



ИНСТИТУТ  
ГЕОЛОГИИ И  
МИНЕРАЛОГИИ  
имени В.С. Соболева СО РАН

# ПРИМЕНЕНИЕ МОРФОМЕТРИЧЕСКОГО АНАЛИЗА И УПРАВЛЯЕМОЙ КЛАССИФИКАЦИИ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ ПРИ ГИС-КАРТОГРАФИРОВАНИИ КРИОГЕННОГО РЕЛЬЕФА ОСТРОВА АРГА (ДЕЛЬТА РЕКИ ЛЕНЫ)

Картозия Андрей Акакиевич  
к.г.-м.н., ученый секретарь / н.с.

Москва  
2024



Локализация подверженных термоденудационным процессам участков земной поверхности, оценка интенсивности этих процессов и прогноз трансформации геолого-геоморфологического каркаса территории обладают высокой степенью актуальности.



Обусловленные термоденудационными процессами изменения растительности, почвенного покрова и увлажненности территории отражаются в спектральных характеристиках земной поверхности и могут наблюдаться на космических снимках.

Развитие облачных технологий, методов обработки больших объемов ДДЗ и применение методов машинного обучения в задачах управляемых классификации и кластеризации позволяют картографировать поверхность Земли с высокой степенью точности в полуавтоматическом режиме.





**Локализация подверженных термоденудационным процессам участков земной поверхности с помощью ГИС-картографирования с применением методов машинного обучения и морфометрического анализа**

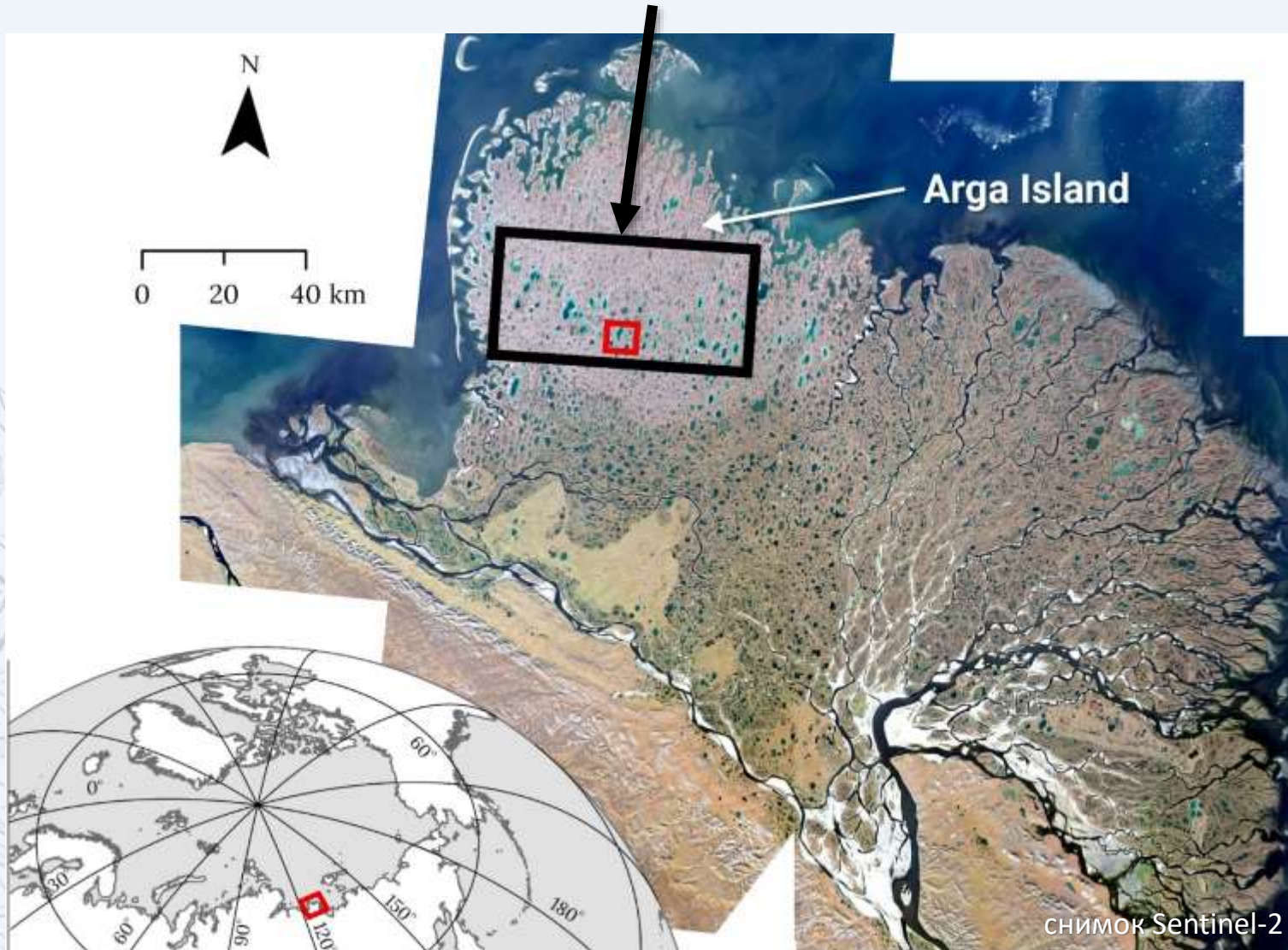


- I. обработка ДДЗ и их подготовка к анализу;
- II. определение классов земной поверхности, отражающих разную степень интенсивности проявления термоденудационных процессов и картографирование эталонных участков;
- III. полуавтоматическое ГИС-картографирование классов земной поверхности различными методами управляемой классификации и оценка точности моделей машинного обучения;
- IV. расчет индекса топографической позиции (ТРП) и определение площадного распространения выделенных классов земной поверхности в областях положительных и отрицательных значений индекса.





Размеры территории исследований – 60\*30 км



Google Earth Engine



Фрагмент территории исследований





Использовался снимок **Sentinel-2 (Level-2A)** от **29 июля 2021** года с процентом облачных пикселей 0,008006%. Далее к исходным каналам снимка (B1-12) были добавлены индексы, которые широко используются для оценки биомассы и увлажненности (<https://www.indexdatabase.de>).

**1) ndVi** |  $(B08 - B04) / (B08 + B04)$ .

**2) ndWi** |  $(B8 - B11) / (B8 + B11)$ .

**3) EVI** |  $2.5 * (B08 - B04) / ((B08 + 6.0 * B04 - 7.5 * B02) + 1.0)$ .

**4) TC wetness** |  $0.1509 * B02 + 0.1973 * B03 + 0.3279 * B04 + 0.3406 * B08 - 0.7112 * B11 - 0.4572 * B12$

**5) TC greenness** |  $-0.2848 * B2 - 0.2435 * B3 - 0.5436 * B4 + 0.7243 * B8 + 0.0840 * B11 - 0.1800 * B12$

После нормализации каналов и добавления индексов к композиту к нему был применен МГК, после чего первые три канала результатов этого преобразования были добавлены к исходному набору данных.

**6) PCA 1, 2, 3** – результаты преобразования методом Principal Component Analysis (методом главных компонент (МГК)) всего набора данных (каналы снимка и вышеперечисленные индексы). Метод используется для преобразования данных входного многомерного атрибутивного пространства и получения некоррелированных данных по трём новым осям. Этим достигается исключение избыточности исходных данных и получение новых трех каналов, которые могут быть использованы для анализа. В GEE реализован специальный инструмент Eigen Analysis.

