

基于梯度提升决策树模型的 Sentinel-1 图像浅海水深反演

黄茂苗¹, 魏永亮^{1,2,3}, 唐泽艳⁴, 刘浩¹, 袁文泉⁵, 袁新哲⁶

(1. 上海海洋大学 海洋科学学院, 上海 201306; 2. 上海河口海洋测绘工程技术研究中心, 上海 201306; 3. 上海海洋大学 国际海洋研究中心, 上海 201306; 4. 自然资源部 东海发展研究院, 上海 201306; 5. 西南科技大学 环境与资源学院, 四川 绵阳 621010; 6. 国家卫星海洋应用中心, 北京 100081)

摘要: 利用合成孔径雷达(Synthetic Aperture Radar, SAR)反演浅海水深在海洋遥感中极具挑战性。本文采用梯度提升决策树(Gradient Boosting Decision Tree, GBDT)为核心的机器学习算法, 使用 Sentinel-1、全球水深数据、风场和流场数据来反演杭州湾和长江口南缘相连的浅海区域的水深。首先分析反演的最佳风速和迭代次数, 再对 0~10 m、10~20 m、20~30 m、30~40 m、40~50 m 的分段水深和 0~10 m、0~20 m、0~30 m、0~40 m、0~50 m 的总体水深用相关系数、均方根误差和平均绝对误差进行精度评价, 最后分析反演水深的空间分布特征。结果表明: 反演的最佳风速约为 3.78 m/s, 并且 GBDT 模型达到最佳精度时的迭代次数远小于其他模型, 最佳迭代次数为 4。分段水深中, 40 m 以内的相关系数都高于 0.8, 其中以 10~20 m 的相关系数最高, 为 0.9; 40~50 m 则最低, 为 0.73。40~50 m 的平均绝对误差和均方根误差均为最大, 分别为 1.89 m 和 2.24 m, 20~30 m 的平均绝对误差和均方根误差均为最小, 分别为 0.75 m 和 0.96 m。在总体水深中, 虽然随水深区间的扩大, 相关系数会逐渐增加, 但是平均绝对误差和均方根误差的精度都随水深区间的扩大而下降, 且在 0~50 m 区间内的平均绝对误差和均方根误差最大, 分别为 1.06 m 和 1.59 m, 因此反演的最佳区间为 0~40 m。该区域的水深从杭州湾海岸线开始由浅及深阶梯增加, 反演结果能够较好的表现研究区内的实际水深分布情况, 比较符合当前区域的水下地形特征。

关键词: 遥感; 合成孔径雷达; 水深; 梯度提升决策树; 迭代

中图分类号: P731.22 文献标识码: A 文章编号: 1000-3096(2024)4-0001-17
DOI: 10.11759/hyxx20221121001

水深是海洋信息的重要组成部分之一, 其中浅海水深是海洋地形、海上航行和浅海自然资源开发利用最基础同时也是最重要的海洋空间地理信息之一。人工探测、船载声呐以及激光点云等传统手段对水深探测步骤复杂并且消耗大量人力、物力、财力。而遥感水深反演可以大尺度、迅速、低成本^[1]获取水深数据, 其中光学遥感受大气条件、水体杂质等影响较大, 微波遥感则不受这些因素干扰, 因此微波遥感探测水下地形在海洋科学中具有重要的意义。

合成孔径雷达(Synthetic Aperture Radar, SAR)^[2], 又称主动式微波成像雷达, 依靠电磁散射进行距离探测和目标信息提取。自 20 世纪 50 年代 SAR 技术出现后, 在科研、经济、军事领域发挥了重要作用。据现有理论体系, 可将 SAR 水下地形成像过程概括为: 海流与水下地形相互调制改变表层流速, 然后风致表面短波与海面表层流相互调制改变表层短波空间分

布, 最后雷达波和表层短波相互调制改变归一化雷达截面积(Normalized Radar Cross Section, NRCS)^[3]。

SAR 浅海水下地形遥感研究可以追溯到 1969 年, 第一颗海洋雷达卫星 Seasat 首次向人们展示了 SAR 浅海水下地形探测的巨大潜力^[4]。1978 年, 海洋卫星发射并传回了大量包含浅海水下特征的雷达影像, 自此, 将 SAR 用于浅海地形研究的潜质引起了广泛的关注^[5]。美军在 1982 年开展的雷达海上观测研究中, 对浅海地形的成像理论进行初步的探究^[6-7]。Alpers 和 Hennings 提出的 SAR 浅海水下地形的成像理论模型为 SAR 浅海水

收稿日期: 2022-11-21; 修回日期: 2023-05-10
基金项目: 国家自然科学基金项目(41976174; 41606196)
[Foundation: National Natural Science Foundation of China, Nos. 41976174, 41606196]
作者简介: 黄茂苗(1998—), 女, 四川乐山人, 研究方向为海洋遥感和遥感图像处理, E-mail: H1213965330@outlook.com; 魏永亮(1981—), 山东潍坊人, 通信作者, 副教授, 研究方向为海浪模式、海洋微波遥感, E-mail: yl-wei@shou.edu.cn

下地形量化研究迈出了重要的一步^[8]。杨劲松等^[9]对文氏海浪谱采取完全非线性积分法对 SAR 图像通过仿真计算,得到了合成孔径雷达在浅海信息反演中的最佳雷达参数、海况的等信息。夏长水等^[10]利用 SAR 仿真模型对渤海塘沽海区的水深进行反演。杨俊钢^[11]结合海浪波高频谱解析表达式和成像机理,设计了多源多时相的浅海地形探测方法,通过最小迭代函数反演台湾岛浅滩不同地形下的水深。范开国等改进传统仿真模型的缺陷,基于 MS4 程序和成像机理建立新的仿真模型,对风速风向与图像的特征关系提出了新的认识^[12]。黄韦良等^[13]基于浅海水下地形 SAR 成像机理提出的水深探测新技术对台湾岛浅滩海域进行了遥感探测实例研究。赵现斌等^[14]提出了一种全极化的合成孔径雷达反演参数设计方法,对辐射分辨率和信噪比等技术参数的最佳取值范围进行研究。郑全安等^[15]介绍了 SAR 在探测浅海地形的研究进展、发展前沿和几个具有代表性的雷达探测地形的实例,为研究更复杂的地形打下了基础。Bian 等^[16]提出了一种基于涌浪参数和水深之间的线性色散关系的多时相合成孔径雷达水深探测方法。

传统的 SAR 水深遥感反演模型中所需的参数过多,且初始精度不高,往往需要加入迭代模型进行至少 10 多次的迭代运算才能完成精度提升^[17],而机器学习法可以简化反演的过程同时提高结果的精度。其中梯度提升决策树模型在遥感的监测中远优于常见的几种模型,将其用于水深反演保持高精度的同时可以降低算法的迭代次数提高反演效率。

梯度提升决策树(Gradient Boosting Decision Tree, GBDT),属于决策树模型的其中一个分支,其核心思想是把残差学习作为基础,优化梯度方向形成大量回归树,所有回归树结果累加得最终模型,通过逐次迭代来减少残差同时获取较高精度的结果,适用于处理各种非线性的数据。由 Friedman^[18]提出的基于 Boosting 的梯度提升算法,已经被广泛运用于各种领域。Jeany 等^[19]将 GBDT 用于跟踪算法中,获得了较高的精度。Xu 等^[20]提出一种基于多层梯度提升树的高光谱遥感图像的分类方法。陈训来等^[21]基于 GBDT 模型用新的气象雷达的三维拼图和观测雨量资料建立降水预测模型。同年,拱健婷等^[22]提出一种新型快速无损药材鉴定方法,将 GBDT 模型结合近红外光谱对药材的伪品进行判别。金秀等^[23]发现 GBDT 算法可以集成多种模型的优势。Wei 等^[24]将高光光谱影像用于河道水污染检测,发现随机森林法虽然和梯度提升树有相似的精度,但是计算更加复杂。李怡静等^[25]把高分一号影像和实测水质数据结合 GBDT 算法

搭建了遥感水质参数反演模型,证明该方法可以实现高精度的遥感信息监测。孟然等^[26]将 GBDT 模型用于水深反演,发现其精度远高于已有的水深反演模型。同年, Yao 等^[27]把 Landsat 图像和地面实测数据与 GBDT 模型结合用于北部湾沿岸的叶绿素 *a* 的遥感监测中,探究叶绿素 *a* 的变化规律和时空分布情况。

综上所述,将 GBDT 算法与遥感探测相结合具有明显的优势。本文提出一种 SAR 图像水深反演方法,在 GBDT 中加入 NRCS、基础水深数据、风场数据、流场数据建立水深反演模型,来计算未知区域的水深。全文结构如下:第一部分介绍研究区域的特点与具体数据,第二部分介绍水深反演过程使用的方法和各项模型。第三部分详细分析反演的结果并对其进行精度评价。第四部分对所得结果进行讨论,并分析影响结果的因素。

1 研究区域与数据

1.1 研究区

本文选取的研究区域为杭州湾和长江口南缘相连的浅海区域(图 1),杭州湾整体形态呈现喇叭状,其底部地形从湾口到乍浦地势较为平坦,乍浦以后的地势按 0.1‰~2‰坡度变化向东下降,水深变化大,由于独特的地形特征使得这里经常出现暴涨潮和涌潮。其北岸为长江三角洲的南部,沿岸滩地分布较广,海岸线曲折绵长,有利于海上贸易和交通运输,现已凭借其独特的先天优势成为了优良的港口。以上原因使得西面水域较为浑浊,呈泥浆色,东面水域较为清亮,

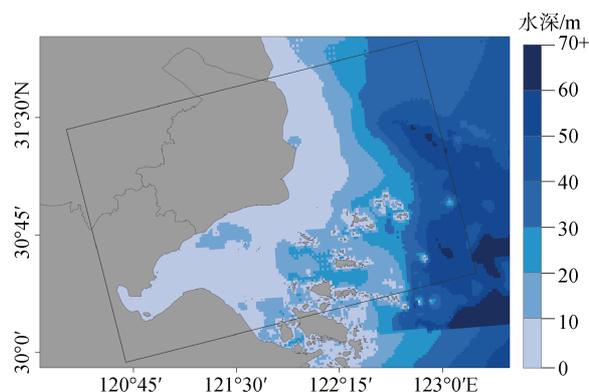


图 1 杭州湾水深示意图(根据审图号 GS(2022)1873 号的地图制作)

Fig. 1 Water depth diagram of Hangzhou Bay (Approval number: GS(2022)1873; ETOPO1 Water depth image data, captured on 2018/06/22 09:55:05UTC, SAR image coverage area is represented by a rectangle)

