J. Lake Sci.(*湖泊科学*), 2026, 38(3):000-000
DOI 10.18307/2026.0342

基于随机森林混合模型的鄱阳湖局部地形遥感反演*

伍 婵[1]，赖格英[1,2**]，黄慧[1]，申 娜[1]
（1:江西师范大学地理与环境学院，南昌 330022）
（2:鄱阳湖湿地与流域研究教育部重点实验室，南昌 330022）

摘 要：湖泊地形作为地理环境的核心要素，在地理研究中具有多维度的重要意义，尤其在地表过程中对水文及水动力过程模拟研究具有关键性的影响。传统大型湖泊地形数据获取手段投入大、耗时长、更新周期慢，因此有必要研究一种基于卫星遥感数据快速获取湖泊地形的方法。本文基于陆地卫星（Landsat）遥感影像与实测高程数据，采用随机森林算法（RF）反演鄱阳湖枯水期局部地形。针对地形特征的空间非平稳性及预测残差的空间自相关性，融合地理加权回归(GWR)与普通克里金法(OK)对反演结果进行优化，并对其误差进行分析。结果表明：(1) 相比 RF，地理加权回归随机森林克里金混合模型（GWR-RF-OK）精度明显提高，两个研究区实测高程与反演高程的决定系数（R^2）均上升，平均绝对误差（MAE）和平均相对误差（MRE）均下降。(2) 混合模型在土地覆盖类型单一的裸滩区和土地覆盖类型相对复杂的鄱阳湖南矶湿地国家级自然保护区（以下简称南矶湿地区）都有较好的反演效果，R^2 分别为 0.71、0.56，MAE 分别为 0.34m、0.35m，MRE 分别为 5.26%、3.06%。经分段分析，模型在地形高程大于 10m 的区域反演效果更佳。(3) 地形起伏程度和土地覆盖类型会影响反演精度，地形平缓区域误差显著小于陡峭区域，同类地表覆被在土地覆盖类型单一区域的地形反演精度明显优于多种土地覆盖类型混合区域。
关键词：地形反演；遥感；地理加权回归；随机森林模型；克里金法；鄱阳湖

Remote sensing inversion of local topography in Lake Poyang by using a hybrid random forest model *

Wu Chan[1], Lai Geying[1,2**], Huang Hui[1]& Shen Na[1]
（1:*School of Geography and Environment*, *Jiangxi Normal University*, *Nanchang 330022*, *P.R.China*）
（2:*Key Laboratory of Poyang Lake Wetland and Watershed Research*, *Nanchang 330022*, *P.R.China*）

Abstract: Lake topography is a fundamental component of the geographical environment, with broad significance in geographical research and foundational influence on surface processes—particularly in hydrological and hydrodynamic modeling. However, conventional methods for acquiring bathymetric data in large lakes are often costly, time-consuming, and updated infrequently, underscoring the need for efficient satellite-based approaches to map lake topography. This study developed a hybrid modeling framework to invert dry-season topography in Lake Poyang by integrating Landsat imagery and measured elevation data using a Random Forest (RF) algorithm. To address spatial non-stationarity in topographic features and autocorrelation in prediction residuals, we further combined Geographically Weighted Regression (GWR) with Ordinary Kriging (OK) to optimize inversion accuracy and analyze associated errors. The results indicate that: (1) the hybrid GWR-RF-OK model significantly outperformed the standalone RF model, with higher coefficients of determination (R^2) and lower mean absolute error (MAE) and mean relative error (MRE) between measured and predicted elevations in two study areas; (2)

* 2025-08-03 收稿；2025-09-09 收修改稿。
国家自然科学基金重点项目（42330108）和国家自然科学基金项目（42161014）联合资助。
**通讯作者；E-mail：Laigeying@126.com

the hybrid model performed well in both a bare beach area with homogeneous land cover and the more complex surface cover of the Nanji Wetland National Nature Reserve, achieving R^2 values of 0.71 and 0.56, MAE of 0.34 m and 0.35 m, and MRE of 5.26% and 3.06%, respectively. Segmented analysis revealed better performance in areas where elevation exceeds 10 m; (3) inversion accuracy was influenced by topographic roughness and land cover heterogeneity, with smaller errors occurring in gently sloping areas and those with uniform land cover compared to steep or mixed-cover terrain. These findings support the utility of remote sensing-based hybrid models for efficient and accurate lake topographic mapping.
Keywords: Topographic inversion；Remote sensing；Geographically Weighted Regression；Random Forest model；Kriging；Lake Poyang

湖泊地形作为自然地理环境的重要组成部分，在现代地理研究中具有重要的基础及实践意义。其形态特征不仅塑造了湖泊生态系统的空间格局，更为水文循环和水动力过程提供了基础载体[1]。从科学研究到工程实践，湖泊地形数据的精确获取与分析已成为理解水域环境演变、预测水资源动态的核心切入点。然而，获取湖泊地形的传统方法，如人工勘测、船载声呐探测等，虽精度高，但难度大，成本高，更新周期长[2]。卫星遥感技术具有长时序、大范围、高效率的监测特点，是识别地物信息的有效手段，在浅水湖泊和潮滩的地形获取方面具有很大优势[3]。

鄱阳湖作为中国最大的淡水湖泊和长江流域最大的通江湖泊，其地形变化不仅影响区域生物多样性与生态稳定性，更关乎全国生态环境变化[4]。受五河、长江来水来沙以及采砂、围垦和水库修建等人类活动影响，鄱阳湖地形发生了不同程度的变化，湖泊面积和库容也随之改变，影响其调节径流和调蓄洪水的能力[5]。泥沙淤积、流水冲刷以及人为干扰使得鄱阳湖地形复杂，勘测面临较大的挑战。政府机构于 1980、1998 和 2010 年对全湖地形进行勘测[6]，于 2020 年对典型断面进行勘测[7]，实测数据更新周期长达十余年。地形数据的滞后，不仅影响水文与水动力模型的精度，而且制约相关研究的开展。有学者利用 ICESat-2 测高数据重建鄱阳湖湖床地形，但该卫星于 2018 年发射，此前数据缺失，而且存在覆盖间隙大、处理过程复杂等问题[8]。因此，探索一种基于遥感技术的新方法对鄱阳湖地形进行估算，作为实测数据更新周期之间的补充十分必要。

大型吞吐性湖泊水域辽阔，除通江水域外，季节性淹没洲滩分布着多样植被和泥滩[9]，全湖尺度的地形研究难以精准刻画不同地表覆盖区域的地形分异特征，因此需开展分区研究以揭示内部差异。目前水下地形的研究多利用光谱特征与水深之间的关系，通过反演水深来获取水下地形，但在浑浊的水域效果不佳[10]。湖滩、潮滩的研究则主要利用水边线法[11, 12]和淹水频率法[13, 14]获取地形，当前针对鄱阳湖的地形研究也多基于此开展[15, 16]，但这两种方法需要获取不同时期的多幅遥感影像，提取的地形等高线可能不一致[17]，而且局部区域提取的边界存在重合，受到不同日期水位差异的影响，反演后的湖区边界会出现高低水位交叉现象，产生较大误差[4]。有学者基于单幅遥感影像，建立地形高程与光谱反射率之间的线性或非线性的关系模型反演地形。例如，陶旭等[18] 和李静等[19] 基于滩地含水量随时间推移与高程呈负相关的关系，建立高程与光谱反射率之间的模型反演了江苏地区的潮滩地形。另有研究基于地表覆盖类型随高程变化，且其光谱特征各异的原理，利用光学遥感影像反演地形，Panagiotou E 等[20] 据此生成了希腊大部分地区的 DEM，Bergsma EW 等 [21]据此估算了美国北卡罗莱纳州达克地区浅海水深和沿岸地形。这些研究表明遥感影像光谱和地形高程之间存在着复杂的关系。但水下地形易受泥沙含量、水质影响，滩地地形则受地表土地覆盖类型等多种因素制约，地形反演的精度提升遭遇瓶颈。