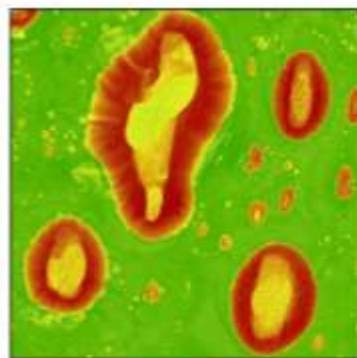
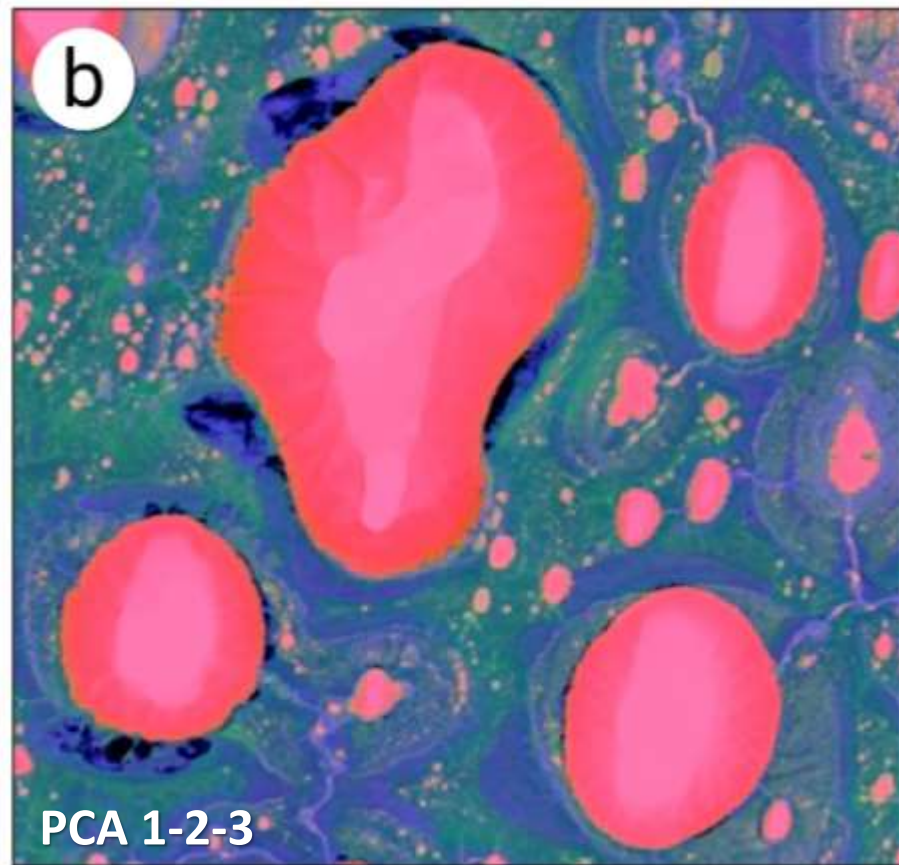
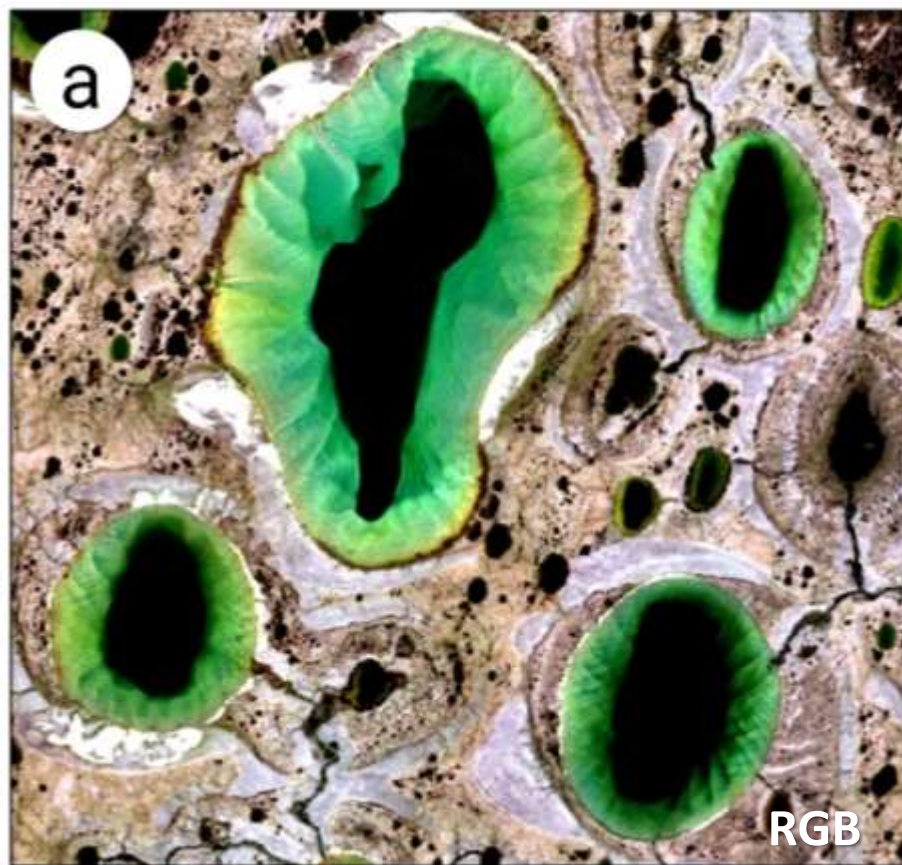


Name	Scale	*Min	*Max
<b>B1</b>	10	0,000	0,978
<b>B2</b>	10	0,010	0,827
<b>B3</b>	10	0,004	0,833
<b>B4</b>	10	0,003	0,843
<b>B5</b>	10	0,000	0,877
<b>B6</b>	10	0,000	0,860
<b>B7</b>	10	0,000	0,864
<b>B8</b>	10	0,000	0,896
<b>B8A</b>	10	0,000	0,898
<b>B9</b>	10	0,000	0,987
<b>B11</b>	10	0,002	0,937
<b>B12</b>	10	0,000	0,915

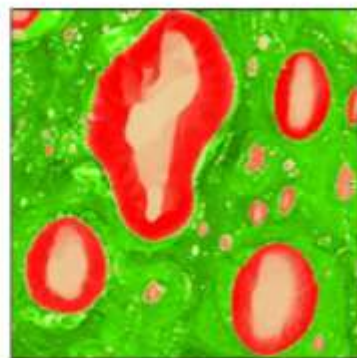
Name	Scale	*Min	*Max
<b>ndvi</b>	10	-0,982	0,826
<b>ndWi</b>	10	-0,728	0,993
<b>evi</b>	10	-0,142	0,597
<b>tc_green</b>	10	-0,065	0,225
<b>tc_wet</b>	10	0,006	0,676
<b>PCA1</b>	10	0,075	0,999
<b>PCA2</b>	10	0,109	0,952
<b>PCA3</b>	10	0,108	0,949

Анализируемое композитное изображение состояло из нормированных каналов исходного снимка Sentinel (B 1-12), а также результатов построения индексов и преобразования PCA. Всего 20 каналов, итоговый размер данных составил около 20 GB.

