Geoinformation Issues
Vol. 9, No 1 (9), 27–38/2017

Mapping tree cover with Sentinel-2 data using the Support Vector Machine (SVM)

Anna Mirończuk
Institute of Geodesy and Cartography, 27 Jacka Kaczmarskiego St., 02-679, Warsaw, Poland
Tel.: +48 22 3291993, Fax: +48 22 3291950, E-mail: anna.mironczuk@igik.edu.pl

Agata Hościło
Institute of Geodesy and Cartography, 27 Jacka Kaczmarskiego St., 02-679, Warsaw, Poland
Tel.: +48 22 3291976, Fax: +48 22 3291950, E-mail: agata.hoscilo@igik.edu.pl

Abstract: The knowledge on forest resources is important for sustainable forest management at local and national level. The aim of this paper is to examine the efficacy of the Support Vector Machine (SVM) approach for tree cover mapping based on Sentinel-2 images and to explore the potential of the Sentinel-2 data for the assessment of tree cover. Sentinel-2 is a constellation of two European satellites providing innovative wide-swath (up to 290 km), high-resolution and multispectral data (13 spectral bands at 10, 20 and 60 m spatial resolution). The study area is located in the Forest Promotion Complex, which is a part of the Knyszyn Forest Landscape Park in Poland. The SVM classification was performed on the single images (spring and summer season) and on multi-date Sentinel-2 images (images from two dates classified simultaneously). In addition, the use of high-resolution bands and a combination of the 10 m and 20 m spatial resolution data was examined. The overall accuracy for all performed classification was very high and reached the level of 96.7%–99.6%, which confirms that SVM classification can be successfully applied for tree cover mapping. The analysis showed that the Sentinel-2 images acquired in the middle of the vegetation season, when the leaves are fully developed are more suitable for tree cover mapping than the images acquired in spring.

Keywords: tree cover, Support Vector Machine, Sentinel-2, classification

Received: 31 January 2018 / Accepted: 10 March 2018

1. Introduction

The knowledge on forest resources is important for sustainable forest management at local and national level and it varies from county to country. The accurate assessment of forest cover area is essential for the country's reporting requirements. Several studies conducted over the Polish forest discussed the disagreement concerning the forest cover in Poland between the official cadastral data and forest status on the ground (Jabłoński, 2015; Hościło et al., 2016). The study conducted by Hościło et al. (2016) revealed that the forest cover is almost 800 thousand hectares larger than the official statistics provided by the Central Statistical Office of Poland. This study was conducted based on the available spatial datasets such as the Digital Forest Map (covering explicitly the State Forests), Topographic Database, Database of Parcel Identification System, National Forest Data Bank and satellite-based High Resolution Layer (HRL – tree cover density) developed within the framework of the Copernicus Land Monitoring programme (https://land.copernicus.eu/pan-european/high-resolution-layers). Forest cover in the first three databases was obtained by the visual interpretation of the national aerial orthophotomap. Mapping forest cover in this way requires regular data acquisition and processing which is costly and time consuming. Therefore, there is an urgent need to come up with an operational methodology for tree cover and forest cover mapping over Poland.

The development of the Earth Observation (EO) technologies as well as the increase in the spatial, spectral and temporal resolution of the satellite images result in the significant growth in the operational services offering EO based products for various applications. For example, Loveland et al.

(2000) developed a 1 km spatial resolution global land cover database (17 general land cover classes) applying unsupervised classification of the monthly Normalized Difference Vegetation Index (NDVI) calculated based on NOAA Advanced Very High Resolution Radiometer (AVHRR) data acquired over the period 1992–1993. Images from the Landsat sensor (spatial resolution of 30 m) have been used to develop the "High-Resolution Global Maps of 21st Century Forest Cover Change". This product consists of the global tree cover extent and annual tree loss and gain between 2000 and 2012 (Hansen et al., 2013). Landsat data have also been used to quantify annual deforestation and degradation across the Brazilian Amazon for the period 2000–2010 (Souza et al., 2013). The authors classified the tree cover using a combination of the spectral mixture analysis, normalized difference fraction index, and knowledge-based decision tree classification, achieving an overall accuracy of 92%. Furthermore, the global forest/non-forest coverage was produced at 25 m spatial resolution using a global mosaic of Phased Array type L-band Synthetic Aperture Radar (ALOS PALSAR) data. This product achieved up to 95% overall agreement with Google Earth images (Shimada et al., 2014).

Some of the existing EO based products, however, require verification and tuning to meet the user requirements. For example, it was estimated that the satellite based Copernicus HRL-tree cover density layer has an overestimation error up to 7.5% (Hościło et al., 2016). In addition, there was a time lag in releasing the Copernicus High Resolution Layers. The HRLs for 2012 were released in 2015, which is too late for the data to be used in national reporting. Some of the existing products refer to tree cover and other to forest cover. The EO based products often refer to the tree cover as it can be detected from space and does not require additional information, for example tree height or minimum forest area size. It has to be stressed that there are around 800 different forest definitions in the World (Archard et al., 2008). Different definitions are required for different purposes and scales. In Poland, the national definition of forest, as defined in the Forest Act of 28 September 1991, refers to a homogeneous tree covered area of minimum 0.1 ha (Ustawa o lasach, 1991).

