Research Article

Combining Near-Infrared Spectroscopy and Chemometrics for Rapid Recognition of an Hg-Contaminated Plant

Bang-Cheng Tang,¹ Hai-Yan Fu,² Qiao-Bo Yin,² Zeng-Yan Zhou,¹ Wei Shi,¹ Lu Xu,^{1,2} and Yuan-Bin She³

¹Research Institute of Applied Chemistry, College of Material and Chemical Engineering, Tongren University, Tongren, Guizhou 554300, China

²The Modernization Engineering Technology Research Center of Ethnic Minority Medicine of Hubei Province, College of Pharmacy, South-Central University for Nationalities, Wuhan 430074, China

³College of Chemical Engineering, Zhejiang University of Technology, Hangzhou 310014, China

Correspondence should be addressed to Wei Shi; wzslwdbd@163.com, Lu Xu; lxchemo@163.com, and Yuan-Bin She; sheyb@zjut.edu.cn

Received 10 March 2016; Revised 29 April 2016; Accepted 17 May 2016

Academic Editor: Paulo R. G. Hein

Copyright © 2016 Bang-Cheng Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The feasibility of rapid recognition of an Hg-contaminated plant as a soil pollution indicator was investigated using near-infrared spectroscopy (NIRS) and chemometrics. The stem and leave of a native plant, *Miscanthus floridulus* (Labill.) Warb. (MFLW), were collected from Hg-contaminated areas ($n_1 = 125$) as well as from regular areas ($n_2 = 116$). The samples were dried and crushed and the powders were sieved through an 80-mesh sieve. Reference analysis of Hg levels was performed using inductively coupled plasma-atomic emission spectrometry (ICP-AES). The actual Hg contents of contaminated and normal samples were 16.2–30.5 and 0.0–0.1 mg/Kg, respectively. The NIRS measurements of impacted sample powders were collected in the mode of reflectance. The DUPLEX algorithm was utilized to split the NIRS data into representative training and test sets. Different spectral preprocessing methods were performed to remove the unwanted and noncomposition-correlated spectral variations. Classification models were developed using partial least squares discrimination analysis (PLSDA) based on the raw, smoothed, second-order derivative (D2), and standard normal variate (SNV) data, respectively. The prediction accuracy obtained by PLSDA with each data preprocessing option was 100%, indicating pattern recognition of Hg-contaminated MFLW samples using NIRS data was in perfect consistence with the ICP-AES results. NIRS combined with chemometrics will provide a tool to screen the Hg-contaminated MFLW, which can be potentially used as an indicator of soil pollution.

1. Introduction

The growing development of agricultural, industrial, and urban activities has largely increased the release of toxic substances such as heavy metals and organic compounds to environmental systems [1–3]. In particular, toxic heavy metals in air, water, soil, and plants have caused severe public environmental concern because of their severe adverse influences on human health [4–6]. It is well known that soil is a major sink for heavy metal pollutants, which can be accumulated and transferred to water, air, plants, and animals. It was estimated that 20% of the total farmland in China had been contaminated, which directly threatens the safety of food production [7].

Numerous research efforts have been devoted to evaluation of the level of soil contamination with heavy metals caused by human activities, including electroplating industry, mining, smelting, coal-fired power stations, steel and iron manufacturing, waste incineration, leather industry, and cement production [8–12]. Most of these researches focused on direct determination of heavy metal levels in the soil. Various analytical methods have been developed and used to quantify the levels of heavy metals in soil, plants, and animals [13, 14], including inductively coupled plasma-atomic emission spectroscopy (ICP-AES), inductively coupled mass spectroscopy (ICP-MS), atomic fluorescence spectrometer (AFS), X-ray fluorescence spectrometer (XRF), neutron

activation analysis (NAA), DC argon plasma multielement atomic emission spectrometer (DCP-MAES), atomic absorption spectrometer (AAS), and scanning electron microscopy with energy dispersive X-ray (SEM-EDX). Although accurate evaluation of heavy metals can be obtained, most of these techniques generally require laborious preconcentration of analytes and sample pretreatment, which have made the analysis time-consuming.

