基于自动监测和 Sentinel-2 影像的钦州湾溶解氧反演模型研究*

钟炜萍1,2,姚焕玫1**,陈华权1

(1. 广西大学资源环境与材料学院,广西南宁 530004;2. 广西壮族自治区海洋环境监测中心站,广西北海 536000)

摘要:钦州湾是广西近岸海域水质较差的区域,为及时了解钦州湾海域水质状况,本文借助遥感水质监测技术 范围广、快速、连续、可视化程度高的优势,以高分辨率遥感卫星 Sentinel-2 影像数据为数据源,结合广西近岸 海域自动监测数据,通过一元与多元,线性与非线性的回归分析方法建立钦州湾溶解氧浓度反演模型。研究表 明,在构建的 370 个波段组合中,最佳波段组合分别是 1/B3、lnB3/(lnB1+lnB2)和 B3/(B1+B2),其 Pearson 相关系数(*R*)分别为 0.905 2,-0.897 0 和-0.889 2。钦州湾溶解氧浓度反演模型中,逐步回归模型拥有最 低的平均相对误差 MRE (6.47%),最低的均方根误差 RMSE (0.584 5),同时模型验证精度 *R*² 为 0.654 3,稳 定性较佳。本文突破了广西近岸海域传统监测的局限性,同时为钦州湾溶解氧遥感监测提供参考。 关键词:钦州湾 Sentinel-2 溶解氧 反演模型 自动监测 卫星遥感

中图分类号:X834 文献标识码:A 文章编号:1002-7378(2020)04-0392-07 DOI:10.13657/j. cnki. gxkxyxb. 20210114. 001

0 引言

传统的海洋环境监测工作方式主要是通过船只 走航,依靠人工采集样品后在实验室进行分析,最后 根据分析结果来判定海洋环境质量。近几十年来,随 着遥感技术的快速发展,以及传感技术和无线数据传 输技术的进步,现代海洋环境监测已经进入立体化的 新时代。要实现海洋环境监测立体化,除不需要大量 人力物力就可在海洋环境恶劣的情况下进行有效工 作,并具有自动、长期、连续、实时收集海洋环境监测 数据的自动监测浮标外,遥感是另一种重要的技术手 段。遥感技术具有快速、连续、范围广、可视化程度高 等特点,在大尺度环境状况和非点源污染监测方面具 有明显优势^[1]。Hakvoort等^[2]使用机载成像光谱数 据结合矩阵反演技术,对 Veluwe 湖泊的有色可溶性 有机物(CDOM)、叶绿素 a 等水质参数进行反演,获 得较好的效果。阎福礼等^[3]使用 MODIS 数据对富 营养化程度较高的太湖进行水质参数反演。郑炎辉 等^[4]利用 GF-1 WFV 遥感影像和 BP 神经网络模型, 建立广东东风水库叶绿素 a、透明度、总磷、总氮等水 质参数反演模型。Pahlevan等^[5]以美国路易斯安那 州 Grand Lake 为研究区域,使用 Sentinel - 2 和 Landsat-8 卫星数据产品进行遥感反射率的定量和定 性验证,反演水质参数总悬浮物浓度。崔文君等^[6]基 于 Landsat-8 OLI 卫星遥感数据,发现 Band 3 和

【作者简介】

- 钟炜萍(1987-),女,硕士,工程师,主要从事海洋自动监测、环境遥感研究,E-mail:569791151@qq.com。
- 【**通信作者】

^{*}广西科技基地和人才专项(桂科 AD17129041)资助。

姚焕玫(1977一),女,博士,副教授,主要从事环境遥感研究,E-mail;yaohuanmei@gxu.edu.cn。

[【]引用本文】

钟炜萍,姚焕玫,陈华权.基于自动监测和 Sentinel-2 影像的软州湾溶解氧反演模型研究[J].广西科学院学报,2020,36(4):392-398.

ZHONG W P.YAO H M.CHEN H Q. Research on Dissolved Oxygen Inversion Model of Qinzhou Bay Based on Automatic Monitoring Data and Sentinel-2 Data [J]. Journal of Guangxi Academy of Sciences, 2020, 36(4), 392-398.

