SPECTROMETERS | LASERS | TOTAL SOLUTIONS

Quantitative Prediction of Tobacco Components using Near-Infrared Diffuse Reflectance Spectroscopy

Kristen Frano
Katherine Bakeev
B&W Tek, Newark, DE

Introduction

Chemical analysis is an extremely important task in the tobacco processing industry. The composition of the tobacco leaf varies depending on where it is grown and the local climate. For a tobacco processing plant, an understanding of the complex chemical composition of the incoming raw tobacco is essential to maintaining product consistency throughout the manufacturing stages and is especially important for new product formulation. In some cases, tobacco processors will not provide full compensation for tobacco growers if their tobacco is found to contain more than the acceptable chlorine content, which can create an additional burden for the testing laboratories. Although automatic analyzers (e.g. continuous segmented flow analyzers) have traditionally been the standard technique for quality control laboratories in the tobacco industry, these analyses require highly skilled personnel and are expensive and time-consuming in both analysis and instrument upkeep. Automatic analyzers can also require the use of hazardous chemical reagents for colorimetric determination of some tobacco components, which generates waste and can be harmful to employees without the proper protection. For instance, nicotine content in tobacco is determined by measuring the color resulting from the tobacco reacting with sulfanilic acid and cyanogen chloride, which is generated *in situ* by the reaction of hazardous chemicals chloramine T (which needs to be prepared fresh every five days) and potassium cyanide (which needs to be prepared fresh every two weeks).

Recently, near-infrared (NIR) spectroscopy has been explored as a complementary analytical technique alongside automatic analyzers for large-volume quality control analysis of tobacco due to its low-cost, speed, ease of analysis, and ability to analyze multiple components with one spectrum using chemometric models. NIR spectroscopy is also a nondestructive analysis, meaning the tobacco samples can be analyzed and then reused, reducing the costs and waste associated with wet chemistry techniques. This is especially beneficial for geographic regions susceptible to drought, where access to water for sample preparation can be restricted, causing delays in analysis and production costs to skyrocket.

In this study, we demonstrate the ability to accurately predict the amount of nicotine, sugar, and chloride in tobacco using the i-Spec® Plus, a portable diffuse reflectance NIR spectrometer with on-board chemometric software from B&W Tek.

Experimental

Instrumentation

Dry, ground tobacco samples were provided over one growing season by a tobacco growing and processing company in South Africa. The company receives tobacco from different regions within South Africa and in various curing conditions, resulting in tobacco samples that can range widely in their nicotine and sugar contents.

NIR measurements were obtained on the i-Spec® Plus 2.2 Diffuse Reflectance NIR Spectrometer (shown in Figure 1) with on-board software, measuring 1100-2200nm with a resolution of 10 nm. All tobacco samples were placed in a petri dish and analyzed on a turntable adapter on top of the i-Spec® Plus in order to take into account any sample heterogeneity. Integration times were 600 µs and averaged 4,000 times per acquisition for a total acquisition time of 2.4 seconds.

Fig. 1 i-Spec® Plus Diffuse Reflectance Portable NIR Spectrometer with turntable

Chemometrics

BWIQ® software was used to create predictive chemometric models for three tobacco components: nicotine, sugar, and chloride. Flue-cured and air-cured tobaccos were used for the creation of the calibration model. The reference values for the percent content of each calibration sample were provided by the tobacco company based on standard British American Tobacco (BAT)-approved methods: segmented flow analysis for nicotine and sugar and potentiometric titration for chloride.

Table 1 shows the information for the tobacco samples used to develop the three chemometric models. The models contain samples that cover the range of typical nicotine, sugar, and chloride concentrations across different tobacco curing methods and were expanded to include some samples with out of specification values.

