Remote Sensing of Environment 113 (2009) 1712–1723

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA) of hyperspectral imagery for urban environments

Jonas Franke [a,*], Dar A. Roberts [b,1], Kerry Halligan [c,2], Gunter Menz [d,3]

[a] University of Bonn, Center for Remote Sensing of Land Surfaces (ZFL), Walter-Flex-Strasse 3, D-53113 Bonn, Germany
[b] University of California, Santa Barbara, Department of Geography, 1832 Ellison Hall, UC Santa Barbara, Santa Barbara, CA 93106-4060, United States
[c] Sanborn Map Company, 610 SW Broadway, Suite 310, Portland, OR 97205, United States
[d] University of Bonn, Department of Geography, Remote Sensing Research Group (RSRG), Meckenheimer Allee 166, 53115 Bonn, Germany

ARTICLE INFO

Article history:
Received 15 December 2008
Received in revised form 27 March 2009
Accepted 28 March 2009

Keywords:
Multiple Endmember Spectral Mixture Analysis (MESMA)
Hyperspectral Mapper (HyMap)
Urban
Land cover
Hyperspectral
Imaging spectrometry
Endmember selection
Hierarchical classification

ABSTRACT

Remote sensing has considerable potential for providing accurate, up-to-date information in urban areas. Urban remote sensing is complicated, however, by very high spectral and spatial complexity. In this paper, Multiple Endmember Spectral Mixture Analysis (MESMA) was applied to map urban land cover using HyMap data acquired over the city of Bonn, Germany. MESMA is well suited for urban environments because it allows the number and types of endmembers to vary on a per-pixel basis, which allows controlling the large spectral variability in these environments. We employed a hierarchical approach, in which MESMA was applied to map four levels of complexity ranging from the simplest level consisting of only two classes, impervious and pervious, to 20 classes that differentiated material composition and plant species. Lower levels of complexity, mapped at the highest accuracies, were used to constrain spatially models at higher levels of complexity, reducing spectral confusion between materials. A spectral library containing 1521 endmembers was created from the HyMap data. Three endmember selection procedures, Endmember Average RMS (EAR), Minimum Average Spectral Angle (MASA) and Count Based Endmember Selection (COB), were used to identify the most representative endmembers for each level of complexity. Combined two-, three- or four-endmember models – depending on the hierarchical level – were applied, and the highest endmember fractions were used to assign a land cover class. Classification accuracies of 97.2% were achieved for the two lowest complexity levels, consisting of impervious and pervious classes, and a four class map consisting of vegetation, bare soil, water and built-up. At the next level of complexity, consisting of seven classes including trees, grass, bare soil, river, lakes/basins, road and roof/building, classification accuracies remained high at 81.7% with most classes mapped above 85% accuracy. At the highest level, consisting of 20 land cover classes, a 75.9% classification accuracy was achieved. The ability of MESMA to incorporate within-class spectral variability, combined with a hierarchical approach that uses spatial information from one level to constrain model selection at a higher level of complexity was shown to be particularly well suited for urban environments.

1. Introduction

Current and accurate information about urban composition is critical for urban planning, disaster response and improved environmental management. Remote sensing has the potential to provide the necessary information about urban infrastructure, socio-economic attributes and environmental conditions at a diversity of scales (Jensen & Cowen, 1999; Small, 2001). As a result, an increasing number of studies have focused on remote sensing of urban environments and their land cover (e.g., Ben-Dor et al., 2001, Herold & Roberts, 2005; Powell et al., 2007; Small, 2001, 2003, 2005).

Urban remote sensing is complicated by the complexity of urban environments which includes considerable spectral diversity at very fine spatial scales (Powell et al., 2007; Small, 2001, 2005). Spectrally, urban areas are complicated by the presence of numerous spectrally unique materials, and the presence of spectrally ambiguous materials, such as dark-shingles and asphalt roads (Herold et al., 2003a). Other factors further complicate analysis, including non-Lambertian behavior of urban materials that leads to high within-class spectral variability (Herold et al., 2004), 3-dimensional heterogeneity of urban areas (Herold et al., 2003a) and material aging, which causes spectral changes (Herold & Roberts, 2005). As a result, urban environments exhibit a high dimensionality in spectral space (Small, 2001, 2005).

* Corresponding author. Tel.: +49 228 73 4023; fax: +49 228 73 6857.
E-mail addresses: jonasfranke@uni-bonn.de (J. Franke), dar@geog.ucsb.edu (D.A. Roberts), halligan.kerry@gmail.com (K. Halligan), g.menz@uni-bonn.de (G. Menz).
[1] Tel.: +1 805 893 2276; fax: +1 805 893 3146.
[2] Tel.: +1 503 228 8708; fax: +1 503 228 8751.
[3] Tel.: +49 228 73 9700; fax: +49 228 73 9702.

0034-4257/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2009.03.018

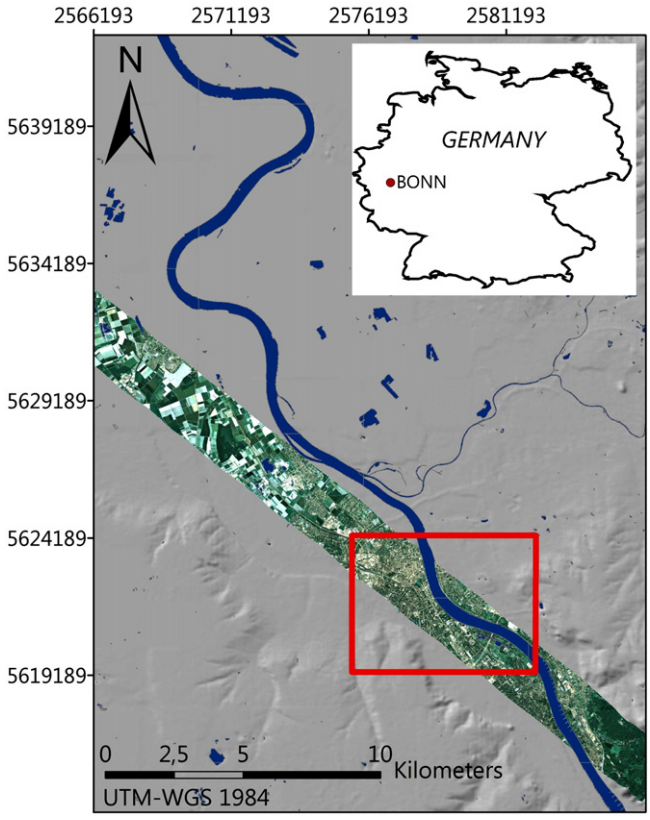

Fig. 1. Location of the study area of Bonn in Germany. Shaded relief data set with the river Rhine given in blue, combined with a swath of true-color HyMap data acquired on May 28, 2005 indicating the landscape structure.

Some studies, therefore, have focused on the spectral separability of urban materials for land cover mapping, using sensors such as the Airborne Visible Near Infrared Imaging Spectrometer (AVIRIS) (Hepner et al., 1998; Hepner & Chen, 2001; Herold et al., 2004) or the Compact Airborne Spectrographic Imager (CASI) (Ben-Dor et al., 2001). Herold et al. (2004) concluded that the current knowledge about urban materials and their separation based on spectral characteristics is inadequate. In addition to the large diversity of urban materials, improved knowledge about spatio-temporal changes in urban vegetation cover is important to determine and model urban environmental conditions (Ridd, 1995; Small, 2001).

