

基于KAN网络的近岸水体高光谱遥感反射率重构

黄廷婷¹, 范冬林^{1,2}, 何宏昌¹, 付波霖^{1,2}, 谭皓原³

1. 桂林理工大学 测绘地理信息学院, 桂林 541004;

2. 桂林理工大学 生态时空大数据感知服务重点实验室, 桂林 541006;

3. 自然资源部 第四海洋研究所, 北海 536015

摘要: 高光谱遥感反射率 R_{rs} (Remote sensing reflectance) 数据在近岸海洋水质参数反演中具有重要应用价值。然而, 高光谱传感器受限于技术复杂性与高成本, 且现场观测的高光谱数据受环境影响显著, 难以实现大范围覆盖。为克服上述限制, 本文提出一种基于KAN (Kolmogorov-Arnold Networks) 的高效高光谱 R_{rs} 重构模型, 该模型利用卫星遥感数据直接进行训练, 从多光谱 R_{rs} 中重建出与真实观测值分布特征高度一致且连续的高光谱 R_{rs} 。本文采用海岸海洋高光谱成像仪HICO (Hyperspectral Imager for the Coastal Ocean) 的2级 R_{rs} 数据为训练样本, 并将其重采样至6种主流多光谱传感器, 实现400—719 nm范围(1 nm间隔)的高光谱 R_{rs} 重构。实验表明, KAN模型在所有传感器上的重构性能均优于传统经验模型Li_2017及深度学习模型DNN (Deep Neural Network)、CNN (Convolutional Neural Network), 其平均均方根误差RMSD (Root Mean Square Deviation) 为 2.25×10^{-4} (sr⁻¹)、平均绝对误差MAE (Mean Absolute Error) 为 1.60×10^{-4} 、平均绝对百分比误差MAPE (Mean Absolute Percentage Error) 为0.0534、决定系数 R^2 (Coefficient of Determination) 为0.9982和偏差(Bias)为 -0.1×10^{-4} , 表现出良好的泛化能力与稳定性。进一步应用验证表明, 基于KAN重构的高光谱 R_{rs} 在叶绿素a (Chlorophyll-a, Chl-a) 浓度反演中展现出更优性能, 显著提升了反演精度, 尤其在高浓度区域表现更加出色。本文提出的KAN高光谱 R_{rs} 重构模型为突破传统模型对实测或模拟数据依赖的瓶颈、提升复杂水体遥感反演性能提供了新思路。

关键词: KAN网络, 高光谱遥感反射率, 遥感反射率重构, 水质参数反演, 近岸海域水体, 水体遥感反射率

中图分类号: TP79/P2

引用格式: 黄廷婷, 范冬林, 何宏昌, 付波霖, 谭皓原. 2026. 基于KAN网络的近岸水体高光谱遥感反射率重构. 遥感学报, 30(2): 385–402

Huang T T, Fan D L, He H C, Fu B L and Tan H Y. 2026. KAN-based hyperspectral remote sensing reflectance reconstruction in coastal waters. National Remote Sensing Bulletin, 30(2): 385–402 [DOI: 10.11834/jrs.20255297]

1 引言

遥感反射率 (R_{rs} , sr⁻¹) 数据是获取海洋光学参数和海洋生物特性的重要基础, 广泛应用于估算海洋、沿海和内陆水域的固有光学特性 (Yang 等, 2013; Kolluru 等, 2021; Jorge 等, 2021; O'Shea 等, 2023), 包括悬浮物浓度 (Shi 等, 2015)、有色溶解有机质CDOM (Colored Dissolved Organic Matter) (Zhu 等, 2014; Chen 等, 2017)、浮游植物Chl-a浓度 (Cui 等, 2020; Pahlevan 等, 2021) 等海洋水色要素。然而, 受技术和成本限制, 当前多数传感器仅提供多光谱 R_{rs} 数据产品,

尤其是在紫外波段和红边区域的观测上存在明显不足 (Wu 和 Sun, 2013), 难以全面刻画水体光学特性。相比之下, 高光谱反射率数据提供了连续且细致的光谱信息, 更有助于构建精度更高的水色参数反演模型 (Wu 和 Sun, 2013)。

因此, 如何基于多光谱数据重建高质量的连续高光谱 R_{rs} 数据, 成为当前遥感领域的重要挑战。针对这一问题, 学界提出了多种 R_{rs} 重构方法。一类方法是基于统计学原理的多元回归模型。例如, Lee_2014模型证实了水体 R_{rs} 在可见光波段间存在极强的线性相关性, 并证明利用简单的多元线性回归模型即可从十几个离散波段中高精度地重建

收稿日期: 2025-08-12; 预印本: 2025-12-10

基金项目: 广西自然科学基金(编号: 2022GXNSFBA035637); 广西科技重大专项(编号: AA23062035-2)

第一作者简介: 黄廷婷, 研究方向为海洋水色遥感。E-mail: 1020232083@glut.edu.cn

通信作者简介: 范冬林, 研究方向为智能计算与海洋遥感。E-mail: dlfan@glut.edu.cn

出完整的高光谱曲线 (Lee 等, 2014)。在此基础上, Sun_2015 模型成功地将该方案应用于光学特性极复杂的浑浊内陆水体, 并通过重新参数化进一步提高了重构精度 (Sun 等, 2015); 而 Li_2017 模型通过引入非线性对数变换, 在减少输入波段数量的同时进一步提高了重构性能 (Li 等, 2018)。然而, 此类方法实现高度依赖大量实测数据进行训练, 导致模型在时空和复杂水体环境中的泛化能力严重受限。另一类方法是基于生物光学传输和稀疏矩阵理论的模型, 如李渊等 (2019) 提出的基于稀疏表达的高光谱重建算法以及郭宇龙等 (Guo 等, 2022) 提出的 BBHR (Bio-optical-Based Hyperspectral Reconstruction) 模型结合生物光学模型与稀疏表示理论, 开创了一种物理驱动的高光谱 R_{rs} 重建方法。这类方法主要针对内陆水域, 通过模拟高光谱 R_{rs} 数据进行重构, 避免了对实测数据的依赖, 但物理模型的精度则受限于底层光学假设, 这使得它们在面对多变水体时泛化能力有限, 在光学特性更为复杂的近岸水域中的适用性仍需进一步验证。此外, 还有部分研究专注于特定波段重构, 例如蓝波段估算算法 (Wei 等, 2020)、紫外 R_{rs} 估算模型 (Liu 等, 2021)、以及基于深度学习的系统 UVISRdl (Ultraviolet and Visible bands Spectral Reconstruction using deep learning) (Wang 等, 2021)。这类方法在特定波段上具有较高重构精度, 但通常无法实现全谱段的高光谱 R_{rs} 重构。综上所述, 无论是依赖实测的统计模型, 还是依赖先验假设的物理模型, 现有主流方法普遍面临着对外部数据的强依赖性与在复杂水体中泛化能力不足两大共性问题。

近年来, 机器学习, 尤其是深度学习, 在地球科学与遥感领域得到了广泛应用 (Lary 等, 2016; Zhu 等, 2017)。凭借强大的特征提取与非线性拟合能力使其成为解决光谱重构等“不适定”问题的理想工具。传统物理方法在高光谱恢复中的适应性和精度受限, 越来越多研究转向采用深度学习等数据驱动方法进行建模 (Alibani 等, 2025)。尤其在陆地遥感等通用场景, 利用多光谱图像重建高光谱图像的深度学习已取得显著进展, 例如 Shi 等 (2018) 提出的 HSCNN+ (Hyperspectral Convolutional Neural Network) 模型, 通过引入深度残差与密集连接等先进的 CNN 架构, 极大地提升了图像光谱重建精度; Li 等 (2020) 提出的

AWAN (Adaptive Weighted Attention Network) 网络, 通过设计新颖的自适应加权注意力机制从 RGB 图像进行高精度光谱重建; 而 Zhao 等 (2024) 提出的混合 Transformer 架构, 通过新颖的行列交叉自注意力机制来专门处理多光谱遥感影像的光谱超分辨率任务。尽管深度学习模型在学习光谱间复杂映射关系方面展现了巨大潜力, 但对于水体 R_{rs} 信号的微弱性、低信噪比以及与陆地场景迥异的光谱特征, 这对模型的精度和鲁棒性提出了更高的要求。因此, 针对水体 R_{rs} 特性的高光谱 R_{rs} 重建研究相对有限。为此, 本文提出一种基于 KAN 网络的水体 R_{rs} 的深度学习重构模型, 旨在提升其在光学复杂近岸水域中的重建精度与适用性, 从而为后续水质参数的精准反演提供更可靠的数据基础。相较于 DNN 和 CNN 为代表的传统深度学习模型, KAN 的核心理论优势在于其将可学习的、由 B 样条参数化的激活函数置于网络连接 (边) 上, 而非节点上。这种“边激活”机制使模型能够自适应地学习输入特征间的最优非线性映射关系, 而非仅依赖固定激活函数的组合。研究表明, 该结构不仅在函数逼近精度和参数效率方面表现优异, 还展现出更佳的神经缩放特性, 从而在科学计算与物理建模等任务中展现出巨大潜力 (Liu 等, 2025)。基于上述优势, 本文首次将 KAN 引入光学复杂近岸水体的高光谱 R_{rs} 重构任务。

本文基于 KAN 神经网络提出了一种高效的近岸水体高光谱 R_{rs} 重构方法, 其核心创新在于构建了一个完全由遥感观测数据驱动的端到端重构范式, 该模型仅利用公开的 HICO 数据进行训练, 彻底摆脱了传统方法对现场实测或模拟数据的依赖, 极大地提升了模型的通用性与实用性。通过引入 KAN 独特的“边激活”机制和卓越的函数逼近能力, 模型能够有效应对水体 R_{rs} 信号微弱、低信噪比及高度非线性等特征, 实现对复杂水体光谱的高保真重建。实验结果进一步验证了该方法在不同传感器条件下的稳定性与鲁棒性明显优于传统深度学习模型。本文不仅为高光谱 R_{rs} 数据的获取提供了新思路, 更为复杂水体环境下的高精度遥感反演奠定了坚实的数据基础。

2 数据与数据处理

2.1 高光谱 R_{rs} 数据集

本文所采用的高光谱 R_{rs} 数据集源自搭载在国

际空间站上的HICO传感器。HICO是首台专为沿海海洋观测而设计的空间成像光谱仪,于2009年9月10日发射,其数据采集时段为2009年9月25日至2014年9月13日。该传感器空间分辨率为90 m,波段覆盖范围为353—1080 nm。其中,400—700 nm波段用于提取水体成分、测深及底部特性的信息(Corson等,2008)。HICO共包含128个波段,光谱分辨率为5.7 nm,信噪比大于200:1(Lucke等,2011),能够精确捕捉沿海区域的复杂光学特性,生成包括水体透明度、海底类型、底部特性、水的固有光学性质、叶绿素及悬浮沉积物等多种沿海环境产品(Corson等,2008)。本文使用2018年国际海洋颜色协调小组 IOCCG (International Ocean-Colour Coordinating Group) 海洋生物光学工作组再处理的HICO二级产品数据,该数据集覆盖了353—719 nm波段范围,共65个波段。为构建高质量训练样本,本文采用Salem等(2023)的数据处理流程:首先剔除光谱波段中含空值、负值以及相似 R_{rs} 值的整条光谱波段;然后对剩余数据进行三次样条插值,将其光谱分辨率提升至1 nm,最终获得6201385 (353—719 nm,共367个波段)个高光谱 R_{rs} 数据集。采用三次样条插值,是因为水体光谱本身具有的平滑特性为插值提供了物理基础,并且光谱插值是水色遥感领域进行算法开发和数据验证时的一项成熟且通行的技术(Fan等,2017; O'Reilly和Werdell,2019),其既能有效避免使用简单线性插值处理稀疏数据时平滑效果不佳的局限性(Salem等,2023)又可以规避高阶多项式拟合中固有的振荡风险(Fornberg和Zuev,2007),从而能更好地保证插值光谱的形态真实性,引导模型学习到正确的光谱映射关系。为平衡计算成本与模型性能,从该高光谱 R_{rs} 数据集中随机选取10万(400—719 nm,共320个波段)个 R_{rs} 光谱,作为后续模型构建与评估的数据集。