随着计算机技术快速发展和大数据时代来临，具备强大数据处理与预测能力的机器学习算法在地形反演领域得到了广泛的关注[22-24]。其中，随机森林算法凭借其能够精准捕捉复杂变量间的非线性关系、较强的抗过拟合能力、较高的运算效率等特点，在地形反演方面具有极大优势[25, 26]，被广泛应用于地形研究[27, 28]。然而，利用机器学习模型进行大范围空间上的预测时，往往难以厘清观测目标之间可能存在的空间相关性，将地理坐标、空间权值等作为特征因子结合机器学习建模是目前有效的解决方法[29, 30]。已有研究将空间分析模型和机器学习模型结合，在水深[31]、生物量[32]、降水[33]，土壤有机碳[34]等领域的估测中取得了优于单一模型的效果。但以往湖泊、浅海的地形反演多着眼于地形与遥感波段之间的相关性而鲜少关注特征因子间

的空间关系，将机器学习模型和空间模型相结合以估测湖泊地形的研究很少见。鉴于此，本文基于光学遥感影像和实测高程数据，结合地理加权回归、随机森林算法与普通克里金法，分析鄱阳湖内部两个不同区域的高程与遥感影像波段关系，开展局部地形研究。旨在构建地形高程反演模型，以期补充实测地形更新周期之间的数据空白，为鄱阳湖水文水动力数值模拟提供局部地形数据参考，服务于相关研究。

1 研究区概况与数据来源

1.1 研究区

鄱阳湖（115° 49′ E~116° 46′ E，28° 24′ N~29° 46′ N）位于长江中下游南岸，江西省北部，湖盆自南向北，由东到西倾斜，湖底较平坦，江西省内修、赣、抚、信、饶五河来水经鄱阳湖调蓄后由湖口注入长江，是典型的大型吞吐性浅水湖泊，具有“丰水为湖，枯水是河”的独特自然地理特征[35]。考虑到鄱阳湖范围较大，不同土地覆盖类型对地形反演的影响可能不同，本文选取了鄱阳湖枯水期同一日的裸滩区和土地覆盖类型相对复杂的鄱阳湖南矶湿地国家级自然保护区（以下简称南矶湿地区）作为研究区，如图 1 所示（主图波段组合为红、绿、蓝波段，子图波段组合为短波红外、近红外和蓝波段）。

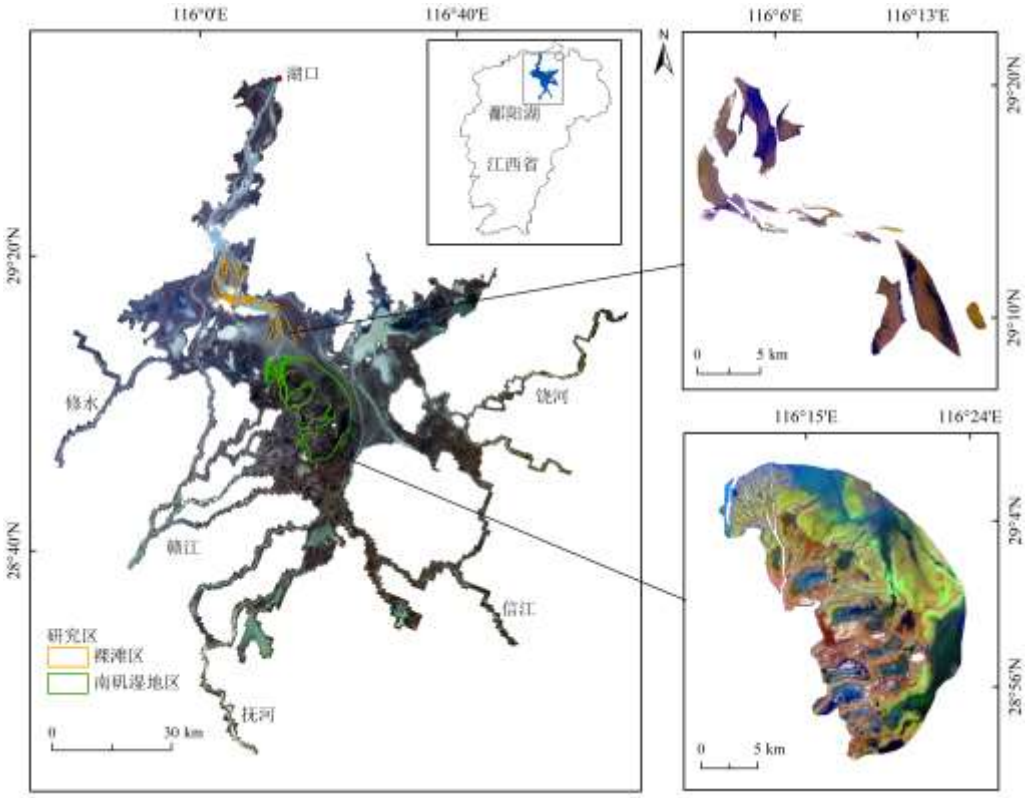

图 1 研究区

Fig.1 Study area

1.2 数据收集与处理

本文使用的数据主要包括光学遥感影像和实测地形数据，如表 1 所示。由于现有实测数据年份（2010 年）的限制，使用的遥感影像数据是从美国地质调查局（United States Geological Survey，USGS) 官网（http://glovis.usgs.gov/）获取的鄱阳湖 2010 年 1 月 14 日 Landsat 5 TM Collection2 Level2 影像数据，遥感影像云量<5%，空间分辨率为 30m。2010 年实测地形高程数据比例尺为 1：10000，来源于江西省水文局。所有的空间数据均使用 WGS84 坐标系和 UTM 投影(50 N 区)。

获取的 Level 2 级遥感影像已完成几何校正、辐射定标和大气校正，在 ENVI5.6 中使用 Band Math 工具

计算归一化差异水体指数（Normalized Difference Water Index，NDWI）[36]，公式（1），提取水体(NDWI>0)，获得水陆划分结果，提取出当日的通江水域并剔除。然后利用南矶湿地矢量数据和裸滩矢量数据，裁剪出两个研究区。

$$NDWI = (Green - NIR)/(Green + NIR) \tag{1}$$

式中：[illegible]、*NIR* 分别指绿色波段、近红外波段的反射率。

本文使用的是 Lai 等[37]在鄱阳湖二维水动力模型中格网化的地形高程数据，其与遥感影像匹配的具体步骤为：首先在 ENVI5.6 中提取出两个研究区遥感影像的各波段反射率，然后在 ArcGIS10.8 中提取出研究区中鄱阳湖二维水动力模型的每一个格网的质点坐标及其相应高程，再将遥感影像各波段反射率提取到每一个点，得到每一个点对应的高程和光谱值。

表 1 研究数据

Tab.1 Research data

数据	时间	精度	来源
LT05_L2SP_121040_20100114_20200825_02_T1	2010/1/14	30m	美国地质调查局
实测地形数据	2010	1：10000	江西省水文局

2 方法

2.1 特征选择

Landsat5 TM 遥感影像共 7 个波段，分别是：蓝色波段 B1（0.45-0.52μm）,绿色波段 B2（0.52-0.60μm），红色波段B3（0.60-0.69μm）,近红外波段B4（0.76-0.9μm）,短波红外波段B5(1.55-1.75μm),热红外波段B6(10.40-12.5μm)，短波红外波段 B7(2.08-2.35μm)。本文选择 Landsat5 TM 的 6 个波段，即 B1-B5 和 B7 波段作为模型的特征变量。

2.2 模型构建

2.2.1 随机森林算法 随机森林算法(Random Forest，RF)是由 Breiman[38]于 2001 年提出的以决策树为学习机器的集成学习算法，其主要思想是通过构造大量独立的决策树，然后对每一棵树的输出结果进行集成，具有较强的泛化能力和数据处理能力。随机森林的参数设置尤为重要，因此采用网格搜索自动优化法对模型进行参数调整，以获得最优的模型参数，并且采用 Hold-out 验证法对模型进行训练和测试。

