

DOI: 10.3969/j.issn.2095-1787.2018.02.007

# MaxEnt 模型参数设置对其所模拟物种地理分布和生态位的影响——以茶翅蝽为例

朱耿平\*, 原雪姣, 范靖宇, 王梦琳

天津师范大学生命科学学院, 天津市动植物抗性重点实验室, 天津 300387

**摘要:**【目的】生态位模型被广泛应用于入侵生物学和保护生物学研究, 现有建模工具中, MaxEnt 是最流行和运用最广泛的生态位模型。然而最近研究表明, 基于 MaxEnt 模型的默认参数构建模型时, 模型倾向于过度拟合, 并非一定为最佳模型, 尤其是在处理一些分布点较少的物种。【方法】以茶翅蝽为例, 通过设置不同的特征参数、调控倍频以及背景拟不存在点数分别构建茶翅蝽的本土模型, 然后将其转入入侵地来验证和比较模型, 通过检测模型预测的物种对环境因子的响应曲线、潜在分布在生态空间中的生态位映射以及潜在分布的空间差异性, 探讨 3 种参数设置对 MaxEnt 模型模拟物种分布和生态位的影响。【结果】在茶翅蝽的案例中, 特征参数的设置对 MaxEnt 模型所模拟的潜在分布和生态位的影响最大, 调控倍频的影响次之, 背景拟不存在点数的影响最小。与其他特征相比, 基于特征 H 和 T 的模型其响应曲线较为曲折; 随着调控倍频的增加, 响应曲线变得圆滑。【结论】在构建 MaxEnt 模型时, 需要从生态空间中考虑物种的生态需求, 分析模型参数对预测物种分布和生态位可能造成的影响。

**关键词:** 生态位模型; MaxEnt 模型; 特征参数; 调控倍频; 背景拟不存在点数

## Effects of model parameters in MaxEnt modeling of ecological niche and geographic distribution: case study of the brown marmorated stink bug, *Halyomorpha halys*

ZHU Gengping\*, YUAN Xuejiao, FAN Jingyu, WANG Menglin

Key Laboratory of Animal and Plant Resistance in Tianjin, College of Life Sciences,  
Tianjin Normal University, Tianjin 300387, China

**Abstract:** 【Aim】 Ecological niche modeling is widely used in studies of biological invasion and conservation. Among available models, MaxEnt is the most popular and widely used model. Recent studies, nonetheless, showed that MaxEnt models tend to overfitting, and it may not be the ideal model when using default setting, especially for species with small sampling size. 【Method】 We investigated the influence of three parameters (i.e. feature combination, regularization multiplier, and maximum number of background points) on MaxEnt modeling of ecological niche and geographic distributions using the marmorated stink bug (*Halyomorpha halys*, Heteroptera) as a case study. Native niche models were calibrated for *H. halys* in eastern Asia and transferred worldwide for comparison and evaluation. Response curves of the species to environmental variables, projections of potential distribution in ecological spaces, and the variation of these potential distributions were used to investigate the effects. 【Result】 Feature combinations had significant effects on MaxEnt modeling outcomes of ecological niches and geographic distributions, followed by the regularization multiplier; the maximum number of background points had the least effect. We found that response curves based on features of hinge and threshold were more complicated than that based on the others; the response curves became more smooth as the regularization multipliers increased. 【Conclusion】 Ecological requirement of the species in its ecological space should be considered when fitting a MaxEnt model, the influence of the Maxent parameters on modeling ecological niches and geographical distributions should also be considered, rather than simply based on default settings.

**Key words:** ecological niche model; MaxEnt; feature combination; regularization multiplier; background points

收稿日期 (Received): 2018-03-30 接受日期 (Accepted): 2018-05-03

基金项目: 国家自然科学基金项目 (31401962)

作者简介: 朱耿平, 男, 副研究员, 博士。研究方向: 入侵生物学和保护生物学

\* 通信作者 (Author for correspondence), E-mail: gengpingzhu@hotmail.com

近年来,生态位模型被广泛应用于入侵生物学、保护生物学、全球气候变化对物种分布的影响,谱系生物地理学及传染病空间传播研究等多个领域(朱耿平等,2013)。随着全球物种分布数据的共享和 GIS 技术的快速发展,相关性方案的生态位模型快速发展(朱耿平和乔慧捷,2016;朱耿平等,2017)。在现有建模工具中,MaxEnt 模型最流行,应用最广泛(Elith *et al.*,2011)。MaxEnt 模型基于最大熵原理,即在满足已知约束的条件下,选择熵最大的模型,它利用物种的存在分布点和环境变量,来推算物种的生态需求和模拟物种的潜在分布(Merow *et al.*,2013)。由于其简单直接的操作、简洁清晰的图形界面及参数自动配置的功能,被广大研究者特别是初学者所青睐(王梦琳等,2017)。

