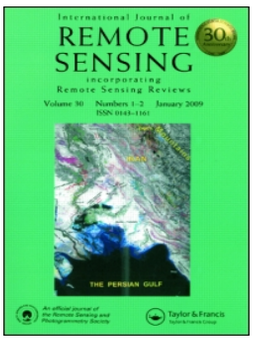
Marcinkowska-Ochtyra, A., Zagajewski, B., Ochtyra, A., Jarocińska, A., Wojtuń, B., Rogaß, C., Mielke, C., Lavender, S. (2017): Subalpine and alpine vegetation classification based on hyperspectral APEX and simulated EnMAP images. - *International Journal of Remote Sensing*, *38*, 7, pp. 1839—1864.

DOI: http://doi.org/10.1080/01431161.2016.1274447



International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: http://www.tandfonline.com/loi/tres20

Subalpine and alpine vegetation classification based on hyperspectral APEX and simulated EnMAP images

Adriana Marcinkowska-Ochtyra, Bogdan Zagajewski, Adrian Ochtyra, Anna Jarocińska, Bronisław Wojtuń, Christian Rogass, Christian Mielke & Samantha Lavender

To cite this article: Adriana Marcinkowska-Ochtyra, Bogdan Zagajewski, Adrian Ochtyra, Anna Jarocińska, Bronisław Wojtuń, Christian Rogass, Christian Mielke & Samantha Lavender (2017) Subalpine and alpine vegetation classification based on hyperspectral APEX and simulated EnMAP images, International Journal of Remote Sensing, 38:7, 1839-1864, DOI: 10.1080/01431161.2016.1274447

To link to this article: http://dx.doi.org/10.1080/01431161.2016.1274447

Published online: 13 Jan 2017.

INTERNATIONAL JOURNAL OF REMOTE SENSING, 2017
VOL. 38, NO. 7, 1839–1864
http://dx.doi.org/10.1080/01431161.2016.1274447

Taylor & Francis
Taylor & Francis Group

Subalpine and alpine vegetation classification based on hyperspectral APEX and simulated EnMAP images

Adriana Marcinkowska-Ochtyra[a], Bogdan Zagajewski[a], Adrian Ochtyra[a,b], Anna Jarocińska[a], Bronisław Wojtuń[c], Christian Rogass[d], Christian Mielke[d] and Samantha Lavender[e]

[a]Department of Geoinformatics, Cartography and Remote Sensing Warsaw, Faculty of Geography and Regional Studies, University of Warsaw, Warsaw, Poland; [b]College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Warsaw, Poland; [c]Department of Ecology, Biogeochemistry & Environmental Protection, Faculty of Biological Sciences, Wrocław University, Wrocław, Poland; [d]Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Telegrafenberg, Potsdam, Germany; [e]Pixalytics Ltd, Plymouth, UK

ABSTRACT

The characterization of vegetation is a very important ecological task, especially in sensitive mountain areas, as alpine regions often respond to small short-term variations of abiotic and biotic components as well as long-term global changes. Spatial techniques, such as imaging spectroscopy, allow for detailed classification of different syntaxonomic categories of vegetation and their status. Based on the Airborne Prism Experiment (APEX) and simulated Environmental Mapping and Analysis Program (EnMAP) data, this study focused on subalpine and alpine vegetation mapping in the eastern part of the Polish Karkonosze National Park (KPN). The spatial resolution of APEX (3.12 m) enabled the classification of 21 vegetation communities, which was generalized into eight vegetation types. These types were identified on scaled-up APEX data, as both 252 bands from most of the spectral range and a spectrally reduced dataset of 30 minimum noise fraction (MNF) transforms, and compared to the simulated (30 m spatial resolution) EnMAP data using test areas extracted from the field survey derived reference non-forest vegetation map. After preprocessing, a pixel purity index (PPI) was calculated using the MNF image and then the training and validation pixels were selected with Support Vector Machine classification of vegetation communities carried out using different kernel functions: linear, polynomial, radial basis function, and sigmoid. The classification accuracy was obtained for 21 base classes, and the best result was achieved by using the linear function and 252 bands (overall accuracy (OA) of 74.39%). The next step was to classify the eight generalized vegetation types, and the OA for the APEX data reached 90.72% while EnMAP reached 78.25%. The results show the potential use of APEX and EnMAP imagery in mapping subalpine and alpine vegetation on a community and vegetation-type scales, within a diverse ecosystem such as the Karkonosze National Park.

ARTICLE HISTORY
Received 15 December 2015
Accepted 5 December 2016

CONTACT Adriana Marcinkowska-Ochtyra adriana.marcinkowska@uw.edu.pl Department of Geoinformatics, Cartography and Remote Sensing Warsaw, Faculty of Geography and Regional Studies, University of Warsaw, 30 Krakowskie Przedmieście Street, Warsaw 00-927, Poland

1. Introduction

Hyperspectral scanners acquire data in a large number of narrow, contiguous, spectral bands at high spatial and spectral resolutions (Goetz et al. 1985), with the discrimination of vegetation communities and the extraction of more precise quantitative information about the terrestrial environment (Jarocińska 2014; Kycko, Zagajewski, and Kozłowska 2014; Waske et al. 2010; Camps-Valls et al. 2004) becoming possible. Mountainous plants adapt themselves to survive difficult conditions through quantitative changes in the relationships between the chlorophyll and carotenoid pigments, and plant tissue structure, e.g. cuticle layer growth, carbon, nutrient, and lignin content. These adaptations depend on species' properties and have a direct impact on the spectral reflectance, which can be quantified by imaging spectroscopy.

One of the first applications of hyperspectral imaging for mountain ecosystems classified 23 forest and 15 non-forest communities in the Yellowstone National Park using Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data. The final result, which used the spectral feature analysis method, identified 16 communities with an overall accuracy (OA) of 74.1% (kappa coefficient of 0.62) (Kokaly et al. 2003). AVIRIS data were also used by Filippi and Jensen (2006), who analysed 22 vegetation classes in Florida (USA) using artificial neural networks. The final map presented 11 classes with an OA ranging between 84% and 86%.

A classification of alpine and subalpine zones within the Polish Tatra National Park (UNESCO's Man and the Biosphere Reserve (M&B)) based on Digital Airborne Imaging Spectrometer (DAIS 7915) images with the Spectral Angle Mapper (SAM), fuzzy neural nets approach and detailed field maps allowed for the identification of 39 dominant vegetation communities (Zagajewski 2010). All 79 bands, covering the electromagnetic spectrum from visible to thermal infrared wavelengths, were scrutinized to identify the most significant spectral bands that allowed for the selection of 40 significant bands; the Adaptive Resonance Theory Map (ARTMAP) simulator was used for plant community classification by applying a fuzzy classification algorithm (fuzzy ARTMAP). Generally, the 40 bands of input data offered 88.6% accuracy and the kappa coefficient was 0.87.

Chan and Paelinckx (2008) classified 16 ecotopes (grasslands, arable areas, and forests) based on the Hyperspectral Mapper (HyMap) data and two decision tree classifiers: Random Forest (RF) and Adaptive Boosting (AdaBoost). OA for classification was 69.5%, but individual classes had much higher accuracies. A problem of poor separation was seen for some classes, and the authors concluded that regrouping or class merging improved the applicability of their classification method. In addition, vegetation mapping at different scales was conducted by Zhang and Xie (2013) for communities and species in the Kissimee River watershed using HyMap data with Support Vector Machine (SVM) and RF combined with object-based image analysis techniques. For 14 communities, the original and the minimum noise fraction (MNF) datasets reached an OA of 74% and 90%, respectively, and for 55 species it was 73% and 85%, respectively. The conclusion was that the more detailed classification, at the species level, was more accurate for the SVM method with the communities-level classification being comparable.

According to Camps-Valls et al. (2004), the SVM method is more efficient than typical neural networks in terms of accuracy, robustness, and simplicity, as well as the speed of

the classification process. They used SVM and the multilayer perceptron for six crop classes with 128 bands of HyMap data that reached a user accuracy of more than 90% and producer accuracy of more than 83% per class. Even noisy bands are successfully detected, and provide a useful result when classifying high-dimensional datasets (Tarabalka et al. 2010).

Based on AVIRIS data and three classification methods, including SVM, vegetation was classified for northern US Indiana; the classes covered corn, soybean, and grass trees. The best results were obtained by SVM for a four class subset scene that achieved an accuracy of 95.9%, with 16 classes covering the full scene achieving an accuracy of 87.3% (Gualtieri and Cromp 1988). Hyperspectral DAIS 7915 data and SVM were used to map nine land-cover types within the La Mancha region (Spain), where SVM achieved the best accuracy – 97.5% (Pal and Mather 2004). The different kernel used by the SVM method, and various multiclass strategies, were tested by Melgani and Bruzzone (2004) with AVIRIS data and nine land-cover classes. The analysis was compared to different classification methods (the radial basis function (RBF) neural networks and *k*-nearest neighbours (*k*-NN) classifier) and, like all of the above studies, showed that SVM was the best algorithm in terms of accuracy, computational time and stability to parameter settings.

The Airborne Prism Experiment (APEX) has been available since 2010 (Demarchi et al. 2014) with articles in the literature primarily oriented towards a general overview about the sensor (Itten et al. 2008), its radiometric calibration (Hueni et al. 2014a; Hueni, Schlaepfer, and Jehle 2014b), geometric (Vreys et al. 2016), and atmospheric correction (Schlaepfer et al. 2008; Sterckx et al. 2016). These data were successfully used to simulate images and products of upcoming space missions (D'Odorico et al. 2013). Thematic APEX applications were oriented towards the analysis of vegetation indices and chlorophyll fluorescence (Damm et al. 2015), burned areas within four vegetation types (Schepers et al. 2014) and also for classifications; detailed urban land cover mapping was developed using two-dimensional reduction techniques: auto-associative neural networks (AANN) and BandClust alongside the full set of bands for classification methods such as RF, AdaBoost, multiple layer perceptron, and SVM (Demarchi et al. 2014). The results showed that the SVM method and full band set combination gave the best OA (82.9%). The APEX data were successfully used for tree species mapping; in the work of Tagliabue et al. (2016) in Forêt de Hardt (France) it was possible to discriminate four species using maximum likelihood (ML) method achieving 74.4% of OA; Raczko et al. (2015) classified five tree species in Karkonosze Mountains (Poland) using SVM, which allowed for reach 78.7% of OA.

