识别区域层面生物多样性数据中的知识缺口及其决定因素

Anaïs Guery[a], Didier Alard[a,b]

a. 波尔多大学，US FAUNA，佩萨克，法国

b. 波尔多大学，法国国家农业、食品与环境研究院，UMR BIOGECO，佩萨克，法国

致谢

我们感谢FAUNA观测站数据库的贡献者（机构和个人），本项目使用了他们提供的数据。我们还感谢FAUNA观测站团队，尤其是Paul Fromage，感谢他们协助进行数据提取和处理。

利益声明

本文作者声明不存在利益冲突。

通讯作者地址

Didier Alard：didier.alard@u-bordeaux.fr

波尔多大学，法国国家农业、食品与环境研究院，UMR BIOGECO，佩萨克，法国

+33 (0)5 40 00 87 74

数据可用性声明

根据自然遗产清查信息系统 (SINP) 的指南，本研究使用的数据包括敏感物种出现数据。除非经确认身份的个人提出有充分理由的申请，否则这些数据不得公开。因此，本研究使用的数据仅可应要求提供。

摘要

生物多样性开放获取数据库是整理物种出现数据并提升其可达性的重要资源。通过汇集不同的数据来源，这些数据库揭示了知识空间分布的不均衡：某些区域或分类群的调查比其他区域或分类群更充分。了解空间和分类学知识缺口的决定因素，有助于指导开放获取数据的使用。在此，我们确定了法国最大行政区内一个法国区域生物多样性数据库中知识缺口的决定因素。我们使用两项指标评估知识缺口：完整度和无知得分；评估对象为 8 个分类群，包括 5 个脊椎动物类群和 3 个无脊椎动物类群。我们分析了整个区域的数据，也在此前的三个次区域层面进行了分析，以确定可能影响知识缺口的因素。研究结果显示，与脊椎动物相比，无脊椎动物的知识缺口更大。总体而言，对于这两项指标，影响知识缺口的因素主要与地点可达性有关，而非生态吸引力。除农业压力对无脊椎动物的影响比对脊椎动物更显著外，所有类群的知识缺口都有相似的决定因素。最终，本研究强调，生物多样性治理通过地方资助和区域政治决策影响开放获取数据库中的知识分布。我们建议将生物多样性资助转向采样不足的分类群和调查不足区域，以减少这些偏差。如果无法做到这一点，使用从这些数据库中提取的数据的用户应借助知识缺口地图校正空间采样偏差 (SSP)，从而更准确地了解物种出现情况。

关键词

生物多样性数据缺口；分类学偏差；空间偏差；开放获取数据；生物多样性治理

1. 引言

随着全球生物多样性不断下降，关于物种分布和生物多样性变化的知识需求对于监测、研究和保护政策而言日益迫切（Butchart et al., 2010年）。近几十年来生物多样性数据库的发展，使生物多样性数据更易于获取且更加结构化（Grassle, 2000年；Reichman et al., 2011年）。全球生物多样性信息网络（GBIF）的扩展便是一个例证：该全球性开放获取数据库创建于2001年，截至2025年已汇集超过3,000,000,000条物种出现记录（GBIF, 2025年）。此类数据库汇集并标准化数据，可用于评估物种分布和种群趋势等生态动态，因而成为研究项目的宝贵资源（Ball-Damerow et al., 2019年；Marcer et al., 2022；Ribeiro et al., 2019年）。最终，它们为公共环境政策框架下的行动提供支持。

然而，开放获取数据库及其用于科学研究的可靠性，尤其是在出现数据方面，因其不准确性和偏差而受到批评；在某些情况下，这使其不适合用于评估物种种群趋势（Gaume & Desquilbet, 2024；Roel van Klink et al., 2020）。必须对数据集进行全面检查，以发现物种误鉴定或投影系统不一致等不准确之处（Colli- Silva et al., 2020；Freitas et al., 2020）。综述指出，由于空间采样偏差和调查努力缺乏协调，数据分布并不均衡，从而导致对全球和地方尺度上物种分布或种群趋势的误读（Beck et al., 2014；Grattarola et al., 2020；Rocha-Ortega et al., 2021；Yang et al., 2013）。开放获取数据库还受到分类学偏差的影响，脊椎动物分类群的抽样情况优于其他分类群（Hugues et al., 2021；Garcia-Rosello et al., 2025；Troudet, Grandcolas, Blin et al., 2017）。在线数据库中的偏差会影响数据再利用于科学项目的潜力，因为这些偏差增加了数据清洗和处理的多个步骤（Hill, 2012；Maldonado et al., 2015；Panter et al., 2020；Zizka et al., 2020）。例如，空间

采样偏差在使用机会性数据进行物种分布模型（SDMs）建模时必须加以校正，以提高模型的可靠性（Baker et al., 2024；Inman et al., 2021；Syfert et al., 2013）。因此，有关使用出现数据的担忧日益增加，因为对某些区域或分类群的调查不足会造成空间采样偏差（Chesshire et al., 2023；Garcia-Rosello et al., 2023）。

改善数据库面临的两项重要挑战是识别调查不足区域和分类群，并了解这些知识缺口的决定因素。可通过完整性估计量和无知度量来识别空间空白（Sousa-Baena et al., 2013；Girardello et al., 2019；Stropp et al., 2016；Ruete, 2015）。空间知识缺口的决定因素可分为两类：与地点可达性相关的变量，以及与地点生态吸引力相关的变量。地点可达性包括道路网密度、与大学的距离、城市化程度等（Correia et al., 2019；Mair & Ruete, 2016；Oliveira et al., 2016）；地点生态吸引力则包括与保护区的距离、预期栖息地丰富度或到访压力（Meyer et al, 2015；Sanchez-Fernandez et al., 2008）。然而，研究通常聚焦于单个物种或分类群，尽管目前没有证据表明不同分类群之间的这些知识缺口在空间上相吻合，尤其是因为地点的生态吸引力取决于物种本身（Prendergast & Eversham, 1997）。事实上，只有少数研究比较过不同分类群的知识缺口决定因素，但研究范围仍局限于某个特定的门或纲（Sanchez-Fernandez et al., 2022；Šmíd, 2022）。

此外，关于生物多样性知识缺口的研究大多在较大的空间尺度上开展，至少是在国家、洲际乃至全球层面。我们认为，因此这类研究只能考察在这些尺度上具有相关性的因素，而会忽略那些反映区域、文化或地理特征，并影响数据收集的更具体的决定因素。例如，在法国，一项名为国家自然遗产编目（法语简称 SINP）的开放获取生物多样性计划于2007年 设立，旨在提高法国公共生物多样性数据的可达性并促进其集中管理（INPN，2025年；Poncet,

2013年）。该计划主要依托区域平台，由各平台管理自己的数据库，并将区域收集的数据转入国家数据库（Robert，2021年）。因此，区域尺度似乎是数据收集和观察者网络管理的首选层级，也与区域自然保护政策直接相关。FAUNA观测站负责收集和管理法国行政区新阿基坦大区的SINP数据。2025年，该站汇编了超过11,100,000条出现数据，涵盖16,520个物种（FAUNA观测站，2025年）。

本文评估了FAUNA数据库中8个分类群的知识缺口，包括常见脊椎动物和昆虫类群。随后，我们旨在确定造成这些缺口的潜在决定因素。我们从三个角度分析其解释模式：（1）分类学角度：我们假设不同分类群的知识缺口在空间分布上并不一致；（2）生物地理学角度：我们研究地点的生态吸引力或可达性在决定知识缺口方面各自发挥的作用；（3）社会—行政角度：我们旨在了解区域数据库中的知识缺口，并探究此前三个大区的特殊作用。这三个大区历史上塑造了数据收集，影响了自然协会的参与，确定了各自辖区内的分类学和空间优先事项，直至它们于2015年合并，成立新阿基坦大区。为考虑方法学问题，我们采用两种不同的知识缺口指标。最终，本研究全面呈现了该地区的空间采样偏差，可为进一步的勘查工作和政治决策提供参考。

2. 材料与方法

2.1. 研究区域

新阿基坦大区是一个法国行政区，于2015年由三个较小的地区——阿基坦、普瓦图-夏朗德和利穆赞——经政治合并而成（. 1）。该地区绵延至西班牙边界以北400多公里，辖有12个

行政　　个行政次级单位，称为省。研究区域总面积为84,000 km ^{2} .

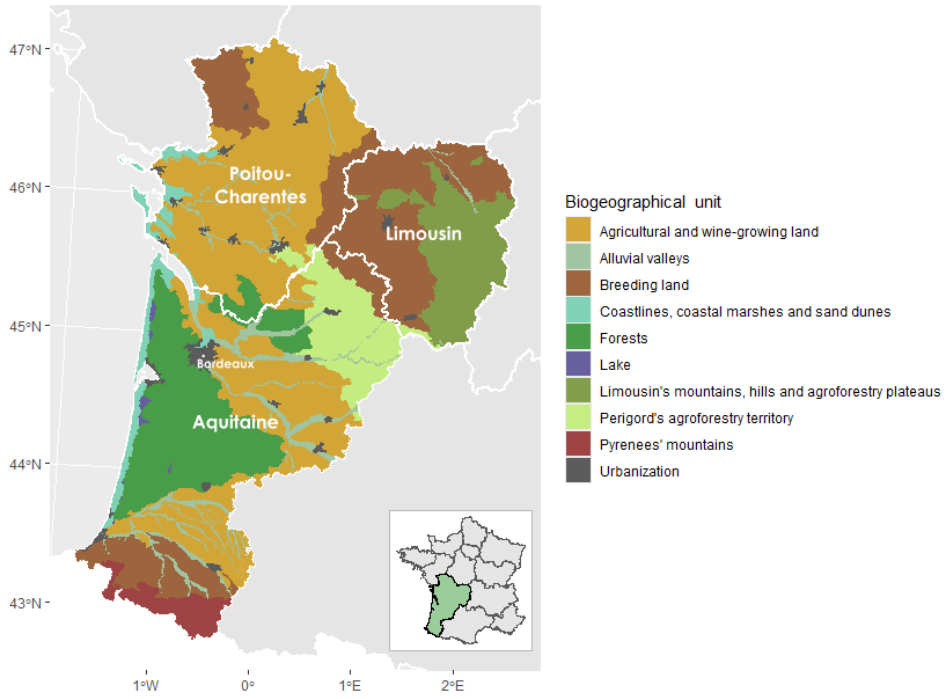


图 1 新阿基坦大区地图，展示其原组成地区和主要生物地理单元。

该地区气候温和，趋于海洋性。自然生境种类繁多，包括比利牛斯山脉、以沙丘为主的海岸线、森林、草地和湿地。然而，覆盖该地区大部分面积的是半自然生境，包括农业用地和畜牧用地。尽管新阿基坦是法国面积最大的行政区，但其人口密度总体较低且分布不均，通常在大西洋沿岸及大城市周边的市镇中较高。

原有各大区之间存在差异。原利穆赞大区以中山地为主，大城市较少。原普瓦图-夏朗德大区以农业和沿海地区为主。最后，原阿基坦大区的景观更加多样，从山区到海岸各不相同。

2.2. 数据筛选

出现记录于20th 2025 年6月从FAUNA数据库中提取，涵盖8个分类群：两栖动物、蝙蝠、陆生鸟类、陆生哺乳动物、爬行动物、日行性蝴蝶、蜻蜓目昆虫（蜻蜓和豆娘）以及直翅目（蚱蜢、蝗虫和蟋蟀）（FAUNA，2025年）。

由于数据质量和精度差异较大，我们使用 FME Workbench 2024.2.3（FME Software, 2025）进行筛选，仅保留符合以下标准的记录：

- 时间：仅使用自 2010年1月1日 起收集的记录。日期跨度超过 1 个月的记录将被剔除。
- 空间：仅保留完全位于新阿基坦大区内的数据。由面积超过 1 km^2 的多边形定位的记录将被剔除；由长度超过 1 公里的线串定位的记录也将被剔除。
- 分类学：仅使用鉴定至种级或更低级别的记录。种下级记录归入相应物种。
- 可靠性：根据 FAUNA 自动验证系统，剔除可靠性等级标记为“存疑”或“非常存疑”的记录。

