



OPEN **Harnessing remote sensing and machine learning techniques for detecting and monitoring the invasion of goldenrod invasive species**

Radek Malinowski^{1✉}, Michał Krupiński², Piotr Skórka³, Łukasz Mikołajczyk³, Karolina Chuda³ & Magdalena Lenda³

Invasive alien species, such as goldenrods (*Solidago* spp.), pose significant threats to biodiversity and ecosystem services across Europe. Effective monitoring of these species is essential for early intervention and informed management, yet traditional ground surveys are often labor-intensive and limited in scale. This study aims to evaluate the potential of remote sensing and machine learning for detecting and monitoring *Solidago* spp. in Kampinos National Park, Poland, using multitemporal imagery from Sentinel-2 and PlanetScope satellites. We compared the performance of Random Forest and One-Class Support Vector Machine classifiers across 17 classification scenarios incorporating spectral bands, vegetation indices, and temporal statistics. Our results showed that Random Forest consistently outperformed One-Class Support Vector Machine (OCSVM) by 1%–15%, achieving the highest F1-score of 0.98 using multitemporal Sentinel-2 data and 2%–29% using PlanetScope imagery. Sentinel-2 data, with its broader spectral range, provided better large-scale detection accuracy, while PlanetScope's higher spatial resolution enhanced local detail. Goldenrod patches are distinctive even in autumn and winter due to living or dry biomass that persists the whole year. In our study autumn imagery (October–November) yielded the most reliable detection due to distinct phenological characteristics of *Solidago* during this period. Importantly, our analysis demonstrates that the added complexity of vegetation indices does not necessarily improve classification accuracy for goldenrod detection. Our findings present high-accuracy invasive species monitoring approach and highlight the critical role of phenological timing in remote sensing-based ecological assessments.

Keywords Invasive alien plants, *Solidago*, PlanetScope, Machine learning, Sentinel-2, Random forest, OCSVM

Invasive alien species (IAS) pose a significant threat to biodiversity in the Anthropocene, alongside agriculture, deforestation, habitat fragmentation, and climate change¹. These species can cause extensive habitat transformation and ecosystem disruption, negatively affecting both natural and human-modified environments². IAS often outcompete native flora and fauna, reduce biodiversity, and alter ecosystem processes such as nutrient cycling and energy flows^{3,4}. One such invasive genus is *Solidago* (goldenrods), particularly *Solidago canadensis* and *Solidago gigantea*, which have become widespread across Europe due to their aggressive dispersal strategies, prolific seed production, and vegetative reproduction^{5–7}. Goldenrods severely impact native biodiversity and ecosystem services. Their dense monocultures displace native plant communities, degrade habitats for pollinators, and reduce agricultural productivity^{5,6,8}. In Poland and Central Europe, they now dominate large portions of post-agricultural land, particularly abandoned fields, where they may occupy up to 70% of the area^{6,9}. Despite being considered invasive in 49 countries, goldenrods are still cultivated by beekeepers for honey production, complicating efforts to control their spread^{2,10}. In Kampinos National Park (KNP) in Poland, the invasion of goldenrods presents a pressing challenge for conservation and land management, particularly in areas recovering from agricultural

¹Wasat Ltd., Warsaw, Poland. ²Space Research Centre, Polish Academy of Sciences, Warsaw, Poland. ³Institute of Nature Conservation, Polish Academy of Sciences, Kraków, Poland. ✉email: radekmali@yahoo.com

abandonment¹¹. Early detection and accurate monitoring are essential for targeted management of these species before they become widespread and difficult to control.

Remote sensing has emerged as a powerful tool for ecological monitoring and invasive species detection, offering efficient, repeatable, and large-scale observation capabilities^{8,9}. It is important so that botanists working on the goldenrod ecology and eradication, as well as practitioners who should manage invasion, plan actions and make decisions are aware that the goldenrod can be easily detected from satellite images. This method is much less time consuming and more cost-effective, comparing to in-situ inventory.

Both passive (e.g., multispectral and hyperspectral) and active (e.g., radar, LiDAR) remote sensing systems have been used to monitor invasive plant species. Among them, multispectral sensors have proven especially useful due to their sensitivity to vegetation spectral characteristics. However, mapping species of the *Solidago* genus using remote sensing has been addressed in only a few studies, most of which have relied exclusively on optical data from the visible and reflective infrared spectra.

The detection of *Solidago altissima* was studied by Ishii and Washitani¹⁰ as well as Rizaludin Mahmud et al.¹¹. In the former case, the authors applied generalized linear regression models to high-resolution airborne hyperspectral data collected in spring, achieving a prediction accuracy of 85% for goldenrod mapping. In the latter study, detailed field spectroscopy data acquired during the flowering period enabled the development of a goldenrod-specific vegetation index. This index was then applied to a WorldView-2 image from autumn, with classification performed using a maximum likelihood classifier and thresholding. Notably, both studies relied on single-date remote sensing imagery. Koco et al.¹² and Zagajewski et al.¹³ mapped occurrences of *Solidago gigantea* and *Solidago canadensis*. Koco et al. used a GNSS system and unmanned aerial vehicle (UAV) imagery to assess goldenrod spread over four consecutive years. Although precise plant positioning was achieved, automated classification of UAV multispectral images resulted in very low accuracy (<50%). In contrast, Zagajewski et al. carried out a more advanced study using multitemporal Sentinel-2 and PlanetScope data representing late summer and autumn. The classification, performed with Random Forest (RF) and support vector machine (SVM) algorithms, yielded high overall accuracies (mostly above 90%), with the best results obtained using Sentinel-2 and RF. Sentinel-2 imagery was also used by Chadoulis et al.¹⁴ to map *Solidago gigantea* in a riparian environment. In this case, classification was based on phenological metrics derived from multitemporal data, using RF, and achieved approximately 80% accuracy. Additional studies attempted to distinguish goldenrods from other invasive or expansive species using airborne hyperspectral data in combination with RF, SVM, or both algorithms^{15–17}. These approaches achieved varying mapping accuracies depending on the composition of the training dataset (specifically, the proportion of the target class within samples) and its size. In several cases, high accuracies above 90% were reported when sufficient training samples were available.

Collectively, these studies confirm the high potential of remote sensing for accurate identification and classification of goldenrod species. However, many of the cited works relied on airborne data, which presents scalability challenges for broad-area mapping. Additionally, the research areas were often limited in spatial extent, which may hinder scalability of their findings. Satellite-based approaches offer a more practical solution for large-scale, operational monitoring. Among the available sensors, Sentinel-2 and PlanetScope stand out due to their high-resolution multispectral capabilities, making automated detection of goldenrod invasions more feasible than ever. Sentinel-2, as part of the Copernicus Programme, provides free, high-resolution (10–60 m) multispectral data with frequent revisit cycles and wide swath coverage, making it suitable for large-scale ecological assessments¹². PlanetScope, a commercial constellation, offers near-daily data at approximately 3-meter spatial resolution, allowing for detailed, site-specific analysis¹³. The complementary characteristics of these systems, spatial resolution, temporal frequency, and availability, make them highly promising for detecting invasive species across heterogeneous landscapes. Importantly, both provide regular and systematic acquisitions that enable temporal analysis of plant phenology, which is particularly relevant in the case of goldenrods.

Despite these opportunities, few studies have performed comparative assessment of Sentinel-2 and PlanetScope use for goldenrod detection or examined how different machine learning algorithms perform on multitemporal datasets. Furthermore, many investigations prioritize algorithmic or sensor comparisons rather than integrating phenological information and considering operational needs for invasive species management.

The aim of this study is twofold. First, we address this research gap by evaluating the performance of two machine learning algorithms, RF and one-class support vector machine (OCSVM), for detecting *Solidago* spp. using Sentinel-2 and PlanetScope imagery. RF is widely recognized for its robustness to overfitting and its ability to process high-dimensional datasets^{18,19} thus it is used in our study as a benchmark method. OCSVM is particularly effective when training data are available for only one class, a common limitation in IAS mapping due to difficulties in obtaining reliable ground truth for different land cover types^{20,21}. Moreover, OCSVM when tested for defect predictions, even if did not outperform RF, it proved to be superior comparing to binary classifiers²². Second, this study investigates the role of phenological timing in classification performance. Goldenrod patches often remain visible even into autumn and winter due to their tall, persistent biomass, which can be green or dry. Therefore, we assess detection performance across the entire growing season, from early spring to late autumn, using multitemporal imagery. Our overarching objective is to evaluate how spatial and temporal resolution, classification strategy, and image timing affect mapping accuracy, with a focus on practical implications for goldenrod monitoring and control.

By focusing on *Solidago* invasion in an ecologically diverse and well-documented site in Poland, this study contributes to the development of scalable, remote sensing-based monitoring frameworks. It also provides valuable insights for optimizing the choice of satellite data, phenological windows, and classification approaches for improved ecological management and conservation planning.

Materials

Study taxon and area

Canadian goldenrods (*Solidago canadensis* L.) and giant goldenrods (*S. gigantea* Aiton) are two highly invasive plant species native to North America that have successfully colonized vast areas of European land²³. Due to their similar morphology and habitat preferences within their secondary range²⁴, we considered these species together in this study and referred to them as goldenrods²⁵. In Central Europe, goldenrods are known to form dense single-species patches that sometimes cover hundreds of hectares^{5,6} (Fig. 1). Despite their relatively late emerging in spring, goldenrods significantly impact on vegetation throughout the growing season due to their rapid growth, high biomass accumulation, substantial litter production, and shading potential, reaching heights of up to 2 meters^{5,6,26}. Additionally, goldenrods release allelopathic chemicals that inhibit the growth of other plant species^{27–29}. Goldenrod shoot density typically ranges from 150 to 200 per square meter, with some exceeding 300 shoots per square meter³⁰. A single goldenrod stem can produce over 10,000 seeds^{31,32} which exhibit high vigor and maintain viability even under diverse temperature and moisture conditions³¹. Approximately 50% of goldenrod seeds successfully germinate³³. After initial seed establishment, goldenrods primarily reproduce vegetatively through rhizomes, forming dense clonal patches that comprise up to 90% of the total plant cover in abandoned fields^{5,30}.

Our study on goldenrod mapping was conducted in the Mazovia Province in central Poland (Fig. 2), specifically in the area west of Warsaw, encompassing the entire KNP and its surrounding neighborhood (SW corner: 52.26°N, 20.28°E; NE corner: 52.42°N, 20.92°E). The study area spans 18 km × 43 km, covering a total area of 774 km². These areas included a variety of land use and land cover (LULC) types, predominantly vegetated areas, such as forests, grasslands, croplands, and settlements. The site was chosen because it represents a mixture of natural ecosystems including protected areas, human-altered environments, and transitional areas.

Over the past centuries the KPN area has been significantly influenced by human activity. Originally dominated by forests, the land was deforested, drained, and finally converted into farmlands and meadows³⁴. Kampinos National Park was established in 1959 to protect its unique forest, wetland, and dune ecosystems, and its species diversity and historical heritage. Mandatory land acquisition began in 1975, leading to the abandonment of extensive agricultural activities in the area. This shift allowed ecological processes, such as secondary succession to take place, which facilitated the encroachment of invasive species onto abandoned rural areas. Goldenrod have notably invaded these fallow and abandoned lands³⁴. Among other types of nature conservation, the KNP area is also designated as a Special Area of Conservation and Special Protection Area - Puszcza Kampinoska (PLC140001) – under the Natura 2000 network, as part of the Bird and Habitats Directives.

