Insights into the BRT (Boosted Regression Trees) Method in the Study of the Climate-Growth Relationship of Masson Pine in Subtropical China

Hongliang Gu 1,2, Jian Wang 1,*, Lijuan Ma 1, Zhiyuan Shang 1 and Qipeng Zhang 1

1 School of Geography, Nanjing Normal University, Nanjing 210023, China; 161301001@stu.njnu.edu.cn (H.G.); wangq306@163.com (L.M.); szy268@163.com (Z.S.); 151301001@stu.njnu.edu.cn (Q.Z.)
2 School of Resources and Environment, Anqing Normal University, Anqing 246011, China
* Correspondence: wangjian@njnu.edu.cn

Received: 8 January 2019; Accepted: 28 February 2019; Published: 5 March 2019

Abstract: Dendroclimatology and dendroecology have entered mainstream dendrochronology research in subtropical and tropical areas. Our study focused on the use of the chronology series of Masson pine (Pinus massoniana Lamb.), the most widely distributed tree species in the subtropical wet monsoon climate regions in China, to understand the tree growth response to ecological and hydroclimatic variability. The boosted regression trees (BRT) model, a nonlinear machine learning method, was used to explore the complex relationship between tree-ring growth and climate factors on a larger spatial scale. The common pattern of an asymptotic growth response to the climate indicated that the climate-growth relationship may be linear until a certain threshold. Once beyond this threshold, tree growth will be insensitive to some climate factors, after which a nonlinear relationship may occur. Spring and autumn climate factors are important controls of tree growth in most study areas. General circulation model (GCM) projections of future climates suggest that warming climates, especially temperatures in excess of those of the optimum growth threshold (as estimated by BRT), will be particularly threatening to the adaptation of Masson pine.

Keywords: Pinus massoniana Lamb.; nonlinear; boosted regression trees; tree ring; general circulation model; subtropical area

1. Introduction

Tree-ring width (TRW) exhibits a high interannual resolution and has become increasingly important for understanding complex climate systems. In the field of dendroclimatology, the development of climate-sensitive chronologies of subtropical and tropical areas remains an important issue. Therefore, studies involving detailed climate-growth analyses in warm and humid subtropical regions of China have been developed in recent decades [1–14]. Despite these achievements, additional studies on the undeniable problems in this area are still necessary. Masson pine (Pinus massoniana Lamb.), one of the most widely distributed tree species in the subtropical regions of China, has high environmental and social impacts. This species grows as well in the lowlands as it does in the highlands and occurs at altitudes ranging from a few meters to approximately 2000 m above sea level (a.s.l.). Because of this, this species occurs in highly varying climatological conditions, ranging from moist lowland river valleys to the dry mountain plateaus in the interior parts of China [15]. Thus, Masson pine is an ideal tree species for subtropical tree-climate research in China. However, using tree rings for dendroclimatology studies in tropical and subtropical areas is often complicated, and the responses of trees to the climate variables can be nonlinear and can even vary across species and ecosystems [16,17]. With respect to the Masson pine, previous ecophysiological or
climatological studies have shown that the climate-growth relationship may exhibit complex patterns. For example, Kuang et al. [1] reported that both temperature and precipitation are significantly linearly related to tree growth. Duan et al. [5] showed that only the January–April average temperature was significantly related to growth. Chen et al. [4] reported a significant negative correlation between growth and the August–October average temperature. Xia et al. [6] reported a significant negative correlation between TRW and May–August average temperature during the growing season. However, some authors reported a significant nonlinear relationship between these two factors. [18,19]. With respect to the majority of dendroclimatological studies, only important climate variables and the tree-growth relationship are well understood, and reliable climate reconstruction works can be implemented [20]. Traditionally, a common practice in dendroclimatology is to correlate TRW with monthly or seasonal climate data in a linear form [21]. However, the tree ring records are likely to be nonlinear biophysical processes and may be complicated [14,21–26]. Therefore, an increasing number of authors have acknowledged nonlinearities in the climate-growth system [17,27–31]. One of the current methods used for calibrating the nonlinear climate-growth relationship is the Vaganov–Shashkin (VS) model [23,32–35]. However, this performs worst in warm and wet environments [22]. Thus, given that our study occurred in a subtropical area, our focus turned to using machine learning methods. Artificial neural network (ANN) was considered as a better choice for careful assessment of complex climate-growth relationships [20,36–50]. However, the learning method process of an ANN is a “black box operation” [51–53], meaning that it is sensitive to overfitting and ANNs have difficulty evaluating the contribution of each variable to the results, from a statistical point of view [53,54]. An ensemble machine learning method (i.e., a boosted regression trees (BRT), developed by Friedman [35], appears to be a promising method for detecting nonlinear relationships, as demonstrated by some recent studies that aimed to model tree growth responses to climate [51,56–58]. To the best of our knowledge, this technique has not been applied in dendroclimatic studies with respect to complex interactions in dendroclimatic relationships in subtropical China. Thus, this paper attempts to examine the application of the BRT for (1) Selecting the relatively important climate predictors for TRW, (2) investigating the complex climate-growth relationship, and (3) evaluating the potential thresholds concerning the effects of climate variables and estimate the influence of the warming climate on the growth of Masson pine.

2. Materials and Methods

2.1. Geography and Climatology of the Study Areas

Tree-ring records from individual sites can also reflect the influence of unobserved localized and nonclimatic influences [21,59]. To systematically explore the relationship between Masson pine and climate on a large spatial scale, we collected tree ring data from five areas, that were subjected to little human disturbance (Table 1).

Fanjing Mountain (FJS), located in the western portion of the Masson pine-dominated area, contains vegetation associated with typical subtropical evergreen broad-leaved forest zonal primitive forest vegetation. The annual average temperature is approximately 17.4 °C. The lowest temperature is approximately 6.1 °C in January, and the highest temperature is approximately 27.8 °C in July. The annual average precipitation is approximately 1129.5 mm, and the average relative humidity is approximately 74%. Three sampling sites were selected here: Tuanlong, Jinchang and Zhaibao.

