

# THE ENMAP CONTEST: DEVELOPING AND COMPARING CLASSIFICATION APPROACHES FOR THE ENVIRONMENTAL MAPPING AND ANALYSIS PROGRAMME – DATASET AND FIRST RESULTS

A. Ch. Braun<sup>a</sup>, M. Weinmann<sup>b\*</sup>, S. Keller<sup>b</sup>, R. Müller<sup>c</sup>, P. Reinartz<sup>c</sup>, S. Hinz<sup>b</sup>

<sup>a</sup> Institute of Regional Science, Karlsruhe Institute of Technology (KIT)

Reinhard-Baumeister-Platz 1, 76131 Karlsruhe, Germany - andreas.ch.braun@kit.edu

<sup>b</sup> Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT)  
Englerstr. 7, 76131 Karlsruhe, Germany - {martin.weinmann, sina.keller, stefan.hinz}@kit.edu

<sup>c</sup> Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR)

82234 Wessling, Germany - {rupert.mueller, peter.reinartz}@dlr.de, \* Corresponding Author

## Commission III, ICWG III/VII

**KEY WORDS:** Hyperspectral Imaging, EnMAP, Classification, Benchmark Datasets, Evaluation Methodology

### ABSTRACT:

The Environmental Mapping and Analysis Programme EnMAP is a hyperspectral satellite mission, supposed to be launched into space in the near future. EnMAP is designed to be revolutionary in terms of spectral resolution and signal-to-noise ratio. Nevertheless, it will provide a relatively high spatial resolution also. In order to exploit the capacities of this future mission, its data have been simulated by other authors in previous work. EnMAP will differ from other spaceborne and airborne hyperspectral sensors. Thus, the assumption that the standard classification algorithms from other sensors will perform best for EnMAP as well cannot be upheld since proof. Unfortunately, until today, relatively few studies have been published to investigate classification algorithms for EnMAP. Thus, the authors of *this* study, who have provided some insights into classifying simulated EnMAP data before, aim to encourage future studies by opening the EnMAP contest. The EnMAP contest consists in a benchmark dataset provided for algorithm development, which is presented herein. For demonstrative purposes, this report also represents two classification results which have already been realized. It furthermore provides a roadmap for other scientists interested in taking part in the EnMAP contest.

## 1 INTRODUCTION

The Environmental Mapping and Analysis Programme, acronym EnMAP, is a spaceborne hyperspectral sensor to be launched during the forthcoming years (Kaufmann et al., 2008; Stuffer et al., 2009). EnMAP is designed as an imaging pushbroom hyperspectral sensor mainly based on modified existing or pre-developed technology. EnMAP offers a spectral range provided by two instruments from 420 nm to 1000 nm (VNIR) and from 900 nm to 2450 nm (SWIR). One important property is the high radiometric resolution and stability in both instruments. Its swath width is 30 km at a spatial resolution of 30 m × 30 m, which of course is high for a spaceborne hyperspectral instrument but low in comparison to airborne instruments. EnMAP will have a fast target revisit of only 4 days (Stuffer et al., 2007; Kaufmann et al., 2006).

Since the launch of EnMAP has not been realized yet, data similar to those expected to be produced by EnMAP have to be simulated. Sophisticated approaches to simulate EnMAP data have been published by Guanter et al. (2009) and Segl et al. (2010, 2012). One dataset simulated by these approaches is the EnMAP Alpine Foreland dataset, showing the Ammersee region in Bavaria, Germany.

Due to these properties, EnMAP data will differ from other hyperspectral data available to users of remote sensing datasets. EnMAP will be the first instrument to provide a radiometric quality largely comparable to airborne instruments (especially in terms of signal-to-noise values) but with a spatial resolution comparable to Landsat data. Since EnMAP will neither be similar to airborne hyperspectral sensors (like HyMap, for instance) nor to other spaceborne instruments (like Hyperion, for instance), it cannot be generally assumed that classification methods appropriate for such instruments will work well for EnMAP too (Braun et al.,

2012). Hence, research is needed to develop appropriate classification techniques even before the launch of EnMAP.

In order to stimulate research on high performance classification, this paper provides a simulated EnMAP benchmark dataset – derived from the EnMAP Alpine Foreland data, which cover 900 km<sup>2</sup> in a 1000 × 1000 pixel image. The dataset comprises 20 land use classes which are spectrally relatively similar and intricate to classify. Evaluation is performed on the basis of overall accuracy, completeness, correctness and quality. Besides describing the benchmark dataset, two results using state-of-the-art classifiers are presented. The results are based on the use of a Support Vector Machine (Cortes and Vapnik, 1995) and a Random Forest (Breiman, 2001).

After presenting the data and results on the conference, the benchmark dataset will be made available on the homepage of the authors. From there, it can be downloaded by further researchers, who will develop their approaches, compare them among one another and publish the results in future publications. This EnMAP contest will provide insight into best practices for EnMAP data classification. It will be helpful to the sensors developers and operators, because the datasets can be delivered with helpful information on data exploitation to interested users. It will furthermore be helpful to users in order to extract better results from their data. Finally, the EnMAP contest will be scientifically interesting by showing common points and differences between traditional hyperspectral classification and classification specifically designed for EnMAP data.