MaxEnt 模型应用中主要有 2 种不同的模型构建方案。当以模拟物种的潜在分布为目的时,所构建的模型不需要转移,为物种分布模型;当以模拟物种的生态位为目的时,所构建的模型需要转移,为生态位模型(Peterson *et al.*,2012)。最新研究表明,使用 MaxEnt 模型的默认参数构建模型时,模型对采样偏差敏感,容易产生过度拟合(overfitting),其转移能力也仅在阈值低的情况下较好(朱耿平和乔慧捷,2016;Morales *et al.*,2017)。此外,多数研究者在地理空间中建立模型,未能从生态空间中去思考所模拟物种的生态需求及其对环境因子的响应,导致所模拟的潜在分布与现实分布差距较大。

MaxEnt 模型中主要有 3 个参数与模型的预测密切相关:(1)特征组合(feature combination,FC)参数,特征参数对应于不同的环境变量,将环境变量进行数学转换,从而使 MaxEnt 可利用复杂的数学关系来推测物种对环境因子的响应;(2)调控倍频(regularization multiplier,RM),调控倍频是在特征参数的基础上新添加给模型的一个约束,通过改变 RM 值调整模型对响应曲线的模拟;(3)最大背景点数(max number of background points,BC),或称拟不存在点(pseudo-absence),即在模型构建区域所选取的不存在分布点,将其所关联的环境变量用于和物种存在分布点所关联的环境变量对比,来分析物种生态需求(Merow *et al.*,2013)。

国内多使用 MaxEnt 默认参数对入侵物种或濒危物种进行个案研究,未能深入分析物种对环境因子的响应和探讨 MaxEnt 的参数设置对所模拟物种

的地理分布和生态位的影响。本研究以茶翅蛾 *Halyomorpha halys* (Stål) 为例,通过设置不同特征参数、调控倍频及最大背景拟不存在点数参数,分别在东亚地区构建茶翅蛾本土模型,然后转移至入侵地,比较和分析这 3 种参数对 MaxEnt 模型所模拟茶翅蛾潜在分布和生态位的影响,阐述 MaxEnt 模型参数对模拟物种地理分布及生态位的影响,以期对物种分布,特别是入侵物种的潜在分布进行合理的预测,促进 MaxEnt 模型在我国更为合理的运用。

## 1 材料与方法

### 1.1 数据来源及处理

茶翅蛾本土分布于东亚地区,自 20 世纪 90 年代相继传入北美洲和欧洲成为世界性害虫,是严重危害果蔬的重要害虫。物种分布点数据和环境变量源自 Zhu *et al.* (2012,2016),已去除采样偏差对模型的影响。环境变量的选取主要考虑其对物种分布的限制作用和变量间的空间相关性(Peterson *et al.*,2011),选取了年平均气温(bio1)、最热月份最高气温(bio5)、最冷月份最低气温(bio6)、年降雨量(bio12)和年平均辐射量(bio20)等对茶翅蛾具有生物学意义的环境变量。

### 1.2 模型构建及比较

MaxEnt 模型的预测结果与特征组合(FC)、调控倍频(RM)以及最大背景点数(BC)参数密切相关。目前,MaxEnt 中有 5 种特征(朱耿平和乔慧捷,2016):线性(linear-L),二次型(quadratic-Q),片段化(hinge-H),乘积型(product-P)和阈值性(threshold-T)。在默认参数设置中,特征组合 FC 的选择与物种分布点数目相关,调控倍频 RM 的值为 1,最大背景拟不存在点数目 BC 为 10000。本研究中,将 RM 值设置为 1 至 4,每次增加 1;将 BC 分别设置为 5000 至 25000,每次增加 5000;分别采用 L、Q、H、T 等 4 个特征参数,构建 MaxEnt 模型(注:在茶翅蛾中,由于环境变量的选择使 MaxEnt 模型不能单独基于特征参数 P 运行)。

