

苹果属山荆子地理分布模拟

王雷宏¹ 杨俊仙² 郑玉红³ 汤庚国⁴

(1 安徽农业大学林学与园林学院 2 安徽农业大学理学院

3 江苏省中国科学院植物研究所,南京中山植物园 4 南京林业大学森林资源与环境学院)

摘要:依据苹果属最新的分类学研究成果,重新鉴定了中国科学院植物研究所等标本馆的山荆子标本,选取了358个有价值的分布记录转换成地理坐标数据,用DIVA-GIS绘制了现状地理分布,这种分布式样主要呈东北—西南方向延伸,西北—东南方向延伸较小,西南、中部、东北地区变种数量相当。利用全球气候栅格数据(分辨率arc 30 s),对落在世界气候变量包络内的分布点的19个气候变量进行了主成分分析,发现年降雨量、日均温差/年温差、暖季雨量是山荆子分布的主导因子,变种间的主导因子生态位有一定间隔或仅离群者重叠。用当前气候数据和未来气候数据对山荆子的当前、未来可能潜在分布进行了预测,结果显示:现状潜在分布区的一些高适区向东南方向迁移,未来气候下潜在分布向北迁移,而最北部的一些分布点消失,分布范围减小。

关键词:分布记录;GIS;潜在分布

中图分类号:S718.3;Q948.2 文献标志码:A 文章编号:1000-1522(2011)03-0070-05

WANG Lei-hong¹; YANG Jun-xian²; ZHENG Yu-hong³; TANG Geng-guo⁴. **Modelling the geographic distribution of *Malus baccata*.** *Journal of Beijing Forestry University* (2011) **33**(3) 70-74 [Ch 21 ref.]

1 School of Forestry & Landscape of Architecture, Anhui Agricultural University, 230036, P. R. China;

2 School of Science, Anhui Agricultural University, 230036, P. R. China;

3 Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing Botanical Garden, Mem. Sun Yat-Sen, 210014, P. R. China;

4 College of Forest Resources and Environment, Nanjing Forestry University, 210037, P. R. China.

Specimens of *Malus baccata* from some herbaria including the Institute of Botany, Chinese Academy of Sciences Herbarium (PE) were re-identified in light of the latest taxonomic outcome of the genus *Malus* Mill. and 358 valuable distribution records were selected and digitized into the geographic coordinate data. Based on these data, the updated geographic distribution is mapped using DIVA-GIS. This distribution presents a pattern extending mainly from the northeast to southwest, and secondarily from northwest to southeast. There are almost equal numbers of varieties in the southwest, central and northeast regions. Nineteen climate variables of the distribution points that fall in the envelope of variables of the world climate (spatial resolution arc 30 s) were analyzed by principal component analysis. Annual precipitation, isothermality, precipitation of the warmest quarter were found to be dominant factors affecting the geographic distribution of *M. baccata*; moreover, there are certain niche intervals of the dominant factors among the varieties, and niche overlaps only to some outliers. The present and future potential distribution maps are predicted by BIOCLIM with current and future climate conditions that are the world climate (spatial resolution arc 30 s). The results show that some highly suitable areas of the present potential distribution migrate southeastward, the whole future potential distribution extends northward, and some most northerly distribution points disappear; the overall range of distribution reduces.

Key words distribution records; geographic information system; potential distribution

收稿日期:2010-11-01

基金项目:安徽省自然科学基金项目(10040606Q18)、“948”国家林业局引进项目(98-04-08)。

第一作者:王雷宏,博士,讲师。主要研究方向:园林植物种质资源。电话:15155187659 Email: leihong_wang@yahoo.cn 地址:210036 安徽省合肥市长江西路130号安徽农业大学林学与园林学院。

责任作者:汤庚国,教授,博士生导师。主要研究方向:园林植物种质资源。电话:15345188711 Email: ggtang1950@yahoo.com.cn 地址:210037 江苏省南京市玄武区159号南京林业大学森林资源与环境学院。

