# Noise Effect on Linear Spectral Unmixing

P. Gong* and A. Zhang†

*Center for Assessment and Monitoring of Forest and Environmental Resources (CAMFER)
Department of Environmental Science, Policy, and Management
151 Hilgard Hall, University of California, Berkeley, CA 94720-3110, USA
†Department of Geomatics Engineering, University of Calgary, Calgary, Alberta
Canada, T2N 1N4

**Abstract**
Using hyperspectral reflectance data collected from six types of surface covers, we synthesized linear mixtures and used them to test the sensitivity of two linear unmixing algorithms to simulated additive noise. We found both algorithms were highly sensitive to noise. This may considerably limit their use in remote sensing.

## I. INTRODUCTION

Pure materials can be measured under controlled conditions in laboratory. For remotely sensed data, however, different components may be included in one measurement causing spectral mixing. In order to identify various components and to determine their spatial proportions from remotely sensed data, one must properly model the spectral mixing process.

A popular method is the linear spectral mixing model [1]. With this model, some surface materials, called endmembers, components, or cover types, are considered to be spectrally mixed together with their areal proportions as weighting factors in one measurement. The physical assumption is that the amount of multiple scattering is not significant among the components in the measurement implying that most photons reaching the sensor interact with just one component. Under this assumption the energy received by a sensor can be approximated by the sum of the energy received from each component in an observation unit such as an image pixel. The inversion of this problem is usually done linearly (see [19][21] for nonlinear solutions), called linear spectral unmixing. A major purpose of unmixing remotely sensed data is to derive subpixel proportions of endmembers.

Linear unmixing was originally applied in geology (e.g., [2,4,8,23]) involving non-vegetative materials whose spectral mixing is at least intuitively close to a linear one. Recently, it was used in ecology, climatology, and urban land cover mapping involving data from sensors on board aircrafts and satellites such as the Landsat, SPOT and NOAA [7,11,12,22,25]. Inclusion of vegetation and human structure as scene components may violate the linear mixing assumption to some extent. Although non-linear models may be more appropriate in ecological and urban studies, they are hard to build, difficult to invert (e.g., [18]) and more dependent on scene components (e.g., models of forest canopies are different from those of buildings).

The simplicity of linear spectral mixing is attractive, but to what extent the linear assumption holds requires the knowledge of the exact form of radiative transfer of scene components. Related to this question is how scene components in neighboring pixels affect the pixel being unmixed. In addition, the atmosphere affects remotely sensed data, particularly in shorter wavelengths. How does the atmospheric interference affect linear unmixing results? Answers to these questions will be helpful for us to understand the reliability of linear unmixing results. In this paper, we present some tests that reveal the sensitivity of two unmixing methods to noise.

## II. LINEAR SPECTRAL MIXING MODELING AND ITS INVERSION - LINEAR SPECTRAL UNMIXING

Suppose there are $p$ endmembers in $m$ bands of a remotely sensed image, and $r_{jk}$ represents the spectral reflectance of $k$th endmember at $j$th band, all the reflectance can be arranged in an $m \times p$ matrix $\mathbf{R}$. The linear spectral mixing model can be expressed as:

$$\begin{bmatrix} d_{i1} \\ d_{i2} \\ \vdots \\ d_{im} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mp} \end{bmatrix} \begin{bmatrix} f_{i1} \\ f_{i2} \\ \vdots \\ f_{ip} \end{bmatrix} \quad (1)$$

1082-4006/99/0501-52$3.00

or $\mathbf{d_i} = \mathbf{R} \bullet \mathbf{f_i}$
with the following constraints:

$$f_{ik} \geq 0 \text{ and } \sum_{k=1}^{p} f_{ik} = 1, \quad \text{for i=1,2,...,n,} \qquad (2)$$

where $n$ is the total number of image pixels, $\mathbf{d_i}$ is the spectral responses measured from the $i$th pixel which is recorded by sensors in $m$ bands, each column vector in $\mathbf{R}$ corresponds to an endmember, and $f_{ik}$ denotes the fractional area of the $k$th endmember in the $i$th pixel. All the fractions of endmembers in pixel $i$ compose a $p$ *x 1* vector $\mathbf{f_i}$ and they should sum to one.

