Title: Hyperspectral vs. Multispectral data: Comparison of the spectral differentiation capabilities of Natura 2000 non-forest habitats

Author: Anna Jarocińska, Dominik Kopeć, Marlena Kycko, Hubert Piórkowski, Agnieszka Błońska

Citation style: Jarocińska Anna, Kopeć Dominik, Kycko Marlena, Piórkowski Hubert, Błońska Agnieszka. (2022). Hyperspectral vs. Multispectral data: Comparison of the spectral differentiation capabilities of Natura 2000 non-forest habitats. „ISPRS Journal of Photogrammetry and Remote Sensing” (2022, vol. 184, s. 148-164), DOI: 10.1016/j.isprsjprs.2021.12.010

ISPRS Journal of Photogrammetry and Remote Sensing 184 (2022) 148–164

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Hyperspectral vs. Multispectral data: Comparison of the spectral differentiation capabilities of Natura 2000 non-forest habitats

Anna Jarocińska [a], Dominik Kopeć [b,c,*], Marlena Kycko [a], Hubert Piórkowski [d], Agnieszka Błońska [e]

[a] Department of Geoinformatics, Cartography and Remote Sensing, Chair of Geomatics and Information Systems, Faculty of Geography and Regional Studies, University of Warsaw, 00-927 Warsaw, Poland
[b] Department of Biogeography, Paleoecology and Nature Conservation, Faculty of Biology and Environmental, University of Lodz, 90-237 Łódź, Poland
[c] MGGP Aero sp. z o.o., 33-100 Tarnów, Poland
[d] Institute of Technology and Life Sciences - State Research Institute, 05-090 Raszyn, Poland
[e] Institute of Biology, Biotechnology and Environmental Protection, Faculty of Natural Sciences, University of Silesia in Katowice, 40-032 Katowice, Poland

ARTICLE INFO

Keywords:
Linear Discriminant Analysis
HySpex
Sentinel-2
Vegetation identification

ABSTRACT

Identification of the Natura 2000 habitats using remote sensing techniques is one of the most important challenges of nature conservation. In this study, the potential for differentiating non-forest Natura 2000 habitats from the other habitats was examined using hyperspectral data in the scope of VNIR (0.4–1 μm), SWIR (1–2.5 μm) and simulated multispectral data (Sentinel-2). The aim of the research was also to determine the most informative spectral ranges from the optical range. Five different Natura 2000 habitats common in Central Europe were analysed: heaths (code 4030), mires (code 7140), grasslands (code 6230) and meadows (codes 6410 and 6510). In order to guarantee the objectivity and transferability of the results each habitat was tested in two areas and in three campaigns (spring, summer, autumn). Hyperspectral data was acquired using HySpex VNIR-1800 and SWIR-384 scanners. The Sentinel-2 data was resampled based on HySpex spectral reflectance. The overflights were performed simultaneously with ground reference data – habitats and background polygons. The Linear Discriminant Analysis was performed in iterative mode based on spectral reflectance acquired from hyperspectral and multispectral data. This resulted in distribution of correctness rate values and information about the most differentiating spectral bands for each habitat. Based on the results of our experiments we conclude that: (i) hyperspectral data (both VNIR and SWIR) obtained from May to September was useful for differentiation of habitats from background with efficiency reaching over 90%, regardless of the area; (ii) the most useful spectral ranges are: in VNIR − 0.416–0.442 μm and 0.502–0.522 μm, in SWIR − 1.117–1.165 μm and 1.290–1.361 μm; (iii) the potential of multispectral data (Sentinel-2) in distinguishing Natura 2000 habitats from the background is diverse; higher for heaths and mires (comparable to hyperspectral data) lower for meadows (6410, 6510) and grasslands (6230); (iv) in case of meadows and grasslands, the correctness rate for the Sentinel-2 data was on average about 20% lower compared to the hyperspectral data.

1. Introduction

The Pan-European Ecological Network Natura 2000 is nowadays one of the most important initiatives supporting biodiversity protection in all of the EU Member States. Based on the Council Directive 92/43/EEC, Natura 2000 areas are created to preserve the most valuable natural habitats in the EU. One of the challenges for nature conservation is monitoring the conservation status of these habitats and searching for new potential areas that should be protected. This is a perfect application where remote sensing tools proved to be very useful for both mapping and monitoring the Natura 2000. The EU supports the identification of habitats based on aerial or satellite images (Ichter et al.,

Abbreviations: HS, Hyperspectral; MS, Multispectral; LDA, Linear Discriminant Analysis; ALS, Airborne Laser Scanner.

* Corresponding author at: Department of Biogeography, Paleoecology and Nature Conservation, Faculty of Biology and Environmental, University of Lodz, 90-237 Łódź, Poland.

E-mail addresses: ajarocinska@uw.edu.pl (A. Jarocińska), dominik.kopec@biol.uni.lodz.pl (D. Kopeć), marlenakycko@uw.edu.pl (M. Kycko), h.piorkowski@itp.edu.pl (H. Piórkowski), agnieszka.blonska@us.edu.pl (A. Błońska).

https://doi.org/10.1016/j.isprsjprs.2021.12.010
Received 1 September 2021; Received in revised form 4 November 2021; Accepted 31 December 2021
Available online 11 January 2022

2014) and has been developing very dynamically in recent years. The first classifications of the Natura 2000 habitat based on remote sensing images was performed in 2008 (Díaz Varela et al., 2008). Consecutively, a number of studies has been conducted in order to identify or model the habitat patches based on remote sensing data acquired at different seasons and from various altitudes: UAV (Belcore et al., 2021), airborne (Sławik et al., 2019) or satellite (Tarantino et al., 2021).

In previous studies, different types of data were also employed to identify habitats, including airborne laser scanning (ALS) data and multispectral (MS) and hyperspectral (HS) images. High classification accuracies were achieved based on different sensors, e.g.: HS data (Haest et al., 2017), fusion of HS and ALS data (Hladik et al., 2013), using multitemporal HS (Marcinkowska-Ochtyra et al., 2019) or MS datasets: RapidEye and Sentinel-2 (Rapinel et al., 2020; Schuster et al., 2015; Stenzel et al., 2014). Moreover, single ALS data was also used successfully (Zlinszky et al., 2014). The acquired classification accuracies displayed high diversity depending on the habitats, but also within a single habitat. Generally, in case of heathlands, accuracies varied from F1 = 28% based on MS data (Haest et al., 2010) to 83% using RGB camera and DTM (Gonçalves et al., 2016); for meadows F = 39% for 6430 based on LiDAR products (Zlinszky et al., 2014) to F1 81–99% for 6410 and 6510 using MS data (Rapinel et al., 2020); for grasslands (habitat 6230) from F1 = 0% based on MS images (Haest et al., 2010) to 93% sing RGB camera and DTM (Gonçalves et al., 2016); and for peatlands F1 from 34% for habitat 7140 using MS images (Stenzel et al., 2014) to around 90% for 7140 and 7230 based on HS data (Szporak-Wasilewska et al., 2021). From the above study, it cannot be concluded that the Natura 2000 habitat classification on MS data will have significantly lower accuracy than on HS data. This is due to the fact that there were no early studies that directly compared the accuracy of Natura 2000 non-forest habitat classifications on HS and MS data. Meanwhile, indirect comparison based on the result of different analysis is not possible due to many differentiating variables: different study sites (area and vegetation diversity); spatial and spectral resolution; training and validation datasets; date of data acquisition; classification method; verification process. All the above-mentioned factors hinder the comparison of the classification results, therefore, preventing from drawing conclusions and giving recommendations regarding issues such as optimal dataset.

Additionally, the habitats are very floristically diverse within different areas, and most studies are solely conducted within a single study site. All the above factors may lead to conclusions that are not universal on the scale of the habitat's range. The motivation for our study was to find out the actual difference in informativeness of HS and MS data for Natura 2000 habitat identification under conditions that ensure correctness and universality of conclusions. For this reason, it was also decided not to conduct the analysis in this research using real Sentinel-2 data but resampled based on HS data. Studied Natura 2000 habitats are characterised by high spectral changeability during the season resulting from agricultural use and phenological variability. Comparability of results from real Sentinel-2 and HS data would require synchronicity of both collections. Taking into account the high number of cloudy days during the growing season in Central Europe, the size of the study areas (number of areas) and the desire to repeat the study in three seasons these analyses were not feasible on real Sentinel-2 data and no other free satellite data. One solution to compare the data was to resample the HS data acquired from the aircraft. It was decided to perform the resampling to the Sentinel-2 data rather than to other satellite data, considering the fact that due to its temporal, spatial and spectral resolution, it is objectively the best potential source of free available satellite data used for monitoring the Pan-European Natura 2000 network. Moreover, Sentinel-2 data is currently most commonly used for Natura 2000 habitats monitoring in many ongoing projects funded by the European Commission; e.g.: ("Possibilities for updating map layers of NATURA 2000 biotopes using advanced remote sensing methods - Faculty of Environmental Sciences CZU Prague," 2020), ("Using satellite images to improve the operation of the Natura 2000 network: A prototype for monitoring Natura2000 sites habitats with Copernicus 2021," 2021).

Hyperspectral data provides substantial information, but at the same time neighbouring bands are correlated, whereas some of the bands contain mostly noise. The feature selection is important in reducing the dimensionality of the data, and leads to better understanding of crucial spectral ranges for vegetation classification. This is also key information for designing new sensors (Chan and Paelinckx, 2008). There have been no studies conducted in order to compare the information in different spectral ranges, VNIR and SWIR, that would make it possible to discriminate Natura 2000 habitats from the background. Only one study determined spectral ranges useful in recognising three non-forest habitats: in the VNIR range (0.4–0.8 μm) and in the SWIR range (1.05–1.10, 1.25–1.40, 1.95–2.05 and 2.25–2.40 μm) (Demarchi et al., 2020). However, it is important to analyse whether these ranges are stable during vegetation season and in different areas.

The aim of our study was to define the potential of HS data (HySpex) in VNIR and SWIR ranges and simulated MS data (HySpex resampled to Sentinel-2 resolution) to identify selected non-forest Natura 2000 habitats. The analyses were performed using Linear Discriminant Analysis (LDA) as a feature selection method to define spectral ranges essential for habitat differentiation from the background. The conclusions were drawn based on two study sites and three seasons for each habitat. Each of the Natura 2000 habitats was studied in two independent areas, which ensured the universality of results for individual habitats. Additionally, in order to take care of the transferability of results, the study covered protected habitats relatively most common at the scale of Central Europe. It was assumed that for monitoring these habitats at the EU scale remote sensing is particularly applicable.

Different methods of bands selection for vegetation analysis were proposed, for example based on mutual information (Guo et al., 2006), principal component analysis, lambda–lambda R^2 models, stepwise discriminant analysis, and derivative greenness vegetation indices (Thenkabail et al., 2004) or Random Forest and Adaboost tree-based selection methods (Chan and Paelinckx, 2008). In this study, the LDA was performed in order to find the spectral ranges differentiating Natura 2000 habitats from the background (He and Wang, 2021). The background includes all habitats not protected by the Natura 2000 network and areas without vegetation. The LDA is a supervised method of dimensionality reduction. It searches for the optimal projection of the analysed classes, at the same time, reducing the dimensions of the data, for example, by selecting the most differentiating spectral bands by reducing their number (Gao and Xu, 2016). The method calculates the directions - linear discriminators that represent axes which maximise separation between multiple signal sources (Gite et al., 2019). The LDA is based on maximising the ratio of inter-class to intra-class spread in the data - Fisher linear discriminator, (Jayaprakash et al., 2018). This method has already been used to find data differentiating selected plant species from the environment (Jarocińska et al., 2021; Zagajewski et al., 2017). In addition, it was used to classify tree species based on hyperspectral data and simulated multispectral data WorldView-3 (Ferreira et al., 2016). With the use of LDA, it was possible to determine spectral ranges that are sufficient to differentiate the Natura 2000 habitats.

2. Materials and methods

The data for habitat has been processed separately and using the same method (Fig. 1). Images and field data were acquired and preprocessed to create the database with spectral reflectance. In order to find the most separative bands, the statistical analyses were performed including removal of the most correlated bands and LDA in the iterative mode. Further visualisation resulted in recommendations concerning the possibility of habitat identification.

