

Doi:10.11840/j.issn.1001-6392.2024.04.003

结合贝叶斯特征选择的近海水深反演方法研究

蓝歆玫^{1,3}, 李佳², 叶杨², 朱洪波^{1,3}, 薛国坤²

(1. 辽宁工程技术大学 测绘与地理科学学院, 辽宁 阜新 123000; 2. 大连黄渤海海洋测绘数据信息有限公司, 辽宁 大连 116000; 3. 辽宁工程技术大学 地理空间信息服务协同创新研究院, 辽宁 阜新 123000)

摘 要: 针对现有近海水深数据获取周期长、危险性大, 难以满足大范围近海水深数据高时间分辨率测量的问题。本文以辽东湾地区为研究区域, 提出了一种结合贝叶斯特征选择的改进 LightGBM 近海水深反演方法。该方法首先依据贝叶斯特征选择方法对 13 个波段及波段比值特征进行优选; 其次基于地理空间加权模块计算输入变量的权重; 最后基于 LightGBM 反演模块对水深数据进行反演研究。本文选用辽东湾 200 个不同空间位置的数据样本作为测试集, 验证本文方法和 LightGBM 模型在近海水深反演任务中的精度。结果表明, 本文方法的皮尔森相关系数 (r) 为 0.946, 均方根误差 ($RMSE$) 为 0.265 m, 偏差 ($bias^2$) 为 0.017, 平均绝对百分比误差 ($MAPE$) 为 0.031。反演精度及稳定性均优于经典 LightGBM 模型, 能够适应于近海水深反演。

关键词: 水深反演; 贝叶斯特征选择; 辽东湾; LightGBM

中图分类号: P731.1; P237 文献标识码: A 文章编号: 1001-6932(2024)04-0463-08

Research on inshore depth inversion method combined with Bayesian feature selection

LAN Xinmei^{1,3}, LI Jia², YE Yang², ZHU Hongbo^{1,3}, XUE Guokun²

(1. School of Geomatics, Liaoning Technical University, Fuxin 123000, China;

2. Dalian Huangbohai Marine Surveying Data Information Co., Ltd, Dalian 116000, China;

3. School of Geospatial Information Service, Liaoning Technical University, Fuxin 123000, China)

Abstract: Aiming at the problems of long acquisition period and high risk of existing offshore water depth data, it is difficult to meet the high time resolution measurement of large-scale offshore water depth data. Taking the Liaodong Bay area as the research area, this paper proposes an improved LightGBM offshore water depth inversion method combined with Bayesian feature selection. Firstly, 13 bands and band ratio features are selected according to Bayesian feature selection method. Secondly, the weight of input variables is calculated based on the geospatial weighting module. Finally, based on the LightGBM inversion module, the inversion of water depth data is studied. In this paper, 200 data samples from different spatial locations in Liaodong Bay are selected as the test set to verify the accuracy of the method and LightGBM model in the offshore water depth inversion task. The results show that the Pearson correlation coefficient (r) value of this method is 0.946, the root mean square error ($RMSE$) is 0.265 meters, the deviation $bias^2$ is 0.017 and the mean absolute percentage error $MAPE$ is 0.031. The inversion accuracy and stability are better than the Classic LightGBM model, which can be applied to the inversion of offshore water depth.

Keywords: water depth inversion; Bayesian feature selection; Liaodong Bay; LightGBM

近海水深是影响近海区域社会经济活动的重要基础要素之一, 在近海航运、近岸工程、资源开发调查等领域发挥着重要的作用^[1-4]。由于近海

水深易受潮汐变化、气象条件、水文与水动力因素等影响, 其存在时空异质性的特点。因此, 以船载声呐走航式测量为代表的传统测量手段虽然

收稿日期: 2023-05-23; 修订日期: 2023-08-07

基金项目: 国家自然科学基金 (42071343)

