АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР
ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ
INFORMATION AND COMMUNICATION TECHNOLOGIES

АҚПАРАТТЫҚ ЖҮЙЕЛЕР
ИНФОРМАЦИОННЫЕ СИСТЕМЫ
INFORMATION SYSTEMS

DOI 10.51885/1561-4212_2024_3_177
IRSTI 20.23.27

M.E. Rakhymberdina[1], G.K. Daumova[1], G.K. Zhussupova[2], R.K. Chettykbayev[1], D.V. Chepashev[2]
[1]D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan
E-mail: MRahymberdina@edu.ektu.kz
*E-mai mail: GDaumova@edu.ektu.kz**
E-mail: rchettykbayev@edu.ektu.kz
[2]Institute of Ionosphere LLP, Almaty, Kazakhstan
E-mail: zhussupova51@gmail.com
E-mail: d.chepashev@ionos.kz

MAPPING VEGETATION TYPES ON DIFFERENT SLOPES AND ASSESSING THE DYNAMICS OF THEIR CHANGE OVER A LONG PERIOD OF TIME

ӘРТҮРЛІ БЕТКЕЙЛЕРДЕГІ ӨСІМДІК ТҮРЛЕРІН КАРТОГРАФИЯЛАУ ЖӘНЕ ОЛАРДЫҢ ҰЗАҚ УАҚЫТ АРАЛЫҒЫНДАҒЫ ӨЗГЕРУ ДИНАМИКАСЫН БАҒАЛАУ

КАРТИРОВАНИЕ ТИПОВ РАСТИТЕЛЬНОСТИ НА РАЗЛИЧНЫХ СКЛОНАХ И ОЦЕНКА ДИНАМИКИ ИХ ИЗМЕНЕНИЯ ЗА ПРОДОЛЖИТЕЛЬНЫЙ ПЕРИОД ВРЕМЕНИ

Abstract. *Vegetation mapping is a key task in remote sensing for environmental protection. Modern remote sensing technologies offer numerous advantages, including significant time savings, the ability to cover large areas, and effective long-term monitoring. These methods can greatly accelerate tasks compared to traditional field-based approaches. This article presents a method for land cover classification using time series analysis of Planet Scope satellite imagery from 2019 to 2023. The aim of the study is to examine how different input data affect classification outcomes and to analyze vegetation dynamics over the selected period. The research was conducted in the northeastern part of East Kazakhstan using the Random Forest (RF) algorithm to determine optimal parameters and track changes in vegetation types. The results include land cover classifications based on five different feature combinations, along with an overall accuracy assessment. The findings indicate that the best results in land cover mapping are achieved by combining spectral bands, topographic indices, and tree crown height layers. Vegetation was classified into three categories: trees, shrubs, and grass cover, with water bodies and bare soil areas also identified. The analysis revealed that from 2019 to 2023, the total area of trees and water bodies decreased, while grass cover, shrublands, and bare areas expanded. These findings can aid in further ecosystem analysis and highlight significant changes in vegetation structure and ecosystem processes.*
Keywords: *Vegetation types, machine learning, Random Forest, remote sensing, change monitoring*

Аңдатпа. *Өсімдік түрлерін картографиялау қоршаған ортаны қорғау саласындағы қашықтықтан зондтаудың маңызды міндеттерінің бірі болып табылады. Заманауи қашықтықтан зондтау технологияларын қолданудың көптеген артықшылықтары бар, оның ішінде уақытты айтарлықтай үнемдеу, үлкен аумақтарды қамту, ұзақ мерзімді тиімді бақылауды қамтамасыз ету. Бұл әдістер дәстүрлі жерүсті әдістермен салыстырғанда тапсырмаларды орындауды айтарлықтай жылдамдата алады.Мақалада 2019-2023 жылдар аралығындағы Planet Scope* *спутниктік суреттерін талдау негізінде өсімдік жамылғысын жіктеу әдісі талқыланады. Бұл зерттеудің мақсаты әртүрлі кіріс деректерінің жіктелу нәтижелеріне және таңдалған кезеңдегі өсімдіктердің өзгеру динамикасына әсерін талдау болып табылады. Шығыс Қазақстанның солтүстік-шығыс бөлігінде «Random Forest» (RF) алгоритмі бойынша ең қолайлы параметрлерді таңдау және өсімдік түрлерінің өзгеруін талдау үшін зерттеу жүргізілді. Топтастыру белгілерінің бес түрлі комбинациясын және жалпы дәлдікті бағалауды пайдалана отырып, жер жамылғысын жіктеу нәтижелері берілген. Зерттеу нәтижелері бойынша спектрлік арналар, топографиялық көрсеткіштер және ағаш қалқасының биіктік қабаты сияқты кіріс параметрлердің комбинациясын пайдалану жер жамылғысының картасын жасауда ең жақсы нәтиже беретіні анықталды. Өсімдік жамылғысы ағаштар, бұталар және шөп жамылғысы болып үш түрге жіктеліп, су ресурстары мен ашық топырақтың аумақтары анықталды. Талдау көрсеткендей, 2019-2023 жылдар аралығында ағаштар мен су қоймаларының жалпы көлемі азайып, ал шөп жамылғысы, бұталар және ашық топырақ алқаптары көбейіп келеді. Осылайша, зерттеу нәтижелері экожүйенің өзгерістерін одан әрі талдау мен жер жамылғысы құрылымындағы және экожүйелік процестердегі маңызды өзгерістерді анықтауға ықпал ете алады.*