Table 1: Tobacco sample information

Tobacco component	# of calibration samples	% content range	# of validation samples
Nicotine	~2,000 flue-cured ~700 air-cured	0.091-7.37	250
Sugar		0.28-25.12	
Chloride	~1,200 flue-cured ~300 air-cured	0.12-8.29	150

Results

Figure 2 shows the raw NIR spectra files acquired on the i-Spec® Plus imported into the BWIQ® software. In order to identify any outlier spectra in the calibration samples, a Leverage test was performed within the software. Identified outliers were removed from the calibration set before the models were developed.

As presented in Figure 2, there is some offset in the baseline of the spectra due to a scattering effect from the tobacco samples. In order to minimize these baseline offsets, a Multiplicative Scatter Correction (MSC) treatment was applied to the calibration sample set in the BWIQ® software. A second derivative pretreatment was also applied to the calibration set in the software in order to further minimize baseline and slope difference and enhance the signal variation in the calibration set. Figure 3 shows the processed NIR spectra after MSC and second derivative pretreatments.
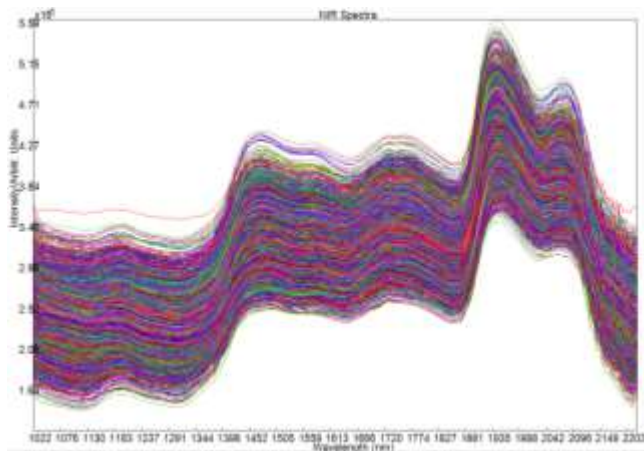

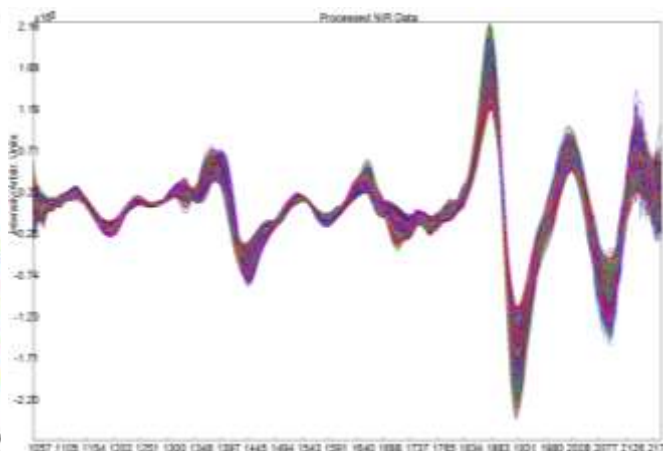Fig. 2 Tobacco NIR raw spectra collected on the i-Spec Plus system

Fig. 3 Processed NIR spectra (MSC and second derivative). Wavelength ranges were isolated to 1198-2050 nm for sugar and nicotine and 1800-2050 nm for chlorine.

