ISPRS Journal of Photogrammetry and Remote Sensing 114 (2016) 24–31

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Review Article

Random forest in remote sensing: A review of applications and future directions

Mariana Belgiu [a,*], Lucian Drăguţ [b]

[a] Department of Geoinformatics – Z_GIS, Salzburg University, Schillerstrasse 30, 5020 Salzburg, Austria
[b] West University of Timisoara, Department of Geography, Vasile Parvan Avenue, 300223 Timisoara, Romania

ARTICLE INFO

Article history:
Received 9 October 2015
Received in revised form 11 December 2015
Accepted 20 January 2016
Available online 12 February 2016

Keywords:
Random forest
Supervised classifier
Ensemble classifier
Review
Feature selection

ABSTRACT

A random forest (RF) classifier is an ensemble classifier that produces multiple decision trees, using a randomly selected subset of training samples and variables. This classifier has become popular within the remote sensing community due to the accuracy of its classifications. The overall objective of this work was to review the utilization of RF classifier in remote sensing. This review has revealed that RF classifier can successfully handle high data dimensionality and multicolinearity, being both fast and insensitive to overfitting. It is, however, sensitive to the sampling design. The variable importance (VI) measurement provided by the RF classifier has been extensively exploited in different scenarios, for example to reduce the number of dimensions of hyperspectral data, to identify the most relevant multisource remote sensing and geographic data, and to select the most suitable season to classify particular target classes. Further investigations are required into less commonly exploited uses of this classifier, such as for sample proximity analysis to detect and remove outliers in the training samples.

1. Introduction

Remote sensing has proved its value in many fields but the success of any image classification depends on various factors, including the choice of a suitable classification procedure (Lu and Weng, 2007). Supervised classifiers are widely used since they are more robust than model-based approaches (Niemeyer et al., 2014). These classifiers are able to learn the characteristics of target classes from training samples and to identify these learned characteristics in the unclassified data. An efficient supervised classifier needs to address the challenges (Millard and Richardson, 2015) involved in (1) mitigating the Hughes phenomenon (i.e. the "curse of dimensionality"), which occurs when the number of variables is much larger than the number of training samples (Ghosh et al., 2014), (2) dealing with the nonlinearity of variables, (3) dealing with imbalanced training samples and noise in both training samples and unlabelled data, and (4) reducing computation time (Gislason et al., 2006).

Over the last two decades the use of the random forest (RF) classifier (Breiman, 2001) has received increasing attention due to the excellent classification results obtained and the speed of processing (Du et al., 2015; Pal, 2005; Rodriguez-Galiano et al., 2012). The RF classifier yields reliable classifications using predictions derived from an ensemble of decision trees (Breiman, 2001). Furthermore, this classifier can be successfully used to select and rank those variables with the greatest ability to discriminate between the target classes. This is an important asset given that the high dimensionality of remotely sensed data makes the selection of the most relevant variables a time-consuming (Körting et al., 2013), error prone, and subjective task (Belgiu et al., 2014a).

A number of studies have systematically investigated the utilization of the RF classifier for hyperspectral data classification (Ham et al., 2005) and land cover (LC) classification of Enhanced Thematic Mapper (ETM+) (Pal, 2005) or Multispectral Scanner (MSS) and Digital Elevation Model (DEM) data (Gislason et al., 2006). There has, however been no publication to date dedicated to summarizing the use of this versatile and efficient classifier in different application scenarios.

The objective of this work has therefore been to summarize the use of the RF classifier in remote sensing, with special attention to its parameterization and its sensitivities to changes in sampling procedures, to the size and representativeness of training sample

* Corresponding author at: Schillerstr. 30, 5020 Salzburg, Austria. Tel.: +43 662 8044 7518; fax: +43 662 8044 7560.
E-mail addresses: mariana.belgiu2@sbg.ac.at (M. Belgiu), lucian.dragut@fulbrightmail.com (L. Drăguţ).

http://dx.doi.org/10.1016/j.isprsjprs.2016.01.011
0924-2716/

sets, and to noise in the data. Following a brief overview of ensemble classifiers in remote sensing, we present the characteristics of the RF classifier (Section 2). The RF description is intended for readers who have limited experience with machine learning classifiers. Section 3 is dedicated to the use of the RF classifier with data from different sensors. In the following section we then address the sensitivity of the RF classifier to sampling procedures and to feature filtering. Section 5 includes an evaluation of the performance of the RF classifier compared to other mainstream classifiers. Possible objectives for future research are then discussed in Section 6 and conclusions are presented in Section 7. This article is concerned specifically with the use of RF classifier for classification tasks, and regression models are therefore not addressed. We use the terms 'variable' and 'feature' interchangeably to refer to class attributes or properties identified in the remotely sensed data.

2. Ensemble classifiers in remote sensing

Supervised parametric classifiers such as Maximum Likelihood Classification (MLC) deliver excellent results when dealing with unimodal data. However, they have limitations when dealing with multi-modal input datasets because these classifiers assume a normal data distribution (Liu et al., 2011). Non-parametric supervised classifiers such as the Classification and Regression Tree (CART), Support Vector Machine (SVM) (Mountrakis et al., 2011), and Artificial Neural Network (ANN) (Mas and Flores, 2007) classifiers do not make any assumptions regarding frequency distribution and have therefore become increasingly popular for classifying remotely sensed data, which rarely have normal distributions.

Because the nature and causes of spatial variation in images are not understood, the analysis has been limited to the empirical association between surface phenomenon and patterns in images (Woodcock et al., 1988), with the implicit assumption that reality has a consistent spectral response in imagery. This assumption is often violated, however, as a consequence of the complex interplay between factors like scene complexity, scale and aggregation (Marceau et al., 1994). Therefore, simple classifiers may reach their limits in many applications.

In the last years the attention of the remote sensing community has turned to ensemble classifiers (Miao et al., 2012; Gislason et al., 2006). These classifiers can be based on an individual supervised classifier or on a number of different supervised classifiers that are trained using bagging (Breiman, 1996) or boosting approaches (Schapire, 1990; Freund and Schapire, 1997), or variations of these approaches. In the bagging approach (also known as the bootstrap aggregation approach) each classifier in the ensemble is trained on a random subset of a training samples set, whereas in the boosting approach the ensemble classifiers are trained iteratively using all of the training samples, increasing the weightings for the incorrectly classified samples during the training procedure. Previous work has shown that using boosting and bagging ensemble methods achieved greater accuracy than using single classifiers such as decision tree classifiers (Briem et al., 2002; Miao et al., 2012), as well as being more stable and robust to noise in the training data (DeFries and Chan, 2000). In an experimental study using data from various application domains, Dietterich (2000) established that boosting is more accurate than bagging. Boosting approaches have been shown to reduce classification variance and bias (Gislason et al., 2006). However, they require large computational resources, overfit if there are insufficient training samples, and are sensitive to any outliers present in the training samples (Xu et al., 2014). Examples of boosting methods commonly used in remote sensing are AdaBoost (Chan and Paelinckx, 2008; Miao et al., 2012) and JointBoost (Guo et al., 2015). In contrast, bagging approaches reduce the classification variance but they have little effect on the classification bias (Briem et al., 2002; DeFries and Chan, 2000).

