IDENTIFYING HYPERSPECTRAL CHARACTERS OF WETLAND SPECIES USING IN-SITU DATA

Y. SUN [a], X. Liu [a,*], Y. Wu [a], C. Liao [a]

[a] Dept. of Environmental Science and Engineering, Tsinghua University, Beijing, 100084 China, xuehua-hjx@tsinghua.edu.cn

KEY WORDS: Hyperspectral Character, Wetland Vegetation, Derivative Reflectance, Continuum Removal, Mahalanobis Distance, Band Selection

ABSTRACT:

Wetland species identification is a very important process when determining wetland ecosystem types and wetland composition. Hyperspectral data with the merit of high spectral resolution have become a powerful tool for ground feature identification compared with multi-spectral data. This paper investigated the hyperspectral characters of wetland plants by considering the two prominent attributes of hyperspectral data: continuity and hyper-bands. In term of the continuity attribute, Derivative Reflectance (DR) and Continuum Removal (CR) method were used to extract key characters from the original spectra. In order to reduce spectral bands, firstly four factors such as slope (K), area of absorption curve (A), area of left wing (AL), and symmetry (S) were constructed, and secondly Mahalanobis Distance (MD) method was used to select the best bands for identifying plant species. The results showed that the First Derivative Reflectance (FDR), Second Derivative Reflectance (SDR) and Continuum Removal (CR) curves illustrated the variances of plant spectra from various ways. There are 5 bands for the original spectra, 5 bands for FDR curves, 6 bands for SDR curves, and 4 bands for CR curves were selected as better bands for species identification. All the selected bands are mainly located in the typical chlorophyll absorption and water absorption region. The constructed factors such as K1, K3, S3, Al, AL1, A2 and AL2 varies for different species which indicated that they can be used for species identification. However, for factors like K2, A3, AL3 can only be used to identify emerged, floating and submerged plant types.

1. INTRODUCTION

Remote sensing data have been widely used for wetland studies, such as vegetation classification and change monitoring (Gress, 1993; Roshier and Rumbachs, 2004; Rebelo et al, in press). Compared with multi-spectral data, hyperspectral data offered more detailed information of the reflectance spectra of different wetland plant species (Yuan and Zhang, 2006).

Airborne or space borne hyperspectral sensors have been more and more used by researchers on wetland classification and species identification, whereas remote sensed imagery is easy to be influenced because of the constraints of image calibration, such as atmospheric correction and geographic correction (Sanders et al, 2001; Secker et al, 2001). Therefore, the study of in-situ hyperspectral data is needed not only for the remote sensed images calibration but also for some precursory research of spectra attributes of wetlands. This could also contribute to the spectral library of species on the Earth (Ben-Dor et al, 2001).

There are two predominant advantages of hyperspectral data, which are continuity and hyper-bands. In term of the continuity attribute, it becomes possible to apply many mathematical methods to analyse hyperspectral data, such as derivative method (Becker et al, 2005) and continuum removal method (Kokaly and Clark, 1999). However, there are large numbers of bands available from hyperspectral data, and thus optimism bands have to be selected for vegetation identification. In addition, with the plenty of potential information, hyperspectral data could be analysed by more methods and from different aspects, such as constructing factors for better describing the spectra character (Li, 2006).

2. METHOD

2.1 Data collection and processing

Eleven wetland plants were selected in this research including three emerged plants, three floating plants, and five submerged plants (Table 1). The data collection work was implemented in the Arboretum of the Institute of Botany, Chinese Academy of Science in Beijing, in the spring of 2008. The spectral reflectance data were obtained between 10.00am and 14.00pm, under sunny, windless and cloudless conditions.

Table 1. The vegetation types sampled with acronym used in this paper, and the number of spectra collected

Species	Type *	Acronym	No. of spectra curves
Cyperus alternifolius	E	1. CA	43
Cyperus papyrus	E	2. CP	52
Pontederia cordata	E	3. PC	55
Nymphaea tetragona	F	4. NT	38
Hydrocleys nymphoides	F	5. HN	50
Nymphoides peltatum	F	6. NP	54
Pistia stratiotes	F	7. PS	35
Azolla imbircata	F	8. AI	43
Vallisneria asiatica	U	9. VA	45
Potamogeton malaianus	U	10. PM	32
Hydrilla verticillata	U	11. HV	36

*Here, E stands for emerged plant, F stands for floating plant, and U stands for submerged plant.

A SVC GER 1500 Hand-held Spectrometer was used to measure in-situ reflectance with a 4° field of view and 0.5m above the species. There were 512 channels displayed in the

---

waveband between 350 to 1050 nm (visible to near infrared), with 3.2 nm spectral resolution and 1.5 nm band width.

