AUTOMATIC FOREST COVER CLASSIFICATION USING SENTINEL-2 MULTISPECTRAL SATELLITE DATA AND MACHINE LEARNING ALGORITHMS IN GOOGLE EARTH ENGINE

Katarína Onačillová[1],Veronika Krištofová[2], Daniel Paluba[3]

1 Pavol Jozef Šafárik University in Košice, Faculty of Science, Institute of Geography, e-mail: katarina.onacillova@upjs.sk

2 HRDLIČKA – SLOVAKIA s.r.o., Moyzesova 46, 040 01 Košice, Slovakia, e-mail: veronika.kristofova@hrdlicka.sk

3 Charles University, Prague, Czech Republic, Faculty of Science, Department of Applied Geoinformatics and Cartography, e-mail: daniel.paluba@natur.cuni.cz

Abstract: Forest cover plays an essential role in maintaining ecological equilibrium, mitigating climate change, and securing a sustainable future for both humanity and the planet. Most countries conduct forest inventory or remote sensing surveys every few years to monitor changes in forest cover. However, only a few initiatives offer more frequent updates, typically weekly or monthly, focusing exclusively on areas experiencing high rates of deforestation or those of significant ecological value. The present study focuses on the classification of forest cover throughout Slovakia, covering the period 2017-2022, using Sentinel-2 multispectral satellite imagery along with the machine learning (ML) algorithms Random Forest (RF) and Support Vector Machine (SVM). The computation was performed in the cloud-based Google Earth Engine (GEE) platform, which offers a versatile interface for a broad range of computational capabilities for geospatial analysis and landscape monitoring. Forest cover change processing and evaluation is based on the RF classification algorithm, which demonstrated higher accuracy than the SVM classifier. The results indicate that RF outperformed SVM by 4% and 21% in 2017 and 2020, respectively. The RF algorithm achieved an overall accuracy (OA) of 95% in both classification cases (2017 and 2020) and F1 score of up to 0.95. The selected RF algorithm revealed an increase in forest cover in Slovakia, particularly notable during the period 2017-2019, with a slight decrease detected between 2019 and 2020. Furthermore, it was determined that the current forest cover is lower than that reported in official state statistics and land cover databases. Additionally, a user-friendly automatic tool for forest cover classification was developed and made freely available in GEE. This tool can benefit foresters, urban planners, and everyday users by detecting subtle changes in forest cover, crucial for forest sustainability and human well-being.

Keywords: Google Earth Engine, Sentinel-2, forest cover, image classification, machine learning, Slovakia

1 INTRODUCTION

Forests represent an irreplaceable natural resource providing a wide range of ecosystem services essential to the cycle of life on Earth. They play an important role in protecting biodiversity, reducing the effects of global warming in the world through the carbon cycle process, the hydrological cycle, mitigating soil erosion, and human recreation (Alkama and Cescatti, 2016). Despite many essential functions, these ecosystems have been exposed to a great danger of threats to their condition, longevity, and quality in recent decades (Hansen et al., 2013). Since 1990, the world has lost more than 178 million forests, which represents approximately 0.20% of the world's forest cover, although the decline rate in net forest loss has slowed significantly between 2010-2020 due to the decline in forest expansion (FAO, 2022). According to official FAO statistics, Europe is the only continent where forest cover increases over a long period of time. However, in the years 2010-2020, significantly lower values of the net increase rate in forest area were recorded in Europe (and Asia) than in the years 2000-2010 (FAO, 2020).

Similarly, within Central Europe, according to official inventories, the territory of Slovakia shows a permanent and long-term increase in forest cover (MŽP, 2021). There are nine national parks dedicated to the protection of extensive ecological processes along with species diversity and ecosystems that are specific to this area (IUCN, 2020; Šebeň, 2017). However, as stated by the Institute for Environmental Policy (IEP, 2017), although the area of forests is increasing according to official data, international satellite images that show the real state of forest land and can objectively assess the condition of forests indicate that the opposite is true, and that since the beginning of the 21st century the area of Slovak forests has decreased by an average of 0.46% annually. The main difference lies in the differences in the methods used and in the understanding of the forest itself.

Today, both in the world and Slovakia, there is still a noticeable inconsistency of methodologies for monitoring forests conditions and the lack of sustainable development strategies to resolve forest loss and degradation, while balancing the economic, social and environmental benefits of forestry. Most countries conduct forest inventory or remote sensing (RS) surveys only every few years to monitor changes and forest cover quality. Several models and datasets have been developed to analyze forest cover changes, but these are usually based on low temporal resolution discrete data and may therefore ignore key factors that influence forest changes (Matthews et al., 2007). Therefore, there is an urgent need for models and applications that help stakeholders and the public to understand the real state, the continuous spatio-temporal dynamics of forest cover, to identify key factors driving forest change and disrupting forest ecosystems over time, to quickly intervene and reduce the harmful, perhaps illegal, effects on forest loss and better design possible future changes (Tang et al., 2019).

Significant potential for investigating forest stands over the past decade has been provided by the Sentinel-1 and Sentinel-2 satellite imagery of the European Space Agency (ESA) Copernicus program, which can effectively monitor changes, even in large areas of forests with high temporal and spatial resolution, with the help of parametric and non-parametric ML classifiers (Viana et al., 2019). The importance of these satellite data in the classification of forest types using Sentinel-2 data was also emphasized by Immitzer et al. (2019), or using a combination of Sentinel-1 and Sentinel-2 data by Lechner et al. (2022). Thanks to the GEE tool (Gorelick et al., 2017), which represents an innovative cloud platform with access to a catalog with a rich collection of satellite images and other RS data, it is now possible to efficiently perform geospatial analysis at regional and global level in high temporal and spatial resolution. The use of this platform by the RS scientific community has increased almost exponentially in recent years (Yang et al., 2022).

The aim of this paper was to develop an automated approach for forest cover classification using Sentinel-2 imagery and supervised ML classification algorithms RF and SVM in a GEE environment. A tool was developed and applied for the territory of Slovakia and its regions for the period 2017-2022. As input data, selected Sentinel-2 satellite bands, normalized difference vegetation index (NDVI) and the digital elevation model SRTM (Shuttle Radar Topography Mission) were used. The combination of the Corine Land Cover (CLC) and Global Forest Change (GFC) databases was used to create validation training samples. The accuracy of classification results using the RF and SVM algorithms was then assesed using the script created, and the changes in the state of the forest cover in the selected area were evaluated. The results of image classification in the GEE environment were also compared to the official data obtained from the Statistical Office of the Slovak Republic (SOSR) and to CLC data on forest area.

2 DATA AND METHODS

The analyses in this paper were performed in the GEE environment. To demonstrate the effective use of this platform for image classification and analysis of large-volume satellite data, even for large areas, the entire territory of Slovakia was chosen as the area of interest. In the first step, input data were prepared, namely Sentinel-2 multispectral satellite data, SRTM elevation model, CLC and GFC database. From the Sentinel-2 dataset, images with a cloud cover below 5% were chosen for a selected time range. The applied input bands were then enhanced with NDVI and SRTM. From the selected time series of images, a median mosaic was created for each pixel and each input band. A forest mask was created with the intersection of the CLC and GFC databases, which was also used to generate 2000 randomly generated training data. In the next step, two ML algorithms, RF and SVM, were tested by evaluating accuracy using error matrices. The algorithm with better results, RF, was used to produce land cover maps and quantify the year-to-year changes in forest area in the period 2017-2022 for the entire Slovakia and individual regions, and then compared with official SOSR data. Figure 1 depicts the work's methodology. A detailed description of the individual steps can be found in the following chapters.

