MARCJANNA JĘDRYCH, BOGDAN ZAGAJEWSKI, ADRIANA MARCINKOWSKA-OCHTYRA
University of Warsaw, Faculty of Geography and Regional Studies
Department of Geoinformatics, Cartography and Remote Sensing
m.jedrych@student.uw.edu.pl; bogdan@uw.edu.pl; adriana.marcinkowska@uw.edu.pl

Application of Sentinel-2 and EnMAP new satellite data to the mapping of alpine vegetation of the Karkonosze Mountains

Abstract. Effective assessment of environmental changes requires an update of vegetation maps as it is an indicator of both local and global development. It is therefore important to formulate methods which would ensure constant monitoring. It can be achieved with the use of satellite data which makes the analysis of hard--to-reach areas such as alpine ecosystems easier.

Every year, more new satellite data is available. Its spatial, spectral, time, and radiometric resolution is improving as well. Despite significant achievements in terms of the methodology of image classification, there is still the need to improve it. It results from the changing needs of spatial data users, availability of new kinds of satellite sensors, and development of classification algorithms. The article focuses on the application of Sentinel-2 and hyperspectral EnMAP images to the classification of alpine plants of the Karkonosze (Giant) Mountains according to the: Support Vector Machine (SVM), Random Forest (RF), and Maximum Likelihood (ML) algorithms. The effects of their work is a set of maps of alpine and subalpine vegetation as well as classification error matrices. The achieved results are satisfactory as the overall accuracy of classification with the SVM method has reached 82% for Sentinel-2 data and 83% for EnMAP data, which confirms the applicability of image data to the monitoring of alpine plants.

Keywords: Sentinel-2, EnMAP, classification, alpine vegetation, satellite systems

1. Introduction

The elements of plant cover have been included on maps for centuries. However, the beginnings of thematic plant mapping date back to as late as the 19th century. The pioneers in this field were the European researchers such as, among others, J.F. Schouw (1823), O. Sendtner (1854), C.F.P. Martius (1858), A. von Humboldt, and A. Bonpland (1895). The boom of vegetation mapping took place after the World War I. During that period, numerous institutions specialising in this field were established. For example, in France – Service de la Carte de la Végétation de la France au 1:200 000, in Switzerland – Stiftung Geobotanisches Forschungsinstitut Rübel, in the then Soviet Union – Komarov Botanical Institute of the Russian Academy of Sciences (A. Küchler, I. Zonneveld 1988). Moreover, the first vegetation maps were created, for instance, the vegetation map of

Switzerland at the scale of 1:20 000 (E. Schmidt 1940), the vegetation map of the Montpellier region (J. Braun-Blanquet, Y.T. Chou 1947), or the fitosociological map of the Montelago vegetation (F. Pedrotti 1967). Intuitive and inaccurate marking of the borders between individual divisions posed a major limitation to these studies (T. Dirnböck et al. 2003).

The development of photointerpretation of aerial pictures and new remote sensing tools increased imaging in order to determine the expanse of single divisions have revolutionised the field of vegetation mapping. Satellite sensors which record electromagnetic radiation at different electromagnetic spectrum bands enabled the identification of single plant patches (J.R. Jensen 1983, F.C. Billingsey 1984, A. Bannari et al. 1995) and precise marking of the borders between individual vegetation communities on the basis of satellite images. In many cases this resulted in the automated creation

of maps. The remote sensing allowed for the monitoring of vegetation with regional or continental coverage (D.A. Quattrochi, J.C. Luvall 1999; A. Jarocińska, B. Zagajewski 2009; B. Zagajewski 2010; S. Autosh 2012; M. Kycko et al. 2017; A. Marcinkowska-Ochtyra et al. 2017). What is more, the increase in popularity of satellite data stems, among other things, from the fact that it is up-to-date, consistent and recurrent. Such data also makes it possible to conduct the analysis in large areas. Since 1972, the most commonly used environmental data has been received from Landsat satellites. However, the spatial and spectral resolution of its equipment (MSS in the case of ERTS/Landsat 1, 2, and 3, TM – Landsat 4 and 5, ETM+ – Landsat 7 and OLI – Landsat 8) was not sufficient to make thorough analysis, including the mapping of heterogeneous vegetation – the category which alpine plants belong to. Since the mid-nineties, researchers have greatly focused on the improvement of spatial, time and spectral resolution in the field of satellite remote sensing – increase in the number of spectral bands with simultaneous narrowing of their bandwidth. The current maximum spatial resolution of civil remote sensing data is 30 cm (panchromatic WorldView-3 and WorldView-4 optical satellites), whereas the shortest recurrence time of data recording in Poland is about few days (COSMOSky-Med radar satellite). Such good sensor parameters are considered the most significant when conducting precise analysis, for example, detection of vegetation stress. Unfortunately, the data is not available free of charge, which greatly limits its applicability.

The changes introduced in this area seem to be the result of the actions taken by the European Space Agency (ESA) within the Copernicus programme, the aim of which is to build the system of environmental monitoring through the strengthening of Europe's potential in terms of sensor and satellite construction as well as provide European recipients with data for its practical use (M. Berger et al. 2012). As part of the mission, the attempts to put devices examining vegetation in orbit have already been made. The Proba V mission, which is the continuation of the SPOT-VÉGÉTATION project carried out in the years 1998–2014, provides everyday data that can be used for vegetation and environmental monitoring (it is used to determine indicators such as NDVI and fAPAR). The FLEX (Fluorescence Explorer) satellite is going to be put in orbit in 2022. Its FLORIS device will take fluorescence measurements, which will make it possible to control the circulation of carbon dioxide between plants and the environment. The proper application of this data will necessitate the formulation of suitable methods of its processing. The effect of the recent work conducted by the European Space Agency and European Union is also the Sentinel-2 satellite mission under the Sentinel project which is a part of the Copernicus programme. The aim of the Sentinel mission is to build a modern system to monitor the world with a global scope. It will be possible thanks to 10 satellites that are going to be put in orbit. They will operate in five separate systems, each of which will consist of two satellites moving synchronously each of which will deliver different kinds of data.

The German EnMAP (Environmental Mapping and Analysis Programme) mission may bring a new quality to environmental research. The EnMAP satellite, which is planned to be launched in 2018, will provide cost-free hyperspectral data. Its spatial resolution is in fact only 30 m but it will record the radiation in 242 narrow bands and 2 bandwidths: 420–1000 nm (88 bands) and 900–2450 nm (154 bands), which is a groundbreaking solution with features comparable to the parameters of devices for aerial recording (H. Kaufmann et al. 2012). Technical details of the EnMAP sensor are presented in table 1.