The Environmental Mapping and Analysis Program (EnMAP) system is not in orbit at the time of writing, but published articles already describe the characteristics of the mission and the anticipated applications (Guanter et al. 2015; Heldens et al. 2011). Applications have been developed and tested using simulated data derived from higher resolution airborne hyperspectral data using the EnMAP End to End Simulator (EEtES) (Segl et al. 2012). EnMAP simulated data based on the Advanced Imaging Spectrometer for Applications (AISA) images were used to map the gradual transition in shrub vegetation in a semi-arid, natural environment in Portugal using the Support Vector Classification (SVC) model (Suess et al. 2015). Simulated EnMAP data were also used to retrieve information on the seasonal development of vegetation parameters, such as leaf

area index (LAI), by generating look-up-table libraries to invert the PROSAIL model – the combination of two radiative transfer models: the leaf optical Properties Spectra (PROSPECT), and Scattering by Arbitrary Inclined Leaves (SAIL, Locherer et al. 2015).

Therefore, in summary, there have been numerous investigations of protected, semi-, and natural ecosystems that used hyperspectral data to determine the extent of different vegetation classes using various classification techniques, with a mixture of outcomes found. Karkonosze Mountains has a complicated vegetation mosaic with high community diversity, mostly with high spatial heterogeneity, so the use of machine learning method to classification was promising to achieve satisfactory results, rather than endmember-based approaches or traditional parametric classifiers (Zhang and Xie 2013). A big advantage of SVM is that it is able to produce higher accuracy than traditional classifiers using a small number of training datasets that characterize the communities of this area. In comparison to different methods described in the literature review, SVM gave the best results for different types of classes, and it was the reason for choosing this method to classify diverse mountain vegetation.

For this research, the SVM classification approach was chosen with different kernel functions tested for non-forest vegetation communities. The aims of this study were to:

- develop a processing method for APEX hyperspectral imagery to identify mountainous vegetation communities for the Karkonosze UNESCO's M&B Reserve using SVM,
- check the capabilities of vegetation-type classification when using data with EnMAP's spatial resolution,
- assess the usefulness of aerial APEX and simulated satellite EnMAP data to classify the vegetation types in the Karkonosze UNESCO's M&B Reserve.

The work was divided into two parts: first, the classification of vegetation communities based on an APEX image subset that's the dataset referred to as 'APEX_comm' (Figure 1). Second, the classified communities were generalized into vegetation types to provide a classification for the greater geographical area covered using the APEX dataset called 'APEX_types' and a new EnMAP dataset called 'EnMAP_types' (Figure 1).

The main research problems were expressed as an identification of each non-forest vegetation community and also a study of how the use of different kernel types and a reduction of spectral dimensionality impacts classification accuracy. In a subsequent step, there was upscaling of the units from communities to types alongside an upscaling of the data from an airborne to satellite spatial resolution. The eastern part of Karkonosze National Park was selected for these upscaling tests because of its high variability in terms of vegetation communities near the Mały Staw lake area.

2. Area and object of the study

2.1. *Study area*

Karkonosze belongs to the Sudety Mountains, a chain of middle mountains shared by the Czech Republic, Poland, and Germany. The study area is situated in the eastern part of Karkonosze in the border region between the Czech Republic and Poland. It was divided

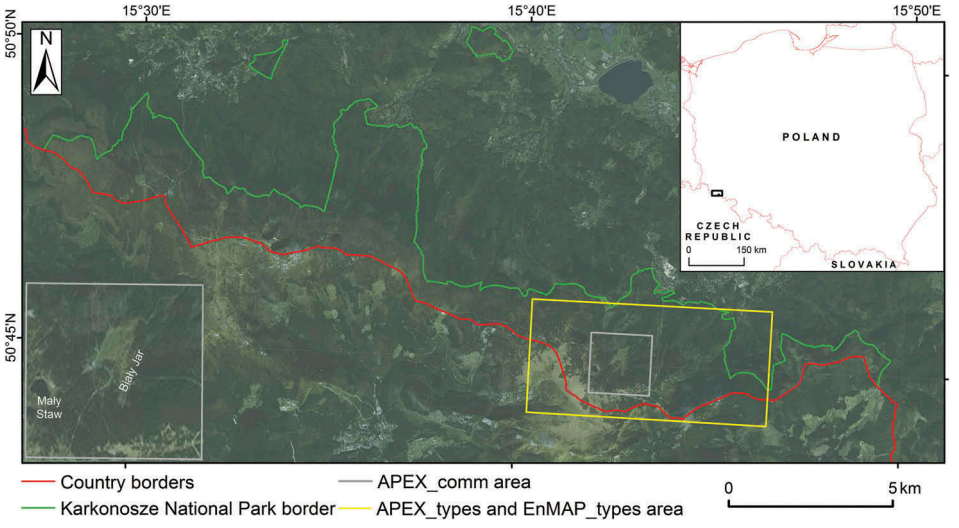


Figure 1. The study area in the eastern part of Karkonosze.

into a greater and smaller part, depending on research objectives, as described in Section 2.2 (Figure 1). The larger area covers 7.33 × 3.46 km (50°46′00″–50°44′09″ N 15°40′20″–15°46′33″ E), and it is located between 1270 and 1602 m above sea level (a.s.l.) while the smaller area covers 1.86 × 1.83 km (50°45′30″ –50°44′30″ N, 15°42′00″–15°43′30″ E), and it is located between about 1270 and 1430 m a.s.l. The bedrock of the eastern Karkonosze is mainly acidic and primarily comprised of granites, gneiss and crystalline schists (Kondracki 2009). The mean annual precipitation in the summit parts is about 1512 mm, with a mean annual temperature variation of between 0.7°C and 4.2°C and the first frosts begin in November (Sobik et al. 2013). Considerable varieties of climate conditions, along with the altitudinal gradient, result in distinct vegetation belts. The middle subalpine belt (1250–1450 m) is a mosaic of Sudetic dwarf pine shrubs, matt-grass meadows and subalpine mires covering mostly flat and gently sloping summit areas. The alpine belt extends to the highest peak of Sniezka, which reaches 1602 m a.s.l. located in greater part of the study area.

The eastern part of subalpine and alpine belts in Karkonosze is exceptional due to its geomorphology and vegetation pattern, with the most remarkable geomorphological landmarks being the summit planation surfaces, nivation hollows, and glacial corries. The arctic–alpine vegetation also constitutes one of the most significant centres of geo-biodiversity for this mountain (Śtursa 2013). Large plantations in the study area are covered by dwarf-pine stands (*Pinetum mughi sudeticum* association), mat-grass communities of the Nardion alliance (mostly *Carici (rigidae)-Nardetum* association), and peat bogs that are dominated by plant communities of the class *Oxycocco-Sphagnetea*. Glacial corries and nivation hollows are the most diverse environments in the Karkonosze, with a large number of plant communities and associated variety of habitats (rock faces, avalanche tracts, snowbeds, subalpine scrubs, subalpine tall forbs, and springs). The Mały Staw Cirque, as the biggest glacial cirque in the Karkonosze, together with Biały Jar nivation hollow (Figure 1), are the locations with the highest plant community diversity in this mountain range.

Table 1. Vegetation communities (21) and types (9) in the study area of the eastern part of Karkonosze National Park.

Vegetation community	Vegetation type
Carici (rigidae)-Nardetum	Alpine grasslands
Empetro-Vaccinietum	Heathlands
Pinetum mugo sudeticum	Subalpine dwarf pine scrubs
Crepido-Calamagrostietum villosae	Subalpine tall-forbs
Pado-Sorbetum	Deciduous shrubs vegetation
Salicetum lapponum	Deciduous shrubs vegetation
Athyrietum distentifolii	Herbs
Adenostyletum alliariae	Herbs
Scheuchzerio-Caricetea nigrae	Fens
Oxycocco-Sphagnetea	Bogs
Cardamino-Montion	Springs
Calamagrostio villosae-Piceetum	Forests
Rhizocarpion alpicolae	Rock and scree vegetation
Umbilicarion cylindricae	Rock and scree vegetation
Vaccinium myrtillus com.	Heathlands
Calluna vulgaris com.	Heathlands
Molinia caerulea com.	Subalpine tall-forbs
Deschampsia flexuosa com.	Alpine grasslands
Deschampsia caespitosa com.	Subalpine tall-forbs
Peucedanum ostruthium com.	Ruderal vegetation
Urtica dioica com.	Ruderal vegetation

2.2. *Subalpine and alpine vegetation*

The analysis included subalpine and alpine vegetation types with the communities located in the highest parts of the Polish side of the Eastern Karkonosze Mountains. The non-forest vegetation map from 2002 acted as a reference, containing 17 vegetation types and 49 vegetation communities identified across the whole of the Karkonosze National Park (Wojtuń and Żołnierz 2002). It was derived by manual interpretation of orthophotomaps and validated by field surveying. The selected area, near Mały Staw lake, contains the vegetation listed in Table 1.

All of the vegetation communities listed in Table 1 are important from a phytosociological point of view, and could be separately distinguished within the non-forest vegetation map (Wojtuń and Żołnierz 2002). However, it was not possible to identify all of them in the image data visually and based on the spectral characteristics alone due to their spectral similarity. For example, it was not possible to differentiate *Pino-mugo Sphagnetum* and *Pinetum mugo sudeticum* communities using APEX data due to a difference only in the soil properties. Therefore, they were merged into a community termed *Pinetum mugo sudeticum*. In addition, the associations of the *Molinio-Arrhenatheretea* class were removed from the final map because this community was dominated by *Deschampsia caespitosa* in the study area, and the *Calamagrostion* alliance was replaced by the communities contained within it – *Carici (rigidae)-Nardetum, Deschampsia flexuosa,* and *Deschampsia caespitosa*. The associations from the *Artemisietea vulgaris* class were merged with *Urtica dioica* comm. because they covered relatively small areas separately and both classes are mostly covered by nettles. Finally, two classifications were used (Table 1): vegetation community and vegetation type. There were 21 vegetation community units, alongside the early stages of vegetation succession (the area located near Biały Jar nivation hollow), and areas without vegetation. For vegetation type, the 21 community units were merged into larger units

according to the non-forest vegetation map (Wojtuń and Żołnierz 2002). Two vegetation type classes: including *Urtica dioica* community – ruderal vegetation, and including *Athyrietum distentifolii* and *Adenostyletum alliariae* – herbs were removed from the final version because classification using 30 m EnMAP data resulted in only 5 pixels for ruderal vegetation and 7 pixels for herbs, so it was considered too small a sample. This resulted in 8 vegetation types alongside areas without vegetation.

3. Data

3.1. *Image data*

The data used in this study were the APEX and simulated EnMAP images. APEX is the Swiss-Belgian dispersive push broom imaging spectroradiometer, operating between 380 and 2500 nm in 288 bands (bandwidths of about 5 nm in the visible and near-infrared (VNIR) and 10 nm in the shortwave infrared (SWIR)) with a spatial ground resolution ranging from 1.75 to 3 m depending on the flying altitude (Itten et al. 2008; Vreys, Iordache, and Meuleman 2014). The spectroradiometer was installed on board a Dornier 228 aircraft operated by DLR (the German Aerospace Center), and the flight campaign covered the whole area of the Karkonosze Mountains within both the National Park borders and their surroundings (Jelének et al. 2014). Data were acquired on 10 September 2012, in the framework of the European Facility for Airborne Research (EUFAR) project entitled Hyperspectral Remote Sensing for Mountains Ecosystems (HyMountEcos) that involved scientists from the Polish University of Warsaw and the Czech Charles University of Prague. The September date of image acquisition allowed the greatest colouration differences for the species within the communities to be captured (Żołnierz and Wojtuń 2013).

