JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Estimating Hydrogen Cyanide in Forage Sorghum (*Sorghum bicolor*) by Near-Infrared Spectroscopy

Glen P. Fox,*[,†,‡] Natalie H. O'Donnell,[§] Peter N. Stewart,[∥,⊥] and Roslyn M. Gleadow[§]

[†]Centre for Nutrition and Food Science, Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Toowoomba, Queensland 4350, Australia

[‡]Queensland Department of Employment and Economic Development and Innovation, Lesley Research Centre, Toowoomba, Queensland 4350, Australia

[§]School of Biological Sciences, Monash University, Clayton, Victoria 3800, Australia

[∥]Pacific Seeds, Toowoomba, Queensland 4350, Australia

ABSTRACT: Hydrogen cyanide (HCN) is a toxic chemical that can potentially cause mild to severe reactions in animals when grazing forage sorghum. Developing technologies to monitor the level of HCN in the growing crop would benefit graziers, so that they can move cattle into paddocks with acceptable levels of HCN. In this study, we developed near-infrared spectroscopy (NIRS) calibrations to estimate HCN in forage sorghum and hay. The full spectral NIRS range (400−2498 nm) was used as well as specific spectral ranges within the full spectral range, i.e., visible (400−750 nm), shortwave (800−1100 nm) and near-infrared (NIR) (1100−2498 nm). Using the full spectrum approach and partial least-squares (PLS), the calibration produced a coefficient of determination (R^2) = 0.838 and standard error of cross-validation (SECV) = 0.040%, while the validation set had a R^2 = 0.824 with a low standard error of prediction (SEP = 0.047%). When using a multiple linear regression (MLR) approach, the best model (NIR spectra) produced a R^2 = 0.847 and standard error of calibration (SEC) = 0.050% and a R^2 = 0.829 and SEP = 0.057% for the validation set. The MLR models built from these spectral regions all used nine wavelengths. Two specific wavelengths 2034 and 2458 nm were of interest, with the former associated with C═O carbonyl stretch and the latter associated with C−N−C stretching. The most accurate PLS and MLR models produced a ratio of standard error of prediction to standard deviation of 3.4 and 3.0, respectively, suggesting that the calibrations could be used for screening breeding material. The results indicated that it should be feasible to develop calibrations using PLS or MLR models for a number of users, including breeding programs to screen for genotypes with low HCN, as well as graziers to monitor crop status to help with grazing efficiency.

KEYWORDS: *Dhurrin, feed quality, hydrogen cyanide, NIRS, wavelengths*

INTRODUCTION

Cyanogenesis is the process whereby plants release hydrogen cyanide (HCN) from *in situ* cyanide-containing compounds. While cyanogenic glycosides are nontoxic, in the presence of certain enzymes, these compounds are hydrolyzed to produce HCN, which is highly toxic. The development of HCN may play a role in plant defense against herbivores.[1] In sorghum, the cyanogenic glycoside dhurrin is synthesized from the amino acid tyrosine in a series of steps catalyzed by two P450s and a UGT-transferase. Dhurrin is broken down to HCN when it is mixed with specific β-glucosidases (dhurrinase).[2] Dhurrin and dhurrinase are spatially separated in the living plant, such that HCN is only released when the tissue is damaged, consistent with its putative role in herbivore defense. The amount of cyanide able to be released from dhurrin is known as the cyanide potential (HCNp).

Near-infrared spectroscopy (NIRS) is an analysis tool used routinely in agricultural sciences. Since its development in the 1950s, it has become the main work-horse for cereal-based plant breeding programs,[3] as well as finding applications in forage testing,[4] remote sensing on plants for growth,[5] and crop nutrition status.[6] Specifically, in cereal plant breeding applications, NIRS has been mainly used in crops, such as wheat or barley, where there is a quality specification on commercially delivered crops.[3] For grain sorghum, no target quality specifications exist, although the opportunity to predict feed traits exists using NIRS.[7] In forage sorghum, NIRS has been used to estimate characteristics, such as chemical composition and feed quality.[8] To date, there has been one published report describing the NIRS estimation of the antinutritional factor, namely, HCNp, in forage sorghum[9] using a partial least-squares method.

