

李宏群,蒲永兰,刘晓莉,等. 基于 GARP 模型对松材线虫在重庆适生区的预测[J]. 江苏农业科学,2017,45(24):82-84.
doi:10.15889/j.issn.1002-1302.2017.24.020

基于 GARP 模型对松材线虫在重庆适生区的预测

李宏群¹,蒲永兰²,刘晓莉³,刘晓梅¹,杨清钰²,丁世敏⁴

(1. 长江师范学院生命科学与技术学院,重庆 408100; 2. 重庆市林业科学研究院森林资源保护与林业信息技术研究所,重庆 400036;
3. 长江师范学院图书馆,重庆 408100; 4. 长江师范学院武陵山片区绿色发展协同创新中心,重庆 408100)

摘要:松材线虫(*Bursaphelenchus xylophilus*)原产于北美洲,是松材线虫病的主要病原生物。依据松材线虫的地理分布数据以及环境因子数据,采用预设预测规则的遗传算法(genetic algorithm for rule-set production,简称 GARP)模型和 ArcGIS 软件进行建模并对松材线虫病在重庆的潜在分布区进行预测。结果表明,年平均温度、月均值昼夜温差、等温性、最冷月最低温度、气温年变化范围、最干旱季度平均温度、最热季度平均温度、年降水量、最湿润月降水量、最湿润季度降水量、最干旱季度降水量和最冷季度降水量对松材线虫病的分布影响显著。松材线虫的潜在入侵区主要集中在中部,即三峡库区腹地,然后向重庆东北、东南以及南方向传播。因此,在最佳和高度适生区应该加强监测及预警。

关键词:松材线虫;潜在生境;GARP 模型;重庆;入侵;生物多样性

中图分类号: S763.16 **文献标志码:** A **文章编号:** 1002-1302(2017)24-0082-03

松材线虫病是由外来入侵种松材线虫(*Bursaphelenchus xylophilus*)引起的,目前被列为我国第一大林业外来有害生物,被称作松树“癌症”“松树艾滋病”“松树禽流感”,已给发生地区造成巨大的经济、生态和社会损失。我国自 1982 年在江苏省南京市中山陵的黑松上首次发现该病以来,短短 30 多年,疫情已经迅速扩展到我国 14 个省(市、区),192 个县(市、区),累计致死松树 5 亿多株,毁灭松林 33 万 hm^2 ,严重影响了疫区经济和社会的可持续发展^[1-2]。目前松材线虫病对庐山、黄山和三峡库区等的生态安全构成了严重威胁,且到目前为止没有十分有效的防治方法,形势十分严峻。

物种潜在分布区预测模型在濒危物种保护、生物多样性保护区设计的优先性评估、外来入侵物种扩散潜能以及全球气候变化对物种分布区的影响等研究中得到越来越广泛的应用^[3]。松材线虫的入侵很难被控制,消耗大量的人力、物力和财力,以致预防是比防治更经济的行为。在过去,国内外对松材线虫治病机制、发生现状以及防治措施的研究较多^[4-6],而对于其潜在生境预测的研究相对较少^[1,7-8]。三峡库区重庆段还没有人研究,且目前重庆的万州区、涪陵区、长寿区、巴南区、云阳县和忠县已经被确定是松材线虫病的高发疫区,已表现出扩大蔓延的趋势,引起了我国学者的重视。用来预测物种潜在分布区的预设预测规则的遗传算法(genetic algorithm for rule-set production,简称 GARP)生态位模型最初由 Stockwell 创建,是预设规则的遗传算法,利用物种的已

知分布数据和环境数据模拟物种的基本生态位需求,探索物种已知分布区的环境特征与研究区域的非随机关系^[9-11]。GARP 模型应用时间较早,是目前最准确的物种潜在分布区预测模型之一,应用较为广泛^[9-11]。因此,本研究以松材线虫在重庆地区的分布数据为基础,运用 GARP 生态模型软件,预测其潜在入侵区域,为制定合理的防治措施提供科学依据。

1 材料与方法

1.1 材料

1.1.1 松材线虫病分布点数据来源 依据重庆市涪陵区林业局病虫害防治站提供的松材线虫发生地分布数据,采用现场观测,手持 GPS 定位的方式,提取发生地的经纬度坐标,核对位置后得到 63 个有效分布点,根据 GARP 软件要求,将分布点经度和纬度储存成 .csv 格式的文件。

1.1.2 软件工具 GARP 软件:从 <http://www.nhm.ku.edu/desktopgarp/> 免费下载,版本为 1.1.6;GIS 软件由中国科学院重庆绿色智能技术研究院生态过程与重建研究中心提供,版本为 ArcGIS 9.3。

