Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Remote Sensing of Environment 106 (2007) 253 – 267
Remote Sensing of Environment
www.elsevier.com/locate/rse

Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil

Rebecca L. Powell [a,*], Dar A. Roberts [a], Philip E. Dennison [b], Laura L. Hess [c]

[a] Department of Geography, University of California-Santa Barbara, Santa Barbara, CA 93106, USA
[b] Department of Geography, University of Utah, 260 S. Central Campus Dr., Room 270, Salt Lake City, UT 84112, USA
[c] Intsitute for Computational Earth System Science, University of California-Santa Barbara, Santa Barbara, CA 93106, USA

Received 3 June 2005; received in revised form 1 September 2006; accepted 3 September 2006

Abstract

The spatial and spectral variability of urban environments present fundamental challenges to deriving accurate remote sensing products for urban areas. Multiple endmember spectral mixture analysis (MESMA) is a technique that potentially addresses both challenges. MESMA models spectra as the linear sum of spectrally pure endmembers that vary on a per-pixel basis. Spatial variability is addressed by mapping sub-pixel components of land cover as a combination of endmembers. Spectral variability is addressed by allowing the number and type of endmembers to vary from pixel to pixel. This paper presents an application of MESMA to map the physical components of urban land cover for the city of Manaus, Brazil, using Landsat Enhanced Thematic Mapper (ETM+) imagery.

We present a methodology to build a regionally specific spectral library of urban materials based on generalized categories of urban land-cover components: vegetation, impervious surfaces, soil, and water. Using this library, we applied MESMA to generate a total of 1137 two-, three-, and four-endmember models for each pixel; the model with the lowest root-mean-squared (RMS) error and lowest complexity was selected on a per-pixel basis. Almost 97% of the pixels within the image were modeled within the 2.5% RMS error constraint. The modeled fractions were used to generate continuous maps of the per-pixel abundance of each generalized land-cover component. We provide an example to demonstrate that land-cover components have the potential to characterize trajectories of physical landscape change as urban neighborhoods develop through time. Accuracy of land-cover fractions was assessed using high-resolution, geocoded images mosaicked from digital aerial videography. Modeled vegetation and impervious fractions corresponded well with the reference fractions. Modeled soil fractions did not correspond as closely with the reference fractions, in part due to limitations of the reference data. This work demonstrates the potential of moderate-resolution, multispectral imagery to map and monitor the evolution of the physical urban environment.

Keywords: Spectral mixture analysis; Urban land cover; V–I–S model; Landsat

1. Introduction

Urban areas are currently among the most rapidly changing types of land cover on the planet. Though covering only a few percent of the global land surface, cities are the loci of human population and activity and are thereby sites of significant natural resource transformation (Lambin et al., 2001). Remote sensing imagery can provide a timely and synoptic view of urban land cover, as well as a means to monitor change in urbanizing landscapes and to compare urban environments globally. However, deriving accurate, quantitative measures from remote sensing imagery over urban areas remains a fundamental research challenge due to the great spectral and spatial variability of the materials that compose urban land cover (Forster, 1985; Lu & Weng, 2004; Xian & Crane, 2005). The highly heterogeneous nature of urban surface materials is problematic at multiple spatial scales, resulting in a high percentage of mixed pixels in moderate resolution imagery and even limiting the utility of high spatial resolution imagery (Myint et al., 2004; Small, 2005).

* Corresponding author. Current address: Department of Geography, University of Denver, 2050 E. Iliff Ave., Denver, CO 80208, USA. Tel.: +1 303 871 2667; fax: +1 303 871 2201.
E-mail address: rpowell8@du.edu (R.L. Powell).

0034-4257/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2006.09.005

The most common approach to characterizing urban environments from remote sensing imagery is land-use classification, i.e. assigning all pixels in the image to mutually exclusive classes such as residential, industrial, recreational, etc. (Carlson & Sanchez-Azofeifa, 1999; Forster, 1985; Ridd, 1995). However, this approach can be problematic for several reasons. First, most urban land-use classes are not spectrally distinct, resulting in considerable confusion between classes (Ridd, 1995; Small, 2005). Second, the physical composition of land-use classes may vary dramatically from region to region due to different building materials and different construction practices, and therefore cross-regional comparisons between urban areas are limited (Small, 2005). In rapidly growing cities, particularly in the developing world, multiple forms of land use may occur within the same geographic space, limiting the usefulness of traditional land-use categories. Finally, urban land-use classes have little correlation with physical parameters–i.e. they are social constructs imposed on the physical urban environment– and partitioning the urban landscape into such classes simplifies the heterogeneity that is a key characteristic of the urban environment (Ridd, 1995).

In contrast, mapping the urban environment in terms of its physical components preserves the heterogeneity of urban land cover better than traditional land-use classification (Clapham, 2003; Ji & Jensen, 1999), characterizes urban land cover independent from analyst-imposed definitions (Jensen, 1983; Ridd, 1995), and more accurately captures change through time (Ji & Jensen, 1999; Rashed et al., 2005). Ridd (1995) proposed that any urban environment can be conceptualized in terms of three primary physical components: vegetation, impervious surfaces, and soil (V–I–S components), in addition to water. Lu and Weng (2004) provide an overview of current research that has applied the V–I–S model to characterize urban environments. However, with few exceptions, these studies have used V–I–S components as input to urban land-use classification, transforming the continuous variables captured by the V–I–S characterization into mutually exclusive land-use classes (e.g. Lu & Weng, 2004). Exceptions to this include Wu and Murray (2003) and Wu (2004), who generated maps of continuous values representing the fraction of impervious surface for the metropolitan area of Columbus, OH. Another exception is the work by Rashed et al. (2003), who mapped the physical gradients of the three V–I–S components in the Los Angeles area and suggested links between the socio-economic characteristics of neighborhoods and their corresponding V–I–S composition. Similarly, Rashed et al. (2005) mapped the V–I–S components for the greater urban area in Cairo, Egypt, and identified change between the physical components aggregated to census tracts. They proposed that monitoring the V–I–S components between dates provided a more accurate estimate of urban land-cover change, as much change occurred within land-use classes.

The goal of this paper is to demonstrate a methodology to map the variation of V–I–S components in an urban environment using moderate-resolution remote sensing imagery. Spatial variability is addressed by mapping the sub-pixel components of land cover using spectral mixture analysis (SMA), which models each pixel as a linear sum of spectrally 'pure' endmembers (e.g. Adams et al., 1993). Traditional SMA uses a fixed number of endmembers to map the entire landscape. Previous studies that applied SMA to map urban environments have noted that the limited number of allowed endmembers cannot adequately capture the high spectral heterogeneity of urban materials (Lu & Weng, 2004; Rashed et al., 2005; Song, 2005; Wu, 2004). In this study, therefore, spectral variability is addressed by applying multiple endmember spectral mixture analysis (MESMA), which allows the number and type of endmembers to vary on a per-pixel basis (Roberts et al., 1998b).

The potential of using MESMA to map the V–I–S components of an urban area was demonstrated by Rashed et al. (2003). We expand their work by developing a methodology to build a regionally specific spectral endmember library from a large collection of reference and image endmembers. Additionally, we utilize aerial videography collected over the study area to assess the accuracy of maps of fractional V–I–S cover, and investigate the question of the appropriate scale of analysis for comparative or change detection studies. The product of our work is a set of maps representing the per-pixel fractional cover of each component—vegetation, impervious surfaces, and soil. These maps are locally specific, capturing the spectral variability that is distinct to the region, yet globally representative of urban land cover, allowing comparison of urban composition across regions and through time. We test our methodology on case study, a rapidly growing city located in the Brazilian Amazon.

2. Methods

2.1. Study site and data

The study area is the region immediately surrounding the city of Manaus, in the state of Amazonas, Brazil (Fig. 1). Manaus is located on the banks of the Rio Negro, approximately 18 km north of the confluence with the Rio Solimões, forming the main stem of the Amazon River. Though located approximately 1200 km up-stream from the mouth of the Amazon, Manaus is currently the most populated city in the Amazon region, with a population of almost 1.4 million recorded in the 2000 national census (IBGE, 2000). Founded as a Portuguese settlement in 1669, Manaus first grew to national prominence during the rubber boom between 1870 and 1920. However, the city's real growth as an urban and industrial center has occurred since 1967, when the federal government of Brazil declared the city and port of Manaus a Free Trade Zone (Zona Franca de Manaus), leading to a dramatic increase in public and commercial infrastructure, manufacturing and industrial facilities, and corresponding population growth (Browder & Godfrey, 1997; Silva-Forsberg, 1999).

