Object-oriented mapping of urban trees using Random Forest classifiers

Article *in* International Journal of Applied Earth Observation and Geoinformation · February 2014

DOI: 10.1016/j.jag.2013.07.002

3 authors:

Anne Puissant
University of Strasbourg

Simon Rougier
University of Strasbourg

Andre Stumpf
GAF AG

International Journal of Applied Earth Observation and Geoinformation 26 (2014) 235–245

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Object-oriented mapping of urban trees using Random Forest classifiers

Anne Puissant[a,*], Simon Rougier[a], André Stumpf[a,b]

[a] Laboratoire Image, Ville, Environnement, UMR 7362 CNRS – Université de Strasbourg, France
[b] Institut de Physique du Globe de Strasbourg, UMR 7516 CNRS – Université de Strasbourg, France

ARTICLE INFO

Article history:
Received 27 February 2013
Accepted 5 July 2013

Keywords:
Urban trees
VHR satellite images
Object based image analysis
Feature selection, Random Forest classifier

ABSTRACT

Since vegetation in urban areas delivers crucial ecological services as a support to human well-being and to the urban population in general, its monitoring is a major issue for urban planners. Mapping and monitoring the changes in urban green spaces are important tasks because of their functions such as the management of air, climate and water quality, the reduction of noise, the protection of species and the development of recreational activities. In this context, the objective of this work is to propose a methodology to inventory and map the urban tree spaces from a mono-temporal very high resolution (VHR) optical image using a Random Forest classifier in combination with object-oriented approaches. The methodology is developed and its performance is evaluated on a dataset of the city of Strasbourg (France) for different categories of built-up areas. The results indicate a good accuracy and a high robustness for the classification of the green elements in terms of user and producer accuracies.

1. Introduction

Demographic, economic and environmental changes in many European cities impose to the local authorities to monitor and control the different urban functions (Jongman et al., 2004). This statement has been set in force within the European Green Infrastructure Policy which concerns the protection of the biodiversity in urban spaces (EU Biodiversity Action Plan, 2006; EU Biodiversity Strategy Plan for 2020, 2011). Similar policies are being developed at the national levels; for instance, in France, the implementation of the Grenelle 1 (2009) and Grenelle 2 (2010) environmental policy aimed at creating a regional scheme of ecological coherence, through a network of ecological corridors. The 'green' part of this network is composed of wetlands and forest (COMOP, 2009; Hubert-Moy et al., 2012).

Different types of vegetation are integrated in the available international, national or local databases according to their scales. At coarse scales (1:100,000–1:50,000) existing land cover/use databases (e.g. Corine Land Cover, Urban Atlas, etc.) integrate vegetation only if the size of the patches exceed 25 ha. At medium scales (1:50,000–1:25,000), the French national topographic database integrate some vegetation classes with patches larger than 0.5 ha. However, the database is updated only every 3 years and is not complete for all the territories (IGN, 2012). At finer scales (1:5000–1:200), urban databases locate individual trees (e.g. the center of the crown) from field observations, but provide only local coverage and are constrained to public domain. The creation of exhaustive (private and public) and up-to-date databases necessitates to acquire sub-meter aerial or satellite imagery that enable a detailed mapping of the forested elements by delineating the areas occupied by tree crowns (or patches outside forest).

The mapping of urban trees suffers from the lack of operational methods for automatic detection from different sensors and platforms (aerial photographs, satellite images). Visual image interpretation is time consuming, expensive and not adapted for larger areas. For the analysis of very high resolution (VHR) satellite images, several more automatic approaches have been proposed using per-pixel unsupervised or supervised (Peterson et al., 2004; Tooke et al., 2009) algorithms. The proposed methods are generally solely based on spectral information (Nichol and Wong, 2007). In VHR satellite images however, basic urban patterns (e.g., houses, gardens, roads) are composed by several pixels with an intrinsically high heterogeneity of the spectral response. Furthermore, the spatial resolution of the sensors was improved at the expense of the spectral resolution (Key et al., 2001). Consequently, per-pixel classifications based on spectral characteristics alone are known as insufficient to fully exploit VHR images and additional features such as texture and neighborhood relations should be included (Tuominen and Pekkarinen, 2005; Sheeren et al., 2009).

Object-based image analysis (OBIA) has been proposed as an alternative to the pixel-based classification approaches (Blaschke and Strobl, 2001; Benz et al., 2004; Zhang and Feng, 2005; Ke et al., 2008). It involves segmenting images into homogeneous regions

* Corresponding author at: 3 rue de l'Argonne, F-67000 Strasbourg, France. Tel.: +33 3 68 85 09 15; fax: +33 3 68 85 09 50.
E-mail address: anne.puissant@live-cnrs.unistra.fr (A. Puissant).

0303-2434/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jag.2013.07.002

and characterizing objects with a set of features related to spectral, spatial and contextual properties (Lang et al., 2006; Youjing and Hengtong, 2007; Mathieu et al., 2007; Pham and He, 2008; Tan and Wang, 2009; Vannier and Hubert-Moy, 2010). Recent studies used OBIA concepts to detect tree crown borders (Iovan et al., 2008; Ardila et al., 2011, 2012).

An efficient and robust classification depends, among other factors, on the identification of image object metrics that can efficiently differentiate urban tree elements from non-tree elements. For example, Zhang et al. (2010) used a decision tree classifier and Principal Component Analysis to eliminate correlated metrics. Most of the experiments use trial and error procedures to select the most relevant metrics in a ruled-based classification (Zhang and Feng, 2005; Lang et al., 2006; Li et al., 2010; Van Delm and Gulinck, 2011), which is time-consuming, subjective and hence prone to errors.

Machine learning algorithms (such as Random Forests, RF; Breiman, 2001) have demonstrated excellent performance for the analyses of many complex remote sensing datasets (Gislason et al., 2006; Lawrence et al., 2006; Watts and Lawrence, 2008) but few have been concerned OBIA (Watts and Lawrence, 2008; Stumpf and Kerle, 2011; Rodriguez-Galiano et al., 2012) or have focused on the extraction of specific vegetation pattern (Dorigo et al., 2012; Peters et al., 2011). RF is based on ensembles of classification trees and exhibits many interesting properties, such as high accuracy, robustness against over-fitting the training data, and integrated measures of variable importance (Diaz-Uriarte and Alvarez de Andres, 2006). However, like many other statistical learning techniques, RF is bias-prone in situations where the number of instances is distributed unequally among the classes of interest. In fact, under class-imbalance, most classifiers tend to be biased in favor of the majority class, and vice versa, and may underestimate the number of cases belonging to the minority class (He and Garcia, 2009).

In this context, the objective of this work is to investigate the applicability and performance of RF learning algorithms in combination with an object-oriented approach to map and inventory areas occupied by urban tree crowns (called here '*wooded elements*') from VHR optical images. Similar issues have been addressed in the context of landslide mapping from VHR imagery (Stumpf and Kerle, 2011) but the transferability of such an approach to a different domain application has yet not been investigated.

The performance of the approach is evaluated at the City of Strasbourg for areas characterized by different degrees of artificialization. The structure of the manuscript is as follows. First, the study site and available dataset are detailed (Section 2). Second, the methods are presented in terms of image segmentation and RF classification (Section 3). Third, the strategy to validate the results is presented and discussed (Section 4). Finally some conclusions and research perspectives are highlighted (Section 5).

2. Study site and data

2.1. Study site and dataset

The study site is the urban area of Strasbourg (North-East France). A Quickbird image (DigitalGlobe©) was acquired in July 2008, which is the ideal season to identify vegetation (Fig. 1). The image consists of four multispectral bands (red, green, blue and near infrared) at 2.4 m spatial resolution and one panchromatic band at 0.6 m spatial resolution. The multispectral bands were pansharpened with the panchromatic band through a Gram-Schmidt method (Amro et al., 2011) to produce a multispectral image at 0.6 m spatial resolution. The image is not orthorectified but only georeferenced (Lambert 1 projection) since the area is located in the flood plain of the Upper Rhine Graben and globally flat.