2.2 Derivative method

Benefited from the high spectral resolution of hyperspectral data, the reflectance could construct a comparatively continuous curve, and thus the derivative method is likely to be used for data analysis. From one hand, the derivative processing could enhance the small variations of spectral curve (Li, 2006). And from the other hand, the derivatives could also be relatively less sensitive to the spectral variations of sunlight and skylight (Tsai and Philpot, 1998), and also eliminate background signals (Becker et al, 2005).

The First Derivative Reflectance (FDR) and Second Derivative Reflectance (SDR) were calculated according to Eq. (1) and Eq. (2). For FDR method, the slope information of the curve could be drawn out of the curve, and for SDR method, the curvature information could be highlighted.

$$FDR_{\lambda_i} = \frac{dR}{d\lambda} = \frac{R_{\lambda_{j+1}} - R_{\lambda_j}}{\Delta\lambda} \tag{1}$$

$$SDR_{\lambda_i} = \frac{d^2R}{d\lambda^2} = \frac{d}{d\lambda}\left(\frac{dR}{d\lambda}\right) = \frac{R_{\lambda_{j+2}} - 2R_{\lambda_{j+1}} + R_{\lambda_j}}{(\Delta\lambda)^2} \tag{2}$$

Where

FDR_{λ_i}	=	The FDR between band j and band j+1, where wavelength equals to *i*;
SDR_{λ_i}	=	The SDR between band j and band j+2, where wavelength equals to *i*;
$R_{\lambda_j}, R_{\lambda_{j+1}}, R_{\lambda_{j+2}}$	=	Reflectance of band j, j+1, j+2;
$\Delta\lambda$	=	Difference of wavelength between band j and band j+1.

2.3 Continuum Removal method

Continuum Removal (CR), which is a normalisation method, could emphasis the location and depth of individual absorption features throwing drawing a curve with the absorption values from 0 to 1 (Clark & Roush, 1984). According to the typical wetland vegetation reflectance spectra, three wavebands were selected to calculate CR, which are 400-530nm (blue, chlorophyll absorption), 550-730nm (red, chlorophyll absorption) and 930-1000nm (Near Infrared, NIR, water absorption) (See Figure 1). The CR values of three wavebands were calculated by the Eq. (3), and the illustration of key factors can be found in Figure 1.

$$CR_i = \frac{R_i}{R_{Hi}} \tag{3}$$

$$R_{Hi} = \frac{R_i}{R_{start} + K \cdot (\lambda_i - \lambda_{start})} \tag{4}$$

$$K = \frac{R_{end} - R_{start}}{\lambda_{end} - \lambda_{start}} \tag{5}$$

Where

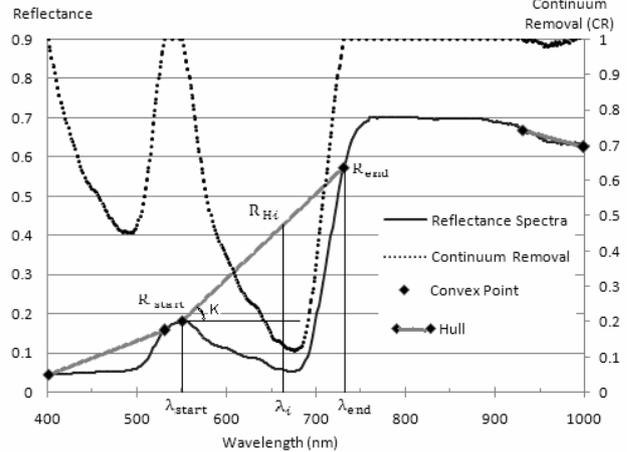|		
CR_i	=	The continuum removal value of band *i*;
R_i	=	The reflectance of band *i*;
R_{Hi}	=	The hull value of band *i*;
R_{end}, R_{start}	=	The reflectance of the start point and end point in an absorption curve;
$\lambda_{end}, \lambda_{start}$	=	The wavelength of the start point and end point in an absorption curve;
K	=	The slope between the start point and end point in an absorption curve;

Figure 1. Three wavebands and key factors used for calculating Continuum Removal (CR). The vegetation spectrum in the figure is from *Cyperus alternifolius* (1.CA).

