



Prediction of yield and quality in medicinal plant *Ligusticum chuanxiong* Hort. using uncrewed aerial vehicle multispectral measurement

Yun-Fan Li^{1,2}, Chen Wu^{1,2}, Hong-Mei Jia^{1,2}, Xi Chen^{1,2}, Jin-Niu Xing^{1,2}, Wei-Ping Gao^{1,2} and Zhu-Yun Yan^{1,2}

¹ State Key Laboratory of Southwestern Chinese Medicine Resources, Chengdu University of Traditional Chinese Medicine, Chengdu, China

² School of Pharmacy/School of Modern Chinese Medicine Industry, Chengdu University of Traditional Chinese Medicine, Chengdu, China

ABSTRACT

Accurate predicting the yield and quality of medicinal materials before harvest can effectively guide post-harvest process, including processing and storage, thereby ensuring the final quality of medicinal materials. Currently, traditional experimental methods for yield and quality estimation are inadequate to offer reliable guidance for harvesting and processing of medicinal plant. Uncrewed aerial vehicle (UAV) multispectral can quickly and accurately estimate the yield and quality of field crops. Based on the UAV multispectral data of *Ligusticum chuanxiong* Hort. obtained about half a month before and near harvest, this study predicted the rhizome yield and the content of active components such as ferulic acid, Z-ligustilide and senkyunolide A. Additionally, the quality discriminant models of chuanxiong rhizoma were constructed according to the ferulic acid content index stipulated in Pharmacopoeia of the People's Republic of China (2020). The results performed on the independent validation set show that the best prediction effects of fresh weight and dry weight of rhizome were NRMSE = 23.76%, MAPE = 14.75% and NRMSE = 34.65%, MAPE = 21.73%, respectively. And the best predictive effects of ferulic acid, Z-ligustilide and senkyunolide A were as follows: NRMSE = 13.35%, MAPE = 10.25%; NRMSE = 34.35%, MAPE = 23.40%; and NRMSE = 45.26%, MAPE = 25.48%. Furthermore, the quality discriminant models XGBoost and AdaBoost had effective performances (Accuracy = 0.7083, AUC = 0.7214). These results suggest that UAV multispectral can be effectively employed to predict both yield and quality before harvest, thereby guiding the harvest and processing of *L. chuanxiong*.

Submitted 13 December 2024

Accepted 14 March 2025

Published 7 April 2025

Corresponding author

Zhu-Yun Yan,
yanzhuyun@cdutcm.edu.cn

Academic editor

Viktor Brygadyrenko

Additional Information and
Declarations can be found on
page 16

DOI 10.7717/peerj.19264

© Copyright
2025 Li et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Agricultural Science, Plant Science, Spatial and Geographic Information Science

Keywords UAV Multispectral, *Ligusticum chuanxiong* Hort., Yield, Quality, Active component

INTRODUCTION

Cultivation of medicinal plants in China has a long history and is still thriving (Song et al., 2024). In 2020, the total planting area of Chinese medicinal plants is about 55, 596 km² (Wang et al., 2022). By 2025, there will be about 1,667 km² cultivation bases of Chinese medicinal plants in China (Wang et al., 2023a). The quality and economic value of Chinese

medicinal materials are directly related, and the quality of Chinese medicinal materials is not only affected by cultivation methods and collection time, but also by processing, transportation and storage processes in the production area (Yang et al., 2024). Therefore, it is necessary to prepare the human and material resources according to the yield and quality of medicine plants to complete those processes timely and efficiently for maximizing the utilization of Chinese medicine resources. How to quickly and accurately estimate the yield and quality of Chinese medicine materials before harvesting medicinal plants is an important issue in the current Chinese medicine planting and production industry.

Uncrewed aerial vehicles (UAVs), a low-altitude independent sensing technology, have become one of the most popular tools in precision agriculture production management applications due to its flexible and convenient operation and quick access to crop-related data in specific regions (Alckmin et al., 2022). The growth of multiple crops can be monitored by high-throughput phenotypic data collected by UAV (Karunathilake et al., 2023), such as corn (Shahhosseini et al., 2019), wheat (Upreti et al., 2019; Wang et al., 2021), rice (Cen et al., 2019), sugarcane (Shendryk et al., 2020), soybean (Maimaitijiang et al., 2017), potato (Jasim et al., 2020), sweet potato (Tedesco et al., 2021) and other crops. The multispectral sensor can obtain the spectral reflectance data of more than two bands on the crop surface without contact, and can calculate the Normalized Difference Vegetation Index (NDVI) and other vegetation indexes (VIs) related to crop growth status and biomass according to the data (Radočaj et al., 2023; Wang et al., 2023a). At present, the machine learning algorithm model combined with VIs derived from multispectral independent sensing images can accurately monitor crop growth, including biomass and quality prediction (Zhang et al., 2022a; Zhang et al., 2022b). For example, the use of UAV multispectral data can not only construct regression prediction models of table beet root weight (Chancia et al., 2021) and cassava underground biomass (Selvaraj et al., 2020), but also effectively evaluate the soluble solids content of grape (Lyu et al., 2023) and industrial poppy thebaine alkaloid content (Iqbal, Lucieer & Barry, 2020). However, UAV multispectral is mostly used for crop yield and quality prediction and rarely for medicinal plants.

The Chinese medicinal material, chuanxiong rhizoma, is the dried rhizome of *L. chuanxiong*, which is widely used in the treatment of cardiovascular and cerebrovascular diseases (Chen et al., 2022; Li et al., 2022). The quality of chuanxiong rhizoma is affected by the timely fresh processing and the storage conditions (Yu et al., 2021; Yan et al., 2019). Predicting the yield and quality of *L. chuanxiong* before harvesting is important to rationally manage the post-harvest handling process. In the previous study, our team used UAV multispectral to construct models for detecting water and nutrient deficiencies of *L. chuanxiong* (Li et al., 2023). Therefore, UAV multispectral can be a potential technical platform for estimating the yield and quality of *L. chuanxiong*. The main purpose of this study are divided into two points: (1) The regression model for predicting the yield of *L. chuanxiong*, the content of three active components in chuanxiong rhizoma, and the quality discriminant model of chuanxiong rhizoma based on FA content were obtained, and the accuracy and fitting degree of the model were tested; (2) the obtained models is used to predict the independent validation set in the same period and evaluate the

universality of the models. It is expected to provide a feasible method for predicting the yield and quality of *L. chuanxiong* before harvest.

MATERIALS & METHODS

Study area

The study areas, Shiyang Town in Dujiangyan City (30.85°N, 103.66°E) and Aoping Town in Pengzhou City (31.10°N, 104.00°E), were in the Chengdu Plain (Fig. 1), which lies in western China and has a subtropical humid monsoon climate. The altitude of Aoping Town is about 581.7 m, the average annual temperature is 15.7 °C, and the average annual rainfall is 924.7 mm, while the altitude of Shiyang Town is about 698.5 m, the average annual temperature is 15.2 °C, and the average annual rainfall is 1,177.5 mm (Yin et al., 2012). In addition, Shiyang Town is the traditional genuine producing area of *chuanxiong* rhizoma, while Aoping Town has the largest planting area and yield (Li et al., 2012; Li et al., 2024).

UAV image collection

The UAV platform was the DJI Phantom 4 Multispectral four-wing UAV (DJI, Shenzhen, China), which was equipped with a multispectral lens (FC6360), including a color sensor for visible light imaging and five monochromatic sensors for multispectral imaging (B: 450 ± 16 nm, G: 560 ± 16 nm, R: 650 ± 16 nm, RE: 730 ± 16 nm, NIR: 850 ± 26 nm), the resolution power of each sensor was 2.08 million pixels. The top of the UAV has a light intensity sensor that could be used for radiometric correction of later data. The flight geographical coordinates were determined by real time kinematic (RTK) GPS system with horizontal and vertical kinematic errors less than one cm and 1.5 cm respectively.

About half a month before the harvesting period (April 27, D1) and near the harvesting period (May 8, D2) in 2023, respectively, the UAV images of *L. chuanxiong* fields in Shiyang Town were obtained. In addition, the UAV images in Aoping Town was collected near the harvesting period (May 3, D3). The images were all obtained from 11:00 to 14:00 on the same day, there was no cloud cover to block the sun, and the low wind speed did not affect the flight. During the flight, the recording angle to the ground was 90° and the flight speed was 6 m/s. The flight altitude was 50 m, with a ground sampling distance (GSD) of 2.65 cm pixel⁻¹. Images were acquired with 85% forward overlap and 60% lateral overlap. Prior to each flight, the image of two radiation calibration panels with reflectance values of 10% and 90% were obtained at an altitude of 1.5 m.

