

DOI: 10.5846/stxb202010192661

巨喜锋, 林峻, 吴建国, 陈吉军, 关靖云, 李漠岩, 郑江华. 新疆典型蝗虫适生区分布预测. 生态学报, 2022, 42(21): 8605-8617.

Ju X F, Lin J, Wu J G, Chen J J, Guan J Y, Li M Y, Zheng J H. Prediction of potential living area of typical locusts in Xinjiang based on species distribution model. Acta Ecologica Sinica, 2022, 42(21): 8605-8617.

新疆典型蝗虫适生区分布预测

巨喜锋¹, 林峻², 吴建国², 陈吉军², 关靖云³, 李漠岩³, 郑江华^{1,3,*}

¹ 新疆大学资源与环境科学学院, 乌鲁木齐 830046

² 新疆治蝗灭鼠指挥部办公室, 乌鲁木齐 830001

³ 绿洲生态教育部重点实验室, 乌鲁木齐 830046

摘要:新疆草原面积广阔, 农牧业地位突出, 蝗灾对当地经济、生态威胁很大, 近年新疆极端天气日渐频发, 蝗灾监测与防治任务艰巨。以意大利蝗、西伯利亚蝗和亚洲飞蝗为代表的蝗虫数据为基础, 综合考虑对蝗虫各生命周期有重要影响的环境因素, 运用 BIOCLIM 模型、领域模型 (DOMAIN)、马氏距离模型 (MAHAL)、广义线性模型 (GLM)、随机森林模型 (RF)、提升回归树模型 (BRT)、支持向量机模型 (SVM)、最大熵模型 (MaxEnt) 等八种经典物种分布模型及集成模型对新疆典型蝗虫适生区展开了预测。结果表明: (1) 不同模型对新疆典型蝗虫适生区预测存在差异, 其中 DOMAIN 最差 (曲线下面积 (AUC) = 0.688, 真实技巧统计 (TSS) = 0.301), 而提升回归树 (BRT) 最佳 (AUC = 0.920, TSS = 0.910), 基于 BRT、SVM 和 MaxEnt 3 个集成模型预测的新疆蝗虫适生区可靠性更高; (2) 新疆典型蝗虫不同等级适生区面积约 56.844 万 km², 占新疆总面积的 36%, 其中高适生区面积 16.568 万 km²; (3) 新疆典型蝗虫适生区主要集中于北疆阿勒泰、塔城地区, 此外东部哈密地区及南疆绿洲边缘地带亦有分布。研究可为新疆草原工作部门推进蝗虫监测防治工作提供支持。

关键词:蝗虫; 物种分布模型; 适生区; 模型集成; 新疆

Prediction of potential living area of typical locusts in Xinjiang based on species distribution model

JU Xifeng¹, LIN Jun², WU Jianguo², CHEN Jijun², GUAN Jingyun³, LI Moyan³, ZHENG Jianghua^{1,3,*}

¹ College Of Resources And Environmental Science, Xinjiang University, Urumqi 830046, China

² Locust And Rodent Control Headquarters Of Xinjiang Uygur Autonomous Region, Urumqi 830001, China

³ Key Laboratory Of Oasis Ecology, Ministry Of Education, Urumqi 830046, China

Abstract: Xinjiang grassland covers a vast area and occupies a prominent position in agriculture and animal husbandry. Locust plague is a great threat to the local economy and ecology. Furthermore, extreme weather in Xinjiang has become more frequent, the task of monitoring and preventing locust plagues is difficult. Based on the data of locusts represented by Italian locust, Siberian locust and Asian migratory locust, we comprehensively considered the environmental factors that had important influence on the different life cycles of locusts, and used eight typical species distribution models of BIOCLIM, DOMAIN, MAHAL, GLM, RF, BRT, SVM, MAXENT, and ensembled models to predict typical locust suitable areas in Xinjiang. The results showed that: (1) different models had differences in the prediction of typical locust suitable areas in Xinjiang, in which DOMAIN was the worst (AUC = 0.688, TSS = 0.301), and the BRT was the best (AUC = 0.920, TSS = 0.910). The prediction of the suitable area based on the three ensembled models of the BRT, SVM and MaxEnt was more reliable. (2) The suitable area of typical locusts of different grades in Xinjiang was about 568440 km², accounting for 36%

基金项目:天山雪松项目 (2020XS04); 新疆草原生物灾害遥感监测项目 (2021)

收稿日期:2020-10-19; **网络出版日期:**2022-06-21

* 通讯作者 Corresponding author. E-mail: Zheng_jianghua@126.com

of the total area of Xinjiang, among which the high adaptable area was 165680 km². (3) The typical locust breeding areas were mainly concentrated in the Altay and Tacheng areas of northern Xinjian. In addition, they were also distributed in Hami and the edge of oasis in southern Xinjiang. This study may provide support to the Xinjiang grassland authorities in their efforts to monitor and control locusts.

Key Words: locust; species distribution model; suitable areas; ensembled model; Xinjiang

草地是自然生态系统的重要组成部分,不仅在水土保持、生态平衡、碳固定方面具有重要生态价值,而且在政治、经济、文化、生态环境建设方面扮演着重要角色^[1-2]。新疆草原草地面积为 57.2588 万 km²,在全国草地面积中排名第三^[3],其健康受包括蝗灾在内的各类草原生物灾害严重影响,其中蝗灾与水灾、旱灾并称为三大天灾,新疆蝗灾发生最早可追溯到东汉时期^[4],1986—1991 年发生面积近 4 万 km²,其中严重危害区面积近 2.8 万 km²,发生面积占可利用面积的 8.5%^[5]。新疆草原蝗虫具有种类多,分布广,数量大,灾害严重的特点。20 世纪陈永林^[6]调查新疆蝗虫有 62 属 157 种;近期于冰洁等^[7]指出新疆蝗虫种类达 65 属 177 种;而徐超民等^[8]认为新疆草原蝗虫发生区蝗虫记载有 8 科 16 属 171 种,其中优势种包括亚洲飞蝗 (*Locusta migratoria migratoria*)、意大利蝗 (*Calliptamus Italicus*)、黑腿星翅蝗 (*C. barbarus*)、西伯利亚蝗 (*Gomphocerus Sibirius*)、宽须蚁蝗 (*Myrmeleotettix palpalis*)、黑条小车蝗 (*Oedaleus decorus*)、红胫戟纹蝗 (*Dociostaurus kraussi kraussi*) 等十余种。近年新疆极端气候事件日渐频繁,气候变暖和极端降水增加与新疆典型蝗虫持续爆发关系紧密^[9],此外肆虐亚非的沙漠蝗 (*Schistocerca gregaria*),使得全球粮食安全和生态安全备受关注,文献也应用多种技术方法研究了沙漠蝗的迁飞通道以及入境中国的风险,此外 Guan 等还通过 MaxEnt 模型研究了在未来气候条件下沙漠蝗的适生区分布情况^[10-11]。在常态化疫情防控条件下,切实做好新疆本地典型蝗虫适生区预测与境外蝗虫迁入入境监测工作异常重要。新疆草原蝗虫在日常监测、防治方面面临防治范围广、防治监测蝗虫类型多的严峻挑战。

