

韩广. 基于Sentinel-2遥感影像的水库库容曲线快速构建[J]. 人民珠江, 2025, 46(7): 86-100.

基于Sentinel-2遥感影像的水库库容曲线快速构建

韩广

(邯郸市水利水电勘测设计研究院, 河北 邯郸 056000)

摘要: 针对利用遥感技术构建水库库容曲线尚未形成一套自动化流程的问题, 基于水体指数法提出利用累积频率曲线拐点法确定最优阈值, 采用区域种子生长算法优化水体边缘, 进而构建了一套完整地自动化流程快速构建水库库容曲线。以Sentinel-2遥感影像为数据源, 以河北省邯郸市武安市车谷水库为研究对象, 采用NDWI、MNDWI、RWI、MBWI和AWEInsh共5种水陆像元度量方法, 提取水体面积, 构建库容曲线, 并和2017年测量成果比较。结果显示: NDWI和RWI水体指数提取精度较高, 提取车谷水体面积平均误差为2.5%, 相对误差分别为-7.3%~5.8%和-6.2%~4.7%, R^2 分别为0.99336和0.99049; 库容相对误差范围分别为-9.4%~-1.8%和-9.3%~-1.7%。研究表明: 累积频率曲线拐点法对水体指数最优阈值的确定具有较好的适用性和稳定性, 且原理简单易于实现。区域种子生长算法可有效解决水体边缘不连续的现象, 减少系统误差, 进一步提高精度。累积频率曲线拐点法为遥感水体提取提供了新思路; 对批量构建全国中小型水库库容曲线提取具有一定的参考意义。

关键词: Sentinel-2; 水体指数; 累积频率曲线拐点法; 生长算法; 库容曲线; 车谷水库

中图分类号: TV123; TP79 **文献标识码:** A **文章编号:** 1001-9235(2025)07-0086-15

Rapid Construction of Reservoir Storage Capacity Curve Based on Sentinel-2 Satellite Multi-Spectral Remote Sensing Image

HAN Guang

(HanDan Design and Research Institute of Water Conservancy and Hydropower, Handan 056000, China)

Abstract: As a set of automated process has not yet been formed for constructing reservoir capacity curves using remote sensing technology, it is proposed to determine the optimal threshold with the inflection point method of the cumulative frequency curve based on the water body index method. The edge of the water body is optimized by the regional seed growth algorithm, so as to construct a complete set of automated process to quickly construct the reservoir capacity curves. With Sentinel-2 remote sensing images as the data source and Chegu Reservoir in Wuan City, Handan City, Hebei Province as the research object, this paper uses five land and water image metrics, namely, NDWI, MNDWI, RWI, MBWI, and AWEInsh to extract the area of the water body, construct the reservoir capacity curve, and compare it with the results of the measurements in 2017. The results are as follows. The extraction accuracy of NDWI and RWI water body index is high. The average error of Chegu water body area extraction is 2.5%. The relative error is -7.3%~5.8% and -6.2%~4.7%, and the R^2 is 0.99336 and 0.99049, respectively. The range of the relative error of the reservoir capacity is -9.4% to -1.8% and -9.3% to -1.7%, respectively. The study shows that the cumulative frequency curve inflection point method is highly applicable and stable in determining the optimal threshold of water body index, with a simple and feasible principle. The regional seed growth algorithm can effectively eliminate the discontinuity at the edge of the water body, reduce systematic error, and

收稿日期: 2024-11-16 修回日期: 2024-12-10 网络首发日期: 2024-12-27

作者简介: 韩广(1989—), 男, 硕士, 主要从事水文水资源研究。E-mail: 451984531@qq.com

further improve the accuracy. The cumulative frequency curve inflection point method provides a new idea for remote sensing water body extraction, serving as a reference for the batch construction of national small and medium-sized reservoir capacity curve extraction.

Keywords: sentinel-2; water Index; inflection point method of the cumulative frequency curve; growth algorithm; storage capacity curve; Chegu Reservoir

水库是流域防洪工程体系的重要组成部分,是国家水网重要结点,是保障国家水安全的重要基础设施。水库库容是水库有效发挥防洪、供水、生态、发电等功能的重要保障,事关工程安全、防洪安全、供水安全和生态安全。截至2023年,全国已建成各类水库94 877座,其中大型水库836座,中型水库4 230座,小型水库89 811座^[1]。中国各类水库的80%修建于20世纪50—70年代,受当时基本图件、测量仪器及方法限制,库容曲线精度偏低,加上水库运行期间受到库区微地貌、淤积、土地利用变化等因素,水库库容曲线难免发生变化,直接影响到水库防洪计算及发电调度运行^[2]。为强化库区、坝体、下游河道监管与问题整改,维护水库库容安全、工程安全、河道安全,科学推进水库减淤清淤,水利部办公厅印发《关于加快开展承担防洪任务的大中型水库库容曲线复核工作的通知》。

目前,中国水库库容曲线复核主要采用陆域地形测量和水域地形测量相结合的方法,其中陆域地形测量主要有航空摄影测量、机载激光扫描和全野外数字测量等方法;水域地形测量主要为GNSS RTK与水深测量相结合的方法。这项工作不仅涉及对水库的实地测量,还需要借助专业的测绘设备和软件,通常需要投入大量的人力、物力和财力。中国水库众多,分布范围广,各地经济状况和水库管理水平各异,寻找一种即实用又能满足测量精度的方法复核水库库容曲线,显得尤为重要。

为贯彻落实《水利部印发关于加强水文情报预报工作指导意见的通知》(水文[2019]203号)《国务院办公厅关于切实加强水库除险加固和运行管护工作的通知》(国办发[2021]8号)要求,各省均开展了中小型水库雨水情测报工作,实现对库水位实时监测,积累了大量的水位观测资料,为使用高分遥

