A Crown Contour Envelope Model of Chinese Fir Based on Random Forest and Mathematical Modeling

Yingze Tian 1,2,3, Baoguo Wu 1,2,3, Xiaohui Su 1,2,3,*, Yan Qi 4, Yuling Chen 1,2,3 and Zhiqiang Min 1,2,3

1 School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China; tianyingze@bjfu.edu.cn (Y.T.); wubg@bjfu.edu.cn (B.W.); chenyuling92@163.com (Y.C.); mzqli026@bjfu.edu.cn (Z.M.)
2 Engineering Research Center for Forestry-oriented Intelligent Information Processing, National Forestry and Grassland Administration, Beijing 100083, China
3 Forestry Information Research Institute, Beijing Forestry University, Beijing 100083, China
4 International College Beijing, China Agricultural University, Beijing 100083, China; icb7@cau.edu.cn

* Correspondence: suxhui@bjfu.edu.cn

Abstract: The tree crown is an important part of a tree and is closely related to forest growth status, forest canopy density, and other forest growth indicators. Chinese fir (Cunninghamia lanceolata (Lamb.) Hook) is an important tree species in southern China. A three-dimensional (3D) visualization assistant decision-making system of plantations could be improved through the construction of crown contour envelope models (CCEMs), which could aid plantation production. The goal of this study was to establish CCEMs, based on random forest and mathematical modeling, and to compare them. First, the regression equation of a tree crown was calculated using the least squares method. Then, forest characteristic factors were screened using methods based on mutual information, recursive feature elimination, least absolute shrink and selection operator, and random forest, and the random forest model was established based on the different screening results. The accuracy of the random forest model was higher than that of the mathematical modeling. The best performing model based on mathematical modeling was the quartic polynomial with the largest crown radius as the variable (R-squared (R^2) = 0.8614 and root mean square error (RMSE) = 0.2657). Among the random forest regression models, the regression model constructed using mutual information as the feature screening method was the most accurate (R^2 = 0.886, RMSE = 0.2406), which was two percentage points higher than mathematical modeling. Compared with mathematical modeling, the random forest model can reflect the differences among trees and aid 3D visualization of a Chinese fir plantation.

Keywords: Chinese fir; crown contour envelope model (CCEM); random forest; tree factors; mutual information (MI)

1. Introduction

The tree crown is an important part of a tree that reflects the growth status of individual tree, and also reflects the adaptation and variation degree of trees to different growth environments [1,2]. Significant physiological processes such as photosynthesis, respiration, and transpiration take place in the tree crown. The ecological environment in the tree crown is also an important component of the forest ecosystem. The shape of the tree crown and the distribution of its leaves affect the interception of rainfall and the utilization of solar energy [3–6]. The tree crown structure affects the growth of trees and also the dynamic changes of forest stands. Therefore, the study of the shape of tree crowns is of great significance. Visualization of the tree crown provides the basis for forest stand dynamic visualization simulation and is a major research topic in forestry informatization. It is also important for a three-dimensional (3D) visualization assistant decision-making system of plantations, because it permits direct observation of the growth status of plantations and
calculations of crown density. The shape of tree crowns is often considered to be a feature of space geometry with boundary \[7,8\]. Using the vertical plane of the trunk to cut the tree crown, the closed intersection line is called the crown contour envelope of the tree crown. Similarly, the crown contour envelope can be rotated around the trunk to the tree crown.

Many methods have been used to study tree crown contour envelope, such as the fractal method, the simple geometry method, and the mathematical model method. The fractal method is a geometry concept, proposed by Mandelbrot in the 1970s, which has been widely used in natural botany \[9\]. Paul Henning, Kangning Lu, and other scholars used the fractal method of tree growth modeling to construct crown contour envelope models (CCEMs) \[10,11\]. However, shortcomings of the fractal method include the difficulty of expressing tree crowns of different tree species and the fact that it cannot reflect the influence of a series of parameters on tree crown growth.

The simple geometry method uses simple geometry, such as cylinder, cone, parabola or their combination, to represent the shape of tree crowns. In early research, Gill, Bigging, Hann, Marshall, and other scholars defined the tree crown shape of different tree species, such as Douglas fir (\textit{Pseudotsuga menziesii} (Mirbel) Franco) and eastern hemlock (\textit{Tsuga canadensis} (L.) Carrière), and different growth stages of the same tree species, as regular geometric bodies such as cone, paraboloid, ellipsoid, and cylinder, and established models to predict tree crown volume and crown radius at any crown height \[12–15\]. This method has long been used to study the relationship between crown size and tree growth. Enying Guo and Han disassembled the entire Chinese fir plantation into cone, circular truncated cone, and cylinder, and then constructed a model for tree crown morphology based on diameter at breast height DBH, tree height, crown radius and crown length. They used visualization technology to realize a 3D growth visual simulation of a Chinese Fir plantation \[16,17\]. Because it is difficult to measure tree crown factor, many researchers have used this simple and convenient method to predict tree crown shape. However, this method cannot be used to completely describe tree crown given that the tree crown shape of trees in stands is irregular.

The mathematical model method involves using mathematical equations to simulate tree crown contour envelope at certain growth stages. Crecente Campo, Hann, Yanrong Guo, and other scholars used simple polynomials and deformation piecewise function equations to define tree crown contour envelope \[18,19\]. Some researchers have used an improved Kozak, Weibull, and other special curves to define the crown shape as well as different methods for building the CCEMs \[20\]. This method can accurately describe the crown shape with a small number of tree parameters. However, because of the small number of variables used, such models do not effectively reflect the differences in tree morphology and do not reflect complex dynamic changes of the crown in the process of tree growth. In addition, reparameterization of these models through the addition of relevant tree factors is complex.

With the development of artificial intelligence technology, machine learning provides a new method for forest growth and harvest prediction \[21\]. Machine learning has several advantages, including the lack of hypothesis for the distribution of input data, the ability to reveal the hidden structure of data, and robust prediction results. Machine learning has been widely used in forest research, including the prediction of tree height, DBH, and volume \[22–26\]. Drawbacks of mathematical modeling fitting methods include the difficulty of determining the form of the model, selecting tree factors, and determining the regression equation; in addition, the model forms or parameters of different tree species or different regions of the same tree species also differ. Thus, much work is required to revise these models. The machine learning method is free of these problems and can quickly generate models in line with research objectives to simulate tree crown.

Random forest is an integrated algorithm based on the classification tree proposed by Leo Breiman \[27\], which is a set of tree classifiers that uses the bootstrap repeated sampling method to extract samples from samples for modeling. The final output is a simple majority voting (classification) or average (regression) of output results of a single tree. The random
forest model is more robust to outliers and noise, is more rapid than boosting algorithm, and is less overfitted. This model has been applied to fire prediction, forest growth, and harvest prediction [28].

Fujian Province is located in North China and it has the highest forest coverage rate (66.8%) in China. Chinese fir is one of the most important plantation tree species in Fujian Province and accounts for 21.35% of the total plantation area in China. To date, few studies have been conducted on the crown contour envelope of Chinese fir plantations. Most CCEMs based on mathematical modeling have only used crown depth as an independent variable to predict crown shape, and few studies have added variables such as tree height and diameter at breast height as covariates to improve model accuracy. In addition, machine learning has not been used to predict Chinese fir crown shape. Thus, the goals of the present study were the following: (1) to collect classical CCEMs suitable for Chinese fir and use Chinese fir crown data in Fujian Province to fit them; (2) to use different feature selection methods to screen tree factors that affect the crown shape of Chinese fir, construct the random forest regression model, and then fit hyperparameters; and (3) to evaluate the CCEMs constructed by mathematical modeling and random forest regression model and compare them.

