

Determination of potassium sorbate and sorbic acid in agricultural products using THz time-domain spectroscopy

Yuying Jiang(蒋玉英), Guangming Li(李广明), Ming Lv(吕明), Hongyi Ge(葛宏义), Yuan Zhang (张元)

Citation: Chin. Phys. B . 2020, 29(9): 098705 . **doi:** 10.1088/1674-1056/ab9f25
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Terahertz polarization conversion and sensing with double-layer chiral metasurface

Zi-Yang Zhang(张子扬), Fei Fan(范飞), Teng-Fei Li(李腾飞), Yun-Yun Ji(冀允允), Sheng-Jiang Chang(常胜江)
Chin. Phys. B . 2020, 29(7): 078707 . **doi:** 10.1088/1674-1056/ab9294

Broadband terahertz time-domain spectroscopy and fast FMCW imaging: Principle and applications

Yao-Chun Shen(沈耀春), Xing-Yu Yang(杨星宇), Zi-Jian Zhang(张子健)
Chin. Phys. B . 2020, 29(7): 078705 . **doi:** 10.1088/1674-1056/ab9296

High birefringence, low loss, and flattened dispersion photonic crystal fiber for terahertz application

Dou-Dou Wang(王豆豆), Chang-Long Mu(穆长龙), De-Peng Kong(孔德鹏), Chen-Yu Guo(郭晨瑜)
Chin. Phys. B . 2019, 28(11): 118701 . **doi:** 10.1088/1674-1056/ab44af

Terahertz coherent detection via two-color laser pulses of various frequency ratios

Xin-Yang Gu(顾新杨), Ke-Jia Wang(王可嘉), Zhen-Gang Yang(杨振刚), Jin-Song Liu(刘劲松)
Chin. Phys. B . 2019, 28(9): 098701 . **doi:** 10.1088/1674-1056/ab37fc

Resonance properties of THz plasmonic dipole-bowtie antenna array: The critical role of the substrate

Yang Yu-Ping, Ranjan Singh, Zhang Wei-Li
Chin. Phys. B . 2014, 23(12): 128702 . **doi:** 10.1088/1674-1056/23/12/128702

SPECIAL TOPIC — Terahertz physics

Determination of potassium sorbate and sorbic acid in agricultural products using THz time-domain spectroscopy*

Yuying Jiang(蒋玉英)[1,2], Guangming Li(李广明)[1,2,3], Ming Lv(吕明)[1,2,3], Hongyi Ge(葛宏义)[1,2,3,†], and Yuan Zhang (张元)[1,2,‡]

[1] *Key Laboratory of Grain Information Processing and Control (Henan University of Technology), Ministry of Education, Zhengzhou* 450001, *China*
[2] *Henan Provincial Key Laboratory of Grain Photoelectric Detection and Control, Zhengzhou* 450001, *China*
[3] *College of Information Science and Engineering, Henan University of Technology, Zhengzhou* 450001, *China*

(Received 31 March 2020; revised manuscript received 2 June 2020; accepted manuscript online 23 June 2020)

The aim of this study was to investigate the feasibility of detecting potassium sorbate (PS) and sorbic acid (SA) in agricultural products using THz time-domain spectroscopy (THz-TDS). The absorption spectra of PS and SA were measured from 0.2 to 1.6 THz at room temperature. The main characteristic absorption peaks of PS and SA in polyethylene and powdered agricultural products with different weight ratios were detected and analyzed. Interval partial least squares (iPLS) combined with a particle swarm optimization and support vector classification (PSO-SVC) algorithm was proposed in this paper. iPLS was used for frequency optimization, and the PSO-SVC algorithm was used for spectrum analysis of the preservative content based on the optimal spectrum ranges. Optimized PSO-SVC models were obtained when the THz spectrum from the PS/SA mixture was divided into 11 or 12 subintervals. The optimal penalty parameter c and kernel parameter g were found to be 1.284 and 0.863 for PS (0.551–1.487 THz), 1.374 and 0.906 for SA (0.454–1.216 THz), respectively. The preliminary results indicate that THz-TDS can be an effective nondestructive analytical tool used for the quantitative detection of additives in agricultural products.

Keywords: THz-TDS, preservative content, quantitative analysis, iPLS, PSO-SVC, nondestructive detection

PACS: 87.50.U– **DOI:** 10.1088/1674-1056/ab9f25

1. Introduction

Consumers are concerned about food quality, especially given the wide use of food additives like potassium sorbate (PS) and sorbic acid (SA), which may leave residues in agricultural products and food. As acidic preservatives, PS and SA are used to inhibit the growth of mold and microorganisms, however, the excessive use of preservatives in agricultural products can create threats to human health.[1] There are several technologies[2,3] used to measure PS and SA content, such as high-performance liquid chromatography, gas chromatography, and mass spectroscopy. Although these technologies are sensitive and accurate, they are time-consuming, expensive, and require sophisticated sample preparation processes performed by specialists. Therefore, a rapid, accurate, nondestructive, and reliable method for measuring PS and SA residues in food is desired.

