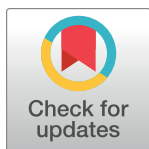


RESEARCH ARTICLE

An integrated UAV growth monitoring model of *Cinnamomum camphora* based on whale optimization algorithmRongxin Gong¹, Haina Zhang^{1*}, Xianghui Lu¹, Haolong Wan¹, Yue Zhang¹, Xin Luo¹, Jie Zhang¹, Rongxiu Xie^{1,2}¹ Jiangxi Provincial Engineering Research Center of Seed-Breeding and Utilization of Camphor Trees, Nanchang Institute of Technology, Nanchang, China, ² Jiangxi Provincial Technology Innovation Center for Ecological Water Engineering in Poyang Lake Basin, Nanchang, China* zhanghaina@nit.edu.cn

OPEN ACCESS

Citation: Gong R, Zhang H, Lu X, Wan H, Zhang Y, Luo X, et al. (2024) An integrated UAV growth monitoring model of *Cinnamomum camphora* based on whale optimization algorithm. PLoS ONE 19(6): e0299362. <https://doi.org/10.1371/journal.pone.0299362>

Editor: Sandeep Samantaray, National Institute of Technology Srinagar, INDIA

Received: September 12, 2023

Accepted: February 9, 2024

Published: June 21, 2024

Copyright: © 2024 Gong et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: National Natural Science Foundation of China (Grant Nos. 52269013 and 32060333) provided experimental sites and financial support. Jiangxi Provincial Natural Science Foundation Project (Grant No. 20232BAB205031) Jiangxi Provincial Key Discipline Academic and Technical Leader Training Program Youth Project (Grant No. 20204BCL23046) provided technical guidance.

Abstract

To explore an effective analysis model and method for estimating *Cinnamomum camphora*'s (*C. camphora*'s) growth using unmanned aerial vehicle (UAV) multispectral technology, we obtained *C. camphora*'s canopy spectral reflectance using a UAV-mounted multispectral camera and simultaneously measured four single-growth indicators: Soil and Plant Analyzer Development (SPAD) value, aboveground biomass (AGB), plant height (PH), and leaf area index (LAI). The coefficient of variation and equal weighting methods were used to construct the comprehensive growth monitoring indicators (CGMI) for *C. camphora*. A multispectral inversion model of integrated *C. camphora* growth was established using the multiple linear regression (MLR), partial least squares (PLS), support vector regression (SVR), random forest (RF), radial basis function neural network (RBFNN), back propagation neural network (BPNN), and whale optimization algorithm (WOA)-optimized BPNN models. The optimal model was selected based on the coefficient of determination (R^2), normalized root mean square error (NRMSE) and mean absolute percentage error (MAPE). Our findings indicate that apparent differences in the accuracy with different model, and the WOA-BPNN model is the best model to invert the growth potential of *C. camphora*, the R^2 of the model test set was 0.9020, the RMSE was 0.0722, and the MAPE was 7%. The R^2 of the WOA-BPNN model improved by 18%, the NRMSE decreased by 33%, and the MAPE decreased by 9% compared with the BPNN model. This study provides technical support for the modern field management of *C. camphora* essential oil and other dwarf forestry industries.

1. Introduction

Crop growth refers to the state and trend of crops at different growth stages. Common crop growth parameters include PH leaf area index LAI, chlorophyll content, water, and AGB [1, 2]. Crop growth is directly related to crop yield and quality, and it must be observed and

Jiangxi Provincial Forestry Bureau Camphor Tree Research Project (Approval No. 202007-01-04)
Jiangxi Provincial Science and Technology Department Major Science and Technology Project of China, (20203ABC28W016-01-04) provided equipment support.

Competing interests: The authors have declared that no competing interests exist.

assessed to effectively support agricultural production [3, 4]. However, the traditional crop growth monitoring methods are based on visual observation or destructive sampling, which is labor-, time-, and resource-consuming and particularly inefficient when monitoring large crop area [5].

Remote sensing technology provides information rapidly and non-destructively on a large scale and has broad applications for crop growth prediction [6, 7]. Liu et al. used the DJI Phantom 4 multispectral UAV to obtain the spectral reflectance of a rice canopy and established regression models of three growth indicators and vegetation indices, including rice SPAD, AGB, and leaf area, discovering that the accuracy of the multi-vegetation index regression model was superior to that of the one-dimensional regression model and that the multi-vegetation index regression model and rice SPAD established using support vector machine had the highest accuracy [6]. However, a single-growth index characterizes the growth condition of a specific crop feature and fails to represent the crop's comprehensive growth to some extent; therefore, several researchers have studied integrating the single-growth index. For example, Y. Xu et al. obtained AGB, PH, SPAD, and plant water content data of winter wheat and constructed the CGMI based on the coefficient of variation (CV) method; they discovered that the inversion model of CGMI and vegetation index of wheat constructed by genetic algorithm optimized neural network have the highest accuracy, with the coefficient of determination R^2 reaching 0.8 [8]. X. Xu et al. used an improved fuzzy comprehensive evaluation (FCE) method to establish a comprehensive yield evaluation index (CYEI) to obtain information, such as wheat LAI, biomass, water content, and nitrogen content, to monitor wheat growth and estimate yield [9].

The above studies have shown that machine learning (ML) effectively builds single or comprehensive crop growth prediction models using UAV multispectral techniques. The ML models' accuracy and efficiency depend heavily on their internal parameters, which can be improved using optimization algorithms to determine model parameters [10]. For example, Han et al. used the XGBoost model optimized by the grasshopper optimization algorithm (GOA) to establish the inversion model of multispectral vegetation index and wheat raw AGB, discovering that the GOA-XGB model (having 37 input factors) had the highest accuracy, with $R^2 = 0.847$ [11]. Jianqiang et al used UAV multispectral remote sensing to establish a model for monitoring SPAD values of jujube leaves found that the PSO-ELM model using the particle swarm optimization (PSO) algorithm was more accurate than the ELM model alone, with a 7% improvement in R^2 and a 52% decrease in RMSE [12]. Mirjalili et al. proposed the novel whale optimization algorithm (WOA), which uses random or optimal search agents to simulate hunting behavior and spirals to simulate the bubble-net attacking mechanism of humpback whales [13]. This algorithm is widely used in economic scheduling, optimal control, photovoltaic systems, image segmentation, and other fields owing to its simple mechanism, fewer parameters, high optimization capacity, and ability to escape local optima [14]. For example, the WOA algorithm can solve deep neural network architecture training and hyper-parameter optimization problems [15]. In the field of remote sensing for agriculture and forestry, Feng et al. used unmanned aerial spectral remote sensing to establish a model for monitoring nitrogen content in rice, and found that the rice nitrogen content inversion model constructed by Whale-optimized Extreme Learning Machine (WOA-ELM) was superior to the rice nitrogen content inversion model constructed by Whale-optimized Extreme Learning Machine (ELM) [16]. This suggests that the WOA algorithm has great potential in the optimization of agroforestry remote sensing models. *Cinnamomum camphora* (Linn.) Presl is an ornamental tree and essential woody oil plant in China. Because of the oil's anti-inflammatory, anti-bacterial, and pain-relieving properties, it is also exported to other countries as a spice and holds significance in the medical and spice industry owing to its high economic value [17–

[19]. *C. camphora* essential oil industry is developing rapidly and has emerged as a leading forestry industry in southern China [20]. *C. camphora* growth directly affects its oil yield and essential oil quality. Therefore, *C. camphora* growth must be rapidly and accurately monitored, and the development of field management systems for forestry production must be timely directed to develop the high-quality *C. camphora* essential oil industry. However, the research on *C. camphora* growth monitoring only stays on the inversion of a single growth index, and some researchers have established the inversion model of spectral index on camphor SPAD and oil yield using UAV multispectral remote sensing images, and found that radial basis function neural network and random forest are the best models for inversion of spectral index on camphor SPAD and oil yield, with the coefficients of decision of 0.79 [21] and 0.87 [22], respectively, and there are few studies on the use of multiple growth indicators to realize comprehensive monitoring of *C. camphora* growth.

In this study, we obtained *C. camphora*'s canopy spectral reflectance using a UVA-mounted multispectral camera and simultaneously measured four single-growth indicators: SPAD, AGB, PH, and LAI. The CV and equal weighting methods were used to construct *C. camphora* comprehensive growth monitoring indicators. The multiple linear regression (MLR), partial least squares regression (PLS), support vector regression (SVR), random forest (RF), radial basis function neural network (RBFNN), back propagation neural networks (BPNN), and WOA-BPNN models were used to establish an efficient multispectral inversion model for the comprehensive growth of *C. camphora* and an efficient method for estimating *C. camphora* growth, providing technical support for modern field management of *C. camphora* essential oil and other dwarf forestry industries.

2. Materials and methods

2.1 Overview of the study area

The study site was located in the *C. camphora* Germplasm Resource Nursery, Nanchang Institute of Technology, Nanchang, Jiangxi Province, (28°41'40.85"N, 116°1'41.18"E). The study area had a subtropical humid seasonal-wind climate with sufficient light and an average annual rainfall of 1600 mm, with precipitation of 147–157 d and an average multi-year temperature of approximately 17°C. The average annual sunshine hours were approximately 1772–1845 h, with a sunshine rate of 40%. Some of the soils in the area are dense and have a reddish soil texture that is weakly acidic. The soil included: organic matter content, 6.39 g·kg⁻¹; total nitrogen content, 0.62 g·kg⁻¹; total phosphorus content 0.30 g·kg⁻¹; total potassium content, 13.00 g·kg⁻¹; alkaline hydrolysis nitrogen content, 47.74 mg·kg⁻¹; fast-acting phosphorus content, 1.49 mg·kg⁻¹; fast-acting potassium content, 61.10 mg·kg⁻¹.