The aim of this paper is to examine the efficacy of the Support Vector Machine (SVM) approach for tree cover mapping based on Sentinel-2 images and to explore the potential of the Sentinel-2 data for the assessment of tree cover. In recent years, there has been a significant increase in using the SVM method in remote sensing. Mountrakis et al. (2011) summarized empirical results from over 100 articles using the SVM image classification algorithm. Several studies confirmed that the SVM gives higher classification accuracy than the traditional classification method using small training datasets. Shao and Lunetta (2012) compared SVM classification of land-cover using MODIS time-series data to multilayer perceptron neural networks (NN) and regression trees (CART). The SVM accuracy reached from 77% to 80% for training sample sizes from 20 to 800 pixels per class, compared to 67–76% and 62–73% for NN and CART, respectively. Melgani and Bruzzone (2004) performed a detailed comparison of SVM, conventional k-Nearest Neighbour (kNN), and a radial basis neural network using hyperspectral remote sensing data. The results of their studies indicated that SVM substantially outperformed the other two classifiers. Noi and Kappas (2018) examined and compared the performances of Random Forest (RF), kNN, and SVM for land use/cover classification using Sentinel-2 data. They tested 14 different training sample size and confirmed that SVM produced the highest overall accuracy with the least sensitivity to the training sample sizes.

Moreover, the SVM approach was successfully used in classification of urban areas (Warner and Nerry, 2009), land cover /land use classification (Shao and Lunetta, 2012; Adam et al., 2014), crop classification (Mathur and Foody, 2008), impervious surface mapping (Inglada, 2007) and for detection of illegal logging (Kuemmerle et al., 2009).

In this study the tree cover mapping was performed using Sentinel-2 images. Sentinel-2 is an Earth Observation (EO) mission implemented by the European Commission (EC) and European Space Agency (ESA) as part of the Copernicus Programme. This mission is dedicated to global land observation in support of services such as land cover changes monitoring and natural disaster management. Sentinel-2 is a constellation of two European satellites providing innovative wide-swath (up to 290 km), high-resolution and multispectral data (13 spectral bands at 10, 20 and 60 m spatial resolution) (Drusch et al., 2012).

2. Study area

The study area is located in the Forest Promotion Complex, which is a part of the Knyszyn Forest Landscape Park (Fig. 1). The study site covers around 16 800 hectares, of which 75% is forest. The area is dominated by the fresh mixed coniferous forest and fresh mixed broadleaved forest. The dominant tree species in the majority of stands are pine and spruce mixed with birch, alder and oak. The largest area is occupied by forest aged 80–90 and 50–60 years. The average age of the forest stands is around 57 years, which is the effect of the reforestation after the Second World War. The tree species composition results from the continental climate prevailing in north-eastern Poland (short growing season – about 200 days; average annual temperature – 8.5°; precipitation – about 150 days a year) and the soil type formed by the last glaciation (clay, sandy soil and gravel) (Łaziuk, 2014).

3. Data and Methods

3.1. Sentinel-2 data and pre-processing

The Sentinel-2 mission consists of two identical satellites, Sentinel-2A (launched on 23 June 2015) and Sentinel-2B (launched on 7 March 2017) collecting data with high revisit frequency up to 5 days at the equator (faster at higher latitudes). The two satellites operate on opposite sides of the orbit. The orbit is Sun-synchronous at 786 km altitude. The orbit inclination is 98.62° and the Mean Local Solar Time at the descending node is 10:30 am. Sentinel-2 acquires 13 spectral bands (Table 1) with 10 m (4 bands), 20 m (6 bands) and 60 m (3 bands) spatial resolution (Drush et al., 2012).

In this study two cloud-free Sentinel-2 images acquired on 31 August 2015 (late-summer image) and 28 March 2016 (spring image), respectively, were used. The Sentinel-2 A images Level 1C were downloaded from the ESA Copernicus Open Access Data Hub website. All images were corrected for atmospheric, terrain and cirrus effects using the ESA Sen2Core package (Mueller-Wilm et al., 2016). The Sentinel-2 10 m and 20 m spectral bands were

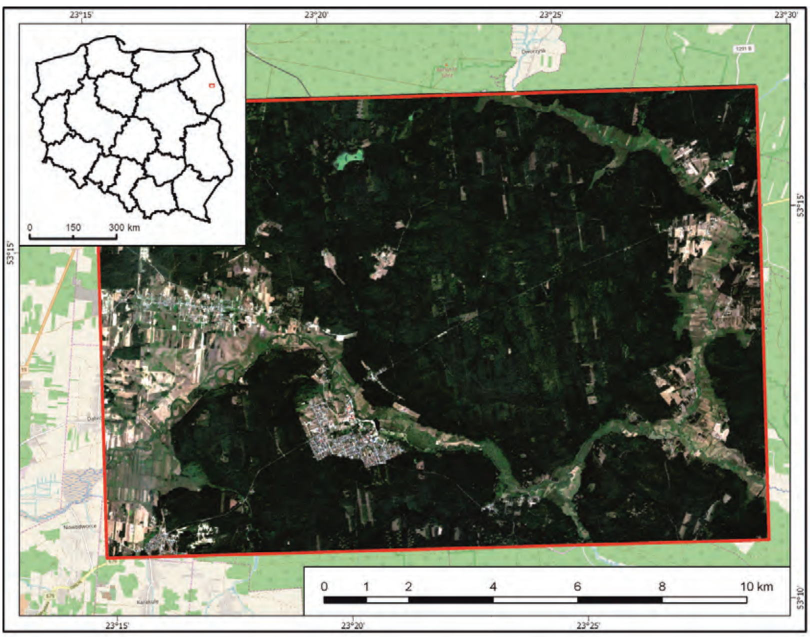

Fig. 1. The location and extent of the study area (red outline) overlaid on the Sentinel-2 image acquired on 31.08.2015, natural colour composite (RGB)

Table 1. Wavelengths and Bandwidths of Sentinel-2 (source: https://sentinel.esa.int/web/sentinel/missions/sentinel-2)

