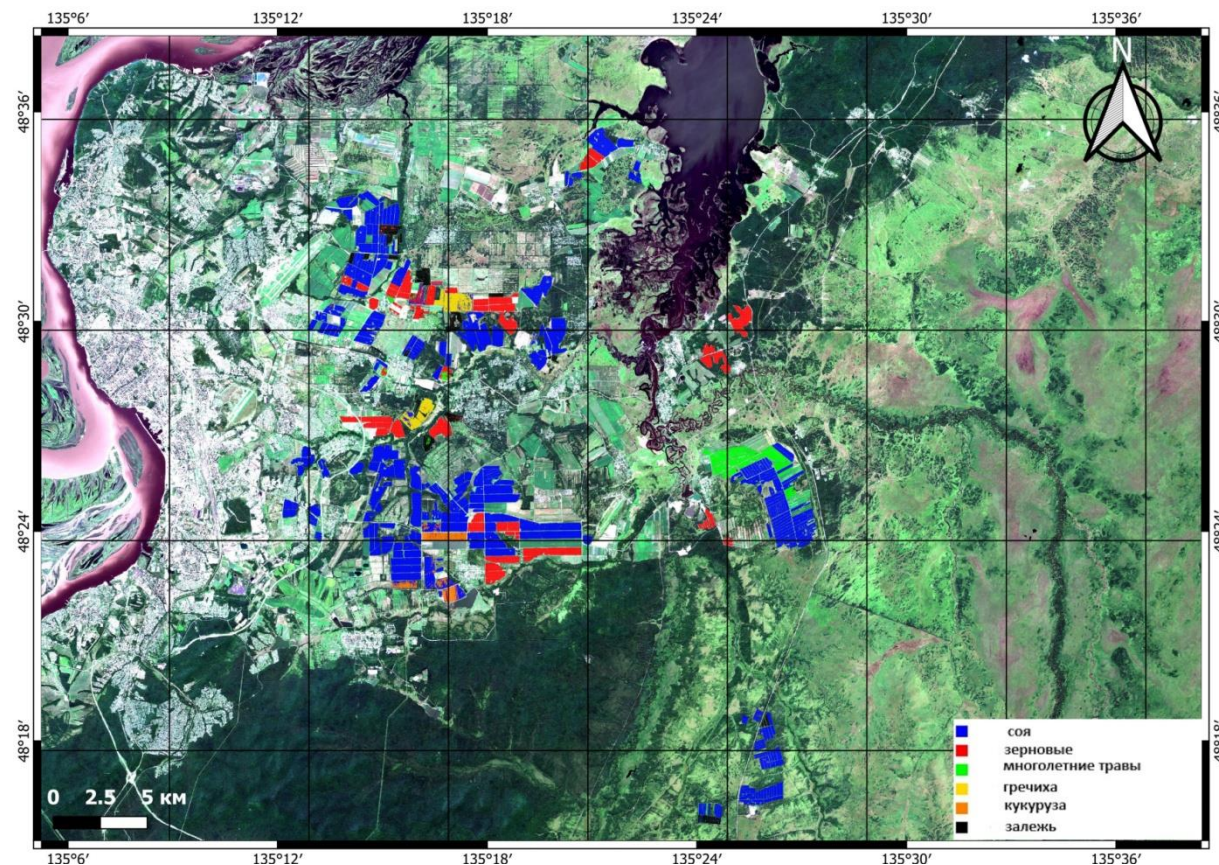
# Оценка точности распознавания

Год	ОА, %	Залежь	Соя	Зерновые	Многолетние травы	Гречиха	Кукуруза
2021	97,3	0,99	0,98	0,86	0,74	0,96	-
2021 (раннее)	92,2	0,97	0,97	0,46	0,41	0,94	-
2022	96,3	0,96	0,98	0,96	0,94	0,94	0,65
2022 (раннее)	88,6	0,93	0,94	0,81	0,81	0,44	0,30
2023	96,3	0,94	0,98	0,97	0,94	0,85	0,75
2023 (раннее)	88	0,9	0,92	0,84	0,89	0,59	0,37

# Карты исследуемых полей



Карта пахотных земель (Хабаровский район, 2022 год)



