

基于机器学习算法的广西桉树适宜性研究

杜雨菲, 吴保国, 陈玉玲

(北京林业大学 信息学院, 北京 100083)

摘要:【目的】探索立地因子与桉树 *Eucalyptus* 适宜性之间的关系, 开展树种适宜性研究, 为桉树适宜性研究提供新思路, 为科学造林提供支持。【方法】以广西桉树人工林为研究对象, 选取广西国有高峰林场的 1 883 个森林资源小班调查数据, 分别运用朴素贝叶斯、支持向量机、随机森林算法作为树种适宜性评价方法, 构建桉树适宜性分类模型。输入为地貌类型、海拔、坡向、坡位、坡度、凋落物厚度、腐殖质层厚度、土层厚度、石砾含量、成土母质, 土壤类型等 11 个立地因子信息, 输出为桉树适宜性。【结果】3 种算法构建的模型拟合精度依次为 63.18%、69.73%、78.03%, 泛化精度依次为 64.33%、67.93%、78.18%。相比于朴素贝叶斯、支持向量机算法, 随机森林算法分类效果更好。立地因子重要性排序由高到低依次为: 海拔、土层厚度、坡向、坡度、石砾含量、凋落物厚度、坡位、腐殖质层厚度、土壤类型、地貌类型、成土母质。200~350 m 海拔、80~100 cm 土层厚度的地区比较适宜桉树生长。【结论】基于机器学习算法构建的桉树适宜性评价模型可以较好地对桉树的适宜性做出预测。图 2 表 4 参 19
关键词: 森林测计学; 适宜性; 机器学习; 朴素贝叶斯; 支持向量机; 随机森林; 桉树
中图分类号: S758.8 **文献标志码:** A **文章编号:** 2095-0756(2020)01-0122-07

Eucalyptus suitability in Guangxi based on machine learning algorithms

DU Yufei, WU Baoguo, CHEN Yuling

(School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China)

Abstract: [Objective] The aim is to provide a new idea for tree species suitability evaluation, provide a support for scientific afforestation, and explore the relationship between site factors and tree suitability. [Method] Take a *Eucalyptus* plantation in Guangxi as the research object, 1 883 forest resource sub-compartment survey data of Guangxi state-owned Gaofeng Forest Farm were selected. Then, Naive Bayesian, Support Vector Machine, and Random Forest algorithm were used to evaluate the suitability of tree species and to construct a suitability classification model for *Eucalyptus*. Eleven site factors, namely, landform type, elevation, aspect, slope position, slope, litter thickness, humus layer thickness, soil layer thickness, gravel content, parent material, and soil type were input with the output being *Eucalyptus* suitability. [Result] The fitting accuracy of the three models was 63.18% for Naive Bayesian, 69.73% for Support Vector Machine, and 78.03% for Random Forest algorithm with a generalization accuracy of 64.33% for Naive Bayesian, 67.93% for Support Vector Machine, and 78.18% for Random Forest algorithm. The order of importance for site factors was elevation>soil layer thickness>aspect>slope>gravel content>litter thickness>slope position>humus layer thickness>soil type>landform type>parent material. Overall, *Eucalyptus* was more suitable for growth in areas of 200–350 m altitude and 80–100 cm soil layer thickness. [Conclusion] Thus, machine learning classification algorithms could be used to fit the non-linear relationship between tree species suitability and site factors. [Ch, 2 fig. 4 tab. 19 ref.]

