

Институт космических  
исследований  
Шестнадцатая Всероссийская  
Открытая конференция  
«Современные проблемы  
дистанционного зондирования  
земли из космоса»



**САМАРСКИЙ** УНИВЕРСИТЕТ  
SAMARA UNIVERSITY

## Пространственная классификация древесных пород по данным Sentinel-2 и таксации леса

**Докладчик:** доцент кафедры геоинформатики  
и информационной безопасности,  
**А.Ю. Денисова**

**Соавторы:** А.Ю.Баврина, Л.М. Кавеленова,  
Е.С. Корчиков, О.А. Кузовенко, Н.В. Прохорова,  
Д.А. Терентьева, В.А. Федосеев

Москва, 2018



## Задача классификации лесных экосистем в Самарской области по данным Sentinel-2 и таксации леса

**Цель исследований:** исследование вариантов построения классификатора древесных пород по данным Sentinel-2 и таксации леса на основе локальных и поэлементных классификаторов

### **Задачи:**

- выбор методов обработки и состава данных ДЗЗ;
- исследование алгоритма классификации и метода обучения классификатора;
- анализ процедур пред- и пост-обработки данных классификации;
- анализ ошибок классификации с целью выявления изменений в данных таксации.

### **Практическое значение классификации лесов:**

- 1) Уточнение данных таксации: оперативность и снижение стоимости за счёт использования данных Sentinel-2;
- 2) Выявление фрагментарно сохранившихся на фоне пахотных угодий коренных лесных экосистем с целью изучения и сохранения биологического разнообразия региона;
- 3) Выявление ценных естественных лесных природных сообществ с возможным приданием им охранного статуса.



## Обзор методов классификации лесных сообществ по данным Sentinel-2

Авторы	Данные	Доп. данные	Количество классов лесных сообществ	Методы
Immitzer, Vuolo, Atberger, 2016	1 снимок 10 каналов	Эталонные наземные площадки	7	LSMS сегментация Random Forest классификатор
Кедров, Тарасов, 2017	1 снимок NIR, Green, Blue	Данные экспертного дешифрирования снимков	4-15	Нейронные сети
Перепечина, Глушенков, Корсиков, 2017	1 снимок NIR, Green, Blue	Эталонные наземные площадки	3	Линейный и квадратичный дискриминантный анализ

Выбранный подход: совмещение пространственной и поэлементной спектральной классификации на основе результатов для гиперспектральных изображений полученных (Борзов, Потатуркин, 2017) и (Федосеев, 2018)