作者简介: 蓝歆玫 (1993—), 博士研究生, 主要从事海洋水质遥感反演研究, 电子邮箱: lxm-work@foxmail.com

通信作者: 李佳, 学士, 助理工程师, 主要从事遥感信息智能提取与分析研究, 电子邮箱: simplaghy@163.com

能够获取较为精确的近海水深数据,但难以满足广域范围内高时间分辨率的近海水深测量工作的需要。

遥感技术的发展为快速获取大范围海洋水深数据提供了可能^[5-7]。基于遥感影像的近海水深反演方法可分为物理模型、统计模型和半理论半经验模型。物理方法通常依据近海水文及水动力原理,建立描述近海水深的半解析物理模型算法和利用全局平差等技术手段实现近海水深反演^[8]。该方法不需要大量数据,减少了对辅助数据的依赖,但普适性不强,难以获取大范围、高精度的水深数据。统计模型通常构建水深数据与遥感反射率之间的关系进而获取反演模型,如线性模型、多项式模型等。该类函数模型在水深反演工作中预测速度较快,但由于水深受影像质量、气象条件等因素的影响较大导致其预测能力不稳定^[9]。其中,半理论半经验模型因其具有反演速度快、模型不易过拟合等特点,较为广泛地应用于近海水深反演领域^[10]。Shah等在单波段水深反演的基础上提出了波段比值算法应用于近海水深反演领域,证明了基于多光谱数据近海水深反演的可行性^[11]。Kutser等通过构建波段比值与实测水深值之间的关系模型进行波段比值水深反演研究^[12]。近年来,随着机器学习的发展,基于机器学习回归分析的水深反演日益受到国内外研究者的关注。刘瑾璐等基于 Landsat 8-OLI 多光谱影像与实测水深数据,对多元线性回归模型、随机森林模型和 BP (Back Propagation) 神经网络模型在近海水深反演工作中的精度进行评估。其中,随机森林模型的反演精度优于其他两种反演方法^[13]。林位衡等针对 BP 神经网络易陷入局部最优解的问题,提出一种基于粒子群优化的 BP 神经网络水深反演模型,获得了较高的反演精度^[14]。Dai 等提出基于 LightGBM (Light Gradient Boosting Machine) 算法的水深反演方法,结果表明,该方法作为随机森林的改进算法,在水深反演任务中获得了更高的精度和速度^[15]。当前基于机器学习方法的近海水深反演研究多集中于反演模型研究及多种机器学习模型之间的比较,关于机器学习水深反演中不同波段组合方法的特征组合对量化模型反演精度的贡献研究较少。

上述方法仅基于遥感反射率及空间位置信息

与实测近海水深数据相结合对广域范围内水深进行反演,未顾及遥感影像各波段反射率及地理空间位置在近海水深反演任务中的贡献程度。贝叶斯特征选择是一种机器学习模型的超参数寻优算法^[16]。利用贝叶斯特征选择构建输入特征选择模块,该模块能够基于遥感反射率及空间地理信息与水深之间的复合因果关系进行最优特征组合选择,本文基于上述关系结合 LightGBM 回归模型提出近海水深加权反演方法。在分析不同特征组合对近海水深反演贡献程度的基础上,构建精确反演近海水深数据的机器学习模型。对上述特征经过加权处理后的水深反演模型,其反演性能和精度也会得到进一步的提高。

因此,本文基于贝叶斯特征选择和 LightGBM 多元加权回归方法对大范围近海海域进行水深反演。首先,基于贝叶斯特征选择分析了遥感反射率和波段比值在内的 13 个特征对水深反演的贡献,然后,基于融合贝叶斯特征选择的改进 LightGBM 多元加权回归模型并结合地理空间加权构建了一个近海水深加权反演模型。本研究的目的一是分析不同特征组合对近海水深反演的贡献;二是构建精确反演辽东湾近海区域水深数据的机器学习模型。

1 材料与方法

1.1 研究区概况

辽东湾为我国渤海三大海湾之一,位于渤海北部,自河北省大清河河口至辽东半岛南端老铁山角以北的海域。本文研究区域(图1)位于辽东湾北部浅海区,介于 $40^{\circ}21'54''\text{N}$ — $40^{\circ}53'17''\text{N}$ 、



图1 研究区地理位置图

121°02'24"E—122°14'46"E。海底地形自湾顶及东西两侧向中央倾斜，湾东侧水深大于西侧，全湾被第三纪以来的厚层沉积物覆盖。湾顶与辽河下游平原相连，水下地形平缓，构成小凌河口到西崴子350 km淤泥质平原海岸。东西两岸与千山、燕山、松岭相邻，水下地形较陡，形成基岩—砂砾质海岸。最深处约32 m，位于湾口的中央部分，河口大多有水下三角洲^[17-19]。因此，辽东湾近海水深情况对该地区的社会经济建设、生态格局变化具有重要意义。

1.2 研究资料

研究时间为2022年4月，采用2022年4月22日覆盖研究区的一景Landsat 9多光谱影像中的B1-B7、B9、B10、B11等10个波段为影像数据源，影像空间分辨率为30 m×30 m，对其进行辐射定标、正射校正、大气校正^[20]等预处理工作。实测水深数据为与影像同期的单波束测深仪采集的辽东湾6 379个测点的水深数据。由于影像成像时间与实测水深数据获取时间不同，对应的潮汐高度也不同，本文根据影像和实测数据获取时间及实测数据基准面信息，查询天文潮位预报表，对近海水深数据进行潮汐改正^[21]。实测水深数据深度范围为6.66~10.36 m。