Limitations of remote sensing techniques for urban mapping in the spatial dimension, as observed in previous studies (Powell et al., 2007; Small, 2001, 2005), resulting from coarse spatial resolutions of sensors such as Landsat Enhanced Thematic Mapper (ETM+), can potentially be overcome by novel analysis techniques. For urban applications, a spatial resolution of at least 5 m is required, in order to adequately capture urban structures (Small, 2003). However, due to the high spatial variability of urban structure with spectrally heterogeneous materials close to each other, mixed pixels are still common in images covering urban areas (Powell et al., 2007). Spectral mixture analyses (SMA) can potentially solve some of the problems associated with the spectral heterogeneity of urban surfaces (Small, 2001). However, simple mixing models, which consist of a single set of endmembers applied to an entire scene, are potentially not appropriate for urban areas because they cannot account for considerable within-class variability (Rashed et al., 2003; Roessner et al., 2001). To overcome this limitation, Song (2005) proposed Bayesian spectral mixture analysis (BSMA), in which endmembers are not treated as constants. Multiple Endmember Spectral Mixture Analysis (MESMA; Roberts et al., 1998) represents an alternative approach, in which the number and types of endmembers are allowed to vary on a per-pixel basis, thereby accounting for urban spectral heterogeneity. For example,

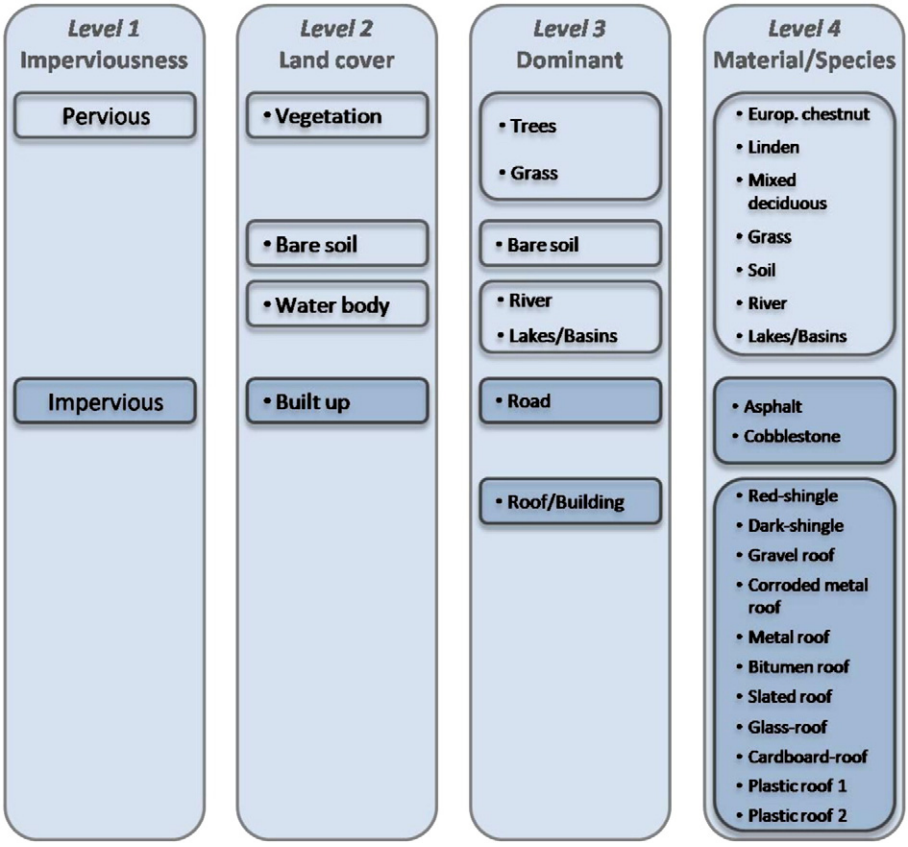

Fig. 2. Hierarchical structure of the Multiple Endmember Spectral Mixture Analysis giving 4 levels of different complexity.

MESMA has been used to map vegetation, impervious and soil fractions in a number of urban areas, including the city of Manaus, Brazil (Powell et al., 2007) and Los Angeles (Rashed et al., 2003).

Object-oriented approaches also have considerable potential for mapping urban areas at high accuracy (Herold et al., 2003b). Benediktsson et al. (2005) also discuss the importance of spatial and spectral information describing a morphological method for a joint spatial/spectral classifier for urban environments. Roessner et al. (2001) incorporated spatial context to improve endmember selection to iteratively unmix hyperspectral data covering an urban area. In this study, we propose hierarchical MESMA, in which different models are used for different levels of complexity, and in which highly accurate models at the lowest level of complexity are used to spatially constrain models at higher levels of complexity. We tested this approach using HyMap data acquired over the city of Bonn, Germany. Our objectives were thereby to (i) demonstrate the potential of MESMA for mapping urban land cover at various levels of detail ranging from imperviousness to material discrimination, (ii) determine materials or land covers with a high degree of spectral confusion (iii) incorporate spatial constraints into MESMA to improve classification accuracies and thus (iv) to prove this hierarchical MESMA approach for urban environments.

2. Methodology

2.1. Study area and data

This study focused on an urban transect in the region of Bonn, Germany. The city of Bonn is located in Western Germany along the river Rhine, approximately 30 km southeast of the city of Cologne. Bonn represents a typical large German city with a population of 320,000. The city dates from Roman times and contains a medieval city center with large 19th and 20th century urban extensions. Bonn is situated in generally level terrain at an average elevation of 55 m above sea level. Although the city includes a few taller buildings (e.g., the 165 m Posttower) it is dominated by 3–5 stories commercial and residential structures as well as 1 and 2 floor family houses, principally located in residential areas that are concentrically arranged around the town center. The Bonn area is characterized by highly diverse land cover types and complex urban material compositions.

Fig. 1 shows the Bonn study site location and an airborne Hyperspectral Mapper (HyMap) subscene acquired on 28/05/2005 covering a representative NW–SE cross-section of the city, emphasizing the highly diverse land use regime present within the study area. A number of urban land covers are shown: residential areas (with differing densities and socio-economic structures), mixed-use areas, and commercial and industrial districts. Non-urban land cover types include water bodies, green vegetation and bare soils. The specific urban materials present in the Bonn region result from centuries of urban development combined with local traditional influences. The diversity of building materials found in the area includes asphalt and cobblestone road surfaces, as well as roofs composed of slate, metal, glass, gravel, bitumen, plastics and a number of types and colors of composite shingles. Old houses in the historic town center typically have shingle roofs composed of dark slate from the nearby Rhenish Slate Mountains. Predominant vegetation types in the study area

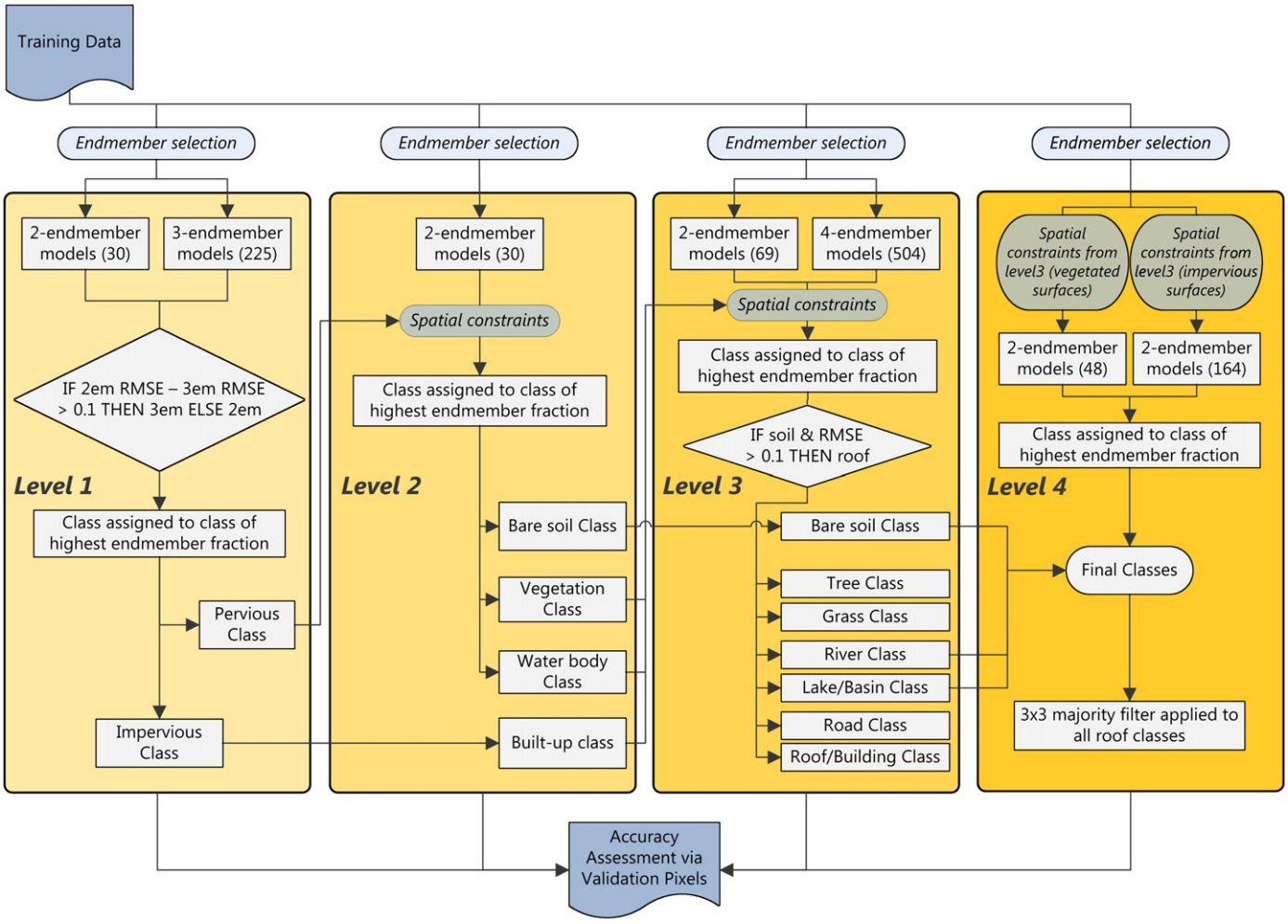

Fig. 3. Workflow of the hierarchical MESMA.

include European chestnut, linden and other mixed deciduous trees, as well as riparian areas that are mostly covered with grass.