Measurement of *L. chuanxiong* rhizome fresh weight and dry weight

L. chuanxiong was mined in Aoping Town (May 4, 2023) and Shiyang Town (May 8, 2023). Each sample consisted of three rhizomes of *L. chuanxiong* in the center of a square sampling area of one m². A total of 125 samples from six different planting fields in Shiyang Town were used as the training and test sets. In addition, a total of 24 samples were collected from eight different planting fields of in Aoping Town as the independent validation set. After harvesting, the above-ground parts, roots and impurities such as soil attached to the rhizomes were removed and fresh weight (FW) was measured. After the rhizomes were brought to the laboratory and dried at 105 °C for 30 min, then dried at 50 °C until the

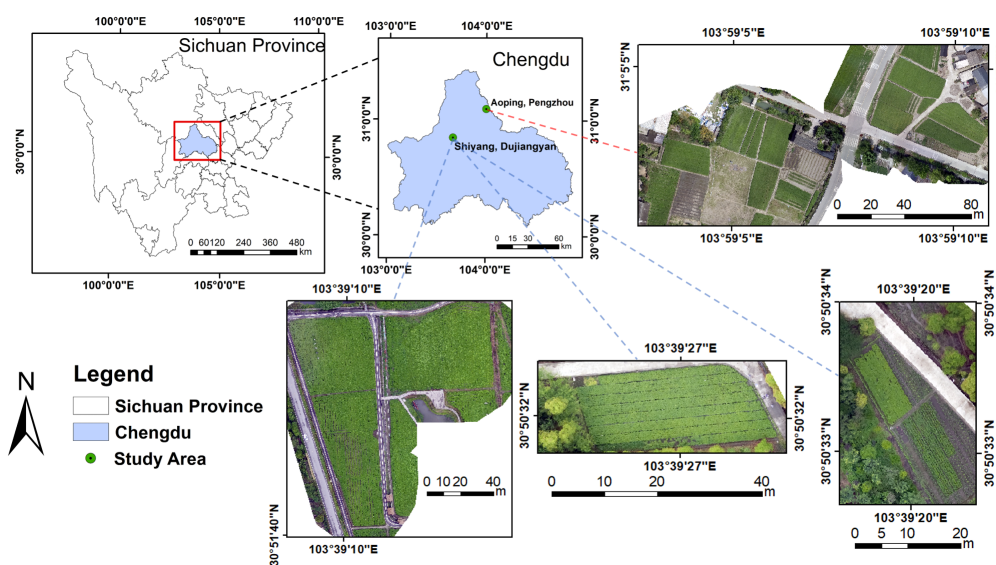


Figure 1 Location of the study area.

Full-size [DOI: 10.7717/peerj.19264/fig-1](https://doi.org/10.7717/peerj.19264/fig-1)

weight remained unchanged, the dry weight (DW) was determined. The FW and DW of each sample were the average of the FW and DW of the three rhizomes in the sample.

Measurement of active component contents

Preparation of standard substance solutions

Ferulic acid (FA), Z-ligustilide (Z-L) and senkyunolide A (SA) standard substances were dissolved in methanol, respectively, as the standard solution mother solutions. Then, they were diluted into standard substance solutions with concentrations of $31.92 \mu\text{g mL}^{-1}$, $275.00 \mu\text{g mL}^{-1}$ and $160.00 \mu\text{g mL}^{-1}$ for quantitative analysis, respectively.

Preparation of sample solutions

The dried *L. chuanxiong* rhizome samples were crushed into powder and collected through a sieve with a pore. About one g of the powder was accurately weighed and put into a conical bottle with a stopper, and then 50 mL of 75% methanol solution was accurately added. Samples were ultrasonically extracted (power: 360 W, frequency: 40 kHz) for 30 min. The supernatant was taken and filtered by $0.22 \mu\text{m}$ microporous membrane, and the filtrate was taken to obtain the sample solution.

Chromatographic conditions

Quantitative analysis was performed using an UltiMate 3000 UPLC system equipped with a DAD detector. The separation of the three active components was performed at 30°C using an Agilent ZORBAX Eclipse Plus C18 column ($2.1 \times 50 \text{ mm}$, $1.8 \mu\text{m}$). Use 0.03% phosphoric acid water (A)-methanol (B) as the gradient solvent system. The gradient elution conditions were as follows: 0~3 min, 15%~25% B; 3~5 min, 25%~28% B; 5~9 min, 28%~30% B; 10~16 min, 48%~50% B; 16~18 min, 50%~80% B. The flow rate and injection volume were 0.40 mL min^{-1} and $1 \mu\text{L}$, respectively. The detection wavelengths

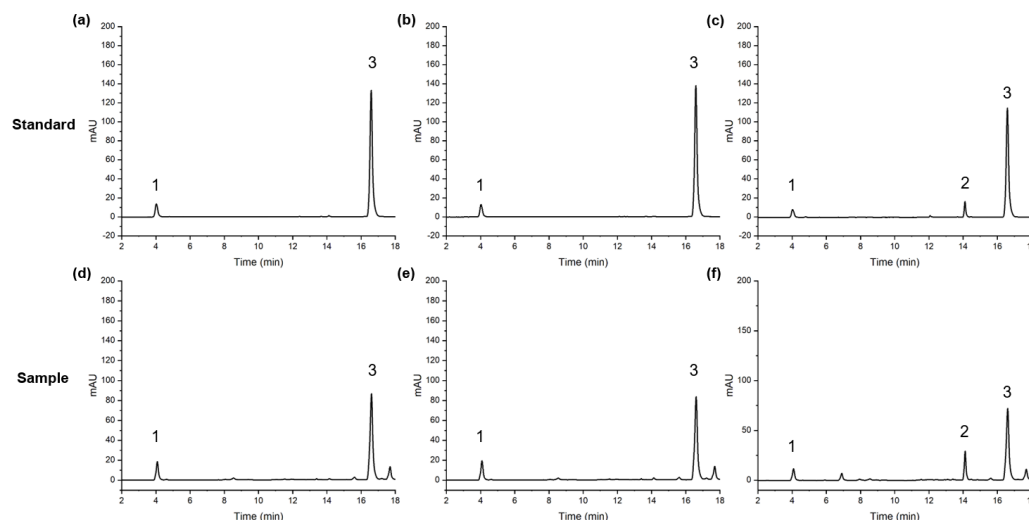


Figure 2 The peaks of FA(1), SA(2) and Z-L(3) in standard and sample. 1, FA; 2, SA; 3, Z-L; (A, D) 322 nm; (B, E) 328 nm; (C, F) 282 nm.

[Full-size](#) DOI: 10.7717/peerj.19264/fig-2

were 282 nm, 322 nm and 328 nm. The chromatographic peak separation results (Fig. 2) show that the chromatographic peak separation was good.

Multispectral image processing

The spectral images of the five bands obtained from each flight mission were imported into Pix4D Mapper 4.5.6 to complete the geometric correction and radiation correction with image information from the corresponding radiation calibration panels. Then, the spectral images were spliced into orthographic mosaic reflectance images. ENVI 5.3 software was used to overlay the orthographic mosaics of each band into a multispectral image. After classifying the canopy of *L. chuanxiong* and the background of soil and weeds by the Support Vector Machine class function of ENVI, a mask was constructed for segmentation, and the multispectral image of *L. chuanxiong* was obtained (Fig. 3). By selecting the location of each field sampling area and editing the corresponding region of interest (ROI), the VIs of each ROI were extracted and calculated to obtain 23 characteristic variables (Table 1).

VIs selection and models construction

The best model input variables were screened by Pearson correlation analysis between variables and the target values of FW, DW, FA, Z-L and SA in the corresponding samples. However, there may be a high degree of correlation and collinearity among variables. Screening only the characteristic variables with high Pearson correlation coefficient may lead to data redundancy and miss some effective information (Xu et al., 2023). Therefore, according to the order of the absolute value of the correlation coefficient between the variables and the target value from large to small, the characteristic variables with the correlation coefficient greater than 0.9 were eliminated sequentially. The remaining characteristic variables with significant correlation ($p < 0.05$) were used as model input variables to construct a prediction model corresponding to each target value.

