Investigation of the Potential of Hyperspectral EnMAP Data for Land Cover and Land Use Classification

S. KELLER[1], A. C. BRAUN[2], S. HINZ[1] & M. WEINMANN[1]

Abstract: In this paper, we focus on the classification of hyperspectral data that is expected to be acquired with the Environmental Mapping and Analysis Program (EnMAP) mission, a hyperspectral satellite mission supposed to be launched into space in the near future. More specifically, we consider the dataset presented with the EnMAP Contest, a benchmark that has recently been initiated with the objective of classifying different land cover and land use classes based on EnMAP-like hyperspectral data. We address the EnMAP Contest by presenting a framework that involves (1) standard approaches for dimensionality reduction and feature selection and (2) classifiers relying on different learning principles. The classification results derived for different configurations of our framework clearly reveal the potential of respective techniques and provide the basis for further improvements in different research directions.

1 Introduction

The automated analysis of hyperspectral imagery has become a topic of major interest in remote sensing as such data is valuable for the classification of land cover and land use (PLAZA et al. 2009; CAMPS-VALLS et al. 2014). However, most of the currently available benchmark datasets for evaluating the performance of respective classification approaches consist of (partially) labeled airborne hyperspectral imagery acquired during low-altitude flights. This allows an analysis of areas with a rather limited extent while a large-scale or even nationwide coverage would often be desirable. Recent attempts towards a large-scale coverage with hyperspectral data have resulted in a conceptual transfer towards the use of satellite remote sensing which is addressed by the *Environmental Mapping and Analysis Program* (*EnMAP*). EnMAP represents a hyperspectral satellite mission supposed to be launched into space in the near future (KAUFMANN et al. 2008). In order to already obtain first impressions about expected future hyperspectral EnMAP data products before the mission is launched, simulated EnMAP-like hyperspectral data with 244 spectral bands covering the range from 420 nm to 2450 nm has been released and can be used for numerous investigations (GUANTER et al. 2009; SEGL et al. 2010).

In this paper, we focus on the classification of simulated EnMAP-like hyperspectral data (Fig. 1). We present a framework which involves (1) standard approaches for dimensionality reduction and feature selection and (2) several classifiers relying on different learning principles (instance-based learning, probabilistic learning and ensemble learning). To evaluate the performance of this framework, we consider nine different configurations of the framework and present the respective classification results for a recently published benchmark dataset with 20 different land

[1] Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, 76131 Karlsruhe, Germany, E-Mail: [sina.keller, stefan.hinz, martin.weinmann]@kit.edu
[2] Karlsruhe Institute of Technology (KIT), Institute of Regional Science, Reinhard-Baumeister-Platz 1, 76131 Karlsruhe, Germany, E-Mail: andreas.ch.braun@kit.edu

cover and land use classes (five water classes, five forest/meadow classes, five agriculture classes, three urban/industrial classes, and two pasture/fallow classes) in tables and figures. These results correspond to an overall accuracy in the range of about 77-84 % and a kappa value in the range of about 76-83 %, whereby the best classification results are obtained with the classifier relying on ensemble learning. The derived results clearly reveal the potential of EnMAP-like hyperspectral data, and they provide a basis for further investigations in different research directions.

This paper represents an extension of our previous work (KELLER et al. 2016), where we investigated the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data. In addition to the quantitative classification results obtained for the subset of labeled pixels, we now also focus on the qualitative results obtained for the whole scene. We provide an in-depth analysis of the results obtained for nine different configurations of our framework by considering global evaluation metrics (overall accuracy, Cohen's kappa coefficient, average completeness, average correctness and average quality) as well as class-wise evaluation metrics (completeness, correctness and quality) and the whole classified scene corresponding to the given simulated EnMAP-like hyperspectral data.

After briefly summarizing related work in Section 2, we explain the single components of our proposed framework for the classification of hyperspectral EnMAP-like data in Section 3. Subsequently, in Section 4, we describe the benchmark dataset used in the scope of our experiments in detail, and we present the respective results derived with our framework and discuss these results with respect to different criteria. Finally, we provide concluding remarks and suggestions for future work in Section 5.