When calculating CR, more constructed factors could be designed for better describing the features of the absorption curve and therefore seizing the key characters of vegetation spectra (Li, 2006). Firstly, the slope (K) could be calculated directly by Eq. (5). Then, CR could be converted to BD according to Eq. (6), and thus integral method could be used to calculate the area of absorption curve (A), the area of left wing (AL), and the symmetry (S) (see Eq. (6)-(8) and Figure 2). Since there are three bands selected for CR analysis (see Section 2.3), the factors will accordingly named A1, AL1, S1 for 400-530nm, A2, AL2, S2 for 550-730nm, and A3, AL3, S3 for 930-1000nm.

$$BD_i = 1 - CR_i \tag{6}$$

$$A = \int_{\lambda_{start}}^{\lambda_{end}} CR \quad , \quad AL = \int_{\lambda_{start}}^{\lambda_{BD_{max}}} CR \tag{7}$$

$$S = \frac{AL}{A} \tag{8}$$

Where

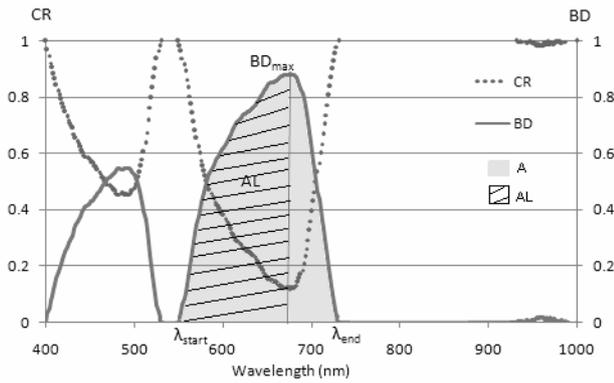|		
BD_i	=	Band depth of band *i*;
CR_i	=	Continuum removal of band *i*;
A, AL	=	Area of absorption curve and its left wing;
$\lambda_{end}, \lambda_{start}$	=	The wavelength of the start point and end point in an absorption curve;
λ_{BDmax}	=	Wavelength where there is a maximal BD;
S	=	Symmetry of absorption curve.

Figure 2. The illustration of BD, A and Al taking the example of *Cyperus alternifolius* (1.CA)

2.4 Band selection through Mahalanobis Distance

With hundreds of spectra bands available for analysing, the hyperspectral data could, from one side, provide plenty of detailed information of the reflectance of vegetation, while from the other side, the redundancies will slow-down the vegetation modelling speed and also the correlation between different bands could disturb the feature identification. Therefore, different methods were used for band selection, including BH and JM distance method (Schmidt and Skidmore, 2003), and statistically testing method (Becker et al, 2007).

A new method - Mahalanobis Distance (MD) was used for band selection in this research. The MD method has been widely used for different purpose (De Maesschalck et al, 2000; Goovaerts et al, 2005), and its merit relies on taking into account the correlation in the data. For selecting the best bands for species identification, firstly, the reflectance differences of paired species were calculated for every band. In this case, there were totally 55 possible pairs for the 11 plant species. Secondly, the MD was calculated by the Eq. (9) for measuring the distance between the difference matrix and the zero point (as the comparison point, CP). The zero point means the least difference between different species, so the larger the MD value, the more significant of the difference between species, and thus the easier to identify different species.

$$MD_{(i)} = \sqrt{\left(X_i - X_{cp}\right)\Sigma^{-1}\left(X_i - X_{cp}\right)^T} \qquad (9)$$

Where

$MD_{(i)}$	=	The MD of band *i*;
X_i	=	The difference matrix of band *i*;
Xcp	=	The defined comparison point;
Σ^{-1}	=	The determinant of the variance–covariance matrix.

3. RESULTS

The spectral reflectance curves of the 11 wetland plant species were shown in Figure 3. The figure showed that different species shared similar shape of curves, while difference can be identified especially in the green reflectance (500-600nm) and NIR reflectance (700-1000nm) regions. Especially in the NIR portion, it is clear that the three emerged species have the highest reflectance (1.CA, 2.CP and 3.PC), and in the contrary,

the three submerged species have the lowest reflectance (9.VA, 10.PM and 11.HV). The four floating plants share medium reflectance. The reason of this phenomenon is that the reflectance of wetland plants could be influenced very easily by water absorption, and water has an absorption peak in the NIR wavebands. However, for getting more information of the curves, detailed analysis was applied in the following sections.
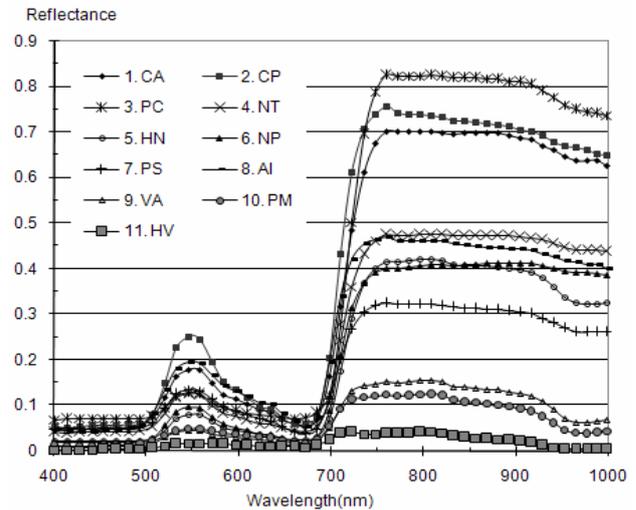

Figure 3. Spectra of 11 wetland plant species collected by GER 1500 Spectrometer.

