

Effects of Spectral Transformations on Leaf C:N Ratio Inversion with Hyperspectral Data

SHI Run-he ZHUANG Da-fang

Data Center, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences
No. 3, Datun Rd., Beijing 100101, China
shirh@reis.ac.cn

QIAN Qiao-jing NIU Zheng

LARSIS, Institute of Remote Sensing Applications, Chinese Academy of Sciences
PO Box 9718, Datun Rd., Beijing 100101, China

Abstract: Leaf C:N ratio is a new factor in the field of biochemical inversion with hyperspectral data. Effects of common-used spectral transformations including $\log(R)$, $\log(1/R)$, $1/R$, etc. from 400nm to 2490nm on its inversion are compared. Results show that their effects on statistical modeling are not apparent. Continuum removal is used on original reflectance in the range of 2030nm to 2220nm, in which exists an apparent absorption peak due to cellulose, lignin, protein, etc. The effect is distinctive and tends to improve the precision of C:N ratio inversion. Further, it is a robust and physically based transformation.

Keywords: C:N, Spectral transformation, Continuum removal.

1. Introduction

With the development of hyperspectral techniques, remote sensing of vegetation reached an amazing atom and molecule level. Spectral reflectance could be measured at nm level in visible and near infrared region, which realizes continuous reflectance spectra for any object. In principle, the spectra are determined by the biochemical compositions in a leaf and their chemical elements, micro-structure and content. That is the reason why we are able to inverse biochemical information from their spectra.

Methods of biochemical inversion with remotely sensed data could be divided into two types as a whole, physical modeling and statistical modeling. The latter is simple and easy to practise, so it is widely used in many researches^{[1][2]}. However, it is site-specific, time-specific, short of robustness and sometimes hard to understand. All these features retard its further applications and development^[3]. Spectral transformations are mathematic transforms of original reflectance, which are always used prior to statistical analyses in order to attain higher coefficient of correlation or regression^[4]. The common-used spectral transformations are simple operations on original reflectance, such as $1/R$, $\log(R)$, $\log(1/R)$, etc. All of them aim to make absorption features in certain wavelengths more distinctive. Then derivative spectra (dR) are developed to eliminate influences caused by backgrounds like soil. Its effect is quite good in many cases. Continuum removal is a numerical method to estimate the absorptions not due to the band of interest and remove their effects. Band depth normalization is another

spectral transformation, which is based on continuum-removed spectra^[5]. Some researches show that it is a good preprocessing of statistical modeling, which makes the model more robust.

In most previous researches, biochemical compositions are treated as inversion objects, such as nitrogen, chlorophyll, cellulose, lignin, etc^{[4][5]}. While in practice, C:N ratio is widely used in many fields such as ecology, global changes, precision agriculture, etc. But it is rarely studied unfortunately^[6]. Given its importance in the nutrition utilizing efficiency (NUE) and plant life processes, we use C:N ratio as an inversion object here.

In this paper, all the transformations mentioned above will be compared for their effects on leaf C:N ratio inversion.

2. Method

Reciprocal, logarithmic and first derivative operations are employed on original reflectance (R) or its reciprocal respectively in the range of 400nm to 2490nm with 10nm interval. Continuum removal is employed on R in one absorption feature centered at 2100nm with 2nm interval. Two end points are set at 2030nm and 2220nm respectively.

Statistical analyses on the transformed spectra include correlation analysis and stepwise multiple linear regression. Coefficients of correlation and coefficients of determination are compared site-by-site and transformation-by-transformation to find the advantages and disadvantages of each transformation and to examine their robustness in statistical modeling. As to $1/R$, $\log(R)$, $\log(1/R)$ and dR , all bands from 400nm to 2490nm with 10nm interval are participated in correlation and regression operation, while as to continuum-removed reflectance (R'), only bands from 2030nm to 2220nm with 2nm interval are participated in, because one continuum represents only one absorption feature. As we can see in Fig. 1, two main carbon-containing compositions, cellulose and lignin both have one absorption peak at 2100nm or so and form a V-shape and a U-shape respectively, while as to protein, a main nitrogen-containing composition in a leaf, has two absorption peaks at this region and forms a distinguishable W-shape^{[7][8]}. So that we express our interests in this region.