Key words: forest mensuration; suitability; machine learning; Naive Bayes; Support Vector Machine; Random

收稿日期: 2019-03-18; 修回日期: 2019-05-30

基金项目: “十三五”国家重点研发计划项目(2017YFD0600906)

作者简介: 杜雨菲, 从事林业信息技术研究。E-mail: 723216546@qq.com。通信作者: 吴保国, 教授, 博士生导师, 从事林业信息技术与信息管理、林业决策支持系统研究。E-mail: wubg@bjfu.edu.cn

Forest; *Eucalyptus*

桉树 *Eucalyptus* 栽培是广西林业生产发展中的优势产业^[1]。作为短周期工业原料林首选的造林树种之一，桉树效益高、周期短，但生长受立地条件影响较大，规模栽培时需要进行适宜性研究。树种适宜性研究是当前开展适地适树、造林决策研究的热点^[2]。王小明等^[3]综合了气候、土壤、地形等环境因子建立 Logistic 回归模型用以确定香榧 *Torreya grandis* ‘Merrillii’ 适宜种植区域，模型检验数据集的总正确率达到了 69.8%。KOO 等^[4]基于环境因子，建立物种分布模型和时间模拟模型研究云杉 *Picea rubens* 适生区，模型验证结果的曲线下面积(area under curve, AUC)达 0.99。胡秀等^[5]基于温度、降雨及海拔等环境因子，采用 MaxEnt 模型软件构建了檀香 *Santalum album* 的适宜性预测模型，AUC 值为 0.98。高若楠等^[6]选取立地因子，利用随机森林模型研究了杉木 *Cunninghamia lanceolata* 的适宜性，模型泛化精度达 89.5%。PIRI-SAHRA GARD 等^[7]利用随机森林模型分别探讨了环境因子与白桉 *Haloxylon persicum* 等 5 种植物分布之间的关系，AUC 值总体在 0.95 以上。相较于传统数学建模技术，物种分布模型和机器学习模型在构建树种适宜性评价模型时效果较好。目前，应用于树种适宜性研究的机器学习分类算法层出不穷，但通过对比分析多种不同的分类算法，从而进行树种适宜性评价的研究相对较少。本研究以广西桉树为对象，使用朴素贝叶斯(naive bayesian, NB)^[8]、支持向量机(support vector machine, SVM)^[9]、随机森林(random forest, RF)^[10]等 3 种机器学习分类算法，探索立地因子与桉树适宜性之间的关系，开展树种适宜性研究，为桉树适宜性研究提供新思路，为科学造林提供支持。

1 材料与方法

1.1 研究区概况

研究区广西国有高峰林场(22°49′~23°15′N, 108°08′~108°53′E)地处广西壮族自治区南宁市，属低丘陵山地地带，平均海拔为 200~500 m。亚热带季风气候，年平均气温为 20.8~21.9 ℃，年均相对湿度 80%以上。海拔 300 m 以下的土壤绝大部分为赤红壤^[11]。地理位置、气候和土壤条件均十分优越，有利于亚热带植物的规模化种植。

1.2 数据来源

数据来源于广西高峰林场森林资源规划设计调查数据中的桉树小班数据，包括立地因子、林分平均年龄、优势木平均高。立地因子包括地貌类型、海拔、坡向、坡位、坡度、凋落物厚度、腐殖质层厚度、土层厚度、石砾含量、成土母质和土壤类型。地貌类型有低山、丘陵 2 种，坡向包括东、南、西、北、东北、东南、西北、西南、无坡向，坡位包括脊部、上坡、中坡、下坡、谷地、平地，成土母质包括砂岩、第四纪红土，土壤类型包括赤红壤、黄壤、红壤。整理数据，剔除缺失严重记录、异常数据，得到桉树小班数据 1 883 个。

1.3 研究方法

1.3.1 朴素贝叶斯算法 朴素贝叶斯算法是一种基于概率论的机器学习分类算法^[8]。针对桉树小班数据训练集 D ，类别集合为 y_j ， y_1 代表适宜桉树生长， y_2 代表不适宜桉树生长， a_i 是待分类的小班，有 $a_1, a_2, \dots, a_{11}$ 共 11 个立地因子。统计在各类别下各立地因子的条件概率估计值，即估计第 i 个立地因子在第 j 个类别中出现的概率 $P(a_i | y_j)$ ，根据特征独立性假设以及贝叶斯定理，桉树适宜性分类结果(h_{nb})可用朴素贝叶斯分类器表示为：

$$h_{nb} = \arg \max_j P(y_j | x) = \arg \max_j P(a_i | y_j) P(y_j) = \arg \max_j P(y_j) \prod_{i=1}^{11} P(a_i | y_j). \quad (1)$$

