

DOI: 10. 20040/j. cnki. 1000-7709. 2023. 20222205

基于函数型数据分析的河流有机碳通量 评估模型及其应用

赵梓屹¹, 刘 凌¹, 闫 峰², 钱 宝³

(1. 河海大学水文水资源学院, 江苏 南京 210098; 2. 南昌大学鄱阳湖环境与
资源利用教育部重点实验室, 江西 南昌 330031; 3. 水利部长江流域委员会水文局, 湖北 武汉 430010)

摘要: 由于总有机碳(TOC)不是常规水文监测数据, 监测频率远低于流量资料, 且常常具有缺失值, 这对评估 TOC 通量和揭示碳在河流的输移过程带来了较大难度。针对不同监测频率下的河流 TOC 通量评价问题, 基于函数型数据分析(FDA), 提出了一种改进的计算模型。首先, 对离散数据集进行函数化处理, 将所有指标均转化为同一评价时域上的光滑连续的函数曲线; 而后通过黎曼积分, 计算评价断面在研究时段的 TOC 通量。对长江朱沱站的评价结果表明, 朱沱站 2018~2020 年 TOC 年通量的均值为 $70.2 \times 10^4 \text{t}$; 朱沱站 TOC 通量年内分配极不均匀, 高度集中在汛期, 特别是夏季。汛期(6~9月)的 TOC 通量年均值为 $43.82 \times 10^4 \text{t}$, 相当于总量的 62.43%; 夏季 TOC 通量年均值为 $34.90 \times 10^4 \text{t}$, 相当于总量的 49.72%。而枯水期(10~5月)的 TOC 通量年均值仅为 $26.37 \times 10^4 \text{t}$, 占全年总量的 37.57%; 冬季的 TOC 通量年均值近 $7.13 \times 10^4 \text{t}$, 占全年总量的 10.15%; 与传统方法相比, 改进后的 TOC 通量模型能更有效地处理缺失值和多指标监测序列不一致问题, 并更准确地评估不同季节和丰枯水期的 TOC 通量变化过程。

关键词: 总有机碳; 通量评估; 函数型数据分析; 长江

中图分类号: [TV11]

文献标志码: A

文章编号: 1000-7709(2023)06-0031-04

1 引言

全球每年通过河流输送入海洋的碳高达 $10 \times 10^8 \text{t}$, 约相当于海洋从大气中所吸收碳总量的 45%^[1]。在河流系统中, 生物固碳的产物主要为有机碳, 因此有机碳是河流水碳研究的重点^[2]。河流中总有机碳(TOC)的特征通常用通量反映, 即某一时段内通过评价断面的 TOC 总量^[3]。FERGUSON R^[4]根据河流年径流量与碳浓度的乘积估算 TOC 通量数值。但由于径流数据和 TOC 浓度随时间的变化很大, 因此估算结果精度相对较低。随后, LIU D 等^[5]建立了河流 TOC 与流域特征变量(流域人口密度、降水和河流悬浮泥沙浓度)之间的关系, 以估计中国进入边缘海洋的河流 1953~2016 年的 TOC 通量, 但这种方法对人口密度、降雨和泥沙的历史数据监测时长和精度要求过高。2014 年, ZHANG L 等^[6]对传统方法进行了改进, 利用碳浓度与月均流量乘积计

算长江三峡的 TOC 通量。值得注意的是, 随着水文自动化系统的发展, 目前我国主要控制性水文断面已可实现流量的逐日甚至逐小时监测。然而 TOC 不是常规监测数据, 监测频率低, 且经常出现缺失值。因此, 针对具有不同监测频率和缺失值的水碳数据, TOC 通量计算依然是河流碳循环研究的难点之一。与传统技术相比, 函数型数据分析(FDA)不仅能较好地反映评价指标的动态变化, 且能有效地处理数据集中的缺失值和监测序列不一致问题, 并已在经济学、管理学和生态学动态评价中得到了广泛应用^[7]。鉴此, 本文基于函数型数据分析理论, 提出一种改进的河流 TOC 通量计算模型, 并以长江朱沱站为例, 计算 TOC 的年际通量, 分析不同季节和丰枯水期的 TOC 通量变化过程, 以期为长江上游的水碳循环研究提供依据。