Airborne Hyperspectral Mapper (HyMap) data were acquired on 28/05/2005. The HyMap-system is a whisk-broom scanner with an ax-head double mirror which acquires 126 spectral bands with a bandwidth of 16 nm (in the VIS and NIR region) in the spectral range between 450 nm and 2480 nm. HyMap is typically operated at altitudes between 2000 and 5000 m agl and has an instantaneous field of view (IFOV) of 2.5 mr along track and 2.0 mr across track by a field of view (FOV) of $\pm 30°$ (Cocks et al., 1998). The chosen configuration resulted in a nominal ground IFOV of 4.0 m. HyVista Corp. and the German Aerospace Center (DLR) carried out the pre-processing of the HyMap data. The validation of the atmospheric corrections using ATCOR 4 (Richter and Schlaepfer, 2002) was performed against *in-situ* measurements obtained with an ASD FieldSpec Pro spectroradiometer (Analytical Spectral Devices Inc., Boulder, Co, USA). During further pre-processing, 26 bands showing high levels of noise, especially those near the water absorption features at 1400 and 1900 nm, were removed from the data to improve overall image quality.

2.2. Endmember selection

The quality of SMA results, in general, is highly dependent on the availability of representative endmembers (Tompkins et al., 1997). Endmembers used in SMA can either be derived from image pixels or from a spectral library that contains reference endmembers derived from measurements taken in the field, laboratory, from radiative transfer models (Sonnentag et al., 2007) or derived from other images. The advantage of image endmembers is that they can be collected at the same scale as the image and are easier to associate with image features (Rashed et al., 2003). In addition, image endmembers have the advantage of being canopy-scale spectra, incorporating the effects of non-linear mixing which may be present, especially in vegetation surfaces, at the leaf-scale. Several approaches for the selection of optimal/representative endmembers from images have been developed. One example is the Pixel Purity Index (PPI), described by Boardman et al. (1995). Recently, several endmember selection approaches have been proposed for MESMA, in which a spectral library is analyzed to identify the subset of spectra that are most representative of a specific class, and least confused with members outside that class. Examples include Count-based Endmember Selection (CoB) (Roberts et al., 2003), Endmember Average Root Mean Square Error (EAR) (Dennison & Roberts, 2003) and the Minimum Average Spectral Angle (MASA) (Dennison et al., 2004). In contrast to the PPI, these approaches require knowledge about the spectral characteristics to assign each spectrum to a particular class. Using CoB, optimal endmembers are identified as those spectra within a spectral library that model the greatest number of spectra within a specific class while meeting other selection criteria, including fractional constraints (i.e. fractions are required to be between 0 and 100%), and fit constraints based on RMSE and spectral residuals (Roberts et al., 2003). CoB provides several quality parameters that allow for a ranking of representative endmembers. The in-CoB parameter reports the total number of spectra modelled within the class, whereas the out-CoB reports the total number of models outside of the class. A high in-CoB with a low out-CoB represents an excellent endmember choice. EAR calculates the average RMSE produced by a spectrum when it is used to model all other members of its class. The optimum spectrum produces the lowest average RMSE.

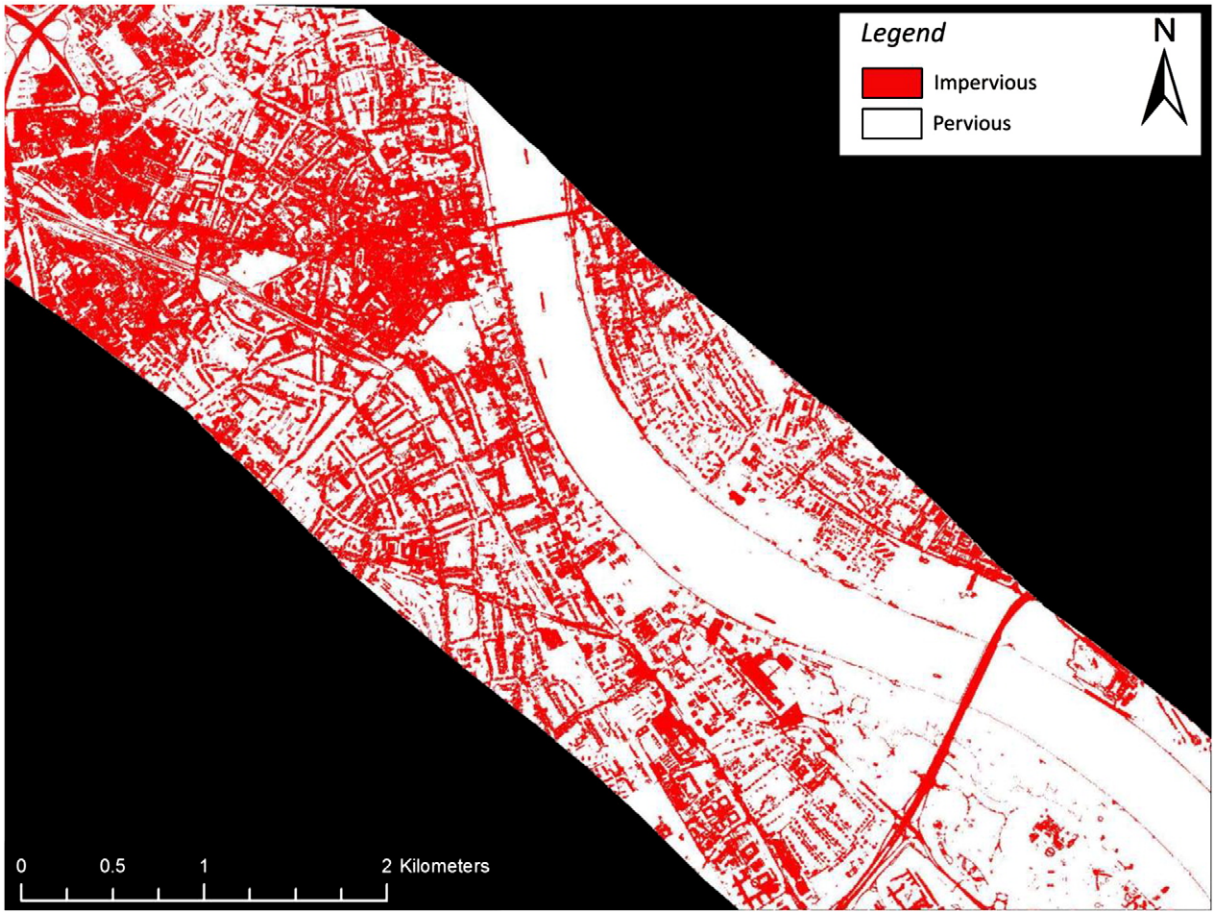

Fig. 4. Merged classification result of level 1 as derived from two- and three-endmember model MESMA, showing impervious and pervious surfaces in the area of Bonn with an overall classification accuracy of 97.2%.

The MASA calculates the average spectral angle between a reference spectrum (candidate model) and all other spectra within the same class. The best endmember is selected as the one that produces the lowest average spectral angle (Dennison et al., 2004).

In this study, a comprehensive survey of the area of Bonn was conducted, in order to identify representative surface types and to find locations suitable for developing a spectral library from the image. A Differential Global Position System (Trimble GeoExplorerXT with a Trimble GeoBeacon receiver) was used to collect reference data with a high spatial accuracy. Three to five Regions of interest (ROI), consisting of around 30 pixels per ROI were selected in the HyMap image for each endmember type. A random sample of all ROIs was extracted, to develop a spectral library consisting of 1521 spectra with 1384 pixels set aside for validation of the MESMA results. The spectral library was developed using 'VIPER-tools', an ENVI add on (www.vipertools.org), and all relevant metadata added. For each hierarchical MESMA level, separate spectral libraries were created containing optimal/most representative endmembers with low probability of confusion with other endmember classes, characterized by low EAR or MASA values or high in-Cob values. These metrics were calculated with the 'VIPER-tools' and endmembers were consecutively sorted by their metric values. The spectra selection should be done on a case by case basis, depending on the users' objective. A number of different strategies may be employed. In our study, most of our selections were from the top candidates of each metric.

