# Prediction of Wood Moisture Content Based on THz Time-Domain Spectroscopy

Ruifeng Duan,[a] Yuan Wang,[b,c,d,*] Lei Zhao,[b,c,d] Xing Da Yun,[b,c,d] and Nan Zhou [e]

A new method for predicting wood moisture content using terahertz (THz) time-domain spectroscopy (TDS) is presented in this paper. The THz wave is a promising method in measuring wood moisture content due to its sensitivity to water, impressive penetration ability in wood, and no destructive effect on wood interior. In this study, the selected wood, Douglas fir (*Pseudotsuga menziesii*), with different moisture content was studied. THz-TDS was used to extract the optical parameter of the sample. The THz refractive index and absorption coefficient spectrum of the wood were calculated. The first and second derivatives of the absorption coefficient spectrum were processed to obtain the first and second derivative spectra. The successive projections algorithm (SPA) was used to select the characteristic frequency for the THz absorption coefficient spectrum and its first and second derivative spectrum of the wood. A regression prediction model of wood moisture content was established by partial least squares regression (PLS). The results showed that the proposed model based on the second derivative spectrum had the best prediction effect for the moisture content of wood.

*Contact information: a: School of Information Science and Technology, Beijing Forestry University, Beijing 100083, P.R. China; b: School of Technology, Beijing Forestry University, Beijing 100083, P.R. China; c: Key Lab of State Forestry and Grassland Administration for Forestry Equipment and Automation, Beijing 100083, P.R. China; d: Joint International Research Institute of Wood Nondestructive Testing and Evaluation, Beijing Forestry University, Beijing 100083, P.R. China; e: State Key Laboratory of Precision Measurement Technology and Instrument, Tianjin University, Tianjin 300072, P.R. China;*
*\* Corresponding author: wangyuan@bjfu.edu.cn*

## INTRODUCTION

Wood drying is an essential procedure to improve wood quality and utilization rate. In this process, wood moisture content is a key parameter because one can choose a scientific drying technology. For instance, when the moisture content of wood is high, it is necessary to use lower temperature and higher humidity. When the moisture content is approximately 30%, the drying must be carried out slowly. Drying without knowing the moisture content of the wood may lead to cracking, warping, and wood defects. Therefore, it is necessary to detect wood moisture content accurately ahead of time.

Traditional methods of wood moisture content detection include the weighing method, conductivity method, capacitance method, and microwave method (Schajer *et al*. 2006; Zhou *et al*. 2011; Koumbi-Mounanga *et al*. 2015). The weighing method can obtain accurate measurement results. However, the testing process is time-consuming and greatly influenced by human factors. The resistance method determines the wood moisture content by measuring the resistance between two electrodes, which is based on the conductivity of

moisture in wood. When measuring the moisture content near the fiber saturation point, the measured value will suddenly deviate from the true value. The microwave method is used to measure the wood moisture content, and the relationship between the microwave and the wood moisture content can be determined. However, the measurement range is limited and affected by the temperature and humidity in the wood drying room, which leads to a large measurement error. Therefore, developing a non-contact and non-destructive detecting approach to obtain accurate wood moisture content during the drying process is an important issue that urgently needs to be solved.

Many researchers resort to spectral analysis methods due to their characteristics of rapidity and non-destructiveness. The nuclear magnetic resonance (NMR) (Almeida *et al*. 2007; Xu *et al*. 2017) method has high precision and speed, but it has a requirement on the size of the sample and is expensive. Near infrared (NIR) spectroscopy (Watanabe *et al.* 2011; Kobori *et al*. 2013) can also be utilized to measure wood water content rapidly and nondestructively, but the poor penetration makes it helpless for measuring the inside of the wood. X-rays (Tanaka *et al*. 2009; Kim *et al*. 2015) have outstanding penetration in wood, but they are harmful to the human body.

Terahertz (THz) spectroscopy has emerged as an important spectral analysis technique in recent years. The THz spectrum contains abundant physical and chemical information of the measured object. The weak intermolecular interactions, such as van der Waals force and hydrogen bond, as well as the molecular structure and related environmental information reflected by the skeleton vibration, dipole vibration, rotation transition of many macromolecules, and the low-frequency vibration of crystal lattice, have obvious response to the absorption intensity and the location of absorption peak in THz absorption spectrum (Zhang *et al*. 2007a). The THz time-domain spectroscopy (THz-TDS) spectrum can be used to calculate the optical parameters, such as the absorption coefficient, refractive index, and dielectric constant, of the sample. These parameters can be used to analyze the composition, structure, physical and chemical properties of the materials. Due to its strong correlation to water and good penetration in wood, THz wave is suitable for wood moisture content detection.

In this paper, THz-TDS technique was used to predict the water content of Douglas fir. The authors collected THz time-domain spectra at different moisture content of the wood and calculated its absorption coefficient and refractive index spectra. It was found that the absorption coefficient spectrum had a strong relationship with water content. Furthermore, the first and second derivatives of the absorption coefficient spectrum were processed. Then, the successive projection algorithm (SPA) was used to select the characteristic frequency and the partial least squares (PLS) was used to establish a regression prediction model. The results showed that the SPA-PLS model can accurately detect the moisture content of wood and the model based on the second derivative spectrum displayed the best prediction effect.

## ALGORITHM THEORY

### Successive Projections Algorithm

The SPA is a forward variable selection algorithm to minimize the collinearity of vector space, which was proposed by Bregman in 1965 (Almeida *et al*. 2018). The advantage of this method is that it can extract several characteristic wavelengths of the whole band, thus eliminating redundant information in the original spectral matrix. The

SPA is a cyclic variable selection method. The projection vectors of the selected frequencies on the unselected frequencies are computed cyclically, and the maximum frequency of the projection module is taken as the selected variable. Therefore, the correlation between the selected frequencies and the previous frequencies is the weakest. The SPA calculation is simple and fast. The method is generally used to select a characteristic wavelength of near infrared spectroscopy (ISO 3130 (1975)). In this study, the characteristic frequency of THz absorption coefficient spectrum and refractive index spectrum were screened by SPA.

Taking the spectral data set $X_{ixj}$ of j frequencies of i samples as an example, the steps of SPA are as follows:

(1) Before iteration, the $m$ column of spectral training set data is assigned to $x_{m(1)}$, $m \in 1$, $2, \ldots j$.

(2) Set the unselected frequencies as $S$ :

$$S = \left\{ m, 1 \leq m \leq p, p \notin \left[ m(1), m(2), \ldots m(i) \right] \right\} \tag{1}$$

(3) Calculate the projection of the selected frequency in the residual vector $x_k$, and select the maximum frequency of the projection vector. Cyclic calculation until all frequencies have been calculated:

$$m(i+1) = \arg \left[ \max \left( \left\| x_{m(i+1)} \right\| \right) \right], i \in S \tag{2}$$

(4) The multiple linear regression model is established by using the variables in the subset, and the subset with the minimum root mean square error is selected to carry out stepwise regression modeling, and the least characteristic variables are selected on the premise of guaranteeing accuracy.

## Partial Least Squares

The PLS algorithm is a regression modeling method of multiple dependent variables to multiple independent variables, which realizes the combination of multiple linear regression, principal component analysis, and canonical correlation analysis.

The principle of PLS regression modeling is as follows: The sets of dependent variables and independent variables are $Y = (y_1, y_2, \ldots y_q)$, $y_j \in R^n$, and $X = (x_1, x_2, \ldots x_p)$, $x_i \in R^n$. Then, PLS extracts component $t_1$ from $X$ and component $u_1$ from $Y$, where $t_1$ is the linear combination of $x_1, x_2, \ldots x_p$, $u_1$ is the linear combination of $y_1, y_2, \ldots y_q$. The extracted components should satisfy the following two requirements: Firstly, $t_1$ and $u_1$ should carry the variation information in their respective data tables to the greatest extent; Secondly, $t_1$ and $u_1$ have the greatest correlation. After the first components $t_1$ and $u_1$ are obtained, the regressions of $X$ to $t_1$ and $Y$ to $u_1$ are calculated. If the accuracy of the regression equation does not meet the requirements, the residual information is used to extract the components. The $m$ components $t_1, t_2, \ldots t_m$ extracted by $X$ are analyzed by $y_k (k = 1, 2 \ldots, q)$ regression method, and finally reduced to the equation of $y_k$ about $x_1, x_2, \ldots x_p$.