The southern sampling area was Jiulian Mountain (JLA), known as the “gene bank of biological resources”, located at the border of the Jiangxi province and Guangdong province. Masson pine trees in this area composes secondary forest. The annual average temperature is approximately 16.4 °C, the lowest temperature in January is approximately 6.8 °C, and the highest temperature is in July, is approximately 24.4 °C. The average annual precipitation is approximately 2155.6 mm, and the average annual relative humidity is approximately 85%. The sampling sites here include Jiantou, Banlingxia and Luxia.
### Table 1. Study area information.

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Sampling Site</th>
<th>Latitude (° N)</th>
<th>Longitude (° E)</th>
<th>Elevation (m a.s.l.)</th>
<th>Soils</th>
<th>Cores/Trees</th>
<th>Slopes (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FJS (W)</td>
<td>Tuanlong</td>
<td>27.92</td>
<td>108.57</td>
<td>1120</td>
<td>Yellow soil</td>
<td>33/16</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Jinchang</td>
<td>28.00</td>
<td>108.72</td>
<td>1053</td>
<td>Yellow soil</td>
<td>21/10</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Zhaihao</td>
<td>27.78</td>
<td>108.72</td>
<td>808</td>
<td>Yellow soil</td>
<td>33/14</td>
<td>35</td>
</tr>
<tr>
<td>JLA (S)</td>
<td>Jianyou</td>
<td>24.63</td>
<td>114.60</td>
<td>446</td>
<td>Oxisol</td>
<td>34/23</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Banlingxia</td>
<td>24.20</td>
<td>114.25</td>
<td>590</td>
<td>Oxisol</td>
<td>14/6</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Luxia</td>
<td>23.90</td>
<td>114.33</td>
<td>419</td>
<td>Yellow soil</td>
<td>20/8</td>
<td>21</td>
</tr>
<tr>
<td>JLN (C)</td>
<td>Wulingyan</td>
<td>29.12</td>
<td>115.28</td>
<td>1023</td>
<td>Yellow soil</td>
<td>32/15</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Wumeishan</td>
<td>28.80</td>
<td>114.72</td>
<td>685</td>
<td>Yellow soil</td>
<td>35/18</td>
<td>35</td>
</tr>
<tr>
<td>TBS (N)</td>
<td>Tongbaishan</td>
<td>32.4</td>
<td>113.27</td>
<td>509</td>
<td>Yellow soil</td>
<td>46/21</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Jigongshan</td>
<td>31.82</td>
<td>114.05</td>
<td>264</td>
<td>Brown soil</td>
<td>40/21</td>
<td>22</td>
</tr>
<tr>
<td>WYS (E)</td>
<td>Dabugang</td>
<td>27.17</td>
<td>117.35</td>
<td>475</td>
<td>Yellow soil</td>
<td>53/24</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Wuyishan</td>
<td>27.77</td>
<td>118.02</td>
<td>276</td>
<td>Latesol</td>
<td>40/20</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Xiaowuyi</td>
<td>27.75</td>
<td>115.03</td>
<td>273</td>
<td>Yellow soil</td>
<td>20/11</td>
<td>24</td>
</tr>
</tbody>
</table>

The five capital letters (W, S, C, N and E) denote the western, southern, central, northern and eastern margins of Masson pine distribution, respectively.

Near the center of the Masson pine distribution, we chose to sample on Jiuling Mountain (JLN), which is the site of the original Masson pine forest. The annual average temperature is approximately 16.7 °C, the lowest temperature in January is approximately 4.4 °C, and the highest temperature in July is approximately 28.2 °C. The annual average precipitation is approximately 1653 mm, and the average relative humidity is approximately 79%. The rainfall is abundant during every season each year, and the climate is warm. The two sampling sites in this area are Wulingyan and Wumeishan.

The fourth sampling area, located at the northern margin of the Masson pine distribution, is Tongbai Mountain (TBS), which is dominated by its primary forest. The highest altitude of the Tongbai Mountain is 1140 m and the climate of the north subtropical region is considered monsoon is humid. The average annual temperature is approximately 15 °C, the average temperature in January is approximately 2 °C, and the average temperature in July is approximately 28 °C. The climate is mild and sunshine is abundant. The northern region lacks an alpine barrier, and the cold wave from the north can continue into winter. The winter temperature is low, the annual precipitation is approximately 1168 mm, and the annual average relative humidity is approximately 74%. Tongbaishan and Jigongshan were the two sampling sites.

We also developed TRW chronologies on Wuyi Mountain (WYS), which is covered with both natural and artificial secondary forests. The annual average temperature is approximately 18 °C, the lowest temperature in January is approximately 7.1 °C, and the highest temperature in July is approximately 27.8 °C. The average annual precipitation is approximately 1818.6 mm, and the average relative humidity is approximately 81%. The sample area here mainly includes mainly three sampling sites: Xiaowuyi, Wuyishan Park and Dabugang.

