Tree Cover Mapping Using Hybrid Fuzzy *c*-Means Method and Multispectral Satellite Images

LINDA GULBE[1]*, ALEKSANDRS KOZLOVS[1], JĀNIS DONIS[2] AND AGRIS TRAŠKOVS[1]

[1] *Ventspils University of Applied Sciences, Inzenieru street 101, Ventspils, LV-3601, Latvia*

[2] *Latvian State Forest Research Institute "Silava", Rigas street 111, Salaspils LV-2169, Latvia*

**Corresponding author; e-mail: linda.gulbe@venta.lv*

Gulbe, L., Kozlovs, A., Donis, J. and Traškovs, A. 2019. Tree Cover Mapping Using Hybrid Fuzzy *c*-Means Method and Multispectral Satellite Images. *Baltic Forestry* 25(1): 113–123.

Abstract

Countrywide up-to-date tree cover maps provide valuable information for planning and management purposes to investigate location of the resources and to identify afforestation and deforestation patterns. Landsat programme offers freely available satellite data with time span more than three decades and it can serve as bases for tree cover map calculation using satellite image classification; however, practical use of classification methods is limited due to lack of user-friendly solutions and complex interpretation of the results. The objective of this study is to evaluate user-friendly hybrid classification scheme for tree cover mapping in Latvia and to explore the nature of the spectral classes and consistency of the results when methodology is applied to images of different dates. Tree cover in this context means the area covered by crown of the tree, which may or may not be considered as forest according to local provisions. Tree cover is estimated using unsupervised fuzzy c-means methods with the stability check to ensure the presence of the same spectral classes in independent tests. Spectral classes are classified into two categories: tree cover and other by employing k-nearest neighbours. Such approach does not require high quality sample data and does not include user defined internal parameters of the algorithms (however, they can be specified if needed). The best overall accuracy achieved for year 2014 was 94.2% with producer's accuracy 98.7% (tree cover), 90.5% (other land cover), user's accuracy 90.0% (tree cover), 98.8% (other land cover) and kappa 0.89. Consistency studies showed high impact (within 10% of overall accuracy) of unique conditions during the image acquisition. Some of the spectral classes represent borderline case between relatively dense tree cover and other land cover types like sparse young stands. Those cases are the main threat to the consistency between the results of different dates and seasons.

Keywords: Landsat and Sentinel satellite images, tree cover, fuzzy c-means, k-nearest neighbours, spectral classes, consistency of the results

Introduction

Forest is one of the most important natural resources to ensure economic needs, well-being of citizens and stability of the climate. Countrywide up-to-date tree cover maps provide valuable information for planning and management purposes to investigate location of the resources and to identify afforestation and deforestation patterns (Franklin 2001). Tree cover in this context means the area covered by crown of the tree, which may or may not be considered as forest according to local provisions. Landsat program offers free of charge satellite images for more than three decades ensuring the most detailed spatial and temporal coverage of land cover observations. Sentinel-2 Earth observation mission by European Space Agency is another free of charge multispectral data provider. Digital format promotes computer processing of those images and preparation of thematic maps for different environmental control purposes (Hansen and Loveland 2012). Meanwhile, despite the long years of the studies of medium spatial resolution satellite data classification, a practical use of satellite images in many fields delays mostly due to user unfriendly solutions (Hansen and Loveland 2012) and complex interpretation of the classification results and accuracy assessment (De Leeuw et al. 2010, Nagendra et al. 2013).

Motivation for this study is based on the afforestation problem in the agriculture lands in Latvia. Checking of the databases of record keeping organizations for agriculture land registry showed differences in the total area of agriculture lands within 25% (Pilvere 2012). Total amount of abandoned agricultural lands was identified as 147.6 thousand hectares from which 6.2 thousand hectares were found as afforested (Pilvere 2012).

Remote sensing tools have a potential to provide accurate tree cover maps for specific date enabling to investigate current forest coverage and to identify

changes in the forest areas like deforestation, forest degradation, forest fragmentation, forest afforestation and forest regeneration (Sloan and Sayer 2015). Comparison between existent forest area databases and remote sensing-based tree cover maps could allow efficient identification of non-registered forest areas.

Emphasis in this study is put on tree cover classification instead of term "forest classification".

Forest definition usually is stated in the local provisions. According to the Law on Forests in Latvia "forest" is an ecosystem in all its stages of the development, where the main producers of the organic matter are trees. Minimum height of the trees has to be at least 5 meters and the present or potential tree crown projection on the ground has to be at least 20% of the whole area of the forest stand. (Law of Forests 2000)

Afforested agricultural lands often do not contain tree stands (for example, several trees growing in bushes or trees being smaller than 5 meters) according to this definition, so, instead of mapping forest according to the definition, tree cover mapping is more crucial. Although the computer methods are the same for both tasks, differences occur in preparing sample and test data sets.

Many Landsat imagery studies include forest area identification in the general land use or cover classification scheme besides agricultural lands, urban areas, water and other classes present in current geographical area (Wilkinson 2005).

Statistical machine learning methods like k-Nearest Neighbours (kNN) (Haapanen et al. 2004), maximum likelihood (Schulz et al. 2010, Srivastava et al. 2012), support vector machine (SVM) (Pal and Mather 2005, Mountrakis et al. 2011) and artificial neural networks (ANN) (Civco 1993, Kavzoglu and Mather 2003) are employed for classification of the pixels in the land cover classes.

Review article of Holmgren and Thuresson (1998) summarized classification accuracy from different studies of forest classes starting from young stand and ending with mature stands. They concluded that typical overall accuracy is in the range from 65% - 85%, but inclusion of other land cover types helped to improve overall accuracy above 90%. Study of Srivastava et al. (2012) compared maximum likelihood classifier, SVM and artificial neural networks and authors concluded that application of ANN results in only 2% accuracy increment.

Some studies have tested unsupervised classification to reduce the importance of user selected sample data (Mather 2004, Fan et al. 2009).

Other type of methods for tree cover identification is based on spectral indices and image thresholding. Ye et al. (2014) proposed special spectral index FI (Forest Index) for separating the forested areas by simple thresholding. They reported overall accuracy for two test sites in the USA and Canada (broad leaf forest, needle leaf) 96.2% (kappa 0.911) and 97.8% (kappa 0.954) using 500 random test points labelled through visual interpretation.

There also have been wide scale studies of forest cover mapping. Potapov et.al (2015) studied forest cover dynamics (annual forest cover loss and decadal forest cover gain) from 1985 to 2012 in Eastern Europe including Latvia, while researchers of University of Maryland prepared world-wide estimates of tree cover (Hansen et al. 2013).