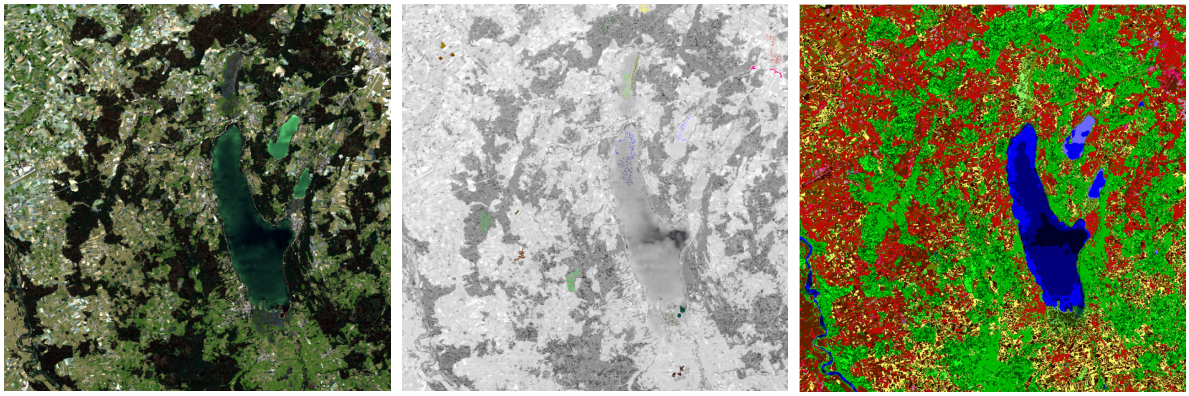

Fig. 1: The *EnMAP Contest* (BRAUN et al. 2015): simulated hyperspectral data covering 244 spectral bands is available for each pixel of the scene (left; image courtesy of Dr. K. Segl). Some pixels within the simulated EnMAP dataset have been labeled (center). The objective is to assign respective class labels to all pixels (right).

2 Related Work

The simplest approach for classifying hyperspectral imagery certainly consists in a per-pixel consideration of the reflectance values of all spectral bands and a subsequent pixel-wise classification based on well-known standard classification approaches which are meanwhile

available in a variety of software packages, e.g. kernel-based methods (CAMPS-VALLS & BRUZZONE 2005), Support Vector Machines (MELGANI & BRUZZONE 2004; CHI et al. 2008), or Random Forests (HAM et al. 2005; JOELSSON et al. 2005). In particular, Support Vector Machines have been widely used for classifying hyperspectral data and related but more sophisticated classification approaches represented by Import Vector Machines and Relevance Vector Machines have also been used for this task (BRAUN et al. 2011; BRAUN et al. 2012).

Despite the classifier, the data representation plays an important role. When classifying hyperspectral data, it should generally be taken into account that the reflectance values of some spectral bands are strongly correlated and that typically not all spectral bands are relevant to the classification task. In particular for high-dimensional data representations as given for hyperspectral data, the Hughes phenomenon (HUGHES 1968) can often be observed. According to this phenomenon, an increase of the number of features over a certain threshold results in a decrease in classification accuracy, given a constant number of training examples (MELGANI & BRUZZONE 2004). Hence, it seems desirable to transfer the given high-dimensional data into a new, more compact representation which encodes almost the same information with respect to the classification task. On the one hand, this can be done via dimensionality reduction (VAN DER MAATEN et al. 2009) which focuses on the mapping of feature vectors from the original feature space onto a specific subspace spanned by meta-features. For this purpose, variants of the well-known Principal Component Analysis or the Independent Component Analysis are commonly applied (LICCIARDI et al. 2012; WANG & CHANG 2006; VILLA et al. 2011). On the other hand, it is possible to apply feature selection (GUYON & ELISSEEFF 2003; SAEYS et al. 2007) which aims at only retaining the relevant features for classification and thus only considers reflectance values of specific spectral bands. Respective methods for instance allow assessing the importance of single spectral bands for land cover classification (LE BRIS et al. 2014) as well as different band selection and band fusion processes (CHEHATA et al. 2014). These studies demonstrate that approaches for both dimensionality reduction as well as feature selection are appropriate to cope with the high degree of redundancy contained in hyperspectral datasets and that the consideration of only a relatively small subspace of the data is still sufficient to derive reasonable classification results.

More recent trends in the classification of hyperspectral imagery mainly address the consideration of spatial context, i.e. the spatial structure within the local image neighborhood of a pixel is considered in addition to the reflectance values of the spectral bands characterizing that pixel. This is commonly referred to as spectral-spatial classification. Respective approaches are for instance given by a probabilistic pixel-wise classification that is followed by either a Markov Random Field regularization (TARABALKA et al. 2010) or a hierarchical optimization (TARABALKA & TILTON 2011). Furthermore, it has been proposed to involve segmentation approaches to support classification, e.g. by using the results of a pixel-wise classification in combination with watershed segments (TARABALKA et al. 2008) which allows a majority voting within watershed segments to account for spatial information. Instead of relying on a segment-based consideration of the results of a pixel-wise classification, spatial context can also be assessed by sampling spectral information within the local surrounding of a pixel, e.g. within adaptive pixel neighborhoods (FAUVEL et al. 2008) or within superpixels (FANG et al. 2015).

