Remote Sensing of Environment 137 (2013) 184–197

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Support vector regression and synthetically mixed training data for quantifying urban land cover

Akpona Okujeni [a,*], Sebastian van der Linden [a], Laurent Tits [b], Ben Somers [c], Patrick Hostert [a]

[a] Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
[b] Department of Biosystems, M3-BIORES, Katholieke Universiteit Leuven, W. de Croylaan 34, BE-3001 Leuven, Belgium
[c] Department Earth and Environmental Sciences, Division Forest, Nature and Landscape Research, Katholieke Universiteit Leuven, Celestijnenlaan 200E - bus 2411, B-3001 Leuven, Belgium

ARTICLE INFO

Article history:
Received 16 October 2012
Received in revised form 24 April 2013
Accepted 9 June 2013
Available online 19 July 2013

Keywords:
Support vector regression
Support vector machines
MESMA
Unmixing
Hyperspectral
Imaging spectroscopy
Urban
Berlin

ABSTRACT

Exploiting imaging spectrometer data with machine learning algorithms has been demonstrated to be an excellent choice for mapping ecologically meaningful land cover categories in spectrally complex urban environments. However, the potential of kernel-based regression techniques for quantitatively analyzing urban composition has not yet been fully explored. To a great extent, this can be explained by difficulties in deriving quantitative training information that reliably represents pairs of spectral signatures with associated land cover fractions needed for empirical modeling. In this paper we present an approach to circumvent this limitation by combining support vector regression (SVR) with synthetically mixed training data to map sub-pixel fractions of single urban land cover categories of interest. This approach was tested on Hyperspectral Mapper (HyMap) data acquired over Berlin, Germany. Fraction estimates were validated with extensive manual mappings and compared to fractions derived from multiple endmember spectral mixture analysis (MESMA). Our regression results demonstrate that the sets of multiple mixtures yielded high accuracies for quantitative estimates for four spectrally complex urban land cover types, i.e., fractions of impervious rooftops and pavements, as well as grass- and tree-covered areas. Despite the extrapolation uncertainty of SVR, which resulted in fraction values below 0% and above 100%, physically meaningful model outputs were reported for a clear majority of pixels, and visual inspection underpinned the quality of produced fraction maps. Statistical accuracy assessment with detailed reference information for 92 urban blocks showed linear relations with R^2 values of 0.86, 0.58, 0.81 and 0.85 for the four categories, respectively. Mean absolute errors (MAE) ranged from 6.4 to 12.8% and block-wise sums of the four individually modeled category fractions were always around 100%. Results of MESMA followed similar trends, but with slightly lower accuracies. Our findings demonstrate that the combination of SVR and synthetically mixed training data enable the use of empirical regression for sub-pixel mapping. Thus, the strengths of kernel-based approaches for quantifying urban land cover from imaging spectrometer data can be well utilized. Remaining uncertainties and limitations were related to the known phenomena of spectral similarity or ambiguity of urban materials, the spectral deficiencies in shaded areas, or the dependency on comprehensive and representative spectral libraries. Therefore, the suggested workflow constitutes a new flexible and extendable universal modeling approach to map land cover fractions.

1. Introduction

Research on the structure and functioning of cities is of great importance, particularly considering the rates of current and projected global urbanization (Alberti, 2005; Grimm et al., 2008; Pickett et al., 2011). Detailed descriptions of urban surface properties, i.e., beyond the broad differentiation of "urban land" and "non-urban environments", are crucial for developing a more thorough understanding of the integrated human–natural ecosystem (Cadenasso, Pickett, & Schwarz, 2007). The spatial distribution, abundances and conditions of different impervious, pervious and vegetation cover types, for example, directly influence processes related to urban climate (Chudnovsky, Ben-Dor, & Saaroni, 2004; Gluch, Quattrochi, & Luvall, 2006), hydrology (Arnold & Gibbons, 1996; Pauleit & Duhme, 2000) or biodiversity (Blair, 1996; McKinney, 2002).

Data from municipal authorities provide detailed information on urban surface characteristics at the scale of urban blocks, a preferred spatial unit for urban management purposes or urban environmental studies (Heiden et al., 2012; Pauleit & Duhme, 2000). Data collections are, however, often based on cost- and time-intensive field surveys and are subject to irregular updates. In this context, remote sensing constitutes a supplemental technique to derive information on the physical composition of urban areas (Jensen & Cowen, 1999; Maktav, Erbek, & Jurgens, 2005; Miller & Small, 2003).

* Corresponding author. Tel.: +49 30 2093 6905; fax: +49 30 2093 6848.
E-mail address: akpona.okujeni@geo.hu-berlin.de (A. Okujeni).

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.06.007

Data acquired by various multispectral spaceborne sensors, which range from very high to moderate spatial resolution and high acquisition frequency, have been applied for mapping urban land cover (Myint, Gober, Brazel, Grossman-Clarke, & Weng, 2011; Powell, Roberts, Dennison, & Hess, 2007; Rashed, Weeks, Roberts, Rogan, & Powell, 2003; Thomas, Hendrix, & Congalton, 2003). Many studies followed the V–I–S model by Ridd (1995), a framework used to universally characterize three major urban surface components, i.e., "vegetation", "impervious surfaces" and "soil". However, the coarse spectral resolution of multispectral data often limits the accurate distinction of different impervious, pervious and vegetation surface cover types (Herold, Gardner, & Roberts, 2003; Myint et al., 2011; Small & Lu, 2006). Therefore, Small (2004) and Small and Lu (2006) proposed the substrate vegetation dark surface (SVD) model, which characterizes the urban environment by biophysical surface reflectance properties rather than by land cover types. Nevertheless, linking spectral with thematic information remains desirable, especially when high resolution spectral information is available. Other studies overcame some of the limitations of multispectral data by analyzing high-resolution images with object-based strategies (Myint et al., 2011; Shackelford & Davis, 2003; Thomas et al., 2003). Still others use complex multi-sensor and multi-temporal approaches that extend the spectral feature space or take advantage of information beyond the spectral domain (Griffiths, Hostert, Gruebner, & van der Linden, 2010; Taubenböck et al., 2012). Even when employing such fusion approaches we need to better understand the opportunities and limitations of high resolution spectral data.

The increasing availability of imaging spectrometer data from urban areas has allowed for a more detailed mapping of thematic surface properties on a purely spectral basis. Several studies have demonstrated the great potential of continuous spectral information for fine-scale classifications, including various man-made materials and vegetation species (Franke, Roberts, Halligan, & Menz, 2009; Herold et al., 2003). Other studies have included rather broad-scale classifications when considering ecologically meaningful sub-categories, e.g., impervious rooftops and pavements, pervious soils, and grass- and tree-covered areas (Roessner, Segl, Heiden, & Kaufmann, 2001; van der Linden & Hostert, 2009). Thus far such investigations have been more experimental in nature because currently available high quality data sets are almost exclusively constrained to airborne acquisitions with limited spatial coverage and temporal frequency. This is likely to change with the advent of new hyperspectral satellite missions, e.g., the German Environmental Mapping and Analysis Program (EnMAP; Stuffler et al., 2007; Stuffler et al., 2009) or the Hyperspectral Infrared Imager (HysplRI; National Research Council, 2007; Roberts, Quattrochi, Hulley, Hook, & Green, 2012). Such systems will provide imaging spectrometer data of a large spatial extent on a timely and frequent basis, which will open up new opportunities for urban remote sensing applications (Heldens, Heiden, Esch, Stein, & Mueller, 2011). Ground sampling distances (GSDs) in the range of 30 to 60 m will, however, complicate studies, especially in spatially heterogeneous environments. Robust analysis approaches with universal applicability and best possible transferability on data of coarser spatial resolution will be needed to best utilize imagery acquired by future spaceborne imaging spectrometers.

Despite this promising development, using imaging spectrometer data for urban applications remains a challenge. The high number of materials, their varying conditions, as well as the surfaces' anisotropic reflectance behavior lead to vast spectral diversity. As a consequence, different spectral phenomena such as high within-class variability of individual materials, complex multi-modal class compositions, and high inter-class similarity or ambiguity complicate urban land cover assessments (Heiden, Segl, Roessner, & Kaufmann, 2007; Herold, Roberts, Gardner, & Dennison, 2004; Herold, Schiefer, Hostert, & Roberts, 2006; Schiefer, Hostert, & Damm, 2006). This complexity is further aggravated by the high amount of spectrally mixed pixels typical for urban remote sensing data. The extent of spectral mixing, i.e., the number of mixed pixels and the abundance of different materials contributing to the mixed signal, strongly depends on the fine-scale spatial patterns of different urban objects, and on the GSD of the respective sensor (Small & Lu, 2006).

Spectrally mixed signatures produce inaccuracies in per-pixel classification approaches, which assign each pixel to a discrete urban land cover category (Powell et al., 2007; Small, 2001). This applies even to investigations on airborne or high resolution spaceborne remote sensing data of 4 m GSD (Small, 2003; van der Linden, Janz, Waske, Eiden, & Hostert, 2007). Hence, quantitative mapping of urban land cover components constitutes an alternative concept that becomes particularly relevant once coarser resolution spaceborne imaging spectrometer data becomes available. In general, quantitative approaches account for sub-pixel mixing by transforming reflectance measurements of individual pixels into both physically meaningful quantities of surface fractions and thematically meaningful land cover types. Multiple endmember spectral mixture analysis (MESMA; Roberts et al., 1998) is probably the most commonly used technique to systematically decompose mixed pixels into fractional abundances of distinct spectrally pure endmembers (EMs). Within-class spectral variability is accounted for by using an extensive iterative procedure with multiple linear mixture models (Somers, Asner, Tits, & Coppin, 2011). In the context of imaging spectrometry of urban areas, MESMA was successfully used to quantify thematically detailed land cover types or to derive material-oriented fraction maps (Franke et al., 2009; Roberts et al., 2012).