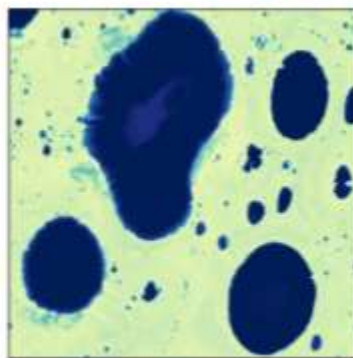




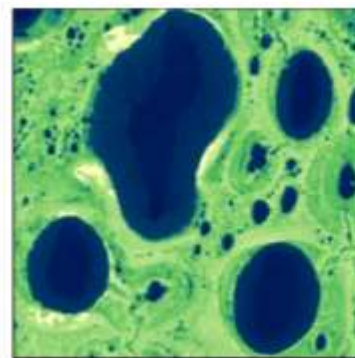
ndVi



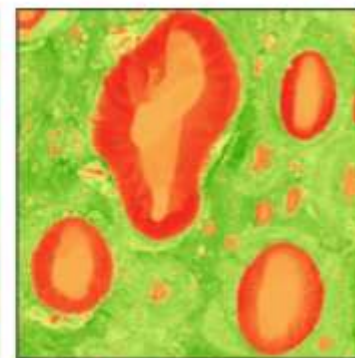
ndWi



EVI



TCw



TCg

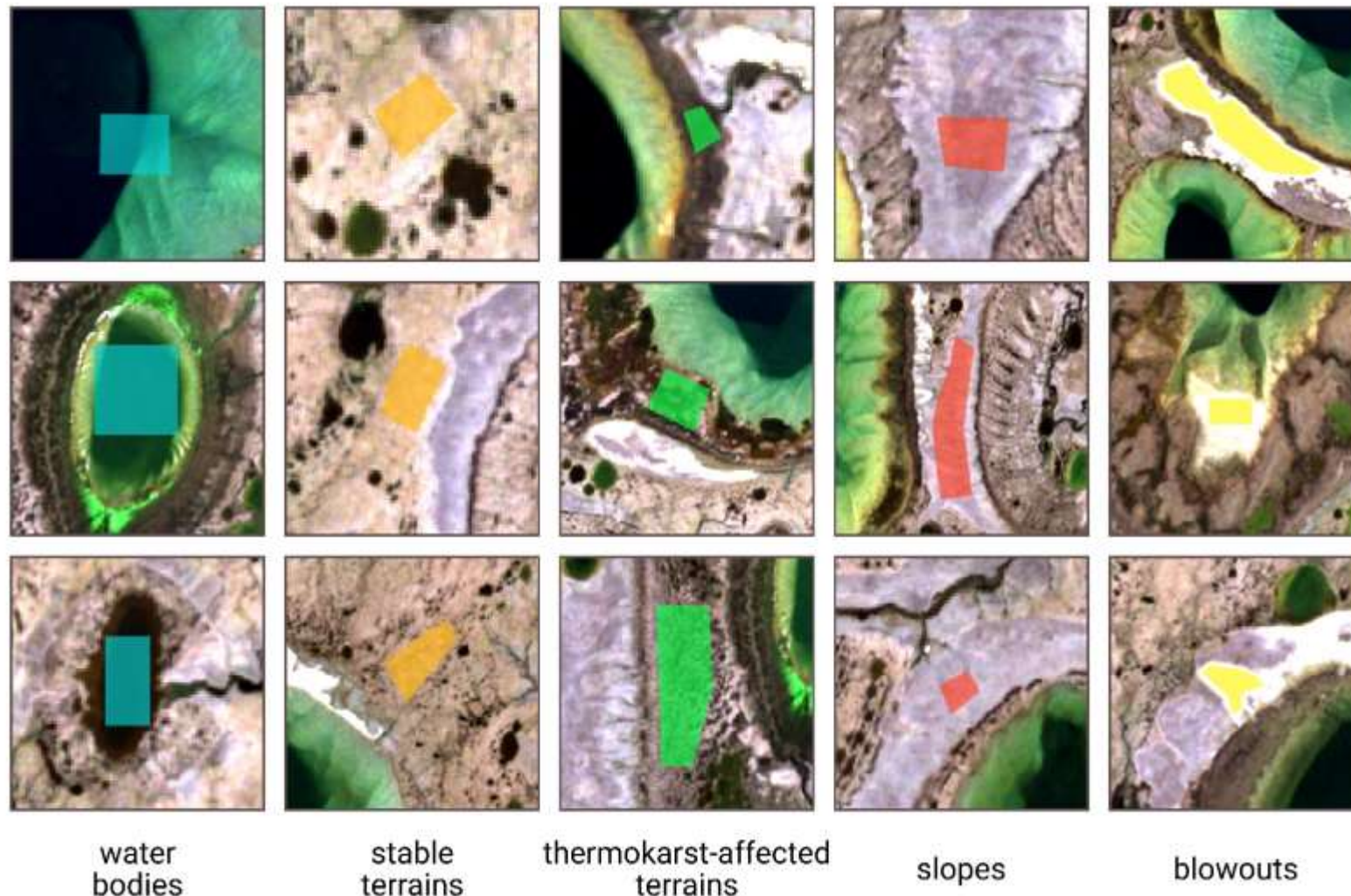




Выделяемые классы поверхностей\*:

- 1) водные объекты;
- 2) стабильные участки (пространства водоразделов);
- 3) участки с деградирующими ММП (плоские пространства дниц термокарстовых котловин и термоэрозионных долин);
- 4) склоны;
- 5) песчаные выдувы.

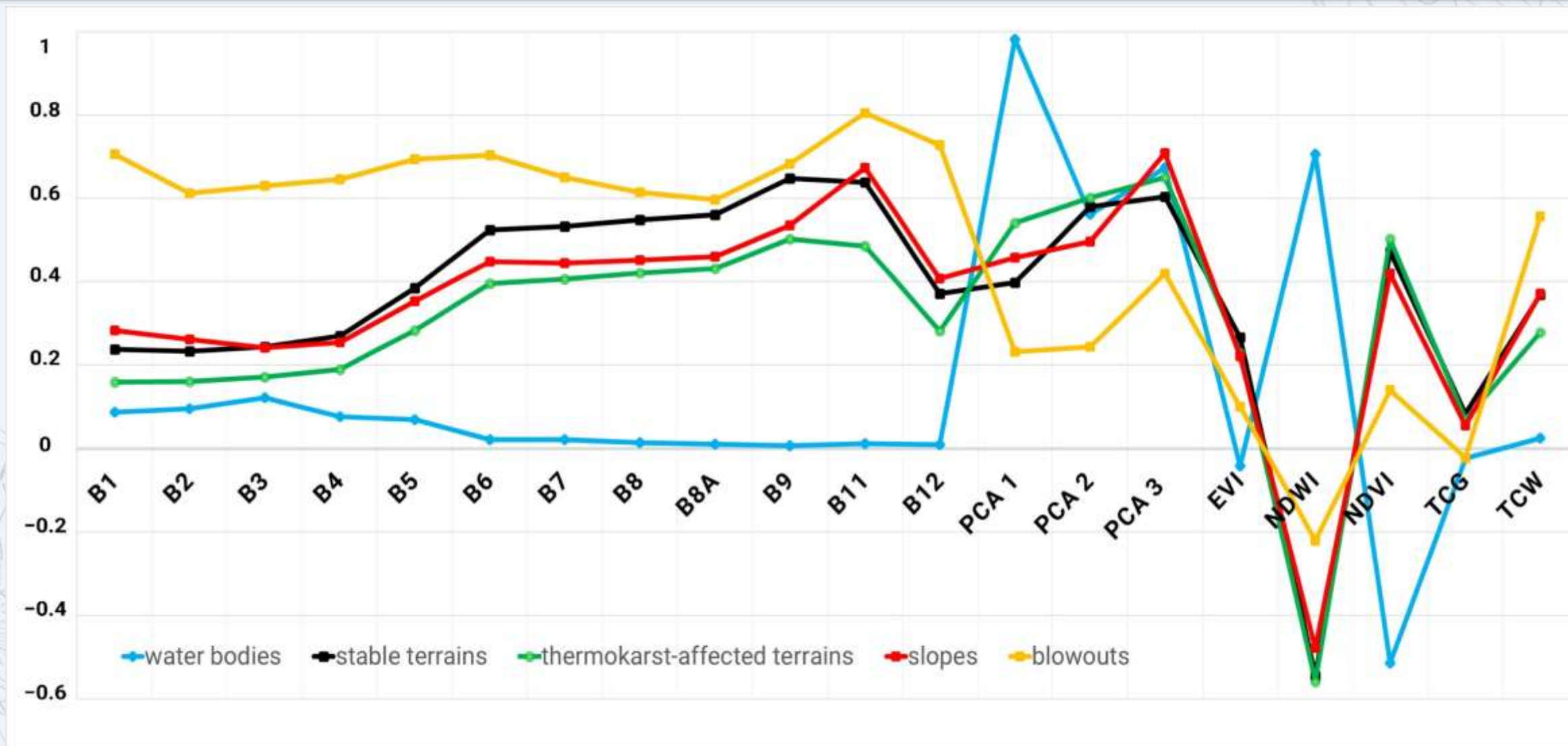
\*land cover classes



100 эталонных полигонов по 20 штук на каждый класс.