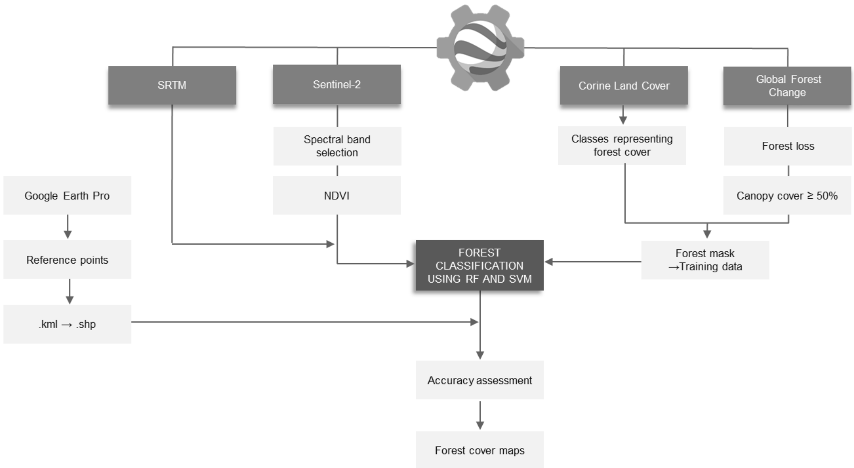

Figure 1 Flowchart of the methodology used in this work. Source: processed by the authors

2.1 Area of interest

Forest cover classification using ML algorithms and Sentinel-2 multispectral data was carried out for the territory of Slovakia with an area of 49,036 km^2 located in Central Europe (Figure 2). Although Slovakia is one of the smaller countries in terms of size, its geographical location and regional differences in altitude allow to observe a variety of natural and climatic conditions, diverse ecosystems and land cover types.

Slovakia's forest cover today reflects long-term development and the effects of both natural and man-made factors. With 22% and 18% of Slovakia's total forest area, respectively, beech and spruce forests make up the largest proportion of the country's forest cover. Compared to the primary forest structure in Slovakia, there is currently a decline in species such as oaks, beeches and firs, but an increase in spruce and pine can be observed (Minďáš et al., 2006). However, this period is also characterized by the intentional planting of new forests, mainly spruce monocultures. The State of Europe's Forests report by Forest Europe (2020) states that Slovakia ranks 13th out of 43 European countries in terms of forest cover. One of the main current challenges of forest management is the need to face the increasing risk of harmful factors that affect overall forest cover. The loss of forest cover due to human activity persists at the expense of the economy, and forests are frequently turned into agricultural land, which not only negatively affects the soil and water cycle but also reduces the ability of forests to regulate the environment. In addition, biotic factors (fungi, insects and bacteria) and, to a large extent, abiotic factors such as weather extremes also affect the forest. Wind calamities, which have occurred frequently in Slovakia's history and are still observed annually, have a particularly detrimental effect on the forests.

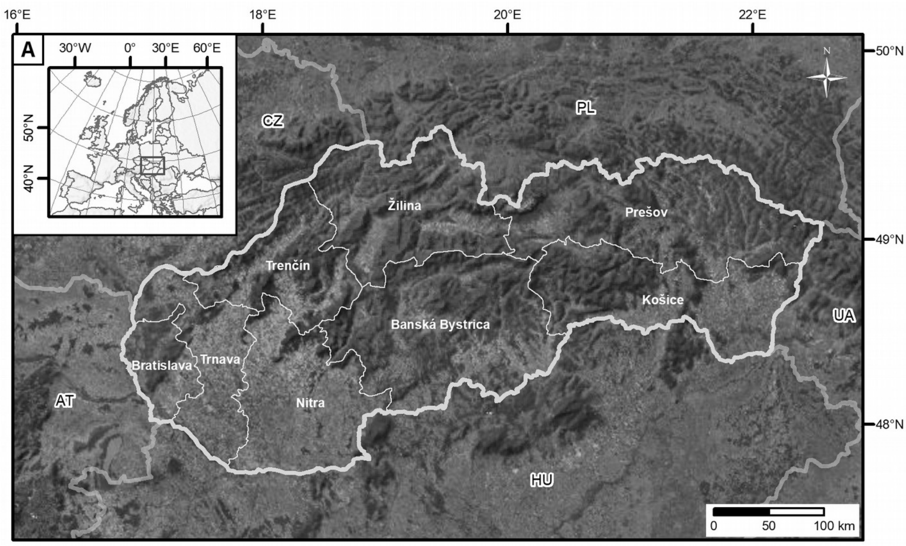

Figure 2 Location of the area of interest within Europe. Source: ©ESRI basemap (2022)

2.2 Input data

The main source of data for the supervised classification was Sentinel-2 multispectral satellite imagery at the Level-2A preprocessing level. The Copernicus Sentinel-2 mission consists of a constellation of two satellites, Sentinel-2A and Sentinel-2B, providing optimal coverage and data for the entire planet. Due to its high spatial, temporal (data acquired on the same location by two satellites approximately every five days), and spectral resolution (10-60 m) this mission provide valuable information particularly for agricultural research and forest vegetation monitoring (Drusch et al., 2012). Sentinel-2 Level-2A data have been available in GEE since 28 March 2017, therefore images from 2017 were used in this study.

Consequently, the time period for satellite image retrieval was defined only for the growing season, i.e. April-October, as this setting allows for a better differentiation of forest cover from other land cover classes with a lower probability of cloud and snow cover. Another important factor was setting the maximum cloud cover criterion at 5%, which eliminated the input of cloud-covered scenes in image classification that could bias the classification result. Testing and then selecting those spectral bands that had the greatest impact on image classification also contributed to achieving high accuracy in forest cover classification. To classify forest cover, only the bands listed in Table 1 were used.

Table 1 Properties of selected bands of Sentinel-2 satellite.

Band	Name / description	Central wavelength (μm)	Spatial resolution (m)
Band 2	Visible blue	0.490	10
Band 3	Visible green	0.560	10
Band 4	Visible red	0.665	10
Band 5	Vegetation Red Edge	0.705	20
Band 6	Vegetation Red Edge	0.740	20
Band 8	Near Infrared (NIR)	0.842	10
Band 11	Short Wavelength Infrared 1 (SWIR)	1.610	20
Band 12	Short Wavelength Infrared 2 (SWIR)	2.190	20

Source: data based on Drusch et al. (2012), European Space Agency (2015)

Based on the testing of the image classification results, we discovered that adding the SRTM and NDVI to the input feature set improves the accuracy of image classification. Also in the study of Svoboda et al. (2022) for Czechia, the results of the land cover classification were improved when NDVI and SRTM were included. The NDVI is calculated as a normalized difference of visible infrared (B8) and visible red (B4) bands, using the following equation (Rouse et al. 1974):

$$NDVI = (B8 - B4)/(B8 + B4)$$

Elevation information from SRTM was added as an input feature for image classification primarily to represent heights that affect the occurrence of forest vegetation. The SRTM data are provided at a resolution of 1 arcsecond (approximately 30×30 m) (Farr et al., 2007). We have used the latest version available in GEE, SRTM90_V4, which, unlike previous versions, has been processed to fill data gaps to ensure ease of use.

The so-called median composite was then used as input into the classification, which was created by calculating the median value for each pixel and each input band across Slovakia from the time series of available Sentinel-2 images for a selected period in a given year. Visual analysis showed that the period between 20 April and 10 October yields the least cloudy composites for each year, so we used this range in our analyses. The median composite was chosen because it is less susceptible to possible outliers due to potential cloud and snow cover effects compared to the mean composite.

2.3 Training data preparation

A sufficient number and quality of training samples are another prerequisite for successful classification. Based on the extent of the area of interest, training plots were created for two classification classes to validate the resulting forest cover classifications. Sites covered by forest vegetation were classified into the forest class, and sites without forest cover into the non-forest class. For the creation of training samples, we relied on the research by Paluba et al. (2021), in which a forest mask was created with an overall accuracy of more than 90%. Specifically, we used two freely available land cover databases in GEE for this purpose, namely the CLC by Büttner et al. (2017) and GFC by Hansen et al. (2013). The CLC product provides land cover layers classified into 44 classes, in the form of five updated databases for 1990, 2000, 2006, 2012 and 2018. These databases were created using satellite image classification and in-situ data, and use a minimum mapping unit (MMU) of 25 hectares for areal phenomena and a minimum width of 100 m for linear phenomena. In this work, the most up-to-date CLC layer (CLC2018) was used, of which 3 classes were selected to represent forest cover: deciduous (311), coniferous (312) and mixed forest (313). Forests are defined in the CLC as forest vegetation covers taller than 5 m with a minimum canopy cover of 30%. For young plantations, a minimum threshold of 500 subjects per hectare is considered (Kosztra, 2019).