3 Methodology

In the scope of this paper, we present a framework for classifying hyperspectral data (Fig. 2). This framework includes three different options for the data representation per pixel (Section 3.1) as well as three different options for supervised classification (Section 3.2).

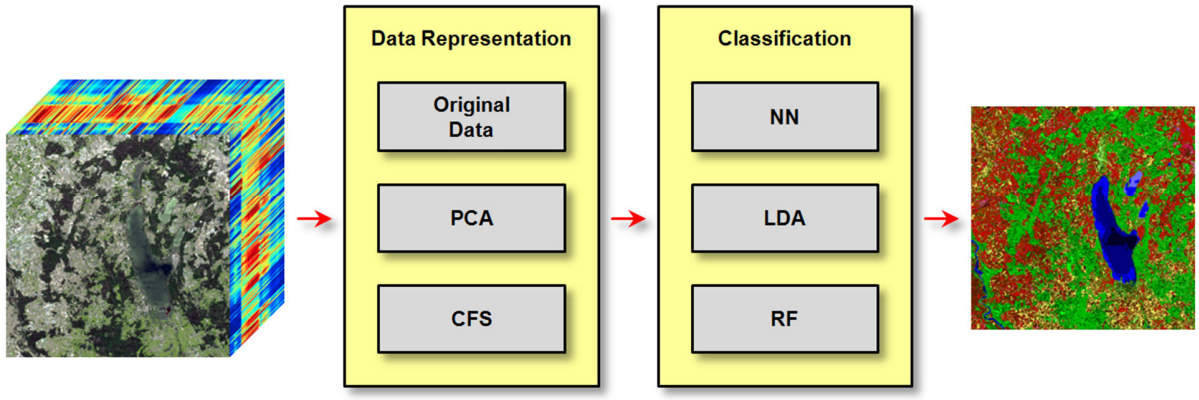

Fig. 2: The proposed framework allowing for three different options for the data representation and three different options for supervised classification.

3.1 Data Representation

In our framework, we focus on different options for deriving feature vectors which serve as input to the subsequent classification procedure. We assume that each pixel is described with a high number of spectral bands and that the spectral information of each spectral band is represented by a reflectance value between 0 % and 100 %. In the following, we focus on three options to derive the feature vectors that serve as input to the subsequent classification procedure.

3.1.1 Original Data Representation

The straightforward approach consists in using the given representation of the data as input to the classification procedure, where we simply concatenate the reflectance values for all spectral bands to obtain the respective feature vectors.

3.1.2 Dimensionality Reduction: Principal Component Analysis

For dimensionality reduction, we focus on using a standard *Principal Component Analysis* (*PCA*) in the scope of this paper and assume that this is appropriate, even though the PCA is not necessarily the best approach for dimensionality reduction (CHERIYADAT & BRUCE 2003). The main idea of the PCA consists in transforming a set of feature vectors from a high-dimensional feature space (which is spanned by possibly correlated features) to a new feature space which is spanned by linearly uncorrelated meta-features, the principal components, which account for the data variability along the respective dimension. Ranking the meta-features with respect to the respectively covered variability allows us to consider the few meta-features that are most

relevant with respect to data variability, while other meta-features are almost irrelevant with respect to data variability. In the scope of our work, we use the most relevant meta-features that cover 99.9 % of the variability of the given training data to define the new feature vectors, and we assume that we do not have a significant loss of relevant information with respect to the classification task when discarding all other, less relevant meta-features.

3.1.3 Feature Selection: Correlation-based Feature Selection

For feature selection, we focus on using *Correlation-based Feature Selection* (*CFS*) (HALL 1999) which takes into account (1) the correlation between features and classes to identify relevant features and (2) the correlation among features to identify and discard redundant features. More specifically, CFS exploits the average correlation $\bar{\rho}_{FC}$ between features and classes as well as the average correlation $\bar{\rho}_{FF}$ among classes to evaluate the relevance R of a feature subset comprising n features according to

$$R = \frac{n\bar{\rho}_{FC}}{\sqrt{n + n(n-1)\bar{\rho}_{FF}}}$$

where the correlation metric is given by the symmetrical uncertainty (HALL 1999). Deriving a suitable feature subset thus corresponds to maximizing the relevance R over the feature subset space. Beginning with an initial feature subset either a feature is added to the feature subset or a feature is removed from the feature subset until the relevance R converges to a stable maximum. The new feature vectors thus correspond to a selection of a few dimensions of the original data representation.

3.2 Supervised Classification

For classification, we integrate three classifiers relying on different learning principles into our framework. This allows us to draw more general conclusions about the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data.