2.2 多光谱 R_{rs} 数据集构建

为获取模型训练所需的多光谱输入数据,将HICO的高光谱 R_{rs} 数据集(100000×320,400—719 nm)依据式(1),通过加权平均的方式重采样至6个多光谱传感器海洋与陆地颜色仪器S3A OLCI (Ocean and Land Colour Instrument)、中分辨率成像光谱仪MERIS (Medium Resolution Imaging Spectrometer)、中分辨率成像光谱仪MODIS

(Moderate Resolution Imaging Spectroradiometer)、宽视场海洋观测传感器SeaWiFS (Sea-viewing Wide Field-of-view Sensor)、Sentinel-2A搭载的多光谱仪器S2A MSI (Multi-Spectral Instrument)以及Landsat 8搭载的陆地成像仪OLI (Operational Land Imager)的相应波段,获得各个传感器对应的多光谱 R_{rs} 数据集(100000×B,B为传感器波段数,光谱响应如图1所示)。鉴于海洋生物光学变量被广泛认为近似服从对数正态分布(Campbell,1995), R_{rs} 也同样符合对数正态分布假设(Matsuoka等,2016)。为了增强模型对数值尺度和分布的适应性,将原始多光谱 R_{rs} 特征与其对数变换后的特征拼接,共同作为模型的输入。

$$R_{rs}(\lambda_i) = \frac{\sum_{j=\lambda_{\min}}^{\lambda_{\max}} R_{rs}(\lambda_j) \times \text{SRF}_i(\lambda_j)}{\sum_{j=\lambda_{\min}}^{\lambda_{\max}} \text{SRF}_i(\lambda_j)} \quad (1)$$

式中, $R_{rs}(\lambda_i)$ 为多光谱传感器在第*i*个波段的重采样遥感反射率, λ_j 高光谱数据的第*j*个波长,范围从 $\lambda_{\min}$ 到 $\lambda_{\max}$ 覆盖传感器第*i*个波段的光谱响应值范围, $\text{SRF}_i(\lambda_j)$ 为多光谱传感器第*i*波段在波长 λ_j 处的光谱响应函数SRF (Spectral Response Function)值, $R_{rs}(\lambda_j)$ 高光谱数据在波长 λ_j 处的遥感反射率。

2.3 实测高光谱 R_{rs} 数据与多光谱传感器 R_{rs} 产品数据

本文以SeaWiFS多光谱传感器为例,利用GLORIA_2022实测数据集与NASA SeaBASS (SeaWiFS Bio-optical Archive and Storage System)提供的卫星产品构建星—地匹配验证集,验证KAN模型在真实卫星数据上的高光谱重构能力。SeaBASS数据库提供严格时空匹配的卫星遥感数据与地面实测数据对,专门用于验证卫星数据产品和评估水色遥感算法的精度。本文从SeaBASS数据库中获取SeaWiFS(更新至2024年5月2日)多光谱 R_{rs} 产品(R2022再处理版本)。GLORIA_2022实测数据集包含7572个高光谱 R_{rs} 数据,波长范围为350—900 nm(1 nm间隔),涵盖超过30年的时间跨度和广泛的地理区域,涉及25个国家的450多个水体,包括湖泊、沿海水域、河口和河流等(Lehmann等,2022)。经过严格质量控制后(Lehmann等,2023),获得68组多光谱 R_{rs} 与实测高光谱 R_{rs} 样本。

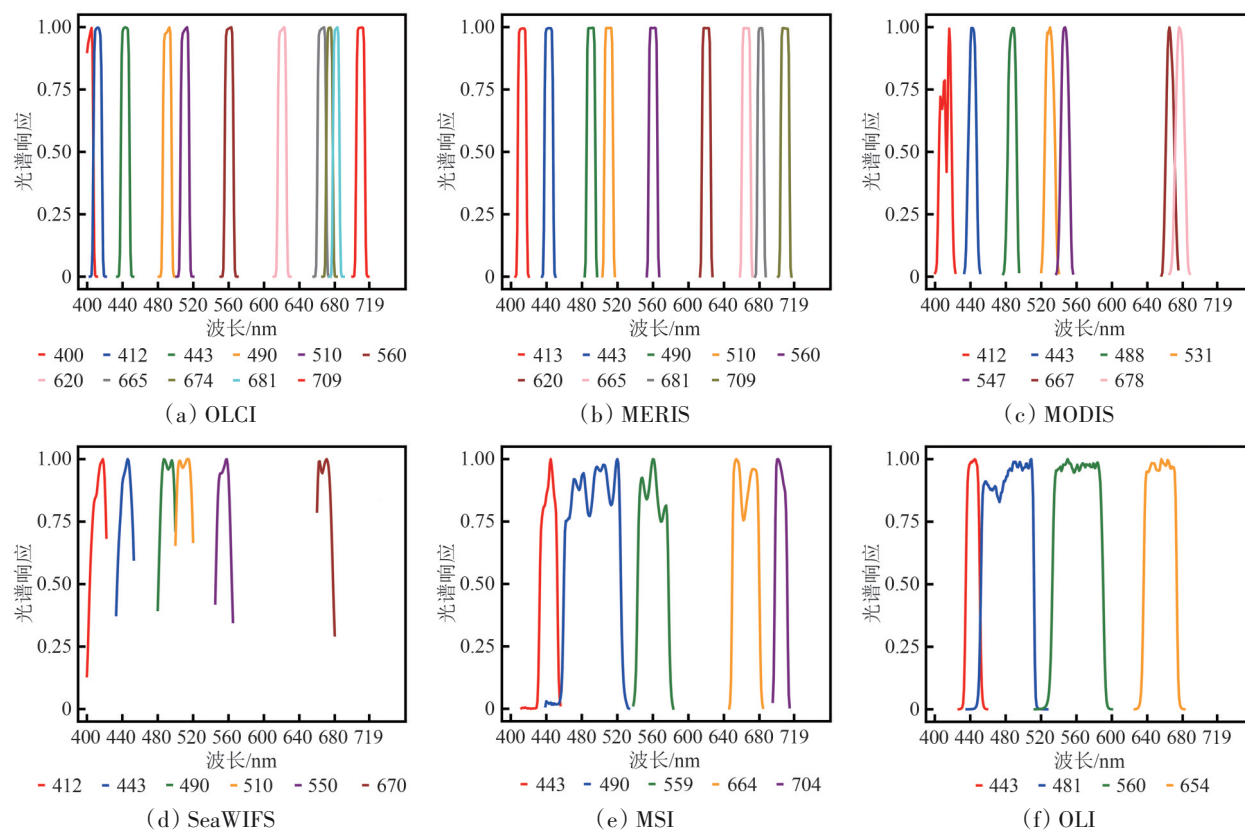


图1 6个传感器的光谱响应

Fig. 1 Spectral responses of six sensors

3 方法

3.1 KAN重构模型

KAN 是一种基于 Kolmogorov-Arnold 定理 (Arnold 等, 2009; Braun 和 Griebel, 2009) 的新型神经网络架构 (Liu 等, 2024), 其通过构建可学习的非线性函数来实现数据拟合 (Yang 等, 2024)。KAN 网络由多个 KAN Linear 层堆叠而成, 每一层的输出由传统的固定非线性函数和基于 B 样条 (B-spline) 函数线性加权组合共同决定, 实现对输入数据的复杂非线性关系表达。与传统神经网络将激活函数置于节点 (“节点激活”) 不同, KAN 创新性地将激活函数作用于网络连接的边上 (“边激活”), 而节点仅执行简单的线性求和。其核心优势在于能够更灵活地控制非线性变换的复杂度和作用范围, 同时保持网络结构的简洁性与可解释性。

本文构建的 KAN 模型结构中, 每个 KAN Linear 层将输入特征并行地送入两个计算分支: 基础分支和样条分支。在基础分支中, 输入特征经过一个

固定的非线性激活函数 SiLU (Sigmoid Linear Unit), 进行非线性变换。在样条分支中, 则由可学习的 B 样条基函数构成自适应的非线性激活函数。这种设计将一个全局的、固定的非线性与一个局部的、自适应的非线性变换相结合, 使得模型能够灵活捕捉输入数据中复杂的局部非线性关系, 从而能够以极高的效率精确拟合高光谱 R_{rs} 的复杂变化。最终, 每个 KAN Linear 层的输出由这两个分支相加而成。网络的最后一层负责将高维特征映射至目标高光谱 R_{rs} 输出维度, KAN 模型的网络架构如图 2 所示。

3.2 KAN 输入与训练策略

本文首先将完整的数据集 (共 10 万个高光谱—多光谱 R_{rs} 样本对) 按 7 : 1.5 : 1.5 的比例随机划分为训练集、验证集和测试集, 分别包含 70000 对、15000 对和 15000 对高光谱—多光谱 R_{rs} 样本。其中, 训练集用于对 KAN 模型进行训练, 验证集用于调整和选择最优超参数, 测试集则用于评估模型的最终性能。随后, 在数据预处理阶段, 为降低异常值对模型训练的影响, 采用 RobustScaler 对

所有输入特征和目标光谱进行归一化。在模型训练阶段, 为确保训练过程的稳定与高效, 采用AdamW (Adam with Decoupled Weight Decay) 优化器, 并结合包含正则项的均方误差损失函数。同时, 引入学习率调度与早停机制, 以提升模型的泛化能力并防止过拟合。考虑到不同多光谱传感器的波段数量和分布存在差异, 本文采用Optuna

超参数优化框架, 为每种传感器独立地、自动化地搜索最优模型配置, 包括网络深度 (1—4)、每层隐藏单元数 (64—512)、学习率与权重衰减系数等, 确保模型结构与特定传感器的输入信息动态匹配, 从而实现更高精度的高光谱 R_{rs} 重构。模型训练完成后, 可直接输入多光谱 R_{rs} 数据, 快速重构出对应的高光谱 R_{rs} 。

