

УДК: 551.435:528.06(571.65):004.7

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ СКЛОНОВ ПО КОСМИЧЕСКИМ СНИМКАМ НА ПРИМЕРЕ НЕКОТОРЫХ РАЙОНОВ СЕВЕРНОГО ПРИОХОТЬЯ

Колегов П. П.

*ФГБУН Северо-Восточный комплексный научно-исследовательский институт
им. Н. А. Шило ДВО РАН, г. Магадан
E-mail: kolegovpp@gmail.com*

В статье представлены алгоритмы машинного обучения, основанные на методах SVM, Random Forest и k-Nearest Neighbors, для классификации спектрозональных космических снимков (Sentinel-2). Проведен сравнительный анализ моделей, полученных с использованием вышеуказанных методов и ручной классификации, представлены параметры точности, полноты и F-меры. Практическое применение включает реклассификацию изображений по степени открытости склонов от растительности и оценку их морфометрических характеристик. Получены данные о площади распространения динамически активных склонов и их экспозиции: горы Дел-Урэкчэн (4.0 %, 240–270°), Ольское плато и его отроги (1.4 %, 165–285°), Туманский хребет (9.4 %, 150–195°).

Ключевые слова: классификация, склоны, машинное обучение, ГИС, Северное Приохотье.

DOI: 10.34078/1814-0998-2023-4-62-71

ВВЕДЕНИЕ

С развитием технологий в области дистанционного зондирования Земли открываются новые перспективы в области геоморфологического анализа и картографии. Космические изображения местности, получаемые с помощью современных спутниковых систем, предоставляют обширные объемы данных о поверхности планеты. Эти данные могут быть использованы для классификации и анализа геоморфологических обстановок, что имеет огромное значение в различных областях, включая геологию, географию, экологию, градостроительство.

Одной из ключевых задач в геоморфологическом анализе является классификация ландшафтных элементов на спутниковых изображениях, чему посвящена статья И. Квона с коллегами (2022). Эти элементы могут включать в себя горы, долины, реки, озёра, леса, поля и многое другое. Выделение таких элементов может быть сложной задачей, но методы машинного обучения (Machine Learning – ML) позволяют справиться с этой задачей быстро, качественно и с высокой точностью.

В работе Д. Бздока с коллегами (2018) приводится сравнение методов классической статистики с машинным обучением. Так, авторами статьи выделяются отличия в способности обрабаты-

вать данные при увеличении числа переменных в каждом наблюдении. Изначально классическое статистическое моделирование было предназначено для данных с ограниченным числом входных переменных и относительно умеренным объемом выборки. В таких условиях модель может заполнять пробелы в неучтенных аспектах системы, однако с увеличением количества переменных и их взаимосвязей статистические выводы становятся менее точными. В отличие от статистики, методы машинного обучения (далее – ML-методы) не требуют подробной информации о конкретных измеренных выборках; их применение ориентировано на выделение полезных свойств для распознавания конкретных аспектов, что особенно ценно в случае большого числа переменных при ограниченном размере выборки.

Кроме того, классические статистические методы, как правило, требуют ручного выделения признаков, ограничены в адаптации к сложным нелинейным зависимостям и могут быть чувствительны к шуму, в то время как ML-методы автоматически извлекают признаки, более гибки и способны моделировать сложные взаимосвязи в данных, обладают устойчивостью к шуму и эффективно работают с большими объемами информации. В контексте классификации космических изображений ML-методы обеспечивают более точные и адаптивные модели, что особенно важно при сложных структурах данных, характерных для этого вида изображений.

Применение ML-методов позволяет автоматизировать процесс классификации и сделать его более точным и быстрым. Модели машинного обучения, такие как нейронные сети (Neural Networks), случайные леса (Random Forest), метод опорных векторов (Support Vector Machine) и др., могут обучаться на размеченных данных, чтобы распознавать различные классы местности на изображениях. Они способны выявлять сложные и нелинейные закономерности, которые могут быть невидимыми для человеческого глаза и/или традиционных методов обработки изображений. С более подробным обзором разнообразных практик, применяемых для анализа спутниковых и аэросъемочных изображений, с описанием моделей и алгоритмов, можно ознакомиться на открытом ресурсе GitHub (Techniques..., 2023).

В свете растущего объема снимков от спутниковых миссий (например, Landsat, Sentinel-1 и -2), покрытие которых охватывает почти всю территорию Земли, а снимки повторно публикуются относительно часто (от недели до года), возникает необходимость в эффективных методах анализа, интерпретации и поиска различий в представляемых данных. Модели машинного обучения способны обрабатывать такие объемы материалов и выявлять различия в изображениях (например, снятых в разные сезоны года или время суток), которые могли бы остаться незамеченными при ручном сопоставлении.

Кроме того, использование ML-методов позволяет улучшить прогнозирование и мониторинг различных процессов, как погодных (Voosen, 2020), так и геоморфологически опасных, например обвалов (Zhou et al., 2018) или по-

жаров (Gizatullin, Alekseenko, 2022). Модели алгоритмов могут анализировать временные данные и выявлять различия (Zhang et al., 2023), что может быть полезным для управления экологическими ресурсами и оценки рисков их использования.

Целью данной работы является знакомство читателя с алгоритмами и результатами применения техник машинного обучения для классификации спектрально-анализованных снимков и выделения различных типов ландшафтов на примере некоторых районов Северного Приохотья.

Основной задачей исследования является классификация ландшафтов на три категории: 1) открытые от растительности склоны, 2) задернованные склоны, 3) склоны с активными склоновыми процессами. Выбранные категории позволяют дать качественную оценку морфодинамической активности исследуемых территорий.

МЕТОДЫ И МАТЕРИАЛЫ ИССЛЕДОВАНИЯ

Все этапы работы с космическими изображениями проводились с помощью геоинформационной системы QGIS (2010) с модулем Orfeo ToolBox (ОТВ, 2012). В качестве основы для классификации использовались следующие материалы: 1) спектрально-анализованные снимки с космического аппарата Sentinel-2; 2) цифровая модель рельефа (ЦМР) Arctic DEM (2018); 3) векторные слои дорог и инженерных сооружений по данным openstreetmap.org.