2.1. The random forest classifier

The RF classifier is an ensemble classifier that uses a set of CARTs to make a prediction (Breiman, 2001). The trees are created by drawing a subset of training samples through replacement (a bagging approach). This means that the same sample can be selected several times, while others may not be selected at all (Fig. 1A).

About two thirds of the samples (referred to as *in-bag* samples) are used to train the trees (Fig. 1A) with the remaining one third (referred to as *out-of-the bag* samples) are used in an internal cross-validation technique for estimating how well the resulting RF model performs (Breiman, 2001).

This error estimate is known as the out-of-bag (OOB) error. Each decision tree is independently produced without any pruning and each node is split using a user-defined number of features (*Mtry*), selected at random. By growing the forest up to a user-defined number of trees (*Ntree*), the algorithm creates trees that have high variance and low bias (Breiman, 2001). The final classification decision is taken by averaging (using the arithmetic mean) the class assignment probabilities calculated by all produced trees. A new unlabelled data input is thus evaluated against all decision trees created in the ensemble and each tree votes for a class membership. The membership class with the maximum votes will be the one that is finally selected (Fig. 1B).

As mentioned above, two parameters need to be set in order to produce the forest trees: the number of decision trees to be generated (*Ntree*) and the number of variables to be selected and tested for the best split when growing the trees (*Mtry*). Theoretical and empirical research has highlighted that classification accuracy is less sensitive to *Ntree* than to the *Mtry* parameter (Ghosh et al., 2014; Kulkarni and Sinha, 2012). Since RF classifier is computationally efficient and does not overfit, *Ntree* can be as large as possible (Guan et al., 2013). The majority of the studies reported in this review set the *Ntree* value to 500 because the errors stabilize before this number of classification trees is achieved (Lawrence et al., 2006). Another reason for this value being commonly used could be the fact that 500 is the default value in the R package for random forests, which is called "*randomForest*" and is the most popular RF implementation. Other investigators have used different values to *Ntree*, such as 5000 (Adelabu et al., 2014a; Díaz-Uriarte and De Andres, 2006; Millard and Richardson, 2015; Nitze et al., 2015; Stumpf and Kerle, 2011), 1000 (Colditz, 2015; Reese et al., 2014; Sesnie et al., 2010), or 100 (Guan et al., 2013). In a study dedicated to polarimetric synthetic aperture radar (PolSAR) image classification, Du et al. (2015) investigated the sensitivity of the RF classifier to the number of trees (from 10 to 200 trees in steps of 10) and showed that this parameter has no influence on the classification results. Similar results were reported by Topouzelis and Psyllos (2012), who used random forests to classify oil spills from synthetic aperture radar (SAR) data and found that an ensemble of 70 trees was sufficient for this classification purpose as the classification no longer improved as the number of trees increased above this threshold. Based on the results published to date, we suggest that the default value of 500 for *Ntree* is an acceptable value when using the RF classifier on remotely sensed data.

The *Mtry* parameter is usually set to the square root of the number of input variables (Gislason et al., 2006). Ghosh et al. (2014) set *Mtry* to the total number of available variables, but this approach increases the computation time as the algorithm has to compute the information gain contributed by all of the variables used to split the nodes.

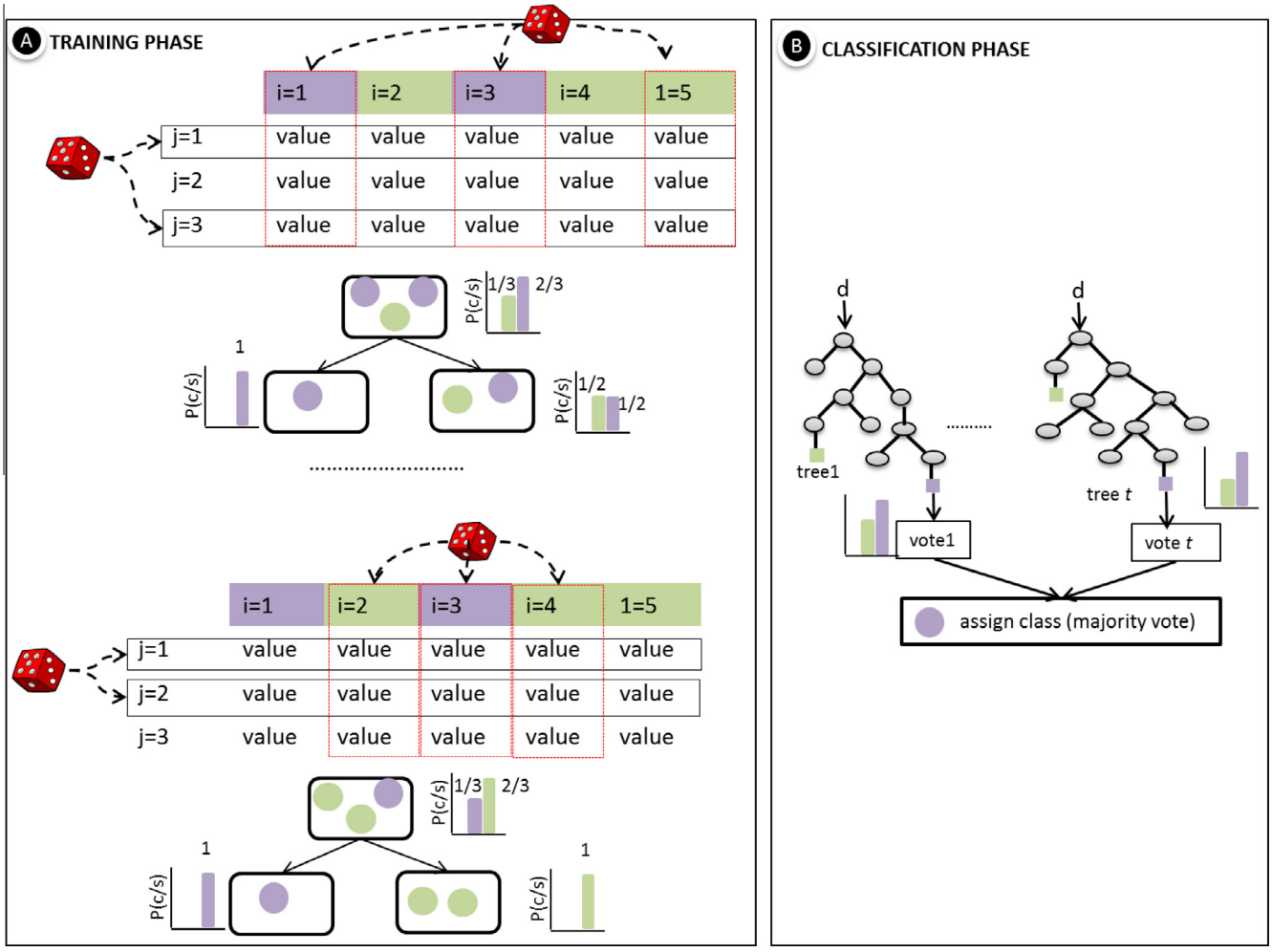

Fig. 1. Training and classification phases of Random Forest classifier: *i* = samples, *j* = variables, *p* = probability, c = class, s = data, *t* = number of trees, d = new data to be classified, and value = the different values that the variable *j* can have.

