Heterogeneous forest classification by creating mixed vegetation classes using EO-1 Hyperion


Manuscript ID: TRES-PAP-2016-0483.R2

Manuscript Type: IJRS Research Paper

Date Submitted by the Author: 30-Dec-2016

Complete List of Authors: Deák, Márton; Eötvös Loránd University, Department of Physical Geography
Telbisz, Tamás; Eötvös Loránd University, Department of Physical Geography
Árvai, Mátyás; Hungarian Academy of Sciences, Department of Soil Mapping and Environmental Informatics
Mari, László; Eötvös Loránd University, Department of Physical Geography
Horváth, Ferenc; Hungarian Academy of Sciences, Centre for Ecological Research
Kohán, Balázs; Eötvös Loránd University, Department of Environmental and Landscape Geography
Szabó, Orsolya; Partium Christian University, Department of Economics
Kovács, József; Eötvös Loránd University, Department of Physical and Applied Geology

Keywords: hyperspectral data, classification, forestry

Keywords (user defined): band selection
Reply to reviewer 1

Question 1:
The sentence in page 11, lines 14-21 (Studies about … full spectra) refers to the number of spectral bands that were used for classification. Spectral indices are based on 2-3 bands in general whereas full spectra is more than 100 bands in general.

Answer 1:
Corrected like the referee suggested – please see in Answer 3.

Question 2:
The next sentence (Classification is … be very effective) in page 11, lines 21-27 refers to the usage of other features in addition to spectral information. Dalponte et al. (Dalponte et al. 2008) utilize LiDAR data whereas Tarabalka et al. (Tarabalka et al. 2009) utilize segmentation map based on spatial distribution. On the other hand, neural network approaches (Gong et al. 1997, Wang et al. 2008, Caiyun 2014) use only spectral information. As I have repeatedly said in the previous comments, a data processing technique (neural network) cannot be compared with types of data (LiDAR data and segmentation map). In that sense, I don’t think referring to neural network in this sentence is appropriate. Considering that the neural networks utilize whole spectral bands, I suppose one option is to move the references (Gong et al. 1997, Wang et al. 2008, Caiyun 2014) to the references with respect to full spectra (Williams et al. 2003, Clark et al. 2005, Boitt et al. 2014, Atzberger et al. 2015, Burai et al. 2015) in page 11, lines 14-21.

Answer 2:
Also corrected like the referee suggested – please see in Answer 3.

Question 3:
2. high resolution (Page 11, line 10, 14) and lower resolution (Page 11, line 12)
Please specify the types of the resolutions, i.e., spectral or spatial resolution.

Answer 3:
Also corrected. Most of the paragraph looks like this now:
“Using this technology species are mostly identified using high spatial resolution hyperspectral data but we can find less papers in the topic of the application of lower spatial resolution hyperspectral satellite images. Studies about the application of high spatial resolution data are often concentrating on the biochemical properties of the plants which are connected to spectral indices or extracted from different parts of the full spectra. Studies using high spectral resolution field and laboratory spectrometry are even more chemistry-dependent. These studies are using different techniques and methods utilizing field/laboratory spectrometer, aerial hyperspectral imagery or both (Gong et al. 1997, Williams et al. 2003, Clark et al. 2005, Boitt et al. 2014, Caiyun 2014, Atzberger et al. 2015, Burai et al. 2015, Wang et al. 2015).”

Question 4:
3. classification accuracy (Page 19, lines 50-53)
At this point, the accuracies have not been shown. Please specify where (i.e., section 5) the evidence (85.78 %) is shown.

Answer 4:
Corrected like this:
“As stated in section 5 in the previous step the highest accuracy was achieved by using our influential band reduction method (84.28%), the rule images created during that classification process were further utilized.”

Question 5:
4. other combinations (Page 20, lines 32-43)
The authors describe that there is a priori information that the possible classes are limited to oak, maple, pine, beech, ground vegetation, oak-pine, oak-maple, oak-beech and maple-beech. But the a priori constraints are not used in the classification process (Page 22, lines 19-33). Without the constraints, there is a possibility that some pixels are classified into other combinations, e.g., maple-beech, pine-beech. However Table 3b excludes the impossible combinations. Were the impossible combinations (e.g., maple-beech, pine-beech) avoided without any constraint?

Answer 5:
In the sample area there were some species which were not mixing due to natural biogeographical reasons – these are maple-pine and beech-pine mixtures (maple-beech was present as stated in the paper). Although we might not have been clear only by listing the possible classes. In this case we used some constraints since there still is a possibility that one pixel contains eg. 50% pine and 50% maple in a non-mixing way (eg. there is a sharp boundary between a homogenous maple and a homogenous pine block which is running in the middle of a pixel), but the probability of that is very low.

Considering this we think that the reviewer is right at this point. We think by studying a larger sample area there is a small possibility that such “false-positive” mixed pixels will be present in small quantities. To avoid detecting them and to clarify we added some more sentences in section 4.3:

“Pine-maple and pine-beech mixtures were not present, so we excluded them from the further steps.”

and the end of the section:

“However pine-maple and pine-beech mixtures in the study area are not present there is still a possibility of classifying some pixels as such (eg. at the boundaries of two touching homogeneous pine and maple blocks). To avoid this, we introduced the constraint: if $P_{pine} \geq 0.25$, then maple and beech classes were not investigated.”

Question 6a:

5. Fig. 3

From the view point of band selection, the comparison should be based on the identical number of bands. For example, the classification accuracy based on 51 (proposed method) should be compared with the classification accuracy based on 51 (SDA). When the number of bands is 47, the classification accuracy based on SDA outperforms the classification based on the proposed method. However, when the number of bands is increased to 51, the classification accuracy based on the proposed method slightly outperforms the classification based on SDA.

