

# 基于 GAM 的中山市薇甘菊环境因子分析及适生区分布预测\*

高嵩<sup>1</sup> 马莹莹<sup>1</sup> 马红丽<sup>1</sup> 谢德志<sup>2</sup>

(1. 广州研美园林工程设计有限公司, 广东 广州 510520; 2. 广州碳汇林业有限公司, 广东 广州 510520)

**摘要** 为分析环境因子对薇甘菊 *Mikania micrantha* 分布的影响, 2020—2021 年间以中山市为研究区, 于薇甘菊盛花期进行实地踏查, 收集薇甘菊分布点数据, 定量分析 10 个环境因子对薇甘菊分布的影响, 基于 GAM 模型 (Generalized additive model) 对中山市薇甘菊适生区分布进行预测。结果显示, (1) 模型结果拟合精度高, TSS (Total sum of squares) 均值为 0.87, AUC (Area under the curve) 均值为 0.93; (2) 10 个环境因子对薇甘菊分布均有贡献, 贡献率最大的为降水量季节性变化 (18.63%), 其次为海拔 (17.90%), 第三为 4 月降水量 (16.47%); (3) 模型预测结果显示中山市约 89.23% 的地区适宜薇甘菊分布。研究构建的 GAM 模型拟合剪度高, 并证明了中山市区域尺度下水、热和海拔为影响薇甘菊分布的主导因子。

**关键词** GAM 模型; 薇甘菊; 分布预测

中图分类号: S765 文献标志码: A 文章编号: 2096-2053 (2024) 06-0122-06

DOI: 10.20221/j.cnki.2096-2053.202406018

## Analysis of Environmental Factors and Prediction of Suitable Distribution Areas of *Mikania micrantha* in Zhongshan City Based on GAM

GAO Song<sup>1</sup> MA Yingying<sup>1</sup> MA Hongli<sup>1</sup> XIE Dezhi<sup>2</sup>

(1. Guangzhou Yanmei Landscape Engineering Design Co., Ltd, Guangzhou, Guangdong 510520, China;

2. Guangzhou Carbon Sequestration Forestry Co., Ltd, Guangzhou, Guangdong 510520, China)

**Abstract** To analyze the effects of environmental factors on the distribution of *Mikania micrantha*, a field survey was conducted during the peak flowering period of *M. micrantha* in Zhongshan city from 2020 to 2021. Data on distribution points of *M. micrantha* were collected, and the effects of 10 environmental factors on its distribution were quantitatively analyzed. Based on the Generalized Additive Model (GAM), the distribution of suitable areas for *M. micrantha* in Zhongshan city was predicted. The results showed that: (1) the model had high fitting accuracy, with a total sum of squares (TSS) mean of 0.87 and an area under the curve (AUC) mean of 0.93; (2) Ten environmental factors contribute to the distribution of chamomile, with the highest contribution rate being seasonal variation in precipitation (18.63%), followed by altitude (17.90%), and the third being precipitation in April (16.47%); (3) The model prediction results show that about 89.23% of the areas in Zhongshan city are suitable for the distribution of *M. micrantha*. The GAM model constructed in the

\* 基金项目: 广东省林学会科技计划项目 (2021-GDFS-KJ-04)。

第一作者: 高嵩 (1987—), 男, 工程师, 主要从事林业有害生物防控研究, E-mail: 174172950@qq.com。

通信作者: 谢德志 (1993—), 男, 工程师, 主要从事外来入侵物种研究, E-mail: 252317552@qq.com。

study has high fitting accuracy and has been proven to be the dominant factor affecting the distribution of *M. micrantha* at the regional scale in Zhongshan city, including water, heat, and altitude.