物种分布模型 (Species Distribution Models, SDMs) 以已知物种分布点及环境数据作为输入,利用数学统计方法或机器学习理论探究物种在生态位空间实际状况^[12],在生物多样性保护研究、物种入侵研究及物种适生区预测等方面得到广泛应用^[13]。在模型发展中形成了基于生态位理论、数理统计和机器学习物种分布模型^[14],其中 BIOCLIM 模型、HABITAT 模型及领域模型 (DOMAIN) 等都是典型的基于生态位量化理论而形成的模型,该类模型基于典型已知物种分布点及环境空间中多变量极值组成的包络体外边缘,预测出物种分布的潜在区域^[15]。在此基础上,随着数学方法的发展,出现了应用数理统计方法探索物种分布点与预测区统计学关系以预测物种潜在分布区的模型算法,该类模型算法包括广义线性模型 (GLM)、随机森林模型 (RF) 和马氏距离 (MAHAL) 等^[16]。近些年伴随着计算机科学的发展,机器学习的理论思想逐步引用到物种分布模型预测中,其中应用相对较为广泛的有人工神经网络 (ANN)、支持向量机 (SVM)、提升回归树 (BRT) 以及最大熵模型 (MaxEnt) 等一大批优秀模型^[17]。但基于不同模型的物种分布区及潜在适生区预测结果具有很强的不确定性,从而导致单个最优模型的选择具有一定困难,因此将多个不同模型预测输出组合为一个集成模型变得尤为重要^[18]。集成建模的基本思想是将各单一模型中捕获到的真实信号予以保留,对数据和模型中产生的错误和不确定更好地分离和降低,从而使得建模的结果更为可靠^[19-20]。

在新疆典型蝗虫发生预测相关研究中,杨洪升等^[21]探究了新疆蝗虫与植被长势、气候因子间的关系,指出蝗虫发生与植被长势息息相关,且温度是影响植被长势的主要因子。Zheng 等^[22]使用基于 GIS 多准则分析方法预测了新疆蝗虫可能爆发的区域,得到较高的预测精度。杨会枫等^[23]使用 MaxEnt 模型预测了西伯利亚蝗在新疆的潜在分布,确定了影响西伯利亚蝗分布的重要环境因子,得出模型精度较高。李培先等^[24]基于 MaxEnt 模型预测了未来不同气候变化情景下新疆意大利蝗潜在分布,指出极端水分条件和水热条件对其潜在分布影响巨大。因不同物种分布模型自身差异,各模型算法对模型不确定性的影响会进一步影响到模型预

测结果的准确度^[25]。基于集成建模思想发展起来的利用多个模型信息的集合模型可以有效降低单一模型预测的不确定性,从而提升模型的稳定性及预测结果的可靠性^[26]。当前也有大量的文献应用多个模型集成对不同生物物种的分布做了预测,如韩欣尧等^[27]通过多个物种分布模型的集合对草地贪夜蛾的潜在分布区做了预测,Lippi 等^[28]应用多个物种分布模型对美国犬蜚的分布区做了预测。

前人的研究集中于单一类型蝗虫在不同气候条件下的潜在适生区展开,但各类主要蝗虫对地理环境、草地类型、气候状况、土壤条件等环境的适宜性存在一致性,因此文献^[29]将新疆草原蝗虫发生区分为八大亚区;此外前人针对草原的相关研究主要通过单一物种分布模型开展,但单一模型预测的不确定性较高,而通过多模型集成预测的方法则受到了很大的关注,相关研究成果也逐步增多;且相聚焦于与蝗虫生长密切相关的气候因素、植被等因素展开,影响蝗虫产卵、繁殖关系密切的土壤环境研究有待深入。基于此,本文综合考虑了新疆草原优势蝗种,而非某单一类型蝗虫;而且集成建模可以有效降低预测不确定性、提升模型的可靠性^[30];研究表明土壤因素对蝗虫的繁殖、虫卵孵化过程具有显著影响^[31—32]。

本文以与蝗虫活动紧密相关的环境数据为基础,应用经典物种分布模型对新疆典型蝗虫的潜在适生区进行研究,探究模型在新疆典型蝗虫预测研究中的性能表现;选择高性能模型构建草原蝗虫预测模型集合对新疆典型蝗虫的适生区展开预测,以期为一线蝗虫监测工作提供可靠的技术支持。

1 研究区概况

新疆位于亚欧大陆腹地,处于我国北方干旱半干旱区,地理位置在东经 $73^{\circ}40'$ — $96^{\circ}23'$,北纬 $34^{\circ}25'$ — $49^{\circ}10'$,面积 166 万 km^2 ,如图 1。其地形特征是“三山夹两盆”——自北向南依次为阿尔泰山、准噶尔盆地、天山、塔里木盆地和昆仑山。因深居内陆,远离海洋而形成典型的干旱气候特征,年均气温 10°C ,年均降水量 167.1 mm,降水空间差异大,北部多于南部,由此形成多种草地类型。草场面积占自治区国土面积一半以上,是当地农牧业重要的生产资料^[33]。

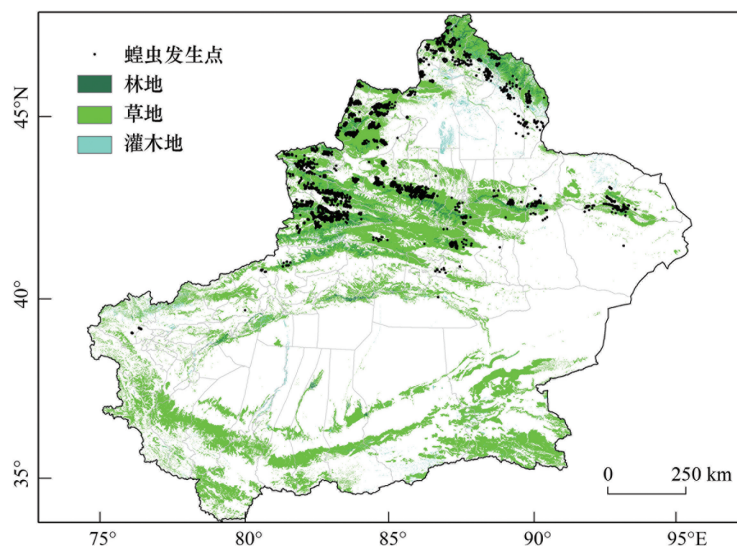


图 1 研究区位置及新疆典型蝗虫发生监测点

Fig.1 The location of the study area and the typical locust occurrence point in Xinjiang

2 数据与方法

2.1 数据来源

2.1.1 蝗虫发生数据

使用由新疆维吾尔自治区蝗虫鼠害预测预报防治中心站提供的全疆范围内各类蝗虫发生历史记录,并通

过收集文献^[23-24,33-36]及全球生物多样性信息网络数据库(GBIF)中新疆典型蝗虫数据记录,此外还有研究团队长期在野外调查中采集记录的蝗虫发生区 GPS 数据。原始数据记录共计 3800 条,通过筛选及预处理得到 3714 条数据参与后期处理与分析。

2.1.2 气象数据

气象数据包括与新疆典型草原蝗虫繁殖、生长密切相关的气温降水数据,其中气温数据为月最高气温和月最低气温。气象数据空间分辨率为 2.5', 该数据时间范围基本与所收集的蝗虫发生记录相一致,为 2000—2018 年。该数据集由 WorldClim(<https://www.worldclim.org/data/monthlywth.html>)提供,数据包括平均最低温度、平均最高温度和总降水量 3 个变量。通过处理与合成生成多年平均最低温度、多年平均最高温度和多年总降水量数据。

