



Оценка распаханности участков личных подсобных хозяйств с использованием методов машинного обучения по данным Sentinel-2



DOI: 10.5281/zenodo.14048405

Артем Александрович Сабаев (artsab@yahoo.com), Елена Владимировна Понькина (ponkinaelena77@mail.ru)
Алтайский государственный университет, Барнаул

Введение

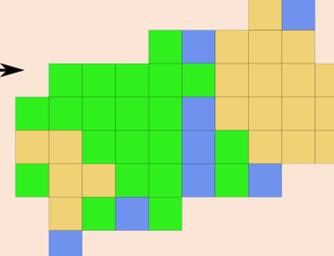
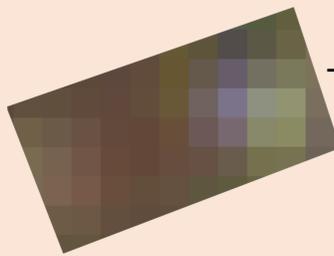
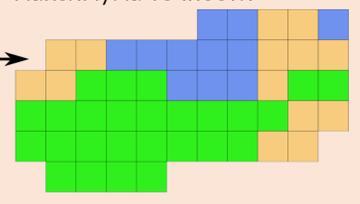
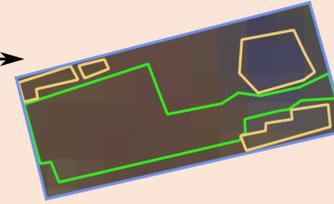
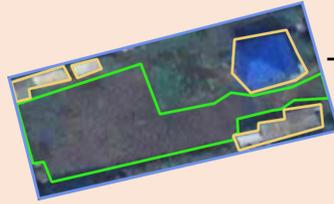
Одним из аспектов исследования режимов использования участков ЛПХ является оценка степени распаханности, базирующаяся на детектировании наличия и площади пашни, показывающая долю культивируемой пашни от общей площади участка.

Методы

На участках ЛПХ выделены полигоны трёх классов: пашня, здания и сооружения, прочие объекты. По данным спутниковых снимков Sentinel-2 вычислены временные ряды индексов GNDVI, MCARI, MSI и NDMI, и их статистики за вегетационный период. Были обучены 4 модели машинного обучения – градиентный бустинг, деревья решений, случайный лес и метод опорных векторов. Классификация выполнена на уровне пикселей. Выбор классификационной модели на основе оценок точности по тестовой выборке. Итоговая предиктивная классификация всех участков территории выполнена методом градиентного бустинга который показал точность классификации выше остальных рассматривавшихся методов.

Алгоритм обработки данных

1. Формирование выборки объектов и разметка данных
2. Выгрузка данных Sentinel-2 и расчёт GNDVI, MCARI, NDMI, MSI
3. Отбор модели классификации по величине максимума точности
4. Выборка объектов для прогностической классификации
5. Прогностическая классификация
6. Оценка степени распаханности



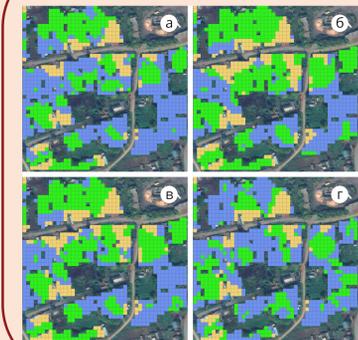
Выборка объектов

Территория – с. Кытманово, пос. Тягун, Кытманушка, Алтайский край



- Класс 1 – пашня (524 объекта)
- Класс 2 – иные земли (116 объектов)
- Класс 3 – здания и сооружения (177 объектов)

Классификация



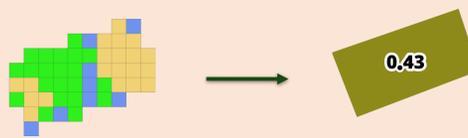
- а) Градиентный бустинг (Gradient Boosting)
 - б) Дерево решений (Decision Tree)
 - в) Случайный лес (Random Forest)
 - г) Метод опорных векторов (Support Vector Machines)
- Распахиваемые земли
 - Здания и сооружения
 - Иные земли

