

# Journal of Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

## Mapping secondary succession species in agricultural landscape with the use of hyperspectral and airborne laser scanning data

Aleksandra Radecka  
Dorota Michalska-Hejduk  
Katarzyna Osińska-Skotak  
Adam Kania  
Konrad Górski  
Wojciech Ostrowski

**SPIE.**

Aleksandra Radecka, Dorota Michalska-Hejduk, Katarzyna Osińska-Skotak, Adam Kania, Konrad Górski, Wojciech Ostrowski, "Mapping secondary succession species in agricultural landscape with the use of hyperspectral and airborne laser scanning data," *J. Appl. Remote Sens.* **13**(3), 034502 (2019), doi: 10.1117/1.JRS.13.034502.

# Mapping secondary succession species in agricultural landscape with the use of hyperspectral and airborne laser scanning data

Aleksandra Radecka,<sup>a,\*</sup> Dorota Michalska-Hejduk,<sup>b</sup>  
Katarzyna Osińska-Skotak,<sup>a</sup> Adam Kania,<sup>c</sup> Konrad Górski,<sup>a</sup>  
and Wojciech Ostrowski<sup>a</sup>

<sup>a</sup>Warsaw University of Technology, Department of Photogrammetry, Remote Sensing and Spatial Information Systems, Faculty of Geodesy and Cartography, Warsaw, Poland

<sup>b</sup>University of Lodz, Department of Geobotany and Plant Ecology, Lodz, Poland

<sup>c</sup>Definity Ltd., Sw. Judy Tadeusza, Wroclaw, Poland

**Abstract.** Secondary succession is a process that is often observed taking place in former agricultural ecosystems. Its characteristics are especially important in protected areas, for the purposes of monitoring and protective measures. Effective mapping of succession is facilitated by the development of automated methodologies based on remote sensing data, which are capable of complementing traditional field research. The objective of this work is to determine whether the classification of high-resolution hyperspectral and light detection and ranging (LiDAR) data with the use of the random forest algorithm enables us to produce an accurate succession species map. First, feature extraction techniques are applied to 1-m hyperspectral images and a  $\sim 7$  point/m<sup>2</sup> dense point cloud. Minimum noise fraction layers and vegetation indices are calculated from the hyperspectral data and geometry related indices from the LiDAR data. Finally, the recursive feature elimination algorithm is applied to the combined dataset and the reference polygons to select the optimal set of features for subsequent classification. The results indicate that the proposed methodology has the potential to be used operationally. The final classification product is characterized by a relatively high Cohen's kappa value of 0.68, with single species classified with various accuracies, expressed by *F1* scores ranging from 0.45 to 0.87. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.13.034502](https://doi.org/10.1117/1.JRS.13.034502)]

**Keywords:** ecological process; Natura 2000 monitoring; hyperspectral imaging; light detection and ranging data; data fusion; random forest.

Paper 190042 received Jan. 18, 2019; accepted for publication Jun. 20, 2019; published online Jul. 16, 2019.

## 1 Introduction

The succession process is currently the subject of a large number of research projects all over the world, for example in Europe,<sup>1–3</sup> North America,<sup>4–9</sup> South America,<sup>10–12</sup> and also in Asia.<sup>13</sup> The reason behind the interest in this topic is undoubtedly connected with its significance, which can be seen from three different perspectives. In some areas, similar to the current research, the secondary succession is perceived as a threat to certain ecosystems. This is because large areas of seminatural nonforest communities, e.g., grasslands and meadows (and especially the less productive parts) have been abandoned, resulting in secondary succession.<sup>14–16</sup> The cessation of mowing or grazing allows species with clonal growth to complete their full development and induces changes in the quantitative and spatial structures of plant communities.<sup>17</sup> One of the results of this process is the disappearance of some groups of species (e.g., heliophilous ones) and the formation of shrub and forest communities created by species that are better adapted to poor light conditions. This situation leads to changes in the composition of species in an ecosystem,<sup>18,19</sup> in not only plant species but also animal communities.<sup>20–22</sup> As a result, secondary

---

\*Address all correspondence to Aleksandra Radecka, E-mail: [aleksandra.radecka@pw.edu.pl](mailto:aleksandra.radecka@pw.edu.pl)

succession also leads to changes at the landscape level. This standpoint is generally taken if the process threatens the existence of protected habitats or the vegetation involves invasive species. This view is represented by Szostak et al.,<sup>2</sup> Szostak et al.,<sup>3</sup> Vanderlinder et al.,<sup>7</sup> and others. In other places, secondary succession is recognized as having a highly positive impact on climate and biodiversity and as an important reservoir of carbon.<sup>8,9,11</sup> In most cases, however, the importance of secondary succession is simply explained by the large size of the terrain encompassed, land use and land cover changes, and the resultant need to assess its influence on ecosystem functions and services.<sup>1,4–6,10,12,13</sup>

Regardless of the reason, studies of secondary succession focus mainly on identifying the places, in which the succession occurs, estimating the size of this area<sup>3,7</sup> and describing the stages of the process,<sup>1,4–6,9–13</sup> and rarely on determining other characteristics such as the stand age or species richness.<sup>8</sup> To the authors' knowledge, few research projects have investigated the subject of the differentiation of succession species. Although many studies have focused on the identification of tree, shrub, or other plant species, only a small number of these have involved the mapping of succession tree species.<sup>23,24</sup> The majority of studies have been carried out in an urban environment<sup>25–27</sup> or in different types of mature forests.<sup>28–30</sup> This indicates a need for further research into succession-specific challenges, e.g., the possibilities of distinguishing different succession species based on their spectral and geometrical characteristics or approaches to the processing of reference data. This study is a response to these needs. Its primary goals are to (i) examine the possibility of effectively mapping succession species present in an agricultural landscape via the use of simultaneously acquired high-resolution hyperspectral and light detection and ranging (LiDAR) data and (ii) to investigate the most important features of tree-species classification using the recursive feature elimination (RFE) algorithm.

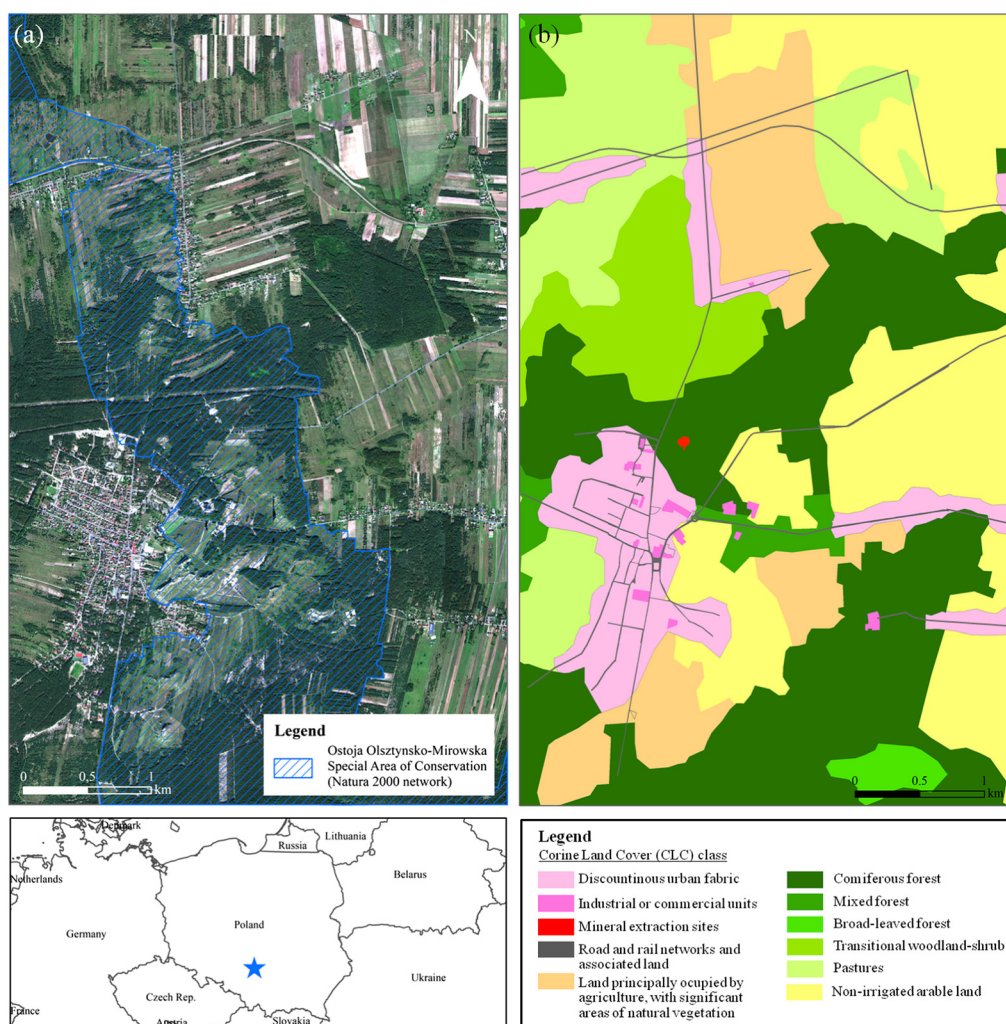
## 2 Study Area

The study area is located in the southern part of Poland, in the Silesian Voivodeship, near Czestochowa city (50°45" N; 19°17" E). It encompasses an area of about 25 km<sup>2</sup>. This area, according to the physiographic regionalization of Poland,<sup>31</sup> is located in the macroregion of the Krakow-Czestochowa Upland, the mesoregion of the Czestochowa Upland, and the microregion of the Mirowsko-Olsztynska Upland. The area is located in the Warta basin, east of Czestochowa, at the northern end of the Czestochowa Upland, being an enclave of natural and seminatural ecosystems, present between the highly urbanized areas of Silesia and Czestochowa industrial districts.