2.1.3 土壤数据

土壤湿度数据来自全球土地数据同化系统(<https://disc.gsfc.nasa.gov/>),使用空间分辨率为 0.25°的表层土壤湿度(0—10 cm)数据,该数据时间范围与气象数据一致。采用表层土壤酸碱度数据(0—30 cm),获取自中国科学院南京土壤研究所(<http://www.soilinfo.cn/>),其空间分辨率为 1 km。

2.1.4 植被数据

使用 MODIS NDVI 月度归一化植被指数(MOD13A3)表征研究区植被生长状况,数据时间范围与气象数据保持一致,空间分辨率为 1km,通过最大值合成方法,形成多年植被指数平均值。

2.1.5 地形数据

本研究使用的地形数据为数字高程模型(DEM),数据获取自中国科学院地理科学与资源研究所资源环境科学与数据中心(<http://www.resdc.cn/>),数据空间分辨率为 500 m,使用 ArcGIS 软件生成研究区坡度、坡向数据。

根据实际状况,在开展数据分析和模型构建前,使用 ArcGIS 软件对所有环境因子数据重投影为 WGS—84 坐标,根据新疆行政区划数据(由国家基础地理信息中心发布的国家基础地理信息数据库提供)裁剪出研究区数据,将所有数据重采样至 1 km,并将处理结果统一转为 Geo TIFF 格式。

2.2 研究方法

为避免环境因子间共线性问题导致模型稳健性降低,采用相关性分析方法和方差膨胀因子进行筛选。Green 团队推荐相关系数($r=0.7$)为最佳阈值^[37]。由于相关性分析有时存在非显性的相关结构,学者^[38]提出使用方差膨胀系数度量多重共线性影响的严重程度,一般认为当方差膨胀系数高于 10 时,则模型会存在较为严重的多重共线性问题^[39]。移除方差膨胀系数最大且相关性高于 0.7 的环境变量,使用其他变量构建模型。

物种分布模型在各领域获得了很高的评价,如 BIOCLIM、DOMAIN 及最大熵(MaxEnt)等经典模型^[40]。但由于数据量的增大及各类模型应用场景差异,通过集成建模可降低模型选择中的误差,同时提供所有候选模型中各变量的相对重要性^[41]。集成建模分简单平均和加权平均两种方法^[20],其中加权平均基于单个模型的精度系数(AUC)分配权重^[42],如式 1:

$$WAp_i = \frac{\sum_j (AUC_j \times p_{ij})}{\sum_j AUC_j} \quad (1)$$

式中, WAp_i 为 i 的集成模型预测, AUC_j 为单一模型的 AUC 值; p_{ij} 为各单一模型对 i 的预测值。

本文使用 R 语言 dismo(<https://cran.r-project.org/web/packages/dismo/index.html>)包实现各典型模型的预测,之后使用加权平均方法实现集成模型对新疆典型蝗虫适生区预测。

众多模型评价方法中接受者操作特性曲线(ROC)曲线下面积(AUC)以及真实技巧统计(TSS)得到了广泛的应用,AUC 将所有阈值范围的值集成到一个单一最终指标中,而非在所有阈值上寻找度量的最大值^[43];TSS 是对 Kappa 系数优点的继承和对物种分布模型响应曲线不足的克服,通过敏感性和特异性度量^[44]。Araujo 等^[45]认为 AUC 低于 0.5 为反预测(即观测值与预测值间存在负相关),AUC 低于 0.6 为失败,低于 0.7

为差,低于 0.8 为可接受,低于 0.9 为良好,高于 0.9 为优秀;张雷等^[46]认为 TSS 低于 0.4 为失败,低于 0.55 为一般,低于 0.70 为好,低于 0.85 为很好,高于 0.85 为极好。

3 结果与分析

3.1 环境变量筛选

通过相关分析得出变量间存在不同相关性,其中土壤湿度与降水变量间相关性最高(0.5027),而坡度与坡向间相关性最小(0.0021),相关性最高的变量未达到最佳阈值。将方差膨胀系数的相关性阈值设定为 0.7,再次对所有变量进行筛选,得出月最低气温变量及地形变量与其他变量间共线性问题更为突出,如表 1,因此予以剔除。在新疆典型蝗虫适生区预测中使用除月最低气温数据及地形外的其余七个环境因子参与建模。

表 1 各环境变量方差膨胀系数
Table 1 VIF of each environmental variable

环境变量 Environment Variables	方差膨胀系数 Variance inflation factor	环境变量 Environment Variables	方差膨胀系数 Variance inflation factor
坡度 Aspect	1.0017	土壤湿度 Soil moisture	1.6610
归一化植被指数 Normalized difference vegetation index	1.1475	土壤酸碱度 Soil pH	1.1876
降水 Precipitation	2.0309	月最高气温 Monthly maximum temperature	1.8300
坡向 Slope	1.3369		

3.2 模型精度评价

将发生监测点数据分为训练集和验证集两部分,其中验证样本集通过随机选取 20% 样点得到。利用 AUC 验证各预测模型的性能得出均在可接受的范围以内,如图 2 和表 2 所示,其中 DOMAIN 模型的 AUC 值和 TSS 值均最低,分别为 0.688 和 0.301;提升回归树模型最高,分别为 0.920 和 0.910。

表 2 各模型真实技巧统计值
Table 2 TSS of each model

模型 Model	TSS 值 TSS Value	阈值 Threshold	模型 Model	TSS 值 TSS Value	阈值 Threshold
BIOCLIM	0.538	0.57	RF	0.725	0.72
DOMAIN	0.301	0.55	BRT	0.910	0.51
MAHAL	0.871	0.12	SVM	0.755	0.69
GLM	0.545	0.43	MaxEnt	0.893	0.36

TSS:真实技巧统计值 True skill statistics; BIOCLIM: Bioclim 模型 Bioclim; DOMAIN: 领域模型 Domain; MAHAL: 马氏距离模型 Mahalanobis distance; GLM: 广义线性模型 Generalized linear model; RF: 随机森林模型 Random forest; BRT: 提升回归树模型 Boosted regression trees; SVM: 支持向量机模型 Support vector machine; MAXENT: 最大熵模型 Maximum entropy

3.3 模型预测结果

首先使用单一模型对新疆典型蝗虫的适生区进行预测,得出高适宜性区域集中于天山以北区域,但预测结果存在一定差异,其中应用 BIOCLIM 模型得出适生区面积最小,DOMAIN 模型得出全疆大范围内均为高度适生区,如图 3。地理分布上,高度适生区集中于北疆各大草原及天山山麓、阿尔泰山附近;行政区划上,则集中于阿勒泰地区、塔城地区、博乐地区、石河子市、昌吉州、乌鲁木齐市一带,此外伊犁州、巴音郭楞蒙古自治州、克孜勒苏柯尔克孜自治州及哈密部分地区也有分布。

由于各模型性能存在差异,为提高预测模型可靠性和稳健度,本研究将模型 AUC 和 TSS 值均更高的提升回归树、支持向量机和最大熵等 3 个高精度模型进行集成,以 3 个模型的加权平均最终预测新疆典型蝗虫适生区,如图 4。通过集成模型预测结果与单一模型相似,但通过集成模型预测的结果结合了三种模型的优势,综合各模型在不同区域的预测结果,从而使得预测的区域达到更合理的范围。在对三种典型蝗虫统一预测的