Review article of Wilkinson (Wilkinson 2005) summarizes more than 500 land cover and land use classification experiments using remote sensing data during 15 years starting from 1990. It was concluded that during 15 years classification accuracy has not significantly improved and average values for overall accuracy are 76.2% with standard deviation 15.59% and for kappa 0.656 with standard deviation 0.198. Typically, land cover classification scheme includes 8 land cover classes in average with standard deviation 4.6 and employs 7.85 features with standard deviation 11.54. It was concluded that spatial resolution does not impact classification accuracy significantly since the high spatial resolution sensors usually possess low spectral resolution (Wilkinson 2005).

Our interests were more related to convenient mapping of tree cover for local scales to promote use of remote sensing in daily work of record keeping and controlling institutions. Therefore, emphasis in requirements for the methods was set on convenient use of the algorithms without a lot of internal parameters and strong requirements for sample data quality.

Hybrid classification approaches aims to combine the advantages of supervised and unsupervised classification. Unsupervised classification typically is employed for optimized preparation of input data for supervised classification: training data clustering, training data selection from spectrally homogeneous areas and image stratification before supervised methods (Kuemmerle et al. 2006). The main reason for the use of unsupervised methods only in the training data preparation phase is a time demand of clustering algorithms (Richards and Jia 2006).

Results of the unsupervised methods depend on the initialization of cluster parameters and challenge is a relationship between spectral classes and land cover classes expected. Robustness of clustering in studies has been achieved by methods based on combining the results of different algorithms and different parameter initializations (Banerjee et al. 2015). Fan et al. (2009) proposed single point iterative weighted fuzzy c-means clus-

tering algorithm which incorporates weighting of data attributes to adjust clustering results to a specific application. Yu et.al (2008) evaluated inclusion of different metrics in FCM and concluded that for remote sensing images ability to discriminate clusters of different shape and size improves for approximately 5%.

The objective of this study is to evaluate user-friendly hybrid classification workflow for tree cover mapping in Latvia and to find answers to the following research questions by the means of case studies:

RQ1: how is stability of unsupervised classification affected by random initialization and number of spectral classes?

RQ2: how are the results of the unsupervised classification affected by unique conditions (like atmosphere, illumination, and season) during the image acquisition?

RQ3: what are the relations between forest spectral classes and forest inventory parameters?

We evaluated a very simple approach: an image is classified using unsupervised fuzzy c-means (FCM) to find the natural spectral groupings of pixels (spectral classes) and spectral classes are labelled and merged in two land cover classes (tree cover, other) using supervised k-nearest neighbours. Unsupervised approach was chosen to explore the actual image content avoiding pressing classification by sample data, which could be subjective and inaccurate. Fuzzy classification was chosen to enable subpixel mapping; however, subpixel mapping has been considered as a complex method due to spectral interactions between neighbouring pixel and non-linear combinations of reflectance signatures (Townshend et. al. 2000). Several case studies were performed to find the relations between spectral classes and sample data to draw conclusions about what exactly is "seen" in the image by statistical machine learning algorithms.

Material and methods

Study area

The study area includes four administrative regions in the central part of Latvia: Salaspils, Ropazu, Stopinu and Ikskiles regions, located between 24.19° E and 25.11° E, 56.6° N and 57.1° N. Total area of these regions is 864.7 km^2 of which approximately 42% are covered by forests, but 33% consists of farmlands according to the ancillary data available for this study. The study area covers 350.2 km^2 of forest owned by the State of Latvia. Dominant tree species in the study area are Scots pine (*Pinus sylvestris* L., dominant in 61% of total area of State forests), Norway spruce (*Picea abies* L., dominant in 13%), birch (*Betula pendula* Roth, dominant in 13%) according to the Regular Forest Inventory data. Agricultural lands typically are afforested with Norway spruce and birch. Mostly all tree species are present in the mixed stands. The landscape is a mosaic of agricultural lands, forests, swamps, urban and water areas.

Remote sensing data

Landsat 8 OLI, Landsat 7 ETM+, Landsat 5 TM and Sentinel-2A images were downloaded for the tree cover identification tasks. Table 1 shows the list of the images employed in this study and the main characteristics of the image acquisition. Images were reprojected to national map projection LKS-1992 Latvia TM (EPSG code: 3059) and geometric accuracy was assessed by visual inspection using uniformly distributed road crossings selected in orthophotomaps. Orthorectification by using second order polynomial transformation, nearest neighbour resampling and 50 ground control points precisely located in orthophotomaps were performed if maximal geometrical shift was larger than two satellite image pixels. No additional pre-processing operations were performed to keep the methodology as simple as possible.

RGB orthophotomaps were acquired from Latvian Geospatial Information Agency to enable visual assessment of the study site. Aerial photos forming orthophotomaps were in 0.4 m spatial resolution and image acquisition was organized in June of 2013.

Table 1. Landsat and Sentinel-2A satellite images employed in the study and meta data information

Landsat scene ID	Acquisition date (dd.mm.yyyy)	Solar elevation (°)	Cloud cover (%)
LC81870202014023LGN00	23.01.14.	12.2	5.87
LC81870202014055LGN00	24.02.14.	22.0	17.83
LC81870202014087LGN00	28.03.14	34.55	0.01
LC81870202014135LGN00	15.05.14.	50.5	1.26
LC81870202014215LGN00	03.08.14.	48.7	0.17
LC81870202014247LGN00	04.09.14.	38.9	0.7
LC81870202014263LGN00	20.09.14.	33.1	5.37
LC81870202015074LGN00	15.03.15.	29.2	14.38
LC81870202015186LGN00	05.07.15.	53.8	36.4
LE71870202002238KIS00	26.08.02.	41.2	0
LT51870202000241FUI00	28.08.00.	39.8	0
LE71870202000137EDC00	16.05.00.	50.3	0
LT51870201988288KIS00	14.10.88.	22.8	9
LT51870201987285XXX03	12.10.87.	23.7	0
S20150824T094301 (part of the full ID)	24.08.15.	43.3	0
S20170316T094021 (part of the full ID)	16.03.17.	30.7	0

Field data

The field measurements were carried out for naturally afforested former agriculture lands. Clusters of sample plots in the test area were established subjectively based on visual comparison of orthophoto images 1994-1999 with images of 2013.

We established 4 sample plots within each cluster. Centres of sample plots were placed by visual assessment in way to characterise particular conditions in overgrown area and plots were chosen to be as homogeneous as possible with respect to the forest inventory vari-

ables. We took at least 100 GPS readings to get coordinates (LKS-1992 Latvia TM) of sample plot centre. For each sample plot we gave description of forest type.