The thermal conditions of the area are typical for the climatic region of the central highlands, with an average annual temperature of 7.6°C. The coldest months of the year are January and February, and the average temperatures during these months are relatively low, approaching –3°C. The average annual air temperature amplitude is typical of western and central Poland, amounting to 20.8°C. Due to the thermal conditions, the vegetation period lasts for about 210 days per year. Snow cover, for which the durability depends on a negative air temperature, persists in the upland area for about 50 to 70 days. Precipitation is relatively high, exceeding 700 mm per year.

The area includes a complex of limestone hills (absolute heights range between 278 and 374 m) with numerous karst forms, such as caves, crosses, fissures, sinkholes, and so on. The hills are covered by natural forest communities (mostly beech and hornbeam forests) or seminatural species-rich grasslands.<sup>32</sup> The areas adjacent to the hills are occupied by cultivated fields and pine forests. The area is characterized by a wide diversity of habitats, with a considerable wealth of plant and animal species,<sup>33</sup> it is, therefore, protected as a Natura 2000 site known as Ostoja Olsztynsko-Mirowska (PLH240015) (Fig. 1).

The most important threatened nonforest habitats in this area are two types of grassland (codes 6120 and 6210), which were analyzed in this research. The first type, habitat 6120, consists of dry, frequently open grasslands on more or less calcareous sands with a subcontinental centre of distribution, comprising seminatural, moderately open to closed, relatively low-grown mesoxeric grasslands on slightly calcareous sands in the lowlands, and medium height mountains throughout temperate Europe. These grasslands are mostly dominated by tussock-forming,

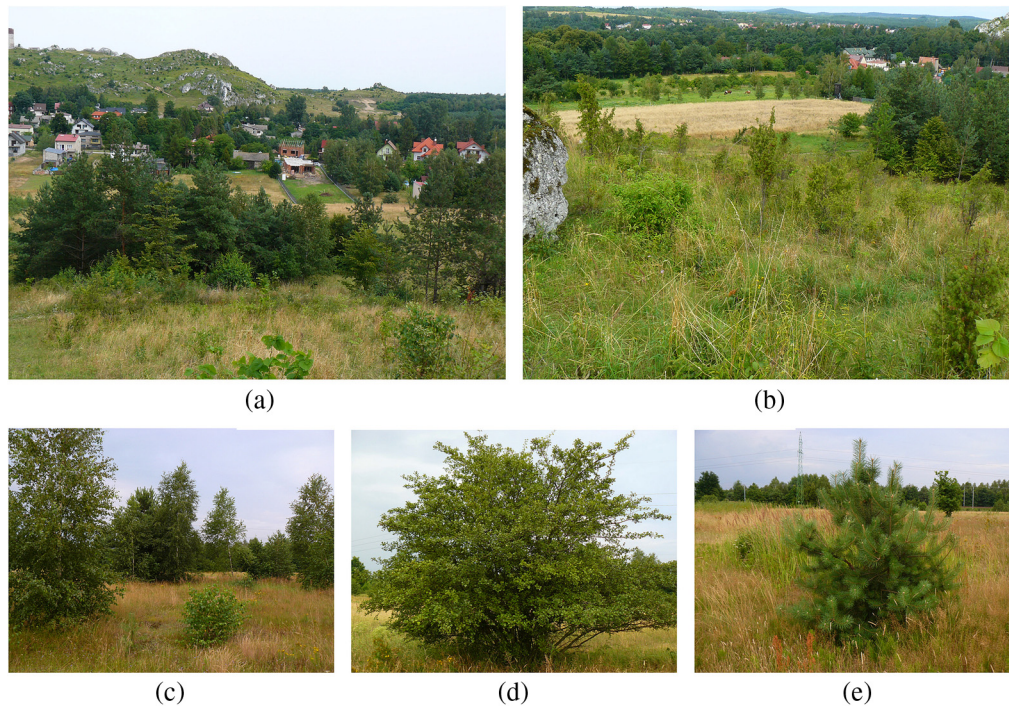


**Fig. 1** An overview of the study site: (a) the area of airborne data acquisition and botanical field measurements and (b) land use/land cover map prepared based on Corine Land Cover 2018 data. A part of the area belongs to the Ostoja Olsztynsko-Mirowska Natura 2000 protected site (hatched blue lines). The location of the study area is presented in the lower image.

narrow-leaved grasses of the *Festuca ovina* aggregate, often accompanied by *Agrostis capillaris*, *Poa angustifolia*, or *Carex praecox*. If they are in good condition, they are relatively rich in herbs, which form an important nectar source for insects.<sup>34</sup>

The grasslands of habitat 6210 are among the most species-rich plant communities in Europe in the number of plant species they support per unit area.<sup>35</sup> The calcareous grasslands of north-west Europe, for instance, consist of up to 80 plant species per square metre.<sup>32,36</sup> They also contain a large number of rare and endangered species, including the priority species listed in Annex II of the Habitats Directive, as well as various bryophytes and lichens.<sup>37</sup> In Poland, the floristic diversity of these grasslands is created by the many rare species of vascular plants, including orchids.<sup>38</sup> The invertebrate fauna associated with this habitat, particularly butterflies, is also rich<sup>37</sup> and includes a number of species listed in the Habitats Directive.<sup>39</sup>

In moderate climate regions, both types of grasslands described above depend on extensive agricultural management, and primarily grazing. The abandonment of agropastoral activities results in the development of “scrubland facies,” represented by thermophile scrub with an intermediate stage of the thermophilous fringe vegetation (*Trifolio-Geranieta*).<sup>40</sup> Calcareous grasslands evolve toward formations such as *Rosa* spp., *Prunus spinosa*, *Rhamnus cathartica*, *Crataegus* spp., *Cornus sanguinea*, *Juniperus communis*, and *Betula pendula* (10 to 15 years) and, in the succession process, toward the formation of more complex forest (of the order



**Fig. 2** Succession promoters present in the Ostoja Olsztynsko-Mirowska Natura 2000 protected site: (a) grasslands being overgrown by Scots pine, *Pinus sylvestris*, silver birch, *Betula pendula*, and common hazel, *Corylus avellana*; (b) common juniper, *Juniperus communis*, and buckthorn, *Rhamnus cathartica*, on a limestone hill; (c) silver birch, *Betula pendula*; (d) hawthorn, *Crataegus* spp.; (e) Scots pine, *Pinus sylvestris*.

*Fagetalia*), over 70 years or more.<sup>41</sup> Scrub encroachment is the most frequently documented reason for change in 6210\* sites; this is considered to be an acute threat because it can result in an increase in soil nutrients and a decline in the richness of grassland species<sup>42</sup> as the succession progresses.

An overview of the landscape and selected succession species of the study area is presented in Fig. 2.

### 3 Materials and Methods

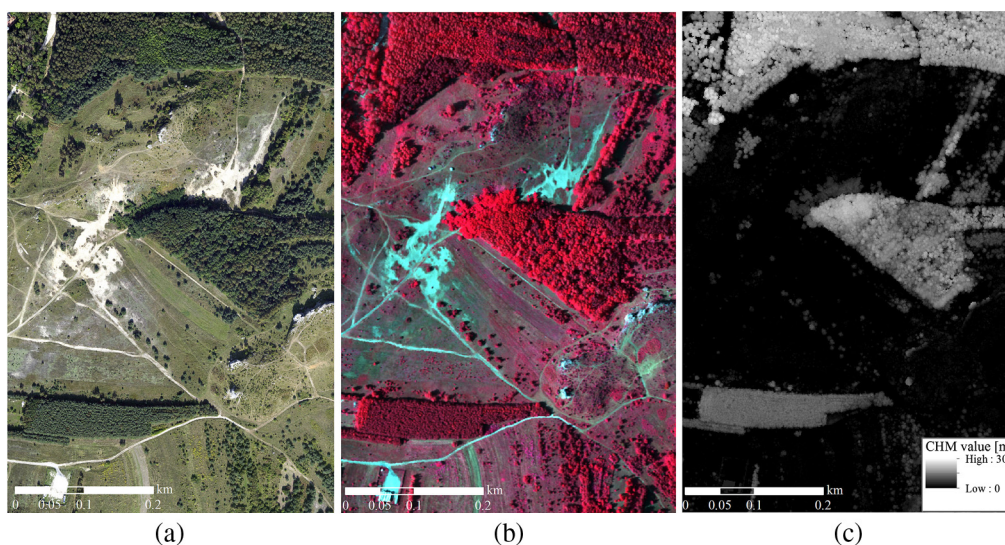
#### 3.1 Data

The research described here was carried out using two types of data: remote sensing and field data. Remotely acquired data consisted of hyperspectral images, RGB images, and LiDAR data (Fig. 3). The technical parameters characterizing the sensors and the data they provided are presented in Table 1. The collected field data could be divided into two groups: the first was spectrometer measurements that were needed for the evaluation of the atmospheric correction of hyperspectral imagery, and the second comprised botanical measurements acquired in the form of GPS points, with supplementary attributes characterizing the trees and shrubs occurring in the area under study (i.e., the height of an object, its crown density, crown radius, level of discoloration, level of defoliation, and neighboring vegetation type). The species were divided into two groups: (i) species playing a role in the early stages of the succession process, referred to as “promoters” (seven species) and (ii) nonsuccessive species referred to here as “other,” which comprised the remainder of a diverse range of species occurring in the study area. The total number of collected polygons was 406 (Fig. 4); after processing, they served as a classification reference.