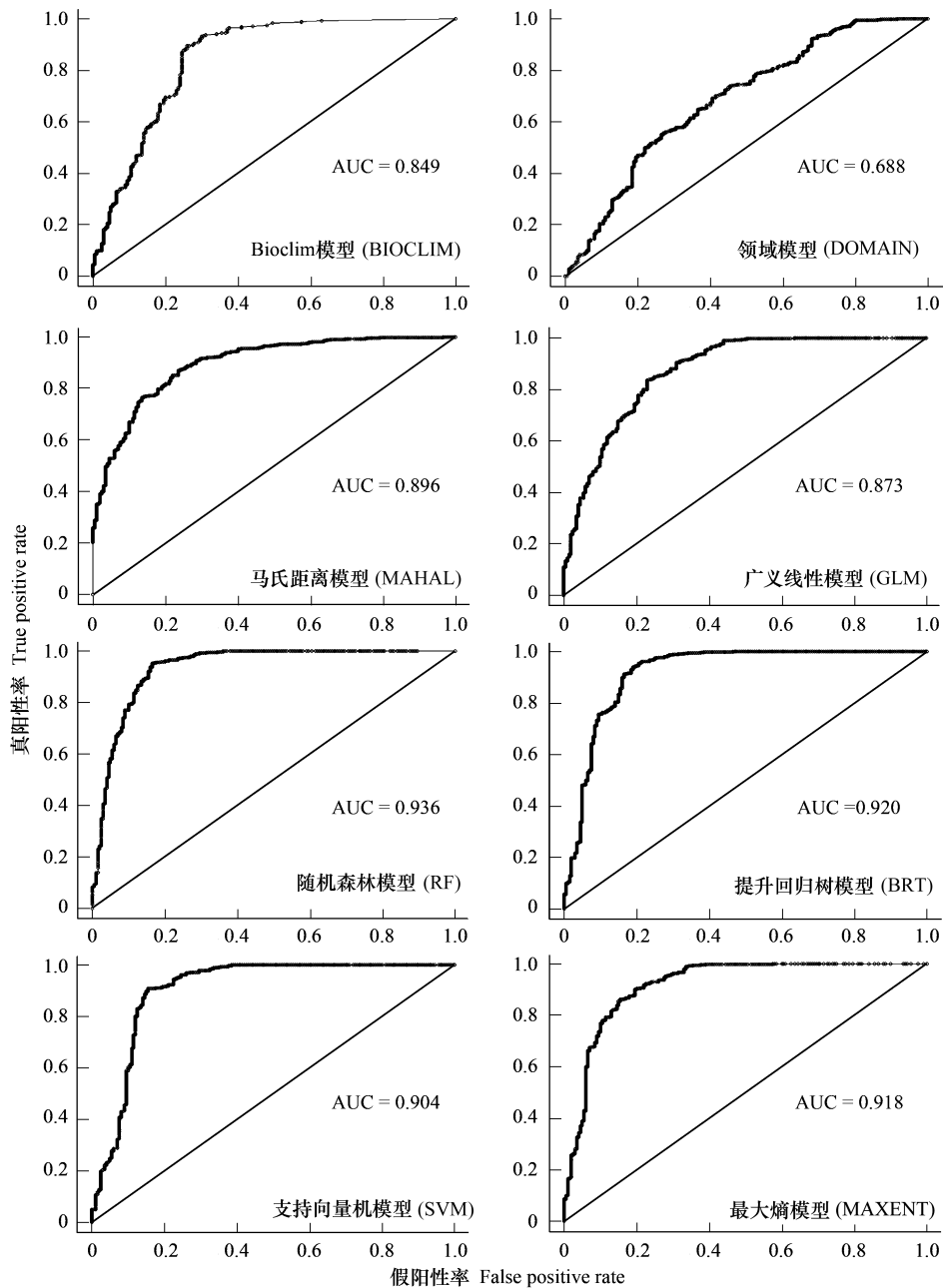


图2 各模型接受者操作特性曲线
Fig.2 Receiver operation characteristic curve of each model
AUC:曲线下面积 Area under curve

基础上,运用集成模型分别针对三种典型蝗虫进行适生区的预测,有利于认识到不同类型蝗虫在全疆各地的地理分布特征和规律,以便在后期工作中对单一蝗虫采取相对应的监测与防控措施。作为新疆典型蝗虫适生区集成预测的重要基础,也分别对亚洲飞蝗、意大利蝗和西伯利亚蝗在全疆的适生区进行了模拟预测。三类蝗虫的适生区具有较大比例的重合区,从图中看,蝗虫的高度适生区基本都集中在新疆北部天山段及伊犁河谷等区域,对三类蝗虫适生区的分别预测将会对蝗虫的监测、防治等各项工作的开展提供便利和支持。

使用 ArcGIS 自然间断法,将集成模型预测的新疆典型蝗虫适生区划分为 5 个等级^[24],如表 3。根据划分标准对集成模型预测结果分类,无论是新疆典型草原蝗虫还是亚洲飞蝗、意大利蝗和西伯利亚蝗单一类型蝗

