University of Kentucky
UKnowledge

International Grassland Congress Proceedings | XXIII International Grassland Congress

Estimating δ15N and δ13C in Barley and Pea Mixtures Using Near-Infrared Spectroscopy with Genetic Algorithm Based Partial Least Squares Regression

Hyo-Won Lee
Korean National Open University, South Korea

Hyo-Jin Lee
Sungkyunkwan University, South Korea

Han-Jong Ko
Korean National Open University, South Korea

Chang Yoon
Chonbuk National University, South Korea

Follow this and additional works at: https://uknowledge.uky.edu/igc

Part of the Plant Sciences Commons, and the Soil Science Commons

This document is available at https://uknowledge.uky.edu/igc/23/2-3-2/6

The XXIII International Grassland Congress (Sustainable use of Grassland Resources for Forage Production, Biodiversity and Environmental Protection) took place in New Delhi, India from November 20 through November 24, 2015.

Proceedings Editors: M. M. Roy, D. R. Malaviya, V. K. Yadav, Tejveer Singh, R. P. Sah, D. Vijay, and A. Radhakrishna

Published by Range Management Society of India

This Event is brought to you for free and open access by the Plant and Soil Sciences at UKnowledge. It has been accepted for inclusion in International Grassland Congress Proceedings by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Paper ID: 492
Theme 2. Grassland production and utilization
Sub-theme 2.1. Quality, production, conservation and utilisation

Estimating δ15N and δ13C in barley and pea mixtures using near-infrared spectroscopy with genetic algorithm based partial least squares regression

Hyo-Won Lee[1]*, Hyo-Jin Lee[2], Han-Jong Ko[1], Chang Yoon[3]
[1]Department of Agriculture, Korean National Open University, Seoul, Korea
[2]Department of Landscape Architecture, Sungkyunkwan University, Suwon, Korea
[3]Department of Animal biology, Chonbuk National University, Jeonju, Korea
*corresponding author e-mail : hyowon@knou.ac.kr

Keywords: Barley and pea mixtures, Carbon and nitrogen isotope, Genetic algorithm, Near-infrared spectroscopy

Introduction

Stable isotope measurements have been increasingly used as a method to obtain information on relationships between plants and their environment (Dawson *et al.*, 2002). Stable isotopes are seen as a powerful tool for advancing our knowledge on stock cycling and, nitrogen and carbon isotopic compositions have provided key insights into biogeochemical interactions between plants, soils and the atmosphere (Robinson, 2001). For the stable isotope measurements, the $\delta^{13}C$ isotopic signature has been used successfully to disentangle physiological, ecological and biogeochemical processes and, $\delta^{15}N$ studies have significantly improved our knowledge on nitrogen cycling pathways and nitrogen acquisition by plants (Vallano and Sparks, 2008).

For the stable isotope measurements, traditional laboratory methods using isotope analysis are accurate and reliable, but usually time-consuming and expensive. Near-infrared spectroscopy (NIRS) analysis provides rapid, accurate and less expensive estimation. NIRS have been made to estimate herbage parameters using statistical methods such as multiple linear regression and partial least square regression (PLSR). PLSR uses all available wavebands in multivariate calibration for quantitative analysis of the spectral data. However, previous studies indicated that PLSR with waveband selection might improve their predictive accuracy in multivariate calibration at laboratory (Leardi, 2000) and the selection of appropriate wavelengths can refine the predictive accuracy of the PLS model by optimizing important spectral wavebands both in laboratory NIRS (Jiang *et al.*, 2002). To optimize important spectral wavebands by wavelength selection, genetic algorithms (GA) is widely used, because GA has the ability to simulate the natural evolution of an individual and GA is well suited for solving variable subset selection problems (Ding *et al.*, 1998).

Barley and pea mixture is one of the most important forage species for livestock farming in Korea. To investigate nitrogen fixation and transfer in barley and pea mixture, stable isotope measurements was widely used. However, there was no research to estimate stable isotope in barley and pea mixture using NIRS in Korea. The aim of this study was to investigate performance of NIRS with PLSR using genetic algorithms based wavelength selection (GA-PLSR) and compare with PLSR without wavelength selection (FS-PLSR) for the estimation of $\delta^{15}N$ and $\delta^{13}C$ in barley and pea mixture.

Materials and Methods

The study site was controlled from May to June in 2008 in Incheon City, Korea. A completely randomized block design with three replications was used for the experiment and one reference plot assigned each treatment for nitrogen fixation evaluation. Seeding mixture was 40 kg barley and 80 kg pea per ha. N rate of 40, 80 and 120 kg/ha as fertilizer was applied at seeding. Forage was harvested from each plot in ripening stage at ground level and 113 samples were collected. Hyperspectral reflectance were collected from 113 samples using Foss-6500 (NIRS system, Inc., USA). Total of 1051 wavebands between 500-2500 nm were collected and spectral data were smoothed. Isotope analysis were carried out to estimate Total N, Total C, $\delta^{15}N$ and $\delta^{13}C$ at the National Instrumentation Center for Environmental Management, College of Agriculture and Life Sciences, Seoul National University.

