NWU-RS-LAND-004

Sensor-independent analysis method for hyper-multispectral data based on the pattern decomposition method

L.F.Zhang, S.Furumi, K.Muramatsu, N.Fujiwara,
M.Daigo and L.P.Zhang

2004.9.15

Sensor-independent analysis method for hyper-multispectral data based on the pattern decomposition method

L.F. Zhang[1], S. Furumi[2], K. Muramatsu[2,4], N. Fujiwara[2], M. Daigo[3], and L.P. Zhang[1]

[1] National Key Laboratory for Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, P. R. China
[2] Laboratory of Nature Information Science, Department of Information and Computer Sciences, Nara Women's University, Nara, Japan
[3] Department of Economics, Doshisha University, Kyoto, Japan
[4] KYOUSEI Science Center for Life and Nature, Nara Women's University, Nara, Japan

Abstract. The pattern decomposition method (PDM) is a type of spectral mixing analysis in which each pixel is expressed as the linear sum of fixed, standard spectral patterns for water, vegetation, and soil, with supplementary patterns included when necessary. For each sensor, these standard patterns are calculated from the same original reflectance spectra. The PDM framework can be applied to any optically sensed data. In this method, however, the normalization of standard spectral patterns depends on how many bands and which wavelengths the sensor observes. As a result, the obtained pattern decomposition coefficients may differ for each sensor, even when the same sample object is observed. This paper describes a modified PDM with a supplementary pattern; the developed sensor-independent method uses the same normalized spectral patterns for any sensor. Fixed multi-band (1,260 bands) spectra served as the universal standard spectral patterns. The resulting pattern decomposition coefficients showed sensor independence. That is, regardless of sensor, the three coefficients had nearly the same values for the same samples. The estimation errors for pattern decomposition coefficients depended on the sensor used. The estimation errors for Landsat/MSS and ALOS/AVNIR-2 were larger than those of Landsat/TM (ETM+), Terra/MODIS, and ADEOS-II/GLI. The latter three sensors had negligibly small errors.

1. Introduction

Satellite remote sensing is important for land cover research and estimating global net primary production (NPP). Numerous satellite sensors collect measurements for land analysis. Recent satellite sensors, such as Terra/MODIS and ADEOS-II/GLI, provide hyper-multispectral data. However, analysis results depend on sensor performance and are especially affected by the number of bands and wavelengths observed. Consequently, it is difficult to compare analysis results obtained using data from different satellite sensors.

We sought to develop a sensor-independent analysis method. Sensor independence means that analysis results for the same sample should be the same or nearly the same, regardless of the sensor used. Most analysis methods are sensor dependent. For example, although the principal component transformation (PCT) method can be applied to data obtained from any type of optical sensor, the results differ depending on the sensor type, so PCT is a sensor-dependent method.

The pattern decomposition method (PDM) uses multispectral data effectively (Fujiwara *et al*., 1996; Muramatsu *et al*., 2000; Daigo *et al*., 2003). The PDM is a type of spectral mixing analysis (Adams *et al*., 1995), which expresses the spectrum of each pixel as the linear sum of three fixed, standard spectral patterns (*i.e*., the patterns of water, vegetation, and soil). For each sensor, these standard patterns are calculated from the same original reflectance spectra. This framework can be used to apply the PDM to data obtained from any optical sensor. However, the resulting pattern decomposition coefficients may differ by sensor, even for the same sample object. Normalization methods account for these differences. The PDM has sensor-independent characteristics and can be made sensor-independent with a well-designed definition. Therefore, we developed a modified PDM for sensor-independent analysis. This method, referred to as the universal PDM (UPDM), uses the same normalization method for all sensors. In this method, the standard patterns are extracted from ground-measured common spectra. The UPDM can be applied to any sensor, and the results obtained from various sensors are almost the same for the same sample.

On average, 95.5% of land-cover spectral reflectance information can be transformed into the three decomposition coefficients and decomposed into the three standard patterns with about 4.2% error per degree of freedom (Daigo *et al*., 2003). However, some objects, such as yellow leaves, have slightly larger decomposition errors. Depending on the research purpose, supplementary standard patterns can be applied. For example, to study vegetation changes in more detail, an additional supplementary spectral pattern can be added to reproduce the spectral reflectance. In this study, a yellow-leaf pattern was used as a supplementary spectral pattern.

To study sensor independence using the UPDM, we analyzed about 600 ground-measured samples, including green-leaf, yellow-leaf, dead-leaf, soil, water, and concrete samples etc.. We made simulated data of Landsat/MSS, ALOS/AVNIR-2, Landsat/ETM+, Terra/MODIS, and ADEOS-II/GLI sensors using ground-measured data, and compared the analyses results of these data. In this paper, we describe the UPDM and simulated results representing several sensors.