Түйін сөздер: *Өсімдік түрлері, машиналық оқыту, Random Forest, қашықтықтан зондтау, өзгерістердің мониторингі.*

Аннотация. *Картографирование растительности является одной из наиболее важных задач дистанционного зондирования в области защиты окружающей среды. Применение современных технологий дистанционного зондирования имеет множество преимуществ, включая существенную экономию времени, охват обширных территорий и обеспечение эффективного длительного наблюдения. Эти методы позволяют значительно ускорить выполнение задач по сравнению с традиционными наземными способами. В статье рассматриваеся метод классификации растительного покрова на основе анализа спутниковых снимков Planet Scope временных рядов за период с 2019 по 2023 год. Цель данного исследования состоит в анализе влияния различных входных данных на результаты классификации и динамики изменений растительности за выбранный период. Было проведено исследование в северо-восточной части Восточного Казахстана с использованием алгоритма Random Forest (RF) для выбора наиболее подходящих параметров и анализа изменений типов растительности. Представлены результаты классификации растительного покрова с использованием пяти различных комбинаций классификационных признаков и оценки общей точности. Согласно результатам исследования, было установлено, что использование комбинации таких входных параметров, как спектральные каналы, топографические индексы и слой высоты крон деревьев, дают наилучшие результаты в картировании растительного покрова. Растительный покров был классифицирован на три типа: деревья, кустарники и травяной покров, а также были определены площади водных ресурсов и открытой почвы. Анализ показал, что в период с 2019 по 2023 годы общая площадь деревьев и водоемов уменьшается, в то время как площади травяного покрова, кустарников и не покрытых растительностью участков увеличиваются. Таким образом, результаты исследования могут способствовать дальнейшему анализу изменений экосистемы и выявлению важных изменений в структуре растительного покрова в экосистемных процессах.*

Ключевые слова: *Типы растительности, машинное обучение, Random Forest, дистанционное зондирование, мониторинг изменений.*

Introduction. The classification and mapping of vegetation using remote sensing are regarded as effective methods for collecting land cover data at various spatial scales (Feddema, 2005). These methods are crucial for understanding the impacts of land cover changes on factors such as agricultural production, avalanche formation, carbon sequestration, water quality, runoff, and biodiversity conservation.

Machine learning has attracted increasing attention in recent years, and its use in land cover classification has been expanding (Sun, 2023). The Random Forest machine learning algorithm proposed by Leo Breiman and Adele Cutler combines two key concepts: Breiman's bagging method and the random subspace method proposed by Tin Kam Ho. With this flexibility, Random Forest can effectively solve a wide range of machine learning problems (Breiman, 2001).

The objective of this study was to select the optimal combination of input variables for highprecision mapping to determine the vegetation dynamics on different slopes. A dataset of topographic (Elevation, Slope, Aspect), spectral indices (NDVI, NDWI, EVI, MSAVI) and canopy height were used as RF input parameters to classify vegetation at a local scale on the GEE platform with Planet Scope satellite images in the selected area. Then, the obtained optimal RF parameters were applied to classify the vegetation type and analyze their dynamics during the period 2019-2023.

Literature review. Machine learning is particularly used in land cover classification (Georganos, 2018), remote sensing image classification (Sheykhmousa, 2020) and soil property mapping (Hengl, 2015). They provide high classification accuracy for different vegetation types, resistance to overfitting and the ability to handle large amounts of data. In a number of previous studies, the Random Forest algorithm has demonstrated better results than other machine learning algorithms. For example, Dino Dobrinić achieved 92% accuracy in his study using Sentinel imagery for a hilly area of Northern Croatia (Dobrinić, 2021). The results of the Thanh Noi Phan study show that all datasets provided medium to high accuracy land cover maps with an overall accuracy of over 84.31% (Phan, 2020). Also, Xueliang Zeng obtained the best results in soil identification using the RF algorithm in his research conducted on the Google Earth Engine (GEE) platform (Zeng, 2024). Samuel Edwin Pizarro created a highly accurate model by combining spectral bands and topographic indices with the random forest algorithm, with the Kappa coefficient being 0.81 (Pizarro, 2022).

Materials and methods. The research methodology includes the following 6 steps: 1) identification and processing of Planet Scope satellite data, 2) feature extraction and dataset preparation, 3) selection of training and testing samples, 4) vegetation classification and accuracy assessment in the selected area, 5) identification of the optimal dataset for vegetation classification, 6) analysis of RF classification accuracy and change dynamics over the selected period. The workflow is shown in Figure 1. The details are described in the following sections.

