arXiv:2304.02632v1 [stat.AP] 5 Apr 2023

用于地块至景观尺度存量变化评估的历史森林生物量制图

预印本

Lucas K Johnson
环境科学研究生项目
纽约州立大学环境科学与林业学院
锡拉丘兹，13210
ljohns11@esf.edu

迈克尔·J·马奥尼
环境科学研究生项目
纽约州立大学环境科学与林业学院
锡拉丘兹，13210
mjmahone@esf.edu

MadeleineL Desrochers纽约州立大学环境科学与林业学院可持续资源管理系 锡拉丘兹, 13210mldesroc@syr.edu

ColinM Beier纽约州立大学环境科学与林业学院可持续资源管理系 锡拉丘兹, 13210cbeier@esf.edu

2023-04-04

摘要

了解历史森林动态，特别是森林生物量和碳储量的变化，对于评估当前森林气候效益以及在各种政策、监管和管护情景下预测未来效益已变得至关重要。仅基于国家森林清查的碳核算框架仅限于大范围估算，但将这些清查与遥感数据相结合的基于模型的方法，可生成景观范围内随时间变化的连续高分辨率森林生物量和碳储量地图。在此，我们描述构建基于地图的储量变化框架的基本步骤：利用免费数据和开源工具，以高时间和空间分辨率（年度、30米）绘制纽约州（美国）1990年至2019年的历史森林生物量。

利用Landsat影像、美国林务局森林清查与分析（FIA）数据以及现成的激光雷达数据集，我们开发了三种用于绘制历史森林地上生物量（AGB）分布的建模方法：以FIA样地级地上生物量估算值为训练数据（直接）、以激光雷达衍生的地上生物量地图为训练数据（间接），以及对直接模型和间接模型的预测结果进行集成平均。模型预测表面（地图）在多个尺度上与FIA估算值进行了比较。三种方法均产生了可行的结果，但在模型复杂度方面存在明显权衡，

地图精度、饱和效应和精细尺度模式表征。由此产生的地图产品有助于确定森林碳储量在何处、何时以及如何因人为和自然驱动因素共同作用而发生变化。因此，这些产品可作为广泛应用的输入，包括存量变化评估、监测、报告与核查框架，以及确定优先保护地块或将其纳入改进管理计划。

关键词 Landsat • 激光雷达 • 地上生物量 • 机器学习 • 土地覆盖

1 引言

森林是最有效的天然碳汇之一，因此对稳定地球气候至关重要，但其提供这一关键服务的能力受到过去和当前人为影响的强烈塑造。随着政策制定者和利益相关者寻求基于自然的解决方案来降低大气温室气体（GHG）浓度并减缓气候变化，理解森林碳相对于人类活动的时空动态变得日益重要（Malmsheimer 等，2008；Fargione 等，2018）。通过更好地把握森林景观中的当地社会和生态条件，决策者可以识别并优先考虑适合不同策略（如重新造林、避免土地转换或强化森林经营）的地块，以维持和/或增加碳封存，并有效抵消其他部门的温室气体排放（Houghton，2005；Houghton 等，2012）。为量化潜在气候效益，通常使用存量变化方法评估碳状况和趋势，这需要历史数据以及通过永久样地网络进行的持续监测工作。

美国农业部的森林清查与分析（FIA）计划等国家森林清查（NFI）依托广泛的抽样设计，提供大范围的森林生物量、碳储量和存量变化估算。过去三十年来，这些计划为认识森林碳动态提供了基础性见解和关键数据（Buendia 等，2019；Woodall 等，2015），但其空间覆盖受样本密度和复测频率限制（McRoberts，2011），因此无法呈现与规划和决策最为相关的精细尺度格局与动态。基于模型的方法将野外数据（如 FIA 数据）与全覆盖遥感数据相结合，可为特定区域内的所有地图单元（像元）生成预测，从而满足这一需求。

可用数据的局限在很大程度上使得采用基于模型的方法来表征森林碳历史时空动态仍充满挑战。遥感数据最能描述森林中最显著的地上部分，因此，地上生物量（AGB）通常被用作初始目标变量（Houghton、Hall 和 Goetz，2009），随后再通过经验转换估算特定碳库（Heath 等，2009；Woodall 等，2011）。机载激光雷达已成为此类地上生物量制图极具价值的遥感数据源，但其采集范围通常是边界不规则的局部至区域尺度；汇集用于大范围应用时，会形成时空上的拼接斑块（Johnson 等，2022；Huang 等，2019；Skowronski 和 Lister，2012）。与机载激光雷达点云相比，遥感光学影像的空间覆盖范围更广、时间一致性更好，但无法以同等细节水平或空间分辨率表征森林结构。光学数据集仍然是目前可获取的最佳历史地表观测数据；尤其是 Landsat 计划在过去四十年间提供了30米分辨率的光谱信息，支持了广泛的历史时间序列制图工作（Hansen 和 Loveland，2012；Banskota 等，2014；Wulder 等，2022）。较新的星载遥感任务利用激光雷达和合成孔径雷达（SAR）采集数据，或可在同样广阔的尺度上量化森林结构，但这些平台无法实现 Landsat 所具备的历史连续性（Dubayah 等，2014；Abdalati 等，2010；Torres 等，2012；Rosenqvist 等，2007）。

少数研究利用Landsat时间序列影像开展多年、高分辨率、大范围的地上生物量制图（Kennedy、Ohmann 等人，2018年；Matasci等，2018年；Hudak 等人，2020年）。这些研究可分为“直接”方法和“间接”方法：前者使用 FIA 野外样地的地上生物量测量值拟合模型（Kennedy、Ohmann 等人，2018年）；后者则使用以激光雷达数据训练的独立模型所生成的地上生物量预测值拟合模型（Matasci等，2018年；Hudak 等人，2020年）。与间接方法相比，直接方法具有一定的简约性，并能减少误差在多阶段建模过程中的传播。由于可利用由基于激光雷达的预测值（像元）构成的更大模型训练样本，间接方法有望产生更准确的预测。理论上，与传统野外样地相比，基于激光雷达的预测值样本覆盖的地上生物量条件范围更广，地理定位精度更高，与 Landsat 像元的空间兼容性也更好（Hudak 等人，2020年）。这两种主要方法（直接和间接）仅在时间快照（单年制图）条件下进行过比较，研究范围包括阿拉斯加北方森林中一片相对较小（820,000公顷）且同质的区域（Strunk 等人，2014年），以及利用墨西哥国家森林清查并加入合成孔径雷达数据的墨西哥地区（Urbazaev等，2018年）。

作为纽约州（NYS）基于地图开展森林碳核算这一更广泛工作的组成部分，本文介绍了将 FIA 的离散样地清查数据转化为覆盖全州、时间跨度为30年（1990—2019年）的年度地上生物量地图的方法，地图分辨率为30米。由此生成的地图产品提供了必要数据，可用于以空间显式的方式复现 FIA 的存量变化核算方法，并灵活生成从单个地块到全州范围不同尺度的成果。我们开发的模型实现了上述目标；据我们所知，这是首次尝试综合直接和间接方法。我们利用 Landsat时间序列影像、FIA样地以及公开可获取的现成激光雷达数据，构建了结合这两种不同建模策略（直接和间接）的集成方法，发挥各自优势并提高整体方法的预测精度。我们在多个尺度上评估了三种方法（直接、间接和集成）的制图预测与一组独立 FIA 估算值之间的一致性。这些基于公开数据和开源工具的方法灵活、高效，并可扩展至不同空间和时间范围，因此为有意开展森林地上生物量动态制图、用于回溯分析和监测的人士提供了一个框架。按照该框架得出的结果不仅为管理相关尺度上的存量变化分析提供数据，还将广泛支持森林管护、未来研究和持续规划。

2 数据与方法

2.1 概述

我们开发了三种建模方法（图 1），用于绘制纽约州（NYS）范围内地上生物量（AGB）的年度分布图。直接方法将美国农业部森林清查与分析（FIA；Grayet al. (2012)）野外样地中的AGB估算值作为因变量。间接方法将Johnsonetal.(2022)开发的基于激光雷达的AGB预测值作为因变量。对于这两种方法，分别将因变量与从时间匹配的Landsat影像和土地覆盖分类中提取的预测变量，以及时间上保持不变的气候、地形和生态图层相关联。我们使用这些组合数据集分别构建了堆叠集成模型，每个模型均由多个机器学习（ML）模型组成。我们对这两种方法的预测结果取平均值，以创建第三种集成方法。三种建模方法均用于以30米分辨率对整个纽约州进行年度（1990-2019年）AGB预测，并使用一组通用的独立FIA样地对所得地图进行评估。

2.2 研究区域

纽约州位于美国东北部，面积为141,297 km^2，截至2019年约59%为森林覆盖（USFS 2020）。森林以北部硬木-铁杉类型为主，但在该州西部和南部分别包括阿巴拉契亚橡树和山毛榉-枫-椴树林（Dyer 2006）。与美国东北部大部分地区一样，纽约州在18和19世纪被广泛砍伐，随后重新造林和保护导致景观以现已超过100年的林分为主（Whitney 1994; Lorimer 2001; Michael J. Mahoney 等人 2022）。纽约州于1885年创建了森林保护区，为几十年后成为阿迪朗达克和卡茨基尔公园奠定了基础。这些公园内的任何州有或收购土地此后被指定为“永远野生”，并在很大程度上受到保护免受木材采伐。更近期的土地利用动态表明，该州总农业用地持续下降，并被类似范围的森林和开发用地所取代（Widmann et al. 2012; Widmann 2016）。总森林面积估计在2012年达到峰值，由于持续人类开发导致的森林损失最近已超过因农业废弃而获得的增益（Widmann 2016; USFS 2020）。采伐活动、天气相关事件和虫害爆发驱动持续森林区域内的干扰和损害模式（Kosiba et al. 2018; USFS 2020）。

2.3 野外数据

为模型开发和地图评估这两个不同目的，分别从纽约州的FIA清查中汇编了两套野外数据集。FIA项目汇编了胸高直径为 $\geq$ 12.7厘米（5英寸）的树木的地上生物量估算值（Gray等，2012），并将其转换为兆克/公顷（Mg ha^{-1}）单位。FIA采用永久性清查样地，这些样地布设于准系统性的六边形网格中，并划分为五个面板；假定每个面板均完整覆盖全州范围，并按照5—7年的周期重新测量（Bechtold和Patterson，2005年）。树木测量数据以及随后基于异速生长方程得出的地上生物量估算值，仅记录于被视为森林的样地部分。根据FIA的定义，一个区域要被视为森林，必须至少有10%的树木蓄积度，面积至少为0.4公顷（1英亩），且宽度至少为36.58米（120英尺）。任何满足这些最低要求、但被开发用于非森林土地利用的土地，均不被视为森林。根据这一定义，某些非森林状况很可能包含未被测量的地上生物量。然而，在缺乏其他信息的情况下，我们假定任何非森林状况均代表0地上生物量。

FIA样地由四个相同的圆形子样地组成，半径为7.32 m（24 ft），其中一个子样地以宏样地质心为中心，另外三个子样地位于36.6 m（120 ft）外，方位角分别为360°、120°和240°（Bechtold and Patterson 2005年）。样地位置由FIA项目以平均坐标的形式提供，收集

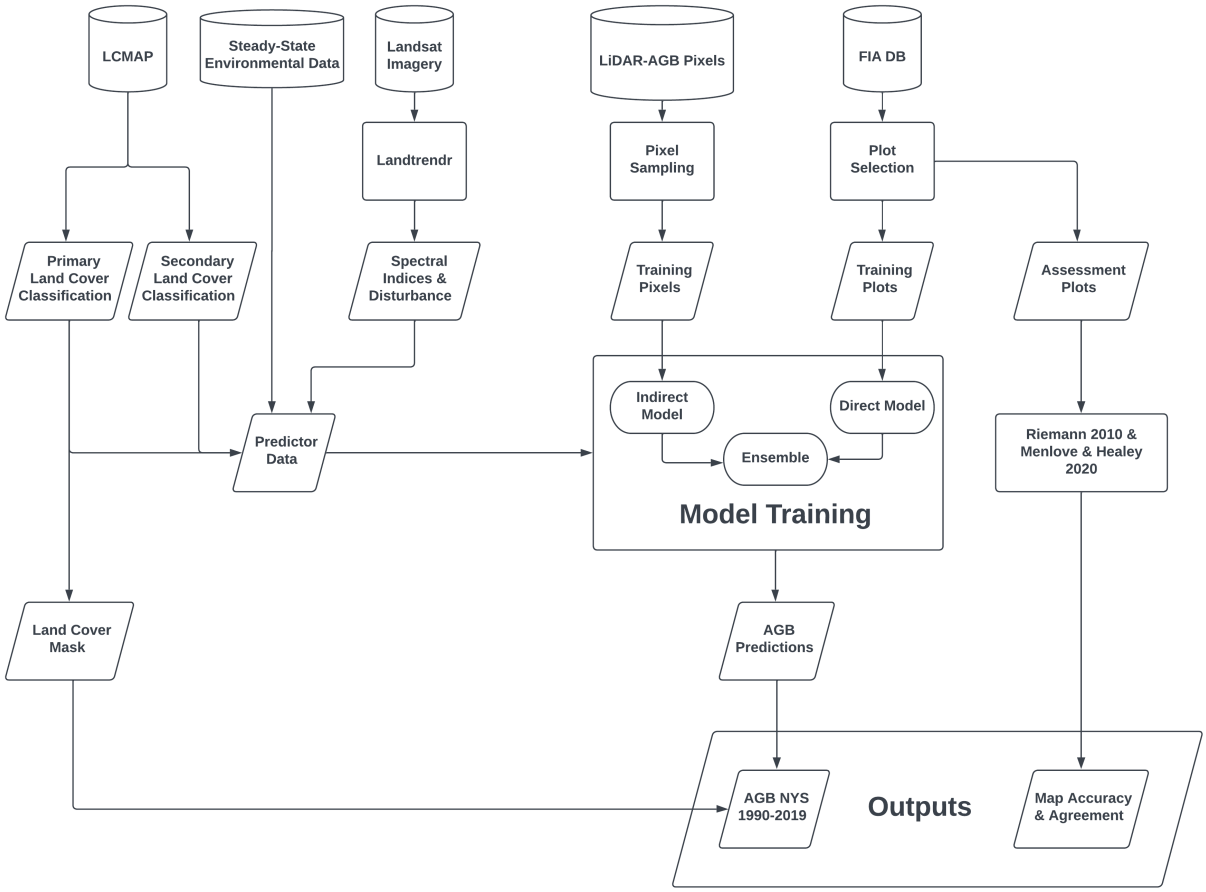


图1：展示建模与制图方法关键要素的流程图。圆柱体表示数据存储库，平行四边形表示数据产品和结果，矩形表示处理步骤，椭圆表示模型。

经过多次重复访问，代表中心样地的质心，随后我们利用这些坐标构建了一个多边形数据集，用于表示包括全部四个子样地在内的整个样地布局。由于最初获取的宏样地质心 GPS 坐标精度不足，因此有必要对坐标进行平均（Hoppus 和 Lister，2005年；Cooke，2000年）。我们使用“FIA 样地”一词来指代四个子样地的集合。

我们仅考虑了按照国家样地设计布设、且所有子样地均被标记为已测量的FIA样地。重要的是，排除未测量样地并不会使FIA的概率抽样失效，因为FIA项目假定这些样地在整个景观中呈随机分布（BechtoldandPatterson 2005年）。此外，对于可用样地被清查多次的情况，我们随机选取其中一次，以避免重复。上述初始筛选标准最终得到5,144个样地，这些样地的清查时间介于2002年和2019年之间。随后，我们利用FIA的面板划分将这组样地分为模型开发数据集和地图评估数据集：随机选取五个面板中的一个，并将所有具有该面板标识的样地分配到地图评估数据集，其余样地则分配到模型开发数据集。通过这种方式，我们将20%的可用样地数据划分出来，用于独立的地图评估，从而获得了一个具有完整空间覆盖的概率样本，并利用该样本生成地图一致性指标的无偏估计（Stehman andFoody 2019年；Riemann et al. 2010年）。

