

Article

Characterization of Wheat Varieties Using Terahertz Time-Domain Spectroscopy

Hongyi Ge ^{1,2,*}, Yuying Jiang ^{1,2}, Feiyu Lian ³, Yuan Zhang ³ and Shanhong Xia ^{1,2}

¹ State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China; E-Mails: jiangyuying11@163.com (Y.J.); shxia@mail.ie.ac.cn (S.X.)

² University of the Chinese Academy of Sciences, Beijing 100080, China

³ Key Laboratory of Grain Information Processing & Control, Ministry of Education, Zhengzhou 450001, China; E-Mails: lfywork@163.com (F.L.); zhangyuan@haut.edu.cn (Y.Z.)

* Author to whom correspondence should be addressed; E-Mail: gehongyi2004@163.com; Tel.: +86-10-8268-1741.

Academic Editor: W. Rudolf Seitz

Received: 19 April 2015 / Accepted: 21 May 2015 / Published: 27 May 2015

Abstract: Terahertz (THz) spectroscopy and multivariate data analysis were explored to discriminate eight wheat varieties. The absorption spectra were measured using THz time-domain spectroscopy from 0.2 to 2.0 THz. Using partial least squares (PLS), a regression model for discriminating wheat varieties was developed. The coefficient of correlation in cross validation (R) and root-mean-square error of cross validation (RMSECV) were 0.985 and 1.162, respectively. In addition, interval PLS was applied to optimize the models by selecting the most appropriate regions in the spectra, improving the prediction accuracy ($R = 0.992$ and RMSECV = 0.967). Results demonstrate that THz spectroscopy combined with multivariate analysis can provide rapid, nondestructive discrimination of wheat varieties.

Keywords: terahertz time-domain spectroscopy; wheat varieties; absorption spectrum; interval partial least squares

1. Introduction

Wheat is one of the most important agricultural products in the world. Due to increasing free trade, wheat varieties from diverse origins are widely available in the markets. However, the nutrition and processing quality of wheat varieties differ, and false wheat seeds can cause great losses for farmers. Thus, there is a demand for rapid analytical methods to discriminate the material properties of wheat. Traditional methods such as morphology analysis, physics, chemical methods, machine vision, and DNA molecular marker analysis enable sensitive classification of wheat varieties [1–4]. However, these methods are time-consuming and require complex operation. Spectroscopic methods such as near infrared, mid-infrared, and Raman spectroscopy are adequate analytical tools that have been used largely for material classification [5–7]. However, few studies have utilized the far infrared or terahertz (THz) band for discriminating wheat varieties.

Terahertz radiation, which occupies frequencies between 0.1 and 10 THz, lies between microwaves and infrared bands in the electromagnetic spectrum. Due to the rapid advances in generating, detecting, and analyzing THz radiation, THz spectroscopy has been recently developed as an analytical method [8,9]. Because many molecules have unique spectral fingerprints in THz band as a result of the vibrational transitions of the molecules, THz spectroscopy can be used for discriminating materials. Furthermore, THz radiation is non-destructive and can penetrate many nonpolar materials. As a result, THz spectroscopy has been successfully applied in detecting explosives [10] and drugs [11], in the field of biological sciences [12], and in food safety control [13].

The aim of the present study is to investigate the potential of THz spectroscopy as a non-destructive method to discriminate wheat varieties. The THz spectra of wheat varieties were measured and analyzed in the frequency of 0.2–2 THz. In addition, chemometric methods were used to evaluate wheat varieties based on THz spectra. The partial least squares (PLS) and interval PLS (iPLS) methods were used to obtain better discrimination results.

2. Materials and Methods

2.1. Experimental Setup

A conventional terahertz transmission spectroscopy system was used in the experiment. The mode-locked Ti-sapphire femtosecond laser, which provided 100-fs pulses at a wavelength of 800 nm and a repeating frequency of 80 MHz, was divided into two beams (pump beam and probe beam) using a polarization beam splitter (PBS). The THz pulses were generated from the low-temperature-grown GaAs photoconductive antenna with an attached silicon hyperhemispherical lens. The THz radiation from the emitter was collected and focused on the sample by a pair of parabolic mirrors (PM). Electro-optic (EO) detection was employed to observe the THz signal. The transmitted THz radiation was focused and collimated by PM onto the ZnTe EO detector crystal. A detailed description of THz-TDS can be found in Ref. [14]. The THz beam path was filled with nitrogen gas to remove absorption of atmospheric water vapor [15]. The samples were placed at the focal point of the THz beam spectroscopy, and the measurements were performed at an ambient temperature of 294 K with a relative humidity of approximately 3%.