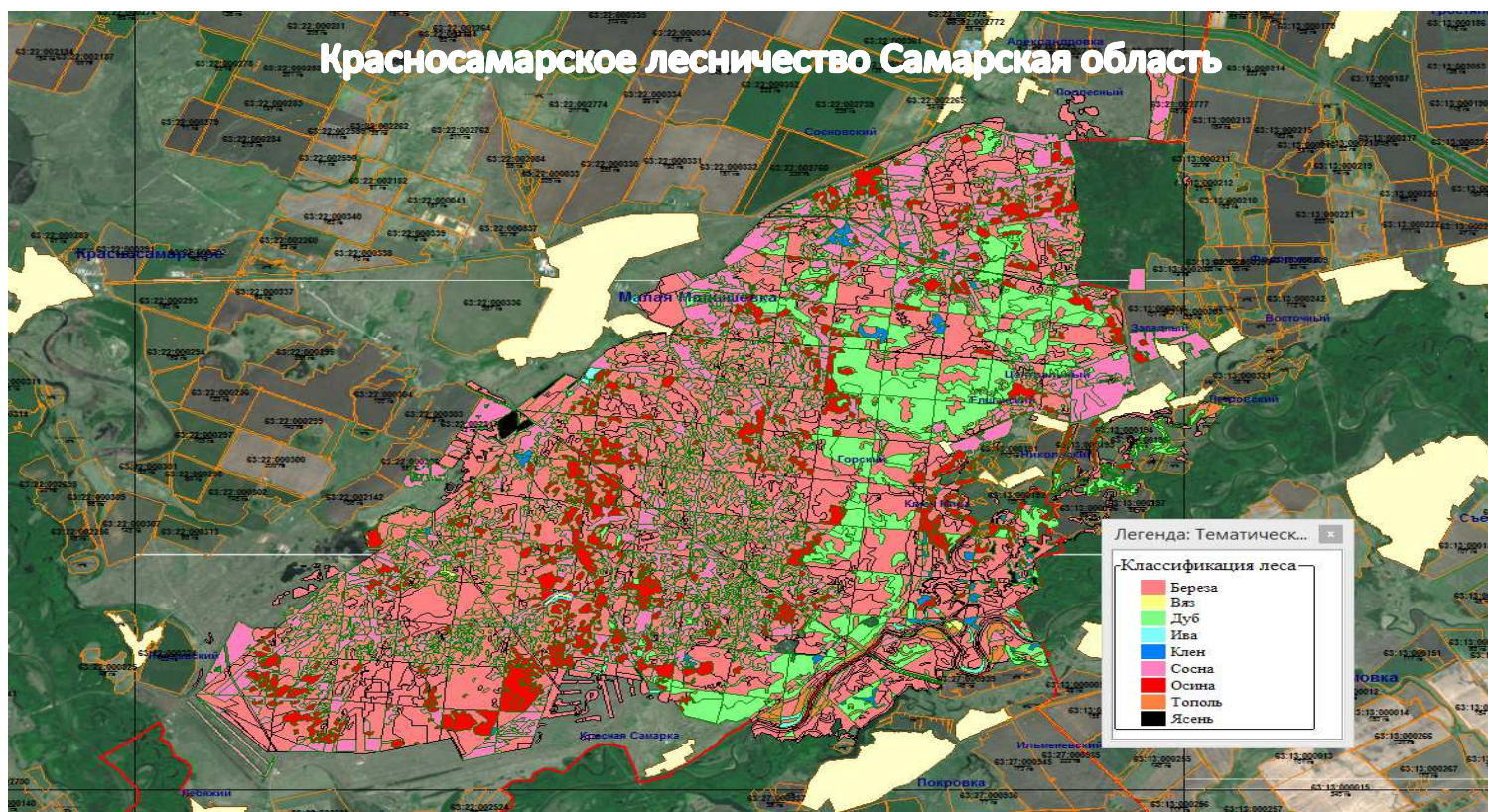
## Данные таксации

### Векторные данные:

- границы таксационных выделов в ГИС «Ингео»;
- таксационная формула: 8Я2С;
- 2013-2014 г.г.

### Отбор участков для обучения и тестирования классификации:

- 8/10 доли участка занято одной породой;
- 9 классов : берёза, вяз, дуб, ива, клён, осина, сосна, тополь, ясень.



Данные предоставлены: Министерством лесного хозяйства и природопользования Самарской области



## Данные дистанционного зондирования Земли

### Композит снимков Sentinel-2:

- 17 снимков;
- 8 спектральных каналов
- уровень обработки L1C и L2A;
- разрешение 10 м\*;
- период: апрель – октябрь 2018 г.

\* Повышение разрешения до 10 м с помощью Scanex Image Processor 4.2;

№ канала	Длина волны, нм	Разр., м
2 (Blue)	496.6	10
3 (Green)	560.0	10
4 (Red)	664.5	10
5 (Vegetation red edge)	703.9	20
6 (Vegetation red edge)	740.2	20
7 (Vegetation red edge)	782.5	20
8 (NIR)	835.1	10
8A (Narrow NIR)	864.8	20