Currently, simulated images based on the set of input data – other multispectral and hyperspectral images – can be received. The si-

Table 1. Technical parameters of EnMAP sensor

Spectral parameters	VNIR	SWIR
Spectral range	420−1000 nm	900−2450 nm
Numer of bands	88	154
Spatial parameters		
Spatial resolution (at the nadir point)	30 m	
Time parameters		
Revisit time	23 days	

mulation is done with the use of the EnMAP end-to-end Simulation Tool (EeteS).

Due to the on-line availability of Sentinel-2 and planned release of EnMAP satellite images, it is proper to formulate optimum methods for its processing and mapping of alpine vegetation, with strong focus on the selection of a classifying algorithm. This would allow for precise mapping of vegetation divisions and help to monitor the natural environment.

Alpine vegetation communities belong to unique ecosystems which are a significant component of the environment as well as a key indicator of the occurring changes since numerous species, often endemic, can be found in a small area. The field research of alpine vegetation is relatively difficult mainly due to a short growing season, its high changeability, significant delevelling and difficulties with exploration of the terrain. It is therefore recommended to apply the data and remote sensing techniques to mapping and researching alpine ecosystems (B. Zagajewski et al. 2005, A. Marcinkowska et al. 2014, A.M. Jarocińska et al. 2016, A. Ochtyra et al. 2016).

The information delivered by the Sentinel-2 system is particularly valuable for vegetation mapping. The system comprises two satellites: Sentinel-2A, launched on 23 June 2015, and Sentinel-2B, put in orbit on 7 March 2017). They are equipped with Multi-Spectral Instrument (MSI) with 13 spectral bands and bandwidths ranging from 443 to 2190 nm (4 of them with a narrower spectral range) and resolution between 10 and 60 m (Thales Alenia Space 2016). The data received from this device is exceptionally valuable in order to conduct the vegetation analysis as the values of reflected radiation within the recorded range make it possible to assess its condition (J. Delegido et al. 2011). Selected parameters of the Sentinel-2 sensor are presented in <u>table 2</u>.

Due to its specificity, Sentinel-2 data has already been successfully applied to the classification and analysis of the land cover (H. Topaloğlu et al. 2016) and aquatic plants (D. Stratoulius et al. 2015), crop and forest tree species (M. Immitzer et al. 2016), and finally alpine vegetation (L. Kupková et al. 2017).

So far, simulated EnMAP data has been tested by researchers who specialise in the identification of minerals (N. Yokoya et al. 2016, C. Mielke et al. 2016, N.K. Bösche et al. 2014) and vegetation. As far as the latter is concerned, the major focus was on the detection of water stress in forests and coastal vegetation (S. Nink et al. 2015, J. Stoffles et al. 2015, S. Doltzer et al. 2015), measurement of plant indicators (M. Locherer et al. 2015, B. Siegamnn et al. 2015), monitoring of vegetation (P. Leitão et al. 2015), and vegetation mapping (A. Marcinkowska--Ochtyra et al. 2017, A. Braun et al. 2015, S. Nink et al. 2015, S. Suess et al. 2015). EnMAP Box software, specifically dedicated to EnMAP data, turned out to be particularly useful for doing the analysis. It contains a number of built-in modules which enable quick and precise analysis. Special classification modules are useful for the classification of vegetation (imageSVM – Support Vector Machine for Classification and Regression, imageRF – Random Forest for Classification and Regression, imagePLSR – Partial Least Square Regression).

Most classification which has been done so far both with the use of Sentinel-2 data and simulated EnMAP data was based on mappers which belong to the group of algorithms representing machine learning – the field that has been dynamically developing recently – applying the methods of so-called Support Vector Machine (SVM) or Random Forest (RF). The development in machine learning results primarily from the growth in the number of variables which

Table 2. Key parameters of Sentinel-2 sensor

Spatial resolution	Spectral resolution				
10 m	490 nm (blue)	560 nm (green)		665 nm (red)	842 nm (NIR)
20 m	Red-Edge			SWIR	
	705 nm	740 nm	783 nm	865 nm	1610 nm	2190 nm
60 m	443 nm		945 nm		1375 nm	
Revisit time	10 days (5 days for whole system)				

are used for classification and the volume of this data (M. Pesaresi et al. 2016). In such conditions, highly good results started to emerge with the application of a machine learning algorithm that is considered artificial intelligence (J. Tomczak 2013).

The Support Vector Machine algorithm, described by V. Vapnik (1995), is an active learning method which is used to divide classified pixels by a mathematical function. If such an operation cannot be performed in a two-dimensional space, classified pixels are transferred to spaces with a larger number of dimensions in which they can be divided. This is the actual advantage of this method. It enables effective classification of heterogeneous pixels in terms of spectral reflectance (so-called mixels), which contain some smaller objects belonging to different classes. This often occurs in a situation when classification is made on the basis of data with a large size of a pixel such as Sentinel-2 and EnMAP images. The SVM method has therefore been often applied by scientists conducting their research based on this data with positive results. High overall accuracy, which is one of the basic measures used to assess the classification reliability that is defined as quotient of correctly classified pixels to all pixels in the error matrix (B. Zagajewski 2010), has been obtained by many researchers. A team led by H. Topaloğlu (H. Topaloğlu et al. 2016), which used Sentinel-2 data, classified the land cover in Turkey with the overall accuracy at the level of 82%. D. Stratoulias and others (2015) also assessed (by visual assessment) their classification of aquatic vegetation of Balaton correctly using the SVM method. S. Suess and others (2015), who made their classification on the basis of EnMAP data concerning shrub vegetation in Portugal, obtained the overall accuracy of 81%, whereas A. Braun and others (2015) classified the land cover in Bavaria, using the SVM method, with the overall accuracy of 85%.

The Random Forest method applies the technique of selecting k observation vectors from an n-element sample which is used to construct a decision tree. A given number of m elements (where k>m) are selected from every tree node. Then, out of this group, n vectors (where m>n) are selected from another node. The selection continues until the node contains elements belonging to the same class. The elements (such as pixels) are classified to the class in which they are most common in all constructed trees. The Random Forest method was successfully applied to the vegetation mapping based on both multispectral (M. Gartzia et al. 2013) and hyperspectral (mainly aerial) data (P. Burai et al. 2016). The RF algorithm, considered relatively more universal than SVM, that is, not requiring long parametrisation (Q. Feng et al. 2015), was also applied to the research on Sentinel-2 and EnMAP data. M. Immitzer and others (2016), using the RF algorithm and achieving high overall accuracy, made the classification of crop species (OA – 72%) and trees (OA – 76%), which is comparable to the overall accuracy of 74% achieved by A. Braun and others (2015) in their classification of land cover.