本刊网址: <http://journal.bjfu.edu.cn>

物种的生态位和地理分布信息是研究其多样性空间分布格局的重要基础^[1],然而大多数物种都缺乏这种生态位和地理分布信息,明显影响科学发展和生产上的应用。物种分布模型就是试图通过已有的一些分布记录信息和地理气候信息来预测物种的地理分布,进而利用生态位模型研究环境参数与物种丰度的关系以及生态学、系统与进化、保护与利用中的多样性问题^[2]。现有模拟物种分布的方法较多,早期的模拟仅基于物种分布存在的信息,用环境变量包络法或者距离测度法进行预测^[3-4],后来把物种分布的存在或不存在都作为数据,分别考查两种数据相应的环境变量建模^[5-6]。如 BIOCLIM 模型主要依据预测点与已知分布点气候变量的一致性进行预测,采用生物气候变量包络法运算^[3],而最大熵模型 MAXENT 依据的生物原理是在无外力作用下生物种总是在约束下争取最大的自由度,运算时把数据分成两组,让机器学习训练数据,运算法则是从符合条件的分布中选择熵最大的分布^[7]。生物学者往往为选择哪一种分布模型而困惑,Elith 等^[2]采用 226 物种的分布数据对 16 个分布模型进行了比较研究,结果表明行之有效的模型是 Generalised Additive models, GARP, BIOCLIM,而机器学习模型和群落模型适用于分布记录信息较少或者分布信息易受各种干扰的物种。因此采用哪种模型,主要依据物种分布信息的数量和质量以及生物生态特性来决定^[8]。物种分布信息的电子化和网络化已经促成了在线物种分布模拟,国际 GBIF 搜集全世界的植物物种分布记录,已实现在线生态位模型模拟物种的地理分布^[9-10]。然而网站对一些物种的分布记录有限,甚至不区分野生或栽培分布,其在线模拟偏差仍较大^[11],所以对分布于亚洲或者中国的植物进行分布模拟仍然有较大的意义。曹铭昌等^[12]利用广义模型及分类回归树模拟了中国 15 个常见树种的分布。左闻韵等^[13]建立以支持向量机 (Support Vector Machine, SVM) 为核心的预测系统,预测杜鹃属植物在中国的潜在分布格局。王娟等^[8]用 BIOCLIM 模型预测了我国北方温带地区 5 种锦鸡儿的可能潜在分布区和未来分布区的变化趋势,并且研究了地理分布与环境变量之间的相关性,获得了科学的结论。

山荆子 (*Malus baccata* (L.) Borkh.) 是苹果属植物中分布范围最广的一个种,属东亚分布型,但集中分布于我国境内,也可认为是中国特有,其形态变异样式已得到了充分的研究^[14],现有的遗传多样性空间分布格局研究仍不充分^[15],地理分布与气候环境的关系也未见研究,用生态位模型模拟其地理

分布,可为遗传多样性研究的系统抽样作好准备,也可用以推测其地理分布与气候因子的关系。本文将其地理分布数据 (358 个分布点) 矢量化,采用 BIOCLIM 模型模拟其潜在自然分布,分析其地理分布与生物气候因子之间的关系,并预测未来气候变化对该植物分布区的影响,为今后该植物的多样性起源以及种质资源利用提供基础资料。

1 材料和方法

1.1 分布数据来源和处理

以苹果属分类学最新资料为标准^[16],对中国科学院植物研究所标本馆 (PE),江苏省中国科学院植物研究所标本馆 (NAS),南京林业大学标本馆 (NF),南京大学标本馆 (N) 的标本重新鉴定,获取分布地点。再查阅中国科学院植物研究所标本馆地名数据库,查找分布点的经纬度坐标,去除无价值的分布信息,共获得 358 个精确分布记录,用以作为分析的基础数据。

1.2 气象栅格数据和地理矢量数据

当前和未来气候环境数据都包括了 19 个环境因子变量, bio1 年均温, bio2 平均日温差, bio3 日均温差 (bio2) 除以年均温差 (bio7), bio4 温度季节变化方差, bio5 最热月温度, bio6 最冷月温度, bio7 温变年均差, bio8 雨季均温, bio9 干季均温, bio10 暖季均温, bio11 冷季均温, bio12 年降雨量, bio13 最湿月降雨量, bio14 最干月降雨量, bio15 季降雨量变异系数, bio16 雨季降水量, bio17 干季降水量, bio18 暖季降水量, bio19 冷季降水量。以上数据空间分辨率为 arc 30 s (相当于 1 km 空间分辨率),均来自 WORLDCLIM^[17] (<http://www.worldclim.org/>)。从 <http://nfgis.nsd.gov.cn/>, <http://www.diva-gis.org> 下载 1:400 万的中国、亚洲矢量地图作为分析底图。