#### 2.2. Tree-Ring Preparation and Measurement

The sampling sites were selected on the south slope, far away from human residents. The average slope was approximately 20 to 35 degrees. There were 2 to 3 sampling subsites in each area. There was no evidence indicating that natural conditions such as soil and water supply had experienced extreme disturbance at each sampling site. According to the sampling standards for the International Tree-Ring Data Bank (ITRDB), a tree-ring core with a diameter of 5.15 mm of Masson pine was drilled on both sides of the tree stem, which were parallel to the direction of the slope. In total, 207 old and dominant tapped trees were selected to maximize the temporal extent of the TRW records. Tree ring cores were preserved in a paper tube and then returned to the laboratory. After drying in a natural state, the tree-ring cores were fixed and sanded with increasingly finer sandpaper, up to 400 grit, to produce visible rings [60]. The skeleton plot was first used to validate the dating, after which the pseudo and
missing tree-rings were initially identified [21]. The preliminary calendar age for each tree ring was determined. Afterward, the ring width was measured using an AcuRite (VoorTech, New York, NY, USA) measuring system, with an precision of 0.001 mm. Finally, the COFECHA program was used to check and control the quality of cross-dating, and samples with week correlation, large singularities and juvenile effects were rejected [61,62]. The ARSTAN program was used to establish chronologies [63]. After cross-dating, the chronologies were subjected to detrending and other statistical procedures were used routinely on the tree-rings [21]. We applied the cubic spline curve (SP), which eliminates low-frequency variations and enhances high-frequency climate signals [64]. The resulting population chronology is often described with various statistical measures to quantify the signal to noise ratio (SNR) [65], the mean sensitivity (MSX) [66,67], standard deviation (SD), the commonly used expressed population signal (EPS) [68], with a threshold of 0.85 [69], explained variance in first eigenvector (PC1) and the Gini coefficient [70,71]. In addition, principal component analysis was used to estimate the common similar and different regional environments that influenced the growth process of the trees. These statistical analyses (Table 2) were performed via R (version 3.5.1), using the dplR package [72–74].

### Table 2. Characteristics of the tree ring chronologies of five study areas.

<table>
<thead>
<tr>
<th></th>
<th>FJS (W)</th>
<th>JLA (S)</th>
<th>JLN (C)</th>
<th>TBS (N)</th>
<th>WYS (E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean width (mm)</td>
<td>2.391</td>
<td>3.193</td>
<td>2.190</td>
<td>2.334</td>
<td>2.759</td>
</tr>
<tr>
<td>Series intercorrelation</td>
<td>0.58</td>
<td>0.53</td>
<td>0.59</td>
<td>0.574</td>
<td>0.537</td>
</tr>
<tr>
<td>Mean Gini coefficient</td>
<td>0.271</td>
<td>0.331</td>
<td>0.278</td>
<td>0.239</td>
<td>0.287</td>
</tr>
<tr>
<td>Mean sensitivity</td>
<td>0.244</td>
<td>0.208</td>
<td>0.234</td>
<td>0.256</td>
<td>0.252</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.307</td>
<td>0.300</td>
<td>0.374</td>
<td>0.259</td>
<td>0.295</td>
</tr>
<tr>
<td>Signal–noise ratio</td>
<td>19.45</td>
<td>13.49</td>
<td>19.05</td>
<td>31.81</td>
<td>14.78</td>
</tr>
<tr>
<td>Expressed population signal</td>
<td>0.951</td>
<td>0.931</td>
<td>0.95</td>
<td>0.97</td>
<td>0.936</td>
</tr>
<tr>
<td>Explained variance in first eigenvector</td>
<td>35.3</td>
<td>24</td>
<td>30.3</td>
<td>37.2</td>
<td>25.2</td>
</tr>
<tr>
<td>1st order autocorrelation</td>
<td>Raw width</td>
<td>f = 0.5, nys = 30</td>
<td>0.587</td>
<td>0.59</td>
<td>0.689</td>
</tr>
<tr>
<td></td>
<td>0.332</td>
<td>0.306</td>
<td>0.316</td>
<td>0.188</td>
<td>0.219</td>
</tr>
</tbody>
</table>

Fanjing Mountain (FJS), Jiulian Mountain (JLA), Jiuling Mountain (JLN), Tongbai Mountain (TBS), Wuyi Mountain (WYS). “f” represents wavelength cutoff. “nys” represents a number giving the rigidity of the smoothing spline.

2.3. Climate Data

Climate data were used for detecting dendroclimatology from the CRUTS 4.01 datasets (https://crudata.uea.ac.uk/cru/data/hrg/) [75]. Meteorological station records were obtained from the National Meteorological Information Center (http://data.cma.cn/). The locations of the five homogenized meteorological stations are as follows: the Sinanxian (SNX) station is close to FJS, the Lianpingxian (LPX) station is near JLA, the Xiushuixian (XSX) station is a short distance away from JLN, the closest station to TBS is in Xinyangshi (XYS), and the Shaowushi (SWS) station is near WYS. Basic climate information is shown in Figures 1 and 2. In Figure 2A,B, both precipitation and temperature display a single peak that essentially belongs to the climate type of the same period of rain and heat. The peak precipitation occurs each year in June, while the peak surface temperature occurs in July; spring, autumn and winter are relatively dry. The precipitation data from the meteorological stations indicate that the annual precipitation in the TBS sampling area is the lowest, and that in the WYS sampling area is the highest. The temperature records from the meteorological stations indicate that the temperature of the TBS sampling area is relatively low, and that of the JLA sampling area is relatively high. Based on the CRUTS 4.01 datasets and instrumental data, the annual temperature and precipitation gradually decreases from southeast to northwest (Figure 2C,D) in our analysis area, and their difference are quite obvious. When viewed alongside the latitude, the precipitation along the western margin of the Masson pine distribution changes little, but the gradient is obvious along the eastern margin. Therefore, the hydroclimatic environment of Masson pine differs in the subtropical wet monsoon climate, where heat is abundant and precipitation is plentiful.
Forests 2018, 9, x FOR PEER REVIEW 5 of 20

Figure 1. The study area and meteorology station. Five study areas include FJS (Fanjing Mountain), JLA (Jiulian Mountain), JLN (Jiuling Mountain), TBS (Tongbai Mountain), and WYS (Wuyi Mountain).

