American Journal of Biochemistry and Biotechnology

Original Research Paper

Near Infrared Spectroscopy Identification Method for Five Chemical Types of *Cinnamomum camphora*

Bailian Tu, Yueting Zhang, Yongjie Zheng, Xinliang Liu and Yanfang Wu

Camphor Engineering and Technology Research Center of National Forestry and Grassland Administration, Jiangxi Academy of Forestry, Nanchang, China

Article history
Received: 19-07-2024
Revised: 09-09-2024
Accepted: 16-09-2024

Corresponding Author:
Yanfang Wu
Camphor Engineering and Technology Research Center of National Forestry and Grassland Administration, Jiangxi Academy of Forestry, Nanchang, China
Email: yanfangwu2012@163.com

Abstract: To more quickly and accurately identify the five chemical types of *Cinnamomum camphora*, a desktop and portable near-infrared spectrometer was used to collect the near-infrared spectra of the leaves of the five chemical types of *C. camphora*. Combined with chemometric methods, the near-infrared spectra of different chemical types of *C. camphora* were discriminated and analyzed. Using a desktop near-infrared spectrometer in the 1000-2000 nm wavelength range, we applied dg2 and SNV preprocessing techniques, combined with the SIMCA algorithm, to build a model. This model successfully identified the three chemical types- Linalool, Borneol, and Camphor-with 100% accuracy. Only two samples of Cineol and Iso-nerolidol were not correctly identified. In the wavelength range of 1600-2400 nm, the portable near-infrared spectrometer was pretreated with S-G smoothing and SNV pretreatment. The optimal model was established after spectral matching and the sample recognition rate of the model was as high as 99.73%. A sample of Cineol and Iso-nerolidol was incorrectly identified. In terms of the external verification test, the recognition rates of the two models reached 99.20 and 98% respectively, showing high recognition ability. This study provides a robust and efficient method for the rapid, field-based identification of five different chemotypes of *C. camphor*a, which could significantly benefit practical applications in the field.

Keywords: *Cinnamomum camphora*, Chemical Type, Near-Infrared Spectroscopy, Desktop, Portable

Introduction

Cinnamomum camphora, as a characteristic species of the Camphor genus and even the Camphor family, has important cultural, ecological, and economic value. According to the different main components of *C. camphora* leaves camphor oil, it can be divided into five different chemical types: Linalool, Borneol, Cineol, Iso-nerolidol, and Camphor (Zhang *et al*., 2020). Different chemical types have different applications. For example, the main content of the essential oil of Linalool leaves is "linalool" and the natural linalool is known as the "king of spices". It is the most frequently used spice variety in the production of perfume essence, household products, and soap essence and is widely used in all aspects of essence and fragrance (HongFu *et al*., 2019; Pereira *et al*., 2018; Dos Santos *et al*., 2018). Natural Borneol is not only a precious traditional Chinese medicine, but also a high-end spice (Li and He, 2017). It has protective effects on organs and tissues such as the heart and brain, antibacterial, anti-inflammatory (Badillo *et al*., 2010), analgesic, and absorption-promoting effects. The differentiation of chemical types is mainly based on the differences in chemical components of the same plant species, which is a manifestation of intra-species biodiversity of plants and the manifestation of morphological differences is not significant (Tian *et al*., 2021). Therefore, it is difficult to accurately distinguish different chemical types of *C. camphora* only by observing their external morphology. Other auxiliary means are needed for accurate identification. The common methods for distinguishing different chemical types of *C. camphora* include: Aroma analysis, essential oil analysis, and molecular biology. The fragrance method mainly relies on the subjective judgment of the discriminator and requires experienced professionals. It is greatly influenced by human factors and has certain limitations. The essential oil analysis method mainly extracts essential oils through traditional methods and then combines GC-MS and other instruments to analyze the chemical composition of essential oils, thereby achieving the study of the chemical type of camphor trees. There are three main drawbacks:

① All samples need to be preprocessed and the procedure is cumbersome
② Instrument testing is expensive, time-consuming, and labor-intensive
③ A large amount of samples is required and the plants are severely damaged. Molecular biology methods are used to study the differences in different chemical types of camphor trees at the gene and protein level, requiring complex experimental procedures and high professional requirements. It is difficult to quickly identify their differences and can usually only be carried out in the laboratory, which cannot be widely promoted in production. The above three methods all have certain limitations. This study proposes a new method for identifying different chemical types of camphor trees - near-infrared spectroscopy.

NIR analysis technology, as a new type of optical detection technology, has the advantages of no damage to the sample, non-destructive testing, preserving the original characteristics of the sample, being pollution-free, and consumable-free for testing, harmless to the environment and human health and reducing long-term testing costs (Qiu *et al*., 2023; 2021; Forte *et al.*, 2022). In recent years, NIR analysis has been widely used in the identification of plant species (Jue *et al*., 2023; Chen *et al.*, 2021; Wu *et al.*, 2020). Wang *et al*. (2014) used a portable spectrometer combined with partial least squares discriminant analysis (PLS-DA) to establish discriminant models for the field-measured spectra of four bamboo species: *Phyllostachys aurea*, *Bashania fargesii*, *Pseudosasa japonica*, and *Phyllostachys glauca*. These established models were used to identify the leaves of the validation set of bamboo species, with a recognition rate of 100%. Wang *et al.* (2019) combined PLS-DA with a multi-column recognition variable matrix to establish 9 leaf discrimination models, with a recognition accuracy of 99.58%.

