

DOI: 10.5846/stxb201411202297

黄敏毅, 孔晓泉, 段仁燕, 吴甘霖, 张中信. 虚拟物种的基本原理及其在物种分布模型评估中的应用. 生态学报, 2016, 36(9): - .

Huang M Y, Kong X Q, Duan R Y, Wu G L, Zhang Z X. The basic principle of virtual species and its application to evaluations of species distribution models. Acta Ecologica Sinica, 2016, 36(9): - .

虚拟物种的基本原理及其在物种分布模型评估中的应用

黄敏毅, 孔晓泉, 段仁燕*, 吴甘霖, 张中信

安庆师范学院生命科学院, 安庆 246011

摘要: 物种分布模型是建立在物种出现或缺失数据的基础上, 但可获得真实分布数据存在着各种各样的缺点(如: 物种识别错误、坐标错误、抽样偏差、数据缺失等), 影响着物种分布模型的预测性能、稳定性及应用, 因此使用物种真实分布数据评估物种分布模型将带来很大的不确定性。为避免这种不确定性, 越来越多的研究使用虚拟物种来评价物种分布模型的性能, 评估新方法的优劣。虚拟物种是一种建立在真实(或虚拟)地理信息系统下人工生命, 是简化和抽象的物种, 它通过模拟物种对环境变量的响应关系, 评估物种在不同环境变量下的出现概率, 人为地给出虚拟的物种分布数据。虚拟物种具有数据容易获得、数据质量可控、避免过度模拟等优势, 目前它被广泛用于评估物种特性、抽样偏差、地理信息、出现/缺失标准等对物种分布模型性能的影响。虚拟物种是大尺度研究中不可或缺的重要工具, 有利于解决真实数据未能解决的科学问题。常用的构成算法有求和法、求积法和综合法, 但这些方法可能存在补偿效应, 扩大了物种的分布范围。考虑到虚拟物种的不足, 本综述提出了未来虚拟物种可能的发展方向(避免过度脱离真实, 完善虚拟物种的构成算法, 构建虚拟的模式生物、群落及生态系统等)。为帮助研究者快速构建虚拟物种, 基于 R 环境开发了一个虚拟物种构成软件包(SDMvspecies)。虚拟物种可以与真实物种相结合, 通过改进模型的构成方法, 有利于解决一些真实数据未能解决的问题; 虚拟物种的应用也将导致一些新理论的产生, 有利于更好地理解生态学原理。

关键词: 物种分布模型; 虚拟物种; 生境适宜度; 抽样偏差; 人工生命

The basic principle of virtual species and its application to evaluations of species distribution models

HUANG Minyi, KONG Xiaoquan, DUAN Renyan*, WU Ganlin, ZHANG Zhongxin

College of Life Sciences, Anqing Normal University, Anqing 246011, China

Abstract: The fascinating scientific questions of how and where species will potentially distribute under current and changing environmental conditions have inspired many biogeographers, ecologists, and managers to develop to predict the potential distributions of plants or animals by quantifying species-environment and develop predictive models. The species distribution model (SDM), an essential modeling tool, has been relationships. A key challenge in using real species data (presence-absence data and/or presence-only data) for SDM is the uncertainty about where and how the thousands of species distribution data records are attained. The majority of species distribution data sets are derived from herbaria, university databases, museums, or even amateur field workers. Therefore, attaining a reasonable explanation for species distribution in the wild is often hindered by the problems inherent in these large data sets, including species-specific properties (e.g., species prevalence, dispersal barriers, interspecific competition, distribution pattern), biased sampling

基金项目: 林业局公益性行业科研专项(201304314); 国家自然科学基金(31300342, 31100311); 安徽省自然科学基金(1408085MC64)

收稿日期: 2014-11-20; **网络出版日期:** 2015- -

* 通讯作者 Corresponding author. E-mail: huang.m.y@163.com

(e.g., reachability of observation sites, visibility or detectability of observation objects), variability among observation methods (e.g., time interval and spatial range), and habitat types, particularly for data collected over a long time interval and a large spatial range. The use of virtual species could provide a suitable unifying framework to select the most appropriate model for such evaluations, by comparing the predictive accuracy and virtual distributions in a geographic information system model of a real landscape. In recent years, virtual species distribution models have become increasingly important tools to study various problems in the fields of conservation biology, ecology, biogeography, climate change research, and evolution. Virtual species have many advantages, including the ease of attaining a large number of data sets for each scenario, ability to fully control the quality of data, prevention of the over-fitted phenomenon inherent to SDMs, and the ability to independently evaluate the predictive power of SDMs regardless of other factors. There are three common methods to generate virtual species: the additive method, multiplicative method, and comprehensive method. Here, we provide an overview of recent advances in the development of virtual species distribution models by using spatially explicit simulated distribution data to represent the ‘true’ species’ distributions. We highlight the four main applications of these models, including species-specific characteristics, sampling bias, geographic information, and threshold standard for species occurrence, in evaluating model performance. Considering the current limitations, we propose future directions for the development of virtual species, including avoiding excessive assumptions that do not reflect reality, optimizing the generation of virtual species to avoid the compensatory effect and reflect true species dynamics and biological characteristics, and generating a virtual model organism, population, community, and ecosystem. To help researchers generate virtual species easily and quickly, our research team has developed a software package, *SDMvspecies*, based on R language. The software package has four methods to create virtual species, including the niche synthesis method, pick mean method, pick median method, and artificial bell-shaped response method. The *SDMvspecies* software can be accessed with a free download from the website <http://cran.r-project.org/web/packages/sdmvspecies/>. We further address the need for better integration of virtual species with ecological theory, which is expected to lead to new questions, theories, and an improved mechanistic understanding of ecological systems.

