SciencePG
Science Publishing Group

Effective Quantification of Tannin Content in Sorghum Grains Using Near-infrared Spectroscopy

Yongsheng Wang[1, 2, *], Jie Li[1, 2], Bo Wang[1], Yuting Zhang[2], Junling Geng[2], Li Xin Wen[3], Aike Li[4]

[1]Animal Nutrition Center of Nutrition & Health Research Institute, China Oil & Foodstuffs Corporation, Beijing, China

[2]Beijing Engineering Research Center of Livestock Products Quality and Safety Source Control, Beijing, China

[3]College of Veterinary Medicine, Hunan Agricultural University, Changsha, China

[4]Institute of Grain Quality and Nutrition, Academy of National Food and Strategic Reserves Administration, Beijing, China

Email address:

wangyongsheng@cofco.com (Yongsheng Wang)

[*]Corresponding author

Abstract: This study was conducted to investigate the feasibility of determining tannin content in sorghum grains with near-infrared reflectance spectroscopy (NIRS). A total of 110 sorghum grain samples were collected. The data matrix of the pretreated NIRS was randomly divided into a calibration set (Nc=77 samples) and a prediction set (Np=33 samples). The analysis of tannin content was based on the colorimetric method of GBT 15686-2008. Diffuse reflectance spectra of 110 sorghum samples were generated on a Fourier-transform NIRS with a scanning range of 12800-4000 cm^{-1} and resolution of 16 cm^{-1} and 64 scans. Several spectra pretreatment methods were compared to for an optimum spectral pretreatment method. The optimal model was determined according to coefficient of determination for calibration (R^2_{CAL}), root mean standard error of calibration ($RMSE_{CAL}$), coefficient of determination for cross-validation (R^2_{CV}), root mean standard error of cross-validation ($RMSE_{CV}$) and the residual predictive deviation (RPD). The results showed that the tannin content of the sorghum grains ranged from 0.01% to 2.12% DM with the average of 0.58%, and first derivative was the optimal spectral pretreatment with the lowest $RMSE_{CV}$ of 0.14. The absorption peaks of the optimal model mainly located at 9402-7492 cm^{-1} and 5452-4244 cm^{-1}. The RPD of calibration, cross-validation and external validation were 6.22, 4.22 and 3.0, respectively. The findings suggest that the established model using NIRS is effective to quantify tannin content in sorghum grains rapidly.

Keywords: Tannin, Near-infrared Spectroscopy, Sorghum

1. Introduction

Sorghum is one of the leading cereal crops worldwide and ranks the fifth highest production of the cereal crops, following maize, wheat, rice, and barley, with 57.6 million tons of annual production globally in 2017 [1]. Sorghum is well known for its outstanding agronomic performance with great adaptability to a variety of environments and has been widely cultivated in tropical and subtropical regions for food/feed production or industrial feedstock [2]. Sorghum grain is high in starch content and is a rich source of nutrients and bioactive compounds such as phenolic acids, flavonoids, and tannins [2, 3]. Because of the superior agronomic

attributes, nutrient value and health potential, sorghum has attracted increasing attention from academia and industries in the past decades. As well, it has been used as feedstuff in livestock production with the advantages of high energy value and relatively low price [4-6]. However, the intrinsic tannin in sorghum grains acts as a double-edged sword in animal nutrition, i.e., a proper dose of dietary tannin would exert a positive effect on animal health and production, whereas exceeding dose may cause certain deleterious effects such as reduced feed intake, nutrient digestibility, growth performance and poor gastrointestinal health, acting as an antinutritional factor [7]. Thus, to optimize the utilization of sorghum in animal production, there is a need for a precise

measurement of tannin content which largely depends on variety, agronomic and climate conditions [8].

