Heliyon 10 (2024) e37650

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Rapid identification of moxa wool storage period based on hyperspectral imaging technology and machine learning

Huiqiang Hu [a,1], Yunlong Mei [a,1], Yunpeng Wei [a], Chang Liu [a], Huaxing Xu [a], Xiaobo Mao [a,b,*], Yuping Zhao [c,*], Luqi Huang [c]

[a] School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou, 450001, China
[b] Research Center for Intelligent Science and Engineering Technology of Traditional Chinese Medicine, Zhengzhou University, Zhengzhou 450001, Henan, China
[c] China Academy of Chinese Medical Sciences, Beijing 100020, China

ARTICLE INFO

Keywords:
Moxa wool
Hyperspectral imaging technology
Storage period
Machine learning

ABSTRACT

Moxa wool (MW), derived from the dried leaves of *Artemisia argyi*, plays a significant role in traditional Chinese medicine. However, the quality of MW varies with its storage period, impacting its therapeutic efficacy. Traditional methods for quality detection are limited and destructive. To address this, we propose a non-destructive detection method using hyperspectral imaging technology and machine learning algorithms to accurately identify the storage period of MW. Nevertheless, hyperspectral data poses challenges due to its high dimensionality and redundancy, leading to increased computational complexity. To overcome this, we employed principal component analysis (PCA), competitive adaptive reweighted sampling (CARS), and successive projection algorithm (SPA) for data dimensionality reduction and wavelength selection. The results demonstrate that these techniques significantly enhance the accuracy of MW storage year identification. For Nanyang MW, the CARS+SVM model achieved the highest accuracy rates of 99.8% in the visible-near-infrared (VNIR) range and 99.55% in the shortwave infrared (SWIR) range. Similarly, for Qichun MW, the SPA+SVM model achieved identification accuracies of 99.78% and 99.47% in the VNIR and SWIR ranges, respectively. This research provides valuable insights into the rapid detection of MW quality by indication of storage years and presents a novel approach for quality control of MW in the field of traditional Chinese medicine. The combination of hyperspectral imaging and machine learning offers a promising solution for efficient and accurate MW identification, contributing to the advancement of traditional medicine practices.

1. Introduction

Traditional Chinese medicine (TCM) is a precious treasure of Chinese culture, encompassing the collective wisdom of diverse ethnic medical practices and reflecting the Chinese people's profound understanding of life, health, and diseases [1,2]. It serves as a valuable resource in terms of healthcare, economy, technology, culture, and ecology. Among the numerous distinctive medical techniques in China, moxibustion therapy stands out as one of the prominent practices. Moxibustion involves the use of moxa sticks

* Corresponding authors.
E-mail addresses: mxb@zzu.edu.cn (X. Mao), zhaoyuping0316@hotmail.com (Y. Zhao).
[1] Huiqiang Hu and Yunlong Mei contributed equally to this work and should be considered co-first authors.

https://doi.org/10.1016/j.heliyon.2024.e37650
Received 20 April 2024; Received in revised form 5 September 2024; Accepted 6 September 2024

Available online 11 September 2024

or cones to ignite, burn, and apply heat to specific acupoints or areas on the body surface [3]. This therapy stimulates the acupoints and facilitates the flow of Qi (vital energy) along the meridians, thus harmonizing disrupted physiological and biochemical functions within the body [4]. As a result, it plays a crucial role in disease prevention, treatment, as well as promoting overall health and well-being. Moxibustion therapy offers a range of effects, including dispelling cold pathogens, promoting meridian circulation, reinforcing Yang energy, and resolving stasis and stagnation [5].

Moxa sticks, commonly used in moxibustion therapy, are crafted from a material called moxa wool (MW), which is obtained by drying and crushing the leaves of the *Artemisia* plant. Depending on the duration of storage for the raw *Artemisia* leaves, MW can be classified as either "new" or "aged." The latter category is further divided into one-year MW, two-year MW, three-year MW, and so on. The ancient Chinese philosopher Mencius expressed the belief that the optimal choice is the three-year MW, stating, "Like a seven-year illness, seek the three-year MW," highlighting the perceived quality associated with the storage period [6]. With the passage of time, the chemical composition of essential oils in the moxa leaves undergoes changes, leading to significant variations in therapeutic effects. The moxa sticks or cones used in moxibustion therapy can be made from different types of MW. Moxa sticks made from new MW have stronger heat intensity and higher volatile oil content, which can potentially cause skin and meridian burns [7]. In contrast, aged MW, having lost a considerable amount of volatile oil, produces milder heat, making it more suitable for medical and healthcare purposes. However, it is important to note that prolonged storage does not always guarantee better quality. Research indicates that the concentration of active ingredients increases each year, reaching its peak in the third year, but subsequently declines rapidly [8,9]. Consequently, moxa sticks made from three-year-old MW are generally preferred in moxibustion therapy.

In terms of origin, historical texts classify moxa into three distinct types: Northern moxa, Qi moxa, and Hai moxa. Qi moxa, specifically produced in Qichun, has gained widespread recognition as an authentic medicinal material with a stellar reputation [8,10]. Moxa sticks or cones derived from Qi moxa are highly esteemed for their exceptional quality. Notably, Qi moxa exhibits potent heat penetration and contains significantly higher levels of active ingredients compared to moxa from other regions [11], rendering it more effective in therapeutic applications. However, the prevalence of market adulteration has led to the substitution of inferior products for genuine Qi moxa, a concerning practice. Traditional identification methods, such as GC-MS [12] and HPLC [13], involve intricate analytical procedures, necessitate skilled operators, and entail subjective assessments. Moreover, these methods can cause damage to the medicinal materials and pose potential risks to human health due to the use of experimental reagents, in addition to being costly. Hence, there is a pressing need for novel technological approaches to efficiently monitor and maintain the quality of MW.

In recent years, it has witnessed significant advancements in non-destructive testing techniques for agricultural products and food quality inspection [14,15]. Among these innovations, hyperspectral imaging technology has emerged as a rapidly developing field [16]. By capturing and analyzing a broad range of spectral information spanning from the visible to the near-infrared region, hyperspectral imaging can unveil the distinctive chemical composition and characteristics of samples [17,18]. This technology combines the strengths of spectroscopy and imaging, enabling the simultaneous acquisition of spatial and spectral information [19]. It empowers researchers to identify and quantify specific components or compounds within samples based on their unique spectral signatures [20]. With the continuous improvement of hyperspectral imaging systems and data analysis techniques, the application of this technology to non-destructive testing of traditional Chinese medicinal materials has become increasingly viable [21].

Through the integration of advanced data analysis methods, hyperspectral imaging technology has demonstrated impressive results in these areas. For instance, Yao et al. [22] utilized hyperspectral image information to extract characteristic wavelengths of Sanqi powder and proposed an optimized least squares support vector machine model using the marine predator algorithm. Their study achieved a remarkable classification accuracy of 95% for Sanqi powder grading. In another study, Wang et al. [23] proposed a near-infrared hyperspectral wavelength selection method based on one-way analysis of variance (ANOVA). Through this approach, they identified 37 feature wavelengths that significantly contributed to the analysis. After conducting 10-fold cross-validation, the average accuracy on the test set reached an impressive 95.25%. Long et al. [24] combined hyperspectral imaging technology with chemometrics to classify the geographical origin of Hangbai chrysanthemum. They employed the BAGCT-RBFN method for origin identification and achieved recognition rates of 96.0% for the training set and 92.0% for the prediction set. Additionally, Wang et al. [25] addressed the identification of production areas for Coix seed by proposing an attention mechanism (AM), convolutional neural networks, and long short-term memory integrated deep learning model. Through the selection of effective wavelengths and the implementation of the ACLSMT model, they achieved improved prediction accuracy. These studies highlight the potential of hyperspectral imaging technology and its successful integration with advanced analytical techniques in traditional Chinese medicine research and applications.

The combination of hyperspectral imaging technology and machine learning methods has emerged as a significant research focus in the field of quality control in TCM. However, the application of this technology for identifying the storage age and origin of MW remains limited. Hence, there is immense importance in leveraging hyperspectral imaging technology for the precise identification of MW storage age and origin. This study presents a non-destructive identification method for different storage ages of MW, utilizing the integration of hyperspectral imaging technology and machine learning. Initially, hyperspectral data from MW samples originating from two different sources and various storage ages were collected. Several preprocessing methods were compared to determine the most effective approach. PCA, CARS, and SPA were employed for dimensionality reduction and wavelength selection. Subsequently, three different classification models were evaluated and compared in terms of their accuracy for storage age identification, with the aim of determining the optimal model.