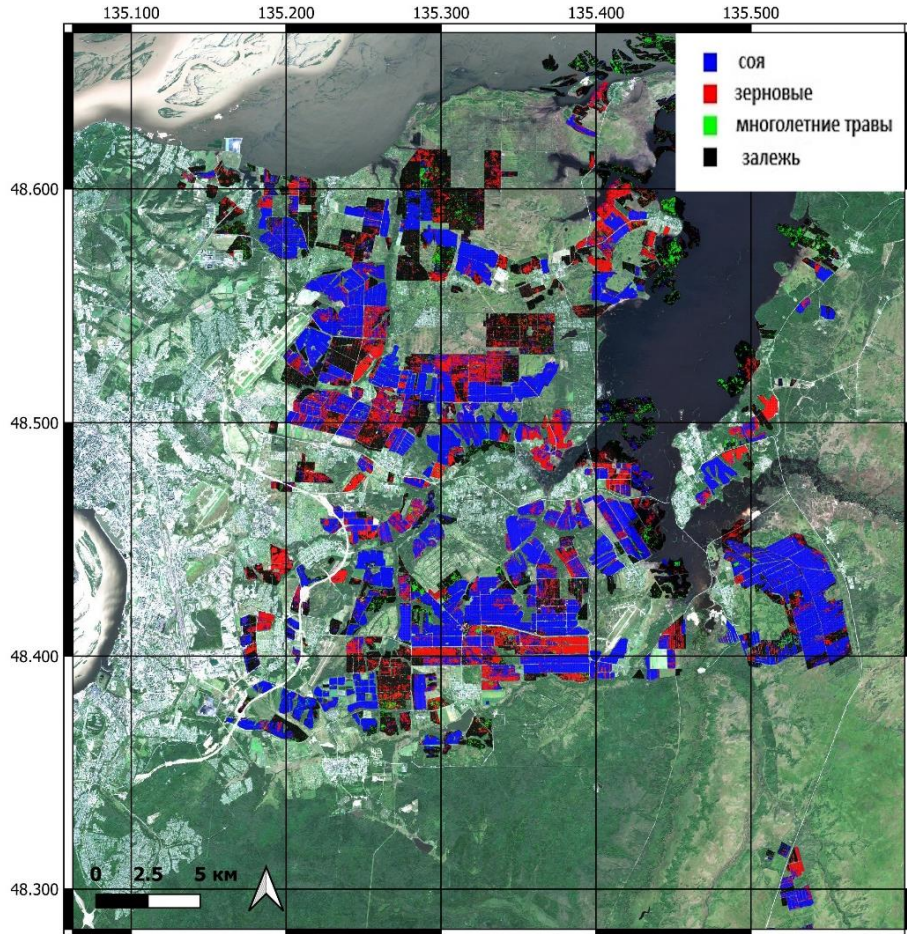
Карта пахотных земель (Хабаровский район, 2023 год)