An algorithm which is not included in the category of machine learning is the Maximum Likelihood method (ML). This algorithm has often been used for the classification of multispectral data to calculate the likelihood of a pixel to belong to a given class with the standard distribution of data, taking variations of spectral reflectance in each class also into account. It requires that sufficient amount of training data is entered (it must exceed the number of imaging channels), which may turn out to be problematic when it comes to classifying hyperspectral data. The algorithm is well applicable to data with standard distribution. What is more, it is relatively fast and available in most programmes for the analysis of raster (ENVI, Saga GIS) and vector data such as ArcGIS (J.B. Campbell, R.H. Wynne 2011). Researches, using this method, achieved satisfactory results in their study of vegetation, including alpine plants, which was based on multispectral data (M. Tobler et al. 2003, M. Gartzia et al. 2013, A. Ochtyra et al. 2016, R. Suchá et al. 2016). Classification based on the ML algorithm has also been applied to Sentinel-2 data. This method was used, among others, by H. Topaloğlu and others (2016), who achieved the overall accuracy of the land cover classification at the level of 76%, as well as D. Traganos and P. Reinartz (2017) who classified aquatic vegetation. EnMAP data has not been classified with the ML method yet, which means that spectral space needs to be sufficiently reduced.

2. Area and subject of research

Area of research which covers the range of the Karkonosze Mountains in the Sudetes Mountains that stretches right in the NW–SE direction along the Polish-Czech border with the highest peak Śnieżka (1602 m a.s.l., <u>fig.1</u>). This mountain range is unique in terms of its

a.s.l.). They comprise primarily of species which are resistant to atmospheric conditions – rock and scree vegetation (mainly different lichens species).

Taking into consideration great environmental diversity of the Karkonosze Mountains, the decision was made to establish the Karkonosze National Park (Karkonoski Park Narodowy/
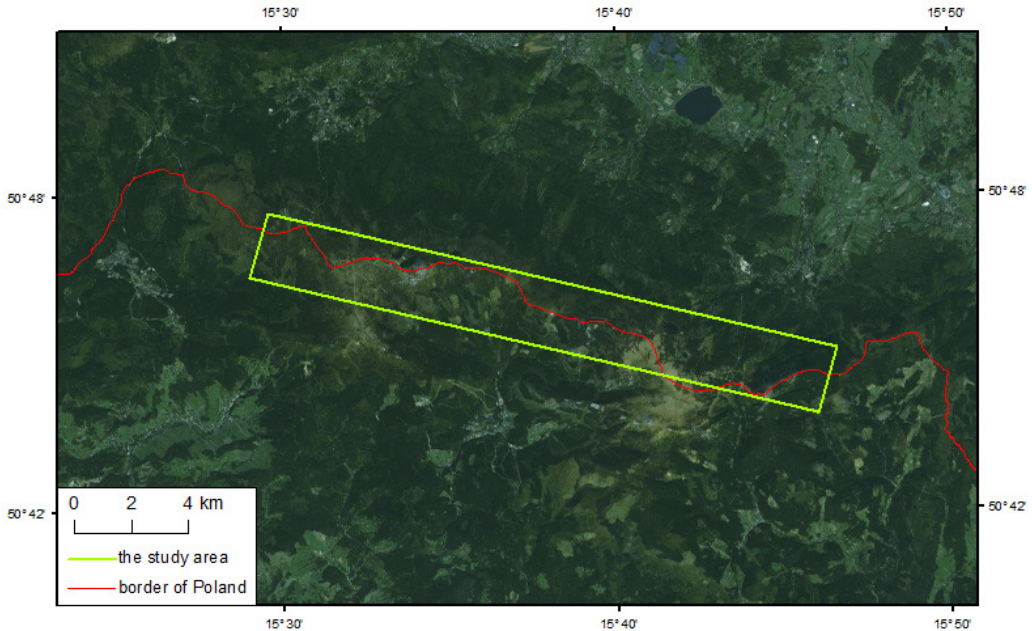

Fig. 1. The study area

geological structure which is characterised with the occurrence of plutonic igneous rocks and metamorphic ones as well as landform with some postglacial formations. The Karkonosze Mountains are also exceptional in terms of their vegetation layers. Areas below 1250 m a.s.l. are mostly covered with forests of lower and upper montane belt, mainly spruce and beech with fir (Sudeten beech). Above this level, up to approximately 1450 m a.s.l., there is subalpine vegetation dominated by dwarf-pine stands. *Calamagrostis villosa g*rasslands, swards with *Nardus stricta* as a dominating species, alpine swards and heathlands are also quite common there. Swamp and mire vegetation appears in the areas rich in water. The upper layers of the Karkonosze Mountains are covered by alpine plants (the boundary is agreed to be at 1450 m

Krkonošský Národní Park) and the UNESCO Karkonosze/Krkonoše Biosphere Reserve in their territory in both Poland and the Czech Republic.

3. Research methods

The classification of alpine vegetation and assessment of its accuracy have been made for seven vegetation classes, according to the model presented in figure 2.

The first phase of work was aimed at acquiring Sentinel-2 and EnMAP input data. They were generated with End-to-End Simulation on the basis of data received from the APEX aerial sensor (on 10 September 2012) in the GFZ German Research Centre for Geosciences in Potsdam. The result was atmospherically and
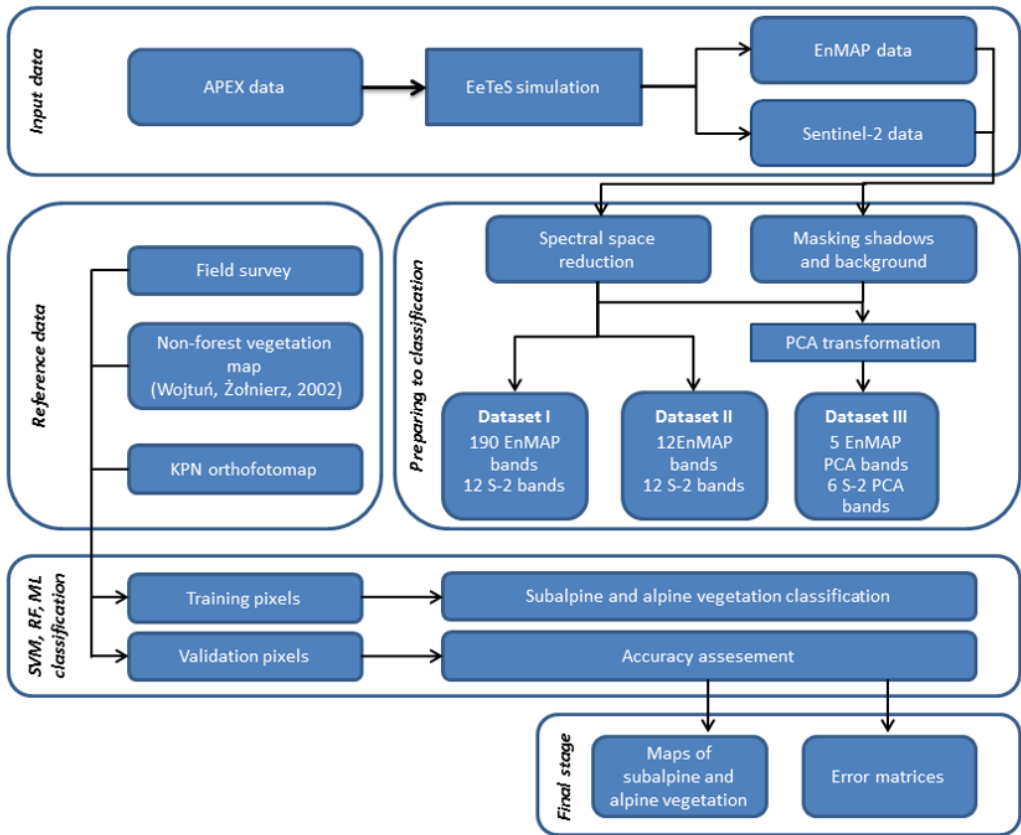
Fig. 2. Classification procedures scheme

geometrically corrected EnMAP and Sentinel-2 images.

