物种分布模型中空间变化系数的使用指南

Jeffrey W. Doser[1, 2], Marc Kéry[3], Sarah P. Saunders[4], Andrew O. Finley[2,5,6], Brooke L. Bateman[4], Joanna Grand[4], Shannon Reault[4], Aaron S. Weed[7], Elise F. Zipkin[1, 2]

[1]密歇根州立大学综合生物学系，美国密歇根州东兰辛[2]密歇根州立大学生态学、进化与行为项目，美国密歇根州东兰辛[3]瑞士鸟类学研究所，瑞士森帕赫[4]美国国家奥杜邦协会科学部，美国纽约州纽约[5]密歇根州立大学林业系，美国密歇根州东兰辛[6]密歇根州立大学统计与概率系，美国密歇根州东兰辛[7]美国国家公园管理局东北温带清查与监测网络，美国佛蒙特州伍德斯托克

通讯作者：Jeffrey W. Doser，电子邮箱：doserjef@msu.edu；ORCID ID：0000-0002-8950-9895

arXiv:2301.05645v2 [stat.AP] 4 Aug 2023

数据可用性声明

与本文相关的所有数据和代码均可在 GitHub 上获取（https://github.com/doserjef/Doser_et_al_2023_SVC），并将在论文被接收后发布于 Zenodo。

致谢

本研究获得美国国家科学基金会（NSF）资助，资助项目编号为 DBI-1954406、DMS-1916395 和 DEB-2213565。

个人简介

Jeffrey W. Doser 是一位统计生态学家，致力于开发和应用层次贝叶斯模型，为不同空间尺度上的野生动物保护和自然资源保护提供参考。

摘要

Aim

物种分布模型（SDMs）正越来越多地应用于宏观尺度，并使用探测—未探测数据。这类模型通常假设，一组固定的回归系数足以描述物种—环境关系和/或种群趋势。然而，这些关系往往呈现非线性和/或空间变化模式，这是不同尺度上与非生物过程和生物过程复杂交互作用的结果。空间变系数（SVC）模型能够轻松地考虑环境协变量效应的变异。然而，由于人们尚未充分了解与较简单的框架相比，SVC模型能够提供哪些推断优势，因此它们在生态学中的应用相对有限。

创新

本文展示了空间变系数物种分布模型的推断优势，尤其关注如何运用这一方法提出并检验生态学假设，以探究种群趋势和物种—环境关系空间变异的驱动因素。我们通过模拟和两个案例研究展示空间变系数物种分布模型的推断优势：一项研究评估了过去二十年美国东部51种森林鸟类物种的空间变化趋势；另一项研究则评估了过去五十年土地覆盖变化对草地鹨（*Ammodramus savannarum*）在美国本土分布影响的空间变异。

主要结论

与更简单的替代方案相比，我们在两个实证案例研究中都发现了有力证据支持空间变系数物种分布模型。空间变系数所捕捉的精细空间尺度因素，是森林鸟类出现趋势空间变异的主要驱动因素。此外，对于草地和农田面积与草原鹨之间复杂的物种—栖息地关系，空间变系数揭示了其中的复杂性，为未来土地利用变化可能如何塑造其分布提供了细致入微的见解。这些应用展示了空间变系数物种分布模型的用途：帮助揭示驱动物种在局地和广阔尺度上分布的环境因素。最后，我们讨论空间变系数物种分布模型在生态学和保护领域的潜在应用。

引言

阐明驱动物种分布的因素是生态学的一项基本目标。物种分布模型（SDMs）是研究物种在时空中何处出现以及为何出现的主要工具（Guisan 和 Zimmermann，2000年）。物种分布模型可以利用多种类型的数据（例如，仅出现数据、丰度数据），但通常会在基于参数化广义线性模型（广义线性模型）的框架中使用出现-缺失（或探测-未探测）数据，从而评估物种—环境关系及局地出现的概率。参数化物种分布模型通常通过一组回归系数（例如线性项和二次项）来描述数据集空间范围内的物种—环境关系，从而针对模型中纳入的所有环境预测因子，得到一组线性或单峰响应曲线（Guisan 等人，2002年）。当空间范围涵盖整个物种分布范围时，这类物种—环境响应曲线的组合刻画了物种的多变量实际环境生态位（Guisan 等人，2017）。然而，在广阔的空间范围内拟合物种分布模型时，单一的一组线性和二次回归系数可能无法充分描述物种—环境关系。这可能是因为真实关系呈现复杂的非线性形式、关系在空间上存在变异（即关系取决于具体情境，或具有非平稳性；Rollinson 等人 2021），也可能仅仅是由于不同地区的栖息地分类不一致（例如，在模型域的不同区域中，“干旱草原”并非指完全相同的栖息地类型）。

复杂、非线性和／或随空间变化的物种—环境关系，可能源于在不同尺度上发挥作用的多种非生物和生物过程（Osborne 和 Suárez-Seoane，2002年；Miller，2012；Clark 和 Luis，2020）。与非生物因素的交互作用，例如历史干扰制度、景观组成／配置、细尺度生境特征（如植被质量）以及环境条件（如土壤含量、温度），可能导致环境因素在物种分布范围内对物种产生不同影响（Rollinson 等人，2021）。鉴于资源可利用性存在空间异质性，与资源更丰富的地区相比，环境因素对物种出现的影响在资源有限地区可能更强（Pease 等人，2022a）。如果气候塑造物种分布，那么气候变化对物种的影响在其气候极限附近的地区应最强（Amburgey 等人，2018年）。例如，Sultaire 等人（2022年）发现，气温升高和积雪持续时间对雪鞋兔（*Lepus ameri-*

canus）出现的影响存在空间变异，这表明气候以不同方式限制着该物种在其分布范围内的分布。或者，随空间变化的物种—环境关系也可能源于生物过程，例如局部遗传适应或物种相互作用的空间变异（如捕食、竞争）。Pease 等人（2022a）发现，在北卡罗来纳州，森林覆盖对白尾鹿（*Odocoileus virginianus*）出现的影响存在空间变异，这在一定程度上是由该州捕食压力的差异所致。若存在此类随空间变化的物种—环境关系却未予以考虑，可能会导致对影响物种出现地点的非生物因素作出误导性推断，并对管理和保护建议产生重要影响（Rollinson 等人，2021）。

除了用于刻画物种—环境关系外，监测数据也常用于物种分布模型，以量化出现趋势。非线性和/或空间变化的出现趋势，主要源于影响目标物种的非生物因素和生物因素发生了时空变化（例如，某物种分布范围内不同区域的土地利用变化程度存在差异）。不过，这类模式也可能源于物种与这些非生物因素和生物因素之间的复杂关系（例如，土地利用变化在物种分布范围的某个区域产生的影响与其他区域不同）。量化物种出现情况随时间的空间变异，是生物多样性监测计划的一项常见目标（Bled 等人，2013；Babcock 等人，2016；Meehan 等人，2019年），因为这类见解有助于提出有关种群变化驱动因素的假设（例如，Crossley 等人 2021），确定保护或恢复的优先区域（例如，Ethier 等人 2017），并深入了解哪些地方需要加大监测力度以减少不确定性（Reich 等人，2018年）。

在生态学文献中，人们使用了多种方法来检验关于物种—环境关系或出现趋势的空间变异和非线性的假设。带有非生物变量和/或生物变量交互作用的广义线性模型简单而直观，可用于评估非线性和/或随空间变化的关系（Spake 等人，2023年）。此外，在预先定义的分层（例如生态区、管理单元）中分别估计斜率（固定效应或随机效应），也是另一种常见的替代方法（例如，Smith 和 Edwards 2021）。此外，非线性函数（例如阈值）也可轻松纳入参数化物种分布模型中（Hostetler 和 Chandler，2015年）。然而，这些方法要求事先了解与关注变量发生交互作用的协变量、此类关系的函数形式，和/或关系中变异的空间分辨率；而这些信息大多在分析前并不清楚。

此外，在跨宏观尺度开展研究时，明确不同空间尺度上多个驱动因素之间的交互作用尤为关键，但使用上述方法很难做到。因此，我们需要更灵活的方法，能够轻松处理复杂的物种—环境关系，而无需事先了解所有交互变量及其交互作用的性质。

因此，人们广泛使用机器学习方法，例如 Max- Ent（Phillips 等人，2006）和随机森林（Liaw 等人，2002年），在考虑复杂物种—环境关系的同时对物种分布建模。MaxEnt 使用协变量上的不同“特征类别”组合（即数学函数）来表征仅出现数据中的非线性关系；随机森林则根据协变量空间划分数据，并在各个分区中拟合由分类树或回归树组成的集成模型（Valavi 等人，2021）。这两种方法在生态学和保护领域应用广泛，但在提供物种—环境关系和/或出现趋势的不确定性估计方面能力有限；它们要求所有交互变量均已知并纳入模型；而且无法处理不完全探测（即某物种存在于某地点，却未能观察到它的情况；Kéry 2011年；Kellnerand Swihart2014）。

本文讨论空间变化系数（SVC）在物种分布模型中的应用。这是一种高度灵活的方法，可用于模拟非线性和/或随空间变化的物种—环境关系，无论这些关系是如何产生的。空间变系数模型是广义线性模型的直观扩展，允许回归系数随空间平滑变化。通过在分层贝叶斯框架中拟合空间变系数模型，我们可以针对感兴趣的空间区域，生成物种—环境效应和/或趋势的预测，并完整传播不确定性，同时考虑普遍存在的观测误差，例如由不完全探测造成的误差。近期研究表明，由于宏观尺度上物种—环境关系普遍存在异质性，这一灵活框架在各种生态学应用中的受关注程度日益提高（例如，Meehan 等人 2019年；Rollinson 等人 2021；Sultaire 等人 2022年）。然而，相较于其他方法，空间变系数所能带来的推断优势尚未得到全面认识。Thorson 等人（2023年）近期指出，有七个生态学问题可以通过空间变化系数加以研究。本文在其工作的基础上，通过模拟和两项针对美国繁殖鸟类的案例研究，明确展示空间变系数物种分布模型相较于较简单方法的推断优势。最后，我们将讨论

何时应使用空间变系数物种分布模型来替代或补充较简单方法的实用指南。

什么是空间变化系数?