Another database used to generate the training data was the GFC global-scale database, which is the result of a time series analysis of Landsat imagery and describes forest cover, loss and gain at a spatial resolution of 30 m. Similar to the CLC database, trees are defined in this database as vegetation taller than 5 meters. Forest loss is defined in the GFC database as a disturbance of the forest area, where the forest-covered area has changed to an area without forests. This loss should represent a change of at least 50% of the canopy cover at pixel level. Forest gain is defined as the exact opposite of the forest loss. It is the phenomenon where a non-forest area is converted to forest (Hansen et al., 2013). For the purpose of this work, GFC version 1.10 (Hansen et al., 2023), containing data from 2000 to 2022, was used.

We selected pixels with a tree cover greater than 50% using the tree cover layer from 2000 (treecover2000) from the GFC database. Subsequently, the treecover2000 layer was masked using the layer representing forest cover loss for each year (lossyear) to produce the resulting GFC layer. The final forest mask (with two classes: "forest" and "non-forest") was created using the intersection of the resulting GFC layer with CLC2018. The "non-forest" class contained all types of land cover except forest. Using the intersection of the two databases, i.e. selecting pixels that belong to the "forest" category in both databases, helps to reduce errors that may appear in each database.

In the next step, the training dataset was created by generating 2000 random points. The values of all input bands were extracted to each point. Using the final forest mask, it was automatically determined for all points whether they represented forest or non-forest area. This is an automatic process whose advantage is that the input parameters for the classification can be changed quickly and efficiently; that is, the GFC forest loss layer can be used for any year period since 2000.

2.4 Classification algorithm selection

To select an appropriate classification algorithm, it was necessary to review all factors that influence the classification. These include the spatial resolution of the

RS data, the quality and number of training samples, availability of classifiers in the selected environment and the source of the reference data (Lu and Weng, 2007).

The GEE platform has a variety of classification algorithms that can be implemented according to the user's needs. To obtain relevant results, two supervised ML classification algorithms, RF and SVM, were used and compared. Since the beginning of the 21st century, the SVM algorithm has represented a possible alternative with higher classification accuracy compared to the numerous algorithms used until then (Pal and Mather, 2005). It is a nonparametric classification algorithm that can achieve good classification results even from complex data that contain noise (Chuvieco, 2020) and also when using fewer training data (Mantero et al., 2005). The RF classification algorithm operates on the basis of building multiple decision trees using a randomly selected subset of training samples and variables (Belgiu and Drăguţ, 2016). It is becoming increasingly used for its high accuracy, reliability and efficiency in processing high-dimensional and non-linearly distributed data. For example, it is often used to map forest stands and complex ecosystems (Waśniewski et al., 2020; Li et al., 2022), arable land (Phalke et al., 2020), mountainous areas (Hościło and Lewandowska, 2019), and has also shown its potential in simulating water salinity (Khan et al., 2020).

In GEE, RF number of trees was set due to the computational speed to 10, which is the default setting of this classifier in the GEE environment. However, the user should keep in mind, that setting greater number of trees tends to stabilize the out-of-bag error, thus also the resulting classification accuracy, while being computationally more expensive. The other parameters were left to the default, while SVM was used with the default parameters, i.e., with a linear kernel. Since no other kernel setting could efficiently classify the forest cover of the entire territory of Slovakia in the GEE, the linear kernel for the SVM was selected. This is due to the computational limitation of GEE – a large area is associated with a greater computational complexity of other types of kernels – this in the free GEE version caused that other types of kernel were out of the computational capacity.

2.5 Accuracy assessment

Evaluation of classification accuracy was particularly important to determine a more reliable classifier for forest cover classification. The accuracy assessment was performed separately for 2017 and 2020 classification results and for the two classification algorithms used. We selected 2017 as the first year with Sentinel-2 data available in GEE and also because the CLC2018 layer was built on 2017 data, so the accuracy of CLC2018 should be most accurate in this year. The year 2020 was chosen due to the noticeable high forest loss recorded between 2017 and 2020 in the Veľká Fatra and Malá Fatra National Parks (MŽP, 2021).

For each chosen year, two independent reference datasets generated based on the high-resolution open-access satellite imagery in Google Earth Pro (GEP) were used to verify the classification's accuracy. The reference dataset consisted of 200 reference points, representing 10% of the training dataset, with 100 points representing the 'forest' class and another 100 points representing the non-forest class. To validate the reference points for 2017, GEP mosaics consisting mostly of imagery from 2017 were used. The reference points for 2020 were validated using GEP mosaic consisting mostly of imagery from 2020. The reference datasets for both years were exported in.kml format, further converted to shapefile (.shp) and then imported into the GEE environment, where the classifications' accuracy of the algorithms were evaluated and expressed in terms of error matrix and metrics such as overall accuracy (OA), User's Accuracy (UA), Producer's accuracy (PA) and kappa coefficient. In addition, the overall $F1$ score was computed according to the formula by Long et al. (2015):

$$F\,1-score=2*\frac{precision*recall}{precision+recall} \quad .$$

3 RESULTS AND DISCUSSION

3.1 Evaluation of the accuracy of the SVM and RF algorithms

First, the accuracy of classifications using the SVM and RF algorithms was compared for 2017 and 2020. Classification outputs are shown in Figure 3. Accuracy metrics of both algorithms using error matrices for both years are displayed in the Table 2.

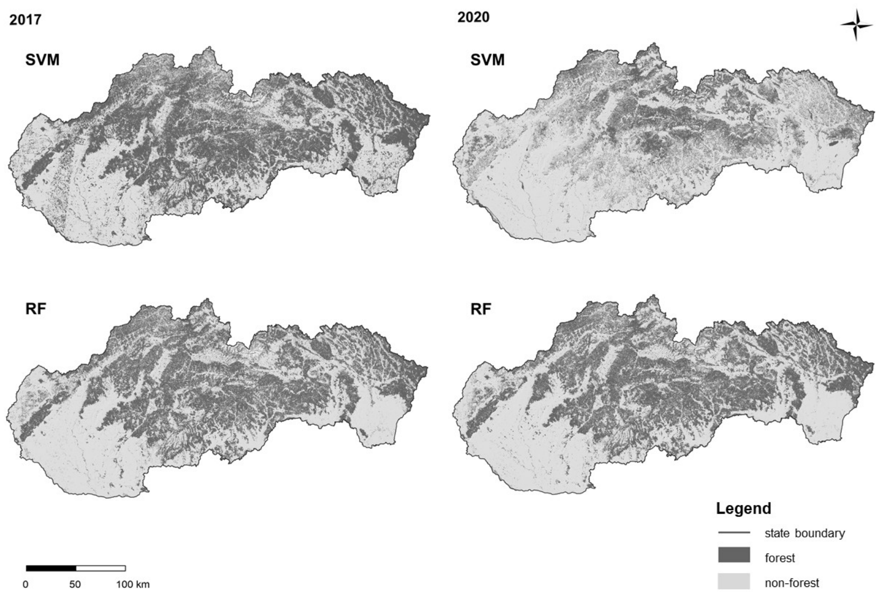

Figure 3 Forest classification using SVM and RF classification algorithms for 2017 and 2020. Source: prepared by the authors

Table 2 Error matrix for image classification using the SVM and RF classifiers for 2017-2022

SVM 2017					
Class	Reference data		Total		UA(%)
	Non-forest	Forest			
Non-forest	91	8	99		91.92
Forest	9	92	101		91.09
Total	100	100	200		
PA (%)	91	92		0.91	OA
F1-score	0.91			Kappa	0.83

RF 2017					
Class	Reference data		Total		UA(%)
	Non-forest	Forest			
Non-forest	98	9	107		91.59
Forest	2	91	93		97.85
Total	100	100	200		
PA (%)	98	91		0.95	OA
F1-score	0.95			Kappa	0.89

SVM 2020					
Class	Reference data		Total		UA(%)
	Non-forest	Forest			
Non-forest	96	48	144		66.67
Forest	4	52	56		92.86
Total	100	100	200		
PA (%)	96	52		0.74	OA
F1-score	0.79			Kappa	0.48

RF 2020					
Class	Reference data		Total		UA(%)
	Non-forest	Forest			
Non-forest	100	11	111		90.09
Forest	0	89	89		100
Total	100	100	200		
PA (%)	100	89		0.95	OA
F1-score	0.95			Kappa	0.89

Source: processed by authors

The OA of the four evaluated classifications using the SVM and RF algorithms varied from 0.74% to 0.95%. The values of the Kappa coefficient ranged from 0.48% to 0.89%. The RF algorithm achieved higher overall accuracy and behaved approximately equally stably, showing an OA of 95% in both classification cases (2017 and 2020). For the forest class, the RF results in 2017 and 2020 produced PA's of 91% and 89%, respectively.