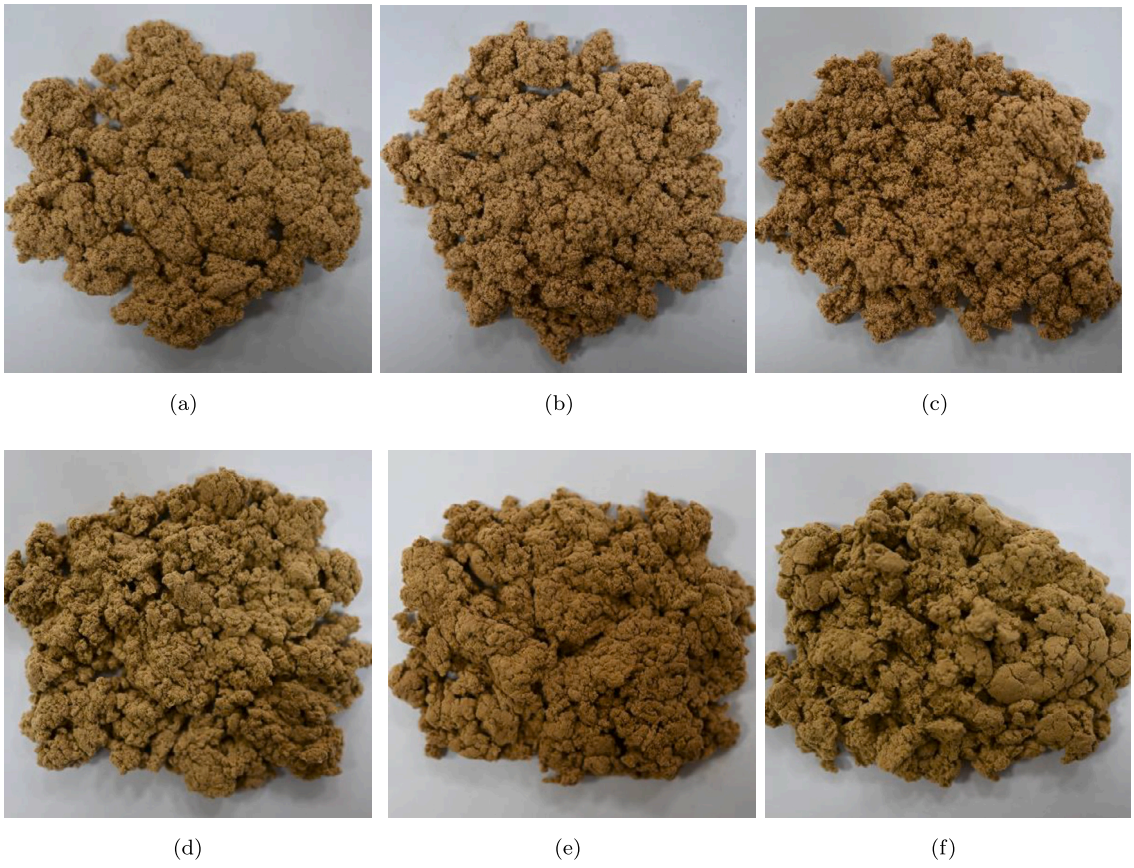

Fig. 1. (**a**) Nanyang MW, 1-year samples; (**b**) Nanyang MW, 2-year samples; (**c**) Nanyang MW, 3-year samples; (**d**) Qichun MW, 1-year samples; (**e**) Qichun MW, 2-year samples; (**f**) Qichun MW, 3-year samples.

2. Materials and methods

2.1. Sample preparation

In this study, MW samples were sourced from two distinct regions: Nanyang in Henan Province and Qichun in Hubei Province. The samples were categorized based on their storage duration, including Nanyang one year, Nanyang two years, Nanyang three years, Qichun one year, Qichun two years, and Qichun three years. The sample pool comprised 12 batches, 9 batches, 8 batches, 6 batches, 9 batches, and 6 batches, respectively, totaling 750 grams (with an average of 15 grams per batch). Fig. 1 illustrates a selection of these samples.

2.2. Hyperspectral imaging system acquisition

The study utilized a high-resolution hyperspectral imaging system featuring the HySpex series hyperspectral imaging spectrometer from Norsk Elektro Optikk A/S, Norway. The system comprised two halogen lamps, a CCD detector, a mobile platform, a visible-near-infrared camera (SN0605 VNIR), and a shortwave infrared camera (N3124 SWIR). The distance between the camera lenses and the samples was set at 25 cm. A linear scanning method was employed to capture spectral images, facilitated by a custom 150 W DC linear light source. The platform had a movement speed of 1.0 mm/s, and the spectral resolution was set at 6 nm. Optimal spectral image quality was achieved by adjusting the integration time and frame period of the camera, considering the diverse forms and appearance characteristics of the samples. During image acquisition, all lights were turned off, and measures were taken to ensure a quiet environment, minimizing the impact of external factors on the image acquisition process. The hyperspectral image (SWIR) of the Nanyang MW sample aged one year is shown in the Fig. 2.

2.3. Data correcting

Before processing the hyperspectral image data, it is crucial to mitigate the influence of external factors, such as instrument and electrical variations, on the sample data [26]. A commonly employed method is dark-current correction, which involves using

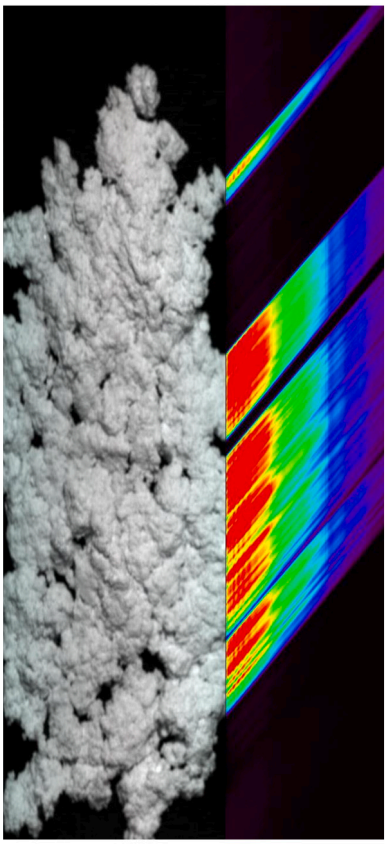

Fig. 2. Hyperspectral Image (SWIR) of One-Year Nanyang MW Sample.

ENVI 5.3 software. Dark-current correction is a well-established technique in spectral data processing, where the relative reflectance is calculated by comparing sample measurements with those obtained from standard white indicate material and dark indicate for method reference panels. The dark reference panel is typically assumed to possess a reflectance of 0, while the standard white reference panel, exhibiting Lambertian characteristics, is assumed to have a reflectance of 1. This correction effectively minimizes the impact of instrument and environmental factors, resulting in more accurate data. The calculation formula for dark-current correction is as follows:

$$I_{new} = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}} \tag{1}$$

The corrected image is denoted by I_{new}, and it is obtained by applying a correction method to the original hyperspectral image I_{raw}, taking into account the dark reference image I_{dark} and the white reference image I_{white}.

2.4. Spectral preprocessing

Spectral preprocessing methods are vital for denoising and extracting relevant information from spectral curves, compensating for noise and interference caused by instrument factors, human operational errors, and environmental variations during spectral data collection. Several commonly used spectral preprocessing methods include First Derivative (FD), Second Derivative (SD), Savitzky-Golay (S-G) smoothing, Multiplicative Scatter Correction (MSC), and Standard Normal Variable Transformation (SNV). Derivative methods involve calculating the first or second derivative of the sample spectral curve, effectively removing baseline swift and measurement errors while highlighting characteristic peaks in the spectrum. S-G smoothing, on the other hand, applies local weighted regression to the spectral curve, smoothing out noise and interference. MSC corrects for multiplicative scatter effects by introducing a scatter correction matrix, enhancing the accuracy and reliability of the spectral data. SNV standardizes the sample data, transforming it into a normal distribution with a mean of 0 and a variance of 1, ensuring comparability among different samples. These preprocessing methods play a crucial role in improving the quality and interpretability of spectral data by reducing noise, compensating for instrument variations, and enhancing the spectral features relevant to the analysis.

