Mapping impervious surface fractions using automated Fisher transformed unmixing

Fei Xu^a, Xin Cao^{b,*}, Xuehong Chen^b, Ben Somers^{a,*}

^a Division of Forest, Nature and Landscape, KU Leuven, Leuven, Belgium

^b State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China

ARTICLE INFO

Edited by: Emilio Chuvieco

Keywords:

Spectral mixture analysis
Fisher discriminant analysis
Spectral variability
Impervious surfaces
Urban
MESMA

ABSTRACT

Spatial explicit monitoring of impervious surface fractions provides essential information for urban planning, urban disaster prevention and mitigation. Moderate spectral and spatial resolution satellite observation has become a tool to perform such monitoring across different spatial and temporal scales. Yet, one major challenge in urban remote sensing is the high spatial heterogeneity causing the presence of image pixels constituting a mixture of different land cover materials. Here we apply an automated unmixing procedure specifically designed to optimally account for the pronounced spectral variability (i.e. high within-class and low between-class variability) prevailing in urban settings. The approach is based on Fisher Discriminant Analysis (FDA), a data dimensionality reduction technique in which between-class variability is maximized and within-class variability is minimized while transforming spectra from the reflectance to the Fisher feature space. We integrated the FDA transformation in the widely-applied Multiple Endmember Spectral Mixture Analysis (MESMA) approach in an attempt to more effectively address the endmember variability problem. We made use of online spectral libraries to train the Fisher transformation parameters and steer the automated selection of image endmembers. Our Fisher transformed MESMA algorithm (F-MESMA) was tested on medium spatial resolution Landsat scenes and impervious fraction maps were generated for five major cities located in different regions of the world. The output of F-MESMA was compared to the output of state-of-the-art spectral mixture analysis approaches which were specifically designed to reduce endmember variability issues. We could demonstrate that compared to other spectral transformation approaches, the ratio of within- vs between-class variability of the endmembers was most strongly reduced after applying the Fisher transformation. As a direct consequence, across all five cities F-MESMA consistently provided the most accurate impervious surface fraction estimates ($RMSE_{F-MESMA} = 0.13$ vs. $RMSE_{alternative\ approaches} = [0.16-0.17]$).

1. Introduction

Impervious surfaces - manmade constructions such as rooftops, drive ways and parking lots - not only comprise a major component of urban land cover, but also play an important role in regulating urban hydrological and heat fluxes (Ward et al., 2018). Together with the expected urban growth, impervious surface fractions will progressively grow, thereby contributing to global urban environmental problems such as urban heat islands (Coseo and Larsen, 2014), urban flooding (Du et al., 2016) and water quality deterioration (Conway et al., 2007); and this will be particularly true for development countries where urban expansion is expected to be most intense (Weng, 2012). Sustainable urban planning therefore requires accurate monitoring of impervious surfaces on both local, regional and global scales.

Spaceborne remote sensing has become the most efficient alternative to monitor impervious surfaces across different scales (Yang et al., 2010; Lu and Weng, 2006). In general, highest mapping accuracies and highest thematic detail are obtained using high spatial resolution satellite imagery (e.g. IKONOS, Quickbird), yet expensive costs limit their feasibility to conduct large scale mapping (Sawaya et al., 2003). By comparison, regular observation times, stable data quality and high accessibility make medium spatial resolution images most suitable to monitor impervious surface distribution across scales (Weng, 2012). Yet, due to high heterogeneity in urban landscapes and limitations in spatial resolution, the presence of mixed pixels has been recognized as a major problem when mapping impervious surfaces using medium spatial resolution images (Weng, 2012). Spectral mixture analysis (SMA) is a common image processing technique to deal with

* Corresponding authors.

E-mail addresses: caoxin@bnu.edu.cn (X. Cao), ben.somers@kuleuven.be (B. Somers).

<https://doi.org/10.1016/j.rse.2019.111311>

Received 24 October 2018; Received in revised form 23 May 2019; Accepted 10 July 2019

Available online 18 July 2019

0034-4257/ © 2019 Elsevier Inc. All rights reserved.

mixed pixels (Burazerović et al., 2013). SMA models a mixed spectrum as a combination of pure spectral signatures of its constituent components (i.e. endmembers), weighted by their sub-pixel fractional cover (Keshava and Mustard, 2002). Once the endmembers have been extracted, the fractional distribution of endmembers within mixed pixels can be generated using approaches such as maximum likelihood (Settle, 2006) or least-square regression analysis (Barducci and Mecocci, 2005). Two types of SMA techniques are widely used: linear SMA (LSMA), assuming that only single reflection happens from each land cover component (or endmember) within a pixel; and nonlinear SMA, accounting for multiple photon interactions among different endmembers (Keshava and Mustard, 2002). Although experimental studies have shown that nonlinear SMA outperforms LSMA if image pixels are composed of ground objects with an expressed three-dimensional structure (e.g. trees, buildings) (Somers et al., 2014; Dobigeon et al., 2014), given its relative simplicity and because of lower sensitivity to collinearity and noise, LSMA has been extensively applied to monitor impervious surface distributions (Weng, 2012). In particular, the V (vegetation)-I (impervious)-S (soil) model (Ridd, 1995) which considers the urban environment a combination of vegetation, impervious surfaces and soil, is widely used when conducting urban subpixel mapping using medium spatial resolution remote sensing images such as Landsat 5 TM, Landsat 8 OLI or Sentinel-2A (Phinn et al., 2002; Ward et al., 2000; Wu and Murray, 2003).

Due to variation in material composition and properties, geometry and illumination conditions, impervious surfaces generally show high levels of spectral variability (Herold and Roberts, 2005), posing additional difficulties in selecting representative impervious surface endmembers and increasing the sensitivity of the LSMA model to noise and collinearity (Somers et al., 2011). Two types of spectral variability are commonly distinguished: the variability within a material (or endmember) class (within-class variability), and the similarity among endmember spectra of different classes (inter-class variability) (Zhang et al., 2004; Somers et al., 2011). It has been proven that the accuracy of sub-pixel fraction estimates linearly decreases with within-class variability (Barducci and Mecocci, 2005), because high within-class variability will lead to a high likelihood that the endmember's actual spectral characteristics in a specific pixel can deviate from predefined standardized and fixed endmembers used in LSMA. Equally, low between-class variability caused by high similarity among the spectral characteristics of different material classes or endmembers is known to lead to a dramatic drop in estimation accuracy; a phenomenon referred to as collinearity. Indeed the more pronounced the spectral similarity among endmember classes, the more sensitive the LSMA model is for white noise in the endmember spectra (Chen et al., 2011; Van der Meer and Jia, 2012). According to Somers et al. (2011) the negative effect of high within-class and low between-class variability is further referred to as the endmember variability problem.

Many solutions have been proposed to address the endmember variability problem in LSMA (see Somers et al., 2011 and Zare and Ho (2014)) for a comprehensive review of all methods). Multiple Endmember Spectral Mixture Analysis (MESMA) (Roberts et al., 1998) is a widely used method to cope with within-class variability. When conducting MESMA, endmembers are allowed to vary in each pixel and an iterative strategy is adopted for selecting an optimal group of endmembers per pixel from a preselected endmember spectral library. MESMA showed to be effective in addressing the within-class variability problem in many ecosystems, including urban areas (Franke et al., 2009; Degerickx et al., 2017) and demonstrated to be effective to unmix both multi- (Quintano et al., 2013) and hyperspectral imagery (Franke et al., 2009). Wu (2004) developed the Normalized Linear Spectral Mixture Analysis algorithm (NSMA) to reduce brightness variation within an endmember class. Zhang et al. (2004) proposed derivative spectral unmixing (DSU), proving that the second derivative of reflectance spectra can reduce the within-class variability while enhance between-class variability. Somers et al. (2010a, 2010b) proposed Stable

Zone Unmixing (SZU), a waveband selection algorithm to address spectral variability problems in LSMA. By selecting only wavelengths showing low(er) within-class variability and high(er) between-class variability SZU allows to increase fraction estimate accuracies compared to traditional unmixing approaches using the full suite of available wavebands. Even though these methods are successfully tested in several case studies, each of them has its constraints. For instance, while NSMA reduces within-class brightness variation, it may increase the collinearity between bright impervious surfaces and soil (Chen et al., 2011). DSU proved to enlarge between-class variability between lichen and rock, but in an operational setting, and in a complex urban environment, the process of calculating spectral derivatives might be negatively impacted by additional signal noise such as shade effects and changing illumination conditions (Wu, 2004). Moreover, DSU has specifically been designed to unmix hyperspectral images. Also SZU, originally developed for hyperspectral data, may have its restriction due to the limited number of available wavebands when using medium spatial resolution imagery (Somers et al., 2010a, 2010b). In comparison with aforementioned endmember variability reduction methods, Fisher discriminant analysis (FDA) conducts a transformation – referred to as Fisher transformation in the remainder of this paper – to search for the projection directions on which the distance of data samples of different endmember classes is maximized while data samples of the same class are minimized. In the resulting new feature space, i.e. the Fisher feature space, the data as such demonstrates higher between-class and lower within-class variability compared to the original spectral reflectance space (Fukunaga, 2013; Sugiyama, 2007). By reducing the ratio of within and between-class variability, FDA based linear spectral unmixing holds as such the potential to improve endmember fraction estimate accuracy, which has indeed been demonstrated by Liu et al. (2017) and Jing et al. (2010) showing better accuracy than conventional SMA techniques for vegetation and crop cover fraction estimates, respectively.