Spatial resolution [m]	Band number		S2A		S2B	
			Central wavelength [nm]	Bandwidth [nm]	Central wavelength [nm]	Bandwidth [nm]
10	2	Blue	496.6	98	492.1	98
	3	Green	560.0	45	559.0	46
	4	Red	664.5	38	665.0	39
	8	NIR	835.1	145	833.0	133
20	5	Vegetation Red Edge	703.9	19	703.8	20
	6		740.2	18	739.1	18
	7		782.5	28	779.7	28
	8a	Narrow NIR	864.8	33	864.0	32
	11	SWIR	1613.7	143	1610.4	141
	12		2202.4	242	2185.7	238
60	1	Coastal aerosol	443.9	27	442.3	45
	9	Water vapour	945.0	26	943.2	27
	10	SWIR (Cirrus)	1373.5	75	1376.9	76

used for further analysis. The open source common architecture for ESA Toolboxes – SNAP (SeNtinel's Application Platform) – was used to subset images, reproject to the national projection and to resample 20 m spatial resolution bands (B5, B6, B7, B8a, B11, B12) to 10 m.

3.2. Forest Cover Map

The Forest Cover Map referring to the national definition of forest as defined in the Forest Act of 28 September 1991, forest by definition is a homogeneous area of 0.1 ha (Ustawa o lasach, 1991), was produced for the reference year 2012. This includes forests under all forms of ownership and areas that are forested, but officially recorded as non-forest (it should be noted that, for example, the clear cuts are classified as forest by the national definition of forest). The assessment of the actual forest cover in Poland was derived based on existing spatial datasets such as the Digital Forest Map (covering explicitly the State Forests), Topographic Database, Database of Parcel Identification System, High Resolution Layer (Copernicus Land Monitoring product based on classification of satellite data) and National Forest Data Bank (Hościło et al., 2016). The weighted raster analysis was applied to derive the Forest Cover Map. A detailed description of the undertaken approach can be found in the paper by Hościło et al. (2016). The minimum mapping unit was equal to 0.1 ha and the minimum mapping width was equal to 10 m.

3.3. Aerial orthophoto

Training samples (described in chapter 3.4) and classification results were verified based on the national aerial orthophoto available through the geoportal.gov.pl in the form of the Web Map Service (WMS).

3.4. Training and validation sample datasets

The Forest Cover Map was used as reference data to select training and validation sampling plots. Applicability of the SVM is only possible when labelled samples of all the land-cover classes present in the scene are available (Munoz-Mari et al., 2010), so two classes: (1) tree and (2) non-tree were examined. It was assumed that samples for training and validation will match to the size and location of pixels of S-2 images (sample = pixel). For this purpose, a net of polygons was created from which the sampling plots for the tree and non-tree classes were selected based on the random sampling method (Fig. 2). In total 240 polygons of 10 × 10 m were selected for tree areas and the same number of polygons for non-tree areas. To avoid geometrical instability of satellite images and to minimize the edge effect, the sampling plots had to be surrounded by at least 3 × 3 pixels (30×30 m) and located within larger polygons. Manual verification was carried out to make sure that all sampling plots are tree-covered. The manual verification of the sampling polygons was applied against the aerial orthophoto. For the non-tree cover class, the training samples were carefully selected and represent various land cover classes i.e. agriculture areas, water, shrubs, buildings, roads, excavation sites.

A benefit of SVM is the ability to successfully handle small training data sets (Mantero et al., 2005). Research by Noi and Kappas (2018) has shown that the accuracy results of SVM were not significantly different among different training sample sizes. According to Mathur and Foody (2008) the accuracy of an SVM classification depends not so much on the size of input training data but more on the location of training data in the feature space. Therefore sampling plots were split equally into the training set (50%) and the validation set (50%).

3.5. Support Vector Machine classification

The tree cover classification was performed using the Support Vector Machine (SVM). The SVM approach was introduced within the Framework of the Statistical Learning Theory developed by Vapnik (Cortes and Vapnik, 1995). The SVM is a supervised non-parametric statistical learning technique (Mountrakis et al., 2011) that was originally designed for binary classification (Mathur and Foody, 2008). The SVM is based on the main hypothesis that the training set is linearly separable. The SVM searches for the optimal line (hyperplane) which separates without errors the training set, and maximises the distance, named the "margin", between the objects of both classes and the hyperplane (Fig. 3). Thus, instead of using the whole available training set to describe classes, SVM uses only those training samples that describe class boundaries (support vectors) (Roli and Fumera, 2001).

The SVM classification and validation of the classification outcomes were performed in the En-MAP BOX version 2.2.1 – freely available, platform-independent software designed to process hyperspectral remote sensing data (van der Linden et al., 2015). The SVM comprises a two-step approach consisting of (1) the parameterization of a Support Vector Classifier (SVC) based on reference data, and (2) the classification of the image data. The Radial Basic Function (RBF) kernel was used to implement the SVM algorithm. Parameterizing SVM in