2.3. Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA)

Linear SMA assumes that a mixed spectrum can be modeled as a linear combination of pure spectra, known as endmembers (Adams et al., 1986). Under ideal conditions, the most accurate fractional estimates can be achieved using the minimum number of endmembers required to account for spectral variability within a mixed pixel (Sabol et al., 1992). Fractional errors occur either when too few endmembers are used, resulting in spectral information that cannot be accounted for by the existing endmembers, or too many, in which case minor departures between measured and modelled spectra are often assigned to an endmember that is used in the model, but not actually present (Roberts et al., 1998). Urban environments are particularly difficult for a simple mixture model because a single endmember cannot account for considerable spectral variation within a class. In contrast, MESMA can account for within-class variability and thus is likely to be more suitable for urban remote sensing. Typically, MESMA is applied by running numerous models for a pixel and selecting one model based on its ability to meet selection criteria and produce the best fit, typically a minimum RMS (Painter et al., 1998). Selection criteria include fractional constraints (minimum and maximum fraction constrains), maximum allowable shade fraction, RMSE constraints and a residual constraint set to remove any model that exceeds a threshold over a range of wavelengths. Using this approach, pixel-scale limits in spectral dimensionality are recognized while also accounting for considerable spectral variability within a scene. The constraints for the models are variably selectable, whereby MESMA can also be run in an unconstrained mode. Roberts et al. (1998) found that with the flexible MESMA approach, a majority of pixels in an image could be modeled with only two-endmember models. Powell and Roberts (2008) found that natural landscapes in Brazil required only two-endmember models, disturbed regions required three- and urban areas required four-endmember models.

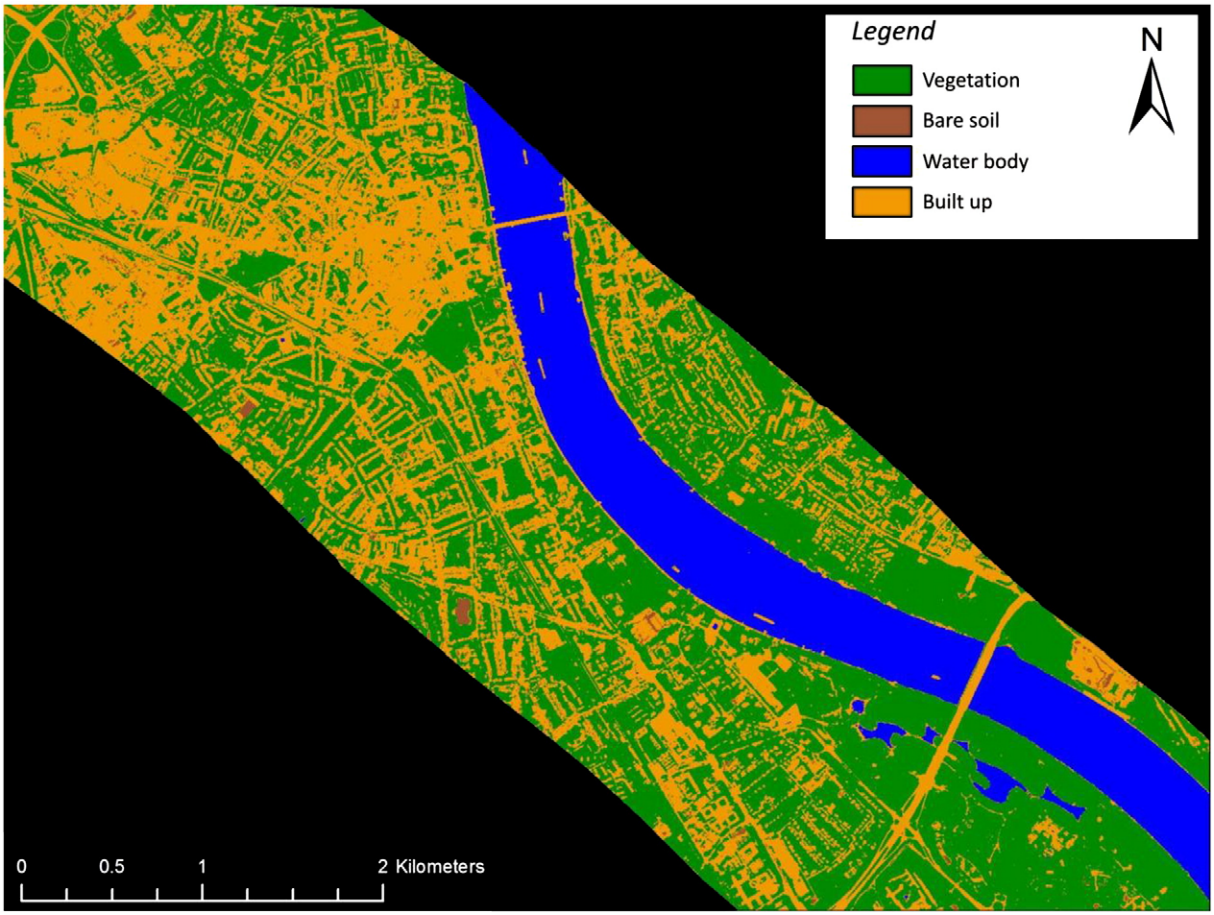

Fig. 5. Classification result of level 2 displays the land cover classes 'Vegetation', 'Built-up', 'Bare soil' and 'Water body' in the urban area of Bonn with an overall classification accuracy of 97.2%.

In this study, a hierarchically structured MESMA was realized, whereby two-, three- or four-endmember models were uniquely tailored for four different levels of complexity (Fig. 2). At the lowest level only two broad classes were mapped, impervious and pervious. At level 2, four land cover classes were mapped; at the third level, the land cover classes were subdivided into up to two categories, such as grass and trees, roads vs. roof materials and lake vs. river. At the highest level of complexity, 20 classes of specific materials or tree species were mapped (Fig. 2).

The basic idea of hierarchical MESMA was to use the result from one level as a spatial constraint for the next level, taking advantage of higher classification accuracies achieved at lower levels of complexity to improve accuracies at higher levels. For example, models for level 2 were constrained by results from level one, where vegetation, bare soil and water bodies are restricted to areas mapped as pervious, and built-up is restricted to the impervious class. In this case, the results for level 1 and level 2 are the same for impervious and built-up. MESMAs of levels 1 and 2 were run unconstrained and as spatial constraints only those masks were used, in cases when classification accuracy was greater than 85% at a lower level. MESMAs of level 3 were run partially constrained (minimum and maximum allowable fractions were constrained). By using spatial information at one level for the analysis of the next level, high classification accuracies achieved at the lower levels should therefore improve accuracies at higher levels. The workflow of this hierarchical MESMA is displayed in Fig. 3.

For MESMA of the imperviousness at level 1, 30 two-endmember models were used, whereby the first spectral library contained 30 representative endmembers (15 representing pervious and 15 representing impervious surfaces) determined by EAR, MASA and Cob and the second spectral library contained shade. In addition, a three-endmember model was applied with 15 pervious endmembers in the first library, 15 impervious endmembers in the second and shade in the third spectral library resulting in 225 models. The RMSE change between the results of the two- and three-endmember model results was calculated afterwards. In cases where the RMSE did not change more than 0.1 between both results, the result of the two-endmember model was chosen. The three-endmember model was only selected where a third endmember was needed to drop the RMSE (RMSE change >0.1). Both MESMA classification results (classes were assigned to the highest endmember fraction) were merged to the two final classes impervious and pervious.

MESMA of the second level discriminated 4 different land cover classes by the use of 30 two-endmember models similar to level 1. MESMA was run in an unconstrained mode, whereas the analysis was spatially constrained, because the land cover classes vegetation, bare soil and water body were only analyzed for pervious areas as identified at the first level. Due to the fact that spatial constraints were used, the land cover class 'Built-up' at level 2 is the same as the impervious class from level 1. Over most parts of the region RMSE values at level 2 were low, which indicates that this level could be successfully modeled with only two-endmember models.

MESMA at the third level determined dominant surface types including trees, grass, bare soils, rivers, lakes/basins, roads and roofs/buildings. First, 69 two-endmember models were used, applied to each land cover mask as derived from the results of level 2 (excepting bare soil). MESMA was run in a partially constrained mode (minimum and maximum allowable fraction values were constrained). A four-endmember model was run additionally in order to improve discrimination between vegetation types as well as roads and roofs/buildings.