1.3 研究方法

采用 DIVA-GIS (<http://www.diva-gis.org>) 软件, BIOCLIM 模型在该软件中自动实现^[18-20]。

2 结果与分析

2.1 地理分布格局

山荆子的现实分布格局见图 1,分布图主要呈东北-西南走向,东南和西北方向有一定的延伸,有少量的点扩散到蒙古、俄罗斯、朝鲜。对分布图层和地理图投影,设置地理尺度单位为 m,考查每单元格 (相当于 10 km²) 的变种丰富度,对应 4 种颜色,黄色 4 个变种,橙黄色 3 个变种,浅绿色 2 个变种,浓绿色 1 个变种,见图 2。从图 2 可以看出四川南部



图1 山荆子实际地理分布图

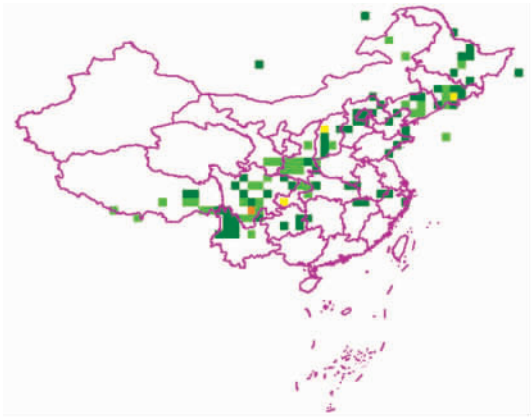
Fig.1 Actual geographic distribution maps of *M. baccata*

图2 山荆子变种丰富度

Fig.2 Varietal richness of *M. baccata*

注:黄色 橙黄色 浅绿 浓绿分别代表单位格子有4、3、2、1个变种。

(以 N 28.5°, E 102.5°) 和东南部 (以 N 29.4°, E 106.6°) 有 3~4 个变种,山西的西北部 (N 38.5°, E 111.6°) 有 3 个变种,吉林省长白山 (N 42.5°, E 127.5°) 也有 3 个变种,显然各变种的分布格局与地理位置也有关联。

2.2 分布与气候变量的关系

BIOCLIM 模型中主要是采用环境包络运算法则进行了模拟,利用落在环境因子变量包络内的有效分布点作为分析的对象,对其所在单元格环境变量做主成分分析,确定主导因子,本次有效点共 199 个,19 个主成分中的前 3 个主成分累计方差百分比达到了 85.347 9% > 85%,反映出占有效分布点的生

表1 前4个主成分方差解释

Tab.1 Variance of first four principal components explained

主成分	特征值	方差百分率/%	累计百分率/%
1	9.137 2	48.090 7	48.090 7
2	4.413 4	23.228 3	71.319 0
3	2.665 5	14.028 9	85.347 9
4	1.829 2	9.627 5	94.975 4

物气候变量信息量的 85.347 9% (见表 1);从表 2 可以看出第一主成分主要反映年降雨量 (bio12),最低气温 (bio6),干季的最低温 (bio9);第二主成分反映的是日均温差/年温差 (bio3),最高温 (bio5),暖季均温 (bio10),湿季均温 (bio8);第三主成分主要反映的是暖季雨量 (bio18),湿季雨量 (bio16),冷季雨量 (bio19)。可见年降雨量、日均温差/年温差、暖季雨量是影响其分布的主导因子或者限制因子。以年降雨量为例,比较变种单位的累积频数分布图,发现山荆子和毛山荆子 (*Malus baccata* var. *mandshurica*) 有生态位间隔 (图 3),在同一累积分布频数处,毛山荆子分布点处的降雨量明显较大,而对于离群点来说,两条曲线起末端有重合,离群点是俄罗斯布赫塔,内蒙古海拉尔西山,伊春市东北林业大学凉水林场,安徽黄山云海。