Tannin content has been determined by standard wet-chemical methods such as Chlorox bleach test, vanillin-HCl colorimetric assay and normal-phase HPLC with fluorescence detection, which are time and effort consuming, costly and destructive [8]. Notably, due to the exploding sample size and the need of on-line decision making in modern production, a faster and easier method for tannin determination is imperative. Of the alternatives, spectroscopic techniques might provide a faster, non-invasive measurement based on the interaction of light and chemical groups. For example, near-infrared reflectance spectroscopy (NIRS) is based on the unique near-infrared absorption properties of different chemical groups in organic compounds [9, 10]. It has been used as a rapid and cost-effective method for the qualitative and quantitative analyses in various industries [11, 12]. Incorporation of NIRS and chemometric techniques has given birth of numerous models for rapid analysis of chemical composition (protein, fiber, oil, minerals, etc.) or structure (texture) of a variety of materials, such as soil, forage, fruits [9, 13]. Dykes et al. [14] reported the possibility of using NIRS to rapidly and non-destructively predict the concentration of condensed tannins in whole grain sorghum. Factors like weather condition and nutrient supply also affect the tannin profile in sorghum grain, so more research is needed to prove the feasibility of NIRS in quantifying tannin content.

The present study compared eight spectral pretreatment methods and established a calibration model for quantitative analysis of tannin in sorghum grain using NIRS method. The model would facilitate the on-line decision making in sorghum production.

2. Materials and Methods

2.1. Sampling and Pretreatment

A total of 110 sorghum grain samples collected from China, USA and Brazil were ground (1.0 mm sieve) with a pulverizer (Retsch ZM200, Germany). The powdered samples were kept in sealed plastic bags (85 × 60 mm) and stored in a dry, dark place. To evaluate the performance of the calibration models, the data matrix of the pretreated NIR spectra (N=110 samples) was randomly divided into a calibration set (Nc=77 samples) and a prediction set (Np=33 samples), with the tannin values in the range of the calibration set.

2.2. Analysis of Tannin Content in Wet Chemical Method

The analysis of tannin content was based on the colorimetric method of GBT 15686-2008 [15]. In brief, sample (1 g) extracted with 20 mL of 75% dimethylformamide solution was centrifuged, and the supernatant was mixed with ammonia solution (8 g/L) and ammonium ferric citrate (3.5 g/L). The absorption was read on 525 nm and Tannic acid (Merck 773) was used as a standard. Samples were measured in triplicate.

2.3. Spectral Scanning

A Bruker Tang-R near infrared spectrometer (Bruker, Karlsruhe, Germany) was used to acquire NIR spectra in reflectance mode. All the samples were successively subjected to spectral scanning in the sample cup (diameter: 75 mm). Spectra were scanned at 12800-4000 cm^{-1} with a resolution of 4 cm^{-1} and 64 scans per sample. Each sample was measured four times and the mean was used for multivariate analysis.

2.4. Calibration Model Establishment and Validation

Spectral data were analyzed using OPUS 7.5 software. Potential outliers were checked by applying principal component analysis and Hotelling's T-squared statistic. Eight mathematical spectral pre-processing methods were compared, including Vector normalization, minimum-maximum normalization, first derivative, second derivative, multivariate scattering correction, first derivative + minus a straight line, first derivative + vector normalization and first derivative + multivariate scattering correction. In order to avoid over-fitting, partial least squares (PLS) method combined with full cross-validation was used to develop the calibration model. The optimal model was determined according to coefficient of determination for calibration (R^2_{CAL}), root mean standard error of calibration (RMSE$_{CAL}$), coefficient of determination for cross-validation (R^2_{CV}), root mean standard error of cross-validation (RMSE$_{CV}$) and the residual predictive deviation (RPD), which was calculated by the standard deviation (SD) of the measured composition divided by RMSE$_{CV}$, i.e., RPD=SD/RMSE$_{CV}$.

Besides the internal cross-validation, external validation was further conducted. The best models obtained for the calibration set were tested using 33 samples of the prediction set randomly selected from 110 samples. The prediction power was evaluated based on coefficient of determination for validation (R^2_{VAL}) and root mean standard error of prediction (RMSEP). All the procedures were performed according to the method outlined by Zhang et al [16].