Материалами для классификации являлись спектрально-анализованные снимки, снятые аппаратом Sentinel-2, которые были получены в центре открытого доступа «Коперник» (Copernicus..., 2014): для гор Дел-Урэкчэн и Ольского плато: мис-

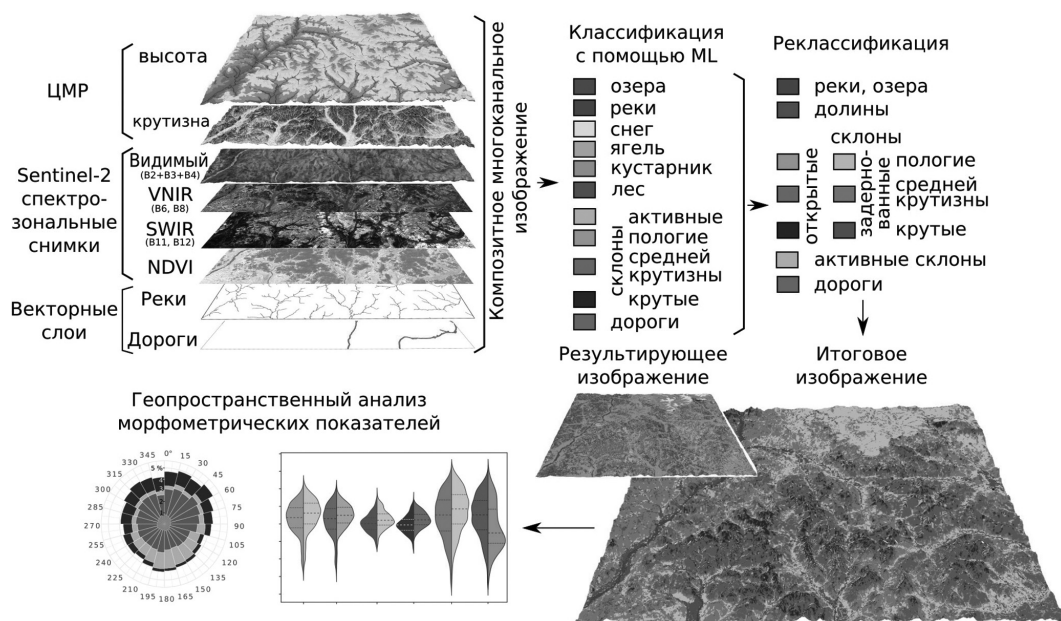


Рис. 1. Общий принцип анализа композитных изображений, использованный в работе.

Fig. 1. General principle of composite image analysis used in the research.

сия – S2A, уровень – MSIL1C, дата – 27.08.2022, тайл – T57VUJ; для Туманского хребта: миссия – S2A, уровень – MSIL1C, дата – 06.09.2022, тайл – T56VPN.

Разрешение снимков в зависимости от канала (спектральный диапазон в нм) составило: B2 (490), B3 (560), B4 (665) и B8 (842) – 10 м/пиксель; B5 (705), B6 (740), B7 (783), 8A (865) и B12 (2190) – 20 м/пиксель (Spatial Resolution, 2012). Из представленных каналов были рассчитаны композитные изображения (Kwong et al., 2022) нормализованного разностного индекса вегетации (Normalized Difference Vegetation Index – NDVI) и расширенного вегетационного индекса (Enhanced Vegetation Index – EVI), используемые для классификации растительности и их ареалов распространенности. Цифровая модель рельефа Arctic DEM, разрешение которой составляет 2 м/пиксель, опубликована в свободном доступе на сайте издателя (Arctic DEM, 2018). Из ЦМР были рассчитаны модели крутизны склонов и водотоков. Векторные данные дорог и инженерных сооружений были загружены с сайта openstreetmap.org (Geofabrik..., 2018). Далее они переводились в растровый формат для последующей работы. Полученные изображения сводились в один композитный снимок, разрешение которого составило 10 м/пиксель. Общий алгоритм работы ML-метода представлен на рис. 1.

Методы машинного обучения, представленные в данном исследовании, доступны в программной среде Orfeo ToolBox, которая разработана Центром научно-технических исследований и разработок в области аэрокосмических технологий (Centre National d'Etudes Spatiales, CNES). Программа реализована для C++ и Python как отдельное приложение, а также в виде отдельного модуля в QGIS. Применялись также стандартные средства данной ГИС, основанные на программных средах GDAL и GRASS GIS.

Преимуществом модуля ОТВ является его способность работать с большим объемом данных и применять множество различных алгоритмов машинного обучения для повышения точности классификации, но для исследователей главным аспектом является его доступность непосредственно в ГИС. ОТВ использует различные алгоритмы машинного обучения, основанные как на нейронных сетях, так и на классических статистических методах – опорные вектора, бустинг, Байес и др. В работе применены такие методы, как: k-ближайших соседей (KNN), случайные леса (Random Forest) и метод опорных векторов (Support Vector Machine). Данные методы выбраны, потому что они хорошо подходят для многоклассовой классификации.

Метод k-ближайших соседей (KNN) является простым и эффективным методом классифи-

кации. Он основан на близости объектов в признаковом пространстве. Хорошо работает в задачах с небольшим объемом данных, где важны локальные зависимости.

Случайные леса (RF) являются эффективным методом в решении задач классификации с большим количеством признаков, он позволяет хорошо обрабатывать данные среднего объема, а также устойчив к переобучению.

Метод опорных векторов (SVM), реализованный через библиотеку LibSVM, является мощным инструментом для решения задач классификации и регрессии. Он эффективен в задачах с относительно небольшим объемом данных, когда важна точность классификации. В работе используется *линейное ядро* с формулировками *C* и *Nu*, применяющимися для больших объемов данных, но могущими быть менее точными. В формулировке *C* штраф за ошибки при обучении модели определяется параметром стоимости *C*, где его увеличение ужесточает границу разделения классов. В случае формулировки *Nu* параметр стоимости *Nu* определяет долю обучающих объектов, допускающихся как опорные векторы и ошибки. Это предоставляет более гибкий контроль над ошибками и позволяет балансировать чувствительность к ошибкам и обобщение модели. Таким образом, различие заключается в том, как определяется параметр стоимости и какие значения он принимает для регулирования штрафа за ошибки и опорных векторов.

В качестве объектов для обучения используются эталонные полигоны типов ландшафтов с указанием соответствующего класса, автоматически преобразующиеся в группу точек при выполнении алгоритма машинного обучения.