筛选后，数据集包含 2,420,654 条出现记录。

我们采用由 856 个网格单元组成的 10 x 10 千米网格，将每条出现记录与相应网格单元关联。对于与多个网格单元相交的记录，将其关联到与多边形面积相交比例最高的网格单元。为避免任何边界效应，我们剔除了与新阿基坦大区相交面积不足 40% 的所有网格单元。

2.3. 空间知识缺口评估

我们使用两项指标评估该地区每个分类群在各网格单元中的知识缺口：基于已观测物种的完整度，以及基于调查努力的无知指数。相关计算使用 R 4.4.3 版完成，并整合到 FME Workbench 2024.2.3 版工作空间中（R核心团队，2025）。多分类群评分通过计算每个网格单元中所有分类群各项指标评分的平均值获得。

2.3.1. 基于完整性指标的知识缺口

我们使用 Chao2 基于出现频率的指数（Chao & Jost, 2012；Chao 等，2020；Sousa-Baena 等，2014；Troia & McMamanay, 2016）计算每个网格单元的完整度。该指数根据被视为“稀有”的物种，估算每个网格单元中未检测到的物种数（S_{und}）。本研究以年份作为样本的替代指标。采用以下无偏公式（Chao & Colwell, 2017；Gotelli & Colwell, 2011；OBIS, 2025）：

$$S_{und}=\left(\frac{m-1}{m}\right)\frac{q_1(q_1-1)}{2(q_2+1)}$$

其中，S_{und}：该网格单元中未检测到的物种数；

m：样本总数（15 年）；

q_1：一个采样单元（1 年）中观测到的物种数；

q_2：恰好在两个采样单元（两年）中观测到的物种数。

然后利用未检测到的物种数来估计每个网格单元的预期物种数（S_{exp}）：

$$S_{\exp}=S_{obs}+S_{und}$$

其中 S_{exp}：网格单元上预期的物种数；

S_{obs}：网格单元上观测到的物种数。

对于出现记录较少的网格单元，Chao2 指数并不可靠。由于缺乏数据，单例种和双例种的数量可能特别高，从而导致对未检测到的物种数量的估计偏高。因此，对于出现记录少于 10 条的网格单元，不计算预期物种数量。

单例种和双例种的数量与采样力度高度相关，这可能导致同一生物地理单元内的预期物种数量出现较大波动。为比较各生物地理单元内的知识缺口，我们计算了每个生物地理区域的预期物种平均数量（图 1）。为此，我们确定了占多数的生物地理单元，并将其关联到每个 10×10 公里网格单元。预期物种平均数量按以下公式计算：

$$m_{déf} = \frac{\sum S_{att.i}}{N_i}$$

其中，m_{def}：该生物地理单元的预期物种平均数量；

$S_{att.i}$：相应生物地理单元中网格单元 i 的预期物种数量；

N_i：属于相应生物地理单元的网格单元总数。

随后，为同一生物地理单元内的每个网格单元赋予相同的预期物种平均数。观测记录少于 10 条的网格单元也会被赋予相应生物地理单元的预期物种平均数。

知　　基于完整度的知识缺口使用以下公式计算：

$$G_{com} = \frac{S_{obs}}{m_{unep}}$$

其中 G_{com}：基于完整性指标的知识缺口。

2.3.2. 基于无知指数指标的知识缺口

为了利用调查努力计算知识缺口，我们采用了半无知指数算法（Ruete，2015）。该指数旨在根据从 GBIF 或 SINP 等公共数据库中提取的出现数据估算无知得分。它基于这样的假设：同一分类群中的物种具有相似的偏差，因为观察者通常专注于特定分类群，而非某个特定物种。根据每个网格单元记录的观测记录数量，该指数可用于识别调查不足的网格单元。

我们首先确定每个网格单元中观测到的物种数与观测记录数之间的比值（物种观测指数）（Ruete，2015）：

$$S_i = N_i / R_i$$

其中 S_i：网格单元 i的物种观测指数；

N_i：网格单元中的出现记录数；

R_i：该网格单元中记录的物种数。

为评估半无知得分，使用观测记录数阈值 $O_{0.5}$ 。该阈值对应于足够数量的观测记录，使我们可以假定某物种无观测记录的情况，有50%的可能性是由于该物种在该网格单元中真实缺失，另有50%的可能性是由于未检测到该物种（Ruete，2015）。由于这是一个渐近曲线，$O_{0.5}$越高，无知得分达到0的速度就越慢（Ruete，2015）。

此处测试了$O_{0.05}$ 的 3 个取值：1、5 和 10。结果随$O_{0.05}$ 取值变化不大，因此采用中间值 5。

一旦确定了物种观测记录指数和参考观测记录数量，即可使用以下公式计算最终的半无知得分（Ruete，2015）：

$$G_{ign}=O_{0.5}/(N_i+O_{0.5})$$

其中，G_{ign}：基于网格单元 *i*的无知指数指标计算得出的知识缺口。

无知得分与完整度采用相反的梯度：使用无知得分时，数值越接近 1，知识缺口越大；使用完整性指标时，数值越接近 0，知识缺口越大。为比较这两种指标，按以下方式计算了无知得分的反向值：

$$G_{inv}=1-G_{ign}$$

在以下所有分析中，均使用无知得分的反向值作为知识得分。因此，随着完整度和知识得分增加并趋近于 1，知识缺口会缩小。

2.4. 解释知识缺口的潜在决定因素

为确定可能导致知识缺口出现空间或分类学差异的变量，我们根据研究区域数据集的可用性，选取了八个潜在决定因素，分别涉及自然度等级、生物多样性得分、保护区占比、平均道路密度、城市化压力、农业压力和到访压力。使用 ArcGIS Pro 3.2.0（ESRI，2024年），为每个网格赋予上述各变量的取值。

自然度最近已在国土范围内以20米分辨率完成制图，其依据是被设计为三个基本组成部分的景观代理指标：生物物理完整性、过程自发性和时空连续性（Guetté等，2021年）。三个栅格数据从UICN网站下载。使用了将这三个组成部分汇总为潜在自然度的数据集。该数据集经裁剪以匹配研究区域的范围。利用ArcGIS“分区统计表”工具，提取了每个10公里×10公里网格单元的自然度最大值。我们没有使用平均值

自然度的平均值，因为与该指标的精细分辨率相比，较大的网格单元尺寸往往会平滑数据，导致各网格单元之间的变化很小。我们假定，可能引起观察者兴趣的小型生境斑块可借助这一景观分析方法加以区分。

生物多样性得分来自新阿基坦热点项目。该项目利用物种出现数据、物种分布建模和空间优先级排序，以1公里分辨率识别保护优先级较高的区域（Collectif，2021年）。本研究采用了Zonation方法的加性效益函数（ABF）结果（Moilanen，2007年），因为该结果能够突出物种丰富区域中的网格单元。利用ArcGIS的“分区统计表”工具，提取了每个10 × 10公里网格单元的累积生物多样性平均值。直观而言，生物多样性热点与生态吸引力的关联比与地点可达性的关联更强，尤其是因为累积物种丰富度是根据建模得到的分布计算的。然而，新阿基坦热点研究的方法深受地点访问情况的影响，倾向于突出到访量最高的地点。它处于到访压力与自然度的交汇处。

平均道路密度被用作地点可达性的替代指标。该指标利用国家BD TOPO数据集中的“VOIE NOMMEE”形状文件计算得出（法国国家地理研究所，2025年）。首先根据新阿基坦边界范围对该文件进行裁剪，然后将“线密度”工具的搜索半径设置为5,000米，输出网格单元大小设置为1公里。随后，使用“分区统计表”工具获取每个网格单元的平均道路密度。

保护区被确定为知识缺口的一个决定因素，因为观察者更有可能前往生态价值较高的区域。此外，保护区可能会提升地点（网格单元）的吸引力，也可能成为生物多样性治理的地方行动者；它们可以在辖区内开展特定的生物多样性调查，并为公民科学研究筹集资金或发起此类研究。不同生态保护类型的数据集被合并为一类：国家自然公园（‘国家公园’）、区域自然公园（‘区域公园’）、

国家级自然保护区（‘国家级自然保护区’）、区域级自然保护区（‘区域级自然保护区’）以及‘具有特殊动植物价值的地点（‘ZNIEFF’）（INPN，2025年；Magellium，2025年；FAUNA观测站，2025年；法国国家公园，2025年）。随后使用“相交汇总”工具，计算每个网格单元中被保护区覆盖的面积百分比。

2021年，博物馆下属PatriNat部门在法国国家研究计划框架下绘制了对生物多样性具有显著影响的人为压力分布图（Cherrier等，2021）。汇总数据集将压力按10千米分辨率划分为三类：农业、城市化和到访量。栅格数据从PatriNat网站下载，并裁剪至新阿基坦大区的范围。由于两项研究使用的是同一10 km x 10 km网格，因此未进行进一步处理。根据共同标识符，将三类压力中每一类的压力值连接至各个网格单元。

2.5. 统计分析

为识别八个分类群之间知识缺口的潜在一致性，针对每个指标分别进行了一次主成分分析，使用[856 网格单元 x 8个得分] 数据矩阵。对于两个指标，即完整度和知识，网格单元得分均采用10x10km网格分辨率。我们考察了前大区层级是否因调查方法和治理方式的差异而呈现不同模式。为根据三个前大区划分网格单元，使用ArGIS的“相交汇总”工具，对网格单元数据集与前大区边界的形状文件进行处理。每个网格单元均被连接至与其表面积相交最多的前大区。随后，在R 2024.12.1中使用*FactoMineR*软件包对数据集进行主成分分析（Le等，2008；R核心团队，2025）。

我们通过对 [856 个网格单元 × 7 个决定因素]进行主成分分析，探究了知识缺口的潜在空间决定因素。该分析使我们能够判断

土地覆盖类型和地理特征的异质性，尤其是前行政区内部及彼此之间的异质性，是否会转化为多样的决定因素组合。

为研究潜在决定因素与各分类群知识缺口之间的关系，我们在区域尺度上针对脊椎动物、无脊椎动物以及所有分类群进行了冗余分析（RDA）。分析使用了*vegan*程序包中的 rda() 函数（Oksanen等，2025）。

根据主成分分析识别出的知识缺口相似性，将分类群合并为两个更广泛的类别：除蝙蝠外的脊椎动物（两栖动物、爬行动物、陆生鸟类和非飞行哺乳动物）以及无脊椎动物（蜻蜓目昆虫、直翅目、日行性蝴蝶）。对于每个类别，使用 FME Workbench 中的“统计计算器”工具（FME Software，2025），计算其中各分类群在每个网格单元的知识缺口得分平均值。一个数据集的平均得分依据完整度计算，另一个则依据知识指数计算。

为研究变量与各个分类群知识缺口之间的关系，使用*stats*程序包进行了线性回归（R核心团队，2025）。针对两种知识缺口指标，分别对每个前子区域和整个区域中的各个分类群，以及脊椎动物和无脊椎动物进行了回归分析。

所有分析均使用 R 2024.12.1（R核心团队，2025）完成。

3. 结果

3.1. 知识缺口的空间和分类学分布

两项指标呈现出相似的空间分布（图 2），完整度和知识得分最低的区域位于该大区西北部，也就是前普瓦图-夏朗德大区的范围内。在每项指标中，脊椎动物和无脊椎动物的得分分布相似。总体而言，无脊椎动物的知识缺口

