

文章编号: 1007-4619(2007)03-0426-07

利用遥感和 GIS 的方法预测 外来入侵物种的潜在分布

陈 浩^{1,2}, 陈利军²

(1. 武汉大学 遥感信息工程学院, 湖北 武汉 430079; 2. 国家基础地理信息中心, 北京 100080)

摘 要: 外来物种入侵已经成为人类历史上最重大的生态事件之一, 它直接威胁到了包括中国在内的世界上很多国家的经济、公众健康、农业生产力和生态的完整性。该研究利用遥感和 GIS 的方法进行了外来入侵物种的相对适生区分析, 预测了外来物种一豚草在中国的潜在分布, 并在信息理论的框架下建立了一种改进的加权平均逻辑回归模型。利用 logit 阈值和频率统计的方法对入侵物种的生境进行相对适应性划分, 减小了生物多样性数据的不对称带来的影响。

关键词: 遥感; GIS 入侵物种; 逻辑回归; 信息理论; 频率统计

中图分类号: P208/TP79 **文献标识码:** A

Predicting the Potential Distribution of Invasive Exotic Species Using GIS and Remote Sensing

CHEN Hao^{1,2}, CHEN Lijun

(1. School of Remote Sensing Information Engineering, Wuhan University, Hubei Wuhan 430079, China;

2. National Geomatics Center of China, Beijing 100080, China)

Abstract: Exotic species invasion has been one of the most dramatic ecological event in human history that threatens our economy, public health and ecological integrity. Explaining the nature of the species and species-environment relationship and predicting the spatial distribution of the invasive exotic plants is of great importance for invasive exotic plants prevention and early warning efforts. One approach to species specific predictions involves the use of habitat suitability or niche-based models whereby environmental conditions suitable for maintenance of populations of a species are identified and mapped onto geographic space. These approaches combine herbarium specimen locations data with a suite of GIS layers e.g. climatic, topographic and land cover to create the ecological models of the species requirements. Coupled with these models, GIS can project the ecology model onto geographic space and mapping the habitat suitability maps in native ranges and exotic ranges.

This paper proposes an improved logistic regression approach in an information theoretic framework to predict the suitability of ragweed in both native and invaded ranges. Information theoretic approaches computed and assessed the modeling choice as well as produced a weighted average model based on the multiple models rather than using the sole model with the lowest AIC value or the highest Akaike weight. This multiple model inference is useful to reduce model selection bias.

Having obtained the weighted average model, the resulting regression equations were applied to the native samples including the present points and pseudo-absence points to produce the output of the logit value. Because

收稿日期: 2006-09-01; 修订日期: 2006-11-02

基金项目: 国家自然科学基金 (编号: 40371084)

作者简介: 陈 浩 (1979—), 男, 武汉大学遥感信息工程学院, 博士研究生, 现从事 GIS 空间建模、环境遥感和数据挖掘方面的研究。

E-mail: chenhaoq87@whu.com

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

of lack in true absence data, we didn't transform the logit value back to regular space scaled from 0—1 representing probability of a pixel containing the species but regard the logit value as the degree of the suitability for the species. So we proposed a new approach specifically to compartmentalize the habitat suitability using logit value thresholds and frequency statistics. At last, we used this habitat suitability model developed in native ranges to project onto the exotic ranges to predict the ragweed's potential distribution in China.

Key words: RS; GIS; invasive exotic species; logistic regression; information theory; frequency statistics

1 引 言

外来物种入侵是一个全球性问题,对任何一个国家来说,外来入侵物种对其国民经济、公众健康、农业生产力和生态完整性都构成了极大的危害^[1]。同时,由于生物本身具有繁殖、扩散和适应能力,当一个新的入侵物种被发现造成重大影响时,对其进行治理是极其困难的,因此预测和预防物种的入侵比控制其爆发更为可行,也更经济^[2]。由于人们对大尺度地域关联性了解的局限和地理空间技术运用不足,现今对入侵物种的研究多半集中于单个物种和植物群落之间关系的研究上,从而忽视了地理空间技术的运用^[3]。潜在的生物入侵本质上是一个地理学问题,对生物入侵的研究应当可以运用地理学知识包括遥感和 GIS 技术手段来解决,这对于开展入侵物种的监测和治理研究具有非常大的意义。

