

基于广义代数差分法和因子选择的杉木人工林立地质量评价

龚宇浩¹, 孙益群², 董晨^{1,3}, 胡彦蓉^{1,3}, 高威芳^{1,3}

(1. 浙江农林大学 数学与计算机科学学院, 浙江 杭州 311300; 2. 浙江省淳安县林业局, 浙江 淳安 311700;
3. 浙江农林大学 林业感知技术与智能装备国家林业和草原局重点实验室, 浙江 杭州 311300)

摘要: 【目的】研究浙西南杉木 *Cunninghamia lanceolata* 人工林立地质量, 探究影响人工林立地质量的主要因素, 为杉木人工林经营提供理论依据。【方法】以浙江省庆元县杉木人工林为研究对象, 采用解析木数据, 利用广义代数差分法, 构建了杉木人工林地位指数模型, 以此为因变量, 使用随机森林算法从二类杉木小班数据中筛选出影响杉木林地位指数的主导因子, 在此基础上, 通过地位指数得分值确定立地质量等级, 对庆元县杉木小班进行立地分类与立地质量评价。【结果】以 Richards 为基础构建的广义代数差分方程最符合庆元县杉木人工林的生长过程, 其决定系数 (R^2) 为 0.969, 均方根误差 (E_{RMS}) 为 1.778, 平均绝对误差 (E_{MA}) 为 0.400, 且模型的多形曲线符合杉木林生长的生理学意义。通过随机森林的因子选择得出: 影响杉木人工林地位指数的主要立地因子有地貌、腐殖质厚度、土层厚度, 与传统的偏相关系数法相比, 基于随机森林的因子选择具有更大的贡献度和更合理的因子非线性关系解释性。结合小班数据, 以地貌、腐殖质厚度、土层厚度等 3 个立地因子将小班划分为 17 个立地类型, 通过计算得分可知研究区立地质量处于中等及以上。【结论】基于广义代数差分法的地位指数模型适用于浙西南杉木人工林, 与基于置换的因子选择方法筛选出的主导因子相结合能够对立地质量进行有效评价。图 4 表 4 参 27

关键词: 杉木人工林; 广义代数差分法; 立地因子选择; 随机森林; 立地类型划分

中图分类号: S757 文献标志码: A 文章编号: 2095-0756(2023)06-1282-10

Site quality evaluation of *Cunninghamia lanceolata* plantation based on generalized algebraic difference approach and factor selection

GONG Yuhao¹, SUN Yiqun², DONG Chen^{1,3}, HU Yanrong^{1,3}, GAO Weifang^{1,3}

(1. College of Mathematics and Computer Science, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China;
2. Chun'an County Forestry Bureau, Chun'an 311700, Zhejiang, China; 3. Key Laboratory of National Forestry and Grass Administration on Forestry Sensing Technology and Intelligent Equipment, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China)

Abstract: [Objective] With an investigation of the site stand of *Cunninghamia lanceolata* plantation in southwest Zhejiang Province, this study is aimed to explore the main factors affecting the site quality of the plantation so as to provide a theoretical basis for *C. lanceolata* plantation management. [Method] First, with *C. lanceolata* plantation in Qingyuan District of Zhejiang Province taken as the research object, the parse tree data and generalized algebraic difference approach was employed to establish the site index (SI) model of *C. lanceolata* plantation. Then, with SI chosen as the dependent variable, random forest algorithm was used to screen out the dominant factors affecting the forest position index from the data of Class II *C. lanceolata* sub class data, on the basis of which, the site quality level was determined by the score value of SI, and the site

收稿日期: 2022-11-23; 修回日期: 2023-05-06

基金项目: 浙江省自然科学基金资助项目 (LQ21C160004)

作者简介: 龚宇浩 (ORCID: 0009-0008-7429-8053), 从事人工林立地质量评价。E-mail: 1037313668@qq.com。通信作者: 董晨 (ORCID: 0009-0005-0835-7965), 讲师, 博士, 从事林分生长收获模拟研究。E-mail: dongchen@zafu.edu.cn

classification and site quality evaluation were carried out for *C. lanceolata* sub classes in Qingyuan District. [Result] The generalized algebraic difference approach based on Richards equation was the most consistent with the growth process of *C. lanceolata* in Qingyuan District with the coefficient of determination (R^2), root mean squared error (E_{RMS}) and mean absolute deviation (E_{MA}) being 0.969, 1.778 and 0.400 respectively and the multishape curve of the model was in accordance with the physiological significance. The main site factors affecting the SI of *C. lanceolata* plantation are landform, humus thickness and soil layer thickness, as was shown in the factor selection of random forest, which had a greater contribution and more reasonable interpretation of the nonlinear relationship between factors when compared with the traditional partial correlation coefficient method. With landform, humus thickness and soil layer thickness selected as the site factors, the sub class was divided into 17 site types, among which the sub classes under investigation were of medium site quality or above. [Conclusion] The site index based on generalized algebraic difference method is suitable for *C. lanceolata* plantation in southwest Zhejiang Province, and when combined with the dominant factors selected by permutation importance method, it can effectively evaluate the site quality. [Ch, 4 fig. 4 tab. 27 ref.]