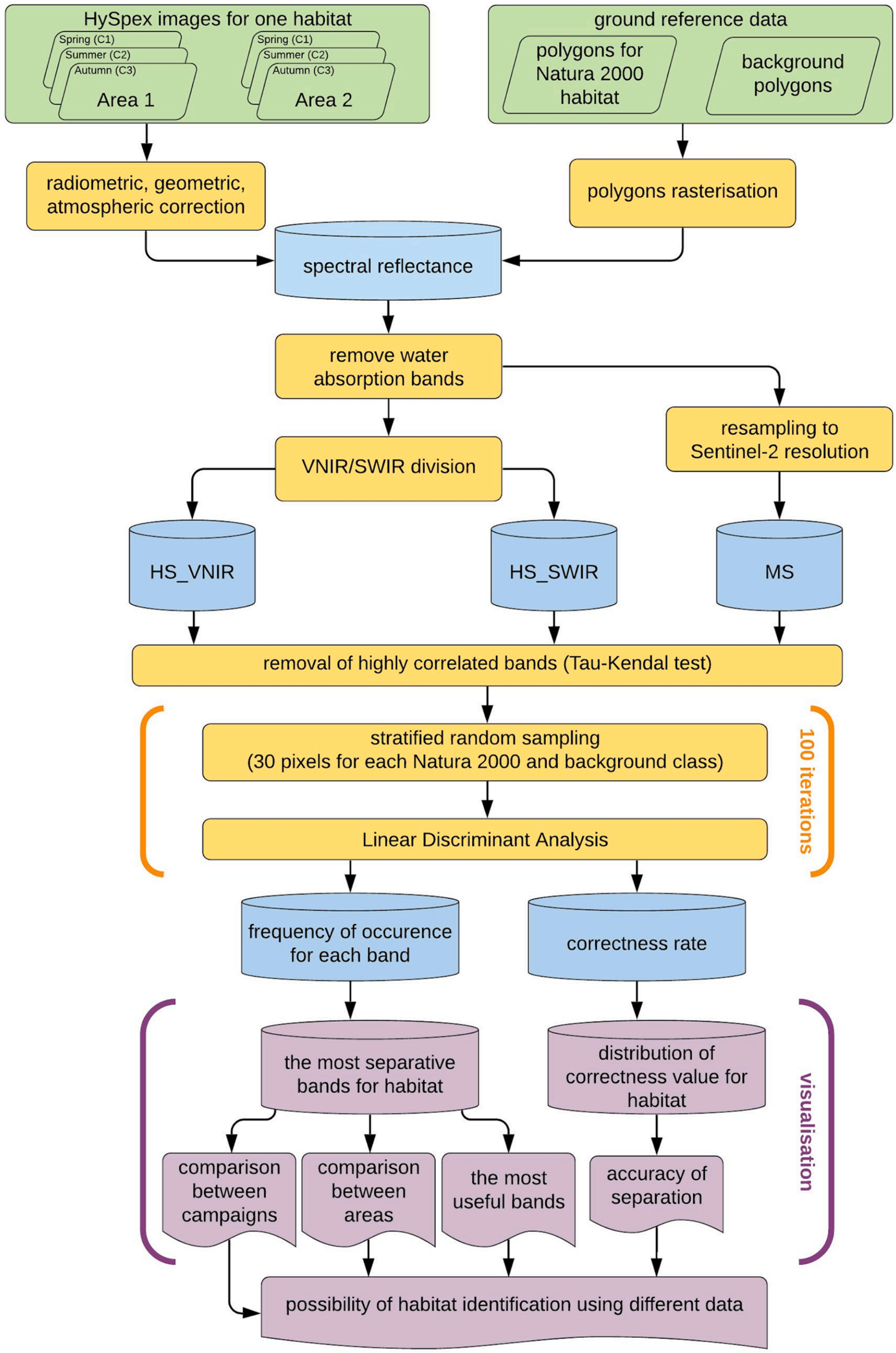

Fig. 1. The overview of data acquiring and analysis for each Natura 2000 habitat.

2.1. Study areas

The study was conducted in some of the most valuable areas in Poland in the terms of conservation status and biodiversity of non-forest ecosystems. All seven study sites are located on the territory of Poland in the continental biogeographic area (Fig. 2) and they are protected as Natura 2000 sites. The total area of investigated areas is 228 km^2.

Study sites were selected in such a way that each of the five habitats was analysed at two geographically distant locations. As a result, the study also took into account the geographic variation of each habitat. Moreover, the study sites differ significantly in the diversity of non-forest ecosystems and intensity of agricultural use. An important feature that differentiates the areas is that in area BI1 and BU4 there are only few other types of natural and semi-natural non-forest ecosystems apart from the studied habitats, whereas in the remaining areas, the diversity of non-forest ecosystems referring physiognomically to the studied habitats is significantly higher (from 7 to 14, see also Table 1.).

2.2. Natura 2000 habitats

Five Natura 2000 non-forest habitats were included in the study. The selected habitats represent different types of non-forest ecosystems (heaths, grasslands, meadows and peatlands), and are among the most common non-forest Natura 2000 habitats in the continental

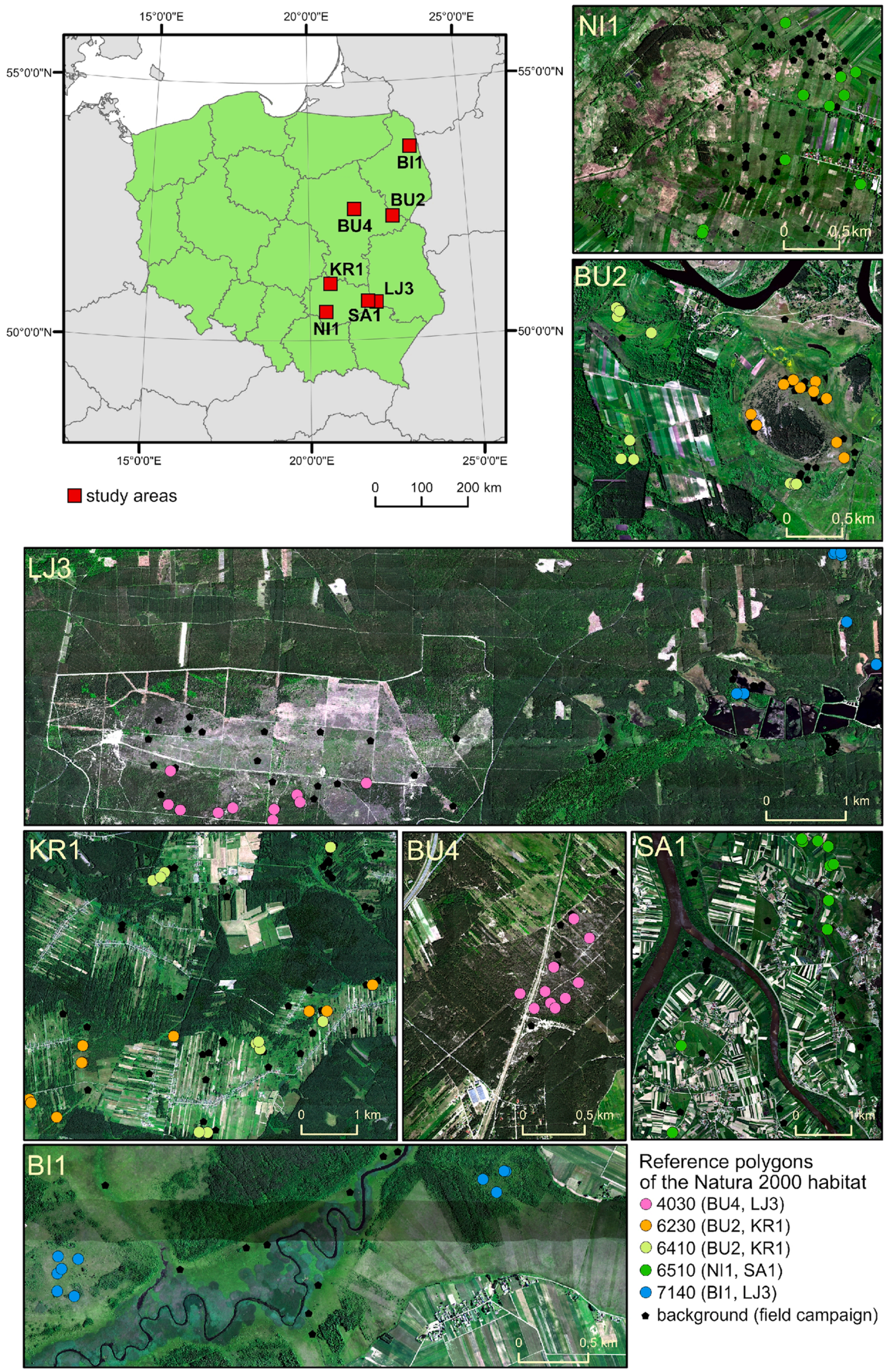

Fig. 2. Location of study areas and distribution of reference polygons of analysed habitats. Each habitat was analysed in two areas: 4030 (BU4, LJ3), 6230 (BU2, KR1), 6410 (BU2, KR1), 6510 (NI1, SA1) and 7140 (BI1, LJ3).

Table 1
The analysed number of pixels for each class on each habitat is an area given for one campaign.

Class/Natura 2000 code	Number of pixels for each class / number of subclasses (only for "background")						
	BI1	BU2	BU4	KR1	LJ3	NI1	SA1
Natura 2000 habitats							
Dry heaths/4030			144		136		
Nardus grasslands/ 6230		138		132			
Molinia meadows/ 6410		130		136			
Hay meadows/ 6510						134	135
Mires/7140	132				140		
Background - other habitats							
Other meadows		538/ 4		551/ 4		669/ 5	270/ 2
Other grasslands		131/ 1	138/ 1		138/ 1	133/ 1	
Rushes	132/ 1	137/ 1			169/ 2	170/ 2	133/ 1
Other mires						132/ 1	
Nitrophilous tall herb		135/ 1		138/ 1	135/ 1	368/ 2	141/ 1
Forest	134	135	137	140	136	131	139
Background - areas without vegetation							
Bare soil	142	136	135	132	124	134	137
Water	134	140		135	123	136	133
Number of pixels for each campaign	674	1620	554	1364	1201	2007	1088

biogeographical region (Fig. 3). As these habitats host many valuable and protected species, the conservation of these habitats is a key to the conservation of biodiversity in the Central European Lowlands. Additionally, the selected habitats differ significantly in their species structure so that the aspect of different internal structure and physiognomy of the habitat could be taken into account in the study.

2.2.1. European dry heaths (Natura 2000 habitat code – 4030)

Shrub communities, with a small number of species and varied cover, associated with oligotrophic sandy substrate, occurring in different topographic positions; most often there is one dominant species *Calluna vulgaris*, less often other species - *Arctostaphyllos uva-ursi.* Patches of *Pohlio-Callunetum* Shimwell 1973 em Brzeg 1981 are identified as habitat code 4030 in area LJ3, while in area BU4 *Arctostaphylo-Callunetum* R.Tx. et Prsg 1940.

2.2.2. Species-rich Nardus grasslands (Natura 2000 habitat code 6230)

A semi-natural grassland community with a loose sward of low species diversity (dominated by *Nardus stricta*), often with a clump structure, developed in various topographic locations on moderately moist soils under conditions of long-term extensive grazing and limited fertilisation. Including the following patches of *Polygalo-Nardetum* Prsg 1953; *Nardo-Juncetum squarrosi* Nordh. 1920; *Calluno-Nardetum strictae* Hrync. 1959. In area BU2 poor patches dominate with a very high and constant share of *Nardus stricta*, area KR1 is distinguished by the significant habitat variability. On site KR1, patches of all three communities forming the habitat were found.

2.2.3. Molinia meadows (Natura 2000 habitat code 6410)

Meadow communities, semi-natural, with sward, rich in dicotyledonous perennials (e.g. *Betonica officinalis*) and clump grasses (*Molinia caerulea*), geographically diversified, developed on deforested areas, on soils of different fertility and trophicity levels and of variable moisture content, under conditions of an extensive mowing. Patches of the following communities are identified as habitat code 6410: *Junco-*

Fig. 3. Analysed Natura 2000 habitats.

Molinietum Prsg 1951; *Selino carvifoliae-Molinietum* Kuhn 1937; *Galio veri-Molinietum* Kącki 2007; *Ranunculo polyanthemi-Filipenduletum vulgari* Hundt 1958. In both areas the habitat shows very high variability with respect to dominant species.

2.2.4. Lowland hay meadows (Natura 2000 habitat code 6510)

Meadow communities developing on rich, fresh mineral soils, maintained as a result of an extensive hay cutting; the sward is composed of grass species admixed with dicotyledonous perennials. Patches of the following communities are identified as habitat code 6510: *Arrhenatheretum elatioris* Br.-Bl. ex Scherr. 1925; communities of *Poa pratensis-Festuca rubra* Fijałk. 1962. In both studied areas dominate patches of the *Arrhenatheretum elatioris* complex, whereas patches belonging to the *Poo-Festucetum rubrae* complex occur considerably less frequently.

2.2.5. Transition mires and quaking bogs (Natura 2000 habitat code 7140)

Peat-forming communities developing on overgrown lakes and ponds and in flooded valleys on peat substrate, under conditions of permanent, stable moisture and oligo- and mesotrophic rainwater supply; varied vegetation, with a well-developed layer of mosses (Sphagnum sp.), numerous low sedges (*Carex* sp.), Cottongrasses (*Eriophoprum* sp.), as well as trees and shrubs. Patches of the association *Menyantho-Sphagnetum* teretis Warén 1926 are identified as habitat code 7410 in the BI1 area, whereas patches of the association occur in the LJ3 area, where the habitat is more diversified: *Caricetum lasiocarpae* Koch 1926, *Eriophoro angustifolii-Sphagnetum recurvi* Jasnowski 1968; *Caricetum rostratae* Rübel 1912; *Rhynchosporetum albae* Koch 1926; *Eriophoro vaginati-Sphagnum recurvi* HUECK 1925. In both areas, patches are related to raised bogs *Oxycocco-Sphagnetea* Br.- Bl. et R.Tx. 1943.

