

# 全球海洋表面叶绿素 a 浓度 21 年 (1998–2018) 月-季-年度数据集研发

李连伟<sup>1</sup>, 付宇轩<sup>1,2</sup>, 薛存金<sup>2,3,\*</sup>, 崔建勇<sup>1</sup>, 张源榆<sup>1,2</sup>, 徐洋峰<sup>1,2</sup>

1. 中国石油大学(华东)海洋与空间信息学院, 青岛 266580;

2. 中国科学院数字地球重点实验室, 北京 100094;

3. 中国科学院空天信息创新研究院, 北京 100094

**摘 要:** 基于 1998 年 1 月至 2018 年 12 月期间的 SeaWiFS、Terra-MODIS、Aqua-MODIS、MERIS 和 VIIRS 共 5 个传感器的叶绿素 a 浓度数据, 设计了基于小波变换与 Kalman 滤波技术相结合的多源遥感数据融合技术和基于查找表法与最大值合成法相结合的融合产品研制技术, 研发了月、季、年共三种尺度的全球海洋表面叶绿素 a 浓度融合数据集 (1998–2018)。该数据集的空间分辨率为 4 km×4 km, 时间分辨率为月/季/年, 数据格式为 TIFF, 由 357 个数据文件组成, 数据量为 50.1 GB (压缩为 64 个数据文件, 压缩后数据量为 19.4 GB)。将该数据产品与实测值和欧空局 GSM 数据产品的对比分析, 结果显示该数据集 2008 年数据产品与实测值的拟合度为 79%, 而 GSM 与实测值的拟合度仅为 35%, 低于本融合数据集的数据精度。

**关键词:** 海洋表面; 叶绿素 a; 数据融合; 数据集; 1998–2018

**DOI:** <https://doi.org/10.3974/geodp.2021.02.15>

**CSTR:** <https://cstr.escience.org.cn/CSTR:20146.14.2021.02.15>

**数据可用性声明:**

本文关联实体数据集已在《全球变化数据仓储电子杂志(中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2021.05.05.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2021.05.05.V1>.

## 1 前言

叶绿素 a 是一个关键的水色要素, 其值的高低能够反映水体初级生产力状况, 是评价水体富营养化程度的指标<sup>[1]</sup>, 因此掌握叶绿素 a 的含量及变化情况对维护生态环境质量具有重要意义。世界上各个国家已相继发射了多颗海洋卫星, 能够实现大范围、高精度的海洋环境观测, 可为各国科研人员研究海洋环境提供海量遥感数据, 也为探测海洋叶绿素 a 浓度提供了坚实的数据基础<sup>[2–4]</sup>。

由于不同的遥感传感器的时空分辨率不同, 单个卫星影像在覆盖率、分辨率和利用率方面存在不同的缺陷。为克服存在的缺陷, 充分利用不同传感器影像的优点, 相关科研人

收稿日期: 2021-04-06; 修订日期: 2021-05-30; 出版日期: 2021-06-25

基金项目: 中国科学院战略性 A 类型先导专项 (XDA19060103)

\*通讯作者: 薛存金 0000-0003-3605-6578, 中国科学院空天信息创新研究院, [xuecj@aircas.ac.cn](mailto:xuecj@aircas.ac.cn)

数据引用方式: [1] 李连伟, 付宇轩, 薛存金等. 全球海洋表面叶绿素 a 浓度 21 年 (1998–2018) 月-季-年度数据集研发[J]. 全球变化数据学报, 2021, 5(2): 219–225. <https://doi.org/10.3974/geodp.2021.02.15>.  
<https://cstr.escience.org.cn/CSTR:20146.14.2021.02.15>.

[2] 李连伟, 付宇轩, 薛存金等. 全球海洋表面叶绿素 a 浓度 4-km 栅格数据集 (1998–2018)[J/DB/OL]. 全球变化数据仓储电子杂志, 2021. <https://doi.org/10.3974/geodb.2021.05.05.V1>.  
<https://cstr.escience.org.cn/CSTR:20146.11.2021.05.05.V1>.