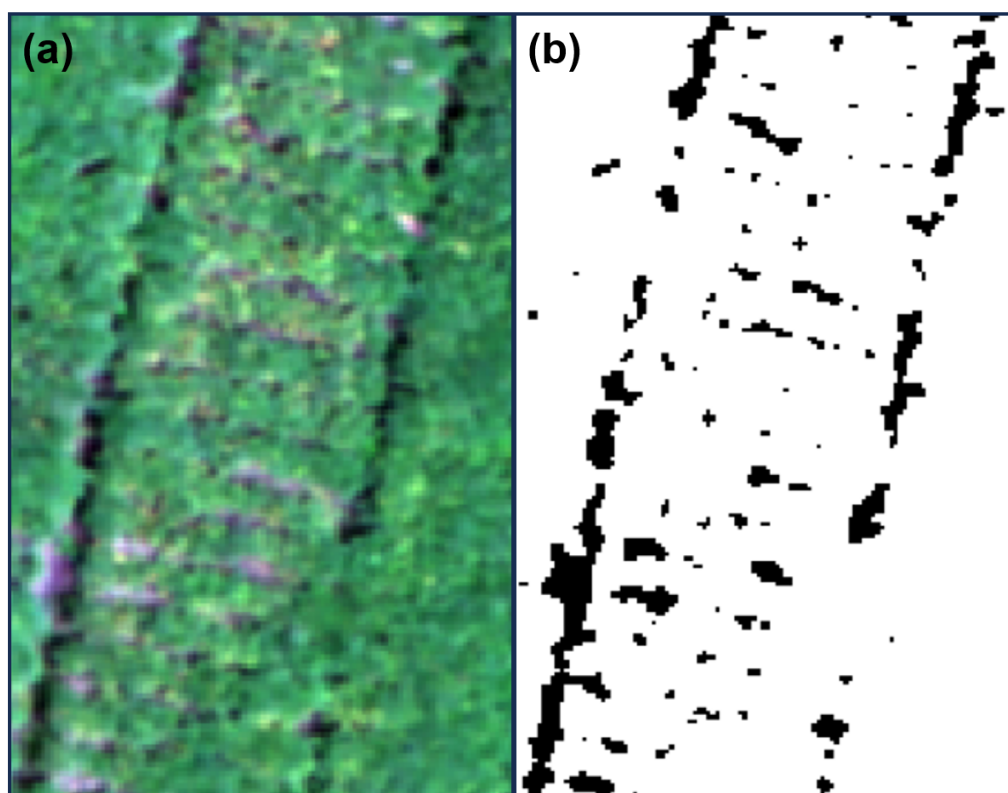


Figure 3 Mask extraction for separating the canopy of *L. chuanxiong* from the background. (A) RGB image; (B) mask.

[Full-size](#) DOI: [10.7717/peerj.19264/fig-3](https://doi.org/10.7717/peerj.19264/fig-3)

Model construction and related data processing were performed in Python 3.11. Four machine learning ensemble algorithms, Random Forest (RF), adaptive boosting (AdaBoost), Gradient Boosting Decision Tree (GBDT) and extreme gradient boosting (XGBoost), were used to construct regression prediction models of FW, DW, FA, Z-L and SA contents, respectively. The quality discriminant models were based on the standard set by the Pharmacopoeia of the People's Republic of China (2020) (ChP 2020) ([Chinese National Pharmacopoeia Commission, 2020](#)) that the FA content of chuanxiong rhizoma should not be less than 0.10% (1.0 mg g^{-1}). Considering that there may be a large difference between the sample sizes that meet the criteria or not, the Border line SMOTE was used to oversample the categories with a small sample size to obtain balanced data. Each data set was randomly divided into a training set and a testing set with a ratio of 7:3 for model building. In the process of model construction, each algorithm determined the optimal model parameters through grid search and five-fold cross validation.

Three indicators were used to evaluate the effectiveness and stability of the model. Coefficient of determination (R^2) is used to measure the fitting degree of the regression model, and the value was 0~1. The larger the value, the better the fitting degree of the model. Normalized root mean squared error (NRMSE) and mean absolute percentage error (MAPE) are indicators of the gap between the predicted value and the true value. The former

Table 1 The characteristic variables used in this study.

Variable	Name	Formula	Reference
R	Red	Band1	\
G	Green	Band2	\
B	Blue	Band3	\
RE	Red Edge	Band4	\
NIR	Near Infrared	Band5	\
r	Normalized Red	$R/(R+G+B)$	<i>Tarbell & Reid (1991)</i>
g	Normalized Green	$G/(R+G+B)$	<i>Tarbell & Reid (1991)</i>
b	Normalized Blue	$B/(R+G+B)$	<i>Tarbell & Reid (1991)</i>
RBRI	Red Blue Ratio Index	R/B	<i>Wu et al. (2023)</i>
GLI	Green Leaf Index	$(2G-R-B)/(2G+R+B)$	<i>Louhaichi, Borman & Johnson (2001)</i>
MSRI	Modified Simple Ratio Index	$(NIR/RE-1)/\sqrt{NIR/RE+1}$	<i>Chen (1996)</i>
OSAVI	Optimized Soil Adjusted Vegetation Index	$(NIR-R)/(NIR+R+0.16)$	<i>Rondeaux, Steven & Baret (1996)</i>
RECI	Red Edge Chlorophyll II Index	$NIR/RE-1$	<i>Gitelson et al. (2005)</i>
RERDVI	Red Edge Re-normalized Different Vegetation Index	$(NIR-RE)/\sqrt{(NIR+RE)}$	<i>Roujean & Breon (1995)</i>
LCI	Leaf Chlorophyll II Index	$(NIR-RE)/(NIR+R)$	<i>Datt (1999)</i>
EVI	Enhanced Vegetation Index	$2.5(NIR-R)/(NIR+6R-7.5B+1)$	<i>Justice et al. (1998)</i>
WI	Woebbecke Index	$(G-B)/(R-G)$	<i>Woebbecke et al. (1995)</i>
GRVI	Green Ratio Vegetation Index	NIR/G	<i>Buschmann & Nagel (1993)</i>
RVI	Ratio Vegetation Index	NIR/R	<i>Jordan (1969)</i>
RERVI	Red Edge Ratio Vegetation Index	NIR/RE	<i>Jasper, Reusch & Link (2009)</i>
BRVI	Blue Ratio Vegetation Index	NIR/B	<i>Kandylakis & Karantzalos (2016)</i>
NGRVI	Normalized Green-Red Vegetation Index	$(G-R)/(G+R)$	<i>Gitelson et al. (2002)</i>
NDVI	Normalized Difference Vegetation Index	$(NIR-R)/(NIR+R)$	<i>Rouse et al. (1974)</i>

is more sensitive to large errors because of the root mean square calculation, while the latter weights all errors equally. Both of them can compare the effects of models composed of different data sets. The performance of the model in the independent validation set was evaluated using NRMSE and MAPE. The calculation formula is as follows (*Costa et al., 2022; Zhu et al., 2023*):

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

$$NRMSE = (1/\bar{y}) \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / n} \cdot 100\% \quad (2)$$

$$MAPE = (1/n) \sum_{i=1}^n |y_i - \hat{y}_i| / y_i \cdot 100\% \quad (3)$$

where n is the sample size of the input data, y_i is the true target value, \bar{y} is the average value of the corresponding true target value, and is \hat{y}_i the predicted target value calculated by the corresponding regression model.

The performance of the quality discriminant model was evaluated using the following indicators: Accuracy is used to evaluate the overall discriminant effect of the discriminant model. Precision, recall and F1-score derived from the confusion matrix can explain the discriminant effect in more detail. The area under curve (AUC) is the area under the receiver operating characteristic (ROC) curve, and the value is 0~1. The larger the AUC, the better the discriminant effect of the model. The performance of the model in the independent validation set was evaluated using accuracy and AUC. The formula is as follows ([Zhang et al., 2022a](#); [Zhang et al., 2022b](#)):

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (4)$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (5)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (6)$$

$$\text{F1 - score} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall}) \quad (7)$$

where TP is the number of samples whose instances are positive and predicted to be positive, TN is the number of samples whose instances are negative and predicted to be negative, FN is the number of samples whose instances are positive and predicted to be negative, FP is the number of samples whose instances are negative and predicted to be positive.