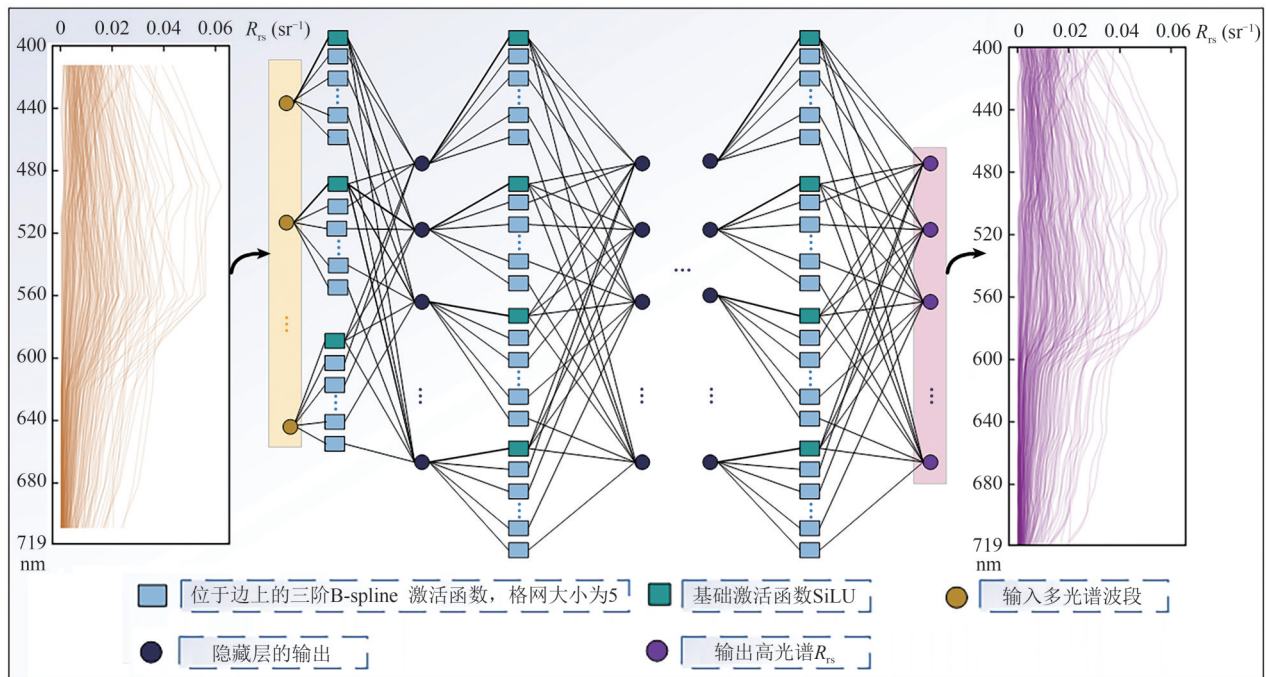


图2 KAN模型的网络架构

Fig. 2 The network architecture of the KAN model

在下游应用阶段, 以MODIS多光谱传感器为研究对象, 将其多光谱 R_{rs} 数据输入至训练好的KAN模型中, 重构出高光谱 R_{rs} 数据。最终, 将重构高光谱 R_{rs} 数据及其对应的MODIS原始多光谱 R_{rs} 数据分别基于混合密度网络MDN (Mixture Density Network) 模型 (Pahlevan 等, 2020) 进行Chl-a浓度反演。

3.3 对比模型

为全面评估和验证KAN模型的重构性能, 本文选取了3种具有代表性的 R_{rs} 重构模型进行对比分析, 涵盖了传统统计模型和深度学习模型不同方法。选择这些模型旨在通过与不同原理和复杂度的基准模型进行比较, 更全面地评估KAN模型在光谱重构精度、非线性关系处理能力以及潜在性能提升等方面的综合优势。

Li_2017模型 (Li 等, 2018) 是对Sun_2015模型 (Sun 等, 2015) 的进一步改进。Sun_2015模型基于多元线性回归模型, 是对Lee_2014方案 (Lee 等, 2014) 的改进, 通过引入新的参数化方法, 有效减少了重建高光谱 R_{rs} 时的不确定性, 从而提高了重建精度。Li_2017模型采用非线性对数模型, 基于实地观测的400—700 nm范围内的高光谱 R_{rs} 数据, 针对不同波段数的组合进行测试, 并通过选择八波段方案的参数化方法, 成功实现了不同水体类型的 R_{rs} 重构。Li_2017模型在解释相邻波段 R_{rs} 之间的相互依赖性方面表现出色, 达到了良好的重构效果, 模型表达式如式 (2) 所示。

$$R_{rs}^{rec}(\lambda_j) = \exp\left(\sum_{i=1}^M K_{ij} \ln(R_{rs}(\lambda_i))\right) \quad (2)$$

式中, $R_{rs}^{rec}(\lambda_j)$ 表示第 j 波段的重建 R_{rs} , $R_{rs}(\lambda_i)$ 表示用于重建的多光谱第 i 个波段, K_{ij} 表示由多元回归

确定的系数, M 表示用于重建的多光谱波段数。

DNN模型作为基准对比模型, 主要由于其作为成熟且灵活的深度学习解决方案, 所具备的强大非线性建模能力和自动特征提取优势及稳定的性能表现。DNN在实现不同传感器间 R_{rs} 产品的直接对比方面显示出良好性能 (Pahlevan 等, 2017b)。通过与DNN的对比, 主要聚焦于重构结果的性能差异, 以验证KAN在解析精度和计算效率上的提升, 确保评估结果在遥感光谱重构领域具有可比性和说服力, 为多光谱到高光谱 R_{rs} 重构任务提供可靠的深度学习参考基准。

CNN模型采用基于残差连接的架构, 主要基于其在光谱序列特征提取方面的独特优势。作为深度学习结构的典型代表, 其广泛应用于光谱图像分类、遥感反演和光谱分析等任务。CNN通过局部感受野和权值共享机制, 能够高效捕捉多光谱数据中相邻波段间的局部相关性。通过对比, 可以深入分析KAN模型在光谱局部特征建模和全局依赖关系学习方面的性能差异, 确保评估结果的全面性和可靠性。通过与这3个模型的深入比较, 可以更客观、全面地展示KAN模型在多光谱到高光谱 R_{rs} 重构任务中的综合表现。

3.4 评价指标

本文采用平均绝对百分比误差 (MAPE)、均方根误差 (RMSD)、平均绝对误差 (MAE)、决定系数 (R^2)、偏差 (Bias) 5个指标计算原始 R_{rs} 和重构的 R_{rs} 光谱波段之间的误差值, 全面评估重构模型的性能, 其中具体公式如下:

$$MAPE_{R_{rs}}(\lambda) = \frac{1}{n} \sum_n \left| \frac{R_{rs}(\lambda) - R_{rs}^{rec}(\lambda)}{R_{rs}(\lambda)} \right| \quad (3)$$

$$MAE_{R_{rs}}(\lambda) = \frac{1}{n} \sum_n |R_{rs}(\lambda) - R_{rs}^{rec}(\lambda)| \quad (4)$$

$$R_{R_{rs}}^2(\lambda) = 1 - \frac{\sum_n (R_{rs}(\lambda) - R_{rs}^{rec}(\lambda))^2}{\sum_n (R_{rs}(\lambda) - \bar{R}_{rs}(\lambda))^2} \quad (5)$$

$$Bias_{R_{rs}}(\lambda) = \frac{1}{n} \sum_n (R_{rs}(\lambda) - R_{rs}^{rec}(\lambda)) \quad (6)$$

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_{rs}(\lambda) - R_{rs}^{rec}(\lambda))^2} \quad (7)$$

式中, λ 为波长, n 为样本数, $R_{rs}(\lambda)$ 为原始的 R_{rs} , $R_{rs}^{rec}(\lambda)$ 为重构的 R_{rs} , $\bar{R}_{rs}(\lambda)$ 为所有原始 R_{rs} 的平均值。

4 实验与结果分析

4.1 基于卫星传感器的高光谱 R_{rs} 重构模型性能评估

为评估各模型的重构性能, 本文从定性和定量多个角度进行了系统性分析, 结果一致表明, KAN模型在重构精度、光谱细节保真度、稳定性及泛化能力上均显著优于对比模型。

首先, 从典型水体样本的定性对比 (图3) 中可以直观地看到各模型的性能差异。所有模型均能在宏观尺度上较好地复现原始反射光谱的关键典型物理特征, 如约 560 nm 的绿光峰、约 675 nm 的叶绿素吸收谷, 以及 680 nm 以上红边区域的反射率快速回升。然而, 模型在细节重构能力上, 尤其是在传感器波段缺失的光谱区间存在差异。如图中各局部波段放大子图展示的缺失区域的重构情况, 包括在紫蓝过渡区 (400—425 nm, 图3 (e)、(f))、蓝绿过渡区的光谱间隙 (450—480 nm, 图3 (c)、(d)) 以及在形态变化剧烈且数据缺失的红边区域 (大于 680 nm, 图3 (a) — (f)), 在这些关键区域中, KAN模型展现出明显优势, 能精准捕捉原始光谱的细微曲率和陡峭斜率, 几乎与原始值重合。相比之下, Li_2017模型在这些关键区域出现了显著的形态偏移, 而DNN和CNN模型的表现则介于二者之间。这初步表明, KAN模型在光谱细节保真度上具有最优的性能。

其次, 表1的定量评估结果系统地证实了KAN模型的卓越精度与鲁棒性。总体来看, KAN模型在所有6种传感器上均表现出最佳性能, 在各项误差指标 (RMSD、MAE、MAPE) 及 R^2 上均优于对比模型, 展现出卓越的重构精度、鲁棒性与泛化能力。在OLCI与MERIS两种光谱分辨率较高的传感器上, KAN模型性能最为突出, MAPE分别为 0.0225 和 0.0268, R^2 高达 0.9998 与 0.9997, RMSD低至 1.02×10^{-4} 与 1.21×10^{-4} , 几乎实现精确拟合。值得注意的是, KAN模型在所有传感器上的RMSD均低于 0.0005 sr^{-1} , 远低于 R_{rs} 数据可接受的不确定性上限 (Pahlevan 等, 2017b; Cetinić 等, 2018; Salem 等, 2023), 证明了其结果的可靠性。同时, 结果也揭示了模型性能与传感器波段数量的普遍关联: 从拥有11个波段的OLCI到仅有4个波段的OLI, 所有模型的重构精度整体呈下降趋

势。然而, KAN模型的精度下降幅度最小, 即便在最具挑战性的OLI上, 其 R^2 (0.9960) 仍远高于

其他模型, 体现了其对输入信息不足的良好适应性。

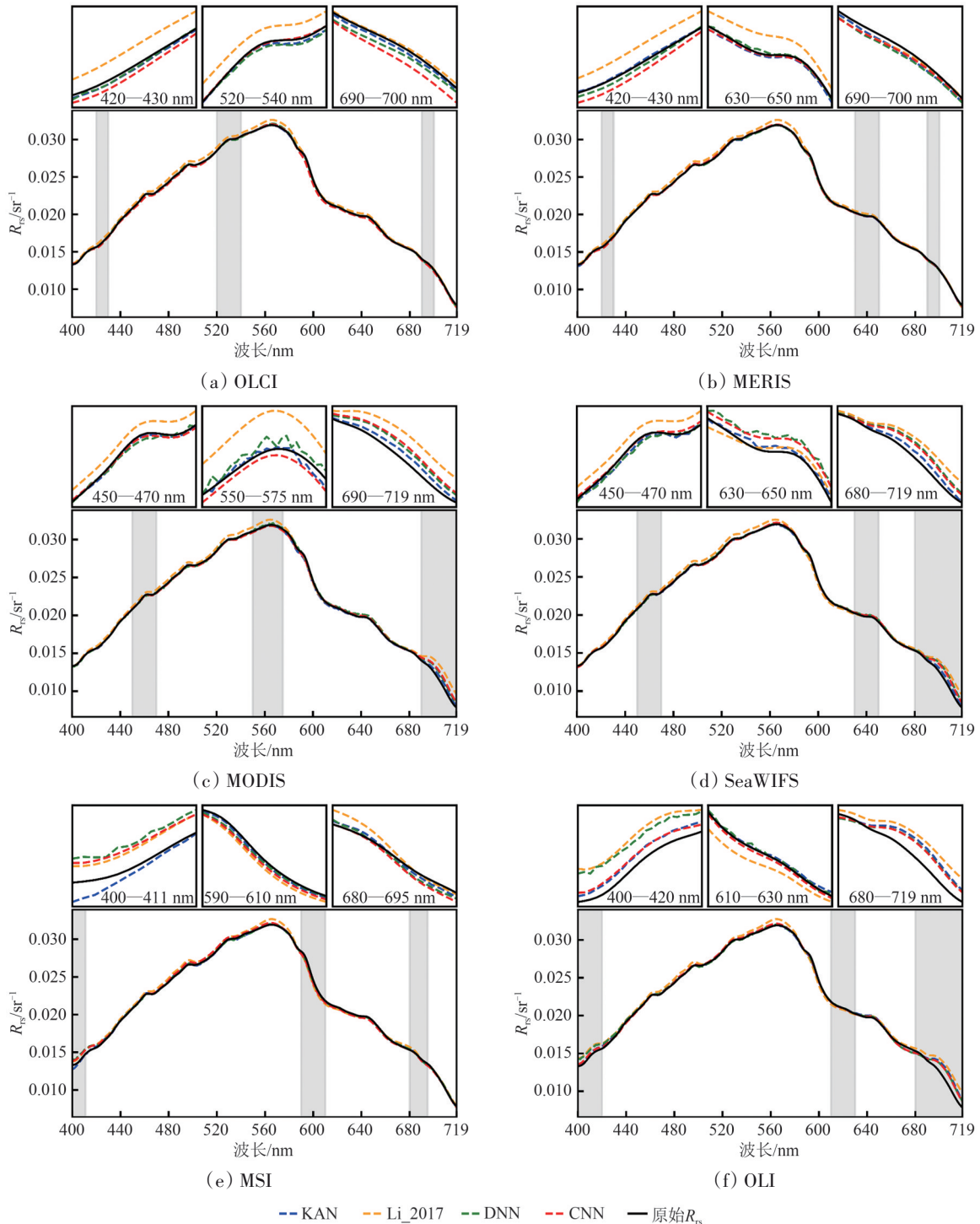


图3 4种模型(KAN, Li_2017, DNN, CNN)在6个传感器上重构的高光谱 R_{rs} 与原始高光谱 R_{rs} 对比

Fig.3 Comparison between reconstructed and original hyperspectral R_{rs} by four models (KAN, Li_2017, DNN, CNN) on six sensors