2. Materials and Methods

2.1. Study Area Data Collection

The study area is located in Fujian Province of China (between 23.3° to 28.3° N and 115.6° to 120.5° E), which features a subtropical monsoon humid climate. Fujian is the core area for Chinese fir central producing, as it has the highest level of cultivation worldwide. The data of this study were collected from the Dali forest farm, the Lanxia forest farm, and the Jiangle state-owned forest farm in Jiangle County, Shunchang County, Fujian Province. The terrain of the land is primarily hilly and mountainous, with forest coverage greater than 90%. The tree species are mainly Chinese fir and Eucalyptus (Eucalyptus robusta Smith). Shunchang County is the central area of Chinese fir production in Fujian Province and is known as the “hometown of Chinese fir”.

In this study, different age groups and different forest stand densities were used to set up temporary sample plots in the Chinese fir plantation. At the Dali forest farm, Lanxia forest farm, and Jiangle forest farm, 65, 23, and 3 standard plots, respectively, with a size of 30 m², were set up in sample plots of Chinese fir plantation. Three to five trees were selected in each standard plot, and a total of 423 trees were studied. Measurements taken on trees included diameter at breast height (DBH, cm), total tree height (HT, m), crown length (CL), height under branch (HBLC), largest crown radius (LCR), and crown radius (CR) at 1/10 CL, 1/4 CL, 1/2 CL, 3/4 CL, and 9/10 CL from crown top to crown bottom (Figure 1). In this study, we measured the crown radius in the north-south direction and the east-west direction and used the mean value as the variable to establish CCEMs.

The tree factors used to establish CCEMs were selected according to the factors affecting the growth of trees. The crown contour envelope is mainly affected by AGE, N, HT, DBH, CL, HBLC, and LCR. In addition, this paper also defines the following composite tree factors: the tree crown length ratio (CH), HT to DBH ratio (HD = HT/DBH), and tree crown diagonal coefficient (CLC = CL/LCR). To facilitate description of the crown envelope, we defined the perpendicular distance from any crown to the horizontal plane of crown top as DINCT, the perpendicular distance from any crown to the horizontal plane of crown bottom as DINCBL, the ratio of DINCT to CL as RDINCT, and the ratio of DINCBL to CL as RDINCBL. All factors related to crown and their descriptions are shown in Table 1.
The tree factors used to establish CCEMs were selected according to the factors affecting the growth of trees. The crown contour envelope is mainly affected by AGE, N, HT, DBH, CL, HBLC, and LCR. In addition, this paper also defines the following composite tree factors: the tree crown length ratio (CH), HT to DBH ratio (HD = HT/DBH), and tree crown diagonal coefficient (CLC = CL/LCR). To facilitate description of the crown contour envelope, we defined the perpendicular distance from any crown to the horizontal plane of crown top as DINCT, the perpendicular distance from any crown to the horizontal plane of crown bottom as DINCB, the ratio of DINCT to CL as RDINCT, and the ratio of DINCB to CL as RDINCB. All factors related to crown and their descriptions are shown in Table 1.

Table 1. Description of tree factors.

| Factors          | Type        | Computational Formula |
|------------------|-------------|-----------------------|
| HT (m)           | measuring factor |                        |
| DBH (cm)         | measuring factor |                        |
| CL (m)           | measuring factor |                        |
| HBLC (m)         | measuring factor |                        |
| LCR (m)          | measuring factor |                        |
| LCD (m)          | composite factor | LCD = 2 × LCR          |
| N (num/ha)       | basic factor  |                        |
| AGE (years)      | basic factor  |                        |
| CH               | composite factor | CH = CL/HT             |
| HD               | composite factor | HD = HT/DBH            |
| CLC              | composite factor | CLC = CL/LCR           |
| CR               | composite factor |                        |
| DINCT (m)        | measuring factor |                        |
| DINC_B (m)       | composite factor | DINC_B = CL − DINC_B   |
| RDINC_T          | composite factor | RDINC_T = DINC_T/CL    |
| RDINC_B          | composite factor | RDINC_B = DINC_B/CL    |

To construct the CCEMs, the data was divided into different datasets to form a training set and test set by the following two principles: First, to ensure the integrity of the tree crown data, the tree crown data of the same tree should be divided into the same dataset. Second, as there are major differences in the tree crown shapes of Chinese fir of different ages, the training set and the test set must contain Chinese fir of the same age, and the proportion of Chinese fir of different ages in the training set and test set should be similar.

According to the two principles above, the datasets were randomly divided 10 times, according to the 70% proportion of the training set. The coefficient of variation of important factors (AGE, CR, DBH, HT, and N) was calculated, and a group of partition results with the most similar coefficients of variation was selected as the training set and test set for the model construction. Basic information of the dataset is shown in Table 2.
### Table 2. Statistics of the fitting dataset.

| Variable | Dataset       | Number | Maximum  | Minimum  | Average | σ  | CV  |
|----------|---------------|--------|----------|----------|---------|----|-----|
|          | All set       | 2115   | 29.00    | 5.00     | 16.50   | 7.11 | 0.4310 |
| AGE      | Training set  | 1480   | 29.00    | 5.00     | 16.40   | 7.05 | 0.4301 |
|          | Test set      | 635    | 29.00    | 5.00     | 16.60   | 7.24 | 0.4352 |
|          | All set       | 2115   | 3.40     | 0.04     | 1.05    | 0.68 | 0.6530 |
| CR       | Training set  | 1480   | 3.40     | 0.04     | 1.03    | 0.67 | 0.6505 |
|          | Test set      | 635    | 3.10     | 0.05     | 1.09    | 0.71 | 0.6562 |
|          | All set       | 2115   | 31.30    | 6.10     | 16.10   | 5.60 | 0.3479 |
| DBH      | Training set  | 1480   | 28.60    | 6.10     | 16.00   | 5.58 | 0.3490 |
|          | Test set      | 635    | 31.30    | 6.60     | 16.50   | 5.62 | 0.3408 |
|          | All set       | 2115   | 3000.00  | 900.00   | 2492.00 | 597.00| 0.2394 |
| HT       | Training set  | 1480   | 3000.00  | 900.00   | 2492.00 | 597.00| 0.2394 |
|          | Test set      | 635    | 3000.00  | 900.00   | 2523.00 | 542.00| 0.2149 |
|          | All set       | 2115   | 22.80    | 3.00     | 12.50   | 3.92 | 0.3144 |
| N        | Training set  | 1480   | 31.30    | 6.10     | 16.10   | 5.60 | 0.3479 |
|          | Test set      | 635    | 3000.00  | 900.00   | 2492.00 | 597.00| 0.2394 |

Note, σ is standard deviation and CV is coefficient of variation.

2.2. Mathematical Modeling of Chinese Fir Crown Contour Envelope Model (CCEM)

In this study, we defined the geometric cross-section formed by any plane across the trunk and tree crown as the max crown profile. In this geometric cross-section, we constructed a plane rectangular coordinate system with the crown top as the origin, the direction of trunk as the X-axis, and the direction perpendicular to the trunk as the Y-axis. The red curve shown in Figure 2 is the crown contour envelope, and the CCEM was constructed with CR as the dependent variable; RDINC as the independent variable; and HT, DBH, HBLC, CL, and DINC as covariates. The crown contour can be rotated 360 degrees around the X-axis to obtain the entire tree crown, therefore, the crown shape can be studied through the CCEM.