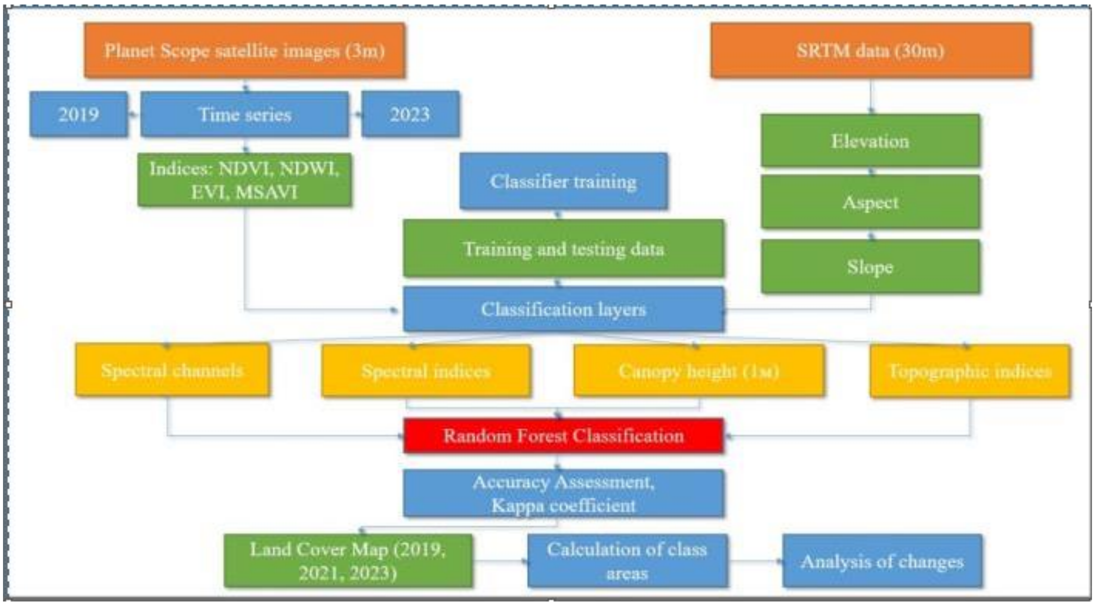

Figure 1. Research method graph

Note – compiled by the authors

The area of research. The research territory is situated in the northeastern part of the East Kazakhstan region (Fig. 2), covering over 2955 km² and featuring elevations ranging from 768 to 4448 meters. The geographic coordinates of the area are 85°30' - 87°00' east longitude and 49°10' - 49°50' north latitude. The majority of the study area lies in the southeastern part of the Listvyaga Ridge, extending to the highest peak, Belukha, at the border with Russia, and reaches the northern section of the Bukhtarma River in the south. The Belukha glaciers serve as the source of the Belaya and Chernaya Berel rivers, and the area also includes lakes such as Rakhmanovskoye, Yazovoye, Chernovoye, and Maralye.

Mountain forests are typically found at elevations ranging from 1,200 to 2,200 meters above sea level. Coniferous trees include evergreen cedar, spruce and fir, as well as larch; deciduous trees include white-trunked birch, poplar, aspen, and many different types of willow. Shrub flora numbers over 50 species (Bel'gibaev, 2007).

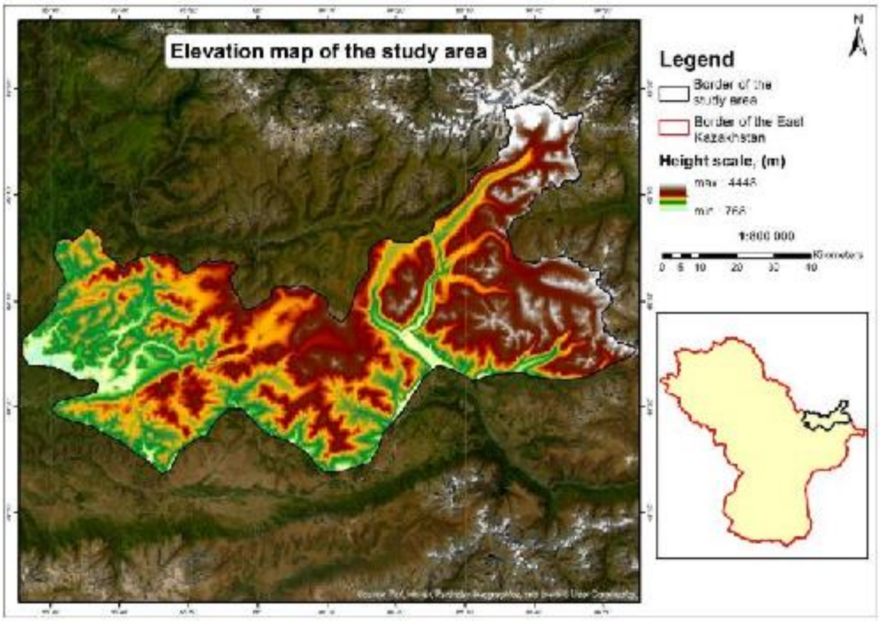

Figure 2. Map of the terrain of the research territory

Note – compiled by the authors

The data used in this study were divided into two groups: Planet Scope satellite data and ancillary data. Brief information on each data group is provided below.

Satellite datasets. The land cover maps were created using commercial data from the PlanetScope satellite constellation, which consists of three groups of satellites: Dove-C, Dove-R , and SuperDove. The PlanetScope satellites provide images with a resolution of 3.0 to 4.2 meters, which capture images in at least four spectral bands: Blue, Green, Red, and NIR.

In addition to providing individual images, the Planet Labs portal allows you to receive composite images. These are stitched images that combine the best parts of several images, removing clouds for analysis. The results were images dated July 13, 19, 2023, which were made up of 20 individual scenes combined into one composite image. Similarly, the composite image for July 3, 2021, was made up of 33 scenes, for 2019, images were taken on July 6, 17, 29 and August 1, 12, and consisted of 63 scenes.