Satellite imagery

Our research explored the feasibility of recognizing *Solidago* spp. based on two sets of multispectral imagery from the Sentinel-2 and PlanetScope constellations. Sentinel-2 offers multispectral data with spatial resolutions ranging from 10 m to 60 m, covering visible to short-infrared wavelengths (Table 1)³⁵. An additional advantage of the Sentinel-2 is its ability to provide multitemporal observations, enhancing the likelihood of obtaining cloud-free images. In mid-latitude regions of Europe, the revisit period ranges from two to three days. Sentinel-2 images downloaded from CREODIAS³⁶ represent the year 2022, corresponding to the reference period for the in-situ data collection. For our study, cloud-free satellite images were selected from the Level-2 A product, which provides the bottom-of-atmosphere reflectance data. The selected image data covered several months of the growing season, including spring, summer, and autumn (Table 2). For images with limited, sporadic cloud cover,



Fig. 1. A vast meadow colonized densely by goldenrod.

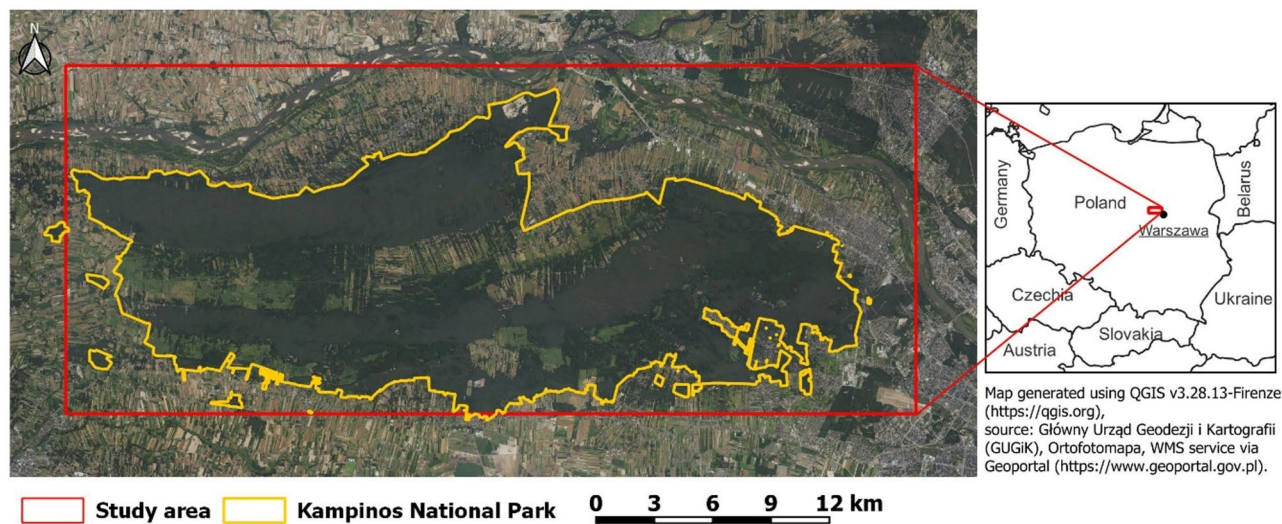


Fig. 2. Illustration of the analyzed area covering the Kampinos National Park and nearby area in the north-west of Warsaw. Orthophotomap from Geoportal WMS service is used as a background in the map.

Sentinel-2				PlanetScope		
Band	Central wavelength (nm)	Bandwidth (nm)	Spatial resolution (m)	Band	Central wavelength (nm)	Bandwidth (nm)
B1	442.7	21	60	B1	443	20
B2	492.4	66	10	B2	490	50
B3	559.8	36	10	B3	531	36
B4	664.6	31	10	B4	563	36
B5	704.1	15	20	B5	610	20
B6	740.5	15	20	B6	665	31
B7	782.8	20	20	B7	705	15
B8	832.8	106	20	B8	865	40
B8a	864.7	21	20			
B9	945.1	20	60			
B10	1373.5	31	60			
B11	1613.7	91	20			
B12	2202.4	175	20			

Table 1. Spectral characteristics of Sentinal-2 and planetscope data. The italics bands indicate those that are interoperable once between the two sensing systems.

Sentinel-2	Percentage of replaced pixels	PlanetScope
2022.03.20	0.0	2022.03.19
2022.06.11 (2022.06.26)	0.6	2022.06.11
2022.08.05 (2022.07.23)	0.2	2022.08.05
2022.09.06 (2022.09.26)	0.2	2022.09.05
2022.11.03 (2022.11.08)	1.1	2022.10.10

Table 2. Acquisition dates (YYYY.MM.DD) of satellite image used in the study. Dates in parentheses indicate additional data used to replace cloudy pixels.

cloudy pixels were replaced with non-cloudy pixels from the nearest available dates (Table 2). Spectral bands of varying resolutions were resampled to a 10 m pixel size using the nearest-neighbor method.

PlanetScope is a satellite constellation developed and operated by Planet Labs, Inc³⁷. The constellation comprises 130 small CubeSat 3U satellites, a number that continues to grow³⁷. It is the only currently operating constellation that enables the daily acquisition of high-resolution satellite imagery for the entire world’s land surface. The latest third generation PlanetScope satellite, SuperDove, acquires multispectral data across eight

BDOT layer name	Class characteristics	Represented LULC class
BUBD	Building outlines	Buildings
KUKO	Transportation hub (only a sub-class of parking spaces used)	Artificial surfaces
PTKM	Transportation network (roads, railways)	Artificial surfaces
PTLZ	Forests, wooded areas	Broadleaf and coniferous tree cover
PTTR	Grasslands	Natural and agricultural grasslands
PTTR	Croplands	Croplands
PTWP	Surface waters	Water bodies

Table 3. Description of land cover and land use (LULC) classes of interest from the topographic objects database (BDOT) database used for Preparation of reference data.

Land cover class	Number of validation samples
Solidago	500
Artificial surfaces	100
Broadleaf trees	1300
Buildings	50
Coniferous trees	1500
Croplands	650
Grasslands	1000
Water bodies	130

Table 4. Analyzed LULC classes and number or validation samples.

spectral bands, six of which align with Sentinel-2 channels (Table 1). PlanetScope images selected for this study represent cloud-free bottom-of-atmosphere surface reflectance (Level-3B) at a spatial resolution of three meters. The PlanetScope data acquisition dates correspond to those of Sentinel-2, except for autumn data acquired in October for PlanetScope and November for Sentinel satellites (Table 2).

Reference data

Accurate reference data is essential for training and validating remote sensing-based classification models, particularly when mapping specific plant taxa such as invasive *Solidago* spp. In this study, we developed a reference dataset using two main sources: the national Topographic Objects Database (Baza Danych Obiektów Topograficznych, BDOT)³⁸ and in-situ field surveys conducted during the 2022 growing season (August–November). We used BDOT polygons to identify and delineate non-*Solidago* LULC classes including buildings, roads, grasslands, croplands, forest stands, and water bodies (Table 3). These polygons were additionally visually verified and confirmed using high-resolution orthophotos and PlanetScope imagery from the same period. To account for possible temporal mismatches and classification ambiguity, grasslands and croplands were additionally confirmed via field visits. Reference data for *Solidago* spp. were collected only during field surveys when polygons with full cover of *Solidago* spp. were recorded in SW Maps software for smartphones (SW Maps, version 2.9.1.1 (06.07.2022), <https://aviyaantech.com/swmaps/index.html>). The reference polygons for all analyzed LULC classes were collected to represent homogenous cover with the minimum area of at least 20 m x 20 m, ensuring that each polygon could be accurately represented by a single pixel in both the PlanetScope and Sentinel-2 images. These reference polygons were designed to provide high-quality samples for training machine learning classifiers and validating classification results. The homogeneity of the cover within each sample (pixels) allowed for use of the same sets of samples for the classification of both satellite datasets.

Next, all polygons (obtained from BDOT and from field surveys) were overlaid with Sentinel-2 pixels grid (10 m x 10 m). Final sets of reference polygons were generated by selecting only those pixels, that were covered in 100% by a polygon of a certain LULC class.

To minimize the potential influence of autocorrelation on classification and accuracy assessment, reference polygons for all classes were divided into two groups: training and validation. This division ensured an unbiased accuracy assessment by preventing the use of samples from a reference polygon dedicated to training the classifiers in the validation process. Finally, to prepare the sets of training and validation samples, stratified random sampling method was applied within the reference polygons. An area representing a polygon of 10 m x 10 m corresponding to the Sentinel-2 pixel was set as a single sample. Details regarding the number of samples used during the training phase are provided in Section “Random forest” for the Random Forest algorithm and in Section “One-class classifier”, for the OCSVM algorithm.

For the validation samples the proportion was estimated from the BDOT dataset, with the assumption that the *Solidago* spp. class constituted one-third of the grassland class. Due to the limited number of *Solidago* polygons, the number of *Solidago* samples was set at 500, and the sample number for the remaining classes were adjusted accordingly, based on their class coverage in the BDOT database. The estimated number of validation samples for each LULC class is listed in Table 4.

Index name	Acronym	Equation	Sentinel-2 bands used	PlanetScope bands used	References
Green normalized difference vegetation index	GNDVI	$(\text{NIR} - \text{Green})/(\text{NIR} + \text{Green})$	B08, B03	B08, B04	40
Merris terrestrial chlorophyll index	MTCI	$(\text{NIR} - \text{Red-Edge})/(\text{Red-Edge} - \text{Red})$	B08, B05, B04	B08, B07, B06	41
Normalized differential red-edge index	NDRE	$(\text{NIR} - \text{Red-Edge})/(\text{NIR} + \text{Red-Edge})$	B08, B05	B08, B07	42
Normalized difference vegetation index	NDVI	$(\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})$	B08, B04	B08, B06	43
Optimized soil adjusted vegetation index	OSAVI	$(\text{NIR} - \text{RED})/(\text{NIR} + \text{RED} + 0.16)$	B08, B04	B08, B06	44
Soil adjusted vegetation index	SAVI	$(\text{NIR} - \text{RED})/(\text{NIR} + \text{RED} + 0.5)$	B08, B04	B08, B06	45
Simple ratio	SR	RED/NIR	B08, B04	B08, B06	46
Vegetation index	VI	NIR/RED	B08, B04	B08, B06	47

Table 5. Vegetation indices used in the classification of *Solidago* spp. Including their equations and information of the spectral bands utilized. NIR – near infrared.

Scenario	Features	Number of Sentinel-2 features	Number of Planet Scope features
1	Image from March	10	8
2	Image from March + VIs	10 + 8	8 + 8
3	Image from June	10	8
4	Image from June + VIs	10 + 8	8 + 8
5	Image from August	10	8
6	Image from August + VIs	10 + 8	8 + 8
7	Image from September	10	8
8	Image from September + VIs	10 + 8	8 + 8
9	Image from October/November	10	8
10	Image from October/November + VIs	10 + 8	8 + 8
11	Spectral bands from 5 dates	5 × 10	5 × 8
12	Spectral bands from 5 dates + Vis	5 × 10 + 5 × 8	5 × 8 + 5 × 8
13	VIs from 5 dates	5 × 8	5 × 8
14	Statistics from 5 dates	5 × 10	5 × 8
15	Statistics of VIs	8 × 5	8 × 5
16	Statistics from 5 dates + statistics of VIs	5 × 10 + 8 × 5	5 × 8 + 8 × 5
17	All possible features: spectral bands from 5 dates + VIs + statistics from 5 dates + statistics from VIs	5 × 10 + 5 × 8 + 5 × 10 + 8 × 5	5 × 8 + 5 × 8 + 5 × 8 + 8 × 5

Table 6. Overview of classification feature combinations used in various classification scenarios. VIs – vegetation indices.