Support vector machines (SVMs) have received increasing attention in the remote sensing community; SVMs are supervised, non-parametric statistical learning techniques designed to solve classification and regression problems (Schölkopf & Smola, 2002; Vapnik, 1995). A recent review of implementations and applications of SVMs was provided by Mountrakis, Im, and Ogole (2011). As applied to remote sensing data, SVMs have shown great capabilities to model complex multi-modal, nonlinear data distributions in high-dimensional spectral feature spaces (Huang, Davis, & Townshend, 2002; Melgani & Bruzzone, 2004; Pal & Mather, 2006). Thus far, the advantage of kernel-based support vector classification (SVC) for a per-pixel-based mapping of spectrally complex urban land cover categories from imaging spectrometer data has been demonstrated (Tuia & Camps-Valls, 2011; van der Linden et al., 2007; Waske, van der Linden, Benediktsson, Rabe, & Hostert, 2010). However, the potential of kernel-based support vector regression (SVR) as a quantitative technique for analyzing the complex hyperdimensional urban feature space with high spectral and spatial heterogeneity has not yet been fully explored. Given a set of adequate training samples, SVR allows us to estimate a continuous output variable, such as sub-pixel fractions of a single land cover category. To date, most studies have used SVR for predicting biophysical/chemical plant parameters by relating spectral information to in situ samples (Camps-Valls, Bruzzone, Rojo-Alvarez, & Melgani, 2006; Tuia, Verrelst, Alonso, Perez-Cruz, & Camps-Valls, 2011; Verrelst et al., 2012). Few studies have adopted SVR to quantify fractions of broad urban categories in 30 m multispectral satellite imagery using reference information from very high resolution land cover maps (Esch et al., 2009; Walton, 2008). This latter approach fully relies on the availability of very accurately co-registered, very high resolution reference. Often such data – and especially accuracy – is not given when high resolution imagery from spatially heterogeneous environments is used as input, e.g., for airborne line scanner data from urban areas. Moreover, reliable sub-pixel fractions cannot be labeled in the data itself or mapped in the field. These difficulties in finding reliable quantitative training information explain the lack of studies that use regression techniques for sub-pixel mapping. To the best of our knowledge, no investigations have been carried out where SVR was used together with imaging spectrometer data to derive sub-pixel fraction maps of spectrally complex urban land cover categories.

In this paper, we propose the combination of support vector regression with synthetically mixed training data to quantify urban land cover. The idea behind our approach was inspired by Foody and Mathur (2006), who suggested generating class-dominated spectral mixtures using spectral averaging to improve SVC training. Based on imaging spectrometer data and a corresponding spectral library, we consecutively derived fraction maps with SVR for single urban land cover categories of interest, so-called target categories. For each target category we generated a synthetically mixed training data set from the spectral library, which was subsequently used for SVR model training. This training set consists of pure original spectra and multiple-binary mixed spectra representing a range of mixing fractions between 0 and 100% of the target category against all remaining background cover types. To generate a universal model that accounts for spectral variability, the spectral variations within each land cover category, e.g., varying illumination and shade, or differences in material composition and condition, were included and learned from appropriate training. We thus fulfilled the need for quantitative training information required for empirical modeling, and we overcame the mentioned difficulties of using regression techniques for sub-pixel mapping. Subsequently, each SVR model was applied to the image data to derive the desired fraction map per target category.

The objective of this study was to quantify spectrally complex and ecologically meaningful urban land cover categories using a combination of SVR and synthetically mixed training data. We thus explore whether (machine learning-based) regression techniques can be used for sub-pixel mapping with training data derived from commonly available sources such as spectral libraries. We used airborne Hyperspectral Mapper (HyMap) data acquired over Berlin, Germany, to consecutively estimate fractions of two different impervious categories, i.e., rooftops (*roof*) and paved surfaces (*pavement*), including roads, sidewalks and parking lots. We further differentiated two vegetation categories, i.e., *grass* and *tree*. Our objective was tested by comparing fraction estimates at pixel and urban block scale to maps from very high resolution reference data. To explore the quality of results compared to established approaches we evaluated our maps to results derived with MESMA.

2. Data

2.1. Imaging spectrometer data and pre-processing

Image data was acquired by the HyMap sensor over Berlin on 20th August 2009 at approximately 11:45 local time. HyMap covers the spectral region between 440 and 2500 nm, with 128 bands and bandwidths between 10 and 20 nm. The sensor's field of view (FOV) is 60°, and the instantaneous field of view (IFOV) is 2.5 mr along track and 2.0 mr across track (Cocks, Jenssen, Stewart, Wilson, & Shields, 1998). The aircraft was flown from south to north at an altitude of 2005 m (AGL), resulting in raw pixels of 5 × 4 m at nadir. Given this setup, an area of 2.3 by 22.8 km was covered, representing a subset of Berlin's urban–rural gradient (Fig. 1). A great variety of urban structure types were captured, including a commercial center, residential areas of varying densities, green spaces, water bodies, sport grounds, industrial areas, as well as agricultural areas and forest patches.

Pre-processing of the HyMap image encompassed system correction (Cocks et al., 1998) and atmospheric correction (Richter & Schläpfer, 2002) by HyVista Corporation and the German Aerospace Center (DLR). Data were delivered in reflectance values with a reduced number of 126 bands. We performed a parametric geocoding (Schläpfer & Richter, 2002) using a 3.6 m resolution digital surface model (DSM; acquired in January 2008 and resampled from 0.5 m). This cell size also constitutes the final spatial resolution of the orthorectified HyMap image. The slightly higher output resolution than that of the raw image was chosen to better preserve a higher portion of the spectral information during geometric processing (Schläpfer & Richter, 2002). Based on 15 independent control points from digital orthophotos (DOPs; acquired in April 2009 with 0.1 m resolution), we calculated a spatial accuracy (RMSE) of 2.91 and 3.08 m in easting and northing.

2.2. Image spectral library

A comprehensive spectral library was developed from the HyMap image. We included major characteristic man-made and natural surface materials relevant for the study area. The library was structured according to a slightly modified hierarchical categorization scheme for urban areas as proposed by Roessner et al. (2001) and Heiden et al. (2007). Levels I and II represent rather broad land cover categories such as impervious, pervious or vegetation. *Roof*, *pavement*, *grass* and *tree* correspond to Level III and represent the more thematically detailed urban land cover categories of interest in this work. Level IV forms the spectral basis for the analysis (Table 1).

The initial collection of around 300 pure material spectra (mean values of manually delineated spectrally homogeneous regions of interest in the HyMap image) was based on expert knowledge and auxiliary information, including field mapping data, DOPs and Google Street View. We considered the spectral variability of materials, e.g., variations in brightness caused by varying illumination and shading effects, or variations due to differences in material condition and composition. In a second step, we iteratively developed the final image spectral library used for processing. To create a subset of spectra from the initial collection that best describes the spectral diversity of the study region, we manually selected a representative standard spectrum (spectrum close to the mean within the collection of multiple material spectra) for materials with frequent spatial occurrence. We adopted this subset of spectra to derive fraction maps using SVR and synthetically mixed training data. After visual inspection and statistical accuracy assessment of results, we added spectra to account for missing materials. Spectra were also added to account for illumination and shading effects or variations in condition and composition. The final image spectral library consists of 41 material spectra, of which 17 were assigned to the *roof* category, 6 to the *pavement* category, 5 to the *grass* category and 7 to the *tree* category. Six spectra were assigned to a so-called *other* category. This additional category, with mainly non-vegetated pervious surfaces of marginal spatial occurrence, did not constitute a category of interest for this analysis.

2.3. Reference data

Most analyses of urban environmental indicators by municipal authorities in Berlin are performed on a block scale, and related maps are presented in the Urban and Environmental Information System (UEIS; SenStadt, 2013a). These urban blocks, therefore, formed the spatial unit to statistically validate the results in three regions of various urban structure types (Fig. 1). Validation area 1 represents a high-density commercial and housing zone (high-density urban area), area 2 represents a medium-density residential and industrial zone (medium-density urban area), and area 3 represents a low-density residential zone with mainly detached and semi-detached housing patterns and private gardens (low-density urban area). The effect of urban surface geometry, e.g., building displacement at large view-angles, constitutes a significant drawback for any wide-swath sensor over urban areas. When compared to reference data (e.g., cadastral information), the three-dimensional geometry produces inaccuracies related not to the spectral information content, but rather to unavoidable geometric distortions. We therefore only selected blocks within a ±10° off-nadir region to minimize effects of the urban 3D-geometry on comparisons between HyMap and reference data. By limiting our evaluation to that area of the scene where almost entirely spectral effects (and not geometric) influence our accuracy measures, we were able to determine

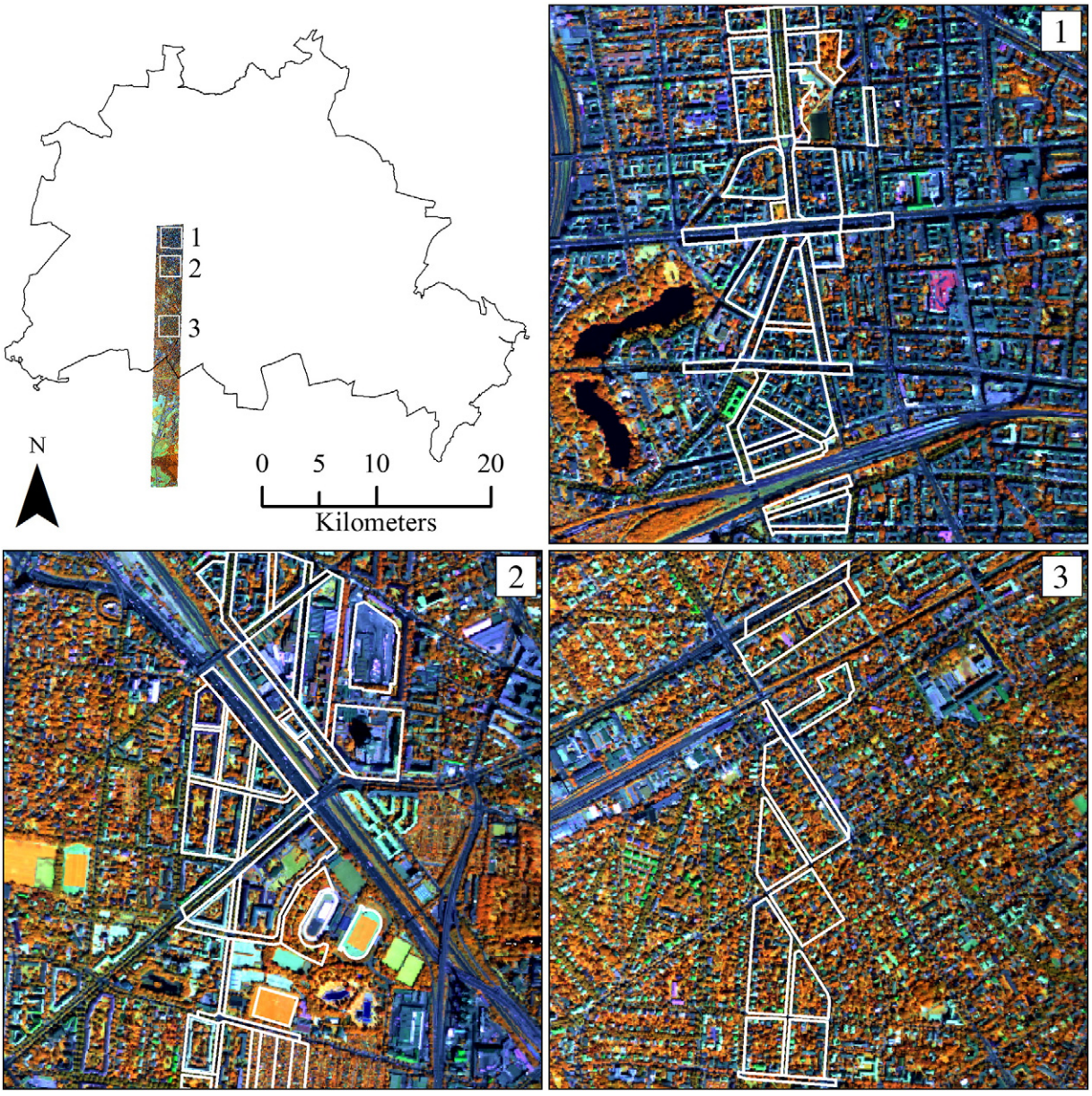

Fig. 1. HyMap footprint covering a subset of Berlin's urban gradient (R = 833 nm, G = 1652 nm, B = 632 nm) and validation areas 1–3 (white polygons indicate urban blocks used for validation).