Using THz-TDS, we can measure both the phase and amplitude of the THz pulses propagating through the sample and reference (nitrogen gas). A reference pulse signal, $E_{ref}(t)$, in the absence of wheat and a sample pulse signal, $E_{sam}(t)$, are recorded. Comparing the sample pulse and reference pulse using a fast Fourier transform, the complex refractive index $N(\omega)$ can be expressed as follows:

$$N(\omega) = n(\omega) - ik(\omega) \quad (1)$$

where $n(\omega)$ and $k(\omega)$ are the real refractive index and extinction coefficient, respectively, describing the dispersion and absorption characteristics of the sample. Here, ω is the cyclic frequency, and i is the imaginary unit. The complex transmittance function $H(\omega)$ of sample is given by [16–18]:

$$H(\omega) = \frac{E_{sam}(\omega)}{E_{ref}(\omega)} = \frac{4N}{(N+1)^2} \exp\left(\frac{i\omega(N-1)d}{c}\right) = \rho(\omega) \exp(i\phi(\omega)) \quad (2)$$

where $E_{sam}(\omega)$ and $E_{ref}(\omega)$ are the complex amplitudes of the Fourier transform of $E_{ref}(t)$ and $E_{sam}(t)$, respectively, c is the speed of light, and $\rho(\omega)$ and $\phi(\omega)$ are the amplitude ratio and related phase difference of the reference and sample, respectively.

The $n(\omega)$ and absorption coefficient $a(\omega)$ were obtained by:

$$n(\omega) = \frac{\phi(\omega)}{\omega d} c + 1 \quad (3)$$

$$a(\omega) = \frac{2}{d} \ln\left(\frac{4n(\omega)}{\rho(\omega)[n(\omega)+1]^2}\right) \quad (4)$$

2.2. Sample Preparation

In this study, eight wheat varieties were prepared for the analysis. Samples were supplied by Henan University of Technology, Zhengzhou, China, and were harvested in 2013. The collection of samples is as diverse as possible and representative of the main production areas. These wheat grains are mixtures with different components and complex structures and can have different chemical and physical properties. The sample properties are shown in Table 1.

Table 1. The Properties of the eight wheat varieties under consideration.

No.	Wheat Variety	Bulk Density (g/L)	Crude Protein Content (%)	Water Content (%)	Imperfect Grain (%)	Gluten Content (%)
1	Zhengmai 9023	756	14.6	12.5	3.2	27.9
2	Zhouyuan 9369	790	14.9	12.5	3.4	33.0
3	Aobiao	845	14.0	11.5	1.0	26.0
4	DNS	840	14.5	11.9	1.6	38.0
5	Jiamai	830	13.8	12.2	1.8	39.0
6	Jinan 17 wheat	773	15.6	11.5	3.0	34.0
7	Zhoumai 27	798	13.2	12.1	3.8	33.0
8	Yunong 416	787	14.3	12.5	4.0	32.5

The eight wheat samples are identified as Zhengmai 9023, Zhouyuan 9369, Aobiao, DNS, Jiamai, Jinan17 wheat, Zhoumai 27, and Yunong 416 wheat. For each variety of wheat, 20 samples were prepared

without further purification before grinding. Generally, the wheat samples can transform to more stable form during the storage and manufacture process, such as grinding and compaction. To form thin, circular slices samples, wheat samples were ground into fine powder for 2 min, which was subsequently sieved by filtering laws using 200-eye sieves; then, the sieved powder was pressed into pellets with a thickness of approximately 1 mm and a diameter of 13 mm under a pressure of 5 tons for 5 min. All sample preparation processes were implemented at room temperature.

2.3. Chemometrics Methods

A chemometric analysis was performed to investigate the THz spectral data of wheat varieties using a partial least squares (PLS) regression. The PLS regression is based on latent variables, which are constructed to identify the maximal covariance between two matrices [19]. In particular, PLS was applied to find the best correlation between the spectral data X and measured parameter designating the class of interest Y . The model is formulated as follows:

$$X = TP^T + E \quad (5)$$

$$Y = UQ^T + F \quad (6)$$

where X and Y are the input and output matrices, T and U are the score matrices, P and Q are the loading matrices that can be regarded as the covariance between X and Y and between Y and U , respectively, and E and F are the residual matrices. In this paper, the absorption coefficient and refractive index of a selected frequency range are used as the input matrix X , while the wheat varieties of samples are used as the output matrix Y , which can be regarded as the number for each wheat variety sample, and the corresponding wheat variety can be described using number 1–8 given in Table 1.

The interval partial least squares (iPLS) method is a variable selection technique [20] used for identifying the important spectral regions and removing interference from other regions. The iPLS method divides the whole spectrum into subintervals of equal width and develops the PLS model for each subinterval spectrum. The prediction performance of PLS and iPLS is compared by the root-mean-square error of cross-validation (RMSECV) values, and the best subinterval can be selected by the lowest model RMSECV [21]. The quality of the calibration model is evaluated using the correlation coefficient between the reference and predicted value (R), the root-mean-square error (RMSE) of the calibration set (RMSEC), and the root mean square error of the prediction set (RMSEP) [22]. A better model has a better prediction accuracy that provides higher R and lower RMSECV values. The parameters are calculated as follows:

$$RMSECV = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_r^i - y_p^i)^2} \quad (7)$$

$$R = \frac{\sum_{i=1}^n (y_r^i - \bar{y}_r)(y_p^i - \bar{y}_p)}{\sqrt{\sum_{i=1}^n (y_r^i - \bar{y}_r)^2 \sum_{i=1}^n (y_p^i - \bar{y}_p)^2}} \quad (8)$$

where n is the number of samples in the calibration sample set, y_r^i is the reference value of the i th sample, y_p^i is the predicted value of the i th sample, $\overline{y_r}$ is the average of the sample reference values, and $\overline{y_p}$ is the average of the predicted values of samples.