2. The pattern decomposition method

2.1. *Review of the conventional pattern decomposition method*

In the conventional PDM, the reflectance (or brightness) data for each observed pixel are decomposed into the standard spectral patterns of water, vegetation, and soil using the following formula (Daigo *et al*., 2003):

$$R_i \rightarrow C_w \cdot P_{iw} + C_v \cdot P_{iv} + C_s \cdot P_{is}, \quad (1)$$

where R_i is the reflectance of band i measured on the ground (or by satellite sensor), C_w, C_v, and C_s are the respective decomposition coefficients, and P_{iw}, P_{iv}, and P_{is} are the normalized standard spectral patterns of water, vegetation, and soil, which relate to the properties of each sensor. The fitted reflectance value is given by

$$R_i^{'} = C_w \cdot P_{iw} + C_v \cdot P_{iv} + C_s \cdot P_{is}, \quad (2)$$

where the residual of the band i reflectance is defined by

$$r_i = R_i - R_i^{'}. \quad (3)$$

The pattern decomposition coefficients obtained using equation (1) are evaluated using the least squares method. When the PDM is used for spectral mixing analysis, a non-negative constraint is put on the coefficients:

$$C_w \geq 0, \quad C_v \geq 0, \quad C_s \geq 0 \quad (4)$$

The standard spectral patterns are normalized as

$$\sum_{i=1}^{N} \left| P_k(i) \right| = 1 \qquad (k = w, v, s). \quad (5)$$

Here, N is the number of sensor bands used. If necessary for more detailed analysis, a supplementary standard spectral pattern can be added.

According to equation (5), the standard patterns depend on the band wavelengths detected and the number of bands used, although the same continuous spectra serve as the standard spectra. As a result, the pattern decomposition coefficient values for the same object depend on the sensor (*i.e.*, the values are sensor dependent).

2.2. *The universal pattern decomposition method*

2.2.1. The universal pattern decomposition method with three components

In response to PDM shortcomings, we developed a universal PDM (UPDM). In the UPDM, we define each standard spectral pattern as a continuous spectral function from 350 to 2,500 nm. The other equations remain the same as in the conventional PDM. Instead of equation (5), however, standard pattern normalizations are defined as follows:

$$\int \left| P_k(\lambda) \right| d\lambda = \int d\lambda \quad (k = w, v, s), \quad (6)$$

where the discrete band number i in equation (5) is changed into continuous wavelength λ, $\int d\lambda$ refers to integration of the total wavelength range, and $P_k(\lambda)$ is defined as

$$P_k(\lambda) = \frac{\int d\lambda}{\int |R_k(\lambda)| d\lambda} R_k(\lambda), \tag{7}$$

where $R_k(\lambda)$ represents the spectral reflectance patterns of standard objects. The shapes and magnitudes of the standard patterns $P_k(\lambda)$ are fixed for all sensors.

For each sensor band, we intercepted $P_k(\lambda)$ values. Thus, the standard patterns for each sensor are defined by

$$P_{ik} = \frac{\int_{\lambda_{si}}^{\lambda_{ei}} P_k(\lambda) d\lambda}{\int_{\lambda_{si}}^{\lambda_{ei}} d\lambda} \quad (k = w, v, s), \tag{8}$$

where λ_{si} and λ_{ei} are the start and end wavelengths for band *i*, respectively, and $\int_{\lambda_{si}}^{\lambda_{ei}} d\lambda$ is the wavelength width of band *i*. The decomposition coefficients C_k were obtained for each sensor by the least squares method using equation (1). In principle, nearly equal values should result for the same object; furthermore, coefficient precision is expected to improve as the number of bands increases.

The remaining decomposition coefficient calculation procedures are the same as in the conventional PDM. However, the non-negative constraint put on coefficients in the conventional PDM is removed in the UPDM. The PDM can be explained using two analysis methods: spectral mixing analysis and multi-dimensional analysis. If we think of the PDM as multi-dimensional analysis, standard patterns are interpreted as an oblique coordinate system, and coefficients are thought of as the coordinates of a pixel's reflectance. In this case, non-negative constraints are inappropriate. We think this interpretation is more general than traditional interpretations. If an application requires non-negative constraints, the UPDM is also compatible with the non-negative constraints.

If we compute an integral of the entire band, we can obtain equation (9) from equation (1), as follows:

$$C_w \int P_w(\lambda) d\lambda + C_v \int P_v(\lambda) d\lambda + C_s \int P_s(\lambda) d\lambda = \int R(\lambda) d\lambda. \tag{9}$$

Considering equation (6), equation (9) can be transformed into

$$C_w + C_v + C_s = \frac{\int R(\lambda) d\lambda}{\int d\lambda}. \tag{10}$$

In principle, nearly the same values should result from measurements of the same object, regardless of sensor type.

The reduced χ^2 is defined as follows:

$$\chi^2 = \sum_{i=1}^{n} r(i)^2 / (n-3), \tag{11}$$

where n refers to the total number of bands for a sensor, 3 is the number of coefficients, and $(n-3)$ represents the degrees of freedom for a dataset of n bands.