A Landsat Enhanced Thematic Mapper (ETM+) scene (path 231, row 62) acquired on August 11, 2001, was used in this study. A 750-km^2 region centered over Manaus was subset from the Landsat image, and the georegistration was refined by co-registering the subset to a 1:100,000 topographic maps published by the Brazilian Institute of Geography and Statistics

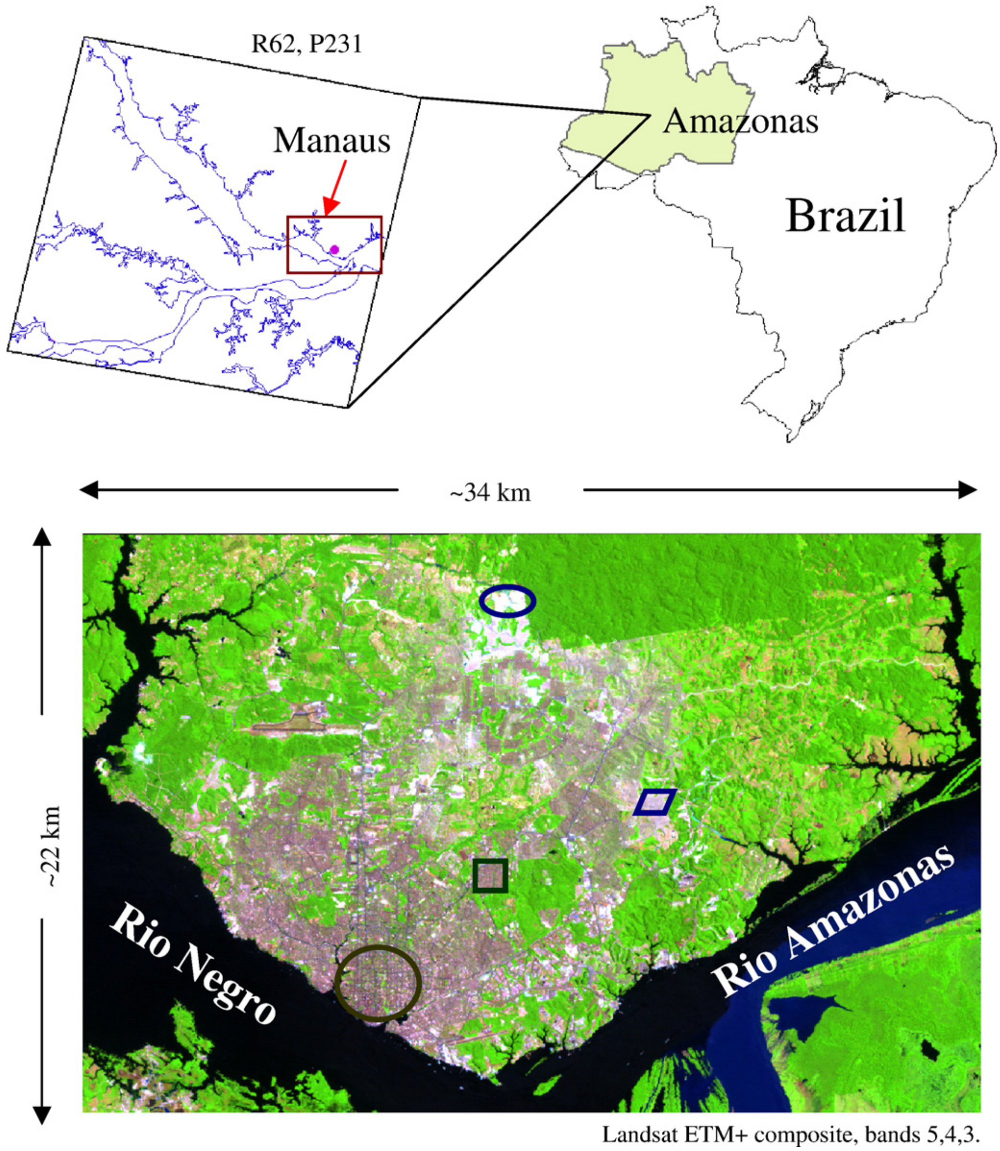

Fig. 1. Study area. Regions highlighted on the Landsat sub-scene are discussed in the Results section.

(IBGE). Twenty-three georegistration points were located, and nearest-neighbor resampling was applied to a first-degree polynomial fit. The average root mean squared error (RMSE) for the registration points was 0.16 pixels. Reference data were produced from digital aerial videography collected over Manaus on June 12, 1999, as part of the Validation Overflights for Amazon Mosaics (Hess et al., 2002). The flight lines were designed to provide almost complete and continuous coverage of the city. Wide-angle and zoom videography were collected simultaneously; however, only the wide-angle videography was used for this study. Average swath width was approximately 1 km; average ground instantaneous field of view (GIFOV) was 1.5 m. Global positioning system (GPS) location and time code, aircraft altitude, and aircraft height data were encoded on each videography frame and used to automatically generate geocoded videography mosaics, with an estimated absolute geolocation error of 5–10 m along the center third of the videography.

2.2. Multiple endmember spectral mixture analysis

Spectral mixture analysis (SMA) is based on the assumption that the reflectance P' measured at pixel i can be modeled as the linear sum of N endmembers, or spectrally 'pure' materials, weighted by the fraction f_{ki} of each endmember within the field of view of pixel i (e.g. Adams et al., 1993; Roberts et al., 1998a). That is, for a given wavelength, λ:

$$P'_{i\lambda} = \sum_{k=1}^{N} f_{ki} * P_{k\lambda} + e_{i\lambda}, \tag{1}$$

where $e_{i\lambda}$ is a residual term indicating the disagreement between the measured and modeled spectra. The modeled fractions are typically constrained by the following:

$$\sum_{k=1}^{N} f_{ki} = 1. \tag{2}$$

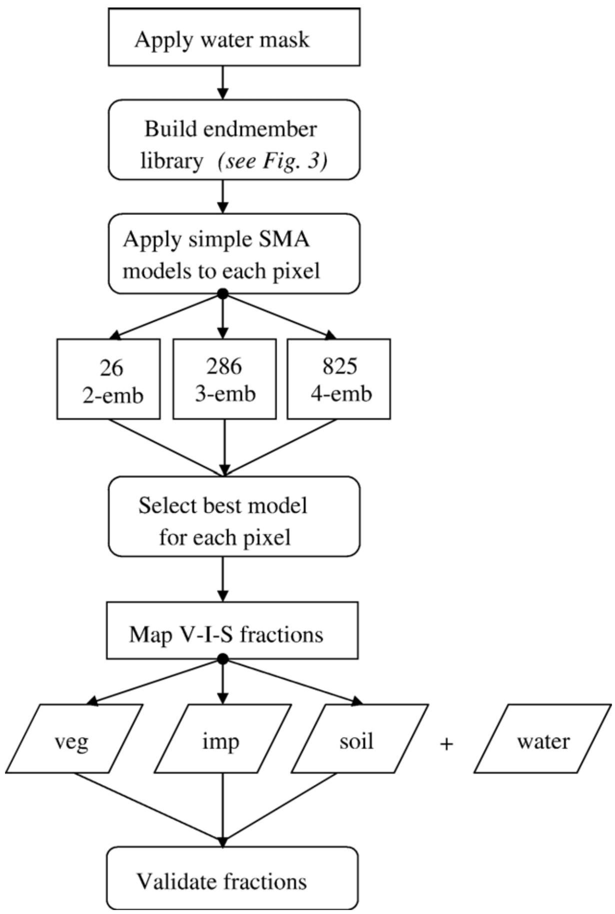

Fig. 2. Methods flowcharts.

Model fit is assessed by calculating the root mean squared error (RMSE) of the residuals for each pixel across all bands (Adams et al., 1993; Roberts et al., 1998a), given by:

$$\mathrm{RMSE}_i = \left(\sum_{k=1}^{\lambda} \left(e_{i\lambda}\right)^2 / N \right)^{1/2}. \tag{3}$$

Endmember spectra can be collected in the field or lab (reference endmembers) or extracted from an image (image endmembers). Constraints for selecting appropriate models for each pixel can be specified in terms of the range of endmember fractions, residuals for each wavelength, and the RMSE (Roberts et al., 1998b).

In a standard application of SMA, a fixed number of representative endmembers–usually between two and five–are selected and the entire image is modeled in terms of those spectral components. However, this procedure is limited because the selected endmember spectra may not effectively model all elements in the image, or a pixel may be modeled by endmembers that do not correspond to the materials located in its field of view. Both cases result in decreased accuracy of the estimated fractions (Sabol et al., 1992). These limitations of simple SMA are particularly problematic in urban environments, which exhibit high degrees of spectral heterogeneity on fine spatial scales. A technique that addresses these limitations is multiple endmember spectral mixture analysis (MESMA), which allows the number and type of endmembers to vary on a per pixel basis (Roberts et al., 1998b).

Table 1
Allowed models by generalized material classes

Two-endmember (26)	Three-endmember (286)	Four-endmember (825)
npv+shade	npv+veg+shade	npv+imp+imp+shade
veg+shade	npv+soil+shade	veg+imp+imp+shade
soil+shade	npv+imp+shade	soil+imp+imp+shade
imp+shade	veg+soil+shade	
	veg+imp+shade	
	soil+imp+shade	
	imp+imp+shade	

Numbers in parentheses indicate the total number of models generated for all permutations of endmembers included in the final library.

In this case study, MESMA was implemented following the procedure illustrated in Fig. 2 and detailed in the sections that follow. First, an endmember library was constructed from candidate image and reference endmembers. Next, a series of simple SMA models–based on all combinations of library endmembers presented in Table 1–was applied to every pixel in the image, and the 'best-fit' model was selected for each pixel. The models were generalized into the land-cover components of interest (i.e. vegetation, impervious surface, soil), and an image of fractional coverage per pixel was generated for each component. To assess how well these fraction images represented the actual fraction of each land-cover component on the ground, fractional cover derived from MESMA was compared to fractional cover measured from classified aerial videography. Agreement between modeled fraction cover and reference fraction cover was used to refine the combinations of endmembers allowed for SMA modeling, as well as to select to most appropriate constraints for MESMA application to Landsat data.