For the unmixing of a mixture pixel, three types of parameters in Equation (1) are of interest: (a) the total number of endmembers in the pixel, $p$; (b) the spectral identity of endmember $k$ in $m$ bands, i.e., all the elements in the $k$th column of $\mathbf{R}$ ($r_{jk}$, j = 1, 2, ..., $m$) that can be used to plot a spectral curve of endmember k; (c) the proportion of endmember $k$ in the pixel, $f_{ik}$.

The solution to Equation (1) is limited by the knowledge of these parameters. If parameters (a) and (b) are known, it is possible to determine (c) pixel by pixel. This is a typical situation used to derive spatial proportions of various endmembers from remotely sensed data. It is easy to get the proportional solution by inverting matrix $\mathbf{R}$. If (a) and (c) are known, (b) can be obtained from $\mathbf{d_i}$ and $\mathbf{f_i}$ [14]. This method is applicable to situations where available ground measurements of $\mathbf{f_i}$ are used to derive $\mathbf{R}$. When only (a) is known, or none of the three types of parameters is known, it is still possible to estimate these mixture parameters from a simultaneous analysis of many pixels though the mixing proportions in these pixels must vary from pixel(s) to pixel(s). Some methods have been proposed to achieve the solution including the use of principal component analysis [24], factor analysis [16], and mathematical programming [19].

When the number and the spectral identities of endmembers are known, there exist three possible cases to extract proportions of each endmember: (a) m+1<p; (b) m+1=p; (c) m+1>p. For case (b), it is a simple case to solve Equation (1) with the constraint Equation (2). In case (a), an image acquired from sensors may have a smaller number of spectral bands than the number of endmembers. Under such circumstances, Equation (1) becomes underdetermined. In order to obtain the spatial proportions for each image pixel, an optimal solution has to be explored. Some researchers recover proportions $\mathbf{f}$ using the geometrical scattering structure of data in multispectral space in combination with evidence such as the tasseled cap behavior of vegetation in a red-infrared scatter plot [15]. For data acquired with imaging spectrometers, dozens to hundreds of bands for each image pixel are available. Under this condition, equation (1) is mostly overdetermined as in case (c). In this study, we focus on case (c). To solve (1), the direct inversion is

$$f = (R^T R)^{-1} R^T \bullet d \qquad (3)$$

To avoid singular results, a Singular Value Decomposition (SVD) algorithm may be used [5][20]. Since these inversion methods do not guarantee that $\mathbf{f}$ in (2) are between 0 and 1, they are called unconstrained methods.

In a constrained approach, proportions $f_i$ can be solved through a nonnegative least squares strategy. The nonnegative least squares method is a special case of least squares problem with linear inequality constraints on the solution. This problem is defined as

Minimize $\| \mathbf{Rf - d} \|$ subject to $0 \leq \mathbf{f} \leq 1$. (4)

Details on solving (4) are found in [17].

In this study, we used both the unconstrained SVD and the nonnegative least squares methods. Since simulated data were used, the true fraction for $f_j$ endmember $j$ was known. We used the root mean squared error, RMS, as a measure to evaluate unmixing results,

$$RMS = \sqrt{\frac{1}{p}\sum_{j=1}^{p}(\bar{f}_j - f_j)^2} \qquad (5)$$

where $\bar{f}_j$ corresponds to the unmixing result for endmember $j$.