Key words: *Cunninghamia lanceolata* plantation; generalized algebraic difference approach; site factors selection; random forest; site classification

立地质量是影响森林生产力的量化指标，是林地生产力评估和森林林分经营管理等林业工作和研究的基础^[1]。科学合理地开展林地立地质量评价，对研究森林生长收获规律，实现森林可持续发展具有重要的指导意义。

在人工林立地质量评价研究中，地位指数法是使用最广泛的方法^[2]。地位指数即林木基准年龄时的优势木平均高度。较早构建的地位指数方法是 BAILEY 等^[3]基于差分方程法 (algebraic difference approach, ADA) 提出的，即通过对方程进行参数削元，建立反映 2 组不同年龄和优势木高的差分方程，从而对基准年龄的优势木高进行预估。此后，不同学者使用 ADA 构建了各种人工林树种的地位指数模型^[4-6]。ADA 构建的模型是固定基准年龄的静态方程，依赖于预先选定的基准年龄^[7]，且在构建过程中只能设置 1 个自由参数，地位指数曲线无法同时满足多形性。对此，CIESZEWSKI 等^[8]提出了广义代数差分法 (generalized algebraic difference approach, GADA)，在推导时可设置树高最大值参数与形状参数，并能够构建与基准年龄无关、具有多条渐近线的多形曲线簇，合理地改进了 ADA 依赖于基准年龄、曲线簇不能满足多形性的不足。大量研究表明：GADA 构建的地位指数模型优于基于固定基准年龄的模型，且 GADA 模型在林木生理生长特性方面具有更合理的解释性^[9-12]。

地位指数模型需要测定林分的优势木高和年龄，然而，在森林资源调查体系中，往往难以确定林分优势木，对于宜林地，更无树高和年龄数据可寻。鉴于此，不少学者借助立地因子来实现立地质量评价^[13-14]。这类模型的关键是选择合理的立地因子，而以往研究中，选取的立地因子不仅数量多且具有较大的主观性，不能很好地解释立地与林分生长的非线性相关性。机器学习随机森林算法能够有效处理因变量和自变量之间的非线性、交互作用等问题，同时能够拥有因子选择功能，被成功应用于林业领域分类和预测问题的研究^[15]，而对于因子选择的后续研究和应用，却鲜有报道。

杉木 *Cunninghamia lanceolata* 是浙江省西南部地区的主要人工用材树种，具有较大的经济价值和生态效益。本研究以浙江省庆元县杉木人工林为研究对象，采用广义代数差分法预测杉木林的地位指数，并以此为因变量，采用随机森林构建因子特征选择模型，筛选出主导因子划分立地类型，以期为提高杉木林生产潜力提供理论指导。

1 研究地区与方法

1.1 研究区概况

研究区位于浙江省丽水市庆元县 (27°25'~27°51'N, 118°50'~119°30'E)。2018 年庆元县森林覆盖率高

达 86.06%，森林蓄积量达 1321.6 万 m^3 。该地区大部分为山区，海拔为 253~1 758 m，地势由东北向西南倾斜。该地区属亚热带季风气候，年均气温为 17.4 $^{\circ}C$ ，年均降水量为 1 760.0 mm，无霜期为 245.0 d。土壤类型以红壤、黄壤为主，有少量潮土和粗骨土。庆元县森林植被种类丰富，由巾子峰、白坎、百丈 3 个独立区块组成了大型森林公园，主要针叶树种为福建柏 *Fokienia hodginsii*、南方红豆杉 *Taxus mairei*、杉木等。

1.2 数据来源

研究数据包括庆元县周边的杉木标准地解析木数据和庆元县杉木二类调查小班数据。解析木数据主要用于地位指数模型的构建。2015 年充分考虑了现有林分的立地质量、年龄、密度等因素，设置了 18 块 20 m \times 20 m 具有代表性的林分标准地。每个标准地挑选 1 株未受损伤的、未受压的、健康的优势木作为研究对象，对其伐倒并截取圆盘进行树干解析。将单株解析木中数据进行排列，共得 117 个树高和年龄数据，图 1 是通过解析木得到的树高和年龄分布图。

解析木具有代表性，通过解析木构建的地位指数能在一定程度上反映整体小班立地水平。为进一步探索地位指数与立地因子之间的关系，从庆元县 2019 年度二类调查数据中获取杉木小班资料。对原始数据进行数据完整性检查，以 3 倍标准差为标准剔除各项异常数据，并删除数据集的缺失值，剔除变量为 0 的小班数据，得到杉木小班数据 4 920 条。杉木小班的立地因子包括海拔 (263~1 652 m)，地貌 (低山、中山)，坡度 (缓坡、斜坡、陡坡、急坡)，坡向 (阴坡、半阴坡、半阳坡、阳坡)，坡位 (上坡、中坡、下坡)，土壤类型 (红壤、黄壤)，土层厚度 (1~115 cm)，腐殖质厚度 (厚、中、薄)。

将地位指数模型应用于二类调查小班的地位指数预测时，需要林分优势木高数据，而二类调查数据中不具备林分优势木高调查数据，因此使用标准地调查数据构建平均木和优势木高模型，计算公式为 $D=1.054H+2.174$ ，决定系数 (R^2) 为 0.94，其中 D 为林分优势木高， H 为林分平均高，从而得出优势木数据用于小班地位指数的预测。

1.3 基于广义代数差分法的地位指数模型构建

1.3.1 基础模型选择 在地位指数模型构建中，需要选择用于建立广义差分方程的基础理论方程。在基础理论方程的选择中，SCOLFORO 等^[16]选择了 Korf、修正 Weibull 等多个理论方程用于巴西桉树 *Eucalyptus robusta* 的地位指数研究。SEKI 等^[17]运用 Richards 和 Korf 方程实现了土耳其克里米亚松 *Pinus nigra* 地位指数的模型研究。综合已有研究成果，本研究采用修正 Weibull、Korf、Richards 方程作为地位指数构建基础模型。

$$\text{修正 Weibull 方程: } h = a(1 - e^{-bt^c}); \quad (1)$$

$$\text{Korf 方程: } h = ae^{-bt^c}; \quad (2)$$

$$\text{Richards 方程: } h = a(1 - e^{-bt})^c. \quad (3)$$

式 (1)~(3) 中： h 为优势木高； t 为林分年龄； a 、 b 、 c 为模型参数。

1.3.2 广义代数差分法的推导过程 对于地位指数而言，需要具备以下性质：①多形性；②基准年龄无关性；③具有多条水平渐近线。但使用代数差分法推导的差分地位指数模型时，无法同时满足水平渐近线与多形性的要求。而广义代数差分法允许在推导时设置多个自由参数，能够满足多条水平渐近线和多形性 2 个性质。