1.1.3 基础地理数据 从国家基础地理信息系统(<http://Mnfgis.nsd.gov.cn/>)下载获得 1:400 万的我国国界和省界以及县界行政区划图,作为分析的底图。

1.1.4 环境数据 本研究选取影响松材线虫病分布的 19 个生物气候变量作为主要环境数据(表 1)。数据来自 Berkeley 大学 Worldclim 下载中心国际网站(<http://www.worldclim.org/>)免费下载空间,分辨率为 30" 当前状态(current conditions 1950—2000)19 个生物气候变量,所有层数据的投影格式均转换为 GCS-WGS-1984,然后通过重庆市的地图对所获得数据进行掩膜获得三峡库区重庆段数据,此过程在 ArcGIS 9.3 软件中进行,并把所有环境数据转换为 ASCII 格式。

1.2 模型模拟

在使用该 GARP 模型时,每次运算随机选取 50% 的数据

收稿日期:2016-06-17

基金项目:教育部春晖计划(编号:Z2015139、Z2015129);重庆市大学生创新创业训练计划(编号:201410647002);长江师范学院科研创新团队建设计划(编号:2014XJTD02、2014XJTD06);重庆高校优秀成果转化资助项目(编号:KJZH17132);重庆市林业重点科技攻关(编号:2016-3)。

作者简介:李宏群(1973—),男,陕西西安人,博士,教授,主要从事动物生态与病虫害防治研究。E-mail:501680655@qq.com。

用于建立模型,剩下 50% 的数据用于验证模型误差。创建模型规则时反复运算 1 000 次或其精度达到 0.01 而终止。在选取最优的环境因子时,首先将 19 个环境因子一起进行 1 000 次模型运算,将所获得的溢出误差 95% 置信区间的算术平均值作为基础溢出误差。然后使用刀切法(jackknife)依次省略每一个环境因子,创建 19 组不同的环境因子组。分析每一组环境因子与溢出误差(omission error)之间的关系,如果一个环境因子的缺失导致溢出误差显著降低(低于基础溢出误差 $\times 95\%$),则该环境因子将不包含在最终的分析中。反之,如果一个环境因子的缺失导致溢出误差高于基础溢出误差,则认为这一因子对预测结果影响显著^[12]。

在最终分析中,重复创建 1 000 次模型,创建模型规则时重复运算次数最大值(max iteration)设为 5 000,收敛极限(convergence limit)为 0.01,规则类型(rule types)使用默认设置,并生成 arc 格式的图层。运算完毕后,选取所有内部遗漏误差和外部遗漏误差为 0 的预测中溢出误差最小的 50 个模型,即为最优模型。利用 ArcGIS 的空间扩展模块(spatial analyst tools),将最优模型的预测结果进行等权重叠加。叠加的结果为每一个栅格的数值为 0~50,0 表示 50 个预测模型都预测该栅格不适合松材线虫繁殖,而 50 表示 50 个预测模型都预测该栅格适合松材线虫繁殖。同时,把每一个栅格的值除以 50 来表示该区域的适宜性,称之为重叠系数(overlap index,简称 OI)。重叠系数数值越大,表示该栅格的环境参数越适合物种的生存。因此,根据重叠系数的大小可以将预测的

区域划分为:最佳适生区($0.9 < OI \leq 1.0$)、高度适生区($0.7 < OI \leq 0.9$)、中度适生区($0.5 < OI \leq 0.7$)和非适生区($OI \leq 0.5$)。

2 结果与分析

2.1 环境因子选择及模型选择

19 个环境因子共同进行模型运算所获得的基础溢出误差为 28.530 7。选取最优模型时,依次去除 19 个环境因子所获得溢出误差的平均值如表 1 所示。其中温度季节性变化、最热月最高温度、最湿润季度平均温度、最冷季度平均温度、最干旱月降水量、降水量季节变化和最热季度降水量被去除时,溢出误差明显下降($< 28.530 7 \times 0.95 = 27.104 2$),而去除年平均温度、月均值昼夜温差、等温性、最冷月最低温度、气温年变化范围、最干旱季度平均温度、最热季度平均温度、年降水量、最湿润月降水量、最湿润季度降水量、最干旱季度降水量和最冷季度降水量则会使溢出误差明显升高。因此,温度季节性变化、最热月最高温度、最湿润季度平均温度、最冷季度平均温度、最干旱月降水量、降水量季节变化和最热季度降水量在最终的分析中被剔除,使用其他 12 个环境因子建立 1 000 个模型,并选取所有内部遗漏误差和外部遗漏误差为 0 的预测中 72 个模型,即为最优模型。以独立外部检验数据集为基础的卡方检验显示这 72 个模型在统计上都高度显著($P < 0.01$),即模型的预测能力远远好于随机模型。