2 研究区域与方法

收稿日期: 2022-10-20, **修回日期:** 2022-11-18

基金项目: 国家自然科学基金项目(52069012)

作者简介: 赵梓屹(1998-), 女, 硕士研究生, 研究方向为生态水利, E-mail: bingtanghuluwa@163.com

通讯作者: 闫峰(1988-), 男, 副教授, 研究方向为水文水资源, E-mail: yfmilan@163.com

2.1 研究区域

朱沱站为长江上游干流最重要的控制性水文站和我国自动化监测程度最高的水文站之一,目前已实现流量的逐日监测。但如图 1 所示,由于 TOC 并非常规指标,监测间隔一般为 2 个月,部分月份还有缺失值,这给朱沱站的碳通量计算带来了极大挑战。

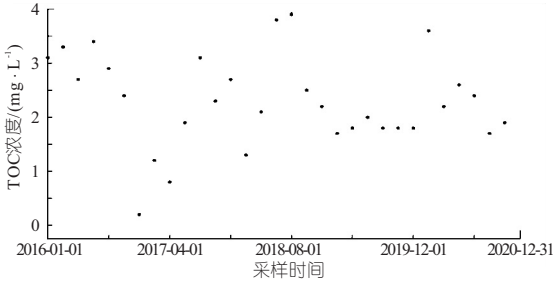


图 1 朱沱站 TOC 实测值序列

Fig. 1 Sequence of measured TOC values at Zhutuo Station

2.2 传统 TOC 通量计算方法

TOC 通量主要根据河流径流量、TOC 碳浓度、采样时间的乘积来估算。设评价期共包括 n d, 则该评价期的 TOC 通量 C 为:

$$C = \lambda \sum_{i=1}^n x_i q_i \quad (1)$$

式中, λ 为量纲转化系数,取 0.086 4; x_i 为第 i d 的水体 TOC 浓度, mg/L; q_i 为第 i d 的流量, m^3/s 。

式(1)的环境意义明确,然而目前我国 TOC 不是常规监测数据,经常出现缺失值的问题,很难直接通过式(1)计算碳通量。为此,假定碳含量在两次采样间隔期内保持恒定,而后可根据评价时段内的径流量与碳平均浓度的乘积计算碳通量^[4]。设评价期内的径流量为 W , 碳平均浓度为 x , 则该评价期的 TOC 通量 C 为:

$$C = \lambda W x \quad (2)$$

式(2)基本解决了各指标监测频率不一致的问题。但对于较为稀疏的监测序列,常常会出现监测数据对评价期的代表性不强的问题。

2.3 基于 FDA 的 TOC 通量计算方法

函数型数据分析(FDA)是传统离散型统计模型向连续时域的拓展和深化。函数型数据分析的内在机制是“化数为形”,即首先根据离散的监测序列构建研究指标在连续时域上随时间变化的函数,而后通过一系列基于函数或泛函的运算,提取目标函数的相关特征信息。

其中离散监测序列的函数化是指将离散的监测数据集 $X = \{x(t_1), x(t_2), \dots, x(t_i), \dots, x(t_s)\}$ 转化为时域上的连续函数 $x(t)$, 它的基本原理见图 2。图 2 中, $\theta_j(t)$ 为基函数; h_k 为 $t_{k+1} - t_k$ 。

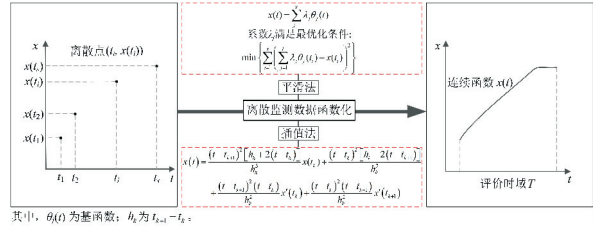


图 2 监测数据函数化原理示意图

Fig. 2 Schematic diagram of principle of functionalization

函数化的方法包括插值法和平滑法两种。当监测数据存在异常值或误差较大时一般使用平滑法;数据误差较小时通常采用插值法。平滑法是指通过基函数的线性组合拟合 $x(t)$, 得到的函数曲线 $x(t)$ 通常并不精确通过每一个监测数据点。插值法则是在离散数据序列的基础上补插连续函数 $x(t)$, 使 $x(t)$ 严格精确通过每一个监测数据点。在 FDA 中,最常用的插值法是三次样条插值。