![Crown contour envelope of Chinese fir.](image)

In early research, Gill and other scholars defined the crown shape of different tree species in different growth stages using some simple geometry, such as cylinder, cone, paraboloid, and so on [12–15]. This simple geometry was obtained by rotating the tree trunk with power function (straight line, linear function, throwing object line, etc.). The tree crown contour envelope based on this is expressed in the form of Model (1). In order to describe the tree crown shape more flexibly and reflect the difference between different tree species at different growth stages, McPherson modified Model (1) slightly to obtain Model (2) as follows:

\[
CR = LCR \times RDINC_T^a 
\]

(1)

\[
CR = LCR(1 - (1 - RDINC_T^2)^a)
\]

(2)

Baldwin constructed a CCEM to predict the crown radius at any position of *Pinus taeda* L. by using RDINC_3 as the variable [29–31]. The model described the vertical distribution of crown radius better. Crescent Campo added the LCR to Model (3), which was used to predict the tree crown shape of Pinus radiata. Chmura [32,33] combined Baldwin’s model
form and McPherson’s [34] reparameterization method to obtain Model (4) to predict the tree crown shape as:

\[
CR = \text{LCR}(a \times \frac{\text{RDINC}_B - 1}{\text{RDINC}_B + 1} + b(\text{RDINC}_B - 1))
\]

(3)
\[
CR = \text{LCR}(a \times \frac{\text{RDINC}_B - 1}{\text{RDINC}_B + 1} + b(1 - \text{RDINC}_B)^2)
\]

(4)

Cresent Campo et al. used RDINC_B as the variable and LCR as a constraint condition to construct CCEM successively by a quadratic polynomial, cubic polynomial, and quartic polynomial to simulate the crown shape of Radiata pine (Pinus radiata D. Don). Chen Dong [35] applied the following Models (5) to (8) to Chinese fir and achieved good results:

\[
CR = \text{LCR}(a + b \times \text{RDINC}_B + c \times \text{RDINC}_B^2)
\]

(5)
\[
CR = \text{LCR}(a + b \times \text{RDINC}_B + c \times \text{RDINC}_B^2 + d \times \text{RDINC}_B^3)
\]

(6)
\[
CR = \text{LCR}(a + b \times \text{RDINC}_B + c \times \text{RDINC}_B^2 + d \times \text{RDINC}_B^3 + e \times \text{RDINC}_B^4)
\]

(7)
\[
CR = \text{LCR}(a(\text{RDINC}_B - 1) + b(\text{RDINC}_B^2 - 1) + c(\text{RDINC}_B^3 - 1) + d(\text{RDINC}_B^4 - 1))
\]

(8)

The variable exponential model can predict the radius of the crown at different positions of the crown by changing exponential form with the change in the relative position of the tree crown. Hann [14,34] and Yanrong Guo [19] described the tree crown contour of Chinese fir using Models (9) and (10), respectively as follows:

\[
CR = \text{LCR} \times \text{Exp}(a + b \times \text{RDINC}_B + c \times \text{RDINC}_B^2)
\]

(9)
\[
CR = \text{LCR} \times \text{Exp}(a + b \times \text{Exp}(c \times \text{RDINC}_B))
\]

(10)

Kozak proposed and revised the trunk equation, a variable exponential equation. Because of the high degree of similarity between the trunk equation and the CCEM, Maguire, Garber, Weiskittle, Huilin Gao, and others modified and reparameterized the Kozak equation by adding tree factors DBH, CH, and HD [36,37]. This equation is more consistent with the characteristics of the tree crown contour. The basic form of the modified model is shown in Model (11) as:

\[
CR = a \times \text{DBH}^b \left( \frac{1 - (1 - \text{RDINC}_T)^{0.5}}{1 - (c \times \text{CH}^d)^{0.5}} \right)
\]

(11)

Wang Chengde added DBH to Model (1) and HT and N to Model (9) to predict the crown contour of Chinese fir and eucalyptus. The modified model is shown in Models (12)–(14) [38] as follows:

\[
CR = \text{LCR} \times \text{RDINC}^{a \times (100 \times \text{DBH})^b}
\]

(12)
\[
CR = \text{LCR} \times \text{Exp}(a + b \times \text{HT}^c \times \text{LCR}^d \times \text{RDINC}_B + e \times \text{RDINC}_B^2)
\]

(13)
\[
CR = \text{LCR} \times \text{Exp}(a + b \times \text{RDINC}_B + (c \times \log 10N) \times \text{RDINC}_B^2)
\]

(14)

The least squares method is widely used in model parameter fitting because it can find the best function matching by minimizing the sum of squares of errors. The above models are all multivariate nonlinear models. Among them, Model (1) and Model (2) can be transformed into simple linear regression models, and Models (3)–(9) can be transformed into multiple linear models for regression [39].

Generally, linear regression can be solved by using the general least squares method, but the nonlinear regression problem, such as that associated with Models (10)–(14), is more difficult to address. The Levenberg–Marquardt (L-M) algorithm is the most widely used nonlinear least squares algorithm. It was proposed by D.W.Marquardt in 1963 [40].
This method is a combination of steepest descent method and linearization method (Taylor series). The steepest descent method is suitable for the initial stage of iteration when the parameter estimation value is far from the optimal value, and the linearization method (Gauss–Newton method) is suitable for the later stage of iteration when the parameter estimation value is close to the optimal value [41]. The optimal value can be found quickly by combining these two methods.

2.3. Chinese Fir CCEM Based on Random Forest

2.3.1. Feature Selection

When the random forest method is used for regression, over fitting is likely to occur if there are too many features in the dataset. Some algorithms can be used to generate a dataset of the importance of each feature. With this dataset, a threshold can be determined and some features that are most helpful for model training can be selected. Then, model training can be carried out after selecting important variables. The common feature selection methods include filter, embedded, and wrapper [42]. In this study, the mutual information method (MI), recursive feature elimination (RFE), least absolute shrink and selection operator (LASSO), and random forest (RF) were used for feature selection. Each method is described below as follows:

**Mutual information method (MI):** MI indicates whether the two variables \(X\) and \(Y\) are related, and the strength of the relationship [43]. If \((X, Y) \sim p (X, Y)\), the mutual information \(I (X, Y)\) between \(X\) and \(Y\) is defined as:

\[
I(X, Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right)
\] (15)

If \(X\) and \(Y\) are closely related, \(I (X, Y)\) will be larger. The minimum value of \(I (X, Y)\) is \(H(Y)\); at this time, \(H (Y|X)\) is 0, meaning that \(X\) and \(Y\) are completely related. When \(X\) is determined, \(Y\) is a constant value, and there is no probability of other uncertain situations; thus, \(H (Y|X)\) is 0. When \(I (X, Y)\) is taken as 0, \(X\) and \(Y\) are independent, and \(H (Y) = H (Y|X)\), which means that the appearance of \(X\) does not affect \(Y\).

**Recursive feature elimination (RFE):** RFE is a greedy algorithm for finding the optimal feature subset. The main idea is to repeatedly build the model (regression model), select the best (or worst) features, separate the selected features, and then repeat the process on the remaining features until all features are traversed [44]. The order in which features are eliminated in this process is the sequence of features.