逐波段误差分析进一步揭示了输入信息不完备性对重构性能的局部影响, 图4直观地展示了各个模型在不同传感器上的指标 (R^2 、MAPE、MAE、Bias、RMSD) 随波长的变化情况。KAN模型在所

有6个传感器的重构性能都较为理想, 误差指标 (MAE图4 (c) 和RMSD图4 (e)) 始终处于所有模型中的最低水平, 且波动最为平缓, 表明其重构误差最小、稳定性最高。同时, 其 R^2 (图4 (a))

最接近于1，而Bias（图4（d））曲线最贴近零值线，共同证明了KAN模型的重构结果具有最高的精度和数据保真度。相比之下，DNN和CNN模型性能相近，但整体逊于KAN。Li_2017模型的性能最弱，其误差指标数值远高于其他模型，且波动剧烈，并存在明显的负向系统偏差。尽管模型间存在性能差异，但在传感器波段覆盖稀疏或缺失的光谱区域，例如短于420 nm、接近480 nm和590 nm，以及大于690 nm的红光与近红外区域，所有模型的重构误差都会一致增大，形成明显的

“误差峰”。这些误差峰的位置和大小因传感器差异而有所不同，与各传感器的波段“间隙”精确对应，例如：MSI和OLI在400—430 nm范围内缺乏有效覆盖且光谱响应较弱；MODIS、MERIS、OLCI和SeaWiFS在约450—480 nm范围内存在波段间隙；MODIS、MSI、OLI和SeaWiFS在600—640 nm及700—719 nm范围内存在光谱信息不足等。此现象充分表明，误差峰值通常对应于波段缺失较多的区段，输入信息不足是影响重构性能的主要原因。

表1 KAN重构模型在测试集上的评估指标值
Table 1 Evaluation metrics for the KAN reconstruction model on the test set

数据集	模型									
	KAN					Li_2017				
	RMSD ($\times 10^{-4}$)	MAE ($\times 10^{-4}$)	Bias ($\times 10^{-4}$)	R^2	MAPE	RMSD ($\times 10^{-4}$)	MAE ($\times 10^{-4}$)	Bias ($\times 10^{-4}$)	R^2	MAPE
OLCI	1.02	0.74	0.00	0.9998	0.0225	3.00	2.04	1.33	0.9990	0.0429
MERIS	1.21	0.88	0.00	0.9997	0.0268	3.21	2.22	1.40	0.9987	0.0526
MODIS	2.21	1.55	-0.02	0.9986	0.0595	4.68	3.19	1.45	0.9946	0.1121
SeaWiFS	2.63	1.79	-0.03	0.9977	0.0668	5.19	3.53	1.49	0.9937	0.1288
MSI	2.90	2.13	0.00	0.9973	0.0616	5.48	3.95	1.38	0.9936	0.1143
OLI	3.50	2.49	0.02	0.9960	0.0830	5.98	4.25	1.16	0.9909	0.1508

数据集	模型									
	DNN					CNN				
	RMSD ($\times 10^{-4}$)	MAE ($\times 10^{-4}$)	Bias ($\times 10^{-4}$)	R^2	MAPE	RMSD ($\times 10^{-4}$)	MAE ($\times 10^{-4}$)	Bias ($\times 10^{-4}$)	R^2	MAPE
OLCI	1.39	1.03	-0.04	0.9997	0.0313	1.95	1.47	-0.53	0.9995	0.0483
MERIS	1.55	1.14	0.03	0.9996	0.0353	1.75	1.30	0.16	0.9995	0.0387
MODIS	2.73	1.95	0.11	0.9981	0.0772	2.90	2.11	0.08	0.9982	0.0727
SeaWiFS	3.11	2.17	0.25	0.9972	0.0806	3.23	2.29	0.31	0.9975	0.0913
MSI	3.21	2.39	0.07	0.9970	0.0706	3.57	2.66	-0.62	0.9969	0.0784
OLI	3.77	2.72	-0.17	0.9956	0.0961	3.78	2.73	-0.02	0.9957	0.0914

最后，图5的箱线图分布从统计学上综合展现了各模型的稳定性和泛化能力。整体来看，KAN模型（图5（a））在性能指标（MAPE，MAE，Bias， R^2 ，RMSD）上，其中位数始终最低，箱体最窄，表明其典型误差最小且分布最集中；Li_2017（图5（b））的误差中位数最高、箱体最宽，误差大且波动明显；DNN（图5（c））和CNN（图5（d））介于二者之间，但离散程度仍明显高于KAN。离群点的分布进一步揭示了模型的鲁棒性差异，相较于其他3种模型，KAN的离群点最少，且主体误差区间紧凑，体现出更稳健的误差控制。在其他指标上差异同样明显，4个模型的 R^2 中位数均接近1，但KAN的箱体边缘与离群点更紧密地

聚集在1附近，拟合优度和稳定性最佳。在Bias方面，KAN的箱体几乎压在零线上，表现出接近完美的无偏性，而Li_2017则存在明显的系统性负偏差。跨传感器对比揭示了模型的泛化能力。此现象凸显了KAN模型在面对不同质量的输入数据时，拥有最强的稳定性和最可靠的泛化能力。

综上所述，KAN模型不仅在平均精度上领先，更在光谱细节保真度、稳定性及泛化能力等关键维度上，均显著优于传统经验模型和主流深度学习模型。同时，本文也系统地验证了输入光谱波段信息的数量和分布是高光谱重构性能的决定性因素。

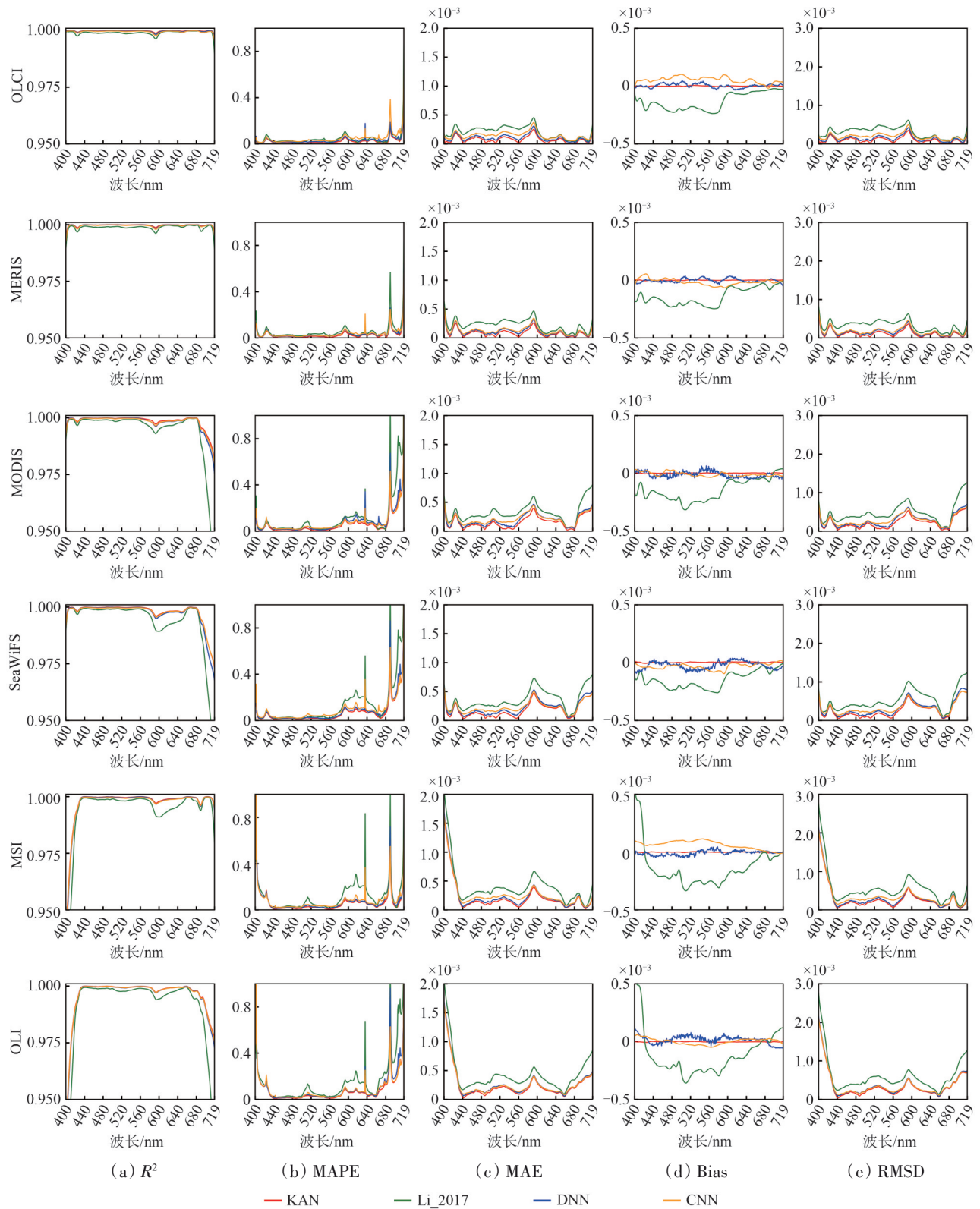


图4 4种模型(KAN, Li_2017, DNN, CNN)在6个传感器各波段上的性能指标对比

Fig. 4 Performance comparison of four models (KAN, Li_2017, DNN, and CNN) across spectral bands of six sensors

4.2 基于重构高光谱 R_{rs} 的近岸水体 Chla 反演

为验证重构高光谱 R_{rs} 数据在下游水质参数反演中的应用价值, 本文以 MODIS 传感器为例, 将其重构的高光谱 R_{rs} 数据与原始多光谱 R_{rs} 数据分别

用于 Chl-a 浓度的反演, 并与气溶胶机器人观测网 AERONET-OC (Aerosol Robotic Network-Ocean Color) 的实测数据进行对比验证。