**Key words** GAM model; *Mikania micrantha*; suitability distribution prediction

薇甘菊 *Mikania micrantha*, 也称小花蔓泽兰或小花假泽兰, 是菊科假泽兰属多年生草本植物或灌木状攀缘藤本<sup>[1]</sup>, 于 20 世纪 80 年代末传入我国广东沿海地区。因其生长速度快, 多缠绕附生于乔灌木植物, 直至压覆附主植物冠层顶部, 争光抢光资源导致附主植物死亡<sup>[2]</sup>, 对生物多样性和生态安全造成严重影响<sup>[3]</sup>。薇甘菊 2003 年被列入中国首批 16 种外来入侵生物之一, 2004 年被列入全国森林植物检疫对象名单, 2022 年被列入重点管理外来入侵物种名录。

薇甘菊适生区分布预测模型的研究, 可以为完善监测体系、预防其分布蔓延提供一定的科学依据<sup>[3]</sup>。有张海娟等<sup>[4]</sup> 基于生态位模型预测薇甘菊在中国的适生区分布; 付小勇等<sup>[5]</sup> 基于 Max-Ent 预测及评价了薇甘菊在云南省的分布; 陈志云等<sup>[6]</sup> 基于 MaxEnt 预测中山市的薇甘菊适生区。与生态位模型和最大熵 MaxEnt 模型对比, GAM 模型 (Generalized additive model) 是非参数化的函数模型, 不需要对函数形式进行假设, 具有对环境因素的高可解释性和分析性能。本研究通过实地调查获取中山市薇甘菊的分布点数据, 利用 SPSS (Statistical package for the social sciences) 筛选出 10 个低相关性的环境变量因子, 定量分析确定了对中山市薇甘菊分布贡献前 3 的主导因子, 运用 Biomod 2 软件包和 ArcGIS 等软件构建预测模型, 并选取 TSS (Total sum of squares) 及 AUC (Area under the curve) 对模型精度进行评价<sup>[7]</sup>, 对中山市薇甘菊监测、预防和除治具有参考意义。

## 1 材料与方法

### 1.1 研究区概况

研究区为中山市全境, 位于广东省中南部, 面积 1 800.14 km<sup>2</sup>, 均在北回归线以南, 地属亚热带季风气候, 光热充足, 雨量充沛, 年平均气温为 22.5 °C, 年平均降水量为 1 886 mm, 日照时长 1 705.4 h, 海拔最高 531 m, 地形主要是平原, 地貌由低山、丘陵、台地、平原和海滩组成, 植被代表类型为热带季雨林型的常绿季雨林, 主要森林类型有常绿阔叶林、常绿针叶林、常绿阔叶混

交林和经济林。

### 1.2 数据收集

1.2.1 薇甘菊分布点数据 于 2020 和 2021 年薇甘菊盛花期, 对研究区农田、水源、林地和道路周边进行实地踏查, 收集薇甘菊发生地分布点位置、面积、群落类型、地形和水分条件等, 共 950 个。

1.2.2 气候数据 包含年平均气温、最高最低气温、年降水量、最高最低降水量等 25 个变量, 数据来源于 1950—2020 年全世界气象台站, 下载自 WorldClim (<http://www.worldclim.org/>), 空间分辨率为 30 arc-seconds。

1.2.3 地形数据 包括海拔、坡度和坡向 3 个变量, 数据来源于 1950—2020 年全世界气象台站, 下载自 WorldClim 数据库 (<http://www.worldclim.org/>), 空间分辨率与气候数据一致。坡度和坡向因子数据, 利用 ArcGIS 中的空间分析模块 (Spatial Analyst), 从海拔数据中提取得到。

1.2.4 土壤数据 包括土壤类型、土壤 pH 值、土壤有机质含量、土壤容重、土壤碳酸钙含量、土壤砂含量、粘土含量、淤泥含量、碎石体积百分比、土壤的阳离子交换能力、可交换钠盐和电导率, 下载自 2023 年世界土壤数据库 (<http://www.fao.org/soils-portal/soil-survey>), 空间分辨率与气候变量数据一致, 共 12 个变量。

1.2.5 植被数据 包括植被类型、郁闭度/覆盖度和归一化植被指数, 空间分辨率与气候变量数据一致, 下载自国家地球系统科学数据中心 (<http://www.geodata.cn/main/>), 共 3 个变量, 为 2001 年植被类型、2011 年郁闭度指数、2023 年归一化指数。