采用广义代数差分法推导步骤为：①选择生长方程为基准方程，并根据理论或经验指定其中的参数为自由参数。②提出一个与立地质量相关的量 X_0 ，并假设自由参数与 X_0 之间具有函数关系。③将上述函

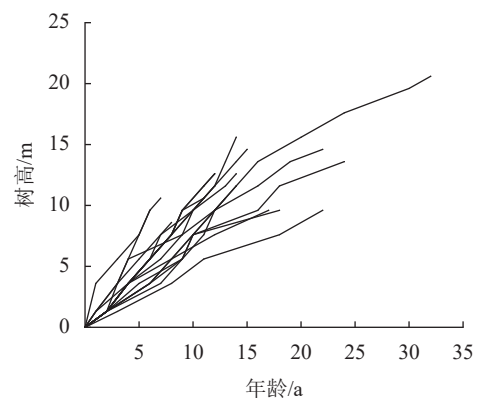


图 1 解析木数据统计

Figure 1 Parse tree data statistics

数关系带入基准方程，从中解出 X_0 的表达式。对于无法求得显式解的 X_0 ，利用迭代法进行求解。④将地位指数和基准年龄分别带入表达式中的树高和年龄，并将表达式带入基础方程的 X_0 ，推导出差分地位指数方程。

以修正 Weibull 为例，参数 a 影响曲线在时间轴上的位置， c 表示曲线的形状参数，影响曲线的弯曲程度，因此选择参数 a 、 c 为自由参数。并假设 a 与 X_0 存在指数关系， c 与 X_0 之间存在线性关系， c_1 、 c_2 为线性关系的参数，将基础方程转换为：

$$h = e^{X_0} \left(1 - e^{-bt^{c_1+c_2X_0}} \right) \tag{4}$$

从中解出 X_0 的表达式为：

$$X_0 = \ln h_1 - \ln \left(1 - e^{-bt_1^{c_1+c_2X_0}} \right) \tag{5}$$

式 (5) 中： t_1 为指定年龄， h_1 为指定年龄下的优势木高，当 t_1 为基准年龄时， h_1 为地位指数。令式 (4) 中， $t = t_2$ ， $h = h_2$ ，其中 t_2 为预测年龄， h_2 为预测年龄下的优势木高，将式 (5) 代入式 (4)，得到差分方程模型，表达式为：

$$h_2 = e^{X_0} \left(1 - e^{-bt_2^{c_1+c_2X_0}} \right) \tag{6}$$

在 3 个基础方程中， b 代表生长率参数，是林木固有属性，与立地关系不紧密，因此将 a 、 c 指定为自由参数。在理论方程基础上，使用广义代数差分法推导出 6 个差分方程，如表 1 所示。

表 1 基础方程与差分地位指数方程

Table 1 Different site index equations and base equations

基础方程	编号	自由参数	X 的初始解或方程	差分方程
修正Weibull方程	M1	$a = e^{X_0}$ $c = c_1 + c_2X_0$	$X_0 = \ln h_1 - \ln \left(1 - e^{-bt_1^{c_1+c_2X_0}} \right)$	$h_2 = e^{X_0} \left(1 - e^{-bt_2^{c_1+c_2X_0}} \right)$
	M2	$a = e^{X_0}$ $c = c_1 + c_2/X_0$	$X_0 = \ln h_1 - \ln \left(1 - e^{-bt_1^{c_1+c_2/X_0}} \right)$	$h_2 = e^{X_0} \left(1 - e^{-bt_2^{c_1+c_2/X_0}} \right)$
Korf方程	M3	$a = e^{X_0}$ $c = c_1 + c_2X_0$	$X_0 = \ln h_1 + bt^{-(c_1+c_2X_0)}$	$h_2 = e^{X_0} e^{-bt_2^{-(c_1+c_2X_0)}}$
	M4	$a = e^{X_0}$ $c = c_1 + c_2/X_0$	$X_0 = \ln h_1 + bt^{-(c_1+c_2/X_0)}$	$h_2 = e^{X_0} e^{-bt_2^{-(c_1+c_2/X_0)}}$
Richards方程	M5	$a = e^{X_0}$ $c = c_1 + c_2X_0$	$X_0 = (\ln h_1 - c_1 F_0) / (1 + c_2 F_0)$ $F_0 = \ln \left(1 - e^{-bt_1} \right)$	$h_2 = e^{X_0} \left(1 - e^{-bt_2} \right)^{c_1+c_2X_0}$
	M6	$a = e^{X_0}$ $c = c_1 + c_2/X_0$	$X_0 = \frac{1}{2} \left[F_0 + \sqrt{F_0^2 - 4c_2 \ln(1 - e^{-bt_1})} \right]$ $F_0 = \ln h_1 - c_1 \ln \left(1 - e^{-bt_1} \right)$	$h_2 = e^{X_0} \left(1 - e^{-bt_2} \right)^{c_1+c_2/X_0}$

说明： a 、 b 、 c 为模型参数； c_1 、 c_2 为线性关系的参数； X_0 表示与立地质量相关的量； t_1 为指定年龄， h_1 为指定年龄下的优势木高； t_2 为预测年龄， h_2 为预测年龄下的优势木高。

1.4 立地因子特征选择

立地质量与立地因子存在密切的关系^[18-20]，从立地因子中提取出主导因子，能够高效地评价林地立地质量。林分的立地质量与立地因子间通常是复杂的非线性关系，传统方法简化了假设条件，难以达到理想的效果。因此，在立地因子的选择上，要选择影响地位指数，又易于测量和获取的数据，同时能保证一定精度的因子。已有研究表明：采用随机森林提取的特征因子能够在不需要先验知识的情况下准确表达因子与林木生长的非线性关联关系^[15]。因此，本研究使用构建好的地位指数模型来计算每个杉木小班的地位指数，使用随机森林算法从立地因子中筛选出影响地位指数的主导因子，并与传统的偏相关分析结果进行比较。