2.3. Airborne data

For each habitat, the airborne and ground reference data were acquired in 2017, in two areas and three campaigns. The first campaign (C1) was carried out in spring (from the middle of May to the end of June), the second campaign (C2) was conducted in full summer (July-August), and the last (C3) in early autumn, in September or in early October. The time of overflights and on-ground botanical surveys were selected based on the phases of vegetation development within the analysed habitats to find the largest differences in spectral signatures of habitats and background, as well as to capture different development phases of the dominant species (Fig. 4).

The hyperspectral images were acquired using two HySpex cameras from the Norwegian Norsk Elektro Optikk (NEO): VNIR-1800 (0.4–0.9 μm) with 182 spectral bands with spectral sampling 3.26 nm, and SWIR-384 (0.9–2.5 μm) with 288 bands with spectral sampling 5.45 nm ("HySpex," 2021). The radiometric resolution was equal to 16 bits. The scanners were placed on the Cessna CT206H aeroplane, and the overflight was performed on the altitude of 730 m a.s.l. (Sławik et al., 2019). During each overflight, Airborne Laser Scanner data with a density of 7 points/m^2 was also acquired. The hyperspectral images (HS) simultaneously with the ALS data was obtained for one area in a single campaign during one overflight. In this way, we performed 21 overflights - on seven areas in three campaigns.

Firstly, radiometric calibration was performed in the HySpex RAD software provided by the manufacturer. Parametric geocoding for each image was established in PARGE software ("PARGE Airborne Image Rectification," 2021) based on parameters acquired during the overflights and Digital Surface Model developed from points clouds from ALS. The images from two scanners were combined into a single one in 0.935 μm and one image with 451 bands with 1 m spatial resolution was created. Atmospheric correction was executed in ATCOR-4 software (ReSe Apps) using the MODTRAN model (Richter and Schlapfer, 2020). A Savitzky-Golay filter with a range of 13 bands was used to smooth the spectral reflectance. Finally, the mosaicking was performed with an OrthoVista product in INPHO software ("OrthoVista," 2020). Due to disturbances in bands where the absorption of radiation by water occurs, the last 21 bands were removed and images with 430 bands were acquired. In some studies, also bands from the beginning of the blue spectrum were removed (Lenhard et al., 2015). On the other hand, the blue light was noticed as important in vegetation analysis (Çimtay and İlk, 2018; Demarchi et al., 2020; Jarocińska et al., 2021). There is no clear conclusion concerning the use of the beginning of visible light using HySpex data, that is why in this study only the end of SWIR bands was removed. Details of pre-processing can be found in previous studies (Jarocińska et al., 2021; Sławik et al., 2019).

2.4. Ground reference data

Simultaneously with the overflights, ground surveys were conducted in each area in three campaigns. The aim was to establish ground reference polygons in each area: for Natura 2000 habitats and background (other non-forest vegetation communities). The field survey was only used to indicate the coordinates of each habitat and background type. The areas thus determined were then used to extract spectral reflectance curves from the data. Field surveys were one possible method for obtaining reference polygons. Localisation of reference polygons by photointerpretation on acquired HS or any types of airborne data was not possible, because the surveyed habitats do not visually differentiate from other types of low semi-natural vegetation (beyond habitat 4030).

The permanent reference plots established during the first campaign (C1) were checked in two other campaigns (C2 and C3) where the polygons were not changed in structure or physiognomy, for example, due to land use changes or agrotechnical treatment. In order to

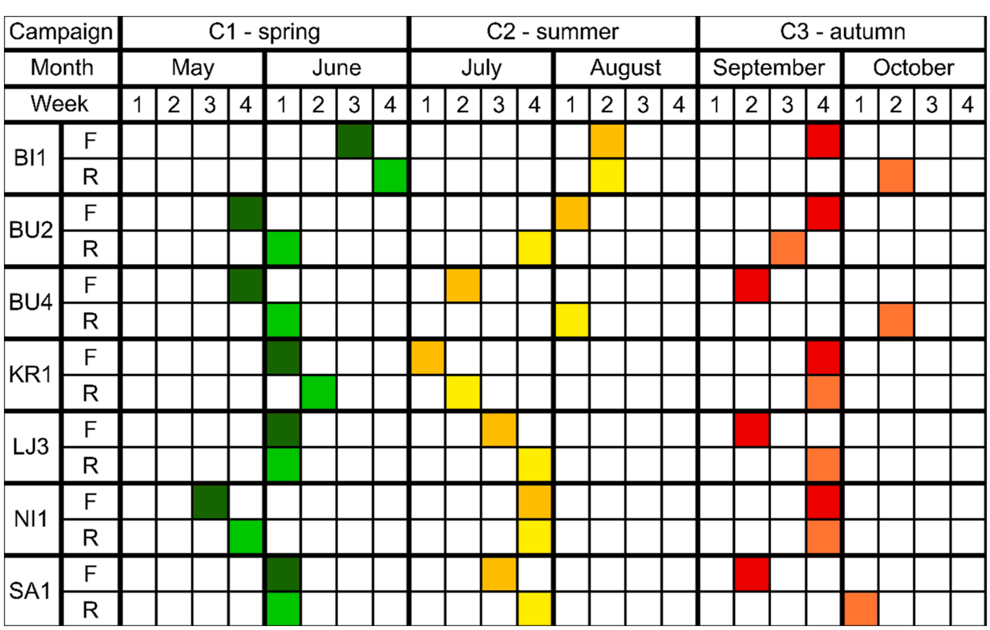

Fig. 4. The date of acquiring data: F - overflight, R - ground reference data.

determine the location, the centre of polygons was registered using a GPRS recorder (model: MobilerMapper 120, Spectra Pecision) with the precision of 0.1–0.5 m. All polygon centres for habitats and background were located at least 4 m away from the border of the vegetation patch. Polygons of both habitats and other plant communities (background) were established only in well-developed patches (with dominance of indicator species). In total, 1448 unique polygons were acquired with 259 covering the analysed Natura 2000 habitats. It was assumed that reference polygons should be as evenly spaced as possible, and therefore, the minimum values of the analysed habitats and types of background were equal to 10. This approach provided a guarantee that the minimum number of pixels for each class would be higher than 100.

Moreover, background polygons (10 polygons for forest, bare soil and water respectively) were located based on visual interpretation of airborne images (see chapter 2.3.). Only in area BU4 two classes were located (forest and bare ground) as there were no reservoirs.

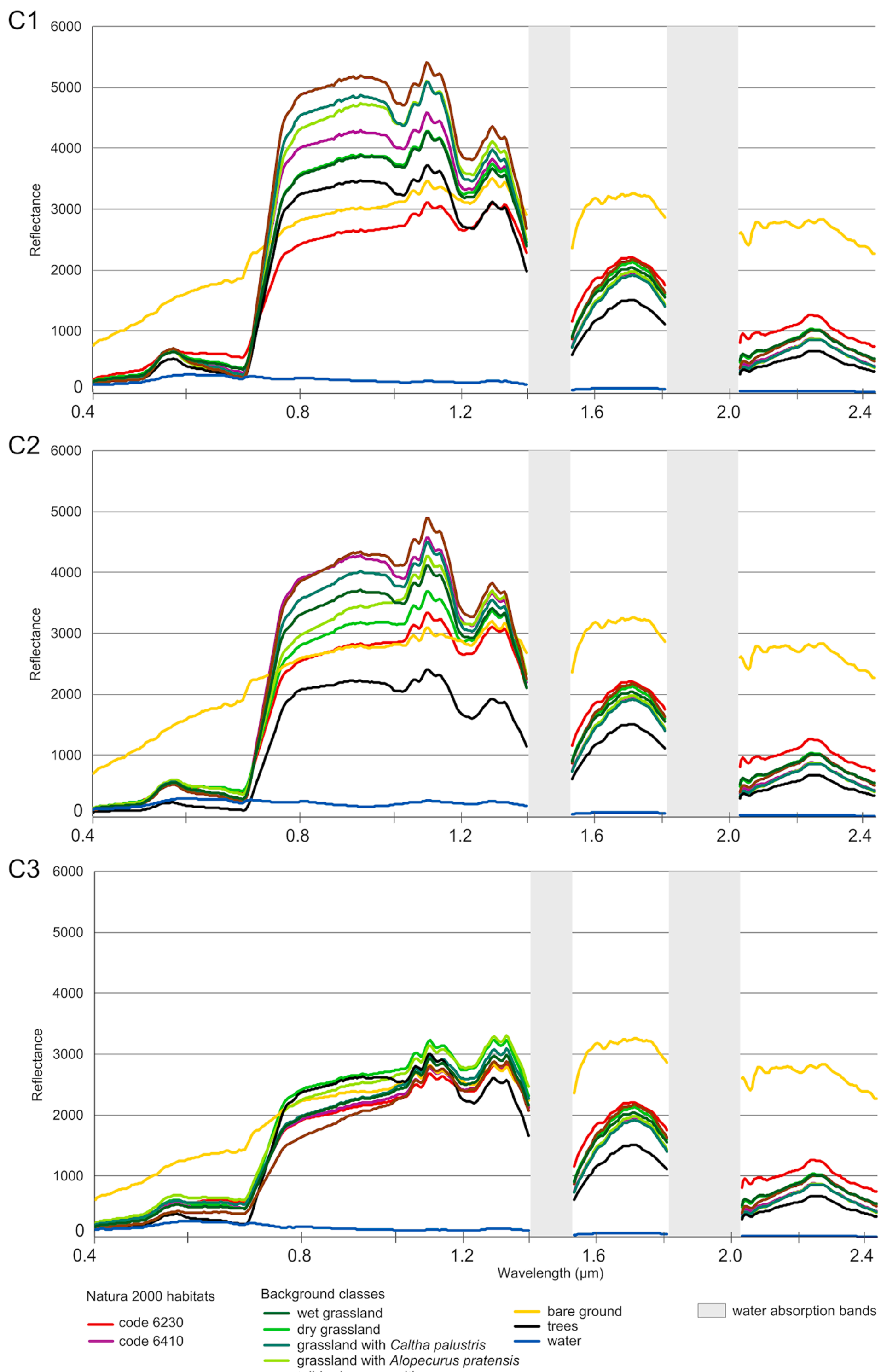

Fig. 5. Spectral characteristics of the objects for the example area KR1 in campaign one (C1), two (C2) and three (C3).

2.5. Spectral reflectance databases

The process of creating the spectral reflectance databases was divided into several stages. Firstly, polygons were rounded, based on the polygon's centre location with a radius of 2 m. Then, the polygons were rasterised using the Region of Interest tool in ENVI 5.3. software. In order to avoid not spectrally clear pixels, only pixels with at least 50% coverage in analysed polygons were added to the database. Finally, the number of pixels for each class varied from 123 to 144 per class (Table 1). In total, the same 8508 pixels were analysed for each campaign. The spectral reflectance was acquired for each pixel from the database using HySpex images with 430 bands and saved in ASCII format.

Then, 64 spectral bands, where water absorbs light, were removed from the HS database: bands from 240 to 263 (1.366–1.490 μm) and from 316 to 355 (1.780–1.990 μm) (Gates, 1970; Jensen, 1983; Richter and Schlapfer, 2020). As a result, 366 hyperspectral bands (HS) were remaining. The HS database was divided into two parts: HS_VNIR (0.4–1 μm) and HS_SWIR (1–2.5 μm). The HS_VNIR consisted of 172 bands and the HS_SWIR – 194 bands. For the purpose of the reference the following graphs show the average spectral reflectance for the sample site KR1, which included two Natura 2000 habitats (*Nardus* grasslands – 6230 and *Molinia* meadows – 6410) and background subclasses used for LDA analysis (dry grassland, wet grassland, grassland with *Caltha palustris*, grassland with *Alopecurus pratensis*, tall-herb communities, bare ground, water and forest). The graphs respectively show the surveyed habitats and background in campaign one (C1), two (C2) and three (C3) respectively (Fig. 5.).

Subsequently, the multispectral dataset (MS) was created based on hyperspectral data. The HS dataset was imported to ENVI 5.3 software as ASCII file and saved as a spectral library file. This spectral library was resampled to Sentinel-2 resolution using ENVI 5.3 filters. In this procedure, only wavelengths of band centres were provided, therefore, the software assumes critical sampling and applies a Gaussian model with an FWHM equal to the band spacing. The last step was the export to ASCII format.

As a result, three databases were further analysed: HS_VNIR, HS_SWIR and MS.