the spectral effectiveness of the SVR approach. The effects of urban 3D-geometry on wide-swath sensor data and the related performance limits of any algorithm in urban environments is beyond the scope of this work and exhaustively described in van der Linden and Hostert (2009). Reference fractions of Level III land cover categories of interest (see Table 1) were calculated for each urban block. The spatial extent of buildings was extracted from Berlin's digital cadastre (SenStadt, 2013b). Vegetation types and paved impervious grounds were manually delineated at 1:500 map scale based on the available DSM and DOPs. Whereas the building data set was regarded as spatiotemporal-invariant, the temporal shift of the DSM and DOPs to the Hymap image may have influenced validation to some extent. The final reference data set encompassed a total number of 92 urban blocks with areas between 2530 and 65,950 m^2, of which there were 45 building blocks of different urban structure type, 37 street polygons including wide main streets and small side roads, and 10 green spaces consisting of parks, allotment gardens or sport grounds.

3. Methods and data analysis

The complete methodological framework of this study is illustrated in Fig. 2. Imaging spectrometer data, a corresponding spectral library and independent reference information formed the data base for processing. Initially, library spectra were used to generate synthetic mixtures, which served, along with their mixing fractions, as input for SVR model training. Cover fraction estimates were then derived by applying the SVR model to the image data. The entire procedure was carried out separately for each of the four land cover categories of interest. Finally, accuracy was assessed with an independent reference data set.

3.1. Generation of synthetically mixed training data

The generation of synthetically mixed training data from a spectral library is illustrated in Fig. 3. For each Level III urban land cover

Table 1
Hierarchically structured image spectral library.

Level III	Level IV	No. spectra
Roof	Red clay tile	4 (a)
	Red cement tile	3 (a)
	Bitumen	5 (b)
	Brown roof tile	1
	Brown roof shingle	1
	White roof material (polyethylene)	1
	White roof material (unknown)	1
	Zinc roof material	1
Pavement	Asphalt	4 (a, b)
	Concrete	2 (b)
Grass	Grass (intensively manicured)	2 (a)
	Grass (extensively manicured)	1
	Grass (dry)	2 (a)
Tree	Deciduous tree	7 (a)
Other	Tartan (sports ground)	1
	Railtrack (concrete sleepers)	1
	Railtrack (wooden sleepers)	1
	Sand (playground)	1
	Soil	1
	Water	1

Multiple spectra per material considered to account for spectral variability due to variations in (a) illumination and shading effects, and (b) material condition or composition.

category of interest, a quantitative training set was generated as follows:

Step 1: The image spectral library was split into two sections: (i) a target category and (ii) a background category, which included all remaining categories.

Step 2: Synthetically mixed spectra between each pure spectrum of the target category (i.e., 100% mixing fraction), and each pure spectrum of the background category (i.e., 0% mixing fraction of the current target category) were calculated. In line with many unmixing approaches (e.g., Roberts et al., 1998; Tompkins, Mustard, Pieters, & Forsyth, 1997), we hereby assumed linear mixing systematics for simplification. While constructing the synthetically mixed training data, decisions on the mixing complexity and the mixing interval had to be made.

- The mixing complexity refers to the number of materials from the target and background category that contribute to a mixed pixel. In general this is not known in advance, and thus requires a flexible strategy to account for both simple and more complex mixtures. Binary mixtures between each pure spectrum of the target category and each pure spectrum of the background category constitute the simplest complexity. This complexity can be arbitrarily increased by allowing intra- and inter-categorical mixtures with more than two materials, e.g., three or four to calculate ternary or quaternary mixtures. In this study, we approximated the mixing complexity inherent in the HyMap image by overlaying a 3.6 m grid on the reference data set. Our calculations indicated that less than 10% of the pixels are composed of more than two land cover types. Although this analysis could not be carried out on a material level, we used this approximation to investigate the use of binary mixtures for deriving reliable fraction estimates from the 3.6 m HyMap image.
- The mixing interval refers to the number of intermediate mixtures within the given fraction range between 0 and 100% when mixing the target category spectra against the background category spectra. This interval can be defined globally or for each land cover type separately, either by a constant step size or by arbitrarily selected percentages. To keep the computing time low, the definition of the mixing interval should further balance the trade-off between the necessary and avoidable number of intermediate mixing steps, while at the same time considering the accuracy of fraction estimates. In this study we trained SVR models with global step sizes of 50%, 25%, 20%, 10% and 5%. Our results showed that accuracies increased with decreasing increments. For step sizes below 20%, no considerable performance improvements were observed. However, a clear increase in processing time was observed. Therefore, we ultimately selected the global step size of 20%.

Step 3: Finally, all pure original and binary-mixed spectra were combined in a single spectral library. This multiple set of spectra, together with the associated mixing fractions between 0 and 100% of the target category, constituted the training data for SVR modeling.

3.2. Support vector regression

The main idea behind SVR is to estimate the linear dependency between pairs of n-dimensional input vectors and 1-dimensional target variables by fitting an optimal approximating hyperplane to a set of training samples. The hyperplane is defined by a linear function that is found based on structural risk minimization. That is, the linear model is optimal when it minimizes a cost function considering a

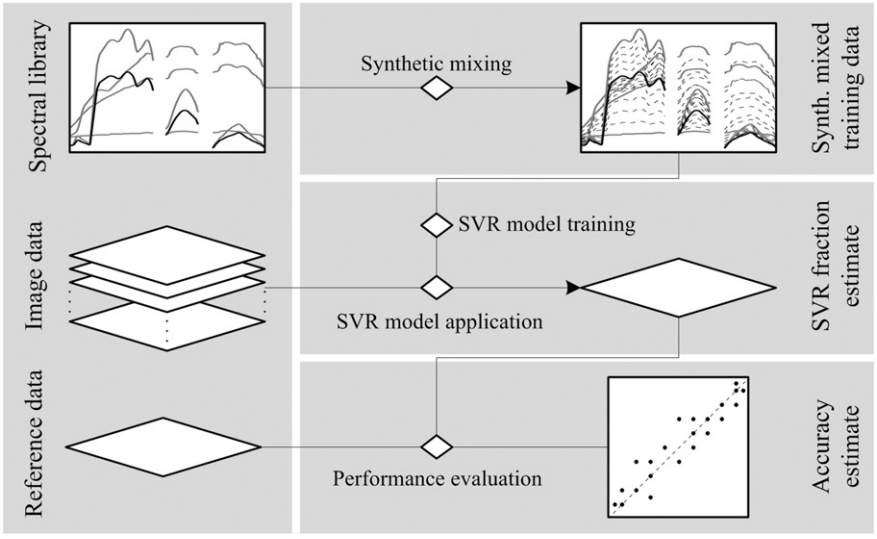

Fig. 2. Workflow for quantifying urban land cover with SVR and synthetically mixed training data.

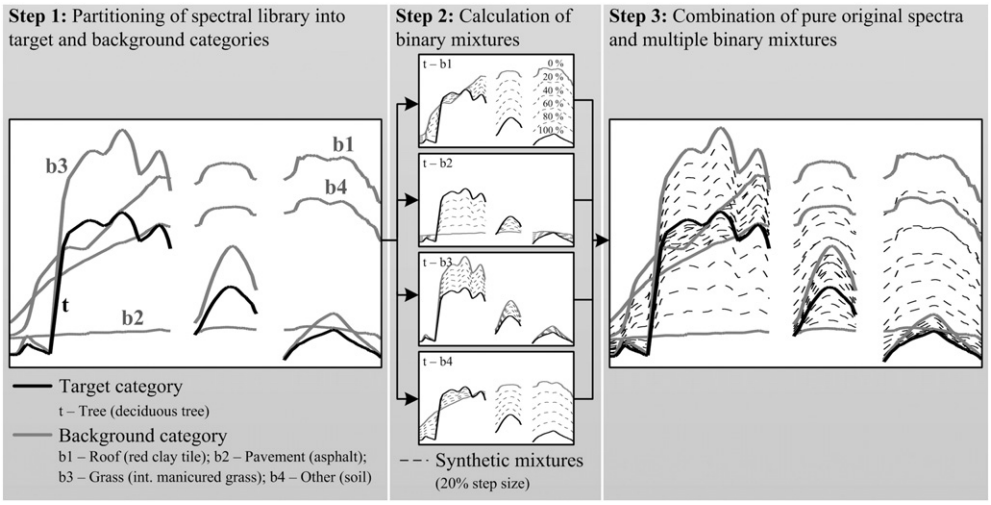

Fig. 3. Generation of synthetically mixed training data.

maximized margin that encloses the samples, and a minimized error of approximation. Approximation is controlled by an ε-insensitive loss function that directly influences the width of the margin. Samples within the margin are not penalized. The error of approximation, i.e., samples located outside of the margin, is measured using slack variables. The influence of samples outside the margin is controlled by a regularization parameter C. Integrated into a kernel framework, SVR is capable of capturing nonlinear data distributions. By using a Gaussian kernel function, the data is implicitly mapped into a higher dimensional feature space, wherein the new data distribution enables a better fitting of a linear function. Hence, kernel-based SVR turns out to be a nonlinear regression in the original input space (Brereton & Lloyd, 2010; Smola & Schölkopf, 2004; Vapnik, 1995).

