

2007–2015 年太湖水体理化监测数据集

ISSN 2096-2223
CN 11-6035/N



文献 DOI:
10.11922/csdata.2019.0037.zh
数据 DOI:
10.11922/sciencedb.833
文献分类: 地球科学

收稿日期: 2019-08-15
开放同评: 2019-09-02
录用日期: 2019-10-30
发表日期: 2020-02-24

闵岫^{1,2}, 钱荣树², 朱广伟^{1,2}, 黄建明², 秦伯强^{1,2*}, 杨宏伟^{1,2},

李宽意^{1,2}, 张运林^{1,2}, 沈睿杰^{1,2}

1. 中国科学院南京地理与湖泊研究所, 南京 210008
2. 中国科学院太湖湖泊生态系统研究站, 江苏无锡 214026

摘要: 本数据集整理了 2007–2015 年中国科学院太湖湖泊生态系统研究站太湖 8 个长期常规监测站点的逐月理化数据, 包括透明度、水深、水温、电导率、悬浮质、溶解氧、pH、硅酸盐、磷酸盐、亚硝酸盐氮、硝酸盐氮、氨态氮、化学需氧量、总磷、总氮和溶解性总有机碳, 共计 16 个理化指标, 13 810 个数值。使用箱线法对数据进行质量检验, 检验结果显示 16 个指标中有 10 个指标的有效数据占比达到 95% 以上, 仅有 1 个指标的有效数据占比在 90% 以下。本数据集的整理出版可为探究太湖生态系统变化特征提供数据基础, 为富营养化防控治理及饮用水安全保障等提供科学支撑。

关键词: 太湖; 富营养; 理化指标; 长期监测

数据库(集)基本信息简介

数据库(集)名称	2007–2015 年太湖水体理化监测数据集
数据作者	闵岫, 钱荣树, 朱广伟, 黄建明, 秦伯强, 杨宏伟, 李宽意, 张运林, 沈睿杰
数据通信作者	秦伯强 (qinbq@niglas.ac.cn)
数据时间范围	2007–2015 年
地理区域	太湖
数据量	169 KB
数据格式	*.xlsx
数据服务系统网址	http://www.sciencedb.cn/dataSet/handle/833
基金项目	中国生态系统研究网络(CERN)野外台站运行服务费; 科技部国家生态系统观测研究网络(CNERN)研究站运行服务费。
数据库(集)组成	本数据集包含数据 864 条, 每条数据包含: 台站代码、年、月、观测站代码、透明度(m)、水深(m)、水温(℃)、电导率(μs/cm)、悬浮质(mg/L)、溶解氧(mg/L)、pH、硅酸盐(mg/L)、磷酸盐(mg/L)、亚硝酸盐氮(μg/L)、硝酸盐氮(mg/L)、氨态氮(mg/L)、化学需氧量(mg/L)、总磷(mg/L)、总氮(mg/L)和溶解性总有机碳(mg/L)。

* 论文通信作者

秦伯强: qinbq@niglas.ac.cn

引言

太湖（119°52′–120°36′E，30°55′–31°32′N）是我国五大淡水湖泊之一，地处长江三角洲，流域面积约 3.69 万 km²。太湖流域是我国人口最密集、经济最发达的地区之一。太湖作为典型的浅水性湖泊，平均水深 1.9 m，水面面积 2338 km²，是整个太湖流域水调节和水生态系统的中心，承担着农业灌溉、工业用水及居民生活用水等重要功能，对长江下游地区的经济发展起着举足轻重的作用^[1]。随着经济发展加快，人类活动增加使得大量的营养物质进入湖泊，导致太湖富营养化水平增高，随之而来的蓝藻水华日益严重，对湖泊生态功能及周边人民供水安全造成了严重危害。

以往研究表明，氮、磷、有机碳及水温等理化因子的综合作用对太湖蓝藻的生长和水华的暴发有着重要的影响^[2-3]。因此，水体理化因子的长期监测数据有助于了解太湖水生态系统的变化情况，可为太湖水质目标管理、及时有效制定应对措施等提供数据基础。