RESULTS

Correlation analysis of VIs

According to the results of the correlation analysis ([Fig. 4](#)) combined with the feature screening method in VIs selection and models construction, the input characteristic variables of the FW prediction model in D1 were red edge (RE), optimized soil adjusted vegetation index (OSAVI), green (G), Woebbecke Index (WI), blue (B), modified simple ration index (MSRI), Green Leaf Index (GLI), normalized blue (b) and red blue ratio index (RBRI), while the input characteristic variables of the DW prediction model were RE, OSAVI, green (G), WI, B, normalized green-red vegetation index (NGRVI), radio vegetation index (RVI), GLI, b, red (R), and RBRI. In addition, red edge chlorophyll II index (RECI), red edge re-normalized different vegetation index (RERDVI), LCI, enhanced vegetation index (EVI), RBRI, WI, G, green ration vegetation index (GRVI), B and r were the input characteristic variables of the FW and DW prediction models in D2.

The input characteristic variables of the FA content prediction model and the quality discriminant model in D1 were b, RBRI and B, while those in D2 were RE and red edge ratio vegetation index (RERVI). The input variables of the SA content prediction model in D1 were b, g, B, BRVI, R and r, while those in D2 were RE, GLI, g, RVI, and RERVI. The input variables of the Z-L content prediction model in D1 were G, RE, OSAVI, WI, RBRI, B and NGRVI, while those in D2 were G, B, RECI, NIR, RERDVI and WI.

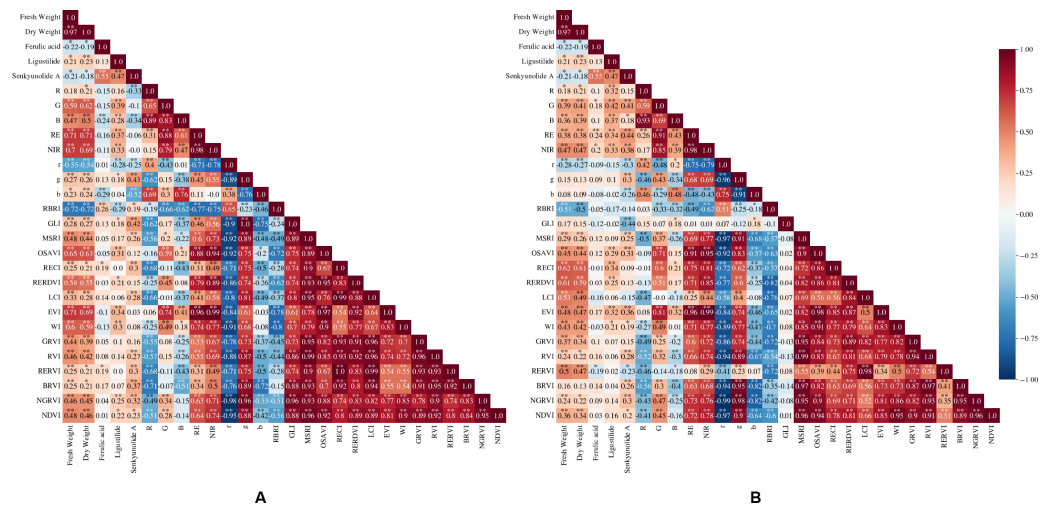


Figure 4 Correlation coefficients between variables and the FW, DW and the content of FA, Z-L and SA in D1 and D2. (A) D1; (B) D2; $p < 0.05$; **, $p < 0.01$.

Full-size [DOI: 10.7717/peerj.19264/fig-4](https://doi.org/10.7717/peerj.19264/fig-4)

Table 2 Evaluation of different FW prediction models.

Period	Model	Training set			Test set		
		R ²	NRMSE	MAPE	R ²	NRMSE	MAPE
D1	XGBoost	0.89	14.15%	14.37%	0.82	19.19%	16.83%
	GBDT	0.78	20.95%	15.49%	0.76	23.93%	16.43%
	AdaBoost	0.79	19.21%	18.45%	0.76	23.43%	23.48%
	RF	0.84	17.22%	17.54%	0.80	20.11%	17.80%
D2	XGBoost	0.87	15.31%	15.68%	0.71	24.38%	18.40%
	GBDT	0.81	19.21%	14.01%	0.72	24.67%	17.22%
	AdaBoost	0.72	23.12%	21.16%	0.69	25.91%	18.11%
	RF	0.72	22.40%	22.04%	0.71	25.36%	17.91%

Notes.

D1 and D2 represent April 27 and May 8 respectively.

Different prediction models

Yield prediction models

The R² (Table 2) of the four algorithm models in the FW training set and the test set data in both periods could reach more than 0.68. Overall, the model with the best prediction accuracy in both periods was XGBoost, the R² of the training set was close to 0.90, and the NRMSE and MAPE were less than 20.00%. The XGBoost over-fitting degree in D1 was lower than that in D2, whose R² on the testing set was 19.54% lower than that on the training set.

The difference between the DW prediction models (Table 3) was similar to that of FW, and the best prediction model was the XGBoost model in D1. The R² on the training and test sets were both greater than 0.80, and the NRMSE and MAPE were within the range from 19.00% to 21.00%. In the model in D2, the XGBoost model performed best in the

Table 3 Evaluation of different DW prediction models.

Period	Model	Training set			Test set		
		R ²	NRMSE	MAPE	R ²	NRMSE	MAPE
D1	XGBoost	0.82	20.04%	20.85%	0.81	20.94%	19.02%
	GBDT	0.77	23.40%	17.52%	0.76	25.16%	20.54%
	AdaBoost	0.79	21.54%	25.03%	0.78	21.94%	25.44%
	RF	0.80	20.82%	22.96%	0.79	21.88%	22.52%
D2	XGBoost	0.82	19.77%	22.49%	0.70	27.15%	22.15%
	GBDT	0.81	21.07%	14.92%	0.74	24.71%	22.18%
	AdaBoost	0.71	25.30%	25.78%	0.68	28.53%	22.36%
	RF	0.70	26.71%	22.27%	0.69	24.14%	30.40%

Notes.

D1 and D2 represent April 27 and May 8 respectively.

Table 4 Evaluation in D1 prediction model of different active component contents.

Period	Model	Training set			Test set		
		R ²	NRMSE	MAPE	R ²	NRMSE	MAPE
FA	XGBoost	0.53	18.88%	15.86%	0.39	20.95%	18.01%
	GBDT	0.81	12.31%	7.94%	0.44	20.30%	17.27%
	AdaBoost	0.58	17.64%	16.38%	0.43	20.23%	17.75%
	RF	0.45	20.63%	18.75%	0.12	24.56%	20.04%
Z-L	XGBoost	0.34	19.01%	17.69%	0.22	23.59%	25.12%
	GBDT	0.59	15.23%	13.49%	0.21	23.73%	23.59%
	AdaBoost	0.49	17.52%	17.31%	0.25	19.75%	21.13%
	RF	0.79	10.70%	9.81%	0.28	22.74%	23.42%
SA	XGBoost	0.81	15.01%	13.34%	0.57	28.64%	27.74%
	GBDT	0.85	13.36%	12.68%	0.68	24.85%	25.27%
	AdaBoost	0.74	17.72%	19.21%	0.70	22.25%	25.80%
	RF	0.89	11.24%	10.63%	0.63	27.00%	26.09%

Notes.

FA, Z-L and SA represent ferulic acid, Z-ligustilide and senkyunolide A, respectively.

training set with an R² of 0.82, but the over-fitting situation was still higher than other algorithm models.

Active component contents prediction models

The effects of FA, Z-L and SA content prediction models in D1 are shown in Table 4. Among the FA content prediction models in D1, the training set R² of GBDT was the largest, but the R² of the test set decreases greatly, up to 45.68%, while the AdaBoost model has a small degree of over-fitting while ensuring a low error. The degree of over-fitting of the Z-L content prediction models was high, and the AdaBoost had the lowest degree of over-fitting that the difference between the R² values of the training set and the test set was 0.24. The R² of all SA content prediction models in D1 performed well and the degree of over-fitting was low. Among them, the RF training set had the highest R² of 0.89, while the R² of the AdaBoost test set was higher and the degree of over-fitting was lower.

Table 5 Evaluation in D2 prediction model of different active component contents.