Another problem that influences the accuracy of LSMA fraction maps is improper endmember selection (Elmore et al., 2000). Selecting endmembers involves identifying both the number and type of endmembers and their corresponding spectral signatures (Somers et al., 2011). A number of advanced analysis techniques to locate and extract endmembers from image data have been described. The Pixel Purity Index (PPI, Boardman et al., 1995), Vertex Component Analysis (VCA, Nascimento and Dias, 2005) and N-FINDR (Winter, 1999) are a few of the more commonly used algorithms generally applied on hyperspectral data. By applying these automated endmember extraction algorithms (EEA) on image subsets, Somers and Asner (2012) obtained endmember bundles to better represent the spectral variability of the image scene components. All these techniques assume that endmembers constitute the vertices of a convex polytope covering the image data points in high dimensional space. PPI randomly produces N number of projection vectors onto which image pixels are projected. The pixels that most often locate at the vertex of the convex polytope are selected as endmembers. VCA iteratively projects data onto a direction orthogonal to the subspace spanned by the endmembers already determined. The endmember spectra of the new class correspond to the new vertex point of the polytope (Nascimento and Dias, 2005). N-FINDR replaces the already determined endmembers with candidate image pixels. If doing so the simplex volume increases the candidates are accepted as new endmembers. Even though these techniques could produce image endmembers automatically, calibrating the input parameters of these techniques is not always straightforward. This is especially true for extracting impervious surface endmembers. Impervious surfaces show serious variability both in type and number of endmembers making it hard to define in advance the optimal number of endmembers for PPI, and the optimal number of iterations for VCA and N-FINDR.

Automated unmixing for deriving impervious surface fractions is thus hindered by high spectral variability (i.e. high within-class and low between-class variability) complicating both the endmember selection

as well as the spectral unmixing process (Weng, 2012). In this study, we evaluate the potential of an automated Fisher transformed unmixing approach, applied on medium spatial and spectral resolution Landsat imagery, to map subpixel impervious surface fractions for five major cities. We hypothesize that the FDA transformation will reduce the ratio of within- and between-class endmember variability and therefore will provide more accurate LSMA derived fraction estimates. Previous studies already introduced the basic concepts of FDA based unmixing (Liu et al., 2017; Jing et al., 2010). Yet, here we complement these studies by (i) providing an operational FDA-based workflow for unmixing of impervious surface fractions. Specific is that in our approach the automatic identification of image endmembers as well as the unmixing step itself (i.e. estimating endmember fractions) is based on FDA. Moreover, we combine FDA with MESMA, a state-of-the-art SMA technique commonly applied to address endmember variability issues in urban environments. We hypothesize that a hybrid FDA-MESMA (F-MESMA) approach is more effective in addressing impervious surface endmember variability and consequently provides more robust subpixel classification results compared to traditional MESMA techniques. To test the generic nature and robustness of our approach we selected five cities located in different regions of the world, with variable population densities and subject to different climate types.

2. Theoretical background

This study proposes an automated unmixing method for extraction

of impervious surfaces from medium spatial resolution satellite imagery. The method consists of three steps as detailed in the flowchart in Fig. 1. In the first step, reflectance images and generic spectral library data of vegetation, impervious surface, soil, and water (V-I-S-W) are transformed into Fisher feature images using a Fisher transformation (Sugiyama, 2007). In the second step, site specific image endmembers for vegetation, impervious surfaces and soil are automatically identified and extracted from the input image. In the third step, Multiple Endmember Spectral Mixture Analysis (MESMA) was adopted to produce V-I-S fractions. By performing the whole unmixing procedure in Fisher feature space a significant reduction in the ratio of within- and between-class endmember variability and consequently a higher fraction estimate accuracy is anticipated. In the subsequent sections each step of the methodology is explained in detail.

2.1. Fisher transformation to simultaneously reduce the ratio of within-class and between-class endmember variability

Based on training samples of vegetation, impervious surfaces, soil and water pixels, the Fisher transformation searches for the projection directions on which the distance between data points of different endmember classes is maximized while the distance between data points of the same class is minimized. So, the original reflectance image projected in the Reflectance space R^{nb} , where nb is the number of spectral bands in the image, is transformed to Fisher feature space R^{c-1} using calculated projection directions. In practice, training samples can

Fig. 1. Flowchart of the proposed method, including three steps. (i) Image transformation; (ii) Endmember selection; (iii) V-I-S fraction estimation.

be selected from existing spectral libraries (See section 3.2) resulting in a set of spectra, $\mathbf{m}_1^c, \mathbf{m}_2^c, \dots, \mathbf{m}_N^c$, belonging to the land cover type $c \in \{\text{vegetation, impervious surface, bare soil and water}\}$. The V-I-S-W training spectra are used for maximizing the objective function (Eq. (1)) whose numerator is the projected between-class variance matrix and denominator is the projected within-class variance matrix. Maximizing the objective function comes down to minimizing the denominator (within-class variability) and simultaneously maximizing the numerator (between-class variability). The projection vector $\mathbf{w}$ which corresponds to the maximum value of the objective function is considered as the well-trained Fisher transformation projection vector. The process of maximizing the objective function, and thus to find $\mathbf{w}$, is equal to obtaining the eigenvectors of the matrix $S_w^{-1}S_b$ (Eq. (2), Okada and Tomita, 1985). Importantly, because the rank of $S_w^{-1}S_b$ (in Eq. (2)) is always less than the number of sample categories (Okada and Tomita, 1985), in this study we build LSMA models with three endmembers (V-I-S), four types of training samples are required (e.g. in this study we used water, as the 4th group of training data).

The objective function of the Fisher transformation is shown in (1), where $\mathbf{w}$ is the projection vector; $\bar{\mathbf{x}}$ is the total sample mean spectrum; $\bar{\mathbf{x}}_v, \bar{\mathbf{x}}_i, \bar{\mathbf{x}}_s$ and $\bar{\mathbf{x}}_w$ are the average spectrum; n_v, n_i, n_s and n_w are the corresponding number of samples and $\mathbf{x}_{v,i}, \mathbf{x}_{i,i}, \mathbf{x}_{s,i}$ and $\mathbf{x}_{w,i}$ are the i^{th} sample spectrum of vegetation, impervious surface, soil and water respectively. S_w and S_b are the within-class scatter matrix and between-class scatter matrix, respectively:

$$J = \frac{\mathbf{w}^T S_b \mathbf{w}}{\mathbf{w}^T S_w \mathbf{w}} \quad (1)$$