## 2 RELATED WORK

Few approaches have been published which provide classification results on simulated EnMAP data. This research gap is especially

deplorable, since intense efforts are being undertaken to develop an EnMAP box which could integrate such approaches and make them readily available to future users (Held et al., 2012). Braun et al. (2012) compare three different state-of-the-art kernel-based classifiers (Support Vector Machine, Import Vector Machine, Relevance Vector Machine) on simulated EnMAP data, concluding that Import Vector Machine and Relevance Vector Machine outperform the Support Vector Machine. They furthermore outline the particular differences of the three methods and point out, why a combination of the three methods could even enhance performance<sup>1</sup>. Dörnhöfer and Oppelt (2015) produce a bio-optical model to analyse profundity and benthos of the coast of Helgoland, Germany. They use simulated EnMAP scenes to produce and test the model, concluding that EnMAP reliably detects micro-variations of structures and that water remote sensing will benefit from EnMAP's properties. Schwieder et al. (2014) use simulated EnMAP data for mapping shrub cover fraction in Southern Portugal, comparing three machine learning techniques (Support Vector Regression, Random Forest, Partial Least Squares Regression). Support Vector Regression performed best and thus, EnMAP and Support Vector Regression is attributed a great potential in quantifying fractional vegetation cover and monitoring gradual land use change processes. Bracken (2014) used EnMAP data to estimate soil erosion in semi arid Mediterranean environments. The Ph.D. thesis is one of the first to fully document the benefits of the new mission for a relevant environmental problem. Faßnacht et al. (2011) present a method to automatically extract tree covered areas from several hyperspectral datasets, including EnMAP. This method is based on the extended Normalized Difference Vegetation Index (NDVI).

In order to stimulate further research on classifying EnMAP data, a benchmark dataset will be introduced in the following. The dataset will be made available to the research community and published results will be regularly compared.

### 3 DATASET

This section introduces the EnMAP contest dataset. The dataset is based on the simulated EnMAP Alpine Foreland image, provided by Guanter et al. (2009). The colleagues produce a  $1000 \times 1000$  pixel datasets, covering  $30 \times 30$  km regions. Hence, the ground sampling distance is 30 m. The datasets cover the 420 to 2450 nm spectral range at a varying spectral sampling of 6.5-10 nm. The images consist of 244 simulated spectral channels. Figure 1 shows a near natural color visualization of the 244 channel simulated EnMAP dataset Alpine Foreland. The image depicts the area around the Ammersee in Bavaria, Germany. Note the diversity of different agricultural, vegetation, urban and industrial, and water classes (cf. Guanter et al. (2009)).

This diversity is represented in the EnMAP contest dataset. The EnMAP contest dataset comprises 20 different land use classes, and it represents a typical scenario for modern remote sensing: high accuracy is aimed for, whereas only small training data sets are provided. The  $N_C = 20$  classes are defined to be spectrally very similar and thus, provide a dataset which is difficult to classify, cf. Figure 2. The mean spectra are shown in Figure 3 to visualize spectral similarity. More specifically, classes have been defined on the screen by focussing on visual differences in the images (considering several channel combinations) but also by checking pixels' individual spectra. Then, typical areas have been assigned a class label  $l \in \{l^1, \dots, l^{N_C}\}$ , where  $N_C$  represents the number of classes and  $N_C = 20$  for this dataset. From

<sup>1</sup>In Braun et al. (2014), such a combination is presented, albeit on other data than on EnMAP.

these areas, a random subsample has been drawn for each class, since the number of pixels within the entire areas was too large to be exploited conveniently with state-of-the-art classifiers (which, while in the developing phase, tend to be time consuming given larger training numbers). The pixels from this subsample within these areas have afterwards been randomly split a *second time*, this time into a training set  $\mathcal{X}$  and a test set  $\mathcal{Y}$  (see also, Section 5). The splitting for each class is similar, 70% of the pixels are in  $\mathcal{X}$  and 30% are in  $\mathcal{Y}$ . In total,  $\mathcal{X}$  comprises 2617 pixels and  $\mathcal{Y}$  1124 pixels.

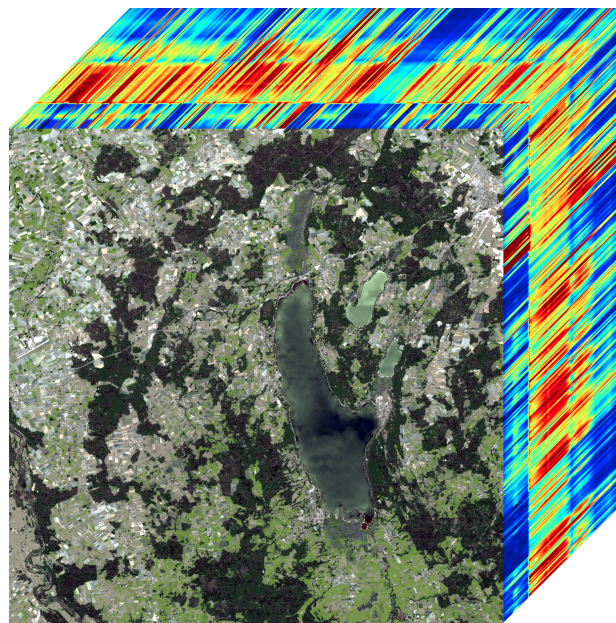


Figure 1. Near natural color visualization of the 244 channel simulated EnMAP dataset Alpine Foreland, showing the area around the Ammersee in Bavaria, Germany. Dataset produced by Guanter et al. (2009), courtesy of Dr. K. Segl. Larger image available at: [www.ipf.kit.edu/code.php](http://www.ipf.kit.edu/code.php)