注: ETOPO1 水深图像数据,时间: 2018/06/22 09:55:05UTC, SAR 图像覆盖区域用矩形表示

呈浅蓝色,同时当前水域时常受大风影响,使得光学遥感水深反演模型在该区域难以发挥作用。

1.2 数据

1.2.1 Sentinel-1 遥感卫星数据

本文使用的哨兵 1 号(Sentinel-1)卫星影像数据来自地球数据网站(<https://search.asf.alaska.edu>),选取 2018 年 1 月至 2018 年 11 月共计 23 景数据,极化方式为 VV、VH,空间分辨率为 10 m×10 m,时间分辨率为 12 d。大量的研究表明,由于海表的特殊性,同极化的回波信号强于交叉极化,对海洋的研究通常采用 VV 极化。

1.2.2 全球水深数据

水深数据来自美国国家海洋大气局(<https://www.ngdc.noaa.gov>)网站中的 ETOPO1 数据,其空间分辨率为 1'。该数据结合了海洋水深和陆地地形,其元数据建立在部分地区和全球数据集的基础上,也可用于冰面和基岩的研究中,将该数据加入水深遥感反演模型前需要根据遥感数据的生成时间进行潮汐校正,便于反演结果更接近真实数据,提高结果的可信度。

1.2.3 风场和流场数据

海表风场数据采用遥感系统网站(<https://remss.com>)提供的 2018 年 1 月 17 日至 2018 年 12 月 19 日的 CCMP 风矢量产品。CCMP 是由微波传感器的海

表风反演和再分析的组合,使用该方法创造的海表风场数据可以保证研究区域内的风场平滑过渡,其时间分辨率为 6 小时,空间分辨率为 0.25°×0.25°。

海表流场数据由 HYCOM 网站(<https://www.hycom.org>)提供的 2018 年 1 月 17 日到 2018 年 11 月 13 日存储在 netCDF 中的数据,共记录了 40 个不同深度层次的海流数据,空间分辨率为 0.08°×0.08°。研究表明,流速对于合成孔径雷达观测水下地形是至关重要的,流速越大越易观测地形,当流向与坡度方向平行时观测效果最佳。

2 方法与反演模型

2.1 图像去噪

SAR 图像中固有噪声往往会最大程度接近或超过要提取的信息,按分辨率的损失情况把噪声的处理方法分为两种:一种是保留原始分辨率的情况下,利用滤波算法抑制噪声,另一种是以损失分辨率为前提对数据进行多视操作抑制斑点噪声^[28]。为保留 SAR 数据高分辨率的特点,本文选用中值滤波平滑图像去除斑点噪声。根据 SAR 的光谱图,利用 IDL (Interactive Data Language)设置不同窗口的中值滤波对图像去噪(将窗口为 3×3 的去噪图像简称为滤波 3),根据表 1 中图像与水深数据的相关系数,判定最佳去噪方式。

表 1 不同去噪窗口对精度的影响

Tab. 1 Influence of different denoising windows on accuracy

图像	原始	滤波 3	滤波 5	滤波 7	滤波 9	滤波 11
相关系数	76.81%	77.59%	77.40%	76.34%	82.25%	76.00%

在探索最佳滤波窗口时,对未去噪和已去噪的数据随机提取浅海的小块区域样本,建立机器学习的水深反演模型。不同数据反演相关系数从高到低依次为:滤波 9、滤波 3、滤波 5、原始、滤波 7、滤波 11。最高为 82.25%,最低为 76.00%。图 2 可以看出,从原始数据到滤波 7 的精度变化不大,经过滤波 3、滤波 5,精度相较原始图像有所提升,滤波 7 精度相较原始图像有所下降,但是整体上精度变化波动较小,未见明显的提升,到滤波 9 精度提升较大(如图 3 所示),从原始的 76.81%提升到 82.25%,滤波 11 精度又明显下降,即反演精度并不会随着滤波窗口的增加而增加,前期会在一个小范围内提升,当到达最佳滤波窗口时,精度提升较为明显,此后继续扩大滤波窗口,精度便会下降。

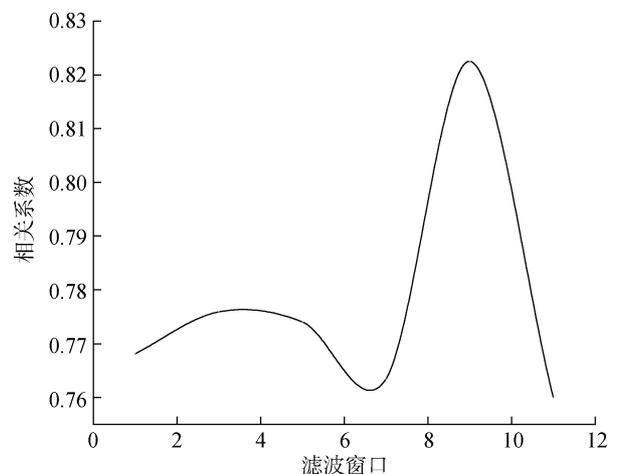


图 2 不同去噪窗口对精度的影响

Fig. 2 Influence of different denoising windows on accuracy

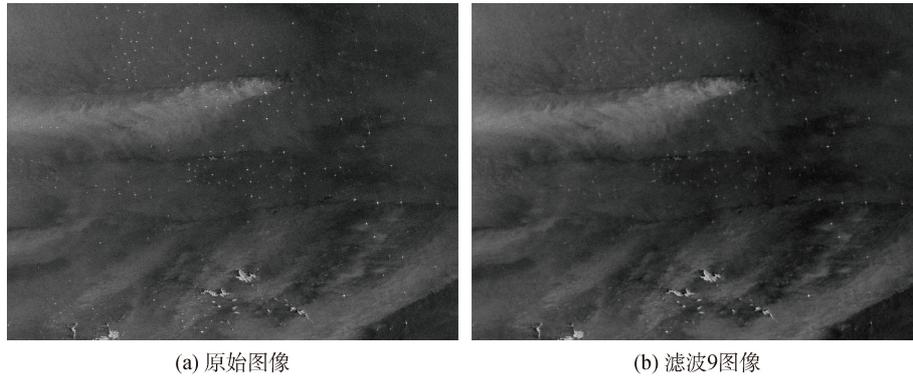


图 3 精度对比图

Fig. 3 Precision comparison chart (Sentinel-1 image data, range: 122.39°E–122.937°E 30.904°N–31.291°N, Time: 2018/06/22 09:55:05UTC)

注: Sentinel-1 图像数据, 范围: 122.39°E~122.937°E, 30.904°N~31.291°N, 时间: 2018/06/22 09:55:05UTC

2.2 SDWI 水体指数模型

雷达水体指数提取法^[29](Sentinel-1 Dual-Polarized Water Index, SDWI), 是遥感数据水陆分割常用的阈值法, 即对图像采取一定预处理算法之后, 假设 K 为图像像元值, T 为阈值。当 $K > T$, 该像元为水体; 当 $K < T$, 为非水体, 对图像进行二值化, 将非水体部分的信息设为 0, 水体部分设为 1, 从而精确提取水体信息。SDWI 算法的基本原理为找出同极化和交叉极化数据之间的初始关系, 将两种极化影像相乘之后再扩大 10 倍来增加水陆像元之间的差异, 在水体信息部分不断拟合、筛选、对比, 最终获取研究区最优的水体信息分割方法, 公式(1)如下:

$$K_{SDWI} = \ln(10 \cdot K_{VV} \cdot K_{VH}) - 8, \quad (1)$$

式中, K_{SDWI} 、 K_{VV} 和 K_{VH} 分别为雷达水体指数计算结果、同极化和交叉极化, 这三张图像的像元值。

2.3 梯度提升决策树水深反演模型

科学家对 SAR 浅海水深反演进行了多年的研究, 提出的反演模型大致可以分为两类, 一是基于 AH 模型改进后的算法, 二是基于 Navier-Stokes 方程的算法。前者需要加入大量实测海洋物理数据, 但结果还需要在迭代模型中进行循环运算, 才能满足一般的精度要求。后者需要解算大量复杂的方程。本文通过对大量 SAR 图像处理与分析, 利用梯度提升决策树算法对包含初步水深信息的 SAR 图像、风场数据、流场数据进行机器学习训练, 从而计算未知区域的水深(如图 4 所示)。

梯度提升决策树是一种基于决策树模型的集成算法, 可以在极少的调参时间下获得较高的精度,

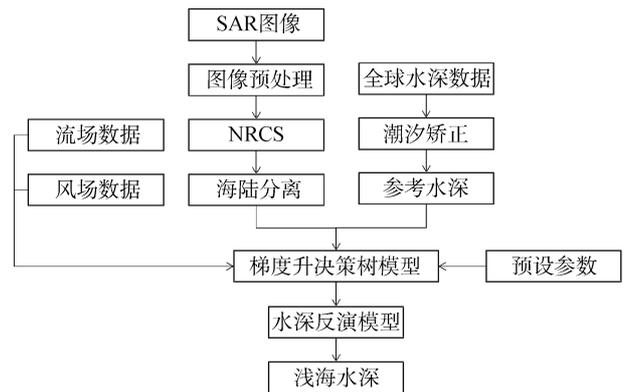


图 4 水深反演流程图

Fig. 4 Water depth inversion flowchart

不管是离散型还是连续型数据都可以灵活处理。在数据预测和分析中使用效果较好, 包含二叉分类树和二叉回归树。通过模型迭代运算, 将分散的弱深度学习器提升组合成一个强深度学习器, 提高水深反演精度, 每次迭代关系密切, 需要上一次模型计算的结果输入到下一次的模型中, 该算法中需要选取最优的迭代次数和最大特征数, 避免在水深精度评价中出现过拟合的状况。

首先建立弱深度学习器, 如公式(2)所示:

$$f_0(x) = \arg \min \sum_{i=1}^N L(y_i, c). \quad (2)$$

对已有水深样本 $i=1, 2, \dots, N$, 计算残差, 如公式(3)所示:

$$r_{im} = - \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)} = f_{m-1}(x). \quad (3)$$

将上一个步骤计算得到的残差(负梯度)作为新

的真实值,把数据 $(x_i, r_{im}), i = 1, 2, \dots, N$ 当作下一决策树的训练数据从而得到新的回归树 $f_m(x)$, 对应的叶子节点区域为 $R_{jm}, j = 1, 2, \dots, J$ 。其中回归树的子叶节点个数为 J 。

子叶区域的拟合值为,如公式(4)所示:

$$\gamma_{jm} = \arg \min_{\gamma} \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma). \quad (4)$$

对公式更新得到强水深度学习器,如公式(5)所示:

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^J \gamma_{jm} I(x \in R_{jm}). \quad (5)$$

其最终生成的水深度学习器,如公式(6)所示:

$$f(x) = f_M(x) = f_0(x) + \sum_{m=1}^M \sum_{j=1}^J \gamma_{jm} I(x \in R_{jm}). \quad (6)$$