比脊椎动物更大（图 2）。所有分类群的得分分布更均匀，因而减小了差异和分布异质性。

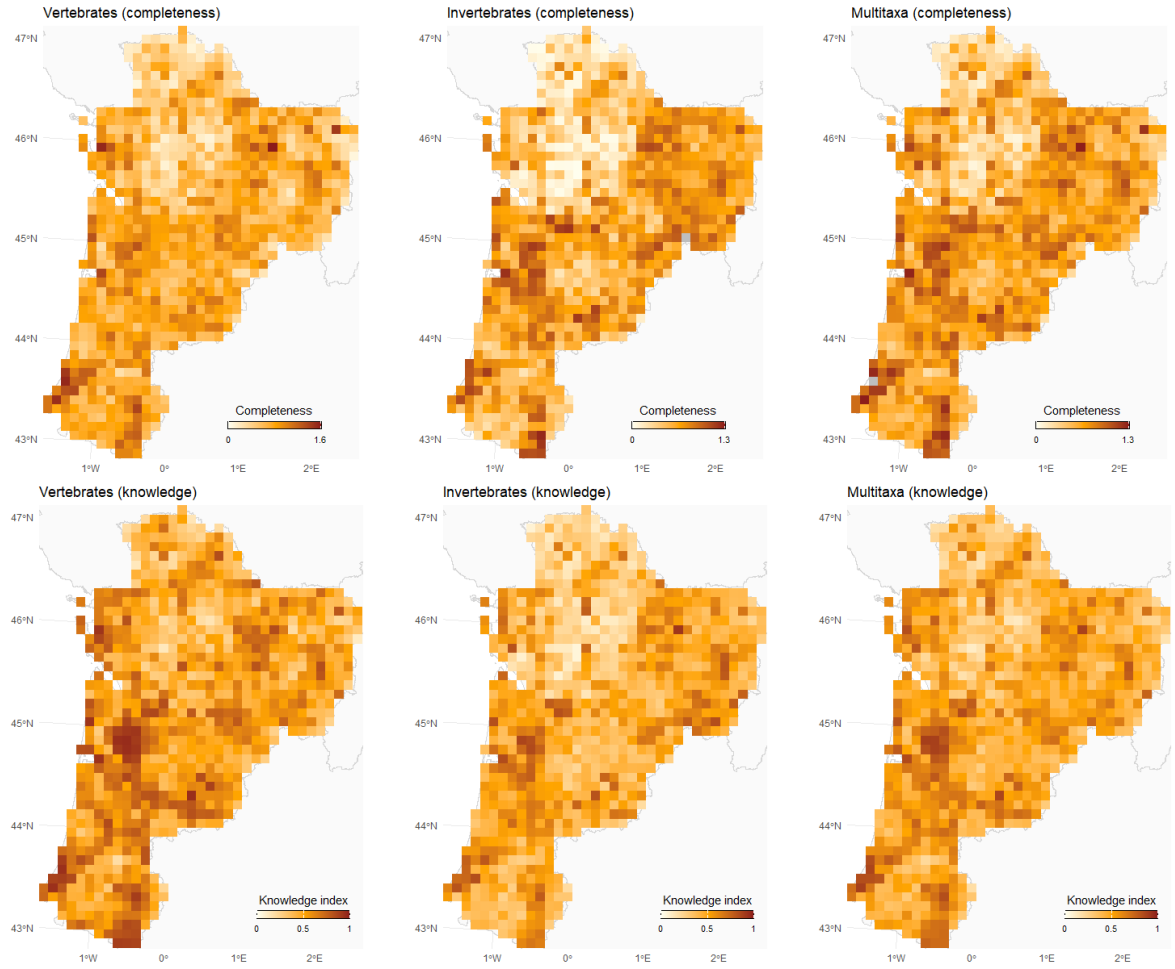


图 2 新阿基坦地区脊椎动物、无脊椎动物及所有分类群的知识度和完整度得分空间分布，采用 10 千米 × 10 千米网格单元表示。完整度得分可能高于 1，因为已观测物种数量可能超过预期物种数量。该梯度既能展示出现这种情况的区域，也能展示采样最少区域。另一方面，知识得分的范围限定在 0 到 1 之间。

知识缺口的模式在前大区内部及彼此之间呈现异质性，表明知识缺口的成因有所不同。在西北部，前大区普瓦图-夏朗德

在所有分类群和指标中，其整体知识度和完整度评分均较低，尤其是无脊椎动物（图 2 / 补充材料图S1和图S2）。位于该地区东部和北部边界的网格单元总体知识缺口较大。该地区不同分类群的分布模式各不相同。例如，海岸线附近的鸟类总体调查更为充分，完整度得分较高，而其他分类群则并非如此（补充材料图S1和图S2）。

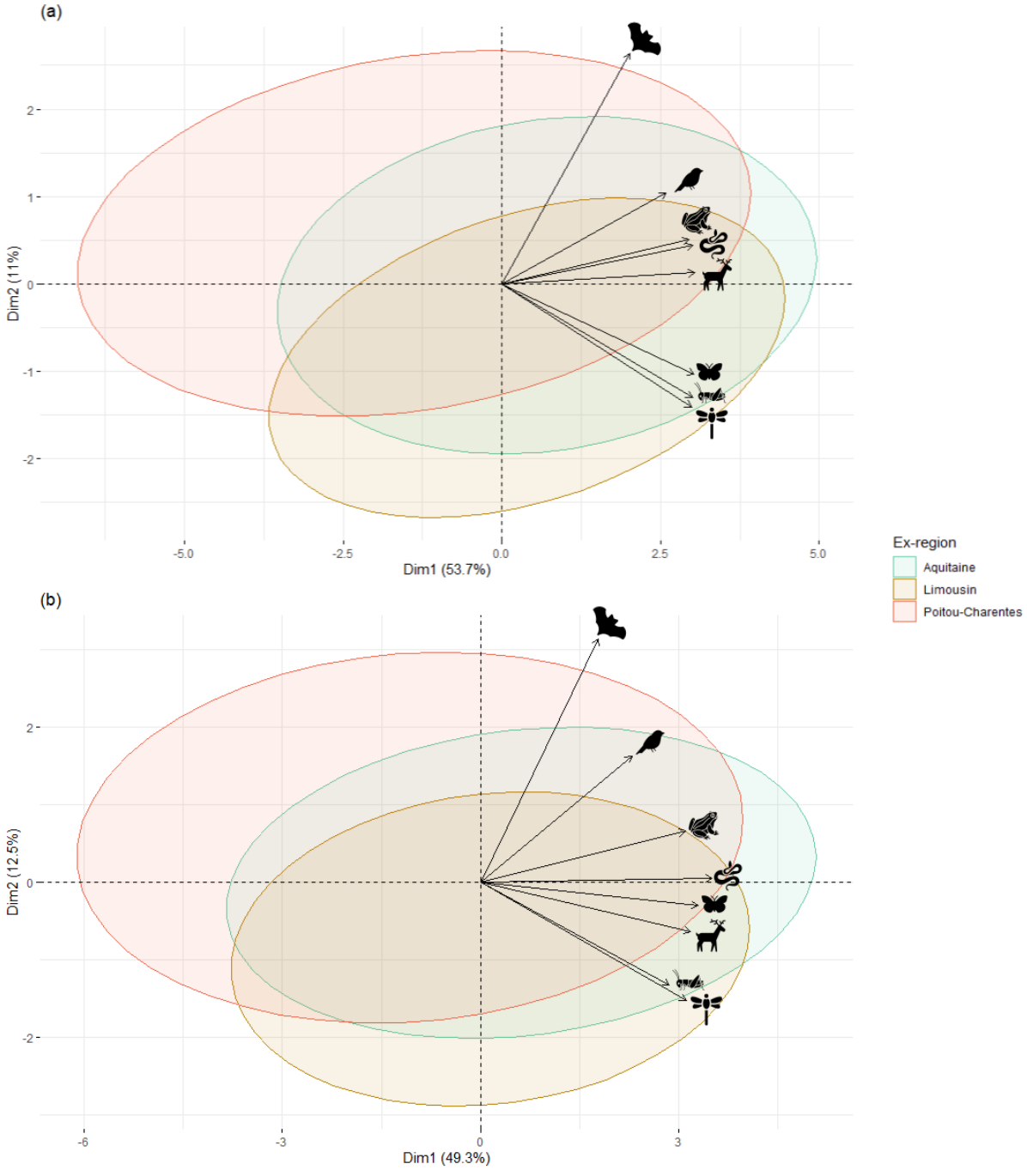


图 3 基于 856 个网格单元的矩阵，对各分类群的完整度得分（图 3a）和知识得分（图 3b）进行主成分分析（PCA）双标图，并按第二维度从高到低排列（a）：蝙蝠、陆生鸟类、两栖动物、爬行动物、陆生哺乳动物、日行性蝴蝶、直翅目、蜻蜓目昆虫。向量箭头的方向和长度表示各分类群得分在主成分上的载荷。椭圆根据质心和 90% 置信区间，表示各前亚区域内网格单元的分布。

对各分类群的知识度和完整度评分进行主成分分析，结果一致（图 3）。在基于完整度得分的第一项主成分分析中（图 3a），第一维度解释了总方差的 53.7%，而第二维度解释了 11%。

对完整度得分的分析可识别出两个分类群簇：第一组为脊椎动物，包括鸟类、两栖动物、爬行动物和哺乳动物；第二组为无脊椎动物，包括蜻蜓目昆虫、直翅目和日行性蝴蝶。无脊椎动物的得分与蝙蝠的得分正交。脊椎动物仅与第一维度呈正相关，而无脊椎动物与第一维度呈正相关、与第二维度呈负相关。

总体而言，网格单元在第一轴上越偏向正向区域，相关知识就越充分。反之，位于该轴负向区域则表明该网格单元存在较多知识缺口。这些缺口的性质，也就是该网格单元中信息不完整的分类群，会影响网格单元在主成分分析第二轴上的位置。因此，这些知识缺口主要涉及位于上方（第二轴正向区域）的蝙蝠，以及位于第二轴负向区域的无脊椎动物（《昆虫》）。

关于用于可视化前大区网格单元分布的椭圆，其在轴 1 上的平均位置描述了完整度的平均水平，而其在

第二轴上的位置则展示了造成知识缺口的分类群的性质。椭圆的大小反映了前大区网格单元中这些水平的变异程度以及缺口的性质。因此，椭圆的分布显示各前大区之间的知识缺口存在差异，其中阿基坦的平均完整度较高。普瓦图-夏朗德的平均得分较低，且不同情况之间的异质性更大。

基于知识得分进行的第二次主成分分析也呈现出相似的模式（图3b）。第一维度解释了总方差的49.3%，第二维度解释了12.5%。与基于完整度的主成分分析略有不同，蝴蝶与第一维度的关联更紧密，因此其知识分布与脊椎动物更为接近。前大区在知识得分和完整度得分主成分分析中的位置相似（图3）。

3.2. 知识缺口的决定因素

基于知识缺口决定因素的主成分分析中，前两个维度解释了约58%的变异（图4）。其中大部分变异由第一维度解释，接近37%。自然度和保护区与第一维度高度相关，且主要与农业压力截然相反；农业压力则与第一维度的负向部分相关。生物多样性热点主要与自然度和保护区的梯度相关，似乎位于到访压力与自然度之间的交汇处。总体而言，与栖息地质量相关的决定因素（自然度、保护区和热点）与人类压力相关的变量（农业和城市化）相对立，合乎逻辑。

与普瓦图-夏朗德这一较为均质的前行政区相比，阿基坦前子区域的网格单元在决定因素空间中的离散程度更高，这与其显著的栖息地异质性相一致。此外，三个不同前行政区的网格单元在决定因素空间中的平均位置相近，因为三个椭圆均