国外学者运用地理空间技术研究生物入侵起步较早。早在 1987 年美国地质调查局 (USGS) 就用遥感和 GIS 技术对生物资源进行调查、编目,实施长期的动态监测,进而评定了特定区域内的生物多样性。在生物入侵研究方面,中国起步较晚,特别是对入侵物种的预测、分析和治理未能充分地发挥作用^[4]。同时,应用遥感技术对物种资源实施监测也存在一定的困难,许多物种很难直接通过卫星遥感影像来解译。所以,有的学者采用精度很高的影像如红外航空像片来解译外来入侵种^[5],但这种方法花费的成本过大,且不宜推广。其实,对许多入侵物种来说,通过确定与之相关的生命体和环境的关系来评定栖息地是有可能的,只是模型的精度取决于对物种和生态环境之间关系的知识把握。如果结合 GIS 辅助数据和生物地理分区成果,利用卫星影像数据进行生物群系和生态区的制图,确定干扰梯级的景观特征,则是研究生物入侵的新的途径。这一过程包括使用物种标本的地理定位信息以及一系列的遥感和环境数据(气候、地形、生态等)进行机器学习和统计分析得出一些规则,通过对规则的优化选择,形成入侵物种最适宜生存的生境模型,然后把它从

多维的生态空间投影到地理空间中,利用遥感和 GIS 技术绘制物种在入侵地的潜在分布图^[6]。

逻辑回归模型是生态生境建模常用到的方法之一,这种方法通过在多维的生态空间建立模拟物种适应性程度的生境适应性概率曲面,从而达到预测物种的目的。然而生物多样性数据与其他的数据相比具有数据不对称的特点,在建立物种适应性生境模型过程中,往往同时需要准确的物种出现以及物种不出现的地理定位信息。实际上,物种不出现信息是很难获得的,在大多数博物馆(或标本馆)数据库中,往往只有物种收集地点的记录,而没有关于物种在哪儿不出现的信息,这种生物多样性数据的不对称性造成了空间建模的困难和局限^[7]。把地理空间中的随机采样点作为“伪物种不出现”数据点,是解决这一问题常用的折中方法^[8]。但是现实世界中,在这些“伪物种不出现”点上物种有可能会存在。因此,在利用逻辑回归进行生境建模的过程中,需要建立一种改进的方法去解决在缺少真实物种不出现信息时进行合理预测的问题。对于逻辑回归而言,传统的方法是基于零假设的显著性检验。这种方法的有效性特别是在不相关零假设的使用,显著性水平设定的人为武断,基于单一模型的模型选择不确定性等方面已经受到了很多学者的质疑^[9-10]。近些年来,基于信息理论的模型选择和推论被人们认为是基于零假设的显著性检验的一种替代方法。它的理论基础是:不存在单一的真实的模型,模型只能无穷的接近现实,目标是确认哪个模型更接近真实,即信息的损失最小(Ku back-Leibler 信息量)。信息理论的方法考虑了模型的拟合程度和模型的复杂程度两个方面的因素,同时还可以进行多模型平均的建模和变量评估,克服了传统显著性检验的一些不足^[11]。

本文利用地理空间技术和统计学的方法生成了一系列的物种适生逻辑回归候选模型,并通过信息理论(AIC 标准)及其推论建立了基于多模型框架下的逻辑加权平均模型。由于获取的物种的标本数据只包括物种出现点的信息,为了减少这种采样偏差,本文提出了一种利用逻辑阈值和频率统计的方

法进行相对适应性生境的划分,减小多样性数据的不对称性带来的影响。

2 入侵物种适生区建模的原理和方法

2.1 Kullback-Leibler信息量与 Akaike信息标准

信息理论是由 Kullback-Leibler信息量发展而来的。1951年 Kullback和 Leibler发表了一篇描述“信息”的科学含义的论文^[12],在文中提出了著名的 Kullback-Leibler(K-L)信息量的概念。K-L信息量描述了概念上的真实模型和现实中的近似模型 g 之间的度量(距离),它的表达是一个连续函数的积分:

$$K(f, g) = \int f(x) \log \left(\frac{f(x)}{g(x|\theta)} \right) dx \quad (1)$$

式中, f 和 g 是 D 维的概率分布, $K(f, g)$ 表示了候选模型 g 近似代替现实模型 f 的信息损失。假设有 R 个候选模型,目标是寻求使 $K(f, g)$ 最小的近似模型。