Recent calibration development strategies use a partial least-squares (PLS) regression approach, i.e., combining all spectral data, of up to 1050 wavelengths. This approach has been used successfully in a number of plant-based agricultural applications, particularly breeding.[3] However, it is possible is to use only a few specific wavelengths [multiple linear regression (MLR)] that are correlated to the trait of interest.[10] The early NIRS instruments used a limited number of specific filters (specific wavelengths) and for very few grain traits, such as moisture, protein, and lipid.[11]

The aim of this study was to determine the suitability of NIRS to estimate HCNp in forage sorghum. A NIRS calibration has been reported, and PLS and MLR models were compared to ascertain if one approach provided a better calibration than

Received: August 8, 2011
Revised: May 14, 2012
Accepted: May 17, 2012
Published: May 17, 2012

the other. The development of a NIRS calibration could be used in a number of applications, including breeding to screen lines, because NIRS is a high-throughput, low-cost technology, as well as on the farm for monitoring forage sorghum during the growing season.

MATERIALS AND METHODS

Sorghum Samples. Forage sorghum [*Sorghum bicolor* (L.) Moench × *Sorghum sudanense* Stapf. cv. Superdan] was tested. All samples were collected in 2008 from near Gatton in the Lockyer Valley, Queensland, Australia, except two from South Australia, Australia. A total of 12 additional samples were from an experimental Sorghum × Sudan line. To ensure that the samples represented a wide range of HCN, a further 24 samples were taken from hay made in the same season. Field samples of live plants were harvested, where the uppermost fully unfurled leaf was collected. The hay samples were all whole plant samples. Samples were dried at 70 °C and then ground in a Christie and Norris 200 mm hammer mill, with a 1 mm screen. Initially, 153 samples were combined and split into calibration (109) and validation (44) sets. A further 15 samples were used as an independent test set to test significant differences ($p < 0.05$) between calibration models.

Measurement of HCNp. Dhurrin in plant tissue was determined as described previously,[12] by hydrolyzing the cyanogenic glycoside and trapping the resultant HCN in a well containing 1 M NaOH. Initially, samples were ground using a Christy and Norris 200 mm cross-arm mill fitted with a 1 mm screen. The dried and ground samples were placed in plastic Ziploc bags and stored in a dehumidified cold room at 4 °C. Hydrolysis was achieved by adding 500 μL of β-glucosidase almond emulsion (0.01%, w/v) (β-D-glucoside glucohydrolase, EC 3.2.1.21, Sigma) in 0.1 M trisodium citrate−HCl buffer (pH 5.5) to approximately 10 mg of ground oven-dried leaf material in a sealed glass vial and incubating at 37 °C for 15 h. Cyanide in the NaOH well was neutralized with acetic acid and assayed. HCN was measured with a FLUOstar OPTIMA UV/vis absorbance spectrophotometer microplate reader (BMG LABTECH, Offenburg, Germany) at 595 nm. A total of 1 g of cyanide (CN) detected by this method is equivalent to 11.6 g of dhurrin and referred to as HCNp. All samples were tested in triplicate.

NIRS. Ground samples were scanned in a small ring cup in the NIRSystem 6500 instrument (Foss NIRSystems, Inc., Silver Spring, MD).

Table 1. Summary of HCN in Forage Sorghum Samples Used

	number of samples	range (%, dry basis)	mean (%, dry basis)	standard deviation
calibration	109	0.021−0.941	0.230	0.156
validation	44	0.036−0.521	0.236	0.129
test set	15	0.022−0.531	0.189	0.160

Table 2. Summary of the Best NIR Calibration Model Statistics

region	spectral range (nm)	math treatment[a]	number of factors[b]	R^2 (cal)	SECV (cal)[c]	R^2 (val)	SEP (val)	RPD (test)[d]
			PLS					
full	400−2498	2,4,4,1	5	0.838	0.044	0.824	0.047	3.4
visible	400−800	2,4,4,1	5	0.712	0.079	0.702	0.092	1.7
SWNIR	802−1098	2,4,4,1	9	0.813	0.056	0.798	0.069	2.3
NIR	1100−2498	1,4,4,1	9	0.826	0.049	0.804	0.051	3.1
			MLR					
full	400−2498	1,4,4,1	9	0.836	0.051	0.821	0.054	3.0
visible	400−800	2,4,4,1	3	0.525	0.086	0.467	0.121	1.3
SWNIR	802−1098	1,4,4,1	9	0.798	0.056	0.701	0.078	2.1
NIR	1100−2498	2,4,4,1	9	0.847	0.050	0.829	0.057	2.8