THz time-domain spectroscopy (THz-TDS) can be used to identify different chemicals, including many nonpolar materials,[4,5] which can be used as a spectroscopic tool for nondestructive testing. Compared to the other spectroscopic technologies, THz spectroscopy can be used to directly measure the absorption coefficient and refractive index of the sample with a high signal-to-noise ratio (SNR) with pump-probe detection and without using Kramers–Kronig relation. Thus, THz-TDS has been used as an analytical tool in various fields,[6,7] such as chemistry, biology, and food quality control. Moreover, a number of recent studies have focused on using THz spectroscopy with chemometrics in agriculture and the food industry for qualitative and quantitative analysis. Nishikiori *et al.* qualitatively studied the THz spectra of L-, D-, and DL-tartaric acid and quantitatively analyzed a tartaric acid mixture using THz-TDS with PLS.[8] Hua and Zhang qualitatively and quantitatively studied solid pesticides using THz-TDS with other chemometric methods. The results indicate that THz-TDS was promising for residual pesticide detection in food.[9] Ma *et al.* measured the spectra of thiabendazole using THz-TDS and quantitatively analyzed the content of residues with PLS, interval PLS (iPLS), backward iPLS, and moving window PLS methods.[10] Qin *et al.* quantitatively studied the THz spectra of TCH and CCH in soil and rice samples using PLS.[11] Those reports presented qualitative and quantitative analyses of materials using THz-TDS,

*Project supported by the National Natural Science Foundation of China (Grant Nos. 61705061 and 61975053), Key Science and Technology Program of Henan Province of China (Grant Nos. 182102110204 and 192102110047), Key Scientific and Research Project of Educational Committee of Henan Province of China (Grant No. 19B510001), and Open Fund Project of Key Laboratory of Grain Information Processing & Control, Ministry of Education, Henan University of Technology (Grant Nos. KFJJ2016108 and KFJJ2017107).

†Corresponding author. E-mail: gehongyi2004@163.com
‡Corresponding author. E-mail: zy-haut@163.com

© 2020 Chinese Physical Society and IOP Publishing Ltd

although a few studies focused on qualitative and quantitative analysis of preservatives in agricultural products, especially in stored wheat. The use of an improved algorithm for optimization support vector machine parameters by particle swarm, which provides higher prediction accuracy of PM2.5 concentration.[12] PSO-SVC is a classification optimization algorithm, which could provide high classification accuracy and generalization ability.

In this study, two kinds of preservatives were measured using THz-TDS, and the main absorption peaks in PS and SA in polyethylene and wheat powder with different weight ratios were detected and analyzed in this paper. PSO-SVC models were constructed based on the full spectrum and feature frequency ranges selected by iPLS to determine preservative contents from THz absorption spectra.

2. Experimental methods and material

2.1. Experimental setup

The THz spectra of preservatives and mixtures were measured using a terahertz spectroscopy system (Zomega Terahertz Corporation, USA) in collimated geometry. The NIR laser (FemtoFiber Pro), which provided 100 fs pulses at a wavelength of 780 nm and a repeating frequency of 80 MHz, was used as a pump source. A low temperature-grown GaAs photoconductive antenna was used for THz generation and a ZnTe electro-optic (EO) crystal was used for THz detection. This system was used to acquire time-domain data by measuring the time delay between generated and detected pulses. The spectral range of the system was 0 to 3.5 THz with 0.03 THz resolution, and its SNR was 5000 : 1. 10 references (without sample) and the sample spectra were collected and averaged to reduce the random error. All measurements were gathered at room temperature at approximately 1% relative humidity. A more detailed description of the system configuration and analysis method can be found in the literature.[13]

The signal transmitted from the sample holder (without a sample) was used as the reference signal. The time-domain pulse is transformed to the frequency domain with a fast Fourier transform (FFT)[14]

$$\tilde{E}(\omega) = A(\omega)\,\mathrm{e}^{-\mathrm{i}\phi(\omega)} = \int \mathrm{d}t E(t)\,\mathrm{e}^{-\mathrm{i}\omega t}\phi(\omega), \qquad (1)$$

where $A(\omega)$ and $\phi(\omega)$ are the amplitude and phase of the electric field, respectively, and $E(t)$ is the time-domain waveform.