对于模型开发数据集，我们进一步选择了1,954个完全森林化的样地，以确保非森林条件下的无响应不会降低预测变量与样地级地上生物量之间的关系。然而，为了用覆盖最广泛可能条件范围的信息来训练和测试我们的模型，我们添加了一组95个完全非森林化的样地，这些样地根据激光雷达导出的最大高度被识别为真正的零值（地上生物量）≤ 1 m（Johnson et al. 2022）。模型开发数据集包含2,049个独特样地（表1）。对于地图评估数据集，我们根据土地覆盖掩膜（第2.7节）过滤了我们映射区域之外的样地，因为这些样地被认为在我们感兴趣的总体之外，结果共有545个样地（表1）。

表1：划分为模型开发和制图评估数据集的FIA样地年度数量。

Year	模型开发	地图评估
2002	172	
2003	188	
2004	98	
2005	106	
2006	157	
2007		207
2008	165	
2009	146	
2010	153	
2011	174	
2012		191
2013	138	
2014	156	
2015	119	
2016	96	
2017	129	
2018	19	96
2019	33	51
总计	**2049**	**545**

2.4 激光雷达数据和激光雷达像元采样

对于我们的间接建模方法，我们使用现有的基于激光雷达的地上生物量预测表面作为模型训练的参考数据（图2）。Johnson et al. (2022)利用覆盖纽约州62.46%（7,835,690公顷）的17个落叶期激光雷达数据集的时空拼接开发了这些30米表面。激光雷达数据从700–5300米的高度采集，脉冲密度范围为每平方米1.54–3.24个脉冲^{2}。从高度归一化点云计算的一组40个预测变量，结合地形、气候、土地覆盖和地籍数据，与FIA样地共置作为模型训练数据。使用机器学习模型的堆叠集成（Wolpert 1992年）在拼接上做出预测；更多细节可在Johnson et al. (2022)中找到。

遵循Johnson et al.(2022)，我们使用基于LCMAP主要分类（Brown等人 2020年；Zhu和Woodcock 2014年）的植被掩膜以及适用区域掩膜（Meyer和Pebesma 2021）限制制图空间。因此，我们的基于LiDAR的地上生物量预测样本仅限于植被覆盖的景观，以及预测所依据的预测因子数据在训练数据中得到充分代表的区域。遵循Hudak等（2020年）所述的间接建模工作，我们从基于LiDAR的地上生物量预测中进行了分层随机抽样，其中各层被定义为从0到制图最大地上生物量值（~330 Mgha^{-1}）范围内的20个等间隔区间。每个层抽取1,000个像元，共得到20,000个具有空间分辨率的地上生物量预测值。

2.5 Landsat 和辅助数据

我们通过在 Google Earth Engine（GEE，Gorelick 等人（2017年））中处理 Landsat Collection 1 数据（C1，美国地质调查局（2018年）），生成了一组包含 16 个年度 Landsat 衍生预测变量的数据集。我们遵循 Michael J. Mahoney 等人（2022年）描述的处理框架，使用基于 Roy 等人（2016年）的系数处理的生长季中值合成影像，以及 GEE 中的 LandTrendr 实现（下称 LT-GEE），生成连续且平滑的 30 年时间序列，其中包含用于描述地表状况和扰动历史的像元级指标（Kennedy、Yang 和 Cohen

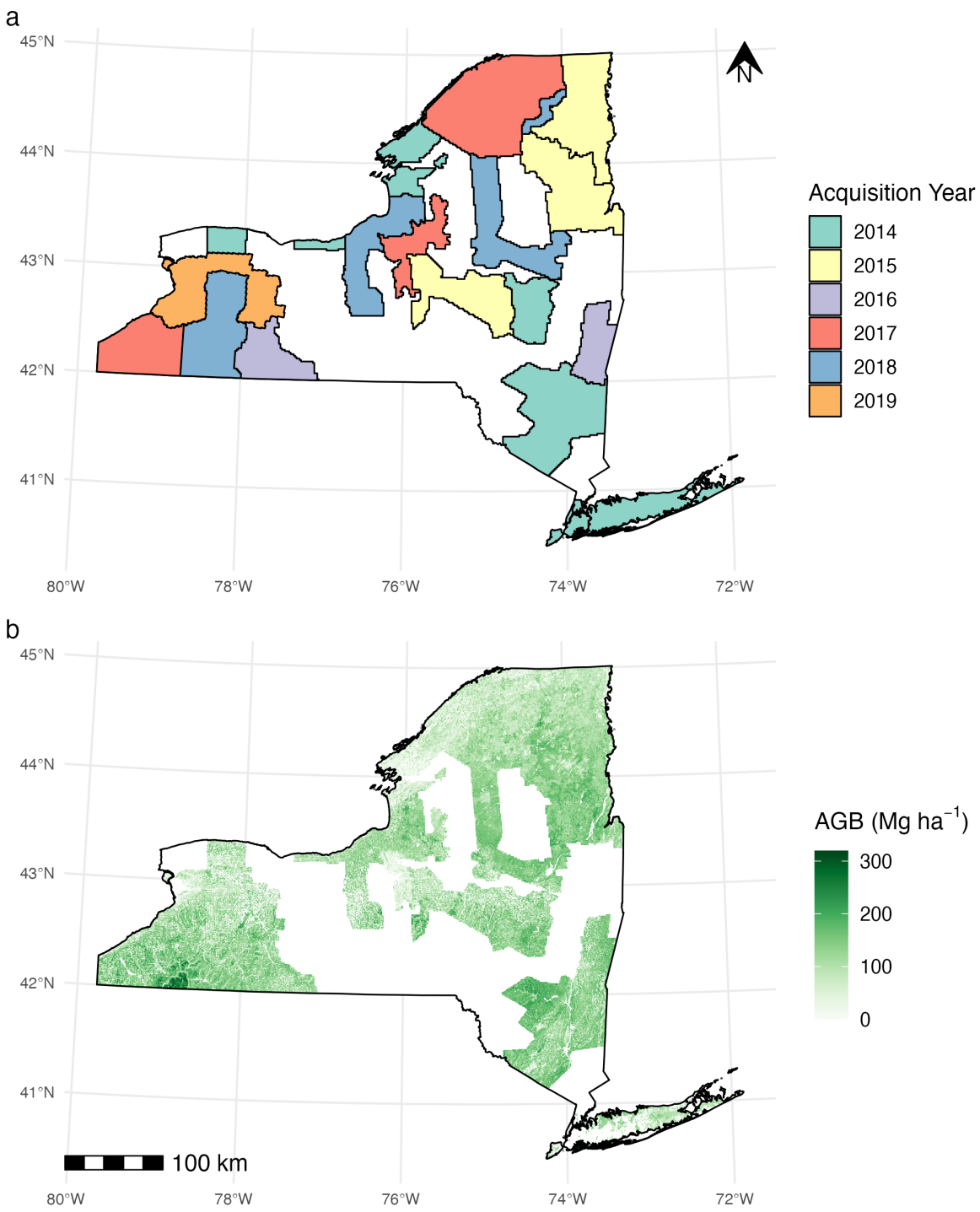


图2：基于激光雷达的地上生物量参考数据。a)按采集年份着色的激光雷达数据集空间覆盖范围。b) 为参考数据采样的基于激光雷达的地上生物量预测的时空拼图。

2010年；Kennedy、Yang 等人，2018年）。所有光谱指数及其相应的差值均采用1年滞后计算（Hudak 等人，2020年），并拟合到按时间分割的归一化燃烧比（NBR）顶点（Kennedy、Yang 等人，2018年）。我们使用 GEE 的“awesome-spectral-indices”JavaScript库计算归一化燃烧比（NBR；Kauth 和 Thomas（1976年））、缨帽湿度、亮度和绿度（TCW、TCB、TCG；Cocke、Fulé 和 Crouse（2005年））、归一化差异植被指数（NDVI；Kriegler 等人（1969年））、简单比值（SR；Jordan（1969年））以及修正简单比值（MSR；Chen（1996年））（Montero 等人，2022年）。扰动指标采用单独的 NBR 分割进行处理，并使用 LT-GEE 参数，以提高对离散扰动事件发生时间的敏感度（Kennedy、Yang 等人，2018年）。我们选择使用 NBR 处理所有其他由 LT-GEE 得出的预测变量，为其他所有指数提供扰动历史和时间断点进行拟合，因为研究表明 NBR 最能表征扰动事件（Kennedy、Yang 和 Cohen，2010年）。补充材料 1 提供了此处使用的 LT-GEE 参数的更多信息。

我们还纳入了美国地质调查局土地变化监测、评估与预测（LCMAP）1.2版的年度主要和次要土地覆盖分类预测结果（Brown 等人 2020年；Zhu 和 Woodcock2014年）。此外，还纳入了一组稳态辅助预测变量，用于表示气候、地形、生态和土地覆盖方面的地理空间变异（Kennedy、Ohmann 等人2018年）。这些预测变量包括来自 PRISM 气候组数据的降水量和温度30年气候常年值（PRISM 气候组 2022年）、高程、坡向、坡度，以及从30米数字高程模型推导出的地形湿度指数（MichaelJ. Mahoney、Beier 和 Ackerman 2022年；美国地质调查局 2019年；Beven 和Kirkby 1979年）；代表2005年状况的全球树冠高度图（Simard等人2011年；Hudak等人 2020年）；由美国人口普查局识别的距最近面状水体和线状水体的距离（米）（Walker2022年；美国人口普查局2013年）；由鱼类及野生动物管理局（FWS）制定的国家湿地清查分类（FWS 2022年；Wilen和 Bates 1995年）；以及美国环境保护署（EPA）的4级生态区（Omernik和 Griffith2014年；CEC 1997年）。如果单个EPA4级生态区未覆盖该州的≥ 2%，则将其合并到所属的3级生态区；如果这种合并仍未覆盖该州的 ≥ 2%，则将这些生态区设为“其他”。所有分类变量（LCMAP、生态区、湿地）均编码为布尔指示变量。

29个预测因子层（表2）中的每一层均被投影以匹配Landsat 30 m像素几何。预测因子的栅格堆栈在构建的FIA样地多边形处被裁剪并聚合（加权平均）（第2.3节），并且还与采样的基于LiDAR的地上生物量预测进行叠加（第2.4节），从而创建了两套不同的数据用于

基于同一组预测因子的模型训练。使用R（R 核心团队 2021）编程语言的exactextractr（Daniel Baston 2022年）和terra（Hijmans 2022年）包来编译训练数据集。

表2：用于模型拟合的预测变量定义。

组	预测因子	定义
光谱指数	TCB、TCW、TCG	缨帽亮度、湿度和绿度，以及 使用 LT-GEE 去除噪声
	NBR	归一化燃烧比，噪声已使用 LT-GEE
	NDVI	归一化差异植被指数，噪声 使用 LT-GEE 去除
	SR	使用LT-GEE去除噪声的简单比值
	MSR	使用去噪后的修正简单比值 LT-GEE
Delta	Delta *_	对所有预测变量以1年滞后计算的变化，在 “光谱指数”组
干扰	YOD, MAG	最新干扰年份及相关 NBR变化幅度，通过NBR识别 LT-GEE中的分割（1985-2019年）
生态	CHM	反映2005年状况的全球冠层高度模型 (Simard 2011)，从1公里降采样至30米 分辨率
	生态区	EPA 4级生态区。若为4级则聚合至3级 区域 < 2%的州面积。如果是第 3 级，则设为“其他”。 州的汇总 < 2% 。
	湿地	FWS国家 湿地清单
	到水体距离_ _	到最近 TIGER/Line Shapefile 的距离（米） 来自美国人口普查局的水
气候	降水量、最高温度、最低温度	降水量、最高 温度和最低温度的 30 年正常值，源自 年度PRISM气候模型
地形	坡向、高程、 坡度、TWI	坡向、高程、坡度和地形湿度指数 源自30米数字高程模型
土地覆盖	LCPRI、LCSEC	LCMAP 一级和二级土地覆盖 分类

2.6 模型开发

我们采用标准训练框架开发了三种不同的建模方法。直接方法使用从FIA野外数据派生的模型数据集中的随机80%划分来训练模型（第 2.3节），间接方法则使用LiDAR地上生物量预测样本中的随机80%划分来训练模型（第 2.4节）。我们分别为这两种方法开发了多组机器学习模型，并将每组模型组合成堆叠集成，以更好地反映模型选择的不确定性（Wintle等，2003年），并降低组成模型预测的泛化误差（Wolpert，1992年）。第三种方法是将直接集成模型和间接集成模型的预测结果进行简单平均，以此构建集成模型。已有研究表明，当数据噪声较大，且预测变量与响应变量之间的关系复杂且大多未知时，模型平均能够改善单个预测结果（Wolpert，1992年；Dormann等，2018年）。对于这三种方法，我们均使用20%的测试划分，根据各自的数据集评估模型性能，并尝试不同的预测变量和模型形式。

直接方法和间接方法均使用了第 2.5 节中描述的全部 29 个预测变量，而集成模型仅使用这些模型的预测结果构建。直接方法和间接方法均结合了随机森林，

由 ranger R 包实现（Breiman 2001；Wright 和 Ziegler 2017）的随机森林，以及由 lightgbmR 包实现的随机梯度提升机（GBM）（Friedman 2002；Ke 等，2017；Shi 等，2022）。直接方法还纳入了由 kernlab R 包实现的支持向量机（SVM）（Cortes 和 Vapnik 1995；Karatzoglou 等，2004）。SVM 的训练时间相对于训练观测数量呈二次至三次方增长（Bottou 和 Lin 2007），因此，使用 16,000 个训练点实现间接方法在计算上不可行。

每个组成机器学习模型均使用上述 80% 训练分区进行调优，并采用迭代网格搜索：先通过五折交叉验证测试较宽范围的超参数，再经过数轮迭代逐步缩小范围，只保留性能最佳的组合。随后，模型在所有其他分析中均使用最准确的超参数组合。各组成模型选定的超参数以及线性回归集成中的系数见补充材料 2。对于训练数据集中的每个观测值 n，均使用 n−1 个观测值和各自的最优超参数拟合所有组成模型，并对第 n 个（留出）观测值进行预测。随后，使用线性回归模型根据这些留一法预测值估算地上生物量（AGB），将各组成机器学习模型组合成线性回归集成，具体如下：

$$\text{AGB} = \beta_0 + \beta_1 \cdot P_1 + \ldots + \beta_n \cdot P_n \tag{1}$$

其中 β_* 是通过普通最小二乘回归估计的系数，而 P_* 是相应的组成模型预测。在抽象层面上，直接方法构建如下：

$$\text{AGB} = ensemble(RF, SVM, LGB) \tag{2}$$

其中 $ensemble$代表 发送 方程 1，以及 RF， SVM，以及 GBM将替代 P_*变量 i n 方程1. 间接 ct方法构建如下：

$$\text{AGB} = ensemble(RF, LGB) \tag{3}$$

且总体集成构建如下：

$$\text{AGB} = \frac{ensemble(RF_{direct}, SVM_{direct}, LGB_{direct}) + ensemble(RF_{indirect}, LGB_{indirect})}{2} \tag{4}$$

2.7 地上生物量制图和后处理

直接和间接方法的线性模型集成，以及总体平均集成，被用于对全州所有30米像元进行预测。鉴于我们的预测最适用于有木本生物量分布的区域，我们将预测结果与LCMAP1.2版主要土地覆盖分类产品进行叠加（Brown等人 2020年；Zhu 和Woodcock2014年）。据报告，该产品在美国东部1985-2018年期间的总体精度为77.4%（Pengra 等人2020年）。LCMAP数据与我们的地上生物量地图具有完全相同的像元几何结构，其年度分辨率使其能够与每个独立制图年份在时间上对齐。我们对地上生物量预测表面进行掩膜处理，移除已开发土地、耕地、水域和荒地像元，然后按剩余的三类植被LCMAP类别——树木覆盖度、草地/灌丛和湿地——统计地上生物量。