中国科学院太湖湖泊生态系统研究站（以下简称太湖站）隶属于中国科学院南京地理与湖泊研究所，成立于 1988 年，并于 1989 年进入中国生态系统研究网络（CERN）。“太湖站”无锡本部位于太湖梅梁湾东岸（31°24′N，120°13′E，图 1），作为太湖生态环境长期监测平台于 1991 年开始对太湖进行连续采样监测，积累了丰富的历史监测数据。本数据集汇集了太湖站 8 个长期观测站点 2007–2015 年的月尺度数据，包括透明度（m）、水深（m）、水温（℃）、电导率（μs/cm）、悬浮质（mg/L）、溶解氧（mg/L）、pH、硅酸盐（mg/L）、磷酸盐（mg/L）、亚硝酸盐氮（μg/L）、硝酸盐氮（mg/L）、氨态氮（mg/L）、化学需氧量（mg/L）、总磷（mg/L）、总氮（mg/L）和溶解性总有机碳（mg/L）。

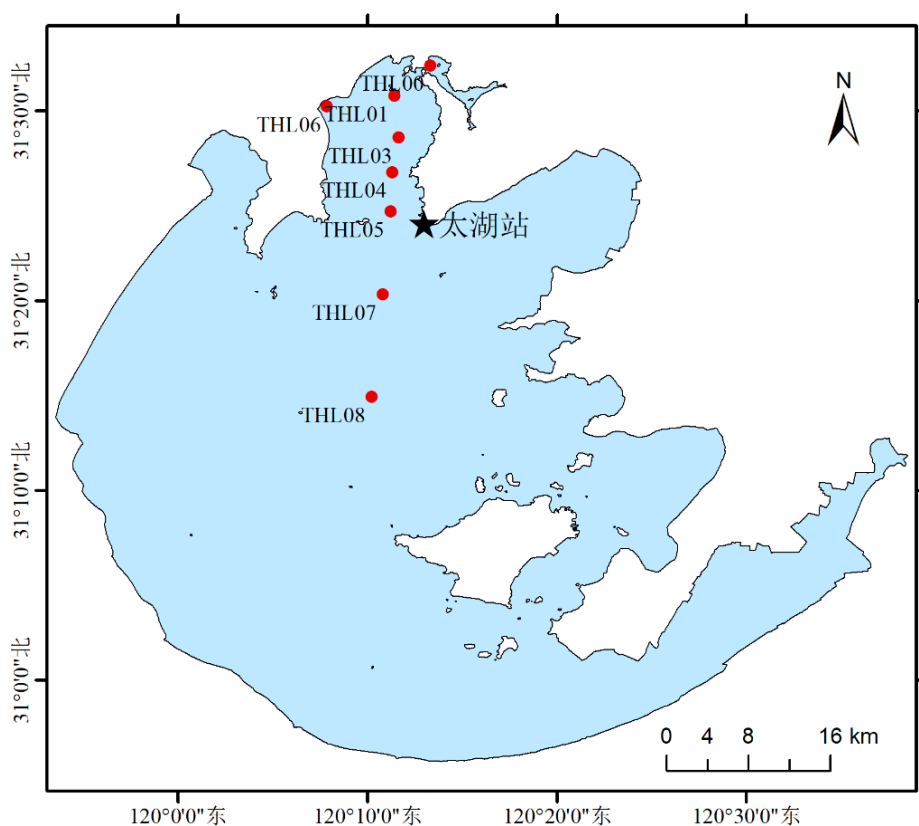


图 1 太湖站位置以及站点分布图

1 数据采集和处理方法

1.1 数据采集

本数据集中涉及的太湖站 8 个常规观测站点分别为 THL00、THL01、THL03、THL04、THL05、THL06、THL07 和 THL08，样地编码沿用 CERN 长期生态联网监测中的编码，具体见表 1。此 8 个观测站点长期定位监测太湖生态环境各项指标，采样频率均为 1 次/月，于每月中旬采样。

表 1 太湖站观测点序号、编码及经纬度

观测站序号	观测站代码	观测站名称	东经 (°)	北纬 (°)
0	THL00	太湖站 0 号观测站	120.22217	31.53983
1	THL01	太湖站 1 号观测站	120.19067	31.51317
3	THL03	太湖站 3 号观测站	120.19433	31.47633
4	THL04	太湖站 4 号观测站	120.18883	31.44583
5	THL05	太湖站 5 号观测站	120.18733	31.41117
6	THL06	太湖站 6 号观测站	120.13117	31.50383
7	THL07	太湖站 7 号观测站	120.18017	31.33833
8	THL08	太湖站 8 号观测站	120.17062	31.24816