2.6. Statistical analysis

The HS and MS databases thus prepared were statistically processed using LDA in order to determine which spectral bands differentiate habitat from background across the study areas in the three campaigns. LDA has been used in this article as a linear feature selection approach to reduce the number of variables (Song et al., 2010). The use of LDA enabled to specify the linear combination of features that characterises or separates two or more classes of objects or events. As a pattern recognition technique, LDA maximises the ratio of between-class variance to the within-class variance in any particular data set in order to obtain a maximum discrimination (Tharwat et al., 2017). The first step of the algorithm is to find the directions (called linear discriminants, which are a linear combination of predictor variables) that maximise the separation between classes, and in the following step the algorithm uses these directions to predict class (Friedman, 1989).

The analyses were performed for three datasets: HS_VNIR, HS_SWIR and MS. The first step was to check all input datasets in order to search for a correlation of spectral bands using the Tau-Kendal test. All spectral bands whose pairwise correlation coefficient was greater than 0.99 were removed. This step was crucial to conduct any analysis using LDA, as the redundancy introduced by highly correlated bands could significantly degrade the ability to correctly select variables.

Afterwards, the LDA was conducted in iteration mode. In 100 iterations stratified, a random sampling was performed: 30 pixels for the analysed habitat and 30 pixels for each background class (Table 1), which were part of the study area. The pixels were drawn independently from the reference polygons, so that each pixel could be located in a different polygon or several pixels in a single polygon. A value was assigned to each spectral band in order to measure how often it was considered differentiating (frequency of occurrence), while for each iteration, a correctness rate value was calculated to determine the correctness of the matching band. The correctness rate refers to the probability of correct detection bands differentiation of the habitat from the background. The LDA correctness rate with a higher value, maximum approaching to 1, shall confirm their highest importance in differentiating between species.

The maximum number of bands per iteration, selected by LDA as differentiating, was limited to 40 in order to speed up the process; on the other hand, the analysis did not identify more than 40 bands as differentiating. As a final result of the statistical analysis, each habitat was characterized by 100 correctness coefficient values, which demonstrates the effectiveness of differentiation, and frequency of occurrence of the bands that were selected by LDA as most differentiating habitat from background in 100 iterations. This information was provided for two study areas and three campaigns. Thus, for each dataset, we were able to rank layers - bands, based on their order of selection by LDA and their overall contribution to the correctness rate. The analyses were performed in an R (James et al., 2013) software environment using caret (Kuhn et al., 2021), klaR (Roever et al., 2020), MASS (Ripley et al., 2021) and vegan (Oksanen et al., 2019) libraries.

For the HS datasets, diagrams were developed to analyse the suitability of individual bands for the differentiation of the habitat from the background. Bands that were marked as differentiating in at least 50 iterations were defined as differentiating bands. Values for occurrence frequency for bands with frequency above 50 were marked on the diagrams with colours varying from green to purple. Therefore, as a summary, these spectral bands were analysed depending on three factors: campaign, area, and universality. The differentiating bands were defined for each campaign to verify if they are dependent on the time of vegetation season. The bands were marked on the diagram accordingly, if they appeared to be differentiating for campaign in both areas.

To analyse the differences between areas, bands that were differentiating in the three campaigns for a given area were marked on the diagram. This analysis was conducted to verify if the spectral range is dependent on the background classes occurring on the study site. The last part demonstrates that some of the bands can be regarded as universal - differentiating the habitat from the background, regardless of the area and campaign. In this section, the ranges that differentiate in both areas and in each of the three campaigns have been marked.

We compare the MS and HS data by the number of differentiating bands within season and study area in the specified spectral ranges. The number of differentiating bands for HS and MS data were summed up within the following ranges: 0.40–0.50 μm (blue light), 0.50–0.60 μm (green light), 0.60–0.68 μm (red light), 0.68–0.75 μm (RedEdge), 0.75–1.00 μm (NIR), and two SWIR ranges: 1.00–1.50 μm and 1.50–2.45 μm. The diagrams show the sum of differentiating bands in the defined ranges in seasons (the band was considered differentiating if it was significant for both areas in a given campaign) and in areas (differentiating, if it was significant for each campaign in a given area) for HS and MS data.

In the results section, in order to simplify the description by referring to the specific bands spectral ranges the band centre was given (of the full width at half maximum), and not to the entire spectral range.

3. Results

3.1. European dry heaths - Natura 2000 habitat code 4030

The highest average correctness rate values were acquired for HS_SWIR data, regardless of the area and campaign. These values were very high, from 0.990 in the BU4 area in C2 to 0.999 in the LJ3 area in C3 (Fig. 6). Lower values, although still high and quite stable, were

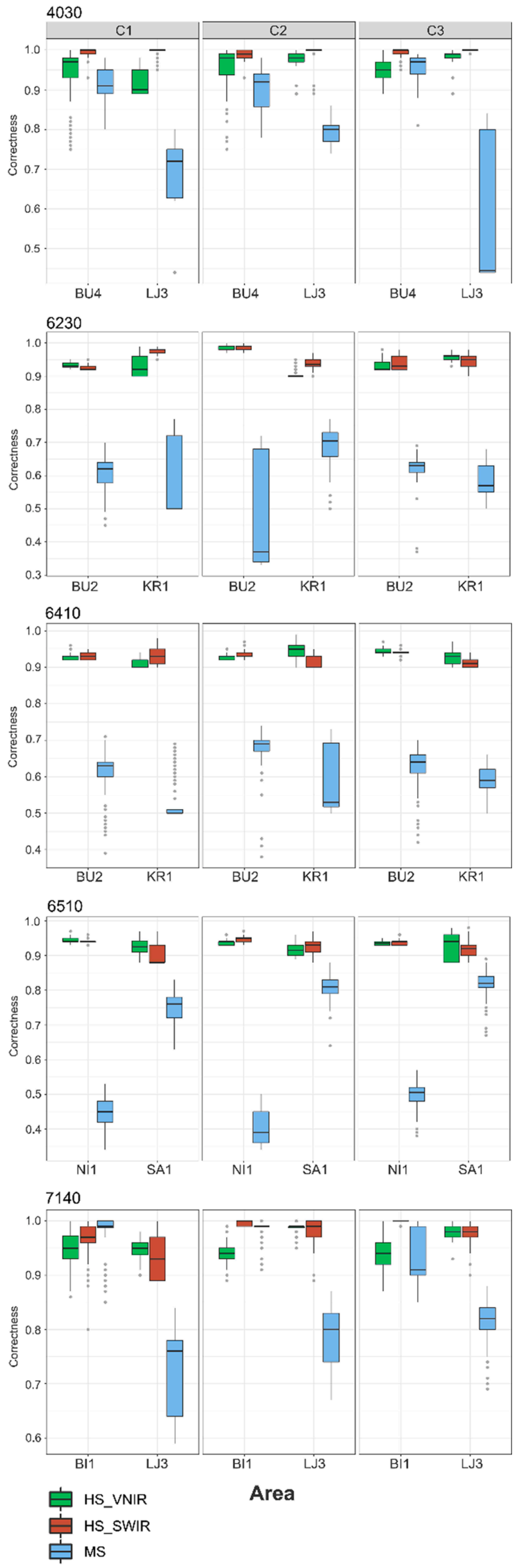

Fig. 6. The values of correctness rate for habitats calculated for hyperspectral (HS_VNIR, HS_SWIR) and multispectral (MS) data for two separate areas and repeated in three seasons (C1 - Spring, C2 - Summer, C3 - Autumn). The values were calculated based on 100 iterations of LDA analysis.

observed for the HS_VNIR data. On the other hand, values of average correctness rate calculated for MS datasets were lower, more diverse and differed within areas. For the BU4 area, the values were high and similar to HS_VNIR data, while for the LJ3 area values were significantly lower for each campaign (the average value for single campaigns was between 0.617 and 0.793).

Based on LDA, seven bands from the HS_SWIR (including five bands from 1.138 to 1.160 μm, band 1.496 μm and band 1.773 μm) were found differentiating habitat 4030 from the background regardless of season and area (Fig. 7). The highest number of differentiating bands (51) was acquired for the C3 autumn campaign for the LJ3 area, whereas the lowest (16) in spring (C1) for BU4.Fig. 8

The stability of acquired results for each campaign (calculated by dividing the number of differentiating bands on both areas to the sum of differentiating bands for at least one area) was the lowest for C1 – 21% and higher for C2 – 27% and C3 – 38%.

The HS_SWIR database presented the highest number of differentiating bands for habitat 4030. Bands differentiating in each campaign were only noticed in HS data. The range from 0.40 to 0.60 μm was more important for the LJ3 area than for BU4.

3.2. Species-rich Nardus grasslands - Natura 2000 habitat code 6230

Regardless of the campaign and the area, the mean value of the correctness rate for MS data did not exceed 0.668 (C2 in the KR1 area). Moreover, the values for MS were significantly more diverse compared to HS databases. Average correctness rate for both HS ranges were high and similar - above 0.9: the lowest average value of 0.903 was observed for VNIR in C2 on KR1; whereas the maximum, at the level of – 0.986 was noted for SWIR in C2 on area BU2. The correctness values obtained do not indicate differences in the effectiveness of the habitat differentiation between neither the areas nor the campaigns.

Only three spectral bands in the 0.416–0.423 μm range (HS_VNIR) were differentiating habitat 6230 from the background for every campaign and both areas (Fig. 9). Based on the LDA analysis, it can be concluded that they were independent of the area and the date of the analysis. More differentiating bands were noticed for the BU2 area than in KR1: 6 bands in the 1.127–1.160 μm range. The most significant bands were noticed for the C3 campaign and only four for C1 (beginning of the blue light). The most differentiating bands were located in the blue range light both for HS and MS databases and also in the SWIR range - in the case of HS data it is was the range of 1.0–1.5 μm, and for the MS – 1.5–2.5 μm.

For the habitat 6230 the correctness rate values were much lower for MS compared to both HS datasets (Fig. 6).

The stability of the results acquired for each campaign (calculated by dividing the number of differentiating bands on both areas to the sum of differentiating bands for at least one area) was the lowest for C1 – 15%, as well as for C2 – 16%, and significantly higher for C3 – 50%.

3.3. Molinia meadows - Natura 2000 habitat code 6410

The differentiation efficiency was distinct for HS and MS data (Fig. 6). The average correctness rate for HS data regardless of the area and campaign was higher than 0.910 for the VNIR and SWIR ranges. The highest values were observed for the KR1 area in the C2 campaign in VNIR – with the average of 0.946, while the lowest for KR1 in the C1 campaign for VNIR – average of 0.910. In the case of the MS database, the correctness rates were significantly lower (not exceeding 0.671) for every campaign and area, showing the mean values ranged between 0.532 and 0.671. In the case of MS data, a difference in the correctness values between the areas was also observed. In C1 and C2 campaigns the mean correctness values were higher for BU2 and lower for KR1. A higher diversity within values was also detected for KR1.

Based on LDA, it can be stated that differentiating ranges are located in 0.40–0.60 μm and 1.00–1.50 μm (Fig. 10). In total, 22 bands were

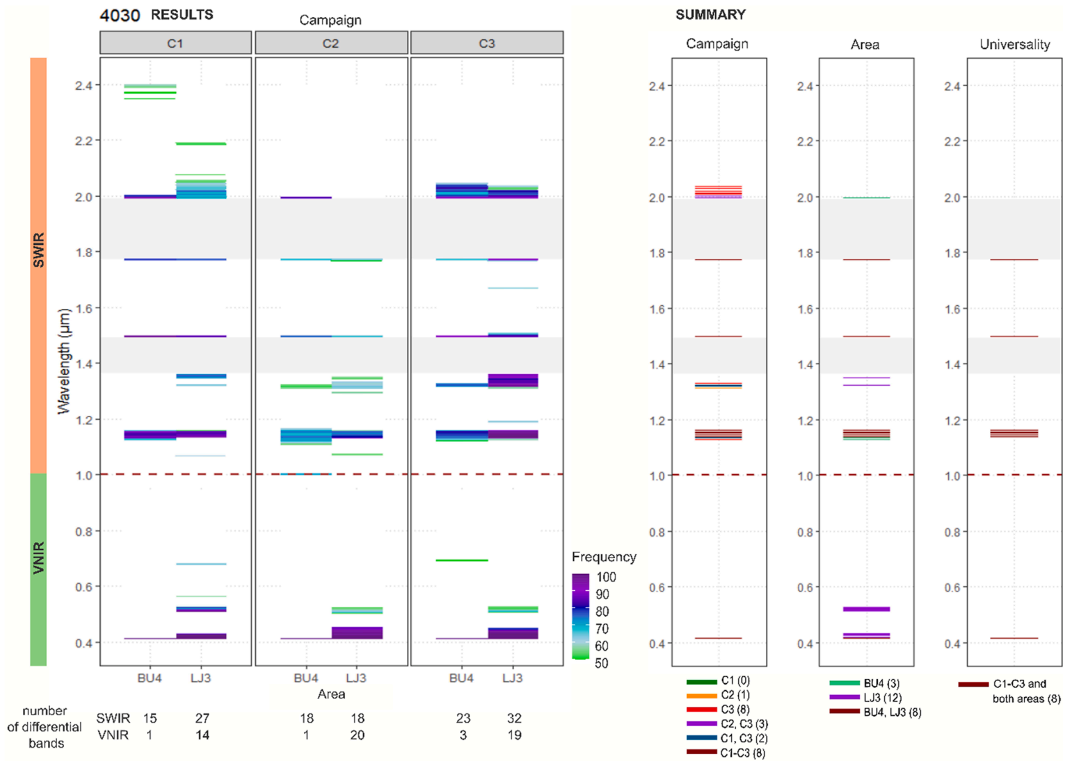

Fig. 7. Bands that differentiate habitat 4030 from the background classes were calculated based on HS_VNIR and HS_SWIR databases for three seasons (C1 - Spring, C2 - Summer, C3 - Autumn) and two areas (BU4 and LJ3). The bands were defined as differentiating when the frequency of occurrence value was above 50 out of 100 LDA iterations. In the summary, the following are presented: Campaign - a band was selected if it was differentiating in both areas for a certain campaign; Area - a band was selected if it was differentiating in three campaigns for a specific area; Universality - a band was selected if it was differentiating in both areas in each campaign. The number of differentiating bands is given at the bottom of the figure.

found, of which 13 were from HS_SWIR and 9 from HS_VNIR. Among these, only 6 were universal, they were therefore differentiating for every campaign and both areas. For the MS database all bands were differentiating, but none was universal. Moreover, significant differences between the areas were observed.