1.4.1 基于随机森林的因子特征选择 随机森林是基于决策树的机器学习算法^[21]。随机森林能够有效地处理非线性、交互作用、共线性等问题，可以用于回归预测，还可以进行因子特征选择。它主要提供了2种因子特征重要性度量方法，分别是基于内置和置换的因子特征选择方法。

①基于内置的因子特征选择。该方法通过计算随机森林中所有回归树的节点平均值减少程度来确定特征的重要性，其使用基尼指数(B)来测量。在本研究中，用 V 表示基于内置的特征重要性方法。立地因子 j 在构成随机森林的回归树子节点 m 上的重要性，即节点 m 分枝前后的基尼指数变化量(V_{jm})为：

$$V_{jm} = B_m - B_l - B_r. \quad (7)$$

式(7)中： B_l 和 B_r 表示分枝后2个新节点 l 、 r 的基尼指数，如果立地因子 j 在第 i 棵回归树中出现的节点在集合 M 中，立地因子 j 在第 i 棵回归树上的重要性(V_{ij})为：

$$V_{ij} = \sum_{m \in M} V_{jm}. \quad (8)$$

因此对拥有 n 棵回归树的随机森林而言，立地因子 j 的重要性评分(V_j)为：

$$V_j = \sum_{i=1}^n V_{ij}. \quad (9)$$

②基于置换的因子特征选择。该方法随机对每个特征进行排序，并计算模型性能的变化，对于特征重要性的评判取决于该特征被随机重排后，模型表现评分的下降程度。过程为：输入训练后的模型 E 和样本集 D ，对于样本集 D 的每个立地因子 j ，在 T 次重复实验中的每次迭代次数 r ，随机重排列立地因子 j ，构造一个随机重排后的数据集 $D_{c_{t,j}}$ ，然后计算模型 E 在数据集 $D_{c_{t,j}}$ 上的参考分数 S_{tj} ，特征 j 的重要性(P_j)计算如下：

$$P_j = s - 1/T \sum_{r=1}^T S_{tj}. \quad (10)$$

式(10)中： s 表示模型的参考分数； T 表示实验次数； r 表示迭代次数； S_{tj} 表示选取第 t 个样本第 j 列计算后的参考分数。基于随机森林的因子特征选择分别采用Python中的feature importance和permutation importance进行计算。

1.4.2 偏相关系数的因子特征选择 当2个变量都与第3个变量相关时，将第3个变量的影响剔除，采用偏相关系数分析另外2个变量之间的相关程度。偏相关系数绝对值越大(越接近1)，表明变量之间的线性相关程度越高，反之越低。本研究运用SPSS 22计算各个立地因子与优势木高的偏相关系数。

1.5 模型拟合与检验指标

使用R语言完成地位指数模型参数的估计和统计检验。在地位指数建模中，使用scorecard包，将数据集分为建模样本与检验样本，其中建模样本占总数据的70%。通过决定系数(R^2)评估地位指数模型拟合优度。

采用均方根误差、平均绝对误差、残差对模型进行检验，均方根误差用于反映地位指数模型拟合的误差，平均绝对误差用于反映模型对独立样本的预测能力，残差反映实际优势木高值与估计值之间的差，具体计算公式参考文献[22]。

2 结果与分析

2.1 地位指数模型拟合结果与检验

如表2所示：通过广义代数差分法构建的差分方程中，除M6外，其余拟合结果均较好。以修正Weibull作为基础方程的广义代数差分法的差分方程中，M1、M2的决定系数均达0.99，且拟合系数均通过显著性检验，其中M2比M1拟合效果更好。因此M2可选作为下一步使用的地位指数模型；以Korf方程为生长方程的广义代数差分法拟合中，M3、M4方程的拟合效果较好，但拟合系数并未通过显著性检验，因此不考虑将M3、M4作为下一步进行比较的方程。以Richards方程为基础的广义代数差分

法拟合中，M5 拟合结果优于 M6，且 M5 的平均绝对误差明显小于 M6，说明 M5 的预测能力更强。因此，综合各项指标，将 M2、M5 作为较优模型，进行进一步检验。

表 2 杉木人工林差分地位指数模型参数估计及拟合情况

Table 2 Parameter estimates and goodness of fit of the difference site index equations for *C. lanceolata*

编号	参数	估计值	标准误	显著性检验	R ²	均方根误差	平均绝对误差
M1	<i>b</i>	0.410	0.123	0.002**	0.998	0.343	0.259
	<i>c</i> ₁	1.471	0.162	<0.001***			
	<i>c</i> ₂	-0.395	0.052	<0.001***			
M2	<i>b</i>	0.403	0.104	<0.001***	0.999	0.261	0.257
	<i>c</i> ₁	-0.580	0.199	0.006**			
	<i>c</i> ₂	2.671	0.401	<0.001***			
M3	<i>b</i>	1.583	0.568	0.008**	0.982	1.336	0.409
	<i>c</i> ₁	2.803	6.226	0.999			
	<i>c</i> ₂	2.799	4.038	0.999			
M4	<i>b</i>	1.074	0.408	0.013*	0.999	0.101	0.232
	<i>c</i> ₁	-0.156	0.248	0.533			
	<i>c</i> ₂	1.881	0.556	0.001**			
M5	<i>b</i>	0.053	0.014	<0.001***	0.969	1.778	0.400
	<i>c</i> ₁	-0.174	0.047	<0.001***			
	<i>c</i> ₂	0.222	0.049	<0.001***			
M6	<i>b</i>	0.072	0.026	0.009**	0.606	6.371	0.631
	<i>c</i> ₁	1.865	1.025	0.078			
	<i>c</i> ₂	-1.763	2.806	0.534			