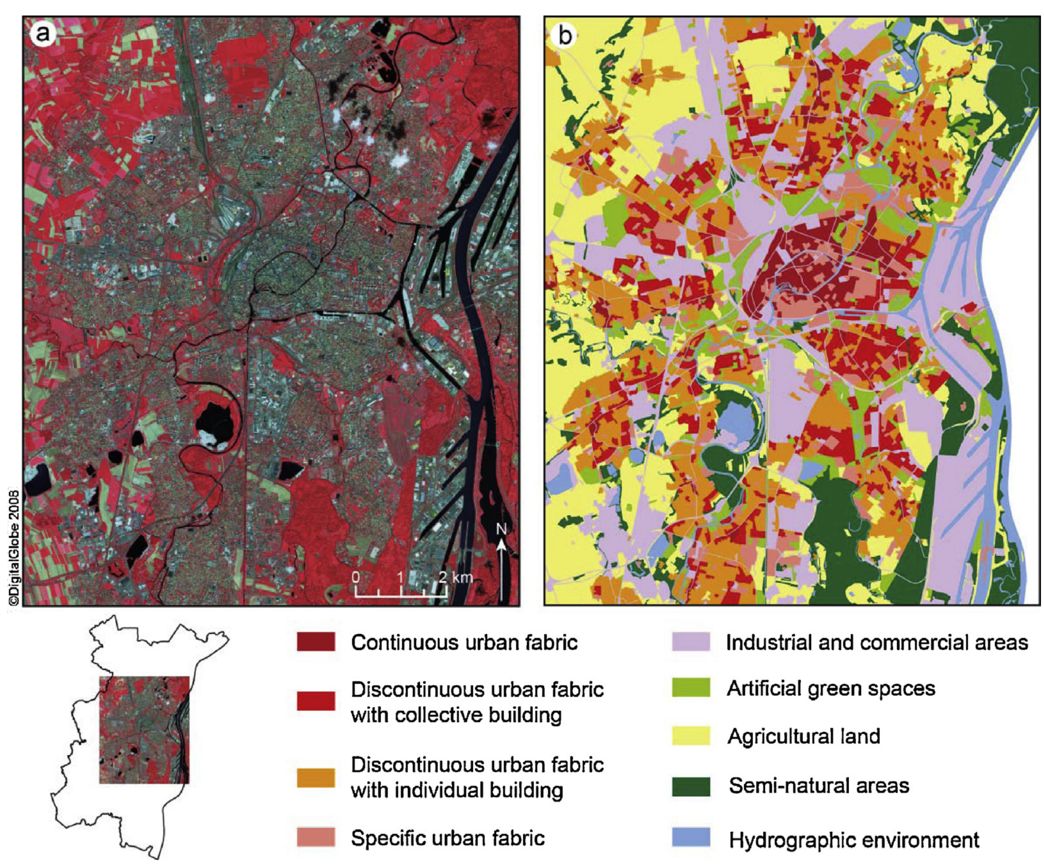

Fig. 1. (a) Study site and Quickbird image (DigitalGlobe 2008, P + MS, 60 cm) extension and (b) regional land cover/use database (BDOCCS©CIGAL, 2008).

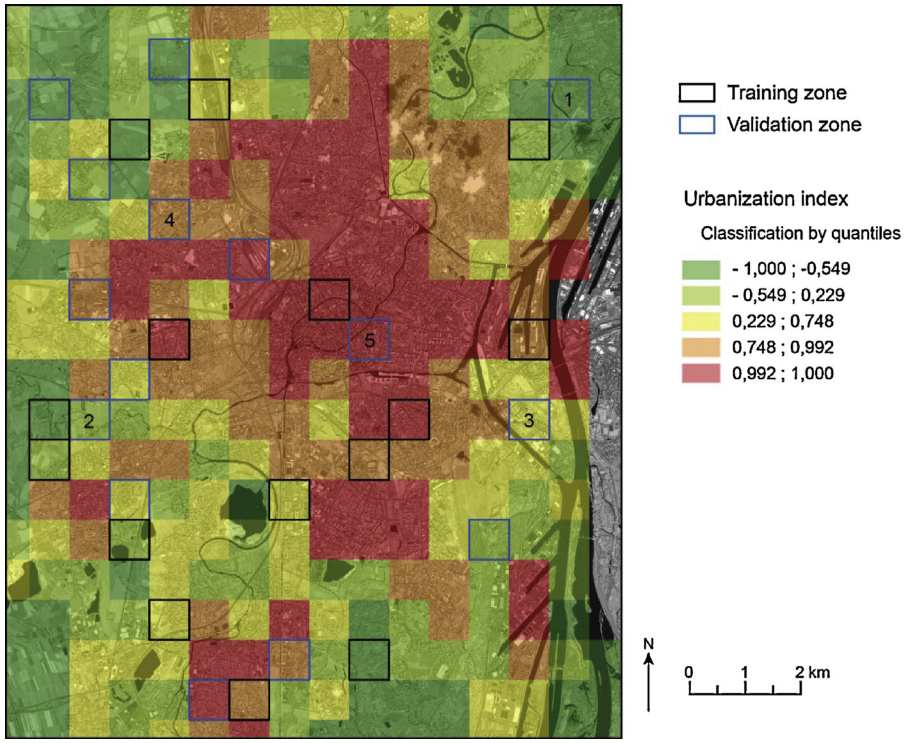

Fig. 2. Urbanization index in 5 classes with the localization of training and validation zones.

The image presents a part (11 km × 13 km) of the urban district of Strasbourg (city center and its inner-ring suburbs), and covers an area of 142 km^2 (Fig. 1a). The study site is characterized by several typical urban features representative of many western European urban areas: a concentric dense city center inherited from the Middle Ages with surroundings organized in rings where the first and second rings are characterized by urban built-up such as discontinuous urban fabric dedicated to housing, leisure activities, commercial and industrial activities (Fig. 1b). The proportion of wooded elements in these areas is composed of some patterns in public and private zones (individual trees, hedgerows, group or isolated trees with different sizes, forest) and crossed by an important river (the Rhine).

2.2. Construction of the training and validation datasets

A rigorous statistical accuracy assessment generally requires suitable reference data, an appropriate sampling scheme and independence between the training and validation datasets (Congalton, 1991). Wooded vegetation extracted from the available vegetation inventory (BDTopo©IGN, 2008) covers 19.8% of the study site. A preliminary assessment of the inventory revealed that this official database does not provide sufficient spatial details and precision to match the information content of the corresponding Quickbird image (Fig. 3, noted b). Consequently, a new database of wooded elements has been extracted by image interpretation of the pansharpened Quickbird image (Fig. 3, noted a). The inventory was elaborated only for the training and validation areas determined through a stratified random sampling scheme.

In this work, the stratified sampling is based on the urbanization index which can be considered as a proxy representing the complexity of a city with several levels of density (Fig. 2). The training dataset covers 5% of the Quickbird image and is distributed over 15 stratified areas according to the urbanization index. The size of the grid cell is typically around 1 km × 1 km (McDonnell and Hahs, 2008) and was adapted in this study to 700 m × 700 m, in order to obtain 15 randomly selected areas representative for different subsections of the study site.

The urbanization index is built by using the binary classes 'artificialized' (A_A) and 'non-artificialized' (A_{NA}) by merging the agricultural and the semi-natural land surfaces in the regional land cover database. The study area was sub-divided into 292 squared plots with a respective area A_T of 0.49 km^2 (700 m × 700 m). For each plot the urbanization index I_U was computed using Eq. (1):

$$I_U = \frac{A_A - A_{NA}}{A_T} \qquad (1)$$

This urbanization index ranges from −1 for the totally vegetated surfaces to 1 for the totally artificialized plots. The proposed index was classified by quantiles in five classes in order to obtain diverse and representative examples of wooded elements for the study site. A classification of this index by quantiles is particularly adapted because it allows each class to be equally represented spatially. Three plots per index class were selected randomly and the outlines of all wooded elements within the 15 resulting plots were digitized based on the visual interpretation of the pansharpened Quickbird image. Non-wooded elements are obtained by the subtraction of each plot surface and the exhaustive digitalization of wooded elements of the plot. The same constraints were used for the selection of 15 validation plots used for the assessment of the final classification. As before the stratified sampling was performed selecting randomly three plots for each class of the urbanization index excluding the plots already considered for training.