It is well known that excessive adsorption and accumulation of certain pollutants can influence the growth and metabolism of native plants [15, 16]. A traditional method to recognize and evaluate soil pollution is by examining the morphological variations of plant indicators caused by soil pollution, which are sensitive to the presence of certain pollutants. Although the level of soil pollution could only be qualitatively evaluated using plant indicators, it is more convenient and economic compared with direct methods by analysis of pollutants in soil. However, the use of plant indicators for soil pollution can be limited for some reasons. Firstly, because the plant species in an area can be influenced by many factors, such as geographical and climatic conditions, usually it is not ready to have a well-studied and suitable plant indicator in certain areas. Secondly, in some seriously polluted areas, the plants sensitive to soil pollution would perish and be gradually replaced with the dominant species, which have adapted to the pollution and whose morphological changes are not significant enough to be exactly recognized by the naked eye. Therefore, rather than examining the plants by the naked eye, it is more reasonable and reliable to characterize the changes in chemical compositions of polluted plants using instrumental techniques.

Near-infrared spectroscopy (NIRS) has been widely applied to analysis of various food and agricultural products [17–22]. The feasibility of using NIRS for quantitative analysis of heavy metals in environmental samples has been extensively evaluated [23–26]. Although in some cases NIRS demonstrates potential for quantitative analysis of heavy metals, in many cases, the sensitivity is lower than by using other methods and time-consuming sample preparation is required to obtain reliable results. Some studies also indicate NIRS is very useful for qualitative analysis of heavy metals [27]. NIRS can provide a powerful tool to simultaneously characterize the multicomponents in a complex system, which could be combined with pattern recognition methods [28, 29] to perform rapid classifications of different types of samples.

Therefore, the objective of this paper was to investigate the feasibility of rapid recognition of a native Hg-contaminated plant *Miscanthus floridulus* (Labill.) Warb. (MFLW) from normal MFLW using NIRS and chemometrics. Special attention was made on the experimental design and data collection to avoid obtaining artifacts caused by factors other than Hg-contamination.

2. Materials and Methods

2.1. Collection and Preparation of MFLW Samples. MFLW samples were collected with leaves cut off from the upper end with a length of 25 cm. The Hg-contaminated MFLW samples ($n_1 = 125$) were collected around a mercury mining factory

FIGURE 1: The preparation of MFLW samples.

in Huashi, Tongren, China, within a range of 3 Kilometers; normal samples ($n_2 = 116$) were collected from an area about 10 Kilometers away (Chuandong, Tongren, China). All the MFLW samples were cleaned with water and kept in a cool, dry, and ventilated place away from direct sunlight to remove the moisture. Each sample (leaves and stalks) was crushed by a disintegrator and then the powders were sieved through an 80-mesh sieve. The dried, crushed, and filtered samples were kept with integrate packaging. An ultraviolet lamp was used to dry each sample for 10 minutes before NIR analysis and Hg reference analysis. The flowchart of sample preparation is shown in Figure 1.

2.2. NIRS Measurements. Impacted MFLW powders were analyzed in a quartz sample cup using an Antaris II Fourier transform-NIR spectrometer (Thermo Electron Co., Waltham, Massachusetts, USA) using the RESTLT 3.0 software in the reflectance mode. The spectra were measured using a PbS detector with an internal gold background as the reference. The working range of spectrometer was $4000\text{--}10000\text{ cm}^{-1}$. Each sample was measured triply while being stirred and impacted before each measurement and the average spectra were obtained. The number of scans for each measurement was 32. The instrumental resolution was 8 cm^{-1} with a scanning interval of 3.857 cm^{-1} , so each raw spectrum had 1557 wavelengths. The temperature was kept at around 25°C and the humidity was kept at a stable level during analysis. In order to avoid artificial spectral variations between different types of samples, the order of analysis for all the samples was permuted randomly.

TABLE 1: The programmed microwave-assisted digestion conditions for ICP-AES analysis of MFLW samples.

Step	Power (W)	Temperature (°C)	Heating time (min)	Holding time (min)
1	800	80	30	5
2	1280	120	30	7
3	1600	160	30	5

2.3. Reference Analysis of Hg Using Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES). The total Hg contents in MFLW were analyzed according to the national standard (GB5009.17-2014). The MFLW powders were digested using the CEM Mars 5 Microwave Accelerated Reaction System (CEM Corp., Matthews, USA). About 0.4 g of homogenized samples was digested in Teflon vessels with 8.0 mL of nitric acid (HNO₃) (V/V, 10%) overnight and kept at 150°C for 5 h. The programmed digestion conditions are summarized in Table 1. Hg contents were analyzed using an Agilent 725 ICP-AES system (Agilent, Victoria, Australia). The precision of ICP-AES analysis was verified by triplication of the samples. Pearson's r of the standard curve was over 0.9999. The average relative standard deviation (RSD) was less than 5.0% and the recovery rate was 96.1~104.5%. The limit of detection (LOD) was calculated to be 0.0025 mg/Kg according to the IUPAC method, where the signal of 3σ of 11 blank solutions was calibrated using the standard curve.

