EXAMINATION OF SPECTRAL TRANSFORMATIONS ON SPECTRAL MIXTURE ANALYSIS

Y. Deng, C. Wu*

Department of Geography, University of Wisconsin-Milwaukee, Milwaukee, USA, Y. Deng (yingbin@uwm.edu), C.Wu
(cswu@uwm.edu)

Commission WG III/1

ABSTRACT:

While many spectral transformation techniques have been applied on spectral mixture analysis (SMA), few study examined their necessity and applicability. This paper focused on exploring the difference between spectrally transformed schemes and untransformed scheme to find out which transformed scheme performed better in SMA. In particular, nine spectrally transformed schemes as well as untransformed scheme were examined in two study areas. Each transformed scheme was tested 100 times using different endmember classes' spectra under the endmember model of vegetation- high albedo impervious surface area-low albedo impervious surface area-soil (V-ISAh-ISAl-S). Performance of each scheme was assessed based on mean absolute error (MAE). Statistical analysis technique, Paired-Samples T test, was applied to test the significance of mean MAEs' difference between transformed and untransformed schemes. Results demonstrated that only NSMA could exceed the untransformed scheme in all study areas. Some transformed schemes showed unstable performance since they outperformed the untransformed scheme in one area but weakened the SMA result in another region.

1. MANUSCRIPT

1.1 Introduction

Spectral mixture analysis (SMA) is majorly applied to estimate land cover class fractions from remotely sensed data, especially for medium and coarse spatial resolution imageries. It assumes that a mixed pixel in a remote sensing image is comprised by several pure land cover types (endmembers).

The main limitation of SMA to acquire accurate result is spectral variability. Within-class variability indicates the spectral differences within a land cover class. Conversely, different land cover types share similar spectral signatures is named between-class variability (similarity). Large within-class variability and small between-class variability are reasons of causing the spectral confusion.

Spectral transformation, focusing on enhancing the spectral characteristics to reduce the within-class variability and to enhance the between-class variability, is one of the widely applied approaches used in SMA to solve spectral variability problems. The discussion of spectral transformation is not enough since few studies examine their effectiveness comprehensively and systematically. The advantages and limitations of each transformed scheme are still unsettled, and the necessity of applying transformed schemes has not been adequately discussed in the literature. Therefore, this study aims to explore the difference between spectrally transformed scheme and untransformed scheme to find out which transformed scheme performed better in SMA. Next section introduces the background of spectral mixture analysis and

spectral transformed techniques. Section 3 presents the experiments as well as the results in two locations with Landsat data. Discussion and conclusions are presented in sections 4 and 5.

2. BACKGROUND

2.1 Spectral Mixture Analysis

The assumption of spectral mixture analysis (SMA) is that more than one land cover classes exist in a mixed pixel. It targets in estimating the land cover classes' fraction within a mixed pixel. SMA can be expressed using equation (1).

$$R_k = \sum_{j}^{n} f_j R_{k,j} + \varepsilon_k \qquad (1)$$

where R_k = spectral reflectance of mixed pixel on band k

f_j = fraction of endmember j within the pixel

$R_{k,j}$ = spectral reflectance of endmember j on band k

ε_k = error of band k; n is the number of endmembers

Mean absolute error (MAE) is used to assess the unmixing accuracy by calculating the absolute difference between the estimated and reference fractions of corresponding land cover class. It can be expressed using equation (2).

$$MAE = \sum_{i=1}^{m} ABS(f_{e,i} - f_{r,i}) / m \qquad (2)$$

where $f_{e,i}$ = estimated fraction of sample i

$f_{r,i}$ = reference fraction of sample i

---

* Corresponding author

m = number of samples

2.2 Spectral transformations

We have examined nine transformed schemes including derivative analysis (DA) (Tsai and Philpot, 1998), principal component analysis (PCA) (Richards and Richards, 1999), independent components analysis (ICA) (Hyvarinen, 1999; Hyvärinen and Oja, 2000), Minimum Noise Fraction (MNF) (Boardman and Kruse, 1994; Green et al., 1988), Tasseled Cap (TC) (Kauth and Thomas, 1976), normalized spectral mixture analysis (NSMA) (Wu, 2004), and Tie spectral (Tie) transformation (Asner and Lobell (2000)). These transformation methods are summarized in Table 1.