Akaike^[9, 11]研究了 K-L信息量(信息和编码领域)和最大释然(统计学领域)的关系,使在单一理论框架(优化)下,模型评估和模型选择的结合成为可能。他通过最大释然函数对相对 K-L信息量进行了估计,并提出了 Akaike信息标准(AIC):

$$AIC = -2 \log(L(\theta | data)) + 2k \quad (2)$$

式中, $\log(L(\theta | data))$ 是给定数据 $data$ 候选模型、未知参数 (θ) 的最大对数似然(log-likelihood), k 为模型中参数个数(包括变量个数和截距)。AIC是一种封装了模型灵活性和模型过度拟合的折衷的评分函数,其中 $-2 \log(L(\theta | data))$ 为模型误差部分, $2k$ 为惩罚项。AIC的大小体现了建立的模型与真实模型接近的程度,其中 AIC 越小,构建的模型越优化^[13]。

2.2 逻辑加权平均模型

在信息理论中除了 AIC值以外,还有两类评价候选模型的方法:一类是 ΔAIC 指标,它代表了候选模型 AIC 值与所有候选模型中最小的 AIC 值(minAIC)之间的差值, ΔAIC 越小,模型越好^[11];另一类是 Akaike 权重 (W_i) 指标,它代表了该候选模型是最佳模型的概率^[13],其中 R 是所有候选模型的数量, Δ_i 是候选模型 i 的 ΔAIC 指标:

$$W_i = \frac{\exp(-\Delta_i/2)}{\sum_{i=1}^R \exp(-\Delta_i/2)} \quad (3)$$

由于现实世界的复杂性,单一的模型(具有最

小的 AIC 或最大 W_i 的模型)往往不能达到完全反映真实模型的目标。本文通过信息理论的推论建立了逻辑加权平均模型^[9]:

$$\hat{\theta} = \sum_{i=1}^R W_i \hat{\theta}_i \quad (4)$$

式中, $\hat{\theta}$ 表示加权平均模型各预测变量的系数, $\hat{\theta}_i$ 表示候选模型 i 的系数。

2.3 基于频率统计的生境适应性划分方法

由于缺少真实的物种不出现的地理定位信息,利用逻辑加权平均模型计算出的生境适应性概率(0—100%)可能没有现实意义。本研究提出了一种基于频率统计的生境相对适应性划分方法。通过加权平均模型,计算出源地每个像素的 \logit 值(逻辑回归模型的线性部分),仅将该值作为衡量物种相对适应性程度的指标,而不是把它转换到代表物种在每个像素范围中可能出现概率的传统的空间尺度上去,利用出现点的最小的 \logit 值作为阈值,将所有源生地的样点(包括物种出现点和“伪物种不出现”点)重新划分为两族,分别对这两族进行四分位数统计,将源地划分为 8 个相对适生类别,最后这 8 个 \logit 值区间成为了相对适应性生境的划分标准(表 1)。这种类似于重分类的方法能减小随机采样生成“伪物种不出现”点造成的误差。

表 1 相对适应性划分标准

Table 1 The habitat suitability classification criteria

相对适应性	划分标准
不适生区域:	$< \logit^a$
1 最不适应	$< 25\%$ 分位数 ^b
2	25% 分位数 ^b — 50% 分位数 ^b
3	50% 分位数 ^b — 75% 分位数 ^b
4 轻度不适应	75% 分位数 ^b — \logit^a
适生区区域:	$> \logit^a$
5 轻度适应	\logit^a — 25% 分位数 ^c
6	25% 分位数 ^c — 50% 分位数 ^c
7	50% 分位数 ^c — 75% 分位数 ^c
8 最适应	$> 75\%$ 分位数 ^c