员提出了多源数据融合技术<sup>[5]</sup>。多源数据融合可增加整体的信息量，可扩展单个卫星影像的时空分辨率，也可提高数据产品的时空连续性、一致性和可靠性<sup>[4]</sup>。

随着海洋遥感卫星的发展，多源数据融合技术被用于海色数据的融合。国内外相关科研人员已不同程度的开展海洋表面叶绿素数据融合产品研制的相关研究<sup>[6-8]</sup>，但大都仍存在着精度低、覆盖率低、时间跨度短等问题。

本文基于国内外已有研究成果，利用 SeaWIFS、Terra-MODIS、Aqua-MODIS、MERIS 和 VIIRS 共 5 个传感器的叶绿素 a 浓度数据，设计了基于小波变换和 Kalman 滤波技术相结合的全球海洋表面叶绿素 a 浓度融合数据集融合算法和基于查找表法和最大值合成法相结合的全球海洋表面叶绿素 a 浓度融合数据集研制方法，生成了 1998–2018 年全球海洋表面叶绿素 a 浓度月尺度、季尺度、年尺度共三种时间尺度的融合数据集，并与实测值、数据源数据值、已有数据集数据值进行了对比分析。

2 数据集元数据简介

《全球海洋表面叶绿素 a 浓度 4-km 栅格数据集（1998–2018）》<sup>[9]</sup>的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 《全球海洋表面叶绿素 a 浓度 4-km 栅格数据集（1998–2018）》元数据简表

条 目	描 述
数据集名称	全球海洋表面叶绿素 a 浓度 4-km 栅格数据集（1998–2018）
数据集短名	Global_Chla_1998-2018
作者信息	李连伟，中国石油大学（华东）海洋与空间信息学院，lilianwei78@163.com 付宇轩，中国石油大学（华东）海洋与空间信息学院，624002974@qq.com 薛存金，0000-0003-3605-6578，中国科学院空天信息创新研究院，中国科学院数字地球重点实验室，xuecj@aircas.ac.cn 崔建勇，中国石油大学（华东）海洋与空间信息学院，xjuzhxcjy@163.com 张源榆，中国石油大学（华东）海洋与空间信息学院，1529142841@qq.com 徐洋峰，中国石油大学（华东）海洋与空间信息学院，xuyf187627@163.com
地理区域	全球海域
数据年代	1998 年 1 月至 2018 年 12 月
时间分辨率	月/季/年
数据量	50.1 GB（压缩后 19.4 GB）
数据集组成	全球海洋表面叶绿素 a 浓度融合月、季、年三个尺度的数据集
基金项目	中国科学院（XDA19060103）
出版与共享服务平台	全球变化科学研究数据出版系统 <a href="http://www.geodoi.ac.cn">http://www.geodoi.ac.cn</a>
地址	北京市朝阳区大屯路甲 11 号 100101，中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据（中英文）、通过《全球变化数据仓储电子杂志（中英文）》发表的实体数据集和通过《全球变化数据学报（中英文）》发表的数据论文。其共享政策如下：（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10%引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的 10%，同时需要对摘取的数据记录标注数据来源 <sup>[10]</sup>

3 数据集研发方法

3.1 数据来源

全球海洋表面叶绿素 a 浓度数据来源于 NASA 的 Ocean Color 官网<sup>[11]</sup>。根据传感器生命周期, 下载了从 1998 年 1 月到 2018 年 12 月共 21 年的全球海洋表面叶绿素 a 浓度数据, 共 2,921 景, 详细情况如表 2 所示。由表 2 可知, 各传感器生命周期如图 1 所示。

表 2 全球海洋表面叶绿素 a 浓度遥感数据集详细情况表

传感器	算法	开始时间	结束时间	景数	时间分辨率	空间分辨率
SeaWiFS	OCI	1997/9/18	2010/11/17	644	8 天	9 km
Terra-MODIS	OCI	2002/7/4	2018/4/13	736	8 天	4 km
Aqua-MODIS	OCI	2002/7/4	2018/4/13	736	8 天	4 km
MERIS	OCI	2002/4/29	2012/3/15	506	8 天	4 km
VIIRS	OCI	2012/1/2	2018/9/21	299	8 天	4 km