Although the OA using RF algorithm is high, the forest cover layers show some EOs (e.g. non-contiguous areas of the forest) and commission (e.g. dense crops, grasslands or dwarf mountain pine classified as forest). Similar results were obtained by Waśniewski et al. (2021), in whose study RF achieved the lowest accuracy in places where the forest was disturbed (UA = 61% and PA = 71%) and also in the study by Conette et al. (2016) where RF showed the lowest accuracy in classifying the degraded forest class, where up to 24% of forest reference points were misclassified as non-forest classes.

Using the same training samples, the SVM classification algorithm showed significant classification inaccuracy for 2020 (OA of 74%, Kappa of 0.48) compared to the RF algorithm. The SVM algorithm's classification results inaccurately expressed significant differences in the forest cover. These errors consisted of incorrectly classified areas, e.g. water bodies or meadows and pastures, for example, in the vicinity of Liptovská Mara and Starina, which the SVM mistakenly classified as "forest" (Figure 4), as can be seen in comparisons of true colour (RGB: B4-B3-B2) and false colour compositions (CIR: B8-B4-B3) from Sentinel-2 satellite imagery and results for image classification using SVM and RF classifier. The higher classification error can also be attributed to the fact that SVM was used with the so-called linear kernel. The choice of this SVM was conditioned by the fact that only this specific type of kernel worked successfully in GEE for the entire territory of Slovakia.

The classification using SVM for 2020 revealed 1.085 million. ha, that is, by more than 1 mil. ha of forest lower status compared to the RF classification. In contrast, according to the RF algorithm, there was an increase in forest cover in Slovakia by almost 24 thousand ha (from 1.810 to 1.834 million ha).

Considering the accuracy assessment, the more reliable was RF algorithm, since it worked with higher accuracy than SVM in both cases. Due to its reliability, the RF algorithm has been preferred in many previous studies. In the work of Li et al. (2022), RF demonstrated a strong potential for forest vegetation classification with an OA of 97.57% and a Kappa value of 0.95. OA equal to 95%, similar to our case, was also achieved by the RF algorithm in the study of Nomura and Mitchard (2018), who monitored tropical forest vegetation using Sentinel-2 satellite images.

This accuracy assessment method has its limitations. The number of reference points set at 200 is not so representative of the entire territory of the Slovak Republic. To enhance the model's generalization, a greater quantity of testing points should be used in the next study for a more comprehensive evaluation.

3.2 Evaluation of changes in forest area in Slovakia

In the further analysis, we exclusively used the outcomes of a more reliable classification using the RF algorithm. To determine the state of forests, we created a layer from the image classification – a mask, which contained only the "forest" class. Subsequently, with the help and flexibility of the created tool, we were able to determine the exact classified forest area and monitor changes at the level of the entire Slovakia, but also at the regional level. However, the calculation of the forest area can be applied to any time period and territory within Europe, the GEE environment also offers a number of datasets containing boundary layers (e.g. FAO district boundaries, boundaries of protected areas NATURA 2000, etc.) and the import of own data that can also be used for the classification of areas of interest and the creation of training samples, even for places that are difficult to access or too large for in-situ data collection.

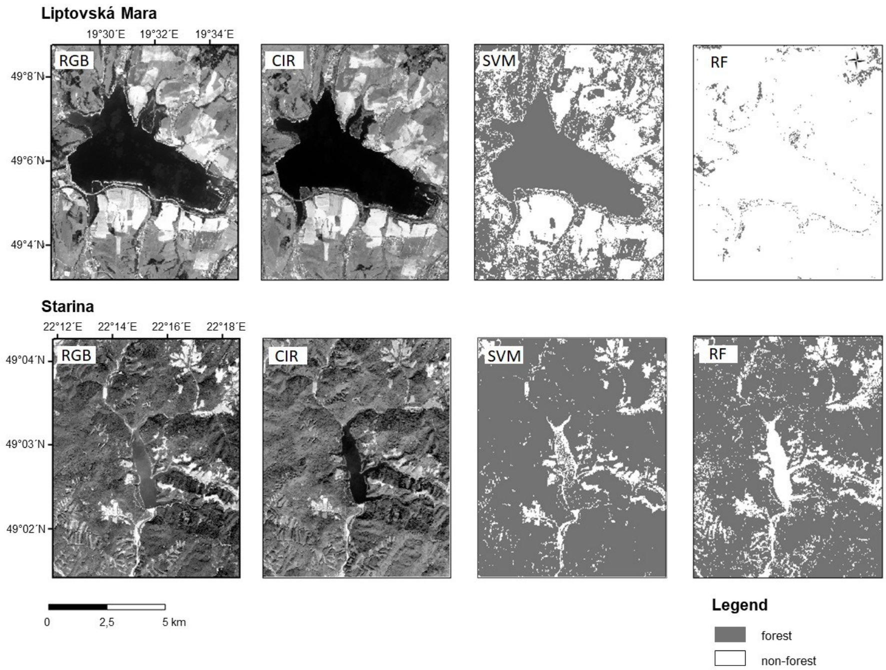

Figure 4 SVM classification results compared to RF for the areas of Liptovská Mara and Starina. From the left: true colour composition (RGB: B4-B3-B2), false colour composition (CIR: B8-B4-B3) (both converted to greyscale), SVM classification results, RF classification results. Source: background colour compositions of Sentinel-2 images from 08/22/2020, processed by the authors

Based on the RF classification results, we can observe the increase in the forest area between 2017 and 2018, by 17,890.48 ha, i.e. 0.99% (Table 3, Figure 5A). The largest increase occurred in 2018-2019 (24,831.68 ha). In 2019, one of the highest levels of forest cover in the monitored period was observed, 1,852,785.6 ha. The decrease by 19,285.11 ha can be observed from 2019 to 2020, resulting in a decrease in forest cover by 1.04% from 2019. In the following years 2021-2022, we can observe an increase in forest area again, with the year 2022 marking the end of our monitoring period and exhibiting the highest forest cover during the 2017-2022 period. This increase can be explained by the effective elimination of calamity wood (a remnant of past wind disasters) and bark beetles. Comparing the beginning and end of the examined period (years 2017 and 2022), the largest gains in the forest cover can be observed southwest of the town Brezno, east of the city Nové Mesto nad Váhom and Trenčín, and also in the area of the High Tatras National Park. On the other hand, the highest losses are visible in the area of Slovak Paradise National Park, west of the city of Košice and near the Prešov city and Prievidza (Figure 5B).

Table 3 Forest area in Slovakia for the period 2017-2022 based on the RF classification

Year	Forest area [ha]
2017	1,810,063.44
2018	1,827,953.92
2019	1,852,785.60
2020	1,833,500.49
2021	1,844,562.41
2022	1,855,912.20

* the median composite was calculated for the period April 1 - October 31 due to the higher effects of cloudiness in the original range.

