

# 基于哨兵 2 数据的 2019–2021 年三亚河水体透明度数据集

ISSN 2096-2223  
CN 11-6035/N



邱瑞婷<sup>1,2,3</sup>, 王胜蕾<sup>1,2,4\*</sup>, 史建康<sup>5</sup>, 李俊生<sup>1,2,4</sup>,  
张方方<sup>2,4</sup>, 张文志<sup>2,4</sup>, 梅悦<sup>1,2,3</sup>

1. 海南空天信息研究院, 海南省地球观测重点实验室, 海南三亚 572029
2. 中国科学院空天信息创新研究院, 数字地球院重点实验室, 北京 100094
3. 中国地质大学(北京), 北京 100049
4. 可持续发展大数据国际研究中心, 北京 100094
5. 海南省生态环境监测中心, 海口 571126

文献 CSTR:  
32001.14.11-6035.noda.2022.0015.zh  
文献 DOI:  
10.11922/11-6035.noda.2022.0015.zh  
数据 DOI:  
10.11922/sciencedb.j00001.00412  
文献分类: 地球科学

收稿日期: 2022-03-21  
开放同评: 2022-4-20  
录用日期: 2022-07-22  
发表日期: 2022-12-21

**摘要:** 水体透明度是环保部门最关心的重要水质参数之一, 反映水体浑浊程度, 对水体初级生产力与水生态系统具有重要影响。作为典型海岛独立水系, 海南省总体地表内陆水体水资源较丰富, 河流径流充沛, 但受干湿季风和地形影响具有时空分布不均的特点, 目前针对海南岛内陆水体水质研究相对较少。本文以海南省三亚市三亚河为研究区, 基于 GEE 云计算平台和 Google 云存储的哨兵 2 地表反射率数据, 采用基于 QA<sub>AV</sub>6 的透明度半解析模型反演 2019–2021 年三亚河时间序列透明度。对于三亚河水体动态提取, 采用归一化水体指数 NDWI 与 OTUS 自动化阈值分割相结合的算法提取细小河流水体。数据以 GeoTiff 栅格格式存储, 同时保存了像元透明度值和坐标信息, 便于相关的地理信息系统软件读取和分析。基于 GEE 云数据库反演水体长时序透明度, 具有高效快速的优势。本数据集可以为三亚河水质监测、水体污染治理及水生态保护等提供重要科学依据。

**关键词:** 三亚河; 哨兵 2; 透明度; QA<sub>AV</sub> 半解析模型

## 数据库(集)基本信息简介

数据库(集)名称	基于哨兵 2 数据的 2019–2021 年三亚河水体透明度数据集
数据作者	邱瑞婷、王胜蕾、史建康、李俊生、张方方、张文志、梅悦
数据通信作者	王胜蕾 (wangsl@radi.ac.cn)
数据时间范围	2019–2021 年
地理区域	地理范围为 18° 09' 34" – 18° 37' 27" N、108° 56' 30" – 109° 48' 28" E, 三亚河水体的空间分布
空间分辨率	10 m
数据量	99.2 MB
数据格式	*.tif、*.csv
数据服务系统网址	<a href="http://doi.org/10.11922/sciencedb.j00001.00412">http://doi.org/10.11922/sciencedb.j00001.00412</a>

\* 论文通信作者  
王胜蕾 wangsl@radi.ac.cn

基金项目	海南省重大科技计划 (ZDKJ2019006) ; 国家自然科学基金 (41901272) ; 中国科学院先导专项 (XDA19080304) 。
数据库 (集) 组成	本数据包括 2019-2021 年近 3 年三亚河透明度的月均值和季度均值的时空分布栅格数据集。其中月均值透明度数据共 36 期, 季度透明度数据每隔 3 个月 1 期, 总共 12 期; 以及日均值、月均值、季度均值透明度的统计数据。所有数据保存为一个压缩文件 (2019-2021 年三亚河透明度长时序变化数据.zip) 。

## 引言

三亚河位于海南省三亚市境南部, 是与三亚城市生活最密切相关的河流, 被称为三亚人民的母亲河。然而, 近年来随着三亚城市化进程的加快, 居民、游客数量急剧增加, 污水排河、垃圾堆岸、船舶排污等问题严重影响三亚内河生态环境, 河流水质出现下降, 引起三亚市政府的高度重视。根据海南省 2019 年生态环境质量和城镇内河 (湖) 水质状况月度报告显示, 三亚河城区部分河段水质一度降至地表水 V 类和劣 V 类, 甚至达到重度污染<sup>[1]</sup>。因此加强三亚河水质实时监测, 对于三亚市水环境治理与保护十分重要。

卫星遥感手段可以克服常规观测方法的不足, 具有宏观、大范围、快速、周期性动态获取水体信息的优点, 可以有效监测大范围水体水质的空间分布和动态变化<sup>[2-3]</sup>。卫星遥感数据已成为重要且低成本的大范围地表水质监测数据源, 通过卫星遥感可监测的水体水质参数有很多, 其中比较重要的是水体透明度。透明度是衡量水体清澈程度的重要指标, 同时也是评价湖库富营养化的一个重要指标, 反映了综合的水质状况, 是水环境保护部门最关心的湖库水质参数之一<sup>[4-7]</sup>。内陆河流作为重要的资源和载体, 是生态系统的绿色生命线, 关系到人类的生存和发展。目前, 国内外学者利用卫星遥感手段对水质监测的研究主要集中在区域跨度较大的研究对象上, 例如大型湖泊水库和河口海岸的监测, 对小尺度和微型态环境问题研究相对较少<sup>[8-10]</sup>。

因此, 本研究以三亚河为研究区, 利用时间、空间分辨率较高的遥感影像哨兵 2 MSIL2A 数据, 基于 GEE 云平台生产 2019-2021 年三亚河水体透明度的长时序变化遥感产品, 达到利用遥感技术对三亚河进行长时间序列水质时空变化状况监测的目的。本数据集可以为三亚河水质监测、水体污染控制与治理及水生态保护等提供重要科学依据。

## 1 数据采集和处理方法

### 1.1 数据采集方法

本研究所采用的数据包括两部分: 一部分是实测水体透明度, 一部分是遥感数据。2022 年 1 月 8 日在三亚市三亚河进行了一次实地调查, 实地调查过程中, 在河道中布设了 17 个采样点, 测量了水面遥感反射率 ( $R_{rs}$ ) 光谱、水体透明度 (又称作塞氏盘深度,  $Z_{SD}$ , 单位 m)、浊度等数据。 $Z_{SD}$  的测定是通过将一个黑白的塞氏盘放到水中, 直到该盘不再可见, 记录此时塞氏盘的深度为透明度<sup>[11]</sup>。实地调查采样点分布如图 1 所示。