2.5. Features extraction

Feature extraction plays a crucial role in hyperspectral data analysis for some reasons, e.g. it reduces the dimensionality of the data by selecting a subset of relevant features, which is essential for efficient processing and interpretation. The high dimensionality of hyperspectral data can lead to computational challenges and may obscure the underlying patterns.

Principal Component Analysis (PCA) is a widely used technique for feature extraction and dimensionality reduction in hyperspectral data analysis. PCA transforms the original high-dimensional spectral data into a new set of variables called principal components, which are linear combinations of the original spectral bands [27]. The primary goal of PCA is to capture the maximum amount of variability in the data using a smaller number of derived variables. These principal components are ordered in terms of their significance, with the first component explaining the most variance in the data, followed by the second, and so on [28]. By retaining only a subset of the principal components that contribute significantly to the variance, PCA effectively reduces the dimensionality of the data. In the context of hyperspectral data, PCA helps to identify the spectral bands that contain the most relevant information. It can highlight the key spectral features and patterns that differentiate different classes or materials. This is particularly useful when dealing with large datasets and allows for efficient representation, visualization, and analysis of the data [29].

2.6. Effective wavelength selection

In the analysis of hyperspectral data, the significance of wavelength selection lies in its ability to reduce data dimensionality and focus on the spectral regions that contain the most discriminative information. By selecting a subset of relevant wavelengths, unnecessary noise and redundant spectral bands can be eliminated, leading to more efficient data processing and analysis.

Competitive adaptive reweighted sampling (CARS): The principle behind CARS involves an iterative process that incorporates two main steps: competitive learning and adaptive reweighting. Initially, each feature is assigned an equal weight [30]. In the competitive learning step, the weights of the features are adjusted based on their ability to discriminate between classes or capture important patterns in the data. Features that contribute more to the discrimination or patterns are assigned higher weights, while less informative features are assigned lower weights.

The application of CARS in hyperspectral data analysis allows for the identification of the most relevant spectral bands for a specific task, such as classification or target detection. By selecting a subset of informative bands, CARS reduces the dimensionality of the data, which can improve computational efficiency and enhance classification accuracy. Additionally, CARS helps to mitigate the curse of dimensionality and suppress the influence of noisy or redundant features, thereby improving the overall performance of hyperspectral analysis algorithms.

Successive projections algorithm (SPA): SPA is a wavelength selection method widely used in the analysis of high-dimensional hyperspectral data. SPA aims to identify a subset of the most informative and representative features (spectral bands) by iteratively projecting the data onto a subspace spanned by selected features [31]. The principle of SPA involves a sequential and greedy search strategy. Initially, all features are considered as candidates. In each iteration, SPA selects the feature that best represents the data in terms of its ability to discriminate between classes or capture important information. This is determined by evaluating a criterion such as the Fisher ratio, Mahalanobis distance, or another relevant measure. The selected feature is added to the subset of chosen features [32].

2.7. Classification models

Extreme learning machine (ELM): ELM is a machine learning algorithm that belongs to the family of feedforward neural networks. It offers efficient and effective solutions for both regression and classification tasks. ELM stands out due to its unique approach to training the neural network, which greatly reduces computational complexity compared to traditional methods. In ELM, the weights and biases of the hidden layer in the network are randomly generated, eliminating the need for iterative weight optimization. This random initialization process enables rapid training, making ELM particularly suitable for large-scale and time-sensitive applications [33]. Despite its simplified training process, ELM consistently demonstrates strong generalization performance, often achieving comparable or superior results compared to other popular machine learning algorithms, such as SVM and Multi-Layer Perceptrons (MLP), while requiring less training time.

Support vector machine (SVM): SVM are powerful and widely used machine learning algorithms for classification and regression tasks. SVM aim to find an optimal hyperplane that maximally separates or fits the data points in a high-dimensional feature space [34]. The key idea behind SVM is to transform the input data into a higher-dimensional space using a kernel function. This transformation allows the data to become linearly separable, even when the original feature space is not. SVMs search for the hyperplane that best separates the data points by maximizing the margin, which is the distance between the hyperplane and the closest data points from each class [35].

Random forest (RF): RF is a popular and effective machine learning algorithm that combines the power of ensemble learning with decision trees. It is widely used for classification, regression, and feature selection tasks. Random Forest constructs multiple decision trees and aggregates their predictions to make accurate and robust predictions [36]. During the training process, each decision tree in the Random Forest is grown by recursively splitting the data based on feature thresholds that optimize specific criteria, such as information gain or Gini impurity. This iterative splitting process allows the trees to capture different aspects of the data and contribute to the ensemble's diversity. Furthermore, the depth and complexity of the decision trees can be controlled to prevent overfitting and ensure generalization to unseen data [37].

2.8. Model evaluation and software

After constructing a classification model, it is crucial to evaluate its performance to ensure its reliability and effectiveness. Predictive accuracy is commonly used as an evaluation criterion to provide an intuitive assessment of the model's quality. The calculation formula for predictive accuracy is as follows:

$$Accuracy(\%) = \frac{TP + TN}{TP + FP + TN + FN} \tag{2}$$

where TP, TN, FP and FN refer to true positive samples, true negatives samples, false positives and false negatives samples, respectively. Accuracy is expressed as percentages, ranging from 0% to 100%, where a higher value indicates better identification performance.

In this experiment, ENVI 5.3 software was utilized for the visualization of the MW hyperspectral image data. The preprocessing of the raw data and sample collection were carried out using MATLAB 2018b. The final experimental environment was implemented using Python 3.7. The dataset was randomly divided into a training set (80% of the dataset) and a test set (20% of the dataset).

3. Results and discussion

3.1. Spectral characteristics of MW

In Fig. 3, the average spectral curves of MW samples from the two sources across different years are displayed. A comparison of the spectral curves in the VNIR and SWIR ranges reveals noticeable differences among MW samples of varying ages. This highlights the potential of hyperspectral technology for accurate identification. For instance, MW samples aged one year from Nanyang exhibit the lowest reflectance in both the VNIR and SWIR regions. On the other hand, MW samples aged three years from Qichun display the highest reflectance. Additionally, substantial differences can be observed between the two geographic sources, with Nanyang generally exhibiting lower overall reflectance compared to Qichun. These variations in spectral characteristics can be attributed to the differing concentrations of plant chemicals, particularly terpineol, within the MW samples. Terpineol is one of the primary active compounds present in MW's volatile oil, possessing pharmacological properties such as antibacterial, anti-inflammatory, and analgesic effects. The concentration of terpineol significantly influences the quality and clinical efficacy of MW. The disparities observed in the spectral curves underscore the potential of hyperspectral technology for precise identification and classification of MW samples based on their storage age and geographic origin. By analyzing the spectral information, it becomes possible to discern subtle differences that may not be apparent through visual inspection alone. This valuable insight can contribute to quality control and ensure the efficacy of moxibustion therapy.

Additionally, it is worth noting that in Fig. 3a, Fig. 3b, Fig. 3c and Fig. 3d, we also observe distinct variations in the shape and intensity of the spectral curves within the VNIR and SWIR ranges. This indicates the significance of spectral features in different wavelength ranges for discriminating MW samples based on their ages and geographic origins. Therefore, when conducting spectral analysis, it is important to consider information from multiple bands to obtain more accurate results. Therefore, MW identification based on hyperspectral technology can assist in accurately assessing the quality and age of MW, providing scientific evidence and guidance for the practice of moxibustion therapy.

3.2. Classification ability when considering full wavelengths

Firstly, SVM, ELM, and RF classification models were employed to classify the full-wavelength data. To enhance the classification performance of these models, several preprocessing methods, including FD, SD, S-G, MSC, and SNV, were applied. The experimental results are summarized in Table 1. It can be observed from the table that the classification accuracy of the preprocessed data generally exhibited improvements compared to the original data. All three classification models achieved accuracy rates exceeding 90%, except for the ELM model in the SWIR range.

Taking a comprehensive view, in the VNIR range, the RF classification model combined with the SNV preprocessing method achieved a relatively high classification accuracy. In the SWIR range, both the SVM and RF classification models, when combined with the first-order derivative and SNV preprocessing methods, yielded good classification results. These findings indicate that the adaptability and effectiveness of different classification models and preprocessing methods may vary across different spectral ranges.