2.3. Library construction

Essential to any application of SMA is the careful selection of endmembers (Dennison & Roberts, 2003; Tompkins et al., 1997), as endmembers must accurately and comprehensively

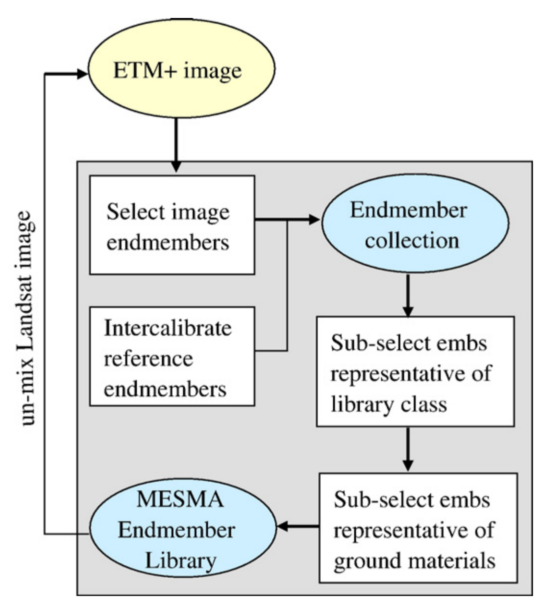

Fig. 3. Summary of MESMA spectral library construction.

represent ground materials if the output of SMA is to be physically meaningful. Endmember selection for standard SMA commonly relies on identifying an extreme spectrum to represent each material of interest (e.g. Adams et al., 1995; Boardman et al., 1995; Smith et al., 1985). In contrast, endmember selection for MESMA focuses on identifying a set of spectra that represents the spectral variation for each material in the scene (Okin et al., 2001; Painter et al., 1998; Roberts et al., 1998b). The particular challenge in building a spectral library for use in MESMA is twofold. First, for each class of materials, the library should include a sufficient number of spectra to adequately represent the spectral variation of the material on the ground. Second, as the total number of endmembers–and therefore potential models–increases, the computational efficiency of the MESMA exponentially decreases (Halligan, 2002), and therefore the endmember library should remain sufficiently small so that the application of MESMA remains computationally viable.

To attain these goals, we followed a three-step process to construct the MESMA library used in this application (Fig. 3). First, we compiled a collection of possible endmembers, including image and reference endmembers. Based on previous applications of SMA in the region (e.g. Roberts et al., 1998a), we knew that to adequately capture the vegetation class we need to include spectra of green vegetation and non-photosynthetic vegetation (NPV), i.e. dry or senesced vegetation. Therefore, we organized the collection of endmembers into four groups: green vegetation, NPV, impervious surfaces, and soil. Second, for each class of materials, a subset of endmembers was selected that best represented the class in the library collection. Finally, from this subset, endmembers were selected that best represented materials on the ground. Each step is discussed in more detail below.

2.3.1. Endmember collection

The first step in library construction consisted of collecting candidate endmembers for each group of interest. To include the broadest range of possible endmembers, we employed three methods of spectral collection: endmembers derived from the 2001 Landsat ETM+ sub-scene, reference endmembers collected in the field, and reference endmembers collected in a laboratory. Reference endmembers were collected with a variety of instruments, including a Beckman DK2a (Beckman Instruments, Fullerton, CA, USA), a Cary-5E (Varian, Sunnyvale, CA, USA), and an ASD-FR spectrometer (Analytical Spectral Devices, Boulder, CO, USA). Water was obviously an important component of the scene, but it was treated separately from the other materials of interest. Dark pixels are highly degenerate (i.e. they can be modeled successfully by a large shade fraction and a small bright fraction of almost any material), and therefore they cannot be modeled very accurately using spectral mixture analysis. A water/dark-pixel mask was generated by applying a threshold to Landsat band 7, and those pixels were removed from further evaluation.

Image endmembers were collected by applying a pixel purity index (PPI) to the 2001 sub-scene. The PPI algorithm identifies extreme pixels in multidimensional space by projecting the spectrum of each pixel on multiple vectors randomly oriented in feature space and recording the number of times the pixel is found to be extreme (Boardman et al., 1995). Pixels identified as extreme were visually inspected, and the spectra of those which could be classified with confidence as one of the materials of interest were collected. In general, all edge pixels–particularly between land and water–were eliminated from further consideration.

Reference endmembers for materials in the NPV and soil classes were collected in the field in 1991, along with calibration targets that could be located on a 1991 Landsat Thematic Mapper (TM) scene. Reflectance calibration targets from 1991 were used to retrieve apparent reflectance from the 1991 Landsat data using an empirical line calibration (Conel, 1990; Roberts et al., 1985). The 1991 reflectance image was, in turn, used as a reference to intercalibrate the 2001 image using a relative radiometric calibration approach, in which encoded radiance was regressed against reflectance using the mean spectra of invariant ground targets (Furby & Campbell, 2001). This technique is based on two assumptions: first, that invariant surfaces which include a range of brightness values can be found between the two dates, and second, that there is a uniform atmosphere over each scene (Souaza et al., 2005). To reduce data volume, all processing was actually accomplished in encoded radiance space by inverting equations for reflectance to convert reflectance to DN.

Vegetation spectra collected in the field do not easily 'scale up' to match spectra measured by a Landsat sensor because field spectra are collected on the scale of individual leaves or branches, and cannot adequately capture the multiple-scattering environment of a canopy. A Landsat pixel, however, captures the reflectance of an entire canopy (or multiple canopies); as a result, the spectrum measured at the sensor is the result of both direct reflection and multiple scattering of incoming solar radiation. The ideal scale for collecting plant spectra, therefore, would be a scale roughly comparable to the sensor scale; however, in this study, field collection of vegetation spectra at the appropriate scale was not possible. To include vegetation spectra of known materials, therefore, we used spectral proxies, derived either from leaf stacks or branches measured in North America, or from canopies measured by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS), to augment our library of vegetation image endmembers. In both cases, the proxy spectra were collected from plant species with the same physiognomy. As the dominant controls on spectra measured by Landsat TM are biophysical, i.e. related to the tree and leaf structure, not species specific, we are confident that such spectra would be representative of the dominant plant spectra found in the Amazon.

Reference endmember spectra for the impervious surface class were collected in the laboratory by importing samples of commonly used roof materials from a construction supply store in the Amazon. As no field spectrometer was available for the collection of urban materials, reference spectra for roads and other paved impervious surfaces were not included in the library. Reflectance spectra for green vegetation and for impervious materials were convolved to TM reflectance spectra and converted to 2001 digital number equivalents based on the conversion values developed for the NPV and soil field spectra above. The resulting collection of endmembers–hereafter referred to as the 'library collection'–consisted of 664 endmembers, 302 image endmembers and 362 reference endmembers. The number of endmembers in the library collection grouped by material class was as follows: 192 green vegetation, 105 NPV, 78 impervious, and 289 soil.

2.3.2. Endmembers representative of library

The next steps in MESMA library construction were to remove spectra that had a high probability of confusion with other material classes and to identify which spectra were most representative of the spectra within their material class. We followed a procedure developed by Dennison and Roberts (2003) in which each spectrum in the library collection was modeled as a series of two-endmember models, using all other spectra in the collection and photometric shade. Including a shade endmember in each SMA model (i.e. a spectrum with a reflectance of zero in all bands) accounts for variation in illumination (Dennison & Roberts, 2003). This resulted in 663 unique two-endmember models for each spectrum, and the RMSE for all models was recorded. Any image spectrum that was successfully modeled by a reference endmember from another class was removed from further consideration, eliminating image spectra that were potentially confused with materials of other classes. To identify the most representative spectra for each material class, a measure termed 'endmember average RMSE,' or EAR, was applied (Dennison & Roberts, 2003). The RMSE for each spectrum unmixing all other spectra within its class was averaged. For each class, the spectrum with the lowest average, or EAR, was considered the endmember that best represented that class.

2.3.3. Endmembers representative of image materials

How well a spectrum models other spectra within the library collection does not necessarily indicate how well that spectrum will model its class of materials in the image (Song, 2005). Furthermore, the best 'average' spectrum of a class in the library collection may not successfully account for the spectral diversity of materials on the ground. Therefore, a link needed to be made between the endmembers that best represented their groups within the library collection and the endmembers which sufficiently represented materials of their class on the ground.

Such a link was established by running a series of two-endmember models for each class of materials on the image and developing a set of criteria for selecting endmembers for the MESMA library. For each class, the procedure started with the spectrum with the lowest EAR, and the image was modeled as a two-endmember model (spectrum(1) + shade). Then, spectrum (1) was removed from the library collection, along with any spectra that it modeled, and EAR was recalculated. A new library, consisting of spectrum(1) + spectrum(2) + shade was used to model the entire image as two-endmember models. Spectrum(2) and any spectra that it modeled were removed from the library collection, EAR recalculated, and the new library of three bright endmember spectra was used to model the image, etc. At each iteration, the following were recorded: (1) the number of pixels modeled by each spectrum included in the library (i.e. number of pixels modeled by each in presence of competition), and (2) the number of pixels that would have been modeled by each spectrum had there been no other spectra in the library (i.e. number modeled by each without competition). The process was repeated until the increase in the number of pixels modeled when a new spectrum was added to the library was negligible.