The absorption coefficient $\alpha(\omega)$ and refractive index $n(\omega)$ of a sample can be extracted from the following equations:[15,16]

$$n(\omega) = \frac{\varphi(\omega)}{\omega d}c + 1, \qquad (2)$$

$$a(\omega) = \frac{2}{d}\ln\left[\frac{4n(\omega)}{\rho(\omega)(n(\omega)+1)^2}\right], \qquad (3)$$

where c is the speed of light, ω is the wave frequency, $\rho(\omega)$ and $\phi(\omega)$ are the amplitude ratio and phase difference between the reference and sample, respectively, and d is the sample thickness.

2.2. Sample preparation

Crystalline PS and SA used in our experiment were purchased from Tianjin Dongda chemical Co., Ltd with 99% purity. The chemical formulas for PS and SA are $C_6H_7O_2K$ and $C_6H_8O_2$, respectively. High-density polyethylene (HDPE) powder was purchased from Sigma-Aldrich (St. Louis, USA). Stored wheat was supplied from Henan university of technology (Zhengzhou, China). All materials were used without further purification.

The PS and SA powder samples were crushed into small particles with a mortar and pestle, and the wheat samples were milled into a fine powder and sieved by filtering through 200-eye sieves to reduce the baseline offsets at higher frequencies. These particles were then mixed with HDPE and sieved the wheat powder in different concentrations, ranging from 0% to 50% by weight. The components in each mixture are shown in Table 1. 20 samples of each mixture were selected, and the water content in these samples was $< 12\%$. The mixture was compressed into 13 mm diameter pellets at 10 MPa for 5 min with a tablet press. The resulting sample thickness ranged from 1 mm to 2 mm, which provides a sufficient path length to eliminate the effect of multiple reflections between the two sample surfaces.[17] All sample preparation processes were performed at room temperature and labeled according to their properties.

Table 1. Component of mixture.

No.	PS/SA/mg	PE/wheat/mg	Concentration/%
1	0	200	0
2	5	195	2.5
3	10	190	5
4	20	180	10
5	30	170	15
6	40	160	20
7	60	140	30
8	80	120	40
9	100	100	50
10	200	0	100

2.3. Chemometric methods

Chemometrics methods, including PCR, SVM, and PLS, have been used to analyze features in THz spectra and investigate the correlation between spectral features and sample characteristics.[18–21] In this study, Matlab and Origin software (Origin Lab, Northampton, MA, USA) were used for spectral processing and multivariate analysis.

2.3.1. Interval PLS

Interval partial least squares (iPLS) is a variable selection technique used for identifying the important spectral regions and for removing interference between different regions.[22] iPLS was used to optimize the models and develop calibration models based on spectral subintervals of equal width, which can then be used to determine preservative contents in different mixtures.

The model performance was evaluated in terms of the correlation coefficient between the reference and predicted value (R), root-mean-square error of cross-validation (RMSECV) values, root-mean-square error (RMSE) of the calibration set (RMSEC), and root mean square error of the prediction set (RMSEP).[23] A more accurate model provides higher R and lower RMSECV values. These parameters are defined as follows:

$$R = \frac{\sum_{i=1}^{n}(y_{\mathrm{r}}^{i}-\overline{y_{\mathrm{r}}})(y_{\mathrm{p}}^{i}-\overline{y_{\mathrm{p}}})}{\sqrt{\sum_{i=1}^{n}(y_{\mathrm{r}}^{i}-\overline{y_{\mathrm{r}}})^{2}\sum_{i=1}^{n}(y_{\mathrm{p}}^{i}-\overline{y_{\mathrm{p}}})^{2}}}, \tag{4}$$

$$\mathrm{RMSECV} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{\mathrm{r}}^{i}-y_{\mathrm{p}}^{i})^{2}}, \tag{5}$$

where n is the number of measurements 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, $\overline{y_{\mathrm{r}}}$ is the average of the sample reference values, and $\overline{y_{\mathrm{p}}}$ is the average of the predicted values of samples.

2.3.2. PSO-SVC

Support vector classification (SVC) can be used for multi-class classification.[24] Particle swarm optimization (PSO) is inspired by the behavior of social organisms in groups, such as ant colonies, and schools of birds and fish. This heuristic algorithm randomly moves a solution population within the solution space.[25] In order to build an optimal SVC model, the RBF was used as the kernel function, and the PSO was used for optimizing the parameters in the SVC.