Answer 6a:
We think that comparing the best results produced by each method shows the real difference between them, but still any other comparisons can be seen in Fig. 3. But for further clarification we added the following sentences to the “6. Discussion” section:

“We also compared the two optimal number of bands (47 for SDA and 51 for influential band selection). By 47 bands SDA outperformed our influential band selection method (82.85% for SDA and 62.48% for our method) but at 51 bands and above our influential band selection method yielded somewhat better results (83.61% for SDA and 84.28% for our method). This difference is small, but remained persistent in case of even more bands.”

**Question 6b:**
Fig. 3 also indicates that, when the numbers of bands exceed some value (50 (proposed) and 46 (SDA)), the classification performances are saturated. If we assume the difference in accuracies between 50 (proposed) and 46 (SDA) is not significant, we could say that SDA could achieve good accuracy with fewer number of bands. Please give more discussion that supports the advantage of the proposed method.

**Answer 6b:**
We don’t really think that the small (1-2%) gap between the results mean that this difference is not significant. If it was random we think that sometimes SDA would produce slightly better results than our method, but in this case both are varying around the accuracy value produced using the optimal number of bands (remaining SDA less accurate in every case than our influential band selection). In conclusion in this case our method managed to find a somewhat better combination, in conclusion SDA ruled out some key wavelengths – probably in the NIR region (e.g. 711.72 nm and 884.7 nm) which could have been important for an even more accurate classification. Since the results are slightly but persistently better, we think that this difference in the band composition is a big enough difference: 5 bands are not used by SDA but used by our method and 1 is used by our method and not by SDA.

We would like to point out that this study is only about mixed vegetation classes. In case of other mixtures (e.g. urban or mineral) the results may be better or worse.

To clarify this we also added the following sentences to the “6. Discussion” section:

“As seen in Table 1. bands corresponding to 711.72 nm, 884.7 nm, 1013.3 nm, 1114.19 nm, 1205.07 nm were selected by our influential band selection method and were not selected...
using SDA while the band corresponding to 742.25 nm was ruled out by influential band selection but it was selected using SDA. In conclusion in this case we managed to find an even better combination of bands for mixed vegetation classes by using more bands in the NIR region and in the starch, protein, oil and H₂O absorption bands (Kumar et al. 2001).

**Other corrections:**

We also corrected some previously unseen typos and added one more band in the 47 bands list (it was accidentally left out after the first revision).

Also one of our author changed his affiliation in the meantime (Mátyás Árvai).
Heterogeneous forest classification by creating mixed vegetation classes using EO-1 Hyperion

Márton Deák

Department of Physical Geography, Eötvös Loránd University, 1117 Pázmány Péter promenade 1/c., Budapest, Hungary, dmarton@elte.hu; corresponding author

Tamás Telbisz

Department of Physical Geography, Eötvös Loránd University, Budapest, Hungary, telbisztom@ludens.elte.hu

Mátyás Árvai

Institute for Soil Sciences and Agricultural Chemistry, Department of Soil Mapping and Environmental Informatics, Hungarian Academy of Sciences, Budapest, Hungary.

arvai.matyas@agrar.mta.hu; matyas.arvai@gmail.com

László Mari

Department of Physical Geography, Eötvös Loránd University, Budapest, Hungary, mari.laci@gmail.com

Ferenc Horváth

Centre for Ecological Research, Hungarian Academy of Sciences, Tihany, Hungary, horvath.ferenc@okologia.mta.hu

Balázs Kohán

Department of Environmental and Landscape Geography, Eötvös Loránd University, Budapest, Hungary, kohan.balazs@gmail.com

Orsolya Szabó
Department of Economics, Partium Christian University, Oradea, Romania, 
orsolya_szabo@yahoo.es

József Kovács

Department of Physical and Applied Geology, Budapest, Hungary, 
kevesolt@geology.elte.hu
Heterogeneous forest classification by creating mixed vegetation classes using EO-1 Hyperion

Hyperspectral satellite data is an efficient tool in vegetation mapping, however previous studies indicate that classifying heterogeneous forests might be difficult. In this study, we propose a mapping method for a heterogeneous forest using the data of NASA’s EO-1 Hyperion supplemented by field survey. We introduced a band reduction method to raise classification accuracy of the Support Vector Machine classification algorithm and compared the results to the PCA-reduced, SDA-reduced and the original dataset. We also used a modified version of the V-I-S model to create mixed vegetation classes consisting of the commonly mixing species in the area and classified them using Decision Tree Classification method. We managed to achieve 84.2% approximately using our band reduction method which is 2.3% increase compared to PCA (81.9%), 1.4% compared to the SDA (82.8%) and 5.5% compared to the original dataset (76.7%). Introducing the mixed vegetation classes raised the overall accuracy even higher (85.8%).

