

一种基于拓扑网络的水质模型的建立

周文春*, 易玉敏, 陶灵娟

云南省生态环境工程评估中心, 云南 昆明

收稿日期: 2022年9月4日; 录用日期: 2022年10月3日; 发布日期: 2022年10月10日

摘要

出于对时间成本、经济成本的考虑, 经济欠发达的区域在对大多数流域环境管理的过程中, 往往只关注污染源强的问题, 忽略了时间、空间对于流域水质的相互关系, 因此并不能回答不同的源强在不同人为干预的影响下水环境质量的变化。本文以环境质量与人类活动之间的响应关系为研究对象, 构建了一套简单、灵活、有效的水质模型, 从而实现在有限的输入条件下流域水质的快速模拟, 为科学准确的环境管理决策提供行之有效的支撑。

关键词

大尺度流域水质模拟, 模拟模型, 拓扑网络, 模型优化, 环境管理

The Establishment of a Water Quality Model Based on Topology Network

Wenchun Zhou*, Yumin Yi, Lingjuan Tao

Yunnan Appraisal Center for Ecological and Environmental Engineering, Kunming Yunnan

Received: Sep. 4th, 2022; accepted: Oct. 3rd, 2022; published: Oct. 10th, 2022

Abstract

For the consideration of time cost and economic cost, the economically underdeveloped regions tend to focus only on the intensity of pollution sources, ignoring the spacial and temporal interrelation with water quality in environmental management of most river basins, so they can not answer the impact of different sources on the quality of water environment under the influence of different man-made interventions. Taking the response relationship between environmental quality and human activities as the research object, this paper constructs a simple, flexible and effective water quality model, so as to realize the rapid calculation of watershed water quality under

*通讯作者。

the condition of limited data entry and provide effective support for scientific and accurate environmental management decision-making.

Keywords

Water Quality Simulation of Large-Scale Watershed, Simulation Models, Topology Networks, Model Optimization, Environmental Management

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

河流、湖泊的水质是一个复杂的过程，其与污染源强、时间、空间、人类活动、产业布局、经济发展等因素之间有很强的依赖性和相关性。水质模型的运用就是要在大数据时代结合爆炸式的信息与严谨的科学规律，做到最精细的决策，从而实现经济与环境的平衡。在当今社会，流域环境管理以及水污染防治对于水质模型的需求越来越强烈，然而现阶段较为常见的水质模型如 Mike, SWAT 等需要构建复杂的输入系统，这大大限制了水质模型在环境管理中的运用和推广，也增加了模型的不确定性和校验难度。同时，由于模型体量太过于庞大，这些模型往往消耗过久的模拟时间，限制了这些模型在较大流域中的运用。根据 Ingrid Keupers 等人的研究[1]，在 PC 端利用 MIKE11-ECOLab 来模拟不到 400 km² 的汇流区域时耗时约 4 天。河网水动力模型是水质模型的基础[2]，河网二叉树的编码方式早已在大尺度流域中有所运用[3] [4]。付立冬基于 WebGIS 技术采用一维水质模型实现了蒲河流域基于河流流域的容量总量计算及排污许可分配[5]，卢诚等人基于 EFDC 模型建立了十堰市神定河的水动力水质模型[6]，两种水质模型的实现方式对于单一河道的水质模拟具有较好的模拟效果，但这两种方法仅适用于单一河道，对于复杂河网并不适用。薛野基于三级解法的分级解法建立了河网一维水动力模型[7]，能够较好地模拟复杂河网的水质，但该模型仅考虑了单一污染源的输入，并未考虑大流域范围内不同污染源的输入过程。为了提高模型模拟的效率，同时减小模型的复杂程度，我们基于水质一维模型构架了更简便快捷的大流域水质模型。

2. 研究方法

河网水质模拟模型的构建以一维水质模型为基础，分为河网拓扑构建、污染源拓扑关系构建、污染源核算、河网流量初始化以及模拟运算几个步骤，模型构建的原理如图 1 所示。1) 河网拓扑构建以 GIS 水文分析为基础，设置一定的阈值后定义并提取河网。李铁键等人提出了一种基于二叉树理论并以二元形式表示的河网编码方法[8]，本文河网拓扑网络的构建与该方法相似，不同之处在于本文所构建拓扑网络除了河网本身还包括与河道关联的污染源的拓扑结构。2) 污染源拓扑关系构建以构建好的河网拓扑以及污染源位置为基础，计算出距离污染源最近计算单元后在该计算单元处设置概念化的“排污口”。面源污染对水质的影响很突出[9] [10]，不同于经济发达地区面源污染通过雨水排口进入水体，云南地形起伏大、管网设施不完善，大部分面源污染无固定排口。为凸显面源污染的影响，除了常规点源污染源，城镇生活源、农村生活源等污染源以所在城镇或乡村为源强概化为点源，农田径流以地块质心概化为点源。3) 污染源强的核算分为点源和面源，点源污染源强数据获取途径较多，面源污染源强与人类活动、