2.2.2 地理加权回归法 地理加权回归（Geographically Weighted Regression, GWR）是 Brunsdon 等[39]提出的一种考虑空间非平稳性的回归方法，公式（2）。它通过空间距离和空间核函数来衡量地理位置的临近性，可以反映每个样本的地形高程与光谱特征之间的相关性，并将其表示为空间权值，权重值通过高斯核函数计算得出。将两个研究区的波段信息分别输入 GWR 模型构建空间矩阵，使用 K 最近邻算法（KNN）为每个波段计算空间权值，作为 RF 模型新增的自变量。较小的 K 值可以更好地捕捉局部特征，因此将 K 设置为 5。

$$y_i = \beta_0(\mu_i, v_i) + \sum_i^k \beta_k(\mu_i, v_i)\chi_{ik} + \varepsilon_i \tag{2}$$

其中，(μ_i, v_i)是样本的位置，ε_i是残差项，假设它服从$\varepsilon_i \sim N(0, \sigma^2)$随机分布。此外，$\beta_0$是常数项，$\beta_k$是位置$(\mu_i, v_i)$处的回归系数。

2.2.3 普通克里金插值法 克里金插值法(Kriging Interpolation)是由 OLIVER 等[40]提出的一种利用已知点的权重值来对未知点进行预测的插值方法。本文使用的是普通克里金法(Ordinary Kriging，OK)，它假设空间数据属性值具有平稳性和自相关性，属性值在每个空间位置上的数学期望都是同一个常数，计算公式如（公式 3）。采用径向量函数（RBF）结合常数核的复合核函数，利用 GWR-RF 模型训练样本中预测值和实测值之间的残差进行训练建模，然后利用训练好的模型对测试集数据进行空间插值预测，将预测的残差加到 GWR-RF 模型的初始预测值上，得到优化后的预测值。

$$\hat{Z}(\chi_0)=\sum_{i=1}^{n}\lambda_i Z(\chi_i) \tag{3}$$

式中：$\hat{Z}$为点的预测值;λ_i为i位置测量值的权重;$Z(\chi_i)$为i位置的测量值;n为数据点个数。

2.2.4 GWR-RF-OK 模型 混合模型集成了三种算法的优势，它分别利用了 GWR 的空间非平稳性建模能力、RF 的非线性拟合能力和抗过拟合能力，以及 OK 对空间自相关残差的优化能力，以克服机器学习模型在处理复杂空间问题时的局限性。传统全局回归方法假定变量关系在空间上恒定，这与现实情况往往不符，为解决该问题， GWR 引入了空间权重矩阵，基于距离衰减原则为每个位置生成一套局部回归系数，从而将原始遥感波段转化为能够反映其局部空间重要性的新特征。RF 则凭借其强大的非线性拟合能力，处理地形高程与遥感波段间复杂的映射关系，它将 GWR 生成的空间特征与原始波段一同输入，构建更精确的地形预测模型，GWR 为 RF 提供“空间语境”，而 RF 负责解析该语境与地形之间的复杂关系。尽管 GWR-RF 模型性能优异，但其预测残差常具有空间自相关性，OK 作为一种地统计方法，可通过变异函数量化并插值残差的空间结构，捕捉 RF 未解释的细微空间变异，从而进一步提升地形反演的精度。利用 GWR-RF-OK 算法进行地形反演的过程如图 2 所示，模型构建的所有操作均在 python3.11.6 中完成。

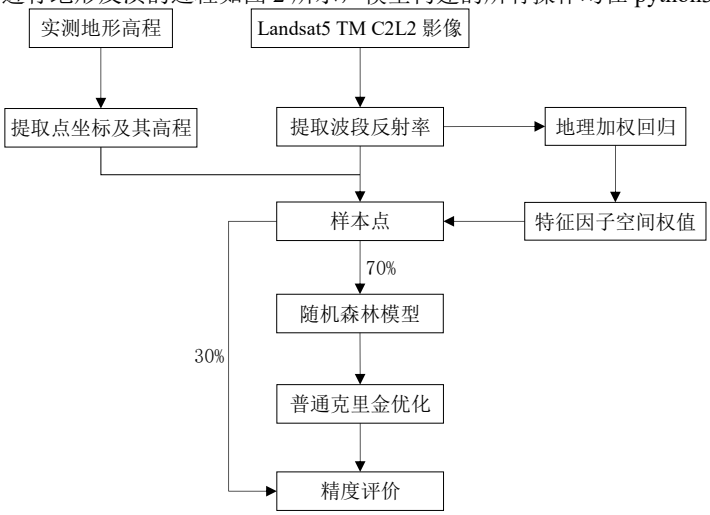

图 2 技术路线图
Fig.2 Technical Roadmap

2.3 评价指标

采用决定系数（R^2），平均绝对误差（MAE）和平均相对误差 (MRE)作为评价指标，见公式 4。

$$\begin{cases} R^2=1-\dfrac{\sum_{i=1}^{n}(\hat{y}_i-y_i)^2}{\sum_{i=1}^{n}(\overline{y}-y_i)^2} \\ MAE=\dfrac{1}{n}\sum_{i=1}^{n}|\hat{y}_i-y_i| \\ MRE=\dfrac{1}{n}\sum_{i=1}^{n}\dfrac{|\hat{y}_i-y_i|}{|y_i|}\times 100\% \end{cases} \tag{4}$$

式中，$\hat{y}_i$为反演模型的预测值，y_i为实测值，$\overline{y}$ 为实测值的均值，n为样本数量。

3 结果与分析

3.1 地形高程与光谱反射率的关系

使用皮尔逊相关系数法分别计算了两个研究区地形高程与遥感影像波段之间的相关性，如图 3 所示。

对于裸滩区，B4 是最优特征因子，与地形高程相关性为 0.5891，这是因为近红外波段对土壤含水量和粘土矿物含量的变化比较敏感，与前人研究相符[7]。对于南矶湿地区，B1 贡献率最大，相关性为 0.3591，这可能是由于该区域水域面积较大，B1 波段对水体有较强的穿透能力，可以获得更多水下信息。但所有特征的重要性都相对偏低，这可能是由于该区域土地覆盖类型相对复杂所致[41]。由于选取的特征因子较少，所以全部用于模型构建。

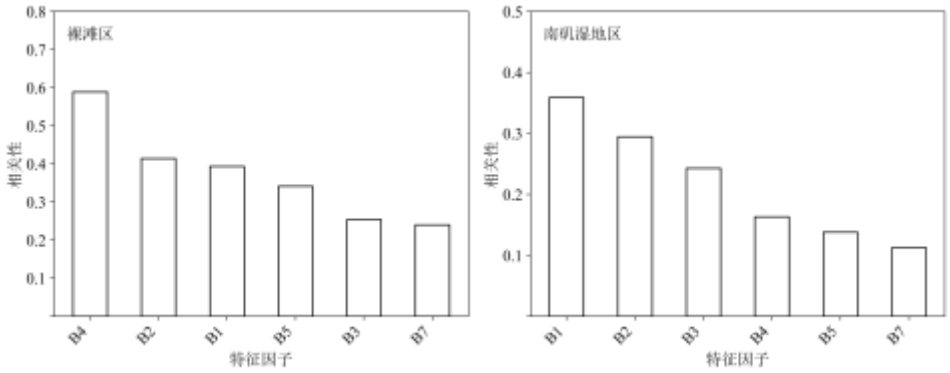

图 3 特征因子相关性

Fig.3 Correlations of characteristic factors

3.2 模型性能比较

在裸滩区随机选取 1886 个样本点，南矶湿地区随机选取 2316 个样本点，其中 70%作为训练数据，30%作为验证数据。模型产生了 162 种可能的参数组合，总共训练了 810 次，以保证模型的稳定性和准确性。模型训练得到的两个研究区最优参数如表 2 所示。

表 2 最优模型参数

Tab.2 Optimal model parameters

参数 研究区	sample	max_depth	Max_ features	min_samples _leaf	min_samples _split	n_estimators
裸滩区	1886	10	Sqrt	2	5	500
南矶湿地区	2316	10	Sqrt	3	5	100

为了验证所构建模型的有效性，本文采用 R^2、MAE、MRE 以及反演高程与实测高程的范围对比，比较了 RF 和 GWR-RF-OK 模型在两个研究区的性能，见表 3。

