

遥感机理与深度学习双驱动的主粮作物叶片叶绿素含量混合反演方法

沈艳艳^{1,2}, 孟冉^{2,3}, 李佳晟³, 赵平^{1,2}, 赵峰⁴, 孙瑞^{1,2}, 张红艳⁵,
倪响⁵, 路利杰^{2,3}, 刘勇⁵, 刘劼^{2,3}

1. 华中农业大学 资源与环境学院, 武汉 430070;

2. 智慧农场技术与系统全国重点实验室, 哈尔滨 150000;

3. 哈尔滨工业大学 计算学部 人工智能研究院, 哈尔滨 150000;

4. 东北林业大学 林学院, 哈尔滨 150000;

5. 北大荒信息有限公司, 哈尔滨 150000

摘要: 准确估算叶片叶绿素含量LCC (Leaf Chlorophyll Content) 对作物生理监测和精准农业管理具有重要意义。然而, 传统基于可见光—近红外冠层反射率的植被指数VI (Vegetation Index) 在LCC反演中面临诸多挑战: 一是LCC与冠层结构信号混淆, 导致光谱响应中目标信息与结构噪声高度耦合; 二是不同作物类型间的冠层结构异质性进一步加剧了VIs对结构参数的敏感性差异, 显著限制了模型的跨作物泛化能力及应用范围。鉴于此, 本研究提出一种遥感机理与深度学习双驱动的LCC反演框架, 旨在削弱冠层结构影响并提升模型在不同主粮作物间的适应性。该方法首先基于PROSAIL辐射传输模型模拟多种叶面积指数LAI (Leaf Area Index) 与LCC组合, 构建低LAI敏感性的植被指数比值特征集VIRFS (Vegetation Index Ratio Feature Set); 随后结合主动学习策略, 优化迁移学习中目标区域样本的选择过程, 在有限标注样本下实现高效模型微调。模型在中国东北(玉米、水稻、大豆)、黄淮平原(小麦)和长江流域(水稻)三大农业区的主粮作物数据集上进行了系统验证。结果表明: (1) 本研究提出的混合方法在异地主粮作物LCC反演中表现出色, 模型在不同主粮作物场景下 R^2 稳定高于0.69, RMSE低于 $4.77 \mu\text{g}/\text{cm}^2$ 。(2) 相比传统植被指数特征集VIFS (Vegetation Index Feature Set), VIRFS在最优微调条件下可显著降低LAI敏感性, 3个地区不同主粮作物的 R^2 提升0.18—0.23, RMSE降低1.85—2.51 $\mu\text{g}/\text{cm}^2$ 。(3) 融合主动学习的迁移学习策略在仅使用30%本地标注样本时, 即可实现不同主粮作物的高精度LCC反演($R^2=0.69$ —0.74, RMSE=4.98—5.76 $\mu\text{g}/\text{cm}^2$), 相较随机采样策略精度 R^2 提升0.02—0.16, RMSE降低0.05—1.42 $\mu\text{g}/\text{cm}^2$ 。综上所述, 本研究构建的耦合物理规律与数据驱动的反演框架, 能够显著提升主粮作物LCC反演的精度与鲁棒性, 为多区域、多作物LCC的无损监测提供了普适性解决方案。

关键词: 叶片叶绿素含量, 无人机多光谱遥感, 冠层结构异质性, 植被指数比值特征, 迁移学习, 主动学习

中图分类号: TP701/P2

引用格式: 沈艳艳, 孟冉, 李佳晟, 赵平, 赵峰, 孙瑞, 张红艳, 倪响, 路利杰, 刘勇, 刘劼. 2026. 遥感机理与深度学习双驱动的主粮作物叶片叶绿素含量混合反演方法. 遥感学报, 30(5): 1392–1412

Shen Y Y, Meng R, Li J S, Zhao P, Zhao F, Sun R, Zhang H Y, Ni X, Lu L J, Liu Y and Liu J. 2026. A hybrid method for major food crops leaf chlorophyll content inversion driven by remote sensing mechanisms and deep learning. National Remote Sensing Bulletin, 30(5): 1392–1412 [DOI:10.11834/jrs.20265150]

1 引言

叶片叶绿素含量LCC (Leaf Chlorophyll Content) 是表征作物光合作用能力和氮素营养状态的重要

指标, 实现其精准反演对农业监测而言具有重要意义 (Croft 等, 2017)。传统的LCC测定方法主要依赖破坏性采样与实验室化学分析, 过程繁琐、成本较高, 且难以实现大尺度、连续动态监测。

收稿日期: 2025-05-28; 预印本: 2026-01-26

基金项目: 国家自然科学基金(编号:42471362); 黑龙江省重点研发计划(编号:2022ZX01A25, JD2023GJ01)

第一作者简介: 沈艳艳, 研究方向为农业遥感。E-mail: shenyy_rs@163.com

通信作者简介: 孟冉, 研究方向为农业遥感。E-mail: mengran@hit.edu.cn

遥感技术的发展为LCC的快速且非破坏性反演提供了可行路径。通过分析作物冠层的光谱响应特征,可以实现LCC的定量估测,为农业信息获取提供了高效手段。特别是近年来无人机平台的广泛应用,显著提升了高时空分辨率光谱数据的获取效率,为作物生理参数的精细化估算提供了技术支撑(Jiang等, 2022; Pan等, 2024; Zhang等, 2024)。

当前LCC遥感反演方法主要分为经验方法、物理模型方法和混合方法。其中,经验方法主要基于冠层反射光谱与LCC之间的统计相关性,为LCC估算提供了一种简便直接的解决方案。最基本的经验方法是构建植被指数,如NDVI、CSI等(谷晨鹏等, 2024; Zhang等, 2022),它们以选定波段间的比值或差值形式表征光谱特征。随着计算能力的提升,基于全光谱信息的多元统计方法和机器学习技术也被广泛应用于LCC反演,如偏最小二乘回归、高斯过程、随机森林及卷积神经网络(Verrelst等, 2015; Zhen等, 2021)。尽管这些方法在特定数据集上表现出较高的精度与稳健性,但其建模过程高度依赖于大量高质量的地面实测数据。在数据获取成本高昂的背景下,这成为制约其大范围推广应用的主要瓶颈。此外,经验方法的模型结构往往缺乏明确的物理机制解释,模型的泛化能力和迁移能力也容易受到观测条件变化的影响。相比之下,基于辐射传输模型RTM(Radiative Transfer Model)的物理方法则通过模拟光与植被之间的相互作用过程,从机理层面建立光谱与LCC之间的定量关系。以PROSAIL模型为代表的RTM方法将叶片尺度的PROSPECT模型与冠层尺度的SAIL模型进行耦合,能够模拟400—2400 nm波段的冠层反射率变化,反映植被结构、生理参数及环境条件的综合影响。这些模型常通过查找表法或迭代优化技术进行参数反演(Qian和Liu, 2020; Sun等, 2021; Xu等, 2019)。然而,RTM反演过程计算复杂,容易受到“病态反演”问题影响,且在高维输入空间中需要生成大量模拟数据,导致模型训练与优化成本较高。为兼顾物理模型的可解释性与经验方法的高效性,混合方法应运而生,并逐渐成为当前LCC反演研究的热点方向。该类方法通常利用RTM(如PROSAIL)生成大规模、已知参数的模拟光谱数

据集,并将其作为训练样本用于构建基于机器学习的反演模型。通过这种方式,一方面,混合方法生成的训练数据克服了实测数据不足的问题,极大地提升了模型的稳健性与迁移性(Li等, 2023; Wan等, 2024; Zhang等, 2021);另一方面,由于训练数据来源于物理模型,其预测结果具备更强的可解释性,便于模型输出与作物生理机制之间的对接。

目前,大多数研究仍集中于单一作物(如水稻、玉米)及其特定生育时期,所构建的LCC反演模型主要依赖于在特定作物生理状态和观测条件下提取的光谱特征(Bhadra等, 2024; Li等, 2020)。这种模型虽然在特定场景下能够取得较高精度,但应用范围受限,难以满足多样化作物类型和大范围区域监测的需求。本研究聚焦于我国主要粮食作物,包括水稻、玉米、小麦和大豆,兼具代表性与数据可得性,旨在探讨不同主粮作物中的LCC反演适应性差异。作为粮食安全的核心作物,这些主粮作物在遥感监测中的精准识别与参数估测,对于指导农业管理和保障国家粮食安全具有重要意义(Zhang等, 2021; Wan等, 2023)。随着遥感监测在农业管理和粮食安全评估中的应用不断拓展,迫切需要建立能够跨作物、跨区域应用的通用型LCC反演模型。因此,如何构建具备良好跨作物适应性和环境鲁棒性的通用型LCC反演模型,已成为实现遥感监测技术规模化、实用化应用的关键问题之一。

跨作物LCC反演所面临的首要挑战在于作物冠层结构异质性所引发的光谱耦合影响。作物冠层光谱不仅受到LCC本身的影响,还受到LAI等结构参数作用,尤其是在可见光至近红外波段,LCC与LAI的光谱响应高度耦合,导致基于冠层反射率的传统反演方法易受结构变化影响,降低反演准确性(Chen等, 2022; Li等, 2025; Xu等, 2019; Zhang等, 2023)。为降低冠层结构对LCC反演精度的影响,当前研究主要从两个方向展开探索:(1)通过构建对冠层结构不敏感的植被指数以提升LCC的反演能力。例如,LICI等指数通过特定波段组合抑制结构变量的影响,从而增强对叶绿素含量的敏感性(Li等, 2020)。然而,此类指数多依赖高光谱数据构建,难以在多光谱遥感平台上直接应用,限制了其实用性与推广性。(2)通

过物理建模或机器学习方法实现结构与生化参数的解耦。在物理建模方面,如PROSAIL等辐射传输模型(RTM)可将冠层结构与生化特征参数分离建模,提高模型的可解释性(Gu等, 2025; Qian和Liu, 2020);在机器学习方面,则可通过多源特征融合来增强对结构与生化信息的辨析能力(Chen等, 2022)。但前者通常需要大量先验参数,计算复杂度高;后者则易受训练样本代表性和泛化能力的限制。近期,有研究尝试采用植被指数组合(如TCARI/OSAVI)以增强LCC反演中对冠层结构变异的鲁棒性,结果表明在冠层尺度下其与LCC表现出更高的相关性(Daughtry等, 2000; Xu等, 2019),但其跨作物的适应性仍有待验证。总体而言,现有方法多聚焦于单一作物或特定场景,仍缺乏能够有效应对多作物冠层结构差异所引起的光谱异质性的通用化反演策略。

跨作物LCC反演的另一核心挑战来自于跨作物场景下的域适应困难。不同作物之间的形态结构、物候节律以及生长环境的差异,导致其在遥感特征空间中存在明显分布偏移,进而削弱了传统机器学习模型的泛化能力。迁移学习作为应对小样本学习与域偏移问题的有效工具,近年来已在农业遥感中广泛应用。其基本思想是利用在源域任务上预训练的模型参数,通过微调适应目标域的新任务,从而提升在数据稀缺情况下的模型表现(Pan和Yang, 2010; Zhao等, 2022)。在LCC反演中,混合方法通过RTM生成大量多样化模拟样本,为迁移学习提供了丰富的训练资源,结合微调—迁移学习策略,可在一定程度上提升模型对新环境与新作物的适应能力(Bhadra等, 2024; Zhang等, 2021)。然而,当前大多数迁移学习研究集中在单一作物的时序迁移或相似生态环境下的局部迁移,较少考虑作物间光谱特征分布差异较大的情况(Li等, 2021; Bhadra等, 2024)。如何优化迁移样本的选择策略,提升特征空间对齐程度,是提高跨作物LCC反演模型鲁棒性的关键。此外,大多数研究在模型微调过程中仍采用随机样本选择策略,其样本代表性不足,影响模型性能。对此,主动学习作为一种智能样本选择策略,能够在候选样本中优先选择信息量更高、代表性更强的数据用于模型训练,从而提高模型的鲁棒性和泛化能力。尽管已有研究在单

一作物参数反演中验证了其有效性(Liu等, 2025; Wan等, 2023),但该方法在跨作物LCC反演迁移学习中的应用研究尚不充分,其泛化潜力仍有待深入挖掘。

针对上述存在问题,本研究拟以中国三大农业区的水稻,玉米,大豆和小麦为研究对象,提出一种减弱不同作物冠层结构异质性影响的跨作物LCC混合反演方法,首先基于辐射传输模型生成模拟数据集用于后续分析,其次构建并筛选减弱冠层结构异质性影响的光谱指数比值特征,最后通过微调—迁移学习策略构建耦合物理规律和深度学习共同驱动的LCC混合反演模型,并基于3个研究区内的不同作物数据集进行验证,以期提高跨作物LCC反演模型精度和适用性。

2 材料与方法

2.1 研究区及研究对象

《中国综合农业区划》划定的九大农业区不仅揭示了农业生产的地域分异规律,更凸显了主要粮食作物与地理环境的高度适配性。为了检验所提出模型在跨区域作物上的准确性与稳定性,本研究选取了三大典型农业区的核心粮食产区:东北区、黄淮海区和长江中下游区,分别为研究区1、2和3。研究区1位于东北区的鹤岗市和佳木斯市,这里的主要气候为温带大陆性季风气候。良好的土壤条件和平坦的地势使得该地区的种植业发展迅速,主要作物包括水稻、玉米和大豆。研究区2位于黄淮海区的泰安市,属于温带大陆性半湿润季风气候区。与研究区1相比,研究区2冬季寒冷但不极端,夏季温暖湿润,非常适合冬小麦的生长。研究区3位于长江中下游区的黄冈市,主要气候为亚热带大陆性季风气候,充足的太阳辐射和水资源为该地区的水稻种植提供了良好的条件。