The SVR analysis was carried out with imageSVM (Rabe, van der Linden, & Hostert, 2010), which is an open source tool for support vector machine applications with remote sensing data. Delivered as part of the EnMAP-Box (Rabe, Jakimow, van der Linden, & Hostert, 2012), imageSVM uses LIBSVM (Chang & Lin, 2011) during the training of support vector models. We followed a two-step training approach: initially, for each target category the synthetically mixed data set was used as input to train an SVR model. Optimal parameter sets for γ, which determines the width of the Gaussian kernel, as well as for C and ε, were determined by a grid search in conjunction with an internal 3-fold cross-validation using the RMSE as a measure of performance. Thus, the effectiveness of SVR to find a single solution, i.e., one global model, for the complex functional relation between the multiple set of spectra and associated mixing fractions was exploited. Subsequently, each SVR model was applied to the HyMap image, which resulted in one fraction image per target category. The model output range covers continuous values between 0 and 100%, i.e., partial interpolations of 20% training intervals, as well as values below 0% or greater than 100%, i.e., extrapolated model outputs.

3.3. Multiple endmember spectral mixture analysis

A more physically-based approach for estimating sub-pixel cover fractions is MESMA (Roberts et al., 1998). MESMA is directly driven by the physically explicit mixture model that describes the mixed pixel signal as a linear combination of the spectral properties of the constituent components (or EMs) weighted by their relative cover in the pixel. Once the EMs and their spectral signatures are known, inversion techniques are used to solve for the fractions with minimal additional error in the model equations, thus obtaining fully constrained cover fractions. As opposed to more traditional unmixing techniques, MESMA allows the EMs to vary on a per-pixel basis. Many mixture models are iteratively calculated for each pixel in an image, thereby evaluating all possible EM combinations from a spectral library and calculating the fit between the measured and modeled mixed signal. The model with the lowest RMSE is assigned to the pixel. Often, the mixing complexity is handled by calculating simple two-EM models and more complex three- and four-EM models per pixel. The final model is selected based on a given criterion, e.g., a model of higher complexity is only selected if the error decrease related to increased complexity exceeds a threshold (Franke et al., 2009; Powell et al., 2007; Roberts et al., 2012).

The MESMA approach aims to achieve a comprehensive estimation of land cover fractions for all urban categories considered in the analysis. The 41 pure library spectra related to Level III of the hierarchical categorization scheme (Table 1) were used as endmembers for processing. As less than 10% of the pixels are composed of more than two land cover types (compare Section 3.1), all possible combinations between two material endmembers were evaluated. Consequently, we do not account for higher order mixing complexities. We do, however, guarantee a fair methodological comparison between both approaches in terms of mixing complexity considered in the analysis. In the initial setup, shadow was not accounted for, as shaded materials were included in the library (Asner & Lobell, 2000). This resulted in a total number of 623 two-EM models (two material EMs). In the second setup, we included a spectrally flat photometric shade spectrum as additional EM, with subsequent shade normalization (Powell et al., 2007; Roberts et al., 2012). This resulted in a total number of 623 three-EM models (two material EMs, one shade EM). For each setup, results of the mixture model with lowest modeling error per pixel were retained, with the corresponding estimated cover fractions for the land cover categories considered in the analysis, i.e., *roof*, *pavement*, *grass*, *tree* and *other*.

3.4. Validation of fraction maps

Pixel-wise accuracy assessment of fraction estimates from empirical modeling is not straightforward. Similar to the dilemma related to acquiring quantitative training data, inaccurate co-registration between mapping results and higher resolution reference information will lead to biased accuracy statistics, especially in such a spatially heterogeneous urban environment. Spatial units larger than a pixel, e.g., a cluster of pixels or polygons, can mitigate this influence (Stehman & Wickham, 2011). This strategy was used, for example, by Powell et al. (2007) and Roberts et al. (2012) to statistically validate MESMA-based urban land cover fraction maps. In addition to a

pixel-wise evaluation, i.e., visual inspection of fraction maps and a consistency check of fraction values due to the extrapolation uncertainty of SVR, we used urban blocks from the UEIS to statistically assess the accuracy of our results.

4. Results

4.1. Performance of SVR at pixel scale

The consistency of pixel values was used to evaluate the reliability of fraction estimates derived by SVR and its ability to generalize. We considered the percentages of physically meaningful (0–100%), negative (<0%) and super-positive (>100%) fractions for all urban blocks and for urban blocks stratified by structure types (Table 2). For all urban blocks, around 60% of *roof* or *pavement* pixels show physically meaningful fractions between 0 and 100% (Table 2a). The percentage of negative fractions for both categories is clearly higher than the percentage of super-positive fractions. With over 80% of pixels, *grass* and *tree* show a higher number of physically meaningful fractions. Stratified results further illustrate that percentages vary between urban structure types (Table 2b). For all land cover categories, the number of physically meaningful fractions increases along with a decreasing density of urban block. At more than 90%, the highest increase can be observed for *grass* and *tree* within urban blocks from low-density urban areas. With approximately 56% of *roof* and *pavement* pixels, street polygons show the lowest numbers of physically meaningful fractions.

An analysis of the spatial distribution of pixels with physically unrealistic fractions values reveals that for all urban land cover types, most of the pixels with super-positive fractions are associated with pure surface coverage of the respective target category (Fig. 4). Vice versa, most of the pixels with negative fraction values are linked to surfaces with target fractions of 0%. Minor deviations from this common trend can be observed for *roof* and *pavement*, e.g., rooftops or parts of a paved bus parking area that are misleadingly mapped with negative fractions.

Table 2
Percentages of physically meaningful, negative and super-positive pixel fractions for a) all urban blocks, and b) urban blocks stratified by structure types.

	Category	0–100%	<0%	>100%
a) All				
	Roof	62.9	27.5	9.6
	Pavement	58.8	24.9	16.3
	Grass	80.6	19.0	0.4
	Tree	81.5	15.2	3.4
b) By urban structure type				
High-density	*Roof*	61.3	21.3	17.4
	Pavement	54.5	30.8	14.8
	Grass	67.2	32.7	0.0
	Tree	73.3	24.6	2.1
Medium-density	*Roof*	64.0	23.3	12.7
	Pavement	59.1	25.7	15.2
	Grass	76.2	23.4	0.3
	Tree	79.8	18.3	1.9
Low-density	*Roof*	66.1	28.1	5.8
	Pavement	64.1	21.8	14.1
	Grass	91.7	8.2	0.2
	Tree	91.5	6.0	2.4
Street	*Roof*	56.5	39.0	4.4
	Pavement	56.6	19.8	23.7
	Grass	85.3	14.7	0.0
	Tree	80.8	11.6	7.6
Green space	*Roof*	75.3	23.7	1.1
	Pavement	63.1	30.8	6.1
	Grass	93.0	3.3	3.7
	Tree	86.7	10.9	2.4

Given this rather meaningful location of outliers, we produced stretched fraction maps with values between 0 and 100%, where negative and super-positive fractions were set to 0 and 100%, respectively. Visual inspection of five selected subsets reveals that spatial patterns of the urban land cover categories in the reference data set are reproduced well by SVR (Fig. 5). For example, *roof* is differentiated from the background cover types with high precision. Minor confusion occurs in the case of paved streets and backyards or shaded areas, i.e., areas covered by cast shadows of buildings and trees or shaded components of trees. In a few cases, particularly for dark roofing materials, *roof* fractions are underrated. The *pavement* estimates are accurate for streets or backyards. Simultaneously, however, a weak quantitative differentiation, especially from dark background cover types is observed. This leads to highly overrated *pavement* fractions for some rooftops and, above all for shaded grass- or tree-covered areas. These effects are particularly strong for highly fragmented low-density urban areas and green spaces, both of which have complex illumination and shadowing effects within vegetation stands. The quantitative differentiation of *grass* and *tree* from background cover types by SVR is highly accurate. Different types of grass-covered surfaces, e.g., lawns consisting of dry or extensively manicured grass, are correctly mapped as having full *grass* coverage of 100%. The precise distinction between *grass* and *tree* can be observed along streets, i.e., where street trees dominate the scenery, and in the example of green space. Confusion between the two vegetation types is relatively low, and mostly appears when tree crowns are highly illuminated.

4.2. Performance of SVR and MESMA at urban block scale

Urban land cover fraction estimates modeled by SVR and MESMA are compared to reference information at the urban block scale. We used the original SVR outputs, i.e., unstretched fraction values, to calculate average block estimates. For a total of 92 urban blocks, *roof*, *grass* and *tree* estimates by SVR show linear relationships close to the 1:1 line, with high accuracies as indicated by MAE values below 9% and R^2 values over 0.8 (Fig. 6). Estimates of *pavement* are less accurate with increased scattering around the 1:1 line, a higher MAE value of 12.8% and a lower R^2 value of 0.58. For most urban blocks, SVR produces physically meaningful fraction values between 0 and 100%. Only five urban blocks show negative values of up to −5%.

For all urban land cover types, accuracy measures of MESMA carried out in the setup with two-EM models (two material EMs) follow similar trends but with slightly reduced accuracies when compared to SVR (Fig. 6). The largest deviations from SVR results are observed for *pavement*, where the MAE increases by 7.2%. In terms of R^2, a quite large decrease of 0.2 is observed for *grass* estimates. As mentioned in Section 3.3, MESMA was additionally carried out in a setup with three-EM models (two material EMs, one shade EM) with subsequent brightness normalization. Results are very similar (scatterplot not shown), with both exhibiting slightly increased and reduced accuracies (*roof*: MAE = 11.7%, R^2 = 0.72; *pavement*: MAE = 20.1%, R^2 = 0.61; *grass*: MAE = 9.5%, R^2 = 0.59; *tree*: MAE = 6.3%, R^2 = 0.75).