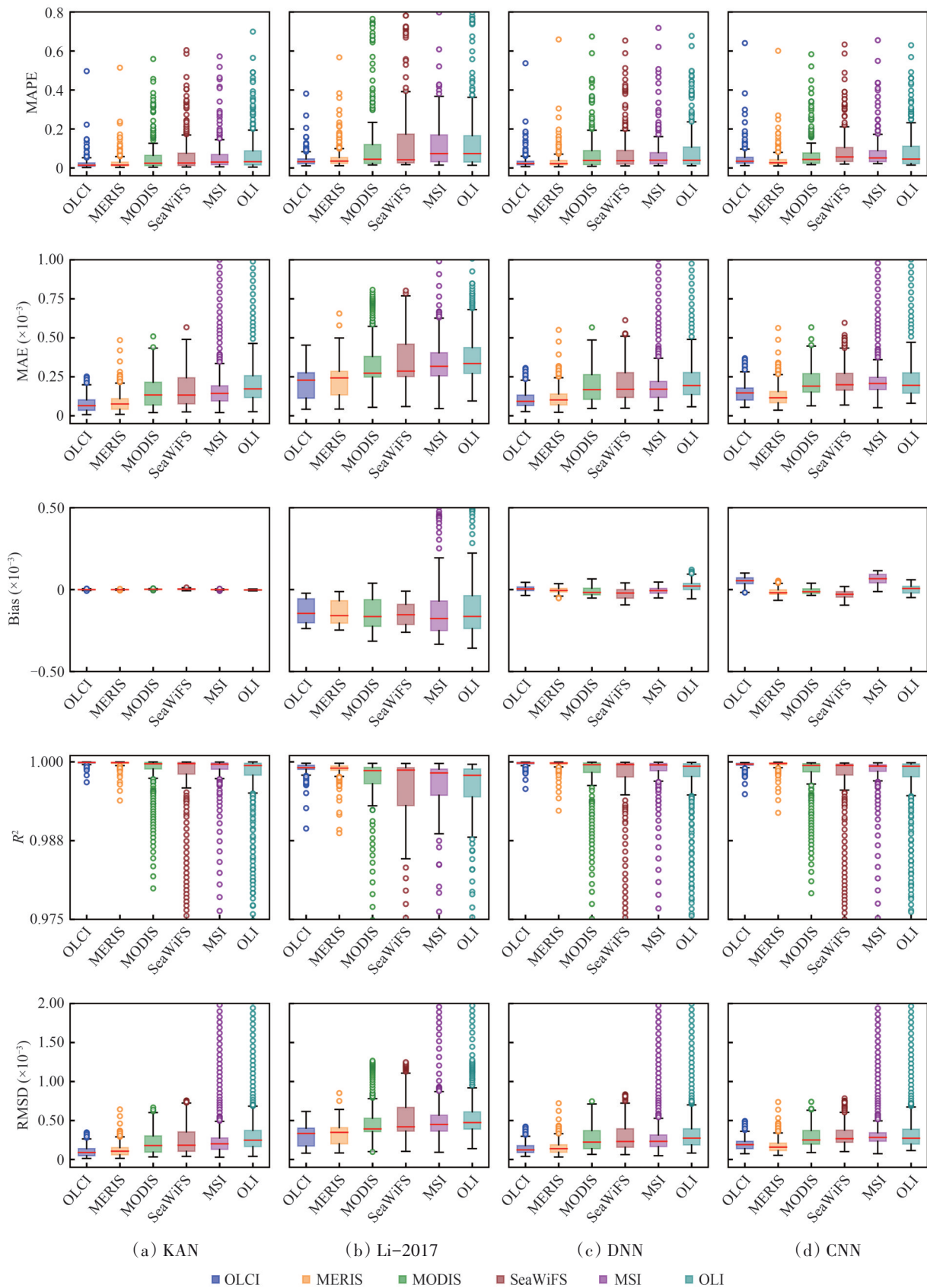


图5 4种模型(KAN, Li_2017, DNN, CNN)模型在各个传感器上的性能指标(MAPE, MAE, Bias, R^2 , RMSD)箱线图对比 ($n=15000$)

Fig. 5 Boxplot comparison of performance metrics (MAPE, MAE, Bias, R^2 , and RMSD) for four models (KAN, Li_2017, DNN, CNN) across various sensors ($n = 15000$)

本文使用的遥感与实测数据包括 MODIS Aqua 3 级产品数据与 AERONET-OC 实测 Chl-a 数据。MODIS Aqua 3 级产品由谷歌地球引擎 GEE (Google Earth Engine) 平台提供, 空间分辨率为 1 km, 时间跨度为 2002 年—2024 年, 广泛用于全球海洋颜色监测与生物地球化学分析 (Oyama 等, 2009; Chen 等, 2013)。AERONET-OC 数据来自美国国家航空航天局 NASA (National Aeronautics and Space Administration) AERONET 官网, 覆盖 2002 年—2024 年, 提供全球近岸站点的高质量实测离水辐

射数据和 Chl-a 浓度, 在遥感产品验证和生物光学建模研究中被广泛应用 (Zibordi 等, 2009; Moore 等, 2015; Pahlevan 等, 2017a; Valente 等, 2019)。本文通过时间 (± 4 h) 和空间匹配策略, 从 GEE 平台提取与 AERONET-OC 站点对应的 MODIS 多光谱 R_{rs} 数据, 获得 1804 个有效配对样本, 按比例 7 : 3 随机划分为训练集和测试集用于重构高光谱 R_{rs} , 最终获得 542 个反演样本, AERONET-OC 实测 Chl-a 数据集信息如表 2 所示。

表 2 AERONET-OC 实测 Chl-a 数据集的信息
Table 2 Information of the AERONET-OC measured Chl-a dataset

数据集	样本数	最小值/(mg/m ³)	最大值/(mg/m ³)	平均值/(mg/m ³)	标准差/(mg/m ³)	中位数/(mg/m ³)
AERONET-OC	1804	0.07	42.59	2.44	3.56	1.23

考虑到 R_{rs} 数据反演 Chl-a 浓度是一个典型的“不适定”问题, 即单一的 R_{rs} 光谱可能对应多个不同的 Chl-a 值 (Defoin-Platel 和 Chami, 2007; Yang 等, 2011)。MDN 模型能够有效缓解反演中的非唯一性问题, 并在内陆和沿海等多种生物光学条件下表现出优于传统物理或经验算法的预测性能 (Pahlevan 等, 2020)。因此, 本文采用 MDN 模型进行 Chl-a 反演, 其架构设计直接使用 Pahlevan 论文中的设计 (Pahlevan 等, 2020), 配置五层隐藏层, 每层 100 个神经元, 高斯混合分量设为 5, 以捕捉 Chl-a 预测值的复杂概率分布。为确保对比的公平性, 所有反演任务均采用统一的 MDN 架构。

图 6 与表 3 共同展示了基于 4 种模型重构的高光谱 R_{rs} 在 Chl-a 浓度反演任务中的应用效果, 并与 MODIS 多光谱 R_{rs} 反演结果进行了对比。从散点分布可以直观地看出, 基于重构高光谱 R_{rs} 的反演结果普遍比直接使用 MODIS 多光谱的结果更紧密地聚集在 1 : 1 对角线附近, 表明光谱重建有效提升了反演精度。在 4 种模型中, KAN 模型 (图 6 (a)) 重构高光谱 R_{rs} 的反演结果分布最为集中, 与对角线拟合度最高, 尤其是在高浓度区域 (如 Chl-a > 10 mg/m³), 其低估现象最轻微, 表现最为稳健。相比之下, Li_2017 模型 (图 6 (b)) 重构高光谱 R_{rs} 的反演结果最为离散, 甚至劣于原始多光谱反演, 而 DNN (图 6 (c)) 与 CNN (图 6 (d)) 的表现则介于两者之间。表 3 的量化指标进一步证实了这些视觉观察结果: 基于 KAN 重构光谱的反演取得了最优精度 (RMSE=0.78 mg/m³, R^2 =0.94), 显著优于 MODIS 多光谱反演 (RMSE=0.93 mg/m³, R^2 =

0.91) 及其他所有模型, 而 Li_2017 模型的重构结果在反演中表现最差, 其 RMSE (1.04 mg/m³) 甚至高于基准值, 表明其重构引入的误差对反演造成了负面影响。与多光谱相比, KAN 重构高光谱 R_{rs} 的应用使 RMSE、MAE 和 RMSLE 分别降低了约 16.13%、11.11% 和 7.14%, 同时 R^2 提升了约 3.30%。综合图表结果可见, 高光谱重建技术显著提升了 Chl-a 反演精度, 尤其在高浓度水体中表现更为突出。其中, KAN 模型在所有对比方法中效果最优。重构的高光谱 R_{rs} 有效弥补了 MODIS 传感器在高浓度区域估算能力不足的局限, 提升了整体反演性能与精度, 展现出更强的适应性与泛化能力。

5 讨 论

5.1 模型重构性能与传感器配置分析

本文对比了 4 种高光谱 R_{rs} 重构模型在 6 个多光谱传感器上的性能精度, 从实验结果来看, OLCI 的重构质量始终最佳, 其次是 MERIS, 然后是 MODIS, 之后依次为 SeaWiFS、MSI, 而 OLI 的重构质量最差。这种性能差异与各传感器的波段数量密切相关。其中, OLCI 拥有最多的 11 个覆盖均匀的波段, MERIS、MODIS 和 SeaWiFS 分别具有 9、7 和 6 个波段, 而 MSI 和 OLI 则分别只有 5 个和 4 个波段。已有大量研究表明, 波段数量的增加不仅提升了光谱信息的维度, 为模型提供了更丰富的特征信息, 也增强了模型捕捉水体光谱特征的能力 (Landgrebe, 2002; Bioucas-Dias 等, 2013; Gholizadeh 等, 2016)。尤其是 KAN 模型, 通过引

入可学习激活函数参数来增强网络学习能力,能够更有效地提取输入数据中的高阶特征,因此在高维光谱输入下表现更优。相反,较少波段数量的传感器可能无法提供足够多的光谱特征,从而限制了高光谱 R_n 的准确重建,这与Li等(2018)的研究结果相一致。此外,波段的具体分布位置对局部波段的重构质量具有重要影响。从图4分析得到,各传感器在不同波长上的误差峰值位置存在显著差异。这主要是因为模型的重构精度高度依赖于输入波段的分布,当某一光谱区域(如红边、近红外区)缺乏邻近的波段输入时,模型对此处的光谱形态约束减弱,导致重构不确定性增加,形成误差峰值。例如,OLCI与MERIS在红边

区域设置了更密集的波段,这使得模型能够更精确地捕捉该区域光谱曲线的陡峭斜率和细微形态,而MODIS和OLI在该区域波段缺失或间隔较大,导致模型难以准确拟合光谱的快速变化,从而造成局部重构误差显著增大。综合来看,本文所提出的KAN模型凭借其灵活的结构,能够适应从OLCI到OLI等光谱信息差异显著的多种传感器条件,不仅在光谱信息丰富的场景下实现高精度重构,即便在光谱信息受限的条件下也能保持较高拟合度,充分展现了其在 R_n 重构任务中的广泛适用性与显著优势。这一高精度的重建结果将有助于后续水体特征分析并为水质参数反演提供更有权威性的数据基础。