3.1 Continuous spectra analysis

The continuity attributes of the species spectra were analysed through FDR, SDR and CR methods.

(1) First Derivative Reflectance (FDR)
The FDR curves of 11 species showed the slope of the original spectra in every band (see Figure 4). The positive FDR means the curve is increasing while the negative FDR represents spectra is decreasing. Zero point happens when there is a peak in the curve. As can be seen from Figure 4, there are two climaxes of spectra increasing which are located in about 520nm and 710nm. The climax of 520nm illustrated the reflectance of chlorophyll, where different species have different FDR. Species 2.CP has the highest FDR compared with other species, which means the highest increasing speed of reflectance. The climax of 710nm indicates the "red edge" phenomenon, which could be used to determine the vegetation health (Gupta et al, 2001) as well as to identify different species.
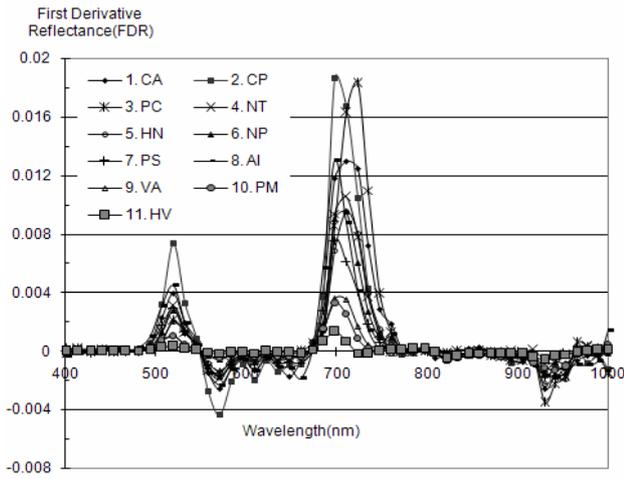
Figure 4. Comparison of the FDR curves for 11 wetland species.

(2) Second Derivative Reflectance (SDR)

The SDR curves illustrated the curvature of the original spectra. When the SDR is positive, it means there is a concave in the original spectra. This kind of points can be found around 510nm, 580nm, 690nm, and 960nm. On the contrary, when the SDR is negative, a convex will be found in the spectra (Becker et al, 2005). This kind of points can be found in 540nm, 560nm, 630nm, 720nm, and 910nm.
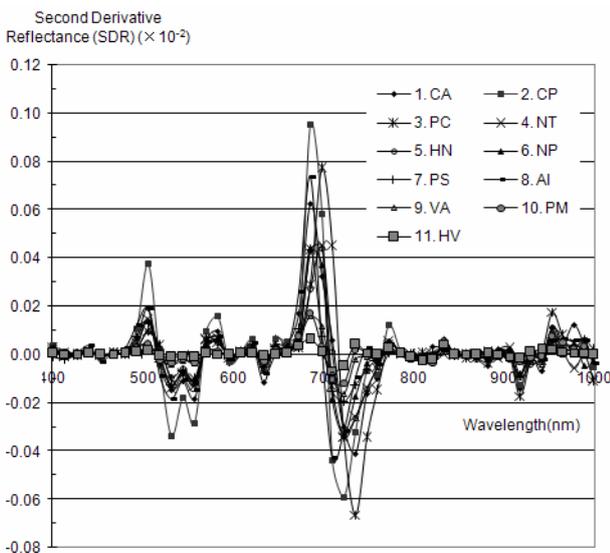

Figure 5. Comparison of the SDR curves for 11 wetland species.