Tree layer was described in concentric circular sample plots, where trees were measured depending on breast height diameter (DBH): more than 14.0 cm in 500 m^2, 6.1-14.0 cm in 100 m^2, and 2.1-6.0 cm DBH in 25 m^2 sample plots. Stems of smaller trees and bushes were counted in 60 m^2 large sample plot. For trees with DBH more than 2.1 cm, height was measured for 3 trees for each species in the sample plot.

Totally, trees were measured in 280 sample plots: 70 clusters in different locations with 4 sample plots for each cluster. Field measurements showed that former agricultural lands are afforested by birch and spruce.

Ancillary data

Since field measurements were performed only for the naturally afforested agriculture lands (only young stands), the regular forest inventory (RFI) data base from State Forest Service was employed as additional data source to understand satellite image classification results in more details with respect to the forest inventory variables. The RFI contains measurements and estimates of forest inventory variables at forest stand level. Average stand area is 1.7 ha and statistics of some forest inventory parameters are summarized in Table 2. Latest updates for our RFI data base version were recorded on 2014.

Tree cover maps derived from orthophotomap processing using decision tree-based method were employed as well to examine impacts of tree cover to satellite data classification. Accuracy of these maps was evaluated during the same research project this study was performed using the same test point set. Accuracy variables for tree cover/other case were the following: overall accuracy 90.9%, kappa coefficient 0.81, user's accuracy 93.0%/88.1% and producer's accuracy 87.2% for stands in the RFI database, 98.2% for afforested lands (according to field measurements) and 90.4% for other class.

Table 2. Statistics of forest inventory data and sample plots

Statistics of the RFI data base			
Forest inventory parameter	Average	Maximum value	Standard deviation
Age (years)	61	217	46
Height (m)	16.2	36	10.5
Diameter at breast height (cm)	18.5	68	12.8
Volume	188.5	752	134.5
Statistics of the field plot data			
Age (years)	12.8	19	3.5
Height (m)	9.4	19.3	4.1
Diameter at breast height (cm)	10.2	26.6	4.5
Volume	58.6	241.6	51.6

Methods

Workflow of the tree cover classification using FCM is shown in Figure 1. For some of the case studies kNN method alone was employed.

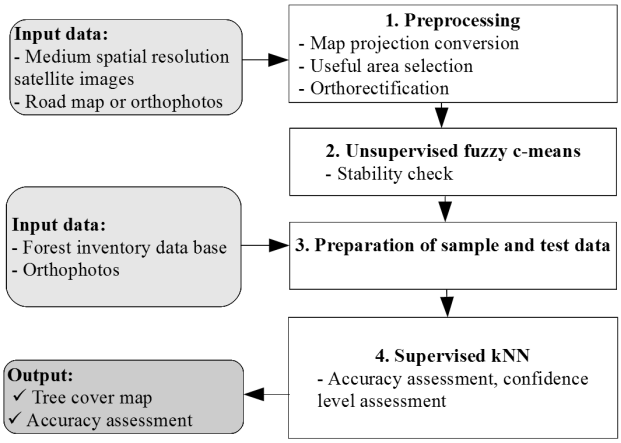

Figure 1. Workflow for tree cover identification

Forest area identification using hybrid fuzzy c-means

Since the statistical machine learning algorithms usually are computationally intense (Duda 2000), it is logical to remove easily identifiable objects from the image. Such objects are clouds (CC), cloud shadows (CS), water areas (W) and "no data" (ND) pixels. Useful area (UA) mask marks only those pixels which are valid candidates for tree cover/other classification:

$$UA = 1 - (CC\ OR\ CS\ OR\ W\ OR\ ND). \quad (1)$$

Useful area mask was calculated using the solution Fmask described in (Zhu et. al. 2015). Output of this step is a binary mask showing useful areas for classification.

Unsupervised approach classifies the image pixels by aggregating them into the natural spectral groupings or clusters (spectral classes). Choice of the hard classification method can be reasonable only if large, homogeneous land cover areas are present in the image. If the landscape is highly fragmented and transition land cover classes are present in the area, then application of the soft classification is more logical. FCM is a method of clustering which allows one pixel to belong to more than one cluster (spectral classes). "Membership grades range in value from 0 to 1 and provide a measure of the degree to which the pixel belongs to or resembles the specified class, just as the fractions of proportions used in linear mixture modelling" (Mather 2005, p. 234.).

Basic FCM algorithm is the following (Ghosh et al. 2011, Park 2012):

1. Randomly initialize the membership matrix U (size $n \times c$, where n is the number of pixels in the image to be

processed, c is the number of spectral classes specified by the user) following to the constraints:

$$\sum_{i=1}^{c} u_{ij} = 1, for\ \forall j = 1, \ldots, n \quad (2)$$

2. Calculate centroids V (size $c \times f$) using:

$$v_i = \sum_{j=1}^{n} (u_{ij})^m x_j / \sum_{j=1}^{n} (u_{ij})^m \quad (3)$$

where m is the fuzzyfication level, $m = 1.2$ in our studies; x_i is the ith pixel vector of size $1 \times f$: $x_i = [x_{i1}, x_{i2}, ..., x_{if}]$; f is the number of satellite image bands.

3. Update membership matrix U values according to equation:

$$u_{ij} = 1 / \sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)} \quad (4)$$

d_{ij} is the Euclidean distance in feature space from pixel x_i to pixel x_j, see formula 7;

4. Compute dissimilarity between iterations:

$$\Delta = \| U_t - U_{t-1} \| \quad (5)$$

where t is the number of the iteration.

5. If $\Delta < \varepsilon$, where ε is the positive constant, then stop?, othewise go to step 2.

The FCM result is represented by a membership matrix.

FCM is applied only to image pixels, which are marked by the useful area mask. This FCM implementation employs random initialization of the membership matrix U. Random initialization can result in different centroids for the same image if we test the same image several times. Because of this reason, FCM is supplemented with the following stability check:

1. FCM is repeated for the same data set 5 times. Centroids are recorded for each of the five tests.

2. Centroids are renumbered using the Euclidean distance. Centroid placement in the first test is used as a reference placement. Then for the next tests Euclidean distances are calculated form the current centroid in the current test to every centroid in the reference placement. Current centroid is placed where the minimum Euclidean distance was found with some reference centroid.