РЕЗУЛЬТАТЫ

Опорными районами Северного Приохотья выбраны: центральная часть гор Дел-Урэкчэн, южная часть Ольского плато и его отрогов, центральная часть Туманского хребта.

Алгоритм работы в ГИС. В начале работы в ГИС создан векторный слой и выделены полигонами эталонные объекты для последующего обучения моделей, которые будут использованы при классификации изображений. В качестве классов объектов выбраны следующие: 1) реки и озёра; 2) дороги; растительный покров: 3) луговая растительность, 4) тундровая растительность; 5) деревья (главным образом лиственница), 6) кустарники (главным образом стланик); открытые от растительности склоны: 7) пологие, 8) средней крутизны, 9) крутые, 10) активные склоны. Всего выделено около 643 объектов, в среднем по 50 ед. на каждый класс. Под «активными склонами» автор работы понимает динамически активные склоны, поверхность которых намного бы-

стрее смещается вниз к тальвегу по сравнению с обычным десерпционным склоном. В результате такого смещения обломки пород постоянно переворачиваются и не приобретают пустынный загар, в связи с чем и выделяются светлым фотоном в видимом диапазоне спектра на космических снимках.

Для проведения машинного обучения было получено композитное многоканальное изображение, содержащее все необходимые источники, а именно: спектрозональные снимки; композитные изображения NDVI и EVI; ЦМР высоты

и углов наклона; реки, рассчитанные из ЦМР; дорожная сеть. Расчет растров с вегетационными характеристиками осуществлен по формулам, приведенным в статье И. Квона и коллег (2022).

Обучение моделей осуществлялось по композитному многоканальному изображению следующими методами – KNN, RF, cSVM и nuSVM; пропорция данных для обучения и проверки модели составила 0.7.

Далее проведены классификации изображений с использованием полученных моделей, пример которых показан на рис. 2.

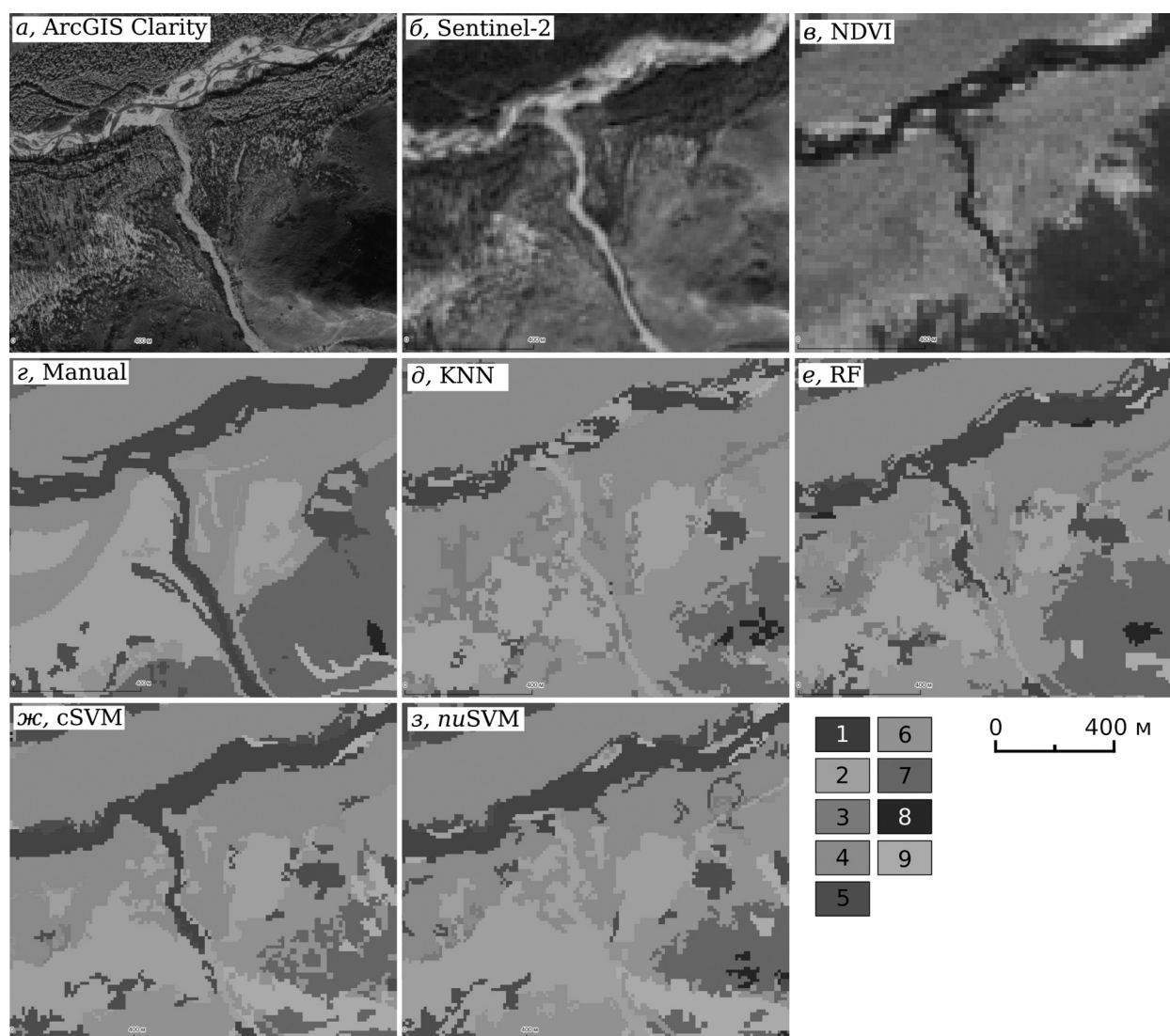


Рис. 2. Пример исходных данных и результаты классификации локального участка. Космическое изображение местности с сервиса ArcGIS Clarity (а) и со спутникового аппарата Sentinel-2 (б), композитное изображение нормализованной разностного индекса вегетации – NDVI (в), ручная классификация (г), метод k-ближайших соседей (д), случайные леса (е), метод опорных векторов с формулировками С (ж) и Nu (з). Условные обозначения: 1 – реки; растительность: 2 – тундровая, 3 – луговая, 4 – древесная, 5 – кустарниковая; склоны: 6 – пологие, 7 – средней крутизны, 8 – крутые, 9 – активные.