表 3 不同模型在两个研究区的反演精度

Tab.3 Inversion accuracy of different models across two study areas

反演效果	模型	R^2	MAE（m）	MRE（%）	实测高程范围（m）	反演高程范围（m）
裸滩区	RF	0.34	0.61	9.89	0.73~11.6	4.26~10.99
	GWR-RF-OK	0.71	0.34	5.26		2.07~11.56
南矶湿地区	RF	0.32	0.48	4.14	5~15.52	10.46~12.98
	GWR-RF-OK	0.56	0.35	3.06		6.19~14.18

对于裸滩区，RF 模型的 R^2 为 0.34，MAE 为 0.61m，MRE 为 9.89%。GWR-RF-OK 的 R^2 为 0.71，MAE 为 0.34m，MRE 为 5.26%。相比于 RF 模型，GWR-RF-OK 模型的 MAE 下降了 29.17%，MRE 下降了 47.24%。裸滩区的实测高程范围为 0.73~11.6m，RF 模型、GWR-RF-OK 模型的预测高程范围分别为 4.26~10.99m、2.07~11.56m，GWR-RF-OK 模型预测的高程范围更接近实测值。对于南矶湿地区，RF 模型的 R^2 为 0.32，MAE 为 0.48m，MRE 为 4.14%。GWR-RF-OK 的 R^2 为 0.56，MAE 为 0.35m，MRE 为 3.06%。相比于 RF 模型，

GWR-RF-OK 模型的 MAE 下降了 27.08%，MRE 下降了 26.09%。南矶湿地区的实测高程范围为 5~15.52m，RF 模型、GWR-RF-OK 模型的预测高程范围分别为 10.46~12.98m、6.19~14.18m，显然 GWR-RF-OK 模型预测效果更佳。相较于单一模型， GWR-RF-OK 模型在裸滩区和南矶湿地区的反演 R 均上升，误差均下降。虽然两个模型预测的高程范围均小于实测高程范围，在最大值和最小值处的预测结果都有偏差，但是混合模型在最值处的预测效果更接近实测值，表明 GWR-RF-OK 模型在地形高程反演方面具有更高的精度。

3.3 鄱阳湖地形反演评价

3.3.1 地形高程反演结果分析 裸滩区 R^2 为 0.71，MAE 为 0.34m，MRE 为 5.26%，南矶湿地区 R^2 为 0.56，MAE 为 0.35m，MRE 为 3.06%。由表 3 可知，虽然混合模型反演结果在高程最值，尤其是最小值处存在偏差，但整体而言精度较高。将 GWR-RF-OK 模型应用于两个研究区，反演结果见图 4。裸滩区地形高程呈现由湖滩向入江水道逐渐降低的变化趋势，其中地形最高处在区域南部，最低处位于北部入江水道沿岸。南矶湿地区地形最高处位于中西部靠近湖泊边缘的一侧，呈现由西南向东北降低的整体趋势，但其西部的入湖河道沿岸和西南部蝶形湖区域的高程明显低于四周。两幅地形高程反演图能够较好地反映两个研究区的高程变化趋势，基本符合当地地形分布特征。

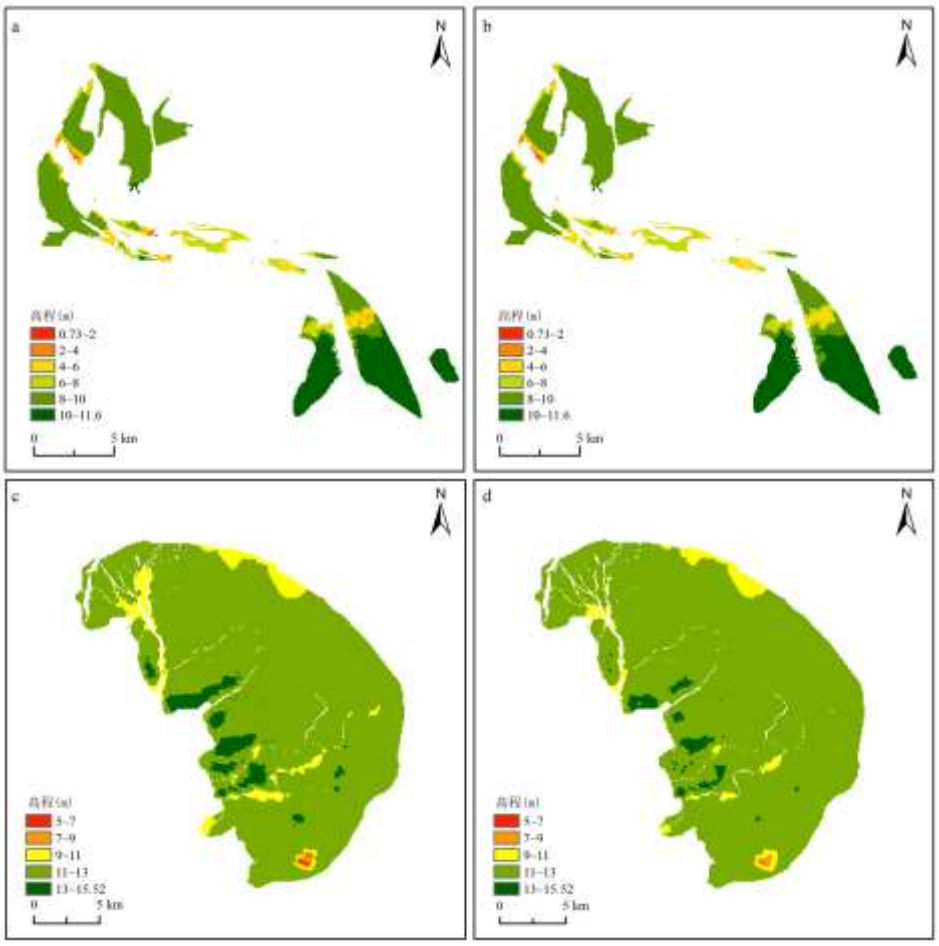

图 4 实测地形（左）与反演地形（右）：裸滩区（a~b）和南矶湿地区（c~d）

Fig.4 Measured topography (left) vs. inverted topography (right): (a~b) Bare Flat Area; (c~d) Nanji Wetlands

3.3.2 地形高程分段精度评价 为了分析 GWR-RF-OK 模型在鄱阳湖不同高程范围的反演效果，以 5m 为间隔，分段统计两个研究区的验证集在不同高程范围内的平均相对误差和平均绝对误差，对不同高程段的反演结果进行评价。

表 4 不同研究区地形高程分段误差

Tab.4 Segmented terrain elevation errors across two study areas

研究区 误差	裸滩区			南矶湿地区	
高程（m）	<5	5~10	>10	5~10	>10
样本点（个）	15	448	103	37	658
MRE（%）	63.03	4.19	1.52	20.41	2.65
MAE（m）	1.44	0.32	0.16	1.84	0.32

从表 4 可知，裸滩区，5m 高程以下的区域反演效果最差，MRE 和 MAE 分别为 63.03%和 1.44m，误差在 5m 以上区域大幅减小，在 5~10m 高程范围内，MRE 和 MAE 为 4.14%和 0.32，相比于高程小于 5m 的区域，分别下降了 93.35%和 77.78%。误差在 10m 以上的区域达到最小，MRE 和 MAE 分别为 1.52%和 0.16m。南矶湿地区，模型在 10m 高程以上区域反演精度大幅提升，MRE 和 MAE 在高程 10m 以下的区域分别为 20.41%和 1.84，误差过大的原因可能和裸滩区高程 5m 以下区域一致，与样本点过少有关。在 10m 以上高程范围降至 2.65%和 0.32m，分别下降了 87%和 83%。整体而言，GWR-RF-OK 模型在两个研究区的 MRE 和 MAE 都是随高程的增大而减小，就高程大于 5m 的范围而言，裸滩区的反演效果显著优于南矶湿地区，而且两个研究区在高程 10m 以上范围的 MRE 和 MAE 都更小，反演效果更优，而鄱阳湖滩地高程多在 12~17m 之间[42]，这进一步证明了本文所提方法在鄱阳湖局部地形估测方面的可行性。