3. Average standard deviation of the centroid values was calculated for five tests. First, the standard deviation was calculated for each image band for the tests, then the average value for each centroid (along image bands) was found and finally the average value for the all centroids was calculated.

4. If $\sigma_{centroids} > 0.01$, then number of spectral classes is reduced by one and the whole procedure is repeated again until the stable solution is found.

Output of this step is an image of the spectral classes.

Spectral classes have to be labelled in two land cover classes (tree cover/other) by using sample data and the supervised k-nearest neighbours' algorithm.

"In a supervised classification, the image analyst "supervises" the pixel categorization process by specifying, to the computer algorithm, numerical descriptors of the various land cover types present in a scene" (Lillesand 2004, pp. 551). kNN assumes that similar land cover types exist within a large reference area covered by a satellite image and that the spectral radiometric responses of the pixels are only dependent on the state of the land cover (Franco-Lopez et al. 2001)

The basic kNN algorithm is the following:

1. Prepare sample data by providing sample pixel vectors and corresponding land cover classes according to the reference data or visual assessment:

$$S = [(x_1, \omega_{1j}), (x_2, \omega_{2j}), \ldots, (x_n, \omega_{nj})]^T \quad (6)$$

where x_j is the jth pixel vector of size $1 \times f$; ω_j is the land cover class for pixel i, j = 1, 2, ... , Z, Z is the number of land cover classes; T is the transposition operation.

2. Calculate the Euclidean distance in feature space from pixel to be classified x_a to all sample pixels x_i from S:

$$d_{ai} = \sqrt{\sum_{g=1}^{f} (x_{ag} - x_{ig})^2} \quad (7)$$

3. Select k most similar sample pixels (nearest neighbours) with the smallest distance d.

4. Land cover class can be estimated as mode of k nearest neighbour's land use classes.

Fifty-one pixels (or all pixels if the spectral class contains less than 51 pixels) are randomly selected from each spectral class and classified using the sample data. Spectral class is assigned to land cover class to which the majority of 51 random pixels belong. Spectral classes are united in two thematic maps (tree cover, other) by using image addition of the spectral class images with the same land cover class recognized. The number of pixels assigned to tree cover and other ones are called confidence level in this study (should not be confused with statistics). If the majority of pixels is assigned to one land cover class, then kNN confidence level is called high. If approximately half of the pixels are assigned to one land cover type, but other pixels to other land cover type, then kNN confidence level is called low.

For some of the case studies kNN was applied independently on pixel basis.

Output of this step is tree cover and other land cover images.

Preparation of the sample and test data

The sample data set was prepared using the RFI database while test data set was prepared only by visual assessment. The sample data set should include com-

prehensive variations in forest inventory variables, therefore RFI data are useful. All updated and valid RFI database records were employed for the sample set, therefore the test data set were prepared independently. The sample data set for kNN supervision included 1292 sample points:

• Class: tree cover. Totally, 711 points were selected from the RFI database. Points were centroids of forest stand polygons, where standing volume was greater than 0 m^3/ha, area of the stand larger than 5 hectares and the stand was inventoried not earlier than 2012.

• Class: other. Totally 581 points. Regular point grid was generated with spacing of 1 km between points. All points overlapping with the RFI database were removed automatically, but the rest of the points were checked visually using orthophotomaps and Google Satellite Open Layer in QGIS for 2014. Points of other land cover type but closer than 60 m to the forest were moved.

Raster values for each image were selected for each point. No additional processing of sample data was performed to keep the methodology simple.

The test data set was prepared only for Ropazi region and included 1000 randomly placed points. The random point placement was chosen to ensure more objective accuracy assessment by picking test points without any preceding considerations which might affect accuracy variables. The test points were labelled with the land cover type labels by using visual assessment of orthophotomaps and Google Satellite Open Layer in QGIS for 2014. Random pixels closer to the border of other land cover type were moved at least 60 m far from the border. Test data set included 465 tree cover points and 535 points of other land cover type. Additional 280 afforested agricultural land points were evaluated to assess the tree cover product application for afforested land identification.

Accuracy assessment

Variables derived from the confusion matrix were employed for the accuracy assessment as follows: overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA) and kappa (Foody 2002).

Description of the case studies

Research question RQ1: how is FCM stability affected by random initialization and the number of spectral classes c?

Resultant spectral classes are dependent on two parameters: 1) the number of spectral classes c usually specified by the user, 2) initialization of the membership matrix.

A case study was performed for image No. LC81870202014215LGN00. This image was chosen since it was the only cloud free image for summer season (with predictable vegetation condition happening on the ground) in the year consistent with RFI data. Since this case study is related to the behaviour of the FCM method, similar results could be obtained for other images and test image results are provided to illustrate hopping of the centroids depending on the number of spectral classes.

If the user specifies wrong number of the spectral classes or some specific case happens in initialization step, then FCM results can be unstable. Meaning, that centroids returned might be different if FCM has been applied to the same data set several times with the same user defined parameters. Therefore, FCM was repeatedly applied for the experiment purposes to the same data. For each number of spectral classes from 2 to 15 FCM was repeated five times to check similarity of centroids among those five tests. Centroid values for each experiment were saved and renumbered since the order of the spectral classes can change. After the renumbering we obtained a matrix with size $C \times F \times N$ for each number of classes c, where F is the number of bands and N is the number of experiments ($N=5$ in our study). Centroid hopping is measured by centroid standard deviation for each number of spectral classes. Low standard deviation shows that centroids are approximately the same in each experiment, but high standard deviations shows that FCM with that number of spectral classes is unstable and results should be analysed more carefully.

Research question RQ2: how are the results of the hybrid classification affected by unique conditions (atmosphere, illumination and season) during the image acquisition?

All Landsat and Sentinel images listed in the section "Remote sensing data" were employed.

Sample data were selected from each image using the same points described in the section "Preparation of the sample data and test data". kNN independently was applied to the Sentinel 2A images to check consistency of the supervised classification. The hybrid classification scheme was run for each image together with stability check starting with $c=8$ and standard deviation allowed was 0.01. These constants were found by experiment and trial method as a solution, which provides results with overall accuracy higher than 80% for each test image in our study. kNN statistics were acquired for the random samples from each spectral class to estimate the confidence level how each spectral class was classified. The coincidence level is expressed as a number of pixels from the spectral class classified as tree cover (TC) and other (O). If almost all pixels are assigned to just one land cover type, then confidence level is high. Otherwise the confidence level is low.

Research question RQ3: what are the relations between forest spectral classes and forest inventory parameters?

The case study was performed using satellite images of year 2014.