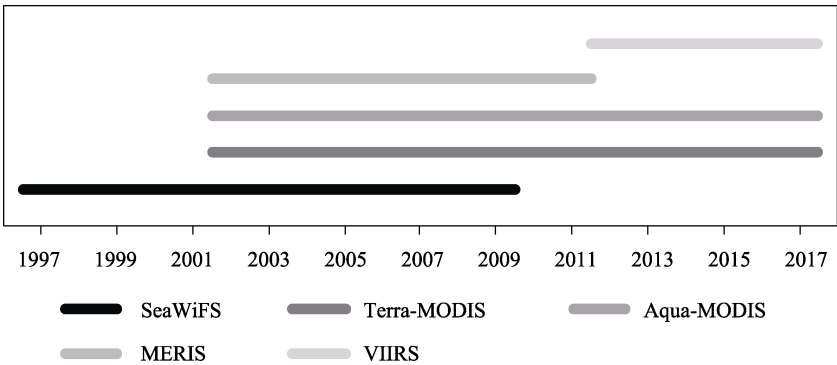


图 1 多传感器叶绿素 a 浓度遥感数据集覆盖时间范围图

全球海洋表面叶绿素 a 浓度实测数据来源于 SeaBass<sup>[12]</sup>, 总共有 5,711 个观测文件, 每个观测文件记录了站点、编号、日期、时间、纬度、经度、水深、叶绿素浓度等信息, 实测数据统计表如表 3 所示。

3.2 算法及流程

基于自适应加权融合算法实现全球海洋表面叶绿素 a 浓度融合数据集, 该算法的核心思想在于: (1) 以区域最优作为权重选择的评价标准, 使每个传感器所采集的数据自适应地找到其相应的最优加权因子; (2) 基于小波变换实现动态加权融合。自适应加权融合算法的设计流程如图 2 所示。

融合算法的具体步骤包括: (1) 首先基于是否有实测值采用阈值迭代法对图像进行分割; (2) 对没有实测值的区域用方差作为划分依据, 将无实测值区域分为相似区和多变区; (3) 相似区采用固定权重加权融合; (4) 多变区中低频信息融合采用加权系数寻优法, 高频信息融合采用保边缘平滑法; (5) 最后利用小波反变换完成全球海洋表面叶绿素 a 浓度

表 3 全球海洋表面叶绿素 a 浓度实测数据统计表

传感器	实测点 (个)
TERRA-MODIS	1,576
AQUA-MODIS	925
SeaWiFS	2,280
MERIS	781
VIIRS	149
总计	5,711

融合。

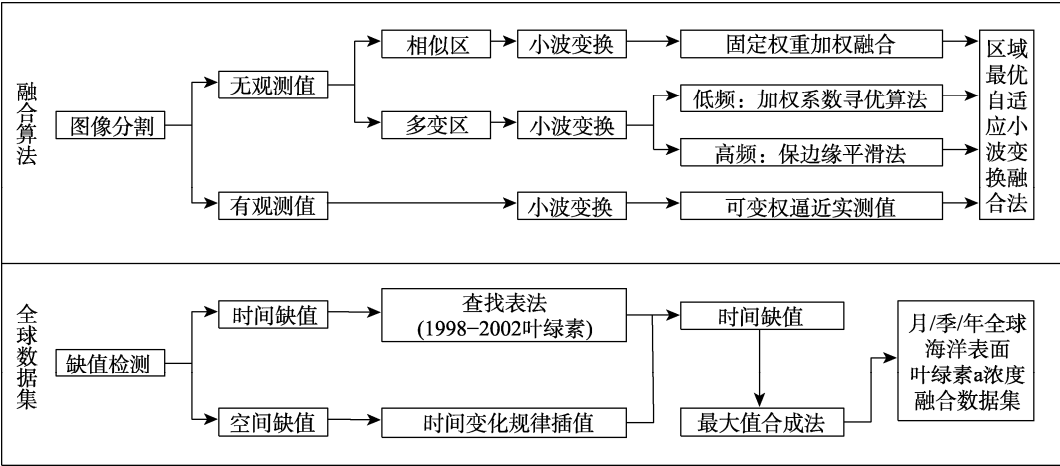


图 2 数据集研发算法流程图

《全球海洋表面叶绿素 a 浓度 4-km 栅格数据集 (1998–2018)》研制采用最大值合成法及查找表法，具体流程包括：(1) 首先对每八天的数据进行融合（全球海洋表面叶绿素 a 浓度数据 Terra、Aqua、MERIS、SeaWiFS、VIIRS），得到一年 46 幅图像；(2) 利用最大值合成法筛选出每个像素位置上每个月、季、年的最大值，作为最终融合数据集；(3) 对某些年份只有唯一数据源的情况（1998–2002 年只有 SeaWiFS 全球海洋表面叶绿素 a 浓度数据），用查找表法对该数据进行计算，使该数据和融合数据集具有一致性。

