

# 水平式ADCP数据过滤的探索

吴亚玲, 秦 永

长江水利委员会上游水文水资源勘测局, 重庆

收稿日期: 2025年5月28日; 录用日期: 2025年6月29日; 发布日期: 2025年7月21日

## 摘 要

水平式ADCP(又名H-ADCP)已广泛应用到水文测站流量的在线监测中, 不但减轻了测站的测量工作, 更能准确的掌握测站流量的实时变化情况, 特别是水库或电站的增多, 导致越来越多的水文测站受到不同程度的影响, 较频繁的出现中低水水位流量关系紊乱, H-ADCP可以解决因受水流脉动的影响, 人工测验无法完整控制流量过程的问题。H-ADCP的应用作为水文监测的核心技术, 数据的处理方法直接影响资料的成果质量。

## 关键词

H-ADCP的应用, 数据过滤处理, 特异值处理

# Application of Horizontal ADCP Data Filtering

Yaling Wu, Yong Qin

Hydrological Bureau of the Yangtze River Water Resources Commission, Upper Yangtze River Hydrological and Water Resources Survey Bureau, Chongqing

Received: May 28<sup>th</sup>, 2025; accepted: Jun. 29<sup>th</sup>, 2025; published: Jul. 21<sup>st</sup>, 2025

## Abstract

Horizontal ADCP (also known as H-ADCP) has been widely used in online monitoring of hydrological station flow, which not only reduces the measurement work of the station, but also accurately grasps the real-time changes in station flow. Especially with the increase of reservoirs or power stations, more and more hydrological stations are affected to varying degrees, resulting in frequent disturbances in the relationship between low and medium water level flow. H-ADCP can solve the problem of incomplete control of flow process by manual testing due to the influence of water flow pulsation. As the core technology of hydrological monitoring, the application of H-ADCP directly affects the quality of data processing methods.

## Keywords

### Application of H-ADCP, Data Filtering Processing, Outlier Processing

Copyright © 2025 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

时代的发展推动着水利监测的高质量发展, 水文也逐步更替了新型生产力, 在线监测仪器既能完整的监测水文数据, 又可以自动完成资料收集工作, 解放生产力, 而 H-ADCP 作为流量监测的主力军已经广泛应用于测站测量, 但是由于设备自身技术状态和受干扰水流自身的特性, 收集到的数据有一定波动性。因此, 亟需寻求特定的方法, 筛选出不合理的监测数据, 并使筛选后的数据满足资料整编的需求。

本文从误差原理入手, 结合实际数据探讨数据优化的方法, 筛选出反映真实河道的 H-ADCP 流量自动监测数据, 进而为自动监测、水情报汛和资料整编提供技术参考。

## 2. H-ADCP 测流原理及安装

ADCP 测速原理是基于以下两个假定条件[1]: 首先假定反射声波信号产生多普勒频移的水中悬浮物或者气泡是和水流等速运动的。同时, 假定在距仪器一定距离内两波速相应测点处的流速大小方向也是相同的。工作时, 探头安装在测流断面的水面下, 按一定规则排列的四个声波换能器向水中发射脉冲声波, 之后接收被水体中颗粒物散射回来的声波[2], 根据速度矢量得到流速矢量, 再根据断面面积, 得到流量。H-ADCP 是将 ADCP 仪器固定在岸边, 测量探头朝向对岸, 定时对单元流速数据进行收集的监测方法。

研究对象为重庆北碚水文站, H-ADCP 选择在基本水尺断面较为平缓的左岸安装, 安装采用垂直钢结构支架, 将该仪器固定在断面起点距 76.0 m 处, 以避开水流紊流区域。仪器的垂直波束高程为 172.52 m, 在历年最枯水位以下, 高于河底淤积层, 安装位置详见图 1。经过现场调整, 仪器表面水平指向对岸, 且与水流方向垂直, 仪器倾斜传感器测量的纵、横摇角度确保在 1° 以内。