2.2 叶片叶绿素含量地面测量数据

本研究通过3个来自不同农业区的地面观测数据集来评估所提出的混合方法在估算冠层水平LCC方面的性能。这些数据集包括大田实验(数据1)和小区实验(数据2和数据3)。数据1采集自研究区1,包括水稻、玉米和大豆。该地区农作物生长季为5—9月。本研究于2023年7月初对

3种作物的种植样地选择了73个样方进行LCC测量。每个样地内的样方间距在10 m以上,样方大小为1 m×1 m。数据2采集自研究区2的冬小麦,该地区冬小麦的生长季为2月下旬至6月上旬。本研究于2024年4月初在冬小麦种植基地选取了113个种植不同品种小麦的小区进行LCC测量,每个小区内的冬小麦品种一致,小区大小为2 m×2 m。数据3采集自研究区3的水稻,数据来自水稻田的小区实验,其中包括了不同的施肥处理。本研究于2020年和2021年采集了两年的数据。其中:2020年采集了32个小区的数据,2021年采集了64个小区的数据,每个小区的大小为5 m×5 m;在2020年的水稻灌浆期和灌浆后期,以及2021年的水稻抽穗期和灌浆期内,研究共采集了192条数据。本研究共采集了378份田间实测作物样本,包括不同农业区和不同作物(水稻,玉米,大豆和小麦)。利用这些样本评估光谱比值特征和迁移学习技术在不同田间测量条件下反演作物LCC的有效性和鲁棒性。

在数据采集过程中,研究使用便携式SPAD-502叶绿素仪测量作物叶片的相对叶绿素含量。每个样方通过五点法选择5片叶片,接着对每个叶片的上、中、下3个部位进行测量,然后取平均值作为单叶片的相对叶绿素含量值,再将所有叶片的平均值作为样方的相对叶绿素含量。LCC由SPAD-502叶绿素仪测定,并通过式(1)将SPAD测量值转换为叶绿素浓度值(Coste等,2010),以匹配辐射传输模型中的叶绿素含量值。具体计算公式如下:

$$LCC = \frac{117.1 \times SPAD}{148.84 - SPAD} \quad (1)$$

2.3 无人机多光谱影像数据

本研究使用DJI Mavic 3多光谱传感器获取了研究区1和研究区2的多光谱影像,这些影像包含4个波段:绿波段G (Green),红波段R (Red),红边波段RE (RedEdge)和近红外波段NIR (Near-infrared)。研究区3的多光谱影像获取自DJI精灵4 (P4M)无人机平台搭载的多光谱传感器,该传感器包含5个光谱波段:蓝波段(Blue, B), G, R, RE, NIR。所有无人机多光谱影像的获取时间均与地面调查数据一致,并且在无云天气条件下进行。在数据采集过程中,无人机的飞行速

度为2 m/s、旁向重叠和航向重叠均为80%。其中,数据1飞行高度为100 m、数据2飞行高度为15 m,数据3飞行高度为20 m。飞行前,在田间均匀布置了若干个地面控制点,并使用RTK测量获取控制点中心的三维坐标,以保证影像几何精度。同时,研究区放置了标准反射率板,用于后期预处理时将无人机多光谱影像的数字量化值转换为实际的地物反射率。本研究使用大疆智图软件对原始多光谱影像进行了辐射定标和影像拼接等预处理。然后通过植被指数和设定的阈值完成背景去除。基于去除背景后的多光谱影像,结合研究区样方范围,利用Python提取各样方的光谱均值用于后续研究。

2.4 基于PROSAIL模型的冠层反射率模拟数据

PROSAIL辐射传输模型被广泛用于模拟作物冠层的反射光谱。本研究使用结合了PROSPECT-5B和4SAIL的PROSAIL-5B模型,该模型能够同时考虑叶片的光谱特性和整个冠层的辐射特性。PROSAIL模型可以根据设定的参数范围生成不同的模拟数据集,本研究模型参数范围设置结合了先验知识以及先前报道的相关研究(Fernández-Guisuraga等,2023; Wan等,2024; Xu等,2019),如表1所示。然而,并非所有参数都有助于LCC的估计,不同参数之间的相互作用可能会增加不确定性。因此,本研究使用局部敏感性分析方法研究了模型性能对不同参数变化的敏感性。本研究根据某一参数的范围和步长,逐步改变该参数的值(例如,LCC的数值依次为10,20,30,40,50,60,70,80 $\mu\text{g}/\text{cm}^2$),同时保持其他参数不变,以分析该参数变化对450—800 nm波段冠层反射率的影响。结合各参数的敏感性分析结果,在生成模拟数据集时不影响冠层反射率的模型参数采用固定值,而对敏感模型参数则选用给定的范围。本研究按照随机分布方式生成6000条从400 nm到2500 nm(间隔1 nm)的冠层反射率模拟数据集,以增强稳健性和泛化能力。并根据无人机多光谱波段的中心波长对模拟数据进行平均,使其匹配无人机光谱分辨率(Zhang等,2024)。此外,为了更好地模拟实际测量条件并避免过拟合现象,将2%的高斯白噪声添加到模拟反射率数据中(García-Haro等,2018)。

表1 PROSAIL-5B模型的参数输入范围
Table 1 The input ranges of model parameters for PROSAIL-5B inversion and simulations

模型参数	单位	取值范围	步长	固定值
结构参数 N	—	1—3	0.5	1.5
叶绿素含量 LCC	$\mu\text{g}/\text{cm}^2$	0—80	10	50
棕色素含量 Cbp	$\mu\text{g}/\text{cm}^2$			0
胡萝卜素含量 Car	$\mu\text{g}/\text{cm}^2$	5—30	5	8
等效水厚度 C_w	g/cm^2			0.01
干物质含量 C_m	g/cm^2	0.002—0.020	0.003	0.010
叶面积指数 LAI	m^2/m^2	1—8	1	5
叶倾角分布 ALA	($^\circ$)	30—80	10	40
热点参数 Hspot	m/m	0.01—0.31	0.06	0.01
土壤亮度系数 Psoil		0—1	0.2	0
太阳天顶角 SZA	($^\circ$)	10—70	12	30
观测天顶角 VZA	($^\circ$)	0—50	10	0
相对方位角 RAA	($^\circ$)			90

2.5 减弱冠层结构异质性的植被指数比值特征构建

作物冠层的光谱特性受多种因素影响，包括生理生化组分、叶片与冠层结构、植株生长状况、土壤背景及天气条件等。在利用遥感数据估算 LCC 时，冠层反射信号受到 LAI 等结构参数的耦合作用，导致 LCC 信息与冠层结构噪声高度混杂，从而影响反演精度。为降低 LAI 变化对 LCC 估算的干扰，本研究基于 PROSAIL 辐射传输模型，模拟不同作物类型下的多种 LAI 与 LCC 组合数据，以提高模型在不同作物条件下的适用性和稳健性。进一步地，选取对 LCC 和 LAI 具有不同敏感度的植被指数，构建植被指数比值特征集 VIRFS (Vegetation Index Ratio Feature Set)，以减少 LAI 变化对 LCC 反演的影响。在构建比值特征时，本研究选取了对 LCC 高度敏感的植被指数（如 CIG、TCI 等）作为分子，以对 LAI 变化具有高响应但饱和特性的植被指数（如 SR、MTVI2 等）作为分母。这种构造形式可被视为将 LCC 与 LAI 在光谱信号中的混合贡献进行“物理层面解耦”，在一定程度上减少了 LAI 在特征中的影响力，从而保留了 LCC 在分子层面的主导效应，进一步达到增强 LCC 响应、抑制 LAI 干扰的效果。随后，采用 Relief-F 算法筛选出对 LCC 敏感性更高的比值特征，以提升 LCC 估算的准确性。为评估 VIRFS 的有效性，本研究同时构建了植被指数特征集 VIFS (Vegetation Index Feature

Set)，该特征集由 10 个对 LCC 敏感的单一植被指数组成。本研究将在不同作物类型下开展 LCC 反演实验，通过将 VIRFS 与 VIFS 的建模效果进行对比分析，以验证 VIRFS 在减弱冠层结构异质性影响方面的优势。VIRFS 的筛选方法包括以下 3 个步骤：

(1) 比值特征构建。基于相关文献选取 10 个对 LCC 敏感的植被指数作为分子，5 个对 LAI 敏感的植被指数作为分母 (Chen 等, 2022; Gao 等, 2024; Liang 等, 2015; Xu 等, 2019)，如表 2 所示。计算所有可能的植被指数比值，共获得 50 个比值特征（如 Datt99/SR, Datt99/LCI, …, TCARI/MTVI2, TCARI/OSAVI）。其中，作为分子的 10 个植被指数共同组成了 VIFS。

表2 参数反演过程中的敏感植被指数
Table 2 Sensitive VIs during parameter estimation

作物参数	敏感植被指数	计算公式
叶绿素含量 LCC	Datt99	$(\text{NIR}-\text{RE})/(\text{NIR}+\text{R})$
	LCI	$(\text{NIR}-\text{RE})/(\text{NIR}+\text{R})$
	RERI	$(\text{RE}-\text{R})/\text{NIR}$
	CIG	$(\text{NIR}/\text{G})-1$
	TVI	$0.5 \times (120 \times (\text{NIR}-\text{G}) - 200 \times (\text{R}-\text{G}))$
	TCI	$1.2 \times (\text{RE}-\text{G}) - 1.5 \times (\text{R}-\text{G})(\text{RE}/\text{R})^{0.5}$
	CIRE	$(\text{NIR}/\text{RE})-1$
	NDRE	$(\text{NIR}-\text{RE})/(\text{NIR}+\text{RE})$
	GNDVI	$(\text{NIR}-\text{G})/(\text{NIR}+\text{G})$
	TCARI	$3 \times ((\text{RE}-\text{R}) - 0.2 \times (\text{RE}-\text{G}) \times (\text{RE}/\text{R}))$
叶面积指数 LAI	SR	NIR/R
	EVI2	$2.5 \times (\text{NIR}-\text{R})/(\text{NIR}+2.4 \times \text{R}+1)$
	NDVI	$(\text{NIR}-\text{R})/(\text{NIR}+\text{R})$
	MTVI2	$1.5 \times (1.2 \times (\text{NIR}-\text{G}) - 2.5 \times (\text{R}-\text{G})) / ((2 \times \text{NIR}+1)^2 - (6 \times \text{NIR}-5 \times \text{R}^{0.5}) - 0.5)^{0.5}$
	OSAVI	$1.16 \times (\text{NIR}-\text{R})/(\text{NIR}+\text{R}+0.16)$

(2) 特征权重计算。利用 PROSAIL 生成的模拟数据，采用 Relief-F 算法计算 50 个比值特征与 LCC 的权重值，以衡量其对 LCC 估算的贡献度。

(3) 最优比值特征筛选。在每组具有相同分子的比值特征中，选取权重值最高的比值特征，最终构建包含 10 个比值特征的 VIRFS。

2.6 迁移学习

2.6.1 主动学习策略的实施

在迁移学习应用中，引入新的本地样本以扩

展模型适用性至关重要。然而, 传统的随机选择策略往往难以充分代表目标场景的数据分布, 影响模型性能。为此, 主动学习策略通过有针对性地选择最具信息量的样本, 成为提高样本利用效率的有效途径。在主动学习框架中, 基于不确定性的算法 (如 BALD) 通常依托贝叶斯推断, 通过计算预测信息增益来选择高不确定性的样本, 适用于模型预测方差较高的场景; 而基于特征空间覆盖的算法 (如 Core-set) 则以样本代表性为核心, 通过最小化已选样本与未选样本之间的特征空间距离, 实现特征空间的广泛覆盖。然而, 这两类方法通常仅关注输入空间或输出空间的单一维度信息, 且面临计算复杂度高、特征冗余干扰显著以及采样稳定性不足等问题 (Kim 和 Shin, 2022; Nabika 等, 2024)。相较之下, 改进的贪心采样策略 iGS (improved Greedy Sampling) 在采样过程中综合考虑样本在输入空间 (特征分布) 与输出空间 (预测响应) 上的代表性, 通过联合相似度量实现信息量与多样性的平衡。该方法无需复杂的贝叶斯推断或嵌套优化, 具有较高的计算效率与实现简便性。因此, 本研究在模型微调阶段采用主动学习中的 iGS 方法作为样本选择策略, 以提升样本选择效率与代表性 (Wu 等, 2019; Wan 等, 2023; Ma 等, 2024)。

(1) 计算地面测量数据中每个样本与训练数据集集中样本的欧氏距离, 找到每个实测样本与训练数据集中最近样本的距离。具体计算公式如下:

$$D_x(i, j) = \sqrt{\sum_{k=1}^m (x_{i,k} - x_{j,k})^2}, i = 1, 2, \dots, N; j = 1, 2, \dots, M \quad (2)$$

$$d_x(i) = \min_j D_x(i, j) \quad (3)$$

式中, $D_x(i, j)$ 表示地面测量数据中第 i 个样本与训练数据中第 j 个样本之间的距离, $x_{i,k}$ 和 $x_{j,k}$ 是样本的特征值, m 是特征的数量。 $d_x(i)$ 表示地面测量数据中第 i 个样本与训练数据中最近样本的最小距离。