函数型数据分析的最大优势在于经函数化处理后,所有指标均转化为同一评价时域上光滑连续的函数曲线,从而有效地解决了缺失值和多指标监测时刻不一致的问题。此外,与 Excel 中的趋势线不同,经函数化得到的曲线有明确、具体的数学表达式,可通过后续计算进一步揭示指标的统计特征。

将水体的 TOC 浓度函数记作 $x(t)$, 将流量函数记作 $q(t)$, 对于研究时域 $T^* \in T$, 根据黎曼积分可知, T^* 时段内的 TOC 通量 C 为:

$$C = \lambda \int_{t \in T^*} x(t) q(t) dt \quad (3)$$

由于监测资料已进行了有效性监测,可认为实测值的误差较小,因此使用三次样条插值法进行监测数据函数化。相应地碳通量的计算流程见图 3。

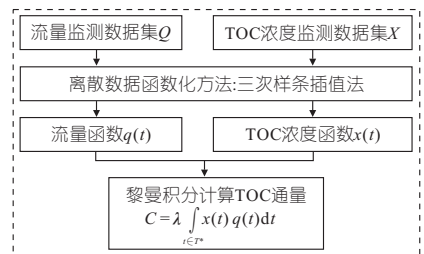


图 3 长江朱沱站 TOC 通量计算模型

Fig. 3 TOC flux calculation model for Zhutuo Station, Yangtze River

3 评价结果与分析

3.1 流量函数与 TOC 浓度函数

通过三次样条函数,拟合朱沱站的流量函数和 TOC 函数见图 4。

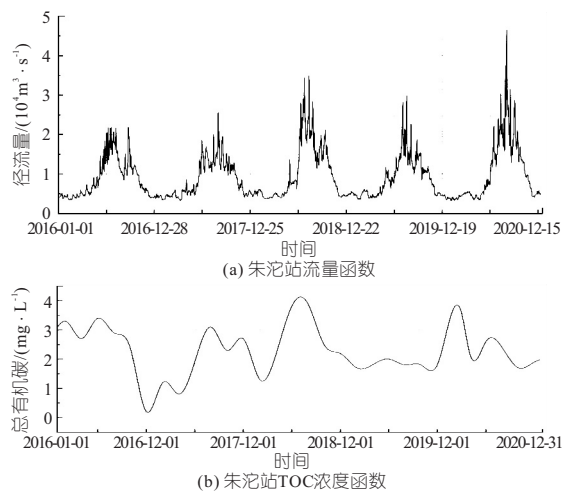


图 4 朱沱站流量函数、TOC 浓度函数

Fig. 4 Flow function and TOC concentration function at Zhutuo Station

由图4(a)可知,朱沱站的多年平均流量为 9 172.51 m³/s,但径流的年内分配极不均匀,汛期

(6~9月)的平均流量为 15 342.57 m³/s;而枯水期(10月至次年5月)的平均流量为 6 079.87 m³/s,仅相当于汛期的 39.63%。

由图4(b)可知,朱沱站的 TOC 多年平均浓度为 2.27 mg/L,其中 2019 年的浓度整体显著低于平均值。朱沱站 TOC 浓度函数的年内分配极不均匀,整体上呈先上升再下降的趋势,局部会出现小幅下降或上升。同时,朱沱站 TOC 浓度呈一定的季节变化,但无非常明显的周期性。

对照图 1 与图 4(b),发现经函数化处理后,流量和 TOC 浓度为同一研究时域(2016~2020 年)上关于时间的连续函数,从而有效地解决了长江朱沱站流量和 TOC 监测时刻不一致的问题。