2.2.2. The Universal pattern decomposition method with four components

As mentioned above, a supplementary yellow-leaf spectral pattern was used for detailed analysis of vegetation changes. Consequently, the general equation of universal PDM was changed to

$$R_i \rightarrow C_w \cdot P_{iw} + C_v \cdot P_{iv} + C_s \cdot P_{is} + C_4 \cdot P_{i4}, \tag{12}$$

where C_4 represents the supplementary coefficients of a yellow leaf, and P_{i4} is the supplementary standard pattern for i bands (i is the number of bands for the sensor), which is intercepted from $P_4(\lambda)$. The yellow-leaf spectrum and three components were used to define $P_4(\lambda)$ as follows:

$$P_4(\lambda) = \frac{r_4(\lambda)\int d\lambda}{\int |r_4(\lambda)| d\lambda}, \tag{13}$$

where $r_4(\lambda)$ is the residual yellow-leaf value relative to i bands:

$$r_4(\lambda) = R_4(\lambda) - \{C_w P_w(\lambda) + C_v P_v(\lambda) + C_s P_s(\lambda)\}. \tag{14}$$

Here, $R_4(\lambda)$ is the measured value for the yellow-leaf sample, and $r_4(\lambda)$ is the residual value. For any sensor, P_{i4} values were calculated with equation (8). The reduced χ^2 were calculated using

$$\chi^2 = \sum_{i=1}^{n} r(i)^2 / (n-4), \tag{15}$$

where n equals the number of bands.

3. Data used in this analysis

3.1. *Standard spectral patterns*

Consistent with the conventional PDM, we used the same water, vegetation, and soil standard spectral patterns. In the UPDM, however, we converted an uninterrupted spectral wavelength range from 350 to 2,500 nm, excluding regions of strong atmospheric absorption. Therefore, the total number of bands equaled 1,260. Sample data were measured outdoors under solar light or indoors under a halogen lamp with a Field Spec FR (Analytical Spectral Devices, Inc.) or MSR7000 (Opto Research Corp.) radio-spectrometer. Both radio-spectrometers give raw spectral values for every 1 nm in wavelength from 350 to 2,500 nm and have a spectral resolution from 3 to 10

nm. The samples used to define standard spectral patterns were the same as the reference (Daigo *et al*., 2003). Table 3.1 shows the wavelength regions used to obtain the standard spectral patterns. The spectral region used had atmospheric transmittance greater than 80%. Figure 3.1 shows the wavelengths selected from the original standard reflectance values for water, vegetation, and soil.

Table 3.1. Wavelength regions used in this analysis

No	Wavelength/nm
1	371.0 ~ 900.0
2	991.0 ~ 1100.0
3	1191.0 ~ 1300.0
4	1521.0 ~ 1750.0
5	2081.0 ~ 2360.0

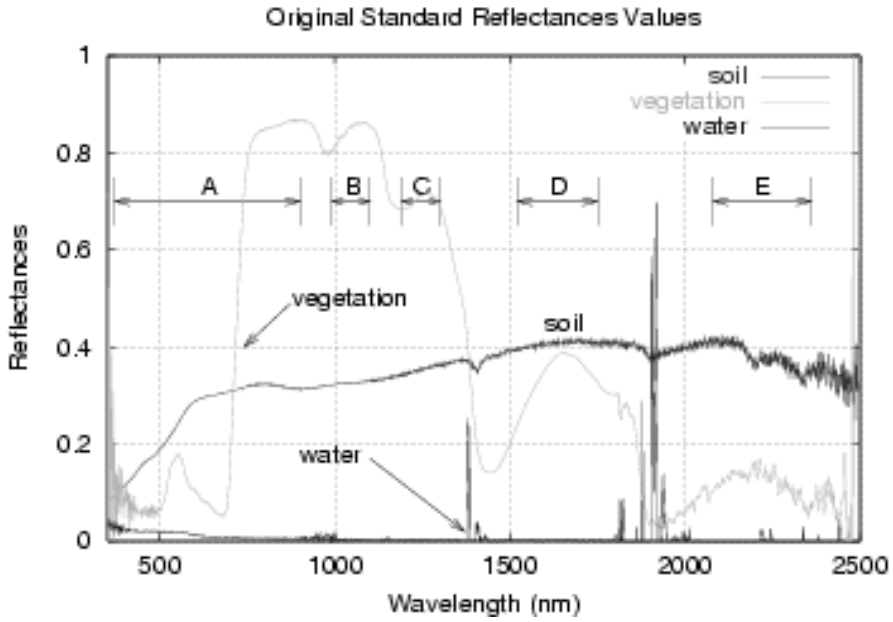


Figure 3.1. Wavelengths selected in this analysis. The figure shows the original measured reflectance values for the three standard patterns (water, vegetation, and soil).

In this analysis, we selected a yellow-leaf sample to serve as the supplementary spectral pattern. Using a yellow leaf rectifies problems associated with the withered-vegetation pattern. Dead-leaf samples are remarkably similar to soil and are therefore not suitable as a supplementary spectral pattern. Figure 3.2 shows the new normalized standard spectral patterns of water (blue ×), vegetation (green ×), soil (red +), and the yellow-leaf supplement (purple □).