2.8 地图一致性评估

[{ "id": 0, "output": "我们依据 Riemann 等人（2010年）以及 Menlove and Healey (2020年) 所规定的方法，评估了我们的地上生物量地图与 FIA参考数据之间的一致性。前者在一系列尺度上评估一致性，并通过仅提取与 FIA样地重合的像元，解决地图聚合估计值（许多像元）与 FIA 聚合估计值（少量样地）之间空间支持范围不匹配的问题。后者则将基于 FIA 得出的地上生物量估计值——这些估计值已针对六边形地图单元内的森林覆盖进行调整，并外推至相应面积——与我们所制地图的分区平均地上生物量进行比较。"}]

遵循Riemann等人（2010年）的做法，我们将三种建模方法各自生成的地上生物量预测表面与地图评估数据集（第2.3节）进行了比较。比较分别在样地到像元尺度上，以及在不同大小的六边形网格内进行；六边形网格质心之间的距离范围为20公里（34,641公顷）至50公里（216,506公顷）。由于样地清查涵盖多个年份（2007年、2012年、2018年、2019年），我们仅从那些与数据集中特定样地清查在时间上相匹配的地图表面中提取预测值。随后，我们将这些数据汇集在一起，生成时间上广义的精度评估。作为对Riemann等人（2010年）方法的扩展，我们通过汇总质心间距为50公里的六边形网格单元内的样地到像元残差和FIA参考数据分布，评估预测误差的空间模式。我们还根据每个样地处占主导地位的 LCMAP分类对样地到像元结果进行分组，以展示各植被土地覆盖类别之间的一致程度。

遵循 Menlove andHealey(2020)的方法，我们将掩膜预测值的平均值（按每个像元与给定六边形网格相交的比例加权）与一组基于 FIA得出的估计值进行了比较；这些估计值对应于64,000公顷的六边形网格，代表FIA针对纽约州最新清查周期（2013年–2019年）所采用的最精细可接受尺度。由于2016年处于 Menlove and Healey(2020)数据所涵盖时间段的中心位置，我们使用了每种方法的2016年地上生物量图进行比较。按照建议，我们通过将 FIA估计值除以每个六边形网格内植被覆盖区域的比例，校正了 FIA估计值与我们的制图估计值之间森林定义的差异；其中植被覆盖区域基于 LCMAP 的树木覆盖度、草地/灌丛和湿地确定。最后，我们将此比较限制在大部分面积位于纽约州边界内的六边形网格中。

评估指标包括以 Mg ha^{-1} 为单位的平均绝对误差（MAE）、相对于参考地上生物量（AGB）均值的 MAE 百分比（% MAE）、以 Mg ha^{-1} 为单位的均方根误差（RMSE）、相对于参考 AGB 均值的 RMSE 百分比（%RMSE）、以 Mg ha^{-1} 为单位的平均误差（ME），以及决定系数（R^2）。各项指标的方程和公式，以及相应的标准误估计值，见补充材料3。本研究所述的所有分析均使用 R编程语言（R 核心团队 2021）中的 exactextractr（Daniel Baston 2022）、sf（Pebesma 2018）和 terra（Hijmans 2022）程序包完成。

2.9 细尺度空间格局的定性比较

我们还对 Huntington Wildlife Forest (HWF) 内部及周边各建模方法的制图预测进行了目视比较。HWF 是位于纽约州 Newcomb 的一片面积达 6,000 公顷的森林，既有保护区，也有正在积极管理的区域；我们的团队通过实地观测和遥感观测，逐渐熟悉了当地的景观。尽管这项比较仅涵盖全州范围的一小部分，但其目的是定性评估各建模方法在不同管理制度和景观状况下刻画地上生物量密度精细空间格局的相对优缺点。我们进行了栅格两两相减，以生成突出显示不同建模方法之间分歧区域的表面；此外，还使用了 1 米激光雷达衍生的冠层高度模型（CHM；Atlantic Inc（2015年））以及来自国家航空影像计划（NAIP；地球资源观测与科学 (EROS) 中心（2017年））的 0.5 米自然彩色影像，作为补充的定性参考信息。CHM 和 NAIP 影像反映的是 2015年的状况，因此我们比较了各建模方法生成的 2015年地上生物量预测表面。

3 结果

3.1 年度地上生物量图

我们使用三种建模方法中的每一种，制作了30年（1990-2019年）全州范围内30米分辨率的地上生物量图。三种建模方法中每一种的全州地上生物量平均值在整个时间段内对每个包含的LCMAP分类均稳步增加（图 3）。然而，与其更高的饱和效应阈值一致（第 3.2节），间接方法产生的平均值显著大于直接方法和集成方法（图 3）。大约在2006年，所有三种模型均在树木覆盖度分类像元的全州平均值上产生了小幅下降；这与东北部大规模虫害爆发的时间相对应（2005-2007年，Kosiba等人（2018年）），特别是森林天幕毛虫（*Malacosoma disstria*）落叶事件，该事件影响了纽约州约120万英亩的土地（美国林务局 2006年）。虽然仅落叶不一定会导致地上生物量损失，但我们的模型对光谱信息的依赖使其无法区分冠层特定变化和结构变化。

使用集成预测结果进行全时序栅格相减（2019年地上生物量 - 1990年地上生物量）后，所得结果反映了这些年度趋势，地图上以地上生物量增加为主（图 4）。这幅30年储量变化图还呈现了由人为影响和地籍边界所驱动的地上生物量变化模式，这些模式与可归因于其他自然过程的变化形成对照。具体而言，储量变化图突显了经营林与阿迪朗达克森林保护区土地交织而成的镶嵌格局，以及伴随这些不同土地利用方式而出现的变化空间格局和幅度差异（图 4 b）；图中还显示，阿勒格尼河以西的私人土地上，地上生物量损失呈斑块状分布，而河流以东的阿勒格尼州立公园内则略有增加且总体相对稳定（图 4 c）；此外，图中还揭示了一条沿卡茨基尔森林保护区边界南北延伸的森林生长带（图4d）。

在林分尺度上，对于纽约州北部有土地所有者提供的经营记录的区域，我们的年度地图准确捕捉了经营林中采伐活动的时间、强度及其后续恢复（再生）情况（图 5）。尤其是查看图5 中的皆伐区域及这些多边形边界内的残余地上生物量时，我们注意到，现有的空间经营记录只是对采伐方案的最佳近似，可能无法反映采伐活动的真实范围。同样，我们的制图预测也捕捉到了这些采伐多边形以外的扰动（见图 5 中从2015年开始的西部区域）；在此情况下，这些扰动可归因于未纳入土地所有者所提供记录的采伐事件。

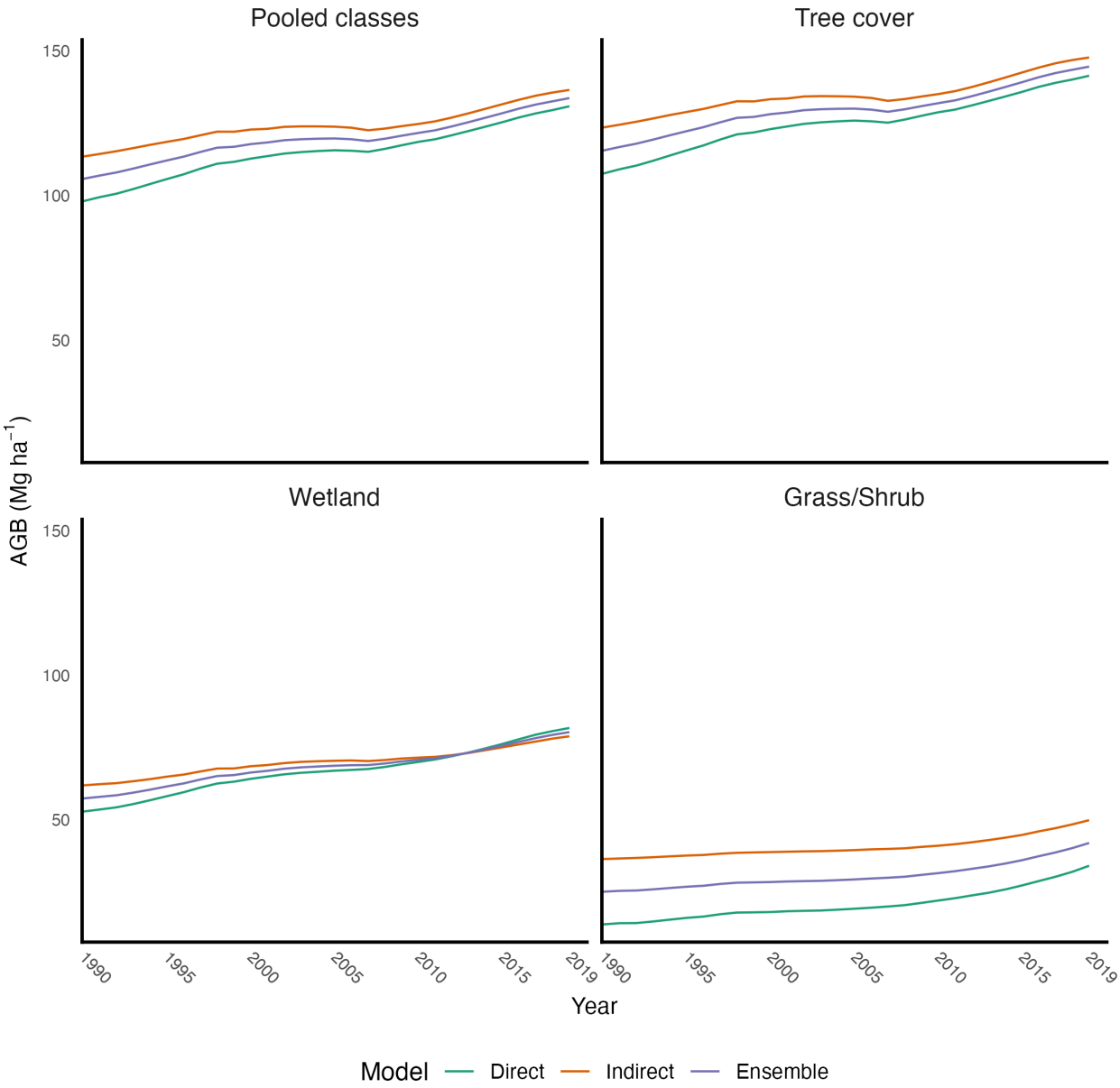


图3：按LCMAP类别划分的每种建模方法的年度全州汇总（平均地上生物量）。

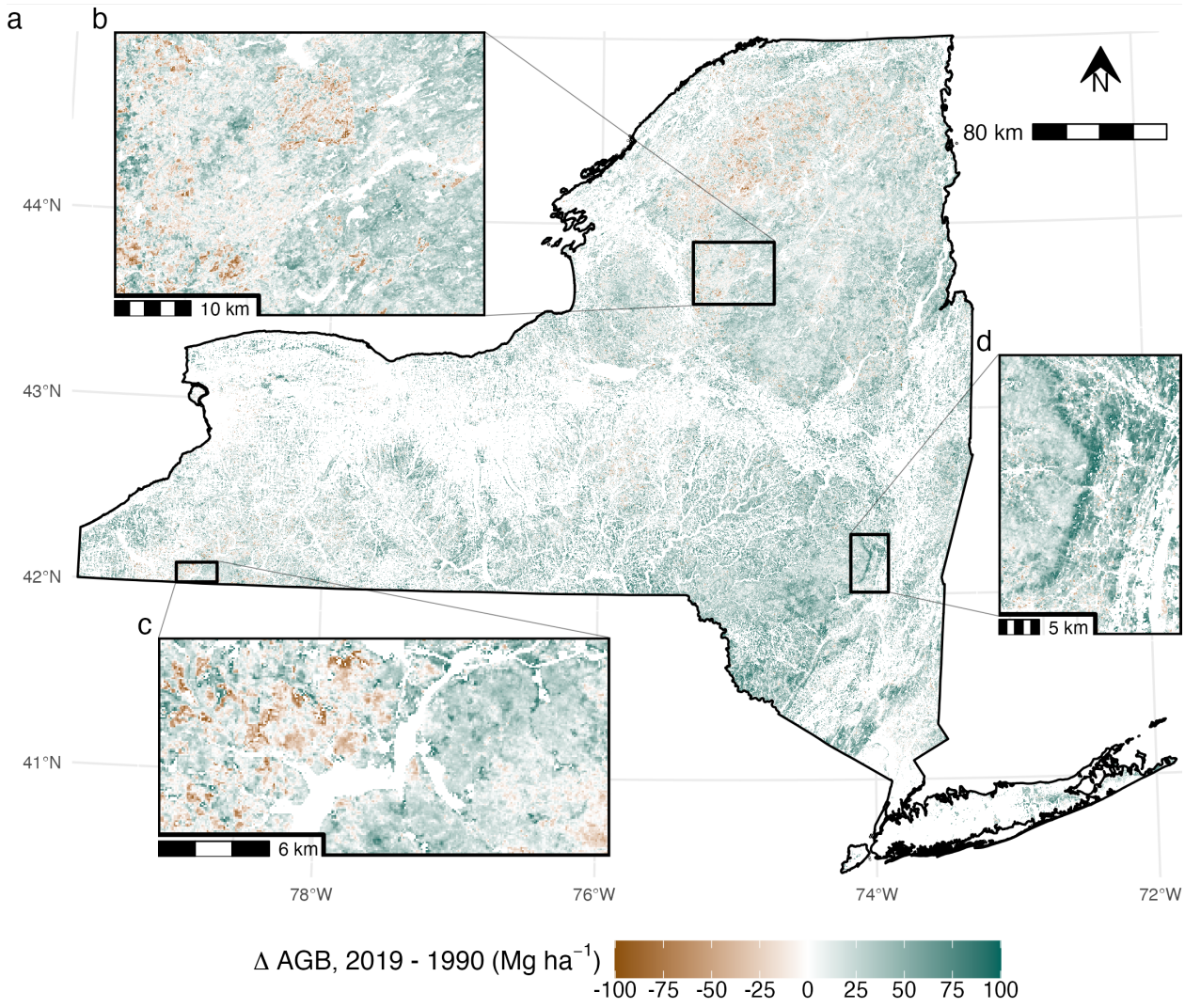


图4：纽约州（美国）地上生物量差值图（2019年地上生物量 - 1990年地上生物量），采用集成模型预测结果。a）全州尺度。b）斯蒂尔沃特水库以南，由经营林和阿迪朗达克森林保护区土地组成的镶嵌区域，位于纽约州。c）阿勒格尼河区域，河流以东为阿勒格尼州立公园的一部分。d）卡茨基尔森林保护区边界沿线的森林生长情况。显示值上限设为 ± 100 Mg ha^{-1} 。