Band 4 是反演珠江口近岸海域溶解氧的最佳波段组合,利用实测光谱一阶微分值构建模型得出的平均相 对误差为 19.2%。王冰等^[7]使用 Landsat TM 影像 结合实测数据获得呼伦湖溶解氧敏感波段,建立反演 模型,约半数点位相对误差超过 10%。总体来说,针 对淡水湖泊叶绿素 a、营养盐类等水质参数反演的卫 星遥感研究较多,而对海水中溶解氧的遥感反演研究 较少,且已有研究主要使用 Landsat 系列卫星影像数 据,分辨率较低,且得到的模型精度一般。

溶解氧(Dissolved Oxygen, DO)指溶解在水中 的分子态氧,溶解氧可能会因为藻类的生长而过饱 和,当水体受有机或无机还原性物质污染时,溶解氧 会降低[8]。广西钦州湾位于北部湾北部,由内湾(茅 尾海)和外湾(钦州湾)构成,中间狭窄,两端宽阔,东、 西、北三面为陆地所环绕,南面与北部湾相通,是一个 半封闭型的天然海湾^[9]。近十年来,钦州湾一直是北 部湾近岸海域中污染最严重的海区之一,尤其是钦州 湾内湾茅尾海海域,局部海域常年为第四类或劣四类 海水水质^[10]。同时,溶解氧作为赤潮发生前后重要 的水质参数之一[11],对其展开研究在海洋生态方面 具有重要意义。目前,虽然广西在近岸海域海洋自动 监测上已有一定的基础,但是缺少以遥感技术为手 段,对广西近岸海域水质进行大面积监测的研究工 作。本文利用遥感手段,基于自动监测数据,建立钦 州湾的溶解氧反演模型,为钦州湾海域的合理开发、 科学管理提供技术支撑。

1 材料与方法

1.1 研究区域与数据获取

研究区域位于广西北部湾北部钦州湾,即 21°33′36″—21°55′12″E,108°24′36″—108°43′48″ N^[12]。溶解氧实测数据来自广西近岸海域水质自动 监测网络布设在钦州湾的3个自动监测站位(Z1-Z3,见图1)。监测仪器为YSI6600V2多参数水质分 析仪,可每0.5h出具一次溶解氧监测结果,高频率 的溶解氧实测数据可在较大程度内与Sentinel-2卫 星影像的成像时间接近。

Sentinel-2 系列卫星中的 Sentinel-2 A 于 2015 年 6 月 23 日发射,携带一枚多光谱成像仪,覆盖 13 个光谱波段,卫星传感器扫描宽度达 290 km,光谱范 围为 0.4—2.4 μ m。其在可见光、近红外到短波红外 具有不同的空间分辨率,分别为 10 m (4 个波段),20 m (6 个波段)和 60 m (3 个波段),其中可见光波段 及近红外波段空间分辨率为 10 m,可用于观测地球 土地覆盖变化,森林、湖水及近海水域污染情况^[12]。 本文选取 2017—2019 年研究区域内云量<30%的影 像作为影像数据源。



Fig. 1 Location of study area and automatic monitoring stations

1.2 影像数据预处理

采用 SNAP 软件中附带的插件 Sen2cor v2.8.0 对影像进行辐射定标和大气校正。校正完成之后,对 影像进行重采样,将蓝、绿、红、近红外(宽)4 幅 10 m 分辨率的单波段影像导出为 ENVI 格式,并以 Layer Stacking 工具重新合成为一幅影像。为方便影像数 据的进一步处理和专注于研究区的影像数据分析,对 研究区域进行裁剪并提取监测点位的波段数据。影 像预处理流程图如图 2 所示。