3.2. *Sensors and sample data used for this analysis*

The sensors simulated in this analysis were MSS, ALOS, ETM+, MODIS, and GLI. The virtual sensors MODEL and CONTINUE were also used. Bands in which output signals from the actual sensor (*i.e.*, MODIS & GLI) were saturated on land areas were removed. In the wavelength region above 2,000 nm, the ground-measured data was poor in quality. Therefore, these data were also removed. For each band of each sensor, average reflectance values were obtained from ground-measured data within the wavelength widths of each band. Table 3.2 lists the bands used in this analysis. The model sensor CONTINUE has 92 bands with wavelengths ranging from 371 to 1,750 nm (table 3.1) and a bandwidth of 10 nm.

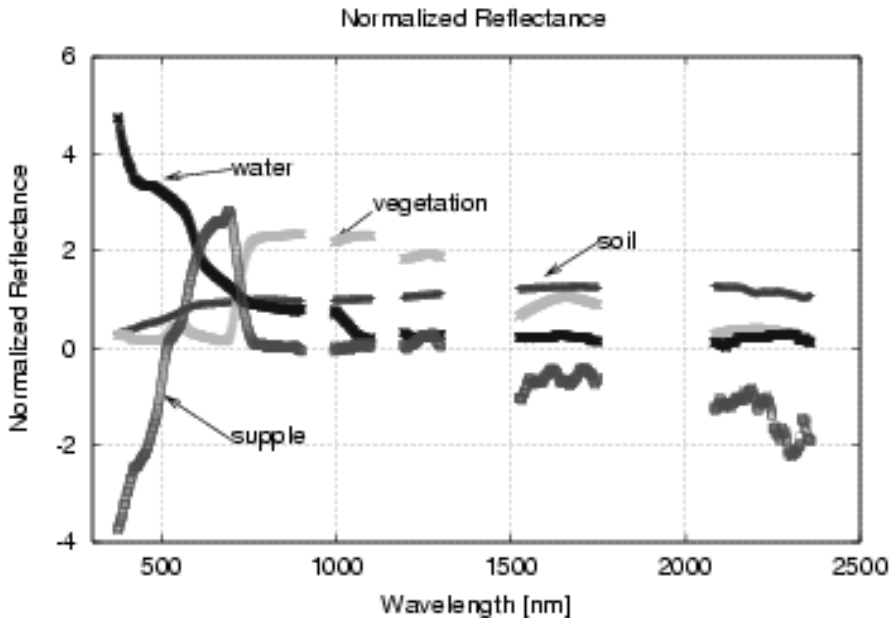


Figure 3.2. Normalized standard patterns of soil, water, vegetation, and the supplementary pattern. This figure shows the normalized values of the standard spectral patterns of water, vegetation, and soil given in figure 3.1; the supplementary standard pattern represents a yellow leaf.

Table 3.2. Spectral bands used in this analysis

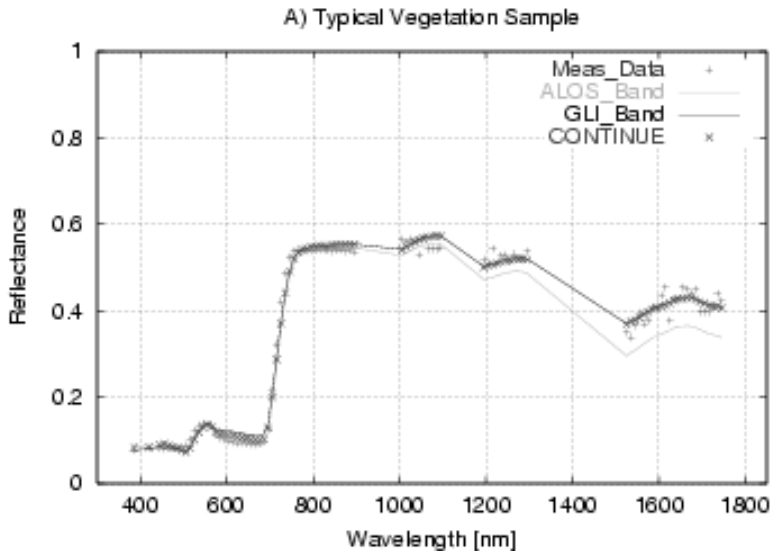
No	MSS/nm	ALOS/nm	ETM+/nm	MODIS/nm	GLI/nm	MODEL/nm
1	500.0～600.0	420.0～500.0	450.0～519.0	459.0～479.0	375.0～385.0	385.0～425.0
2	600.0～700.0	520.0～600.0	520.0～600.0	545.0～565.0	455.0～465.0	455.0～465.0
3	700.0～800.0	610.0～690.0	630.0～690.0	620.0～670.0	540.0～550.0	540.0～550.0
4	800.0～1,100.0	760.0～890.0	760.0～900.0	841.0～876.0	673.0～683.0	673.0～683.0
5			1,550.0～1,750.0	1,230.0～1,250.0	705.0～715.0	705.0～715.0
6			2,080.0～2,350.0	1,628.0～1,652.0	759.0～767.0	759.0～767.0
7				2,105.0～2,155.0	855.0～875.0	855.0～875.0
8					1,040.0～1,060.0	991.0～1,010.0
9					1,230.0～1,250.0	1,040.0～1,060.0
10					1,540.0～1,740.0	1,200.0～1,250.0
11					2,100.0～2,320.0	1,540.0～1,640.0
12						1,650.0～1,740.0
13						2,100.0～2,320.0