表2 相对于各变量的主成分得分系数

Tab.2 Principal components score coefficient related to each variable

气候变量	主成分 1	主成分 2	主成分 3
bio1	0.284 0	0.027 5	-0.297 3
bio2	-0.216 9	-0.195 9	0.075 8
bio3	0.096 7	-0.421 1	-0.050 4
bio4	-0.231 9	0.314 1	0.135 6
bio5	0.065 1	0.394 6	-0.251 9
bio6	0.292 3	-0.099 9	-0.247 2
bio7	-0.259 3	0.253 8	0.141 3
bio8	0.108 4	0.344 0	-0.246 1
bio9	0.291 3	-0.136 5	-0.222 1
bio10	0.125 6	0.369 1	-0.252 4
bio11	0.285 6	-0.144 9	-0.244 1
bio12	0.294 2	0.018 1	0.250 7
bio13	0.228 0	-0.032 7	0.301 0
bio14	0.224 5	0.223 5	0.250 3
bio15	-0.212 4	-0.138 2	-0.006 6
bio16	0.248 5	-0.070 8	0.269 8
bio17	0.231 1	0.202 6	0.257 6
bio18	0.234 4	-0.088 6	0.284 2
bio19	0.232 8	0.190 0	0.266 7

2.3 可能潜在分布或未来分布预测

对整个亚洲地区用 BIOCLIM 模型预测了现状可能潜在分布区 (图 4),未来可能潜在分布区 (图 5),设置分布概率 4 个等级,对应 4 种颜色,分别是浅绿色 2.5%~5%,黄色 5%~10%,桔黄 10%~20%,红色 20%~35%,颜色越深分布概率越大。比较实际分布图,现状可能潜在分布区,未来潜在分布预测图,发现现状可能潜在分布的一些高适分布点中心向东南方向迁移,未来潜在分布预测图的主体部分向北移动,然最北部的一些离群分布点减少或消失,整体来看,未来向北迁移,分布范围减小。

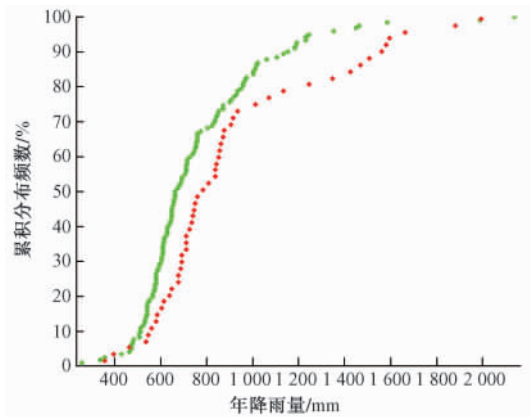


图 3 山荆子和毛山荆子的累积分布频数与年降雨量

Fig. 3 Cumulative distribution frequency of *M. baccata* and

M. baccata var. *mandshurica* with annual precipitation

注: 绿色. 山荆子; 红色. 毛山荆子。

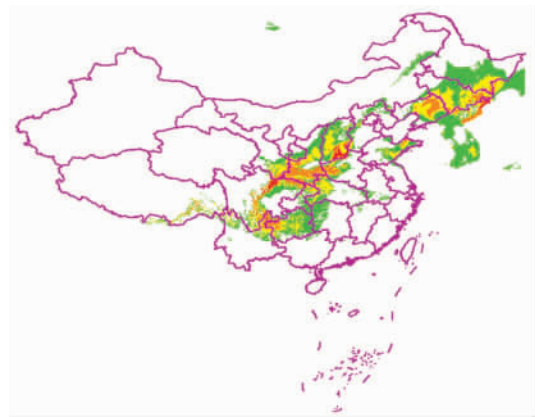


图 4 山荆子现状可能潜在分布

Fig. 4 Present potential distribution map of

M. baccata predicted by DIVA-GIS

注: 浅绿色. 2.5 ~ 5%; 黄色. 5 ~ 10%;