说明：***表示 $P < 0.001$ ，**表示 $P < 0.01$ ，*表示 $P < 0.05$ 。

从图 2 可知：模型 M2、M5 的残差分布均较为均匀，大部分集中分布在 X 轴附近。在杉木优势木高较低时，模型 M2、M5 的拟合残差绝对值较大，为 0.5~1.5，而随着优势木高的增大，残差分布的范围呈趋势性减少，因此，M2、M5 可用于进行地位指数模型构建。在残差分析上，M2、M5 之间精度差距很小，通过绘制以 20 a 为基准年龄的地位指数曲线簇对方程进一步分析，设定级距为 2 m，地位指数为 10~26 m (图 3)。模型 M2 具有多条水平渐近线，但图像为非 S 型的单型生长曲线，在低龄高速生长后，林木生长趋于平缓，近成熟时，生长速率趋近 0，这与杉木达到成熟的实际不一致。模型 M5 的地位指数曲线簇具备多形性，满足地位指数越高，低龄优势木生长越快的实际情况，并呈现林木年龄越趋近 20 a，曲线越趋近于平缓的情况，是理想的地位指数模型。在模型检验上，M2、M5 之间的差异性不大，而在预测能力上，M5 更符合实际的状况。因此选用 M5 作为最优的地位指数模型，进行下一步研究。将预估参数代入模型 M5 的 X_0 中，并将所得代入差分方程，其表达式为：

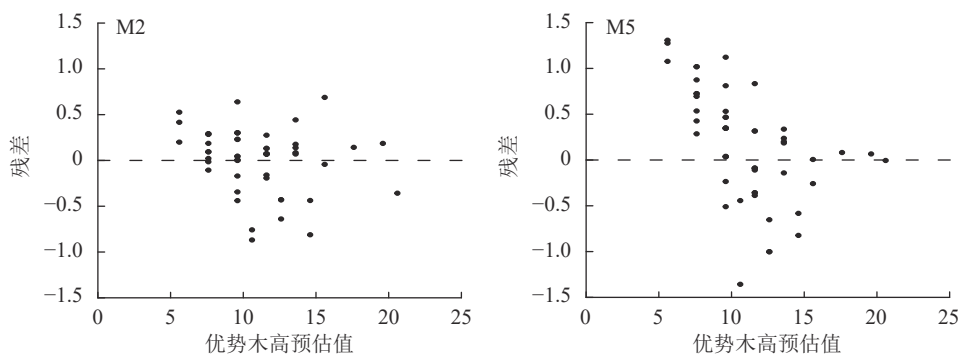


图 2 模型 M2 和 M5 地位指数残差图

Figure 2 Residuals against predicted dominant tree height for model M2 and M5

$$h_2 = \exp(X_0)[1 - \exp(-0.053t_2)]^{-0.174+0.222X_0};$$

$$X_0 = \{\ln h_1 + 0.174 \ln[1 - \exp(-0.053t_1)]\} / \{1 + 0.222 \ln[1 - \exp(-0.053t_1)]\}.$$

其中, t_1 为指定年龄, h_1 为指定年龄下的优势木高, 当 t_1 为基准年龄时, h_1 为地位指数; t_2 为预测年龄, h_2 为预测年龄下的树高。

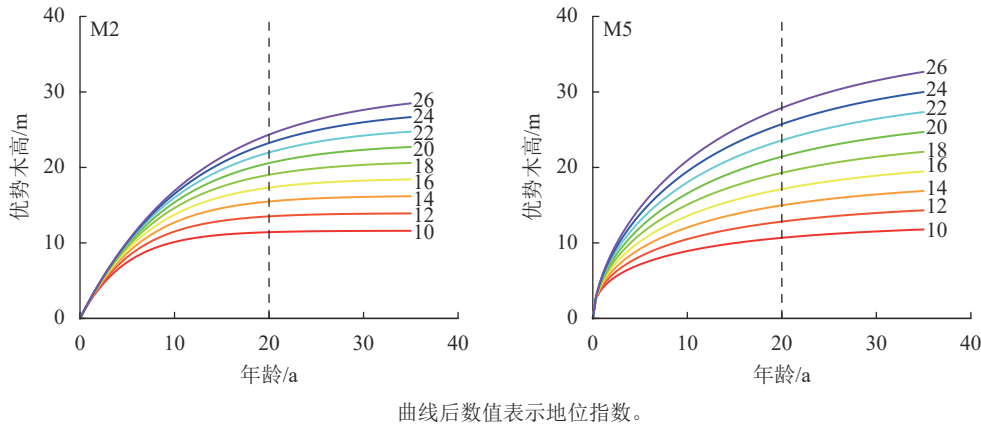


图3 模型 M2 和 M5 地位指数曲线簇
Figure 3 Site index curves generated with difference site index model M2 and M5

2.2 因子特征选择

2.2.1 基于随机森林的因子特征选择 依据基于内置的因子特征得出: 各因子对立地质量的影响程度从大到小依次为地貌、腐殖质厚度、坡向、土层厚度、坡位、坡度、海拔、土壤类型。其中重要性占比地貌为 0.471, 腐殖质厚度为 0.111, 坡向为 0.094, 土层厚度为 0.090, 坡位为 0.078, 坡度为 0.058, 海拔为 0.057, 土壤类型为 0.041。依据基于置换的因子特征得出: 各因子对立地质量的影响程度从大到小依次为地貌、腐殖质厚度、土层厚度、土壤类型、海拔、坡位、坡度、坡向。其中重要性占比地貌为 0.468, 腐殖质厚度为 0.168, 土层厚度为 0.160, 土壤类型为 0.148, 海拔为 0.117, 坡位为 0.117, 坡度为 0.109, 坡向为 0.101 (图 4)。