我们的案例研究聚焦于观测不完全的鸟类物种。因此，我们在占域模型（MacKenzie 等人，2002年；Tyre 等人，2003年）的框架下讨论空间变系数。这类模型是用于刻画不完全探测的层次广义线性模型的特定形式，不过，我们的研究发现可直接推广到所有参数化物种分布模型（即广义线性模型）。下面简要概述空间变系数占域模型。有关完整的统计细节，请参阅 Doser 等人（2023年）。

设 $\boldsymbol{s}_j$ 表示地点 j的空间坐标，其中 $j = 1$、……、J分别在 $t = 1$、……、T个主要时间段（以下称为“季节”）内进行采样。请注意，数据可能仅在一个季节内获得（即 $T = 1$），也可能来自多季节数据集；在这类数据集中，地点无需每个季节都进行采样（即允许存在缺失值）。为考虑不完全探测，在季节 t期间，于地点 j进行 $k = 1$、……、K_t（$\boldsymbol{s}_j$）次重复采样，以判断目标物种未被探测到是否确实意味着其不存在（MacKenzie 等人，2002年；Tyre 等人，2003年）。请注意，不同地点／季节组合的重复采样次数可能不同，而且并非每个地点在每个季节都必须进行多次重复采样（例如，von Hirschheydt 等人 2023年）。此类重复采样通常是在一个季节内的短时间段里多次访问同一地点，但也可以采用其他重复方式，例如空间子抽样和由多名观察者进行观测（MacKenzie 等人，2017年）。设 $y_{t,k}$（$\boldsymbol{s}_j$）表示在季节 t的调查 k 期间，于地点 j 观测到研究物种（1）或未探测到（0）；设 z_t（$\boldsymbol{s}_j$）表示该物种在季节 t期间于地点 j 真实存在（1）或不存在（0）。注意，我们假设在同一季节内的重复调查中，z_t（$\boldsymbol{s}_j$）不会发生变化（即“闭合”假设）。我们根据该物种在季节 t（z_t（$\boldsymbol{s}_j$））期间于地点 j 的真实出现状态，对观测数据 $y_{t,k}$（$\boldsymbol{s}_j$）进行建模。具体而言，我们有

$$y_{t,k}(\boldsymbol{s}_j) \sim \begin{cases} 0, & z_t(\boldsymbol{s}_j) = 0 \\ \text{Bernoulli}(p_{t,k}(\boldsymbol{s}_j)), & z_t(\boldsymbol{s}_j) = 1 \end{cases}, \tag{1}$$

其中，$p_{t,\ k}(\boldsymbol{s}_j)$ 表示在地点 j 、季节 t的第 k次重复调查中探测到该物种的概率。我们将探测概率建模为地点、季节和/或调查层级

（即观测层级）协变量的函数，具体如下：

$$\text{logit}(p_{t,k}(\boldsymbol{s}_j)) = \boldsymbol{v}_{t,k}(\boldsymbol{s}_j)\boldsymbol{\alpha}, \tag{2}$$

其中 $\boldsymbol{\alpha}$ 是回归系数向量（包括截距），用于描述协变量 $\boldsymbol{v}_{t、k}(\boldsymbol{s}_j)$ 对探测的影响。

真实出现状态 $z_t(\boldsymbol{s}_j)$ 是一个部分可观测变量，因此，如果 $y_{t、k}(\boldsymbol{s}_j) = 1$，我们就知道 $z_t(\boldsymbol{s}_j) = 1$（因为我们假设不存在假阳性）；但如果 $y_{t、k}(\boldsymbol{s}_j) = 0$，我们就无法确定该物种在该地点是真正缺失，还是我们未能探测到它。因此，我们将$z_t(\boldsymbol{s}_j)$建模为

$$z_t(\boldsymbol{s}_j) \sim \text{Bernoulli}(\psi_t(\boldsymbol{s}_j)), \tag{3}$$

其中 $\psi_t(\boldsymbol{s}_j)$ 是该物种在地点 j、季节 t的出现概率。拟合占域模型时，我们可以在广义线性模型框架下，通过协变量对出现概率 $\psi_t(\boldsymbol{s}_j)$ 的影响来估计物种—环境关系。为简单起见，考虑一个环境变量 $\text{x}_t(\boldsymbol{s}_j)$，它随各空间位置 $\boldsymbol{s}_j$ 和季节 t变化（例如温度、降水）。空间变系数占域模型的形式为

$$\text{logit}(\psi_t(\boldsymbol{s}_j)) = \beta_0 + \eta_t + \text{w}_0(\boldsymbol{s}_j) + \beta_1 \cdot \text{x}_t(\boldsymbol{s}_j) + \text{w}_1(\boldsymbol{s}_j) \cdot \text{x}_t(\boldsymbol{s}_j), \tag{4}$$

其中，β_0 是截距，η_t 是时间随机效应（即特定于季节的截距），用于解释出现概率中未建模的时间自相关（仅适用于 $T > 1$）；$\text{w}_0(\boldsymbol{s}_j)$是空间随机效应，用于解释出现概率中未建模的空间变异；β_1 是协变量 $\text{x}_t(\boldsymbol{s}_j)$的非空间效应；$\text{w}_1(\boldsymbol{s}_j)$则是该协变量在每个空间位置 $\boldsymbol{s}_j$处的空间变化效应。

SVC模型中的空间变化效应可以通过多种方式估计，包括广义加性模型（广义加性模型；Wood 2006）和高斯过程（Banerjee 等人，2014年）。鉴于高斯过程在空间统计学中的广泛应用、其相对较高的预测性能（Golding 和 Purse，2016），以及广义加性模型可能导致关系过度平滑（Stein，2014年），我们重点讨论后者。具体而言，我们有

$$\text{w}_1(\boldsymbol{s}) \sim N(\boldsymbol{0}, \boldsymbol{C}(\boldsymbol{s}, \boldsymbol{s}', \boldsymbol{\theta})), \tag{5}$$

其中 $\boldsymbol{C}(\boldsymbol{s},\ \boldsymbol{s}',\ \boldsymbol{\theta})$ 是一个 $J \times J$ 协方差矩阵，它是任意一对站点坐标之间距离的函数 $\boldsymbol{s}$ 以及 $\boldsymbol{s}'$ 和一组参数（$\boldsymbol{\theta}$），这些参数根据空间相关函数控制空间过程。这里我们采用最近邻高斯过程（Datta 等人， 2016）作为完整高斯过程的高效近似，并使用指数相关函数，通过两个参数对协变量空间变化效应的相关性进行建模，参数为 $\boldsymbol{\theta} = \{\sigma^2,\ \phi\}$，其中 σ^2 是空间方差参数，而 ϕ 是空间衰减参数。关于空间变系数占域模型的详细信息，请参见 Doser 等人 (2023 年)。

空间变系数模型与更简单的替代模型相比如何?

为了与其他模型进行比较，请考虑一个更一般化的出现概率模型 $\psi_t(\boldsymbol{s}_j)$，其形式为

$$\text{logit}(\psi_t(\boldsymbol{s}_j)) = \beta_0 + \eta_t + \text{w}_0(\boldsymbol{s}_j) + f(\text{x}_t(\boldsymbol{s}_j), \boldsymbol{\beta}), \tag{6}$$

其中，β_0、η_t 和 $\text{w}_0(\boldsymbol{s}_j)$ 的定义见 (4)，而 $f(\text{x}_t(\boldsymbol{s}_j)\ \boldsymbol{\beta})$ 是一个通用函数，通过一组参数 $\boldsymbol{\beta}$ 将协变量 $\text{x}_t(\boldsymbol{s}_j)$ 与出现概率联系起来。请注意，如果 $\text{w}_0(\boldsymbol{s}_j)$ 和 η_t 不适用于目标物种/数据集，则可将其从 (6) 中删除。我们考虑五种函数关系来描述 $\psi_t(\boldsymbol{s}_j)$ 与 $\text{x}_t(\boldsymbol{s}_j)$ 之间的物种-环境关系；这些关系在估计非线性和/或空间变化的物种—环境关系方面的能力各不相同：

1. **线性**：$f(\text{x}_t(\boldsymbol{s}_j)\ \boldsymbol{\beta}) = \beta_1 \cdot \text{x}_t(\boldsymbol{s}_j)$