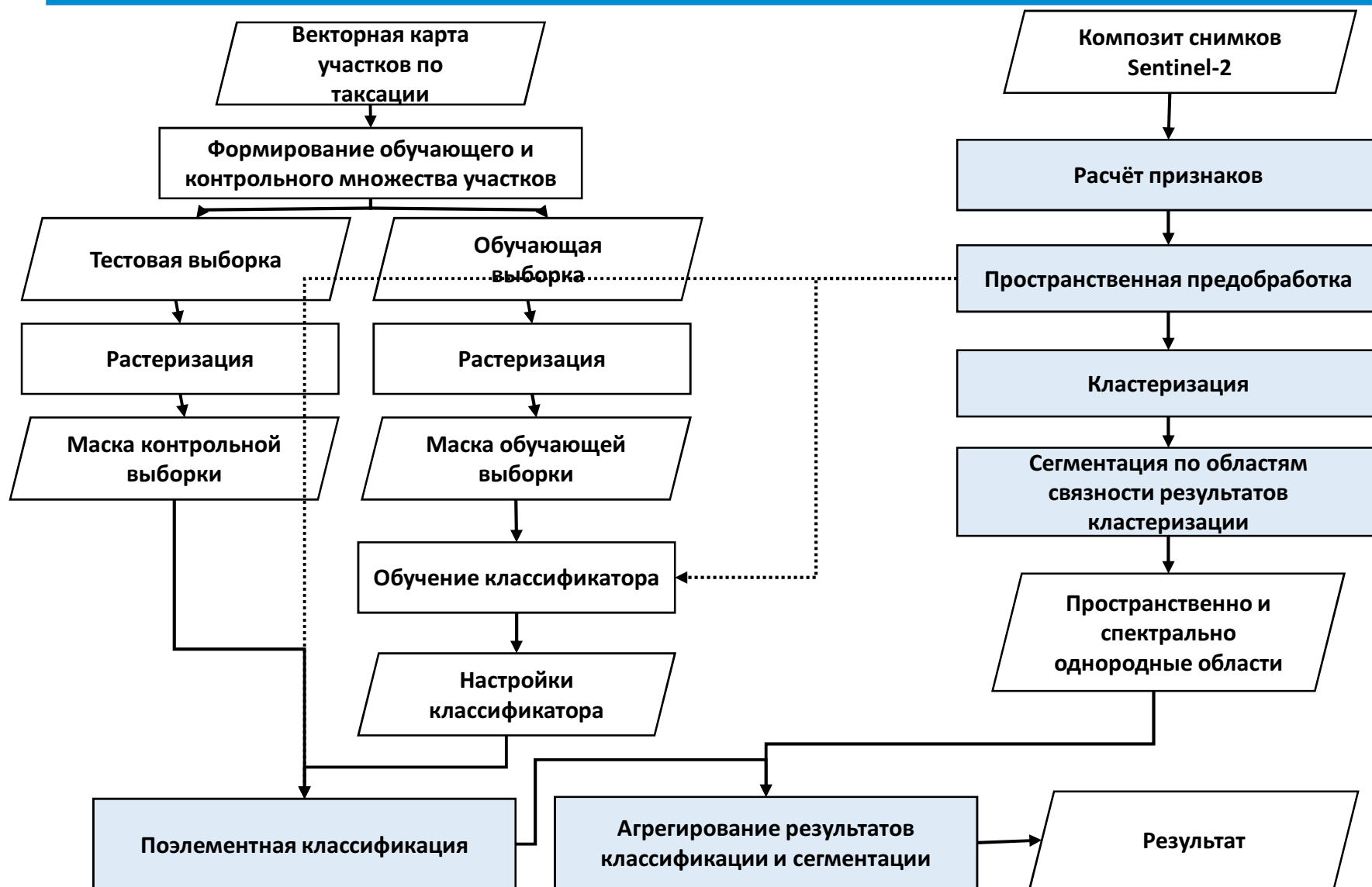


№	Дата	№	Дата	№	Дата
1	12.04.2018	4	04.05.2018	7	27.05.2018
2	17.04.2018	5	12.05.2018	8	23.06.2018
3	02.05.2018	6	22.05.2018	9	03.07.2018

№	Дата	№	Дата	№	Дата
9	03.07.2018	12	07.08.2018	15	04.09.2018
10	31.07.2018	13	20.08.2018	16	16.09.2018
11	05.08.2018	14	01.09.2018	17	19.10.2018



## Общая схема процедуры классификации





## Методы и алгоритмы, используемые в процедуре классификации

<b>Расчёт признаков</b>	<ul style="list-style-type: none"><li>• нормализация – приведение к диапазону <math>[-1,1]</math></li><li>• преобразование методом главных компонент (PCA)</li></ul>
<b>Пространственная предобработка</b>	<ul style="list-style-type: none"><li>• медианная фильтрация окном <math>N \times N</math></li></ul>
<b>Кластеризация</b>	<ul style="list-style-type: none"><li>• k-means++</li><li>• обобщенный EM алгоритм (Generalized EM - GEM)*</li></ul>
<b>Сегментация</b>	<ul style="list-style-type: none"><li>• выделение областей связности по результатам кластеризации</li></ul>
<b>Поэлементная классификация</b>	<ul style="list-style-type: none"><li>• метод опорных векторов с радиальными базисными функциями (SVM-RBF).**</li></ul>
<b>Агрегирование результатов классификации</b>	<ul style="list-style-type: none"><li>• поэлементная классификация векторов медиан для сегментов (медианный способ); **</li><li>• голосование большинства по сегменту;</li><li>• пост обработка локальным окном.</li></ul>

\* Denisova AY, Sergeyev VV. EM clustering algorithm modification using multivariate hierarchical histogram in the case of undefined cluster number. In: Proceedings of SPIE [Internet]. International Society for Optics and Photonics; 2018 [cited 2018 Sep 14]. p. 108064H.