In the context of identifying MW samples from two different geographic sources and various ages, the SNV preprocessing method achieved the highest classification accuracy. As a result, in subsequent experiments, the SNV algorithm was selected for spectral preprocessing. The overall results demonstrated satisfactory classification performance, emphasizing the effectiveness of integrating hyperspectral imaging technology and machine learning in accurately differentiating MW samples of different ages. However, it is worth noting that while the majority of the classification models achieved high accuracy rates, the ELM model in the SWIR range exhibited relatively lower performance. This observation suggests that further investigation and fine-tuning may be necessary to optimize the performance of the ELM model in this specific range.

3.3. Classification ability when considering effective feature and wavelengths

In the field of traditional Chinese medicine quality testing, the use of hyperspectral imaging enables the acquisition of information-rich, highly continuous, and high-resolution images. It also allows for the simultaneous capture of spatial and spectral information.

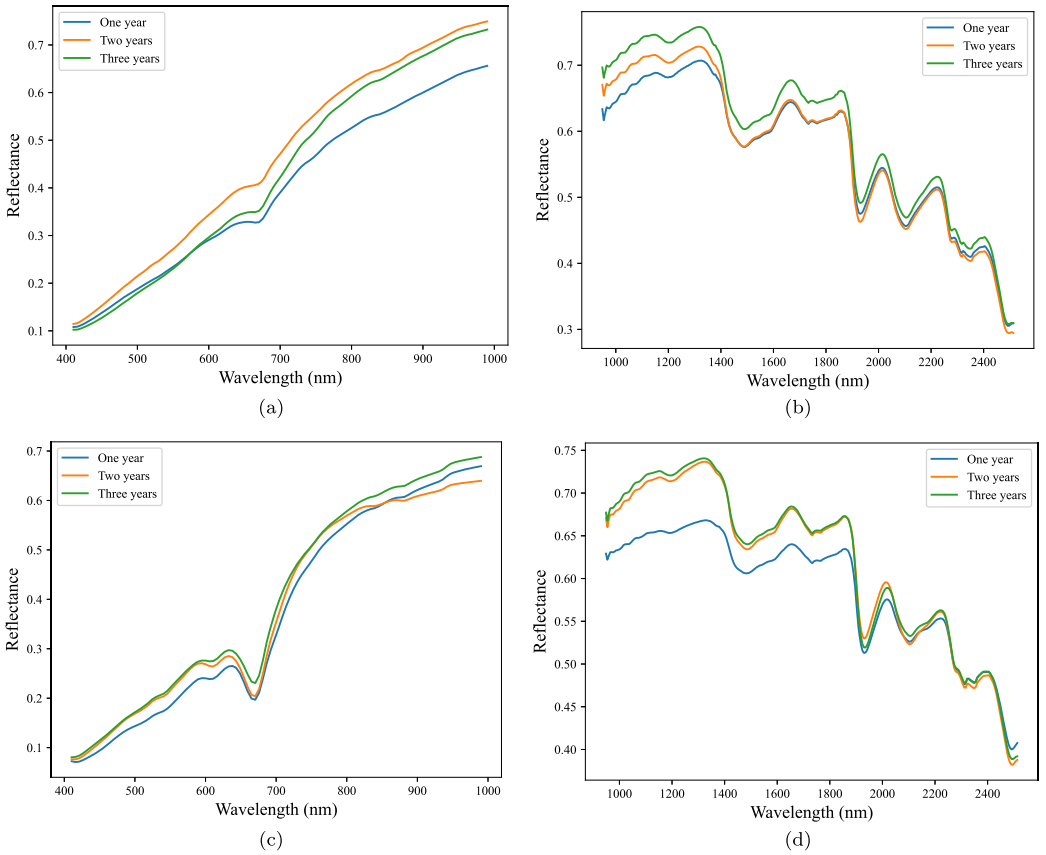

Fig. 3. (a) and (b), average spectra reflectance of Nanyang MW of different storage period under VNIR and SWIR; (c) and (d), average spectra reflectance of Qichun MW of different storage period under VNIR and SWIR.

Table 1
The classification results (Accuracy in%) of the MW's growing year using the three models with full wavelengths.

Models	Preprocessing	VNIR (400-1000 nm)		SWIR (1000-2500 nm)	
		NYMW	QCMW	NYMW	QCMW
SVM	Raw	90.27	95.96	94.55	94.68
	FD	93.76	98.20	97.27	98.40
	SD	96.82	99.10	96.82	97.87
	S-G	98.76	97.53	96.25	98.94
	MSC	97.01	98.65	96.64	97.87
	SNV	**98.94**	**99.33**	**98.18**	**97.87**
ELM	Raw	96.46	86.29	72.73	94.68
	FD	93.45	98.65	68.64	77.66
	SD	81.59	85.17	61.36	61.17
	S-G	99.82	95.73	76.23	94.68
	MSC	97.88	97.98	77.86	96.28
	SNV	**98.23**	**98.65**	**92.38**	**97.87**
RF	Raw	98.23	97.08	86.36	87.23
	FD	99.65	98.65	98.08	97.84
	SD	99.29	97.98	89.09	87.23
	S-G	98.94	97.98	88.18	92.55
	MSC	94.48	98.88	96.82	98.40
	SNV	**99.29**	**99.10**	**98.18**	**97.87**

However, hyperspectral data often contains a large amount of redundant information, which can slow down model computation. Therefore, it is necessary to perform data dimensionality reduction and select relevant spectral wavelengths. In this study, three methods, namely PCA, SPA, and CARS, were employed for dimensionality reduction and feature wavelength selection. The detailed methods and procedures can be found in Section 2.5 and 2.6

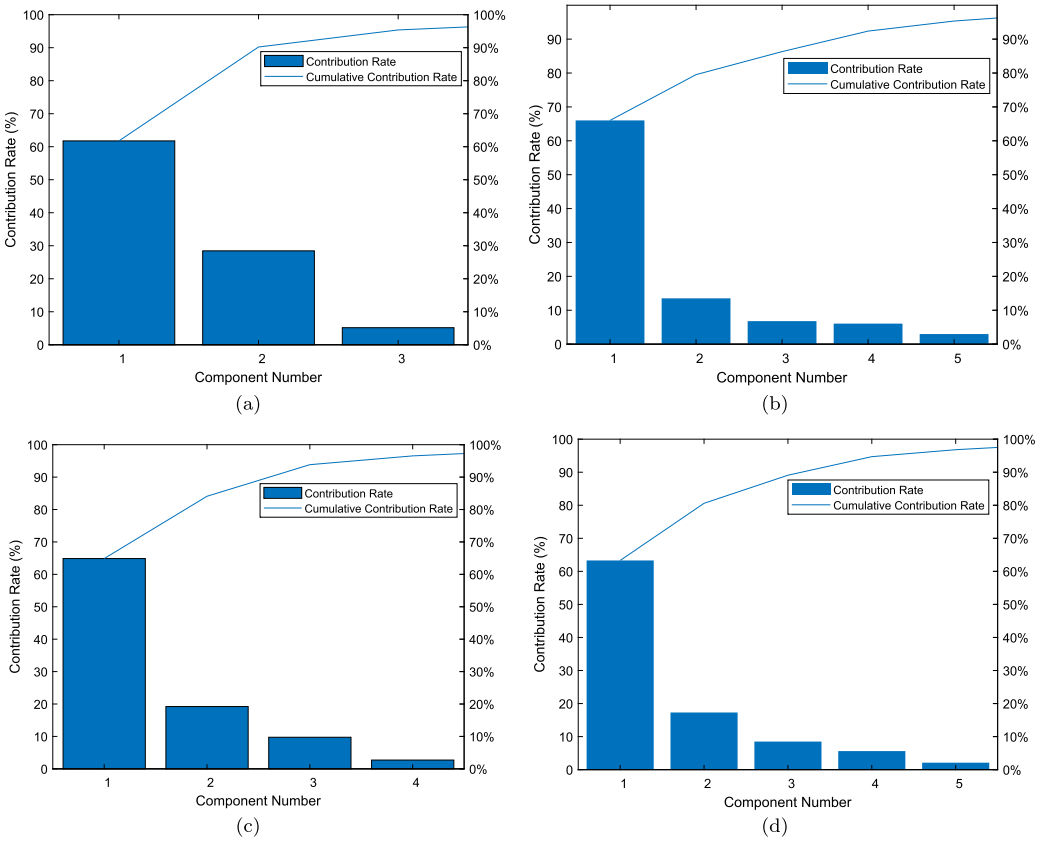

Fig. 4. (a) and (b), contribution and cumulative contribution of principal components based on Nanyang MW VNIR and SWIR data; (c) and (d), contribution and cumulative contribution of principal components based on Qichun MW VNIR and SWIR data.