2. **二次型**：$f(\text{x}_t(\boldsymbol{s}_j)\ \boldsymbol{\beta}) = \beta_1 \cdot \text{x}_t(\boldsymbol{s}_j) + \beta_2 \cdot \text{x}_t^2(\boldsymbol{s}_j)$

3. **层**：$f(\text{x}_t(\boldsymbol{s}_j)\ \boldsymbol{\beta}) = \beta_1 \cdot \text{x}_t(\boldsymbol{s}_j) + \beta_{2,\text{STRATUM}_j} \cdot \text{x}_t(\boldsymbol{s}_j)$

4. **交互作用**：$f(\text{x}_t(\boldsymbol{s}_j)\ \boldsymbol{\beta}) = \beta_1 \cdot \text{x}_t(\boldsymbol{s}_j) + \beta_2 \cdot \text{x}_t^*(\boldsymbol{s}_j) \cdot \text{x}_t(\boldsymbol{s}_j)$

5. **空间变系数**：$f(\text{x}_t(\boldsymbol{s}_j)\ \boldsymbol{\beta}) = \beta_1 \cdot \text{x}_t(\boldsymbol{s}_j) + \text{w}_1(\boldsymbol{s}_j) \cdot \text{x}_t(\boldsymbol{s}_j)$

线性模型仅假设物种-环境关系呈线性。二次模型则在此基础上扩展，允许物种-环境关系在某个最优水平达到峰值，并随着偏离该最优水平而逐渐下降（即

凹关系），或在环境变量的极端值处达到峰值（即凸关系）。分层模型会估计环境预测变量的总体线性效应，以及该效应在一组分层（例如生态区、管理单元、保护区）中的特定层调整项。请注意，层特定偏差既可作为固定效应，也可作为随机效应进行估计。交互作用模型同样会估计线性的物种-环境关系，但允许这种关系通过与另一个在各 J 采样点之间变化的协变量发生交互作用而产生空间变异（$\mathbf{x}_t^*(\boldsymbol{s}_j)$）。SVC模型会估计总体线性的物种-环境关系，同时允许这种线性关系在数据集中的每个地点各不相同。SVC模型中的空间变化调整项（$\mathbf{w}_1(\boldsymbol{s}_j)$）用于对总体效应 β_1所表示的物种-环境关系进行局部调整。SVC模型可视为分层模型的扩展，其中各层现在是单个采样点；也可视为交互作用模型的扩展，其中参与交互作用的“协变量”未知，并在模型拟合过程中进行估计。

我们开展了一项模拟研究，比较上述五种模型，并评估它们估计不同形式的物种—环境关系的能力。简而言之，我们模拟了一个 20×20 网格上的探测-未探测数据，其中出现概率由单个协变量决定；该协变量在模拟区域南部取负值，在北部取正值。我们设置了六种不同的函数形式来表示真实的物种—环境关系（图1最左列）：(1) 线性；(2) 二次型；(3) 在九个分层中分别具有线性效应；(4) 与沿水平方向变化的第二个协变量存在交互作用；(5) 与一个未知（“缺失”）协变量存在交互作用；以及 (6) 上述五个部分之和（即“完整”效应）。对于每种物种—环境关系函数形式，我们模拟了 50 个数据集，并使用 `spOccupancy` R 软件包（Doser 等人，2022年）将五种模型分别拟合到每个数据集；其中交互作用模型仅纳入假定已知的协变量，而不纳入未知协变量。我们使用广泛适用信息准则（WAIC；Watanabe 2010年）比较模型表现。完整模拟细节见补充信息 S1。

在所有六种物种-环境关系形式中，SVC模型始终能够捕捉真实关系，而较简单的模型仅在部分情景下表现良好（图1）。当真实的物种-环境关系为线性时，所有模型都取得了近乎完美的表现

乎完全相同，且都与真实情况非常接近的物种—环境关系估计值（补充信息 S1，表 1）。较简单且预先指定了函数形式的模型（例如二次型、层、已知交互作用）在数据由确切的物种—环境关系生成时，根据 WAIC 的表现优于空间变系数模型，这表明当物种—环境关系的真实形式已知时，应使用较简单的模型，而非更复杂的空间变系数模型。然而，当真实的物种—环境关系不同于拟合模型所采用的形式时，较简单的模型捕捉潜在模式的能力有限（图 1），且根据 WAIC 的表现明显较差（补充信息 S1：表 S1）。即使数据并非由空间变系数模型生成，该模型仍能准确捕捉关系中的真实模式，这表明当物种—环境关系在拟合模型前未知时，空间变系数模型能够揭示其较简单的函数形式。此外，当物种—环境关系与某个未知变量发生交互作用（第 5 行，图1）或由多个组成部分决定时（第 6 行，图 1）。

案例研究1：美国东部森林鸟类空间变化的出现趋势

量化空间明确的趋势有助于识别保护关注区域（例如气候变化避难所），并深入了解物种分布范围动态（例如后缘与前缘趋势）。在本案例研究中，我们利用北美繁殖鸟类调查（BBS；$\sim$4.04 百万平方公里^2 ）地区（即西经100度线以东的美国本土）2000—2019年（$T = 20$ 年）的北美繁殖鸟类调查（BBS；Pardieck 等（2020））探测-未探测数据，评估了美国东部森林鸟类物种的出现趋势。我们依据Bateman 等人（2020）的栖息地分类，将分析范围限定为由66种东部森林鸟类组成的群落。随后，我们评估了其中51种鸟类的趋势；这些物种的繁殖分布范围（根据国际鸟盟（2021）确定）与研究区域至少有50%的重叠（物种名称见补充信息S2：表3）。本案例研究的目标是：（1）绘制美国东部这51种物种出现趋势的空间明确地图；（2）将空间变系数占域模型与三种代表分布变化驱动因素简单假设的替代模型进行比较

。具体而言，我们提出了以下四种假设及相应模型：

1. T该物种在美国东部呈现恒定的线性趋势（即线性模型 l)

2. 由于鸟类群落、栖息地类型和管理举措存在差异，该物种的趋势在生态上各异的广阔分层（鸟类保护区）之间有所不同（即分层模型）。

3. 该物种的趋势与30年（1981-2010）的气候常值相互作用（即温度模型）。如果气候塑造了物种分布，那么随着年气温日益升高，我们预计气候极端附近的物种趋势会有所不同（例如，分布范围北界呈正向趋势、分布区南界呈负向趋势，表明分布范围正向北迁移）。

4. 该物种的趋势随多种相互作用的非生物与生物变量而变化，这些变量共同导致该物种分布范围内呈现复杂的空间变化趋势（即空间变系数模型）。

我们使用了以下数据： $J = 1846$ 至少在2000年至2019年间调查过一次的繁殖鸟类调查路线（即地点；每条路线平均调查年数为 $= 15$）。繁殖鸟类调查观察员在每条路线沿线的50个停留点分别进行了为时三分钟的定点计数调查，记录半径0.4公里范围内所有看到或听到的鸟类。我们将每个地点的每个物种数据汇总为 $K = 5$ 个空间重复样本（每个样本包含50个停留点中10个点的数据）；若在某个重复样本的10个停留点中的任意一点检测到该物种，则该样本取值为1，否则取值为0。尽管以往研究曾采用这种方法处理繁殖鸟类调查数据（例如，Rushing等，2020），但在占域建模框架中使用空间重复样本很可能会违反封闭性假设（Kendall和White，2009年），因此我们将响应变量称为物种特异性出现（或“利用”），而非占域。

对于这 51 个物种中的每一个，我们都拟合了一个多季节占域模型，其中将每年 t 各地点 j 的出现概率建模为

$$\text{logit}(\psi_t(\boldsymbol{s}_j)) = \beta_0 + \text{w}_0(\boldsymbol{s}_j) + \beta_1 \cdot \text{TMAX}(\boldsymbol{s}_j) + f(\text{YEAR}_t, \boldsymbol{\beta}_{\text{TREND}}) + \eta_t, \quad (7)$$

其中 β_0 和 $\text{w}_0(\boldsymbol{s}_j)$共同表示空间变化截距， β_1 是.....的效应 t

30年（1981—2010年）最高气温气候常年值对出现概率的影响，η_t 是用于处理残差时间自相关的 AR(1)随机年份效应，而$f(\mathrm{YEAR}_t, \beta_{\mathrm{TREND}})$是估计的趋势参数，其形式在四种模型间有所不同。对于所有模型，我们将探测概率表示为年份的线性效应和二次效应、调查日的线性效应和二次效应（用于考虑探测概率的季节变化），以及调查重复的线性效应和二次效应（用于考虑一天中不同时段探测概率的变化）的函数。对于每个物种，我们采用国际鸟盟公布的既有分布范围（国际鸟盟，2021），并且在拟合占域模型时，仅纳入位于物种分布范围50公里缓冲区内的样线。利用 TerraClimate 数据（Abatzoglou 等，2018年）计算每条样线的30年平均最高温度。

我们使用 WAIC 评估模型简约性，从而比较这四个模型。我们利用 2021 年繁殖鸟类调查的实际观测值，通过受试者工作特征曲线下面积（AUC；Hosmer Jr 等人 2013）比较各模型预测（即预报）物种在 2021 年未来出现情况的能力。具体方法参照Zipkin 等人（2012）。完整详情见补充信息 S2。