遥感数据来源于 GEE 平台, 使用 2022 年 1 月 7 日的哨兵 2 MSI Level-2A 数据进行分析。哨兵 2 MSI Level-2A 数据, 是地表反射率数据, 来源于 Google Earth Engine (GEE) 云平台的遥感数据库。哨兵 2 环境监测卫星是欧洲航天局“哥白尼计划”发射的多光谱对地观测卫星, 由哨兵 2A 和哨兵

2B 双星组成, 单星重访周期为 10 天, 双星交替实现 5 天完全覆盖赤道的同一地区, 周期时间较短。哨兵 2 MSI Level-2A 数据是在哨兵 2 Level-1C 数据基础上利用 Sen2cor 大气校正插件进行大气校正后得到的大气底层反射率数据, 对 Level-1C 数据进行了大气、地形和卷云校正, 并且数据空间分辨率均重采样到了 10 m, 覆盖可见光、近红外和短波红外共 12 个光谱波段。哨兵 2 Level-2A 数据分块图像为 100 km×100 km 的 UTM/WGS84 投影的正射图像, 分块图像大小约为 500 MB。

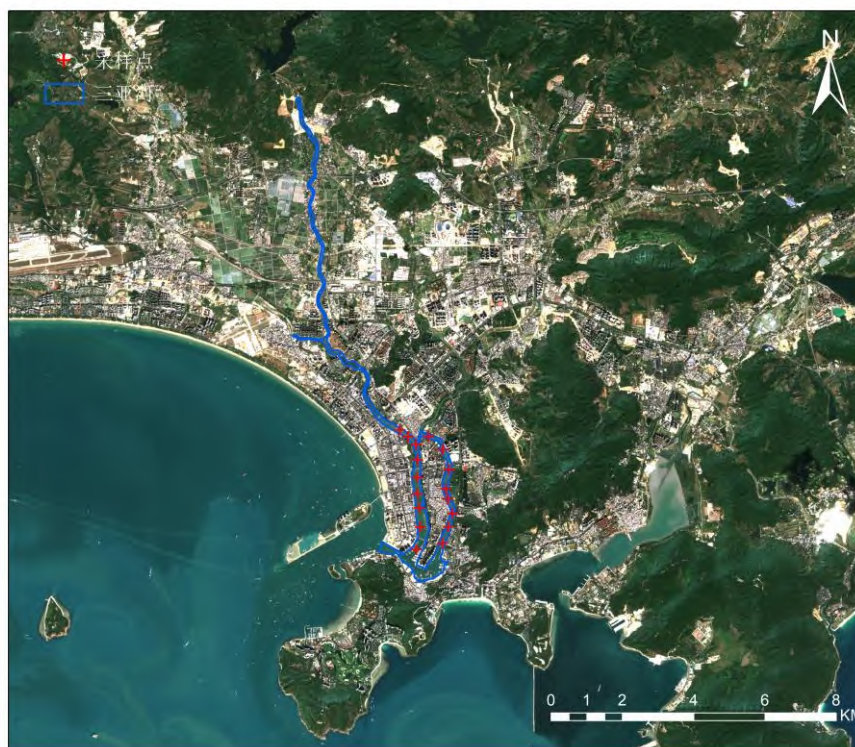


图 1 三亚河实测透明度采样点分布图及采样准同步哨兵 2 真彩色 (B4(R),B3(G),B2(B)) 合成影像

Figure 1 The distribution map of the measured transparency sampling points in Sanya River and the composite image of sampling quasi-synchronized Sentinel-2 true color (B4(R), B3(G), B2(B))

## 1.2 数据处理方法和流程

基于 Google Earth Engine (GEE) 计算云平台上的哨兵 2 MSI Level-2A 数据, 研究生产了海南省三亚市三亚河的月平均透明度数据集。数据生产的具体流程如下图 2 所示: 首先, 基于 GEE 云计算平台, 查找现场实测透明度时空同步的无云高质量哨兵 2 MSI Level-2A; 然后, 构建适用于三亚河的透明度遥感算法, 并利用实测透明度数据集对透明度遥感反演算法进行评估; 最后, 将构建好的遥感算法应用于 GEE 云平台存储的长时序哨兵 2 数据集, 并利用 ArcGIS 等相关软件对透明度月均值产品进行快速分析与制图。其中, 遥感数据处理中涉及的主要步骤包括云掩膜处理、三亚河细小水体提取、透明度遥感算法构建、透明度遥感精度评估、季度平均透明度计算等。

### 1.2.1 水体提取

本文采用归一化水体指数 NDWI (Normalized Difference Water Index) [12]和 OTSU 自动化阈值分割[13]相结合的算法, 实现面向水体的自动化阈值分割提取水体。通常情况, 基于水体指数提取水体的方法是通过计算最优阈值或人为选取经验阈值来区分每景遥感影像的水体与非水体部分。本文根

据水体的一般矢量边界向外扩展 1.5 倍面积，在此范围内确定阈值。针对大量长时序遥感数据，手动确定阈值需要耗费大量时间，采用统一的经验阈值又无法使每景影像的水体提取精度达到最佳。因此，本文结合 OTSU 算法对 NDWI 进行双峰阈值分割，实现逐景影像水体阈值自动化确定，更加高效、精确、自动化地区分每景遥感影像的水体与非水体部分。OTSU 算法又称最大类间方差法，是在最小二乘原理的基础上推导而来的。其基本思想是基于影像的直方图计算各灰度级的发生概率，并以某一阈值变量  $t$  将构成图像的所有像素分为两类，然后求取每一类的类间方差，选取使得两组类间方差最大时的  $t$  值，作为二值化处理的最佳阈值<sup>[4]</sup>。该算法能根据不同的影像的直方图自适应地确定每景影像的最佳阈值。NDWI 计算公式如下：

$$NDWI = \frac{B8 - B3}{B8 + B3} \quad (1)$$

式中的 B8、B3 分别表示哨兵 2 MSI Level-2A 数据第 8 波段（近红外）和第 3 波段（绿波段）的地表反射率。

此外，为排除水体提取中的岸边地物的影响和云干扰，对水体提取结果向内腐蚀两个像元，剔除岸边地物的干扰并利用 GEE 官方提供的哨兵 2 云掩膜算法对 2019–2021 年三亚河数据进行云掩膜处理，该算法利用 QA60 波段来标记是否有云。