FCM responses for the spectral classes were selected for each centroid of forest stand polygon from the RFI data base. For each spectral class those forest stands were chosen, where this spectral class contained FCM value higher than 0.5. Statistics of forest inventory parameters (standing volume, age and tree species) were calculated for each spectral class to investigate relations between spectral classes and forest inventory parameters. Since the RFI data base can contain data errors, only averaged values were analysed.

We compared image LC81870202014023LGN00 (the highest overall accuracy in this study) with TC maps derived from orthophotomaps to find relationships between satellite image classification results and tree cover. For each spectral class found in Landsat image percentage of TC was calculated.

Results

Results of the case studies

Research question RQ1: how is FCM stability affected by random initialization and number of spectral classes c?

Table 3 shows the averaged standard deviation of the centroid values in five experiments for each number of spectral classes. Low standard deviation indicates that centroid values almost do not change in the independent experiments and stability is achieved, but high standard deviations mean significant changes in the centroid values and unstable FCM result. We can observe that FCM provides stable solutions in some cases. This stability can be observed in the case when number of spectral classes *c* coincides with the actual number of the spectral classes in the image (the actual number represents actual image content and it is not known by the user). For example, if there are 4 spectral classes present in the image and we run FCM with *c=5*, then one class will be separated artificially and since it would not be related to the actual spectral class, the splitting can be different in the different tests. According to the results for a specific Landsat image, *c=2*, *c=4*, *c=6*, *c=7* and *c=8* can achieve the stable solution (indicated by low standard deviation) with the presence of random initialization. *c* values higher than 8 result in unstable separation of the spectral classes. So for further studies *c=8* was set as an initial number of spectral classes. This number can be corrected by stability check.

Research question RQ2: how are the results of the hybrid classification affected by unique conditions (atmosphere, illumination, and season) during the image acquisition?

In case, when kNN was applied independently on pixel basis without FCM based grouping of pixels, then OA and kappa were 94.7%/0.89 for Sentinel-2A image acquired in March and 96.1%/0.92 for the image acquired in August. Classification results differ in 0.12 km^2 per each square kilometre in average. Differences in this context means, that different land cover types were assigned for the same area in images of March and August. After the filtering of potentially noisy differences due to geometrical shifts with the median spatial filter of size 3 by 3, only 0.09 km^2 per each square kilometre in average differed. Visual inspection of the remaining areas resulted in conclusion that classification results differ due to: 1) the actual changes in tree coverage (like clear cutting), 2) the borders of the stands what could be explained by both geometrical shifts and changes in the solar elevation and azimuth, and 3) the land cover classes with high spectral variation because of the season, like bogs and marshes. For practical use, bogs and marshes could be masked out by using temporal features.

Next, we will consider the results for FCM based workflow. Table 4 shows experiment results as a kNN statistics. There are spectral classes, which are classified with a very high confidence level meaning that almost all random pixels of the spectral class are assigned to the same land cover class, but some spectral classes represent some kind of transition classes between those two desired land cover types. For example, low tree density forest stands and shrubs can form the transition classes. Since the spectral classes represent the actual image information, this experiment shows how critical is the selection of sample data in the supervised classification. In non-homogeneous areas such transition classes are very common and it requires extra knowledge how to handle those pixels.

There are one or two transition spectral classes on average for each image and kNN confidence level which do not show clear trends according to the season or image acquisition date. Exception is image

Table 3. Stability of the centroids in five experiments for each number of spectral classes *c*

c	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\sigma_{centroids}$	0.00	203.5	0.004	135.7	0.1	0.4	0.01	199.0	164.2	72.5	131.4	160.1	108.6	68.0

LT51870202000241FUI00 that shows four transition classes, but the reason for this could be intensive cloud cover and lack of the number of useful pixels.

Differences among classification results of different dates are very small for the high confidence level classes, but are quite significant for the transitional classes. Visual assessment of satellite images showed that the transitional classes are affected by image acquisition season due to changing response for green vegetation.

Table 4. Supervised classification statistics for the stable solution of FCM. 51 random pixels were selected for each class and classified using kNN. Statistics show how many pixels in each spectral class were classified as tree cover (TC) and other (non-forest) (O) land cover type. The kNN statistics for classes with low confidence level are coloured bold. If most of the 51 pixels are assigned just to one land cover type, then confidence level is high

Image No.	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
LC81870202014023LGN00	**29 O**	11 O	1 O	**21 O**	0 O	46 O	51 O	51
	22 TC	40 TC	50 TC	**30 TC**	51 TC	5 TC	0 TC	0
LC81870202014055LGN00	7 O	46 O	**28 O**	51 O	0 O	43 O	9 O	0 O
	44 TC	5 TC	**23 TC**	0 TC	51 TC	8 TC	42 TC	51 TC
LC81870202014087LGN00	51 O	21 O	47 O	0 O	3 O	9 O	-	-
	0 TC	30 TC	4 TC	51 TC	48 TC	42 TC	-	-
LC81870202014135LGN00	**18 O**	1 O	49 O	5 O	**17 O**	42 O	-	-
	33 TC	50 TC	2 TC	46 TC	**34 TC**	9 TC	-	-
LC81870202014215LGN00	51 O	**15 O**	0 O	50 O	6 O	46 O	**24 O**	-
	0 TC	**36 TC**	51 TC	1 TC	45 TC	5 TC	**27 TC**	-
LC81870202014247LGN00	**24 O**	42 O	49 O	3 O	10 O	48 O	0 O	-
	27 TC	9 TC	2 TC	48 TC	41 TC	3 TC	51 TC	-
LC81870202014263LGN00	48 O	8 O	1 O	2 O	6 O	**28 O**	51 O	**36 O**
	3 TC	43 TC	50 TC	49 TC	45 TC	**23 TC**	0 TC	**15 TC**
LC81870202015074LGN00	0 O	0 O	**32 O**	6 O	40 O	3 O	51 O	-
	51 TC	51 TC	**19 TC**	45 TC	11 TC	48 TC	0 TC	-
LC81870202015186LGN00	47 O	**21 O**	42 O	0 O	3 O	1 O	45 O	**14 O**
	4 TC	**30 TC**	9 TC	51 TC	48 TC	50 TC	6 TC	**37 TC**
LE71870202002238KIS00	4 O	51 O	**34 O**	51 O	0 O	1 O	3 O	49 O
	47 TC	0 TC	**17 TC**	0 TC	51 TC	50 TC	48 TC	2 TC
LT51870202000241FUI00	47 O	**24 O**	7 O	**22 O**	**31 O**	5 O	**19 O**	10 O
	4 TC	**27 TC**	44 TC	**29 TC**	**20 TC**	46 TC	**32 TC**	41 TC
LE71870202000137EDC00	11 O	51 O	9 O	51 O	**27 O**	0 O	1 O	48 O
	40 TC	0 TC	42 TC	0 TC	**24 TC**	51 TC	50 TC	3 TC
LT51870201988288KIS00	51 O	2 O	0 O	0 O	51 O	51 O	**15 O**	50 O
	0 TC	49 TC	51 TC	51 TC	0 TC	0 TC	**36 TC**	1 TC
LT51872021987285XXX03	51 O	50 O	51 O	1 O	**25 O**	51 O	0 O	0 O
	0 TC	1 TC	0 TC	51 TC	**26 TC**	0 TC	51 TC	51 TC
S20150824T094301	12 O	51 O	51 O	51 O	**33 O**	**35 O**	1 O	0 O
	39 TC	0 TC	0 TC	0 TC	**18 TC**	**16 TC**	50 TC	51 TC
S20170316T094021	**15 O**	47 O	51 O	51 O	9 O	0 O	-	-
	36 TC	4 TC	0 TC	0 TC	42 TC	51 TC	-	-