The PLS regression model needs principal component analysis. The determination of principal component has a great impact on the accuracy of the model. Predictive residual sum of squares ($S_{spre}$) and cross validation are usually used to find the optimal number of principal components. The smaller the $S_{spre}$ value, the higher the prediction accuracy of the model is. The $S_{spre}$ can be calculated by:

$$S_{\mathrm{spre}} = \sum (y_i - \hat{y}_i)^2 \tag{3}$$

where $y_i$ is the standard data and $\hat{y}_i$ is the predicted data.

## Model Evaluation Index

The accuracy of the regression prediction model needs to be judged after it is established. The accuracy of the model is mainly evaluated through fitting degree, correlation coefficient, and root mean square error (RMSE) in this paper.

Fitting degree is expressed by $R^2$, whose range is [0, 1]. The closer its value is to 1, the better the fitting degree of the model is. Fitting degree is given as follows,

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (\overline{y}_i - y_i)^2} \tag{4}$$

where $n$ is the number of sample, $y_i$ is the standard value, $\hat{y}_i$ is the predicted value, and $\overline{y}_i$ is the average of the standard value.

Variable $r(y_i, \hat{y}_i)$ stands for the correlation coefficient. If $r(y_i, \hat{y}_i) > 0$, it indicates that the two variables are positively correlated. In contrast, if $r(y_i, \hat{y}_i) < 0$ it means that they are negatively correlated. The closer $|r(y_i, \hat{y}_i)|$ approaches 1, the closer the linear relationship between the two variables is. The closer $|r(y_i, \hat{y}_i)|$ approaches 0, the weaker the linear correlation between the two variables is. Correlation coefficient is defined as Eq. 5,

$$r(y_i, \hat{y}_i) = \frac{\mathrm{Cov}(y_i, \hat{y}_i)}{\sqrt{\mathrm{Var}[y_i]\mathrm{Var}[\hat{y}_i]}} \tag{5}$$

where $\mathrm{Cov}(y_i, \hat{y}_i)$ is the covariance between $y_i$ and $\hat{y}_i$, and $\mathrm{Var}[y_i]$ and $\mathrm{Var}[\hat{y}_i]$ are the variances of $y_i$ and $\hat{y}_i$, respectively.

The *RMSE* is used to measure the deviation between the observed value and the true value. The smaller the *RMSE* of the samples in the prediction set, the better the fitting and prediction effect of the model is. The *RMSE* is expressed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)} \tag{6}$$

## EXPERIMENTAL

Douglas fir was selected to be the experimental material. The authors made 5 pieces of wood samples. The samples were cut transversely and the sizes of each sample were 50 mm × 30 mm × 5 mm.

## Standard Moisture Content Determination Method

According to the method for determination of the density of wood (ISO 3131 (1975)), the wood samples were placed in a drying box and dried at $103 \pm 2$ °C for 8 h (Zhao *et al*. 2018). Then, an electronic balance with accuracy of 0.001 g was used to weigh the quality. During the drying course, wood samples were weighed every 2 h and THz time-domain spectrum was acquired after each weighing. The test ended when the change

of sample weight was less than the sample mass, which means the wood was absolutely dried. The process is shown in Fig. 1.
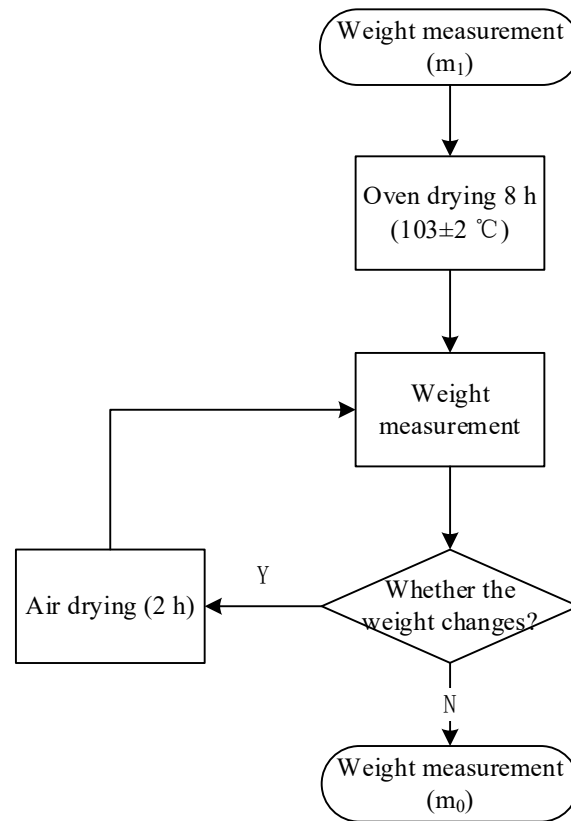
```
        ┌─────────────────────┐
        │  Weight measurement │
        │        (m₁)         │
        └─────────────────────┘
                  │
                  ▼
        ┌─────────────────────┐
        │   Oven drying 8 h   │
        │    (103±2 ℃)        │
        └─────────────────────┘
                  │
                  ▼
        ┌─────────────────────┐
   ┌───▶│       Weight        │
   │    │    measurement      │
   │    └─────────────────────┘
   │              │
   │              ▼
┌──────────┐  ◇ Whether the ◇
│Air drying│◀─◇ weight changes?◇  Y
│  (2 h)   │   ◇            ◇
└──────────┘      │
                  │ N
                  ▼
        ┌─────────────────────┐
        │  Weight measurement │
        │        (m₀)         │
        └─────────────────────┘
```

**Fig. 1.** Wood moisture content measurement process

The moisture content of wood was calculated according to the following Eq. 7,

$$W = \frac{m_1 - m_0}{m_0} \times 100 \ \%$$  (7)

where $m_1$ is the sample weight (g) in the experiment and $m_0$ is the sample weight (g) when absolutely dried.

## THz Time-Domain Spectrum Acquisition

The THz-TDS is a coherent measurement technology of wideband THz pulse, which can detect the amplitude and phase information of THz pulse at the same time, and the spectral information can be achieved by Fourier transform infrared spectroscopy (Gribenyukov *et al*. 2018). It is convenient to extract the refractive index, absorption coefficient, extinction coefficient, and other optical parameters of the sample from the THz-TDS spectrum. Then, some physical properties and chemical information of samples can be obtained by analyzing these optical parameters.

According to the collection mode of THz spectrum, the THz-TDS system can be divided into transmission type and reflection type. In practical operation, the appropriate THz spectrum measurement method should be varied according to different samples and experimental conditions. When the tested samples are thin or the absorption intensity to

THz pulses is low, the THz time-domain spectra of samples can be better obtained by transmission type. The THz time-domain spectroscopy system used in this paper is the transmission type.

In this paper, the THz spectra of wood samples were collected by the THz-TDS equipment (MenloSystems, Martinsried, Germany) in the State Key Laboratory of Precision Measurement Technology and Instruments of Tianjin University. The operating wavelength range of the equipment is 780 to 1650 nm, pulse width is 90 fs, total average output power is 500 mw, and repetition rate is 100 mHz. The block diagram of the THz-TDS system is shown in Fig. 2.
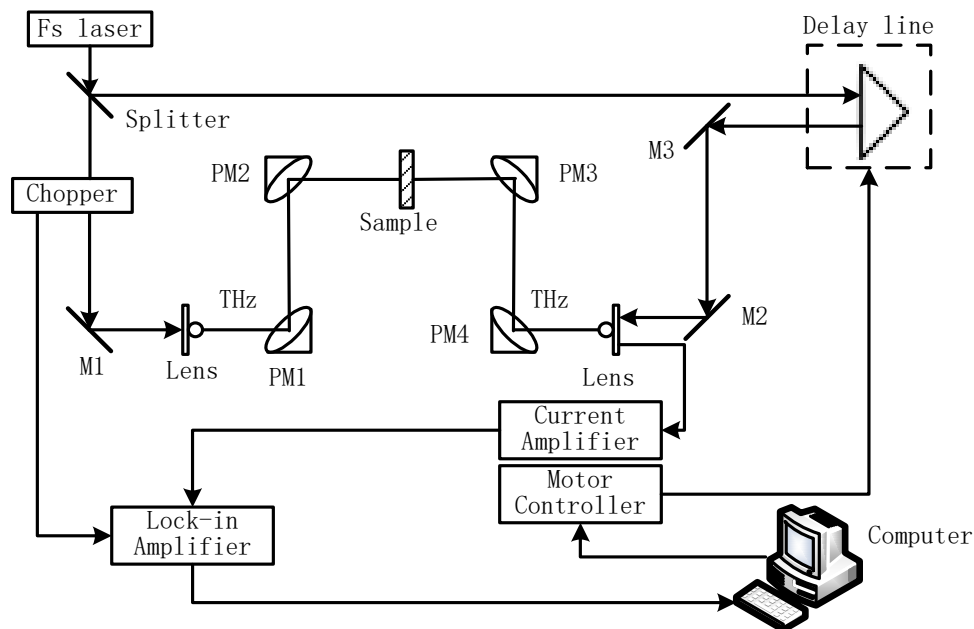


**Fig. 2.** THz-TDS system diagram

The wood samples were placed at the communal focus of parabolic mirror PM2 and PM3 of THz-TDS system, and the THz wave was perpendicular to the wood texture. The transmission spectra of samples can be obtained by a scanning operation. A set of reference spectra with no samples were measured before the scanning operation. The frequency-domain spectra of the samples and reference signals were obtained after fast Fourier transform (FFT) was applied to the time-domain signal (TERA K15; MenloSystems, Martinsried, Germany).