Key Words: species distribution models; virtual species; habitat suitability; biased sampling; artificial life

物种分布模型(SDM, Species distribution model)是一个重要的建模工具,主要用于预测物种在当前或特定的环境空间的潜在分布^[1-14]。自从 Nix 等^[6]开始模拟澳大利亚物种的分布后,物种分布模拟已经渗透到生态学、保护生物学和生物地理学等多个领域,并被广泛应用于研究进化和生态假说^[7],评估物种的入侵和扩散^[8-9],模拟生物多样性的组成^[10],预测气候变化、栖息地破碎和其他环境变化的影响^[1,11-12],搜寻未被发现物种和稀有物种可能的分布区域^[13]和发展管理规划和保护策略^[14]等。物种分布模型已经成为大尺度研究中不可或缺的重要工具,包括物种分布机制的理解、物种与环境适应过程的再现、物种存在(或丰度)和环境条件之间的量化等,这有利于预测环境变化对物种分布的影响。

传统的物种分布模型主要建立在真实数据的基础上。但这些真实物种的出现/缺失数据(PA, Presence/absence data)或者仅仅物种的出现数据(PO, Presence-only data)存在不确定性,如大量的数据是如何获得的,在哪获得的^[15]。现有的物种分布数据主要来自于数据库、博物馆、标本馆及公开发表的论文等,甚至包括业余野外工作者收集的资料^[16]。而且,即使是大规模的调查,也很难完整地了解物种的真实分布,因为物种分布受到很多因素的制约,例如,物种属性(如生活史、稀有性、物种分布范围、可观测能力、传播障碍、种间竞争、扩散能力等)^[8,17];分布特征(如分布范围和空间自相关性)^[18-19];抽样偏差^[20];观察方法(如时间间隔和空间范围的)的制约;观测地点的可到达性(有的栖息地容易观察,如开阔地、草地、林缘、低空等;而有的栖息地较难观察,如森林、水体、湿地等)^[21]。特别是,许多物种分布数据是在 GPS 技术广泛应用之前获得的^[22]或来自于一些不熟悉物种分类调查者的调查资料^[23]。可见,任何关于现实自然的理解是基于有限的数据(或偏

差的数据),而研究方法的比较和对数据集的分析都受到数据量完整性的约束及研究者抽样偏差的影响^[24-27],最后造成转化的物种与环境关系的偏差或不足^[27]。

为了比较不同模型算法的性能差异,需要评估这些方法是否合理,预测结果是否正确等,这些都要求将模型的预测结果与真实的分布进行科学地比较。然而,无法获得物种的真实分布范围。为避免真实物种数据无所不在的误差,越来越多的科研工作者开始用虚拟物种(Virtual species)来评估模型算法的有效性^[26-30]。通过虚拟物种的方法可以在任何时候获得所有相关信息。因此,可用虚拟物种来评估物种分布模型的性能及选择合适的模型。

事实上,早在 20 世纪 50 年代就有科学家使用虚拟数据进行科学研究。最早的模拟数据应用可以追溯到 Stickel^[31]测试动物运动因素(用于评估标志重捕数据的质量)和 Swan^[32]提出的植物分类计数评估(模拟物种种群在单一环境梯度下的数量变化)。此后,Fortin 等^[33]使用密集采样的数据点(在 0.5 km² 的范围里人工设置 200 分布点)作为“真实值”,模拟研究不同抽样大小及物种分布格局与空间自相关的关系。随着计算机技术和统计学的快速发展,预测变得更加真实可行,也能提供更多有用的信息。其中, Hirzel 等^[26]最早用虚拟物种来评估物种广义线性模型(GLM, Generalised linear model)和生态位因子模型(ENF, Ecological niche factor analysis)的预测性能,他们认为虚拟物种的使用提供了一个高效的工具,允许研究者完全控制输入数据的质量来评估预测精度,为选择最适宜的模型提供了一个可信的方法。

近年来,虚拟物种逐渐成为重要的工具来研究保护生物学、生态学、生物地理学、气候变化及进化等方面的科学问题^[26,34]。本综述中,系统总结了虚拟物种最新的研究进展及一些主要的应用,并在此基础上探讨了虚拟物种存在的不足及下一步可能的发展方向。

1 虚拟物种的概念及优点

虚拟物种是通过模拟物种对环境变量的响应关系,评估物种在不同环境变量下的出现概率,人为地给出的具有一定分布范围的虚拟化的物种^[28]。虚拟物种又被称作“人工数据”(Artificial data)或“人造物种”(Artificial species)^[27,35]，“虚拟生态学家”(Virtual ecologist)^[24,28]，“模拟数据”(Simulated data)^[18]，“虚拟生态”(Virtual ecology)^[24,29]等。在 Wiley 数据库检索了 2000—2014 论文题目或摘要中出现这 6 个词中任意一个词的文章数。检索结果证实,关于虚拟物种的相关论文数量呈明显增加趋势(图 1)。

通过构建虚拟物种来评估物种的分布模型性能可避免未知生物因素的影响,如个体水平(性别、年龄、迁移能力、耐受能力等),种群水平(分布模型、种内竞争)和群落水平(种间竞争、食物资源的可利用性、捕食关系)等^[34]。虚拟物种的优点主要包括:1)数据很容易获得,不受时空的限制^[27-28];2)可完全控制数据质量,避免随机抽样误差^[36];3)避免物种分布模型中的过度模拟现象^[26,28,34];4)可独立的评估不同环境因子对物种分布模型预测能力的影响等^[26,28,34]。

2 虚拟物种构成的基本原理及构成软件

Hutchinson^[37]认为物种的生态位是物种在多维环境空间中的超体积,在这个多维环境条件下物种能够存活、繁殖和分布。一些重要的非生物限制因子(多个环境变量)决定着物种的分布范围,可通过这些环境变量来评估物种的真实生态位。虚拟物种是建立在 Hutchinson 的超体积生态位基础上,通过人为地设定限制物种分布的环境因子,根据超体积生态位理论,反推其在真实地图上的分布。虚拟物种主要通过物种分布与环境