The test *C. camphora* species was "Ganfeng No. 1". The study area was divided into 66 plots (Fig 1), each planted with nine *C. camphora* seedlings in three rows and three columns with a plant spacing of 1 m. Each sample plot was 3 m × 3 m, with 594 one-year-old cutting seedlings planted in April 2021.

2.2 Data acquisition

2.2.1 Multispectral data acquisition. An MS600pro multispectral camera platform (Yusense, Inc., Qingdao, China), manufactured by Changguang Yuchen Information Technology and Equipment (Qingdao) Co., Ltd., mounted on DJI Matrice M300RTK Quadrotor UAV was used to acquire remote sensing data. The multispectral camera had six spectral channels. The multispectral image was taken on the morning of September 25, 2022, in clear and cloudless weather. The flight route was planned based on the study area, and whiteboard correction was performed. The following conditions were applied: flight altitude, 30 m; setting speed,

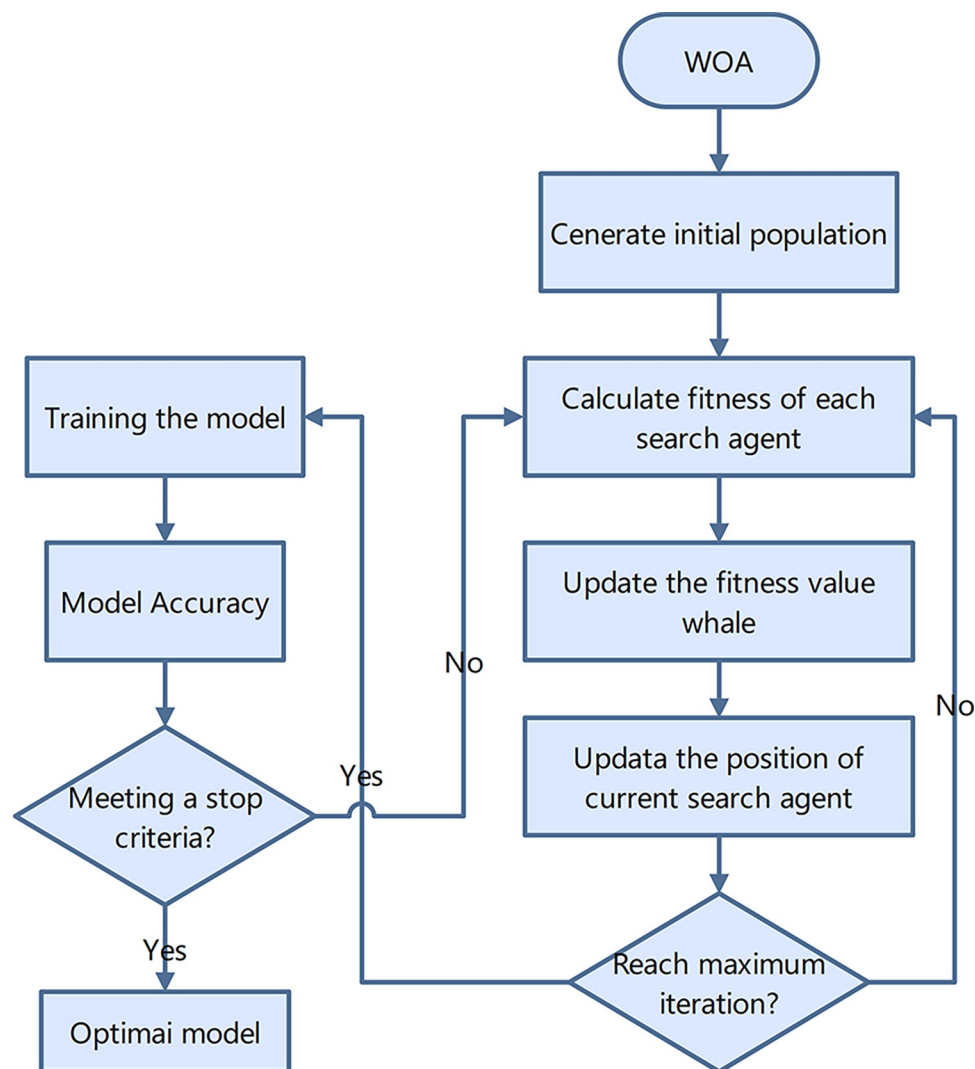


Fig 1. Flowchart of whale flock optimization algorithm.

<https://doi.org/10.1371/journal.pone.0299362.g001>

2.5 m·s⁻¹; resolution of the image element, 4.09 cm; automatic capture mode, with a 75% heading overlap and 65% side overlap.

The obtained UAV images were spliced and preprocessed with geometric and radiometric corrections using Ysense Map aerial remote sensing preprocessing software. Subsequently, the preprocessed UAV multispectral image information was imported into ENVI software, each plot in the study area was considered a region of interest (ROI), and *C. camphora* leaf images were intercepted to remove the impact bands of soil and shadows (Fig 1). The average reflectance spectrum of *C. camphora* leaf samples in each ROI was used as the plot spectral reflectance.

2.2.2 *C. camphora* growth index collection. The *C. camphora* growth index was simultaneously obtained with the UAV flight. Three good-growing *C. camphora* trees were selected in each plot, and five healthy mature leaves from the upper and middle parts of each tree were randomly selected to measure their SPAD using a SPAD-502 handheld chlorophyll meter. Each leaf part was measured thrice to obtain an average value, and the average value of five leaf measurements was calculated as the tree SPAD. A tape was used to measure the height of nine

C. camphora plants in each plot, which was the distance of the plant from the ground to the highest point of the branches, and the average was calculated as the PH of the plot. The LAI data were measured using the LAI-2200C plant canopy analyzer (USA). Three representative *C. camphora* trees were selected for each sample plot, and LAI values were measured and recorded at four selected sites for each tree; the average of the four measurements was calculated as the LAI of the tree, and the average of the three trees was calculated as the LAI of the plot. After bringing three representative *C. camphora* trees from each sample plot to the laboratory, they were first heated at 105°C for 30 min, then baked at 80°C until constant mass, and weighed using an electronic balance. The mean value was calculated for each plot to obtain the AGB.

2.3 Vegetation index construction

Vegetation indices are quantitative values derived from multispectral remote sensing data about the Earth's vegetation cover status, typically obtained through mathematical operations using the red and near-infrared bands. Table 1 shows the calculation formulae for various common vegetation indices, where B, G, R, RE, and NIR are the spectral reflectance of 450, 550, 660, 750, and 840 nm bands, respectively.

2.4 Comprehensive growth index construction

To better understand the growth status of *C. camphora* in the study area, four vegetation growth indicators: SPAD, PH, LAI, and AGB, were selected to construct a comprehensive growth monitoring index (CGMI). The four single-growth indicators were normalized to eliminate the dimension and were used to construct CGMI₁ with equal weights using the following formula [34]:

$$U_i = \frac{x_i - \min x_i}{\max x_i - \min x_i} \quad (1)$$

$$CGMI_1 = \frac{1}{4} \sum_{i=1}^4 U_i \quad (2)$$

Table 1. Vegetation index calculation formula.

Vegetation Index	Formula	Literature
Normalized Differential Vegetation Index (NDVI)	$(NIR - R)/(NIR + R)$	[23]
Greenness Normalized Differential Vegetation Index (GNDVI)	$(NIR - G)/(NIR + G)$	[23]
Modified Simple Ratio (MSR)	$((NIR/R) - 1)/\sqrt{(NIR/R) + 1}$	[24]
Transformed Chlorophyll Absorption Reflectance Index (TCARI)	$3 \times (RE - R) - 0.2 \times (RE - G) \times (RE/R)$	[25]
Red-edge Chlorophyll Index (CIrededge)	$NIR/(RE - 1)$	[26]
Difference Vegetation Index (DVI)	$NIR - R$	[27]
Optimized Soil Adjusted Vegetation Index (OSVAVI)	$(NIR - R)/(NIR + R + 0.16)$	[28]
Visible Difference Vegetation Index (VDVI)	$(2G - R - B)/(2G + R + B)$	[29]
Structural Insensitivity Index (SIPI)	$(NIR - B)/(NIR + B)$	[30]
Modified Triangular Vegetation Index (MTVI)	$\frac{1.5(1.2(NIR - G) - 2.5(R - G))}{\sqrt{(2NIR + 1)^2 - 0.5 - (NIR - 5\sqrt{R})}}$	[31]
Modified Nonlinear Vegetation Index (MNVI)	$1.5(NIR^2 - 2)/(NIR^2 + R + 0.5)$	[32]
Atmospheric Resistance Vegetation Index (ARVI)	$(NIR - (2R - B))/(NIR + (2R - B))$	[33]

<https://doi.org/10.1371/journal.pone.0299362.t001>

where i is the index category, U_i is the normalized index of i , and $maxx_i$ and $minx_i$ are the maximum and minimum values of the index of i category, respectively.

The CV method is a common weight determination method that uses the degree of variation of each evaluation index to determine its weight; the given index weight increases or decreases with increasing or decreasing CV, respectively [35]. Because the CV reflects the relative fluctuation of indicators, indicators with large coefficients of variation reflect more important information and have higher differentiation and stability, which are given larger weights; conversely, the evaluation indicators are given smaller weights. The CV, i.e., the standard deviation of the indicator divided by the mean, is calculated as [36]:

$$S_i = \sqrt{\frac{1}{4} \sum_{i=1}^4 (U_i - \bar{x}_i)^2} \quad (3)$$

$$V_i = \frac{S_i}{\bar{x}_i} \quad (4)$$

where $\bar{x}_i$ is the average of index i , S_i is the standard deviation of index i , and V_i is the CV of index i .