Spectra were included in the final MESMA library based on two criteria: (1) the spectrum modeled a minimum number of pixels −0.01% of the number of land pixels in the image, or at least 633 pixels, and (2) a substantial proportion of pixels that could be modeled by spectrum(i) with no competition had to be modeled by spectrum(i) in the face of competition. That is, the ratio of the number of pixels modeled by spectrum(i) in the last run of image unmixing to the number of pixels that could be modeled by spectrum(i) with no competition must be greater than a specified threshold. For this case study, a threshold of 20% was applied. For example, to select which endmembers of the NPV class to include in the final MESMA library, 10 spectra were tested based on their EAR ranking (Table 2). Note that when spectra(9) and (10) were added to the library, the number of new pixels modeled was negligible, and so no further spectra were considered. Even with no competition, spectrum(7) modeled fewer than the threshold value ($<0.01\%$), and therefore was not included in the final MESMA library. Of the remaining spectra, those with EAR ranks of 1, 2, 3, 5, and 6 successfully met the second criterion, and therefore were considered necessary to capture the spectral diversity of this class of materials in the image. Thus, the final MESMA library included five endmembers for the NPV class. The procedure was repeated to determine the endmembers for the other three classes of materials.

2.4. SMA models

The endmember library developed above was used to map each pixel in the image in terms of the fractional abundance of each class. Two-, three-, and four-endmember models were applied and compared to determine the best model for each pixel. Allowed combinations of material classes are given in Table 1; all possible permutations of spectra for each model in Table 1 were tested. In an effort to increase computational efficiency by reducing the total number of models tested, all models were limited to different classes of materials, with the exception of allowing two types of impervious surfaces in a single model. This exception was deemed necessary based on visual inspection of the videography reference data, as well as preliminary analyses indicating that accuracy of all the modeled fractions increased if multiple impervious surfaces were allowed. To account for variations in brightness of surfaces due to viewing angle, topography, and other forms of shading, all models included the shade endmember (Dennison & Roberts, 2003).

Every allowed model was evaluated for every pixel in the image. Candidate models were selected based on the following constraints: (a) The bright fraction values were constrained between −0.10 and 1.10. (b) The shade fraction values were constrained between −0.10 and 0.50. (c) A RMSE threshold of 0.025

Table 2
Criteria used to select NPV endmembers for the MESMA library

Emb. name	No. map	Lib. Vers.	1	2	3	4	5	6	7	8	9	10	Total map
PLRA0004	16,009	1	16,009										16,009
DIMBARK6	14,406	2	13,197	11,029									24,226
mc190441	6062	3	12,517	10,278	2423								25,218
mc229416	1672	4	12,517	9940	2144	1072							25,673
mc326221	9126	5	10,811	9821	2099	1061	4870						28,662
ma417303	2326	6	10,811	9680	2095	244	4859	1895					29,584
ma107476	598	7	10,804	9673	2088	231	4841	1894	62				29,593
mc192444	5767	8	10,804	9622	1335	226	4841	1894	62	868			29,652
i355103	1461	9	10,804	9622	1334	213	4841	1883	62	868	56		29,683
ARME0020	1061	10	10,804	9617	1334	211	4841	1883	62	825	56	128	29,753
	map(v10)/no. map:		0.675	0.668	0.220	0.126	0.530	0.810	0.104	0.143	0.038	0.121	

Potential endmembers for the NPV material class were linked to materials on the ground by running a series of two-endmember models. The first run used a library with only one NPV endmember–that identified as the most representative of its class based on the EAR measure–and photometric shade. For each subsequent version, an additional NPV endmember was added to the library; the order endmembers were added was determined by the EAR measure (see text for full description). The number of pixels that could be mapped by each endmember if it were the only NPV endmember in the library (i.e. no competition) is indicated in the column labeled *No. map*. The total number of pixels successfully modeled by two-endmember models using all of the NPV endmembers for that version of the library is indicated in the column labeled *Total map*. Each row indicates the number of pixels mapped by each endmember for that version of the library. For example, Library Version 3 consisted of three NPV endmembers, and all 3 two-endmember models were tested for all pixels in the scene. The result was that endmember(1) modeled 12,517 pixels; endmember(2) 10,278; and endmember(3) 2423. However, in Library Version 10, 10 NPV endmembers were competing to model each pixel, and the number of pixels modeled by endmember(3) in that competitive environment dropped to 1334. The last row of the table indicates the ratio of the number of pixels modeled by each endmember in Library Version 10 to the number of pixels that could be modeled by that endmember in the absence of competition. For endmember(3), that ration is 1334/6062, or approximately 22%.

maximum reflectance (equivalent to approximately 6.4 DN) was applied. The constraints for the bright fractions and maximum shade fraction were tested on this data set and selected as the best compromise between the goal of physically reasonable fractions and the imprecision of per-pixel analysis introduced by the modular transfer function of Landsat data (Forster, 1985; Huang et al., 2002; Townshend et al., 2000). The RMSE constraint is most commonly accepted in the literature (e.g. Dennison & Roberts, 2003; Roberts et al., 1998b).

For each pixel, the best model at each level of model complexity (i.e. two-endmember, three-endmember, or four-endmember) was selected by comparing all models which met the constraints above and selecting the model with the lowest RMS error—i.e. assumed best fit (Painter et al., 1998). If no model met all constraints, the pixel was left unmodeled. At this stage, each pixel was associated with up to three potential models. Previous work has shown that there is a negative correlation between increasing model complexity and accuracy (Halligan, 2002; Sabol et al., 1992). Additionally, there is a positive correlation between model complexity and computational expense (Halligan, 2002; Roberts et al., 1998b). We therefore assumed the optimal model was the model with lowest complexity, though we note that this assumption is not tested empirically in this case study and it is possible that we lose information by choosing the simplest model. Specifically, model selection was based on the following steps: if pixel(i) could be modeled by a two-endmember model, the two-endmember model was selected as the optimal model. Otherwise, if pixel(i) could be modeled by a three-endmember model, the three-endmember model was selected, etc. Thus, pixel(i) was modeled as a four-endmember model only if no other option existed within the model constraints detailed above. The output of this selection process was an image indicating the optimal model per pixel, as well as the fractional value of each endmember associated with that model.

2.5. Mapping

The final product of our analysis was a set of fractional abundance maps for each class of materials (i.e. vegetation, impervious surfaces, soil, and water). Because shade was not considered a land-cover component, but rather a variant on endmember brightness, the fractions associated with each optimal model had to be converted to fractions that represented the physical abundance of the material present within each pixel. The fractions of each pixel were therefore shade-normalized, that is, each non-shade fraction was divided by the sum of the non-shade fractions for that pixel (Adams et al., 1993). For two-endmember models, the resulting shade-normalized fraction is 100%. After shade normalization, the endmembers were re-labeled as their corresponding generalized class. The NPV and green vegetation fractions were combined into a single vegetation class, as these two categories of spectra effectively represent all states of vegetation, i.e. NPV spectra model senesced vegetation and bark/stems, while green vegetation spectra model live, leafy vegetation. These shade-normalized, class-generalized fractions were combined to generate an image of each land-cover material, with values representing the physical abundance of that material. Pixels included in the water mask were assigned a fraction value of 100% water. An image of the RMSE values corresponding to the optimal model selected for each pixel was also generated.

2.6. Accuracy assessment

The videography time code was divided in to 15-s intervals, which were randomly sampled. Sampled intervals were mosaicked, georeferenced and inspected for cloud cover and unacceptable distortion due to aircraft movement. This resulted in a total of fifty-five 15-s mosaics. Each reference image was classified into the four material classes of interest–vegetation, impervious, soil, and water–by applying a multi-resolution segmentation algorithm and subsequent supervised classification as implemented in eCognition software (Baatz et al., 2003). The segmentation parameters were held constant for the processing of all 15-s mosaics. Classified segments were visually inspected for accuracy, and errors were manually corrected. The center timecode of each mosaic was used to locate the sample center on the reference image, and the UTM coordinates were extracted to locate the sample on the Landsat image. For each sample, the size of the sampling unit was varied, and the fractions of materials within each window size were extracted from the classified videography mosaics and from the corresponding window on the Landsat image. Unmodeled pixels were excluded from the average fraction calculations.

Window sizes corresponded to blocks of Landsat pixels, ranging from a single Landsat pixel (30 m × 30 m) to a 17 × 17 block of Landsat pixels (510 m × 510 m) in increments given in Table 3. Window sizes were variable for several reasons: (1) No assessment of the geometric error between the Landsat ETM+ image and the videography mosaics was conducted, and such error could potentially negatively impact the accuracy measures (Powell et al., 2004); (2) The precision of estimating per-pixel land cover may be limited by the modulation transfer function (MTF) of the Landsat sensor; in other words, the signal measured by the sensor for a given pixel is partially influenced by the land cover of surrounding pixels (Forster, 1985; Townshend et al., 2000; Huang et al., 2002); and (3) Though we derived our results from 30 × 30-m resolution data, we had no real knowledge of the appropriate spatial scale for interpreting the fractions generated. Varying the window size allowed us to investigate how measures of accuracy were impacted based on the size of the sampling unit, as well as to consider how confidence in the estimates of land-cover fractions might vary with spatial scale.