在茶翅蛾本土地区构建模型,然后将其转入入侵地来比较和检验模型。在探讨 MaxEnt 模型参数对所模拟物种生态位的影响时,基于响应曲线和所模拟物种地理分布在生态空间中的生态位映射。在生态位映射时,基于 10% 训练集阈值(10th training threshold)将预测结果进行二维化,即将单元预测值转换成分布和不分布(presence/absence),然后在

NicheA 3.0(Qiao *et al.*, 2016)中将地理分布的栅格图在生态空间中进行生态位映射。在探讨 MaxEnt 模型参数对所模拟物种地理分布影响时,基于不同参数条件下,模拟茶翅蜡潜在分布的变异程度。

## 2 结果与分析

### 2.1 不同参数设置对响应曲线预测的影响

基于 4 个特征参数的 MaxEnt 模型,茶翅蜡对 5 个气候因子的响应曲线呈现较大差异(图 1)。在茶翅蜡对年平均气温的响应曲线中,当特征为 Q、H 时,基于二者的响应曲线均表现出线性上升趋势;当特征为 L、T 时,对应的响应曲线近似正太分布,其中基于特征 T 的响应曲线较 L 曲折。在茶翅蜡对最热月份最高气温的响应曲线中,当特征为 L 和 Q 时,基于二者的响应曲线均表现出近似线性的上升趋势;当特征为 H、T 时,其对应的响应曲线近似正态分布,且基于特征 T 的响应曲线较 H 曲折。在茶翅蜡对最冷月份最低气温的响应曲线中,当特征为 L 时,响应曲线呈现近似线性的上升趋势;当特征为 Q、H 和 T 时,基于三者的响应曲线均近似正态分布,且基于特征 Q 的响应曲线较圆滑,基于特征 T 的响应曲线最曲折。在茶翅蜡对年降雨量的响应曲线中,当特征为 L、Q 时,基于二者的响应曲线呈近似线性的上升趋势;当特征为 H、T 时,响应曲线近似正太分布,且基于特征 T 的响应曲线较曲

折。在茶翅蜡对年平均辐射量的响应曲线中,当特征为 L、Q 时,响应曲线呈近似线性的下降趋势;当特征为 H、T 时,响应曲线不规则,总体为下降趋势,其中基于特征 T 的曲线较曲折。

与特征参数对响应曲线造成的影响比较,调控倍频 RM 值对 MaxEnt 模型预测的响应曲线的影响相对较小,但 RM 值对响应曲线的平滑程度影响较大,总体表现为:随着 RM 值的增大,响应曲线愈加平滑(图 2)。特别是在对最热月份最高气温、最冷月份最低气温和年平均辐射量的响应曲线中,随着 RM 值的增加,响应曲线变得尤为平滑。

与特征和调控倍频参数对响应曲线影响相比较,背景点数设置对 MaxEnt 模型所模拟的茶翅蜡对 5 个气候因子响应曲线的影响最小。基于 5 个不同背景点数的响应曲线,表现出类似的形状,差异不明显(图 3)。

### 2.2 不同参数设置对生态位预测的影响

特征参数、调控倍频和背景拟不存在点数对 MaxEnt 模型所模拟的茶翅蜡生态位具有影响,其中特征参数的影响较大,基于不同特征所预测的生态位表现出较大差异;调控倍频对生态位预测的影响相对较小;背景拟不存在点数对生态位预测的影响最小,基于不同背景点数 MaxEnt 模型所模拟的生态位重叠(图 4)。