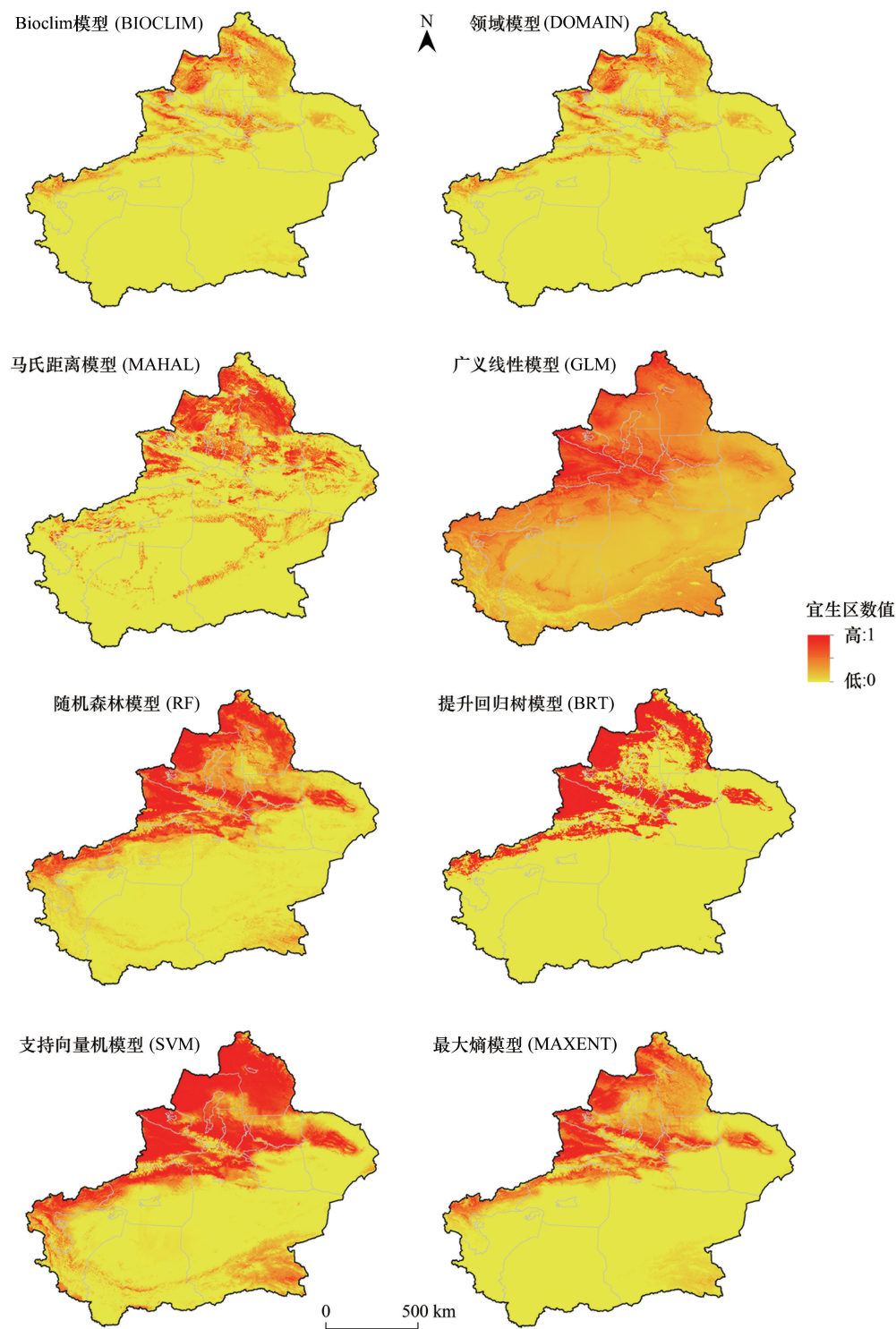


图 3 各模型对典型蝗虫预测结果
Fig.3 Prediction result of each model for typical locusts

虫优势种而言,北疆各地州高适生区面积占比较大,南疆各地州则相对较少。而在北疆的地理分布则主要集中于天山一带、阿尔泰山一带、东部的吐哈盆地和西部的伊犁河谷一带,南疆的适生区则呈带状分布于塔里木盆地南缘的绿洲一带,如图 5。

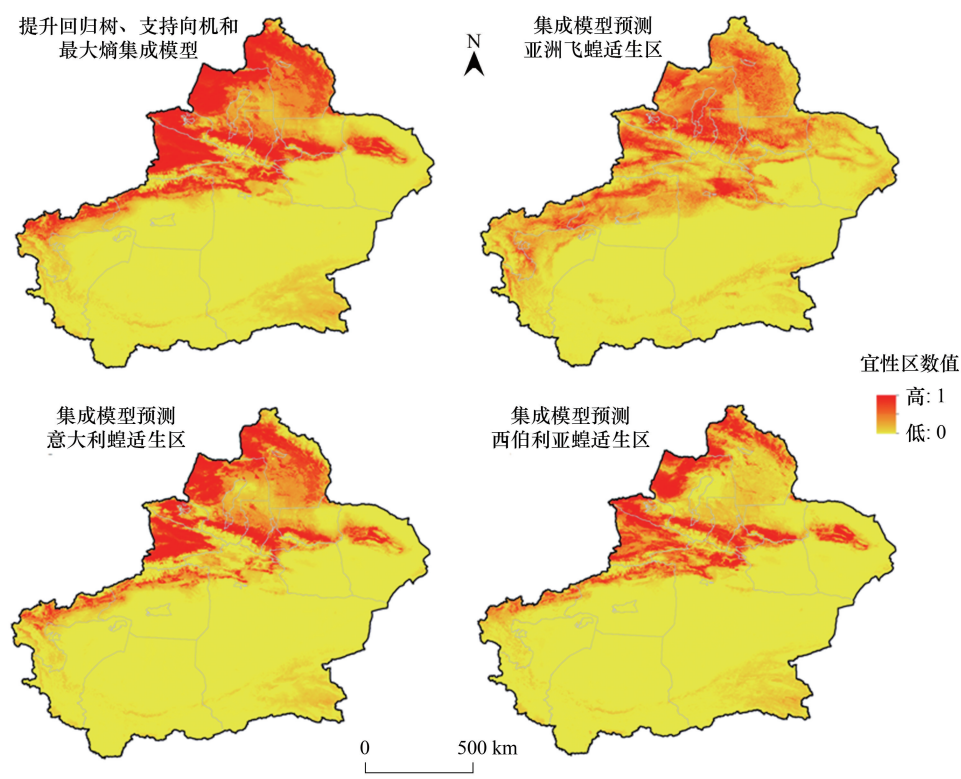


图 4 集成模型对不同蝗虫类型预测结果

Fig.4 Prediction results of different locust types by BRT,SVM、MaxEnt

表 3 新疆典型蝗虫适生区划分

Table 3 The division of suitable area of Xinjiang typical locust

分布概率(<i>P</i>) Probability distribution	适宜性等级 Suitability grade	分布概率(<i>P</i>) Probability distribution	适宜性等级 Suitability grade
$0 < P < 0.08$	非适生区	$0.49 \leq P < 0.75$	适生区
$0.08 \leq P < 0.26$	低适生区	$0.75 \leq P < 1.0$	高危区
$0.26 \leq P < 0.49$	边缘适生区		

统计新疆典型蝗虫不同适生区分区面积和比例,如表 4,各类适生区(低适生区、边缘适生区、适生区、高危区)面积共约 59.844 万 km²,占新疆所有面积的 36%。其中低适生区面积为 28.284 万 km²,所占比例最大,占有所有适生区面积的近二分之一(47%),占有总面积的 17%;适生区面积 1.735 万 km²,所占比例最小,仅为不到百分之一(0.87%);边缘适生区所占面积较低适生区低,但也高于其他适生区分区;而高危区面积为 16.568 万 km²,所占面积仅次于低适生区,所占比例也高达 10%。

表 4 不同适生区分区面积及占比

Table 4 Area and proportion of different suitable areas

适生区分区 Suitable area division	面积 Area/(万 km ²)	占比 Proportion/%	适生区分区 Suitable area division	面积 Area/(万 km ²)	占比 Proportion/%
非适生区 Unsuitable	105.436	63.637	适生区 Medium	1.735	0.817
低适生区 Low	28.284	17.071	高危区 High	16.568	10.472
边缘适生区 Marginal	13.257	8.002			