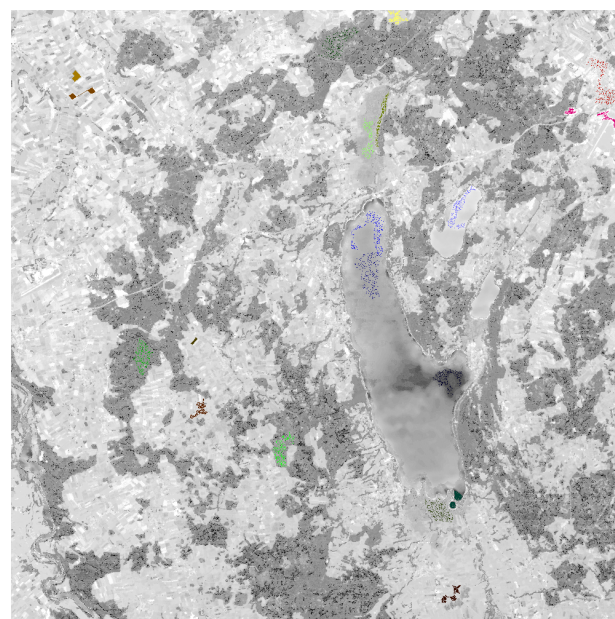


Figure 2. Location of labelled data within the simulated EnMAP dataset from Figure 1. Larger image available at: [www.ipf.kit.edu/code.php](http://www.ipf.kit.edu/code.php)