注: \logit^a 物种出现点的最小 \logit 值

分位数^b: \logit 值小于 \logit^a 值的“伪不出现点”的基于 \logit 值的分位数统计

分位数^c: 包括 \logit 值大于 \logit^a 值的“伪不出现点”和所有物种出现点的基于 \logit 值的分位数统计

3 研究物种和 GIS 数据准备

3.1 研究物种

豚草 (*Ambrosia artemisiifolia* L.) 属菊科一年生草本植物, 别名艾叶破布草, 原产美国和加拿大。近 200 年来, 豚草迅速在世界范围内蔓延, 由于其花粉是过敏性鼻炎和季节性哮喘的主要病源^[14], 因此严重影响了人类的健康, 豚草生长繁殖迅速, 与周围的农田争夺养分, 造成了作物减产甚至绝收, 是世界公认的公害性杂草之一。20 世纪 30 年代传入中国以后, 豚草就在中国迅速扩张, 并在东北, 华北, 华中和华东等地约 15 个省 (直辖市) 有分布, 而且有继续扩张的势头。

3.2 物种数据

豚草的适生区建模所需物种数据, 包括从美国和加拿大各大博物馆或标本数据库收集的 243 个源生地 (美国和加拿大) 的植物标本记录, 和从中国科学院植物研究所标本馆及中国科学院昆明植物研究所标本馆收集的 83 个豚草在入侵地 (中国) 的植物标本记录。鉴于本研究缺乏关于豚草不出现的地理位置的信息, 因此在豚草的源生地陆地部分 (美国和加拿大南部) 随机采了 1000 个样点, 对所有的物种点 (包括标本点和随机采样点) 按照精度为 0.167° 建立经纬度地理参考。运用缓冲区分析从 1000 个随机点中去除标本点附近 8 个邻域所包含的随机采样点, 剩下的 852 个随机样点作为“伪物种不出现”数据点。

3.3 GIS 数据的准备

本文选择了 6 个气候或遥感预测变量图层, 它们分别是高程、年平均降水、最冷月份的最低温、太阳辐射、单位像素内农业和城镇密度、单位像素内森林覆盖密度图层。这些来源于不同地理尺度的数据都转换为 $0.167^\circ \times 0.167^\circ$ 空间分辨率的栅格图层。三个气候学数据 (年平均降水、最冷月份的最低温、太阳辐射) 来源于 WorldClim 1.4 数据集。高程数据来源于美国地质调查局 HYDRO-1K 数据集 (<http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html>)。人类活动的干扰是入侵物种扩散和生存的重要因素, 本文用农业和城镇密度代表了人类干扰的程度。农业和城镇密度图层通过美国地质调查局全球土地利用数据集^[15] (USGS Global Landcover

Characterization dataset) 计算得来, 其中每个像素值代表了在该像素内农业和城镇所占比例 (取值范围从 0—100%)。森林覆盖密度是基于高分辨率辐射计 (AVHRR) 数据和 TM 影像数据进行回归分析计算得出的, 该图层来源于联合国森林资源评估数据集 (UN/FAO Forest Resources Assessment 2000 dataset), 其中每个像素值代表了在该像素内森林所占比例 (取值范围从 0—100%)。

3.4 模型变量的选择

逻辑回归建模过程中, 为了使模型结果同时具有统计学和生态学的意义, 本文建立了一种“核心变量加边缘变量”的模型变量选择策略。基于豚草生长和扩散的生物学先验知识, 选择了三个变量及其各自的平方项作为模型的核心预测变量, 它们包括: 最冷月份的最低温 (T)、太阳辐射 (S)、年平均降水 (P)。剩下的三个变量及其各自的平方项加上两个交互项成为模型的边缘变量, 它们包括: 高程 (E)、单位像素内农业和城镇密度 (AU)、单位像素内森林覆盖密度 (F)、最冷月份的最低温 \times 太阳辐射 (TS)、最冷月份的最低温 \times 年平均降水 (TP)。模型变量的描述及其缩写形式见表 2。首先建立包含所有核心变量及其平方项的“核心”模型 (Core), 然后采用前向和后向逐步选择的方法, 依次把边缘变量的不同组合加入到“核心”模型中, 组成更为复杂的候选逻辑模型。