3.3.3 不同土地覆盖类型的地形高程反演评价 南矶湿地区的土地覆盖类型较为复杂，但主要是水体、泥滩和植被三类。为了分析模型在不同土地覆盖类型的反演效果，通过目视解译，基于 GEE 平台，使用 2010 年 1 月 14 日 Landsat 5 TM Collection2 Level2 影像数据完成南矶湿地区地表覆盖的监督分类（水体、泥滩和植被）。其中 70%的样本作为训练集，30%作为测试集， 分类的 Kappa 精度为 0.85，总体精度为 0.90，分类结果具有较高的精度。将其与当日的遥感影像（波段组合为 7、4、1）进行对比，如图 5 所示，分类结果与当日实际土地覆盖类型具有较高的一致性，可以作为南矶湿地区不同土地覆盖类型高程反演评价的支撑数据。

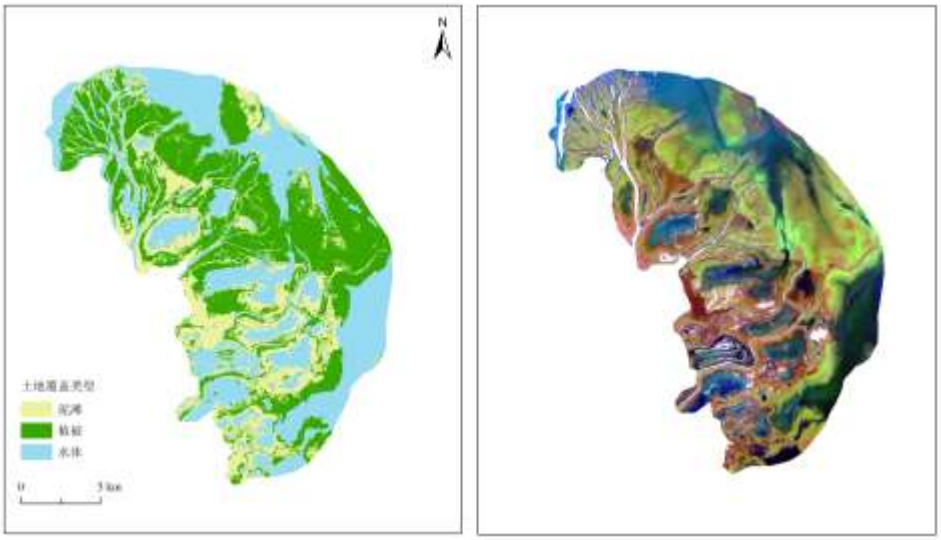

图 5 南矶湿地区土地覆盖类型分类与遥感影像对比图

Fig.5 Comparison of land cover classification and remote sensing imagery in Nanji Wetlands

表 5 不同土地覆盖类型地形高程反演精度对比

Tab.5 Comparison of topographic elevation inversion accuracy across different land cover types

土地覆盖类型	样本点（个）	MAE（m）	MRE	高程范围（m）
水体	700	0.31	2.85%	6.5~14
植被	700	0.41	3.53%	6~15.24
泥滩	700	0.61	5.45%	6~15.52

将地形高程点数据与各类地表覆盖矢量数据在 ArcGIS 中进行相交处理，得到各土地覆盖类型中的地形高程样本点。为了避免因样本量的差异而影响精度，分别在不同土地覆盖类型中都随机选取 700 个样本点，计算了不同土地覆盖类型的 MAE 和 MRE。如表 5 所示，反演地形高程与实测地形高程具有较高的一致性，水体、泥滩和植被区域的 MAE 分别为 0.31m、0.61m 和 0.41m，MRE 分别为 2.85%、5.45%和 3.53%。所构建的模型在水体、泥滩、植被区域反演地形高程均表现出较高的精度，这表明所提方法既能有效应用于包含多种不同土地覆盖类型的区域，也能有效应用于同一区域内不同的土地覆盖类型。与水体相比，植被覆盖区的地形反演精度略低，这可能是由于反演得到高程是植被表面高度，而非地面真实高程，但此区域多为低矮的草本植被，对反演精度影响较小。泥滩区的误差较大，这可能是由于其介于水体和植被的过度地带，空间分布离散且面积较小，导致模型反演精度较低。

3.3.4 *地形高程反演误差分析* 为了更好地分析 GWR-RF-OK 模型反演鄱阳湖局部地形的效果，利用实测地形高程的验证点，分别对裸滩区和南矶湿地区的地形反演的残差进行分析。由图 6 可知，两个研究区的残差分布都比较集中，主要分布在±1m 之间，在此范围内，裸滩区的残差占比为 91.33%，南矶湿地区占总数的 91.8%，其中裸滩区和南矶湿地区分别有约 84%、76%的离散点的残差介于-0.5~0.5m 之间。表明，GWR-RF-OK 模型在预测鄱阳湖局部地形高程时，裸滩区的反演效果优于南矶湿地区，虽两个研究区中有小部分区域反演效果不佳，但绝大部分区域的误差均在可接受范围内。

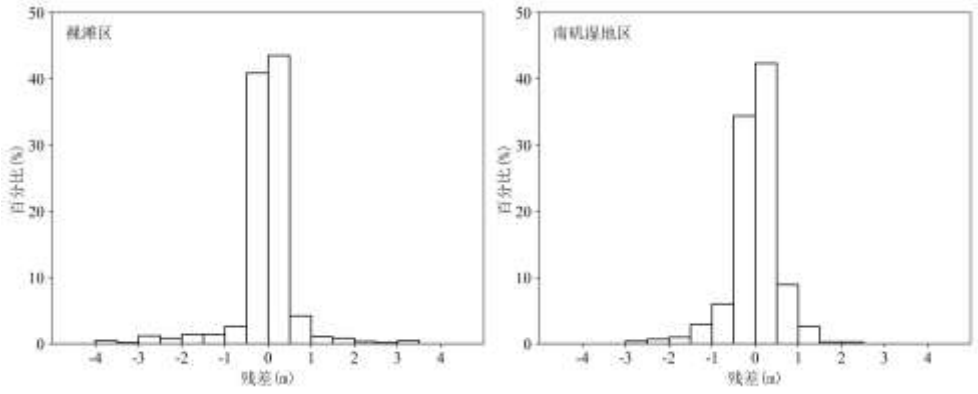

图 6 残差柱状图

Fig.6 Histogram of residuals

裸滩区有 4.1%的区域误差超过 2m，其中 1.4%的区域残差差值超过 3m。南矶湿地区有 1.4%的离散点残差差值超过 2m。为了探究误差产生的原因，分别对两个研究区的残差空间分布进行分析。图 7 表明，裸滩区残差差值大于 2m 的区域零散分布在入江水道沿岸，西北部、中部和南部靠近水体的极小部分区域反演误差较大，差值大于 1m。0.5~1m 的残差差值主要集中分别于南部。裸滩区残差整体空间分布相对规律，主要沿鄱阳湖入江水道分布。这可能是由于利用 NDWI 剔除通江水域时，水体与非水体的界定不够清晰，在提取裸滩区光谱特征时，囊括了部分水体信息，反演结果受到了水体干扰。此外，土壤干湿度也会影响反演结果，靠近入江水道的区域土壤湿度相对较大，湿润土壤比干燥土壤吸收更多的电磁波，导致在可见光、近红外波段反射率明显降低，导致这部分区域反演效果不佳。南矶湿地区残差差值大于 2m 的区域零星分布在西部、南部，大于 0.5m 的区域主要分布在西南部，此外，在除东北部外的整个南矶湿地区都有

零散分布，残差空间分布较为离散。不难发现，残差大的区域主要分布在不同土地覆盖类型的相邻区域。为了探究其误差产生原因，在裸滩区随机抽样 700 个样本点与南矶湿地的泥滩区进行比较。结果表明，南矶湿地区的泥滩区域（R^2：0.37，MAE：0.61，MRE:5.45%）比土地覆盖类型单一的裸滩区（R^2：0.56，MAE：0.37，MRE:5.74%）的精度低，这可能是由于南矶湿地区水体、泥滩、植被交错分布，在不同土地覆盖类型的相邻区域，遥感影像无法精确地提取地物特征，较为复杂的土地覆盖类型影响了波段反射率提取的准确度，从而降低了模型在此区域的反演精度。