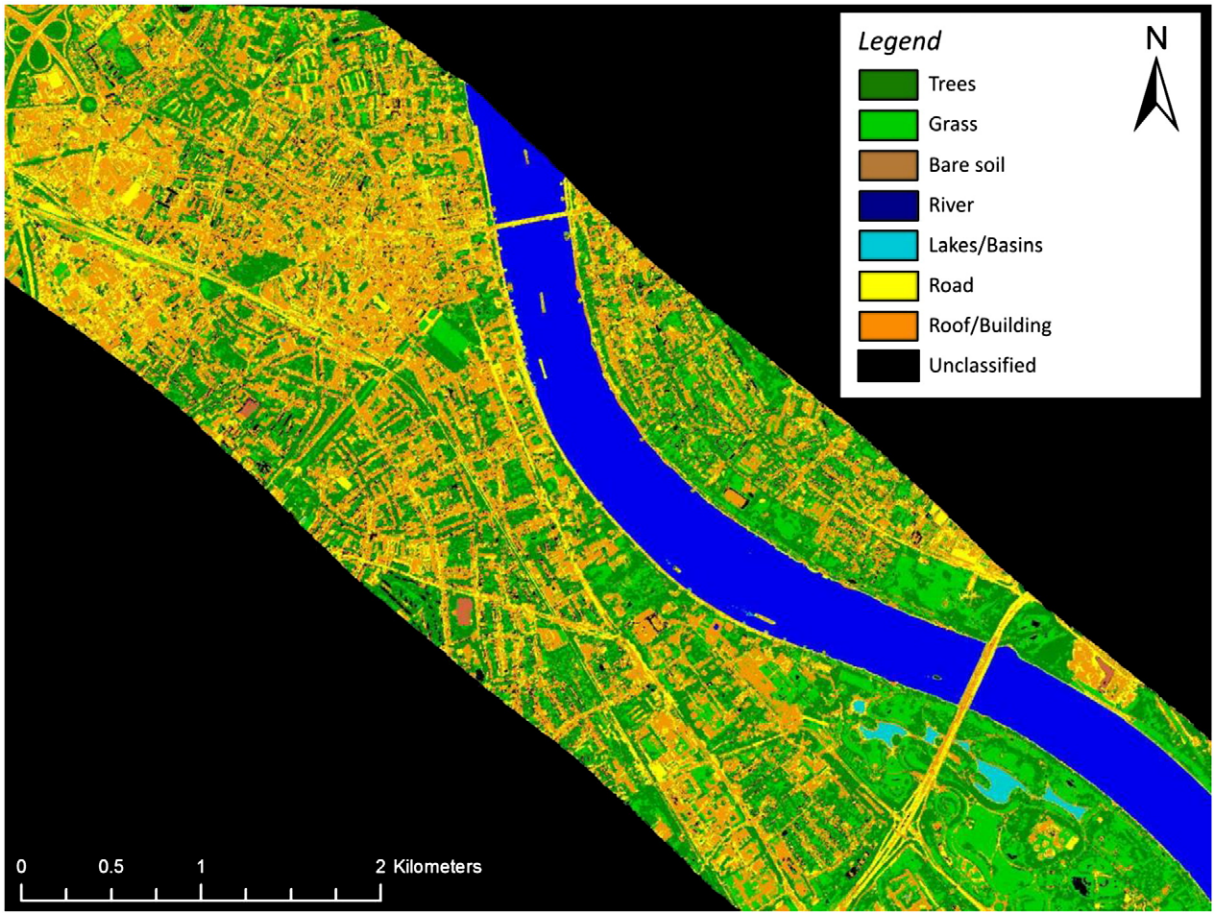

Fig. 6. Classification result of level 3 as derived from two- and four-endmember model MESMA that gives the dominant surface types trees, grass, bare soils, rivers, lakes/basins, roads and roofs/buildings with an overall classification accuracy of 81.7%.

The so called V–I–S model was proposed by Ridd (1995), which divides urban areas into three physically based classes, vegetation, impervious surfaces and soil. In the present study, the first spectral library contained 12 vegetation endmembers, the second library contained 14 endmembers representing impervious surfaces, the third library contained 3 soil endmembers and the fourth spectral library contained shade resulting in 504 models. Classes were assigned to the highest endmember fractions. The RMSE was used as a constraint for each class. Due to a significant confusion between soil and red-shingle roofs, a maximum RMSE of 0.025 was applied as a constraint for the soil class. All pixels with dominant soil fractions and RMSE values greater than 0.1, were assigned to the roof class. If no model met all these constraints, the pixel was left unmodeled/unclassified. The classification result of level 3 was merged from the results of the two-endmember and four-endmember models.

MESMA of level 4 discriminated 20 different materials or vegetation species as shown in Fig. 2. Due to the fact that the classes 'River', 'Lakes/Basins' and 'Bare soil' were already final classes at level 2 or 3 with classification accuracies higher than 85%, the information for these classes was taken from those levels, respectively. 212 two-endmember models were used for the MESMA applied at level 4. MESMA was run in an unconstrained mode similar to level 2. 48 two-endmember models were used for all vegetated pixels as identified at level 3 and 164 two-endmember models were used for all pixels not classified as water, vegetated or bare soil at level 3. Results of the discrimination of vegetation species and urban materials as obtained from MESMAs of level 4 were merged with final classes already obtained at levels 2 and 3. Minor classification errors were present in some buildings, represented by individual pixels of a different class imbedded within an otherwise compact building object. To reduce this type of error, the building classes were smoothed using a 3*3 majority filter to remove single-pixel errors within buildings. All other classes remained unfiltered.

Classification results of each hierarchical level were compared to the random sample of validation pixels, in order to assess classification accuracy, whereby the total sample size of 1384 pixel splits – depending on the hierarchical level – to the final classes.

3. Results and discussion

Figs. 4–7 show the classification results of the 4 hierarchical levels. The MESMA results of levels 1 to 3 (Figs. 4–6) reveal the spatial structure of the urban area of Bonn with mostly impervious areas in the central business district (CBD) in the northern part of the scene close to the Rhine River and in the strongly industrial area in the northwest. Residential areas, in contrast, are clearly distinguishable by a higher percentage of vegetated areas. In the southern part of the scene, the densely vegetated recreation area 'Rheinaue' is obvious, which also acts as a natural floodplain for the Rhine River, which occasionally floods. Observing the result of hierarchical MESMA level 4 (Fig. 7), a detailed insight into the urban spatial structure is given. Considering object sizes and roof materials, the industrial area in the northwest is clearly distinguishable from the CBD. In addition, larger objects with different roof materials are obvious in the southern part of the scene as well, that indicate the area of governmental buildings, museums, headquarters of organizations and companies etc. The spatial distribution of vegetation species also gives detailed information about urban environmental condition. The V–I–S model as