Based on HS data, the stability of acquired results for each campaign (calculated by dividing the number of differentiating bands on both areas to the sum of differentiating bands for at least one area) was relatively high: for C1 – 52%, for C2 – 63% and for C3 – 67%.

3.4. Lowland hay meadows - Natura 2000 habitat code 6510

The average correctness values for 6510 were considerably lower for MS data compared to HS - and equal to those of 6230 and 6410 habitats (Fig. 6). For the VNIR database, the values ranged from 0.902 (SWIR, SA1, C1) to 0.947 (SWIR, NI1, C2). In general, there were no differences between the parameter values for VNIR and SWIR. On the contrary, the differences between areas based on the MS dataset were substantial: the values were higher for SA1 - the average from 0.747 for C1 to 0.815 for C3, compared to NI1 - from 0.402 for C2 to 0.501 for C3. In the case of HS data, the correctness values for SA1 were lower and more diverse, but the differences between the two areas were not significant.

Only five bands at the beginning of blue light were found universal, thus differentiating regardless of area and campaign, based on the LDA (Fig. 11). Moreover, an additional eight bands were defined for the NI1 area: in the HS_VNIR four at 0.432–0.442 µm, and in the SWIR: one band 1.003 µm and three in the range 1.138–1.149 µm. Almost all differentiating bands were noticed in the spring campaign (C1). Apart from blue light, two bands were the most differentiating: from the 0.502 µm and 1.496 µm.

Comparing the results of MS and HS analysis, different conclusions can be drawn. For MS data no universal bands were noticed. In the SA1 area all bands except bands 5 (0.703 µm) and 12 (2.203 µm) were differentiating the habitat from the background. Apart from that, all bands were found differentiating for the C3 campaign. It should be noted that the results for MS are less reliable than for HS based on the correctness values.

By comparing the number of differentiating bands in both areas to the sum of differentiating channels in a given campaign at least in one area, the stability of conclusions for C1 can be determined at 25%, C2 – 30% and C3 – 22%.

3.5. Transition mires and quaking bogs - Natura 2000 habitat code 7140

The results for correctness for habitat 7140 were largely diverse and different, particularly between the two areas (Fig. 6). The tendency in correctness rate values for LJ3 was similar to previously described habitats: the values for HS were high (the average varied from 0.933 for SWIR in C1 to 0.988 for VNIR in C2), and considerably lower and more diverse for MS (the average varied from 0.726 for C1 to 0.813 for C3). For HS, the results were similar for each campaign. Completely different relationships were noticed for the BI1 area: the values for HS and MS were relatively similar and high, especially in C1 (from 0.954 for VNIR to 0.970 for MS) and C2 (from 0.942 for VNIR to 0.997 for SWIR).

Two bands were found universal for differentiating habitat 7140 from background for each campaign in the two areas: 0.416 µm and 1.138 µm (Fig. 12). The highest number of differentiating bands were discovered in C3: two bands in the range 0.419–0.423 µm, one with a wavelength of 0.439 µm, and five bands in the range of 1.127–1.149 µm. Another four bands were noticed for C1: 0.419 µm, 0.429 µm and two in the 0.512–0.515 µm range. The habitat 7140 was the only one where universal bands were noticed, regardless of the term and area for the MS dataset: first three bands in the 0.444–0.559 µm. For the MS data for LJ3 five more bands in the range 0.419–0.432 µm differentiate the habitat.

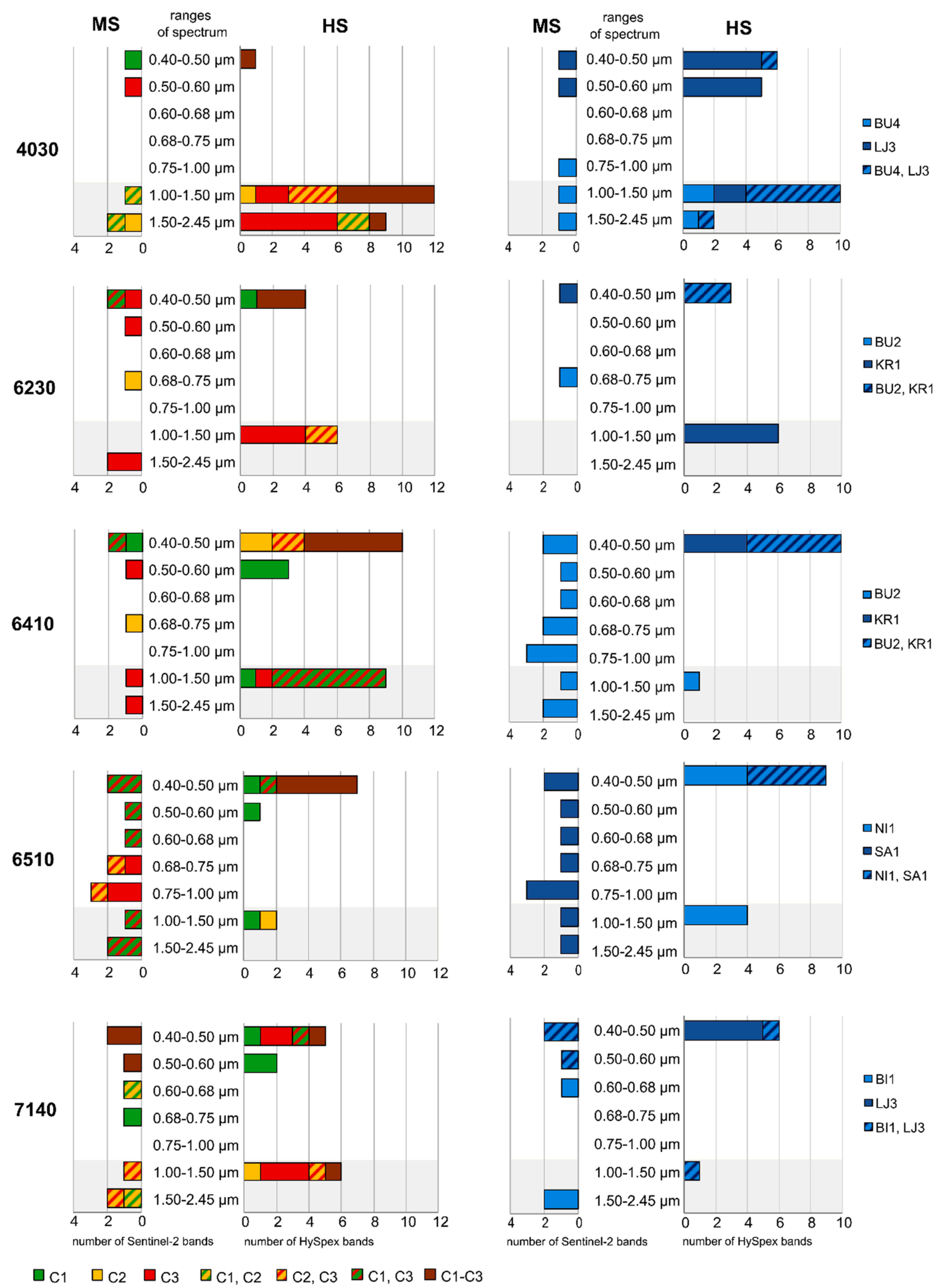

Fig. 8. The number of bands differentiating each habitat from the background calculated for the HS and MS dataset dependent on spectral range. The analysis was calculated for three seasons (C1 - Spring, C2 - Summer, C3 - Autumn; marked as differentiating, if it was significant for both areas in a given campaign) and two areas (differentiating if it was significant for each campaign in a given area).

In the case of MS data, the results were similar: three bands in the range 0.444–0.559 μm were universal, while the next 3: 0.664 μm, 1.613 μm and 2.203 μm differentiated the habitat from the background in the LJ3 area.

By comparing the number of differentiating bands in both areas to the sum of differentiating channels in a given campaign in at least one area, the stability of conclusions for C1 can be determined at 38%, C2 – 11% and C3 – 22%.

4. Discussion

4.1. The efficiency of separation and the most useful spectral ranges

Based on the research conducted, it can be concluded that for each Natura 2000 habitat studied in each area and campaign, bands which differentiate habitats from background were noticed (Fig. 13). Most of the HySpex bands (271 out of 366) did not differentiate any habitat from the background, which means that only 95 bands (26%) were useful

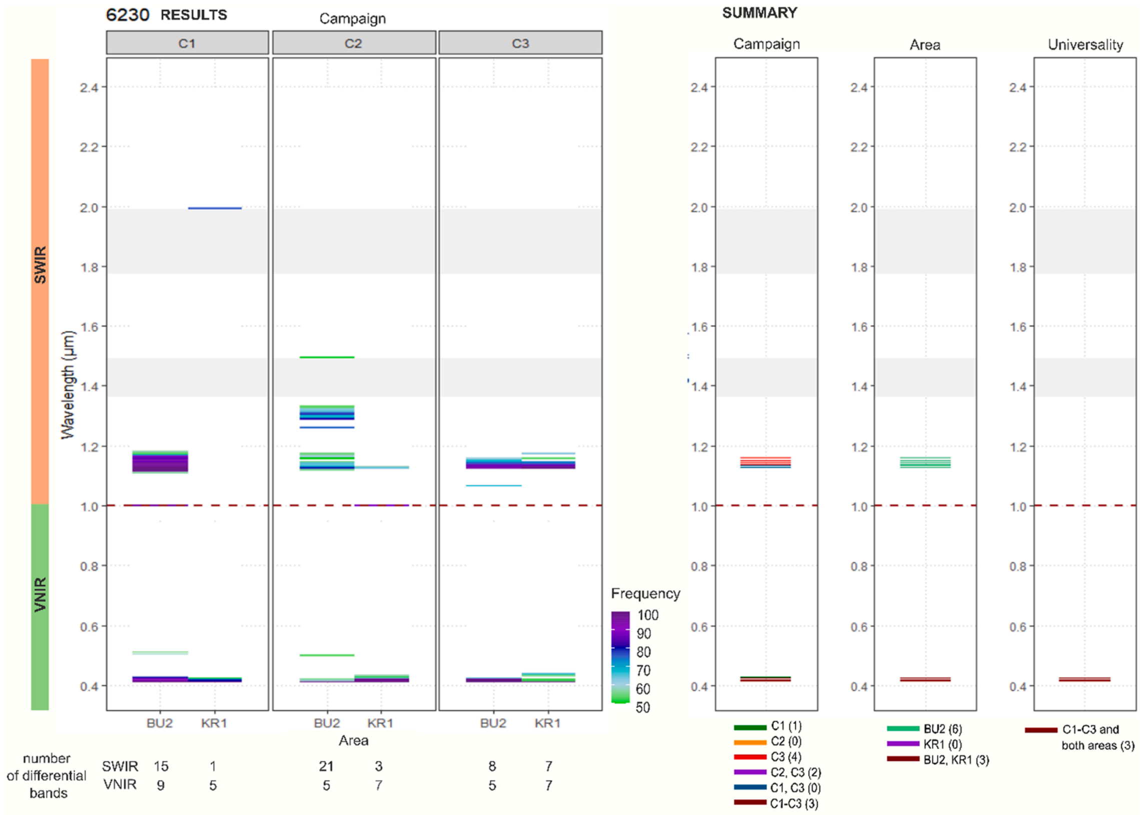

Fig. 9. Bands that differentiate habitat 6230 from the background classes were calculated based on HS_VNIR and HS_SWIR databases for three seasons (C1 - Spring, C2 - Summer, C3 - Autumn) and two areas (BU2 and KR1). The bands were defined as differentiating when the frequency of occurrence value was above 50 out of 100 LDA iterations. In the summary, the following are presented: Campaign - a band was selected if it was differentiating in both areas for a certain campaign; Area - a band was selected if it was differentiating in three campaigns for a specific area; Universality - a band was selected if it was differentiating in both areas in each campaign. The number of differentiating bands is given at the bottom of the figure.