3. Results and Discussion

3.1. Analyzed Value of Tannin Content in Sorghum Grain

As illustrated in Figure 1, tannin content of the sorghum grains used in the present study ranged from 0.01% to 2.12% DM with the mean of 0.58% DM (SD=0.06). The distribution of tannin contents of the samples exhibited a normal pattern with the most frequent zone of 0.28%-0.35% DM accounting for 40.17% of the sample population, followed by the zones of 0.35%-0.43% DM and 0.21%-0.28% DM. The results of Zhang et al. [16] investigating the tannin content of 90 sorghum germplasm resources showed that the range of tannin content was 0.05%-1.35% DM with an average of 0.58% DM. Similarly, the study of Mu et al. [17] revealed that the average tannin content of 17 sorghum varieties was 0.72% DM with a distribution range of 0.08%-1.26%. These findings suggested that our samples cover a full range of

tannin content commonly reported in varieties and germplasms of sorghum, implying a wide adaptability of the established model.
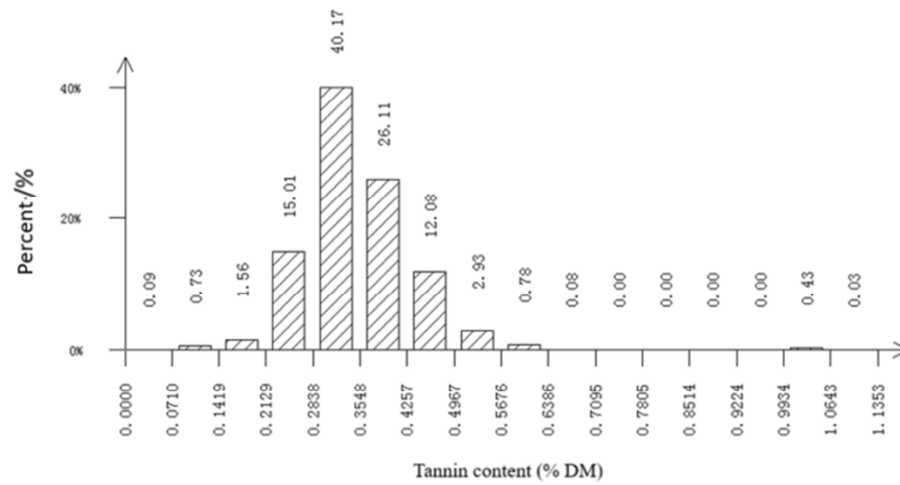

Figure 1. The distribution of tannin content in sorghum grain samples.
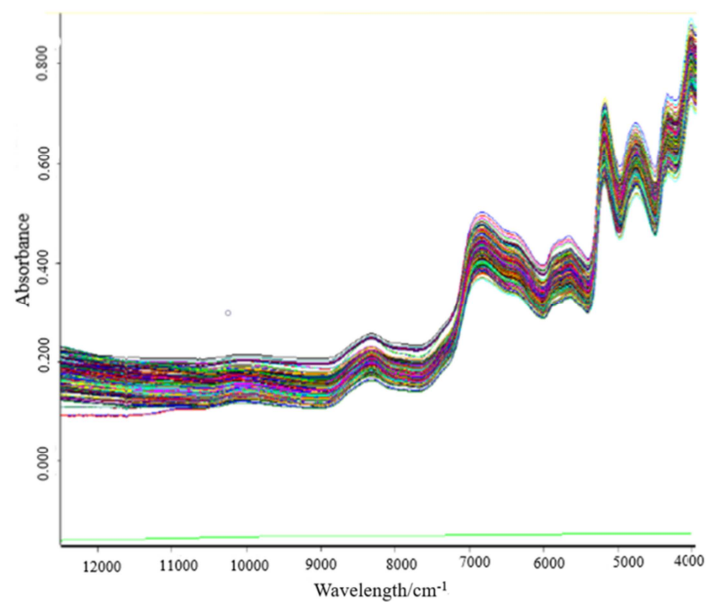
3.2. Interpretation of NIR Spectra