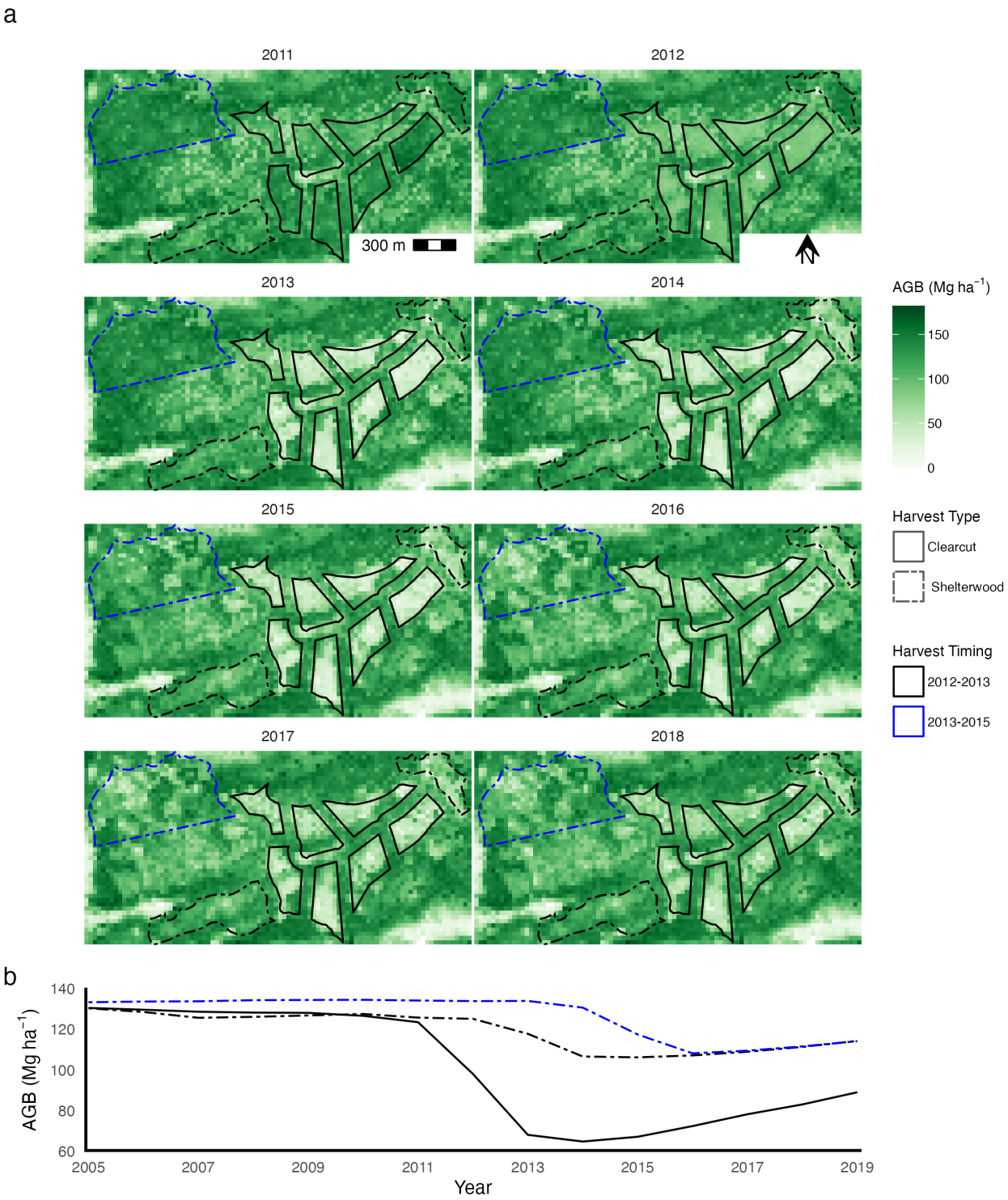


图 5：量化纽约州北部（美国）采伐及后续再生导致的地上生物量变化。a）集成模型对选定年份的年度地上生物量预测，并叠加按记录的采伐类型和时间标示的采伐记录。b）按采伐类型和时间分组的采伐多边形年度面积级制图预测汇总（平均地上生物量），其中轨迹符号系统与 a）中的多边形符号系统相对应。

3.2 地图一致性

尽管我们三种建模方法的精度指标估计值差异不大，但集成模型的精度最高（表 3）。另一方面，间接方法在这些指标上的精度最低，这可能是由于初始激光雷达-地上生物量预测中的像元级误差产生了累加效应（Johnson等，2022）。随着聚合单元规模增大，我们观察到三种建模方法的制图预测与FIA估计值之间的一致性均有所提高：直接方法的%平均绝对误差从34.1降至19.46%，间接方法从35.69降至20.23%，集成方法从33.88降至19.23%（表 3）。平均绝对误差、均方根误差、%均方根误差和R^2, 也呈现出一致性随之提高的相似趋势，但ME估计值在所有聚合尺度上基本保持稳定且为正。

这三个模型在地上生物量为零或接近零的参考观测值上往往会产生过高预测，尤其是在样地:像素尺度和20千米尺度的比较中（图 6），从而导致平均误差估计值为正且显著（表3）。许多过高预测可以归因于我们依赖基于树的模型（RF、GBM），其预测值是终端节点内的平均值（Baccini 等人，2008年；Urbazaev等，2018年）。然而，这些过高预测也可能是由我们的地图评估数据集中的结构性零值造成的：在该数据集中，FIA地上生物量被假定为零，但实际上由于FIA的严格森林定义而未被测量（第2.3节；Johnson et al. (2022)）。将FIA样地分类为草地/灌丛后出现的较大相对误差，进一步证明了森林定义差异对地图一致性结果的影响（表 4）。遗憾的是，在没有额外数据的情况下，我们无法识别包含结构性零值的样地，也无法将其与其他确实存在过高预测和误差的样地区分开来。

对最大参考观测值的低估（即饱和效应）是使用光学影像对森林结构进行建模时的常见问题（Lu 2005年；Duncanson、Niemann 和Wulder 2010年），在三种建模方法中均很明显，但程度各不相同（图 6）。直接方法最先出现饱和，无法预测超过 204Mg ha^{-1}的值，而间接方法在这方面表现最佳，预测值最高可达 289 $Mgha^{-1}$，且参考数据中仅有 1%超出其上限。总体而言，随着聚合尺度增大，高估和低估的模式有所减弱，系统一致性得到改善，这一点可从所有模型的 GMFR 线与 1:1 线逐渐趋于一致中看出（图 6）。尽管在估计指标方面的准确性最低（图6；表3），间接方法在所有尺度上均取得了最佳的系统一致性（GMFR 与 1:1 的比较）。

与FIA的小区域估计值（Menlove andHealey(2020)）进行的地图比较同样显示，在参考AGB分布的极端值处存在高估和低估这两种模式，同时也显示了三种建模方法各自的饱和效应（补充材料4）。尽管与Menlove and Healey (2020)的估计值相比始终存在低估，直接方法所得估计值落在所提供的95%置信区间内的比例更高（90.31%），高于集成方法（88.27%）和间接方法（85.2%）。同样，在对样地到像元残差进行制图时，局部误差（高估和低估）更多地与每个六边形单元内的参考AGB数量有关，而不是与空间或区域模式有关（补充材料 4）。

各模型的树木覆盖度一致性（表 4）总体上与表3中的样地到像元一致性相符，因为绝大多数地图评估样地都属于这一分类。湿地和草地/灌丛分类样地数量较少，其地图一致性较差；ME估计值表明，草地/灌丛分类样地的生物量被显著高估，而湿地分类样地的生物量被低估（表 4）。植被类别间一致性的这种差异，很可能是因为各土地覆盖分类在我们的参考数据集中的代表程度不一，同时也因为我们基于LCMAP定义的植被掩膜（第 2.7节）与FIA采用的严格森林定义（第 2.3节）并不一致。

3.3 精细空间格局的定性比较

在Huntington Wildlife Forest (HWF)内、HWF以北的森林保护区（High Peaks Wilderness；Pataki和Cahill（1999））以及HWF西南方的经营林中，与其他两种方法相比，间接地上生物量地图最能呈现整个景观中已知的格局，且空间异质性最强（图 7）。在地图间差异最大的区域，即该区域的东北角和西北角，这一点尤为明显。在东北角，针叶林主导的湿地（NAIP图 7）包含该区域一些最高的植被（CHM图 7 ≥ 225 $Mg\ ha^{-1}$），间接方法得出的生物量预测值较高（Mg ha），与这些景观特征相符。在地图的西北角，高海拔云杉冷杉林分布于此，间接方法相应地得出较低的地上生物量预测值；而直接方法则无法将这些区域与景观的其他部分区分开来。按定义，集成地图在精细空间格局和预测值大小两方面，兼具直接地图和间接地图的特征。

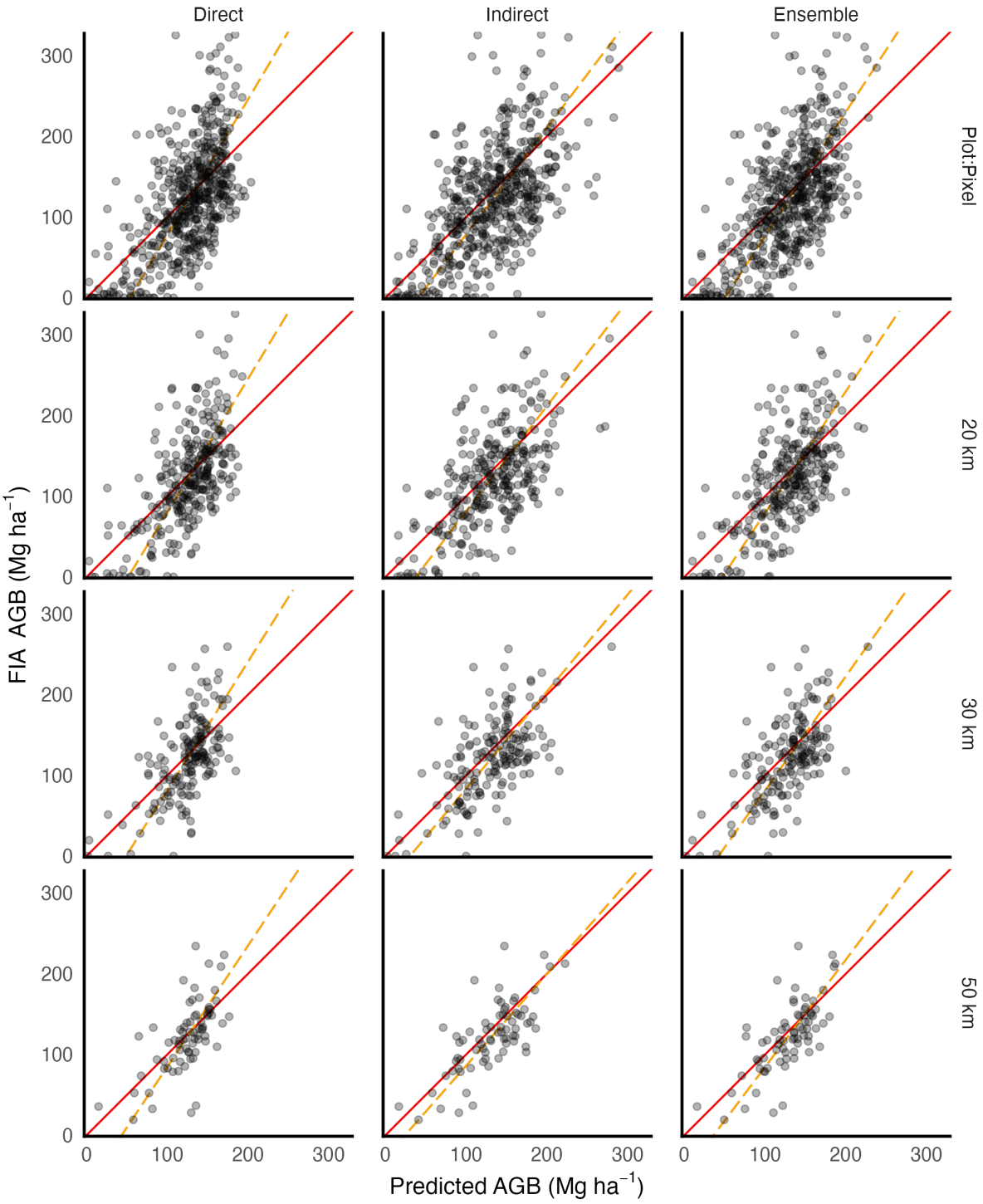


图6：在选定尺度下，将制图得到的地上生物量与 FIA估算的地上生物量进行比较，这些尺度由六边形质心之间的距离表示（样地:像素、10公里、25公里和50公里）。几何平均函数关系（GMFR）趋势线以虚线（橙色）表示，1:1 线以实线（红色）表示

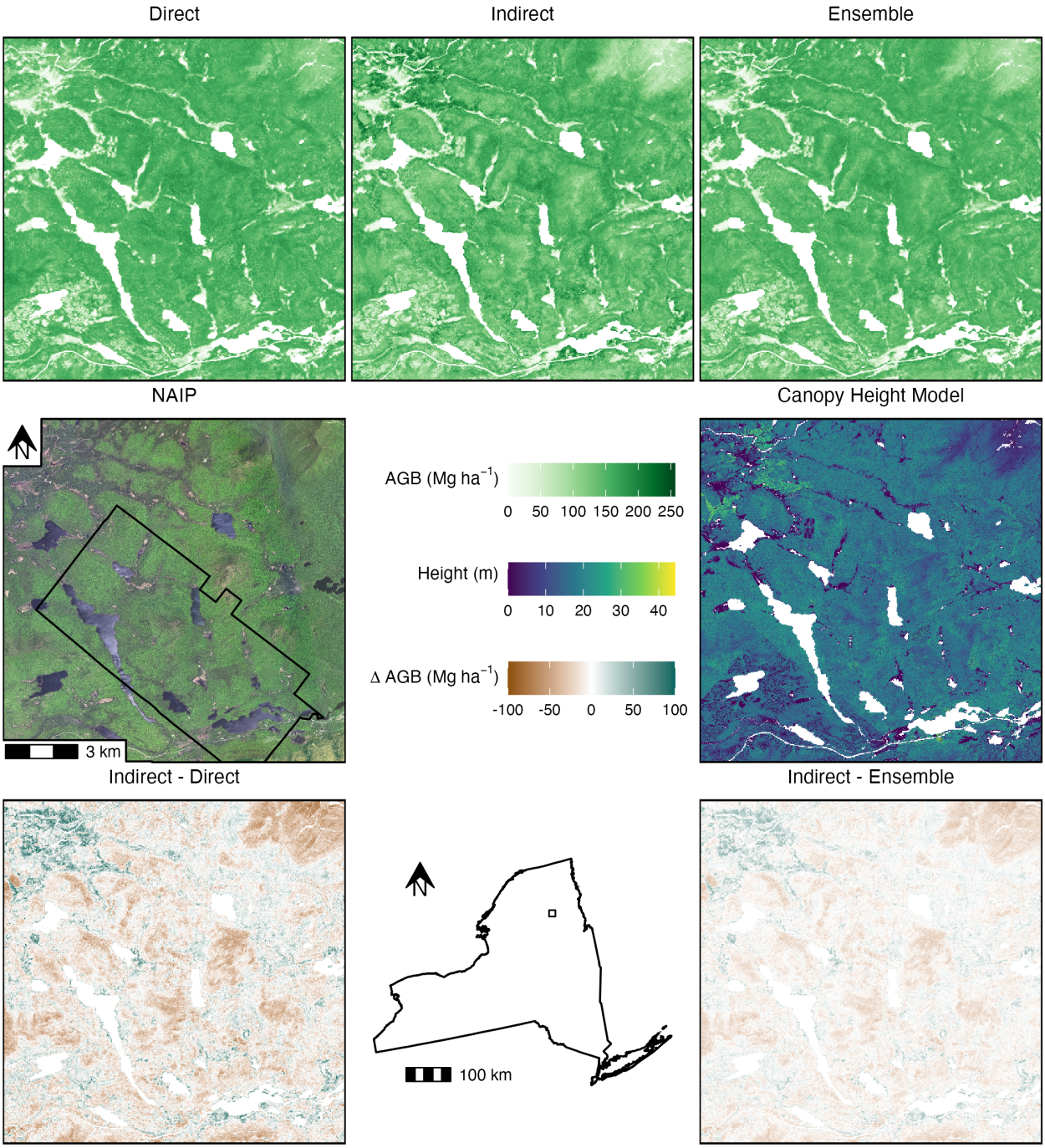


图7：Huntington Wildlife Forest 内各建模方法所得地图的定性比较（NAIP 面板中的黑框标示其边界），以及纽约州 Newcomb 周边区域的比较（纽约州面板中的黑框标示完整区域范围）。成对栅格相减（显示值上限为 ± 100 Mg ha^{-1}）突显了模型预测之间差异的空间分布模式及幅度。集成 - 直接 未显示，因为它与间接 - 集成重复。另提供由 1 m 激光雷达生成的冠层高度模型，以及分辨率为 0.5 m 的自然色国家航空影像计划 (NAIP) 正射影像，供参考。所有表面均表示2015年的状况。