(differentiating at least one habitat from the background in at least one area and in at least one campaign). On the other hand, 22 bands were useful for every habitat.

Most of the bands differentiating the habitats from the background were located in the blue light 0.416–0.442 μm (VNIR) and 1.117–1.165 μm (SWIR). The beginning of green light, 0.502–0.522 μm, as well as around 1.290–1.361 μm, were also useful. Additionally, 37 bands differentiated only one habitat from the background; these bands covered most of the optical domain except the near infrared (0.7–1.0 μm). Similar spectral ranges were noticed as useful for vegetation differentiation, especially in green light and around 1.3 μm (Thenkabail et al., 2004). Also, quite similar results were acquired for ecotope mapping using two methods: Adaboost and Random Forest (Chan and Paelinckx, 2008).

The number of universal bands (which differentiate certain habitats in each campaign and in both areas) was low: the minimum of 2 for habitat 7140 and the maximum of 8 for 4030. Studies on the importance of bands were previously conducted for habitats 6120, 6440, and 6510, however, the goal was to identify habitats, not to determine useful bands (Demarchi et al., 2020). The most important bands were defined as 0.4–0.8 μm, mainly the blue and the red range, part of NIR (1.05–1.10 μm), and of SWIR (1.25–1.40 μm, 1.65–1.80 μm, 1.95–2.05 μm and 2.25–2.40 μm). These ranges mostly overlap with ranges mentioned in this study.

Most bands deemed useful were observed for habitat 4030: 76 out of 366 bands for the whole optical domain, 25 out of 172–15% in VNIR and 51 out of 194–26% in SWIR. For heaths high accuracies of classification were also acquired (Gonçalves et al., 2016; Haest et al., 2017). For mires 7140, 15% of the bands from HS datasets were differentiating it from the background, and 20% of the SWIR and 10% of the VNIR range. The lowest number of differentiating bands was noticed for habitat 6410: 10% in the entire optical range, 12% of VNIR and only 9% of SWIR. In general, it can be concluded that in the case of meadows and grasslands, fewer bands were defined as differentiating than in the case of mires and heaths. Also, classification accuracies for the meadows achieved in other studies were not always high (Buck et al., 2015). Thus, it can be concluded that there is a positive relationship between the number of differentiating bands with the classification accuracies.

More differentiating bands were found in the SWIR range (66 out of 194 bands – 34%) than in the VNIR range (29 out of 172 channels – 17%). Based on the correctness values it can be assumed that only one of the ranges, VNIR or SWIR, allow for the habitat differentiation. Previously performed classifications were mostly based on images from the whole optical range or its parts: VIS and NIR, i.e. using the RapidEye (Schuster et al., 2015; Stenzel et al., 2014). While no studies using only the SWIR range data were previously published, our results indicate that the SWIR range is also useful for differentiating (Figs. 6, 8 and 13). Based on correctness values, the differences between VNIR and SWIR differentiation results are relatively minor. Only for habitat 4030, the highest and more stable correctness values were noticed for SWIR compared to VNIR (Fig. 6), therefore, using the images from SWIR resulted in better classification accuracies achieved. On the other hand, results for VNIR were also satisfying, so using only VNIR data could also give good classification results for heaths. Similarly, mires 7140 are characterised by higher correctness values for the SWIR range than for the VNIR for one area BI1 in C2 and C3 campaigns (Fig. 6). This may be related to differences in the background classes (Table 1). The background classes are different on areas LJ3 and BI1: on LJ3 polygons are dominated by dry vegetation, whereas on BI1 by hydrogenic habitats. At LJ3, the SWIR range, which is an indicator of water content, is therefore more

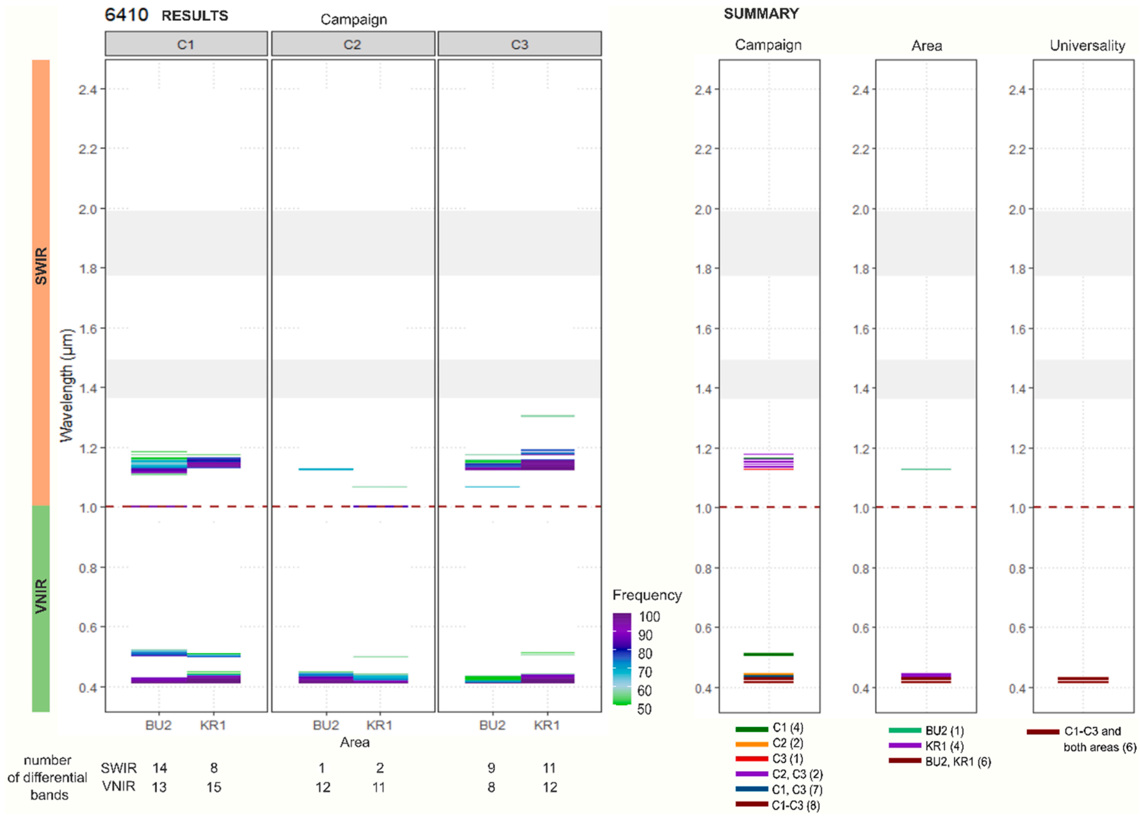

Fig. 10. Bands that differentiate habitat 6410 from the background classes were calculated based on HS_VNIR and HS_SWIR databases for three seasons (C1 - Spring, C2 - Summer, C3 - Autumn) and two areas (BU2 and KR1). The bands were defined as differentiating when the frequency of occurrence value was above 50 out of 100 LDA iterations. In the summary, the following are presented: Campaign - a band was selected if it was differentiating in both areas for a certain campaign; Area - a band was selected if it was differentiating in three campaigns for a specific area; Universality - a band was selected if it was differentiating in both areas in each campaign. The number of differentiating bands is given at the bottom of the figure.

influential in differentiating habitat from background because, unlike BI1, the background at LJ3 is mainly composed of dry vegetation.

Results from the two areas are generally consistent for most of the habitats, so it can be assumed that the background vegetation is not affecting the results in a significant way and the conclusions drawn will be universal. On the other hand, there are differences in the differentiating spectral ranges between individual areas. For habitat 6410 no differences between spectral ranges were noticed, whereas for habitat 7140 were the highest. This can be related to the fact that in case of 7140 vegetation, also the background classes differ much between the two analysed areas (see Sections 2.1 and 2.2).

Based on the correctness results, habitats differentiate from the background that was similar in individual campaigns, and differentiating bands were found. It can be stated that there is a possibility to distinguish habitat from the background using HS data. Similar conclusions were drawn from the previous studies focusing on classification of Natura 2000 habitats based on HS and ALS data (Demarchi et al., 2020; Szporak-Wasilewska et al., 2021).

4.2. The efficiency of habitats differentiating using simulated multispectral data

When entering the study, we assumed that HS would have had better accuracies than MS. Such a general statement could have been assumed in advance to be true and did not require testing. The new knowledge which our study provides is related to the variable potential to identify individual Nature 2000 habitats. The highest values were acquired for habitat 4030, and for 7140 in the BI1 area. Previous studies showed that it is possible to identify mires and heaths based on multispectral images. High accuracies (determination coefficient $R^2 = 0.94$, OA = 84%, and kappa = 63%) were achieved for habitat 4030 using combined data Sentinel-2 with Sentinel-1 (Schmidt et al., 2018). Mires were successfully classified using RapidEye data, e.g. 7120 with accuracy of F1 equal 91% (Stenzel et al., 2014). It can be concluded that these habitats, under some conditions, can be classified based on the MS data at a high level of accuracy.

On the other hand, it should be stated that the efficiency in habitat differentiation using MS data is relatively diverse and dependent on the areas, especially for 6510 and 7140 (Fig. 14). Significant differences between the differentiating bands were noticed for 6410 and 6510. As well as for the area, the differences were noticed for three campaigns within one habitat, therefore, it is impossible to define the most efficient campaign. The classification of Natura 2000 habitats based on the MS data is less efficient compared to the HS data and can be dependent on data and background classes.

The differentiating MS bands were covering most of the analysed bands. On the other hand, the correctness values were much lower compared to the HS data. Perhaps, the number of MS bands is too low for an accurate habitat differentiation, whereas the number of HS bands is too high. Based on the frequency of occurrence of the MS data, no universal differentiating bands could be determined for any habitat. The determined bands vary within habitats, areas, and campaigns. Generally, the most useful is the band 0.444, which appears to be similar to the results of the HS data (Fig. 13).

There are no published analyses which would compare HS and MS data for Natura 2000 habitat classification, therefore, the conducted studies are pioneering in this research area. For almost all cases, the HS (HS_VNIR or HS_SWIR) data was more efficient in habitat differentiation from the background compared to the MS data, which was proven based on correctness values. The only exception was habitat 7140 in the BI1

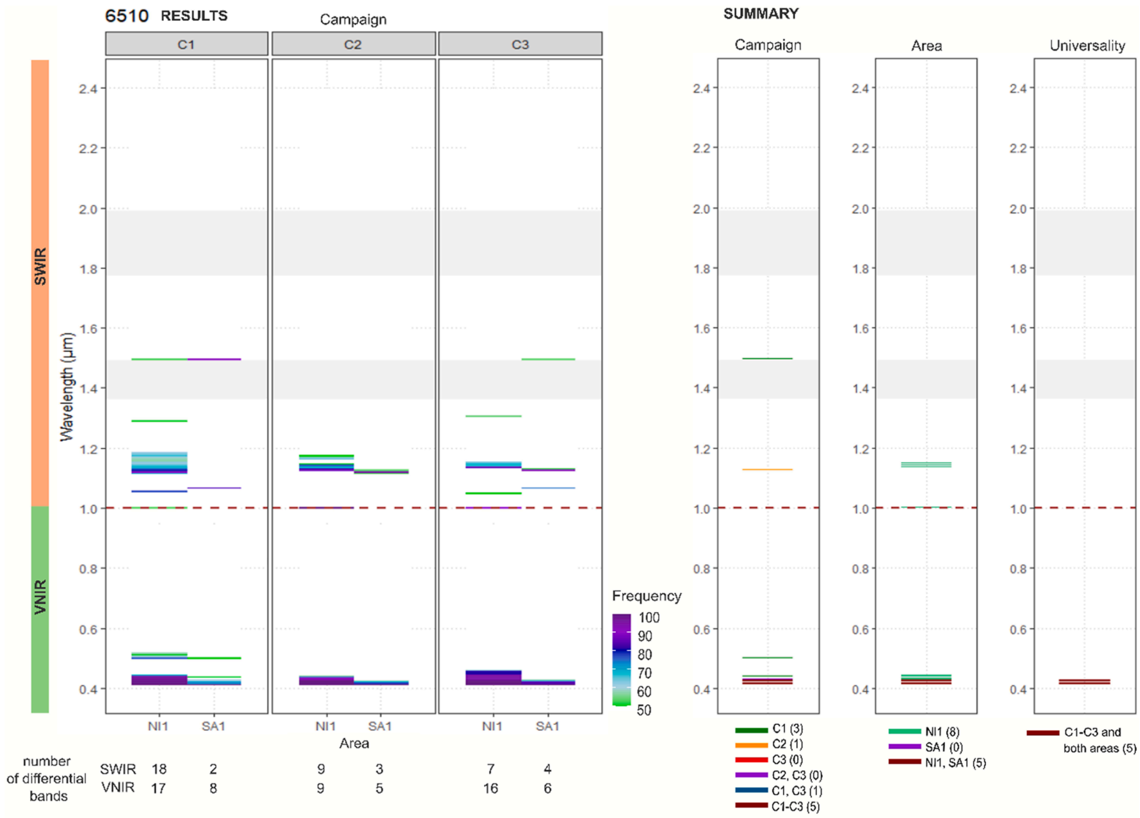

Fig. 11. Bands that differentiate habitat 6510 from the background classes were calculated based on HS_VNIR and HS_SWIR databases for three seasons (C1 - Spring, C2 - Summer, C3 - Autumn) and two areas (NI1 and SA1). The bands were defined as differentiating when the frequency of occurrence value was above 50 out of 100 LDA iterations. In the summary, the following are presented: Campaign - a band was selected if it was differentiating in both areas for a certain campaign; Area - a band was selected if it was differentiating in three campaigns for a specific area; Universality - a band was selected if it was differentiating in both areas in each campaign. The number of differentiating bands is given at the bottom of the figure.

area in C1 and C2 campaigns (Fig. 14). One of the possible reasons for this result is lower quality of HS images caused by unstable lighting conditions in this area in C1 and C2, compared to other HS images. It is likely that resampling to Sentinel-2 resolution reduced the issue, whereas in narrow bands of HS images the problem was visible even after the pre-processing.