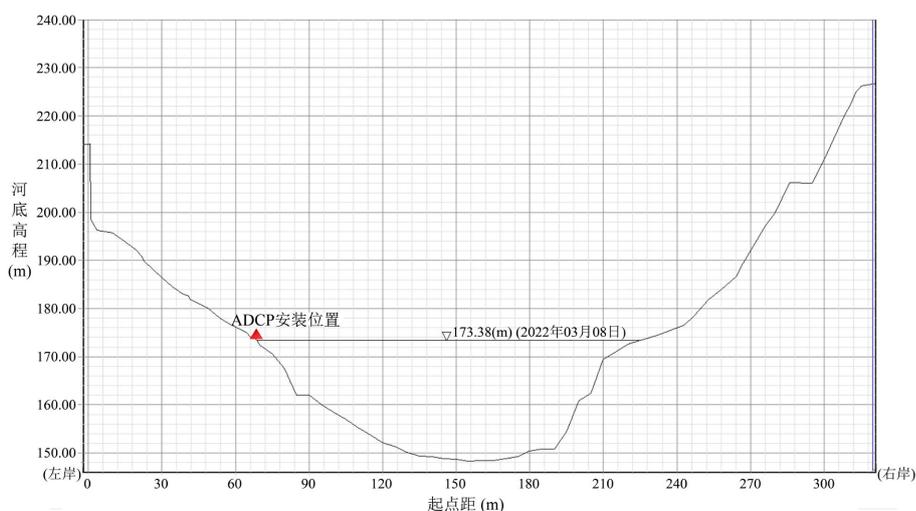


Figure 1. Schematic diagram of H-ADCP installation location at Beibei Station  
图 1. 北碚站 H-ADCP 安装位置示意图

### 3. 测站概况

北碚水文站是嘉陵江流域出口控制站, 集水面积: 156,736 km<sup>2</sup>, 是国家基本水文站, 为三江汇合后嘉陵江下段基本控制站。

该站测验河段顺直, 两岸平缓而稳定, 站址位于嘉陵江观音峡内, 但是中低水位时期经常受草街电站日调节下泄和三峡顶托的影响, 水位流量关系较为紊乱, 流量过程控制比较困难, 测站河床为典型的 U 型, 且河床为乱石组成, 较为稳定, 其流速分布一般遵循抛物线型分布; 该河道虽通航, 但是中低水位时期通航船只较少, 且通航位置更靠右岸, 对 H-ADCP 影响不大; 该情况适合使用 H-ADCP 进行在线监测。

### 4. 仪器应用情况

目前, 北碚站在 2019 年通过数据比测, 已在中低水位的时段投产使用, 实现流量在线监测, 使用状况基本良好, 主要方法如下:

H-ADCP 将横剖面分成若干单元, 根据北碚站情况每 15 分钟采集一次单元流速, 其中一起的数据采集时间为 3 分钟, 考虑脉冲的影响, 采集程序采用算数平均法对这 3 分钟的数据进行平均作为某个时间点流速。采用试错的方法计算出不同单元的流速  $V_{Pi}$ , 通过同时间比测得到断面平均流速  $V_d$ , 两者建立关系, 得到相关关系最优的区间单元和关系, 进而通过水位查算断面得到断面流量。

在枯水期, 仪器收集的数据中存在一些极低的流量数据, 这些数据有变化过程, 但是一般情况下基本不会出现如此低流量数值。以北碚水文站 2024 年 12 月份数据为例, 滤波前数据过程图见图 2。

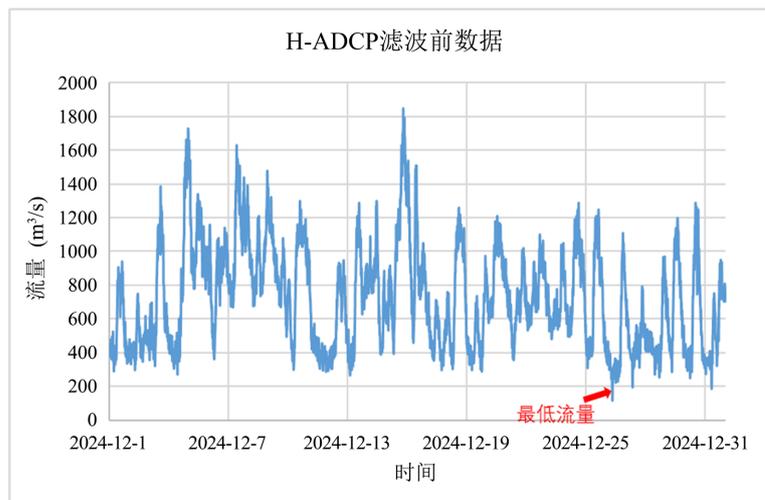


Figure 2. Data before H-ADCP filtering at Beibei Station

图 2. 北碚站 H-ADCP 滤波前数据

通过滤波前流量数据发现, 北碚站 12 月的最低流量为 117 m<sup>3</sup>/s, 出现时间在 12 月 26 日; 月最高流量 1850 m<sup>3</sup>/s, 出现时间在 12 月 15 日; 最低流量已低于北碚水文站建站以来的最低流量, 同时也低于 12 月的缆道施测最低流量, 因此, 有必要怀疑此最低流量是否异常, 进一步检查流量变化过程发现, 波形和强度也无明显的异常, 所以, 确认此最低流量为异常值, 需进一步探索此异常值出现的原因。

