ISPRS Journal of Photogrammetry and Remote Sensing 101 (2015) 163–173

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Incorporating land use land cover probability information into endmember class selections for temporal mixture analysis

Wenliang Li, Changshan Wu *

Department of Geography, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53201-0413, USA

ARTICLE INFO

Article history:
Received 15 November 2014
Received in revised form 3 December 2014
Accepted 4 December 2014

Keywords:
Logistic regression
Classification tree
Land use land cover probability
Endmember class
Temporal mixture analysis
Impervious surfaces

ABSTRACT

As a promising method for estimating fractional land covers within a remote sensing pixel, spectral mixture analysis (SMA) has been successfully applied in numerous fields, including urban analysis, forest mapping, etc. When implementing SMA, an important step is to select the number, type, and spectra of pure land covers (also termed endmember classes). While extensive studies have been conducted in addressing endmember variability (e.g. spectral variability of endmember classes), little research has paid attention to the selection of an appropriate number and types of endmember classes. To address this problem, in this study, we proposed to automatically select endmember classes for temporal mixture analysis (TMA), a variant of SMA, through incorporating land use land cover probability information derived from socio-economic and environmental drivers. This proposed model includes three consecutive steps, including (1) quantifying the distribution probability of each endmember class using a logistic regression analysis, (2) identifying whether each endmember class exists or not in a particular pixel using a classification tree method, and (3) estimating fractional land covers using TMA. Results indicate that the proposed TMA model achieves a significantly better performance than the simple TMA and a comparable performance with the METMA with an SE of 2.25% and an MAE of 3.18%. In addition, significantly better accuracy was achieved in less developed areas when compared to that of developed areas. This may indicate that an appropriate endmember class set might be more essential in less developed areas, while other factors like endmember variability is more important in developed areas.

1. Introduction

Spectral mixture analysis (SMA) is a promising method for estimating fractional land covers within a remote sensing image pixel. As a physical based approach, SMA has proven effective in addressing mixed pixel problem and been successfully applied in urban analysis (Weng and Hu, 2008; Wu and Murray, 2003; Wu, 2004), vegetation mapping (Liu and Yang, 2013; Small, 2001; Small and Lu, 2006; Song, 2005), geological mapping (Bedini, 2009; Johnson et al., 1993; Ramsey and Christensen, 1998; Sunshine and Pieters, 1993), natural hazard risk assessment (Jia et al., 2006; Eckmann et al., 2010), etc. SMA assumes the spectra of a pixel as a linear/nonlinear combination of the spectra of comprised pure land cover types (also called endmember classes), and the areal fraction of each endmember class can be estimated through an inversion model such as the inverse least squares deconvolution method. For successful applications of SMA, an important step is the selection of the number, type, and spectral signatures of endmember classes (Tompkins et al., 1997; Elmore et al., 2000). Within these, the choice of appropriate spectral signatures has been extensively studied in the past decades (Somers et al., 2011). In particular, Asner and Lobell (2000) introduced an autoSWIR approach to extract endmembers through highlighting the variations among different land cover types. In addition to band selection, spectral transformations have also been applied in SMA. Wu (2004) proposed a normalized spectral mixture analysis (NSMA) to reduce within-class brightness variations of the reflectance spectra. Zhang et al. (2004) developed a derivative spectral unmixing (DSU) method to highlight between-class variability while reducing within-class variability, with which spectral unmixing was implemented with the second derivative endmember spectra. Further, wavelet analysis techniques, including discrete wavelets and continuous wavelets, were proposed by Li (2004). In addition to the spectral transformation method, spectral weighting is another alternative. Instead of considering all bands as equally important, the weighted spectral mixture analysis (WSMA) (Chang and Ji, 2006) assumes that the spectral bands with the

* Corresponding author. Tel.: +1 (414)229 4860; fax: +1 (414)229 3981.
E-mail addresses: wenliang@uwm.edu (W. Li), cswu@uwm.edu (C. Wu).

http://dx.doi.org/10.1016/j.isprsjprs.2014.12.007
0924-2716/

lowest endmember variability are more important, thereby assigned higher weights accordingly. Further, Somers et al. (2009a) found that the traditional SMA generally overlooked the correlation of error variance and reflectance energy, and then developed a two-step WSMA through taking the variations of the reflected energy into consideration. Moreover, Somers et al. (2010) developed a stable zone unmixing (SZU) technique to select spectral bands with the lowest within-class variations in SMA for generating corresponding land cover fractions. In addition, the multiple endmember spectral mixture analysis (MESMA) method was also proposed to address endmember variability (Roberts et al., 1998). With MESMA, a spectra library with a wide range of endmembers is constructed, and different sets of endmember combinations, instead of a fixed endmember set, are employed to derive the fractional land covers for each pixel. Recently, Deng and Wu (2013) developed a spatially adaptive spectral mixture analysis (SASMA) to mitigate endmember variability through only including spatially adjacent endmember candidates into consideration.