## RESULTS AND DISCUSSION

### THz Spectral Analysis

The original THz time-domain spectra of wood were obtained by transmission of THz-TDS. In order to reduce the influence of ambient noise, each sample was measured three times and the average value was taken as the detection result. A total of 200 groups of wood THz spectra below fiber saturation point were obtained in the test, including moisture content ranged from 1.29% to 35.49%.

Five kinds of wood THz spectra with different moisture contents were selected to display how the spectral waveform change with moisture content, which is shown in Fig. 3. Compared with the reference spectra, the THz time-domain spectra of wood samples had a certain delay in time and an attenuation in amplitude. The time delay increased along with the moisture content of wood samples, while the amplitude decreased with it. In order to eliminate the measurement errors caused by factors such as equipment accuracy, water vapor in air, and the oscillation caused by reflection and refraction of THz wave, the wood THz time-domain spectral data were intercepted by Windowing.
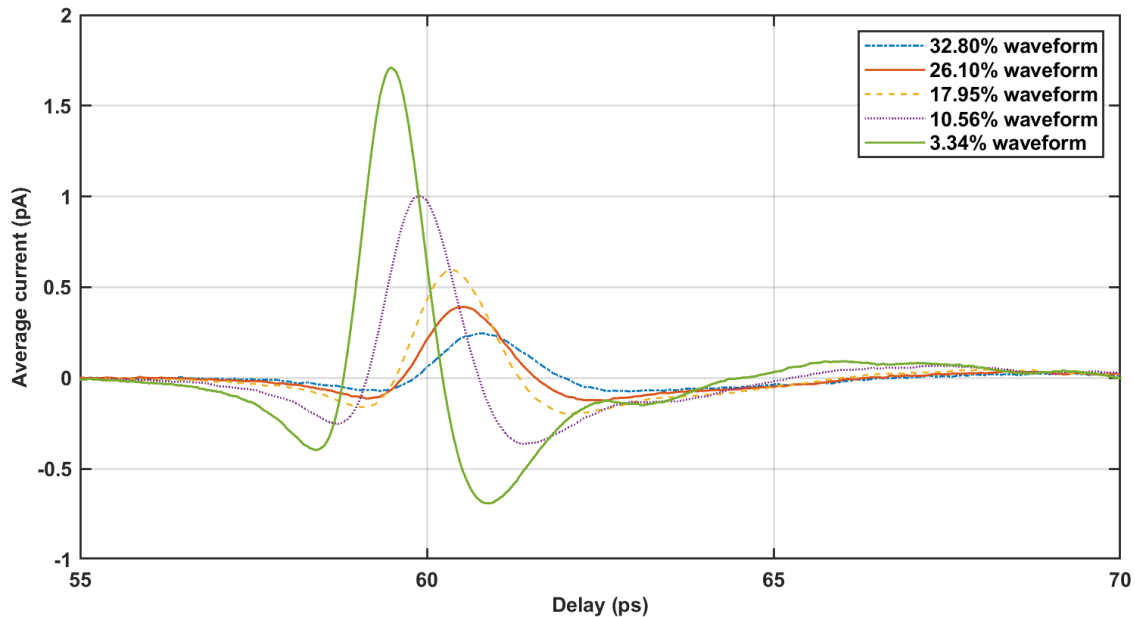


**Fig. 3.** THz time-domain waveform of wood samples with different moisture content
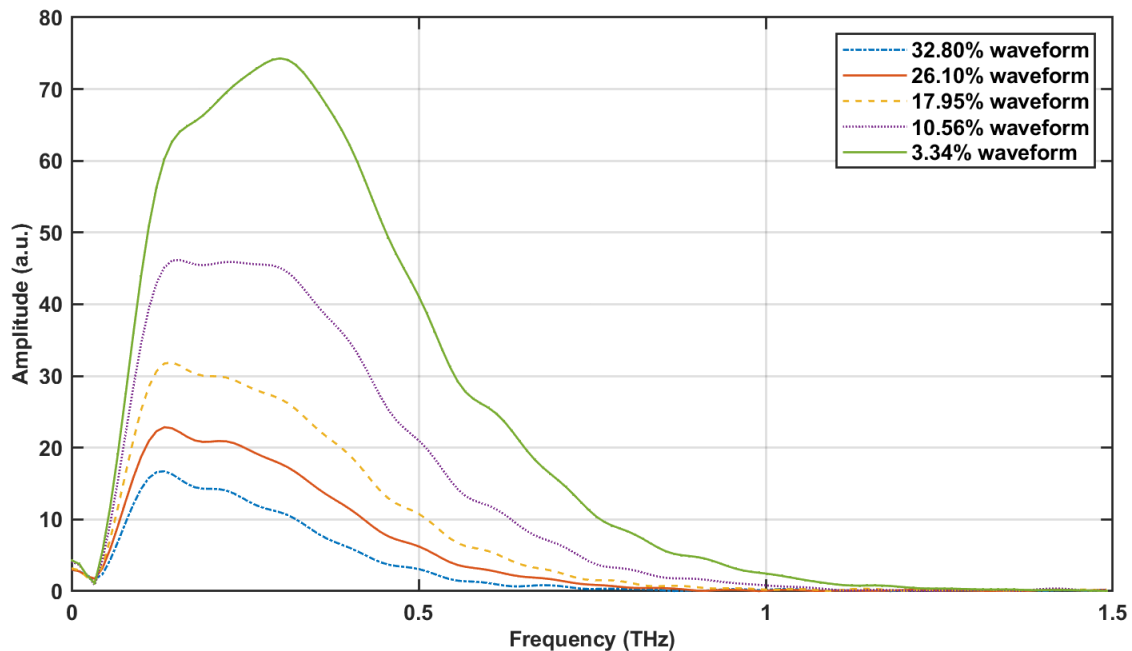


**Fig. 4.** THz frequency-domain waveform of wood samples with different moisture content

The THz frequency-domain spectra of wood samples can be obtained by performing FFT to the time-domain spectra, which is shown in Fig. 4. It can be seen that the trend of the spectral waveform of wood samples with different moisture content was similar. However, the amplitudes were quite different. With the increase of moisture content of wood samples, the THz spectrum amplitude of wood samples decreased.

In conclusion, the THz time-domain waveform of wood samples with different moisture content showed different time delay and amplitude attenuation. The amplitude was negatively correlated with the moisture content of wood samples, while the time delay was positively correlated with it. In frequency-domain, the THz waveform of wood samples further showed the amplitude-frequency characteristics with different moisture content. The amplitude was negatively correlated with the moisture content of wood samples. It showed that it was feasible to use THz wave to predict wood moisture content.

## THz Optical Parameter Extraction

According to the optical constants extracting model based on THz time-domain spectroscopy proposed by Dorney and Duvillaret, the absorption coefficient and refractive index can be calculated based on THz frequency (Li *et al*. 2014).

Macroscopic optical properties of samples can be described by complex refractive index,

$$\tilde{n} = n - ik \tag{8}$$

where $\tilde{n}$ is complex refractive index and $n$ is real refractive index, which is used to describe the dispersion characteristics of samples, and $k$ is extinction coefficient, which is used to describe the absorption characteristics of samples.

The refractive index of THz spectrum can be expressed as follows,

$$n(\omega) = \frac{\varphi(\omega)c}{\omega d} + 1 \tag{9}$$

and the expression of the absorbance is shown below in Eq. 10,

$$\alpha(\omega) = \frac{2}{d} \ln \frac{4n(\omega)}{\rho(\omega)(n(\omega)+1)^2} \tag{10}$$

where $n(\omega)$ is the real part of refractive index, $d$ denotes the sample thickness (m), $c$ is the velocity (m/s) of the THz wave propagating in a vacuum, $\omega$ is angular frequency, and $\rho(\omega)$ and $\varphi(\omega)$ represent the amplitude ratio and phase difference between sample signal and reference signal, respectively.