4. Results of the universal PDM

4.1. *Reproducibility of the observed spectra with universal PDM and reduced* χ^2

Figure 4.1 shows the reflectance spectra of 92 bands and the reconstructed spectra calculated from the pattern decomposition coefficients obtained for various sensors. In the figure, red rhombuses show the 92 bands of the ground-measured CONTINUE data. The blue broken line, green broken line, and purple crosses represent reconstructed spectra for the ALOS, GLI, and CONTINUE sensors, respectively. The original spectra are reproduced well, except those from ALOS. The reconstructed spectra for ETM+, MODIS, and MODEL are almost the same as the spectra for GLI and CONTINUE. For the ALOS sensor, the data points corresponding to the ALOS band wavelengths (499.5 ~ 1,100.5 nm) are reproduced exactly. However, above wavelengths of 1,000 nm, the reproduced values differ from the original data.

Since we have four fitting parameters, C_w, C_v, C_s, and C_4, the reflectance values at each band wavelength for the four-band sensors (MSS and ALOS) were reproduced exactly. The reduced χ^2 could not be calculated. Here, we compare the reconstructed spectrum of the 92 bands for each sensor with ground-measured CONTINUE data. Figure 4.2 shows the results of the reduced χ^2 calculated for the 92 bands. Approximately 600 samples were used. The values decrease as the band number increases and converge at 0.00066 for band numbers larger than five (TM). Table 4.3 lists the average values of the reduced χ^2 for each sensor. The square root of 0.00062 is 0.025 (2.5%), which is the fitting error per degree of freedom.



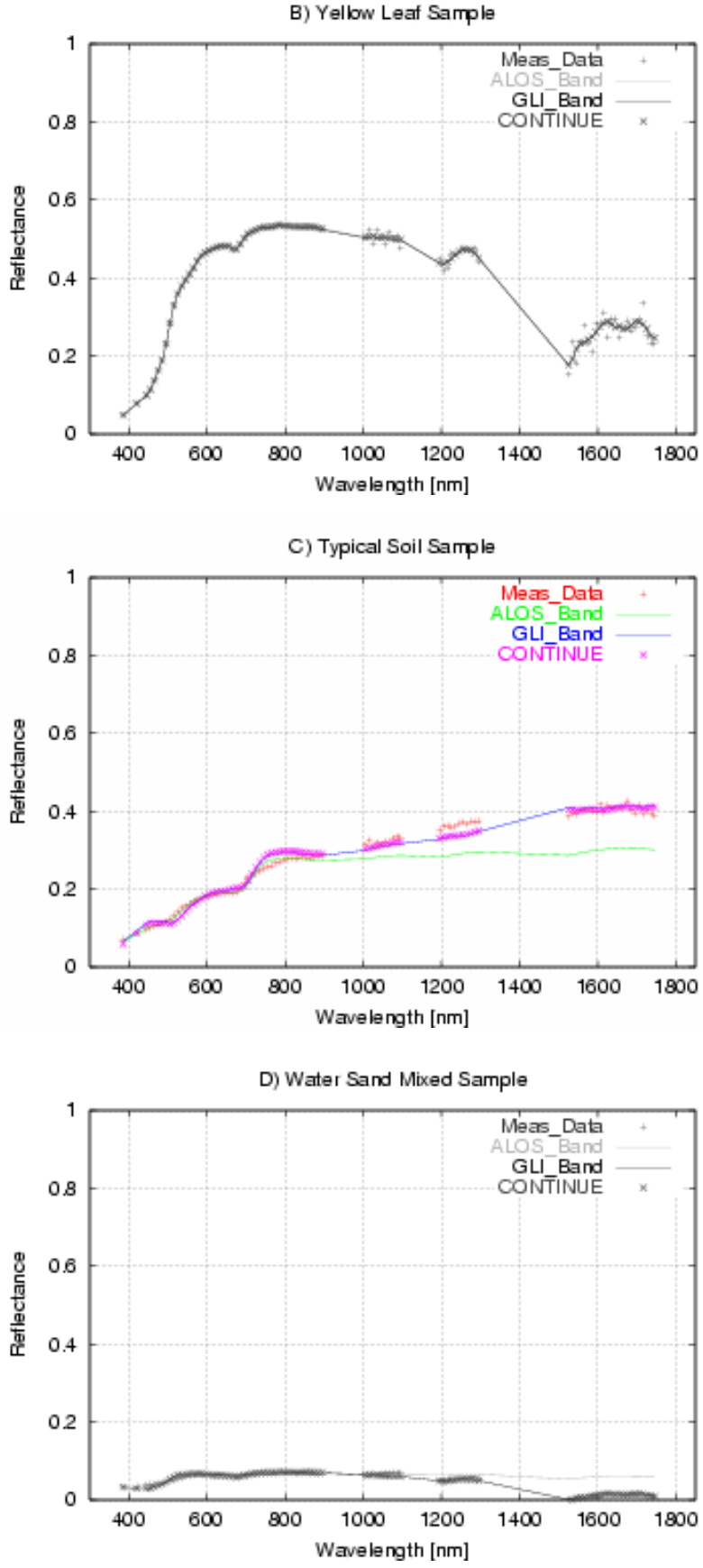


Figure 4.1. Original reflectance and reconstructed reflectance spectra. A) Typical vegetation

sample, B) Yellow leaf sample, C) Typical soil sample, D) Water sand mixed sample.