The CV method was used to determine the weights of the normalized SPAD, PH, LAI, and AGB vegetation growth indicators and construct CGMI₂ using the formula [34]:

$$W_i = \frac{V_i}{\sum_{i=1}^4 V_i} \quad (5)$$

$$CGMI_2 = \sum_{i=1}^4 U_i W_i \quad (6)$$

where W_i is the index weight of i .

2.5 Model construction

2.5.1 Multiple linear regression (MLR). MLR is a widely used statistical model for predicting a continuous dependent variable from multiple independent variables. Unlike simple linear regression, which utilizes only one single independent variable, MLR can accommodate multiple independent variables, increasing the complexity and accuracy of the prediction model. It is important to note that the MLR model performs better for cases where there is no high degree of covariance (i.e., there is no linear correlation between the independent variables) between the independent variables. When there is a high degree of covariance between the independent variables, the model results may be misleading. In practice, regression diagnostics and model validation are usually required to assess the applicability and accuracy of MLR models. [37]. The MLR model for this study was built using Matlab software, setting Modelspec as linear.

2.5.2 Support vector regression (SVR). Support Vector Machines (SVMs) were originally proposed for binary classification problems. It is a powerful supervised learning algorithm designed to find an optimal hyperplane for classification in a high-dimensional feature space. SVR uses the same principles as SVM, but focuses on modeling and predicting continuous variables rather than category labels. It is particularly well suited for solving nonlinear regression problems. In SVR, a kernel function (e.g. radial basis function RBF or polynomial kernel function) is used to transform the feature space into a high-dimensional space. By projecting the data into this high-dimensional space, SVR can capture more complex relationships between features and target variables. The choice of kernel function plays a key role in determining the

mapping from the original feature space to the high-dimensional space. Overall, SVR is an important extension of Support Vector Machines that applies a kernel function to transform the feature space into a nonlinear regression model.[38, 39]. In this study, a Gaussian RBF was used as the kernel function, and a network search method was used to select optimal parameters, with an RBF parameter of 30 and a penalty factor of 20.

2.5.3 Random Forest (RF). Random Forest (RF) is an integrated learning method that consists of multiple decision trees, and its basic principle is to improve prediction accuracy and robustness through the integration of multiple decision trees. Each decision tree uses randomly selected features and samples during training, so that each decision tree learns different features and laws from different subsets of data. When a prediction is needed, each decision tree makes an independent prediction and finally the final prediction is determined by voting. Overall, Random Forest is a powerful integrated learning method that can effectively deal with classification and regression problems. It improves the prediction performance and generalization ability through the mechanism of decision tree integration and random attribute selection, with good robustness and interpretability.[40]. A 10-fold cross-validated grid search method was used to adjust RF parameters, setting the number of leaf nodes to 4 and the number of decision trees to 100.

2.5.4 Partial least squares (PLS). Partial Least Squares Regression (PLS) combines the advantages of typical correlation analysis, principal component analysis, and multiple linear regression to deal with high-dimensional data and multiple covariates, and to solve multivariate problems. The core idea of PLS is to find the maximum covariance between the independent and dependent variables by projecting them in the Direction. [41]. The PLS model for this study was built using Matlab software, setting NCOMP as 2.2.5.5 Radial basis function neural network (RBFNN)

Radial Basis Function Neural Network (RBFNN) is a nonparametric regression method that uses radial basis functions as basis functions to fit continuous nonlinear functions. The core idea of RBFNN is to use radial basis functions as activation functions in the hidden layer of the network. These radial basis functions are basis functions parameterized by the distance between the input variables and their respective centers. The neurons in the hidden layer are activated using these radial basis functions and their outputs are linearly combined with the input variables as weights. Finally, the output layer is linearly combined to obtain the final prediction.[42, 43]. RBFNN has only one linear layer in its hidden layer, which reduces the number of iterations and computation time and improves the training speed compared with traditional feedforward neural networks. The expansion speed of RBF was set to 100.

2.5.5 Back propagation neural network (BPNN). BPNN is an artificial neural network based on an error backpropagation algorithm. Its basic structure comprises an input layer, a hidden layer, and an output layer. BPNN can build a nonlinear mapping relationship by combining multiple layers of neurons. By changing the weights and bias of the hidden layer, BPNN can approximate the nonlinear relationship between the input and output [44]. Its most prominent feature lies in the signal's forward propagation and the error's backward propagation, and this propagation mechanism shifts the output data closer to the expected value. BPNN can be improved by increasing intermediate layers and the number of neurons to fit different application areas and task types [45, 46]. The BPNN estimation model was constructed using MATLAB's Neural Network toolbox. Based on the numerical optimization theory, the transfer function of the implicit layer was set to the Tansig function, and the Levenberg–Marquardt (L–M) algorithm was used for network training. After several training sessions, the number of intermediate layer neurons was set to 12, the maximum number of iterations to 1000, the training target to $e \times 10^{-4}$, and the learning rate to 0.1. The size of the batch data used for

Table 2. Model parameters.

Models	Tuning parameter models
MLR	MODELSPEC = 'linear'
PLS	NCOMP = 2
SVR	penalty factor: 20 RBF parameter: 30
RF	leaf nodes: 4 decision trees: 100
RBFNN	expansion speed: 100
BPNN	intermediate layer neurons: 12 maximum number of iterations: 1000 training target: $\times 10^{-4}$

<https://doi.org/10.1371/journal.pone.0299362.t002>

gradient descent each time was 9, and the simulated values were obtained after the neural network training.

For a better presentation of the model, the model parameters are shown in the following Table 2.

2.6 Whale optimization algorithm (WOA)

Humpback whales have a special hunting method called bubble-net attacking method. The WOA simulates the humpback whale's unique hunting method and seizure mechanism, which is mainly divided into searching and encircling the prey, bubble-net foraging, and obtaining the global optimal solution. First, the searching and encircling prey are simulated as follows [13]:

$$B = |2rx_{best}^t - x_i^t| \quad (7)$$

$$x_i^{t+1} = x_{best}^t - AB \quad (8)$$

$$A = 2ar - a \quad (9)$$

$$a = 2 - \frac{2t}{T} \quad (10)$$

where A is the coefficient vector, B is the encircling step, x_{best}^t is the position of the individual whale with the global optimal solution within the population when iterated to generation t, x_i^t is the position of the individual when iterated to generation t, r is a random number within [0,1], a is a linear decreasing coefficient from 2 to 0, and T is the maximum number of iterations.

The bubble-net attacking is simulated as follows:

$$C = |x_{best}^t - x_i^t| \quad (11)$$

$$x_i^{t+1} = x_{best}^t + Ce^{bl} \cos(2\pi l) \quad (12)$$

where C denotes the distance between the whale and the prey, b is a constant controlling the logarithmic spiral shape, and l is a random number within [-1,1]. The flowchart of the whale swarm optimization algorithm is shown in Fig 1.

2.7 Model accuracy evaluation

To determine the most ideal model and test its reliability and accuracy of prediction results, R^2 , normalized root mean square error (NRMSE) and Mean Absolute Percentage Error (MAPE) were calculated using the following formulae [47]:

$$R^2 = \frac{\sum_{i=1}^n (\hat{a}_i - \bar{b})^2}{\sum_{i=1}^n (a_i - \bar{b})^2} \quad (13)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{a}_i - a_i)^2}{n}} \quad (14)$$

$$NRMSE = \frac{RMSE}{\hat{a}_{imax} - \hat{a}_{imin}} \quad (15)$$

$$MAPE = \frac{100\%}{n} \sum_i^n \left| \frac{\hat{a}_i - a_i}{a_i} \right| \quad (16)$$

where n is the number of samples, $\hat{a}_i$ is the estimated value of *C. camphora* growth, $\hat{a}_{imax}$ and $\hat{a}_{imin}$ are the maximum and minimum values of the estimated *C. camphora* growth, respectively, a_i is the measured value of the photosynthetic parameter of *C. camphora*, and $\bar{b}$ is the mean value of actual measurements of photosynthetic parameters. The model is considered accurate when R^2 is close to 1, NRMSE and MAPE are close to 0, and prediction results are close to the measured results. In this study, the accuracy parameters of the models were statistically compared to identify the optimal model more intuitively.

The research framework is shown in the Fig 2 below.

3. Results and analysis

3.1 Data analysis

According to the CV method, the four single index weights of SPAD, PH, LAI, and AGB were 0.130514, 0.236403, 0.302251, and 0.330832, respectively; therefore, based on Eq (6), the CGMI was:

$$CGMI_1 = 0.25 (U_1 + U_2 + U_3 + U_4) \quad (17)$$

$$CGMI_2 = 0.130514U_1 + 0.236403U_2 + 0.302251U_3 + 0.330832U_4 \quad (18)$$

where U_1 is the normalized SPAD value, U_2 is the normalized PH, U_3 is the normalized biomass, and U_4 is the normalized plant water content. The statistical results of *C. camphora* growth data are shown in Tables 2 and 3. The sample size was 66, and the coefficients of variation of SPAD, PH, LAI, and AGB were 0.09, 0.17, 0.22, and 0.23, respectively. Except for SPAD, the coefficients of variation of single-growth indices were greater than 0.15, indicating that the growth data were significantly affected by the treatments. The coefficients of variation of comprehensive growth indices were over 0.30, at least 0.10 higher than that of single-growth indices, with greater dispersion, which had less impact on the modeling analysis for the algorithm with large noise tolerance.