Fig. 2. Example of the original data and results of local site classification. Satellite imagery taken from ArcGIS Clarity (a) and the Sentinel-2 satellite (б), composite image of the normalized difference vegetation index (NDVI) (в), manual classification (г), method of k-nearest neighbors (д), random forests method (е), method of support vector with C formulation (ж), and with Nu formulation (з). Legend: 1 – rivers. Vegetation: 2 – tundra, 3 – meadowy, 4 – woody, 5 – shrubby. Slopes: 6 – gentle, 7 – moderate, 8 – steep, 9 – active.

ОБСУЖДЕНИЕ

В данном разделе приводится сравнение результатов, полученных по разным классификационным моделям и ручным методом, оценка их эффективности, а также показан пример дальнейшего использования полученных схем.

Анализ эффективности классификаций заключается в том, чтобы выявить, насколько успешно методы справились с задачей разделения различных классов местности на спектральных изображениях, а именно оценить такие показатели, как точность классификации, полнота и F-мера.

Точность (Precision) – характеристика, определяющая долю объектов, отнесенных моделью к выбранному классу и на самом деле принадлежащих к нему. Рассчитывается по формуле:

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives},$$

где *True Positives* – количество объектов, правильно классифицированных как выбранный класс, и *False Positives* – количество объектов, ошибочно классифицированных (не принадлежащих к выбранному классу).

Полнота (Recall) – определяет, какая доля объектов из искомого класса была правильно обнаружена моделью. Формула:

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives},$$

где *True Positives* – количество правильно обнаруженных объектов класса и *False Negatives* – количество объектов из класса, которые модель не обнаружила.

F-мера (F-score) – гармоническое среднее между точностью и полнотой, используемое для оценки общей производительности классификационной модели. Формула для F-меры:

$$F-score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}.$$

Рассчитанные представленные показатели по разным моделям классификации отображены в таблице 1.

Стоит отметить, что результаты классификации моделей в большинстве случаев оптимистичны, поскольку эталонные классы основывались на объектах, выделенных вручную. С другой стороны, имеются и неточности в первичной модели, которые обусловлены в первую очередь малым количеством хорошо классифицированных объектов (в среднем 40 эталонов на класс), а во вторую – пропорцией между обучающей и проверочной выборками (которые, в свою очередь, берутся в случайном порядке). Все эти факторы являются актуальными для построения моделей.

С другой стороны, подтвердить правильность моделей классификации можно в сравнении с реальными данными или, как принято в работах по машинному обучению, при проведении сравнительного анализа между моделями. Ручная классификация ландшафтов в большей степени отражает те обстановки, которые в действительности представлены на изучаемой территории, но данный метод хотя и является точным, однако подвержен субъективной оценке тех классов, которые выделяет исследователь. Например, можно ли точно определить границу между пологим склоном ($> 5^\circ$ крутизны поверхности), покрытым растительностью, и долиной ($< 5^\circ$) на космических снимках? В связи с этим не следует принимать ручную классификацию как эталон – нужно воспринимать ее как одну из допустимых (более точную, но и трудозатратную).

Для сравнения результатов классификаций ML-методами была выполнена ручная обработка космического снимка для локальной части исследуемой территории (рис. 2, а). Полученные данные преобразованы в растровый формат в таком же разрешении, что и остальные результаты классификаций. Затем проводился подсчет занимаемой площади каждого класса в каждой модели. Результаты приводились к 100 % и показаны в таблице 2.

Таблица 1. Оценка точности (P), полноты (R) и F-меры (F) моделей первичной классификации

Table 1. Assessment of Precision (P), Recall (R), and F-score (F) of the primary classification models

Метод		KNN			RF			cSVN			nuSVN		
Первичный класс (номер)		P	R	F	P	R	F	P	R	F	P	R	F
Реки (2)		1.00	0.29	0.45	0.90	0.88	0.89	0.91	0.97	0.94	0.91	0.97	0.94
Снег (3)		0.96	0.85	0.90	0.96	0.85	0.90	1.00	1.00	1.00	1.00	0.85	0.92
Дороги (5)		1.00	0.85	0.92	1.00	0.85	0.90	1.00	0.82	0.90	1.00	0.88	0.93
Растительность	тундровая (15)	0.5	0.23	0.32	0.91	0.91	0.91	1.00	0.94	0.97	0.91	0.94	0.92
	луговая (10)	0.53	1.00	0.69	1.00	0.94	0.96	0.93	0.91	0.92	1.00	0.97	0.98
	кустарниковая (12)	0.92	0.70	0.80	0.94	0.91	0.92	0.77	0.82	0.80	0.86	0.91	0.88
	древесная (11)	0.80	0.97	0.88	0.86	0.94	0.90	0.80	0.94	0.86	0.91	0.91	0.91
Склоны	пологие (30)	0.38	0.85	0.53	0.92	0.73	0.67	0.80	0.85	0.82	0.88	0.94	0.91
	средние (31)	0.51	0.52	0.52	0.56	0.82	0.66	0.77	0.79	0.78	0.75	0.61	0.67
	крутые (36)	0.64	0.64	0.64	0.79	0.67	0.73	0.81	0.76	0.78	0.68	0.91	0.78
	активные (35)	0.64	0.26	0.37	0.76	0.58	0.66	0.81	0.76	0.78	0.87	0.82	0.84

Обобщая полученные данные по разным методам классификации, мы наблюдаем схожую картину между результатами. Лучше всего, по субъективной оценке автора, справился метод опорных векторов, а именно cSVN. У него чуть хуже происходит классификация склонов, но хорошо диагностируются водотоки. Стоит также учитывать, что первичные модели всегда имеют ошибки и что некоторые категории необходимо вторично пересчитать, основываясь на других алгоритмах. Всё это приводит к применению реклассификации получившегося изображения.