3. Results and Discussion

3.1. Spectra of Wheat Samples

To remove the random error and increase the signal-to-noise ratio (SNR), each sample is measured three times; the sample spectrum was the average of three scanning spectra in the range of 0.2–2.5 THz, and the reference was measured between every three samples. Figure 1a,b show the time-domain spectra of the eight wheat samples and the corresponding frequency-domain spectra obtained using a fast Fourier transform algorithm. Furthermore, the refractive indices and the absorption coefficients of the eight wheat samples are calculated using Equations (3) and (4), respectively, and the calculated results are shown in Figure 2a,b.

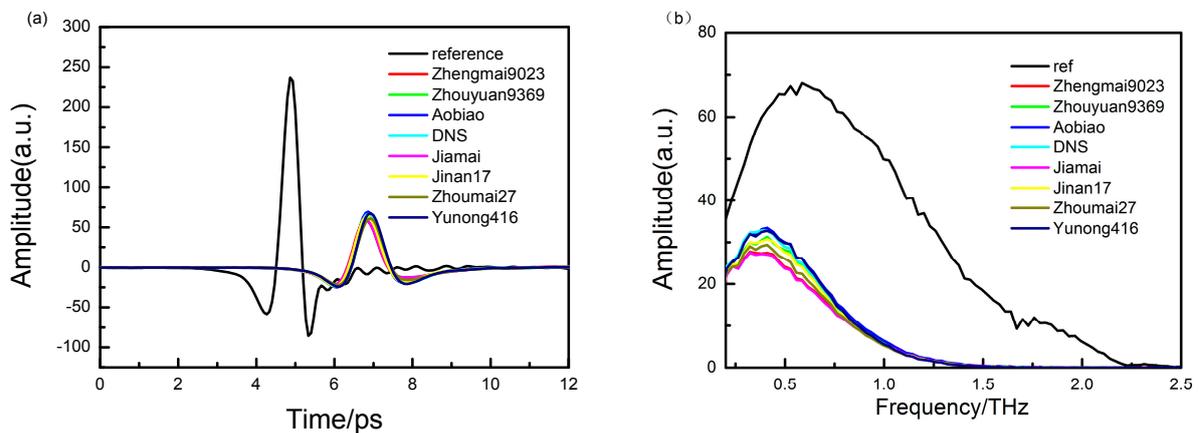


Figure 1. (a) Time-domain THz spectra of the eight wheat samples and reference; and (b) the frequency spectra of the eight wheat samples and the reference in the range of 0.2–2.5 THz.

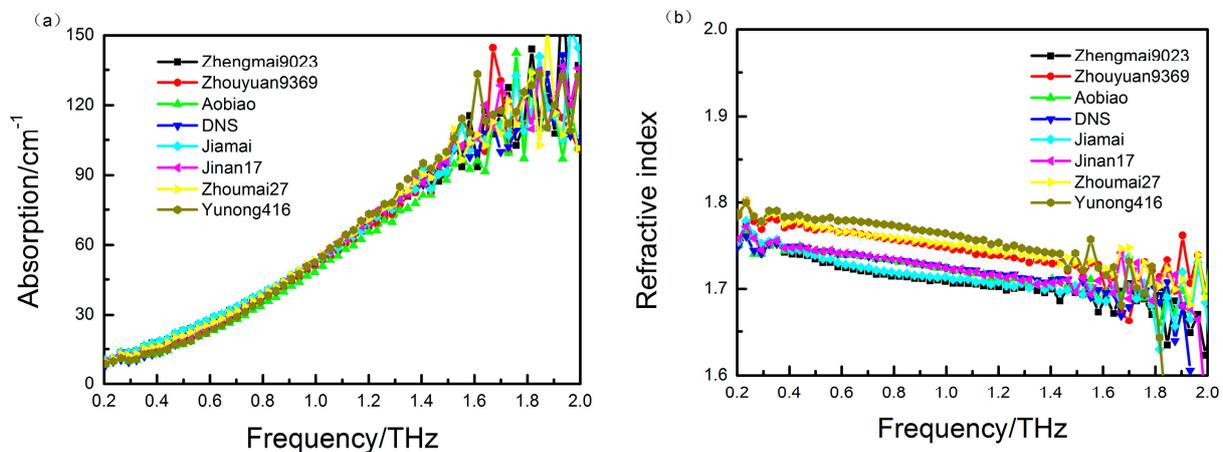


Figure 2. (a) The absorption coefficient and (b) refractive index for the eight wheat samples in the range of 0.2–2.0 THz.