\*\* Борзов СМ, Потатуркин ОИ. Классификация типов растительного покрова по гиперспектральным данным дистанционного зондирования Земли. Вестник НГУ Серия: Информационные технологии. 2014;12(4):13–22;



## Обобщенный EM (GEM) алгоритм кластеризации

### Предположения:

- целочисленные M-мерные векторы-признаков  $x_n$ ,  $n=0, \dots, N-1$ , где  $x_{nm} \in [0, 2^B - 1]$ ;
- кластеры описываются смесью гауссовых распределений с неизвестным числом кластеров  $G$  и неизвестными параметрами распределения  $\mu_g$  и  $\Sigma_g$ ,  $g=0, \dots, G-1$ .

$$p(x) = \sum_{g=0}^{G-1} \tau_g \frac{1}{(2\pi)^{M/2} |\Sigma_g|^{1/2}} \exp \left\{ -\frac{(x - \mu_g)^T \Sigma_g^{-1} (x - \mu_g)}{2} \right\}$$

### Обобщенный EM алгоритм кластеризации

1. Оценка и усечение по порогу (MVF) иерархической гистограммы

2. Извлечение приближенного списка мод гистограммы по  $D$  верхним уровням иерархии

3. Инициализация параметров кластеров по списку мод

4. Кластеризация слиянием с использованием стандартных EM итераций

### Результат:

- оценка количества кластеров и параметров распределений в кластерах;
- изображение меток кластеров.





### Этапы:

1. Выбор способа алгоритма кластеризации и способа агрегации результатов классификации
2. Анализ процедуры расчёта признаков и пространственной предобработки
3. Анализ влияния периода наблюдений на качество классификации
4. Результаты анализа ошибок классификации с использованием данных наземных обследований

### Метод формирования обучающей выборки:

- равный объем выборки для каждого класса (2000 пикселей)
- обучение – северная часть лесничества
- контроль - южная часть лесничества

**Реализация:** MATLAB R2017b, ResLook

**Точность классификации** – доля правильно классифицированных объектов контрольной выборки



## Выбор базовой части пространственной обработки

### Зафиксировано:

- уровень продукта L2A
- предобработка – нормализация+136 главных компонент
- даты снимков: 12.04 - 19.10

### Исследуется:

- метод кластеризации
- эффект от постобработки локальным окном
- способ агрегации результатов

Кластеризация	Количество кластеров	Количество сегментов	Агрегация	Постобработка локальным окном	Точность 10-fold	Точность на тестовом фрагменте
-	-	-	-	-	0.8500	0.7473
-	-	-	-	+		0.7607
GEM	9	2495	Голосование	-		0.8187
GEM	9	2495	Медиана	+		0.8158
<b>GEM</b>	<b>9</b>	<b>2495</b>	<b>Голосование</b>	<b>+</b>		<b>0.8216</b>
GEM	9	2495	Медиана	+		0.8187
kmeans++	3	803	Голосование	+		0.7942
kmeans++	6	1897	Голосование	+		0.7931
kmeans++	10	4300	Голосование	+		0.7914



## Анализ процедуры расчёта признаков и пространственной предобработки

### Зафиксировано:

- даты снимков: 12.04 - 19.10
- GEM кластеризация
- голосование большинства
- постобработка локальным окном

### Исследуется:

- варианты окна предобработки
- число признаков
- наличие атмосферной коррекции
- использование NDVI

Состав композита	Расчёт признаков	Количество компонент PCA	Размер окна предобработки	Атмосферная коррекция	Точность на тестовом фрагменте
яркости	PCA + нормализация	136 (все)	3	+	0.8216
<b>Окно предобработки</b>					
яркости	PCA + нормализация	136 (все)	1	+	0.8103
яркости	PCA + нормализация	136 (все)	5	+	0.8136
<b>Варианты расчёта признаков</b>					
яркости	PCA	136 (все)	3	+	0.8029
яркости	нормализация	136 (все)	3	+	0.8135
яркости	PCA + нормализация	40	3	+	0.8160
яркости	нормализация	40	3	+	0.7154
яркости	PCA + нормализация	20	3	+	0.7772
<b>Атмосферная коррекция и NDVI</b>					
яркости	PCA + нормализация	136 (все)	3	-	0.7798
яркости+NDVI	PCA + нормализация	153 (все)	3	+	0.8158
NDVI	PCA + нормализация	17 (все)	3	+	0.7535