1.3 构建波段组合

本文以溶解氧浓度(DO)作为溶解氧反演模型的 因变量,以波段组合作为自变量,通过建立数学模型 设法找出最能代表溶解氧浓度与波段组合之间关系 的数学表达式。Sentinel-2 搭载的传感器上有4种分 辨率为10 m的波段,为表述方便,本文以 B1—B4 来 分别表示 Sentinel-2 传感器的蓝、绿、红、近红外波段 (表1)。将自动监测站位所在像元的4个波段反射 率值导出,通过相加、相比、取倒数、取对数等方式进 行波段组合的构建。



Fig. 2 Sentinel-2 image preprocessing flow

表1 Sentinel-2 产品的4种波段

Table 1 Four bands in Sentinel-2 products

编号 No.	波段 Band	中心波长 Central wavelength (nm)	分辨率 Resolution (m)	带宽 Bandwidth (nm)
B1	蓝光 Blue light	490	10	65
B2	绿光 Green light	560	10	35
B3	红光 Red light	665	10	30
B4	近红外光(宽) Near infrared light (wide)	842	10	115

1.4 Pearson 相关性分析

Pearson 相关系数(*R*)是一种线性相关系数,反 映 2 个变量间的线性相关程度^[13],可用来分析不同 波段组合与自动监测站位溶解氧实测值的相关程度, 从中筛选出敏感程度最高的波段组合。其公式如下:

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

式中,R 表示自变量 x 和因变量 y 的相关系数,n 为 样本量, x_i 、 y_i 分别为 x 和 y 的第 i 个样本值, \overline{x} 、 \overline{y} 分别为 x 和 y 的平均值。R>0,说明 2 个变量具有 正相关关系;R<0,说明 2 个变量呈负相关关系。通 常来说,波段组合与溶解氧浓度的 |R| 越高,用该波 段组合来构建的反演模型的模型精度越高。

1.5 回归分析

回归分析是指确定自变量与因变量间相互依赖 的定量关系的统计分析过程。虽然自变量 *x* 和因变 量 *y* 之间可能没有严格、确定的函数关系,但可以通 过回归分析建立起变量之间的相关关系近似数学表达式^[14]。

(1)一元线性回归:采用相关性最高的波段组合 依次作为自变量 *x* 来进行一元线性回归。

(2)一元非线性回归:采用相关性最高的波段组合的对数值依次作为自变量 x 来进行回归建模,筛选出模型精度最高的模型。

(3)二元线性回归:以相关性最高的两组波段组 合共同作为自变量 *x*。

(4) 三元非线性回归: 以相关性最高的三组波段 组合共同作为自变量 *x*。

4 种回归模型均以溶解氧浓度实测值作为因变量 y,最后筛选出显著性较强、验证精度最高的回归 模型。

1.6 逐步回归分析

逐步回归分析是一个删除不显著因素、保留显著 因素的过程。逐步回归分析在引入和剔除变量的过 程中可以保证最后保留的变量是互相独立、无多重共 线性的,非常适用于多个自变量的多元回归模型构 建^[15]。基于 Sentinel-2 影像数据构建的波段组合和 溶解氧实测数据进行逐步回归分析,剔除不显著波段 组合,最终得到的逐步回归模型。

1.7 模型精度验证

反演模型的验证非常重要,要判断模型是否稳定 可靠,必须从多个角度来对反演模型的精度进行验 证。本研究采用决定系数(R²)、平均相对误差 (MRE)、均方根误差(RMSE)进行精度验证。

(1)*R*²:将实测值与模型输出值进行线性拟合, 线性拟合的*R*² 值越大,表明模型输出值与实测值之 间的相关性越高。值得注意的是,相关性高并不意味 着模型输出值与实测值相近,因此不能仅凭*R*² 值来 判断模型精度。

(2) MRE:表明模型输出值与实测值的偏离程度。计算公式如下:

MRE(%) =

(3)RMSE:亦称标准误差,均方根误差对一组测量中的特大或特小误差非常敏感,所以能够很好地反映测量的精密度。计算公式如下:

$$RMSE = \sqrt{\frac{\sum (_{x}) @ (_{x}) @$$

2 结果与分析

2.1 最佳波段组合

本文共获得 370 种波段组合,以下选取了 Pearson 相关性(*R*)较高的 24 个波段组合(表 2),研究区 表 2 相关性较高的波段组合

Table 2 Band combinations with high correlation

域中有 3 个组合波段与溶解氧浓度相关性效果较好, 分别是 1/B3、lnB3/(lnB1+lnB2)和 B3/(B1+B2), 其 Pearson 相关系数(R)分别为 0.905 2,-0.897 0 和-0.889 2,已基本满足钦州湾溶解氧浓度反演的 需求。

序号	正相关 Positive correlation		序号	负相关 Negative correlation	
No.	波段组合 Band combinations	R	No.	波段组合 Band combinations	R
1	1/B3	0.905 2	13	lnB3/(lnB1+lnB2)	-0.897 0
2	$\ln B1/(\ln B1 + \ln B3)$	0.882 3	14	B3/(B1+B2)	-0.8892
3	$\ln B2/(\ln B2 + \ln B3)$	0.881 0	15	$\ln[B3/(B1+B2)]$	-0.8882
4	lnB1/lnB3	0.880 9	16	B3/B2-B1/B3	-0.886 0
5	lnB2/lnB3	0.880 4	17	lnB3/lnB1	-0.883 4
6	$(1/B3 - 1/B2) \times B1$	0.876 3	18	$\ln B3/(\ln B1 + \ln B3)$	-0.8823
7	$\ln[B1/(B1+B3)]$	0.874 8	19	lnB3/lnB2	-0.8810
8	B1/(B1+B3)	0.873 6	20	$\ln B3/(\ln B2 + \ln B3)$	-0.881 0
9	ln(B1/B3)	0.873 0	21	$B3 \times B3/(B1 \times B2)$	-0.880 5
10	lnB1-lnB3	0.873 0	22	B3/B1 - B2/B3	-0.8790
11	B2/B1 + B1/B3	0.871 3	23	B1/B2-B1/B3	-0.8763
12	B1/B2 + B2/B3	0.869 3	24	$(1/B2 - 1/B3) \times B1$	-0.876 3

2.2 建模结果

在相关性分析结果的基础之上,通过筛选得到 48个影像质量良好的点位监测数据,随机抽取其中 的40个点位分别进行回归建模尝试,并获得4个回 归反演模型公式:

(1)一元线性回归模型 y = 1100.345x + 4.925, x = 1/B3, $R^2 = 0.837 8$;

(2)一元非线性回归方程模型 $y = 1.5995x + 17.209, x = \ln(B1/B3), R^2 = 0.7529;$

(3) 二元线性回归模型 $y = 603.192x_1 - 29.39x_2 + 19.745, x_1 = 1/B3, x_2 = \ln B3/(\ln B1 + \ln B2), R^2 = 0.8573;$

(4) 三元非线性回归模型 y = 605.842x₁ 表 3 基于 Sentinel-2 数据的溶解氧反演模型精度验证

27. $8376x_2 - 0.3227x_3 + 19.12, x_1 = 1/B3, x_2 = \ln B3/(\ln B1 + \ln B2), x_3 = B3/(B1 + B2), R^2 = 0.8573;$

以及逐步回归分析模型: $y = 771.854x_1 - 1.476x_2 + 6.435, x_1 = 1/B3, x_2 = B3 × B3/(B1 × B1), R^2 = 0.863。$

随机抽取剩余的 8 个点位对以上 5 个模型公式 进行精度验证,结果如表 3 所示。通过对 5 种回归模 型的综合比较可知(表 3),逐步回归模型的决定系 数、平均相对误差和均方根误差是 5 种模型中最小 的。平均相对误差和均方根误差越小,说明模型输出 结果与实际值更贴近,反演模型的稳定性也更有 保证。

Table 3 Accuracy verification of dissolved oxygen inversion model based on Sentinel-2 data

