NIR TECHNIQUE IN THE CLASSIFICATION OF COTTON LEAF GRADE

Yongliang Liu
USDA, ARS, Cotton Structure & Quality Research Unit
New Orleans, LA
Jonn Foulk
FX-Fibers LLC
Clemson, SC

Abstract

NIR spectroscopy, a useful technique due to the speed, ease of use, and adaptability to on-line or off-line implementation, has been applied to perform the qualitative classification and quantitative prediction of cotton quality characteristics, including trash index. Cotton classification is performed using the HVI^{TM} system and one term to assess the degree of trash amount is leaf grade. In this study, visible/NIR spectra were collected to explore the feasibility for the discrimination of cotton samples with various leaf grade categories. The results suggested that the optimal model in the 1105-1700 nm region could determine the cotton leaf grade with a success rate of ~95.0%.

Introduction

As a universal testing method and official classification, the USDA's Agricultural Marketing Service (AMS) has implemented the high volume instrument (HVI^{TM}) cotton classification procedures to identify the number of non-lint particles and to measure the surface area covered by non-lint particles (ASTM, 1997). Qualified AMS human classers manually determine the leaf grade and extraneous matter through a process of visually examining the cotton samples and them comparing them to the Universal Cotton Standards.

In July 2011, AMS recommended the instrumental leaf grade as the official leaf grade. The instrumental leaf grade was determined by an equation that utilizes HVI^{TM} trash readings of percent area and particle count. On trials during the 2009 and 2010 cotton classing seasons, AMS thoroughly analyzed the comparisons of instrumental results to the official manual classifications and the Universal Cotton Standards. The results showed that the HVI^{TM} was capable of determining leaf grade as defined by the Universal Cotton Standards for leaf grade more accurately than cotton classers, as factors such as grader fatigue and central tendency were eliminated (AMS, 2012).

Near infrared (NIR) spectroscopy has been attempted for the quantitative prediction of HVI^{TM} trash parameters (Liu et al., 2010). Because of high diversification of trash types and their heterogeneous distribution, the reported results have not yet proven satisfactory. Thus, a 90% confidential interval was subjectively applied to remove outlier samples and the re-developed NIR models indicated the promise for practical applications.

To examine the determination of leaf grade via NIR technique (including the visible (400-750 nm) region), this study explored the approach of soft independent modeling of class analogy of principal component analysis (SIMCA/PCA) to interpret the spectra of diversified commercial cottons.

Materials and Methods

Cotton Fibers and Official Instrumental Leaf Grade

A total of 650 raw cotton samples (350 from the 2010 crop year and 300 from the 2011 crop year) and their HVI^{TM} leaf grade assignments were utilized. These samples represented diversities in upland cotton varieties, growth locations, and ginning practices. Spectra were acquired following a standard procedure of conditioning at a constant relative humidity of 65 ± 2% and temperature of 21 ± 2 ºC for at least 48 hours.

Visible/NIR Reflectance Spectral Measurement

About 10 g of cotton fibers was loaded into a Foss coarse granular cell (3.8 cm-wide x15.2 cm-long x 4.8 cm-depth) and scanned on a FOSS XDS rapid content analyzer (FOSS NIRSystems Inc., Laurel, MD). To keep a good contact between sample and optical window, During NIR testing, 750 g of extra weight was placed on top of a sample creating ~0.18 PSI for good contact between the sample and optical window. A background was recorded with a built-in internal reference before acquiring the samples. The log (1/R) readings were collected over the visible/NIR range of 400 - 2500 nm at 0.5 nm interval and 32 scans. Mean spectrum of averaging three spectra for each sample by repacking was taken for model development.

SIMCA/PCA Classification Model Development

All visible/NIR spectra were imported into PLSplus/IQ package in Grams/AI (Version 7.01, Galactic Industrious Corp., Salem, NH, current part of Thermo Fisher Scientific) for discriminant model development. Classification models were generated using seven classes (representing leaf grade 1 through 7) with mean centering (MC) and the first (1^{st}) derivative spectral pretreatment in various spectral regions. For each of the seven classes in different models, the optimal number of factors was used as suggested by the software package. Applying the SIMCA/PCA models to validation (and calibration) samples and employing the class assignment rule of the least spectral residual values, the sample was identified as belonging in the class being modeled, i.e., one of seven groups ranging from leaf grade 1 to 7 categories. During the SIMCA/PCA process, leaving one-out cross-validation method was used.

Results and Discussion

Visible/NIR Spectral Response to Leaf Grade

Figure 1 shows the average visible/NIR log (1/R) spectra of cotton fibers with leaf grade 1, 3, 5, and 7, respectively. Although cotton fibers with low leaf grade have the visible/NIR bands in common with ones having high leaf grade, there are some intensity variations induced by the leaf grade elevating. For instance, the spectra of those with high leaf grade show great log (1/R) intensity increase in the spectral region of visible/short-wavelength (SW) NIR region (< 1100 nm) and also relatively weak intensity reduction in the range from 1100 to 2500 nm. This is expected, because high leaf grade samples should have different spectral characteristics compared to low leaf grade cottons.