PLSR analysis was performed in the reflectance to estimate Total N, Total C, $\delta^{15}N$ and $\delta^{13}C$ using PLS toolbox version 6.5 (Eigenvector Research, Inc., WA) in Matlab software version 8.10 (Mathworks Inc., Sherborn, MA, USA). The GA-PLS regression was conducted using a GA program in Matlab (Leardi, 2000), which was designed to minimize a risk of over-fitting due to the large number of variables in hyperspectral dataset. The suitability in each dataset for applying GA bands selection was conducted by fitness function (Leardi, 2000; Leardi *et al.*, 2002) before the GA run, and five GA runs were performed because each GA-PLS gave a slightly different model. The predictive ability of the models was evaluated by the cross-validated coefficient of determination (R^2_{cv}) and the root mean square error of calibration ($RMSE_{cv}$).

Results and Discussion

Coefficient of determination in FS-PLSR and GA-PLSR to predict Total N, Total C, $\delta^{15}N$ and $\delta^{13}C$ using reflectance are shown in Table 1. GA-based wavelength selection in PLS calibration suggested that the important hyperspectral wavebands for estimating $\delta^{15}N$ are 8.8% and $\delta^{13}C$ are 15.8% of all 2001 wavebands in the 500–2500 nm range. In every parameter, GA-PLSR models showed higher R^2_{cv} (Total N, R^2_{cv} = 0.70; $\delta^{15}N$, 0.69; Total C, 0.55; $\delta^{13}C$, 0.66) than that of FS-PLSR (Total N, R^2_{cv} = 0.62; $\delta^{15}N$, 0.66; Total C, 0.39; $\delta^{13}C$, 0.23), and GA-PLSR models showed lower RMSEcv (Total N, $RMSE_{CV}$ = 0.53; $\delta^{15}N$, 1.04; Total C, 1.11; $\delta^{13}C$, 0.51) than that of FS-PLSR (Total N, $RMSE_{CV}$ = 0.69; $\delta^{15}N$, 1.37; Total C, 1.39; $\delta^{13}C$, 0.65). These results confirm that the Total N, Total C, $\delta^{15}N$ and $\delta^{13}C$ of barley and pea mixture can be more accurately predicted by GA-PLSR than FS-PLSR in the spectroscopy.

Table 1. Coefficient of determination (R^2), root mean square error of calibration (RMSEcv), number of selected waveband (NW) and percentage of number of selected waveband (NW%) from FS-PLS and GA-PLS regression to predict Total N, $\delta^{15}N$, Total C and $\delta^{13}C$.

Herbage variable	FS-PLS		GA-PLS			
	R^2_{cv}	$RMSE_{CV}$	R^2_{cv}	$RMSE_{CV}$	NW	NW%
Total N	0.62	0.69	0.70	0.53	199	19.0
$\delta^{15}N$	0.66	1.37	0.69	1.04	92	8.8
Total C	0.39	1.39	0.55	1.11	190	18.1
$\delta^{13}C$	0.23	0.65	0.66	0.51	161	15.3

Conclusion

We investigated the performance of the GA-PLSR model in the spectral assessment of Total N, Total C, $\delta^{15}N$ and $\delta^{13}C$ of barley and pea mixtures. The results indicated that the wavelength selection procedure using GA-PLSR refined the predictive ability expected by optimization of the wavelength subset, and Total N, Total C, $\delta^{15}N$ and $\delta^{13}C$ of barley and pea mixtures can be more accurately estimated by GA-PLS than by FS-PLS.

References

Dawson, T. E., S. Mambelli, A. H. Plamboeck, P. H. Templer and K.P. Tu. 2002. Stable isotopes in plant ecology. *Annu. Rev. Ecol Syst.* 33: 507-559.

Ding, Q., G. W. Small and M. A. Arnold. 1998. Genetic algorithm based wavelength selection for the near-infrared determination of glucose in biological matrixes: initialization strategies and effects of spectral resolution. *Anal. Chem.* 70**:** 4472-4479.

Jiang, J. H., R. J. Berry, H. W. Siesler and Y. Ozaki. 2002. Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data. *Anal. Chem.* 74**:** 3555-3565.

Leardi, R. 2000. Application of genetic algorithm-PLS for feature selection in spectral data sets. *J. Chemometrics* 14**:** 643-655.

Leardi, R., M. B. Seasholtz and R. J. Pell. 2002. Variable selection for multivariate calibration using a genetic algorithm: prediction of additive concentrations in polymer films from Fourier transform-infrared spectral data. Anal. Chim. Acta. 461**:** 189-200.

Robinson, D. 2001. Delta N-15 as an integrator of the nitrogen cycle. *Trend. Ecol. Evol.* 16**:** 153-162.

Vallano, D. M. and J. P. Sparks. 2008. Quantifying foliar uptake of gaseous nitrogen dioxide using enriched foliar delta N-15 values. *New Phytol.* 177**:** 946-955.