1.2 数据处理方法

本数据集包含的各项指标中，有些指标为采样现场测定，如水深、水温、透明度。其中，使用测深仪测量水深，水温计测量水温，测量仪器定期校准；透明度使用塞氏圆盘，每个采样点测定 3 次取均值。现场测定的指标通常是采样人员在现场测量之后，将数据记录在采样表中，并填写其他辅助信息，如采样点、采样时间以及天气等。

其余指标在对采集的水样进行妥善保存之后带回实验室按照湖泊生态调查观测与分析标准方法^[4]进行测定。其中，pH 使用 PHS-3TC 仪测定，使用前使用标准缓冲液校准；电导率使用 DDS-11C 电导率仪进行测定；悬浮物使用烘箱和十万分之一的电子天平进行测定；溶解氧、化学需氧量使用 10 ml 微量滴定管进行测定，使用前进行浓度标定；磷酸盐、亚硝酸盐氮、硝酸盐氮、氨态氮、总磷、总氮使用紫外分光光度计进行测定，分析时带标样控制分析结果；硅酸盐使用 SKALAR 分析仪，溶解性总有机碳使用 TOC-VCPC 仪进行测定，分析时均带标样控制分析结果（表 2）。

表 2 太湖站水质监测使用仪器及质量控制表^[5]

项目	分析方法	使用仪器	质量控制
水深	测深仪法	测深仪	测深仪每年校准
水温	水温计法	表层水温计	水温计每年校核
透明度	塞氏盘法	塞氏圆盘	每个点测定 3 次取均值
溶解氧	碘量法	10 ml 微量滴定管	每次标定硫代硫酸钠使用液浓度
pH	pH 复合电极法	PHS-3TC	每年更换电极，每次使用前用标准缓冲液校准

项目	分析方法	使用仪器	质量控制
硅酸盐	硅钼蓝自动比色法	SKALAR 流动分析仪	分析时带标样控制分析结果
磷酸盐	钼锑抗分光光度法	UV-2450PC 紫外分光光度计	分析时带标样控制分析结果
亚硝酸盐氮	N-(1-萘基)-乙二胺光度法	UV-2450PC 紫外分光光度计	分析时带标样控制分析结果
硝酸盐氮	酚二磺酸光度法	UV-2450PC 紫外分光光度计	分析时带标样控制分析结果
氨态氮	纳氏试剂光度法	UV-2450PC 紫外分光光度计	分析时带标样控制分析结果
化学需氧量	酸性法	10 ml 微量滴定管	每次标定高锰酸钾使用液浓度
总磷	钼锑抗分光光度法	UV-2450PC 紫外分光光度计	分析时带标样控制分析结果
总氮	紫外分光光度法	UV-2450PC 紫外分光光度计	分析时带标样控制分析结果
溶解性总有机碳	燃烧法	TOC-VCPN	分析时带标样控制分析结果
电导率	电导率仪法	DDS-11C 电导率仪	定期标定
悬浮物	烘干法	烘箱, 十万分之一电子天平	烘 3 小时后在干燥器中冷却称重, 再烘半小时后在干燥器中冷却称重; 误差小于万分之一

2 数据样本描述

太湖水体理化监测数据集为表格型数据, 包含太湖站 8 个常规观测站点 2007–2015 年 16 个理化指标, 观测频率为 1 次/月。其具体字段名称、类型等参见表 3。数据集中缺测数据用 “/” 表示, 使用箱线法对数据进行质量评估得到的离群数据用绿色字体标注, 异常数据用红色字体标注。各指标分析精度见示例。

表 3 太湖水体理化监测数据表内容

字段名称	数据类型	量纲/说明	示例
台站代码	字符型	THL, “TH” 太湖拼音首字母, “L” 代表湖泊站	THL
年份	数字型	观测年度, 4 位数字	2008
月份	数字型	观测月份, 1–2 位数据	5
观测站代码	字符型	观测样地编码, 5 位	THL00
透明度	数字型	m	0.00
水深	数字型	m	3.5
水温	数字型	°C	21.4
电导率	数字型	μs/cm	730
悬浮质	数字型	mg/L	873.00
溶解氧	数字型	mg/L	10.48
pH	数字型	无量纲	6.81
硅酸盐	数字型	mg/L	1.69
磷酸盐	数字型	mg/L	0.015