Тренировочный набор - 61 833 точки с уникальными значениями;  
набор для валидации - 38 972 точки. Соотношение: 0,63





Отметим разнонаправленность кривых значений классов в различных каналах и индексах, а также значительный разброс этих значений в некоторых каналах, необходимый для адекватной классификации.



В ходе исследования мы сравнили результаты классификации, полученные с помощью следующих методов:

- ✓ CART – дерево классификации и регрессии,
- ✓ GTB – градиентный бустинг,
- ✓ RF – случайного леса,
- ✓ KNN – K-ближайших соседей,
- ✓ SVM – опорных векторов,
- ✓ MIND – минимального расстояния.



Google Earth Engine



*Smile - Statistical Machine Intelligence and Learning Engine*

Kartozia, A. Using Google Earth Engine to Assess the Current State of Thermokarst Terrain on Arga Island (the Lena Delta). *Earth* **2024**, 5, 228-243.  
<https://doi.org/10.3390/earth5020012>





## Пример матрицы несоответствий для модели **Random forest**

		Classification (predicted)					Total
		1	2	3	4	5	
Ground Truth (Actual)	1	9656	0	117	0	0	9773
	2	0	6158	139	104	0	6401
	3	60	9	6205	3	0	6277
	4	0	52	37	9344	5	9438
	5	0	0	0	0	7083	7083
	Total	9716	6219	6498	9451	7088	38972

True Positive (TP)	False Positive (FP)	True Negative (TN)	False Negative (FN)
9656	60	29139	117
6158	61	32510	243
6205	293	32402	72
9344	107	29427	94
7083	5	31884	0

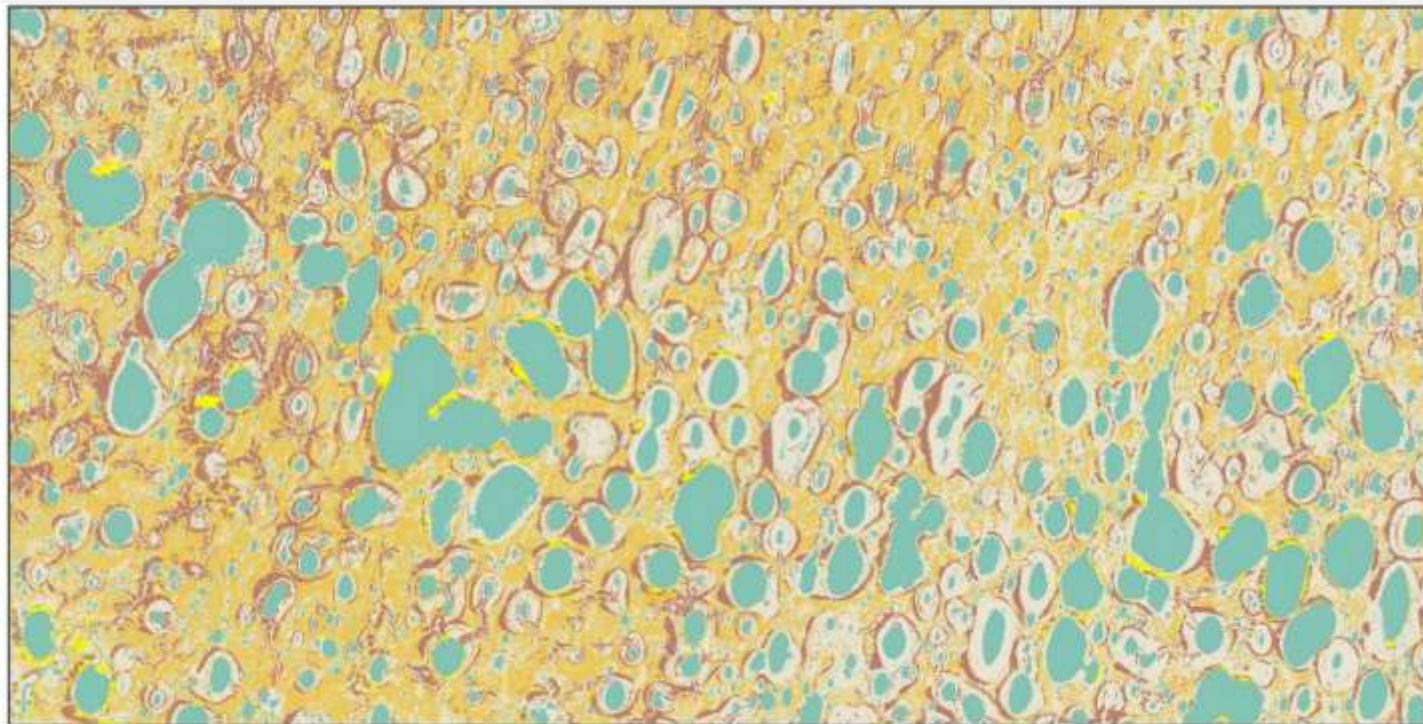
- 1) водные объекты;
- 2) стабильные участки;
- 3) участки с деградирующими ММП;
- 4) склоны;
- 5) песчаные выдувы.

Для оценки точности классификаторов мы посчитали их матрицы несоответствий (confusion matrix). На их основе анализировался ряд метрик качества модели классификации.



Model	Recall	Precision	F1 Score	OA	Kappa	Kulczynski's	CSI	BER
CART	97,82%	97,86%	97,82%	97,81%	97,26%	97,84%	95,68%	2,16%
GTB	98,24%	98,25%	98,23%	98,23%	97,79%	98,24%	96,48%	1,76%
KNN	98,68%	98,69%	98,68%	98,67%	98,34%	98,68%	97,37%	1,32%
MIN_D	98,03%	98,07%	98,04%	98,03%	97,54%	98,05%	96,10%	1,95%
RF	98,64%	98,63%	98,62%	98,62%	98,28%	98,63%	97,27%	1,37%
SVM	98,13%	98,18%	98,12%	98,12%	97,65%	98,16%	96,31%	1,84%