The operations were thereafter undertaken, independently of each other, to prepare the data for classification and collect source materials. The preparation of data for classification, due to the fact that the data had been previously corrected both atmospherically and geometrically, was done by means of reducing superfluous bands and creating a mask of shadows and background characteristic for diversified terrain. The outcome of that phase of work were three sets of data for classification, including both EnMAP and Sentinel-2 image, which were to be further classified and analysed separately. The first set included an EnMAP image that consisted of 190 bands and a Sentinel-2 image with 12 bands. The second one consisted of Sentinel-2 data with 12 bands, whereas EnMAP

data was spectrally adjusted to the Sentinel-2 resolution. The third one included the data which was obtained as a result of the Principal Component Analysis (PCA) – five principal components (PCA bands) with the highest informativeness were chosen for EnMAP data, whereas for Sentinel-2 data – six principal components.

Source data was obtained through the access to the vector version of *Mapa roślinności nieleśnej Karkonoskiego Parku Narodowego* (*Map of non-forest vegetation of the Karkonosze National Park*; B. Wojtuń, L. Żołnierz 2002) and maps from geoportals of the Polish and Czech national parks. The team of experts from the University of Warsaw also conducted field observations in order to identify traverses representing separate communities on their itinerary with GPS. The results of the identifi-

cation were saved as a shapefile. All the data was then analysed in order to distinguish training and verification traverses.

The next phase was the classification of vegetation. First, independent training and verification traverses were designated using the ENVI 5.3 software based on the collected testing data. The result was 1060 training patches and 931 verification ones for EnMAP data and 1593 training polygons and 1440 verification ones for Sentinel-2 data. Second, the training polygons served to classify seven types of vegetation in the EnMAP Box software applying SVM and RF algorithms and in the ENVI 5.3 using the ML algorithm. The next step was the verification of the classification results based on the established verification traverses. The assessment of accuracy and applicability of the classifications for vegetation mapping was made on the basis of the generated error matrices and visualisation of the classification in the ENVI 5.3 software.

The final phase was drawing the maps of the alpine and subalpine vegetation of the Karkonosze Mountains in the ArcMap 10.3 software.

4. Results

The best results of all the six classifications made on Sentinel-2 data and eight classifications based on EnMAP data, in terms of overall accuracy and accuracy for individual classes as well as in visual terms, were achieved with the use of the SVM mapper (fig. 3, tables 3 and 4). For Sentinel-2 data, the highest overall accuracy (OA) was obtained for the set of data consisting of six bands as a result of PCA

Table 3. Error matrix for Sentinel-2 classification using SVM (overall accuracy 78,33%)

Class	Correctly classified pixels (%)							
	1	2	3	4	5	6	7	8
1. Subalpine dwarf pine scrubs	**90**	2	0	10	1	2	0	5
2. Heathlands	1	**54**	2	7	23	5	2	1
3. Grasslands	0	18	**79**	3	20	5	0	0
4. Bogs, fens and springs	3	6	10	**73**	8	5	0	0
5. Subalpine tall-forbs	1	18	7	6	**41**	9	0	0
6. Areas without vegetation	0	3	3	0	6	**64**	6	0
7. Rock and scree vegetation	0	0	1	1	1	11	**91**	1
8. Forests	5	0	0	0	0	0	0	**91**

Table 4. Error matrix for EnMAP classification using SVM (overall accuracy 82,92%)

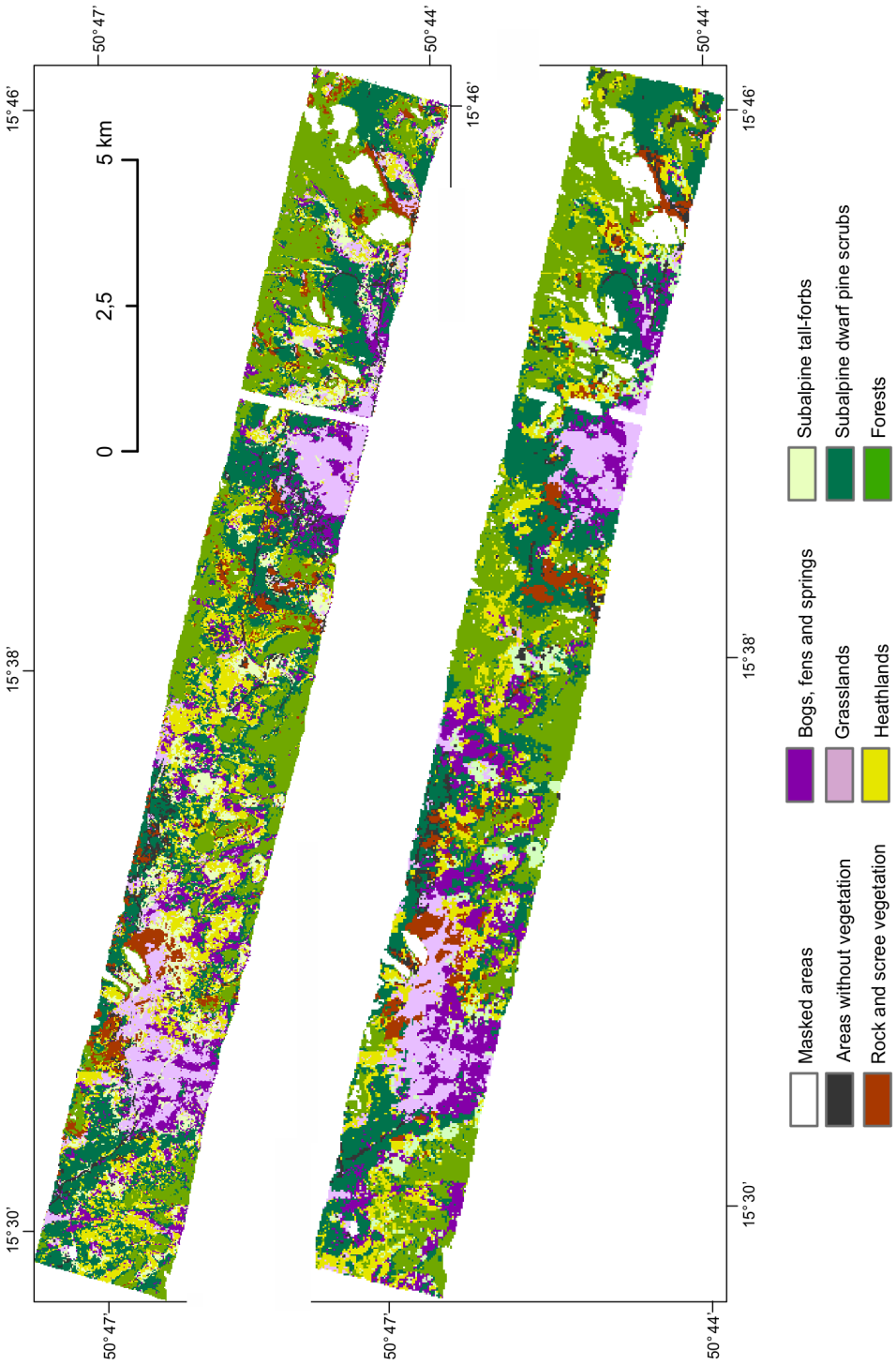| Class	Correctly classified pixels (%)							
	1	2	3	4	5	6	7	8
1. Subalpine dwarf pine scrubs	**81**	0	0	0	4	4	11	0
2. Heathlands	0	**68**	0	8	13	0	4	7
3. Grasslands	0	0	**96**	0	0	0	0	4
4. Bogs, fens and springs	0	2	0	**90**	2	0	6	0
5. Subalpine tall-forbs	0	4	1	6	**69**	1	19	0
6. Areas without vegetation	0	0	0	4	0	**96**	0	0
7. Rock and scree vegetation	6	7	0	5	11	0	**71**	0
8. Forests	0	13	3	1	0	1	1	**80**