$$\begin{aligned} S_b &= \frac{1}{N} \left[n_v (\bar{\mathbf{x}}_v - \bar{\mathbf{x}})(\bar{\mathbf{x}}_v - \bar{\mathbf{x}})^T + n_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^T + n_s (\bar{\mathbf{x}}_s - \bar{\mathbf{x}})(\bar{\mathbf{x}}_s - \bar{\mathbf{x}})^T + n_w (\bar{\mathbf{x}}_w - \bar{\mathbf{x}})(\bar{\mathbf{x}}_w - \bar{\mathbf{x}})^T \right], \\ S_w &= \frac{1}{N} \left[\sum_{i=1}^{n_v} (\mathbf{x}_{v,i} - \bar{\mathbf{x}}_v)(\mathbf{x}_{v,i} - \bar{\mathbf{x}}_v)^T + \sum_{i=1}^{n_i} (\mathbf{x}_{i,i} - \bar{\mathbf{x}}_i)(\mathbf{x}_{i,i} - \bar{\mathbf{x}}_i)^T + \sum_{i=1}^{n_s} (\mathbf{x}_{s,i} - \bar{\mathbf{x}}_s)(\mathbf{x}_{s,i} - \bar{\mathbf{x}}_s)^T + \sum_{i=1}^{n_w} (\mathbf{x}_{w,i} - \bar{\mathbf{x}}_w)(\mathbf{x}_{w,i} - \bar{\mathbf{x}}_w)^T \right] \end{aligned}$$

the projection vector $\mathbf{w}$ (also the eigenvector of the matrix $S_w^{-1}S_b$) that maximizes J in Eq. (1) can be obtained by solving the following equation:

$$S_w^{-1}S_b \mathbf{w} = \lambda \mathbf{w} \quad (2)$$

in which λ is a nonzero eigenvalue. After solving Eq. (2), $c-1$ nonzero eigenvalues can be obtained if there are c number of classes as the rank of $S_w^{-1}S_b$ is $c-1$ and a new feature space $\mathbf{R}^{c-1}$ is spanned by the projection vector $\mathbf{w}$ (Okada and Tomita, 1985). In this study, three nonzero eigenvalues can be used for calculating three Fisher projection vectors ($\mathbf{w} = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$). With these projection vectors, reflectance space can be linearly transformed into Fisher space using Eq. (3).

2.2. Fisher-transformation based Spectral Mixture Analysis

2.2.1. Fisher-transformed Linear Spectral Mixture Analysis

Once Fisher projection vectors are obtained, LSMA can be performed in Fisher feature space:

$$FM = FR_v f_v + FR_i f_i + FR_s f_s + F\varepsilon \quad (3)$$

$$f_v + f_i + f_s = 1$$

$$0 \leq f_v, f_i, f_s \leq 1$$

where $M_{nb \times 1}$ is the mixed reflectance spectrum of V-I-S endmembers, $\mathbf{R}_v, \mathbf{R}_i, \mathbf{R}_s$ are reflectance endmembers, and f_v, f_i, f_s are the fractions of vegetation, impervious surface and soil respectively. $\mathbf{F}$ is a $3 \times nb$ matrix, each row of $\mathbf{F}$ is a Fisher projection vector. In order to obtain physically

meaningful fraction estimates the solution of model (3) is constrained by: $f_v + f_i + f_s = 1$ and $0 \leq f_v, f_i, f_s \leq 1$.

In practice and as outlined in Fig. 1 the major steps for solving Fisher-transformation based LSMA can be summarized as follows:

- Step1: Projection vector $\mathbf{w} = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is obtained by maximizing J in (1), and $\mathbf{w}$ is termed as the $\mathbf{F}$ matrix with a dimension of $3 \times NB_{OLI}$;
- Step 2: V-I-S endmembers are selected in Fisher space (procedure explained in Section 2.3);
- Step 3: f_v, f_i and f_s are solved from model (3) under restrictions of $f_v + f_i + f_s = 1$ and $0 \leq f_v, f_i, f_s \leq 1$.

2.2.2. Fisher-transformed Multiple Endmember Spectral Mixture Analysis (F-MESMA)

Although Fisher transformation can reduce the endmembers' within-class variability while enhancing between-class variability, the accuracy of modeled proportions could still be improved by implementing Multiple Endmember Spectral Mixture Analysis (MESMA). MESMA is an unmixing technique that addresses within-class endmember variability by allowing the number and type of endmembers to vary on a per pixel basis (Roberts et al., 1998). The algorithm evaluates all possible combinations of available endmembers and selects the combination resulting in the smallest modelling error for each pixel. The "goodness of fit" between the LSMA-modeled vs the observed pixel spectrum is expressed as the root mean square of the residual error (RMSE = $\sqrt{\frac{\sum_{b=1}^{NB} (\varepsilon_b)^2}{NB}}$, where NB is the number of spectral (or Fisher feature) bands and ε_b is the residual error for each band b). A relative RMSE threshold of 0.05 (Franke et al., 2009; Demarchi et al., 2012; Fan and Deng, 2014) was set for comparing the best 2 and 3 endmember models. In particular, the 2-endmember models refer to the LSMA models accounting for combinations of only two endmembers (i.e. "V + I", "V + S", "I + S"), whereas the 3-endmember model refers to the LSMA model including all three endmember classes (i.e. "V + I + S"). This later is necessary since more complex models tend to become more sensitive to collinearity (Chen et al., 2011). Previous work also verified that there is a negative correlation between increasing model complexity and fraction accuracy (Halligan, 2002; Roberts et al., 1998). Therefore, the optimal endmember combination is thus supposed to be as simple as possible while producing acceptable estimation accuracy.

Note that here, in order to obtain V-I-S sub-pixel cover fractions, the MESMA model consisting of vegetation, impervious surfaces, and soil endmembers, is applied not in original reflectance but in Fisher feature space. Water pixels are masked out using a threshold of NDWI before the unmixing process (McFeeters, 1996; Lu et al., 2011).

2.3. Endmember selection

Here we apply a two-step approach to extract image endmember spectra, locally adapted to the study site or city. First, a cluster-based KNN classifier (Jia and Richards, 2005) is implemented to identify a set of V-I-S candidate endmember corner pixels in the Fisher feature space. To initiate the classification procedure, a set of comprehensive V-I-S training samples are required. These can be derived from online spectral libraries (see Section 3.2). These training samples act as the initial classification centers for the KNN classifier. To account for the high spectral variability encountered among impervious materials, the training samples should include a wide range of construction materials as well as vegetation and soil spectra. Note that both the input image and the training endmember samples are transformed into Fisher feature space (see Section 2.2) prior to train the KNN classifier.

Second, in order to select the final set of pure image endmembers, the average spectra of neighborhood pixels close to the KNN cluster centers (in Fisher feature space) are selected as endmember spectra. In

particular, the “neighborhood pixels” are defined as the pixels surrounding cluster centers within a certain distance in the feature space. This pre-set distance is referred to as the feature space radius (FSR) defined by a Euclidean distance from a cluster center to other data point in feature (or reflectance) space. A Euclidean distance array (EDA) between each pixel belonging to the same category and the cluster center is calculated to help determine the FSR. In practice, in this study we determined an empirical optimal FSR value as 0.3 times the standard deviation of EDA (results not shown) (Eq. (4)).

$$EDA = \{x_1^c, x_2^c, \dots, x_N^c\}$$

$$\sigma_{EDA} = \sqrt{\frac{\sum_{i=1}^{N,c} (x_i^c - u_{EDA}^c)^2}{N_c}} \quad (4)$$

$$FSR = n \times \sigma_{EDA}$$

where x_i^c represents the Euclidean distance in Fisher feature space between pixel i and its corresponding classification center c , N_c represents the total number of pixels belonging to class c , u_{EDA}^c represents the Fisher feature vector of classification center c , σ_{EDA} represents the standard deviation of EDA. Note that the total number of selected endmembers is constraint to the number of the corresponding training spectra (or Fisher feature vectors; section 3.2). The image endmember selection procedure is graphically illustrated in Fig. 2.

Note that the above mentioned endmember selection approach is suited to extract endmember bundles for impervious materials and soil but the approach is less suited for extracting representative vegetation spectra. Whereas the spectral variability encountered in impervious surface materials and soil can be comprehensively represented by spectra from online available spectral libraries that are used as input in the initial classification process, this is only to a lesser extent the case for vegetation. The spectral characteristics of vegetation differ not only among different vegetation types but also between individuals of the same species (Cochrane, 2000). On top of that, phenological differences cause major spectral changes among and between plant species (e.g. Somers and Asner, 2012). To capture the full variability in vegetation spectra, an alternative strategy was therefore adopted to extract vegetation endmembers. First, a NDVI threshold (i.e. 0.70) (Myint, 2006)

Fig. 3. Illustration of the protocol to select vegetation image endmembers used to unmix individual image pixels. Two examples are illustrated: (a) to unmix the orange pixel (M1) the vegetation endmember is defined as the Fisher feature vector of pixel V1, being the closest pure vegetation pixel; (b) to unmix the orange pixel (M2) the vegetation endmember is defined as the average Fisher feature vector of V3 and V4. Pixels V2 and V5 are not selected since they are not the vegetation pixels closest to the pixel to be unmixed.

was used for extracting pure vegetation pixels. During the unmixing process each mixed pixel is now unmixed using the geographically nearest pure vegetation pixel's reflectance spectrum or feature vector as the vegetation endmember. If there are multiple pure vegetation pixels selected as endmember candidates, an average reflectance spectrum or feature vector is calculated based on these candidates. If there is no pure vegetation pixel present in the search window, the window will be iteratively extended until a pure vegetation pixel is found. The procedure for selecting vegetation endmembers is illustrated in Fig. 3.