Figure 2. Climate information on the Masson pine distribution area. (A) The monthly mean precipitation total from five meteorology stations (“SNX”, “LPX”, “XSX”, “XYS” and “SWS” represent the Sinanxian, Lianpingxian, Xiushuixian, Xinyangshi, and Shaowushi station, respectively) nearby the five study areas. (B) The monthly mean temperature from five meteorology sites nearby the five study areas. (C) The annual mean precipitation total and (D) the annual mean temperature from the CRUTS 4.01 grid database.

2.4. BRT Methods

The generalized boosting model (GBM) has the ability to characterize nonlinear changes. Elith et al. [76] produced a detailed paper concerning the applications of a boosted regression trees (BRT) method under the (GBM) framework in relation to ecological data. The strengths of the GBM model (including BRT) were threefold: (1) It facilitated the processing of multiple interactive variables; (2) characterized partial dependence, i.e., potential nonlinear and nonmonotonic changes in response functions; and (3) had a high predictive accuracy and abundant flexibility. BRT models are robust in
relation to small-size or absent data [77]. In this paper, BRT models were implemented using the GBM package (version 2.1.4) [76,78].

The first consideration of BRT models is distribution. For continuous predictors, the Gaussian distribution is selected (to minimize the squared error), which provides a better correlation between the observed value and the GBM model estimate [56,78]. We took the train function in the caret package to explore the optimal values for the foremost parameters [79]. The function tries a number of different parameters, compares the error rates, and then suggests the smallest parameter that generates an appreciable decrease in the error rate. The first two parameters generated were shrinkage (learning rates) and the number of trees (nt). A good rule of thumb for an ideal shrinkage value is 0.001, which is close to the point of diminishing return. Shrinkage is also related to the number of trees. The optimal number of trees (minimizing the prediction bias) was estimated using 10-fold cross-validation [51,76]. To increase the model accuracy, randomness is included using a bagging fraction of 0.5 [76].

The expected sensitivity of the tree growth to the climate variables is quantified based on two factors: the relative influence of the variables and the partial dependence plots [56]. With respect to the BRT models, an extension of a variable’s “relative influence” (here, the relative influence refers to the contribution of each variable to the minimization of the loss function) was developed by Friedman [55]. The relative importance is calculated by averaging the number of times a variable is selected for splitting, weighted by the squared improvement to the BRT model, as a result of each split, it is then scaled so the values sum to 100. Variables whose relative influence exceeds the median of each model are classified as “highly important”, and those below the median are classified as having “low importance” [56]. The partial dependence plots indicate the relationship of the predictor variable “X” with the response variable “Y”, as well as the former’s dependence on the latter. In current applications, a change in the quantitative tree growth index of partial dependence plots is quantified as a function of the season climate signal.

In southeastern China, the growth season of coniferous trees generally occurs from early April to November [80]. Cai and Liu [81] selected the March–October season length for *Pinus taiwanensis* Hayata. Zhang et al. [82] accurately quantified the length of the growing season of Masson pine. These authors have shown that December is the beginning of the dormancy period of Masson pine. The cambium is inactive in January and is likely to be semidormant. February is the end of the dormancy period of the Masson pine or the beginning of its germination, and buds appear in March (i.e., the beginning of rapid growth). Hardly any growth occurs in the summer, after which rapid growth begins again. Therefore, in terms of tree physiology relevant to studying the dendroclimatic relationships, the Masson pine growth season selected in this study is listed in Table 3. To avoid overfitting, we used the “vif” function in the car package (version 3.0-2) [79], to exclude any collinearity variables (vif >5). We found that the selected climate variables were not collinear.

<table>
<thead>
<tr>
<th>Climate Variable</th>
<th>FJS</th>
<th>JLA</th>
<th>JLN</th>
<th>TBS</th>
<th>WYS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior June–August precipitation</td>
<td>10.921</td>
<td>11.041</td>
<td>7.137</td>
<td>5.381</td>
<td>8.108</td>
</tr>
<tr>
<td>Prior September–November precipitation</td>
<td>4.688</td>
<td>7.607</td>
<td>10.479</td>
<td>5.470</td>
<td>16.034</td>
</tr>
<tr>
<td>February–May precipitation</td>
<td>7.755</td>
<td>2.958</td>
<td>8.976</td>
<td>3.393</td>
<td>8.842</td>
</tr>
<tr>
<td>June–August precipitation</td>
<td>6.313</td>
<td>3.643</td>
<td>7.845</td>
<td>7.418</td>
<td>2.994</td>
</tr>
<tr>
<td>September–November precipitation</td>
<td>4.713</td>
<td>13.950</td>
<td>9.777</td>
<td>25.311</td>
<td>5.056</td>
</tr>
<tr>
<td>Prior June–August temperature</td>
<td>9.278</td>
<td>6.228</td>
<td>7.386</td>
<td>25.629</td>
<td>8.171</td>
</tr>
<tr>
<td>Prior September–November temperature</td>
<td>15.057</td>
<td>6.400</td>
<td>7.736</td>
<td>25.311</td>
<td>11.972</td>
</tr>
<tr>
<td>September–November temperature</td>
<td>6.061</td>
<td>11.574</td>
<td>4.571</td>
<td>4.585</td>
<td>20.539</td>
</tr>
</tbody>
</table>

Fanjing Mountain (FJS), Julian Mountain (JLA), Jiuling Mountain (JLN), Tongbai Mountain (TBS), Wuyi Mountain (WYS). The four most important variables in each area are highlighted in bold. “Prior” represents the previous year.
3. Results

3.1. Tree Ring Chronology Statistics

By the use of COFECHA quality test, which included several trials of different frequency responses and wavelength cutoff years, the standard and residual chronologies of the TRW in five study areas were ultimately established with an SP (a 50% frequency response and wavelength cutoff at 30 years) (Figure 3). The chronology statistics of each tree ring chronology series indicate that the chronology can be used for subsequent analyses (Table 2). The high mean sensitivity obtained (0.208–0.256) is greater than 0.2. The mean sensitivity is approximately 0.2, which is generally accepted as a series that is sensitive enough for climate reconstruction [83]. The first autocorrelation value for the raw width series of this study ranged from 0.587 to 0.689. The mean Gini coefficient shows that it satisfies three hypotheses related to variability in the ring-width increment: (1) it would decrease as the latitude increased; (2) it would increase with the longitude (i.e., west to east); and (3) it would decrease as elevation increased [70]. EPS values ranged from 0.931 to 0.970, indicating that each chronology was dominated by a coherent stand-level signal. The series intercorrelation, SNR, PC1, and EPS of the same sample area were all relatively high, which indicated that the variation in the ring width of each sequence at the same sampling site was synchronized, indicating that they might be affected by consistent changes in the external climate and environment.