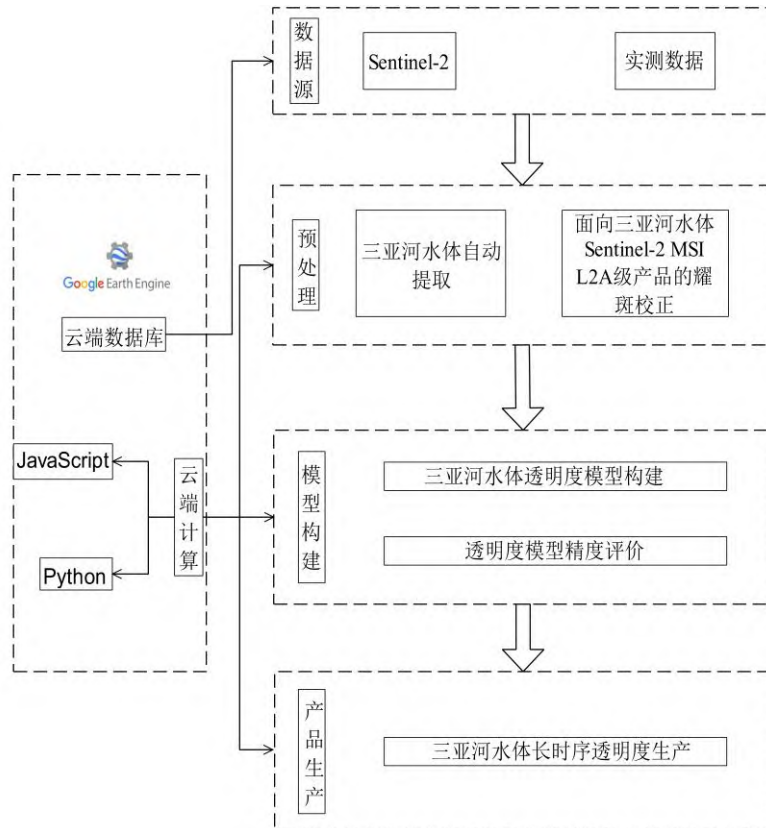


图 2 基于 GEE 平台和哨兵 2 (Sentinel-2) 数据的三亚河透明度反演流程图

Figure 2 Flow chart of Sanya River transparency inversion based on GEE platform and Sentinel-2 data

为评价水体提取精度，在 GEE 上随机选点通过人工目视解译识别水体，并将其作为水体的真值，然后与基于 NDWI 和 OTSU 的算法提取的水体进行对比，最终构建混淆矩阵，并统计水体提取精度。结果表明，基于 NDWI 和 OTSU 的水体提取算法的用户精度为 97.4%，生产者精度为 99.3%，总体精度为 98.3%，Kappa 系数为 0.97，提取效果相对较好。

### 1.2.2 离水反射校正

由于海南哨兵 2 地表反射率影像数据没有对水面天空光反射与耀斑进行校正，因此需要进一步的数据校正，并将其转化为离水反射率  $R_{rs}(\lambda)$ ，单位  $\text{sr}^{-1}$ 。本研究采用基于像素的校正方法<sup>[15-16]</sup>，即在像元每个波段值减去近红（NIR）短波红外（SWIR）波段的最小值，以去除水面天空光反射和太阳耀斑对反射率数据的影响。校正公式如下：

$$R_{rs}(\lambda) = \frac{R(\lambda) - \min(R_{NIR} : R_{SWIR})}{\pi} \quad (2)$$

其中  $R(\lambda)$  表示哨兵 2 波段的反射率值， $\min(R_{NIR} : R_{SWIR})$  表示哨兵 2 近红外和短波红外波段的反射率最小值。

### 1.2.3 透明度反演方法

在过去的几十年里，随着遥感技术的发展，国内外学者已开发许多水体透明度遥感算法，主要包括两种：经验算法、半解析算法<sup>[2]</sup>。本研究采用基于 QAAv6 的最新透明度半解析模型反演三亚河透明度<sup>[4]</sup>。该透明度半解析模型改进了  $Z_{SD}$  物理解释模型，基于水体最佳穿透窗口的漫衰减系数  $K_d$ （单位  $\text{m}^{-1}$ ）求解  $Z_{SD}$ ；而  $K_d$  可以通过水体吸收系数  $a$ （单位  $\text{m}^{-1}$ ）和后向散射系数  $b_b$ （单位  $\text{m}^{-1}$ ）计算得到<sup>[17]</sup>； $a$  和  $b_b$  则基于面向较浑浊 II 类水体的 QAAv6 算法通过水体离水反射率计算得到<sup>[18-19]</sup>。

首先，根据 QAAv6 算法，通过离水反射率  $R_{rs}(\lambda)$  计算水体吸收系数  $a$  和后向散射系数  $b_b$ ：

$$r_{rs}(\lambda) = R_{rs}(\lambda) / (0.52 + 1.7R_{rs}(\lambda)) \quad (3)$$

$$r_{rs}(\lambda) = g_0 u(\lambda) + g_1 u^2(\lambda) \quad (4)$$

$$u(\lambda) = \frac{b_b(\lambda)}{\alpha(\lambda) + b_b(\lambda)} \quad (5)$$

式中， $r_{rs}(\lambda)$  为水面以下遥感反射率，单位  $\text{sr}^{-1}$ ； $g_0$  和  $g_1$  是模型常数； $u$  为 Gordon 参量。根据 QAAv6 算法，参考波长处的水体吸收系数  $a(\lambda_0)$  可以通过离水反射率  $R_{rs}(\lambda)$  或水面以下遥感反射率  $r_{rs}(\lambda)$  的经验模型计算得到。其中，当  $R_{rs}(665) < 0.0015 \text{ sr}^{-1}$  时，在 665 nm 红波段的水体吸收系数  $a(665)$  估算公式为：

$$\chi = \log \left( \frac{r_{rs}(443) + r_{rs}(490)}{r_{rs}(560) + 5 \frac{r_{rs}(670)}{r_{rs}(490)} r_{rs}(670)} \right) \quad (6)$$

$$\alpha(\lambda_0) = \alpha(665) = \alpha_w(665) + 10^{-1.146 - 1.366\chi - 0.469\chi^2} \quad (7)$$