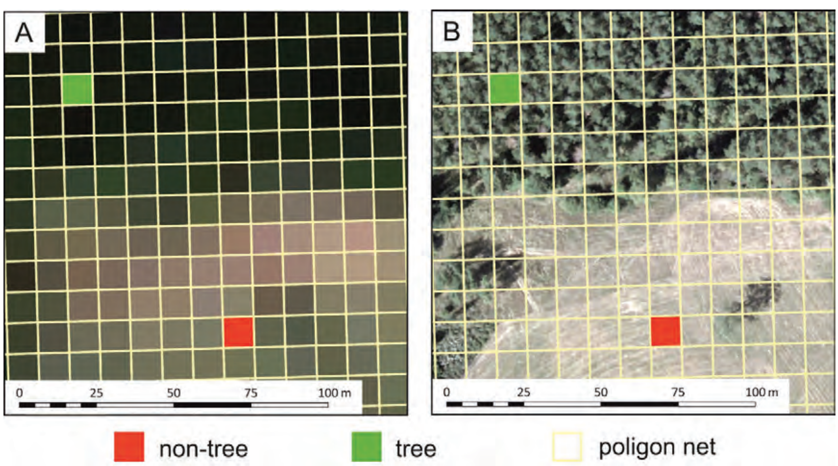

Fig. 2. Example of the sampling plots overlaid on the S-2 image (A) and on the orthophoto (B)

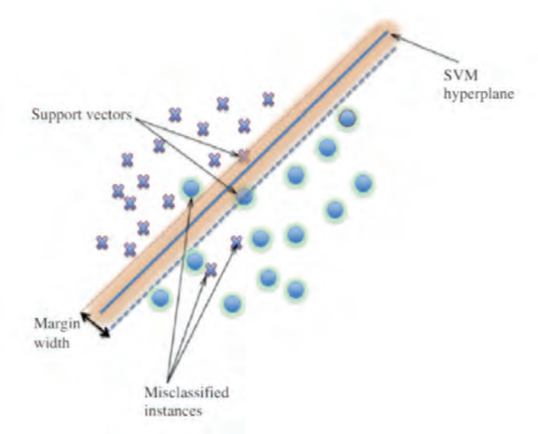

Fig 3. Example of the separation of the classes using the Linear Support Vector Machine approach (source: Mountrakis et al., 2011)

this way requires setting the parameter γ, which defines the width of the Gaussian kernel function and regularizing parameter C, which controls the trade-off between the maximization of the margin between the training data vectors and the decision boundary plus the penalization of training errors (van der Linden et al., 2014). The following parameters were used for the classification γ ranging from 0.01 to 1000 with a linear kernel, and C ranging from 0.01 to 1000 with 3-fold cross validation on the training data.

To assess the potential of the Sentinel-2 data for tree cover mapping, the SVM classification was performed on single images (from spring and summer season separately) and on multi-date Sentinel-2 images (stack of two images classified simultaneously). In addition, the classification was carried out on the high-resolution bands (blue: B2, green: B3, red: B4, and NIR: B8) at 10 m spatial resolution and on a combination of the 10 m and 20 m bands (Vegetation Red Edge: B5, B6, B7, Narrow NIR: B8A and SWIR: B11, B12). Table 2 presents the description of executed SVM classifications.

The accuracy of the SVM classification was assessed based on the 240 validation sampling plots (120 tree and 120 for non-tree), not used in the classification process. The accuracy of each classification was assessed using the same testing set. The accuracy measures are the overall accuracy including the 95% confidential level, Kappa accuracy and classification confusion matrix including a number of correctly and incorrectly classified pixels. In addition, the omission error, which is the share of reference pixels in the class that have been omitted in the classification and the commission error, referring to the percentage of the class pixels in the classification which are falsely classified, were calculated.

Table 2. Number of SVM classifications carried out based on a single image (CL1, CL3, CL5) and combination of S-2 images (CL2, CL4, CL6)

Sentinel-2 dates	Spatial resolution 10 m	Spatial resolution 10 m + 20 m
31.08.2015	CL1	CL2
28.03.2016	CL3	CL4
31.08.2015 + 28.03.2016	CL5	CL6

Based on the results of the accuracy assessment, the best tree cover map was selected and compared to the Forest Cover Map. For these purposes the tree patches less than 0.1 ha had to be eliminated from the tree cover map to meet the national forest definition criteria and to be comparable with the content of the Forest Cover Map.

4. Results and discussion

In general, the overall accuracy for tree/non-tree SVM classifications was very high and reached the level of 96.7%–99.6% and a Kappa accuracy of 93.3%–99.2% (Table 3). The highest overall accuracy and Kappa value was obtained for the stack of Sentinel-2 images (CL5&6) and for the single late-summer image (CL1&2). The best classification was obtained for the stack of 10 m band spatial resolution S-2 images – CL5. By contrast, the lowest overall accuracy (96.7–97.1%) and Kappa accuracy (93.3%–94.1%) was achieved for the single S-2 image acquired in spring (CL3&4).

The analysis of the classification accuracy for separate classes (tree and non-tree) revealed that the tree cover class was classified with the highest accuracy using the single S-2 image acquired in the summer period (CL1) and using the multi-date approach (CL5). The lowest Kappa accuracy values (90.3% and 93.4%) were achieved for the S-2 image

Table 3. Overall accuracy and Kappa accuracy of performed classifications

	S-2 31.08.2015		S-2 28.03.2016		S-2 stack	
S-2 spatial resolution	CL1 10 m	CL2 10m & 20m	CL3 10 m	CL4 10m & 20m	CL5 10 m	CL6 10m & 20m
Overall Accuracy [%]	99.16	98.74	96.65	97.07	99.58	98.74
Kappa Accuracy [%]	98.33	97.49	93.30	94.14	99.16	97.49

acquired in the spring 28.03.2015 (CL3&4). In this case, the commission and omission errors for tree and non-tree classes reached the highest values (Table 3). Comparison of other classification results revealed that in the case of CL1, the commission error of the non-tree class is larger than in the CL 5 case, which means some of the non-tree areas were misclassified as trees in CL1. However, the commission errors for the tree class were higher for the classification based on multi-date images (CL5&6) (Table 4). The omission errors for the tree class were larger for the classification on the single S-2 images acquired in the summer period, whereas the omission errors for the non-tree cover class were larger for the multi-date classification (CL5&6).