![Figure 3](image-url)

**Figure 3.** The standard chronologies (STD) and residual chronologies (RES) detrended with cubic spline methods. “FJS”, “JLA”, “JLN”, “TBS” and “WYS” represent Fanjing Mountain, Jiulian Mountain, Jiuling Mountain, Tongbai Mountain, and Wuyi Mountain, respectively.

As shown in Figure 4, the cumulative proportion of the first three principal components is as high as 83.1%, the value of PC1 is 40.01%, and the other three principal components are also higher than 10%. The loadings of JLA, WYS and JLN are relatively large in the first principal component, while they are relatively small for TBS and FJS. On the other hand, the loadings of FJS, TBS and JLN are relatively large in the second principal component. These results show that the regional climate conditions affecting the growth of Masson pine may be relatively consistent, and that niche factors have an impact on tree growth.
With respect to the regression of continuous data, the mean absolute error (MAE) and the root mean squared error (RMSE) are easy to understand and are widely used to measure the performance of machine learning problems. If these values are small, then the models can make accurate predictions when faced with “new” data [85]. The R-squared values ($R^2$) for the BRT model ultimately applied across all five study areas are 0.491 (FJS), 0.559 (JLA), 0.608 (JLN), 0.597 (TBS), and 0.569 (WYS). The BRT model performs better for TBS and JLN than for the other areas. In addition, we compared it with other models, including the generalized linear model (GLM), random forests (RF), and bayesian regularization neural network (BRNN) models (Figure 5). The RMSE and MAE values among the four models are relatively small in the five study areas and give slightly more accurate results. However, the nonlinear models (BRT, BRNN and RF) slightly outperform the linear model. These results show that the BRT model can be a useful tool for dendroclimatic studies.

### 3.2. Performance of the BRT Models

To evaluate the performance of the BRT model, resampling and cross-validation were used [84]. The results of principal component analysis (PCA) with chronologies from five study areas (Fanjing Mountain (FJS), Jiulian Mountain (JLA), Jiuling Mountain (JLN), Tongbai Mountain (TBS), Wuyi Mountain (WYS)). Total variance was explained by all five principal components. “PC1” denotes first principal component, and so on.

#### Performance of different models across the five study areas

<table>
<thead>
<tr>
<th>Model</th>
<th>Proportion of Variance</th>
<th>Standard Deviation</th>
<th>Cumulative Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>0.401</td>
<td>1.415</td>
<td>0.402</td>
</tr>
<tr>
<td>PC2</td>
<td>0.226</td>
<td>1.063</td>
<td>0.627</td>
</tr>
<tr>
<td>PC3</td>
<td>0.204</td>
<td>1.006</td>
<td>0.831</td>
</tr>
<tr>
<td>PC4</td>
<td>0.146</td>
<td>0.852</td>
<td>0.970</td>
</tr>
<tr>
<td>PC5</td>
<td>0.023</td>
<td>0.345</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Figure 4.** The results of principal component analysis (PCA) with chronologies from five study areas (Fanjing Mountain (FJS), Jiulian Mountain (JLA), Jiuling Mountain (JLN), Tongbai Mountain (TBS), Wuyi Mountain (WYS)). Total variance was explained by all five principal components. “PC1” denotes first principal component, and so on.

**Figure 5.** Performance of different models across the five study areas (Fanjing Mountain (FJS), Jiulian Mountain (JLA), Jiuling Mountain (JLN), Tongbai Mountain (TBS), Wuyi Mountain (WYS)). Four models include generalized linear model (GLM), random forests (RF), bayesian regularization neural network (BRNN), and boosted regression trees (BRT). “MAE” and “RMSE” represent mean absolute error and represents root mean squared error, respectively.
3.3. Relative Influence of Seasonal Climate Factors

One major advantage of BRT models is that they can select variables of relative importance. By ranking climate predictors based on how much they influence TRW, the values of relative influence from the five study areas can be determined (Table 3 and Figure 6). Some aspects of the growth response to climate were relatively consistent across all five study areas. The September–November precipitation and February–May temperature have a comparatively large influence on tree growth in the five study areas, and spring temperatures tended to be more influential than summer temperatures. Despite these similarities, some differences occurred between the study areas. For example, winter temperatures were generally unimportant, except at JLA, where the December–January temperature has a relatively strong effect on growth. The June–August temperature was relatively important at JLA and JLN, but less important at FJS, TBS and WYS.