**Least absolute shrink and selection operator (LASSO):** LASSO was first proposed by Robert Tibshirani in 1996. This method is a type of compressed estimation [45]. By constructing a penalty function, a more refined model is obtained, which compresses some regression coefficients, that is, the sum of absolute values of mandatory coefficients is less than a fixed value, and some regression coefficients are set to 0. Therefore, this method retains the advantage of subset contraction and provides a biased estimation for data with multicollinearity. Lasso adds a norm LP as a penalty constraint in the calculation of minimizing the sum of squares of residuals (RSS). The advantage of lp-norm is that when \(\lambda\) is sufficiently large, some coefficients to be estimated can be reduced to 0 accurately. The determination of \(\lambda\) is carried out by cross-validation method by first cross verifying the given value of \(\lambda\) and selecting the minimum error of cross-validation. According to the obtained value of \(\lambda\), the model can be refitted with all of the data.

**Random forest:** Random forest consists of multiple decision trees. Each node in the decision tree is a condition about a feature that divides the dataset into two according to different response variables. The node can be determined by using the impurity, for regression problems, the least squares method is typically used to fit models. When training the decision tree, we can calculate how much Gini impurity is reduced by each feature in the decision tree. For a decision tree forest, we can calculate how much Gini impurity is
reduced by each feature, and then take the average reduction of Gini impurity as the value of feature selection. The formula of Gini impurity is as follows:

\[ GI_m = 1 - \sum_{k=1}^{|k|} p_{mk}^2 \]  \hspace{1cm} (16)

where \( k \) indicates the number of categories and \( p_{mk} \) is the proportion of node \( k \) in node \( m \). Intuitively, the Gini impurity represents the probability of obtaining different categories from two samples that are randomly selected from node \( m \).

2.3.2. Hyperparameter Optimization

The random forest regression model is a set of \( n \) decision trees \( \{T_1(X), \ldots, T_n(X)\} \), where \( x = [x_1, \ldots, x_p] \) is the \( p \)-dimensional vector of the features related to the target variables; the result is the output value of \( n \) trees \( \{Y_1 = T_1(X), \ldots, Y_n = T_n(X)\} \), and \( Y_n \) is the predicted value of the \( n \)th tree. For regression problems, \( Y \) is the average of the predicted values of a single tree [46–48]. The random forest model has the following five important parameters: the maximum depth of the tree (max_depth), the number of features in the feature subset (max_features), the minimum number of leaf node samples (min_samples_leaf), the minimum number of node partition samples (min_samples_split), and the number of decision trees (n_estimators). With the root mean square error (RMSE) as the evaluation index and \( M = \{(X_1, Y_1), \ldots, (X_n, Y_n)\} \) as the training set, the training process of the random forest regression model is as follows:

1. Randomly generate \( m \) variables for the binary tree on the node, the choice of binary tree variables satisfies the principle of minimum Gini impurity.
2. Use bootstrap combination method to sample with replacement \( k \) sample sets in \( M \) to form \( k \) decision trees; then, the unsampled samples are used for the prediction of a single decision tree.
3. According to the random forest composed of \( k \) decision trees, the final result is the average output of each decision tree.

GridSearchCV and RandomizedSearchCV are two commonly used methods in hyperparametric optimization. The principle of GridSearchCV is simple, i.e., the program tries each set of hyperparameters one by one, and then selects the best group. This process is time-consuming and is a dimensional disaster. In 2012, James Bergstra and Yoshua Bengio proposed the RandomizedSearchCV method for parametric optimization [47]. The introduction of RandomizedSearchCV can effectively improve the efficiency of optimization, but the solution is not necessarily the optimal solution. This is the first study to use RandomizedSearchCV to obtain the approximate range of the optimal solution, and then GridSearchCV is used to obtain the optimal solution by specifying a small range for the obtained results.

2.4. Model Evaluation and Validation

For the goodness of fit of each model, \( R^2 \), MAE, MSE and RMSE were used as the test indexes in this study. The calculation formula of each test index is as follows:

\[ R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \]  \hspace{1cm} (17)

\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |(y_i - \hat{y}_i)| \]  \hspace{1cm} (18)

\[ MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]  \hspace{1cm} (19)
\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}
\]  

(20)

In the above formulas, \( n \) is the number of observed samples, \( y_i \) is the actual crown radius of the \( i \)th observed tree, \( \hat{y}_i \) is the predicted crown radius of the \( i \)th observed tree, and \( \bar{y} \) is the average of the actual crown radius of all observed samples.

3. Results

3.1. Mathematical Modeling

The transformed Models (1)–(9) were regressed by the ordinary least squares algorithm, and the Models (10)–(14) were regressed by the least squares method, the parameter fitting results of each model are shown in Table 3.

Table 3. Fitting results of mathematical modeling parameters.

| Model  | a    | b    | c    | d    | e    | f    |
|--------|------|------|------|------|------|------|
| Model 1| 0.473| /    | /    | /    | /    | /    |
| Model 2| 0.858| /    | /    | /    | /    | /    |
| Model 3| 2.488| -3.174| /    | /    | /    | /    |
| Model 4| 0.265| 1.133| 0.480| /    | /    | /    |
| Model 5| 0.793| 0.573| -1.303| /    | /    | /    |
| Model 6| 0.922| -0.705| 1.529| -1.702| /    | /    |
| Model 7| 0.836| 0.504| -3.158| 4.925| -3.066| /    |
| Model 8| 0.582| -3.561| 5.695| -3.548| /    | /    |
| Model 9| -0.324| 1.547| -2.992| /    | /    | /    |
| Model 10| -0.178| -0.003| 6.940| /    | /    | /    |
| Model 11| 0.302| 0.589| 3.998| 31.535| 0.477| -0.036|
| Model 12| 1.570| -0.423| /    | /    | /    | /    |
| Model 13| -0.328| 1.054| 0.106| 0.176| -3.012| /    |
| Model 14| -0.318| 1.511| -0.889| /    | /    | /    |

Note: the symbol “/” means the parameter is not included in the model.

Table 4 shows the evaluation results of the mathematical modeling Chinese fir CCEM. The models with the best fitting effect in the training set are Models (7), (6), (4), (10), and (5); in the test set, the models with the best fitting effect are Model (7), (6), (4), (12), and (10). Under the constraint of LCR, quartic polynomial and cubic polynomial with RDINC as the variable have the best fitting effect. With the addition of DBH to Model (1) (i.e., Model (12)), the fitting precision is slightly improved; with the addition of HT and N to Model (9) (i.e., Models (13) and (14), respectively), the improvement in fitting precision is not obvious. Therefore, the modified model suitable for specific areas or specific tree species is difficult to apply to other regions or other tree species. To obtain a more accurate CCEM suitable for Chinese fir, the basic model needs to be revised, and the process of adding variables and parameter fitting is difficult.