Worse differentiation results using MS compared to HS data may be caused by too low number of bands and especially lack of SWIR bands in the Sentinel-2 spectral resolution. The obtained results showed that it is possible to differentiate habitats from background even if using only SWIR data. Different ranges, apart from blue light, were determined as useful for HS and MS datasets. On the other hand, it is not fully correct to compare ranges for HS and MS data due to completely different spectral resolution of the Sentinel-2 and HySpex data. The reflectance values in the band of both sensors can be related to different vegetation biochemical and biophysical properties.

What is more, the LDA analysis was performed on typical polygons (conservation status - Favourable - FV) for each habitat, without transition zones, patches in poor condition or degraded. All mentioned habitat patches should be identified as a habitat in the classification process. This can cause difficulties in classification, even for the habitats with high correctness values and many differentiating bands. The MS data can be used for classification under certain conditions:

- the method identifies a single habitat, not the mixture of different habitats,
- the background on classified area has different spectral properties then the habitat,
- habitat has uniform structure and species composition.

4.3. Application of acquired results

Conclusions regarding the significance of specific spectral regions (VNIR, SWIR) or comparison of HS and MS data were based on studies conducted in two separate areas (Table 1). Due to the methodology adopted in this way, conclusions concerning the possibility of identifying individual Natura 2000 habitats are universal.

Based on the results, it can be stated that acquiring HS images from May to September is optimal to differentiate the habitat from the background. The differences in the average correctness rate between individual campaigns are low and vary from 90% to 100% (Fig. 14). Correctness values for the MS data are much more diverse during the vegetation season, the minimum average correctness value was around 40% (NI1 area, campaign 2, habitat 6510). It is not known yet what is the reason for such high variability in the results for the MS data. This may be caused by different spectral similarity of the habitat to the background in different areas. The NI1 area is characterized by very different meadows communities, which can be similar to the habitat patches (Table 1). In this case, use of MS data seems to be ineffective. This limits the use of the MS data to identify Natura 2000 habitats, especially, in the case of commercial satellite data, where the use of multitemporal data is limited. Identification results based on only one MS image may result in low classification accuracy.

However, it should be noted that for MS data, an increase in classification accuracy can occur as a result of using multi-temporal data. The effective use of multitemporal data fusion has been proven, for example, on satellite RapidEye data (Schuster et al., 2015) or Sentinel-2 (Tarantino et al., 2021). Multitemporal fusion using HS data was used less frequently due to limited availability of HS data and large amount of datasets, which make classification long lasting (Marcinkowska-Ochtyra

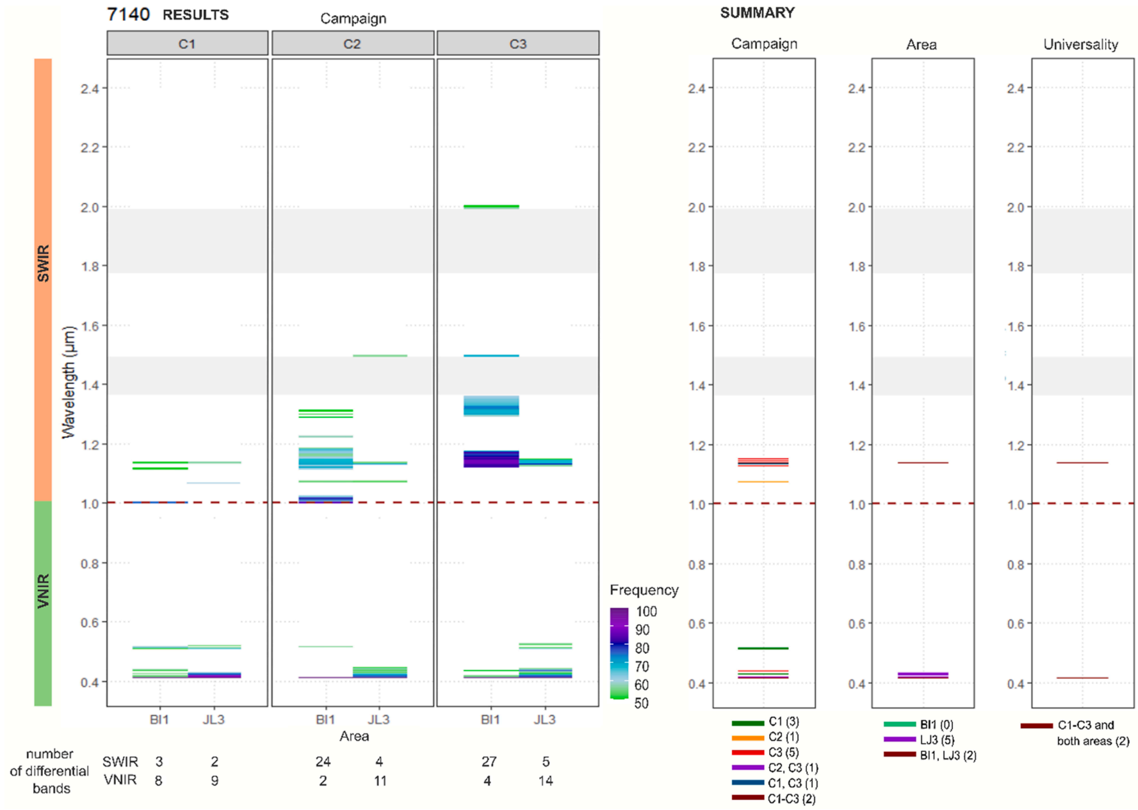

Fig. 12. Bands that differentiate habitat 7140 from the background classes were calculated based on HS_VNIR and HS_SWIR databases for three seasons (C1 - Spring, C2 - Summer, C3 - Autumn) and two areas (BI1 and LJ3). The bands were defined as differentiating when the frequency of occurrence value was above 50 out of 100 LDA iterations. In the summary, the following are presented: Campaign - a band was selected if it was differentiating in both areas for a certain campaign; Area - a band was selected if it was differentiating in three campaigns for a specific area; Universality - a band was selected if it was differentiating in both areas in each campaign. The number of differentiating bands is given at the bottom of the figure.

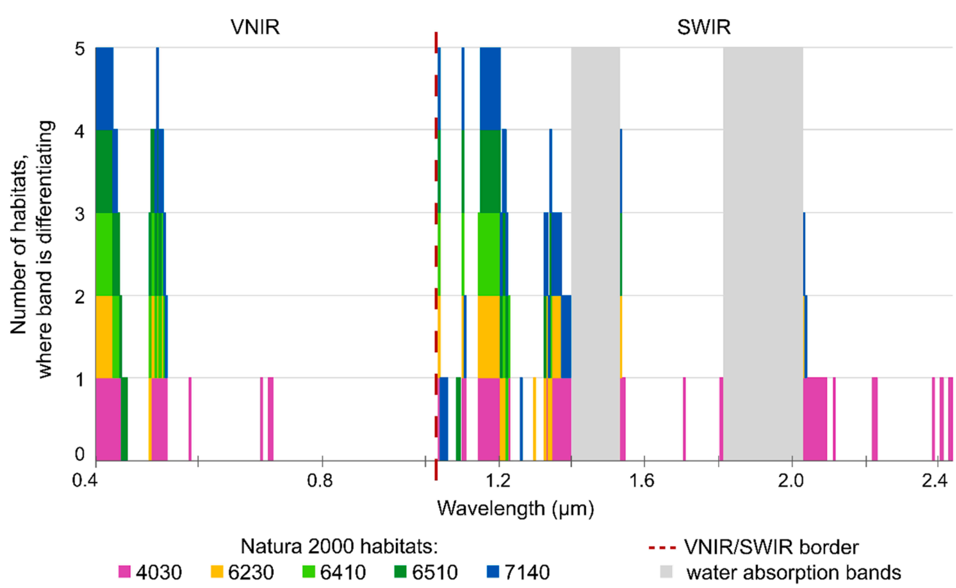

Fig. 13. The HS bands useful for five habitats differentiation from the background. The band was marked as useful if it was differentiating in at least one campaign for at least one area.

et al., 2019). Acquiring hyperspectral data is also expensive.

The conducted research provided knowledge about the relation between analysed habitats and reflectance. Such information can be useful for the development of new hyperspectral sensors dedicated to specific applications.

5. Conclusions

The above study analysed the discrimination of five selected non-forest Natura 2000 habitats (heaths, mires, meadows and grasslands) using hyperspectral data in the range of VNIR (0.4–1.0 μm) and SWIR

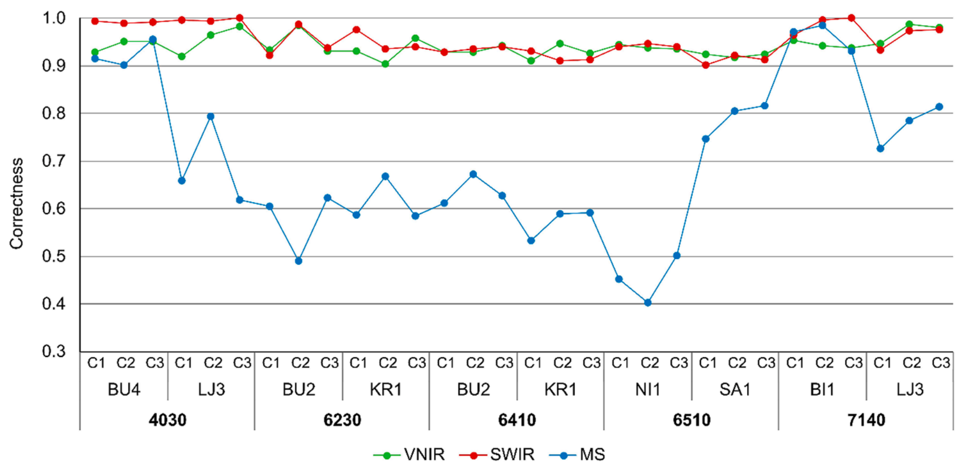

Fig. 14. Average correctness rate values according to the data used: HS_VNIR, HS_SWIR, and MS for each data set analysed (five habitats in three campaigns in two areas).

(1.0–2.5 μm) as well as multispectral data resampled to the Sentinel-2 spectral resolution.

The research described above, related to Natura 2000 habitats identification, resulted in the following conclusions:

- The potential of the MS data (Sentinel-2) is diverse and dependent on the habitat and background classes present in the analysed area. The MS data (Sentinel-2) could be efficient in differentiation of 4030 (heaths - average correctness rate 0.81) and 7140 (mires - average correctness rate – 0.87). In case of meadows (6410, 6510 - average correctness value 0.61) and grasslands (6230 - average correctness value 0.59), the efficiency of differentiation is significantly lower.
- For the HS data (both VNIR and SWIR range) acquired from May to September, very high and stable (reaching over 90%) discrimination of Natura 2000 habitat from the background was observed, regardless of the examined area.
- Differentiation spectral ranges were identified, regardless of the habitat type, season, and area. It included the following ranges: 0.416–0.442 μm 0.502–0.522 μm for VNIR and 1.117–1.165 μm, 1.290–1.361 μm for SWIR.
- The analyses conducted in this study were performed on typical polygons (with favourable conservation status) for each Natura 2000 habitat, excluding transition zones, patches in poor condition or degraded. Therefore, the acquired correctness value could possibly be lower while analysing this sort of patches.