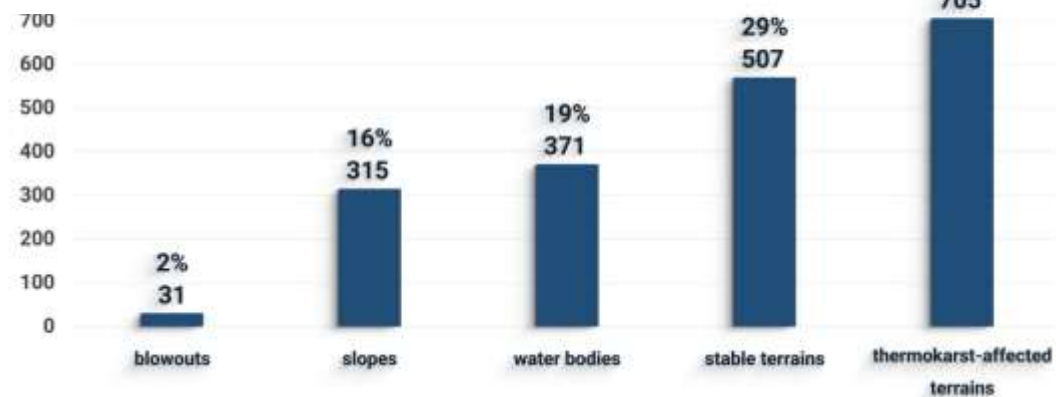


0 5 km



Карта распространения  
выделенных классов  
земной поверхности

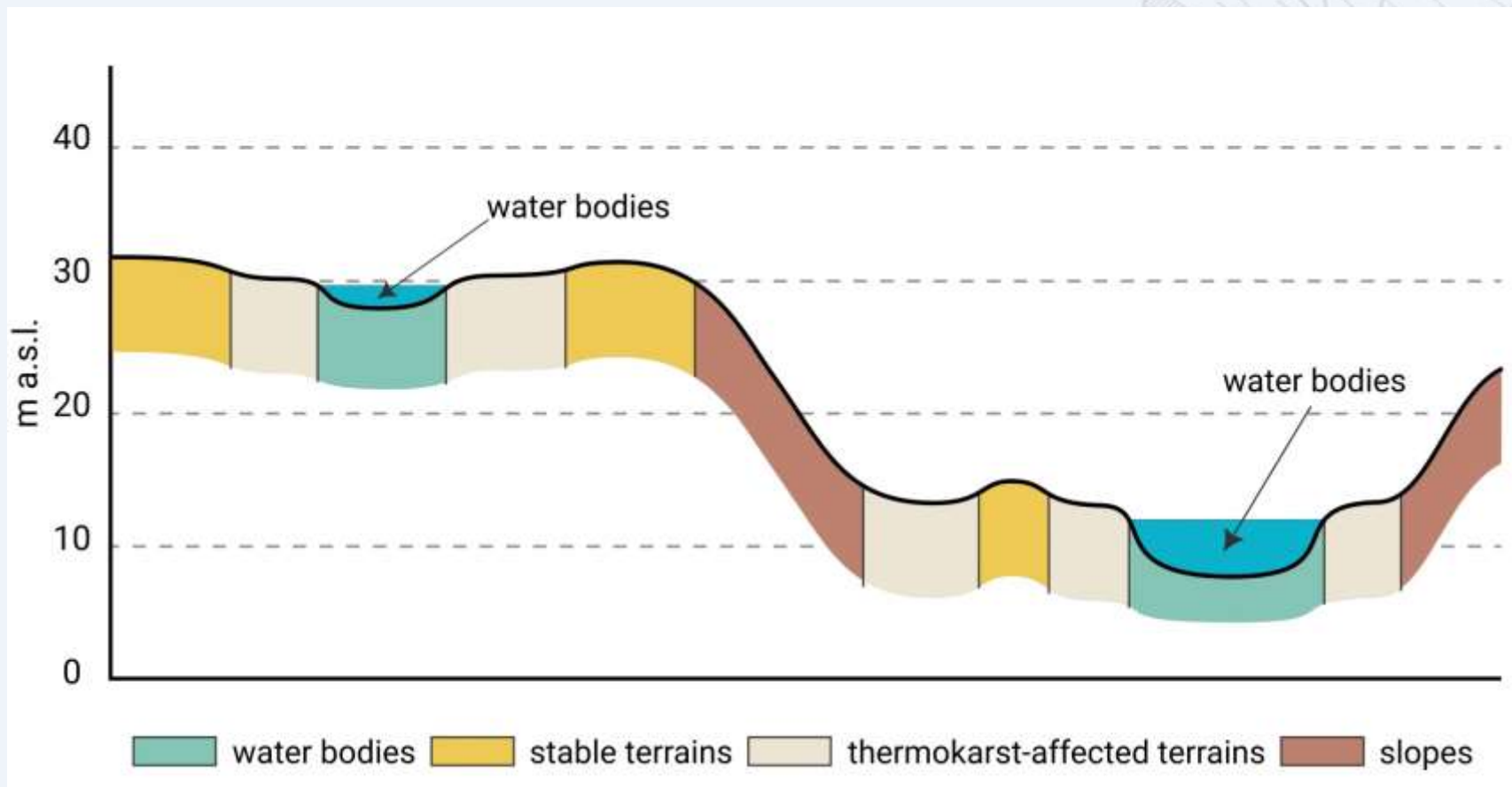
Площадное распространение  
выделенных классов  
в % от анализируемой  
территории

Area (% , km<sup>2</sup>)

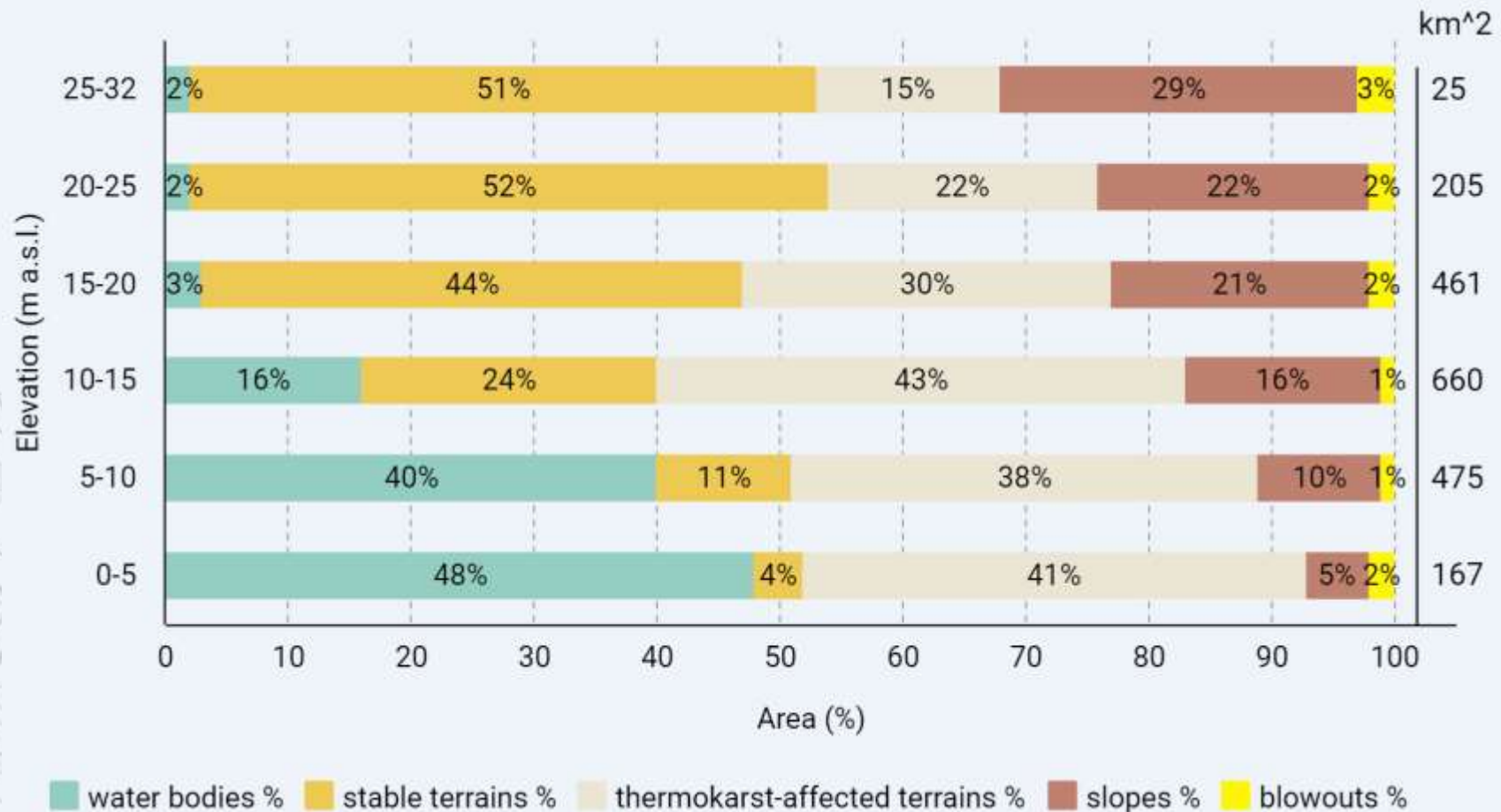


Участки, относящиеся к классу земной поверхности, эталонные полигоны которого находились на пространствах водоразделов с недеградировавшими ММП, могут располагаться на днищах термокарстовых котловин и термоэрозионных долин.