The forest/non-forest was also successfully classified by the SVM classifier by Kuemmerle et al. (2009) – overall accuracies reached 97.1%–98.0%. This study was conducted in the Ukrainian Carpathians on the Landsat images in the period from 1980 to 2007. Huang et al. (2008) also used Landsat images and the SVM algorithm for developing an automated solution to forest cover change detection – overall accuracy values were approximately 90%.

The SVM classification on the early spring images achieved the lowest accuracy, which may be due to the misclassification of the broadleaf trees as non-forest areas (underestimation) and misclassification of, for example, wet meadows as forest (Fig. 4). In total, 6 forest verification plots were misclassified as non-tree class. The analysis showed that the S-2 images acquired in the middle of the vegetation season, when the leaves are fully developed, are more suitable for tree cover mapping. The discrepancy in the accuracy of the classification results between the single image (CL1&2) and the combination of two images (CL5&6) may be influenced by the low accuracy of the spring image classification.

The largest inconsistency in the classification results is visible on the edge of the forest and over the young forest (Fig. 5).

The comparison of the classification results shows that the classification performed on the single S-2 image acquired in summer and multi-date images gave the best results. The single image approach is less time-consuming as compared to the multi-date approach. Additionally, the multi-date approach is more sensitive to the geometrical precision of the

Table 4. Performance of classifications

		S-2 31.08.2015		S-2 28.03.2016		S-2 stack	
		CL1 10 m	CL2 10 m & 20 m	CL3 10 m	CL4 10 m & 20 m	CL5 10 m	CL6 10 m & 20 m
Kappa Accuracy [%]	Non-tree	96.71	95.10	96.54	94.94	100.00	100.00
	Tree cover	100.00	100.00	90.28	93.36	98.34	95.10
Commission Error [%]	Non-tree	1.65	2.46	1.74	2.54	0.00	0.00
	Tree cover	0.00	0.00	4.84	3.31	0.83	2.44
Omission Error [%]	Non-tree	0.00	0.00	5.04	3.36	0.84	2.52
	Tree cover	1.67	2.50	1.67	2.50	0.00	0.00

Geoinformation Issues
Vol. 9, No 1 (9), 27–38/2017

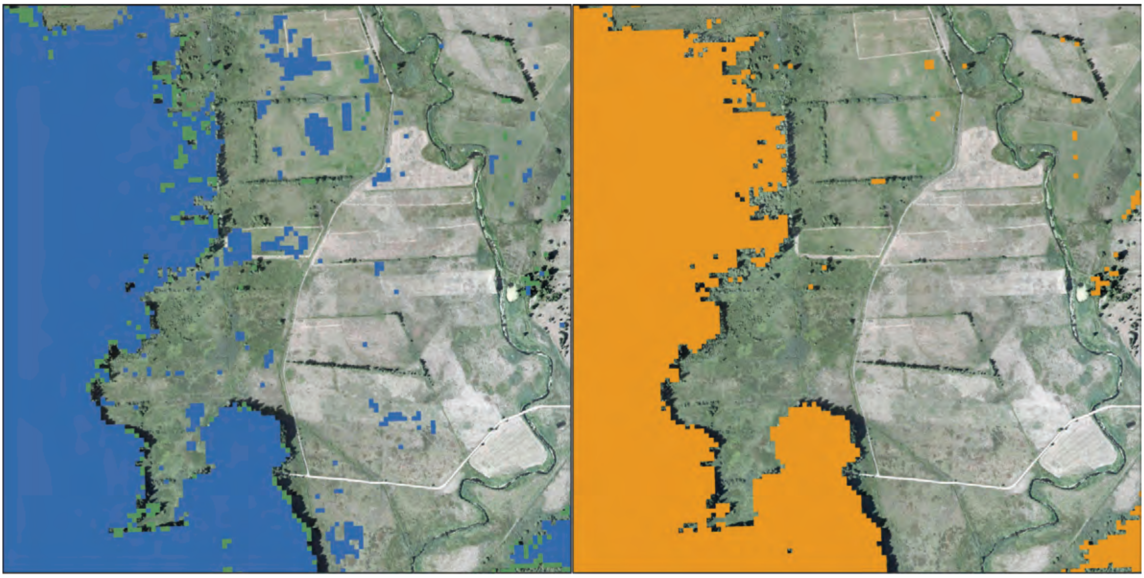

Fig. 4. The results of the tree cover SVM classification on (left) the single S-2 spring image (CL3, blue colour) and (right) on the multi-date S-2 images (CL5, orange colour); note the overestimation of the tree cover CL3 over the meadows and arable land