3. Method

The proposed method has three main steps (Fig. 4): (1) a segmentation and computation of object-attributes, (2) the classification step, and (3) the validation step. In the first step, the

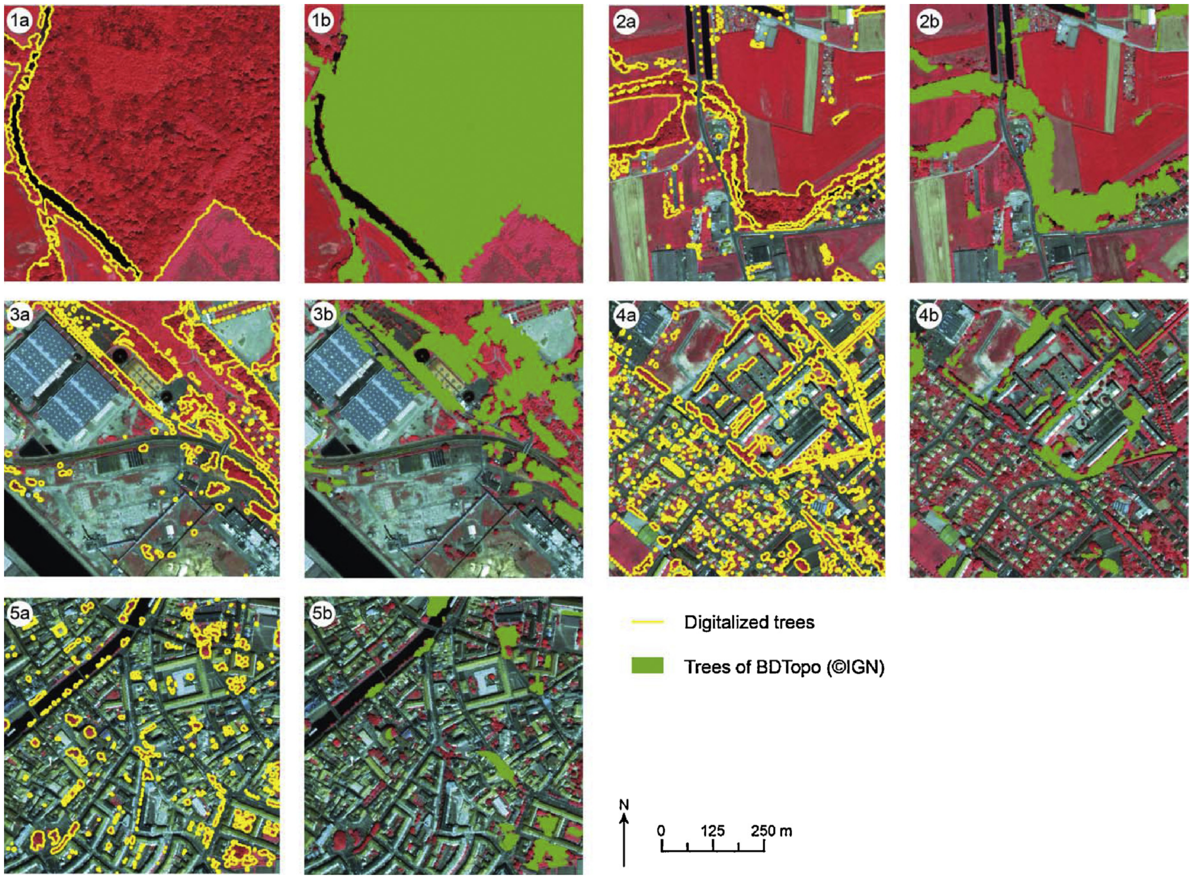

Fig. 3. For each class of urbanization index defined in Fig. 2, (a) wooded elements identified with visual interpretation and (b) wooded elements of the BDTopo.

segmentation algorithms were applied using the four multispectral bands and the MSAVI index (Section 3.1). To increase performances of the segmentation, the introduction of a vegetation index in this process can be useful (Zhang and Feng, 2005; Pham et al., 2011). In this study, the MSAVI index (Qi et al., 1994) is choose because it is well known to be particularly adapted for differentiation of vegetation in urban spaces (Pham et al., 2011; Van Delm and Gulinck, 2011). The objective is to improve and to optimize the classification procedure with RF with three protocols of tests. The test A is performed to correct the bias resulting from the imbalance between the two classes by adjusting the proportion of the classes in the training set (Section 3.3.1). The test B targets to select the more suitable metrics related to the variable importance computed with RF (Section 3.3.2). For each node of each decision tree, RF uses only a small part of total variables randomly chosen. The last test (Test C) modifies the number of variables used for these nodes (Section 3.3.3). In the third step, the results of the RF classification are evaluated by using cross-validation based on validation dataset detailed in Section 2.2.

3.1. Image segmentation

Image segmentation generates regions, and the delineation quality of the target objects has a direct influence on the accuracy of the image classification. Several segmentation algorithms have been developed in the last decades and applied in remote sensing image analysis (Pal and Pal, 1993; Dey et al., 2010), all of them aiming at the delineation of relatively homogeneous and meaningful segments.

The multi-resolution image segmentation (MRIS) and the spectral difference algorithm (SDA) segmentation implemented in eCognition® software, frequently used in Earth science studies (Blaschke, 2010), were used for image segmentation.

The MRIS is a region growing algorithm merging adjacent pixels or regions based on a heterogeneity criterion (Benz et al., 2004). This criterion composed by a scale parameter controlling the object size, weights for color and shape of the segments, and further weights for the importance of each spectral band, as well as for the importance of the compactness and smoothness of the segments. Considerable research has already been dedicated to the optimization of the scale parameter preceding rule-based classification (Espindola et al., 2006; Drăguţ et al., 2010; Martha et al., 2011). However, a number of recent studies have shown that an over-segmentation is often preferable in the context of supervised classification since it enables to capture also relatively small objects and provides a better sampling of fine spatial boundaries (Duro et al., 2012a; Smith, 2010; Stumpf and Kerle, 2011).

The SDA-based segmentation is employed frequently in studies of urban vegetation (Schöpfer et al., 2005; Lang et al., 2006; Pham et al., 2011). It merges neighboring objects according to their mean layer intensity values and a maximum threshold defined by the user (Defienens, 2011). Segments are then fused if the difference of these spectral values is below this threshold. Here SDA was applied on top of the first over-segmentation (obtained here by using MRIS) to improve the final segmentation (Johnson and Xie, 2011).

For each segment, 100 objects features 'potentially useful' for the classification of wooden elements are computed (eCognition Software®Definiens, 2011). They include spectral, shape, texture and contextual attributes (Table 1). The resulting image segments and their associated features were subsequently used as an input into the RF algorithm.

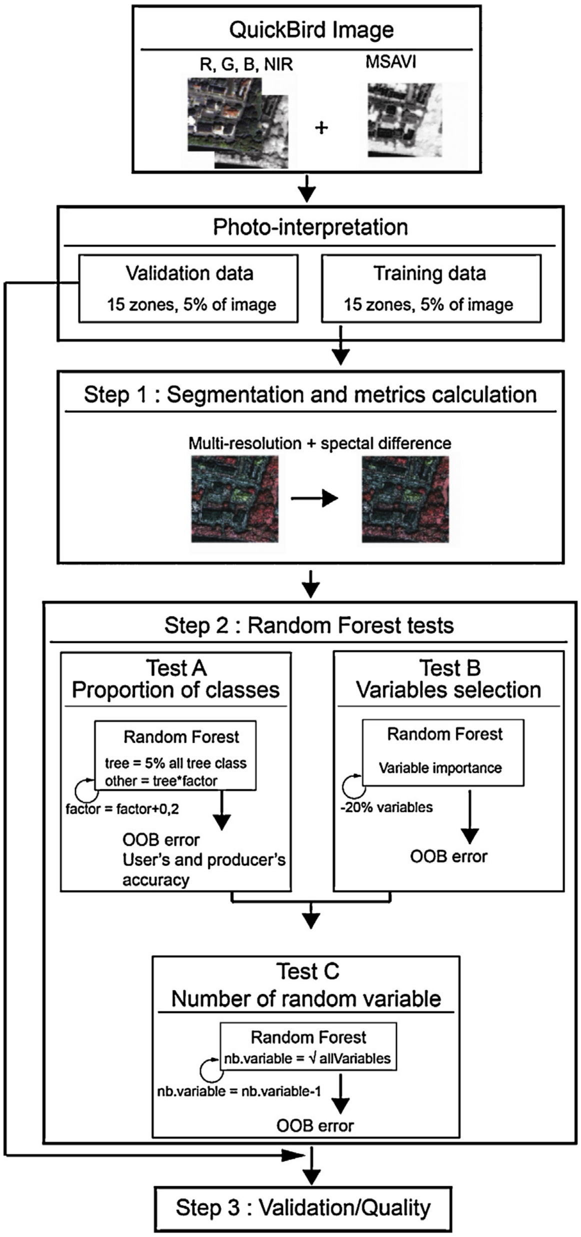

Fig. 4. Flowchart of the proposed methodology in three steps.