The residual plots of Models (4), (6), (7), and (10) are shown in Figures 3 and 4. Figure 3 shows the residual diagram of the training set models, and Figure 4 shows the residual diagram of the test set models. The residual error of both the training set and the test set models shows a clear trumpet shape and wide fluctuation range.
Table 4. Mathematical modeling evaluation results.

| Model  | Training Set | Test Set |
|--------|--------------|----------|
|        | $R^2$ | MAE | MSE | RMSE | $R^2$ | MAE | MSE | RMSE |
| Model 1 | 0.8421 | 0.1765 | 0.0712 | 0.2668 | 0.8487 | 0.1878 | 0.0771 | 0.2776 |
| Model 2 | 0.8044 | 0.1975 | 0.0882 | 0.2969 | 0.8127 | 0.2118 | 0.0954 | 0.3089 |
| Model 3 | 0.8327 | 0.1969 | 0.0754 | 0.2746 | 0.8276 | 0.2144 | 0.0878 | 0.2963 |
| Model 4 | 0.8578 | 0.1869 | 0.0640 | 0.2531 | 0.8594 | 0.1926 | 0.0716 | 0.2677 |
| Model 5 | 0.8521 | 0.1829 | 0.0666 | 0.2582 | 0.8496 | 0.1956 | 0.0766 | 0.2768 |
| Model 6 | 0.8593 | 0.1771 | 0.0634 | 0.2518 | 0.8610 | 0.1831 | 0.0708 | 0.2661 |
| Model 7 | 0.8602 | 0.1768 | 0.0630 | 0.2510 | 0.8614 | 0.1833 | 0.0706 | 0.2657 |
| Model 8 | 0.8579 | 0.1869 | 0.0640 | 0.2531 | 0.8593 | 0.1928 | 0.0717 | 0.2677 |
| Model 9 | 0.8099 | 0.2259 | 0.0857 | 0.2927 | 0.8038 | 0.2469 | 0.0999 | 0.3161 |
| Model 10 | 0.8547 | 0.1856 | 0.0655 | 0.2559 | 0.8536 | 0.1938 | 0.0746 | 0.2731 |
| Model 11 | 0.7376 | 0.2514 | 0.1182 | 0.3439 | 0.7429 | 0.2560 | 0.1310 | 0.3619 |
| Model 12 | 0.8480 | 0.1721 | 0.0685 | 0.2617 | 0.8545 | 0.1818 | 0.0741 | 0.2723 |
| Model 13 | 0.8160 | 0.2198 | 0.0829 | 0.2879 | 0.8081 | 0.2396 | 0.0978 | 0.3127 |
| Model 14 | 0.8103 | 0.2229 | 0.0855 | 0.2923 | 0.8058 | 0.2438 | 0.0989 | 0.3145 |

3.2. Random Forest

The results of MI feature screening are as follows: The value of $I(\text{CL}, \text{CR})$; $I(\text{CH}, \text{CR})$; $I(\text{HD}, \text{CR})$; and $I(\text{CLC}, \text{CR})$ are 0; the value of $I(\text{RDINC}_T, \text{CR})$; $I(\text{DINC}_T, \text{CR})$; $I(\text{RDINC}_B, \text{CR})$; and $I(\text{LCR}, \text{CR})$ are the four highest values; and other variables show correlations with CR (Figure 5). Therefore, $\text{RDINC}_T$, $\text{DINC}_T$, $\text{RDINC}_B$, $\text{LCR}$, $\text{DINC}_B$, $\text{AGE}$, $\text{DBH}$, $\text{N}$, $\text{HT}$, and $\text{HBLC}$ were used as reserved features.

Figure 3. Mathematical modeling residual plot (training set).
The RFE feature screening results are as follows: RMSE is positively correlated with the number of variables. The RMSE is lowest and the accuracy is the highest with 14 feature combinations, indicating that model accuracy does not continually improve as more variables are used for modeling. When the number of features is greater than nine, RMSE does not significantly change (Figure 6). Therefore, $DINC_T$, $LCR$, $RDINC_B$, $DINC_B$, $RDINC_T$, $AGE$, $HBLC$, $CLC$, and $DBH$ were used as the modeling features.
The RFE feature screening results are as follows: RMSE is positively correlated with the number of variables. The RMSE is lowest and the accuracy is the highest with 14 feature combinations, indicating that model accuracy does not continually improve as more variables are used for modeling. When the number of features is greater than nine, RMSE does not significantly change (Figure 6). Therefore, DINC_T, LCR, RDINC_B, DINC_B, RDINCT, AGE, HBLC, CLC, and DBH were used as the modeling features.

The LASSO screening results are as follows: The variables RDINC_T, CLC, HD, CH, LCR, HBLC, HT, and AGE are positively correlated with CR; DINC_B, and CL are negatively correlated with CR; and RDINC_B, DINC, DBH, and N are not correlated with CR (Figure 7). Therefore, RDINC_T, CLC, HD, CH, LCR, HBLC, HT, AGE, DINCD_B, and CL were used as modeling features.

The random forest screening results are as follows: The variables DINC_T, LCR, RDINC_B, RDINC_T, and DINC_B have higher scores; the variables AGE, CLC, DBH, HD, and HBLCLC have intermediate scores; and CH, HT, N, and CL have lower scores (Figure 8). Therefore, DINC_T, LCR, RDINC_B, RDINC_T, DINC_B, AGE, CLC, and DBH were used as the modeling features.
Figure 7. Variable importance of least absolute shrink and selection operator (LASSO).

The random forest screening results are as follows: The variables DINC T, LCR, RDINCB, RDINCT, and DINCB have higher scores; the variables AGE, CLC, DBH, HD, and HBLC have intermediate scores; and CH, HT, N, and CL have lower scores (Figure 8). Therefore, DINCT, LCR, RDINCB, RDINCT, DINCB, AGE, CLC, and DBH were used as the modeling features.

Figure 8. Variable importance of random forests.

According to the characteristics of the tree crown dataset and the parameter optimization process of the random forest algorithm, the number of iterations was set to 100, the RMSE was used as the evaluation standard, and five-fold cross-validation was used for hyperparameter optimization. Table 5 shows the setting of the hyperparameter range and the results of RandomizedSearchCV and GridSearchCV.

Table 5. Hyperparameter optimization of random forest.

| Parameter          | First | RDCV | Second | GSCV |
|--------------------|-------|------|--------|------|
| max_depth          | 5–14  | 14   | 13–14  | 14   |
| max_features       | none, auto, sqrt | auto | auto | auto |
| min_samples_leaf   | 10–20 | 10   | 9–11   | 10   |
| min_samples_split  | 2, 4  | 4    | 3, 4, 5| 4    |
| n_estimators       | 100–2000 | 800  | 700–900| 700  |

Note: RDCV, RandomizedSearchCV; GSCV, GridSearchCV; First, initial parameters, according to the result of RandomizedSearchCV to set Second parameters.

Therefore, the random forest regression model was established by setting max_depth = 5, max_features = auto, min_samples_leaf = 10, min_samples_split = 4, and n_estimators = 700.

Table 6 shows the results of CCEMs based on random forest regression model. Using $R^2$ as the evaluation indicator, the order of the accuracy of the models established by the four feature screening methods in the training set from high to low is MI, RF, RFE, and LASSO; in the test set, the order of the accuracy of the model established by the four feature screening methods from high to low is MI, RFE, RF, and LASSO. For both the training set and the test set, the random forest regression model established by MI as the feature screening method has the highest accuracy. Using RMSE as the evaluation indicator, the model established by MI as the feature screening method has a minimum value of RMSE in the training set and the test set. In summary, the random forest regression model established by MI as the feature screening method has the best effect for predicting the crown radius, which has some significance and practical value.