The APEX data were pre-processed at the Flemish Institute for Technological Research VITO's Centre for Remote Sensing and Earth Observation Processes. First, it was radiometrically calibrated into radiance cubes, and then it was geometrically corrected using the light detection and ranging (LiDAR) derived digital terrain model (DTM) tiles, followed by an atmospheric correction using the MODTRAN® 4 radiative transfer code and reflectance retrieval algorithms from de Haan et al. (1991). Reference spectra of homogenous areas (e.g. concrete and asphalt) collected using the ASD FieldSpec® 3 spectroradiometer, during the flight campaign, were used to validate the atmospheric correction. As a result, the image data were provided as geo-referenced at-surface-reflectance having a ground sampling distance of 3.12 m.

The area near Mały Staw lake, and its surroundings, was chosen for the purposes of vegetation community classification (APEX_comm, Table 2). A greater area was selected for vegetation-type classification, because this area included joined vegetation communities (APEX_types, Table 2). For the APEX_comm and APEX_types datasets, bands linked to atmospheric water vapour absorption (1335–1422 nm and 1759–1955 nm) were removed with 252 bands retained.

EnMAP is a future German hyperspectral satellite mission that enables a characterization of the Earth's surface, due to be launched in 2018. This push broom prism instrument will cover a 30 km swath width with a spectral range from 420 to 2450 nm using two spectrometers: VNIR (420–1000 nm, with 89 bands) and SWIR (900–2450 nm, with 155 bands) (Guanter et al. 2015).

Table 2. Specifications on spatial and spectral properties of three datasets.

	APEX_comm	APEX_types	EnMAP_types
Spectral bands	252[a]/30[b]	252[a]	190[a]
Spectral range (nm)	413–2448	413–2448	423–2439
Spatial resolution (m)	3.12	3.12	30
Area size (pixels)	656 × 646	2767 × 1542	261 × 146

[a]Removed water vapour absorption bands.
[b]MNF transforms.

In order to simulate an EnMAP image over the study area, the spectrally similar APEX characteristics were resampled to 30 m pixel size. This was performed by convolving the APEX image with the wavelength-dependent point spread functions (PSF) of EnMAP and extracting every *n*th data point of the convolution result, where *n* was the ratio between the average EnMAP and the average APEX ground sampling distance. As occurred for APEX, bands in the water vapour absorption region (1335–1476 nm and 1773–2007 nm) were removed and so, finally, 190 bands were used (EnMAP_types, Table 2).

3.2. *Reference data*

The information about existing vegetation communities and types was extracted from a 2002 non-forest vegetation map (described in Section 2.2) alongside visual interpretation of the high, 12 cm resolution, orthophotos and field documentation. The vector layer contained 49 vegetation communities for the whole of the Karkonosze National Park (Wojtuń and Żołnierz 2002), but was resized to the study area that contained 21 communities. The pixel purity index (PPI) (Theiler et al. 2000) was calculated from the MNF transformed APEX data without the noisy bands, in order to find the most spectrally pure pixels to support training sample selection. It required human interaction to manually select the final set of pure pixels (Chang and Plaza 2006).

A field survey was conducted from 20 to 30 August 2013 and from 30 to 31 August 2014. The validation sample locations were chosen in homogenous polygons and also with major variation within each vegetation community class. Around 14,000 samples were collected, and a stratified random sampling scheme was applied in order to avoid the effects of spatial autocorrelation, which is dependent on the size of each vegetation community's area. Separate random samples were generated to perform the accuracy assessment on the pixel wise mapping unit, and as the classification had a large number of categories (defined as more than 12) more than 75–100 samples per category (Congalton 1991) was optimal. However, this was not realized in the case of five vegetation community classes: *Adenostyletum alliariae, Calluna vulgaris, Deschampsia flexuosa, Peucedanum ostruthium,* and *Urtica dioica* that occupy smaller areas, but for SVM smaller training samples are not obstacles. Finally, a total of 9562 pixels were used for training and 4608 pixels were used for validation of vegetation community classification, which were labelled according to the finest level of a non-forest vegetation map developed by Wojtuń and Żołnierz (2002).

For the vegetation-type classification, samples created in the previous analysis step were merged. The information about the non-forest vegetation types was extracted from the reference non-forest vegetation map database, and generalized from vegetation communities into vegetation types with class names assigned according to the 2002 non-forest vegetation

map. In total, 10,149 training and 5022 validation pixels were collected for the APEX_types dataset with 329 and 191 pixels, respectively, collected for the EnMAP_types data.

4. Methods

The stages used to prepare the image data and then undertake the classification are presented in Figure 2 and explained in Sections 4.1–4.3.

4.1. *Dimensionality reduction of APEX_comm dataset*

A commonly used method of dimensionality reduction in hyperspectral analysis is the MNF transformation, which is used to determine inherent dimensionality and segregate noise (Green et al. 1988; Boardman and Kruse 1994; Underwood, Ustin, and DiPietro

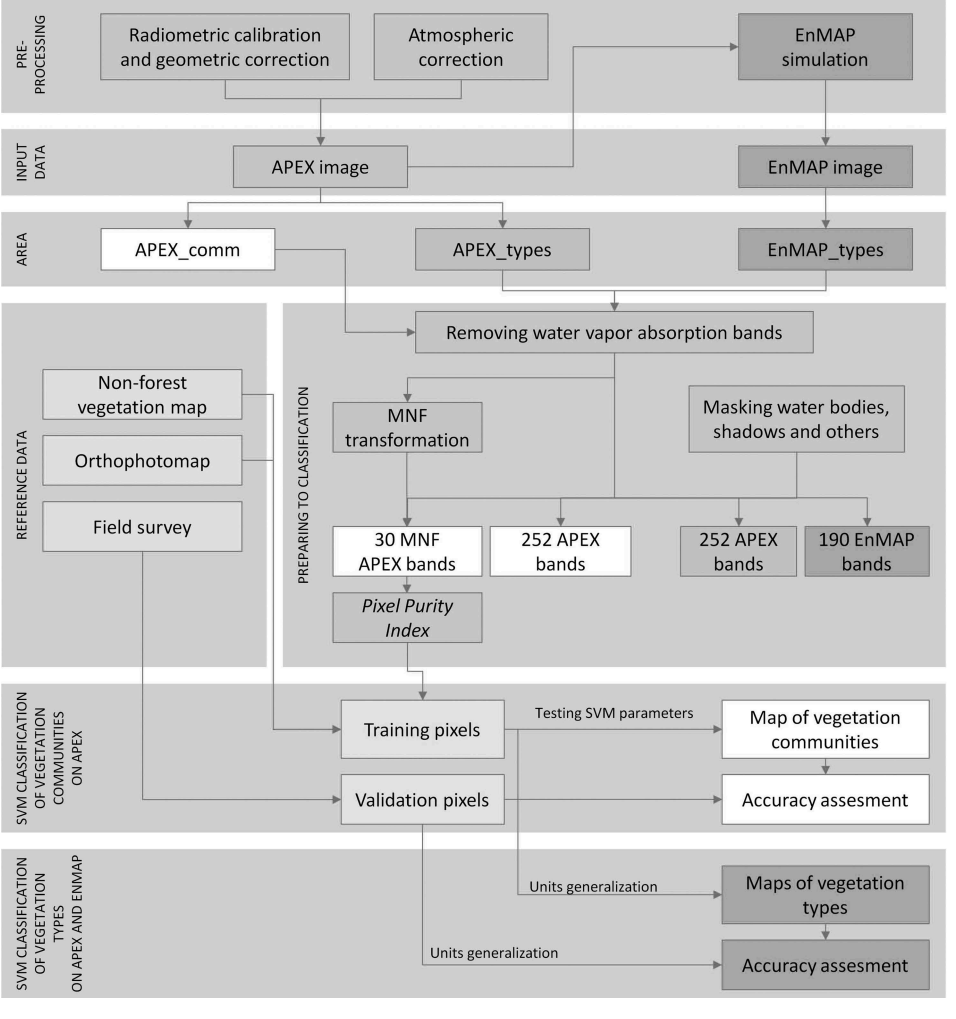


Figure 2. Classification procedures for the APEX and EnMAP data.

2003). The MNF tool in ENVI was used to apply the approach, and then analyse the output image using eigenvalues; from which the first 30 MNF transforms were used. This approach resulted in two datasets: 30 MNF transforms and original 252 band dataset.

4.2. *Masking areas*

In order to extract only subalpine and alpine vegetation, the forests and mid-forest clearings were masked, as well as the Czech side of National Park as this region was not studied, using a mask based upon pre-existing non-forest vegetation map that included the gaps of shelters. The next step was the suppression of shadows and waterbodies. Although the data were atmospherically corrected and the topographic correction was not performed, it was not possible to eliminate shadows in areas of large height difference. Therefore, shadows were masked using the reflectance in the 106th spectral band (952 nm) for APEX with a 6% threshold value and 94th band (95 nm) with a 7% threshold value for EnMAP. However, some areas of water were also detected as shadows and so had to be adjusted manually. Finally the forest, sheltered area and shadow masks were merged into a single mask that was applied to the images (Figures 3(*a*) and (*b*)).