字段名称	数据类型	量纲/说明	示例
亚硝酸盐氮	数字型	μg/L	14.6
硝酸盐氮	数字型	mg/L	0.35
氨态氮	数字型	mg/L	0.61
化学需氧量	数字型	mg/L	140.25
总磷	数字型	mg/L	6.478
总氮	数字型	mg/L	86.27
溶解性总有机碳	数字型	mg/L	14.996
备注	字符型	必要的说明文字	水华很多, 100% 覆盖

3 数据质量控制和评估

3.1 数据质量控制

为了保证数据质量进而实现有效共享, CERN 形成了严谨的质量控制体系, 通过计划、执行、评估三个步骤, 采取前端控制和后端质控的管理模式, 对数据进行审核、检验和评估^[6]。质量控制管理具体可分为三级:

第一级为台站层级的质量控制。太湖站作为 CERN 站, 在样品采集、分析化验和数据处理过程中, 严格按照 CERN 统一制定的水环境观测规范^[7]、湖泊生态调查观测与分析标准方法^[4]以及水环境观测质量控制规范^[6]来开展相关工作。太湖站仪器分析数据使用的各种仪器设备由专人负责管理使用, 并要求厂商定期上门维护。所有监测分析项目, 由专人长期严格按照规范操作进行, 在分析测试过程中, 需要带标样控制分析结果, 并要求做到数据真实、记录规范、书写清晰、数据及辅助信息完整等。分析测试工作完成之后, 由数据管理员进行了各种源数据的集成、整理、转换、格式统一; 通过一系列质量控制方法, 去除随机及系统误差, 并与历史数据信息进行比较, 评价数据的完整性、准确性、可比性和连续性。

第二级为专业分中心的质量控制。太湖站分属 CERN 水体分中心, 每年太湖站按时保质保量向水体分中心提交监测数据, 水体分中心通过阈值范围、完整性、一致性等方面的检查进行数据审核, 汇总形成数据质量评估报告, 并将数据整理上报至 CERN 综合中心。

第三级为 CERN 综合中心的质量控制。综合中心除了用软件检查数据质量外, 重点检查数据格式的规范性, 尽可能减少数据入库的错误率。

3.2 数据质量评估

本数据集内共有数据 864 条, 每条数据包含 16 个指标共计 13 810 个数值, 其中缺失数值 14 个(表 4)。为检验数据质量, 我们使用箱线图法计算数据集中各指标的四分位间隔范围($IQR=Q3-Q1$, 其中 $Q3$ 和 $Q1$ 分别为上、下四分位数)来检验数据集中各指标的离群值(Outlier)和异常值(Extreme)^[8]。我们将高于上四分位数或低于下四分位数 1.5 倍 IQR 的数据标记为离群值; 高于上四分位数或低于下四分位数 3 倍 IQR 的数据标记为异常值, 除去离群值和异常值之外的数据称为有效数据。本数据集 16 个指标中有 10 个指标的有效数据占比达到 95% 以上, 仅有 1 个指标的有效数据占比在 90%

以下（表 4）。本数据集中没有剔除筛选出来的离群值和异常值，只是对可能存在问题的数据进行标记，供使用该数据集的研究人员参考。

表 4 数据质量评估表

项目	数值个数	缺失值个数	缺失值占比	离群值个数	离群值占比	异常值个数	异常值占比	有效数据占比
透明度	861	3	0.35%	20	2.31%	16	1.85%	95.49%
水深	853	11	1.27%	30	3.47%	2	0.23%	95.02%
水温	864	0	0.00%	0	0.00%	0	0.00%	100%
电导率	864	0	0.00%	8	0.93%	1	0.12%	98.96%
悬浮质	864	0	0.00%	37	4.28%	16	1.85%	93.87%
溶解氧	864	0	0.00%	6	0.69%	0	0.00%	99.31%
pH	864	0	0.00%	35	4.05%	2	0.23%	95.72%
硅酸盐	864	0	0.00%	7	0.81%	0	0.00%	99.19%
磷酸盐	864	0	0.00%	45	5.21%	19	2.20%	92.59%
亚硝酸盐氮	864	0	0.00%	43	4.98%	32	3.70%	91.32%
硝酸盐氮	864	0	0.00%	19	2.20%	0	0.00%	97.80%
氨态氮	864	0	0.00%	50	5.79%	50	5.79%	88.43%
化学需氧量	864	0	0.00%	13	1.50%	23	2.66%	95.83%
总磷	864	0	0.00%	23	2.66%	29	3.36%	93.98%
总氮	864	0	0.00%	33	3.82%	9	1.04%	95.14%
溶解性总有机碳	864	0	0.00%	38	4.10%	13	1.50%	94.10%