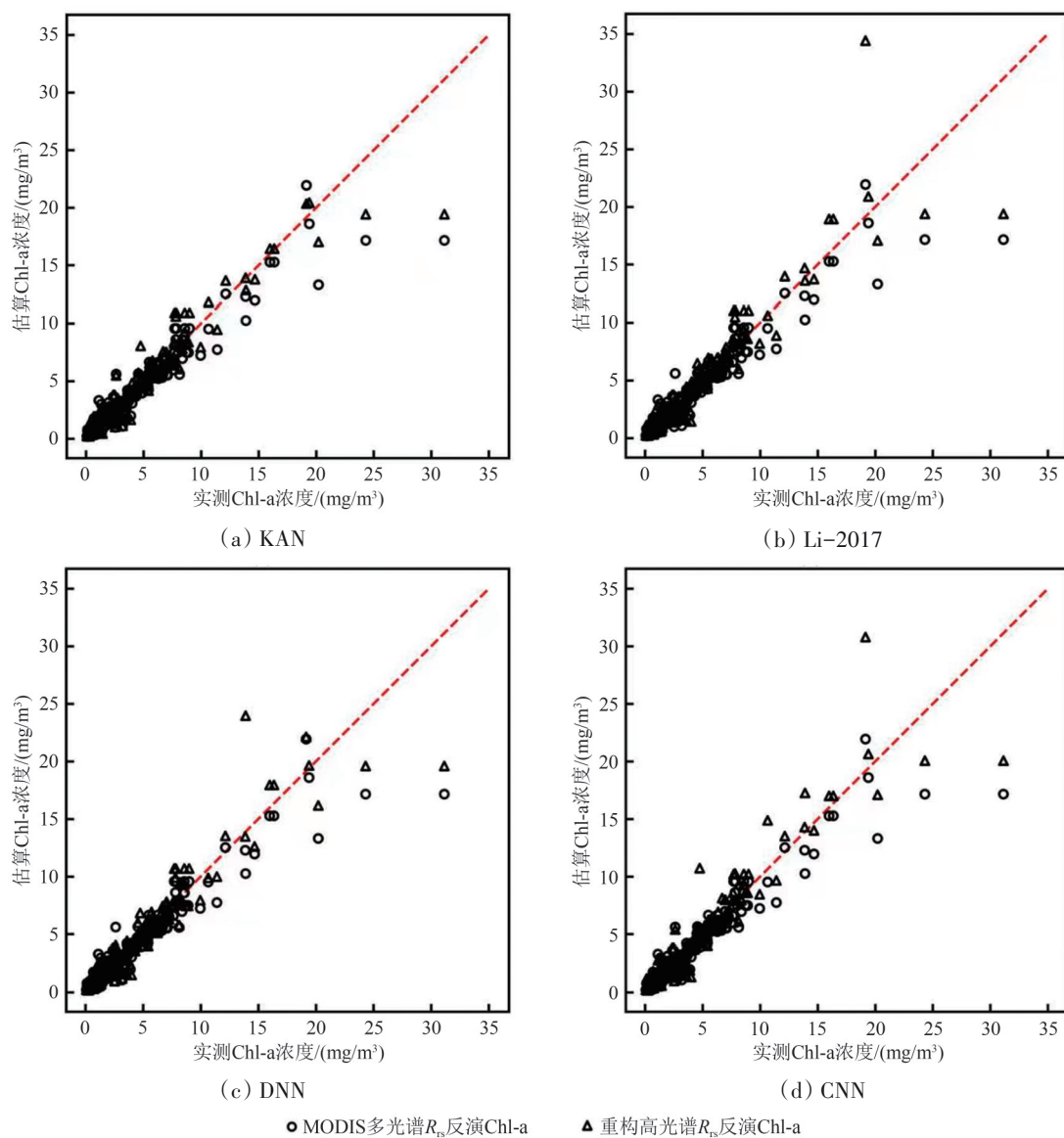


图6 4种模型重构的高光谱 R_n 与MODIS多光谱 R_n 反演Chl-a浓度的散点图对比

Fig.6 Scatter plot comparison of Chl-a concentrations retrieved from reconstructed hyperspectral R_n and MODIS multispectral R_n by four models

表3 基于MDN模型反演Chl-a浓度的误差值
Table 3 Error values of Chl-a concentration retrieval based on the MDN model

重构模型	重构高光谱 R_{rs} 反演Chl-a浓度				MODIS多光谱 R_{rs} 反演Chl-a浓度			
	RMSE	MAE	R^2	RMSLE	RMSE	MAE	R^2	RMSLE
KAN	0.78	0.32	0.94	0.13	0.93	0.36	0.91	0.14
Li_2017	1.04	0.37	0.89	0.14	—	—	—	—
DNN	0.89	0.34	0.92	0.13	—	—	—	—
CNN	0.95	0.35	0.91	0.14	—	—	—	—

5.2 重构高光谱数据在水质参数反演中的应用价值

既有研究已充分论证，高光谱遥感数据在水质参数反演中的精度优势，主要体现在敏感波段的识别及光谱混叠问题区分方面（Sudduth等，2015；Hestir等，2015；Kudela等，2015；Yang等，2022）。通过重构高光谱 R_{rs} 数据反演Chl-a浓度，得到了比多光谱数据更高的精度与稳定性。这一性能提升的关键在于重构高光谱数据所提供的连续、细粒度光谱信息，使模型能够更准确地表征水体中复杂的光学非线性响应。随着水体中Chl-a浓度的升高，色素的吸收效应与颗粒物的散射效应呈现出愈发复杂的非线性关系，这在光谱上体现为红边位置、峰值形态等特征的微妙变化。多光谱传感器受限于其较宽的波段设置和离散的光谱采样，难以有效解析此类细微的非线性特征，在高浓度条件下易导致信号饱和或光谱混淆，从而引起反演误差的急剧增加。与此相对，重构后的高光谱数据凭借其光谱连续性和高分辨率，能够更完整地保留这些细粒度的光谱信息，为反演模型提供了在高浓度水平下进行精确判别的关键依据，因此在高浓度区间依然维持了较优的反演性能。这种优势不仅体现在Chl-a，还可推广至悬浮物浓度、CDOM等其他典型水质参数的反演任务中。更重要的是，高光谱 R_{rs} 重构为下游多任务联合建模提供了更强的光谱支撑。例如，在多参数协同反演、光谱衍生指数构建、以及基于物理模型的水色分类等任务中，重构高光谱数据均可作为模型提供更丰富、更连续的谱段信息，从而增强模型的判别能力和鲁棒性。

5.3 KAN模型在真实卫星数据上的性能分析与局限性

为深入评估KAN模型在真实卫星数据上的性能与局限性，本文将精确匹配到的68个SeaWiFS

多光谱 R_{rs} 数据直接输入对应训练好的KAN模型获得重构高光谱 R_{rs} ，并与实测高光谱数据对比。量化结果（表4）显示，模型整体表现稳健，展现出强大的泛化能力（ $R^2=0.8556$ ， $RMSE=0.0013\text{ sr}^{-1}$ ），图7进一步展示了3组不同Chl-a浓度下的重构与实测高光谱 R_{rs} 对比。

表4 SeaWiFS重构高光谱 R_{rs} 的误差值
Table 4 Errors in hyperspectral R_{rs} reconstructed from SeaWiFS

样本数	RMSD	MAE	Bias	R^2
68	0.0013	0.0009	-0.0001	0.8556

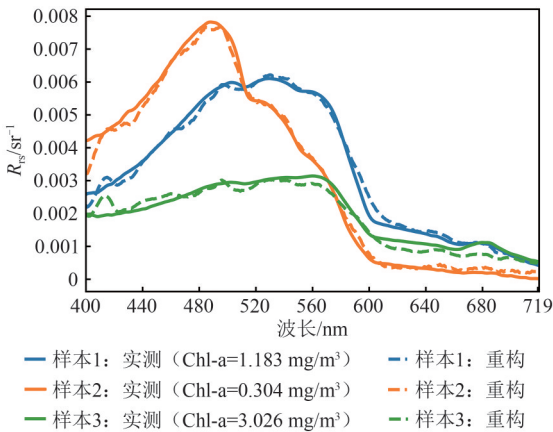


图7 基于KAN模型的SeaWiFS多光谱 R_{rs} 数据重构高光谱 R_{rs} 与实测高光谱 R_{rs} 对比

Fig. 7 Comparison between the hyperspectral R_{rs} reconstructed from SeaWiFS multispectral R_{rs} data using the KAN model and the in-situ measured hyperspectral R_{rs} .

结果显示，KAN重构模型整体上能够较为准确地再现高光谱 R_{rs} 的形态与数值特征，说明其已经学习到了普适的生物光学规律，并能够将其迁移到真实多光谱传感器数据上。然而，在400—440 nm蓝光区、560—660 nm波段缺失区域以及大于680 nm红边区域，模型在不同浓度水体中仍表现出一定的低估或高估，尤其在低浓度（0.304 mg/m³）和中高浓度样本（3.026 mg/m³）偏差更为明显。

这一现象主要源于输入波段信息稀疏所导致的约束不足：在沿海和内陆水域，短蓝波段的卫星 R_{rs} 本身存在较大不确定性，且信号易受 CDOM 强吸收与大气瑞利散射干扰（Gordon, 1997; Wei 等, 2020），红边区域则由 Chl-a 的强吸收与散射主导（Gitelson 等, 2007）。已有研究表明 SeaWiFS 的较低信噪比进一步放大了不确定性，其在蓝光波段的不确定性约为 5%，在绿色波段增加至 10%—15%，而红光波段则超过 20%，不确定性普遍偏大（Hu 等, 2013）。这些结果与本文的误差表现相互印证，表明即便是 KAN 这样性能卓越的模型，输入数据的稀疏性与噪声水平仍会直接转化为重构精度的下降和系统性偏差。总体来看，KAN 模型能够显著提升多光谱数据的光谱解析度，并在与真实数据中表现稳健，但其精度仍受限于源数据的质量与信息量。未来研究应探索多传感器数据融合，利用互补波段弥补单一传感器的不足，从而实现更高精度和更广适用性的水体 R_{rs} 重构。

在模型效率方面，KAN 模型引入了基于 B 样条的可学习激活函数，其计算成本体现在前期的超参数搜索阶段，约需 6 h 才能确定最优架构。与 DNN 相比，这一调优过程耗时更长；但在实际训练阶段，KAN 的训练速度优于 CNN，体现出较高的参数效率与收敛速度。实验证明，最优的 KAN 模型具有卓越的学习效率，仅需少量训练轮次即可收敛，远少于传统深度学习模型，从而以前期时间成本换取了学习效率和重构精度上的双重优势，实现了性能与效率的平衡。

6 结 论

针对现有高光谱 R_{rs} 数据获取困难，以及传统重构方法对现场实测或模拟数据依赖强、泛化能力有限的问题，本文提出了一种基于 KAN 的高光谱 R_{rs} 重构新方法。该方法通过引入可学习的 B 样条核函数，在端到端框架下直接从多光谱卫星数据学习水体光谱特性的复杂非线性映射关系，有效重建出与真实观测值分布特征一致的高质量高光谱 R_{rs} 数据。与传统经验模型 Li_2017 以及 DNN、CNN 深度学习模型相比，KAN 模型在 6 种主流多光谱传感器上均展现出更优的光谱重构性能，尤其是在输入光谱波段信息受限时仍能保持较强的鲁棒性。在下游的 Chl-a 反演应用中，基于 KAN 重构高光谱数据的 Chl-a 反演精度显著优于直接使

用多光谱数据反演结果，该优势在高浓度水体中尤为突出，验证了高光谱信息对于提升复杂水体参数遥感反演性能的重要价值。KAN 模型虽提升了表达能力，但同时引入了更多可调超参数，导致模型训练过程在时间和调参成本上较传统模型更高，仍需进一步优化其结构与计算效率。总体而言，本文提出的 KAN 高光谱 R_{rs} 重构方法克服了传统方法对现场实测或模拟数据的依赖较大的局限性，具有较高的适用性与可迁移性，为现有卫星多光谱数据的深度利用开辟了新途径。该方法在六种主流多光谱传感器上均取得优异的重构性能，展现出良好的跨传感器适用性；同时，在全球范围的近岸水体样本上表现出显著的泛化能力，表明该方法不仅适用于不同传感器的数据，也具备跨地域的应用潜力。随着多光谱传感器种类和数据量的不断增加，该重构方法具备广泛应用于不同地区近岸海洋与水质监测的潜力，能够为海洋水色遥感技术的发展提供新的数据支持。未来将进一步拓展该重构方法在悬浮物浓度、有色溶解有机质等水质参数反演中的应用，并结合多源多尺度遥感数据与物理先验机制，进一步提升模型的可解释性与工程推广能力。

志 谢 本次实验所使用的数据包括：HICO Level-2 R_{rs} 数据集由美国国家航空航天局海洋生物学处理小组提供；AERONET-OC 实测数据由 NASA 和法国国家科学研究中心（CNRS）长期提供；GLORIA_2022 实测高光谱 R_{rs} 数据由地球与环境科学数据发布平台发布；MODIS Aqua 3 级产品通过 GEE 平台获取。谨此向所有为本研究提供数据支持的机构及其工作人员致以诚挚的感谢！