(3) Continuum Removal (CR)

The CR curves highlighted the absorption spectra of original reflectance spectra (see Figure 6) (Kokaly and Clark, 1999). In general, the absorption region 2 (550-730nm) has the highest absorption rate among three regions. This region is in the red-edge and it has been applied chlorophyll identification (Zarco-Tejada et al, 2003). Besides, the species 3.PC has the weakest absorption in region 1 (400-540nm) and species 11.HV has the strongest absorption in region 3 (930-1000nm). The curves in three regions are hardly symmetrical and varied in different band.
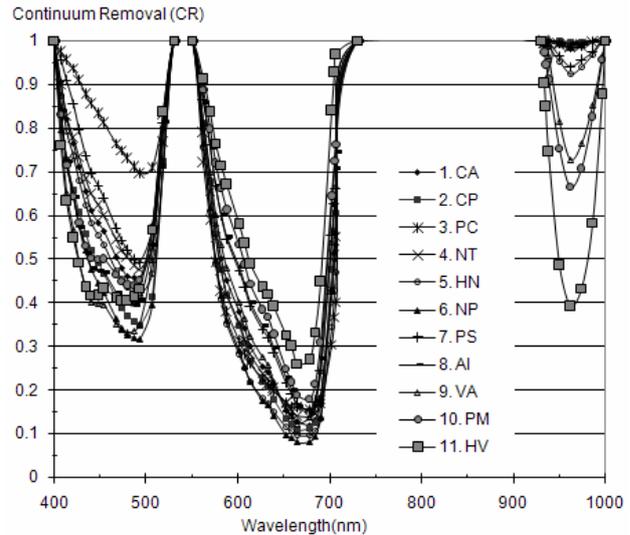

Figure 6. Comparison of the CR curves for 11 wetland species.

According to the analysis above, FDR, SDR and CR curves illustrated different attributes of original spectra and the attributes were emphasised in different ways. Therefore, band selection method should be applied to extract the most dominant attributes.

3.2 Constructed Factors Analysis

Compared with the continuous factors, the constructed factors means for every species, there is only one value for every factor. The constructed factors analysed in this research are K, A, AL, and S, as explained in section 2.3.

The K values of three absorption regions are illustrated in Figure 7. The K values are varied for different wetland plant species. For absorption region 1, species 2.CP and 8.AJ have the highest K values while the submerged species (9.VA, 10.PM and 11.HV) have comparatively lower K than the other species. For the absorption region 2, K values could be grouped in 3 parts, which are emerged species with the highest K, floating species with the medium K, and submerged species with the lowest K. And for region 3, species 3.PC and 5.HN showed higher K than other species. To sum up, K1, K2 and K3 are all varied for species which is possible to be used for species identification. And K2 showed better ability to identify wetland plant species into emerged, floating, and submerged groups
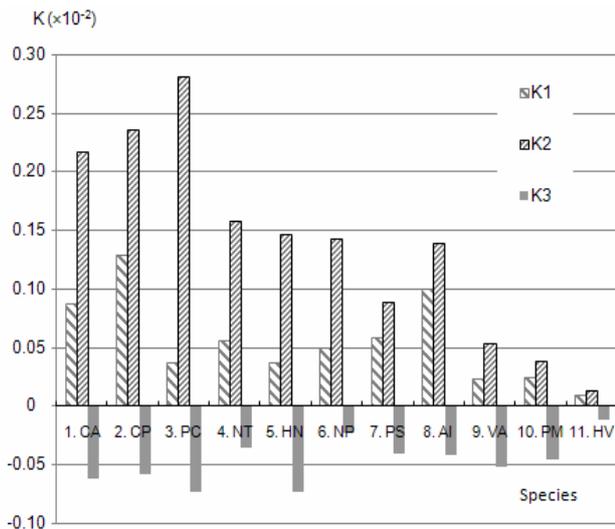
462

Figure 7. K values of three wavebands (1: 400-530nm, 2: 550-730nm, 3: 930-1000nm) for the 11 wetland plant species.

The A, AL and S values were shown in Figure 8. Ddifferent factors behaved differently for 11 wetland plant species. Factors S1 and S2 are comparatively stable for different species, which means these factors are probably the specific attributes for wetland plants. On the contrary, for the factors such as S3, Al, AL1, A2 and AL2, different species behaved differently which means these factors could probably be used for species identification. For the factor A3 and AL3, submerged species have comparatively higher values than the other species, which

is because of the absorption of water but could also be used for identifying submerged plant species.

3.3 Analysis of MD-base Band Selection

MD method was used to select bands from the original spectra, as well as FDR, SDR and CR curves, for better identifying wetland plant species. The results were shown in Figure 9. The selected bands should have higher MD values than the average, and also maintain the continuous high MD for at least 10nm. To use 10nm as the least band region is because the hyperspectral images usually have the spectral resolution of 3-10nm. Therefore, the selected bands could also be applied on remote sensing images, which is actually one of the objectives for applying in-situ hyperspectral research.