Besides the selection of an appropriate endmember sets, the choices of type and number of endmember classes also play an essential role in SMA and a few studies have examined this issue (Somers et al., 2009b). One early attempt was made by Smith et al. (1990), who proposed a three-endmember model: vegetation–soil–shade (V–S–S) model for mapping vegetation cover fractions in a desert environment. Furthermore, Ridd (1995) proposed a vegetation–impervious surface–soil (V–I–S) model for characterizing urban environments. Moreover, Small (2001) considered that urban environments are composed by vegetation, low-albedo materials, and high-albedo materials, and developed another three-endmember model (V-L-H) for urban applications. Wu and Murray (2003) developed a four-endmember model: vegetation-low albedo-high albedo-soil (V-L-H-S) model to characterize the urban environments of Columbus, Ohio, United States. Zhang et al. (2014) applied the V-L-H model in high-density urban areas, and the V-L-H-S model in low-density urban areas. In addition to these models with a fixed endmember set for each pixel, MESMA has taken the endmember class types and number into consideration, and chosen the optimal endmember class types based on a set of error terms (Dennison and Roberts, 2003).

While these models provide valuable references in selecting endmember classes for parameterizing biophysical compositions, several problems still remain. Specifically, endmember classes are typically identified from vertices of n-dimensional scatter plots and then applied to the entire study area. This approach applies the same endmember class set to each individual pixel for the whole image, and neglects their spatial distributions. Consequently, if a non-existing endmember class is included in the unmixing process, the estimated abundance for this endmember class is usually over zero, thereby inevitably generating large over-estimation errors. On the contrary, if an endmember class is mistakenly ignored, the resultant abundance of that class is zero, thereby leading to severe under-estimation errors. While MESMA has taken the selections of endmember class types and number into consideration, it neglects the spatial associations of land use land covers, and only employs pixel-wise estimation errors (e.g. root mean square error (RMSE)) as the criteria.

To address this problem, in this study, we proposed to automatically select endmember classes for SMA through incorporating land use land cover probability information derived from socio-economic and environmental drivers. A logistic regression model is typically applied to estimate the distribution probability of a land use land cover (LULC) type through analyzing the relationships among LULCs and their socio-economic and environmental factors, such as distance to the nearest city, distance to water, elevation, and slope. For instance, urban infrastructures are likely to be found in areas with flat or gentle slopes and with convenient access to transportation networks; vegetation may be found near to lakes and river streams. This logistic regression analysis has been applied in explaining the process of urbanization, and simulating future urban development scenarios (Verburg et al., 2004; Li and Wu, 2013; Li et al., 2014). Such analysis, therefore, may provide a physical means of explaining the existence and spatial distribution of endmember classes, and its successful applications in land use studies suggest that it might be a better alternative for addressing the endmember class selection issue in SMA. In this study, we incorporated LULC probability information into the selections of endmember class type and number, and further inputted to spectral unmixing for better deriving impervious surface estimates. This proposed model includes three consecutive steps, including (1) quantifying the distribution probability of each endmember class using a logistic regression analysis, (2) identifying whether each endmember class exists or not in a particular pixel using a classification tree method, and (3) estimating fractional land covers using temporal mixture analysis (TMA). TMA is a variant of SMA, with which NDVI values, instead of the original reflectance spectra, are employed for spectral unmixing (Knight and Voth, 2011; Yang et al., 2012, 2014; Li and Wu, 2014). One major advantage of TMA is the incorporation of NDVI time series profiles into spectral unmixing. With NDVI time series profiles, different phenophases (green up, maturity, senescence, and dormancy) of green vegetation can be identified and the significant amount of phenological information can be of great help for distinguishing one land cover type from another, thereby effectively addressing the issues of endmember variability. For instance, impervious surfaces only experience the dormancy phenological stage during the entire year with very low NDVI values. Bare soil, on the other hand, is with very low NDVI values in fall and winter (e.g. dormancy), but with relatively high NDVI values in spring and summer due to the growth of vegetation. Therefore, NDVI time series profiles are essential for distinguishing impervious surfaces and bare soil. Further, NDVI time series profiles can also help estimating the percent of impervious surfaces due to its close relation to vegetation fractions. Therefore, TMA, instead of SMA, has been employed in this study for estimating large-scale impervious surface distribution. The rest of the paper is organized as follows. The study area and data are introduced in Section 2. Details of the proposed TMA approach, as well as comparative analysis and accuracy assessment, are described in Section 3. Results of this study are reported in Section 4. Finally, discussion and conclusions are provided in Sections 5 and 6.

2. Study area and data

The State of Wisconsin, located in the north-central part of the United States of America, was chosen as the study area. Wisconsin is within the longitudes of 86°46′–92°53′W and latitudes of 42°37′–47°05′N (see Fig. 1), with a geographic area of 169,639 square kilometers that covering 72 counties. The total population of Wisconsin was 5.68 million in 2010, and the majority of the population resides in the southeastern part of the state. Approximately 68% of the population in Wisconsin resides in urbanized areas, including Milwaukee, Madison, and Green Bay, etc. Wisconsin has experienced fast population growth and urbanization in the past decades. The population has increased almost 16%, from 4.89 million in 1990 to 5.68 million in 2010, and it is expected that the growth trend will continue in the future. This rapid urbanization has caused challenging issues, including environmental pollution, traffic congestion, etc. As impervious surfaces are served as one of the major components of urban areas and widely considered as a key indicator of urbanization intensity and environmental quality, accurate estimates of impervious surfaces are essential for Wisconsin.