As shown in Figure 1, the effective measurement range is 0.2–2.5 THz, and the sample spectra have narrow bandwidths. Although there is a minor difference in the shift among the samples, there is remarkable shift between the samples and the reference, which indicates that the refractive indices of the samples are different. Moreover, the amplitude changes of the samples indicate that the sample absorption coefficients differ.

Because of the similar components in the samples and the chemical complexity of wheat, the optical parameters of the measured wheat samples are very similar, as shown in Figure 2. In order to represent the variation in measurement effectively, the average absorption coefficients of 20 samples of eight wheat samples are shown in Figure 3. The average absorption coefficients are 37.1492 (zhengmai9023), 39.4354 (zhouyuan9369), 35.3358 (aobiao), 42.1133 (DNS), 37.5468 (jiamai), 42.2520 (jinan17), 39.8226 (zhoumai27) and 39.9409 (yunong416), respectively. In addition, the spectra of wheat samples at higher frequencies (above 1.5 THz) produce lower SNR due to the limitation of the dynamic range of the measurement system. Because there are no obvious absorption peaks in the spectra and because the differences in the absorption spectra for the eight wheat samples are not significant, we employ PLS regression to investigate the relationship between the minor spectral differences and the measured wheat varietal properties.

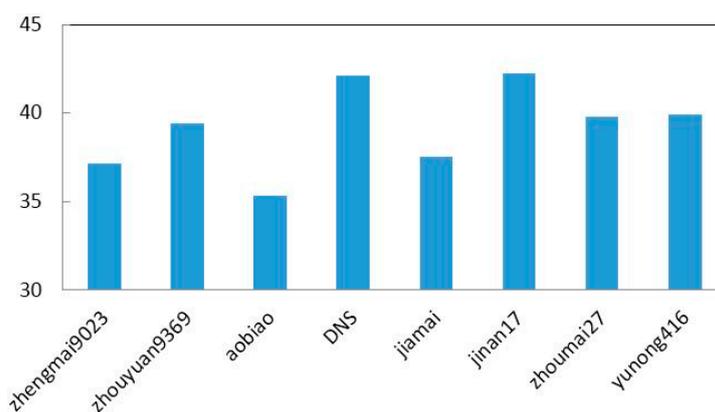


Figure 3. The average absorption coefficient of eight wheat samples.

3.2. PLS Analysis

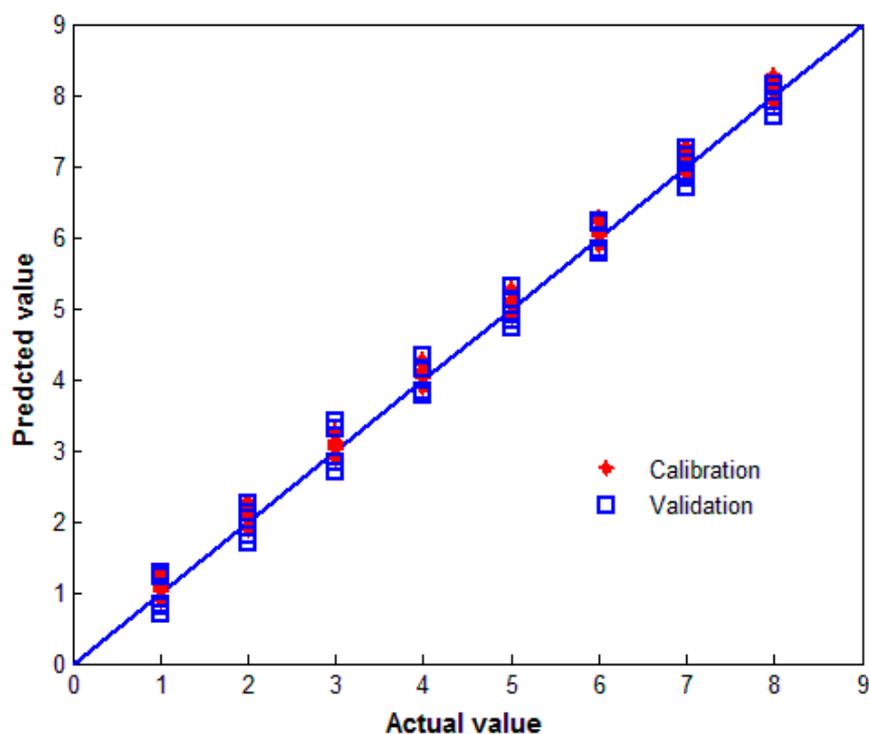
The PLS analysis was performed on THz spectral data with the MATLAB software package (Version 2012a, Mathworks Inc., Natick, MA, USA) using the PLS toolbox (Version 4.0, Eigenvector Research Inc., Wenatchee, USA) and user-written scripts. The original spectra with 128 frequency variables were analyzed to establish the correlation between the spectral data and relevant wheat varieties. The eight wheat samples were set as output variables using numbers 1 through 8 (Table 1). The PLS calibration model was developed using a leave-one-out cross-validation calculation. In this study, both the absorption spectra and the refractive spectra of wheat samples from 0.2 to 1.5 THz were applied to obtain the best model. Then, the models were employed to predict the eight varieties of samples. A set of eight varieties of wheat (160 samples in total) was used in this experiment. All the samples were divided into two sets randomly, the calibration set (96 samples) and the prediction set (64 samples). The performance of the cross-validation models is shown in Table 2.