Keywords: hyperspectral data; classification; forestry; band selection

1. Introduction

In case of vegetation classification it is very important to have high spatial resolution and high accuracy while including as many classes as possible. Fire hazard modeling (Thonicke et al. 2010), ecological (Hufkens et al. 2012) or landscape research (Varga et al. 2015), silvi- and agricultural surveys (Martin-Queller et al. 2012, Nelson et al. 2009) and even agricultural forecasts (Ines et al. 2013) all demand the precise knowledge of the vegetation composition in an area. The traditional ways of forest mapping can be time-consuming and often costly and relying only on field survey, data collection may be difficult in some areas (eg. private properties). Official databases are often outdated and might not contain high resolution data.
Remote sensing might be an ideal choice for forest estimation, since it is fast, cost-efficient and data can be collected about large areas at the same time. The application of hyperspectral remote sensing data is a good option. Using this technology species are mostly identified using high spatial resolution hyperspectral data but we can find less papers in the topic of the application of lower spatial resolution hyperspectral satellite images. Studies about the application of high spatial resolution data are often concentrating on the biochemical properties of the plants which are depending connected to on spectral indices or extracted from different parts of the on the full spectra. Studies using high spectral resolution field and laboratory spectrometry are even more chemistry-dependent. These studies are using different techniques and methods utilizing field/laboratory spectrometer, aerial hyperspectral imagery or both (Gong et al. 1997, Williams et al. 2003, Clark et al. 2005, Boitt et al. 2014, Caiyun 2014, Atzberger et al. 2015, Burai et al. 2015, Gong et al. 1997, Wang et al. 2015).

Classification is also possible through multi-data (Dalponte et al. 2008, Caiyun 2014) approaches, but neural network (Gong et al. 1997, Wang et al. 2015) and spectral spatial methods are also tend to be very effective. On the other hand, hyperspectral aerial images are costly and cover only a small area, while spaceborne data is much more cost-efficient and can provide information about a large area at the same time.

NASA’s Earth Observing-1 was one of the first satellites equipped with a hyperspectral sensor, the Hyperion. It was originally designed as an experiment for testing spaceborne hyperspectral sensors but is still active and new methods come up in various topics. Its applications range from volcanology (Abrams et al. 2013) to vegetation analysis and classification (Pu et al. 2005, Mitri and Gitas 2010) and forest monitoring (Goodenough et al. 2002). The application of hyperspectral data usually
produced better results in forest classification than multispectral data (Goodenough et al. 2003, Pu et al. 2012). In earlier studies, confusion was high regarding the deciduous species (George et al. 2014, Dalponte et al. 2012) and most studies aimed at classifying forests consisting of (spectrally) well separable, homogeneous patches dominated by a single species in the upper canopy sector (eg. Thenkabail et al. 2003, Galidaki and Gitas 2015).

The Hyperion’s spatial resolution is 30 meters/pixel making spectral mixture one of the most serious problems. This challenge can tackled by using only a subset of bands which can be used effectively for spectral mixture and sub-pixel analysis. Further classification accuracy also tends to be higher, since the spectral separability of the surface cover types becomes also usually higher. Principal component analysis (PCA), minimum noise fraction (MNF) and stepwise discriminant analysis (SDA) turned out to be the most effective methods and they are the most widely used (eg. Qian and Fowler 2007, Licciardi et al. 2011, Xu et al. 2012, Licciardi and Chanussot 2015). Their application however might not be ideal if the endmembers have high correlation or the goal of the study is to separate a group of spectrally similar endmembers from each other in an otherwise spectrally heterogeneous area (Ifarragueri and Chang 2002, Farrell and Mersereau 2005, Prasad and Bruce 2005).

In our study area, the vegetation is extremely mixed and spectrally similar. Consequently, the band reduction and classification processes needed a slightly different approach. Our goal was to create a band reduction method which has proved to be successful using the Support Vector Machine (SVM) algorithm. We chose SVM because earlier studies demonstrated that it produces better results in forest classification than Spectral Angle Mapper, Maximum Likelihood or other algorithms (Huang et al. 2002, Pal and Mather 2005, Vyas D. et al. 2011, Shao and Lunetta 2012)
and it is. As a verification, we compared our band reduction results to the ones produced by the PCA method and by using the original full-dimensional data. We also designated mixed vegetation classes by using an altered version of the V-I-S (Vegetation-Impervious-Soil) model presented by Ridd (1995). We substituted the urban endmembers with the commonly appearing tree species and created a vegetation map showing mixed classes using sub-pixel information. This method is similar to some spectral umixing methods which can be linear or non-linear (Keshava and Mustard 2002, Keshava 2003), but instead of determining the ratio of every possible endmember in one pixel we only tried to determine if given class is present in given pixel or not.

2. Study area

Our study area is located in the Budai Hills, which are topographically part of relatively low mountains (the highest peak is 559 m) next to Hungary’s capital city, Budapest (Figure 1). The study field is relatively flat and is bordered by the Csillébérc, Széchenyi-hegy and Farkas-árok morphological units. The vegetation is diverse and mainly arboreal.

This site is a member of the Pannonicum Floristic province, within the Pannonicum Floristic province which belongs to the Bakonyicum Floristic district. The macroclimate is continental considerably affected by submediterranean climate, thus characteristically Eurasian, continental and submediterranean elements determine the flora in this region. Ideally, the region’s typical plant community is the Pannonian - Balkanic turkey oak–sessile oak forests, but the study area is a protected territory, thus faster growing tree
species are the dominant ones. In the past decades, the forestry responsible for silviculture only rarely interfered in the natural succession of this woody area, therefore newly planted patches are only sparsely found in the area. It follows that the prime trees (sessile oak - Quercus petraea, common hornbeam - Carpinus betulus) are unable to do substance. Nowadays, the plant community of the area represents a mixture of Tilia (lime-tree), Quercus (oak) and Sorbus (sorb) species. The study area also includes homogeneous plantings of Pinus nigra (European black pine) substance. Fagus sylvatica (Beech) appears in high numbers in some gorges or on slopes with northerly aspects, because the secondary tree species – such as lime-trees and sorbs – can’t grow fast enough where the number of sunny hours is too low.