2.4. Chemometrics Analysis. The data analysis was performed on MATLAB 7.0.1 (Mathworks, Sherborn, MA). In order to remove the unwanted variation in NIRS data, smoothing [30], taking second-order derivative (D2) [30], and standard normal variate (SNV) [31] were performed on the raw data. The DUPLEX algorithm [32] was used to divide the measured samples into a representative training set and test set.

Partial least squares discriminant analysis (PLSDA) [33] was used to develop classification models to distinguish the Hg-contaminated from the regular samples. For PLSDA, a dummy response vector was constructed using +1 and -1 to represent the regular and Hg-contaminated samples, respectively. The number of PLSDA components was estimated using Monte Carlo cross validation (MCCV) [34]. The number of PLS components was determined as to obtain the lowest error rate of MCCV (ERMCCV):

$$\text{ERMCCV} = \frac{\sum_{i=1}^B m_i}{\sum_{i=1}^B N_i}, \quad (1)$$

where B is the number of MCCV data splitting and m_i and N_i are the numbers of misclassified and leave-out samples, respectively. For prediction, a cutoff value of zero was used to assign a new sample to one of the two classes.

For prediction, the overall accuracy (ACCU) was computed to evaluate the performance of classification models:

$$\text{ACCU} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}, \quad (2)$$

where TP, TN, FN, and FP represent the numbers of true positives, true negatives, false negatives, and false positives, respectively. In this work, regular and Hg-contaminated MFLW samples were seen as “positives” and “negatives,” respectively. Another two usually used indices, sensitivity (SENS) and specificity (SPEC), were also adopted to evaluate the classification performance:

$$\begin{aligned} \text{SENS} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{SPEC} &= \frac{\text{TN}}{\text{TN} + \text{FP}}. \end{aligned} \quad (3)$$

SENS and SPEC describe the model ability to correctly accept the “positives” and to correctly reject the “negatives,” respectively.

3. Results and Discussion

According to the analytical results of ICP-AES, the Hg contents of regular and contaminated MFLW objects ranged from 0.0 to 0.1 mg/Kg and 16.2 to 30.5 mg/Kg, respectively, indicating an obvious Hg-contamination of soil surrounding the mercury mining areas. The NIR spectra of regular and Hg-contaminated MFLW samples are shown in Figure 2. Seen from Figure 2, the raw spectra of regular and Hg-contaminated MFLW samples have very similar absorbance peaks in the range of 4000–10000 cm⁻¹. The peaks can be mainly assigned as follows [35]: 8377 cm⁻¹ (the second overtones of C–H stretching), 6823 cm⁻¹ (overlapping of the first overtone of O–H stretching and N–H stretching), 5662 cm⁻¹ (the first overtones of C–H stretching), 5184 cm⁻¹ (the combination of the baseband of O–H stretching and the first overtone of C–O deformation), and 4748 cm⁻¹ (combination of N–H stretching and deformation of peptide groups). Some bands (8377 cm⁻¹, 5662 cm⁻¹, and 4748 cm⁻¹) are very weak and the peak resolution is very low. Figure 2 also demonstrates the NIRS data preprocessed by smoothing and taking D2 and SNV transformation. Even with data preprocessing, the spectral difference between regular and Hg-contaminated MFLW samples is still very subtle and is difficult to be distinguished by the naked eye. Therefore, it is necessary to develop chemometric models to extract the relevant information for classification of regular and Hg-contaminated MFLW samples.

In order to obtain representative data sets for developing and validating classification models, the DUPLEX algorithm was adopted to divide the collected samples into training and prediction objects. Considering the regular and Hg-contaminated MFLW samples have different distributions, the DUPLEX algorithm was performed separately on the two classes. The 116 regular samples were split into 80 training and 36 test samples; the 125 Hg-contaminated objects were split into 85 training and 40 test samples. For model building, the training and test samples from the two classes were combined to form the final training and test sets, so 165 (80+85) training samples and 76 (36+40) test samples were obtained.