降雨情势、下垫面类型之间有着直接的相关性。径流曲线数模型(SCS-CN 模型) [11]和长期水文影响评价模型(L-THIA 模型) [12]是连接降雨、人类活动和下垫面类型的关键。4) 通常在水质模型的运用过程中, 支流河道通过概化的形式简化计算流程, 河道流量的赋值只能逐段进行或者求解圣维南微分方程[13] [14], 这样的实现方式并不利于在大流域水质模拟中运用。河网流量初始化以 GIS 汇流分析所得各栅格累计流量以及流域实际出流点流量为基础分配。图遍历是求解拓扑网络常用的手段, 通常包括深度优先算法和广度优先算法[15] [16]。广度优先算法指的是从图的一个未遍历的节点出发逐层遍历相邻节点, 遵从先进先出的原则通常采用队列的形式实现; 深度优先算法从图的一个未遍历节点出发, 沿着某个分支遍历到叶子后再遍历另一分支, 深度优先算法遵从先进后出的原则, 通常通过栈的形式实现。本文对河网的遍历采用深度优先算法中的中序遍历形式实现。模拟运算过程以河网拓扑构建、污染源拓扑构建、河网流量初始化三个准备工作的结果为输入, 以中序遍历的方式实现河网水质的模拟运算。

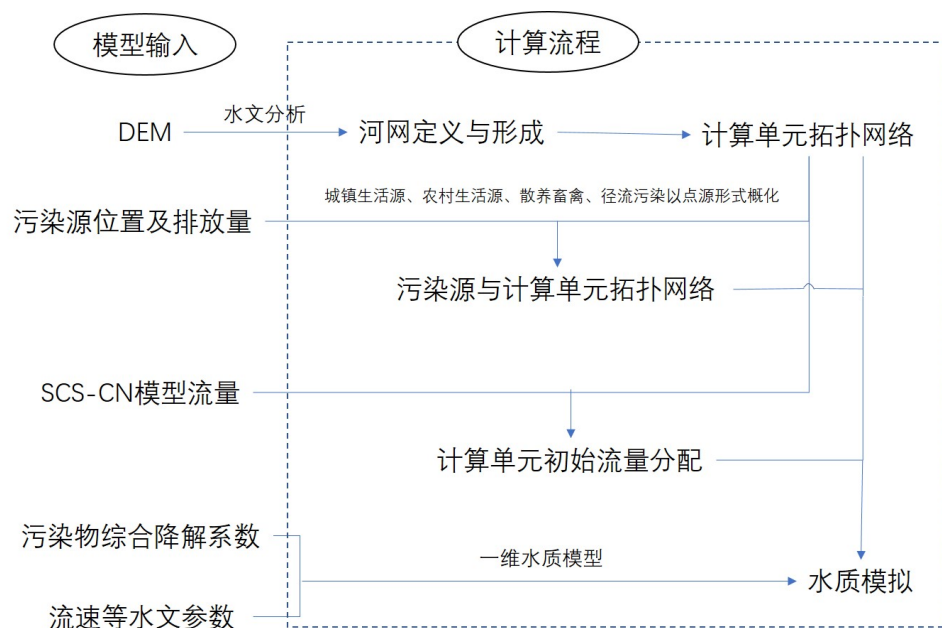


Figure 1. Schematic diagram of model construction principle
图 1. 模型构建原理示意图

2.1. 一维水质模型

在忽略离散作用时, 描述河流污染物一维稳态衰减规律的微分方程为:

$$u \frac{dc}{dx} = -Kc \quad (2-1)$$

将 $u = \frac{dx}{dt}$ 代入, 得到

$$\frac{dc}{dt} = -Kc \quad (2-2)$$

积分解得

$$C = C_0 e^{-\frac{Kx}{86400u}} \quad (2-3)$$

式中： u ——河流断面平均流速，m/s；
 x ——沿程距离，m；
 K ——综合降解系数，1/d；
 C ——沿程污染物浓度，mg/L；
 C_0 ——前一个节点后污染物浓度，mg/L。