It is especially important that all of the data mentioned already were acquired simultaneously. Acquiring all remote sensing data at the same time enabled subsequent uncomplicated and

**Table 1** Technical parameters for the sensors used and the data they provided. <sup>43,44</sup>

Technical parameters	Sensor type			
	Airborne laser scanner (FWF) Riegl LMS-Q680i	Hyperspectral camera HySpex		Medium format camera RGB
		VNIR-1800	SWIR-384	
Flight altitude (m)		735		
Point density (points per square meter)	7	—	—	—
Spatial resolution (m)	—	0.5	1	0.1
Spectral resolution	1.55 $\mu\text{m}$	430 bands encompassing 0.4- to 2.4- $\mu\text{m}$ spectral range		Red, green, and blue bands
Spectral sampling (nm)	—	3.26	5.45	—
FOV max (deg)	60	34	34	52
Overlap (%)	62.7	30	30	55.8
Overlap width (m)	855	450	450	720

**Fig. 3** (a) RGB orthophotomap, (b) hyperspectral imagery, and (c) DSM generated based on LiDAR data for a part of the study area.

precise (low registration error) data fusion. Remote sensing data were acquired in early autumn, between September 10 and 13, 2016, that is, at a time of intensive discoloration of the leaves, which provided the opportunity to discriminate between species. Field data were gathered as close as possible to the date of the flight: botanical field campaign measurements were made on September 29 and 30, 2016.

### 3.2 Overview of Workflow

The research involved several steps, which are presented schematically in Fig. 5.

First, the raw hyperspectral and LiDAR data were the subject of initial processing. The accuracy of the atmospheric correction of hyperspectral imagery was evaluated, making use of spectrometer measurements made on surfaces that were assumed to be unchangeable, such as asphalt. Raw RGB images were also orthorectified and the final product served for visual interpretation



Fig. 4 Location of measured trees and shrubs as referenced data.

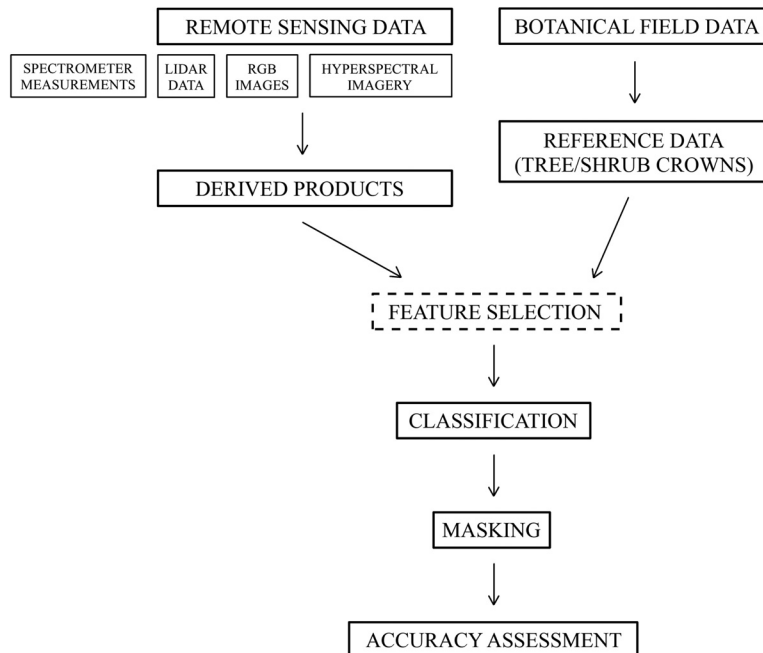


Fig. 5 The main steps of the research described in this paper.

throughout all the stages of the research. In the second step, dedicated to feature extraction, preprocessed hyperspectral and LiDAR data were used to prepare new informative products. The hyperspectral imagery alone and the datasets created by combining it with the LiDAR data are known to be powerful sources of information about vegetation and have been used in many research studies involving species discrimination.<sup>23,25,26,28–30,45</sup>

In the next stage, botanical field measurements were manually processed, which made it possible to determine the outlines of the research objects. The polygons created at this stage were later used as a reference for species classification. The supervised feature selection done in the fourth step made it possible to choose an optimal set of features for subsequent classification. The result of the classification process was then masked based on three criteria: height, land cover, and the presence of shadows. In the final step, the result was evaluated using statistical measures.

The whole process is described in detail in the following sections, which discuss each of the aforementioned issues.

### 3.2.1 Preparation of reference data

The source data for preparation of the reference polygons were botanical field measurements indicating the location and characteristics of the research objects, which were trees and shrubs. Points were collected using a differential global positioning system (DGPS) receiver, with error values of between 0.3 and 0.5 m on average and never exceeding 1 m. Although the measurement accuracy was high, the creation of a reference polygon in the form of a buffer around a point would not be an ideal approach. Bearing in mind the relationship between the measurement accuracy and the size of the object, it could be expected that several valid pixels would be omitted and some invalid ones added to a reference set in this way.

The procedure chosen in the first step assumed the creation of a buffer whose radius was determined to be the object's radius, as measured by the botany expert in the field, plus a few (2 to 3) meters. The additional enlargement of the buffer was done to account for a DGPS measurement accuracy, as well as possible little misalignments between the measurements and remote sensing data. The prepared vectors, i.e., circles, were then projected onto the pixel grid of the hyperspectral data, dividing each circle into a number of elements reflecting pixels. In the last step, these elements were visually evaluated against the presence of an object (a tree or a shrub), mixed pixels on the border of an object and its surroundings, and the presence of shadows. This work was carried out by a botany expert, who utilized many different materials: orthophotos, hyperspectral mosaic, and a crown height model (CHM). An example of the resultant polygons can be seen in Fig. 6.

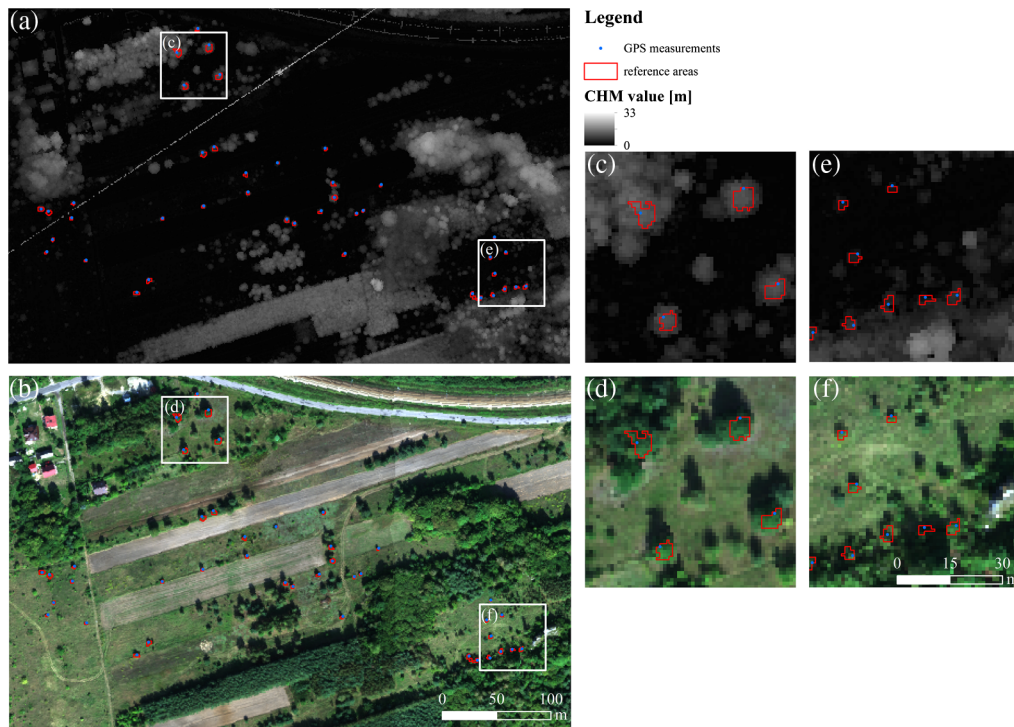
A few nonsuccessive species present on the research area occurred mostly in high forest stands, and therefore, reference polygons created for these specimens could not be based on precise dGPS measurements. However, they were highly different spectrally and in crown shape, thus easy to identify visually and vectorize on the remote sensing materials.

Finally, the prepared reference polygons were divided into training and validation samples using stratified random sampling. Each stratum was a combination of a species and the values of parameters that in the authors' opinion characterized the most important of its features: its height, crown density, and size defined by the radius. This operation ensured diversification of both the training and the validation sets for each species. The final number of reference polygons used in the classification process is presented in Table 2.