1.3 研究方法

系统流程图（图2）描述了本文提出的近海水深反演方法的主要步骤。具体流程：利用Landsat 9多光谱卫星传感器获取辽东湾近海地区

影像，并对其进行辐射定标、大气校正等预处理。然后获取与实测水深数据相同空间位置的多波段遥感反射率数据，将遥感反射率与实测水深数据组成近海水深反演数据集。利用该训练集对改进LightGBM网络模型进行训练，通过贝叶斯特征选择、地理空间加权、参数调优，选择最优化模型。利用最优训练模型对Landsat 9影像进行反演，获得辽东湾近海水深数据并进行性能评估进而制作辽东湾部分近海区域水深图。

1.3.1 贝叶斯特征选择

众所周知，由于与反演目标相关性较低的特征中包含大量的误差和干扰信息，使用更多的特征并不总会得到较好的反演结果^[22]。此外，使用更多相关性较低的特征将会出现模型拟合更加困难、计算量加大、反演时间增加等问题。本文为探寻最优的输入特征组合任务，提出一种基于改进贝叶斯优化算法的特征选择方法。贝叶斯优化算法是一种机器学习模型的超参数寻优算法^[16]。该算法的主要原理是通过构造具有已知有限样本点函数所输出的后验概率分布来寻找函数的最佳特征选择。贝叶斯优化算法中有两大核心过程：先验函数过程和采集函数过程。先验函数过程基于高斯过程回归（Gaussian Process Regression, GPR）^[23]从而拟合目标函数，而采集函数过程主要使用置信度上限（Upper Confidence Bound, UCB）、预期改善（Expected Improvement, EI）和改善概率（Probability of Improvement, PI）方法，用于估

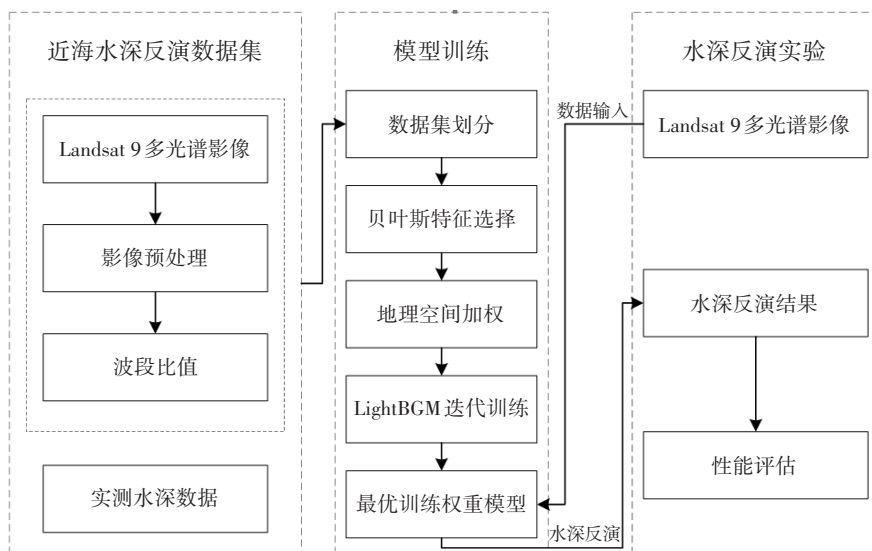


图2 系统流程图

计在当前已知数据条件下,下一个最优点最有可能出现的位置,从而减少不必要的目标函数评估。与此同时,为了防止优化过程中陷入局部极端现象,贝叶斯优化算法通常会加入一定的随机性,在随机搜索和基于后验分布取值之间进行权衡,以在短时间内实现可接受的局部最优。本文以贝叶斯超参数寻优算法为基础,将超参数寻优任务改造为模型寻找最优输入特征组合任务,改进后的贝叶斯特征选择方法如式(1)所示。

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} (y^* - y) \frac{P(x|y)P(y)}{P(x)} dy \quad (1)$$

式中: x 为 LightGBM 近海水深反演模型中的输入特征集合, y 为使用该特征集合的风险值, y^* 为风险值 y 的阈值,即当 y 的风险概率小于 y^* 时以 y^* 表示当前风险概率。 $P(x|y)$ 为观测风险的概率分布,当 y 的风险概率小于 y^* 时以当前超参数 x 形成的概率密度 $e(x)$ 代替 $P(x|y)$ 。反之,则以除当前特征之外的剩余特征形成的概率密度 $g(x)$ 表示。因此, $P(x) = \int_{-\infty}^{+\infty} P(x|y)P(y)dy = y^*e(x) + (1 - y^*)g(x)$ 。可以看出,为获得最佳反演精度,输入特征组合 x 在 $e(x)$ 的概率要尽量大, $g(x)$ 的概率要尽量小。通过计算 $g(x)/e(x)$ 评估特征集合中各特征组合在反演模型迭代上的表现,并选出最优的模型输入