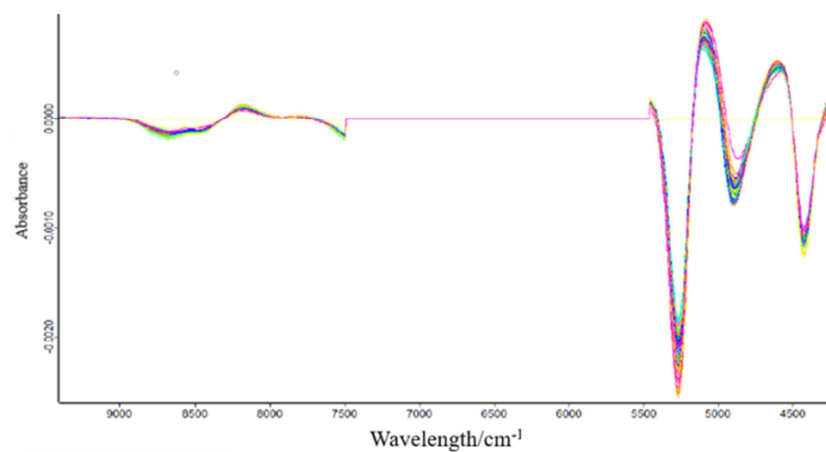*Figure 2. Raw near-infrared reflectance spectra of sorghum grain.*

Figure 3. First derivative pretreated NIR spectra of sorghum grain.

The original NIR spectra of the 110 sorghum grain samples were similar (Figure 2). In general, the absorption values of the NIR spectra of sorghum grain remained around 0.200 in the wavenumber of 12800-7500 cm^{-1} and then progressively enhanced in the range of 7500-4000 cm^{-1} with several absorption peaks high up to 0.800 of absorption. These spectra are generally known to consist of many overlapping narrow bands of different vibrational modes for various of functional groups. Thus, the spectra need to be pretreated for establishing a model. For example, first derivative is useful in resolving overlapping bands and minimizing the effect of particle size. Savitzky-Golay smoothing can wipe out high frequency information and enhance the ratio of signal to noise. Standard normal variate can reduce spectral and particle size variability [9, 18]. In the present study, eight pretreatments including vector normalization, minimum-maximum normalization, first derivative, second derivative, multivariate scattering correction, first derivative + minus a straight line, first derivative + vector normalization and first derivative + multivariate scattering correction, were applied to process the spectra. Compared with Figure 1, the spectra appeared more informative and distinctive (Figure 3 and Figure 4). Hence, these processing enhanced the resolution and improved NIRS prediction power owing to the reduction of multi-collinearity and baseline shift. Relatively, first derivative was the best spectrum pretreatment method with a lowest $RMSE_{CV}$ of 0.14. Thus, first derivative pretreated spectra were further used to build up calibration model, with apparent absorption peaks locating in 9402-7492 cm^{-1} and 5452-4244 cm^{-1}.
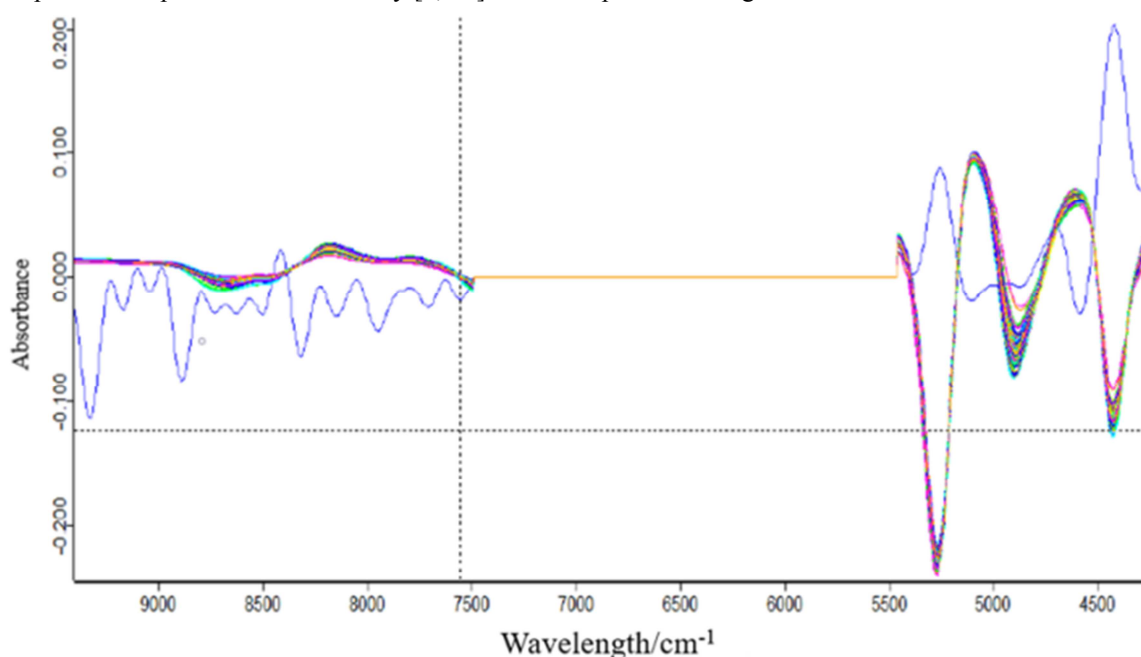