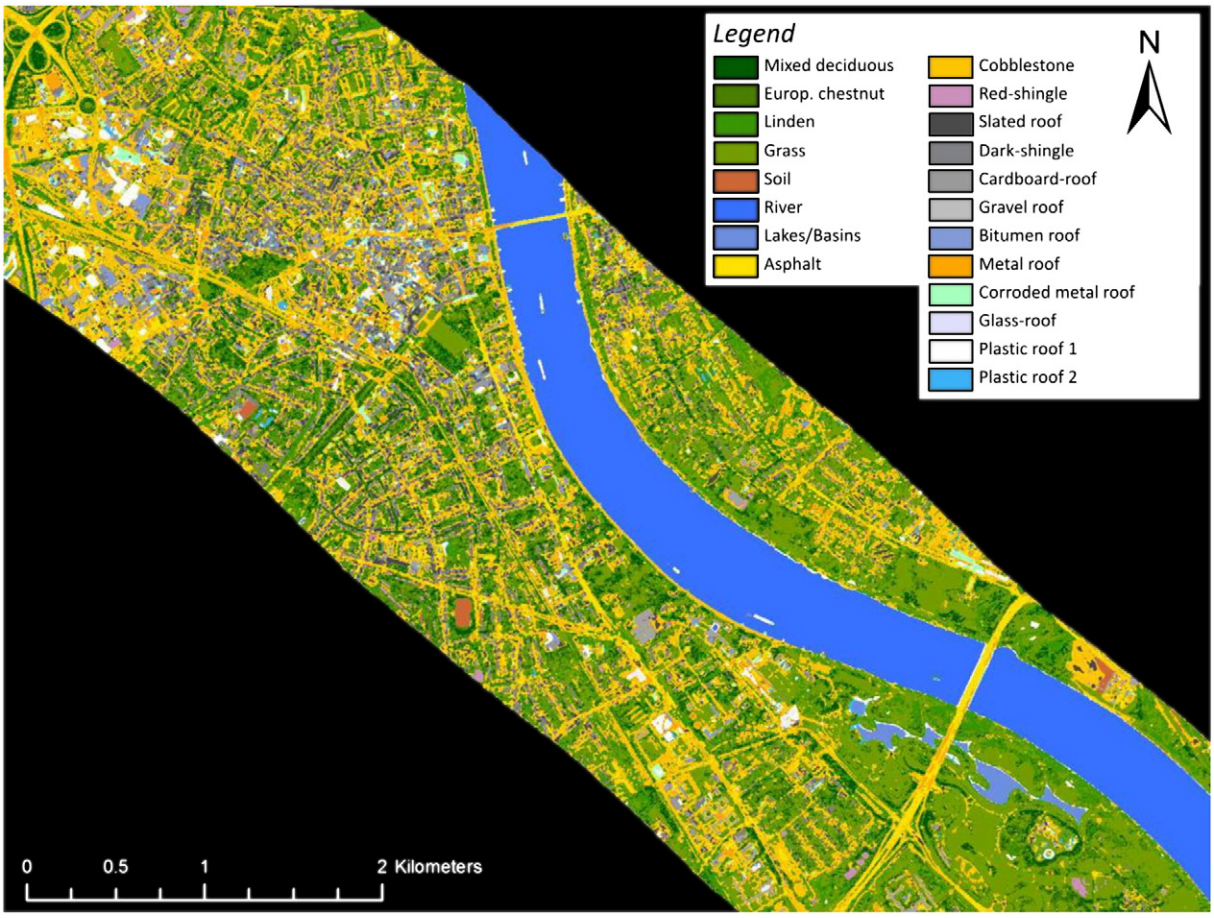

Fig. 7. Classification result of level 4 that shows different materials and vegetation species in 20 classes of the urban area of Bonn with an overall classification accuracy of 75.9%.

Fig. 8. False-color image giving the fractions of vegetation, impervious surfaces and soil as derived from MESMA by using four-endmember models (V–I–S model) at level 3. Black pixels give areas where the maximum allowable fractions or the RMSE exceed the set constraints.

described above is displayed in Fig. 8, which shows the fractions of vegetation, impervious surfaces and soil as a false-color RGB as derived by the four-endmember models at level 3. A mask considering maximum allowable fractions and RMSE constraints was thereby applied.

Error matrices were calculated using ground reference data for each hierarchical level (Tables 1–4). The overall classification accuracies were 97.2% for level 1 (kappa coefficient of 0.94) and level 2 (kappa coefficient of 0.95), 81.7% for level 3 (kappa coefficient of 0.75) and 75.9% (kappa coefficient of 0.74) for hierarchical level 4. Only minor confusion occurred at level 1 between the classes impervious and pervious (Table 1). The high accuracies of 95.4% and 100% respectively could be on the one hand achieved due to the selection of highly representative endmembers using EAR, MASA and CoB and on the other hand due to the fact that results of the two- and three-endmember models were merged, depending on the RMSE change. The endmember selection approaches selected the following representative endmembers: for pervious surfaces, 10 vegetation endmembers, 2 water and 3 soil endmembers (Fig. 9a); for impervious surfaces, 3 road endmembers and 12 roof endmembers (Fig. 9b). Using the RMSE change between the results of the two- and three-endmember models was very suitable for identifying pixels that required a third endmember to drop the RMSE and improved accuracies. In particular, the models cardboard roof/shade as well as grass/shade in the two-endmember MESMA had the highest frequency within these pixels with high RMSE change.

At level 2, 4 different land cover classes were discriminated by 30 two-endmember models, whose endmembers were specifically selected by the mentioned endmember selection procedures. Some confusion was evident for bare soil, in which 23.4% of the cases were mis-classified as built-up area (Table 2). The class 'Bare soil' had a comparatively low sample size in the validation data (47) since it already is a final class at level 2 (total sample size splits down to the final classes). Almost no confusion occurred for the other classes at level 2, whereby the classes 'Vegetation' and 'Water body' were almost perfectly classified with accuracies of 99.8% and 100%.

Table 1
Error matrix of the level 1 classification result and ground truth data gives the percentage of classification accuracy and misfit.

Classified/ground truth	Pervious	Impervious
Sample size	513	871
Pervious	100.0	4.6
Impervious	0.0	95.4

Table 2
Error matrix of the level 2 classification result and ground truth data gives the percentage of classification accuracy and misfit.

Classified/ground truth	Vegetation	Bare soil	Built-up	Water body
Sample size	398	47	871	68
Vegetation	99.7	0.0	0.9	0.0
Bare soil	0.0	76.6	2.2	0.0
Built-up	0.3	23.4	96.9	0.0
Water body	0.0	0.0	0.0	100.0

Table 3
Error matrix of the level 3 classification result and ground truth data gives the percentage of classification accuracy and misfit.

Classified/ ground truth	Road	Grass	Trees	River	Lakes/basins	Roof/building	Bare soil
Sample size	161	141	257	50	18	710	47
Unclassified	1.9	0.0	0.0	0.0	0.0	7.9	0.0
Road	78.3	0.0	0.0	0.0	0.0	15.8	12.8
Grass	0.0	85.8	7.0	0.0	0.0	0.6	0.0
Trees	0.0	14.2	93.0	0.0	0.0	0.1	0.0
River	0.0	0.0	0.0	100.0	0.0	0.0	0.0
Lakes/basins	0.0	0.0	0.0	0.0	100.0	0.0	0.0
Roof/building	19.8	0.0	0.0	0.0	0.0	75.5	2.1
Bare soil	0.0	0.0	0.0	0.0	0.0	0.1	85.1

In order to determine different surface types, two- and four-endmember models were run at level 3 using specifically selected endmembers. Results of the four-endmember models were picked where classification accuracies resulting from the two-endmember models were low. The classes 'River' and 'Lakes/Basins' could be perfectly discriminated by the two-endmember model, whereas results of the four-endmember model were used for the other classes. Major confusion in the merged results at level 3, particularly occurred between the classes 'Roof/Building' and 'Road' with mis-classification rates of 19.9% and 15.8% respectively. This result is consistent with previous work by Herold et al. (2004) and Herold and Roberts (2006), who found considerable confusion between roofs and roads due to spectral ambiguity. In that study, vertical height information derived from LiDAR data was used to remove this confusion (Herold & Roberts, 2006). In addition, confusion between 'Grass' and 'Trees' occurred with mis-classification rates of 14.2% and 7.0%, respectively. However, a high overall classification accuracy of 81.7% was achieved in a range from 75.5% (Roof/Building) to 100% (River and Lakes/Basins). Benediktsson et al. (2005) as well as Roessner et al. (2001) reported similar accuracies for some classes defined in their studies that are similar to the land cover classes defined in the present study. As discussed previously in the methods, considerable confusion occurred between soil and the red-shingle roof endmembers, resulting in mis-classification of soil as an impervious surface and a high RMSE. Detailed analysis of endmembers selected in the high RMSE pixels showed that a four-endmember model consisting of *grass/gravel-roof/soil/shade* was most common in these areas.

At hierarchical level 4, a further discrimination of 20 classes of different urban materials or vegetation species was accomplished. 48 two-endmember models were applied to all vegetated pixels as identified at level 3 and 164 two-endmember models to those areas which were not defined as final classes at previous levels. Using a spatially constrained MESMA, an overall accuracy of 75.9% was achieved, whereby for 12 classes, accuracies exceeded 80%, with 8 exceeding 90%. By applying a majority filter to all building classes, the overall accuracy could be improved by 2.2%. For the vegetation classes, major confusion occurred between the tree species European chestnut and mixed deciduous (55.7% misfit), whereas Linden and grass were mapped at accuracies exceeding 84%. One challenge in developing a spectral library of optimal tree species endmembers in urban areas was changes in below canopy substrates, with the same tree species occurring on different backgrounds with different amounts of sub-pixel mixing, resulting in a change in spectral shape. Of the impervious materials, slate showed the greatest confusion with other materials such as asphalt, cobblestone, dark-shingle and metal roofs. Slate is widely used roof material in the area of Bonn. It is also very dark, which most likely led to confusion with other low reflectance, spectral similar materials including asphalt, cobblestone and dark-shingle (Herold et al., 2004).