Table 3
Accuracy measures for MESMA fraction images (n = 55)

Win_size	1	3	5	9	13	17
Vegetation						
Slope	0.66	0.79	0.84	0.97	1.01	1.03
Intercept	13.5	7.8	4.4	−3.0	−5.2	−5.8
RSQ	0.43	0.65	0.71	0.85	0.89	0.91
Impervious						
Slope	0.58	0.81	0.83	0.93	0.96	0.96
Intercept	19.7	13.6	13.7	12.3	12.5	12.5
RSQ	0.30	0.66	0.70	0.85	0.89	0.90
Soil						
Slope	0.23	0.19	0.24	0.45	0.51	0.54
Intercept	5.8	6.7	5.8	2.2	0.8	−0.1
RSQ	0.07	0.05	0.08	0.31	0.46	0.57
MAE						
Vegetation	23.9	16.1	14.5	10.4	9.3	8.3
Impervious	25.1	15.7	15.3	12.2	11.7	11.8
Soil	15.5	13.3	11.2	10.0	8.1	7.8
Bias						
Vegetation	−4.3	−3.5	−4.3	−4.7	−4.9	−4.4
Impervious	7.2	7.9	8.8	10.5	11.4	11.5
Soil	−5.3	−4.4	−4.5	−5.8	−6.4	−7.0

Window size refers to the sample dimensions in units of Landsat ETM+ pixels.

Accuracy measures were based on the correlation between modeled fractions and reference fractions from the videography. The 'goodness' of the correlation was assessed by the slope, intercept and R-squared of the relationship, where, in an ideal case, the slope of the relationship would equal one, the intercept zero, and the R-squared value would approach one. Two types of error measurement were also used to evaluate the accuracy of the fraction estimations, mean absolute error and bias. Mean absolute error (MAE) is average absolute value of the difference between modeled and measured fraction values, while bias is the average of the error, indicating trends in over- or under-estimation (Schwarz & Zimmermann, 2005).

Those equations are given as:

$$\mathrm{MAE} = \left(\sum_{i=1}^{n} |\bar{Z}_{ki} - Z_{ki}| \right) / n \qquad (4)$$

$$\mathrm{bias} = \left(\sum_{i=1}^{n} (\bar{Z}_{ki} - Z_{ki}) \right) / n \qquad (5)$$

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 number of samples.

3. Results and discussion

3.1. MESMA library and models

The final endmember library was selected based on the criteria described above and consisted of 26 endmembers in addition to photometric shade (Fig. 4). The number of endmembers per generalized class was as follows: non-photosynthetic vegetation—five (2 reference endmembers, 3 image endmembers), green vegetation—two (image endmembers), soil—eight (1 reference, 7 image endmembers), and impervious—11 (3 reference, 8 image endmembers). The allowed models consisted of all possible permutations of the generalized combinations listed in Table 1. This resulted in 26 two-endmember models, 286 three-endmember models, and 825 four-endmember models.

Within the study area, 95.9% of the land pixels were successfully modeled. The distribution of model complexity for the final MESMA output is given in Table 4. Over 50% of the non-water pixels in the study area were modeled by two-endmember models, while an additional 45.5% of non-water pixels were modeled by three-endmember models. Four-endmember models were not included in the final product for several reasons. The inclusion of four-endmember models resulted in a very small increase in the number of successfully modeled pixels, approximately 1% of land pixels; yet, the

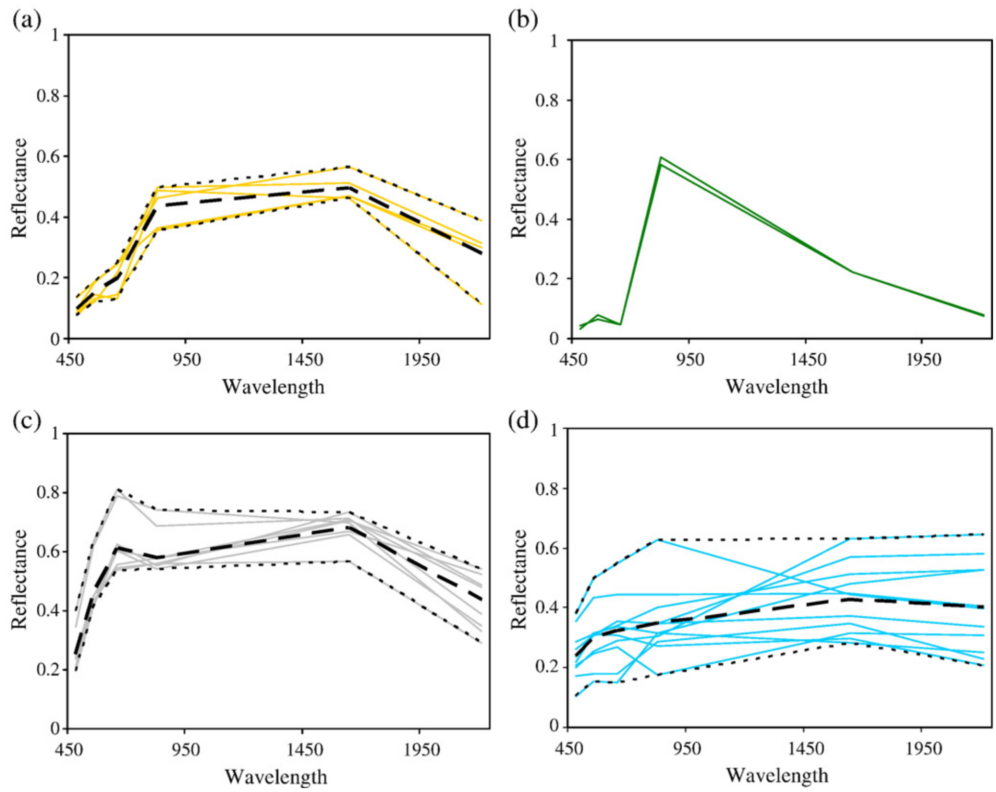

Fig. 4. Spectra included in the final endmember library: (a) NPV, (b) green vegetation, (c) soil, (d) impervious. Average reflectance values for each class–except green vegetation–are represented by the dashed line, and minimum values for each band by the dotted lines. (For interpretation of the references to colour in this figure legend, the reader is referrred to the web version of this aticle.)

computational expense of including four-endmember models is quite high. Additionally, including four-endmember models did not appreciably improve accuracy statistics, and visual inspection of the pixels modeled by four-endmember models indicated substantial errors. The majority of four-endmember models were either edge pixels between water and land or were located in topographically shaded areas of primary forest. Therefore, the final MESMA product that was used to generate maps of V–I–S fractions included only two- and three-endmember models.

3.2. Fraction maps

Maps of the generalized fractions are shown in Fig. 5. Bright areas represent higher fractions and dark areas lower fractions, while black pixels indicate that the material is not present. The oldest part of the city, which coincides with the central business district highlighted by the circle in Fig. 1, has very little vegetation and very high impervious values. In fact, for the heavily built-up areas of the city, the vegetation fraction map is almost the inverse of the impervious fraction map. Major roads are highlighted on the impervious map, as is the international airport located towards the upper left corner on the map. In the upper right portion of the map, there is a sharply defined area that has no impervious or soil fractions present; this is the southern edge of the Reserva Ducke, a natural reserve in which no development is permitted. The distribution of the soil fraction is much spottier than the vegetation and impervious fractions. Bare soil tends to be most prominent in areas undergoing construction and expansion, such as the area in the north, central portion of the map, adjacent to the Reserva Ducke, indicated by the oval in Fig. 1. Some roads are also highlighted on the soil fraction map. One source of spectral confusion stands out on the impervious map: some lowland areas that are seasonally flooded are mapped as having a relatively high impervious fraction, for example, along the edges of the two islands in the lower right portion of the map.

Table 4
Number of pixels successfully modeled for each level of model complexity

Model	Pixels modeled	% Non-water area	% Total area
Water mask	200,690	–	24.1
Two-endmember	319,323	50.4	38.3
Three-endmember	288,188	45.5	34.5
Four-endmember	6,485	1.0	0.8

The composition of V–I–S fractions can be compared between different neighborhoods. For example, we sampled four neighborhoods that represent a gradient of ages (Fig. 1). The oldest–highlighted by a circle–is the central business district (Centro), which was built up during the 19th Century. The second–highlighted by a square–is the neighborhood of Coroado, founded in the early 1970s (Silva-Forsberg, 1999). The third–highlighted by a parallelogram–is a part of the São José Operário neighborhood that has been constructed since 1996. The fourth–highlighted by an oval–an area north of the neighborhood of Cidade Nova, represents the most recent development, as construction began only after 1999. The dates for the

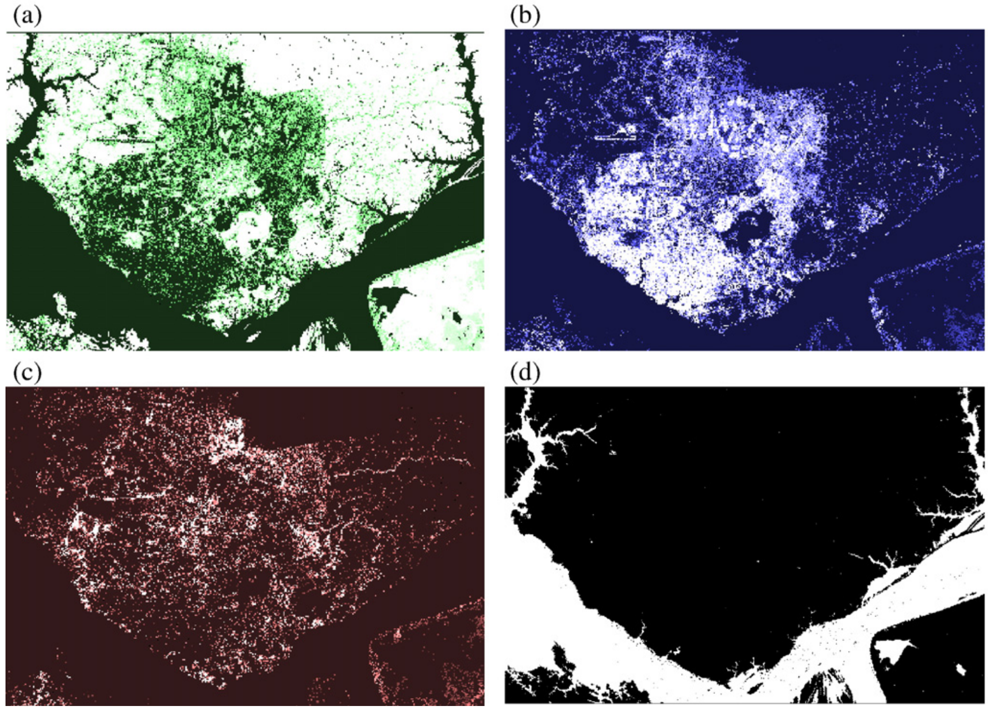

Fig. 5. Fraction images generated from MESMA: (a) vegetation fraction, (b) impervious fraction, (c) soil fraction, and (d) water mask. Brighter areas indicate higher fractions, while darker areas indicate lower fractions.