特征组合。

1.3.2 改进 LightGBM 近海水深反演模型

为提升获取稳定、大范围、高精度的水深获取能力,本文提出一种融合贝叶斯特征选择与 LightGBM 的近海水深加权反演模型,如图3所示。本文模型由输入层 (DATA)、贝叶斯特征选择模块 (Bayesian optimization)、地理空间加权模块 (Geospatial Weight module) 以及 LightGBM 反演输出层 (Prediction module) 组成。在经典 LightGBM 的基础上,增加了贝叶斯特征选择模块 (Bayesian optimization) 和地理空间加权模块 (Geospatial Weight module)。图3中 DATA 层输入分为波段遥感反射率、波段比值 (Bands & Band ratios) 及实测水深值 (Water depth) 两部分贝叶斯特征选择模块对输入层输入的13个波段及波段比值特征进行自适应特征选择,选出最适用于当前水深反演任务的最优输入特征组合并导入至地理空间加权模块。由于近海水深数据具有空间连续性,且各测点水深并不相互独立,因此本文将各测点的水深数据及空间位置数据进行地理空间加权,如式(2)所示。

$$y_i = \beta_0(u_i, v_i) + \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (2)$$

式中: y_i 为在位置 i 处的水深值, x_{ik} ($k = 1, 2, \dots, m$) 为位置 i 处的 k 个特征, (u_i, v_i) 为点 i 的坐标,

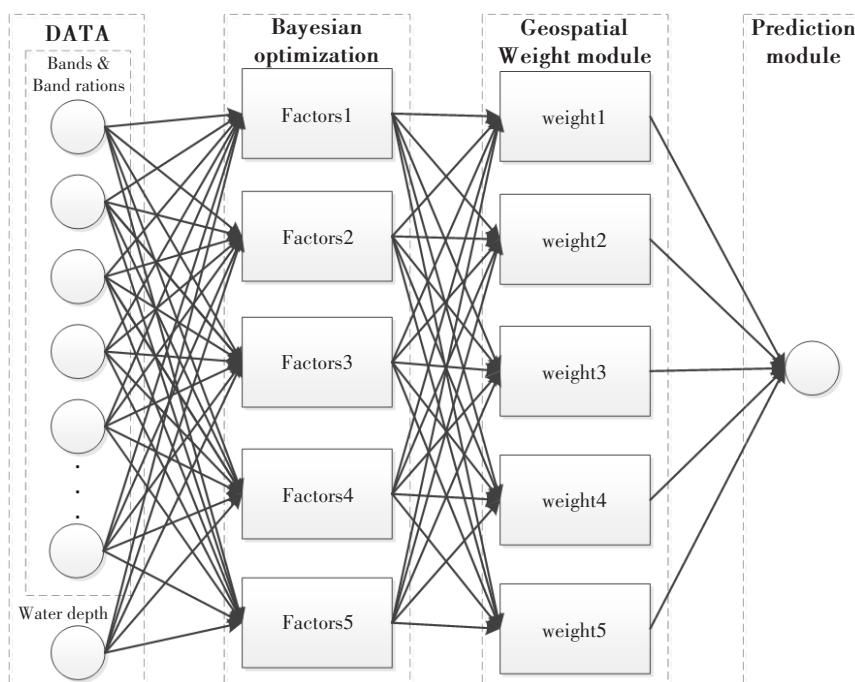


图3 改进 LightGBM 近海水深反演模型结构

$\beta_0(u_i, v_i)$ 为截距项, $\beta_k(u_i, v_i) (k = 1, 2, \dots, m)$ 为回归系数。本文将不同的特征 k 分别输入至式(2)中, 通过计算经地理空间加权以后的各个特征与实测水深值之间的皮尔森相关系数(r)作为模型权重, 输入到LightGBM反演输出层, 实现近海水深信息的高精度反演。

1.3.3 性能评估

为实现对自适应水域提取方法的客观评价, 本文构建检测近海水深反演精度的数学模型, 对该方法的反演准确性进行评估。本文采用4个性能指标对近海水深反演方法进行评价, 即和皮尔森相关系数(r)、均方根误差($RMSE$)、偏差($bias^2$)、平均绝对百分比误差($MAPE$)^[24-25]定义见式(3)~式(6)。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4)$$

$$bias^2 = \frac{1}{n} \sum_{i=1}^n (\bar{y} - x_i)^2 \quad (5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{x_i} \right| \quad (6)$$

式中: x_i 为各采样点的实测水深, $\bar{x}$ 为测区各采样点平均水深, y_i 为各采样点的反演水深, $\bar{y}$ 为上述各采样点的反演水深平均值。其中, $RMSE$ 的值越小, r 值越大, 即模型反演水深精度越高。

2 实验与分析

2.1 水深反演实验

改进LightGBM近海水深反演实验基于机器学习框架scikit-learn 0.24.2构建。网络模型采用Geforce RTX 3090 TI进行训练, 结合Python 3.8环境下的地理空间数据库(Geospatial Data Abstraction Library, GDAL)模块, 对输入特征进行选择及参数调整。