Таблица 2. Площадное распределение классов на локальном участке местности, полученные разными ML-методами

Table 2. Area distribution of classes in a local area of the terrain obtained by different ML methods

Класс		Ручная	KNN	RF	cSVM	nuSVM
Растительность	реки (2)	10.85	4.25	8.61	9.71	8.13
	луговая (10)	0.00	7.10	4.87	0.27	0.85
	древесная (11)	33.62	38.83	31.83	36.91	33.42
	кустарниковая (12)	6.54	3.07	5.55	6.39	7.54
	тундровая (15)	20.92	19.04	15.31	18.08	24.69
Склоны	пологие (30)	6.73	19.29	20.41	16.56	12.72
	средние (31)	19.68	7.52	11.83	6.60	6.97
	активные (35)	1.35	0.30	0.94	5.48	5.32
	крутые (36)	0.31	0.60	0.65	0.00	0.36

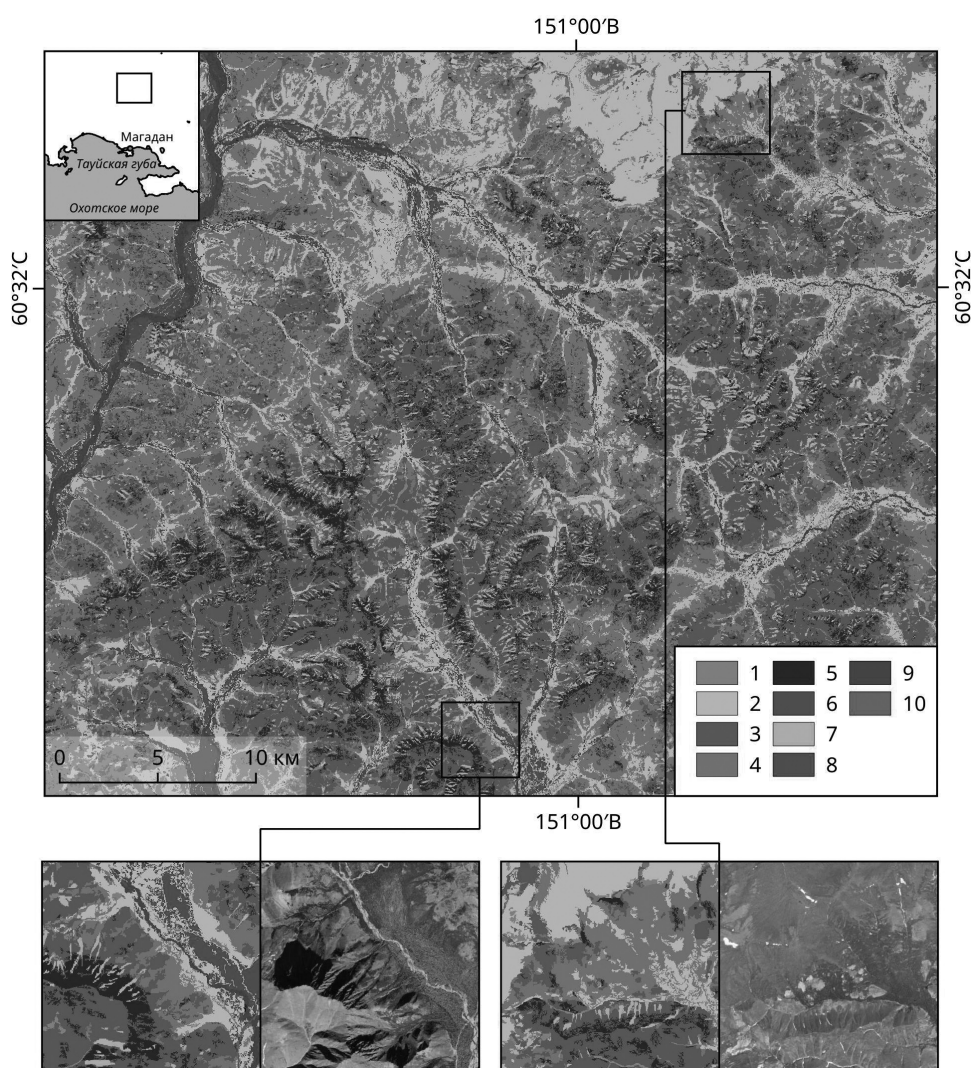


Рис. 3. Схема типов склонов района гор Дел-Урэкчэн и Ольского плато. Условные обозначения: пологие склоны открытые (1) и задернованные (2), склоны средней крутизны открытые (3) и задернованные (4), крутые склоны открытые (5) и задернованные (6), активные склоны (7), долины (8), реки и озёра (9), дороги (10). На верхней врезке – положение района исследования в регионе, на нижних врезках – крупные части схемы и космоснимки местности (сервис Arcgis Imagery).

Fig. 3. Diagram of slope types in the Del-Urekchen Mountains and the Ola Plateau area. Legend: gentle slopes, open (1) and grassy (2); moderate slopes, open (3) and grassy (4); steep slopes, open (5) and grassy (6); active slopes (7); valleys (8); rivers and lakes (9); roads (10). Upper inset: research area position in the region; lower insets: large parts of the scheme and satellite images of the area (ArcGIS Imagery service).

Практическое применение. В представленной работе принципы реклассификации основаны на крутизне склонов и наличии растительного покрова, первое и второе достаточно точно можно получить из данных ЦМР и композитного изображения NDVI. Что касается выделения класса активных склонов, то был применен подход, заключающийся в первую очередь в диагностировании их по светловому фототону в видимом диапазоне спектра и крутизне наклона поверхности, а во вторую – в отличии активных склонов от некоторых метасоматических пород, проявляющемся в инфракрасном спектре (SWIR).

Полученная схема поверхности подверглась реклассификации на следующие типы по правилам: *долины* – все задернованные типы поверхности, а также крутизна меньше 5° и абсолютная высота меньше 900 м (для района гор Дел-Урэкчэн) и 400 м (для района Туманского

хребта); *пологие склоны задернованные* – все задернованные поверхности с крутизной от 5 до 15° ; *средней крутизны склоны задернованные* – то же, с крутизной от 15 до 35° ; *крутые задернованные склоны* – то же, с крутизной больше 35° ; *открытые от растительности склоны* подразделялись по крутизне на пологие, средней крутизны и крутые согласно углу наклона, также в эти классы включены теневые участки; *активные склоны* – ранее классифицированные активные склоны с крутизной более 15° , а также включающие площадь дешифрированных осыпей, в противном случае присваивается класс пологих склонов; перенос без изменений классов *рек и дорог*.

В ходе применения алгоритмов машинного обучения по первичной классификации многоканальных изображений и их реклассификации были получены схемы типов склонов, отображенные на рис. 3 и 4.