感影像提取水库库容曲线提供了前提条件。

随着遥感技术的快速发展和普及,高分遥感影像逐步被应用于水体遥感解译中,在水利方面的应用越来越成熟广泛^[3]。陈曦等^[4]使用ERDAS IMAGINE遥感图像处理系统,利用CBERS-02B、资源一号02C以及资源三号卫星高分辨率遥感影像,推求桐仁桥小水库库容曲线。张莉芳等^[5]采用Landsat卫星遥感影像,提取了柘林水库库容曲线。王金鑫等^[6]提出了多源时相遥感影像数据耦合的库容曲线重构方法,并重构了深圳市径心水库库容曲线。

虽然目前基于遥感技术对水库库容曲线提取已得到应用,但仍存在批量提取水体面积效率低等问题,难以形成一套自动化流程。因此,本文基于水体指数法,提出利用累积频率曲线拐点法自动确定最优阈值,应用区域种子生长算法优化水体边缘,以实现高效、批量地提取水体面积,并采用JAVA语言,使用GDAL库编制相关计算程序,构建了一套完整地自动化流程来快速构建水库库容曲线。

1 研究方法

1.1 遥感影像水体提取

目前使用遥感影像提取水体的算法很多,一般而言,这些方法大致可分为单波段法、谱间关系法、影像分类法和水体指数法等几个类别,其中水体指数法被研究人员广泛采用^[7]。水体指数是一种简单且高效地提取水体信息的方法。但由于受区域自然条件、成像时大气条件、传感器差异等因素影响,水体指数法的阈值随不同地区和季节影像的不同而有差异。阈值的不稳定性和设定的繁琐性常常是水体遥感指数被诟病之处^[8]。

为实现大批量、自动化、高效地提取水体信息,本文采用水体指数法。同时为实现阈值的自动确定,从统计学角度,按聚类分类的思想,绘制水体指数累积频率曲线,经数学分析提出最优阈值位于曲线上由凹变凸的拐点处,进而提出一种高效、易编程的自动化阈值确定方法,即累积频率曲线拐点法。

1.1.1 水体指数法

水体指数的构建基本可分为2种类型:比值型和差值型。比值型指数通过比值运算扩大水体和非水体信息的区别,分离度较强,并能抵消地形的影像。由于其变量常不带系数,因此具有较强的普适性。差值型指数则多带有经验系数,在不同卫星、不同地区和不同环境中表现不太稳定^[8]。本文选取国内外比较有影响力的水体指数,其中比值型选取归一化差异水体指数(Normalized Difference Water Index, NDWI)^[9]和改进的归一化差异水体指数(Modified Normalized Difference Water Index, MNDWI)^[10];差值型指数选取多波段水体指数(Multi-band Water Index, MBWI)^[11]和自动水体提取指数(Automated Water Extraction Index, AWEInsh)^[12];以及专门用于Sentinel-2提取细小水体的植被红边水体指数算法(Redside Water Index, RWI)^[7]。基于Sentinel-2遥感影像的各水体指数计算见式(1)~(5):

$$\text{NDWI} = (B_3 - B_8) / (B_3 + B_8) \quad (1)$$

$$\text{MNDWI} = (B_3 - B_{11}) / (B_3 + B_{11}) \quad (2)$$

$$\text{MBWI} = 2 \times B_3 - B_4 - B_{8A} - B_{11} - B_{12} \quad (3)$$

$$\text{AWEInsh} = 4 \times (B_3 - B_{11}) - (0.25 \times B_8 + 2.75 \times B_{12}) \quad (4)$$

$$\text{RWI} = [(B_3 + B_5) - (B_8 + B_{8A} + B_{12})] / (B_3 + B_5 + B_8 + B_{8A} + B_{12}) \quad (5)$$

式中: B_3 、 B_4 、 B_5 、 B_8 、 B_{8A} 、 B_{11} 和 B_{12} 分别为Sentinel-2的第3、4、5、8、8A、11和12波段的遥感反射率。

1.1.2 累积频率曲线拐点法

1.1.2.1 理论分析

利用水体指数法提取水体信息时,自动阈值的确定非常关键。常用的方法有最大类间方差法

(Otsu法)、迭代法、直方图双峰法等,以及相关的改进算法^[13-14]。每种方法的核心思想均是利用一定的原则,进行分类,且每种方法都有各自的优缺点和适用范围。如:Otsu方法已被许多研究证明其确定的最优阈值并不可靠^[15-17];直方图双峰法常常会面临单峰或多峰的情况,导致获取阈值结果无意义。

水体指数法原理是基于水体与其他地表覆盖类型在特定波段反射的差异,通过波段运算,经归一化处理后,将影像二值化,以区分水体和其他地表覆盖物。该方法本身具有较强的物理基础和理论依据,且该方法的本质是通过指数值的差异来分类。为更好有效地利用水体指数法,建议应先对水体指数图进行聚类再分类。统计分析中累积频率曲线常用于分析数据的分布特征、中心趋势和离散程度等。由于累积频率曲线生成过程中需要先排序,在一定程度上完成了聚类。本文通过绘制水体指数图的累积频率曲线,提出最优阈值位于曲线上由凹变凸的拐点处。

以 $P(I \geq i) = \sum_{n=1}^N S(I \geq i) / N \times 100$ 为横坐标,以水体指数 I 为纵坐标,绘制累积频率曲线图。其中 $P(I \geq i)$ 表示水体指数图中像元数值大于等于 i 的百分比; $S(I \geq i)$ 为指示函数,当满足 $I \geq i$ 时取1,否则取0; N 为水体指数图的像元总数。

