RANDOM FORESTS FOR CLASSIFYING MULTI-TEMPORAL SAR DATA

Björn Waske [(1)], Vanessa Heinzel [(1)], Matthias Braun [(1)], Gunter Menz [(2)]

[(1)]Center for Remote Sensing of Landsurfaces (ZFL),University of Bonn,
Walter-Flex-Str. 3, 53113 Bonn, Germany, bwaske@uni-bonn.de
[(2)]Remote Sensing Research Group (RSRG), University of Bonn,
Meckenheimer Allee 166, 53115 Bonn, Germany

ABSTRACT

The accuracy of supervised land cover classifications depends on several factors like the chosen algorithm, adequate training data and the selection of features. In regard to multi-temporal remote sensing imagery statistical classifier are often not applicable. In the study presented here, a Random Forest was applied to a SAR data set, consisting of 15 acquisitions. A detailed accuracy assessment shows that the Random Forest significantly increases the efficiency of the single decision tree and can outperform other classifiers in terms of accuracy. A visual interpretation confirms the statistical accuracy assessment. The imagery is classified into more homogeneous regions and the noise is significantly decreased. The additional time needed for the generation of Random Forests is little and can be justified. It is still a lot faster than other state-of-the-art classifiers.

1. INTRODUCTION

In context of land cover classification of regions that are dominated by agricultural land use, mono-temporal approaches are often inefficient, due to great temporal differences in the crop phenology. Multi-temporal techniques are more adequate for this purpose and can improve the classification accuracy. On the other hand the efficiency of optical imagery is often limited by weather conditions and the generation of an adequate time series can be realized best by synthetic aperture radar (SAR). In regard to upcoming missions, with increased revisit times and better spatial resolutions multi-temporal concepts become more attractive, particularly for operational monitoring systems.

However, data sets with high temporal and spatial resolution might become very large and complex. In addition such imagery contains noise (i.e., speckle), irrelevant information and unnecessary details. Regarding this, the use of an adequate classifier algorithm is essential. Conventional statistical approaches as the maximum likelihood classifier are not applicable for multi-temporal imagery, because in most cases the data cannot be modeled by an appropriate statistical data model. Thus, non-parametric techniques like e.g., neural networks, support vector machines and self-learning decision trees seem more appropriate in this context. These methods are not constrained to prior assumptions on the distribution of input data as is the case for the maximum likelihood classifier and seems more suited for classifying time series.

Decision trees are applied successfully to remote sensing imagery. By producing efficient rules a decision tree is successively partitioning the training data into an increasing number of smaller homogenous regions (i.e., classes). A set of rules at each node is leading to the final leaf, i.e., the land cover class. In contrast to neural networks and support vector machines the training time of decision trees is very low and their handling is rather simple. This makes decision trees particularly interesting in regard to operational classifier systems. Besides the classifier algorithm the training data and input features (e.g., a specific image acquisition) have a dominant impact on the performance of the supervised classification [1],[2]. On the other hand the availability of ground truth data and remote sensing imagery is often limited and can not be influenced by the user. Furthermore neither the training samples nor the selected features can be assumed to be ideal for a reliable training process.

In several studies the accuracy is increased by multiple classifier ensembles. In contrast to a conventional single classifier that generates one output a classifier ensemble consists of several independent outputs (see Figure 1). To generate a classifier ensemble, one aspect of the input data, which the base-classifier is sensitive to, needs to be modified between individual training procedures. It is assumed that each individual classifier produces independent errors, which are not produced by the majority of the other classifiers. The improvement of the accuracy depends on the independency of the different classifiers. Afterwards the outputs of all classifiers are combined for a final result.

Boosting is a well known method for the creation of classifier ensembles [3],[4], which have been applied successfully to time series of multispectral imagery. By boosting the distribution of the training samples is iteratively changed during the classifier training. In the initial classifier training phase, all samples are equally weighted. Afterwards misclassified samples are assigned a higher weight than those classified correctly and the next classifier in the ensemble is based on the modified training data.

Proc. 'Envisat Symposium 2007', Montreux, Switzerland
23–27 April 2007 (ESA SP-636, July 2007)

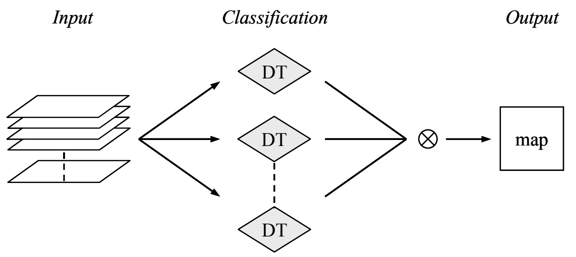

Figure 1. Schematic diagram of a classifier ensemble

The selection of feature subsets (i.e., specific images of a time series) is another effective method for the generation of a classifier ensemble. In contrast to boosting the training samples remain unchanged by this approach. In Waske *et al.* [5] a decision tree ensemble that is based on *random feature selection* is applied successfully to a set of multi-temporal SAR data. The overall accuracy is significantly increased compared to the accuracy achieved by a single decision tree.