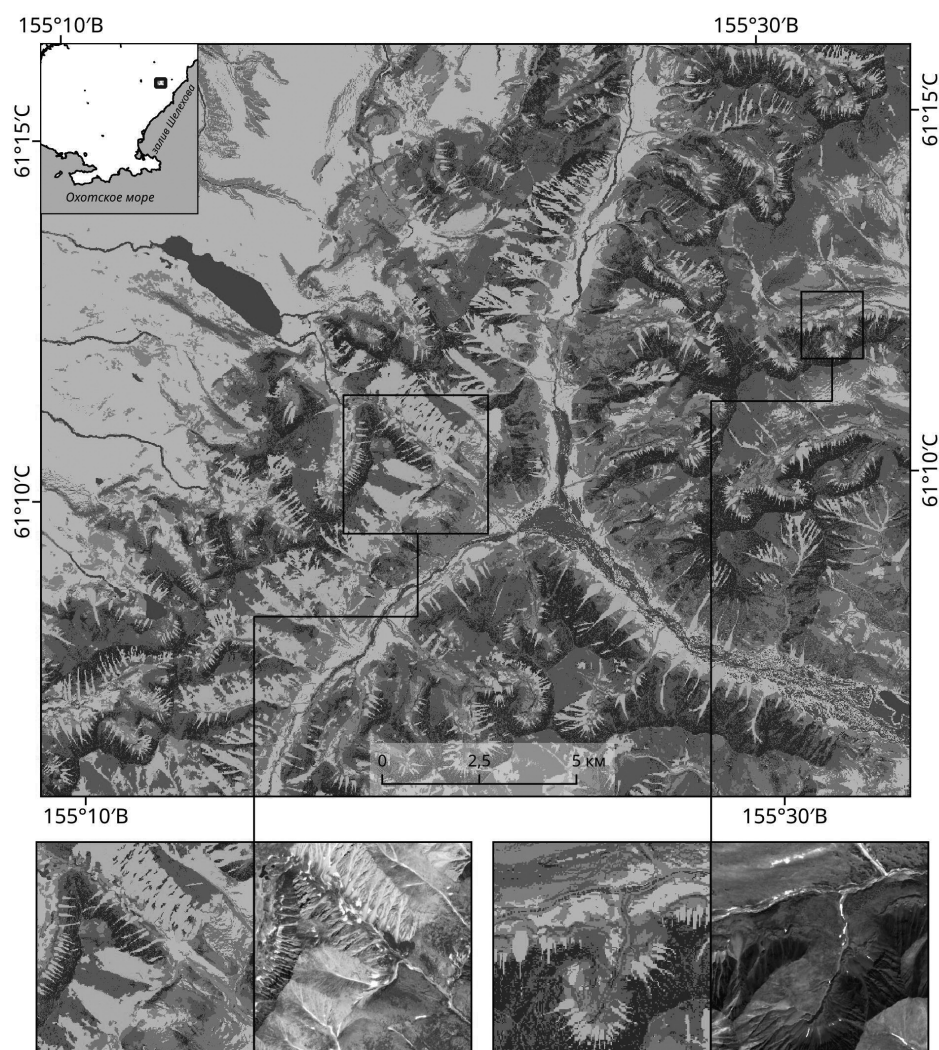


Рис. 4. Схема типов склонов района Туманского хребта. Условные обозначения см. на рис. 3, космоснимки сервисов Arcgis Imagery и Bing.

Fig. 4. Diagram of slope types in the Tumansky Ridge area. For Legend, see Fig. 3; satellite images from the ArcGIS Imagery and Bing services.

Таблица 3. Распределение типов склонов по району исследования

Table 3. Distribution of slope types in research areas

Тип склонов Район	Реки и озёра	Долины	Пологие ($< 15^\circ$)		Средней крутизны		Крутые ($> 35^\circ$)		Активные
			задерн.	откр.	задерн.	откр.	задерн.	откр.	
горы Дел-Урэкчэн	0.7	4.2	13.1	1.9	30.6	36	2.1	7.5	3.9
Ольское плато	1.6	3.8	38.2	1.7	33.8	15.9	1.1	2.5	1.4
Туманский хребет	1.6	1.1	27.6	8.8	11.0	28.2	1.7	10.6	9.4

Дальнейшие операции включали в себя расчет занимаемой площади (см. табл. 3) и экспозиции различных типов склонов. Последняя рассчитывалась по сети точек (100×100 м) с извлечением данных из ЦМР и схемы склонов.

Площадное распространение разных типов склонов приведено в табл. 3.

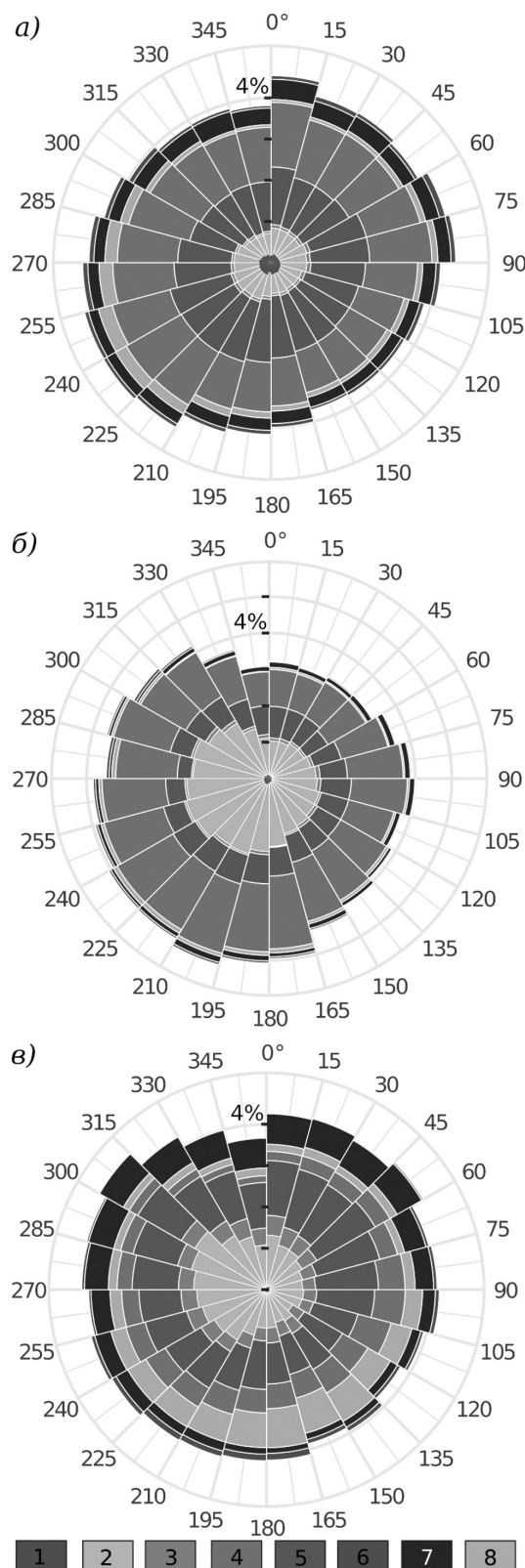
В горах Дел-Урэкчэн можно отметить, что половина района исследования приходится на область, покрытую растительностью; пространственно эта область, главным образом, затрагивает левобережье р. Армань, где распространены склоны с высотами меньше 1000 м. Долю свободной от растительности поверхности, на которой протекают склоновые процессы, можно оценить в 47.5 %, при этом на активные склоны приходится чуть меньше 4 % территории района.