The specific steps for optimization of an SVC model with PSO are as follows:

(1) Determine training, test sets for the SVC and define a search interval for the penalty parameter c and kernel parameter g;

(2) Determine the value of the objective function, suppose there are N particles in the D-dimensional, initialize the position and speed of particle i as: $X_i = (X_{i1}, X_{i2}, \ldots, X_{iD})$ and $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$;

(3) Calculate the fitness function $f(X_i)$, find the best position experienced by the particle i: $pbest_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$ and the best position experienced by the population: $gbest_i = (g_1, g_2, \ldots, g_D)$;

Usually, the position and speed of d-th dimension are range within $[X_{\min,d}, X_{\max,d}]$ and $[-V_{\max,d}, V_{\max,d}]$ respectively; if V_{id} and X_{id} exceeds the boundary value, then the V_{id} and X_{id} are defined as the maximum speed and then boundary position respectively;

(4) Update the d-th dimension speed and position of particle i:

$$V_{id}^{k} = WV_{id}^{k-1} + c_1 r_1 (pbest_{id} - X_{id}^{k-1}) + c_2 r_2 (gbest_d - X_{id}^{k-1}), \tag{6}$$

$$X_{id}^{k} = X_{id}^{k-1} + V_{id}^{k-1}, \tag{7}$$

where k represents the number of iterations; c_1, c_2 represent the acceleration constant, which adjust the maximum learning steps; r_1, r_2 are two random functions used to increase search randomness; W is an inertial weight used to adjust the search range of the solution.

(5) When reached the maximum number of iterations or the global best position met the minimum limit, go to step (5), otherwise, return to step (3);

(6) Output the optimal parameters c and g, then SVC classification using the optimal parameters c and g.

3. Results and analysis

3.1. THz spectra from pure compounds

THz time-domain spectra of PS and BA are shown in Fig. 1, and the absorption coefficients and refractive indices are shown in Fig. 2. Each spectrum is presented as an average of 5 individual spectra, and the reference was measured between every two samples. Frequencies below 0.2 THz or above 1.6 THz had very low signal-to-noise ratio (SNR) and would not be considered effective data, and it is clear that the useful measurement range is 0.2 to 1.6 THz.

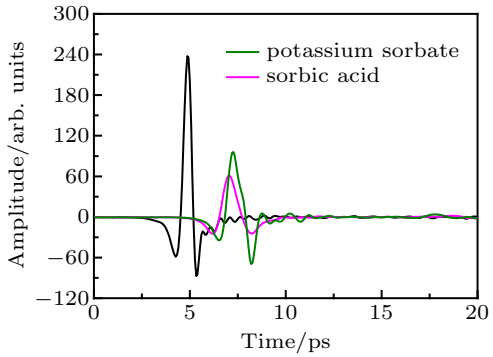


Fig. 1. THz time-domain spectra of the PS and SA.

As shown in Fig. 2, the two preservatives exhibit their own distinctive THz absorption characteristics. The obvious absorption peak is located at 0.97 THz for PS, and at 0.88 THz for SA, while the other absorption peaks are relatively weak; the average refractive indices of the two preservatives are 1.81

and 1.69, respectively. Because THz is sensitive to the inter- or intramolecular vibrational states of the preservatives, these unique absorption peaks in PS and SA could be used as fingerprint features for quick qualitative analysis of preservatives.

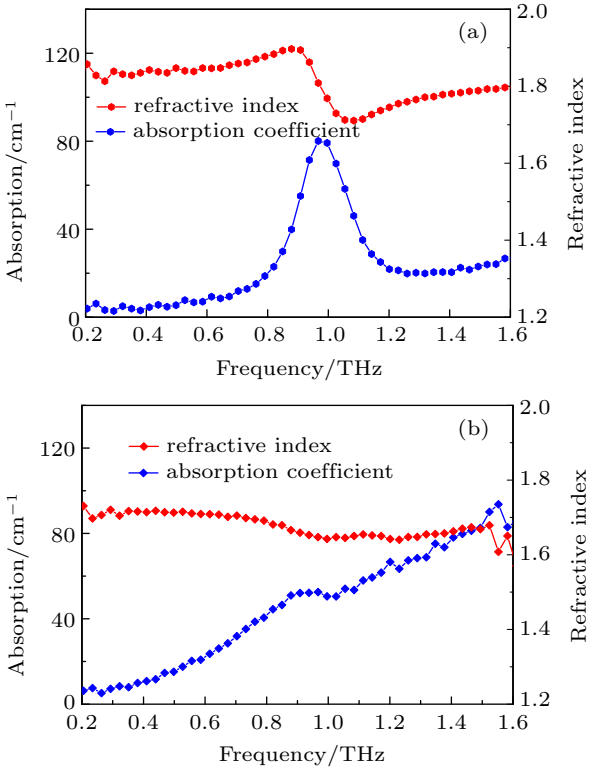


Fig. 2. Absorption coefficients and refractive indices of the (a) PS and (b) SA.

