WAVELENGTH SELECTION OF HYPERSPECTRAL LIDAR BASED ON FEATURE WEIGHTING FOR ESTIMATION OF LEAF NITROGEN CONTENT IN RICE

Lin Du [a,b], Shuo Shi [b,c], Wei Gong [b,c,*], , Jian Yang [b], Jia Sun [b], Feiyue Mao[d]

[a] School of Physics and Technology, Wuhan University, Wuhan, China - conradlinyu@whu.edu.cn
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China - (conradlinyu, shishuo, sunjia, wind_yang, weigong)@whu.edu.cn
[c] Collaborative Innovation Centre of Geospatial Technology, Wuhan, China - (shishuo, weigong)@whu.edu.cn
[d] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China - maofeiyue@whu.edu.cn

Commission I, WG I/2

KEY WORDS: Hyperspectral LiDAR, Feather Weighting, Leaf Nitrogen Content, Wavelength Selection, Partial Least Squares, Support Vector Machine

ABSTRACT:

Hyperspectral LiDAR (HSL) is a novel tool in the field of active remote sensing, which has been widely used in many domains because of its advantageous ability of spectrum-gained. Especially in the precise monitoring of nitrogen in green plants, the HSL plays a dispensable role. The exiting HSL system used for nitrogen status monitoring has a multi-channel detector, which can improve the spectral resolution and receiving range, but maybe result in data redundancy, difficulty in system integration and high cost as well. Thus, it is necessary and urgent to pick out the nitrogen-sensitive feature wavelengths among the spectral range. The present study, aiming at solving this problem, assigns a feature weighting to each centre wavelength of HSL system by using matrix coefficient analysis and divergence threshold. The feature weighting is a criterion to amend the centre wavelength of the detector to accommodate different purpose, especially the estimation of leaf nitrogen content (LNC) in rice. By this way, the wavelengths high-correlated to the LNC can be ranked in a descending order, which are used to estimate rice LNC sequentially. In this paper, a HSL system which works based on a wide spectrum emission and a 32-channel detector is conducted to collect the reflectance spectra of rice leaf. These spectra collected by HSL cover a range of 538 nm – 910 nm with a resolution of 12 nm. These 32 wavelengths are strong absorbed by chlorophyll in green plant among this range. The relationship between the rice LNC and reflectance-based spectra is modeled using partial least squares (PLS) and support vector machines (SVMs) based on calibration and validation datasets respectively. The results indicate that **I**) wavelength selection method of HSL based on feature weighting is effective to choose the nitrogen-sensitive wavelengths, which can also be co-adapted with the hardware of HSL system friendly. **II**) The chosen wavelength has a high correlation with rice LNC which can be retrieved by using PLS and SVMs regression methods.

1. INTRODUCTION

Nitrogen is an important nutrient element mainly linked to the crop photosynthesis. Nitrogen availability is also a determinative factor of crop productivity (Van, K. H. et al., 1989). Adequate supply of nitrogen for crops insure the optimized yields, however it can in turn contribute to some other issues we should deal with, such as nitrate contamination of surface and ground water (Errebhi, M. et al., 1998; Zvomuya, F. et al., 2003). Additionally, over-fertilization with nitrogen will be a waste of resource and can also increase the cost of nitrogen fertilizer. Thus, efficient detection of nitrogen content and appropriate management of nitrogen fertilizer of crop in both canopy and leaf levels are essential to balance these factors.

In the past decades, the efficient methods of detecting growing status of crops is the multi- and hyperspectral remote sensing (Dalponte et al., 2013; Schlemmer et al., 2013). This method can offer researchers about biochemical parameters of crops such as the nitrogen content. Nevertheless, this passive technology can be effected by outside illumination conditions and is strongly sensitive with its change (Nevalainen, O. et al., 2013). In addition, the growing structure which is another important parameter used to indicate the crop growing status,

can not be got by this method. LiDAR has been used in many fields since the laser has been first introduced. It can be a powerful tool for crop structure and biomass detection because it can penetrate the crop canopy and receive some more useful information reflected by numerous layers of crops (Morsdorf et al., 2006). However LiDAR with only single wavelength is weak in collecting spectral properties, which limits its application in estimation of crop biochemical parameters. In recent years, attempts to combine the passive multi- and hyperspectral remote sensing technologies and LiDAR have been made in the field of remote sensing for precision nitrogen management (Wei, Gong et al., 2012; Nevalainen, O. et al., 2014; Niu Zheng et al., 2015).