2.2. 河网拓扑构建

河网拓扑可以利用数字高程模型(DEM)数据提取，并以一定的编码方式形成计算机可操作的拓扑模型。GIS 水文分析工具以在计算区域模拟无限降雨的方式形成汇流，因而所形成的河道并不一定是真实的河道，需要设置一定的汇流面积为阈值定义河道。DEM 所提取河网本身已经包含拓扑信息，为了在关联污染源的时候更易操作，需要将拓扑关系以拓扑表的形式存储。存储拓扑关系的最小单元也即计算单元，计算单元可以以任意精度将河网进行切割，然而为了提高计算的效率，并为以 DEM 计算形成的汇流栅格分配初始流量做准备，需要以 DEM 的分辨率为最小切割单元进行切割。如图，假设 DEM 分辨率为 90 m，则以 90 m 为一个单元将河网分为 11 个计算单元，其中 1、7、9 号单元为源头单元，11 号为出流单元，2 号单元为 1 号单元下游，同时也是 3 号单元的上游，由于支流的存在，10 号单元存在 4 号和 5 号两个上游，以此类推可以形成图中所示的河网拓扑关系表。

2.3. 污染源拓扑关系构建

长久以来，在规划或者决策过程中往往仅考虑所管理区域污染物的排放量、入河量，然而对于污染物入河以后对河网水质的影响欠缺系统性的分析。与工业源、集中式污水处理厂等具有实际排放口的点源污染不一样，面源污染没有固定的排放口，为了更加系统性地分析面源污染带来的影响，需要将面源污染概化为点源纳入到模型计算中。如图 2 所示，若 1 号、3 号污染源代表农村生活源，村落往往傍水而建，则可用该村落产生的生活污水推算出入河量，并将该村落投影到附近河道作为一个“虚拟”排口。

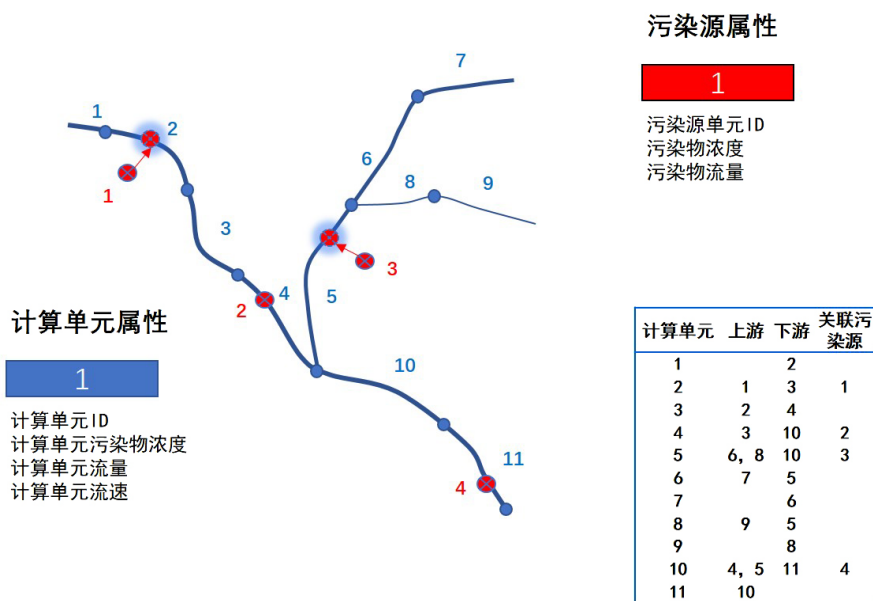


Figure 2. Schematic diagram of topological relationship construction of river network and pollution sources

图 2. 河网以及污染源拓扑关系构建示意图

排口设置完成后需要将污染源与计算单元关联, 将被关联的污染源编入拓扑关系表, 至此拓扑关系表相当于一个索引表, 以上游、下游、计算单元以及污染源的 ID 为键值实现计算单元的关联。计算单元和污染源数据应分别封装到不同的类内, 其中计算单元类应包含计算单元 ID、计算单元污染物浓度、计算单元流量和计算单元流速四个属性, 污染源类应包含污染源 ID、污染物浓度和污染物流量三个属性。

2.4. 污染源源强的核算

点源污染源可以通过调查统计的方式获得污染源强的数据, 面源污染源的源强与流域范围内人类活动及产汇流的过程相关。面源污染源强的确定分两部分进行, 首先通过年平均降雨量在区域内的分布情况以及下垫面的覆盖类型利用 SCS-CN 模型计算各计算栅格累计流量, 再根据各下垫面的平均事件浓度计算在相应降雨事件条件下的污染源强。面源污染仅考虑 COD 以及 TP 的污染负荷, $\text{NH}_3\text{-N}$ 负荷不计。