So for each image pixel F-MESMA is evaluating all 2 and 3-endmember models of the V-I-S scheme. As mentioned before, to improve the efficiency of MESMA, the spectral vector (or feature vector) of the vegetation pixel closest to the pixel under consideration was selected as vegetation endmember, meanwhile the selected soil candidate endmembers are averaged into one soil endmember based on the

Fig. 2. Illustration of the procedure for selecting impervious surface image endmembers (Step 2 in Fig. 1). Note that the protocol for selecting soil image endmembers is the same.

Fig. 4. Five experimental sites selected across the cities of Beijing, Brussels, New York City, Las Vegas, and Sydney; Bing maps of the experimental sites (downloadable from <http://www.rivermap.cn/>) used to provide impervious surface fraction reference data; corresponding Landsat 8 OLI images in true colour composition. During the delineation of the study sites in each city we looked for optimal overlap in acquisition time between Bing Maps and Landsat 8 OLI. As a result the study area extent can differ between cities.

assumption that soil spectra within a city are less variable compared to vegetation and impervious surfaces. For the impervious surface class an endmember bundle is extracted according to the above described method and for each pixel all endmember spectra are iteratively evaluated as separate models in the MESMA procedure.

3. Material & methods

3.1. Study sites

Cities across the world are characterized by different land cover compositions. Indeed, the vegetation, soil and impervious materials present in one city are not necessarily the same as in another city. Also, the spatial configuration of land cover patches will differ across cities. The key requirement for this study was thus to test the robustness of the proposed F-MESMA across different cities. Five cities were selected (Fig. 4): Beijing (China), Brussels (Belgium), Las Vegas (United States), New York city (United States), and Sydney (Australia). This suit of cities was selected as they accounted for variation in geographic regions, climate type and population (Schneider and Woodcock, 2008) allowing as such to comprehensively test the effectiveness and robustness of the F-MESMA approach. The tests were performed based on cloud free, atmospherically and geometrically corrected Landsat 8 OLI scenes, acquired in June, July or August (see details in Table 1). Water pixels were masked out using a NDWI threshold (0.05) before conducting spectral mixture analysis (McFeeters, 1996).

3.2. Spectral library data

To initiate the Fisher transformation (Section 2.1) and the KNN classifier (Section 2.3) in total 63 training spectra (including vegetation, impervious surface, soil, and water) were selected from three

publically available online spectral libraries: the ASTER spectral library (<http://speclib.jpl.nasa.gov>; Baldridge et al., 2009), the London Urban Micromet data Archive (LUMA) spectral library (<http://micromet.reading.ac.uk/spectral-library/>; Kotthaus et al., 2014), and the USGS Digital Spectral Library (<https://speclab.cr.usgs.gov/spectral-lib.html>; Clark et al., 2007). More specifically, 17 vegetation spectra were included, ranging from grass (i.e. lawn grass and dry grass), shrub (i.e. sagebrush and saltbush), to tree (i.e. aspen and maple leaves) spectra; 24 impervious surface subclasses of representative construction materials (Kotthaus et al., 2014) were extracted (listed in Fig. 5), including among others concrete, asphalt and tar; 19 soil spectra were selected ranging from dark silty loam, gray silty clay, to reddish sandy; finally, three water training spectra were selected including a spectrum representative for coastal seawater, open ocean water, and river water. Note that the same group of training spectra was applied across all five study sites. For further analysis the spectra were resampled to the spectral resolution of Landsat 8 OLI using the appropriate spectral response function (Trishchenko et al., 2002).

3.3. Reference data

For each study site, summer date Bing-map orthophotos which were acquired in the same year of the Landsat OLI images (spatial resolution of 2.39 m, <http://www.rivermap.cn/>) were used to generate ground reference data of the impervious surface fractions. Training pixels for vegetation, soil and impervious land cover were randomly selected via visual interpretation of the orthophotos. The training samples were used to classify the orthophotos in vegetation, soil and impervious pixels according a supervised classification tool as implemented in the eCognition software (Baatz et al., 2004). An independent set of validation pixels (150 to 200 for each land cover class) were used to assess the accuracy of the supervised classification resulting in a Kappa index

Table 1

Indication of the selected Landsat scene, climate type and population size for the selected study sites.

Study sites	Landsat 8 OLI image	Climate type ^a (Main climate; precipitation; temperature)	Population size(million)
Beijing	July 10, 2017	snow; winter dry; hot summer	21.71
Brussels	June 22, 2016	warm temperate; fully humid; warm summer	1.17
New York city	August 21, 2017	snow; fully humid; warm summer	8.53
Las Vegas	August 15, 2017	arid; desert; hot arid	0.63
Sydney	July 11, 2017	warm temperate; fully humid; warm summer	5.37

^a Kottek et al., 2006.

which was > 0.92 for all study sites. The classification result was subsequently resampled to 30 m spatial resolution to generate impervious surface reference fractions matching the Landsat 8 OLI spatial resolution.

3.4. Analysis

The suitability and effectiveness of deriving impervious surface fraction maps using F-MESMA was evaluated across the five study areas.

The estimated subpixel fractional abundance maps were matched to the reference fractions acquired from the Bing-map orthophotos (Section 3.3.). We evaluated both the absolute fraction estimate error and the goodness of fit between the estimated fractions and the reference fractions. We used three commonly used metrics (Wu and Murray, 2003; Deng and Wu, 2013; Somers and Asner, 2012): the root-mean-square-error (RMSE), the mean-absolute-error (MAE), and the Pearson correlation coefficient (R). RMSE and MAE both quantify the absolute fraction error, while R quantifies the goodness of linear relationship:

$$R = \frac{COV(\bar{Z}, Z)}{\sqrt{D(\bar{Z})} \sqrt{D(Z)}}$$

$$MAE = \left(\sum_{i=1}^n |\bar{Z}_{ki} - Z_{ki}| \right) / n$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\bar{Z}_{ki} - Z_{ki})^2}{n}} \quad (5)$$

where $\bar{Z}_{ki}$ is the modeled fractional value of land-cover component k measured at pixel i , Z_{ki} is reference fractional value, and n is the total number of pixels. Specifically, Pearson correlation coefficient (R)

measures how good the linear relationship the estimated fractions have with reference fractions. Higher R values mean better linear relationship between estimated values and reference values. MAE measures the average difference between estimates and reference fractions. Lower MAE values indicate estimates are relatively consistent with reference fractions. RMSE also measures differences between estimated values and reference values. Low RMSE values are expected when comparing estimated values and reference values. Although both the orthophotos, from which the ground reference fractions are derived, and the Landsat data are geometrically corrected, slight geometric errors when overlapping both data sources is inevitable. In order to reduce the negative influence of geometric errors, a sampling unit of 3×3 pixels (Wu and Murray, 2003) was therefore used, and the data pairs of estimated fractions vs reference fractions were randomly sampled from a low to high fraction intervals, e.g. 0.1–0.2, 0.2–0.3, ..., 0.9–1.0. Note that 20 data pairs were randomly selected per 10% fraction interval ($n = 200$ per city).

Recall that in F-MESMA the selection of endmembers is performed in Fisher feature space after which also the MESMA unmixing is performed in that same Fisher space (i.e. both endmembers and spectral image to be unmixed are projected in Fisher space prior to unmixing). So this implies that in F-MESMA both the process of endmember selection and the process of unmixing is influenced by FDA. To analyze the influence on both processes independently, two alternative MESMA approaches were compared against our proposed F-MESMA: (i) the traditional MESMA approach, T-MESMA, in which endmembers were selected in Reflectance space and the endmembers are applied to MESMA in the form of reflectance vectors (so both endmember selection and unmixing are performed in Reflectance space); and (ii) R-MESMA, in which endmembers are selected in Fisher feature space but the endmembers are applied to MESMA in form of Reflectance vectors

Fig. 5. The selected V-I-S spectra from online available spectral libraries. The number in bracket is the corresponding number of the spectra of the impervious surface subclass.