In the early stage, our team studied the ability of desktop near-infrared spectrometers to identify four species of camphor plants: *Cinnamomum bodinieri*, *C. parthenoxylon*, *C. longepaniculatum* and *C. septentrionale*. The results showed that the constructed model can effectively identify these four species of camphor plants (Tu *et al*., 2022). At the same time, the ability of a portable near-infrared spectrometer to identify 10 different species of Lauraceae plants was also explored and good experimental results were obtained. The constructed model was able to successfully identify different species of Lauraceae plants. It has been confirmed that NIR technology can be applied to species identification of *Camphora* and Lauraceae plants and the recognition effect is significant. In the selection of modeling instruments, both desktop and portable near-infrared spectrometers can effectively identify different species of Lauraceae. However, there is no report on whether near-infrared spectroscopy can be used to detect the chemical composition of *C. camphora* leaves camphor oil and whether it can distinguish different chemical types of the same species. This study used the Fourier desktop near-infrared spectrometer from Buchi, Switzerland and the handheld near-infrared spectrometer from Thermo Fisher Scientific, USA (microPHAZIRTM Rx) to collect near-infrared spectral information of five different chemical types of *C. camphora* leaf samples. Combined with classification algorithms, a recognition model for different chemical types of *C. camphora* was established to explore the feasibility of applying NIR analysis technology to the identification of *C. camphora* chemical types, thereby solving the problem of rapid identification of the main chemical types of *C. camphora*.

Materials and Methods

Test Instruments and Materials

In August 2021, five different chemical types of leaves were collected from the Lauraceae Plant Resources Conservation Base of Jiangxi Academy of Forestry (28°44′41″N, 115°48′46″E): Linalool, Cineol, Iso-nerolidol, Borneol, and Camphor. Five 3-year-old individual plants were randomly selected for each chemical type to collect samples. The mature functional leaves of the year with intact appearance, healthy growth, without wormholes and yellowing were selected from four directions: East, South, West, and North at the time of collection. The collected samples were brought back to the laboratory for near-infrared spectroscopy collection in time. Benchtop NIR spectroscopy and portable NIR spectroscopy were used to collect leaf spectral information and the sample information is shown in Table (1).

The desktop near-infrared spectrometer is the Fourier transform near-infrared spectrometer NIRFlex N-500, equipped with a solid measurement cell (BUCHI, Switzerland), with a spectral range of 1000-2500 nm and a resolution of 13.38 nm. The main analysis software includes Operator spectral acquisition software, NIRCal spectral analysis software, and MATLAB R2018b data processing software.

The portable near-infrared spectrometer is the American Thermo Fisher handheld near-infrared spectrometer (micro PHAZIRTM Rx), which operates on the principle of non-destructive chemical analysis in near-infrared spectroscopy. The sampling mode is diffuse reflection and the light source is a halogen tungsten lamp, which is harmless to the sample and human body. The collection spectrum range is 1600-2400 nm and the measurement time is generally only 3 seconds or even shorter. It can work in ambient temperatures of 5-40°C.

Table 1: Sample Collection Information

Samples	Sampling number of samples/plant	Desktop			Portable		
		Total number of leaves	Modeling set sample	Validation set sample	Total number of leaves	Modeling set sample	Validation set sample
Linalool	5	200	134	66	210	200	10
Cineol	5	200	134	66	200	190	10
Iso-nerolidol	5	200	134	66	201	191	10
Borneol	5	198	132	66	206	195	10
Camphor	5	200	134	66	155	145	10

Near Infrared Spectroscopy Collection

The collected fresh leaf samples are directly used for near-infrared spectroscopy collection. Research has shown that dust on the surface of leaves can have a certain impact on the reflectance of spectral collection (Zhang *et al*., 2023; Jung *et al.*, 2022). Therefore, before spectral collection, it is necessary to wipe off the dust on the surface of leaves to avoid its impact on the model performance. The sample collection method of the desktop near-infrared spectrometer is to collect the front spectrum of the leaf and each leaf is collected to obtain 3 spectra. The average spectrum of these 3 spectra is taken as the characteristic spectrum of the leaf. The portable is to take the aluminum foil as the background, lay the blades flat on the aluminum foil, and align the instrument light source with the collection site for spectral collection. The front spectrum of the leaf is collected and the midpoint of the main vein of the leaf vein is taken as the dividing line and the spectra of the upper, middle (midpoint and lower part of the main vein of the leaf vein) are collected. 3 spectra are collected from each leaf and the average spectrum is taken as the characterization spectrum of the leaf sample, so as to ensure that the part of the collection spectrum of each leaf is consistent.

Spectral Preprocessing

Savitzky Golay (S-G) smooth differentiation, which includes differentiation and smoothing, is the most commonly used method for signal denoising in NIR analysis (Ngeljaratan *et al*., 2021; Yin *et al.*, 2023). It can reduce the impact of instrument noise on spectra and improve signal-to-noise ratio (Melese *et al.*, 2023). Normalization processing eliminates the adverse effects of spectra with significant differences during analysis by performing dimensionless processing on the spectra, thereby improving the robustness and accuracy of the analysis model (Folli *et al*., 2022). The standard normal transformation of SNV eliminates scattering effects by normalizing the spectral matrix into rows (Németh and Rákosa, 2023). Normalize Range is normalized by finding the maximum and minimum values of each spectrum and then dividing each point in the spectrum by that value (maximum-minimum) and finally the total range covered by each spectrum is exactly 1. A combination of S-G smoothing, SNV, or normalized range was selected to pretreat the sample spectra. The three parameters of S-G smoothing: Derivative, smooth pts, and degree were adjusted. Taking the total mismatch of the model as the evaluation index, the preprocessing effect under different parameter settings was compared to select the best preprocessing method and establish the optimal model.

Discriminant Model Establishment

Desktop

The family-independent soft pattern method SIMCA (Soft Independent Modeling of Class Analogy) is selected for modeling and the class distance between samples in principal component space and samples of different categories is used to distinguish the attributes of samples (Khanban *et al.*, 2022). It is realized through two steps: First, a modeling set with known attributes is selected for Principal Component Analysis (PCA) on the spectral matrix of the modeling set samples, establishing a SIMCA principal component analysis model for known categories; second, the constructed model is used to calculate the spectral residuals of the SIMCA principal component analysis model for each category of unknown samples. The category with the smallest spectral residuals is the category of unknown samples (Li *et al.*, 2022; Foschi *et al.*, 2022; Yan *et al.*, 2023).