In our study area the different species are rarely separated in homogeneous groups. Instead, oak species are commonly mixed with black pine, maple and rarely with beech, but we also found maple-beech mixed patches.

3. Data

3.1 Pre-processing

The Hyperion instrument is capable of resolving 220 spectral bands between 400 to 2500 nm. The images are covering a 7.5 km wide and 100 km long area on the surface. Since it was originally designed for aerial applications the signal-to-noise ratio is relatively low and there are many errors in the data (Datt et al. 2003, Khurshid et al. 2006). Most of these can be corrected using pre-processing methods or avoided by using a selection of bands.

Data pre-processing steps were made according to Apan et al. (2004). We corrected striping, marked some spectral ranges as “bad bands” in the atmospheric regions and also ruled out the bands which had a very low signal-to-noise ratio. In the
end, we used only 141 bands. For the image pre-processing and data processing we used EXELIS VIS’s ENVI 5.2 software.

The image was captured on the 14th July 2013 (ID: EO1H1880272013195110KA_SG1_01). The data was Level1Gst and was in radiance units. We selected a mid-summer image to avoid seasonal, temporary spectral differences which are caused by (for example) different vegetation stages so the method can focus more on the constant, physiological properties of the trees. Cloud coverage was low (less than 1%) but it affected some pixels in the study area. The clouds and built-up areas were masked out during the process.

3.2. **Field survey**

Homogeneous and heterogeneous vegetation blocks were digitized using a Trimble Nomad GPS (min. accuracy was 7 m). We selected vegetation blocks partly on species level (European black pine; beech) and partly on genus level (oak: sessile oak and turkey oak; maple: field maple, Norway maple). This way we classified altogether 588 pixels in the field as homogeneous groups and used 314 of them as training samples (their average spectra was used for classification and band reduction) and 274 as control points. We calculated their average spectra, using 38 points for ground vegetation, 114 for pine, 83 for beech, 24 for maple and 55 for oak. We also collected 203 pixels for different mixed scenarios (62 pixels for maple-beech mixed, 5 for oak-beech mixed – which was present only in a very few areas –, 71 for oak-pine mixed and 35 for oak-maple mixed) from which we used 115 as training (to extract the average rule values) and 58 as control.
4. Methods

4.1. Band reduction

Principal component analysis is a multivariate statistical method to orthogonally transform the data and reduce the number of dimensions (bands) by converting the points to uncorrelated data. It is considered as a standard application in many fields and is still the most popular band reduction method in (hyperspectral) remote sensing. One of its drawbacks is, when applied it does not discriminate between the training spectra and the rest. As a result, when two or more training spectra have high similarity related to the rest of the image, the small differences which make them different from each other will likely disappear (Farrell and Merserau 2005, Prasad and Bruce 2005) – which was also the case in this study.

To avoid this, we propose a band reduction method which separates the spectra by selecting only the bands in case of which the minimum difference between the average training samples in one band is larger than average minimum difference value of all the training samples measured in all bands. This method will be done in respect of the bands where the probability of confusion is the highest, eliminating the bands where one or more training samples are more similar, naturally raising the class separability.

First, we calculated the average (radiance) spectra from the training areas and used them as input. Second, we calculated the differences between all of the above mentioned training spectra in each band, then the minimum value was assigned to the corresponding band (this became the minimum intensity difference). Third, we calculated the average of all minimum intensity differences. We called a band influential, if the smallest intensity difference was larger than the average minimum intensity difference calculated using all bands.
For Peer Review Only

The formula seen in Equation (1) describes the method of finding the influential bands \((ib(b))\).

Equation (1)

\[
mid(b) = \min_{i,j,k} \mid I(j,b) - I(k,b)\mid
\]

\[
ib(b) = \begin{cases} 
1, & \text{if } \mid(b) > \frac{1}{n} \sum_{i=1}^{n} \mid(b) \\
0, & \text{if } \mid(b) \leq \frac{1}{n} \sum_{i=1}^{n} \mid(b)
\end{cases}
\]

Where \(b\) is the band number, \(\mid(b)\) is the “minimum intensity difference” at band \(i\), \(j\) and \(k\) are the IDs of the classes – therefore \(I(j,b)\) is the average intensity of class \(j\) at band \(i\). \(ib(b)\) is a Boolean-type variable. Therefore \(j, k \in (1...m)\) and \(i \in (1...n)\) where \(m\) is the number of classes and \(n\) is the number of bands.

Using the method we selected 51 bands which can be seen in Table 1. The original and band reduced spectra are in Figure 2a and 2b.

Figure 2 here

Table 1 here

The selected bands formed three different groups. The first group is between 752.43 nm and 884.7 nm (including 711.72 nm separately) in the NIR clearly corresponding to the chlorophyll reflectance ranges (Curran et al. 1991, Kumar et al. 2001). The second group is between 1013.3 nm and 1114.19 nm (including 993.17 nm separately), which in case of vegetation may be connected to starch, protein, oil and and H\(_2\)O absorption (Kumar et al. 2001). The third group is 1164.68 nm to 1305.96 nm,
which corresponds to lignin, cellulose, starch and H$_2$O content (Serrano et al. 2002, Kumar et al. 2001). Using this method we selected 6 bands which were not present using the SDA reduction and 1 which was only included there. The 6 bands only selected by our method are: 711.72 nm, 884.7 nm, 1013.3 nm, 1114.19 nm, 1194.97 nm and 1205.07 nm and the only band which wasn’t selected by us but was included using SDA was 742.25 nm.