![Figure 6. Median importance values of each climate variable (see Table 3). Upper and lower limit of boxes represent the third quartile (Q3) and first quartile (Q1). Black dots represents the outliers.](image)

3.4. Nonlinearity of the Data

A frequency distribution, which is a statistical tool, was generated to visualize the distribution of one seasonal climate variable (Figure 7). The interannual variations of climatic conditions in the five study areas are quite different, and extreme climate phenomena occurred in some years. For example, at WYS, for the prior September–November precipitation variable, the chart shows a skewed frequency distribution, with variation ranging from 35 to 150 mm; moreover, the frequency in some years is high, whereas in others is low, but there is a high probability of a higher frequency or concentration of 60 mm. Combining partial dependence plots and climate annual dynamics will help elucidate these complex climate–growth relationships. We found that with changes in climatic conditions, the sensitivity of growth to climatic conditions will increase or decrease rapidly. Afterward, the partial dependence curve eventually flattens. When the partial dependence curve was associated with fluctuation, there was a general asymptotic growth response model, i.e., tree growth is relatively insensitive to a wide range of climatic conditions [56]. Therefore, the asymptotic growth response to climate is highly common in the study areas.
The upper-right plots of Figure 7 show that the February–May temperature is somewhat directly related to the tree growth at FJS. The interannual variations of the February–May temperature range from 10 to 13.3 °C and center on 11–12 °C. The partial dependence plots of FJS on February–May temperature show that, at 11 °C, the value is 0.92, after which there is an upward trend followed by the peak (0.95). Tree growth was relatively insensitive to the colder years, but during the warmer years, the sensitivity of tree ring growth to the temperature of the time period increased rapidly. Note that February is the beginning of the germination of Masson pine. At this time, low temperatures may slow the germination process and delay growth.

Residual plots are a useful graphical tool for identifying nonlinearity. The presence of a pattern may indicate a problem with some aspects of a linear model. Figure 8 displays a residual plot from the linear regression. The residuals exhibit obvious trends in some study areas, which may indicate nonlinearity of the data. For example, an upward growth trend in the bottom-right panel of Figure 8 shows that the relationship between tree growth and the September–November precipitation of the previous year may be nonlinear at WYS.
3.5. Threshold of Climate-Growth in the Future

To assess the potential impact of future climate change on tree growth, we estimated the optimal seasonal temperature or precipitation range of seasons for tree growth. These optimal thresholds were defined within the partial dependency plots derived by the BRT model [56]. These optimal growth ranges are then compared with the mean distribution of the climatic values of the seasons in the CMIP5 datasets for the period of 2006–2100 (https://esgf-node.llnl.gov/projects/cmip5/). A comparison of the CMIP3 and CMIP5 datasets gave a slightly more accurate result for simulating future climate change in China [86]. Through comparison with the CRUTS 4.01 datasets, we can quantify the sensitivity of tree growth to future climate change and determine what variables will lead to significant changes in Masson pine, i.e., relative optimal growth hydroclimate conditions (Figures 9 and 10). Understanding how climate change will affect the cambial growth of Masson pine in the future is of practical significance.

The average temperature in the next 100 years will be higher than that of the historical datasets (Figure 9). The box chart shows that the median values of the seasonal temperature of the CMIP5 datasets are higher than the optimum growth temperature, as estimated by the BRT models. These temperatures correspond to the relatively cool temperature values to be experienced for the next 100 years at most sites and suggest that the optimum conditions for growth here have been relatively rare. It can be seen that the median temperature of each season during the 1901–2016 period is essentially no more than that of the optimum growth temperature, but in some years, the temperature exceeds the optimum temperature threshold. These years, which exceed the optimum growth temperature, may strongly influence on the growth of Masson pine. In some study areas, there are abnormal temperature occurrences in some years, i.e., the abnormally low June–August temperature at JLA and the abnormally high February–May temperature at TBS.

Generally, the optimum precipitation range suitable for Masson pine growth varied substantially among the study areas (Figure 10). The median values of seasonal precipitation during the next 100 years were essentially equal to those within the CRUTS 4.01 datasets. Even some seasonal precipitation is lower than the historical precipitation data. For example, at JLA, the future median February–May precipitation is 101 mm, but the historical median value is 194 mm. The interannual June–August precipitation (regardless of the current year or prior year) varies greatly in the five study areas. Moreover, the interannual February–May precipitation variations are relatively large. In different areas, the threshold of precipitation, suitable for tree growth in different seasons, is very different, and the relative influence of different climatic variables varies from place to place. These differences are not explained by the location of the study area or its topography, and the reasons for this nonuniformity deserve further study.
Forests were defined within the partial dependency plots derived by the BRT model [56]. These optimal seasonal temperature or precipitation ranges for tree growth. These optimal thresholds in the future are of practical significance. (Figures 9 and 10). Understanding how climate change will affect the cambial growth of Masson pine lead to significant changes in Masson pine, i.e., relative optimal growth hydroclimate conditions future climate change in China [86]. Through comparison with the CRUTS 4.01 datasets, we can comparison of the CMIP3 and CMIP5 datasets gave a slightly more accurate result for simulating the CMIP5 datasets for the period of 2006–2100 (https://esgf-node.llnl.gov/projects/cmip5/). A different areas, the threshold of precipitation, suitable for tree growth in different seasons, is very different, and the relative influence of different climatic variables varies from place to place. These, among the study areas (Figure 10). The median values of seasonal precipitation during generally, the optimum precipitation range suitable for Masson pine growth varied substantially among the study areas (Figure 10). The median values of seasonal precipitation during the 1901–2016 period is 101 mm, but the historical median value is 194 mm. The interannual precipitation is lower than the historical precipitation data. For example, at JLA, the future median the next 100 years were essentially equal to those within the CRUTS 4.01 datasets. Even some seasonal temperatures correspond to the relatively cool temperature values to be experienced for the next 100 years, i.e., the abnormally low June–August temperature at TBS. It can be seen that the median temperature of each season during the 1901–2016 period is essentially no more than that of the optimum growth temperature, but in some years, the temperature exceeds the optimum temperature threshold. These years, which exceed the optimum growth essentially no more than that of the optimum growth temperature, as estimated by the BRT models. These datasets are higher than the optimum growth temperature, as estimated by the BRT models. These year, the temperature distribution of the total precipitation from 2006 to 2100, based on the CMIP5 datasets (https://esgf-node.llnl.gov/projects/cmip5/).