From Figure 9 it can be seen that 5 bands were selected for the original spectra, which are 410-424nm (blue chlorophyll absorption), 523-566nm (reflectance of green), 670-766nm (red-NIR edge), 796-821nm and 927-996nm (NIR, water absorption). There are also 5 bands were better in the FDR curves, which are 507-530 nm (chlorophyll absorption), 680-743 nm and 751-768 nm (red-NIR edge), 912-962 nm and 976-990 nm (NIR, water absorption). And for SDR curves, 6 bands were considered to be better than the others, which contains three bands in the chlorophyll absorption (496-516 nm, 521-540 nm and 544-561 nm), one band in the red-NIR edge (667-758 nm), and two in the NIR region (950-964 nm and 985-999nm). Finally, for CR curves, 4 bands were selected which contains one band in the absorption increasing wing (410-432 nm), and three bands in the decreasing wing (504-527 nm, 694-724 nm, and 956-973 nm).
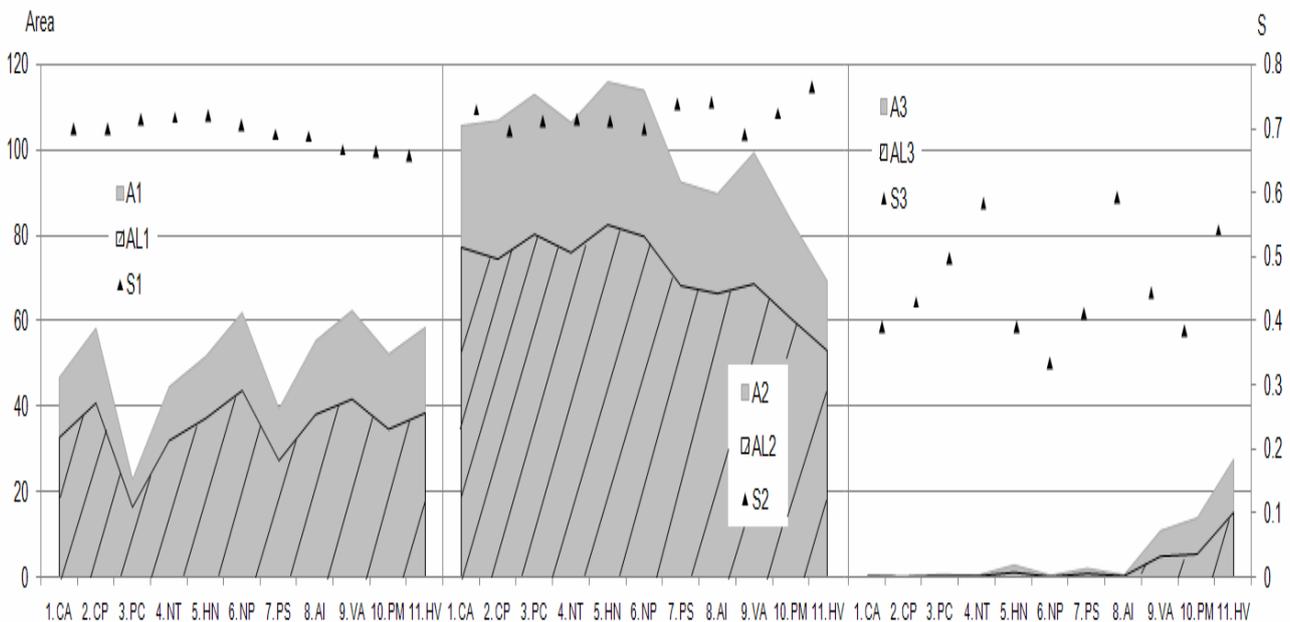