完全二值化影像图的累积频率曲线为不连续的分段函数(图1),曲线中 i_w 代表水体像元值, i_o 代表非水体像元值。

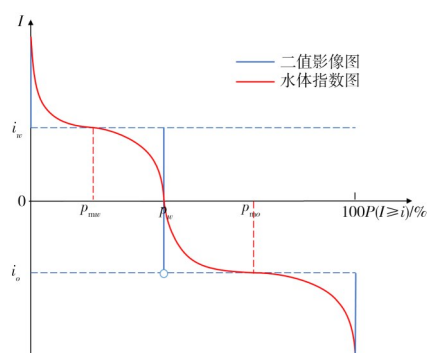


图1 累积频率曲线

Fig. 1 Cumulative frequency curves

实际中,对于归一化的水体指数图,其水体指数为在某范围内变动的离散函数,但整体呈二值化趋势。水体指数图的累积频率曲线为非严格单调递减函数,即 $I'(p) \leq 0$ 。实际应用中,经统计分析得到的累积频率曲线图是离散的点集。为方便分析,用光滑的曲线将点集连接,并假定曲线是连续的,二阶可导的。

曲线上 (p_{mw}, i_w) 代表水体像元密度最大的点。在该点指数值 i 的小幅减少,将引起水体累积频率 p 的大幅增加,即 $I'(p_{mw})$ 为附近邻域 $[p_{mw} - \varepsilon, p_{mw} + \varepsilon]$ 内的最大值。由于 $I'(p)$ 连续可导,因此 $I'(p_{mw})$ 为极大值点,即 $I''(p_{mw}) = 0$ 。由于 $I'(p_{mw}) > I'(p_{mw} - \varepsilon), I'(p_{mw}) > I'(p_{mw} + \varepsilon)$,因此区间 $[p_{mw} - \varepsilon, p_{mw}]$ 内 $I''(p) > 0$,区间 $[p_{mw}, p_{mw} + \varepsilon]$ 内 $I''(p) < 0$ 。由函数凹凸性和拐点定义可知,点 p_{mw} 为曲线上由凸变凹的拐点。曲线上 (p_{mo}, i_o) 代表非水体像元密度最大的点,在该点指数值 i 的小幅减少,将引起非水体累积频率 p 的大幅增加。同理点 p_{mo} 也为曲线由凸变凹的拐点。

由于 $I'(p_{mw})$ 和 $I'(p_{mo})$ 均为附近邻域内最大值(极值点)。依据连续函数的有界性定理,设 $I'(p)$ 在区间 $[p_{mw}, p_{mo}]$ 内的最小值为 $I'(p_w)$,则存在区间 $[p_w - \varepsilon, p_w + \varepsilon]$,使得 $I'(p_w - \varepsilon) > I'(p_w)$,且 $I'(p_w + \varepsilon) > I'(p_w)$ 。由于 $I'(p)$ 在闭区间 $[p_w - \varepsilon, p_w + \varepsilon]$ 上是连续可导的,因此 p_w 为 $I'(p)$ 的极值点,即 $I''(p_w) = 0$ 。由于区间 $[p_w - \varepsilon, p_w]$ 上 $I''(p) < 0$,区间 $[p_w, p_w + \varepsilon]$ 上 $I''(p) > 0$,因此 p_w 为由凹变凸的拐点,即在 p_w 点附近随累积频率的增加,指数值的变化速度先减少后增加。

对于无其他地物干扰,应用效果较好的水体指数,水体和非水信息分离度较大且不重合,即累积频率曲线上 (p_{mw}, i_w) 和 (p_{mo}, i_o) 是唯一的,区间 $[p_{mw}, p_{mo}]$ 内必然存在一个拐点 p_w 。累积频率曲线图在区间 $[p_{mw}, p_{mo}]$ 内的变化趋势是水体指数二值化效果的直接体现。在 p_w 点附近存在某个频率范围,

表达了水体向非水体的过渡过程。该频率范围越小,凹凸变化越明显,说明在较小的累积频率区间水体指数的变化范围越大,二值化效果越好。

综上所述,提出点 p_w 为水体和非水体的分界点, $I(p_w)$ 即为最优阈值。因此,可以通过寻找累积频率曲线上由凹变凸的拐点来确定最优阈值。此外,水体指数图的累积频率曲线突出了水体向非水体的转变过程(区间 $[p_{mw}, p_{mo}]$),弱化了其他影响,即使影像中水体占比较小,也可以较好地地区分水体和非水体。但同时也要有足够的非水体像元,以保证累积频率曲线的完整性,才能充分体现水体向非水体的转变过程。

1.1.2.2 算法实现

水体指数最优阈值的确定可以转换为寻找累积频率曲线上由凹变凸的拐点 p_w 。实际应用中,由于受区域自然条件、大气条件和季节等因素影响,相同区域不同时期采用相同的水体指数法对水体提取的效果可能不同。体现在累积曲线上:一是可能不存在由凹变凸的拐点,二是可能存在多个由凹变凸的拐点。因此,本文给出以下建议:①若累积频率曲线上不存在由凹变凸的拐点,说明水体和非水体的区分界面不明显,这种情况下最优阈值的选取较为敏感,即阈值的较小变动,将引起较大的误差。此时水体指数法已无法有效区分水体和非水体;②若累积频率曲线上存在多个由凹变凸的拐点,说明水体指数图中存在较多类水体的干扰信息,如云层、冰雪等,基于水体指数二值化的原理,累积频率曲线上除两端某区间范围外,最优阈值处 $I'(p_w)$ 最小,由于 $I'(p) \leq 0$,即 $|I'(p_w)|$ 最大。此时,建议先利用简化线算法进行处理,在保持曲线形状的同时,减少干扰信息,之后再寻求由凹变凸的拐点。若仍存在多个凹变凸的拐点,则判断各拐点处的 $|I'(p_w)|$ 值,选取 $|I'(p_w)|$ 值最大处的阈值作为最优阈值。