The scatterplots (Fig. 6) reveal regions of agreement and disagreement between both approaches. The *roof* estimates of most urban blocks from medium- and high-density urban areas are highly underrated by MESMA. For both methods, street polygons and green spaces are misleadingly mapped as having *roof* proportions of up to 20%. However, *pavement* estimates by SVR and MESMA show opposite tendencies. For building blocks from low-density urban areas and green spaces, *pavement* proportions derived by SVR are highly overestimated. In contrast, MESMA highly underrates intermediate to high *pavement* fractions, especially for street polygons. The *grass* and *tree* estimates of both approaches follow similar patterns. The reduced accuracies of MESMA compared to SVR for both categories can be attributed to small differences. The *grass* fractions, particularly of building blocks from low-density urban areas, are slightly

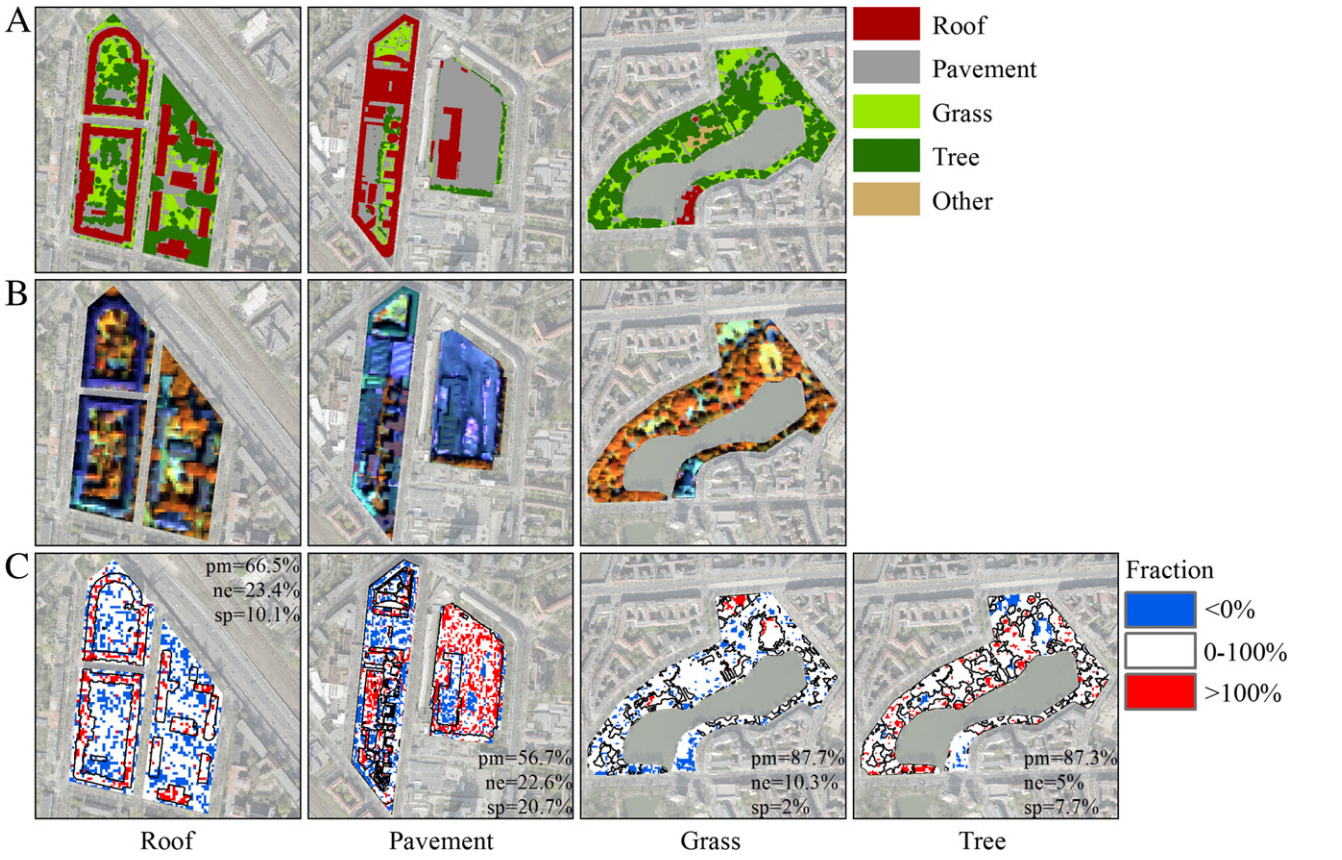

Fig. 4. Subsets showing the spatial distribution of pixels with physically meaningful and unrealistic fractions values for different urban blocks. (A) Reference data, (B) HyMap image (R = 833 nm, G = 1652 nm, B = 632 nm) and (C) aggregated fractions modeled by SVR. The percentage of physically meaningful (pm), negative (ne) and super-positive (sp) fractions is reported in each image chip.

underestimated by SVR, and to a greater extent by MESMA. The effect of misleadingly mapped *grass* fractions within street polygons is higher for MESMA. Further, *tree* fractions within street polygons are underrated by MESMA. For both approaches, the largest underestimates of *grass* and overestimates of *tree* fractions are observed for four green spaces with highly fragmented allotment gardens. Also, in this case deviations from the 1:1 line are larger for MESMA than for SVR.

The reliability of the proposed SVR approach for quantitatively mapping urban land cover at the urban block scale was further evaluated by investigating whether estimated fractions that were consecutively derived by four different SVR models sum to unity. For 45 building blocks, block-wise sums of *roof*, *pavement*, *grass* and *tree* estimates produce reasonable fractions with a mean value of 111.4% and a standard deviation of 10.9% (Fig. 7a). Block-wise sums of street polygons are slightly less accurate with a mean value of 116.7% and a standard deviation of 13.1% (Fig. 7b). The reasons for super-positive sums accord with previously observed results. For building blocks, physically unrealistic fractions clearly above 100% mainly occur when *pavement* is highly overestimated. This effect is particularly strong in low-density urban areas. For street polygons, block-wise sums clearly above 100% are mainly caused by misleadingly estimated *roof* and *grass* fractions, which are either inexistent or negligibly small in the reference data set.

5. Discussion

5.1. Urban land cover fractions estimates by SVR at pixel scale

Accurate percentages of physically meaningful fraction values at the pixel scale highlight the predictive power of the SVR models and prove the capability of the sets of pure and synthetically mixed spectra to account for large parts of the spectral diversity within the HyMap image. The fact that spectrally complex, i.e., multi-modal, urban land cover categories such as *roof* and *pavement*, as well as spectrally similar classes, e.g., *tree* and *grass*, can be quantitatively separated shows the strength of SVR to extract almost all spectral information. Although statistical assessment cannot be performed at the pixel scale, the visual inspection of maps (see Fig. 5) strengthens the proposed SVR approach.

Nevertheless, two major types of errors exist during SVR modeling. First, unconsidered spectra of potential surface materials or material variability can lead to improper extrapolation, i.e., the estimation beyond the known training space. Second, the spectral ambiguity of training data can lead to interpolation uncertainties when fitting the hyperplane. Both errors lead to reduced mapping performances, either by inaccurate land cover fraction estimates in the physically meaningful range between 0 and 100%, or by unrealistic fraction value estimates, i.e., negative or super-positive values (see Table 2).

Accordingly, a high number of physically unrealistic fraction values are found where spectral variability of man-made materials is high, e.g., for the *roof* category or for dense urban areas where complex illumination and shadowing effects amplify spectral variations. Similarly, negative and super-positive *roof* and *pavement* fraction values often occur in areas where spectral ambiguity between the target and background category is high. This is particularly the case for urban blocks that include large paved areas, e.g., parking lots (see Fig. 4) or streets (see Table 2). In contrast, *grass* and *tree* predominantly show physically meaningful fraction values and only very few super-positive fractions. High accuracies for vegetated surfaces underline the approach's capability to quantitatively describe diverse categories with often subtle spectral differences.

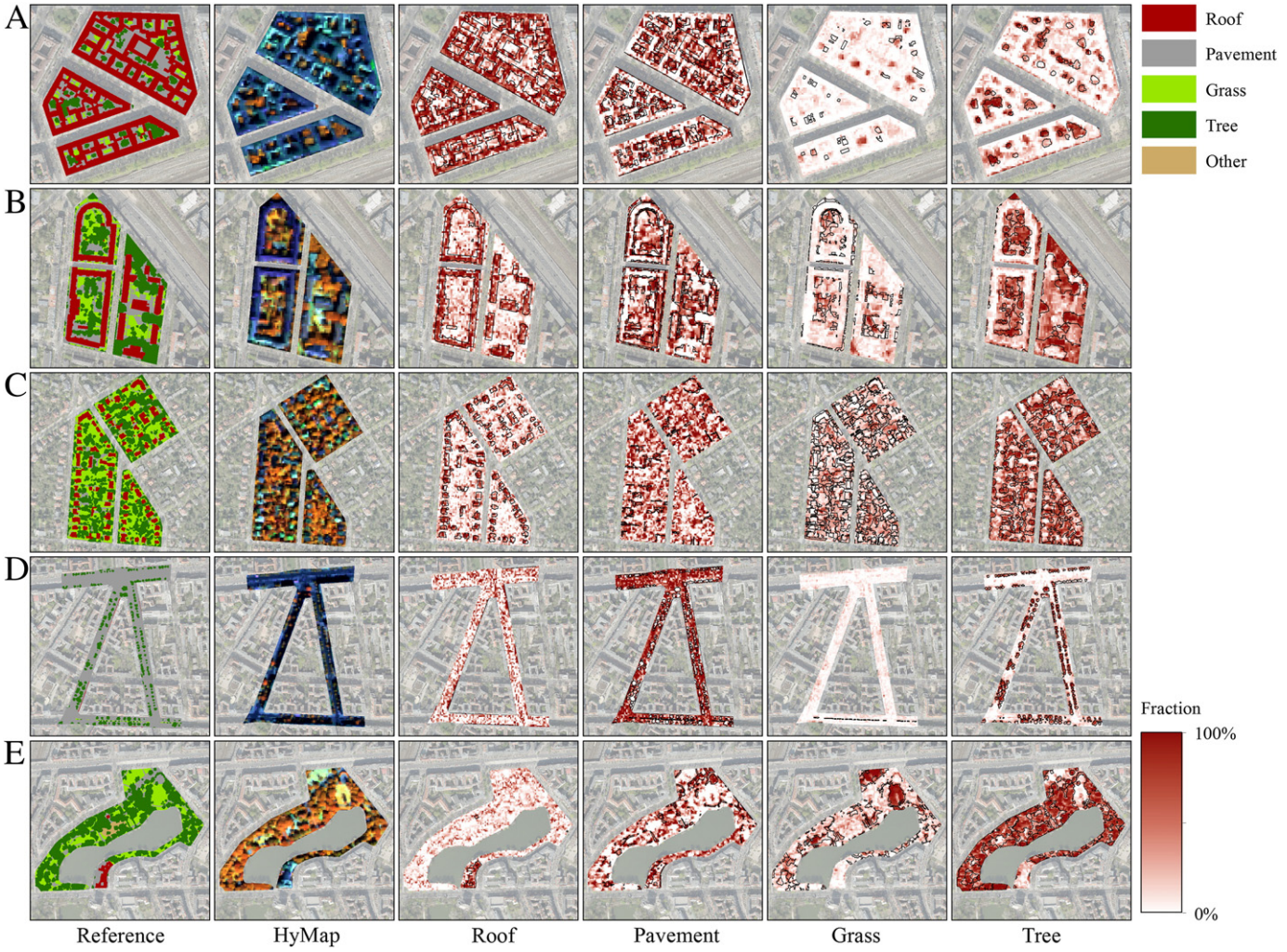

Fig. 5. Subsets showing urban blocks of different structure types from the reference data, HyMap image (R = 833 nm, G = 1652 nm, B = 632 nm) and fraction maps modeled with SVR. (A) High-density-, (B) medium-density-, (C) low-density urban area, (D) street polygons and, (E) green space. Fraction images are stretched between 0 and 100% (values <0% set to 0%, values >100% set to 100%).