The calculation steps are as follows:

All samples in a dataset are separated and numbered according to different categories. PCA calculates each type of data separately. Taking one type of spectral data X as an example, a PCA model is established as follows:

$$X = \bar{X} + TP \quad (1)$$

$\bar{X}$: Sample means; T: Score matrix; P: Load matrix.

Cross-validation determines the main component score A:

$$X = \bar{X} + T_A P_A + E \quad (2)$$

E: Model residuals.

The sum of residuals for all samples with principal component A is calculated as follows:

$$S_0 = \frac{\sum_{i=1}^{n} E^2}{(n-A-1)\times(M-A)} \quad (3)$$

M: The dimensions of the P matrix.

The residual of unknown samples in this situation is calculated as follows:

$$S_i = \frac{(X_i-\bar{X}-T_A P_A)^2}{M-A} \quad (4)$$

F the residual of unknown samples in this situation is calculated as follows:

$$F = \frac{S_i}{S_0} \quad (5)$$

The calculated value and critical values of F (degrees of freedom are $(M-A)$ and $((n-A-I)\times(M-A))$, respectively) are compared. If the calculated value is less than the critical value, the unknown sample belongs to that category; if it is greater than the critical value, it does not belong to that category.

Portable

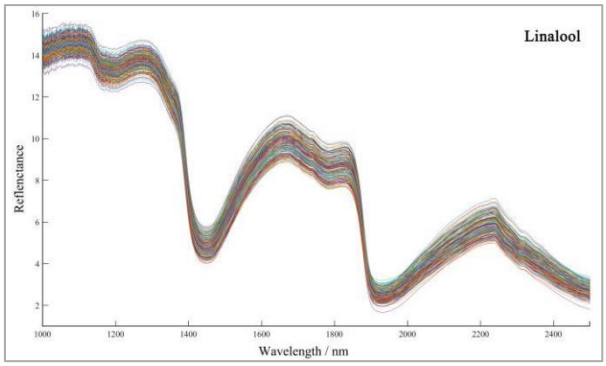
The instrument-supporting analysis software Method Generator and MATLAB R2018b software are used to complete average spectrum, spectral preprocessing, model establishment, and validation. Spectral Matching is used for the calculation of the model. Spectral matching is used for material identification or verification, which refers to comparing the shape of each spectrum in the dataset, using the spectrum with the highest matching value with an unknown sample to identify that sample. The effectiveness of the model is evaluated based on the value of Total Mismatches. When it is 0, it indicates that all samples can be completely distinguished from each other, otherwise, it is not. Qualitative models can be divided into two modes: Identification (ID) and Verification (Verification). The former is used for identification, which can identify a certain sample as a certain substance in the model; The latter is used for verification, verifying whether a certain sample is the required sample. According to the experimental design, this experiment establishes an ID identification model. A threshold is set for the constructed model, which means that when the spectrum of the scanned sample matches the spectrum in the model, it is judged as the corresponding sample. The threshold in this study is consistently set to 0.99. The model with external validation set samples is validated to test its recognition ability for unknown samples.

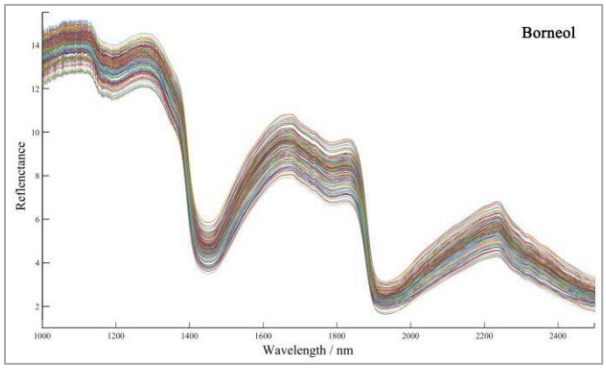
Results and Analysis

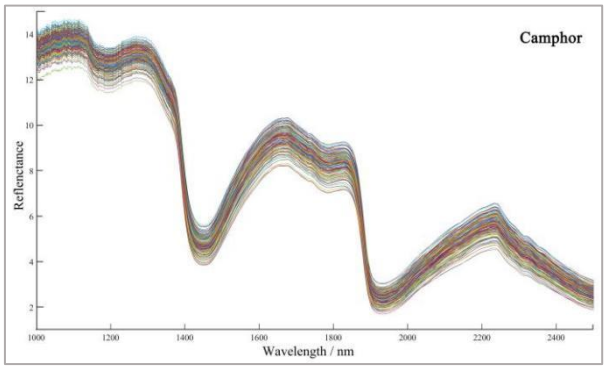
Desktop Near-infrared Spectroscopy Discrimination Model

Near Infrared Spectroscopy Analysis

The original spectra of five different chemical types of *C. camphora* leaf samples are shown in Fig. (1). From Fig. (1), it can be seen that the changes in the near-infrared spectra of the five chemical types of *C. camphora* leaf samples are basically consistent. However, different samples have certain differences in reflectivity at the same wavelength and are not completely identical. This may be due to the unique chemical composition and structure of the substances inside the leaves, indicating that there are certain differences in the internal composition of the leaves.