Table 2. Calibration and validation results obtained with the effective spectrum PLS Model.

Input Variable	Frequency Range (THz)	Factors	Calibration		Cross Validation	
			R	RMSEC	R	RMSECV
Absorption coefficient	0.2–1.5 THz	5	0.987	0.759	0.983	1.028
Refractive index	0.2–1.5 THz	5	0.982	1.472	0.979	1.684

The absorption coefficient, which has a higher R-value and lower RMSE and RMSECV values, demonstrates better performance than the refractive index for the full-spectrum PLS model. Thus, this indicates that the absorption spectrum-based PLS model is a better model for prediction of wheat varieties compared to the refractive index-based PLS model in the frequency range of 0.2–1.5 THz.

Figure 4 shows the calibration and validation results for wheat varietal discrimination using the absorption spectrum PLS regression model. In the model, the reference line indicates the zero residuals between the predicted and the actual values. Figure 4 shows that the predicted values for all varieties of samples agree with the actual values, indicating that the PLS model can identify wheat varieties.

**Figure 4.** The calibration and validation results for wheat discrimination using the PLS model.

3.3. iPLS Analysis

For comparison, iPLS analysis was also performed using the same absorption spectral data sets to improve the model performance. First, the full spectrum was divided into 16 equal subintervals with eight variables. Calibration models were developed for each of the 16 intervals. Then, cross-validation was performed for each of the 16 models. Figure 5 presents the iPLS variable selection results for the discrimination of wheat varieties.

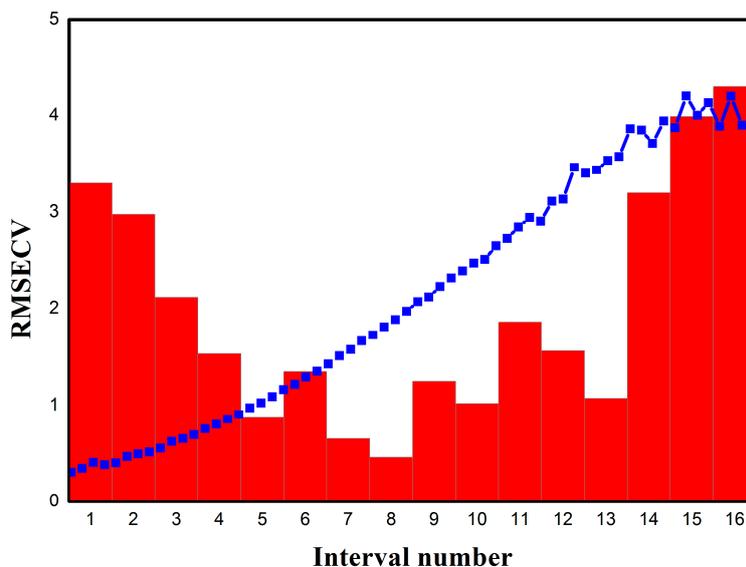


Figure 5. The iPLS results for the THz spectra data. The columns indicate the RMSECV in each subinterval, and the mean absorption spectrum of the wheat samples is overlaid on the plot.

From Figure 5, the width of each column is the same, while the height of each column indicates the RMSECV value calculated by the subinterval PLS model. The eighth column with the interval 57–64, corresponding to the frequency range of 787.5–900 GHz, shows the lowest RMSECV and is selected for developing the model. Therefore, the resulting model should have more precision in discriminating the wheat varieties.

In addition, iPLS selects the most relevant part of the spectrum, which can improve the performance of the model by removing the noise and interference from other regions. In the experiment, a different number of intervals were used to explore the best spectral region by the lowest RMSECV. The interval widths of 4 and 16 were also employed to construct the regression model, dividing the spectrum into 32 and 8 subintervals, respectively. Table 3 shows the calibration and validation results of the optimal iPLS models on THz absorption spectra.

Table 3. Calibration and validation results obtained with the optimal iPLS regression model.

Interval Variables	Frequency Range	R_{cal}	RMSEC	RMSECV
4	0.731–0.956 THz	0.991	0.768	1.260
8	0.787–0.900 THz	0.992	0.573	0.967
16	0.675–1 THz	0.984	0.837	1.237

As shown in Table 3, the iPLS model with eight interval variables has the lowest RMSECV value (0.967), the lowest RMSEC value (0.573), and the highest R value (0.992), compared to those of the full spectrum PLS model (Table 2). It is clear that the performance of the subinterval PLS model was better than the full-spectrum PLS model.