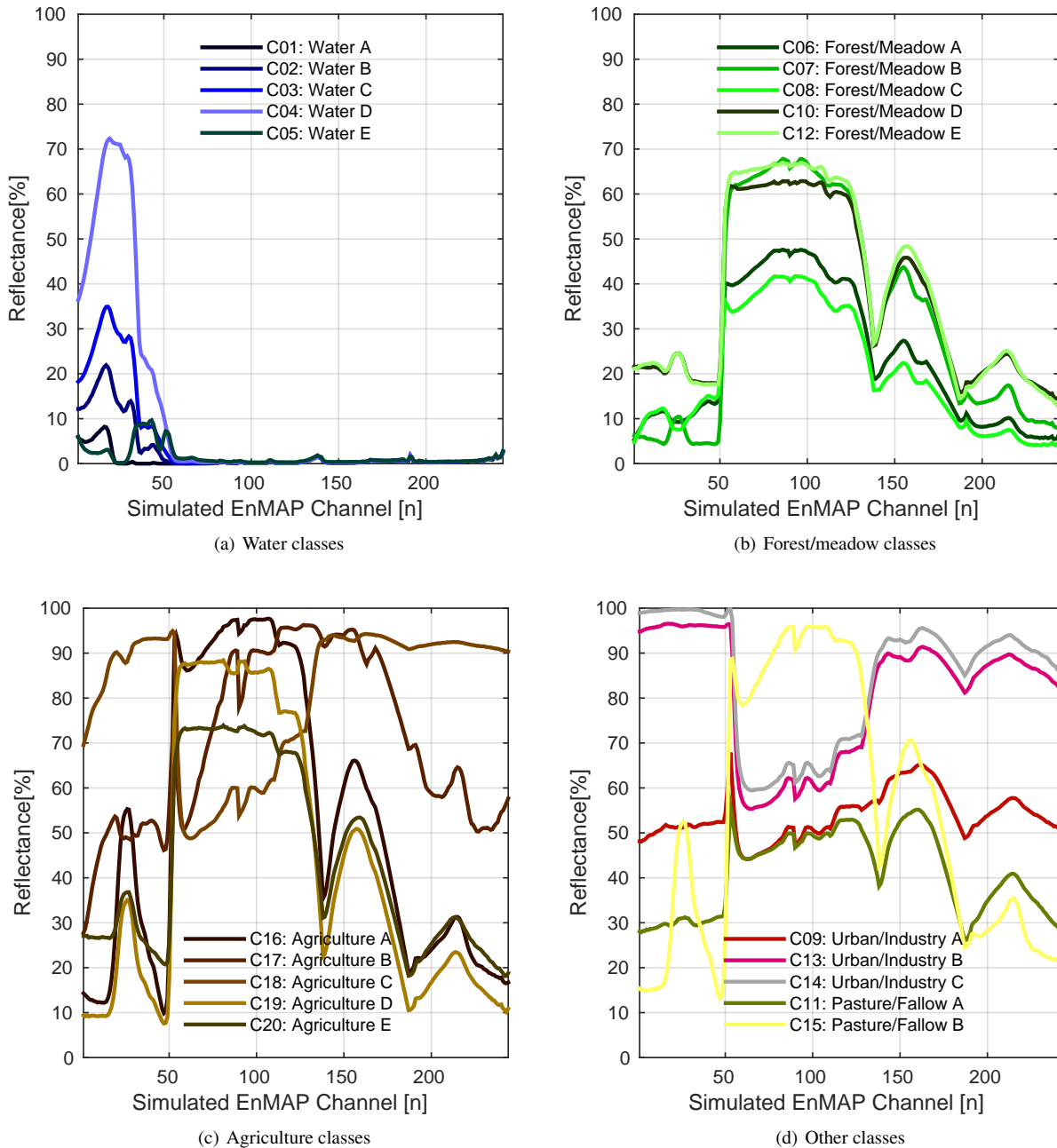


Figure 3. Average spectra for the 20 land use classes of the EnMAP contest dataset: (a) water classes, (b) forest/meadow classes, (c) agriculture classes, (d) other classes.

#### 4 EVALUATION

Our evaluation focuses on a comparison of the performance of different approaches for classifying the provided dataset. For each approach, the test set is classified, and the resulting labels are compared to the reference labels on a per-pixel basis. We determine the respective confusion matrices and derive a measure indicating the overall performance as well as per-class measures indicating the class-wise performance. More specifically, the confusion matrix  $\mathbf{C} = [c_{ij}]$  is defined in a way that the reference is given in row direction, while the prediction is given in column direction. Based on the confusion matrix, we consider the *overall accuracy*

$$\text{overall accuracy} = \frac{\sum_i c_{ii}}{\sum_i \sum_j c_{ij}} \quad (1)$$

in order to argue about the overall effectiveness of a specific approach. For the class-wise considerations, we assign the  $i$ -th class the following measures:

- True Positive (TP): 
$$TP_i = c_{ii} \quad (2)$$

- False Positive (FP): 
$$FP_i = \sum_{j, j \neq i} c_{ij} \quad (3)$$

- False Negative (FN): 
$$FN_i = \sum_{j, j \neq i} c_{ji} \quad (4)$$

- True Negative (TN):

$$TN_i = \sum_{i,j} c_{ij} - TP_i - FP_i - FN_i \quad (5)$$

Based on these measures, we derive the *completeness (recall)* given by

$$\text{completeness}_i = \frac{TP_i}{TP_i + FN_i} \quad (6)$$

as well as *correctness (precision)* given by

$$\text{correctness}_i = \frac{TP_i}{TP_i + FP_i} \quad (7)$$

and we furthermore derive the measure of *quality* which represents a compound metric indicating a good trade-off between omission and commission errors (Heipke et al., 1997):

$$\text{quality}_i = \frac{TP_i}{TP_i + FP_i + FN_i} \quad (8)$$

$$= \frac{1}{\text{completeness}_i^{-1} + \text{correctness}_i^{-1} - 1} \quad (9)$$

Further evaluation measures, that better capture the topological properties of the classified regions (see e.g. (Weidner, 2008)), could be introduced at a later stage.

## 5 CLASSIFICATION

In the scope of this paper, we focus on a supervised classification of individual pixels by using the training data in order to train a classifier which should afterwards be able to generalize to new, unseen data. Introducing a formal description, the training set  $\mathcal{X} = \{(\mathbf{x}_i, l_i)\}$  with  $i = 1, \dots, N_{\mathcal{X}}$  consists of  $N_{\mathcal{X}}$  training examples. Each training example encapsulates a feature vector  $\mathbf{x}_i \in \mathbb{R}^d$  in a  $d$ -dimensional feature space and the respective class label  $l_i \in \{l^1, \dots, l^{N_C}\}$ , where  $N_C$  represents the number of classes. In contrast, the test set  $\mathcal{Y} = \{\mathbf{x}_j\}$  with  $j = 1, \dots, N_{\mathcal{Y}}$  only consists of  $N_{\mathcal{Y}}$  feature vectors  $\mathbf{x}_j \in \mathbb{R}^d$ . If available, the respective class labels  $l_j \in \{l^1, \dots, l^{N_C}\}$  may be used for evaluation (this is the case for the test set of the EnMAP contest dataset, see Section 3). For multi-class classification, we involve two classifiers represented by a Support Vector Machine (Cortes and Vapnik, 1995) and a Random Forest (Breiman, 2001).

### 5.1 Classification Based on a Support Vector Machine

Given the training set  $\mathcal{X}$ , the Support Vector Machine – just as comparable kernel-based models like the Import Vector Machine or Relevance Vector Machine – optimizes a linearly solvable classification problem depending only on the input features  $\mathbf{x}_i$ , a weight vector  $w$  and a bias  $b$ . Kernel-based methods introduce non-linear functions  $\phi(\mathbf{x}_i)$  to easily find a linear solution in a Reproducing Kernel Hilbert Space (RKHS), these higher dimensional feature spaces are induced implicitly by kernel functions  $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$ . Finally, all methods look for a subset  $\mathcal{V} \subset \mathcal{X}$  of training samples to sparsely induce these spaces.