Table 6. Random forest regression model evaluation results.

| Selection Method | Training Set | Test Set |
|------------------|--------------|----------|
|                  | $R^2$ | MAE | MSE | RMSE | $R^2$ | MAE | MSE | RMSE |
| MI               | 0.9190 | 0.1277 | 0.0365 | 0.1910 | 0.0864 | 0.1629 | 0.0579 | 0.2406 |
| RFE              | 0.9175 | 0.1289 | 0.0372 | 0.1928 | 0.0842 | 0.1647 | 0.0590 | 0.2429 |
| LASSO            | 0.9170 | 0.1281 | 0.0374 | 0.1934 | 0.0836 | 0.1641 | 0.0593 | 0.2435 |
| Random Forest    | 0.9188 | 0.1278 | 0.0366 | 0.1913 | 0.0841 | 0.1651 | 0.0590 | 0.2429 |
Figures 9 and 10 show the residual plots of the training set and test set of the random forest regression models established by four feature screening methods. The random forest regression model established by LASSO as the feature screening method has a clear trumpet shape and shows heteroscedasticity in the training set. In the test set, the heteroscedasticity of all models is similar. For combined model evaluation accuracy and residual plot, the random forest regression model constructed based on the MI as the feature screening method performs best for predicting tree crown contour envelope of Chinese fir.

![Figure 9. Random forest regression model residual plot (training set).](image1)

![Figure 10. Random forest regression model residual plot (test set).](image2)
4. Discussion

The tree crown is important for evaluating the growth vigor of trees and the status of competition with adjacent trees. Forest stand 3D visualization is also an important part of the decision-making system for plantation growth and harvest. In the early 3D visualization of forest stands, trees were only defined using some simple geometry, such as cylinders and cones. Such an approach could not accurately capture the actual growth of trees. In the 1980s, some researchers applied the concept of fractals to the visualization of the tree crown contour. Although this method could capture the shape of the tree crown contour to a certain extent, there was no otherness in the tree crown contour based on this method, and the fractal parameters were not easy to determine. The method of taking the tree crown contour as a continuous and complete line segment and expressing it with a specific function expression has been considered. In the early stage, there were only two parameters in this equation: \( \text{DINC}_T \) (or \( \text{DINC}_B \)) and \( \text{CR} \). Therefore, the model based on this model was used to describe tree crown uniformity, but the shape of the tree crown contour was different in different growth stages. Consequently, some researchers tried to add some variables, such as \( \text{AGE} \), \( N \), \( \text{DBH} \), and \( \text{CL} \) into the equation. However, \( \text{AGE} \) is strongly correlated with the \( \text{DBH} \), \( \text{CL} \), and other variables. Therefore, adding these variables to modify the mathematical modeling can improve the accuracy of the model and better reflect the differences among trees, but the determination and modification of the model form are difficult; furthermore, the model forms of different tree species and different ages need to be considered comprehensively. Among the models mentioned in this paper, the \( \text{HT} \), \( N \), \( \text{AGE} \), \( \text{CH} \), and other variables do not show noticeable improvement in the model accuracy.

The results of random forest regression showed that the addition of multiple tree characteristic factors improved the fitting accuracy of Chinese fir crown contour envelope. In addition, the precision of random forest regression model constructed by different combinations of tree characteristics was also different. Therefore, using a single factor such as \( \text{HT} \) and \( \text{AGE} \), and composite factors such as \( \text{CR} \) and \( \text{CLC} \) to predict Chinese fir crown contour envelope could prove to be useful. In both the training set and test set, the simulation accuracy and model interpretation were higher for the random forest regression model than the mathematical regression model, and the overall effect of random forest regression model was better. The results of variable importance analysis showed that the main factors affecting the Chinese fir crown contour envelope in Chinese fir plantation were \( \text{LCR} \), \( N \), \( \text{AGE} \), \( \text{DBH} \), and \( \text{HT} \). Among these factors, \( \text{LCR} \) had the most significant effect on the Chinese fir crown contour envelope.

The CCEM based on the random forest regression method does not need to consider the correlation between variables, and the process is relatively simple. Therefore, we can select different forms of variable combinations to select the best group to build a random regression forest model. In this study, the random forest regression models constructed by four feature selection methods showed high performance; the best was the random forest model constructed by MI. The reserved features of this method were \( N \), \( \text{AGE} \), \( \text{DBH} \), \( \text{HT} \), \( \text{HBLC} \), \( \text{LCR} \), \( \text{CLC} \), \( \text{DINC}_T \), \( \text{DINC}_B \), \( \text{RDINC}_T \), and \( \text{RDINC}_B \). Among these variables, \( N \) and \( \text{AGE} \) were the initial factors, and the \( \text{DBH} \), \( \text{HT} \), \( \text{HBLC} \), and \( \text{LCR} \) had mature growth models with \( \text{AGE} \) and \( N \) and SI (site index) as variables and its distribution model; the other composite factors could be calculated from the above single factor [49–52]. Therefore, the CCEM based on the random forest has higher accuracy than the CCEMs based on mathematical modeling, and it can describe different shapes of tree crown at various stages of growth. Therefore, the random forest CCEM can accurately reflect differences in tree crown morphology among forests. Thus, the forest stand 3D model is of great significance for a 3D visualization of a plantation and for the management of plantation growth and harvest.

Figure 11 shows the crown contour envelope of a Chinese fir plantation with 5-year, 10-year, 15-year, 20-year, and 25-year standard trees. The X-axis is \( \text{DINC}_T \), and the Y-axis is \( \text{CR} \). The mathematical modeling regression model and random forest regression model selected
the best performing Model (7), i.e., the random forest regression model based on MI. The 5-year, 10-year, and 15-year prediction results show that the CCEM based on random forest regression model is superior to the mathematical modeling regression model. For 20-year-old Chinese fir, the random forest prediction result is slightly better than the mathematical modeling, whereas the prediction results of 25-year-old Chinese fir are close. In general, the random forest method has higher fitting accuracy than mathematical modeling.

Figure 11. Crown contour envelope of 5-year-old to 25-year-old Chinese fir. Error1 is the absolute error between true value and random forest; Error2 is the absolute error between true value and Model (7). The black points are crown radius of 1/10 CL, 1/4 CL, 1/2 CL, 3/4 CL, and 9/10 CL from crown top to crown bottom.

One advantage of the mathematical modeling approach is that it is highly generalized; consequently, the CCEM constructed by mathematical modeling is relatively simple. All of the trees it describes in a stand have the same crown shape, but this does not apply to the requirements of modern forestry precision management. A CCEM based on random forest can accurately reflect the differences among trees in a stand combined with existing stand distribution models, such as the HT distribution model and the DBH distribution model. Covariables such as HT and N can also be added to the mathematical model to improve the prediction accuracy. However, the form of the model is extremely difficult to determine, the fitting is more difficult, and its generalization is also reduced to some extent. For example, Chengde Wang added covariables DBH, HT, and N to Models (1) and (9) to obtain Models (12)–(14), and the results showed that adding covariables effectively improved the fitting accuracy [38]. However, in our study, improvements associated with adding covariables were small. If more covariables are added, the model form becomes more difficult to control. For example, Model (7) with three covariables had the lowest prediction accuracy among all of the CCEMs. Another advantage of random forest is that it can help to identify the tree factors most closely related to the crown in the process of feature screening, which aids the study of the crown shape. According to the four feature screening methods in this study, AGE, DBH, N, and HT significantly affect crown shape. Aiming at extended the study case problem of machine learning black box, several random sampling tests were carried out for further discussion in this paper. The splitting of the dataset was the common method; the dataset was randomly split 200 times into the model training set (70%) and
After selecting, validating, and comparing the parameters, the final parameters showed powerful stability in multiple iterations. The ultimate goal establishes that CCEMs aid plantation management. Therefore, CCEMs that are constructed based on random forest can be tailored to specific areas. As the amount of sample data increases, the prediction accuracy of the CCEM increases, because the random regression forest provides a robust method for dealing with the similar samples. For Chinese fir in other areas or other tree species, the feature combination form or hyperparameter optimization scheme used in this study may not be optimal, but if sample data are sufficient, the method described in this paper could still be used to construct CCEMs based on random forest in target areas or tree species.