During recent years there has been an increased interest in additional RF functions, such as using the variable importance (VI) to optimize feature space (Belgiu et al., 2014b; Corcoran et al., 2013; Pedergnana et al., 2013), measuring the correlation between high dimensional datasets on the basis of internal proximity matrix measurements (Peerbhay et al., 2015), or identifying outliers in the training samples by explorative analysis of sample proximities (Corcoran et al., 2013). The VI can be calculated internally in a number of different ways, such as using the Mean Decrease in Gini (MDG), or the Mean Decrease in Accuracy (MDA) (Breiman, 2001). The MDG measures how much a variable reduces the Gini Impurity metric in a particular class, while the MDA takes into account the difference between the OOB error resulting from a data set obtained through random permutations of the values of the different variables and the OOB error resulting from the original data set (Breiman, 2001). The majority of studies reported in this review used the MDA to determine the VI (Schmidt et al., 2014).

The computing time required to establish the RF classification model is:

$$T\sqrt{MN\log(N)} \quad (1)$$

where T is the number of trees, M is the number of variables used in each split, and N is the number of training samples (Breiman, 2001).

At the time of writing the RF classifier had been implemented in a number of software packages, such as eCognition (Trimble, 2013), R software (R-Development-Core-Team, 2005), Weka (Holmes et al., 1994), the scikit-learn package (Pedregosa et al., 2011), imageRF (Waske et al., 2012), Ranger (originally known as Random Jungle) (Schwarz et al., 2010), STATISTICA (http://www.statsoft.com), Willows (Zhang et al., 2009), and Matlab (https://github.com/karpathy/Random-Forest-Matlab).

3. Random forest developments in remote sensing

3.1. Use of random forests to extract information from multispectral, radar, LiDAR and thermal remote sensing imagery

The RF classifier has been successfully used to map LC classes (Colditz, 2015; Haas and Ban, 2014; Stefanski et al., 2013; Tsutsumida and Comber, 2015) and urban buildings (Belgiu and Drăguț, 2014), to classify insect defoliation levels using the red-edge band of a RapidEye image (Adelabu et al., 2014a), to map boreal forest habitats using WorldView-2 imagery (Räsänen et al., 2013), to map biomass using Landsat temporal data (Frazier et al., 2014), to classify urban impervious surfaces from single-date MODIS data (Deng and Wu, 2013), to identify tree health using IKONOS data (Wang et al., 2015), and to map tree canopy cover and biomass using uni-temporal and multi-temporal Landsat 8 imagery (Karlson et al., 2015). RF classifier has also been successfully used for oil spill mapping from SAR data (Topouzelis and Psyllos, 2012) and for LC classifications from both multitemporal

SAR data (Waske and Braun, 2009) and PolSAR data (Uhlmann and Kiranyaz, 2014).

A number of studies have looked into the potential of RF classification to improve urban object classification from airborne LiDAR data (Chehata et al., 2009; Niemeyer et al., 2014), and also for mapping reforested landslides using variables calculated either for each pixel (Chen et al., 2014) or for image objects delineated by segmentation (Li et al., 2015). A few studies have also explored the use of random forests in the classification of Unmanned Aerial Vehicle (UAV) data (Ma et al., 2015), and in thermal remote sensing (Sun and Schulz, 2015).

3.2. Application of the random forest classifier to hyperspectral imagery

The classification of hyperspectral data is challenged by its high dimensionality (Ham et al., 2005; Junshi et al., 2015; Zhang, 2014). This problem is amplified by the relatively small amount of training data available, which results in an increase in the number of misclassifications (Chan and Paelinckx, 2008; Zhang, 2014). Hyperspectral remote sensing applications therefore require solutions that are able to remove redundant and noisy wavebands.

RF classifier has been tested as a feature space optimization technique with satisfactory results in numerous applications such as the mapping of tree species (Cavallaro et al., 2015; Dalponte et al., 2013; Fassnacht et al., 2014; Shang and Chisholm, 2014), LC classes (Chan and Paelinckx, 2008; Demarchi et al., 2014), invasive plants (Lawrence et al., 2006; Peerbhay et al., 2015), pine trees struck by lightning (Abdel-Rahman et al., 2014), grass species (Mansour et al., 2012), and insect defoliation levels in mopane woodland (Adelabu et al., 2014b).

3.3. Random forest classifier and multi-source data

Remote sensing data on its own is not always sufficient for classifying target objects (Gislason et al., 2006; Guan et al., 2013). A number of studies have therefore evaluated the impact of using either imagery acquired by different sensors, or remotely sensed data combined with ancillary geographic data such as information derived from digital elevation data (Gislason et al., 2006) or soil classes (Corcoran et al., 2013) on classification accuracy. Given the diversity of input datasets that could be used to improve classification, it is important to include only the most relevant datasets in order to reduce the computational burden without sacrificing the accuracy of the results (Corcoran et al., 2013). In this context, RF classifier has been used to evaluate the contribution of each data source to the classification results (Gislason et al., 2006; Corcoran et al., 2013), and/or for supervised classification purposes (Sesnie et al., 2010; Bechtel and Daneke, 2012; Ahmed et al., 2015; Naidoo et al., 2012; Guo et al., 2011). Gislason et al. (2006) used RF procedures to identify the most important variables for LC classification using MSS, elevation (10 m contour intervals), slope, and aspect datasets. They found that elevation was the only topographic variable of any importance for identifying target land cover classes. Corcoran et al. (2013) evaluated the influence of multisource and multi-temporal remotely sensed and ancillary data (such as Thematic Mapper (TM) data, radar data, terrain variables (curvature and elevation), and soil descriptors) on the accuracy of RF-based classification for wetlands mapping in northern Minnesota. The authors found that a classification model that consisted of elevation, curvature, soil descriptors, and remotely sensed imagery acquired in spring achieved the best accuracy.

The investigations described above concluded that the RF classifier can be successfully used to classify multisource remote sensing and geographic data, mainly because of its computational speed (Gislason et al., 2006; Corcoran et al., 2013) and because it helps image analysts to optimize the classification model by using only those input datasets identified as being important to the classification objective.

4. Sensitivity of the random forest classifier to training samples and data dimensions

The samples used to train the supervised classifiers need to fulfil a number of requirements: (1) training and validation data must be statistically independent, (2) training samples must be class balanced, (3) training samples must be representative of the target classes, and (4) the training sample needs to be large enough to accommodate the increasing number of data dimensions (to mitigate the Hughes phenomenon).