4 数据价值

富营养化是全球湖泊普遍面临的生态灾害问题之一。太湖是我国蓝藻水华问题出现最早、治理时间最长的大型湖泊之一。2007 年 5 月无锡发生饮用水危机事件，根本原因就是富营养化导致的蓝藻水华在水源地积聚^[9]。自此之后，太湖治理攻坚战已开展 10 余年，整体水质有所改善^[10-11]，但蓝藻水华灾害问题依然存在^[12]。

水体理化因子是衡量气候变化及人类活动对湖泊生态系统影响的重要指标。本数据集包含了太湖站 8 个长期监测站点 2007–2015 年的理化数据，对深入了解太湖水体生态系统的结构、功能和演替规律具有重要意义，可为揭示水体内能量流转及物质循环过程、探索大型富营养化浅水湖泊蓝藻水华爆发机制提供数据基础，为蓝藻水华预测预警、富营养化防控治理及饮用水安全保障等提供科学支撑。

5 数据使用方法和建议

本数据集可通过 Science Data Band 在线服务网址（<http://www.sciencedb.cn/dataSet/handle/833>）

获取数据；也可登录江苏太湖湖泊生态系统国家野外科学观测研究站数据资源服务平台 (<http://thl.cern.ac.cn/meta/detail/LC02>)，通过申请获取数据服务。

数据作者分工职责

闵岫（1984—），女，江西南昌人，博士，工程师，数据管理员。主要承担工作：数据整理和论文撰写。

钱荣树（1967—），男，江苏泰州人，工程师，实验员。主要承担工作：水体化学指标分析。

朱广伟（1972—），男，河南郑州人，博士，研究员。主要承担工作：监测方案设计、水样采集以及数据质量管理。

黄建明（1977—），男，江苏南京人，助理工程师，实验员。主要承担工作：溶解性总有机碳分析。

秦伯强（1963—），男，江苏苏州人，博士，研究员。主要承担工作：总体工作部署和技术指导。

杨宏伟（1967—），男，江苏常州人，博士，副研究员。主要承担工作：野外观测指导和水样采集。

李宽意（1971—），男，湖南常德人，博士，研究员。主要承担工作：数据质量控制。

张运林（1976—），男，湖南邵阳人，博士，研究员。主要承担工作：水样采集。

沈睿杰（1985—），男，江苏无锡人，硕士，工程师，实验员。主要承担工作：数据质量评估。

参考文献

- [1] 王苏民, 窦鸿身. 中国湖泊志[M]. 北京: 科学出版社, 1998.
- [2] 孙小静, 秦伯强, 朱广伟. 蓝藻死亡分解过程中胶体态磷、氮、有机碳的释放[J]. 中国环境科学, 2007, 27(3): 341-345.
- [3] 黄钰铃, 陈明曦, 刘德富. 不同氮磷营养及光温条件对蓝藻水华生消的影响[J]. 西北农林科技大学学报: 自然科学版, 2008, 36(9): 93-100.
- [4] 黄祥飞, 陈伟民, 蔡启铭. 湖泊生态调查观测与分析[M]. 北京: 中国标准出版社, 2000.
- [5] 秦伯强, 胡春华. 中国生态系统定位观测与研究数据集—湖泊湿地海湾生态系统卷[M]. 中国农业出版社, 2010.
- [6] 袁国富, 张心昱, 唐新斋, 等. 陆地生态系统水环境观测质量保证与质量控制[M]. 北京: 中国环境科学出版社, 2012.
- [7] 中国生态系统研究网络科学委员会. 陆地生态系统水环境观测规范[M]. 北京: 中国环境科学出版社, 2007.
- [8] RICHARDSON A D, HOLLINGER D Y, SHOEMAKER J K, et al. Six years of ecosystem-atmosphere greenhouse gas fluxes measured in a sub-boreal forest[J]. Scientific data, 2019, 6: 117.
- [9] QIN B Q, ZHU G W, GAO G, et al. A drinking water crisis in Lake Taihu, China: Linkage to climatic variability and lake management[J]. Environ Manage, 2010, 45: 105-112.
- [10] 易娟, 徐枫, 高怡, 等. 2007 年以来环太湖 22 条主要河流水质变化及其对太湖的影响[J]. 湖泊科学, 2016, 28(6): 1167-1174.