Table 1. Spectral transformation

Transformation	Linearity	Reference
DA(1-3)	Linear	(Tsai and Philpot, 1998)
PCA	Linear	(Richards and Richards, 1999)
ICA	Linear	(Hyvarinen, 1999; Hyvärinen and Oja, 2000)
MNF	Linear	(Boardman and Kruse, 1994; Green et al., 1988)
TC	Linear	(Kauth and Thomas, 1976)
NSMA	Nonlinear	(Wu, 2004)
Tie	Nonlinear	(Asner and Lobell, 2000)

3. EXPERIMENTS

3.1 Study areas and data sources

Two regions were examined in this study, including Janesville, WI, United States, and Asheville, NC, United States. Janesville is located in the western shore of Lake Michigan, presenting as the humid continental climate. Flat plain is the major landscape in Janesville. Asheville is the largest city in western North Carolina. It is located in the Blue Ridge Mountains where two rivers, the Swannanoa River and French Broad River, merge together. Mountainous characteristics are significant in Asheville area. Residential buildings are constructed based on its local terrain. All study areas are mainly covered by commercial buildings, freeways, parking lots, residential houses, soil, and vegetation (tree and grass).

A scene of Landsat 8 Operational Land Imager (OLI) imagery (Janesville: June 3rd, 2014) and a scene of Landsat 5 Thematic Mapper (TM) (Asheville: June 2nd, 2009) imagery were employed in this study. Image preprocessing, including radiometric calibration, atmospheric correction using Fast Line-of-sight Atmospheric Analysis of Hypercube (FLAASH) with corresponding parameters, and reprojection to Universal Transverse Mercator (UTM) (Janesville: Zone 16; Asheville and Columbus: Zone 17), were applied. Historical high spatial resolution images (Janesville: June 12th, 2014; Asheville: May 30th, 2009) acquired on Google Earth were utilized to assess the mapping accuracy.
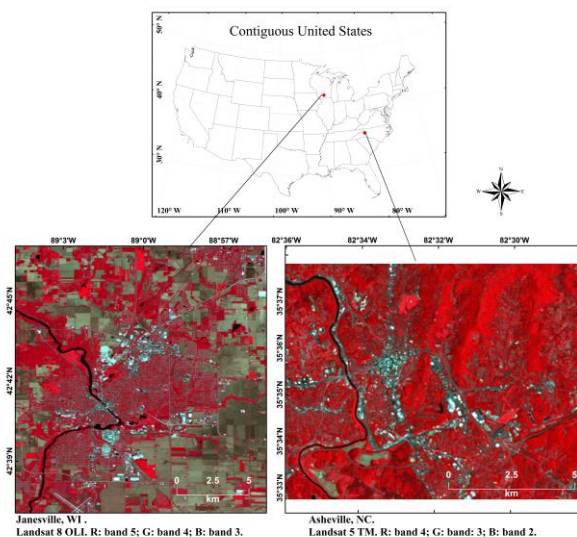

Figure 1. Study areas of Janesville WI, and Asheville NC.

3.2 Method

3.2.1 Spectral transformation: Spectral transformations were applied to the original data using the corresponding methods. In particular, DA1-3 (first to third derivative analysis), PCA, ICA, MNF, and TC were calculated directly from ENVI and its extension models. NSMA was computed using the method described in Wu (2004). Tie was created with the method stated in Asner and Lobell (2000).

3.2.2 Sample selection: Training samples (pure land cover classes' pixels), including vegetation (V), high albedo impervious surface area (ISAh), low albedo impervious surface area (ISAl), and soil (S), were selected carefully from the transformed images with the assistance of high spatial resolution images. The number of training samples of V, ISAh, ISAl, and S are 50, and 50 in study areas of Janesville, and Asheville respectively. Spectral libraries were constructed from the corresponding training samples.

64, and 60 testing samples of Janesville, and Asheville regions were collected to access each scheme's performance. Each testing sample is 3 pixels×3 pixels (90m ×90m) to avoid the geometric error impact acquired from reprojection and data acquisition. Fractions of impervious surface area within the testing samples were derived through digitizing the corresponding areas in high spatial resolution images.