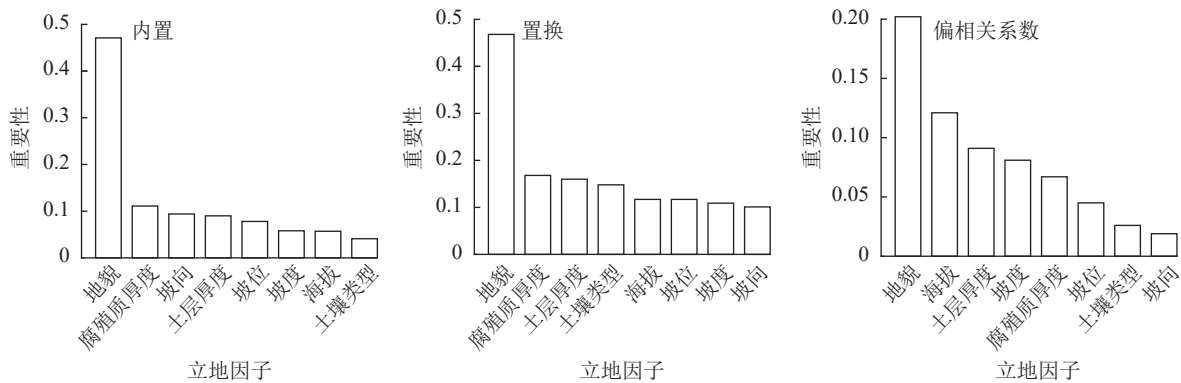


图4 基于内置、置换和偏相关系数的特征重要性排序
Figure 4 Site factors importance ranking based on feature importance, permutation importance and partial correlation

2.2.2 基于偏相关系数的因子特征选择 依据偏相关系数对立地因子进行特征性选择得出: 各因子对立地质量的影响程度从大到小依次为地貌、海拔、土层厚度、坡度、腐殖质厚度、坡位、土壤类型、坡向。其中重要性占比地貌为 0.202, 海拔为 0.121, 土层厚度为 0.091, 坡度为 0.081, 腐殖质厚度为 0.067, 坡位为 0.045, 土壤类型为 0.026, 坡向为 0.019 (图 4)。

通过基于偏相关系数的因子特征选择方法得到前 3 个因子的贡献度为 41.4%; 通过基于内置的因子特征选择方法得到前 3 个因子的贡献度为 67.6%; 通过基于置换的因子特征选择方法得到前 3 个因子的贡献度为 79.6%(表 3)。根据因子选择贡献最高的立地类型进行划分, 选取地貌、腐殖质厚度、土层厚度作为主导因子, 进行立地类型划分与立地质量评价。

2.3 立地类型划分与立地质量评价

采用地位指数模型预测每个小班的地位指数，根据地位指数值，采用四等分，根据需要将地位指数数值划分为优 (>14.31)、良 (>9.54~14.31)、中 (>4.76~9.54)、差 (≤ 4.76) 4 个等级。

按照因子特征对庆元县杉木立地进行分类，分别采用地貌、腐殖质厚度、土层厚度来进行立地类型划分，共划分出 17 个立地类型 (表 4)。由表 4 可以看出：庆元县杉木立地质量等级整体属于良好，共有 2 564 块小班立地质量处于良及以上，占比约 52%。说明研究样地的立地水平良好，总体适宜杉木生长，可以进行一定的人工抚育提高林分生产力。

表 3 贡献度细则

因子特征 选择方法	第1贡献因子	第2贡献因子	第3贡献因子	贡献度/%
偏相关系数	地貌	海拔	土层厚度	41.4
内置	地貌	腐殖质厚度	坡向	67.6
置换	地貌	腐殖质厚度	土层厚度	79.6

表 4 立地类型与立地质量等级表

Table 4 Site classification and evaluation of *C. lanceolata*

立地类型区	立地类型组	土层厚度	立地得分值	小班数量/块	立地质量	立地类型区	立地类型组	土层厚度	立地得分值	小班数量/块	立地质量
		厚土层	12.39	365	良			厚土层	12.21	52	良
	厚腐殖质层	中土层	12.76	15	良		厚腐殖质层	中土层	10.46	48	良
		厚土层	11.81	284	良			薄土层	10.11	1	良
低山	中腐殖质层	中土层	11.65	953	良	中山	厚土层	10.31	225	良	
		薄土层	11.10	17	良		中腐殖质层	中土层	9.13	1763	中
		厚土层	9.84	162	良			薄土层	9.08	324	中
	薄腐殖质层	中土层	10.97	371	良			厚土层	8.53	24	中
		薄土层	10.40	71	良		薄腐殖质层	中土层	9.22	175	中
								薄土层	9.51	70	中

3 讨论与结论

在地位指数模型构建中，广义代数差分法是当前较为理想的方法。庆元县地处浙江省西南地区，杉木生长速度良好，具有一定的代表性。通常认为，分布在长江以南的杉木在 20 a 左右已经达到生长高峰，这与本研究构建的地位指数模型相一致，而在趋近 20 a 基准年龄下，曲线逐渐趋于平缓，曹元帅等^[7]、牛亦龙等^[22]也得到了相似的研究结果。本研究选取以 Richards 为基础方法的广义代数差分方程，构建可用于预测庆元县杉木林达到预测年龄下的优势木高的地位指数模型，模型精度较高，形成的多形地位指数曲线符合杉木的生长特性，可以较好地预估当地的杉木林地位指数。