Fig. 3. Alpine and subalpine vegetation map derived from: A – Sentinel-2 dataset (SVM classifier), B – EnMAP dataset (SVM classifier)

(OA – 82%), yet worse for the total set of bands (OA – 78%). In the case of EnMAP data, the best overall classification accuracy was obtained for the whole set of 190 bands (OA – 83%). Significantly lower overall accuracy was obtained, however, for the set of EnMAP data which was spectrally reduced to five principal components (OA – 60%). In visual terms, the SVM classifications with the highest accuracy results were very favourable and showed only slight differences from the model map by B. Wojtuń i L. Żołnierz (2002). It should be noted that the classification results were also satisfactory with respect to the areas with complex topography where the occurrence of vegetation is mosaic.

Satisfactory classification results were achieved with the use of the ML mapper. For complete Sentinel-2 data (12 bands), the overall accuracy was at the level of 81%, whereas for spectrally reduced data (6 PCA bands) – 79%. For EnMAP data, the overall accuracy obtained for the data spectrally reduced and adjusted to Sentinel-2 data (12 bands) equalled 75%. However, lower accuracy was obtained for five EnMAP bands as a result of PCA – 63%. In visual terms, the classification was equally satisfactory, yet some of the classes were overestimated at the cost of others (mainly heathlands at the cost of grasslands).

The results of the Sentinel-2 data classification based on RF method were characterised by high overall accuracy (OA for the complete set of bands – 81.9%), however, the visualisation of the classification is incorrect. There is also the salt-and-pepper effect which ought to be corrected. The results of the classification made on EnMAP data were not satisfactory either – the highest overall accuracy (for five EnMAP bands after PCA transformation) was 58.7%. Its visualisation is also insufficient. The areas covered with mosaic vegetation are incorrectly classified and there are classification mistakes in the areas dominated by one homogenous type of vegetation (dwarf-pine stands are classified as forests).

As far as the accuracy at the level of classes is concerned, two measures were applied: producer's accuracy, which is the ratio of the number of correctly classified pixels to total number of verification pixels, and user's accuracy that shows the ratio of the number of correctly classified pixels to the number of training

pixels (R.G. Congalton 1991). The best classification results for Sentinel-2 data (complete data set – 12 bands) were the following: rock and scree vegetation (producer's accuracy: SVM – 92%, ML – 95%, RF – 99%, user's accuracy: SVM – 87%, ML – 99%, RF – 99%) and dwarf-pine stands and forests. The worst results in terms of both producer's and user's accuracy characterised the grasslands class (producer's accuracy: SVM – 41%, ML – 38%, RF – 52%; user's accuracy: SVM – 54%, ML – 45%, RF – 64%).

The results concerning EnMAP data were relatively comparable to the ones obtained for Sentinel-2 data. The best classified species were rock and scree vegetation (producer's accuracy: SVM – 95%, ML – 38%, RF – 96%; user's accuracy: SVM – 86%, ML – 45%, RF – 95%) and forests; classification results for dwarf-pine stands were slightly worse than in the case of Sentinel-2 data. The worst results, as in the case of Sentinel-2 data, were obtained for the grasslands class (producer's accuracy: SVM – 71%, ML – 32%, RF – 28%; user's accuracy: SVM – 68%, ML – 35%, RF – 41%).

5. Discussion

The analysis of the achieved results implies that EnMAP and Sentinel-2 data can be applicable to the classification of plants based on adequate classification algorithms, including SVM and ML (for spectrally reduced multispectral and hyperspectral data). The results of the classification with these methods should be considered good (overall accuracy for most classifications >70%) and may be used to draw reliable alpine vegetation maps. It is confirmed with other studies which used comparable data and classification algorithms and produced similar results to those achieved by the authors. The team of A. Marcinkowska-Ochtyra (2017) classified the alpine vegetation of the Karkonosze mountains based on EnMAP data, using the SVM method, and achieved the classification OA at the level of 78%, whereas the team of L. Kupková (2017), working on Sentinel-2 data, achieved 78% for the ML algorithm and 71% for the SVM one.

The SVM mapper can be effectively applied to the classification of pixels that are not spectrally clean (E. Raczko, B. Zagajewski 2017), which is relatively common when classifying

alpine vegetation and may eventually lead to obtaining very good overall accuracy results of the classification made with this method. The large size of a pixel of the Sentinel-2 and EnMAP sensor should also be taken into account as it may contribute to the appearance of numerous mixels.

The ML classification results may be considered good. Comparable results with the application of the same method were also obtained by other researchers (R. Suchá et al. 2016 – OA 78%, M. Gartzia et al. 2013 – Kappa index 0.9). The ML classification on hyperspectral data usually requires that its spectral space is reduced (the number of traverses should be greater by one than the number of bands), which may sometimes produce negative results (ML classification OA based on five PCA principal components for EnMAP data – 63%).