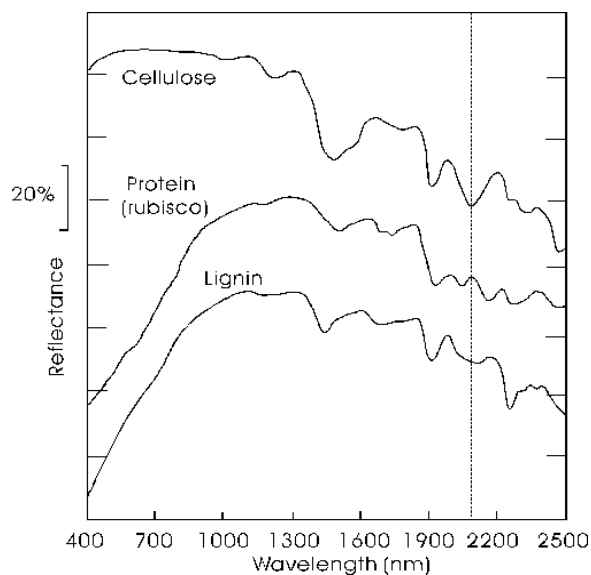


Fig. 1. Reflectance spectra of cellulose, lignin and protein

3. Data Set

Reflectance data and biochemical data used in this study are both from NASA ACCP (Accelerated Canopy Chemistry Program) data set [9], which was compiled by an interdisciplinary team to investigate the feasibility of making estimates of canopy biochemical compositions from remote sensing observation. More than 500 dry leaf samples from three eastern U.S. forests (Harvard Forest, Massachusetts; Blackhawk Island, Wisconsin; and Howland, Maine) are examined. All samples were oven dried at 70 °C for 48h, ground through a 1 mm mesh, and homogenized prior to measurement [10]. Spectral reflectance was measured with a NIR-Systems Model 6250 scanning monochromator with spinning sample cup module [11]. Reflectance data were gathered over the wavelength range from 400 nm to 2498 nm at a 2 nm interval with a 10 nm bandpass and recorded in the form of $\log(1/R)$. Descriptive statistics of the chemical concentrations and C:N ratio at each site are shown in Table 1. We can see that standard deviations of carbon concentrations and nitrogen concentrations of the 3 sites are relatively small compared to those of C:N ratio.

Table 1. Descriptive statistics of the chemical concentrations and C:N ratio at each site

	Site	No.	Max.	Min.	Mean	Std. Dev.
C	BHI	182	43.87	52.66	49.60	1.96
	HF	187	47.38	52.68	50.41	0.94
	HOW	186	47.52	53.12	50.85	1.22
N	BHI	182	1.09	3.51	2.37	0.54
	HF	187	0.93	3.19	1.85	0.54
	HOW	186	0.69	2.67	1.34	0.44
C:N	BHI	182	13.13	47.55	22.46	7.21
	HF	187	15.63	54.33	30.10	10.11
	HOW	186	18.99	74.94	41.85	12.50

4. Results

1) Correlation analysis

The change of correlation coefficients (r) about C:N ratio and different transformed reflectance spectra are shown in Fig. 2, which includes three pictures representing three sites respectively. Four curves are almost overlapping, they are R , $1/R$, $\log(R)$ and $\log(1/R)$, in which the curve of $\log(R)$ and $\log(1/R)$ are overlapping absolutely for $\log(1/R)$ equals $-\log(R)$. The zigzag green curve shows the correlation of C:N ratio and first derivative reflectance spectra. We can see that the effects of transformations such as $1/R$, $\log(R)$ and $\log(1/R)$ on correlation coefficients are not apparent. The maximums of r of those transformations locate at the same wavelengths,

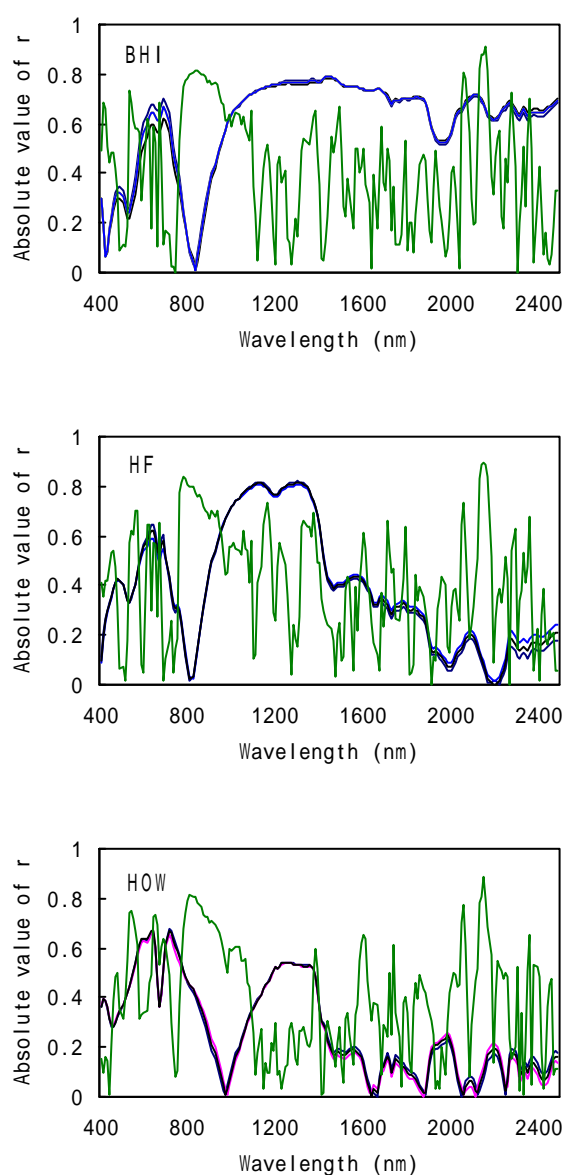


Fig. 2. Correlation coefficients of C:N ratio and different spectra in three sites

which are the same as those of the original reflectance. While the derivative transformation could make the maximum coefficients higher and the wavelengths of maximums are not the same as R. We should pay attention to two regions where coefficients are distinctively high, they are 800nm region and 2100nm region. The former is also called the “red edge” and the latter is where we discussed in Part 2.