径流曲线数模型(SCS-CN, Soil Conservation Service Curve Number Model)是由美国农业部自然资源保护局开发的确定特定区域降雨事件直接径流近似量的有效方法。

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \quad (2-4)$$

式中: Q ——累计径流量, mmH_2O ;

P ——降雨量, mmH_2O ;

I_a ——初损, 包括: 地表蓄水、截留和产流前的下渗, mmH_2O ;

S ——潜在最大保留量, mmH_2O 。

初损 I_a 通常按 $0.2S$ 估算, 则径流量为:

$$Q = \begin{cases} 0, & P < 0.2S \\ \frac{(P - 0.2S)^2}{P + 0.8S}, & P \geq 0.2S \end{cases} \quad (2-5)$$

$$S = 25.4 \left(\frac{1000}{CN} - 10 \right) \quad (2-6)$$

式中 CN 为径流曲线数, 是反映降雨前流域特征的一个综合参数, 它与流域前期土壤湿润程度、坡度、植被、土壤类型和土壤利用现状等有关。

利用美国土壤保持局长期水文影响评价模型(Long-Term Hydrologic Impacts Assessment of landuse changes, L-THIA)对非点源负荷估算。非点源污染包括城市径流、农村径流、农田径流。

$$L = R * C * A * 10^{-6} \quad (2-7)$$

式中: L ——污染负荷量, kg ;

R ——径流量, mm ;

C ——事件平均浓度(Event Mean Concentration, EMC), mg/L ;

A ——集水区面积, m^2 。

2.5. 河网流量初始化

水是污染物降解和扩散的媒介, 只有让河网中所有计算单元的计算单元流量属性赋值后河网水质模拟才可以展开。通常以圣维南方程为基础, 推导单一河道下一维水动力学数值解, 这无疑增加了水质模拟的复杂程度。为了减少模型运行的成本和复杂度, 可以利用 GIS 汇流分析计算计算范围内所有像元点

会累计流量与流域出流点累计流量的占比，从而依据出流点实际流量推演河网各计算单元的流量。由于GIS 汇流分析时假设降雨无限大，因而累计流量与流域面积成正比，累计流量之比也即汇流面积之比。如图3所示，11号计算单元端点的汇流面积为 A_{11} ，4号计算单元汇流面积为 A_4 ，7号计算单元的汇流面积为 A_7 ，则11号、4号、7号计算单元的初始流量分别为 Q_{11} ， $Q_{11} \cdot A_4 / A_{11}$ ， $Q_{11} \cdot A_7 / A_{11}$ ，以此类推可以将整个河网赋值初始流量。

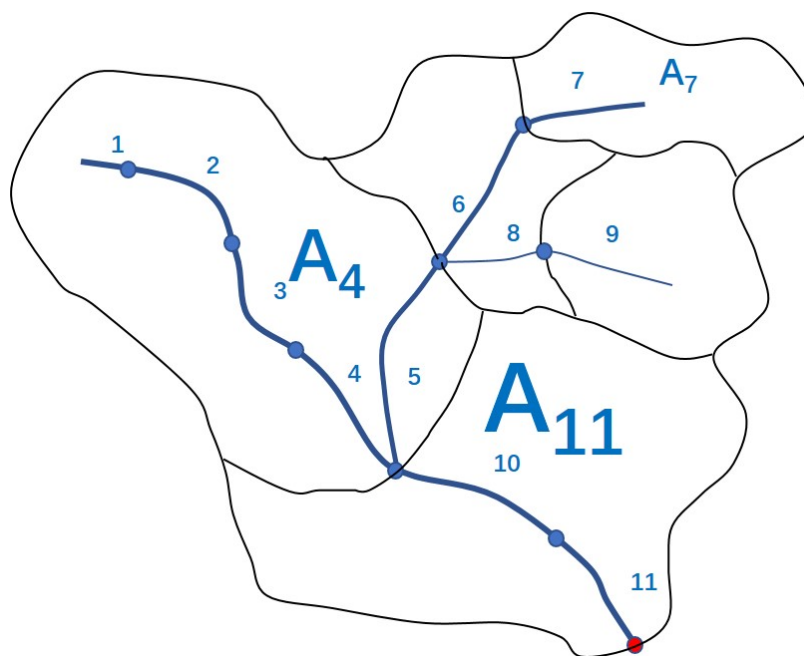


Figure 3. Schematic diagram of flow distribution principle of river network
图3. 河网流量分配原理示意图