In Breiman's Random Forests [6] the training samples as well as the input features are modified. For the training phase of each decision tree within the ensemble a set of training samples is selected randomly. Furthermore at each split node of an individual tree a random feature selection is performed. Afterwards the various outputs are combined to a final result, using a simple majority vote. The approach was applied successfully to hyperspectral imagery, using a limited training sample set [7]. Gislason *et al.* [8] have applied the method to a multisource data set, consisting of Landsat MSS data and topographical data. In this case the approach performs better than a single decision tree and comparable to other ensembles methods, whereas their computation time of the Random Forests is much faster. Pal [9] has used the method for the classification of Landsat ETM+ data from an agricultural region.

Regarding these results, it seems worthwhile to apply the Random Forests on a multi-temporal data set. In the presented study the concept is used for classifying a time series, consisting of ENVISAT ASAR imagery from an agricultural area. The classification results are compared with other parametric and non-parametric methods, as a boosted decision tree and a maximum likelihood classifier.

2. METHODS

Our test site is located near Bonn in the German state of North Rhine-Westphalia (NRW). The flat landscape is predominantly used by agriculture, with cereals and sugar beets as the main crops. The data set includes 15 Envisat ASAR PRI images between the period of January and November, 2005. The imagery contains alternating polarization and image mode data from different tracks and swaths (see Table 1 and Figure 2).

Table 1. Multi-temporal SAR data

Date	Swath/Track	Polariz.	Orbit
6-Jan.	2 / 337	VV	des
10-Feb.	2 / 337	VV	des
12-Apr.	6 / 208	HH / HV	des
21-Apr.	2 / 337	VV	des
26-May.	2 / 337	VV	des
30-Jun.	2 / 337	VV	des
10-Jul.	2 / 487	HH / HV	asc
22-Jul.	7 / 158	HH / HV	asc
4-Aug.	2 / 337	VV	des
14-Aug.	2 / 487	HH / HV	asc
8-Sep.	2 / 337	VV	des
18-Sep.	2 / 487	HH / HV	asc
30-Sep.	7 / 158	HH / HV	asc
13-Oct.	2 / 337	VV	des
17-Nov.	2 / 337	VV	des

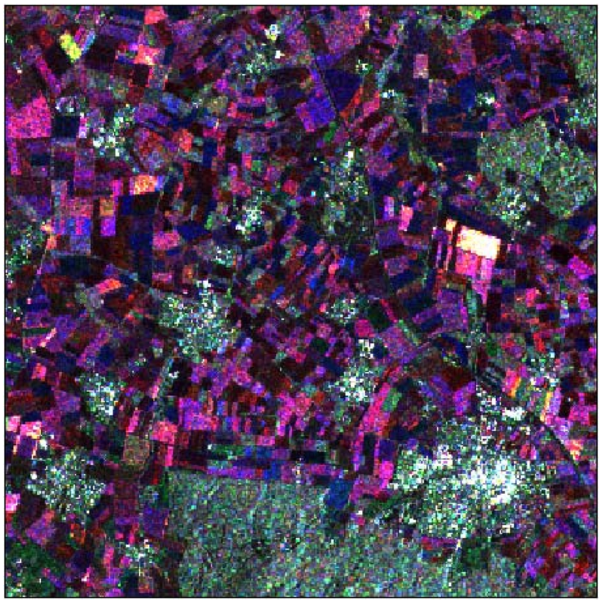

Figure 2. Multi-temporal Envisat ASAR composite

For the derivation of the calibrated backscatter intensity a common procedure was performed [10]. For the noise reduction a multi-temporal speckle filter was applied [11]. The data set was co-registered and terrain corrected, using orbital information and a digital elevation model. For the classifier training and the accuracy assessment an extensive ground truth campaign was conducted in summer 2005. For the classifier training a sample set was generated using equalized random sampling and detailed ground truth data as reference information. In doing so a training set was generated that contains 500 pixel per land cover class: *arable crops, cereals, forest, grassland, orchards,* *canola, root crops and urban*. Using the same method as before, an independent validation set of 1000 pixels per class was generated. An adequate feature subset size (i.e., number of selected images at each tree node) was chosen and a Random Forest was generated, using [12]. In addition the decision tree algorithm *C4.5* [13] was used to create a common decision tree and a boosted decision tree.