### 3.2.2 LiDAR data processing

Raw LiDAR data were preprocessed and oriented according to the standard procedure in Riegl's dedicated software RiProcess (RIEGL Laser Measurement Systems GmbH, Horn, Austria).<sup>46</sup> Full-wave signal decomposition was also performed in Riegl's software (RIEGL Laser Measurement Systems GmbH, Horn, Austria),<sup>47</sup> using a Gaussian decomposition technique to describe the recorded signal, its extracted echoes and their descriptive parameters such as the echo width and amplitude. Next, the preprocessed point clouds were classified using Axelsson's algorithm<sup>48</sup> into ground, vegetation, and unclassified, according to the ASPRS





**Fig. 6** (a) CHM, (b) hyperspectral imagery, and (c)–(f) overlaid with the reference polygons and close-up of the study area fragment.

**Table 2** The number of reference polygons used in the classification process.

Species name	Number of calibration polygons	Number of validation polygons
Silver birch, <i>Betula pendula</i>	14	18
Scots pine, <i>Pinus sylvestris</i>	15	19
Common juniper, <i>Juniperus communis</i>	20	25
Buckthorn, <i>Rhamnus cathartica</i>	13	20
Hawthorn, <i>Crataegus spp.</i>	15	19
Blackthorn, <i>Prunus spinosa</i>	13	17
Common pear, <i>Pyrus communis</i>	13	20
Other species	71	94

LAS classification standard, in Terrasolid software (Terrasolid Ltd., Helsinki, Finland).<sup>49</sup> The result of the automatic classification was manually checked using cross sections views of the point cloud in TerraScan and all detected errors were manually corrected by the airborne data provider.

In the next step, LiDAR-based rasters, an input dataset for species classification, were generated. The processing of point clouds was performed using TU Wien OPALS software<sup>50</sup> and standard rasterization methods. To calculate several raster models, several additional calculations were needed directly in the point clouds, for example for the normalized height or echo ratio parameter<sup>51</sup> of each point. The final raster datasets were divided into three groups based on the point cloud class used to create each raster and according to the scheme used in BCAL software:<sup>52</sup> an ALL group (generated based on the points of all classes except noise), a GRD group (generated based on ground points only), and a VEG group (generated based on vegetation

**Table 3** LiDAR-based raster dataset used in the experiment.

Feature type	Statistics generated	ALS classes
Amplitude	Min, max, mean, and variance	ALL, GRD, and VEG
Echo width	Min, max, mean, and variance	ALL, GRD, and VEG
nDSM	Min, max, mean, and variance	ALL and VEG
nDSM	Range, quantiles 0.05 to 0.95	VEG
DTM	Min, max, mean, and variance	GRD
Echo ratio	Min, max, mean, and variance	ALL, GRD, and VEG
Intensity	Min, max, mean, and variance	ALL, GRD, and VEG
Point density	Number of pts/m <sup>2</sup>	ALL, GRD, and VEG
Total vegetation density	Vegetation to ground points ratio	—
Vegetation cover	Vegetation to all points ratio	—
IQR, interquartile range of height	Range between quantiles 0.75 and 0.25	—
Slope	DSM-based slope map	ALL

classes points). This kind of division not only makes it possible to describe the parameters of the vegetation itself more effectively but also to gather information about other parameters depending on penetration (i.e., the number of ground returns) and the ratio of vegetation to nonvegetation points. These models include, for instance, total vegetation density, vegetation cover, or Interquartile range of height (IQR).<sup>53</sup> All of the products generated in this way are listed in Table 3. Similar features characterizing the structure of the analyzed objects were also used in research studies by Chance et al.,<sup>25</sup> Alonzo et al.,<sup>26</sup> and Shen and Cao.<sup>28</sup>

### 3.2.3 Hyperspectral data processing

The processing of hyperspectral data started with the transformation of single raw images into the useful product: a true orthomosaic. First, the images were radiometrically calibrated using Hypspec RAD software (Norsk Elektro Optikk AS, Skedsmokorset, Norway),<sup>54</sup> and a geometric correction was then applied using PARGE software (ReSe Applications GmbH, Wil, Switzerland).<sup>55</sup> Next, the images were subjected to atmospheric correction, which was done with the use of ATCOR4 software (ReSe Applications GmbH, Wil, Switzerland)<sup>56</sup> omitting corrections for terrain topography using variable water vapor and visibility estimation and predetermined aerosol type (rural). The bands covering wavelengths longer than 2.35  $\mu\text{m}$  contained high noise, were cropped, leaving 430 bands that were subjected to the spectral polishing process using the Savitzky–Golay algorithm with a range of 13 bands.<sup>57</sup> In the last step, the single corrected images were merged to produce a mosaic. The mosaicking process was carried out in PARGE (ReSe Applications GmbH, Wil, Switzerland),<sup>55</sup> using the “center cropped” option that finds the middle of the overlapping areas between the images as a cut line. As mentioned in the overview above (Sec. 3.2), the results of the atmospheric correction were evaluated based on field measurements done using ASD FieldSpec 4 spectroradiometer (ASD Inc., Longmont, Colorado). All of this work was done by the airborne data provider.

The output of the processing specified already was used in the next step to produce new, informative products. The choice on the set of products followed extensive research studies which aimed at determining the datasets that perform best in distinguishing between species. Feature extraction involved calculating the minimum noise fraction (MNF)<sup>58</sup> and vegetation indices. MNF is a linear transformation consisting of principal component analysis rotation, smoothing, rejection of the most noisy components, and finally retransformation to the original spectral space.<sup>58</sup> It was performed using ENVI software (Harris Geospatial Solutions,

Broomfield, Colorado)<sup>59</sup> and the resultant layers were put through a preselection process, in which the noisy layers were deleted and only informative layers were kept. Vegetation indices were calculated using both ENVI (Harris Geospatial Solutions, Broomfield, Colorado)<sup>59</sup> and EnMAP-Box software (The Environmental Mapping and Analysis Program, Earth Observation Center EOC of DLR, Germany),<sup>60</sup> and many different vegetation features were characterized.

### 3.2.4 Feature selection

In the next step, the products of the LiDAR and hyperspectral data processing underwent a feature selection procedure. As described by Archibald and Fann<sup>61</sup> and Singh et al.,<sup>62</sup> the application of feature selection is an important step that should be implemented before classification, as it can reduce the time taken to build the learning model and increase the accuracy of the classification result, improving its generalization ability as well. These benefits can be achieved by removing redundant and insignificant features. In this research, the feature selection process was performed using the RFE algorithm based on a random forest estimate of feature importance. The chosen algorithm is categorized in the literature as a wrapper approach.<sup>62</sup>

As a standard part of the random forest learning algorithm, an internal assessment of the variable importance is calculated. For every model learned using the random forest classifier, an estimation of the importance of each feature used for classification is available in the form of a value between 0.0 and 1.0.

When used in an RFE workflow, this possibility is used in the following way. First, a classifier is learned using the full set of features. The resulting importance of the variable (or feature) is then analyzed and typically one feature (or sometimes more) with the lowest score (i.e., the least useful in discriminating between target classes) is eliminated from the feature set. The next round of learning is then performed, eliminating the next weakest feature, and this procedure is repeated until all the features are exhausted or the accuracy of the model starts dropping significantly, which happens as the number of features gets too small to discriminate the classes.

A full set of validation metrics was calculated at each stage of this procedure (i.e., number of features, Kappa values calculated in both hard and fuzzy way, validation accuracies on the training, as well as on validation set), then the metrics from the whole procedure were plotted on a graph. Example of the selected metrics is presented in Fig. 7. The number of features considered “best” was selected after manual analysis, aiming for relatively small number of most informative features still resulting in good learning/validation performance for a given dataset.

These properties of the algorithm made it possible to choose an optimal set of features for classification, a combination of hyperspectral and LiDAR-based layers, making use of Cohen’s kappa value<sup>63</sup> to characterize each of the models.

### 3.2.5 Classification and postprocessing

As indicated in the previous sections, the classification involved making use of both the set of layers determined during the feature selection stage and the reference polygons. The classification was done using a pixel-based random forest algorithm, treating the succession species as separate classes, and the rest of the species present in the research area as one class, in a similar way to Colgan et al.<sup>45</sup> and Graves et al.<sup>23</sup>

The result obtained from a classifier was an image indicating one chosen class for each pixel, meaning that the research area was completely classified. Given the way in which the algorithm works, there was no measure that specified the degree of similarity between the model (training data) and each pixel analyzed. Therefore, areas that were outside the interest of the research could not be eliminated automatically in the same step.

The exclusion process was realized by preparing a specific mask. Four types of deletion criteria were defined:

- Anthropogenic objects (i.e., nonvegetative areas outside the analysis in the research) were selected using 0.4 threshold value of the normalized difference vegetation index (NDVI).