PlanetScope images require additional processing in the form of image normalization, due to the difference between the band ranges of the Dove-C, Dove-R and SuperDove satellites. Therefore, the Planet Labs portal provides the ability to harmonize images with Sentinel-2 images: blue - band 1, green - band 3, red - band 4, NIR - band 8a. Thus, we received harmonized Planet Scope images with time intervals for the months of July-August 2019, 2021, 2023 (URL: https://earth.esa.int/eogateway/missions/planetscope).

Topographic data. As topographic variables for land cover classification from the Shuttle Radar Topography Mission (SRTM) 30 m resolution data available from the US Geological Survey (USGS) website (URL: https://earthexplorer.usgs.gov/), we determined the slope (in degrees from 0 to 90°) and aspect (in degrees from 0 to 360°) from the Elevation layer (Figure 1) and transformed the aspect using trigonometric sinusoidal values to avoid cyclical data values. The obtained sinusoidal aspect values reflect the degree of tilt toward the east, ranging from 1 (oriented to the east) to -1 (oriented to the west).

Spectral Indices. Vegetation indices like MSAVI, EVI, NDVI, and NDWI are employed to differentiate between areas with and without vegetation, assess vegetation health while accounting for biomass and atmospheric conditions, and enhance the detection of water bodies (Huete, 2002; Gilang, 2021; McFeeters, 1996).

Using the following equations, spectral indices were calculated – NDVI, NDWI, EVI, MSAVI.

$$NDVI = \frac{(\mathrm{NIR+Red})}{(\mathrm{NIR-Red})} \tag{1}$$

$$NDWI = \frac{(\mathrm{Green-NIR})}{(\mathrm{Green+NIR})} \tag{2}$$

$$MSAVI = \frac{2*\mathrm{NIR}+1-\sqrt{(2*\mathrm{NIR}+1)^2-8(\mathrm{NIR-Red})}}{2} \tag{3}$$

$$EVI = 2.5 * \frac{(\mathrm{NIR-Red})}{(\mathrm{NIR}+6*\mathrm{Red}-7.5*\mathrm{Blue}+1)} \tag{4}$$

where: NIR – infrared range; Red – red range; Green – green range; Blue – blue range.

Canopy Height. The Global Canopy Height Maps dataset provides detailed information on tree canopy heights worldwide at 1 m resolution, covering the period from 2009 to 2020. The majority of the data (80%) is based on imagery taken between 2018 and 2020, providing a current and accurate representation of tree canopy height and distribution over this period. This layer was used to distinguish vegetation from each other based on tree canopy height (Tolan, 2024).

As a result, 9 auxiliary data sets were obtained, on the basis of which 5 combinations were formed for studying the optimal RF parameters and for vegetation classification (Table 1.)

Table 1. Combinations and their input layers

Combination number	Input layers
N1	spectral channels + spectral indices + topographic indices
N2	spectral channels + spectral indices + crown height
N3	spectral channels + topographic indices + crown height
N4	spectral indices + topographic indices + crown height
N5	spectral channels + spectral indices + topographic indices + crown height
Note – compiled by the authors	

Collecting training data. Extracting training and validation samples from the original map is a critical step in the digital vegetation mapping process. The quality of the training data directly affects the accuracy of the updated vegetation map. We extracted the area data from the GEE base map layer, which is built on high-resolution satellite data. Information from Planet Scope satellite images was also used to select training data. Training data for various land cover types were generated by selecting pixels through visual analysis, with each pixel being assigned a known land cover classification. This enabled the machine learning algorithm to learn and distinguish these types based on their spectral properties.

The training data was divided into 6 classes: trees, grass, bare soil, shrubs, snow, and water. By comparing data from two different sources, 4594 training samples were collected for 2019, 2021, and 2023.

Random Forest Algorithm. Today, RF is considered one of the most widely used algorithms for land cover classification using remote sensing data (Phan, 2020; Zeng, 2024; Pizarro, 2022; Mercier, 2019). Random Forest is a machine learning method that uses multiple decision trees for classification or regression analysis. Its advantages include high accuracy and resistance to overfitting when working with different types and high-dimensional datasets (Amani, 2017). The RF classifier requires two key parameters to operate: the number of randomly selected variables (mtry) used to build the decision tree at each splitting step, and the number of decision trees (ntree) (Fu, 2017). It is based on the concept of ensemble learning, in which multiple decision trees are built and their results are combined to produce the final prediction (Breiman, 2001).

RF uses each vote on the decision tree to produce results:

$$\mathrm{A}(\mathrm{a}) = \mathrm{B}_{\mathrm{arg}}^{x}\max\sum_{y-1}^{\mathrm{C}} \mathrm{d}_{y}^{x}(\mathrm{a}) \tag{5}$$

where: A(a) is the model based on RF extraction algorithm; *Bargx* is the *x*-labeling of the extracted class; *C* is the number of voting decision trees in the RF extraction algorithm forest; *dy* is the *y*-th voting decision tree in the RF extraction algorithm forest (Zeng, 2024).

All 5 sets of machine learning algorithm combinations were run using code written in GEE. During the training process, the training data set was split into training and validation data (8:2).