3.2 基于 FDA 的 TOC 通量评估

根据基于 FDA 的 TOC 通量评估模型,计算朱沱站 2016~2020 年不同季节和年度的 TOC 通量见表 1。

表 1 基于 FDA 的 TOC 通量评估结果

Tab. 1 Results of FDA-based TOC flux assessment

| 时期 | 2016 年 | | 2017 年 | | 2018 年 | | 2019 年 | | 2020 年 | |
|-----|---------------------------|-------------|---------------------------|-------------|---------------------------|-------------|---------------------------|-------------|---------------------------|-------------|
| | TOC 通量 /10 ⁴ t | 占全年总量 百分比/% | TOC 通量 /10 ⁴ t | 占全年总量 百分比/% | TOC 通量 /10 ⁴ t | 占全年总量 百分比/% | TOC 通量 /10 ⁴ t | 占全年总量 百分比/% | TOC 通量 /10 ⁴ t | 占全年总量 百分比/% |
| 春季 | 13.35 | 18.69 | 4.60 | 7.60 | 11.18 | 11.54 | 8.73 | 17.34 | 9.25 | 12.89 |
| 夏季 | 35.24 | 49.33 | 29.00 | 47.99 | 55.81 | 57.57 | 20.18 | 40.07 | 34.28 | 47.76 |
| 秋季 | 15.56 | 21.78 | 21.40 | 35.41 | 23.80 | 24.55 | 14.37 | 28.53 | 18.57 | 25.87 |
| 冬季 | 7.29 | 10.21 | 5.44 | 9.00 | 6.16 | 6.35 | 7.07 | 14.05 | 9.68 | 13.48 |
| 汛期 | 44.19 | 61.85 | 37.76 | 62.48 | 66.75 | 68.85 | 26.74 | 53.11 | 43.66 | 60.82 |
| 枯水期 | 27.26 | 38.15 | 22.67 | 37.52 | 30.20 | 31.15 | 23.61 | 46.89 | 28.12 | 39.18 |
| 全年 | 71.45 | 100.00 | 60.44 | 100.00 | 96.95 | 100.00 | 50.35 | 100.00 | 71.79 | 100.00 |

由表 1 可知,长江朱沱站 2018~2020 年 TOC 年通量的均值为 70.2×10⁴t,但年内分配极不均匀。汛期(6~9月)的 TOC 通量年均值为 43.82×10⁴t,相当于总量的 62.43%;而枯水期(10月至次年5月)的 TOC 通量年均值仅为 26.37×10⁴t,相当于总量的 37.57%。在四个季度中,夏季 TOC 通量年均值为 34.90×10⁴t,相当于总量的 49.72%;秋季 TOC 通量年均值为 18.74×10⁴t,相当于总量的 26.70%;春季 TOC 通量年均值为 9.42×10⁴t,相当于总量的 13.42%;冬季 TOC 通量年均值为 7.13×10⁴t,相当于总量的 10.15%。

从枯水期和汛期来看,TOC 通量高度集中在汛期;而从四季来看,TOC 通量显著集中在夏季,结合流域的水文气象调查可知,造成朱沱站 TOC 通量年内分布不均匀的原因主要为:①受季风气候的影响,流域内汛期降雨集中,导致土壤有机质的侵蚀、陆地植物残体等陆源 TOC 输入增多。②夏季气温上升,光照增强,水生植物的固碳能力

提高。相应的水中内源性 TOC 也随之增加。

3.3 基于传统方法的 TOC 通量评估

假定 TOC 含量在两次监测间隔间保持不变,采用传统方法(式(2))计算 2016~2020 年长江朱沱站的 TOC 通量分别为 70.42×10⁴、53.06×10⁴、83.24×10⁴、49.92×10⁴、74.98×10⁴t。

对比表 1 可发现,与传统方法相比,基于函数型数据分析的碳通量计算得到的年碳通量有一定程度的区别,但总体上无数量级的差别,且结果较接近,表明函数型数据分析在 TOC 评估中可行、准确,得到的数据可信。

由于碳浓度监测数据的缺失,逐日计算的传统碳通量计算方法(式(1))并不适用。应用年径流量与碳平均浓度的乘积计算碳通量的估算方法可从年际方面分析碳通量的变化,而基于 FDA 的方法可更进一步地分析年内碳通量的变化特征,并能剖析碳通量变化与季节之间的关联,汛期、枯水期的碳通量的分布,更清晰地阐明流域碳通量的影响因素,如降雨输入、温度等。