The generation of training samples is a time intensive, expensive and subjective task (Ghimire et al., 2012). It is therefore important to evaluate the sensitivity of RF classification to sampling design and imbalanced training samples (Colditz, 2015; Dalponte et al., 2013; Jin et al., 2014; Mellor et al., 2015). Colditz (2015) explored the sensitivity of RF classifier to different sampling designs such as random sampling, the allocation of samples in proportion to the expected area, or the allocation of an equal number of samples to each class. The results of this study revealed that area-proportional allocation of training samples per class achieved the best classification results because the classes that occupy large areas require more samples than those that occupy small areas (Colditz, 2015).

When investigating forest classification from Landsat data, Mellor et al. (2015) found that RF classification was relatively insensitive to mislabelled training data and that imbalanced training data can be introduced to reduce the errors in those classes that pose the greatest challenges to classifications. In contrast, investigations by Millard and Richardson (2015) and Dalponte et al. (2013) found that RF classifier is sensitive to spatial autocorrelation of the training classes and to the proportions of the different classes within the training samples. RF classifier fails to cope with imbalanced training data and tends to favor the most representative classes (Dalponte et al., 2013). Jin et al. (2014) showed that the proportionally allocated training sample design reduces the commission error of the under-represented classes, and that the equally allocated training sample schema reduces the omission error of the under-represented classes. Thus the results of investigations into the impact of sampling design on RF classification results appear to be contradictory and an analysis of the sensitivity of RF classifier to training samples is therefore recommended when using this classifier for remote sensing data classification.

The sizes of the training samples sets have been found to influence the performance of the RF classifier (Du et al., 2015; Stumpf and Kerle, 2011). Investigations to date have reported that the sizes of the training samples should represent approximately 0.25% of the total study area (Colditz, 2015). Deng and Wu (2013) used the RF classifier to classify large-scale urban impervious areas and concluded that the method performed better when a large number of samples was available. The same conclusion was reached by Du et al. (2015).

RF classification may be attractive when a small sample size is used in combination with high dimensional data inputs (Chan et al., 2012). Although RF classifier can handle high dimensional data, a number of studies have investigated various alternative methods for feature space optimization (Díaz-Uriarte and De Andres, 2006; Fassnacht et al., 2014; Karlson et al., 2015; Millard and Richardson, 2015; Stumpf and Kerle, 2011). There are a number of feature selection methods used in combination with RF classification, such as filter methods, embedded methods, and the wrapper methods (Saeys et al., 2007; Weinmann et al., 2015).

A detailed review of the existing feature selection methods can be found in Saeys et al. (2007).

Filter methods are able to identify the most relevant features independently of a specific classifier. Principal Component Analysis (PCA), Independent Component Analysis (ICA), Minimum Noise Fraction (MNF) analysis (Zhang, 2014), Wilks's Lambda stepwise discriminant analysis (Thenkabail et al., 2013), linear discriminant analysis with stepwise feature selection, and Spearman's rank order filter are just a few examples of methods that can be used to remove redundant, noisy and irrelevant variables. However, these filter methods are rarely used in combination with an RF classifier. One possible reason is that no significant differences are obtained between the RF classification results obtained by applying these pre-filtering methods and those obtained using the feature selection methods embedded in the RF classifier (Adelabu et al., 2014b; Dalponte et al., 2013). For example, Millard and Richardson (2015) applied Spearman's rank order correlation to determine pair-wise correlations between 15 important variables identified by the RF classifier (out of 28 variables). The classification accuracies obtained using correlated variables were similar to those obtained using uncorrelated variables. Adelabu et al. (2014b) evaluated the performance of the RF classifier when features filtering procedures such as one-way Analysis of Variance (ANOVA) and the procedure based on the OOB error were used. The authors reported that both features filtering procedures identified red edge and near-infrared regions as being important for identifying defoliation levels. Unlike ANOVA, OOB also identified the green region as being relevant to the classification objective, and the OOB-based feature space optimization yielded better classification results.

For embedded methods, features relevance is assessed using the learning classifier itself. It has been reported that these methods perform well, but are specifically tuned to the applied classifier (Weinmann et al., 2015). The MDA error estimation embedded in the RF classifier is frequently used to pre-select the most important variables (Karlson et al., 2015; Li et al., 2015). Díaz-Uriarte and De Andres (2006) proposed an iterative backward feature elimination procedure to reduce the number of less relevant variables, until the internal accuracy calculated on the basis of OOB error stabilises. This procedure has been incorporated within the R software (R-Development-Core-Team, 2005) as the "*varSelRF*" package, and has been successfully used to classify landslide objects from high resolution optical data such as QuickBird, IKONOS, Geoeye-1 and aerial photographs (Stumpf and Kerle, 2011) and reforested landslides from LiDAR data (Chen et al., 2014; Li et al., 2015), as well as to map urban classes from orthoimagery and LiDAR (Guan et al., 2013) and grass species from hyperspectral data (Mansour et al., 2012).

The Boruta wrapper feature selection method built into the RF algorithm (Kursa and Rudnicki, 2010) has been successfully used to reduce noisy and highly correlated variables derived from LiDAR and hyperspectral data, in order to model plant compositions and diversity in forested areas (Leutner et al., 2012). A combination of embedded and wrapper feature selection methods was reported by Chan and Paelinckx (2008). These authors examined the relevance of hyperspectral bands (HyMap data) for identifying LC classes by applying the OOB error and the wrapper method with best-first search algorithms. The bands in the vicinity of the 1.3 and 1.9 μm wavebands were identified as the most relevant variables for mapping the LC of interest.

5. Comparing the random forest classifier with other machine learning classifiers

A number of investigations have explored the differences between the RF classifier and other state-of-the art machine learning classifiers in terms of the accuracy of the classification results, the training time required (Gislason et al., 2006), and the stability of the classifiers if the training samples (Chan and Paelinckx, 2008) or the study areas (Vetrivel et al., 2015) are changed. The RF classifier was found to outperform decision tree classifiers (Ghimire et al., 2012; Gislason et al., 2006; Han et al., 2015), the Binary Hierarchical Classifier (BHC) (Ham et al., 2005), Linear Discriminant Analysis (LDA) (Shang and Chisholm, 2014), and ANN classifiers (Chan and Paelinckx, 2008) in terms of classification accuracy. The advantages of the RF classifier over SVM classifiers, object-oriented methods, and MLC are highlighted in Chutia et al. (2015). RF and SVM classifiers are equally reliable according to Dalponte et al. (2013), Ghosh et al. (2014), and Sesnie et al. (2010), with the RF classifier results being slightly better for high dimensional input data such as hyperspectral imagery (Abdel-Rahman et al., 2014; Chan et al., 2012; Shang and Chisholm, 2014). Ghosh and Joshi (2014) reported that SVM classification performs better than RF classification in Object Based Image Analysis (OBIA). SVM classification seems, however, to be more sensitive to feature selection (Li et al., 2015; Vetrivel et al., 2015) and is less user-friendly because of the number of critical parameters that have to be set (Chan et al., 2012).