Accuracy assessment. In the course of the work, the following metrics were used to evaluate the accuracy: Overall Accuracy (OA) and Kappa Coefficient (KC), which were calculated using the functions “confusionMatrix.accuracy”, “confusionMatrix.kappa” on the GEE platform (Huang, 2017). The validation dataset was used to evaluate the accuracy of vegetation maps created on the basis of base maps from GEE, as well as digital maps obtained using various classification methods.

Overall Accuracy measures the effectiveness of the algorithm and is determined by the ratio of correctly classified samples to the total number of samples tested. Its values range from 0 to 1, where values approaching 1 indicate higher classification accuracy. Similarly, the Kappa coefficient, which ranges from -1 to 1, reflects the consistency of classification across all map types, with values closer to 1 indicating stronger agreement (Zeng, 2024).

Importance of Input Parameters. To determine the importance of input variables and the degree of their influence on classification, the MDI (Mean Decrease Impurity) method of variable importance assessment implemented in Google Earth Engine is used. MDI is a method for determining the importance of features in the RF algorithm. It measures how much each feature reduces the diversity in the trees of the RF algorithm and calculates the average value of this value for all trees in the algorithm. The greater the reduction in the diversity of a feature in the trees, the higher its importance (Agarwal, 2023).

Results and discussion. Based on the RF classification, this study generated 15 land cover maps, each showing a certain OA and QC accuracy value. The classification results were visualized as a map, where each vegetation type was displayed in its own color. As mentioned in the Materials and Methods section, in addition to the spectral channels of Planet Scope images, we also used additional variables to test whether they improve the accuracy of the land cover maps and to find an effective option for combining variables.

Accuracy evaluation results. The classification accuracy evaluation results shown in Table 2 show that the highest overall accuracy (OA) and kappa coefficient (KC) in 2023 were demonstrated by the N3 combination, which includes spectral channels, topographic indices and tree crown height layer, with OA estimates of 0.94 and KC of 0.92.

Table 2. Accuracy evaluation results for five combinations

Combination number	Accuracy assessment	year 2023	year 2021	year 2019
N1	OA KK	0.87 0.86	0.9 0.88	0.91 0.89
N2	OA KK	0.86 0.82	0.86 0.83	0.85 0.81
N3	OA KK	0.94 0.92	0.93 0.91	0.9 0.87
N4	OA KK	0.88 0.85	0.93 0.91	0.85 0.8
N5	OA KK	0.92 0.89	0.92 0.91	0.92 0.9
Note – compiled by the authors				

In 2021, the N3 and N4 combinations showed similar results, with a slight difference in OA of 0.0033 and KC of 0.0028, respectively. In 2019, the best combination was N5, which achieved the results of OA - 0.92 and KK - 0.9. The lowest indicators were shown by the combination N2, which included spectral channels, spectral indices and the tree crown height layer. According to Table 2, the combination N3 shows better results compared to the combination N5, although the difference between them is insignificant.

The results show that the use of topographic indices can provide a moderate to high fit to the training data. This is supported by the fact that all combinations including these variables showed high performance, while their absence resulted in worse results.

This increase can be explained by the order of importance of the variables in Figure 3 for all input features. In all three time series, the Elevation layer was always rated as the most important for classification. Among the spectral indices, the EVI index had the greatest influence, and among the spectral channels, the NIR and Blue bands played the greatest role. Overall, the topographic indices and spectral channels showed significant influence and were active, while the Canopy Height layer had the least influence.

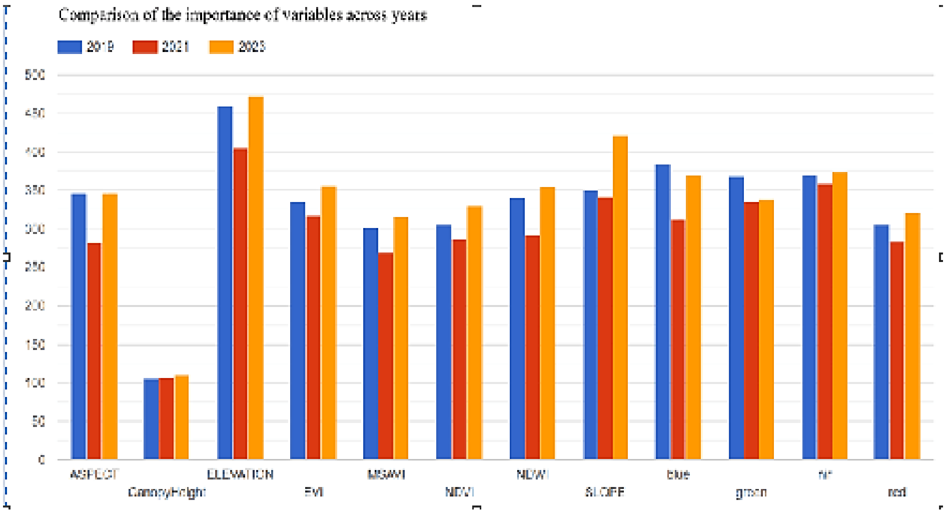

Figure 3. Comparison of the importance of variables for 2019, 2021, 2023

Note – compiled by the authors

Results of classification and analysis of change. Change detection is the process of identifying the differences between multiple raster datasets after performing classification and calculating vegetation indices. Based on the previous results, we selected the best classification combination N3 (Figure 4), and calculated the total area for each class in 2023 to analyze the changes in vegetation (Table 4). The spatial distribution of each class of vegetation cover is shown in the figure (map). According to the table, the largest territory is occupied by class 2, representing herbaceous vegetation, followed by class 4, open soil, which includes the entire earth's surface not covered by vegetation and water resources.

