

固态存储水位精简摘录的研究

赵良民

(福建省水文水资源勘测局,福建 福州 350001)

摘要: 讨论了固态存储水位数据精简摘录存在问题及所需满足的要求。介绍了曲线矢量压缩的 Douglas-Peucker 算法,对它做了改进并应用于水位过程线的精简摘录。应用实例表明该法能够以精简的数据反应出水位变化完整过程,是使用计算机完成固态存储水位摘录的有效方法。

关键词: 固态存储;水位;精简摘录

中图分类号: P208

文献标识码: A

文章编号: 1000-0852(2014)02-0084-03

随着水文现代化的大力推进,各种水位自动监测设备在我国的河流、湖泊、水库、感潮河段的水文水位站得到广泛应用。水位自动监测仪器种类较多,按照传感器工作原理分为浮子式、压力式、超声波式、接触式等。仪器一般按定时段采样记录,时段长可设为 5、10、15、30min 等,与传感器连接的固态存储器记录存储每次采样的时间和水位数据。设置了适当采样周期的自动水位计能观测记录到完整的水位变化过程,同时也生成了大量的数字化时间、水位数据。以我国大量使用的 DT350 水位数据采集器为例,设置 5min 采样记录一次,一天有 288 组的时间、水位数据,一年约有 10 万组数据。大量原始观测数据即不适合直接提供水文资料整编软件使用,成果也不能满足资料整编特别是水文年鉴汇编刊印的要求。如何使用计算机对原始固态存储水位数据进行预处理,完成合理的精简摘录,是本文所要介绍的内容。

1 摘录要求

根据 GB/T 50138《水位观测标准》7.1.11 条及附录 B.2.2,可以认为固态存储水位精简摘录与纸介质模拟自记水位计记录的人工摘录一样,必须满足以下要求:

(1) 零时、8 时、24 时及测流、测沙等时间的水位必须摘录。

(2) 月最高、最低水位必须摘录,以满足统计特征值的需要。

(3) 摘录水位变化转折点,摘录成果应能反应水位变化的完整过程,并应满足计算日平均水位和推算流量的需要。

上述(1)、(2)点要求可根据时间或水位极值在原始观测数据中挑选。对于(3)点要求可以理解为水位精简摘录必须做到以尽量少的数据来尽可能完整反应水位变化过程,因此本文引入曲线矢量数据压缩的 Douglas-Peucker 算法对水位过程线进行精简摘录。

2 摘录方法

曲线矢量数据压缩是计算机图形学、地理信息系统等学科中的一个常见问题。它是从组成曲线的数据集合 A 中抽取一个子集 A',用这个子集作为一个新的信息源,在规定的精度范围内该子集能够从内容上尽可能近似反映原集合 A,而在数量上则尽可能精简。

Douglas-Peucker 算法(以下简称 D-P 算法)是曲线矢量数据压缩的一种常用而有效的经典算法。其基本思想是:找出与曲线两端点连线垂直距离最大的曲线中间点,判断其到两端点连线的距离是否小于允许限差 ϵ ,若是,舍弃两端点间所有中间点。否则保留该点,并以该点为界将曲线分为两部分,用同样的方法对它们的中间点进行检测,以确定下一批压缩后的保留点。如此反复进行,直至所有的最大垂直距离均小于允许限差 ϵ 为止。D-P 算法能够取得曲线特征点,再现曲线形态特征且数据压缩效果好。

D-P 算法处理过程中计算的曲线中间点至两端点连线垂直距离(如图 1 的 P_8B),在以时间、水位表示的水位过程线上无实际物理意义,不便于允许限差 ε 的取值。因此本文将该法移用于固态存储水位精简摘录时,改为计算过程线中间点水位与该点在过程线两端点连线的同时刻线上水位的绝对差值(如图 1 的 P_8A),相应地把允许限差 ε 改称为水位摘录允许限差 $\angle Z$ 。水位摘录允许限差 $\angle Z$ 可以理解为:使用精简摘录后的水位过程线内插其中任何时刻的水位与原始观测值最大相差不超过 $\angle Z$, $\angle Z$ 表示了精简摘录成果的精度。参考《水位观测标准》4.2.4 条对自记水位计允许测量误差的规定, $\angle Z$ 可考虑在 2cm 附近取值。