Combining these two technologies requires different types of instruments such as the laser sources, waveform receiving components, which maybe increase the development cost of a new attempt to estimate the crop nitrogen content. Du Lin et al. (Du Lin et al., 2016) developed a hyperspectral LiDAR (HSL) by using a laser with wide spectrum range of about 600 nm to 2000 nm, and 100 mW average out-put power, which was applied in some practical fields, such as in estimating the vegetation nitrogen content. This HSL system has a multi-

---

channel detector for receiving echo signals and can precisely measure distance and abundant spectral properties.

Although this HSL system has a wide emission range, receiving all of those echo is usually difficult and unnecessary. Seeing from the biological knowledge, the reflectance spectrum can perform different properties in visible and infrared bands dependent on the absorption and reflectance ratio by chlorophyll (a and b). In addition, the so-called " red-edge " located in 700 nm can be more useful for detection of crop nitrogen content (Gitelson et al., 1996). Thus, selecting some nitrogen-sensitive wavelengths from these feature spectral ranges can be a efficient method of finding the great potential and flexibility of HSL in biochemical parameters estimation.

In this study, a wavelength selection method based on feather weighting (Huang, Rui et al., 2005) is proposed to characterize each channel of the HSL system. By this way some feature wavelengths related to nitrogen content can be picked up. This wavelength selection method can avoid some inherent limitations of the HSL hardware, such as signals obstructing by a geometry gap of the detector, " edge effect " of the detector channel and the signals interference in the adjacent or even the same channel. The regression results by using PLS and SVMs methods indicate that wavelength selected based on the feather weighting algorithm is strongly sensitive with nitrogen content and these feature wavelength can be used to estimate leaf nitrogen content (LNC) of rice.

2. MATERIALS AND METHODS

2. 1 Materials and experiments

The rice samples were grown in an agricultural production base located in SuiZhou city of China. Reflectance spectral properties of rice leaf differ from one growing stage to another after fertilization, especially in the stage of booting and heading. Thus, 240 samples (120 samples each stage) in these two stages were collected on July 16, 2014 and August 1, 2014 respectively.

The instrument used to collect reflectance spectrum of rice in leaf level is the HSL system developed by Du Lin et al (Du Lin et al., 2016). The HSL emits a wide-band "white" laser and collected echo signals with a high-precision grating spectrometer. The rice leaf samples are measured with 32 channels whose centre wavelength are listed in Table 1. Reflective factor (RF) of a white reference panel (>99% reflectance, Spectralon, Labsphere, Inc., North Sutton, NH) is first captured, then getting the reflectance of the rice leaf samples calculated as R_λ=leaf radiance / reference reflectance radiance at wavelength λ. Three separate spectra are measured and averaged as the characteristic spectra of each sample in a dark laboratory condition. On the same point of each measurement, five separate scans are made. After that the nitrogen contents of rice leaf samples are measured with a chemical extraction solvent.

Channel number	1	2	3	4	5	6	7	8
Central wavelength (nm)	538	550	562	574	586	598	610	622
Channel number	9	10	11	10	13	14	15	16
Central wavelength (nm)	634	646	658	670	682	694	706	718
Channel number	17	18	19	20	21	22	23	24
Central wavelength (nm)	730	742	754	766	778	790	802	814
Channel number	25	26	27	28	29	30	31	32
Central wavelength (nm)	826	838	850	862	874	886	898	910

Table 1. Central wavelengths of the HSL system.

2. 2 Feature weightings for each channel of HSL system

The wavelengths position located on the detector array is related to the type of grating spectrograph and the blazed wavelength of the grating. By controlling the rotator of the grating, wavelengths will be adapted to the detector channel well. Meanwhile, some other factors should be considered when using this HSL system in practise. The first considerable factor is that some sensitive wavelengths appear near the edge or be obstructed by the inherent gap of the detector array. Another is the signals interference from the neighbour or even the same channel. To solve these problems above, a wavelength selection method based on the feature weighing (Huang, Rui et al., 2005) is adopted in this study. By characterising each channel of HSL system with their corresponding feature weightings, the wavelengths being adverse to spectrum collection will be replaced by those that have a high feature weighting.