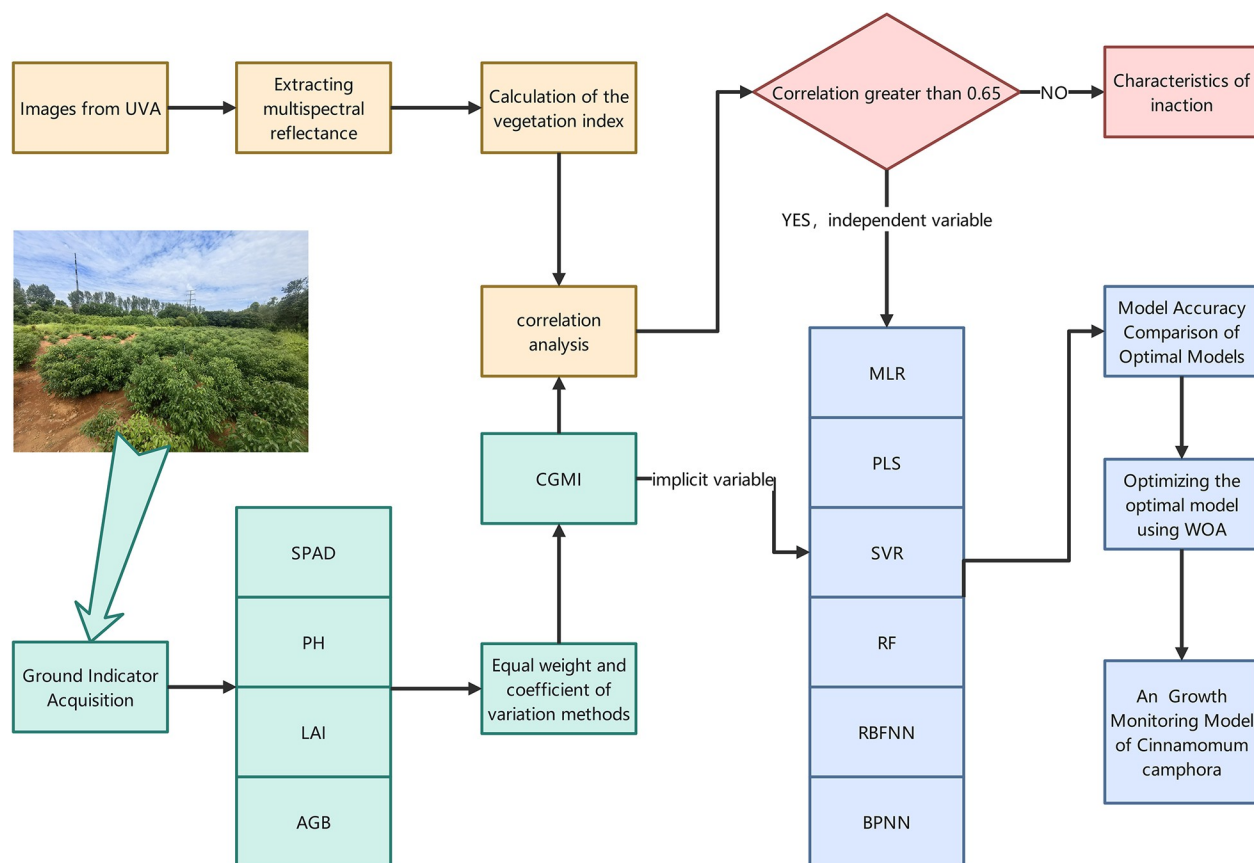


Fig 2. Research framework diagram.

<https://doi.org/10.1371/journal.pone.0299362.g002>

3.2 Correlation analysis of vegetation index and growth index

The Pearson correlation coefficients of vegetation index and *C. camphora* growth index calculated based on UAV multispectral reflectance are shown in Fig 3, and the correlation coefficients take values in the range of $[-1,1]$. The correlation coefficients were positively correlated when the two variables increase simultaneously.

It can be seen that some vegetation indices are highly correlated with each other, so in order to reduce the information sink and covariance, only five vegetation indices with high correlation with *C. camphora* and significant at 0.01 level, namely, TCART, $Cl_{rededge}$, OSAVI, DVI and MNVI, are selected as model inputs in this study. Among them, TCART is the vegetation index with the highest correlation with *C. camphora*'s growth index, and its correlation

Table 3. Sample statistics.

Name	Minimum value	Maximum value	Average value	Standard deviation	Coefficient of variation
SPAD	33.50	54.78	40.18	3.75	0.09
PH	0.58	1.57	1.18	0.20	0.17
LAI	1.25	4.28	2.90	0.63	0.22
AGB	8720.09	2095.28	6347.70	1515.69	0.23
CGMI ₁	0.06	0.90	0.52	0.16	0.32
CGMI ₂	0.06	0.93	0.55	0.18	0.33

<https://doi.org/10.1371/journal.pone.0299362.t003>

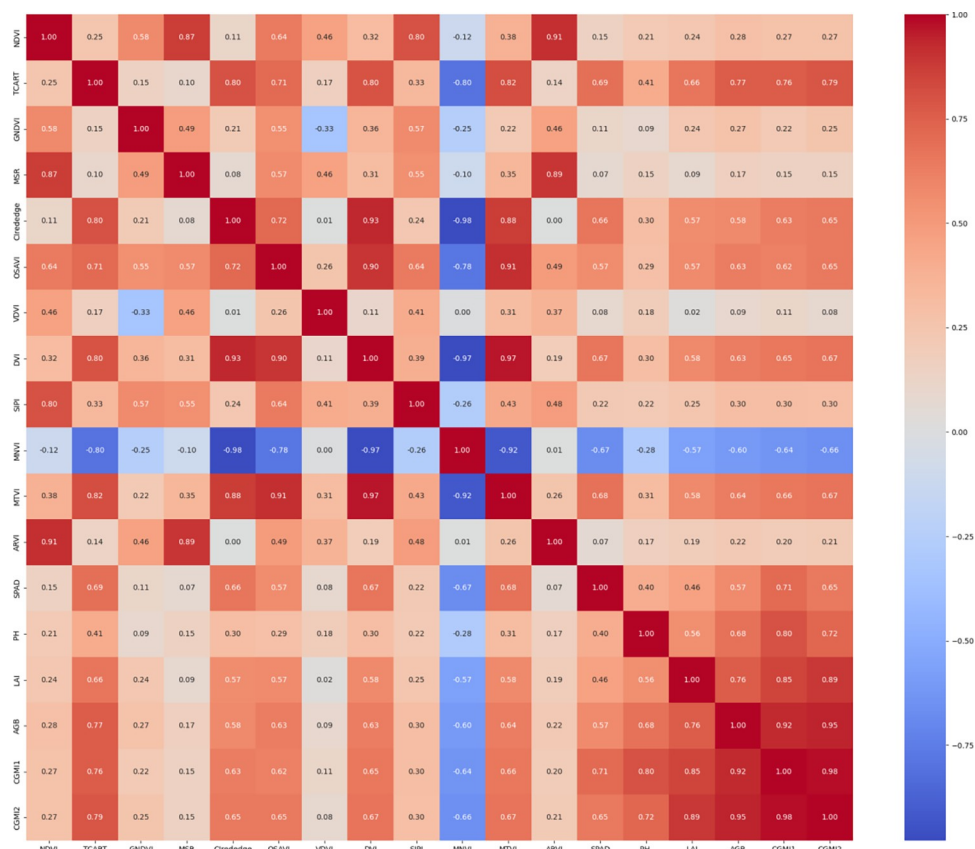


Fig 3. Correlation analysis.

<https://doi.org/10.1371/journal.pone.0299362.g003>

with four single growth indexes of SPAD, PH, LAI and AGB, as well as two composite growth indexes of CGMI₁ and CGMI₂ are 0.69, 0.41, 0.66, 0.77, 0.76, and 0.79, respectively. compared with the single growth indexes, the correlation of CGMI with the five vegetation indexes which have passed the significance Compared with the single growth index, the correlation between CGMI and the five vegetation indices that passed the significance test was significantly higher, and the improvement of CGMI₂ was more significant than that of CGMI₁. The correlation between CGMI₂ and DVI was 123.33% higher than that between PH and DVI, and the correlation between CGMI₂ and TCART was 92.45% higher than that between PH and TCART, indicating that the selected vegetation indices could better respond to CGMI₂.

3.3 Inversion model construction

Fifty-two samples were selected as the training set, and 14 as the validation set to estimate CGMI₁ and CGMI₂ using six models: MLR, PLS, SVR, RF, RBFNN, and BPNN, and The model accuracy is shown below (Table 4 and Fig 4) and the model fitting results are shown below (Figs 5 and 6):

In the monitoring model of comprehensive growth index of *C. camphora* constructed by equal weight method, only the R^2 of training set and validation set of BPNN model were greater than 0.6, which were 0.7259 and 0.6332, respectively, NRMSE 0.1513 and 0.2371, and MAPE 15% and 17%, respectively. The second is RBFNN, RF and SVR models, R^2 of training set and validation set is both greater than 0.5. And the RBFNN model has low NRMSE and

Table 4. Model accuracy evaluation.

	Model Type	R^2		NRMSE		MAPE	
		Training set	Test set	Training set	Test set	Training set	Test set
CGMI ₁	MLR	0.6291	0.3408	0.2172	0.2714	22%	18% [^]
	PLS	0.5398	0.4632	0.2196	0.2086	21%	22%
	SVR	0.6671	0.5416	0.1510	0.1646	15%	26%
	RF	0.6610	0.5742	0.2275	0.4483	22%	44%
	RBFNN	0.6977	0.5649	0.1498	0.1779	14%	17%
	BPNN	0.7259	0.6332	0.1513	0.2371	15%	17%
CGMI ₂	MLR	0.6465	0.6403	0.2149	0.2265	21%	17%
	PLS	0.6264	0.5102	0.1931	0.1952	19%	19%
	SVR	0.7384	0.6000	0.1505	0.2418	12%	24%
	RF	0.7098	0.6842	0.2084	0.4381	20%	43%
	RBFNN	0.7550	0.6646	0.1292	0.1661	12%	16%
	BPNN	0.8803	0.7761	0.0878	0.1642	9%	16%

<https://doi.org/10.1371/journal.pone.0299362.t004>

MAPE, with NRMSE of 0.1498 and 0.1779 for the training and validation sets, and MAPE of 14% and 17%, respectively. However, the R^2 of MLR and PLS model validation set was lower than 0.5, and the NRMSE was higher, which could not effectively invert the comprehensive growth of *C. camphora* dwarf forest.