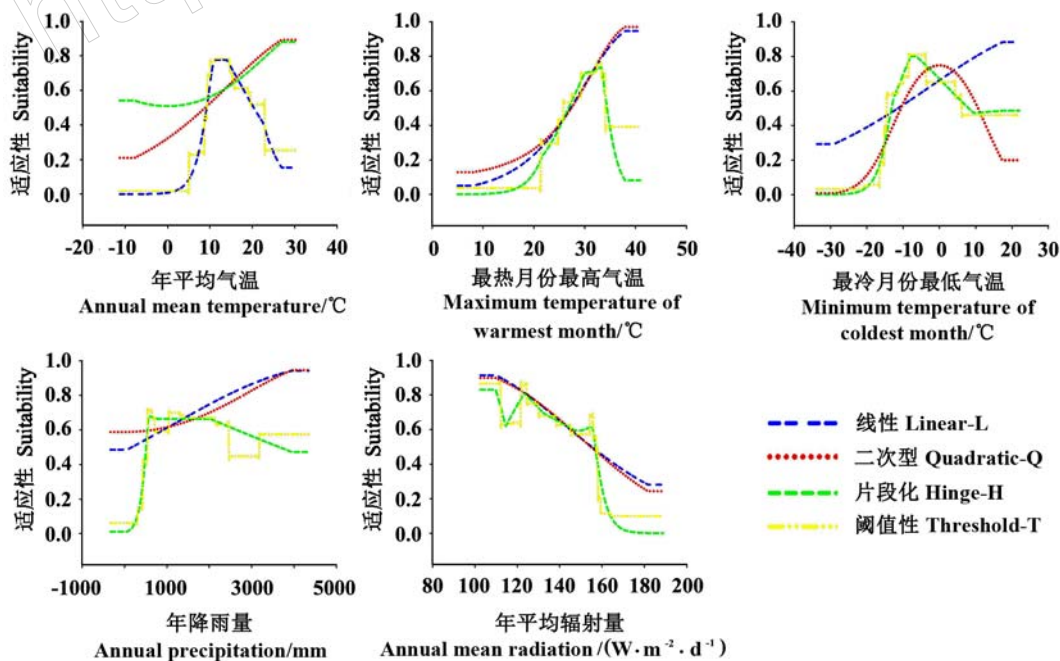


图 1 基于 4 个特征组合的 MaxEnt 模型预测的茶翅蜡对 5 个气候因子的响应曲线

Fig.1 Response curves of *H. halys* to five bioclimatic variables based on four features in MaxEnt model

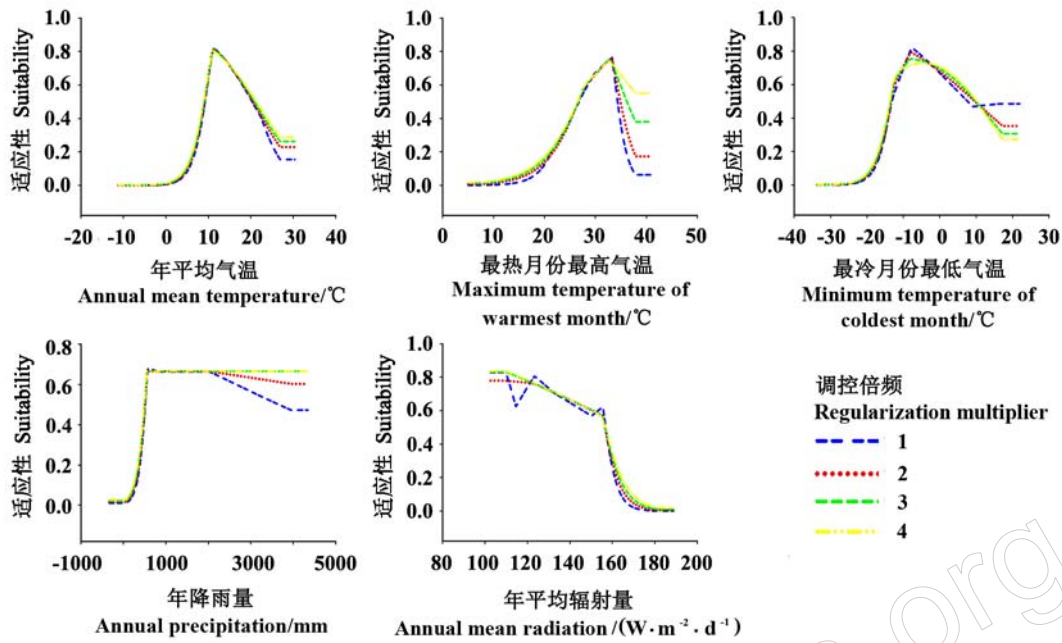


图 2 基于 4 种调控倍频的 MaxEnt 模型预测的茶翅蜡对 5 个气候因子的响应曲线

Fig.2 Response curves of *H. halys* to five bioclimatic variables based on four regularization multipliers in MaxEnt model