(2) 使用随机森林算法基于训练数据集构建回归模型, 然后对地面测量数据进行预测, 并计算预测值与每一个训练样本目标变量之间的绝对差值, 找到每个实测样本与训练集中最近目标变量的最小绝对差。具体计算公式如下:

$$D_y(i, j) = |f(x_i) - y_j|, i = 1, 2, \dots, N; j = 1, 2, \dots, M \quad (4)$$

$$d_y(i) = \min_j D_y(i, j) \quad (5)$$

式中, $D_y(i, j)$ 表示地面测量数据中第 i 个样本的预测值 $f(x_i)$ 与训练数据中第 j 个样本的真实目标值之间的绝对差。 $d_y(i)$ 表示地面测量数据中第 i 个样本与训练数据中目标值最近的样本的最小绝对差。

(3) 将前两个步骤得到的距离相乘, 得到一个综合的评分, 根据样本选择数量依次选择具有最大评分的实测样本作为模型微调的样本数据。

$$d_{xy}(i) = d_x(i) * d_y(i), i = 1, 2, \dots, N \quad (6)$$

本研究通过该方法确保在目标场景中选择信息量较大的样本, 从而提升模型在新场景中的适应性和估算精度。为平衡采样成本与模型性能, 并确定最优微调数据量, 本研究设置微调数据比例为本底实测数据的 10%、20%、30%、40% 和 50%。此外, 为定量评估 iGS 方法与随机选择方法的效果, 研究计算样本的信息熵 (Entropy) 和变异系数 CV (Coefficient of Variation), 分析不同样本选择策略对 LCC 反演精度的影响, 以评估 iGS 在提升模型适应性和稳健性方面的优势。Entropy 是信息论中的关键指标, 用于衡量数据的不确定性和复杂性, 具体见式 (7)。在数据选择过程中, 较高的 Entropy 意味着样本集能够更充分地代表目标数据分布。CV 作为衡量数据离散程度的相对指标, 具体见式 (8), 用于评估选定样本相对于整体数据的变异程度。CV 值越大, 说明样本的分布范围越广, 信息量越丰富, 有助于提高模型的泛化能力。

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (7)$$

$$CV = \frac{SD}{Mean} \times 100 \quad (8)$$

式中, n 是随机样本值 X 可以取的不同状态的数量, $p(x_i)$ 是样本值 X 取值 x_i 的概率。SD 为数据的标准差, Mean 为数据的平均值。

2.6.2 基于微调的迁移学习方法

机器学习方法的建模训练过程可认为是知识学习过程, 而不同场景和任务的知识存在相关性, 在解决新任务场景时, 通过迁移学习方法可以将原始模型已学习的相关知识快速应用于新模型, 以提高新模型的基础认知水平 (Ma 等, 2024; Safonova 等, 2023)。本研究迁移学习的总体思路可以概括为预训练模型、微调模型和应用模型。

研究首先基于辐射传输模型生成的大量、多样化的模拟数据来预训练模型,使模型通过学习反演任务下的不同光谱数据处理使用能力,来提高模型泛化特征表示水平和鲁棒性。然后在实际LCC反演场景下,利用少量具有代表性的地面实测样本对预训练模型进行微调学习,使模型拥有新的LCC反演抽象信息表征能力,不仅减少了对实际地面测量数据的需求,降低了训练成本,同时提高了模型的LCC反演能力。最后通过实际场景应用验证了基于微调后的新模型的LCC高精度反演效果。

模型构建方法基于深度神经网络DNN (Deep Neural Networks) 实现,DNN是一种多层次的神经网络模型,它通过多个隐藏层的非线性变换来进行复杂的数据建模和特征提取。与传统的机器学习方法相比,DNN能够自动从大量数据中学习更加抽象和高级的特征,使其在各种任务中表现

出色(Yosinski等,2014; Zhang等,2021)。在本研究中,DNN采用Pytorch框架进行构建与训练,模型输入为10维特征向量,输出为1维LCC预测值。为增强DNN的特征提取能力及复杂非线性关系建模能力,同时缓解梯度消失问题,采用了隐藏层维度递增的设计策略。具体而言,模型包含5个全连接隐藏层,神经元数量依次设置为64、128、256、512和1024,每一隐藏层后均接入ReLU激活函数以引入非线性。训练过程中,batch size设定为128,采用Adam优化器进行参数更新,损失函数选择均方误差(MSE),以最小化预测值与实测LCC之间的偏差。本研究采用决定系数(R^2)、均方根误差(RMSE)和相对均方根误差(RRMSE)定量评价模型的准确性,具体公式详见Wan等(2021)。本研究采用的LCC反演方法技术路线图如图1所示。

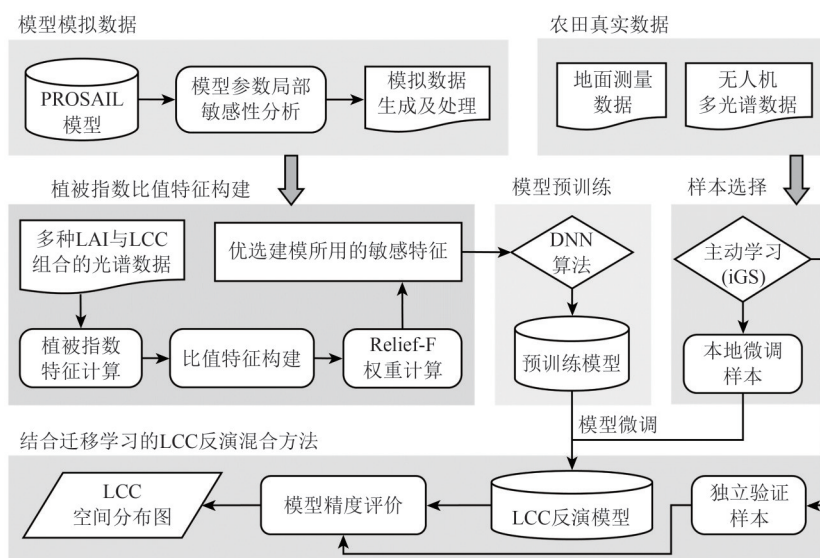


图1 LCC混合反演方法技术路线图

Fig. 1 Workflow of the hybrid method for LCC inversion

3 结果与分析

3.1 PROSAIL 模型参数敏感性

在叶片参数设置中,考虑到 C_w 对LCC敏感的波段只有边际影响,且研究对于干枯叶片忽略不计,所以 C_w 和 C_{bp} 被设置为固定值。本研究对其他叶片参数(N 、LCC、 C_{ar} 和 C_m)的敏感性分析如图2所示。可见:450—800 nm的冠层反射率对 N 和LCC的变化高度敏感,随着 N 值从1—3的增加,

冠层反射率在可见光(尤其是550 nm附近)显著降低,而在近红外波段(700—800 nm)略有上升(图2(a));LCC的升高显著降低了冠层在450—750 nm的反射率,尤其在500—700 nm区域影响最为敏感(图2(b))。 C_m 的升高对可见光波段的影响较少,但在近红外(720—800 nm)区间的冠层反射率显著降低(图2(c)); C_{ar} 含量的变化几乎不影响450—800 nm的冠层反射率(图2(d))。因此,使用PROSAIL模型生成模拟数据集时考虑了

N、LCC 和 *Cm* 的影响，使用表 1 的参数范围进行模拟数据生成，而对影响较小的 *Car* 采用固定值以减少冗余变量。在冠层参数设置中，考虑到 RAA 对 LCC 反演的影响较小，故设为固定值。本研究对其他冠层参数（LAI、ALA、Hotspot、Psoil、SZA 和 VZA）的敏感性分析如图 3 所示。可见，冠层反射率对 LAI（图 3（a））、ALA（图 3（b））和 Hspot（图 3（c））的变化较为敏感，特别是在 700—800 nm；

相比之下，Psoil（图 3（d））、SZA（图 3（e））和 VZA（图 3（f））对光谱响应的影响较弱，主要体现在整体幅度上的轻微变动，说明在本研究范围内，这些参数对目标叶绿素反演特征的干扰有限。因此，使用 PROSAIL 模型生成模拟数据集时考虑了 ALA、LAI 和 Hspot 的影响，使用表 1 的参数范围进行模拟数据生成，而对影响较小的 Psoil、SZA 和 VZA 采用固定值以简化模型输入维度，减少无效扰动。

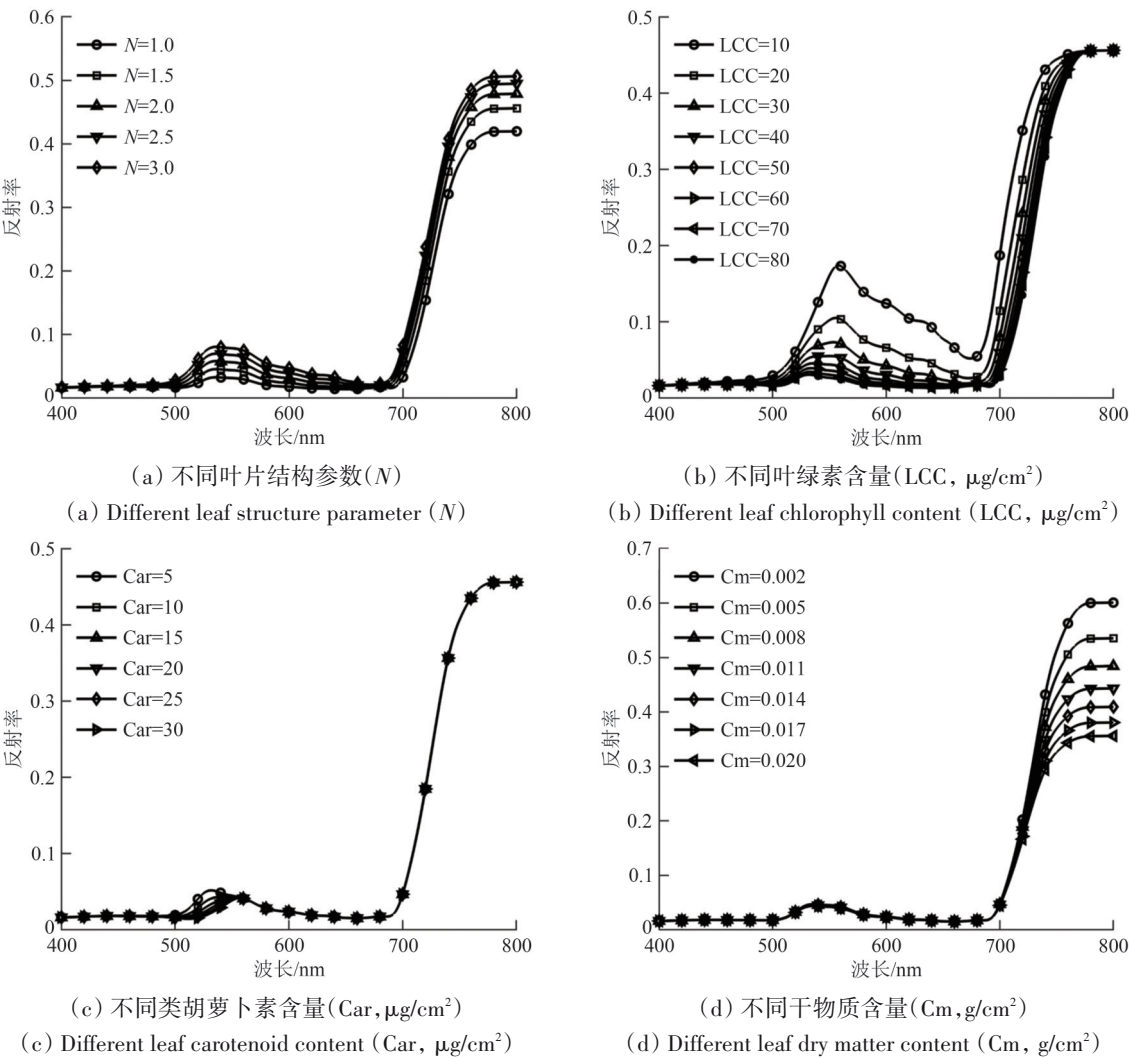


图2 450—800 nm波段冠层反射率对叶片参数变化的敏感性分析
Fig. 2 Sensitivity analysis of canopy reflectance at 450—800 nm to variations in leaf parameter

3.2 植被指数比值特征在缓解 LAI 干扰中的表现与优势分析

本研究基于辐射传输模型生成的涵盖多种 LAI 与 LCC 组合数据，对构建的植被指数比值特征进行深入分析。通过应用 Relief-F 算法，计算了 50 个比值特征对 LCC 的权重值，并在每组具有相同分

子的比值特征中，筛选出权重值最高的 1 个用于 LCC 反演。最终，构建了由 10 个植被指数比值特征组成的 VIRFS 及其基于 Relief-F 算法的权重值，详见表 3。可见，在所选的 10 个比值特征中，有 80% 的分母为 SR，表明 SR 相比其他植被指数更能有效表征 LAI 信号，从而有助于削弱冠层结构对光谱特征的干扰，提升比值特征对 LCC 的响应能力。