Source: processed by authors

These classification results are accessible also through the forest cover database we created for the entire territory of Slovakia covering all years from 2017 to 2022. The database stored in raster format has a spatial resolution of 30 m and is freely available as a GEE Image Collection and can be imported using the following code:

ee. ImageCollection('users/danielp/Slovakia_forests_2017-2022').

3.3 Comparison of the classification results with official SOSR data on forest land

This section compares the results of our RF classification with the official SOSR data on forest land for each region and the entire SR during the monitored period (Table 4).

Compared to the SOSR data, the forest cover value calculated from the RF classification was lower in all cases. SOSR reports that the forest cover in Slovakia gradually increased during the monitoring period, by 0.05% in total. Our RF classification shows that the amount of forest cover increased between 2017 and 2019, decreased slightly in 2020, and then increased again at the end of the monitoring period. In 2022, according to the RF classification, forest areas accounted for 37.86% of Slovakia's total area – approximately 4% less than according to SOSR data. According to the RF classification results, forest cover increased by 0.95% between 2017 and 2022, with the highest increase in 2017-2019 (0.88%).

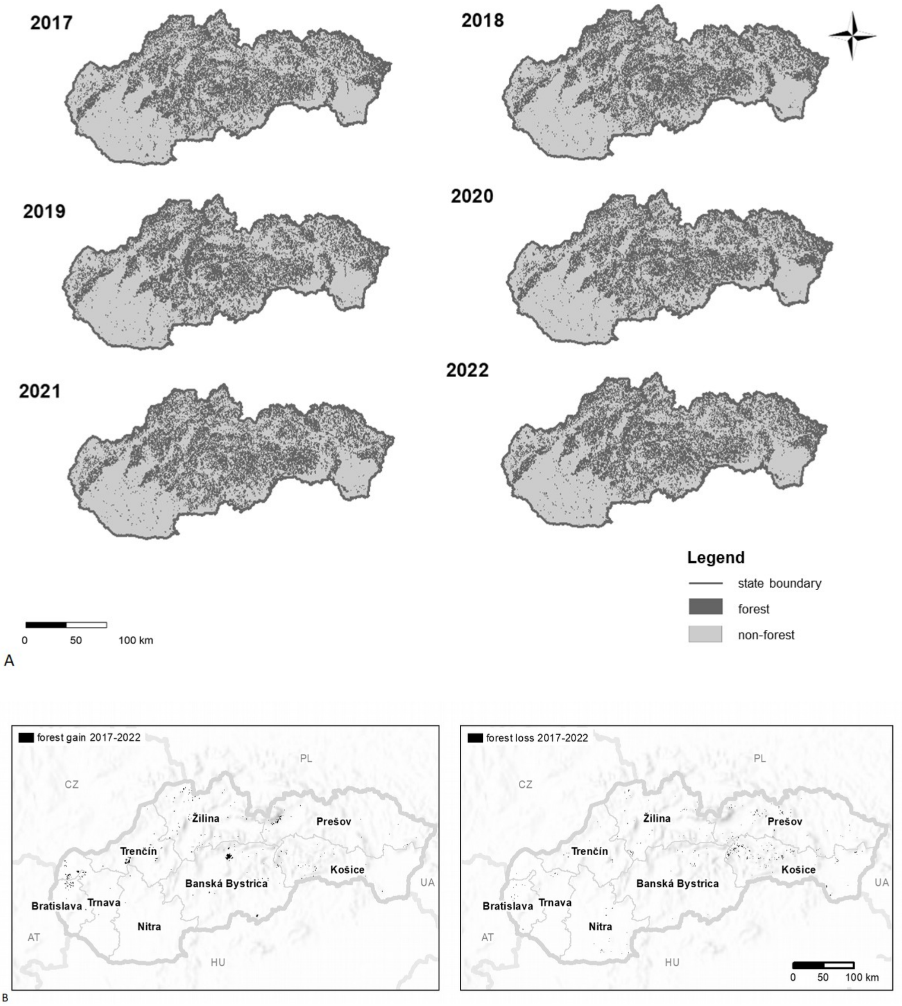

Figure 5 A) Change in forest cover in Slovakia according to RF classification during the period 2017-2022. B) The largest areas of forest loss and forest gain – difference raster for years 2017 and 2022. Source: processed by the authors

The reason for the different forest cover values between the RF classification results and SOSR data is primarily due to the fact that the official SOSR data actually provide information on forest land, not forest cover. As forest land, the SOSR define an area that is used for forest production that can also be temporarily deforested but restored and takes into account also areas that indirectly fulfill forestry functions, such as forest roads, warehouses and nurseries, as yet unforested land that was taken from the agricultural land fund and subsequently assigned to forestry (SOSR, 2022). The input of our classification was freely available satellite images that contrary to SOSR data, provides objective information on the actual extent of the forest cover. Naturally, even the results from RS data work with a certain degree of accuracy (95% in our case), which can be affected by snow and cloud cover, or other phenomena that prevent effective analysis. However, we also must draw attention to the spatial resolution of our results (30 m), due to which it is not possible to detect forest areas below 0.1 ha (900 m^2).

Table 4 Forest cover of Slovakia and its regions according to the RF classification and SOSR data

Region	Forest cover for the monitored year (%)											
	2017		2018		2019		2020		2021		2022	
	SOSR	RF	SOSR	RF	SOSR	RF	SOSR	RF	SOSR	RF	SOSR	RF
Bratislava	36.59	29.97	36.59	30.16	36.59	30.08	36.58	30.36	36.59	28.83	36.58	29.07
Trnava	15.79	11.85	15.77	12.48	15.77	12.46	15.77	12.45	15.77	12.24	15.77	12.67
Trenčín	49.95	46.36	49.43	45.73	49.45	47.10	49.46	46.98	49.46	46.71	49.46	46.72
Nitra	15.28	12.52	15.28	12.85	15.29	13.00	15.29	12.82	15.29	12.83	15.29	13.18
Žilina	56.36	49.09	56.39	46.91	56.34	46.01	56.35	46.16	56.39	45.69	56.41	45.32
Banská Bystrica	49.40	47.94	49.43	46.98	49.47	45.96	49.49	47.22	49.50	46.24	49.52	46.11
Prešov	49.53	47.02	49.65	46.51	49.74	46.42	49.77	47.60	49.80	44.56	49.82	47.00
Košice	39.85	37.53	39.87	35.50	39.88	35.03	39.92	35.92	39.93	35.79	39.92	37.00
SR	**41.33**	**36.91**	**41.32**	**37.28**	**41.34**	**37.79**	**41.36**	**37.36**	**41.37**	**37.62**	**41.38**	**37.86**

Source: RF data processed by the authors, SOSR (2017-2022)

Thanks to effective calculation of forest cover with the help of the developed tool, it is also possible to monitor the forest cover e.g. at the level of individual regions. Forest cover is expressed by the share of the area represented by forests in the total area of the area area under study – the region (NUTS 3). The district layer was added using the dataset FAO GAUL (2015) (note that the FAO GAUL boundary area layer may differ slightly from the district boundary layer used by SOSR).

The largest share of the total forest area in 2022 was concentrated in the Banská Bystrica (432,206.63 ha) and Prešov region (427,956.76 ha). We observed a higher than 40% share of forests in the regions of Banská Bystrica and Prešov for the entire 2017-2022 period. The highest forest coverage with 49.09% was observed in the Žilina region in 2017. Six out of nine national parks in Slovakia are located on the territory of these three regions. Throughout the analyzed period, the lowest forest cover was observed in the Trnava region in 2017 (11.85%). The highest forest area gain (by 4,269.48 ha) was observed between 2017 and 2022 in the Nitra region, while the largest decrease occurred in the Žilina region (by 25,888.38 ha). MŽP SR (2021) states that large deforestation occurred in this region (Žilina) between 2017-2020 with the greatest forest loss manifested in the national parks of Veľká Fatra and Malá Fatra. Additionally, the MŽP SR emphasized the high rate of forest loss in the area of Kysuce.