2.6. 模拟运算

栈是一种数据结构，遵循先进后出的规则，利用栈可以实现对河网树形结构的遍历。当沿着河网计算计算单元的污染物浓度时，若沿途不存在支流或者计算单元上游已经完成所有计算时，通过检索拓扑关系表查找对应下游计算单元顺序计算即可完成一连串的污染物浓度计算。若出现支流或者上游存在未完成计算的单元时需要将未完成计算的支流或者计算单元推入栈中等待计算，直到推至河流源头时再逐一从栈中取出计算单元进行计算。计算单元的污染物浓度与上游计算单元污染物浓度之间符合一维动力学关系，其中源头污染物的浓度可以是上游控制单元来水的污染物浓度或者水体本地污染物浓度，当计算单元处排口设置，则该计算单元的污染物浓度等于一维水质模型运算结果与排口污染物均匀混合后的浓度。

如图4所示，依据前述计算规则，若计算起点选择3号计算单元，则由于3号计算单元并非所在河道源头，上游存在未完成计算的计算单元，则依次将3号、2号计算单元推入栈中，2号计算单元上游单元为1号，1号已至源头，则依次按照1→2→3的顺序从栈中取出元素并依照一维水质模型计算各计算单元水质，其中2号单元上有1号污染源，故将1号污染源与2号计算单元均匀混合后作为2号计算单元污染物浓度继续参与计算。计算至4号计算单元时，上游计算单元皆以完成计算，且不存在支流，则继续往下游计算，4号计算单元下游10号单元存在支流5号计算单元，故而将10号计算单元及其未完成支流推入栈中等待计算。如此整个过程的计算顺序应该是1→2→3→4→9→8→7→6→5→10→11。

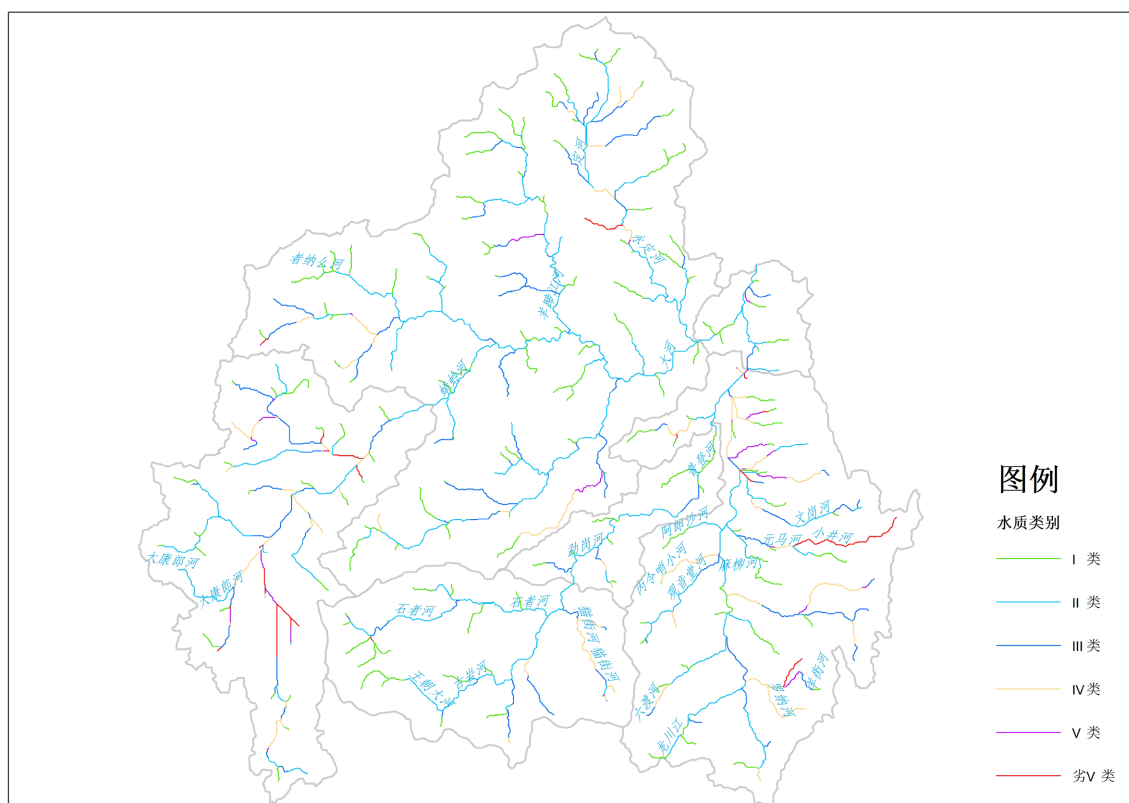


Figure 9. Water quality simulation results of each reach in the basin
图 9. 流域内各河段水质模拟结果