确定主导立地因子是立地质量评价与立地分类中的一项重要工作。张沛健等^[23]研究得出：影响雷州半岛-海南岛沿海台地地区尾细桉 *Eucalyptus urophylla* × *E. tereticornis* 生长的主要立地因子是土壤质地、成土母岩和土壤密度；郭艳荣等^[13]研究表明：影响杉木人工林生长的主导因子为土层厚度、坡向、腐殖质厚度；杨海宾等^[20]认为：影响杉木的主导因子是坡向、坡位和土壤类型。本研究得出地貌、腐殖质厚度、土层厚度为杉木生长的主导因子。地貌条件直接影响气候和土壤，局部地形对杉木生长有很大影响，随着地势升高，杉木分布高度也随之上升，冻害和断梢现象更为频繁，地貌高低与杉木树高生长呈明显的负相关^[24-25]。因此，地貌为影响杉木生长的主导因子之一。腐殖质是土壤有机质的一种，含有多种养分，能提高土壤的保肥、保水性，并能缓解土壤的酸碱度，有利于微生物活动，促进林木生长，适宜的腐殖质厚度为杉木根系发展提供了良好的环境与空间^[26]，因此腐殖质厚度同样是影响杉木生长的主导因子之一；土层厚度影响土壤养分、水分总储量和根系分布的空间范围，尤其是杉木喜肥沃、深厚、疏松、湿润且排水良好的土壤，对土层厚度的要求较高，因此土层厚度也是影响杉木生长的主导因子之一。林区整体气候温润，导致部分因子差异性不大，这也是坡向、坡位等在因子选择中不明显的原因。本研究采用主导因子对庆元县地区进行立地划分，通过得分值的计算，将立地类型划分为优、良、中、差共 4 个等级，其中 2 564 块研究区小班的立地质量为良及以上，表明庆元县立地质量条件良好。

因子特征选择中, 偏相关系数是在消除其他变量影响的条件下, 计算某2个变量之间的相关系数; 基于内置的因子特征选择方法为在树结构的内部节点中, 因子按方差减少进行选择, 但存在当一个特征被选择后, 与其相关的其他特征的重要度则会变得很低等问题。正如研究中当第1个立地因子地貌被确认后, 其余因子的重要性都呈现较低的水平; 基于置换的因子特征选择方法用于克服使用方差减少计算出的默认特征重要性的缺点, 对模型的特征没有偏见, 也不局限于特定的模型类别, 适用性较广。在以上3种方法中, 地貌均被选择, 说明地貌的变化导致预测结果波动性最大。而腐殖质厚度、土层厚度在3种方法中均各有体现, 在基于内置的因子特征选择中, 腐殖质厚度被选择, 是其基于树子节点划分时, 方差的变化程度较小且导致预测结果变化较大。而在偏相关系数的因子特征选择方法中, 消除了其他变量的影响后, 其表现性不如土层厚度。立地质量评价中通常使用偏相关系数选择主导因子, 而包括随机森林在内的众多机器学习算法均提供了因子选择的方法或参数, 因此在主导因子划分这一问题上, 可以不局限于偏相关系数去探索更适用于立地质量评价的主导因子选择方法。此外, 由于气候、水文等条件难以量化^[27], 因此使用间接因子进行拟合, 难免在确信度上有所缺失。在今后运用立地因子对立地质量进行评价时, 可以使用降维手段筛选出参与建模的立地因子, 进一步提高评价结果的预测精度。

4 参考文献

- [1] 徐罗, 亢新刚, 郭韦韦, 等. 天然云冷杉针阔混交林立地质量评价[J]. 北京林业大学学报, 2016, **38**(5): 11–22.
XU Luo, KANG Xin'gang, GUO Weiwei, *et al.* Site quality evaluation of natural spruce-fir and broadleaf mixed stands [J]. *Journal of Beijing Forestry University*, 2016, **38**(5): 11–22.
- [2] SKOVSGAARD J P, VANCLAY J K. Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands [J]. *Forestry*, 2008, **81**(1): 13–31.
- [3] BAILEY R L, CLUTTER J L. Base-age invariant polymorphic site curves [J]. *Forest Science*, 1974, **20**(2): 155–159.
- [4] 倪成才, 于福平, 张玉学, 等. 差分生长模型的应用分析与研究进展[J]. 北京林业大学学报, 2010, **32**(4): 284–292.
NI Chengcai, YU Fuping, ZHANG Yuxue, *et al.* Application analysis and recent advances of projection growth models [J]. *Journal of Beijing Forestry University*, 2010, **32**(4): 284–292.
- [5] 王志波, 季蒙, 李永乐, 等. 华北落叶松人工林差分地位指数模型构建[J]. 林业资源管理, 2021(1): 156–163.
WANG Zhibo, JI Meng, LI Yongle, *et al.* Construction of difference site index model for *Larix principis-rupprechtii* plantation [J]. *Forest Resources Management*, 2021(1): 156–163.
- [6] 李斌成, 许业洲, 袁慧, 等. 湖北省日本落叶松差分型立地指数模型构建[J]. 森林与环境学报, 2020, **40**(4): 433–441.
LI Bincheng, XU Yezhou, YUAN Hui, *et al.* Construction of a differential site index model for a *Larix kaempferi* plantation in Hubei Province [J]. *Journal of Forest and Environment*, 2020, **40**(4): 433–441.
- [7] 曹元帅, 孙玉军. 基于广义代数差分法的杉木人工林地位指数模型[J]. 南京林业大学学报(自然科学版), 2017, **41**(5): 79–84.
CAO Yuanshuai, SUN Yujun. Generalized algebraic difference site index model for Chinese fir plantation [J]. *Journal of Nanjing Forestry University (Natural Sciences Edition)*, 2017, **41**(5): 79–84.
- [8] CIESZEWSKI C J, BAILEY R L. Generalized algebraic difference approach: theory based derivation of dynamic site equations with polymorphism and variable asymptotes [J]. *Forest Science*, 2000, **46**(1): 116–126.
- [9] MCKENNEY D W, PEDLAR J H. Spatial models of site index based on climate and soil properties for two boreal tree species in Ontario, Canada [J]. *Forest Ecology and Management*, 2003, **175**(1): 497–507.
- [10] CALAMA R, MONTERO G. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain [J]. *Canadian Journal of Forest Research*, 2004, **34**(1): 150–163.
- [11] GONZÁLEZ-RODRÍGUEZ M A, DIÉGUEZ-ARANDA U. Exploring the use of learning techniques for relating the site index of radiata pine stands with climate, soil and physiography [J/OL]. *Forest Ecology and Management*, 2020, **458**(C): 117803[2022-11-01]. doi: 10.1016/j.foreco.2019.117803.
- [12] NUNES L, PATRÍCIO M, TOMÉ J, *et al.* Modeling dominant height growth of maritime pine in Portugal using GADA methodology with parameters depending on soil and climate variables [J]. *Annals of Forest Science*, 2011, **68**(2): 311–323.