表3：选定尺度的地图一致性结果。RMSE、MAE、ME 单位为 Mg ha^{-1}。尺度 = 六边形质心之间的距离，单位为 km；PPH = 每个六边形的样地数；n = 比较单元数量（样地或六边形）。所有精度指标定义见补充材料3。括号内为标准误，最小值下限为 0.01。

尺度	n	PPH	模型	% MAE	MAE	% RMSE	RMSE	ME	R^2
样地:像素	545		直接	34.10	41.20 (1.38)	43.29	52.31 (3.06)	4.83 (2.23)	0.38 (0.01)
			间接	35.69	43.13 (1.44)	45.30	54.73 (3.14)	11.15 (2.30)	0.32 (0.01)
			集成	33.88	40.94（1.36）	42.84	51.76 (2.98)	7.99 (2.19)	0.39 (0.01)
20 千米	302	1.8	直接	29.17	35.53 (1.69)	37.81	46.06 (4.19)	3.42 (2.65)	0.40（0.01）
			间接	31.51	38.39 (1.71)	39.80	48.48 (3.92)	9.22 (2.74)	0.33 (0.01)
			集成	29.57	36.02 (1.64)	37.65	45.86 (4.21)	6.32 (2.62)	0.40 (0.01)
30 千米	172	3.17	直接	25.35	30.76 (1.87)	32.41	39.32 (6.12)	3.66（2.99）	0.37 (0.01)
			间接	27.40	33.25 (1.86)	33.97	41.21 (5.34)	10.03 (3.06)	0.30 (0.01)
			集成	25.79	31.30 (1.76)	32.02	38.86 (5.01)	6.85 (2.93)	0.38 (0.01)
50公里	73	7.47	直接	19.46	23.85 (2.70)	26.97	33.05 (12.80)	2.58（3.88）	0.43 (0.01)
			间接	20.23	24.80 (2.39)	26.14	32.04 (8.54)	9.46 (3.61)	0.46 (0.01)
			集成	19.23	23.57 (2.39)	25.38	31.10 (10.31)	6.02 (3.60)	0.49 (0.01)

表4：按 LCMAP分类分组的样地到像元尺度地图一致性。均方根误差、平均绝对误差、平均误差，单位为 Mg ha^{-1}。n= 样地数量。所有精度指标的定义见补充材料3。括号内为标准误，数值上限为 0.01 和 1.00。

LCMAP	n	模型	% 平均绝对误差	MAE	% 均方根误差	RMSE	ME	R^2
草地/灌丛	14	直接	87.58	34.07 (7.50)	111.81	43.49 (66.44)	3.76（12.02）	0.41 (1.00)
		间接	101.20	39.36 (10.95)	143.30	55.74 (217.09)	33.85 (12.28)	0.03 (1.00)
		集成	93.92	36.53（6.65）	112.34	43.69 (51.15)	18.80（10.94）	0.40 (1.00)
湿地	57	直接	40.19	34.56 (4.34)	55.15	47.43 (28.20)	-9.21 (6.22)	0.40 (0.01)
		间接	43.47	37.38 (4.26)	57.12	49.12 (29.95)	-10.29 (6.42)	0.35 (0.02)
		集成	40.87	35.15 (4.22)	54.94	47.24 (31.20)	-9.75（6.18）	0.40 (0.01)
树木覆盖度	474	直接	33.13	42.21 (1.48)	41.67	53.10 (3.44)	6.55 (2.42)	0.31 (0.01)
		间接	34.47	43.93 (1.55)	43.42	55.34 (3.46)	13.06 (2.47)	0.26 (0.01)
		集成	32.78	41.77（1.46）	41.19	52.50 (3.23)	9.80（2.37）	0.33 (0.01)

4 讨论

在本研究中，我们将经过时间平滑、分段处理和空缺填补的Landsat影像，与一组基于激光雷达的地上生物量（AGB）预测值样本以及美国农业部森林清查与分析（FIA）野外样地相结合，为美国纽约州（NYS）生成了年度全覆盖地上生物量地图。为此，我们开发了三种独立的建模方法，包括直接、间接和集成方法。总体而言，我们发现三种建模方法的表现相近，这表明每种方法单独使用都可以取得令人满意的效果；但在模型复杂度、地图精度、饱和效应以及精细空间格局的表现方面，均存在明显的权衡。与目标相似、但位于具有不同干扰和经营制度的温带地区的现有研究进行比较后发现，本文所述的基本方法可用于追踪不同生态域以及经营林中的森林生物量动态，而不受占主导地位的林业实践影响。由每种建模方法生成的地图，为了解30年来森林结构、发展、干扰和变化的时空格局提供了宝贵的洞见，并可作为多种应用的输入数据，包括基于地图的存量变化评估、筛选或确定优先纳入自然气候项目的森林地块，以及纽约州范围内未来的监测、报告和核查（MRV）系统。

4.1 不同建模方法之间的权衡

三种建模方法并无绝对优胜者，而是各有优势，能够满足不同项目特定的限制与目标。总体而言，集成方法生成的地图最为准确（表 3），它在精细尺度模式表征和模型饱和效应方面结合了直接方法与间接方法的特点。尽管它是三种方法中最复杂的，但同时也缓解了基于样地的方法（第2.3节）和基于激光雷达的方法（第 2.4节）各自的局限，并发挥了它们的优势。

训练数据集。这些结果总体上支持在生态应用中采用模型集成，尤其是在数据存在噪声且自然变异是误差重要来源的情况下（Dormann等，2018年）。

按定义，直接方法最为简约，只需一个模型阶段，并且生成地上生物量地图产品所需的时间和精力投入最少。间接方法和集成方法需要对太字节级的激光雷达数据进行计算量很大的管理和分析（Johnson等，2022），不过，如果使用激光雷达条带或样本来代替全覆盖制图，这一工作量可以减少（Wulder等，2012年；Matasci等，2018年；Urbazaev等，2018年）。此外，间接模型和集成模型中更高的复杂性使得预测不确定性的估计比直接方法更具挑战性（Saarela等，2016年）。

间接方法受饱和效应的影响最小，因此在所有尺度上与FIA参考数据均达成了最佳的系统性一致（图 6）。由于仅有1%的参考地上生物量样地超出间接模型的预测上限，该方法最适合追踪成熟林分的持续生长。这一特点对于纽约州及更广泛区域尤为重要，因为历史土地利用动态表明，大多数林分要么已经达到成熟，要么正在接近成熟（第2.2节）。未能准确量化这些林分中的地上生物量，将导致在当地和全州尺度上对碳储存和碳汇的显著低估。此外，我们发现，间接方法生成的地图与我们对当地森林状况的认识最为一致，也最能代表景观中的细尺度特征（图7）。间接模型的优势可归因于其拥有规模大得多的参考数据样本；从理论上讲，还归因于其通过大范围的基于LiDAR的地上生物量地图（第 2.4节）获取了对纽约州地上生物量分布和景观状况更广泛的覆盖。

4.2 与现有研究的比较

跨研究比较模型性能和地图一致性时应谨慎，因为景观、数据采集规程、遥感数据产品和地上生物量分布可能存在很大差异，并对最终的一致性指标产生重大影响。不过，我们仍以相对的方式进行比较，以便将我们方法的成效置于目标相似的现有研究中加以考察。Kennedy、Ohmann等（2018年）、Hudak等（2020年）和Matasci等（2018年）均利用Landsat时间序列数据，以30米分辨率绘制以下地区和时期的年度地上生物量地图（依次为）：俄勒冈州西喀斯喀特省和北加利福尼亚州，2000—2016年；华盛顿州、俄勒冈州、爱达荷州和蒙大拿州，1990—2012年；加拿大森林主导生态系统，1984—2016年。Kennedy、Ohmann等（2018年）仅采用直接建模，针对涵盖广泛地上生物量值（0—1000 Mg ha ~103 Mg ha^{-1} ）的模型训练样地，得到的均方根误差为-1；Hudak等（2020年）和Matasci等（2018年）则仅采用间接建模，针对独立FIA样地的均方根误差分别为64%，针对基于LiDAR的地上生物量预测的均方根误差为66%。尽管此类直接比较存在局限，但它们表明，依托Landsat时间序列影像刻画森林动态的类似方法适用于多个领域——从具有同龄林干扰制度的美国西部和加拿大针叶林主导的森林（Kennedy、Ohmann等，2018年；White等，2017年），到以异龄林干扰制度为主的美国东部北方硬木林和混交林（第2.2节）。要确保在大尺度碳核算框架中，所有经营性森林所有者及其景观都得到准确、公正的对待，就需要具备这种追踪不同干扰模式和管理制度下森林变化的能力（Desrochers等，2022年）。

4.3 年度地上生物量地图的应用

我们经过严格评估的地图产品具有广泛的应用场景，在这些场景中，需要了解森林生物量（以及由此延伸出的森林碳库）的时空格局。最直接的是，鉴于我们在模型开发（第 2.3节、第 2.4节）和地图评估（第2.8节）中广泛使用了FIA 样地级信息，我们的年度地图为将FIA 信息转化为空间显式存量变化核算方法的输入提供了途径。这种基于地图的框架能够按照FIA 的核算方法汇总储量变化和固存速率，同时还具有更大的灵活性：可针对纽约州内任意面积单元，在这 30 年期间的任意时间窗口内进行汇总（图 4）。这一提升的分辨率使我们能够识别具有明显时空特征的AGB 损失和增加，并将其归因于整个景观中的保护、监管和所有权格局（图 4b、c、d）。基于样本的存量变化方法能够总体上捕捉这些结果，而我们的地图则可以更精确地识别整个景观中人类和自然过程在何处、何时以及如何影响森林碳储量。

尽管模型数据不应取代直接测量、清查或实地知识，但我们的地图所提供的历史视角使我们能够填补管理记录或森林清查数据不可用之处的空白（图 5）。过去的管理信息以及历史地上生物量或碳储量信息均可获得，这为量化各种管理制度的成果打开了诸多机会之门（Kaarakka 等，2021；Patton等，2022年）。此外，参与碳抵消项目所需证明额外性的责任，取决于建立可信的照常经营基线；而没有历史数据，这些基线不可能形成（Gillenwater 等，2007）。诸如此处开发的地图数据集可以同时为潜在参与者和项目管理者填补这一空白，从而最大限度地减少许多原本高得令人望而却步的前期成本和要求（Charnley,

Diaz和Gosnell2010年；Kerchner和Keeton 2015年）。更广泛地说，这些历史数据集可以提供基线，以更好地理解当前和未来的森林状况如何响应多重变化驱动因素，包括快速变化的气候（Cohen等 2016年；White等，2017年）。

由于我们主要依赖联邦资助且公开可获取的数据源，以及开源软件和工具，因此可以灵活运用为这一历史背景开发的相同方法，以满足持续监测（MRV）的需求。我们的建模流程只需结合年度 Landsat影像 和FIA 清查数据，并 opportunistically 加入激光雷达数据采集（Sugarbaker等，2014年、2017年），即可构建一个极具成本效益、广泛可复现且可扩展的景观监测框架。通过整合有望提高相对于仅使用 Landsat 训练的模型预测精度的新信息流，还可以进一步增强这一方法（例如，ESA 的 BIOMASS任务——Quegan 等（2019年）；NASA 的 GEDI 任务——Dubayah 等（2014年））。

最新的地上生物量（AGB）和碳储量地图将使决策者能够确定优先地块，既可通过购买完整产权或保护地役权来加以保护，也可将其纳入改进的森林管理项目和碳市场（Merenlender等，2004；Malmsheimer 等，2008；Kelly、Germain 和Stehman，2015年；Kerchner 和Keeton，2015年）。同样，及时的年度地上生物量（AGB）地图可以支持全面覆盖的MRV和采伐监测，但并不一定取代必要的实地考察，而是可用于筛选出可能合规的地块与需要进一步仔细检查的地块（Gillenwater 等，2007）。在此类系统下，监测成本不仅会显著降低，可能还会降低潜在项目的财务盈亏平衡阈值（Charnley、Diaz 和 Gosnell，2010；Kerchner和 Keeton，2015年）；此外，当定期普查各处房产或土地持有情况在其他条件下不可行时，严格随机的现场考察也将不再必需。除年度监测之外，根据我们30年的地图所推导出的高分辨率地上生物量（AGB）变化轨迹，还可为时间序列预测和景观模拟研究提供信息，以预测各种政策和管理情景下的碳影响（MacLean 等，2021）。

5 结论

历史森林动态的高分辨率地图可作为空间显式存量变化核算框架的输入，为预测地块到景观尺度土地管理决策所产生的碳效应提供关键信息。目前亟需能够在近期交付这些历史数据集，并提供可复现、一致且广泛适用的数据产品的方法。我们展示了三种基于模型的方法，利用开源数据、软件和工具，以30米分辨率预测过去三十年（1990–2019年）纽约州（NYS）范围内的年度地上生物量（AGB）。结果表明，这三种方法均能提供有效输出，并且相互之间各具独特优势，从而为纽约州提供了一系列可选方案。在纽约州，预计森林将作为碳汇发挥重要作用，助力到2050年实现净零碳经济。更广泛而言，本研究制作的地图产品能够帮助管理者和决策者最大限度地发挥森林景观在自然气候解决方案和相关政策中的作用。

6 致谢

我们谨向美国农业部FIA项目致以感谢，感谢其提供数据共享与合作；感谢荷兰养老基金和F &W Forest Management,LLC 分享管理记录；感谢纽约州GPO 整理并提供激光雷达数据；感谢纽约州环境保护部气候变化办公室提供资金支持。我们还要感谢 Grant M. Domke、Stephen V. Stehman和 Eddie Bevilacqua 在整个项目期间给予的支持与指导。