综上所述,水体指数最优阈值确定的关键步骤如下。

a)采用道格拉斯-普克抽稀算法(DP算法)^[18]进

行简化线处理。选择累积频率曲线首尾两点连接,遍历两点之间的所有点,找到距离首尾连线垂线距离最大的点,若最大垂距小于设定阈值,则删除中间所有点;若最大垂距大于设定阈值,则保留此点,并将其与首尾相连,得到2段新的线段。分别对新线段进行如上操作,直至判断完整个曲线上的点。由于构建累积频率曲线时,两轴的意义并不相同,因此建议采用垂距在I轴上投影长度和设定阈值进行比较。此外,不同水体指数得到的指数范围是不同的。为具有普适性,本文将阈值的确定转化为简化后至少保留的点数,即由大到小试算阈值,当保留点数满足要求时便停止。

b)利用离散型函数拐点算法^[19]分析由凹变凸的拐点区间。设累积频率曲线共有 n 个散点,分别为 $p_0, p_1, p_2, \dots, p_k, \dots, p_{n-1}$,对应的纵坐标为 $I(p_0), I(p_1), I(p_2), \dots, I(p_k), \dots, I(p_{n-1})$ 。如果拐点处于点 p_{k-1} 与 $p_k(k=3, 4, 5 \dots n-3)$ 之间,则有:

$$F(k) = \left(\frac{I(p_{k-1}) - I(p_{k-2})}{p_{k-1} - p_{k-2}} - \frac{I(p_{k-2}) - I(p_{k-3})}{p_{k-2} - p_{k-3}} \right) \times \left(\frac{I(p_{k+2}) - I(p_{k+1})}{p_{k+2} - p_{k+1}} - \frac{I(p_{k+1}) - I(p_k)}{p_{k+1} - p_k} \right) < 0 \quad (6)$$

若在 $p_{k-1} \sim p_k$ 区间,累积频率曲线由凹变凸,则有:

$$\frac{I(p_{k-1}) - I(p_{k-2})}{p_{k-1} - p_{k-2}} > \frac{I(p_k) - I(p_{k-1})}{p_k - p_{k-1}} \quad \text{且} \quad \frac{I(p_{k+1}) - I(p_k)}{p_{k+1} - p_k} > \frac{I(p_k) - I(p_{k-1})}{p_k - p_{k-1}} \quad (7)$$

若存在多个由凹变凸的区间,则判断 $\left| \frac{I(p_k) - I(p_{k-1})}{p_k - p_{k-1}} \right|$ 值,选取该值最大的区间。

c)判断 $p_k - p_{k-1}$ 是否满足精度要求。若满足,则最优阈值取 $I(p_w) = (I(p_k) + I(p_{k-1})) / 2$ 。若不满足,则按精度要求,对 $p_k \sim p_{k-1}$ 区间进行细化,选取 $|I'(p)|$ 最大点对应的水体指数为最优阈值。

1.1.3 区域种子生长算法

研究中发现:基于水体指数采用累积频率曲线拐点法确定最优阈值,提取的水体面积整体偏小,存在明显系统误差。经初步分析主要原因为:受影像空间分辨和水体指数法自身限制(水体边缘处指数存在突变),导致水体边缘提取不完整。为进一步提高水体提取精度,本文采用区域种子生长算法^[20]对水体边缘进行优化。

应用区域种子生长算法是对水体指数法的补充,为实现方法间的互补,生长准则的制定不易再基于水体指数。参考人工目视的原理,考虑到水对红外波段的吸收系数值非常大^[21],将R(红)G(绿)B(蓝)色彩模式中的R(红)替换成NIR(近红外),形成伪色域(NIRGB),并基于标准色差公式^[22]来制定生长准则。

通过归一化处理,得到每个像元的NIR、G和B的亮度值,见式(8)~(10)。区域生长时,以种子像元为基准,向外逐层生成缓冲区,搜寻连续的目标像元,并判断目标像元和种子像元的色差,见式(11)。当色差小于给定阈值时,将目标像元归为水体,实现增长。

$$L_{\text{NIR}} = \frac{R_{8,(x,y)} - R_{8,\min}}{R_{8,\max} - R_{8,\min}} \times 255 \quad (8)$$

$$L_g = \frac{R_{3,(x,y)} - R_{3,\min}}{R_{3,\max} - R_{3,\min}} \times 255 \quad (9)$$

$$L_b = \frac{R_{2,(x,y)} - R_{2,\min}}{R_{2,\max} - R_{2,\min}} \times 255 \quad (10)$$

$$\Delta L_c = \sqrt{\Delta L_{\text{NIR}}^2 + \Delta L_g^2 + \Delta L_b^2} \quad (11)$$

式中: x 为影像栅格的行号; y 为影像栅格的列号; $L_{\text{NIR}}、L_g$ 和 L_b 分别为 (x,y) 处近红外、绿和蓝光的亮度值; $R_8、R_3$ 和 R_2 分别为Sentinel-2的第8、3和2波段的遥感反射率; $R_{8,\max}、R_{3,\max}$ 和 $R_{2,\max}$ 分别为影像栅格中相应波段的最大遥感反射率值; $R_{8,\min}、R_{3,\min}$ 和 $R_{2,\min}$ 分别为影像栅格中相应波段的最小遥感反射率值; $R_{8,(x,y)}、R_{3,(x,y)}$ 和 $R_{2,(x,y)}$ 分别为 (x,y) 处相应波段的遥感反射率值; $\Delta L_c、\Delta L_{\text{NIR}}、\Delta L_g$ 和 ΔL_b 分别为目标像元与种子像元的色差、近红外、绿和蓝光亮度差值。

实际应用时可能会遇到生长不可控的情况,即将非水体误判为水体。其原因一是由于阈值设置过于宽泛;二是由于水体指法误将边缘处非水体确定为水体。因此区域种子生长算法运行前,需要进一步判断水体边缘处和孤立的水体像元是否为真实的水体。

1.2 库容曲线快速构建

1.2.1 库容计算

本文库容计算采用台柱体积法^[5]。首先把水体按不同水位分成 n 层,依据台柱体积公式计算每层的体积,经累加得到库容。

台柱体积公式:

$$V_i = \frac{1}{3} \Delta h (S_{i-1} + \sqrt{S_i S_{i-1}} + S_i) \quad (12)$$

累计库容计算公式:

$$V = \sum_{i=1}^n V_i + V_0 \quad (13)$$

式中: V_i 为相邻水位之间的库容差值,万 m^3 ; Δh 为相邻两水位的水位差, m; S_{i-1} 和 S_i 分别为两相邻水位对应的水面面积, $10^4 m^2$; i 为序数; n 为累计个数; V_0 为初始库容,万 m^3 ; V 为累计库容,万 m^3 。