Figure 6 shows plots of the actual values compared to the values predicted by the PLS models based on the full THz absorption spectra and the optimal interval in the 0.787–0.9 THz range. The RMSEP of the full-spectrum PLS model is 0.845, while the RMSEP of the iPLS model is improved to 0.642.

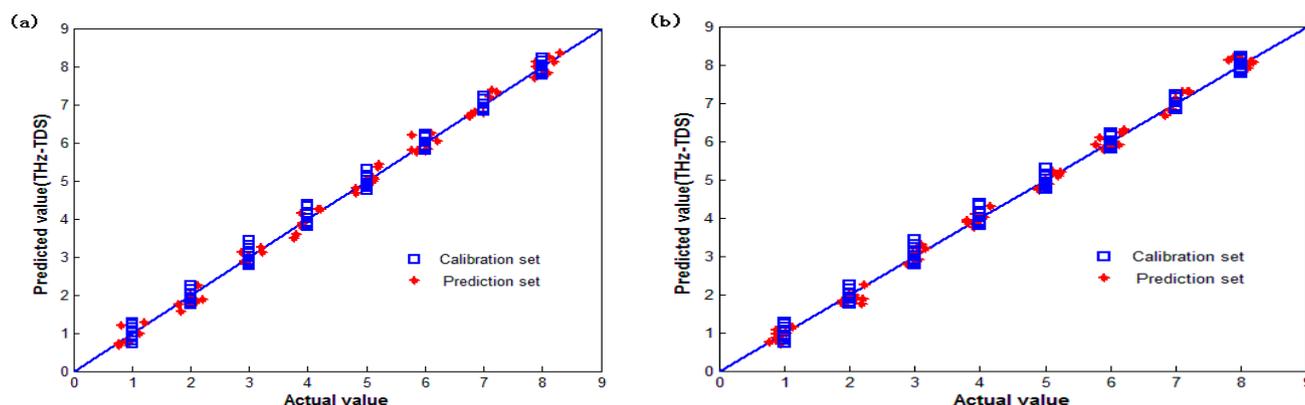


Figure 6. Scatter plots of the actual value vs. the predicted value for the discrimination of wheat varieties using (a) the full-spectrum PLS calibration model and (b) the iPLS calibration model based on the selected subinterval in the range 0.787–0.9 THz.

In Figure 6, the eight subintervals along the X-axis represent the eight wheat varieties as follows: intervals from 0.5–1.5, 1.5–2.5, 2.5–3.5, 3.5–4.5, 4.5–5.5, 5.5–6.5, 6.5–7.5, and 7.5–8.5 represent samples 1–8, respectively. The Y-axis represents the predicted model value for each wheat variety. The maximum range of values predicted by the PLS model for the discrimination wheat grains is 0.67–1.28, while that for the iPLS model is 0.70–1.15, which produces satisfactory results in the evaluation of the wheat varieties. The dispersion degree of the predicted value for Aobiao wheat and Zhengmai 9023 wheat is relatively high, as shown in Figure 6a, but both varieties can be discriminated correctly.

Comparing Figure 6a,b, improved prediction accuracy of the optimized PLS model for the discrimination of wheat varieties is observed, and the relative prediction error in the full-spectrum model is reduced. These results indicate that both the PLS and iPLS models can obtain good predictions for discriminating wheat varieties using THz spectroscopy; however, the iPLS model improves the prediction accuracy through optimal spectral region selection technology.

3.4. Discussion

Compared with other spectroscopic methods, such as NIR, MIR, and Raman, THz spectroscopy is a non-destructive rapid method to discriminate wheat varieties. The authors of [6] report that both MIR and NIR techniques can be applied in the industry as a rapid analytical tool to measure the quality parameters. In our experiment, the total time required for preparation, measurement, and analysis of wheat samples was within 5 min, which is also equal to that in [6,7] using data fusion with multiple analytical measurements to increase the model performance and decrease the modeling error for the prediction of quality parameters of crude oil, and the average RMSEP for the fully fused model (which included the IR, Raman, and NMR data) was calculated to be 0.307%. However, few studies compare the iPLS and PLS methods on the spectroscopic data of sample for the improvement in model performance.

Due to complex compositions in the wheat, the proportion of the characteristic components is small, and the overlapping signals obscure the THz absorption spectra. The absorption spectra for the wheat samples obtained in this experiment are featureless. In fact, featureless spectral absorption in the THz range that increases with increasing frequency is expected for many disordered amorphous materials [23–25]. Even though the characteristic absorption peaks are lacking, variations in the composition of different

wheat samples generate changes in the THz absorption curves. The absorption spectra of various samples have markedly different THz spectra. Thus, chemometric analysis was employed using PLS regression to discriminate the eight wheat samples.