1.3.2 支持向量机算法 支持向量机是一种二分类机器学习算法^[12]。在由立地因子构成的特征空间中寻找 1 个分类超平面对桉树小班数据训练样本进行归类(适宜或不适宜)，分类超平面遵循间隔最大化原则。设有 2 类线性可分的样本集合(g_i, h_i)， $i=1, \dots, n$ ； $h_i \in \{+1, -1\}$ ；线性判别函数表示为：

$$f(g) = \omega x + b. \quad (2)$$

式(2)中： ω 为平面的法向量， b 为截距。通过最大化间隔，得到最优分类面函数式(3)。对线性不可分的数据，也可以通过核函数将其映射到高维空间，使得样本线性可分。

$$f(g)=\text{sgn}\{(\omega g)+b\}=\text{sgn}\left\{\sum_{i=1}^n a_i^* h_i(gg)+b^*\right\}。 \quad (3)$$

式(3)中: a_i^* 是不为零的样本, 即支持向量, b^* 是分类阈值。

1.3.3 随机森林算法 随机森林是一种集成机器学习算法^[13]。采用 Bootstrap 重抽样法对桉树小班数据训练集 D 进行 n 次抽样, 得到 D_1 、 D_2 、 $\dots$ 、 D_n 共 n 个训练子集; 各训练子集分别训练 1 棵决策树, 组成随机森林。在单棵树的训练过程中, 随机选出部分立地因子用以确定决策树的分割节点, 得到 n 种结果; 使用简单投票法, 得到最多票数的类别或者类别之一为最终的桉树适宜性评价模型, 输出结果见式(4)。

$$H(x)=\arg \max_Y \sum_{i=1}^n I[h_i(x)=Y]。 \quad (4)$$

式(4)中: $H(x)$ 为组合分类模型; $h_i(x)$ 为单个决策树分类模型; Y 为输出桉树适宜性的变量; $I(\cdot)$ 为示性函数。

1.3.4 模型评价指标 混淆矩阵(confusion matrix)也称误差矩阵(error matrix), 是评价模型分类效果的常用的指标^[14]。如表 1 所示: 混淆矩阵的每一列代表了桉树适宜性评价模型的预测类别, 每一行代表该小班真实的归属类别, 主对角线元素的总和为被正确分类的小班总数(N)。模型的精度(A , 包括拟合精度和泛化精度)可用小班数与小班总数的比值来表示:

$$A=(N_{TP}+N_{TN})/N。 \quad (5)$$

式(5)中: N_{TP} 为正类预测为正类的小班数; N_{TN} 为负类预测为负类的小班数。分类误差率(classification error rate)为该类别预测错误的小班数与该类别小班总数的比值, 包括模型对于桉树生长适宜性的分类误差率[式(6)]和不适宜性的分类误差率[式(7)]。精度、生长适宜性的分类误差率(ε_1)和不适宜性的分类误差率(ε_2)通常作为衡量桉树适宜性评价模型判定能力的指标^[6]。

$$\varepsilon_1=N_{FN}/(N_{TP}+N_{FN}); \quad (6)$$

$$\varepsilon_2=N_{FP}/(N_{FP}+N_{TN})。 \quad (7)$$

式(6)和式(7)中: N_{FN} 为正类预测为负类的小班数; N_{FP} 为负类预测为正类的小班数。

2 结果与分析

2.1 模型实现

树种适宜性评价标准最常用的是地位指数(site index, SI)^[6], 各小班地位指数可通过林分平均年龄和优势木平均高得到^[15]; 地位指数小于平均值的小班判定为不适宜桉树生长, 大于或等于平均值的判定为适宜桉树生长^[6]。本研究的 1 883 个样本数据中, 适宜桉树生长的样本有 1 005 个, 不适宜的有 878 个, 样本量存在一定的不平衡性。利用机器学习算法解决分类问题时, 数据集不平衡会对模型效果造成影响, 因此需要进行平衡化处理。在不损失原始样本的前提下, 通过 SMOTE 算法^[16]对样本构成做平衡化处理, 共得到样本量 3 512 个, 其中适宜桉树生长的样本 1 756 个, 不适宜的 1 756 个; 将实验数据按 70% 和 30% 的比例分为训练样本和测试样本, 分别用于模型的训练和测试。