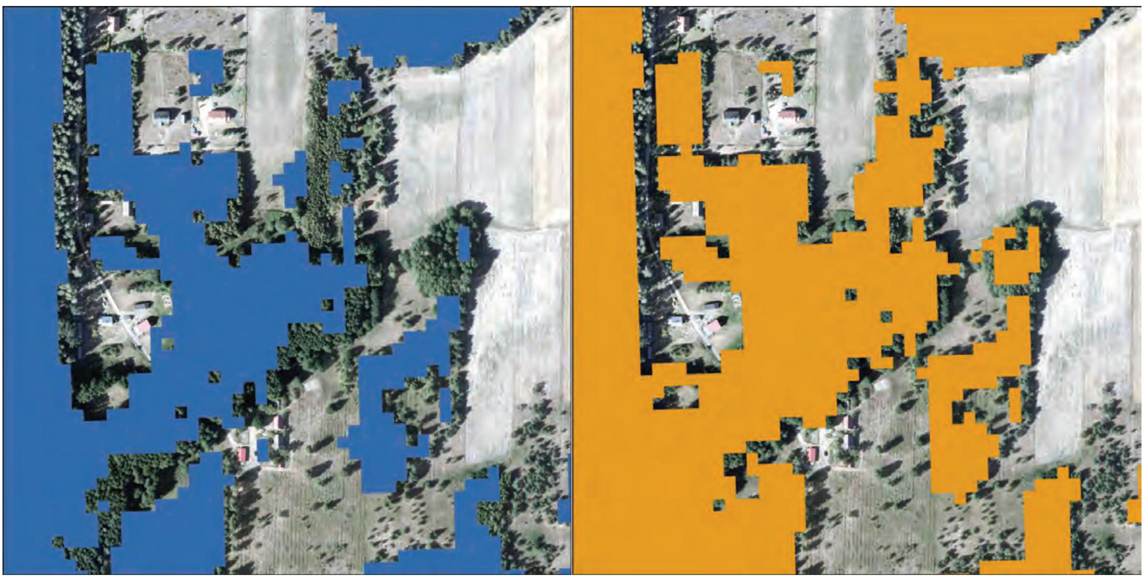

Fig. 5. The results of the tree cover SVM classification on (left) the single S-2 spring image (CL 3, blue colour) and (right) on the multi-date S-2 images (CL 5, orange colour); note the discrepancy on the forest edges

input data and often requires additional co-registration to assure that the images are perfectly overlaid. On the other hand, the single image approach may be less reliable in the case of the more complex forest structure and landscape variability. It should also be stressed that the acquisition date of the satellite image used in the classification of tree coverage is very important.

The classification performed on the S-2 single summer image at 10 m spatial resolution (CL1) gave lower omission error (1.7%) compared to a combination of 10 m & 20 m bands (CL2 – 2.5%). In both cases, however, the Kappa accuracy was very high (100%). The performance of the classification of the multi-resolution data may be improved by the application of the super-resolution method described by Brodu (2017), which allows all bands to be brought from 20 m down to 10 m spatial resolution. In this study, the traditional resampling method was applied.

In the final step, the results of the best SVM classification – CL5 was cross-checked against the Forest Cover Map. To be able to compare both maps, the tree patches less than 0.1 ha had to be eliminated from the SVM tree cover map. This is because the Forest Cover Map refers to the national forest definition. Generally, there was a quite good agreement between the two maps, especially over the large forest patches. The SVM accurately delineated small forest patches and forest on the edges that were omitted in the Forest Cover Map (Fig. 6). The underestimation of forest areas in the SVM map predominantly occurred within the larger forest patches, which is mainly associated with the clearcuts and open areas that are by national forest definition classified as forest but are currently not covered by trees.

The classification of remotely sensed data provides information on the tree coverage regardless of the official national definition, thus further post-classifications analysis is required to come up with the extent of the national forest. Nevertheless, the tree cover mapping using a time series of Sentinel-2 images can be both more accurate and efficient than the conventional methods based on the visual interpretation of the aerial orthophoto or derived statistically from field sampling. Secondly, satellite data enable more frequent updates of the extent of tree cover and monitoring of the tree cover changes over a larger area. In addition to that, the fully automated classification of the satellite images allows reliable products to be produced in a short time.

5. Conclusions

The overall accuracy for all classification performed was very high and reached above 96%, which confirms that it can be successfully applied for tree cover mapping. The results of this study confirmed that the classification performed explicitly on the 10 m spectral resolution bands gave slightly

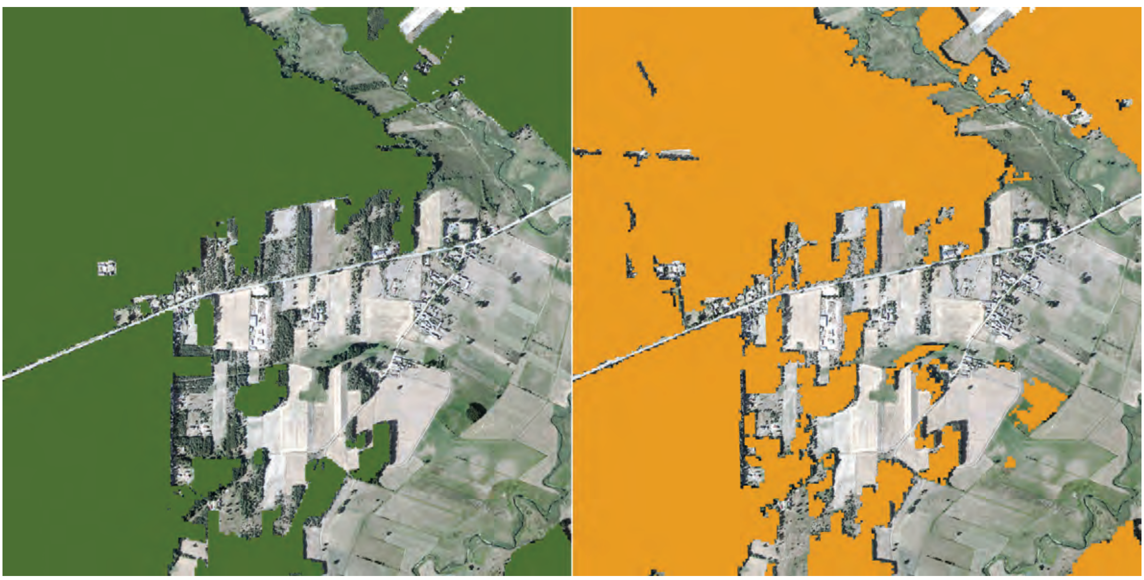

Fig. 6. Forest Cover Map (green colour) and the SVM forest map (orange colour)- result of the CL5