Previous studies employing NIR for tobacco processing have demonstrated that wavelength selection is extremely important in building predictive models, as the selected wavelengths must directly correspond to the components under investigation.[1] Furthermore, although second derivative pre-processing enhances the variation in the data, it also amplifies noise present in the spectra. Applying a manual wavelength selection in BWIQ® allows the ability to only include specific wavelengths of the spectrum, and to eliminate the spectral region above 2050 nm where the noise has been enhanced. In this study, nicotine and sugar contents were modeled with a wavelength range of 1198-2050 nm. Chloride was modeled with a wavelength range of 1800-2050 nm.
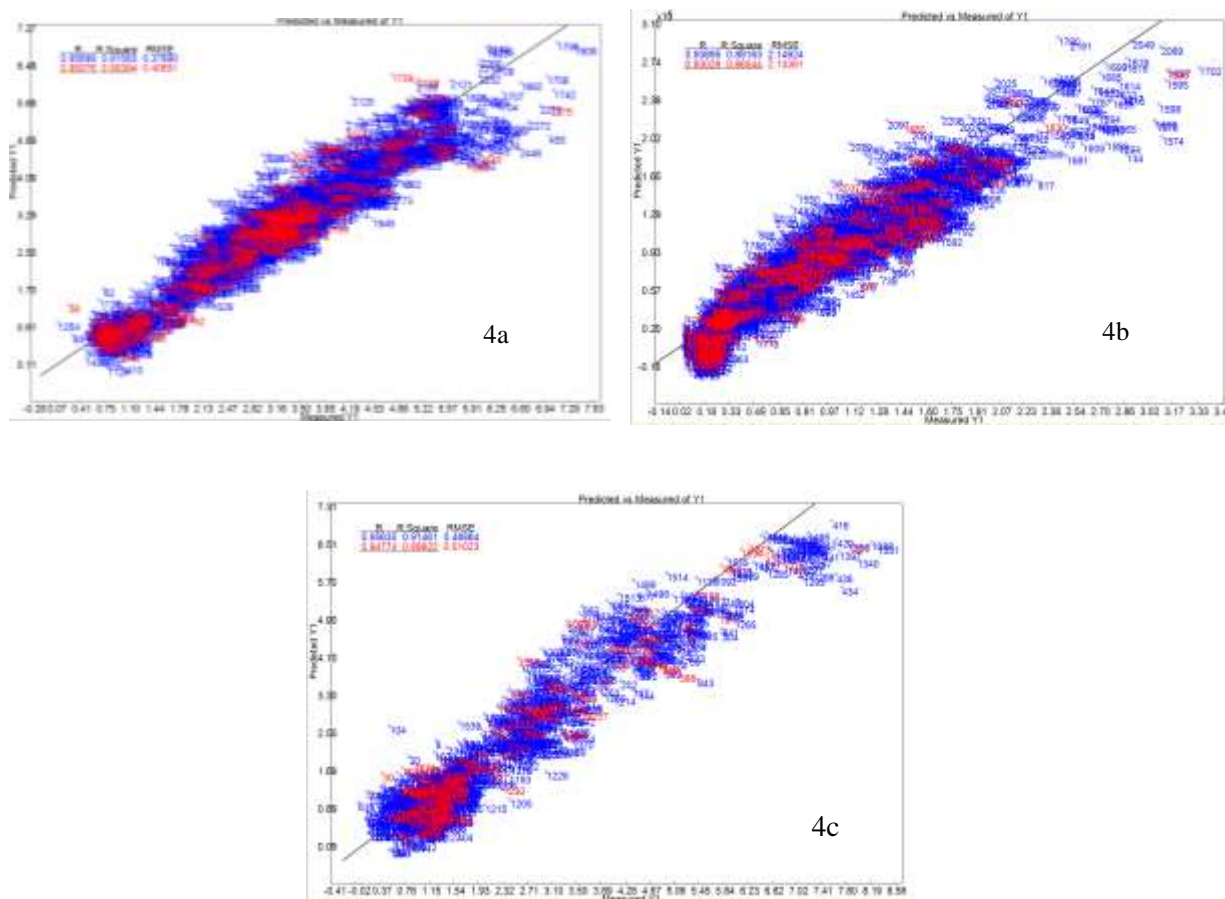

Fig. 4 (a) PLS calibration curve for nicotine (b) PLS calibration curve for sugar and (c) PLS calibration curve for chloride. Blue indicates calibration samples, red indicates validation samples.

[1] Diffee, J. T. In *Handbook of Near-Infrared Analysis*; Burns, D. A., Ciurczak, E. W., Eds.; Practical Spectroscopy Series; Marcel Dekker, Inc: New York, NY, 1992; Vol. 13, pp 433–473.