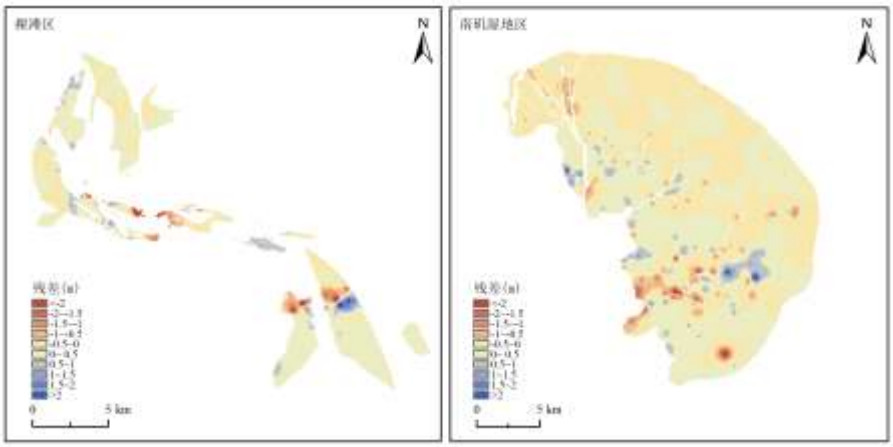

图 7 残差空间分布图

Fig.7 Spatial distribution of residuals

为了进一步确定两个研究区误差产生的原因，分别在裸滩区和南矶湿地区选取 2 个断面，对比分析反演地形和实测地形。图 8 为两个研究区断面所在的位置，对比验证结果见图 9。不难发现，地形平缓处的高程反演误差较小，在地形陡峭的区域，高程反演误差较大。表明模型在大部分平缓区域的反演精度较高，但在地形起伏大的复杂区域仍需提高精度。

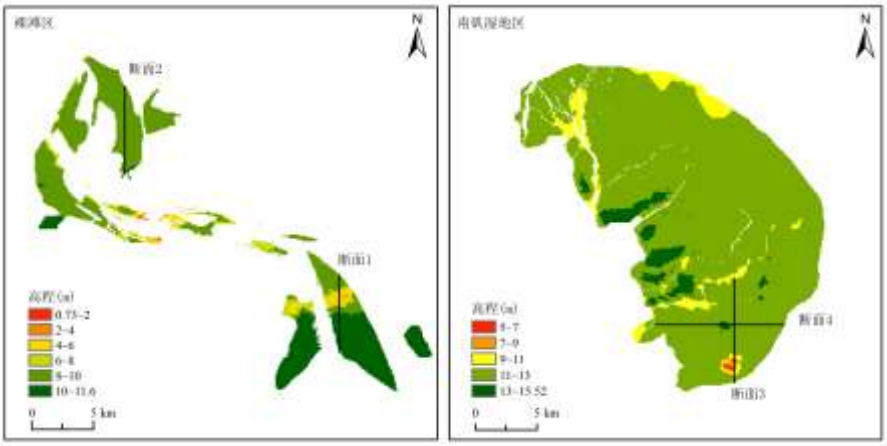

图 8 研究区及验证横断面位置

Fig.8 Study areas and checked cross-section

4 讨论

GWR-RF-OK 混合模型在估测鄱阳湖枯水期的季节性洲滩的土地覆盖类型单一区域和较复杂区域的地形时均表现出较高的精度。与利用水边线法和淹水频率法研究湖泊地形高程的方法相比，通过地形高程

数据与单幅遥感影像的光谱信息反演鄱阳湖局部地形高程，既降低了数据获取难度，又避免了提取多幅遥感影像水体时可能产生的误差。与直接利用 ICESat-2 卫星数据构建地形高程相比，基于更成熟的光学卫星遥感，数据的获取和预处理更便捷，可应用的空间范围更大。ICESat-2 卫星在其发射（2018 年）之前的数据缺失，而光学遥感拥有丰富的存档数据，可应用的时间范围更长。理论上当研究区域没有发生大的地物覆盖类型和高程变化时，所提方法适用于其他所有具有光学遥感数据年份枯水期的地形高程反演。然而，本文仅基于枯水期单一时相数据构建模型，模型在平水期、丰水期的适用性有待进一步探究。此外，由于现有实测地形数据年份的限制，为了实现遥感影像和实测数据的精准匹配，本文只采用了 Landsat5 遥感影像，它无法反映小于 30m 的微型地貌特征，对于陡峭地形，会因变化范围小于 30m 而无法被有效区分。在所提方法中，Landsat 数据可以被其他类似的更高分辨率的遥感数据代替，如 Sentinel-2 数据，也可以增加其他类型的遥感数据和光学遥感数据一起作为特征变量，如 Sentinel-1 数据，以提升地形反演精度。

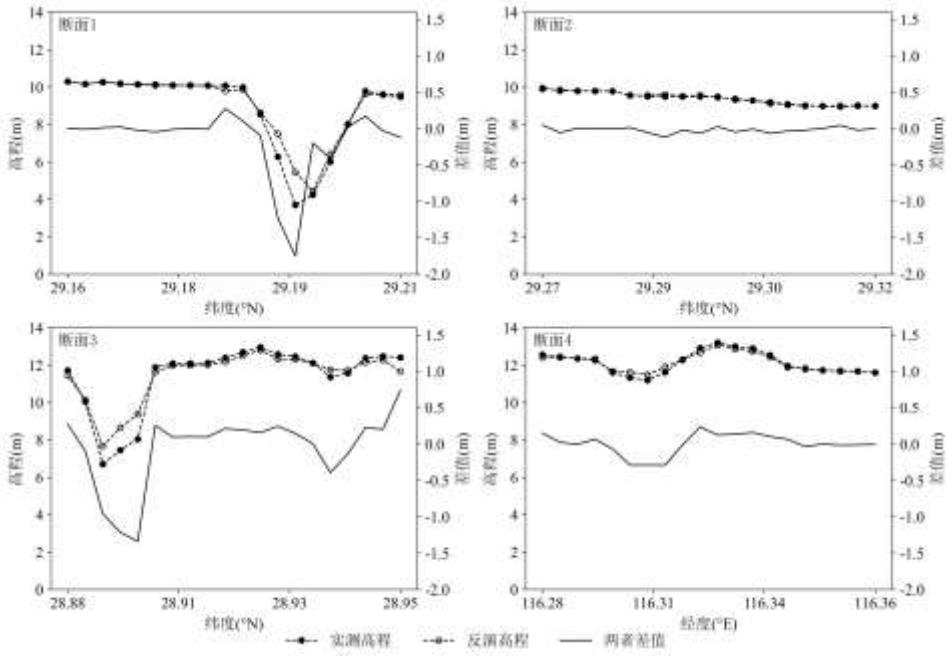

图 9 地形精度验证的断面对比

Fig.9 Comparison of measured and extracted cross-sections

以往在反演地形高程时，多数研究仅对单一地表覆被地区的地形或海拔进行估算[43-45]。本文考虑到土地覆盖类型不同的问题，在具有多种土地覆盖类型的区域以及区域内的水体、植被和泥滩区域均获得良好的地形反演精度，表明所提方法在不同的土地覆盖类型都获得良好的预测性能，可以预测季节性洲滩和浅水的地形高程。然而，由于遥感光谱波段无法穿透浑浊水体到达水底，也无法穿透植被到达地面，所以此方法在浑浊水域和高大植被覆盖区的预测精度可能不如人意，如何提高此方法在浑浊水域和高大植被覆盖区域的反演精度有待进一步研究。另外，在提取水体时仅使用 NDWI，未考虑低矮草本等植被可能因绿光反射低、近红外反射高导致 NDWI 假阳性（即 NDWI>0 但实际为植被），在后续研究中，可引入归一化植被指数（NDVI）进行联合判别，以减少此类误判。