Methods
Classification procedure

Two machine learning algorithms were tested for the identification of *Solidago* spp.: RF and OCSVM. The characteristics of these algorithms, their usage descriptions, and the applied parameters are provided in the following subsections.

For both algorithms and image data sources, various combinations of classification features were utilized. This included the classification of spectral bands from single and multiple images within the prepared time series, as well as vegetation indices and statistics derived from these features. The selected indices were among the most commonly used, facilitating the analysis of different vegetation characteristics³⁹ (Table 5).

- Summing up, four types of features have been usedSpectral reflectance data from ten bands of Sentinel-2 and eight bands of PlanetScope. Classifications on Sentinel-2 were performed only on bands with 10 and 20 m spatial resolutions (strategy 1), excluding the 60 m spatial resolution bands due to their limited usefulness for land surface analyzes^{48,49}. The exception was the direct comparison of Sentinel-2 and PlanetScope data, where all corresponding bands, including band 1 (coastal blue) from Sentinel-2 (strategy 2) were used. PlanetScope imagery was analyzed with native 3 m spatial resolution.
- Vegetation indices;
- Multitemporal statistical parameters from spectral bands: minimum, maximum, mean, median, and standard deviation values from five dates.
- Multitemporal statistical parameters from vegetation indices: minimum, maximum, mean, median, and standard deviation values from five dates.

Different combinations of these features were used to design classifications for 17 scenarios for each satellite dataset, as summarized in Table 6.

Finally, to directly compare the satellite datasets, an additional test was performed (strategy 2). This test involved classifying *Solidago* spp. using the RF classifier on the interoperable spectral bands of Sentinel-2 and PlanetScope (six bands). For each of the five data acquisition dates, classification was performed on the equivalent bands from both sensing systems. Bands 1 and 2 correspond between both sensors. Band 3 from Sentinel-2 corresponds to band 4 from PlanetScope and band 4 to band 6. Finally, band 5 and 8a of Sentinel-2 sensors are equivalent to bands 7 and 8 of PlanetScope data, respectively (Table 1).

Random forest

RF is a well-tested and frequently used machine learning algorithm for various classification tasks, particularly in EO analyses^{18,19,50–52}.

For the RF classification, two main parameters were set. The first was the number of decision trees that built the forest (500). Such high number of trees may be time consuming but ensures the repeatability of results in the context of accuracy performance⁵¹. the second parameter determined the number of features selected for each split in the tree. The square root of the number of input features was selected to optimize the classification computation. Based on previous research indicating improved performance of the RF classifier with a larger number of training samples^{13,16} we used 1000 samples per class for all RF classification tests. These samples were randomly selected from the previously prepared reference datasets and represented eight different LULC classes.

One-class classifier

The second classifier used was the OCSVM algorithm²⁰ a machine-learning technique for anomaly detection. The OCSVM models a one-class problem, to identify unusual and infrequent occurrences in a dataset. It constructs a hyperplane to maximize the margin around the majority data points, effectively isolating potential outliers or anomalies. This method is particularly useful for unbalanced datasets with a large proportion of normal cases relative to anomalies. The OCSVM has a variety of applications, including fraud detection, network security, and quality control, due to its ability to classify instances as either normal or abnormal. In satellite data classification, OCSVM has been successfully applied, for example, for building detection with Sentinel-2 images²¹ or sea ice detection on Sentinel-1 SAR images⁵³. Unlike multiclass classifiers such as RF, it does not require sample collection for classes of no interest. For instance, in the current tests using RF, samples from seven additional classes were used.

The performance of the OCSVM classifier can be optimized by adjusting two specific parameters, γ (gamma) and ν (nu). γ indicates the kernel width, that is, the Gaussian radial basis function while ν describes the fraction of outliers in the training set. To determine the optimal combination of both parameter values, a grid search was conducted as applied by Krupinski et al.²¹. This search tested 15 γ values and 10 ν parameter values, resulting in a total of 150 combinations. The applied parameter values are listed in Table 7.

In all 17 scenarios (Table 6), *Solidago* was classified using each possible combination of γ and ν parameters. The same procedure was applied for Sentinel-2 and PlanetScope imagery.

Each classification process was divided into two steps. First, parameter tuning was conducted using a training set (500 *Solidago* samples only) and a testing set (7500 samples: 500 *Solidago* and 7000 non-*Solidago*) for each of the 150 parameter combinations within each scenario. This step aimed to identify the optimal parameter combination.

Figure 3 illustrates the parameter tuning procedure for one of the 17 scenarios, showing 150 OCSVM classifications in the form of an accuracy board. Each rectangle on the board displays accuracy metrics for a combination of OCSVM parameters. Values in red indicate the F1-score of the *Solidago* class, and values in blue indicate the producer's (top of rectangle) and user's (bottom of rectangle) accuracy. The green line marks the best result for that scenario. The board also included values of the γ and ν parameters used on X and Y axis.

The second step is an accuracy assessment based on a validation set. Although the training of the OCSVM classifier required only samples from one class, non-*Solidago* samples were also required to perform the complex accuracy assessment. In this phase 5230 samples (500 from *Solidago* and 4730 from non-*Solidago*) were used, as presented in Table 4.

Results

Comparison of classification results

The accuracy of *Solidago* spp. classification was assessed using a standard remote sensing method for computing the accuracy matrix and deriving accuracy measures, including F1-score, user's, producer's and overall accuracy. We used mainly F1-score as an indicator of the quality of classification because it represents the balance between the measures of the user's and producer's measures for each analyzed class. Figures 4 and 6 present the F1-score values obtained for the *Solidago* spp. class in 17 classification scenarios using two types of satellite images (Sentinel-2 and Planet-Scope) for the RF and OCSVM classifiers, respectively. The accuracy measures of the remaining (opponent) classes were of little interest in this study.

Random Forest classification.

Parameter	Values
Gamma (γ)	$2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}$
Nu (ν)	0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9

Table 7. List of values used in the grid search of OCSVM classifications.

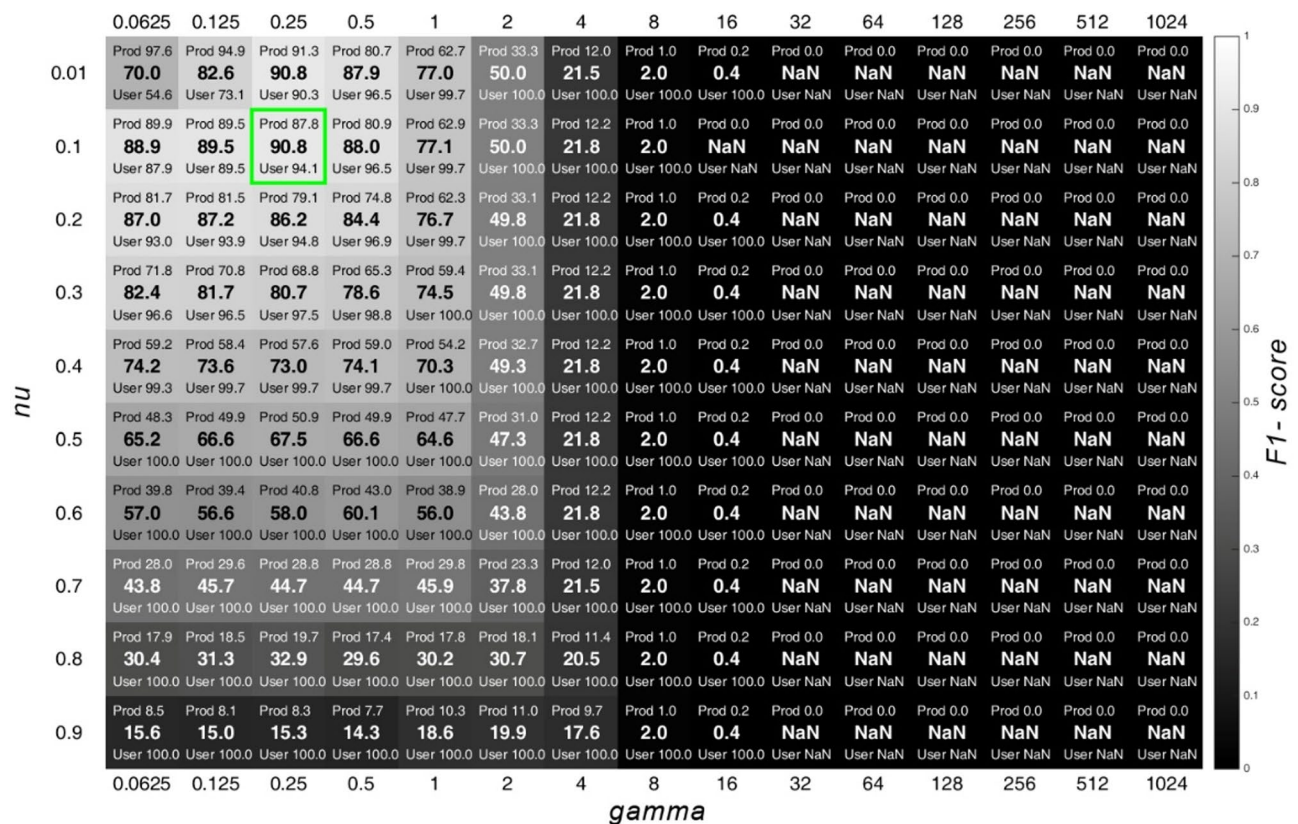


Fig. 3. Example of a grid search results - accuracy board for the classification of Sentinel-2 dataset (the first step of analysis - the parameters tuning).

For the RF classifier, the results (F1-scores) demonstrated relatively high accuracy across all scenarios and for both imagery sources. The highest F1-score of 0.98 was achieved using all possible combinations of Sentinel-2 data and their derivatives (scenario 17). In contrast, the lowest score of 0.71 was observed with the classification based on a single date (June) of PlanetScope data (Fig. 4; Table 8).

Sentinel-2 data generally yielded more consistent accuracy across different dates, except for the September image, which had the lowest values. PlanetScope data provided more differentiated results between the analyzed dates, with the June data providing the weakest result. For both datasets, autumn imagery (October or November) was crucial for achieving a high classification accuracy for *Solidago* spp. Additionally, F1-scores were higher for Sentinel-2 data compared to the PlanetScope data, with differences ranging from approximately 1 and 16%, depending on the scenario (Fig. 4). Overall accuracy for all scenarios was also higher for Sentinel-2 classifications (Table 8), reaching over 0.98 in the best scenario (17), compared to slightly lower values for PlanetScope data in the same scenario (0.97).

Scenarios utilizing multitemporal data generally resulted in higher accuracy scores compared to those based on single date images. Most F1-scores exceeded 0.9, except for scenario 15, which used yearly statistics of vegetation indices and yielded the lowest values from all classifications utilizing multitemporal data (0.87 and 0.8 for Sentinel-2 and PlanetScope data, respectively). The best results for the Sentinel-2 data were obtained using all possible data sources (scenarios 17–0.98), while the optimal classification for PlanetScope data was achieved with spectral bands from multitemporal imagery (scenario 11–0.97).