Figure 4. First derivative + vector normalization pretreated NIR spectra of sorghum grain.

3.3. Establishment of Calibration Model and Its Evaluation

The performance of the calibration models was evaluated using coefficient of determination for calibration (R^2_{CAL}), root mean standard error of calibration ($RMSE_{CAL}$), coefficient of determination for cross-validation (R^2_{CV}), root mean standard error of cross-validation ($RMSE_{CV}$) and the residual predictive deviation (RPD). The performance of our calibration model with the best spectrum pretreatment method (first derivative) was represented in Table 1, Figure 5 and Figure 6. The R^2_{CV} of calibration, cross-validation and external validation was 97.50, 94.37 and 88.86 along with the $RMSE_{CV}$ of 0.09, 0.13 and 0.18, respectively. Generally, an optimal model is indicated by high R^2 and RDP, low $RMSE_{CAL}$ and $RMSE_{CV}$ values. In fact, NIRS is very difficult to attain the regression relationship and repeatability as well as those of the wet chemical measurement ($R^2 > 0.99$), especially when the model is adaptable to a wide range of samples with different characteristics influenced by various factors. Practically, the thumb rule is that RPD values > 10 are considered equivalent effectiveness to the reference method, whereas RPD values < 2.5 indicate that NIRS is not suitable for the quantitative determination of the components [9]. For the present study, the RPD values of the best model were ≥ 3.0 (Table 1), implying that the model is effective for the analytical purposes. The performance of the model is owed to the wide range of tannin content and the optimization of spectrum pretreatment as mentioned above. Taken together, a rapid determination of tannin content in sorghum grain using NIRS is feasible and effective.

Table 1. Optimization results for calibration model of tannin content in sorghum grain.