Published comparisons between the RF classifier and other ensemble classifiers such as AdaBoost have reported different results. Chan and Paelinckx (2008) reported that RF and AdaBoost classifiers both yielded similar classification results but that the RF algorithm was faster to train and more stable. Miao et al. (2012) compared four basic ensemble classification methods (the bagging tree, RF, AdaBoost tree, and AdaBoost random tree methods) in investigations that concentrated on the classification of ecological zones using TM/ETM+ images. CART was selected as the individual classifier for each of these ensemble approaches. The RF and AdaBoost random tree methods achieved similar accuracies to the bagging tree and AdaBoost tree approaches, but with greater efficiency (i.e. a reduced computation burden). The results of these investigations indicated also that the AdaBoost tree and AdaBoost random tree methods yielded better classification results than the bagging tree and RF methods. Xu et al. (2014) and Waske and Braun (2009) showed that the RF method outperformed the boosting decision tree ensemble methods. However, as emphasized by Miao et al. (2012), the base classifier for the AdaBoost method was C5.0 whereas for the RF method it was CART.

In conclusion therefore, the RF classifier achieves better classification results when multi-dimensional data such as hyperspectral or multi-source data are used; it also requires the setting of fewer parameters (Chan et al., 2012; Shao et al., 2015) and is faster than the SVM classifier or other ensemble classifiers such as AdaBoost.

6. Future directions

6.1. Stability of the RF classifier

The stability of the RF classifier is an important criterion for its integration into operational settings. Previous studies have reported that the overall classification accuracy of the RF classifier decreases when the algorithm is trained on different study areas (Vetrivel et al., 2015). Juel et al. (2015) tested the transferability of RF classification models created to map vegetation from aerial images and Digital Elevation Models (DEMs), concluding that the resulting classification model was not transferable to new areas. The development of hybrid classification methodologies that integrate the RF classifier (or other mainstream supervised classifiers) with explicitly specified models that convey the objects' semantics (Du et al., 2015), and the utilization of spectral indices (which have

been shown to be more stable when applied to new study areas) might provide a solution to the aforementioned problem.

6.2. Variable importance and robustness

Determining optimized feature space is an important step toward developing consistent and operational classification systems (Corcoran et al., 2013); it may also be of assistance in designing future application-oriented sensors (Chan and Paelinckx, 2008; Nitze et al., 2015). Further investigations are required to evaluate the robustness of feature space created using filter or embedded feature selection methods. A number of robustness indices that have been proposed in published literature (such as the consistency index developed by Kuncheva, 2007 or entropy-based measures proposed by Křížek et al., 2007) can be used to assess the stability of the selected features.

There have been numerous valuable investigations into the predictive power of a variety of variables in different classification scenarios, but the results are hardly likely to benefit the remote sensing community because the identified variables are not systematically organized into online features catalogues that can be easily reused by those interested in pursuing similar investigations.

The reported investigations into RF classifications have mainly focused on per-pixel classifications, with research into the coupling of OBIA with RF classification being relatively poorly represented. The increasing number of variables computed for the image objects generated through segmentation makes OBIA classification a highly subjective and time-intensive task (Belgiu and Drăguţ, 2014; Belgiu et al., 2014a; Guan et al., 2013). OBIA might therefore benefit from using the VI measurement to optimize feature space. For example, Guan et al. (2013) examined the contributions made by variables calculated from objects delineated from LiDAR and orthoimagery data to map urban objects, and a similar approach was used by Mishra and Crews (2014) for vegetation mapping in savannah regions.

6.3. Classification accuracy evaluation

The evaluation of classification accuracy involves additional effort and time to collect validation samples. Lawrence et al. (2006) and Zhong et al. (2014) reported that the OOB error in the RF method could be used as a reliable measure of classification accuracy. This assertion needs to be further tested using a variety of datasets in different application scenarios.

6.4. Identification of outliers in training samples

The sample proximity measurement available with the RF algorithm identifies outliers in training samples. The proximity between two samples is measured by the ratio of the number of trees retained in the same terminal node to the total number of trees generated in the ensemble (Gislason et al., 2006). As mentioned by Touw et al. (2012), the use of the proximity measurement allows the identification of variable prototypes, i.e. variables that are representative of a group of samples. There is a need for further investigations into the use of proximity measurements and their possible combination with other mainstream supervised classifiers such as the SVM or ANN classifiers.

6.5. Robustness of variable importance (VI) embedded in the random forest classifier

Additional questions requiring future investigation concern the robustness of feature selection (and the resulting RF classification model) when the number of training samples is either increased or reduced (Chan and Paelinckx, 2008) or when noise is added to the features, the stability of the VI measure with iterative classifications (Mellor et al., 2015; Millard and Richardson, 2015), and the sensitivity of the *Ntree* parameter to the feature space (i.e. to the number of variables). Mellor et al. (2015) and Rodriguez-Galiano et al. (2012) reported that the correlation between trees (which is proportional to the classification error) increases as the number of variables for splitting the trees nodes increases.

6.6. Novel approaches to improving the random forest classifier

A number of different approaches have been proposed for improving the RF classifier. Menze et al. (2011) investigated RF classification using oblique model trees as base learners rather than orthogonal model trees. The authors found that a learned node model performed better than a random split model and that the oblique the RF classifier identifies variables that are not considered to be relevant variables by orthogonal RF classification (i.e. by univariate split models). Oblique RF classification has not yet been used with remote sensing data. Zhang and Suganthan (2014) proposed rotation RF classification as a new way of increasing the diversity of each tree in the RF classification by concatenating different rotation spaces into a higher space at the root node during the training phase. Junshi et al. (2014) applied rotation RF classification using various feature space transformation methods such as PCA, MNF, ICA and showed that rotation RF classification with PCA transformation outperformed the RF classifier, SVM, AdaBoost method, and "logistic regression via variable splitting and augmented Lagrangian (LORSAL)" classifiers. Du et al. (2015) used RF and rotation RF classifiers (Rodriguez et al., 2006) to classify PolSAR data using both polarimetric and spatial features (texture measurements and morphological profiles) and compared the classification results with those obtained using SVM classification. Rotation RF classification splits the variables set into subsets and applies PCA to each subset of variables in order to reconstruct a full feature space for the trees in the ensemble classifier. The authors concluded that rotation RF classification outperformed RF and SVM classifications, but at the cost of increased computational requirements. Ham et al. (2005) proposed a combination of RF with "adaptive random subspace feature selection" within a binary hierarchical classifier (BHC) to identify LC classes in seasonal swamps, occasional swamps, and drier woodland from Hyperion and AVIRIS data.