以变量空间为中心，这表明，尽管某些前行政区中的差异更大，但它们在每个前行政区内具有相同的性质。

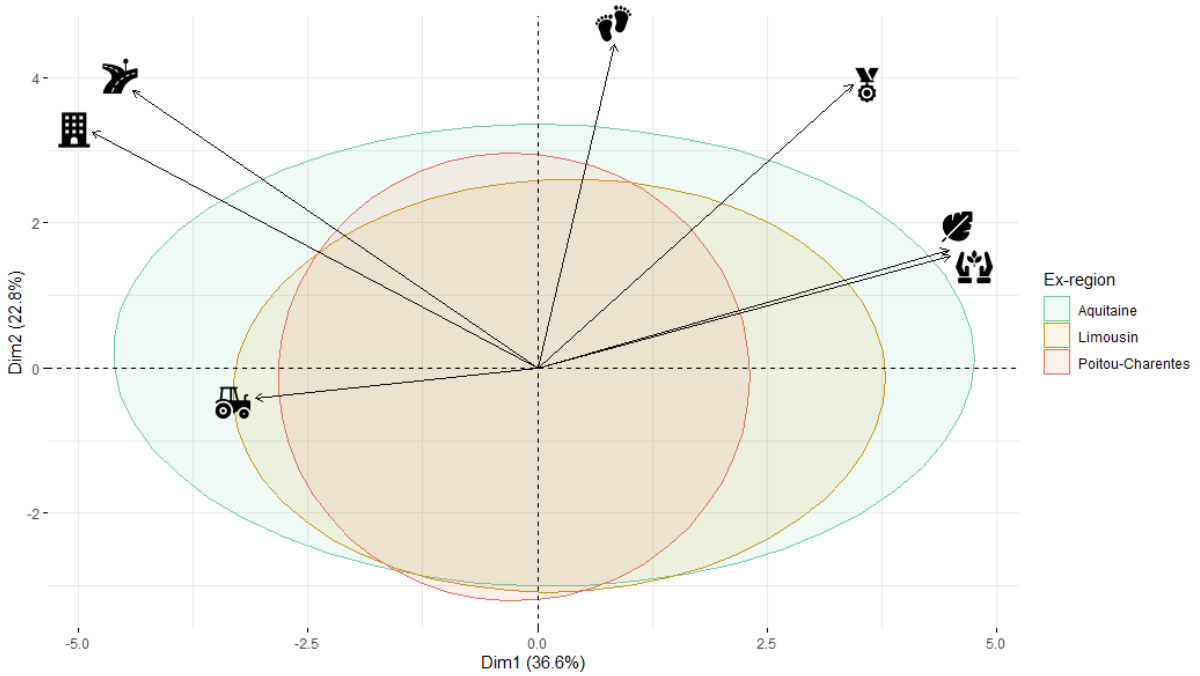


图4基于856个网格单元与新阿基坦地区知识缺口的7项决定因素矩阵的主成分分析双标图，在第一维度上从左到右依次为：城市化压力（建筑物）、道路密度（道路）、农业压力（拖拉机）、到访压力（足迹）、热点（奖章）、自然度（叶子）和保护区面积（手）。向量箭头的方向和长度显示各决定因素对主成分的载荷。椭圆利用质心和90%置信区间表示各前亚区域内网格单元的分布。

脊椎动物和无脊椎动物类群的知识缺口决定因素呈现出不同的模式（图 5、图 6），证实了两个亚门之间存在不同的观测偏差。各分类群中不同前区域之间，关系强度以及变量的统计显著性存在差异（图 6）。调整后的 R^2 值较低，表明选定变量只能解释知识缺口变化的一部分。

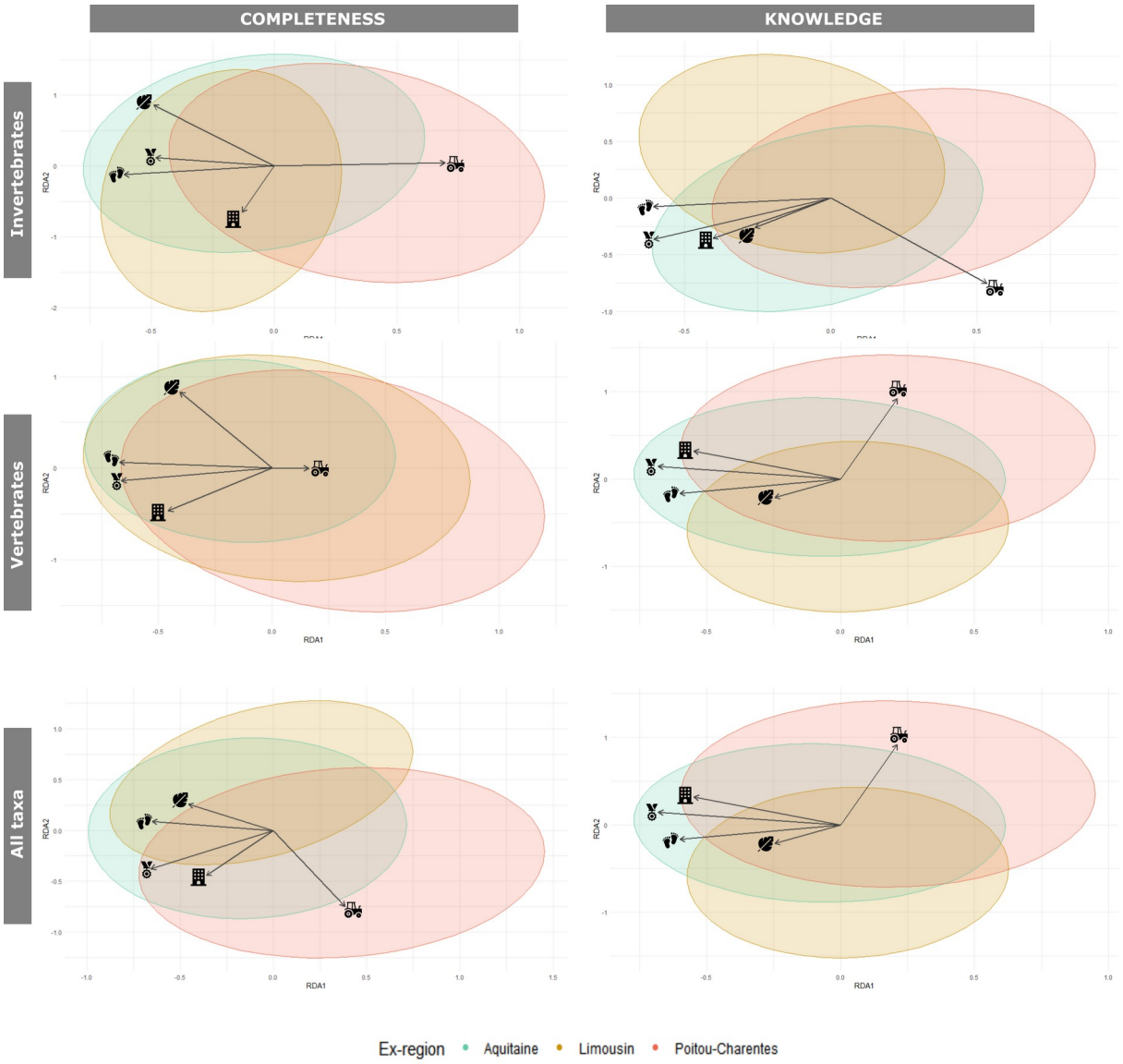


图 5 冗余分析（RDA）展示了知识缺口的决定因素对脊椎动物、无脊椎动物及所有分类群的完整度和知识得分的贡献。向量箭头的方向和长度表示知识缺口得分与决定因素之间相关性的强弱。

知识缺口主要与热点密度和城市化压力呈负相关。总体而言，生物多样性热点对知识缺口的显著贡献更大，介于 0.3 至 0.6 之间，这意味着生物多样性热点中的知识缺口较少（图 6）。对于大多数

分类群而言，城市化压力也是一个显著的决定因素，知识缺口在城市化程度较高的地区有所减少。其贡献通常较为适中，介于 0.02 至 0.07 之间。所有分类群和前大区均存在到访压力，尽管其影响较弱，但对知识缺口的贡献为正。

其他变量总体上是完整度和知识得分较不显著的决定因素。在达到显著水平时，农业压力是 RDA 1 正向端唯一的变量，但它对知识得分和完整度得分有负向贡献；随着农业压力增加，这两项得分会降低（图 5）。事实上，农业压力对第一维度有正向影响，因此与其他决定因素正交；其他决定因素均位于 RDA 1 的负向一侧（图 5）。不过，利穆赞的两栖动物知识得分是个例外，农业压力对其有正向贡献（图 6）。其他变量的大部分变异则体现在 RDA 2 上，例如自然度和城市化压力，它们对完整度得分和知识得分有正向影响（图 5）。

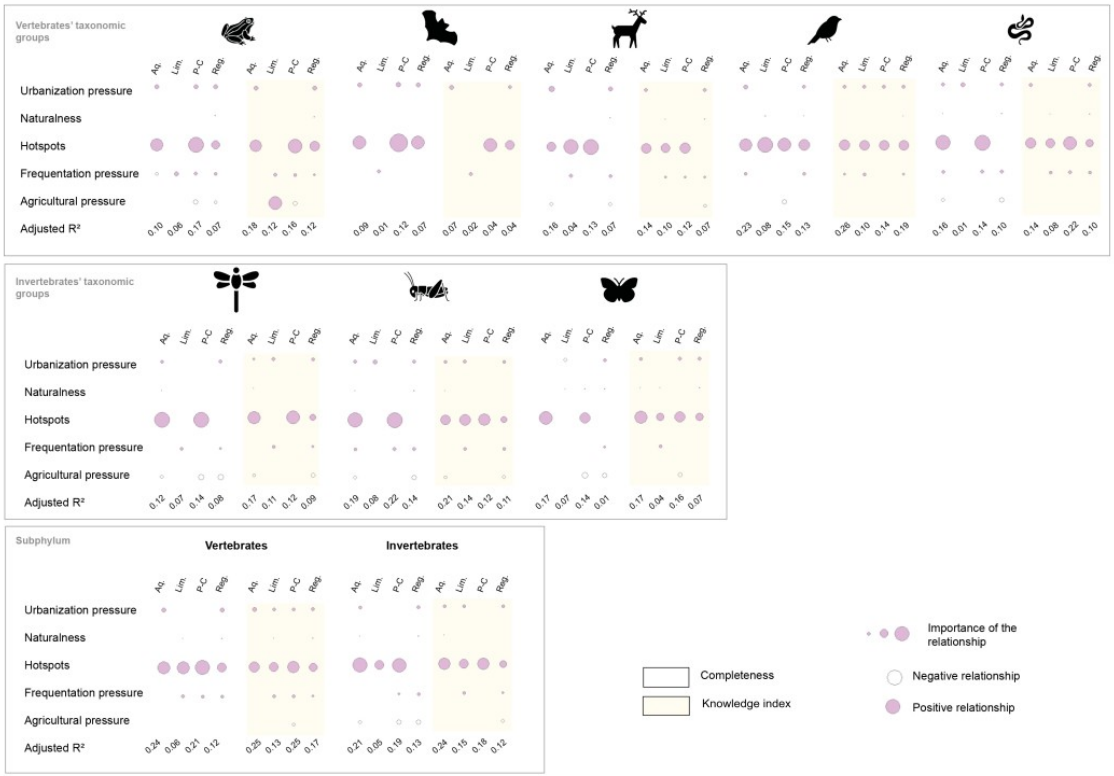


图 6知识缺口得分与线性回归模型的结果

各分类群和亚门的决定因素。此处仅显示 p值 <0.05 的变量。圆圈大小表示关系强度。粉色圆圈表示正相关关系，白色圆圈表示负相关关系。

知识缺口的决定因素在不同分类群之间存在差异。就脊椎动物而言，爬行动物与两栖动物之间存在相似性，其主要受热点和到访压力驱动（图 6）。鸟类和哺乳动物表现出更多差异，因为它们主要受热点和城市化压力影响。无脊椎动物的得分受热点和城市化压力驱动，同时也受农业压力驱动，而农业压力对无脊椎动物的显著贡献强于对脊椎动物的贡献（图 5，图 6）。