[a]All calibrations produced the best models with SNV detrend scatter correction. [b]Number of factors used in the PLS calibration and number of wavelengths used in the MLR calibration. [c]Standard error of cross-validation for the PLS calibration and standard error of calibration in the MLR calibration. [d]Ratio of SEP (test set) and standard deviation (SD) (test set).

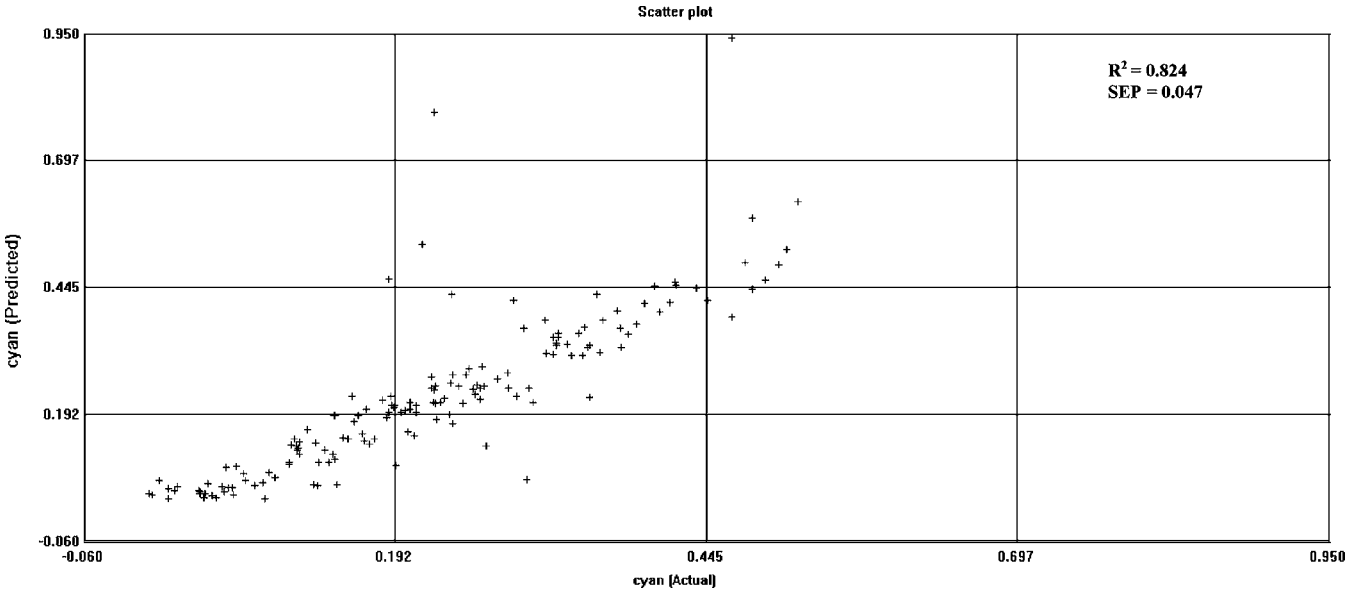

Figure 1. Scatter plot of actual versus predicted values using the best PLS calibration model.

Table 3. Wavelengths Selected in MLR Calibrations

region	spectral range (nm)	wavelength[a] (nm)	chemical assignment[b]
full	400−2498	776	N−H third overtone (OT) NH_2
		628	chlorophyll
		2068	NH_2
		2314	CH_2
		446	unassigned
		546	unassigned
		1704	oil
		2034	C═O second OT
		2316	CH_2
SWNIR	700−1098	736	O−H third OT OH
		888	C−H third OT CH_2
		948	O−H second OT OH
		1056	C−H CB CH_2
		1012	CONH second OT amide
		788	C≡CH third OT
		858	C−H third OT aromatic
		1000	N−H second OT NH_2
		864	C═C third OT
NIR spectra only	1100−2498	2394	no assignment
		1152	C−H second OT CH_3
		1836	C−H CB CH_2
		1398	C−H CB CH_2 CH stretch + CH bend
		1256	no assignment
		2430	no assignment
		2302	C−H bend second OT protein
		1228	C−H second OT CH
		2458	C−N−C stretch first OT protein

[a]Wavelengths are listed in order of selection for the MLR model. [b]Chemical assignments from the WinISI software.