Table 4. Classroom area volume for 2019, 2021, 2023

Class designation (km2)	2023	2021	2019
0 – trees	560,108502	535,915764	642,222153
1 – shrubs	50,115681	140,416371	43,2504
2 – grass	1598,585184	1627,936533	1554,5142
3 – water	14,397363	15,856929	16,528482
4 – bare soil	675,131895	555,190722	642,925458
5 – snow	57,408606	80,430912	56,306538
Note – compiled by the authors			

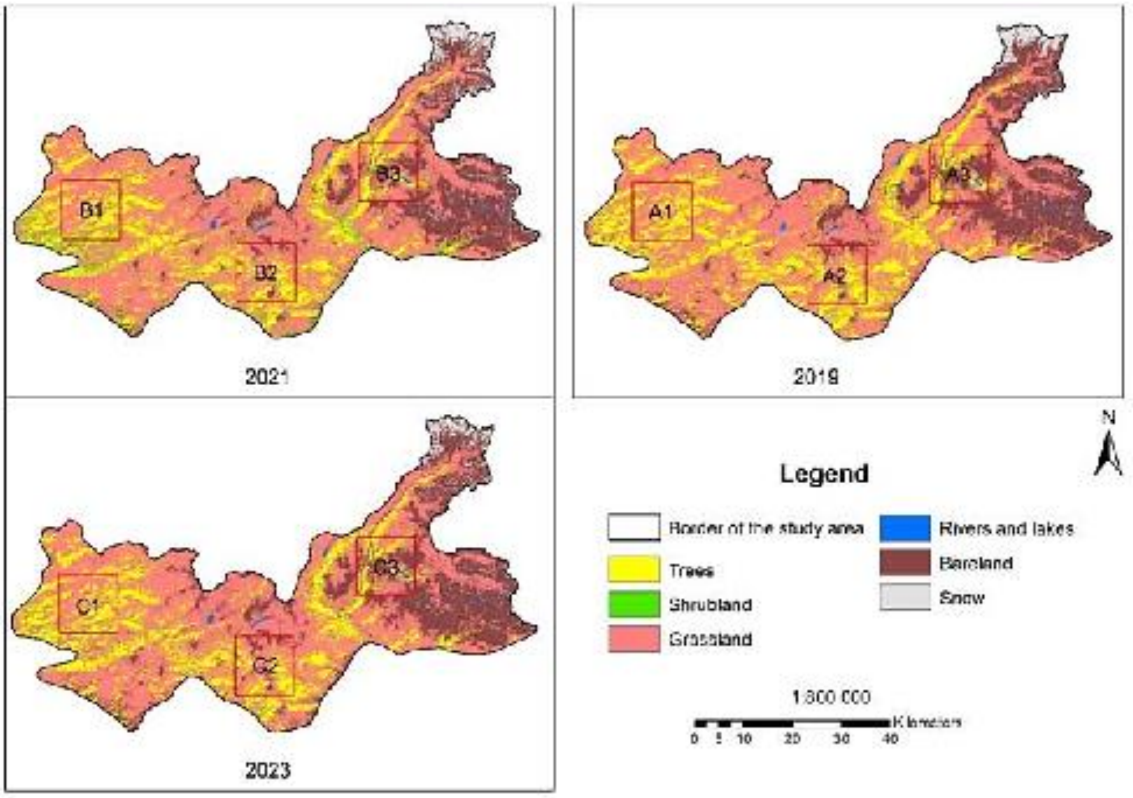

Figure 4. Map of vegetation cover of the study area for 2019, 2021, 2023 (combination N3)
Note – compiled by the authors

It is important to note that the time interval of the Planet Scope satellite data images has a significant impact on the classification results. As mentioned, the 2021 images are dated early July, while the 2019 images cover the period from early July to mid-August, which may lead to differences in snow and vegetation conditions between years. In Figure 5 with options A, B, C, you can clearly see the differences in the classifications for 2019, 2021 and 2023. In the "trees" class, the differences between the options are not particularly pronounced, while the "shrubs" and "open soil" classes demonstrate noticeable differences between 2021 and 2019, 2023 (the exact figures for the areas of the classes are presented in Table 4).