Most materials were mapped at relatively high accuracies at level 4 using only two endmembers to map materials and discriminate

Table 4
Error matrix of the level 4 classification result and ground truth data gives the percentage of classification accuracy and misfit.

Classified/ ground truth	Asphalt	Cobblestone	Bitumen roof	Cardboard roof	Corroded metal	Dark-shingle	Glass roof	Gravel roof	Metal roof	Plastic roof 1	Plastic roof 2	Red-shingle	Slate roof	River	Lakes/ basins	Grass	Mixed deciduous	European chestnut	Linden	Bare soil
Sample size	81	80	81	39	32	84	26	73	80	86	80	61	68	50	18	141	96	79	82	47
Asphalt	80.3	2.5	0.0	0.0	0.0	0.0	0.0	2.7	3.7	0.0	1.3	0.0	19.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cobblestone	7.3	67.5	2.5	5.1	0.0	6.0	0.0	17.8	6.3	0.0	1.2	0.0	17.6	0.0	0.0	0.0	0.0	0.0	0.0	14.9
Bitumen roof	0.0	0.0	79.0	0.0	0.0	0.0	0.0	0.0	7.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cardboard roof	0.0	8.8	0.0	94.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Corroded metal	0.0	0.0	0.0	0.0	93.7	0.0	0.0	0.0	1.3	0.0	0.0	0.0	5.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Dark-shingle	0.0	0.0	9.9	0.0	0.0	57.1	0.0	11.0	9.9	0.0	0.0	0.0	7.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Glass roof	0.0	0.0	0.0	0.0	6.3	0.0	80.8	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Gravel roof	0.0	1.3	4.9	0.0	0.0	0.0	0.0	64.4	7.5	0.0	0.0	0.0	1.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Metal roof	12.4	2.5	3.7	0.0	0.0	0.0	19.2	2.7	58.8	0.0	0.0	0.0	25.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Plastic roof 1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	1.3	100.0	2.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Plastic roof 2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	95.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Red-shingle	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Slate roof	0.0	17.4	0.0	0.0	0.0	32.1	0.0	0.0	0.0	0.0	0.0	0.0	20.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
River	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
Lakes/basins	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
Grass	0.0	0.0	0.0	0.0	0.0	4.8	0.0	0.0	0.0	0.0	0.0	0.0	2.9	0.0	0.0	100.0	11.5	0.0	1.2	0.0
Mixed deciduous	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	58.3	55.7	8.5	0.0
European chestnut	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	44.3	6.1	0.0
Linden	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	28.1	0.0	84.2	0.0
Bare soil	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	85.1

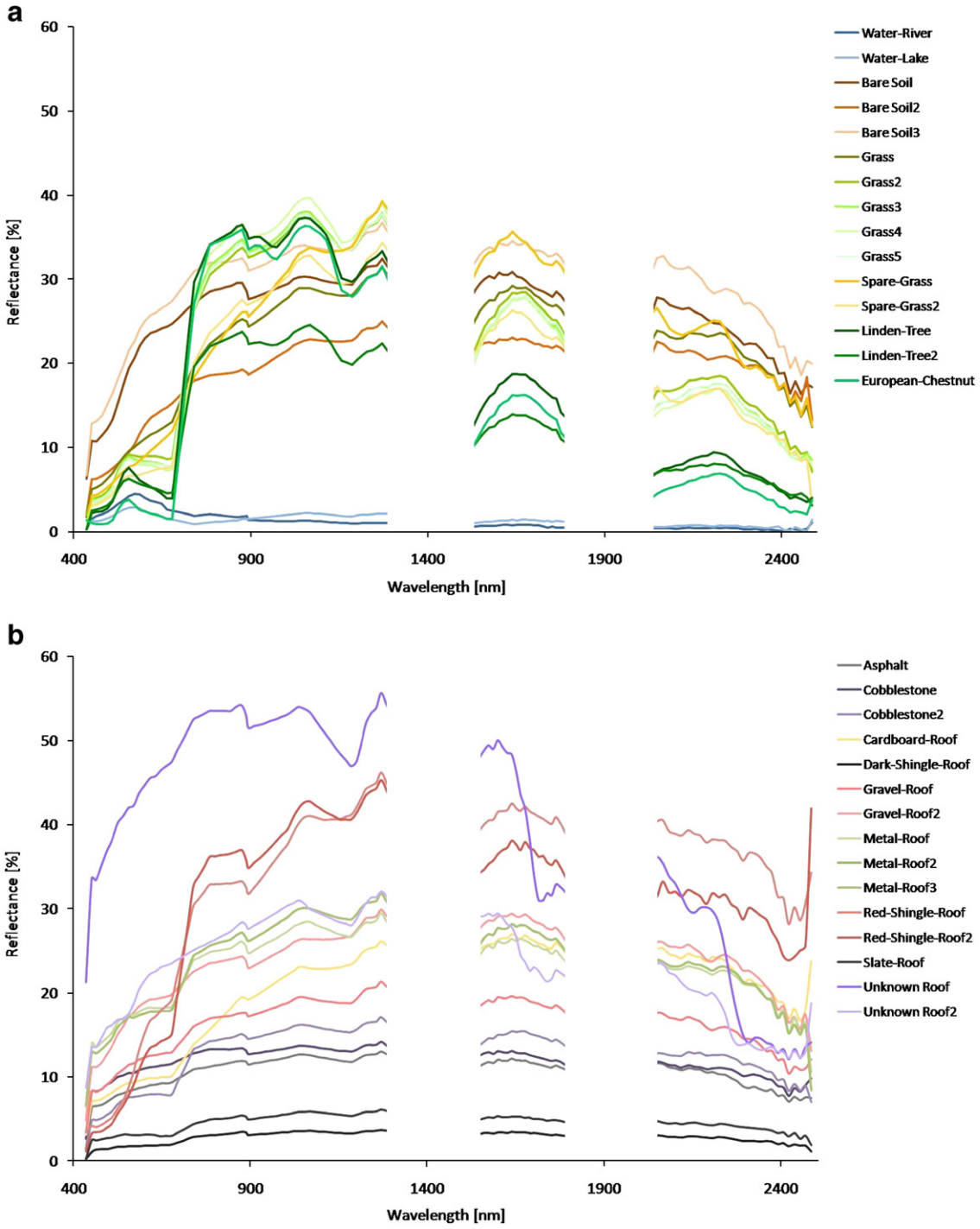

Fig. 9. a: Spectral library as used in level 1 for MESMA showing the selected pervious endmembers. b: Spectral library as used in level 1 for MESMA showing the selected impervious endmembers.

vegetation species. The advantage of two-endmember models has already been discussed by Rashed et al. (2003). Uncertainties in these results can be caused by various factors. First of all, even though the endmembers were derived from the image, it can still be very difficult to locate enough endmembers to represent the heterogeneity of urban surfaces over the entire image. The same material can, in some cases, have varying spectral characteristics due to, for instance, aging or atmospheric effects (Rashed et al., 2003). Song (2005) stated that a spectrum that adequately models other spectra in the library does not necessarily model its class of material in the image adequately. The wide field of view (FOV) of the most airborne hyperspectral imaging sensors leads to brightness artifacts caused by anisotropic, bidirectional reflectance that is exacerbated by large viewing angles. This effect is particularly prevalent when the flight direction is perpendicular to the sun-target-sensor-plane resulting in an across-track brightness gradient (Schiefer et al., 2006). In addition, shade fractions are highly sensitive to topographic effects and solar zenith angle (Rashed et al., 2003). In particular in urban areas, shade fractions show considerable variability due to the presence of sub-pixel shadows and numerous steeply inclined surfaces. To overcome this problem, Powell et al. (2007) shade normalized each image pixel. In contrast to previous studies, brightness-effects do not appear to have had a major impact in this study, since MESMA incorporates shade as an endmember while allowing the model to select a different

endmember if needed based on spectral shape. The presented hierarchical MESMA approach improved classification accuracies, since spatial constraints were used for MESMA that minimized spectral confusion. By including a spatial component in MESMA, accuracies of 20 distinct urban material/vegetation species classes were achieved at level 4 that are comparably high (e.g., Herold et al., 2003a). The novel approach of applying simple two-endmember models and – depending on the requirements – additionally three- or four-endmember models to produce high classification accuracies at low complexity levels, allows using these results to improve accuracies for higher complexity levels. Hence, this hierarchical MESMA approach accounts for the spectral pixel-to-pixel variability of urban areas by the simultaneous control of the spatial dimension. Contemplating the discrete classes that can be derived from MESMA as done in the present study, the importance of the per-pixel-fractions of the endmembers as shown in Fig. 8 has to be recalled. In contrast to classification algorithms, MESMA provides physical measurement of material or vegetation abundance that is useful information for urban studies. Further studies are nevertheless desirable, to evaluate the performance of other classifiers and compare them to classification results derived from hierarchical MESMA. In addition, future studies might focus on improving model choice in MESMA depending on complexity levels as well as testing our hierarchical MESMA in different urban areas.