研究的地区为湖泊季节性洲滩，地形相对平坦。理论上，所提出的方法可以应用于其他高程差异较小的地区。但当所提方法应用于其他具有不同土地覆盖类型和高程特征的地区时，需要对模型参数进行调整，所选的特征重要性也可能会发生变化。此外，两个研究区中样本点少的高程段误差最大，样本点的数量可能会影响模型的精度，过少的样本点反演的地形高程精度明显低，因此在未来研究中应尽可能多地选择样本点以避免这一问题。

与 RF 模型相比，GWR-RF-OK 混合模型在反演鄱阳湖裸滩区和南矶湿地区的地形高程时都表现出更高的精度，与现有研究侧重反演单一地表覆被的地形高程相比，所提方法在不同地表覆被下具有较高的精度和令人满意的鲁棒性，为估测湖泊局部地形高程提供了一个补充方法。但本文只使用了随机森林模型，在未来工作中，可以应用不同的机器学习模型或深度学习模型进行研究，以进一步提高反演精度。

5 结论

通过耦合地理加权回归、随机森林算法和普通克里金法，本文提出了一种反演鄱阳湖局部地形高程的新方法，可以得出以下结论：

1）相较于单一模型，GWR-RF-OK 混合模型既考虑了地形高程与遥感影像光谱的空间异质性，又考虑了它们之间复杂的相关性，有效地提高了地形高程的反演精度。

2）地形高程数据可以直接用于鄱阳湖洲滩地形的反演，混合模型在土地覆盖类型单一和相对复杂的区域均取得了较高的反演精度。通过分段分析发现，模型在 10m 以上高程范围的反演效果更佳，在鄱阳湖地形反演方面具有极大可行性。

3）利用遥感技术研究出露水面的湖滩地形时，地形起伏和土地覆盖类型会影响反演结果，在地形起伏大和不同土地覆盖类型相邻区域，会产生较大的误差。地形平缓区域误差显著小于陡峭区域，同类地表覆被在土地覆盖类型单一区域的地形反演精度明显优于多种土地覆盖类型混合区域。

6 参考文献

[1] Long YN, Yan SX, Jiang CB, et al. A new method for extracting lake bathymetry using multi-temporal and multi-sour-ce remote sensing imagery: A case study of Dongting Lake. *Acta Geographica Sinica*, 2019, **74**(07):1467-1481. DOI:10.11821/dlxb201907015 [隆院男, 闫世雄, 蒋昌波, 等. 基于多源遥感影像的洞庭湖地形提取方法. 地理学报, 2019, **74**(07):1467-1481.]

[2] Igbinenikaro OP, Adekoya OO, Etukudoh EA. Review of modern bathymetric survey techniques and their impact on offshore energy development. *Engineering Science & Technology Journal*, 2024, **5**(4):1281-1302. DOI:10.51594/estj/v5i4.1018.

[3] Zhang B, Li JS, Shen Q, et al. Recent research progress on long time series and large scale optical remote sensing of inland water. *National Remote Sensing Bulletin*, 2021, **25**(1):37-52. DOI:10.11834/jrs.20210570.

[4] Zhang GQ, Yue JP, Liu SN, et al. Terrain extraction method of Poyang Lake based on remote sensing image. *Bulletin of Surveying and Mapping*, 2022, (05):62-66+73. DOI:10.13474/j.cnki.11-2246.2022.0142. [张龚泉, 岳建平, 刘胜男, 等. 基于遥感影像的鄱阳湖地形提取方法. 测绘通报, 2022, (05):62-66+73.]

[5] Hu ZP, Wang SG. Evolution of scour and sedimentation and its hydrological and ecological effects in Poyang Lake. *Water Resources and Hydropower Engineering*, 2022, **53**(06):66-78. DOI:10.13928/j.cnki.wrahe.2022.06.006. [胡振鹏, 王仕刚. 鄱阳湖冲淤演变及水文生态效应. 水利水电技术(中英文), 2022, **53**(06):66-78.]

[6] Wu GP, Liu YB, Fan XW. Bottom topography change patterns of the Lake Poyang and their influence mechanisms in recent 30 years. *Journal of Lake Sciences*, 2015, **27**(06):1168-1176. DOI:10.18307/2015.0623 [吴桂平, 刘元波, 范兴旺. 近 30 年来鄱阳湖湖盆地形演变特征与原因探析. 湖泊科学, 2015, **27**(06):1168-1176.]

[7] Ou Yang QL. Characteristics and causes of scouring and deposit in Poyang Lake. *Jiangxi Hydraulic Science & Technology*, 2021, **47**(05):335-340. DOI:10.3969/j.issn.1004-4701.2021.05-05 [欧阳千林. 鄱阳湖近 10 年来冲淤变化特征及成因分析. 江西水利科技, 2021, **47**(05):335-340.]

[8] Li M, Kuang R. Topographic Reconstruction of Poyang Lake Based on ICESat-2 ATL08 Data. 2025 6th International Conference on Geology, Mapping and Remote Sensing (ICGMRS): IEEE, 2025, :359-362.

[9] Hu ZP, Lin YR. Analysis of Evolution Process and Driving Factors for Aquatic Vegetations of Poyang Lake in 30 Years. *Resources and Environment in the Yangtze Basin*, 2019, **28**(08):1947-1955. DOI:CNKI:SUN:CJLY.0.2019-08-018. [胡振鹏, 林玉茹. 鄱阳湖水生植被 30 年演变及其驱动因素分析. 长江流域资源与环境, 2019, **28**(08):1947-1955.]

[10] Wei LJ, Yang H, Wang CF, et al. Using remot esensing to compute water depth in coastal water of North Jiang su based on Landsat8 satellite. *Transactions of Oceanology and Limnology*, 2022, **44**(06):23-32. DOI:10.13984/j.cnki.cn37-1141.2022.06.004. [李经纬,

杨红, 王春峰, 等. 基于 Landsat 8 卫星的江苏北部近岸海域水深遥感反演研究. 海洋湖沼通报, 2022, **44**(06):23-32.]

[11] Wu J, Ke CQ, Cai Y, et al. Monitoring multi-temporal changes of lakes on the tibetan plateau using multi-source remote sensing data from 1992 to 2019: A case study of lake Zhari Namco. *Journal of Earth Science*, 2024, **35**(5):1679-1691. DOI:10.1007/s12583-022-1639-8.

[12] Jain A, Ramakrishnan R, Thomaskutty A, et al. Topography and morphodynamic study of intertidal mudflats along the eastern coast of the Gulf of Khambhat, India using remote sensing techniques. *Remote Sensing Applications: Society and Environment*, 2022, **27**:100798. DOI:10.1016/j.rsase.2022.100798.

[13] Ma SX, Wang N, Zhou LL, et al. Inversion of Tidal Flat Topography Based on the Optimised Inundation Frequency Method—A Case Study of Intertidal Zone in Haizhou Bay, China. *Remote Sensing*, 2024, **16**(4):685. DOI:10.3390/rs16040685.

[14] Chen CP, Zhang C, Tian B, et al. Tide2Topo: A new method for mapping intertidal topography accurately in complex e-stuaries and bays with time-series Sentinel-2 images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, **200**:55-72. DOI:10.1016/j.isprsjprs.2023.05.004.

[15] Feng L, Hu C, Chen X, et al. MODIS observations of the bottom topography and its inter-annual variability of Poyang Lake. *Remote sensing of environment*, 2011, **115**(10):2729-2741.

[16] Chen C, Wu GP, Niu HL, et al. Combining multisource satellite data to estimate lake basin bathymetry for seasonal water bodies. *National Remote Sensing Bulletin*, 2024, **28**(10):2673-2687. DOI:10.11834/jrs.20232588. [陈闯, 吴桂平, 牛汇林, 等. 面向季节性水体的湖盆地形多源遥感协同定量估算. 遥感学报, 2024, **28**(10):2673-2687.]

[17] Liu D, Li Y. The Calculation of Area and Storage of Poyang Lake Based on Remote Sensing Technology. *Remote Sensing Information*, 2012, (02):57-61. DOI:10.3969/j.issn.1000-3177.2012.02.011 [刘东, 李艳. 基于遥感技术的鄱阳湖面积库容估算. 遥感信息, 2012, (02):57-61.]