3.2. THz spectroscopic analysis of preservative mixtures

Twenty samples were examined at each concentration listed in Table 1, and a total of 200 samples THz spectra were randomly divided into a calibration set (120 bands) and a prediction set (80 bands). Five absorption spectra were gathered from each sample and averaged to calculate the stability of the system and the robustness of the model. A Fourier transform was used to convert the time-domain spectra to frequency-domain spectra. The absorption coefficients of different concentrations of PS/polyethylene/wheat mixtures and SA/polyethylene/wheat mixtures were calculated, as shown in Figs. 3 and 4 respectively.

Figure 3 shows a distinct absorption peak at 0.97 THz, which coincides with the characteristic absorption peak from pure PS, as shown in Fig. 2(a). The absorption peaks at 1.44 and 0.7 THz are not obvious and can only be discerned at high PS content (greater than 5 %). Figure 4 shows a distinct absorption peak at 0.88 THz, which coincides with the absorption peak in pure SA, as shown in Fig. 2(b). The absorption peaks at 1.6 and 0.64 THz are not obvious and can only be discerned at high SA content (higher greater than 5%).

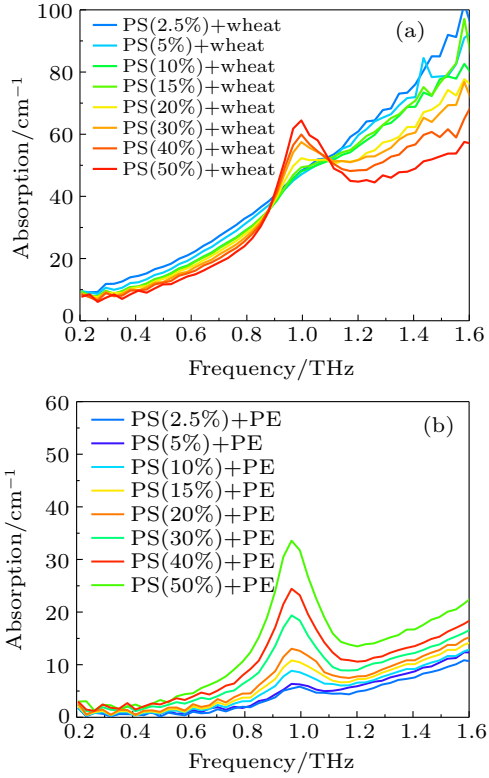


Fig. 3. THz absorption spectrum of PS and wheat/polyethylene mixture.

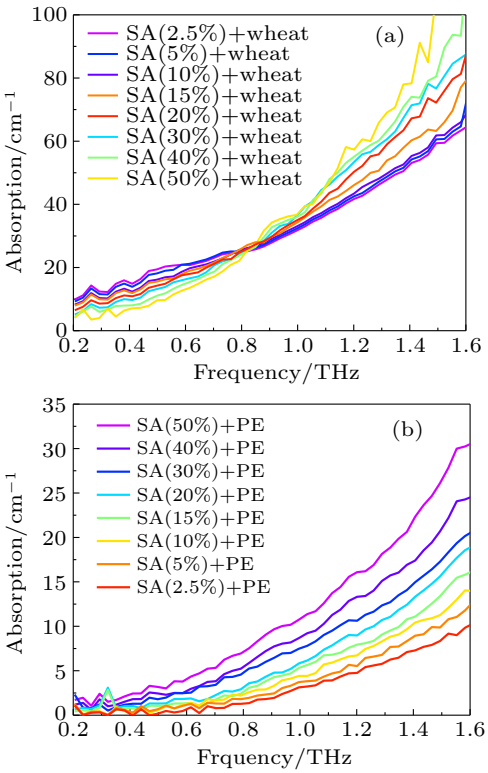


Fig. 4. THz absorption spectrum in SA and wheat/polyethylene mixtures.

As shown in Figs. 3(a) and 4(a), due to strong absorption and scattering of high-frequency THz radiation by PS/SA and polyethylene/wheat mixtures, the absorption spectra of these mixtures oscillated above 1.6 THz. As expected, similar trends were observed between 0.2 THz and 1.6 THz. As

shown in Fig. 3(a), near the characteristic absorption peaks, the absorption intensities and absorption coefficients of the PS/polyethylene/wheat mixtures increased as the PS concentration increased, while the trends of other parts are opposite. As shown in Fig. 4(a), the absorption intensities and coefficients from 0.2 THz to 0.88 THz in the SA/polyethylene/wheat mixtures increased as the SA concentration increased, while the opposite trend was observed from 0.88 to 1.6 THz. In addition, the absorption spectrum of PS/SA and wheat powder mixtures increases approximately linearly as the frequency increased.