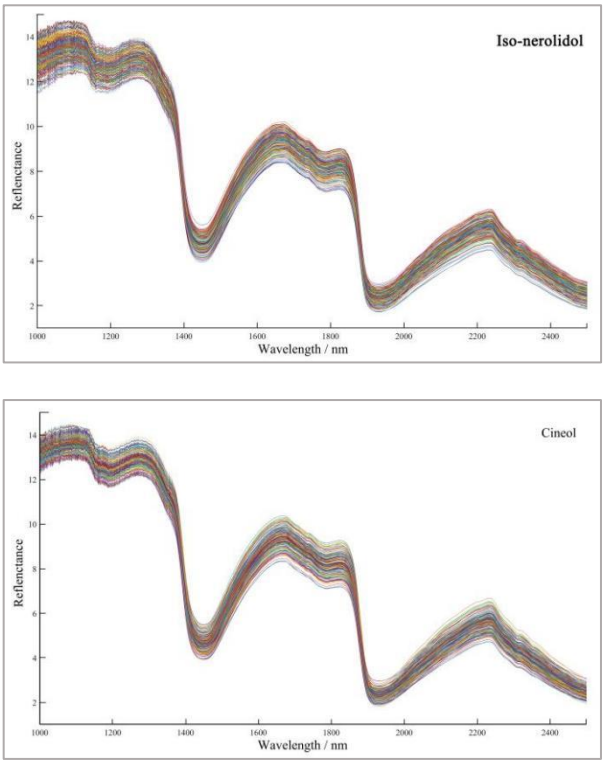


Fig. 1: Original near infrared spectrogram of sample

Principal Component Analysis

The PCA method is one of the classic feature extraction and dimensionality reduction techniques, which can classify unknown samples without any relevant knowledge background (Németh and Rákosa, 2023; Yuan *et al.*, 2023; Francisco and Martín, 2022). Figure (2) shows the sample distribution map of the original spectral data after principal component decomposition, retaining the first three principal components. As shown in Fig. (2), the cumulative contribution rate of the first three principal components can reach 98.80%. Camphor tree samples with similar chemical compositions are closer in multidimensional space and have a certain clustering trend. This lays the foundation for the establishment of different chemical types of camphor tree recognition models in the future. On this basis, some spectral data preprocessing methods can be used to improve the resolution of the model. Based on this, some spectral data preprocessing methods can be used to improve the resolution ability of the model.

Spectral Preprocessing and Modeling Band Selection

Table (2) shows relevant parameters of NIR discrimination models for different chemical types of *C. camphora* established under different conditions. According to Table (2), the classification model established using the original and preprocessed spectra has a recognition rate of 100% for the modeling set samples, while the recognition rate for the validation set samples varies. It can be clearly seen from Table (2) that the sample recognition rates of the model validation set established without preprocessing are all low, while those preprocessed with dg2+ SNV show a significant improvement. The identification rates for the samples of Linalool, Borneol, and Camphor are 100% and those of Cineol and Iso-nerolidol are 96.97 and 95.45%, respectively. Only 2 spectra of Cineol and 3 spectra of Iso-nerolidol are not correctly identified.

This may be due to some differences in sample processing and preservation in the early stage, which led to its incorrect identification. It needs to be optimized and improved in subsequent research.

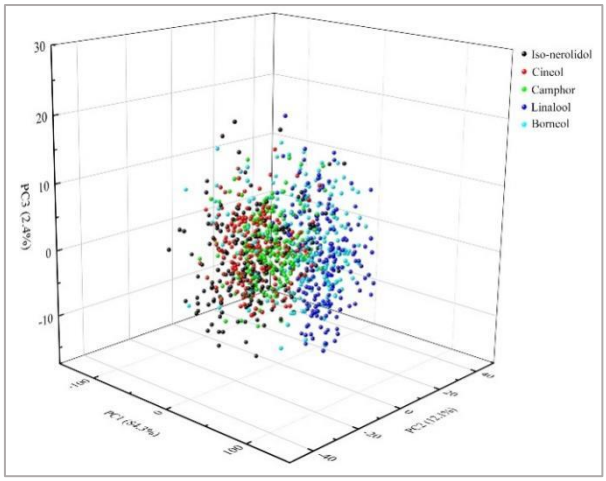


Fig. 2: 3D score for principal component analysis

Table 2: Calibration results of the SIMCA model after spectral preprocessing

C. camphora chemical type	Spectral preprocessing methods	Modeling band/nm	Principal components	Recognition rate/ %	
				calibration set	Validation set
Linalool	dg2+ SNV	1000-2000	6	100	100
Cineol	dg2+ SNV	1000-2000	5	100	96.97
Iso-nerolidol	dg2+ SNV	1000-2000	6	100	95.45
Borneol	dg2+ SNV	1000-2000	6	100	100
Camphor	dg2+ SNV	1000-2000	6	100	100
Linalool	-	1000-2500	15	100	46.97
Cineol	-	1000-2500	15	100	43.08

Iso-nerolidol	-	1000-2500	15	100	24.24
Borneol	-	1000-2500	15	100	12.90
Camphor	-	1000-2500	15	100	40.91

Note: dg2: 2st Savitzky-Golay 9 Points; SNV: Standrad Normal Variate

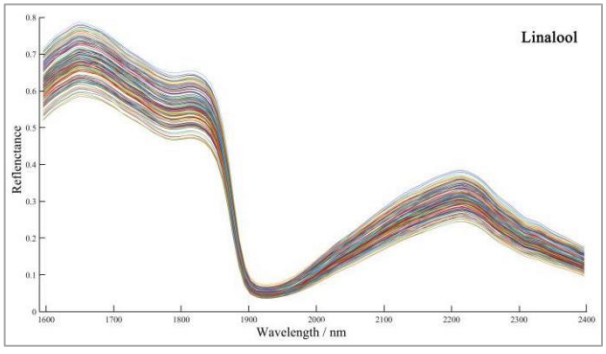
Establishment and Validation of Models

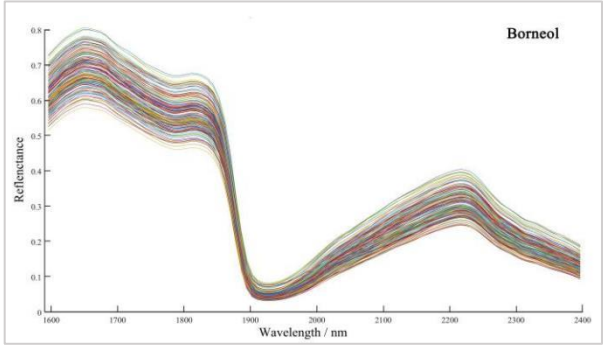
SIMCA method is selected as the modeling algorithm, with 1000-2000 nm as the modeling wavelength, to establish identification models for different chemical types of *C. camphora* after dg2+SNV preprocessing. Further validation of the established model is carried out by substituting 250 spectra of 5 chemical-type samples that were not involved in the model establishment into the SMICA model to calculate the recognition rate. The results show that only two samples, Cineol and Iso-nerolidol, are not correctly recognized, with a correct recognition rate of 99.20%.