Since the training areas for the groups we used for the average calculation (thus for finding the influential bands as well) not only differ in the species composition but also in many forest physiological properties (eg. closure, height, age), the selected influential bands do not only show the difference between the physiochemical difference between the different tree species, but also the physiological difference of the forest segments. Thus the final training spectra are consisting of the mixture of photosynthetic and non-photosynthetic vegetation parts, soil and reflective properties of the ground vegetation (if present) as well, having relatively large intra-class variability. This variability is expected to be lower if the forest segments have similar age, closure and the ratio of the dominant species is closer to 100% (since then one pixel will correspond to one type of surface cover), but because of the spectral differences between photosynthetic and non-photosynthetic parts it’s expected to be higher at low closure (Guerschman et al. 2015).

4.2. Classification

Support Vector Machine is a supervised classification algorithm which performs well compared to other methods if the dimensionality is high and the training spectra are highly correlated (Pal and Mather 2006). SVM was performed firstly on the full-dimensional, secondly on the PCA-reduced, thirdly on the SDA reduced, and fourthly on the 51 influential bands data.
In order to achieve the highest accuracy with PCA-reduction, we used the first 3 principal components (only the first 3 principal components were selected because increasing the number of PCA bands lowered classification accuracy).

We also compared our band selection and classification results to the bands selected by the Stepwise Discriminant Analysis method as used by George at al. (2014) and Vyas et al. (2011) and described by Green and Caroll (1978). Using this method we selected 47 spectral bands which are included in Table 1 as comparison to our selection method.

The training areas were selected by taking into consideration the vegetation patterns. Both homogeneous and heterogeneous patches were selected. The homogeneous patches were used as training areas for all the classification methods and the heterogeneous ones were utilized later, while introducing mixed vegetation classes.

Our first goal was to classify the image for the dominant species using the rule images (which contained probability values of each pixel corresponding to a class) of the SVM, classifying a pixel into the class with the highest probability (this is the default setting in most image analysis software). In order to achieve the highest accuracy, the grid search method described by Hsu et al. (2010) was used. Using this method the user should try different $\gamma$ and penalty parameter ($C$) values while using the radial basis function (RBF) kernel and constantly cross-validate the results. According to this in the final classification we used radial basis function kernel, gamma value was 0.25 and the penalty parameter was 100. The highest results were achieved by using these settings.

The designated classes corresponded to three deciduous genera (beech, maple, oak) and one coniferous (pine). Considering the typical surface cover types of the area we had to introduce ground vegetation (opening areas without trees) as a new class.
There were also a lot of built-up areas near the study area but as stated earlier they were masked out prior to analysis.

4.3. Assigning mixed vegetation classes

Our second goal was the application of the rule images created during the classification process and the use of them to create mixed vegetation classes. As stated in section 5. Since in the previous step the highest accuracy was achieved by using our influential band reduction method (51 bands, 84.28%), the rule images created during that classification process were further utilized.

We used an alternative version of the V-I-S model, modifying the method of Lu and Weng (2004), who applied subpixel classification for their Indianapolis study area by substituting the urban endmembers (vegetation, soil, impervious) with the dominant species in the study area (beech, pine, maple, oak). We used the same idea as the above mentioned scientists. The basis of this method is to classify the pixels by using the rule image values in a way, that if a pixel reaches a given value of classification probability it can be classified as a mixture of two or more classes. It is also possible to classify each pixel to its dominant class (assigning the one where the rule layers have the highest probability value). Since at the mixed vegetation areas the rule values were similar for several classes and the rule value of the dominant class was only slightly higher, the introduction of mixed classes was more realistic. Using the same idea, we classified the image for the dominant species and classified each pixel regarding the certainty of the membership to one or two classes.

As stated earlier, the commonly mixing species were oak with pine, maple and beech as well as maple with beech, thus we introduced oak-pine, oak-maple, oak-beech and maple-beech classes beside the original pure classes (oak, maple, pine, beech, ground vegetation). Pine-maple and pine-beech mixtures were not present, so we
We excluded them from the further steps. We extracted the mean rule values of each significant (mixing) class according to the field survey data of the mixed training areas. They can be seen in Table 2.

Excluded them from the further steps. We extracted the mean rule values of each significant (mixing) class according to the field survey data of the mixed training areas. They can be seen in Table 2.

Table 2 here

Since the mean values of all significant species in the mixed classes were above 0.25 (as it can be seen in Table 2), we selected this threshold value to create mixed classes. If a pixel reached the above threshold value only in one class it was classified solely into that class. But if the pixel reached it in two (or more) classes, it was classified as a member of all. However, we note that triple (or higher) membership was very uncommon—in these cases, the pixels were finally classified into the two classes with the highest probability values.

This process was implemented by applying a decision tree classification method where the input data sources were the rule images used for classification. This classification method is widely used to aid other classification methods as well as it is used by its own (eg. Friedl and Brodley 1997, Tooke et al. 2009). A pixel \( P \) is classified into class \( a \), if 
\[
P_a = \max \{ P_i | i \in (1,n) \},
\]
where \( P_i \) is the corresponding rule image value of a given pixel in class \( i \), and \( n \) is the number of classes. In the present case the classes were oak, maple, pine, beech and ground vegetation. This way, each pixel is assigned to the dominant class present in the corresponding area. Thereafter, if \( P_x >= 0.25 \), then \( P \) also becomes a member of class \( x \) beside its dominant class. In case a pixel is assigned to more than two classes, only the two highest probability classes are left.