序号 No.	模型公式 Model	决定系数 R^2	平均相对误差 MRE(%)	均方根误差 RMSE
1	y = 1100.345x + 4.925	0.6267	7.56	0.584 5
2	y = 1.5995x + 17.209	0.600 1	8.39	0.611 7
3	$y = 603.192x_1 - 29.39x_2 + 19.745$	0.639 5	8.21	0.617 5
4	$y = 605.842x_1 - 27.8376x_2 - 0.3227x_3 + 19.12$	0.639 5	8.21	0.616 5
5	$y = 771.854x_1 - 1.476x_2 + 6.435$	0.654 3	6.47	0.521 1

综上考虑,选择逐步回归模型 $y = 771.854x_1 - 1.476x_2 + 6.435(模型精度 <math>R^2 = 0.863$)作为钦州湾 溶解氧的反演模型,对应的波段组合为 1/B3 和 B3×B3/(B1×B1)。

另外,为避免随机抽取点位进行验证可能带来的 验证结果的偶然性,进一步对逐步回归模型所有点位 的输出值与实测值进行比较分析。由图3可知,实测 值曲线与模型输出值曲线保持了较高的吻合度,在前 30个点位(约占总点位数的2/3)中,模型输出值与实 测值之间的相对误差均低于5%,表明模型的稳定性 较强,展现了较高的模型精度。

2.3 钦州湾溶解氧时空分布特征

通过前面建立的溶解氧反演模型计算,得到基于 Sentinel-2 A 卫星 2017—2019 年遥感数据的钦州湾 溶解氧浓度时空分布结果(图 4)。从反演结果上看, 2017—2019 年钦州湾溶解氧浓度存在明显的时间变 化,按照枯水期(3,4 月)、丰水期(7—9 月)、平水期 (11,12 月)划分,研究区域每年都基本呈现溶解氧浓 度丰水期<枯水期<平水期的规律。按四季划分,钦 州湾溶解氧浓度基本呈现夏季最低、春季适中、秋冬 最高的特点。在空间分布上,2017—2019 年钦州湾 溶解氧浓度均呈现内湾<外湾的特点,且内湾溶解氧 浓度分布不均匀,越靠近河口的位置溶解氧浓度 越低。











3 结论

本研究基于 Sentinel-2 系列卫星遥感影像并结 合自动监测实测溶解氧浓度数据,建立广西钦州湾溶 解氧反演模型,得出 1/B3、lnB3/(lnB1+lnB2)和 B3/ (B1+B2)3个波段组合与溶解氧浓度相关性效果较 好。采用最佳敏感波段和回归模型构建钦州湾的溶 解氧浓度反演模型,逐步回归模型稳定性较佳,模型 精度 R²达到 0.863。通过反演结果得出 2017— 2019年部分时段钦州湾溶解氧浓度分布图,直观显 示钦州湾溶解氧的空间分布状况,可为钦州湾溶解氧 实时遥感监测提供理论依据与技术支撑。

以往水质反演所使用的实测数据大多为卫星过 境当天,通过人工采集样品的方式得到水质监测数 据,而本研究使用的 0.5 h/次的高频率自动监测数 据能在较大限度内与遥感影像数据接近,在很大程度 上提高了模型精度。因此可以预见,自动监测与遥感 技术相结合的方式,在未来会有较好的发展潜力。采 用最佳敏感波段和回归模型方式构建钦州湾的溶解 氧浓度反演模型,优势在于模型简单、易于实现,且最 终得到了较好的反演效果。

本文实现了基于遥感技术对钦州湾溶解氧浓度 的大范围动态监测,得出的模型精度较佳,是一次较 好的尝试。不过,影响钦州湾水环境的重要指标有很 多,在未来的研究中仍需对其他指标进行研究,例如 石油类、营养盐等,使反演模型更加全面和实用。

参考文献

- [1] 陈令新,王巧宁,孙西艳.海洋环境分析监测技术[M]. 北京:科学出版社,2018.
- [2] HAKVOORT H, DE-HAAN J, JORDANS R, et al. Towards airborne remote sensing of water quality in the Netherlans - validation and error analysis [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2002, 57(3):171-183.
- [3] 阎福礼,林亚森,王世新,等.太湖富营养化水体比辐射 率测量及 MODIS 水体温度反演应用[J].中国科学院大 学学报,2019,36(6):784-793.
- [4] 郑炎辉,张园波,何艳虎.基于GF-1WFV影像与神经网

络模型的水库水质多参数反演研究[J].人民珠江, 2020,41(7):57-62,84.