1.2.6 人类活动变量数据 包括人类活动足迹、人口密度和土地利用类型, 其中人类活动足迹数据源于社会经济数据与应用中心 (<https://earthdata.nasa.gov/centers/sedac-daac>) 2009 年数据、人口密度和土地利用类型数据 (1 : 1 000 000) 源于中国科学院资源环境科学与数据中心 (<https://www.resdc.cn/>) 2020 年数据, 空间分辨率与气候变量数据一致, 共 3 个变量。

### 1.3 数据预处理

1.3.1 薇甘菊分布数据预处理 将薇甘菊分布点位置信息导入 ArcGIS 软件, 利用缓冲区分析法进行校对、筛选, 缓冲半径取 1.5 km, 任意 2 个分布点距离小于 3 km 时只保留 1 点<sup>[8]</sup>, 排除因距离近造成空间关联性较大的导致过拟合模拟的分布点, 最终获得 101 个分布数据。

1.3.2 变量数据筛选及预处理 为科学分析环境变

量对薇甘菊分布的影响, 消除环境变量间的数据冗余和多重共线性问题<sup>[9]</sup>, 利用 SPSS26.0 软件进行相关性分析, 从 5 种数据集 (气候、地形、土壤、植被和人类活动) 共 46 个变量中, 筛选出  $|R| < 0.8$  的 10 个环境变量 (表 1) 用于建模分析。

利用 ArcGIS 软件的掩膜工具, 以中山市行政区划范围为掩膜, 裁剪气候因子、地形因子、土壤因子、植被因子和人类活动因子等栅格数据。

表 1 中山市薇甘菊分布影响变量因子

Table 1 Variable factors affecting the distribution of *Mikania micrantha* in Zhongshan city

地理环境变量 Geographic environmental variables	变量因子 Variable factor	含义 Meaning
气候变量 Climate variables	bio_ 3 (Isothermality)	等温性
	bio_ 11 (Mean temperature of coldest quarter)	最冷季平均温度
	bio_ 15 (Precipitation seasonality)	降水量的季节性变化
	bio_ prec04	4 月降水量
地形变量 Terrain variables	bio_ prec12	12 月降水量
	bio_ elev	海拔
	bio_ slope	坡度
土壤变量 Soil variables	bio_ ph	土壤 pH 值
植被变量 Vegetation variables	bio_ ndvi	归一化植被指数
人类活动变量 Human activity variables	bio_ pdensity	人口密度

### 1.4 模型构建

1.4.1 Biomod 2 软件包 Biomod 2 软件包是基于 R 语言平台的一个免费、开源的程序包, 是一种新型的计算框架, 能处理模型中物种与环境之间的关系, 并降低不确定性, 最终预测物种的分布。Biomod 2 的功能有: 模拟物种分布、模型测试、预测不同环境条件下物种的分布及扩散 (例如气候变化、土地覆被变化及土地利用变化等)。在运行之后, 可得到的结果有: 物种分布, 环境变量响应曲线, 模型运行效果以及环境变量与物种之间的相互关系, 能够准确预测物种的潜在分布<sup>[10-11]</sup>。

1.4.2 GAM 模型 GAM 模型是在回归模型的基础上, 对每个因变量采用平滑函数进行变换, 采用加和操作, 构造趋近于训练数据集的的曲面, 以此来降低线性关系设定引起的僵化的模型结构。其公式如下<sup>[12]</sup>:

$$g(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_i(x_i)$$

其中  $g$  为关联函数 (Link function),  $Y$  为自变量,  $\beta_0$  是截距项,  $f_i(x_i)$  为光滑样条函数, 由公式可以看出, 采用光滑样条函数进行因变量变换已达到局部优化的效果, 所以其残差小、精度高, 在因变量和自变量关系复杂时候具有优势。在物