表 2 模型参数变量名称及重要性权重

Table 2 The important weights for each of the variables

模型参数	变量名称及缩写 ^a	权重 ^b
核心变量: Core	最冷月份的最低温 (T_1, T_2)	1.000
	太阳辐射 (S_1, S_2)	1.000
	年平均降水 (P_1, P_2)	1.000
边缘变量: Marg	高程 (E_1, E_2)	0.999
	单位像素内农业和城镇密度 (AU_1, AU_2)	0.753
	单位像素内森林覆盖密度 (F_1, F_2)	0.981
	最冷月份的最低温 \times 太阳辐射 (TS)	0.950
	最冷月份的最低温 \times 年平均降水 (TP)	0.444

^a下标 1 表示一次项, 下标 2 表示二次项

^b权重计算时包括所有的一次项和二次项

4 结果与分析

4.1 信息理论框架下的逻辑回归模型

根据“核心变量加边缘变量”的模型变量选择策略,通过 R 统计工具建立了候选二元逻辑回归模型集,并计算出每个候选模型相应的信息理论指标,它们包括: Akaike 信息标准 (AIC), Δ_{AIC} 指标, Akaike 权重 (W_i) 以及累积 Akaike 权重 (W_s)。根据 AIC 从小到大的顺序选择了 9 个较优模型,见表 3。

表 3 9 个最佳候选逻辑回归模型的信息理论统计结果

Table 3 Results of information theoretic statistics for the best nine models for ragweed

模型结构 ^a	AIC	Δ_i	W_i	W_s
Core E ₁ AU ₁ AU ₂ F ₁ F ₂ TS	938.4	0.00	0.246	0.246
Core E ₁ AU ₁ AU ₂ F ₁ F ₂ TP	939.0	0.6	0.186	0.432
Core E ₁ F ₁ F ₂ TS TP	940.2	1.8	0.103	0.535
Core E ₁ E ₂ AU ₁ AU ₂ F ₁ F ₂ TS	940.4	2.0	0.091	0.626
Core E ₁ F ₁ F ₂ TS	940.6	2.2	0.081	0.707
Core E ₁ E ₂ F ₁ F ₂ AU ₁ AU ₂ TS TP	940.9	2.5	0.070	0.777
Core E ₁ F ₁ F ₂ AU TS	941.5	3.1	0.052	0.829
Core E ₁ E ₂ F ₁ F ₂ TS TP	942.0	3.6	0.041	0.870
Core E ₁ E ₂ F ₁ F ₂ TS	942.6	4.2	0.031	0.901

^a 模型变量缩写形式见表 2

表 3 中第一个模型 (模型结构为“Core E₁, AU₁, AU₂, F₁, F₂, TS”)具有最小的 AIC 值及最大的 Akaike 权重,这说明它在整个候选逻辑模型集中是相对最优的,但是它的 Akaike 权重为 0.246 表明用该模型近似描述真实模型的置信概率只有 24.6%,把这个单一模型作为最终模型缺乏足够的说服力。在此基础上通过信息理论的推论计算了多模型加权平均。所选择的 9 个模型的 Δ_{AIC} 指标都小于或接近 4 可以认为它们有助于反映真实模型^[13]。这 9 个模型的累积 Akaike 权重为 0.901,用这种多模型结构去近似描述现实世界的真实模型,比其他的单一模型具有更高的置信概率 (90.1%)。根据方程 (4),得出的最终模型是基于 Akaike 权重的加权平均,包括了上述 9 个逻辑回归候选模型中所包含的核心变量和所有的边缘变量。

在信息理论框架下,通过对所有包含任一变量的模型的 Akaike 权重统计求和来评估气候环境变量对物种入侵的重要性 (表 2)。除了三个核心变量 (权重为 1) 以外,高程 (0.999) 和单位像素内森林覆盖密度 (0.981) 都具有很高的权重,并出现在所

有 9 个候选逻辑回归模型中,说明这是两个对豚草入侵非常重要的因素。同时温度和太阳辐射的交互作用 (0.950) 也是影响物种入侵特别是短日照植物豚草引种生长的重要因子。