The absorption coefficient and refractive index spectrum are shown in Fig. 5 and Fig. 6. It can be seen from Fig. 5 that THz absorption coefficient spectrum was greatly influenced by water content. Though the waveform trends and the position of absorption peaks were basically the same, there were quite differences in the absorption intensity. The THz absorption intensity of wood enhanced with the increase of moisture content, indicating that the THz absorption intensity of wood was positively correlated with the moisture content of wood. In Fig. 6, the refractive index spectra of different water content seemed similar in appearance, no clear pattern was found. Therefore, the authors chose the THz absorption coefficient spectrum to predict the moisture content of wood in this paper.
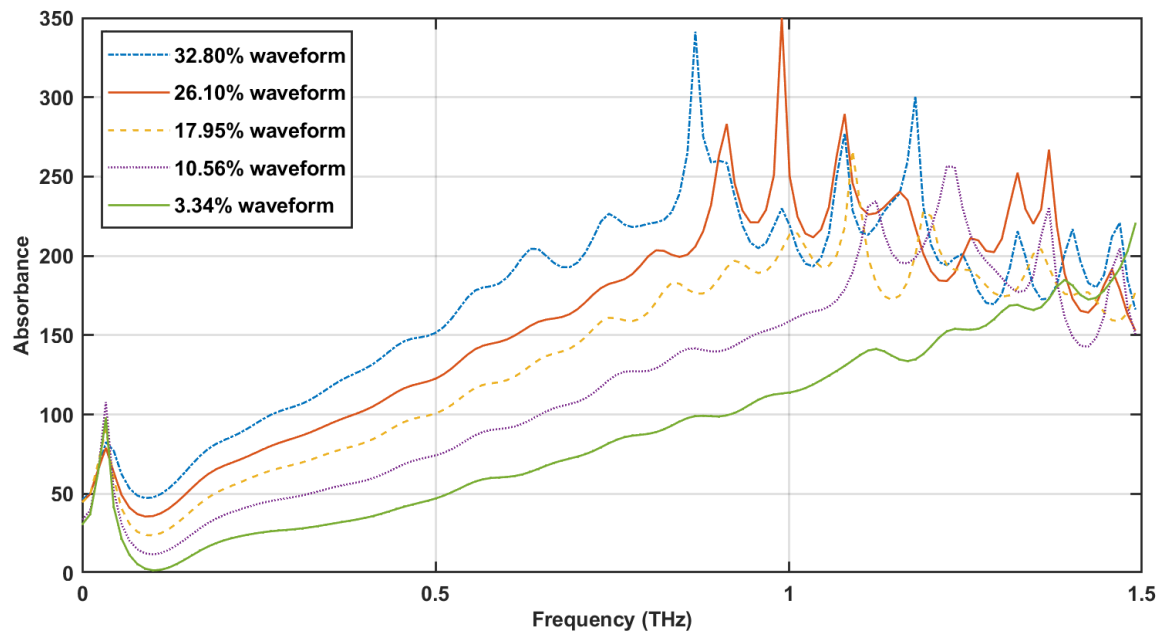
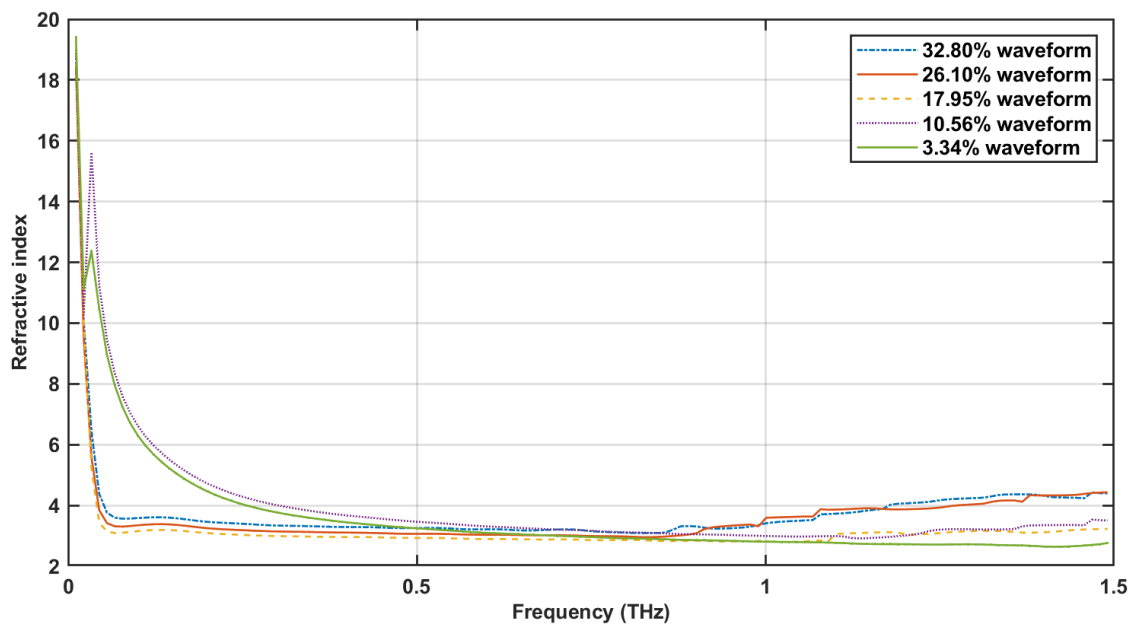**Fig. 5.** Waveform of THz absorption coefficient of wood samples at different moisture contents



**Fig. 6.** THz refractive index waveform of wood samples with different moisture content

### Spectral Derivative Processing

To eliminate the baseline drift of the spectrum, enhance the spectral characteristics, and reduce the noise influence of the instrument itself, the first derivative and the second derivative of the absorption coefficient spectrum were processed to obtain the first derivative spectrum and the second derivative spectrum, as shown in Figs. 7 and 8.
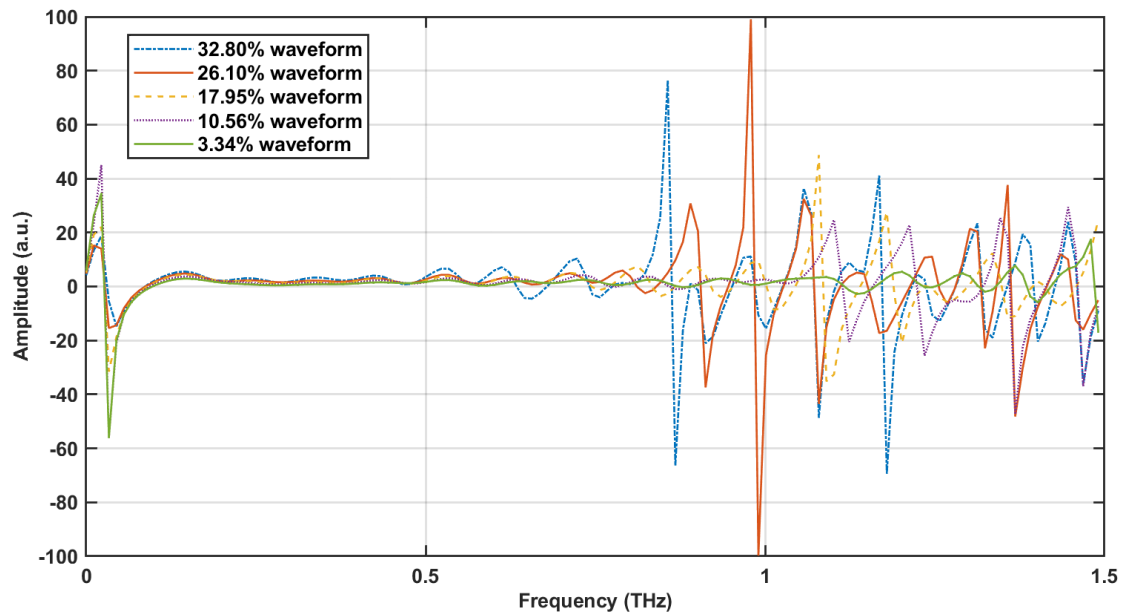
**Fig. 7.** First derivative waveform
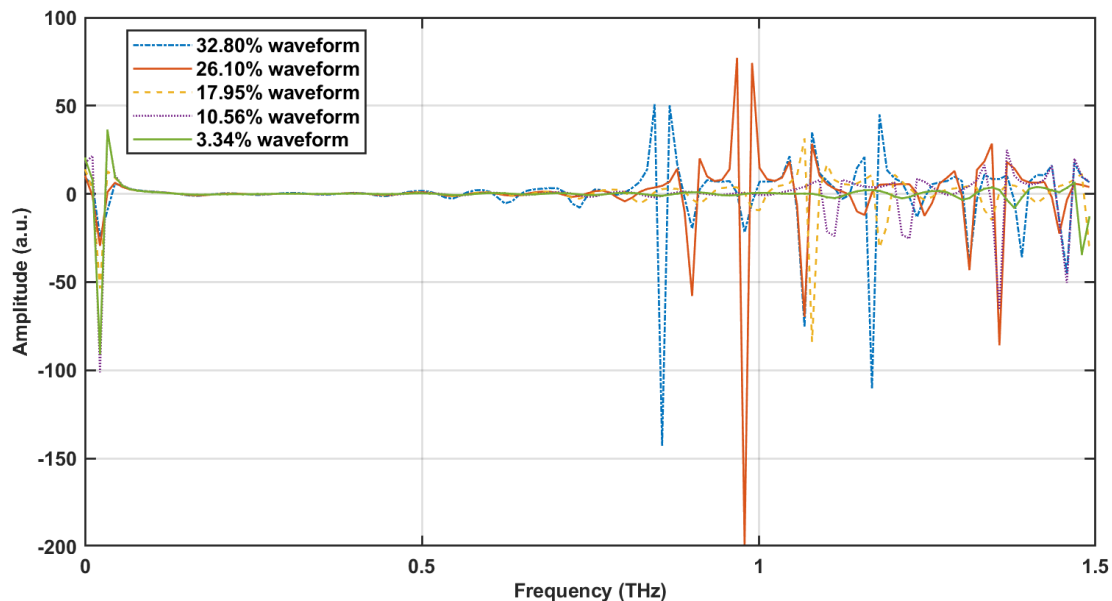