3.2. *Image classification: Random Forest (RF)*

RF is a machine learning algorithm proposed by Breiman (2001) for regression and classification. This supervised method is used in many domains like genetics (Diaz-Uriarte and Alvarez de Andres, 2006), medical imaging (Genuer et al., 2010), pharmacology (Svetnik et al., 2004) or ecology (Prasad et al., 2006). More specifically in remote sensing this method has already been employed to map landslides (Stumpf and Kerle, 2011), urban spaces (Guo et al., 2011), agricultural lands (Duro et al., 2012a), and natural areas (Chan et al., 2010).

RF is a multiple decision tree classifier based on classification and regression tree (CART, Breiman et al., 1984). For each tree this method performs bootstrap sampling and enables the calculation of an error estimate based on the instances remaining "out-of-bag" (OOB). RF, unlike CART, does not consider all variables at each node to determine the best split threshold but a random subset of the original set of features. The number of variables per node is typically set to the square root of the total number of variables but can be adjusted by the user. Those two mechanisms, sampling and the use of random variables for each node, create very different uncorrelated trees. Another user-defined parameter is the number of trees, which must be sufficiently large to capture the full variability of the training data and yield good classification accuracy. RF is assigning the final class to an object based on the majority vote of all trees in the forest. The RF package (Liaw and Wiener, 2002) for the open source statistical language R (R Development Core Team, 2012) was used for all experiments in the study.

Table 1
Overview of computed features for each region.

Type	Features	Band or index used
Spectral	Mean of spectral bands	R, G, B, NIR, MSAVI
	Brightness	/
	Maximum difference	/
	Standard deviation of pixel values	R, G, B, NIR, MSAVI
	Minimum pixel value	R, G, B, NIR, MSAVI
	Maximum pixel value	R, G, B, NIR, MSAVI
	Ratio (band/Brightness)	R, G, B, NIR
	Skewness of pixel value	R, G, B, NIR, MSAVI
	Mean difference to neighbors	R, G, B, NIR, MSAVI
	Mean difference to brighter neighbors	R, G, B, NIR, MSAVI
	Mean difference to darker neighbors	R, G, B, NIR, MSAVI
Geometry	Area, perimeter, length, width, length/width, asymmetry, border index, compactness, density, radius of largest enclosed ellipse, radius of smallest enclosed ellipse, rectangular fit, roundness, shape index	/
Textural	GLCM (all bands) (all directions, 0°, 45°, 90°, 135°)	Homogeneity, contrast, dissimilarity, entropy, angular 2nd moment, mean, standard deviation, correlation

RF has several advantages. It is a non-parametric method and therefore does not require that values of variables do follow a particular statistical distribution. Moreover it is not sensitive to noise or over fitting, is relatively quickly compared to other classification method like boosting techniques (Breiman, 2001). The calculation time for training RF is defined by Eq. (2):

$$cT\sqrt{MN}\log N \quad (2)$$

where c is a constant dependant of the data complexity, T the number of tree, M the number of variables and N the number of instances (Breiman, 2003). In comparison with support vector machines (SVM) whose the complexity varies between N^2 and N^3 (Bottou and Lin, 2007), RF is, therefore, better adapted for large datasets. Furthermore, RF needs less tuning (Rodriguez-Galiano et al., 2012; Gislason et al., 2006; Pal, 2005), and provides intrinsic measures of variable is calculated by permuting the value of a variable on the OOB sample and measuring the difference of OOB errors before and after the permutation. Those measures allow to analyze and interpret a classification (Rodriguez-Galiano et al., 2012), and to know which sensor (Guo et al., 2011) or which segmentation scale (Duro et al., 2012b) is more appropriate to identify a particular geographic object.

3.3. *Random Forest tuning*

To reduce the computational load, each tree was built from a stratified bootstrap sample comprising only 5% of the entire training set. This has the advantage to considerably reduce the computation time and memory requirements while at the same time the full training dataset is exploited through the repeated random sampling from the underlying distribution.

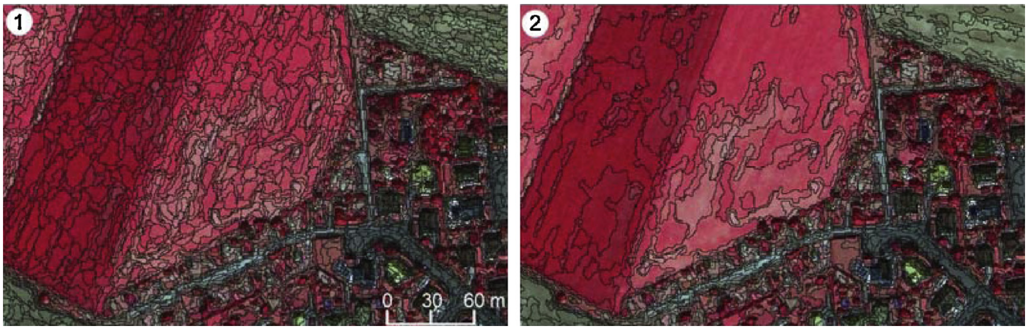

Fig. 5. Segmentation results with (1) MRIS and (2) MRIS + SDA.

3.3.1. Proportion of classes (Test A)

The wooded elements at the study site are underrepresented relative to the second class representing all mineral surfaces and other vegetation elements (called non-wooded elements). Indeed the wooded elements correspond to 17% of total surface of the training areas. Using a statistical classification method like RF when a class is underrepresented can lead to a bias in the classification, and underestimate the presence of the minority class (Stumpf and Kerle, 2011). The magnitude of this bias depends on the magnitude of the class-imbalance and the class-overlap (García et al., 2007) and is a priori unknown. By iteratively training and predicting on subsamples of the training set, it is however possible to estimate a class ratio for the training set that provides nearly balanced user's and producer's accuracies on the test set. For all the classification realized in this test, bootstrap samples corresponding to 5% of wooded segments and the same number of non-tree segments multiplied by a factor β were used for each tree in the RF. In the first iteration β is set to 1.0 and is than iteratively increased by 0.2 until reaching the original class-ratio. In each iteration a prediction is made on the entire training set to assess the balance between user's and producer's accuracy. The objective of this tuning step is to reach a balance between both.

3.3.2. Variables selection (Test B)

The second test targeted the selection of the most important variables for the classification and an assessment of the effect of their removal on the results. For this test, the method introduced by Diaz-Uriarte and Alvarez de Andres (2006) was employed. In a first step, this method builds a RF with a large number of trees (n = 5000), to measure the variable importance. Here the variable importance measured in the first part of the methodology (Section 3.1) was reused. The method proceeds by iteratively constructing new RFs, each with 2000 trees, disregarding in each iteration the least important 20% of the variables. This process is repeated until only the two most important variables are left. For each RF the OOB error is calculated and serves a criterion to select the model with the smallest OOB error. This variables selection simplifies the model and consequently helps to accelerate the classification. Furthermore, since the feature selection is performed on a subset of the data (i.e. the training set) only the selected subset of all attributes needs to be calculated for the validation areas and the whole image, which significantly decreases the computational load.

3.3.3. Number of random variables (Test C)

Finally the last tuning step performed before the final classification model concerns the adjustment of the number of random variable used at each node. This number is by default the square root of the total number of variables and it hence depends on the variable selected with the previous test (Section 3.3.2). Gislason et al. (2006) found empirically that a small number of random variables with a sufficiently large number of trees yielding higher classification accuracy on a remote sensing dataset. Indeed a smaller number of random variables could less correlated trees and enhanced performance when the trees are aggregated. Training was performed by varying iteratively the number of random variables for each node using RFs with 1000 trees. Tests were performed for all possible values between the default value and one.