3.2.3 SMA and accuracy assessment: Fully constraint linear spectral mixture analysis was applied to transformed and untransformed data 100 times with randomly selected spectra in corresponding spectral library under the endmember model of V-ISAh-ISAl-S.

The performance of each transformed scheme was evaluated through the mean absolute error (MAE). Estimated fractions of ISA are calculated by the sum of ISAh's and ISAl's fractions in the same pixel.

Boxplots were applied to illustrate the general performance of each scheme. Significant difference of the mean MAEs between transformed and untransformed scheme was tested by Paired-Samples T test. Mean differences, which indicate the general

performance between transformed and untransformed scheme, is calculated by using mean MAEs of untransformed subtracting the transformed schemes'. Positive value means the mean MAE of untransformed scheme is larger than transformed scheme's while negative value represents the opposite consequence. Significant values indicate the significance of the test result.

3.3 Results

The detailed performance of transformed and untransformed schemes were presented in the boxplots (Figures 2-3). Paired-Sample T test results were displayed in Table 2.

3.3.1 Janesville area: The median MAE of the untransformed scheme is around 0.11, which is less than most of transformed schemes except NSMA. NSMA has the lowest median MAE of 0.085. Median MAE values of PCA and TC are similar to that of the untransformed scheme (0.11). Transformed schemes of DA1-3, ICA, MNF, and Tie have slightly higher median MAEs than the untransformed scheme, varying from 0.12 to 0.14.

3.3.2 Asheville area: The median MAE of the untransformed scheme is the same as that in Janesville (0.10). NSMA still has a much lower median MAE than untransformed scheme. Schemes of DA1-3, MNF, and Tie have slightly lower median MAEs compared to the untransformed scheme (0.10) with values around 0.09.
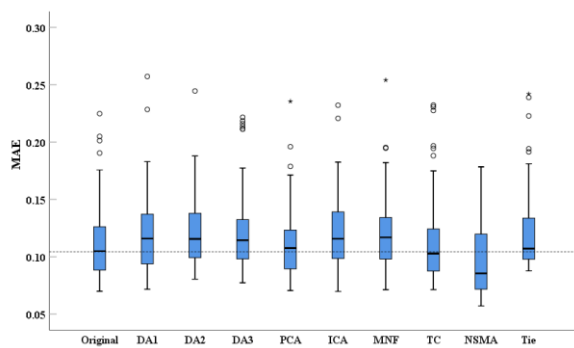

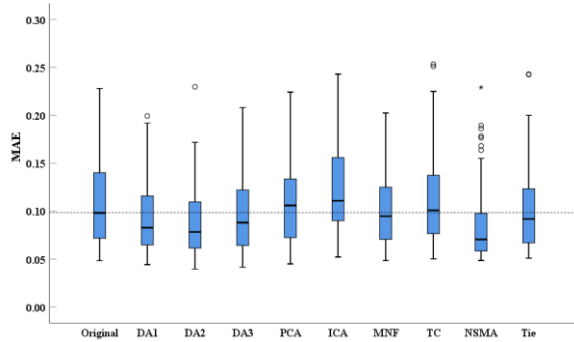Figure 2. Boxplot of MAE of Janesville

Figure 3. Boxplot of MAE of Asheville

Table 2 reveals that only NSMA demonstrates significant improvement with the mean MAEs lower than the untransformed scheme in all study areas (positive values in two areas). Paired-Samples T test also illustrates that the differences between NSMA and the untransformed scheme are significant as their p values are less than 0.05. Paired-Samples T test only indicates the significant difference of DA1-2 in Asheville while schemes of DA3, MNF, and Tie have p values larger than 0.05.

Other transformed schemes, such as TC and ICA are with lower accuracy as the mean differences are negative in all study areas. Moreover, significant difference between untransformed scheme and TC and ICA cannot be obtained since their p values are larger than 0.05.

Table 2. Results of Paired-Samples T test and comparison

Schemes	Mean difference		Sig. (2 tailed)	
	Jane.	Ash.	Jane.	Ash.
DA1	-0.008	**0.015**	0.054	**0.011**
DA2	-0.008	**0.020**	0.043	**0.001**
DA3	-0.009	**0.012**	0.059	0.058
PCA	**0.001**	-0.001	0.894	0.864
ICA	-0.009	-0.015	0.031	0.024
MNF	-0.009	**0.007**	0.069	0.219
TC	-0.001	-0.002	0.908	0.763
NSMA	**0.015**	**0.022**	**0.001**	**0.000**
Tie	-0.009	**0.008**	0.071	0.152

4. DISCUSSION

This study compared nine spectral transformations in two different study areas respectively. 100 times repeated test with different endmembers' spectra were employed to reveal the reliability of each scheme.