**Fig. 8.** Second derivative waveform

## Spectral Characteristic Selection

The characteristic frequencies of THz absorption coefficient spectrum, first derivative spectrum, and second derivative spectrum were selected by SPA. As shown in Fig. 9, seven characteristic frequency points were selected from the 135 frequency points of absorption coefficient spectrum, accounting for 5.18% of all frequency bands. Figure 10 shows the variation of RMSE value with the increase of the number of variables. It can be seen that the RMSE decreased with the increase of variables at the beginning. When the number was 7, the RMSE reduced to 0.02421 and tended to be stable.
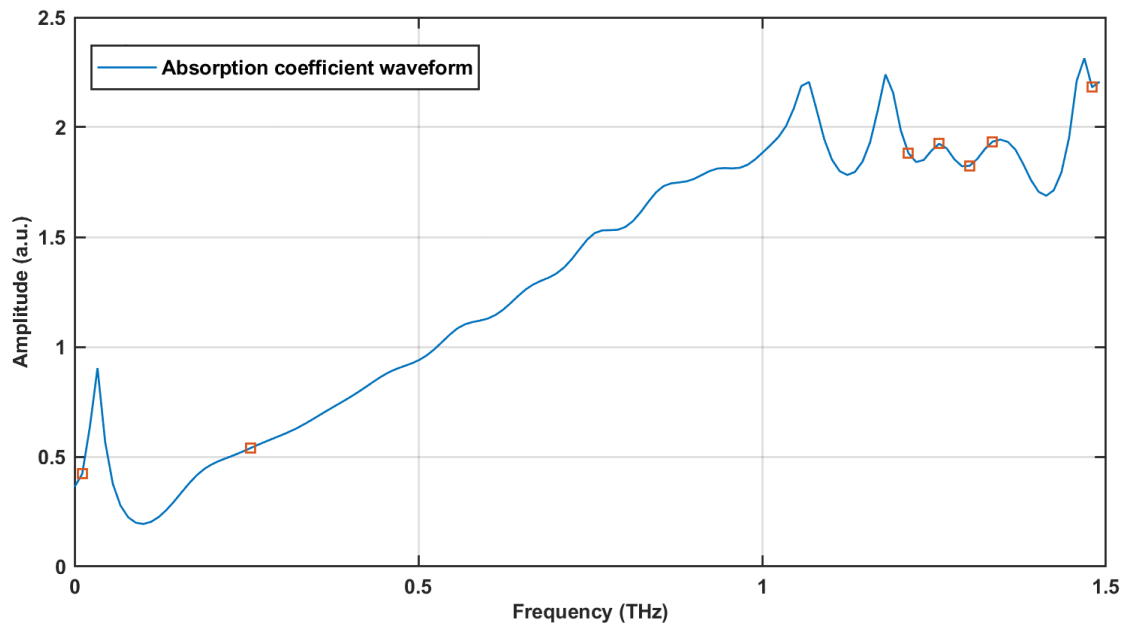
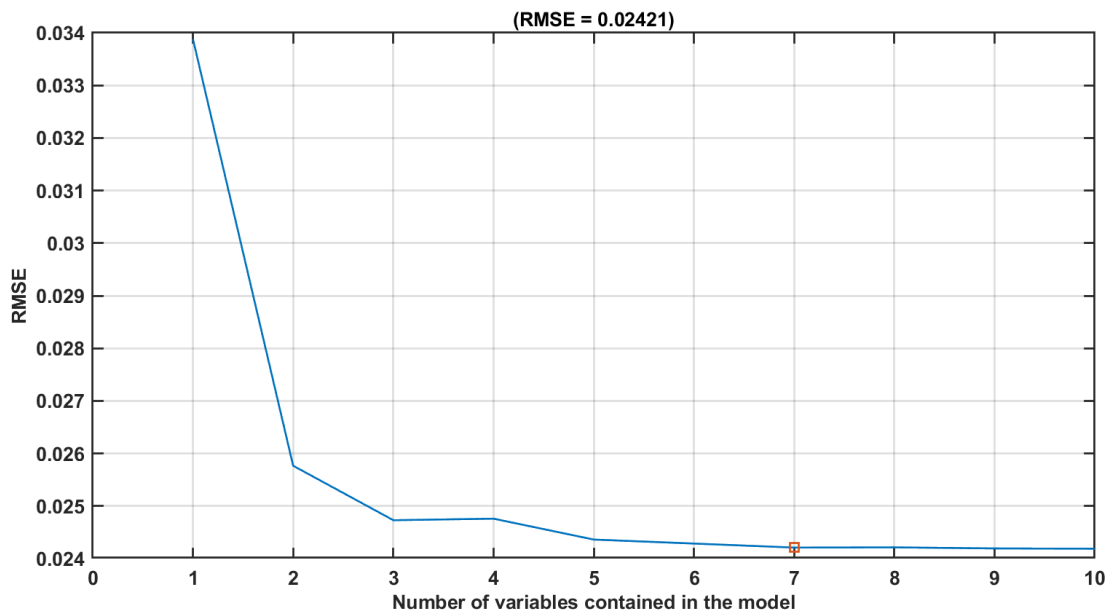**Fig. 9.** Characteristic frequencies selection of absorption coefficient spectrum