## Влияние периода наблюдений на качество классификации

### Зафиксировано:

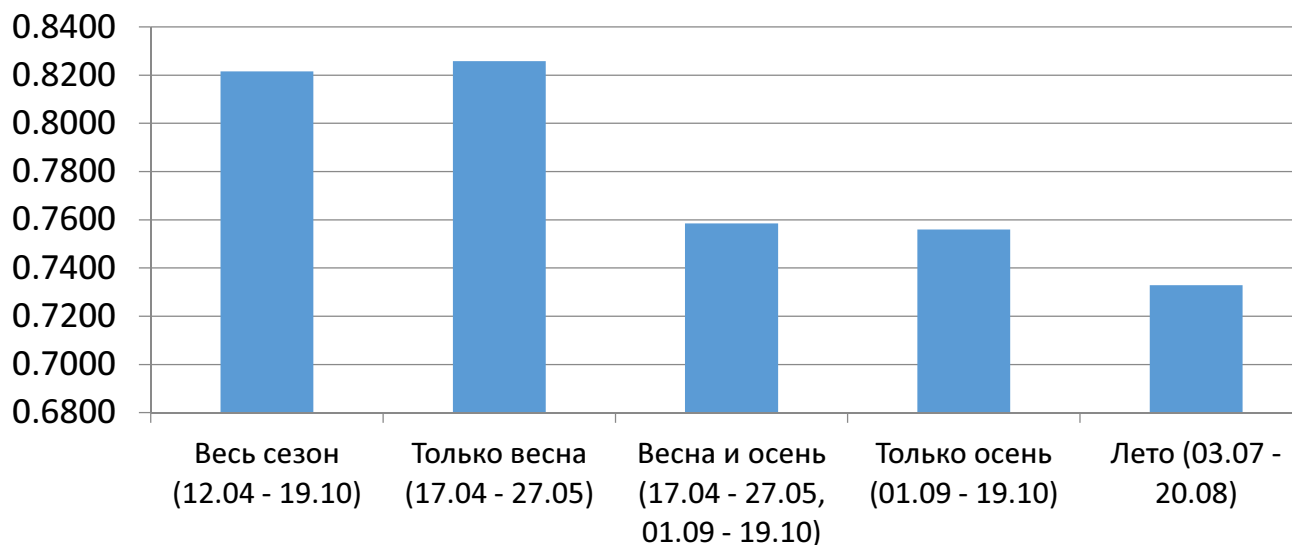
- GEM кластеризация
- голосование большинства
- постобработка локальным окном