All these tests were evaluated with OOB error and a confusion matrix based on the entire training set. Using the training data is useful to compare and optimize the algorithm performance with different settings but it is clear that it cannot substitute the validation on an isolated test set. This final accuracy assessment is carried out comparing the classification results at the 15 test areas against mappings resulting from visual image interpretation. Three accuracy measures are used: the users' accuracy (UA), the producer's accuracy (PA) and the F-measure which allows to measure classification performance independent class-imbalance. This measure is expressed in Eq. (3):

$$F\text{-measure} = \frac{2 * \mathrm{UA} * \mathrm{PA}}{\mathrm{UA} + \mathrm{PA}} \quad (3)$$

4. Results and discussion

4.1. Segmentation results

Considering that small wooded elements may correspond to only a few pixels in the Quickbird image, the scale parameter for the MRIS segmentation (step 1) was set to 15 resulting in a relative over-segmentation of the image (Fig. 5(1)). The weights of the color criteria were set to 0.7 and to 0.3 for the shape to give more importance to spectral intensities. Compactness was weighted with 0.3 and smoothness with 0.7. The first segmentation yielded less than 800,000 segments (Table 2).

For the second step of the segmentation based on SDA, the scale parameter was set to 20 and reduced the number of segments of the initial segmentation by fusing more than 55,000 segments. The mean size of segments increased slightly while the maximum size of segments has increased strongly (Table 2). This segmentation in two steps allowed fusing homogeneous areas such as agricultural zones (Fig. 5(2)).

The segmentation framework (MRIS and SDA) offers the possibility to assign different weights to each band. The four spectral bands were kept with a weight of one and the MSAVI index is assigned to four in order to give the same importance than the multispectral bands.

The number of segments does not exactly correspond to the sum of segments of both classes because the classes were assigned subsequent to the segmentation in a GIS (Geographic Information System). The objects exported after segmentation and the

Table 2
Parameters used for segmentation framework and results.

		MRIS	SDA
Segmentation parameters	Scale parameter	15	20
	Color/shape	0.7/0.3	/
	Compactness/smoothness	0.3/0.7	/
	Weight of spectral bands (B, G, R, NIR, MSAVI)	1, 1, 1, 1, 4	
Number of segments (%)	Total	792,614	736,583
	Wooded elements	/	105,756 (13%)
	Non wooded elements	/	688,651 (87%)
Size of segments (in pixels)	Mean	26.44	28.45
	Minimum	1	1
	Maximum	2012	142,986

calculation of features were intersected with the ground truth. Objects resulting from the segmentation have boundaries exactly following pixels whereas the photo-interpretation necessarily generalizes objects. Consequently, some additional objects are generated during this process. The wooded elements represent 13% of the total number of segments.

4.2. Random Forest: results of the three tests

4.2.1. Results for proportion of classes (Test A)

The results obtained from the segmentation of the training set and its RF-based classification show, as expected, imbalance between wooded elements and non-wooded elements. This is due to the overrepresentation of the non-wooded class compared to the wooded class as described in Section 3.3.1. The wooded class counts 105,756 segments while the non-wooded class comprises up to 688,651 producing imbalance between user's and producer's accuracies. Thereby wooded elements are underrepresented relative to the non-wooded by a factor approximately 6.5.

This test consists in varying the proportion of the "non-wooded" class in the construction of each tree of the forest. From the same number of segments is used for each class that is a number of "other" segments equal to the "wooded" segment multiplied by a factor 1. Then this factor is iteratively increased by 0.2 until the original class-ratio ($\beta = 6.5$) is reached. The user's and producer's accuracies curves presented in Fig. 6 demonstrate that a balance of both accuracies is reached at $\beta = 5.6$, which is relatively close to the inherent class-imbalance of the training data ($\beta = 6.5$).

4.2.2. Results for variable selection (Test B)

The importance of the different features (Fig. 7) represents their potential to differentiate wooded and non-wooded elements. In general measures derived from MSAVI are dominant in the variable importance ranking. Indeed seven out of eight MSAVI-based features are in the twenty most important variables. These attributes largely relegate those calculated with the NIR bands to the lower ranks beyond the 50th place. Indeed the NIR is commonly described to be crucial for the vegetation extraction, whereas our results indicate that combined index such as MSAVI is more relevant for the presented application. Only the ratio of NIR and the brightness is ranked as a relatively important feature. The textural attributes are also important in the classification process, and in particular the GLCM mean computed on all bands. This texture measure, calculated in four possible directions and in all direction, is present among the twenty most important features, each one providing information to distinguish wooded elements from other objects. Relatively few shape features are present among the most important variables; the first of them is the density occupying the 31st place. This is consistent with the rather small size of segments resulting from the scale of the segmentation. Contextual features of neighborhood relationships take an important place in the classification process (Fig. 7). Indeed, seven of these features rank among the 20 most relevant features.

The variable selection is performed by removing 20% of the remaining features at each iteration. Fig. 8 displays the evolution of OOB error when removing the less significant features. This error increases only slightly in the first iteration and then strongly when switching from three to two attributes. At the first iteration the suppression of twenty features has little impact on the classification performance and the OOB error increases only about 0.01. This RF model with 80 features was therefore considered one

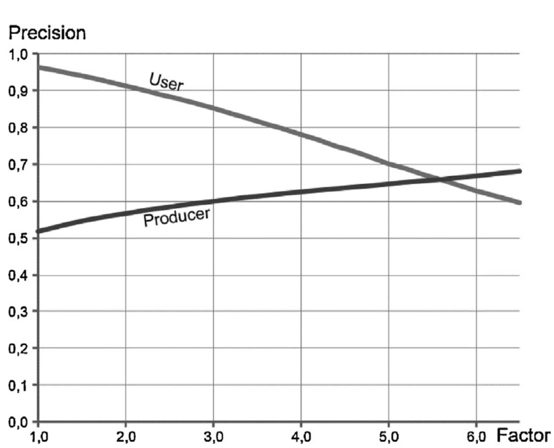

Fig. 6. Dependency of the user and producer accuracy to class imbalance.

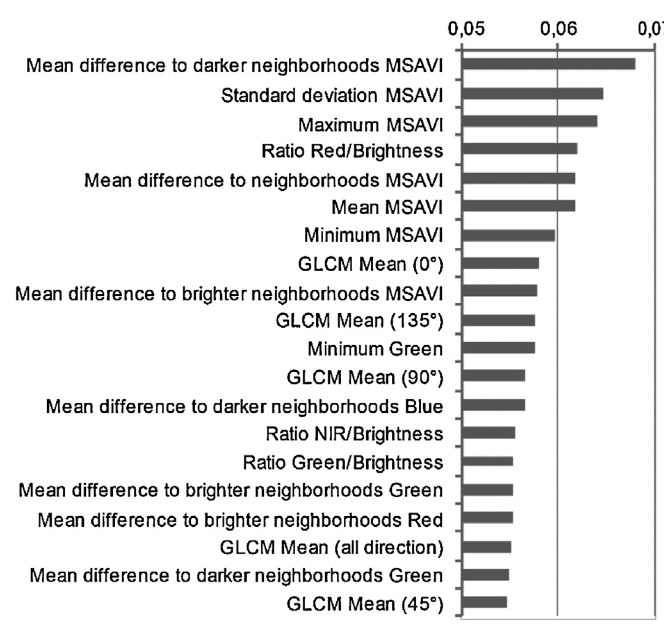

Fig. 7. Features importance.

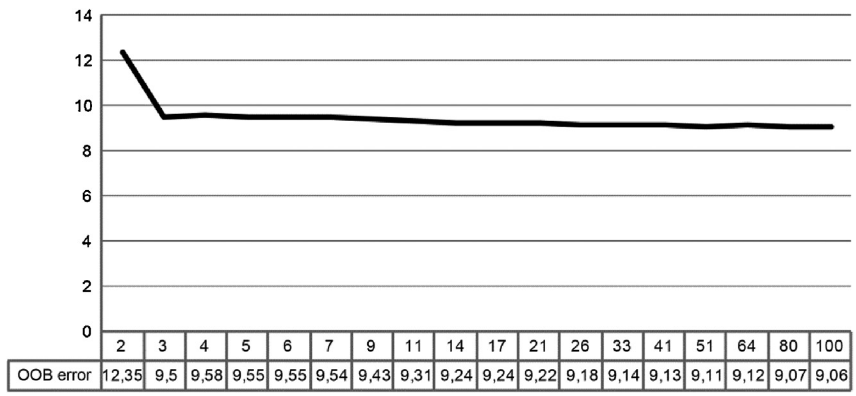

Fig. 8. Evolution of OOB error related to the number of attributes.

possible solution for obtaining a high accuracy prediction on the test set but bares the disadvantage of high computational costs for the computation of 80 features for the entire study site. Consequently, an alternative solution that might provide a better trade of between accuracy and the number of features necessary for its implementation. The RF model with 33 attributes was selected because a majority of attributes were removed, whereas the loss of accuracy compared to the full model is less than 0.1. For the subsequent test, both models with 33 and 80 features will be tested and evaluated. After this test both final models will be used to classify the validation data set and accuracy assessment performed.