**Fig. 10.** Variation of RMSE with the number of variables in absorption coefficient spectrum

The SPA screened nine characteristic frequency points from a total 135 frequency points of first derivative spectrum, accounting for 6.67% of all frequency bands, which is shown in Fig. 11. The RMSE was 0.0362, and the RMSE variation with the number of characteristic variables is shown in Fig. 12.
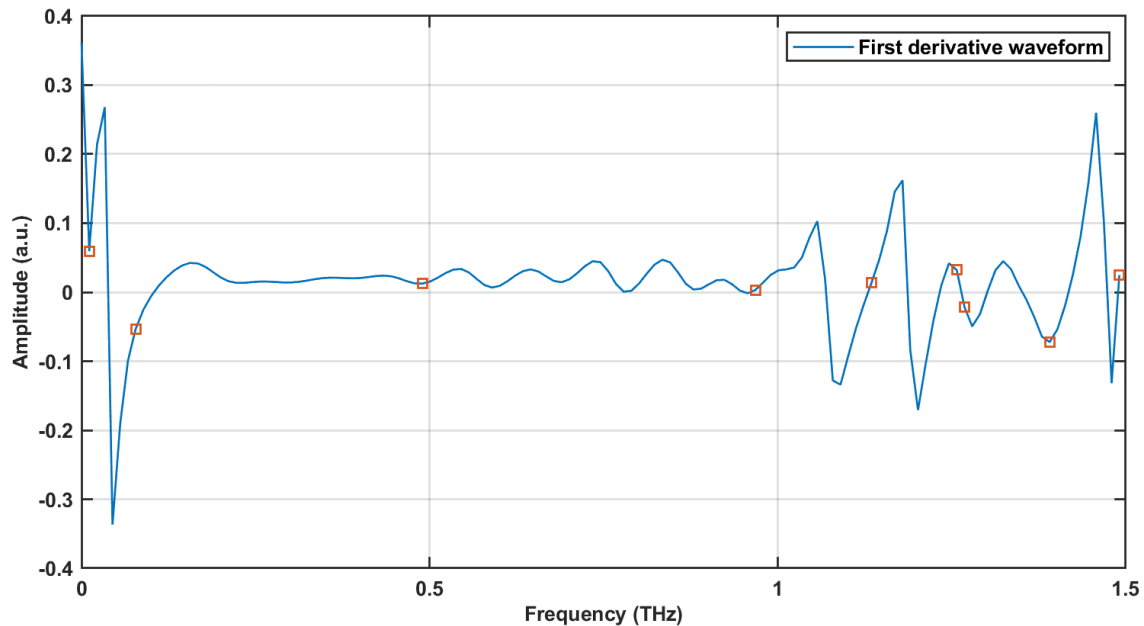
**Fig. 11.** Characteristic frequency selection of first derivative spectrum
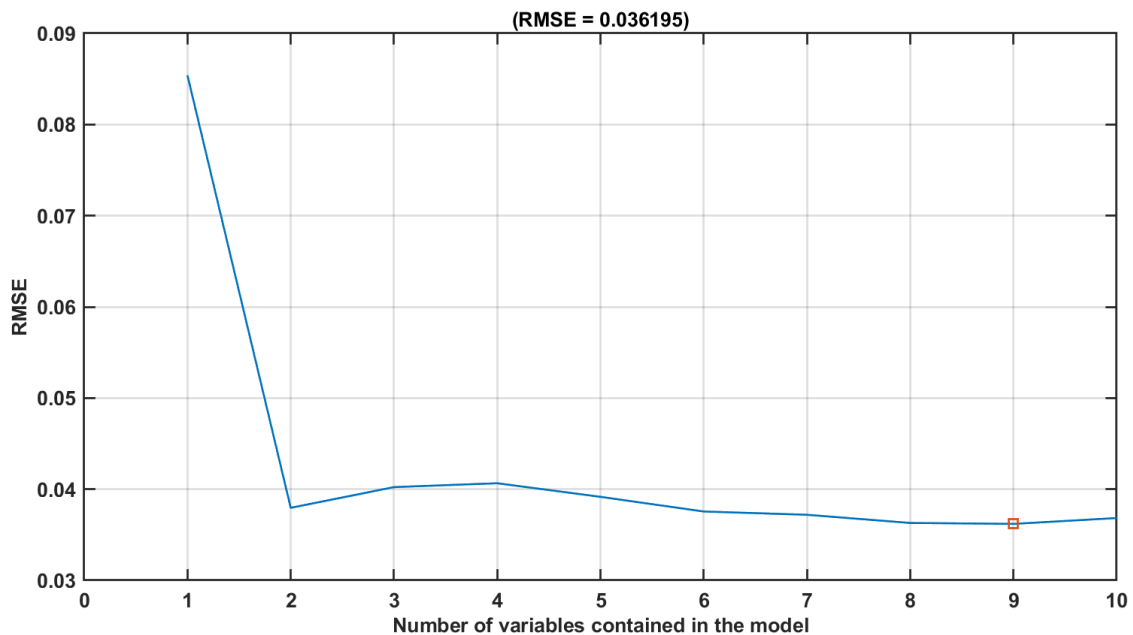


**Fig. 12.** Variation of RMSE with the number of variables in first derivative spectrum

As shown in Fig. 13, approximately 16 characteristic frequency points were screened from 135 frequency points of second derivative spectrum, accounting for 11.85% of all frequency bands. The RMSE was 0.0364, and the RMSE variation with the number of characteristic variables is shown in Fig. 14.
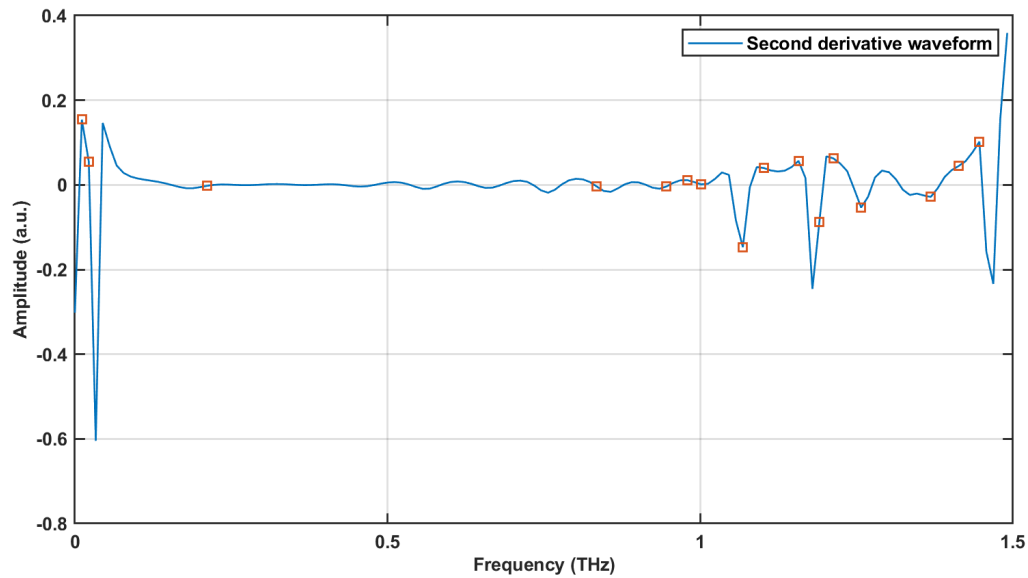
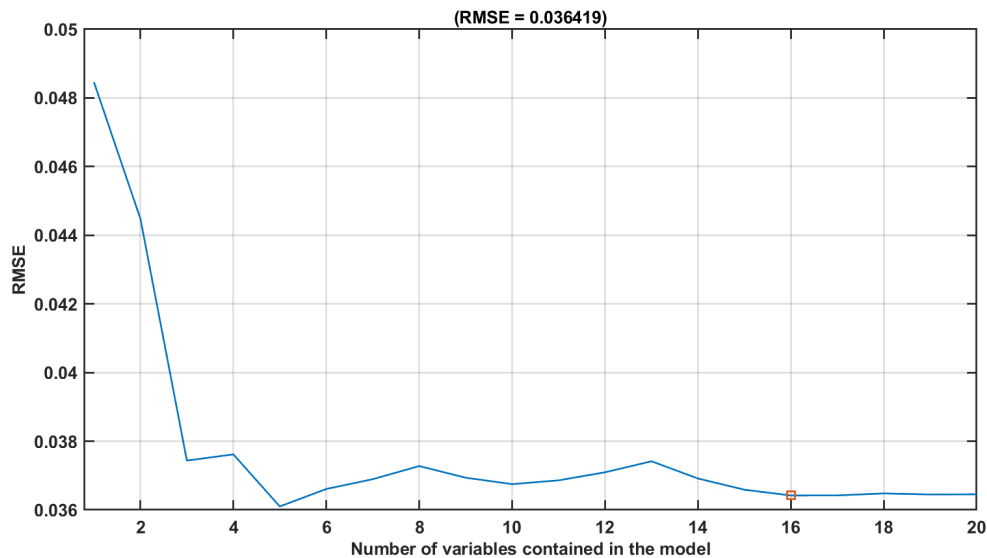**Fig. 13.** Characteristic frequency selection of second derivative spectrum



**Fig. 14.** Variation of RMSE with the number of variables in second derivative spectrum

## PLS Modeling and Evaluation

The authors respectively divided 200 sets of THz absorption coefficient spectra, first derivative spectra, and second derivative spectra into two groups, 150 groups as training set and 50 groups as test set. The PLS regression prediction model was established using training set, then the data of test set are predicted.