## III. ENDMEMBERS, SIMULATED ADDITIVE NOISE AND UNMIXING EXPERIMENTS

Spectral reflectance of six endmembers was extracted from a hyperspectral image acquired with a compact airborne spectrographic imager (casi) in Oregon [13]. Raw casi data were calibrated [3] and converted from radiance to reflectance [10]. The spectral interval between successive bands is approximately 1.8 nm. After smoothing the data with a 1 X 7 average filter, spectral curves of the six endmembers are shown in Figure 1 with a total of 210 spectral bands. The six endmembers include three types of gravel at different locations, two types of grass and ponderosa pines. Gravel_1 is relatively pure while Gravel_2 and Gravel_3 are slightly different from Gravel-1 and contaminated by different coverages of tuft grass. The two types of grass, Grass_1 and Grass_2, are snow grass and tuft grass, respectively. In this study, the six endmembers serve only as an example data set for the purpose of testing the sensitivity of linear unmixing methods to noise.

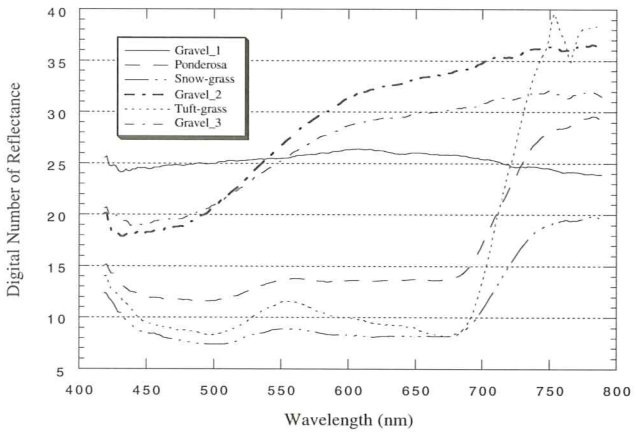


**Figure 1**. Smoothed CASI reflectance curves of six endmembers

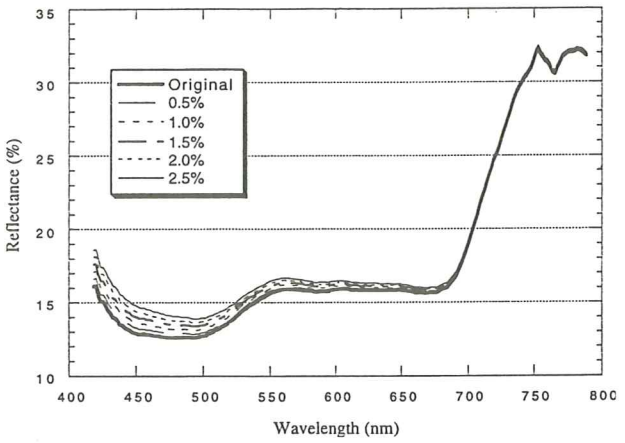


**Figure 2.** Simulated Rayleigh scattering added to an artificially created linear mixture spectra

Using these endmembers, we constructed **R** in equation (1) and tested the reliability of the two algorithms. Replacing $\mathbf{d_i}$ with each endmember and then solving equation (1) using either method, we can obtain a 1 corresponding to the endmember used in place of $\mathbf{d_i}$ and 0 proportions for other endmembers. Simulating an observation vector **d** by linearly mixing the six endmembers in **R** using proportions of 0.05, 0.27, 0.11, 0.07, 0.35, and0.15, respectively, the two methods can unmix **d** to recover these proportions correctly. **d** is presented using a bold solid line in Figure 2. Using a smaller number of spectral bands and repeating the above tests, we can still get the correct results as long as the number of bands is greater than the number of endmembers. These tests indicate that if no noise exists in the observation vector **d**, both algorithms are accurate. The question is how reliable these methods are when noise exists in **d**.

We only experimented with observation vectors containing small additive noise approximating upward atmospheric path radiance (Figure 2). When reflectance of endmembers is measured in field or in laboratory, the atmospheric effect on such measurements may be ignored. Remotely sensed data, however, are affected by atmospheric scattering particularly at shorter wavelengths if the sky condition is clear. We simulated five upward atmospheric path radiance curves in quantities of reflectance based on Rayleigh scattering [6,9,26]. At 420 nm, the simulated additive noise is at the level of 0.5%, 1%, 1.5%, 2%, and 2.5%, respectively. While noise patterns in radiance may be different from those in reflectance, it is not a concern in this study because the simulated noise levels are lower than what might be in reality. The curves in Figure 2 presented in thin or dashed lines represent the simulated noise added to the observation vector along the bold solid spectral curve.