Among the comprehensive growth index monitoring models constructed using the CV method, the BPNN model exhibited the best accuracy, with the highest training and validation set R^2 of 0.8803 and 0.7761, respectively, NRMSE and MAPE are the lowest, with NRMSE of 0.0878 and 0.1642 for the training and validation sets, respectively, and MAPE of 9%h and 16%. respectively. The RF and RBFNN models displayed the next-best accuracy, with training set R^2 of 0.7098 and 0.7550, respectively, and validation set R^2 of 0.6842 and 0.6646 respectively. The NRMSE and MAPE of the validation sets of the RBFNN model were lower (0.1661 and 16%, respectively), whereas those of the RF model were higher (0.4381 and 0.43%,

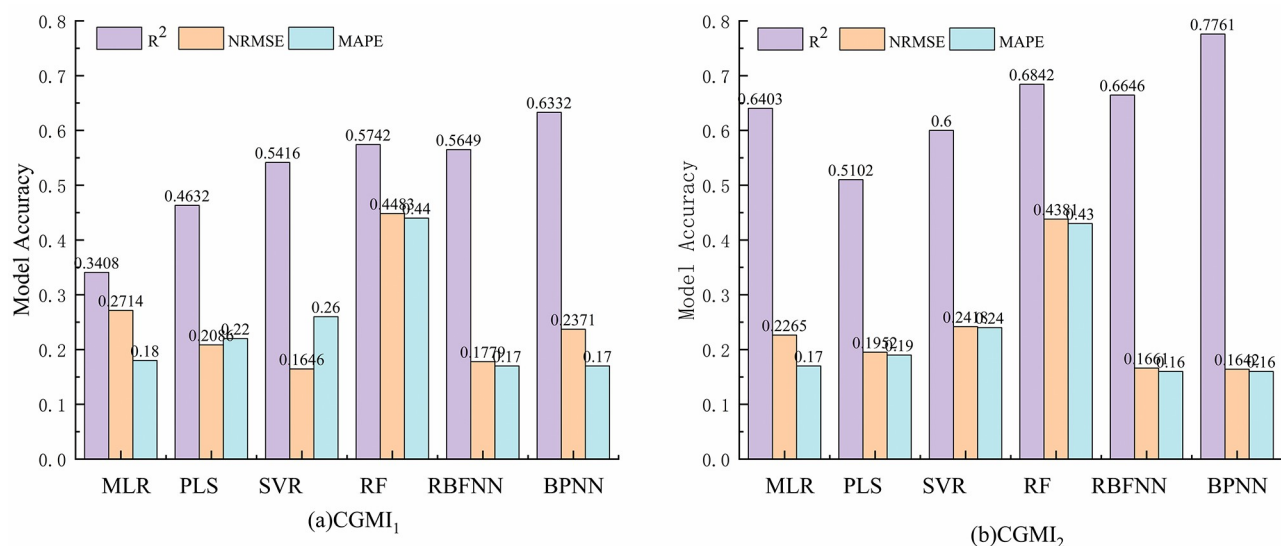


Fig 4. Model evaluation indicators. (a) CGMI₁ model accuracy; (b) CGMI₂ model accuracy.

<https://doi.org/10.1371/journal.pone.0299362.g004>

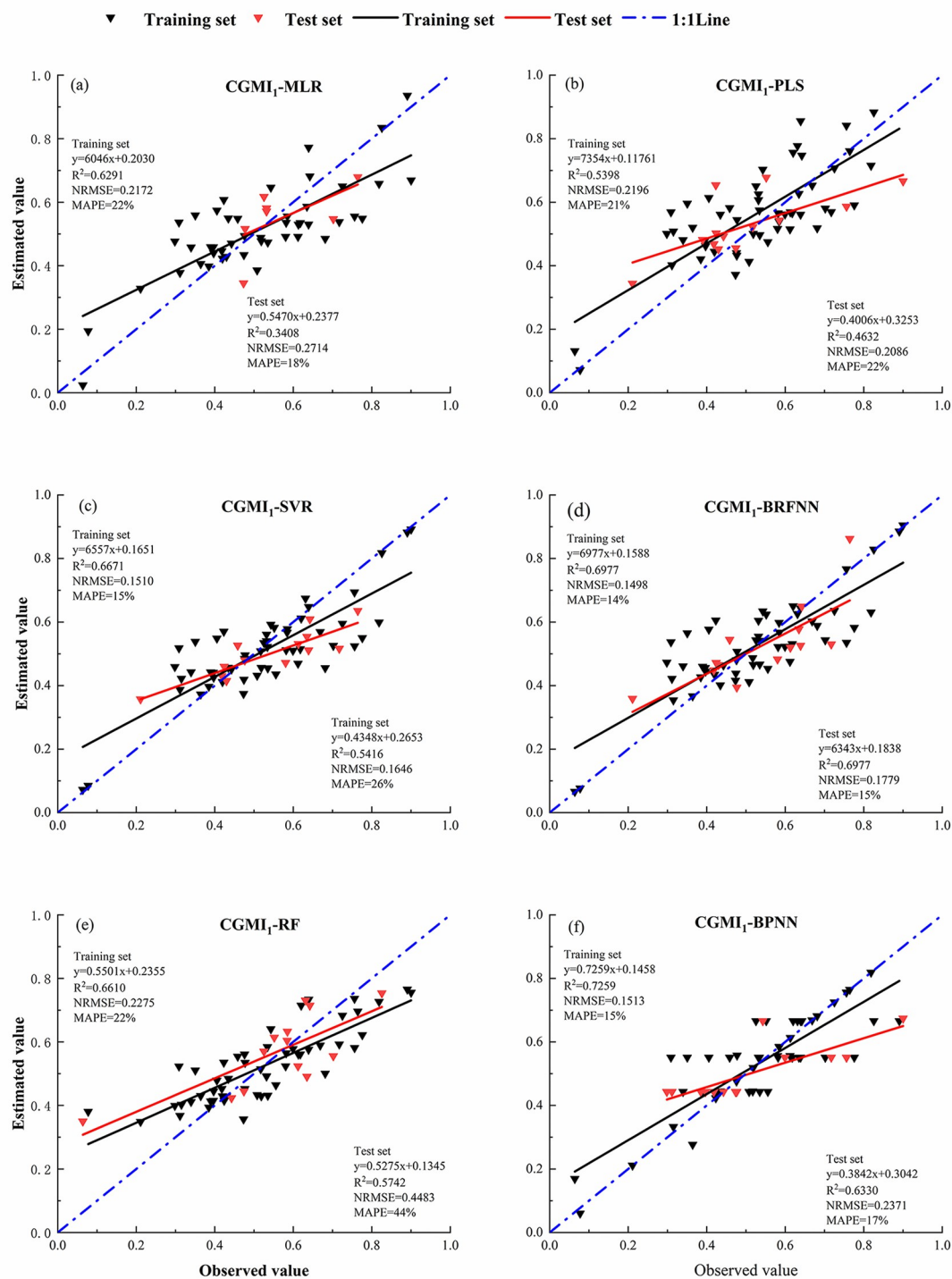


Fig 5. Inversion CGMI₁ model fitting results. (a) CGMI₁ is inverted based on MLR model; (b) CGMI₁ is inverted based on PLS model; (c) CGMI₁ is inverted based on SVR model; (d) CGMI₁ is inverted based on BRFNN model; (e) CGMI₁ is inverted based on RF model; (f) CGMI₁ is inverted based on BPNN model.