Period	Model	Training set			Test set		
		R ²	NRMSE	MAPE	R ²	NRMSE	MAPE
FA	XGBoost	0.68	15.75%	13.77%	0.35	20.59%	17.95%
	GBDT	0.45	20.74%	19.08%	0.24	20.18%	19.60%
	AdaBoost	0.35	23.76%	19.34%	0.30	20.15%	16.76%
	RF	0.78	13.00%	11.40%	0.33	20.78%	18.74%
Z-L	XGBoost	0.69	13.00%	12.45%	0.34	21.97%	24.38%
	GBDT	0.63	14.25%	13.43%	0.31	22.32%	23.67%
	AdaBoost	0.51	17.25%	15.39%	0.39	18.06%	20.70%
	RF	0.82	10.62%	10.45%	0.32	19.62%	16.93%
SA	XGBoost	0.86	13.38%	13.10%	0.75	21.17%	21.49%
	GBDT	0.90	11.54%	6.56%	0.76	21.59%	17.56%
	AdaBoost	0.74	19.06%	17.95%	0.74	20.58%	17.97%
	RF	0.89	11.68%	10.81%	0.67	24.44%	23.76%

Notes.

FA, Z-L and SA represent ferulic acid, Z-ligustilide and senkyunolide A, respectively.

The effect of the three active component content prediction models in D2 is shown in Table 5. The FA and Z-L content prediction models of D2 had higher degree of over-fitting, except for AdaBoost. However, each training set R² of AdaBoost was lower, 0.35 and 0.51, respectively. The R² of the SA content prediction model in D2 was higher, and GBDT was the best, whose R² of the training set and the test set were 0.90 and 0.76, respectively. While the AdaBoost, with training set R² of 0.74 for both training and test sets, had the lowest degree of over-fitting.

Chuanxiong rhizoma quality discriminant models

ChP 2020 stipulates that the FA content in chuanxiong rhizoma should not be less than 0.10% (1.00 mg g⁻¹) (*Chinese National Pharmacopoeia Commission, 2020*). Therefore, we constructed quality discriminant models (Table 6) to evaluate whether the FA content reached the standard. The accuracy of XGBoost and GBDT models in D1 reached 0.9020, and the AUC was 0.9031 and 0.9095, respectively. The Accuracy of XGBoost, AdaBoost and RF constructed by D2 data was better and all of them were 0.8431. Among them, the AUC of RF was larger so the discriminative efficiency was the best.

Independent validation

Independent validation of yield prediction models

Since D2 and D3 are close to the harvesting periods of Shiyang Town and Aoping Town respectively, the VIs of D3 were used as the characteristic variable to be input according to the requirements of the yield prediction models in D2 to evaluate the performance of models on the independent validation set. As shown in Fig. 5, the relationship between the predicted value and the actual value is obtained. When applied to the independent validation set, each model underestimates FW and DW. Among them, RF was the best among the FW prediction models with NRMSE of 23.76%, and MAPE of 14.75%, while

Table 6 Evaluation of different discriminant models.

Period	Model	Whether achieved	Accuracy	Precision	Recall	F1 Score	AUC
D1	XGBoost	No	0.9020	0.8571	0.9600	0.9057	0.9031
		Yes		0.9565	0.8462	0.8900	
	GBDT	No	0.9020	0.8333	0.9524	0.8889	0.9095
		Yes		0.9630	0.8667	0.9123	
	AdaBoost	No	0.8235	0.7308	0.9048	0.8085	0.8357
		Yes		0.9200	0.7667	0.8364	
D2	RF	No	0.8431	0.7778	0.9130	0.8400	0.8494
		Yes		0.9167	0.7857	0.8462	
	XGBoost	No	0.8431	0.8387	0.8966	0.8667	0.8346
		Yes		0.8500	0.7727	0.8095	
	GBDT	No	0.8235	0.8636	0.7600	0.8085	0.8223
		Yes		0.7931	0.8846	0.8364	
	AdaBoost	No	0.8431	0.8276	0.8889	0.8571	0.8403
		Yes		0.8636	0.7917	0.8261	
	RF	No	0.8431	0.7917	0.8636	0.8261	0.8456
		Yes		0.8889	0.8276	0.8571	

Notes.

D1 and D2 represent April 27 and May 8 respectively.

AdaBoost was the best among the DW prediction models with NRMSE of 34.60% and MAPE of 21.73%.

Independent validation of active component content prediction models

The prediction models of FA, Z-L and SA content in D2 were applied to predict the content of active components in the independent validation set, and the results were shown in Fig. 6. AdaBoost performs best in predicting FA content. Compared to the other models, the RF has a more concentrated distribution of the predicted values and smaller relative errors. The predicted Z-L content of each model was generally lower than the true values, among which the RF's predicted values were the most concentrated with NRMSE of 34.35% and MAPE of 23.40%. The predicted SA content of each model was generally lower than the true values. Among them, XGBoost was the best at predicting with NRMSE of 45.26% and MAPE of 25.48%.

Independent validation of chuanxiong rhizoma quality discriminant models

To verify the effect of the constructed quality discriminant model based on the FA content standard of chuanxiong rhizoma stipulated in ChP 2020, the corresponding characteristic variables were entered into the quality discriminant model of D2, and the discriminant results were shown in Fig. 7. The best performing models were XGBoost and RF, with accuracy and AUC of 0.7083 and 0.7214, respectively, which had effective discriminative ability.

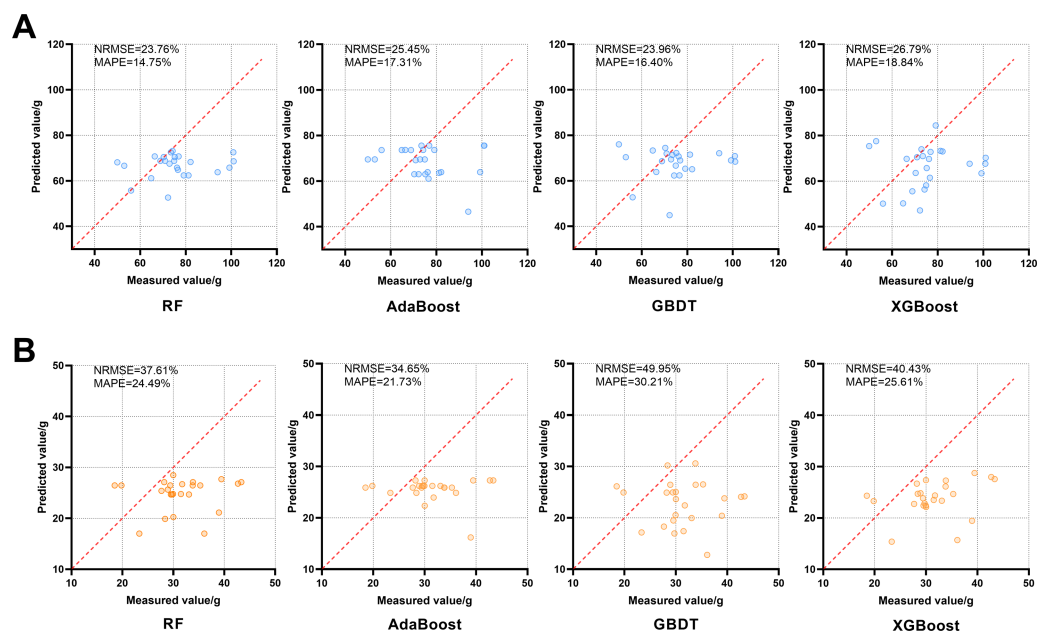


Figure 5 Effects of four FW and DW prediction models, RF, AdaBoost, GBDT and XGBoost, on the independent validation set. The red dotted line in each image is a 1:1 line. (A) FW; (B) DW.