However, pine-maple and pine-beech mixtures in the study area are not present there is
still a possibility of classifying some pixels as such (e.g., at the boundaries of two 
touching homogeneous pine and maple blocks). To avoid this, we introduced the 
constraint: if $P_{\text{pine}} \geq 0.25$, then maple and beech classes were not investigated.

5. Results

Overall accuracy of assigning the dominant species was 76.67% when we used the full-
dimensional (141 bands) data. In case of the PCA reduced dataset, the overall accuracy 
proved to be higher (81.92%) and while using SDA to select a subset of bands (47 
bands) it was even higher (82.85%) but the highest accuracy was achieved when the 
influential bands (51 bands) were selected by our method (84.28%). The introduction of 
mixed vegetation classes raised the accuracy even further to 85.78%. The error matrices 
in Table 3 show the overall accuracy, kappa statistic values and confusion for each 
classification result. We also indicated the user’s accuracy values (User’s accuracy is a 
measure indicating the probability that a pixel is Class A given that the classifier has 
labeled the pixel into Class A – Story and Congalton 1986).

Table 3a and 3b here

Since the accuracy of our method produced only a slight raise in accuracy compared to 
other methods we also made a comparison by using a different number of influential 
and SDA reduced bands to eliminate the probability of random noise affecting the 
results. While the number of bands is determined in both methods we checked the 
overall classification accuracy using 46-56 influential bands (sorted by the minimum 
intensity difference value), and on 42-52 SDA reduced bands (sorted by the individual 
Wilk’s lambda value of each bands as seen in Figure 3. On Figure 3 it is visible,
that in case of both methods a significant raise in accuracy occurred just before the optimal number of bands. It’s also visible that the accuracy value remained relatively constant (with small changes) around value produced by each method’s proposed combination of bands.

*Figure 3 here*

*Figure 4 shows the false colour composite image of the area with the classification results. Altogether, 9 classes were assigned. 5 of them for the dominant species (beech, maple, oak, pine and ground vegetation) and 4 of them, in the last image, for the mixed classes (beech-maple, oak-beech, oak-maple and oak-pine).*

*Figure 4 here*

In each classification image, oak is dominant on the southwestern slopes, while beech and clearings appear commonly on the northeastern slopes, thus the impact of aspect on species distribution is clearly recognizable. European black pine is present only in planted patches or at the margins of the area, where it is commonly mixed with sessile oak.

When we used only the dominant species for classification, the confusion between beech and oak was the highest in the full dimensional and 51 bands data. In the classification based on the full dimensional data 41.67% of the beech control points were misclassified as oak, while in case of the 51 bands data this proportion was only 30.56%. By using the PCA, the results were more accurate regarding the beech (94.2%
was classified correctly), but 59.26% of the maple was also misclassified as oak, making it the least accurate result (25.93%).

The introduction of mixed classes raised the overall accuracy compared to the 51 bands data, but the accuracy values of single-species classes dropped (except for beech). However, the relatively good accuracy (>80%) of mixed classes raised both the kappa and overall accuracy values. The accuracy of maple was fairly low (70.37%), and 22.22% of the maple control points were misclassified as maple-beech.

6. Discussion

We applied the SVM classification method for a small study area with a mixed forest based on the same hyperspectral image using three different sets of bands. First, full dataset (141 bands) was used as input to SVM, second, the PCA reduced spectra were used, and third, we introduced a new band reduction method to select the influential bands based on the training spectra. In this latter case we found 51 influential bands and concluded that these “influential” bands resulted the highest overall accuracy (84.28%). We also introduced mixed vegetation classes based on the V-I-S model and assigned pixels with high rule values in more than one class into mixed classes using a decision tree classifier. Accuracy was the highest (and confusion the lowest) by using this method (85.79%).

We also compared the two optimal number of bands (47 for SDA and 51 for influential band selection). By 47 bands SDA outperformed our influential band selection method (82.85% for SDA and 62.48% for our method) but at 51 bands and above influential band selection yielded somewhat better results (83.61% for SDA and 84.28% for our method). This difference is small, but remained persistent in case of even more bands.
As seen in Table 1, bands corresponding to 711.72 nm, 884.7 nm, 1013.3 nm, 1114.19 nm, 1205.07 nm were selected by our influential band selection method and were not selected using SDA while the band corresponding to 742.25 nm was ruled out by influential band selection but it was selected using SDA. In conclusion in this case we managed to find an even better combination of bands for mixed vegetation classes by using more bands in the NIR region and in the starch, protein, oil and H$_2$O absorption bands (Kumar et al. 2001).

However, the results are good, they still show a certain degree of confusion which can be explained by the high proportion of mixing trees in the field, resulting high rule values for several classes. In case of the PCA reduced dataset, another explanation for confusion might be the phenomenon described earlier: the PCA usually removes the small differences between the spectrally similar classes.

As it can be seen on Figure 34, many of the pixels classified as oak based on the full dataset and based on the PCA, were classified as maple based on the 51 bands data. After the introduction of the mixed classes, most of those pixels were classified as into the mixed oak-maple class. These points were verified via ground-truth points, and we found that the PCA and the full dataset based processing classified the “unstable” pixels (having high probability values in both maple and oak classes) as oak, whereas the 51 influential bands data resulted that these points were classified as maple, but in fact, they are actually mixed areas.