3.4 Comparison of classification results with the CLC database

To compare and highlight the potential of the developed forest research method, we compared the classification results with all the CLC classes containing information on forests (classes 311, 312, 313). These classes were then combined into a uniform mask for the forest, and the forest area and the percentage of forest cover in Slovakia was calculated. Since the CLC2018 database is mainly composed of 2017 satellite imagery and the availability of Sentinel-2 data in the GEE archive begins from 2017, we used the RF classification for the first available year, 2017, for comparison (Table 5).

Table 5 Comparison of the RF classification results with the CLC database

Data	Forest area (ha)	Forest cover (%)
RF (2017)	1,810,063.44	36.89
CLC (2018)	2,053,389.72	41.84
Difference	243,326.28	4.96

Source: RF data processed by the authors, CLC (2018)

Figure 6 illustrates the differences between the CLC data and the generated results of the RF classification for the forest area and forest cover. The difference in values can be mainly attributed to the resolution of the input data entering the classifications (MMU 25 ha for CLC2018 and 10×10 m, i.e. 0.01 ha for our results) and the accuracy of the maps (Figure 7).

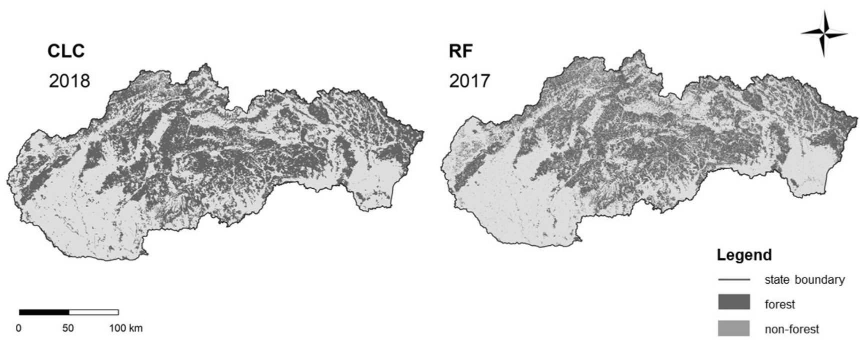

Figure 6 Forests of Slovakia according to the CLC2018 classification (mainly based on satellite data from 2017) and the classification using the RF algorithm. Source: CLC (2018), RF classification processed by the authors

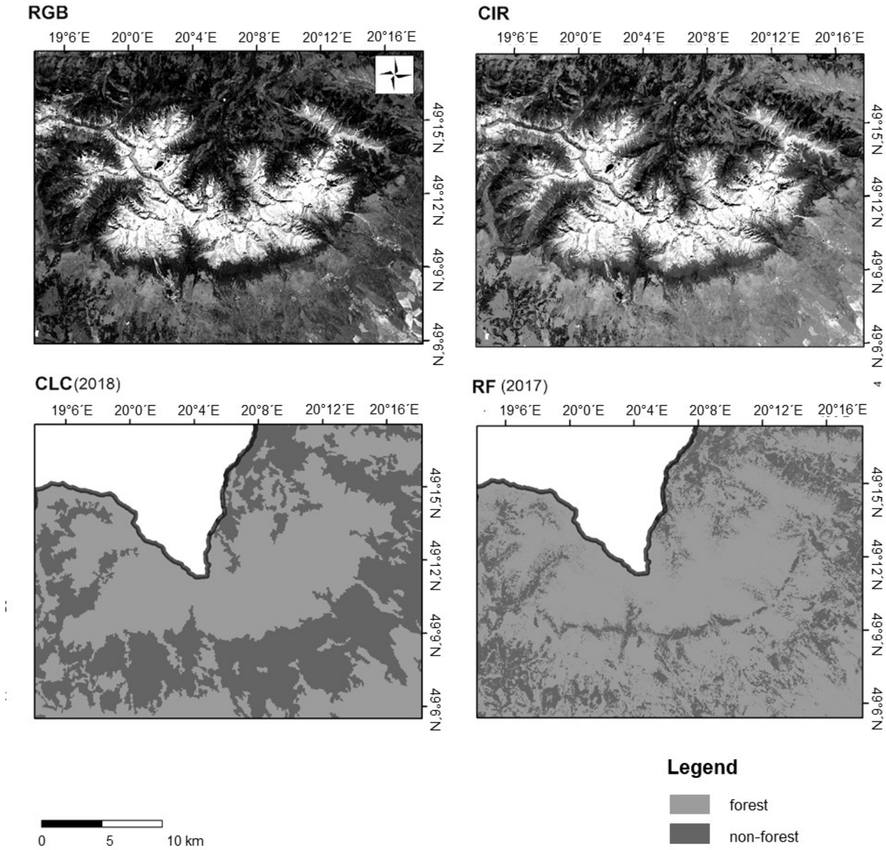

Figure 7 Forest area of RF classification compared to CLC2018. Source: CLC (2018), colour compositions of Sentinel-2 images from 28 May 2017 – processed by the authors

The created RF classification classified only areas that are covered by forest vegetation, while the CLC also listed areas that are not covered by forest vegetation as forests due to the high MMU. According to CLC data, the forest area was 243,326.28 ha larger than according to the RF classification. This is also related to a higher forest coverage according to CLC.

The CLC database provides sufficient information for the creation of training areas, but the spatial and temporal resolution of this inventory is insufficient to monitor forest status. With this comparison, we wanted to point out the potential of the image classification method used in the GEE cloud environment and the high performance of the RF algorithm and Sentinel-2 satellite images, which provided us with detailed information about the forest cover and could be used for detailed analyses.

3.5 A tool for automatic forest cover classification in GEE

In this paper, we also created a GEE tool that enables automatic forest cover classification using Sentinel-2 satellite imagery. A great advantage of this tool and the GEE platform is the possibility of changing the input parameters and modifying individual parts of the code by the users for their own needs. The tool was created using JavaScript using GEE's Code Editor interface and is available in the GitHub repository via the link:

https://github.com/palubad/Automatic-Forest-Classification-GEE

In GEE:

https://code.earthengine.google.com/c2f07a9161037480b5fbf8f11a6acaf2

In the first step, the user of the tool chooses the area of interest for automatic forest cover classification – the user can select any country or lower administrative unit in the EU from the LSIB 2017 database or create its own area in GEE. As training data are prepared using the CLC database, which only covers EU countries, the selection is limited to EU territory. In the next step, the year of analysis is selected. Currently, the years 2017-2022 can be selected due to the availability of data in the GFC database. The range of months for the input data for the median composite can also be specified. Another advantage of the created tool is the possibility to choose the upper limit of cloud cover for Sentinel-2 scenes and any number of training samples to be used in image classification. Advanced users can also modify other parts of the code, e.g. input bands, spectral indices (e.g. NDVI) and additional inputs to the classification (e.g. SRTM).

In the next parts, the methodology described in this paper is carried out, i.e., automatic creation of training dataset using the intersection of the CLC2018 and GLC databases, supervised classification using RF using the generated training data. The following layers are displayed in the map window: a median composite of the Sentinel-2 time series as an RGB composite, the intersection of the GLC and CLC2018 databases, and the resulting classification for the entire area of interest. The classified result can be exported as a so-called asset to GEE or downloaded as a GeoTiff to Google Disk via the "Tasks" tab on the right.

All the codes used in this work, including the comparison of the classification algorithms, the accuracy assessment and the validation points, are available in the mentioned GitHub repository.

3.6 Direction of further research

The following steps will focus on the thorough verification of the obtained outcomes and the implementation of the created technique beyond Slovakian borders. Application of sophisticated methods to mask out clouds in Sentinel-2 imagery, such as S2Cloudless (Zupanc, 2017) or Cloud Score+ (Pasquarella et al., 2023), would probably contribute to achieving even more accurate results. Another goal is to improve the method for automatic forest cover classification outside the EU by using global land cover databases, such as Copernicus Global Land Cover Layers (Buchhorn et al., 2020), ESA WorldCover (Zanaga et al., 2022), etc. It will be important to focus not only on the binary classification of forest/non-forest, but also on the classification of forest types. The high potential for forest monitoring is currently also in the use of freely available Sentinel-1 radar satellite data, which could be used in fusion with Sentinel-2 multispectral data.