参考文献 (References)

- Alibani M, Pastorino M, Moser G and Acito N. 2025. Investigating the potential of deep learning approaches in the reconstruction of VNIR-SWIR hyperspectral data from multispectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18: 14215-14227 [DOI: 10.1109/JSTARS.2025.3575518]
- Arnold V I. 2009. On the representation of functions of several variables as a superposition of functions of a smaller number of variables. Vladimir I. Arnold - Collected Works, 1: 25-46 [DOI: 10.1007/978-3-642-01742-1_5]

- Bioucas-Dias J M, Plaza A, Camps-Valls G, Scheunders P, Nasrabadi N and Chanussot J. 2013. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, 1(2): 6-36 [DOI: 10.1109/MGRS.2013.2244672]
- Braun J and Griebel M. 2009. On a constructive proof of Kolmogorov's superposition theorem. *Constructive Approximation*, 30(3): 653-675 [DOI: 10.1007/s00365-009-9054-2]
- Campbell J W. 1995. The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research: Oceans*, 100(D7): 13237-13254 [DOI: 10.1029/95JC00458]
- Cetinić I, McClain C R and Werdell P J. 2018. Pre-Aerosol, Clouds, and ocean Ecosystem (PACE) Mission Science Definition Team Report. PACE Technical Report Series, 2: NASA/TM-2018-219027.
- Chen J, Zhang M W, Cui T W and Wen Z H. 2013. A review of some important technical problems in respect of satellite remote sensing of chlorophyll-a concentration in coastal waters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(5): 2275-2289 [DOI: 10.1109/JSTARS.2013.2242845]
- Chen J, Zhu W N, Tian Y Q and Yu Q. 2017. Estimation of colored dissolved organic matter from landsat-8 imagery for complex inland water: case study of Lake Huron. *IEEE Transactions on Geoscience and Remote Sensing*, 55(4): 2201-2212 [DOI: 10.1109/TGRS.2016.2638828]
- Corson M R, Korwan D R, Lucke R L, Snyder W A and Davis C O. 2008. The hyperspectral imager for the coastal ocean (HICO) on the international space station//IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium. Boston: IEEE, IV-101-IV-104 [DOI: 10.1109/IGARSS.2008.4779666]
- Cui T W, Zhang J, Wang K, Wei J W, Mu B, Ma Y, Zhu J H, Liu R J and Chen X Y. 2020. Remote sensing of chlorophyll a concentration in turbid coastal waters based on a global optical water classification system. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163: 187-201 [DOI: 10.1016/j.isprsjprs.2020.02.017]
- Defoin-Platel M and Chami M. 2007. How ambiguous is the inverse problem of ocean color in coastal waters? *Journal of Geophysical Research: Oceans*, 112(C3): C03004 [DOI: 10.1029/2006JC003847]
- Fan Y Z, Li W, Gatebe C K, Jamet C, Zibordi G, Schroeder T and Stamnes K. 2017. Atmospheric correction over coastal waters using multilayer neural networks. *Remote Sensing of Environment*, 199: 218-240 [DOI: 10.1016/j.rse.2017.07.016]
- Fornberg B and Zuev J. 2007. The Runge phenomenon and spatially variable shape parameters in RBF interpolation. *Computers & Mathematics with Applications*, 54(3): 379-398 [DOI: 10.1016/j.camwa.2007.01.028]
- Gholizadeh M H, Melesse A M and Reddi L. 2016. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors*, 16(8): 1298 [DOI: 10.3390/s16081298]
- Gitelson A A, Schalles J F and Hladik C M. 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment*, 109(4): 464-472 [DOI: 10.1016/j.rse.2007.01.016]
- Gordon H R. 1997. Atmospheric correction of ocean color imagery in the Earth observing system era. *Journal of Geophysical Research: Atmospheres*, 102(D14): 17081-17106 [DOI: 10.1029/96JD02443]
- Guo Y L, Huang C C, Li Y M, Du C G, Shi L F, Li Y, Chen W Q, Wei H J, Cai E X and Ji G X. 2022. Hyperspectral reconstruction method for optically complex inland waters based on bio-optical model and sparse representing. *Remote Sensing of Environment*, 276: 113045 [DOI: 10.1016/j.rse.2022.113045]
- Hestir E L, Brando V E, Bresciani M, Giardino C, Matta E, Villa P and Dekker A G. 2015. Measuring freshwater aquatic ecosystems: the need for a hyperspectral global mapping satellite mission. *Remote Sensing of Environment*, 167: 181-195 [DOI: 10.1016/j.rse.2015.05.023]
- Hu C M, Feng L and Lee Z. 2013. Uncertainties of SeaWiFS and MODIS remote sensing reflectance: implications from clear water measurements. *Remote Sensing of Environment*, 133: 168-182 [DOI: 10.1016/j.rse.2013.02.012]
- Jorge D S F, Loisel H, Jamet C, Dessailly D, Demaria J, Bricaud A, Maritorena S, Zhang X D, Antoine D, Kutser T, Bélanger S, Brando V O, Werdell J, Kwiatkowska E, Mangin A and d'Andon O F. 2021. A three-step semi analytical algorithm (3SAA) for estimating inherent optical properties over oceanic, coastal, and inland waters from remote sensing reflectance. *Remote Sensing of Environment*, 263: 112537 [DOI: 10.1016/j.rse.2021.112537]
- Kolluru S, Gedam S S and Inamdar A B. 2021. Subcomponent inherent optical properties retrieval from total absorption coefficient and remote sensing reflectance measured in coastal waters. *Journal of Earth System Science*, 130(3): 159 [DOI: 10.1007/s12040-021-01661-8]
- Kudela R M, Palacios S L, Austerberry D C, Accorsi E K, Guild L S and Torres-Perez J. 2015. Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. *Remote Sensing of Environment*, 167: 196-205 [DOI: 10.1016/j.rse.2015.01.025]
- Landgrebe D. 2002. Hyperspectral image data analysis. *IEEE Signal Processing Magazine*, 19(1): 17-28 [DOI: 10.1109/79.974718]
- Lary D J, Alavi A H, Gandomi A H and Walker A L. 2016. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1): 3-10 [DOI: 10.1016/j.gsf.2015.07.003]
- Lee Z, Shang S L, Hu C M and Zibordi G. 2014. Spectral interdependence of remote-sensing reflectance and its implications on the design of ocean color satellite sensors. *Applied Optics*, 53(15): 3301-3310 [DOI: 10.1364/AO.53.003301]
- Lehmann M K, Gurlin D, Pahlevan N, Alikas K, Anstee J M, Balasubramanian S V, Barbosa C C F, Binding C, Bracher A, Bresciani M, Burtner A, Cao Z G, Conroy T, Dekker A G, Di Vittorio C, Drayson N, Errera R M, Fernandez V, Ficek D, Fichot C G, Gege P, Giardino C, Gitelson A A, Greb S R, Henderson H, Higa H, Ira-

- ni Rahaghi A, Jamet C, Jiang D L, Jordan T, Kangro K, Kravitz J A, Kristoffersen A S, Kudela R, Li L, Ligi M, Loisel H, Lohrenz S, Ma R H, Maciel D A, Malthus T J, Matsushita B, Matthews M, Minaudo C, Mishra D R, Mishra S, Moore T, Moses W J, Nguyen H, Novo E M L M, Novoa S, Odermatt D, O'Donnell D M, Olmanson L G, Ondrusek M, Oppelt N, Ouillon S, Pereira Filho W, Plattner S, Ruiz Verdú A, Salem S I, Schalles J F, Simis S G H, Siswanto E, Smith B, Somlai-Schweiger I, Soppa M A, Spyarakos E, Tessin E, van der Woerd H J, Vander Woude A J, Vandermeulen R A, Vantrepotte V, Wernand M R, Werther M, Young K and Yue L W. 2022. GLORIA - a global dataset of remote sensing reflectance and water quality from inland and coastal waters [dataset]. PANGAEA [DOI: 10.1594/PANGAEA.948492]
- Lehmann M K, Gurlin D, Pahlevan N, Alikas K, Conroy T, Anstee J, Balasubramanian S V, Barbosa C C F, Binding C, Bracher A, Bresciani M, Burtner A, Cao Z G, Dekker A G, Di Vittorio C, Drayson N, Errera R M, Fernandez V, Ficek D, Fichot C G, Gege P, Giardino C, Gitelson A A, Greb S R, Henderson H, Higa H, Rahaghi A I, Jamet C, Jiang D L, Jordan T, Kangro K, Kravitz J A, Kristoffersen A S, Kudela R, Li L, Ligi M, Loisel H, Lohrenz S, Ma R H, Maciel D A, Malthus T J, Matsushita B, Matthews M, Minaudo C, Mishra D R, Mishra S, Moore T, Moses W J, Nguyễn H, Novo E M L M, Novoa S, Odermatt D, O'Donnell D M, Olmanson L G, Ondrusek M, Oppelt N, Ouillon S, Pereira Filho W, Plattner S, Verdú A R, Salem S I, Schalles J F, Simis S G H, Siswanto E, Smith B, Somlai-Schweiger I, Soppa M A, Spyarakos E, Tessin E, van der Woerd H J, Vander Woude A, Vandermeulen R A, Vantrepotte V, Wernand M R, Werther M, Young K and Yue L W. 2023. GLORIA - a globally representative hyperspectral *in situ* dataset for optical sensing of water quality. Scientific Data, 10: 100 [DOI: 10.1038/s41597-023-01973-y]
- Li J J, Wu C X, Song R, Li Y S and Liu F. 2020. Adaptive weighted attention network with camera spectral sensitivity prior for spectral reconstruction from RGB images//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle: IEEE: 1894-1903 [DOI: 10.1109/CVPRW50498.2020.00239]
- Li Y, Li Y M, Guo Y L, Zhang Y L, Zhang Y B, Hu Y D and Xia Z. 2019. Reconstruction of water hyperspectral remote sensing reflectance based on sparse representation and its application. Environmental Science, 40(1): 200-210 (李渊, 李云梅, 郭宇龙, 张运林, 张毅博, 胡耀躲, 夏忠. 2019. 基于稀疏表达的水体遥感反射率高光谱重构及其应用. 环境科学, 40(1): 200-210) [DOI: 10.13227/j.hjlx.201804210]
- Li Z X, Sun D Y, Qiu Z F, Xi H Y, Wang S Q, Perrie W, Li Y M and Han B. 2018. A reconstruction method for hyperspectral remote sensing reflectance in the visible domain and applications. Journal of Geophysical Research: Oceans, 123(6): 4092-4109 [DOI: 10.1029/2017JC013734]
- Liu H Z, He X Q, Li Q Q, Kratzer S, Wang J J, Shi T Z, Hu Z W, Yang C, Hu S B, Zhou Q M and Wu G F. 2021. Estimating ultraviolet reflectance from visible bands in ocean colour remote sensing. Remote Sensing of Environment, 258: 112404 [DOI: 10.1016/j.rse.2021.112404]
- Liu Z M, Wang Y X, Vaidya S, Ruehle F, Halverson J, Soljačić M, Hou T Y and Tegmark M. 2024. KAN: Kolmogorov-Arnold networks. arXiv:2404.19756 [DOI: 10.48550/arXiv.2404.19756]
- Lucke R L, Corson M, McGlothlin N R, Butcher S D, Wood D L, Korwan D R, Li R R, Snyder W A, Davis C O and Chen D T. 2011. Hyperspectral imager for the coastal ocean: instrument description and first images. Applied Optics, 50(11): 1501-1516 [DOI: 10.1364/AO.50.001501]
- Matsuoka A, Babin M and Devred E C. 2016. A new algorithm for discriminating water sources from space: a case study for the southern Beaufort Sea using MODIS ocean color and SMOS salinity data. Remote Sensing of Environment, 184: 124-138 [DOI: 10.1016/j.rse.2016.05.006]
- Moore T S, Campbell J W and Feng H. 2015. Characterizing the uncertainties in spectral remote sensing reflectance for SeaWiFS and MODIS-aqua based on global *in situ* matchup data sets. Remote Sensing of Environment, 159: 14-27 [DOI: 10.1016/j.rse.2014.11.025]
- O'Reilly J E and Werdell P J. 2019. Chlorophyll algorithms for ocean color sensors - OC4, OC5 and OC6. Remote Sensing of Environment, 229: 32-47 [DOI: 10.1016/j.rse.2019.04.021]
- O'Shea R E, Pahlevan N, Smith B, Boss E, Gurlin D, Alikas K, Kangro K, Kudela R M and Vaičiūtė D. 2023. A hyperspectral inversion framework for estimating absorbing inherent optical properties and biogeochemical parameters in inland and coastal waters. Remote Sensing of Environment, 295: 113706 [DOI: 10.1016/j.rse.2023.113706]
- Oyama Y, Matsushita B, Fukushima T, Matsushige K and Imai A. 2009. Application of spectral decomposition algorithm for mapping water quality in a turbid lake (Lake Kasumigaura, Japan) from landsat TM data. ISPRS Journal of Photogrammetry and Remote Sensing, 64(1): 73-85 [DOI: 10.1016/j.isprsjprs.2008.04.005]
- Pahlevan N, Schott J R, Franz B A, Zibordi G, Markham B, Bailey S, Schaaf C B, Ondrusek M, Greb S and Strait C M. 2017a. Landsat 8 remote sensing reflectance (Rrs) products: evaluations, intercomparisons, and enhancements. Remote Sensing of Environment, 190: 289-301 [DOI: 10.1016/j.rse.2016.12.030]
- Pahlevan N, Smith B, Binding C, Gurlin D, Li L, Bresciani M and Giardino C. 2021. Hyperspectral retrievals of phytoplankton absorption and chlorophyll-a in inland and nearshore coastal waters. Remote Sensing of Environment, 253: 112200 [DOI: 10.1016/j.rse.2020.112200]
- Pahlevan N, Smith B, Binding C and O'Donnell D M. 2017b. Spectral band adjustments for remote sensing reflectance spectra in coastal/inland waters. Optics Express, 25(23): 28650-28667 [DOI: 10.