4. 讨论

通过评估开放获取在线区域数据库中的知识缺口，并采用两种不同的采样力度衡量指标，我们从三个视角识别出其潜在规律：地理视角，考察知识缺口及其决定因素的空间分布；分类学视角，探究不同分类群之间的差异；以及社会行政视角，考察前大区的影响及生物多样性治理。研究发现显著的地理和分类学差异，且两项指标得出的结果一致。知识缺口的空间分布在各前大区和分类群之间并不均衡，导致若干区域和分类群的采样不足。属于同一亚门的分类群（脊椎动物或无脊椎动物）之间往往更为相似。在所有研究的分类群中，新阿基坦的知识缺口受地点可达性的影响，大于受地点生态吸引力的影响。生物多样性热点和城市化压力对采样力度的影响最强，也最为稳定。然而，脊椎动物和无脊椎动物的决定因素呈现不同模式，尤其是在与农业压力的关系方面。

4.1. 知识缺口的规律：分类学偏差的影响

知识缺口在不同分类群之间的比较会因所选尺度而呈现不同结果。

当将分类群划分为脊椎动物和无脊椎动物时，两类群的知识缺口分布呈现出相似的空间格局。无脊椎动物的知识度和完整度评分总体上弱于脊椎动物，这很可能是分类偏差所致。事实上，脊椎动物物种总体上能得到专业调查以及公民科学项目更好的抽样，因为它们更能引发兴趣（Petersen等, 2021年；Phaka等, 2022）。它们也被认为比无脊椎动物更容易采样；无脊椎动物的物种可能需要特定的采样技术和实验室分析才能鉴定（Doxon等, 2011年；Sittenthaler等, 2023）。这导致无脊椎动物在开放获取数据库中的知识缺口更为突出，科学文献的结果也证实了这一点（Clark & May, 2002；Collen等, 2008；Correia et al., 2019年；Titley等, 2017）。

然而，对各分类群采样偏差的比较表明，完整度和知识得分的幅度以及知识缺口分布存在更大的差异，这证实了某些分类群的抽样程度优于其他分类群（Rocha-Ortega et al., 2021年）。例如，总体而言，鸟类的调查程度优于蝙蝠和直翅目，其中最佳抽样区域靠近大西洋海岸线；而两栖动物的较高得分则位于城市化地区附近。观察者的样点选择偏差可能对此产生影响，因为观察者更倾向于抽样适宜生境，而这些生境在不同分类群之间未必位于相同区域（De Araujo & Ramos, 2021年；Troia & McManamay, 2016年；Mentges et al., 2020）。

总体而言，除农业压力外，各分类群的空间采样决定因素相似。农业压力仅与少数前大区中的少数脊椎动物分类群存在显著关联，却会影响三个无脊椎动物分类群在部分前大区中的知识缺口。农业压力与知识度和完整度评分大多呈负相关；随着农业

压力增大，评分会降低。蝴蝶和蜻蜓等无脊椎动物大多与开阔生境相关，包括草地和田野。农业区属于开阔生境，人们会在此进行抽样，以研究无脊椎动物的丰富度和多样性，这表明农业区的知识缺口可能较小（Dolny等，2021）。然而，化学处理以及人类对这些半自然生境的开发利用，也已被证明会对蝴蝶和其他无脊椎动物的种群产生负面影响（Rands和Sotherton，1986；Sanchez-Bayo，2021）。因此，在可进入的地点内，博物学家很可能会优先选择有利于相关分类群的生境（Andrade-Silva等，2022；Boakes等，2010；Martin等，2012）。这意味着，无脊椎动物调查会选择人类影响较小的开阔生境，而脊椎动物调查则会选择森林等封闭生境。

蝙蝠呈现出不同于其他分类群的模式，这一独特性有助于展示新阿基坦大区知识缺口动态的变化。研究表明，蝙蝠的抽样深受空间偏差影响，导致调查努力分布不均（Fisher-Phelps等，2017）。知识空缺分布并未呈现任何区域外的结构性特征，这表明还有其他偏差在发挥作用。脊椎动物和无脊椎动物的观察者可能会对多个分类群进行抽样，而蝙蝠主要由蝙蝠学家采样，他们会使用声学探测等物种采样技术（O’Farrell和Gannon，1999）。根据生境选择偏差，他们主要在已知适宜蝙蝠栖息的区域进行采样，例如洞穴、建筑物或树木，这也解释了为何城市化压力对蝙蝠知识缺口的影响略强。因此，由于采用了群落特异性采样方法，蝙蝠呈现出不同的知识缺口模式。

4.2. 空间采样偏差与知识空缺分布

令人惊讶的是，在知识度和完整度评分之间，知识缺口的分布或决定因素并未观察到显著差异。由于采样力度受观察者偏差的影响，观察者往往更关注稀有物种和关键种，而非常见物种（Boakes等，2010；Garcillan等，2008；Gardiner等，2012；Minteer

等，2014）。这会导致某些区域只有少数物种的观测记录数量很高，从而造成两种指标之间的差异。在开放获取数据库中，与种群规模相比，稀有物种的观测记录数量会不成比例地偏高。然而，勘查偏差似乎不会影响对知识缺口的评估，因为我们的研究结果符合物种累积曲线所揭示的规律：已观测物种的数量随样本数量增加而增加（Ugland等，2003）。这表明，观测记录数量和物种数量可以互换，用作空间采样偏差的代理指标。

在所有指标和分类群中，都发现了知识缺口区域分布的一般规律。该地区北部，尤其是普瓦图-夏朗德，整体上知识缺口更为严重；而海岸线附近及比利牛斯山脉的区域则调查得更充分。普瓦图-夏朗德以农业用地为主，其生态重要性通常低于山区和沿海栖息地，而后两者的特有物种和易危物种数量更多。因此产生了生境选择偏差：观察者往往倾向于前往预期物种丰富度、多样性和生态重要性较高的地区（Dennis & Thomas, 2000年；Mentges et al., 2020）。

在更精细的尺度上探究知识缺口的决定因素时，研究发现，在所有前述区域和指标中，空间采样偏差均由地点可达性而非地点生态吸引力所驱动。城市化压力、热点和到访压力与知识缺口的关系最强且最为一致，并且随着知识缺口增加而增强。即使使用其他变量作为代理指标，可达性仍被证明会显著影响空间采样（Andrade-Silva等，2022；Correia et al.，2019；Girardello et al.，2018；Mair & Ruete，2016；Sobral-Souza et al.，2024）。靠近城市、道路网络发达且人口密度较高的区域，整体上得到更充分的调查。然而，可达区域往往与人为改造区域相关，导致在生态潜在价值较低的地区进行更充分的采样（Collen et al.，2008；Ronquillo et al.，2020）。归根结底，显著的知识水平与地点所受关注度有关，

这些地点可能吸引访客，也可能因交通便利且靠近城市而有更多人到访。由于区域丰富度被用于评估空间和保护优先事项，这些优先事项可能因空间采样偏差而被高估或低估（Grand et al.，2007；Grattarola, Martinez-Lanfranco, Botto et al.，2020；Engemann et al.，2015）。有研究建议，通过评估空间知识缺口来校正物种分布，从而做出依据更充分的决策。

这些决定因素只能解释知识缺口变化的一部分，因为所有分类群和前大区的调整后 R^2 值都相对较低。其他探讨空间采样偏差决定因素的研究也得出了类似结果（Correia et al., 2019 年；Mair & Ruete, 2016 年）。使用无知得分时，研究表明，无知得分的 $O_{O.O5}$ 参数会导致 R^2 发生变化；随着 $O_{0.05}$ 增大，这种变化会减小（Correia et al., 2019 年；Mair & Ruete, 2016 年）。不过，这些变化大多较小，因为 R^2 值处于相近范围，仅绝对值有所不同。

4.3. 生物多样性治理如何塑造区域知识缺口

该地区部分无法解释的方差，可能与新阿基坦SINP计划在政治和历史层面的组织结构有关。尽管在线公共数据库收集的数据日益增多，前大区的动态仍强烈影响FAUNA数据库的数据收集（Touroult et al., 2020）。合并之前，各前大区的博物学家网络在生物多样性数据方面的组织方式和运作机制各不相同。有些网络早已开始与SINP计划共享数据，另一些则与私人数据库共享数据，或创建自己的数据库来整理数据。前大区的合并并未促成一个统一的新阿基坦数据库的建立；要建立这样的数据库，需要对数百万条历史数据进行格式整理和传输。因此，大多数博物学家网络仍沿用最初的数据整理方式，这也解释了它们在向FAUNA数据库传输数据方面为何做法各异。

普瓦图-夏朗德的博物学家组织向FAUNA数据库传输的数据，大多由公共资金资助；而阿基坦的博物学家也更可能传输由私人资金资助的数据。例如，陆生哺乳动物每个网格单元的平均记录数，在普瓦图-夏朗德为288条，在利穆赞为378条，在阿基坦为464条（FAUNA, 2025年）。因此，知识缺口的空间分布可能受到观察者在区域SINP平台上共享数据意愿的影响。新阿基坦地区各地的数据共享差异，很可能是导致前大区之间知识缺口差异的最主要决定因素。本研究未计入共享差异，因为要估算记录者传输数据集的意愿，既困难又存在争议。

生物多样性调查不仅受兴趣驱动，资助在决定调查哪些物种和地区方面也发挥着重要作用。与动物群相关的项目获得的资助并不均衡，往往聚焦于特定地区和分类群，这可能导致不同分类群和地区之间出现差异（Martin-Lopez等，2009年；Titley等，2017）。在法国和新阿基坦，鸟类调查较为充分，历史悠久的公民参与项目（如 STOC）确保了长期数据集的积累，也使数据的总体空间覆盖更广（Vimont等，2025年）。对于其他分类群，尤其是无脊椎动物，调查资助在时间和空间上的分布更可能零散，因而造成区域和次区域尺度上的知识缺口。研究表明，在同一分类群内，保护资金比起受威胁物种，更倾向于投向旗舰物种，因为后者更容易获得资金和公众关注（Guénard等，2025年）。如果数据获取项目的资助也受相同机制驱动，那么生物多样性治理者和利益相关方就亟须在资助分配决策中纳入物种脆弱性评估和知识缺口评估。

在空间层面，资助也会影响知识缺口的变化：对于大多数组别，位于比利牛斯国家公园内的网格单元，其抽样情况优于其他区域，这不仅是因为该地点颇具吸引力，也得益于科学项目的内部资助。总体而言，在新阿基坦大区，国家公园和区域公园是生物多样性数据的重要

贡献者，这一点可从知识空缺分布的空间分布中看出。事实上，法国保护区会定期获得大量生物多样性项目资助，尤其是来自法国政府的资助。这导致知识空缺分布存在显著差异：获得资助、抽样较为充分的区域与几乎没有获得资助的区域之间形成了鲜明对比。

最后，研究表明，除了资助之外，生物多样性调查和保护也与政治决策密切相关（Le Velly 等人，2024年；Magnusson 等人，2018年）。保护区边界最终由利益相关方决定，这对生物多样性保护和调查的地点有着重大影响。此外，短暂的政治任期与长期保护政策需求之间的时间错配，会影响生物多样性计划及其持续时间（Steinberg，2009年）。因此，应从社会行政驱动因素的角度进一步探讨知识缺口的决定因素，以提升生物多样性知识。

正如所示，知识缺口及其决定因素复杂、错综且涉及多方面。为更深入地探索和理解其潜在动态，还需开展进一步研究。应研究数据获取的时间动态，以识别更多采样偏差。事实上，数据可能均匀分布于整个时间段，也可能集中在一两年，这会极大影响我们对物种分布或生态学的理解。按年份绘制知识空缺图，有助于探索此类变化。同样，物种的物候会随时间而变化，许多野外调查都在物种繁殖期进行。因此，在这一时段内的知识缺口可能少于一年中的其他时段。最后，在分类学偏差与时间偏差的交叉处，研究表明，博物学家往往关注特定生命阶段；例如，对于蜻蜓目昆虫，采集的成体个体多于蜕壳（Raebel 等人，2010）。这些不同的动态可能对知识缺口的分布及其决定因素产生重大影响。加深对知识缺口复杂模式的理解，将有助于更好地利用开放获取数据。归根结底，要理解生物多样性知识空白的关键