latter two sites were determined by inspecting Landsat imagery collected in 1996 and 1999. The average values for V–I–S fractions within each neighborhood are displayed in Fig. 6. Among these samples, the fraction of soil tends to decrease with age of the neighborhood, while the fraction of impervious surface tends to increase with age. This simple analysis suggests that trajectories of physical landscape change within a city may be characterized by mapping V–I–S fractions for multitemporal observations.

3.3. Accuracy assessment

The correlation between reference fractions and modeled fractions are reported for each window size in terms of slope, intercept, and correlation coefficient (i.e. R-squared value, see Table 3) and presented graphically for the 9×9 window size (Fig. 7). Note that the modeled vegetation fraction for each pixel is the sum of the shade-normalized green vegetation and NPV

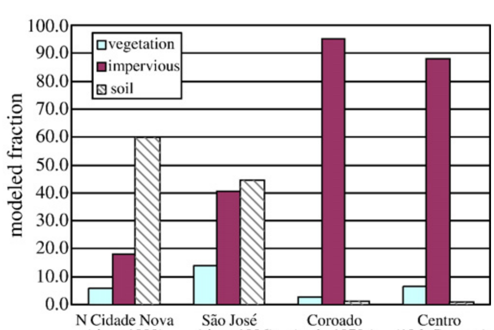

Fig. 6. V–I–S distribution for neighborhoods of different ages.

(i.e. senesced vegetation) fractions. Correspondence between reference and modeled fractions when the sampling unit corresponds to a single pixel (i.e. 1×1 window size) was low for all land-cover categories. However, as the sample unit increased in area, the correlation increased; i.e. the slope and R-squared value of the relationship both approached 1.0 for all fractions. Based on the values reported in Table 3, we propose that the minimum window size for an acceptable correlation is 9×9 pixels (i.e. 270 m×270 m). At that window size, slopes for vegetation and impervious were equal to 0.97 and 0.93 respectively; the intercept values were near zero for the vegetation and reasonably constrained for the impervious fractions, and both fractions had an R-squared value equal to 0.85. The mean absolute error (MAE) for the 9×9 window size ranged between 9.98 for soil and 12.23 for impervious. However, the correlation between reference and modeled soil fractions at this window size was still quite poor; in fact, agreement for soil fractions remained low across all window sizes.

Analysis of residuals in terms of the bias indicated that the impervious fractions were consistently over-estimated by the modeled output, while vegetation and soil were consistently under-estimated relative to the reference fractions (Table 3). The over-estimation of impervious as indicated by the positive bias was approximately equal to the under-estimation of the soil and vegetation fractions, a result of the SMA constraint that the fractions for any pixel must sum to 1.0 (Eq. (2)). A graph of the bias for each sample as a function of the reference fraction revealed patterns of over- and under-estimation (Fig. 7). The impervious fractions were estimated more accurately for samples with high fractions of impervious materials (>70%), and least accurately for the samples with low fractions of impervious

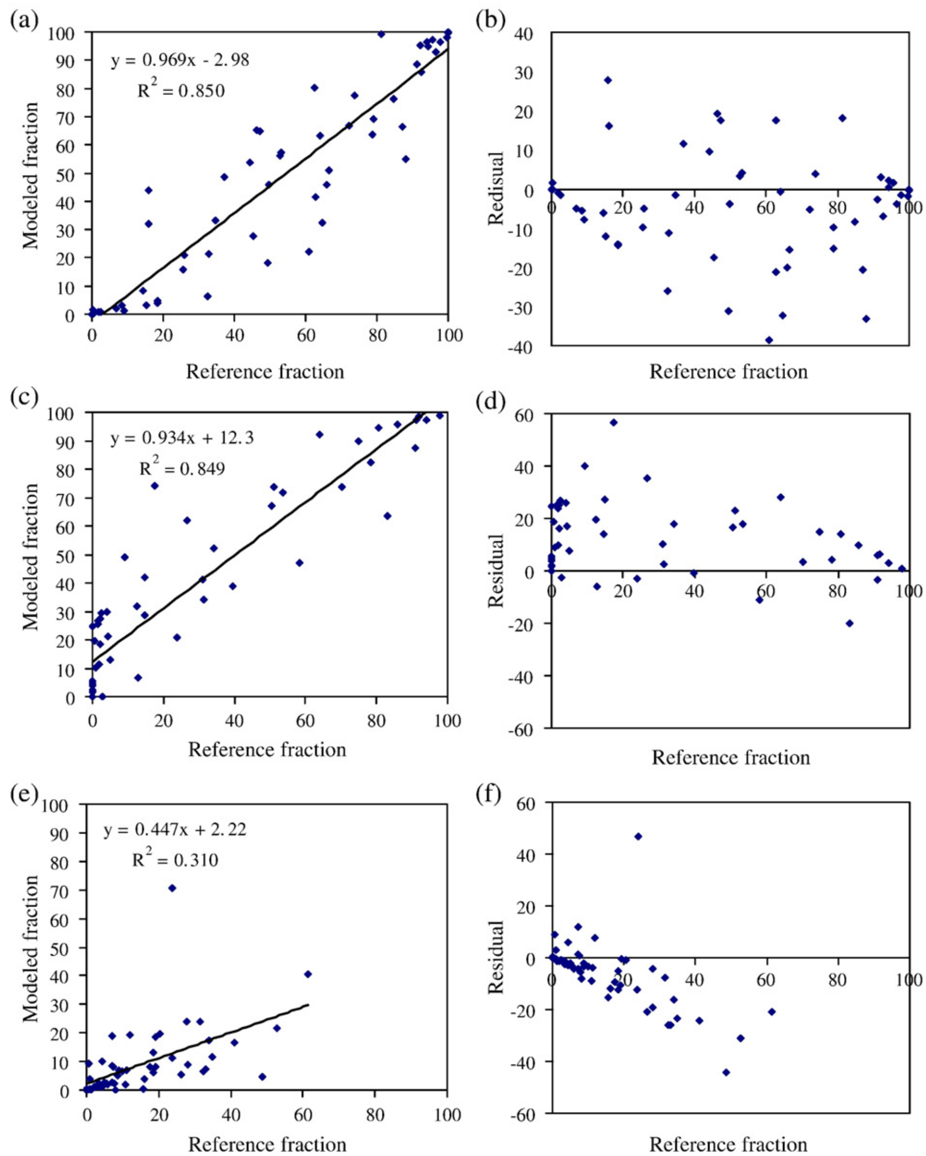

Fig. 7. Comparisons between reference and modeled fractions, window size = 9 × 9 pixels: (a) vegetation fraction scatter plot, (b) vegetation fraction residuals, (c) impervious fraction scatter plot, (d) impervious fraction residuals, (e) soil fraction scatter plot, (f) soil fration residuals.

surfaces (<30%). In only a handful of samples was the impervious fraction underestimated. The negative bias reported for vegetation suggested that overall vegetation fractions were under-estimated. Agreement was highest for very low (<10%) and very high (>90%) fractions, while the residuals for fractional cover in between were rather randomly distributed, with a few extremely negative values. The soil fraction was consistently under-estimated, except for one positive outlier, and residuals became increasingly negative as the reference fraction increases.