4 CONCLUSIONS

Forests are important prerequisites for preserving biodiversity and protecting the climate, soil and water resources of every country. Preventive measures, such as monitoring forest conditions and early detection of factors disrupting forest cover, are necessary to protect forest ecosystems. Nowadays, the image classification using advanced RS methods, especially multispectral satellite data, offers a fast and effective way of obtaining information on the condition and changes in land cover.

The aim of this paper was to analyze changes in forest cover in Slovakia using methodology developed applying Sentinel-2 satellite data and ML algorithms in the GEE environment. In the article, we pointed out the potential of combining freely available Sentinel-2 imagery with other freely available data, such as SRTM, CLC2018 and the GFC database.

The main product of our work is the forest cover classification for Slovakia, while the developed tool for automatic classification of the forest cover can be easily applied for individual regions or districts of Slovakia, or modified for any other area within the EU. The ML algorithms RF and SVM were chosen for image classification, while the RF algorithm showed higher reliability with OA equal to 95% for both validation years. In a further analysis using the RF algorithm, we found that the most forest cover increased in Slovakia in the period 2018-2019 (by 24,831.68 ha). On the contrary, there was a 1.04% decrease in the amount of forest cover in 2019-2020. On a regional level, the highest levels of forest cover were observed in the Žilina, Banská Bystrica and Prešov region.

In comparison to official SOSR and CLC2018 data, the forest cover classification using ML algorithm RF and Sentinel-2 satellite data in GEE demonstrated high classification accuracy and efficiency in monitoring forest cover changes in high temporal and spatial resolution. We have made the mentioned procedure of automatic forest cover classification in the form of a GEE code freely available so that it can be applied to any other territory within Europe, from a local to global scale. In addition, the forest cover raster layers of Slovakia for 2017-2022 created in this work are also freely available in GEE.

Acknowledgements

This research was funded by the Slovak Academy of Sciences (VEGA) under the contract no. VEGA 1/0085/23 "Modeling urban heat islands using geospatial tools", and by the Charles University Grant Agency (GAUK): "Evaluation of forest disturbances and recovery using radar and optical satellite data in Czechia", nr. 412722.

References

ALKAMA, R., CESCATTI, A. 2016. Biophysical climate impacts of recent changes in global forest cover. *Science*, 351, 600-604.

BELGIU, M., DRĂGUŢ, L. 2016: Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24-31.

BUCHHORN, M., SMETS, B., BERTELS, L., ROO, B. D., LESIV, M., TSENDBAZAR, N. E., HEROLD, M., FRITZ, S. 2020. *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2017: Globe (V3.0.1) [Data set]*. Zenodo. https://doi.org/10.5281/zenodo.3518036

BÜTTNER, G., KOSZTRA, B., SOUKUP, T., SOUSA, A., LANGANKE, T. 2017. *CLC2018 Technical Guidelines*. Wien: European Environment Agency.

CONETTE, G., OSWALD, P., SOGNER, M., LEIMGRUBER, P. 2016: Mapping distinct forest types improves overall forest identification based on multi-spectral landsat imagery for myanmar's Tanintharyi region. *Remote Sensing*, 8, 882.

DRUSCH, M., BELLO, U. D., CARLIER, S., COLIN, O., FERNANDEZ, V., GASCON, F., HOERSCH, B., ISOLA, C., LABERINTI, P., MARTIMORT, P., MEYGRET, A., SPOTO, F., SY, O., MARCHESE, F., BARGELLINI, P. 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25-36.

EUROPEAN SPACE AGENCY, 2015. *Sentinel-2 User Handbook*. [online] [cit. 2023-12-06]. Available at: <https://sentinel.esa.int/documents/247904/685211/sentinel-2_user_handbook>

FAO, 2020. *Global Forest Resources Assessment 2020*. Rome: FAO.

FAO, 2022. *The State of the World's Forests 2022*. Rome: FAO.

FAO GAUL, 2015. *The Global Administrative Unit Layers (GAUL)*. Available at: <https://developers.google.com/earth-engine/datasets/catalog/DataLicenseGAUL2015.pdf>

FARR, T. G., ROSEN, P. A., CARO, E., CRIPPEN, R. et al. 2007. The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45, RG2004.

FOREST EUROPE, 2020. *State of Europe's Forests 2020*. Bratislava: Ministerial Conference on the Protection of Forests in Europe. [online] [cit. 2023-06-21]. Available at: <https://foresteurope.org/wp-content/uploads/2016/08/SoEF_2020.pdf>

GORELICK, N., HANCHER, M., DIXON, M., ILYUSHCHENKO, S., THAU, D., MOORE, R., 2017. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.*, 202, 18-27.

HANSEN, M. C., POTAPOV, P. V., MOORE, R., HANCHER, M., TURUBANOVA, S. A., TYUKAVINA, A., THAU, D., STEHMAN, S. V., GOETZ, S. J., LOVELAND, T. R., KOMMAREDDY, A., EGOROV, A., CHINI, L., JUSTICE, C. O., TOWNSHEND, J. R. G., 2023. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342, 850-853.

HANSEN, M., POTAPOV, P., MARGONO, B., STEHMAN, S., TURUBANOVA, S., TYUKAVINA, A., 2013. *User Notes for Version 1.6 of Global Forest Change 2000–2022*. [online] [cit. 2023-08-20]. Available at: <https://developers.google.com/earth-engine/datasets/catalog/UMD_hansen_global_forest_change_2022_v1_10>

HOŚCIŁO, A., LEWANDOWSKA, A., 2019. Mapping Forest Type and Tree Species on a Regional Scale Using Multi-Temporal Sentinel-2 Data. *Remote Sensing*, 11, 929.

IEP, 2017. *V národných parkoch bola strata lesa oproti iným územiam dvojnásobná.* Bratislava: Ministerstvo životného prostredia SR.

IMMITZER, M., NEUWIRTH, M., BÖCK, S., BRENNER, H., VUOLO, F., ATZBERGER, C., 2019. Optimal Input Features for Tree Species Classification in Central Europe Based on Multi-Temporal Sentinel-2 Data. *Remote Sensing*, 11, 2599.