Full-size [DOI: 10.7717/peerj.19264/fig-5](https://doi.org/10.7717/peerj.19264/fig-5)

DISCUSSION

Different sensitivity of variables

The four machine learning algorithms selected in this study could all use the variables selected in the two periods to train effective prediction models for predicting the yield and quality of *L. chuanxiong*. There were differences in the input variables of the yield prediction models in the two periods (Fig. 4), but the most advanced variables (RE, OSAVI, RECI, RERDVI, etc.) were related to the red edge and near-infrared bands, which indicated that the variables based on the reflectance of the red edge and near-infrared bands could better reflect the FW and DW of rhizome of *L. chuanxiong*. This is consistent with the results obtained by Mutanga, Adam & Cho (2012) who constructed a wetland biomass estimation model. In addition, the input variables of the three active component content prediction models selected in the two periods were quite different (Fig. 4). For example, the top input variables of the FA and SA content prediction models in the D1 variables (b, B, BRVI, RBRI, etc.) are mainly related to the blue band, while the D2 variables (RE, RERVI, etc.) are mainly related to the red edge band. The difference in this short-term interval may be due to the fact that April to May is the period when *L. chuanxiong* changes from the end of the vigorous growth period of the aboveground part to the rapid expansion of the underground part. The growth rate of the new leaves on the aboveground part gradually stops, the old leaves gradually wither and the leaf water content changes greatly (Zhao, Fu & Fan, 2008), resulting in a certain change in the canopy nutrient structure, which affected the reflectance of each band. However, the sensitivity of *L. chuanxiong* to the reflectance of

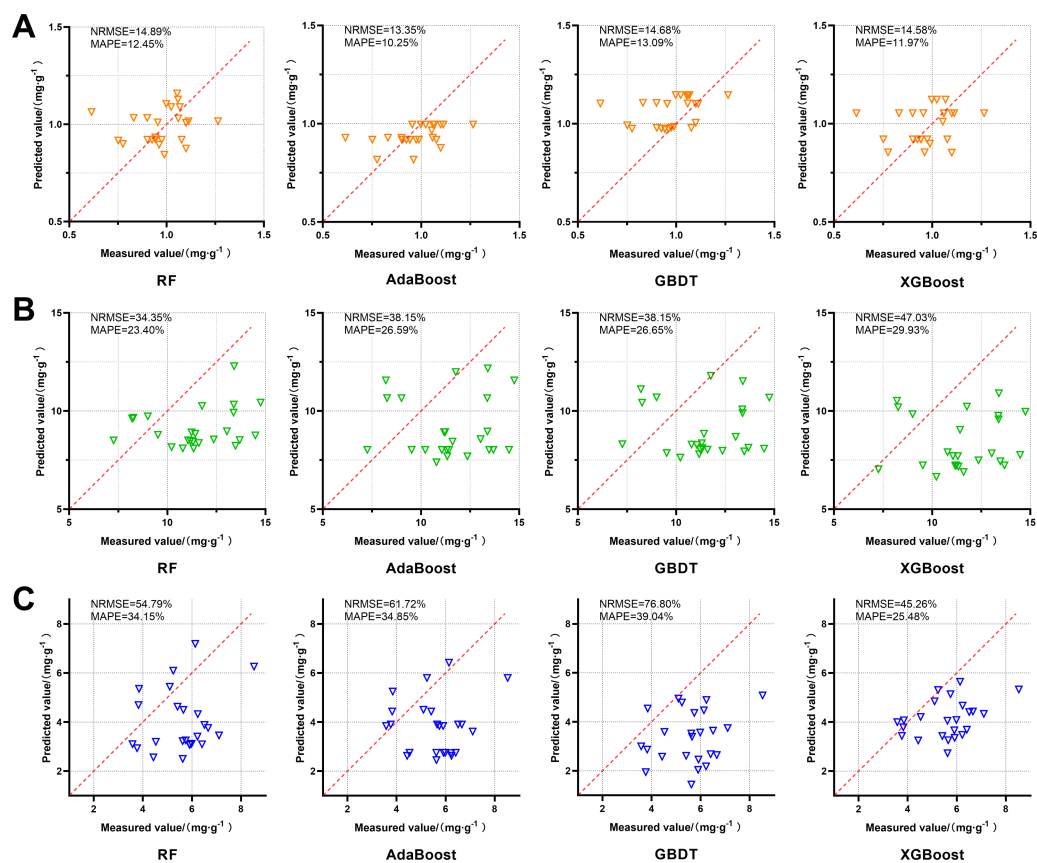


Figure 6 Effects of four FA, Z-L and SA prediction models, RF, AdaBoost, GBDT and XGBoost, on the independent validation set. The red dotted line in each image is a 1:1 line. (A) FA; (B) Z-L; (C) SA.

Full-size [DOI: 10.7717/peerj.19264/fig-6](https://doi.org/10.7717/peerj.19264/fig-6)

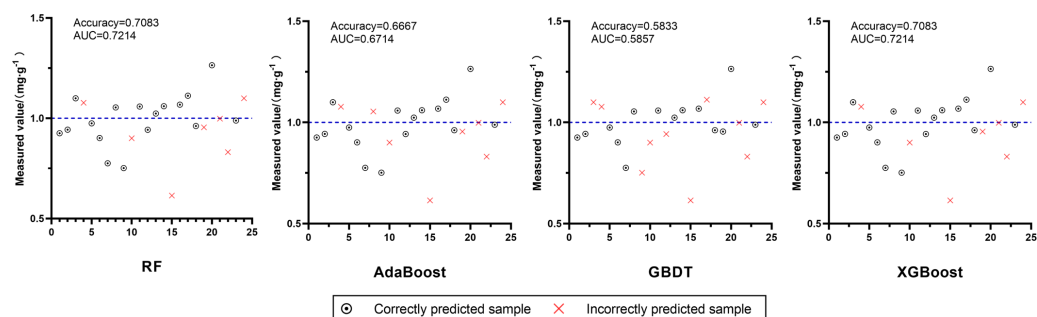


Figure 7 Effects of four quality discriminant models, RF, AdaBoost, GBDT and XGBoost, on independent validation set. The blue dotted line in each image is Measured value = 1.0 mg g⁻¹.

Full-size [DOI: 10.7717/peerj.19264/fig-7](https://doi.org/10.7717/peerj.19264/fig-7)

different bands in this study is different from the results of *Fan et al. (2024)*. They found that although there were differences in the variables sensitive to potato yield at different growth stages, the blue band was always closely related to yield, and the near-infrared band

could reflect the difference in growth between plants. This may be due to the difference in different crops with the different sensitivity of variables.

Different performances of models

The XGBoost performed best in predicting the yield data of Shiyang Town in both periods, and the degree of over-fitting was low (Tables 2 and 3). Similarly, XGBoost performed better than other models in studies such as [Shahi et al. \(2023\)](#), which used UAV images of five growth stages of peanut to predict yield, and [Wang et al. \(2023b\)](#), which used UAV multispectral data to evaluate the above-ground biomass of camphor trees. However, in this study, RF and AdaBoost had the best predictive effect on FW and DW in the independent validation set, and XGBoost performed slightly worse (Fig. 5). This may be because the RF and AdaBoost have fewer hyperparameters, and the model is simpler and has better generalization ability than XGBoost. In addition, the performance of the all models on the independent validation set shows that the predicted value is generally lower than the true value. This is contrary to the results of [Liu et al. \(2019\)](#), who applied the winter oilseed rape biomass prediction models to the independent validation data set of another year's and they believe that this is due to the large climatic differences between the experimental sites in different years. The collection period of the independent validation set of this study is similar to that of the data set collected to construct the models, therefore, the main difference in the sets reflected in the spatial difference. The low predicted values may be mainly related to the differences between the soil ([Peng et al., 2021](#)) and ecological environmental factors ([Yin et al., 2012](#)) in the two locations.

In the performance of predicting the content of the three active components, AdaBoost had a lower degree of over-fitting in D1 and D2. This is similar to the results of [Yoon et al. \(2023\)](#) using different algorithms combined with hyperspectral prediction of mustard metabolites. In that study, AdaBoost predicts secondary metabolites such as phenolics and flavonoids better than XGBoost and LightGBM. In the performance of the independent validation set (Fig. 6), the best models for predicting FA and Z-L content were AdaBoost (NRMSE = 13.35%, MAPE = 10.25%) and RF (NRMSE = 34.35%, MAP = 23.40%), respectively, and they had lower relative errors. The best model for SA content prediction is XGBoost, but there is still a large relative error (NRMSE = 45.26%, MAPE = 25.48%). The results also reflect the advantages of the model with fewer hyperparameters in universality, and show that different models have different performances in predicting the content of different active components.

In addition, we also constructed the quality discriminant models of chuanxiong rhizoma according to the standard that the FA content should not be less than 0.10% (1.00 mg g⁻¹), which is stipulated in ChP 2020. The accuracy and AUC of the models in each period (Table 6) were greater than 0.8 in the test set. Among them, the Accuracy and AUC of XGBoost and GBDT models in D1 could reach more than 0.9, and the those of XGBoost, AdaBoost and RF models in D2 could reach more than 0.84. The models in D2 were applied to the independent validation set, and the effects of XGBoost and RF performed best that the Accuracy and AUC were both 0.7083 and 0.7214, respectively (Fig. 7). Although the performance of the models decreased to some extent in the independent validation set,

they still had effective discriminatory ability, indicating that it was feasible to use VIs to construct a discriminant model based on FA content to evaluate the quality of chuanxiong rhizoma.