3.3.1. Feature extraction with PCA

To further enhance classification accuracy and reduce computational costs, the PCA algorithm was employed in this study for dimensionality reduction on the full-band MW data. The contribution rates and cumulative contribution rates of the principal components for MW samples from two different geographic sources and various ages in the VNIR range are depicted in Fig. 4a and Fig. 4c, respectively. The results demonstrate that, for Nanyang MW samples, the cumulative contribution rate of the first 3 principal components reaches 98.35%. Therefore, the first 3 principal components were selected as the feature data. Similarly, for Qichun MW samples, the cumulative contribution rate of the first 5 principal components computed using PCA was 98.99%. Thus, the first 5 principal components were chosen as the feature data. Regarding the SWIR range MW samples from Nanyang and Qichun, PCA calculated cumulative contribution rates of 98.46% and 98.19% for the first 5 and 5 principal components, respectively, as illustrated in Fig. 4b and Fig. 4d. Consequently, the first 8 principal components for Nanyang MW and the first 10 principal components for Qichun MW were selected as the feature data. By leveraging PCA for dimensionality reduction, the number of features was significantly reduced while preserving the most informative aspects of the original data. This reduction in dimensionality not only facilitated computational efficiency but also retained the essential discriminatory characteristics for accurate classification.

3.3.2. Effective wavelengths extraction with CARS

Furthermore, the CARS algorithm was utilized to select effective wavelengths for the MW data. The MW sample's spectral were preprocessed using the SNV method, and the resulting spectral curve data were used as input variables, while the storage years were used as the output variable. The objective was to identify the feature wavelengths that are correlated with the storage years.

For Nanyang MW, the algorithmic process of wavelength selection in the VNIR range is illustrated in Fig. 5A. Fig. 5A illustrates the variable change trend, five-fold cross-validated root mean squared error (RMSECV), and regression coefficient values for each variable during the feature variable selection process at different thresholds. Specifically, Fig. 5A(a) displays the variation trend of wavelength variables during the iteration process of the CARS algorithm. As the number of iterations increases, the curve of the selected feature variables rapidly declines initially, reaches a threshold, and then gradually stabilizes. The left end of the curve indicates faster computation speed and a rapid reduction in the number of wavelengths, while the right end indicates slower computation speed with a gradual flattening. Additionally, Fig. 5(b) presents the RMSECV values, which exhibit a "valley" shape as the number of iterations increases. The minimum RMSECV value, only 0.1335, is achieved at a threshold of 19. Furthermore, Fig. 5A(c) presents the regression coefficient values for each variable. Combined with Fig. 5(b), it becomes evident that a subset of 20 feature wavelengths was selected.

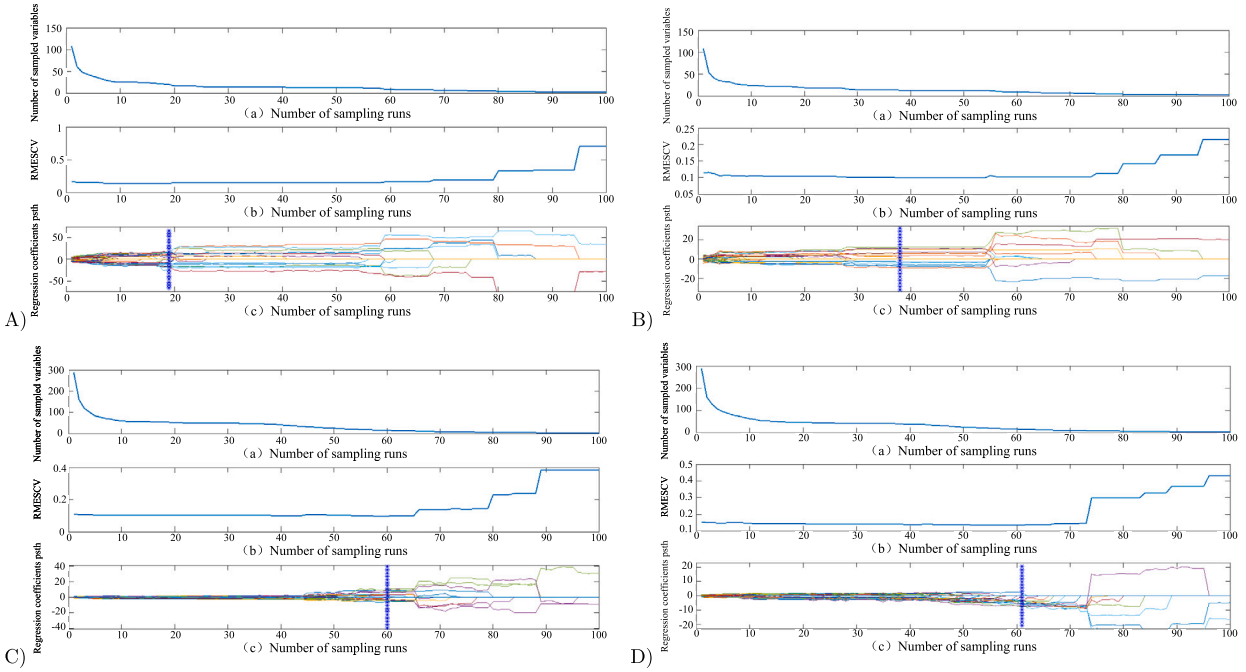

Fig. 5. CARS Wavelength Selection Process. (A) and (B): Based on Nanyang MW VNIR and SWIR data; (C) and (D): Based on Qichun MW VNIR and SWIR data.

Regarding the SWIR range spectral data, the algorithmic process of wavelength selection is shown in Fig. 5B. The optimal number of iterations for variable selection was determined to be 60, and the RMSECV value reached 0.0988 when the threshold was met. A total of 15 feature wavelengths were selected in this process.

For Qichun MW, the variable change trend, five-fold cross-validated root mean squared error (RMSECV), and regression coefficient values during the feature variable selection process are illustrated in Fig. 5C(a), (b), and (c), respectively. According to Fig. 5C(b) and (c), it can be observed that the minimum RMSECV value of 0.0988 is achieved at a threshold of 38, indicating the efficient removal of wavelengths unrelated to storage years on the left side of the threshold. On the right side of the threshold, the increasing RMSECV values indicate the removal of variables with higher correlation. Additionally, Fig. 5(c) demonstrates that a subset of 13 feature wavelengths was selected based on the correlation coefficients. Regarding the SWIR range, the algorithmic process of wavelength selection is depicted in Fig. 5D. The optimal number of iterations for variable selection was determined to be 61, and the RMSECV value reached 0.1331 when the threshold was met. A total of 14 feature wavelengths were selected in this process. The application of the CARS algorithm for wavelength selection allowed for the identification of feature wavelengths that exhibit strong correlations with the storage years of MW samples. By removing irrelevant wavelengths and focusing on the selected subset, the subsequent classification models can achieve improved accuracy and efficiency.

3.3.3. Effective wavelengths extraction with SPA

In addition to the CARS algorithm, this study also utilized the SPA algorithm for effective wavelength selection. The wavelength selection process and results for Nanyang and Qichun MW samples in the VNIR and SWIR ranges are presented. For Nanyang MW in the VNIR range, the feature wavelength selection process using the SPA algorithm is illustrated in Fig. 6a. It can be observed that the error curve stabilizes after reaching the threshold selection point, where the error value reaches its minimum of 0.1572. Therefore, this threshold point was determined as the optimal feature wavelength subset selected by the SPA algorithm, containing 12 feature wavelengths. For Nanyang MW in the SWIR range, the results of the SPA algorithm for wavelength selection are shown in Fig. 6b. The error value at the threshold was found to be 0.0968, and a total of 16 feature wavelengths were selected.

Regarding Qichun MW, the feature wavelength selection process in the VNIR range is depicted in Fig. 6c. It can be observed that the error curve remains stable after reaching the threshold selection point, with an error value of 0.1037. Therefore, it can be inferred that the SPA algorithm obtained the optimal feature wavelength subset at this point, consisting of 11 feature wavelengths. The algorithmic results for wavelength selection in the SWIR range are shown in Fig. 6d. According to the figure, the error value at the threshold was found to be 0.1476, and a total of 16 feature wavelengths were selected.