使用 naiveBayes() 函数构建朴素贝叶斯模型、svm() 函数构建支持向量机模型、randomForest() 函数构建随机森林模型。3 种模型的输入均为地貌类型、海拔、坡向、坡位、坡度、凋落物厚度、腐殖质层厚度、土层厚度、石砾含量、成土母质, 土壤类型, 输出均为桉树生长适宜性。利用模型评价指标对比不同模型, 取最优模型确定为桉树适宜性评价模型并进行桉树生长适宜性预测。对给定立地因子的小班, 将立地因子输入选取的模型, 输出该小班适宜桉树生长的概率, 判断该小班是否适宜桉树生长。进行立地因子重要性评估, 分析立地因子对桉树生长的影响, 得出适宜桉树生长的立地条件。桉树适宜性预测模型构建流程如图 1 所示。

2.2 模型精度分析

多次训练发现: 朴素贝叶斯、支持向量机、随机森林算法构建的桉树适宜性评价模型误差变化均较稳定, 混淆矩阵(表 2)拟合精度为 63.18%、69.73% 和 78.03%, 使用测试数据集对模型进行检验, 混淆矩阵泛化精度分别为 64.33%、67.93% 和 78.18%。与朴素贝叶斯、支持向量机算法相比, 随机森林算法

表 1 混淆矩阵

Table 1 Confusion matrix		
实际类别	预测类别	
	适宜/个	不适宜/个
适宜	N_{TP}	N_{FN}
不适宜	N_{FP}	N_{TN}

表 2 3 种模型混淆矩阵

		拟合分析				预测分析			
模型名称	实际类别	适宜/个	不适宜/个	分类误差率	拟合精度/%	预测适宜	预测不适宜	分类误差率	泛化精度/%
朴素贝叶斯算法	适宜	737	491	0.399 837 1	63.18	329	202	0.380 414 3	64.33
	不适宜	414	816	0.336 585 4		174	349	0.332 696 0	
支持向量机算法	适宜	752	476	0.387 622 1	69.73	334	197	0.370 998 1	67.93
	不适宜	268	962	0.217 886 2		141	382	0.269 598 5	
随机森林算法	适宜	911	317	0.258 143 3	78.03	402	129	0.242 937 9	78.18
	不适宜	223	1 007	0.181 300 8		101	422	0.193 116 6	

预测精度更高，可作为桉树适宜性评价的模型。

2.3 桉树适宜性预测

利用随机森林算法构建桉树适宜性评价模型并对桉树进行生长适宜性预测，预测数据为广西桉树固定样地数据，各样地的地位指数可通过查阅地位指数表得到^[15]。该数据中桉树的地位指数的平均值为 15.09，将地位指数小于平均值的样地判定为不适宜桉树生长；将地位指数大于或等于平均值的样地判定为适宜桉树生长。随机选取 5 个样本进行模型验证，将立地因子输入模型，输出桉树适宜性概率及适宜性判断结果；通过与地位指数的比对(表 3)可知：本研究使用随机森林算法构建的桉树适宜性评价模型在实际中是可以使用的。

2.4 立地因子重要性评估

利用随机森林算法对立地因子进行重要性评估^[10]。对某个立地因子*j*随机取值，通过评估桉树适宜性评价模型分类准确性下降的程度来评估*j*的重要性，分类准确性下降程度越大，说明*j*越为重要。计算方法如公式(8)所示：