# Восстановление сведений о севообороте по

## МНОГОЛЕТНИМ ДАННЫМ

Обучение по данным за 2021-2023 годы

Карта пахотных земель района за 2019 год

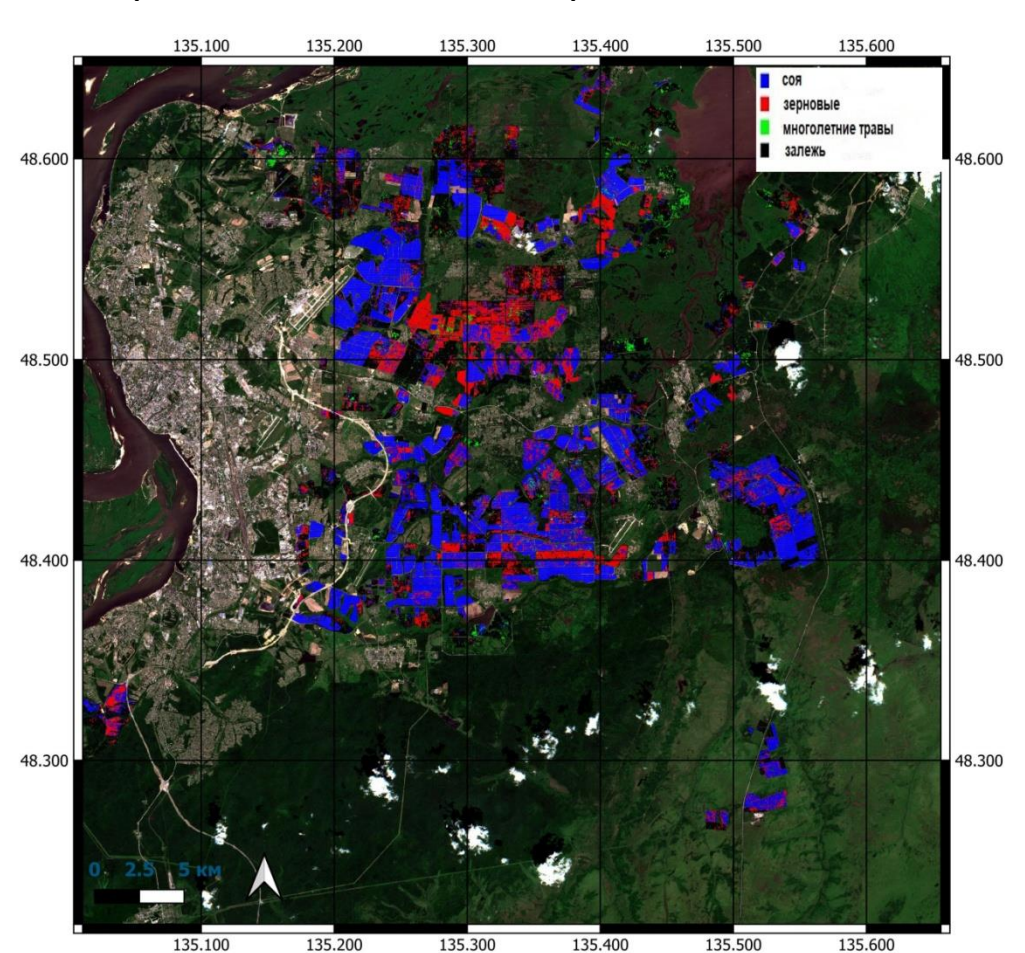


Классификация в  
контурах полей  
ЕФИС  
ЗСН Хабаровского  
района

1439 полей  
28550 га  
Более 2,5 млн.  
пикселей  
4 класса

Более 1,5 млн.  
пикселей для  
обучения

Карта пахотных земель района за 2020 год



# ОСНОВНЫЕ ВЫВОДЫ

- Уделяйте внимание подготовке обучающей выборки (это – большая часть вашего успеха!)
- Изучайте объект в динамике (спутниковый мониторинг, временные ряды снимков или индексов)
- Обращайте внимание на сбалансированность выборок (подбирайте наиболее подходящий алгоритм)
- При подготовке научных статей уделяйте особое внимание оценке точности модели (кросс-валидация, метрики для классов, матрицы неопределенности)
- Визуализируйте результаты классификации (рисунки, карты-схемы)
- Пробуйте, экспериментируйте, творите!

# Основные публикации

- Dubrovin, K. Cropland mapping using Sentinel-1 data in the southern part of the Russian Far East / K. Dubrovin, A. Stepanov, A. Verkhoturov // Sensors. – 2023. – Vol. 23, No. 18. – P. 7902. – DOI 10.3390/s23187902.
- Multi-Year Cropland Mapping Based on Remote Sensing Data: A Case Study for the Khabarovsk Territory, Russia / K. Dubrovin, A. Verkhoturov, A. Stepanov, T. Aseeva // Remote Sensing. – 2024. – Vol. 16, No. 9. – P. 1633. – DOI 10.3390/rs16091633.
- Enhancement of Comparative Assessment Approaches for Synthetic Aperture Radar (SAR) Vegetation Indices for Crop Monitoring and Identification—Khabarovsk Territory (Russia) Case Study / A. Sorokin, A. Stepanov, K. Dubrovin, A. Verkhoturov // Remote Sensing. – 2024. – Vol. 16, No. 14. – P. 2532. – DOI 10.3390/rs16142532.
- Опыт использования данных прибора КМСС КА «Метеор-М» № 2 для задач мониторинга сельскохозяйственных угодий юга Хабаровского края / Л. В. Илларионова, А. С. Степанов, К. Н. Дубровин [и др.] // Современные проблемы дистанционного зондирования Земли из космоса. – 2025. – Т. 22, № 1. – С. 81-92. – DOI 10.21046/2070-7401-2025-22-1-81-92.

# Спасибо за внимание!

Дубровин Константин Николаевич

Научный сотрудник

Вычислительный центр ДВО РАН

г. Хабаровск



8 (909) 859-08-81



nobforward@gmail.com