It is evident that despite the different principles guiding the tow wavelength selection methods (CARS and SPA), they all converge on similar important wavelength ranges. These selected wavelengths are primarily concentrated in regions that exhibit prominent and representative features, as shown in Fig. 7a, Fig. 7b, Fig. 7c and Fig. 7d. For example, in the VNIR range of Nanyang MW data, both the CARS and SPA algorithms selected wavelengths at 540 nm, 600 nm, 675 nm, 702 nm, and 730 nm as the feature wavelengths. In the SWIR range of Nanyang MW data, both algorithms selected wavelengths at 2207 nm and 2485 nm as the feature wavelengths. Similarly, for Qichun MW, in the VNIR range, both the CARS and SPA algorithms chose wavelengths at 421 nm and 621 nm as the

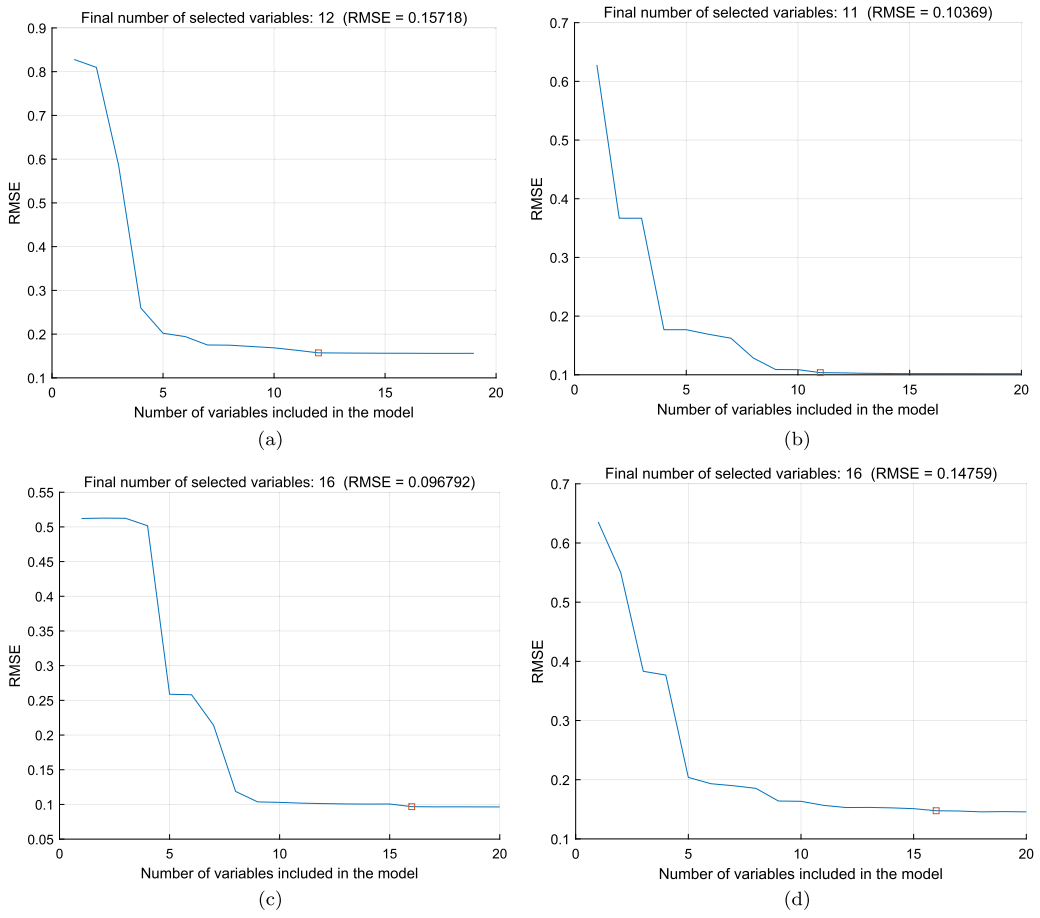

Fig. 6. Optimizing Wavelength Combinations using SPA. (a) and (b): RMSE trend with the number of wavelengths based on Nanyang MW VNIR and SWIR data, respectively; (C) and (D): RMSE trend with the number of wavelengths based on Qichun MW VNIR and SWIR data, respectively.

feature wavelengths. In the SWIR range, both algorithms selected wavelengths at 1951 nm and 2474 nm as the feature wavelengths. These selected feature wavelengths can be considered as informative and representative spectral bands for the classification and identification of MW samples from different ages and geographic origins.

3.3.4. Classification results of feature extraction and wavelength selection

After performing PCA dimensionality reduction and wavelength selection using CARS and SPA algorithms, the classification results of different models for identifying the storage years of MW from different origins are presented in Table 2. It can be observed from the table that the identification accuracy of all models has improved after dimensionality reduction and feature wavelength extraction. For Nanyang MW, the CARS-SVM model demonstrates the best performance in both the VNIR and SWIR ranges. After PCA dimensionality reduction, CARS and SPA feature selection, as well as wavelength selection, the classification accuracy of this model exceeds 99%. This indicates the effectiveness of the combined approach in accurately identifying the storage years of Nanyang MW. Similarly, for Qichun MW, the SPA-SVM model performs the best in both the VNIR and SWIR ranges. The classification accuracy on the test set reaches 99.78% and 99.47%, respectively. The SVM model, with the integration of feature extraction and wavelength selection, achieves satisfactory results in accurately identifying the storage years for both origins. These findings highlight the importance of feature extraction and wavelength selection in improving the classification accuracy of MW samples. The selected features and wavelengths provide valuable information for distinguishing MW samples of different storage years and geographic origins. The best models for Nanyang MW and Qichun MW are VNIR-CARS-SVM and VNIR-SPA-SVM, respectively and the confusion matrix is presented in the Fig. 8.

3.4. Discussion

The findings of this study provide valuable insights into the application of hyperspectral imaging and machine learning techniques for the quality assessment of MW, a traditional Chinese medicinal herb. The results demonstrate the effectiveness of wavelength selection such as PCA, CARS, and SPA in improving the classification accuracy of the models. By reducing the dimensionality of the data and selecting relevant feature wavelengths, the models were able to accurately discriminate between MW samples of different

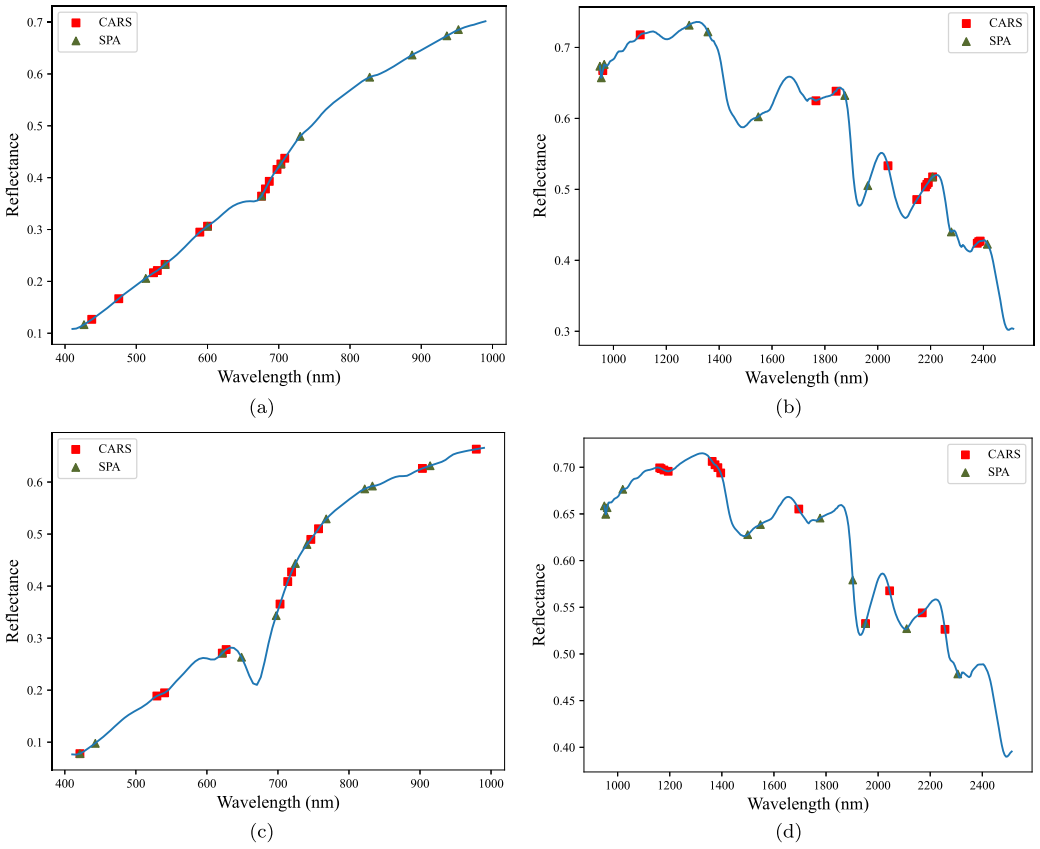

Fig. 7. The specific wavelength positions extracted for (a) VNIR data and (b) SWIR data of Nanyang Moxa, as well as (c) VNIR data and (d) SWIR data of Qichun Moxa using CARS and SPA methods.