通过对 ADCP 的测流原理研究发现, 出现此异常值的原因极可能是水流的脉动影响, 由于 H-ADCP 收集的流量数据都会在一定范围内上下波动, 同段时间收集的数据越多, 数据上下波动的情况就越明显, 不仅体现在针对月极值, 而且对整个流量数据都存在影响。

## 5. 特异值的处理

### 5.1. 特异值的概念

在统计学上, 一组平行测定所得到的分析数据中, 有时会出现个别测定值与其他数据相差较远, 这些数据称为特异值或逸出值(outlier) [1]-[3]; 特异值的产生原因大致有两点: 一是总体固有变异的极端表现, 这些变化是真实且正常的的数据, 只是这次表现显得极端, 这类值和其余值属于同一总体; 二是由于外在条件和方法的改变, 产生一种非正常的、明显有错误的的数据。

而特异值的检验也大致分为两大类: 一是标准偏差预先已知的场合; 另一种是标准偏差未知的场合, 只能利用待检验的一组分析数据本身来检验特异值是否异常。

### 5.2. 特异值处理的理论依据

#### 1) 标准偏差预先已知的场合

如果某测验断面在已知流速水平分布( $x=f(y)$ )的前提下(可通过正常值, 根据最小二乘法确定[3]), 可采用《水文资料整编规范》里“关系曲线检验”中的偏离数值检验法, 以确定去年的公式是否适用于今年的关系, 具体方法如下:

(1) 计算相对偏差  $p_i$  和平均相对偏差  $\bar{p}$ 。

$$p_i = \frac{x_i - f(y_i)}{f(y_i)}, i = 1, 2, \dots, n$$

$$\bar{p} = \frac{\sum p_i}{n}$$

(2) 计算检验值  $t$

$$t = \frac{\sum \bar{p}}{\sqrt{\sum (p_i - \bar{p})^2 / [n(n-1)]}}$$

判定是否通过检验。《水文资料整编规范》规定, 对于偏离数值检验, 置信区间  $\alpha$  采用 0.10~0.20, 临界值按表 1 确定。

**Table 1.** Deviation from the critical value  $t_{1-\alpha/2}$  of the numerical method

**表 1.** 偏离数值法临界值  $t_{1-\alpha/2}$

$\alpha \backslash k$	6	8	10	15	20	30	60	$\infty$
0.05	2.45	2.31	2.23	2.13	2.09	2.04	2.00	1.96
0.10	1.94	1.86	1.81	1.75	1.73	1.70	1.67	1.65
0.20	1.44	1.37	1.37	1.34	1.33	1.31	1.30	1.28
0.30	1.13	1.09	1.09	1.07	1.05	1.05	1.05	1.04

注: 表中  $k$  为自由度, 取  $k = n - 1$ 。

但是 H-ADCP 流量是实时变化的, 且用 H-ADCP 监测的流量多数存在较紊乱的情况, 不适合标准差预先已知的场合。

#### 2) 标准偏差未知的场合

在中低水水流较紊乱的情况下, 特异值多处出现, 如何判断数据是否变化合理呢? 根据《水文统计学》的抽样分布, 一般用样本数字特征来描述。

样本平均值:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

相对误差:

$$E = \frac{1}{n} (x_i - \bar{x})$$

标准差:

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

在判断数据能否处理时, 首先我们要检查样本均值, 如果日均值流量小于测站的历年最低流量, 就要查看这段时间的收集数据是否正常, 要考虑这段时间的流量数据是否能用, 在流量数据有效的情况下才能进行滤波处理。滤波一般方法是用 3 倍的标准差进行数据剔除, 如果某个值与均值的差异大于 3 倍标准差, 认为该值为特异值, 并将其剔除。

解决 H-ADCP 里面水流脉动带来的影响, 一般有两种方法, 一个是算术平均平滑另一个就是特异值滤波。算术平均值平滑是时间向前向后偏移, 用一段时间的平均值作为此时段点流速[4], 这种方法虽能消除一定的水流脉动, 但是极值往往被平滑处理掉了。而特异值滤波是以日均值近似为真值, 跟一日内收集的数据进行比较, 偏差超过标准差的 3 倍的点就作为特异值剔除, 这种方法在保留瞬时极值的同时, 能剔除一些异常的极大极小值, 使处理后的数据更加合理。