- [13] 郭艳荣, 刘洋, 吴保国. 福建省宜林地立地质量的分级与数量化评价[J]. 东北林业大学学报, 2014, 42(10): 54 – 59.
GUO Yanrong, LIU Yang, WU Baoguo. Evaluating dividing rank and quantification of site quality of suitable land for forest in Fujian Province, China [J]. *Journal of Northeast Forestry University*, 2014, 42(10): 54 – 59.
- [14] 石梦婷, 冯自茂, 郝红科, 等. 陕西洛南油松人工林立地质量评价[J]. 西北林学院学报, 2021, 36(6): 127 – 133.
SHI Mengting, FENG Zimao, HAO Hongke, et al. Site quality evaluation of *Pinus tabulaeformis* plantations in Luonan, Shaanxi [J]. *Journal of Northwest Forestry University*, 2021, 36(6): 127 – 133.
- [15] 高若楠, 谢阳生, 雷相东, 等. 基于随机森林模型的天然林立地生产力预测研究[J]. 中南林业科技大学学报, 2019, 39(4): 39 – 46.
GAO Ruonan, XIE Yangsheng, LEI Xiangdong, et al. Study on prediction of natural forest productivity based on random forest model [J]. *Journal of Central South University of Forestry & Technology*, 2019, 39(4): 39 – 46.
- [16] SCOLFORO H F, FERNANDO D C N, SCOLFORO J R S, et al. Modeling dominant height growth of eucalyptus plantations with parameters conditioned to climatic variations [J]. *Forest Ecology and Management*, 2016, 380: 182 – 195.
- [17] SEKI M, SAKICI O E. Dominant height growth and dynamic site index models for Crimean pine in the Kastamonu-Takprü region of Turkey [J]. *Canadian Journal of Forest Research*, 2017, 47(11): 1441 – 1449.
- [18] 雷相东, 符利勇, 李海奎, 等. 基于林分潜在生长量的立地质量评价方法与应用[J]. 林业科学, 2018, 54(12): 116 – 126.
LEI Xiangdong, FU Liyong, LI Haikui, et al. Methodology and applications of site quality assessment based on potential mean annual increment [J]. *Scientia Silvae Sinicae*, 2018, 54(12): 116 – 126.
- [19] 沈剑波, 雷相东, 雷渊才, 等. 长白落叶松人工林地位指数及立地形的比较研究[J]. 北京林业大学学报, 2018, 40(6): 1 – 8.
SHEN Jianbo, LEI Xiangdong, LEI Yuancai, et al. Comparison between site index and site form for site quality evaluation of *Larix olgensis* plantation [J]. *Journal of Beijing Forestry University*, 2018, 40(6): 1 – 8.
- [20] 杨海滨, 张茂震, 丁丽霞, 等. 基于最大胸径生长率的浙江省杉木人工林立地质量评价[J]. 浙江农林大学学报, 2020, 37(1): 105 – 113.
YANG Haibin, ZHANG Maozhen, DING Lixia, et al. Site quality evaluation of *Cunninghamia lanceolata* plantations in Zhejiang Province based on maximum DBH growth rate [J]. *Journal of Zhejiang A&F University*, 2020, 37(1): 105 – 113.
- [21] BREIMAN L. Random forests [J]. *Machine Learning*, 2001, 45(1): 5 – 32.
- [22] 牛亦龙, 董利虎, 李凤日. 基于广义代数差分法的长白落叶松人工林地位指数模型[J]. 北京林业大学学报, 2020, 42(2): 9 – 18.
NIU Yilong, DONG Lihu, LI Fengri. Site index model for *Larix olgensis* plantation based on generalized algebraic difference approach derivation [J]. *Journal of Beijing Forestry University*, 2020, 42(2): 9 – 18.
- [23] 张沛健, 卢万鸿, 徐建民, 等. 雷琼地区尾细桉人工林立地类型划分及其质量评价[J]. 林业科学研究, 2021, 34(6): 130 – 139.
ZHANG Peijian, LU Wanhong, XU Jianmin, et al. Site classification and quality evaluation of *Eucalyptus urophylla* × *E. tereticornis* plantation in Hainan Island and Leizhou Peninsula Region [J]. *Forest Research*, 2021, 34(6): 130 – 139.
- [24] ALVES L F, VIEIR S A, SCARANELLO M A, et al. Forest structure and live aboveground biomass variation along an elevational gradient of tropical Atlantic moist forest (Brazil) [J]. *Forest Ecology and Management*, 2010, 260(5): 679 – 691.
- [25] ASNER G P, HUGHES R F, VARGA T A, et al. Environmental and biotic controls over aboveground biomass throughout a tropical rain forest [J]. *Ecosystems*, 2009, 12(2): 261 – 278.
- [26] QUICHIMBO P, JIMÉNEZ L S, VEINTIMILLA D, et al. Forest site classification in the southern Andean region of Ecuador: a case study of pine plantations to collect a base of soil attributes [J/OL]. *Forests*, 2017, 8(12): 473 [2022-11-01]. doi:10.3390/f8120473.
- [27] ALBERT M, SCHMIDT M. Climate-sensitive modelling of site-productivity relationships for Norway spruce (*Picea abies* (L.) Karst.) and common beech (*Fagus sylvatica* L.) [J]. *Forest Ecology and Management*, 2009, 259(4): 739 – 749.