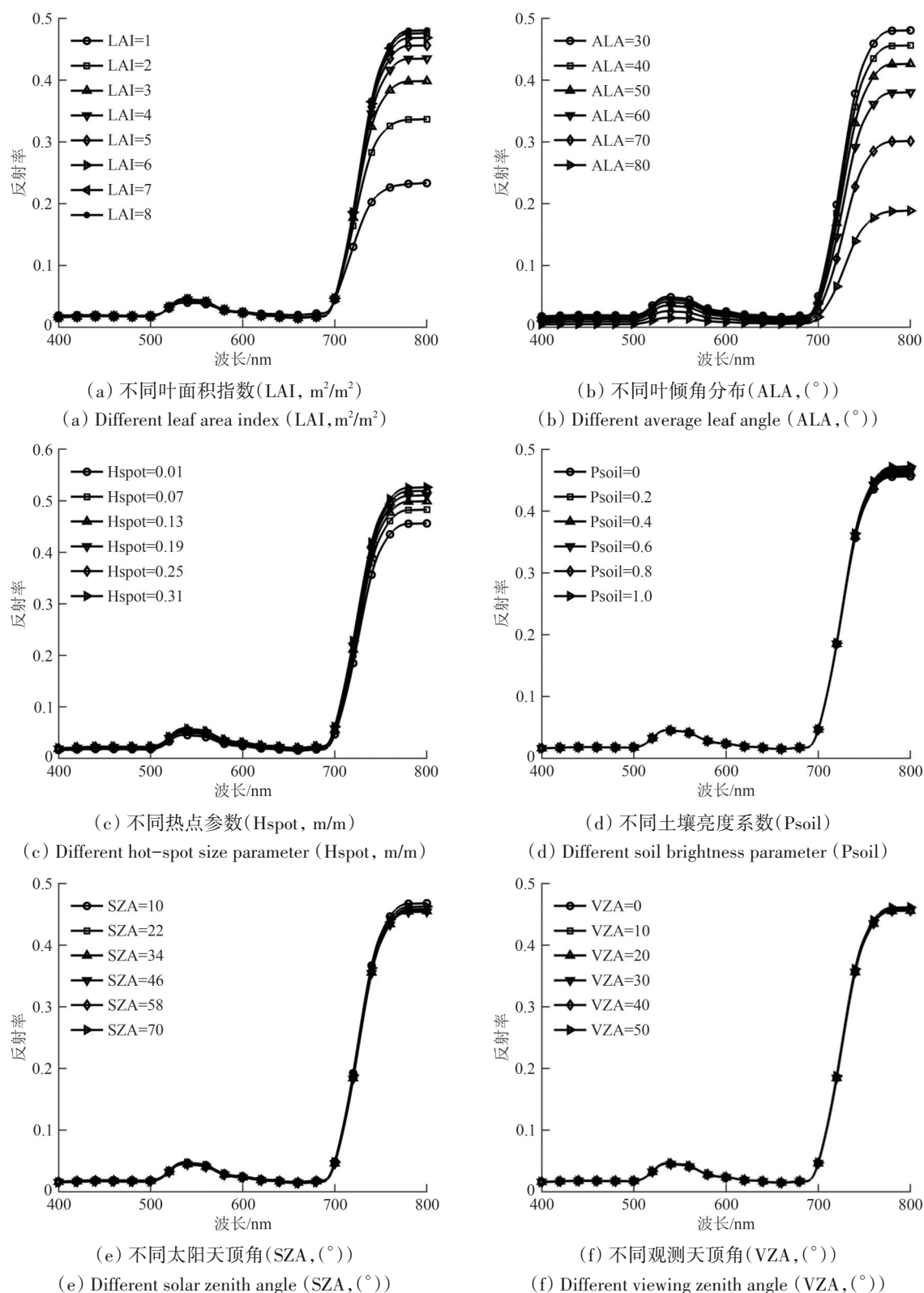


图3 450—800 nm波段冠层反射率对冠层参数变化的敏感性分析

Fig. 3 Sensitivity analysis of canopy reflectance at 450—800 nm to variations in canopy parameters

图4展示了VIRFS与对比特征集VIFS在模拟数据中与LCC和LAI的Pearson相关性分布情况。可见在与LCC的相关性方面，VIRFS整体表现出更高的一致性与显著性，其平均相关系数高于

VIFS；而在与LAI的相关性中，VIRFS特征相关性较低，明显低于VIFS，表明其受冠层结构影响较小。这一结果验证了基于LCC敏感性构建的VIRFS特征集，能够在保留对叶绿素含量高响应能

力的同时，有效抑制与LAI的相关性干扰，体现了其在叶绿素含量建模中的优势与针对性。

表3 本研究反演LCC的两个光谱特征集合
Table 3 Two spectral feature sets used for LCC estimation in this study

植被指数特征集 VIFS (Vegetation Index Feature Set)	植被指数比值特征集 VIRFS (Vegetation Index Ratio Feature Set)	VIRFS权重值
Datt99	Datt99/SR	0.031
LCI	LCI/NDVI	0.256
RERI	RERI/SR	1.000
CIG	CIG/SR	0.566
TVI	TVI/SR	0.648
TCI	TCI/SR	0.679
CIRE	CIRE/OSAVI	0.219
NDRE	NDRE/SR	0.192
GNDVI	GNDVI/SR	0.709
TCARI	TCARI/SR	0.849

为了进一步验证 VIRFS 特征在 LCC 建模中的有效性及其对 LAI 变化的鲁棒性，本研究优先从 VIRFS 特征集中选取了3个与 LAI 和 LCC 相关性分析中表现出明显区分性的比值特征，作为具有代表性的敏感指标；同时，从 VIFS 特征集中选取结构相近的对照特征，以便系统对比两类特征对 LAI 与 LCC 变化的响应差异。VIRFS 特征的选取标准包括以下内容：(1) 统计显著性。选取与 LCC 具有高相关性 (Pearson 相关系数 $|r|>0.7$, $P<0.01$)，且与 LAI 相关性较弱 ($|r|<0.2$, $P<0.01$) 的特征，体现其对叶绿素含量变化具有较强敏感性，同时受冠层结构变化影响较小；(2) 光谱结构代表性。

所选特征多为 VIRFS 特征集中波段组合典型、结构互异性强的比值型指数，能代表该特征集对 LCC 变化响应最显著、且在不同模拟阶段中表现稳定的关键子集。如图 5 (a) 所示，VIRFS 中的比值特征与 LCC 之间呈现出稳定的响应关系，且几乎不受 LAI 变化的干扰，表明其能够有效抑制 LAI 对光谱反射的影响，从而增强对 LCC 的表征能力。相比之下，如图 5 (b) 所示，VIFS 中的植被指数在不同 LAI 水平下与 LCC 的关系存在显著差异，说明其对 LCC 的响应受 LAI 变化的影响较大。因此，若直接使用这类植被指数反演 LCC，易受到冠层结构变化的干扰，可能降低估算精度。通过对比分析这些代表性特征与植被参数之间的响应关系，本研究进一步论证了 VIRFS 特征集在抑制冗余结构信息、增强光谱敏感性的有效性。

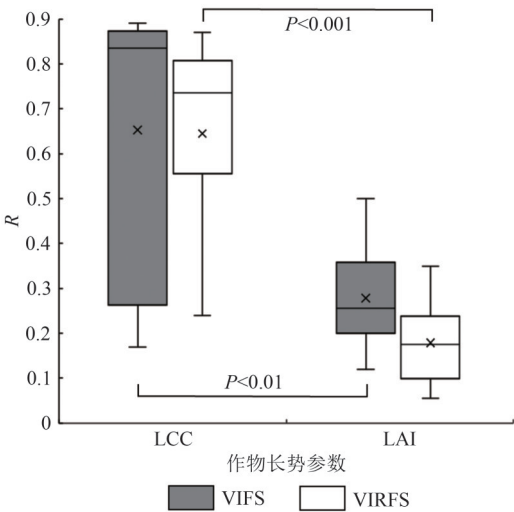
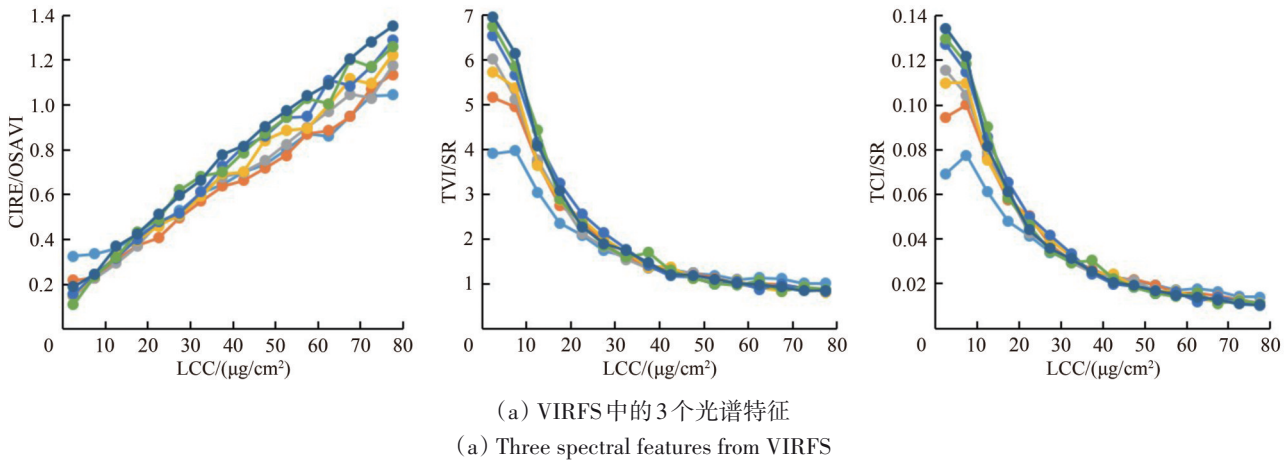


图4 VIFS和VIRFS光谱特征集与LCC和LAI的相关性
Fig. 4 Correlations of the two spectral feature sets VIFS and VIRFS with LCC and LAI



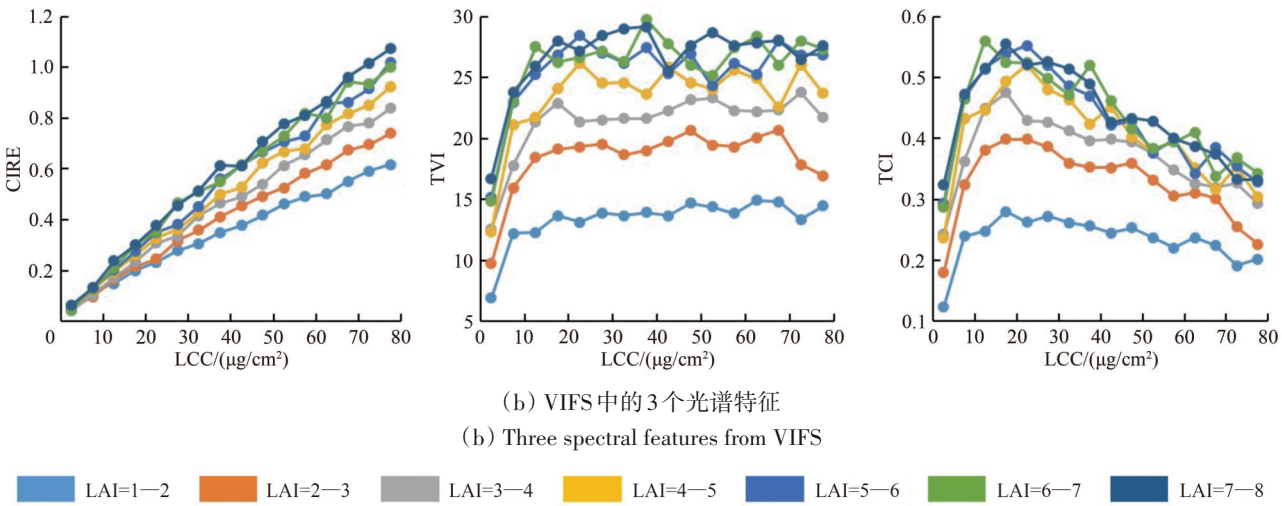


图5 基于PROSAIL模拟数据分析光谱特征在不同LAI下对LCC变化的敏感性
Fig. 5 Sensitivity of spectral features to changes in LCC under different LAI based on PROSAIL-simulated data

3.3 结合地面实测数据的LCC混合反演模型验证

3.3.1 基于不同特征集的迁移学习LCC反演结果

本研究比较两个特征集—VIRFS与VIFS的LCC反演表现，以验证VIRFS在削弱冠层结构异质性方面的优势及其对提升反演精度的有效性。研究采用iGS方法分别选取不同数量的地面调查样本进行模型微调，剩余样本用于模型验证。基于不同特征集的微调迁移学习模型在3个研究区的反演结果如表4所示。可见，在3个研究区中，基于VIRFS的模型在所有微调样本比例下均优于基于VIFS的模型，表

现出更高的精度与稳定性。其中，在最佳微调样本数量下，3个研究区的最佳精度分别为：（1）研究区1。 $R^2=0.82$ ， $RMSE=4.65\text{ }\mu\text{g}/\text{cm}^2$ ， $RRMSE=9.54\%$ ；研究区2。 $R^2=0.69$ ， $RMSE=4.53\text{ }\mu\text{g}/\text{cm}^2$ ， $RRMSE=6.40\%$ ；研究区3。 $R^2=0.75$ ， $RMSE=4.77\text{ }\mu\text{g}/\text{cm}^2$ ， $RRMSE=17.88\%$ 。此外，由表4可见，随着微调样本数量的增加，模型反演精度逐步提升。多数情况下，仅使用30%的地面样本即可实现稳定且可靠的反演效果，此时模型的 R^2 超过或接近0.7， $RMSE$ 维持在4—5 $\mu\text{g}/\text{cm}^2$ 。