Table 4.1. Average values and root mean squares (rms) of the reduced χ^2

Sensor	MSS	ALOS	ETM+	MODIS	GLI	MODEL	CONTINUE
χ^2	0.20861	0.01404	0.00083	0.00069	0.00068	0.00065	0.00062
rms	0.457	0.118	0.029	0.026	0.026	0.025	0.025

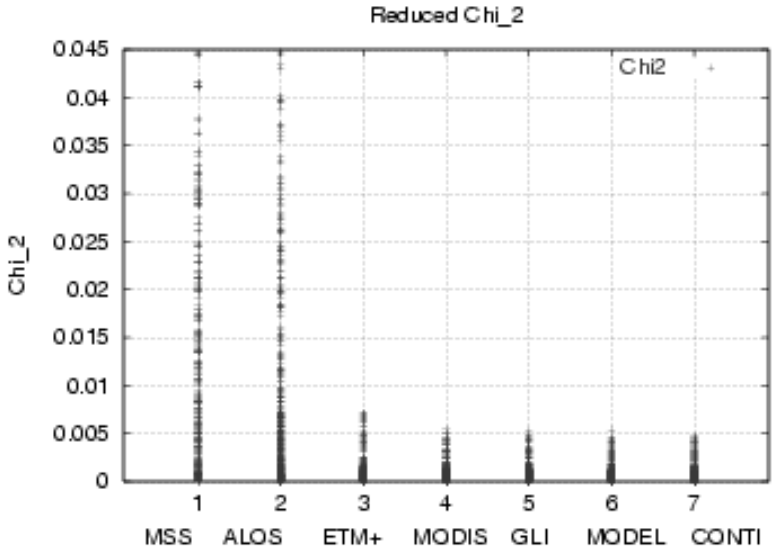
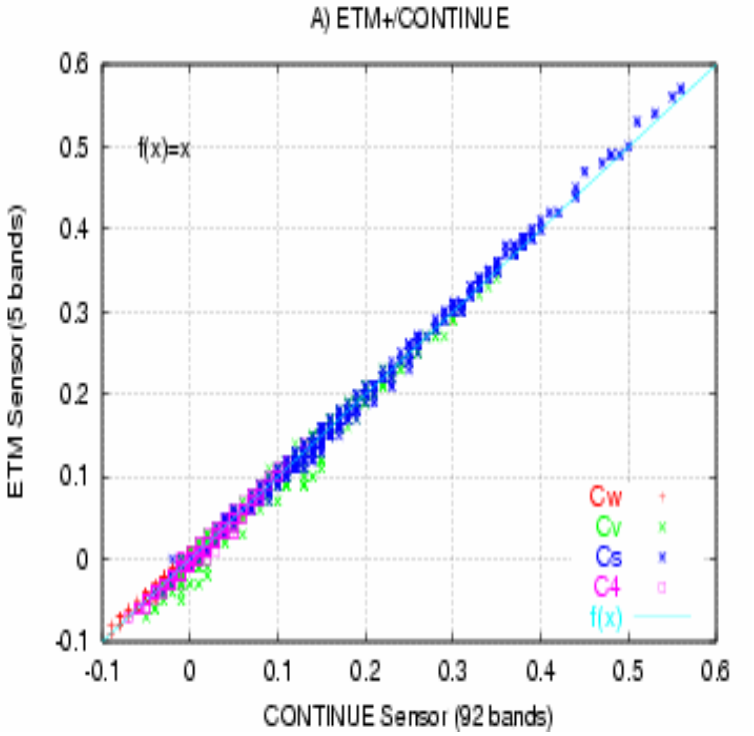
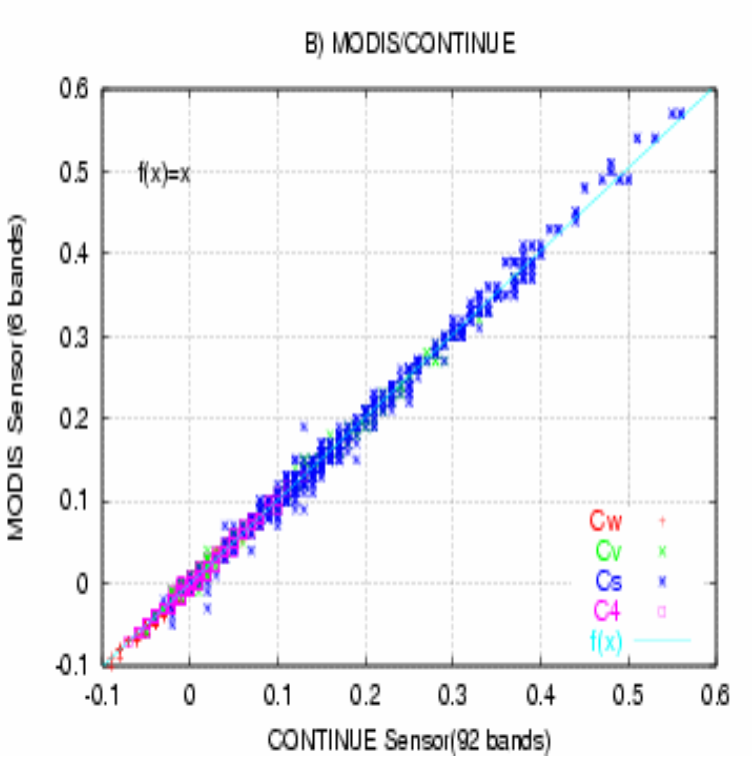
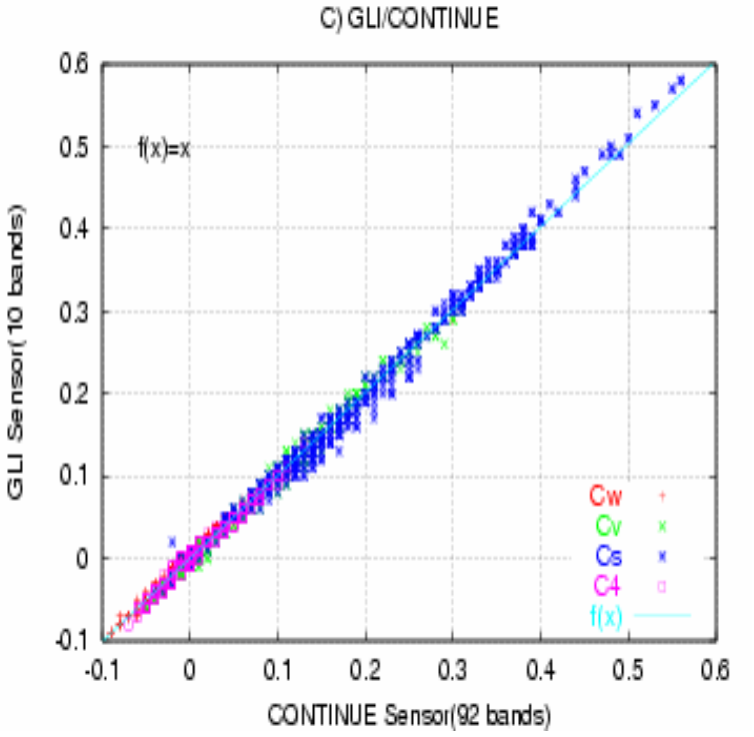