4.3. *SVM classification*

The SVM approach is a supervised learning method of classification that employs optimization algorithms to locate optimal boundaries between classes to separate them with a hyperplane, which maximizes the margin between them in high dimensional space (Huang, Davis, and Townshend 2002). Data points closest to the hyperplane are support vectors, which are the most important elements of the training set (Burges 1988), with kernel functions converting nonlinear boundaries into linear ones in high dimensional space. To appoint a discriminate function for locating boundaries, SVM uses a different type of kernel functions. In the ENVI software, four types of kernel functions $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ are available: linear (1), polynomial (2), RBF (3), and sigmoid (4). The equations for each of them are as follows, with detailed mathematical formulations given in the works of Gualtieri and Cromp (1988), Huang, Davis, and Townshend (2002), Burges (1988), Vapnik (1995):

linear

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i{}^{\mathrm{T}}\boldsymbol{x}_j \tag{1}$$

polynomial

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \left(\gamma \boldsymbol{x}_i{}^{\mathrm{T}}\boldsymbol{x}_j + r\right)^d, \gamma < 0 \tag{2}$$

RBF

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\gamma \left|\left|\boldsymbol{x}_i - \boldsymbol{x}_j\right|\right|^2\right), \gamma > 0 \tag{3}$$

sigmoid

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \tanh\left(\gamma \boldsymbol{x}_i{}^{\mathrm{T}}\boldsymbol{x}_j + r\right), \tag{4}$$

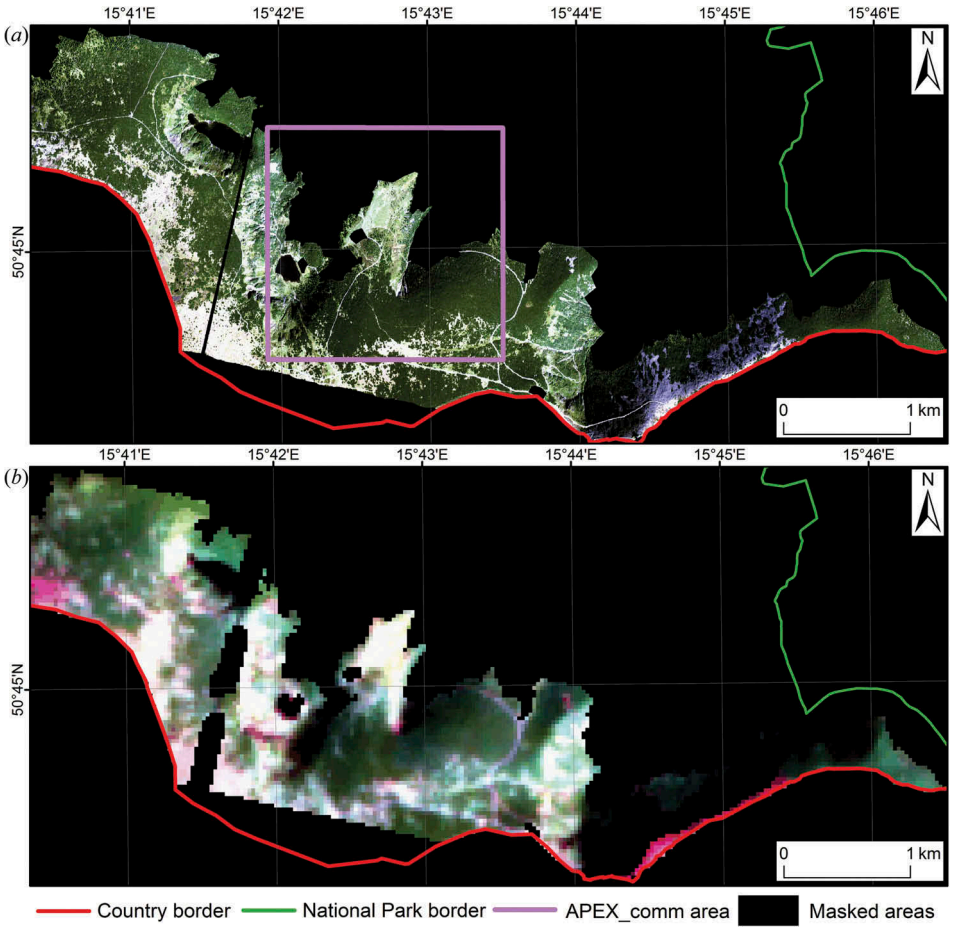


Figure 3. Masking of unwanted areas within the APEX (*a*) and EnMAP (*b*) images.

where $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ are the $\boldsymbol{x}$ data vectors ($_{i,j}$ – pixel window coordinates), γ is the Gamma term in all kernel functions except linear, d is the polynomial degree term in the kernel polynomial function, r is the bias terms in the kernel polynomial and sigmoid functions, and tanh is the hyperbolic tangent.

To find an optimal SVM kernel function in ENVI, all available functions and the parameters were tested and based on the results the classification accuracy was assessed. The polynomial degree (2) was set at 2.00 for the polynomial function (it was based on the literature: increasing the degree of kernel polynomial causes decreasing classification accuracy, Shah et al. 2003), the biases in the kernel polynomial (2) and sigmoid (4) functions were set to the default of 1.00 as this was the most optimal value. If the kernels required a gamma parameter, it was tuned with the C parameter that represented the cost of the penalty in order to find the best combination. Different gamma parameter values (from 0 to 1000) for the polynomial (2), RBF (3), and sigmoid (4) functions were tested and finally set as the inverse of the number of used original bands and it was 0.004 for 252 original and 0.033 for MNF APEX bands. If the penalty

parameter is too large the classification accuracy is high in the training phase but low in the phase of testing, however, when this value is too small, the classification accuracy is too low, making the classification model useless. After testing different penalty parameter values, from 0 to 1000, the same value (100.00) was used for all kernel functions, and the classifications were carried out without the probability threshold.

4.4. *Accuracy assessment*

The accuracies of the obtained maps were compared by constructing an error matrix and calculating the kappa coefficient (Congalton 1991). OA is computed by dividing the total number of correct pixels by the total number of pixels in the matrix, and the kappa statistic is based on the difference between the actual and the chance agreement in the matrix; both of these values indicate how well the classification result agrees with reference data. For each class the producer and user accuracies were also calculated, where the producer accuracy (PA) is the probability of a reference pixel being correctly classified and the user accuracy (UA) indicates the probability that a pixel classified on the map actually represents that category on the ground.

For each classification result, described accuracies were calculated. In the first step for the APEX_comm dataset of 252 bands and 30 MNF transforms and also for different kernel functions of SVM. In the second step, classification performance has been assessed at vegetation-type level for the APEX_types and EnMAP_types band datasets for 252 and 190 bands sets, respectively.

5. Results

The results are reported in two sections: in Section 5.1, APEX vegetation community classification results were obtained by using the band and MNF transform datasets, where the classification accuracy impact of using different SVM kernel functions was assessed; and in Section 5.2, APEX and EnMAP vegetation-type classification as the classification maps and accuracies was obtained for both datasets using the same parameters.

5.1. *APEX vegetation community classification*

The SVM classification was performed for both the original APEX bands (252) and the MNF transforms (30), with 23 classes (21 vegetation communities) extracted. Table 3

Table 3. Overall accuracies (OAs) for the four kernel functions and two different band sets for APEX_comm image.

	OA (%)	
SVM kernel function	252 bands	30 MNF transforms
Linear	74.39	74.35
Polynomial	62.28	73.57
Radial basis function	73.57	73.57
Sigmoid	46.54	71.78

presents the OA of the classification for all four kernel functions and the different datasets.

The best OA was achieved for the linear function using the original APEX bands (74.39%), while the MNF transforms had a slightly lower accuracy (74.35%). Second best was the RBF, achieving the same results for original and MNF bands datasets (73.57%). The worst results were observed for the sigmoid function using the full set of bands (46.54%).

In this study, the performance of the original band dataset depended on the chosen kernel function: for a linear approach, the OA was slightly higher (0.4%) but the MNF transformation gave better results for polynomial (about 11% higher accuracy) and sigmoid (about 25% higher); for RBF, the results were exactly the same. Figure 4(*b*) presents the classification image with the best result, with 21 classified vegetation communities (23 classes including water and mask) against a reference non-forest vegetation vector map (Figure 4(*a*)) with Table 4 showing the confusion matrix for this result.

The list of accuracies for each class is presented in Table 5.

Based on Tables 4 and 5, the best classified class was the early stages of succession that has PA and UA of about 95%. This unit occurs in bare and uncovered areas that are very well identified by the SVM classifier, though the area of the class is quite small. Very high accuracies were also achieved for the *Oxycocco-Sphagnetea* with a 97.4% UA and

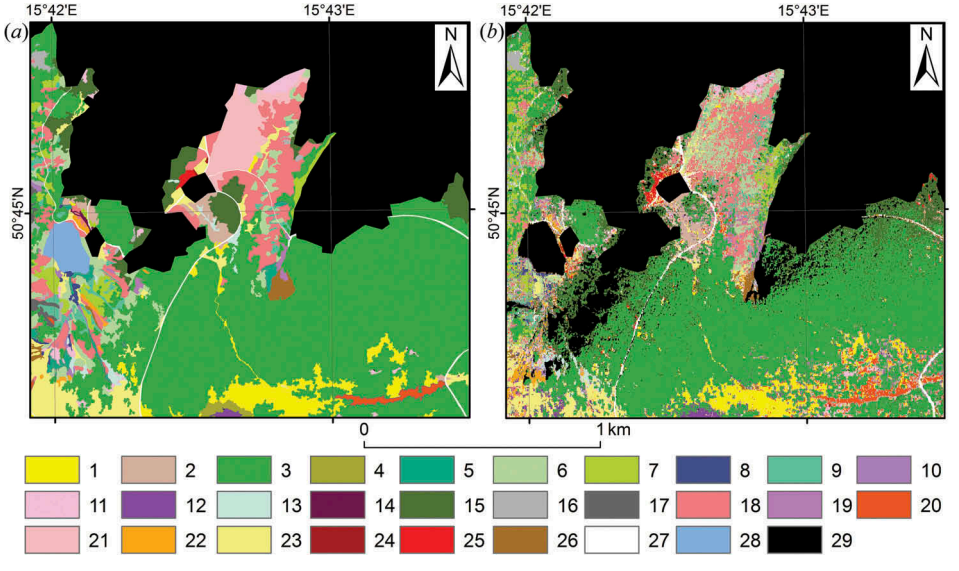


Figure 4. Reference non-forest vegetation map (*a*) and post-classification map of vegetation communities (*b*). Abbreviations developed in Table 1. 1: *Carici (rig.)-Nard.*; 2: *Empetro-Vacc.*; 3: *Pinetum mugo sudet.*; 4: *Pino mugo-Sphagnetum*; 5: *Calamagrostion*; 6: *Crepido-Calamagrost. vill.*; 7: *Pado-Sorb.*; 8: *Salic. lapp.*; 9: *Athyrietum dist.*; 10: *Adenostyletum all.*; 11: *Scheuchzerio-Caric. nigr.*; 12: *Oxycocco-Sphagn.*; 13: *Cardamino-Mont.*; 14: *Artemisietea vulg.*; 15: *Calamagr. villosae-Piceetum*; 16: *Rhizocarp. alp.*; 17: *Umbilic. cyl.*; 18: *Vacc. myrt.* com.; 19: *Calluna vulg.* com.; 20: *Molinia caer.* com.; 21: *Molinio Arrhenatheretea*; 22: *Deschamp. flex.* com.; 23: *Deschamp. caesp.* com.; 24: *Peucedanum ostr.* com.; 25: *Urtica dioica* com.; 26: early stages of succ.; 27: areas without veg.; 28: lakes; 29: masked areas.

Table 4. Confusion matrix – OA 74.39%, kappa coefficient 0.72.