表4 基于不同特征集的微调迁移学习LCC反演模型在3个研究区的精度验证结果

本地训练 样本比例/%		特征集	研究区1			研究区2			研究区3		
			R^2	RMSE/ ($\mu\text{g}/\text{cm}^2$)	RRMSE/%	R^2	RMSE/ ($\mu\text{g}/\text{cm}^2$)	RRMSE/%	R^2	RMSE/ ($\mu\text{g}/\text{cm}^2$)	RRMSE/%
10		VIFS	0	12.93	26.12	0.34	7.42	10.63	0	12.14	47.31
		VIRFS	0.58	7.02	14.24	0.30	7.40	10.58	0.62	5.78	21.89
20		VIFS	0.56	7.35	14.92	0.06	9.10	13.09	0.19	8.59	33.65
		VIRFS	0.74	5.53	11.24	0.64	5.29	7.52	0.72	5.05	19.32
30		VIFS	0.63	6.48	13.23	0.24	8.52	12.27	0.34	7.86	30.71
		VIRFS	0.74	5.76	11.66	0.69	4.82	6.86	0.73	4.98	18.82
40		VIFS	0.55	7.07	14.34	0.42	7.59	10.96	0.52	6.62	26.05
		VIRFS	0.78	5.27	10.73	0.70	4.68	6.65	0.75	4.77	17.88
50		VIFS	0.64	6.17	12.48	0.49	7.04	10.06	0.57	6.38	25.42
		VIRFS	0.82	4.65	9.54	0.69	4.53	6.40	0.73	4.98	18.61

图6进一步展示了在30%样本微调条件下，不同特征集在3个研究区的LCC反演结果。从散点的分布可以直观地看出，VIRFS在不同地区与作物类

型中均显著优于VIFS，充分证明其在缓解冠层结构效应、提升LCC反演精度方面的潜力。如图6（a）和（b）的左图所示，在研究区1的多种作物LCC

反演中，基于 VIRFS 建模的 R^2 较 VIFS 提高了 0.11，RMSE 降低 $0.72\text{ }\mu\text{g}/\text{cm}^2$ ，RRMSE 下降 1.57%。如图 6 (a) 和 (b) 的中图所示，在研究区 2 的小麦反演中，VIRFS 带来的改进更为显著， R^2 提高 0.45，RMSE 降低 $3.70\text{ }\mu\text{g}/\text{cm}^2$ ，RRMSE 减少 5.41%。如图 6 (a) 和 (b) 的研究区 3 所示，在研究区 3 的水稻反演结果中，VIRFS 相较于 VIFS 同样表现优越， R^2 提高 0.39，RMSE 降低 $2.88\text{ }\mu\text{g}/\text{cm}^2$ ，RRMSE 减少 11.89%。这些结果清晰地表明，VIRFS 在减

弱冠层结构异质性后能够显著提高 LCC 反演精度。在所有研究区域和作物类型中，VIRFS 展现出了优于 VIFS 的性能，特别是在小麦（研究区 2）以及水稻（研究区 3）的反演中，改进效果最为显著。这种一致性的改进表明，通过采用植被指数组合构建比值特征的方式能够有效降低 LAI 参数变化的干扰，可以在不同农业环境中提高 LCC 估算的准确性。

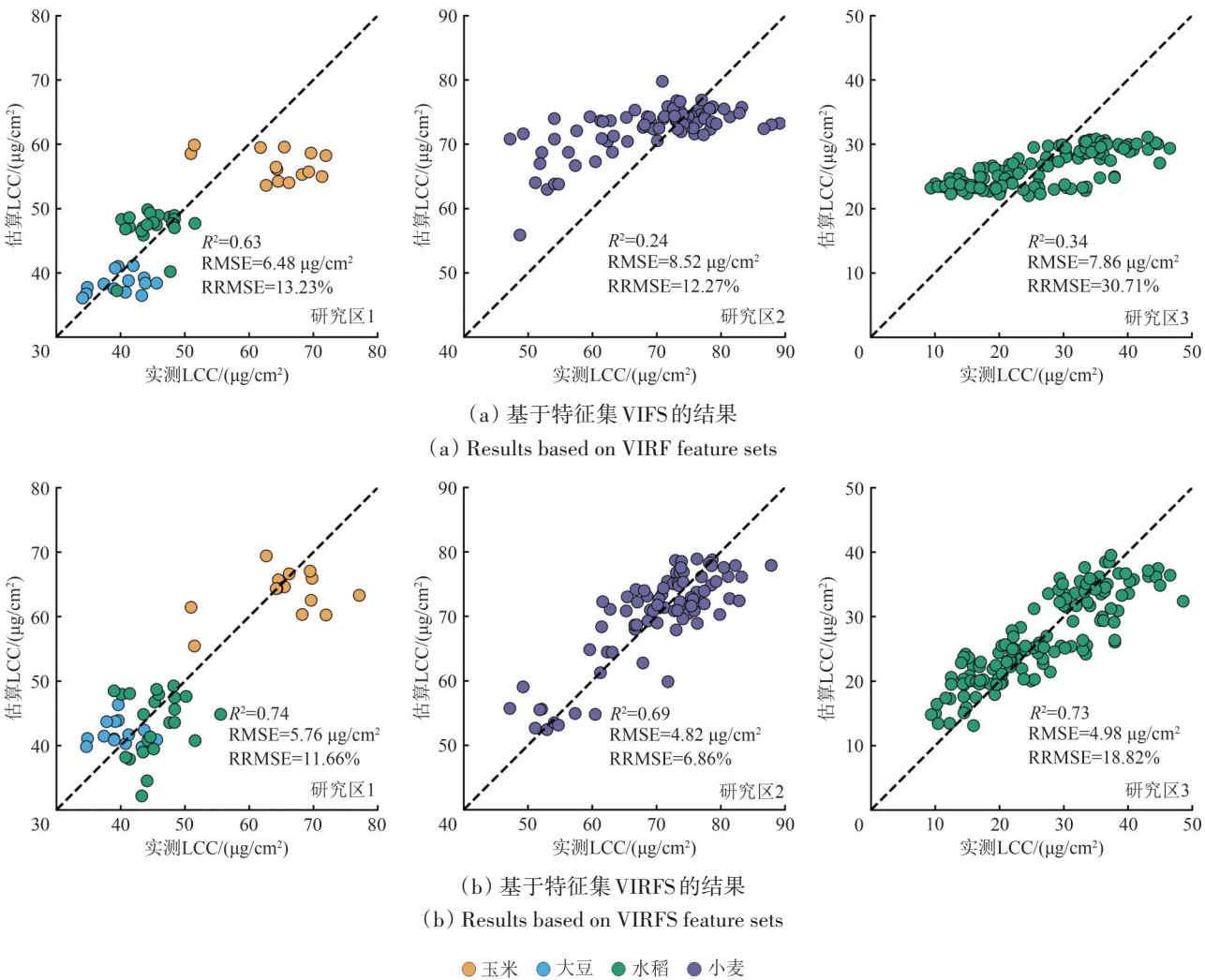


图6 基于不同特征集使用本地 30% 样本微调的 LCC 反演模型在 3 个研究区的地面验证结果

Fig. 6 Ground validation results of LCC inversion models in the three study areas , fine-tuned with 30% local samples based on two feature sets

为进一步评估所选比值植被指数特征在反演 LCC 中的贡献情况，研究采用 SHAP (SHapley Additive exPlanations) 方法，对各区域微调模型在 30% 训练数据上的特征重要性进行了可解释性分析，如图 7 所示。可见，4 个区域在特征重要性排序方面既存在一致性，也表现出一定差异。整体

而言，CIG/SR、CIRE/OSAVI 和 TVI/SR 3 个比值植被指数特征在各区域中均排名靠前，表明其在不同区域下均具有稳定且显著的预测能力。其中，CIG/SR 在模拟数据与实测数据中均表现为最具影响力的特征，进一步验证了其对 LCC 变化具有高度敏感性和强解释力。在局部特征贡献方面，不

同区域的 SHAP 值分布存在一定差异, 模拟数据中特征贡献跨度更大, 而实测区域中模型更依赖于少数关键特征。值得注意的是, 部分特征 (如 Datt99/SR、TCARI/SR 等) 在不同区域中排名波动

较大, 表明其在某些情境下可能受限于 LAI 变化、土壤背景条件或传感器噪声等干扰因素, 泛化能力相对较弱。

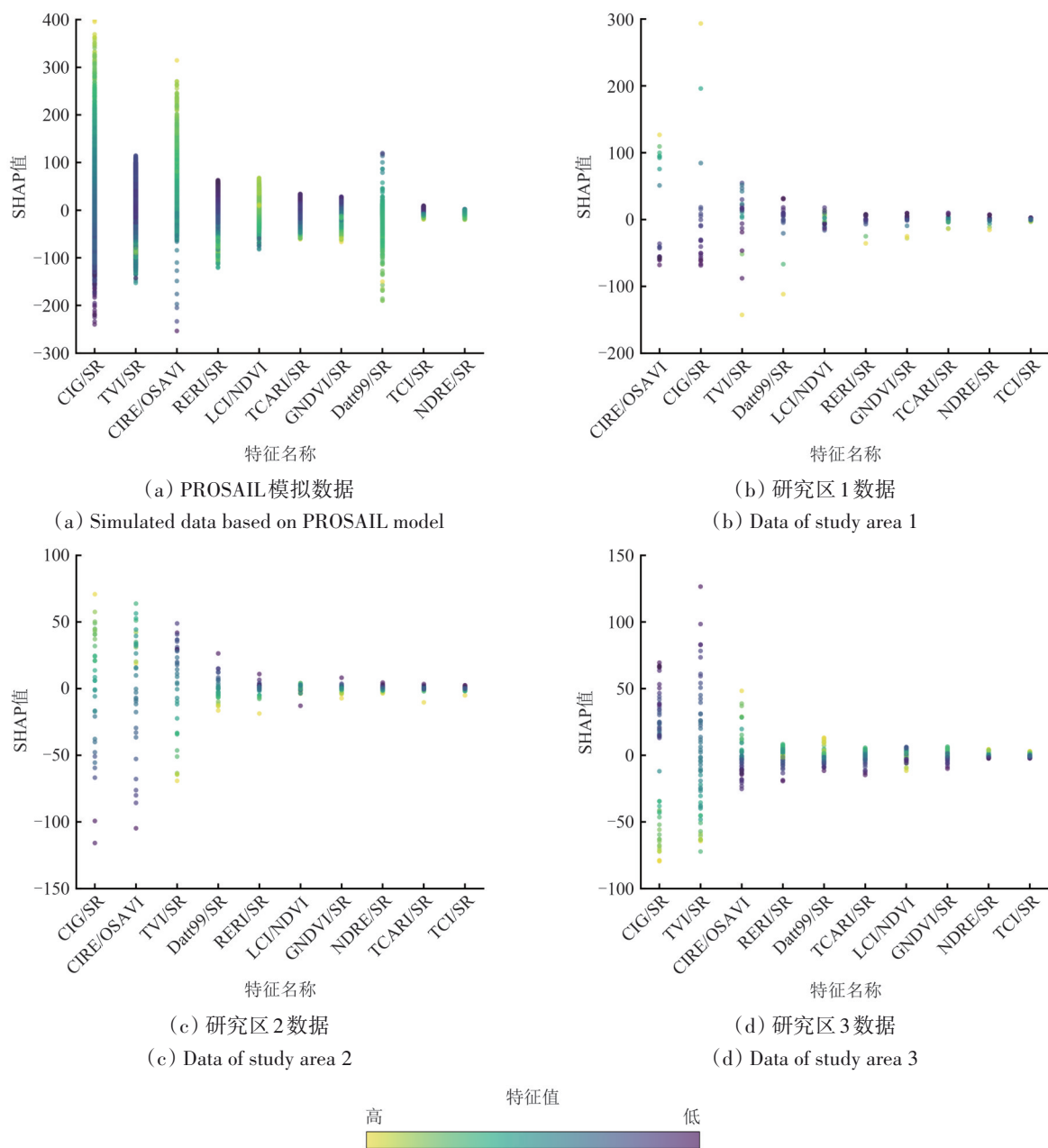


图7 基于SHAP方法的各区域训练数据特征对反演LCC贡献度的可解释性分析结果

Fig. 7 Results of interpretability analysis of the feature contribution to the inversion LCC in the training data of each region based on SHAP method

3.3.2 基于不同样本选择方法的迁移学习LCC反演结果

本部分旨在探讨样本选择策略对LCC反演模型微调效果的影响。以特征集 VIRFS 为基础, 分别采用 iGS 方法和随机选择方法, 从地面标记样本

中按比例选取 10%、20%、30%、40% 和 50% 作为模型微调样本, 其余部分作为验证集, 用于评估在不同微调样本量下模型的反演性能 (图8)。为保证结果的稳定性与代表性, 随机选择方法在每种比例下均重复实验 10 次并取平均值。