式中， $r_{rs}(443)$ 、 $r_{rs}(490)$ 、 $r_{rs}(560)$ 、 $r_{rs}(665)$  分别为 443nm、490nm、560nm、665nm 处的水面以下遥感反射率； $\alpha_w(560)$  为 560nm 处的水体吸收系数。

当  $R_{rs}(665) \geq 0.0015 \text{ sr}^{-1}$  时，在 665 nm 红波段的水体吸收系数  $a(665)$  估算公式为：

$$\alpha(\lambda_0) = \alpha(665) = \alpha_w(665) + 0.39 \left( \frac{R_{rs}(665)}{R_{rs}(443) + R_{rs}(490)} \right)^{1.14} \quad (8)$$

式中,  $R_{rs}(443)$ 、 $R_{rs}(490)$ 、 $R_{rs}(665)$ 分别为 443 nm、490 nm、665 nm 处的遥感反射率。

由上式 (3) (4) (5) 可知, 对于给定的离水反射率  $R_{rs}(\lambda)$ 在  $a(\lambda)$ 已知的情况下能求得相应的  $b_b(\lambda)$ :

$$b_b(\lambda) = \frac{u(\lambda)\alpha(\lambda)}{1 - \mu(\lambda)} \quad (9)$$

并且, 水体后向散射系数  $b_b(\lambda)$ 是纯水后向散射系数  $b_{bw}$ 和颗粒物的后向散射系数  $b_{bp}$ 之和, 且  $b_{bp}(\lambda)$ 与波长  $\lambda$ 之间存在幂指数关系:

$$b_{bp}(\lambda_0) = \frac{u(\lambda_0)\alpha(\lambda_0)}{1 - u(\lambda_0)} - b_{bw}(\lambda_0) \quad (10)$$

$$b_{bp}(\lambda) = b_{bp}(\lambda_0)\left(\frac{\lambda_0}{\lambda}\right)^\eta \quad (11)$$

$$\eta = 2.0 \left( 1 - 1.2 \exp\left(-0.9 \frac{r_{rs}(443)}{r_{rs}(560)}\right) \right) \quad (12)$$

因此在获取了参考波长处的  $b_b(\lambda_0)$ 后, 结合上式从影像中得  $u(\lambda)$ 就能推断出其余波段的  $b_b(\lambda)$ 和  $a(\lambda)$ 。

此后, 基于得到的水体吸收系数  $a(\lambda)$ 和后向散射系数  $b_b(\lambda)$ 可以通过下式计算得到各个波段的漫衰减系数  $K_d(\lambda)$ :

$$K_d(\lambda) = (1 + m_0 \times \theta_s) a(\lambda) + (1 - \gamma) \frac{b_{bw}(\lambda)}{b_b(\lambda)} \times m_1 \times (1 - m_2 \times e^{-m_3 \times a(\lambda)}) b_b(\lambda) \quad (13)$$

式 (13) 中,  $m_{0-3}$ 和  $\gamma$ 是模型参数, 值分别为 0.005、4.26、0.52、10.8 和 0.265。这些模型参数是常数, 不随水体的性质或波长而变化, 主要与太阳天顶角和风速有关, 通过模拟数据集求得<sup>[20]</sup>。

$\theta_s$ 是空气中的太阳天顶角。再依据最佳穿透窗口的漫衰减系数  $K_d$ 即各个波段最小漫衰减系数进行透明度半解析反演, 其模型为:

$$Z_{SD} = \frac{1}{2.5 \text{Min}(K_d(443, 490, 560, 665))} \ln\left(\frac{|0.14 - R_{rs}^r|}{0.013}\right) \quad (14)$$

式 (14) 中,  $R_{rs}^r$ 为最小漫衰减系数  $K_d$ 对应波段的遥感反射率。

## 2 数据样本描述

### 2.1 数据组合

本数据集由两部分组成, 栅格数据和统计数据。栅格数据包括 2019–2021 年的三亚河月均透明度空间分布数据, 共 36 期数据, 以及每隔 3 个月一期, 共 12 期三亚河季度透明度的空间分布数据, 数据的空间分辨率为 10 m, 数据存储为 tiff 栅格数据格式, 投影系统为 WGS84 地理坐标投影。统计数据包括 2019–2021 年透明度日均值透明度、月均值透明度、季度均值透明度统计数据, 保存为 csv 格式。将数据保存为 1 个压缩文件, 包含 2019 年、2020 年、2021 年 3 个文件以及统计数据表。

## 2.2 数据样本

首先基于 GEE 云平台得到三亚河 2019–2021 年的月平均透明度数据，2019–2021 年月均值的透明度时序变化曲线如图 3 所示。从曲线中可以看出三亚河平均透明度在年内波动较大，波动范围在不足 0.5 m 到 2 m 之间变化，每年 3 月份透明度较低，从 4 月份到 5 月份透明度逐渐升高，到 6 月份和 7 月份达到高峰，从 8 月份开始逐渐下降，在 10 月份左右又开始上升。4 到 5 月份透明度升高，正值海南开始进入雨季，降水量增多可能导致河流水量增加透明度上升。据相关数据显示，2019 年 5 月份降水量远大于 2020、2021 年 5 月份的降水量，这很可能是 2019 年 4–5 月三亚河透明度比 2020 和 2021 年 4–5 月份透明度上升幅度大的主要原因<sup>[21-23]</sup>。2020–2021 年 8–10 月份台风的影响使得降雨量增大，但是透明度逐渐降低，很可能是由于台风风浪带来的泥沙，使得水体的透明度降低<sup>[24-25]</sup>，而 11 月份台风结束，水体透明度开始逐渐升高。