参考文献

Abdalati, Waleed, H.Jay Zwally,Robert Bindschadler, Bea Csatho, Sinead Louise Farrell, HelenAmanda Fricker, David Harding等。2010年。“ICESat-2 激光测高任务。”《IEEE 论文集》98（5）：735–51。https://doi.org/10.1109/jproc.2009.2034765。Atlantic Inc. 2015年。“NY_WarrenWashingtonEssex_Spring2015。”ftp://ftp.gis.ny.gov/elevation/LIDAR/。Baccini, A, N Laporte, SJ Goetz,MSun,andHDong.2008年。“基于卫星影像绘制的热带非洲首幅地上生物量地图。”《环境研究快报》3（4）：045011。https://doi.org/10.1088/1748-9326/3/4/045011。Banskota,Asim, Nilam Kayastha,MichaelJ. Falkowski, Michael A. Wulder, RobertE. Froese, and Joanne C. White. 2014年。“利用Landsat时间序列数据进行森林监测：综述。”《加拿大遥感杂志》40（5）：362–84。https://doi.org/10.1080/07038992.2014.987376。Bechtold,William A,andPaul L Patterson. 2005年。《增强型森林清查与分析计划——国家抽样设计与估算程序》。第80卷。美国农业部林务局南方研究站。https: //doi.org/10.2737/SRS-GTR-80。Beven, KeithJ., andMike J. Kirkby. 1979年。“基于物理机制的流域水文可变贡献面积模型。”《水文科学通报》24（1）：43–69。https://doi.org/10.1080/02626667909491834。Bottou, Léon,andChih-JenLin. 2007年。“支持向量机求解器。”《大规模核机器》3（1）：301–20。https://web2.qatar.cmu.edu/~gdicaro/10315-Fall19/additional/SVM-solvers.pdf。Breiman,Leo. 2001年。“随机森林。”《机器学习》45（1）：5–32。https://doi.org/10.1023/A:1010933404324。Brown,Jesslyn F., Heather J.Tollerud, Christopher P. Barber, Qiang Zhou, John L. Dwyer, James E.Vogelmann, Thomas R.Loveland等。2020年。“实施美国国家土地变化持续监测业务能力的经验教训：土地变化监测、评估与预测（LCMAP）方法。”《环境遥感》238：111356。https://doi.org/10.1016/j.rse.2019.111356。Buendia, E, KTanabe, AKranjc, J Baasansuren, M Fukuda, S Ngarize, A Osako, Y Pyrozhenko, PShermanau, andSFederici. 2019年。“《2006年IPCC国家温室气体清单指南》修订。”IPCC：瑞士日内瓦 5：194。CEC.1997年。《北美生态区域：迈向共同视角》。环境合作委员会（魁北克省蒙特利尔）；秘书处。Charnley, Susan,David Diaz,andHannahGosnell.2010年。“通过美国小规模林业减缓气候变化：机遇与挑战。”《小规模林业》9（4）：445–62。https://doi.org/10.1007/s11842-010-9135-x。Chen, JingM.1996年。“植被指数和修正简单比值在北方地区应用中的评估。”《加拿大遥感杂志》22（3）：229–42。https://doi.org/10.1080/07038992.1996.10855178。Cocke, AllisonE., Peter Z.Fulé,andJoseph E.Crouse. 2005年。“利用差值归一化燃烧比和地面数据评估燃烧严重程度的比较。”《国际野火杂志》14：189–98。https://doi.org/10.1071/WF04010。Cohen, Warren B., Zhiqiang Yang, Stephen V. Stehman, Todd A. Schroeder,David M. Bell, Jeffrey G. Masek, Chengquan Huang, and Garrett W.Meigs.2016年。“1985–2012年美国本土森林扰动：森林衰退日益占主导地位。”《森林生态与管理》360（一月）：242–52。https://doi.org/10.1016/j.foreco.2015.10.042。Cooke,William H. 2000年。“利用Landsat专题制图仪数据对乔治亚州进行森林/非森林分层。”载于 McRoberts, Ronald e.; Reams, Gregory a.; vanDeusen, Paul c., Eds.《第一届年度森林清查与分析研讨会论文集》；《普通技术报告》NC-213。明尼苏达州圣保罗：美国农业部林务局北方中央研究站：28-30。Cortes, Corinna,and VladimirVapnik. 1995年。“支持向量网络。”《机器学习》20（3）：273–97。https://doi.org/10.1007/bf00994018。DanielBaston. 2022年。Exactextractr：使用多边形从栅格数据集中快速提取。https://CRAN.R- project.org/package=exactextractr。Desrochers, Madeleine L, Wayne Tripp,Stephen Logan,Eddie Bevilacqua, LucasJohnson,andColin M Beier. 2022年。“在美国东北部经营林中对森林变化检测算法进行实地验证。”《林业杂志》120（5）：575–87。https://doi.org/10.1093/jofore/fvab075。Dormann,Carsten F., Justin M. Calabrese, Gurutzeta Guillera-Arroita, Eleni Matechou, VolkerBahn, Kamil Barto´n, ColinM.Beale 等。2018年。“生态学中的模型平均：贝叶斯、信息论和战术方法用于预测推断的综述。”《生态学专论》88（4）：485–504。https://doi.org/10.1002/ec m.1309。Dubayah, Ralph, Scott J Goetz,James Bryan Blair, TE Fatoyinbo, Matthew Hansen, Sean PHealey, MichelleAHofton 等。2014年。“全球生态系统动力学调查。”载于《AGU秋季会议摘要》，2014：U14A–07。Duncanson, LI,KONiemann, and MA Wulder.2010年。“整合GLAS和LandsatTM数据估算地上生物量。”《加拿大遥感杂志》36（2）：129–41。https://doi.org/10.5589/m10-037。

Dyer, James M. 2006. “重访北美东部的落叶林。” BioScience 56 (4): 341–52. https://doi.org/10.1641/0006-3568(2006)56%5B341:RTDFOE%5D2.0.CO;2. 地球资源观测与科学 (EROS) 中心. 2017年. “国家农业影像计划（NAIP）.” 美国地质调查局. https://doi.org/10.5066/F7QN651G. Fargione, Joseph E., Steven Bassett, Timothy Boucher, Scott D. Bridgham, Richard T. Conant, Susan C. Cook- Patton, Peter W. Ellis, et al. 2018年. “美国的自然气候解决方案。” Science Advances 4 (11). https://doi.org/10.1126/sciadv.aat1869. Friedman, Jerome H. 2002. “随机梯度提升。” Computational Statistics and Data Analysis 38 (4): 367–78. https://doi.org/10.1016/S0167-9473(01)00065-2. FWS. 2022年. 国家湿地清单. 美国内政部, Fish; Wildlife Service, 华盛顿, D.C. https://fwsprimary.wim.usgs.gov/wetlands/apps/wetlands-mapper/. Gillenwater, Michael, Derik Broekhoff, Mark Trexler, Jasmine Hyman, and Rob Fowler. 2007. “监管自愿碳市场。” Nature Climate Change 1 (711): 85–87. https://doi.org/10.1038/climate.2007.58. Gorelick, Noel, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. 2017年. “Google Earth Engine：人人可用的行星尺度地理空间分析。”《环境遥感》 202: 18–27. https://doi.org/10.1016/j.rse.2017.06.031. Gray, Andrew N, Thomas J Brandeis, John D Shaw, William H McWilliams, and Patrick Miles. 2012. “美国森林清查与分析数据库（FIA）。” Biodiversity and Ecology 4: 225–31. https://doi.org/10.7809/b-e.00079. Hansen, Matthew C., and Thomas R. Loveland. 2012. “使用Landsat数据进行大面积土地覆盖变化监测综述。”《环境遥感》 122 (July): 66–74. https://doi.org/10.1016/j.rse.2011.08.024. Heath, Linda S, Mark Hansen, James E Smith, Patrick D Miles, and Brad W Smith. 2009. “使用生物量扩展因子方法计算FIADB中树木生物量和碳的调查。” In Proceedings of the Forest Inventory and Analysis (FIA) Symposium, 21–23. Hijmans, Robert J. 2022年. Terra: 空间数据分析. https://CRAN.R-project.org/package=terra. Hoppus, Michael, and Andrew Lister. 2005年. “使用全球定位系统精确定位森林清查与分析样地的现状。” In Proceedings of the Seventh Annual Forest Inventory and Analysis Symposium. https://www.nrs.fs.fed.us/pubs/7040. Houghton, R. A. 2005年. “地上森林生物量与全球碳平衡。” Global Change Biology 11 (6): 945–58. https://doi.org/10.1111/j.1365-2486.2005.00955.x. Houghton, R. A., Forrest Hall, and Scott J. Goetz. 2009. “生物量在全球碳循环中的重要性。”《地球物理研究杂志：生物地球科学》 114 (G2): n/a–. https://doi.org/10.1029/2009jg000935. Houghton, R. A., J. I. House, J. Pongratz, G. R. van der Werf, R. S. DeFries, M. C. Hansen, C. Le Quéré, and N. Ramankutty. 2012. “土地利用和土地覆盖变化的碳排放。” Biogeosciences 9 (12): 5125–42. https://doi.org/10.5194/bg-9-5125-2012. Huang, Wenli, Katelyn Dolan, Anu Swatantran, Kristofer Johnson, Hao Tang, Jarlath O’Neil-Dunne, Ralph Dubayah, and George Hurtt. 2019年. “美国马里兰州、宾夕法尼亚州和特拉华州三州地区森林碳监测系统地上生物量高分辨率制图。” Environmental Research Letters 14 (9): 095002. https://doi.org/10.1088/1748-9326/ab2917. Hudak, Andrew T, Patrick A Fekety, Van R Kane, Robert E Kennedy, Steven K Filippelli, Michael J Falkowski, Wade T Tinkham, et al. 2020年. “美国西北部区域年度地上生物量制图的碳监测系统。” Environmental Research Letters 15 (9): 095003. https://doi.org/10.1088/1748-9326/ab93f9. Johnson, Lucas K., Michael J. Mahoney, Eddie Bevilacqua, Stephen V. Stehman, Grant M. Domke, and Colin M. Beier. 2022年. “使用激光雷达覆盖时空拼贴进行高分辨率景观尺度生物量制图。” International Journal of Applied Earth Observation and Geoinformation 114: 103059. https://doi.org/10.1016/j.jag.2022.103059. Jordan, Carl F. 1969年. “从森林地面光质量推导叶面积指数。” Ecology 50 (4): 663–66. https://doi.org/10.2307/1936256. Kaarakka, Lilli, Meredith Cornett, Grant Domke, Todd Ontl, and Laura E. Dee. 2021. “改进的森林管理作为自然气候解决方案：综述。” Ecological Solutions and Evidence 2 (3). https://doi.org/10.1002/2688- 8319.12090. Karatzoglou, Alexandros, Alex Smola, Kurt Hornik, and Achim Zeileis. 2004. “Kernlab – R中核方法的S4包。”《统计软件杂志》 11 (9): 1–20. https://doi.org/10.18637/jss.v011.i09. Kauth, Richard J., and G. S. P. Thomas. 1976年. “穗帽变换 - Landsat所见农作物光谱-时间发展的图形描述。” In Symposium on Machine Processing of Remotely Sensed Data. Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017年. “LightGBM：高效的梯度提升决策树。” In Advances in Neural Information Processing Systems, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett. Vol.

30. Curran Associates, Inc.
https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa- Paper.pdf. Kelly, Matthew C, René HGermain, and Stephen VStehman. 2015年。“家庭森林所有者对森林保护项目的偏好：纽约州案例研究。”《森林科学》61 (3):597–603.https://doi.org/10.5849/forsci.1 3-120.Kennedy, Robert E, Janet Ohmann, Matt Gregory, Heather Roberts, ZhiqiangYang, David M Bell, VanKane,等。2018年。“经验性综合森林生物量监测系统。”《环境研究快报》13(2):025004. https://doi.org/10.1088/1748-9326/aa9d9e. Kennedy, Robert E, ZhiqiangYang, and WarrenB. Cohen. 2010年。“利用年度Landsat时间序列检测森林扰动与恢复趋势：1. LandTrendr——时间分割算法。”《环境遥感》114 (12): 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008. Kennedy, RobertE, Zhiqiang Yang, Noel Gorelick, Justin Braaten, Lucas Cavalcante, WarrenB.Cohen, and Sean Healey. 2018年。“在GoogleEarthEngine上实现LandTrendr算法。”《遥感》10 (5)。
https://doi.org/10.3390/rs10050691. Kerchner, CharlesD.,andWilliam S. Keeton.2015年。“加利福尼亚州监管性森林碳市场：对东北部土地所有者的可行性。”《森林政策与经济》50(1月): 70–81.
https://doi.org/10.1016/j.forpol.2014.09 .005.Kosiba, Alexandra M., Garrett W. Meigs, James A. Duncan, Jennifer A. Pontius, William S. Keeton, andEmmaR. Tait. 2018年。“美国东北部森林损害与扰动的时空格局：2000–2016年。”《森林生态与管理》430 (12月): 94–104. https://doi.org/10.1016/j.foreco.2018.07.047. Kriegler, Frank J., William A. Malila, Richard F.Nalepka,andW.Richardson. 1969年。“预处理变换及其对多光谱识别的影响。”载于。Lorimer, Craig G. 2001年。“北美东部森林中扰动的历史与生态作用：9,000年的变化。”《野生动物学会公报》（1973-2006）29 (2): 425–39. http://www.jstor.org/stable/3784167. Lu, Dengsheng.2005年。“利用Landsat TM数据估算巴西亚马逊地区的地上生物量。”《国际遥感杂志》26 (12):2509–25.
https://doi.org/10.1080/01431160500142145. MacLean, Meghan Graham, Matthew J. Duveneck, Joshua Plisinski, Luca L. Morreale, Danelle Laflower, andJonathan R. Thompson. 2021。“森林碳轨迹：新英格兰不同土地利用情景的后果。”《全球环境变化》69 (7月): 102310. https://doi.org/10.1016/j.gloenvcha.2021.102310. Mahoney, MichaelJ., ColinM. Beier, and Aidan C.Ackerman.2022年。“terrainr：用于创建沉浸式虚拟环境的R软件包。”《开放源码软件期刊》7 (69): 4060. https://doi.org/10.21105/joss.04060. Mahoney, Michael J, Lucas K Johnson, Abigail ZGuinan, andColin M Beier.2022年。“美国东北部农业后景观中低矮灌丛土地覆盖类型的分类与制图。”《国际遥感杂志》43 (19-24):7117–38. https://doi.org/10.1080/01431161.2022.2155086. Malmsheimer, Robert W., Patrick Heffernan, Steve Brink, Douglas Crandall, Fred Deneke, ChristopherGalik, Edmund Gee,等。2008年。“美国应对气候变化的森林管理解决方案。”《林业杂志》106(3): 115–17.https://doi.org/10.1093/jof/106.3.115. Matasci, Giona, Txomin Hermosilla, Michael A. Wulder, Joanne C. White, Nicholas C. Coops, Geordie W. Hobart, Douglas K. Bolton, Piotr Tompalski, andChristopherW.Bater. 2018年。“利用Landsat时间序列和激光雷达样地研究加拿大森林生态系统三十年的森林结构动态。”《环境遥感》216(10月):697–714. https://doi.org/10.1016/j.rse.2018.07.024.McRoberts, Ronald E. 2011年。“基于卫星影像的地图：科学推断还是漂亮图片？”《环境遥感》115(2): 715–24. https://doi.org/10.1016/j.rse.2010.10.013. Menlove,James, andSean P.Healey.2020年。“美国综合森林生物量数据集支持对遥感生物量估算进行定制化验证。”《遥感》12 (24)。
https://doi.org/10.3390/rs12244141. Merenlender,A. M.,L.Huntsinger, G.Guthey, andS.K. Fairfax.2004年。“土地信托与保护地役权：谁在为谁保护什么？”《保护生物学》18 (1): 65–76.https://doi.org/10.1111/j.1523-1739.2004.00401.x.Meyer, Hanna,and EdzerPebesma.2021。“预测未知空间？估算空间预测模型的适用区域。”《生态学与进化方法》12 (9):1620–33. https://doi.org/10.1111/2041-210x.13650. Montero, David, Cesar Aybar, Miguel D.Mahecha, and Sebastian Wieneke. 2022年。“Spectral：通过Google Earth EngineJavaScriptAPI部署的优秀光谱指数。”《国际摄影测量、遥感与空间信息科学档案》XLVIII-4/W1-2022: 301–6. https://doi.org/10.5194/isprs-archives- XLVIII-4-W1-2022-301-2022. Omernik,JamesM., and Glenn E.Griffith.2014年。“美国本土生态区：分层空间框架的演变。”《环境管理》54 (6):1249–66.https://doi.org/10.1007/s00267- 014-0364-1. Pataki, George E, and John P Cahill. 1999年。“High Peaks Wilderness Complex单元管理计划。”