1.2.2 库容曲线构建

本文研究目的是利用多时序的遥感影像快速自动构建水库库容曲线,更注重水库库容曲线构建的整体精度,并不需要每幅影像都能高精度地成功提取水体面积。通常受各种条件影响,水体指数法偶尔会出现不适用的情况。因此,当时序影像足够多时,可适当舍弃误差较大的,以保证整体的精度。库容曲线快速构建流程如下:①依据各景影像生成的时间,从水库观测资料中提取相应的水位;②使用水体指数法对预处理后的各景影像进行归一化处理,得到水体指数图;③对水体指数图进行统计分析,得到累积频率曲线;④对累积频率曲线进行简化线处理,以减少干扰;⑤在处理后的累积频率曲线上查找由凹变凸的拐点区间,并确定最优阈值,查找拐点前,可先依据水位变动情况,估算或通过人工目视解译选取影像中水体的最大和最小面积,并作为先验条件输入,以减少搜寻范围,提高效

率,若不存在由凹变凸的拐点区间,则舍弃该景影像;⑥依据最优阈值提取水体,利用区域种子生长算法对水体边缘进行优化,计算水体面积,建议输入库区最大矢量范围作为限制条件,以减少库外区域噪声的影响;⑦生成水位-水面面积散点图,并采用多项式^[6]对水位-面积曲线进行拟合;⑧判断水位-水面面积散点和拟合曲线的误差关系,舍弃误差较大的点,并重新拟合,反复进行该步骤直至各散点误差均满足精度要求为止;⑨依据拟合的水位-面积曲线得到等米间距的水面面积;⑩采用台柱体积公式计算库容差,并经累积得到水位-库容曲线。

2 案例分析

2.1 研究区概况

以河北省邯郸市武安市洺河流域的车谷水库为分析对象。该水库是一座综合利用的中型水利枢纽工程,总库容3 315万 m^3 ,坝址以上集水面积为124 km^2 ,集水区范围为36. 816~36. 995°N, 113. 724~113. 857°E,上游主要河流为管陶川,见图2。



图2 研究区概况

Fig. 2 Overview of research area

2.2 数据来源

2017年武安市水利局对车谷水库库容曲线进行复核,得到水库水位、水面面积和蓄水量等特征数据,见表1。

表1 车谷水库特征数据

Tab. 1 Characteristic data of Chegu Reservoir

水位/ m	面积/ km ²	相对蓄水量/ 万 m ³	水位/ m	面积/ km ²	相对蓄水量/ 万 m ³
681	0.413	0	695	0.782	838
683	0.475	89	697	0.853	1 002
685	0.529	189	699	0.911	1 179
687	0.574	300	701	0.991	1 368
689	0.621	419	703	1.058	1 573
691	0.672	548	705	1.132	1 792
693	0.731	688	707	1.203	2 026

本文选取 Sentinel-2 卫星多光谱遥感数据,共有 13 个波段,空间分辨率为 10、20、60 m, A 和 2B 两颗卫星同时运行,时间分辨率 5 d,时间和空间分辨率较高。选取 2016 年 6 月至 2020 年 12 月期间云量小于 10% 的 71 景影像,以及云量大于 15% 但库区无云的 2019 年 6 月 2 日和 7 月 2 日 2 景影像,共计 73 景,见图 3。所选影像均为经大气校正后的 L2A 级产品。分析前,使用欧空局推荐的哨兵数据应用平台(Sentinel Application Platform,简称 SNAP,版本 10. 0. 0),对各景影像进行重采样和裁剪,得到水库长时序的 10 m

分辨率影像。分析范围共 21. 62 km²。

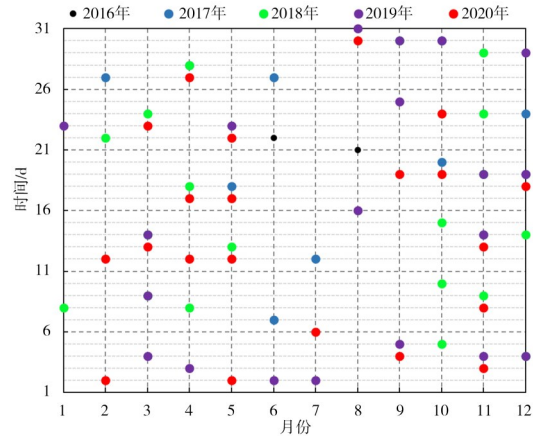


图3 选取卫星遥感影像的时间矩阵

Fig. 3 Temporal matrix of selected satellite remote sensing images

2.3 水体面积提取评估

以武安市车谷水库为研究对象,采用 JAVA 语言,使用 GDAL 库编制相关计算程序,分别计算 73 景影像的 NDWI、MNDWI、RWI、MBWI 和 AWEInsh 指数,先利用累积频率曲线拐点法提取水体,再采用区域种子生长算法对水体边缘进行优化,最后计算水体面积,并和 2017 年实测水位-面积进行比较,分析提取精度,见表 2 和图 4。依据水库蓄水情况,选取 2016-06-22(水位 681. 01 m)、2017-05-18(水位 704. 51 m)、2018-04-08(水位 698. 21 m)和 2020-09-04(水位 688. 99 m)作为典型时期进行分析,见表 3、图 5—7。