2.1.1 近海水深反演数据集

基于实测水深数据及Landsat 9多光谱影像遥感反射率构建适用于辽东湾近海地区的水深反演数据集。

由于遥感影像在成像过程中电磁波穿过大气

会收到大气、传感器姿态等影响, 为此首先对Landsat 9多光谱影像进行辐射定标、正射校正、大气校正等预处理, 在一定程度上减少影像遥感反射率误差对近海水深反演的影响。

近海水深数据在分布上具有空间连续性, 且受洋流、天气、空间位置等因素也会导致水深数据在遥感影像上反射率的差异。为验证本文提出的方法对于自然状态下近海水深的反演效果, 采用了辽东湾不同空间位置的6379个单波束测深仪的实测水深数据组成数据真值样本。

Landsat 9多光谱影像的空间分辨率为30 m×30 m, 为构建近海水深反演数据集, 首先将上述6379个单波束测深仪测点及栅格影像统一定义为CGCS2000坐标系。其次基于python3.8版本的地理空间数据库(Geospatial Data Abstraction Library, GDAL)模块生成各测点的缓冲区, 缓冲区半径为15 m。最后利用缓冲区图层对栅格影像图层进行裁剪, 选择在实测水深数据测点邻域内最邻近的影像栅格, 并将该栅格的各波段遥感反射率与测点的实测水深值组成近海水深反演数据集。数据集共有4778组样本数据, 并在4778组样本数据中随机选取200组样本构成独立测试集, 其余样本按照9:1比例将其划分为训练集和验证集^[26-27]。

2.1.2 模型训练

本文基于python3.8版本的scikit-learn模块进行建模, 将训练样本中的B1-B7、B9、B10、B11波段和海洋遥感中常用的波段组合B1/B2、B1/B3、B2/B3等13个遥感影像波段特征作为输入的自变量数据, 实测水深数据为因变量^[10]。首先, 将上述数据导入贝叶斯特征选择模块, 依据贝叶斯特征选择自适应选择最优的输入波段特征组合。如图3所示, 在本研究中, 贝叶斯特征选择自适应选取的最优输入特征组合为B1、B2、B3、B5和B1/B2。其次, 由于近海水深数据具有空间连续性, 本文获取每个数据样本的地理空间信息并基于地理空间加权模块对各输入特征进行加权, 如表1所示。最后将上述数据输入至LightGBM近海水深反演模块中进行迭代训练, 确定最优权重模型, 并通过水深反演实验制作研究区水深图。目前在机器学习领域有多种算法可以对输入特征进行选择, 但容易陷入局部最优解。本文基于

scikit-learn 0.24.0 模块构建贝叶斯特征选择算法, 实现近海水深反演任务中的影像最优输入特征组合的自适应选择。该方法不仅当前输入的特征组合与水深真值之间的关系, 而且考虑除当前输入特征组合外的其他特征组合方式与水深真值之间的关系, 同时加入一定的随机性使算法在特征选择的过程中避免陷入不可接受的局部最优情况。Landsat 9 影像的波段及波段组合特征经贝叶斯特征优选后, 利用地理空间加权模块依据水深数据的空间连续性对各不同波段及波段组合赋予地理空间权重, 表 1 展示了优选的波段特征对水深反演的贡献大小。特征选择后的输入变量权重及最优化训练参数如表 1、表 2 所示。经贝叶斯特征选择和地理空间加权后的影像特征与实测水深真值之间的相关性更强, 反演特征更加突出, 为大范围、高精度的近海水深反演奠定了基础。

表 1 模型最优化输入变量及权重

自变量	B1	B2	B3	B5	B1/B2
r	0.577 2	0.540 4	0.542 8	0.506 9	0.454 2
权重	1	0.876	0.89	0.771	0.62

表 2 模型最优化训练参数

参数指标	max_depth	num_leaves	学习率	迭代次数	损失值
参数	20	85	0.118	500	0.1

利用训练集对改进 LightGBM 近海水深反演模型进行训练, 其训练过程如图 4 所示。“Iterations”表示训练样本被迭代的次数, “I2”表示反演模型在训练集上的预测结果与实测水深数据之间的差异, 其被计算为损失并通过损失函数量化表示, 模型在第 500 个 Epochs 达到最优化权重模型。