Figure 5 shows the F1-score for *Solidago* spp. classification using the RF algorithm performed on a reduced number of spectral bands (strategy 2). Only the bands that correspond between the two sensing systems were used. The analysis did not reveal obvious trends or advantages of either system for mapping *Solidago* spp., and the received F1-score values differed by a maximum of 5.6% (Table 9). For August and September, Sentinel-2 data yielded higher F1-scores, while for the three remaining dates, the PlanetScope system received better results. The highest F1-score was achieved with the PlanetScope data for the image from October (0.87), and the best result for Sentinel-2 was only slightly worse for the image from August (0.86). The lowest values for both sensors were observed for June with PlanetScope scoring 0.71 and Sentinel-2 scoring 0.68. Both sensors, demonstrated low user's accuracy of less than 0.6 (Table 9).

Features importance

As part of the RF classification, Gini importance was calculated, expressing the mean decrease in impurity¹⁹. Such measures can be considered as indications of the level of importance of the features used in the classification, though should be interpreted with caution⁵⁴. The Gini index was compared for classification based

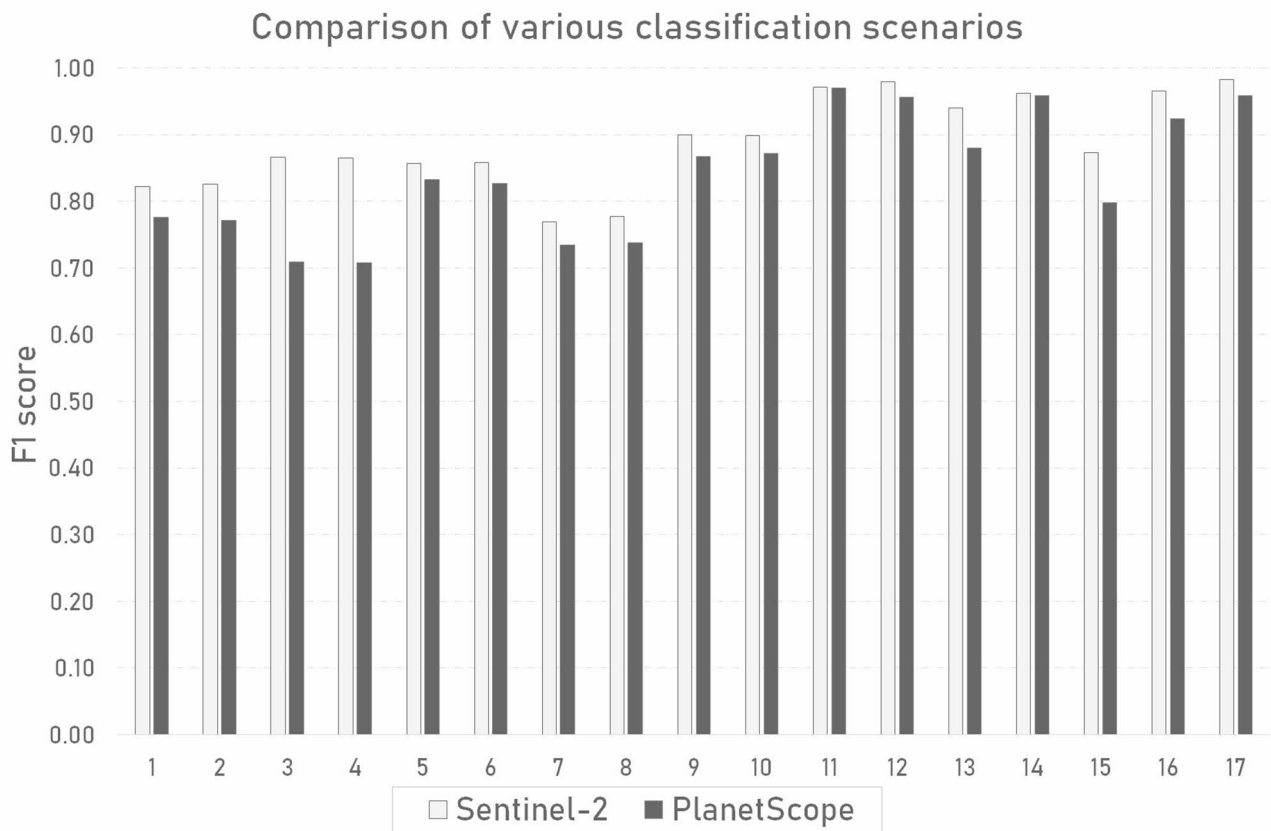


Fig. 4. Graph representing F1-scores of the *Solidago* class for 17 classification scenarios using Random Forest classifier (strategy 1).

Scenario	Sentinel-2				PlanetScope			
	UA	PA	F1-score	OA	UA	PA	F1-score	OA
1	0.795	0.852	0.822	0.924	0.833	0.726	0.776	0.897
2	0.882	0.776	0.826	0.925	0.843	0.710	0.771	0.891
3	0.808	0.934	0.866	0.946	0.588	0.892	0.709	0.901
4	0.802	0.938	0.865	0.937	0.586	0.894	0.708	0.864
5	0.810	0.910	0.857	0.905	0.765	0.912	0.832	0.886
6	0.806	0.916	0.858	0.908	0.745	0.928	0.826	0.887
7	0.674	0.894	0.769	0.911	0.657	0.834	0.735	0.909
8	0.689	0.892	0.778	0.913	0.677	0.810	0.738	0.899
9	0.939	0.864	0.900	0.933	0.874	0.860	0.867	0.914
10	0.941	0.860	0.899	0.934	0.888	0.856	0.872	0.902
11	0.956	0.988	0.972	0.976	0.970	0.970	0.970	0.955
12	0.965	0.994	0.979	0.977	0.947	0.966	0.956	0.962
13	0.949	0.930	0.939	0.943	0.823	0.946	0.880	0.924
14	0.953	0.970	0.961	0.971	0.942	0.976	0.959	0.963
15	0.853	0.894	0.873	0.934	0.706	0.918	0.798	0.912
16	0.948	0.984	0.966	0.979	0.886	0.964	0.923	0.965
17	0.971	0.994	0.982	0.982	0.944	0.974	0.959	0.972

Table 8. Accuracy measures for the *Solidago* class obtained from the random forest classification scenarios used in strategy 1. UA: User’s accuracy; PA: Producer’s accuracy; OA: Overall accuracy.

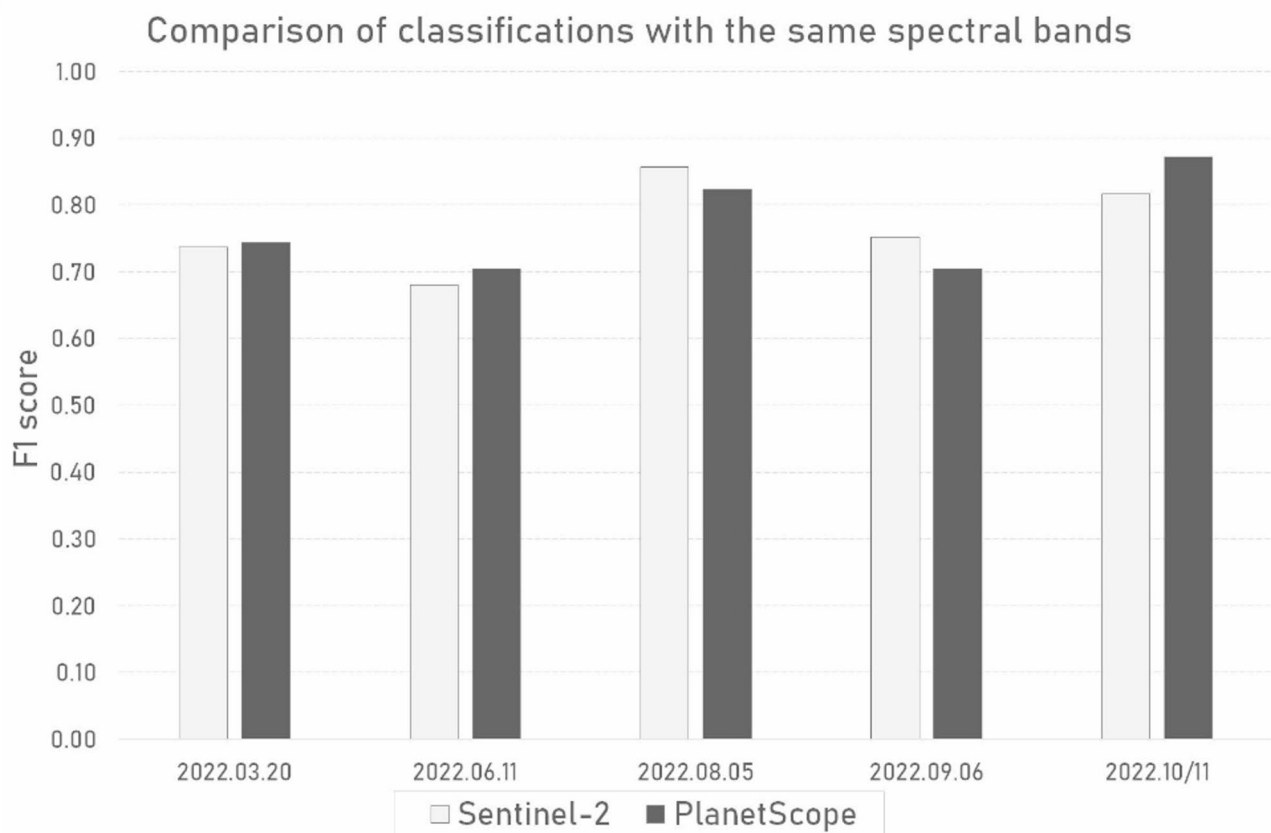


Fig. 5. Graph of F1-score values for the RF classification of *Solidago* spp. based only on the corresponding spectral bands.

Acquisition date	Sentinel-2				PlanetScope			
	UA	PA	F1-score	OA	UA	PA	F1-score	OA
2022.03.20	0.691	0.788	0.736	0.901	0.788	0.704	0.743	0.885
2022.06.11	0.554	0.882	0.681	0.890	0.580	0.898	0.705	0.897
2022.08.05	0.801	0.920	0.857	0.914	0.760	0.898	0.823	0.892
2022.09.06	0.671	0.854	0.752	0.876	0.606	0.842	0.705	0.895
2022.10.11	0.915	0.736	0.816	0.911	0.886	0.858	0.872	0.907

Table 9. Accuracy measures derived for the *Solidago* class in the random forest classification at different acquisition dates in strategy 2.

on the spectral bands for each acquisition date. It was not analyzed for classification using multitemporal data or spectral indices because of the extensive combination of classification features. Tables 10 and 11 present the relevance of particular bands of the Sentinel-2 and PlanetScope data, respectively, from the classifications in which all available spectral bands were used. For Sentinel-2 data, the SWIR bands B11 and B12 were consistently the most informative for all separate acquisition dates used in the study, followed by the NIR band (B8). Band B5, representing the red edge part of the electromagnetic spectrum, was the least important, considering the mean Gini index values from all images (dates). In the classifications based on PlanetScope imagery, the NIR band (B8) outperformed the other channels, whereas B1, the coastal blue band, was the least useful. In the second classification strategy (Tables 12 and 13), in which only the corresponding bands of both sensing systems were used, the NIR bands were found to be the most crucial for both EO datasets. The B3 band (green) of Sentinel-2 and B1 band of PlanetScope were the least informative.