Patton, Ry M、Diane H Kiernan、Julia I Burton 和 John E Drake。2022年。“阿迪朗达克中部森林碳储量、固存速率与结构复杂性之间的管理权衡。”《森林生态与管理》525：120539。https://doi.org/10.1016/j.foreco.2022.120539。Pebesma, Edzer。2018年。“R 的简单要素：为空间矢量数据提供标准化支持。”《R 期刊》10（1）：439–46。https://doi.org/10.32614/RJ-2018-009。Pengra, Bruce、Steve Stehman、Josephine A Horton、Daryn J Dockter、Todd A Schroeder、杨志强、Alex J Hernandez 等。2020年。“LCMAP 参考数据产品 1984-2018 土地覆盖、土地利用和变化过程属性（1.2 版，2021年11月）。”https://doi.org/10.5066/P9ZWOXJ7。PRISM 气候组。2022年。“PRISM 气候数据。”https://prism.oregonstate.edu。Quegan, Shaun、Thuy Le Toan、Jerome Chave、Jorgen Dall、Jean-Francois Exbrayat、Dinh Ho Tong Minh、Mark Lomas 等。2019年。“欧洲航天局 BIOMASS 任务：从太空测量森林地上生物量。”《环境遥感》227：44–60。https://doi.org/10.1016/j.rse.2019.03.032。R 核心团队。2021年。R：统计计算语言与环境。奥地利维也纳：统计计算 R 基金会。https://www.R-project.org/。Riemann, Rachel、Barry Tyler Wilson、Andrew Lister 和 Sarah Parks。2010年。“利用美国林务局森林清查与分析（FIA）数据对森林特征连续地理空间数据集进行有效评估的评估规程。”《环境遥感》114（10）：2337–52。https://doi.org/10.1016/j.rse.2010.05.010。Rosenqvist, A.、M. Shimada、N. Ito 和 M. Watanabe。2007年。“ALOS PALSAR：全球环境监测先导任务。”《IEEE 地球科学与遥感汇刊》45（11）：3307–16。https://doi.org/10.1109/tgrs.2007.901027。Roy, D. P.、V. Kovalskyy、H. K. Zhang、E. F. Vermote、L. Yan、S. S. Kumar 和 A. Egorov。2016年。“Landsat-7 至 Landsat-8 反射波长和归一化差异植被指数连续性的表征。”《环境遥感》185（11月）：57–70。https://doi.org/10.1016/j.rse.2015.12.024。Saarela, Svetlana、Sören Holm、Anton Grafström、Sebastian Schnell、Erik Næsset、Timothy G. Gregoire、Ross F. Nelson 和 Göran Ståhl。2016年。“利用三种信息来源进行森林清查的基于分层模型的推断。”《森林科学年刊》73（4）：895–910。https://doi.org/10.1007/s13595-016-0590-1。Shi, Yu、Guolin Ke、Damien Soukhavong、James Lamb、Qi Meng、Thomas Finley、Taifeng Wang 等。2022年。LightGBM：轻量梯度提升机。https://github.com/Microsoft/LightGBM。Simard, Marc、Naiara Pinto、Joshua B Fisher 和 Alessandro Baccini。2011年。“利用星载激光雷达全球绘制森林冠层高度。”《地球物理研究杂志：生物地球科学》116（G4）。https://doi.org/10.1029/2011 JG001708。Skowronski, Nicholas S 和 Andrew J Lister。2012年。“激光雷达对大面积森林清查应用的效用。”载于：Morin, Randall s.；Liknes, Greg c.，编。《从现状走向趋势：森林清查与分析（FIA）研讨会 2012》；2012年12月4–6日；马里兰州巴尔的摩。NRS-p-105 通用技术报告。宾夕法尼亚州纽敦斯克韦尔：美国农业部、林务局、北方研究站。*[光盘]*：410-413，410–13。https://www.fs.usda.gov/treesearch/pubs/42792。Stehman, Stephen V. 和 Giles M. Foody。2019年。“土地覆盖产品严格精度评估中的关键问题。”《环境遥感》231（9月）：111199。https://doi.org/10.1016/j.rse.2019.05.018。Strunk, Jacob L.、Hailemariam Temesgen、Hans-Erik Andersen 和 Petteri Packalen。2014年。“利用野外样地、Landsat 和激光雷达条带样本预测森林属性。”《摄影测量工程与遥感》80（2）：143–50。https://doi.org/10.14358/pers.80.2.143-150。Sugarbaker, Larry J.、Eric W. Constance、Hans Karl Heidemann、Allyson L. Jason、Vicki Lukas、David L. Saghy 和 Jason M. Stoker。2014年。“三维高程计划倡议：行动呼吁。”美国地质调查局。https://doi.org/10.3133/cir1399。Sugarbaker, Larry J.、Diane F. Eldridge、Allyson L. Jason、Vicki Lukas、David L. Saghy、Jason M. Stoker 和 Diana R. Thunen。2017年。“三维高程计划现状，2015年。”美国地质调查局。https://doi.org/10.3133/ofr20161196。Torres, Ramon、Paul Snoeij、Dirk Geudtner、David Bibby、Malcolm Davidson、Evert Attema、Pierre Potin 等。2012年。“全球环境与安全监测计划 Sentinel-1 任务。”《环境遥感》120（5月）：9–24。https://doi.org/10.1016/j.rse.2011.0 5.028。Urbazaev, Mikhail、Christian Thiel、Felix Cremer、Ralph Dubayah、Mirco Migliavacca、Markus Reichstein 和 Christiane Schmullius。2018年。“在墨西哥整合野外测量、机载激光雷达以及合成孔径雷达和光学卫星数据估算森林地上生物量及其不确定性。”《碳平衡与管理》13（1）。https://doi.org/10.1186/s13021-018-0093-5。美国地质调查局。2019年。“三维高程计划 1米分辨率数字高程模型。”https://www.usgs .gov/the-national-map-data-delivery。美国人口普查局。2013年。“TIGER/Line 形状文件。”https://www.census.gov/geographies/mapping-files/time- series/geo/tiger-line-file.html。

USFS。2006年。《2006年美国森林虫害与病害状况》。美国农业部林务局。https://www.fs.usda.gov/foresthealth/publications/ConditionsReport_2006.pdf。———。2020年。《纽约州森林，2019年》。美国农业部，林务局；美国农业部，林务局，北方研究站。https://doi.org/10.2737/fs-ru-250。美国地质调查局。2018年。《Landsat数据集》。美国地质调查局。https://doi.org/10.3133/fs20183049。Walker, Kyle。2022年。Tigris：加载人口普查TIGER/Line形状文件。https://CRAN.R-project.org/package=tigris。White, Joanne C.、Michael A. Wulder、Txomin Hermosilla、Nicholas C. Coops 和 Geordie W. Hobart。2017年。《利用Landsat时间序列对加拿大25年森林扰动与恢复进行全国年度表征》。《环境遥感》194（6月）：303–21。https://doi.org/10.1016/j.rse.2017.03.035。Whitney, Gordon G.。1994年。从沿海荒野到果实平原：1500年至今温带北美环境变化史。剑桥，英国：剑桥大学出版社。Widmann, RichardH.。2016年。《纽约州森林，2015年》。美国农业部，林务局；美国农业部，林务局，北方研究站。https://doi.org/10.2737/fs-ru-96。Widmann, Richard H.、Sloane Crawford、CharlesBarnett、Brett J.Butler、Grant M. Domke、Douglas M.Griffith、MarkA. Hatfield等。2012年。《纽约州森林，2007年》。美国农业部，林务局；美国农业部，林务局，北方研究站。https://doi.org/10.2737/nrs-rb-65。Wilen,B.O. 和Bates,M. K.。1995年。《美国鱼类及野生动物管理局的国家湿地清查项目》。载于《世界湿地分类与清查》，153–69。施普林格荷兰。https://doi.org/10.1007/978- 94-011-0427-2 13. Wintle, B. A.、M. A. McCarthy、C. T. Volinksy 和R. P. Kavanagh。2003年。《利用贝叶斯模型平均更好地表示生态模型中的不确定性》。《保护生物学》17（6）：1579–90。https://doi.org/10.111 1/j.1523-1739.2003.00614.x。Wolpert, David H.。1992年。《堆叠泛化》。《神经网络》5（2）：241–59。https://doi.org/10.1016/S0893-6080(05)80023-1。Woodall,Christopher W.、John W.Coulston、Grant M. Domke、Brian F. Walters、David N. Wear、James E.Smith、Hans-ErikAndersen等。2015年。《美国森林碳核算框架：存量与存量变化，1990-2016年》。美国农业部，林务局，北方研究站。https://doi.org/10.2737/nrs-gtr-154。Woodall, Christopher W.、Linda S. Heath、Grant M.Domke和 MichaelC.Nichols。2011年。《用于估算2010年美国森林清查中树木地上材积、生物量和碳的方法与方程》。美国农业部，林务局，北方研究站。https://doi.org/10.2737/nrs-gtr-88。Wright,Marvin N.和Andreas Ziegler。2017年。《ranger：在 C++ and R中针对高维数据快速实现随机森林》。《统计软件杂志》77（1）：1–17。https://doi.org/10.18637/jss.v077.i01。Wulder, Michael A.、David P.Roy、VolkerC. Radeloff、Thomas R. Loveland、Martha C. Anderson、David M. Johnson、SeanHealey等。2022年。《Landsat科学研究与影响五十年》。《环境遥感》280：113195。https://doi.org/10.1016/j.rse.2022.113195。Wulder, Michael A.、Joanne C. White、Ross F. Nelson、ErikNæsset、Hans Ole Ørka、Nicholas C. Coops、Thomas Hilker、Christopher W. Bater和TerjeGobakken。2012年。《用于大面积森林表征的激光雷达抽样：综述》。《环境遥感》121（6月）：196–209。https://doi.org/10.1016/j.rse.2012.02.001。Zhu, Zhe 和Curtis E. Woodcock。2014年。《利用所有可用Landsat数据持续检测和分类土地覆盖变化》。《环境遥感》144：152–71。https://doi.org/10.1016/j.rse.2014.01.011。

arXiv:2304.02632v1 [stat.AP] 5 Apr 2023

补充材料

目录

补充材料 1：LandTrendR 参数.2 补充材料 2：模型开发.3 补充材料3：精度/一致性指标.6 补充材料 4：地图一致性.7 参考文献.12

补充材料 1：LandTrendr 参数

表1：16个基于 Landsat 派生的预测变量的 LandTrendrGoogle Earth Engine（LT-GEE）分割参数（Kennedy 等，2018年）。

参数	年度反射率（NBR、NDVI、SR、MSR、TC*）	干扰（YOD, MAG）
maxSegments	5	10
尖峰阈值	0.5	0.9
顶点数量超限	3	3
防止单年恢复	true	true
恢复阈值	0.25	0.75
p值阈值	0.05	0.05
最佳模型比例	0.75	0.75
所需最少观测数	6	6

表 2：LT-GEE 中的扰动预测因子（YOD、MAG）参数。

参数	值	运算符
Delta	Loss	
Sort	最新	
Year	1985-目标年份	
幅度	50	大于
持续时间	4	少于
扰动前光谱值	300	大于
最小制图单元	7	

补充材料 2：模型开发

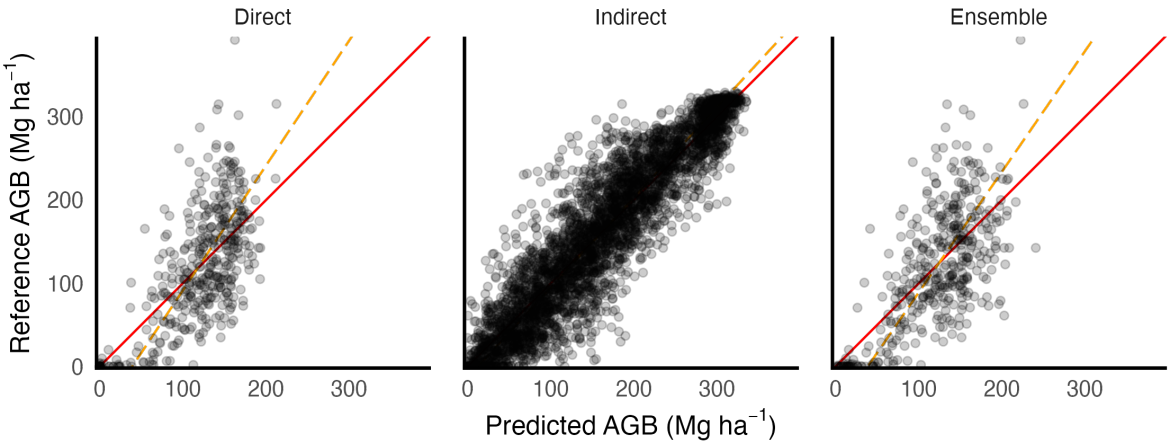


图1：模型数据集中20%测试部分的参考与预测地上生物量散点图（直接 $n = 410$；间接 $n = 4000$；集成 $n = 410$）。地上生物量值单位为 Mg ha^{-1}。几何平均函数关系（GMFR）趋势线以橙色虚线表示，1:1线以红色实线表示。

表3：模型性能指标（如第2.6节所定义），针对模型数据集的20%测试划分（直接 n = 631；间接 n = 4000；集成 n = 631）。

	直接	间接	集成
RMSE	52.50	36.91	53.37
% 均方根误差	44.19	22.28	44.92
MAE	40.23	27.51	41.75
% 平均绝对误差	33.86	16.61	35.14
ME	-0.54	-0.46	1.70
R^2	0.48	0.85	0.46

表4：最终随机森林（RF）选定的超参数。

模型	树的数量	Mtry	最小节点大小	抽样比例	有放回抽样
直接	1,500	31	2	0.85	True
间接	1,000	13	3	1	否

表5： 最终随机梯度提升机（LGB 模型的选定超参数).

模型	学习率	轮数	叶片数量	最大深度	极端随机树	叶片中的最小数据	装袋比例	装袋频率	特征比例	箱内最少数据量	L1	L2
直接	0.05	100	16	24	True	16	0.8	6	0.8	14	0.1	0.1
间接	0.10	4,000	43	24	True	3	1.0	5	0.7	15	8.0	5.0

表 6：最终支持向量机 (SVM) 的选定超参数。

模型	核函数	Type	西格玛	C	伊普西隆
直接	拉普拉斯算子	ε 支持向量回归	0.0019531	36	0.0441942

表7：直接线性模型集成。每个组成模型预测的系数，四舍五入至最接近的千分位。

Term	系数
截距	-12.223
RF	0.733
LGB	0.091
SVM	0.277

表8：间接线性模型集成。每个组成模型预测的系数，四舍五入到千分位。

Term	系数
截距	-4.067
RF	0.335
LGB	0.688

补充材料3：精度/一致性指标

$$\mathrm{RMSE} = \sqrt{(\frac{1}{n})\sum_{i=1}^{n}(y_i - \hat{y_i})^2} \tag{1}$$

$$\%\ \mathrm{RMSE} = 100 \cdot \frac{\mathrm{RMSE}}{\bar{y}} \tag{2}$$

$$\mathrm{MAE} = (\frac{1}{n})\sum_{i=1}^{n}(|y_i - \hat{y_i}|) \tag{3}$$

$$\%\ \mathrm{MAE} = 100 \cdot \frac{\mathrm{MAE}}{\bar{y}} \tag{4}$$

$$\mathrm{ME} = (\frac{1}{n})\sum_{i=1}^{n}(y_i - \hat{y_i}) \tag{5}$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6}$$

其中， n 为数据集中包含的参考观测值数量，$\hat{}y_i$ 为预测的地上生物量值， y_i 为相应的参考地上生物量值， $\bar{y}$ 为参考地上生物量平均值。

R^2 和均方根误差的标准误计算如下：

$$\mathrm{SE}_{\mathrm{boot}} = \sqrt{\frac{Var_{boot}}{n}} \tag{7}$$

其中， n 为数据集中包含的参考观测值数量， Var_{boot} 计算为对自助法重采样进行 1000次迭代所得的 R^2 和均方根误差估计值的方差。平均绝对误差和平均误差的标准误计算如下：

$$\mathrm{SE} = \sqrt{\frac{\sum_{i=1}^{n}(e_i - \bar{e})^2}{n-1}} \tag{8}$$

其中 e_i 是给定参考观测的误差， $\bar{e}$ 是所有包含的参考观测的平均误差。在平均绝对误差的情况下，$e_i = |y_i - \hat{}y_i|$，给定参考观测误差的绝对值。