以图 1 所示为例,介绍应用改进后的 D-P 算法精简摘录水位的处理过程。固态存储水位过程线由点列 $(P_0, P_1, \cdots, P_{12})$ 表示,首先在过程线两端点 P_0 和 P_{12} 间连一条直线,探测与该直线上同时刻水位差值最大的点(如图 1 所示为 P_8),如果水位差值不大于水位摘录允许限差 $\angle Z$ (事先给定),则 P_0 和 P_{12} 之间所有点全部删除,用直线 P_0P_{12} 代替原过程线,摘录结束。如果水位差值大于允许限差 $\angle Z$,则摘录该点,并使用该点将过程线分两段,图 1 中分为 $(P_0, \cdots, P_8)$ 和 $(P_8, \cdots, P_{12})$,将 $(P_0, \cdots, P_8)$ 和 $(P_8, \cdots, P_{12})$ 分别作为独立过程线重复上述

步骤,最终摘录的特征点为 $(P_0, P_1, P_3, P_8, P_{10}, P_{12})$ 。

3 应用实例

渡里(二)水文站为集水面积 76.5km^2 的小河站,枯水期受水力发电等人类活动影响,水位起伏较明显,洪水期水位暴涨暴落。该站自 2002 年起在自记井内安装了 WFH-2 型机械编码水位传感器,与 SW40 型日记水位计并行观测。WFH-2 型编码水位计外接 DT350 型水位数据采集器,采样时间间隔为 5min。

取用记录完整的 2004 年渡里(二)站固态存储水位资料,分别按允许限差 $\angle Z=2.5, 2.0, 1.5\text{cm}$,使用前述方法精简摘录后重新进行水位、流量资料整编。该站当年整编使用的水位则是由人工在 SW40 型日记水位自记纸上摘录取得,流量采用临时曲线法推求。

不同允许限差 $\angle Z$ 的全年固态存储水位摘录总数和当年最大一次洪水水位摘录个数见表 1。其中 $\angle Z=1.5\text{cm}$ 时,全年、次洪水水位摘录个数与当年整编成果(分别为 5041 和 82)最为接近,数据精简压缩率分别为 4.7%、5.2%。

使用不同允许限差 $\angle Z$ 精简摘录的水位计算全年日平均水位,同时采用全部原始固态存储水计算日平均水位,两者差值见表 1。日平均水位误差在 $\pm 1\text{cm}$ 以内,由精简摘录的水位数据能计算得到令人满意的日平均水位成果。

使用不同允许限差 $\angle Z$ 精简摘录的水位进行推流计算,各项流量成果与当年整编数值的误差见表 2。其中 $\angle Z=1.5\text{cm}$ 时,各项误差在 $\pm 3\%$ 以内。考虑 WFH-2 水位计与 SW40 水位计观测到的水位有时会略有差别,可以认为两者推算的流量成果还是十分接近。

图 2 为 $\angle Z=1.5\text{cm}$ 时,精简摘录数据点绘的水位过程线与全部原始固态存储数据点绘的水位过程线对照图。图形直观地表明精简摘录的数据能取得水位

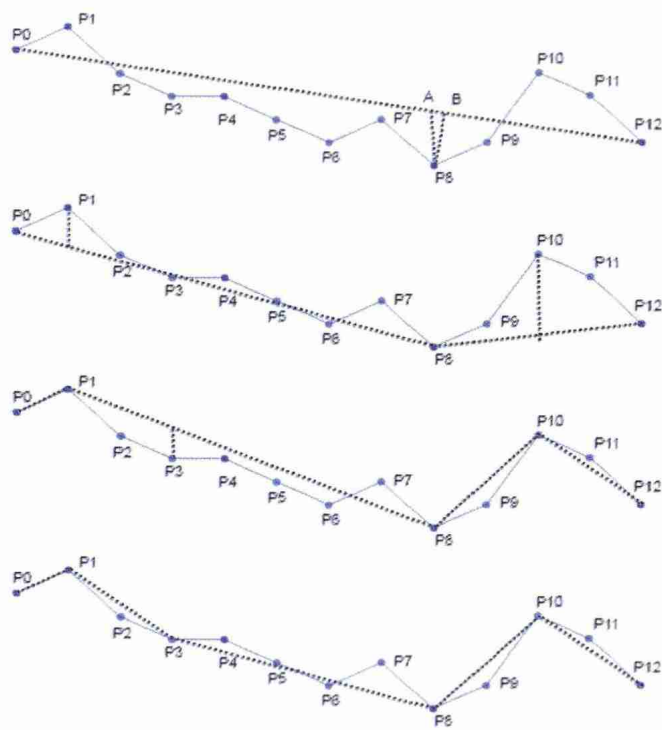


图 1 D-P 算法示意图
Fig.1 Schematic diagram of Douglas-Peucker algorithm

表1 不同 $\angle Z$ 的水位摘录数及日平均水位计算误差
Table1 The stage data and errors of daily mean stage with different $\angle Z$