To obtain the best prediction model, both PLS and iPLS analysis methods were applied to construct regression models for discriminating the wheat varieties based on the THz absorption spectra. The above-mentioned analysis indicates that the model performance is primarily affected by the selection of the subinterval range. An appropriate interval width can provide the best prediction results. If the subinterval width is too large, noise and irrelevant information as well as the input sample spectra cause the model to predict the wheat varieties with lower accuracy.

Furthermore, correlations between neighboring regions in the spectrum cannot be obtained by the iPLS model [26]. In particular, the PLS model has only one subinterval (the full spectrum), which includes all spectral information without removing the noise or irrelevant intervals. Thus, the PLS performance is not better than the iPLS model performance. If the interval width is too low, and the corresponding subinterval is too broad, high prediction error will be caused by the loss of small signature information in the spectrum.

Moreover, the prediction accuracy of the constructed model is also affected by the THz absorption of sample. The measured THz spectra of sample are dependent on the experimental environment, experimental setup, background noise, uncertainty of the time-domain signal system [27], and sample preparation process. Moreover, the chemical constituents of wheat grains are determined by many factors [28], such as the wheat variety, plant environment, climate, harvest year, and so on. Even though wheat samples from different harvest years are prepared, the absorption spectra of samples are likely different from each other due to the difference between the chemical compositions of samples. To construct an accurate PLS regression model for wheat varietal discrimination using THz absorption spectra, the influencing factors of the spectra, such as wheat samples from different harvest years in the same plant area and wheat samples from different harvest years and different plant areas, will need to be considered and explored in further systemic studies.

In our work, to avoid the influence of these factors on the performance of the prediction model, the average spectra of the samples were produced by repeated measurements under the same conditions. The PLS regression model was developed to discriminate the wheat varieties based on the THz spectra, and iPLS was used to remove irrelevant information from the spectra. The RMSECV and RMSEC values were used to evaluate the performance of the PLS model. The best subinterval in the iPLS model, which was determined using the lowest RMSECV, was selected for the PLS calibration model. The prediction results for the wheat varieties were improved by using the iPLS model with eight interval variables and the lowest RMSECV value of 0.967.

4. Conclusions

In this study, the potential of THz spectra as a new, non-destructive technology for wheat varietal discrimination has been demonstrated. First, the THz spectra from 0.2 to 2 THz were measured and analyzed for eight wheat varieties using the transmission configuration. Then, PLS regression was employed to obtain a prediction model for wheat varieties. For comparison, iPLS was also employed to improve the performance of the calibration model using the proper subinterval in the spectrum. The

results showed that good prediction models could be obtained with lower RMSECV and higher R-values in the two regression methods. In particular, the prediction accuracy for wheat grain varieties was improved with the iPLS calibration model ($R = 0.992$ and $RMSECV = 0.967$). As a result, THz spectroscopy associated with chemometric techniques showed useful measurement and discrimination of wheat varieties. However, further studies should be considered to account for the large number of mixed varieties and for varieties harvested in different years.

Acknowledgments

This work was supported by the National High-Tech R&D Program of China (863 Program) (Grant No. 2012AA101608) and the National Natural Science Foundation of China (Grant No. 61071197).

Author Contributions

In this paper, Hongyi Ge, Yuying Jiang and Feiyu Lian conducted the algorithm design, experiments and analysis. Technical discussion and data processing were performed by Hongyi Ge and Yuying Jiang. Yuan Zhang and Shanhong Xia have supervised the whole work. All authors contributed to the writing and the revision of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Zapotoczny, P., Discrimination of wheat grain varieties using image analysis: Morphological features. *Eur. Food Res. Technol.* **2011**, *233*, 769–779.
2. Pourreza, A.; Pourreza, H.; Abbaspour-Fard, M.H.; Sadrnia, H., Identification of nine Iranian wheat seed varieties by textural analysis with image processing. *Comput. Electron. Agric.* **2012**, *83*, 102–108.
3. Levandi, T.; Pussa, T.; Vaher, M.; Ingver, A.; Koppel, R.; Kaljurand, M. Principal component analysis of HPLC-MS/MS patterns of wheat (*Triticum aestivum*) varieties. *Proc. Est. Acad. Sci.* **2014**, *63*, 86–92.
4. Paliwal, J.; Visen, N.S.; Jayas, D.S.; White, N.D.G. Cereal grain and dockage identification using machine vision. *Biosyst. Eng.* **2003**, *85*, 51–57.
5. Miralbes, C. Discrimination of European wheat varieties using near infrared reflectance spectroscopy. *Food Chem.* **2008**, *106*, 386–389.
6. Ferreira, D.S.; Galao, O.F.; Pallone, J.A.L.; Poppi, R.J. Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples. *Food Control* **2014**, *35*, 227–232.
7. Dearing, T.I.; Thompson, W.J.; Rechsteiner, C.E.; Marquardt, B.J. Characterization of Crude Oil Products Using Data Fusion of Process Raman, Infrared, and Nuclear Magnetic Resonance (NMR) Spectra. *Appl. Spectrosc.* **2011**, *65*, 181–186.