3.3. Quantitative analysis of preservative mixtures

As shown in Fig. 2, the absorption peaks of PS appeared at 0.7 THz, 0.97 THz, and 1.44 THz. Then the feature frequency ranges should be selected as 0.5–1.6 THz. As for other frequency ranges, the spectral curve is flat and no absorption peaks were observed. The iPLS algorithm can be used to obtain the proper feature frequency ranges. The full spectrum can be divided into 10 to 16 subintervals in turn, and a PLS regression model was established in each interval. Figure 5 shows a spectrum divided into 12 subintervals. The black dashed line shows the RMSECV value for PLS based on the full-spectrum model, which is called the threshold for interval selection. The columns show the RMSECV value in each subinterval. The RMSECV values in the 3rd through 10th intervals are significantly lower than the threshold. The iPLS wave screening results for 12 subintervals as shown in Table 2. From Table 2, we can see that the feature frequency range could be set to 0.551–1.487 THz.

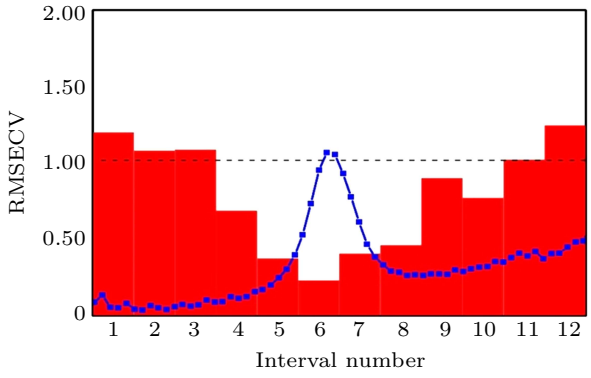


Fig. 5. iPLS results for 12 subintervals in the absorption spectrum from PS.

The iPLS wave screening results of different subintervals was shown in the Table 3. The model performance is evaluated using the average RMSECV and average correlation coefficient R. As shown in Table 3, the model with a PS spectrum divided 12 subintervals produced the lowest average RMSECV (0.7521) and the highest average R (0.9637) values. Meanwhile, the model with an SA spectrum divided into 11 subintervals produced RMSECV and R values of 0.7631 and 0.9614, respectively, when the selected feature frequency range was 0.454 THz to 1.216 THz.

Table 2. iPLS wave screening results for 12 subintervals of the absorption spectrum from PS.

No.	Frequency range/THz	RMSECV	R
Full spectrum	0.2–1.6	1.0198	0.932
The 1st subinterval	0.2–0.317	1.2135	0.898
The 2nd subinterval	0.317–0.434	1.0832	0.942
The 3rd subinterval	0.434–0.551	1.0863	0.991
The 4th subinterval	0.551–0.668	0.6882	0.989
The 5th subinterval	0.668–0.785	0.3754	0.996
The 6th subinterval	0.785–0.902	0.2319	0.998
The 7th subinterval	0.902–1.019	0.4073	0.997
The 8th subinterval	1.019–1.136	0.4621	0.996
The 9th subinterval	1.136–1.253	0.9003	0.935
The 10th subinterval	1.253–1.37	0.7712	0.990
The 11th subinterval	1.37–1.487	1.0214	0.936
The 12th subinterval	1.487–1.6	1.2465	0.896

Table 3. iPLS wave screening results for different subintervals of the PS spectra.

No. of subintervals	Feature frequency ranges/THz	Average RMSECV	Average R
10	0.48–1.46	0.8653	0.9513
11	0.454–1.47	0.8464	0.9598
12	0.551–1.487	0.7521	0.9637
13	0.546–1.496	0.7907	0.9621
14	0.5–1.5	0.8763	0.9465
15	0.479–1.502	0.8583	0.95587
16	0.552–1.52	0.8906	0.9476

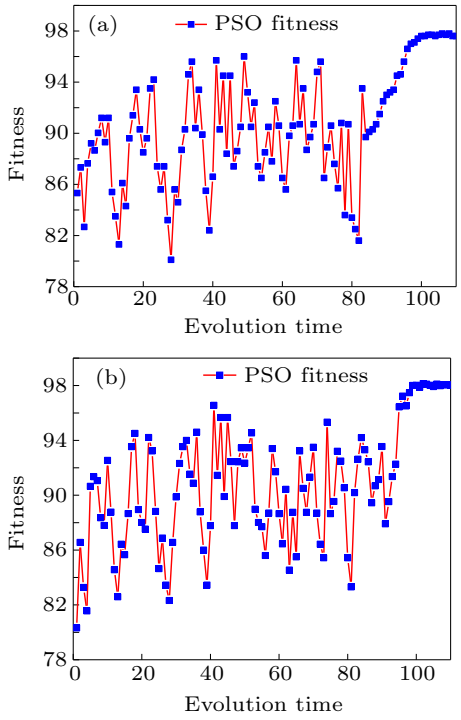


Fig. 6. PSO fitness function for (a) PS and wheat mixtures, and (b) SA and wheat mixtures.