The calibration software (WinISI V1.6, Foss NIRSystems, Inc., Silver Spring, MD) for the NIRS instrument was used for all calibration models using absorbance as a log(1/*R*) function. Two calibration methods were used in developing a calibration model for HCN. The first was a typical approach using the full spectrum (400−2498 nm at 2 nm intervals) in a modified PLS regression model. A number of scatter corrections were tested, including no treatment, multiplicative scatter correction, and standard normal variate detrend. Math pretreatments tested included none, first, and second derivatives. MLR models were also assessed with the above scatter corrections and premath treatment combinations using a stepwise regression with a maximum number of 10 terms (9 wavelengths) based on the WinISI software. For both calibration approaches, full spectra, visible spectra (400−750 nm), shortwave near-infrared (SWNIR) spectra (800−1100 nm), and near-infrared (NIR) spectra (1100−2500 nm) were tested. In addition, for all calibration models, the spectral outlier (global H) was set at 5.000 GH and cross-validation settings were used, i.e., four groups.

To calculate which model was most accurate from all models developed, the modified Fearn test was used.[13] This method tests the significance of the differences in standard deviation and bias between two models based on the residuals for the references and predicted values for each sample in the tested population. A ratio of standard error of prediction to standard deviation (RPD)[14] was also calculated to ascertain the potential application of the NIRS models for application in breeding or industry screening. RPD values of 1 or less are an indication of an inadequate model. Values greater than 2.5 indicate that the model could be useful for initial screening purposes, whereas a RPD greater than 5 indicates that the model would be good for quality control and prediction.

RESULTS

One of the main objectives in developing NIRS calibrations is to ensure that a suitable range of the trait of interest is sampled and the level of precision in the reference method is acceptable. Table 1 shows the descriptive statistics and number of samples for the calibration, validation, and test sets used in this study. Both sets included samples below and above the HCN range acceptable for cattle (0.6%). However, the average values for each set was below the 0.6% threshold, suggesting the range of values could be expanded to include more samples above the threshold.

PLS Approach. Two calibration methods were used in developing a calibration model for HCN. The first was a typical approach using the full spectrum [400−2498 nm (1050 data points)] with a modified PLS regression model. The best PLS model was using the NIR spectra (400−2500 nm), giving a coefficient of determination (R^2) = 0.837 and standard error of cross-validation (SECV) = 0.044%, using a standard normal variance (SNV) detrend pretreatment with the first derivative, 4 nm gap, and 4 nm smoothing (Table 2). Figure 1 shows the scatter plot for the actual versus predicted values from this model. The use of NIR and SWNIR spectra both produced reasonable calibrations (Table 2). The R^2 for the validation set using the full spectrum was 0.824, with a standard error of prediction (SEP) = 0.047%. The RPD was 3.4, indicating that the predictive model could be used for screening in a breeding program. All other PLS models produced $R^2 > 0.66$ and RPD above 1.6. As suggested by Williams and Sobbering,[14] these calibration statistics may be useful for screening in a breeding program but not for quality testing applications.

MLR Approach. The second calibration model was developing using a MLR approach, in this case a stepwise wavelength selection process. The NIR spectra produced the most accurate calibration with a 2,4,4,1 math treatment and SNV detrend scatter correction. The R^2 = 0.847, and a standard error of calibration = 0.050%. For the validation set, a R^2 = 0.829, with a SEP = 0.057%. The RPD for this model was 2.8, suggesting that it could also be used for screening in a breeding program. However, the full spectra and NIR spectra regions showed potential for providing useful predictions from a MLR calibration approach with R^2 greater than 0.79, SEPs less than 0.080%, and RPDs greater than 2.0.

