RETRIEVAL OF FUEL MOISTURE CONTENT FROM HYPERSPECTRAL DATA VIA PARTIAL LEAST SQUARE

Jie Zhang[1,2,3], Jian-Jun Wu[1,2,3,*], Lei Zhou[1,2,3]

1. State Key Laboratory of Earth Surface Process and Resource Ecology, Beijing Normal University, Beijing 100875, China; 2. Key Laboratory of Environment Change and Natural Disaster, MOE, Beijing Normal University, Beijing 100875, China; 3. Academy of Disaster Reduction and Emergency Management, MOCA/MOE, Beijing Normal University, Beijing 100875, China

1. INTRODUCTION

Accurate quantitative estimation of vegetation moisture status plays an important role in agriculture and forestry. Changes of vegetation water content can result in changes of spectral reflectance characteristics, which provides more potential to explore vegetation moisture status using spectral information. And vegetation water content expressed as Fuel Moisture Content (FMC), among vegetation moisture parameters, is commonly used for predicting vulnerability to wild fire [1-4]. Currently, FMC estimations employing spectral information are mainly carried out based on spectral indices derived from limited moisture feature bands and are more likely to overlook the underneath water information lying in the entire spectrum. Partial Least Square (PLS) is a new multivariate statistical method that can effectively reduce multi-fold linearity and it is commonly used in chemistry. This paper mainly explores the potential of PLS method for FMC retrieval based on LOPEX dataset. Firstly, the FMC retrieval model is established via PLS using the entire spectra and the performance of PLS to estimate FMC over various vegetation types is evaluated. Next, to lower model complexity, two band selection methods (Selection by Correlation Coefficient and Genetic Algorithm) are coupled with PLS (respectively called CC-PLS and GA-PLS in this paper). Also, the ability of CC-PLS and GA-PLS to extract FMC from LOPEX dataset is examined.

[*]Corresponding Author, E-mail: wjj@ires.cn

This research received financial support from the National High-tech R & D Program (863 Program) of China (grant number: 2006AA120108).

2. DATA

In order to have a wide range of water content variation, the data used in this paper is the LOPEX dataset established by the JRC in 1993. Samples of single leaf are selected for analysis and data involved here mainly include reflectance data & water content data(both fresh and dry weight).

3. METHODOLOGY

3.1. Fuel Moisture Content

Fuel Moisture Content (FMC) is defined as the ratio of vegetation water content (fresh weight minus dry weight) versus fresh weight [5].

$$FMC = \frac{FW - DW}{FW}$$

Where FM and DW respectively represent the fresh and dry weight of leaves.

3.2. Partial Least Square

Partial Least Square (PLS), first brought out by S. Wold and C. Albano in 1983, has been successfully applied in chemistry. It is a new multivariable statistical method and it can simultaneously accomplish regression modeling, data structure simplification and correlation analysis between two groups of variables. Compared to traditional multivariable linear regression, PLS regression model can still be developed even when independent variables have severe multi-fold correlation. One important thing in PLS analysis is to determine the optimal number of factors, and in this paper this is implemented via cross-validation. Besides, the auto-scale is adopted as the data preprocessing method .More detailed description about PLS can be found in [6].

3.3. Band Selection-Partial Least Square Coupled Method

3.3.1. CC-PLS

There exists correlation between leaf spectra and water content. However, this correlation varies with different bands and those band regions with good correlation can be called moisture sensitive bands. Thus, the correlation coefficients can be used for band selection. In this paper, 0.4 is defined as the threshold for band selection. Then, the selected bands enter into the following PLS analysis for FMC estimation.

3.3.2. GA-PLS

Genetic algorithms (GA) are a computer model to simulate natural selection. During this process, duplication, transformation & mutation are used to select the good individual and get rid of the bad, and eventually discover the best solution. Given the ability of genetic algorithms to simulate an individual's natural evolution, GA is well suitable for solving the problems of variable subset selection. In this paper, GA is used for band selection and then combined with PLS for FMC retrieval. More detailed description of GA-PLS method can be found in [7].

4. CONCLUSIONS

Table.1. Statistical Features of Sample FMC

	Min	*Max*	μ	σ	*CV*
FMC	0.0862	0.9264	0.6525	0.1385	21.23%

Table.2. Performance of FMC Estimation using PLS/CC-PLS/GA-PLS

	Calibration		Validation	
	R^2	RMSE	R^2	RMSE
PLS	0.890	0.045852	0.865	0.050992
CC-PLS	0.930	0.036724	0.922	0.038687
GA-PLS	0.893	0. 045311	0.869	0.050142

PLS is applied on LOPEX dataset to retrieve FMC from fresh leaf reflectance and the main conclusions are:

PLS shows great potential to estimate FMC from spectral data for both calibration and validation models. For PLS, the R^2 & RMSE is 0.890 & 0.045852 for calibration and 0.865 & 0.050992 for validation. When coupled with band selection methods, the models also show high estimation for both CC-PLS and GA-PLS (Despite influenced by GA, there may exist negligible variation between GA-PLS estimation accuracy corresponding to different GA runs). The CC-PLS model resulted in a determination coefficient (R^2) of 0.930 and a RMSE of 0.036724 for calibration, and 0.922 & 0.038687 for validation. Also, the GA-PLS model resulted in a determination coefficient (R^2) of 0.893 and a RMSE of 0. 045311 for calibration, and 0.869 & 0.050142 for validation. Besides, the band selection and PLS coupled model incorporates less bands, lowering the complexity of model. Thus, the high estimation accuracy and much easier modeling make the band selection-PLS coupled methods superior to original PLS. Moreover, compared to CC-PLS, GA can be automatically implemented in software package, making it much convenient for analysis despite relatively lower estimation accuracy than CC-PLS in this paper. The results demonstrate that PLS and PLS coupled method provide a new approach to extract water content from vegetation spectra and further the hyperspectral imagery.

5. REFERENCES

[1] S. L. Ustin, et al., “Estimating Canopy Water Content of Chaparral Shrubs Using Optical Methods,” *Remote Sensing of Environment,* vol. 65, no. 3, pp. 280-291, 1998.

[2] S. J. Pyne, P. L. Andrews, R. D. Laven, et al., “ Introduction to wild land fire,” Wiley: New York, 1996.

[3] E. Chuvieco, et al., “Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating,” *Remote Sensing of Environment,* vol. 92, no. 3, pp. 322-331, 2004.

[4] D. Riao, P. Vaughan, E. Chuvieco, P. J. Zarco-Tejada, S. L. Ustin, “Estimation of fuel moisture content by inversion of radiative transfer models to simulate equivalent water thickness and dry matter content. Analysis at leaf and canopy level,” *IEEE Transactions on Geo-Science and Remote Sensing,* vol. 43, no.4, pp. 819-826, 2005.

[5] M. Maki, M. Ishiahra, M. Tamura, “Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data,” *Remote Sensing of Environment,* vol. 90, no. 4, pp. 441-450, 2004.

[6] H. W. Wang, “Partial least-squares regression method and its application,” *National Defense Industry Press,* Beijing, 1999.

[7] L. Li, et al. "Retrieval of vegetation equivalent water thickness from reflectance using genetic algorithm (GA)-partial least squares (PLS) regression." *Advances in Space Research*, vol. 41, no. 11, pp. 1755-1763, 2008.