3.4. Discussion

3.4.1. Representative endmembers

Other researchers have noted the potential of applying MESMA to urban environments because the limited number of endmembers used in standard SMA studies does not sufficiently capture the spectral heterogeneity of urban environments (Lu & Weng, 2004; Rashed et al., 2003; Wu, 2004). In other words, a single endmember per category of material is not sufficient. Our work also indicates that it is not sufficient to build a spectral library based on the most representative endmembers of each category. The 'purest' endmember fractions are not necessarily representative of materials within the scene, and representative spectra may not necessarily be selected as 'pure' endmembers (Song, 2005). An endmember that is the 'most representative' of its class, in this case determined by EAR, may not capture land cover with distinct spectra that occupy small areas within the scene. For example, in the case of selecting the NPV endmembers (Table 2), spectrum(6) maps a relatively small area of the scene as a two-endmember model, but more than 80% of the pixels potentially mapped by spectrum(6) were mapped in the presence of competition. In other words, spectrum(6) represented a spectrally distinct material with small areal extent. Therefore, while spectrum(6) was not identified as most representative of the

Table 5
Accuracy measures for MESMA library constructed from top 3 EAR spectra for each class ($n=55$)

Win_size	1	3	5	9	13	17
Vegetation						
Slope	0.76	0.78	0.86	0.99	1.03	1.05
Intercept	12.2	14.2	7.8	0.2	−1.9	−2.4
RSQ	0.54	0.65	0.73	0.84	0.88	0.91
Impervious						
Slope	0.34	0.46	0.53	0.57	0.61	0.62
Intercept	3.1	3.5	2.8	0.9	0.3	0.1
RSQ	0.19	0.54	0.65	0.85	0.92	0.94
Soil						
Slope	0.03	0.05	0.12	0.30	0.35	0.35
Intercept	6.9	7.8	6.0	2.7	1.5	1.3
RSQ	0.00	0.01	0.04	0.23	0.44	0.53
MAE						
Vegetation	43.0	41.5	39.4	38.5	38.0	36.5
Impervious	33.4	30.8	31.5	29.8	29.2	28.0
Soil	18.9	16.1	13.5	13.5	12.5	11.6
Bias						
Vegetation	−0.4	2.6	0.3	−0.4	−0.1	0.2
Impervious	−16.6	−12.5	−10.6	−10.9	−10.3	−10.2
Soil	−7.2	−5.2	−5.9	−7.4	−8.0	−8.4

endmembers in the library, it was included as an endmember in the final MESMA library because it could model materials that would be not be captured by more ‘representative’ endmembers.

To compare the improvement resulting from the multi-step endmember selection method presented in this paper to a spectral library composed of the ‘most representative’ endmembers, we built a spectral library consisting of the three most representative endmembers for each material class within the library, as measured by EAR. This ‘EAR library’ therefore consisted of 12 spectra in addition to photometric shade, and MESMA was applied to the image using the same constraints and models as detailed in the Methods section above. The EAR library successfully modeled 80.3% of the non-water pixels, a significant decrease (−15.6%) compared to the pixels modeled by the MESMA library built by linking spectra to the scene. The agreement between reference fractions and fractions modeled by the EAR library is presented in Table 5. The correlation for the vegetation fraction is essentially equivalent to the correlation reported for the final MESMA library because two of the three vegetation endmembers were identical for the two libraries. However, the number and combinations of endmember included in each library for the other material classes were quite different, and correlations for the impervious and soil fractions are greatly improved when endmembers were selected based on the spectral variability of materials within the scene.

3.4.2. Accuracy assessment

While other studies have applied SMA to map the fractional abundance of the physical components in an urban environment, this is the first study to our knowledge that provides a quantitative accuracy assessment of the continuous fraction values for all components. However, it should be remembered that the measures reported for any accuracy assessment, including those discussed above, refer to the agreement between reference data and modeled data, and do not necessarily reflect agreement between modeled data and ‘truth’ (Foody, 2002). An additional step in qualitatively assessing the accuracy of the modeled fractions, therefore, is to consider the quality of the reference data, in particular, whether the over- and under-estimation attributed to the modeled fractions might in part be due to systematic errors in the reference data classification (Powell et al., 2004). For example, analysis of residuals indicated that the vegetation fractions tended to be under-estimated by the MESMA output. Close inspection of the classified videography, however, suggested conditions that may have resulted in a systematic over-estimation of the vegetation fraction in the reference data. Specifically, areas that were highly shaded, such as spaces between buildings, were classified as vegetation in the videography, but more likely consisted of impervious surfaces or bare soil. Similarly, very dark roofs were classified as vegetation, resulting in an over-estimate of vegetation and corresponding under-estimate of the impervious area in the reference data. This could partly account for some of the biases noted above.

Two issues associated with classifying the reference data also provide insight concerning the low agreement between reference and modeled soil fractions. First, visually distinguishing between impervious and soil cover was difficult, especially in areas that were highly built-up. As a result, it was often difficult to verify the accuracy of the eCognition classification. Second, comparing the fractional cover generated by MESMA with the traditional exclusive and exhaustive classes imposed on the videography mosaics introduced a fundamental problem, that of equating two different types of data-continuous and discrete. While the segmentation algorithm implemented by eCognition provided an efficient means of classifying the complex videography mosaics, classifying the segments into discrete classes of materials was problematic in two cases.

The first type of problem occurred because the segments generated did not always correspond to a single class of land-cover materials. For example, in densely built-up areas, bare soil tended to occur in tiny slivers, and the segments generated often included soil slivers in addition to some surrounding impervious materials. If these segments were consistently classified as soil on the videography, the reference soil fractions would have been over-estimated. The second type of problem occurred when two classes of materials existed as a mixture within the segment. For example, a vacant lot could consist of soil mixed with senesced grasses. The segment which included that lot had to be classified as one material or the other–the segment was assigned to the material that was most visually abundant–while the MESMA output captured and quantified the presence of both materials. If, the videography segments in such cases were consistently classified as soil, the reference soil fractions would again have been over-estimated.

Finally, presenting accuracy assessment in terms of correlations had a distinct disadvantage compared to presenting a traditional error matrix. The former provides no specific information concerning the confusion between classes of materials, while the latter displays counts of agreement and disagreement between classes (Foody, 2002). While the correlation measures indicate the level of agreement between reference and modeled data, there is no way to unravel specifically where the disagreement lies. In other words, we may know that impervious surfaces were in general over-estimated, but we do not know, without looking at each sample individually, whether impervious surfaces were generally confused with NPV, with green vegetation, or with soil. From previous work we have conducted in this region, we do know that common sources of confusion for Landsat data include some NPV and soil spectra, as well as some soil and impervious spectra. Both sources of spectral confusion could also contribute to the low accuracy of the soil fraction.

We investigated five of the largest residuals across all materials that resulted from the accuracy assessment using a 9×9 window size. In four of the five cases, the samples were located on a sharp and long boundary between two contrasting land cover types, such as between a highly built-up area (i.e. dominant impervious fraction) and a tract of forest (i.e. dominant vegetation fraction). Because the boundary between land-cover types cut through the entire accuracy sample, the impact of a georegistration error between the reference and modeled data is quite severe. For example, if a 9×9 sample is centered over a border between two homogeneous land-cover types, the fractions recorded for that sample are 50% for each material class. If the sample window is shifted one pixel in a direction perpendicular to the boarder, the fractions change to 39% and 61%. If the window is shifted two pixels, the fractions change to 28% and 72%, resulting in a significant over-estimate of one fraction and corresponding under-estimate of the other, even if the modeled and reference data would have agreed exactly had they been accurately georeferenced. The fifth case, with an extremely large residual, was clearly the result of land-cover change that occurred between the date the videography was collected and the date the ETM+ image was acquired, a difference of 2 years.

The samples resulting in the five largest residuals were removed from the pool, and accuracy assessment measures recalculated using the remaining 50 samples (Table 6). For all materials and at all window sizes, the accuracy measures improved, especially for the soil fraction. In all cases, slope and R-squared values increase (i.e. approach 1.0), and the intercept shifted closer to zero. The MAE and the bias also decreased in all cases. We can conclude, therefore, that a truer measure of modeled fraction accuracy would require more careful co-registration of the reference and image data, screening reference samples for possible change between dates, and perhaps using higher resolution reference data so that more accurate discrimination of land-cover materials was possible. However, the fundamental problem of comparing the exclusive and exhaustive reference data classes with the continuous values of the modeled fraction data remains, and may ultimately limit the quality of accuracy assessment for continuous data, a challenge also noted by Rashed et al. (2005).

Table 6
Accuracy measures with revised samples (n=50)

Win_size	1	3	5	9	13	17
Vegetation						
Slope	0.75	0.85	0.89	0.98	1.00	1.02
Intercept	11.6	6.2	3.5	−2.0	−4.1	−5.0
RSQ	0.53	0.73	0.81	0.90	0.90	0.92
Impervious						
Slope	0.63	0.85	0.87	0.95	0.97	0.97
Intercept	15.4	10.1	10.0	9.9	11.0	11.4
RSQ	0.36	0.73	0.78	0.91	0.92	0.92
Soil						
Slope	0.48	0.43	0.42	0.48	0.49	0.50
Intercept	2.1	3.2	3.1	1.3	0.8	0.4
RSQ	0.26	0.25	0.28	0.63	0.65	0.71
MAE						
Vegetation	21.5	14.0	11.9	8.6	8.6	8.1
Impervious	22.9	13.8	13.1	10.2	10.2	10.7
Soil	11.4	9.9	8.5	7.0	6.7	6.8
Bias						
Vegetation	−1.4	−1.8	−2.1	−3.0	−4.1	−4.0
Impervious	3.3	5.2	6.0	8.5	10.0	10.5
Soil	−4.5	−3.6	−4.0	−5.4	−5.8	−6.3

The spectral library applied was the final MESMA library (compare to Table 3).

3.4.3. Unit of analysis

The effect of increasing correlation between reference and modeled fractions as window size increased was an expected result for several reasons: (a) increasing the window size should decrease the impact of geolocation error, though not in all cases, as discussed above; (b) the signal recorded at the sensor for a single pixel is affected by the spectral properties of surrounding pixels, and aggregating the modeled data reduces this effect (e.g. Forster, 1985; Huang et al., 2002; Townshend et al., 2000); and (c) the process of averaging fractions over larger areas should reduce the variance of each data set (Woodcock & Strahler, 1987), and thereby increase correlation between fractions, assuming the means of the two datasets are similar. Given the challenges of accuracy assessment for continuous fractions across the urban landscape and the subsequent limitations on confidence in relating reference and modeled fractions, generating accuracy measures with different sample sizes provides a way to assess the appropriate unit of analysis for a given application (Small, 2001). For example, an analyst might want to consider the level of agreement that would be required for confidence in the results of change detection between two dates. Based on the accuracy measures for this case study as reported in Tables 3 and 6, we suggest that the minimum resolution required for such a comparison is the 9×9 window size. While aggregating the modeled fractions to this resolution (270 m) results in considerable loss of heterogeneity and specificity, it is still a resolution that allows assessment of the general physical composition of the urban landscape, as well as comparisons between subdivisions of the city, such as neighborhoods.