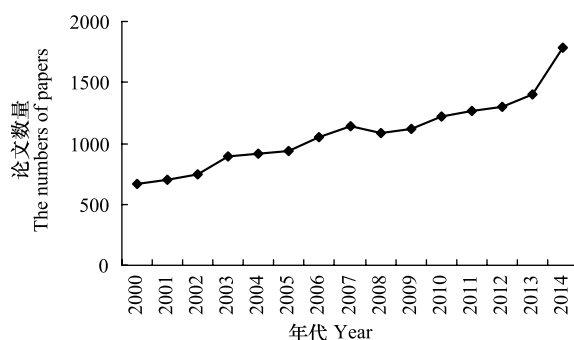


图 1 2000—2014 年关于虚拟物种论文发表数量

Fig. 1 The numbers of papers on virtual species from 2000 to 2014

变量的关系及其生境适宜性(在不同范围上物种出现的概率),转换为存在/不存在分布,虚拟产生物种的分布范围,是在模拟环境下构成的虚拟分布空间。具体包括三步来构建虚拟物种^[2,26,34,35,38-40]。这三步包括:1)根据超体积生态位理论,产生合适的生境适宜度指数来反映物种与环境变量的关系;2)转化生境适宜度指数为适宜的二进制地图(物种出现/缺失地图);3)通过抽样这些模拟数据,评估特定的目的。比如,抽样的偏差和敏感度、预测精度等。其中,如何产生合适的生境适宜度指数对后期构建的虚拟物种模拟至关重要,而产生生境适宜度指数主要包括三个环节^[35]:1)通过主成分分析或者人为指定合适的环境变量;2)选择适宜的方法(物种与环境变量的响应函数关系)来转化环境变量为生境适宜度指数。常用的函数包括钟形曲线(高斯曲线)、线性曲线(线性递增和线性递减)、截线性曲线(递增和递减)及逻辑斯谛曲线。最后,将每个变量的生境适宜度指数转化为总的生境适宜度指数,常用的三种不同的方法为求和法、求积法及综合法(表1)。已有的研究表明,环境变量与物种的不同关系及构成算法会对物种分布模型的性能造成影响^[35,41],应根据研究目的评估这些差异及选择合适的方法。

表 1 构成虚拟物种的常用的三种算法
Table 1 The three common ways to generate virtual species

方法 Methods	计算公式 Calculation formula	参考文献 Reference
求和法 Additive method	$H = \frac{1}{\sum_{i=1}^n W_i} \sum_{i=1}^n W_i H_i + \varepsilon \quad (1)$	[26,34]
求积法 Multiplicative method	$H = \sum_{i=1}^n H_i + \varepsilon \quad (2)$	[35]
综合法 Comprehensive method	$H = \prod_{i=1}^n W_i H_i + \varepsilon \quad (3)$ $P(X_i) = \frac{1}{1 + e^{-f(X_i)}} \quad (4)$ 其中, $f(X_i) = a_0 + \sum_{j=1}^n a_j x_{ij} + \sum_{j,k=1}^n b_{jk} x_{ij} x_{ik} \quad (4)$	[35,38] [2]

H_i :第*i*个环境变量的生境适宜度指数; W_i :生境适宜度指数的权重; ε :随机误差。 a_j 和 b_{jk} 为0—1正态分布的随机值; X_i 为物种在第*i*个环境变量的生境适宜度指数; $f(X_i)$ 是总的生境适宜度指数; P 是物种在特定格子下的出现概率。

为便于研究者快速构建虚拟物种,Duan等^[42]基于R环境开发了一个虚拟物种构成软件包SDMvspecies。这个软件包具有简单、直观和标准化的特点,甚至对那些不熟悉R环境的研究者也能很方便的运用。软件是开源的,软件开发者可在此基础上进行改进和升级,其免费下载地址为<http://cran.r-project.org/web/packages/sdmvspecies/>。安装R包之前需要先安装栅格化地图软件,如GIS、Raster。目前的软件包主要包含了4个构成虚拟物种的算法:1)生态位综合法(Niche syntheses method):每一个环境变量均被转化为偏生境适宜度指数,然后加权为总的生境适宜度指数^[26]。2)平均值法(Pick mean method):通过主成分分析确定简化的不具有相互关系的环境变量,如果每个环境变量的范围是位于被选择的生境适宜度指数的变化范围内(平均值±标准差),该位置作为虚拟物种真实的分布位点^[36]。3)中值法(Pick median method):整个计算过程与平均值法类似,通过主成分分析获得关键的环境因子。如果环境因子处于环境变化的四分位中心,则认为该区域适宜虚拟物种存在^[43]。4)人工钟型曲线法(Artificial bell-shaped response method):通过人为给出特定环境变量的平均值和标准差,利用乘积的方法获得最后的生境适宜度^[30]。

3 虚拟物种在评估物种分布模型性能中的应用

3.1 评估物种特性的影响

多数物种分布模型的问题是未考虑物种性能差异(比如,物种是否容易识别、栖息地是否被完整调查或是否存在物种识别错误等)造成物种分布数据的不准确^[8,17]。即使进行高效足够大的采样工作,也不能完全

掌握真实的分布,这会导致对物种分布预测结果的不正确。Lahoz-Monfort 等^[17]通过虚拟物种模拟了 1000 个位点环境与物种出现的关系(2 个不同强度的正线性逻辑斯谛函数关系),4 个环境与物种可探测概率之间关系(不变的、正相关、负相关及独立的),评估位置的探测概率及物种的探测概率对模型性能的影响,结果证实如果忽视物种的可探测能力差异会对预测结果造成明显的影响。

物种分布模型构建的前提假说是物种分布与环境处于平衡阶段,即物种出现的地点代表适宜的环境空间,没有出现的地点代表不适宜的空间。但扩散受限制的物种及外来物种不符合这个平衡假说,特别是针对处于扩展阶段的外来物种而言。Václavík 和 Meentemeyer^[8]用一个真实的入侵物种(*Phytophthora ramorum*)和构建的虚拟物种评估了入侵过程中的不同阶段(初级、中级、平衡)对模型预测精度的影响,结果证实基于真实物种和虚拟物种的预测结果均有较高的预测精度,但是处于不同入侵阶段的物种分布数据对模型的预测结果会造成明显的影响。他们认为研究者所获得真实入侵物种分布数据受入侵阶段的影响,而建立在不完整分布数据上的预测结果会造成对外来物种入侵风险理解的偏差。比如,一个处于入侵初期的物种,其分布范围远低于真实的分布空间,根据其分布数据所模拟构建的环境生态位仅代表了其真实适宜生态位的一小部分,弱化了外来物种的入侵范围及入侵风险。而要真实评估外来物种的入侵风险,需要评价外来物种所处的入侵阶段。但对真实物种而言,很难评估物种的入侵阶段,因此需要通过虚拟物种来评价不同入侵阶段的预测偏差及选择合适的预测模型。