种分布建模实践中, 物种的响应曲面往往是不对称的, GAM 模型则相对更加灵活, 可以处理更加复杂的响应曲面。同时, GAM 根据数据特征驱动模型, 也就是说数据的结构决定模型拟合方式, 因此其建模更为灵活, 实用性更高<sup>[13]</sup>。

1.4.3 运行设置 将预处理后的薇甘菊分布点数据先平分为 3 组, 再把每组 75% 的数据作为训练集、25% 的数据作为测试集, 输入 GAM 模型以相同的模式运行 10 次。最后, 用 TSS 值<sup>[14]</sup> 和 AUC 值对模型统计学精度进行评价, 以评估预测结果的准确性。TSS 值和 AUC 值区间为 0~1, 值越大表示模型预测准确性越高、效果越好<sup>[4,14]</sup>。Biomod 2 软件包输出结果为薇甘菊在中山市的存在概率, 结果值在 0~1 之间, 值越接近 1 表示物种越可能存在。

### 1.5 适宜分布区划分及制图

1.5.1 适宜分布区划分 利用 IPCC (Intergovernmental panel on climate change) 评估可能性的划分方法<sup>[13]</sup>, 结合薇甘菊分布点数据, 将 Biomod 2 软件包输出结果分为 4 个等级, 即高适生区、中适生区、低适生区和非适生区。

1.5.2 制图 将 Biomod 2 软件包输出结果导入 Arc-GIS 软件, 利用重分类工具、密度间隔方法,

用4种颜色表示4个等级分布, 编制中山市的薇甘菊适生区等级分布图。

## 2 结果与分析

### 2.1 模型精度检验

模型运行结果显示 TSS 均值为 0.87, 最大值为 0.88, 最小值为 0.85 (图 1)。AUC 均值为 0.93, 最大值为 0.94, 最小值为 0.92。

### 2.2 环境变量的贡献率

依据 GAM 模型运算结果, 得到不同环境变量对薇甘菊在中山市分布的贡献百分率和累计贡献百分率 (表 2)。结果显示, 10 个环境变量都有贡献率, 贡献率最大的为降水量季节性变化 (18.63%), 其次为海

拔 (17.90%), 第三为 4 月降水量 (16.47%)。

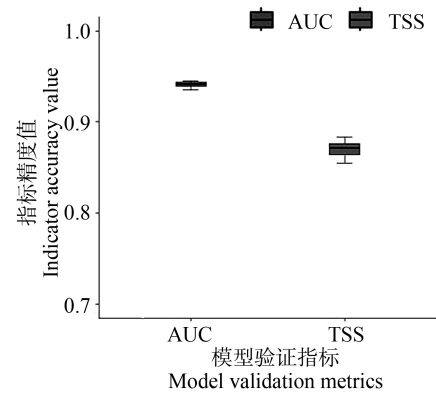


图 1 中山市薇甘菊分布预测模型精度  
Fig. 1 Accuracy of the distribution prediction model of *Mikania micrantha* in Zhongshan city

表 2 影响薇甘菊分布的环境变量贡献百分率和累积贡献百分率

Table 2 Accumulated contribution percentage and cumulative contribution percentage of environmental variables affecting the distribution of *Mikania micrantha*

环境变量 Environment variable	贡献百分率/% Contribution percentage	累积贡献百分率/% Cumulative contribution percentage
bio_15	18.63	18.63
bio_elev	17.90	36.52
bio_prec04	16.47	52.99
bio_3	10.61	63.60
bio_ndvi	9.45	73.04
bio_pdensity	7.96	81.00
bio_slope	8.17	89.17
bio_11	6.01	95.18
bio_ph	3.84	99.01
bio_prec12	0.99	100

### 2.3 影响薇甘菊在中山市分布的主导因子

依据环境变量对薇甘菊在中山市分布的贡献百分率, 确定了对薇甘菊在中山市分布影响较大的 3 个环境变量依次为降水量的季节性变化 (bio\_

15)、海拔 (bio\_elev)、4 月降水量 (bio\_prec04)。主导因子影响曲线图中横坐标代表的是环境因子变量值, 纵坐标代表的是薇甘菊发生概率 (图 2)。