2.4 评价指标

为更加准确描述反演结果的精确性,本文采用相关系数(Correlation Coefficient, CORR, 记为 C_{corr})、平均绝对误差(Mean Absolute Deviation, MAD, 记为 D_{ma})以及均方根误差(Root Mean Square Error, RMSE, 记为 E_{rms})来检验水深反演的精度,表达式分别为公式(7)~公式(9),如下:

$$C_{corr} = \frac{\sum_{i=1}^m (x - x_i)}{\sum_{i=1}^m (x - \bar{x}_i)}, \quad (7)$$

$$D_{ma} = \frac{\sum_{i=1}^m |x_i - x|}{m}, \quad (8)$$

$$E_{rms} = \sqrt{\frac{\sum_{i=1}^m |x - x_i|^2}{m}}. \quad (9)$$

3 结果分析与精度评价

3.1 最佳风速

研究表明, SAR 图像上条纹的明暗情况与风速具有较强的相关性^[30],在低风速下 SAR 图像偏暗,高风速下则偏亮^[12]。因此风场是雷达成像过程中至关重要的一个因素,风速过低使得后向散射信号太弱无法被雷达系统探测,风速过高则会出现白沫效应和波浪破碎,使得 NRCS 出现饱和现象,直接掩盖水下地形对信号的影响^[31]。本文在选择水深反演的最佳风速时,参考盛琳等^[32]在利用非线性模型对礁盘的水深反演中使用的逐步回归法,该方法的优势

是可以迅速、运算次数最少同时最大程度保持变量的特性。分析反演初始精度与风速的关系,可以快速选取最佳风速。

水深反演的相关系数与风速的关系如图 5 所示。两者之间并非简单的线性关系,反演的精度随风速的变化一开始是先增加后减小;随后会出现一个精度变化滞缓区间,在该区间内,精度并不会随着风速的增加有明显的变化,一旦风速超过该区间,精度就继续下降。当风速小于 3.78 m/s 时,精度随着风速的增加而增加,当风速在 3.78~6 m/s 时,精度随着风速的增加而减小,当风速在 6~7.5 m/s 时,精度几乎不变,当风速大于 7.5 m/s 时,精度急速下降,最佳风速约为 3.78 m/s。

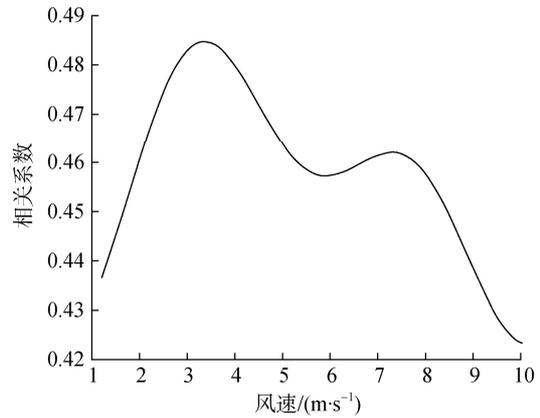


图 5 相关系数与风速

Fig. 5 Correlation and wind speeds

3.2 最佳迭代次数

为研究 GBDT 算法在水深反演中的精度,预先反复试验,确定算法中各项核心参数。对雷达、风场、流场数据随机采样,选定内符合精度进行精度验证,将采样点一部分作为训练集加入到 GBDT 算法中建立水深反演模型,另一部分作为样本集用于精度检验。样本集所得的水深数据与实际水深进行对比,利用相关系数、平均绝对误差、均方根误差作为精度评定标准,并将结果绘制统计图表,如表 2 所示。

表 2 浅海水深反演迭代精度比较

Tab. 2 Comparison of iterative accuracy in shallow water depth inversion

	迭代 1	迭代 2	迭代 3	迭代 4	迭代 5
相关系数	0.94	0.93	0.96	0.98	0.98
平均绝对误差	1.41	1.36	1.18	1.06	1.05
均方根误差	3.17	3.40	2.55	1.59	1.58

GBDT 模型中最重要的参数之一就是迭代次数 (迭代 1 次简称迭代 1, 以此类推)。选取最优迭代次数不仅可以减小算法的运行时间, 还可以提高算法的精度。将样本集逐次增加迭代次数, 记录 3 种评价指标并进行分析。由图 6 可知, 经过第 1 次迭代后, 样本集的大部分反演结果都非常接近甚至等于

原始水深, 但仍然存在个别数据点的反演值与实际值差距较大, 经过第 2 次、第 3 次迭代后, 偏差较大的水深点仍然存在, 第 4 次迭代后, 偏差值较大的水深点全部消失, 第 5 次迭代的散点图和第 4 次的散点图未见明显变化, 且各项精度统计数据也几乎相等。

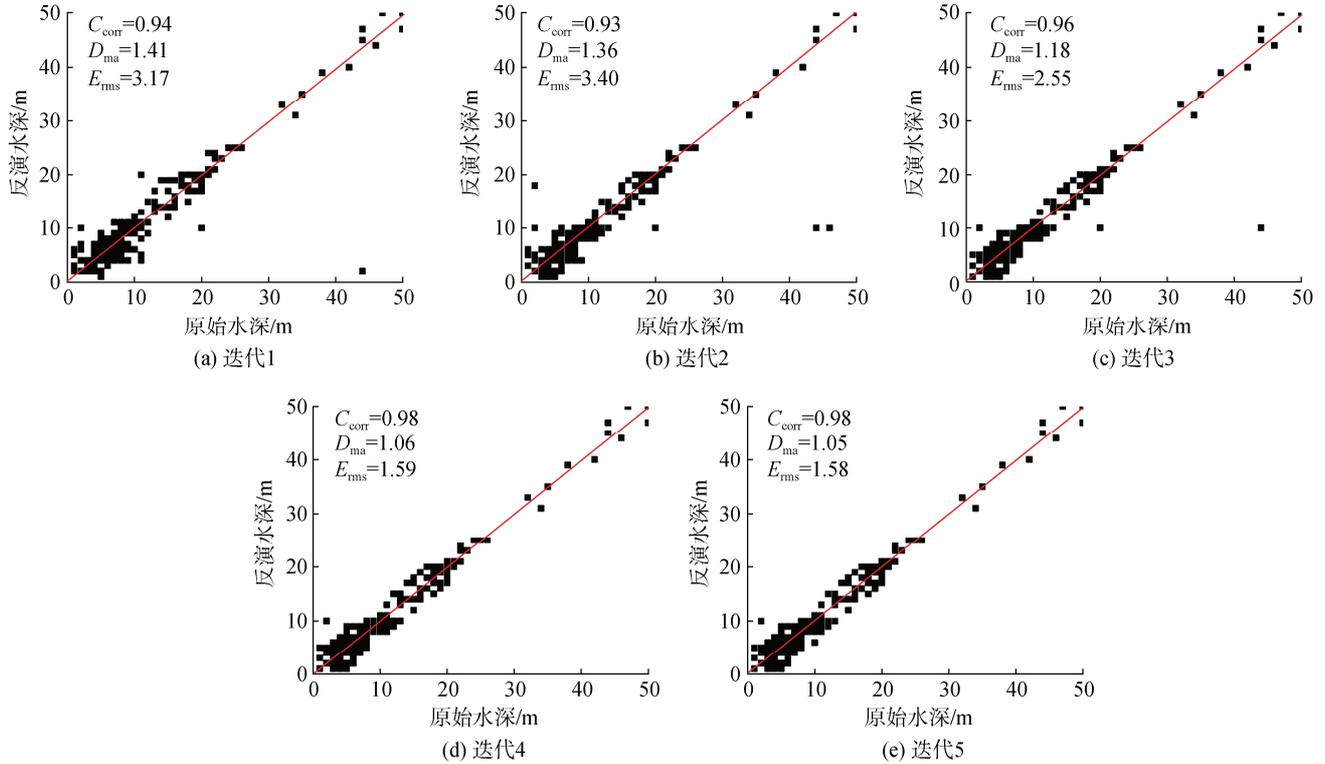


图 6 浅海水深反演迭代散点图

Fig. 6 Iterative scatter diagram of shallow water depth inversion

注: 红线是相关性分析的回归线

由图 7 可知, 第 1 次迭代后, 相关系数和平均绝对误差都取得了较好的精度, 但是均方根误差达到了 3 m; 第 2 次迭代后, 相关系数有所提升, 但变化幅度小, 平均绝对误差明显减小, 但均方根误差却增大, 远超 3 m; 第 3 次迭代后, 相关系数继续提升, 平均绝对误差和均方根误差均下降, 第 4 次迭代和第 5 次迭代后, 3 项指标的数据几乎一致。综上所述, 综合考虑算法运行的时间和计算机的硬件情况, 最佳迭代次数为 4。

3.3 分段水深反演精度分析

为探究反演结果中误差分布情况, 将水深划分为 0~10 m、10~20 m、20~30 m、30~40 m、40~50 m, 表 3 记录了 3 种评价指标, 并将结果分水深制作散点图(图 8)和折线统计图(图 9)。

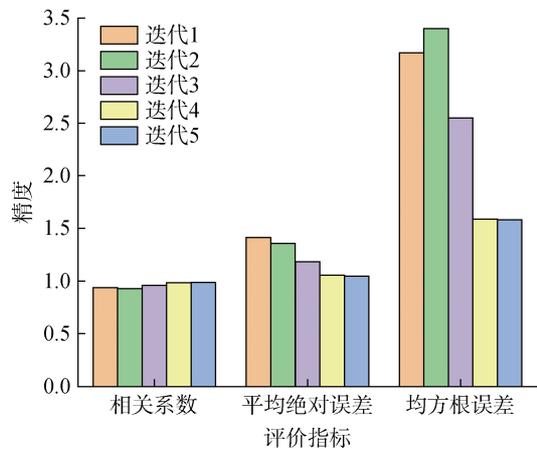


图 7 浅海水深反演迭代精度比较

Fig. 7 Comparison of iterative accuracy in shallow water depth inversion

表 3 浅海水深反演分段精度比较

Tab. 3 Comparison of piecewise accuracy in shallow water depth inversion

分段水深/m	0~10	10~20	20~30	30~40	40~50
相关系数	0.83	0.90	0.87	0.84	0.73
平均绝对误差/m	0.95	1.30	0.75	1.25	1.89
均方根误差/m	1.53	1.72	0.96	1.66	2.24