3.2 评估抽样偏差的影响

物种的出现数据受到种群自身动态及环境复杂性的影响。物种真实的分布很难完整地获得,因为其分布比较复杂,而且随着时间和空间尺度变化而处于动态变化之中。由于原始数据来源于不同地点、不同时间及不同调查者,任何抽样都存在未知程度的抽样误差^[20,30]。尽管这些数据同样可以提供有价值的信息,但研究发现抽样偏差与模型的预测精度密切相关^[25]。考虑到物种性能的差异及物种的动态变化,不可能知道物种所有的真实分布数据,所以很难应用真实物种来评估抽样偏差对模型预测结果的影响。但虚拟物种不受时间、空间的限制,可根据研究的需要进行人为地设定,也可人为地设定各种可能的抽样方法,为评估抽样偏差对模型性能的影响及探讨改进抽样方法提供了可能。例如,Fourcade 等^[25]通过构建虚拟物种比较了 4 种偏差类型(偏差出现的位置、梯度、偏差的中心、据城市或道路的距离)及 3 个水平的偏差强度(低强度、中强度、高强度)对最大熵模型(Maxent, Maximum entropy)预测性能的影响。结果证实,不同的偏差类型均导致模型的预测精度下降,下降的程度取决于物种的种类及偏差类型。Fourcade 等^[25]并将虚拟物种的预测结果与真实物种(*Chrysemys cylindraceus*, *Chrysemys picta*)进行对比,发现抽样偏差对物种 *C. picta* 分布模型预测精度的影响明显高于其他 2 个物种(*P. cylindraceus* 和虚拟物种)。如对物种 *C. picta* 而言,所有的评估测定均受到处于偏差地点(偏差的中心)的影响,其中预测精度下降超过 5%,与假定的没有偏差的预测模型重叠范围仅在 26%到 49%之间。而对物种 *P. cylindraceus* 而言,其他类型的偏差影响比较小,与假定的没有偏差模型之间没有明显的差异;但是偏差地点(偏差的中心)与没有偏差的模型重叠范围的存在明显差异。随后,他们比较了 5 种不同偏差校正方法(改进的系统抽样法、评估抽样偏差的软件包、改进的缺失数据设置、进行聚类分析和分散抽样)的效果,结果证实,校正效果依赖于偏差类型、偏差强度及物种。可见,忽视抽样偏差的影响可能会导致不正确的预测结果。

其他学者也探讨了其他修正偏差方法对模型预测精度的影响。例如,Barbet-Massin 等^[38]使用高斯曲线构建了 2 个虚拟物种,探讨了 1 个由气候引起的偏差(由高斯分布转化为二项式分布)和 2 种由空间引起的偏差(在物种散布范围的边缘人为地移除一些数据和仅仅选择出现在交通线(如公路和铁路)周围的数据)的影响。结果发现,抽样偏差降低了所有模型的预测精度,但受影响的程度受到出现点数量、质量、出现-缺失数据的比值及选择的出现-缺失数据方法的影响。比如,当出现地理偏见时,对三个回归模型(GLM, GAM (Generalized additive models) 和 MARS (Multivariate adaptive regression splines) 和 MDA (Mixture discriminant analysis))而言,随机选择出现-缺失数据的方法表现最优越。当出现气候偏见时,对三个回归模型而言,当出

现点较少时,在研究区域之外选取不出现点的方法最优越;当出现点很多时,2°法(在远离出现数据点至少 2° 的范围内附近选取不出现点的方法)会得到更好的结果;但对其他模型(MDA、CTA(Classification tree analysis)、BRT(Boosted regression trees)、RF(Random forest))而言,不论出现点多少,2°法均表现的最好。Rapacciuolo 等^[20]通过构建虚拟物种提出了一个两次使用物种分布数据来降低抽样偏差的方法,并与真实鸟类分布数据进行对比分析,结果证实其新提出的方法能有效地提高模型的预测能力。此外,Varela 等^[30]通过在伊比利亚半岛构建虚拟物种,比较了地理和气候过滤对降低抽样偏差的影响。主要的方法是使用 R 软件栅格包中的抽样格子函数(Grid-sample function)进行地理和气候过滤:在地理过滤过程中,x 轴为经度,y 轴为纬度,同一格子中如果出现位点超过 2 个,重复的出现点将被去掉,这样在地理空间上就缺少了重复的聚集出现点。在气候过滤中,x 轴为最暖月的最高温度,y 轴为最干旱季节的降雨量,气候上相似的格子将被去掉。结果发现,地理过滤并没有提高模型的预测精度,甚至有时候变得更糟糕;而气候过滤则能稳定地获得很好的预测结果。这些结果均证明,抽样偏差会对模型的预测结果造成明显的影响,应根据不同的偏差类型、偏差强度及物种的分布特点,采用合适的修正方法。

3.3 评估地理信息的影响

当邻近位点之间空间变量的数值彼此不独立的时候,就会出现空间自相关(Spatial autocorrelation)^[18-19]。造成的主要原因包括:1)一些依赖距离的生物学过程,诸如物种进化、灭绝、分布或物种的相互作用;2)环境与物种出现的关系多为非线性的,但常被模拟为线性的;3)模型未能考虑重要的环境因素,比如空间异质性及对空间异质性的反应^[18]。真实物种的分布受限于特定的地理空间,很难评估空间自相关的影响程度。而人工产生的虚拟物种可虚拟产生具有不同程度空间自相关程度的分布数据,有利于评价空间自相关对模型性能的影响。例如,Saas 和 Gosselin^[19]通过虚拟物种方法来计算规则和不规则的空间位点的自相关关系,评估了模型的预测准确性和常见的错误类型,结果发现频率法(Frequentist methods)明显增加了 I 型错误(Type I errors)的概率,而贝叶斯法(Bayesian method)可以获得相对满意的预测结果;同样,对真实物种而言,基于距离的贝叶斯法预测结果也是最好的。但是其他的研究结果证实空间自相关对模型结果的影响程度反而很低^[40,44]。例如,Naimi 等^[40]使用虚拟物种检测了物种分布模型的预测结果是否受位置不确定性(空间自相关)的影响。Thibaud 等^[44]通过构建虚拟物种测定了 5 个因素(没有协变量、物种的出现和缺失过程导致的空间自相关、抽样大小、抽样设计和模拟技术)的相对重要性。结果,他们发现空间自相关并不是时刻存在的,而是受到物种所分布的地理空间的影响。可见,为提高模型的预测精度,应首先评估是否存在空间自相关及空间自相关程度。