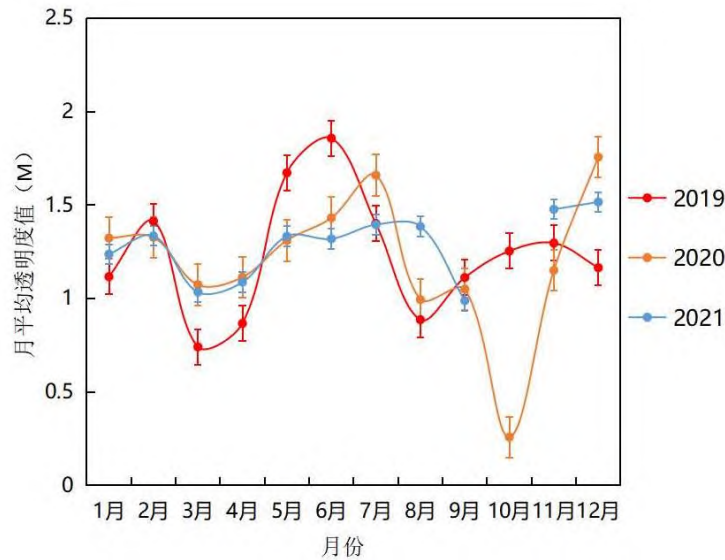


图 3 遥感反演的 2019–2021 年三亚河月平均透明度时间序列（断点处表示该月缺数据）

Figure 3 Time series of monthly average transparency of Sanya River from 2019 to 2021 retrieved by remote sensing (Breakpoints indicate a lack of data in this month)

结合三亚热带干湿季气候，每年 5 月到 10 月为雨季，11 月到次年 4 月为干季，将一年分为四期，第一期为 2–4 月，第二期为 5–7 月，第三期为 8–10 月，第四期为 11 月至次年 1 月。然后基于 GEE 得到三亚河 2019–2021 年的季度平均透明度的栅格数据，总共生成 12 期的透明度数据，然后利用 ArcGIS 等相关软件进行处理制图，得到 2019–2021 年三亚河长时序透明度的空间分布变化数据。从图 4、图 5 可以看出，在 2019–2021 年三亚河的季度平均透明度上有显著的时空差异。在空间上，入港口透明度较高，而上游的透明度较低。在时间上 5–7 月份透明度达到一个峰值。

## 3 数据质量控制和评估

本研究基于采集的实测透明度对同步哨兵 2 数据反演的透明度数据进行精度和可靠性评估。实测透明度数据与遥感反射率数据采用以下策略进行时空匹配：(1)时间窗口设置为实测数据测量时间  $\pm 1d$ 。(2)空间窗口为  $3 \times 3$  像元，取采样点相邻  $3 \times 3$  的像元的中值作为该采样点的反射率<sup>[26]</sup>。精度评价指标包括平均相对误差 MRE (Mean Relative Error)、平均绝对误差 MAE (Mean Absolute Error)、

均方根误差 RMSE (Root Mean Squared Error)、平均无偏相对误差 AURD (Averaged Unbiased Relative Difference), 其计算公式如下所示。基于实测透明度, 与星地同步的模型反演透明度进行比较, 得到透明度反演的平均相对误差 MRE 为 13.6%, 平均绝对误差 MAE 为 0.12, 均方根误差 RMSE 为 0.02 m, 平均无偏相对误差 AURD 为 15.2%, 结果如图 6 所示。

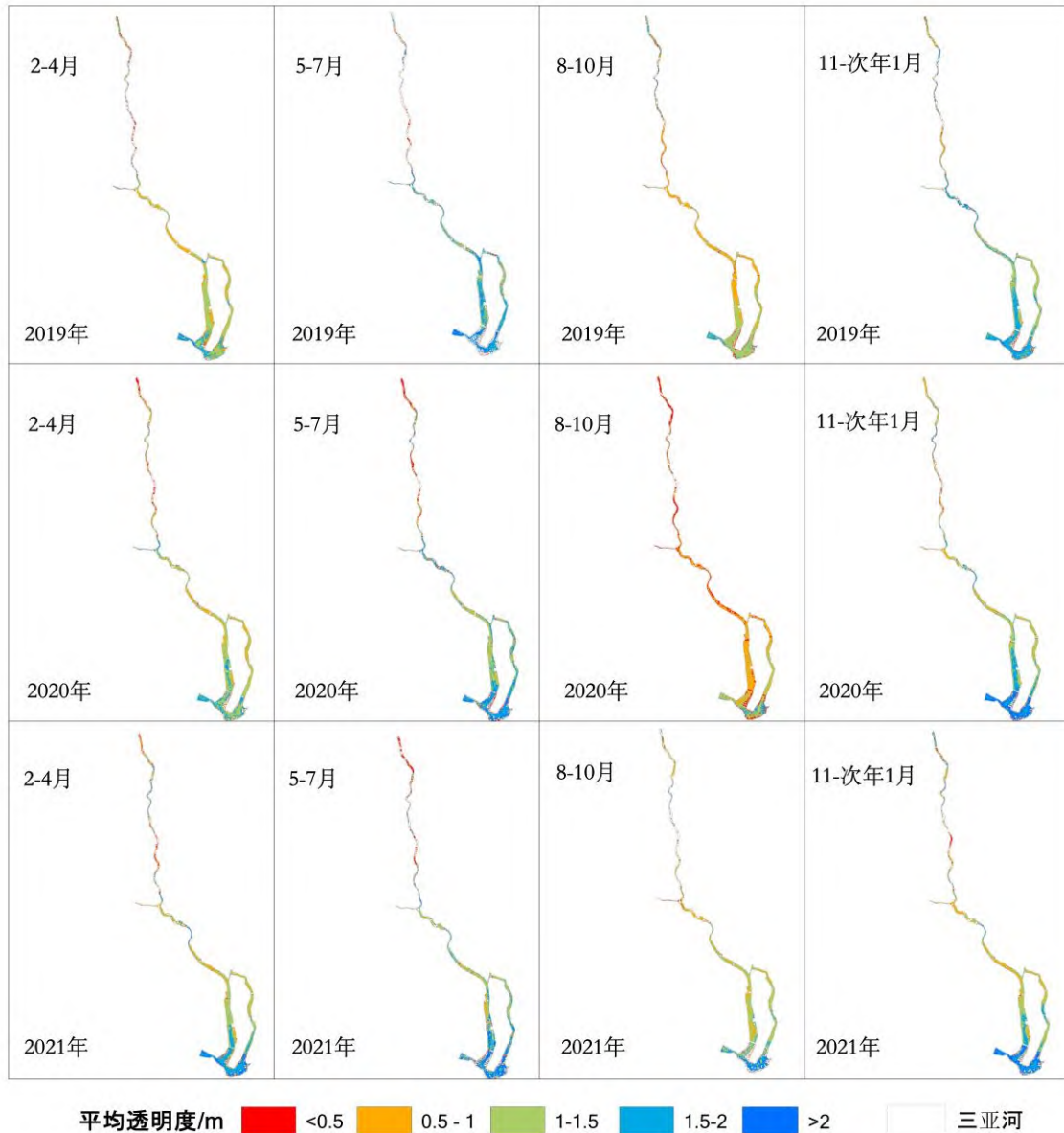


图 4 遥感反演的三亚河 2019–2021 年季度平均透明度 (河流界限范围内空白表示非水体覆盖区域或无有效观测)