4.2. 模型校验

河网水质的模拟同样利用 Python 代码实现, 由于 COD 数据缺失, 此处仅以 TP 和 $\text{NH}_3\text{-N}$ 的实测浓度来校验模型, 校验结果如图 10 所示。分别在蜻蛉河、龙川江和勐岗河(又名普登河)选取断面计算 2017 年 TP 和 $\text{NH}_3\text{-N}$ 的年平均浓度, 共选取 5 个断面。TP 和 $\text{NH}_3\text{-N}$ 实测值和模拟结果之间具有较好的相关性, 其中 TP 皮尔逊相关系数为 0.91, $\text{NH}_3\text{-N}$ 为 0.98。 $\text{NH}_3\text{-N}$ 实测值和模拟值之间的线性拟合结果较好, 斜率接近与 1, 且 Nash-Sutcliffe 系数为 0.95, 表明模型对 $\text{NH}_3\text{-N}$ 具有较好的模拟能力; TP 实测值与模拟值之间的线性拟合结果相比较于 $\text{NH}_3\text{-N}$ 较差一些, 拟合曲线斜率约为 1.36, 略大于 1, Nash-Sutcliffe 系数为 0.82, 表明一维水质模型对 TP 具有一定的模拟能力, 但模拟的精度还不够高。

4.3. 模型运行效率

模型运行的时间与河网密度以及河网长度呈正相关。模型运行效率见图 11, 根据模型运行时间与河网上游河段数之间的关系可以将河段分为 5 个模式, 其中模式 1 河段数量占比最高, 占总河段数量的 92.4%, 模型运行时间随上游河段数量的增长模式为线性增长, 增长速率最大; 模式 2 模型运行时间随上游河段数量的增长无明显相关性, 该部分河段数量占总河段数量的 0.58%, 主要为无支流汇入的干流河段; 模式 3 模型运行时间随上游河段的增长呈指数型增长, 该部分河段占总河段数量的 1.97%, 空间分布形式无明显规律; 模式 4 和模式 5 增长方式均为线性, 增长速率略小于模式 1, 该部分河段分别占总河段数量的 3.67%和 1.34%, 空间分布无明显规律。若以 Shreve 法河网级别作为河网密度的表征形式, 模型运行时间与 Shreve 河网级别之间的关系和模型运行时间与上游河段数量之间的关系一致。