4.2.3. Results for the number of randomly selected variables (Test C)

In this test, the parameters defined as optimal with the two previous tests are conserved to measure the effects of the reduction in the number of random variables used for the creation of nodes in the RF model. For the construction of each of the 1000 trees in all performed classifications, 5% of "wooded" class segments are used and the same number multiplied by a factor 5.6 of "other" class segments to balance user's and producer's accuracy. These parameters are applied for both models, one with 33 and one with 80 features. The number of random variables used for each node starts with the default value. Consequently, five variables are used for the first model and nine for the second model. This number was decreased iteratively by one until only one variable remained.

For both models using 33 or 80 features, results of this test are similar (Fig. 9). With the exception of small local minima, the OOB error decreased as the number of randomly selected variables per split was reduced. The overall error reduction is rather small for both configurations and is approximately 0.15. A possible explanation is that the use of only one variable per split yields a stronger decorrelation of the individual trees in the forest, which slightly enhances the classification results.

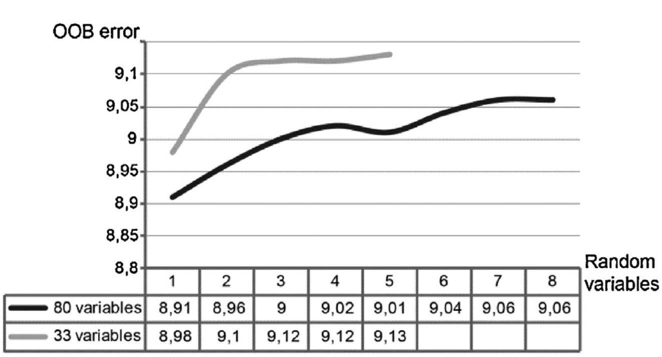

Fig. 9. Evolution of OOB error related to the number of randomly selected variables.

All described tests regarding the segmentation and the classification allowed determining the optimal parameters for these two steps. They were carried out using the training data only and the determined parameters were subsequently used to implement the classification of the validation zones and the entire image.

4.3. Validation step

An accuracy assessment of both models (comprising 33 and 80 features) was performed using the validation datasets for which plots were selected depending on the urbanization index (Section 2.2) to take better into account the landscape diversity. Accuracy measures calculated per segments present similar values for the two models (Table 3). The producer's accuracy for the 'wooded' class is higher than the user's accuracy by approximately 5% for both classifications, while the *F*-measures for these two cases were equivalent. Indeed the difference was less than 0.15%. Thereby just the model using 33 features was conserved for the next step of the study because it presented better results and allowed reducing the necessary time to calculate attributes for the entire image.

The ground truth provides objects that are generalized at the scale of the visual image interpretation, whereas the segmentation process creates edges of segments that follow limits of pixels. During the class attribution by intersecting ground truth and classification results in a GIS, many small objects are created along intersecting borders. These objects lead to a bias when the accuracy is assessed in terms of number of objects. Indeed a divided segment yields several segments with the same values of attributes that will be classified in the same class by the RF model causing a multiplication of correct and incorrect classifications affecting the accuracy measures. To correct this bias, the use of area-based accuracy measures is more and corresponds better to the traditional calculation of confusion matrix realized with pixels.

The different measures of accuracy are all higher when the segment area is used, compared to using the number of segments. Indeed the user's accuracy is equal to 0.871, producer's accuracy to 0.759 and *F*-measure to 0.813. This augmentation demonstrates

Table 3
Results of classifications calculated with segments.

	User accuracy (%)	Producer accuracy (%)	*F*-measure (%)
80 variables classification	62.20	67.42	67.70
33 variables classification	62.53	67.32	64.84

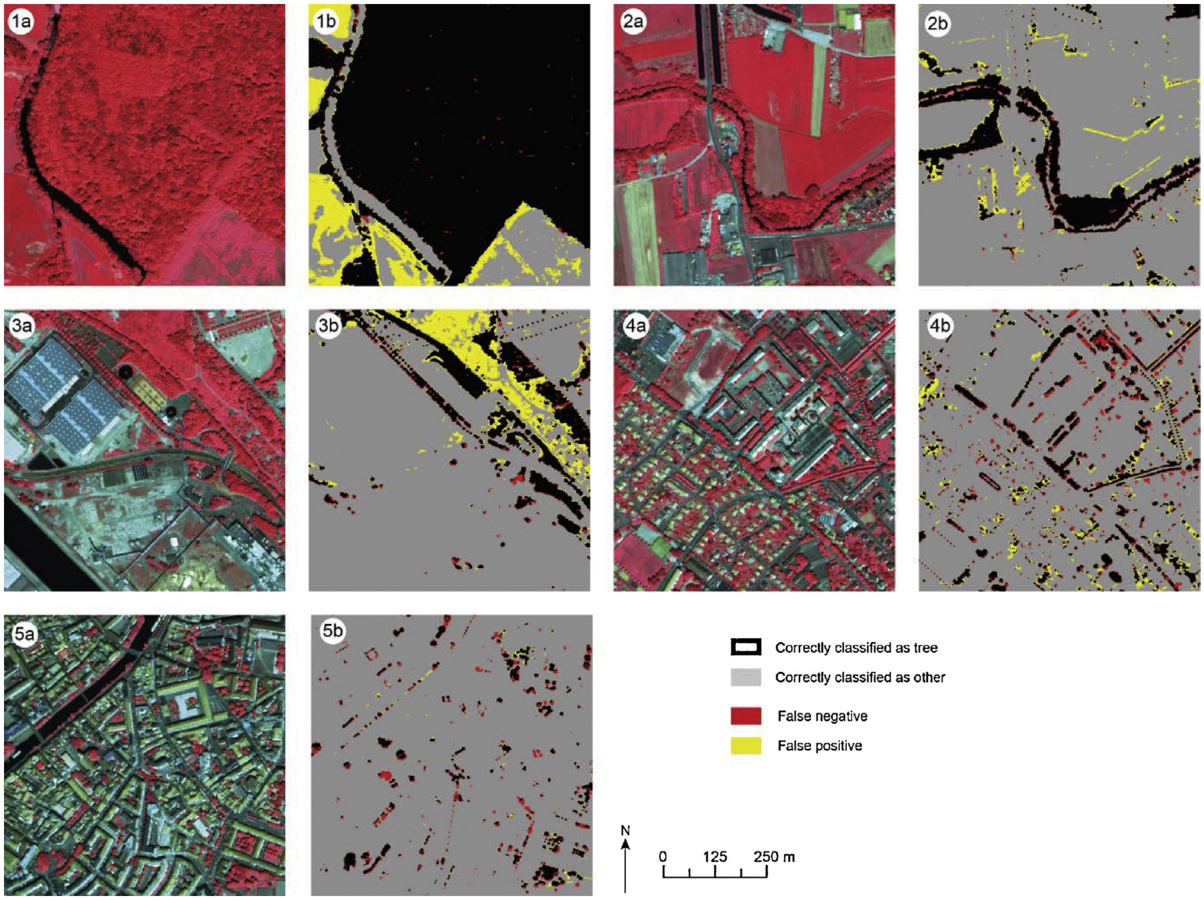

Fig. 10. (a) Quickbird image and (b) results with the SDA segmentation algorithm and 33 variables.

that the number of misclassified segments is not representative since it over-represents the influence of relatively small misclassified objects. Moreover producer's accuracy is lower than the user's accuracy, while the opposite is true when the accuracies are calculated using the number of objects. Some commission errors correspond to large plots created with the spectral difference segmentation (Fig. 10) at agricultural land. These errors reflect the results obtained according to land cover/use for the agricultural area (Table 4). Shadows due to the tree density yield some mistakes in both of those classes and create omission errors (Figs. 1b and 10).