Despite the unfavorable existence of outliers, the comparison of reference data and the pixel-wise spatial distribution of physically unrealistic fraction values indicate that this problem can be compensated for in a post-processing step: by setting negative and super-positive fraction values to 0% and 100%, respectively, physically meaningful cover fractions of the target category are achieved (see Fig. 5).

5.2. Urban land cover fraction estimates by SVR and MESMA at urban block scale

Accurate maps of urban surface properties, i.e., fractional abundances of various types of impervious and vegetated surface cover types, serve as an important data basis for proper urban planning or ecological studies (Cadenasso et al., 2007; Heiden et al., 2012; Pauleit & Duhme, 2000). Aggregated land cover information at the block scale constitutes a suitable spatial unit for such purposes (Heiden et al., 2012; Pauleit & Duhme, 2000). We statistically evaluated sub-pixel fraction estimates by SVR at the urban block level. The quality of results was further evaluated by means of a comparison to MESMA. By (i) using the same image spectral library, and (ii) considering the same mixing complexity, we created identical starting points for a fair methodological comparison.

The high accuracies achieved at the urban block scale underscore the capability of SVR, along with the synthetically mixed training data, to quantitatively map urban land cover (see Fig. 6). For a total number of 92 urban blocks we differentiated four spectrally complex and similar target categories with high precision. Reduced accuracies of *pavement* estimates generally indicate the critical mapping of paved surface types like roads, sidewalks or parking lots. Effective discrimination between surfaces from the target category and from the remaining background cover types is illustrated by the reproduced spatial patterns of the SVR-based fraction maps (see Fig. 5). The high performance of the proposed SVR approach is particularly noteworthy for the spectrally complex *roof* category, which includes a diversity of roofing materials at different illumination conditions. The same applies for the approach's capability to handle both vegetation categories, *grass* and *tree*. For example, grass surfaces with differences in physiological condition were mapped as desired with maximal fractions. Further, trees along street corridors or within green spaces were in large part unambiguously quantified from background cover types. Inaccurate mappings of vegetation types were low and occasionally occurred when brightly illuminated tree crowns were confused with photosynthetically active grass. This applies, for example, to individual sun-exposed trees along streets or in backyards, parking lots and the like. Inaccurate mappings of both vegetation types for the allotment gardens can be, to a great extent, explained by uncertainties in the reference data, where the interactive mapping of highly fragmented patches of grass and trees appears critical. Adapting the synthetically

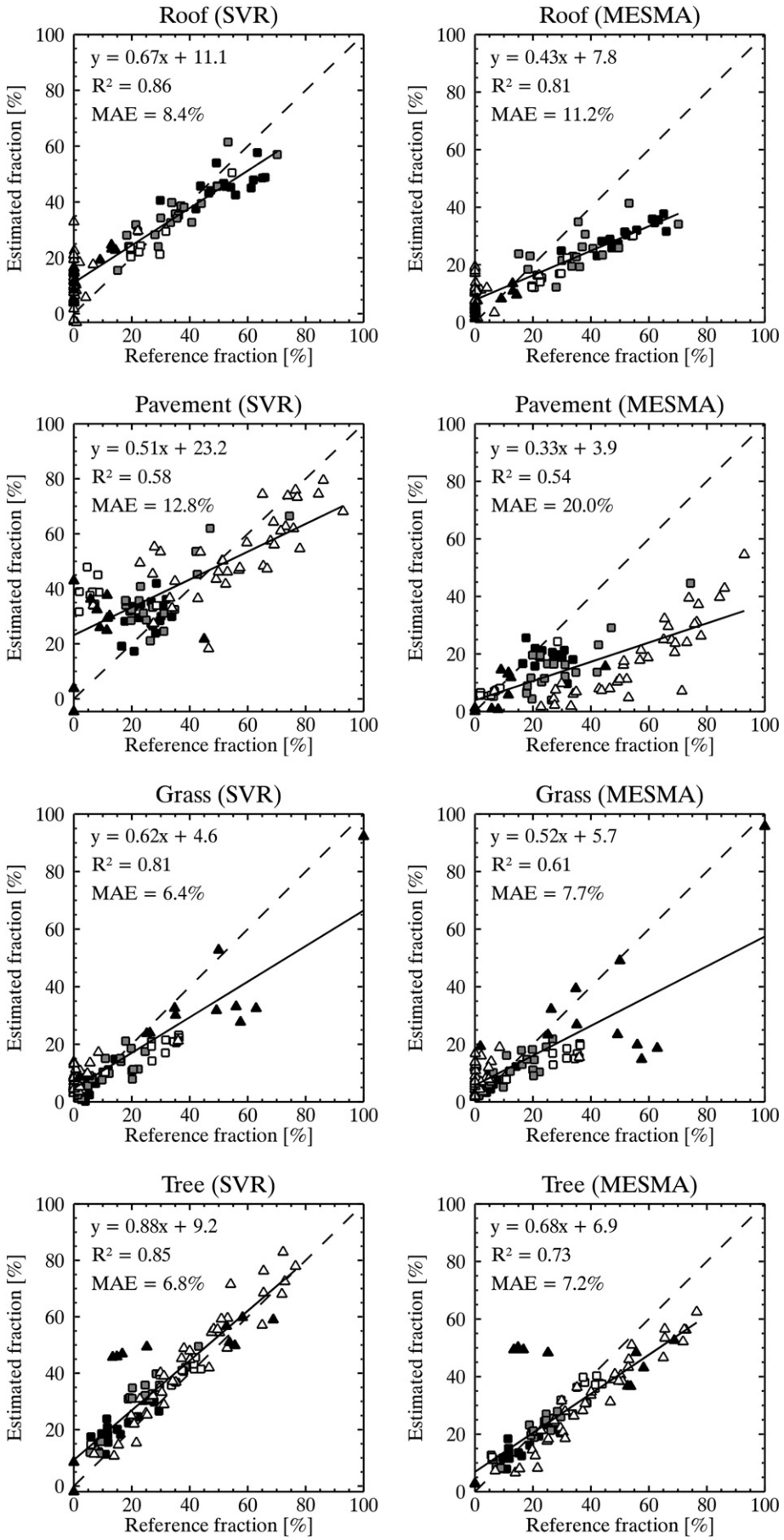

Fig. 6. Scatterplots of *roof*, *pavement*, *grass* and *tree* fractions for 92 urban blocks from the reference data set compared to modeled SVR and MESMA (two-EM models) fractions. Square symbols = building blocks (black = high-density-, gray = medium-density-, white = low-density urban area); triangle symbols = green spaces (black), street polygons (white).

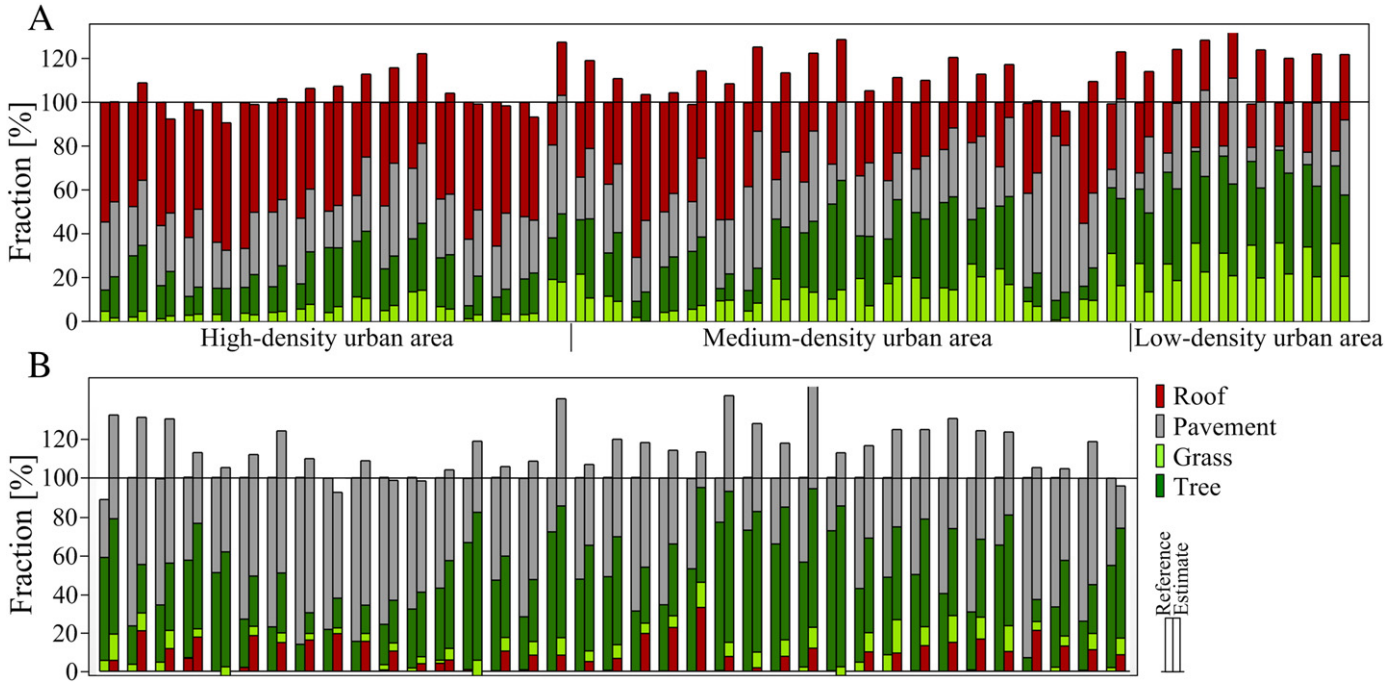

Fig. 7. Combined *roof*, *pavement*, *grass* and *tree* fraction estimates for (A) 45 building blocks and (B) 37 street polygons.

mixed training data to more detailed vegetation characteristics, i.e., distinguishing green vegetation from non-photosynthetic vegetation, or including vegetation species, could be easily incorporated but was beyond the scope of this work.