水位允许限差 $\angle Z$ /cm	水位摘录数		日平均水位误差/cm
	全年	次洪	
2.5	3107	66	-0.69~0.48
2.0	3718	71	-0.62~0.67
1.5	4995	89	-0.28~0.47

表2 不同 ΔZ 流量计算误差Table2 The errors of discharge calculated with different ΔZ

水位允许限差 ΔZ /cm	年径流量误差 /%	月平均流量误差 /%	洪量误差 /%
2.5	1.7	-0.9~4.9	-0.3~1.9
2.0	1.7	-0.8~1.0	0.0~1.7
1.5	1.1	-0.1~2.8	-0.6~1.3

变化转折点,反应出水位变化的完整过程。

4 结语

应用以水位摘录允许限差 ΔZ 为控制的D-P算法能有效地对固态存储水位进行精简摘录,摘录成果可以满足有关规范要求。水位摘录允许限差 ΔZ 表示摘录数据最大误差,实际使用中可取1.5~2.0cm。该法即适用于定时采集,也适用于事件采集取得的水位数据精简摘录。另外,该法也可考虑用于由自动监测仪器取得的其它水文要素过程的精简摘录,如固定式AD-CP测得的流速过程。

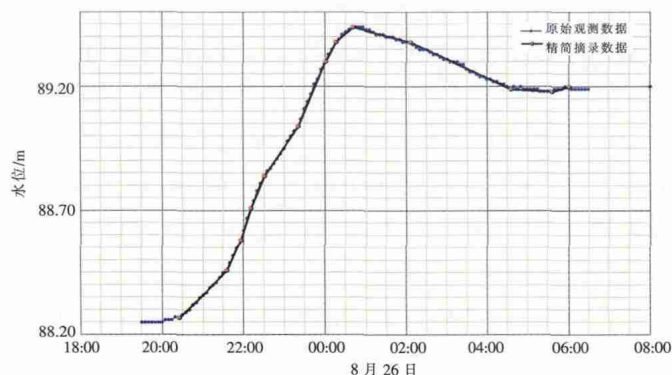


图2 水位精简摘录成果的软件截图

Fig.2 The hydrograph of stages extracted by applying the improved Douglas-Peucker algorithm

参考文献:

- [1] 林祚顶. 水文现代化与水文新技术 [M]. 北京: 中国水利水电出版社, 2008. (LIN Zuoding. Modernization and New Technology of Hydrology [M]. Beijing: China WaterPower Press, 2008. (in Chinese))
- [2] GB/T 50138-2010, 水位观测标准[S]. (GB/T 50138-2010, Standard for Stage Observation [S]. (in Chinese))
- [3] Douglas D H, Peucker T K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature [J]. The Canadian Cartographer, 1973, 10(2): 112-122.

Extraction of Simplified Stage Data from Solid State Storage

ZHAO Liangmin

(Hydrology and Water Resources Survey Bureau of Fujian Province, Fuzhou 350001, China)

Abstract: This paper discussed the issues and requirements about how to extract the simplified stage data inform solid state storage, and introduced to Douglas-Peucker algorithm for curve vector compression. The algorithm was made better to accommodate simplification of stage data in this paper. A practical example of improved Douglas-Peucker algorithm was described, which indicates the effectiveness of this method.

Key words: solid state storage; stage; extraction

(上接第74页)

A Model for Standard U-shaped Cross-section Surface Curve Calculation

TENG Kai

(Water Authority of Qiqihar City, Qiqihar 161006, China)

Abstract: Summation method with conventional staging to complete the standard U-shaped cross-section surface curve calculation not only exist cumulative error, but also does calculation of the heavy workload with low efficiency. Based on the theory of optimum fitting, with the standard minimum residual difference as the objective function, in the scope of the engineering applicable parameters, a non-integrable function was obtained by a simplified formula to replace the segments in the original integral. By integration, an analytical model for the cross-sectional surface curve was obtained. Because the positive and negative fitting errors offset each other the accuracy of the calculation was higher.

Key words: standard U-shaped cross section; normal depth; optimization fitting; approximate calculation