Figure 9. The box-and-whisker plots for the optimum temperature range for Masson pine growth. The box-and-whisker plots indicate the median (bold lines) and the 25th and 75th percentiles (lower and upper edges of the box). Light grey boxes denote the optimum range from the BRT models. The observed distribution of average temperature values for each month and season from 1901 to 2016 is expressed by the steel blue box-and-whisker plots. The orange box-and-whisker plots indicate the projected distribution of temperature values from 2006 to 2100, based on CMIP5 datasets (https://esgf-node.llnl.gov/projects/cmip5/).

Figure 10. The box-and-whisker plots concerning the optimum precipitation range for Masson pine growth. Light grey boxes denote the optimum range derived from the BRT models. The observed distribution of the total precipitation for each month and season from 1901 to 2016, is expressed by the steel blue box-and-whisker plots. The orange box-and-whisker plots indicate the projected distribution of total precipitation from 2006 to 2100, based on the CMIP5 datasets (https://esgf-node.llnl.gov/projects/cmip5/).
4. Discussion

4.1. BRT as a Potential Tool for Analysing Climate-Growth Relationships

Linear regression (correlation analysis, multiple regression, principal component regression, etc.) is commonly used to analyze the statistical relationship between the TRW (as a proxy) and climate data [21,87]. Pearson’s product-moment correlation has become a workhorse for analyzing complex dendroclimatic relationships. However, Pearson’s coefficient is very sensitive to outliers; even a single outlier can change the direction of the coefficient. Additionally, this technique cannot capture the nonlinear relationship between two variables. Because the BRT model is insensitive to outliers and multicollinearity, it can fit complex nonlinear relationships and can automatically handle the interactive effects between predictors [76,77,88]. We preliminarily used this technique to determine the most important variables, evaluate potential thresholds regarding the effects of these climate variables, and examine their functional responses. Based on two evaluation metrics (RMSE and MAE), our analyses confirmed the results of a previous study showing that the BRT, BRNN and RF models performed better than the generalized linear models (GLMs) did [45,89,90], with less variance and bias. Moreover, the BRT model performed better for TBS and JLN than for the other areas. Considering the ecological and climate factors, both TBS and JLN are dominated by primeval Masson pine forest, and the highest temperature in July is higher than 28 °C. On the other hand, in conjunction with WYS, a lower altitude may be another factor that affects the performance of BRT methods. However, this result deserves further study. To our knowledge, some studies have shown that there are nonlinear climate-growth relationships in subtropical areas of China [18,19,43,91]. Thus, it is necessary to further explore this issue.

The frequency distribution of the historical climate data shows that some outliers occur during some seasons of different years. This phenomenon indicates that extreme climate records are also frequent in subtropical regions [92], which would cause a divergent problem and a nonlinear response pattern [43,91,93].

In addition, the residual plots exhibit obvious trends in some study areas, which may provide an indication of nonlinearity of the data [94]. We assessed the partial dependence plots and found that the asymptotic growth response to climate was relatively common in the Masson pine growth models. We speculated that this pattern of an asymptotic climate-growth relationship may exhibit nonlinear patterns [56,95]. Until a certain threshold, the climate-growth relationship may be linear. Once beyond this threshold, tree growth will be insensitive to some climate factors, after which a nonlinear relationship may occur. Some authors have stated that the growth of Pinus taiwanesis exhibits some nonlinear characteristics associated with temperature and precipitation, and the linear relationship only holds true within a certain threshold of subtropical China [91]. The widespread occurrence of growth thresholds indicates that small climate changes may have great impacts on the growth of Masson pine. In fact, tree growth is a continuously dynamic process. While tree-ring datasets provide insight into a limited interval of climate conditions, the long-term impact of tree growth on the ontogenetics of Masson pine remains unknown. The assumption that the dendroclimatic relationship is constant across the whole spectrum of climate-growth responses in linear regression is not always correct [22,46]. The response of tree growth to climate is a complex process and there may be interactive and compensatory effects between climate variables in different seasons and years, which makes for an unstable relationship between tree growth and climate [37,91].

Last, factor–discrete can be used to explain the nonlinearities in physiological responses to climate [56,95]. Discrete climate factors are revealed in cases where the response threshold is relatively consistent among study areas. In our studies, thresholds in tree responses to the February to May temperature were reached from 11 to 18 °C in the majority of our study areas. The temperature thresholds of Masson pine in the early part of the growing season are similar to those in some previous studies [96,97]. Therefore, this phenomenon probably represents a real physiological threshold in the growth response to temperature.
The enhanced regression model (BRT) allows us to clearly identify the threshold of the climate-growth response and develop a statistical model of that response, and this model has a greater explanatory power than does the linear statistical model. The universality of nonlinear responses to climate is important for understanding how trees respond to climate variability in subtropical China, and our ability to identify and describe these threshold responses is crucial for developing accurate predictions of the impact of climate warming on subtropical forests.

4.2. Seasonal Features in Growth-Climate Relationships

Our relative influence analysis using the BRT technique showed that the response patterns of Masson pine to climate change exhibit both similarities and differences. The beginning of each phenology of Masson pine coincides with the month of the temperature variable that affects the growth of the species, as revealed by the BRT model. Principal component analysis reveals that the regional environment affecting the growth of Masson pine was relatively consistent, but the niche factors impact the growth of trees.

Notably, with the exception of wet areas with relatively low elevation (JLA and WYS), the February–May temperature significantly impacts Masson pine growth. In addition, the growth of Masson pine in wet areas at high elevation was affected by the temperature in February [98]. In warm-temperate forests, stem radial growth at the beginning of growing seasons is generally limited by temperature [99,100]. Our similar results at JLN confirm previous findings in which, compared with other seasons, the January–April temperature has a greater impact on the growth of Masson pine in Hunan and Jiangxi provinces [5]. Other dendroclimatological studies in southeastern China have confirmed that temperatures during the spring and early summer strongly influence coniferous forests [101,102]. With respect to northern region of the Masson pine distribution zone, growth is related to the squares of the temperature during February, May and August [18,19]. In fact, in our northern study area (TBS), the February–May temperature had an important influence on the growth of Masson pine. The increase of the temperature in the spring is conducive to the division and growth of early wood cells, and early wood is the main component of radial tree growth.