Table 1 shows the noise effect on the linear unmixing results. Indices 1 - 5 correspond to the above five levels of additive noise in an increasing order. As one would expect, when the magnitude of additive noise increases, the derived proportions of most endmembers become less accurate. From the RMS errors, we can see that the unconstrained and the constrained unmixing methods resulted in similar levels of errors. Surprisingly, a less than 2.5% reflectance of additive noise has led to a greater than 20% of RMS error in proportion (or an error greater than 15% in arithmetic average of proportions) with either method. For individual endmembers, it can be seen that with noise at the level of index 5 contaminating the original mixture the proportion for endmember Grass_1 is almost 4 times overestimated. Examining the change of curve shapes in Figure 2, we can see that the shape at shorter wavelengths becomes more and more similar to Grass_1 (snow grass in Figure 1) as the noise level increases. This indicates that noise causing the shape changes of spectral curves can dramatically influence the unmixing results. The implication of this experiment is that accurate radiometric calibration and atmospheric correction must be done in order to obtain reliable results using these methods. Unfortunately, atmospheric correction errors smaller than 2.5% in reflectance are difficult to achieve in practice. This places a strong restriction on the use of endmembers measured in laboratory or in field for linearly unmixing remotely sensed data acquired from aircraft or satellite platforms.

Thus far, we used all 210 spectral bands in the experiment. They are highly redundant for estimating the proportions of six endmembers. Five band selection methods have been developed and tested [26]. In this study, we used some of band selection results from two of those methods, method A and method B,

**Table 1.** Linear unmixing results from data contaminated by additive noise

| Endmember | Gravel1 | P. Pine | Grass1 | Gravel2 | Grass2 | Gravel3 | RMS |
|---|---|---|---|---|---|---|---|
| Real Prop. | 0.05 | 0.27 | 0.11 | 0.07 | 0.35 | 0.15 | 0.000 |
| Index | Unconstrained Unmixing Method | | | | | | |
| 1 | 0.06 | 0.23 | 0.19 | 0.04 | 0.34 | 0.17 | 0.040 |
| 2 | 0.08 | 0.18 | 0.28 | 0.02 | 0.33 | 0.19 | 0.084 |
| 3 | 0.09 | 0.14 | 0.36 | -0.01 | 0.32 | 0.21 | 0.124 |
| 4 | 0.11 | 0.10 | 0.45 | -0.04 | 0.31 | 0.23 | 0.167 |
| 5 | 0.12 | 0.06 | 0.53 | -0.07 | 0.30 | 0.24 | 0.206 |
| Index | Constrained Unmixing Method | | | | | | |
| 1 | 0.06 | 0.23 | 0.19 | 0.04 | 0.34 | 0.17 | 0.040 |
| 2 | 0.08 | 0.18 | 0.28 | 0.02 | 0.33 | 0.19 | 0.084 |
| 3 | 0.10 | 0.14 | 0.37 | 0.00 | 0.32 | 0.19 | 0.125 |
| 4 | 0.12 | 0.09 | 0.46 | 0.00 | 0.31 | 0.17 | 0.167 |
| 5 | 0.15 | 0.04 | 0.55 | 0.00 | 0.31 | 0.15 | 0.209 |

Note: P. Pine— Ponderosa Pine; Real Prop.—Real Proportion

as presented in Figures 3. Figure 3a shows 20 selected bands clustered at the longer wavelengths whereas Figure 3b shows 20 selected bands spread over the entire spectral range. We used the selected bands to unmix the simulated noisy observation vectors in Figure 2. The results are presented in Tables 2 and 3 for the unconstrained and constrained unmixing methods, respectively. The five rows next to each method of band selection in Tables 2 and 3 correspond to the five indices in Table 1. By comparing the RMS errors listed in Tables 2 and 3, it seems that the two unmixing methods again have similar performances. When these are compared with the RMS errors in Table 1, it seems that the reduced data set produced better unmixing results. This is particularly true for the results obtained with band selection method A due to the fact that the magnitude of noise in the bands of longer wavelengths selected by method A is much smaller than that in many bands selected by method B.