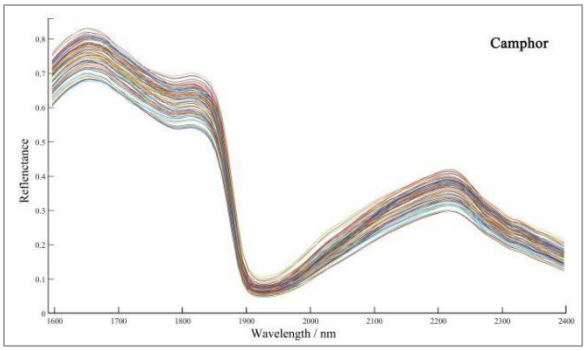
Portable Near-Infrared Spectral Discrimination Model

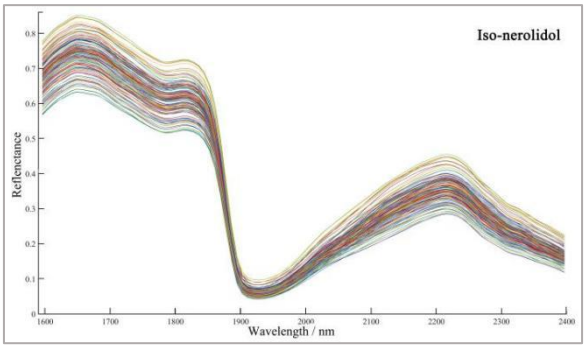
Original Spectrum of the Sample

The near-infrared spectra of five different chemical types of *C. camphora* leaf samples are shown in Fig. (3). From Fig. (3), it can be seen that the overall trend of near-infrared spectra of different chemical types of *C. camphora* leaf samples is basically consistent, with high reflectivity in the range of 1600-1850 nm. In this band, the highest reflectivity is observed in the leaves of different *C. camphora*, followed by Cineol, Camphor, and Borneol, with Linalool having the lowest reflectivity. Overall, it can be seen that the spectral absorbance of the camphor leaf sample is slightly lower than the other four chemical types in the range of 1600-2400 nm.









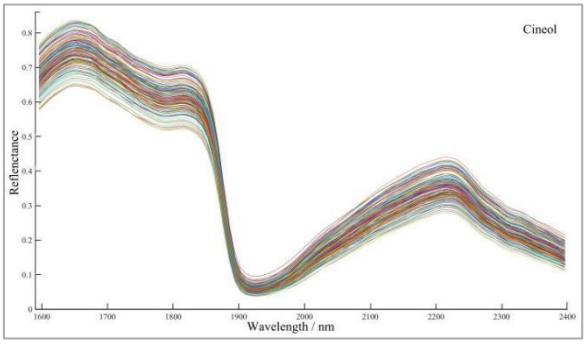


Fig. 3: Original spectra of all samples

Spectral Preprocessing

The modeling results of spectra processed under different preprocessing conditions are shown in Table (3). From Table (3), it can be seen that the spectral modeling results without preprocessing are not ideal, with a sample recognition rate of only 67.70%. After preprocessing, the sample recognition rate of the constructed model has been significantly improved, especially after the first derivative preprocessing, the sample recognition rate is above 95%. The model constructed by combining S-G smoothing (1, 3, 2) and SNV preprocessing shows the best performance, with only 2 samples not correctly recognized and a recognition rate of 99.73%. The preprocessing effect is shown in Fig. (4).

Table 3: Performance of model established by different preprocessing methods

Preprocessing method				Total mismatches	Recognition rate %
No				248	67.70
S-G + Normalize Range	1	3	2	6	99.21
	2	3	2	157	79.55
	1	5	2	27	96.48
S-G + SNV	1	5	2	29	96.22
	1	3	2	2	99.73
	2	3	3	157	79.55

Note: Pre-processing method, the different numbers represent the derivative order, the number of smoothing points, and the number of polynomials for the S-G smoothing method, respectively

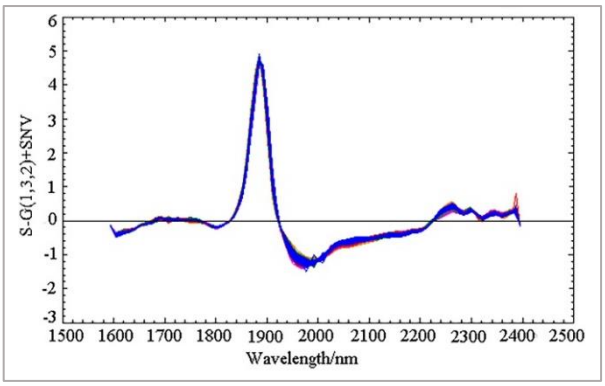


Fig. 4: Effect diagram of S-G smoothing + SNV

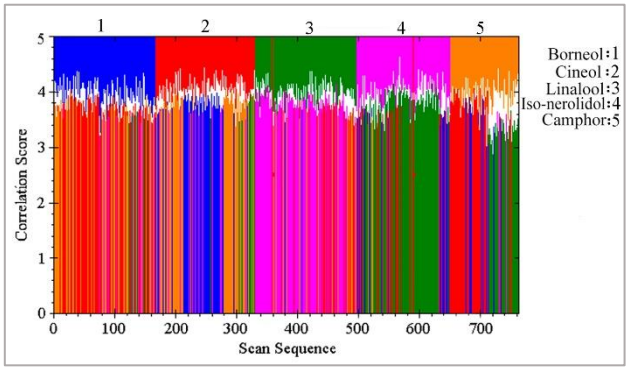


Fig. 5: Model spectra matching diagram

Establishment and Validation of Recognition Models

After processing with the best preprocessing method, the optimal recognition model is established, with a threshold of 0.99. The spectral matching effect of the model is shown in Fig. (5). From Fig. (5), it can be seen that there is a vertical red line in both the 3 (Linalool) and 4 (Iso-nerolidol) regions. This red line indicates that one Linalool sample and the Iso-nerolidol sample are incorrectly identified, while the other samples are correctly identified, with a sample recognition rate of 99.73%. The recognition of the model is tested using an external validation set of samples. The model is able to correctly recognize samples of Linalool, Cineol, Borneol, and Camphor. Only one sample of Iso-nerolidol is mistakenly identified as Linalool and the recognition rate for unknown samples reaches 98%.