表2 水体面积提取成果

Tab. 2 Results of water area extraction

水体指数	类型	原始景数	累积频率曲线拐点法				应用区域种子生长算法后					
			遥感影像利用		误差分析/%		R ²	遥感影像利用		误差分析/%		R ²
			利用景数	利用率/%	相对	平均		利用景数	利用率/%	相对	平均	
NDWI			73	100	-9. 4 ~ 2. 0	4. 6	0. 992 40	73	100	-7. 3 ~ 5. 8	2. 5	0. 993 36
MNDWI	比值		66	90. 4	-16 ~ -0. 1	9. 4	0. 982 54	69	94. 5	-7. 3 ~ 4. 2	2. 9	0. 992 19
RWI		73	71	97. 3	-12. 3 ~ 1. 4	5. 9	0. 984 85	72	98. 6	-6. 2 ~ 4. 7	2. 5	0. 990 49
MBWI	差值		70	95. 9	-13. 4 ~ 4. 2	7. 6	0. 979 02	70	95. 9	-6. 7 ~ 5. 5	2. 6	0. 991 80
AWEInsh			68	93. 2	-17. 3 ~ 0. 6	9. 2	0. 974 97	71	97. 3	-7. 7 ~ 7. 9	3. 1	0. 989 83

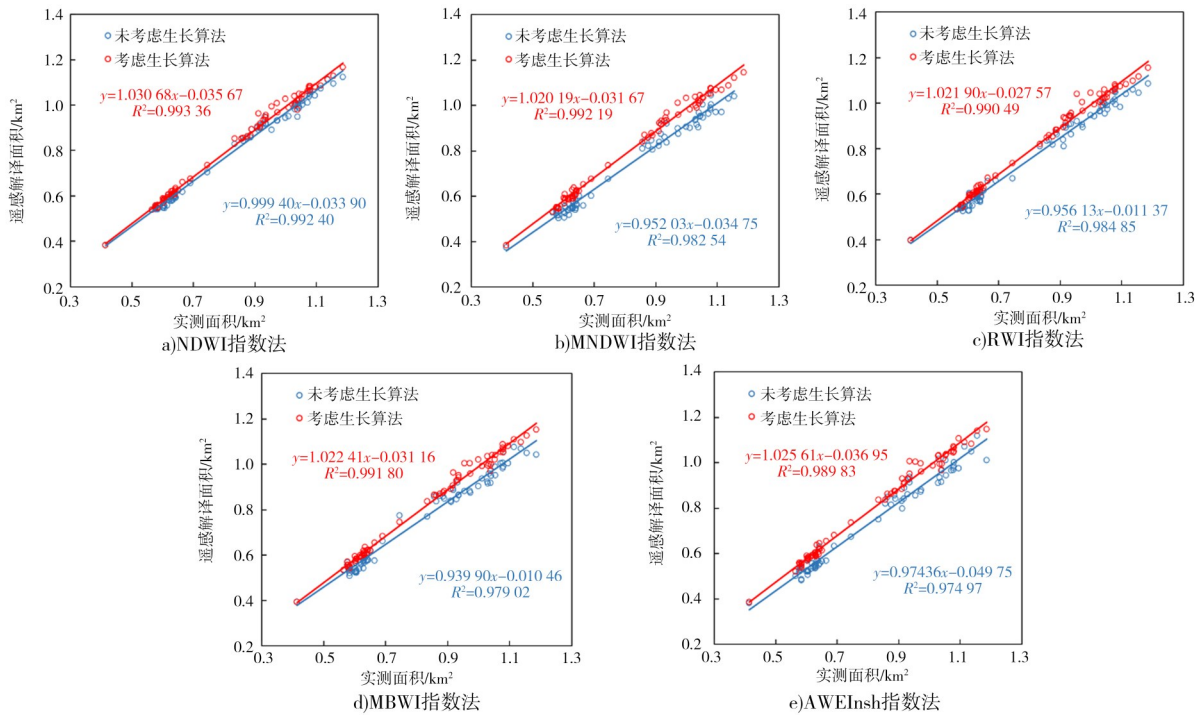


图 4 遥感解译水体面积和实测面积对比

Fig. 4 Comparison of remote sensing interpreted water body area and actual measured area

表 3 典型时期水体面积提取成果

Tab. 3 Results of water area extraction for typical periods

日期	水库特征		水体指数	累积频率曲线拐点法			应用边缘生长算法	
	水位/m	面积/km ²		阈值	面积/km ²	相对误差/%	面积/km ²	相对误差/%
20160622	681.01	0.413	NDWI	0.061	0.384	-7.1	0.385	-6.9
			MNDWI	-0.010	0.376	-9.1	0.385	-6.9
			RWI	-0.201	0.397	-4.0	0.398	-3.7
			MBWI	-4.040	0.394	-4.7	0.388	-6.0
			AWEInsh	-5.822	0.383	-7.4	0.387	-6.5
20170518	704.51	1.113	NDWI	0.074	1.074	-3.5	1.086	-2.4
			MNDWI	0.008	0.972	-12.6	1.070	-3.9
			RWI	-0.162	1.042	-6.4	1.086	-2.4
			MBWI	-3.848	1.076	-3.3	1.111	-0.2
			AWEInsh	-6.603	1.072	-3.6	1.092	-1.9
20180408	698.21	0.889	NDWI	-0.037	0.895	0.7	0.896	0.8
			MNDWI	0.018	0.807	-9.2	0.869	-2.2
			RWI	-0.209	0.864	-2.9	0.883	-0.7
			MBWI	-3.591	0.849	-4.5	0.884	-0.6
			AWEInsh	-5.313	0.842	-5.3	0.875	-1.6
20200904	688.99	0.621	NDWI	0.013	0.582	-6.2	0.597	-3.8
			MNDWI	0.042	0.547	-11.9	0.595	-4.2
			RWI	-0.208	0.580	-6.6	0.602	-3.1
			MBWI	-2.901	0.559	-9.9	0.597	-3.9
			AWEInsh	-3.143	0.542	-12.7	0.593	-4.5