		Reference data (accuracy, %)																						
	Class	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23
Classification (accuracy, %)	#1	**62**	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0
	#2	1	**91**	1	2	0	4	0	0	2	3	0	2	0	0	0	2	3	0	2	0	0	0	1
	#3	1	0	**91**	0	3	14	1	0	11	0	1	33	0	2	5	8	4	0	1	0	1	0	1
	#4	2	0	0	**68**	0	1	25	0	5	0	0	0	0	0	7	0	0	0	1	0	0	0	2
	#5	0	0	0	0	**86**	3	2	0	0	0	0	0	2	0	8	1	0	0	0	0	0	0	1
	#6	0	0	0	1	6	**34**	2	0	0	0	0	0	0	1	0	0	7	0	1	0	0	0	0
	#7	0	0	0	3	2	3	**39**	0	0	0	0	0	1	1	6	0	0	0	0	0	0	0	0
	#8	0	0	0	0	0	1	0	**68**	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	#9	0	0	1	1	0	0	1	0	**66**	1	0	0	0	0	3	0	2	0	3	0	3	0	0
	#10	0	0	0	0	0	0	0	0	0	**93**	0	0	0	0	1	0	0	0	0	0	0	0	0
	#11	1	3	0	0	0	3	0	21	0	0	**85**	0	0	0	1	0	0	0	1	0	0	0	0
	#12	0	0	3	1	0	10	1	0	0	0	1	**55**	0	4	6	0	0	0	0	0	1	0	0
	#13	0	0	0	0	0	0	0	0	0	0	0	0	**83**	0	0	0	0	1	0	0	30	0	0
	#14	0	0	0	0	0	3	4	0	0	0	0	0	0	**85**	1	0	0	0	0	0	0	0	0
	#15	0	1	0	10	1	5	21	0	1	0	0	1	15	0	**55**	0	0	0	2	0	0	0	0
	#16	0	0	2	3	1	6	2	0	0	0	1	2	0	3	1	**75**	1	1	1	0	0	1	6
	#17	4	0	0	0	1	0	0	0	5	0	0	1	0	0	0	0	**80**	0	1	0	1	0	0
	#18	3	0	0	1	0	9	1	8	0	0	1	1	0	1	2	0	0	**87**	2	30	3	0	0
	#19	24	2	0	11	0	2	0	0	8	3	1	3	0	0	2	0	2	7	**72**	10	7	0	1
	#20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**60**	3	0	0
	#21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	**49**	0	0
	#22	0	1	0	0	0	0	2	3	0	0	0	0	0	2	1	12	0	2	0	0	0	**95**	2
	#23	0	0	0	0	0	1	0	0	0	0	9	0	0	0	0	1	0	1	9	0	0	4	**85**
	Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

#1: *Carici (rig.)-Nard.*; #2: *Empetro-Vacc.*; #3: *Pinetum mugo sudet*; #4: *Crepido-Calamagrost. vill.*; #5: *Pado-Sorb.*; #6: *Salic. lapp.*; #7: *Athyrietum dist.*; #8: *Adenostyletum all.*; #9: *Scheuchzerio-Caric. nigr.*; #10: *Oxycocco-Sphagn.*; #11: *Cardamino-Mont.*; #12: *Calamagr. villosae-Piceetum*; #13: *Rhizocarp. alp.*; #14: *Umbilic. cyl.*; #15: *Vacc. myrt.* com.; #16: *Calluna vulg.* com.; #17: *Molinia caer.* com; #18: *Deschamp. flex.* com.; #19: *Deschamp. caesp.* com.; #20: *Peucedanum ostr.* com.; #21: *Urtica dioica* com.; #22: early stages of succ.; #23: areas without veg. (com. = community). Bold values indicates classified pixels in whole matrix (it is a standard procedure of presenting error matrix for the classes).

Table 5. Producer and user accuracy for classified vegetation communities (com. = community).

Class	Producer's accuracy (%)	User's accuracy (%)
Carici (rigidae)-Nardetum	61.9	94.4
Empetro-Vaccinietum	91.2	77.5
Pinetum mugo sudeticum	91.5	70.5
Crepido-Calamagrostietum villosae	67.8	51.1
Pado-Sorbetum	86.0	86.9
Salicetum lapponum	34.2	56.1
Athyrietum distentifolii	38.6	73.9
Adenostyletum alliariae	68.4	89.7
Scheuchzerio-Caricetea nigrae	66.1	75.7
Oxycocco-Sphagnetea	93.2	97.4
Cardamino-Montion	84.9	68.9
Calamagrostio villosae-Piceetum	55.1	52.6
Rhizocarpion alpicolae	82.5	83.2
Umbilicarion cylindricae	85.1	83.3
Vaccinium myrtillus com.	55.2	52.7
Calluna vulgaris com.	75.0	47.7
Molinia caerulea com.	79.6	69.5
Deschampsia flexuosa com.	87.2	54.7
Deschampsia caespitosa com.	72.0	82.4
Peucedanum ostruthium com.	60.0	75.0
Urtica dioica com.	49.3	50.8
Early stages of succession	95.2	94.4
Areas without vegetation	85.3	62.6

93.2% PA and for *Empetro-Vaccinietum* with 91.2% PA and 77.5% UA. These classes are a mixture of different minor species, which can be strongly distinguished from the surroundings that are homogenous grasslands or tall-forbs. High PA and UA were also obtained for *Pinetum mugo sudeticum*, which is the dominant class within the data and has 91.5% and 70.5%, respectively. It is supported by studies of Zagajewski (2010) where the best classified communities of Tatra Mountains in Poland, were also among others *Empetro-Vaccinietum* in a complex with *Pinetum mugho* and mountain-pine scrub on silicate substrate. Generally, the PA and UA of the results can be considered as high and range from 60% to about 95% for 18/23 classes of PA and for 16/23 classes of UA.

The poorest classification result was obtained for the *Salicetum lapponum*, which was greatly overestimated (34.2% PA and 56.1% UA). It is a shrub that is about 1 m in height, and not very widely located on steep shady slopes. So, if it was not present in groups then it was difficult to map due to both the terrain and APEX pixel size. Lower levels of both of accuracies were also reached for the *Urtica dioica* community with 49.3% PA and 50.8% UA which was represented by smaller test sample sizes such as *Athyrietum distentifolii* with 38.6% PA but higher UA of 73.9% and for *Calluna vulgaris* with 47.7% UA and 75% PA. The first two classes were difficult to map because of their temporal variability between the field survey dates, in 2013 and 2014, and the 2002 non-forest vegetation map. *Urtica dioica* belongs to ruderal vegetation located near shelters and its range is changeable and depends on the degree of human pressure which is also confirmed by studies in Tatras where one of the worst classified were communities dominated by ruderal plants and modified by seasonal grazing by animals (Zagajewski 2010). In 2008, *Athyrietum distentifolii* was observed to experience dieback connected with weevil attacks on the leaves. Both accuracies achieved about 50% for the

Calamagrostio villosae-Piceetum community. It is not really a non-forest community, and it is located on the border of the reference non-forest vegetation map (Wojtuń and Żołnierz 2002). It consists of single trees surrounded by other communities and was mostly confused with *Pinetum mugo sudeticum* because of the spectral similarity.

5.2. *APEX and EnMAP vegetation types classification*

Vegetation-type classification was conducted for both the APEX_types and EnMAP_types datasets. The results of this classification are presented in Figures 5(*a*) and (*b*) and Tables 6–8.

The OA for the APEX classification reached 90.72% and for EnMAP it was 78.25%, and the kappa statistics were 0.85 and 0.74, respectively. Very high classification accuracies in

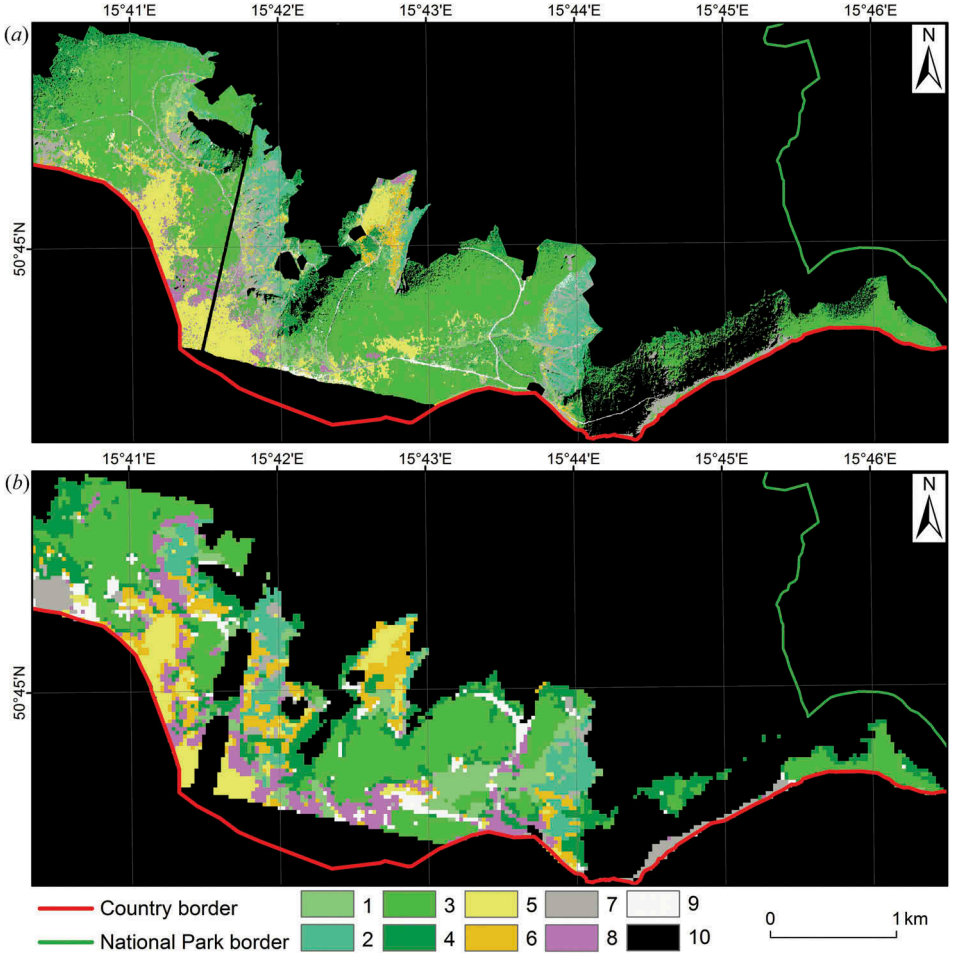


Figure 5. Vegetation-type classification images derived from the APEX (*a*) and EnMAP (*b*) datasets. 1: subalpine tall-forbs; 2: deciduous shrubs vegetation; 3: subalpine dwarf pine scrubs; 4: forests; 5: grasslands; 6: heathlands; 7: rock and scree vegetation; 8: bogs, fens, and springs; 9: areas without vegetation; 10: masked areas.