3.2.1 Nearest Neighbor Classifier

The *Nearest Neighbor* (*NN*) *classifier* relies on the principle of instance-based learning. For each feature vector belonging to a sample in the test set, this classifier performs a comparison to the feature vectors belonging to the samples in the training data and selects the class label associated with the most similar feature vector in the training data. Accordingly, the induction process is delayed to the prediction stage, i.e. there is no training stage and the prediction stage might correspond to a higher computational effort due to the required comparisons of vectors (here with respect to the Euclidean distance).

3.2.2 Linear Discriminant Analysis Classifier

The *Linear Discriminant Analysis* (*LDA*) *classifier* relies on the principle of probabilistic learning. In the training stage, it is assumed that the instances of different classes exhibit a Gaussian distribution in the feature space. Accordingly, the training of the LDA classifier consists in fitting a multivariate Gaussian distribution to the given training data, whereby the parameters of a Gaussian distribution have to be estimated for each class. Due to a lack of

knowledge about the behavior of single classes in the feature space, the same covariance matrix is assumed for each class so that only the means vary for the different classes. In the prediction stage, the class probabilities are evaluated for each feature vector belonging to a sample in the test set and the label of the class with the maximum probability is selected.

3.2.3 Random Forest Classifier

The *Random Forest* (*RF*) *classifier* (BREIMAN 2001) relies on the principle of ensemble learning. In the training stage, an ensemble of randomly trained decision trees is generated via bagging (BREIMAN 1996). Thereby, subsets of the training data are randomly drawn with replacement and an individual decision tree is trained for each subset. In the prediction stage, for each feature vector belonging to a sample in the test set, each decision tree casts a vote for one of the defined classes and the majority vote is selected to obtain the most probable class label.

4 Experimental Results

In the following, we first introduce the used benchmark dataset (Section 4.1). Subsequently, we present the derived results (Section 4.2) and we also provide a short discussion with respect to different aspects (Section 4.3).

4.1 Dataset

To evaluate the performance of our framework, we use the dataset presented in the scope of the recent EnMAP contest (BRAUN et al. 2015). This dataset is based on the simulated *EnMAP Alpine Foreland dataset* (GUANTER et al. 2009; SEGL et al. 2010), but it additionally contains a subset of pixels that are labeled and thus allow a benchmarking of different approaches for the classification of hyperspectral data.

Like the simulated EnMAP Alpine Foreland dataset, the considered dataset covers an area of about 30 km × 30 km around the Ammersee in Bavaria, Germany. It is represented by an image of 1000 × 1000 pixels (i.e. the ground sampling distance is 30 m), where each pixel is associated with values on 244 simulated spectral bands covering the spectral range of 420-2450 nm with a varying spectral sampling of 6.5-10 nm. A near natural color visualization of the considered image is provided in Fig. 1 (left) and reveals a diversity of water, vegetation, agricultural, and urban/industrial classes.

The additional labeling has been performed for a subset of pixels as shown in Fig. 1 (center) and with respect to 20 different land cover and land use classes, whereby the manual annotation has been done by considering visual differences in the image (considering several band combinations) as well as the individual spectra of selected pixels. More specifically, the labeling takes into account five water classes, five forest/meadow classes, five agriculture classes, three urban/industrial classes, and two pasture/fallow classes. The average spectra of these classes are shown in Fig. 3, and it can be observed that some of the classes reveal a spectrally very similar behavior, which makes classification more challenging. With the annotation, a split of the labeled pixels into training data and test data is already provided, whereby the training data comprises hyperspectral data for 2,617 labeled pixels and the test data contains hyperspectral data for 1,124 labeled pixels.