Figure 8. A, AL and S values of three absorption wavebands (1: 400-530nm, 2: 550-730nm, 3: 930-1000nm) for the 11 wetland plant species.
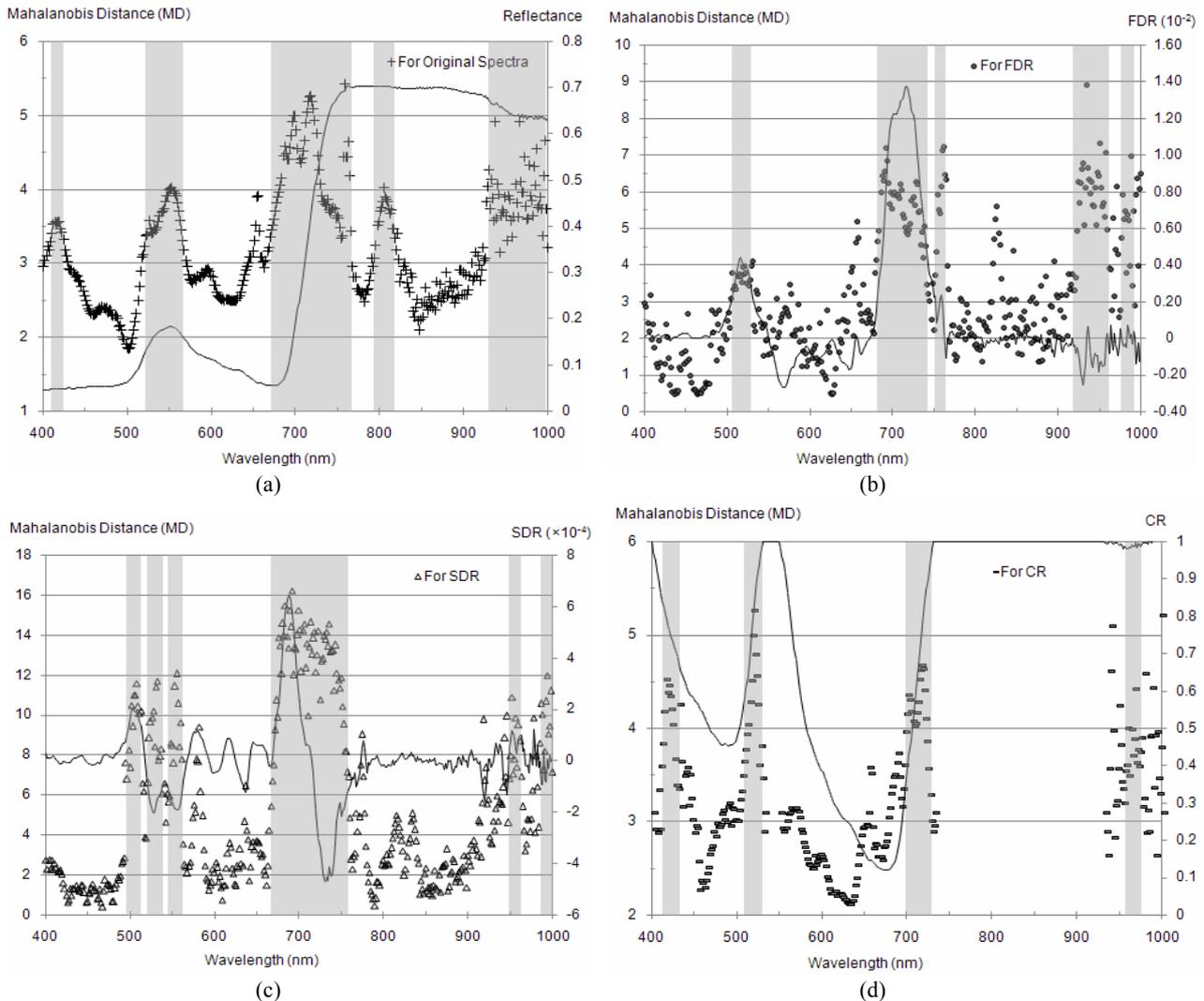
Figure 9. Band selection results through Mahalanobis Distance for original spectra (a), FDR (b), SDR (c), and CR (d). (The data of *Cyperus alternifolius* (1.CA) were used as an example in the figure.)

4. DISCUSSION

Compared with multi-spectral data, hyperspectral data offered more detailed information of the reflectance spectra of different wetland plant species. There are mainly two advantages of hyperspectral data, which are continuity and hyper-bands.

For making use of the continuity attribute, derivative analysis and continuum removal analysis could be applied on the original reflectance spectra. The First Derivative Reflectance (FDR) could illustrate the slope of original spectra and the Second Derivative Reflectance (SDR) could indicate the curvature of spectra. The Continuum Removal (CR), from another point of view, constructed the absorption spectra based on the original spectra. All three methods could actually magnify the small variances in the original spectra which might be used for species and vegetation identification.

The hyper-bands attribute can provide more opportunities for better illustrating plant species while it can also be redundant sometimes. To solve this problem, two methods could be applied. From one hand, constructed factors could be constructed from the continuous curves, such as slope (K) of reflectance within a fixed waveband, area of absorption curve

(A), area of left wing (AL), and symmetry (S). From the other hand, band selection could also be very helpful to reduce the bands with week diagnostics or correlated. All selected bands and constructed factors could also be used for species and vegetation identification.

5. CONCLUSION

This research studied the spectra characters of wetland plant species, based on continuity and hyper-bands attributes of hyperspectral data. The results showed that different wetland plants have similar spectra curves while they are still possible to be distinguished in some wavebands in visible and NIR in hyperspectral data.