Table 5 shows the accuracy assessment for each image. In the case of year 2014 overall accuracy is from 83.7% to 94.2%. Fluctuations of the accuracy are not clearly related with the image acquisition seasons, meaning that accuracy is affected strongly by other factors like atmosphere and illumination conditions and would require more reference data on atmosphere conditions to analyse it in details. However, it could be expected to have higher accuracy for non-green vegetation seasons, because the spectral separability between deciduous trees and grasslands is better in winter and early spring. Partial snow cover should not make the analysis of spectral classes more complex, because the spectral response of snow is quite distinctive comparing with the tree cover. Decrease in accuracy for years 2002, 2000, 1988 and 1987 could be explained by outdated reference data (reference data were acquired for year 2014). Those results are shown only for comparison of outdating rate for the tree cover related reference data.

Table 5. Accuracy assessment of tree cover classification

Image No.	Test pixels available due to cloud cover	OA (%)	PA (%) forest/other	UA (%) forest/other	kappa
LC81870202014023LGN0	1000	94.2	98.7/90.5	90.0/98.8	0.89
LC81870202014055LGN0	1000	91.2	95.5/87.5	86.9/95.7	0.82
LC81870202014087LGN0	1000	93.6	97.8/90.0	89.6/98.0	0.87
LC81870202014135LGN0	1000	83.7	98.9/70.5	74.4/98.7	0.68
LC81870202014215LGN0	1000	86.2	98.9/75.1	77.6/98.8	0.73
LC81870202014247LGN0	973	88.4	98.9/79.2	80.6/98.8	0.77
LC81870202014263LGN0	936	92.8	98.2/88.1	88.0/98.2	0.86
LC81870202015074LGN0	1000	86.6	99.6/75.3	77.8/99.5	0.74
LC81870202015186LGN0	992	86.0	98.5/75.0	77.6/98.3	0.72
LE71870202002238KIS00	1000	83.8	97.2/75.2	72.1/96.7	0.68
LT51870202000241FUI00	656	80.9	95.5/69.2	71.3/95.0	0.63
LE71870202000137EDC0	982	83.2	97.8/70.8	74.2/97.4	0.67
LT51870201988288KIS00	1000	79.9	99.1/63.2	70.2/98.8	0.61
LT51872021987285XXX0	1000	78.8	99.6/60.9	68.9/99.4	0.59
S20150824T094301	1000	91.1	89.7/92.3	91.0/91.1	0.82
S20170316T094021	1000	89.3	85.0/93.1	91.4/87.7	0.78

Research question RQ3: what are the relations between forest spectral classes and forest inventory parameters?

Table 6 shows the statistics of forest inventory variables for spectral classes. It can be seen that in most cases spectral classes are sensitive to standing volume not to tree species. The spectral classes with standing volume more than 100 m^3/ha were identified in all images with high confidence level. Only in three images the spectral classes with average standing volume smaller than 70 m^3/ha were identified with the high level of confidence.

Forest classes with lower level of confidence included lower standing volume values and higher level of tree species mixture. Other land cover type classes with lower confidence level included standing volume values smaller than 45 m^3/ha or strong mixture of tree species.

Table 6. Average forest inventory parameter values for the spectral classes Cl.: V is the standing volume (m^3/ha), A is the age of the stand (years), TS is the dominant tree species, TSF is the percentage of dominant tree species stands

Image No.	Forest inventory variable	Cl.1	Cl.2	Cl.3	Cl.4	Cl.5	Cl.6	Cl.7	Cl.8
LC81870202014023LGN0	TSF	**0.53**	0.57	0.61	**0.49**	0.51	-	-	-
	TS	**1**	1	1	**1**	1	-	-	-
	V	**30**	66	194	**112**	240	-	-	-
	A	**11**	28	64	**40**	76	-	-	-
LC81870202014055LGN0	TSF	0.52	-	**0.28**	-	0.61	-	0.59	0.65
	TS	1	-	**1**	-	1	-	1	1
	V	247	-	**133**	-	233	-	48	162
	A	82	-	**41**	-	78	-	20	57
LC81870202014135LGN0	TSF	**0.58**	0.65	-	-	**0.42**	-	-	-
	TS	**1**	1	-	-	**4**	-	-	-
	V	**48**	246	-	-	**136**	-	-	-
	A	**18**	82	-	-	**38**	-	-	-
LC81870202014215LGN0	TSF	-	**0.33**	0.71	-	0.38	-	**0.57**	-
	TS	-	**1**	1	-	1	-	**1**	-
	V	-	**93**	246	-	164	-	**65**	-
	A	-	**28**	84	-	53	-	**26**	-
LC81870202014247LGN0	TSF	**0.54**	-	-	0.47	0.33	-	-	-
	TS	**1**	-	-	1	1	-	-	-
	V	**58**	-	-	172	100	-	-	-
	A	**20**	-	-	57	31	-	-	-
LC81870202014263LGN0	TSF	-	0.37	0.62	0.66	0.39	**0.52**	-	**0.46**
	TS	-	1	1	1	1	**1**	-	**1**
	V	-	54	204	258	123	**44**	-	**17**
	A	-	17	70	87	41	**17**	-	**7**

Fuzzification level was set relatively low (m=1.2) to promote searching for more spectral classes and using fuzzy logic only to describe pixels on the borders among the spectral classes. The lower fuzzification level makes result more similar to hard k-means clustering. In future research with more accurate reference data, the higher fuzzification level values could be investigated in the context with forest density.