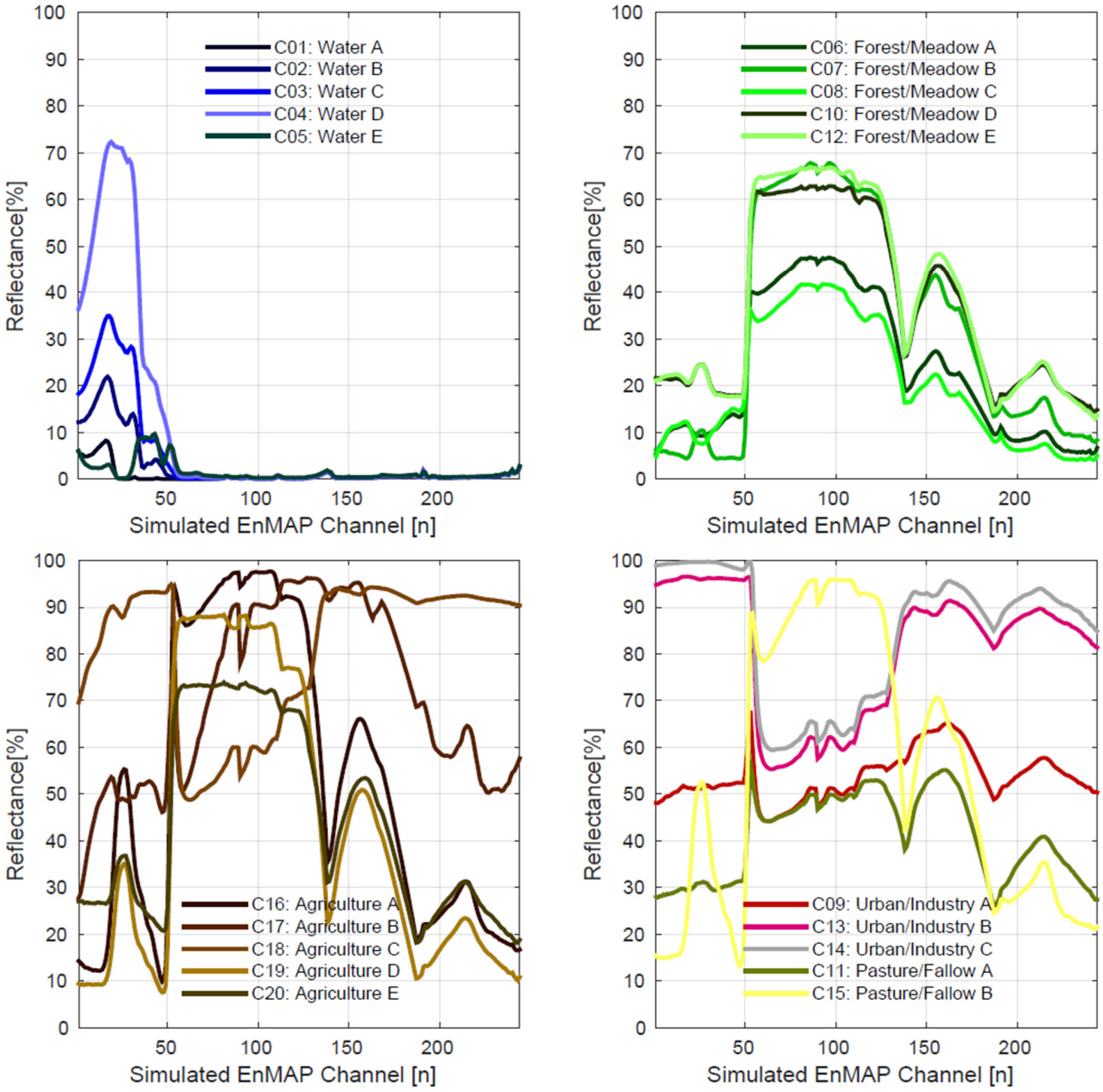

Fig. 3: Average spectra for the 20 land cover and land use classes considered in the scope of the EnMAP contest (BRAUN et al. 2015): water classes (top left), forest/meadow classes (top right), agriculture classes (bottom left) and other classes (bottom right)

4.2 Results

To judge about the performance of different configurations of our framework for classifying hyperspectral EnMAP data, we first consider the general evaluation metrics represented by overall accuracy OA, Cohen's kappa coefficient κ, average completeness $\overline{\mathrm{CMP}}$, average correctness $\overline{\mathrm{COR}}$ and average quality $\overline{\mathrm{Q}}$.

In total, we consider nine different configurations of our framework, resulting from all conceivable combinations that are possible when considering three options for classification (1:

NN classifier; 2: LDA classifier; 3: RF classifier) and three options for the data provided to the respective classifier (1: use of the values of all 244 spectral bands per pixel; 2: use of meta-features derived via PCA-based dimensionality reduction; 3: use of the values of those spectral bands that are selected via CFS). The derived classification results are visualized in Fig. 4.