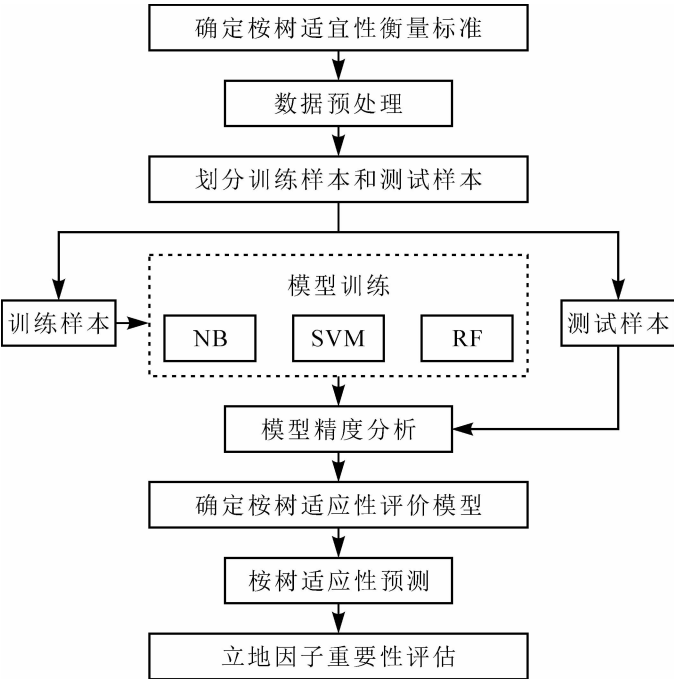


图 1 模型构建流程图
Figure 1 Flowchart of model building

$$V_j=\sum_{r=1}^{N_T}(E_r'-E_r)。$$

(8)

式(8)中：*E_r*为*j*的值随机后的袋外(out of bag, OOB)误差，*E_r*为*j*的值随机前的 OOB 误差，*N_T*为分类树的数量。标准化处理后得到的平均准确度降低程度(mean decrease accuracy, MDA)可用来描述立地因

表 3 随机森林算法模型判断结果

		Table 3 Predicted results of random forest models												
序号	地貌类型	海拔/m	坡向	坡位	坡度/°	凋落物厚度/cm	腐殖质层厚度/cm	土层厚度/cm	石砾含量/%	成土母质	土壤类型	模型预测结果		地位指数
1	丘陵	295	无坡向	平地	2	3	5	95	10	-	红壤	0.851 7	适宜	22
2	低山	780	东北	上坡	10	1	3	80	0	-	红壤	0.098 3	不适宜	12
3	丘陵	90	北	下坡	25	5	5	100	8	-	赤红壤	0.082 1	不适宜	10
4	丘陵	195	东	中坡	28	2	8	100	5	-	赤红壤	0.951 7	适宜	26
5	低山	230	西南	脊	23	5	25	110	5	-	赤红壤	0.976 9	适宜	28

说明：“-”表示数据未获取

子 j 的重要性。

对分类树节点作 t 分割, 计算使用立地因子 j 前与使用后基尼指数的减小值(D_{Gj}), 对所有节点的 D_{Gj} 求和后再对所有分类树 N_T 取平均, 得到平均基尼指数降低程度(mean decrease Gini, MDG)。MDG越大, 立地因子 j 越重要。基尼指数 $Gini(t)$ 的计算方法如公式(9)所示。

$$Gini(t)=1-\sum_{i=1}^k [p(i|t)]^2. \quad (9)$$

式(9)中: $p(i|t)$ 为类别 i 在节点 t 处的概率, k 为分类结果数, 在本研究中取值为2。

使用 varImpPlot() 函数对 11 个立地因子进行重要性评估。由表 4 可知: 不同重要性评估方法对立地因子的排序结果基本一致; 立地因子的重要性排序由高到低依次为: 海拔、土层厚度、坡向、坡度、石砾含量、凋落物厚度、坡位、腐殖质层厚度、土壤类型、地貌类型、成土母质。