Nine wavelengths were selected using the MLR model for all three spectral regions, including the NIR region (Table 3). The visible region produced a very poor calibration model (Table 2), with only four wavelengths selected (not shown). The majority of wavelengths selected were associated with CH bonding (methyl CH_2 or methylene CH_3) (Table 3). However, in the full spectral region and the SWNIR region calibrations, there was one or more wavelengths associated with NH bonding. These included 776, 2034, and 2068 nm in the full spectral region and 1000 and 1012 nm in the SWNIR region. In the "NIR spectra only" calibration, six of the nine wavelengths were associated with CH bonding, while for the remaining three, there was no chemical assignment (Table 3).

As seen in the loading plots (Figure 2), there was a strong influence from the visible region in the full spectrum model (Figure 2a). The first loading explained 60% of the full spectral calibration. However, when using the visible region only, a poor calibration was developed, where the first loading (Figure 2b) explained only 34% of the visible region calibration. The first

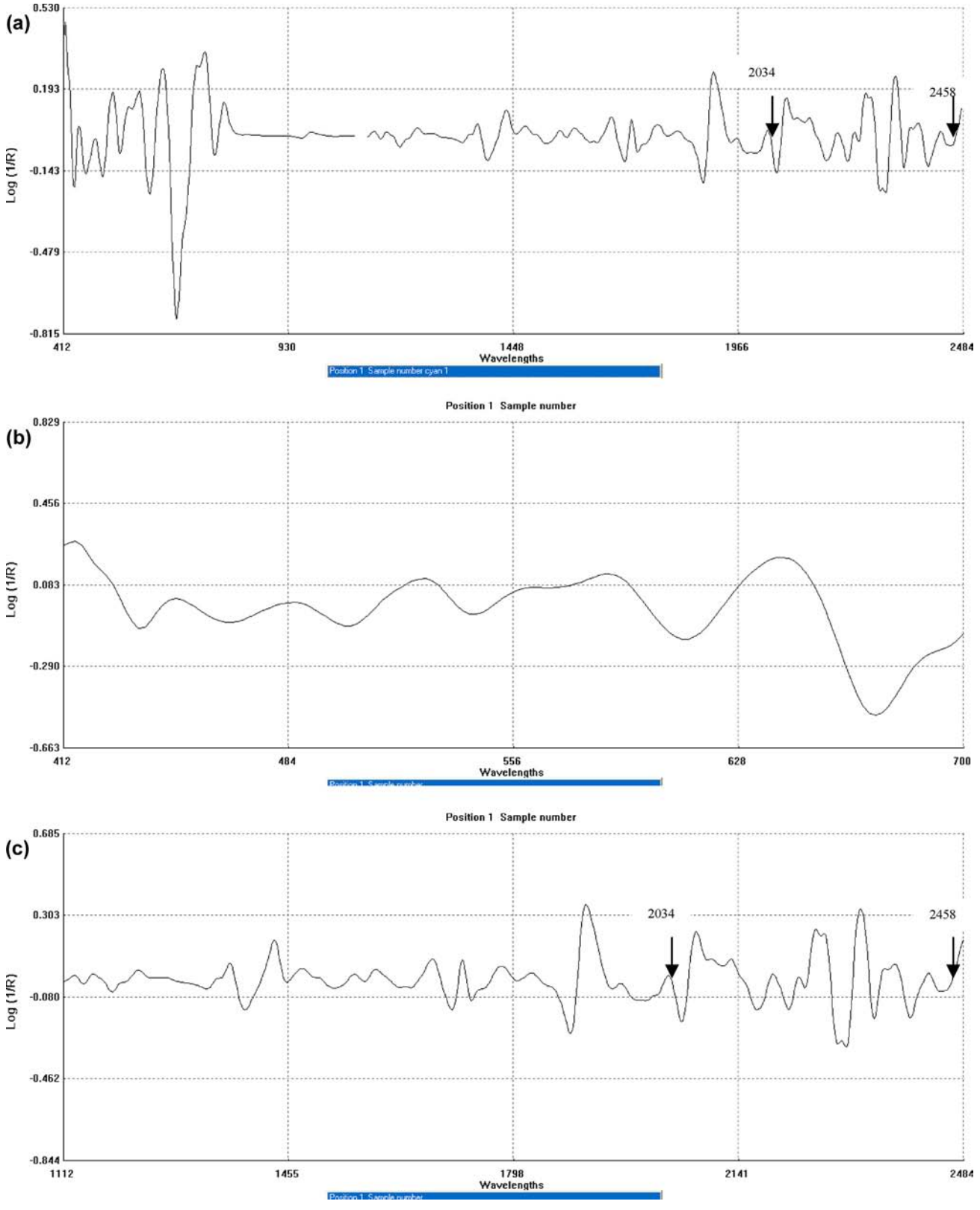

Figure 2. Loading line plot 1 from (a) full spectrum, (b) visible spectrum only, and (c) NIR spectrum only second derivative PLS model.

loading plot for the NIR region is shown in Figure 2c. This loading explained 63% of the NIR spectrum calibration.

While both PLS and MLR calibrations used wavelengths associated with nitrogen bonding, two wavelengths from MLR calibrations were of specific interest. In the full spectrum MLR calibration, the 2034 nm wavelength was selected (Table 3). This wavelength is associated with the C═O bond and assigned as a carbonyl stretch second overtone associated with amide I (urea) (as per the WinISI software). Because the chemical formula for urea is O═C−$(NH_2)_2$, it is possible that the covalent bonding

dx.doi.org/10.1021/jf205030b | J. Agric. Food Chem. 2012, 60, 6183−6187

between the C≡N atoms in HCN could result in a similar chemical stretch to C═O. In the NIR spectrum only calibration, the 2458 nm wavelength was selected (Table 3). This wavelength was assigned as a C−N−C stretch first OT protein in the WinISI software. These two specific regions have been highlighted in panels a and c of Figure 2 for the full spectra and NIR spectra PLS calibrations, respectively.

The modified Fearn test[13] was used to note any significant differences ($p < 0.05$) between calibrations when predicting the test set. The results indicated that there was no significant difference ($p < 0.05$) between the predicted HCN levels for the full spectral, "NIR spectra only", and "SWNIR spectra only" for both PLS and MLR calibrations. There was a significant difference ($p < 0.05$) between the visible spectral region and all of the other regions from both PLS and MLR calibrations. There was also no significant difference ($p < 0.05$) between the "best" PLS (full spectra) and MLR (full spectra) models.