## 4 数据结果与验证

### 4.1 数据集组成

《全球海洋表面叶绿素 a 浓度 4-km 栅格数据集 (1998–2018)》包括月、季、年三个尺度的数据集，所包含的产品个数与数据量如表 4 所示。

表 4 全球海洋表面叶绿素 a 浓度遥感融合数据集组成

分类	产品个数	数据量
全球海洋表面叶绿素 a 浓度融合月尺度数据集	252	35 GB
全球海洋表面叶绿素 a 浓度融合季尺度数据集	84	11.8 GB
全球海洋表面叶绿素 a 浓度融合年尺度数据集	21	2.9 GB

### 4.2 数据预处理

数据预处理负责建立统一时空尺度的时空数据集，为融合算法设计和产品算法研制提供基础，包括数据格式的统一、空间分辨率的统一和坐标系统的统一。

#### (1) 统一数据格式

下载的全球海洋表面叶绿素 a 浓度遥感数据中 MODIS/Terra、MODIS/Aqua、SeaWiFS、VIIRS 的数据都是 NetCDF（后缀名.nc）格式的文件，MERIS 的数据为 HDF4 格式文件。

基于 ENVI-IDL 语言和 Python 语言编写数据格式转换程序，将来源于多种传感器的全球海洋表面叶绿素 a 浓度遥感数据的格式统一转换为 TIFF 格式。