表 4 立地因子重要性评估

Table 4 Importance assessment of site factors

立地因子	MDA		MDG		综合排序
	评估值	排序	评估值	排序	
地貌类型	6.532 154	10	2.917 723	10	10
海拔	64.038 512	1	119.872 133	1	1
坡向	57.933 214	3	90.212 541	2	3
坡位	29.452 163	7	35.412 112	7	7
坡度	45.712 589	5	87.012 457	3	4
凋落物厚度	40.893 212	6	41.215 115	6	6
腐殖质层厚度	24.314 587	8	28.732 522	8	8
土层厚度	60.091 123	2	84.971 211	4	2
石砾含量	52.331 421	4	67.124 189	5	5
成土母质	3.158 456	11	2.312 457	11	11
土壤类型	14.796 521	9	7.354 123	9	9

选取对桉树生长影响最大的 2 个立地因子(海拔和土层厚度), 利用随机森林算法进行单因素分析。由图 2 可知: 研究区海拔为 200~350 m、土层厚度为 80~100 cm 的地区比较适合桉树生长。

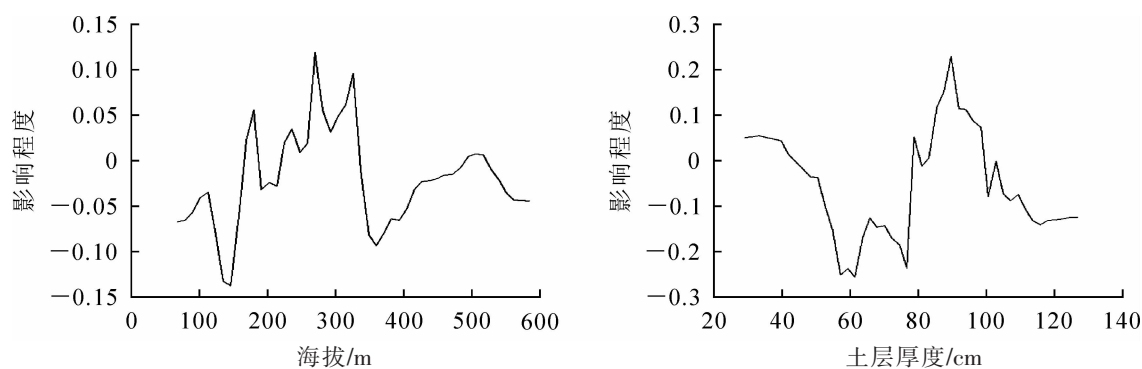


图 2 海拔、土层厚度对桉树生长的影响

Figure 2 Effects of altitude and soil thickness on growth of *Eucalyptus*

3 结论与讨论

基于朴素贝叶斯算法、支持向量机算法、随机森林算法 3 种算法构建的模型拟合精度分别为 63.18%、69.73%和 78.03%, 泛化精度分别为 64.33%、67.93%和 78.18%; 相较于朴素贝叶斯、支持向量机算法, 随机森林算法对缺失数据不敏感, 在训练的过程中能检测到特征与特征之间的互相影响, 模型泛化能力强, 具有更高的预测精度, 在本研究中分类效果最好。缺少特征独立性假设和立地因子数据是朴素贝叶斯算法分类效果差的原因; 而缺少通用的解决方案, 对缺失数据敏感, 受核函数的影响较大