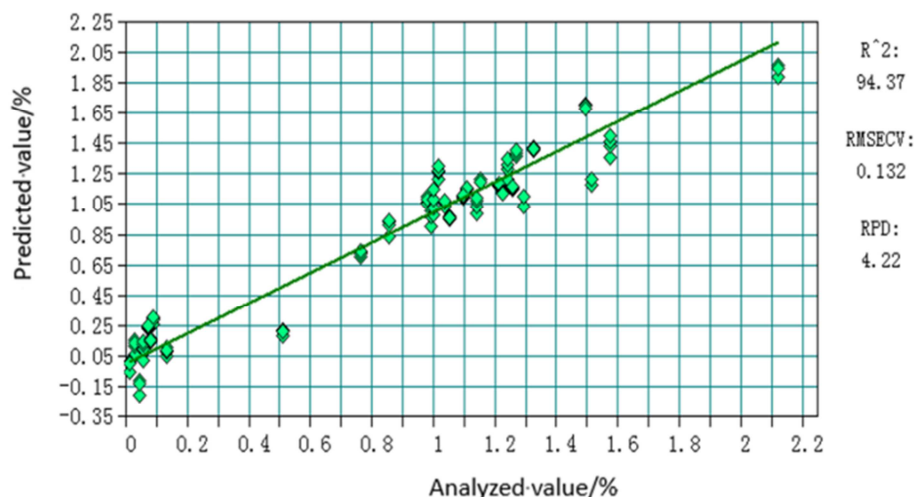| Item	R^2	RMSE	RPD
Calibration	97.50	0.09	6.33
Cross-validation	94.37	0.13	4.22
External validation	88.86	0.18	3.00

Figure 5. *NIRS predicted value VS analyzed value by calibration and cross-validation statistics.*
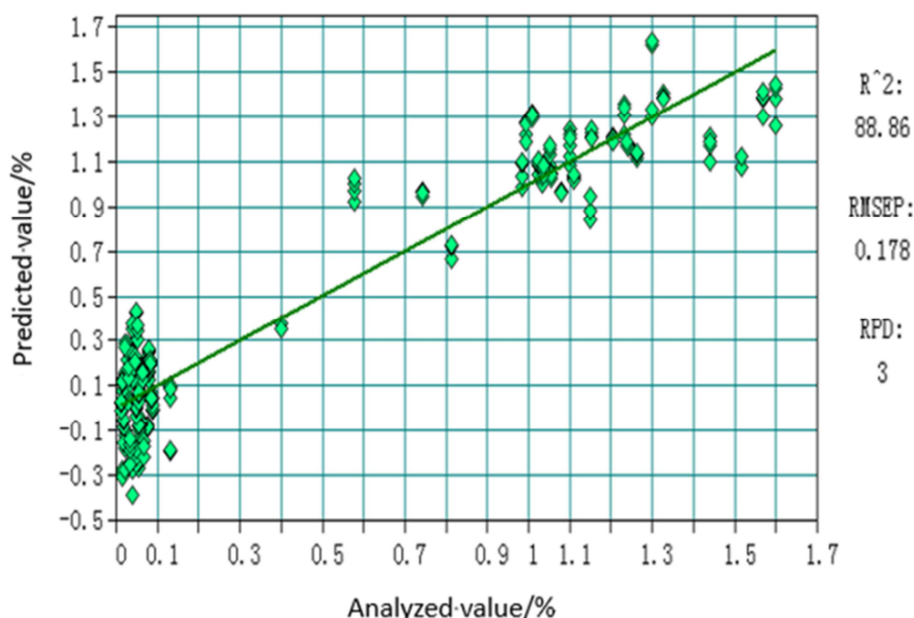

Figure 6. *NIRS predicted value VS analyzed value by external validation statistics.*

4. Conclusions

Chemical analysis showed that the tannin content of the 110 sorghum samples ranged 0.01%-2.12% DM, covering almost of the reported tannin values in sorghum germplasms. The comparison of various spectra pretreatment methods revealed that first derivative was the optimal method for processing sorghum spectra as the absorption peaks of the optimal model mainly located in 9402-7492 cm^{-1} and 5452-4244 cm^{-1}. The residual predictive deviation (RPD) values of calibration, cross-validation and external validation were 6.22, 4.22 and 3.0, respectively, higher than the threshold (RPD > 2.5) for the admission of quantitative analysis. Therefore, the established model using NIRS is effective to conduct rapid determination of tannin content in sorghum grains, and our model would facilitate the on-line decision making of sorghum quality.