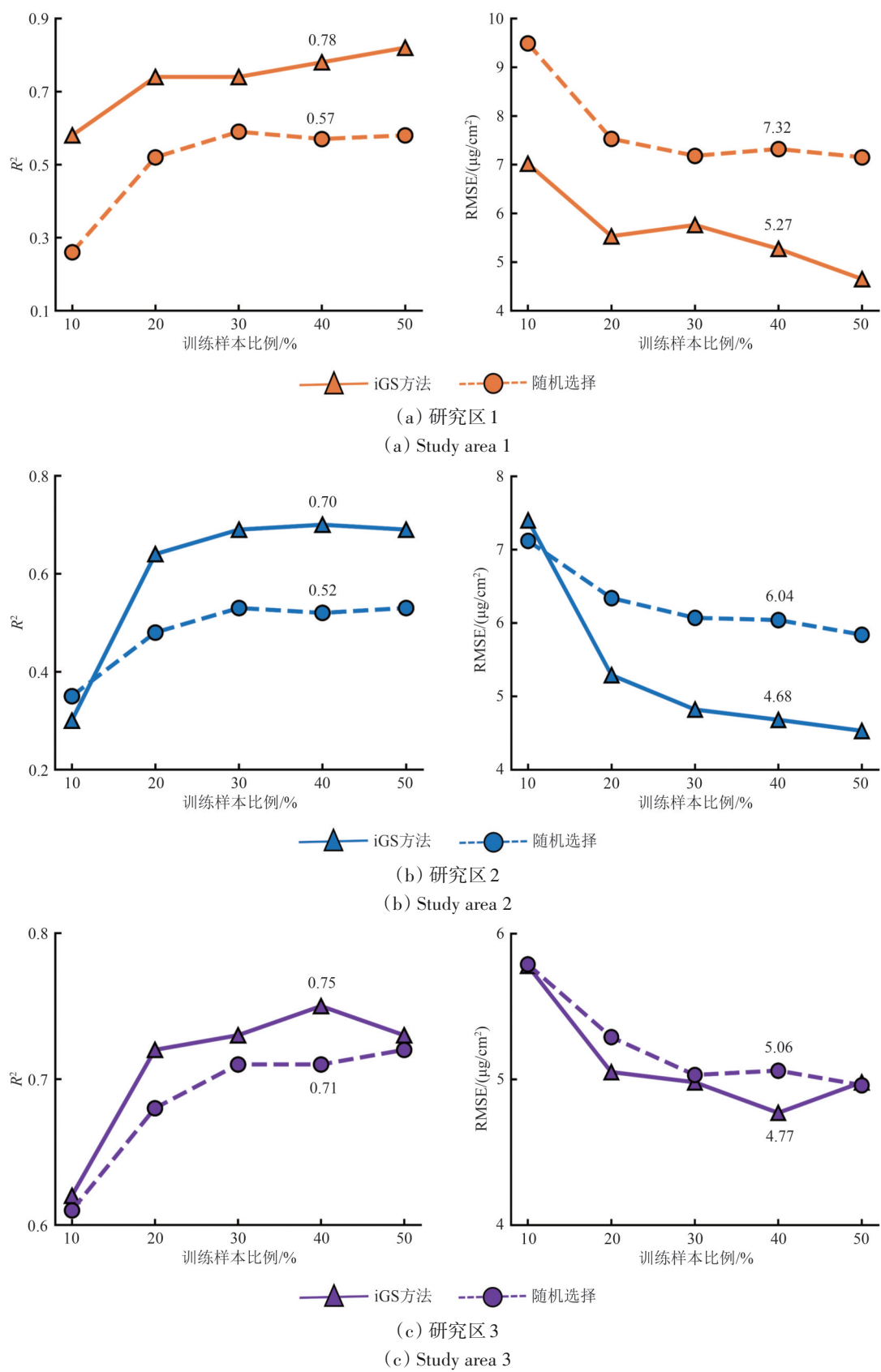


图8 基于不同样本选择方法的微调迁移学习LCC反演模型在3个研究区的地面验证结果

Fig. 8 Ground validation results of transfer learning-based LCC inversion models fine-tuned using different sample selection methods in the three study areas

由图 8 可见，在所有样本比例条件下，基于 iGS 方法的 LCC 反演精度均优于随机选择方法。尤其在微调样本比例为 40% 时，反演性能差异尤为显著。具体而言：在研究区 1（多作物）中，iGS 方法相比随机选择使 LCC 反演精度 R^2 提高 0.21，RMSE 降低 $2.05\text{ }\mu\text{g}/\text{cm}^2$ （图 8（a））；在研究区 2（小麦）中， R^2 提高 0.18，RMSE 降低 $1.36\text{ }\mu\text{g}/\text{cm}^2$ （图 8（b））；在研究区 3（稻田）中， R^2 提高 0.04，RMSE 降低 $0.29\text{ }\mu\text{g}/\text{cm}^2$ （图 8（c））。上述结果表明，iGS 方法选取的样本在稳定性与代表性

方面更具优势，能够在不同作物类型及区域中提升 LCC 反演的整体精度。特别是在多作物和小麦区域，改进效果最为显著。此外，从图 8 进一步发现，即便仅使用 20% 的 iGS 选择进行模型微调，其反演效果也优于随机选择方法使用 50% 样本的结果。这一发现表明，iGS 方法能够在数据量有限的情况下构建高效模型，显著降低数据采集和标注成本。最后，选取各研究区中反演精度最高的模型生成 LCC 空间分布图，结果如图 9 所示。

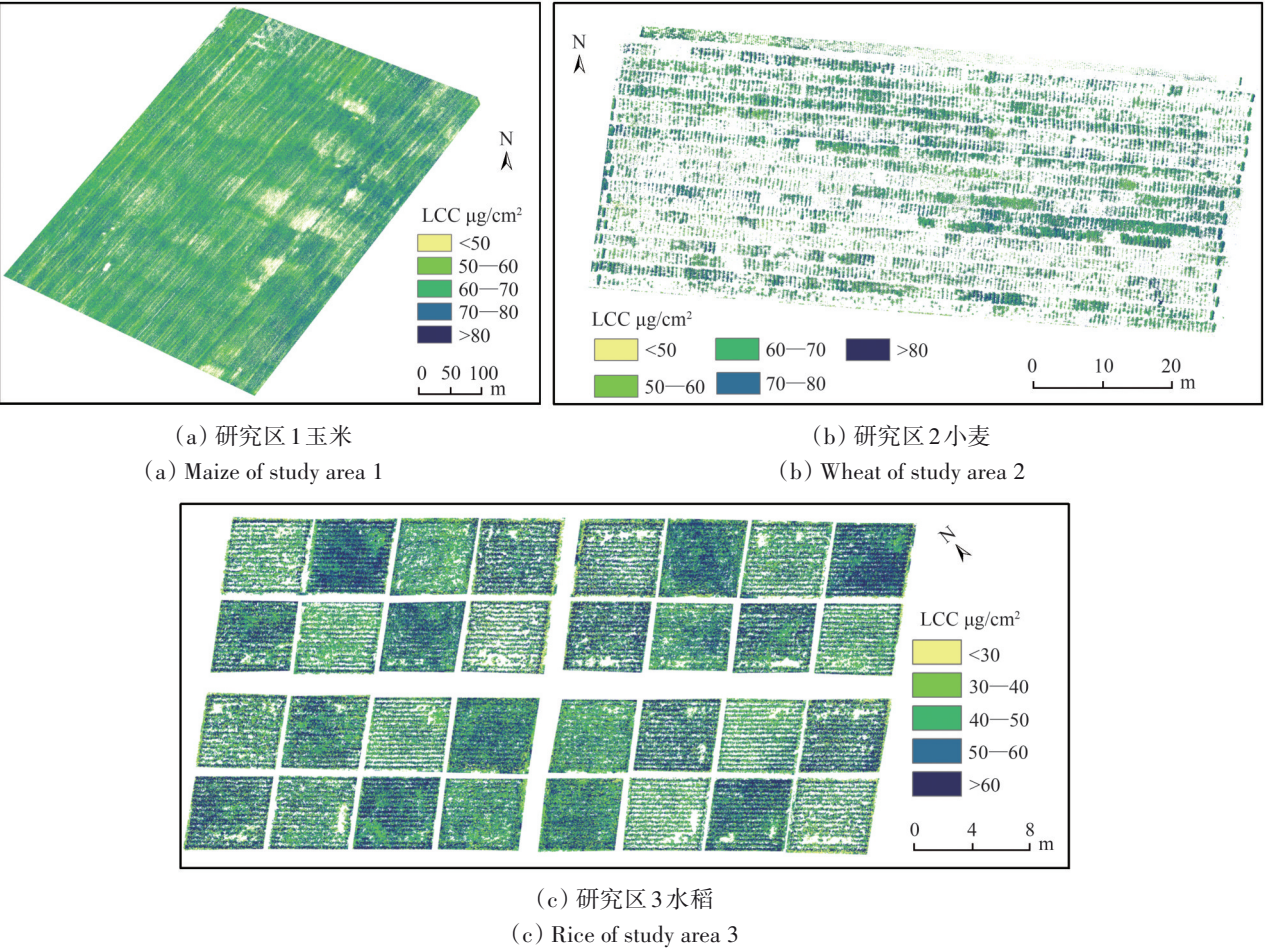


图 9 基于无人机多光谱影像的 3 个研究区 LCC 空间分布图
Fig. 9 LCC spatial distribution maps of three study areas derived from UAV multispectral imagery

在评估主动学习方法相较随机选择方法在 LCC 反演中的有效性后，为进一步分析两种样本选择策略的代表性与多样性，本研究引入了 t-SNE 与 UMAP 两种非线性降维方法，对所选样本在高维空间中的分布特征进行可视化分析。t-SNE 强调保持局部邻域结构，而 UMAP 则在保持局部结构的同时兼顾全局拓扑结构，从而提供更全面

的低维嵌入结果。本研究以各研究区中由 iGS 方法和随机选择方法分别选取的 40% 样本为对象进行降维可视化分析，其中随机选择方法设定随机种子为 42，以保证实验结果的可重复性。图 10 中展示了 3 种区域在 t-SNE 与 UMAP 空间中的样本分布情况，并计算了样本的信息熵 (H) 与变异系数 (CV)，作为样本选择效果的定量评价指标。

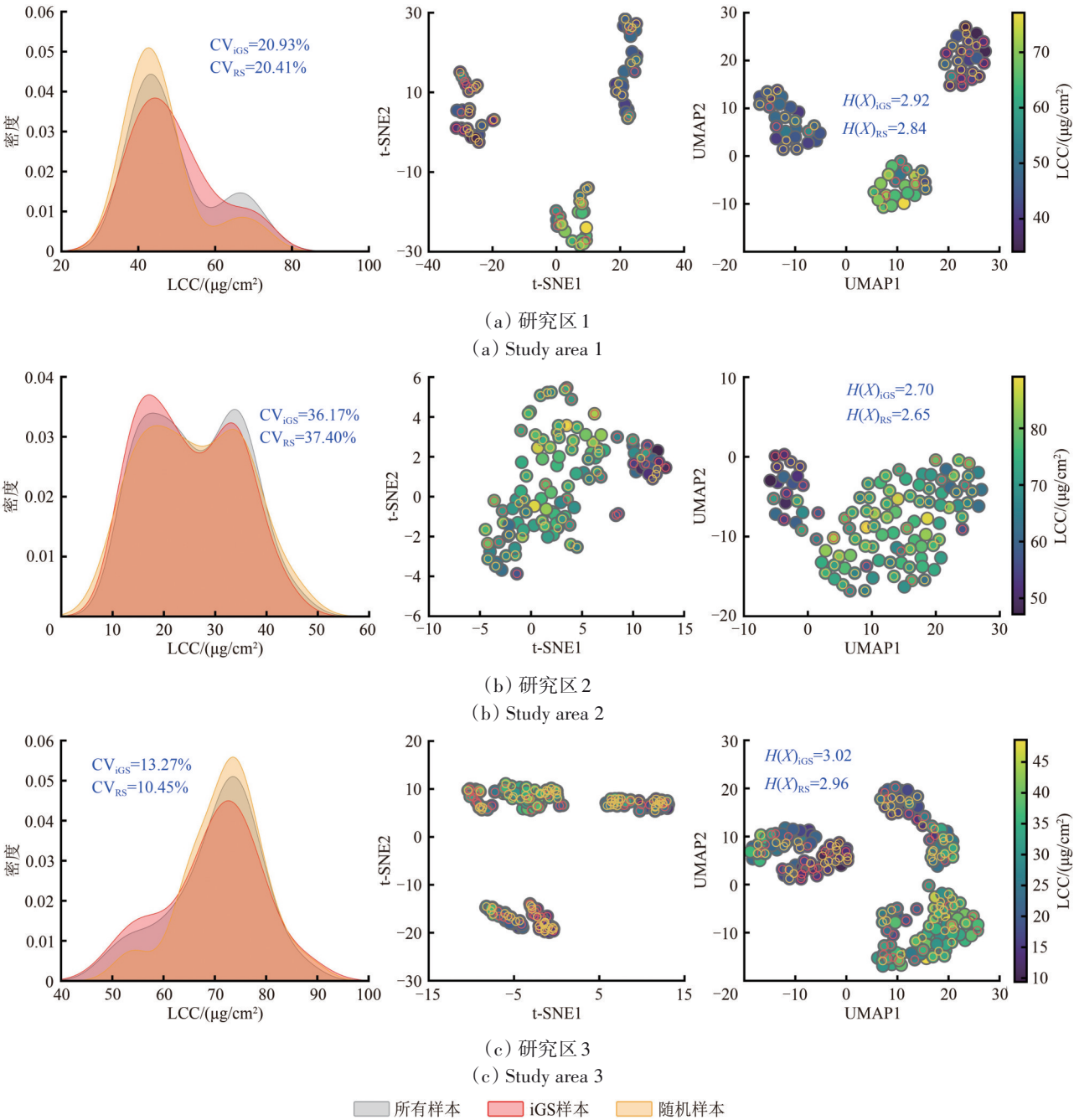


图 10 不同研究区 LCC 密度分布、t-SNE 和 UMAP 特征分布

Fig. 10 The LCC density distribution, the visualization map of t-SNE and UMAP feature distribution of different study areas