CONCLUSIONS

In this study, the models for predicting the fresh weight, dry weight of *L. chuanxiong* rhizome and the contents of ferulic acid, Z-ligustilide and senkyunolide A in chuanxiong rhizoma were constructed on the basis of UAV multispectral images. And the quality discriminant models of chuanxiong rhizoma based on the standard that the FA content should not be less than 0.10% (1.00 mg g^{-1}) stipulated in the Pharmacopoeia of the People's Republic of China (2020) were constructed for the first time. In addition, the verification of the models on the independent validation set was completed, and the results proved that each prediction model had a certain degree of universality. According to the results of this study, we believe that UAV multispectral has great potential in non-destructively predicting the yield and quality of medicinal plants before harvesting. It can further provide a reasonable reference for the appropriate specification and classification of medicinal materials in production, as well as the management methods of processing, transporting and storing, so as to ensure the quality of medicinal materials and make more effective use of the resources.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was funded by the Science and Technology Department of Sichuan Province (No. 2021YFS0045) and Chengdu University of Traditional Chinese Medicine (No. XJ2023001601). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Science and Technology Department of Sichuan Province: No. 2021YFS0045.
Chengdu University of Traditional Chinese Medicine: No. XJ2023001601.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Yun-Fan Li conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Chen Wu performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Hong-Mei Jia performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

- Xi Chen performed the experiments, prepared figures and/or tables, and approved the final draft.
- Jin-Niu Xing performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Wei-Ping Gao performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Zhu-Yun Yan conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The raw yield and quality-related data are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.19264#supplemental-information>.

REFERENCES

- Alckmin GT, Lucieer A, Rawnsley R, Kooistra L. 2022.** Perennial ryegrass biomass retrieval through multispectral UAV data. *Computers and Electronics in Agriculture* **193**:106574 DOI [10.1016/j.compag.2021.106574](https://doi.org/10.1016/j.compag.2021.106574).
- Buschmann C, Nagel E. 1993.** In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *International Journal of Remote Sensing* **14**:711–722 DOI [10.1080/01431169308904370](https://doi.org/10.1080/01431169308904370).
- Cen H, Wan L, Zhu J, Li Y, Li X, Zhu Y, Weng H, Wu W, Yin W, Xu C, Bao Y, Feng L, Shou J, He Y. 2019.** Dynamic monitoring of biomass of rice under different nitrogen treatments using a lightweight UAV with dual image-frame snapshot cameras. *Plant Methods* **15**:32 DOI [10.1186/s13007-019-0418-8](https://doi.org/10.1186/s13007-019-0418-8).
- Chancia R, Van Aardt J, Pethybridge S, Cross D, Henderson J. 2021.** Predicting table beet root yield with multispectral UAS imagery. *Remote Sensing* **13**:2180 DOI [10.3390/rs13112180](https://doi.org/10.3390/rs13112180).
- Chen JM. 1996.** Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing* **22**:229–242 DOI [10.1080/07038992.1996.10855178](https://doi.org/10.1080/07038992.1996.10855178).
- Chen H, Zhao Y, Qin G, Bi Y, Yue G, Zhang M, Chang X, Qiu X, Luo L, Yang C. 2022.** Antifungal effects and active components of *Ligusticum chuanxiong*. *Molecules* **27**:4589 DOI [10.3390/molecules27144589](https://doi.org/10.3390/molecules27144589).
- Chinese National Pharmacopoeia Commission. 2020.** *Pharmacopoeia of the People's Republic of China: Part I*. Beijing: China Medical Science Press 42 (in Chinese).
- Costa L, Kunwar S, Ampatzidis Y, Albrecht U. 2022.** Determining leaf nutrient concentrations in citrus trees using UAV imagery and machine learning. *Precision Agriculture* **23**:854–875 DOI [10.1007/s11119-021-09864-1](https://doi.org/10.1007/s11119-021-09864-1).

- Datt B. 1999. A new reflectance index for remote sensing of chlorophyll content in higher plants: tests using eucalyptus leaves. *Journal of Plant Physiology* 154:30–36 DOI 10.1016/S0176-1617(99)80314-9.
- Fan Y, Liu Y, Yue J, Jin X, Chen R, Bian M, Ma Y, Yang G, Feng H. 2024. Estimation of potato yield using a semi-mechanistic model developed by proximal remote sensing and environmental variables. *Computers and Electronics in Agriculture* 223:109117 DOI 10.1016/j.compag.2024.109117.
- Gitelson AA, Kaufman YJ, Stark R, Rundquist D. 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment* 80:76–87 DOI 10.1016/S0034-4257(01)00289-9.
- Gitelson AA, Viña A, Ciganda V, Rundquist DC, Arkebauer TJ. 2005. Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters* 32:2005GL022688 DOI 10.1029/2005GL022688.
- Iqbal F, Lucieer A, Barry K. 2020. Prediction of poppy thebaine alkaloid concentration using UAS remote sensing. *Precision Agriculture* 21:1045–1056 DOI 10.1007/s11119-020-09707-5.
- Jasim A, Zaeen A, Sharma LK, Bali SK, Wang C, Buzza A, Alyokhin A. 2020. Predicting phosphorus and potato yield using active and passive sensors. *Agriculture-Basel* 10:564 DOI 10.3390/agriculture10110564.
- Jasper J, Reusch S, Link A. 2009. Active sensing of the N status of wheat using optimized wavelength combination: impact of seed rate, variety and growth stage. In: *Precision agriculture '09*. Wageningen: Wageningen Academic, 21–30 DOI 10.3920/9789086866649_003.
- Jordan CF. 1969. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* 50:663–666 DOI 10.2307/1936256.
- Justice CO, Vermote E, Townshend JRG, Defries R, Roy DP, Hall DK, Salomonson VV, Privette JL, Riggs G, Strahler A, Lucht W, Myneni RB, Knyazikhin Y, Running SW, Nemani RR, Zhengming W, Huete AR, Van Leeuwen W, Wolfe RE, Giglio L, Muller J, Lewis P, Barnsley MJ. 1998. The moderate resolution imaging spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing* 36:1228–1249 DOI 10.1109/36.701075.
- Kandylakis Z, Karantzas K. 2016. Precision viticulture from multitemporal, multispectral very high resolution satellite data. *ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLI-B* 8:919–925 DOI 10.5194/isprsarchives-XLI-B8-919-2016.
- Karunathilake EMBM, Le AT, Heo S, Chung YS, Mansoor S. 2023. The path to smart farming: innovations and opportunities in precision agriculture. *Agriculture* 13:1593 DOI 10.3390/agriculture13081593.
- Li Q, He P, Zhang D, Pei Z, Chen Y, Li Z, Zheng C. 2024. Analysis of industrial status and development strategy of *Ligusticum chuanxiong* Hort. *Chinese Traditional and Herbal Drugs* 55:2771–2783 (in Chinese).

- Li D, Long Y, Yu S, Shi A, Wan J, Wen J, Li X, Liu S, Zhang Y, Li N, Zheng C, Yang M, Shen L. 2022. Research advances in cardio-cerebrovascular diseases of *Ligusticum chuanxiong* hort. *Frontiers in Pharmacology* 12:832673 DOI 10.3389/fphar.2021.832673.
- Li W, Tang Y, Chen Y, Duan J-A. 2012. Advances in the chemical analysis and biological activities of Chuanxiong. *Molecules* 17:10614–10651 DOI 10.3390/molecules170910614.
- Li W, Wang K, Han G, Wang H, Tan N, Yan Z. 2023. Integrated diagnosis and time-series sensitivity evaluation of nutrient deficiencies in medicinal plant (*Ligusticum chuanxiong* Hort.) based on UAV multispectral sensors. *Frontiers in Plant Science* 13:1092610 DOI 10.3389/fpls.2022.1092610.
- Liu Y, Liu S, Li J, Guo X, Wang S, Lu J. 2019. Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Computers and Electronics in Agriculture* 166:105026 DOI 10.1016/j.compag.2019.105026.
- Louhaichi M, Borman MM, Johnson DE. 2001. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto International* 16:65–70 DOI 10.1080/10106040108542184.
- Lyu H, Grafton M, Ramilan T, Irwin M, Wei H-E, Sandoval E. 2023. Using remote and proximal sensing data and vine vigor parameters for non-destructive and rapid prediction of grape quality. *Remote Sensing* 15:5412 DOI 10.3390/rs15225412.
- Maimaitijiang M, Ghulam A, Sidike P, Hartling S, Maimaitiyiming M, Peterson K, Shavers E, Fishman J, Peterson J, Kadam S, Burken J, Fritsch F. 2017. Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS Journal of Photogrammetry and Remote Sensing* 134:43–58 DOI 10.1016/j.isprsjprs.2017.10.011.
- Mutanga O, Adam E, Cho MA. 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation* 18:399–406 DOI 10.1016/j.jag.2012.03.012.
- Peng F, Wu Y, Tao S, Yuan C, Shi T, Chen Y, Sha X, Liao X, Liao X, Li S, Zhang C. 2021. Spatial variability and abundance and deficiency evaluation of soil nutrients in the main producing areas of *Ligusticum chuanxiong* Hort. *Modern Chinese Medicine* 23:666–673 (in Chinese).
- Radočaj D, Šiljeg A, Marinović R, Jurišić M. 2023. State of major vegetation indices in precision agriculture studies indexed in web of science: a review. *Agriculture* 13:707 DOI 10.3390/agriculture13030707.
- Rondeaux G, Steven M, Baret F. 1996. Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment* 55:95–107 DOI 10.1016/0034-4257(95)00186-7.
- Roujean J-L, Breon F-M. 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment* 51:375–384 DOI 10.1016/0034-4257(94)00114-3.
- Rouse JW, Haas RW, Schell JA, Deering DW, Harlan JC. 1974. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation.