Under normal climatic conditions, growth at an altitude of 300 m was affected by the mean monthly precipitation during the previous October [98]. In Fujian Province, the autumn and winter temperature and precipitation are highly important for width of tree-rings [4,8,103,104], at WYS (mean altitude of 275 m a.s.l.), that the September–November temperature and previous September–November precipitation yield the greatest relative importance value. In the southern distribution region of Masson pine (JLA), the tree growth is highly sensitive to the winter temperature and autumn precipitation. These results are consistent with those of a previous study that showed that the winter temperature negatively affects on tree growth [103]. Compared with the findings in a previous study in the tropical area [17], our findings showed that the September–November precipitation is a growth-limiting factor at JLA. In addition, the September–November precipitation of the current years is important for tree growth at TBS and JLN. In the Lushan Mountains, near to our study area (JLN), the precipitation during the dry season (July–November) greatly promoted stem growth [105]. In the northern region of the distribution zone of Masson pine, the August–November precipitation positively affected on tree growth [18].

Consequently, the relatively important climate factors do vary across the whole distribution range of Masson pine, but the spring and autumn climate conditions mainly affect tree growth. These results may explain the bimodal radial growth pattern of Masson pine stems, as two growth peaks during the transitional seasons (spring–early summer and autumn) and a decreased growth rate in the summer were observed [105]. Zhang et al. [82] also mentioned that the growth of Masson pine during dry seasons is obviously stronger than during the wet seasons.
4.3. Influence of Global Warming

We briefly considered the possible impacts of future climate change on the growth of Masson pine tree-rings based on general circulation model (GCM) projections. In the context of future climate change, subtropical monsoon climate will be characterized by milder winters and hotter summers. When the CMIP5 datasets were compared, the optimum growth temperature occurred within a relatively low range. These findings suggest that warming during the next 100 years is likely to lead to significant changes in Masson pine. Temperatures exceeding the thermal optimum for growth will reduce net photosynthesis due to increased photorespiration [106,107]. In addition, based on the CMPI5 datasets, future precipitation during the season from September–November will dip below the optimum precipitation threshold. In the future, the growth of Masson pine may not benefit from future global warming and will exhibit increased sensitivity to conditions of the dry season (September–November).

Trees growing at temperatures higher than those for the optimum growth conditions tend to distribute less biomass to their roots [93]. This phenomenon could lead to a decrease in or a movement of Masson pine. Some authors have confirmed that the declining trend and the distribution area of Masson pine will likely gradually shift northward with the warming climate [82,103]. Thus, the hydrological combination during the dry season will play an important role in the normal growth of Masson pine indicating that the growth of this species will be affected not only by environmental factors (temperature and precipitation) but also by net photosynthetic energy.

5. Conclusions

In this paper, climate-growth correlation analyses were carried out on a large spatial scale, based on the tree-ring chronologies of Masson pine. Compared with the results of previous studies, our results showed that, by the use of a nonlinear machine learning algorithm (a BRT model), spring and autumn conditions were relatively highly influential on Masson pine. GCM projections of the future climate suggest that the warming climate, especially temperatures in excess of those of the optimum growth threshold (as estimated by BRT), will be particularly threatening to the adaptation of Masson pine. The relationship between the Masson pine TRW and climate is unstable when the habitat exceeds the optimal threshold value and may be nonlinear. The BRT method can be a variable selection tool for identifying the important linear and/or nonlinear independent climate variables, which can then be used in more classic predictive models.

Author Contributions: Conceptualization, J.W.; Data curation, L.M.; Investigation, Z.S. and Q.Z.; Writing—original draft, H.G.; Writing—review and editing, H.G.

Funding: This work was supported by the National Natural Science Foundation of China (grant No. 41271204), and the Social Science Foundation of Anhui, China (grant No. SK2016A0544).

Acknowledgments: We are grateful for the helpful and constructive comments of three anonymous reviewers. We also thank the assistant editor (MS Winter Wang) for her help.

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

References


34. Speed, J.D.M.; Austrheim, G.; Hester, A.J.; Mysterud, A. Browsing Interacts with Climate to Determine Tree-Ring Increment: Browsing Interacts with Climate. Funct. Ecol. 2011, 25, 1018–1023. [CrossRef]
38. Jevšenak, J.; Džeroski, S.; Levañič, T. Predicting the Vessel Lumen Area Tree-Ring Parameter of Quercus Robur with Linear and Nonlinear Machine Learning Algorithms. Geochronometria 2018, 45, 211–222. [CrossRef]
40. Zhang, Q.-B.; Hebda, R.J.; Zhang, Q.-J.; Alfaro, R.I. Modeling Tree-Ring Growth Responses to Climatic Variables Using Artificial Neural Networks. For. Sci. 2000, 46, 229–239. [CrossRef]
41. Woodhouse, C.A. Artificial Neural Networks and Dendroclimatic Reconstructions: An Example from the Front Range, Colorado, USA. Holocene 1999, 9, 521–529. [CrossRef]


70. Bunn, A.G. Statistical and Visual Crossdating in R Using the DplR Library. *Dendrochronologia* 2010, 28, 251–258. [CrossRef]


92. Duan, J.; Zhang, Q.; Lv, L.-X. Increased Variability in Cold-Season Temperature since the 1930s in Subtropical China. *J. Clim.* 2013, 26, 4749–4757. [CrossRef] [PubMed]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).