4.2 相对适生区制图

中国和北美 (特别是美国和加拿大南部) 具有相似的地理纬度和植物分类体系,利用遥感软件 ERDAS 8.6 把在源生地建立的逻辑加权平均模型运用到物种的源生地,并结合基于源生地物种出现点的 logit 阈值计算的适应性生境划分标准 (8 个 logit 值区间) 绘制了豚草在北美的分布图 (图 1(a)); 同时“投影”到物种入侵地——中国,绘制了豚草在中国的潜在分布图 (图 1(b))。图 1 中随着黄色的逐步加深,豚草在该区域不出现的可能性加大;随着蓝色的逐步加深,豚草在该区域出现的可能性加大。

根据已知的豚草在入侵地 (中国) 的植物标本记录对模型准确性进行了评估: 其中有 0 个在第 1 到第 5 类中; 有 4 个在第 6 类中; 有 9 个在第 7 类中; 剩下的 70 个点在第 8 类中,也就是最适应区域内; 所有的 83 个已知豚草出现点都落在了预测适生区域内 (包括 5 6 7 8 四个类)。如果以是否落在适生区内作为评估预测准确性的标准,那么,预测豚草在中国的分布的准确率为 100%; 如果分别以第 5 6 7 类的适生区界限作为评估标准,那么,预测准确率分别为 100%, 95.2%, 84.3%。由于缺少真实的物种不出现的地理定位数据,上述的预测准确率只能反映模型具有较低的漏判率 (分别为 0.048%, 15.7%) 而不能评估预测模型的错判率。同时,图 1(b) 中豚草表现出在入侵地——中国具有强大的扩散潜力,甚至表现出了一种“过度预测” (Overprediction) 的趋势,除了西藏和青海等区域外,豚草几乎在全国都能生长,这可能是在建模过程中没有考虑到在不同地理空间 (北美和中国) 中入侵物种生境的变化原因。

考虑到在不同地理空间中生境的变化,特别是在入侵初期当地生物因素对入侵物种种群建立的制约 (如竞争、捕食等), 本文建立了两种相对适应性生境的划分标准。一种是利用源生地物种出现点的 logit 阈值将所有源生地的样点 (包括出现点和“伪物种不出现”点) 重新划分为两族,然后分别进行四分位数统计,结果如图 1 所示; 另一种是利用入侵地已知物种出现点的 logit 阈值进行划分得出的相对适应性生境的划分标准,结果如图 2 所示。由于

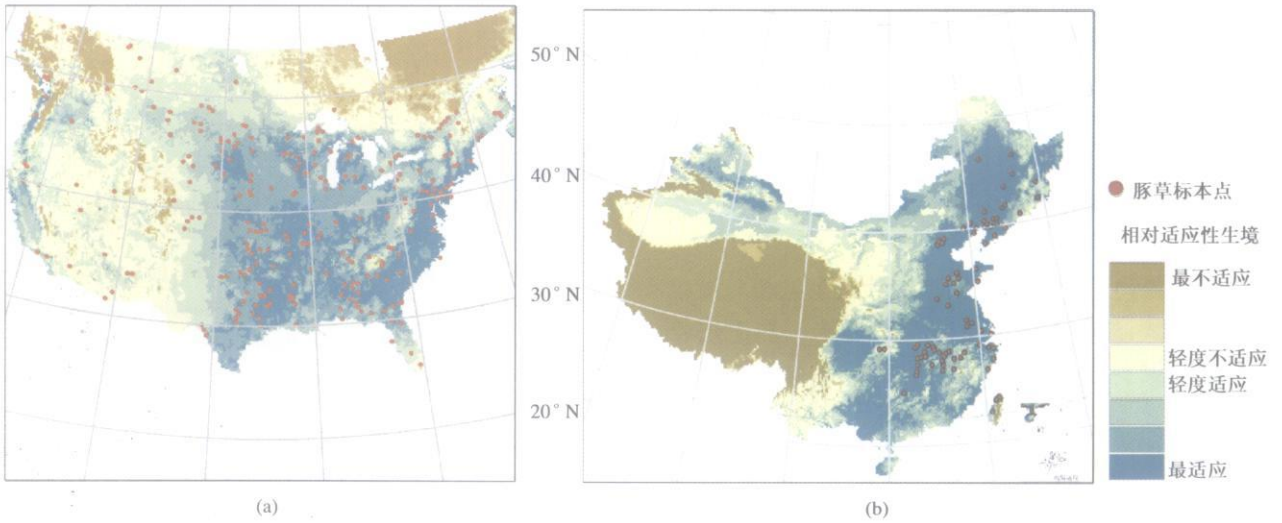


图 1 基于源地 logit 阈值划分标准绘制的豚草在源地(a)和入侵地(b)的相对适应性图
 Fig.1 Habitat-suitability maps based on logit thresholds derived from the native herbarium specimen records for ragweed in both native (a) and invaded ranges (b)