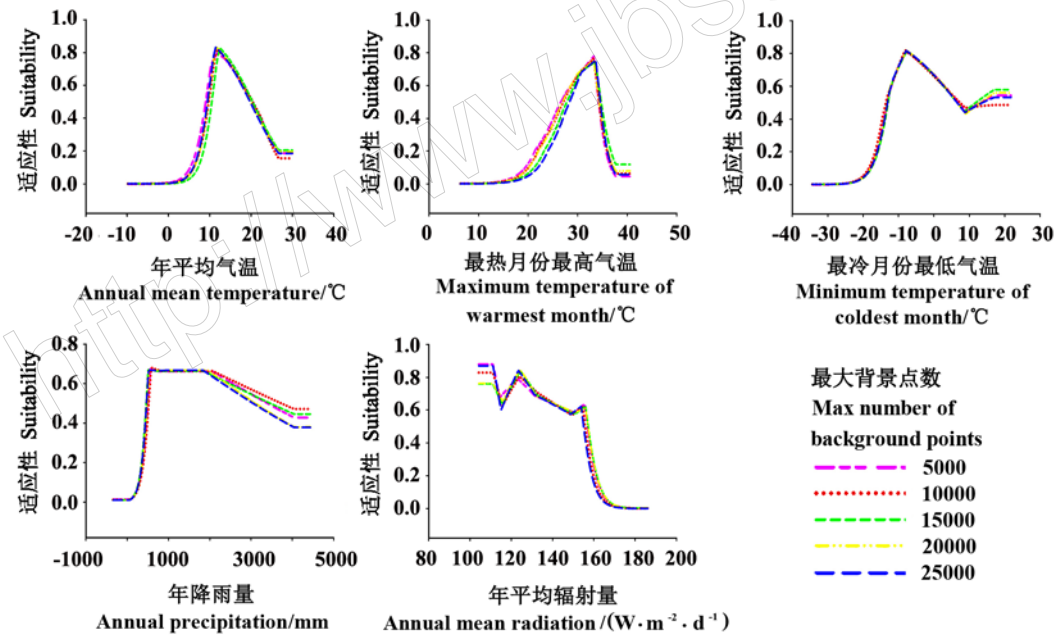


图 3 基于 5 种最大背景不分布点数的 MaxEnt 模型预测的茶翅蜡对 5 个气候因子的响应曲线

Fig.3 Response curves of *H. halys* to five bioclimatic variables based on five sets of background points in MaxEnt model

### 2.3 不同参数设置对潜在分布预测的影响

特征参数, 调控倍频和背景拟不存在点数对 MaxEnt 模型所预测的茶翅蜡潜在分布具有影响。在不同参数下构建 MaxEnt 模型, 其所预测潜在分布在全球不同区域表现出不同程度的差异性, 其中特征参数对潜在分布预测的影响最大, 调控倍频的影响次之, 背景点数的影响最小(图 5)。特征参

数, 调控倍频和背景点数对潜在分布预测造成的差异性在空间分布上是不均一的, 如在北美洲西部、南美洲、非洲和亚洲北部地区、中东地区以及澳洲地区, 特征参数对潜在分布的影响较大, 而在南美洲北部和非洲中部, 调控倍频和背景点数的影响相对较大。

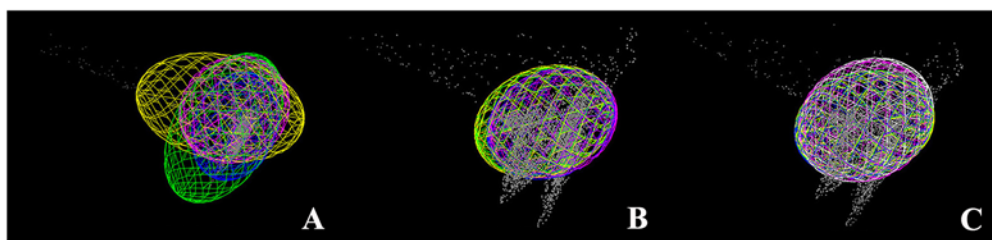


图 4 基于不同特征参数、调控倍频和最大背景点数的 MaxEnt 模型对茶翅蜡生态位的预测

Fig.4 Ecological niche modeling based on MaxEnt model predictions for *H. halys* using different feature combinations, regularization multipliers and background points

A: 特征组合参数; B: 调控倍率; C: 背景点数。图中不同椭圆代表 MaxEnt 所模拟的潜在分布在生态空间中的映射。

A: Feature combination; B: Regularization multiplier; C: Background points.

Ellipsoids represent the niche based on MaxEnt model predictions under different parameters.