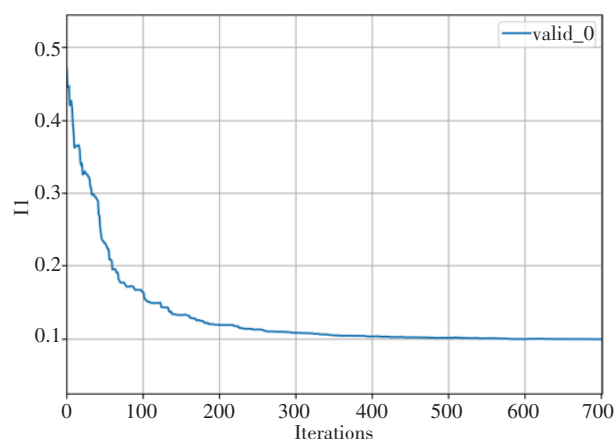


图 4 模型训练过程

2.2 实验结果分析

本文依据辽东湾地区 Landsat 9 影像及与之相匹配的同期实测水深数据, 并从中选择 200 个不同空间位置的数据样本组成独立测试集, 作为反演模型的精度评估。利用上述测试集以实测水深数据为真值, 对本文方法与 LightGBM 模型在自然状态下的水深反演能力进行客观评价。得到上述测试集中上述两种反演模型的 r 、 $RMSE$ 、 $bias^2$ 及 $MAPE$ 性能评估指标均值 (图 5、图 6)。此外, 为评估水深反演模型的计算效率, 本文统计模型运算时长, 如表 3 所示。本文模型在经典 lightGBM 模型的基础上进行由于增加了特征选择模块, 选择 5 个特征作为最佳输入特征组合, 减少了模型运算量, 进而导致模型运算效率的提高。

表 3 近海水深反演模型计算时长

反演模型	运算时长/s	反演模型	运算时长/s
经典 lightGBM 模型	81.25	本文模型	60.41

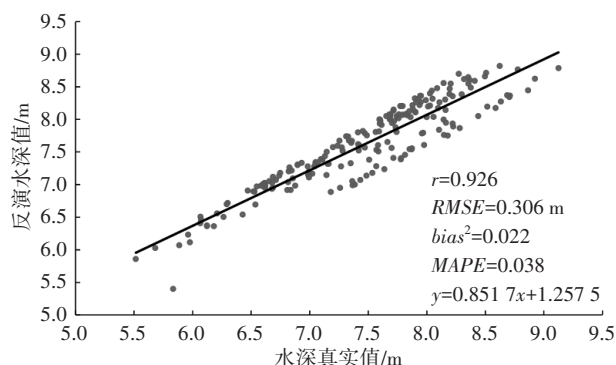


图 5 lightGBM 模型反演水深与真实值关系示意图

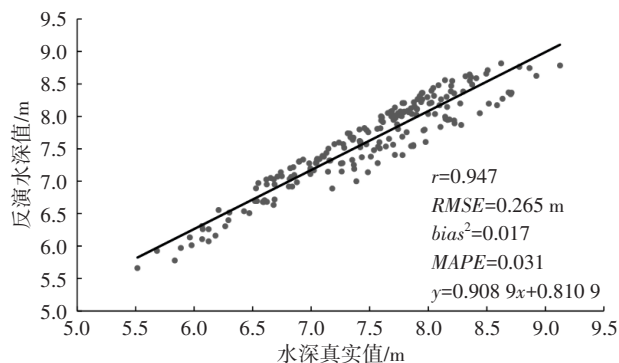


图 6 本文模型反演水深与真实值关系示意图

如图 5、图 6 所示, 两者分别为经典 LightGBM 模型及本文方法输入 B1-B7、B9、B10、B11 以及波段组合 B1/B2、B1/B3、B2/B3 等 13 个特征对测试集辽东湾 200 张不同空间位置的数据样本进行水深

反演实验。其中图5为经典LightGBM模型直接获取的反演水深结果；图6为本文方法在输入上述13个特征经贝叶斯特征选择模块自适应选取最佳输入特征组合的基础上，导入数据样本的地理空间信息并对其进行地理空间加权最终输入LightGBM模型获取反演水深结果。结果表明：结合贝叶斯特征选择和地理空间加权的水深反演方法选择的自变量误差较小，与实测水深值相关性较高，数据特征更接近于自然状态下辽东湾水深分布情况。因而本文方法的皮尔森相关系数较LightGBM提升了0.021，RMSE误差减少了0.041 m， $bias^2$ 偏差降低了0.005，MAPE减少了0.007。在水深反演实验中，随着深度的增大，前述两种水深反演模型的反演误差逐渐增大，这与光在水中的透射率随深度的增大而减小的定律相符合。但在此基础上，本文方法与LightGBM模型相比，水深反演效果更加趋于稳定。图7为利用本文方法反演的辽东湾部分近海海域2022年4月水深专题图。由图7可知，在水深专题图中，水深值随图中色带颜色的加深而减小。在研究时段该海域东南部鲅鱼圈港、西北部锦州港葫芦岛港、东北部盘锦港等部分地区为保证大型船只通道航行长期进行挖泥工作，水深值较大，最深处达14.72 m，而海域中东部大辽河河口、北部辽河河口附近等地区较浅，河流入海导致了淤泥堆积，水深值最浅处为1.17 m。水深反演结果与该地区实际水深情况基本一致，因此，本文方法可适用于该地区近海水深的大范围、高精度反演工作。为验证本文方法在辽东湾近海水深反演中的正确性（图7），本文选择部分非研究时段的辽东湾近海水深数据，通过潮差改正将其改正为研究时刻辽东湾水深值。基于上述实测水