我们使用贝叶斯推断和 `spOccupancy` R 软件包中实现的马尔可夫链蒙特卡罗（MCMC）算法拟合所有模型（Doser 等人，2022，2023）。简而言之，我们为模型中的所有非空间参数指定了宽松的先验分布，并为空间参数指定了弱信息先验分布，详情见补充信息 S2。每个模型运行三条链，进行 100,000 次 MCMC 迭代，其中预热期为 50,000 次迭代，抽稀率为 50，最终得到 3000 个后验样本。我们使用潜在尺度缩减因子（即$\hat{}$R；Brooks和Gelman 1998）评估收敛情况。拟合完四种备选模型后，我们针对研究区域内各物种的分布范围进行预测，绘制了二十年间的出现趋势图。

结果

对于我们纳入分析的大多数森林鸟类物种，美国东部二十年间的出现趋势存在空间变异，这一结论得到了强有力的支持；许多物种在其繁殖分布范围内同时呈现正向和负向趋势（图 2）。对于几乎所有物种（51种中的49种，即96%），SVC模型的表现都显著优于（即 $\Delta\mathrm{WAIC} > 2$）线性趋势模型、分层模型

和最高温度交互作用模型（根据WAIC判断），这表明SVC模型带来的额外灵活性提高了模型简约性。在三个备选模型中，根据WAIC，BCR模型分别在78%和90%的物种中优于温度模型和线性模型；而温度模型在这51个物种中总体支持度较低，仅在57%的物种中优于线性模型。对SVC模型支持度较低的两种物种（美洲丘鹬和东美角鸮）原始出现概率都非常低（即 < 0.05）。与较简单的分层模型相比，SVC模型需要更多观测值，才能获得不确定性合理的估计值；这可能是较简单的分层模型在这两种物种中表现更好的原因。相较于三个备选模型，SVC模型在预测性能方面的提升所获得的支持程度差异较大。总体而言，SVC模型在预测未来的物种特异性出现（即2021年）方面通常表现最佳；对于69%的物种（51种中的35种），其AUC最高（补充信息 S2表3）。

SVC模型揭示了出现趋势中的空间异质性，而另外三种较简单的替代模型未能显示出这种异质性。图3 展示了四种模型中三个示例物种（灰猫鹊（GRCA）、东菲比霸鹟（EAPH）和林鹪（WOTH））的趋势估计值。SVC模型显示，灰猫鹊在其分布范围南部（路易斯安那州和密西西比州除外）总体呈负向趋势，而在北部各州呈正向趋势或没有明确方向的趋势，这表明其分布范围可能正在向北迁移。温度模型和分层模型都能捕捉到美国东部南部区域趋势更偏负向这一总体模式，但都未能捕捉到SVC模型揭示的更精细尺度变异（即路易斯安那州和密西西比州的正向趋势，以及明尼苏达州北部的负向趋势）。东菲比霸鹟则呈现相反的模式：分层模型和温度模型都能较好地捕捉到其分布范围南部总体呈正向趋势、北部呈负向趋势的特点，但未能捕捉到SVC模型揭示的趋势中更精细尺度的变异。林鹪在其分布范围东部的趋势强烈为负向，而在西北部则更偏正向；SVC模型和分层模型在很大程度上都捕捉到了这一模式。相比之下，由于最高温度与林鹪趋势之间不存在关联，温度模型所呈现的趋势基本没有空间变异

。这清楚地说明，如果交互变量不能充分解释趋势中的空间变异，交互模型就缺乏灵活性。

温度模型揭示了最高温度与年度趋势之间交互作用的方向和显著性存在较大异质性（补充信息 S2：图 1），这几乎无法支持气候生态位位置是东部森林鸟类出现趋势空间变异的稳定驱动因素。在总计 51 个物种中，18 个物种存在显著的负向交互作用（即在温度较高的区域，趋势的正向程度较低／负向程度较高），而 8 个物种存在显著的正向交互作用（即在温度较高的区域，趋势的正向程度较高／负向程度较低）。尽管这些显著趋势表明，趋势中的某些空间变异可能与一个分布范围内的气候位置有关，但与温度模型相比，空间变系数模型的模型拟合度（以及在较小程度上的预测）大幅提升，这表明这种关系并非东部森林鸟类出现趋势空间变异的主要驱动因素。同样，与恒定线性趋势模型相比，分层模型性能有所提升，这表明趋势的空间变异可能部分归因于不同鸟类保护区之间栖息地和气候条件的广域尺度差异。然而，与分层模型相比，空间变系数模型性能进一步提升，这表明在更精细空间尺度上起作用的其他因素，也是出现趋势空间变异的重要成因。

案例研究 2：土地覆盖变化对草地鹨的影响

量化土地覆盖变化及其造成的栖息地可利用性变化对物种分布的影响，对于理解导致大规模鸟类种群衰退的主要驱动因素至关重要（Rosenberg 等人，2019年）。在北美所有鸟类类群中，草原鸟类的种群衰退幅度最为显著，这主要被认为是栖息地丧失、农业集约化以及农业景观中农药使用增加所致（Stanton 等人，2018年）。然而，土地覆盖变化对草原鸟类物种的影响不太可能在空间上保持恒定，因为它会与局地气候、农田管理措施（例如农药使用）、捕食压力及其他胁迫因素产生复杂的交互作用。在本案例研究中，我们展示了空间变系数模型能够提供

关于栖息地变化对草地鹨（*Ammodramus savannarum*）在美国本土分布产生的空间变化效应的见解，研究范围为1970—2019年。我们再次使用北美繁殖鸟类调查（Pardieck 等人，2020）的数据，这些数据是在50年间于草地鹨分布范围内的 $J = 2542$ 条调查路线收集的（分布范围依据国际鸟盟（国际鸟盟，2021）的数据确定，方法如案例研究1所述）。本案例研究的目标是：（1）绘制空间明确地图，展示草地面积和农田面积变化对草地鹨出现情况的影响；（2）比较一系列空间变系数模型和更简单的替代模型，以提出并检验有关栖息地变化效应空间变异驱动因素的假设。

我们采用与案例研究1相同的方法，将繁殖鸟类调查数据汇总为重复的探测-未探测数据，使用 $K = 5$ 每条繁殖鸟类调查路线上的空间重复样本，对麻雀分布范围内各路线层级的出现情况进行建模。我们利用美国地质调查局地球资源观测与科学中心（USGS EROS）的数据，计算每条繁殖鸟类调查路线周围1公里范围内每年的草地面积（天然草地覆盖类别）和农田面积（农田和干草地/牧场覆盖类别的合计）（Sohl 等人，2016）。我们通过从每个地点的各年度数值中减去其50年平均值，计算草地和农田面积的年度偏差，以明确评估栖息地随时间变化的影响，并将其与栖息地可利用性的空间变异区分开来（Clement 等人，2019年；Saunders 等人，2022年）。

我们假设，草地鹨的出现与草地面积变化呈正相关，并且由于以下原因，这种影响会随空间而变化：(1) 鉴于草地鹨偏好由大量连片草地组成的景观，其与草地平均面积之间存在正向交互作用（Shaffer 等人，2021）；(2) 与温度存在交互作用，即在该物种分布范围的气候极端附近（即分布范围边界），草地面积变化的影响最强；以及 (3) 与精细尺度栖息地质量（例如草地高度、枯落物数量；Shaffer 等人 2021）和管理措施存在额外交互作用，而这些因素无法作为模型的协变量。我们预计，在大平原地区，草地鹨的出现与农田覆盖变化呈负相关，因为该地区农田覆盖增加很可能对应着草地面积减少（草地是该地区的主要土地覆盖类型；补充信息S3：图1）。然而，在草地面积很少的地区（即西经100度线以东），我们预测草地鹨的出现与

农田面积呈正相关，因为此前研究发现，当原生草地面积有限时，草地鹨的出现与干草地和耕地有关（Shaffer 等人，2021）。

我们拟合了五个候选模型，这些模型在草地面积变化和农田面积变化效应的函数形式上有所不同，以检验我们的假设（完整细节见补充信息 S3）。具体而言，五个模型分别为：(1) 线性模型，其中草地变化和农田变化具有恒定的线性效应；(2) 生境交互模型，其中草地变化和农田变化具有线性效应，草地变化与50年平均草地面积之间存在交互作用，农田变化与50年平均农田面积之间也存在交互作用；(3) 温度交互模型，其中草地变化和农田变化均具有线性效应，且二者均与平均温度条件存在交互作用（即如案例研究1所述，采用 TerraClimate 计算的30年最高气温）；(4) 空间变系数模型；(5) “完整”模型，包含空间变系数、温度的交互效应，以及50年平均土地覆盖面积的交互效应。

对于五个候选模型中的每一个，我们都使用以下方法拟合了贝叶斯多季节占域模型：`spOccupancy` R 软件包（Doser 等人，2022年）。在五个模型中，探测概率均以线性和二次型序数日期、年份随机效应，以及调查重复的单独截距为协变量进行建模，以解释同一繁殖鸟类调查路线内五个空间重复样本之间探测概率的差异。鉴于我们重点推断物种-环境关系，我们使用 WAIC 比较候选模型。先验分布采用非信息性或弱信息性设定（补充信息 S3）。每个模型均运行三条链，进行100,000次 MCMC迭代，其中预热期为50,000次迭代，抽稀率为50，最终得到3000个后验样本。我们使用潜在尺度缩减因子评估收敛情况，并通过目视检查以下软件包生成的轨迹图进行评估：`coda`软件包（Plummer 等人，2006）。

