Modeling applications:

The model was validated using data from 8006 fir plantations, which came from Qingyuan County Type II survey data. Eight site factors, which were altitude, slope gradient, slope direction, slope position, canopy closure, soil thickness, soil type, and landform were selected as the input variables. Site form was used as an output variable and the classification was carried out using the random forest model. For this model application, we divided the training set and test set in a ratio of 7:3, and the results of the model application are as follows:

(1) Using the Schumacher equation as the base equation, the ADA method was used to obtain the expression for the site form as follows:

$$SF = \frac{H_D e^{-\frac{9.0725}{15}}}{e^{-\frac{9.0725}{D}}}$$

where D denotes the mean diameter at breast height and H_D denotes the dominant height at breast diameter D;

(2) The random forest model was used to model and analyze the fir data in Qingyuan County, and the confusion matrix of the model classification results on the test set is shown in Table S1;

Table S1. Predicted results of grades on the model test set.

Actual type	Predictive type		
	Grade I	Grade II	Grade III
Grade I	29	9	0
Grade II	24	1498	202
Grade III	0	56	584

(3) The model was tested with an accuracy of 0.75 and the classification of the model is reported in Table S2;

Table S2. Performance of classification models for classification levels I, II, and III in Qingyuan County.

Actual Type	Precision	Recall	Accuracy
Grade I	0.55	0.76	0.64
Grade II	0.96	0.87	0.91
Grade III	0.74	0.91	0.82
Total	0.75	0.85	0.79

(4) A comparison of the true and predicted values of the cedar classification classes in the Qingyuan County are plotted in Figures S1 and S2, respectively. From these figures, the estimated and real values of the model show a high degree of consistency, indicating that the random forest model used in this study can effectively classify the quality of the stand in practical applications, and has a high degree of accuracy;

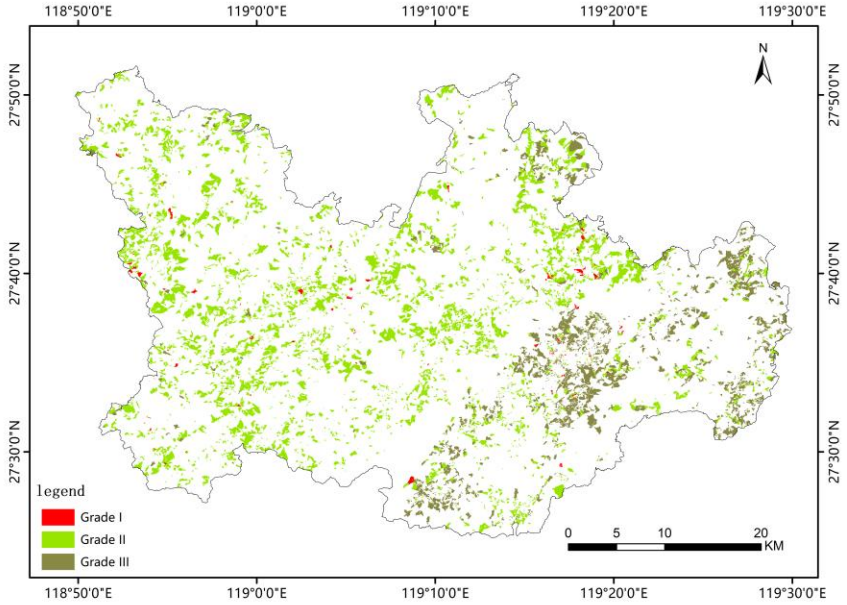

Figure S1. The map of fir grades in Qingyuan County.