Figure 4 Quarterly average transparency of Sanya River from 2019 to 2021 retrieved by remote sensing (Blanks within the river boundary indicate non-water coverage areas or no effective observations)



$$\begin{aligned}
 MAE &= \frac{1}{n} \sum_1^n |X - X'| \\
 MRE &= \frac{1}{n} \sum_1^n \frac{|X - X'|}{X} \\
 RMSE &= \sqrt{\frac{\sum_1^n (X - X')^2}{n-1}} \\
 AURE &= \frac{1}{n} \sum_1^n \frac{2|X - X'|}{|X + X'|}
 \end{aligned}
 \tag{15}$$

其中,  $X$  表示实测值,  $X'$  表示校正后数值,  $n$  表示星地同步匹配数据对个数。

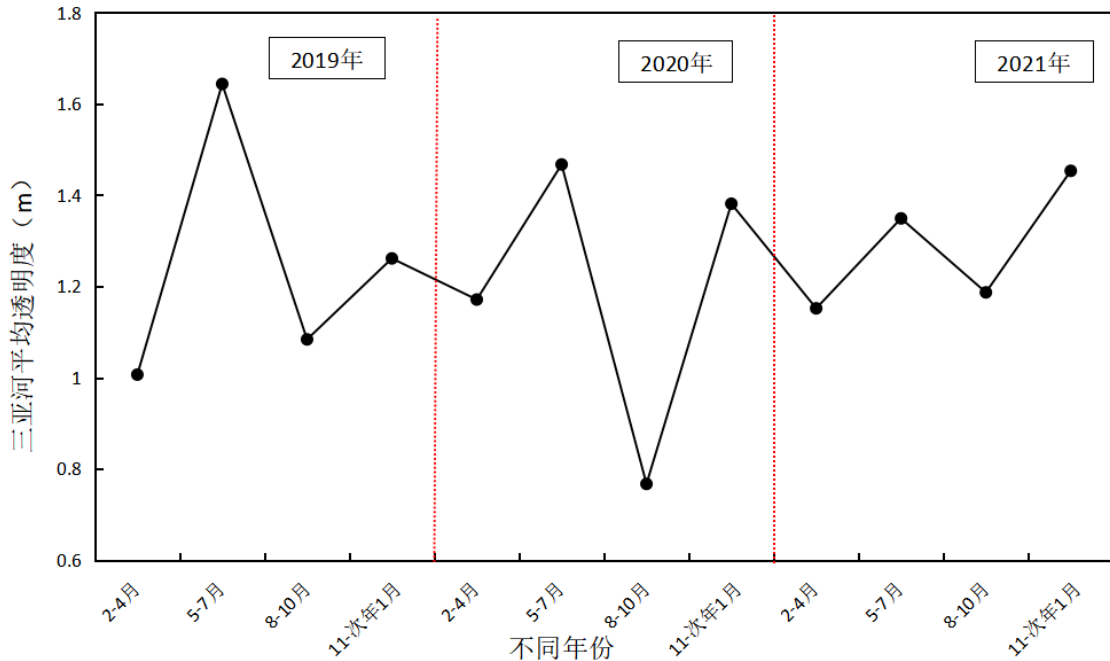


图 5 2019 年—2021 年三亚河季度平均透明度变化曲线

Figure 5 Change curve of the quarterly average transparency of Sanya River from 2019 to 2021

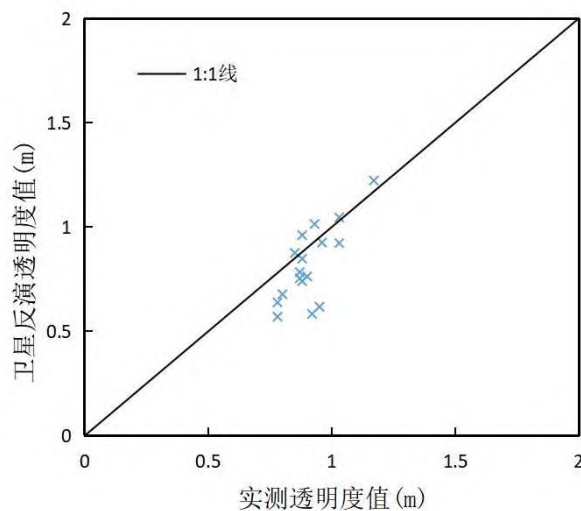


图 6 卫星反演透明度与同步实测透明度比较散点图

Figure 6 Scatter chart of the comparison between the satellite inversion transparency and the synchronous measurement transparency

## 4 数据价值

本研究以水质时空数据缺乏且受关注度高的三亚河为研究区域，基于 GEE 平台的 Sentinel-2 遥感数据集生产 2019–2021 年 36 期月度透明度遥感产品以及 12 期季度透明度遥感产品，监测其长时序时空变化规律。本数据集立足于海南省河流生态监管业务需求，有助于从宏观上把握三亚河水质动态变化趋势，为进一步加强三亚河水环境监测提供客观可靠的遥感监测数据集，为后续水环境治理以及周边生态环境建设提供数据和理论依据。

## 5 数据使用方法和建议

2019–2021 年三亚河水体透明度长时序空间分布数据保存格式均为 TIFF 格式，可以在 ArcGIS、ENVI 等地理信息系统软件中对该数据集进行读取、分析、制图等相关操作。目前针对海南岛内陆水体透明度的研究相对较少，本文提供的数据集可以为后续研究海南岛水质提供研究参考。