决定因素，就需要更好地理解数据本身存在的知识缺口。

参考文献

Andrade-Silva, J., Baccaro, F. B., Prado, L. P., Guénard, B., Warren, D. L., Kass, J. M., Economo, E. P., & Silva, R. R. (2022). 对巴西亚马逊流域蚂蚁多样性的全面评估：整合地理、生态和形态驱动因素，分析采样偏差。《生态地理学》, *2022*(9), e06295. https://doi.org/10.1111/ecog.06295

生态连续性援助（ACE）新阿基坦。（无日期）。生态景观单元。访问日期：2025年3月26日，网址：https://www.ace-aquitaine.fr/

Baker, D. J., Maclean, I. M. D., & Gaston, K. J. (2024年). 在没有独立测试数据的情况下，校正物种分布模型空间采样偏差的有效策略。《多样性与分布》, *30*(3), e13802. https://doi.org/10.1111/ddi.13802

Ball-Damerow, J. E., Brenskelle, L., Barve, N., Soltis, P. S., Sierwald, P., Bieler, R., LaFrance, R., Ariño, A. H., & Guralnick, R. P.（2019年）。数字时代初级生物多样性数据库的研究应用。《*PLOS ONE*》，*14*（9），e0215794。https://doi.org/10.1371/journal.pone.0215794

Beck, J., Böller, M., Erhardt, A., & Schwanghart, W.（2014）。GBIF数据库中的空间偏差及其对物种地理分布建模的影响。《生态信息学》，*19*，10-15。https://doi.org/10.1016/j.ecoinf.2013.11.002

Boakes, E. H., McGowan, P. J. K., Fuller, R. A., Chang-qing, D., Clark, N. E., O'Connor, K., & Mace, G. M.（2010）。生物多样性的失真图景：物种出现数据中的时空偏差。*PLoS Biology*，*8*（6），e1000385。https://doi.org/10.1371/journal.pbio.1000385

Butchart, S. H. M., Walpole, M., Collen, B., Van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M.,

Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., ... Watson, R. (2010). 全球生物多样性：近期下降指标。《科学》，*328*(5982),1164-1168. https://doi.org/10.1126/science.1187512

Chao, A., & Colwell, R. K. (2017). Chao不等式衍生三十年：利用出现数据和不完整抽样估计并比较丰富度。*SORT*-统计与运筹学汇刊，*41*(1)，3-54。

Chao, A., & Jost, L. (2012). 覆盖度-基础的稀疏化与外推：按完整度而非样本量对样本进行标准化。《生态学》，*93*(12)，2533-2547。https://doi.org/10.1890/11-1952.1

Chao, A., Kubota, Y., Zelený, D., Chiu, C., Li, C., Kusumoto, B., Yasuhara, M., Thorn, S., Wei, C., Costello, M. J., & Colwell, R. K. (2020). 量化样本完整度并比较群落间的多样性。《生态学研究》，*35*(2)，292-314。https://doi.org/10.1111/1440-1703.12102

Cherrier, O., Prima, M.-C., & Rouveyrol, P. (2021). 法国本土陆地人为压力制图—保护地网络诊断目录（第 118 页 + 21 页附录）。PatriNat（法国生物多样性办公室 - 法国国家科学研究中心 - 法国国家自然历史博物馆）。

Chesshire, P. R., Fischer, E. E., Dowdy, N. J., Griswold, T. L., Hughes, A. C., Orr, M. C.,Ascher, J. S., Guzman, L. M., Hung, K. J., Cobb, N. S., & McCabe, L. M. (2023).Completeness analysis for over 3000 United States bee species identifies persistent datagap. 《生态地理学》，*2023*（5），e06584。https://doi.org/10.1111/ecog.06584

Clark, J. A., & May, R. M. (2002). 保护研究中的分类学偏差。《科学》，*297*（5579），191-192。https://doi.org/10.1126/science.297.5579.191b

Collectif.（2021年）。新阿基坦大区生物多样性热点与保护空缺制图（第105页）。国家南大西洋植物学保护中心与鸟类保护组织联盟

（协调）。obv-na.fr/ofsa/ressources/dispositifs/hotspot/Rapport_热点_2021.pdf

Collen, B.、Ram, M.、Zamin, T. 和 McRae, L.（2008）。热带生物多样性数据缺口：应对全球监测中的差异。《热带保护科学》，*1*（2），75-88。https://doi.org/10.1177/194008290800100202

Colli-Silva, M.、Reginato, M.、Cabral, A.、Forzza, R. C.、Pirani, J. R. 和 Vasconcelos, T. N. D. C.（2020）。评估大西洋森林——巴西最具多样性且受威胁最严重的植物地理区——生物多样性记录的缺失及其空间准确性。《分类学》，*69*（3），567-577。https://doi.org/10.1002/tax.12239

Correia, R. A.、Ruete, A.、Stropp, J.、Malhado, A. C. M.、Dos Santos, J. W.、Lessa, T.、Alves, J. A. 和 Ladle, R. J.（2019）。利用无知得分探索卡廷加多个分类群的生物多样性记录力度。《生态指标》，*106*，105539。https://doi.org/10.1016/j.ecolind.2019.105539

De Araujo, M. L., & Ramos, F. N.（2021）。聚焦调查力度：生物多样性热点地区附生植物采样中的缺口与偏差。《森林生态与管理》，*498*，119544。https://doi.org/10.1016/j.foreco.2021.119544

Dennis, R. L. H., & Thomas, C. D.（2000年）。蝴蝶分布图中的偏差：热点地区与记录者的活动范围的影响。《昆虫保护杂志》，*4*(2), 73-77。https://doi.org/10.1023/A:1009690919835

Dolný, A., Ožana, S., Burda, M., & Harabiš, F.（2021年）。景观格局及其变化对蜻蜓目昆虫（昆虫纲）的物种丰富度、物种组成和保护价值的影响。《昆虫》，*12*(6), 478。https://doi.org/10.3390/insects12060478

Doxon, E. D., Davis, C. A., & Fuhlendorf, S. D. (2011年). 两种无脊椎动物抽样方法的比较：吸尘法和扫网采样：抽样方法

无脊椎动物。《野外鸟类学杂志》，*82*(1), 60-67。
https://doi.org/10.1111/j.1557-9263.2010.00308.x

Engemann, K., Enquist, B. J., Sandel, B., Boyle, B., Jørgensen, P. M., Morueta-Holme, N., Peet, R. K., Violle, C., & Svenning, J.（2015）。有限的抽样阻碍了对热带生物多样性热点物种丰富度的"大数据"估算。*Ecology and Evolution*，*5*(3), 807-820.https://doi.org/10.1002/ece3.1405

ESRI.（2024年）。*ArcGISPro*（3.2.0版）[软件]。https://www.esrifrance.fr/fr-fr/homeFisher-Phelps, M., Cao, G., Wilson, R. M., & Kingston, T.（2017）。保护工作中的偏差：跨越时间和生态系统，开源蝙蝠地点数据受到与保护区距离的严重影响。《生态信息学》，*40*，22-34。
https://doi.org/10.1016/j.ecoinf.2017.05.003

Freitas, T. M. S., Montag, L. F. A., De Marco, P., & Hortal, J. (2020). 生物多样性大数据中的物种鉴定有多可靠？评估在线数据库中新热带鱼类科的记录。《系统学与生物多样性》，*18*(2), 181-191.https://doi.org/10.1080/14772000.2020.1730473

Garcia-Rosello, E., Gonzalez-Dacosta, J., Guisande, C., & Lobo, J. M. (2023). GBIF未能如实呈现昆虫的全球分布。《系统昆虫学》, *48*(4), 489-497.
https://doi.org/10.1111/syen.12589

García-Roselló, E., González-Dacosta, J., & Lobo, J. M. (2025年). 地理偏差削弱了欧洲生物多样性数据的环境代表性。《整体环境科学》，*996*，180178。
https://doi.org/10.1016/j.scitotenv.2025.180178

Garcillán, P. P., Ezcurra, E., & Vega, E. (2008). 瓜达卢佩岛：失落的天堂重现？

通过估算过度放牧对偏远海洋岛屿灭绝的影响得出的结论

累积函数估算偏远海洋岛屿上过度放牧对灭绝的影响。《生物多样性与保护》，*17*(7), 1613-1625.

https://doi.org/10.1007/s10531-008-9370-z

Gardiner, M. M., Allee, L. L., Brown, P. M., Losey, J. E., Roy, H. E., & Smyth, R. R. (2012). 从瓢虫研究中获得的启示：美国和英国公民科学项目监测数据的准确性-。《生态学与环境前沿》， *10*(9), 471-476。https://doi.org/10.1890/110185

Gaume, L., & Desquilbet, M.（2024年）。InsectChange：评论。*Peer Community Journal*，*4*，e97。https://doi.org/10.24072/pcjournal.469

GBIF.org（2025年）。*GBIF*主页。https://www.gbif.org/

Girardello, M., Chapman, A., Dennis, R., Kaila, L., Borges, P. A. V., & Santangeli, A.（2019年）。蝴蝶名录数据中的空缺：全球分析。《生物保护》，*236*，289-295。https://doi.org/10.1016/j.biocon.2019.05.053

Girardello, M., Martellos, S., Pardo, A., & Bertolino, S.（2018年）。生物多样性出现信息中的空缺可能阻碍国际生物多样性目标的实现：来自跨分类群分析的见解。《环境保护》，*45*(4), 370-377。https://doi.org/10.1017/S0376892918000115

Gotelli, N. J. 和 Colwell, R. K.（2011年）。估算物种丰富度。载于《生物多样性：测量与评估前沿》（牛津大学出版社，第39-53页）。

Grand, J.、Cummings, M. P.、Rebelo, T. G.、Ricketts, T. H. 和 Neel, M. C.（2007年）。有偏数据降低保护区网络的效率与有效性。《生态学快报》，*10*（5），364-374。https://doi.org/10.1111/j.1461-0248.2007.01025.x

Grassle, F.（2000年）。海洋生物地理信息系统（OBIS）：一个用于在多维地理背景中获取、建模和制图海洋生物学数据的在线全球图集。《海洋学》，*13*（3），5-7。https://doi.org/10.5670/oceanog.2000.01

Grattarola, F., Martínez-Lanfranco, J. A., Botto, G., Naya, D. E., Maneyro, R., Mai, P., Hernández, D., Laufer, G., Ziegler, L., González, E. M., Da Rosa, I., Gobel, N., González, A.,

González, J., Rodales, A. L., & Pincheira-Donoso, D. (2020a). 四足动物生物多样性热点的多种形式及开放获取数据匮乏带来的挑战。《科学报告》， *10*(1), 22045。https://doi.org/10.1038/s41598-020-79074-8

Grattarola, F., Martínez-Lanfranco, J. A., Botto, G., Naya, D. E., Maneyro, R., Mai, P., Hernández, D., Laufer, G., Ziegler, L., González, E. M., Da Rosa, I., Gobel, N., González, A., González, J., Rodales, A. L., & Pincheira-Donoso, D. (2020b). 四足动物生物多样性热点的多种形式及开放获取数据匮乏带来的挑战。《科学报告》， *10*(1), 22045. https://doi.org/10.1038/s41598-020-79074-8

Guénard, B., Hughes, A. C., Lainé, C., Cannicci, S., Russell, B. D., & Williams, G. A. (2025). 全球保护资金有限且分配不均，导致大多数受威胁物种得不到支持。《美国国家科学院院刊》， *122*(9), e2412479122.https://doi.org/10.1073/pnas.2412479122