等导致了支持向量机算法分类效果欠理想。本研究采用的多模型对比为以后其他树种适宜性研究选取模型提供了参考。

海拔、土层厚度、坡向、坡度等立地因子对桉树生长影响较大，地貌类型、成土母质等则较小。原因可能是海拔高度、坡向、坡度的改变造成空气温度、空气湿度、太阳辐射等变化^[6]，从而影响桉树生长；土层厚度与土壤养分、矿元素等密切相关^[17]，研究区的桉树种植区域，地貌类型和成土母质均比较单一，因此对桉树生长的影响并不明显。对海拔、土层厚度等立地因子的单因素分析发现：桉树适宜生长地区多数海拔为 200~350 m，土层厚度为 80~100 cm。研究认为^[18]：海拔高度低于 350 m，桉树径生长随海拔升高而增粗，当海拔大于 350 m，环境热量不足桉树容易引发低温冻害^[19]。土层厚度对桉树的影响体现在土壤的营养状况和给树木生长提供的养分上^[17]，本研究发现土层越厚，土壤营养条件越好，也越适宜桉树生长。总的来说，不同的立地因子对桉树生长的影响程度不同，选择桉树种植区域时应客观考虑各个立地因子的影响程度，从而合理地调整立地条件的组合，最大程度满足桉树生长。

基于机器学习算法构建的树桉树适宜性评价模型可以较好地对桉树的适宜性做出预测，为科学造林提供依据。树种适宜性分析不仅要立地分为适宜该树种生长以及不适宜该树种生长，还可以进一步对其进行细分，从二分类问题转变为多分类问题，进一步研究机器学习算法在树种适宜性分析中的应用。

4 参考文献

- [1] 黄国勤, 赵其国. 广西桉树种植的历史、现状、生态问题及应对策略[J]. 生态学报, 2014, **34**(18): 5142 – 5152.
HUANG Guoqin, ZHAO Qiguo. The history, status quo, ecological problems and countermeasures of *Eucalyptus* plantations in Guangxi [J]. *Acta Ecol Sin*, 2014, **34**(18): 5142 – 5152.
- [2] 郭艳荣, 吴保国, 刘洋, 等. 立地质量评价研究进展[J]. 世界林业研究, 2012, **25**(5): 47 – 52.
GUO Yanrong, WU Baoguo, LIU Yang, *et al.* Research progress of site quality evaluation [J]. *World For Res*, 2012, **25**(5): 47 – 52.
- [3] 王小明, 敖为赳, 陈利苏, 等. 基于 GIS 和 Logistic 模型的香榧生态适宜性评价[J]. 农业工程学报, 2010, **26**(增刊 1): 252 – 257.
WANG Xiaoming, AO Weijiu, CHEN Lisu, *et al.* Suitability assessment of Chinese torreyia based on GIS and Logistic regression [J]. *Trans Chin Soc Agric Eng*, 2010, **26**(suppl 1): 252 – 257.
- [4] KOO K A, MADDEN M, PATTEN B C. Projection of red spruce (*Picea rubens* Sargent) habitat suitability and distribution in the Southern Appalachian Mountains, USA [J]. *Ecol Model*, 2014, **293**: 91 – 101.
- [5] 胡秀, 吴福川, 郭微, 等. 基于 MaxEnt 生态学模型的檀香在中国的潜在种植区预测[J]. 林业科学, 2014, **50**(5): 27 – 33.
HU Xiu, WU Fuchuan, GUO Wei, *et al.* Identification of potential cultivation region for *Santalum album* in China by the MaxEnt ecologic niche model [J]. *Sci Sil Sin*, 2014, **50**(5): 27 – 33.
- [6] 高若楠, 苏喜友, 谢阳生, 等. 基于随机森林的杉木适生性预测研究[J]. 北京林业大学学报, 2017, **39**(12): 36 – 43.
GAO Ruonan, SU Xiyu, XIE Yangsheng, *et al.* Prediction of adaptability of *Cunninghamia lanceolata* based on random forest [J]. *J Beijing For Univ*, 2017, **39**(12): 36 – 43.
- [7] PIRI-SAHRAHAGARD H, AJORLO M, KARAMI P. Modeling habitat suitability of range plant species using random forest method in arid mountainous rangelands [J]. *J Mount Sci*, 2018, **15**(10): 2159 – 2171.
- [8] 孙玫, 张森, 聂培尧, 等. 基于朴素贝叶斯的网络查询日志 session 划分方法研究[J]. 南京大学学报(自然科学), 2018, **54**(6): 1132 – 1140.
SUN Mei, ZHANG Sen, NIE Peiyao, *et al.* Research on session segmentation of web search query logs based on naive Bayes [J]. *J Nanjing Univ Nat Sci*, 2018, **54**(6): 1132 – 1140.
- [9] 张广群, 李英杰, 汪杭军. 基于词袋模型的林木业务图像分类[J]. 浙江农林大学学报, 2017, **34**(5): 791 – 797.
ZHANG Guangqun, LI Yingjie, WANG Hangjun. Classification of forestry images based on the BoW Model [J]. *J Zhejiang A&F Univ*, 2017, **34**(5): 791 – 797.