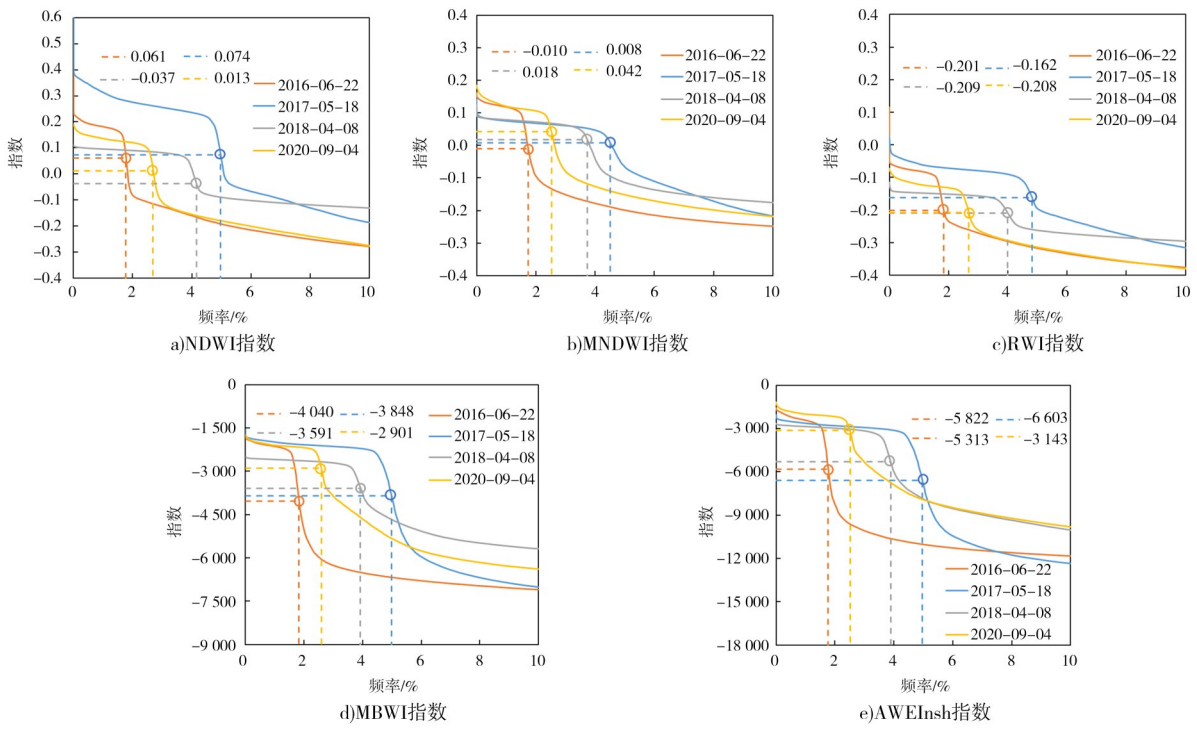


图5 典型时期累积频率曲线

Fig. 5 Cumulative frequency curves for typical periods

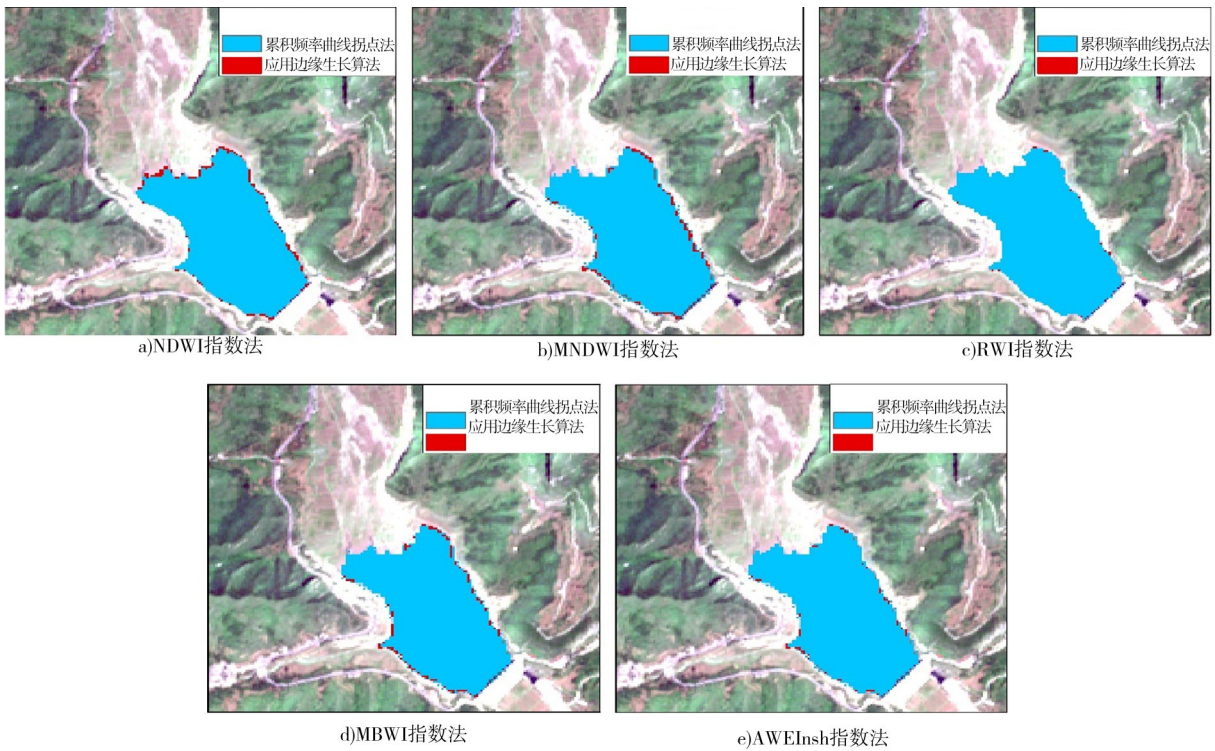


图6 2016年6月22号水体提取对比

Fig. 6 Comparison of water body extraction on June 22, 2016

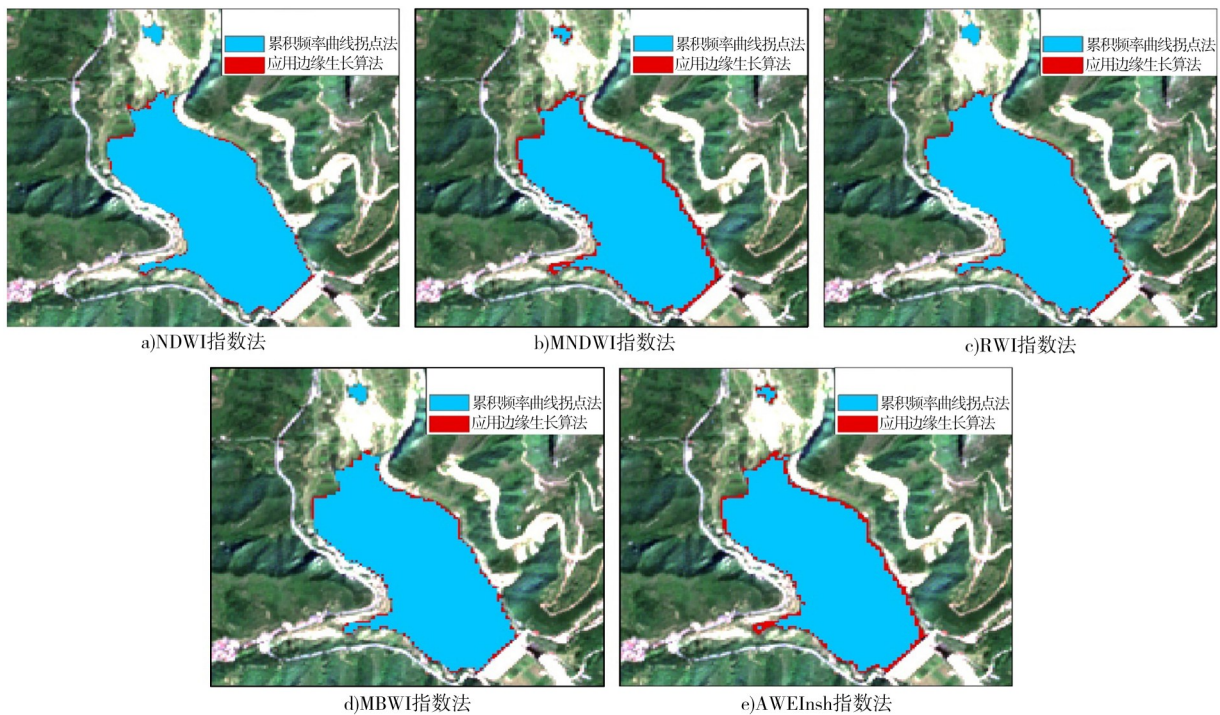


图 7 2020 年 9 月 4 号水体提取对比

Fig. 7 Comparison of water body extraction on September 4, 2020

程序运行过程中有 3 处主要参数需要设置:一是将累积频率曲线等分成多少份,均采用 500;二是由 DP 算法处理后,累积频率曲线至少保留多少个点,以水位-面积拟合较好,且影像利用率尽量高为条件进行试算,本文 NDWI、MNDWI、RWI、MBWI 和 AWEInsh 指数分别采用 200、250、150、250 和 200;三是生长准则的阈值,车谷水库水质较好,库区颜色比较一致,均采用 15。

经分析,采用频率曲线拐点法能够较好地确定最优阈值,各水体指数均能较好地利用遥感影像,NDWI 指数利用率最高为 100%,MNDWI 利用率最少为 90.4%。NDWI 指数和 RWI 指数提取水体比较完整,稳定性好,且 RWI 指数对细小水体的提取有更好的效果;MNDWI 指数法虽然在描述水体边缘具有明显优势,但水体边缘指数的明显变化,往往导致水体提取不完整,水体具有明显的不连续现象,导致误差较大。同样,MBWI 和 AWEInsh 指数也存在不同程度的边缘不连续现象,见图 6、7。