DISCUSSION

In this study, we built NIR calibrations to predicted HCNp. Two calibration approaches were used, with these being the typical full-spectral calibration model and a selected wavelength calibration model. The resultant calibrations showed that both approaches were similar in predicting HCNp in forage sorghum samples. Our results, especially for the PLS model for the NIR spectral region only were comparable to the only other reported NIRS calibration for HCN,[9] in which that study used only the NIR spectral region only.

From our study, when the full, visible spectra, SWNIR spectra, and NIR spectra were compared, the best calibration in terms of R^2, SEP, and RPD values was using the full spectrum. This would suggest that using all wavelengths from the visible, SWNIR, and NIR regions was the best strategy for building a PLS calibration for HCN. The loading plots showed the positive and negative contributions of these regions to the PLS model. However, despite using the 1050 wavelengths available, it was possible to build a predictive calibration with a similar level of accuracy using only nine wavelengths. While for this study, the best MLR calibration was using the NIR spectra, the full spectrum calibration, which included visible spectrum wavelengths, was similar. This would suggest that, for this type of sample, including wavelength from the visible and NIR spectral regions would provide accurate calibrations, regardless of whether choosing MLR or PLS calibration strategies. The use of fewer wavelengths provides an opportunity for the development of specific instruments, which could provide fast estimation in the field-based instrument to predict HCNp levels or potentially using remote sensing technologies.

The approach of using fewer wavelengths was considered specifically for this small chemical molecule. It may not work for some traits, especially where there may be strong interactions between different chemical species in a complex sample matrix, for example, wheat dough, which has starch, protein, and lipid components. However, for chemicals such as HCN, with a simple chemical structure and, hence, associations with specific wavelengths, the MLR approach has merit. This has already been shown where both PLS and MLR models proved suitable.[10]

The possibility to use NIRS to predict HCN in forage sorghum breeding could help in selecting genotypes that inherently produce lower concentrations of dhurrin. Other potential applications include the monitoring of crops to gauge HCNp levels prior to feeding cattle as well as assessing crops to understand how the environment influences the expression of dhurrin. The results of this study have shown the potential to use NIR to predict the HCNp content in forage sorghum, which has good economic potential for graziers and plant breeders alike.