- [5] PAHLEVAN N, SARKAR S, FRANZ B A, et al. Sentinel-2 Multi Spectral Instrument (MSI) data processing for aquatic science applications: Demonstrations and validations [J]. Remote Sensing of Environment, 2017, 201:47-56.
- [6] 崔文君,夏丽华,解学通,等.基于实测光谱的珠江口溶 解氧反演模型[J].广州大学学报:自然科学版,2017, 16(6):84-92.
- [7] 王冰,安慧君,吕昌伟.基于多源遥感数据的呼伦湖溶解 氧反演模型[J].生态学杂志,2013,32(4):993-998.
- [8] 国家环境保护总局《水和废水监测分析方法》编委会.水 和废水监测分析方法[M].4版.北京:中国环境科学出版社,2002.

- [9] 张荣灿,姜发军,陈宪云,等.钦州湾近岸海域水质污染 状况评价[J].广西科学,2014,21(4):403-410.
- [10] 蓝文陆,李天深,罗金福.广西钦州湾环境与生态研究 [M].北京:科学出版社,2020.
- [11] 韦蔓新,何本茂,赖廷和.廉州湾赤潮形成期间 pH 值 和溶解氧的时空分布及其与环境因素的关系[J].广西 科学,2004,11(3):221-224.
- [12] 中国海湾志编纂委员会.中国海湾志 第十二分册 广西 海湾[M].北京:海洋出版社,1993.
- [13] 龚燃.哨兵-2A光学成像卫星发射升空[J].国际太空, 2015(8):36-40.
- [14] 郭立焕.回归分析方法的原理及应用[J].山西财经学 院学报,1982(4):54-60.
- [15] 胡松伟.变深度发射性能可靠性影响因素多元逐步回 归分析[J]. 兵工学报,2017,38(5):986-994.

Research on Dissolved Oxygen Inversion Model of Qinzhou Bay Based on Automatic Monitoring Data and Sentinel-2 Data

ZHONG Weiping^{1,2}, YAO Huanmei¹, CHEN Huaquan¹

(1. College of Resources, Environment and Materials, Guangxi University, Nanning, Guangxi, 530004, China; 2. Marine Environmental Monitoring Center of Guangxi Zhuang Autonomous Region, Beihai, Guangxi, 536000, China)

Abstract: The Qinzhou Bay is an area with poor water quality in the coastal waters of Guangxi. To understand the status of the seawater quality in Qinzhou Bay in time, a dissolved oxygen concentration inversion model of Qinzhou Bay is necessary. The inversion model was set up with the aid of remote sensing water quality monitoring technology which had advantages of wide rage, fast speed, high continuity and high visual degree, using the high resolution remote sensing satellite Sentinel-2 image data as the data source, and combining with the automatic monitoring data of the coastal waters of Guangxi. The study showed that among the 370 band combinations constructed, the best band combinations were 1/B3, lnB3/(lnB1+lnB2) and B3/(B1+B2), and the Pearson correlation coefficient (R) were 0.905 2, -0.897 0 and -0.889 2, respectively. The inversion model of dissolved oxygen concentration in Qinzhou Bay was constructed based on the best sensitive band and linear and nonlinear regression analysis methods. The stepwise regression model had the lowest mean relative error (MRE, 6, 47%) and the lowest root mean square error (RMSE, 0, 584 5), and the model verification accuracy R^2 was 0.654 3, and it showed better stability. This article breaks through the limitation of traditional monitoring in Guangxi coastal waters and provides a reference for remote sensing monitoring of dissolved oxygen in Qinzhou Bay.

Key words: Qinzhou Bay, Sentinel-2, dissolved oxygen, inversion model, automatic monitoring, remote sensing