结果表明，iGS方法在样本分布方面显著优于随机选择方法。t-SNE 与 UMAP 可视化结果均显示，iGS 所选样本在嵌入空间中的分布更为均匀，覆盖范围更广，具有更好的代表性与结构覆盖性。特别是在 UMAP 图中，iGS 样本更好地保留了原始特征空间中的拓扑关系，展现出较强的区域渗透性。从定量指标来看，iGS 样本在各区域中均表现出更高的信息熵，反映出其在捕捉数据多样性与关键特征方面的优越性。此外，样本 LCC 值的密度图与 CV 分析结果表明，iGS 样本的 CV 值普遍高

于随机样本，进一步印证了其在样本空间覆盖性与代表性上的优势。综上所述，主动学习方法 iGS 不仅能够提升 LCC 反演模型的精度，在样本空间分布层面亦具备更高的有效性，为模型提供了更有价值的微调训练样本支持。

4 讨论

4.1 比值特征在减弱冠层结构异质性中的优势

本研究结果充分验证了比值特征集（VIRFS）

在削弱冠层结构异质性干扰方面的显著优势。相较于 VIFS, VIRFS 通过构建基于植被指数间组合的比值特征,能够在一定程度上抑制由冠层结构变化(如 LAI)引发的光谱响应偏差,从而增强模型对 LCC 的敏感性与稳定性。这一思路在早期已有研究中被提出并初步应用于提高遥感参数反演的稳健性(Daughtry 等, 2000; Xu 等, 2019),而本研究在多作物、多区域的实证基础上进一步拓展了其适用范围与方法体系。

由表 3 可知, VIRFS 中有 80% 的比值特征以 SR 为分母,反映出其在构建稳健光谱特征方面的重要性。SR 作为一种结构简单且物理意义明确的植被指数,已被证实表征 LAI 变化方面具备良好的稳定性(Daughtry 等, 2000; Guo 等, 2023)。因此,通过 SR 作为分母抑制 LAI 变化的影响,能够在复杂多变的冠层条件下,有效提升光谱特征对 LCC 的响应能力。鉴于本研究采用的深度神经网络模型具备较强的非线性拟合与特征冗余适应能力,特征间可能存在的共线性对建模性能影响有限,因此在特征选择阶段保留了全部比值特征,以充分发挥模型的表达能力与泛化潜力。从表 4 结果可知,模型在多种作物类型及不同区域条件下,基于 VIRFS 的反演模型均表现出更优的精度和一致性,显示出其良好的跨作物泛化能力和对结构异质性的适应性。比值特征通过增强对目标生化参数的选择性,同时抑制非目标变量的干扰,为提升 LCC 反演的准确性提供了更加稳健的特征基础。综上所述,基于比值特征构建的反演模型不仅提升了对 LCC 的估算精度,也为应对多样化作物冠层结构所带来的遥感反演挑战提供了有效解决思路。

4.2 结合微调迁移学习的混合方法在 LCC 反演中的优势

在迁移学习与 PROSAIL 模型应用于作物 LCC 反演方面,已有研究表明该策略具有一定的有效性。例如, Zhang 等(2021)和 Bhadra 等(2024)分别采用基于 PROSAIL 模拟数据和无人机观测数据构建的迁移学习框架,实现了对 LCC 的准确估算。然而,这两项研究仍存在一定局限性。首先,它们均仅针对单一作物构建了迁移策略,未充分考虑不同作物之间在光谱特性、冠层结构等方面的差异对迁移效果的影响,导致方法的泛化性受

限,难以直接推广至其他作物类型。其次,这两项研究主要基于高光谱数据进行建模,尽管高光谱具有丰富的光谱信息,但其在数据获取、处理复杂性以及成本控制方面存在一定障碍,限制了该方法在主流多光谱遥感平台上的实际应用。本研究基于 PROSAIL 模拟数据构建特征集 VIRFS,并训练初始的 DNN 模型,通过引入部分地面观测样本对模型进行微调,从而实现对不同作物类型的无人机多光谱 LCC 反演。从表 4 和图 8 的结果可以看出,仅使用 20% 及以上的实测样本进行微调,模型在多个地点和不同作物条件下均表现出较高的 LCC 反演精度与稳定性。与完全依赖实测样本构建模型相比,该方法显著缓解了因实测数据稀缺导致的模型欠拟合问题(Zhang 等, 2021; Zhao 等, 2022),有效扩展了训练样本空间,提高了模型的泛化能力。

此外,微调过程引入了针对目标区域作物特性和环境条件的本地信息,使得模型能够在保持模拟数据所蕴含的结构知识基础上,更加适应真实观测场景的复杂性与不确定性,从而提升反演结果的可靠性和精度。总体来看,本研究结合迁移学习策略,提出了一种融合模拟数据与有限地面观测数据的 LCC 反演方法,在提高反演模型精度的同时,兼顾了多作物适应性与实际可操作性。

4.3 未来工作的展望

尽管本研究在多作物 LCC 反演中取得了较为理想的结果,但仍存在一些值得进一步深入研究方向。一方面,当前构建的迁移学习框架主要基于 PROSAIL 模型生成的模拟数据,其对实际环境的刻画能力在一定程度上受限于模型假设和输入参数的准确性。目前三维辐射传输模型已经在遥感领域得到了应用,尤其是在生物性状反演研究中展现出了巨大的应用潜力(Dong 等, 2024; Zhao 等, 2024; Qi 等, 2022)。未来可结合三维辐射传输模型更好地处理作物冠层的空间异质性,使得反演参数更加准确。另一方面,虽然本研究验证了主动学习策略(如 iGS)在提升样本选择效率方面的有效性,但随着地块尺度从地块级扩展到县级或更大区域,面临的样本池急剧增大,使得每轮需计算所有未标注样本的信息量(如特征距离与预测误差的乘积),由此带来较高的计算开

销。在实际应用中, 其计算复杂度与运行效率问题仍是亟待解决的挑战。未来研究可进一步优化信息量计算的效率, 已有研究表明, 主动学习中的核心集选择问题可转化为 k -Center 问题, 并借助近似算法获得高效解 (Sener 和 Savarese, 2018)。为减少特征距离计算的开销, 可探索特征降维 (如 PCA) 或近似距离 (如哈希编码距离) 等手段。此外, 自适应次模性的效率优化可大幅提升深度模型中子任务的计算效率, 包括理论近似比保证、计算加速技术及应用场景的泛化能力等方面 (Golovin 和 Krause, 2011)。这些优化方向有望为主动学习在大尺度遥感数据自动化标注中的推广提供支撑。同时, 随着深度学习技术的不断发展, 未来还可尝试将图神经网络、注意力机制等新型神经网络结构引入 LCC 反演任务中, 以更充分地挖掘空间—光谱信息之间的复杂关系, 从而推动作物长势参数遥感估测方法向更智能、更泛化的方向发展。

5 结 论

针对多种作物条件下冠层结构异质性干扰作物 LCC 估测精度这一关键问题, 以三大典型农业区的多种作物为研究对象, 采用 PROSAIL 模拟数据、低 LAI 敏感性比值特征集 (VIRFS)、主动学习策略优化迁移学习及深度学习融合的方法, 提出了一种融合物理模拟与深度学习的主粮作物 LCC 反演框架, 并依据作物实测地面标注数据进行实验验证。主要结论如下:

(1) 所提方法具有优异的跨作物适应性与估测精度。在三大典型农业区的多种作物验证中, 模型 R^2 达到 0.69—0.82, RMSE 控制在 4.53—4.77 $\mu\text{g}/\text{cm}^2$, 有效克服了冠层结构异质性的干扰。

(2) 比值特征集 (VIRFS) 显著提升了模型的鲁棒性。构建的特征集 VIRFS 在降低冠层结构参数干扰方面表现出显著优势, 是提升模型在不同作物间泛化能力与稳定性的关键因素。

(3) “物理模拟+主动学习”策略有效降低了样本需求。通过结合辐射传输模拟数据集与主动学习策略, 模型在仅使用 30% 地面标注样本的情况下仍能实现较高反演精度, 成功缓解了传统遥感反演中对大量实测数据的依赖问题。

志 谢 此次野外实验的数据获取得到了华

中农业大学、山东农业大学和黑龙江七星农场的大力支持, 在此表示衷心的感谢!

参考文献 (References)

- Bhadra S, Sagan V, Sarkar S, Braud M, Mockler T C and Eveland A L. 2024. PROSAIL-Net: a transfer learning-based dual stream neural network to estimate leaf chlorophyll and leaf angle of crops from UAV hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 210: 1-24 [DOI: 10.1016/j.isprsjprs.2024.02.020]
- Chen Z L, Jia K, Wei X Q, Liu Y, Zhan Y L, Xia M, Yao Y J and Zhang X T. 2022. Improving leaf area index estimation accuracy of wheat by involving leaf chlorophyll content information. *Computers and Electronics in Agriculture*, 196: 106902 [DOI: 10.1016/j.compag.2022.106902]
- Coste S, Baraloto C, Leroy C, Marcon É, Renaud A, Richardson A D, Roggy J C, Schimann H, Uddling J and Hérault B. 2010. Assessing foliar chlorophyll contents with the SPAD-502 chlorophyll meter: a calibration test with thirteen tree species of tropical rainforest in French Guiana. *Annals of Forest Science*, 67(6): 607-607 [DOI: 10.1051/forest/2010020]
- Croft H, Chen J M, Luo X Z, Bartlett P, Chen B and Staebler R M. 2017. Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology*, 23(9): 3513-3524 [DOI: 10.1111/gcb.13599]
- Daughtry C S T, Walthall C L, Kim M S, de Colstoun E B and McMurtrey J E III. 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, 74(2): 229-239 [DOI: 10.1016/S0034-4257(00)00113-9]
- Dong M X, Liu S Y, Jiang R B, Qi J B, de Solan B, Comar A, Li L Y, Li W J, Ding Y F and Baret F. 2024. Comparing and combining data-driven and model-driven approaches to monitor wheat green area index with high spatio-temporal resolution satellites. *Remote Sensing of Environment*, 305: 114118 [DOI: 10.1016/j.rse.2024.114118]
- Fernández-Guisuraga J M, Calvo L, Quintano C, Fernández-Manso A and Fernandes P M. 2023. Fractional vegetation cover ratio estimated from radiative transfer modeling outperforms spectral indices to assess fire severity in several Mediterranean plant communities. *Remote Sensing of Environment*, 290: 113542 [DOI: 10.1016/j.rse.2023.113542]
- Gao S, Yan K, Liu J X, Pu J B, Zou D X, Qi J B, Mu X H and Yan G J. 2024. Assessment of remote-sensed vegetation indices for estimating forest chlorophyll concentration. *Ecological Indicators*, 162: 112001 [DOI: 10.1016/j.ecolind.2024.112001]
- García-Haro F J, Campos-Taberner M, Muñoz-Marí J, Laparra V, Camacho F, Sánchez-Zapero J and Camps-Valls G. 2018. Derivation of global vegetation biophysical parameters from EUMETSAT Polar System. *ISPRS Journal of Photogrammetry and Remote Sensing*, 139: 57-74 [DOI: 10.1016/j.isprsjprs.2018.03.005]