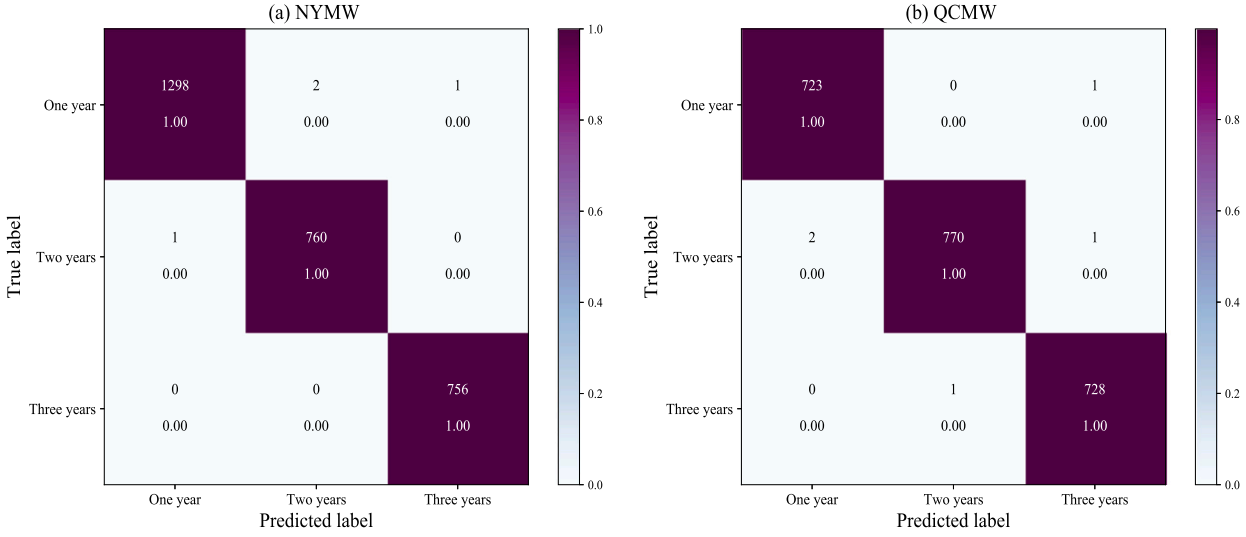

Fig. 8. Confusion matrix of the best results obtained for Nanyang MW and Qichun MW.

storage years and origins. PCA played a crucial role in extracting key principal components that captured the most significant variance in the data. This dimensionality reduction technique not only improved computational efficiency but also retained the most relevant information for accurate classification. Additionally, the CARS and SPA algorithms effectively selected optimal feature wavelengths, focusing on regions of the spectrum that exhibited distinct and representative characteristics related to MW's storage years. The performance of the CARS-SVM model for Nanyang MW and the SPA-SVM model for Qichun MW showcased the efficacy of the selected

Table 2
The classification results (Accuracy in%) of the MW's growing year using the three models with effective wavelengths.

Models	Methods	VNIR (400-1000 nm)		SWIR (1000-2500 nm)	
		NYMW	QCMW	NYMW	QCMW
SVM	PCA	99.12	99.55	99.09	98.94
	CARS	99.82	99.33	99.55	98.40
	SPA	99.65	99.78	99.09	99.47
ELM	PCA	99.65	99.10	99.09	98.94
	CARS	99.29	99.10	99.55	98.40
	SPA	99.47	99.55	99.09	98.40
RF	PCA	98.76	99.33	98.64	98.40
	CARS	99.29	99.55	99.09	98.40
	SPA	98.41	99.33	98.18	97.87

preprocessing techniques and wavelength selection methods. These models achieved high classification accuracies, exceeding 99%, demonstrating the potential of hyperspectral imaging and machine learning approaches for accurately identifying MW's storage years. The success of the SVM model in both origin classification tasks suggests that hyperspectral imaging and machine learning can be utilized as reliable tools for quality assessment and authentication of MW. The integration of spectral analysis and machine learning offers a robust framework for evaluating the quality and ensuring the efficacy of traditional medicine practices.

It is important to acknowledge that further research is necessary to validate the performance of the models on larger and more diverse datasets. Exploring advanced preprocessing techniques and feature selection methods could further enhance the accuracy and robustness of the classification models. Overall, this study demonstrates the potential of hyperspectral imaging and machine learning as valuable tools for the quality assessment and authentication of traditional Chinese medicinal herbs like MW. The findings contribute to the development of objective and efficient methods for evaluating the quality and ensuring the effectiveness of traditional medicine practices, opening avenues for future advancements in this field.

4. Conclusions

This study demonstrated the effectiveness of hyperspectral imaging technology and machine learning algorithms in accurately identifying the storage years of MW samples. By analyzing the spectral characteristics of MW, we were able to extract valuable features and wavelengths that correlated with its age. The integration of machine learning algorithms, such as SVM, ELM, and RF, further improved the classification accuracy. Preprocessing methods, including FD, SD, S-G smoothing, MSC, and SNV, played a crucial role in enhancing the performance of the models. Dimensionality reduction techniques, such as PCA, CARS, and SPA, effectively reduced computational complexity and improved efficiency. These findings provide valuable insights into MW quality control and contribute to the advancement of moxibustion therapy.

In conclusion, the successful integration of hyperspectral imaging technology and machine learning algorithms offers a promising approach for non-destructive identification of MW samples. This technology has the potential to enhance the assessment of MW quality and ensure the safety and effectiveness of moxibustion therapy. Future research can focus on expanding the sample size, refining feature extraction techniques, and exploring the impact of other factors on MW quality. The findings of this study provide a solid foundation for further advancements in the field of traditional Chinese medicine and contribute to the overall understanding and application of moxibustion therapy.

Ethics approval and consent to participate

No ethics involved.

Consent for publication

Not applicable.

Funding

This work was supported by the Scientific and technological innovation project of China Academy of Chinese Medical Sciences [grant numbers CI2021A05209] and the Innovation Team and Talents Cultivation Program of National Administration of Traditional Chinese Medicine [grant numbers ZYYCXTD-D-202205] and the Key Project at Central Government Level [grant numbers 2060302-2101-26].

CRediT authorship contribution statement

Huiqiang Hu: Writing – original draft, Software, Methodology, Investigation. **Yunlong Mei:** Methodology, Investigation. **Yunpeng Wei:** Software, Resources. **Chang Liu:** Software, Formal analysis. **Huaxing Xu:** Writing – review & editing, Resources. **Xiaobo**

Mao: Writing – review & editing, Resources, Data curation. **Yuping Zhao:** Resources, Methodology. **Luqi Huang:** Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and materials

Data will be made available upon request.

Acknowledgements

Thanks to the data analysis platform provided by the School of Electrical and Information Engineering at Zhengzhou University, and thanks to Yueyue Liang for collecting the data.