References

[1] FAOSTAT. http://www.fao.org/faostat/en/#data/QC. Last access on Oct. 20, 2019.

[2] Xiong Y, Zhang P, Warner R D, et al. Sorghum grain: From genotype, nutrition, and phenolic profile to its health benefits and food applications. Compre. Rev. Food Sci. F. 2019; 34: 1541-1550.

[3] Shen S, Rui H, Charlie L, et al. Phenolic compositions and antioxidant activities differ significantly among sorghum grains with different applications. Molecules 2018; 23: 1203-1214.

[4] Zhao J F, Hu G, Tang L. Effects of supplementation with compound enzymes and probiotics on sorghum diets on growth performanc, serum antioxidant indexes and intestinal morphology of *Liangfenghua* broilers. Chinese J. Anim. Nutr. 2018; 30: 2318-2327.

[5] Paulk C B, Hancock J D, Fahrenholz A C, et al. Effects of sorghum particle size on milling characteristics and growth performance in finishing pigs. Anim. Feed Sci. Tech. 2015; 202: 75-80.

[6] Wang T, Hou M, Shang J. Characteristics of muscle fibers and relative meat quality of fattening sheep fed sweet sorghum silage and corn silage diets. Pratacultural Sci. 2018; 35: 2722-2727.

[7] Ronda V A C V. Sorghum for Animal Feed. In: Breeding Sorghum for Diverse End Uses, UK: Woodhead Publishing, 2019; pp. 229-238.

[8] Dykes L. Tannin Analysis in Sorghum Grains: Methods and Protocols. Sorghum, 2019.

[9] Zhang J, Li S, Lin M, et al. A near-infrared reflectance spectroscopic method for the direct analysis of several fodder-related chemical components in drumstick (*Moringa oleifera Lam.*) leaves. Biosci. Biotech. Biochem. 2018; 82: 1-8.

[10] Rungpichayapichet P, Mahayothee B, Nagle M, et al. Robust NIRS models for non-destructive prediction of postharvest fruit ripeness and quality in mango. Postharvest Biol. Tech. 2016; 111: 31-40.

[11] Mohamed E S, Saleh A M, Belal A B, et al. Application of near-infrared reflectance for quantitative assessment of soil properties. Egypt J. Remote Sensing Space Sci. 2018; S1168123289.

[12] De Girolamo A, Cervellieri S, Cortese M, et al. Fourier transform near-infrared and mid-infrared spectroscopy as efficient tools for rapid screening of deoxynivalenol contamination in wheat bran. J. Sci. Food Agri. 2019; 99: 1946-1953.

[13] Jaconi A, Vos C D. Near infrared spectroscopy as an easy and precise method to estimate soil texture. Geoderma. 2019; 337: 906-913.

[14] Dykes L, Hoffmann Jr. L, Portillo-Ridriguez O, et al. Prediction of total phenols, condensed tannins, and 3-deoxyanthocyanidins in sorghum grain using near-infrared (NIR) spectroscopy. J. Cereal Sci. 2014; doi: 10.1016/j.jcs.2014.02.002.

[15] GB. Sorghum-Determination of tannin content GB/T 15686-2008. General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China; National standardization management Council, Chinese Standard Publishing. 2008.

[16] Zhang Y, Zhou F, Zhang X, et al. Cluster analysis and determination of protein and tannin content in sorghum germplasms. Tianjin Agri. Sci. 2017; 54: 65-74.

[17] Mu Z, Tian X, Shi Y. Determination method of tannin content in sorghum. J. Shanxi Agri. Sci. 2014; 23: 42-46.

[18] Towett E K, Alex M, Shepherd K D, et al. Applicability of near-nfrared reflectance spectroscopy (NIRS) for determination of crude protein content in cowpea (*Vigna unguiculata*) leaves. Food Sci. Nutr. 2013; 1: 45-53.