<https://doi.org/10.1371/journal.pone.0299362.g005>

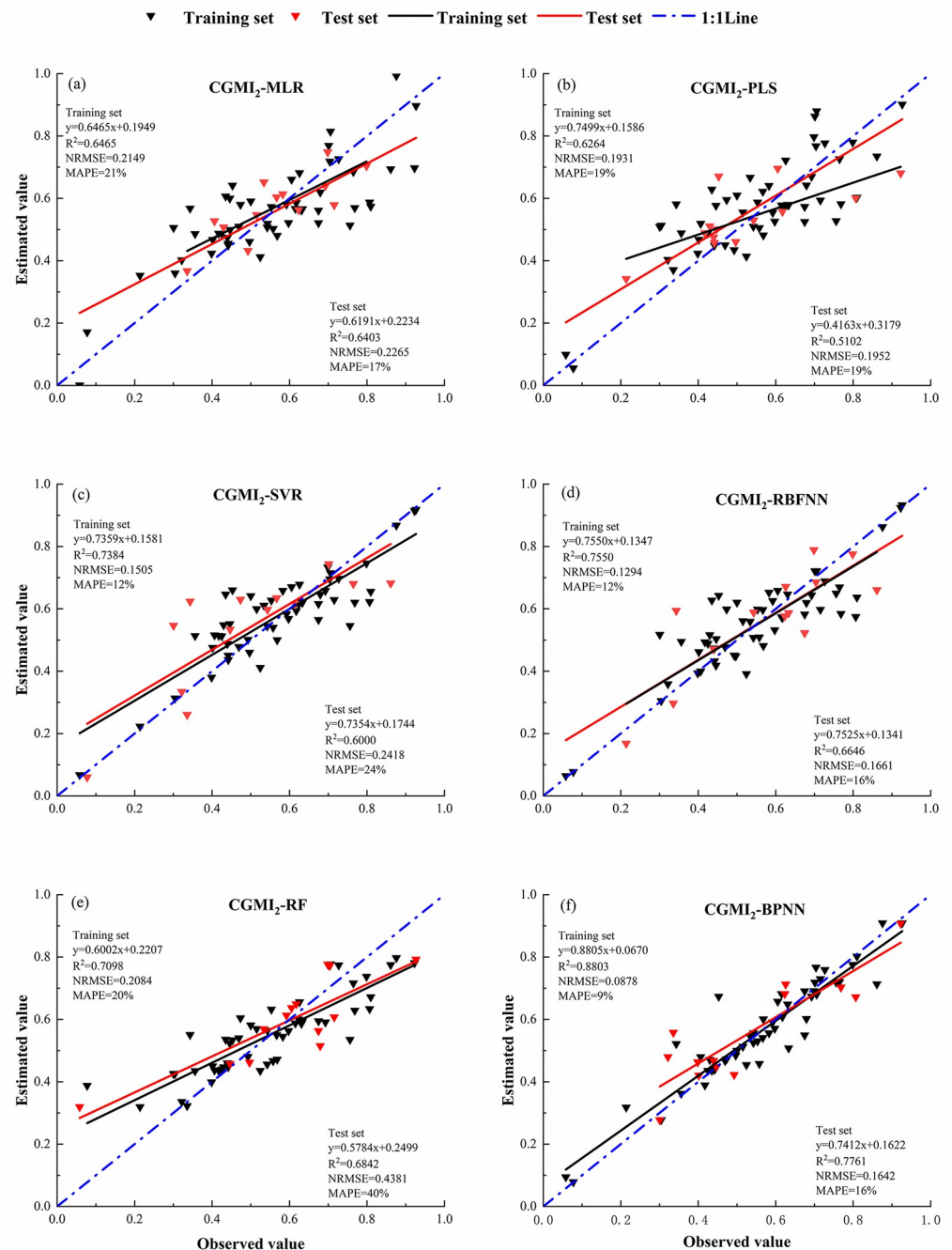


Fig 6. Inversion CGMI₂ model fitting results. (a) CGMI₂ is inverted based on MLR model; (b) CGMI₂ is inverted based on PLS model; (c) CGMI₂ is inverted based on SVR model; (d) CGMI₂ is inverted based on RF model; (e) CGMI₂ is inverted based on BRFNN model; (f) CGMI₂ is inverted based on BPNN model.

<https://doi.org/10.1371/journal.pone.0299362.g006>

respectively). The accuracy of the MLR and SVR models were similar, with both the training and validation set R^2 being >0.6 . Only the PLS model validation set R^2 was <0.6 , with its training and validation set R^2 being 0.6264 and 0.5102, respectively, NRMSE being 0.1931 and 0.1952, respectively, and MAPE of 19% and 19%, respectively.

The comprehensive growth index monitoring model's accuracy was significantly higher when constructed using the CV method than the equal weighting method, with the MLR model's accuracy showing the most significant improvement. The validation set R^2 and NRMSE of the MLR model constructed using CGMI₂ improved by 87% and decreased by 16%, respectively, compared with that constructed using CGMI₁. Following that, compared with the BPNN model constructed using CGMI₁, that constructed using CGMI₂ increased both the training and validation set R^2 by 21%, decreased NRMSE by 41% and 12%, respectively, and MAPE was reduced by 6% and 1%, respectively. The PLS model exhibited the least change; compared with the PLS model constructed using CGMI₁, the PLS model constructed using CGMI₂ increased the training and validation set R^2 by 16% and 10%, respectively, decreased NRMSE by 12% and 6%, respectively, and MAPE was reduced by 2% and 3%.

3.4 Optimizing BPNN

The inversion model of *C. camphora* growth constructed by combining BPNN with the CV method had high prediction accuracy and satisfactory modeling effect. However, because the training process in deploying the BPNN architecture is challenging owing to the long training time required and low efficiency, this study used the WOA algorithm to optimize the weights and thresholds of the BPNN model while maintaining the remaining BPNN model's parameter settings intact. The optimization parameters included population size, number of evolutions, number of neurons in the input layer, number of neurons in the output layer, weights from the input layer to the hidden layer, and weights from the hidden layer to the output layer. The initial population size of the WOA-optimized BPNN model was 10, the number of iterations was 60, and the optimized model fitting results are shown in Fig 7.

In the residual plot, it can be observed that the relationship between the residuals of the WOA-BPNN model and the sample data shows a more uniform distribution than the BPNN model. Therefore, according to the performance of the residual plot, it can be concluded that the WOA-BPNN model has better performance than the BPNN model.

The accuracy of WOA-BPNN model optimized by WOA algorithm is significantly higher than that of BPNN model. The R^2 of training set and validation set of camphor growth inversion model constructed by WOA-BPNN combined with coefficient of variation method are 0.9960 and 0.9020, respectively, NRMSE are 0.0124 and 0.0722, respectively, and MAPE are 1% and 7%, respectively. The WOA-BPNN model is 13% and 18% higher than the BPNN model training set and validation set R^2 , respectively, the NRMSE is reduced by 85% and 33%, respectively, and the MAPE is reduced by 7% and 9%. This shows that the WOA algorithm can be used to optimize the growth inversion model of *C. camphora* dwarf forest constructed by BPNN and coefficient of variation method. This algorithm can be applied to the establishment of comprehensive growth monitoring model of *C. camphora* dwarf forest.

4. Discussion

The correlation between the growth and multispectral vegetation indices of *C. camphora* revealed that among the four single-growth indices, SPAD, AGB, and LAI had high and similar correlation with the multispectral index, whereas PH had a low correlation with the multispectral vegetation index. The multispectral vegetation index can reflect information on plant chlorophyll content, cover, and density, although PH had no direct relationship with these data; therefore, PH had a low correlation with the multispectral vegetation indices [48]. CGMI₁ and CGMI₂ comprehensive growth indices were more correlated with multispectral vegetation indices than the four single-growth indices, presumably because comprehensive growth indices provide more information than single-growth indices. The five vegetation indices:

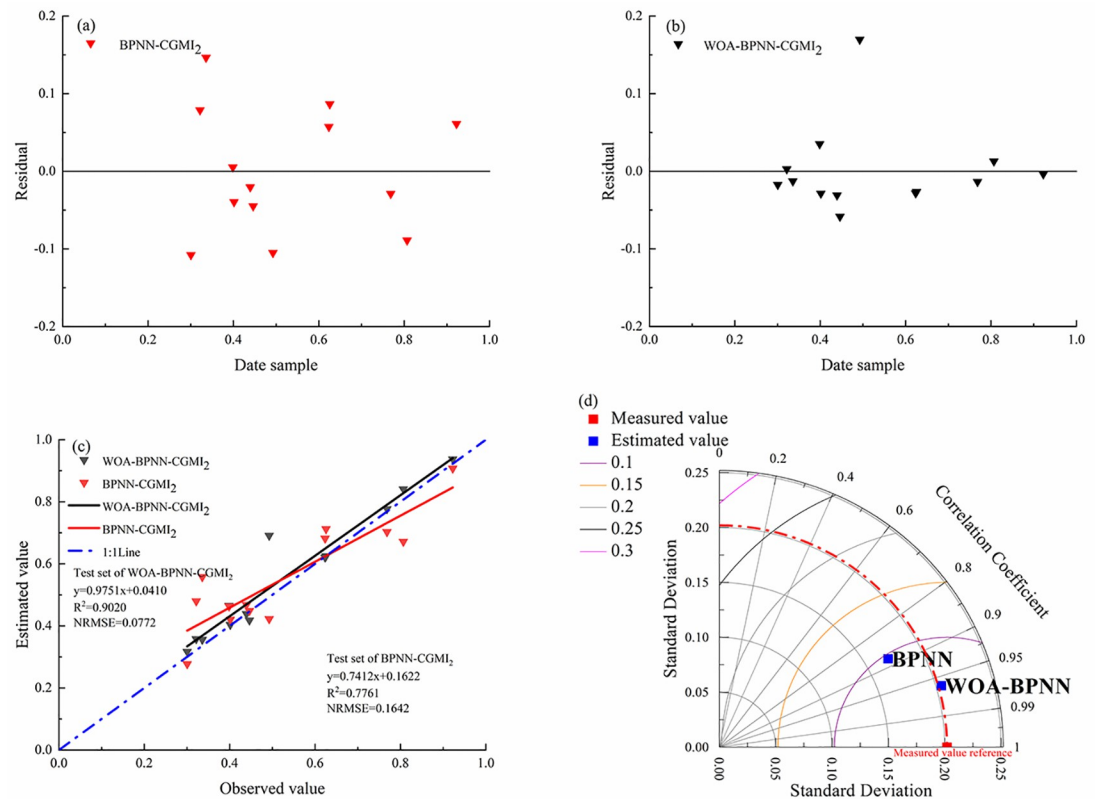


Fig 7. Optimized model fitting results. (a) Residual plot of the BPNN-CGMI₂; (b) Residual plot of the WOA-BPNN-CGMI₂; (c) Comparison of the fitting results of the validation set; (d) Taylor diagram of the test set.

<https://doi.org/10.1371/journal.pone.0299362.g007>

TCART, $Cl_{rededge}$, OSaVI, DVI, and MNVI, were more correlated with *C. camphora* growth, with a correlation with CGMI₂ of 0.79, 0.65, 0.65, 0.67, and 0.67, respectively, indicating that CGMI₂ correlated more significantly with NIR band multispectral reflectance and chlorophyll-related vegetation indices.