结果

我们发现，草地和农田覆盖变化的影响存在空间变异，且有强有力的证据支持这一点：所有纳入空间变系数和／或交互作用的模型，其表现都显著优于影响为常数的模型（即，$\Delta\text{WAIC} > 2$）（补充信息 S3：表1）。温度交互模型优于生境交互模型（$\Delta\text{WAIC} = 18.65$），这表明最高温度对于解释以下方面的空间变异更为重要：

栖息地变化的影响，而非栖息地数量。值得注意的是，与常数模型相比，纳入草地和农田变化影响的空间变系数，大幅降低了 WAIC（即， ΔWAIC $= 774.90$）；其降幅显著大于任一交互模型的降幅（ΔWAIC $= 20.69$对应生境交互模型， ΔWAIC $= 39.34$ 对应温度交互模型）。同时纳入空间变系数、生境交互作用和温度交互作用的模型，表现略优于仅纳入空间变系数的模型（ΔWAIC $= 3.32$）。总之，这些结果表明，与温度和栖息地的交互作用可以解释栖息地变化对草地鸟类出现的影响所存在的部分空间变异，但这些影响的大部分变异源于未能解释的空间变异，而空间变系数考虑了这种变异。对不同候选模型预测的草地和农田变化影响绘制成图后发现，与不含空间变系数的模型相比，含有空间变系数的模型能捕捉到更多栖息地变化影响的空间变异（图4）。

候选模型估计的效应对我们的假设提供了不一的支持。表现最佳的模型显示，在草地鹨的繁殖分布范围内，栖息地变化既有正向效应，也有负向效应。正如预测的那样，在北部大平原（草地面积较大的地区），草地变化的效应强烈为正，这表明在以草地为主的景观中，草地面积减少会导致草地鹨出现概率下降。生境交互模型也支持这一点，该模型显示草地变化与平均草地面积之间存在正向交互作用（图 4A）。出乎意料的是，我们发现，在南部大平原，草地面积变化的效应接近于零或为负，这表明该地区草地增加不会影响草地鹨出现概率，甚至可能导致其下降。鉴于该地区的草地和牧场面积相对较大（补充信息 S3 图 1），这可能表明草地面积对草地鹨出现概率的最佳值在该区域达到峰值，这与以往研究结果一致，即草地鹨偏好的草地面积因生态区而异（Johnson andIgl，2001年）。温度交互模型显示，这一模式部分与最高温度的负向交互作用有关（图 4B；Gorzo 等，2016）。在草地鹨分布区南界的大部分地带，其出现概率与农田面积呈负相关，这部分归因于与最高温度之间的负向交互作用（图4）。另一方面，农田变化的效应总体上

在美国东北部和美国中西部均为正向（图 4），这部分与农田变化和农田面积之间的正向交互作用有关（图 4D）。在美国东北部和中西部各州，原生草地所剩无几，农田面积变化的正向效应表明，农田面积增加会提高草地鹨出现概率。此外，美国东北部以森林为主，因此农田覆盖增加可能与森林覆盖减少有关，这或许可以部分解释农田覆盖变化的正向效应，因为草地鹨完全避开森林（Grant et al.，2004）。尽管这类评估带有推测性，但正是空间变系数模型揭示物种与环境之间精细、多层面关系的能力，才使这些洞见成为可能；进而，这些洞见可用于指导地方和区域层面的管理与保护优先事项。

讨论

随着生态学研究的空间和时间范围不断扩大，考虑复杂的空间变化和/或非线性物种—环境关系变得愈发重要（Rollinson 等人，2021）。交互作用、分层以及非线性模型等常用统计方法可以部分解释此类模式，但在估计由多个相互作用且自身也随空间变化的因素所产生的空间变化物种—环境关系时，其灵活性有限；尤其是在这些因素无法作为协变量获取的情况下。本文探讨了在物种分布模型中使用空间变系数（SVC），这种方法能够在层次广义线性模型框架内有效地建模非线性和/或空间变化的物种—环境关系，同时处理野生动物和植物数据集中常见的观测偏差（例如，Chen 等，2013）。通过模拟以及对美国鸟类开展的两项案例研究，我们强调了空间变系数物种分布模型的推断优势：它能够生成并检验生态学假设，探究驱动估计关系和/或出现趋势空间变异的因素。

我们的模拟研究表明，空间变系数物种分布模型能够准确捕捉不同形式的复杂空间变异物种—环境关系；如果数据不支持这种额外的复杂性，它们也能揭示更简单的物种—环境关系（例如线性、二次型；图1）。因此，当人们对

模型拟合前对物种—环境关系的形式知之甚少时，可以使用空间变系数物种分布模型来生成关于关系真实形式以及哪些非生物和/或生态因素影响该关系的假设。当真实的物种—环境关系先验已知时，鉴于空间变系数模型的复杂性更高，根据简约原则建立的信息准则，较简单的参数化广义线性模型可能优于空间变系数物种分布模型。在这种情况下，我们的模拟研究表明，空间变系数物种分布模型不会得出错误结论，而是会揭示真实关系较简单的参数化形式。在宏观尺度上开展研究时，模型拟合前不太可能已知和/或获得所有交互变量；在这种情况下，空间变系数物种分布模型将优于较简单的替代模型（第 5、6 行；图 1），并有助于揭示这些模式背后的生态驱动因素。

如两个实证案例研究所示，空间变系数物种分布模型的一项关键优势在于，能够检验并进一步提出假设，探究物种—环境关系和出现趋势中空间变异的驱动因素。在评估物种—环境关系和/或趋势的空间变异时，我们建议将空间变系数物种分布模型与代表明确假设的简单参数化物种分布模型进行比较，因为这类比较有助于揭示不同空间变化效应/趋势驱动因素获得支持的程度（Pease 等人，2022b；Sultaire 等人，2022年）。例如，在东部森林鸟类案例研究中，温度模型显示，18 个物种的出现趋势与繁殖季最高气温之间存在显著的负向交互作用，另有 8 个物种存在显著的正向交互作用（补充信息 S3 图 1）。然而，对于所有 26 个温度交互作用显著的物种，SVC模型都是表现最佳的模型。这表明，尽管繁殖季温度通常能解释出现趋势的一部分变异，但还有其他因素（例如栖息地质量变化、气候变化）对解释出现趋势的精细尺度变异也很重要。

空间变系数模型是否比仅包含空间变化截距的模型具有更好的预测性能，仍是一个悬而未决的统计学问题。在我们的东部森林鸟类案例研究中，与较简单的模型相比，空间变系数模型在预测2021年的出现概率时，预测性能提升相对有限。四个模型都纳入了空间随机效应，以解释出现概率的空间变异；鉴于从2019年（建模数据集的最后一年）到2021年森林鸟类分布可能发生的变化较小，它们预测整个研究区域出现概率的能力相近。

在其他生态学和自然资源应用中，与仅包含空间变化截距的模型相比，空间变系数模型在预测方面的优势表现不一：一些研究发现空间变系数模型的预测性能有所提高（Sultaire 等人，2022年；May，2023），而另一些研究则表明，性能提升取决于物种（Pease 等人，2022a；Doser 等人，2023）或区域（Babcock 等人，2015年）。无论如何，我们认同Thorson 等人（2023）的观点，即空间变系数物种分布模型的主要优势在于，它们能够更好地检验和提出假设，并回答与物种—环境关系及其趋势的空间变异相关的生态学问题。

除理论贡献外，空间变系数物种分布模型的结果（例如，物种分布范围内物种—环境关系和趋势的估计值）还可用于多尺度保护和管理决策。例如，在草地鹨案例研究中，我们发现草地面积减少最有可能对北部大平原造成最大影响（即导致出现概率下降），这凸显了提供大片连片天然草原的重要性，以防止该地区该物种进一步衰退（Shaffer 等人，2021）。对多种草原鸟类物种开展类似分析后，可将空间变系数物种分布模型的输出共同用作空间优先排序分析的输入。这有望显著改进保护区规划，并帮助资源管理者确定栖息地恢复最有利于整体鸟类群落的具体地点（Grand 等人，2019）。此外，物种趋势的估计值可作为确定最需保护物种保护状况的基础。空间变系数物种分布模型所揭示的空间变异趋势（例如东部森林鸟类案例研究中得出的趋势），可用于局部（即州级）、区域（即生态区）和大陆尺度的评估（Smith 等人，2023年）。最终，这有助于我们更好地理解不同地区的保护行动为何会产生不同结果。为管理者提供多尺度出现趋势，有助于制定因地制宜的行动计划，从而更有效地恢复受关注的保护物种。利用空间变系数物种分布模型估计局部趋势，还可改进对物种分布变化的预测（Barnett 等人，2021），包括预测其对入侵物种（Thorson 等人，2023年）以及未来气候和/或土地利用变化（Gonthier 等人，2014年）的响应。