CRediT authorship contribution statement

Anna Jarocińska: Conceptualization, Methodology, Validation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Dominik Kopeć:** Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft, Writing – review & editing. **Marlena Kycko:** Visualization, Writing – review & editing. **Hubert Piórkowski:** Acquisition of ground reference data. **Agnieszka Błońska:** Acquisition of ground reference data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The study was co-financed by the Polish National Centre for Research and Development (NCBR) and MGGP Aero under the programme "Natural Environment, Agriculture and Forestry" BIOSTRATEG II.: The innovative approach supporting monitoring of non-forest Natura 2000 habitats, using remote sensing methods (HabitARS), project number: BIOSTRATEG2/297915/3/NCBR/2016. The Consortium Leader is MGGP Aero. The project partners include the University of Lodz, the University of Warsaw, Warsaw University of Life Sciences, the Institute of Technology and Life Sciences, the University of Silesia in Katowice, Warsaw University of Technology.

References

Belcore, E., Pittarello, M., Lingua, A.M., Lonati, M., 2021. Mapping Riparian Habitats of Natura 2000 Network (91E0*, 3240) at Individual Tree Level Using UAV Multi-Temporal and Multi-Spectral Data. Remote Sens. 13, 1756. https://doi.org/10.3390/rs13091756.

Buck, O., Millán, V.E.G., Klink, A., Pakzad, K., 2015. Using information layers for mapping grassland habitat distribution at local to regional scales. Int. J. Appl. Earth Obs. Geoinf. 37, 83–89. https://doi.org/10.1016/j.jag.2014.10.012.

Chan, J.-W., Paelinckx, D., 2008. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sens. Environ. 112 (6), 2999–3011. https://doi.org/10.1016/j.rse.2008.02.011.

Çimtay, Y., İlk, H.G., 2018. A novel bilinear unmixing approach for reconsideration of subpixel classification of land cover. Comput. Electron. Agric. 152, 126–140. https://doi.org/10.1016/j.compag.2018.07.003.

Demarchi, L., Kania, A., Ciężkowski, W., Piórkowski, H., Oświecimska-Piasko, Z., Chormański, J., 2020. Recursive Feature Elimination and Random Forest Classification of Meadows and Dry Grasslands in Lowland River Valleys of Poland Based on Airborne Hyperspectral and LiDAR Data Fusion (preprint). Earth Sci. https://doi.org/10.20944/preprints202004.0144.v1.

Díaz Varela, R.A., Ramil Rego, P., Calvo Iglesias, S., Muñoz Sobrino, C., 2008. Automatic habitat classification methods based on satellite images: A practical assessment in the NW Iberia coastal mountains. Environ. Monit. Assess. 144 (1-3), 229–250. https://doi.org/10.1007/s10661-007-9981-y.

Ferreira, M.P., Zortea, M., Zanotta, D.C., Shimabukuro, Y.E., de Souza Filho, C.R., 2016. Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. Remote Sens. Environ. 179, 66–78. https://doi.org/10.1016/j.rse.2016.03.021.

Friedman, J.H., 1989. Regularized Discriminant Analysis. J. Am. Stat. Assoc. 84 (405), 165–175. https://doi.org/10.1080/01621459.1989.10478752.

Gao, J., Xu, L., 2016. A Novel Spatial Analysis Method for Remote Sensing Image Classification. Neural Process. Lett. 43 (3), 805–821. https://doi.org/10.1007/s11063-015-9447-0.

Gates, D.M., 1970. Physical and physiological properties of plants, in: National Research Council Agricultural Board (Ed.), Remote Sensing with Special Reference to Agriculture and Forestry. National Academy of Sciences, Washington, pp. 224–252.

Gite, H.R., Solankar, M.M., Surase, R.R., Kale, K.V., 2019. Comparative Study and Analysis of Dimensionality Reduction Techniques for Hyperspectral Data. In: Santosh, K.C., Hegadi, R.S. (Eds.), Recent Trends in Image Processing and Pattern Recognition. Communications in Computer and Information Science. Springer, Singapore, pp. 534–546. https://doi.org/10.1007/978-981-13-9181-1_47.

Gonçalves, J., Henriques, R., Alves, P., Sousa-Silva, R., Monteiro, A.T., Lomba, Â., Marcos, B., Honrado, J., Rocchini, D., 2016. Evaluating an unmanned aerial vehicle-based approach for assessing habitat extent and condition in fine-scale early successional mountain mosaics. Appl. Veg. Sci. 19 (1), 132–146. https://doi.org/10.1111/avsc.12204.

Guo, B., Gunn, S.R., Damper, R.I., Nelson, J.D.B., 2006. Band Selection for Hyperspectral Image Classification Using Mutual Information. IEEE Geosci. Remote Sens. Lett. 3 (4), 522–526. https://doi.org/10.1109/LGRS.2006.878240.

Haest, B., Vanden Borre, J., Spanhove, T., Thoonen, G., Delalieux, S., Kooistra, L., Mücher, C., Paelinckx, D., Scheunders, P., Kempeneers, P., 2017. Habitat mapping and quality assessment of NATURA 2000 heathland using airborne imaging spectroscopy. Remote Sens. 9 (3), 266. https://doi.org/10.3390/rs9030266.

Haest, B., Thoonen, G., Vanden Borre, J., Spanhove, T., Delalieux, S., Bertels, L., Kooistra, L., Mucher, S., Scheunders, P., 2010. An object-based approach to quantity and quality assessment of heathland habitats in the framework of natura 2000 using hyperspectral airborne ahs images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, p. 38.

He, T., Wang, S., 2021. Multi-spectral remote sensing land-cover classification based on deep learning methods. J. Supercomput. 77 (3), 2829–2843. https://doi.org/10.1007/s11227-020-03377-w.

Hladik, C., Schalles, J., Alber, M., 2013. Salt marsh elevation and habitat mapping using hyperspectral and LIDAR data. Remote Sens. Environ. 139, 318–330. https://doi.org/10.1016/j.rse.2013.08.003.

HySpex [WWW Document], 2021. URL https://www.hyspex.com/ (accessed 7.7.21).

Ichter, J., Evans, D., Richard, D., 2014. Terrestrial habitat mapping in Europe: an overview., EEA Technical report 1/2014. Publications Office of the European Union, Luxembourg.

James, G., Witten, D., Hastie, T., Tibshirani, R. (Eds.), 2013. An introduction to statistical learning: with applications in R, Springer texts in statistics. Springer, New York.

Jarocińska, A., Kopeć, D., Tokarska-Guzik, B., Raczko, E., 2021. Intra-Annual Variabilities of Rubus caesius L. Discrimination on Hyperspectral and LiDAR Data. Remote Sens. 13, 107. https://doi.org/10.3390/rs13010107.

Jayaprakash, C., Damodaran, B.B., Viswanathan, S., Soman, K.P., 2018. Randomized independent component analysis and linear discriminant analysis dimensionality reduction methods for hyperspectral image classification. J. Appl. Rem. Sens. 14, 1. https://doi.org/10.1117/1.JRS.14.036507.

Jensen, J.R., 1983. Biophysical Remote Sensing. Ann. Assoc. Am. Geogr. 73 (1), 111–132. https://doi.org/10.1111/j.1467-8306.1983.tb01399.x.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., 2021. Package "Caret" Classification and Regression Training.

Lenhard, K., Baumgartner, A., Schwarzmaier, T., 2015. Independent Laboratory Characterization of NEO HySpex Imaging Spectrometers VNIR-1600 and SWIR-320m-e. IEEE Trans. Geosci. Remote Sens. 53 (4), 1828–1841. https://doi.org/10.1109/TGRS.2014.2349737.

Marcinkowska-Ochtyra, A., Gryguc, K., Ochtyra, A., Kopeć, D., Jarocińska, A., Sławik, Ł., 2019. Multitemporal Hyperspectral Data Fusion with Topographic Indices—Improving Classification of Natura 2000 Grassland Habitats. Remote Sens. 11, 2264. https://doi.org/10.3390/rs11192264.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Simpson, G., Solymos, P., Stevenes, M., Wagner, H., 2019. Vegan: Community Ecology Package. R package version 2.5-6.

OrthoVista [WWW Document], 2020. URL http://www.amigooptima.com/trimble-inpho/ortho-vista.php (accessed 12.18.20).

PARGE Airborne Image Rectification [WWW Document], 2021. URL http://www.rese-apps.com/software/parge/index.html (accessed 7.7.21).

Possibilities for updating map layers of NATURA 2000 biotopes using advanced remote sensing methods - Faculty of Environmental Sciences CZU Prague [WWW Document], 2020. URL https://www.fzp.czu.cz/en/r-9411-projects-and-partnerships/r-9880-projects/r-16288-possibilities-for-updating-map-layers-of-natura-2000-biotopes-using-advanced-remote-sensing-methods (accessed 11.3.21).

Rapinel, S., Rozo, C., Delbosc, P., Bioret, F., Bouzillé, J.-B., Hubert-Moy, L., 2020. Contribution of free satellite time-series images to mapping plant communities in the Mediterranean Natura 2000 site: the example of Biguglia Pond in Corse (France). Mediterranean Botany 41, 181–191. https://doi.org/10.5209/mbot.66535.

Richter, R., Schlapfer, D., 2020. Atmospheric / Topographic Correction for Airborne Imagery 196.

Ripley, B., Venables, B., Bates, D.M., Hornik, K., Gebhardt, A., Firth, D., 2021. Support Functions and Datasets for Venables and Ripley's MASS.

Roever, C., Raabe, N., Luebke, K., Ligges, U., Szepannek, G., Zentgraf, M., Meyer, D., 2020. klaR: Classification and Visualization.

Schmidt, J., Fassnacht, F.E., Förster, M., Schmidtlein, S., Nagendra, H., Atzberger, C., 2018. Synergetic use of Sentinel-1 and Sentinel-2 for assessments of heathland conservation status. Remote Sens. Ecol. Conserv. 4 (3), 225–239. https://doi.org/10.1002/rse2.68.

Schuster, C., Schmidt, T., Conrad, C., Kleinschmit, B., Förster, M., 2015. Grassland habitat mapping by intra-annual time series analysis – Comparison of RapidEye and TerraSAR-X satellite data. Int. J. Appl. Earth Obs. Geoinf. 34, 25–34. https://doi.org/10.1016/j.jag.2014.06.004.

Sławik, Ł., Niedzielko, J., Kania, A., Piórkowski, H., Kopeć, D., 2019. Multiple Flights or Single Flight Instrument Fusion of Hyperspectral and ALS Data? A Comparison of their Performance for Vegetation Mapping. Remote Sens. 11, 970. https://doi.org/10.3390/rs11080970.

Song, F., Mei, D., Li, H., 2010. Feature Selection Based on Linear Discriminant Analysis, in: 2010 International Conference on Intelligent System Design and Engineering Application. Presented at the 2010 International Conference on Intelligent System Design and Engineering Application (ISDEA), IEEE, Changsha, Hunan, China, pp. 746–749. https://doi.org/10.1109/ISDEA.2010.311.

Stenzel, S., Feilhauer, H., Mack, B., Metz, A., Schmidtlein, S., 2014. Remote sensing of scattered Natura 2000 habitats using a one-class classifier. Int. J. Appl. Earth Obs. Geoinf. 33, 211–217. https://doi.org/10.1016/j.jag.2014.05.012.

Szporak-Wasilewska, S., Piórkowski, H., Ciężkowski, W., Jarzombkowski, F., Sławik, Ł., Kopeć, D., 2021. Mapping Alkaline Fens, Transition Mires and Quaking Bogs Using Airborne Hyperspectral and Laser Scanning Data. Remote Sens. 13, 1504. https://doi.org/10.3390/rs13081504.

Tarantino, C., Forte, L., Blonda, P., Vicario, S., Tomaselli, V., Beierkuhnlein, C., Adamo, M., 2021. Intra-Annual Sentinel-2 Time-Series Supporting Grassland Habitat Discrimination. Remote Sens. 13, 277. https://doi.org/10.3390/rs13020277.

Tharwat, A., Gaber, T., Ibrahim, A., Hassanien, A.E., 2017. Linear discriminant analysis: A detailed tutorial. AI Commun. 30 (2), 169–190. https://doi.org/10.3233/AIC-170729.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S., Van Der Meer, B., 2004. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. Remote Sens. Environ. 91 (3-4), 354–376. https://doi.org/10.1016/j.rse.2004.03.013.

Using satellite images to improve the operation of the Natura 2000 network: A prototype for monitoring Natura2000 sites habitats with Copernicus 2021 [WWW Document], 2021. URL http://www.cop4n2k.eu/ (accessed 11.3.21).

Zagajewski, B., Tømmervik, H., Bjerke, J.W., Raczko, E., Bochenek, Z., Kłos, A., Jarocińska, A., Lavender, S., Ziółkowski, D., 2017. Intraspecific Differences in Spectral Reflectance Curves as Indicators of Reduced Vitality in High-Arctic Plants. Remote Sens. 9, 1289. https://doi.org/10.3390/rs9121289.

Zlinszky, A., Schroiff, A., Kania, A., Deák, B., Mücke, W., Vári, Á., Székely, B., Pfeifer, N., 2014. Categorizing Grassland Vegetation with Full-Waveform Airborne Laser Scanning: A Feasibility Study for Detecting Natura 2000 Habitat Types. Remote Sens. 6, 8056–8087. https://doi.org/10.3390/rs6098056.