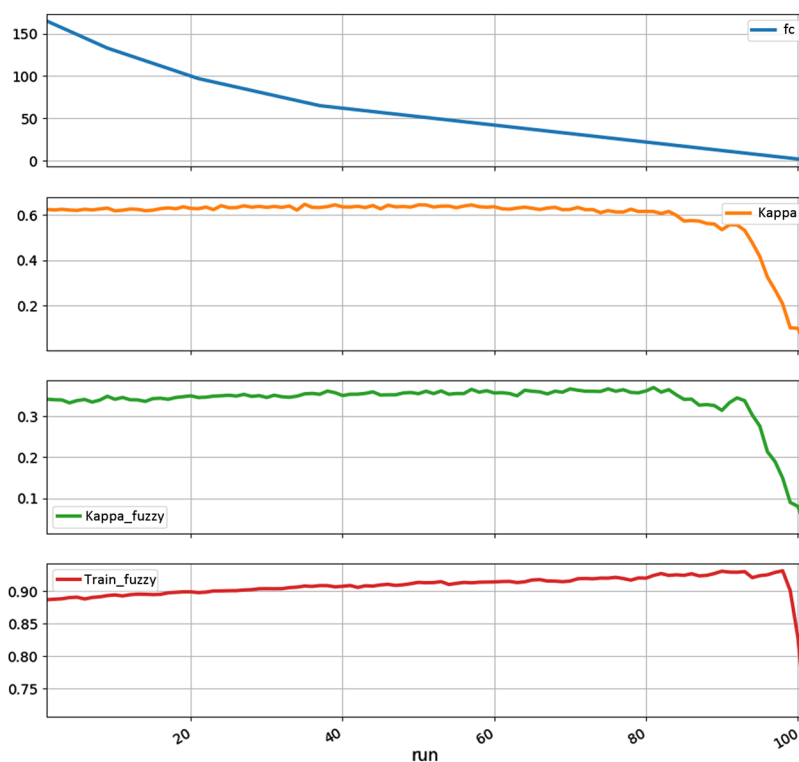


Fig. 7 Selected validation metrics used in the feature selection procedure.

- Shadows (i.e., places without proper spectral information, where a correct classification result could not be expected) were delineated using 0.2 threshold value of a selected near-infrared band (band 140–0.86  $\mu\text{m}$ ).
- Very short vegetation, that is, mainly grasses, short herbal, and agricultural species, but sometimes also young individuals representing succession species (the first group do not take part in an ongoing succession process, whereas the second could not be properly classified using data characterized by the parameters specified in Sec. 3.1), was delineated using 0.3-m threshold value of the CHM.
- High vegetation including old trees, some of which could also originate from succession but which were not of primary interest in the research, was delineated using 7-m threshold value of the CHM.

The defined threshold values were determined visually in an expert approach.

## 4 Results

### 4.1 Feature Selection Result

The procedures described in the previous section gave rise to the following products. The first one was an optimal set of features (hyperspectral and LiDAR-based raster products) for species discrimination and thus for use in classification. From the 177 features entered into the feature selection algorithm, a set of 24 were chosen, as listed in Table 4.

The set comprised 11 LiDAR-based products and 13 hyperspectral ones (five vegetation indices and eight MNF products) suggesting that both types of information are important and a diversified set is needed. Similar conclusions were drawn by Alonzo et al.,<sup>26</sup> Shen and Cao,<sup>28</sup> Colgan et al.,<sup>44</sup> and Dalponte et al.<sup>29</sup>

LiDAR-based products express the geometrical characteristics of the analyzed objects. The ability of airborne laser scanning (ALS) to penetrate the vegetation and register multiple echoes allows the determination of many features of trees and shrubs, such as the heights and shapes of

**Table 4** Set of features chosen at the feature selection stage.

No.	Feature category	Feature name (for LiDAR-based: class, type, and statistic)
1	LiDAR-based	ALL nDSM max
2	LiDAR-based	ALL nDSM mean
3	LiDAR-based	ALL nDSM min
4	LiDAR-based	ALL nDSM variance
5	LiDAR-based	ALL Echo ratio mean
6	LiDAR-based	nDSM sigma 3 sigma 0
7	LiDAR-based	nDSM sigma 5 sigma 0
8	LiDAR-based	VEG nDSM mean
9	LiDAR-based	VEG echo ratio mean
10	LiDAR-based	VEG echo ratio min
11	LiDAR-based	VEG intensity mean
12	ENVI vegetation indices	ARI2
13	EnMap-box vegetation indices	LCI
14	EnMap-box vegetation indices	LWVI2
15	EnMap-box vegetation indices	PSSrc
16	EnMap-box vegetation indices	SRtot
17	MNF	Band 3
18	MNF	Band 5
19	MNF	Band 6
20	MNF	Band 10
21	MNF	Band 11
22	MNF	Band 13
23	MNF	Band 14
24	MNF	Band 21

individuals, the relative proportions of different parts of the objects, and, indirectly, leafage-type characteristics. The use of these kinds of variables in the feature selection process makes it possible to find the unique characteristics of the species or ranges of values with which the species may be described. As the results show, discrimination between succession species in the research area was mainly achieved in the geometrical domain by the use of features representing different statistics of the normalized digital surface model (nDSM), and therefore focused on the height of the objects. Additionally, three of the chosen features were statistics of the echo ratio, which is a measure of local transparency and roughness.<sup>51</sup> Two other features that were used were also connected with roughness. Sigma 0, which is the standard deviation of the distance between the points and a fitted plane, was calculated from the nDSM with a three- or five-pixel kernel. The set of features chosen in the feature selection process also included a variable characterizing the intensity of the LiDAR signal. Since the amount of reflected and registered laser energy depends on the character of an object, this phenomenon also follows the shape and complexity of a tree and is consequently relevant in the differentiation of species.

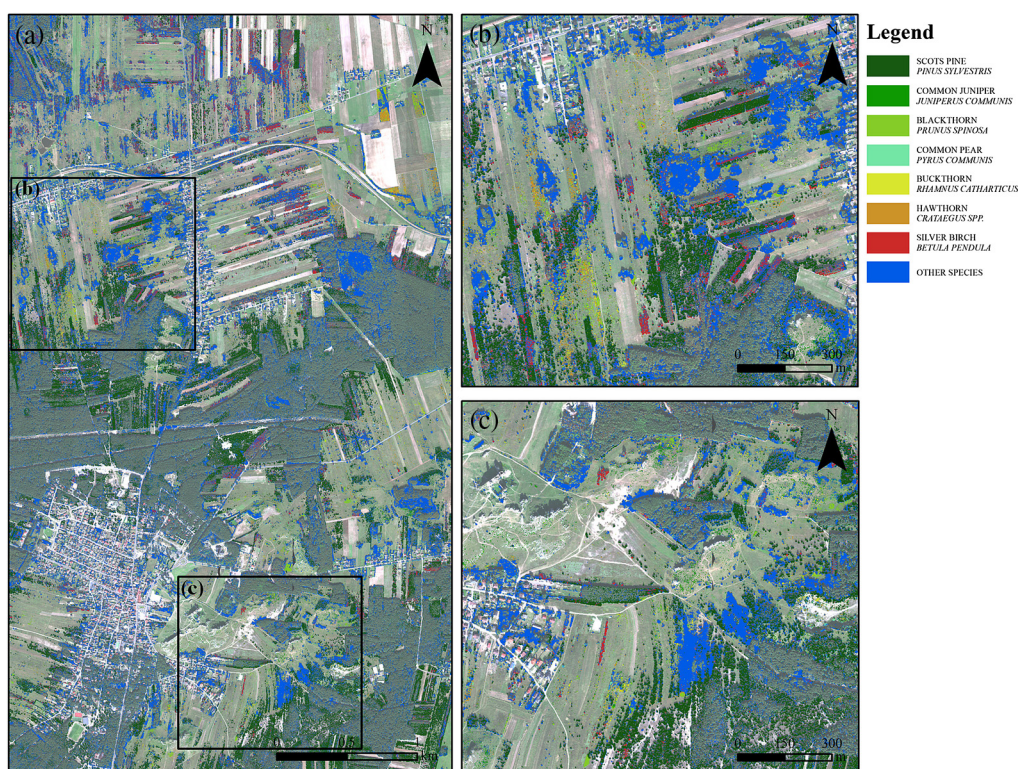
The chosen hyperspectral raster products were complementary to the characterized LiDAR-based features. Five vegetation indices turned out to be important for species differentiation: anthocyanin reflectance index 2 (ARI2),<sup>64</sup> leaf chlorophyll index (LCI),<sup>65</sup> leaf water vegetation index (LWVI2),<sup>66</sup> simple ratio (SRtot),<sup>67</sup> and pigment-specific simple ratio for carotenoids (PSSRc).<sup>68</sup> Four of these refer to vegetation characteristics that are related to the visible part of the electromagnetic spectrum, i.e., the presence of leaf pigments (chlorophylls and carotenoids). The amounts of the specified pigments are functions of time and may change differently for different species; these are, therefore, significant in the analysis. The results of the study by Alonzo et al.,<sup>26</sup> which focused on urban tree species discrimination, confirm the high information content of the visible spectral range when compared to other parts of the electromagnetic spectrum.

The final index chosen LWVI2 is connected with the amount of water present in leaves, a feature that is distinguished in the short-wave infrared part of the spectrum. The amount of water (leaf turgor) is strongly dependent on the air (temperature and humidity) and soil (primarily moisture) conditions, in which the plant grows. For individuals growing relatively close to each other in similar habitat conditions, it can be expected that the amount of water in the leaves is partially a function of the species.