（2）统一空间分辨率

为了保证数据集的精度和利用率，采用双线性插值法对较低分辨率的数据进行空间插值，插值后全球海洋表面叶绿素 a 浓度遥感数据的空间分辨率统一为 4 km。

（3）统一坐标系

将来源于多种传感器的全球海洋表面叶绿素 a 浓度遥感数据统一为 WGS-84 地理坐标系。

4.3 数据结果

数据结果包括全球海洋表面叶绿素 a 浓度融合月尺度数据集，季尺度数据集和年尺度数据集，其中，年尺度数据集示例如图 3 所示。

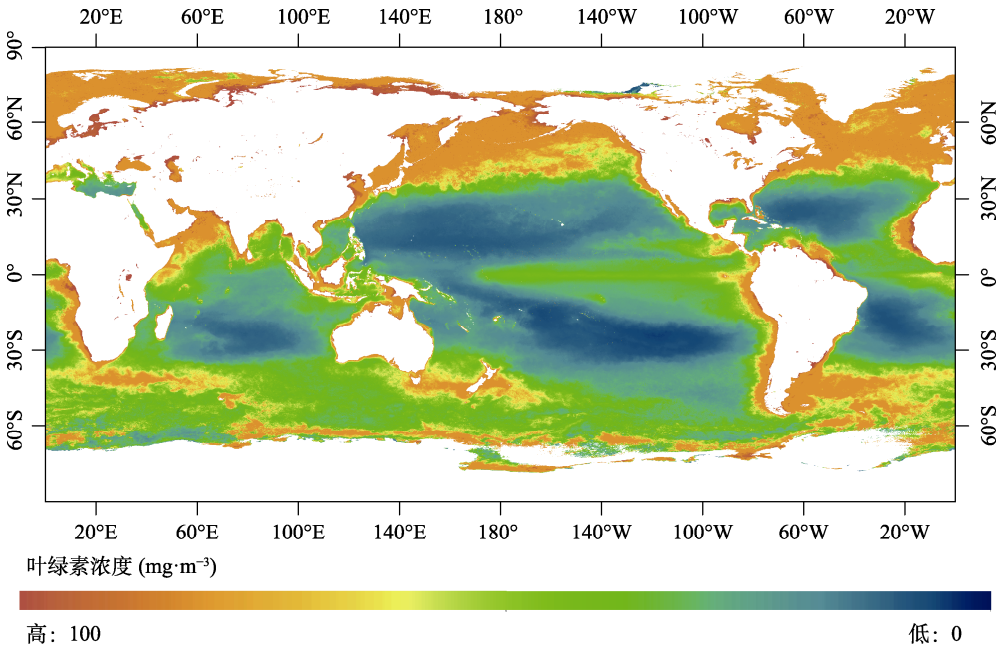


图 3 全球海洋表面叶绿素 a 浓度融合年尺度数据可视化图（2008 年）

4.4 数据验证

（1）融合数据集数据值与实测值对比分析

分析融合数据集数据值与实测值的对比，首先从实测数据集中选择每八天最大值和对应坐标值，从融合数据集图像中找到对应坐标和时间匹配的叶绿素 a 浓度融合值；再计算以 10 为底的对数值，其中叶绿素 a 浓度小于 1 的计算后变为负值；最后计算两种数据点的相关性。经计算两者数据点的匹配度较高，近 87% 的点拟合度很好。

（2）融合数据集数据值与数据源数据值对比分析

为对比分析融合数据集数据值与数据源数据值，一组选用 2005–2010 年的融合数据集、Aqua、Terra、MERIS、SeaWIFS 等数据，另一组选用 2012–2016 年的融合数据集、Aqua、Terra、VIIRS 等数据。分别统计出与实测最大值对应的遥感值，然后再筛选出同一天的融

合数据集数据值和多源传感器数据值，并开展相关性分析，其结果如图 4、图 5 所示。

从第一组数据对比中发现，与融合数据集相关性最好的是 MERIS 数据，拟合度达到了近 87%，拟合度最差的是 Terra 数据，拟合度仅为 66%。

从第二组数据对比中发现，Aqua、Terra、VIIRS 等三种数据与融合数据集的拟合程度要好于第一组，最小拟合度近为 83%，最大拟合度近为 93%，说明这三种传感器数据与融合数据拟合性较好。

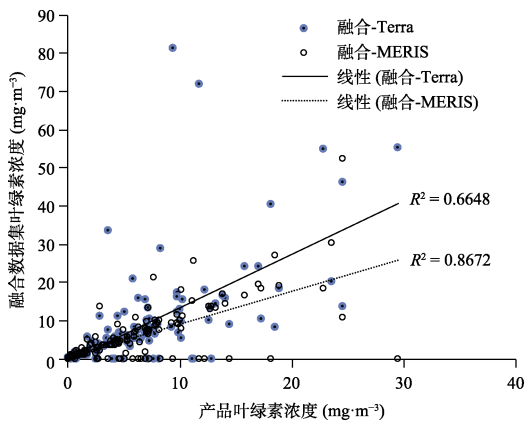


图 4 融合数据集数据值与源数据数据值对比分析图

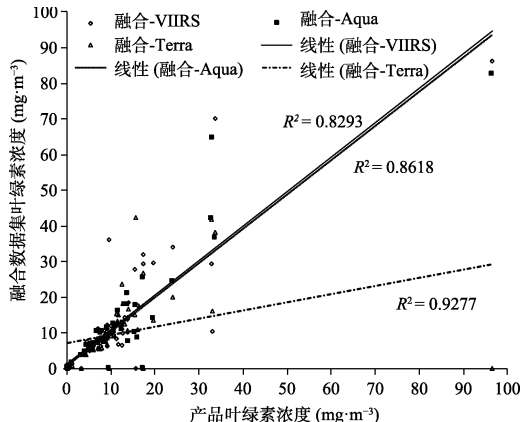


图 5 融合数据集数据值与源数据数据值对比分析分析图

(3) 融合数据集数据值与已有数据集数据值对比分析

选择 2008 年融合数据集数据值、GSM 产品数据值、2008 年实测最大值进行对比，共匹配 24 对点，分别以融合-实测、GSM-实测进行对比分析，如图 6 所示。从图 6 可知，融合数据集数据值与实测值的拟合度为 79%，而 GSM 数据值与实测值的拟合度仅为 35%。由此可知，融合数据集数据值与实测数据值的拟合度远远高于 GSM 产品，说明融合数据集质量较高，与实测点匹配较好。