## 数据作者分工职责

邱瑞婷（1998—），女，河北邯郸人，硕士研究生，研究方向为水环境遥感。主要承担工作：基础数据的收集、数据处理、数据集生产和实地调查。

王胜蕾（1988—），女，河北邢台人，博士，助理研究员，研究方向为水环境遥感。主要承担工作：数据集设计、数据统计与分析。

史建康（1980—），男，安徽利辛人，高级工程师，研究方向为生态调查与评估。主要承担工作：数据集质量控制。

李俊生（1979—），男，吉林人，博士，研究员，研究方向为水环境遥感。主要承担工作：数据集设计、数据集质量控制。

张方方（1986—），男，河南人，博士，副研究员，研究方向为水环境遥感。主要承担工作：数据集质量控制。

张文志（1998—），男，河南商丘人，硕士研究生，研究方向为水环境遥感。主要承担工作：野外数据采集。

梅悦（1996—），女，山东德州人，硕士研究生，研究方向为水环境遥感。主要承担工作：野外数据采集。

## 参考文献

- [1] 海南省生态环境厅. 2019 年海南省城镇内河（湖）水环境质量月报[EB/OL]. (2019-03-02). [http://hnsthb.hainan.gov.cn/xxgk/0200/0202/hjzl/shjzl/201903/t20190321\\_2447936.html](http://hnsthb.hainan.gov.cn/xxgk/0200/0202/hjzl/shjzl/201903/t20190321_2447936.html). [Department of Ecology and Environment of Hainan Province. 2019 Hainan Urban Inland River (Lake) Water Environment Quality Monthly Report[EB/OL]. (2019-03-02). [http://hnsthb.hainan.gov.cn/xxgk/0200/0202/hjzl/shjzl/201903/t20190321\\_2447936.html](http://hnsthb.hainan.gov.cn/xxgk/0200/0202/hjzl/shjzl/201903/t20190321_2447936.html).]
- [2] 宫鹏. 拓展与深化中国全境的环境变化遥感应用[J]. 科学通报, 2012(16): 1379–1387. [GONG P. Expand and deepen the application of environmental change remote sensing in China[J]. Chinese Science

Bulletin, 2012(16): 1379–1387.]

[3] 张兵, 李俊生, 申茜, 等. 长时序大范围内陆水体光学遥感研究进展[J]. 遥感学报, 2021, 25(1): 37–52. [ZHANG B, LI J S, SHEN Q, et al. Recent research progress on long time series and large scale optical remote sensing of inland water[J]. National Remote Sensing Bulletin, 2021, 25(1): 37–52.]

[4] LEE Z P, SHANG S L, HU C M, et al. Secchi disk depth: a new theory and mechanistic model for underwater visibility[J]. Remote Sensing of Environment, 2015, 169: 139–149. DOI: 10.1016/j.rse.2015.08.002.

[5] LIU D, DUAN H T, LOISELLE S, et al. Observations of water transparency in China's Lakes from space[J]. International Journal of Applied Earth Observation and Geoinformation, 2020, 92: 102187. DOI:10.1016/j.jag.2020.102187.

[6] SONG K S, LIU G, WANG Q, et al. Quantification of lake clarity in China using Landsat OLI imagery data[J]. Remote Sensing of Environment, 2020, 243: 111800. DOI:10.1016/j.rse.2020.111800.

[7] WANG S L, LI J S, ZHANG B, et al. Changes of water clarity in large lakes and reservoirs across China observed from long-term MODIS[J]. Remote Sensing of Environment, 2020, 247: 111949. DOI:10.1016/j.rse.2020.111949.

[8] WANG X H, TIAN H, JI T M, et al. Remote sensing monitoring method for comprehensive water quality index in rivers based on sentinel-2 satellite[J]. Aerospace Shanghai (Chinese & English), 2020, 37(5): 92–97, 104. DOI:10.19328/j.cnki.1006-1630.2020.05.014.

[9] MATTHEWS M W. Eutrophication and cyanobacterial blooms in South African inland waters: 10 years of *MERIS* observations[J]. Remote Sensing of Environment, 2014, 155: 161–177. DOI:10.1016/j.rse.2014.08.010.

[10] HOU X J, FENG L, DUAN H T, et al. Fifteen-year monitoring of the turbidity dynamics in large lakes and reservoirs in the middle and lower basin of the Yangtze River, China[J]. Remote Sensing of Environment, 2017, 190: 107–121. DOI:10.1016/j.rse.2016.12.006.

[11] QING S, CUI T W, LAI Q, et al. Improving remote sensing retrieval of water clarity in complex coastal and inland waters with modified absorption estimation and optical water classification using Sentinel-2 MSI[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 102: 102377. DOI:10.1016/j.jag.2021.102377.

[12] MCFEETERS S K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features[J]. International Journal of Remote Sensing, 1996, 17(7): 1425–1432. DOI:10.1080/01431169608948714.

[13] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62 – 66. DOI:10.1109/TSMC.1979.4310076.

[14] 黄广才, 成其换. 近 30a 杞麓湖面积变化遥感监测分析[J]. 地球科学前沿(汉斯), 2019(11): 1064–1070. [HUANG G C, CHENG Q H. Remote sensing monitoring and analysis of the area change of nearly 30a qilu lake[J]. Advances in Geosciences (Hans), 2019(11): 1064–1070.]

[15] WANG S L, LI J S, ZHANG B, et al. Trophic state assessment of global inland waters using a MODIS-derived Forel-Ule index[J]. Remote Sensing of Environment, 2018, 217: 444–460. DOI:10.1016/j.rse.2018.08.026.