表 1 生态位模型参数

环境因子	溢出误差	状态	环境因子	溢出误差	状态
年平均温度	29.562 5	+	最冷季度平均温度	26.789 8	-
月均值昼夜温差	30.179 0	+	年降水量	30.177 1	+
等温性	29.484 5	+	最湿润月降水量	29.367 2	+
温度季节性变化	26.671 3	-	最干旱月降水量	28.267 3	-
最热月最高温度	27.024 5	-	降水量季节变化	27.052 2	-
最冷月最低温度	30.043 4	+	最湿润季度降水量	31.646 1	+
气温年变化范围	29.306 6	+	最干旱季度降水量	28.985 8	+
最湿润季度平均温度	26.466 1	-	最热季度降水量	25.572 2	-
最干旱季度平均温度	29.813 6	+	最冷季度降水量	29.362 8	+
最热季度平均温度	28.601 0	+			

注:“+”、“-”分别表示该环境因子包含、不包含在最终的分析中。

2.2 潜在分布区的预测

把这 72 个模型作为最优的模型,利用 ArcGIS 的空间扩展模块,将从最优模型中随机选取的 50 个预测结果进行等权重叠加。模型预测图与我国重庆行政区图叠加的结果如图 1 所示,红、橙、绿、白 4 种颜色依次表示潜在分布的可能性由高到低。由图 1 可看出,松材线虫病适生区主要集中在重庆中部,即三峡库区腹地,然后向重庆东北、东南以及南方向传播。具体分布为:最佳适生区有涪陵区、长寿区、丰都县、垫江县、忠县、梁平区和万州区;高度适生区有南川区、武隆区、万盛区、綦江区、南岸区、巴南区、江北区、渝北区、彭水县、石柱县、奉节县、开县和云阳县;中度适生区有巫山县、巫溪县、酉阳县、黔江区、江津区、永川区、永昌县;非适生区有沙坪坝区、北碚区、合川区、九龙坡区、大渡口区、大足县、璧山县、铜梁县、潼南县、秀山县等。

3 讨论

3.1 环境因子分析

在筛选用于入侵生物的最优环境因子集合时,通常会采用一系列的刀切法分析,通过多次重复计算,每次省略 1 个环境参数,然后分析某一环境参数与遗漏误差之间的相关性,确定最优的环境参数,即如果 1 个环境参数的存在与遗漏误差呈正相关关系,表明该环境参数会降低模型预测的准确性,该环境参数会在最终的分析中被去掉。此方法有普遍适用性^[10-12]。对于 GARP 模型来说,虽然可以产生成功率较高且具有显著统计意义的结果,但预测产生了破碎化的过预测生境^[13]。有研究表明,GARP 可以成功地预测到物种大部分的分布点,同时也会预测到远远大于物种已知分布区的分布范围^[14-15]。为使预测范围更加精确,以便实施重点地段的监测

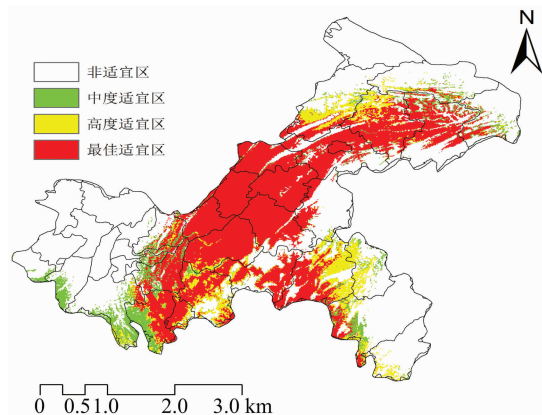


图1 基于 GARP 模型预测松材线虫在重庆的适生度空间格局

和预警,本研究在对环境因子进行筛选时,并未使用遗漏误差,而是使用溢出误差^[12]。在 Desktop Garp 模型运算的结果列表中,溢出误差代表的是预测结果超出实际分布区所占的百分比。本研究在筛选最优环境因子集合时,根据刀切法分析,温度季节变化、最热月最高温度、最湿润季度平均温度、最冷季度平均温度、最干旱月降水量、降水量季节变化和最热季度降水量被去除时溢出误差明显下降,表明这 7 个环境参数与溢出误差呈正相关,说明这 7 个环境参数会降低模型预测的准确性。因此,这 7 个环境参数在最终的分析中被去掉。其他的 12 个环境参数被用来建立最佳模型,也说明这 12 个环境参数对其分布影响显著,分别为年平均温度、月均值昼夜温差、等温性、最冷月最低温度、气温年变化范围、最干旱季度平均温度、最热季度平均温度、年降水量、最湿润月降水量、最湿润季度降水量、最干旱季度降水量和最冷季度降水量。Graham 在研究大量化石生物的基础上认为,物种的生境更多的与温度和降水的极值有关,而不是与平均值有关^[16],以及外来物种入侵往往具有较强的抗逆性^[10-11]。本研究结果正好验证了上述研究结果,也符合三峡大坝修建后极端气候在重庆反复出现的事实。另外,最热季度平均温度、降水季节性变化以及最温暖月最高温度是影响松材线虫传媒昆虫——松墨天牛 (*Monochamus alternatus*) 的主要环境因子^[1],因此,在研究松材线虫的潜在生境时,可以综合考虑寄主植被和传媒昆虫的生态环境因子。