Similar to previous studies (Franke et al., 2009; Roberts et al., 2012), MESMA is demonstrated to be a powerful technique for the quantitative mapping of urban land cover. This is underscored by the accurate results, particularly for *roof*, *grass* and *tree* estimates (see Fig. 6). Likewise, the quantitative description of the *pavement* category appears critical. Performances are very similar for two different setups, i.e., two-EM models (two material EMs) and three-EM models (two material EMs, one shade EM) with subsequent shade normalization.

When directly comparing the results of both approaches, SVR achieves higher accuracies at the urban block scale (see Fig. 6). However, considering the reference data, both negative and super-positive fraction values at the pixel scale (see Fig. 4) have a positive balancing effect on the calculation of urban block mean values. This must be taken into account when interpreting the SVR-based fraction estimates at the urban block scale. The SVR and MESMA mappings show common sources of errors, which are generally linked to the limitations of urban land cover assessments on a spectral basis, even when using imaging spectrometer data. Misleadingly mapped *roof* fractions along street corridors can be attributed to the spectral similarity or ambiguity of many impervious surface types, particularly to the well-known confusion between darker bituminous roofing materials and street asphalt (Franke et al., 2009; Herold et al., 2004; van der Linden et al., 2007), or to the high numbers of cars, which can be spectrally more similar to roofing materials (van der Linden et al., 2007). Similarly, mapped *roof* fractions within green spaces can be attributed to the confusion between spectra of dark roofing materials and spectra of shaded parts of vegetation stands. The lower accuracies of *pavement* can be ascribed to two major sources of uncertainty. First, paved surfaces contain a higher portion of spectrally ambiguous materials, leading to more spectral confusion. In contrast, the higher number of distinct spectra in the *roof*, *grass* and *tree* categories leads to unambiguous fraction estimates. Second, the high spectral similarity between materials from the *pavement* category, shaded surfaces and water bodies plays an important role, but with fundamental differences between both methods. Using SVR and synthetically mixed training data, shadow-covered grass surfaces or shaded components of trees are misleadingly mapped as having high *pavement* fractions. This effect is particularly visible in urban blocks with high proportions of vegetation, i.e., in low density urban areas or green spaces, and can be attributed to the spectral confusion between the characteristic dark flat spectral signature of asphalt and the low reflectance of shaded areas. In contrast, MESMA underestimates intermediate and high *pavement* fractions, particularly along streets. This is due to the almost zero reflectance water spectrum within the *other* category, which is universally used by MESMA to model non-water pixels with low flat reflectance signatures, e.g., asphalt. Hence, underrated *pavement* fractions are erroneously compensated by high *other* fractions. Including a spectrally flat photometric shade spectrum in the three-EM model (two material EMs, one shade EM) with subsequent shade normalization led to slightly improved *pavement* results, but did not compensate for the observed underestimation along streets. This is due to the similarity of the water and the photometric shade signature. Hence, the use of an additional shade EM is only effective when excluding the water spectrum from the library.

Given these observations, several similarities can be identified for both approaches. In general, the quality of the SVR and MESMA mappings depends on the availability of appropriate training information during modeling. The comprehensiveness of the library spectra ensures the reliable estimation of land cover fractions from the image. Further, common sources of uncertainty for both SVR- and MESMA-based mappings are often related to spectral ambiguities. This shortcoming cannot be overcome by any analysis technique when using spectral information alone. Moreover, both approaches assume that spectrally mixed pixels can be approximated by linear mixing systematics. The high achieved accuracies support this simplification, as well as the omission of possible nonlinear mixing effects occurring from multiple-scattering between objects in the sensors' field of view (Borel & Gerstl, 1994; Somers, Cools, et al., 2009). Finally, both methods can be flexibly adapted to the spectral complexity of the analyzed image data. This is of great importance with respect to the coarser spatial resolution of future spaceborne imaging spectrometer data, where mixtures beyond two urban land cover categories

will prevail in most pixels. In terms of SVR, synthetically mixed training data can be created both for simple binary mixtures and for more complex ones, e.g., ternary or quaternary mixtures. Similarly, MESMA can exploit EM models of different complexities. In addition, expert knowledge can be used to optionally exclude specific mixtures that do not occur in nature.

Despite these similarities, SVR is conceptually different from MESMA. These differences have implications on the implementation and performance of the respective technique, and thus on the ultimate selection of the appropriate method for quantifying urban land cover. First, modeling urban land cover with SVR requires quantitative training information, which is derived by synthetically mixed training data from the spectral library. This enables empirical modeling with regression, as alternative reference information is difficult to collect. MESMA uses pure spectra as EM for estimating fractions by model inversion. Accordingly, no further interaction by the user is necessary. Second, with SVR we generally aim to estimate land cover fractions of specific target categories. Depending on the study purpose, this can be either one individual urban land cover category such as "urban tree", or several subsequently mapped land cover categories to comprehensively describe the physical composition of urban areas. The quality of the proposed SVR approach for comprehensive mapping is demonstrated by the feasibility of combining the fraction maps produced by four different SVR models without imposing constraints (see Fig. 7). One general advantage of SVR – the ability to model the complex functional relations between the set of multiple binary-mixtures and associated mixing fractions in a one-step approach – is exploited. A single global SVR model per target category is trained to map urban land cover, while at the same time accounting for spectral variability. In contrast, MESMA aims to achieve a comprehensive estimation of cover fractions for all urban categories considered in the analysis. Spectral variability is accounted for by iteratively calculating all possible two-EM mixture model combinations on a per-pixel basis. The model with the best fit (i.e., RMSE between observed and modeled spectra) is selected. As all land cover categories in the analysis are considered, MESMA generates an additional fraction map visualizing the fractional abundances of the *other* category, a land cover category of minor interest in this work. Third, accuracies of modeling results depend on the capability of the respective technique to exploit the spectral information given during training. Thus, the general strengths of SVR to cope with spectral complexity and to effectively exploit subtle spectral differences between otherwise spectrally similar materials by individually weighting spectral bands are utilized. This may to a great extent explain the more accurate SVR-based fraction estimates, particularly for *roof*, *grass* and *tree* estimates. In contrast, MESMA equally weights all wavebands, and thus is subject to the risk of underemphasizing these possibly decisive spectral features (Somers, Delalieux, Stuckens, Verstraeten, & Coppin, 2009). The difference in performance between both methods generally indicates the strengths of SVR to optimally exploit the spectral information provided by the training input. Therefore, the excellent suitability of SVMs for mapping complex urban land cover types is again demonstrated.

6. Conclusion

In this paper we demonstrate the potential of the machine learning-based SVR to be used for the quantitative mapping of four spectrally complex and ecologically meaningful urban land cover types on a purely spectral basis. Imaging spectrometer data was used to estimate fractions of impervious rooftops and pavements, as well as grass- and tree-covered areas along the urban gradient of Berlin, Germany. We propose the integration of synthetically mixed training data into the training stage of the analysis. This poses a straightforward and repeatable strategy to overcome the dilemma related to the difficult acquisition of quantitative training information that is needed for sub-pixel mapping with empirical regression techniques.

Constructing synthetically mixed data for SVR model training only requires a spectral library that represents the spectral diversity of the study area. Using an image spectral library allows easy associations to be made with spectral features of the image. Universal spectral databases, e.g., consisting of laboratory or field spectra of man-made urban materials and vegetation species, can be used alternatively, providing that radiometric consistency between library spectra and the image is given. This becomes important when the number of pure spectra present in the image is low, e.g., as expected for future spaceborne imaging spectrometer data of urban areas.

The objective of this study, i.e., to quantify the four spectrally complex and ecologically meaningful urban land cover categories using the proposed SVR approach, was tested by comparing fraction estimates at the pixel and the urban block scale to manual field mappings, and to fraction maps derived by MESMA. Although conceptually different, both methods demonstrate their great potential to quantify the land cover of spectrally heterogeneous urban environments. The proposed SVR approach aims to quantify a single target land cover category in a universal global model that accounts for spectral variability. The subsequent analysis of several target categories allows for a reliable comprehensive mapping of urban land cover. This is different from the traditional MESMA method, which uses pure endmembers in an extensive iterative procedure with multiple linear mixture models for the comprehensive mapping of all land cover categories considered in the analysis. The high accuracies achieved in this study demonstrate that the combination of SVR and synthetically mixed training data can be recommended as an alternative quantitative mapping method, particularly when users are only interested in mapping specific land cover categories of interest. Moreover, improved fraction estimates at the urban block scale by SVR indicate the general strengths of machine-learning-based techniques to effectively cope with the spectral complexity inherent in imaging spectrometer data from urban areas.

Most of the observed uncertainties can be related to the known phenomena of spectral ambiguity or similarity of urban materials and the spectral deficiencies in shaded areas. Whereas spectral ambiguities cannot be overcome by any analysis technique when using spectral information alone, strategies to cope with spectral similarities or shadow effects must be developed, particularly when analyzing impervious surface cover types. Future work should therefore focus on data transformation (e.g., Wu, 2004) or spectral feature extraction techniques (Herold et al., 2004; e.g., Heiden et al., 2007) prior to SVR processing.

The presented method was tested in a spectrally and spatially challenging urban environment (van der Linden & Hostert, 2009). The robustness and accuracy of the approach can therefore be expected to be of similar or better quality in more homogeneous areas. Given the powerful separation of two vegetations types, the quantitative mapping of natural ecosystems appears especially worthwhile. By separating the model training process from the model application and validation, the general idea of a universal quantitative mapping approach is given. This is important for future studies, when accuracy, reliability and transferability of SVR models to image data with different spatial extents, spatial resolution and varying acquisition dates must be tested.

Acknowledgments

This research is funded by the German Research Foundation (DFG) under project no. HO 2568/2-2 and partly funded by the German Aerospace Centre (DLR) and the Federal Ministry of Economics and Technology (BMWi) as part of the EnMAP Core Science Team activities (FKZ 50EE0949). The authors would like to thank the two anonymous reviewers for their valuable comments. The authors are grateful to M. Cierpinski for generating the reference data, A. Rabe, B. Jakimow and S.