Результат I

Точность классификации, объекты тренинговой выборки/объекты тестовой выборки

Номер класса	Метрики точности	Градиентный бустинг (Gradient Boosting)	Случайный лес (Random Forest)	Дерево решений (Decision Tree)	Метод опорных векторов (Support Vector Machines)
	Overall Accuracy	0.973 / 0.877	0.986 / 0.870	0.891 / 0.856	0.924 / 0.853
1	Producer's Accuracy	0.984 / 0.984	0.959 / 0.973	0.928 / 0.963	0.981 / 0.992
	User's Accuracy	0.977 / 0.857	0.987 / 0.856	0.924 / 0.848	0.981 / 0.825
2	Producer's Accuracy	0.954 / 0.656	0.976 / 0.664	0.864 / 0.634	0.840 / 0.584
	User's Accuracy	0.956 / 0.945	0.988 / 0.919	0.846 / 0.892	0.948 / 0.967
3	Producer's Accuracy	0.934 / 0.712	0.935 / 0.680	0.584 / 0.696	0.677 / 0.560
	User's Accuracy	0.968 / 0.918	0.975 / 0.904	0.694 / 0.837	0.916 / 0.946

Расчёт степени распаханности



$$\text{Степень распаханности: } R = \frac{N1}{N}$$

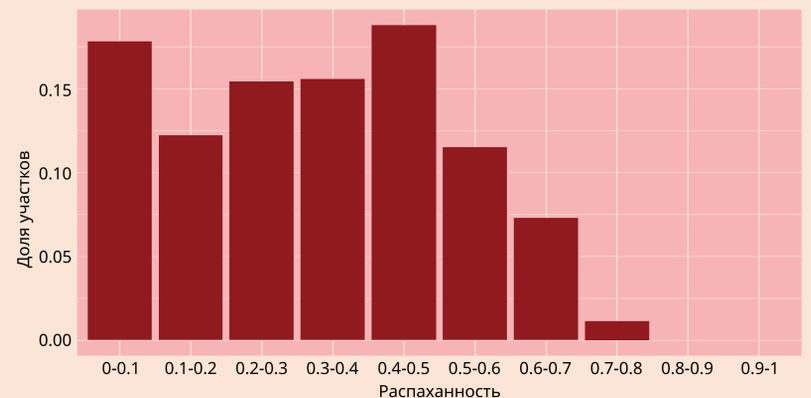
N – общее количество пикселей участка
N1 – количество пикселей класса 1 (распахиваемые земли)



- Распаханность
- 0.0 – 0.1
 - 0.1 – 0.2
 - 0.2 – 0.3
 - 0.3 – 0.4
 - 0.4 – 0.5
 - 0.5 – 0.6
 - 0.6 – 0.7
 - 0.7 – 0.8
 - 0.8 – 0.9
 - 0.9 – 1.0

Результат II

Оценка степени распаханности участка выполнена на основе попиксельной классификации как отношение количества пикселей, классифицированных как класс «пашня» к общему количеству пикселей участка, попадающего в полигон, выраженной в процентах. Результаты классификации были картированы с использованием цветовой шкалы степени распаханности от 0 до > 100%. На территории исследования по состоянию на 2022 год около 18% всех участков имеет распаханность менее 10%, низкий уровень распахки (от 0 до 20%) имеет около 30% участков, средний уровень (от 20 до 50%) – около 50% участков, около 20% участков имеет высокий уровень распахки (более 50%).



Заключение

Разработанный инструментариум представляется перспективным для внедрения в систему мониторинга режимов использования участков ЛПХ на уровне сельских населенных пунктов, в частности детектирования пашни и забрасываемости участков. Анализ пространственных паттернов как и поиск факторов культивации пашни на участках ЛПХ в населенных пунктах представляет интерес для дальнейшего исследования.