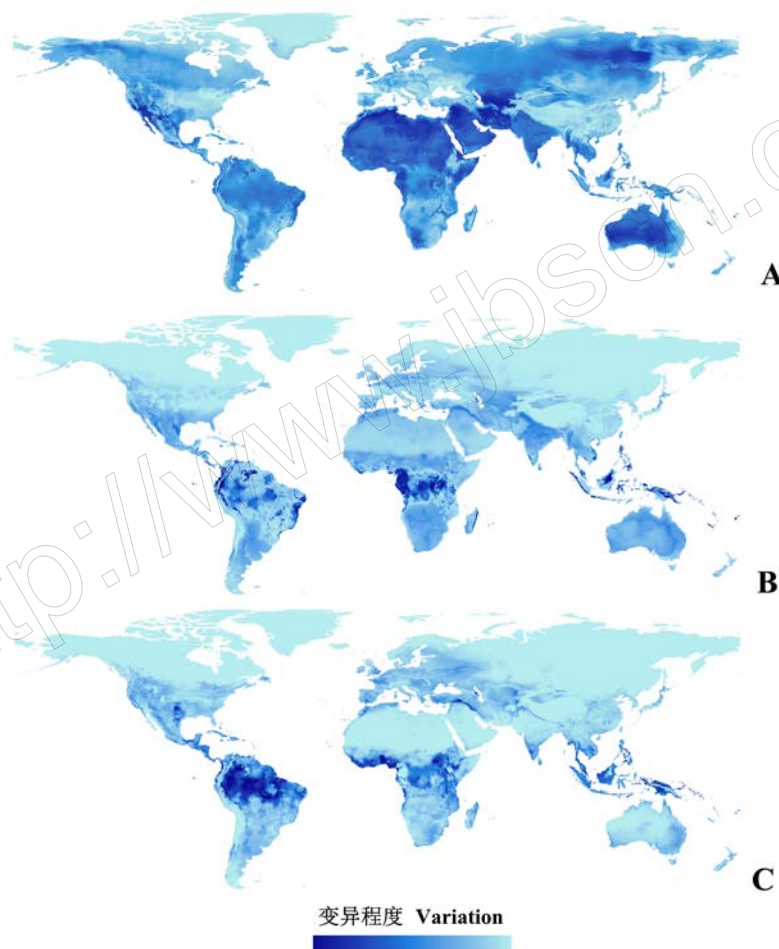


图 5 基于不同特征参数、调控倍频和最大背景点数的 MaxEnt 模型对茶翅蜡潜在分布预测的变异程度

Fig.5 Variations of MaxEnt distributional predictions for *H. halys* based on different feature combinations, regularization multipliers and background points

A: 特征组合参数; B: 调控倍率; C: 背景点数。

A: Feature combination; B: Regularization multiplier; C: Background points.

### 3 讨论

本研究中,基于 4 个特征参数的 MaxEnt 模型预测茶翅蜡对 5 个气候因子的响应曲线表现出较大的差异,其中基于特征 L 和 Q 的响应曲线较为平滑。在不同气候因子中,基于特征 H 的响应曲线略

有差异:在年平均气温下,其响应曲线最为平滑,而在年平均辐射量下,其响应曲线较为曲折。与特征 L、H 和 Q 相比,基于特征 T 的响应曲线最为曲折。在茶翅蜡研究中,特征参数的选择对 MaxEnt 模型预测的生态位影响较大,所预测的潜在分布在空间

上亦呈现差异性。基于特征 L 和 Q 建立模型所模拟的响应曲线相对圆滑,笔者推测其所模拟的生态位可能更接近于基础生态位(fundamental niche),模型相对简单,具有较高转移能力;而依据特征 H 和 T 建立模型其响应曲线相对曲折,所模拟的生态位倾向于现实生态位(realized niche),模型相对复杂,对分布数据拟合较好,具有较高的预测能力。简单模型和复杂模型各具有优势,一般认为简单模型的准确率低、可转移性高;而复杂模型恰好相反,准确率高、可转移性差(乔慧捷等,2013)。

在 MaxEnt 模型中,调控倍频通过给特征参数增添约束起作用,其默认设置为 1(Elith *et al.*, 2011)。基于 4 种调控倍频,MaxEnt 模型对茶翅蝽 5 个气候因子响应曲线的预测表现出差异性,总体趋势是随着 RM 值的增大,响应曲线变得平滑。调控倍频对 MaxEnt 模型预测生态位具有影响,调控倍频的变化将会影响 MaxEnt 模型所推测的生态位在现实生态位和基础生态位之间变化(Jiménez-Valverde *et al.*, 2011)。调控倍频对 MaxEnt 模型预测茶翅蝽潜在分布亦具有影响,特别是在南美洲北部以及非洲中部地区的预测变异程度较大。在模型构建中, RM 值的设置需要考虑物种对环境因子的影响。