The feature weighting is calculated in two steps. The first is to calculate the divergence (D) of the same classes. Because the first principal components have explained the most information of the target, thus D can be modified as

$$D = \sum_{r=1}^{N} \eta \times D(r) \qquad (1)$$

where N = channel number
$D(r)$ = divergence of the rth channel, ($r = 1, 2, 3...N$)
η = function related to the rth eigenvalue of correlation matrix of this class

Secondly, calculate the feature weighting of each wavelength in this class.

$$\alpha_c(r) = \frac{1}{N} \sum_{j=1}^{R} \sigma_j \times (U_{rj})^2 \qquad (2)$$

where $\sigma_j = \dfrac{D(j)}{D}$ = ratio of divergence, $j = 1, 2, 3...N$
U_{rj} = a eigenvector matrix element of the correlation matrix of the cth class, $c = 1, 2, 3...n$

Equation (2) expresses the significant coefficient of the band r corresponding to class c, which is then sorted in a descending order. Seeking the rth channel position in this descending sequence in all classes (the number of class is n), the rth channel feature weight can be written as

$$\omega(r) = \frac{1}{N}\sum_{c=1}^{n}(N - V_c^r + 1) \qquad (3)$$

where $\quad V_c^r =$ position of band r in the descending sequence $\alpha_c(r)$ of class c.

In this study, the wavelengths with a higher $\omega(r)$ represent that this wavelengths are high correlated to LNC, and then they will be used to estimate the nitrogen status of rice.

2.3 Regression methods

Vapnick's theory stated that SVMs could find a hyper-plane which was not only used to separate all of classes but also used for regression problems with a ε function (Vapnick, Vladimir N., 1998). A small ε corresponds to a small regression error, which can be optimised by choosing penalty parameter(c) and the kernel parameter(γ) in the sets [-5, -4,…, 0,…, 4, 5] and [-10, -9,…, 0,…, 9, 10] respectively. Unlike SVMs, PLS is based on the principal component analysis (PCA), and it relates the input variables (X) and the responding variables (Y) with a linear relationship using a few independent principal components (PCs) (P. Hansen et al., 2003; Li Fei et al., 2014).

3. RESULTS AND DISCUSSION

Nitrogen effects the process of vegetation photosynthesis, and is also related to chlorophyll synthesis. Thus, in different nitrogen levels, the reflectance spectrum of rice leaf are different (Gitelson et al., 1996; Broge and Leblanc, 2001; Zhu et al., 2008). According to this relationship, LNC of rice can be retrieved from reflectance spectra collected by the 32-channel HSL system. Figure 1 shows the reflectance spectrum in visible and near infrared band collected by this HSL system. The 32 channels position showed in Figure 1 is separated with dashed lines.
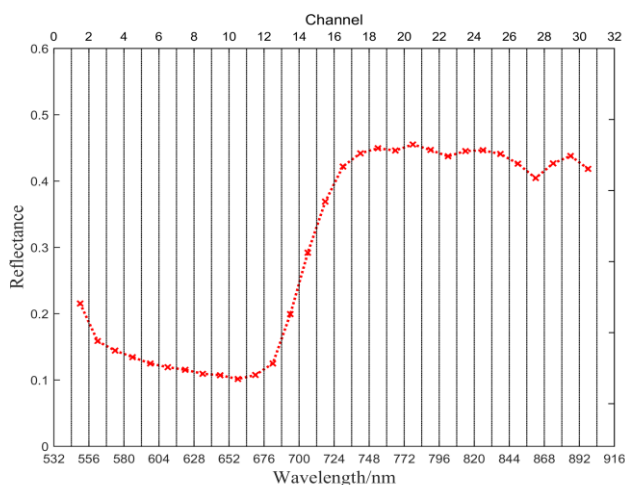

Figure 1. The reflectance spectrum and the channel position of the HSL system.

Figure 2 shows the coefficient of determination (R^2) between the observed and estimated LNC by using SVMs and PLS methods after selecting feature wavelengths from 32 wavelengths of the HSL system. The results indicate that these selected feature wavelengths are high correlated to LNC of rice,

and they can be used to estimate biochemical parameters, especially the nitrogen. Generally, the R^2 of regression models are > 0.7 with PLS and >0.9 with SVMs.
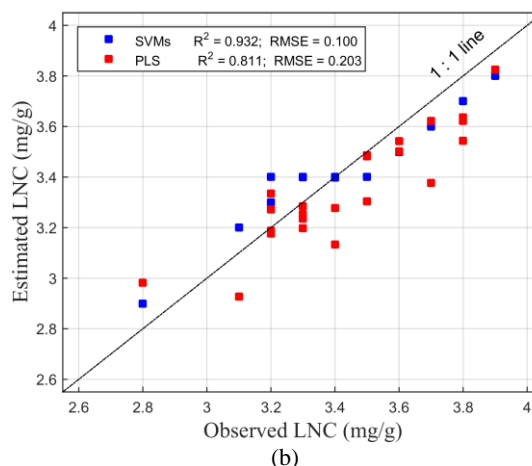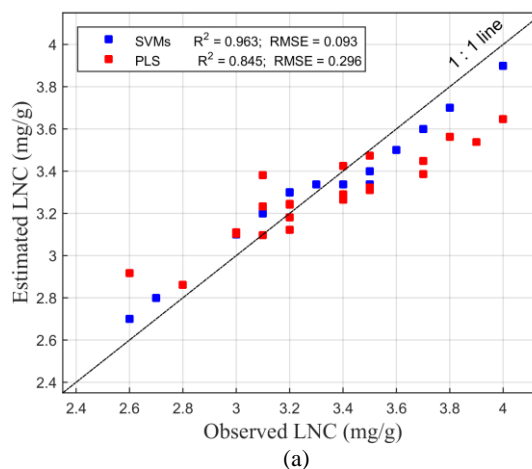