better accuracy than the classification based on a combination of the 10 and 20 m spatial resolution bands. This could be due to the applied resampling method. The performance of the classification of the multi-resolution data may be improved by the application of the super-resolution method described by Brodu (2017), which allows all bands to be brought from 20 m down to 10 m spatial resolution. Further work is required to address this issue. The analysis showed that the Sentinel-2 images acquired in the middle of the vegetation season, with a presence of leaves are more suitable for the tree cover mapping. The vegetation phenology is crucial not only for tree cover mapping but also for tree type analysis. The next step of this research will be: i) to test the SVM methods over a larger area using different time series of Sentinel-2 data and to investigate the synergy of the Sentinel-2 and Sentinel-1 data, and ii) to investigate the performance of different machine learning approaches for tree cover mapping.

Acknowledgements

The research was funded by the Ministry of Science and Higher Education through the Award for Young Scientists.

The Forest Cover Map was produced by the Institute of Geodesy and Cartography within the framework of the project financed by the General Directorate of The State Forest.

The SVM classification was performed using the EnMAP software developed by Humboldt Universität zu Berlin, Prof. Hostert, Humboldt Innovation GmbH, by an order of Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences. For more information see http://www.gfz-potsdam.de.

The authors thank the anonymous reviewers for useful and constructive comments.

References

Adam E., Mutanga O., Odindi J., Abdel-Rahman E., (2014): *Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers*, International Journal of Remote Sensing, Vol. 35, No 10, pp. 3440–3458.

Archard F., et al., (2008): *Forest definition and extent*, Vital Forest Graphic: Stopping the downswing?, The United Nations Environment Programme, The Food and Agricultural Organization of the United Nations & The United Nations Forum on Forests Secretariat. Online at: https://www.cbd.int/doc/meetings/for/wscb-fbdcc-01/other/wscb-fbdcc-01-oth-13-en.pdf

Brodu N., (2017): *Super-Resolving Multiresolution Images with Band-Independent Geometry of Multispectral Pixels*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 55, No 8, pp. 4610–4617.

Cortes C., Vapnik V., (1995): *Support-vector network. Machine Learning*, Vol. 20, pp. 273–297.

Drusch M., Del Bello U., Carlier S., Colin O., Fernandez V., Gascon F., Hoersch B., Isola C., Laberinti P., Martimort P., Meygret A., Spoto F., Sy O., Marchese F., Bargellini P., (2012): *Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services*, Remote Sensing of Environment, Vol. 120, pp. 25–36.

Hansen M., Potapov P., Moore R., Hancher M., Turubanova S.A., Tyukavina A., Thau D., Stehman S.V., Goetz S., Loveland T., Kommareddy A., Egorov A., Chini L., Justice C., Townshend J., (2013): *High-Resolution Global Maps of 21st Century Forest Cover Change*, Science, Vol. 342, No 6160, pp. 850–853. Online at: http://earthenginepartners.appspot.com/science-2013-global-forest.

Hościło A., Mirończuk A., Lewandowska A., (2016): *Określenie rzeczywistej powierzchni lasów w Polsce na podstawnie dostępnych danych przestrzennych*, Sylwan, Vol. 160, No 8, pp. 627–634.

Huang C., Song K., Kim S., Townshend J., Davis P., Masek J., Goward S., (2008): *Use of a dark object concept and support vector machines to automate forest cover change analysis*, Remote Sensing of Environment, Vol. 112, No 3, pp. 970–985.

Inglada J., (2007): *Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features*, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 62, No 3, pp. 236–248.

Jabłoński M., (2015): *Powierzchnia gruntów leśnych – przyczyny zmian i spójność źródeł* *danych*, Wiadomości Statystyczne, Vol. 60, No 11, pp. 54–68.

Kuemmerle T., Chaskovskyy T., Knorn J., Radeloff V., Kruhlov I., Keeton W., Hostert P., (2009): *Forest cover change and illegal logging in the Ukrainian Carpathians in the transition period from 1988 to 2007,* Remote Sensing of Environment, Vol. 113, No 6, pp. 1194–1207.

van der Linden S., Rabe A., Held M., Wirth F., Suess S., Okujeni A., Hostert P., (2014): *imageSVM Classification, Manual for Application: imageSVM version 3.0*, Humboldt-Universität zu Berlin, Germany.

van der Linden S., Rabe A., Held M., Jakimow B., Leitão P.J., Okujeni A., Schwieder M., Suess S., Hostert P., (2015): *The EnMAP-Box—A Toolbox and Application Programming Interface for EnMAP Data Processing*, Remote Sensing, Vol. 7, No 5, pp. 11249–11266. Online at: http://www.enmap.org/?q=enmapbox

Loveland T., Reed B., Brown J., Ohlen D., Zhu Z., Yang L., Merchant J., (2000): *Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data*, International Journal of Remote Sensing, Vol. 21, No 6–7, pp. 1303–1330.

Łaziuk K., (2014): *Zasoby leśne*, Online at: http://www.suprasl.bialystok.lasy.gov.pl/zasoby-lesne/-/asset_publisher/1M8a/content/zasoby-lesne/pop_up?_101_INSTANCE_1M8a_viewMode=print

Mantero P., Moser G., Serpico S.B., (2005): *Partially supervised classification of remote sensing images through SVM-based probability density estimation,* IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, No 3, pp. 559–570.