3.4 环境因子重要性

新疆各典型蝗虫对不同环境因子具有不同耐受性,本文综合考虑影响意大利蝗、西伯利亚蝗和亚洲飞蝗

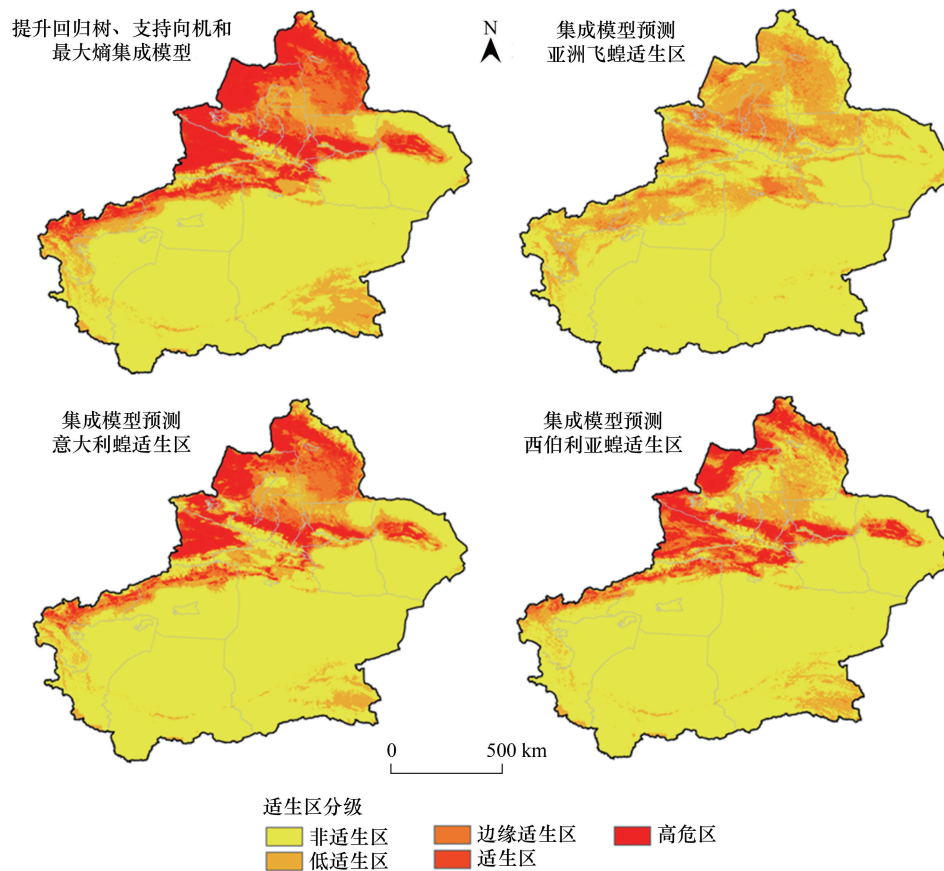


图 5 新疆典型蝗虫适生区分级

Fig.5 Typical locust Distribution in Xinjiang Classified

生存与繁殖等活动的重要生态环境变量,各环境变量对预测新疆蝗虫适生区分布的贡献率如表 5。综合而言,在新疆三类典型蝗虫适生区预测的建模中,不同环境变量在各主要预测模型中的贡献率存在较大差异性,但是各个环境变量贡献率的排名整体上较为一致。具体而言,从 3 个预测模型的表现上看,在对三类蝗虫预测中,降水、归一化植被指数及月最高气温 3 个变量始终是对蝗虫适生区预测贡献最大的几类,其中在提升回归树模型中,对亚洲飞蝗、意大利蝗和西伯利亚蝗适生区预测贡献率最高的变量均为降水,其表现分别为 20.8%、53.7%和 48.2%;支持向量机模型中,对亚洲飞蝗、意大利蝗和西伯利亚蝗适生区预测贡献率最高的变量分别为土壤酸碱度、月最高气温和降水,其表现分别为 20.13%、26.28%和 23.55%;最大熵模型中,对亚洲飞蝗、意大利蝗和西伯利亚蝗适生区预测贡献率最高的变量均为降水,其表现分别为 31%、62.2%和 67.3%。

土壤湿度和土壤酸碱度及地形因子也同样对新疆草原典型蝗虫的适生区分布有较大影响,而坡向则在新疆典型蝗虫发生和分布中的影响较小。环境变量在各预测模型分别针对三类典型蝗虫适生区预测中的贡献率存在一定差别,但根据 3 个模型的综合表现月最高气温、降水和植被 3 个变量始终是蝗虫生长和繁育的重要且直接的环境因素,对于其生境的选择具有很大的影响。

4 讨论与结论

4.1 讨论

物种分布模型准确预测基础之一是环境变量筛选,研究表明^[47]环境变量会对模型预测精度造成较为显著的影响。当各环境变量间存在较高相关性时,会难以避免地带来相关和多重共线性等问题,从而在建模中造成不必要的信息冗余,影响模型的稳健性。本研究采用相关性及方差膨胀系数移除月最低气温变量构建

预测模型,这与众多学者在相关研究中采用的变量筛选方法基本一致,如学者 Pradhan^[48]使用皮尔逊相关系数对环境变量进行了相关性分析,其后计算方差膨胀系数,对南亚和东南亚的生物气候变量进行了多重共线性和冗余的筛选。

表 5 环境变量贡献率

Table 5 Contribution rate to environmental variables

环境变量 Environment variables		蝗虫类型/预测模型 Type of Locust/Model								
蝗虫类型 Type of Locust		亚洲飞蝗			意大利蝗			西伯利亚蝗		
预测模型 Model		BRT	SVM	MaxEnt	BRT	SVM	MaxEnt	BRT	SVM	MaxEnt
坡度 Aspect		10.9	8.31	6.6	2.3	24.96	4.6	8.5	6.03	9.7
归一化植被指数 Normalized difference vegetation index		11.5	18.78	2.8	13.9	9.54	13.5	9.6	16.92	5.5
降水 Precipitation		20.8	13.39	31	53.7	18.83	62.2	48.2	23.55	67.3
坡向 Slope		10.8	18.02	11.8	3.2	1.09	0.5	3	3.16	0.5
土壤湿度 Soil moisture		19.4	10.54	25.7	4.8	13.74	2.4	6.3	18.39	4.5
土壤酸碱度 soil pH		17.1	20.13	2.6	3.2	5.56	1.1	7.9	10.12	1.8
月最高气温 Monthly maximum temperature		9.5	10.83	19.5	18.9	26.28	15.7	16.5	21.83	10.7
总计 Total		100	100	100	100	100	100	100	100	100

对模型性能度量的重要标准之一是接受者操作特性曲线(即 AUC 值),AUC 受阈值影响较低,可对所有阈值范围内的性能评价结果予以显示,因此也受到生态环境领域学者重视^[49]。AUC 值越接近 1 即表明模型的性能越好,本研究中各模型的 AUC 基本都保持在 0.7 以上,说明各模型的性能可较好地预测新疆典型蝗虫的适生区,但模型间存在一定差异,其中 DOMAIN 性能相对最差,对新疆典型蝗虫的预测能力最低;而提升回归树、支持向量机、最大熵 3 个模型的 AUC 均大于 0.9,其性能在新疆典型蝗虫的预测中高于其他各模型,这也与相关研究中各模型在不同物种分布预测模型中的性能表现基本一致^[50]。具体如 Duan 等^[51]使用 BIOCLIM、DOMAIN、MAHAL、RF、MaxEnt 和 SVM 等六个模型预测了五种树种的潜在分布区,以标准差、变异系数、Kappa 值和 AUC 值评估模型的稳定性,指出 MAHAL、RF、MaxEnt 和 SVM 具有更高预测精度。

物种分布模型模拟出物种生存的可能性空间依赖于基本生态位的稳定性及空间分布的平衡^[52]。通过物种分布模型预测的物种潜在适生区与真实适生区之间存在一定差异,但预测的潜在适生区基本覆盖了其实际的分布状况^[53]。本文参考陈健等^[54]关于蝗虫生境研究应用进展中不同生境因子对蝗虫影响的研究成果,结合新疆草原典型蝗虫具体实际,综合考虑了对蝗虫有显著影响的气温、地形、植被和土壤四个方面的因素,使预测的潜在适生区较真实蝗虫适生区范围最大化^[55]。在新疆典型蝗虫适生区预测中各生境因子对预测模型的贡献率表现差异显著,其中降水量对蝗虫适生区的影响尤为显著,这与王晗等^[56]在新疆意大利蝗适生区与降水量、最高气温和最低气温的变化相关研究中相关的结论有一定的关联性。因本文选取的环境变量并非在不同气候情境下的生物气候变量,预测变量的贡献情况与李培先^[24]等关于意大利蝗潜在分布中各变量贡献率和杨会枫^[23]关于西伯利亚蝗在新疆潜在分布中各变量贡献率存在一定差异。