3.2 潜在分布区分析

用 GARP 生态位模型模拟外来入侵物种在入侵地区的潜在分布是在展开较深入的生物学研究之前探知入侵物种可能分布的有效手段,它不需要物种的生物学资料,因此对具有潜在入侵威胁或初侵入的检疫性物种的风险分析和预防控制具有重要的参考意义^[11]。松材线虫生态位模型的预测结果主要集中在重庆中部,即三峡库区腹地,然后向重庆东北、东南以及南方向传播,与其在重庆已入侵地区的地理生态式样基本相符,所有已经入侵的地区几乎均被生态位模型预测为最佳适生区域和高度适生区域,其中涪陵区、长寿区、丰都县、垫江县、忠县、梁平区和万州区等均处在最佳适生区内,即三峡库区腹地,并与高度适生区紧密相连,即南川区、武隆区、万盛区、綦江区、南岸区、巴南区、江北区、渝北区、彭水县、石柱县、奉节县、开县和云阳县,在这些高度适生区中也发现零星松材

线虫病适生区,说明这些高度适生区就是未来松材线虫病可能爆发的区域。因此在最佳适生区和高度适生区域应该进行严格的检疫和防控措施。另外,控制松材线虫病的传媒昆虫——松墨天牛,因为控制松墨天牛是切断松材线虫自然传播途径的有效方法。

参考文献:

- [1] 巨云为,李明阳,吴文浩. 江苏省松材线虫发生的预测方法[J]. 林业科学,2010,46(12):91-96.
- [2] Robinet C, Roques A, Pan H, et al. Role of human-mediated dispersal in the spread of the pinewood nematode in China[J]. PLoS One,2009,4(2):e4646.
- [3] Guisan A, Thuiller W. Predicting species distribution; offering more than simple habitat models[J]. Ecology Letters,2005,8(9):993-1009.
- [4] 徐福元,郑华英,刘云鹏,等. 马尾松种源对松褐天牛成虫取食、松材线虫病的抗性分析[J]. 林业科学,2011,47(1):101-106.
- [5] Kanetani S, Kikuchi T, Akiba M, et al. Detection of *Bursaphelenchus xylophilus* from old discs of dead *Pinus armandii* var. *amamiana* trees using a new detection kit[J]. Forest Pathology,2011,41(5):387-391.
- [6] 何龙喜,吴小芹,吉静,等. 不同松树感染松材线虫后 NO 和核酶变化与其抗病性的关系[J]. 林业科学,2012,48(1):109-114.
- [7] 吴文浩,李明阳. 基于生态位模型的松材线虫潜在生境预测方法研究[J]. 林业调查规划,2009,34(5):33-38.
- [8] Robinet C, Opstal N V, Baker R, et al. Applying a spread model to identify the entry points from which the pine wood nematode, the vector of pine wilt disease, would spread most rapidly across Europe[J]. Biological Invasions,2011,13(12):2981-2995.
- [9] Stockwell D, Peters D. The GARP modelling system: problems and solutions to automated spatial prediction[J]. International Journal of Geographic Information Systems,1999,13(2):143-158.
- [10] 余岩,陈立立,何兴金. 基于 GARP 的加拿大一枝黄花在中国的分布区预测[J]. 云南植物研究,2009,31(1):57-62.
- [11] 周先容,余岩,周颂东,等. 巴山榧树地理分布格局及潜在分布区[J]. 林业科学,2012,48(2):1-8.
- [12] 王瑞. 我国严重威胁性外来入侵植物入侵与扩散历史过程重建及其潜在分布区的预测[D]. 北京:中国科学院研究生院,2006.
- [13] 马松梅,张明理,张宏祥,等. 利用最大熵模型和规则集遗传算法模型预测孑遗植物裸果木的潜在地理分布及格局[J]. 植物生态学报,2010,34(11):1327-1335.
- [14] Hernandez P A, Graham C H, Master L L, et al. The effect of sample size and species characteristics on performance of different species distribution modeling methods[J]. Ecography,2006,29(5):773-785.
- [15] Pearson R G, Raxworthy C J, Nakamura M, et al. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar[J]. Journal of Biogeography, 2007,34(1):102-117.
- [16] Graham A. The current status of the legume fossil record in the Caribbean region[M]. London: the Royal Botanic Gardens Press, 1992:161-167.