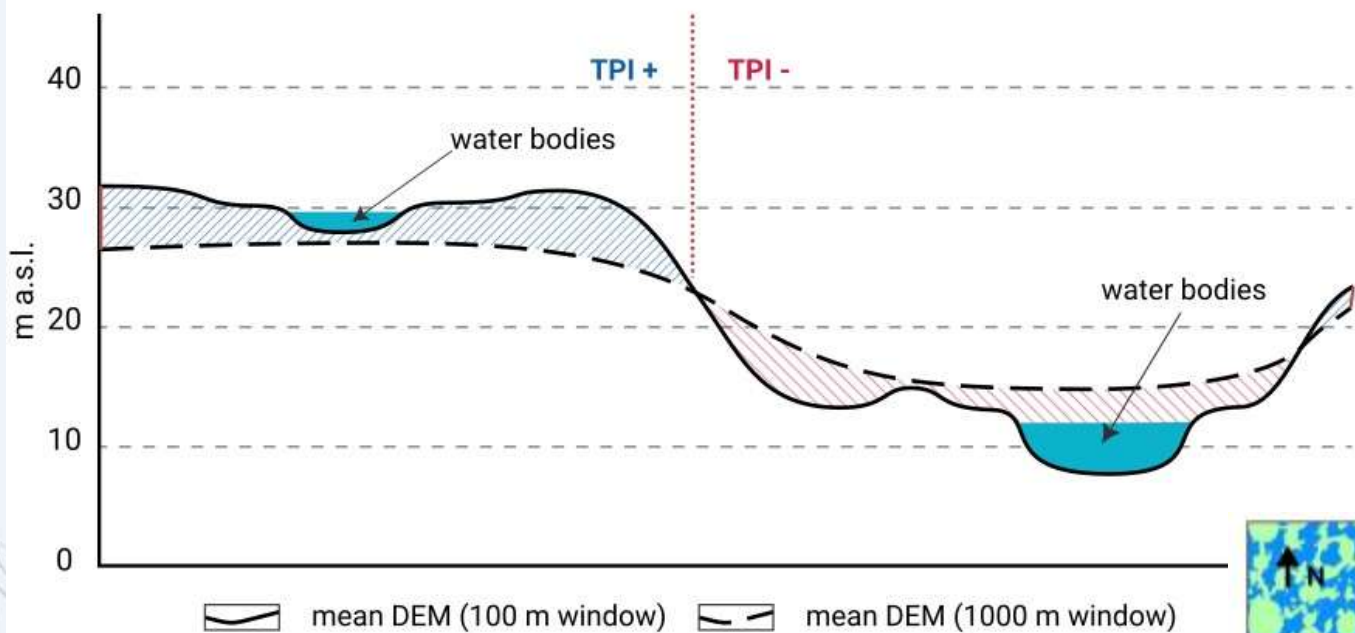
В свою очередь, третий класс, эталонные полигоны которого выбирались на участках, претерпевающих активное воздействие термоденудационных процессов, определяются также и на водораздельных поверхностях.





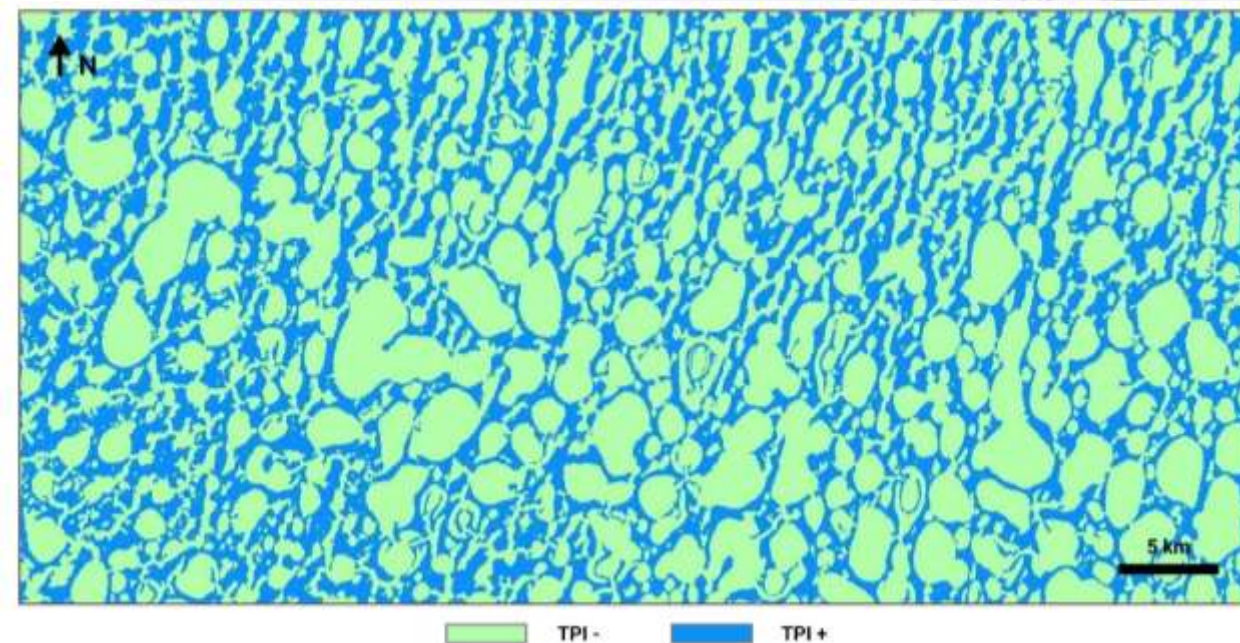


Таким образом, было необходимо отличить эти участки друг от друга применив морфометрический анализ. То есть выделить на территории водораздельные поверхности и термокарстовые котловины с термоэрозионными долинами, что было сделано посредством расчета ТРІ.