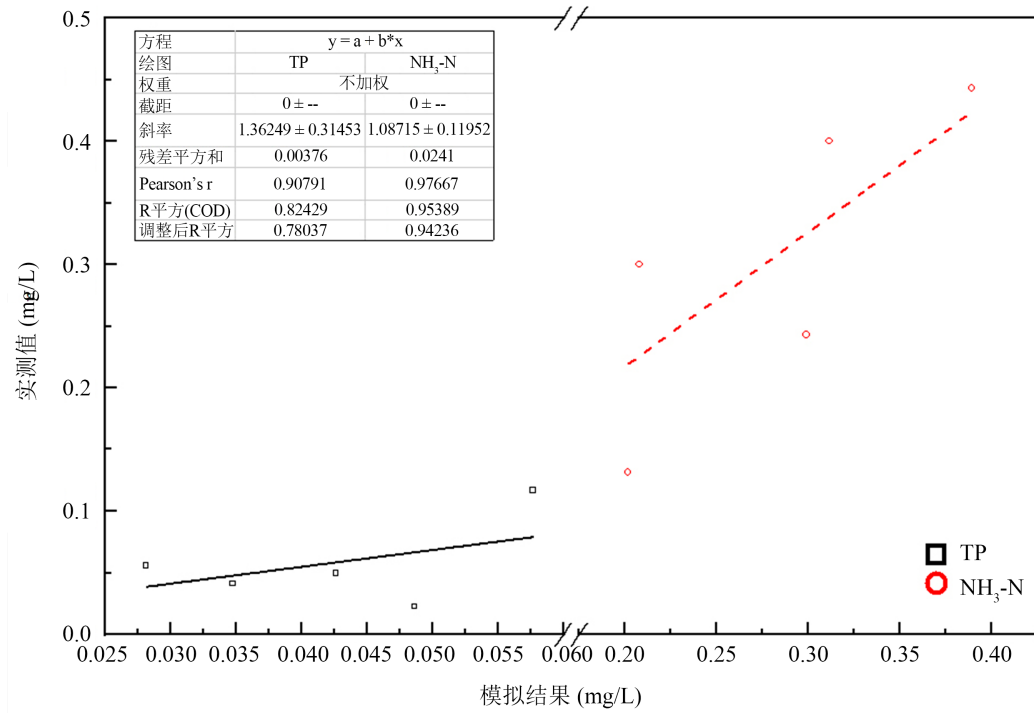
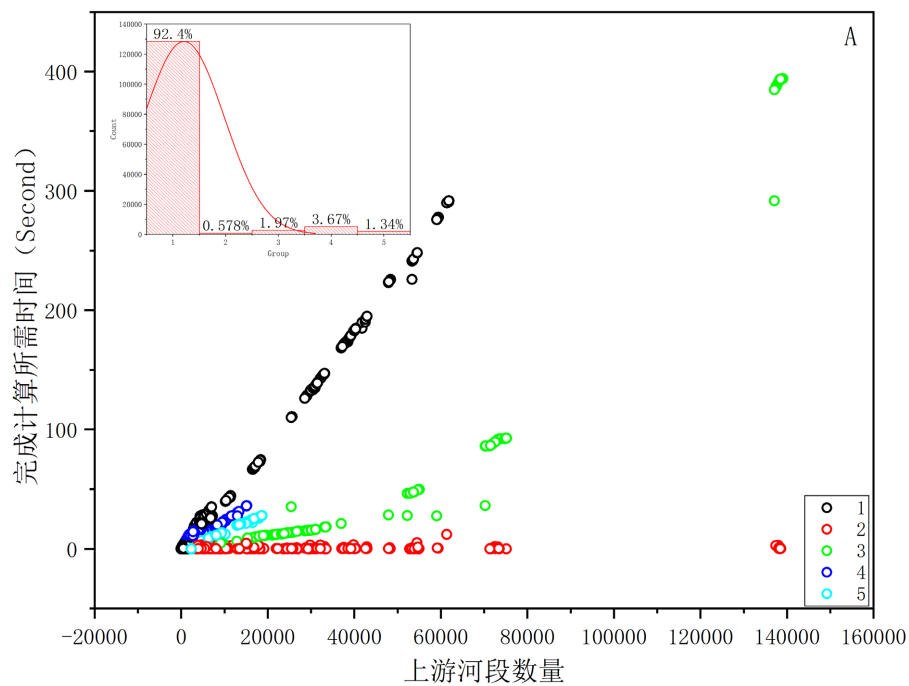


Figure 10. Regression analysis between model running results and measured results
 图 10. 模型运行结果与实测结果回归分析

综上，我们利用模式 1 表征模型的复杂程度。

$$Z = 0.0068X + 4.87 \times 10^{-5}Y - 0.26 \quad (R^2 = 0.997) \quad (4-1)$$

式中 Z 为模型运行时间，单位 s；X 为河网 Shreve 级别；Y 为上游河网长度，单位 m。



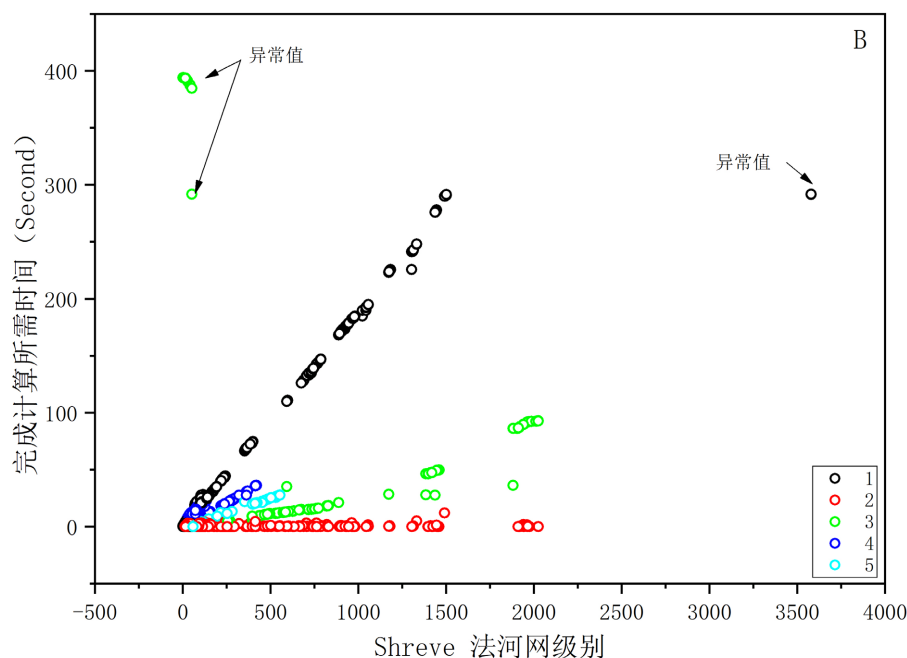


Figure 11. Model operation efficiency. (A. Relationship between model operation time and the number of river reaches. B. Relationship between model running time and river network density)