When applying RF classification on its own the classes of interest are classified without taking into account any contextual information (such as topology), which can be of considerable importance when classifying objects in complex environments. To address this problem, Niemeyer et al. (2014) integrated an RF classifier into a Conditional Random Field (CRF) framework for urban objects classification. This method yielded promising results and therefore warrants further investigation, especially in scenarios where objects with high intra-class variability are to be identified.

7. Summary

The RF classifier is less sensitive than other streamline machine learning classifiers to the quality of training samples and to overfitting, due to the large number of decision trees produced by randomly selecting a subset of training samples and a subset of variables for splitting at each tree node. The RF classifier has been shown to be suitable for classifying hyperspectral data, where the curse of dimensionality and highly correlated data pose major challenges to other available classification methodologies. Additional internal measurements provided by this classifier (such as

the VI) have also been extensively used to select the optimal variables computed from a single sensor, from several distinct sensors, or from a combination of remote sensing and geographic data. A number of authors have reported using the VI measurement to identify the most suitable seasons for identifying desired classes.

A large number of the investigations reviewed herein have looked into the sensitivity of the RF classifier to the two parameters that need to be set. The most common recommendation is to set the *nTree* parameter to 500 and *Mtry* to the square root of the number of input variables. Despite the fact that RF classification is robust to high data dimensionality, many authors have recommended an iterative backward feature elimination procedure to reduce the number of less relevant variables until the internal accuracy (calculated on the basis of the OOB error) no longer varies. Using this approach significantly increases the classification accuracy. The sensitivity of RF classification to the sampling design needs to be taken into account in order to reduce misclassification.

Acknowledgement

We are grateful to the two anonymous reviewers for providing comments and suggestions that greatly improved the article.

References

Abdel-Rahman, E.M., Mutanga, O., Adam, E., Ismail, R., 2014. Detecting Sirex noctilio grey-attacked and lightning-struck pine trees using airborne hyperspectral data, random forest and support vector machines classifiers. ISPRS J. Photogramm. Remote Sens. 88, 48–59.

Adelabu, S., Mutanga, O., Adam, E., 2014a. Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels. ISPRS J. Photogramm. Remote Sens. 95, 34–41.

Adelabu, S., Mutanga, O., Adam, E., Sebego, R., 2014b. Spectral discrimination of insect defoliation levels in mopane woodland using hyperspectral data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 177–186.

Ahmed, O.S., Franklin, S.E., Wulder, M.A., White, J.C., 2015. Characterizing stand-level forest canopy cover and height using landsat time series, samples of airborne LiDAR, and the random forest algorithm. ISPRS J. Photogramm. Remote Sens. 101, 89–101.

Bechtel, B., Daneke, C., 2012. Classification of local climate zones based on multiple earth observation data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5, 1191–1202.

Belgiu, M., Drăguț, L., 2014. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. ISPRS J. Photogramm. Remote Sens. 96, 67–75.

Belgiu, M., Drăguț, L., Strobl, J., 2014a. Quantitative evaluation of variations in rule-based classifications of land cover in urban neighbourhoods using WorldView-2 imagery. ISPRS J. Photogramm. Remote Sens. 87, 205–215.

Belgiu, M., Tomljenovic, I., Lampoltshammer, T., Blaschke, T., Höfle, B., 2014b. Ontology-based classification of building types detected from airborne laser scanning data. Remote Sens. 6, 1347–1366.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Breiman, L., 2001. Random forest. Mach. Learn. 45.

Briem, G.J., Benediktsson, J.A., Sveinsson, J.R., 2002. Multiple classifiers applied to multisource remote sensing data. IEEE Trans. Geosci. Remote Sens. 40, 2291–2299.

Cavallaro, G., Dalla Mura, M., Benediktsson, J.A., Bruzzone, L., 2015. Extended self-dual attribute profiles for the classification of hyperspectral images. IEEE Geosci. Remote Sens. Lett. 12, 1690–1694.

Chan, J.C.-W., Beckers, P., Spanhove, T., Borre, J.V., 2012. An evaluation of ensemble classifiers for mapping Natura 2000 heathland in Belgium using spaceborne angular hyperspectral (CHRIS/Proba) imagery. Int. J. Appl. Earth Obs. Geoinf. 18, 13–22.

Chan, J.C.-W., Paelinckx, D., 2008. Evaluation of random forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sens. Environ. 112, 2999–3011.

Chehata, N., Guo, L., Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. Int. Arch. Photogramm. Remote Sens. Spatial Inform. Sci. 39, 207–212.

Chen, W., Li, X., Wang, Y., Chen, G., Liu, S., 2014. Forested landslide detection using LiDAR data and the random forest algorithm: a case study of the Three Gorges, China. Remote Sens. Environ. 152, 291–301.

Chutia, D., Bhattacharyya, D.K., Sarma, K.K., Kalita, R., Sudhakar, S., 2015. Hyperspectral remote sensing classifications: a perspective survey. Trans. GIS. http://dx.doi.org/10.1111/tgis.12164.

Colditz, R., 2015. An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms. Remote Sens. 7, 9655.

Corcoran, J., Knight, J., Gallant, A., 2013. Influence of multi-source and multi-temporal remotely sensed and ancillary data on the accuracy of random forest classification of wetlands in northern Minnesota. Remote Sens. 5, 3212–3238.

Dalponte, M., Orka, H.O., Gobakken, T., Gianelle, D., Naesset, E., 2013. Tree species classification in boreal forests with hyperspectral data. IEEE Trans. Geosci. Remote Sens. 51, 2632–2645.

DeFries, R.S., Chan, J.C.-W., 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sens. Environ. 74, 503–515.

Demarchi, L., Canters, F., Cariou, C., Licciardi, G., Chan, J.C.-W., 2014. Assessing the performance of two unsupervised dimensionality reduction techniques on hyperspectral APEX data for high resolution urban land-cover mapping. ISPRS J. Photogramm. Remote Sens. 87, 166–179.

Deng, C., Wu, C., 2013. The use of single-date MODIS imagery for estimating large-scale urban impervious surface fraction with spectral mixture analysis and machine learning techniques. ISPRS J. Photogramm. Remote Sens. 86, 100–110.

Díaz-Uriarte, R., De Andres, S.A., 2006. Gene selection and classification of microarray data using random forest. BMC Bioinformatics 7, 3.

Dietterich, T., 2000. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. Mach. Learn. 40, 139–157.

Du, P., Samat, A., Waske, B., Liu, S., Li, Z., 2015. Random forest and rotation forest for fully polarized SAR image classification using polarimetric and spatial features. ISPRS J. Photogramm. Remote Sens. 105, 38–53.

Fassnacht, F.E., Neumann, C., Forster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P.K., Koch, B., 2014. Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 2547–2561.

Frazier, R.J., Coops, N.C., Wulder, M.A., Kennedy, R., 2014. Characterization of aboveground biomass in an unmanaged boreal forest using landsat temporal segmentation metrics. ISPRS J. Photogramm. Remote Sens. 92, 137–146.

Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55, 119–139.

Ghimire, B., Rogan, J., Galiano, V.R., Panday, P., Neeti, N., 2012. An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. GISci. Remote Sens. 49, 623–643.

Ghosh, A., Fassnacht, F.E., Joshi, P.K., Koch, B., 2014. A framework for mapping tree species combining hyperspectral and LiDAR data: role of selected classifiers and sensor across three spatial scales. Int. J. Appl. Earth Obs. Geoinf. 26, 49–63.

Ghosh, A., Joshi, P.K., 2014. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. Int. J. Appl. Earth Obs. Geoinf. 26, 298–311.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27, 294–300.

Guan, H., Li, J., Chapman, M., Deng, F., Ji, Z., Yang, X., 2013. Integration of orthoimagery and lidar data for object-based urban thematic mapping using random forests. Int. J. Remote Sens. 34, 5166–5186.

Guo, B., Huang, X., Zhang, F., Sohn, G., 2015. Classification of airborne laser scanning data using JointBoost. ISPRS J. Photogramm. Remote Sens. 100, 71–83.

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 J. Photogramm. Remote Sens. 66, 56–66.

Haas, J., Ban, Y., 2014. Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta. Int. J. Appl. Earth Obs. Geoinf. 30, 42–55.

Ham, J., Chen, Y., Crawford, M.M., Ghosh, J., 2005. Investigation of the random forest framework for classification of hyperspectral data. IEEE Trans. Geosci. Remote Sens. 43, 492–501.

Han, H., Lee, S., Im, J., Kim, M., Lee, M.-I., Ahn, M., Chung, S.-R., 2015. Detection of convective initiation using meteorological imager onboard communication, ocean, and meteorological satellite based on machine learning approaches. Remote Sens. 7, 9184.

Holmes, G., Donkin, A., Witten, I.H., 1994. Weka: a machine learning workbench, intelligent information systems, 1994. In: Proc. of the second Australian and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia, 29 November – 2 December, pp. 357–361.

Jin, H., Stehman, S.V., Mountrakis, G., 2014. Assessing the impact of training sample selection on accuracy of an urban classification: a case study in Denver, Colorado. Int. J. Remote Sens. 35, 2067–2081.

Juel, A., Groom, G.B., Svenning, J.-C., Ejrnæs, R., 2015. Spatial application of random forest models for fine-scale coastal vegetation classification using object based analysis of aerial orthophoto and DEM data. Int. J. Appl. Earth Obs. Geoinf. 42, 106–114.

Junshi, X., Peijun, D., Xiyan, H., Chanussot, J., 2014. Hyperspectral remote sensing image classification based on rotation forest. IEEE Geosci. Remote Sens. Lett. 11, 239–243.

Junshi, X., Wenzhi, L., Chanussot, J., Peijun, D., Guanghan, S., Philips, W., 2015. Improving random forest with ensemble of features and semisupervised feature extraction. IEEE Geosci. Remote Sens. Lett. 12, 1471–1475.

Karlson, M., Ostwald, M., Reese, H., Sanou, J., Tankoano, B., Mattsson, E., 2015. Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using landsat 8 and random forest. Remote Sens. 7, 10017.

Körting, T.S., Fonseca, L.M. Garcia, Câmara, G., 2013. GeoDMA—geographic data mining analyst. Comput. Geosci. 57, 133–145.

Křížek, P., Kittler, J., Hlaváč, V., 2007. Improving stability of feature selection methods, computer analysis of images and patterns. In: Proc. of the 12th