Table 6. APEX_types confusion matrix – OA 90.72%, kappa coefficient 0.85.

Classification (accuracy, %)	Reference data (accuracy, %)								
Class	Subalpine tall- forbs	Deciduous shrubs vegetation	Subalpine dwarf pine scrubs	Forests	Grasslands	Heathlands	Rock and scree vegetation	Bogs, fens and springs	Areas without vegetation
Subalpine tall-forbs	**86**	1	0	1	1	4	0	1	8
Deciduous shrubs vegetation	1	**67**	0	0	0	5	1	0	5
Subalpine dwarf pine scrubs	3	23	**97**	13	0	2	1	4	5
Forests	2	4	2	**85**	0	22	29	0	1
Grasslands	5	0	0	0	**98**	1	0	0	5
Heathlands	3	1	0	0	0	**62**	1	1	0
Rock and scree vegetation	1	1	0	0	0	2	**67**	1	12
Bogs, fens and springs	1	1	0	0	0	1	0	**93**	0
Areas without vegetation	0	1	0	1	0	0	1	0	**66**
Total	100	100	100	100	100	100	100	100	100

Bold values indicates classified pixels in whole matrix (it is a standard procedure of presenting error matrix for the classes).

Table 7. EnMAP_types confusion matrix – OA 78.25%, kappa coefficient 0.74.

Classification (accuracy, %)	Reference data (accuracy, %)								
Class	Subalpine tall- forbs	Deciduous shrubs vegetation	Subalpine dwarf pine scrubs	Forests	Grasslands	Heathlands	Rock and scree vegetation	Bogs, fens and springs	Areas without vegetation
Subalpine tall-forbs	**79**	0	0	0	0	3	0	3	22
Deciduous shrubs vegetation	13	**80**	0	0	0	0	0	10	0
Subalpine dwarf pine scrubs	0	10	**91**	5	0	0	0	7	0
Forests	4	0	9	**90**	0	21	2	3	0
Grasslands	0	0	0	0	**94**	21	2	3	0
Heathlands	0	0	0	0	0	**45**	14	10	0
Rock and scree vegetation	0	0	0	0	0	7	**72**	3	22
Bogs, fens and springs	0	10	0	0	6	0	0	**60**	9
Areas without vegetation	4	0	0	0	0	3	2	0	**47**
Total	100	100	100	100	100	100	100	100	100

Bold values indicates classified pixels in whole matrix (it is a standard procedure of presenting error matrix for the classes).

Table 8. Producer and user accuracies for the APEX and EnMAP classified vegetation types.

	APEX_types		EnMAP_types	
Class	Producer's accuracy (%)	User's accuracy (%)	Producer's accuracy (%)	User's accuracy (%)
Subalpine tall-forbs	86.2	76.1	79.2	67.9
Deciduous shrubs vegetation	67.4	69.9	80.0	80.0
Subalpine dwarf pine scrubs	97.5	96.8	91.1	94.1
Forests	84.9	57.3	89.7	63.6
Grasslands	98.4	97.0	94.2	86.0
Heathlands	62.3	94.0	44.8	52.0
Rock and scree vegetation	66.6	87.9	71.9	82.1
Bogs, fens and springs	93.4	91.4	60.0	66.7
Areas without vegetation	65.8	90.9	46.9	83.3

both datasets were reached for subalpine dwarf pine scrubs (about 90% of PA and UA) and grasslands (for APEX about 98% UA and PA and for EnMAP about 94% PA and 86% UA). Both classes are represented by large homogeneous patches and, even for the 30 m EnMAP pixels, finding representative samples was straightforward. In general, most of the classes were well classified with five of the APEX_types classes having PA of more than 70% and seven for UA, alongside six of the EnMAP_types classes also having PA and five of UA of more than 70%.

However, classes were classified better with the APEX_types dataset and the change was the greatest in case of bogs, fens and springs (about 33% more of PA and about 25% more of UA). APEX data vegetation community classification, which included *Oxycocco-Sphagnetea*, reached an accuracy of more than 90% for both PA and UA as this unit is easier to classify with the finest spatial resolution data because of the mosaic of different species. Less changeable, but also improved for APEX-based classification, were areas without vegetation, that were primarily mountain trails less than 30 m in width and composed of pixels with a mixture of trail and surrounding vegetation (about 19% difference between PA for the APEX and EnMAP data). However, deciduous shrubs vegetation and forests were classified better using the EnMAP data (for deciduous shrubs vegetation it was about 10–12% more of PA and UA and for forests about 5–6% more of PA and UA) because of the 30 m spatial resolution allowed for including shrubs or trees with surrounding vegetation into one class. A slightly better PA for EnMAP was also obtained for homogeneous rock and scree vegetation (about 5% more), but UA was better for APEX data. This class was well identified because it was spectrally homogeneous and easy to find on the image and in the terrain. The same class was classified the worst for both datasets – heathlands. In the APEX_types dataset 62.3% of heathlands have been correctly identified (PA), while UA being better (94%). A similar situation was also present in the EnMAP_types dataset where the PA was only 44.8% and the UA was slightly higher and it was 52%. This unit contains different types of bilberries and heathers, so it is heterogeneous and covers smaller areas. Especially for EnMAP, the spatial resolution makes it difficult to determine the exact boundaries of the training samples and so the mixed nature of this vegetation type made it a challenge to classify; potentially, more training pixels could improve the results.

6. Discussion

An OA for the 252 band APEX dataset reached about 75% and kappa coefficient was 0.72, which is a satisfactory result. These results are supported by preliminary results for a non-forest and forest vegetation communities classification using raw APEX data in the western part of Karkonosze (Marcinkowska et al. 2014), where an OA reached 79.1% and kappa 0.77. The OA and kappa obtained for forest species classification using APEX in Karkonosze was 78.6% and 0.71, respectively (Raczko et al. 2015), which also confirms the accuracy of this article. It is also supported by studies of vegetation communities in Tatras (Zagajewski 2010), where, admittedly, the OA was higher at 88.6% and the kappa 0.87 but the accuracies for individual classes were comparable, e.g. more than 90% PA for *Empetro-Vaccinietum* or *Pinetum mugo* and the worst classified complex communities as willow thicket or *Vaccinium myrtilus* community in a complex with tall herb communities.

It was difficult to map heterogeneous mountain vegetation, especially at the community level. To reduce the mixed pixel classification problem, Belluco et al. (2006) suggests using a high spatial resolution dataset. Clark, Roberts, and Clark (2005) discriminated tropical rain forest tree species at leaf to crown levels using three scales: field spectral characteristics, HYDICE airborne image and also simulated multispectral broadband imagery with fixed spatial resolution. They observed a decrease in classification accuracy from fine to coarser scales, which is also confirmed by the present study. Ghosh et al. (2014), where the three spatial data resolutions were 4, 8 (HyMap data), and 30 m (Hyperion data), concluded that 8 m resolution was the best for tree species classification. It was explained as a compromise between over-generalization and spectral mixing within 30 m pixels, contrasting with higher spectral variance within a species class resulting in lower spectral separability amongst species in 4 m pixel. Overall, the differences between the best results of 4 and 8 m were small and not negatively influencing accuracies. It should be noted that tree species have different physiognomy compared to non-forest vegetation communities or types, with the compromise between a too general and too detailed spatial resolution being consistent with the results presented, i.e. why the classification of APEX data for vegetation types achieved the best result. In this context, the result obtained for vegetation communities should be considered as satisfactory, because a fine spatial resolution dataset (e.g. 12 cm orthophotomaps) could cause communities containing mosaics of different species not to be distinguished properly. A comparative study of different scales performed by Zhang and Xie (2013) found that the community-level classification was easier and better performed than species-level, with the RF method achieving an OA of 36% for the original dataset and 79% for the MNF transformed data. The authors used also segmentation, which improved discrimination of larger and more complex communities than a big number of different species. Comparison of OA and kappa coefficient in vegetation-type classification of presented study, for both the APEX and EnMAP datasets, revealed that using finer spatial scales was slightly better. However, comparing vegetation-type classifications between both datasets, showed that the data are characterized by a significantly different spatial resolution and therefore it could be concluded that the OA differences and kappa values are not high and also PA and UA for classes are not very varied.

The most common approach in the literature is to use the RBF kernel (Foody and Mathur 2004; Chan et al. 2012; Lin et al. 2008), which was in second position for the vegetation community classification accuracy. There are some situations where the RBF is not the best, when the number of features is very large, which is where one may just use the linear kernel (Hsu, Chang, and Lin 2010). In this case, the best OA was achieved for this approach when using the original APEX spectral bands. The worst classification accuracy was obtained for sigmoid function (about 28% worse than for linear) which is supported by studies of Shah et al. (2003). Testing the penalty and gamma parameters allowed for choose the most optimal combinations for the functions requiring these parameters leading to reach the best classification accuracy. The results for all functions were more stable using data without noise (MNF transforms).

Hyperspectral data give an opportunity to find even slight differences between similar classes, which is not possible using multispectral data with wider ranges and less number of bands. Some classes of vegetation community or type differentiate more in the red edge region, another classes are distinguishable significantly in the SWIR region. For that reason, the data with many bands bring higher probability to find differences between classes, especially when the object of the classification is a mosaic of different species as a vegetation community. Demarchi et al. (2014), for detailed urban mapping (22 classes), also achieved the best accuracy when all the original bands were used. Using dimensionality reduction techniques is varying depending on used data and methods. Pal and Mather (2006) found that MNF transformation resulted in decreased accuracy, but in the studies of Zhang and Xie (2013), the authors observed increased classification accuracy. The fact is that this transformation decreases the computational time, and so speeds up the classification process, which would be a potential trade-off against OA. So, in the presented work, combining the MNF method and SVM classifier is promising for reducing the high data volume with no significant loss of accuracy, but for presenting the full potential of APEX and forthcoming EnMAP data, it was better to present the whole spectrum of available spectral bands.

Visual interpretation supported the accuracy assessment results. The vegetation community map was similar to a reference non-forest vegetation map, but attention should be paid on its actuality and the information from the 2013 and 2014 field survey. The maps of vegetation type were similar in most areas, especially in large, and homogeneous patches. For EnMAP_type dataset, some breaks were observed in linear trails labelled as areas without vegetation and confused with rock and scree vegetation because of their spectral similarity. Too large a spread was observed for bogs, fens, and springs which is confirmed by its commission in the error matrix.

The accuracy results could be improved by performing topographic correction of the APEX data, which might allow the classification of shadow affected areas. Including these areas would be important because of their very high diversity, especially near Mały Staw Cirque. For EnMAP, this problem was less significant because of its spatial resolution, which could not distinguish linear-shaped thin vegetation such as herbs, so the area would have been classified as the majority of the neighbouring pixels – subalpine dwarf pine scrubs. The rest of the masked area is located near Sniezka peak, where there is only rock and scree vegetation alongside subalpine dwarf pine scrubs. Therefore, the lost information comes only from these two classes, whose location is possible to predict so the problem is less serious than in the classification of vegetation communities.