### Исследуется:

- информативность сезонов съемки
- информативность каналов

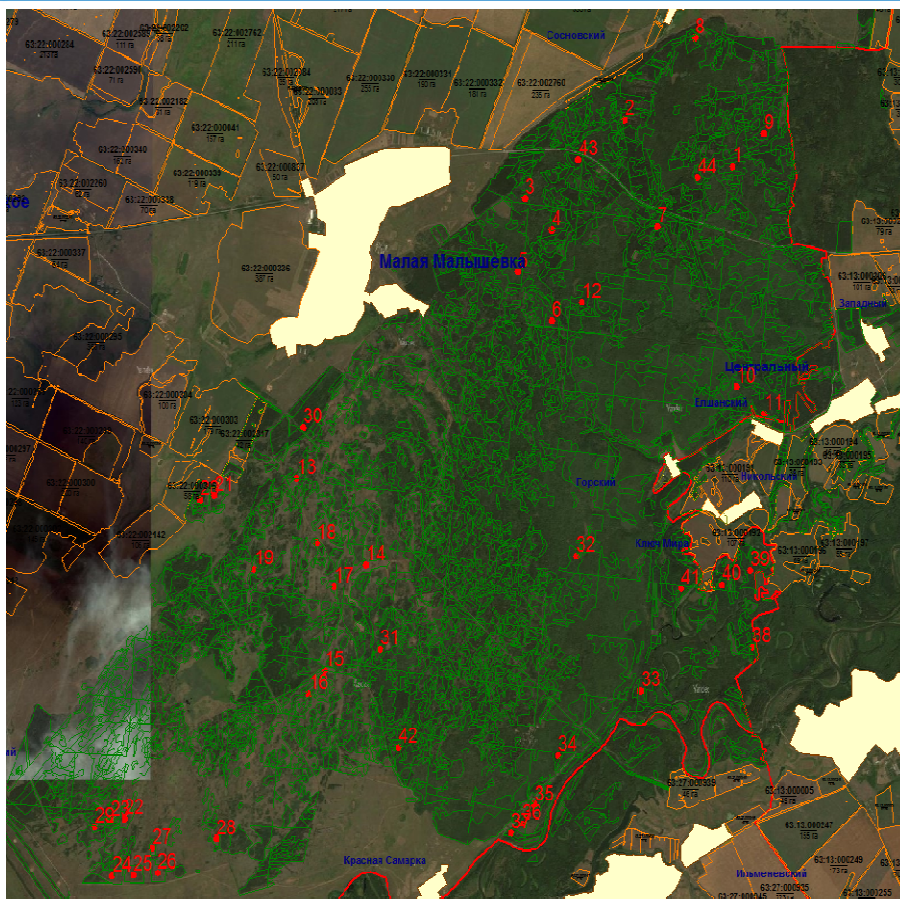
Каналы для составления композита	Расчёт признаков	Количество компонент PCA	Размер окна предобработки	Атмосферная коррекция	Период съемки	Точность на тестовом фрагменте
все	PCA + нормализация	136 (все)	3	+	12.04 - 19.10	0.8216
только 10 м	PCA + нормализация	68 (все)	3	+	12.04 - 19.10	0.8162
только 20 м	PCA + нормализация	68 (все)	3	+	12.04 - 19.10	0.8000

### Точность классификации для разных периодов





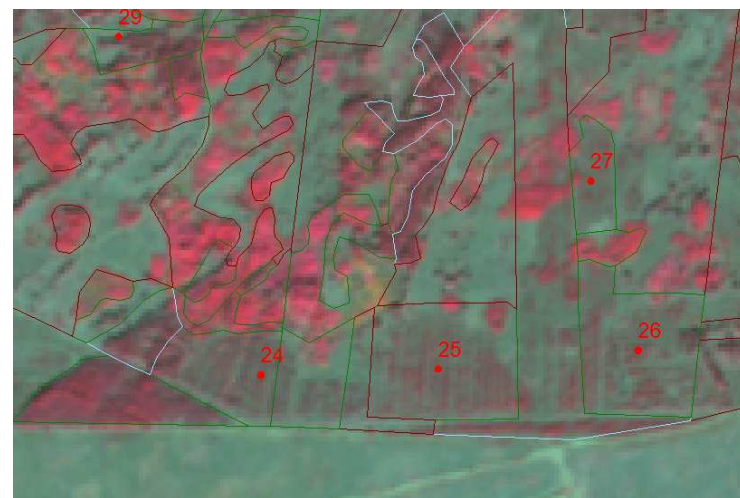
## Результаты анализа ошибок классификации с использованием данных наземных обследований



Выявлены 45 локализованных участков с ошибочной классификацией  
Из них по 13 произведены наземные обследования