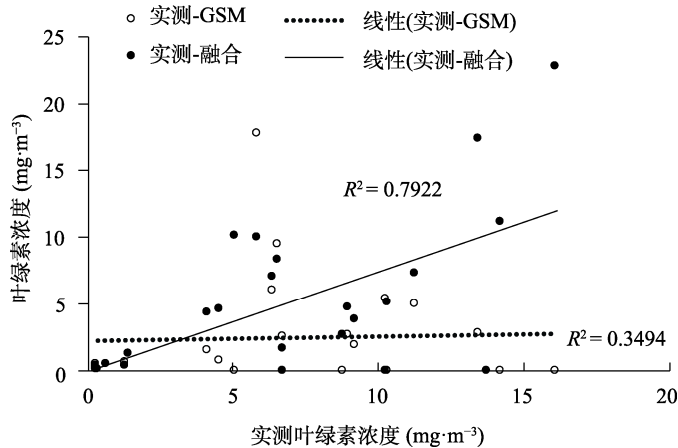


图 6 融合数据集、GSM 产品与实测数据值对比分析图

## 5 讨论和总结

本文基于 SeaWiFS、Terra-MODIS、Aqua-MODIS、MERIS 和 VIIRS 共 5 个传感器的叶绿素 a 浓度数据, 设计了数据融合算法和融合数据集研制方法, 研制了 1998–2018 年期间的月、季、年三种时间尺度的全球海洋表面叶绿素 a 浓度融合数据集, 空间分辨率为  $4\text{ km} \times 4\text{ km}$ , 时间分辨率为月/季/年, 数据量 50.1 GB (压缩后 19.4 GB)。基于实测值和欧空局的 GSM 数据值, 对本融合数据集进行了验证, 结果表明本融合数据集的拟合度高于 GSM 产品。因需大量实测数据对融合算法进行优化, 但由于实测数据的不足, 本融合数据集的精度还有进一步提升空间。

**作者分工:** 李连伟负责数据结果验证; 薛存金负责数据集总体规划与设计; 崔建勇负责数据集融合算法设计与算法实现; 付宇轩、张源榆参与数据下载与预处理; 徐洋峰负责数据集融合; 李连伟和付宇轩负责数据论文撰写。

**利益冲突声明:** 本研究不存在研究者以及与公开研究成果有关的利益冲突。

## 参考文献

- [1] 王卷乐, 张永杰, 杨飞等. 鄱阳湖叶绿素 a 浓度数据集(2009–2012)[J]. 全球变化数据学报, 2017, 1(2): 208–215.
- [2] 李新星, 张亭禄, 田林等. 多卫星传感器南海叶绿素 a 浓度遥感数据融合[J]. 遥感学报, 2015, 19(4): 680–689.
- [3] 施英妮, 张亭禄, 石立坚等. 基于客观分析的多源卫星叶绿素 a 浓度产品融合方法研究[J]. 海洋学报(中文版), 2016(3): 82–87.
- [4] 崔建勇, 刘晓东, 岳增友等. 多源海洋遥感叶绿素数据融合[J]. 遥感信息, 2020, 35(3): 31–36.
- [5] 陈芸芝, 汪小钦, 吴波等. 基于自适应加权平均的水色遥感数据融合[J]. 遥感技术与应用, 2012, 27(3): 333–338.
- [6] Gregg, W. W., Conkright, M. E. Global seasonal climatologies of ocean chlorophyll: blending *in situ* and satellite data for the coastal zone color scanner era [J]. *Journal of Geophysical Research*, 2001, 106(C2): 2499–2516.
- [7] 曲利芹, 管磊, 贺明霞. SeaWiFS 和 MODIS 叶绿素浓度数据及其融合数据的全球可利用率[J]. 中国海洋大学学报(自然科学版), 2006, 36(2): 321–326.
- [8] 陈芸芝, 郑高强, 汪小钦等. 基于 GSM01 融合的多传感器数据叶绿素 a 浓度反演[J]. 地球信息科学学报, 2013, 15(6): 911–917.
- [9] 李连伟, 付宇轩, 薛存金等. 全球海洋表面叶绿素 a 浓度 4-km 栅格数据集(1998–2018)[J/DB/OL]. 全球变化数据仓储电子杂志, 2021. <https://doi.org/10.3974/geodb.2021.05.05.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2021.05.05.V1>.
- [10] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策[OL]. <https://doi.org/10.3974/dp.policy>. 2014.05 (2017 年更新).
- [11] MERIS Reprocessing Information. <https://oceancolor.gsfc.nasa.gov/cgi/browse.pl>.
- [12] SeaBASS. <https://seabass.gsfc.nasa.gov/>.