OCSVM classification

The highest accuracy of the OCSVM classification, measured by F1-score was 0.9 for Sentinel-2 and 0.91 for PlanetScope data (Table 14). For the Sentinel-2 data the best accuracy was achieved for scenario 16, which utilized statistics from the spectral bands and vegetation indices across five dates. The lowest result for the Sentinel-2 data (F1-score of 0.7) was obtained for the September image combined with the VIs (scenario 8).

Spectral band	March	June	August	September	November	Mean value of Gini importance
B2	0.101	0.112	0.066	0.077	0.086	0.088
B3	0.089	0.054	0.099	0.071	0.063	0.075
B4	0.088	0.123	0.124	0.118	0.090	0.109
B5	0.044	0.056	0.082	0.077	0.041	<i>0.060</i>
B6	0.090	0.100	0.062	0.061	0.082	0.079
B7	0.077	0.075	0.056	0.079	0.072	0.072
B8	0.075	0.128	0.104	0.113	0.125	0.109
B8A	0.096	0.079	0.089	0.142	0.081	0.097
B11	0.155	0.159	0.202	0.153	0.175	<u>0.169</u>
B12	0.185	0.115	0.116	0.110	0.185	0.142

Table 10. Classification feature relevance expressed by Gini importance for RF classifications with Sentinel-2 imagery. The highest mean value is in bold and underlined and the lowest is in italics.

Spectral band	March	June	August	September	October	Mean value of Gini importance
B1	0.095	0.074	0.086	0.070	0.054	<i>0.076</i>
B2	0.116	0.082	0.085	0.092	0.112	0.097
B3	0.079	0.088	0.099	0.088	0.064	0.083
B4	0.078	0.086	0.100	0.095	0.066	0.085
B5	0.077	0.104	0.081	0.072	0.110	0.089
B6	0.173	0.146	0.155	0.137	0.140	0.150
B7	0.140	0.169	0.136	0.142	0.169	0.151
B8	0.241	0.252	0.258	0.304	0.284	<u>0.268</u>

Table 11. Classification features relevance expressed by Gini importance for RF classifications using planetscope imagery. The highest mean value is bolded and underlined, while the lowest is in italics.

Spectral band	March	June	August	September	November	Mean value of Gini importance
B1	0.188	0.088	0.101	0.104	0.125	0.121
B2	0.129	0.151	0.114	0.130	0.102	0.125
B3	0.093	0.120	0.142	0.099	0.091	<i>0.109</i>
B4	0.165	0.191	0.186	0.159	0.225	0.185
B5	0.160	0.186	0.181	0.188	0.179	0.179
B8A	0.265	0.265	0.275	0.319	0.279	<u>0.281</u>

Table 12. Classification features relevance expressed by Gini importance for RF classifications using Sentinel-2 imagery based on spectral bands that corresponds to planetscope data. The highest mean value is bolded and underlined, while the lowest is in italics.

Spectral band	March	June	August	September	October	Mean of Gini importance
B1	0.115	0.091	0.104	0.083	0.071	<i>0.093</i>
B2	0.137	0.101	0.104	0.112	0.148	0.121
B4	0.110	0.135	0.139	0.129	0.095	0.121
B6	0.219	0.180	0.182	0.170	0.184	0.187
B7	0.165	0.193	0.168	0.157	0.185	0.174
B8	0.254	0.299	0.302	0.350	0.318	<u>0.304</u>

Table 13. Classification feature relevance expressed by Gini importance for RF classifications using planetscope imagery based on spectral bands that corresponds to Sentinel-2 data. The highest mean value is bolded and underlined, while the lowest is in italics.

Scenario	Sentinel-2				PlanetScope			
	UA	PA	F1-score	OA	UA	PA	F1-score	OA
1	0.708	0.800	0.751	0.949	0.555	0.892	0.685	0.921
2	0.677	0.840	0.750	0.946	0.764	0.870	0.814	0.962
3	0.857	0.780	0.817	0.967	0.795	0.720	0.756	0.955
4	0.822	0.880	0.851	0.971	0.579	0.818	0.678	0.926
5	0.795	0.644	0.712	0.950	0.705	0.444	0.545	0.929
6	0.824	0.702	0.758	0.957	0.672	0.540	0.599	0.931
7	0.723	0.762	0.742	0.949	0.709	0.620	0.662	0.921
8	0.826	0.606	0.699	0.950	0.646	0.616	0.631	0.931
9	0.950	0.728	0.824	0.970	0.813	0.584	0.680	0.947
10	0.853	0.682	0.758	0.958	0.813	0.504	0.622	0.941
11	0.979	0.746	0.847	0.974	0.995	0.802	0.888	0.958
12	0.940	0.816	0.874	0.977	0.998	0.826	0.904	0.983
13	0.905	0.836	0.869	0.976	0.905	0.824	0.863	0.975
14	0.891	0.848	0.869	0.976	0.819	0.808	0.814	0.965
15	0.811	0.902	0.854	0.971	0.739	0.794	0.766	0.954
16	0.937	0.868	0.901	0.982	0.823	0.826	0.824	0.966
17	0.957	0.836	0.892	0.981	0.964	0.860	0.909	0.984

Table 14. Accuracy measures derived for the *Solidago* class for the OCSVM classification scenarios used in this study (strategy 1). UA: knowledge based; PA: Random Forest; OA = Overall accuracy.

The highest accuracy for PlanetScope data was obtained for the combined use of spectral bands from all dates and their derivatives (scenario 17) (Fig. 6). The lowest accuracy for PlanetScope was seen with the spectral bands from the August image (0.55). With only a single Sentinel-2 image used, the highest F1-score of 0.86 was achieved from June data, where spectral bands were combined with vegetation indices (scenario 4). For PlanetScope, the best results for a single date were obtained for images from March (0.81), which were based purely on the spectral bands.

A comparison of both classifiers shows that RF enables more accurate classification than the OCSVM across the available datasets. The differences in the F1-score values for particular scenarios ranged from 1 to 28%. In all scenarios, the Sentinel-2 data yielded higher accuracies with the RF classifier. For PlanetScope, RF outperformed OCSVM in all scenarios except scenarios 3 and 4. For both the classifiers, the Sentinel-2 image from September provided the weakest results. Similarly, for both algorithms, multitemporal datasets (scenarios 11–17) tended to ensure more accurate results.

Solidago maps

Classification scenarios that performed best for both classifiers and datasets were used to prepare maps of the extent of goldenrods within the KNP. Figure 7 shows the distribution of *Solidago* spp. derived from Sentinel-2 and PlanetScope imagery using the RF algorithm. The maps resulting from the RF classification presented a very similar distribution of *Solidago* spp. in both datasets, with patches of different sizes scattered in the analyzed area. Slightly larger fragmentation is visible in the map based on PlanetScope data (Fig. 7b), particularly in the western part, which may originate from the different classification results and pixel sizes of the dataset used. From the perspective of the entire study area, good differentiation was observed between *Solidago* spp. and forested areas (as seen on the orthophoto map in the background). Most *Solidago* spp. occurs in grasslands or agricultural areas.

The maps representing the extent of *Solidago* spp. derived from the OCSVM classifier and the two datasets (Fig. 8) were even more similar to each other than those derived from the RF classifier (Fig. 7). It can also be observed that the plant extent was larger in the OCSVM classification map than in the RF maps. To some extent, this may result from the lower classification accuracy of the OCSVM results as compared to RF, with a slightly larger overestimation of the *Solidago* extent (Tables 8 and 14). Additional computations confirm this fact. The area covered by *Solidago* spp. classified by OCSVM on Sentinel-2 data (9895 ha) was 2180 ha larger than that derived from RF. Similarly, the area estimated from PlanetScope imagery and OCSVM (10650 ha) exceeded the RF result by 4250 ha. In general, the distribution of *Solidago* spp. on the maps was similar between the classification algorithms and satellite datasets, with the occurrence of invasive plants dominated either grassland or agricultural land. Figure 9 illustrates subsets of classification results for two areas (AOIs) for both satellite datasets and two classification algorithms. The illustrations present differences in the classification results and allow comparison with orthophoto (Fig. 9a and f).

Discussion

This study demonstrates the applicability of remote sensing and machine learning techniques for detecting and monitoring of *Solidago* spp. in a real-world conservation setting. Unlike many previous studies that focus on algorithm performance in controlled or small-scale settings, our research evaluated classification accuracy for a

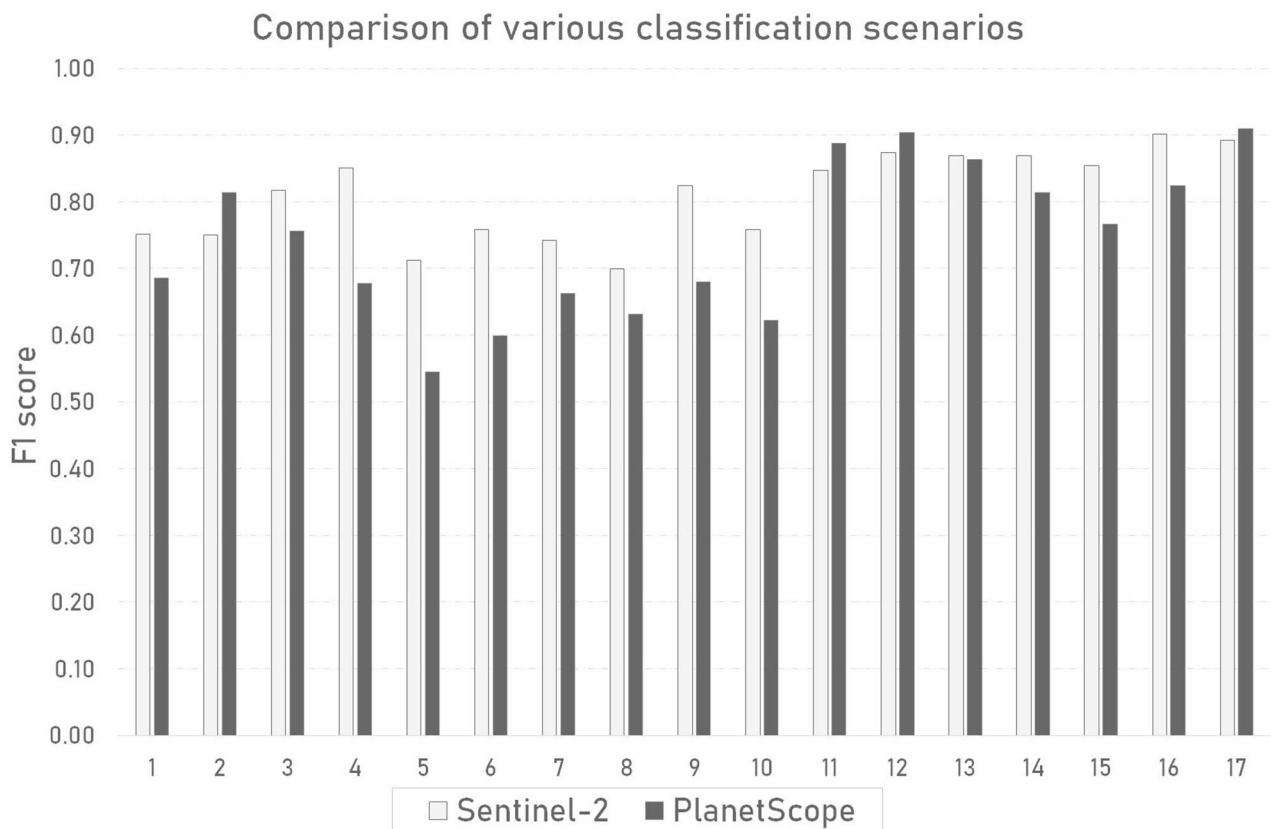


Fig. 6. Graph illustrating F1-score values of the *Solidago* class for 17 classification scenarios using OCSVM classifier (strategy 1).

well-established IAS across a large, heterogeneous protected area. Using operationally available multitemporal satellite data, we integrated phenological insights with multi-sensor observations to map *Solidago* spp. within KNP, Poland, a landscape experiencing increasing invasion pressure. The results confirmed that the RF classifier outperformed OCSVM across most scenarios, particularly in multitemporal data were included.