虽然生态学中常用其他统计方法（如广义加性模型）和机器学习方法（如随机森林、MaxEnt）

来处理复杂的物种—环境关系，但本文所述的贝叶斯空间变系数模型是一种颇具吸引力的替代方案，因为它们：(1) 无需先验了解哪些变量存在交互作用；(2) 能够轻松提供所有估计值的不确定性度量；(3) 易于嵌入用于处理生态学数据中普遍存在的观测偏差的分层建模框架（即占域模型）。尽管如此，贝叶斯空间变系数物种分布模型所具备的灵活性，可能会给模型的实现带来计算和实际操作方面的困难。贝叶斯框架能够将不确定性完整地传播至所有估计值和预测结果，但模型运行可能需要相当长的时间。例如，在草地鹨案例研究中，完整的 SVC模型所用数据集包含近 400,000 个观测值，使用 `spOccupancy` (Doser 等人，2022年) 运行一条包含 100,000 个样本的马尔可夫链蒙特卡洛链，耗时约 10 小时。 `R` 软件包 `VAST` (Thorson，2019年) 和 `sdmTMB`(Anderson 等人，2022年) 提供了用于拟合空间变系数物种分布模型的最大似然（即频率学派）替代方法。它们的运行速度快得多，但这些模型并未明确考虑不完全探测。此外，与较简单的替代方法（即分层、交互作用）相比，空间变系数物种分布模型估计复杂的空间变化物种—环境关系，可能需要较大的样本量才能达到合理的不确定性水平。使用探测-未探测数据时，这一点尤为突出：与许多物种分布模型使用的计数（如丰度）或连续型（如生物量）数据源相比，这类数据所能提供的用于估计空间变系数的信息相对较少。在补充信息 S4 中，我们提供了更多模拟研究，以深入了解空间变系数估计值的可靠性如何随数据集中空间位置的数量和季节数而变化。样本量有限时，分层或交互作用等较简单的方法可能更有用，能够在不确定性不过高的情况下估计物种—环境关系。最后，估计的空间变化截距与空间变化系数之间可能存在混杂，尤其是在处理规模适中的数据集（如包含 500 个数据点的数据集）时，这可能导致误导性结论。不过，在估计随时间变化的协变量（如时间趋势，或如两个案例研究中所示的栖息地变化）的空间变系数时，由于多季节数据额外包含时间成分，混杂程度会降低。我们发现，近期关于如何最大限度减少空间混杂并理解其在空间显式物种分布模型中影响的指南，同样适用于空间变系数物种分布模型（Mäkinen 等人，2022年），不过仍需进一步研究，以了解此类混杂何时可能

发生，并如何最好地减轻其影响。

物种—环境关系中的空间变异在整个生态学领域普遍存在（Rollinso n等，2021年），这是与非生物与生物变量复杂交互作用的结果，而在统计分析前，这些变量很少全部已知或可供测量。正如本研究所示，在物种分布模型中使用空间变化系数，有助于阐明驱动物种分布动态在局部和广阔空间尺度上变化的环境因素。这有助于更有效地检验生态学假设，并为多尺度保护和管理举措提供依据。在宏观尺度上拟合物种分布模型时，我们鼓励将空间变系数物种分布模型与更简单的替代方案进行比较，以增进我们对物种—环境关系在空间上的驱动因素的理解。

参考文献

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.（2018年）。TerraClimate：一个高分辨率全球数据集，包含1958–2015年逐月气候和气候水量平衡。*Scientific Data*，5(1)：1–12。

Amburgey, S. M., Miller, D. A., Campbell Grant, E. H., Rittenhouse, T. A., Benard, M. F., Richardson, J. L., Urban, M. C., Hughson, W., Brand, A. B., Davis, C. J., 等（2018年）。分布范围位置与气候敏感性：种群间人口统计响应的结构及其对气候变化的响应。*Global Change Biology*，24(1):439–454。

Anderson, S. C., Ward, E. J., English, P. A., 和 Barnett, L. A.（2022年）。sdmTMB：一个快速、灵活且易于使用的 R 软件包，用于拟合包含空间和时空随机场的广义线性混合效应模型。*bioRxiv*。

Babcock, C., Finley, A. O., Bradford, J. B., Kolka, R., Birdsey, R., 和 Ryan, M. G.（2015年）。基于激光雷达，利用具有空间变系数的分层模型预测森林生物量。*Remote Sensing of Environment*，169:113–127。

Babcock, C., Finley, A. O., Cook, B. D., Weiskittel, A., and Woodall, C. W.（2016）. 森林生物量与生长建模：将长期清查与激光雷达数据相结合。*Remote Sensing ofEnvironment*，182：1–12。

Banerjee, S., Carlin, B. P., and Gelfand, A. E.（2014年）。空间数据的层次建模与分析。Chapman and Hall/CRC。

Barnett, L. A., Ward, E. J., and Anderson, S. C. (2021). 通过纳入局部趋势改进物种分布变化的估计值。《生态地理学》，44(3):427–439.

Bateman, B. L., Wilsey, C., Taylor, L., Wu, J., LeBaron, G.S., andLangham, G.(2020). 北美鸟类需要采取减缓和适应措施，以降低其对气候变化的脆弱性。*Conservation Science and Practice*，2(8):e242.

国际鸟盟（2021）。《2021年世界鸟类手册》。世界鸟类物种分布图，2021版。http://datazone.birdlife.org/species/requestdis。

Bled, F., Sauer, J., Pardieck, K., Doherty, P., and Royle, J. A. (2013). 基于北美繁殖鸟类调查数据的趋势建模：一种空间显式方法。*PLoS One*，8(12):e81867.

Brooks, S. P. 和 Gelman, A.（1998）。监测迭代模拟收敛的一般方法。*Journal of Computational and Graphical Statistics*，7(4):434–455。

Chen, G., Kéry, M., Plattner, M., Ma, K. 和 Gardner, B.（2013）。在植物分布研究中，不完全探测是常态而非例外。*Journal of Ecology*，101(1):183– 191。

Clark, T. 和 Luis, A. D.（2020）。非线性种群动态在动物中普遍存在。*Nature Ecology & Evolution*，4(1):75–81。

Clement, M. J., Nichols, J. D., Collazo, J. A., Terando, A. J., Hines, J. E. 和 Williams, S. G.（2019）。剖析全球变化：评估气候和土地覆盖变化对鸟类分布变化的相对重要性。*Ecology and Evolution*，9(4):1985–2003。

Crossley, M. S., Smith, O. M., Davis, T. S., Eigenbrode, S. D., Hartman, G. L., Lagos-Kutz, D., Halbert, S. E., Voegtlin, D. J., Moran, M. D. 和 Snyder, W. E.（2021）。复杂的生活史使蚜虫容易遭遇近期丰度下降。*Global Change Biology*, 27(18):4283–4293。

Datta, A., Banerjee, S., Finley, A. O. 和 Gelfand, A. E.（2016）。用于大型地统计数据集的层次最近邻高斯过程模型。*Journal of the American StatisticalAssociation*，111(514)：800–812。

Doser, J. W., Finley, A. O., Kéry, M. 和 Zipkin, E. F.（2022年）。spOccupancy：用于单物种、多物种及综合空间占域模型的 R 软件包。《生态学与进化方法》，13(8)：1670–1678。

Doser, J. W., Finley, A. O., Saunders, S. P., Kéry, M., Weed, A. S. 和 Zipkin, E. F.（2023年）。利用空间变化系数占域模型刻画复杂的物种—环境关系。*arXiv* 预印本。

Ethier, D. M., Koper, N. 和 Nudds, T. D.（2017年）。驱动受威胁草原鸟类区域尺度种群动态的机制中的时空变异。*Ecology and Evolution*，7(12)：4152–4162。

Golding, N. 和 Purse, B. V.（2016）。利用高斯过程快速、灵活地进行贝叶斯物种分布建模。《生态学与进化方法》，7(5)：598–608。

Gonthier, D. J., Ennis, K. K., Farinas, S., Hsieh, H.-Y., Iverson, A. L., Batáry, P., Rudolphi, J., Tscharntke, T., Cardinale, B. J. 和 Perfecto, I.（2014年）。农业中的生物多样性保护需要多尺度方法。《英国皇家学会学报*B*辑：生物科学》，281(1791)：20141358。

Gorzo, J. M., Pidgeon, A. M., Thogmartin, W. E., Allstadt, A. J., Radeloff, V. C., Heglund, P. J., 和 Vavrus, S. J.（2016）。利用北美繁殖鸟类调查评估大陆上濒危程度最高的鸟类群落——草原鸟类——对天气变异的广泛尺度响应。《秃鹰：鸟类学应用》，118(3):502–512。

Grand, J., Wilsey, C., Wu, J. X., 和 Michel, N. L.（2019年）。北美草原鸟类的未来：将持续性威胁和新出现的威胁纳入全年周期保护优先事项。*Conservation Science and Practice*, 1(4):e20。

Grant, T. A., Madden, E., 和 Berkey, G. B.（2004年）。北部混合草原中的乔木和灌木入侵：对繁殖期草原鸟类的影响。《野生动物学会通报》，32(3):807–818。

Guisan, A., Edwards Jr, T. C., 和 Hastie, T.（2002年）。研究物种分布时使用广义线性模型和广义加性模型：背景介绍。《生态建模》，157(2-3)：89–100。

Guisan, A., Thuiller, W., 和 Zimmermann, N. E.（2017年）。栖息地适宜性与分布模型：及其在 *R* 中的应用。剑桥大学出版社。