References

[1] F. Cheung, Tcm: made in China, Nature 480 (7378) (2011) S82–S83.
[2] H.Y. Xu, Y.Q. Zhang, Z.M. Liu, T. Chen, C.Y. Lv, S.H. Tang, X.B. Zhang, W. Zhang, Z.Y. Li, R.R. Zhou, et al., Etcm: an encyclopaedia of traditional Chinese medicine, Nucleic Acids Res. 47 (D1) (2019) D976–D982.
[3] J.-I. Kim, J.Y. Choi, H. Lee, M.S. Lee, E. Ernst, Moxibustion for hypertension: a systematic review, BMC Cardiovasc. Disord. 10 (1) (2010) 1–6.
[4] J.f. Zhang, Y.c. Wu, Modern progress of mechanism of moxibustion therapy, J. Acupunct. Tuina Sci. 4 (2006) 257–260.
[5] H. Deng, X. Shen, et al., The mechanism of moxibustion: ancient theory and modern research, Evid.-Based Complement. Altern. Med. 2013 (2013).
[6] S.Y. Xue, C.Y. Wang, T. Li, S.M. Liu, Y. Shi, Therapeutic efficacy observation on moxibustion with moxa of different storage years for moderate-to-severe primary knee osteoarthritis, J. Acupunct. Tuina Sci. (2020) 345–351.
[7] Y. Shao, D. Zhu, Y. Wang, Z. Zhu, W. Tang, Z. Tian, Y. Peng, Y. Zhu, Moxa wool in different purities and different growing years measured by terahertz spectroscopy, Plant Phenomics (2022) 2022.
[8] D. Liu, Y. Chen, X. Wan, N. Shi, L. Huang, D. Wan, Artemisiae argyi folium and its geo-authentic crude drug qi ai, J. Tradit. Chin. Med. Sci. 4 (1) (2017) 20–23.
[9] S.Y. Kim, Y. Chae, S.M. Lee, H. Lee, H.J. Park, et al., The effectiveness of moxibustion: an overview during 10 years, Evid.-Based Complement. Altern. Med. 2011 (2011).
[10] M. Li, X. Chai, L. Wang, J. Yang, Y. Wang, Study of the variation of phenolic acid and flavonoid content from fresh artemisiae argyi folium to moxa wool, Molecules 24 (24) (2019) 4603.
[11] R. Jin, M.M. Yu, B.X. Zhao, X.T. Fu, Y.G. Chen, H.Z. Guo, Analysis on chemical compositions of artemisia argyi from qichun of different years and moxa wool refined in different proportions, Zhongguo Zhen Jiu (Chin. Acupunct. Moxib.) 30 (5) (2010) 389–392.
[12] F. Li, S. Tang, N.T. Tsona, L. Du, Kinetics and mechanism of oh-induced α-terpineol oxidation in the atmospheric aqueous phase, Atmos. Environ. 237 (2020) 117650.
[13] F. Thibaud, M. Courregelongue, P. Darriet, Contribution of volatile odorous terpenoid compounds to aged cognac spirits aroma in a context of multicomponent odor mixtures, J. Agric. Food Chem. 68 (47) (2020) 13310–13318.
[14] T. Yi, C. Lin, J. En Ci, Y. Ji Zhong, Application and prospects of hyperspectral imaging and deep learning in traditional Chinese medicine in context of ai and industry 4.0, Zhongguo Zhongyao Zazhi (China J. Chin. Mater. Med.) 45 (22) (2020) 5438–5442.
[15] C. Dong, C. Yang, Z. Liu, R. Zhang, P. Yan, T. An, Y. Zhao, Y. Li, Nondestructive testing and visualization of catechin content in black tea fermentation using hyperspectral imaging, Sensors 21 (23) (2021) 8051.
[16] C. Zhang, W. Wu, L. Zhou, H. Cheng, X. Ye, Y. He, Developing deep learning based regression approaches for determination of chemical compositions in dry black goji berries (lycium ruthenicum murr.) using near-infrared hyperspectral imaging, Food Chem. 319 (2020) 126536.
[17] G. Lu, B. Fei, Medical hyperspectral imaging: a review, J. Biomed. Opt. 19 (1) (2014) 010901.
[18] B. Lu, P.D. Dao, J. Liu, Y. He, J. Shang, Recent advances of hyperspectral imaging technology and applications in agriculture, Remote Sens. 12 (16) (2020) 2659.
[19] M.A. Calin, S.V. Parasca, D. Savastru, D. Manea, Hyperspectral imaging in the medical field: present and future, Appl. Spectrosc. Rev. 49 (6) (2014) 435–447.
[20] G. ElMasry, D.W. Sun, Principles of hyperspectral imaging technology, in: Hyperspectral Imaging for Food Quality Analysis and Control, Elsevier, 2010, pp. 3–43.
[21] Y. Pan, H. Zhang, Y. Chen, X. Gong, J. Yan, H. Zhang, Applications of hyperspectral imaging technology combined with machine learning in quality control of traditional Chinese medicine from the perspective of artificial intelligence: a review, Crit. Rev. Anal. Chem. (2023) 1–15.
[22] K. Yao, J. Sun, N. Tang, M. Xu, Y. Cao, L. Fu, X. Zhou, X. Wu, Nondestructive detection for panax notoginseng powder grades based on hyperspectral imaging technology combined with cars-pca and mpa-lssvm, J. Food Process Eng. 44 (7) (2021) e13718.
[23] L. Wang, J. Li, H. Qin, J. Xu, X. Zhang, L. Huang, Selecting near-infrared hyperspectral wavelengths based on one-way anova to identify the origin of lycium barbarum, in: 2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), IEEE, 2019, pp. 122–125.
[24] W. Long, S.R. Wang, Y. Suo, H. Chen, X. Bai, X. Yang, Y.P. Zhou, J. Yang, H. Fu, et al., Fast and non-destructive discriminating the geographical origin of hangbaiju by hyperspectral imaging combined with chemometrics, Spectrochim. Acta, Part A, Mol. Biomol. Spectrosc. 284 (2023) 121786.
[25] Y. Wang, F. Xiong, Y. Zhang, S. Wang, Y. Yuan, C. Lu, J. Nie, T. Nan, B. Yang, L. Huang, et al., Application of hyperspectral imaging assisted with integrated deep learning approaches in identifying geographical origins and predicting nutrient contents of coix seeds, Food Chem. 404 (2023) 134503.
[26] B. Jia, W. Wang, X. Ni, K.C. Lawrence, H. Zhuang, S.C. Yoon, Z. Gao, Essential processing methods of hyperspectral images of agricultural and food products, Chemom. Intell. Lab. Syst. 198 (2020) 103936.
[27] C. Rodarmel, J. Shan, Principal component analysis for hyperspectral image classification, Surv. Land Inf. Sci. 62 (2) (2002) 115–122.
[28] S. Karamizadeh, S.M. Abdullah, A.A. Manaf, M. Zamani, A. Hooman, An overview of principal component analysis, J. Inf. Hiding Multimed. Signal Process. 4 (3B) (2013) 173.
[29] M.P. Uddin, M.A. Mamun, M.A. Hossain, Pca-based feature reduction for hyperspectral remote sensing image classification, IETE Tech. Rev. 38 (4) (2021) 377–396.
[30] H. Li, Y. Liang, Q. Xu, D. Cao, Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration, Anal. Chim. Acta 648 (1) (2009) 77–84.
[31] C. Zhang, H. Jiang, F. Liu, Y. He, Application of near-infrared hyperspectral imaging with variable selection methods to determine and visualize caffeine content of coffee beans, Food Bioprocess Technol. 10 (2017) 213–221.

[32] M.C.U. Araújo, T.C.B. Saldanha, R.K.H. Galvao, T. Yoneyama, H.C. Chame, V. Visani, The successive projections algorithm for variable selection in spectroscopic multicomponent analysis, Chemom. Intell. Lab. Syst. 57 (2) (2001) 65–73.
[33] G. Huang, G.B. Huang, S. Song, K. You, Trends in extreme learning machines: a review, Neural Netw. 61 (2015) 32–48.
[34] M.A. Chandra, S. Bedi, Survey on svm and their application in image classification, Int. J. Inf. Technol. 13 (2021) 1–11.
[35] V.K. Chauhan, K. Dahiya, A. Sharma, Problem formulations and solvers in linear svm: a review, Artif. Intell. Rev. 52 (2) (2019) 803–855.
[36] M. Belgiu, L. Drăguţ, Random forest in remote sensing: a review of applications and future directions, ISPRS J. Photogramm. Remote Sens. 114 (2016) 24–31.
[37] J.L. Speiser, M.E. Miller, J. Tooze, E. Ip, A comparison of random forest variable selection methods for classification prediction modeling, Expert Syst. Appl. 134 (2019) 93–101.