(a)

(b)

Figure 2. R^2 and RMSE of estimation models using SVMs and PLS to estimate the LNC of rice in (a) booting and (b) heading stages.

Du Lin et al. (Du Lin et al., 2016) adopted SVMs to retrieve rice LNC and the R^2 of retrieval models could be more than 0.7 by using all of the feature wavelengths. According to the method of wavelength selection introduced in this study, 32 feature wavelengths are characterised with their corresponding feature weightings. Instead using all of the feature wavelengths, those with a higher weighting are chosen as the inputs of regression models. Both in the booting and heading stages, SVMs with BRF kernel function behave great ability in estimating rice LNC, namely, R^2=0.963, RMSE=0.093 in booting stage, and R^2=0.932, RMSE=0.1 in heading stage. In addition, the predicted LNC values disperse on both sides of the 1:1 line, which means that SVMs and PLS models with the characterised wavelengths have a higher reliability in rice LNC estimation. The maximum R^2 of PLS models are >0.8. Compared with the results of PLS models, it can be found the SVMs with RBF function show a good estimation performance of rice LNC, thus a non-linear relationship maybe more suitable between the LNC and reflectance spectrum.

Some wavelength being poorly correlated with the rice LNC will not be helpful for LNC retrieval, thus they have been

picked up when modeling the relationship between LNC of rice and reflectance spectrum. The remaining feature wavelength combinations with high feature weightings are conducted to estimate the rice LNC. After experiments the chosen wavelength combinations using SVMs method with a R^2 of >0.8 in booting stage are 754 nm and 694 nm. In heading stage, 910 nm and 634 nm can be used to retrieve LNC of rice with a R^2 of >0.9. When using the PLS method with a maximum R^2, the wavelength combinations become 682 nm, 850 nm, 754 nm, 838 nm, 802 nm, 634 nm, 658 nm, 730 nm, 646 nm in booting stage and 910 nm, 589 nm, 802 nm, 838 nm, 706 nm, 622 nm, 874 nm, 646 nm, 586 nm, 682 nm, 658 nm, 610 nm, 694 nm in heading stage. These bands located in visible and near-infrared range, in which exists strong absorption by chlorophyll of green plant.

These results indicate that although the centre wavelengths of HSL are chosen based on the sensitivity to LNC of rice, we can pick up some of them to estimate LNC without decrease the retrieval accuracy using a weighting criterion. Thus, the selection method of feature wavelength introduced in this study are efficient for finding nitrogen-sensitive bands for LNC retrieval. As we all known, chlorophyll in plant is also an important indicator of crops growing status, and the nitrogen content and chlorophyll are linked each other, so to find some spectral indices related to chlorophylls absorption maybe helpful in LNC estimation. Besides, the relationship between LNC and reflectance spectrum is complex, a typical model, such as the PROSPECT (Gastellu-Etchegorry and Bruniquel-Pinel 2001; Tucker, 1979) maybe tell us more about the biophysical process in plant.

In the future, HSL will be developed and applied in many more fields because of its flexibility and great advantage of spectrum collection combined with structure parameters of the targets. Like the passive remote sensing technology, such as some airborne sensors, MODIS and AVHRR, HSL may be carried by a airborne platform and used for global environment monitoring.

4. CONCLUSION

HSL system has great potential in LNC estimation of rice. This study assigns a feature weighting to each channel of HSL system by using matrix coefficient analysis and divergence threshold. By re-ranking the feature wavelength of HSL system, the wavelengths high-correlated to the LNC will be used to estimate rice LNC sequentially by the means of SVMs and PLS regression. The experiment gain a R^2 of >0.8, which shows that these chosen wavelength has a high correlation with rice LNC, and the wavelength selection method of HSL based on feature weighting is effective in choosing the nitrogen-sensitive wavelengths, which can also be co-adapted with the hardware of HSL system friendly.

ACKNOWLEDGEMENTS

REFERENCES

Broge, Niels Henrik, Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sens. Environ. 76 (2), pp. 156－172.