AUTHOR INFORMATION

Corresponding Author

*E-mail: g.fox1@uq.edu.au.

Present Address

[⊥]School of Agriculture and Food Science, The University of Queensland, Brisbane, Queensland 4072, Australia.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Gleadow, R. M.; Woodrow, I. E. Constraints on effectiveness of cyanogenic glycosides in herbivore defense. *J. Chem. Ecol.* **2002**, *28*, 1301−1313.

(2) Nielsen, K. A.; Tattersall, D. B.; Jones, P. K.; Møllera, B. L. Metabolon formation in dhurrin biosynthesis. *Phytochemistry* **2008**, *69*, 88−98.

(3) Osborne, B. G. Applications of near infrared spectroscopy in quality screening of early-generation material in cereal breeding programmes. *J. Near Infrared Spectrosc.* **2006**, *14*, 93−101.

(4) Shenk, J. S.; Westerhaus, M. O. Accuracy of NIRS instruments to analyze forage and grain. *Crop Sci.* **1985**, *25*, 1120−1122.

(5) Yang, C.; Everitt, J. H.; Bradford, J. M.; Escobar, D. E. Mapping grain sorghum growth and yield variations using airborne multispectral digital imagery. *Trans. ASAE* **2000**, *43*, 1927−1938.

(6) Gole, R. S.; Patil, V. D.; Shinde, P. V.; Chuare, J. S. Studies on spectral reflectance under normal and nitrogen−phosphorus stress condition in soybean (*Glycine max* L.) crop. *J. Soils Crops* **2008**, *18*, 317−322.

(7) van Barneveld, R. J.; Nuttall, J. D.; Flinn, P. C.; Osborne, B. G. Near infrared reflectance measurement of the digestible energy content of cereals for growing pigs. *J. Near Infrared Spectrosc.* **1999**, *7*, 1−7.

(8) Bruno-Soares, A. M.; Murray, I.; Paterson, R. M.; Abreu, J. M. F. Use of near infrared reflectance spectroscopy (NIRS) for the prediction of the chemical composition and nutritional attributes of green crop cereals. *Anim. Feed Sci. Technol.* **1998**, *75*, 15−25.

(9) Goff, B. M.; Moore, K. L.; Falesa, S. L.; Pedersen, J. F. Comparison of gas chromotography, spectrophotometry and near infrared spectroscopy to quantify prussic acid potential in forages. *J. Sci. Food Agric.* **2011**, *91*, 1523−1526.

(10) Fox, G. P.; Onley-Watson, K.; Osman, A. Multiple linear regression calibrations for barley and malt protein based on the spectra of hordein. *J. Inst. Brew.* **2002**, *108*, 155−159.

(11) Norris, K.; Hart, J. R. Direct spectrophotometric determination of moisture content of grain and seeds. In *Principles and Methods of Measuring Moisture Content in Liquids and Solids*; Wexler, A., Ruskin, R. E., Eds.; Reinhold: New York, 1965; Vol. *4*, pp 19−25.

(12) Blomstedt, C. K.; Gleadow, R. M.; O'Donnell, N.; Naur, P.; Jensen, K.; Laursen, T.; Olsen, C. E.; Stuart, P.; Hamill, J. D.; Møller, B. L.; Neale, A. D. A combined biochemical screen and TILLING approach identifies mutations in *Sorghum bicolor* L. (Moench) resulting in acyanogenic forage production. *Plant Biotechnol. J.* **2012**, *10*, 54−66.

(13) Reid, D.; Guthrie, J. Comparing two competing calibrations for a given data set using Fearn's criteria: Version 2.0. *NIR Spectrum* **2004**, *1*, 7.

(14) Williams, P. C.; Sobbering, D. C. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spectrosc.* **1993**, *1*, 25−32.