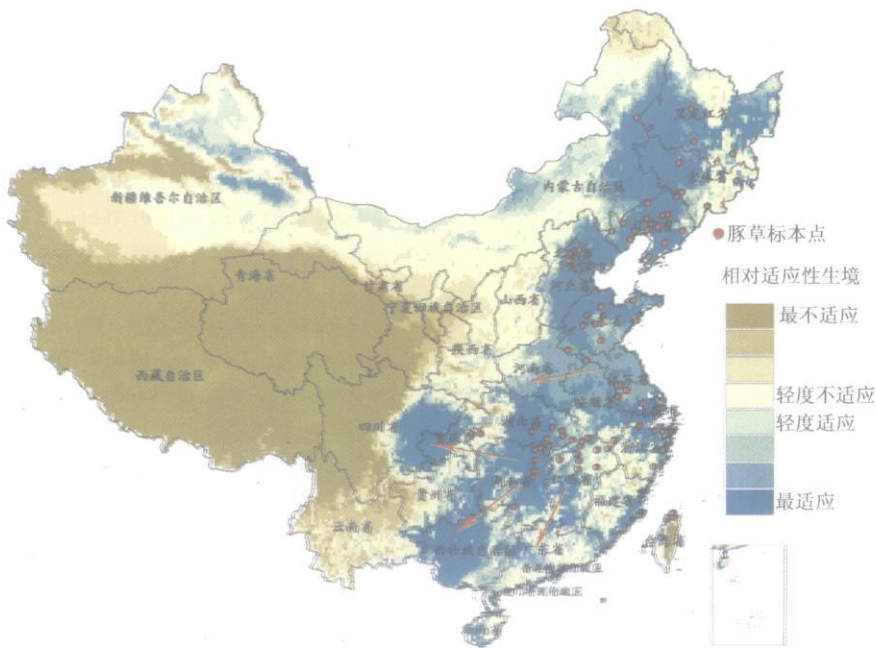


图 2 基于入侵地 logit 阈值划分标准绘制的豚草在入侵地的相对适应性图
 Fig.2 Habitat-suitability maps based on logit thresholds derived from the invaded herbarium specimen records for ragweed in invaded ranges

图 2 通过加入入侵地已知的物种出现点的数据, 将入侵地内生物因素的制约融入进模型中, 结果更接近豚草在中国的实际分布。从豚草在中国的潜在分布图可以看出: 除了已知的豚草入侵区域外, 四川盆地、河南、广西、广东等省份都是适合豚草生长的区域, 表现出了豚草在中国强大的继续扩散的潜力 (图 2 中绿色箭头表示豚草可能的扩散方向) 如果该区域没有采取足够的预防措施, 在不久的将来豚草这种危害性巨

大的恶性杂草将可能蔓延到该区域。

5 结 论

(1) 潜在的生物入侵本质上是一个地理学问题, 因此, 对生物入侵的研究可以运用地理学知识包括遥感和 GIS 技术手段来解决。本文中, 利用遥感和 GIS 的方法, 在信息理论框架下建立了一种改进

加权平均逻辑回归模型, 对外来入侵物种——豚草在中国的潜在分布进行了预测, 能帮助政府的管理和决策者进行入侵物种的监测和防治。

(2) 信息理论的方法考虑了模型的拟合程度和模型的复杂程度两个方面的因素, 同时可以进行多模型平均的建模和变量评估, 克服了传统的显著性检验的不足。

(3) 提出了一种基于频率统计的生境相对适应性划分方法, 这种类似于重分类的方法能减小生物多样性数据不对称造成的影响, 在只有物种出现数据的情况下, 进行入侵物种的潜在分布预测。