Discussion

Influence of Different Times on Chemical Type Identification Models

As discussed above, the main basis for dividing *C. camphora* into different chemical types is the content of the main components of essential oils. Essential oils are secondary metabolites of plants and their relative content of main components is closely related to the metabolic activities of plants. The unique metabolic mechanism is a fundamental characteristic of the chemical type. The metabolic mechanism of *C. camphora* changes in different growth seasons or developmental stages, leading to changes in the relative percentage content of chemical or principal components in leaf essential oils (Tian *et al.*, 2021). The main principle of NIR technology is the doubling and combining absorption of hydrogen-containing groups X-H (X = C, N, O, etc.). Scanning the near-infrared spectrum can obtain information of hydrogen-containing groups in the sample and then combining chemometrics methods, the relationship between the near-infrared spectrum of the sample and its property data is established, thereby establishing a NIR correction model. In addition, qualitative identification methods using chemometrics are used to extract the near-infrared absorption spectrum features of samples and establish corresponding class models. Comparing the absorption spectra of class models with unknown samples can identify the category of unknown samples. This study establishes NIR recognition models for different chemical types by scanning spectral information of *C. camphora* leaf samples and combining their chemical type properties. The discrimination is based on the unique near-infrared absorption spectral information of different chemical types of leaves, that is the differences in the essential oils contained (Different Content of Main Components of Essential Oils). Even though the relative percentage content of the main components in the leaf essential oil of *C. camphora* varies at different developmental stages and growth seasons, regardless of the changes, the proportion of its main components is still the highest and it still has a unique near-infrared absorption spectrum. This indicates that the recognition model will not be limited by different periods.

Differences in Identification Models Built by Different Spectroscopic Instruments

This study uses a desktop near-infrared spectrometer and a portable near-infrared spectrometer to establish identification models for different chemical types of *C.*

camphora. The two models have differences in modeling bands, model preprocessing, and modeling algorithm selection, but the models have achieved good results. In terms of external validation of the models, the correct recognition rate of the models built by the desktop near-infrared spectrometer is slightly higher than that of the portable spectrometer. This may be because the portable spectrometer (Spectral Range: 1600-2400 nm) has a narrower spectral range than the desktop spectrometer (Spectral Range: 1000-2500 nm) and collects less information. However, portable near-infrared spectrometers are compact, lightweight, and easy to carry. They do not have strict requirements for the detection environment and can work in temperatures of 5-40°C, meeting the needs of workers for rapid identification in the field.

The essential oil of camphor leaves has a wide range of uses, among which the natural borneol and linalool in the essential oil of camphor leaves are widely used in essence, perfume, medicine, the chemical industry, and more. Natural borneol and linalool mainly exist in Borneol trees and Linalool trees and these chemical types are difficult to accurately distinguish from each other in terms of morphology. They are hybridized during the planting process and cannot collect high content and purity natural borneol and linalool. They need to go through a complex distillation process to obtain the target product, which seriously increases production costs. Research has found (Zhang *et al.*, 2023) that the essential oil composition of the leaves of sexually propagated seedlings of *C. camphora* varies greatly, making it difficult to maintain the characteristics of the maternal parent. Asexual reproduction offspring can better preserve the characteristics of the maternal parent. Therefore, it is necessary to identify different chemical types of *C. camphora* among dispersed and mixed populations in the wild and select target chemical types containing high content and high-purity essential oils in a planned and targeted manner according to needs and reproduce offspring through asexual reproduction (Liu *et al*., 2002; 2023 Chen *et al*., 2024). The NIR recognition model established can help workers quickly and accurately identify different chemical types in the field, laying a good foundation for selecting high content and high purity "double high" *C. camphora* varieties in the future.

Conclusion

A desktop near-infrared spectrometer is used to establish identification models for different chemical types of *C. camphora*. The establishment conditions for optimal identification models are: In the modeling wavelength range of 1000-2000 nm after dg2 and SNV preprocessing, 6 principal components are selected and the SIMCA algorithm is used to calculate the model. The constructed model can 100% recognize three chemical types of Linalool, Borneol, and Camphor, but only fails to correctly identify two samples of Cineol and Iso-nerolidol, indicating that the performance of the constructed model is optimal and can accurately identify the measured samples. The external validation results show that the constructed model has a correct recognition rate of 99.20% for unknown samples, demonstrating high recognition ability. This provides an effective method for quickly and accurately identifying five different chemical types of *C. camphora*.

A portable near-infrared spectrometer is used to establish a discrimination model for different chemical types of *C. camphora* and a wavelength range of 1600-2400 nm is selected. After S-G smoothing SNV preprocessing and spectral matching, the optimal model is established. The recognition rate of the constructed model sample is 99.73% and one sample of Cineol and Iso-nerolidol is incorrectly identified. The model is tested using unknown samples that are not involved in the model and the results show that the model has high practical application ability, with a recognition rate of 98% for unknown samples, and can basically correctly identify all samples. The above indicates that the recognition model established using a portable near-infrared spectrometer can quickly and accurately identify different chemical types of *C. camphora*, providing a new method for identifying different chemical types of *C. camphora* in the field.

The successful establishment of two models strongly proves the feasibility of using near-infrared spectroscopy technology to identify different chemical types of camphor trees. It provides a non-destructive, fast, and accurate analysis method for the identification of camphor tree chemical types, providing a scientific basis for the rational utilization and protection of camphor tree resources. In subsequent research, on the one hand, it is necessary to continue to expand the sample size, increase the number of camphor tree samples from different regions and growth environments, and improve the universality and accuracy of the study; On the other hand, it is also possible to deeply explore the potential chemical and physical mechanisms that lead to differences in near-infrared spectra, providing theoretical support for further optimization of the model. The established model is applied to actual production to verify its effectiveness and optimize and adjust the model based on feedback. At the same time, it can strengthen the interdisciplinary integration of chemistry, physics, biology, and other disciplines and jointly recommend the application and development of near-infrared spectroscopy technology in the identification of different chemical types of camphor trees.