When mixed classes were used, the accuracy was somewhat lower for the homogeneous patches. This might be explained by the presence of other mixing species, which remained unobserved during the field survey. This also might be caused by other individual trees of non-listed species (eg. lime-tree, sorb) as well.
While our results were relatively good compared to most of the other studies, there are even better ones in the literature. For example, George et al. (2014) used SDA to reduce the bands for forest classification. The accuracy of their classification was 82.27% which is somewhat lower than the results produced by our influential band selection method but it is about the same as what we’ve got while using the SDA reduced dataset (82.85%) on our own sample area. Goodenough et al. (2002) produced 92.9% overall accuracy in Canada by using the full dataset and Maximum Likelihood classification (however, they also had easily classifiable, non-vegetation classes that have raised the overall accuracy value). In addition, we’d like to point out that none of the above mentioned studies used mixed classes and the vegetation they studied was fairly homogeneous.

Our relatively good results might also have been affected by different closure parameters. In this case, soil and ground vegetation reflectance had different effects on the different classes.

In general, our conclusion was that both the band reduction method and the application of the V-I-S model in forest classification are suitable for characterizing and classifying mixed vegetation areas.

67. References:


https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf (accessed October 22 2015.)


Shao, Y. and R.S. Lunetta. 2012. “Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited


### Tables and table captions

Table 1. Selected influential bands

<table>
<thead>
<tr>
<th>Band no.</th>
<th>Wavelength (nm)</th>
<th>Band no.</th>
<th>Wavelength (nm)</th>
<th>Band no.</th>
<th>Wavelength (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>711.72</td>
<td>57</td>
<td>884.7</td>
<td>102</td>
<td>1164.68</td>
</tr>
<tr>
<td>40</td>
<td>752.43</td>
<td>78</td>
<td>993.17</td>
<td>103</td>
<td>1174.77</td>
</tr>
<tr>
<td>41</td>
<td>762.6</td>
<td>82</td>
<td>1013.3</td>
<td>104</td>
<td>1184.87</td>
</tr>
<tr>
<td>42</td>
<td>772.78</td>
<td>83</td>
<td>1016.98</td>
<td>105</td>
<td>1194.97</td>
</tr>
<tr>
<td>43</td>
<td>782.95</td>
<td>84</td>
<td>1023.4</td>
<td>106</td>
<td>1205.07</td>
</tr>
<tr>
<td>44</td>
<td>793.13</td>
<td>85</td>
<td>1027.16</td>
<td>107</td>
<td>1215.17</td>
</tr>
<tr>
<td>45</td>
<td>803.3</td>
<td>86</td>
<td>1033.49</td>
<td>108</td>
<td>1225.17</td>
</tr>
<tr>
<td>46</td>
<td>813.48</td>
<td>87</td>
<td>1037.33</td>
<td>109</td>
<td>1235.27</td>
</tr>
<tr>
<td>47</td>
<td>823.65</td>
<td>88</td>
<td>1043.59</td>
<td>110</td>
<td>1245.36</td>
</tr>
<tr>
<td>48</td>
<td>833.83</td>
<td>89</td>
<td>1047.51</td>
<td>111</td>
<td>1255.46</td>
</tr>
<tr>
<td>49</td>
<td>844</td>
<td>90</td>
<td>1053.69</td>
<td>112</td>
<td>1265.56</td>
</tr>
<tr>
<td>50</td>
<td>851.92</td>
<td>91</td>
<td>1057.68</td>
<td>113</td>
<td>1275.66</td>
</tr>
<tr>
<td>51</td>
<td>854.18</td>
<td>92</td>
<td>1063.79</td>
<td>114</td>
<td>1285.76</td>
</tr>
<tr>
<td>52</td>
<td>862.01</td>
<td>93</td>
<td>1073.89</td>
<td>115</td>
<td>1295.86</td>
</tr>
<tr>
<td>53</td>
<td>864.35</td>
<td>94</td>
<td>1083.99</td>
<td>116</td>
<td>1305.96</td>
</tr>
<tr>
<td>54</td>
<td>872.1</td>
<td>95</td>
<td>1094.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>874.53</td>
<td>96</td>
<td>1104.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>882.19</td>
<td>97</td>
<td>1114.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bands selected by SDA (46-47 bands)

<table>
<thead>
<tr>
<th>Band no.</th>
<th>Wavelength (nm)</th>
<th>Band no.</th>
<th>Wavelength (nm)</th>
<th>Band no.</th>
<th>Wavelength (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>742.25</td>
<td>78</td>
<td>993.17</td>
<td>105</td>
<td>1194.97</td>
</tr>
<tr>
<td>40</td>
<td>752.43</td>
<td>83</td>
<td>1016.98</td>
<td>108</td>
<td>1225.17</td>
</tr>
<tr>
<td>41</td>
<td>762.6</td>
<td>84</td>
<td>1023.4</td>
<td>109</td>
<td>1235.27</td>
</tr>
<tr>
<td>42</td>
<td>772.78</td>
<td>85</td>
<td>1027.16</td>
<td>110</td>
<td>1245.36</td>
</tr>
<tr>
<td>43</td>
<td>782.95</td>
<td>86</td>
<td>1033.49</td>
<td>111</td>
<td>1255.46</td>
</tr>
<tr>
<td>44</td>
<td>803.3</td>
<td>87</td>
<td>1037.33</td>
<td>112</td>
<td>1265.56</td>
</tr>
<tr>
<td>45</td>
<td>813.48</td>
<td>88</td>
<td>1043.59</td>
<td>113</td>
<td>1275.66</td>
</tr>
<tr>
<td>46</td>
<td>823.65</td>
<td>89</td>
<td>1047.51</td>
<td>114</td>
<td>1285.76</td>
</tr>
<tr>
<td>47</td>
<td>833.83</td>
<td>90</td>
<td>1053.69</td>
<td>115</td>
<td>1295.86</td>
</tr>
<tr>
<td>48</td>
<td>844</td>
<td>91</td>
<td>1057.68</td>
<td>116</td>
<td>1305.96</td>
</tr>
<tr>
<td>49</td>
<td>851.92</td>
<td>92</td>
<td>1063.79</td>
<td>117</td>
<td>1315.96</td>
</tr>
</tbody>
</table>
Table 2. Mean rule values on the mixed areas