International Analysis on Computer Analysis of Images and Patters, Vienna, Austria, 27–29 August, pp. 929–936.
Kulkarni, V.Y., Sinha, P.K., 2012. Pruning of random forest classifiers: a survey and future directions. In: Proc. of International Conference on Data Science Engineering (ICDSE) Cochin, Kerala, India, 18–20 July, pp. 64–68.
Kuncheva, L.I., 2007. A stability index for feature selection. Artif. Int. Appl., 421–427
Kursa, M.B., Rudnicki, W.R., 2010. Feature selection with the Boruta package. J. Stat. Softw. 36.
Lawrence, R.L., Wood, S.D., Sheley, R.L., 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). Remote Sens. Environ. 100, 356–362.
Leutner, B.F., Reineking, B., Müller, J., Bachmann, M., Beierkuhnlein, C., Dech, S., Wegmann, M., 2012. Modelling forest α-diversity and floristic composition — on the added value of LiDAR plus hyperspectral remote sensing. Remote Sens. 4, 2818–2845.
Li, X., Cheng, X., Chen, W., Chen, G., Liu, S., 2015. Identification of forested landslides using LiDar data, object-based image analysis, and machine learning algorithms. Remote Sens. 7, 9705.
Liu, K., Shi, W., Zhang, H., 2011. A fuzzy topology-based maximum likelihood classification. ISPRS J. Photogramm. Remote Sens. 66, 103–114.
Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 28, 823–870.
Ma, L., Cheng, L., Li, M., Liu, Y., Ma, X., 2015. Training set size, scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery. ISPRS J. Photogramm. Remote Sens. 102, 14–27.
Mansour, K., Mutanga, O., Everson, T., Adam, E., 2012. Discriminating indicator grass species for rangeland degradation assessment using hyperspectral data resampled to AISA eagle resolution. ISPRS J. Photogramm. Remote Sens. 70, 56–65.
Marceau, D.J., Howarth, P.J., Gratton, D.J., 1994. Remote Sensing and the measurement of geographical entities in a forested environment. 1. The scale and spatial aggregation problem. Remote Sens. Environ. 49, 93–104.
Mas, J.F., Flores, J.J., 2007. The application of artificial neural networks to the analysis of remotely sensed data. Int. J. Remote Sens. 29, 617–663.
Mellor, A., Boukir, S., Haywood, A., Jones, S., 2015. Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. ISPRS J. Photogramm. Remote Sens. 105, 155–168.
Menze, B.H., Kelm, B.M., Splitthoff, D.N., Koethe, U., Hamprecht, F.A., 2011. On oblique random forests. In: Proc. of the European Conference Machine Learning and Knowledge Discovery in Databases, Athens, Greece, 5–9 September, pp. 453–469.
Miao, X., Heaton, J.S., Zheng, S., Charlet, D.A., Liu, H., 2012. Applying tree-based ensemble algorithms to the classification of ecological zones using multi-temporal multi-source remote-sensing data. Int. J. Remote Sens. 33, 1823–1849.
Millard, K., Richardson, M., 2015. On the importance of training data sample selection in random forest image classification: a case study in peatland ecosystem mapping. Remote Sens. 7, 8489.
Mishra, N.B., Crews, K.A., 2014. Mapping vegetation morphology types in a dry savanna ecosystem: integrating hierarchical object-based image analysis with random forest. Int. J. Remote Sens. 35, 1175–1198.
Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. ISPRS J. Photogramm. Remote Sens. 66, 247–259.
Naidoo, L., Cho, M.A., Mathieu, R., Asner, G., 2012. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a random forest data mining environment. ISPRS J. Photogramm. Remote Sens. 69, 167–179.
Niemeyer, J., Rottensteiner, F., Soergel, U., 2014. Contextual classification of lidar data and building object detection in urban areas. ISPRS J. Photogramm. Remote Sens. 87, 152–165.
Nitze, I., Barrett, B., Cawkwell, F., 2015. Temporal optimisation of image acquisition for land cover classification with random forest and MODIS time-series. Int. J. Appl. Earth Obs. Geoinf. 34, 136–146.
Pal, M., 2005. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 26, 217–222.
Pedergnana, M., Marpu, P.R., Mura, M.D., Benediktsson, J.A., Bruzzone, L., 2013. A novel technique for optimal feature selection in attribute profiles based on genetic algorithms. IEEE Trans. Geosci. Remote Sens. 51, 3514–3528.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Peerbhay, K.Y., Mutanga, O., Ismail, R., 2015. Random forests unsupervised classification: the detection and mapping of solanum mauritianum infestations in plantation forestry using hyperspectral data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 3107–3122.
R-Development-Core-Team, 2005. R: A Language and Environment for Statistical Computing. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2013. <http://www.r-project.org/>.
Räsänen, A., Rusanen, A., Kuitunen, M., Lensu, A., 2013. What makes segmentation good? A case study in boreal forest habitat mapping. Int. J. Remote Sens. 34, 8603–8627.
Reese, H., Nyström, M., Nordkvist, K., Olsson, H., 2014. Combining airborne laser scanning data and optical satellite data for classification of alpine vegetation. Int. J. Appl. Earth Obs. Geoinf. 27 (Part A), 81–90.
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 J. Photogramm. Remote Sens. 67, 93–104.
Rodriguez, J.J., Kuncheva, L.I., Alonso, C.J., 2006. Rotation forest: a new classifier ensemble method. IEEE Trans. Pattern Anal. Mach. Int. 28, 1619–1630.
Saeys, Y., Inza, I., Larrañaga, P., 2007. A review of feature selection techniques in bioinformatics. Bioinformatics 23, 2507–2517.
Schapire, R.E., 1990. The strength of weak learnability. Mach. Learn. 5, 197–227.
Schmidt, A., Niemeyer, J., Rottensteiner, F., Soergel, U., 2014. Contextual classification of full waveform lidar data in the Wadden Sea. IEEE Geosci. Remote Sens. Lett. 11, 1614–1618.
Schwarz, D.F., König, I.R., Ziegler, A., 2010. On safari to random jungle: a fast implementation of random forests for high-dimensional data. Bioinformatics 26, 1752–1758.
Sesnie, S.E., Finegan, B., Gessler, P.E., Thessler, S., Ramos Bendana, Z., Smith, A.M.S., 2010. The multispectral separability of Costa Rican rainforest types with support vector machines and random forest decision trees. Int. J. Remote Sens. 31, 2885–2909.
Shang, X., Chisholm, L.A., 2014. Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 2481–2489.
Shao, Y., Campbell, J.B., Taff, G.N., Zheng, B., 2015. An analysis of cropland mask choice and ancillary data for annual corn yield forecasting using MODIS data. Int. J. Appl. Earth Obs. Geoinf. 38, 78–87.
Stefanski, J., Mack, B., Waske, B., 2013. Optimization of object-based image analysis with random forests for land cover mapping. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6, 2492–2504.
Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using random forests. Remote Sens. Environ. 115, 2564–2577.
Sun, L., Schulz, K., 2015. The improvement of land cover classification by thermal remote sensing. Remote Sens. 7, 8368.
Thenkabail, P.S., Mariotto, I., Gumma, M.K., Middleton, E.M., Landis, D.R., Huemmrich, K.F., 2013. Selection of hyperspectral narrowbands (HNBs) and composition of hyperspectral twoband vegetation indices (HVIs) for biophysical characterization and discrimination of crop types using field reflectance and hyperion/EO-1 data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6, 427–439.
Topouzelis, K., Psyllos, A., 2012. Oil spill feature selection and classification using decision tree forest on SAR image data. ISPRS J. Photogramm. Remote Sens. 68, 135–143.
Touw, W.G., Bayjanov, J.R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., van Hijum, S.A., 2012. Data mining in the life sciences with random forest: a walk in the park or lost in the jungle? Briefings Bioinform., bbs034
Trimble, 2013. eCognition Developer 8.7.2. Reference Book, München.
Tsutsumida, N., Comber, A.J., 2015. Measures of spatio-temporal accuracy for time series land cover data. Int. J. Appl. Earth Obs. Geoinf. 41, 46–55.
Uhlmann, S., Kiranyaz, S., 2014. Integrating color features in polarimetric SAR image classification. IEEE Trans. Geosci. Remote Sens. 52, 2197–2216.
Vetrivel, A., Gerke, M., Kerle, N., Vosselman, G., 2015. Identification of damage in buildings based on gaps in 3D point clouds from very high resolution oblique airborne images. ISPRS J. Photogramm. Remote Sens. 105, 61–78.
Wang, H., Zhao, Y., Pu, R., Zhang, Z., 2015. Mapping *Robinia pseudoacacia* forest health conditions by using combined spectral, spatial, and textural information extracted from IKONOS imagery and random forest classifier. Remote Sens. 7, 9020.
Waske, B., Braun, M., 2009. Classifier ensembles for land cover mapping using multitemporal SAR imagery. ISPRS J. Photogramm. Remote Sens. 64, 450–457.
Waske, B., van der Linden, S., Oldenburg, C., Jakimow, B., Rabe, A., Hostert, P., 2012. ImageRF – a user-oriented implementation for remote sensing image analysis with random forests. Environ. Modell. Softw. 35, 192–193.
Weinmann, M., Jutzi, B., Hinz, S., Mallet, C., 2015. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. ISPRS J. Photogramm. Remote Sens. 105, 286–304.
Woodcock, C.E., Strahler, A.H., Jupp, D.L.B., 1988. The use of variograms in remote sensing: I. Scene models and simulated images. Remote Sens. Environ. 25, 323–348.
Zhang, C., 2014. Combining hyperspectral and LiDAR data for vegetation mapping in the Florida Everglades. Photogramm. Eng. Remote Sens. 80, 733–743.
Zhang, H., Wang, M., Chen, X., 2009. Willows: a memory efficient tree and forest construction package. BMC Bioinformatics 10, 130.
Zhang, L., Suganthan, P.N., 2014. Random forests with ensemble of feature spaces. Pattern Recogn. 47, 3429–3437.
Zhong, L., Gong, P., Biging, G.S., 2014. Efficient corn and soybean mapping with temporal extendability: a multi-year experiment using Landsat imagery. Remote Sens. Environ. 140, 1–13.
Xu, L., Li, J., Brenning, A., 2014. A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery. Remote Sens. Environ. 141, 14–23.