Mathur A., Foody G., (2008): *Crop classification by a support vector machine with intelligently selected training data for an operational application,* International Journal of Remote Sensing, Vol. 29, No 8, pp. 2227–2240.

Melgani F., Bruzzone L., (2004): *Classification of hyperspectral remote sensing images with support vector machines,* IEEE Transactions on Geoscience and Remote Sensing, Vol. 42, No 8, pp. 1778–1790.

Mountrakis G., Im J., Ogole C., (2011): *Support vector machines in remote sensing: A review,* ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 66, pp. 247–259.

Mueller-Wilm U., Devignot O., Pessiot L., (2016): *S2 MPC Sen2Cor Configuration and User Manual.*

Noi P.T., Kappas M., (2018): *Comparison of Random Forest, k-Nearest Neighbor, and Supper Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery,* Sensors, Vol. 18, No 1.

Roli F., Fumera G., (2001): *Support Vector Machines for Remote-Sensing Image Classification*, Proceedings of SPIE – The International Society for Optical Engineering, Vol. 4170, pp. 160–166. Online at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.5830&rep=rep1&type=pdf

Shao Y., Lunetta R., (2012): *Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points,* ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 70, pp. 78–87.

Shimada M., Takuya I., Motohka T., Watanabe M., Tomohiro S., Thapa R., Lucas R., (2014): *New global forest/non-forest maps from ALOS PALSAR data (2007–2010),* Remote Sensing of Environment, Vol. 155, pp. 13–31.

Souza Jr., Siqueira J., Sales M., Fonseca A., Ribeiro J., Numata I., Cochrane M., Barber C., Roberts D., Barlow J., (2013): *Ten-Year Landsat Classification of Deforestation and Forest Degradation in the Brazilian Amazon,* Remote Sensing, Vol. 5, No 11, pp. 5493–551.

Ustawa o lasach z dnia 28 września 1991r. Dz. U. nr 101, poz. 444 z późn. zm.

Warner T.A., Nerr F., (2009): *Does single broadband or multispectral thermal data add information for classification of visible, near-and short wave infrared imagery of urban areas?,* International Journal of Remote Sensing, Vol. 30, No 9, pp. 2155–2171.

WWW References

https://land.copernicus.eu/pan-european/high-resolution-layers; accessed: 08.2017

https://sentinel.esa.int/web/sentinel/missions/sentinel-2 ; accessed: 01.2018

Geoinformation Issues
Vol. 9, No 1 (9), 27–38/2017

Kartowanie terenów zadrzewionych z wykorzystaniem metody Support Vector Machine (SVM) na podstawie danych z Sentinel-2

Anna Mirończuk
Instytut Geodezji i Kartografii, ul. Jacka Kaczmarskiego 27, PL 02-679 Warszawa
Tel.: +48 22 3291993, Fax: +48 22 3291950, E-mail: anna.mironczuk@igik.edu.pl

Agata Hościło
Instytut Geodezji i Kartografii, ul. Jacka Kaczmarskiego 27, PL 02-679 Warszawa
Tel.: +48 22 3291976, Fax: +48 22 3291950, E-mail: agata.hoscilo@igik.edu.pl

Streszczenie: Wiedza na temat terenów zadrzewionych jest istotna zarówno ze względu na zarządzanie lasami, jak i z punktu widzenia poprawności raportowania danych na potrzeby krajowych i międzynarodowych statystyk. Zobrazowania satelitarne są wykorzystywane do określania zasięgu terenów zadrzewionych, szacowania aktualnego stanu zdrowotnego lasów oraz do ciągłego monitorowania zmian zachodzących w lasach.

Głównym celem badań była analiza możliwości wykorzystania metody wektorów nośnych Support Vector Machine (SVM) do kartowania powierzchni zadrzewionej na podstawie zobrazowań z europejskiego satelity Sentinel-2. Misja Sentinel-2 to konstelacja dwóch satelitów: Sentinel-2A i Sentinel-2B, rejestrujących promieniowanie w zakresie optycznym, bliskiej i dalszej podczerwieni. Największym atutem misji Sentinel-2 jest skrócony czas rewizyty (ok. 5 dni), szeroki pas obrazowania (290 km) oraz zwiększona rozdzielczość przestrzenna (10 m, 20 m i 60 m). Teren badań zlokalizowany był na terenie Leśnego Kompleksu Promocyjnego Puszczy Knyszyńskiej. W celu określenia przydatności zobrazowań Sentinel-2 do kartowania terenów zadrzewionych analizy wykonano na pojedynczych zobrazowaniach S-2 zarejestrowanych wczesną wiosną (28.03.2016) i latem (31.08.2015) oraz na kombinacji danych wieloczasowych, pochodzących z dwóch dat. Dodatkowo testowano wpływ liczby kanałów spektralnych na wynik klasyfikacji. W tym celu wykonano klasyfikację na czterech 10 m kanałach spektralnych oraz na kombinacji 10 m i 20 m kanałach spektralnych.

Wyniki przeprowadzonych badań potwierdziły potencjał metody SVM do kartowania terenów zadrzewionych. W każdym przypadku całkowita dokładność wykonanych klasyfikacji osiągnęła wartość powyżej 96%. Największą dokładność osiągnięto w przypadku klasyfikacji obrazu letniego (dokładność całkowita 99.2%, Kappa 98.3%), zaś najniższą w przypadku obrazu wiosennego (dokładność całkowita 96.6, Kappa 93.3%). Wyniki klasyfikacji wykonanej na pojedynczym obrazie S-2 były nieco lepsze niż na wieloczasowych obrazach.

Słowa kluczowe: tereny zadrzewione, Support Vector Machine, Sentinel-2, klasyfikacja