Figure 4.2. Reduced χ^2 for reconstructed spectra

4.2. *Correlation of pattern decomposition coefficients*




B) MODIS/CONTINUE
f(x)=x
MODIS Sensor(6 bands)
CONTINUE Sensor(92 bands)
Cw
Cv
Cs
C4
f(x)
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6


C) GLI/CONTINUE
f(x)=x
GLI Sensor(10 bands)
CONTINUE Sensor(92 bands)
Cw
Cv
Cs
C4
f(x)
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6

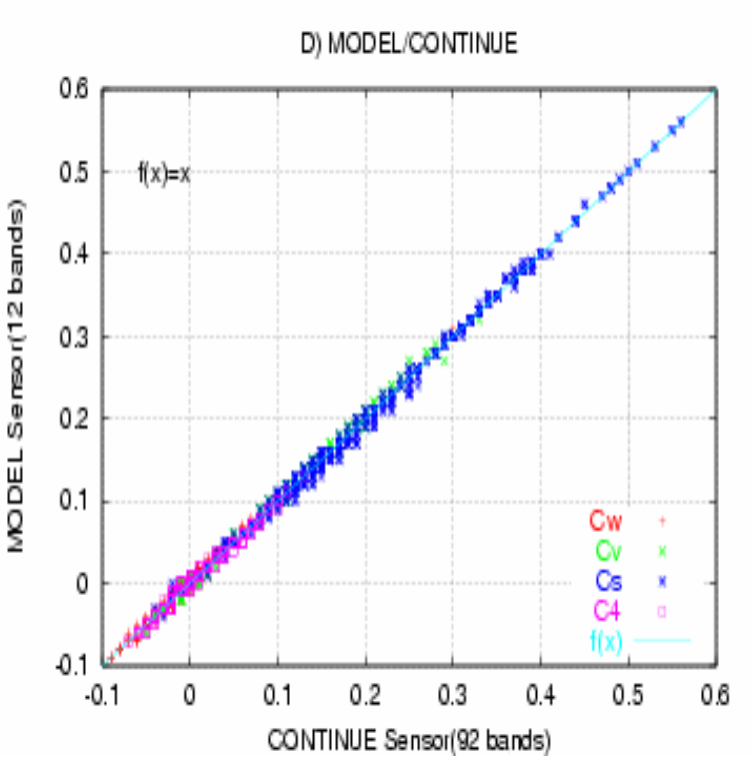


Figure 4.4. Correlation of the universal pattern decomposition coefficients. A) ETM sensor and CONTINUE sensor, B) MODIS sensor and CONTINUE sensor, C) GLI sensor and CONTINUE sensor, D) MODEL sensor and CONTINUE sensor,