Nevertheless, the results of all classifications made with the RF method should be assessed negatively. Despite their high overall accuracy for Sentinel-2 data, they were incorrect.

Another important aspect which is worth noting is a possibility of reducing spectral space in order to accelerate the classification process and limit the volume of data. Such reduction allows for the classification of EnMAP hyperspectral data with the ML algorithm. The effects of the PCA transformation in the case of the Sentinel-2 data classification were satisfactory – there was increase in overall accuracy for the SVM method by 4%, whereas, in the case of other classifications there was slight fall by approximately 2%. For EnMAP data in the SVM and ML classifications, the fall in overall accuracy together with the reduction of spectral space after PCA transformation was clearly visible (for the SVM classification – 18%, ML – 9%), which could have been connected with the selection of too few principal components which were obtained as a result of PCA transformation.

As far as the classification results for specific vegetation types are concerned, it should be noted that spatial resolution of Sentinel-2 and EnMAP data is not sufficient to map divisions with relatively small area or large heterogeneity. This may be confirmed with very low both producer's and user's accuracy for the grasslands class which, in most conducted classifications, did not exceed 50% on Sentinel-2 data and 40% on EnMAP data. Only the results of the SVM classification on complete EnMAP data (190 bands) were satisfactory in terms of producer's and user's accuracy for grasslands (71% and 68%, respectively). This type of vegetation is often distributed mosaically or occurs in patches. The major grasslands species – *Calamagrostis villosa* – can also be found in heathlands. Other researchers have also encountered problems with proper mapping of this type of vegetation species (L. Kupková et al. 2017, R. Suchá et al. 2016). Better results for this class were achieved by A. Marcinkowska-Ochtyra and others (2017) – producer's accuracy for grasslands reached the level of 79%, yet the study resulted in low accuracy for heathlands (PA – 44%).

Moreover, it appears problematic to identify representative pixels for some of the classes if a pixel is relatively large, which makes it necessary to generalise classifications to just a few divisions. It was highlighted in the research by L. Kupková and others (2017), A. Marcinkowska-Ochtyra and others (2017), and S. Suchá and others (2016). Those teams reduced the number of divisions in comparison to similar classifications based on higher spatial resolution data. Too low spatial resolution of Sentinel-2 and EnMAP data also poses a problem with respect to the mapping of other vegetation types based on Sentinel-2 data, such as aquatic (D. Stratoulias et al. 2015), forest and land cover vegetation (M. Immitzer et al. 2016), as well as shrub vegetation – based on the EnMAP data (S. Suess et al. 2015).

6. Summary and conclusions

The conducted analyses show the applicability of Sentinel-2 and EnMAP data to preliminary identification of alpine vegetation based on the classification that may be further used as framework for drawing alpine vegetation maps. The achieved results confirm the reliability of classifications based on the SVM and ML methods. This particularly applies to overall accuracy which, in most of the cases, exceeded 70%. This is actually very good news for researchers, especially that both Sentinel-2 and EnMAP data is available free of charge.

In order to achieve positive results of alpine vegetation classification based on Sentinel-2 and EnMAP data, divisions must be subject to generalisation as the spatial resolution of sensors (Sentinel-2–10 m; EnMAP – 30 m) is in-

sufficient to do the effective mapping of small heterogeneous classes. This is confirmed, among other things, by the results of grass-lands and heathlands classification for which the smallest number of pixels in both training and verification traverses was specified. These vegetation classes were the most problematic in terms of classification. Producer's and user's accuracy for these classes was below 50%.

The choice of an adequate classification algorithm seems essential in order to conduct correct classification. The difference between the highest and lowest overall accuracy value for EnMAP data was 25%, whereas it was lower for Sentinel-2 data. This study has revealed that the support vectors method – an algorithm belonging to the group of Machine Learning algorithms – is the most useful for the classifi-cation of alpine vegetation. A different algorithm from this group, namely the Random Forest method has turned out to be inapplicable. Re-latively good results have been achieved for the data classified with the Maximum Likelihood

method. It should be noted, however, that the classification based on the ML method requires the reduction of the spectral space for hyper-spectral data, which means that more work and time may be needed.

Future research should focus on the possibi-lities of integrating EnMAP data with data of higher spatial resolution, such as Sentinel-2, since 242 bands of the German sensor are capable of distinguishing the nuances in the spectral reflectance curve, which allows for precise mapping of vegetation. Such research was successfully conducted in order to identify minerals (N. Yokoya et al. 2016).

Future tests should also relate to the question of reducing spectral space of EnMAP data in order to streamline the research process with the smallest possible loss in the information capacity. The effectiveness of different methods of reducing spectral space, such as Minimum Noise Fraction (MNF), and different combina-tions of principal components that are eventually identified should also be studied closely.

Literature

Ashutosh S., 2012, *Monitoring forests: A new para-digm of remote sensing & GIS based change de-tecti*on. "Journal of Geographic Information Systems" Vol. 4, pp. 470–478.

Bannari A., Morin D., Bonn F., Huete A.R., 1995, *A re-view of vegetation indices*. "Remote Sensing Re-view" Vol. 13, no. 1–2, pp. 95–120.

Beger M., Moreno J., Johannessen J., Levelt P., Hans-sen R., 2012, *ESA's Sentinel missions in support of earth system science*. "Remote Sensing of En-vironment" Vol. 120, pp. 84–90.

Billingsley F.C., 1984, *Remote sensing for monitoring vegetation: an emphasis on satellites.* In: *The Role of Terrestrial Vegetation in the Global Carbon Cycle*. Edited by G.M. Woodwell. New York: John Wiley and Sons, pp. 161–180.

Bösche N.K., Rogaß C., Mielke C., Kaufmann H., 2014, *Hyperspectral digital image analysis and geochemi-cal analysis of a rare earth elements mineralized intrusive complex (Fen carbonatite Complex in Tele-mark Region, Norway.* In: *Proceedings of 34th EARSeL Symposium*, pp. 4.1–4.6, DOI: 10.12760/03-2014-07.

Braun A., Weinmann M., Keller S., Muller R., Reinartz P., Hinz S., 2015, *EnMAP contest: developing and comparing classification approaches for the envi-ronmental mapping and analysis programme – data-set and first results*. "Remote Sensing and Spatial Information Sciences" Vol. XL-3/W3, pp. 169–175.

Braun-Blanquet J., Chou Y.T., 1947*, Carte des groupe-ments végétaux de la France, region nordouest de*

Montpellier. Station internationale de geobota-nique mediterraneenne et alpine, Montpellier.

Buddenbaum H., Rock G., Hill J., Werner W., 2015, *Measuring stress reactions of beech seedlings with PRI, fluorescence, temperatures and emissivity from VNIR and thermal field imaging spectroscopy*. "European Journal of Remote Sensing" Vol. 48, pp. 263–282.