同样,不同的空间尺度(空间分辨率)会改变物种的分布类型(聚集分布、随机分布和均匀分布),造成预测结果的不确定性。因此,有必要明确空间分辨率是否会造成模型预测结果的变化。为评估空间分辨率的影响程度及精度,Bombi 和 D'Amen^[39]在真实的南非地图上模拟不同分辨率地图(5'、15'、30'、60'和 120')下 10 个虚拟物种分布,来评估广义线性模型和广义推进模型(GBM, Generalized boosted models)预测结果差异;并与真实的地中海地区的爬行动物预测结果进行对比。结果发现,随着分辨率的下降,预测结果精度降低,其中 15'和 30'差异不显著,60'变化显著,120'甚至会得到一个令人不解的结果。整体而言,在不同尺度下,使用真实物种数据得到的预测精度均比较高,研究者推测可能是物种特性的差异造成的。因为虚拟物种是根据人为设定的环境变量来反推其分布范围的,是一种简化的物种;而真实物种与环境存在着复杂的相互关系,同样的变量随着分辨率的改变和地理位置的变化可能起着不同的作用,这限制了真实物种用于研究分辨率改变的有效性。

一些限制物种扩散的因素(生物作用、地理阻碍等)及一些促进其扩散的因素均会对物种分布模型性能造成影响^[21]。扩散限制会造成物种在适宜的区域不能出现,因为其得不到有效补充。此外,物种有限的散布能力也阻碍着物种完全占据未知的空间。因此,需要真实地评估扩散限制的影响程度。比如,生物个体的迁移能力取决于物种的散布能力和繁殖力,这些指标可通过控制实验获得。但是,物种的迁移能力又取决于个

体特征(大小、性别、健康状态等),环境变化的快慢,种内和种间竞争,迁移地点的可到达性,环境空间的多变性等,这些指标很难一一获得。这都限制了使用真实物种来研究扩散限制的影响。可通过虚拟物种的方法,人为地设定虚拟物种的扩散特性及环境的可到达性。比如,Saupe 等^[45]使用一种虚拟物种框架来探讨非生物因素和扩散限制对模型预测结果的影响。他们假定生物的相互作用并不限制物种的分布,通过建立一个简化的 BAM 模型(B 代表了生物的相互作用,A 代表了物种生存与增长所需要的非生物空间,M 代表了物种可到达的区域)来研究扩散限制对模型预测结果的影响。结果发现,传统的方法如果不考虑扩散限制的影响,所预测的潜在分布范围远超过物种真实的分布范围。研究者认为,虚拟物种简化了真实物种复杂的情景,这是未受控制的真实物种很难实现和比较的。因为,即使最简单的真实物种,也存在不清晰的难以通过实验解决的问题(如物种扩散动态、物种之间的相互作用、物种与环境之间的相互关系)。

3.4 评估物种出现判定标准的影响

物种分布模型的预测结果通常为出现概率分布图,常用的做法是通过选择一个阈值将连续出现概率转化为二进制地图,即超过这个阈值代表物种出现,低于这个阈值代表物种消失。阈值的选择对模型的预测效果具有重要作用,不同的阈值选取方法影响了模型的评估结果^[36,46-47]。常用的阈值标准是 0.5,但是这并不一定符合真实的情况。比如,当大量物种出现位点数据(出现数据)没有被记录的时候,出现/缺失的平衡点就低于 0.5,而如果不通过比较就选择 0.5 作为阈值时将导致很多的出现位点被误认为缺失位点被预测(高的数据遗失率),增加了模型的特异度(Specificity),降低了模型的敏感度(Sensitivity)。而如果选择低于 0.5 作为阈值,又会增加模型的敏感度,降低模型的特异度。所以,应根据不同的出现/缺失比值选择合适的阈值标准。阈值标准的选择取决于假阳性(False Positive)和假阴性(False Negative)的比值,也依赖于所选择的模型。因此,有必要对不同模型的阈值标准进行对比分析。

考虑到真实物种并不能完全知道物种所有的出现点,也不可能知道所有的不出现点,因为虚假的真实和虚假的缺失点均会影响模型的敏感度和特异度。为避免真实物种数据源错误对阈值标准选择的影响(如避免对出现/缺失数据源分类错误、避免未考虑的因素或未知的因子影响),确保模型的预测精度,有必要使用虚拟物种进行比较分析。例如,Jiménez-Valverde 和 Lobo^[36]在欧洲区域构建了虚拟物种,比较了四种阈值方法(0.5、最大 Kappa、敏感度与特异度之差最小值、敏感度与特异度之和最大值)。结果发现,敏感度与特异度之和最大值,敏感度与特异度之差最小值这 2 个指标作为阈值得到的预测精度最高,而广泛使用的阈值 0.5 和最大 Kappa 是最差的。他们指出对于真实物种而言,阈值的选择取决于物种出现概率,应根据研究目的选择合适的阈值标准。此外,Meynard 和 Kaplan^[48]通过虚拟物种评价了阈值法和概率法的差异,结果发现阈值法的结果与概率法的结果并不一致。其原因可能归因于二者强调不同的方面,阈值法强调了物种能否出现问题,而概率法主要强调了物种出现概率高低问题,后者更注重随机因素的作用。因此,在研究随机因素的影响时,应优先选择概率法。

4 虚拟物种构成算法的不足及未来的发展

4.1 避免过度脱离真实与开发新的算法

当然,虚拟物种有可能不正确地模拟现实,引入错误或偏差的结果^[48-50],应通过以下几种方式降低这种风险:1)使用真实的生态地理数据。模拟预测在某些情况下可能是有趣的,如探索模型灵敏度、评估预测精度及评价特定的方法等。但物种分布可能受到特定地理因素的限制,建立在真实地理信息系统上的虚拟物种分布可在一定程度上避免过度脱离真实的风险。但是,不同地理图层存在差异,如特定空间的气候特征、空间自相关性、特定位置可到达性等,导致建立在不同地理图层上的研究可能会得到不一致的结果,未来有必要加强真实背景地图对模拟效果的评价。2)进一步细化物种与环境的线性或非线性函数关系。物种分布与环境的关系存在着线性、截性及高斯等多种关系^[35],只有进一步细化特定环境因子在具体的环境空间与物种分布的关系,才能更精确地虚拟现实。3)引入随机因素(统计随机性、遗传随机性和环境随机性)^[51-52]。复合种群