Suess for the support during SVR analysis and helpful comments. The contribution of the students from Humboldt-Universität zu Berlin, who assisted with the field survey, is greatly appreciated. The HyMap data set was acquired and pre-processed by the DLR in the framework of the HyEurope 2009 campaign. HyEurope 2009 was funded by the German Federal Ministry of Education and Research (BMBF). We thank the Berlin Senate Department for Urban Development for providing data from the Urban and Environmental Information System (UEIS) and the DSM.

References

Alberti, M. (2005). The effects of urban patterns on ecosystem function. *International Regional Science Review*, *28*, 168–192.

Arnold, C. L., & Gibbons, C. J. (1996). Impervious surface coverage — The emergence of a key environmental indicator. *Journal of the American Planning Association*, *62*, 243–258.

Asner, G. P., & Lobell, D. B. (2000). A biogeophysical approach for automated SWIR unmixing of soils and vegetation. *Remote Sensing of Environment*, *74*, 99–112.

Blair, R. B. (1996). Land use and avian species diversity along an urban gradient. *Ecological Applications*, *6*, 506–519.

Borel, C. C., & Gerstl, S. A. W. (1994). Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment*, *47*, 403–416.

Brereton, R. G., & Lloyd, G. R. (2010). Support vector machines for classification and regression. *Analyst*, *135*, 230–267.

Cadenasso, M. L., Pickett, S. T. A., & Schwarz, K. (2007). Spatial heterogeneity in urban ecosystems: Reconceptualizing land cover and a framework for classification. *Frontiers in Ecology and the Environment*, *5*, 80–88.

Camps-Valls, G., Bruzzone, L., Rojo-Alvarez, J. L., & Melgani, F. (2006). Robust support vector regression for biophysical variable estimation from remotely sensed images. *IEEE Geoscience and Remote Sensing Letters*, *3*, 339–343.

Chang, C. -C., & Lin, C. -J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*.

Chudnovsky, A., Ben-Dor, E., & Saaroni, H. (2004). Diurnal thermal behavior of selected urban objects using remote sensing measurements. *Energy and Buildings*, *36*, 1063–1074.

Cocks, T., Jenssen, R., Stewart, A., Wilson, I., & Shields, T. (1998). The HyMap™ airborne hyperspectral sensor: The system, calibration and performance. *1st Earsel workshop on imaging spectroscopy* (pp. 37–42).

National Research Council (2007). *Earth science and applications from space: National imperatives for the next decade and beyond*. Washington, D.C.: National Academies Press.

Esch, T., Himmler, V., Schorcht, G., Thiel, M., Wehrmann, T., Bachofer, F., et al. (2009). Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data. *Remote Sensing of Environment*, *113*, 1678–1690.

Foody, G. M., & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, *103*, 179–189.

Franke, J., Roberts, D. A., Halligan, K., & Menz, G. (2009). Hierarchical multiple endmember spectral mixture analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment*, *113*, 1712–1723.

Gluch, R., Quattrochi, D. A., & Luvall, J. C. (2006). A multi-scale approach to urban thermal analysis. *Remote Sensing of Environment*, *104*, 123–132.

Griffiths, P., Hostert, P., Gruebner, O., & van der Linden, S. (2010). Mapping megacity growth with multi-sensor data. *Remote Sensing of Environment*, *114*, 426–439.

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J. G., Bai, X. M., et al. (2008). Global change and the ecology of cities. *Science*, *319*, 756–760.

Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T., & Mueller, A. (2012). Urban structure type characterization using hyperspectral remote sensing and height information. *Landscape and Urban Planning*, *105*, 361–375.

Heiden, U., Segl, K., Roessner, S., & Kaufmann, H. (2007). Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sensing of Environment*, *111*, 537–552.

Heldens, W., Heiden, U., Esch, T., Stein, E., & Mueller, A. (2011). Can the future EnMAP mission contribute to urban applications? A literature survey. *Remote Sensing*, *3*, 1817–1846.

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

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.

Herold, M., Schiefer, S., Hostert, P., & Roberts, D. A. (2006). Applying imaging spectrometry in urban areas. In D. W. Quattrochi, & Q. (Eds.), *Urban remote sensing* (pp. 137–161). New York: CRC Press Inc.

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*, 725–749.

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

Maktav, D., Erbek, F. S., & Jurgens, C. (2005). Remote sensing of urban areas. *International Journal of Remote Sensing*, *26*, 655–659.

McKinney, M. L. (2002). Urbanization, biodiversity, and conservation. *Bioscience*, *52*, 883–890.

Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, *42*, 1778–1790.

Miller, R. B., & Small, C. (2003). Cities from space: Potential applications of remote sensing in urban environmental research and policy. *Environmental Science & Policy*, *6*, 129–137.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*, 247–259.

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, *115*, 1145–1161.

Pal, M., & Mather, P. M. (2006). Some issues in the classification of DAIS hyperspectral data. *International Journal of Remote Sensing*, *27*, 2895–2916.

Pauleit, S., & Duhme, F. (2000). Assessing the environmental performance of land cover types for urban planning. *Landscape and Urban Planning*, *52*, 1–20.

Pickett, S. T. A., Cadenasso, M. L., Grove, J. M., Boone, C. G., Groffman, P. M., Irwin, E., et al. (2011). Urban ecological systems: Scientific foundations and a decade of progress. *Journal of Environmental Management*, *92*, 331–362.

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.

Rabe, A., Jakimow, B., van der Linden, S., & Hostert, P. (2012). EnMAP box, version 1.4 [online]. Available from: http://indus.caf.dlr.de/forum/ (accessed April 2013).

Rabe, A., van der Linden, S., & Hostert, P. (2010). Imagesvm, version 2.1. [online]. Available from: http://www.imagesvm.net/ (accessed April 2013).

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

Richter, R., & Schläpfer, 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., 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.

Roberts, D. A., Quattrochi, D. A., Hulley, G. C., Hook, S. J., & Green, R. O. (2012). Synergies between VSWIR and TIR data for the urban environment: An evaluation of the potential for the hyperspectral infrared imager (HyspIRI) decadal survey mission. *Remote Sensing of Environment*, *117*, 83–101.

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*, 1525–1532.

Schiefer, S., Hostert, P., & Damm, A. (2006). Correcting brightness gradients in hyperspectral data from urban areas. *Remote Sensing of Environment*, *101*, 25–37.

Schläpfer, D., & Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: Parametric orthorectification. *International Journal of Remote Sensing*, *23*, 2609–2630.

Schölkopf, B., & Smola, A. J. (2002). *Learning with kernels — Support vector machines, regularization, optimization, and beyond*. Cambridge, Massachusetts: MIT Press.

SenStadt (2013a). Berlin urban and environmental information system (UEIS) [online]. Available from: www.stadtentwicklung.berlin.de/umwelt/umweltatlas (accessed April 2013).

SenStadt (2013b). Automatisierte Liegenschaftskarte (ALK-Berlin) [online]. Available from: www.stadtentwicklung.berlin.de/geoinformation/liegenschaftskataster/ (accessed April 2013).

Shackelford, A. K., & Davis, C. H. (2003). A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 2354–2363.

Small, C. (2001). Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing*, *22*, 1305–1334.

Small, C. (2003). High spatial resolution spectral mixture analysis of urban reflectance. *Remote Sensing of Environment*, *88*, 170–186.

Small, C. (2004). The Landsat ETM + spectral mixing space. *Remote Sensing of Environment*, *93*, 1–17.

Small, C., & Lu, J. W. T. (2006). Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis. *Remote Sensing of Environment*, *100*, 441–456.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*, 199–222.

Somers, B., Asner, G. P., Tits, L., & Coppin, P. (2011). Endmember variability in spectral mixture analysis: A review. *Remote Sensing of Environment*, *115*, 1603–1616.

Somers, B., Cools, K., Delalieux, S., Stuckens, J., van der Zande, D., Verstraeten, W. W., et al. (2009). Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards. *Remote Sensing of Environment*, *113*, 1183–1193.

Somers, B., Delalieux, S., Stuckens, J., Verstraeten, W. W., & Coppin, P. (2009). A weighted linear spectral mixture analysis approach to address endmember variability in agricultural production systems. *International Journal of Remote Sensing*, *30*, 139–147.

Stehman, S. V., & Wickham, J. D. (2011). Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment*, *115*, 3044–3055.

Stuffler, T., Förster, K., Hofer, S., Leipold, M., Sang, B., Kaufmann, H., et al. (2009). Hyperspectral imaging — An advanced instrument concept for the EnMAP mission (Environmental Mapping and Analysis Programme). *Acta Astronautica*, *65*, 1107–1112.

Stuffler, T., Kaufmann, C., Hofer, S., Förster, K. P., Schreier, G., Mueller, A., et al. (2007). The EnMAP hyperspectral imager — An advanced optical payload for future applications in Earth observation programmes. *Acta Astronautica*, *61*, 115–120.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., & Dech, S. (2012). Monitoring urbanization in mega cities from space. *Remote Sensing of Environment*, *117*, 162–176.

Thomas, N., Hendrix, C., & Congalton, R. G. (2003). A comparison of urban mapping methods using high-resolution digital imagery. *Photogrammetric Engineering and Remote Sensing*, *69*, 963–972.

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.

Tuia, D., & Camps-Valls, G. (2011). Urban image classification with semisupervised multiscale cluster kernels. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *4*, 65–74.

Tuia, D., Verrelst, J., Alonso, L., Perez-Cruz, F., & Camps-Valls, G. (2011). Multioutput support vector regression for remote sensing biophysical parameter estimation. *IEEE Geoscience and Remote Sensing Letters*, *8*, 804–808.

van der Linden, S., & Hostert, P. (2009). The influence of urban structures on impervious surface maps from airborne hyperspectral data. *Remote Sensing of Environment*, *113*, 2298–2305.

van der Linden, S., Janz, A., Waske, B., Eiden, M., & Hostert, P. (2007). Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines. *Journal of Applied Remote Sensing*, *1*.

Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer.

Verrelst, J., Munoz, J., Alonso, L., Delegido, J., Pablo Rivera, J., Camps-Valls, G., et al. (2012). Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sensing of Environment*, *118*, 127–139.

Walton, J. T. (2008). Subpixel urban land cover estimation: Comparing cubist, random forests, and support vector regression. *Photogrammetric Engineering and Remote Sensing*, *74*, 1213–1222.

Waske, B., van der Linden, S., Benediktsson, J. A., Rabe, A., & 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*, 2880–2889.

Wu, C. (2004). Normalized spectral mixture analysis for monitoring urban composition using ETM + imagery. *Remote Sensing of Environment*, *93*, 480–492.