В свою очередь, на территории Ольского плато и его отрогов доминируют покрытые растительностью склоны, занимающие около 77 % территории. Более 40 % поверхности приходится на пологие склоны, которые и формируют плоскогорье плато, а на отроги без растительности, на которых протекают склоновые процессы, приходится всего 18.4 %, при этом доля активных склонов составляет 1.4 %.

В Туманском хребте наблюдается другая картина: поверхность склонов с произрастающей растительностью составляет всего 41.4 %, тогда как к свободным от почвенного слоя относится 57 %, а активные склоны составляют 9.4 % от общей поверхности территории. Пространственно растительность приурочена, главным образом,

Рис. 5. Экспозиция разных типов склонов по районам исследования: а) горы Дел-Урэкчэн, б) Ольское плато и его отроги и в) Туманский хребет. Условные обозначения: долины (1); пологие склоны задернованные (2) и открытые (3); склоны средней крутизны задернованные (4) и открытые (5); крутые склоны задернованные (6) и открытые (7); активные склоны (8).

Fig. 5. Exposure of different slope types in the research areas: а) Del-Urekchen Mountains, б) Ola Plateau and its offshoots, and в) Tumansky Ridge. Legend: valleys (1); gentle slopes, grassy (2) and open (3); moderate slopes, grassy (4) and open (5); steep slopes, grassy (6) and open (7); active slopes (8).



к Верхне-Купкинской впадине, где она произрастает на пологих склонах.

Как видим, анализ процентного соотношения разных типов склонов позволяет сделать выводы о том, что в Туманском хребте, по сравнению с двумя другими регионами, наибольшую долю площади занимают открытые склоны и особенно активные склоны, что свидетельствует о высокой морфодинамической активности в этом регионе, по-видимому обусловленной новейшими движениями вдоль Ланково-Омолонской зоны разломов.

Экспозиция разных типов склонов приведена на рис. 5.

Анализируя экспозиции склонов, покрытых растительностью, можно выделить следующие моменты. В горах Дел-Урэкчэн растительность развита по всем направлениям, с доминированием в юго-западном и северо-восточном простираниях, что вызвано, в свою очередь, северо-западным простиранием осей водоразделов. Ольское плато с отрогами обладает асимметричным распределением, с доминированием в юго-западном направлении, проявление растительности на крутых склонах редкое. Туманский хребет обладает также асимметрией, только в южном направлении. К этому направлению приурочена и растительность на крутых склонах, что связано с климатическим фактором (Охотский склон) и большей долей солнечной активности.

Свободные от растительности поверхности обладают следующими выявленными закономерностями. Пологие склоны имеют равнозначное распространение по сторонам света во всех трех районах. Склоны средней крутизны также развиты равномерно, но имеются некоторые доминирующие направления. Так, на Ольском плато и его отрогах менее распространены склоны с западной экспозицией, а в Туманском хребте доминируют северо-восточные направления. Крутые склоны в трех районах развиты главным образом в северных направлениях, что, видимо, приурочено к ледниковому рельефу и обилию каров и цирков этих экспозиций.

Активные склоны имеют другую ориентировку относительно остальных типов. Так, в горах Дел-Урэкчэн они, главным образом, распространены в азимутах 225–285°. На Ольском плато и его отрогах основная часть приурочена к юго-западным экспозициям, а в Туманском хребте – к азимутам 150–210°. Исходя из представленных роз диаграмм можно заметить, что наиболее активные склоны приурочены к юго-западным экспозициям рельефа. Это может быть связано с тем, что представленные направления получают больше солнечного излучения, что способствует быстрому таянию снега и льда, а также интенсивности физического и химического выветривания, что может приводить к более интенсивному

перемещению обломочного материала и образованию активных склонов.

ВЫВОДЫ

Применение методов машинного обучения при анализе спектрональных снимков способствует увеличению точности и автоматизации классификации местности, что ускоряет процесс обработки спутниковых данных и позволяет исследователям быстрее анализировать морфологические особенности форм рельефа и/или ландшафтов. В данном исследовании были выявлены следующие особенности рельефа в некоторых районах Северного Приохотья:

Процент территории, подверженной активным склоновым процессам, различается в зависимости от региона: горы Дел-Урэкчэн – 4.0 %, Ольское плато и его отроги – 1.4 %, Туманский хребет – 9.4 %.

Экспозиция активных склонов: горы Дел-Урэкчэн – 240–270°, Ольское плато и его отроги – 165–285°, Туманский хребет – 150–195°.

Процент задернованных (стабильных) склонов составил: горы Дел-Урэкчэн – 45.8, Ольское плато и его отроги – 73.1, Туманский хребет – 40.3.

Лучшая модель классификации, полученная на основе ML-методов, для представленных районов основывается на методе опорных векторов (cSVN).