3.5. Conclusion

In this paper, we have presented: a) an application of MESMA in an urban environment using regionally specific endmembers to map the physical abundance of generalized urban materials; b) a methodology for endmember selection which incorporates multiple sources of spectra and links the most representative spectra in each library class to the spectral variability of materials present on the landscape; and c) accuracy assessment for fraction images corresponding to each physical component. These techniques adequately characterized the diversity of materials that compose land cover within a diverse urban area, and at the same time provided a conceptual structure for grouping the specific materials into three general classes—vegetation, impervious, and soil. These generalized classes can characterize urban land cover regardless of specific construction materials or local environmental variation (Ridd, 1995), facilitating comparison of urban data sets on a global scale. We have demonstrated the feasibility of deriving these measures from moderate spectral/spatial resolution imagery. Because of the global availability and historic archive of such data (i.e. over 30 years of Landsat data), regional comparisons of urban development through time are possible.

Future research directions include improving the quality of the reference data to better assess the accuracy of the modeled fractions. In addition, more systematic documentation of the confusion between specific material classes might provide insight into the limitations of spectral mixture analysis applied to data of moderate spectral resolution due to spectral confusion. Analysis of spectral confusion could also lead to development of techniques to screen spectra from the MESMA library based on potential confusion with other material classes. Another issue to investigate is whether different selection rules for determining model complexity lead to higher fraction accuracy, as it is possible that selecting for computational efficiency does not produce the most accurate results. Finally, we plan to apply the methodology developed herein to other cities with different populations, different development histories, and different natural environments to assess the generality of these techniques.

Acknowledgements

This research was funded primarily by a NASA Earth System Science Graduate Student Fellowship. Support was also provided by NASA grant NCC5-282 as part of LBA-Ecology. Digital videography was acquired as part of LBA-Ecology investigation LC-07. The 2001 Landsat ETM+ image was acquired from the Tropical Rain Forest Information Center (TRFIC). We also thank the reviewers for their helpful comments.

References

Adams, J. B., Sabol, D. E., Kapos, V., Almeida Filho, R., Roberts, D. A., Smith, M. O., et al. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, *52*, 137−154.

Adams, J. B., Smith, M. O., & Johnson, P. E. (1993). Imaging spectroscopy: Interpretation based on spectral mixture analysis. In C. M. Pieters, & P. A. J. Englert (Eds.), *Remote geochemical analysis: Elemental and mineralogical composition* (pp. 145−164). Cambridge, England: Cambridge University Press.

Baatz, M., Benz, U., Dehghani, S., Heynen, M., Höltje, A., Hofmann, P., et al. (2003). *eCognition user guide 3: Object oriented image analysis.* Munich, Germany: Definiens Imaging GmbH.

Boardman, J. W., Kruse, F. A., & Green, R. O. (1995). *Mapping target signatures via partial unmixing of AVIRIS data. Summaries of the Fifth Annual JPL Airborne Geoscience Workshop* (pp. 23−26). Pasadena, CA: Jet Propulsion Laboratory Publications.

Browder, J. O., & Godfrey, B. J. (1997). *Rainforest cities: Urbanization, development, and globalization of the Brazilian Amazon.* New York, New York: Columbia University Press.

Carlson, T. N., & Sanchez-Azofeifa, G. A. (1999). Satellite remote sensing of land use changes in and around San José, Costa Rica. *Remote Sensing of Environment*, *70*, 247−256.

Clapham, W. B., Jr. (2003). Continuum-based classification of remotely sensed imagery to describe urban sprawl on a watershed scale. *Remote Sensing of Environment*, *86*, 322−340.

Conel, J. E. (1990). Determination of surface reflectance and estimates of atmospheric optical depth and single scattering albedo from Landsat Thematic Mapper data. *International Journal of Remote Sensing*, *11*, 783−828.

Dennison, P. E., & Roberts, D. A. (2003). Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sensing of Environment*, *87*, 123−135.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*, 185−201.

Forster, B. C. (1985). An examination of some problems and solutions in monitoring urban areas from satellite platforms. *International Journal of Remote Sensing*, *6*, 139−151.

Furby, S. L., & Campbell, N. A. (2001). Calibrating images from different dates to 'like-value' digital counts. *Remote Sensing of Environment*, *77*, 186−196.

Halligan, K.Q. (2002). *Multiple endmember spectral mixture analysis of vegetation in the northwest corner of Yellowstone National Park*. Master's Thesis, University of California, Santa Barbara.

Hess, L. L., Novo, E. M. L. M., Slaymaker, D. M., Holt, J., Steffen, C., Valeriano, D. M., et al. (2002). Geocoded digital videography for validation of land cover mapping in the Amazon basin. *International Journal of Remote Sensing*, *23*, 1527−1556.

Huang, C., Townshend, J. R. G., Liang, S., Kalluri, S. N. V., & DeFries, R. S. (2002). Impact of sensor's point spread function on land cover characterization: assessment and deconvolution. *Remote Sensing of Environment*, *80*, 203−212.

IBGE (2000). *Censo Demográfico.* Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística.

Jensen, J. R. (1983). Biophysical remote sensing. *Annals of the Association of American Geographers*, *73*, 111−132.

Ji, M., & Jensen, J. R. (1999). Effectiveness of subpixel analysis in detecting and quantifying urban imperviousness from Landsat Thematic Mapper Imagery. *Geocarto International*, *14*, 31−39.

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., et al. (2001). The causes of land-use and land-cover change: Moving beyond the myths. *Global Environmental Change*, *11*, 261−269.

Lu, D., & Weng, Q. (2004). Spectral mixture analysis of the urban landscape in Indianapolis with Landsat ETM+ imagery. *Photogrammetric Engineering and Remote Sensing*, *70*, 1053−1062.

Myint, S. W., Lam, N. S. -N., & Tyler, J. M. (2004). Wavelets for urban spatial feature discrimination: comparisons with fractal, spatial autocorrelation, and spatial co-occurrence approaches. *Photogrammetric Engineering and Remote Sensing*, *70*, 803−812.

Okin, G. S., Roberts, D. A., Murray, B., & Okin, W. J. (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment*, *77*, 212−225.

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

Powell, R. L., Matzke, N., de Souza, C., Jr., Clark, M., Numata, I., Hess, L. L., et al. (2004). Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon. *Remote Sensing of Environment*, *90*, 221−234.

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.

Rashed, T., Weeks, J. R., Stow, D., & Fugate, D. (2005). Measuring temporal compositions of urban morphology through spectral mixture analysis: Towards a soft approach to change analysis in crowded cities. *International Journal of Remote Sensing*, *26*, 699−718.

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., Batista, G. T., Pereira, L. G., Waller, E. K., & Nelson, B. W. (1998). Change identification using multitemporal spectral mixture analysis: Applications in Eastern Amazonia. In R. S. Lunetta, & C. D. Elvidge (Eds.), *Remote sensing change detection: Environmental monitoring methods and applications* (pp. 137−161). Chelsea, Michigan: Ann Arbor Press.

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., Yamaguchi, Y., & Lyon, R. J. P. (1985). Calibration of airborne imaging spectrometer data to percent reflectance using field spectral measurements. *Proceedings of the 19th International Symposium on Remote Sensing of Environment* (pp. 679−688). Ann Arbor, Michigan: Environmental Research Institute of Michigan.

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

Schwarz, M., & Zimmermann, N. E. (2005). A new GLM-based method for mapping tree cover continuous fields using regional MODIS reflectance data. *Remote Sensing of Environment*, *95*, 428−443.

Silva-Forsberg, M. C. (1999). *Protecting an Urban Forest Reserve in the Amazon: A Multi-Scale Analysis of edge Effects, Population Pressure, and Institutions*. Dissertation, Center for the Study of Institutions, Population and Environmental Change, Indiana University, Bloomington.

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

Small, C. (2005). A global analysis of urban reflectance. *International Journal of Remote Sensing*, *26*, 661−681.

Smith, M. O., Johnson, P. E., & Adams, J. J. (1985). Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis. *Journal of Geophysical Research*, *90*, C797−C804.

Song, C. (2005). Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporated endmember variability. *Remote Sensing of Environment*, *95*, 248−263.

Souza, C. M., Jr., Roberts, D. A., & Cochrane, M. A. (2005). Combining spectral and spatial information to map canopy damage from selective logging and forest fires. *Remote Sensing of Environment*, *98*, 329−343.

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.

Townshend, J. R. G., Huang, C., Kalluri, S. N. V., DeFries, R. S., Liang, S., & Yang, K. (2000). Beware of per-pixel characterization of land cover. *International Journal of Remote Sensing*, *21*, 839−843.

Woodcock, C. E., & Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, *21*, 311−332.

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

Wu, C., & Murray, A. T. (2003). Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, *84*, 493−505.

Xian, G., & Crane, M. (2005). Assessments of urban growth in the Tampa Bay watershed using remote sensing data. *Remote Sensing of Environment*, *97*, 203−215.