The attributes of spectra curves could be strengthened when using derivative method and continuum removal method. The 5 bands for the original spectra, 5 bands for FDR curves, 6 bands for SDR curves, and 4 bands for CR curves, have been selected as better bands for species and vegetation identification. All the selected bands mainly located in the chlorophyll absorption and water absorption region. The constructed factors showed different behaviours comparing among different species. Some

factors such as K1, K3, S3, Al, AL1, A2 and AL2 have different values for different species, so they could be used for species identification. However, for factors like K2, A3, AL3 could possible be used to identify different groups of species such as emerged plant, floating plant, and submerged plant types.

This research studied the hyperspectral curve of wetland plants from different aspects, and also identified key factors for species identification. The study of in-situ hyperspectral data could be very helpful for building the spectral library of wetland species and vegetation. Also, in-situ data could be applied for remote sensing images, for example the selected bands could be applied in the image classification for vegetation mapping. Combining with airborne and spaceborne images, hyperspectral data will help us to better monitor wetlands and thereafter improve the wetlands protection ability.

REFERENCES

Becker, B.L., D.P. Lusch, J. Qi, 2005. Identifying optimal spectral bands from in situ measurements of Great Lakes coastal wetlands using second-derivative analysis. Remote Sensing of Environment, 97(2), pp. 238-248.

Becker, B.L., D.P. Lusch, J. Qi, 2007. A classification-based assessment of the optimal spectral and spatial resolutions for Great Lakes coastal wetland imagery. Remote Sensing of Environment, 108(1), pp.111-120.

Ben-Dor, E., N. Levin., H. Saaroni, 2001. A spectral based recognition of the urban environment using the visible and near-infrared spectral region. A case study over Tel-Aviv, Israel. International Journal of Remote Sensing,22 (11), pp.2193-2218.

Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. Journal of Geophysical Research, 89(B7), pp. 6329-6340.

De Maesschalck, R., D. Jouan-Rimbaud, D. L. Massart, 2000. The Mahalanobis distance. Chemometrics and Intelligent Laboratory Systems, 50 (1), pp.1-18.

Goovaerts, P., G.M. Jacquez, A. Marcus, 2005. Geostatistical and local cluster analysis of high resolution hyperspectral imagery for detection of anomalies. Remote Sensing of Environment, 95(3), pp.351-367.

Gress, T. A., D. Kettler, C. Mao, L. Wright, and S. Berta. 1993. Wetland recertification and sampling using satellite remote sensing. Advances in Space Research, 13(11), pp.83-90.

Gupta, R.K., D. Vijayan, T.S. Prasad, 2001. Characterization of red-near infrared transition for wheat and chickpea using 3 nm bandwidth data. Advances in Space Research,28(1), pp.189-194.

Kokaly, R.F., R.N. Clark. 1999. Spectroscopic Determination of leaf biochemistry using Band-Depth analysis of absorption features and stepwise multiple linear regression. Remote Sensing of Environment, 67(3), pp.267-287.

Li, M.Z., 2006.   Technologies and Applications of Spectral Analysis. Science Press, Beijing, pp. 122-124, 163-173.

Rebelo, L. M., C. M. Finlayson, N. Nagabhatla, In Press. Remote sensing and GIS for wetland inventory, mapping and change analysis. Journal of Environmental Management.

Roshier, D.A., R.M. Rumbachs. 2004. Broad-scale mapping of temporary wetlands in arid Australia. Journal of Arid Environments, 56(2), pp. 249-263.

Sanders, L.C., J.R. Schott, R. Raqueno, 2001. A VNIR/SWIR atmospheric correction algorithm for hyperspectral imagery with adjacency effect. Remote Sensing of Environment, 78(3), pp. 252-263.

Secker, J., K. Staenz, R.P. Gauthier, P. Budkewitsch, 2001. Vicarious calibration of airborne hyperspectral sensors in operational environments. Remote Sensing of Environment, 76(1), pp.81-92.

Schmidt, K.S., A.K. Skidmore, 2003. Spectral discrimination of vegetation types in a coastal wetland. Remote Sensing of Environment. 85 (1), pp. 92-108.

Tsai, F., W. Philpot. 1998. Derivative Analysis of Hyperspectral Data. Remote Sensing of Environment, 66 (1), pp. 41-51.

Yuan, L., L.Q. Zhang, 2006. Identification of the spectral characteristics of submerged plant Vallisneria spiralis. Acta Ecologica Sinica, 26(4), pp. 1005-1010.

Zarco-Tejada,P.J., J.C. Pushnik, S. Dobrowski, S.L. Ustin, 2003. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. Remote Sensing of Environment, 84 (2), pp. 283-294.