- [10] 李欣海. 随机森林模型在分类与回归分析中的应用[J]. 应用昆虫学报, 2013, **50**(4): 1190 – 1197.
LI Xinhai. Using “random forest” for classification and regression [J]. *Chin J Appl Entomol*, 2013, **50**(4): 1190 – 1197.
- [11] 莫文希, 陈进宁, 唐璇. 高峰林场林下经济发展的思考与建议[J]. 林业经济, 2018(7): 106 – 110.
MO Wenxi, CHEN Jinning, TANG Xuan. Thoughts and suggestions on the development of under-forest economy in Gaofeng Forest Farm [J]. *For Econ*, 2018(7): 106 – 110.
- [12] CORTES C, VAPNIK V. Support-Vector networks [J]. *Mach Learn*, 1995, **20**(3): 273 – 297.
- [13] BREIMAN L. Random forests [J]. *Mach Learn*, 2001, **45**(1): 5 – 32.
- [14] 李婉华, 陈宏, 郭昆, 等. 基于随机森林算法的用电负荷预测研究[J]. 计算机工程与应用, 2016, **52**(23): 236 – 243.
LI Wanhua, CHEN Hong, GUO Kun, *et al.* Research on electrical load prediction based on random forest algorithm [J]. *Comput Eng Appl*, 2016, **52**(23): 236 – 243.
- [15] 陈少雄, 王观明, 修贵金, 等. 广西丘陵台地巨尾桉、尾叶桉(实生)地位指数表的编制[J]. 林业科学研究, 1995, **8**(2): 177 – 181.
CHEN Shaoxiong, WANG Guanming, XIU Guijin, *et al.* Site-index tables of *Eucalyptus grandis* × *E. urophylla* and *E. urophylla* in Guangxi hilly areas [J]. *For Res*, 1995, **8**(2): 177 – 181.
- [16] CHAWLA N V, BOWYER K W, HALL L O, *et al.* SMOTE: synthetic minority over-sampling technique [J]. *J Artif Intell Res*, 2002, **16**(1): 321 – 357.
- [17] 赖文豪, 席沁, 武海龙, 等. 内蒙古兴和县低山丘陵立地类型划分与林草适宜性评价[J]. 浙江农林大学学报, 2018, **35**(2): 331 – 339.
LAI Wenhao, XI Qin, WU Hailong, *et al.* Site classification type and vegetation suitability evaluation for hilly land in Xinghe, Inner Mongolia [J]. *J Zhejiang A&F Univ*, 2018, **35**(2): 331 – 339.
- [18] 罗美娟, 李宝福, 魏影景, 等. 闽南山地桉树生长与立地因子间的典型相关分析[J]. 福建林业科技, 2000, **27**(1): 14 – 17.
LUO Meijuan, LI Baofu, WEI Yingjing, *et al.* Typical correlation analysis between the growth and the site factor of *Eucalypti* in the mountain area of south Fujian [J]. *J Fujian For Sci Technol*, 2000, **27**(1): 14 – 17.
- [19] 韦红, 邢世和. 基于 GIS 技术的区域主要桉树树种用地适宜性评价[J]. 林业勘查设计, 2009(2): 61 – 63.
WEI Hong, XING Shihe. Evaluation on suitability of land resources for growing staple *Eucalyptus* based on GIS [J]. *For Invest Design*, 2009(2): 61 – 63.