5. Conclusions

Crown attributes are key components of growth and yield models because crown shape and size influence production efficiency, which is directly related to growth and mortality [53,54]. HT, DBH, CL, and other tree factors have a substantial effect on crown shape. The main objective of the CCEME based on random forest in this study was to predict the crown radius at any crown depth, which might improve the prediction accuracy. Compared with previous research on CCEM, we considered using random forest regression model and adding more basic tree factors and compound tree factor to predict the crown contour envelope of Chinese fir, which showed a significant improvement as compared with the mathematical modeling. This model has higher accuracy and more easily describes differences among trees. We can use different feature combinations to construct the random forest regression model to predict the crown contour envelope. The CCEM constructed by using MI, LASSO, RFE, and RF as feature selection methods performed well, but MI was the best approach. The newly developed CCEMs were important for predicting crown shape and also provided a new method for the study of crown shape of Chinese fir plantation. In future research, we plan to build an automated variable screening and hyperparameter optimization program to rapidly construct the CCEMs of target tree species.

Author Contributions: Conceptualization, Y.T. and X.S.; methodology, Y.T., X.S., and B.W.; validation, Y.T.; formal analysis, Y.T.; investigation, Y.T., Z.M., and Y.C.; data curation, Y.T.; writing—original draft preparation, Y.T.; writing—review and editing, Y.T., X.S., Y.Q., and B.W.; visualization, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Key National Research and Development Program of China (project no. 2017YFD0600906). The authors would also like to thank the reviewers for their comments, which were helpful in improving the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from the Fujian Agriculture and Forestry University and the Jiangle state-owned forest farm in Fujian Province and are available from the Yingze Tian with the permission of the Fujian Agriculture and Forestry University and the Jiangle state-owned forest farm in Fujian Province.

Acknowledgments: We are grateful to Deqiang Zheng’s team at the Fujian Agriculture and Forestry University for supplying valuable modeling data and the Jiangle state-owned forest farm in Fujian Province for supplying data collection guidance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Depauw, L.; Perring, M.P.; Landuyt, D.; Maes, S.L.; Blondeel, H.; De Lombarde, E.; Brümelis, G.; Brunet, J.; Clooset-Kopp, D.; Decoq, G.; et al. Evaluating structural and compositional canopy characteristics to predict the light-demand-signature of the forest understory in mixed, semi-natural temperate forests. *Appl. Veg. Sci.* 2020. [CrossRef]

2. Carlson, A.R.; Sibold, J.S.; Negron, J.F. Negron. Canopy structure and below-canopy temperatures interact to shape seedling response to disturbance in a Rocky Mountain subalpine forest. *For. Ecol. Manag.* 2020, 472, 118234. [CrossRef]
3. Klimenko, D.E.; Cherepanova, E.S.; Khomyleva, A.A. Spatial modeling of maximum capacity values of irrecoverable rainfall retention by forests in a small watershed. *Forests 2020*, 11, 641. [CrossRef]

4. Klimenko, D.; Ostakhova, A.; Tuneva, A. Experimental data on maximum rainfall retention on crowns of deciduous tree species of the middle Urals (Russia). *Forests 2019*, 10, 183. [CrossRef]

5. Li, G.; Wan, L.; Cui, M.; Wu, B.; Zhou, J. Influence of canopy interception and rainfall kinetic energy on soil erosion under forests. *Forests 2019*, 10, 509. [CrossRef]

6. Ouyang, S.; Xiao, K.; Zhao, Z.; Xiang, W.; Xu, C.; Lei, P.; Deng, X.; Li, J. Stand transpiration estimates from recalibrated parameters for the granier equation in a Chinese Fir (Cunninghahmia lanceolata) plantation in southern China. *Forests 2018*, 9, 162. [CrossRef]

7. Sinoquet, H.; Stephan, J.; Sonohat, G.; Lauri, P.E.; Monney, P. Simple equations to estimate light interception by isolated trees from canopy structure features: Assessment with three-dimensional digitized digitized apple trees. *New Phytol.* 2010, 175, 94–106. [CrossRef]

8. Jiménez-Brenes, F.M.; López-Granados, F.; De Castro, A.I.; Torres-Sánchez, J.; Serrano, N.; Peña, J.M. Quantifying pruning impacts on olive tree architecture and annual canopy growth by using UAV-based 3D modelling. *Plant Methods* 2017, 13, 55. [CrossRef]

9. Bruno, O.M.; de Oliveira Plotze, R.; Falvo, M.; de Castro, M. Fractal dimension applied to plant identification. *Inf. Sci. Int. J.* 2008, 178, 2722–2733. [CrossRef]

10. Beyer, R.; Bayer, D.; Letort, V.; Pretzsch, H. Validation of a functional-structural tree model using terrestrial Lidar data. *Ecol. Model.* 2017, 357, 55–57. [CrossRef]

11. Lu, K.; Zhang, H.; Liu, M.; Ouyang, G. Design and implementation of individual tree growth visualization system of cunninghahmia lanceolata. *For. Res.* 2012, 25, 207–211.

12. Gill, S.J.; Biging, G.S. Autoregressive moving average models of conifer crown profiles. *J. Agric. Biol. Environ. Stats* 2002, 7, 558–573. [CrossRef]

13. Gill, S.J.; Biging, G.S.; Murphy, E.C. Modeling conifer tree crown radius and estimating canopy cover. *For. Ecol. Manag.* 2000, 126, 405–416. [CrossRef]

14. Hann, D.W.; Hanus, M.L. Evaluation of nonspatial approaches and equation forms used to predict tree crown recession. *Rev. Can. Rech. For.* 2004, 34, 1993–2003. [CrossRef]

15. Crescenceto-Campo, F.; Marshall, P.; LeMay, V.; Diéguez-Aranda, U. A crown profile model for Pinus radiata D. Don in northwestern Spain. *For. Ecol. Manag.* 2009, 257, 2370–2379. [CrossRef]

16. Guo, E. *Study of the Stand Growth Model for Eucalyptus Plantation*. Diss. Fujian Agriculture and Forestry University: Fujian, China, 2009.

17. Chen, D.; Wu, B.; Han, Y.; Liu, J. Prediction model of single tree crown radius of Chinese Fir Plantation Based on modified function. *J. Northeast For. Univ.* 2015, 43, 49–53.

18. Crescenceto-Campo, F.; Álvarez-González, J.G.; Castedo-Dorado, F.; Gómez-García, E.; Diéguez-Aranda, U. Development of crown profile models for Pinus pinaster Ait. and Pinus sylvestris L. in northwestern Spain. *Forestry 2013*, 86, 481–491. [CrossRef]

19. Guo, Y.; Wu, B.; Zheng, X.; Zheng, D.; Liu, Y.; Dong, C.; Zhang, M. Simulation model of crown profile for chinese fir (Cunninghahmia lanceolata) in different age groups. *J. Beijing For. Univ.* 2015, 37, 40–47.