理论强调了随机灭绝过程和随机拓展过程对破碎化生境下物种分布的影响。比如,个别物种可能在特定的斑块或者残留的生境保存,但是随机灭绝过程可能导致物种的缺失。特别是当物种的数量很低或者小生境出现波动(如捕食者或者寄生病的出现)时,物种随机灭绝的概率更加大。因此,可引入复合种群模型的随机理论来校正和测试现有的物种分布模型,因为随机灭绝过程对小种群的分布是至关重要的。4)通过真实物种与虚拟物种的结合来解决特定的科学问题。尽管一些研究仅仅使用了虚拟物种,但是更多的研究认为使用真实与虚拟的结合才能更加真实地模拟自然过程^[8,39]。一些研究证实^[8],真实物种与虚拟物种虽然在预测精度及预测范围具有一致性,但是二者之间依然存在一定的差异。虚拟物种是一种简化的物种,忽略了物种的复杂性和多样性。而真实的世界更加复杂,所以从虚拟物种中得到的结论可能是有限的。比如,Wunder 等^[53]发现当使用虚拟物种来确定必要的样本大小以提高模型的预测精度时,可能会得到较窄的下限阈值。Berger 等^[54]模拟了蚱蜢的随机运动,发现运动的随机偏差会造成不同的预测结果。考虑到虚拟物种构建过程中存在的不确定性因素(如:输入的数据、模型的假设前提、参数设置和模拟程序),因此,应通过真实物种与虚拟物种的结合来解决特定的科学问题。

虚拟物种是基于重要生态假设和理论来解决实际问题的一种简化物种,但是现有的构成算法(如加法、乘法和综合法)不符合特定的生态学原理及规律,存在一定的局限性,这可能成为限制虚拟物种方法推广的一个重要因素。在后期的研究中,应开发新的算法,避免现有算法的缺陷,主要包括:

1)避免因子间的补偿效应。从理论上讲,特定的限制因子,如过强或过弱的光照,过高或过低的温度(如极端温度、年最低温及最冷月最低温),稀少的降水(年降水量及最干旱季节的降水量)都可能限制某个物种的分布,但现有的方法会造成补偿效应,弱化了限制因子的作用^[35]。比如,假设一个物种出现的生境适宜度阈值为 0.5,而如果通过 2 个因子来确定物种最后的生境适宜度(假设 2 个因子具有相同的权重),其适宜度分别为 0.8 和 0.3。从理论上讲,由于其中一个因子的生境适宜度低于 0.5,该物种在该区域不能分布,但求和法获得的算术平均数为 0.5,可认为该物种可以在该区域分布。这就造成了特定的补偿效应,即适宜的环境因子弱化了不适宜因子的限制作用。同样,求积法假定环境因子之间存在着相互作用^[55],也会造成类似的补偿效应。这些算法缺乏充足的生态学依据,偏离了一些基本的生态学理论,如不符合生态学的最小因子定律^[56]。李比希的最小因子定律^[56]认为,物种的生长不受可用资源总量的控制,而是受最为稀缺资源(限制因素)的控制。

2)构建能反映物种特性的动态算法。现有方法无法反映物种的动态变化,未能考虑一些生物学特性(比如,繁殖代价及成本、扩散强度及扩散方向、对环境的适应力及对危险的躲避力、进化水平及方向)等^[57-58]。近年来,随着计算机图形学的发展,建立生命模型并动态的模仿生物的生长、发育、死亡等生命活动已经成为当今虚拟现实领域的研究热点,虚拟生物的生长建模与可视化引起了国内外学者广泛关注。若能把物种自身的特性与各种外部因素融入到虚拟生物模型中,将会有助于更真实有效地仿真生命^[59]。比如,可在植物生长建模中,融入植物自身的生理特性(光合作用特性、耐阴性、根系吸水特性等)及环境影响因素,建立具有更广泛生态学意义的虚拟植物生长模型,能更真实地反映了植物在共享资源条件下相互作用的规律。而开发动态的虚拟物种算法,有助于研究者解决一些过去很难研究或者研究周期特别长的科学问题,比如模拟入侵物种的扩散路径,模拟物种的生长、繁殖及扩散过程等。

当前还没有提出一个评估虚拟物种算法好坏的标准,这限制了虚拟物种的发展。如评价模拟对象的适用性(是否适合动物、植物、微生物,广布种或局域种,水生生物或陆生生物等);研究尺度的适宜性(大尺度下和微尺度下构建的虚拟物种是否应该不同)等。为虚拟物种算法提供一个评价体系或者框架,能很好的指导和评价这些算法。

4.2 虚拟模式生物、群落和生态系统

生命科学研究常选用受到广泛研究,对其生物现象有深入了解的物种作为模式生物,根据模式生物所得的研究结果,归纳出一些具有普遍意义的生物模型。如植物的拟南芥(*Arabidopsis thaliana*)、水稻(*Oryza*

sativa)、烟草(*Nicotiana tabacum*)等;动物的非洲爪蟾(*Xenopus laevis*)、斑马鱼(*Danio rerio*)、小鼠(*Rattus norvegicus*)、线虫(*Caenorhabditis elegans*)、黑腹果蝇(*Drosophila melanogaster*)等。未来虚拟物种的发展也可构建不同类群的模式物种,这些模式物种应该具有鲜明的特点,如虚拟的广布种或局域种;虚拟的水生生物或陆生生物;虚拟的入侵植物及入侵动物等,具有开源的标准程序等。