4. Conclusions

In the present study, a hierarchical MESMA was applied to airborne HyMap data by the use of spatial constraints, in order to analyze the urban environment of Bonn in Germany. The high potential of MESMA for mapping urban land cover at various levels of detail was demonstrated. The hierarchical structure of MESMA, whereby results with appropriate classification accuracies from one level were used as spatial constraints for the next level, further improved accuracy. The control of the spatial dimension for different land cover classes in MESMA at higher levels of complexity allowed for a more focused and spatially adjusted spectral unmixing procedure and resulted in a minimization of spectral confusion. The most crucial factor in any SMA approach is the selection of representative endmembers. With the applied endmember selection approaches EAR, MASA and CoB, various quality parameters were available to filter candidate endmembers from a spectral library that contained 1521 potential endmembers. This selection procedure proved to be highly suitable for most endmember classes. Only optimal tree endmembers were difficult to derive from image spectra, due to a high variability of the background surfaces in urban environments, ranging from vegetated-over bare soil to impervious surfaces, which caused a high variability of mixed fractions. For an improved MESMA – or SMA in general – of vegetated areas in urban environments, superior endmembers maybe required derived from other sources. Another spectral confusion was evident between low reflectance materials such as slate, asphalt, cobblestone and dark-shingle. This did not result from the lack of optimal endmembers, but rather from a lack of diagnostic absorption features. However, uncertainties in endmember fraction estimates that are usually caused by brightness differences due to the wide FOV of airborne sensors or particularly due to the complex structure of urban objects were minimized by MESMA. Shade fractions are always included in every MESMA model. Classification results derived from endmember fractions of hierarchical levels 1 to 4 with accuracies of 97.2%, 97.2%, 81.7% and 75.9%, respectively, demonstrated that the flexible MESMA approach combined with a spatially hierarchical structure is particularly suitable to analyze complex urban environments. Different levels of spatial information about urban infrastructure and environment could be derived, in 20 land cover classes at the highest level, which can be valuable for urban management.

Acknowledgements

Data acquisition with the Hyperspectral Mapper (HyMap) was carried out by the German Aerospace Center (DLR). Special thanks go to Matthias Braun at the Center for Remote Sensing of Land Surfaces (ZFL), University of Bonn, Germany, for the coordination of the flight campaign.

References

Adams, J. B., Smith, M. O., & Johnson, P. E. (1986). Spectral mixture modelling: A new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research*, *91*(B8), 8090—8112.

Ben-Dor, E., Levin, N., & Saaroni, H. (2001). A spectral based recognition of the urban environment using the visible and near-infrared spectral region (0.4–1.1 m). A case study over Tel-Aviv. *International Journal of Remote Sensing*, *22*, 2193—2218.

Benediktsson, J. A., Palmason, J. A., & Sveinsson, J. R. (2005). Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(3), 480—491.

Boardman, J. W., Kruse, F. A., & Green, R. O. (1995). *Mapping target signatures via partial unmixing of AVIRIS dataSummaries of the 5th JPL Airborne Earth Science Workshop*, *Vol. 95(1)* (pp. 23-26). Pasadena, CA: Jet Propulsion Laboratory.

Cocks, T., Jenssen, R., Steward, A., Wilson, I., & Shields, T. (1998). The HyMap™ airborne hyperspectral sensor: The system, calibration and performance. *Proceedings of the 1st EARSEL Workshop on Imaging Spectroscopy*. Switzerland: Zurich.

Dennison, P. E., Halligan, K., & Roberts, D. A. (2004). A comparison of error metrics and constraints for Multiple Endmember Spectral Mixture Analysis and Spectral Angle Mapper. *Remote Sensing of Environment*, *93*, 359—367.

Dennison, P. E., & Roberts, D. A. (2003). Endmember selection for Multiple Endmember Spectral Mixture Analysis using Endmember Average RSME. *Remote Sensing of Environment*, *87*, 123—135.

Hepner, G. F., & Chen, J. (2001). Investigation of imaging spectroscopy for discriminating urban land covers and surface materials. *AVIRIS Earth Science and Applications Workshop, Palo Alto, CA*.

Hepner, G. F., Houshmand, B., Kulikov, I., & Bryant, N. (1998). Investigation of the integration of AVIRIS and IFSAR for urban analysis. *Photogrammetric Engineering and Remote Sensing*, *64*(8), 813—820.

Herold, M., Gardner, M., & Roberts, D. (2003). Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(9), 1907—1919.

Herold, M., Goldstein, N. C., & Clarke, K. C. (2003). The spatio-temporal form of urban growth: Measurement, analysis and modeling. *Remote Sensing of Environment*, *86*, 286—302.

Herold, M., & Roberts, D. A. (2005). Spectral characteristics of asphalt road aging and deterioration: Implications for remote-sensing applications. *Applied Optics*, *44*(20), 4327—4334.

Herold, M., & Roberts, D. A. (2006). Multispectral Satellites – Imaging Spectrometry – LIDAR: Spatial–spectral tradeoffs in urban mapping. *International Journal of Geoinformatics*, *2*(1), 1—14.

Herold, M., Roberts, D. A., Gardner, M. E., & Dennison, P. E. (2004). Spectrometry for urban area remote sensing—Development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, *91*, 304—319.

Jensen, J. R., & Cowen, D. C. (1999). Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing*, *65*(5), 611—622.

Painter, T. H., Roberts, D. A., Green, R. O., & Dozier, J. (1998). The effect of grain size on spectral mixture analysis of snow-covered area from AVIRIS data. *Remote Sensing of Environment*, *65*, 320—332.

Powell, R. L., & Roberts, D. A. (2008). Characterizing variability of the urban physical environment for a suite of cities in Rondônia, Brazil. *Earth Interactions*, *12*, 1—32.

Powell, R. L., Roberts, D. A., Dennison, P. E., & Hess, L. L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment*, *106*, 253—267.

Rashed, T., Weeks, J. R., Roberts, D. A., Rogan, J., & Powell, R. L. (2003). Measuring the physical composition of urban morphology using multiple endmember spectral mixture models. *Photogrammetric Engineering and Remote Sensing*, *69*(9), 1011—1020.

Richter, R., & Schlaepfer, D. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction. *International Journal of Remote Sensing*, *23*, 2631—2649.

Ridd, M. K. (1995). Exploring a V–I–S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: Comparative anatomy for cities. *International Journal of Remote Sensing*, *16*, 2165—2185.

Roberts, D. A., Dennison, P. E., Gardner, M., Hetzel, Y., Ustin, S. L., & Lee, C. (2003). Evaluation of the potential of Hyperion for fire danger assessment by comparison to the Airborne Visible/Infrared Imaging Spectrometer. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(6), 1297—1310.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains using Multiple Endmember Spectral Mixture Models. *Remote Sensing of Environment*, *65*, 267—279.

Roessner, S., Segl, K., Heiden, U., & Kaufmann, H. (2001). Automated differentiation of urban surfaces based on airborne hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(7), 1525—1532.

Sabol, D. E., Adams, J. B., Jr., & Smith, M. O. (1992). Quantitative subpixel spectral detection of targets in multispectral images. *Journal of Geophysical Research*, *97*(E2), 2659—2672.

Schiefer, S., Hostert, P., & Damm, A. (2006). Correcting brightness gradients in hyperspectral data from urban areas. *Remote Sensing of Environment*, *101*, 25−37.
Small, C. (2001). Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing*, *22*, 1305−1334.
Small, C. (2003). High resolution spectral mixture analysis of urban reflectance. *Remote Sensing of Environment*, *88*, 170−186.
Small, C. (2005). A global analysis of urban reflectance. *International Journal of Remote Sensing*, *26*, 661−681.
Song, C. H. (2005). Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? *Remote Sensing of Environment*, *95*, 248−263.
Sonnentag, O., Chen, J. M., Roberts, D. A., Talbot, J., Halligan, K. Q., & Govind, A. (2007). Mapping tree and shrub leaf area indices in an ombrotrophic peatland through multiple endmember spectral unmixing. *Remote Sensing of Environment*, *109*, 342−360.
Tompkins, S., Mustard, J. F., Pieters, C. M., & Forsyth, D. W. (1997). Optimization of endmembers for spectral mixture analysis. *Remote Sensing of Environment*, *59*, 472−489.