DA can get rid of unnecessary signal components and highlight minor absorption features by using spectral smoothing and feature reduction method. It, however, can also raise the possibility of ignoring essential spectral features (Youngentob et al., 2011). the performance of DA varied from places to places because of the existence of spectral variability. Unstable performance of the DA was concluded as no consistent results were occurred in this study.

PCA, MNF, and TC did not perform well in this study. The limitation of PCA and MNF may be attributed to the reason that many subtle material substances in Landsat imageries cannot be identified by second-order statistic (Wang and Chang, 2006), resulting in confusion between classes. Moreover, the last three bands of PCA, TC and MNF contain little variance. They may reduce the between-class variance and increase the within-class variance, adding more confusion during the fraction calculation.

ICA's results were opposite to the study of Wang and Chang (2006). It weakened the accuracy compared to second-order statistics-based methods like PCA and MNF. It may be attribute to the reason that ICA only conserves crucial and critical information such as anomalies, endmembers, and small targets instead of variance which preserved by PCA and MNF (Wang and Chang, 2006). However, the pattern of PCA and MNF are also not clear. Because results indicated that their performance varied from places to places as well.

NSMA is effectively addressing the confusion between impervious surface area and soil effectively since between-class variance between soil and impervious surface increased after applying the NSMA. Moreover, the shade effect was removed by brightness normalization. NSMA has a similar performance

in two study areas, proving the stability of NSMA in urban and suburban environment.

5. CONCLUSIONS

This study examined the performance of nine linearly and nonlinearly spectral transformations in Janesville and Asheville respectively by comparing their MAE with untransformed scheme. Many transformed schemes except NSMA could not illustrated a stable and exceeding performance in two study areas. NSMA showed a consistent performance in two study areas. Paired-Samples T test also indicated NSMA's significance in reducing the MAE. DA1-3, MNF, and Tie only improved the accuracy in one area but weakened the result in another region. TC and ICA weakened the performance in all study areas. Therefore, we concluded that NSMA can be applied in the urban areas to improve the accuracy of SMA.

REFERENCES

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

Boardman, J.W., & Kruse, F.A. (1994). Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada. In, Proceedings of the Thematic Conference on Geologic Remote Sensing (pp. I-407): Environmental Research Institute of Michigan

Green, A.A., Berman, M., Switzer, P., & Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing, 26, 65-74

Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE transactions on Neural Networks, 10, 626-634

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. Neural Networks, 13, 411-430

Kauth, R.J., & Thomas, G. (1976). The tasselled cap--a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In, LARS Symposia (p. 159)

Li, S., Kwok, J.T., & Wang, Y. (2002). Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images. Information Fusion, 3, 17-23

Portigal, F., Holasek, R., Mooradian, G., Owensby, P., Dicksion, M., & Fene, M. (1997). Vegetation classification using red edge first derivative and green peak statistical moment indices with the Advanced Airborne Hyperspectral Imaging System(AAHIS). In, International Airborne Remote Sensing Conference and Exhibition- Development, Integration, Applications & Operations, 3 rd, Copenhagen, Denmark

Richards, J.A., & Richards, J. (1999). Remote sensing digital image analysis. Springer

Tsai, F., & Philpot, W. (1998). Derivative analysis of hyperspectral data. Remote Sensing of Environment, 66, 41-51

Wang, J., & Chang, C.-I. (2006). Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. IEEE Transactions on Geoscience and Remote Sensing, 44, 1586-1600

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

Youngentob, K.N., Roberts, D.A., Held, A.A., Dennison, P.E., Jia, X., & Lindenmayer, D.B. (2011). Mapping two Eucalyptus subgenera using multiple endmember spectral mixture analysis and continuum-removed imaging spectrometry data. Remote Sensing of Environment, 115, 1115-1128

Zhang, J., Rivard, B., Sánchez-Azofeifa, A., & Castro-Esau, K. (2006). Intra-and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. Remote Sensing of Environment, 105, 129-141