IUCN, 2020. *Protected Areas*. [online] [cit. 2023-05-07]. Available at: <https://www.iucn.org/theme/protected-areas/about/protected-areas-categories/category-ii-national-park>
KHAN, M. A., SHAH, M. I., JAVED, M. F., KHAN, M. I., RASHEED, S., EL-SHORBAGY, M. A., 2022. Application of random forest for modelling of surface water salinity. *Ain Shams Engineering Journal*, 13, 101635.
KOSZTRA, B., BÜTTNER, G., HAZEU, G., ARNOLD, S. 2019. *European topic centre on urban, land and soil systems: Updated CLC illustrated nomenclature guidelines*. Kodaň: European Environment Agency.
LECHNER, M., DOSTÁLOVÁ, A., HOLLAUS, M., ATZBERGER, C., IMMITZER, M., 2022. Combination of Sentinel-1 and Sentinel-2 Data for Tree Species Classification in a Central European Biosphere Reserve. *Remote Sening*, 14, 2687.
LI, R., FANG, P., XU, W., WANG, L., OU, G., ZHANG, W., HUANG, X., 2022. Classifying Forest Types over a Mountainous Area in Southwest China with Landsat Data Composites and Multiple Environmental Factors. *Forests*, 13, 135.
LONG, J., SHELHAMER, E., DARNELL, T. 2015. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3431-3440.
LU, D., WENG, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28, 823-870.
MANTERO, P., MOSER, G., SERPICO, S. B., 2005. Partially Supervised classification of remote sensing images through SVM-based probability density estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 559-570.
MATTHEWS, R. W., ROBERTSON, K., MARLAND, G., MARLAND, E., 2007. Carbon in wood products and product substitution. *Forestry and climate change*, 91-14.
MINĎÁŠ, J., KONÔPKA, J., NOVOTNÝ, J., JENDEK, S., 2006. *Lesy Slovenska*. Zvolen: Národné lesnícke centrum.
MŽP, 2021. *Správa o stave životného prostredia Slovenskej republiky v roku 2021*. Bratislava: Ministerstvo životného prostredia Slovenskej republiky.
NOMURA, K., MITCHARD, E. T. A., 2018. More Than Meets the Eye: Using Sentinel-2 to Map Small Plantations in Complex Forest Landscapes. *Remote Sensing*, 10, 1693.
PAL, M., MATHER, P. M., 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 5, 1007-1011.
PALUBA, D., LAŠTOVIČKA, J., MOURATIDIS, A., ŠTYCH, P. 2021. Land Cover-Specific Local Incidence Angle Correction: A Method for Time-Series Analysis of Forest Ecosystems. *Remote Sensing*, 13, 1743.
PASQUARELLA, V. J., BROWN, C. F., CZERWINSKI, W., RUCKLIDGE, W. J., 2023. Comprehensive Quality Assessment of Optical Satellite Imagery Using Weakly Supervised Video Learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2124-2134.
PHALKE, A., OZDOGAN, M., THENKABAIL, P., ERICKSON, T., GORELICK, N., 2020. Mapping croplands of Europe, Middle East, Russia, and Central Asia using Landsat 30-m data, machine learning algorithms and Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 104-122.
ROUSE, J. W., HAAS, R. H., SCHELL, J. A., DEERING, D. W., 1974. Monitoring vegetation systems in the Great Plains with ERTS, In: Freden, S. C., Mercanti, E. P., Becker, M. (eds.) *Third Earth Resources Technology Satellite–1 Symposium. Volume I: Technical Presentations, NASA SP-351*. Washington, D. C.: NASA, 309-317.
SVOBODA, J., ŠTYCH, P., LAŠTOVIČKA, J., PALUBA, D., KOBLIUK, N., 2022. Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data – A Case Study of Czechia. *Remote Sens*., 14, 1189.
STATISTIC OFFICE SR 2022. [online] [cit. 2022-03-08]. Available at: <https://slovak.statistics.sk>

ŠEBEŇ, V., 2017. Národná inventarizácia a monitoring lesov Slovenskej republiky 2015-2016. *Lesnícke štúdie*, 65, 43.
TANG, X., BULLOCK, E. L., OLOFSSON, P., ESTEL, S., WOODCOCK, C. E., 2019. Near real-time monitoring of tropical forest disturbance: New algorithms and as-sessment framework. *Remote Sensing of Environment*, 224, 202-218.
VIANA, C. M., GIRÃO, I., ROCHA, J., 2019. Long-Term Satellite Image Time-Series for Land Use/Land Cover Change Detection Using Refined Open Source Data in a Rural Region. *Remote Sensing*, 11, 1104.
WAŚNIEWSKI, A., HOŚCIŁO, A., ZAGAJEWSKI, B., MOUKÉTOU-TARAZEWICZ, D., 2020. Assessment of Sentinel-2 Satellite Images and Random Forest Classifier for Rainforest Mapping in Gabon. *Forests*, 11, 941.
YANG, L., DRISCOL, J., SARIGAI, S., WU, Q., CHEN, H., LIPPITT, C. D., 2022. Google Earth Engine and Artificial Intelligence (AI): A Comprehensive Review. *Remote Sens.*, 14, 3253.
ZANAGA, D., VAN DE KERCHOVE, R., DAEMS, D., DE KEERSMAECKER, W., BROCKMANN, C., KIRCHES, G., WEVERS, J., CARTUS, O., SANTORO, M., FRITZ, S., LESIV, M., HEROLD, M., TSENDBAZAR, N.-E., XU, P., RAMOINO, F., ARINO, O., 2022. *ESA WorldCover 10 m 2021 v200 (Version v200) [Data set]*. Zenodo. https://doi.org/10.5281/zenodo.7254221
ZUPANC, A., 2017. *Improving Cloud Detection with Machine Learning*. [online] [cit. 2023-10-25]. Available at: <https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13>

Automatická klasifikácia lesnej pokrývky pomocou multispektrálnych satelitných dát Sentinel-2 a algoritmov strojového učenia v Google Earth Engine

Súhrn

S rastúcimi hrozbami a tlakom na lesné zdroje v posledných desaťročiach začína dochádzať k strate a degradácii lesov. Úbytok lesov je celosvetovým fenoménom a jedným z kľúčových faktorov globálnej zmeny klímy. V súčasnosti sa preto v regionálnom, ale aj celosvetovom meradle, zvyšujú požiadavky na získavanie včasnejších a presnejších informácií o stave lesa, jeho fungovaní a udržateľnosti, pričom na monitorovanie časopriestorových zmien lesa sa využívajú v posledných desaťročiach aj metódy diaľkového prieskumu Zeme.

Z hľadiska lesnatosti sa Slovenská republika radí medzi popredné krajiny v Európe. Avšak aj lesy na tomto území čelia v posledných rokoch rastúcemu riziku škodlivých činiteľov. K poklesu lesnatosti dochádza vplyvom človeka, ktorý premieňa les na poľnohospodárske plochy za účelom hospodárstva, ale aj vplyvom biotických a abiotických faktorov, akými sú extrémy počasia a veterné kalamity.

Cieľom tohto článku bolo vyvinúť automatizovaný postup klasifikácie lesnej pokrývky pomocou snímok Sentinel-2 a klasifikačných algoritmov Random Forest (RF) a Support Vector Machine (SVM) v prostredí Google Earth Engine (GEE). Pre monitorovanie zmien lesa na území Slovenska a v jeho krajoch v časovom období rokov 2017-2022 bol vytvorený a aplikovaný skript. Ako vstupné údaje boli použité vybrané pásma družice Sentinel-2, normalizovaný diferenčný vegetačný index NDVI a digitálny výškový model SRTM. Na vytvorenie validačných trénovacích vzoriek bola použitá kombinácia databáz Corine Land Cover (CLC) a Global Forest Change (GFC). Následne boli porovnané výsledky klasifikácie s použitím klasifikačných algoritmov riadenej klasifikácie RF a SVM. Pomocou vytvoreného

skriptu sa vyhodnotila správnosť klasifikácií a vyhodnotili sa zmeny stavu lesnej pokrývky na vybranom území, ktoré boli porovnané aj s oficiálnymi údajmi získanými zo Štatistického úradu SR (ŠÚ SR) a údajmi CLC o výmere lesov.

Hlavným produktom sú klasifikácie lesnej pokrývky SR, pričom vytvorený nástroj – kód – je možné jednoducho aplikovať na klasifikáciu lesnej pokrývky pre jednotlivé kraje alebo okresy SR, a tiež modifikovať pre akékoľvek iné záujmové územie v rámci EÚ. Algoritmus RF vykazoval vyššiu spoľahlivosť s celkovou presnosťou 95 % pre oba validačné roky (2017 a 2020), v ktorých sa správal približne rovnako stabilne. Pri ďalších analýzach pomocou RF algoritmu sme zistili, že najviac lesnej pokrývky pribudlo na území SR v období rokov 2018 až 2019. Výmera lesa sa za toto obdobie zvýšila o 24 831,68 ha. Naopak, úbytok lesa bolo možné pozorovať medzi rokmi 2019 a 2020, kedy sa stav lesa znížil o 1,04 %. Na regionálnej úrovni bola v sledovanom období najvyššia lesnatosť pozorovaná v Žilinskom, Banskobystrickom a Prešovskom kraji.

Vykonaná klasifikácia lesnej pokrývky pomocou algoritmu strojového učenia RF a satelitných dát Sentinel-2 v GEE preukázala vysokú presnosť klasifikácie a efektivitu pri sledovaní zmien lesnej pokrývky na území Slovenska vo vysokom časovom a priestorovom rozlíšení v porovnaní s oficiálnymi údajmi ŠÚ SR a CLC. Uvedený postup automatickej klasifikácie lesnej pokrývky a samotné klasifikované rastrové vrstvy lesnej pokrývky Slovenska pre roky 2017 až 2022 sú dostupné vo forme voľne dostupných kódov pre GEE.