The *C. camphora* comprehensive growth monitoring model's accuracy based on the comprehensive growth index CGMI₂ of the CV method was significantly higher than that of the model based on the comprehensive growth index CGMI₁ of the equal weighting method. This result was similar to that of a previous study using the CV method to model the comprehensive growth of winter wheat [9]. This may be because the CV method assigns higher weights to a single index with larger coefficients of variation, whereas the equal weight method assigns the same weight to every vegetation index, making the information in CGMI₂ more distinguishable and stable than that in CGMI₁, leading to higher model accuracy built using CGMI₂.

The BPNN model's accuracy in the vegetation index was higher than that of the MLR, PLS, SVR, RF, and RBFNN models, indicating that the BPNN model more effectively estimates *C. camphora* CGMI than other models. The MLR and PLS models are regression models based on linear assumptions. Although these methods can control the relationship between multiple independent variables and one dependent variable, they cannot adequately capture the nonlinear relationship, resulting in underfitting or overfitting. However, in the case of a nonlinear relationship between vegetation index and comprehensive growth of *C. camphora*, the accuracy of the MLR and PLS models may be affected [49, 50]. RF could not solve regression problems as adequately as classification problems and failed to provide a continuous output. When

regression analysis is performed, some specific noisy data are prone to overfitting in modeling [51]. SVR and RBFNN models can control nonlinear problems better and have faster training speeds than the BPNN model; however, they are mainly applicable to low-dimensional datasets, and their training time and computational cost are higher for high-dimensional datasets, which may affect the accuracy of SVR and RBFNN models [51, 52]. The higher BPNN model accuracy may be attributed to its high nonlinear function approximation capability, self-learning and self-adaptive capabilities, and fault tolerance to cope with measurement errors, making it more suitable for the inversion of CGMI of *C. camphora* than other models [53].

The integrated growth model based on WOA optimization improved R^2 by 12% compared with the previous model for SPAD inversion only [21], indicating that CGMI based on WOA optimization is more advantageous than a single growth index for dwarf forest arborvitae growth inversion. After replacing the BPNN randomly assigned weights and thresholds with the optimal weights and thresholds obtained by the WOA algorithm while keeping the remaining model parameters consistent, the model R^2 and stability were significantly improved, and NRMSE was significantly decreased. Furthermore, the estimated values of the WOA-BPNN model were closer to the measured values compared with those of the BPNN model after plotting scatter plots with the same data, and the model accuracy was improved, indicating that the WOA algorithm can improve the prediction accuracy of the CGMI2-BPNN model for *C. camphora*.

The inversion model derived from this study only applies to the present measurement data, and further research is needed to estimate the CGMI of *C. camphora* in different growth cycles, species, regions, and periods. In addition, the approximation of calculation formulae and some spectral information, which may produce numerous highly correlated vegetation indices when six narrowband multispectral sensors are used, may cause information redundancy and co-linear effects in the model. This study attempted to avoid this problem through vegetation indices screening. Therefore, future model optimization studies can increase the sample size and use hyperspectral technology to screen the spectral bands and vegetation indices that are more sensitive to the *C. camphora* growth index. Furthermore, we discussed and developed models for different *C. camphora* species and growth periods to obtain a more accurate and comprehensive *C. camphora* growth inversion model.

5. Conclusions

In this study, we constructed CGMI based on field and laboratory data, analyzed the correlation between CGMI and selected vegetation indices using the basic principles of the CV method and traditional weighting method, and combined the constructed CGMI₁, CGMI₂, and UAV multispectral data with various growth analysis models and WOA model to achieve effective monitoring of *C. camphora* growth in the study area. Specific conclusions were obtained as follows:

The correlation analysis of the obtained data of SPAD, PH, LAI, and AGB of *C. camphora*, combined with different weighting methods, showed that the correlation of the constructed CGMI increased to different degrees compared with single-growth indices, and most of them showed significant correlation. The TCART, $Cl_{rededge}$, OSAVI, DVI, and MNVI were highly correlated with *C. camphora* growth indices, their correlations with CGMI₂ were 0.79, 0.65, 0.65, 0.67 and 0.67, respectively.

The CV method could better reverse the comprehensive growth of *C. camphora* than the equal weight method. In the model of comprehensive growth indicators of *C. camphora* constructed by the equal weight method, the validation set R^2 of MLR, PLS, SVR, RF, RBFNN and BPNN models were 0.3408, 0.4632, 0.5416, 0.5742, 0.5649, and 0.6332, respectively; and in the

model of comprehensive growth indicators of arborvitae constructed by combining with the method of coefficient of variation. The coefficients of determination R^2 were 0.6403, 0.5102, 0.6000, 0.6840, 0.6646, and 0.7761, respectively.

The training set and validation set R^2 of the WOA-BPNN model with the coefficient of variation method were 0.9960 and 0.9020, the NRMSE was 0.0124 and 0.0722, respectively, and MEAP are 1% and 7%, respectively. The WOA-BPNN model's accuracy optimized using the WOA algorithm was significantly improved compared with the BPNN model. The coefficient of determination increased by 18% and the NRMSE decreased by 33% compared with the BPNN model. Therefore, the WOA-BPNN model is the best model to invert the growth potential of *C. camphora*.

Future modeling studies could increase the sample size and discuss the model for different growth periods of *C. camphora* and try new algorithms to optimize the model as well as screening vegetation and texture indices that are more sensitive to camphor pine growth indices.

Supporting information

S1 Data.

(XLSX)

S2 Data.

(XLSX)

S3 Data.

(XLSX)

S1 File.

(ZIP)

Author Contributions

Conceptualization: Rongxin Gong, Haina Zhang, Xianghui Lu, Jie Zhang.

Data curation: Rongxin Gong, Xin Luo.

Formal analysis: Rongxin Gong, Xianghui Lu, Rongxiu Xie.

Funding acquisition: Haina Zhang, Xianghui Lu, Jie Zhang.

Investigation: Rongxin Gong, Xianghui Lu, Haolong Wan, Yue Zhang.

Methodology: Rongxin Gong.

Project administration: Xianghui Lu, Jie Zhang.

Resources: Rongxin Gong.

Software: Rongxin Gong.

Supervision: Haina Zhang, Xianghui Lu, Jie Zhang.

Validation: Rongxin Gong, Xianghui Lu.

Visualization: Rongxin Gong.

Writing – original draft: Rongxin Gong.

Writing – review & editing: Rongxin Gong, Haina Zhang, Xianghui Lu.

References

1. Oré G, Alcántara MS, Góes JA, Oliveira LP, Yepes J, Teruel B, et al. Crop growth monitoring with drone-borne DInSAR. *Remote Sens (Basel)*. 2020;12. <https://doi.org/10.3390/rs12040615>
2. Wang L, Zhang G, Wang Z, Liu J, Shang J, Liang L. Bibliometric analysis of remote sensing research trend in crop growth monitoring: A case study in China. *Remote Sens (Basel)*. 2019;11. <https://doi.org/10.3390/rs11070809>
3. Shammi SA, Meng Q. Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling. *Ecol Indic*. 2021;121. <https://doi.org/10.1016/j.ecolind.2020.107124>
4. Wang Z, Ye L, Jiang J, Fan Y, Zhang X. Review of application of EPIC crop growth model. *Ecological Modelling*. 2022. <https://doi.org/10.1016/j.ecolmodel.2022.109952>
5. Zhao J, Zhang X, Gao C, Qiu X, Tian Y, Zhu Y, et al. Rapid mosaicking of unmanned aerial vehicle (UAV) images for crop growth monitoring using the SIFT algorithm. *Remote Sens (Basel)*. 2019;11. <https://doi.org/10.3390/rs11101226>
6. Liu S, Zhang B, Yang W, Chen T, Zhang H, Lin Y, et al. Quantification of Physiological Parameters of Rice Varieties Based on Multi-Spectral Remote Sensing and Machine Learning Models. *Remote Sens (Basel)*. 2023;15. <https://doi.org/10.3390/rs15020453>
7. Barrero O, Perdomo SA. RGB and multispectral UAV image fusion for Gramineae weed detection in rice fields. *Precis Agric*. 2018;19. <https://doi.org/10.1007/s11119-017-9558-x>
8. Xu Y, Cheng Q, Wei X, Yang B, Xia S, Rui T, et al. Monitoring of winter wheat growth under UAV using variation coefficient method and optimized neural network. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*. 2021;37. <https://doi.org/10.11975/j.issn.1002-6819.2021.20.008>
9. Xu X, Nie C, Jin X, Li Z, Zhu H, Xu H, et al. A comprehensive yield evaluation indicator based on an improved fuzzy comprehensive evaluation method and hyperspectral data. *Field Crops Res*. 2021;270. <https://doi.org/10.1016/j.fcr.2021.108204>
10. Roscher R, Bohn B, Duarte MF, Garcke J. Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access*. 2020;8. <https://doi.org/10.1109/ACCESS.2020.2976199>
11. Han Y, Tang R, Liao Z, Zhai B, Fan J. A Novel Hybrid GOA-XGB Model for Estimating Wheat Above-ground Biomass Using UAV-Based Multispectral Vegetation Indices. *Remote Sens (Basel)*. 2022;14. <https://doi.org/10.3390/rs14143506>
12. Lu J, Qiu H, Zhang Q, Lan Y, Wang P, Wu Y, et al. Inversion of chlorophyll content under the stress of leaf mite for jujube based on model PSO-ELM method. *Front Plant Sci*. 2022;13. <https://doi.org/10.3389/fpls.2022.1009630> PMID: 36247579
13. Mirjalili S, Lewis A. The Whale Optimization Algorithm. *Advances in Engineering Software*. 2016;95. <https://doi.org/10.1016/j.advengsoft.2016.01.008>
14. Mostafa Bozorgi S, Yazdani S. IWOA: An improved whale optimization algorithm for optimization problems. *J Comput Des Eng*. 2019;6. <https://doi.org/10.1016/j.jcde.2019.02.002>
15. Brodzicki A, Piekarski M, Jaworek-Korjakowska J. The whale optimization algorithm approach for deep neural networks. *Sensors*. 2021;21. <https://doi.org/10.3390/s21238003> PMID: 34884004
16. Yu F, Bai J, Jin Z, Zhang H, Guo Z, Chen C. Research on Precise Fertilization Method of Rice Tillering Stage Based on UAV Hyperspectral Remote Sensing Prescription Map. *Agronomy*. 2022;12. <https://doi.org/10.3390/agronomy12112893>
17. Du Y, Zhou H, Yang L, Jiang L, Chen D, Qiu D, et al. Advances in Biosynthesis and Pharmacological Effects of *Cinnamomum camphora* (L.) Presl Essential Oil. *Forests*. 2022. <https://doi.org/10.3390/f13071020>
18. Xu Y, Qin J, Wang P, Li Q, Yu S, Zhang Y, et al. Chemical composition and larvicidal activities of essential oil of *cinnamomum camphora* (L.) leaf against *anopheles stephensi*. *Rev Soc Bras Med Trop*. 2020;53. <https://doi.org/10.1590/0037-8682-0211-2019> PMID: 31994661
19. Lee SH, Kim DS, Park SH, Park H. Phytochemistry and Applications of *Cinnamomum camphora* Essential Oils. *Molecules*. 2022. <https://doi.org/10.3390/molecules27092695> PMID: 35566046
20. Zhang Z, Wu X, Lai Y, Li X, Zhang D, Chen Y. Efficient Extraction of Bioenergy From *Cinnamomum camphora* Leaves. *Front Energy Res*. 2020;8. <https://doi.org/10.3389/fenrg.2020.00090>
21. Lu X, Wang Q, Zhang H, Gong R, Zhang J, Yang B. Inversion of SPAD of *Cinnamomum camphora* Dwarf Forest Based on UAV Multispectral Remote Sensing. *Nongye Jixie Xuebao/Transactions of the Chinese Society for Agricultural Machinery*. 2023;54. <https://doi.org/10.6041/j.issn.1000-1298.2023.05.020>
22. Lu X, Yang B, Zhang H, Zhang J, Wang Q, Jin Z. Inversion of Leaf Essential Oil Yield of *Cinnamomum camphora* Based on UAV Multi-spectral Remote Sensing. *Nongye Jixie Xuebao/Transactions of the*