图 11. 模型运行效率。(A. 模型运行时间与河段数量关系。B. 模型运行时间与河网密度关系)

5. 结论

相比较于 Mike、SWAT 等商用模型, 本模型对于大流域河网水质的模拟效率较高, 能够大幅度削减模拟消耗的时间, 对于流域面积约 8000 km² (河网长度 5571.72 km, Shreve 河网级别 3579) 的流域, 模拟时间仅为 4.86 min。同时模型输入参数较少, 模型复杂度较低, 可控性较强。模型对于 TP 的模拟效果相比较于 NH₃-N 稍差, 这与 TP 复杂的削减过程有关, 需要进一步完善 TP 在复杂水环境条件下的过程模拟, 但总体来说对河网水质的模拟效果较好。在模型运用前景方面, 在现阶段, 模型可运用于大型流域综合决策过程中较大时间尺度下的水质响应关系的模拟以及河流水功能区环境容量的精细化计算, 对于较小时间尺度(如应急响应)的运用有待下一步的开发。

参考文献

- [1] Keupers, I. and Willems, P. (2017) Development and Testing of a Fast Conceptual River Water Quality Model. *Water Research*, **113**, 62-71. <https://doi.org/10.1016/j.watres.2017.01.054>
- [2] 张明亮, 沈永明. 河网水动力及综合水质模型的研究[J]. 中国工程科学, 2008, 10(10): 78-83.
- [3] 王皓, 李铁键, 高洁, 傅旭东, 王光谦. 大尺度流域河网二叉树编码方法[J]. 河海大学学报(自然科学版), 2009, 37(5): 499-504.
- [4] 李满, 王山东, 杨松, 刘超楠. 基于 DEM 的数字流域河网编码方法研究[J]. 计算机与数字工程, 2014, 42(2): 210-212+226.
- [5] 付立冬, 王艳, 夏广锋. 基于 WebGIS 的一维水质模型实现[J]. 电脑编程技巧与维护, 2017(14): 78-80.
- [6] 卢诚, 安望达, 张晓彤, 马淑平, 马占军. 基于 EFDC 模型的神定河水水质模拟[J]. 中国环境监测, 2020, 36(4): 106-114.
- [7] 薛野. 河网一维水动力水质模型研究及系统实现[D]: [硕士学位论文]. 武汉: 华中科技大学, 2018.

-
- [8] 李铁键, 王光谦, 刘家宏. 数字流域模型的河网编码方法[J]. 水科学进展, 2006(5): 658-664.
- [9] 覃露, 叶维丽, 韩旭, 郭送军, 陈荣志, 张帆. 基于北京清河流域水质提升的入河排污口排放要求确定策略研究[J]. 环境污染与防治, 2020, 42(9): 1102-1107.
- [10] 周睿, 王博, 林豪栋, 杨继洲, 韦大明. 一维水质模型结合改进的输出系数法在流域非点源污染负荷估算中的应用[J]. 吉林大学学报(地球科学版), 2021, 51(2): 505-515.
- [11] Młynski, D., Wałęga, A., Książek, L., Florek, J. and Petroselli, A. (2020) Possibility of Using Selected Rainfall-Runoff Models for Determining the Design Hydrograph in Mountainous Catchments: A Case Study in Poland. *Water*, **12**, Article No. 1450. <https://doi.org/10.3390/w12051450>
- [12] 万术海. 大沽河流域农业面源污染信息的提取与发布[D]: [硕士学位论文]. 阜新: 辽宁工程技术大学, 2012.
- [13] 龚慧, 国静, 李骏, 姚敏. 一维水质模型在排污口对水质影响分析中的应用[J]. 江苏水利, 2017(6): 24-27+31.
- [14] 邢成, 张芊芊, 蔡雅雅, 烟晓婷, 应光国. 河网连续动态模型构建及其在典型杀生剂时空迁移模拟中的应用[J]. 环境科学, 2021, 42(7): 3147-3155. <https://doi.org/10.13227/j.hjx.202011202>
- [15] 严蔚敏. 数据结构[M]. 第二版. 北京: 清华大学出版社, 1992: 178.
- [16] 金绍乾, 贺细平. 基于 Python 的线段树实现与优化[J]. 电脑与信息技术, 2020, 28(4): 33-36.