Features at frequencies ranging from 0.551 THz to 1.487 THz for PS and 0.454 THz to 1.216 THz for SA were se-

lected as the input of the PSO-SVC model for quantitative detection of PS/SA mixtures. Figures 6(a) and 6(b) show the fitness curves with optimized parameters determined using PSO, where the population number is 30. As shown in Fig. 6, the PSO fitness function tends to flatten when the iteration times are over 100. Meanwhile, the prediction accuracy with the PSO-SVC model for PS/SA and wheat mixtures is 97.765% and 98.32%, respectively. The optimal penalty parameter is $c = 1.284$ and 1.374, and the kernel parameter is $g = 0.863$ and 0.906 for PS and SA mixtures with wheat, respectively.

Figure 7 shows a comparison of the actual values and those predicted with the PSO-SVC from 0.551 THz to 1.487 THz and 0.454 THz to 1.216 THz, respectively. The red straight line has a slope of 1, meaning the actual value is equal to the predicted value. Figure 7 shows that the predicted value of the PS/SA and PE mixtures is more accurate than the predicted value for the PS/SA and wheat mixtures. The reason is that there are less scattering and absorption in the PE mixture than these in the wheat powder mixture.

PSO-SVC models were also constructed from the full PS/SA absorption spectrum, which were divided into 12 and 11 subintervals, respectively, in order to verify the reliability of the model. The prediction results with different models are shown in Table 4. Compared to the prediction results based on the full spectrum, the optimized PSO-SVC model provides higher prediction accuracy for PS/SA mixtures. The results indicate that the model provides higher prediction accuracy when the optimal spectral region is selected by iPLS.

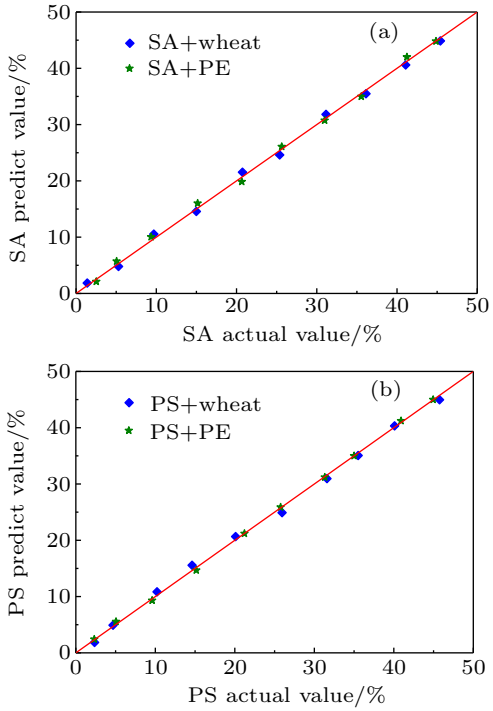


Fig. 7. Scatter plots of the predicted vs. actual value in (a) PS and wheat/PE mixture, and (b) SA and wheat/PE mixture with the PSO-SVC model.

Table 4. Prediction results with different PSO-SVC models.

Sample	Frequency range/THz	c	g	Prediction	
				Correction set/%	Prediction set/%
PS + wheat	0.2–1.6	2.163	1.032	96.653	93.672
	0.551–1.487	1.284	0.863	97.765	95.356
PS + PE	0.2–1.6	2.026	0.965	97.356	95.685
	0.551–1.487	1.206	0.843	98.65	96.657
SA + wheat	0.2–1.6	1.982	1.103	97.473	95.890
	0.454–1.216	1.374	0.906	98.32	97.242
SA + PE	0.2–1.6	1.865	1.214	98.656	97.356
	0.454–1.216	1.179	0.867	99.01	98.013

These results indicate the preservative content in stored grain can be measured with THz-TDS combined with a chemometric method. The iPLS algorithm reduces the acquisition of redundant information of spectral data, simplifies the training model with fewer wavelength points, and improves the efficiency and reliability of the model. PSO-SVC combined with the iPLS method obtains higher prediction accuracy than that of full-frequency modeling without preprocessing of denoising and dimensionality reduction. Generalization, it reduces overfitting and improves the generalization ability of the model.

There still are certain prediction errors during detection, especially when the preservative content is low. In order to increase the quantitative prediction accuracy and further decrease the detection limit, including the choice of classification models and feature spectra, more samples number, and varieties should be used in further research. Meanwhile, the system and background noise should be considered, and more complex situations, such as the existence of more than two components in the mixture samples, should be taken into account for the quality control of stored grain and agriculture products.