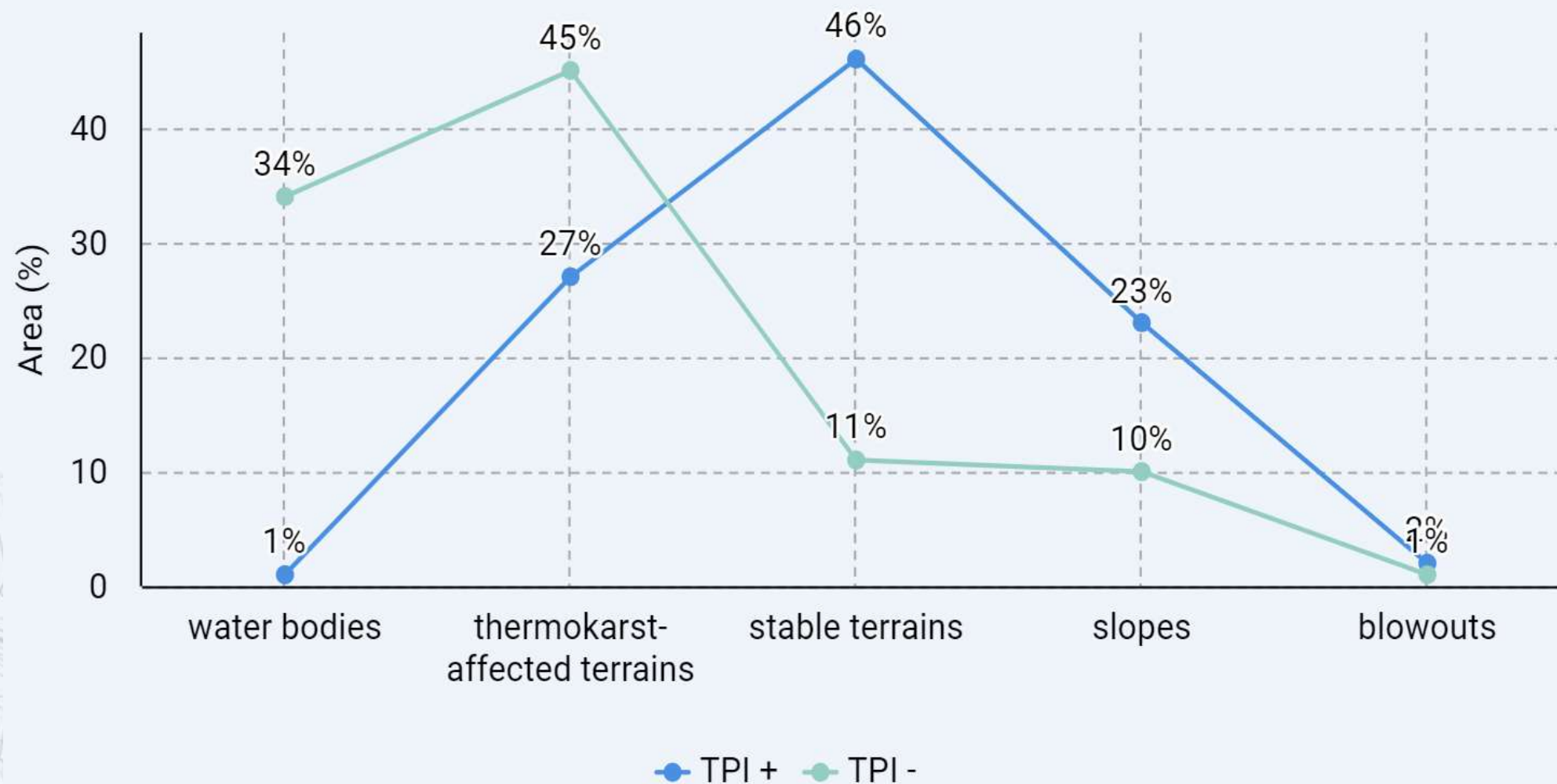


Topographic position index (TPI) рассчитывается как разница между высотой точки и средним значением высоты ее окрестностей в определенном окне.

После ряда пробных расчетов мы воспользовались ТПИ, рассчитанным как разница между ЦМР, усредненной в 100 м окне, и ЦМР усредненной в 1000 м окне. Исходной ЦМР послужил ArcticDEM.

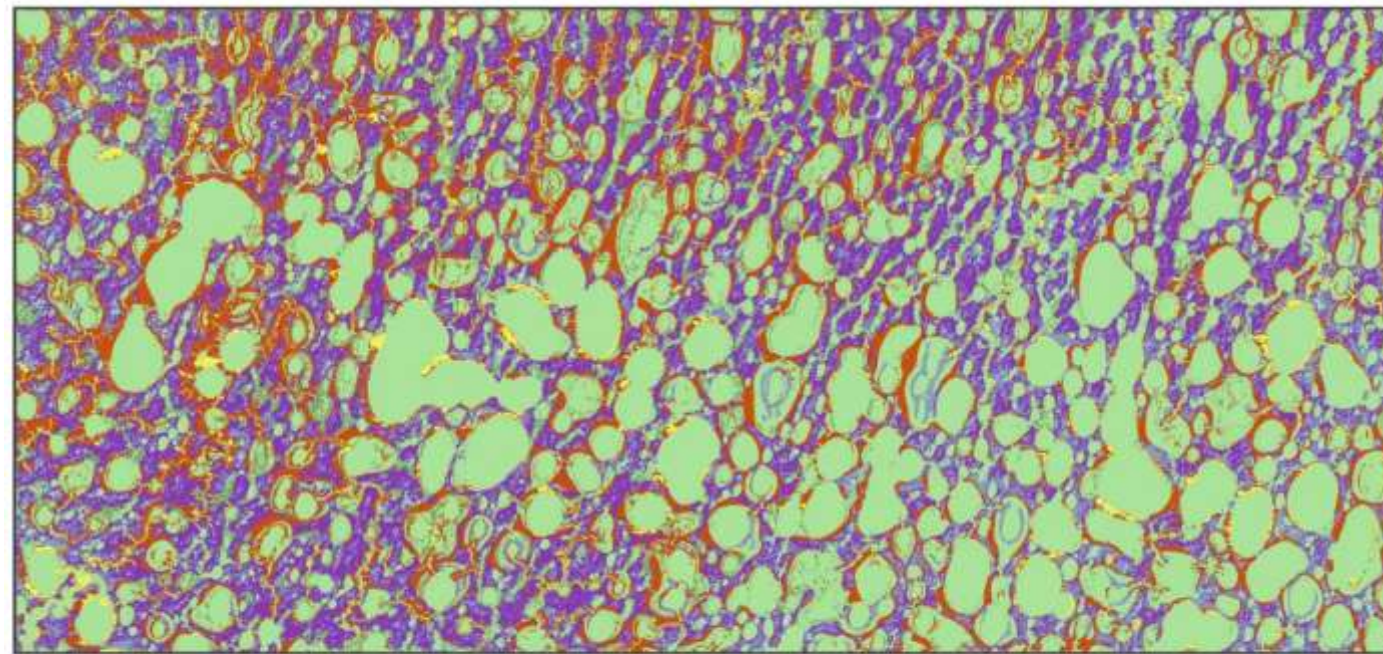
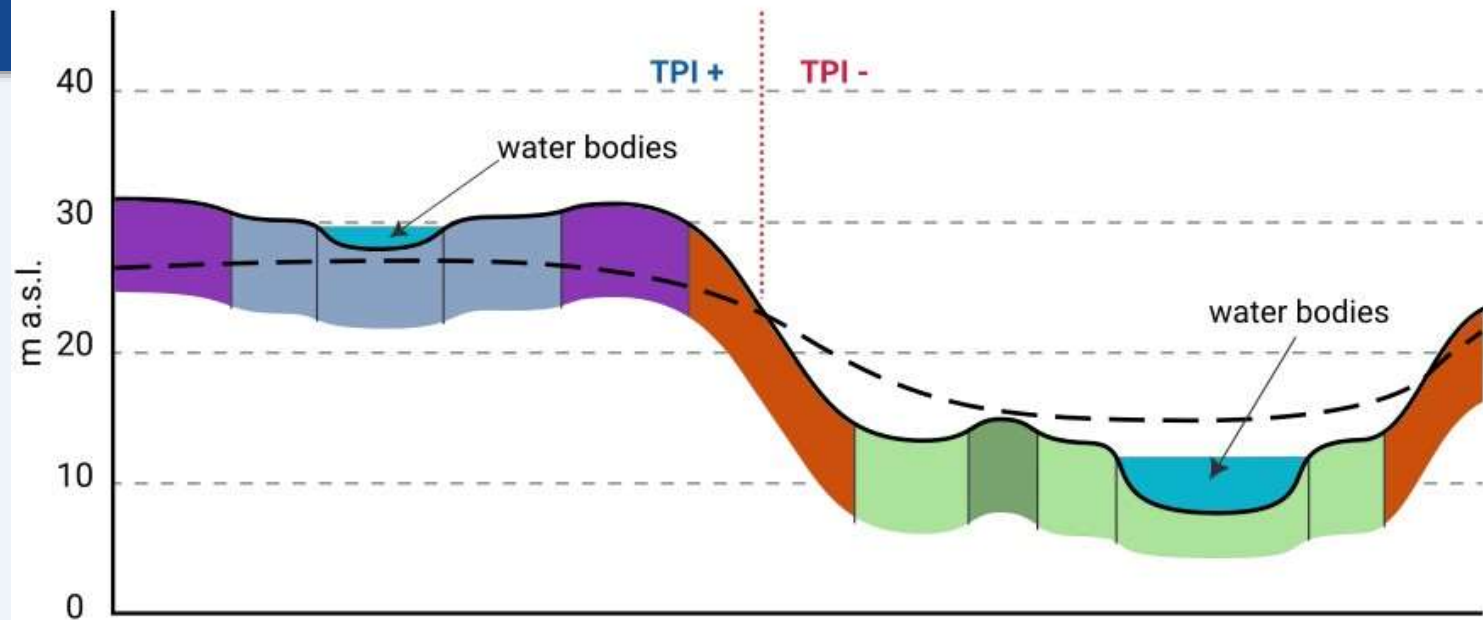