ЛИТЕРАТУРА

- Arctic DEM* – Polar geospatial center // University of Minnesota. 2018. URL: <https://www.pgc.umn.edu/data/arcticdem> (дата обращения 04.05.2020).
- Bzdok D., Altman N., Krzywinski M.* Statistics versus machine learning // *Nature Methods*. 2018. Vol. 15. P. 233–234. URL: <https://doi.org/10.1038/nmeth.4642> (дата обращения 30.01.2023).
- Copernicus* open access hub // European Space Agency. 2014. URL: <https://scihub.copernicus.eu/dhus> (дата обращения 30.01.2023).
- Geofabrik* download server. Karlsruhe, 2018. URL: <https://download.geofabrik.de/russia/far-eastern-fed-district.html> (дата обращения: 19.09.2023).
- Gizatullin A. T., Alekseenko N. A.* Prediction of wildfires based on the spatio-temporal variability of fire danger factors // *Geography, Environment, Sustainability*. 2022. Vol. 15, No. 2. P. 31–37. DOI: 10.24057/2071-9388-2021-139.
- Kwong I. H. Y., Wong F. K. K., Fung T., Liu E. K. Y., Lee R. H., Ng T. P. T.* A Multi-stage approach combining very high-resolution satellite image, GIS database and post-classification modification rules for habitat mapping in Hong Kong // *Remote Sensing*. 2022. Vol. 14. P. 131–143. DOI: 10.3390/rs14010067.
- OTB (Orfeo ToolBox)* – Open source processing of remote sensing images. 2012. URL: <https://www.orfeo-toolbox.org/> (дата обращения 30.01.2023).
- QGIS* – A free and open source geographic information system. 2010. URL: <https://www.qgis.org/ru/site/index.html> (дата обращения 23.12.2016).

Spatial resolution // Sentinel Online. 2012. URL: <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial> (дата обращения 30.01.2023).

Techniques for deep learning on satellite and aerial imagery // GitHub. 2023. URL: <https://github.com/satellite-image-deep-learning/techniques> (дата обращения 19.09.2023).

Voosen P. Europe builds 'digital twin' of Earth to hone climate forecasts // *Science*. 2020. Vol. 370, No. 6512. P. 16–17. DOI: 10.1126/science.370.6512.16.

Zhang X., Yu W., Pun M.-O., Shi W. Cross-domain landslide mapping from large-scale remote sensing im-

ages using prototype-guided domain-aware progressive representation learning // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2023. Vol. 197. P. 1–17. DOI: 10.1016/j.isprsjprs.2023.01.018.

Zhou C., Yin K., Cao Y., Ahmed B., Li Y., Catani F., Pourghasemi H. R. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China // *Computers & Geosciences*. 2018. Vol. 112. P. 23–37. DOI: 10.1016/j.cageo.2017.11.019.

Поступила в редакцию 27.09.2023.

Поступила после доработки 23.11.2023.

APPLICATION OF MACHINE LEARNING METHODS FOR SLOPE CLASSIFICATION BASED ON SATELLITE IMAGERY, ILLUSTRATED USING SELECTED AREAS IN NORTHERN PRIOKHOTYE

P. P. Kolegov

North-East Interdisciplinary Scientific Research Institute n. a. N. A. Shilo, FEB RAS, Magadan

The paper presents machine learning algorithms based on methods such as SVM, Random Forest, and k-Nearest Neighbors for classifying spectral-zone satellite images (Sentinel-2). Models generated using these methods and manual classification have been comparatively analyzed, parameters such as accuracy, recall, and F-measure presented. Practical applications include image reclassification the degree of slope openness from vegetation, and assessment of their morphometric characteristics. Data on the area distribution of dynamically active slopes and their exposure are obtained for the Del-Urekchen Mountains (4.0 %, 240–270°), the Ola Plateau and its foothills (1.4 %, 165–285°), and the Tumansky Ridge (9.4 %, 150–195°).

Keywords: classification, slope, machine learning, GIS, Northern Priokhotye.

REFERENCES

Arctic DEM – Polar Geospatial Center, 2018, *University of Minnesota*. URL: <https://www.pgc.umn.edu/data/arcticdem> (accessed 04.05.2020).

Bzdok, D., Altman, N., Krzywinski, M., 2018. Statistics Versus Machine Learning, *Nature Methods*. 15, 233–234. URL: <https://doi.org/10.1038/nmeth.4642> (accessed 30.01.2023).

Copernicus Open Access Hub, 2014, *European Space Agency*. URL: <https://scihub.copernicus.eu/dhus> (accessed 30.01.2023).

Geofabrik Download Server, 2018. Karlsruhe. URL: <https://download.geofabrik.de/russia/far-eastern-fed-district.html> (accessed 19.09.2023).

Gizatullin, A. T., Alekseenko, N. A., 2022. Prediction of Wildfires Based on the Spatio-Temporal Variability of Fire Danger Factors, *Geography, Environment, Sustainability*. 15 (2), 31–37. DOI: 10.24057/2071-9388-2021-139.

Kwong, I. H. Y., Wong, F. K. K., Fung, T., Liu, E. K. Y., Lee, R. H., Ng, T. P. T., 2022. A Multi-Stage Approach Combining Very High-Resolution Satellite Image, GIS Database and Post-Classification Modification Rules for Habitat Mapping in Hong Kong, *Remote Sensing*. 14, 131–143. DOI: 10.3390/rs14010067.

OTB (Orfeo ToolBox) – Open Source Processing of Remote Sensing Images, 2012. URL: <https://www.orfeo-toolbox.org/> (accessed 30.01.2023).

QGIS – A Free and Open Source Geographic Information System, 2010. URL: <https://www.qgis.org/ru/site/index.html> (accessed 23.12.2016).

Spatial Resolution, 2012, *Sentinel Online*. URL: <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial> (accessed 30.01.2023).

Techniques for Deep Learning on Satellite and Aerial Imagery, 2023, *GitHub*. URL: <https://github.com/satellite-image-deep-learning/techniques> (accessed 19.09.2023).

Voosen, P., 2020. Europe Builds 'Digital Twin' of Earth to Hone Climate Forecasts, *Science*. 370 (6512), 16–17. DOI: 10.1126/science.370.6512.16.

Zhang, X., Yu, W., Pun, M.-O., Shi, W., 2023. Cross-Domain Landslide Mapping from Large-Scale Remote Sensing Images Using Prototype-Guided Domain-Aware Progressive Representation Learning, *ISPRS Journal of Photogrammetry and Remote Sensing*. 197, 1–17. DOI: 10.1016/j.isprsjprs.2023.01.018.

Zhou, C., Yin, K., Cao, Y., Ahmed, B., Li, Y., Catani, F., Pourghasemi, H. R., 2018. Landslide Susceptibility Modeling Applying Machine Learning Methods: A Case Study From Longju in the Three Gorges Reservoir Area, China, *Computers & Geosciences*. 112, 23–37. DOI: 10.1016/j.cageo.2017.11.019.