Acknowledgment

Thanks to the strong support of various scientific research institutions in Jiangxi Province. Meanwhile, the

author sincerely expresses heartfelt gratitude to the respected editors and anonymous reviewers for their help.

Funding Information

The work was supported by the Key Project of Jiangxi Natural Science Foundation (20232ACB205018), Jiangxi Forestry Science and Technology Innovation Special Project (Innovation Special Project No.1 [2021]), and Key R&D Program Projects in Jiangxi Province (20202BBF63006).

Author's Contributions

Bailian Tu and Yueting Zhang: Designed the experiment, collated the data, and prepared the paper.

Yongjie Zheng and Xinliang Liu: Performed the experiment, analyzed the data, and wrote the paper.

Yanfang Wu: Participated in the experiment, collated the paper, and revised the manuscript.

Ethics

Authors should address any ethical issues that may arise after the publication of this manuscript.

Conflict of Interest

No potential conflict of interest was reported by the authors.

References

Badillo, L. M. D., Muñoz, R. E. M., Garciglia, R. S., & Pacheco, M. M. M. (2010). In Vitro Antioomycete Activity of Artemisia Iudoviciana Extracts Against Phytophthora Spp. *Boletín Latinoamericano y Del Caribe de Plantas Medicinales y Aromáticas*, *9*(2), 136-142.

Chen, J., Wang, W., Wang, Q., & Chen, S. (2021). Application of Near Infrared Spectroscopy in Detection of Resilience of Pit-Entry Fermented Grains in Chinese. *Liquor-Making Science & Technology*, *328*(10), 57-60.

Chen, Y., Zhu, Z., Yang, W., Zhang, Y., Yang, J., Huang, L., Hu, H., Li, J., & Chen, L. (2024). Investigating the Emission of Volatile Organic Compound from *Cinnamomum camphora*. *Industrial Crops and Products*, *219*, 119134-118957. https://doi.org/10.1016/j.indcrop.2024.119134

dos Santos, É. R. Q., Maia, C. S. F., Fontes Junior, E. A., Melo, A. S., Pinheiro, B. G., & Maia, J. G. S. (2018). Linalool-Rich Essential Oils from the Amazon Display Antidepressant-Type Effect in Rodents. *Journal of Ethnopharmacology*, *212*, 43-49. https://doi.org/10.1016/j.jep.2017.10.013

Folli, G. S., Santos, L. P., Santos, F. D., Cunha, P. H. P., Schaffel, I. F., Borghi, F. T., Barros, I. H. A. S., Pires, A. A., Ribeiro, A. V. F. N., Romão, W., & Filgueiras, P. R. (2022). Food analysis by portable NIR spectrometer. *Food Chemistry Advances*, *1*, 100074-100042. https://doi.org/10.1016/j.focha.2022.100074

Forte, M., Currò, S., Van de Walle, D., Dewettinck, K., Mirisola, M., Fasolato, L., & Carletti, P. (2022). Quality Evaluation of Fair-Trade Cocoa Beans from Different Origins Using Portable Near-Infrared Spectroscopy (NIRS). *Foods*, *12*(1), 4-228. https://doi.org/10.3390/foods12010004

Foschi, M., D'Addario, A., Antonio D'Archivio, A., & Biancolillo, A. (2022). Future Foods Protection: Supervised Chemometric Approaches for the Determination of Adulterated Insects' Flours for Human Consumption by Means of ATR-FTIR Spectroscopy. *Microchemical Journal*, *183*, 108021-108052. https://doi.org/10.1016/j.microc.2022.108021

Francisco, J., & Martín, G. (2022). Potential of Near-Infrared Spectroscopy for the Determination of Olive Oil Quality. *Sensors*, *22*(8), 2831-3100. https://doi.org/10.3390/s22082831

HongFu, Z., Liang, L. S., GuoRong, Fan, Zong, W., Xing, Chen Shang, & Yan, Si Hong. (2019). Extraction of *Cinnamomum camphora* Essential Oil by Steam Distillation and Its Antifungal Activity. *Chemistry and Industry of Forest Products*, *39*(03), 108-114. https://doi.org/10.3969/j.issn.0253-2417.2019.03.015

Jue, W., Yanjie, L., Yicun, C., Ming, G., Yunxiao, Z., Liwen, W., Shiqing, H., Yongzhi, Z., Kangshuo, Z., & Yangdong, W. (2023). The application of near-infrared spectroscopy in forestry. *Journal of Nanjing Forestry University Natural*, *47*(3), 237-246. https://doi.org/10.12302/j.issn.1000-2006.202203001

Jung, D.-H., Kim, J. D., Kim, H.-Y., Lee, T. S., Kim, H. S., & Park, S. H. (2022). A Hyperspectral Data 3D Convolutional Neural Network Classification Model for Diagnosis of Gray Mold Disease in Strawberry Leaves. *Frontiers in Plant Science*, *13*, 837020. https://doi.org/10.3389/fpls.2022.837020

Khanban, F., Bagheri Garmarudi, A., Parastar, H., & Toth, G. (2022). Evaluation of FT-IR Spectroscopy Combined with SIMCA and PLS-DA for dDetection of Adulterants in Pistachio Butter. *Infrared Physics & Technology*, *127*, 104369. https://doi.org/10.1016/j.infrared.2022.104369

Li, L., & He, H. (2017). Research Status and Development Trend on Utilization of Borneol Camphor Resource (In Chinese. *Hunan Forestry Science & Technology*, *44*(03), 90-96.