相关系数、均方根误差和平均绝对误差在 0~10 m 依次为 0.83、0.95 m、1.53 m；在 10~20 m 相关系数

提升了 0.07, 均方根误差和平均绝对误差分别增加 0.35 m、0.19 m；在 20~30 m, 除相关系数轻微下降, 另外两种指标的精度明显增加, 均方根误差和平均绝对误差分别下降 0.55 m、0.76 m, 即当前水深区间的反演精度明显优于 0~10 m 和 10~20 m 的水深; 在 30~40 m, 相关系数下降 0.03, 均方根误差和平均绝对误差分别增加 0.5 m、0.7 m, 与 20~30 m 相比, 精度下降较为明显; 在 40~50 m, 相关系数减少 0.11, 均方根误差和平均绝对误差大幅增长。

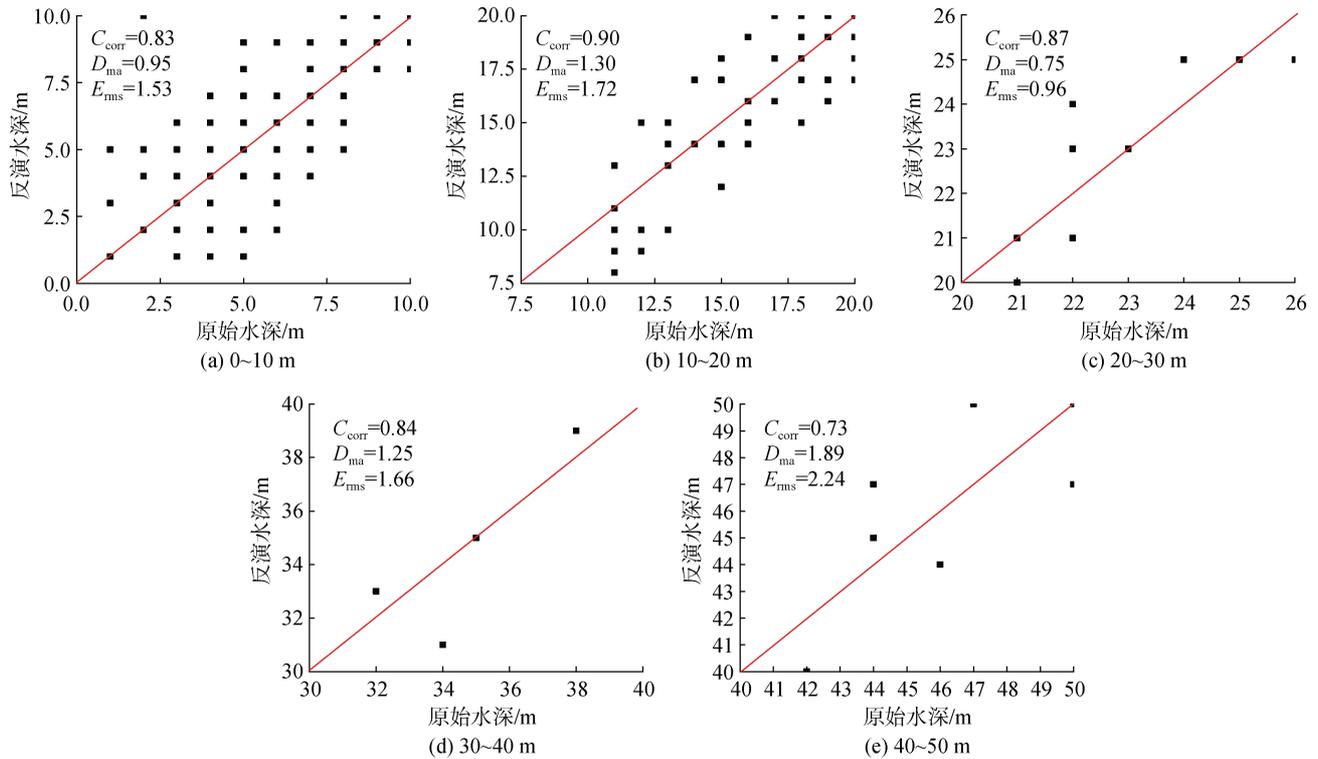


图 8 浅海水深反演分段精度比较

Fig. 8 Comparison of piecewise accuracy in shallow water depth inversion

注: 红线是相关性分析的回归线

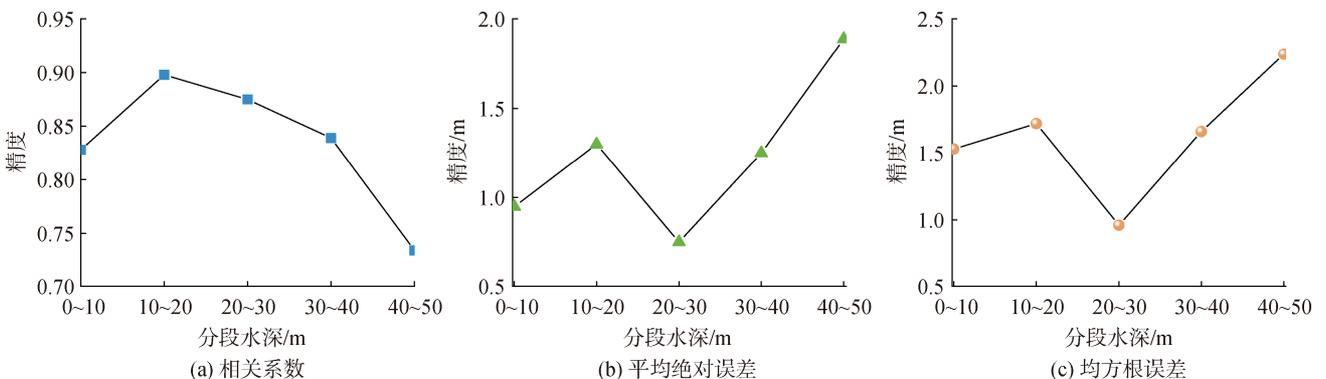


图 9 浅海水深反演分段精度比较

Fig. 9 Comparison of piecewise accuracy in shallow water depth inversion

综上所述,在分段水深中,相关系数随水深段的增加先增大后减小,从高到低依次为 10~20 m、20~30 m、30~40 m、0~10 m、40~50 m,最高为 0.9,最低为 0.73。在 40~50 m 的深水区域相关系数最低,是因为流场与水下地形相互作用反映的水深信息在雷达影像上的表现力有限,随着水深的增加,雷达影像获取的水深信息逐渐减少,因此在较深水域相关系数下降趋势较为明显。平均绝对误差随水深的增加先增加后减小再增加曲折变化,从低到高依次为 20~30 m、0~10 m、30~40 m、10~20 m、40~50 m,其中误差最低为 0.75 m,最高为 1.89 m。均方根误差随水深的变化情况和平均绝对误差一致,从低到高依次为 20~30 m、0~10 m、30~40 m、10~20 m、40~50 m,其中误差最低为 0.96 m,最高为 2.26 m。从 3 种精度指标来看,各水深段的反演精度都比较高,不管是水深探测深度还是精度都远大于光学遥感水深反演模型。

图 9 中相关系数在 0~10 m 水深区间出现精度反常低值,猜测可能是由波浪破碎导致。作为近海表面的一种常见现象,波浪破碎是一种强烈的波与

波相互作用现象多出现在岸边,海浪由于海水底部的阻尼作用使得波峰前倾卷倒,破碎时有大量空气被卷入水中,致使波峰出现白浪,海表产生泡沫、空中溅起大量水滴,次表层海水形成气泡云。波浪破碎在一定程度上会影响着雷达信号的强度大小从而对获取的 NRCS 产生干扰,最终影响水深反演的精度。

3.4 总体水深反演精度分析

为了探究不同深度区间内的反演精度,将水深值按照 0~10 m、0~20 m、0~30 m、0~40 m、0~50 m 进行划分,表 4 中记录了 3 种评价指标,并将结果分水深制作散点图(图 10)和折线统计图(图 11)。

表 4 浅海水深反演总体精度比较

Tab. 4 Comparison of overall accuracy in shallow water depth inversion

水深/m	0~10	0~20	0~30	0~40	0~50
相关系数	0.83	0.96	0.97	0.97	0.98
平均绝对误差/m	0.95	1.04	1.03	1.03	1.06
均方根误差/m	1.53	1.58	1.56	1.56	1.59