<table>
<thead>
<tr>
<th></th>
<th>Beech-Maple</th>
<th>Oak-Beech</th>
<th>Oak-Maple</th>
<th>Oak-Pine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beech</td>
<td>0.364</td>
<td>0.251</td>
<td>0.081</td>
<td>0.007</td>
</tr>
<tr>
<td>Maple</td>
<td>0.297</td>
<td>0.141</td>
<td>0.334</td>
<td>0.056</td>
</tr>
<tr>
<td>Oak</td>
<td>0.124</td>
<td>0.401</td>
<td>0.424</td>
<td>0.489</td>
</tr>
<tr>
<td>Pine</td>
<td>0.066</td>
<td>0.097</td>
<td>0.117</td>
<td>0.442</td>
</tr>
<tr>
<td>Ground vegetation</td>
<td>0.147</td>
<td>0.112</td>
<td>0.042</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Table 3a. Accuracy and confusion matrices of the classification results for single species classes

<table>
<thead>
<tr>
<th></th>
<th>Full dataset (Kappa Coefficient = 0.6871; Overall accuracy = 76.67%)</th>
<th>PCA (Kappa Coefficient = 0.7611; Overall accuracy = 81.92%)</th>
<th>SDA (Kappa Coefficient = 0.7702; Overall accuracy = 82.85%)</th>
<th>Influential bands (Kappa Coefficient = 0.7890; Overall accuracy = 84.28%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>76.67</td>
<td>3.45</td>
<td>14.81</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>3.33</td>
<td>96.55</td>
<td>3.70</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>6.67</td>
<td>0</td>
<td>55.56</td>
<td>5.68</td>
</tr>
<tr>
<td></td>
<td>13.33</td>
<td>0</td>
<td>3.70</td>
<td>89.77</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>22.22</td>
<td>3.41</td>
</tr>
</tbody>
</table>

User's accuracy
Table 3b. Accuracy and confusion matrix of the classification results for mixed classes

<table>
<thead>
<tr>
<th></th>
<th>Oak</th>
<th>Ground veg.</th>
<th>Maple</th>
<th>Pine</th>
<th>Beech</th>
<th>Oak-Maple</th>
<th>Oak-Pine</th>
<th>Oak-Beech</th>
<th>Beech-Maple</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oak</td>
<td>83.33</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.35</td>
<td>2.41</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Ground veg.</td>
<td>0</td>
<td>82.76</td>
<td>0</td>
<td>0</td>
<td>9.09</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maple</td>
<td>0</td>
<td>0</td>
<td>70.37</td>
<td>0</td>
<td>2.17</td>
<td>1.20</td>
<td>0</td>
<td>1.45</td>
<td>0</td>
</tr>
<tr>
<td>Pine</td>
<td>6.67</td>
<td>0</td>
<td>3.70</td>
<td>88.41</td>
<td>0</td>
<td>3.61</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Beech</td>
<td>3.33</td>
<td>10.34</td>
<td>3.70</td>
<td>0</td>
<td>81.82</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.35</td>
</tr>
<tr>
<td>Oak-Maple</td>
<td>0</td>
<td>0</td>
<td>1.45</td>
<td>0</td>
<td>89.13</td>
<td>4.82</td>
<td>0</td>
<td>4.35</td>
<td>0</td>
</tr>
<tr>
<td>Oak-Pine</td>
<td>6.67</td>
<td>0</td>
<td>0</td>
<td>10.14</td>
<td>0</td>
<td>2.17</td>
<td>86.75</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oak-Beech</td>
<td>0</td>
<td>3.45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maple-Beech</td>
<td>0</td>
<td>3.45</td>
<td>22.22</td>
<td>0</td>
<td>9.09</td>
<td>2.17</td>
<td>1.20</td>
<td>0</td>
<td>89.86</td>
</tr>
<tr>
<td>User’s accuracy</td>
<td>83.33</td>
<td>92.31</td>
<td>86.36</td>
<td>91.04</td>
<td>69.23</td>
<td>83.67</td>
<td>87.8</td>
<td>80</td>
<td>84.93</td>
</tr>
</tbody>
</table>
Figure captions

Figure 1. The study area

Figure 2. Radiance spectra – a: Full dataset; b: 51 selected influential bands

Figure 3. Classification accuracy in case of influential bands selection and SDA reduction in relation to the optimal number of bands according to each method

Figure 4. Classification results – a: Image composite; b: Full dataset; c: PCA reduced; d: SDA reduced, e: 51 influential bands; f: 51 bands, mixed classes
Figure 1. The study area

889x482mm (72 x 72 DPI)
Figure 2. Radiance spectra – a: Full dataset; b: 51 selected influential bands

352x137mm (72 x 72 DPI)
Figure 3. Classification accuracy in case of influential bands selection and SDA reduction in relation to the optimal number of bands according to each method.

430x269mm (72 x 72 DPI)
Figure 4. Classification results – a: Image composite; b: Full dataset; c: PCA reduced; d: SDA reduced, e: 51 influential bands; f: 51 bands, mixed classes

289x228mm (200 x 200 DPI)