Buddenbaum H., Stern O., Paschmionka B., Hass E., Gattung T., Stoffels J., Hill J., Werner W., 2015, *Using VNIR and SWIR field imaging spectroscopy for drought stress monitoring of beech seedlings*. "International Journal of Remote Sensing" Vol. 36, pp. 4590–4605.

Burai P., Deak B., Valko O., Tomor T., 2016, *Classifi-cation of Herbaceous vegetation using hyper-spectral imagery*. "Remote Sensing" Vol. 7, no. 2, pp. 2046–2066.

Campbell J., Wynne R., 2011, *Introduction to remote sensing*. New York: The Guilford Presss, pp. 317–360.

Congalton R.G., 1991, *A review of assessing the ac-curacy of classifications of remotely sensed data*. "Remote Sensing of Environment" Vol. 37, pp. 35–46.

Delegido J., Verrelst J., Alonso L., Moreno J., 2011, *Evaluation of Sentinel-2 red-edge bands for empi-rical estimation of green LAI and chlorophyll con-tent*. "Sensor" Vol. 11, no. 7, pp. 7063–7081.

Dirnböck T., Dullinger S., Gottfried M., Ginzler C., Grabherr G., 2003, *Mapping alpine vegetation based on image analysis, topographic variables and Cano-*

nical Correspondence Analysis. "Applied Vegetation Science" Vol. 6, no. 1, pp. 85– 96.

Dotzler S., Hill J., Buddenbaum H., Stoffels J., 2015, *The potential of EnMAP and Sentinel-2 data for detecting drought stress phenomena in deciduous forest communities.* "Remote Sensing" Vol. 7, no. 10, pp. 14227–14258.

Dragozi E., Gitas I.Z., Stavrakoudis D.G., Theocharis J.B., 2014, *Burned area mapping using Support Vector Machines and the FuzCoC feature selection method on VHR IKONOS imagery.* "Remote Sensing" Vol. 6, no. 12, pp. 12005–12036.

Feng Q., Gong J., Liu J., Li Y., 2015, *Flood mapping based on multiple endmember spectral mixture analysis and Random Forest classifier – the case of Yuyao*, *China.* "Remote Sensing" Vol. 7, pp. 12539−12562.

Gartizia M., Alados C., Perez-Cabello F., Bueno C., 2013, *Improving the accuracy of vegetation classifications in mountainous areas. A case study in Spanish Pyrenees.* "Mountain Research and Development" Vol. 33, no. 1, pp. 63–74.

Humbolt, A. von, Bonpland A., 1895, *Geographie des plantes equinoxiales: tableau physique des Andes et pays voisins.* In*: Essai sur la géographie des plantes*, Paris: Levrault, Schoell et Co.

Immitzer M., Vuolo F., Atzberger C., 2016*, First experience with Sentinel-2 Data for crop and tree species classifications in central Europe.* "Remote Sensing" Vol. 8, no. 3, pp. 166–193.

Jarocińska A., Zagajewski B., 2008, *Korelacje naziemnych i lotniczych teledetekcyjnych wskaźników roślinności dla zlewni Bystrzanki.* „Teledetekcja Środowiska" T. 40, pp. 100–125.

Jarocińska A., Zagajewski B., 2009, *Remote sensing tools for analyzing state and condition of vegetation.* "Annals of Geomatics", Polish Association for Spatial Information, Vol. 7, no. 2, pp. 47–54.

Jarocińska A., Kacprzyk M., Marcinkowska-Ochtyra A., Ochtyra A., Zagajewski B., Meuleman K., 2016, *The application of APEX images in the assessment of the state of non-forest vegetation in the Karkonosze Mountains.* "Miscellanea Geographica – Regional Studies on Development" Vol. 20, no. 1, pp. 21–27.

Jensen J.R., 1983, *Biophysical remote sensing – review article.* "Annals of the Association of American Geographers" Vol. 73, no. 1, pp. 111−132.

Kaufmann H., Forster S., Wulf H., Segl K., Guanter L., Bochow M., Heiden U., Muller A., Heldens W., Scheneidehan T., Leitão P.J., van der Linden S., Hostert P., Hill J., Buddenbaum H., Mauser W., Hank T., Krasemann H., Rottgers R., Oppelt N., Heim B., 2012*, EnMAP Technical Report, GFZ Data Services.* Potsdam, pp. 1–44.

Khorram S., Nelson S., Koch F., van der Wiele C., 2012, *Remote sensing.* New York: Springer US, pp. 1−37.

Küchler A., Zonneveld I., 1988, *Vegetation mapping.* Berlin: Springer.

Kupková L., Cervená L., Suhá R., Jakesová L., Zagajewski B., Brezina S., Alberchtova J., 2017, *Classification of tundra in the Karkonose Mountains National Park, using APEX, AISA Dual and Sentinel-2A Data.* "European Journal of Remote Sensing" Vol. 50, no. 1, pp. 29–46.

Kycko M., Zagajewski B., Zwijacz-Kozica M., Cierniewski J., Romanowska E., Orłowska K., Ochtyra A., Jarocińska A., 2017, *Assessment of hyperspectral remote sensing for analyzing the impact of human trampling on Alpine wards.* "Mountain Research and Development" Vol. 37, no. 1, pp. 66–74.

Leitão P., Schwieder M., Suess S., Okujeni A., Galvão L., Linden S., Hostert P., 2015, *Monitoring natural ecosystem and ecological gradients: perspectives with EnMAP.* "Remote Sensing" Vol. 7, no. 10, pp. 13098–13119.

Locherer M., Hank T., Danner M., Mauser W., 2015, *Retrieval of seasonal leaf area index from simulated EnMAP data through optimized LUT-Based inversion of the PROSAIL model.* "Remote Sensing" Vol. 7, no. 8, pp. 10321–10346.

Marcinkowska A., Zagajewski B., Ochtyra A., Jarocińska A., Raczko E., Kupková L., Stych P., Meuleman K., 2014, *Mapping vegetation communities of the Karkonosze National Park using APEX hyperspectral data and Support Vector Machines.* "Miscellanea Geographica" Vol. 18, no. 2, pp. 23–29.

Marcinkowska-Ochtyra A., Zagajewski B., Ochtyra A., Jarocińska A., Wojtuń B., Rogass C., Mielke C., Lavender S., 2017, *Subalpine and alpine vegetation classification based on hyperspectral APEX and simulated EnMAP images.* "International Journal of Remote Sensing" Vol. 38, no. 7, pp. 1839–1864.

Martius C.F.P., 1858, *Flora brasiliensis.* Leipzig: Oldenburg Verlag.