3. RESULTS

The accuracy assessment shows that the two ensembles, the boosted DT and the Random Forest, significantly improve the overall accuracies compared to a single decision tree and a maximum likelihood classifier (Table 2). The accuracy of the simple decision tree is 54.0%, whereas the boosted DT and the Random Forest achieve much higher overall accuracies of 71.0% and 72.5% respectively.

Table 2. Overall accuracies [%]

Classifier Algorithm	Overall Accuracy
Maximum Likelihood	49.9
Decision Tree	54.0
DT (boosting)	71.0
Random Forest	72.5

As the overall accuracy, the assessment of the producer and user accuracies shows the positive impact of the Random Forest, which outperforms the common decision tree as well as the boosted decision tree in the most cases (Table 3 and Table 4). The accuracies for *arable crops* and *orchards* are lower, compared to the user and producer accuracies achieved for other land cover classes. A reason for this could be the variability within the class *arable crops*. In general the ground beneath the relatively clear *orchards* is covered by grassland. Hence the class might appear as a mixture between grassland and forest. In contrast to this, land cover classes as *cereals* and *root crops* are less heterogeneous.

Table 3. Producer Accuracies [%], bold numbers indicates best results

Land cover class	Decision Tree	DT (boosting)	Random Forest
Arable crops	47.6	61.8	**66.8**
Cereals	61.7	76.4	**77.0**
Forest	63.6	83.4	**85.7**
Grassland	62.2	79.1	**80.7**
Orchard	37.7	55.0	**58.6**
Canola	54.7	**73.1**	71.7
Root crops	49.1	**72.7**	71.2
Urban	56.5	66.8	**68.6**

Table 4. User Accuracies [%], bold numbers indicates best results

Land cover class	Decision Tree	DT (boosting)	Random Forest
Arable crops	40.3	**66.0**	63.0
Cereals	57.4	**73.4**	73.1
Forest	58.6	70.0	**70.8**
Grassland	65.5	74.3	**76.3**
Orchard	39.6	61.2	**62.2**
Canola	64.3	76.7	**80.4**
Root crops	54.6	67.6	**73.3**
Urban	56.1	79.6	**84.5**

The visible assessment of the output maps shows the advantage of the classifier ensembles (see Figure 3). The map from the single decision tree shows the main structures of the classified area. On the other hand areas that are homogeneous in reality (e.g., field plots) appear noisy in the classification map. Under some circumstances the true classes are hard to define (i.e., confusion) and boundaries between natural features seem blurred. The visual interpretation of the classification output of the Random Forests is in accordance with the statistical accuracy assessment. The noise is significantly reduced and consequently the confusion between land cover classes is decreased. Almost all pixels within a homogeneous region are assigned to the same land cover class.

4. CONCLUSION

The study clearly shows that results from classifier ensembles are superior to those from a simple decision tree. Regardless of ensemble concept a higher accuracy is achieved. The Random Forest outperforms the other concepts in terms of the overall and class-specific accuracies. The overall processing time of the Random Forests – including the training of several classifiers within the ensemble and the final majority vote – is below those of the sophisticated approaches like neural networks or support vector machines.

The visual assessment confirms the positive impact of classifier ensembles: the degree of noise is significantly decreased and the image is classified into more homogeneous areas. Perhaps, the differences between the classification maps from a conventional (single) decision tree and a Random Forest are comparable to the differences between pixel-based and object-based classification results.

The approach taken thus appears very well suited for SAR data. Looking at the remaining errors, the confusion between the classes forest and urban is particularly obvious. In subsequent studies the influence of data set size and individual land cover classes will be investigated to further optimize the concept for future applications. The good results for separating agricultural classes, which are hard to be described by mono-temporal analyses, underline the importance of multi-temporal approaches in this context. In several other studies the classification accuracy is increased by multi-sensor imagery. For future work the data set will thus be extended by optical data from the same growing period.