RF's ability to handle high-dimensional input and minimize overfitting has been demonstrated in many remote sensing applications^{18,19} and this study confirms its strength in the context of seasonal mapping of invasive plants. In contrast, the OCSVM classifier showed greater variability and tended to overestimate the presence of *Solidago* spp., especially when based on early- or mid-season data. This is consistent with known limitations of one-class classifiers, which lack information on background heterogeneity^{20,21}. The analysis also provides practical insight into the relative utility of Sentinel-2 and PlanetScope imagery. Sentinel-2 consistently achieved higher classification accuracy, particularly in multitemporal scenarios, and its broader spectral range (SWIR bands) proved valuable for detecting phenological shifts in *Solidago* spp. populations. Bands B11 and B12 were repeatedly identified as key features in RF classifications. These results support some previous work⁵⁵ but our study adds the important observation that these spectral advantages translate into measurable differences in seasonal detection performance, especially in autumn. Importantly, Sentinel-2 is freely available, making it more feasible option for national-scale IAS monitoring. Although PlanetScope data produced slightly lower accuracies, its finer spatial resolution (3 m) allowed for more detailed detection of fragmented or narrow *Solidago* patches that might be missed at Sentinel-2's coarser resolution. This has implications for targeted control efforts or site-scale habitat restoration. However, PlanetScope's commercial licensing presents a significant barrier to wide-scale use, particularly for public agencies or conservation Non-Governmental Organizations with limited budgets. These findings highlight that the practical trade-off between spatial detail and cost must be considered, and that Sentinel-2 data alone may be sufficient for most management applications when classification is optimized with temporal and spectral information.

One of the novel contributions of this study is the detailed assessment of phenological timing in classification accuracy. While phenology is frequently cited as important in vegetation studies^{9,56} our work quantifies how seasonal differences affect the spectral separability of *Solidago* spp. The highest classification accuracy was achieved using autumn imagery (October–November), surpassing even the flowering period in August–September. This outcome contradicts some earlier findings^{13,16} that emphasized the blooming phase as the optimal detection window. Instead, our results suggest that the senescent stage — when seed heads are prominent and co-occurring vegetation is declining — offers better spectral distinction for *Solidago*. These findings provide operational guidance on when to acquire imagery for mapping this species. Additionally, March and June images

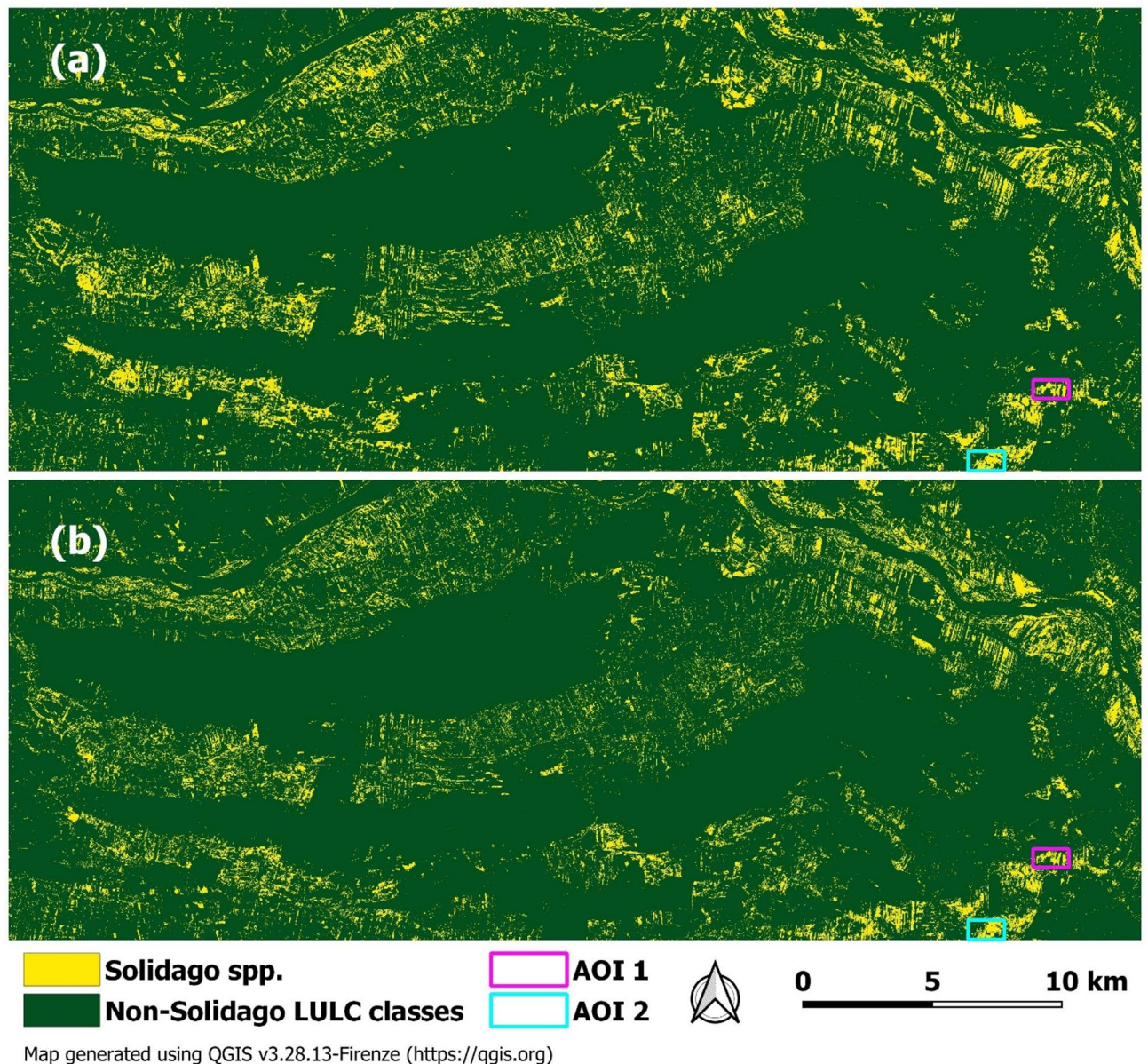


Fig. 7. Maps of the study area with classified extent of *Solidago* spp. using the RF classifier; (a) classification of Sentinel-2 data (scenario 17) and (b) classification of PlanetScope imagery (scenario 11). AOI 1 and AOI 2 indicate locations of subsets presented in Fig. 9.

also yielded high classification accuracy under some scenarios. Early spring imagery likely captured the dry biomass from the previous season, which remains visible before native vegetation regenerates. Meanwhile, June corresponds to a rapid biomass accumulation phase of *Solidago* spp., producing strong spectral signals in the NIR and red-edge bands. This finding underscores the importance of understanding species-specific growth stages rather than assuming that general vegetation vigor (e.g., NDVI peaks) aligns with optimal detection. Future monitoring efforts would benefit from aligning image acquisition with known phenological milestones of target IAS.

Another important finding concerns the relative utility of vegetation indices. While indices like NDVI and NDRE are widely used to enhance vegetation classification, their addition in this study did not consistently improve classification accuracy. In fact, several scenarios showed slightly lower performance when VIs were included, compared to using raw spectral bands alone. This suggests that the spectral signature of *Solidago* spp., especially when tracked across multiple dates, is already sufficiently distinct, and that the computational effort of calculating additional indices may be unnecessary in this case. These results align with recent observations that spectral richness and multitemporal statistics can outperform indices under certain conditions¹⁷. The visual comparison of classification maps between sensors and classifiers revealed that both systems captured a similar spatial pattern of *Solidago* spp. distribution, although PlanetScope-based classifications showed slightly more fragmentation. OCSVM-derived maps tended to overestimate species extent, highlighting the need for careful