8. Ferguson, B.; Zhang, X.C. Materials for terahertz science and technology. *Nat. Mater.* **2002**, *1*, 26–33.
9. Vieira, F.S.; Pasquini, C. Determination of Cellulose Crystallinity by Terahertz-Time Domain Spectroscopy. *Anal. Chem.* **2014**, *86*, 3780–3786.
10. Melinger, J.S.; Laman, N.; Grischkowsky, D. The underlying terahertz vibrational spectrum of explosives solids. *Appl. Phys. Lett.* **2008**, *93*, doi:10.1063/1.2949068.
11. Liang, M.Y.; Shen, J.L.; Wang, G.Q. Identification of illicit drugs by using SOM neural networks. *J. Phys. D Appl. Phys.* **2008**, *41*, doi:10.1088/0022-3727/41/13/135306.
12. Ueno, Y.; Rungsawang, R.; Tomita, I.; Ajito, K. Quantitative measurements of amino acids by terahertz time-domain transmission spectroscopy. *Anal. Chem.* **2006**, *78*, 5424–5428.
13. Ge, H.Y.; Jiang, Y.Y.; Xu, Z.H.; Lian, F.Y.; Zhang, Y.; Xia, S.H. Identification of wheat quality using THz spectrum. *Opt. Express* **2014**, *22*, 12533–12544.
14. Pupeza, I.; Wilk, R.; Koch, M. Highly accurate optical material parameter determination with THz time-domain spectroscopy. *Opt. Express* **2007**, *15*, 4335–4350.
15. Xiao-li, Z.; Jiu-sheng, L. Diagnostic techniques of talc powder in flour based on the THz spectroscopy. *J. Phys. Conf. Ser.* **2011**, *276*, doi:10.1088/0022-3727/41/13/135306.
16. Zhang, Y.; Peng, X.H.; Chen, Y.; Chen, J.; Curioni, A.; Andreoni, W.; Nayak, S.K.; Zhang, X.C. A first principle study of terahertz (THz) spectra of acephate. *Chem. Phys. Lett.* **2008**, *452*, 59–66.
17. Dorney, T.D.; Baraniuk, R.G.; Mittleman, D.M. Material parameter estimation with terahertz time-domain spectroscopy. *J. Opt. Soc. Am. A* **2001**, *18*, 1562–1571.
18. Duvillaret, L.; Garet, F.; Coutaz, J.L. A reliable method for extraction of material parameters in terahertz time-domain spectroscopy. *IEEE J. Sel. Top. Quantum Electron.* **1996**, *2*, 739–746.
19. Nicolai, B.M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K.I.; Lammertyn, J. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* **2007**, *46*, 99–118.
20. Norgaard, L.; Saudland, A.; Wagner, J.; Nielsen, J.P.; Munck, L.; Engelsen, S.B. Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Appl. Spectrosc.* **2000**, *54*, 413–419.
21. Kachrimanis, K.; Braun, D.E.; Griesser, U.J. Quantitative analysis of paracetamol polymorphs in powder mixtures by FT-Raman spectroscopy and PLS regression. *J. Pharm. Biomed.* **2007**, *43*, 407–412.
22. Hua, Y.F.; Zhang, H.J. Qualitative and Quantitative Detection of Pesticides with Terahertz Time-Domain Spectroscopy. *IEEE Trans. Microw. Theory Tech.* **2010**, *58*, 2064–2070.
23. Naftaly, M.; Miles, R.E. Terahertz time-domain spectroscopy for material characterization. *IEEE Proc.* **2007**, *95*, 1658–1665.
24. Zhang, F.; Kambara, O.; Tominaga, K.; Nishizawa, J.; Sasaki, T.; Wang, H.W.; Hayashi, M. Analysis of vibrational spectra of solid-state adenine and adenosine in the terahertz region. *R. Soc. Chem. Adv.* **2014**, *4*, 269–278.
25. McIntosh, A.I.; Yang, B.; Goldup, S.M.; Watkinson, M.; Donnan, R.S. Crystallization of amorphous lactose at high humidity studied by terahertz time domain spectroscopy. *Chem. Phys. Lett.* **2013**, *558*, 104–108.

26. Andersen, C.M.; Bro, R. Variable selection in regression—a tutorial. *J. Chemometr.* **2010**, *24*, 728–737.
27. Withayachumnankul, W.; Fischer, B.M.; Lin, H.Y.; Abbott, D. Uncertainty in terahertz time-domain spectroscopy measurement. *J. Opt. Soc. Am. B* **2008**, *25*, 1059–1072.
28. Triboi, E.; Abad, A.; Michelena, A.; Lloveras, J.; Ollier, J.L.; Daniel, C. Environmental effects on the quality of two wheat genotypes: 1. quantitative and qualitative variation of storage proteins. *Eur. J. Agron.* **2000**, *13*, 47–64.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).