[18] Tao X, Zhang D. Inversion of Tidal Flat Elevation Based on Hyperspectral Remote Sensing. *Advances in Marine Science*, 2013, **31**(04):498-507. DOI:10.3969/j.issn.1671-6647.2013.04.008 [陶旭, 张东. 潮滩滩面高程的高光谱遥感反演研究. 海洋科学进展, 2013, **31**(04):498-507.]

[19] 李静, 方位达, 龚明劼. 基于含水量空间分布的露滩地形遥感反演研究. 江苏省测绘地理信息学会 南京：江苏省测绘地理信息学会年学术年会, 2018.

[20] Panagiotou E, Chochlakis G, Grammatikopoulos L, et al. Generating elevation surface from a single RGB remotely sensed image using deep learning. *Remote Sensing*, 2020, **12**(12):2002. DOI: DOI: 10.3390/rs12122002.

[21] Bergsma EW, Almar R, Rolland A, et al. Coastal morphology from space: A showcase of monitoring the topography-bathymetry continuum. *Remote sensing of environment*, 2021, **261**:112469. DOI:10.1016/j.rse.2021.112469.

[22] Dou M, Shi YX, Qu LB, et al. Zonal Inversion of Underwater Topography of Large Water Bodies Based on Landsat Images. *Journal of Zhengzhou University(Engineering Science)*, 2025, :1-8. DOI:10.13705/j.issn.1671-6833.2025.05.020. [窦明, 史玉仙, 屈凌波, 等. 基于 Landsat 影像的大型水体水下地形分区反演. 郑州大学学报(工学版), 2025, :1-8.]

[23] Zhao YJ, Zhao LY, Zhang HG, et al. LSTM-based remote sensing inversion of largescale sand wave topography of the Taiwan banks. *Remote Sensing*, 2021, **13**(16):3313. DOI:10.3390/rs13163313.

[24] Hu KX, Lu SL, Li Q, et al. Simulation of lake underwater terrain based on the XGBoost model: a case study of typical lakes on the Tibetan Plateau. *Big Earth Data*, 2025, :1-26. DOI:10.1080/20964471.2025.2515713.

[25] Liu JL, Sun DY, Kong DY, et al. Shallow Water Depth Inversed Using Multispectral Satellite Based on Machine Learning. *Laser & Optoelectronics Progress*, 2023, **60**(10):405-413. DOI:10.3788/LOP220584. [刘瑾璐, 孙德勇, 孔得雨, 等. 机器学习视角下的多光谱卫星浅海水深反演研究. 激光与光电子学进展, 2023, **60**(10):405-413.]

[26] Ji Q, Shen W, Rao YL, et al. Contrastive study on remote sensing retrieval of water depth and machine learning model. *Transactions of Oceanology and Limnology*, 2022, **44**(02):84-89. DOI:10.13984/j.cnki.cn37-1141.2022.02.011. [纪茜, 沈蔚, 饶亚丽, 等. 机器学习模型与遥感反演水深的对比研究. 海洋湖沼通报, 2022, **44**(02):84-89.]

[27] Tang W, Zhao CY, Lin J, et al. Inversion of intertidal zone topography based on optimized random forest regression characteristic parameters. *Geocarto International*, 2023, **38**(1):2213196. DOI:10.1080/10106049.2023.2213196.

[28] Yao SQ, Tan K, Wang YJ, et al. Estimating terrain elevations at 10 m resolution by Integrating random forest machine learning model and ICESat-2, Sentinel-1, and Sentinel-2 satellite remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation*, 2024, **132**:104010. DOI:10.1016/j.jag.2024.104010.

[29] Georganos S, Grippa T, Niang Gadiaga A, et al. Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. *Geocarto International*, 2021, **36** (2):121-136. DOI:10.1080/10106049.2019.1595177.

[30] Hengl T, Nussbaum M, Wright MN, et al. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 2018, **6**:e5518. DOI:10.7717/peerj.5518.

[31] Li X, Wu ZQ, Shen WJS. Addressing challenges in port depth analysis: integrating machine learning and spatial information for accurate remote sensing of turbid waters. *Sensors*, 2024, **24**(12):3802. DOI:10.3390/s24123802.

[32] Su HY, Shen WJ, Wang JR, et al. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. *Forest Ecosystems*, 2020, **7**:1-20. DOI:10.1186/s40663-020-00276-7.

[33] Song LY, Zhang K, Chao LJ, et al. Research on zoning multi-source fusion precipitation simulation method of the Yangtze River Basin based on random forest model. *Water Resources Protection*, 2024, **40**(03):125-132. DOI:10.3880/j.issn.10046933.2024.03.015. [宋蕾玥, 张珂, 晁丽君, 等. 基于随机森林模型的长江流域分区多源融合降水模拟方法研究. 水资源保护, 2024, **40**(03):125-132.]

[34] Adeniyi OD, Brenning A, Maerker MJG. Spatial prediction of soil organic carbon: Combining machine learning with residual kriging in an agricultural lowland area (Lombardy region, Italy). *Geoderma*, 2024, **448**:116953. DOI:10.1016/j.geoderma.2024.116953.

[35] Ma YL, Xiong CY, Yi WP. Sedimentary characteristics and developing trend of sediments in Poyang Lake Jiangxi province. *East China Geology*, 2003, (01):29-37. DOI:CNKI:SUN:HSDZ.0.2003-01-005. [马逸麟 熊彩云, 易文萍. 鄱阳湖泥沙淤积特征及发展趋势. 资源调查与环境, 2003, (01):29-37.]

[36] McFeeters SK. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 1996, **17**(7):1425-1432. DOI:10.1080/01431169608948714.

[37] Lai GY, Zhang ZY, Wang P, et al. Simulation of impacts of Poyang Lake Hydraulic Project on the flow of the Yangtze River using an EFDC model. *Journal of Lake Sciences*, 2017, **29**(03):521-533. DOI:10.18307/2017.0301. [赖格英, 张志勇, 王鹏, 等. 拟建鄱阳湖水利枢纽工程对长江干流流量影响的模拟. 湖泊科学, 2017, **29**(03):521-533.]

[38] Breiman LJMl. Random forests. *Machine learning*, 2001, **45**:5-32. DOI:10.1023/A:1010933404324.

[39] Brunsdon C, Fotheringham AS, Charlton MEJGa. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis*, 1996, **28**(4):281-298. DOI:10.1111/j.1538-4632.1996.tb00936.x.

[40] Oliver MA, Webster RJIJoGIS. Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information System*, 1990, **4**(3):313-332. DOI:10.1080/02693799008941549.

[41] Li Y, Yuan L, Zhao ZY, et al. Inversion of tidal flat topography based on unmanned aerial vehicle low - altitude remote sensing and field surveys. *Remote Sensing for Natural Resources*, 2021, **33**(3):80-88. DOI:10.6046/zrzyyg.2020336 [李阳, 袁琳, 赵志远, 等. 基于无人机低空遥感和现场调查的潮滩地形反演研究. 自然资源遥感, 2021, **33**(3):80-88.]

[42] 江西省水利厅. 江西河湖大典: 长江出版社, 2010.

[43] Xi ZL, Xu HD, Xing YQ, et al. Forest canopy height mapping by synergizing icesat-2, sentinel-1, sentinel-2 and topographic information based on machine learning methods. *Remote Sensing*, 2022, **14**(2):364. DOI:10.3390/rs14020364.

[44] Xu N, Ma Y, Yang J, et al. Deriving tidal flat topography using ICESat-2 laser altimetry and Sentinel-2 imagery. *Geophysical Research Letters*, 2022, **49**(2):e2021GL096813. DOI:10.1029/2021GL096813.

[45] Wang ZF, MA ZC, Xiong ZZ, et al. Assessment of Multi-Spectral Imagery and Machine Learning Algorithms for Shallow Water Bathymetry Inversion. *Tropical Geography*, 2023, **43**(9):1689-1700. DOI:10.13284/j.cnki.rddl.003742. [王照翻, 马梓程, 熊忠招, 等. 多光谱遥感数据与多类型机器学习算法的浅海水深反演方法评价. 热带地理, 2023, **43**(9):1689-1700.]