基于信息理论和频率统计的方法在以下方面还有待进一步深入研究。第一, 由于缺少真实的物种不出现数据及基于最小 \log 阈值的频率统计方法对样本整体的敏感性, 缺乏对预测模型准确率进行全面评估的方法, 在今后的研究中, 需要加以改善。第二, 通过源生地的模型去预测物种在入侵地的潜在分布, 会出现生境在不同地理空间的变化, 也就是当前国内外学者提出的模型适应性的问题。在本文中, 模型的建立只考虑到了物理 (如气候) 方面的因素, 只是把已知的入侵地物种的状况作为所有生物学因素 (如源生地和入侵地的生态差异, 特别是竞争、捕食和该区域的生物多样性情况等) 的简单、间接的近似, 缺少对生物学因素的科学、直接的数学描述, 在今后需要进一步加以研究。第三, 人类商贸和文化活动对于入侵物种的分布和扩散也有一定的影响, 把这些相关因素 (特别是人工道路) 加入模型中可能会对模型的性能有所改进。但是, 由于大尺度 (较粗的分辨率) 下, 中美两国的道路网模式 (尤其是高级别道路) 有着较大的差异, 因此在本文中, 没有考虑道路对物种入侵的影响。不同尺度下的预测模型将是今后研究的重点, 在较小尺度 (较高的分辨率) 的区域预测中, 道路和传播路径将作为重要因子加以考虑。

参 考 文 献 (References)

- [1] Guo Q. Perspectives on Trans-pacific Biological Invasion [J]. *Acta Phytocologica Sinica* 2002 26(6): 724—730
- [2] Waagbø T K, Reaser J K. A Global Strategy to Defeat Invasive Species [J]. *Science* 2001 292(5521): 1468—1486
- [3] Guo Q. Ecological Comparisons between Eastern Asia and North America: Historical and Geographical Perspectives [J]. *Journal of Biogeography* 1999 26(22): 199—206
- [4] Xie Y, Li Z, Gregg W P et al. Invasive Species in China: an Overview [J]. *Biodiversity and Conservation* 2000 10(8): 1317—1341
- [5] Stow D, Hope A, Richardson D et al. Potential of Colour-infrared Digital Camera Imagery for Inventory and Mapping of Alien Plant Invasions in South African Shrublands [J]. *International Journal of Remote Sensing* 2000 21(15): 2965—2970
- [6] Peterson A T. Predicting the Geography of Species Invasions via Ecological Niche Modeling [J]. *Quarterly Review of Biology* 2003 78(4): 419—433
- [7] Manel S, Dias J M, Omerod S J. Comparing Discriminant Analysis, Neural Networks and Logistic Regression for Predicting Species Distributions: a Case Study with a Himalayan River Bird [J]. *Ecological Modelling* 1999 120(2/3): 337—347
- [8] Phillips S J, Anderson R P, Schapire R E. Maximum Entropy Modeling of Species Geographic Distribution [J]. *Ecological Modelling* 2006 190: 231—259
- [9] Anderson D R, Burnham K P. Avoiding Pitfalls when Using Information-theoretic Methods [J]. *Journal of Wildlife Management* 2002 66(33): 912—918
- [10] Anderson D R, Burnham K P, Thompson W L. Null Hypothesis Testing: Problems, Prevalence, and an Alternative [J]. *Journal of Wildlife Management* 2000 64(4): 912—923
- [11] Anderson D R, Burnham K P, White G C. Kullback-Leibler Information in Resolving Natural Resource Conflicts when Definitive Data Exist [J]. *Wildlife Society Bulletin* 2001 29(4): 1260—1270
- [12] Kullback S, Leibler R A. On Information and Sufficiency [J]. *Annals of Mathematical Statistics* 1951 22(1): 79—86
- [13] Johnson J B, Omernik K S. Model Selection in Ecology and Evolution [J]. *Trends in Ecology and Evolution* 2004 19(2): 101—108
- [14] Creticos P S, Reed C E, Norman P S et al. Ragweed Immunotherapy in Adult Asthma [J]. *J Allergy Clin Immunol* 1996 334(8): 501—506
- [15] Loveland TR, Reed B C, Brown J F et al. Development of a Global Land Cover Characteristics Database and IGBP Discover from 1-km AVHRR Data [J]. *International Journal of Remote Sensing* 2000 21(6/7): 1303—1330