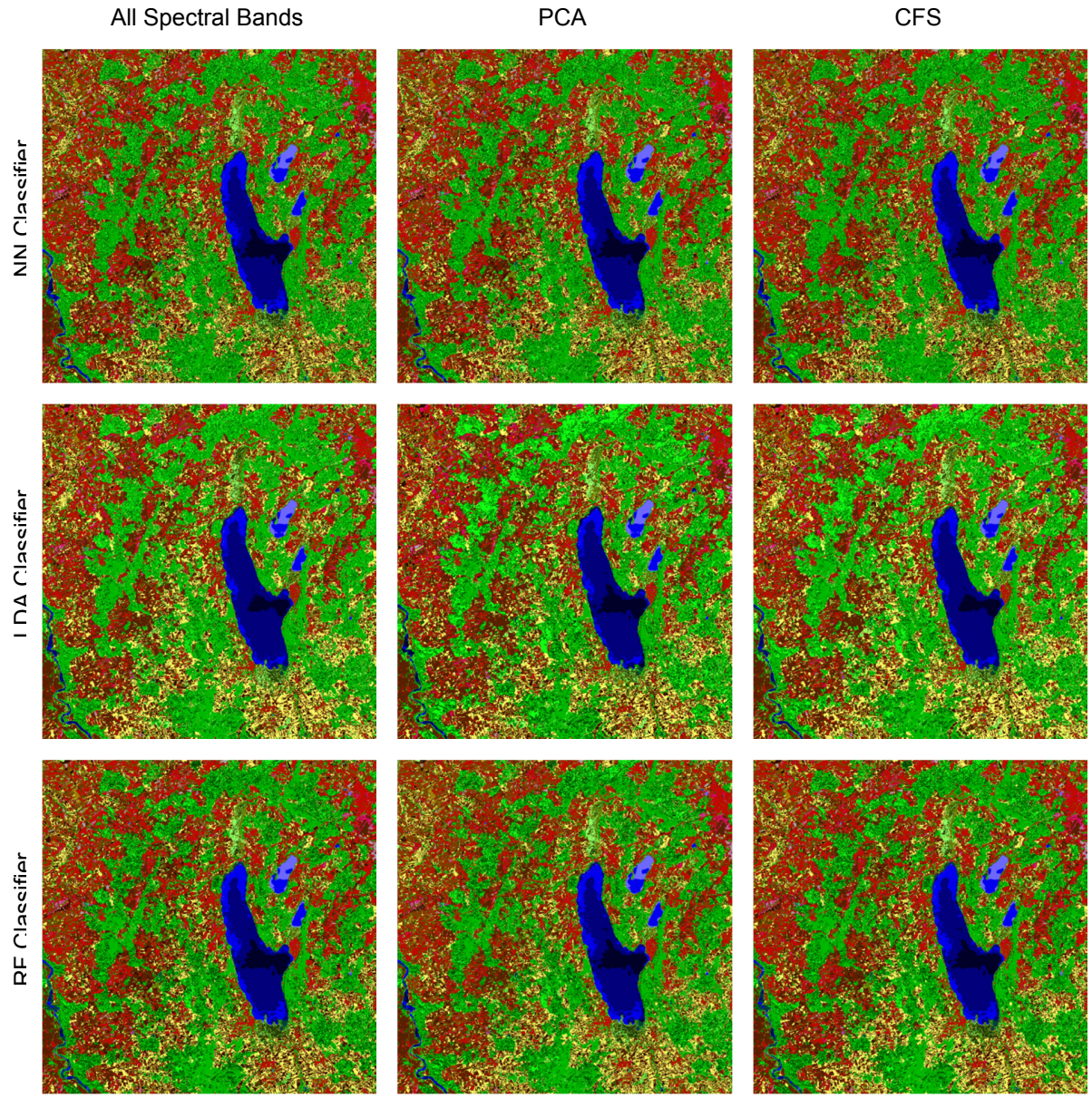

Fig. 4: Classification results obtained for the NN classifier (top row), the LDA classifier (center row) and the RF classifier (bottom row) when using all spectral bands (left column), PCA-based dimensionality reduction (center column) and feature selection via CFS (right column): the color encoding of the different classes is explained in Fig. 3.

The corresponding values for the general evaluation metrics are provided in Tab. 1 and reveal an overall accuracy in the range of about 77-84 % for the nine different configurations of our framework. The kappa value is in the range of about 76-83 % for the considered configurations. The table also contains the average values of completeness, correctness and quality across all 20 classes.

Tab. 1: Classification results obtained for nine different configurations of our framework

Classifier	DR/FS	OA [%]	κ [%]	$\overline{\mathrm{CMP}}$ [%]	$\overline{\mathrm{COR}}$ [%]	$\overline{\mathrm{Q}}$ [%]
NN	-	80.07	78.98	80.68	80.04	69.94
NN	PCA	80.96	79.92	81.52	80.99	71.08
NN	CFS	79.89	78.80	80.29	79.95	69.81
LDA	-	78.65	77.48	79.21	78.37	68.18
LDA	PCA	77.40	76.17	77.47	77.11	65.98
LDA	CFS	79.45	78.34	80.32	78.33	68.18
RF	-	83.19	82.27	83.69	83.19	73.70
RF	PCA	82.74	81.80	83.31	82.72	73.30
RF	CFS	82.65	81.71	83.25	82.55	72.94

Whereas the general evaluation metrics allow statements about the complete classification results, the class-wise evaluation metrics of completeness, correctness and quality can be used to assess statistics about the behavior of single classes. With a focus on distinguishing between 20 different land cover and land use classes, we provide the minimum and maximum values of the three class-wise evaluation metrics for all nine configurations of our framework in Tab. 2. It can be observed that the maximum is always reached at 100 % which indicates that some classes are not too hard to distinguish, even when using different configurations of our framework.