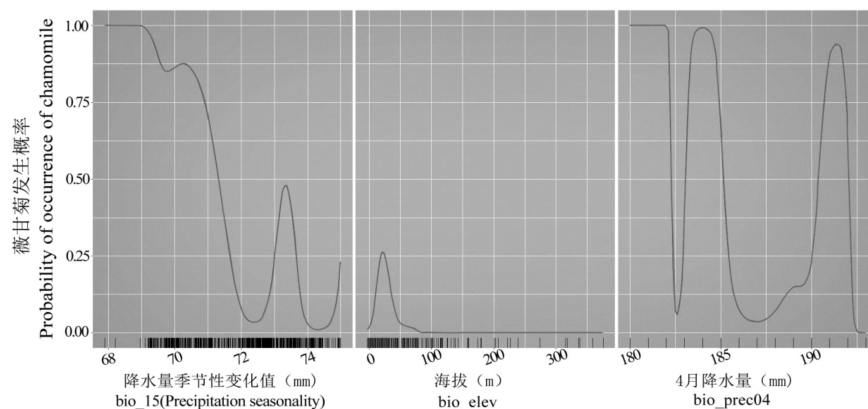


图 2 中山市薇甘菊分布主导因子及发生概率

Fig. 2 Dominant distribution factors and occurrence probability curve of *Mikania micrantha* in Zhongshan city

曲线图表明, 薇甘菊发生概率与环境因子变量值存在很强的响应关系。68.0 $\leq$ 降水量季节性变化 (mm) <71.1 时, 薇甘菊发生概率为 [0.66, 1]; 71.1 $\leq$ 降水量季节性变化 (mm) <71.5 和 73.1 $\leq$ 降水量季节性变化 (mm) <73.6 时, 薇甘菊发生概率为 [0.33, 0.66]; 71.5 $\leq$ 降水量季节性变化 (mm) <73.1 和 73.6 $\leq$ 降水量季节性变化 (mm) <75.0 时, 薇甘菊发生概率为 [0.05, 0.33)。

0 $\leq$ 海拔 (m) <75 时, 薇甘菊发生概率为 [0.05, 0.33); 海拔 (m)  $\geq$ 75 时, 薇甘菊发生概率 [0, 0.05)。

180.0 $\leq$ 4月降水量 (mm) <182.2、182.9 $\leq$ 4月降水量 (mm) <184.9 和 190.5 $\leq$ 4月降水量 (mm) <192.2 时, 薇甘菊发生概率为 [0.66, 1.0]; 182.2 $\leq$ 4月降水量 (mm) <182.4、184.9 $\leq$ 4月降水量 (mm) <185.2 和 190.1 $\leq$ 4月降水量 (mm) <190.5 和 192.2 $\leq$ 4月降水量 (mm) <192.3 时, 薇甘菊发生概率为 [0.33, 0.66); 182.4 $\leq$ 4月降水量 (mm) <182.9、185.2 $\leq$ 4月降水量 (mm) <190.2 和 190.5 $\leq$ 4月降水量 (mm) <192.2 时, 薇甘菊发生概率为 [0.05, 0.33); 4月降水量 (mm)  $\geq$ 192.4 时, 薇甘菊发生概率为 [0, 0.05)。

## 2.4 薇甘菊在中山市适宜性分布预测

2.4.1 适宜分布区划分 结合薇甘菊分布点数据, 将 Biomod 2 软件包输出结果分为高适生区、中适生区、低适生区和非适生区。划分标准为: [0, 0.05) 为非适生区, [0.05, 0.33) 为低适生区, [0.33, 0.66) 为中适生区, [0.66, 1] 为高适生区。