The Support Vector Machine was designed to solve large margin classification problems as an implementation of statistical learning theory. It establishes a separating hyperplane and a maximal margin free of training data by choosing a subset  $\mathcal{SV} \subset \mathcal{X}$  called support vectors (SVs). The optimization problem is given by Equations 10 and 11. The SVs are used to calculate the normal vector  $w$  on the hyperplane and the bias  $b$  to fulfil the constraint

on the optimization problem.

$$\min \frac{\|w\|^2}{2} + C \sum_{i=1}^n \xi_i \quad (10)$$

$$\text{subject to: } l_i(w \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad (11)$$

It can be shown that minimizing Equation 10 is equal to maximizing the *margin*. The slack variables  $\xi_i$  allow for falsely assigned training data in favour of generalization.

### 5.2 Classification Based on a Random Forest

A Random Forest (Breiman, 2001) is an ensemble of randomly trained decision trees. In the training phase, a pre-defined number  $N_T$  of individual decision trees are trained on different subsets of the given training data, where the subsets are randomly drawn with replacement. Thus, the decision trees are all randomly different from one another which results in a de-correlation between individual tree predictions. In the classification phase, the feature vectors  $\mathbf{x}_i$  are classified by each tree, i.e. each tree casts a vote for one of the class labels  $l^k$  with  $k = 1, \dots, N_C$ . Thus, the posterior probability  $p(l_i = l^k | \mathbf{x}_i)$  of a class label  $l_i$  belonging to class  $l^k$  given the feature vector  $\mathbf{x}_i$  may be expressed as the ratio of the number  $N_k$  of votes cast for class  $l^k$  across all decision trees and the number  $N_T$  of involved decision trees:

$$p(l_i = l^k | \mathbf{x}_i) = \frac{N_k}{N_T} \quad (12)$$

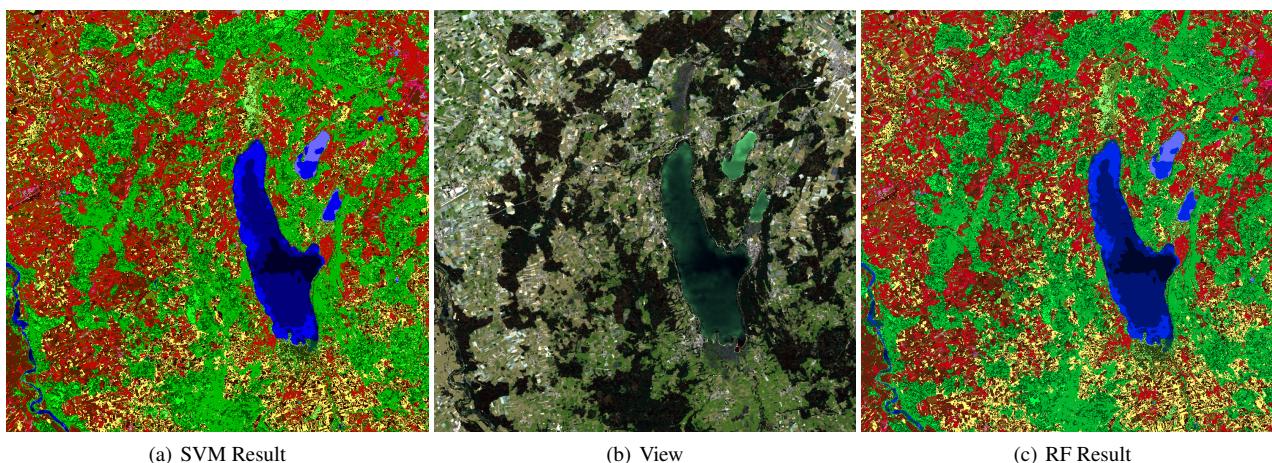
Instead of a probabilistic consideration, the assignment of a respective class label  $l_i$  to an observed feature vector  $\mathbf{x}_i$  is typically based on the majority vote across all decision trees which results in an improved generalization and robustness (Criminisi and Shotton, 2013).

## 6 RESULTS

The data described in Section 3 have been classified by a Support Vector Machine and a Random Forest. For the Support Vector Machine, a one-against-one approach and a Radial Basis Function (RBF) kernel have been used. The kernel parameter  $\gamma$  has been optimized in the range between  $\gamma = 10^{-15}, \dots, 10^5$  and the cost parameter  $C$  in the range between  $C = 10^{-5}, \dots, 10^{15}$  by cross validation with five fold grid search and exponent 5 increments in each parameter, for fully comprehensive instructions cf. Braun et al. (2010, 2012). The Random Forest has been trained using a maximum of 500 trees. For each classifier, the entire dataset  $\mathcal{X} = \{(\mathbf{x}_i, l_i)\}$  with all 244 channels has been used.

The classification results of the Support Vector Machine and the Random Forest are presented in Figure 4, a near-natural color visualisation is also given in order to facilitate visual comparison. Obviously, both classifiers produce visually rather similar results. It should be kept in mind though, that an area of 900 square kilometres is observed and that both classifiers are state-of-the-art methods not expected to fail on large numbers of pixels. Hence, such similar results had to be expected beforehand. The largest failure of both classifiers is that they over-estimate the appearance of class 9, which relates to urban and industrial areas and confuse them with agricultural areas.