Tab. 2: Classification results obtained for nine different configurations of our framework

Classifier	DR/FS	$\mathrm{CMP}_{\mathrm{min}}$ [%]	$\mathrm{CMP}_{\mathrm{max}}$ [%]	$\mathrm{COR}_{\mathrm{min}}$ [%]	$\mathrm{COR}_{\mathrm{max}}$ [%]	$\mathrm{Q}_{\mathrm{min}}$ [%]	$\mathrm{Q}_{\mathrm{max}}$ [%]
NN	-	40.00	100.00	48.00	100.00	27.91	100.00
NN	PCA	45.00	100.00	56.25	100.00	33.33	100.00
NN	CFS	45.00	100.00	45.76	100.00	29.35	100.00
LDA	-	30.00	100.00	42.86	100.00	21.43	100.00
LDA	PCA	16.67	100.00	43.48	100.00	13.70	100.00
LDA	CFS	20.00	100.00	50.00	100.00	16.67	100.00
RF	-	45.00	100.00	58.18	100.00	38.03	100.00
RF	PCA	48.33	100.00	60.00	100.00	36.71	100.00
RF	CFS	48.33	100.00	54.72	100.00	34.52	100.00

4.3 Discussion

As expected, the RF classifier provides the highest quality of the derived classification results. This might be due to the fact that it provides a better generalization capability in case of strongly correlated features. When transforming the original feature vectors via PCA and keeping only the most relevant meta-features that cover 99.9 % of the variability of the given training data, it can be observed that only eight meta-features and therefore a significantly more compact data representation are used, while the quality of the derived classification results is similar with a slight decrease in most cases, but also a slight increase for one case. When using CFS for feature selection, the reflectance values of particular spectral bands are selected. A closer look at the selected bands reveals that reflectance values of 33 spectral bands are considered for classification instead of reflectance values of all 244 spectral bands. Thereby, referring to Fig. 3, 15 spectral bands are considered within the interval [1,50], ten within the interval [51,100], four within the interval [101,150], one within the interval [151,200] and tree within the interval [201,244]. Again, the quality of the derived classification results is similar with a slight decrease in most cases and a slight increase for one case.

5 Conclusions

In this paper, we have investigated the potential of hyperspectral EnMAP data for land cover and land use classification. We have presented a framework that involves (1) standard approaches for dimensionality reduction and feature selection and (2) classifiers relying on different learning principles. For nine different configurations of our framework, we have presented the results of a performance evaluation on the benchmark dataset presented in the scope of the recent EnMAP contest, a simulated hyperspectral dataset with 20 different land cover and land use classes. The derived results clearly reveal that a variety of land cover and land use classes can be distinguished based on hyperspectral EnMAP data. Furthermore, we have demonstrated that the transfer of the original, high-dimensional data representation corresponding to 244 spectral bands to a significantly more compact data representation allows to derive classification results of similar and partially even better quality.

6 References

BRAUN, A. C., WEIDNER, U. & HINZ, S., 2011: Support vector machines, import vector machines and relevance vector machines for hyperspectral classification. Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-4.

BRAUN, A. C., WEIDNER, U. & 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), 436-443.

BRAUN, A. C., WEINMANN, M., KELLER, S., MÜLLER, R., REINARTZ, P. & HINZ, S., 2015: The EnMAP contest: developing and comparing classification approaches for the Environmental Mapping and Analysis Programme – Dataset and first results. The

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **XL** (3/W3), 169-175.

BREIMAN, L., 1996: Bagging predictors. Machine Learning **24**(2), 123-140.

BREIMAN, L., 2001: Random forests. Machine Learning **45**(1), 5-32.

CAMPS-VALLS, G. & BRUZZONE, L., 2005: Kernel-based methods for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing **43**(6), 1351-1362.

CAMPS-VALLS, G., TUIA, D., BRUZZONE, L. & BENEDIKTSSON, J. A., 2014: Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. IEEE Signal Processing Magazine **31**(1), 45-54.

CHEHATA, N., LE BRIS, A. & NAJJAR, S., 2014: Contribution of band selection and fusion for hyperspectral classification. Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-4.

CHERIYADAT, A. & BRUCE, L. M., 2003: Why principal component analysis is not an appropriate feature extraction method for hyperspectral data. Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS) **6**, 3420-3422.