## IV. SUMMARY AND CONCLUSIONS

As linear spectral unmixing algorithms become widely used, there are few studies focused on their validity and sensitivity to noise existing in the data to be unmixed. We selected two linear unmixing algorithms, an unconstrained method based on singular value decomposition and a constrained method based on nonnegative least squares, to test their sensitivity to noise. By linearly mixing spectral curves of six types of surface covers acquired with a compact airborne spectrographic imager, we obtained a spectral mixture. We then added simulated noise to the linearly mixed data set.

By unmixing the noisy data, we found that the two linear unmixing algorithms are very sensitive to noise. The effect of additive noise at less than 2.5% reflectance level is strong enough to cause an average error of greater than 15% in estimated endmember pro-

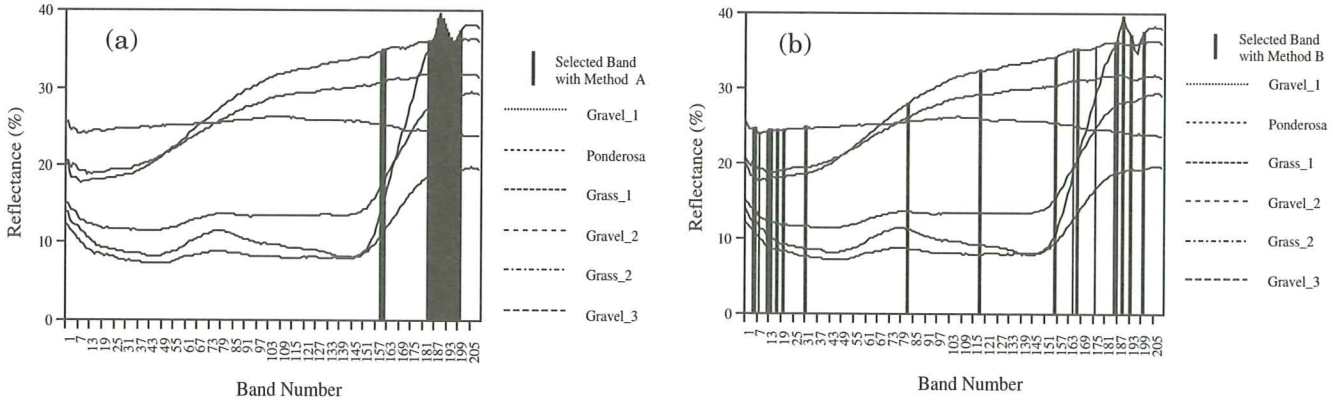


**Figure 3.** Spectral bands selection. a) Selected bands with method A; b) Selected bands with method B.

**Table 2.** Linear unmixing results from the unconstrained method with selected bands

| Endmember | Gravel1 | P. P. | Grass1 | Gravel2 | Grass2 | Gravel3 | RMS |
|---|---|---|---|---|---|---|---|
| Method A | 0.07 | 0.28 | 0.11 | 0.05 | 0.35 | 0.15 | 0.012 |
| | 0.09 | 0.28 | 0.10 | 0.04 | 0.35 | 0.16 | 0.022 |
| | 0.10 | 0.29 | 0.10 | 0.02 | 0.35 | 0.16 | 0.031 |
| | 0.12 | 0.29 | 0.09 | 0.01 | 0.35 | 0.17 | 0.040 |
| | 0.14 | 0.30 | 0.09 | -0.01 | 0.35 | 0.17 | 0.052 |
| Method B | 0.07 | 0.23 | 0.18 | 0.05 | 0.34 | 0.16 | 0.035 |
| | 0.09 | 0.20 | 0.25 | 0.02 | 0.34 | 0.17 | 0.070 |
| | 0.11 | 0.16 | 0.32 | 0.00 | 0.33 | 0.18 | 0.105 |
| | 0.13 | 0.13 | 0.39 | -0.03 | 0.33 | 0.19 | 0.139 |
| | 0.16 | 0.09 | 0.45 | -0.05 | 0.32 | 0.20 | 0.172 |