- [11] 戴秀丽, 钱佩琪, 叶凉, 等. 太湖水体氮、磷浓度演变趋势(1985-2015 年) [J]. 湖泊科学, 2016, 28(5): 935-943.
- [12] 朱广伟, 秦伯强, 张运林, 等. 2005–2017 年北部太湖水体叶绿素 *a* 和营养盐变化及影响因素[J]. 湖泊科学, 2018, 30(2): 279-295.

论文引用格式

闵岫, 钱荣树, 朱广伟, 等. 2007–2015 年太湖水体理化监测数据集[J/OL]. 中国科学数据, 2020, 5(1). (2019-11-07). DOI: 10.11922/csdata.2019.0037.zh.

数据引用格式

闵岫, 钱荣树, 朱广伟, 等. 2007–2015 年太湖水体理化监测数据集[DB/OL]. Science Data Bank, 2019. (2019-10-28). DOI: 10.11922/sciencedb.833.

A physical and chemical monitoring dataset of Taihu Lake from 2007 to 2015

Min Shen^{1,2}, Qian Rongshu², Zhu Guangwei^{1,2}, Huang Jianming², Qin Boqiang^{1,2*}, Yang Hongwei^{1,2}, Li Kuanyi^{1,2}, Zhang Yunlin^{1,2}, Shen Ruijie^{1,2}

1. Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, P. R. China

2. Taihu Laboratory for Lake Ecosystem Research, Chinese Academy of Sciences, Wuxi 214026, P. R. China

* Email: qinbq@niglas.ac.cn

Abstract: This dataset collates the monthly physical and chemical data of eight long-term monitoring stations in Taihu Lake from 2007 to 2015 by TLLER (Taihu Laboratory for Lake Ecosystem Research, Chinese Academy of Sciences), incorporating total of 16 indicators (13,810 values), including transparency, water depth, water temperature, electrical conductivity, suspended solids, dissolved oxygen, pH, silicate, orthophosphate, nitrite nitrogen, nitrate nitrogen, ammonia nitrogen, chemical oxygen demand, total phosphorus, total nitrogen and dissolved total organic carbon. Based on the Box-whisker plot to test the quality of data, the test results show that of the 16 indicators, the proportions of valid data for 10 indicators are over 95%, except for only one indicator whose valid data account for less than 90%. The collation and publication of this dataset can provide a data basis for exploring the characteristics of ecosystem changes in Taihu Lake, and provide scientific support for the prevention and control of eutrophication and the safety of drinking water.

Keywords: Taihu Lake; eutrophication; physical and chemical indicators; long term monitoring

Dataset Profile

Title	A physical and chemical monitoring dataset of Taihu Lake from 2007 to 2015
Data corresponding author	Qin Boqiang (qinbq@niglas.ac.cn)
Data authors	Min Shen, Qian Rongshu, Zhu Guangwei, Huang Jianming, Qin Boqiang, Yang Hongwei, Li Kuanyi, Zhang Yunlin, Shen Ruijie
Time range	2007–2015
Geographical scope	Taihu Lake
Data volume	169 KB
Data format	*.xlsx
Data service system	< http://www.sciencedb.cn/dataSet/handle/833 >
Sources of funding	Supported by Chinese Ecosystem Research Network of Chinese Academy of Sciences and National Ecosystem Research Network of China, Ministry of Science and Technology.
Dataset composition	This dataset contains 864 pieces of data, and each piece of data contains station code, year, month, monitoring station code, transparency (m), water depth (m), water temperature (°C), electrical conductivity (μs/cm), suspended solids (mg/L), dissolved oxygen (mg/L), pH, silicate (mg/L), orthophosphate (mg/L), nitrite nitrogen (mg/L), nitrate nitrogen (mg/L), ammonia nitrogen (mg/L), chemical oxygen demand (mg/L), total phosphorus (mg/L), total nitrogen (mg/L) and dissolved total organic carbon (mg/L).