FIGURE 2: Some of the raw, smoothed, D2, and SNV spectra of regular and Hg-contaminated MFLW samples. An artificial shift was included to distinguish the spectra of regular and Hg-contaminated samples.

FIGURE 3: The predicted responses by PLSDA with different data preprocessing; test samples 1–36 are regular MFLW samples (positives) and 37–76 are Hg-contaminated (negatives).

PLSDA models were developed with the raw and preprocessed spectra. With different numbers of PLSDA components, ERMCCV was computed and the model complexity was determined as to minimize the ERMCCV value. The number of MCCV data splitting was set to be 100 in this work. Considering the size of training set is moderate, in each MCCV data splitting, 30% of the training set was randomly left out for prediction and the other 70% training samples were used for model development. Based on different data preprocessing options, the model parameters and prediction performance are shown in Table 2. It can be seen that, with each data preprocessing option and even without data preprocessing, PLSDA could obtain perfect classification of regular and Hg-contaminated samples and the accuracy, sensitivity, and specificity were all 1, indicating data preprocessing was not necessary to develop an accurate model. Moreover, all the PLSDA models had 2 latent variables

and the low model complexity means that the models would provide good generalization performance. The prediction results by PLSDA with different data preprocessing are shown in Figure 3, indicating distinct classification of regular and Hg-contaminated MFLW samples by PLSDA despite the kind of data preprocessing. By examining and comparing the predicted responses by PLSDA models with different preprocessing methods, the results by PLSDA with raw data and smoothed spectra were very similar, which were obviously different from those obtained by PLSDA with D2 and SNV spectra. Moreover, the prediction errors (with references to the dummy response vector of +1 and -1) obtained by PLSDA with D2 and SNV spectra were much lower than those obtained by PLSDA with the raw and smoothed spectra. Although all the four PLSDA models could achieve a classification accuracy of 1, D2 and SNV were still necessary to remove some unwanted spectral variations

TABLE 2: Discrimination of regular and Hg-contaminated MFLW by PLSDA of NIRS data.

Data	LV	ERMCCV	Accuracy	SENS	SPEC
Raw data	2	0	1	1	1
Smoothing	2	0	1	1	1
D2	2	0	1	1	1
SNV	2	0	1	1	1

to ensure the generalization performance of PLSDA when predicting new samples.

4. Conclusions

The feasibility of using NIRS for rapid classification of regular and Hg-contaminated MFLW samples was investigated. Classification accuracies of 1 were obtained with low model complexity despite the option of data preprocessing. D2 and SNV were demonstrated to be useful to improve training accuracy by removing unwanted spectral variations. Rapid recognition of the Hg levels in the native plant MFLW would provide a useful alternative indicator of Hg-contaminated soil, which can be used for rapid and economic screening of Hg-contamination. Our future research would be focused on investigating the feasibility of other plants as soil indicators as well as on developing the relationship between the levels of heavy metals in plants and soil.

Ethical Approval

This paper does not involve any animal or human experiments.

Competing Interests

Bang-Cheng Tang, Hai-Yan Fu, Qiao-Bo Yin, Zeng-Yan Zhou, Wei Shi, Lu Xu, and Yuan-Bin She declare no competing interests.

Authors' Contributions

Hai-Yan Fu and Bang-Cheng Tang equally contributed to this work.

Acknowledgments

Bang-Cheng Tang is grateful to the financial support from the Research Projects of Guizhou Science and Technology (no. QKHJZLKT[2012]15). Lu Xu is financially supported by the Open Research Program (no. GCTKF2014007) of State Key Laboratory Breeding Base of Green Chemistry Synthesis Technology (Zhejiang University of Technology), the Research Fund for the Doctoral Program of Tongren University (no. trxyDH1501), the Open Research Program (no. 2015ZY006) from the Modernization Engineering Technology Research Center of Ethnic Minority Medicine of Hubei Province (South-Central University for Nationalities), and the research funds from the Education Department of

Guizhou Province (no. QJHKYZ[2015]498). Zeng-Yan Zhou is financially supported by Science and Technology Department of Guizhou Province (no. QKHLHZ[2015]7245).