When preparing this paper, the authors have evaluated several other classifiers, like Import Vector Machines, Relevance Vector Machines, AdaBoost, Neural Networks, Gaussian Mixture Models, but also more traditional techniques like Spectral Angle Mapper and Maximum Likelihood (see webpage for details). Some



(a) SVM Result

(b) View

(c) RF Result

Figure 4. Results of the classified EnMAP contest dataset with two state-of-the-art classifiers. Left: Support Vector Machine (SVM); Centre: view for visual evaluation; Right: Random Forest (RF). Larger image available at: [www.ipf.kit.edu/code.php](http://www.ipf.kit.edu/code.php)

Table 1. Confusion matrix values for Support Vector Machine and Random Forest. Average, minimum and maximum values of class-specific measures represented by quality (QLT), correctness (COR) and completeness (CMP).

Classifier	avrg.QLT	avrg.COR	avrg.CMP
Support Vector Machine	75.23%	84.54%	85.02%
Random Forest	75.11%	83.44%	84.10%
Classifier	min.QLT	min.COR	min.CMP
Support Vector Machine	38.35%	66.66%	46.66%
Random Forest	37.72%	52.72%	48.33%
Classifier	max.QLT	max.COR	max.CMP
Support Vector Machine	100.00%	100.00%	100.00%
Random Forest	100.00%	100.00%	100.00%

of those produced obvious visual differences. However, the main goal of this report is to concentrate on quantitative figures. Thus, two of the most high ranking techniques in terms of overall accuracy have been selected for this report.

On the basis of the results visible in Figure 4, confusion matrices have been computed for both results, they are found in Figure 5. With an overall accuracy of 84.6% for Support Vector Machine, and 83.2% for Random Forest, both classifiers performed particularly well on the EnMAP contest dataset. For comparison, values of class-specific quality figures are provided in Table 1. As can be seen, the Support Vector Machine outperforms the Random Forest approach for the average and minimum class-specific quality figures also. The slight exception for minimum completeness does not neglect the general trend. Of course, there are some individual classes for which Random Forest values were higher than Support Vector Machine values, however, since there were no interpretable trends, comparing classes individually is omitted here.

Hence, in total, it can be concluded that both classifiers are well suited for hyperspectral EnMAP data, with Support Vector Machine being slightly superior to Random Forest, a finding confirmed by Pal (2006), Waske et al. (2010) and Chi et al. (2008) for other hyperspectral data.

## 7 THE ENMAP CONTEST

This contribution has described the dataset for the EnMAP contest and some first results. Now, it encourages the scientific com-

munity to take part in the EnMAP contest to promote scientific cooperation on producing high performance algorithms even before the launch of EnMAP. Therefore, it is required to provide some details about how scientist can take part in the contest. Fully instructive information will be given in a Portable Document File at [www.ipf.kit.edu/code.php](http://www.ipf.kit.edu/code.php). Interested scientists will have to realize the following steps.

1. Go to [www.ipf.kit.edu/code.php](http://www.ipf.kit.edu/code.php) and read the instructions (PDF)
2. Download the training data  $\mathcal{X}$  and the test data  $\mathcal{Y}$  (provided as \*.mat and \*.txt (ASCII) data)
3. Download the entire image  $\mathcal{I}$  (provided as \*.mat and \*.txt (ASCII) data)
4. Develop a classification algorithm  $\mathcal{A} : f(\mathbf{x}_i) \rightarrow l_i$
5. Train the algorithm  $\mathcal{A}$  on  $\mathcal{X} = \{(\mathbf{x}_i, l_i)\}$
6. Apply the algorithm  $\mathcal{A}$  to  $\mathcal{Y}$
7. Calculate the quality figures overall accuracy, completeness<sub>i</sub>, correctness<sub>i</sub> and quality<sub>i</sub> based on  $\mathcal{Y}$
8. Apply the algorithm  $\mathcal{A}$  to  $\mathcal{I}$
9. Report the quality figures to the corresponding author and provide the classified image  $\mathcal{I}$  (for control)

The authors of this paper will evaluate the quality figures reported for  $\mathcal{Y}$  by checking them on the provided results for  $\mathcal{I}^2$ . The first EnMAP contest will go until 31<sup>st</sup> of December 2015. Then, the results will be submitted to the ISPRS Journal of Photogrammetry and Remote Sensing in a condensed manner. Scientist whose results are among the ten highest overall accuracy values confirmed will be invited as co-authors in the ranking of their results.

<sup>2</sup>The authors possess an image with the  $\mathcal{X}$  and  $\mathcal{Y}$  pixels' positions in the image. These positions are not available to contestants and cannot be made available. Thus, falsification of results is avoided.

## 8 DISCUSSION AND CONCLUSION

This contribution has elaborated a benchmark dataset for hyperspectral simulated EnMAP data. The benchmark is made available at [www.ipf.kit.edu/code.php](http://www.ipf.kit.edu/code.php), where results will also be available at higher resolution for a more detailed comparison. Thus, the paper aims to promote research on classifying EnMAP data before its launch. Given a set of comprehensive studies comparing classification approaches, the authors believe that the EnMAP mission will be an even greater success, since confusion of future users about which algorithm to use is reduced. Similar benchmarks have been provided for other hyperspectral datasets, for instance, the ROSIS Pavia dataset, the AVIRIS Indian Pines and Salinas datasets, or the HYDICE Washington D.C. dataset. These datasets have provided some objective insights into the performance expected from individual algorithms.