众多的生态学过程(包括生态位分化、环境过滤、有限的传播、生态位保守和收敛)影响了群落的构成^[60],可通过虚拟物种的方法研究这些过程对群落结构的影响和测试其对零假说的偏离(如本地群落是否是中性)^[61]。也有一些研究通过构建虚拟物种来模拟群落和生态系统,研究特定的生态学规律。例如,Galiana等^[62]进行了一个电脑模拟实验,通过生态角色和营养级的划分,引入虚拟物种构建了一个虚拟的食物网。通过改变一些入侵物种的特性,包括他们的食物宽度、捕食者数量、能量阈值(低于这个阈值,捕食者会灭绝)发现,入侵物种可通过降低本地种的丰富度和每个物种食物资源的数量,改变食物网的结构,增加食物资源单一物种灭绝风险;当面临外来物种入侵的时候,结构简单的食物网很容易被破坏。因此,充分考虑多群落和物种相互作用的整体分析方法有利于更好地理解生物入侵对复杂群落的影响。Zurell等^[29]利用构建的虚拟生态系统(蝴蝶-寄主植物-捕食者),测试瞬态动力学和重要生态过程(生态位宽度、分布和繁殖、种间生态过程,如竞争、捕食、环境特性转化和气候变化)对物种分布预测精度的影响,探讨气候变化下物种分布模型的性能。结果证实模型的预测性能依赖于主要的生态学过程(如物种的传播能力及物种的灭绝速率)和时间动态,快速传播的物种能快速适应气候的变化,其预测精度高于传播能力有限的物种;当考虑蝴蝶与寄主之间的相互作用时,模拟的结果更加精确;在气候变化下,没有寄生虫及低的种群增长速率导致模型高的预测精度。Zurell等^[29]强调,物种的分布模型应当考虑物种在特定时空尺度下的散布规律,强化散布、出现及灭绝边缘的研究,模型必须考虑生物间的相互作用(比如,寄主、寄生虫等)。

可见,通过虚拟数据可以模拟这些重要的生态学过程(模拟种群的增长、随机扩散、生态位分化),研究种群的动态、群落结构及装配规则,探讨中性理论和生态位理论在生物多样性维持中的作用。也可以通过模拟食物链及食物网关系(食物链或食物网的长度、复杂性、捕食者及猎物寡食性或广食性等),能量关系(能量级的损耗、转化效率及每个级别中所需的最低能量)等构建虚拟的生态系统。未来,科学家可在虚拟群落或生态系统基础上,进行很多实用高效的模拟,如模拟生态多样性锐减给生态系统结构及稳定性问题,全球气候变暖是否是造成濒危物种灭绝的主要因素等。当然,这些研究也依赖于完善的软件系统,而开发功能强大、灵活、算法种类全面并方便研究者自行创建新算法或者改变旧算法的虚拟物种应用程序将有效地推动虚拟物种的研究。

参考文献(References):

- [1] Franklin J. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge: Cambridge University Press, 2010.
- [2] Liu C R, White M, Newell G. Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, 2013, 40(4): 778-789.
- [3] 李国庆,刘长成,刘玉国,杨军,张新时,郭柯. 物种分布模型理论研究进展. *生态学报*, 2013, 33(16): 4827-4835.
- [4] 乔慧捷,胡军华,黄继红. 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*, 2013, 43(11): 915-927.
- [5] 许仲林,彭焕华,彭守璋. 物种分布模型的发展及评价方法. *生态学报*, 2015, 35(2): 557-567.
- [6] Nix H A. A biogeographic analysis of Australian elapid snakes// Longmore R, ed. *Snakes: Atlas of Elapid Snakes of Australia*. Cornell: Australian Government Pub. Service, 1986: 4-15.
- [7] 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.
- [8] Václavík T, Meentemeyer R K. Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 2012, 18(1): 73-83.
- [9] 朱耿平,刘强,高玉葆. 提高生态位模型转移能力来模拟入侵物种的潜在分布. *生物多样性*, 2014, 22(2): 223-230.
- [10] Ferrier S. Mapping spatial pattern in biodiversity for regional conservation planning: where to from here?. *Systematic Biology*, 2002, 51(2):

331-363.

- [11] Peterson A T. Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, 2011, 38(5): 817-827.
- [12] 张雷, 刘世荣, 孙鹏森, 王同立. 气候变化对马尾松潜在分布影响预估的多模型比较. *植物生态学报*, 2011, 35(11): 1091-1105.
- [13] Engler R, Guisan A, Rechsteiner L. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 2004, 41(2): 263-274.
- [14] Araújo M B, Peterson A T. Uses and misuses of bioclimatic envelope modeling. *Ecology*, 2012, 93(7): 1527-1539.
- [15] Guralnick R P, Hill A W, Lane M. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters*, 2007, 10(8): 663-672.
- [16] Feeley K J, Silman M R. Modelling the responses of Andean and Amazonian plant species to climate change: the effects of georeferencing errors and the importance of data filtering. *Journal of Biogeography*, 2010, 37(4): 733-740.
- [17] Lahoz-Monfort J J, Guillera-Aroita G, Wintle B A. Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, 2014, 23(4): 504-515.
- [18] Dormann C F, McPherson J M, Araújo M B, Bivand R, Bolliger J, Carl G, Davies R G, Hirzel A, Jetz W, Kissling W D, Kühn I, Ohlemüller R, Peres-Neto P R, Reineking B, Schröder B, Schurr F M, Wilson R. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 2007, 30(5): 609-628.
- [19] Saas Y, Gosselin F. Comparison of regression methods for spatially autocorrelated count data on regularly and irregularly spaced locations. *Ecography*, 2014, 37(5): 476-489.
- [20] Rapacciuolo G, Roy D B, Gillings S, Purvis A. Temporal validation plots: quantifying how well correlative species distribution models predict species' range changes over time. *Methods in Ecology and Evolution*, 2014, 5(5): 407-420.
- [21] Royle J A, Nichols J D, Kéry M. Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, 2005, 110(2): 353-359.
- [22] Wieczorek J, Guo Q H, Hijmans R. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science*, 2004, 18(8): 745-767.
- [23] Boakes E H, McGowan P J K, Fuller R A, Ding C Q, Clark N E, O'Connor K, Mace G M. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology*, 2010, 8(6): e1000385.
- [24] Grimm V, Revilla E, Berger U, Jeltsch F, Mooij W M, Railsback S F, Thulke H H, Weiner J, Wiegand T, DeAngelis D L. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science*, 2005, 310(5750): 987-991.
- [25] Fourcade Y, Engler J O, Rödder D, Secondi J. Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS One*, 2014, 9(5): e97122.
- [26] Hirzel A H, Helfer V, Metral F. Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, 2001, 145(2): 111-121.
- [27] Austin M P, Belbin L, Meyers J A, Doherty M D, Luoto M. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. *Ecological Modelling*, 2006, 199(2): 197-216.
- [28] Zurell D, Berger U, Cabral J S, Jeltsch F, Meynard C N, Münkemüller T, Nehrbaß N, Pagel J, Reineking B, Schröder B, Grimm V. The virtual ecologist approach: Simulating data and observers. *Oikos*, 2010, 119(4): 622-635.
- [29] Zurell D, Jeltsch F, Dormann C F, Schröder B. Static species distribution models in dynamically changing systems: How good can predictions really be?. *Ecography*, 2009, 32(5): 733-744.
- [30] Varela S, Anderson R P, García-Valdés R, Fernández-González F. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography*, 2014, 37(11): 1084-1091.
- [31] Stickel L F. A comparison of certain methods of measuring ranges of small mammals. *Journal of Mammalogy*, 1954, 35(1): 1-15.
- [32] Swan J M A. An examination of some ordination problems by use of simulated vegetational data. *Ecology*, 1970, 51(1): 89-102.
- [33] Fortin M J, Drapeau P, Legendre P. Spatial autocorrelation and sampling design in plant ecology. *Vegetatio*, 1989, 83(1/2): 209-222.
- [34] Jiménez-Valverde A, Lobo J M, Hortal J. The effect of prevalence and its interaction with sample size on the reliability of species distribution models. *Community Ecology*, 2009, 10(2): 196-205.
- [35] Meynard C N, Quinn J F. Predicting species distributions: A critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, 2007, 34(8): 1455-1469.
- [36] Jiménez-Valverde A, Lobo J M. Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecologica*, 2007, 31(3): 361-369.
- [37] Hutchinson M. Methods for generation of weather sequences // Bunting A H, ed. *Agricultural Environments: Characterisation, Classification and Mapping*. Wallingford: CAB, International, 1987: 149-157.
- [38] Barbet-Massin M, Jiguet F, Albert C H, Thuiller W. Selecting pseudo-absences for species distribution models: how, where and how many?.