References

- [1] B. J. B. Nyarko, S. B. Dampare, Y. Serfor-Armah, S. Osae, D. Adotey, and D. Adomako, "Biomonitoring in the forest zone of Ghana: the primary results obtained using neutron activation analysis and lichens," *International Journal of Environment and Pollution*, vol. 32, no. 4, pp. 467–476, 2008.
- [2] J. Aznar-Márquez and J. R. Ruiz-Tamarit, "Environmental pollution, sustained growth, and sufficient conditions for sustainable development," *Economic Modelling*, vol. 54, pp. 439–449, 2016.
- [3] R. van Stigt, P. P. J. Driessen, and T. J. M. Spit, "Steering urban environmental quality in a multi-level governance context. How can devolution be the solution to pollution?" *Land Use Policy*, vol. 50, pp. 268–276, 2016.
- [4] J. A. Acosta, A. Faz, S. Martínez-Martínez, R. Zornoza, D. M. Carmona, and S. Kabas, "Multivariate statistical and GIS-based approach to evaluate heavy metals behavior in mine sites for future reclamation," *Journal of Geochemical Exploration*, vol. 109, no. 1–3, pp. 8–17, 2011.
- [5] Y. Ma, P. Egodawatta, J. McGree, A. Liu, and A. Goonetilleke, "Human health risk assessment of heavy metals in urban stormwater," *Science of the Total Environment*, vol. 557–558, pp. 764–772, 2016.
- [6] Y.-G. Gu, Y.-P. Gao, and Q. Lin, "Contamination, bioaccessibility and human health risk of heavy metals in exposed-lawn soils from 28 urban parks in southern China's largest city, Guangzhou," *Applied Geochemistry*, vol. 67, pp. 52–58, 2016.
- [7] C. Y. Wei and T. B. Chen, "Hyperaccumulators and phytoremediation of heavy metal contaminated soil: a review of studies in China and abroad," *Acta Ecologica Sinica*, vol. 21, no. 7, pp. 1196–1203, 2001.
- [8] G. S. Senesi, M. Dell'Aglio, R. Gaudiuso et al., "Heavy metal concentrations in soils as determined by laser-induced breakdown spectroscopy (LIBS), with special emphasis on chromium," *Environmental Research*, vol. 109, no. 4, pp. 413–420, 2009.
- [9] X. Zeng, X. Xu, H. M. Boezen, and X. Huo, "Children with health impairments by heavy metals in an e-waste recycling area," *Chemosphere*, vol. 148, pp. 408–415, 2016.
- [10] T. Sarkar, M. M. Alam, N. Parvin et al., "Assessment of heavy metals contamination and human health risk in shrimp collected from different farms and rivers at Khulna-Satkhira region, Bangladesh," *Toxicology Reports*, vol. 3, pp. 346–350, 2016.
- [11] C. Gutiérrez, C. Fernández, M. Escuer et al., "Effect of soil properties, heavy metals and emerging contaminants in the soil nematodes diversity," *Environmental Pollution*, vol. 213, pp. 184–194, 2016.
- [12] R. A. Wuana and F. E. Okieimen, "Heavy metals in contaminated soils: a review of sources, chemistry, risks and best available strategies for remediation," *ISRN Ecology*, vol. 2011, Article ID 402647, 20 pages, 2011.
- [13] R. K. Soodan, Y. B. Pakade, A. Nagpal, and J. K. Katnoria, "Analytical techniques for estimation of heavy metals in soil ecosystem: a tabulated review," *Talanta*, vol. 125, no. 11, pp. 405–410, 2014.