Mielke C., Muedi T., Papenfuß A., Bösche N., Rogaß C., Gauert C., Altenberger U., de Wit M., 2016, *Multi- and hyperspectral spaceborne remote sensing of the Aggeneys base metal sulphide mineral deposit sites in the Lower Orange River region, South Africa.* "South African Journal of Geology" Vol. 119, no. 1, pp. 63–76.

Nink S., Hill J., Buddenbaum H., Stoffels J., Sachtleber T., Langshausen J., 2015, *Assessing the suitability of future multi- and hyperspectral satellite systems for mapping the spatial distribution of Norway spruce timber volume.* "Remote Sensing" Vol. 7, pp. 12009–12040.

Ochtyra A., Zagajewski B., Kozłowska A., Marcinkowska-Ochtyra A., Jarocińska A., 2016, *Ocena kondycji drzewostanów Tatrzańskiego Parku Narodowego za pomocą metody drzewa decyzyjnego oraz wielospektralnych obrazów satelitarnych Landsat 5 TM.* „Sylwan" T. 160, nr 1, pp. 256–264.

Pedrotti F., 1967, *Carta fitosociologica della vegetazione de Montelago*. Camerino: Instituto di Botanica, Universita di Camerino.

Pesaresi M., Corbane C., Julea A., Florczyk A., Syrris V., Soille P., 2016, *Assessment of the added-value of Sentinel-2 for detecting built-up areas*. "Remote Sensing" Vol. 8, pp. 299–316.

Quattrochi D.A., Luvall J.C., 1999, *Thermal infrared remote sensing for analysis of landscape ecological processes: methods and applications*. "Landscape Ecology" Vol. 14, no. 6, pp. 577–598.

Raczko E., Zagajewski B., 2017, *Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images*. "European Journal of Remote Sensing" Vol. 50, no. 1, pp. 144–154.

Schmid E., 1940, *Die Vegetationskartierung der Schweiz im Masstab 1:200,000*. Geobotanisches Forschungsinstitut Rübel in Zürich, Bericht für das Jahr 1939, pp. 76−85.

Schouw J.F., 1823, *Grundzige einer allgemeinen Pflanzengeographie (mit Atlas)*. Berlin.

Sendtner O., 1854, *Die Vegetationsverhältnisse Südbayerns nach den Grundsätzen der Pflanzengeographie und mit Bezugnahme auf die Landescultur geschildert*. München.

Shweider M., Leitão P., Suess S., Senf C., Hostert P., 2014, *Estimating fractional shrub cover using simulated EnMAP Data: a comparision of three machine learning tehniques.* "Remote Sensing" Vol. 6, no. 4, pp. 3427–3445.

Siegmann B., Jarmer T., Beyer F., Ehlers M., 2015, *The potential of pan-sharpened EnMAP data for the assessment of wheat LAI*. "Remote Sensing" Vol. 7, no. 10, pp. 12737–12762.

Stoffels, J., Sachtleber, T., Mader, S., Buddenbaum, H., Stern, O., Langshausen, J., Dietz, J., 2015, *Satellite-based derivationof high-resolution forest information layers for operational forest management*. "Forests" Vol. 6, pp. 1982–2013.

Stratoulias D., Balzter H., Zlinszky A., Tóth V.R., 2015, *Assessment of ecophysiology of lake shore reed vegetation based on chlorophyll fluorescence, field spectroscopy and hyperspectral airborne imagery.* "Remote Sensing of Environment" Vol. 157, pp. 72–84.

Suchá R., Jakešová L., Kupková L., Červená L., 2016, *Classification of vegetation above the tree line in the Krkonoše Mts. National Park using remote sensing multispectral data*. "AUC Geographica" Vol. 51, no. 1, pp. 113–129.

Suess S., van der Linden S., Okujeni A., Leitão P., Shweider M., Hostert P., 2015, *Using class probabilities to map gradual transitions in shrub vegetation from simulated EnMAP data.* "Remote Sensing" Vol. 7, no. 8, pp. 10668–10688.

Thales Alenia Space, 2016, *Sentinel-2 Products specification document (PSD)*. European Space Agency (ESA), https://sentinel.esa.int/documents/ 247904/685211/Sentinel-2-Product-Specifications-Document, pp. 41–53 (access 6.01.2017).

Tobler M., Cochard R., Edwards P., 2003, *The impact of cattle ranching on large-scale vegetation patterns in a coastal savanna in Tanzania.* "Journal of Applied Ecology" Vol. 40, no. 3, pp. 430–444.

Tomczak J., 2013, *Wprowadzenie do sztucznej inteligencji*, https://www.ii.pwr.edu.pl/~tomczak/PDF/ si1.pdf (access 01.09.2017).

Topaloğlu R., Sertel E., Musaoglu N., 2016, *Assessment of classification accuracies of Sentinel-2 and Landsat-8 data for land cover/use mapping*. "The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences" Vol. XLI-B8, pp. 1055–1059.

Traganos D., Reinartz P., 2017, *Mapping Mediterranean seagrasses with Sentinel-2 imagery*. "Marine Pollution Bulletin" (article in print).

Vapnik, V.N., 1995, *The nature of statistical learning theory*. New York: Springer.

Wijaya A., Gloaguen R., 2007, *Comparison of multisource data support vector machine classification for mapping of forest cover*. In: *Geoscience and Remote Sensing Symposium 2007. IGARSS 2007*. IEEE International, pp. 1275–1278.

Wojtuń B., Żołnierz L., 2002, *Plan ochrony ekosystemów nieleśnych – inwentaryzacja zbiorowisk*. W: *Plan Ochrony Karkonoskiego Parku Narodowego*. Brzeg: Biuro Urządzania Lasu i Geodezji Leśnej, Oddział w Brzegu, pp. 67 and 2 maps.

Xie Y., Sha Z., Yu M., 2008, *Remote sensing imaginery in vegetation mapping: a review*. "Journal of Plant Ecology" Vol. 1, no. 1, pp. 9–23.

Yokoya N., Cheung-Wai Chan J., Segl K., 2016, *Potential of resolution-enhanced hyperspectral data for mineral mapping using simulated EnMAP and Sentinel-2 images.* "Remtote Sensing" Vol. 8, no. 3, pp. 172–190.

Zagajewski B., 2010, *Ocena przydatności sieci neuronowych i danych hiperspektralnych do klasyfikacji roślinności Tatr Wysokich*. „Teledetekcja Środowiska" T. 43, 113 pp.

Zagajewski B., Folbrier A., Kozłowska A., Sobczak M., Wrzesień M., 2005, *Zintegrowane pomiary roślinności wysokogórskiej*. „Teledetekcja Środowiska" T. 36, pp. 61–68.

http://www.enmap.org/?q=box_applications (access 1.09.2017)

http://land.copernicus.eu/global/products (access 1.09.2017)

https://earth.esa.int/web/guest/missions/esa-future--missions/flex (access 1.09.2017)