To complete this analysis, the performance of the RF classifier has been evaluated for different types of urban fabrics. The accuracy assessment shows that the performance strongly depends on the respective type of urban fabric (Table 4). User accuracy for continuous urban fabric is low with only 57% of wooded elements being detected. These omission errors correspond to small trees or little wide hedges unidentified or only partially recognized. Similar errors are present to a lesser extent within the discontinuous urban fabric with collective buildings such as hospitals and schools. These three land cover/use classes present better producer's than user's accuracies. Similar omission errors occur within the industrial and commercial areas but the producer's accuracy is significantly lower than for the three previous classes. Large areas are classified as wooded but correspond to shrubby vegetation being visually and spectrally similar to trees. Shrubby vegetation also yields errors in other land cover/use classes like artificial green spaces or semi-natural areas, however, due to the presence of large quantities of forest in those zones its relative contribution is less important. The discontinuous urban fabric with individual houses shows relatively low user's and producer's accuracies. The small garden matrixes present in this particular urban morphology create important texture with the small trees and parcels of grass yielding important confusion. Similar sources of confusions such as allotments resembling gardens of individual houses exist within the zones of artificial green.

Table 4
Accuracy assessment of results obtained by SDA segmentation algorithm and 33 variables for land cover/use classes.

Landcover/use	User accuracy (%)	Producer accuracy (%)	*F*-measure (%)
Continuous urban fabric	59.08	89.11	71.05
Discontinuous urban fabric with collective building	68.66	85.51	76.16
Discontinuous urban fabric with individual houses	69.46	65.32	67.33
Specific urban fabric	67.70	78.68	72.78
Industrial and commercial areas	74.86	67.09	70.76
Artificial green spaces	81.46	67.84	74.03
Agricultural land	92.07	48.91	63.88
Semi-natural areas	98.65	88.90	93.52
Hydrographic environment	83.61	91.51	87.38

5. Conclusion

Wooded surfaces extraction in urban areas is an important domain of remote sensing because an accurate cartography is crucial in order to analyze and to improve knowledge about their role in ecological processes and landscape ecology studies. To help end-users to map those objects, this study investigated the use of image segmentation and the Random Forest framework for feature selection and image classification from a VHR remote sensing image. This analysis reveals several interesting points.

The different indexes derived from the MSAVI are crucial for the classification of the Quickbird image as demonstrated by the RF-based variables importance. MSAVI is therefore important information for the segmentation and the classification process to extract wooded elements in urban context especially and should be potentially valuable when applying the method to other urban areas.

The different tests performed with RF revealed that this classifier is very robust to extract wooded vegetation. Overall accuracies and OOB errors did not vary considerably when changing the class balance or the included features. Altering the number of randomly selected variables slightly increased the performance. User's and producer's accuracies could be balanced with a factor controlling the number of each training objects in each class during the construction of RF trees. The value of this factor could not be used in other cities since it depends on the study site and the segmentation scale (Stumpf and Kerle, 2011). However, the iterative resampling scheme used to determine β can be easily repeated for similar studies.

In the presented work RF seemed not sensitive to over fitting and in principle all available features could have been used. The feature selection on the training set, however, enables to disregard unimportant object attributes and thereby can save considerable computational resources when extracting object attributes from large image datasets. Thereby unsuspected features can reveal their importance for the classification through measures of variable importance. These measures enable to perform a variable selection that can delete a large part of features calculated in a first step without classification results being affected. In this study spectral and textural attributes appeared necessary to extract wooded vegetation unlike spatial attributes that did not provide essential information for this step. As for the factor β, importance of variables will be very dependent upon the segmentation and the study site but the method of Diaz-Uriarte and Alvarez de Andres (2006) allows to select automatically relevant features out of a high number of calculated attributes.

The RF classifier does not require the tuning of many parameters, is hence simple to use, and allows to identify vegetation in urban context efficiently. Only the parameters used for the segmentation approach and the OOB error threshold in the variable selection step have to be chosen. All other parameters (imbalance factor β, the overall number of variables, the number of variables selected at each tree node) can be determined automatically based on the available training data. It would be interesting to apply this method in other cities, with different urban structures to verify the reproducibility of this method. While the trained classifier resulting from such a supervised approach is typically specific to the sampled study area, the algorithm can be easily adapted for other studies using training data that represent the study site under consideration.

Since the sampling of training data set is very time consuming methods that allow to decrease the amount of required training data have recently gained greater attention and two different solutions could be interesting in this context. The first is active learning (Tuia et al., 2011; Stumpf et al., 2013) where the classifiers recommends iteratively the most valuable training samples or areas to be labeled by the user. The second possibility could be to employ transfer learning (Pan and Yang, 2010) in order to reuse the classification and the ground truth for the classification of another image by adjusting the classifier to the changes in the underlying distributions.

Acknowledgements

The work was supported by the VALI-URB project (Urban Sprawl and ecological corridors: Consolidation of VHRS remote sensing methods) co-financed by the Belgian Science Policy Office (BELSPO) and the French Centre National d'Etudes Spatiales (CNES).