- [14] Y. N. Vodyanitskii and I. O. Plekhanova, "Biogeochemistry of heavy metals in contaminated excessively moistened soils (Analytical review)," *Eurasian Soil Science*, vol. 47, no. 3, pp. 153–161, 2014.
- [15] A. R. A. Usman, S. S. Lee, Y. M. Awad, K. J. Lim, J. E. Yang, and Y. S. Ok, "Soil pollution assessment and identification of hyperaccumulating plants in chromated copper arsenate (CCA) contaminated sites, Korea," *Chemosphere*, vol. 87, no. 8, pp. 872–878, 2012.
- [16] W. Meng, Z. Wang, B. Hu, Z. Wang, H. Li, and R. C. Goodman, "Heavy metals in soil and plants after long-term sewage irrigation at Tianjin China: a case study assessment," *Agricultural Water Management*, vol. 171, pp. 153–161, 2016.
- [17] A. Alishahi, H. Farahmand, N. Prieto, and D. Cozzolino, "Identification of transgenic foods using NIR spectroscopy: a review," *Spectrochimica Acta A: Molecular and Biomolecular Spectroscopy*, vol. 75, no. 1, pp. 1–7, 2010.
- [18] L. E. Agelet, P. R. Armstrong, J. G. Tallada, and C. R. Hurburgh Jr., "Differences between conventional and glyphosate tolerant soybeans and moisture effect in their discrimination by near infrared spectroscopy," *Food Chemistry*, vol. 141, no. 3, pp. 1895–1901, 2013.
- [19] P. M. Santos, E. R. Pereira-Filho, and L. E. Rodriguez-Saona, "Rapid detection and quantification of milk adulteration using infrared microspectroscopy and chemometrics analysis," *Food Chemistry*, vol. 138, no. 1, pp. 19–24, 2013.
- [20] J. U. Porep, D. R. Kammerer, and R. Carle, "On-line application of near infrared (NIR) spectroscopy in food production," *Trends in Food Science & Technology*, vol. 46, no. 2, pp. 211–230, 2015.
- [21] M. Schmutzler and C. W. Huck, "Simultaneous detection of total antioxidant capacity and total soluble solids content by Fourier transform near-infrared (FT-NIR) spectroscopy: a quick and sensitive method for on-site analyses of apples," *Food Control*, vol. 66, pp. 27–37, 2016.
- [22] S. D. Afandi, Y. Herdiyeni, L. B. Prasetyo, W. Hasbi, K. Arai, and H. Okumura, "Nitrogen content estimation of rice crop based on near infrared (NIR) reflectance using artificial neural network (ANN)," *Procedia Environmental Sciences*, vol. 33, pp. 63–69, 2016.
- [23] T. Kemper and S. Sommer, "Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy," *Environmental Science & Technology*, vol. 36, no. 12, pp. 2742–2747, 2002.
- [24] D. Cozzolino, "Near infrared spectroscopy as a tool to monitor contaminants in soil, sediments and water—state of the art, advantages and pitfalls," *Trends in Environmental Analytical Chemistry*, vol. 9, pp. 1–7, 2016.
- [25] L. Galvez-Sola, J. Morales, A. M. Mayoral et al., "Estimation of parameters in sewage sludge by near-infrared reflectance spectroscopy (NIRS) using several regression tools," *Talanta*, vol. 110, pp. 81–88, 2013.
- [26] D. F. Malley, "Near-infrared spectroscopy as a potential method for routine sediment analysis to improve rapidity and efficiency," *Water Science and Technology*, vol. 37, no. 6-7, pp. 181–188, 1998.
- [27] C. Palmborg and A. Nordgren, "Partitioning the variation of microbial measurements in forest soils into heavy metal and substrate quality dependent parts by use of near infrared spectroscopy and multivariate statistics," *Soil Biology and Biochemistry*, vol. 28, no. 6, pp. 711–720, 1996.
- [28] R. G. Brereton, "Pattern recognition in chemometrics," *Chemometrics and Intelligent Laboratory Systems*, vol. 149, pp. 90–96, 2015.
- [29] P. K. Hopke, "Chemometrics applied to environmental systems," *Chemometrics and Intelligent Laboratory Systems*, vol. 149, pp. 205–214, 2015.
- [30] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.
- [31] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Applied Spectroscopy*, vol. 43, no. 5, pp. 772–777, 1989.
- [32] R. D. Snee, "Validation of regression models: methods and examples," *Technometrics*, vol. 19, no. 4, pp. 415–428, 1977.
- [33] M. Barker and W. Rayens, "Partial least squares for discrimination," *Journal of Chemometrics*, vol. 17, no. 3, pp. 166–173, 2003.
- [34] L. Xu, J.-H. Jiang, Y.-P. Zhou, H.-L. Wu, G.-L. Shen, and R.-Q. Yu, "MCCV stacked regression for model combination and fast spectral interval selection in multivariate calibration," *Chemometrics and Intelligent Laboratory Systems*, vol. 87, no. 2, pp. 226–230, 2007.
- [35] B. M. Nicolai, K. Beullens, E. Bobelyn et al., "Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review," *Postharvest Biology and Technology*, vol. 46, no. 2, pp. 99–118, 2007.