Fig. 6. Distributions of image pixels, classified impervious surface and soil pixels, and V-I-S classification centers in Fisher feature space (only first two Fisher features are presented here).

(so endmember selection is performed in Fisher feature space but unmixing is performed in the traditional Reflectance space). Recall that in F-MESMA both endmembers are selected and the SMA model is solved in Fisher feature space.

The performance of F-MESMA was further compared to two other widely used SMA approaches designed to minimize spectral variability: (i) NSMA (Wu, 2004) and (ii) SZU (Somers et al., 2010a, 2010b). To allow for a fair comparison between the effectiveness of these SMA approaches, NSMA and SZU were also integrated in MESMA. All three the SMA approaches try to minimizing the spectral variability by either conducting Fisher transformation (F-MESMA), normalizing the reflectance spectra (NSMA), or masking out the bands with high spectral variability (SZU). Additionally, the confusion between bright soil and impervious surface endmembers results in severe collinearity which can significantly reduce the accuracy of unmixing methods. Because of the arid climate, bright soil is prevalent in the Las Vegas urban area but rare in other areas, a group of bright soil samples was therefore selected from Landsat8 OLI image of Las Vegas via visual interpretation for comparing the performance of four unmixing approaches (F-MESMA, T-MESMA, SZU, and NSMA) in dealing with collinearity problem (see Fig. 13).

The effectiveness by which each of the SMA approaches minimizes spectral variability was assessed using the InStability Index (Somers et al., 2009). ISI is defined as the ratio of the within-class variability (Δ_{within}) and the between-class variability ($\Delta_{between}$) (Eq. (5)).

$$ISI_i = \frac{m}{m(m-1)} \sum_{z=1}^{m-1} \sum_{j=z+1}^m \frac{1.96(\sigma_{z,i} + \sigma_{j,i})}{|R_{mean,z,i} - R_{mean,j,i}|} \quad (5)$$

$$ISI = \frac{\sum_{i=1}^N ISI_i}{N}$$

where $R_{mean, z, j}$ and $R_{mean, z, i}$ are the mean reflectance values (or feature values) at wavelength i (or band i) for endmember class z and j , respectively, $\sigma_{z, i}$ and $\sigma_{j, i}$ are the standard deviations at the same wavelengths (or bands) of class z and j , respectively, N represents number of wavelengths (or bands). An ISI value exceeding one indicates within-class variability exceeds the between-class variability. When it comes to

spectral mixture analysis, higher between-class variability and lower within-class variability are preferred (Somers et al., 2011; Barducci and Mecocci, 2005).

The accuracy of impervious surface fraction estimates is significantly influenced by the completeness/comprehensiveness of the spectral library used as input in MESMA (Roberts et al., 1998; Somers et al., 2011). To further evaluate the robustness of F-MESMA its sensitivity against the completeness/comprehensiveness of the spectral input library was tested. Therefore, 100 Landsat image pixels were randomly selected within each of the five study sites and unmixed using the different SMA methods. Subsequently impervious surface endmembers were gradually and randomly removed from the original spectral library used as input in the MESMA models and the sensitivity in terms of fraction estimate accuracy was quantified. The experiment was repeated 30 times and the sensitivity was expressed as the change in maximum RMSE.

The quality of V-I-S-W training spectra determines the result of the Fisher transformation, which directly influences the unmixing process of F-MESMA. To evaluate the effect of the comprehensiveness of the training spectral libraries on the estimate accuracy of F-MESMA, the algorithms was iteratively applied in a Fisher feature space produced by V-I-S-W training spectra with varying sizes. For the example of impervious surfaces, we progressively and randomly selected 1, 2, ..., 24 spectra from the impervious surface spectral library shown in Fig. 5, while fixing the sizes of training spectra of vegetation, soil, and water. Note that each random selection of training spectra was repeated 30 times so that every spectrum could be selected with a high probability. The reflectance images were then transformed into Fisher feature images by Fisher transformation trained by the randomly selected training data set. F-MESMA was applied in this Fisher feature space and the estimate accuracy was evaluated in terms of RMSE.

4. Results

4.1. Endmember selection

In order to automatically extract endmember bundles that are

Fig. 7. Selected V-I-S image endmembers respectively in Fisher feature form (top) and reflectance form (bottom). Red lines represent impervious surface endmembers, green lines represent vegetation endmembers, and brown lines soil endmembers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

representative for the image under consideration, a small set of reference endmember spectra was used for calibrating the KNN driven endmember selection procedure described in Section 2.3. Therefore, as shown in Fig. 5, 24 spectra of impervious surface materials, 19 spectra of soil, 17 spectra of vegetation, and 3 spectra of water - selected from online available spectral libraries - were used as cluster centers in the KNN classifier. As expected, all endmember classes showed a significant within-class variability.

The result of the KNN classification is shown in Fig. 6, illustrating, for each of the five study sites, how V-I-S pixels are distributed in the Fisher feature space (only the first two Fisher features are shown), and which image pixels are identified/classified as impervious surface and soil endmember candidates. Note that the image pixels which are classified in Fisher feature space are defined as “classified pixels” and the V-I-S data points which are scattered by the spectral library data are defined as “classification centers”.

The selected V-I-S Fisher feature image endmembers and their corresponding reflectance spectra are presented in Fig. 7. Endmembers are extracted in Fisher space and both the Fisher feature vector (top panels) and reflectance spectra (bottom panels) are shown. Overall, it is clear that the Fisher transformation indeed decreased the variability within the V-I-S endmember classes while the V-I-S Fisher feature characteristics are easier to distinguish against each other as compared to the original reflectance spectra. Moreover, when comparing the reflectance spectra of the extracted endmembers with the calibration spectra shown in Fig. 5, similar spectral characteristics can be observed with for instance highest spectral variability for the impervious surface class across the whole spectral range. These results indicate that the proposed endmember selection method succeeded in automatically selecting V-I-S image endmembers from the Fisher feature space.

4.2. MESMA in Fisher feature vs reflectance space

The selected image endmembers (Fig. 7) were used to run F-MESMA, R-MESMA and T-MESMA on the five study sites. The resulting impervious surface fractions maps are shown in Fig. 8.

A visual inspection of the maps revealed an overall good spatial match with the ground reference fractions for all MESMA methods; except for some consistent difference between the reference map and all the MESMA fraction maps. This latter is most likely due to differences in the acquisition time of the OLI images and the reference maps causing differences in phenology or land cover changes. Despite the overall good match between all MESMA fraction maps and the reference map, the scatterplots comparing estimated vs reference fractions (Fig. 9)

clearly revealed that R-MESMA (average $R_{R-MESMA} = 0.8229$, $RMSE_{R-MESMA} = 0.1605$, and $MAE_{R-MESMA} = 0.1290$) and T-MESMA (average $R_{T-MESMA} = 0.8002$, $RMSE_{T-MESMA} = 0.1648$, and $MAE_{T-MESMA} = 0.1324$) appear to underestimate high impervious surface fractions in some areas of Brussels and New York City while overestimating low impervious surface fractions in Las Vegas. In contrast results for F-MESMA (average $R_{F-MESMA} = 0.8653$, $RMSE_{F-MESMA} = 0.1346$, and $MAE_{F-MESMA} = 0.1062$) appear to have a higher and more consistent fit across the full stretch of impervious surface fractions and across different cities.

4.3. Comparison with alternative endmember variability reduction techniques

Also, when F-MESMA was compared to SZU and NSMA, the Fisher based MESMA approach performed best (Table 2). Although SZU significantly improved the fraction estimate accuracy compared to T-MESMA ($RMSE_{SZU} = 0.1607$ vs $RMSE_{T-MESMA} = 0.1648$; $R_{SZU} = 0.8133$ vs $R_{T-MESMA} = 0.8002$; $MAE_{SZU} = 0.1284$ vs $MAE_{T-MESMA} = 0.1324$), the approach did not reach the accuracy of F-MESMA ($RMSE_{F-MESMA} = 0.1346$). Also, NSMA, with average RMSE of 0.1749, R of 0.8449 and MAE of 0.1368, outperformed T-MESMA in Brussels and New York City but clearly failed to do so in Las Vegas and Sydney. The relative performance of SZU vs NSMA was not consistent across study sites.