Li, L., Zhao, Y., Li, Z., & Wang, Y. (2022). Multi-Information Based on ATR-FTIR and FT-NIR for Identification and Evaluation for Different Parts and Harvest Time of Dendrobium Officinale with Chemometrics. *Microchemical Journal*, *178*, 107430. https://doi.org/10.1016/j.microc.2022.107430

Liu, X., Dai, X., Zhang, T., Wen, S., & Wu, Y. (2023). Principal Component Analysis of Essential Oil from *Cinnamomum camphora* Leaves with Different Propagation Modes (In Chinese. *South China Forestry Science*, *51*(01), 21–25. https://doi.org/10.16259/j.cnki.36-1342/s.2023.01.005

Liu, Y., Li, C., & Li, Y. (2002). Superior Individual Tree Selection on *Cinnamomum camphora* (In Chinese. *Journal of Forestry Engineering*, *16*(01), 26–28. https://doi.org/10.3969/j.issn.1000-8101.2002.01.009

Melese, A. T., Ayele, D. T., Aljerf, L., Al-Fekaiki, D. F., & Akele, M. L. (2023). Investigating the phytoavailability of metals in roots of Croton macrostachyus and Phytolacca dodecandra: induced rhizosphere processes. In *BioMetals* (Vol. 36, Issue 6, pp. 1347-1359). https://doi.org/10.1007/s10534-023-00522-9

Németh, Z. I., & Rákosa, R. (2023). Congruence Concept for Comparison of Spectra. *Applied Spectroscopy*, *77*(4), 350-359. https://doi.org/10.1177/00037028231152497

Ngeljaratan, L., Moustafa, M. A., & Pekcan, G. (2021). A Compressive Sensing Method for Processing and Improving Vision-Based Target-Tracking Signals for Structural Health Monitoring. *Computer-Aided Civil and Infrastructure Engineering*, *36*(9), 1203-1223. https://doi.org/10.1111/mice.12653

Pereira, I., Severino, P., Santos, A. C., Silva, A. M., & Souto, E. B. (2018). Linalool Bioactive Properties and Potential Applicability in Drug Delivery Systems. *Colloids and Surfaces B: Biointerfaces*, *171*, 566-578. https://doi.org/10.1016/j.colsurfb.2018.08.001

Qiu, J., Xiao, Z., Li, J., Lai, P., & Zhang, A. (2021). Applications of Near-Infrared Spectroscopy Technology in the Woody Oil Detection (In Chinese. *Hunan Forestry Science & Technology*, *48*(01), 83-88. https://doi.org/10.3969/j.issn.1003-5710.2021.01.014

Qiu, J., Zhang, A., Li, J., Xiao, Z., Li, C., & Lai, P. (2023). Quantitative Models Study on Rapid Detection of Main Quality Indexes of Litsea cubeba Fruits by Near-Infrared Spectroscopy (In Chinese. *Journal of the Chinese Cereals and Oils Association*, *38*(2), 165-171. https://doi.org/10.3969/j.issn.1003-0174.2023.02.025

Tian, Z., Luo, Q., & Zuo, Z. (2021). Seasonal Emission of Monoterpenes from Four Chemotypes of *Cinnamomum camphora*. *Industrial Crops and Products*, *163*, 113327. https://doi.org/10.1016/j.indcrop.2021.113327

Tu, B., Xie, Y., Wu, Y., Zheng, Y., Liu, X., Zhang, Y., & Xu, H. (2022). Discrimination of four Cinnamomum plants based on near infrared spectroscopy (In Chinese. *South China Forestry Science*, *50*(6), 54-60. https://doi.org/10.16259/j.cnki.36-1342/s.2022.06.013

Wang, Y., Dong, W., Li, Y., & Andrew, K. (2014). Discrimination of Bamboo Species by NIR and Partial Least Squares Discriminant Analysis (PLS-DA) in the Field (In Chinese. *Journal of Bamboo Research*, *33*(4), 16-20. https://doi.org/10.3969/j.issn.1000-6567.2014.04.003

Wang, Z., Yin, S., Li, C., & Li, Y. (2019). Identification of Tree Leaf and Species by Vis/NIR Spectroscopy (In Chinese. *Journal of Northwest Forestry University*, *34*(01), 229-236. https://doi.org/10.3969/j.issn.1001-7461.2019.01.35

Wu, G., Lai, M., Wang, Y., & Lian, H. (2020). Review on Application of Near Infrared Spectroscopy in Plant Leaves (In Chinese. *Forestry and Environmental Science*, *36*(01), 118-127. https://doi.org/10.3969/j.issn.1006-4427.2020.01.017

Yan, J., & Li, X. (2023). Rapid Identification of the Third Generation Hybrid Rice Sterile Lines by Near Infrared Spectroscopy and SIMCA (In Chinese. *Seed*, *42*(6), 146-148. https://doi.org/10.16590/j.cnki.1001-4705.2023.06.146

Yin, X., & Gu, W. (2023). Near Infrared Spectrum Denoising Based on Lifting Wavelet Transform and Savitzky-Golay Filter (In Chinese. *Chinese Journal of Sensors and Actuators*, *36*(03), 404-410. https://doi.org/10.3969/j.issn.1004-1699.2023.03.010

Yuan, L., Meng, X., Xin, K., Ju, Y., Zhang, Y., Yin, C., & Hu, L. (2023). A Comparative Study on Classification of Edible Vegetable Oils by Infrared, Near Infrared and Fluorescence Spectroscopy combined with chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *288*, 122120. https://doi.org/10.1016/j.saa.2022.122120

Zhang, B., Ou, Y., Yu, S., Liu, Y., Liu, Y., & Qiu, W. (2023). Gray Mold and Anthracnose Disease Detection on Strawberry Leaves Using Hyperspectral Imaging. *Plant Methods*, *19*(1), 148. https://doi.org/10.1186/s13007-023-01123-w

Zhang, B., Xiao, Z., Wang, Y., Zhang, H., Li, F., Cheng, H., & Jin, Z. (2020). Review on Chemotypes and Essential Oil Accumulation of *Cinnamomum camphora* (In Chinese. *Guangxi Forestry Science*, *49*(4), 623-630. https://doi.org/10.3969/j.issn.1006-1126.2020.04.028