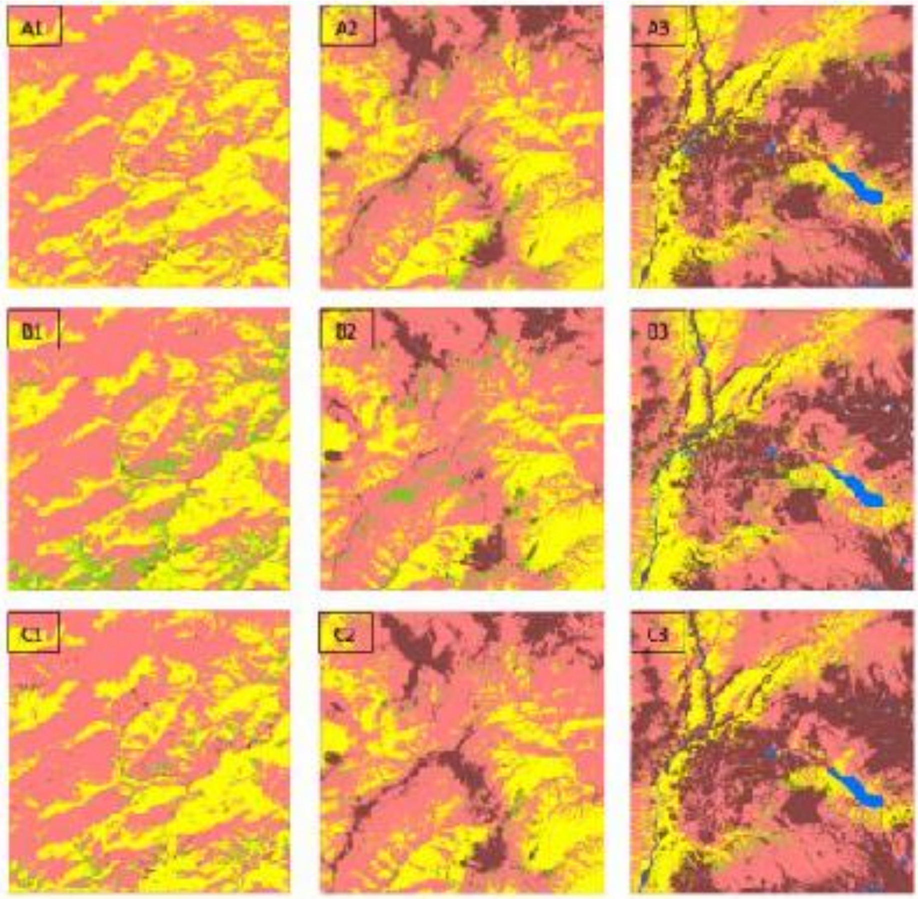

Figure 5. Comparison of RF classification results with the N3 combination set for 2019, 2021, 2023
Note – compiled by the authors

Sources of Error and Limitations of the Study. In this study, different feature variables had different effects on the land cover classification process. The main limitations of this work mainly relate to the differences between the satellite images obtained by Planet Scope and the reference data. Due to the limited data availability, including the lack of SWIR bands in the Planet Scope satellite images, spectral indices such as NDBI, NBR, NDSI, SVVI, which could improve the classification accuracy, were not calculated in this study (Pizarro, 2022). It is worth noting that mountainous areas often have cloud cover, which can make it difficult to obtain cloud-free images for the study area). Also, the final classification result was affected by the differences in the characteristics of the Planet Scope satellite sensors, which was reflected in the final composite images.

The choice of training data plays a key role in class assignment (Zeng, 2024). The amount and accuracy of training data used in the classifier training and testing phases were limited due to remote accessibility. This created difficulties in analyzing the classification results in combination with the available training dataset.

Conclusion. In this study, we present a comparative evaluation of five combinations of the RF algorithm using Planet Scope satellite imagery to generate vegetation type maps. Our analysis confirms that using different approaches to selecting input parameters results in different classification performances (OA varies from 0.8 to 0.94). The best performance was demonstrated by the N3 combination, which included topographic data, a tree crown height layer and four spectral channels, with an overall accuracy of 0.94 and a kappa coefficient of 0.92 for 2023. However, the performance of different combinations may vary depending on the characteristics of the study area and the data used.

The results of the study show that the use of cloud computing on the GEE platform and machine learning methods allows tracking the dynamics of vegetation change with further improvements to the method. Future research should focus on modern deep learning methods such as convolutional neural networks, which require extensive training data from field surveys and Planet Scope imagery. Further development of the methodology will reveal the full potential of digital vegetation mapping for addressing key ecosystem challenges.

Conflict of interest. The authors declare that there is no conflict of interest.

Acknowledgment. This research was supported by project IRN BR21882022 “Research of avalanche activity in the East Kazakhstan region for the development of monitoring systems and scientific justification for their placement”, funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

References

Agarwal, A., Kenney, A., Tan, Y., Tang, T., Yu, B. (2023). MDI+: A flexible random forest-based feature importance framework, https://doi.org/10.48550/arXiv.2307.01932.

Amani M., Salehi B., Mahdavi S., Granger J.E., Brisco B., Hanson A. (2017). Wetland classification using multi-source and multi-temporal optical remote sensing data in Newfoundland and Labrador, Canada. Canadian Journal of Remote Sensing, 43(4), 360-373. https://doi.org/10.1080/07038992.2017.1346468.

Breiman L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324.

Dobrinić D., Gašparović M., Medak D. (2021). Sentinel-1 and 2 time-series for vegetation mapping using random forest classification: A case study of northern Croatia. Remote Sensing, 13(12), 2321. https://doi.org/10.3390/rs13122321.

Feddema J.J., Oleson K.W., Bonan G.B., Mearns L.O., Buja L.E., Meehl G.A., Washington W.M. (2005). The importance of land-cover change in simulating future climates. Science, 310(5754), 1674-1678. https://doi.org/10.1126/science.1118160.

Fu B., Wang Y., Campbell A., Li Y., Zhang B., Yin S., Xing Z., Jin X. (2017). Comparison of object-based and pixel-based random forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data. Ecological Indicators, 73, 105-117. https://doi.org/10.1016/j.ecolind.2016.09.029.