Eight products of MNF transformation were difficult to analyze in detail; the spectral origin of each MNF band is not easy to trace, as it results from complex processing of source spectral bands. It can only be observed that the chosen set did not comprise the most informative layers, that is, the first ones. On the contrary, it included layers that were characterized by different information contents, proving that the features needed to discriminate between species are very specific.

## 4.2 Classification Result

As shown in the overview of the workflow (Sec. 3.2), the set of features chosen as a result of implementing the feature selection algorithm became the subject of classification. Maps showing the spatial distribution of the succession species in the research area are presented in Fig. 8.



**Fig. 8** (a) Spatial distribution of succession species in the Ostoja Olszynsko-Mirowska research area and (b) and (c) close-up of the study area fragment.

The maps reveal that the most frequently occurring succession species are *Pinus sylvestris* and *Betula pendula* (mainly in the northern part of the study area). A large part of the research area is also occupied by nonsuccessive species, referred to here as “other.” Both of the above species, *Pinus sylvestris* and *Betula pendula*, occur in compact single-species patches. In most cases, *Pinus sylvestris* is present close to the border between arable land and dense coniferous stands, whereas *Betula pendula* occurs in various places. The rest of the classified succession species do not form regularly shaped patches and are mostly scattered. In general, the trees and shrubs shaping the succession process form elongated patches corresponding to the fallows on or close to which they grow. This is obvious given that fallows, as unused land, are places with mostly undisturbed ecological processes.

### 4.3 Accuracy Assessment

The classification result was evaluated using statistical measures as part of an accuracy assessment. The report included the value of Cohen’s kappa, which characterizes the global accuracy of the product; *F1* scores, which provide information about the classification accuracy of each species (Table 5); and a confusion matrix (Table 6).

The value of Cohen’s kappa of 0.681 suggests that discrimination between succession species was relatively successful; however, individual species were classified with various levels of accuracy. The highest *F1* score of 0.919 characterized the “other” class, which could have been predicted in view of the large amount of reference data in the class and the broad range of values it encompassed. The confusion matrix indicates that this class had a very low omission error (1.9%) and a higher commission error (13.5%), spreading into other classes, especially *Pyrus communis*, *Betula pendula*, and *Prunus spinosa*. The same tendencies in the commission error of the “other” class were observed by Colgan et al.<sup>45</sup> and Graves et al.<sup>23</sup>

It can also be noticed that three succession species, *Betula pendula*, *Pinus sylvestris*, and *Prunus spinosa*, were classified with high accuracy, with *F1* scores of 0.756, 0.874, and 0.694, respectively. The four remaining species were classified with average accuracy, ranging from *F1* = 0.451 for *Rhamnus cathartica* to *F1* = 0.603 for *Juniperus communis*.

## 5 Discussion

In this study, a methodology for identifying succession species in an agricultural landscape was proposed and evaluated. The results obtained here indicate that the selected approach is valid, although there are several points that can be further developed.

**Table 5** *F1* scores and Cohen’s kappa.

Species name	<i>F1</i> score
Silver birch, <i>Betula pendula</i>	0.756
Scots pine, <i>Pinus sylvestris</i>	0.874
Common juniper, <i>Juniperus communis</i>	0.603
Buckthorn, <i>Rhamnus cathartica</i>	0.451
Hawthorn, <i>Crataegus spp.</i>	0.456
Blackthorn, <i>Prunus spinosa</i>	0.694
Common pear, <i>Pyrus communis</i>	0.531
Other species	0.919
Cohen’s kappa	0.681

**Table 6** Confusion matrix.

	Other species	Common juniper	Scots pine	Hawthorn	Buckthorn	Common pear	Silver birch	Blackthorn	Total	Producer accuracy (%)
Other species	3909	8	3	15	2	3	17	27	<b>3984</b>	98.1
Common juniper	13	73	1	9	4	0	0	11	<b>111</b>	65.8
Scots pine	45	1	264	3	0	6	0	1	<b>320</b>	82.5
Hawthorn	94	3	2	95	4	6	0	2	<b>206</b>	46.1
Buckthorn	28	10	1	28	48	1	0	27	<b>143</b>	33.6
Common pear	161	0	2	37	3	129	7	1	<b>340</b>	37.9
Silver birch	116	0	0	0	0	1	219	0	<b>336</b>	65.2
Blackthorn	154	36	11	24	9	0	0	344	<b>578</b>	59.5
Total	<b>4 520</b>	<b>131</b>	<b>284</b>	<b>211</b>	<b>70</b>	<b>146</b>	<b>243</b>	<b>413</b>		
User accuracy (%)	86.5	55.7	93.0	45.0	68.6	88.4	90.1	83.3		

The first of these is the reference set, which plays a critical role in the classification. The way in which this is prepared affects all stages of processing. The reference polygons determine the training (class definition) and validation (selection of an optimal set of features during the feature selection step, determination of accuracy assessment validity) data. The level of difficulty characterizing the preparation of an appropriate reference set in the agricultural environment can be realized by comparing the nature and size of research objects (trees and shrubs) with the economically viable spatial resolution of the airborne imagery. The monitoring of succession is mainly focused on small vegetation identified at an early stage of the process. The sizes of the objects to be classified are predominantly equal to four or nine times the size of a pixel and represent crown diameters of 1, 2, or 3 m, respectively. Additionally, every individual can be characterized by a different crown density, usually ranging from 30% to 100% at the time of airborne data acquisition. As a result of these two issues and the fact that borders between neighboring image elements (pixel grid) are created randomly in relation to the research objects, a relatively small number of pure pixels and a large number of mixed ones are present. The spectral response from the mixed pixels is a different proportion combination of the signal from the analyzed tree or shrub and other objects present below or next to the crown, such as the ground or shorter or neighboring vegetation. The selection of pixels that have been sufficiently well defined in the spectral and geometrical domains for addition to the reference set is, therefore, very problematic, and it is important for future studies to specify the minimal values for key parameters of the research objects (e.g., crown density) that will allow for a good definition. Determining these lower thresholds will help make the consumer of the final product aware of the map's information content. It can also support botany experts in planning field campaigns by identifying objects that are suitable or unsuitable for measurement.

As explained in Sec. 3.2.1, the reference polygons used in this research were created manually by selecting groups of valid pixels (in the expert's opinion) based on field data (points indicating the object's location) and remote sensing data. This approach has both advantages and drawbacks. On the one hand, the manually prepared outline of a crown is likely to be more precise than one created with the use of an automatic tool. On the other hand, the manual



treatment of mixed pixels may be inconsistent between different individuals, while an automated tool would act in a reliable way and would create an approach that is better fitted to the data. Future research on this subject should be focused on testing different segmentation methods or other automated procedures. The results of research studies by Alonzo et al.,<sup>26</sup> Graves et al.,<sup>23</sup> Colgan et al.,<sup>45</sup> and Shen and Cao<sup>28</sup> suggest that these techniques can be used for the correct and efficient delineation of single trees based on the remote sensing data.

The most problematic issue concerning the classification stage seems to be the definition of classes. The approach applied in the current work is directly related to the project's needs: each succession species was treated as a separate class, whereas the rest of the vegetation occurring in the study area was placed in the "other" class. The results of the accuracy assessment suggest that a few species are very similar to each other, based on remote sensing measures, and that merging these should probably be considered in the future. Incorporating an automatic analysis of the separability of the classes into the methodology may be an optimal way of dealing with this problem. This separability analysis needs to be done individually, not only for different groups of species but also for other individuals (due to intraspecies diversity) and obviously for remote sensing data acquisition dates, as a result of the vegetation's phenological stages over time.

The postprocessing of the final product is also a challenging task. An image resulting from the application of the random forest classifier characterizes each pixel as belonging to a single chosen class; that is, the research area is entirely classified. The correct delineation of the spatial extent of the species, therefore, requires further processing. This can be done in several different ways:

- By collecting the reference polygons for each of the classes present in the research area, this may be time-consuming and therefore expensive, especially if the study site is characterized by high biodiversity, i.e., a large number of species.
- By creating a special mask, i.e., a layer containing the objects and areas for which a reference was not provided and which are therefore to be deleted from the classification map. This method has been proven to be effective for eliminating objects and areas that are relatively easy to define, such as anthropogenic areas or water, but is often useless for deleting nonsuccessive species that are very similar to the research objects.
- Using a measure of similarity between the model (training set) and an analyzed pixel. The calculated values can be thresholded to select pixels with low similarity values, i.e., those that should not be included in any of the analyzed classes. This is the most straightforward approach, but is not always available for a given classifier, for example, the random forest algorithm used in this research. This problem was also discussed by Alonzo et al.<sup>26</sup>
- By utilizing a combination of the aforementioned approaches.

As indicated in Sec. 3.2.5, the classification map was postprocessed using a combined approach. Different nonvegetative objects and areas were delineated using a special mask, while pixels representing nonsuccessive species were chosen by creating an "other" class from a selection of representative species. This approach is easy to implement: it does not require any complicated calculations and can be applied without collecting a reference for each species. However, it often overestimates the "other" class because of its high variance. The issue of the definition of the "other" class should be studied further in the future.