Table 7 shows percentage of tree cover for each spectral class found in the image LC818702020140 23LGN00. This table confirms assumption that the spectral classes classified as forest are strongly related with the tree crown coverage. The lowest tree crown coverage classified as TC was 57.5%. Due to the low solar elevation in winter, the spectral classes are also impacted by illumination forming separate spectral classes for stands in sunlight and shadowed areas. Under logical assumptions, the tree crown coverage increases as standing volume increases confirming impacts of standing volume on the spectral classes.

Table 7. Average tree crown coverage for each spectral class found in image LC81870202014023LGN00. Total area is 551 km^2. Sum of the area occupied by spectral classes is different from total area because threshold T=0.45 was applied to prepare binary mask for each FCM spectral class

Class No.	Area occupied by class (km^2)	Land cover class estimated by kNN	Average tree cover calculated from ancillary data (%)	Notes
1	**47**	**O**	**31.9**	**Very nonhomogeneous class**
2	65	TC	57.5	Young stands, clear cuts, borders of forest stands in sunlit side
3	211	TC	91.2	
4	**32**	**TC**	**67.8**	**Borders of forest stands in shadowed side**
5	108	TC	94.6	
6	36	O	20.2	
7	34	O	10.0	
8	20	O	5.9	

Identification of afforested agricultural lands

Accuracy of afforested agricultural land plots were calculated for images LC81870202014023LGN00 and S20150824T094301.

Only 52 afforested agricultural land points out of 280 were recognized as a forest in Landsat image. Correctly recognized plots showed average standing volume 77 m^3/ha, while unrecognised showed average standing volume 44 m^3/ha. The same tendency was observed in the previous case study, so it can be concluded that Landsat satellite images are not appropriate for identification of afforested lands because standing volume and stand density for these areas in most cases is too small. Due to the fact, that afforested lands are very non-homogeneous, location errors can also affect the results.

Use of the Sentinel data resulted in 135 correctly recognized afforested land points for August, showing increase of accuracy when spatial resolution increased.

Comparison with other tree cover related maps

The best classification result based on image LC81870202014023LGN0 was compared with different freely available maps including tree cover information. Confusion matrix characteristics were calculated for the same test points. Table 8 shows the results. High classification accuracy in our case could be explained by up-to-date, specially selected sample data and good spectral separability of forest class in winter images.

Table 8. Accuracy assessment of different data sets based on test points

Data set identificator	OA (%)	PA (%) forest/other	UA (%) forest/other	kappa
LC81870202014023LGN0 result achieved in this study.	94.2	98.7/90.5	90.0/98.8	0.89
Global tree cover map (Hansen et. al. 2013). Tree cover higher than 50% is considered as tree cover compatible with this study.	85.9	96.3/76.8	78.3/96.0	0.72
European Commission's Joint Research Centre Pan-European Forest/Non-Forest Map 2000.	80.7	87.7/74.6	75.0/87.5	0.62
Tree cover map in CORINE 2012 (first row: including code 29 as tree cover: transitional woodland/shrub, second row: excluding code 29 as tree cover).	84.5 85.7	98.9/71.7 86.5/85.0	75.4/98.7 83.4/87.8	0.69 0.71

Conclusions

The aim of this study was to evaluate a very simple methodology not requiring accurate user input: there is no parameter input required from the user and the classification is based on unsupervised classification so the user does not have to select sample data very carefully. Classes are estimated using only image information and labelled using the supervised classification. The supervised classification results were employed also for setting a confidence level that the spectral class belongs to the tree cover class.

Experiments showed that there is consistency between the spectral classes categorised by high level of coincidence, however, the transition classes can differ within this approach. Despite the high accuracy values, consistency between the results obtained for different dates can be unsatisfactory. A change detection analysis, however, can be performed for the forest classes with standing volume greater than 100 m^3/ha. Tree cover maps generated by this workflow could be employed for reliable estimation of significant tree cover decrease. For afforestation studies, images with higher spatial resolution would be recommended. The main advantage of this workflow is simplicity and user friendliness and opportunity to select reliable spectral classes by means of the confidence level. Of course, reference data with low quality can also affect confidence level evaluation, but in this case, the user must pay less attention to the sample data than using the supervised classification alone. Computational complexity for this solution is very high, so classical solutions for hybrid classification as sample data selection from unsupervised clustering results could be suggestible for map preparation in everyday practice.

For practical application, image acquisition season is not critically important, because classification results of all seasons are affected by the transitional spectral classes. However, winter images can provide higher spectral separability between forests and other areas covered with vegetation.

Another lesson learned was related with accuracy assessment. Despite the uniform distribution of validation points in the test area, it is very easy to acquire high accuracy when points are located into the homogeneous parts of land cover. Therefore, it would be recommendable to perform accuracy assessment using the reference maps generated from high spatial resolution remote sensing data instead of test points to completely understand reliability for different land cover scenarios.

Acknowledgements

This study was performed during project "Aizaugusu lauksaimniecibas platibu un neinventarizetu meza zemju meza inventarizacijas raditaju noteiksana, izmantojot talizpetes metodes" (English translation: "Estimation of forest inventory parameters for afforested agricultural lands and non-inventoried forest lands using remote sensing data"), No. P-10, ERDF activity 2.1.2.1.1. "Competence Centres".

We appreciate providing valuable support and free data of the State Forest Service, Latvian Geospatial Information Agency and USGS.

References

Anderson, J. R., Hardy, E.E., Roach, J. T. and Witmer, R. E. 1976. A land use and land cover classification system for use with remote sensor data. Geological Survey Professional Paper 964. Department of the Interior, US Government Printing Office, Washington, D.C., 34 pp. Available online at: https://pubs.usgs.gov/pp/0964/report.pdf

Banerjee, B., Bovolo, F., Bhattacharya, A., Bruzzone, L., Chaudhuri, S. and Mohan, B.K. 2015. A new self-training-based unsupervised satellite image classification technique using cluster ensemble strategy. *IEEE Geoscience and Remote Sensing Letters* 12(4): 741-745.

Civco, D. L. 1993. Artificial neural networks for land-cover classification and mapping. *International Journal of Geographical Information Science* 7(2): 173-186.

De Leeuw, J., Georgiadou, Y., Kerle, N., De Gier, A., Inoue, Y., Ferwerda, J. and Narantuya D. 2010. The function of remote sensing in support of environmental policy. *Remote Sensing* 2(7): 1731-1750.