Note: P. P.--Ponderosa Pine

**Table 3.** Linear unmixing results from the constrained method with selected bands

| Endmember | Gravel1 | P. P. | Grass1 | Gravel2 | Grass2 | Gravel3 | RMS |
|---|---|---|---|---|---|---|---|
| Method A | 0.07 | 0.28 | 0.11 | 0.05 | 0.35 | 0.15 | 0.012 |
| | 0.09 | 0.28 | 0.10 | 0.04 | 0.35 | 0.16 | 0.022 |
| | 0.10 | 0.29 | 0.10 | 0.02 | 0.35 | 0.16 | 0.031 |
| | 0.12 | 0.29 | 0.09 | 0.01 | 0.35 | 0.17 | 0.040 |
| | 0.14 | 0.30 | 0.09 | 0.00 | 0.35 | 0.16 | 0.049 |
| Method B | 0.07 | 0.23 | 0.18 | 0.05 | 0.34 | 0.16 | 0.035 |
| | 0.09 | 0.20 | 0.25 | 0.02 | 0.34 | 0.17 | 0.070 |
| | 0.11 | 0.16 | 0.32 | 0.00 | 0.33 | 0.18 | 0.105 |
| | 0.14 | 0.11 | 0.40 | 0.00 | 0.33 | 0.15 | 0.143 |
| | 0.17 | 0.06 | 0.49 | 0.00 | 0.32 | 0.13 | 0.187 |

Note: P. P.--Ponderosa Pine

portions. Unmixing results obtained from the two methods tested in this study could be rather unreliable. We suggest that one must be very careful when choosing to use either of the two methods in remote sensing. Our examples illustrate that band selection may be useful to reducing unmixing errors. Selecting bands that are less affected by noise helps to improve the reliability of unmixing results.

## ACKNOWLEDGMENTS

This research was partially supported by a grant from Natural Science and Engineering Research Council of Canada and a start-up grant from the University of California, Berkeley, USA.

## REFERENCES

[1] Adams, J. B., Smith, M.O., Gillespie, A.R., 1989, Simple models for complex natural surfaces: a strategy for the hyperspectral era of remote sensing, *Proceedings of the IGARSS'89*, Vancouver, British Columbia, Canada, pp.16-21.

[2] Adams, J. B., Smith, M.O., P.E. Johnson, 1986, Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site, *Journal of Geophysical Research*, 91: 8098-8112.

[3] Babey, S. K. and R.J. Soffer, 1992, Radiometric calibration of the compact airborne spectrographic imager (CASI), *Canadian Journal of Remote Sensing*, 18(4):233-242.

[4] Blunt, G., M.O. Smith, J.B. Adams, R. Greeley, P.R. Christensen, 1990, Regional aerolian dynamics and sand mixing in the Gran Desierto: evidence from Landsat Thematic Mapper images, *Journal of Geophysical Research,* 95: 15463-15482.

[5] Boardman, J. W., 1989, Inversion of imaging spectrometry data using singular value decomposition, *Proceedings of the IGARSS'89*, Vancouver, British Columbia, Canada, pp.2069-2072.

[6] Chavez, P. S., Jr., 1989, Radiometric calibration of Landsat thematic mapper multispectral images, *Photogrammetric Engineering & Remote Sensing,* Vol. 55, No. 9, pp. 1285 - 1294.

[7] Cross, A. M., J.J. Settle, N.A. Drake, Paivinen, 1991, Subpixel measurement of tropical forest cover using AVHRR data. *International Journal of Remote Sensing,* 12: 1119-1129.