The effectiveness by which the SMA techniques reduced spectral variability was quantified using the ISI index (Fig. 10). The Fisher feature transformation applied in F-MESMA showed the lowest ISI value while the reflectance endmembers used in T-MESMA showed the highest ISI values. It could also be noted that the spectral feature selection, applied in SZU, as well as spectral normalization, applied in NSMA, resulted in significantly reduced ISI values. Surprisingly NSMA obtained an ISI value very close to the one of F-MESMA especially in New York city and Sydney, yet its fraction estimate accuracy was significantly lower than F-MESMA in most cities.

4.4. Robustness against spectral library quality

By progressively and iteratively increasing the number of candidate endmembers to form the input library for SMA we could evaluate the robustness of the different SMA models with regard to the comprehensiveness/completeness of the spectral library (or library quality). Fig. 11 summarizes how the RMSE of the different SMA models changed while progressively increasing candidate endmembers. Note that the

Fig. 8. Impervious surface fraction images for each study site respectively generated by F-MESMA, R-MESMA and T-MESMA. The reference fractions are derived from Bing-map orthophotos. The fractions range from 0 (blue) to 1 (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

RMSE in Fig. 11 is the maximum RMSE observed across 30 iterations. Results clearly showed consistent higher accuracies for F-MESMA when compared to the other methods; yet F-MESMA showed the accuracy was significantly improved from the point that 12 or more impervious surface endmembers were retained. T-MESMA, SZU and NSMA as such showed an overall lower sensitivity to the endmember library size but

even when > 20 endmembers were retained from the original library, T-MESMA, SZU and NSMA all provided much lower impervious fraction estimate accuracies compared to F-MESMA.

Fig. 9. Scatterplot of the MESMA derived vs reference impervious surface fractions for the five study sites. Note that “N” represents the number of samples in each plot, and “R” represents the Pearson's r value.

Table 2

Accuracy metrics for five alternative spectral unmixing approaches. The average, standard deviation (SD) and range of the metrics are shown.

		Beijing	Brussels	New York City	Sydney	Las Vegas	average	SD	range
F-MESMA	RMSE	0.1647	0.1478	0.1146	0.1100	0.1359	0.1346	0.0228	[0.1100–0.1647]
	MAE	0.1319	0.1095	0.0935	0.0833	0.1128	0.1062	0.0187	[0.0833–0.1319]
	R	0.8072	0.8524	0.8808	0.8926	0.8937	0.8653	0.0365	[0.8072–0.8937]
T-MESMA	RMSE	0.1747	0.1635	0.1580	0.1261	0.2019	0.1648	0.0275	[0.1261–0.2019]
	MAE	0.1426	0.1286	0.1257	0.0965	0.1685	0.1324	0.0262	[0.0965–0.1685]
	R	0.7689	0.8659	0.7608	0.8863	0.7193	0.8002	0.0721	[0.7193–0.8863]
R-MESMA	RMSE	0.1961	0.1537	0.1487	0.1218	0.1820	0.1605	0.0292	[0.1218–0.1961]
	MAE	0.1602	0.1180	0.1195	0.0960	0.1513	0.1290	0.0263	[0.0960–0.1602]
	R	0.7713	0.8772	0.7884	0.8993	0.7783	0.8229	0.0605	[0.7713–0.8993]
SZU	RMSE	0.1732	0.1619	0.1580	0.1167	0.1936	0.1607	0.0282	[0.1167–0.1936]
	MAE	0.1406	0.1275	0.1257	0.0902	0.1581	0.1284	0.0250	[0.0902–0.1581]
	R	0.8035	0.8677	0.7608	0.8877	0.7467	0.8133	0.0628	[0.7467–0.8677]
NSMA	RMSE	0.1840	0.1589	0.1389	0.1415	0.2510	0.1749	0.0462	[0.1389–0.2510]
	MAE	0.1487	0.1206	0.1143	0.1135	0.1870	0.1368	0.0315	[0.1135–0.1870]
	R	0.8864	0.8718	0.8070	0.9177	0.7385	0.8449	0.0716	[0.7385–0.9177]

4.5. Robustness against the quality of Fisher transformation

By iteratively changing the sizes of V-I-S-W training spectra we could evaluate the stability of estimate accuracy of F-MESMA to the Fisher transformation quality. Fig. 12 summarizes how the RMSE of the F-MESMA models associated with different sizes of Fisher transformation training spectra changed while progressively adding training spectra. Similar to Fig. 11, the RMSE in Fig. 12 is the maximum RMSE observed across 30 iterations. Note that the RMSE used here is referred to the impervious surface fraction accuracy. Results clearly showed that the estimate accuracy of F-MESMA was negatively affected by the incomplete Fisher transformation training spectra of impervious surface, vegetation, and soil. In contrast, removing or adding water training spectra did not significantly influence the estimate accuracy of F-MESMA. On the other hand, Fig. 12 also indicates that the estimate accuracy of F-MESMA is proportional to the completeness of V-I-S-W training spectra.

5. Discussion and conclusions

The main advantage of F-MESMA is to suppress the spectral variability of endmembers (i.e. high within-class and low between-class variability) by applying a Fisher transformation. The Fisher transformation searches for the projection directions on which the data points of different classes are far from each other and simultaneously where data points of the same class are close to each other (Gao et al., 2012). It is as such not surprising that the endmember library built in Fisher space showed lower within-class and higher between-class variability compared to the endmember set selected in reflectance space (Fig. 7).

Fig. 11. Evolution in RMSE as a function of the endmember library size for four MESMA methods. The x-axis indicates the number of impervious surface endmembers that are remained from the original endmember library.

Especially the expressed spectral variability typical of impervious surface materials, is a well-acknowledged challenge for accurate impervious surface fraction estimation (Weng, 2012). Indeed, the endmembers selected in reflectance space (Fig. 7, bottom panels red lines) clearly illustrated the high spectral variability in impervious surface materials present in our study areas. However, by selecting the

Fig. 10. Bar plots of ISI values of the selected V-I-S image endmembers respectively in form of Fisher features (F-MESMA), Reflectance (T-MESMA), SZU-reflectance (SZU-SMA) and normalized reflectance (NSMA), expressed per study site.

Fig. 12. Evolution in RMSE as a function of the training data size for four MESMA methods. The x-axis indicates the number of training spectra respectively for vegetation, impervious surface, soil, and water.

Fig. 13. Relative relationship between V-I-S endmembers for Las Vegas respectively in three forms: (left) Fisher feature, (center) reflectance and (right) normalized reflectance. The red lines represent impervious surface endmembers, the green lines represent vegetation endmembers, the brown lines represent soil lines, and the gray lines represent bright soil pixels which are manually selected from Landsat 8 OLI image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

endmembers in Fisher space we observed a significant reduction in endmember variability compared to the reflectance endmember sets (also expressed in a significantly lower ISI value, Fig. 10) which consequently also explains the more accurate impervious fraction estimate accuracy for F-MESMA across all study sites. The most expressed improvement in impervious surface fraction estimate accuracy was observed in Las Vegas. This latter can be explained by the prevalence of bright (sandy) soil surfaces in this desert city causing increased collinearity between the soil and impervious surface endmembers as shown in Fig. 13. In Fisher space (Fig. 13, left panel), on the contrary, we clearly observed a lower ambiguity between endmember spectra of both classes explaining the significant improvements in impervious fraction estimate accuracy.