桔黄色. 10 ~ 20%; 红色. 20 ~ 35%。

3 结论与讨论

运用 GIS 技术绘制物种的地理分布图是植物地理分布研究的一个新手段, 准确性较高, 而传统绘制分布点图只是示意某个地区存在, 连线图也极易扩大植物的自然界分布情况。在分布图矢量化的基础上, 用生物气候模型研究影响分布的一些环境因子, 可能阐明物种分布的一些生态学规律, 本次研究发现了年降雨量、日温差/年均温差、暖季雨量是影响山荆子地理分布的主导因子, 而且发现变种间的主导因子生态位有一定的间隔, 变种的形成可能就是在如降雨量、日均温差/年温差、暖季雨量等气候因子的长期选择压力下形成的。

本次研究气候数据采用了最新的全球生物气候栅格数据 (分辨率 arc 30 s), 未采用中国地面气候数据标准值生成栅格数据, 主要原因在于一方面全球生物气候数据精确度已较高, 它本身包含了我国气象台站的数据; 另一方面, 山荆子有一些分布点邻近外国, 如采用中国的气候数据, 就无法准确模拟山荆子的潜在分布和未来可能潜在分布。从本次现状可能潜在分布的结果来看, 日本琉球岛、朝鲜半岛、印度西北部、不丹都可能分布。登录 GBIF (<http://data.gbif.org>) 查找记录信息, 发现在亚洲地区不丹国延布市有 1 个分布点, 韩国 5 个分布点, 但日本未有记录; 总体来看, 本次分布模拟较为准确。现状可能潜在分布图应该接近了山荆子真实地理分布。未来可能潜在分布表明山荆子整体的分布向北迁徙, 且整个分布区的面积有一定的收缩。

模拟物种潜在地理分布影响因素较多^[21], 影响较大的因素仍然是标本鉴定的正确性和分布记录的真实性^[11], 高分辨的环境变量数据也难以保证高海拔地区气候与物种实际分布地点的气候数据吻

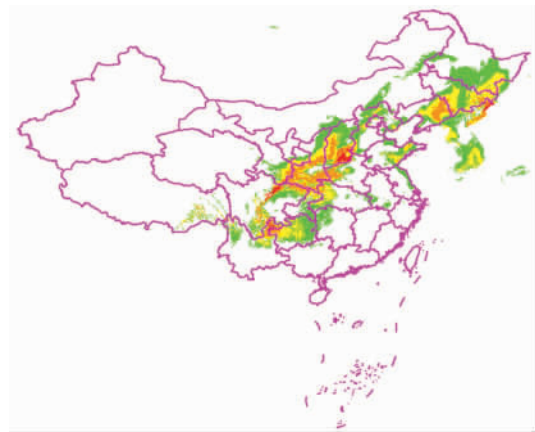


图 5 山荆子未来可能潜在分布

Fig. 5 Future potential distribution map of

M. baccata predicted by DIVA-GIS

注: 浅绿色. 2.5% ~ 5%; 黄色. 5% ~ 10%;

桔黄色. 10% ~ 20%; 红色. 20% ~ 35%。

合^[17], 此外物种本身有某种固有的扩展分布策略。本研究原始标本分布信息接近于系统抽样, 选择简单实用的 BIOCLIM 模型应该是有效的; 而且基于分布与环境生态位理论, 有利于研究植物和环境的关系, 能够较好地模拟预测物种的分布。虽然该模型不能揭示气候变量间的关系以及与物种分布的主导关系, 但从软件中提取有效分布点的气候变量数据, 所以利用这些数据再另外进行主成分分析, 就可以揭示影响物种地理分布的主导气候因子。本研究对环境变量包络内的有效分布点对应的气候变量进行了主成分分析, 揭示出了主导因子。此外, 本次研究未考虑土壤、植被群落的种间关系, 环境变量因子与分布的时滞效应以及随机扰动因子等因素的影响, 这是极其复杂的生物学、生态学过程和作用, 可能影响其地理分布格局及其与气候的关系^[21], 但