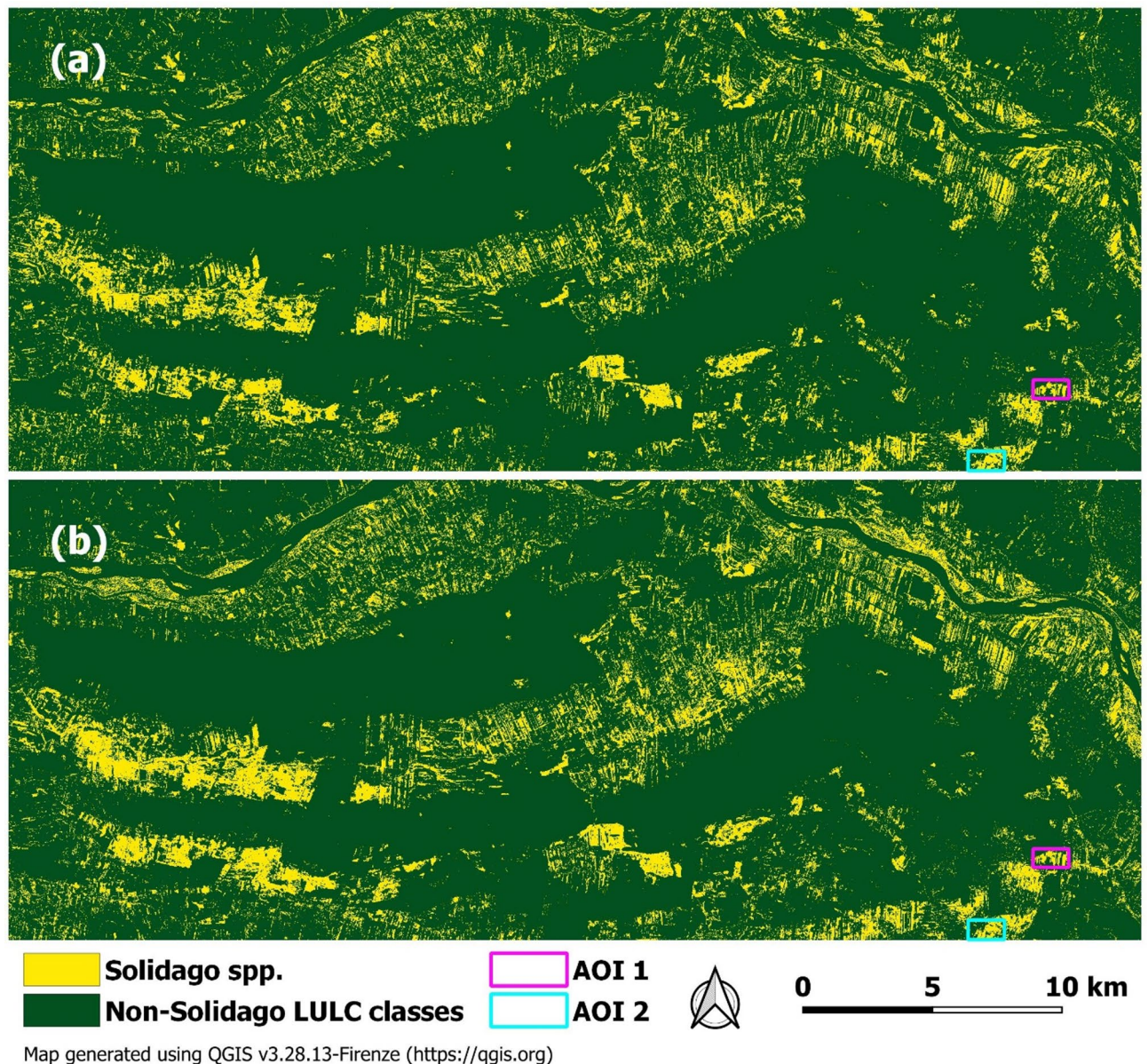


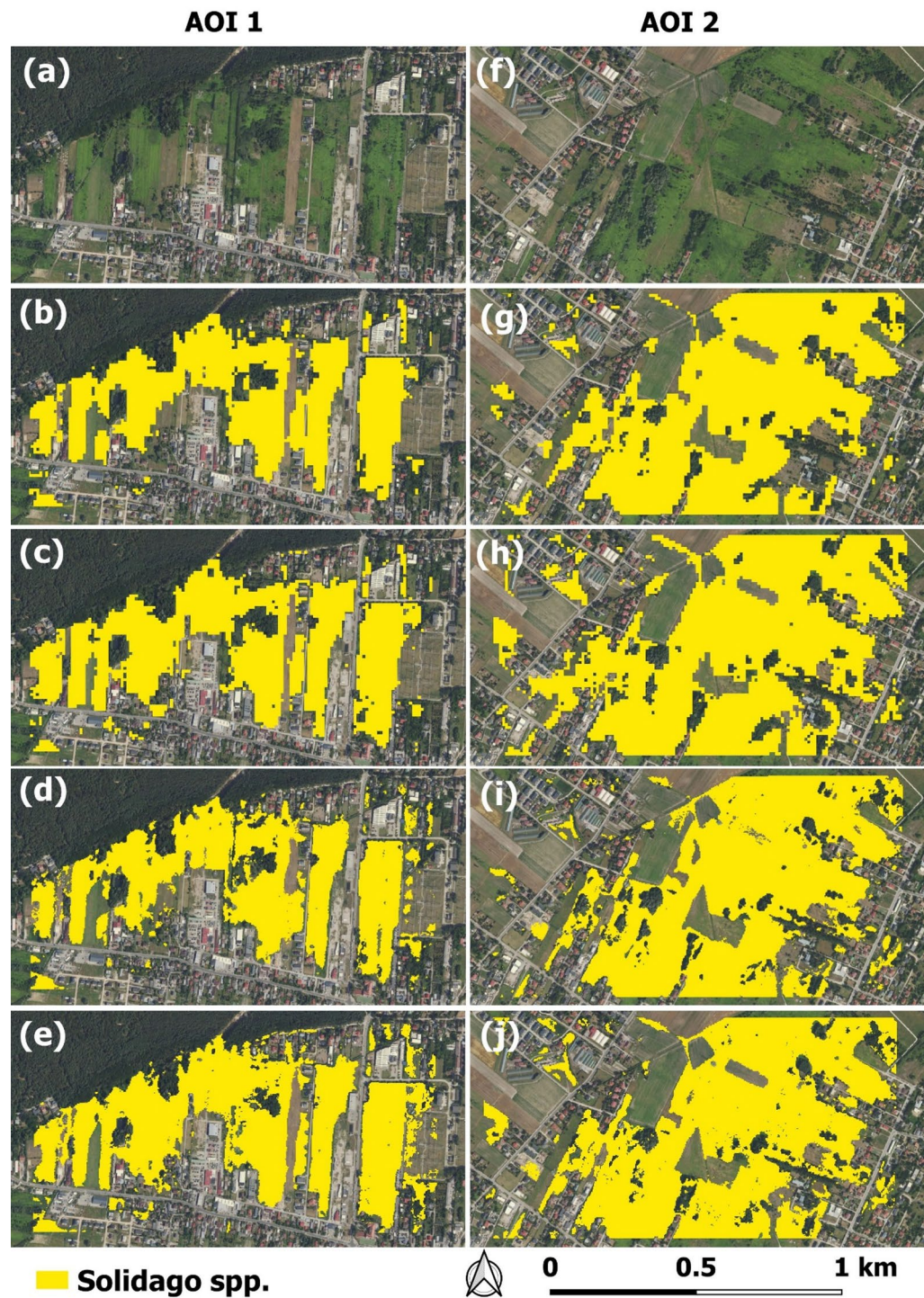
Fig. 8. Maps of the study area with classified extent of *Solidago* spp. with OCSVM classifier; (a) classification of Sentinel-2 data (scenario 16) and (b) classification of PlanetScope data (scenario 17).

interpretation when using one-class methods without extensive field validation. While pixel-level accuracy metrics (F1-score, UA, PA) are commonly used, future work could consider incorporating area-adjusted accuracy assessments^{57,58} to better estimate the actual extent of invasion. This is especially relevant when classification outputs are used to quantify infested area for planning eradication or restoration measures.

It is worth noting that other open satellite data sources, such as Sentinel-1, provide a different type of information (Synthetic Aperture Radar) about objects on the Earth's surface. These data could potentially enhance the performance of algorithms for *Solidago* spp. detection by contributing complementary features. However, verifying this hypothesis requires further research, and thus represents a promising direction for developing more advanced and accurate goldenrod classification methods. Notably, our literature review did not identify studies that have explored the use of Sentinel-1 data for *Solidago* spp. detection.

Practical recommendations

The results have broader implications for conservation planning and invasive species management. First, the study confirms that combining multitemporal satellite imagery with machine learning provides a viable, scalable approach to IAS monitoring. In areas like KNP, where goldenrods are widespread and expanding, timely and accurate maps are essential for prioritizing management actions. Second, the identification of optimal detection periods enables more efficient monitoring campaigns by aligning image acquisition with phenological windows. Finally, the methods tested here should be transferable to other invasive plant species with similar



Map generated using QGIS v3.28.13-Firenze (<https://qgis.org>),
Background orthophoto source: Główny Urząd Geodezji i Kartografii (GUGiK), Ortofotomapa, WMS service via Geoportal
(<https://www.geoportal.gov.pl>).

Fig. 9. Subsets of *Solidago spp.* classification in two locations (AOI 1 and AOI 2 indicated in Fig. 7) representing result from different classifier and datasets; **a** and **f** orthophotomaps from Geoportal WMS service of the selected AOI 1 and AOI 2 (also used as a background in all subsets maps); **b** classification of Sentinel-2 data with RF algorithm in AOI 1; **c** classification of Sentinel-2 data with OCSVM algorithm in AOI 1; **d** classification of PlanetScope data with RF algorithm in AOI 1; **e** classification of PlanetScope data with OCSVM algorithm in AOI 1; **g** classification of Sentinel-2 data with RF algorithm in AOI 2; **h** classification of Sentinel-2 data with OCSVM algorithm in AOI 2; **i** classification of PlanetScope data with RF algorithm in AOI 2; **j** classification of PlanetScope data with OCSVM algorithm in AOI 2.

growth dynamics, especially those dominating disturbed or abandoned land. In conclusion, this study provides a practical and ecologically relevant framework for detecting and monitoring *Solidago* spp. using remote sensing and machine learning. The integration of phenological knowledge, multitemporal imagery, and sensor comparison advances the field beyond algorithm benchmarking toward real-world application. The findings can support environmental agencies and land managers in making informed decisions about where and when to act against *Solidago* invasions, while also contributing to the development of more effective remote sensing tools for biodiversity conservation.

Based on our findings, we recommend to use multitemporal Sentinel-2 data (in our study, 5 images from March to November) for *Solidago* spp. mapping, especially for large-scale monitoring. For finer-scale early detection or site-specific interventions, PlanetScope data may offer additional spatial detail, however, its cost limits its applicability for broad-scale use. The highest classification accuracy was achieved when training was performed on stratified samples representing various land cover and land use classes. However, acquiring such comprehensive reference data is often time- and cost-intensive. If only *Solidago* spp. samples are available, a decrease in classification accuracy can be expected - from an F1-score of 0.98 (using a Random Forest classifier) to 0.90 (using a One-Class SVM classifier). When using only a single image, the maximum achievable F1-scores were 0.90 (RF) and 0.85 (OCSVM), respectively.

Our results indicate that the timing of image acquisition throughout the season plays a more critical role in detection performance than the use of vegetation indices. These insights can support protected area managers, ecologists, and policymakers in designing cost-effective and timely invasive species monitoring strategies.

Conclusion

This study demonstrates the practical potential of combining satellite-based remote sensing and machine learning techniques to support the monitoring of invasive goldenrod (*Solidago* spp.) in Central Europe. By integrating multitemporal imagery from Sentinel-2 and PlanetScope with two classification approaches, RF and OCSVM, we provide a comparative assessment of methods suited to large-area and fine-scale detection of this ecologically impactful species. Our findings highlight the importance of selecting appropriate phenological windows for image acquisition. Contrary to common assumptions, autumn (October - November) emerged as the most effective period for *Solidago* detection, likely due to increased spectral contrast with native vegetation. While PlanetScope's higher spatial resolution improved the detail of local mapping, Sentinel-2 consistently delivered superior classification accuracy, especially when using its shortwave infrared bands and multitemporal data, making it a cost-effective tool for regional monitoring. Among classifiers, RF outperformed OCSVM in nearly all scenarios, offering greater stability and accuracy when sufficient reference data were available. OCSVM, while useful in data-limited contexts, tended to overestimate species presence and requires careful tuning. From a management perspective, this study supports the operational use of Sentinel-2 for broad-scale *Solidago* monitoring and suggests that combining spatial and temporal features can improve detection outcomes without relying heavily on complex vegetation indices. These results can inform the design of early-warning systems and targeted control efforts in protected and invaded areas. The approach might be adaptable to other invasive plant species with distinct seasonal growth patterns, contributing to scalable, satellite-based solutions for biodiversity conservation and ecological restoration.

Data availability

Satellite imagery from Sentinel-2 was accessed from CREODIAS (<https://explore.creodias.eu/search>) where it is openly available. PlanetScope imagery was accessed under Education and Research Program and its use is limited by Planet Ltd. licensing. In-situ dataset collected for this research is available from the corresponding author on reasonable request.