4 结 论

a. 改进的 TOC 通量模型首先将离散数据集函数化处理,所有指标均转化为同一评价时域上光滑连续的函数曲线,从而有效地解决了实测数据缺失及多指标监测时刻不一致的问题;而后通过黎曼积分,计算研究时段的 TOC 通量。

b. 长江朱沱站 2018~2020 年 TOC 年通量的均值为 $70.2 \times 10^4 \text{ t}$ 。其中 2018 年 TOC 年通量最大,2019 年最小。TOC 通量的年内分配极不均匀,从枯水期和汛期来看,TOC 通量高度集中在汛期(6~9 月),汛期的 TOC 通量年均值相当于总量的 62.43%。而从四季来看,TOC 通量显著集中在夏季,夏季 TOC 通量年均值最大,相当于总量的 49.72%。

c. 与传统方法相比,利用基于 FDA 理论改进的模型的年碳通量计算结果与其相近,这表明该模型的结果可信。此外,改进模型还能更有效地处理缺失值和多指标监测序列不一致的问题,并能更准确地评估不同季节和丰枯水期的 TOC 通量变化过程。

d. 对于流量或碳含量监测误差较大的测站,当采用平滑法进行函数拟合时,使用不同基函数构建的曲线往往有一定差异。因此,在后续基于函数型数据分析进行 TOC 通量的研究中,基函数的选择仍值得进一步探讨。

参 考 文 献:

- [1] REIMAN J, XU Y J. Dissolved carbon export and CO_2 outgassing from the lower Mississippi River—Implications of future river carbon fluxes [J]. *Journal of hydrology*, 2019, 578: 124093.
- [2] AITKENHEAD J A, MCDOWELL W H. Soil C: N ratio as a predictor of annual riverine DOC flux at local and global scales [J]. *Global biogeochemical cycles*, 2000, 14(1): 127-138.
- [3] LI M X, PENG C H, WANG M, et al. The carbon flux of global rivers: A re-evaluation of amount and spatial patterns [J]. *Ecological indicators*, 2017, 80: 40-51.
- [4] FERGUSON R I. Accuracy and precision of methods for estimating river loads [J]. *Earth surface processes and landforms*, 1987, 12(1): 95-104.
- [5] LIU D, BAI Y, HE X, et al. Changes in riverine organic carbon input to the ocean from mainland China over the past 60 years [J]. *Environment international*, 2020, 134: 105258.
- [6] ZHANG L, XUE M, WANG M, et al. The spatio-temporal distribution of dissolved inorganic and organic carbon in the main stem of the Changjiang (Yangtze) River and the effect of the Three Gorges Reservoir [J]. *Journal of geophysical research: Biogeosciences*, 2014, 119(5): 741-757.
- [7] 苏为华, 孙利荣, 崔峰. 一种基于函数型数据的综合评价方法研究 [J]. *统计研究*, 2013, 30(2): 88-94.

Evaluation Model of River Organic Carbon Flux Based on Functional Data Analysis and Its Application

ZHAO Zi-yi¹, LIU Ling¹, YAN Feng², QIAN Bao³

(1. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China; 2. Key Laboratory of Poyang Lake Environment and Resource Utilization, Ministry of Education, Nanchang University, Nanchang 330031, China; 3. Hydrological Bureau of Yangtze River Basin Commission, Ministry of Water Resources, Wuhan 430010, China)

Abstract: Since total organic carbon (TOC) is not a routine hydrological monitoring data, the monitoring frequency is much lower than that of flow data and it often has missing values, which makes it difficult to evaluate TOC flux. This study proposes an improved computational model based on functional data analysis (FDA) for the evaluation of riverine TOC fluxes. First, the discrete data set is functionalized, and all the indicators are transformed into s function curves in the same evaluation time domain. And then the TOC flux of the evaluation section during the study period is calculated by Riemann integral. The results of the TOC evaluation at the Zhutuo station on the Yangtze River show that, The average value of the annual TOC flux at Zhutuo Station from 2018 to 2020 is 702 000 t; The intra-annual distribution of TOC fluxes is extremely uneven, with a high concentration in the flood season, especially in summer; The annual average value of TOC flux during the flood season (June to September) is 438 200 t, equivalent to 62.43% of the total; The annual average TOC flux in summer is 349 000 t, equivalent to 49.72% of the total; While the annual average value of TOC flux during the dry season (October to May) is 263 700 t, equivalent to 37.57% of the total; The annual average TOC flux in winter is 71 300 t, equivalent to 10.15% of the total; Compared with the traditional method, the improved TOC flux model is also able to deal with missing values and inconsistencies in multi-indicator monitoring sequences more effectively, and to accurately assess the course of TOC fluxes over different seasons and periods of abundance and depletion.

Key words: total organic carbon; flux; functional data analysis; Yangtze River