2.4.2 制图 中山市的薇甘菊适生区等级分布见图3, 计算各等级区域面积及占比, 得到非适生区面积约 193.88 km<sup>2</sup>, 约占 10.77%。这说明, 中山市约 1 606.26 km<sup>2</sup> 对薇甘菊的分布具有不同程度的适宜性, 约占 89.23%。其中, 高适生区面积约 1 039.13 km<sup>2</sup>, 约占 57.73%, 分布于中山市所有镇 (街道); 中适生区面积约 296.25 km<sup>2</sup>, 约占 16.46%; 低适生区面积约 270.88 km<sup>2</sup>, 约占 15.05%。综上所述, 中山市大部分地区适合薇甘菊的生长, 且高适生区面积大、分布广。

## 3 结论与讨论

通过模型精度检验, 得出 GAM 模型预测中山

## GAM Model

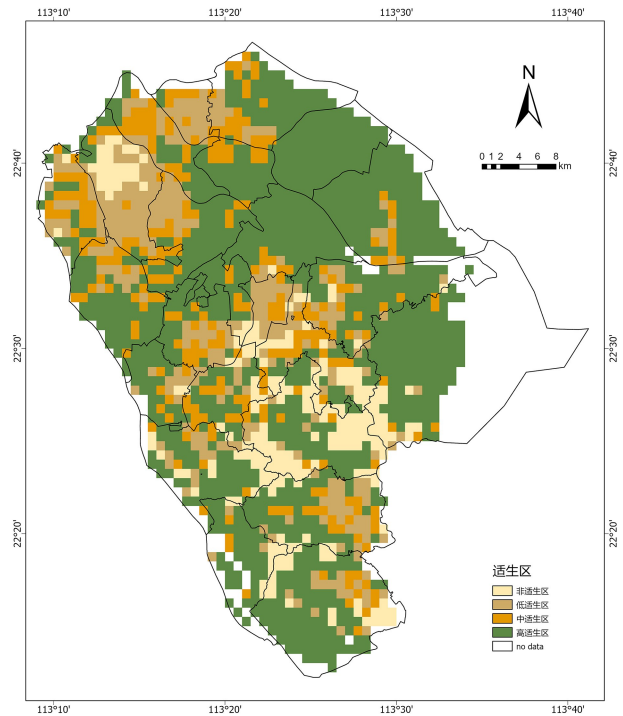


图3 中山市薇甘菊适生区等级分布  
Fig. 3 Grade distribution of *Mikania micrantha* suitable in Zhongshan city

市薇甘菊适生区的 AUC 均值为 0.93 (最大值为 0.94、最小值为 0.92), 高于陈志云等<sup>[6]</sup> MaxEnt 模型的训练集 AUC 均值 0.808 和测试集 AUC 均值 0.873, 验证了 GAM 模型残差小、精度高的优势, 在物种分布预测模型中更有优势。

在进行 GAM 模型构建时, 为消除环境变量间的数据冗余和多重共线性问题, 同时简化预测模型, 提高模型预测准确度, 因此本研究从 5 个数据集 46 个环境变量中选取低相关性的 10 个变量进行分析, 得出降水量的季节性变化 (bio\_15)、海拔 (bio\_elev)、4 月降水量 (bio\_prec04) 为本次研究的主导因子, 累计贡献率为 52.99%。进一步证明在区域尺度范围内制约物种分布范围的气候因子, 水热条件和海拔具有主导作用<sup>[16]</sup>。

本研究虽进行了数据集之间环境变量的相关性分析筛选, 但无法完全消除环境变量间的关联影响, 比如 4 月降水量可以影响春季土壤水分含量及 pH 值, 直接影响到薇甘菊在春季的萌芽<sup>[17]</sup>。另外, 受研究尺度局限性影响, 环境变量主导因子变化幅度较小<sup>[18]</sup>, 且未考虑无薇甘菊分布点数据<sup>[19]</sup> 和人为活动防治因子等变量, 但不代表其对