## 6 Conclusion

The aim of this research was to determine the potential of an accurate identification of succession species present in an agricultural landscape with the use of high-resolution aerial remote sensing data. Promising results were obtained from classifying a dataset comprised of hyperspectral and LiDAR-based products using the random forest algorithm. The use of the RFE algorithm also enabled us to observe that both data sources are important in the correct discrimination of species. The operational use of the methodology presented here requires several elements to be developed further; these are the preparation procedure for the reference polygons and the definition of the classes used in classification.

## Acknowledgments

The authors would like to express their gratitude to MGGP Aero Sp. z o.o. for preprocessing the hyperspectral and LiDAR data and to Justyna Wylazłowska for helping with the botanical field campaign. This research was partially funded by the Poland National Centre for Research and Development. Three types of experts are involved in the project's implementation: airborne remote sensing data provider (Consortium Leader: company MGGP Aero), botany experts (project partners, scientific units: University of Lodz, Institute of Technology and Life Sciences, and University of Silesia in Katowice), and remote sensing specialists (project partners, scientific units: Warsaw University of Technology, University of Warsaw, and Warsaw University of Life Sciences). The authors declare no conflict of interest.

## References

1. N. Kolecka et al., "Mapping secondary forest succession on abandoned agricultural land with LiDAR point clouds and terrestrial photography," *Remote Sens.* **7**, 8300–8322 (2015).
2. M. Szostak et al., "Monitoring the secondary forest succession and land cover/use changes of the Błędów Desert (Poland) using geospatial analyses," *Quaestiones Geographicae* **35**(3), 1–13 (2016).
3. M. Szostak, P. Hawryło, and D. Piela, "Using of Sentinel-2 images for automation of the forest succession detection," *Eur. J. Remote Sens.* **51**(1), 142–149 (2017).
4. S. Martinuzzi et al., "Quantifying tropical dry forest type and succession: substantial improvement with LiDAR," *Biotropica* **45**(2), 135–146 (2013).
5. M. J. Falkowski et al., "Characterizing forest succession with lidar data: an evaluation for the Inland Northwest, USA," *Remote Sens. Environ.* **113**, 946–956 (2009).
6. M. Castillo et al., "LiDAR remote sensing for secondary tropical dry forest identification," *Remote Sens. Environ.* **121**, 132–143 (2012).
7. M. S. Vanderlinder et al., "Use of remote sensing to assess changes in wetland plant communities over an 18-year period: a case study from the Bear River Migratory Bird Refuge, Great Salt Lake, Utah," *West. N. Am. Nat.* **74**(1), 33–46 (2014).
8. J. A. Gallardo-Cruz et al., "Predicting tropical dry forest successional attributes from space: is the key hidden in image texture?" *PLoS One* **7**(2), e30506 (2012).
9. U. J. Sánchez-Reyes et al., "Assessment of land use-cover changes and successional stages of vegetation in the natural protected area Altas Cumbres, Northeastern Mexico, using Landsat satellite imagery," *Remote Sens.* **9**(7), 712 (2017).
10. M. Batistella and D. Lu, "Integrating field data and remote sensing to identify secondary succession stages in the Amazon," 2002, [https://www.researchgate.net/publication/255609982\\_Integrating\\_Field\\_Data\\_and\\_Remote\\_Sensing\\_to\\_Identify\\_Secondary\\_Succession\\_Stages\\_in\\_the\\_Amazon](https://www.researchgate.net/publication/255609982_Integrating_Field_Data_and_Remote_Sensing_to_Identify_Secondary_Succession_Stages_in_the_Amazon) (accessed 15 August 2018).
11. I. C. G. Vieira et al., "Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia," *Remote Sens. Environ.* **87**, 470–481 (2003).
12. V. E. Garcia-Millan, G. A. Sanchez-Azofeifa, and G. C. Malvárez, "Mapping tropical dry forest succession with CHRIS/PROBA hyperspectral images using nonparametric decision trees," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **8**(6), 3081–3094 (2015).
13. G. Çakir et al., "Mapping secondary forest succession with geographic information systems: a case study from Bulanikdere, Kirklareli, Turkey," *Turk J. Agric. For.* **31**(1), 71–81 (2006).
14. B. Barabasz-Krasny, "Vegetation differentiation and secondary succession on abandoned agricultural large-areas in south-eastern Poland," *Biodiv. Res. Conserv.* **41**, 35–50 (2016).
15. P. Poschlod, J. P. Bakker, and S. Kahmen, "Changing land use and its impact on biodiversity," *Basic Appl. Ecol.* **6**, 93–98 (2005).
16. D. Van der Hoek and K. V. Sykora, "Fen meadow succession in relation to spatial and temporal differences in hydrology and soil conditions," *Appl. Veg. Sci.* **9**, 185–194 (2006).
17. Á. J. Albert et al., "Secondary succession in sandy old-fields: a promising example of spontaneous grassland recovery," *Appl. Veg. Sci.* **17**, 214–224 (2014).

18. Z. Kaćki and D. Michalska-Hejduk, "Assessment of biodiversity in Molinia Meadows in Kampinoski National Park based on biocenotic indicators," *Polish J. Environ. Stud.* **19**(2), 351–362 (2010).
19. A. Suder, "Purple-moor grass meadows (alliance Molinion caeruleae Koch 1926) in the eastern part of Silesia Upland: phytosociological diversity and aspects of protection," *Nat. Conserv.* **65**, 63–77 (2008).
20. F. Neves et al., "Successional and seasonal changes in a community of dung beetles (Coleoptera: Scarabaeinae) in a Brazilian tropical dry forest," *Nat. Conserv.* **8**, 160–164 (2010).
21. M. Quesada et al., "Succession and management of tropical dry forests in the Americas: review and new perspectives," *For. Ecol. Manage.* **258**, 1014–1024 (2009).
22. J. Tews, "Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures," *J. Biogeogr.* **31**, 79–92 (2004).
23. S. J. Graves et al., "Tree species abundance predictions in a tropical agricultural landscape with a supervised classification model and imbalanced data," *Remote Sens.* **8**, 161 (2016).
24. N. E. Zimmermann et al., "Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah," *J. Appl. Ecol.* **44**, 1057–1067 (2007).
25. C. M. Chance et al., "Invasive shrub mapping in an urban environment from hyperspectral and LiDAR-derived attributes," *Front. Plant Sci.* **7**, 1–19 (2016).
26. M. Alonzo, B. Bookhagen, and D. A. Roberts, "Urban tree species mapping using hyperspectral and LIDAR data fusion," *Remote Sens. Environ.* **148**, 70–83 (2014).
27. L. Liu et al., "Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data," *Remote Sens. Environ.* **200**, 170–182 (2017).
28. X. Shen and L. Cao, "Tree-species classification in subtropical forests using airborne hyperspectral and LiDAR data," *Remote Sens.* **9**, 1180 (2017).
29. M. Dalponte, L. Bruzzone, and D. Gianelle, "Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas," *IEEE Trans. Geosci. Remote Sens.* **46**(5), 1416–1427 (2008).
30. G. P. Asner et al., "Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR," *Remote Sens. Environ.* **112**, 1942–1955 (2008).
31. J. Kondracki, *Geografia Regionalna Polski*, p. 440, Wydawnictwo Naukowe PWN, Warszawa (1998).
32. M. Ruskiewicz-Michalska and D. Michalska-Hejduk, "Grzyby pasożytnicze roślin charakterystycznych dla muraw kserotermicznych Festuco-Brometea w okolicach Olsztyna koło Częstochowy [Phytopathogenic fungi of characteristic plant species of Festuco-Brometea xerothermic grasslands in the vicinity of Olsztyn near Częstochowa]," *Ziemia Częstochowska* **30**, 195–208 (2003).
33. "Standard data form for Natura 2000 site PLH240015," *Olsztyńsko-Mirowska Refuge*, <http://n2k-ws.gdos.gov.pl/wyszukiwarkaN2k/webresources/pdf/PLH240015> (accessed 22 January 2018).
34. J. A. M. Janssen and J. Dengler, "E1.9a oceanic to subcontinental inland sand grassland on dry acid and neutral soils," European Red List of Habitats—Grasslands Habitat Group, [https://www.researchgate.net/publication/312297575\\_E19a\\_Oceanic\\_to\\_subcontinental\\_inland\\_sand\\_grassland\\_on\\_dry\\_acid\\_and\\_neutral\\_soils](https://www.researchgate.net/publication/312297575_E19a_Oceanic_to_subcontinental_inland_sand_grassland_on_dry_acid_and_neutral_soils) (accessed 31 May 2017).
35. B. Calaciura and O. Spinelli, "Management of Natura 2000 habitats. 6210 Seminal dry grasslands and land facies on calcareous substrates (Festuco-Brometalia) (\*important orchid sites)," European Commission. Technical Report 12(24) (2008).
36. M. F. W. DeVries, P. Poschlod, and J. H. Willems, "Challenges for the conservation of calcareous grasslands in Northwestern Europe: integrating the requirements of flora and fauna," *Biol. Conserv.* **104**, 265–273 (2002).
37. *The Convention on the Conservation of European Wildlife and Natural Habitats*, European Treaty Series 104, Join Nature Conservation Committee, Peterborough, <https://rm.coe.int/1680078aff> (accessed 22 January 2018).
38. W. Mróz, *Monitoring siedlisk przyrodniczych. Przewodnik metodyczny [Natura 2000 Habitat Monitoring], Part I–III*, Biblioteka Monitoringu Środowiska, Warszawa, Poland (2010–2012).