Duda, R. O., Hart, P. E. and Stork, D.G. 2000. Pattern classification. 2nd edition. Wiley-Interscience New York, NY, USA, 680 pp.

Fan, J., Han, M. and Wang, J. 2009. Single point iterative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation. *Pattern Recognition* 42(11): 2527-2540.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80(1): 185-201.

Franklin, S. E. 2001. Remote sensing for sustainable forest management. CRC Press, Boca Raton, Fla, 424 pp. Available online at: https://www.google.com.ua/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEwjet7GKlbDiAhXB86YKHcf2CwIQFjABegQIAhAC&url=https%3A%2F%2Fwww.researchgate.net%2Ffile.PostFileLoader.html%3Fid%3D5285f7c5d039b1d4288b4580%26assetKey%3DAS%253A272177188802562%2540144190350612121&usg=AOvVaw28H0oZbYC-cafYGO5SgyMG&cshid=1558563567101122

Franco-Lopez, H., Ek, A. R. and Bauer, M.E. 2001. Estimation and mapping of forest stand density, volume, and cover type using the *k*-nearest neighbors method. *Remote Sensing of Environment* 77(3): 251-274. Available online at: http://taurus.gg.bg.ut.ee/kalle_r/KAUGSEIRE2/2007/Publikatsioonid/Franco-Lopez%20et%20al%202001.pdf

Ghosh, A., Mishra, N. S. and Ghosh, S. 2011. Fuzzy clustering algorithms for unsupervised change detection in remote sensing images. *Information Sciences* 181(4): 699-715.

Hansen, M. C. and Loveland, T. R. 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment* 122: 66-74.

Hansen, M. C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A. and Kommareddy, A. 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342(6160): 850-853. Available online at: https://pdfs.semanticscholar.org/d4e9/8078fbc0767ebc341f442f64ca7ce9e008db.pdf

Haapanen, R., Ek, A. R., Bauer, M. E. and Finley, A. O. 2004. Delineation of forest/non-forest land use classes using nearest neighbor methods. *Remote Sensing of Environment* 89(3): 265-271.

Holmgren, P. and Thuresson, T. 1998. Satellite remote sensing for forestry planning—a review. *Scandinavian Journal of Forest Research* 13(1-4): 90-110.

Kavzoglu, T. and Mather, P. M. 2003. The use of backpropagating artificial neural networks in land cover classification. *International Journal of Remote Sensing* 24(23): 4907-4938.

Kuemmerle, T., Radeloff, V.C., Perzanowski, K. and Hostert, P. 2006. Cross-border comparison of land cover and landscape pattern in Eastern Europe using a hybrid classification technique. *Remote Sensing of Environment* 103(4): 449-464.

Law on Forests. 2000. Legalisation of the Republic of Latvia. To take effect from 17.03.2000. Published in "Latvijas Vēstnesis", 98/99 (2009/2010), 16.03.2000, Riga.

Lillesand, T. M., Kiefer, R. W. and Chipman J. W. 2004. Remote Sensing and Image Interpretation. 5th edition. John Wiley and Sons, Hoboken. 763 pp.

Mather, P. M. 2005. Computer Processing of Remotely-Sensed Images. 3rd edition. John Wiley and Sons, Hoboken. 324 pp.

Mountrakis, G., Im, J. and Ogole, C. 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(3): 247-259.

Nagendra, H., Lucas, R., Honrado, J. P., Jongman, R. H., Tarantino, C., Adamo, M. and Mairota, P. 2013. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators* 33: 45-59.

Pal, M. and Mather P. M. 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing* 26(5): 1007-1011.

Park, D. C. 2012. Satellite image classification using a divergence-based fuzzy c-means algorithm. In: 5th International Conference on Image and Signal Processing, ICISP 2012, Agadir, Morocco, June 28-30, 2012. Proceedings: 555-561. Available online at: https://link.springer.com/chapter/10.1007/978-3-642-31254-0_63

Pilvere, I. 2012. Par projektu "Lauksaimniecības zemes izmantošanas efektivitātes un iespēju novērtējums" [About project "Use of agricultural land: Assessment of efficiency and opportunities"]. Atskaite-Ziņojums. Latvijas Valsts augļkopības institūts, 64 pp. Available online at: http://www.lvm.lv/images/lvm/Atskaite_zemesefekt_LVM_2012_1.pdf Last accessed on: September 25, 2015 (in Latvian).

Potapov, P. V., Turubanova, S. A., Tyukavina, A., Krylov, A. M., McCarty, J. L., Radeloff, V. C. and Hansen, M. C. 2015. Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sensing of Environment* 159: 28-43.

Richards, J. A. and Jia, X. 2006. Remote sensing digital image analysis. An Introduction. Springer-Verlag, Berlin Heidelberg. 454 pp. Available online at: http://148.206.53.84/tesiuami/S_pdfs/Remote%20Sensing%20Digital%20Image%20Analysis.pdf

Schulz, J. J., Cayuela, L., Echeverria, C., Salas, J. and Benayas, J. M. R. 2010. Monitoring land cover change of the dryland forest landscape of Central Chile (1975–2008). *Applied Geography* 30(3): 436-447.

Sloan, S. and Sayer, J. A. 2015. Forest Resources Assessment of 2015 shows positive global trends but forest loss and degradation persist in poor tropical countries. *Forest Ecology and Management* 352: 134-145.

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Bray, M. and Islam, T. 2012. Selection of classification techniques for land use/land cover change investigation. *Advances in Space Research* 50(9): 1250-1265.

Townshend, J. R. G., Huang, C., Kalluri, S. N. V., Defries, R. S., Liang, S. and Yang, K. 2000. Beware of per-pixel characterization of land cover. *International Journal of Remote Sensing* 21(4): 839-843.

Wilkinson, G. G. 2005. Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing* 43(3): 433-440.

Ye, W., Li, X., Chen, X. and Zhang, G. 2014. A spectral index for highlighting forest cover from remotely sensed imagery. In: SPIE Asia-Pacific Remote Sensing, 8 November, 2014, Beijing, China; Conference Proceedings, Volume 9260, Land Surface Remote Sensing II; 92601L. doi: 10.1117/12.2068775. Available online at: https://doi.org/10.1117/12.2068775

Yu, J., Guo, P., Chen, P., Zhang, Z. and Ruan, W. 2008. Remote sensing image classification based on improved fuzzy c-means. *Geo-Spatial Information Science*, 11(2): 90-94.

Zhu, Z., Wang, S. and Woodcock, C. E. 2015. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment* 159: 269-277.