### Итоговые классы земной поверхности:

- 1) участки водоразделов с недеградировавшими ММП (2 класс, **положительный TPI**);
- 2) участки водоразделов с ММП на начальной стадии деградации (1 и 3 классы, **положительный TPI**);
- 3) днища долин и котловин с деградировавшими ММП, но претерпевающими влияние термоденудационных процессов (1 и 3 классы, **отрицательный TPI**);
- 4) стабильные участки днищ долин и котловин с деградировавшими ММП (2 класс, **отрицательный TPI**);
- 5) склоны; 6) песчаные выдувы.



0 5 km

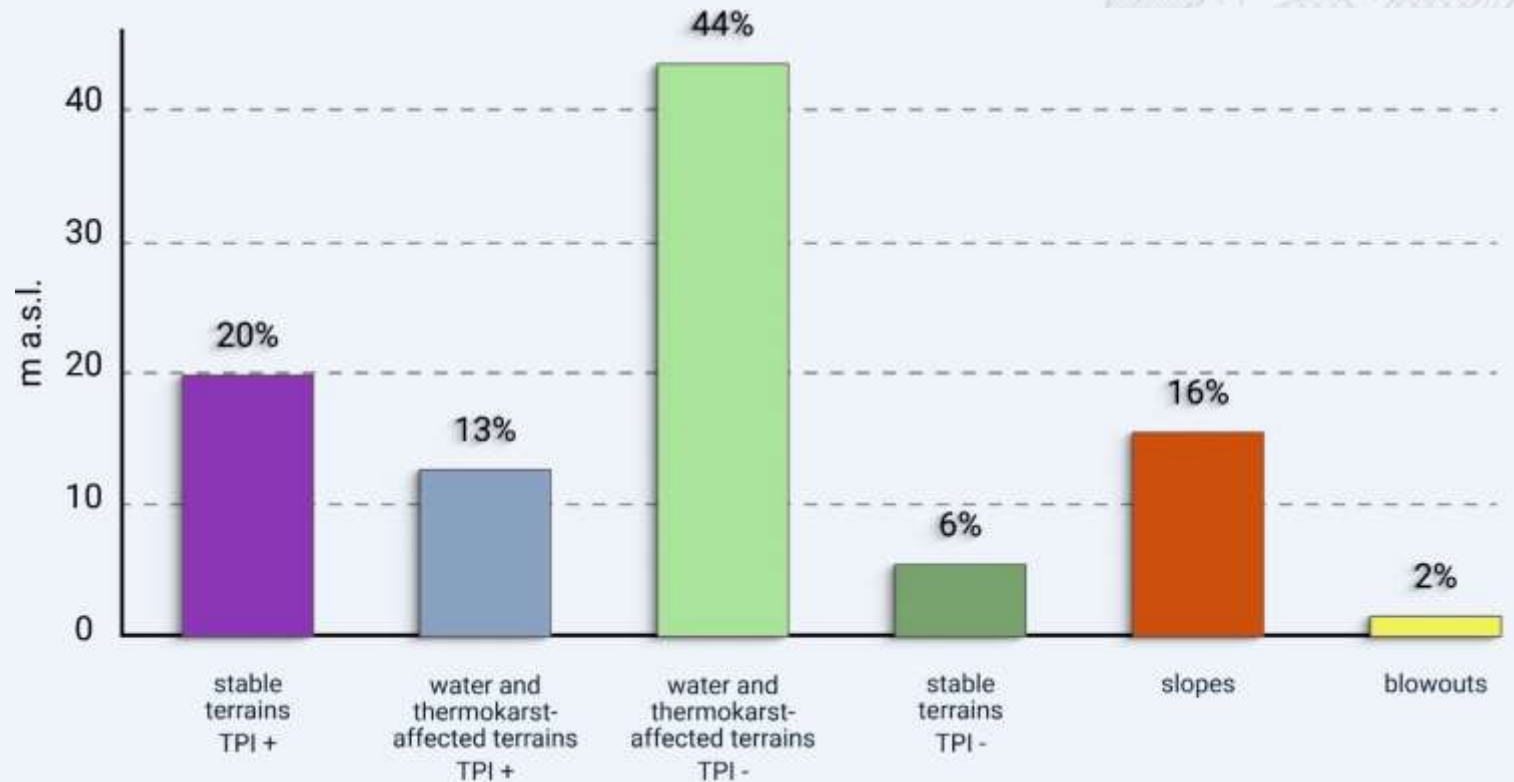






## Итоговые классы земной поверхности:

- 1) участки водоразделов с недеградировавшими ММП (2 класс, **положительный TPI**);
- 2) участки водоразделов с ММП на начальной стадии деградации (1 и 3 классы, **положительный TPI**);
- 3) днища долин и котловин с деградировавшими ММП, но претерпевающими влияние термоденудационных процессов (1 и 3 классы, **отрицательный TPI**);
- 4) стабильные участки днищ долин и котловин с деградировавшими ММП (2 класс, **отрицательный TPI**);
- 5) склоны; 6) песчаные выдувы.





- I. подготовлен набор ДДЗ, который может быть использован во время исследований по различным предметным областям.
- II. определены классы земной поверхности, отражающие разную степень интенсивности проявления термоденудационных процессов, и проведено картографирование эталонных участков;
- III. выполнено полуавтоматическое ГИС-картографирование классов земной поверхности различными методами управляемой классификации и оценена точность каждой из моделей;
- IV. выполнен расчет индекса топографической позиции (TPI) и определено площадное распространение выделенных классов земной поверхности в областях положительных и отрицательных значений индекса.

Таким образом, с помощью разработанной методики ГИС-картографирования выполнена локализация участков земной поверхности, которые относятся к разным стадиям деградации ММП и трансформации геолого-геоморфологического каркаса территории.

Данные о площадном распространении выделенных классов земной поверхности иллюстрируют современное состояние криогенного рельефа территории и их можно использовать для мониторинга изменений природной среды под воздействием изменения климатических условий.

Разработанную методику можно использовать для решения широкого спектра задач как во время научных исследований, в том числе междисциплинарных, так и во время производственных работ.

*Исследование выполнено за счет гранта Российского научного фонда № 23-77-01029, <https://rscf.ru/project/23-77-01029/>.*





# СПАСИБО ЗА ВНИМАНИЕ!



**Картозия Андрей Акакиевич**

Ученый секретарь ИГМ СО РАН

К.г.-м.н., н.с. лаб. «ГИТ и ДЗ» ИГМ СО РАН

**Email:** andrei.kartoziaa@igm.nsc.ru

**Тел.:** +7 993 010 04 86

**ResearchGate:** [www.researchgate.net/profile/Andrei-Kartoziaa](http://www.researchgate.net/profile/Andrei-Kartoziaa)