2) Regression analysis

Regression results are listed in Table 2. R^2 is coefficient of definition. We choose the one when two wavelengths are selected into the regression equation on the case of R, 1/R, log(R), log(1/R) and dR. As to R', we choose the R^2 when three wavelengths are selected, because the wavelengths used in regression are limited in only one absorption feature. If we want to describe its shape, we need at least three points.

Table 2. Regression results

	Dependent	R^2	Selected wavelengths (nm)
BHI	R	0.775	1430, 2040
	1/R	0.756	1430, 2040
	log(R)	0.767	1430, 2040
	log(1/R)	0.767	1430, 2040
	dR	0.869	530, 2160
	R'	0.902	2052, 2110, 2188
HF	R	0.791	1190, 1300
	1/R	0.763	720, 1120
	log(R)	0.785	1190, 1300
	log(1/R)	0.785	1190, 1300
	dR	0.871	780, 2150
	R'	0.910	2054, 2084, 2180
HOW	R	0.744	710, 1140
	1/R	0.709	630, 1140
	log(R)	0.735	710, 1140
	log(1/R)	0.735	710, 1140
	dR	0.827	780, 2150
	R'	0.889	2058, 2086, 2210

5. Conclusions

Spectral transformation is a common-used operation prior to statistical modeling. 1/R, log(R) and log(1/R) are three main transformations which are simple non-linear transform on original reflectance, but they do not change the monotonicity. So we can find from the results of correlation and regression that their effects on C:N ratio inversion are not apparent. Derivative spectra express the change of reflectance at adjacent wavelengths, which is another kind of information compared to original reflectance. According to Fig. 2 and Table 2, we can conclude that the derivative transformation is good for CN ratio inversion. But its spectral curve always vibrates heavily, especially in infrared region, which is not good for following analysis. Sometimes, we have to run a risk of losing valuable information to smooth it by low-pass

filter. As to continuum-removed transformation, it is also an effective preprocessing for statistical modeling of C:N ratio inversion. The wavelengths of selected bands are uniform, physically meaningful and not site-specific. Its regression effects are satisfactory. We only consider one absorption feature in this study, and will pay attention to more features in future researches. All in one, derivative transformation and continuum removal are better than common-used reciprocal and logarithmic transformations in the statistical modeling of C:N ratio.

Acknowledgement

This research was supported by the National Natural Science Foundation of China (40271086) and the China's Special Funds for Major State Basic Research Project (G2000077900).

References

- [1] Yoder, B.J., Pettigrew-Crosby, R.E., 1995. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500nm) at leaf and canopy scales, *Remote Sensing of Environment*, 53:199-211.
- [2] Serrano, L., Penuelas, J., Ustin, S.L., 2002. Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals, *Remote Sensing of Environment*, 81:355-364.
- [3] Martin, M.E., Aber, J.D., 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen and ecosystem processes, *Ecological Applications*, 7:431-443.
- [4] Niu, Z., Chen, Y.H., Sui, H.Z., et al., 2000. Mechanism analysis of leaf biochemical concentration by high spectral remote sensing *Journal of Remote Sensing (China)*, 4(1): 46-50.
- [5] Kokaly, R.F., Clark, R.N., 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression, *Remote Sensing of Environment*, 67:267-287.
- [6] Shi, R.H., Niu, Z., Zhuang, D.F., 2003. Feasibility on estimate of leaf C:N ratio with hyperspectral data, *Remote Sensing Technology and Application (China)*, 18(2): 76-80.
- [7] Elvidge, C.D., 1990. Visible and near infrared reflectance characteristics of dry plant materials. *Int. J. Remote Sensing*, 11(10): 1775-1795.
- [8] Kokaly, R.F., 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration, *Remote Sensing of Environment*, 75: 153-161.
- [9] ACCP, 1994. Accelerated Canopy Chemistry Program Final Report to NASA-EOS-IWG. National Aeronautics and Space Administration, Washington, DC.
- [10] Newman, S. D., Soulia, M. E., Aber, J. D., Dewey, B., and Ricca, A., 1994. Analyses of forest foliage I: Laboratory procedures for proximate carbon fractionation and nitrogen determination, *J. Near-Infrared Spectrosc.*, 2: 5-14.
- [11] Bolster, K. L., Martin, M. E., and Aber, J. D., 1996. Determination of carbon fraction and nitrogen concentration in tree foliage by near infrared reflectance: a comparison of statistical methods. *Can. J. For. Res.*, 26: 590-600.