- Golovin D and Krause A. 2011. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *The Journal of Artificial Intelligence Research*, 42(1): 427-486 [DOI: 10.1613/jair.3278]
- Gu C P, Li J, Liu Q H, Zhang H, Huete A, Fang H L, Liu L Y, Mumtaz F, Lin S R, Wang X H, Dong Y D, Zhao J, Bai J H, Yu W T, Liu C and Guan L. 2025. Deriving leaf-scale chlorophyll index (CIleaf) from canopy reflectance by correcting for the canopy multiple scattering based on spectral invariant theory. *Remote Sensing of Environment*, 322: 114692 [DOI: 10.1016/j.rse.2025.114692]
- Gu C P, Li J, Liu Q H, Zhang H, Zhang Z X, Wen Y and Wang X H. 2024. Retrieval and validation of leaf chlorophyll content using GF-6 WFV Imageries. *National Remote Sensing Bulletin*, 28(10): 2553-2564 (谷晨鹏, 李静, 柳钦火, 张虎, 张召星, 文远, 王晓函. 2024. 高分六号宽幅相机叶片叶绿素含量反演方法与验证. *遥感学报*, 28(10): 2553-2564) [doi: 10.11834/jrs.20242484]
- Guo A T, Ye H C, Huang W J, Qian B X, Wang J J, Lan Y B and Wang S Z. 2023. Inversion of maize leaf area index from UAV hyperspectral and multispectral imagery. *Computers and Electronics in Agriculture*, 212: 108020 [DOI: 10.1016/J.COMPAG.2023.108020]
- Jiang J Y, Weiss M, Liu S Y and Baret F. 2022. Effective GAI is best estimated from reflectance observations as compared to GAI and LAI: demonstration for wheat and maize crops based on 3D radiative transfer simulations. *Field Crops Research*, 283: 108538 [DOI: 10.1016/J.FCR.2022.108538]
- Kim Y and Shin B. 2022. In defense of core-set: a density-aware core-set selection for active learning//*Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. Washington: ACM: 804-812 [DOI: 10.1145/3534678.3539476]
- Li D, Chen J M, Zhang X, Yan Y, Zhu J, Zheng H B, Zhou K, Yao X, Tian Y C, Zhu Y, Cheng T and Cao W X. 2020. Improved estimation of leaf chlorophyll content of row crops from canopy reflectance spectra through minimizing canopy structural effects and optimizing off-noon observation time. *Remote Sensing of Environment*, 248: 111985 [DOI: 10.1016/j.rse.2020.111985]
- Li J T, Wijewardane N K, Ge Y F and Shi Y Y. 2023. Improved chlorophyll and water content estimations at leaf level with a hybrid radiative transfer and machine learning model. *Computers and Electronics in Agriculture*, 206: 107669 [DOI: 10.1016/j.compag.2023.107669]
- Li W, Li D, Warner T A, Liu S Y, Baret F, Yang P Q, Jiang J L, Dong M X, Cheng T, Zhu Y, Cao W X and Yao X. 2025. Improved generality of wheat green LAI models through mitigation of the effect of leaf chlorophyll content variation with red edge vegetation indices. *Remote Sensing of Environment*, 318: 114589 [DOI: 10.1016/j.rse.2024.114589]
- Li Y X, Liu H J, Ma J C and Zhang L X. 2021. Estimation of leaf area index for winter wheat at early stages based on convolutional neural networks. *Computers and Electronics in Agriculture*, 190: 106480 [DOI: 10.1016/j.compag.2021.106480]
- Liang L, Di L P, Zhang L P, Deng M X, Qin Z H, Zhao S H and Lin H. 2015. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. *Remote Sensing of Environment*, 165: 123-134 [DOI: 10.1016/j.rse.2015.04.032]
- Liu Y B, Pei J, Zou Y P, Tan S F, He Y N, Zheng X P, Wang T X, Fang H J, Wang L and Huang J X. 2025. A novel Hybrid-DCNN-based framework for enhanced rice aboveground biomass estimation under limited samples. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 4405716 [DOI: 10.1109/tgrs.2025.3544343]
- Ma Y C, Chen S, Ermon S and Lobell D B. 2024. Transfer learning in environmental remote sensing. *Remote Sensing of Environment*, 301: 113924 [DOI: 10.1016/j.rse.2023.113924]
- Nabika T, Nagata K, Mizumaki M, Katakami S and Okada M. 2024. Bayesian active learning with model selection for spectral experiments. *Scientific Reports*, 14(1): 3680 [DOI: 10.1038/s41598-024-54329-w]
- Pan S J and Yang Q. 2010. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10): 1345-1359 [DOI: 10.1109/TKDE.2009.191]
- Pan Y Y, Wu W X, He J Y, Zhu J, Su X, Li W Y, Li D, Yao X, Cheng T, Zhu Y, Cao W X and Tian Y C. 2024. A novel approach for estimating fractional cover of crops by correcting angular effect using radiative transfer models and UAV multi-angular spectral data. *Computers and Electronics in Agriculture*, 222: 109030 [DOI: 10.1016/j.compag.2024.109030]
- Qi J B, Xie D H, Jiang J Y and Huang H G. 2022. 3D radiative transfer modeling of structurally complex forest canopies through a lightweight boundary-based description of leaf clusters. *Remote Sensing of Environment*, 283: 113301 [DOI: 10.1016/j.rse.2022.113301]
- Qian X J and Liu L Y. 2020. Retrieving crop leaf chlorophyll content using an improved look-up-table approach by combining multiple canopy structures and soil backgrounds. *Remote Sensing*, 12(13): 2139 [DOI: 10.3390/rs12132139]
- Safonova A, Ghazaryan G, Stiller S, Main-Knorn M, Nendel C and Ryo M. 2023. Ten deep learning techniques to address small data problems with remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 125: 103569 [DOI: 10.1016/j.jag.2023.103569]
- Sener O and Savarese S. 2018. A geometric approach to active learning for convolutional neural networks. *arXiv:1708.00489* [DOI: 10.48550/arXiv.1708.00489]
- Sun B, Wang C F, Yang C H, Xu B D, Zhou G S, Li X Y, Xie J, Xu S J, Liu B, Xie T J, Kuai J and Zhang J. 2021. Retrieval of rapeseed leaf area index using the PROSAIL model with canopy coverage derived from UAV images as a correction parameter. *International Journal of Applied Earth Observation and Geoinformation*, 102: 102373 [DOI: 10.1016/j.jag.2021.102373]
- Verrelst J, Rivera J P, Veroustraete F, Muñoz-Mari J, Clevers J G P W, Camps-Valls G and Moreno J. 2015. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – A comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108: 260-272 [DOI: 10.1016/j.isprsjprs.2015.04.013]

- Wan L, Liu Y F, He Y and Cen H Y. 2023. Prior knowledge and active learning enable hybrid method for estimating leaf chlorophyll content from multi-scale canopy reflectance. *Computers and Electronics in Agriculture*, 214: 108308 [DOI: 10.1016/j.compag.2023.108308]
- Wan L, Ryu Y, Dechant B, Lee J, Zhong Z L and Feng H Z. 2024. Improving retrieval of leaf chlorophyll content from Sentinel-2 and Landsat-7/8 imagery by correcting for canopy structural effects. *Remote Sensing of Environment*, 304: 114048 [DOI: 10.1016/J.RSE.2024.114048]
- Wan L, Zhu J P, Du X Y, Zhang J F, Han X Z, Zhou W J, Li X P, Liu J L, Liang F, He Y and Cen H Y. 2021. A model for phenotyping crop fractional vegetation cover using imagery from unmanned aerial vehicles. *Journal of Experimental Botany*, 72(13): 4691-4707 [doi: 10.1093/jxb/erab194]
- Wu D R, Lin C T and Huang J. 2019. Active learning for regression using greedy sampling. *Information Sciences*, 474: 90-105 [DOI: 10.1016/j.ins.2018.09.060]
- Xu M Z, Liu R G, Chen J M, Liu Y, Shang R, Ju W M, Wu C Y and Huang W J. 2019. Retrieving leaf chlorophyll content using a matrix-based vegetation index combination approach. *Remote Sensing of Environment*, 224: 60-73 [DOI: 10.1016/j.rse.2019.01.039]
- Yosinski J, Clune J, Bengio Y and Lipson H. 2014. How transferable are features in deep neural networks?//*Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: MIT Press: 3320-3328
- Zhang H, Li J, Liu Q H, Lin S R, Huete A, Liu L Y, Croft H, Clevers J G P W, Zeng Y L, Wang X H, Gu C P, Zhang Z X, Zhao J, Dong Y D, Mumtaz F and Yu W T. 2022. A novel red-edge spectral index for retrieving the leaf chlorophyll content. *Methods in Ecology and Evolution*, 13(12): 2771-2787 [DOI: 10.1111/2041-210X.13994]
- Zhang Y, Hui J, Qin Q M, Sun Y H, Zhang T Y, Sun H and Li M Z. 2021. Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data. *Remote Sensing of Environment*, 267: 112724 [DOI: 10.1016/j.rse.2021.112724]
- Zhang Z W, Jin W J, Dou R Y, Cai Z W, Wei H D, Wu T Z, Yang S, Tan M L, Li Z J, Wang C, Yin G F and Xu B D. 2023. Improved estimation of leaf area index by reducing leaf chlorophyll content and saturation effects based on red-edge bands. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 4403314 [DOI: 10.1109/TGRS.2023.3270712]
- Zhang Z Y, Li Y, Xie L L, Li S Q, Feng H, Siddique K H M and Lin G Z. 2024. Game analysis of future rice yield changes in China based on explainable machine-learning and planting date optimization. *Field Crops Research*, 317: 109557 [DOI: 10.1016/j.fcr.2024.109557]
- Zhao F, Sun R, Zhong L H, Meng R, Huang C Q, Zeng X X, Wang M Y, Li Y X and Wang Z Y. 2022. Monthly mapping of forest harvesting using dense time series Sentinel-1 SAR imagery and deep learning. *Remote Sensing of Environment*, 269: 112822 [DOI: 10.1016/J.RSE.2021.112822]
- Zhao X, Qi J B, Jiang J Y, Liu S B, Xu H F, Lin S M, Yu Z X, Li L Y and Huang H G. 2024. Fine-scale retrieval of leaf chlorophyll content using a semi-empirically accelerated 3D radiative transfer model. *International Journal of Applied Earth Observation and Geoinformation*, 135: 104285 [DOI: 10.1016/J.JAG.2024.104285]
- Zhen J N, Jiang X P, Xu Y, Miao J, Zhao D M, Wang J J, Wang J Z and Wu G F. 2021. Mapping leaf chlorophyll content of mangrove forests with Sentinel-2 images of four periods. *International Journal of Applied Earth Observation and Geoinformation*, 102: 102387 [DOI: 10.1016/j.jag.2021.102387]

A hybrid method for major food crops leaf chlorophyll content inversion driven by remote sensing mechanisms and deep learning

SHEN Yanyan^{1,2}, MENG Ran^{2,3}, LI Jiasheng³, ZHAO Ping^{1,2}, ZHAO Feng⁴, SUN Rui^{1,2},
ZHANG Hongyan⁵, NI Xiang⁵, LU Lijie^{2,3}, LIU Yong⁵, LIU Jie^{2,3}

1. College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China;

2. National Key Laboratory of Smart Farm Technologies and Systems, Harbin 150000, China;

3. School of Computer Science and Technology, Institute for Artificial Intelligence, Harbin Institute of Technology, Harbin 150000, China;

4. College of Forestry, Northeast Forestry University, Harbin 150000, China;

5. Beidahuang Information Company Limited, Harbin 150000, China

Abstract: The accurate estimation of Leaf Chlorophyll Content (LCC) is of considerable significance for crop physiological monitoring and precision agricultural management. However, traditional Vegetation Indices (VIs) based on visible-near infrared canopy reflectance experience notable challenges in LCC retrieval, as follows: (1) the spectral response is highly coupled with target information (LCC) and structural noise due to the confounding effects of canopy structural signals; and (2) canopy structural heterogeneity across different crop types further exacerbates the sensitivity of VIs to structural parameters, significantly limiting the generalization and applicability of

inversion models for major food crops, such as rice, wheat, and maize, under diverse growth conditions.

To address these issues, this study proposes a hybrid LCC retrieval framework that synergistically integrates remote sensing physical mechanisms with deep learning techniques, aiming to mitigate effectively the confounding influence of canopy structure and enhance model generalizability across diverse crop types. Specifically, a set of Vegetation Index Ratio Features (VIRFS) that exhibit low sensitivity to Leaf Area Index (LAI) is constructed by simulating a wide range of LAI-LCC combinations by using the PROSAIL radiative transfer model. Furthermore, an active learning-based strategy is incorporated into the transfer learning pipeline to optimize the selection of informative samples, enabling efficient model fine-tuning with limited labeled field data. The proposed method is systematically validated on multi-crop, multiregional datasets that are collected from three major agricultural zones in China: Northeast China (maize, rice, soybean), the Huang-Huai-Hai Plain (wheat), and the Yangtze River Basin (rice).

The results show the following: (1) the proposed hybrid method, which integrates radiative transfer mechanisms with deep learning, achieves excellent performance in cross-regional LCC retrieval for major staple crops. It exhibits strong stability and generalizability across diverse crop types with R^2 consistently exceeding 0.69 and root mean square error (RMSE) lower than $4.77 \mu\text{g}/\text{cm}^2$. (2) Compared with the conventional vegetation index feature set, the newly constructed VIRFS is designed to be less sensitive to LAI, significantly mitigating canopy structural interference under optimal fine-tuning conditions. Across the three major agricultural regions (Northeast China, the Huang-Huai-Hai Plain, and the Yangtze River Basin), VIRFS improves R^2 by 0.18–0.23 and reduces RMSE by 1.85 – $2.51 \mu\text{g}/\text{cm}^2$ for various staple crops. (3) The transfer learning framework that incorporates active learning enables high-accuracy LCC estimation by using only 30% of locally labeled samples, achieving R^2 values of 0.69–0.74 and RMSE of 4.98 – $5.76 \mu\text{g}/\text{cm}^2$ across different staple crops. This result represents a notable improvement over random sampling-based transfer learning, with R^2 gains of 0.02–0.16 and RMSE reductions of 0.05 – $1.42 \mu\text{g}/\text{cm}^2$, substantially enhancing model adaptability and inversion efficiency under limited label conditions.

In conclusion, the proposed inversion framework, which coupled physical principles with data-driven methods, significantly improves the accuracy and robustness of multi-crop LCC estimation. It provides a universal solution for nondestructive LCC monitoring across diverse crops and regions.

Key words: leaf chlorophyll content, UAV-based multispectral remote sensing, canopy structural heterogeneity, vegetation index ratio features, transfer learning, active learning

Supported by National Natural Science Foundation of China (No. 42471362); Key Research and Development Program of Heilongjiang, China (No. 2022ZX01A25, JD2023GJ01)