- Methods in Ecology and Evolution, 2012, 3(2): 327-338.
- [39] Bombi P, D'Amen M. Scaling down distribution maps from atlas data: A test of different approaches with virtual species. *Journal of Biogeography*, 2012, 39(4): 640-651.
- [40] Naimi B, Hamm N A S, Groen T A, Skidmore A K, Toxopeus A G. Where is positional uncertainty a problem for species distribution modelling?. *Ecography*, 2014, 37(2): 191-203.
- [41] Elith J, Graham C H. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 2009, 32(1): 66-77.
- [42] Duan R Y, Kong X Q, Huang M Y, Wu G L, Wang Z G. SDMvspecies: a software for creating virtual species for species distribution modelling. *Ecography*, 2015, 38(1): 108-110.
- [43] Lobo J M, Tognelli M F. Exploring the effects of quantity and location of pseudo-absences and sampling biases on the performance of distribution models with limited point occurrence data. *Journal for Nature Conservation*, 2011, 19(1): 1-7.
- [44] 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.
- [45] Saupe E E, Barve V, Myers C E, Soberón J, Barve N, Hensz C M, Saupe E E, Barve V, Myers C E, Soberón J, Barve N, Hensz C M, Peterson A T, Owens H L, Lira-Noriega A. Variation in niche and distribution model performance: the need for a priori assessment of key causal factors. *Ecological Modelling*, 2012, 237-238: 11-22.
- [46] Franklin J. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge: Cambridge University Press, 2009.
- [47] Liu C R, Berry P M, Dawson T P, Pearson R G. Selecting thresh-olds of occurrence in the prediction of species distributions. *Ecography*, 2005, 28(3): 385-393.
- [48] Meynard C N, Kaplan D M. Using virtual species to study species distributions and model performance. *Journal of Biogeography*, 2013, 40(1): 1-8.
- [49] Miller J A. Virtual species distribution models using simulated data to evaluate aspects of model performance. *Progress in Physical Geography*, 2014, 38(1): 117-128.
- [50] Hanberry B B, He H S, Palik B J. Pseudoabsence generation strategies for species distribution models. *PLoS One*, 2012, 7(8): e44486.
- [51] Beale C M, Lennon J J. Incorporating uncertainty in predictive species distribution modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2012, 367(1586): 247-258.
- [52] Marini L, Bruun H H, Heikkinen R K, Helm A, Honnay O, Krauss J, Kühn I, Lindborg R, Pärtel M, Bommarco R. Traits related to species persistence and dispersal explain changes in plant communities subjected to habitat loss. *Diversity and Distributions*, 2012, 18(9): 898-908.
- [53] Wunder J, Reineking B, Bigler C, Bugmann H. Predicting tree mortality from growth data: how virtual ecologists can help real ecologists. *Journal of Ecology*, 2008, 96(1): 174-187.
- [54] Berger U, Wagner G, Wolff W F. Virtual biologists observe virtual grasshoppers: an assessment of different mobility parameters for the analysis of movement patterns. *Ecological Modelling*, 1999, 115(2): 119-127.
- [55] Prasad A M, Iverson L R, Liaw A. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, 2006, 9(2): 181-199.
- [56] Liebig J, Playfair L P B. Chemistry in Its Application to Agriculture and Physiology. Cambridge: TB Peterson, 1847.
- [57] 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.
- [58] Wisz M S, Pottier J, Kissling W D, Pellissier L, Lenoir J, Damgaard C F, Dormann C F, Forchhammer M C, Grytnes J A, Guisan A, Heikkinen R K, Høye T T, Kühn I, Luoto M, Maiorano L, Nilsson M C, Normand S, Öckinger E, Schmidt N M, Termansen M, Maiorano L, Nilsson M C, Normand S, Öckinger E, Schmidt N M, Termansen M, Timmermann A, Wardle D A, Aastrup P, Svenning J C. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, 2013, 88(1): 15-30.
- [59] Mitchell M, Taylor C E. Evolutionary computation: an overview. *Annual Review of Ecology and Systematics*, 1999, 30: 593-616.
- [60] Baiser B, Buckley H L, Gotelli N J, Ellison A M. Predicting food-web structure with metacommunity models. *Oikos*, 2013, 122(4): 492-506.
- [61] Hardy O J. Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *Journal of Ecology*, 2008, 96(5): 914-926.
- [62] Galiana N, Lurgi M, Montoya J M, López B C. Invasions cause biodiversity loss and community simplification in vertebrate food webs. *Oikos*, 2014, 123(6): 721-728.