The PLS prediction model of wood moisture content was established based on THz absorption coefficient spectrum, first derivative spectrum, and second derivative spectrum. Figures 15, 16, and 17 show the scatter plots of PLS model based on training set and test set of three kinds of spectra, respectively.
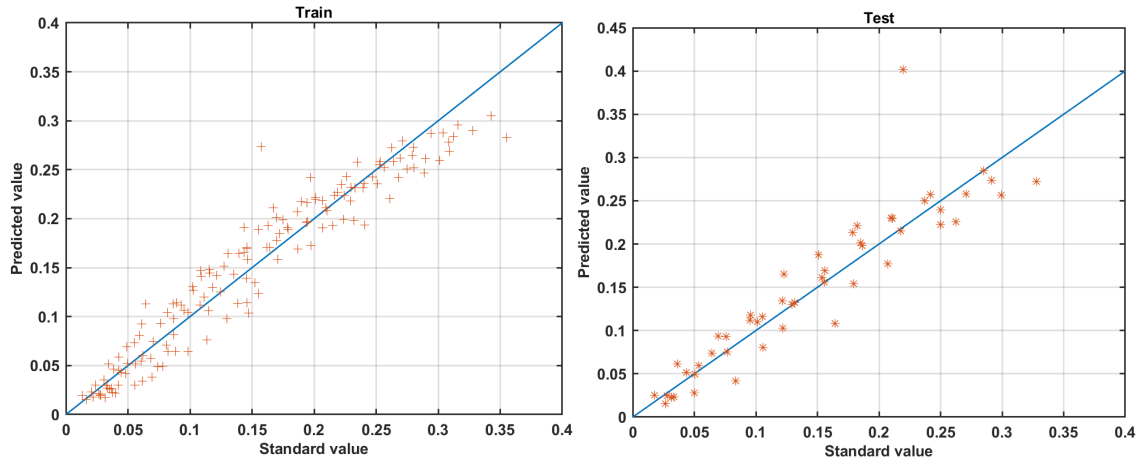
**Fig. 15.** Scattered plot of absorption coefficient SPA-PLS prediction results based on training set and test set
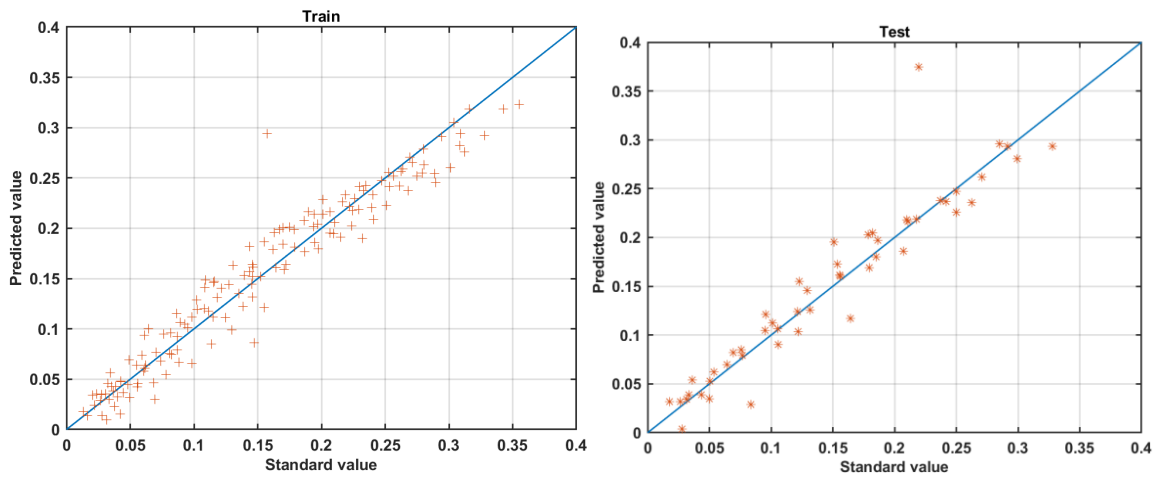


**Fig. 16.** Scattered plot of first derivative SPA-PLS prediction results based on training set and test set
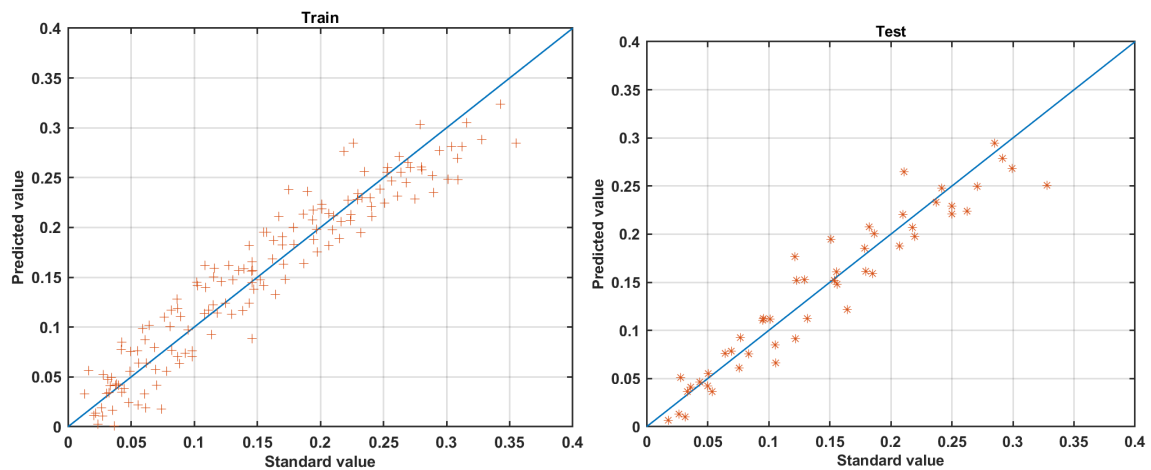


**Fig. 17.** Scattered plot of second derivative SPA-PLS prediction results based on training set and test set

By analyzing the fitting results of scatter plots, the SPA-PLS prediction model based on the second derivative spectrum had the best fitting effect on the training set, while the absorption coefficient spectrum had the worst performance. The results showed that the derivative processing had an obvious effect on reducing noise and enhancing spectral characteristics, thus the prediction accuracy of the model was improved.

The fitting degree, correlation coefficient, and RMSE were used to evaluate the model, which is shown in Tables 1, 2, and 3. The results of training set and test set of PLS model were compared in different frequency bands of 0.2 to 1 THz and the feature frequencies screened by SPA.

**Table 1.** PLS Prediction Model Based on Absorption Coefficient

| Frequency (THz) | Number of Principal Components | Training | | | Testing | | |
|---|---|---|---|---|---|---|---|
| | | $R^2$ | Correlation Coefficient | RMSE | $R^2$ | Correlation Coefficient | RMSE |
| 0.2 to 0.4 | 2 | 0.8871 | 0.9419 | 0.0296 | 0.9148 | 0.9374 | 0.0170 |
| 0.4 to 0.6 | 3 | 0.8877 | 0.9422 | 0.0295 | 0.9011 | 0.9398 | 0.0166 |
| 0.6 to 0.8 | 3 | 0.8355 | 0.9141 | 0.0358 | 0.7952 | 0.8496 | 0.0259 |
| 0.8 to 1.0 | 3 | 0.7965 | 0.8925 | 0.0398 | 0.8007 | 0.7902 | 0.0305 |
| SPA | 2 | 0.9245 | 0.9615 | 0.0242 | 0.9221 | 0.9613 | 0.0133 |

**Table 2.** PLS Prediction Model Based on First Derivative

| Frequency (THz) | Number of Principal Components | Training | | | Testing | | |
|---|---|---|---|---|---|---|---|
| | | $R^2$ | Correlation Coefficient | RMSE | $R^2$ | Correlation Coefficient | RMSE |
| 0.2 to 0.4 | 3 | 0.8294 | 0.9107 | 0.0364 | 0.8599 | 0.8798 | 0.0236 |
| 0.4 to 0.6 | 2 | 0.7605 | 0.8721 | 0.0432 | 0.8592 | 0.7736 | 0.0324 |
| 0.6 to 0.8 | 3 | 0.7011 | 0.8373 | 0.0482 | 0.7862 | 0.8255 | 0.0284 |
| 0.8 to 1.0 | 3 | 0.6111 | 0.7817 | 0.0550 | 0.3621 | 0.7378 | 0.0344 |
| SPA | 2 | 0.9332 | 0.9660 | 0.0228 | 0.9527 | 0.9751 | 0.0107 |

By comparing and analyzing Tables 1, 2, and 3, the fitting and prediction results of the three spectra in PLS screening band were better than those in 0.2 to 1 THz band. The principal component of SPA-PLS model based on absorption coefficient spectrum was 2, the fitting degree of training set was 0.9245, the correlation coefficient was 0.9615, the RMSE was 0.0242. In the test set, the fitting degree was 0.9221, the correlation coefficient was 0.9613, and the RMSE was 0.0133.