Otherwise, the lack of topographic correction did not significantly affect the classification results.

7. Conclusions

This study showed that hyperspectral aerial and satellite data can be used to classify subalpine and alpine vegetation, for satellite at vegetation types and for aerial at both, vegetation community and types scale. For the vegetation community classification, the 252 band set offered the highest accuracies when compared to the MNF input (30 transforms) that underlines the necessity for full spectrum analyses and the use of imaging spectroscopy overall; the 252 band set achieved an OA of 74.39% while the MNF transforms reached 74.35%. The difference is not significant, however, the MNF input did provide a classification time that was more than two times shorter. Overall, the SVM linear kernel function provided the best results (74.39%), with the worst being from the sigmoid kernel function (46.54% for the 252 band set). Very similar results were achieved for the polynomial and RBF kernels with MNF transforms dataset, with OA of around 73%, for original dataset the difference between these two function was greater (about 10% less for polynomial function).

These classifications provided maps of 21 vegetation communities, where 13 classes were classified with accuracies of higher than 70% (PA), and three classes at less than 50% accuracy. Furthermore, SVM was especially beneficial for heterogeneous classes, for which only a few training samples could be identified and when the number of samples was large (around 400 pixels). Overall, the achieved results were acceptable (more than 70%).

This research also demonstrated the potential of APEX data and the SVM method for producing detailed vegetation maps, which represent a large number of classes that were difficult to identify in the terrain; especially in mountainous areas. The best classified vegetation communities, excluding early stages of succession which were not a community, but very well classified (about 95% PA and UA) were the *Oxycocco-Sphagnetea* and *Empetro-Vaccinietum* (more than 90% PA), which were smaller units and a mosaics of different minor species. In addition, high PA was obtained for *Pinetum mugo sudeticum*, which is the dominant class on the image. Also, the reference non-forest vegetation map included some classes that are (from a phytosociological point of view) important, but not identified by the unmixing of pixels within the APEX image. The *Pino-mugo Sphagnetum* and *Pinetum mugo sudeticum* communities were combined into one because they were very much spectrally similar, the same situation was for *Artemisietea vulgaris* and *Urtica dioica* so it should be one class in the hyperspectral mapping. EnMAP classifications of ruderal vegetation and herbs classes were not distinguishable because too small an area was covered by these classes, and so it was not possible to classify and assess their accuracy properly.

The forthcoming EnMAP data will be a very powerful tool for vegetation identification and monitoring. The overall accuracies of the EnMAP and APEX classifications reached high values, with most classified to an accuracy of more than 65% in both datasets. Some classes reached higher accuracy values for APEX (e.g. deciduous shrubs vegetation), because of the complexity and smaller geographical range of these classes. Noting, that the spatial resolution of EnMAP is significantly lower than

APEX, this did not have a significant impact on the overall accuracies and kappa results. However, it did have an impact in the case of small and non-compact patches of vegetation types as in the case of ruderal vegetation that was removed from the final version.

According to this research, hyperspectral airborne data are a suitable solution for obtaining information about vegetation, because the high spatial and spectral resolution allows for the detailed classification of communities. Unfortunately, airborne data can be costly and often has only limited availability. Hence, spaceborne imaging spectroscopy enables a higher temporal resolution that benefits the overall monitoring approach, and even enables the investigation of short-term trends. Therefore, it is concluded that hyperspectral remote sensing techniques can be used to reduce the amount of field-work, especially when searching for and mapping vegetation communities and types. In addition, it allows consistent and spatiotemporal extensive monitoring of sensitive and difficult to access mountainous regions.

Acknowledgements

The authors would like to thank EUFAR, DLR and VITO for the HyMountEcos 2012 project, KPN for the opportunity of conducting the research and GFZ for simulation the EnMAP data. Special thanks to Koen Meuleman, Johan Mijnendonckx, Kristin Vreys, Sindy Sterckx, Marian-Daniel Iordache from VITO, Lidia Przewoźnik from KPN, and Nina Kristina Boesche and Karl Segl from GFZ.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Belluco, E., M. Camuffo, S. Ferrari, L. Modenese, S. Silvestri, A. Marani, and M. Marani. 2006. "Mapping Salt-Marsh Vegetation by Multispectral and Hyperspectral Remote Sensing." *Remote Sensing of Environment* 105 (1): 54–67. doi:10.1016/j.rse.2006.06.006.

Boardman, J. W., and F. A. Kruse. 1994. "Automated Spectral Analysis: A Geological Example Using AVIRIS Data, North Grapevine Mountains, Nevada." In *Proceedings of the ERIM Tenth Thematic Conference on Geologic Remote Sensing, Ann Arbor*, I407–I418. Ann Arbor, MI: Environmental Research Institute of Michigan, ERIM.

Burges, C. J. C. 1988. "A Tutorial on Support Vector Machines for Pattern Recognition." *Data Mining and Knowledge Discovery* 2: 121–167. doi:10.1023/A:1009715923555.

Camps-Valls, G., L. Gomez-Chova, J. Calpe-Maravilla, J. D. Martin-Guerrero, E. Soria-Olivas, L. Alonso-Chorda, and J. Moreno. 2004. ""Robust Support Vector Method for Hyperspectral Data Classification and Knowledge Discovery." *IEEE Transactions on Geoscience and Remote Sensing* 42 (7): 1530–1542. doi:10.1109/TGRS.2004.827262.

Chan, J., and D. Paelinckx. 2008. "Evaluation of Random Forest and Adaboost Tree-Based Ensemble Classification and Spectral Band Selection for Ecotope Mapping Using Airborne Hyperspectral Imagery." *Remote Sensing of Environment* 112 (6): 2999–3011. doi:10.1016/j.rse.2008.02.011.

Chan, J. C.-W., P. Beckers, T. Spanhove, and J. V. Borre. 2012. "An Evaluation of Ensemble Classifiers for Mapping NATURA 2000 Heathland in Belgium Using Spaceborne Angular Hyperspectral (Chris/Proba) Imagery." *International Journal of Applied Earth Observation and Geoinformation* 18: 13–22. doi:10.1016/j.jag.2012.01.002.

Chang, C.-I., and A. Plaza. 2006. "A Fast Iterative Algorithm for Implementation of Pixel Purity Index." *IEEE Geoscience and Remote Sensing Letters* 3 (1): 63–67. doi:10.1109/LGRS.2005.856701.

Clark, M. L., D. A. Roberts, and D. B. Clark. 2005. "Hyperspectral Discrimination of Tropical Rain Forest Tree Species at Leaf to Crown Scales." *Remote Sensing of Environment* 96 (3): 375–398. doi:10.1016/j.rse.2005.03.009.

Congalton, R. G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37: 35–46. doi:10.1016/0034-4257(91)90048-B.

D'Odorico, P., A. Gonsamo, A. Damm, and M. E. Schaepman. 2013. "Experimental Evaluation of Sentinel-2 Spectral Response Functions for NDVI Time-Series Continuity." *IEEE Transactions on Geoscience and Remote Sensing* 51: 1336–1348. doi:10.1109/TGRS.2012.2235447.

Damm, A., L. Guanter, W. Verhoef, D. Schlaepfer, S. Garbari, and M. E. Schaepman. 2015. "Impact of Varying Irradiance on Vegetation Indices and Chlorophyll Fluorescence Derived from Spectroscopy Data." *Remote Sensing of Environment* 156: 202–215. doi:10.1016/j.rse.2014.09.031.

de Haan, J. F., J. W. Hovenier, J. M. M. Kokke, and H. T. C. Van Stokkom. 1991. "Removal of Atmospheric Influences on Satellite-Borne Imagery: A Radiative Transfer Approach." *Remote Sensing of Environment* 37: 1–21. doi:10.1016/0034-4257(91)90046-9.

Demarchi, L., F. Canters, C. Cariou, G. Licciardi, and J. C. W. Chan. 2014. "Assessing the Performance of Two Unsupervised Dimensionality Reduction Techniques on Hyperspectral APEX Data for High Resolution Urban Land-Cover Mapping." *ISPRS Journal of Photogrammetry and Remote Sensing* 87: 166–179. doi:10.1016/j.isprsjprs.2013.10.012.

Filippi, A. M., and J. R. Jensen. 2006. "Fuzzy Learning Vector Quantization for Hyperspectral Coastal Vegetation Classification." *Remote Sensing of Environment* 100: 512–530. doi:10.1016/j.rse.2005.11.007.

Foody, G. M., and A. Mathur. 2004. "A Relative Evaluation of Multi-Class Image Classification by Support Vector Machines." *IEEE Transactions on Geoscience and Remote Sensing* 42: 1335–1343. doi:10.1109/TGRS.2004.827257.

Ghosh, A., F. E. Fassnacht, P. K. Joshi, and B. Koch. 2014. "A Framework for Mapping Tree Species Combining Hyperspectral and Lidar Data: Role of Selected Classifiers and Sensor across Three Spatial Scales." *International Journal of Applied Earth Observation and Geoinformation* 26: 49–63. doi:10.1016/j.jag.2013.05.017.

Goetz, A. F. H., G. Vane, J. E. Solomonson, and B. N. Rock. 1985. "Imaging Spectrometry for Earth Remote Sensing." *Science* 228: 1147–1153. doi:10.1126/science.228.4704.1147.

Green, A. A., M. Berman, P. Switzer, and M. D. Craig. 1988. "A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal." *IEEE Transactions on Geoscience and Remote Sensing* 26: 2665–2674. doi:10.1109/36.3001.

Gualtieri, J. A., and R. F. Cromp. 1988. "Support Vector Machines for Hyperspectral Remote Sensing Classification." In *Proceedings of the 27th AIPR Workshop: Advances in Computer Assisted Recognition, Washington, DC, Oct. 27, 1988*, 221–232. Washington, DC: International Society for Optics and Photonics, SPIE.

Guanter, L., H. Kaufmann, K. Segl, S. Foerster, C. Rogass, S. Chabrillat, T. Kuester, et al. 2015. ""The Enmap Spaceborne Imaging Spectroscopy Mission for Earth Observation." *Remote Sensing* 7: 8830–8857. doi:10.3390/rs70708830.

Heldens, W., U. Heiden, T. Esch, E. Stein, and A. Mueller. 2011. "Can the Future Enmap Mission Contribute to Urban Applications? A Literature Survey." *Remote Sensing* 3: 1817–1846. doi:10.3390/rs3091817.

Hsu, C. W., C. C. Chang, and C. J. Lin 2010. "A Practical Guide to Support Vector Classification". *National Taiwan University, Taiwan*. Available from: http://ntu.csie.org/~cjlin/papers/guide/guide.pdf.