39. S. Colas and M. Hébert, “Guide d’estimation des coûts de gestion des milieux naturels ouverts,” [Guide for estimating the costs of managing open natural environments], Espaces Naturels de France (2000).
40. *Habitats Directive Annex I Habitats*, European Commission, Interpretation Manual of European Union Habitats, version EUR 28, 2013, [http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/Int\\_Manual\\_EU28.pdf](http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/Int_Manual_EU28.pdf) (accessed 22 January 2018).
41. M. Partel et al., “Restoration of species-rich limestone grassland communities from overgrown land: the importance of propagule availability,” *Ecol. Eng.* **10**, 275–286 (1998).
42. The Invertebrate Conservation Trust, “Advice on managing BAP habitats,” *Upland Calcareous Grassland*, 2007, <https://www.buglife.org.uk/advice-and-publications/advice-on-managing-bap-habitats/upland-calcareous-grassland> (accessed 22 January 2018).
43. “HySpex products,” <https://www.hyspex.no/products/> (accessed 22 February 2018).
44. “Datasheet Riegl LMS-Q680i,” 2012, [http://www.riegl.com/uploads/tx\\_pxpriegldownloads/10\\_DataSheet\\_LMS-Q680i\\_28-09-2012\\_01.pdf](http://www.riegl.com/uploads/tx_pxpriegldownloads/10_DataSheet_LMS-Q680i_28-09-2012_01.pdf) (accessed 22 January 2018).
45. M. S. Colgan et al., “Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data,” *Remote Sens.* **4**, 3462–3480 (2012).
46. *RiProcess Data Sheet for RIEGL Scan Data*, [http://www.riegl.com/uploads/tx\\_pxpriegldownloads/11\\_Datasheet\\_RiProcess\\_2016-09-16\\_01.pdf](http://www.riegl.com/uploads/tx_pxpriegldownloads/11_Datasheet_RiProcess_2016-09-16_01.pdf) (20 September 2018).
47. *RiAnalyze Data Sheet for Automated Resolution of Range Ambiguities*, [http://www.riegl.com/uploads/tx\\_pxpriegldownloads/11\\_DataSheet\\_RiMTA-ALS\\_2015-08-24\\_03.pdf](http://www.riegl.com/uploads/tx_pxpriegldownloads/11_DataSheet_RiMTA-ALS_2015-08-24_03.pdf) (accessed 20 September 2018).
48. P. Axelsson, “Processing of laser scanner data—algorithms and applications,” *ISPRS J. Photogramm. Remote Sens.* **54**(2), 138–147 (1999).
49. “TerraSolid Terrascan user guide,” <http://www.terrasolid.com/guides/tscan/index.html> (accessed 20 July 2018).
50. G. Mandlbürger et al., “OPALS—a comprehensive laser scanning software for geomorphological analysis,” in *Geophys. Res. Abstr.*, Vol. 13, EGU2011-7933, EGU General Assembly (2011).
51. B. Höfle et al., “Detection of building regions using airborne LiDAR: a new combination of raster and point cloud based GIS methods,” in *Proc. GI-Forum—Int. Conf. Appl. Geoinf.*, Salzburg, Austria, pp. 66–75 (2009).
52. BCAL, *BCAL Lidar Tools*, Boise, Idaho, 2016, <https://bcal.boisestate.edu/tools/lidar> (accessed 22 January 2018).
53. J. S. Evans et al., “Discrete return Lidar in natural resources: recommendations for project planning, data processing, and deliverables,” *Remote Sens.* **1**(4), 776–794 (2009).
54. HySpex RAD, <https://www.hyspex.no/> (accessed 26 October 2018).
55. D. Schläpfer, “Parametric Geocoding, PARGE User Guide, Version 3.3, ReSe Applications Schläpfer, p. 270 (2016).
56. R. Richter and D. Schlapfer, “ATCOR4 manual,” ReSe Applications, [https://www.rese-apps.com/pdf/atcor4\\_manual.pdf](https://www.rese-apps.com/pdf/atcor4_manual.pdf) (accessed 26 October 2018).
57. Ł. Sławik et al., “Multiple flights or single flight instrument fusion of hyperspectral and ALS data? A comparison of their performance for vegetation mapping,” *Remote Sens.* **11**, 970 (2019).
58. A. A. Green et al., “A transformation for ordering multispectral data in terms of image quality with implications for noise removal,” *IEEE Trans. Geosci. Remote Sens.* **26**(1), 65–74 (1988).
59. “ENVI API programming guide,” Harris Geospatial Solutions Documentation Center, <http://www.harrisgeospatial.com/docs/ProgrammingGuideIntroduction.html> (accessed 21 September 2018).
60. A. Rabe et al., “EnMAP-Box 3 documentation,” Earth Observation Center EOC of DLR, Germany, <http://www.enmap.org/enmapbox.html> (accessed 12 December 2018).
61. R. Archibald and G. Fann, “Feature selection and classification of hyperspectral images with support vector machines,” *IEEE Geosci. Remote Sens. Lett.* **4**(4), 674–677 (2007).
62. A. A. G. Singh, A. Balamurugan, and J. Leavline, “Literature review on feature selection methods for high-dimensional data,” *Int. J. Comput. Appl.* **136**(1), 9–17 (2016).

63. R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, CRC Press, Taylor & Francis Group, Boca Raton (2008).
64. A. A. Gitelson, M. N. Merzlyak, and O. B. Chivkunova, "Optical properties and non-destructive estimation of anthocyanin content in plant leaves," *Photochem. Photobiol.* **74**(1), 38–45 (2001).
65. B. Datt, "A new reflectance index for remote sensing of chlorophyll content in higher plants: tests using eucalyptus leaves," *J. Plant Physiol.* **154**, 30–36 (1999).
66. L. S. Galvao, A. R. Formaggio, and D. A. Tisot, "Discrimination of sugarcane varieties in Southeastern Brazil with EO-1 Hyperion data," *Remote Sens. Environ.* **94**(4), 523–534 (2005).
67. B. Datt, "Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+b and total carotenoid content in eucalyptus leaves," *Remote Sens. Environ.* **66**, 111–121 (1998).
68. G. A. Blackburn, "Spectral indices for estimating photosynthetic pigment concentration: a test using senescent tree leaves," *Int. J. Remote Sens.* **19**(4), 657–675 (1998).

**Aleksandra Radecka** received her BSc and MSc degrees in spatial planning from Warsaw University of Technology (WUT) in Poland, in 2015 and 2017, respectively. She is currently pursuing her PhD in geodesy and cartography at WUT. Since 2016, she has worked as a research assistant in the Department of Photogrammetry, Remote Sensing and Spatial Information Systems, WUT. Her research interest includes the usage of thermal and hyperspectral remote sensing techniques for environmental studies.

**Dorota Michalska-Hejduk** received her MSc degree in biology and her PhD in ecology from the University of Lodz in Poland where she is working in the Department of Geobotany and Plant Ecology. Her research interests include plant science, nature protection, monitoring of Natura 2000 habitats and species, and ecological processes. In 2014, she received an award for services to "Environmental Protection and Water Management" from the Ministry of the Environment in Poland. She is a member of the Society of Ecological Restoration.

**Katarzyna Osińska-Skotak** received her MSc degree in photogrammetry and cartography in 1994, her MSc degree in meteorology and atmosphere protection in 1997, her PhD in 2001, and her DSc degree in photogrammetry and remote sensing in 2011 from Warsaw University of Technology (WUT) in Poland. Since 2012, she has been a head of the Department of Photogrammetry, Remote Sensing and Spatial Information Systems at WUT. Her research interests include the applications of thermal, multi- and hyperspectral remote sensing techniques in environmental and urban monitoring. She is a board member of the Remote Sensing and Geoinformatics Group of the Polish Geographical Society.

**Adam Kania** has experience in private consultancy for business in areas such as environmental management, quality management systems, IT management, and software development. His areas of interest also included data processing, statistics, machine learning, optimization, and AI algorithms. He is currently leading the development of the Vegetation Classification Studio—a scientific software tool for fast and effective classification of remote sensing data. Last year, he founded Definity Ltd., a company working in research and development of algorithms and tools for AI, machine learning, remote sensing, and scientific computing.

**Konrad Górski** received his MSc degree in photogrammetry and remote sensing in 2016 from Warsaw University of Technology (WUT) in Poland. He is now a PhD student at the WUT Faculty of Geodesy and Cartography. From 2016 to 2018, he has been working in the Department of Photogrammetry, Remote Sensing and Spatial Information Systems at WUT. His research interests include employing new sensors and processing techniques in a wide range of specialized applications, especially modern ultralight sensors compatible with unmanned platforms, to enhance industry capabilities.

**Wojciech Ostrowski** received his MSc degree in geodesy and cartography (specialization: photogrammetry and remote sensing) in 2013 from Warsaw University of Technology (WUT) in Poland, where he has been working as a research-teaching assistant in the Department of

Photogrammetry, Remote Sensing and Spatial Information Systems, since 2013. His research interests include LIDAR data applications in archaeology prospection and remote sensing as well as photogrammetric processing of UAV and aerial oblique images. He is a member of the Polish Society for Photogrammetry and Remote Sensing, and the Computer Applications and Quantitative Methods in Archaeology and Aerial Archaeology Research Group.