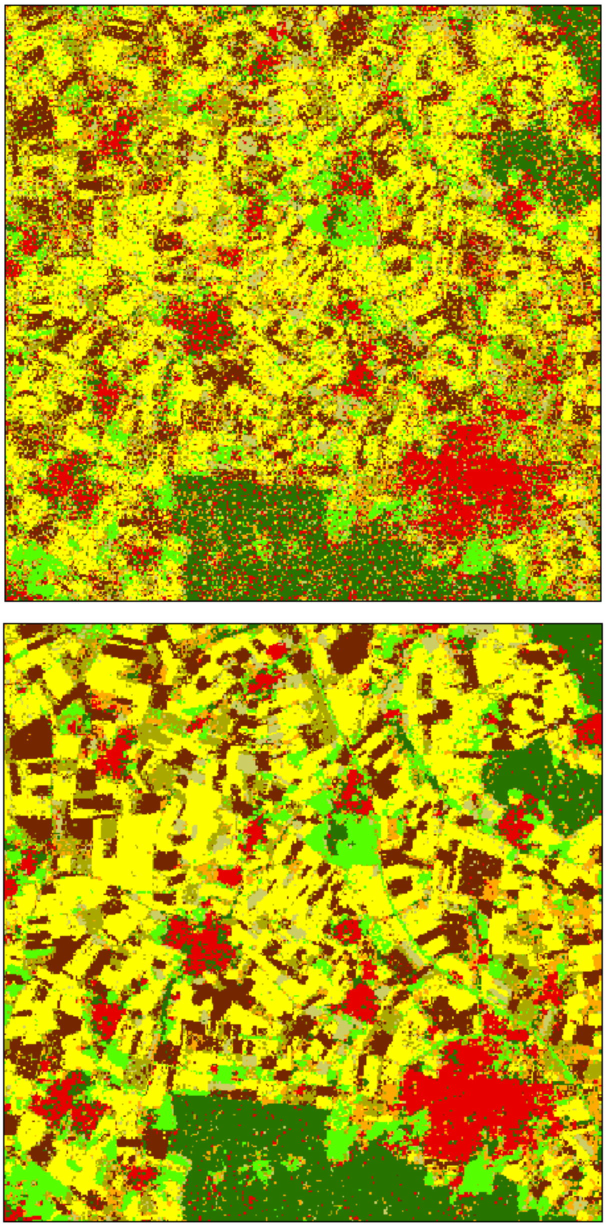

Figure 3, Classification result, using a simple DT (top) and a Random Forest (bottom)

ACKNOWLEDEGMENTS

The authors wish to thank the European Space Agency for providing remote sensing data through a CAT 1 proposal (C1.3115). The images were acquired within the ENVILAND research project (FKZ 50EE0404), funded by the German Aerospace Center (DLR) and Federal Ministry of Economics and Technology (BMWi).

REFERENCES

1. Foody, G.M. & Arora, M.K. (1997). An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *Intern. Journal of Remote Sensing* **18**, 799-810.
2. Pal, M. & Mather, P.M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment* **86**, 554-565.
3. Breiman, L. (1996). Bagging predictors. *Machine Learning*, **24**, 123-140.
4. Freund, Y. & Schapire, R.E (1996). Experiments with a new boosting algorithm. *Proceedings 13th Intern. Confe. Machine Learning*.
5. Waske, B. Schiefer, S. & Braun, M. (2006). Random feature selection for decision tree classification of multi-temporal SAR data. *Proceedings, IGARSS 2006.*
6. Breiman, L. (2001). Random Forests. *Machine Learning*, **45**, 5-32.
7. Ham, J., Chen, Y. & Crawford, M. (2005). Investigation of the Random Forest Framework for Classification of Hyperspectral Data. *IEEE Trans. on Geosciences and Remote Sensing*, **43**, 492-501.
8. Gislason, P.O., Benediktsson, J.A. & Sveinsson, J.R. (2006). Random Forests for land cover classification. *Pattern Recog. Letters*, **27**, 294-300.
9. Pal, M. (2005). Random Forest classifier for remote sensing classification. *Intern. Journal of Remote Sensing*, **26**, 217-222.
10. Laur, H., Bally, P., Meadows, P., Sanchez, J., Schaettler, B:, Lopinto, E. & Esteban, D. (2002). Derivation of the backscattering coefficient σ0 in ESA ERS SAR PRI products. *ESA Document ES-TN-RE-PM-HL09*, Issue 2, Rev. 5d.
11. Quegan, S., Le Toan, T., Yu, J.J., Ribbes, F. & Floury, N. (2000). Multitemporal ERS SAR analysis applied to forest mapping, *IEEE Trans. on Geosciences and Remote Sensing,* **38,** 741-753.
12. Breiman, L. and Culter (2004): http://www.stat.berkeley.edu/~breiman/RandomForests/cc_software.htm
13. Quinlan, J.R (1993). *Programs for Machine Learning*. San Fransisco, CA: Morgan Kaufmann.