笔者认为对于广布种在地理区域尺度上其影响程度较少。

参 考 文 献

- [1] GRAHAM C H , MORITZ C , WILLIAMS S E. Habitat history improves prediction of biodiversity in rainforest fauna [J]. *Proceedings of the National Academy of Sciences of the United States of America* , 2006 , 103 (3) : 632-636.
- [2] ELITH J , GRAHAM C H , ANDERSON R P , *et al.* Novel methods improve prediction of species' distributions from occurrence data[J]. *Ecography* , 2006 , 29 (2) : 129-151.
- [3] BUSBY J R. BIOCLIM—a bioclimate analysis and prediction system [C] // MARGULES C R , AUSTIN M P. *Nature conservation: Cost effective biological surveys and data analysis.* CSIRO , 1991 : 64-68.
- [4] CARPENTER G , GILLISON A N , WINTER J. DOMAIN: A flexible modelling procedure for mapping potential distributions of plants and animals[J]. *Biodiv Conserv* , 1993 , 2 : 667-680.
- [5] PEARCE J , FERRIER S. Evaluating the predictive performance of habitat models developed using logistic regression [J]. *Ecol Model* , 2000 , 133 : 225-245.
- [6] PEARCE J , FERRIER S. An evaluation of alternative algorithms for fitting species distribution models using logistic regression [J]. *Ecol Model* , 2000 , 128 : 127-147.
- [7] PHILLIPS S J , DUDIK M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation [J]. *Ecography* , 2008 , 31 (2) : 161-175.
- [8] 王娟,倪健. 中国北方温带地区5种锦鸡儿植物的分布模拟[J]. *植物生态学报* , 2009 , 33 (1) : 12-24.
- [9] CARSTENS B C , RICHARDS C L. Integrating coalescent and ecological niche modeling in comparative phylogeography [J]. *Evolution* , 2007 , 61 (6) : 1439-1454.
- [10] RICHARDS C L , CARSTENS B C , KNOWLES L L. Distribution modelling and statistical phylogeography: An integrative framework for generating and testing alternative biogeographical hypotheses [J]. *Journal of Biogeography* , 2007 , 34 : 1833-1845.
- [11] 马克平. 生物多样性信息学研究进展 [J]. *生物多样性* , 2010 , 18 (5) : 429-443.
- [12] 曹铭昌,周广胜,翁恩生. 广义模型及分类回归树在物种分布模拟中的应用与比较 [J]. *生态学报* , 2005 , 25 (8) : 2031-2040.
- [13] 左闻韵. 应用 SVM 预测杜鹃属 (*Rhododendron* L.) 在中国的潜在分布中心 [D]. 北京:中国科学院植物研究所 2006.
- [14] 王雷宏. 山荆子变异式样研究 [D]. 南京:南京林业大学, 2008.
- [15] 王雷宏,郑玉红,汤庚国. 8 个山荆子居群遗传多样性的 ISSR 分析 [J]. *西北植物学报* , 2010 , 30 (7) : 1337-1343.
- [16] 钱关泽. 苹果属 (*Malus* Mill.) 分类学研究 [D]. 南京:南京林业大学 2005.
- [17] HIJMANS R J , CAMERON S E , PARRA J L , *et al.* Very high resolution interpolated climate surfaces for global land areas [J]. *International Journal of Climatology* , 2005 , 25 (15) : 1965-1978.
- [18] HIJMANS R J , SPOONER D M. Geographic distribution of wild potato species [J]. *American Journal of Botany* , 2001 , 88 (11) : 2101-2112.
- [19] JARVIS A , GUARINO L , WILLIAMS D *et al.* The use of GIS in the spatial analysis of wild peanut distributions and the implications for plant genetic resources conservation [J]. *Plant Genetic Resources Newsletter* , 2002 , 131 : 29-35.
- [20] JARVIS A , FERGUSON M E , WILLIAMS D E , *et al.* Biogeography of wild arachis: Assessing conservation status and setting future priorities [J]. *Crop Science* , 2003 , 43 (3) : 1100-1108.
- [21] 王运生,谢丙炎,万方浩,等. ROC 曲线分析在评价入侵物种分布模型中的应用 [J]. *生物多样性* , 2007 , 15 (4) : 365-372.

(责任编辑 赵 勃)