Although its main goal is to introduce the dataset, this paper has also presented two results of state-of-the-art classifiers for the simulated EnMAP data. It has shown that classifiers know to perform well on other hyperspectral data, i.e. kernel-based and ensemble techniques are also applicable to EnMAP. A finding which is not necessarily expected in the first place as argued above. The Support Vector Machine performed slightly better than the Random Forest in both global and (generally) class-specific figures. More classifiers have been evaluated by the authors and the respective results will be published in future work.

## 9 ACKNOWLEDGEMENT

The authors acknowledge the simulation of the EnMAP dataset and, perhaps more importantly, the kind support provided by Guanter et al. (2009) and especially Dr. Karl Segl of the GFZ German Research Centre for Geosciences who have helped in understanding the properties of and using the dataset.

## References

Bracken, A. H., 2014. Detecting soil erosion in semi-arid Mediterranean environments using simulated EnMAP data. Thesis (M.Sc.), Department of Geography, University of Lethbridge, Lethbridge, Canada.

Braun, A. C., Rojas, C., Echeverria, C., Rottensteiner, F., Bähr, H.-P., Niemeyer, J., Aguayo Arias, M., Kosov, S., Hinz, S. and Weidner, U., 2014. Design of a spectral-spatial pattern recognition framework for risk assessments using Landsat data – A case study in Chile. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7(3), pp. 917–928.

Braun, A. C., Weidner, U. and Hinz, S., 2010. Support vector machines for vegetation classification – A revision. *Photogrammetrie – Fernerkundung – Geoinformation* 2010(4), pp. 273–281.

Braun, A. C., Weidner, U. and Hinz, S., 2012. Classification in high-dimensional feature spaces – Assessment using SVM, IVM and RVM with focus on simulated EnMAP data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5(2), pp. 436–443.

Breiman, L., 2001. Random forests. *Machine Learning* 45(1), pp. 5–32.

Chi, M., Feng, R. and Bruzzone, L., 2008. Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. *Advances in Space Research* 41(11), pp. 1793–1799.

Cortes, C. and Vapnik, V., 1995. Support-vector networks. *Machine Learning* 20(3), pp. 273–297.

Criminisi, A. and Shotton, J., 2013. *Decision forests for computer vision and medical image analysis*. Advances in Computer Vision and Pattern Recognition, Springer, London, UK.

Dörnhöfer, K. and Oppelt, N., 2015. Anwendung eines biooptischen Modells zur Erfassung von Benthos und Wassertiefen in Küstengewässern – ein Test mit simulierten EnMAP Daten. *DGPF Tagungsband* 24, pp. 384–391.

Faßnacht, F., Weinacker, H. and Koch, B., 2011. Automatic forest area extraction from imaging spectroscopy data using an extended NDVI. *Proceedings of the 7th EARSeL Workshop on Imaging Spectroscopy*.

Guanter, L., Segl, K. and Kaufmann, H., 2009. Simulation of optical remote-sensing scenes with application to the EnMAP hyperspectral mission. *IEEE Transactions on Geoscience and Remote Sensing* 47(7), pp. 2340–2351.

Heipke, C., Mayer, H., Wiedemann, C. and Jamet, O., 1997. Evaluation of automatic road extraction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIII/3-4W2, pp. 151–160.

Held, M., Jakimow, B., Rabe, A., van der Linden, S., Wirth, F. and Hostert, P., 2012. EnMAP-Box Manual: Version 1.4. Technical Report, Humboldt-Universität zu Berlin, Berlin, Germany.

Kaufmann, H., Segl, K., Chabrilat, S., Hofer, S., Stuffer, T., Mueller, A., Richter, R., Schreier, G., Haydn, R. and Bach, H., 2006. EnMAP – A hyperspectral sensor for environmental mapping and analysis. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1617–1619.

Kaufmann, H., Segl, K., Guanter, L., Hofer, S., Foerster, K.-P., Stuffer, T., Mueller, A., Richter, R., Bach, H., Hostert, P. and Chlebek, C., 2008. Environmental Mapping and Analysis Program (EnMAP) – Recent advances and status. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 4, pp. 109–112.

Pal, M., 2006. Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. *International Journal of Remote Sensing* 27(14), pp. 2877–2894.

Schwieder, M., Leitão, P. J., Suess, S., Senf, C. and Hostert, P., 2014. Estimating fractional shrub cover using simulated EnMAP data: a comparison of three machine learning regression techniques. *Remote Sensing* 6(4), pp. 3427–3445.

Segl, K., Guanter, L., Kaufmann, H., Schubert, J., Kaiser, S., Sang, B. and Hofer, S., 2010. Simulation of spatial sensor characteristics in the context of the EnMAP hyperspectral mission. *IEEE Transactions on Geoscience and Remote Sensing* 48(7), pp. 3046–3054.

Segl, K., Guanter, L., Rogass, C., Kuester, T., Roessner, S., Kaufmann, H., Sang, B., Mogulsky, V. and Hofer, S., 2012. EeteS – The EnMAP end-to-end simulation tool. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5(2), pp. 522–530.