补充材料 4：地图一致性

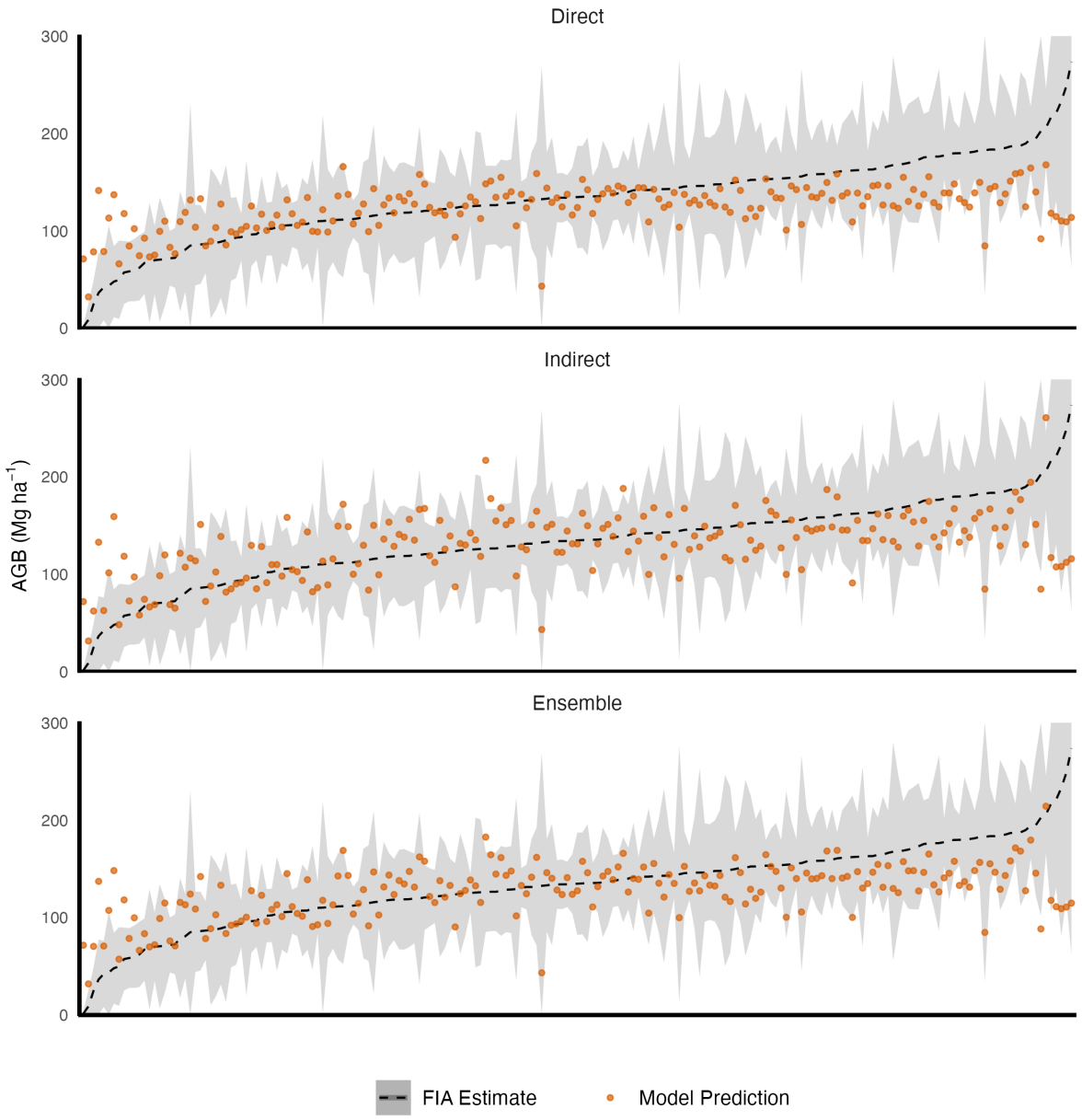


图 2：2016年地上生物量（AGB）预测与（Menlove and Healey 2020年）估计值（虚线）及其相关的95%置信区间（灰色阴影区域）在64,000公顷聚合六边形网格内的比较。FIA估计值按LCMAP 2016年树木覆盖度、湿地和草地/灌丛分类像元所指示的森林覆盖比例进行缩放。总面积中有 $< 50\%$位于纽约州边界外的六边形网格被排除在本分析之外。排除了一个大的异常值（466.51 Mg ha^{-1}），并且为了显示目的，将上限CI边界限制在300 Mg ha^{-1} 。观测值按FIA估计值沿x轴递增排序。

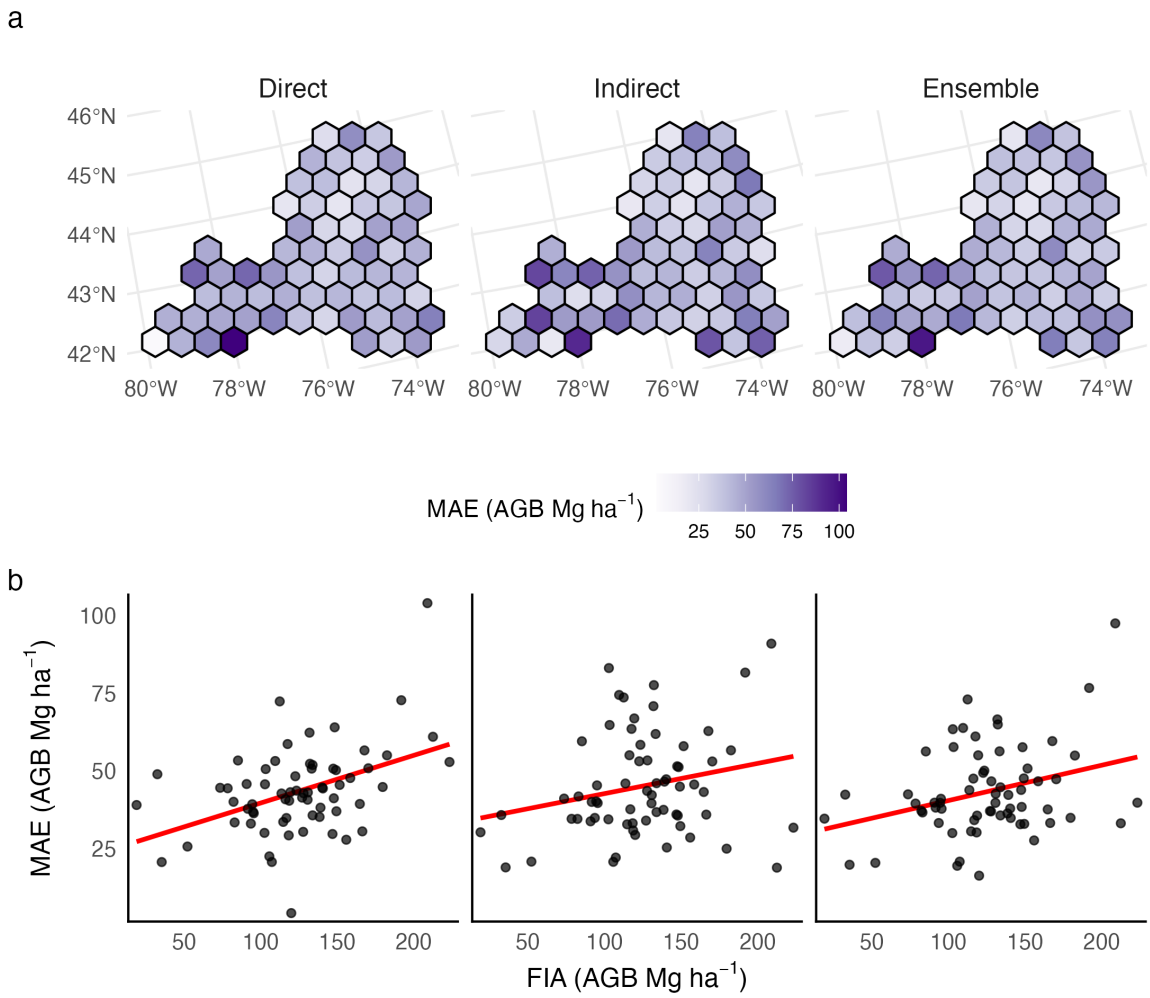


图3：空间平均绝对误差（MAE）。a)根据样地到像元残差计算的平均绝对误差（MAE）（Mgha^{-1}），并汇总至间隔50公里的单元。仅含一个参考样地的六边形网格已移除。b)六边形网格级平均绝对误差（MAE）与平均FIA地上生物量的关系。趋势线采用普通最小二乘回归生成。

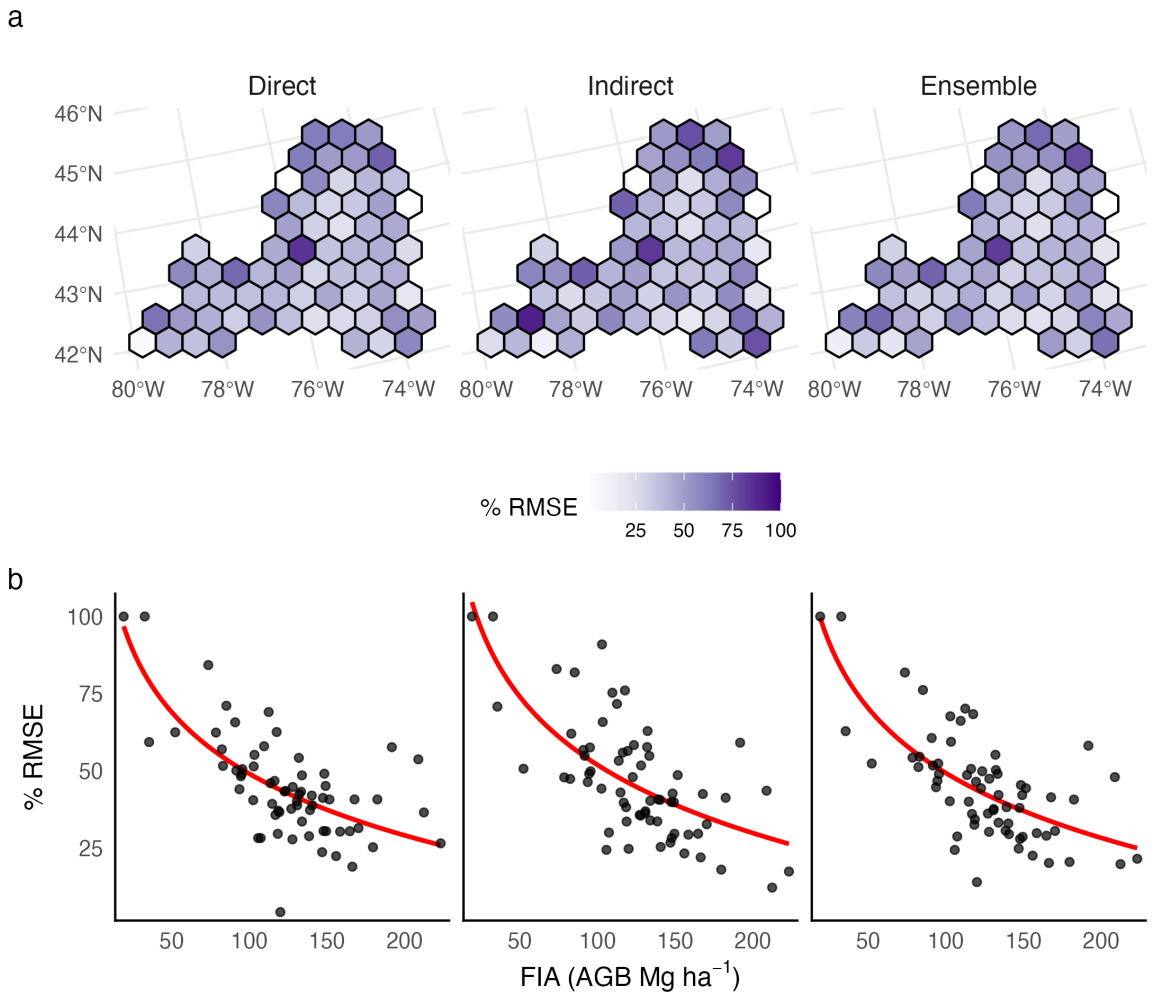


图4：空间百分比均方根误差（RMSE）。a）根据在间隔50公里的单位处汇总的样地到像元残差计算得出的百分比RMSE。移除了仅包含一个参考样地的六边形网格。b）六边形网格层面的百分比RMSE与平均FIA地上生物量的关系。趋势线使用对数回归生成。出于显示目的，a）和b）中的百分比RMSE值均封顶为100。

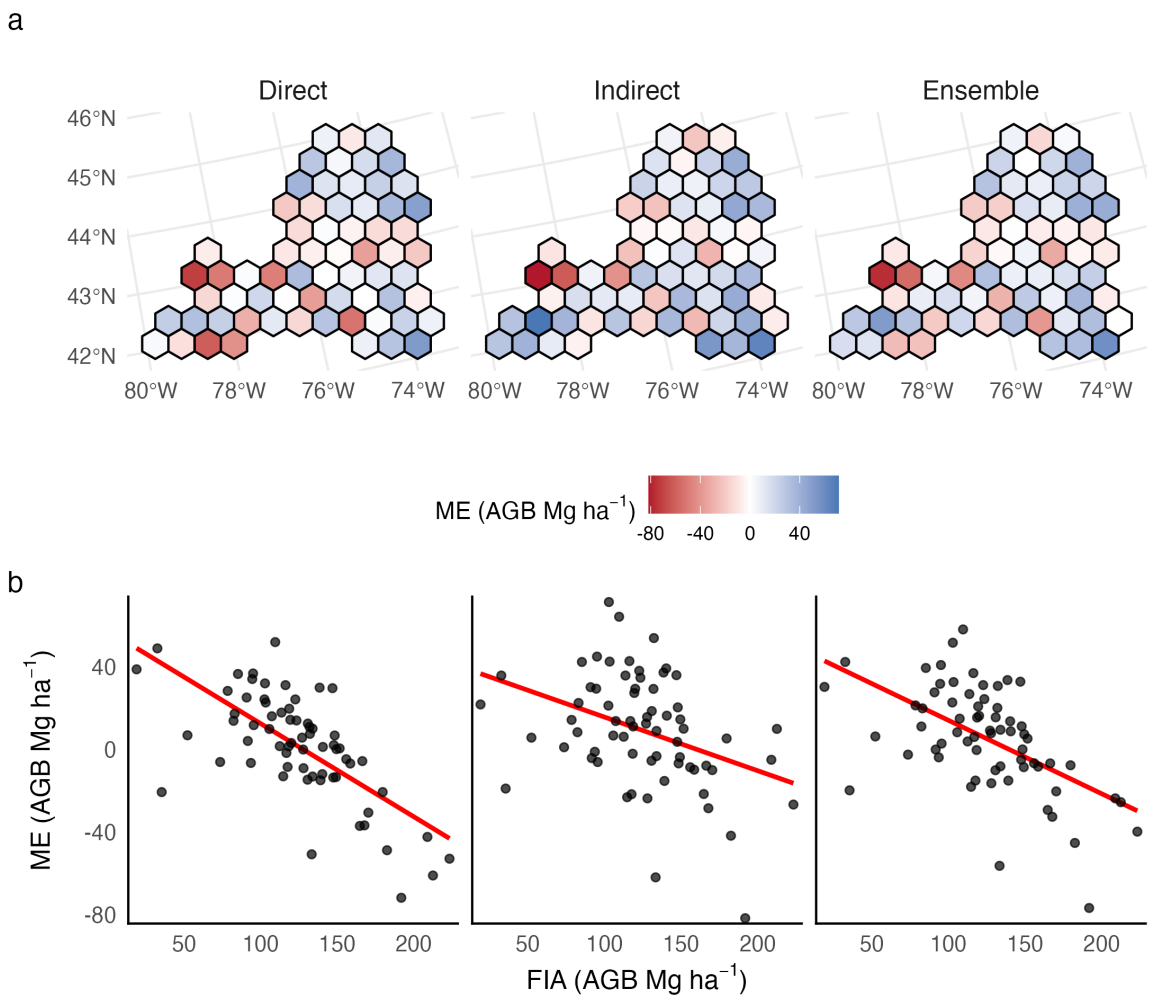


图5：空间平均误差（平均误差）。a) 平均误差（Mg ha^{-1}）由间隔50公里的单元汇总的样地到像元残差计算得出。仅含一个参考样地的六边形网格已被移除。b) 六边形网格级平均误差作为平均FIA地上生物量的函数。趋势线使用普通最小二乘回归生成。

参考文献

Menlove、James 和SeanP.Healey。2020年。“一份全面的美国森林生物量数据集可对遥感生物量估算进行定制化验证。”遥感12 (24)。https://doi.org/10.3390/rs12244141。