预测结果无影响。

本研究基于 GAM 模型预测中山市薇甘菊适生区分布, 通过筛选分布点数据和环境变量数据, 定量分析相关环境因子贡献百分率, 确定了影响中山市薇甘菊适生区分布主导因子, 预测了中山市各适生区和非适生区分布, 与中山市薇甘菊实际分布拟合精度高, 预测结果可信度高, 对中山市完善薇甘菊监测体系、预防薇甘菊分布蔓延工作提供科学依据。

### 参考文献

- [1] 孔国辉, 吴七根, 胡启明, 等. 薇甘菊 (*Mikania micrantha* H. B. K.) 的形态, 分类与生态资料补记 [J]. 2000, 8 (2): 128-130.
- [2] 李国瑞, 李宏德, 惠安萍. 镇康县薇甘菊发生现状及防控对策 [J]. 农业科技通讯, 2022 (4): 217-219.
- [3] 许少嫦, 高亿波, 林绪平, 等. 广东薇甘菊发生现状, 防治和研究进展 [J]. 广东林业科技, 2013, 29 (4): 83-89.
- [4] 张海娟, 陈勇, 黄烈健, 等. 基于生态位模型的薇甘菊在中国适生区的预测 [J]. 农业工程学报, 2011, 27 (S1): 413-418, 420.
- [5] 付小勇, 泽桑梓, 周晓, 等. 基于 MaxEnt 的云南省薇甘菊分布预测及评价 [J]. 广东农业科学, 2015, 42 (12): 159-162, 2.
- [6] 陈志云, 马红岩, 王玲, 等. 广东中山市薇甘菊地理分布及适生区预测 [J]. 亚热带植物科学, 2021, 50 (3): 222-226.
- [7] WISZ M, HIJMANS R J, LI J, et al. Effects of sample size on the performance of species distribution models [J]. Diversity and distributions, 2008, 14 (5): 763-773.
- [8] 王茹琳, 李庆, 封传红, 等. 基于 MaxEnt 的西藏飞蝗在中国的适生区预测 [J]. 生态学报, 2017, 37 (24): 8556-8566.
- [9] 黄君玥, 朱报著, 张卫华, 等. 基于 MaxEnt 和 ArcGIS 的美丽异木棉适生区与主导环境因子分析 [J]. 林业与环境科学, 2024, 40 (1): 1-8.
- [10] PEARSON R G, DAWSON T P, BERRY P M, et al. SPECIES: A Spatial Evaluation of Climate Impact on the Envelope of Species [J]. Ecological modelling, 2002, 154 (3): 289-300.
- [11] THUILLER W, LAFOURCADE B, ENGLER R, et al. BIOMOD: a platform for ensemble forecasting of species distributions [J]. Ecography, 2009, 32 (3): 369-373.
- [12] HASTIE T J, TIBSHIRANI R J. Generalized Additive Models [M]. Routledge, 2017: 136-173.
- [13] 赵泽芳. 气候变化下物种分布模型建构与模型比较: 以羌活为例 [D]. 西安: 陕西师范大学, 2018.
- [14] 桑满杰. 环境变量对物种分布模型的影响与评价: 以山茶萸潜在地理分布预测为例 [D]. 西安: 陕西师范大学, 2015.
- [15] 段居琦, 周广胜. 中国水稻潜在分布及其气候特征 [J]. 生态学报, 2011, 31 (22): 6659-6668.
- [16] 李晓辰, 贡璐, 魏博, 等. 气候变化对新疆雪岭云杉潜在适宜分布及生态位分化的影响 [J]. 生态学报, 2022, 42 (10): 4091-4100.
- [17] 李小平. 土壤物理学家: 邵明安院士 [J]. 西北农林科技大学学报 (自然科学版), 2024, 52 (1): 7.
- [18] 吴卉晶, 咎启杰, 曾辉. 区域尺度薇甘菊入侵分布的影响因子 [J]. 生态学报, 2009, 29 (10): 5442-5449.
- [19] CHEFAOUI M R, HORTAL J, LOBO M J. Potential distribution modelling, niche characterization and conservation status assessment using GIS tools: a case study of Iberian Copris species [J]. Biological conservation, 2005, 122 (2): 327-338.