- 1364/OE.25.028650]
- Pahlevan N, Smith B, Schalles J, Binding C, Cao Z G, Ma R H, Alikas K, Kangro K, Gurlin D, Hà N, Matsushita B, Moses W, Greb S, Lehmann M K, Ondrusek M, Oppelt N and Stumpf R. 2020. Seamless retrievals of chlorophyll-a from sentinel-2 (MSI) and sentinel-3 (OLCI) in inland and coastal waters: a machine-learning approach. *Remote Sensing of Environment*, 240: 111604 [DOI: 10.1016/j.rse.2019.111604]
- Salem S I, Higa H, Ishizaka J, Pahlevan N and Oki K. 2023. Spectral band-shifting of multispectral remote-sensing reflectance products: insights for matchup and cross-mission consistency assessments. *Remote Sensing of Environment*, 299: 113846 [DOI: 10.1016/j.rse.2023.113846]
- Shi K, Zhang Y L, Zhu G W, Liu X H, Zhou Y Q, Xu H, Qin B Q, Liu G and Li Y M. 2015. Long-term remote monitoring of total suspended matter concentration in Lake Taihu using 250 m MODIS-aqua data. *Remote Sensing of Environment*, 164: 43-56 [DOI: 10.1016/j.rse.2015.02.029]
- Shi Z, Chen C, Xiong Z W, Liu D and Wu F. 2018. HSCNN+: advanced CNN-based hyperspectral recovery from RGB images// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Salt Lake City: IEEE: 1052-1060 [DOI: 10.1109/CVPRW.2018.00139]
- Sudduth K A, Jang G S, Lerch R N and Sadler E J. 2015. Long-term agroecosystem research in the central Mississippi river basin: hyperspectral remote sensing of reservoir water quality. *Journal of Environmental Quality*, 44(1): 71-83 [DOI: 10.2134/jeq2014.02.0060]
- Sun D Y, Hu C M, Qiu Z F and Wang S Q. 2015. Reconstruction of hyperspectral reflectance for optically complex turbid inland lakes: test of a new scheme and implications for inversion algorithms. *Optics Express*, 23(11): A718-A740 [DOI: 10.1364/OE.23.00A718]
- Valente A, Sathyendranath S, Brotas V, Groom S, Grant M, Taberner M, Antoine D, Arnone R, Balch W M, Barker K, Barlow R, Bélanger S, Berthon J F, Beşiktepe Ş, Borsheim Y, Bracher A, Brando V, Canuti E, Chavez F, Cianca A, Claustre H, Clementson L, Crout R, Frouin R, García-Soto C, Gibb S W, Gould R, Hooker S B, Kahru M, Kampel M, Klein H, Kratzer S, Kudela R, Ledesma J, Loisel H, Matrai P, McKee D, Mitchell B G, Moisan T, Muller-Karger F, O'Dowd L, Ondrusek M, Platt T, Poulton A J, Repecaud M, Schroeder T, Smyth T, Smythe-Wright D, Sosik H M, Twardowski M, Vellucci V, Voss K, Werdell J, Wernand M, Wright S and Zibordi G. 2019. A compilation of global bio-optical in situ data for ocean-colour satellite applications - version two. *Earth System Science Data*, 11(3): 1037-1068 [DOI: 10.5194/essd-11-1037-2019]
- Wang Y C, Lee Z, Wei J W, Shang S L, Wang M H and Lai W D. 2021. Extending satellite ocean color remote sensing to the near-blue ultraviolet bands. *Remote Sensing of Environment*, 253: 112228 [DOI: 10.1016/j.rse.2020.112228]
- Wei J W, Yu X L, Lee Z, Wang M H and Jiang L D. 2020. Improving low-quality satellite remote sensing reflectance at blue bands over coastal and inland waters. *Remote Sensing of Environment*, 250: 112029 [DOI: 10.1016/j.rse.2020.112029]
- Wu D and Sun D W. 2013. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review — Part II: applications. *Innovative Food Science and Emerging Technologies*, 19: 15-28 [DOI: 10.1016/j.ifset.2013.04.016]
- Yang H B, Kong J L, Hu H H, Du Y, Gao M Y and Chen F. 2022. A review of remote sensing for water quality retrieval: progress and challenges. *Remote Sensing*, 14(8): 1770 [DOI: 10.3390/rs14081770]
- Yang W, Matsushita B, Chen J and Fukushima T. 2011. Estimating constituent concentrations in case II waters from MERIS satellite data by semi-analytical model optimizing and look-up tables. *Remote Sensing of Environment*, 115(5): 1247-1259 [DOI: 10.1016/j.rse.2011.01.007]
- Yang W, Matsushita B, Chen J, Yoshimura K and Fukushima T. 2013. Retrieval of inherent optical properties for turbid inland waters from remote-sensing reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6): 3761-3773 [DOI: 10.1109/TGRS.2012.2220147]
- Yang Z Q, Zhang J S, Luo X L, Lu Z and Shen L L. 2024. Activation space selectable Kolmogorov-Arnold networks. *arXiv*: 2408.08338 [DOI: 10.48550/arXiv.2408.08338]
- Zhao G P, He Y D, Wang Z W and Wu H. 2024. Hybrid transformer architecture for spectral super-resolution reconstruction of multispectral images//IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium. Athens: IEEE: 9468-9471 [DOI: 10.1109/IGARSS53475.2024.10642439]
- Zhu W N, Yu Q, Tian Y Q, Becker B L, Zheng T and Carrick H J. 2014. An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments. *Remote Sensing of Environment*, 140: 766-778 [DOI: 10.1016/j.rse.2013.10.015]
- Zhu X X, Tuia D, Mou L C, Xia G S, Zhang L P, Xu F and Fraundorfer F. 2017. Deep learning in remote sensing: a comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4): 8-36 [DOI: 10.1109/MGRS.2017.2762307]
- Zibordi G, Berthon J F, Mélin F, D'Alimonte D and Kaitala S. 2009. Validation of satellite ocean color primary products at optically complex coastal sites: northern Adriatic Sea, Northern Baltic proper and gulf of Finland. *Remote Sensing of Environment*, 113(12): 2574-2591 [DOI: 10.1016/j.rse.2009.07.013]

KAN-based hyperspectral remote sensing reflectance reconstruction in coastal waters

HUANG Tingting¹, FAN Donglin^{1,2}, HE Hongchang¹, FU Bolin^{1,2}, TAN Haoyuan³

1. School of Geomatics and Geoinformation, Guilin University of Technology, Guilin 541004, China;

2. Key Laboratory of Spatiotemporal Big Data Perception and Service of Ecology, Guilin University of Technology, Guilin 541006, China;

3. The Fourth Institute of Oceanography, Ministry of Natural Resources, Beihai 536015, China

Abstract: To address the challenges in acquiring hyperspectral remote sensing reflectance (R_{rs}) data and the limitations of existing spectral reconstruction methods, including their reliance on in-situ measurements, weak generalization ability, and insufficient accuracy in optically complex coastal waters, this study proposes a novel deep learning model based on the Kolmogorov-Arnold Network (KAN). The model is designed to directly exploit widely available multispectral satellite observations to efficiently reconstruct continuous hyperspectral R_{rs} that closely match the characteristics of true measurements, thereby overcoming the bottlenecks of traditional approaches and improving remote sensing inversion performance in complex nearshore environments.

The proposed method adopts an end-to-end nonlinear modeling framework within the KAN architecture, incorporating learnable nonlinear activation functions to flexibly capture complex local nonlinear relationships in the input data. This approach enables accurate reconstruction of continuous hyperspectral R_{rs} from multispectral inputs, with spectral distributions highly consistent with actual observations. In this work, Level-2 R_{rs} products from the Hyperspectral Imager for the Coastal Ocean were used as training samples. These products were resampled in accordance with the spectral response functions of six mainstream multispectral sensors (S3A OLCI, MERIS, MODIS, SeaWiFS, S2A MSI, and OLI) to generate “multispectral-hyperspectral” data pairs for model training. The approach eliminates the need for in-situ or optical simulation data, relying solely on remote sensing observations for training and modeling, thus greatly enhancing model generality and practicality.

Experimental results demonstrate that the KAN model achieved superior reconstruction accuracy (mean $R^2 > 0.9982$) and robustness across all six sensors compared with benchmark models. It not only reproduced the overall spectral shape with high fidelity but also exhibited strong detail-capturing capability in critical regions with missing or sparse sensor bands, such as the red-edge region beyond 680 nm, producing reconstructed curves nearly identical to the original spectra. In downstream chlorophyll-a inversion applications, the use of KAN-reconstructed data significantly improved retrieval accuracy over original multispectral inputs, reducing RMSE by approximately 16.13% and increasing R^2 by 3.30%. The advantages were particularly evident in high-concentration waters, effectively overcoming the limitations of multispectral sensors.

Overall, the proposed KAN-based hyperspectral R_{rs} reconstruction model successfully breaks the dependency bottleneck of traditional methods on in-situ or simulated data, offering a highly accurate and generalizable universal solution. By generating high-quality continuous spectra from existing multispectral datasets, it substantially enhances the performance and accuracy of water quality parameter retrieval in complex aquatic environments. The model serves as a powerful data augmentation tool for aquatic remote sensing and a novel technical pathway for leveraging vast archives of historical multispectral satellite data to enable reliable global water environment monitoring.

Key words: KAN network, hyperspectral remote sensing reflectance, remote sensing reflectance reconstruction, water quality parameter retrieval, coastal water bodies, water remote sensing reflectance

Supported by Guangxi Natural Science Foundation (No. 2022GXNSFBA035637); Guangxi Science and Technology Major Project (No. AA23062035-2)