Dalponte, Michele, Orka, Hans Ole, Gobakken, Terje, Gianelle, Damiano, Næsset, Erik, 2013. Tree species classification in boreal forests with hyperspectral data. IEEE Trans. Geosci. Remote Sens. 51 (5), pp. 2632－2645.

Du Lin, Wei, Gong, Shi, Shuo , Yang, Jian, Sun, Jia, Zhu, Bo, Song, Shalei, 2016. Estimation of rice leaf nitrogen contents based on hyperspectral LIDAR. International Journal of Applied Earth Observation and Geo-information, 44, pp. 136-143.

Errebhi, M.; Rosen, C. J.; Gupta, S. C.; Birong, D. E, 1998. Potato yield response and nitrate leaching as influenced by nitrogen management. Agron. J. , 90, pp. 10-15.

Gastellu-Etchegorry, J.P., Bruniquel-Pinel, V., 2001. A modeling approach to assess the robustness of spectrometric predictive equations for canopy chemistry. Remote Sens. Environ. 76 (1), pp. 1－15.

Gitelson A., M. N. Merzlyak, and H. K. Lichtenthaler, 1996 Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. Journal of Plant Physiology, 148, pp. 501-508.

Gong Wei, Song, Shalei, Zhu, Bo, Shi, Shuo, Li, Faquan, Cheng, Xuewu, 2012. Multi-wavelength canopy LiDAR for remote sensing of vegetation: design and system performance. ISPRS J. Photogramm. Remote Sens. 69, pp. 1－9.

Gitelson, Anatoly A., Merzlyak, Mark N., Lichtenthaler, Hartmut K., 1996. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. J. Plant Physiol. 148 (3), 501－508.

Hansen P. and J. Schjoerring, 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression, Remote sensing of environment, vol. 86, pp. 542-553.

Huang, Rui, He, Mingyi, 2005. Band selection based on feature weighting for classification of hyperspectral data. IEEE Geosci. Remote Sens. Lett. 2 (2), pp. 156－159.

Li Fei, B. Mistele, Y. Hu, X. Chen, and U. Schmidhalter, 2014. Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression, European Journal of Agronomy, 52, pp. 198-209.

Morsdorf, Felix, Kötz, Benjamin, Meier, Erich, Itten, K.I., Allgöwer, Britta, 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. Remote Sens. Environ. 104 (1), pp. 50－61.

Nevalainen, Olli, Hakala, Teemu, Suomalainen, Juha, Mäkipää, Raisa, Peltoniemi, Mikko, Krooks, Anssi, Kaasalainen, Sanna, 2014. Fast and non-destructive method for leaf level

chlorophyll estimation using hyperspectral LiDAR. Agr. Forest Meteorol. 198, pp. 250‒258.

Nevalainen, Olli, Hakala, Teemu, Suomalainen, Juha, Kaasalainen, S., 2013. Nitrogen concentration estimation with hyperspectral LiDAR. ISPRS Ann. Photogram. Remote Sens. Spatial Inf. Sci. II-5/W2, pp. 205 – 210, http://dx.doi.org/10.5194/isprsannals-II-5-W2-205-2013.

Niu Zheng, Z. Xu, G. Sun, W. Huang, L. Wang, M. Feng, et al., 2015. Design of a New Multispectral Waveform LiDAR Instrument to Monitor Vegetation. IEEE Geoscience and Remote Sensing Letters, 12(7), pp. 1506-1510.

Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J.,Rundquist, D., 2013. Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. Int. J. Appl. Earth Obs. Geoinf. 25, pp. 47‒54.

Tucker, Compton J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote sens. Environ. 8 (2), pp. 127‒150.

Van, K. H.; Goudriaan, J.; Seligman, N. G, 1989. Modeling the effects of nitrogen on canopy development and crop growth. Semin. Ser. Soc. Exp. Biol., 31, pp. 83-104.

Vapnick, Vladimir N., 1998. Statistical Learning Theory. J. Wiley and Sons Inc., Nova York.

Zhu, Yan, Yao, Xia, Tian, YongChao, Liu, XiaoJun, Cao, WeiXing, 2008. Analysis of common canopy vegetation indices for indicating leaf nitrogen accumulations in wheat and rice. Int. J. Appl. Earth Obs. Geoinf. 10 (1), pp. 1‒10.

Zvomuya, F.; Rosen, C. J.; Russelle, M. P.; Gupta, S. C, 2003. Nitrate leaching and nitrogen recovery following application of polyolefin-coated urea of potato. J. Environ. Qual., 32, pp. 480-489.