[8] Ehrlich, R. and W.E. Full, 1987, Sorting out geology - unmixing mixtures. In W.B. Size, edited, *Use and*

*Abuse of Statistical Methods in the Earth Sciences.* Oxford Press: New York. 1987. pp. 33-46.

[9] Forster, B. C., 1984, Derivation of atmospheric procedures for Landsat MSS with particular reference to urban data. *International Journal of Remote Sensing*, 5:799-817.

[10] Freemantle, J. R., R. Pu, J.R. Miller, 1992, Calibration of imaging spectrometer data to reflectance using pseudo-invariant features, *Proceedings of the 15th Canadian Symposium on Remote Sensing,* June, 1992, Toronto, pp.452 - 455.

[11] Gong, P., J.R. Miller, J. Freemantle, B. Chen, 1991, Spectral decomposition of Landsat TM data for urban land-cover mapping, *Proceedings of the14th Canadian Symposium on Remote Sensing,* University of Calgary, Calgary, May 1991. pp. 458-461.

[12] Gong, P., J.R. Miller, M. Spanner, 1994, Forest canopy closure from classification and spectral unmixing of scene components - multisensor evaluation of an open canopy. *IEEE Trans. on Geos. and Remot. Sens.*. 32(5): 1067-1080.

[13] Gong., P., R. Pu, J.R. Miller, 1992, Correlating leaf area index of ponderosa pine with hyperspectral CASI data, *Canadian Journal of Remote Sensing*, Vol. 18, No. 4, pp.275-282.

[14] Hanan, N. P., S.D. Prince, P.H.Y. Hiernaux, 1991, Spectral modelling of multicomponent landscapes in Sahelm, *International Journal of Remote Sensing*, 12, pp. 1243-1258.

[15] Jansinski, M. F., P.S. Eagleson, 1990, Estimation of subpixel vegetation cover using red infrared scattergrams, *IEEE Transactions on Geoscience and Remote Sensing*, 28, pp.153-267.

[16] Klovan, J. E., A.T. Miesch, 1975, Extended CABFAC and QMODEL computer programs for Q-Mode factor analysis of compositional data, Computer and Geosciences, pp.161-178.

[17] Lawson, C. L., R.J. Hanson, 1974, Solving least squares problems, Prentice-Hall, Inc., Englewood Cliffs, N.J.

[18] Liang, S. and A.H. Strahler, 1993, An analytic BRDF model of canopy radiative transfer and its inversion. *IEEE Trans. Geosci. Remote Sens.*, 31(5):1081-1092.

[19] Liang, S., A.H. Strahler, X. Li, 1991, Simultaneous inversion of subpixel proportions and signatures of mixed pixels: two components, in *Proc. of ACSM-ASPRS Fall convertion*. Atlanta, Georgia, pp. B108-117.

[20] Press, W. H., S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, 1992, *Numerical recipes in C: the art of scientific computing*, Cambridge University Press: New York.

[21] Sasaki, K., S. Kawata, and S. Minami, 1983, Constrained nonlinear method for estimating component spectra from multicomponent mixtures. *Applied Optics*, 22(22):3599-3603.

[22] Shimabukuro, Y. E., J.A. Smith, 1991, The least squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Trans. on Geos. and Remote Sensing*. 29(1): 16-20,

[23] Singer, R. B., T.B. McCord, 1979, Mars: large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance, *Proceedings of the 10th Lunar and Planetary Science Conference*, pp.1835-1848.

[24] Smith, M. O., P.E. Johnson, J.B. Adams, 1985, Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis, *Journal of Geophysical Research,* 90:C797-C804.

[25] Smith, M. O., S.L. Ustin, J.B. Adams, A.R. Gillespie, 1990, Vegetation in deserts: 1. A regional measure of abundance from multispectral images, *Remote Sensing of Environment*, 31: 1-25.

[26] Zhang, A., *Hyperspectral Image Analysis: A Study on Band Selection Methods and the Sensitivity of Linear Spectral Unmixing.* Report in the Department of Geomatics Engineering, The University of Calgary, Alberta, Canada, 77p.