Also, when compared to alternative unmixing techniques specifically designed to reduce the effects of endmember variability (i.e. SZU-SMA and NSMA), we observed a systematic higher (Beijing, Las Vegas, and New York City) or similar (Brussels and Sydney) unmixing accuracy for F-MESMA across all study sites compared to the traditional MESMA approaches. In particular, the estimate accuracies of F-MESMA in Las Vegas and New York City were much higher than for the other methods. The outstanding performance of F-MESMA in Las Vegas is mainly caused by the ability of Fisher transformation to simultaneously minimize the within-class variability while extend the between-class variability (Fig. 10). The bright soil Fisher feature vectors in Fig. 13 which are well differentiated from Fisher transformed impervious surface

endmembers further support this. Besides, previous work has revealed that different impervious surface subclasses produce different estimate error levels when applying unmixing methods (Okujeni et al., 2015; Degerickx et al., 2019). The higher estimate accuracy of F-MESMA in New York City could be partly explained by the fact that some impervious surface subclasses which induce relatively low estimate errors are prevalent in New York City. Because no impervious surface subclass reference data is available in this study, this could not be verified in this study. It is also worthwhile to note that nearly all SMA methods produced the lowest estimate error for Sydney compared to the other study areas. This could be explained by the observation that the endmember set for Sydney represented the lowest ISI values across all study areas (Fig. 10) and thus represent an endmember set with low within-class and high between-class variability, resulting in the lowest estimate errors. Both SZU-SMA and NSMA indeed reduced endmember variability as was evident from the low(er) ISI values compared to the endmember library used in the traditional MESMA method (T-MESMA; Fig. 10). Yet, the unmixing performance of both techniques was sub-optimal and not consistent across study areas. For SZU-SMA this can be explained by the limited spectral resolution of Landsat, only providing spectral reflectance in six wavebands. SZU-SMA, selecting only those spectral bands least sensitive to endmember variability, was specifically designed for hyperspectral unmixing applications and the removal of spectral bands from the already limited number of Landsat bands might compromise the unmixing accuracy as was previously demonstrated by

Liu et al. (2017) and Somers et al. (2010a, 2010b). The relatively poor performance of NSMA is more surprising especially for the Las Vegas case. Although the ISI value for NSMA was indeed similar to that of F-MESMA (i.e. the ISI values of New York City and Sydney) (Fig. 10), meaning that overall the within-class variability was reduced, the left panel of Fig. 13 also reveals that at the same time the similarity between the soil and impervious surface endmembers is increased by the spectral normalization. The confusion between soil and impervious surface fractions is likely the reason for the poor impervious fraction estimate accuracies obtained by NSMA.

F-MESMA also showed to be relatively robust with respect to the completeness/comprehensiveness of the endmember library used for calibrating the model. When > 50% of the endmembers were randomly removed from the original endmember library the unmixing performance of F-MESMA decreased, yet remained better than all other tested SMA models (Fig. 11). This could be explained by its ability of suppressing within-class variability. In other words, the variability within a specific endmember class can be addressed with fewer endmembers, since in Fisher space the within-class variability is strongly reduced and already with a few endmembers the spectral variability of the population can be modeled. Therefore, SZU and T-MESMA also rely more on the comprehensiveness of the spectral library than F-MESMA resulting in a consistent lower level of fraction estimate accuracy. Due to high collinearity between soil and impervious surfaces, NSMA showed the worst estimations no matter how many endmembers were excluded, and its low but stable performance independent of the input library can again be attributed to its ability to significantly reduce within-class variability.

Besides, the performance of F-MESMA appears to be sensitive to the varying sizes of training spectra of vegetation, impervious surface, and soil (Fig. 12), but tends to be good when the training spectra is becoming comprehensive. It indicates that the result of Fisher transformation trained by a limited or incomplete V-I-S training spectral library may not be reliable for further analysis and attention should have such be paid to select a proper training library. In contrast, because of the high within-class spectral similarity of water (Fig. 5), the estimate accuracy of F-MESMA did not respond significantly towards changes in water training spectra.

Overall, we can conclude that the automated F-MESMA model allows to produce reliable impervious surface fraction estimates across a variety of urban ecosystems and that F-MESMA produces more stable results compared to established SMA approaches. In particular in urban environments with expressed spectral variability among and spectral similarity across material classes F-MESMA seems to outperform more traditional SMA approaches. Indeed, especially F-MESMA's ability of simultaneously minimizing within-class variability and maximizing between-class similarity makes it a robust alternative to existing SMA approaches. An important feature of our automated F-MESMA workflow is its ability to automatically select site-specific image endmembers, in other words the endmember sets used to unmix image scenes are automatically calibrated for the scene under investigation. As input for the workflow only a handful of material spectra derived from online available spectral libraries are used. This latter is a major asset when operational, large-scale urban mapping applications are considered. However, given the limited sample size in our study (only five cities), future efforts should focus on testing the F-MESMA approach on a larger population of cities across the world.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. 41871224), the Fund for Creative Research Groups of National Natural Science Foundation of China (No.41621061) and China Scholarship Council (No. 201706040077).

References

- Baatz, M., Benz, U., Dehghani, S., Heynen, M., Höltje, A., Hofmann, P., Willhauck, G., 2004. eCognition user guide. Definiens Imaging GmbH, Munich, Germany.
- Baldrige, A.M., Hook, S.J., Grove, C.I., Rivera, G., 2009. The ASTER spectral library version 2.0. *Remote Sens. Environ.* 113 (4), 711–715.
- Barducci, A., & Mecocci, A. (2005). Theoretical and experimental assessment of noise effects on least-squares spectral unmixing of hyperspectral images. *Opt. Eng.*, 44 Art. N° 087008.
- Boardman, J.W., Kruse, F.A., Green, R.O., 1995. Mapping target signatures via partial unmixing of AVIRIS data. In: Summaries, Fifth JPL Airborne Earth Science Workshop. JPL Publication 95-1. pp. 23–26.
- Burazerović, D., Heylen, R., Geens, B., Sterckx, S., Scheunders, P., 2013. Detecting the adjacency effect in hyperspectral imagery with spectral unmixing techniques. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 6 (3), 1070–1078.
- Chen, X., Chen, J., Jia, X., Somers, B., Wu, J., Coppin, P., 2011. A quantitative analysis of virtual endmembers' increased impact on the collinearity effect in spectral unmixing. *IEEE Trans. Geosci. Remote Sens.* 49 (8), 2945–2956.
- Clark, R.N., Swayze, G.A., Wise, R., Livo, K.E., Hoefen, T., Kokaly, R.F., Sutley, S.J., 2007. USGS Digital Spectral Library splib06a. US Geological Survey, Digital Data Series. 2007. pp. 231.
- Cochrane, M.A., 2000. Using vegetation reflectance variability for species level classification of hyperspectral data. *Int. J. Remote Sens.* 21 (10), 2075–2087.
- Conway, K.W., Barrie, J.V., Hill, P.R., Austin, W.C., Picard, K., 2007. Mapping sensitive benthic habitats in the Strait of Georgia, coastal British Columbia: deep-water sponge and coral reefs. *Geol. Surv. Can.* 2, 1–6.
- Coseo, P., Larsen, L., 2014. How factors of land use/land cover, building configuration, and adjacent heat sources and sinks explain urban heat islands in Chicago. *Landsc. Urban Plan.* 125, 117–129.
- Degerickx, J., Okujeni, A., Iordache, M.D., Hermy, M., van der Linden, S., Somers, B., 2017. A novel spectral library pruning technique for spectral unmixing of urban land cover. *Remote Sens.* 9 (6), 565.
- Degerickx, J., Roberts, D.A., Somers, B., 2019. Enhancing the performance of Multiple Endmember Spectral Mixture Analysis (MESMA) for urban land cover mapping using airborne lidar data and band selection. *Remote Sens. Environ.* 221, 260–273.
- Demarchi, L., Canters, F., Chan, J.C.W., Van de Voorde, T., 2012. Multiple endmember unmixing of CHRIS/Proba imagery for mapping impervious surfaces in urban and suburban environments. *IEEE Trans. Geosci. Remote Sens.* 50 (9), 3409–3424.
- Deng, C., Wu, C., 2013. A spatially adaptive spectral mixture analysis for mapping sub-pixel urban impervious surface distribution. *Remote Sens. Environ.* 133, 62–70.
- Dobigeon, N., Tits, L., Somers, B., Altmann, Y., Coppin, P., 2014. A comparison of non-linear mixing models for vegetated areas using simulated and real hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 1869–1878.
- Du, H., Wang, D., Wang, Y., Zhao, X., Qin, F., Jiang, H., Cai, Y., 2016. Influences of land cover types, meteorological conditions, anthropogenic heat and urban area on surface urban heat island in the Yangtze River Delta Urban Agglomeration. *Sci. Total Environ.* 571, 461–470.
- Elmore, A.J., Mustard, J.F., Manning, S.J., Lobell, D.B., 2000. Quantifying vegetation change in semiarid environments: precision and accuracy of spectral mixture analysis and the normalized difference vegetation index. *Remote Sens. Environ.* 73 (1), 87–102.
- Fan, F., Deng, Y., 2014. Enhancing endmember selection in multiple endmember spectral mixture analysis (MESMA) for urban impervious surface area mapping using spectral angle and spectral distance parameters. *Int. J. Appl. Earth Obs. Geoinf.* 33, 290–301.
- Franke, J., Roberts, D.A., Halligan, K., Menz, G., 2009. Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sens. Environ.* 113 (8), 1712–1723.
- Fukunaga, K., 2013. Introduction to Statistical Pattern Recognition. Elsevier.
- Gao, Q., Liu, J., Zhang, H., Hou, J., Yang, X., 2012. Enhanced Fisher discriminant criterion for image recognition. *Pattern Recogn.* 45 (10), 3717–3724.
- Halligan, K.Q., 2002. Multiple Endmember Spectral Mixture Analysis of Vegetation in the Northeast Corner of Yellowstone National Park. Doctoral dissertation. University of California, Santa Barbara.
- Herold, M., Roberts, D., 2005. Spectral characteristics of asphalt road aging and deterioration: implications for remote-sensing applications. *Appl. Opt.* 44 (20), 4327–4334.
- Jia, X., Richards, J.A., 2005. Fast k-NN classification using the cluster-space approach. *IEEE Geosci. Remote Sens. Lett.* 2 (2), 225–228.
- Jing, J., Wang, B., Zhang, L., 2010. A novel approach based on Fisher discriminant null space for decomposition of mixed pixels in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* 7 (4), 699–703.
- Keshava, N., Mustard, J.F., 2002. Spectral unmixing. *IEEE Signal Process.* 19, 44–57.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* 15 (3), 259–263.
- Kothaus, S., Smith, T.E., Wooster, M.J., Grimmer, C.S.B., 2014. Derivation of an urban materials spectral library through emittance and reflectance spectroscopy. *ISPRS J. Photogramm. Remote Sens.* 94, 194–212.
- Liu, M., Yang, W., Chen, J., Chen, X., 2017. An orthogonal Fisher transformation-based unmixing method toward estimating fractional vegetation cover in semiarid areas. *IEEE Geosci. Remote Sens. Lett.* 14 (3), 449–453.
- Lu, D., Weng, Q., 2006. Use of impervious surface in urban land use classification. *Remote Sens. Environ.* 102 (1–2), 146–160.
- Lu, D., Moran, E., Hetrick, S., 2011. Detection of impervious surface change with multi-temporal Landsat images in an urban-rural frontier. *ISPRS J. Photogramm. Remote Sens.* 66 (3), 298–306.

- McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17 (7), 1425–1432.
- Myint, S.W., 2006. Urban vegetation mapping using sub-pixel analysis and expert system rules: a critical approach. *Int. J. Remote Sens.* 27 (13), 2645–2665.
- Nascimento, J.M., Dias, J.M., 2005. Vertex component analysis: a fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* 43 (4), 898–910.
- Okada, T., Tomita, S., 1985. An optimal orthonormal system for discriminant analysis. *Pattern Recogn.* 18 (2), 139–144.
- Okujeni, A., van der Linden, S., Hostert, P., 2015. Extending the vegetation–impervious–soil model using simulated EnMAP data and machine learning. *Remote Sens. Environ.* 158, 69–80.
- Phinn, S., Stanford, M., Scarth, P., Murray, A.T., Shyy, T., 2002. Monitoring the composition and form of urban environments based on the vegetation–impervious surface–soil (VIS) model by sub-pixel analysis techniques. *Int. J. Remote Sens.* 23 (20), 4131–4153.
- Quintano, C., Fernández-Manso, A., Roberts, D.A., 2013. Multiple Endmember Spectral Mixture Analysis (MESMA) to map burn severity levels from Landsat images in Mediterranean countries. *Remote Sens. Environ.* 136, 76–88.
- Ridd, M.K., 1995. Exploring a VIS (vegetation–impervious surface–soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *Int. J. Remote Sens.* 16 (12), 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 Sens. Environ.* 65 (3), 267–279.
- Sawaya, K.E., Olmanson, L.G., Heinert, N.J., Brezonik, P.L., Bauer, M.E., 2003. Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery. *Remote Sens. Environ.* 88 (1–2), 144–156.
- Schneider, A., Woodcock, C.E., 2008. Compact, dispersed, fragmented, extensive? A comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. *Urban Stud.* 45 (3), 659–692.
- Settle, J., 2006. On the effect of variable endmember spectra in the linear mixture model. *IEEE Trans. Geosci. Remote Sens.* 44 (2), 389–396.
- Somers, B., Asner, G.P., 2012. Hyperspectral time series analysis of native and invasive species in Hawaiian rainforests. *Remote Sens.* 4, 2510–2529.
- Somers, B., Cools, K., Delalieux, S., Stuckens, J., Van der Zande, D., Verstraeten, W.W., Coppin, P., 2009. Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards. *Remote Sens. Environ.* 113 (6), 1183–1193.
- Somers, B., Delalieux, S., Verstraeten, W.W., Van Aardt, J.A.N., Albrigo, G.L., Coppin, P., 2010a. An automated waveband selection technique for optimized hyperspectral mixture analysis. *Int. J. Remote Sens.* 31 (20), 5549–5568.
- Somers, B., Verbesselt, J., Ampe, E.M., Sims, N., Verstraeten, W.W., Coppin, P., 2010b. Spectral mixture analysis to monitor defoliation in mixed aged Eucalyptus globules Labill plantations in southern Australia using Landsat 5TM and EO-1 Hyperion data. *Int. J. Appl. Earth Obs. Geoinf.* 12, 270–277.
- Somers, B., Asner, G.P., Tits, L., Coppin, P., 2011. Endmember variability in spectral mixture analysis: a review. *Remote Sens. Environ.* 115 (7), 1603–1616.
- Somers, B., Tits, L., Coppin, P., 2014. Quantifying nonlinear spectral mixing in vegetated areas: computer simulation model validation and first results. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 1956–1965.
- Sugiyama, M., 2007. Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis. *J. Mach. Learn. Res.* 8 (May), 1027–1061.
- Trishchenko, A.P., Cihlar, J., Li, Z., 2002. Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. *Remote Sens. Environ.* 81 (1), 1–18.
- Van der Meer, F.D., Jia, X., 2012. Collinearity and orthogonality of endmembers in linear spectral unmixing. *Int. J. Appl. Earth Obs. Geoinf.* 18, 491–503.
- Ward, D., Phinn, S.R., Murray, A.T., 2000. Monitoring growth in rapidly urbanizing areas using remotely sensed data. *Prof. Geogr.* 52 (3), 371–386.
- Ward, H.C., Tan, Y.S., Gabey, A.M., Kotthaus, S., Grimmond, C.S., 2018. Impact of temporal resolution of precipitation forcing data on modelled urban-atmosphere exchanges and surface conditions. *Int. J. Climatol.* 38 (2), 649–662.
- Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. *Remote Sens. Environ.* 117, 34–49.
- Winter, M.E., 1999. N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In: *Imaging Spectrometry V*. 3753. International Society for Optics and Photonics, pp. 266–276 October.
- Wu, C., 2004. Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sens. Environ.* 93 (4), 480–492.
- Wu, C., Murray, A.T., 2003. Estimating impervious surface distribution by spectral mixture analysis. *Remote Sens. Environ.* 84 (4), 493–505.
- Yang, F., Matsushita, B., Fukushima, T., 2010. A pre-screened and normalized multiple endmember spectral mixture analysis for mapping impervious surface area in Lake Kasumigaura Basin, Japan. *ISPRS J. Photogramm. Remote Sens.* 65 (5), 479–490.
- Zare, A., Ho, K.C., 2014. Endmember variability in hyperspectral analysis: addressing spectral variability during spectral unmixing. *IEEE Signal Process. Mag.* 31 (1), 95–104.
- Zhang, J., Rivard, B., Sanchez-Azofeifa, A., 2004. Derivative spectral unmixing of hyperspectral data applied to mixtures of lichen and rock. *IEEE Trans. Geosci. Remote Sens.* 42 (9), 1934–1940.