本文对新疆各类典型蝗虫的适生区预测结果显示,典型蝗虫的适生区具有显著的区域差异,各类适生区面积近 60 万 km²,适生区地理分布主要集中在新疆天山一带、北部阿尔泰山及伊犁河谷一带,此外在东部的吐哈盆地和南疆塔里木盆地边缘也有适生区,这些基本与文献有关新疆蝗虫地理分布的表述基本一致。同时本文预测的适生区地理分布及其面积也与郑淑丹对 2012 年新疆草地蝗虫发生分布拟合结果一致,但本文预测结果无论是从预测方法上还是预测结果上均有了较大的提升和优化;更近一步,与杨会枫等人使用 MaxEnt 模型对西伯利亚蝗潜在分布区的预测结果同样具有相似性,这也印证了本文构建的新疆典型蝗虫预测模型及其预测结果对新疆典型蝗虫适生区预测存在一定合理性和可靠性。

本文基于新疆典型蝗虫发生数据和相关环境变量,使用经典物种分布模型对新疆典型蝗虫适生区进行了

预测,但还存在一些不足:(1)在环境变量选择方面未全面考虑人类活动等非自然因素影响,因此可能对预测结果造成一定影响;(2)仅选用部分经典物种分布模型进行了预测,暂未将如生态位因素分析模型(ENFA)、广义加性模型(GAM)、人工神经网络(ANNs)等其他具有同样代表性的模型考虑在内,可能也会对更精准预测结果的展示产生一定影响。(3)仅对新疆全区尺度典型蝗虫适生区进行了预测,未对地州甚至区县尺度状况展开,另一方面也未对新疆周边省区以及国外地区展开研究,在未来条件允许的情况下也可考虑在不同尺度下推进。总之,本研究对新疆典型蝗虫适生区预测一方面于区域相关研究深入开展具有一定探索意义,另一方面于基层实际蝗虫监测、防治工作的推进亦可提供一定建议和支持。

4.2 结论

基于物种分布模型对新疆典型蝗虫适生区预测,得出:

(1)新疆典型蝗虫适生区受多种环境因素共同作用,各因素相互影响对新疆蝗虫适生区影响存在差异。本研究选取的众多影响蝗虫繁殖、生长等不同生命活动阶段的环境因子中,月最低气温变量及地形相较于其他变量对蝗虫适生区的影响最小。降水、气温和植被对物种分布预测模型的贡献率最高,此外土壤湿度和土壤酸碱度对新疆典型蝗虫适生区预测也有较高贡献率。

(2)各预测模型性能存在显著差异,其中 DOMAIN 模型在新疆典型蝗虫适生区预测中性能最低($AUC = 0.688$, $TSS = 0.301$),而提升回归树模型的预测性能最佳($AUC = 0.920$, $TSS = 0.910$);综合预测性能最优的 3 个经典模型集成最终预测模型对新疆典型蝗虫潜在适生区的预测具有更高的准确性与合理性。

(3)通过集成预测模型的预测,新疆典型蝗虫的潜在适生区分布具有区域性,其中天山以北的广大草原地区存在各类蝗虫适生区,此外在东部的哈密地区、南疆各绿洲边缘地带亦有零星分布;各类适生区面积 56.844 万 km^2 ,其中高危区面积 16.568 万 km^2 。

(4)新疆草原典型蝗虫适生区存在高度集中、区域差异显著、地理分布结构复杂等多种特点,不同区域、不同地州草原部门对蝗虫的日常监测和防治工作需紧密结合当地工作实际,分区、分时制定严密的蝗虫日常监测和防治计划,在做好本地蝗虫监测防治工作的同时,各边境地区草原部门还需做好跨境蝗虫的迁飞入境的监测、预防工作。

致谢:自治区级产学研联合培养研究生基地——新疆维吾尔自治区治蝗灭鼠指挥部办公室提供蝗虫发生点数据,国家基础地理信息数据库提供本文所有基础底图数据。

参考文献(References):

- [1] 喻锋,李晓波,张丽君,徐卫华,符蓉,王宏. 中国生态用地研究:内涵、分类与时空格局. 生态学报, 2015, 35(14): 4931-4943.
- [2] 孙启忠,玉柱,徐春城. 我国草业新世纪 10 年取得的成就和未来 10 年发展的重点. 草业科学, 2011, 28(12): 2215-2220.
- [3] 严杜建,周启武,路浩,吴晨晨,赵宝玉,曹丹丹,马烽,刘晓学. 新疆天然草地毒草灾害分布与防控对策. 中国农业科学, 2015, 48(3): 565-582.
- [4] 赵玲,王登元,赵莉,肖宏伟. 新疆蝗虫的研究进展. 新疆农业科学, 2012, 49(7): 1212-1223.
- [5] 熊玲. 新疆草原以生物防治为主的蝗虫综合防治技术应用. 新疆畜牧业, 2011(3): 59-63.
- [6] 陈永林. 新疆维吾尔自治区的蝗虫研究蝗虫的分布(续). 昆虫学报, 1981, 24(2): 166-173, 238.
- [7] 于冰洁,陈吉军,季荣. 新疆蝗虫及其监控技术研究进展. 应用昆虫学报, 2019, 56(5): 927-933.
- [8] 徐超民,王加亭,李霜,刘路路,姚贵敏,伟军,王坤芳,王文成,涂雄兵,张泽华. 蝗虫综合防控技术研究进展. 植物保护学报, 2021, 48(1): 73-83.
- [9] 王冬梅. 温度对意大利蝗呼吸代谢及取食和交配行为的影响研究[D]. 乌鲁木齐:新疆师范大学, 2016.
- [10] 赵紫华,涂雄兵,张泽华,李志红. 警惕沙漠蝗种群持续增加和入侵我国边境地区的风险. 植物保护学报, 2021, 48(1): 5-12.
- [11] Guan J Y, Li M Y, Ju X F, Lin J, Wu J G, Zheng J H. The potential habitat of desert locusts is contracting: predictions under climate change scenarios. PeerJ, 2021, 9: e12311.
- [12] Elith J, Leathwick J R. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics, 2009, 40: 677-697.