- [16] WANG S L, LI J S, BING Z, et al. A simple correction method for the MODIS surface reflectance product over typical inland waters in China[J]. *International Journal of Remote Sensing*, 2016, 37(24): 6076–6096. DOI:10.1080/01431161.2016.1256508.
- [17] LEE Z P, HU C M, SHANG S L, et al. Penetration of UV-visible solar radiation in the global oceans: insights from ocean color remote sensing[J]. *Journal of Geophysical Research: Oceans*, 2013, 118(9): 4241–4255. DOI:10.1002/jgrc.20308.
- [18] LEE Z P, CARDER K L, ARNONE R A. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters[J]. *Applied Optics*, 2002, 41(27): 5755–5772.
- [19] LEE Z P, et al. Update of the Quasi-Analytical Algorithm (QAA\_v6) [R/OL]. International Ocean Color Group Software Report [2013-04-03]. [http://www.ioccg.org/groups/Software-OCA/OAA\\_v5](http://www.ioccg.org/groups/Software-OCA/OAA_v5). 2014.
- [20] LEE Z P, HU C M, SHANG S L, et al. Penetration of UV-visible solar radiation in the global oceans: insights from ocean color remote sensing[J]. *Journal of Geophysical Research: Oceans*, 2013, 118(9): 4241–4255. DOI:10.1002/jgrc.20308.
- [21] 海南气象信息服务网. 2019 海南月气候影响评价[EB/OL]. (2019-06-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>. [Hainan meteorological information service website. 2019 Hainan monthly climate impact assessment[EB/OL]. (2019-06-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>.]
- [22] 海南气象信息服务网. 2020 海南月气候影响评价[EB/OL]. (2020-06-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>. [Hainan meteorological information service website. 2020 Hainan monthly climate impact assessment[EB/OL]. (2020-06-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>.]
- [23] 海南气象信息服务网. 2021 海南月气候影响评价[EB/OL]. (2021-06-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>. [Hainan meteorological information service website. 2021 Hainan monthly climate impact assessment[EB/OL]. (2021-06-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>.]
- [24] 海南气象信息服务网. 2020 海南月气候影响评价[EB/OL]. (2020-11-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>. [Hainan meteorological information service website. 2020 Hainan monthly climate impact assessment[EB/OL]. (2020-11-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>.]
- [25] 海南气象信息服务网. 2021 海南月气候影响评价. [EB/OL] (2021-11-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>. [Hainan meteorological information service website. 2021 Hainan monthly climate impact assessment[EB/OL]. (2021-11-10). <http://www.hainanqx.cn/UI/WeatherService/WeatherService20200526.aspx?type=2&tab=5>.]
- [26] 禹定峰, 周燕, 邢前国, 等. 基于实测数据和卫星数据的黄东海透明度估测模型研究[J]. *海洋环境科学*, 2016, 35(5): 774 – 779. DOI:10.13634/j.cnki.mes.2016.05.023. [YU D F, ZHOU Y, XING Q G, et al. Retrieval of Secchi disk depth using MODIS satellite remote sensing and *in situ* observations in the Yellow Sea and the East China Sea[J]. *Marine Environmental Science*, 2016, 35(5): 774–779. DOI:10.13634/j.cnki.mes.2016.05.023.]

## 论文引用格式

邱瑞婷, 王胜蕾, 史建康. 等. 基于哨兵 2 数据的 2019–2021 年三亚河水体透明度数据集[J/OL]. 中国科学数据, 2022, 7(4). (2022-12-21). DOI: 10.11922/11-6035.noda.2022.0015.zh.

## 数据引用格式

邱瑞婷, 王胜蕾, 史建康. 等. 基于哨兵 2 数据的 2019–2021 年三亚河水体透明度数据集[DS/OL]. Science Data Bank, 2022. (2022-12-21). DOI: 10.11922/sciencedb.j00001.00412.

## A dataset of water transparency of Sanya River based on Sentinel-2 data during 2019–2021

QIU Ruiting<sup>1,2,3</sup>, WANG Shenglei<sup>1,2,4\*</sup>, SHI Jiankang<sup>5</sup>, LI Junsheng<sup>1,2,4</sup>,  
ZHANG Fangfang<sup>2,4</sup>, ZHANG Wenzhi<sup>2,4</sup>, MEI Yue<sup>1,2,3</sup>

1. Key Laboratory of Earth Observation of Hainan Province, Hainan Aerospace Information Research Institute, Hainan 572029, P. R. China
2. Key Laboratory of Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, P. R. China
3. China University of Geoscience (Beijing), Beijing 100049, P. R. China
4. International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, P. R. China
5. Center of Ecology Environment Monitoring of Hainan Province, Haikou 571126, P. R. China

\*Email: wangsl@radi.ac.cn

**Abstract:** As one of the most important water quality parameters on the radar screen of environmental protection sectors, water transparency reveals the turbidity degree of water and plays an important role in the primary productivity of water body and water ecosystem. As an independent island water system, Hainan Province has abundant surface inland water resources and plentiful river runoff. However, due to the influence of dry and wet monsoons and topography, the aquatic systems are characterized by uneven spatial and temporal distribution, and there are few studies on the water quality of inland water bodies on Hainan Island. In this study, we took Sanya River in Sanya, Hainan Province as the study area, and used the QAAv6-based semi-analytic model to retrieve the water transparency of Sanya River in time series from 2019 to 2021 based on the GEE cloud computing platform and the massive Sentinel-2 surface reflectance data stored in Google Cloud. With regard to the extraction of dynamic water area from Sanya River, we adopted the algorithm combining the normalized water body index NDWI with OTUS automatic threshold segmentation to extract the small river water. The data are stored in GeoTiff raster format, and the pixel transparency value and coordinate information are stored at the same time for easy reading and analysis by relevant GIS software. The inversion of long time series transparency based on the GEE cloud database is highly efficient. The dataset can serve as valuable scientific evidence for the water quality monitoring, water pollution control,

and aquatic ecological protection of Sanya River.

**Keywords:** Sanya River; Sentinel 2; Transparency; QAA semi-analytical Model

### Dataset profile

<b>Title</b>	A dataset of water transparency of Sanya River based on Sentinel-2 data during 2019–2021
<b>Data corresponding author</b>	WANG shenglei (wangsl@radi.ac.cn)
<b>Data author(s)</b>	QIU Ruiting, WANG Shenglei, SHI Jiankang, LI Junsheng, ZHANG Fangfang, ZHANG Wenzhi, MEI Yue
<b>Time range</b>	2019–2021
<b>Geographical scope</b>	Sanya River (18°09'34"–18°37'27"N, 108°56'30"–109°48'28"E)
<b>Spatial resolution</b>	10 m
<b>Data volume</b>	99.2 MB
<b>Data format</b>	*.tif, *.csv
<b>Data service system</b>	< <a href="http://doi.org/10.11922/sciencedb.j00001.00412">http://doi.org/10.11922/sciencedb.j00001.00412</a> >
<b>Sources of funding</b>	Hainan Provincial Department of Science and Technology (Grant No. ZDKJ2019006); National Natural Science Foundation of China (Grant No. 41901272); Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA19080304).
<b>Dataset composition</b>	This dataset contains the transparency data of the spatial and temporal distribution of monthly and quarterly averages of Sanya River for the last 3 years during 2019–2021, with a total of 36 periods of monthly average transparency data and 12 periods of quarterly transparency data (one period every 3 months). The daily, monthly and quarterly average transparency data are also available. All data are saved as one compressed file, namely “A dataset of water transparency of Sanya River based on Sentinel-2 data during 2019–2021”.