### Причины неверной классификации по данным наземных наблюдений:

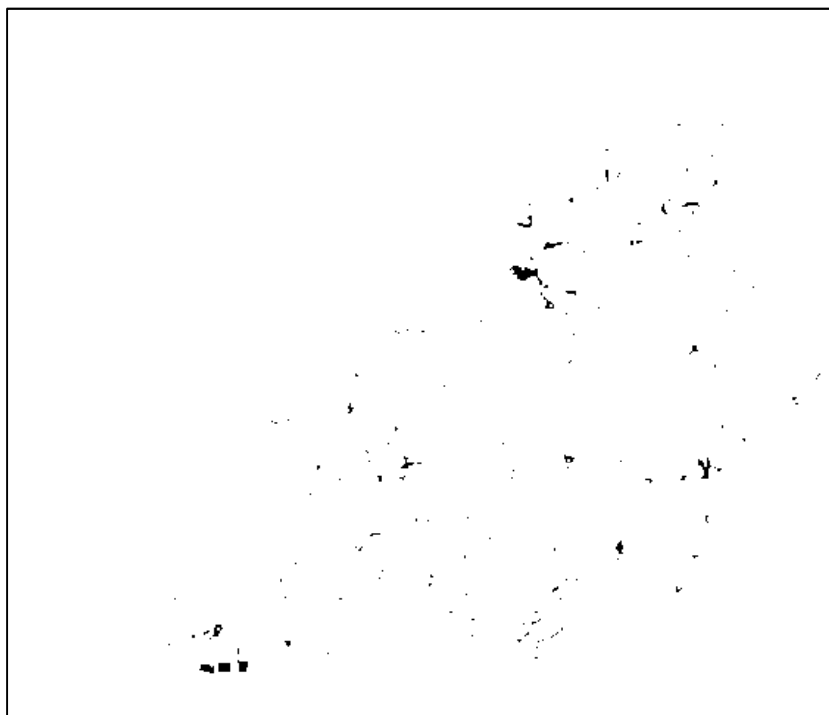
- 1) Разрастание дубовых и сосновых сообществ на территории бывших осиновых участков,
- 2) Разрастание сосновых участков на территорию березняков
- 3) Гари
- 4) Молодые посевы сосны
- 5) Неточности исходных данных таксации



Участки молодых посевов сосны



## Результаты классификации после уточнения таксации



Ошибка классификации



Результат классификации после  
уточнения таксации  
Точность 0.8385



## ВЫВОДЫ

1. Построение эффективного классификатора по таксации и снимкам Sentinel-2 принципиально возможно, и он может быть применим в вышеуказанных прикладных задачах, в частности - для уточнения таксации;
2. Пространственная обработка повышает точность классификации от 2-10% и даже выше при использовании только весенних данных;
3. Наилучший результат классификации при базовом варианте пространственной обработки составляет 0,8216 и соответствует использованию: атмосферной коррекции, всех снимков за сезон, нормализации признаков и PCA для расчёта признаков;
4. Периоды съёмки в порядке убывания информативности: весна, весь сезон, весна и осень, осень, лето;
5. Данные NDVI достаточны для кластеризации GEM, но совершенно недостаточны для качественной классификации. Дополнение 10-метровых каналов 20-метровыми немного повышает качество классификации. Добавление ко всем каналам данных NDVI не даёт прироста точности классификации. Самые информативные каналы - B, NNIR(20м), R, G. Примечательно, что широкий NIR (10 м) менее информативен.
6. Ошибки классификации сосредоточены в местах гарей, разрастания дубовых и сосновых пород по березнякам и осинникам, неточной таксации и молодых посадок леса.



## Список литературы

1. Федосеев ВА. Классификация гиперспектральных спутниковых изображений по эталонным участкам территории. In: Информационные технологии и нанотехнологии (ИТНТ-2018): сборник трудов IV международной конференции и молодежной школы. Самара: Новая техника; 2018. p. 966–75.
2. Кедров АВ, Тарасов АВ. Классификация лесной растительности методом нейронных сетей. Вестник ПНИПУ. 2017;22:44–54.
3. Борзов СМ, Потатуркин ОИ. Классификация типов растительного покрова по гиперспектральным данным дистанционного зондирования Земли. Вестни НГУ Серия: Информационные технологии. 2014;12(4):13–22.
4. Перепечина ЮИ, Глушенков ОИ, Корсиков РС. Учет и оценка лесов, возникших на сельскохозяйственных землях с использованием данных дистанционного зондирования Земли. Лесной журнал. 2016;(4):71–80.
5. Immitzer M, Vuolo F, Atzberger C, Immitzer M, Vuolo F, Atzberger C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. Remote Sensing. 2016;8(3):166.





**САМАРСКИЙ** УНИВЕРСИТЕТ  
SAMARA UNIVERSITY

## БЛАГОДАРЮ ЗА ВНИМАНИЕ

доцент кафедры геоинформатики и  
информационной безопасности,  
**А.Ю. Денисова,**  
[denisova\\_ay@geosamara.ru](mailto:denisova_ay@geosamara.ru)

ул. Московское шоссе, д. 34, г. Самара, 443086  
Тел.: +7 (846) 335-18-26 , факс: +7 (846) 335-18-36  
Сайт: [www.ssau.ru](http://www.ssau.ru), e-mail: [ssau@ssau.ru](mailto:ssau@ssau.ru)