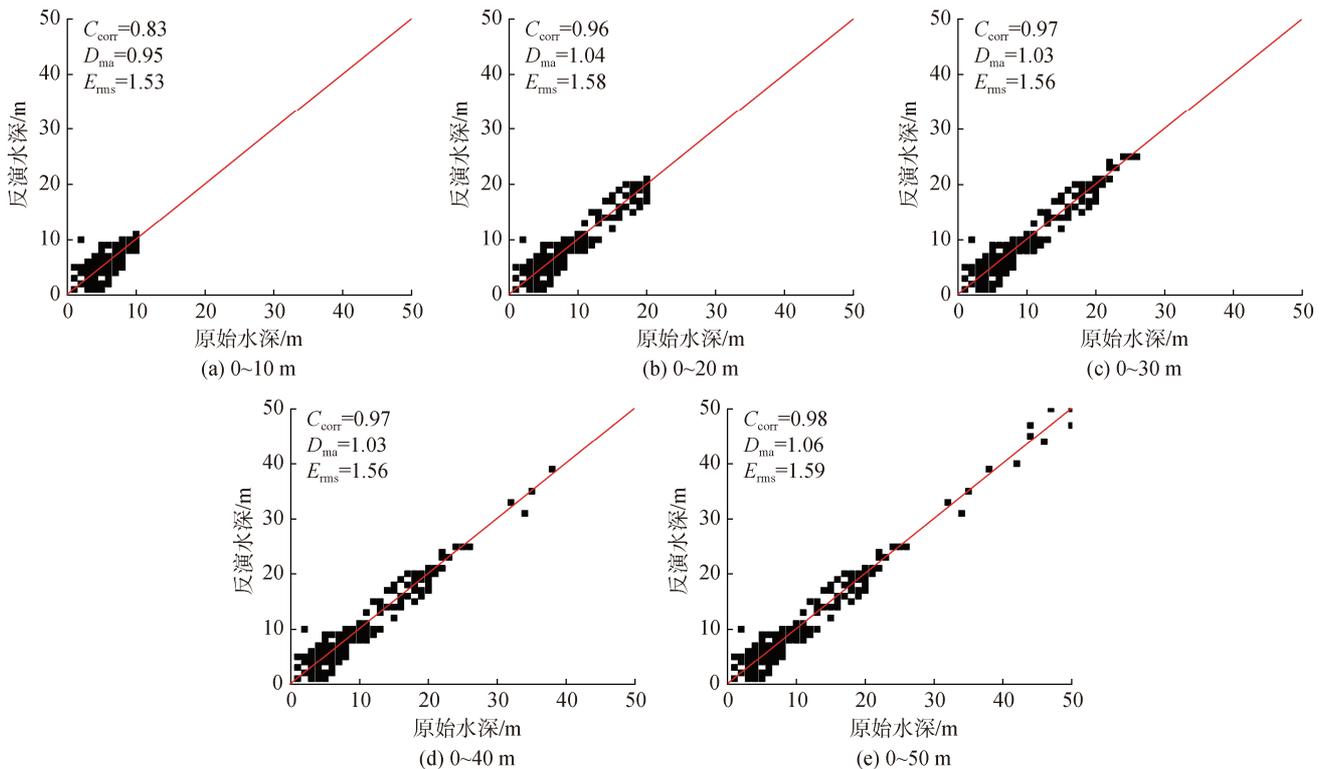


图 10 浅海水深反演总图精度散点图

Fig. 10 Accuracy scatter diagram of general bathymetric inversion maps

注: 红线是相关性分析的回归线

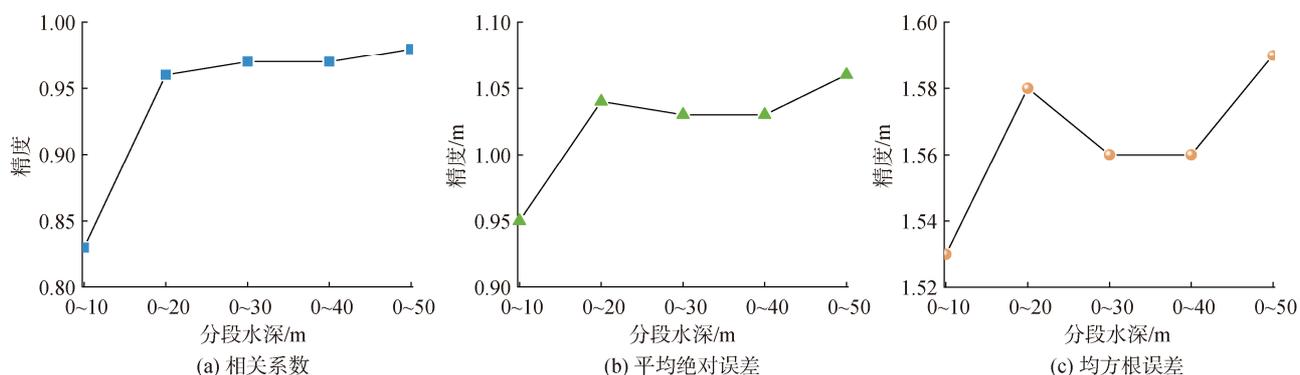


图 11 浅海水深反演总体精度比较

Fig. 11 Comparison of overall accuracy in shallow water depth inversion

当水深区间第一次扩大后, 相关系数会增加, 虽然均方根误差和平均绝对误差的数值也随之增加, 但该区间的精度仍然较好。在 0~30 m, 相关系数较前一个分段增加了 0.1, 均方根误差和平均绝对误差分别下降了 0.01 m、0.02 m, 即该区域的精度轻微上升; 在 0~40 m, 反演精度和 0~30 m 水深区间一致; 在 0~50 m, 除了相关系数增加, 另外两项指标的精度都明显下降, 并且此时的均方根误差已接近 1.6 m, 即反演精度较差。

综上所述, 随着水深增加, 反演的相关系数也会提升, 但是深水区域的实际值和反演值的平均绝对误差和均方根误差都会增加。到达 50 m 左右水深的均方根误差已经接近 1.6 m, 反演值与实际数值的

偏差较大, 即雷达数据提取深水域地形信息的能力较弱, 不可以仅根据相关系数的增加而盲目扩大反演水深区间, 此次反演的最佳水深区间为 0~40 m。

3.5 反演水深空间分布特征

由于研究区内水深像元点的数据量达到千万级别, 在将反演水深可视化时对计算机的配置具有较高的要求, 可利用金字塔降采样算法降低反演的数据量。而风场和流场数据都存在一定的缺失情况, 为提高反演的精度, 在提取训练集和样本集时都排除了缺失数据所在的区域, 使得反演的图像会存在空洞, 利用插值算法填补缺损像元得到反演水深图像图 12(b)。将反演水深与实际水深图 12(a)进行空间分布特征对比分析。

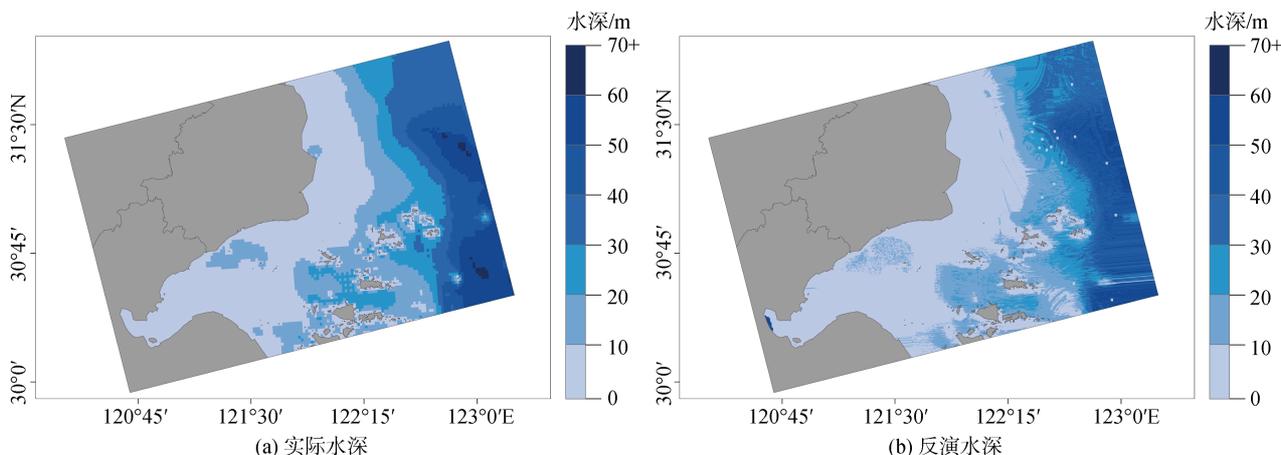


图 12 水深对比图(根据审图号 GS(2022)1873 号的地图制作)

Fig. 12 Comparison of water depth

该区域的水深从杭州湾海岸线开始由浅及深阶梯增加, 浅水区占据大量海域面积。从 120°45'E 到 122°15'E 都是深度为 10 m 左右的浅水区域, 几乎占

据整片海域的 1/3, 同时近岸浅水域内还存在个别零星分布的深水域。122°15'E 到 122°40'E 内的区域水深是 10~20 m, 裹挟了几个零星的浅水域, 几乎包围

了整个舟山群岛。该区域向东是水深为 20~30 m 的区域,从 122°50'E 向东继续前进,经度每增加几分,水深就会增加 10 m。

两幅图像的水深由近岸向深水的延伸趋势几乎完全一致,都是梯度增加。图 12(a)和图 12(b)在 0~10 m 和 10~20 m 水深的分界线明显,边缘过渡清晰;10~20 m 与 20~30 m 水深边界略模糊,但仍可看见较清晰的分界线;20~30 m 与 40~50 m 的水深界线逐渐模糊,需要仔细分辨,再往东水深界线逐渐融合在一起,肉眼难以分辨。反演水深图能较为完整地展现出研究区内的地形特点。图 12(a)在 0~10 m 水深海域内的 30°45'N 附近有一小块明显的深水区域,东西方向狭长分布,而反演水深图中该特征也被提取出来,其分布的范围和形状相似,只是中间夹杂着一些浅水域像元点。在 0~10 m 水深的最南面靠近海岸线的区域,也存在一小块深水区域,在图 12(b)中也体现了该特征。由此可见,反演结果在 0~40 m 水深区间内,能够较好的反应研究区内的实际水深分布情况,比较符合当前区域的水下地形特征。

表 5 最佳风速和前人结果对比

Tab. 5 Comparison between optimal wind speeds and previous results

研究区域	年份	方法	最佳风速/(m·s ⁻¹)	参考文献
任意浅海	2000	M4S 仿真	3~9	[34]
北海	2001	BAS	3~5	[33]
基隆岛	2008	M4S 仿真	4~6	[12]
台湾岛浅滩	2009	M4S 仿真	4	[35]
杭州湾浅滩	2022	GBDT	3.78	本文

综上所述,在采用相同的模型估算最佳风速时,根据研究区域的不同,所得的结论仍会存在小范围的差异,但是整体结果都非常接近,使用不同模型估算最佳风速时,结果虽然不一样,但也是十分接近。由此可知,本文所得的结论具有较高的准确性和较好的适用性。

通过表 6 可以看出利用 GBDT 模型训练不同数据到达最佳精度时的迭代次数差距较大,少的仅为几次,多的能达到上千次。计算机在大多数情况下进行迭代运算时消耗的时间和硬件配置的要求都是成倍增长的,在保证数据达到提取要求的前提下,迭代次数越少越好。本文使用的 NRCS、风场、流场数据和水深建立的水深反演模型仅仅需要进行 4 次迭代就能达到较好的反演效果。

现有的遥感水深反演方法主要分为两类:光学、微波。光学遥感的优势是可以直接穿透水层到达水

4 讨论

4.1 研究结果与前人结论的对比

本文以杭州湾和长江口南缘相连的浅海区域为例,探究最佳风速、最佳迭代次数和不同水深段的精度。

通过参考前人所用非线性模型中的逐步回归法得到研究区内反演的最佳风速约为 3.78 m/s。此次最佳风速研究结果和前人结论对比见表 5。有三组研究人员采用 M4S 系统仿真研究 SAR 浅海地形成像过程中的最佳风速。研究人员和最佳风速分别是:黄韦良等^[34], 3~9 m/s; 范开国等^[12], 4~6 m/s; 范开国等^[35], 4 m/s。即使是采取相同的方法,研究区不同,所得的最佳风速区间也不完全一样,但是都在中等风速区间内,并且较为接近。而 Calkoen 等^[33]使用 BAS 水深估算系统获取的最佳风速区间为 3~5 m/s。和 M4S 仿真系统相比,研究所得的最佳风速仍大致相同。本文的最佳风速为 3.78 m/s,风速结果符合黄韦良等^[34]、Calkoen 等^[33]所得的结论,并且与范开国等^[35]人所得的最佳风速区间只相差 0.22 m/s。

表 6 使用 GBDT 模型迭代次数对比

Tab. 6 Comparison of model iteration times

年份	最佳迭代次数	参考文献
2017	360	[36]
2019	70	[37]
2021	7	[38]
2021	2 500	[39]
2022	300	[40]
2022	4	本文

底并且返回能量,通过计算穿透水层消耗能量与水深之间的关系,从而得出研究区的水深,并且现有的光学遥感水深反演理论模型较多^[49],算法较为成熟,在众多学者的研究改进下,其精度也在不断提升,但是仍存在着无法忽视的重大缺陷:1)受天气和大气条件影响较大。阴雨天气下,研究区内数据含云量较高,目前的云处理技术一般是利用高通滤波弱