- [13] 王茹琳, 李庆, 封传红, 石朝鹏. 基于 MaxEnt 的西藏飞蝗在中国的适生区预测. 生态学报, 2017, 37(24): 8556-8566.
- [14] Booth T H, Nix H A, Busby J R, Hutchinson M F. Bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. Diversity and Distributions, 2014, 20(1): 1-9.
- [15] Huntley B, Barnard P, Altwegg R, Chambers L, Coetzee B W T, Gibson L, Hockey P A R, Hole D G, Midgley G F, Underhill L G, Willis S G. Beyond bioclimatic envelopes: dynamic species' range and abundance modelling in the context of climatic change. Ecography, 2010; 33(3): 621-626.
- [16] Guisan A, Zimmermann N E. Predictive habitat distribution models in ecology. Ecological Modelling, 2000, 135(2/3): 147-186.
- [17] Yu H, Cooper A R, Infante D M. Improving species distribution model predictive accuracy using species abundance: application with boosted regression trees. Ecological Modelling, 2020, 432: 109202.
- [18] Araújo M B, New M. Ensemble forecasting of species distributions. Trends in Ecology & Evolution, 2007, 22(1): 42-47.
- [19] Dormann C F, Calabrese J M, Guillera-Arroita G, Matechou E, Bahn V, Bartoń K, Beale C M, Ciuti S, Elith J, Gerstner K, Guelat J, Keil P, Lahoz-Monfort J J, Pollock L J, Reineking B, Roberts D R, Schröder B, Thuiller W, Warton D I, Wintle B A, Wood S N, Wüest R O, Hartig F. Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. Ecological Monographs, 2018, 88(4): 485-504.
- [20] Hao T X, Elith J, Guillera-Arroita G, Lahoz-Monfort J J. A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Diversity and Distributions, 2019, 25(5): 839-852.
- [21] 杨洪升, 季荣, 王婷. 新疆蝗虫发生的大气环流背景及长期预测. 生态学杂志, 2008, 27(2): 218-222.
- [22] Zheng S D, Zheng J H, Mu C, Ni Y F, Dawuti B, Wu J G. GIS-based multi-criteria analysis model for identifying probable sites of locust outbreak in Xinjiang, China. 2015 23rd International Conference on Geoinformatics. June 19-21, 2015, Wuhan. IEEE, 2015: 1-6.
- [23] 杨会枫, 郑江华, 吴秀兰, 穆晨, 林俊, 许仲林. 基于 MaxEnt 模型的西伯利亚蝗虫在新疆潜在分布预测研究. 新疆农业科学, 2016, 53(1): 43-50.
- [24] 李培先, 林峻, 麦迪·库尔曼, 吴建国, 刘永强, 郑江华. 气候变化对新疆意大利蝗潜在分布的影响. 植物保护, 2017, 43(3): 90-96.
- [25] Thibaud E, Petitpierre B, Broennimann O, Davison A C, Guisan A. Measuring the relative effect of factors affecting species distribution model predictions. Methods in Ecology and Evolution, 2014, 5(9): 947-955.
- [26] 郭彦龙, 赵泽芳, 乔慧捷, 王然, 卫海燕, 王璐坤, 顾蔚, 李新. 物种分布模型面临的挑战与发展趋势. 地球科学进展, 2020, 35(12): 1292-1305.
- [27] 韩欣娆, 朱耿平, 门永亮, 杨琢. 基于集合模型的草地贪夜蛾的潜在分布预测. 生物安全学报, 2021, 30(1): 65-71.
- [28] Lippi C A, Gaff H D, White A L, St John H K, Richards A L, Ryan S J. Exploring the niche of *Rickettsia montanensis* (Rickettsiales: Rickettsiaceae) infection of the American dog tick (Acari: Ixodidae), using multiple species distribution model approaches. Journal of Medical Entomology, 2020, 58(3): 1083-1092.
- [29] 全国畜牧总站. 中国草原蝗虫生物防治实践与应用. 北京: 中国农业出版社, 2014.
- [30] 刘尊雷, 杨林林, 袁兴伟, 金艳, 严利平, 程家骅. 基于集成模型的小黄鱼越冬群体适宜生境及其环境影响因素. 应用生态学报, 2020, 31(6): 2076-2086.
- [31] 刘琼, 何立志, 张永军, ROMAN Jashenko, 季荣. 中哈边境蝗区蝗虫孳生和发生地的重要生态学特征. 环境昆虫学报, 2017, 39(2): 365-371.
- [32] 张艳, 李红梅, 刘路路, 王广君, 尚素琴. 不同土壤类型及其含水量对亚洲小车蝗卵孵化的影响. 环境昆虫学报, 2020, 42(3): 559-565.
- [33] 郑淑丹. 基于 MCA 和马尔可夫链新疆的草地蝗虫灾害时空分布预测模型研究[D]. 乌鲁木齐: 新疆大学, 2015.
- [34] 范福来, 王元信. 亚洲飞蝗在中国新疆维吾尔自治区的发生与防治. 生态学报, 1995, 15(2): 134-141.
- [35] 罗迪, 刘孝贤, 季荣. 气候变暖背景下新疆中哈边境亚洲飞蝗适生区预测及变化[J/OL]. 环境昆虫学报; 1-15 [2022-02-24]. <http://kns.cnki.net/kcms/detail/44.1640.Q.20210720.1326.006.html>.
- [36] 族米娜, 牙森·沙力, 阿孜古丽·阿布力孜, 徐光青, 古丽孜热·曼合木提, 涂雄兵. 新疆部分生态区亚洲飞蝗地理种群形态变异研究. 新疆农业科学, 2018, 55(5): 888-900.
- [37] Kempton R A, Green R H. Sampling design and statistical methods for environmental biologists. Biometrics, 1981, 37(1): 202.
- [38] Hair J F, Black B, Babin B J, Anderson R. Multivariate Data Analysis//: Multivariate Data Analysis, 2011.
- [39] Harrell F E Jr, Lee K L, Mark D B. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine, 1996, 15(4): 361-387.
- [40] 段义忠, 王驰, 王海涛, 杜忠毓, 贺一鸣, 柴乖强. 不同气候条件下沙冬青属植物在我国的潜在分布——基于生态位模型预测. 生态学报, 2020, 40(21): 7668-7680.
- [41] Lauzeral C, Grenouillet G, Brosse S. Dealing with noisy absences to optimize species distribution models: an iterative ensemble modelling approach.

- PLoS One, 2012, 7(11): e49508.
- [42] Hao T X, Elith J, Lahoz -Monfort J J, Guillerá -Arroita G. Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography*, 2020, 43(4): 549-558.
- [43] Fielding A H, Bell J F. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 1997, 24(1): 38-49.
- [44] Allouche O, Tsoar A, Kadmon R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 2006, 43(6): 1223-1232.
- [45] Araújo M B, Pearson R G, Thuiller W, Erhard M. Validation of species— climate impact models under climate change. *Global Change Biology*, 2005, 11(9): 1504-1513.
- [46] 张雷, 刘世荣, 孙鹏森, 王同立. 气候变化对物种分布影响模拟中的不确定性组分分割与制图——以油松为例. *生态学报*, 2011, 31(19): 5749-5761.
- [47] 陈新美, 雷渊才, 张雄清, 贾宏炎. 样本量对 MaxEnt 模型预测物种分布精度和稳定性的影响. *林业科学*, 2012, 48(1): 53-59.
- [48] Pradhan P. Strengthening MaxEnt modelling through screening of redundant explanatory Bioclimatic Variables with Variance Inflation Factor analysis. *Researcher*, 2016, 8(5): 29-34.
- [49] Shabani F, Kumar L, Ahmadi M. Assessing accuracy methods of species distribution models: AUC, specificity, sensitivity and the true skill statistic. *Acta Scientiarum Human and Social Sciences*, 2018
- [50] 李白尼, 魏武, 马骏, 张润杰. 基于最大熵值法生态位模型 (Maxent) 的三种实蝇潜在适生性分布预测. *昆虫学报*, 2009, 52(10): 1122-1131.
- [51] Duan R Y, Kong X Q, Huang M Y, Fan W Y, Wang Z G. The predictive performance and stability of six species distribution models. *PLoS One*, 2014, 9(11): e112764.
- [52] 王文婷, 高思雨, 王淑璠. 甘肃草地 4 种毒杂草潜在入侵区预测研究. *生态学报*, 2019, 39(14): 5301-5307.
- [53] Sillero N. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, 2011, 222(8): 1343-1346.
- [54] 陈健, 王秀君. 遥感与 GIS 在蝗虫生境研究中的应用进展. *生态环境学报*, 2012, 21(5): 970-976.
- [55] 黄文江, 董莹莹, 赵龙龙, 耿芸, 阮超, 张弼尧, 孙忠祥, 张寒苏, 叶回春, 王昆. 蝗虫遥感监测预警研究现状与展望. *遥感学报*, 2020, 24(10): 1270-1279.
- [56] 王晗, 于非, 扈鸿霞, 季荣. 新疆意大利蝗适生区的气候变化特征分析. *中国农业气象*, 2014, 35(6): 611-621.