在实际工作中我们进行数据滤波前, 首先要对数据进行预处理, 因为均值作为判断去除明显错误的基准, 要确保一日内大部分是正确且有效的, 只有个别时段有异常数据, 这种情况下才能采用 3 倍标准差进行滤波处理, 滤波后的数据检查也较为合理, 12 月滤波后的数据对比图见图 3。

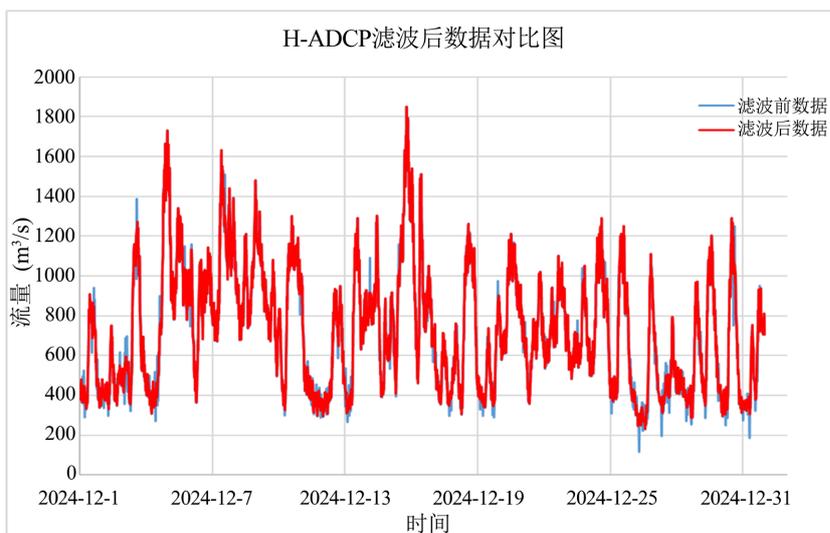


Figure 3. Comparison of H-ADCP filtered data at Beibei Station  
图 3. 北碚站 H-ADCP 滤波后数据对比图

由图 3 可以明显看到滤波筛掉了一部分波动异常大的数据, 月最高流量进行比较后判定是正常的数, 给予保留, 而月最低流量滤波后流量从 117 m<sup>3</sup>/s 增大为 230 m<sup>3</sup>/s, 为了更加直观的进行对照, 把数据列成下列表 2 进行对比。

**Table 2.** Comparison of data before and after H-ADCP filtering at Beibei Station**表 2.** 北碚站 H-ADCP 滤波前后数据对比表格

	数量(个)	最低流量(m <sup>3</sup> /s)	最高流量(m <sup>3</sup> /s)
滤波前数据	2989	117	1850
滤波后数据	2181	230	1850

滤波后数据保留了相对误差范围内月最高值, 过滤后的月最低值更加合理, 消除了一定程度上的数据波动, 从而使资料成果更加准确。

## 6. 结论

本文根据北碚站的水情特性, 基于统计学中的特异值的定义和抽样分布误差, 以 H-ADCP 的原理和自动监测数据为基础, 通过特异值滤波对数据进行有效处理, 筛选出 70%左右的可用数据, 在最低值通过滤波处理更贴合测站实际水情。

H-ADCP 数据容易受到水流脉动的影响, 需要进行数据预处理, 要保证一日内收集的流速大部分数据都是准确且有效的, 流量异常极值的数据是否合理需要重点关注, 如出现比测站历年最低流量还要低的数据时, 要仔细判断该时间段 H-ADCP 的运行状态是否正常, 数据是否能用, 预处理后的数据要进行数据选定方式的滤波处理, 过滤后的数据可以有效削弱水流的脉动影响, 使得资料更加准确。对 H-ADCP 数据处理来说, 特异值滤波比算术平均平滑更好, 保留的值更加贴近测量真实情况, 极值不会被平滑处理掉, 提高数据质量, 为 H-ADCP 应用于水文监测、水文报汛和水资源管理等领域提供可靠数据支持。

## 参考文献

- [1] 朱晓原, 张留柱, 姚永熙. 水文测验实用手册[M]. 北京: 中国水利水电出版社, 2013.
- [2] 冯建军. ADCP 原理及数据处理方法[J]. 港工技术, 2007(3): 53-55.
- [3] 黄振平. 水文统计学[M]. 南京: 河海大学出版社, 2003.
- [4] 胡尊乐, 仲兆林, 郭红丽, 等. 定点式 H-ADCP 自动监测数据的处理[J]. 水文, 2020, 40(6): 46-50.