CHI, M., FENG, R. & BRUZZONE, L., 2008: Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. Advances in Space Research **41**(11), 1793-1799.

FANG, L., LI, S., KANG, X. & BENEDIKTSSON, J. A., 2015: Spectral-spatial classification of hyperspectral images with a superpixel-based discriminative sparse model. IEEE Transactions on Geoscience and Remote Sensing **53**(8), 4186-4201.

FAUVEL, M., CHANUSSOT, J. & BENEDIKTSSON, J. A., 2008: Adaptive pixel neighborhood definition for the classification of hyperspectral images with support vector machines and composite kernel. Proceedings of the IEEE International Conference on Image Processing (ICIP), 1884-1887.

GUANTER, L., SEGL, K. & 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), 2340-2351.

GUYON, I. & ELISSEEFF, A., 2003: An introduction to variable and feature selection. Journal of Machine Learning Research **3**, 1157-1182.

HALL, M. A., 1999: Correlation-based feature subset selection for machine learning. PhD thesis, Department of Computer Science, University of Waikato, New Zealand.

HAM, J., CHENG, Y., CRAWFORD, M. M. & GHOSH, J., 2005: Investigation of the random forest framework for classification of hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing **43**(3), 492-501.

HUGHES, G. F., 1968: On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory **14**(1), 55-63.

JOELSSON, S. R., BENEDIKTSSON, J. A. & SVEINSSON, J. R., 2005: Random forest classifiers for hyperspectral data. Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS), 160-163.

KAUFMANN, H., SEGL, K., GUANTER, L., HOFER, S., FOERSTER, K.-P., STUFFLER, T., MUELLER, A., RICHTER, R., BACH, H., HOSTERT, P. & CHLEBEK, C., 2008: Environmental Mapping

and Analysis Program (EnMAP) – Recent advances and status. Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS) **4**, 109-112.

KELLER, S., BRAUN, A. C., HINZ, S. & WEINMANN, M., 2016: Investigation of the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data. Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-6.

LE BRIS, A., CHEHATA, N., BRIOTTET, X. & PAPARODITIS, N., 2014: Use intermediate results of wrapper band selection methods: a first step toward the optimization of spectral configuration for land cover classifications. Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-4.

LICCIARDI, G., MARPU, P. R., CHANUSSOT, J. & BENEDIKTSSON, J. A., 2012: Linear versus nonlinear PCA for the classification of hyperspectral data based on the extended morphological profiles. IEEE Geoscience and Remote Sensing Letters **9**(3), 447-451.

MELGANI, F. & BRUZZONE, L., 2004: Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing **42**(8), 1778-1790.

PLAZA, A., BENEDIKTSSON, J. A., BOARDMAN, J. W., BRAZILE, J., BRUZZONE, L., CAMPS-VALLS, G., CHANUSSOT, J., FAUVEL, M., GAMBA, P., GUALTIERI, A., MARCONCINI, M., TILTON, J. C. & TRIANNI, G., 2009: Recent advances in techniques for hyperspectral image processing. Remote Sensing of Environment **113**, S110-S122.

SAEYS, Y., INZA, I. & LARRANAGA, P., 2007: A review of feature selection techniques in bioinformatics. Bioinformatics **23**(19), 2507-2517.

SEGL, K., GUANTER, L., KAUFMANN, H., SCHUBERT, J., KAISER, S., SANG, B. & 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), 3046-3054.

TARABALKA, Y., CHANUSSOT, J., BENEDIKTSSON, J. A., ANGULO, J. & FAUVEL, M., 2008: Segmentation and classification of hyperspectral data using watershed. Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS) **3**, 652-655.

TARABALKA, Y., FAUVEL, M., CHANUSSOT, J. & BENEDIKTSSON, J. A., 2010: SVM- and MRF-based method for accurate classification of hyperspectral images. IEEE Geoscience and Remote Sensing Letters **7**(4), 736-740.

TARABALKA, Y. & TILTON, J. C., 2011: Spectral-spatial classification of hyperspectral images using hierarchical optimization. Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-4.

VAN DER MAATEN, L. J. P., POSTMA, E. O. & VAN DEN HERIK, H. J., 2009: Dimensionality reduction: a comparative review. Technical Report TiCC-TR 2009005, Tilburg University, Tilburg, The Netherlands.

VILLA, A., BENEDIKTSSON, J. A., CHANUSSOT, J. & JUTTEN, C., 2011: Hyperspectral image classification with independent component discriminant analysis. IEEE Transactions on Geoscience and Remote Sensing **49**(12), 4865-4876.

WANG, J. & CHANG, C.-I., 2006: Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. IEEE Transactions on Geoscience and Remote Sensing **44**(6), 1586-1600.