Huang, C., L. S. Davis, and J. R. G. Townshend. 2002. "An Assessment of Support Vector Machine for Land Cover Classification." *International Journal of Remote Sensing* 23: 725–749. doi:10.1080/01431160110040323.

Hueni, A., D. Schlaepfer, M. Jehle, and M. E. Schaepman. 2014a. "Impacts of Dichroic Prism Coatings on Radiometry of the Airborne Imaging Spectrometer APEX." *Applied Optics* 53: 5344–5352. doi:10.1364/AO.53.005344.

Hueni, D., D. Schlaepfer, and M. Jehle. 2014b. "APEX: Radiometry under Spectral Shift Conditions." In *Geoscience and Remote Sensing Symposium (IGARSS). Proceedings of the 2014 IEEE International*

Geoscience and Remote Sensing Symposium, Quebec, Canada, 13-18 Jul. 1381–1384. doi: 10.1109/IGARSS.2014.6946692

Itten, K. I., F. Dell'Endice, A. Hueni, M. Kneubühler, D. Schläpfer, D. Odermatt, F. Seidel, et al. 2008. "APEX – the Hyperspectral ESA Airborne Prism Experiment." *Sensors* 8: 6235–6259. doi:10.3390/s8106235.

Jarocińska, A. M. 2014. "Radiative Transfer Model Parametrization for Simulating the Reflectance of Meadow Vegetation." *Miscellanea Geographica – Regional Studies on Development* 18 (2): 5–9. doi:10.2478/mgrsd-2014-0001.

Jelének, J., L. Kupková, B. Zagajewski, S. Březina, A. Ochytra, and A. Marcinkowska. 2014. "Laboratory and Image Spectroscopy for Evaluating the Biophysical State of Meadow Vegetation in the Krkonoše National Park." *Miscellanea Geographica – Regional Studies on Development* 18 (2): 15–22. doi:10.2478/mgrsd-2014-0014.

Kokaly, R. F., D. G. Despain, R. N. Clark, and K. E. Livo. 2003. "Mapping Vegetation in Yellowstone National Park Using Spectral Feature Analysis of AVIRIS Data." *Remote Sensing of Environment* 84: 437–456. doi:10.1016/S0034-4257(02)00133-5.

Kondracki, J. 2009. *Geografia Regionalna Polski*. Warsaw, Poland: PWN.

Kycko, M., B. Zagajewski, and A. Kozłowska. 2014. "Variability in Spectral Characteristics of Trampled High-Mountain Grasslands." *Miscellanea Geographica – Regional Studies on Development* 18 (2): 10–14. doi:10.2478/mgrsd-2014-0003.

Lin, S. W., K. C. Ying, S. C. Chen, and Z. J. Lee. 2008. "Particle Swarm Optimization for Parameter Determination and Feature Selection of Support Vector Machines." *Expert Systems with Applications* 35 (4): 1817–1824. doi:10.1016/j.eswa.2007.08.088.

Locherer, M., T. Hank, M. Danner, and W. Mauser. 2015. "Retrieval of Seasonal Leaf Area Index from Simulated Enmap Data through Optimized LUT-Based Inversion of the PROSAIL Model." *Remote Sensing* 7 (8): 10321–10346. doi:10.3390/rs70810321.

Marcinkowska, A., B. Zagajewski, A. Ochtyra, A. Jarocińska, E. Raczko, L. Kupková, P. Stych, and K. Meuleman. 2014. "Mapping Vegetation Communities of Karkonosze National Park Using APEX Hyperspectral Data and Support Vector Machines." *Miscellanea Geographica – Regional Studies on Development* 18 (2): 23–29. doi:10.2478/mgrsd-2014-0007.

Melgani, F., and L. Bruzzone. 2004. "Classification of Hyperspectral Remote Sensing Images with Support Vector Machines." *IEEE Transactions on Geoscience and Remote Sensing* 42 (8): 1778–1790. doi:10.1109/TGRS.2004.831865.

Pal, M., and P. M. Mather. 2004. "Assessment of the Effectiveness of Support Vector Machines for Hyperspectral Data." *Future Generation Computer Systems* 20 (7): 1215–1225. doi:10.1016/j.future.2003.11.011.

Pal, M., and P. M. Mather. 2006. "Some Issues in the Classification of DAIS Hyperspectral Data." *International Journal of Remote Sensing* 27: 2895–2916. doi:10.1080/01431160500185227.

Raczko, E., B. Zagajewski, A. Ochtyra, A. Jarocińska, A. Marcinkowska-Ochtyra, and M. Dobrowolski. 2015. "Określenie Składu Gatunkowego Lasów Góry Chojnik (Karkonoski Park Narodowy) Z Wykorzystaniem Lotniczych Danych Hiperspektralnych APEX." *Sylwan* 159 (7): 593–599.

Schepers, L., B. Haest, S. Verraverbeke, T. Spanhove, J. V. Borre, and R. Goossens. 2014. "Burned Area Detection and Burn Severity Assessment of a Heathland Fire in Belgium Using Airborne Imaging Spectroscopy (APEX)." *Remote Sensing* 6 (3): 1803–1826. doi:10.3390/rs6031803.

Schlaepfer, D., J. Biesemans, A. Hueni, and K. Meuleman. 2008. "Evaluation of the Atmospheric Correction Procedure for the APEX Level 2/3 Processor". In *Proc. SPIE 7107. Remote Sensing of Clouds and the Atmosphere XIII, Cardiff, Wales, United Kingdom*,September 2008 15-17. doi: 10.1117/12.799884.

Segl, K., L. Guanter, C. Rogass,, T. Küster, S. Roessner, H. Kaufmann, B. Sang, V. Mogulsky, and S. Hofer. 2012. "Eetes – the Enmap End-To-End Simulation Tool." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5: 522–530. doi:10.1109/JSTARS.2012.2188994.

Shah, C. A., P. Watanachaturaporn, P. K. Varshney, and M. K. Arora. 2003. "Some Recent Results on Hyperspectral Image Classification." In *Advances in Techniques for Analysis of Remotely Sensed Data, 2003 IEEE Workshop*, *Maryland, USA*, October 27-28, 346–353. doi:10.1109/WARSD.2003.1295214.

Sobik, M., M. Błaś, K. Migała, M. Godek, and T. Nasiółkowski. 2013. "Klimat." In *Przyroda Karkonoskiego Parku Narodowego*, edited by R. Knapik and A. Raj, 147–186. Jelenia Góra, Poland: Karkonoski Park Narodowy.

Sterckx, S., K. Vreys, J. Biesemans, M. D. Iordache, L. Bertels, and K. Meuleman. 2016. "Atmospheric Correction of APEX Hyperspectral Data." *Miscellanea Geographica – Regional Studies on Development* 20 (1): 16–20. doi:10.1515/mgrsd-2015-0022.

Śtursa, J. 2013. "The Development of Opinions on the Geo-Biodiversity of the Giant Mountains' Arctic-Alpine Tundra and Its Conservation." *Opera Corcontica* 50: 55–74.

Suess, S., S. Van Der Linden, A. Okujeni, P. J. Leitão, M. Schwieder, and P. Hostert. 2015. "Using Class Probabilities to Map Gradual Transitions in Shrub Vegetation from Simulated Enmap Data." *Remote Sensing* 7: 10668–10688. doi:10.3390/rs70810668.

Tagliabue, G., C. Panigada, R. Colombo, F. Fava, C. Cilia, F. Baret, K. Vreys, K. Meuleman, and M. Rossini. 2016. "Forest Species Mapping Using Airborne Hyperspectral APEX Data." *Miscellanea Geographica – Regional Studies on Development* 20 (1): 28–33. doi:10.1515/mgrsd-2016-0002.

Tarabalka, Y., M. Fauvel, J. Chanussot, and J. A. Benediktsson. 2010. "SVM and MRF-Based Method for Accurate Classification of Hyperspectral Images." *IEEE Geoscience and Remote Sensing Letters* 7 (4): 736–740. doi:10.1109/LGRS.2010.2047711.

Theiler, J. P., D. D. Lavenier, N. R. Harvey, S. J. Perkins, and J. J. Szymanski. 2000. ""Using Blocks of Skewers for Faster Computation of Pixel Purity Index". *Proc. SPIE 4132*." *Imaging Spectrometry VI* 61: 61–71.

Underwood, E., S. Ustin, and D. DiPietro. 2003. "Mapping Nonnative Plants Using Hyperspectral Imagery." *Remote Sensing of Environment* 86 (2): 150–161. doi:10.1016/S0034-4257(03)00096-8.

Vapnik, V. N. 1995. *The Nature of Statistical Learning Theory*. New York, United States of America: Springer-Verlag.

Vreys, K., M. D. Iordache, J. Biesemans, and K. Meuleman. 2016. "Geometric Correction of APEX Hyperspectral Data." *Miscellanea Geographica – Regional Studies on Development* 20 (1): 11–15. doi:10.1515/mgrsd-2016-0006.

Vreys, K., M. D. Iordache, and K. Meuleman. 2014, April,11. "APEX 2012 – September Campaign Hymountecos Data Delivery Report". VITO Center of Rem. Sens. & Earth Obs. Proc. Belgium. Mol. Rep. APX-VTO-HYME2012 .

Waske, B., S. Van Der Linden, J. A. Benediktsson, A. Rabe, and P. Hostert. 2010. "Sensitivity of Support Vector Machines to Random Feature Selection in Classification of Hyperspectral Data." *IEEE Transactions on Geoscience and Remote Sensing* 48 (7): 2880–2889. doi:10.1109/TGRS.2010.2041784.

Wojtuń, B., and L. Żołnierz. 2002. "Plan Ochrony Ekosystemów Nieleśnych – Inwentaryzacja Zbiorowisk." In *Plan Ochrony Karkonoskiego Parku Narodowego*. Brzeg, Poland: Bureau for Forest Management and Geodesy.

Zagajewski, B. 2010. "Ocena Przydatności Sieci Neuronowych I Danych Hiperspektralnych Do Klasyfikacji Roślinności Tatr Wysokich [Assessment of Neural Networks and Imaging Spectroscopy for Vegetation Classification of the High Tatras]". *Teledetekcja Środowiska 43*. Warsaw, Poland: Klub Teledetekcji Środowiska, Polskie Towarzystwo Geograficzne

Zhang, C., and Z. Xie. 2013. "Object-Based Vegetation Mapping in the Kissimmee River Watershed Using Hymap Data and Machine Learning Techniques." *Wetlands* 33 (2): 233–244. doi:10.1007/s13157-012-0373-x.

Żołnierz, L., and B. Wojtuń. 2013. "Roślinność Subalpejska I Alpejska." In *Przyroda Karkonoskiego Parku Narodowego*, edited by R. Knapik and A. Raj, 241–278. Jelenia Góra, Poland: Karkonoski Park Narodowy.