- Chinese Society for Agricultural Machinery. 2023;54. <https://doi.org/10.6041/j.issn.1000-1298.2023.04.018>
23. Rahman MM, Robson A. Integrating landsat-8 and sentinel-2 time series data for yield prediction of sugarcane crops at the block level. *Remote Sens (Basel)*. 2020;12. <https://doi.org/10.3390/RS12081313>
24. Chen JM. Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*. 1996;22. <https://doi.org/10.1080/07038992.1996.10855178>
25. Zhang L, Zhang H, Niu Y, Han W. Mapping maize water stress based on UAV multispectral remote sensing. *Remote Sens (Basel)*. 2019;11. <https://doi.org/10.3390/rs11060605>
26. Ma Y, Jiang Q, Wu X, Zhu R, Gong Y, Peng Y, et al. Monitoring hybrid rice phenology at initial heading stage based on low-altitude remote sensing data. *Remote Sens (Basel)*. 2021;13. <https://doi.org/10.3390/rs13010084>
27. Guha S, Govil H. An assessment on the relationship between land surface temperature and normalized difference vegetation index. *Environ Dev Sustain*. 2021;23. <https://doi.org/10.1007/s10668-020-00657-6>
28. Fern RR, Foxley EA, Bruno A, Morrison ML. Suitability of NDVI and OSAVI as estimators of green biomass and coverage in a semi-arid rangeland. *Ecol Indic*. 2018;94. <https://doi.org/10.1016/j.ecolind.2018.06.029>
29. Du M, Noguchi N, Ito A, Shibuya Y. Correlation Analysis of Vegetation Indices Based on Multi-temporal Satellite Images and Unmanned Aerial Vehicle Images with Wheat Protein Contents. *Engineering in Agriculture, Environment and Food*. 2021;14. https://doi.org/10.37221/EAFF.14.3_86
30. Tan C, Zhang Z, Xu C, Ma Y, Yao Z, Wei G, et al. Soil water content inversion model in field maize root zone based on UAV multispectral remote sensing. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*. 2020;36. <https://doi.org/10.11975/j.issn.1002-6819.2020.10.008>
31. Yao X, Wang N, Liu Y, Cheng T, Tian Y, Chen Q, et al. Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. *Remote Sensing*. 2017. <https://doi.org/10.3390/rs9121304>
32. Bhunia GS, Shit PK, Pourghasemi HR, Edalat M. Prediction of Soil Organic Carbon and its Mapping Using Regression Analyses and Remote Sensing Data in GIS and R. *Spatial Modeling in GIS and R for Earth and Environmental Sciences*. 2019. <https://doi.org/10.1016/b978-0-12-815226-3.00019-3>
33. Soulopoulou P, Padilla Blanco MA, Molina JAA, Damos P. Relation between Climate, Agricultural Land Cover and Bird Richness in a Mediterranean Natura 2000 Wetland. *European Journal of Environment and Earth Sciences*. 2022;3. <https://doi.org/10.24018/ejgeo.2022.3.2.280>
34. Pei H, Feng H, Li C, Jin X, Li Z, Yang G. Remote sensing monitoring of winter wheat growth with UAV based on comprehensive index. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*. 2017;33. <https://doi.org/10.11975/j.issn.1002-6819.2017.20.010>
35. Li Q, Meng XX, Liu YB, Pang LF. Risk Assessment of Floor Water Inrush Using Entropy Weight and Variation Coefficient Model. *Geotechnical and Geological Engineering*. 2019;37. <https://doi.org/10.1007/s10706-018-0702-9>
36. Jalililal Z, Amiri A, Castagliola P, Khoo MBC. Monitoring the coefficient of variation: A literature review. *Comput Ind Eng*. 2021;161. <https://doi.org/10.1016/j.cie.2021.107600>
37. Korkmaz M. A study over the general formula of regression sum of squares in multiple linear regression. *Numer Methods Partial Differ Equ*. 2021;37. <https://doi.org/10.1002/num.22533>
38. Chauhan VK, Dahiya K, Sharma A. Problem formulations and solvers in linear SVM: a review. *Artificial Intelligence Review*. 2019. <https://doi.org/10.1007/s10462-018-9614-6>
39. Zhang FO'Donnell LJ. Support vector regression. *Machine Learning: Methods and Applications to Brain Disorders*. 2019. <https://doi.org/10.1016/B978-0-12-815739-8.00007-9>
40. Schonlau M, Zou RY. The random forest algorithm for statistical learning. *Stata Journal*. 2020;20. <https://doi.org/10.1177/1536867X20909688>
41. Monecke A, Leisch F. SemPLS: Structural equation modeling using partial least squares. *J Stat Softw*. 2012;48. <https://doi.org/10.18637/jss.v048.i03>
42. Chouhan SS, Kaul A, Singh UP, Jain S. Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology. *IEEE Access*. 2018;6. <https://doi.org/10.1109/ACCESS.2018.2800685>
43. Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*. 2004;42. <https://doi.org/10.1109/TGRS.2004.831865>
44. Jin C, Jin SW, Qin LN. Attribute selection method based on a hybrid BPNN and PSO algorithms. *Applied Soft Computing Journal*. 2012;12. <https://doi.org/10.1016/j.asoc.2012.03.015>

45. Peng Y, Zhao L, Hu Y, Wang G, Wang L, Liu Z. Prediction of soil nutrient contents using visible and near-infrared reflectance spectroscopy. *ISPRS Int J Geoinf*. 2019;8. <https://doi.org/10.3390/ijgi8100437>
46. Chen N, Xiong C, Du W, Wang C, Lin X, Chen Z. An improved genetic algorithm coupling a back-propagation neural network model (IGA-BPNN) for water-level predictions. *Water (Switzerland)*. 2019;11. <https://doi.org/10.3390/w11091795>
47. Khadka M, Paz A. Comprehensive Clusterwise Linear Regression for Pavement Management Systems. *Journal of Transportation Engineering, Part B: Pavements*. 2017;143. <https://doi.org/10.1061/jpeodx.0000009>
48. Xue J, Su B. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*. 2017. <https://doi.org/10.1155/2017/1353691>
49. Abdikan S, Sekertekin A, Narin OG, Delen A, Balik Sanli F. A comparative analysis of SLR, MLR, ANN, XGBoost and CNN for crop height estimation of sunflower using Sentinel-1 and Sentinel-2. *Advances in Space Research*. 2023;71. <https://doi.org/10.1016/j.asr.2022.11.046>
50. Farahani HA, Rahiminezhad A, Same L, Immanezhad K. A comparison of Partial Least Squares (PLS) and Ordinary Least Squares (OLS) regressions in predicting of couples mental health based on their communicational patterns. *Procedia—Social and Behavioral Sciences*. 2010. <https://doi.org/10.1016/j.sbspro.2010.07.308>
51. Ahmad MW, Reynolds J, Rezgui Y. Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *J Clean Prod*. 2018;203. <https://doi.org/10.1016/j.jclepro.2018.08.207>
52. Zhang G, Xia B, Wang J, Ye B, Chen Y, Yu Z, et al. Intelligent state of charge estimation of battery pack based on particle swarm optimization algorithm improved radical basis function neural network. *J Energy Storage*. 2022;50. <https://doi.org/10.1016/j.est.2022.104211>
53. Hairuddin NL, Yusuf LM, Othman MS. Gender classification on skeletal remains: Efficiency of meta-heuristic algorithm method and optimized back propagation neural network. *Journal of Information and Communication Technology*. 2020;19. <https://doi.org/10.32890/jict2020.19.2.5>