Guisan, A. 和 Zimmermann, N. E.（2000年）。生态学中的预测性栖息地分布模型。《生态建模》，135(2-3):147–186。

Hosmer Jr, D. W.、Lemeshow, S. 和 Sturdivant, R. X.（2013）。应用逻辑回归，第398卷。约翰·威立父子出版公司。

Hostetler, J. A. 和 Chandler, R. B.（2015年）。改进的状态空间模型：利用计数数据推断丰度的空间和时间变异。生态学，96(6):1713–1723。

Johnson, D. H. 和 Igl, L. D.（2001年）。草原鸟类的面积需求：区域视角。《海雀》，118(1):24–34。

Kellner, K. F. 和 Swihart, R. K.（2014年）。在生态学中考虑不完全探测：定量综述。*PLoS One*，9(10):e111436。

Kendall, W. L. and White, G. C.（2009年）。关于在地点占据研究中以空间子单元替代重复时间采样的警示。《应用生态学杂志》，46(6):1182–1188。

Kéry, M.（2011年）。迈向真实物种分布的建模。《生物地理学杂志》，38(4):617–618。

Liaw, A., Wiener, M., et al.（2002年）。使用 randomForest 进行分类与回归。《*R*新闻》，2(3):18–22。

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., and Langtimm, C. A.（2002年）。当探测概率低于1时估计地点占据率。生态学，83(8):2248–2255。

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., and Hines, J. E.（2017年）。占据估计与建模：推断物种出现的模式与动态。爱思唯尔。

Mäkinen, J., Numminen, E., Niittynen, P., Luoto, M., 和 Vanhatalo, J.（2022年）。贝叶斯物种分布建模中的空间混杂。《生态地理学》，2022年(11)：e06183。

May, P., McConville, K. S., Moisen, G. G., Bruening, J., 和 Dubayah, R.（2023年）。用于估计美国本土连续地区生物量密度小区域估计值的空间变化模型。*Remote Sensing of Environment*，286：113420。

Meehan, T. D., Michel, N. L., and Rue, H.（2019）。奥杜邦圣诞节鸟类调查的空间建模揭示了相对丰度趋势的精细尺度模式及其驱动因素。《生态圈》，10(4):e02707。

Miller, J. A.（2012）。物种分布模型：空间自相关与非平稳性。《自然地理学进展》，36(5):681–692。

Osborne, P. E. and Suárez-Seoane, S.（2002）。构建大尺度分布模型之前，是否应对数据进行空间分区？《生态建模》，157(2-3):249–259。

Pardieck, K., Ziolkowski Jr, D., Lutmerding, M., Aponte, V., and Hudson, M.-A.（2020）。北美繁殖鸟类调查数据集 1966–2019年。美国地质调查局数据发布，*https://doi.org/10.5066/P9J6QUF6*。

Pease, B. S., Pacifici, K., and Kays, R.(2022a). 探究四种哺乳动物的空间非平稳性，揭示环境关系的区域差异。《生态圈》, 13(8):e4166.

Pease, B. S., Pacifici, K., Kays, R., and Reich, B. (2022b). 是什么驱动了广布物种中随空间变化的生态关系？《多样性与分布》.

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). 物种地理分布的最大熵建模。《生态建模》, 190(3-4):231–259.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA：用于马尔可夫链蒙特卡罗的收敛诊断与输出分析。《*R*新闻》, 6(1):7–11.

Reich, B. J., Pacifici, K., and Stallings, J. W. (2018年). 将辅助数据纳入物种分布建模的最优空间设计。《生态学与进化方法》, 9(6):1626–1637.

Rollinson, C. R., Finley, A. O., Alexander, M. R., Banerjee, S., Dixon Hamil, K.-A., Koenig, L. E., Locke, D. H., DeMarche, M. L., Tingley, M. W., Wheeler, K., 等（2021）。跨越

时空：生态学研究与应用中的非平稳性。《生态学前沿与环境》，19(1)：66–72。

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., 等（2019年）。北美鸟类的衰退。《科学》，366(6461):120–124。

Rushing, C. S., Royle, J. A., Ziolkowski Jr, D. J., 和 Pardieck, K. L.（2020）。迁徙行为和越冬地理分布导致东部鸟类对近期气候变化的分布范围变化各不相同。《美国国家科学院院刊》，117(23):12897– 12903。

Saunders, S. P., Meehan, T. D., Michel, N. L., Bateman, B. L., DeLuca, W., Deppe, J. L., Grand, J., LeBaron, G. S., Taylor, L., Westerkam, H., 等（2022年）。揭示一个世纪以来全球变化对美国东部鸟类越冬分布的影响。*GlobalChange Biology*，28:2221–2235。

Shaffer, J. A., Igl, L. D., Johnson, D. H., Sondreal, M. L., Goldade, C. M., Nenneman, M. P., Wooten, T. L., 和 Euliss, B. R. (2021). 管理措施对草原鸟类的影响——草地鹨（Ammodramus savannarum）。美国地质调查局专业论文。

Smith, A. C., Binley, A., Daly, L., Edwards, B. P., Ethier, D., Frei, B., Iles, D., Meehan, T. D., Michel, N. L., 和 Smith, P. A. (2023年)。用于估计鸟类种群状况与趋势的空间显式贝叶斯层级模型。*EcoEvoRxiv*。

Smith, A. C. 和 Edwards, B. P. (2021)。北美繁殖鸟类调查的状况与趋势估计值，可用于满足广泛的保护需求；该调查采用灵活的贝叶斯层级广义加性模型。《秃鹰》，123(1):duaa065。

Sohl, T., Reker, R., Bouchard, M., Sayler, K., Dornbierer, J., Wika, S., Quenzer, R., and Friesz, A. (2016). 美国本土48州的历史土地利用与土地覆盖模拟结果。《土地利用科学杂志》，11(4):476–499.

Spake, R., Bowler, D. E., Callaghan, C. T., Blowes, S. A., Doncaster, C. P., Antao, L. H., Nakagawa, S., McElreath, R., and Chase, J. M.（2023年）。理解生态学中的“视情况而定”：关于如何提出假设、可视化和解读统计交互作用的指南。

学：关于如何提出假设、可视化和解读统计交互作用的指南。《生物学评论》。

Stanton, R., Morrissey, C. A., and Clark, R. G.（2018年）。北美农田鸟类衰退趋势及农业驱动因素分析：综述。《农业、生态系统与环境》，254:244–254。

Stein, M. L.（2014年）。空间数据协方差矩阵低秩近似的局限性。《空间统计学》，8:1–19。

Sultaire, S. M., Humphreys, J. M., Zuckerberg, B., Pauli, J. N., and Roloff, G. J.（2022年）。适应雪地的物种在不连续的分布区南界上的生物气候关系空间变异。《生物地理学杂志》，49(1):66–78。

Thorson, J. T.（2019年）。使用时空向量自回归（VAST）软件包进行种群、生态系统、栖息地和气候评估的决策指南。《渔业研究》，210:143–161。

Thorson, J. T., Barnes, C. L., Friedman, S. T., Morano, J. L., and Siple, M. C.(2023年). 空间变系数能够提高物种分布模型的简约性和描述能力。《生态地理学》, 第 e06510 页。

Tyre, A. J., Tenhumberg, B., Field, S. A., Niejalke, D., Parris, K., and Possingham, H. P.(2003年). 改进生物调查的精确度并减少偏差：估计假阴性错误率。《生态学应用》，13(6):1790–1801。

Valavi, R., Elith, J., Lahoz-Monfort, J. J., and Guillera-Arroita, G. (2021). 使用随机森林对物种仅出现数据进行建模。《生态地理学》, 44(12):1731–1742。

von Hirschheydt, G., Stofer, S., and Kéry, M. (2023年). “混合”占域设计：何时增加单次访问数据能够改进标准多次访问模型的推断？《基础与应用生态学》, 67:61–69。

Watanabe, S.（2010年）。奇异学习理论中贝叶斯交叉验证与广泛适用信息准则的渐近等价性。《机器学习研究杂志》，11(12)。

Wood, S. N.（2006）。广义加性模型：*R*语言入门。Chapman and Hall/CRC。

Zipkin, E. F., Grant, E. H. C. 和 Fagan, W. F. (2012). 在考虑不完全探测的同时，使用 AUC 评估群落占域模型的预测能力。《生态学应用》，22(7):1962–1972。

图



图 1：不同物种-环境关系模式（各行）下，不同模型（各列）对模拟物种-环境关系的估计值。红色表示负向效应，白色表示无效应，蓝色表示正向效应，颜色越深表示效应越强。模拟协变量从网格底部的负值逐渐变为顶部的正值。真实的物种-环境关系模拟为线性效应（第 1 行）、二次效应（第 2 行）、跨九个分层的独立线性效应（第 3 行）、与沿水平轴变化的第二个协变量的交互作用（第 4 行）、与未知（“缺失”）协变量的交互作用（第 5 行），以及上述所有成分之和（“完全”效应，第 6 行）。图中显示了五种候选模型相对于真实关系的估计值（第 1 列），包括含有以下效应的模型：线性效应（第 2 列）、二次效应（第 3 列）、层特异性效应（第 4 列）、交互作用（第 5 列）和空间变系数（第 6 列）。