Guetté, A., Carruthers-Jones, J., & Carver, S. J.（2021年）。CARTNAT项目——自然性制图。技术说明。法国世界自然保护联盟委员会，12。CartNat（法国本土潜在自然性制图）——法国世界自然保护联盟

Hill, M. O.（2012）。在记录投入未知时，局部频率是解读物种出现数据的关键。《生态学与进化方法》， *3*（1），195-205。https://doi.org/10.1111/j.2041-210X.2011.00146.x

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H.（2021年）。采样偏差塑造了我们对自然界的认知。《生态地理学》， *44*（9），1259-1269。https://doi.org/10.1111/ecog.05926

Inman, R., Franklin, J., Esque, T., & Nussear, K.（2021年）。使用虚拟物种比较物种分布建模的样本偏差校正方法。《生态圈》， *12*(3), e03422。https://doi.org/10.1002/ecs2.3422

l国家地理研究所。（2025年）。*BD TOPO*® 版本*3.5*。 [数据集]。

Le, S., Josse, J., & Husson, F. (2008). FactoMineR：用于多变量分析的 R 软件包。《统计软件杂志》， *25*，1-18。 https://doi.org/10.18637/jss.v025.i01

Le Velly, G., Delacote, P., Golden Kroner, R., Keles, D., & Pfaff, A. (2024年). 推动美国减少生物多样性保护工作的政治因素。《生态学与社会》， *29*(3), art27。 https://doi.org/10.5751/ES-15338-290327

Magellium. (2025年). 区域自然公园（*PNR*）—法国本土 [数据集]。

Magnusson, W. E., Grelle, C. E. V., Marques, M. C. M., Rocha, C. F. D., Dias, B., Fontana, C. S., Bergallo, H., Overbeck, G. E., Vale, M. M., Tomas, W. M., Cerqueira, R., Collevatti, R., Pillar, V. D., Malabarba, L. R., Lins-e-Silva, A. C., Neckel-Oliveira, S., Martinelli, B., Akama, A., Rodrigues, D., ... Fernandes, G. W. (2018年). 巴西政治危机对生物多样性保护所需科学研究的影响。《生态学与进化前沿》， *6*，163。 https://doi.org/10.3389/fevo.2018.00163

Mair, L., & Ruete, A.（2016年）。解释多个分类群公民科学数据记录投入的空间变异。《*PLOS ONE*》， *11*(1), e0147796。https://doi.org/10.1371/journal.pone.0147796

Maldonado, C., Molina, C. I., Zizka, A., Persson, C., Taylor, C. M., Albán, J., Chilquillo, E., Rønsted, N., & Antonelli, A.（2015）。大数据时代如何估算物种多样性和分布：我们在多大程度上可以信赖公共数据库？《全球生态学与生物地理学》， *24*(8), 973-984。 https://doi.org/10.1111/geb.12326

Marcer, A., Chapman, A. D., Wieczorek, J. R., Xavier Picó, F., Uribe, F., Waller, J., & Ariño, A. H. (2022). 不确定性至关重要：确定自然历史收藏中标本的来源及其对预测物种分布的影响。《生态地理学》， *2022*(9),e06025. https://doi.org/10.1111/ecog.06025

Martin, L. J., Blossey, B., & Ellis, E. (2012). 绘制生态学家的工作地点：陆地生态观测全球分布中的偏差。《生态学与环境前沿》， *10*(4), 195-201. https://doi.org/10.1890/110154

Martín-López, B., Montes, C., Ramírez, L., & Benayas, J. (2009年). 什么驱动与物种保护相关的政策决策？ 《生物保护》, *142*(7),1370-1380. https://doi.org/10.1016/j.biocon.2009.01.030

Mentges, A., Blowes, S. A., Hodapp, D., Hillebrand, H., & Chase, J. M. (2021年). 样地-选择偏差对生物多样性变化估计的影响。 《保护生物学》, *35*(2), 688-698.https://doi.org/10.1111/cobi.13610

Meyer, C., Kreft, H., Guralnick, R., & Jetz, W. (2015). 生物多样性分布有效信息基础的全球优先事项。 《自然·通讯》, *6*(1), 8221.https://doi.org/10.1038/ncomms9221

Minteer, B. A., Collins, J. P., Love, K. E., & Puschendorf, R. (2014). 避免（再度）灭绝。《科学》， *344*(6181), 260-261. https://doi.org/10.1126/science.1250953

Moilanen, A. (2007). 景观分区、效益函数与基于目标的规划：统一保护区遴选策略。《生物保护》， *134*(4), 571-579.https://doi.org/10.1016/j.biocon.2006.09.008

法国国家自然历史博物馆。（2025年）*INPN*。https://inpn.mnhn.fr/

OBIS,（海洋生物多样性信息系统）。（2025年）。生物多样性摘要。

https://obis.org/indicators/documentation/

FAUNA观测站。（2025a）。*FAUNA*观测站。https://observatoire-fauna.fr/

FAUNA观测站。（2025b）。新阿基坦地区 *I* 类和 *II* 类 *ZNIEFF* [数据集]。

O'Farrell, M. J., & Gannon, W. L. (1999). 声学技术与捕捉技术的比较

蝙蝠调查。《哺乳动物学杂志》， *80*(1)，24-30。

https://doi.org/10.2307/1383204

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O' Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., ... Borman, T.（2025年）。*vegan*：群落生态学软件包。*R* 软件包版本 *2.6-10* [软件]。https://CRAN.R-project.org/package=vegan

Oliveira, U., Paglia, A. P., Brescovit, A. D., De Carvalho, C. J. B., Silva, D. P., Rezende, D. T.,Leite, F. S. F., Batista, J. A. N., Barbosa, J. P. P. P., Stehmann, J. R., Ascher, J. S., DeVasconcelos, M. F., De Marco, P., Löwenberg-Neto, P., Dias, P. G., Ferro, V. G., & Santos, A. J.（2016年）。采集偏差对巴西陆地生物多样性知识缺口的强烈影响。《多样性与分布》，*22*(12), 1232-1244。https://doi.org/10.1111/ddi.12489

Panter, C. T., Clegg, R. L., Moat, J., Bachman, S. P., Klitgård, B. B., & White, R. L.（2020）。清理还是不清理：清理开放-源数据可改善受威胁植物物种的灭绝风险评估。《保护科学与实践》，*2*(12), e311。https://doi.org/10.1111/csp2.311

法国国家公园。（2025年）。法国*11*个国家公园边界 [数据集]。

Petersen, T. K., Speed, J. D. M., Grøtan, V., & Austrheim, G. (2021年)。用于理解生物多样性动态的物种数据：物种出现数据收集的内容、地点和时间。《生态学解决方案与证据》，*2*(1), e12048。https://doi.org/10.1002/2688-8319.12048

Phaka, F. M., Vanhove, M. P. M., Du Preez, L. H., & Hugé, J.（2022）。回顾超高生物多样性国家的分类学偏差：南非动物的原始生物多样性数据、文化显著性和科学关注度。《环境评论》，*30*（1），39-49。https://doi.org/10.1139/er-2020-0092

Poncet, L.（2013）。法国生物多样性信息的传播：国家自然遗产清查（INPN）的案例。*Netcom*，*27-1/2*，181-189。https://doi.org/10.4000/netcom.1306

Prendergast, J. R., & Eversham, B. C. (1997). 高阶分类群中的物种丰富度协变：生物多样性指标概念的实证检验。《生态地理学》，*20*(2), 210-216.https://doi.org/10.1111/j.1600-0587.1997.tb00363.x

R 核心团队。（2025年）。*R*：一种用于统计计算的语言和环境（版本 2024.12.1）[软件]。R 统计计算基金会。www.R-project.org/

Raebel, E. M., Merckx, T., Riordan, P., Macdonald, D. W., & Thompson, D. J. (2010). 蜻蜓谬误：为何必须采集蜕壳以避免有偏调查。《昆虫保护杂志》，*14*(5), 523-533。https://doi.org/10.1007/s10841-010-9281-7

Rands, M. R. W., & Sotherton, N. W.（1986）。英格兰耕地上谷类作物的农药使用与蝴蝶数量变化。《生物保护》，*36*(1), 71-82。https://doi.org/10.1016/0006-3207(86)90102-3

Reichman, O. J., Jones, M. B., & Schildhauer, M. P.（2011）。生态学开放数据的挑战与机遇。《科学》，*331*(6018), 703-705。https://doi.org/10.1126/science.1197962

Ribeiro, I., Proença, V., Serra, P., Palma, J., Domingo-Marimon, C., Pons, X., & Domingos, T.（2019年）。利用遥感指标和开放获取的生物多样性数据评估地中海乡村景观中的鸟类多样性格局。《科学报告》，*9*(1), 6826。https://doi.org/10.1038/s41598-019-43330-3

Robert, S. (2021年). SINP 2020 年总结. 法国国家自然历史博物馆 - *HAL Science*.

Rocha-Ortega, M., Rodriguez, P., & Córdoba-Aguilar, A. (2021). GBIF 昆虫数据中生物多样性与灭绝的地理、时间和分类学偏差。《生态昆虫学》，*46*(4), 718-728. https://doi.org/10.1111/een.13027

Rodríguez-Rey, M., & Grenouillet, G. (2022). 欧洲淡水鱼类采样偏差的驱动因素解析。*Fishes*，*7*(6), 383. https://doi.org/10.3390/fishes7060383Ronquillo, C., Alves-Martins, F., Mazimpaka, V., Sobral-Souza, T., Vilela-Silva, B., G. Medina, N., & Hortal, J. (2020). 评估伊比利亚苔藓公开可得分布信息中的时空偏差与缺口。*Biodiversity Data Journal*，*8*, e53474。https://doi.org/10.3897/BDJ.8.e53474

Ruete, A.（2015）。利用认知空白地图展示从生物多样性数据库获取的数据在采样力度上的偏差。*Biodiversity Data Journal*，*3*，e5361。https://doi.org/10.3897/BDJ.3.e5361

Safe Software。（2025）。*FME Software* [软件]。http://www.safe.com/

Sánchez-Bayo, F.（2021）。农药对昆虫及其他节肢动物的间接影响。*Toxics*，*9*（8），177。https://doi.org/10.3390/toxics9080177

Sánchez-Fernández, D., Lobo, J. M., Abellán, P., Ribera, I., & Millán, A.（2008）。淡水生物多样性调查中的偏差：以伊比利亚水甲虫为例。《多样性与分布》，*14*（5），754-762。https://doi.org/10.1111/j.1472-4642.2008.00474.x

Sánchez-Fernández, D., Yela, J. L., Acosta, R., Bonada, N., García-Barros, E., Guisande, C.,Heine, J., Millán, A., Munguira, M. L., Romo, H., Zamora-Muñoz, C., & Lobo, J. M.（2022）。采样力度和清单完整度的模式是否一致？使用伊比利亚半岛五个昆虫分类群数据库进行检验。*Insect Conservation and Diversity*，*15*（4），406-415。https://doi.org/10.1111/icad.12566

Sittenthaler, M., Fischer, I., Chovanec, A., Koblmüller, S., Macek, O., Sattmann, H., Szucsich, N., Zangl, L., & Haring, E. (2023). 对中欧豆娘和蜻蜓（昆虫纲：蜻蜓目）的蜕壳进行 DNA 条形码分析，以鉴定物种。《昆虫保护杂志》，*27*(3), 435-450。https://doi.org/10.1007/s10841-023-00467-x

Šmíd, J. (2022a). 脊椎动物生命之树中的地理和分类偏差。*Journal ofBiogeography*, *49*(12), 2120-2129。https://doi.org/10.1111/jbi.14491

Šmíd, J.（2022b）。脊椎动物生命之树中的地理与分类学偏差。*Journal ofBiogeography*, *49*（12）， 2120-2129。https://doi.org/10.1111/jbi.14491