- NASA/GSFCT type III final report. Greenbelt, MD, USA, 164. Available at <https://ntrs.nasa.gov/citations/19750020419>.
- Selvaraj MG, Valderrama M, Guzman D, Valencia M, Ruiz H, Acharjee A. 2020. Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (*Manihot esculenta* Crantz). *Plant Methods* 16:87 DOI 10.1186/s13007-020-00625-1.
- Shahhosseini M, Martinez-Feria RA, Hui G, Archontoulis SV. 2019. Maize yield and nitrate loss prediction with machine learning algorithms. *Environmental Research Letters* 14:124026 DOI 10.1088/1748-9326/ab5268.
- Shahi TB, Xu C-Y, Neupane A, Fleischfresser DB, O'Connor DJ, Wright GC, Guo W. 2023. Peanut yield prediction with UAV multispectral imagery using a cooperative machine learning approach. *Era* 31:3343–3361 DOI 10.3934/era.2023169.
- Shendryk Y, Sofonia J, Garrard R, Rist Y, Skocaj D, Thorburn P. 2020. Fine-scale prediction of biomass and leaf nitrogen content in sugarcane using UAV LiDAR and multispectral imaging. *International Journal of Applied Earth Observation and Geoinformation* 92:102177 DOI 10.1016/j.jag.2020.102177.
- Song G, Yang Z, Sun W, Gao A, Meng Y, Chen Y. 2024. Current situation and development trend of mechanized planting. *Equipment for Chinese Forestry Machinery & Woodworking Equipment* 52:13–19 (in Chinese) DOI 10.13279/j.cnki.fmwe.2024.0048.
- Tarbell KA, Reid J. 1991. Computer vision system for characterizing corn growth and development. *Transactions of the ASAE* 34:2245–2255 DOI 10.13031/2013.31864.
- Tedesco D, De Almeida Moreira BR, Barbosa Junior MR, Papa JP, Da Silva RP. 2021. Predicting on multi-target regression for the yield of sweet potato by the market class of its roots upon vegetation indices. *Computers and Electronics in Agriculture* 191:106544 DOI 10.1016/j.compag.2021.106544.
- Upreti D, Huang W, Kong W, Pascucci S, Pignatti S, Zhou X, Ye H, Casa R. 2019. A comparison of hybrid machine learning algorithms for the retrieval of wheat biophysical variables from sentinel-2. *Remote Sensing* 11:481 DOI 10.3390/rs11050481.
- Wang F, Chen S, Liu Y, Chen H, Hu Y, Li S, Chen L. 2023a. Progress and prospect of chinese herbal medicine planting industry in China. *Modern Chinese Medicine* 25:1163–1171 (in Chinese) DOI 10.13313/j.issn.1673-4890.20230202002.
- Wang Q, Lu X, Zhang H, Yang B, Gong R, Zhang J, Jin Z, Xie R, Xia J, Zhao J. 2023b. Comparison of machine learning methods for estimating leaf area index and aboveground biomass of cinnamomum camphora based on UAV multispectral remote sensing data. *Forests* 14:1688 DOI 10.3390/f14081688.
- Wang H, Zhang X, Wang J, Shi T, Li M, Wan X, Huang L. 2022. Statistical analysis of the planting area of Chinese medicinal materials in 2020. *China Food & Drug Administration Magazine* 1:4–9 (in Chinese).
- Wang J, Zhou Q, Shang J, Liu C, Zhuang T, Ding J, Xian Y, Zhao L, Wang W, Zhou G, Tan C, Huo Z. 2021. UAV- and machine learning-based retrieval of wheat SPAD values at the overwintering stage for variety screening. *Remote Sensing* 13:5166 DOI 10.3390/rs13245166.

- Woebbecke DM, Meyer GE, Von Bargen K, Mortensen DA. 1995. Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE* 38:259–269 DOI 10.13031/2013.27838.
- Wu Q, Zhang Y, Xie M, Zhao Z, Yang L, Liu J, Hou D. 2023. Estimation of Fv/Fm in spring wheat using UAV-based multispectral and RGB imagery with multiple machine learning methods. *Agronomy* 13:1003 DOI 10.3390/agronomy13041003.
- Xu S, Xu X, Zhu Q, Meng Y, Yang G, Feng H, Yang M, Zhu Q, Xue H, Wang B. 2023. Monitoring leaf nitrogen content in rice based on information fusion of multi-sensor imagery from UAV. *Precision Agriculture* 24:2327–2349 DOI 10.1007/s11119-023-10042-8.
- Yan Y, Xu S, Liang Y, Chen H, Liu Y. 2019. Content determination of 10 kinds of chemical components in *Ligusticum chuanxiong* under different storage conditions by HPLC. *China Pharmacy* 30:807–812 (in Chinese).
- Yang C, Zhang L, Zhuang H, Zhang G. 2024. Research on the development status and quality management methods of traditional Chinese medicine decoction pieces. *Journal of North Pharmacy* 21:190–192 (in Chinese) DOI 10.3969/j.issn.1672-8351.2024.05.064.
- Yin L, Peng Y, Liu R, Chen H, Zhao Z, Liu Y. 2012. Study on the correlation between ecological environment factors and the quality of traditional Chinese medicine. *Pharmacy and Clinics of Chinese Materia Medica* 3:9–14 (in Chinese).
- Yoon HI, Lee H, Yang J-S, Choi J-H, Jung D-H, Park YJ, Park J-E, Kim SM, Park SH. 2023. Predicting models for plant metabolites based on PLSR, AdaBoost, XGBoost, and LightGBM algorithms using hyperspectral imaging of *Brassica juncea*. *Agriculture-Basel* 13:1477 DOI 10.3390/agriculture13081477.
- Yu W, Wang M, Zhang D, Yu X, Li M, Liu W. 2021. Multiple indices based quality assessment of *Ligusticum chuanxiong* in fresh processing. *Chinese Traditional Patent Medicine* 43:2109–2114 (in Chinese) DOI 10.3969/j.issn.1001-1528.2021.08.024.
- Zhang M, Shen M, Pu Y, Li H, Zhang B, Zhang Z, Ren X, Zhao J. 2022a. Rapid identification of apple maturity based on multispectral sensor combined with spectral shape features. *Horticulturae* 8:361 DOI 10.3390/horticulturae8050361.
- Zhang Y, Zhao D, Liu H, Huang X, Deng J, Jia R, He X, Tahir MN, Lan Y. 2022b. Research hotspots and frontiers in agricultural multispectral technology: bibliometrics and scientometrics analysis of the Web of Science. *Frontiers in Plant Science* 13:955340 DOI 10.3389/fpls.2022.955340.
- Zhao Y, Fu T, Fan Q. 2008. Growth dynamic analysis of various organs of *Ligusticum chuanxiong* Hort. *Southwest China Journal of Agricultural Sciences* 4:1089–1093 (in Chinese).
- Zhu Y, Liu J, Tao X, Su X, Li W, Zha H, Wu W, Li X. 2023. A three-dimensional conceptual model for estimating the above-ground biomass of winter wheat using digital and multispectral unmanned aerial vehicle images at various growth stages. *Remote Sensing* 15:3332 DOI 10.3390/rs15133332.