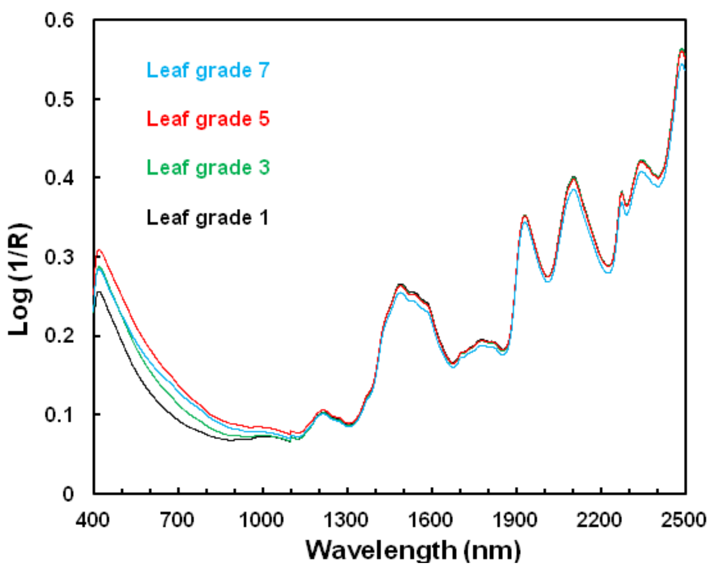

Figure 1. Typical visible/NIR log (1/R) spectra of cotton fibers with various leaf grades.

Calibration and Validation Samples

Table 1 summarizes the assignment of 442 calibration and 208 validation samples. Briefly, the 2010 and 2011 crop year cottons were tested at two ARS locations (Clemson, SC and New Orleans, LA). All samples having the same leaf grade were assigned to different identification numbers indiscriminately. On the increasing order of sample identification numbers, every 3^{rd} sample was used for model validation while remaining spectra were used for SIMCA/PCA classification model development.

Table 1. Sample distribution and assignment in calibration and validation set.[a]

Leaf grade	1	2	3	4	5	6	7	total
Year 2010	50	50	50	50	50	50	50	350
Year 2011	50	50	50	50	50	50	0	300
Calibration no.	68	68	68	68	68	68	34	442
Validation no.	32	32	32	32	32	32	16	208

[a] Leaf grade 8 and partial leaf grade 7 samples were not available

SIMCA/PCA Classification Models

The models were developed using different combinations of full / narrow spectral regions and spectral pre-treatment of MC + 1st derivative. For each model, the classification efficiency was assessed by the sample numbers corrected classified, and the statistics in calibration and validation sets from various spectral regions are compared in Table 2.

Table 2. Seven-class classification of leaf grade 1 to 7 in calibration and validation set.[a]

Spectral region (nm)	405-2495	405-1095	1105-2495	**1105-1700**	900-1700
No. correct classified (valid)	170	161	175	**187**	178
% correct classified (valid)	81.7	77.4	84.1	**89.9**	86.6
No. correct classified (calib)	435	421	432	**442**	441
% correct classified (calib)	98.4	95.2	97.7	**100**	99.8
% average [b]	90.0	86.3	90.9	**95.0**	93.2

[a] The number of calibration (calib) and validation (valid) samples was 442 and 208, respectively.
[b] Mean of % correct classification in calibration and validation set.

Use of discriminant model from the 405-2495 nm region yielded the correct identification of 170 and 435 samples in respective validation and calibration set, giving a 90.0% of overall classification. The 1105-2495 nm model slightly improved the accuracy of correct identification from 90.0% to 90.9%, but the model from the 405-1095 nm region compromised the identification to 86.3%. However, it is encouraging to observe a more enhanced separation of over 93.0% by the models from the 1105-1700 nm and 900-1700 nm region. Notably, the 1105-1700 nm model revealed a classification success of ~ 95.0%, with an 89.9% correct identification in validation set and a 100% determination in calibration set (shown in **Bold**). Despite more apparent intensity variations in the 400-1100 nm region than in the 1100-2500 nm region (Fig.1), the optimal model in Table 2 utilized the portion bands from the latter range. A probable rational might be due to the increase in non-leaf trash among higher leaf grade cottons.

Summary

The result demonstrates the usefulness and effectiveness of visible/NIR technique in determining leaf grade of commercial cotton fibers. Comparison of seven-class SIMCA/PCA models in different spectral regions revealed the optimal classification efficiency of ~95.0% from the 1105-1700 nm NIR region. The observation is most promising in the development of optical spectral sensing system for *in-situ* rapid measurement of leaf grade at cotton facilities.

Acknowledgements

Portion of work was done at ARS Clemson facility (officially closed in November 2011). We sincerely thanked Mr. J. Knowlton (USDA, AMS, Memphis, TN) for meeting our request of providing the diversified cotton samples and also Ms. M. Morris (retired from ARS) for the technical assistance.

Disclaimer

Mention of a product or specific equipment does not constitute a guarantee or warranty by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products that may also be suitable.

References

Agricultural Marketing Service (AMS). 2012. Revision of cotton classification procedures for determining cotton leaf grade. Federal Register. 77 (66): 20503-20505.

American Society for Testing and Materials (ASTM). 1997. *Standard test method for measurement of physical properties of cotton fibers by high volume instruments* (Designation: D5867-12). ASTM International, West Conshohocken, PA.

Liu, Y., G. Gamble, and D. Thibodeaux. 2010. Evaluation of 3 cotton trash measurement methods by visible / near-infrared reflectance spectroscopy (ASABE Paper No. 1008708), American Society of Agricultural and Biological Engineers, St. Joseph, MI.