Received: 22 September 2024; Accepted: 25 August 2025

Published online: 01 September 2025

References

- Pyšek, P. et al. Scientists' warning on invasive alien species. *Biol. Rev.* **95**, 1511–1534 (2020).
- Lenda, M. et al. Misinformation, internet honey trading and beekeepers drive a plant invasion. *Ecol. Lett.* **24**, 165–169 (2021).
- Paini, D. R. et al. Global threat to agriculture from invasive species. *Proc. Natl. Acad. Sci.* **113**, 7575–7579 (2016).
- Walsh, J. R., Carpenter, S. R. & Vander Zanden, M. J. Invasive species triggers a massive loss of ecosystem services through a trophic cascade. *Proc. Natl. Acad. Sci.* **113**, 4081–4085 (2016).
- Moroň, D. et al. Wild pollinator communities are negatively affected by invasion of alien goldenrods in grassland landscapes. *Biol. Conserv.* **142**, 1322–1332 (2009).
- de Groot, M., Kleijn, D. & Jogan, N. Species groups occupying different trophic levels respond differently to the invasion of semi-natural vegetation by *Solidago canadensis*. *Biol. Conserv.* **136**, 612–617 (2007).
- Skórka, P. & Lenda, M. Abandoned Fields as Refuges for Butterflies in the Agricultural Landscapes of Eastern Europe. In *Insect Habitats: Characteristics, Diversity and Management* (eds Harris, E. L., Davies, N. E.), 83–103. (New York: Nova Science Publishers, Inc., 2009).
- Huang, C. & Asner, G. P. Applications of remote sensing to alien invasive plant studies. *Sens. (Switzerland)* **9**, 4869–4889 (2009).
- Vaz, A. S., Alcaraz-Segura, D., Campos, J. C., Vicente, J. R. & Honrado, J. P. Managing plant invasions through the lens of remote sensing: A review of progress and the way forward. *Sci. Total Environ.* **642**, 1328–1339 (2018).
- Ishii, J. & Washitani, I. Early detection of the invasive alien plant *Solidago altissima* in moist tall grassland using hyperspectral imagery. *Int. J. Remote Sens.* **34**, 5926–5936 (2013).
- Rizaludin Mahmud, M., Numata, S. & Hosaka, T. Mapping an invasive goldenrod of *Solidago altissima* in urban landscape of Japan using multi-scale remote sensing and knowledge-based classification. *Ecol. Ind.* **111**, 105975 (2020).
- Koco, Š., Dubravská, A., Vilček, J. & Grulová, D. Geospatial approaches to monitoring the spread of invasive species of *Solidago* spp. *Remote Sens.* **13**, 4787 (2021).

13. Zagajewski, B., Kluczek, M., Zdunek, K. B. & Holland, D. Sentinel-2 versus planetscope images for goldenrod invasive plant species mapping. *Remote Sens.* **16**, 636 (2024).
14. Chadoulis, R. T. et al. Phenological metrics derived from Sentinel-2 data for *Solidago gigantea* mapping. In *IGARSS 2023–2023 IEEE International Geoscience and Remote Sensing Symposium* 445–447. <https://doi.org/10.1109/IGARSS52108.2023.10282732> (2023).
15. Kopeć, D. et al. Using airborne hyperspectral imaging spectroscopy to accurately monitor invasive and expansive herb plants: Limitations and requirements of the method. *Sensors* **19**, 2871 (2019).
16. Sabat-Tomala, A., Raczko, E. & Zagajewski, B. Comparison of support vector machine and random forest algorithms for invasive and expansive species classification using airborne hyperspectral data. *Remote Sens.* **12**, 516 (2020).
17. Sabat-Tomala, A., Raczko, E. & Zagajewski, B. Mapping invasive plant species with hyperspectral data based on iterative accuracy assessment techniques. *Remote Sens.* **14**, 64 (2021).
18. Belgiu, M. & Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **114**, 24–31 (2016).
19. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
20. Chang, C. C. & Lin, C. J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**, 1–27 (2011).
21. Krupinski, M., Lewiński, S. & Malinowski, R. One class SVM for Building detection on Sentinel-2 images. In *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019* (eds Romaniuk, R. S. & Linczuk, M.). <https://doi.org/10.1117/12.2535547> (SPIE, Wilga, 2019).
22. Moussa, R., Azar, D. & Sarro, F. On the effectiveness of one-class support vector machine in different defect prediction scenarios. <https://doi.org/10.48550/ARXIV.2202.12074> (2022).
23. Weber, E. Current and potential ranges of three exotic goldenrods (*Solidago*) in Europe. *Conserv. Biol.* **15**, 122–128 (2001).
24. Perera, P. C. D., Szymura, T. H., Zajac, A., Chmolewska, D. & Szymura, M. Drivers of *Solidago* species invasion in central Europe—Case study in the landscape of the Carpathian mountains and their foreground. *Ecol. Evol.* **11**, 12429–12444 (2021).
25. Kotowska, D., Skórka, P., Pärt, T., Auffret, A. G. & Zmihorski, M. Spatial scale matters for predicting plant invasions along roads. *J. Ecol.* **112**, 305–318 (2024).
26. Herr, C., Chapuis-Lardy, L., Dassonville, N., Vanderhoeven, S. & Meerts, P. Seasonal effect of the exotic invasive plant *Solidago gigantea* on soil pH and P fractions. *J. Plant Nutr. Soil Sci.* **170**, 729–738 (2007).
27. Abhilasha, D., Quintana, N., Vivanco, J. & Joshi, J. Do allelopathic compounds in invasive *Solidago canadensis* s.l. Restrained the native European flora? *J. Ecol.* **96**, 993–1001 (2008).
28. Bing-yao, S. U. N., Jian-zhong, T. N. & Fu-gen, G. U. Zhi-gang, W. a. N., & Ming-de, Z. H. U. Allelopathic effects of extracts from *Solidago canadensis* L. against seed germination and seedling growth of some plants. *J. Environ. Sci.* **18**, 304–309 (2006).
29. Chapuis-Lardy, L., Vanderhoeven, S., Dassonville, N., Koutika, L. S. & Meerts, P. Effect of the exotic invasive plant *Solidago gigantea* on soil phosphorus status. *Biol. Fertil. Soils* **42**, 481–489 (2006).
30. Weber, E. Biological flora of central Europe: *Solidago altissima* L. *Flora* **195**, 123–134 (2000).
31. Bochenek, A. et al. Do the seeds of *Solidago gigantea* Aiton have physiological determinants of invasiveness? *Acta Physiol. Plant.* **38**, 159 (2016).
32. Werner, P. A. & Platt, W. J. Ecological relationships of co-occurring goldenrods (*Solidago*: Compositae). *Am. Nat.* **110**, 959–971 (1976).
33. Szymura, M., Szymura, T. H., Wolski, K. & Świerszcz, S. Can native grass species outcompete invasive goldenrods? Results of a replacement series experiment. *Weed Res.* **58**, 304–317 (2018).
34. Michalska-Hejduk, D. & Otręba, A. *Inwazyjne gatunki roślin w Kampinoskim Parku Narodowym iw jego sąsiedztwie*. Kampinoski Park Narodowy (2014).
35. Gascon, F. et al. Copernicus Sentinel-2A calibration and products validation status. *Remote Sens.* **9**, 584 (2017).
36. CREODIAS. CloudFerro: CREODIAS – Copernicus Data and Information Access Services. <https://explore.creodias.eu/search> (2025).
37. Planet Team. Planet Application Program Interface: In Space for Life on Earth. <https://api.planet.com/> (2017).
38. BDOT. Database of Topographic Objects (pol. Baza Danych Obiektów Topologicznych). <https://www.geoportal.gov.pl/dane/baz-a-danych-obiektow-topograficznych-bdot>.
39. Jensen, J. R. *Remote Sensing of the Environment: An Earth Resource Perspective*, 2nd edn. (Pearson Prentice Hall, 2007).
40. Gitelson, A. A., Kaufman, Y. J. & Merzlyak, M. N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **58**, 289–298 (1996).
41. Dash, J., Curran, P. J. & and The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* **25**, 5403–5413 (2004).
42. Gitelson, A. & Merzlyak, M. N. Quantitative Estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B* **22**, 247–252 (1994).
43. Rouse, J. W. Jr., Haas, R. H., Schell, J. A. & Deering, D. W. Monitoring vegetation systems in the great plains with ERTS. In *Proceedings of the Third Earth Resources Technology Satellite Symposium* Vol. 1, 309–317, (NASA, Goddard Space Flight Center, 1973).
44. Rondeaux, G., Steven, M. & Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **55**, 95–107 (1996).
45. Huete, A. R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **25**, 295–309 (1988).
46. Jordan, C. F. Derivation of Leaf-Area index from quality of light on the forest floor. *Ecology* **50**, 663–666 (1969).
47. Pearson, R. L. & Miller, L. D. Remote mapping of standing crop biomass for estimation of the productivity of the short-grass prairie, Panee National grasslands, Colorado. In *Proceedings of the eighth international symposium on remote sensing of environment*, October 2–6, 1357–1381, 1972, (Ann Arbor, Michigan, 1972).
48. Ettehad Osgouei, P., Kaya, S., Sertel, E. & Alganci, U. Separating built-up areas from bare land in mediterranean cities using Sentinel-2A imagery. *Remote Sens.* **11** (2019).
49. Zhang, T. X., Su, J. Y., Liu, C. J. & Chen, W. H. Potential bands of Sentinel-2A satellite for classification problems in precision agriculture. *Int. J. Autom. Comput.* **16**, 16–26 (2019).
50. Malinowski, R., Heckrath, G., Rybicki, M. & Eltnier, A. Mapping rill soil erosion in agricultural fields with UAV-borne remote sensing data. *Earth Surf. Process Land.* **48**, 596–612 (2023).
51. Malinowski, R. et al. Automated production of a land cover/use map of Europe based on sentinel-2 imagery. *Remote Sens.* **12**, 1–27 (2020).
52. Woźniak, E. et al. Multi-temporal phenological indices derived from time series Sentinel-1 images to country-wide crop classification. *Int. J. Appl. Earth Obs. Geoinf.* **107**, 102683 (2022).
53. Krupinski, M. et al. Satellite monitoring of sea ice dynamics at local scale. In *Earth Resources and Environmental Remote Sensing/ GIS Applications XIV* (eds Schulz, K., Nikolakopoulos, K. G. & Michel, U.) <https://doi.org/10.1117/12.2684340>. (SPIE, 2023).
54. Strobl, C., Boulesteix, A. L., Zeileis, A. & Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **8**, 25 (2007).
55. Thenkabail, P. S., Lyon, J. G. & Huete, A. Advances in hyperspectral remote sensing of vegetation and agricultural crops. In *Fundamentals, Sensor Systems, Spectral Libraries, and Data Mining for Vegetation*. (CRC Press, 2018).
56. Bradley, B. A. et al. Species detection vs. habitat suitability: Are we biasing habitat suitability models with remotely sensed data? *Ecol. Model.* **244**, 57–64 (2012).
57. Stehman, S. V. Estimating area from an accuracy assessment error matrix. *Remote Sens. Environ.* **132**, 202–211 (2013).

58. Olofsson, P. et al. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **148**, 42–57 (2014).

Acknowledgements

The authors acknowledge the possibility of using PlanetScope data from the Planet Lab under the Education and Research Program. We are grateful to two anonymous referees for constructive comments on earlier drafts of the manuscript. This work was supported by the European Union's Horizon 2020 research and innovation program under grant agreement no. 952111 entitled Earth Observation Training in Science and Technology at the Space Research Centre of the Polish Academy of Sciences (EOTiST). M.L., K.C. and Ł. M. were financed by project "Sonata" 2021/43/D/NZ9/02990 from the National Science Center, Poland.

Author contributions

All authors greatly contributed to this work. Conceptualization: R.M., M.L. Methodology: R.M., M.K., Data processing: R.M., M.K. In situ data collection: K.C., Results validation: M.K., Writing – original draft preparation: R.M., M.K., M.L., P.S., K.C., Ł.M.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025