Georganos S., Grippa T., Vanhuysse S., Lennert M., Shimoni M., Wolff E. (2018). Very high resolution object-based land use–land cover urban classification using extreme gradient boosting. IEEE Geoscience and Remote Sensing Letters, 15(5), 607-611, https://doi:10.1109/lgrs.2018.2803259.

Gilang N., Dian A. A. (2021). Comparison of soil adjusted vegetation index (SAVI) and modified soil adjusted vegetation index (MSAVI) methods to view vegetation density in Padang city using Landsat 8 image. International Remote Sensing Journal, https://doi: 10.24036/irsaj.v2i1.23.

Huang D., Xu S., Sun J., Liang S., Song W., Wang Z. (2017). Accuracy assessment model for classification result of remote sensing image based on spatial sampling. Journal of Applied Remote Sensing, 11(1), 046023. https://doi.org/10.1117/1.jrs.11.046023.

Huete A., Didan K., Miura T., Rodriguez E., Gao X., Ferreira L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment, 83(1-2), 195-213. https://doi.org/10.1016/S0034-4257(02)00096-2.

Hengl T., Heuvelink G.B.M., Kempen B., Leenaars J.G.B., Walsh M.G., Shepherd K.D., et al. (2015) Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. PLoS ONE 10(6): e0125814. https://doi.org/10.1371/journal.pone.0125814.

Mercier A., Betbeder J., Rumiano F., Baudry J., Gond V., Blanc L., Bourgoin C., Cornu G., Ciudad C., Marchamalo M., et al. (2019). Evaluation of Sentinel-1 and 2 time series for land cover classification of forest–agriculture mosaics in temperate and tropical landscapes. Remote Sensing, 11(8), 979. https://doi.org/10.3390/rs11080979.

McFeeters S.K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425-1432. https://doi.org/10.1080/01431169608948714.

Pizarro S.E., Pricope N.G., Vargas-Machuca D., Huanca O., Ñaupari J. (2022). Mapping land cover types for highland Andean ecosystems in Peru using Google Earth Engine. Remote Sensing, 14(7), 1562. https://doi.org/10.3390/rs14071562.

Phan T.N., Kuch V., Lehnert L.W. (2020). Land cover classification using Google Earth Engine and random forest classifier – the role of image composition. Remote Sensing, 12(15), 2411. https://doi.org/10.3390/rs12152411.

Sheykhmousa M., Mahdianpari M., Ghanbari H., Mohammadimanesh F., Ghamisi P., Homayouni S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 6308-6325, https://doi:10.1109/jstars.2020.3026724.

Sun J., Ongsomwang S. (2023). Optimal parameters of random forest for land cover classification with suitable data type and dataset on Google Earth Engine. Frontiers in Earth Science, 11, 1188093. https://doi.org/10.3389/feart.2023.1188093.

Tolan J., Yang H.I., Nosarzewski B., Couairon G., Vo H. V., Brandt J., Spore J., Majumdar S., Haziza D., Vamaraju J., Moutakanni T. (2024). Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. Remote Sensing of Environment, 300, 113888, https://doi.org/10.1016/j.rse.2023.113888.

URL: https://earth.esa.int/eogateway/missions/planetscope

Zeng X., Guo X., Jiang Y., Li W., Guo J., Zhou Q., Zou H. (2024). High-accuracy mapping of soil parent material types in hilly areas at the county scale using machine learning algorithms. Remote Sensing, 16(1), 91. https://doi.org/10.3390/rs16010091.

Бельгибаев М.Е., Нурпеисова А. (2007). Шығыс Қазақстан облысының қорықтары. География Казахстана и экология, (5), 3-7. // Bel'gibaev M. E., Nurpeisova A. (2007). Shygys Kazakhstan oblysynyn koryktary. Geografiya Kazakhstana i ekologiya, (5), 3-7.

Information about authors

Rakhymberdina Marzhan Yessenbekovna – D. Serikbayev East Kazakhstan Technical University, 070010, Ust-Kamenogorsk, Serikbayev st, 19, Dean of the School of Earth, tel.87083698133, e-mail: MRahymberdina@ektu.kz

Daumova Gulzhan Kamalbekkyzy – D. Serikbayev East Kazakhstan Technical University, 070010, Ust-Kamenogorsk, Serikbayev st 19, Professor of the School of Earth, tel. 8-777-396-12-47, e-mail: GDaumova@edu.ektu.kz*

Zhussupova Gulsezim Kainarovna – Ionosphere Institute LLP, 050000, Almaty, Kamenskoye Plateau, Horticultural Association "Ionosphere", 117, engineer, tel. +7(771)1810428, e-mail: zhussupova51@gmail.com

Chettykbayev Ruslan Kairatovich – D. Serikbayev East Kazakhstan Technical University, 070010, Ust-Kamenogorsk, Serikbayev st 19, Head of the Center for Competencies and Technology Transfer in the field of information technology, tel. 8-705-650-01-01, e-mail: rchettykbayev@edu.ektu.kz

Chepashev Daniker Vasilievich – Ionosphere Institute LLP, 050000, Almaty, Kamenskoye Plateau, Horticultural Association "Ionosphere", 117, Head of the Laboratory for Space Monitoring of Emergency Situations, tel. +7(707)4648943, e-mail: d.chepashev@ionos.kz