References

Amro, I., Mateos, J., Vega, M., Molina, R., Katsaggelos, A., 2011. A survey of classical methods and new trends in pansharpening of multispectral images. EURASIP Journal on Advances in Signal Processing 79, 1–22.
Ardila, J.P., Tolpekin, V.A., Bijker, W., Stein, A., 2011. Markov-random-field-based super-resolution mapping for identification of urban trees in VHR images. ISPRS Journal of Photogrammetry and Remote Sensing 66, 762–775.
Ardila, J.P., Bijker, W., Tolpekin, V.A., Stein, A., 2012. Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images. International Journal of Applied Earth Observation and Geoinformation 15, 57–69.
Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multiresolution object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing 58, 239–258.
Blaschke, T., Strobl, J., 2001. What's wrong with pixels? Some recent developments interfacing remote sensing. GIS – Journal of Geographic information Systems 6, 12–17.
Blaschke, T., 2010. Object-based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65, 2–16.
Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Chapman & Hall, New York.
Breiman, L., 2001. Random forests. Machine Learning 45, 5–32.
Breiman, L., 2003. RF tool – a class of two eyed algorithm. In: SIAM Workshop.
Bottou, L., Lin, C.J., 2007. Support Vector Machine Support. Large Scale Kernel Machine., pp. 301–320.
Chan, J.C.W., Spanhove, T., Ma, J., Vanden Borre, J., Paelinckx, D., Canters, F., 2010. Natura 2000 identification and conservation status assessment with superresolution enhanced hyperspectral (CHRIS/Proba) imagery. Proceedings of Geographic Object-Based Image Analysis 3 (Part 4/C7), Ghent, Belgium.
Comop, 2009. Trame Verte et Bleue: Orientations nationales pour la préservation et la restauration des continuités écologiques. Guide 2 – Appui méthodologique à l'élaboration regional de la TVB, Rapport du comité opérationnel TVB du Grenelle de l'Environnement, Paris, 82p.
Congalton, R., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37, 35–46.
Defienens, 2011. eCognition 8.64.1 Reference Book. Trimble Documentation, Munich, Germany.
Dey, V., Zhang, Y., Zhong, M., 2010. A review on image segmentation techniques with remote sensing perspective. In: Proceedings of the International Society for Photogrammetry and Remote Sensing Symposium (ISPRS10), vol. XXXVIII (Part 7A), Austria, pp. 31–42.
Diaz-Uriarte, R., Alvarez de Andres, S., 2006. Gene selection and classification of microarray data using random forest. BMC Bioinformatics, 7.
Dorigo, T., Lucieer, A., Podobnikar, T., Čarni, A., 2012. Mapping invasive Fallopia japonica by combined spectral, spatial, and temporal analysis of digital orthophotos. International Journal of Applied Earth Observation and Geoinformation 19, 185–195.
Drăguţ, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. International Journal of Geographical Information Science 24, 859–871.
Duro, D.C., Franklin, S.E., Dubé, M.G., 2012a. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sensing of Environment 118, 259–272.
Duro, D.C., Franklin, S.E., Dubé, M.C., 2012b. Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests. International Journal of Remote Sensing 33, 4226–4502.
Espindola, G.M., Camara, G., Reis, I.A., Bins, L.S., Monteiro, A.M., 2006. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. International Journal of Remote Sensing 27, 3035–3040.
García, V., Mollineda, R., Sánchez, J., Alejo, R., Sotoca, J., 2007. When overlapping unexpectedly alters the class imbalance effects. In: Martí, J., Benedí, J., Mendonça, A., Serrat, J. (Eds.), Lecture Notes in Computer Science. Springer, Berlin/Heidelberg, pp. 499–506.
Genuer, R., Michel, V., Eger, E., Thirion, B., 2010. Random Forest based feature selection for decoding fMRI data. In: Proceeding of the 19th COMPSTAT, Paris, France, pp. 1079–1087.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recognition Letters 27, 294–300.
Guo, L., Chehata, N., Mallet, C., Boukir, S., 2011. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. ISPRS International Journal of Photogrammetry and Remote Sensing 66, 56–66.
He, H., Garcia, E.A., 2009. Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering 21, 1263–1284.
Hubert-Moy, L., Nabucet, J., Lecerf, R., Dufour, S., Burel, F., 2012. Green and blue belt networks: mapping connectivity areas at a regional scale using medium and high spatial resolution satellite images. In: IEEE International Geoscience and Remote Sensing Symposium IGARSS 2012, pp. 6095–6098.
IGN, 2012. Données – Services – Prestations de l'Institut national de l'Information géographique et forstière, 57p.
Iovan, C., Boldo, D., Cord, M., 2008 September. Detection, characterization and modeling vegetation in urban areas from high resolution aerial imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1, 206–213.
Jongman, R.H.G., Külvik, M., Kristiansen, I., 2004. European ecological networks and greenways. Landscape and Urban Planning 68, 305–319.
Johnson, B., Xie, Z., 2011. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. ISPRS Journal of Photogrammetry and Remote Sensing 66, 473–483.
Ke, C.Q., Xia, L., Tang, G.D., Cao, X., 2008. Detecting urban vegetation using a object-oriented method with QuickBird imagery. In: IEEE International Geoscience and Remote Sensing Symposium IGARSS 2008, vol. 3, pp. 138–141.
Key, T., Warne, T.A., McGraw, J.B., Fajvan, M.A., 2001. A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. Remote Sensing of Environment 75, 100–112.
Lang, S., Jekel, T., Hölbling, D., Schöpfer, E., Prinz, T., Kloyber, E., Blaschke, T., 2006. Where the grass is greener – mapping of urban green structures according to relative importance in the eyes of the citizens. In: 1st EARSeL Workshop of the SIG Urban Remote Sensing, Berlin, Germany, pp. 1–8.
Lawrence, R.L., Wood, S.D., Sheley, R.L., 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). Remote Sensing of Environment 100, 356–362.
Li, C., Yin, J., Zhao, J., 2010. Extraction of urban vegetation from high resolution remote sensing image. In: IEEE International Conference on Computer Design and Application, ICCDA 2010, vol. 4, pp. 403–406.
Liaw, A., Wiener, M., 2002. Classification and Regression by random forest. R News 2, 18–22.
Martha, T.R., Kerle, N., van Westen, C.J., Jetten, V., Kumar, K.V., 2011. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. IEEE Transactions on Geoscience and Remote Sensing 49, 4928–4943.
Mathieu, R., Freeman, C., Aryal, J., 2007. Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery. Landscape and Urban Planning 81, 179–192.
McDonnell, M.J., Hahs, A.K., 2008. The use of gradient analysis studies in advancing our understanding of the ecology of urbanizing landscapes: current status and future directions. Landscape Ecology 23, 1143–1155.
Nichol, J., Wong, M.S., 2007. Remote sensing of urban vegetation life form by spectral mixture analysis of high-resolution Ikonos satellite images. International Journal of Remote Sensing 28, 985–1000.
Pal, M., 2005. Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26, 217–222.
Pal, N.R., Pal, S.K., 1993. A review on image segmentation techniques. Pattern Recognition 29, 1277–1294.
Pan, S.J., Yang, Q., 2010. A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering 22, 1345–1359.
Peters, J., Van Coillie, F., Westra, T., De Wulf, R., 2011. Synergy of very high resolution optical and radar data for object-based olive grove mapping. International Journal of Geographical Information Science 25, 971–989.
Peterson, J., Pussa, K., Liira, J., 2004. Issues related to delineation of forest boundaries on Landsat Thematic Mapper winter images. International Journal of Remote Sensing 25, 5617–5628.
Pham, T.T.H., He, D.C., 2008. How do people perceive the city's green space? A view from satellite imagery (in Hanoi, Vietnam). In: IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2008, vol. 3, pp. 1228–1231.
Pham, T.T.H., Apparicio, P., Séguin, A.M., Gagnon, M., 2011. Mapping the greenscape and environmental equity in Montreal: an application of remote sensing and GIS. In: Caquart, S.B., Vaughan, L., Cartwright, W. (Eds.), Mapping Environmental Issues in the City: Arts and Cartography Cross Perspectives. Springer-Verlag, Berlin, pp. 30–48.
Prasad, A.M., Iverson, L.R., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9, 181–199.
Qi, J., Chehbouni, A., Huete, A., Kerr, Y., Sorooshian, S., 1994. Modified soil adjusted vegetation index (MSAVI). Remote Sensing of Environment 48, 119–126.
R Development Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 67, 93–104.
Schöpfer, E., Lang, S., Blaschke, T., 2005. A “green index” incorporating remote sensing and citizen's perception of green space. International Archives of Photogrammetry. Remote Sensing and Spatial Information Sciences, 1–6.
Sheeren, D., Bastin, N., Ouin, A., Balent, G., Lacombe, J.P., 2009. Discriminating small wooded elements in rural landscape from aerial photography: a hybrid pixel/object-based analysis approach. International Journal of Remote Sensing 30, 4979–4990.
Smith, A., 2010. Image segmentation scale parameter optimization and land cover classification using the random forest algorithm. Journal of Spatial Science 55, 69–79.
Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using random forests. Remote Sensing of Environment 115, 2564–2577.
Stumpf, A., Lachiche, N., Malet, J.-P., Kerle, N., Puissant, A., 2013. Active learning in the spatial domain for remote sensing image classification. IEEE Transactions on Geoscience and Remote Sensing, http://dx.doi.org/10.1109/TGRS.2013.2262052 (in press).
Svetnik, V., Liaw, A., Tong, C., Wang, T., 2004. Application of Breiman's random forest to modeling structure–activity relationships of pharmaceutical molecules. In: Roli, F., Kittler, J., Windeatt, T. (Eds.), Multiple Classifier Systems. Springer, Berlin, pp. 334–343.
Tan, Q., Wang, J., 2009. Mapping urban tree coverage using object-oriented image analysis method: a case study. In: IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2009, vol. 5, pp. 342–345.
Tooke, T.R., Coops, N.C., et al., 2009. Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. Remote Sensing of Environment 113 (2), 398–407.
Tuia, D., Volpi, M., Copa, L., Kanevski, M., Munoz-Mari, J., 2011. A survey of active learning algorithms for supervised remote sensing image classification. IEEE Journal of Selected Topics in Signal Processing 5, 606–617.
Tuominen, S., Pekkarinen, A., 2005. Performance of different spectral and textural aerial photograph features in multi-source forest inventory. Remote Sensing of Environment 94, 256–268.
Van Delm, A., Gulinck, H., 2011. Classification and quantification of green in the expanding urban and semi-urban complex: application of detailed field data and Ikonos-imagery. Ecological Indicators 11, 52–60.
Vannier, C., Hubert-Moy, L., 2010. Wooded hedgerows characterization in rural landscapes using very high spatial resolution satellite images. IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2010, 347–350.
Watts, J.D., Lawrence, L., 2008. Merging Random Forest classification an object-oriented approach for analysis of agricultural lands. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (Part 4/B7), 579–582.
Youjing, Z., Hengtong, F., 2007. Identification scales for urban vegetation classification using high spatial resolution satellite data. In: IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2007, pp. 1472–1475.
Zhang, X., Feng, X., 2005. Detecting urban vegetation from Ikonos data using an object-oriented approach. Proceedings, IEEE International Geoscience and Remote Sensing Symposium, IGARSS'05 2, 1475–1478.
Zhang, X., Feng, X., Jiang, H., 2010. Object-oriented method for urban vegetation mapping using Ikonos imagery. International Journal of Remote Sensing 31, 177–196.