Stuffer, T., Förster, K., Hofer, S., Leipold, M., Sang, B., Kaufmann, H., Penné, B., Mueller, A. and Chlebek, C., 2009. Hyperspectral imaging – An advanced instrument concept for the EnMAP mission (Environmental Mapping and Analysis Programme). *Acta Astronautica* 65(7-8), pp. 1107–1112.

Stuffer, T., Kaufmann, C., Hofer, S., Förster, K. P., Schreier, G., Mueller, A., Eckardt, A., Bache, H., Penné, B., Benz, U. and Haydn, R., 2007. The EnMAP hyperspectral imager – An advanced optical payload for future applications in Earth observation programmes. *Acta Astronautica* 61(1-6), pp. 115–120.

Waske, B., van der Linden, S., Benediktsson, J. A., Rabe, A. and Hostert, P., 2010. Sensitivity of support vector machines to random feature selection in classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* 48(7), pp. 2880–2889.

Weidner, U., 2008. Contribution to the assessment of segmentation quality for remote sensing applications. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVII-B7, pp. 479–484.

SVM	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	QLT	COR	CMP
C01	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100,00	100,00	100,00
C02	0	55	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	91,67	100,00	91,67
C03	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92,31	92,31	100,00
C04	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100,00	100,00	100,00
C05	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100,00	100,00	100,00
C06	0	0	0	0	0	33	5	18	0	0	3	0	0	0	0	1	0	0	0	0	43,42	67,35	55,00
C07	0	0	0	0	0	5	55	0	0	0	0	0	0	0	0	0	0	0	0	0	84,62	91,67	91,67
C08	0	0	0	0	0	10	0	50	0	0	0	0	0	0	0	0	0	0	0	0	64,10	73,53	83,33
C09	0	0	0	0	0	1	0	0	56	0	2	0	1	0	0	0	0	0	0	0	80,00	84,85	93,33
C10	0	0	0	0	0	0	0	0	0	54	1	4	0	0	0	0	0	0	0	1	70,13	76,06	90,00
C11	0	0	0	0	0	0	0	0	2	2	56	0	0	0	0	0	0	0	0	0	82,35	87,50	93,33
C12	0	0	0	0	0	0	0	0	0	13	2	40	0	0	1	0	0	0	2	2	62,50	90,91	66,67
C13	0	0	0	0	0	0	0	0	7	0	0	0	28	18	0	0	1	6	0	0	38,36	68,29	46,67
C14	0	0	0	0	0	0	0	0	1	0	0	0	9	49	0	0	0	1	0	0	61,25	71,01	81,67
C15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	19	0	0	1	0	50,00	66,67	66,67
C16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	42	0	0	1	0	51,85	66,67	70,00
C17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	98,36	98,36	100,00
C18	0	0	0	0	0	0	0	0	0	0	0	0	3	2	0	0	0	43	0	0	76,79	84,31	89,58
C19	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	1	37	0	80,43	90,24	88,10
C20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	13	76,47	81,25	92,86

(a) SVM Result

RF	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	QLT	COR	CMP
C01	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100,00	100,00	100,00
C02	0	56	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93,33	100,00	93,33
C03	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93,75	93,75	100,00
C04	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100,00	100,00	100,00
C05	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100,00	100,00	100,00
C06	0	0	0	0	0	29	6	21	0	0	2	0	0	0	0	1	0	0	0	0	33,72	52,73	48,33
C07	0	0	0	0	0	7	53	0	0	0	0	0	0	0	0	0	0	0	0	0	80,30	89,83	88,33
C08	0	0	0	0	0	19	0	41	0	0	0	0	0	0	0	0	0	0	0	0	50,62	66,13	68,33
C09	0	0	0	0	0	0	0	0	54	0	4	0	1	0	0	0	0	0	0	0	81,82	90,00	90,00
C10	0	0	0	0	0	0	0	0	0	52	2	4	0	0	0	0	0	0	0	1	66,67	74,29	86,67
C11	0	0	0	0	0	0	0	0	0	2	55	3	0	0	0	0	0	0	0	0	78,57	84,62	91,67
C12	0	0	0	0	0	0	0	0	0	14	2	39	0	0	1	0	0	0	2	2	58,21	84,78	65,00
C13	0	0	0	0	0	0	0	0	4	0	0	0	31	18	0	0	1	6	0	0	44,29	75,61	51,67
C14	0	0	0	0	0	0	0	0	1	0	0	0	9	49	0	0	0	1	0	0	58,54	68,57	80,00
C15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	19	0	0	1	0	52,38	64,71	73,33
C16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	42	0	0	1	0	49,35	69,09	63,33
C17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	96,77	96,77	100,00
C18	0	0	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0	43	0	0	77,19	83,02	91,67
C19	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	1	37	0	84,44	92,68	90,48
C20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	13	82,35	82,35	100,00

(b) RF Result

Figure 5. Confusion matrices of the Support Vector Machine (above) and Random Forest (below) result. Rows: known classes, Columns: predicted classes. Overall accuracy and class-specific quality measures included. Larger image available at: [www.ipf.kit.edu/code.php](http://www.ipf.kit.edu/code.php)