4. Conclusions

In this paper, we present a new classification method for the quantitative determination of preservative content in stored grain. First, two kinds of preservatives (potassium sorbate and sorbic acid) were measured and analyzed. An iPLS algorithm

was used for spectral optimization and regression analysis, and the PSO-SVC algorithm was used to measure the preservative content based on both the full spectrum and feature frequency range. Combined with iPLS, the PSO-SVC algorithm produces higher accuracy when the feature frequency range is used for classification than when the full spectrum is used for classification. The experimental results demonstrate the feasibility of using iPLS and the PSO-SVC algorithm to determine preservative content from THz-TDS data. The method presented here is also suitable for THz spectrum analysis and identification of other additives in agricultural products.

References

[1] Mohammadzadeh-Aghdash H, Akbari N, Esazadeh K and Ezzati Nazhad Dolatabadi J 2019 *Food Chem.* **293** 491

[2] de Oliveira Arias J L, Rocha C B, Santos A L Q S, Marube L C, Kupski L, Caldas S S and Primel E G 2019 *Food Chem.* **293** 112

[3] Dehghan P, Mohammadi A, Mohammadzadeh-Aghdash H and Ezzati Nazhad Dolatabadi J 2018 *Trends Food Sci. Tech.* **80** 123

[4] Ferguson B and Zhang X C 2002 *Nat Mater* **1** 26

[5] Wang K, Sun D W and Pu H 2017 *Trends Food Sci. Tech.* **67** 93

[6] Smolyanskaya O A, Chernomyrdin N V, Konovko A A, Zaytsev K I, Ozheredov I A, Cherkasova O P, Nazarov M M, Guillet J P, Kozlov S A, Kistenev Y V, Coutaz J L, Mounaix P, Vaks V L, Son J H, Cheon H, Wallace V P, Feldman Y, Popov I, Yaroslavsky A N, Shkurinov A P and Tuchin V V 2018 *Prog. Quantum Electron.* **62** 1

[7] Ge H, Jiang Y, Lian F, Zhang Y and Xia S 2016 *Food Chem.* **209** 286

[8] Nishikiori R, Yamaguchi M, Takano K, Enatsu T, Tani M, de Silva U C, Kawashita N, Taragi T, Morimoto S, Hangyo M and Kawase M 2008 *Chem. Pharm Bull.* **56** 305

[9] Hua Y F and Zhang H J 2010 *Ieee T Microw Theory* **58** 2064

[10] Ma Y H, Wang Q and Li L Y 2013 *J. Quant Spectrosc. Ra* **117** 7

[11] Wang Y, Wang Q, Zhao Z, Liu A, Tian Y and Qin J 2018 *Talanta* **190** 284

[12] Liu W, Guo G, Chen F and Chen Y 2019 *Atmos. Pollution Res.* **10** 1482

[13] Jiang Y, Ge H and Zhang Y 2020 *Food Chem.* **307** 125533

[14] Arora A, Luong T Q, Kruger M, Kim Y J, Nam C H, Manz A and Havenith M 2012 *Analyst* **137** 575

[15] Liu C, Yue L Y, Wang X K, Sun W F and Zhang Y 2012 *Spectrosc. Spect. Anal.* **32** 1471

[16] Duvillaret L, Garet F and Coutaz J L 1996 *Ieee J. Sel Top. Quant* **2** 739

[17] Ge H Y, Jiang Y Y, Xu Z H, Lian F Y, Zhang Y and Xia S H 2014 *Opt. Express* **22** 12533

[18] Zhao Y, Chu B, Fang S, Zhao J, Zhang H and Yu K 2020 *Spectrochim. Acta Part. A: Mol. Biomolecular Spectrosc.* **230** 118048

[19] Zhang X, Lu S, Liao Y and Zhang Z 2017 *Chemometr Intell. Lab* **164** 8

[20] Du C, Zhang X and Zhang Z 2019 *Vib. Spectrosc.* **100** 64

[21] El Haddad J, de Miollis F, Bou Sleiman J, Canioni L, Mounaix P and Bousquet B 2014 *Anal Chem.* **86** 4927

[22] Nørgaard L, Saudland A, Wagner J, Nielsen J P, Munck L and Engelsen S B 2000 *Appl. Spectrosc.* **54** 413

[23] Chen Z, Zhang Z, Zhu R, Xiang Y, Yang Y and Harrington P B 2015 *J. Quant. Spectrosc. Radiat. Transfer* **167** 1

[24] He M, Yang G L and Xie H Y 2013 *Opt. Express* **21** 6346

[25] Kennedy J and Eberhart R 1995 *Particle swarm optimization*, Proceedings of ICNN'95-International Conference on Neural Networks; 27 November–1 December, 1995, vol. 1944 pp. 1942–1948