20. Gao, H.; Dong, L.; Li, F. Crown contour prediction model of Pinus sylvestris var. mongolica based on modified Kozak equation. *Sci. Silvae Sin.* 2019, 55, 84–94.

21. Lei, X. Application of machine learning algorithm in forest growth and harvest prediction. *J. Beijing For. Univ.* 2019, 41, 23–36.

22. Zhang, B.; Saijad, S.; Chen, K.; Zhou, L.; Zhang, Y.; Yong, K.K.; Sun, Y. Predicting tree height-diameter relationship from relative competition levels using quantile regression models for Chinese fir (Cunninghahmia lanceolata) in Fujian province, China. *Forests 2020*, 11, 183. [CrossRef]

23. Sharma, R.P.; Vacek, Z.; Vacek, S.; Kučera, M. A nonlinear mixed-effects height-to-diameter ratio model for several tree species based on czech national forest inventory data. *Forests 2019*, 10, 70. [CrossRef]

24. Pogoda, P.; Ochal, W.; Orzel, S. Performance of kernel estimator and johnson sb function for modeling diameter distribution of black alder (Alnus glutinosa (L.) Gaertn.) stands. *Forests 2020*, 11, 634. [CrossRef]

25. Zhou, R.; Wu, D.; Zhou, R.; Fang, L.; Zheng, X.; Lou, X. Estimation of DBH at forest stand level based on multi-parameters and generalized regression neural network. *Forests 2019*, 10, 778. [CrossRef]

26. Liu, M.; Niklas, K.J.; Niinemets, Ü.; Hölscher, D.; Chen, L.; Shi, P. Comparison of the scaling relationships of leaf biomass versus surface area between spring and summer for two deciduous tree species. *Forests 2020*, 11, 1010. [CrossRef]

27. Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32. Available online: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf (accessed on 31 December 2020). [CrossRef]

28. Surhone, L.M.; Tennoe, M.T.; Henssonow, S.F.; Breiman, L. Random Forest. *Mach. Learn.* 2010, 45, 5–32.

29. Peterson, K.D. Predicting the crown shape of loblolly pine trees. *Can. J. For. Res.* 2011, 27, 102–107. Available online: https://www.researchgate.net/publication/237870522_Predicting_the_crown_shape_of_loblolly_pine_trees (accessed on 31 December 2020).

30. Baldwin, V.C.; Peterson, K.D.; Iii, A.C.; Ferguson, R.B.; Strub, M.R.; Bower, D.R. The effects of spacing and thinning on stand and tree characteristics of 38-year-old loblolly pine. *For. Ecol. Manag.* 2000, 137, 91–102. [CrossRef]

31. Baldwin, V.C.; Peterson, K.D. Predicting the crown shape of loblolly pine trees. *Can. J. For. Res.* 1997, 27, 102–107. Available online: https://cdnrssciencepub.com/doi/10.1139/x96-100 (accessed on 31 December 2020). [CrossRef]
32. Chmura, D.J.; Rahman, M.S.; Tjoelker, M.G. Crown structure and biomass allocation patterns modulate aboveground productivity in young loblolly pine and slash pine. *For. Ecol. Manag.* 2007, 243, 219–230. [CrossRef]

33. Chmura, D.J. Linking Morphology and Physiology as Predictors of Productivity in Elite Families of Southern Pines. Ph.D. Thesis, TEXAS A&M UNIVERSITY, College Station, TX, USA, 2008.

34. Hann, D.W. An Adjustable Predictor of Crown Profile for Stand-Grown Douglas-Fir Trees. *For. Sci.* 1999, 45, 219–230. [CrossRef]

35. Chmura, D.J. Linking Morphology and Physiology as Predictors of Productivity in Elite Families of Southern Pines. Ph.D. Thesis, TEXAS A&M UNIVERSITY, College Station, TX, USA, 2008.

36. Hann, D.W. An Adjustable Predictor of Crown Profile for Stand-Grown Douglas-Fir Trees. *For. Sci.* 1999, 45, 217–225. [CrossRef]

37. Weiskittel, A.R.; Maguire, D.A.; Garber, S.M.; Kanaskie, A. Influence of Swiss needle cast on foliage age-class structure and vertical foliage distribution in Douglas-fir plantations in north coastal Oregon. *Rev. Can. Rech. For.* 2006, 36, 1497–1508. [CrossRef]

38. Wang, C. Study on Crown Growth Simulation and Density Control decision-making Technology of Planation. Ph.D. Thesis, Beijing Forestry University, Beijing, China, 2019.

39. Semenova, T.; Wu, S.F. least-squares method. *J. Acoust. Soc. Am.* 2005, 117, 701. [CrossRef]

40. Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Industrial Appl. Math.* 1963, 11, 431–441. [CrossRef]

41. Ladha, L.; Deepa, T. Feature selection methods and algorithms. *Int. J. Adv. Trends Comput. Sci. Eng.* 2011, 3, 1787–1797.

42. Wang, H.; Sun, H.B.; Zhang, B.M. PG-HMI: Mutual information based feature selection method. *Pattern Recognit. Artif. Intell.* 2007, 20, 55–63.

43. Louw, N.; Steel, S.J. Variable selection in kernel Fisher discriminant analysis by means of recursive feature elimination. *Comput. Stat. Data Anal.* 2005, 51, 2043–2055. [CrossRef]

44. Colombani, C.; Legarra, A.; Fritz, S.; Guillaume, F.; Croiseau, P.; Ducrocq, V.; Robert-Granié, C. Application of Bayesian least absolute shrinkage and selection operator (LASSO) and BayesCπ methods for genomic selection in French Holstein and Montbéliarde breeds. *J. Dairy Sci.* 2013, 96, 575–591. [CrossRef] [PubMed]

45. Bernard, S.; Heutte, L.; Adam, S. *Influence of Hyperparameters on Random Forest Accuracy*; International Workshop on Multiple Classifier Systems Springer: Berlin/Heidelberg, Germany, 2009.

46. Bengio, Y. Gradient-based optimization of hyperparameters. *Neural Comput.* 2000, 12, 1889–1900. [CrossRef]

47. Ou, Q.; Lei, X.; Shen, C. Individual tree diameter growth models of larch–spruce–fir mixed forests based on machine learning algorithms. *Forests* 2019, 10, 187. [CrossRef]

48. Ayuga-Téllez, E.; Mauro-Gutiérrez, F.; García-Abril, A.; González-García, C.; Martínez-Falero, J.E. Comparison of estimation methods to obtain ideal distribution of forest tree height. *Comput. Electron. Agric.* 2014, 108, 191–199. [CrossRef]

49. De Chao, T.; Li, F.; Dong, L. Prediction model of potential maximum crown width of Larix olgensis based on quantile regression. *J. Northeast For. Univ.* 2019, 47, 41–46. [CrossRef]

50. Soares, P.; Margarida, T. A tree crown ratio prediction equation for eucalypt plantations. *Ann. For. Sci.* 2001, 58, 193–202. [CrossRef]

51. Jack, S.B.; Long, J.N. Forest production and the organization of foliage within crowns and canopies. *For. Ecol. Manag.* 1992, 49, 233–245. [CrossRef]