基于 5 个背景拟不存在点数,MaxEnt 模型对茶翅蝽 5 个气候因子的响应曲线和生态位的预测变化较小,但在全球潜在分布预测中变化较大,特别是在南美洲北部以及非洲中部地区的预测变异程度较大。相关性方案的生态位模型,采用存在分布点和不存在分布点所关联的环境变量去推测物种的生态需求和模拟物种分布,因此不存在物种分布点的选择对模拟生态位和潜在分布存在影响(Barve *et al.*, 2011; Zhu *et al.*, 2013)。笔者认为,由于限定了模型构建区域,不存在分布点只能在模型构建区域进行选择,所以本次背景点数的选取对 MaxEnt 模型所模拟的响应曲线和生态位的影响相对较小。

### 参考文献

乔慧捷, 胡军华, 黄继红, 2013. 生态位模型的理论基础、发展方向与挑战. *中国科学: 生命科学*, 43(11): 915-927.  
王梦琳, 范靖宇, 李敏, 朱耿平, 2017. 入侵蔗扁蛾在我国的潜在分布区. *生物安全学报*, 26(2): 129-133.  
朱耿平, 刘国卿, 卜文俊, 高玉葆, 2013. 生态位模型的基本原理及其在生物多样性保护中的应用. *生物多样性*,

21(1): 90-98.  
朱耿平, 乔慧捷, 2016. Maxent 模型复杂度对物种潜在分布区预测的影响. *生物多样性*, 24(10): 1189-1196.  
朱耿平, 范靖宇, 王梦琳, 陈敏, 乔慧捷, 2017. ROC 曲线形状在生态位模型评价中的重要性——以美国白蛾为例. *生物安全学报*, 26(3): 184-190.  
BARVE N, BARVE V, JIMÉNEZ-VALVERDE A, LIRA-NORIEGA A, MAHER S P, PETERSON A T, SOBERÓN J, VILLALOBOS F, 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222(11): 1810-1819.  
ELITH J, PHILLIPS S J, HASTIE T, DUDÍK M, CHEE Y E, YATES C J, 2011. A statistical explanation of Maxent for ecologists. *Diversity & Distributions*, 17(1): 43-57.  
JIMÉNEZ-VALVERDE A, PETERSON A T, SOBERÓN J, OVERTON J M, ARAGÓN P, LOBO J M, 2011. Use of niche models in invasive species risk assessments. *Biological Invasions*, 13(12): 2785-2797.  
MEROW C, SMITH M J, SILANDER J A, 2013. A practical guide to Maxent for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36(10): 1058-1069.  
MORALES N S, FERNÁNDEZ I C, BACA-GONZÁLEZ V, 2017. Maxent's parameter configuration and small samples: are we paying attention to recommendations? A systematic review. *PeerJ*, 5(2): e3093.  
PETERSON A T, SOBERÓN J, PEARSON R G, ANDERSON R P, NAKAMURA M, MARTÍNEZ-MEYER E, ARAUJO M B, 2011. *Ecological niches and geographic distributions*. New Jersey: Princeton University Press.  
QIAO H J, PETERSON A T, CAMPBELL L P, SOBERÓN J, JI L Q, ESCOBAR L E, 2016. NicheA: creating virtual species and ecological niches in multivariate environmental scenarios. *Ecography*, 39(8): 805-813.  
ZHU G P, BU W J, GAO Y B, LIU G Q, 2012. Potential geographic distribution of brown marmorated stink bug invasion (*Halyomorpha halys*). *PLoS ONE*, 7(2): e31246.  
ZHU G P, GAO Y B, ZHU L, 2013. Delimiting the coastal geographic background to predict potential distribution of *Spartina alterniflora*. *Hydrobiologia*, 717(1): 177-187.  
ZHU G P, GARIEPY T D, HAYE T, BU W J, 2016. Patterns of niche filling and expansion across the invaded ranges of *Halyomorpha halys* in North America and Europe. *Journal of Pest Science*, 90(4): 1-13.