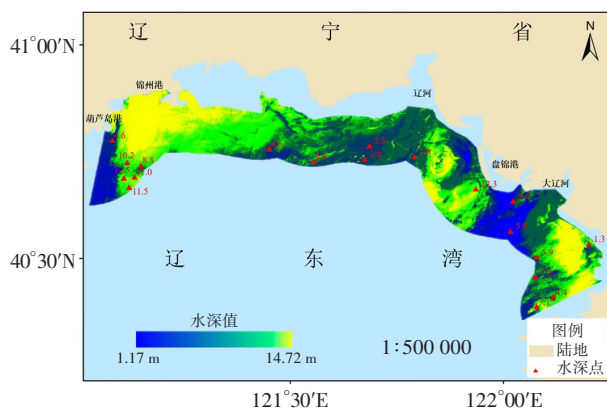


图7 辽东湾部分海域水深反演结果图

深值对本文提出的结合贝叶斯特征选择的近海水深反演方法进行精度评价，如图8所示。

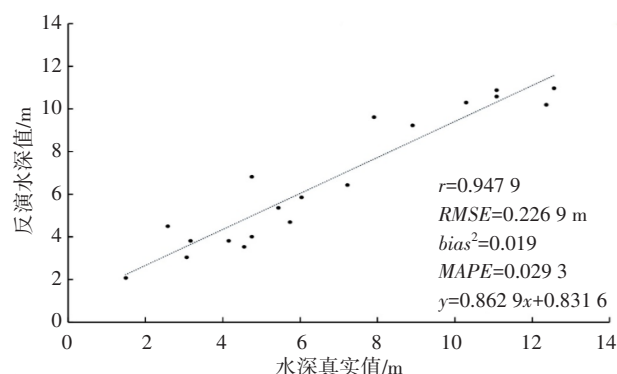


图8 本文方法反演水深外符合精度示意图

3 结语

(1) 本文根据辽东湾2022年4月Landsat 9多光谱影像与同期实测水深数据构建了适用于机器学习水深反演的大规模近海水深反演数据集。基于该数据集结合贝叶斯特征选择和地理空间加权对辽东湾部分海域水深进行反演。

(2) 通过贝叶斯特征选择，将更适合于研究区近海水深反演的影像波段及波段比值输入到地理加权模块之中，一定程度上突出了适合于近海水深反演的波段及波段比值特征，减少了因数据误差过大导致反演精度难以提升的问题。

(3) 由于水深数据存在空间连续性，因此本文基于地理加权回归思想构建地理空间加权模块，将由贝叶斯特征选择后的波段及波段比值分别输入该模块。并将输出的各个特征与实测水深值之间的皮尔森相关系数作为权重，进一步突出数据特征，使输入反演模型的自变量与实测水深数据之间的相关性加强。

(4) 本文提出的改进LightGBM近海水深反演模型输入由13个波段及波段比值特征和各训练样本的地理空间信息组成。经贝叶斯特征选择，B1、B2、B3、B5、B1/B2为模型最佳输入特征组合。并基于地理空间加权模块对上述5个输入特征赋予不同的波段权重实现研究区近海水深的大范围、高精度反演。本文基于LightGBM模型与本文方法在由200个不同空间位置的测试集上进行水深反演实验并制作辽东湾部分海域水深专题图。本文

反演方法的皮尔森相关系数 (r) 为 0.946, 均方根误差 ($RMSE$) 为 0.265 m, 偏差 ($bias^2$) 为 0.017, 平均绝对百分比误差 ($MAPE$) 为 0.031, 均优于 LightGBM 模型。结果表明本文方法是一种适合于大范围、高精度的近海水深反演方法。但由于本文使用影像空间分辨率较低会产生一定程度的反演误差, 在后续研究中, 拟使用高分辨率遥感影像进一步提升反演精度。