A partial least squares (PLS) regression algorithm was applied to all calibration sets for quantitative analysis. Figure 4 shows the generated calibration curves for (a) nicotine, (b) sugar and (c) chloride components. BWIQ software reports several chemometric parameters that are indicative of the performance of a calibration model. Pearson's R coefficient describes the linearity of the calibration curve, while the RMSE refers to the Root Mean Square Error of the calibration and validation. Table 2 presents the chemometric parameters that describe the quality of the created calibration curves for nicotine, sugar, and chloride in tobacco.

Table 2 Regression coefficients for nicotine, sugar, and chloride models

Component	Number of factors	Pearson's R coefficient	RMSEC	RMSEV
Nicotine	6	0.95699	0.37890	0.42200
Sugar	8	0.93895	2.14924	2.13391
Chloride	8	0.95635	0.48863	0.51023

Validation sets are used to show the quality of predictions using the created calibration curve. In order to visualize the difference in the prediction value versus the reference value of the validation samples, the residual distribution plots were analyzed. Figure 5 shows the residual plots for the nicotine, sugar, and chloride models. In models with good linear agreement, the samples should ideally fall close to the zero line in the residual plots (red line), and be evenly be distributed around it.
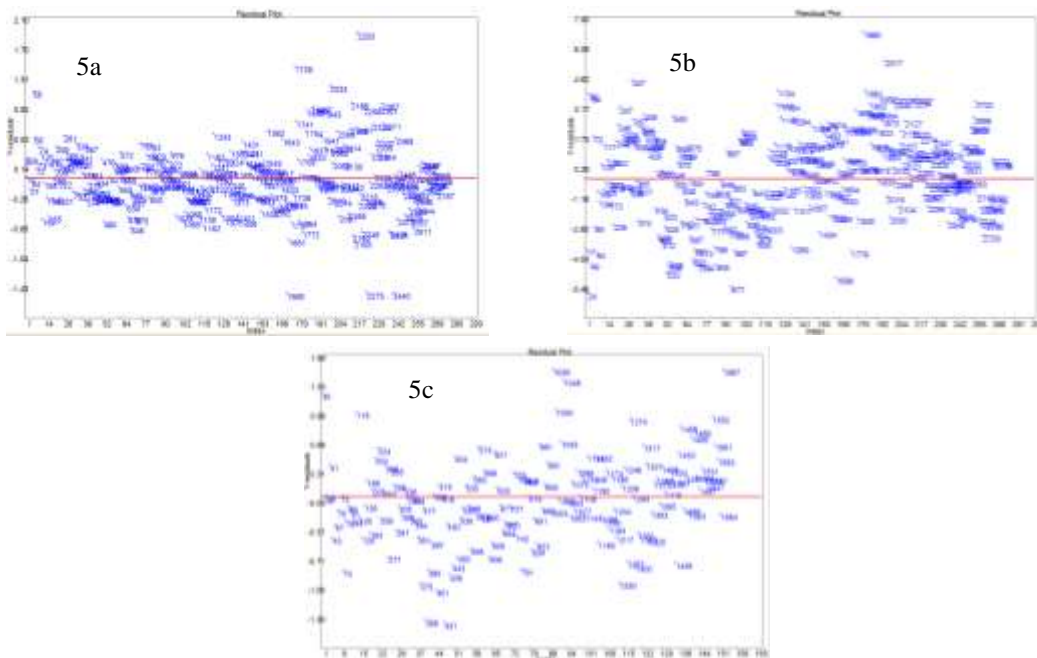

Fig. 5 Residual plots for (a) nicotine, (b) sugar, and (c) chloride models

Conclusion

B&W Tek's i-Spec® Plus portable NIR spectrometer can be used to rapidly and accurately determine the contents of nicotine, sugar, and chloride in ground tobacco using models developed with BWIQ® chemometric software. Once a chemometric model is created for the components of interest, it can be used to accurately predict the content within a tobacco sample, allowing for rapid, on-line quality control analysis. The ability to quickly analyze high volumes of tobacco samples using a NIR system eliminates the use of hazardous chemicals and waste generated with traditional wet chemistry techniques.