The principal component of SPA-PLS model based on first derivative spectrum was 2, the fitting degree of training set was 0.9332, the correlation coefficient was 0.9660, the

RMSE was 0.0228. In the test set, the fitting degree was 0.9527, the correlation coefficient was 0.9751, and the RMSE was 0.0107.

**Table 3.** PLS Prediction Model Based on Second Derivative

| Frequency (THz) | Number of Principal Components | Training | | | Testing | | |
|---|---|---|---|---|---|---|---|
| | | $R^2$ | Correlation Coefficient | RMSE | $R^2$ | Correlation Coefficient | RMSE |
| 0.2 to 0.4 | 2 | 0.8406 | 0.9168 | 0.0352 | 0.8809 | 0.8703 | 0.0262 |
| 0.4 to 0.6 | 3 | 0.8575 | 0.8986 | 0.0387 | 0.9020 | 0.9009 | 0.0216 |
| 0.6 to 0.8 | 3 | 0.7260 | 0.8520 | 0.0462 | 0.8619 | 0.8241 | 0.0281 |
| 0.8 to 1.0 | 3 | 0.6307 | 0.7942 | 0.0536 | 0.4532 | 0.5186 | 0.0432 |
| SPA | 4 | 0.9636 | 0.9816 | 0.0168 | 0.9921 | 0.9587 | 0.0143 |

The principal component of SPA-PLS model based on second derivative spectrum was 4, the fitting degree of training set was 0.9636, the correlation coefficient was 0.9816, the RMSE was 0.0168. In the test set, the fitting degree was 0.9921, the correlation coefficient was 0.9587, and the RMSE was 0.0143.

By comparing the optimal models under three data sets, it was found that the prediction effect of SPA-PLS model based on second derivative spectrum was slightly better than that of SPA-PLS model based on first derivative spectrum, and it was better than that of SPA-PLS model based on absorption coefficient spectrum. The results showed that the selection of characteristic frequency points by SPA can effectively improve the model fitting prediction effect, and the derivative processing can enhance the characteristics of THz spectrum and reduce the impact of noise. Therefore, the SPA-PLS wood THz moisture prediction model based on second derivative spectrum had the best fitting effect and the highest prediction accuracy.

## CONCLUSIONS

1. In this paper, terahertz time-domain spectroscopy (THz-TDS) was used to predict the moisture content of Douglas fir. THz time-domain and frequency-domain spectral data of Douglas fir under different moisture content were obtained, and it can be seen that the moisture content had an obvious effect on THz spectra. The THz wave was sensitive to the moisture content of Douglas fir.

2. For further analysis, the optical parameters were extracted from the THz spectrum. Different moisture content of Douglas fir was found to have a significant effect on terahertz absorption intensity, and there were large differences in absorption intensity. In summary, the terahertz absorption intensity increases with increasing moisture content, and there is a positive correlation between the terahertz absorption intensity and the moisture content of Douglas fir.

3.  The THz absorption coefficient spectra of Douglas fir samples with different moisture content were processed by first derivative and second derivative. The PLS regression prediction models were established based on absorption coefficient spectra, first derivative spectra, and second derivative spectra for prediction. The predictive ability of the model was compared by using fitness, correlation coefficient, and root mean square error. The results showed that the SPA-PLS model based on the second derivative spectrum had the best prediction effect on the moisture content of Douglas fir. Therefore, THz-TDS had a good performance in the prediction of moisture content of Douglas fir.

## ACKNOWLEDGMENTS

## REFERENCES CITED

Almeida, G., Gagné, S., and Hernández, R. E. (2007). "A NMR study of water distribution in hardwoods at several equilibrium moisture contents," *Wood Science and Technology* 41(4), 293-307. DOI: 10.1007/s00226-006-0116-3

Almeida, G., Chen, X., and Li, C. (2018). "Detection of nitrogen content in rubber leaves using near-infrared (NIR) spectroscopy with correlation based successive projections algorithm (SPA)," *Applied Spectroscopy* 72(5), 740-749. DOI: 10.1177/0003702818755142

Gribenyukov, A. I., Dorozhkin, K. V., Morozov, A. N., and Suslyaev, V. I. (2018). "Influence of after-growth treatments on the optical parameters of teraherz $ZnGeP_2$, crystals," *Russian Physics Journal* 60(11), 2000-2003. DOI: 10.1007/s11182-018-1314-9

ISO 3130 (1975). "Wood-determination of moisture content for physical and mechanical tests," International Organization for Standardization, Geneva, Switzerland.

ISO 3131 (1975). "Wood-determination density for physical and mechanical tests," International Organization for Standardization, Geneva, Switzerland.

Kim, C. K., Oh, J. K., Hong, J. P., and Lee, J. J. (2015). "Dual-energy X-ray absorptiometry with digital radiograph for evaluating moisture content of green wood," *Wood Science and Technology* 49(4), 713-723. DOI: 10.1007/s00226-015-0726-8

Kobori, H., Gorretta, N., Rabatel, G., Rabatel, G., BEllon-Maurel, V. R., and Tsuchikawa, S. (2013). "Applicability of Vis-NIR hyperspectral imaging for monitoring wood moisture content (MC)," *Holzforschung* 67(3), 307-314. DOI: 10.1515/hf-2012-0054

Koumbi-Mounanga, T., Groves, K., Leblon, B., Gao, Z., and Cooper, P. A. (2015). "Estimation of moisture content of trembling aspen (*Populus tremuloides* Michx.) strands by near infrared spectroscopy (NIRS)," *European Journal of Wood and Wood Products* 73(1), 43-50. DOI: 10.1007/s00107-014-0856-y

Li, W. X., Li, F., Zhao, G. L., Tang, S. J., and Liu, X. Y. (2014). "Non-destructive and fast identification of cotton-polyester blend fabrics by the portable near-infrared spectrometer," *Spectroscopy & Spectral Analysis* 34(12), 3246-3252. DOI: 10.3964/j.issn.1000-0593(2014)12-3246-07

Schajer, G. S., and Orhan, F. B. (2006). "Measurement of wood grain angle, moisture content and density using microwaves," *Holz als Roh- und Werkstoff* 64(6), 483-490. DOI: 10.1007/s00107-006-0109-9

Tanaka, T., Avramidis, S., and Shida, S. (2009). "Evaluation of moisture content distribution in wood by soft X-ray imaging," *Journal of Wood Science* 55(1), 69-73. DOI: 10.1007/s10086-008-0997-x

Tang, R. N, Chen, X. P, and Li, C. (2018). "Detection of nitrogen content in rubber leaves using near-infrared (NIR) spectroscopy with correlation based successive projections algorithm (SPA)," *Applied Spectroscopy* 72(5), 740-749. DOI: 10.1177/0003702818755142

Watanabe, K., Mansfield, S. D., and Avramidis, S. (2011). "Application of near-infrared spectroscopy for moisture-based sorting of green hem-fir timber," *Journal of Wood Science* 57(4), 288-294. DOI: 10.1007/s10086-011-1181-2

Xu, K., Lv, J. X., Gao, Y. L., Wu, Y. Q., and Li, X. J. (2017). "Determination of moisture content and moisture content profiles in wood during drying by low-field nuclear magnetic resonance," *Drying Technology* 35(15), 1909-1918. DOI: 10.1080/07373937.2017.1291519

Zhang, C., Jiang, X., Mu, K., Zhang, X. C., Zhang, L., Zhou, Q., Zhang, Y., Shen, Y., Zhao, G., and Jiao, Y. (2007). "Identification of explosives and drugs and inspection of material defects with THz radiation," *Proceedings of SPIE Terahertz Photonics* 6840, Article ID 68400S. DOI: 10.1117/12.760133

Zhao, J., Zhang, X., Li, S., Liu, C., Chen, Y., Peng, Y., and Zhu, Y. (2018). "Detecting the propagation effect of terahertz wave inside the two-color femtosecond laser filament in the air," *Applied Physics B* 124(3), article no. 45. DOI: 10.1007/s00340-018-6913-1

Zhou, Q. F., Cai, Y. C., Xu, Y., and Zhang, X. L. (2011). "Preliminary exploration of on-line measuring layer moisture content with electrical resistance method during wood drying," *Advanced Materials Research* 261-263, 480-484. DOI: 10.4028/www.scientific.net/AMR.261-263.480