化一部分干扰或是直接掩膜去除云区域。无论是哪种方法,都会对水深反演造成严重的影响。2)对研究区内水质情况有严格的要求。如表 7 所示,光学数据

的研究区多选取在水质条件好、人类活动少的地方。3)水深探测能力较弱。从探测深度一栏可知,反演的大部分水深区间处于 10~25 m。

表 7 水深探测能力和精度

Tab. 7 Depth detection capability and accuracy

研究区域	年份	数据	探测深度/m	精度	参考文献
江苏近海	2007	TM	15	相关系数=69.23%	[41]
计算机仿真浅海	2015	SAR	60	未标注	[42]
南海某市某群岛	2016	WorldView-2	22	均方根误差=1.76 m	[43]
西沙群岛	2017	高分一号	20	均方根误差=1.8 m	[44]
太平沙	2018	SAR	18	相关系数=75.42%	[45]
南海某岛礁	2020	Planet 卫星影像	20	均方根误差=1.53 m	[46]
甘泉岛	2021	WorldView-2	20	均方根误差=0.95 m	[26]
海南东部博鳌等	2021	Landsat、Sentinel-2	25	均方根误差=1.42 m	[47]
南海某岛礁	2021	WorldView-2	10	均方根误差=0.96 m	[48]
杭州湾浅滩	2022	SAR	40	相关系数=97% 均方根误差=1.56 m	本文

然而利用合成孔径雷达反演水深恰好可以弥补这些不足。并且 SAR 数据在水深反演的潜力仍在探索中。本文最佳探测深度为 40 m, 远超光学遥感的探测能力。由精度一列可知: 王小珍^[45]、田庆久等^[41]的反演结果的相关性均低于 80%, 而本文的相关性高达 97%。表 7 中其余研究人员在 20 m 水深处的均方根误差最低为 0.95 m, 最高为 1.8 m。而本文反演所得的 40 m 水深的均方根误差仅为 1.56 m。综上所述, 利用 GBDT 模型进行水深反演能保持较强的水深探测能力的同时也具有较高的反演精度。期待在将来的研究中可以改进模型, 进一步提高水深探测能力。

4.2 潮汐数据对结果的影响

虽然 SAR 数据具有分辨率高的特点, 但正因如此利用该数据参与反演的数据量远大于低分辨率的遥感数据, 当研究区选择过大就会使得运算时间过长, 因此本文的研究区仅选择了杭州湾附近的浅滩。

基础水深数据是由重力卫星反演后的数据, 在不考虑潮汐影响下获得的水深, 若利用该数据与实际水深进行对比, 需要查找离研究区最近的潮汐检测站数据。中国浅海沿岸设有许多潮汐观测站, 因此同一个研究区附近可能有多个检测站。目前对海洋数据的潮汐校正方法按数据量分类, 可以分为单一

站点校正和多站点校正。单一站点校正常用于小研究区, 选取离研究区最近或者位于研究区正中心的站点进行校正, 多站点校正常用于大尺度海洋遥感中, 即选取多个潮汐点, 不同的算法进行矫正。如崔丹丹等^[50]利用潮汐分带技术校正遥感图像。

本文数据来自大鱼潮汐网(<https://www.chaoxibiao.net>), 该网站数据仅记录每个站点的整点数据, 而遥感图像的生成时间不能与此完全对应。观测数据如图 13 所示, 整点潮汐之间的数学关系接近于一阶线性函数,

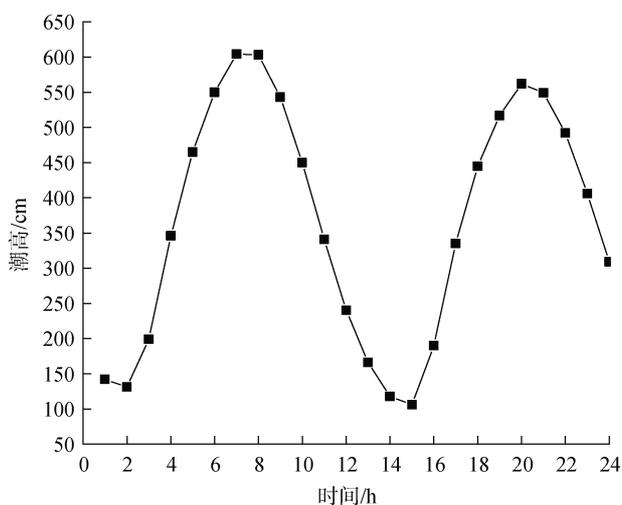


图 13 研究区潮汐数据

Fig. 13 Tidal data of the study area

可据此得出遥感图像对应时间的潮汐数值。不同站点潮汐数据对应的潮高基准面和潮高数据不同,需要计算两者的差值得出与海平面的距离用于潮汐校正。本文的研究区仅为杭州湾附近的浅海区域,且选取的最深水域仅为 50 m,因此选用小区域的单一站点进行校正便可。

4.3 流场数据对水深结果的影响

虽然研究区内水深反演图与实际水深图相比已达到较高的相似性,能较好展示该区域内的水下地形细节,但是仍有个别区域出现了较为明显的误差,如图 14 所示。同一列图表示相同的水域,第一行是实际水深图中的 3 处细节,对应第二行反演水深图的 3 处细节。

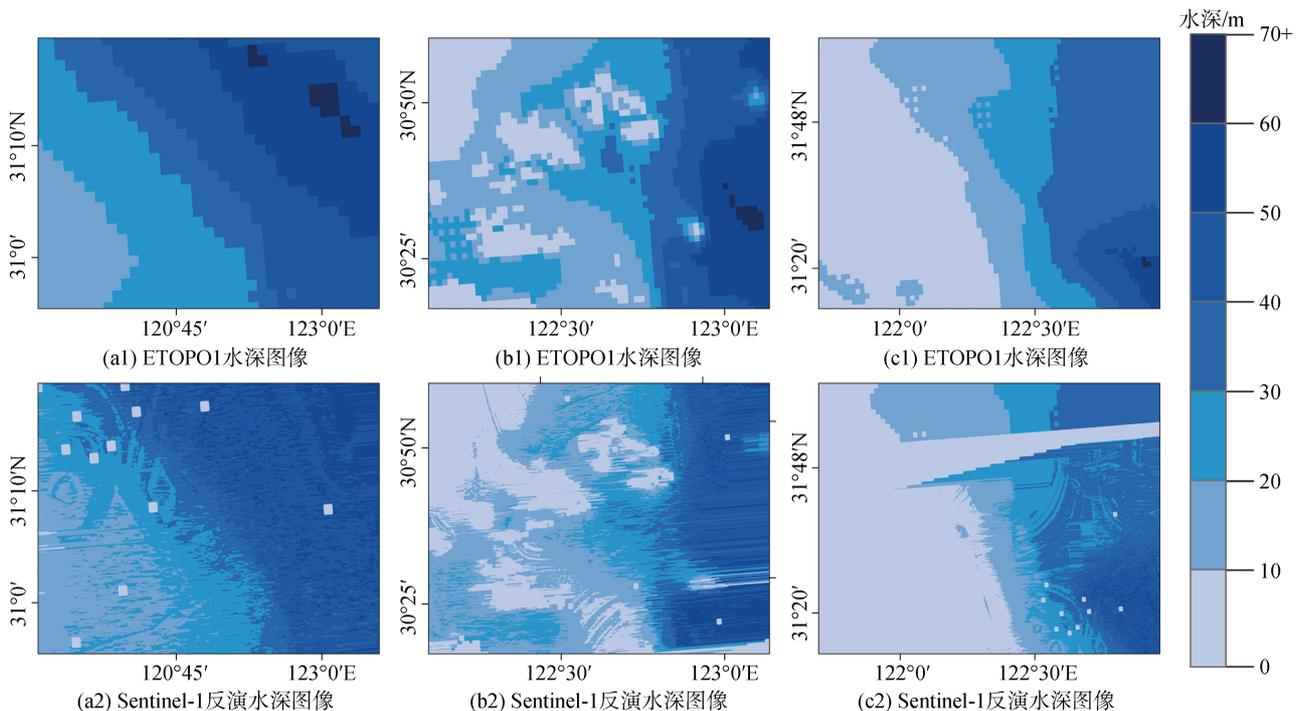


图 14 水深细节对比图

Fig. 14 Comparison chart of water depth details

流场数据对反演结果影响最大,主要是以下两点: 1)流场数据只存在于水体区域,而流场数据的分辨率远低于 SAR 图像,因此两张图像在海陆边缘处的分界线存在一些差异,使得 0~10 m 水域的反演结果较低。2)流场数据存在缺失像元,而 GBDT 水深反演模型中该参数是必不可少的,故在缺少流场数据的区域,无法使用模型进行水深反演,只能利用插值算法获得近似水深数据。第一点可以利用水体指数建立掩膜区间,从而降低边缘差异对反演结果的影响。从图 14 可以看出,该方法处理后的边缘与实际水深边缘无明显差异。而第二点,使反演结果图与实际水深图中出现明显的差异,例如: (a1)和(a2)的右上角实际为 50~70 m 水深,但是(a2)中表现为 40~50 m 水深; (b1)和(b2)的右下角在实际水深图中表现为边界圆润的椭圆状深水水域,而反演水深图中并未展现出这一细节; (c1)和(c2)图的左上居中区域的反演水深图中该

区域有一小块三角状的浅水区,但实际水深图中并无这一细节,且在右下角实际水深图中明显存在 60~70 m 的水深,而反演水深图中也未出现这一细节。由此可知,流场数据缺失对反演结果的影响更大。

4.4 ETOPO1 数据的精确性

本文采用 ETOPO1 水深数据构建模型,该数据海洋部分是用卫星测高技术获得,深海区域具有较高的准确性,而部分浅海的水深则通过参考实测数据进行修改。据文献记载,该数据在经度 105~120°E 水深小于 200 m 的沿海地区容易出现地形虚假起伏变化且广西沿海个别区域的数据也会出现异常平滑的状况^[51]。因此将数据应用于浅海区域进行水深反演时,最好是能借助实测水深点对研究区的数据进行修正,可以大大提高反演水深的准确性。本次研究并未找到该区域的实测数据,希望之后的研究中可以增强数据的可信度和可用性。

5 结论

本次 SAR 浅海水深反演中, 根据后向散射模型对图像进行处理, 从分段和总体两个方面用相关系数、均方根误差和平均绝对误差对反演精度进行分析, 选取最佳风速、迭代次数和数据处理方式, 利用机器学习法进行水深反演。结果如下:

(1)最佳风速约为 3.78 m/s。当风速小于 3.78 m/s 时, 精度随着风速的增加而增加, 当风速在 3.78~6 m/s 时, 精度随着风速的增加而减小, 当风速在 6~7.5 m/s 时, 精度几乎不变, 超过该区间后, 精度直线下降。

(2)最佳迭代次数为 4。前三次迭代中, 三项精度指标均有所变化, 第四次迭代和第五次迭代后, 三项指标的数据几乎一致, 精度不再提升。

(3)在分段水深中, 受近岸波浪破碎影响, 使得最浅水深段精度反常下降, 但各项数据均远大于光学遥感水深反演精度。

(4)最佳反演的水深区间为 0~40 m。相关系数会随着水深的增加而增加, 但是均方根误差和平均绝对误差在 0~50 m 水深范围有较高的误差, 不可根据相关系数的增加而扩大反演水深区间。

(5)该区域水深从杭州湾海岸线开始由浅及深阶梯增加。在 0~40 m 水深的反演结果能够较好地反映研究区内的实际水深分布情况, 比较符合当前区域的水下地形特征。

本文还存在着一些不足之处, 未考虑波浪破碎对水深反演的干扰, 此后的研究中应该考虑建立 NRCS 与波浪破碎参数之间的关系, 更加精确地获取水深信息。

参考文献:

[1] 李凉海, 刘善伟, 周鹏, 等. SAR 卫星组网观测技术与海洋应用研究进展[J]. 海洋科学, 2021, 45(5): 145-156.
LI Lianghai, LIU Shanwei, ZHOU Peng, et al. Research progress of SAR satellite network observation technology and ocean application[J]. Marine Sciences, 2021, 45(5): 145-156.

[2] 张明, 王子龙, 吕晓琪, 等. 一种增强型 YOLOv3 的合成孔径雷达(SAR)舰船检测方法[J]. 海洋科学, 2021, 45(12): 1-7.
ZHANG Ming, WANG Zilong, LV Xiaoqi, et al. An enhanced YOLOv3 method for synthetic aperture radar(SAR) ship detection[J]. Marine Sciences, 2021, 45(12): 1-7.

[3] 杨涛, 赵启钊, 孙光才, 等. 基于多处理板联合 SAR 成像的分布式计算模拟系统研究[J]. 海洋科学, 2021, 45(5): 54-61.

YANG Tao, ZHAO Qizhao, SUN Guangcai, et al. Distributed computing simulation system based on a multi-processing board combined with SAR imaging[J]. Marine Sciences, 2021, 45(5): 54-61.

[4] 范开国, 黄韦良, 贺明霞, 等. SAR 浅海水下地形遥感研究进展[J]. 遥感技术与应用, 2008, 23(4): 479-485.
FAN Kaiguo, HUANG Weigen, HE Mingxia, et al. Progress on remote sensing of the shallow sea bottom topography by SAR[J]. Remote Sensing Technology and Application, 2008, 23(4): 479-485.

[5] FU L L, HOLT B. Seasat views oceans and sea ice with synthetic-aperture radar[R]. Pasadena: California Institute of Technology(Jet Propulsion Laboratory), 1982.

[6] VALENZUELA G R, CHEN D T, GARRETT W D, et al. Shallow water bottom topography from radar imagery[J]. Nature, 1983, 303(5919): 687-689.

[7] GORDON C, GREENWALT D, WITTING J. Surface-wave expression of bathymetry over a sand ridge[C]// NIHOUL J C J. Elsevier oceanography series. New York: Elsevier, 1984, 38: 159-185.

[8] ALPERS W, HENNINGS I. A theory of the imaging mechanism of underwater bottom topography by real and synthetic aperture radar[J]. Journal of Geophysical Research: Oceans, 1984, 89(C6): 10529-10546.

[9] 杨劲松, 黄韦良, 周长宝. 星载合成孔径雷达海浪图像谱仿真研究[J]. 东海海洋, 1999, 17(4): 39-46.
YANG Jinsong, HUANG Weigen, ZHOU Changbao. Satellite-based SAR image spectrum simulation with Wen's wave spectrum[J]. Donghai Marine Science, 1999, 17(4): 39-46.

[10] 夏长水, 袁业立. 塘沽海区海底地形的 SAR 影像仿真与反演研究[J]. 海洋科学进展, 2003, 21(4): 437-445.
XIA Changshui, YUAN Yeli. Simulation and inversion study on the submarine topography in the Tanggu Sea area using SAR Images[J]. Advances in Marine Science, 2003, 21(4): 437-445.

[11] 杨俊钢. 多源多时相 SAR 资料反演水下地形的同化模型[D]. 青岛: 中国科学院研究生院(海洋研究所), 2007.
YANG Jungang. The assimilation model of underwater topography detection with SAR images acquired in different times by different satellites[D]. Qingdao: Graduate School of the Chinese Academy of Sciences (Institute of Oceanology), 2007.

[12] 范开国, 黄韦良, 贺明霞, 等. 风速风向对 SAR 浅海水下地形成像影响的仿真研究[J]. 遥感学报, 2008, 12(5): 743-749.
FAN Kaiguo, HUANG Weigen, HE Mingxia, et al. Simulation study on the effect of wind on SAR imaging of shallow water bathymetry[J]. Journal of Remote Sensing, 2008, 12(5): 743-749.

[13] 范开国, 黄韦良, 傅斌, 等. 台湾浅滩浅海水深 SAR 遥

- 感探测实例研究[J]. 地球物理学报, 2012, 55(1): 310-316.
- FAN Kaiguo, HUANG Weigen, FU Bin, et al. SAR shallow water bathymetry surveys: a case study in Taiwan Shoal[J]. Chinese Journal of Geophysics, 2012, 55(1): 310-316.
- [14] 赵现斌, 严卫, 王迎强, 等. 基于海面散射模型的全极化合成孔径雷达海洋环境探测关键技术参数设计仿真研究[J]. 物理学报, 2014, 63(21): 2018401. ZHAO Xianbin, YAN Wei, WANG Yingqiang, et al. Simulation study on the design of key technical parameters in marine environment sounding with fully polarimetric synthetic aperture radar based on ocean surface scattering model[J]. Acta Physica Sinica, 2014, 63(21): 2018401.
- [15] 郑全安, 谢玲玲. 卫星合成孔径雷达探测海底地形研究进展[J]. 海洋科学进展, 2016, 34(2): 147-161. ZHENG Quanan, XIE Lingling. Progress in research of satellite SAR detection of ocean bottom topography [J]. Advances in Marine Science, 2016, 34(2): 147-161.
- [16] BIAN X L, SHAO Y, WANG S A, et al. Shallow water depth retrieval from multitemporal sentinel-1 SAR data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(9): 2991-3000.
- [17] 李泽军, 王小青, 于祥祯, 等. SAR 图像反演浅海地形的一种改进方法[J]. 电子测量技术, 2012, 35(4): 86-89. LI Zejun, WANG Xiaoqing, YU Xiangzhen, et al. Improved retrieval method for shallow sea topography using SAR image[J]. Electronic Measurement Technology, 2012, 35(4): 86-89.
- [18] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. The Annals of statistics, 2001, 29(5): 1189-1232.
- [19] SON J, JUNG I, PARK K, et al. Tracking-by-Segmentation with Online Gradient Boosting Decision Tree[C].Santiago. IEEE International Conference on Computer Vision (ICCV), 2015: 3056-3064.
- [20] XU S Y, LIU S Z, WANG H, et al. A hyperspectral image classification approach based on feature fusion and multi-layered gradient boosting decision trees[J]. Entropy (Basel, Switzerland), 2020, 23(1): 20.
- [21] 陈训来, 陈元昭, 赵春阳, 等. 梯度提升决策树在雷达定量降水估测中的应用研究[J]. 气象科技进展, 2019, 9(3): 132-137. CHEN Xunlai, CHEN Yuanzhao, ZHAO Chunyang, et al. Application of gradient boost decision tree in radar quantitative precipitation estimation [J]. Advances in Meteorological Science and Technology, 2019, 9(3): 132-137.
- [22] 拱健婷, 李莉, 邹慧琴, 等. 基于近红外光谱和梯度提升决策树建立当归药材及伪品的定性判别模型[J]. 世界科学技术-中医药现代化, 2019, 21(10): 2237-2243. GONG Jianting, LI Li, ZOU Huiqin, et al. Identification of angelica sinensis and its adulterants by NIRS and GBDT[J]. Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology, 2019, 21(10): 2237-2243.
- [23] 金秀, 朱先志, 李绍稳, 等. 基于梯度提升树的土壤速效磷高光谱回归预测方法[J]. 激光与光电子学进展, 2019, 56(13): 131102. JIN Xiu, ZHU Xianzhi, LI Shaowen, et al. Predicting soil available phosphorus by hyperspectral regression method based on gradient boosting decision tree[J]. Laser & Optoelectronics Progress, 2019, 56(13): 131102.
- [24] WEI L F, HUANG C, WANG Z X, et al. Monitoring of urban black-odor water based on nemerow index and gradient boosting decision tree regression using UAV-Borne hyperspectral imagery[J]. Remote Sensing, 2019, 11(20): 2402.
- [25] 李怡静, 孙晓敏, 郭玉银, 等. 基于梯度提升决策树算法的鄱阳湖水环境参数遥感反演[J]. 航天返回与遥感, 2020, 41(6): 90-102. LI Yijing, SUN Xiaomin, GUO Yuyin, et al. Remote sensing retrieval of water quality parameters in poyang lake based on the gradient boosting decision tree algorithm[J]. Spacecraft Recovery & Remote Sensing, 2020, 41(6): 90-102.
- [26] 孟然, 沈蔚, 纪茜, 等. GBDT 模型在遥感水深反演中的应用[J]. 环境生态学, 2021, 3(5): 1-5. MENG Ran, SHEN Wei, JI Qian, et al. The application of GBDT model in remote sensing water depth introverse[J]. Environmental Ecology, 2021, 3(5): 1-5.
- [27] YAO H M, HUANG Y, WEI Y M, et al. Retrieval of chlorophyll-a concentrations in the coastal waters of the Beibu Gulf in Guangxi using a gradient-boosting decision tree model[J]. Applied Sciences, 2021, 11(17): 7855.
- [28] 傅斌. SAR 浅海水下地形探测[D]. 青岛: 中国海洋大学, 2005. FU Bin. Shallow sea bottom topography mapping by SAR[D]. Qingdao: Ocean University of China, 2005.
- [29] 贾诗超, 薛东剑, 李成绕, 等. 基于 Sentinel-1 数据的水体信息提取方法研究[J]. 人民长江, 2019, 50(2): 213-217. JIA Shichao, XUE Dongjian, LI Chengrao, et al. Study on new method for water area information extraction based on Sentinel-1 data [J]. Yangtze River, 2019, 50(2): 213-217.
- [30] 万勇, 时晓磊, 戴永寿. 针对 C 波段组网 SAR 卫星的风场优化反演方法研究[J]. 海洋科学, 2021, 45(5): 74-86. WAN Yong, SHI Xiaolei, DAI Yongshou. Wind field optimization inversion method for C-band net-worked SAR satellite [J]. Marine Sciences, 2021, 45(5): 74-86.
- [31] 于鹏. 浅海水下地形雷达成像理论研究及应用[D].