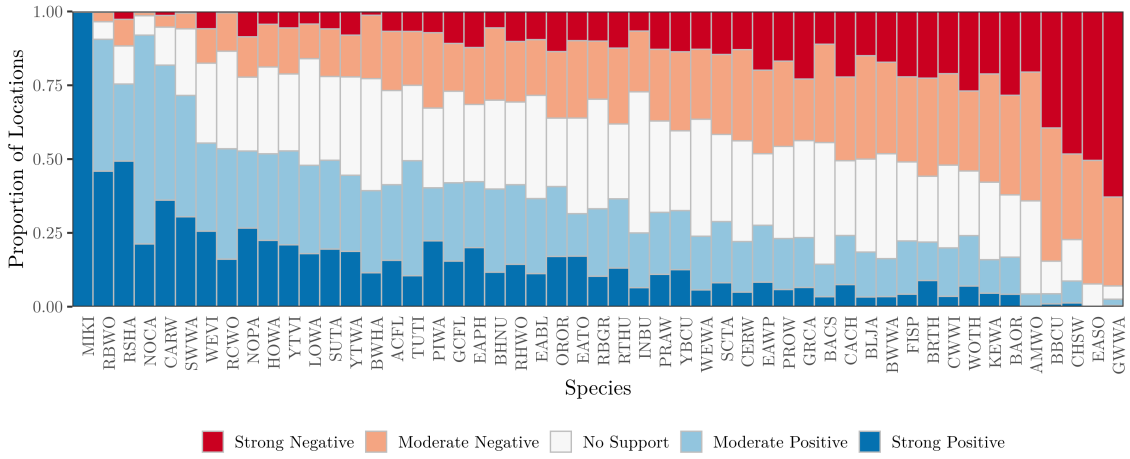


图 2：51 种森林鸟类物种空间变化趋势的汇总。每个柱子的高度对应于指定物种的地点中，其趋势具有颜色所示符号（即正向、负向、无效应）和强度（即强、中等）的地点比例。蓝色表示支持正向趋势，红色表示支持负向趋势。具体而言：(1) 深蓝色 $=$ 强正向：$(\mathrm{P(trend} > 0) > 0.8)$；(2) 浅蓝色 $=$ 中等正向：$(0.6 < \mathrm{P(trend} > 0) \leq 0.8$；(3) 白色 $=$ 无效应：$(0.4 < \mathrm{P(trend} > 0) \leq 0.6$；(4) 浅红色 $=$ 中等负向：$(0.2 < \mathrm{P(trend} > 0) \leq 0.4$；(5) 深红色 $=$ 强负向：$(\mathrm{P(trend} > 0) < 0.2)$。

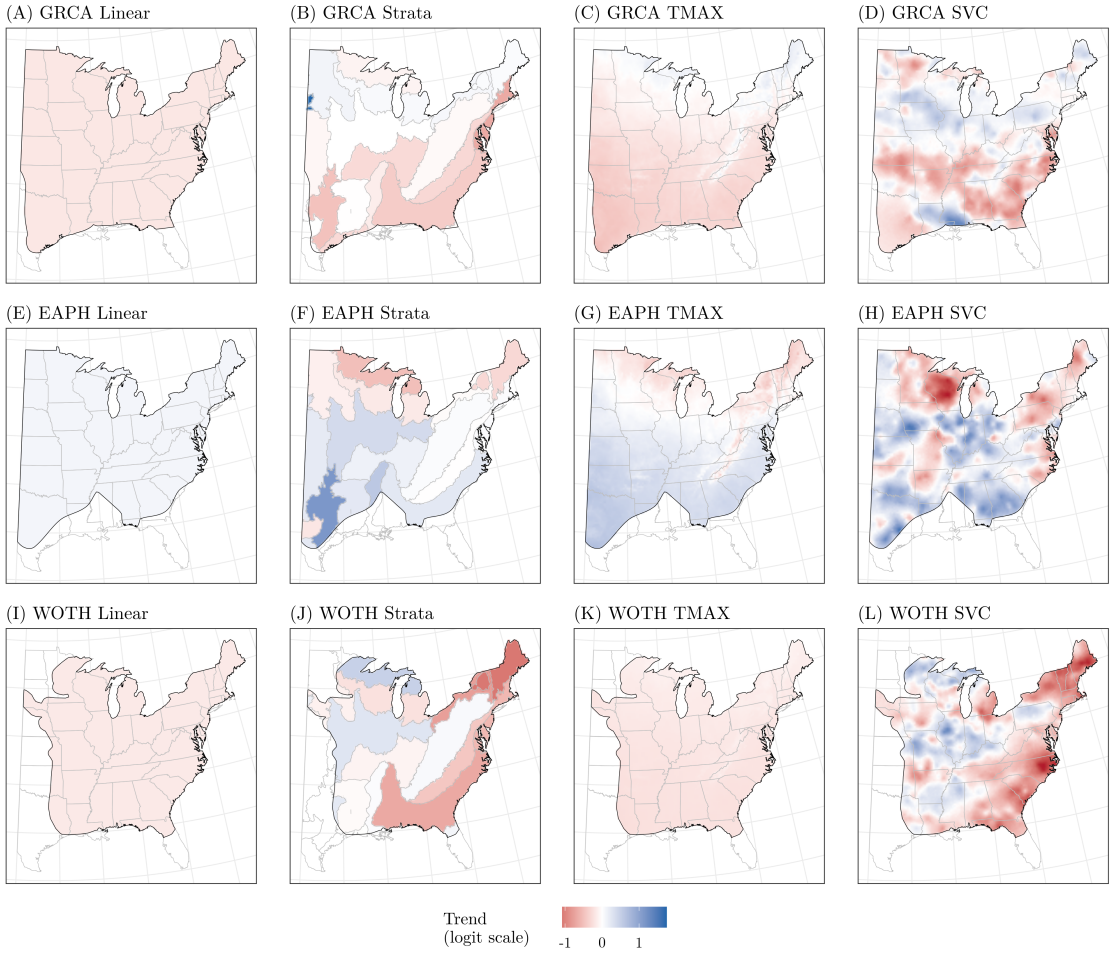


图3：四种模型对2000—2019年三种示例物种出现趋势的中位数预测：在整个物种分布范围内采用恒定线性趋势的空间占据模型（线性）、针对每个鸟类保护区分别设定趋势的空间占据模型（分层）、趋势与30年平均最高气温（TMAX）交互的空间占据模型，以及用于估计空间变化趋势的空间变系数占域模型（SVC）。A-D面板：灰猫鹊（379万平方公里^2）；E-H面板：东菲比霸鹟（355万平方公里^2）；I-L面板：林鸫（312万平方公里^2）。

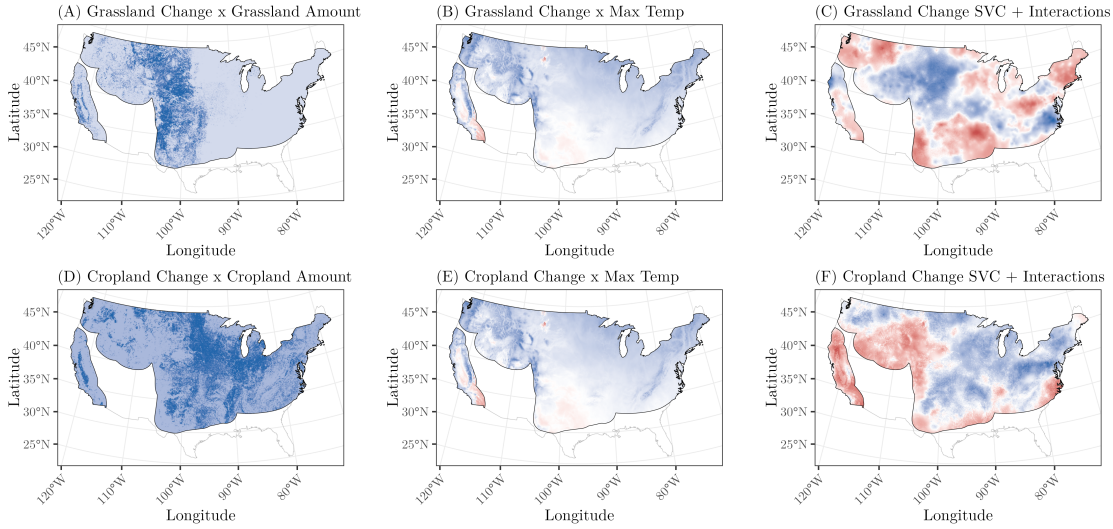


图 4：三个候选模型的中位数预测，显示草地变化（上排）和农田变化（下排）对草地鹨（*Ammodramus savannarum*）出现概率的影响。面板 (A) 和 (D) 显示了一个模型的估计值，该模型包含土地覆盖变化与50年期间平均土地覆盖面积之间的交互作用。面板 (B) 和 (E) 显示了一个模型的估计值，该模型包含土地覆盖变化与30年平均最高气温之间的交互作用。面板 (C) 和 (F) 显示了一个模型的估计值，该模型为土地覆盖变化设置了空间变化系数，并包含其与平均土地覆盖面积和最高温度之间的交互作用。蓝色表示正向影响，白色表示无效应，红色表示负向影响。