参考文献

- [1] 陈义兰, 唐秋华, 刘晓瑜, 等. 多源水深数据融合的近海数字水深模型构建[J]. 海洋科学进展, 2021, 39(3): 461–469.
- [2] 段福楼, 邹庆军, 王玉峰, 等. 近海区域多波束水深测量的水位改正方法[J]. 海洋测绘, 2012, 32(4): 44–46+70.
- [3] LIU Y, GAO X, WANG G, et al. A method of water depth inversion in coastal area considering temperature information[J]. Photogramm Remote Sens Spatial Inf Sci, 2021, 3: 2–28.
- [4] ZOU K, LANGLO L T, RONHOLT G, et al. Full-waveform inversion in a shallow water environment: A North Sea 3D towed-streamer data example[C]// SEG Technical Program Expanded Abstracts 2013, 2013.
- [5] 王燕茹, 张利勇, 刘文, 等. 基于高空间分辨率遥感影像的水深反演有效性评估[J]. 海洋学报, 2023, 45(3): 136–146.
- [6] 楚森森, 程亮, 程俭, 等. 基于光谱分层的浅海水深遥感反演方法[J]. 海洋学报, 2023, 45(1): 125–137.
- [7] ALI W H, BHABRA M S, LERMUSIAUX P F J, et al. Stochastic oceanographic-acoustic prediction and Bayesian inversion for wide area ocean floor mapping[C]//OCEANS 2019 MTS/IEEE SEATTLE. IEEE, 2019: 1–10.
- [8] HUANG R, YU K, WANG Y, et al. Bathymetry of the coral reefs of Weizhou Island based on multispectral satellite images[J]. Remote Sensing, 2017, 9(7): 750.
- [9] 迟昊宇, 郭宝峰, 徐文结, 等. 一种改进的半分析模型的高光谱遥感水深反演方法[J]. 光电子·激光, 2022, 33(12): 1296–1305.
- [10] 沈蔚, 孟然, 栾奎峰, 等. 4种遥感水深反演机器学习模型比较[J]. 海洋测绘, 2022, 42(5): 68–72.
- [11] SHAH A, DESHMUKH B, SINHA L K. A review of approaches for water depth estimation with multispectral data[J]. World Water Policy, 2020, 6(1): 152–167.
- [12] KUTSER T, HEDLEY J, GIARDINO C, et al. Remote sensing of shallow waters – A 50 year retrospective and future directions[J]. Remote Sensing of Environment, 2020, 240: 111619.
- [13] 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[J]. 激光与光电子学进展, 2023, 60(10): 405–413.
- [14] 林位衡, 黄文骞, 李广会, 等. 基于粒子群优化神经网络的水深反演模型[J]. 海洋测绘, 2020, 40(5): 26–29.
- [15] DAI J, LIU T, ZHAO Y, et al. Remote sensing inversion of the Zabuye Salt Lake in Tibet, China using LightGBM algorithm[J]. Frontiers in Earth Science, 2023, 10: 1022280.
- [16] TAN C J, LIM C P, CHEAH Y N. A multi-objective evolutionary algorithm-based ensemble optimizer for feature selection and classification with neural network models[J]. Neurocomputing, 2014, 125: 217–228.
- [17] 王彬, 刘修泽, 王小林, 等. 基于双频识别声纳(DIDSON)的辽东湾近海沙海蜇监测的应用研究[J]. 中国水产科学, 2022, 29(8): 1223–1235.
- [18] 柯丽娜, 徐佳慧, 王楠, 等. 基于遥感生态指数的滨海湿地生态质量变化评价——以辽东湾北部区为例[J]. 生态环境学报, 2022, 31(7): 1417–1424.
- [19] 李东辉, 晁雷, 赵丰泽, 等. 潮汐驱动下的辽东湾水动力及入海污染物输移特征[J]. 海洋科学进展, 2022, 40(3): 423–434.
- [20] 魏英杰, 朱兰艳. 高分辨率遥感影像变化检测的频域分析预处理方法[J]. 测绘学报, 2017, 46(6): 743–752.
- [21] 常伟纲. WorldView-2 立体像对浅海水深反演[D]. 青岛: 山东科技大学, 2020.
- [22] XIE Q, WANG J, LOPEZSANCHE J M, et al. Crop Height Estimation of Corn from Multi-Year RADARSAT-2 Polarimetric Observables Using Machine Learning[J]. Remote Sensing, 2021, 13(3): 392.
- [23] RSAMUSSE C E, WILLIAMS C K I. Gaussian processes in machine learning[J]. Lecture notes in computer science, 2004, 3176: 63–71.
- [24] OUPARDIN A, MICHELE M D, RAUCOULES D, et al. Water depth inversion from satellite dataset[C]//2014 IEEE Geoscience and Remote Sensing Symposium. IEEE, 2014: 2277–2280.
- [25] 张源榆, 黄荣永, 余克服, 等. 基于卫星高光谱遥感影像的浅海水深反演方法[J]. 地球信息科学学报, 2020, 22(7): 1567–1577.
- [26] 彭新月, 张吴明, 钟若飞. 改进 YOLOv3 模型的 GF-2 卫星影像车辆检测[J]. 测绘科学, 2021, 46(12): 147–154.
- [27] 朱洪波, 张在岩, 秦育罗, 等. 农村路面多类型病害检测方法研究[J]. 测绘科学, 2022, 47(9): 170–180.

(本文编辑: 袁泽轶)