Sobral-Souza, T.、Bosco, N. S.、Candelária, L. P.、Collevatti, R. G.、Layme, V. M. G. 和 Rodrigues, D. D. J.（2024年）。大西洋森林小型啮齿动物抽样中的空间偏差：景观与可达性视角。《生态学与保护展望》， *22*（3）， 297-305。https://doi.org/10.1016/j.pecon.2024.07.004

Sousa-Baena, M. S.、Garcia, L. C. 和 Peterson, A. T.（2014）。巴西植物数字化可获取知识的完整度，以及调查和清查的优先事项。《多样性与分布》， *20*（4）， 369-381。https://doi.org/10.1111/ddi.12136

Steinberg, P. F. (2009年). 政治变革中的制度韧性：生物多样性保护案例。 《全球环境政治》, *9*(3), 61-81.https://doi.org/10.1162/glep.2009.9.3.61

Stropp, J., Ladle, R. J., M. Malhado, A. C., Hortal, J., Gaffuri, J., H. Temperley, W., Olav Skøien, J., & Mayaux, P. (2016年). 无知制图：非洲开花植物采集的300年。 《全球生态学与生物地理学》, *25*(9), 1085-1096.https://doi.org/10.1111/geb.12468

Syfert, M. M., Smith, M. J., & Coomes, D. A. (2013年). 采样偏差与模型复杂度对MaxEnt物种分布模型预测性能的影响。 《*PLoS ONE*》, *8*(2), e55158. https://doi.org/10.1371/journal.pone.0055158

Titley, M. A., Snaddon, J. L., & Turner, E. C. (2017). 关于动物生物多样性的科学研究系统性地偏重脊椎动物和温带地区。《*PLOS ONE*》， *12*(12), e0189577。https://doi.org/10.1371/journal.pone.0189577

Touroult, J., Grégoire, L., Régnier, C., Gigot, G., & Barnier, F. (2020). 自然主义知识悖论：物种名录不断变长，生物多样性却日渐衰退。*Bourgogne-Franche-Comté Nature*，《第六次物种灭绝。接下来呢？》(31), 259-264。

Troia, M. J., & McManamay, R. A.（2016年）。填补 GAPS：评估美国开放-获取生物多样性数据库的完整度和覆盖范围。*Ecology and Evolution*， *6*(14), 4654-4669。https://doi.org/10.1002/ece3.2225

Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F.（2017）。生物多样性数据中的分类学偏差与社会偏好。《科学报告》， *7*(1), 9132。https://doi.org/10.1038/s41598-017-09084-6

Ugland, K. I., Gray, J. S., & Ellingsen, K. E.（2003）。物种累积曲线与物种丰富度的估算。*Journal of Animal Ecology*， *72*(5), 888-897。https://doi.org/10.1046/j.1365-2656.2003.00748.x

Van Klink, R., Bowler, D. E., Gongalsky, K. B., Swengel, A. B., Gentile, A., & Chase, J. M. (2020). 荟萃分析揭示陆地昆虫丰度下降，而淡水昆虫丰度上升。《科学》，*368*(6489), 417-420. https://doi.org/10.1126/science.aax9931

Vimont, M., Bartholus, L., Bas, Y., Fontaine, B., Fontaine, C., Julliard, R., Loïs, G., Lorrillière, R., Martin, G., & Porcher, E. (2025). 基于公民科学数据常规编制种群趋势：法国常见鸟类和植物物种动态的见解。*Comptes Rendus. Biologies*， *348*(G1), 229-247. https://doi.org/10.5802/crbiol.184

Yang, W., Ma, K., & Kreft, H. (2013). 大型分布数据库中的地理采样偏差及其对物种丰富度—环境模型的影响。*Journal of Biogeography*， *40*(8), 1415-1426. https://doi.org/10.1111/jbi.12108

Zizka, A., Antunes Carvalho, F., Calvente, A., Rocio Baez-Lizarazo, M., Cabral, A., Coelho, J. F. R., Colli-Silva, M., Fantinati, M. R., Fernandes, M. F., Ferreira-Araújo, T., Gondim Lambert

Moreira, F., Santos, N. M. C., Santos, T. A. B., Dos Santos-Costa, R. C., Serrano, F. C., Alves Da Silva, A. P., De Souza Soares, A., Cavalcante De Souza, P. G., Calisto Tomaz, E., ... Antonelli, A. (2020). 清理 GBIF 没有放之四海皆准的解决方案。*PeerJ*，*8*，e9916。
https://doi.org/10.7717/peerj.9916

声明

资助
本研究由新阿基坦大区环境、规划与住房区域局通过资助 US FAUNA 提供支持。

利益冲突
作者没有需要披露的相关财务或非财务利益。

数据可用性
本研究使用的数据包含敏感物种出现数据，具体遵循自然遗产清查信息系统 (SINP) 的指南。除非经确认身份的个人提出有充分理由的申请，否则不得公开这些数据。因此，本研究使用的数据仅可依申请提供。

图注

图1 新阿基坦大区地图，显示其原组成地区和主要生物地理单元。

图 2新阿基坦地区脊椎动物、无脊椎动物及所有分类群的知识空间分布和完整度得分，采用 10 千米 × 10 千米网格单元表示。完整度得分可能高于 1，因为已观测物种数可能超过预期物种数。该梯度可显示出现这种情况的区域，以及采样最少区域。另一方面，知识得分的范围限定在 0 至 1 之间。

图3 基于856个网格单元矩阵，对各分类群的完整度（图3a）和知识得分（图3b）进行主成分分析（PCA）双标图，按第二维度（a）从高到低排列：蝙蝠、陆生鸟类、两栖动物、爬行动物、陆生哺乳动物、日行性蝴蝶、直翅目、蜻蜓目昆虫。向量箭头的方向和长度表示各分类群得分对主成分的载荷。椭圆基于质心和90%置信区间，反映每个前亚区域内网格单元的分布。

图 4 新阿基坦856个网格单元与知识缺口的7个决定因素矩阵的主成分分析（PCA）双标图，沿第一维度从左至右依次为：城市化压力（建筑物）、道路密度（道路）、农业压力（拖拉机）、到访压力（足迹）、热点（奖章）、自然度（叶子）和保护区面积（手）。向量箭头的方向和长度表示载荷

各决定因素对主成分的载荷。椭圆表示各前亚区域内网格单元的分布，使用质心和90%置信区间绘制。

图 5 冗余分析（RDA）展示了知识缺口的决定因素对脊椎动物、无脊椎动物及所有分类群的完整度和知识得分的贡献。向量箭头的方向和长度表示知识缺口得分与决定因素之间相关性的强弱。

图 6 各分类群和亚门的知识缺口得分与决定因素之间线性回归模型的结果。此处仅显示 p值 <0.05 的变量。圆圈大小表示关系的强弱。粉色圆圈表示正相关关系，白色圆圈表示负相关关系。

图 S1 使用完整性指标展示各分类群知识缺口的分布。

图 S2 使用知识指标展示各分类群的知识缺口分布。

补充材料

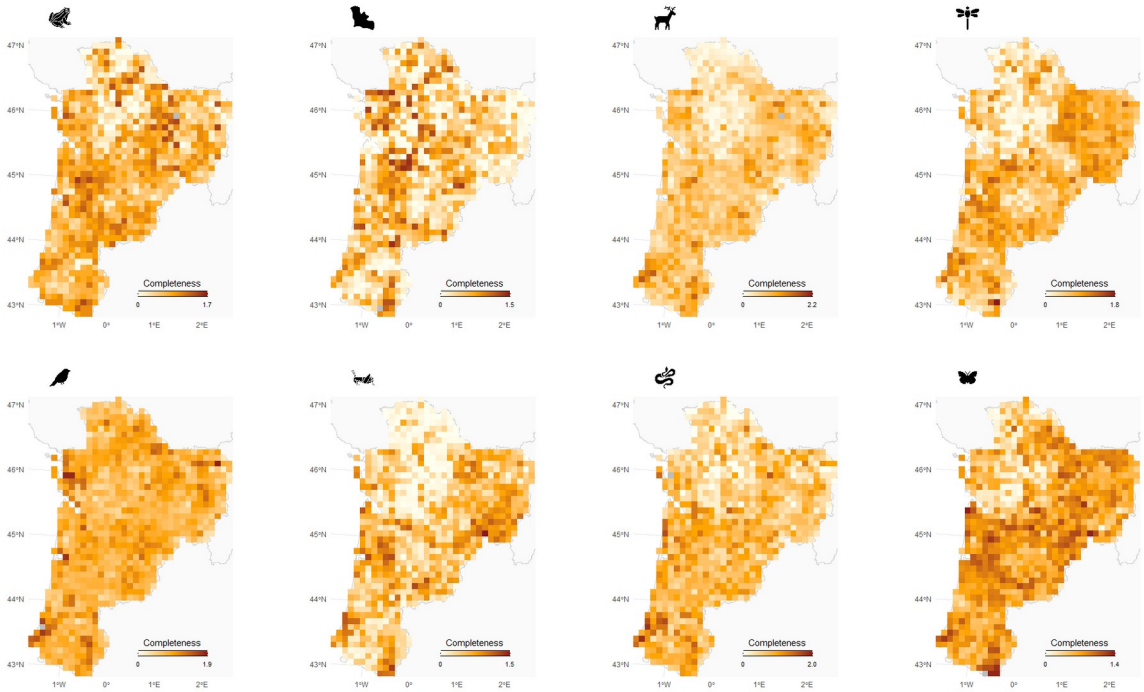


图 S1 使用完整度指标展示各分类群知识缺口的分布指标。

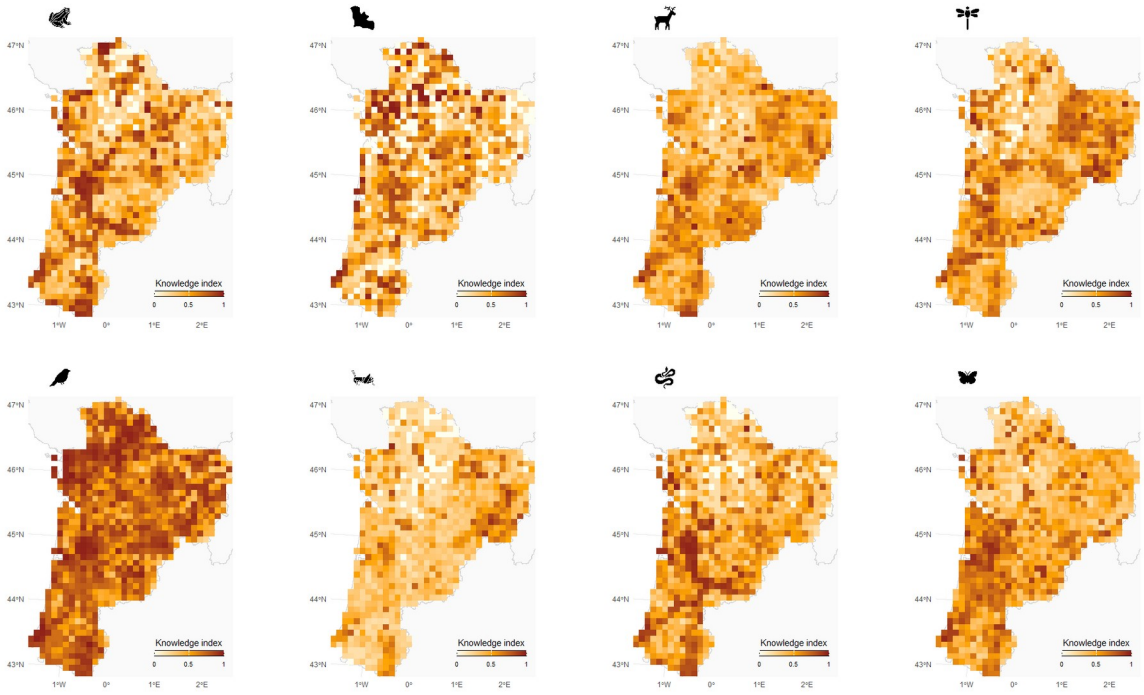


图 S2 使用知识

指标。