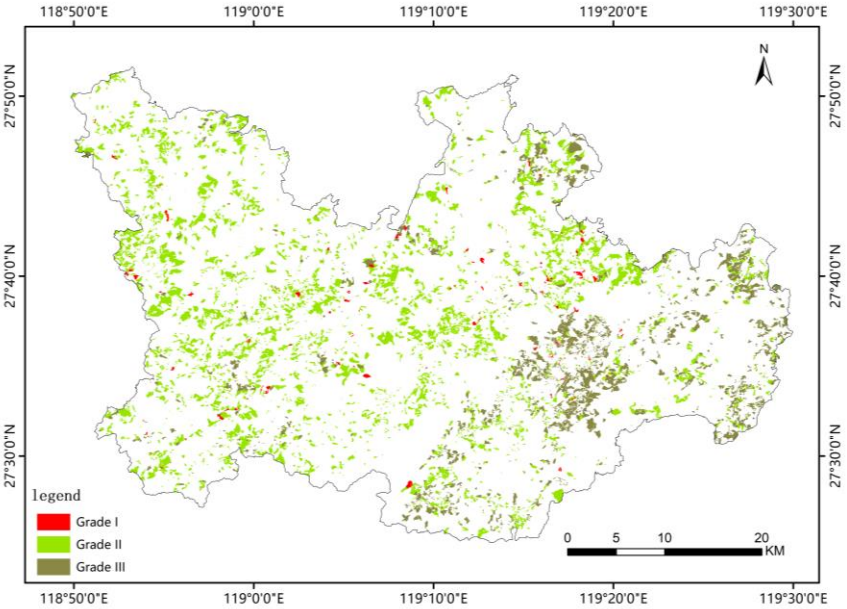

Figure S2. The predictive map of fir grade classification in Qingyuan County.

(5) The importance ranking map of the stand factors based on the random forest model is shown in Figure S3. Figure S3 shows that among the eight site factors, canopy closure and altitude had the greatest effect on fir growth, respectively, followed by soil thickness, slope position, slope direction, slope gradient, and soil taxon. The factor with the least effect on the growth of fir was landform.

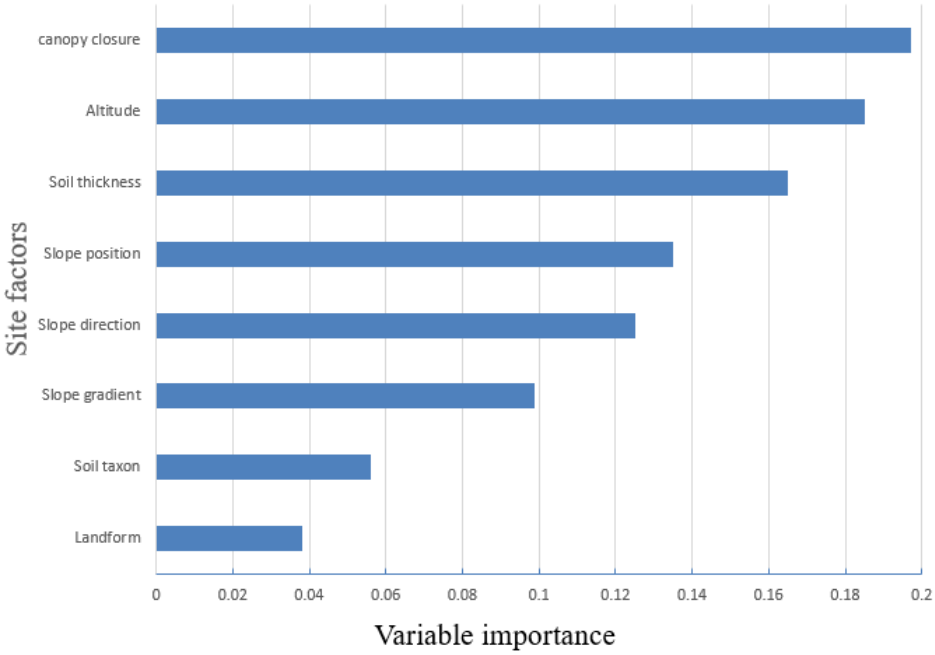

Figure S3. The relative importance of variables affecting site form in Qingyuan County.

The classification accuracy of the model reached 0.75 using Qingyuan County Class II survey data, and the model worked well, indicating that the model is universal and does not require the measurement of forest age, which improves the efficiency of data acquisition. In general, when using the model for classification, the accuracy of the Class II survey data may be lower than that of the Class I survey data, which may be due to the following reasons:

1. Feature differences: the Class II survey data may have large differences in features from the Class I data. If these differences are more difficult for the model to capture and learn and are relevant to the classification problem, it may be difficult for the model to accurately classify the Class II data;
2. Noise differences: noise or interference that may be present in the Class II survey data is greater or more difficult to deal with than noise in the Class I survey data. This will increase the difficulty in predicting and classifying the second type of survey data for the model;
3. Data quality: the quality of the Class II survey data may be relatively low, including problems such as mislabeling and missing or incomplete data. These problems can affect the training and classification results of the model.