Figure 4.4 shows the correlation of the UPDM coefficients obtained using the universal pattern decomposition method. The horizontal-axis shows the UPDM coefficients for the CONTINUE sensor, and the vertical-axis shows the ETM+, MODIS, GLI, or MODEL sensor coefficients. As expected, the coefficients obtained for each sensor nearly equal the coefficients obtained from the CONTINUE sensor. Table 4.2 lists the linear regression coefficients of $f(x) = a_i x \ (i = 1,2,3,4)$ for correlating the UPDM coefficients of each sensor, and the root mean square (rms) of the residuals of coefficient C_k ($k = w, v, s, 4$). Total values were obtained from all UPDM coefficients by linear regression. The coefficients of a_i were nearly 1 for all sensors, and the rms values were very small compared with the coefficient values. This means that the UPDM coefficients are sensor-independent.

Table 4.2. Coefficients of the linear regression function and the root mean square (rms) values of the residuals

	GLI		ETM+		MODIS		MODEL	
	a_1	rms	a_2	rms	a_3	rms	a_4	rms
C_w	1.0122	0.0047	0.9615	0.0036	1.0390	0.0045	1.0125	0.0038
C_v	1.0032	0.0073	0.9801	0.0114	1.0009	0.0059	1.0087	0.0048
C_s	1.0083	0.0112	1.0057	0.0074	1.0105	0.0115	0.9969	0.0061
C_4	0.9925	0.0052	0.9980	0.0057	1.0077	0.0046	0.9735	0.0042
total	1.0070	0.0075	0.9992	0.0077	1.0087	0.0073	0.9995	0.0049

5. Summary and conclusions

We developed a universal pattern decomposition method (UPDM) to obtain sensor-independent pattern decomposition coefficients for reflectance data in the 371 to 1,750 nm wavelength range. For this, we analyzed about 600 ground-measured samples, including green-leaf, yellow-leaf, dead-leaf, soil, water, and concrete samples. A supplementary spectral pattern was used to counteract the influence of withering vegetation. The standard spectral patterns of water, vegetation, soil, and the supplementary pattern were normalized for the entire spectral band from 350 to 2,500 nm, excluding the region of strong vapor absorption. The total number of bands was 1,260 with 1-nm spacing.

We verified that the pattern decomposition coefficients obtained using the UPDM are nearly sensor independent. That is, each sensor showed nearly the same coefficient values within 0.77% of the rms. Exceptions were four-band sensors, such as MSS and ALOS. The reduced χ^2 values for Landsat/ETM+, Terra/MODIS, and ADEOS-II/GLI were around 0.00070. The square root of 0.00070 is 0.026 (2.6%). Using this method, we can analyze satellite data independent of the sensor used and compare the results of analyses directly, using multispectral data effectively.

To further validate this UPDM, we will apply data observed over Wuhan, China, and the Kii Peninsula, Japan with the ETM+, MODIS, and GLI satellites. We will define a sensor-independent vegetation index as a linear function of the UPDM coefficient and will study sensor-independent ground classifications and NPP estimation in future research.

Acknowledgments

The authors thank A. Ono of Otemae University for encouraging us to write this paper. This work was supported under the ADEOS-II/GLI Project of the Japan Aerospace Exploration Agency (JAXA), the Academic Frontier Promotion Project of the Ministry of Education, Science, Sports, and Culture of Japan, and the 973 Project of the People's Republic of China (Project Number 2003CB415205).

References

Adams, J. B., Sabol, D. E., Kapos, V, Filho, R. A., Roberts, D. A., Smith, M. O., and Gillespie, A. R. (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.

Daigo, M., Ono, A., Fujiwara, N, and Urabe, R. (2004). "Pattern decomposition method for hyper-multi-spectral data analysis," *International Journal of Remote Sensing*, 25(6), 1153-1166.

Fujiwara, N., Muramatsu, K., Awa, S., Hazumi, A. and Ochiai, F. (1996). "Pattern expansion method for satellite data analysis," (in Japanese). *Journal of Remote Sensing Society of Japan*, 17(3), 17-37.

Furumi, S., Hayashi, A., Murumatsu, K., Fujiwara, N. (1998). "Relation between vegetation vigor and new vegetation index based on pattern decomposition method," *Journal of Remote Sensing Society of Japan*, 18(3), 17-34.

Hayashi, A., Muramatsu, K., Furumi, S., Shiono, Y., Fujiwara, N., Daigo, M. (1998). "An algorithm and a new vegetation index for ADEOS-II/GLI data analysis," *Journal of Remote Sensing Society of Japan*, 18(2), 28-50.

Muramatsu, K., Furumi, S., Fujiwara, N., Hayashi, A., Daigo, M., and Ochiai, F. (2000). "Pattern decomposition method in the albedo space for Landsat TM and MSS data analysis," *International Journal of Remote Sensing*, 21(1), 99-119.