- 上海: 华东师范大学, 2017.
- YU Peng. Theory and application of radar imaging of shallow water bathymetry[D]. Shanghai: East China Normal University, 2017.
- [32] 盛琳, 王双亭, 周高伟, 等. 非线性模型岛礁礁盘遥感水深反演[J]. 测绘科学, 2015, 40(10): 43-47.
SHENG Lin, WANG Shuangting, ZHOU Gaowei, et al. Research on reefs bathymetry estimation by remote sensing based on nonlinear model [J]. Science of Surveying and Mapping, 2015, 40(10): 43-47.
- [33] CALKOEN C J, HESSELMANS G, WENSINK G J, et al. The bathymetry assessment system: efficient depth mapping in shallow seas using radar images[J]. International Journal of Remote Sensing, 2001, 22(15): 2973-2998.
- [34] 黄韦良, 傅斌, 周长宝, 等. 星载 SAR 遥感浅海水下地形的最佳海况模拟仿真[J]. 自然科学进展, 2000, 10(7): 643-649.
- [35] 范开国, 傅斌, 黄韦良, 等. 浅海水下地形的 SAR 遥感仿真研究[J]. 海洋学研究, 2009, 27(2): 79-83.
FAN Kaiguo, FU Bin, HUANG Weigen, et al. Simulation study on SAR shallow water bathymetry [J]. Journal of Marine Sciences, 2009, 27(2): 79-83.
- [36] 郑佳欢. 网络信贷个人信用评估实证分析-基于决策树组合模型和社交网络[D]. 广州: 暨南大学, 2017.
ZHENG Jiahuan. Empirical analysis on network credit personal credit evaluation-based on decision tree combination model and social network[D]. Guangzhou: Jinan University, 2017.
- [37] 张洋, 叶月, 张宗翔, 等. 基于 GBDT 的学术会议替代计量学评价模型研究[J]. 情报学报, 2019, 38(11): 1150-1159.
ZHANG Yang, YE Yue, ZHANG Zongxiang, et al. Research on altmetrics evaluation model of academic conference based on GBDT[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(11): 1150-1159.
- [38] 冀全伟, 王文磊, 刘治博, 等. 一种基于机器学习算法的岩性填图方法[J]. 地质力学学报, 2021, 27(03): 339-349.
JI Quanwei, WANG Wenlei, LIU Zhibo, et al. A machine learning-based lithologic mapping method [J]. Journal of Geomechanics, 2021, 27(3): 339-349.
- [39] 吴桂龙, 杨志敏, 黄昱. 电力通信运行管理中典型业务数据的智能关联方法[J]. 电信科学, 2021, 37(2): 164-172.
WU Guilong, YANG Zhimin, HUANG Yu. Intelligent correlation method of typical business data in power communication operation management [J]. Telecommunications Science, 2021, 37(2): 164-172.
- [40] 孟昭亮, 张泽涛, 杨媛, 等. 改进的 XGBoost 杂散电
- 流预测及可解释模型[J]. 激光与光电子学进展, 2022, 59(12): 1215001.
- MENG Zhaoliang, ZHANG Zetao, YANG Yuan, et al. Improved XGBoost stray current prediction and interpretable model[J]. Laser & Optoelectronics Progress, 2022, 59(12): 1215001.
- [41] 田庆久, 王晶晶, 杜心栋. 江苏近海岸水深遥感研究[J]. 遥感学报, 2007, 11(3): 373-379.
TIAN Qingjiu, WANG Jingjing, DU Xindong. Study on water depth extraction from remote sensing imagery in Jiangsu coastal zone [J]. Journal of Remote Sensing, 2007, 11(3): 373-379.
- [42] 王静, 陈永强, 李宁, 等. 多视角多波段 SAR 浅海地形反演研究[J]. 雷达科学与技术, 2015, 13(4): 375-383, 389.
WANG Jing, CHEN Yongqiang, LI Ning, et al. Multiple points of view and multiband sar for underwater simulation research[J]. Radar Science and Technology, 2015, 13(4): 375-383, 389.
- [43] 曹彬才, 邱振戈, 朱述龙, 等. 高分辨率卫星立体双介质浅水水深测量方法[J]. 测绘学报, 2016, 45(8): 952-963.
CAO Bincai, QIU Zhenge, ZHU Shulong, et al. Shallow water bathymetry through two-medium photogrammetry using high resolution satellite imagery[J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(8): 952-963.
- [44] 陈本清, 杨燕明, 罗凯. 基于高分一号卫星多光谱数据的岛礁周边浅海水深遥感反演[J]. 热带海洋学报, 2017, 36(2): 70-78.
CHEN Benqing, YANG Yanming, LUO Kai. Retrieval of island shallow water depth from the GaoFen-1 multi-spectral imagery[J]. Journal of Tropical Oceanography, 2017, 36(2): 70-78.
- [45] 王小珍. 浅海典型水下地形 SAR 遥感成像机理和反演研究[D]. 杭州: 浙江大学, 2018.
WANG Xiaozhen. Research on SAR remote sensing imaging mechanism and inversion of typical shallow water topography [D]. Hangzhou: Zhejiang University, 2018.
- [46] 韩中含, 徐白山, 杨成林, 等. 基于 Planet 多光谱影像的南海岛礁水深反演研究[J]. 测绘与空间地理信息, 2020, 43(12): 139-142, 146.
HAN Zhonghan, XU Baishan, YANG Chenglin, et al. Research on reef depth retrieval of South China Sea island based on Planet multispectral image [J]. Geomatics & Spatial Information Technology, 2020, 43(12): 139-142, 146.
- [47] 王久, 邹金. 卫星遥感反演水深技术在偏远地区海图测量中的应用[J]. 北京测绘, 2021, 35(8): 1032-1037.
WANG Jiu, ZOU Jin. Nautical chart surveying in remote areas using satellite-derived bathymetry [J]. Beijing

- Surveying and Mapping, 2021, 35(8): 1032-1037.
- [48] 王鑫, 贝祎轩, 李连阳, 等. 基于机器学习的水深反演方法研究[J]. 国土资源信息化, 2021(5): 22-28.
WANG Xin, BEI Yixuan, LI Lianyang, et al. Research on bathymetry retrieval method based on machine learning[J]. Land and Resources Informatization, 2021(5): 22-28.
- [49] 胡鹏, 赵露露, 高磊, 等. XGBoost 算法在多光谱遥感浅海水深反演中的应用[J]. 海洋科学, 2021, 45(4): 83-89.
HU Peng, ZHAO Lulu, GAO Lei, et al. Application of XGBoost algorithm on multispectral shallow water bathymetry retrieval [J]. Marine Sciences, 2021, 45(4): 83-89.
- [50] 崔丹丹, 张东, 吕林, 等. 基于潮汐分带校正的海岸线遥感推算研究[J]. 海洋测绘, 2017, 37(5): 52-55.
CUI Dandan, ZHANG Dong, LV Lin, et al. Coastline remote sensing prediction based on tide zoning correction[J]. Hydrographic Surveying and Charting, 2017, 37(5): 52-55.
- [51] UEHARA K. Compilation of bathymetric data for the South China Sea: 2: High resolution dataset based on multiple sources[J]. Engineering Sciences Reports Kyushu University, 2016, 37(2): 12-18.

Retrieval of shallow water depths from Sentinel-1 images based on a gradient boosting decision tree model

HUANG Maomiao¹, WEI Yongliang^{1, 2, 3}, TANG Zeyan⁴, LIU Hao¹, YUAN Wenxiao⁵, YUAN Xinzhe⁶

(1. College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China; 2. Engineering Research Center on Estuarine and Oceanographic Mapping, Shanghai Municipal Ocean Bureau, Shanghai 201306, China; 3. International Center for Marine Studies, Shanghai Ocean University, Shanghai 201306, China; 4. East China Sea Institute for Development Research, Ministry of Natural Resources, Shanghai 201306, China; 5. School of Environment and Resource, Southwest University of Science and Technology, Mianyang 621010, China; 6. National Satellite Ocean Application Service, Beijing 100081, China)

Received: Nov. 21, 2022

Key words: remote sensing; Synthetic Aperture Radar; water depth; gradient boosting decision tree; iteration

Abstract: Retrieving shallow water depths using synthetic aperture radar poses significant challenges in the field of ocean remote sensing. In this study, we employed a machine-learning algorithm centered around the gradient boosting decision tree to estimate water depths in shallow sea areas between Hangzhou Bay and the southern margin of the Yangtze Estuary. This was achieved using Sentinel-1 images, global water depth data, and wind and ocean-current data. Initially, we determined the optimal wind speed and number of iterations required for accurate retrieval. The accuracy of these estimates was then assessed across different water depth ranges, namely, 0–10 m, 10–20 m, 20–30 m, 30–40 m, and 40–50 m, for segmented and overall depths of up to 50 m. This evaluation utilized correlation coefficients, root mean square error (RMSE), and mean absolute error (MAE) as the metrics. Finally, the characteristics of the spatial distribution of the retrieved water depths were analyzed. Our findings revealed that the most effective wind speed for retrieval was approximately 3.78 m/s. Notably, the GBDT model required significantly fewer iterations to achieve optimal accuracy than the other models, with the best results obtained after just 4 iterations. In terms of segmented water depth, the correlation coefficients of 0–10 m, 10–20 m, 20–30 m, and 30–40 m exceeded 0.8, with the 10–20-m range showing the highest correlation at 0.9. The 40–50-m range exhibited the largest MAE at 1.89 m and the highest RMSE at 2.24 m, while the 20–30 m range had the lowest MAE and RMSE at 0.75 and 0.96 m, respectively. When analyzing the overall water depths, we observed a gradual increase in the correlation coefficients as the depth range expanded. However, the accuracy, as indicated by the MAE and RMSE, decreased with broader depth intervals. The largest errors were recorded at 0–50 m intervals, with MAE and RMSE values of 1.06 and 1.59 m, respectively, suggesting that an interval of 0–40 m is most reliable for depth retrieval. The water depth in this area gradually increased from shallow to deeper levels, moving away from the coastline of Hangzhou Bay. These results accurately represent the actual water depth distribution within the study area, closely aligning with the underwater terrain characteristics of the region.

(本文编辑: 丛培秀)