利用区域种子生长算法对提取水体边缘进行优化后,可明显提高水体提取质量,各指数法水体

面积提取的平均误差由 4.6%~9.4% 降低至 2.5%~3.1%。优化后 NDWI 指数提取水体效果最好,有效利用 73 景,水体面积相对误差-7.3%~5.8%,平均误差 2.5%, R^2 为 0.993 36;其次是 RWI 指数,有效利用 72 景,水体面积相对误差-6.2~4.7,平均误差 2.5%, R^2 为 0.990 49;MNDWI 指数仅利用 69 景,水体面积相对误差-7.3~4.2, R^2 为 0.992 19。

通过对典型时期水库水体提取分析可知,2016 年 6 月 22 日应用区域种子生长算法后水体面积变化不大,但与实测值比较相对误差在-3.7%~-6.9%。经人工目视检查 RWI 指数提取成果比实际水面略大,其他指数和实际水面基本吻合。因此,考虑该景影像提取的误差主要是由于水位观测导致的,其原因为:本次采用水位为日均水位,而影像生成时间为一天中某个时段。车谷水库承担下游供水任务,当水位较低蓄水量较少时,供水将导致一天内水位出现明显的波动。

综上所述,频率曲线拐点法对各类型指数最优阈值的确定均有较好的适用性,结合区域种子生长算法可以有效消除水体边缘不连续的现象,减少系

统误差,进一步提高水体提取精度。

2.4 水库水位-面积曲线拟合

水库水位-面积曲线常采用多项式进行拟合。利用不同时期的水位和提取的水体面积绘制散点

图,选取效果较好的二次多项式进行拟合,见图8。使用不同水体指数所拟合的水位-面积曲线的相关系数均大于0.97,说明得到的关系式能较好反映水库水位和水面面积之间的关系。

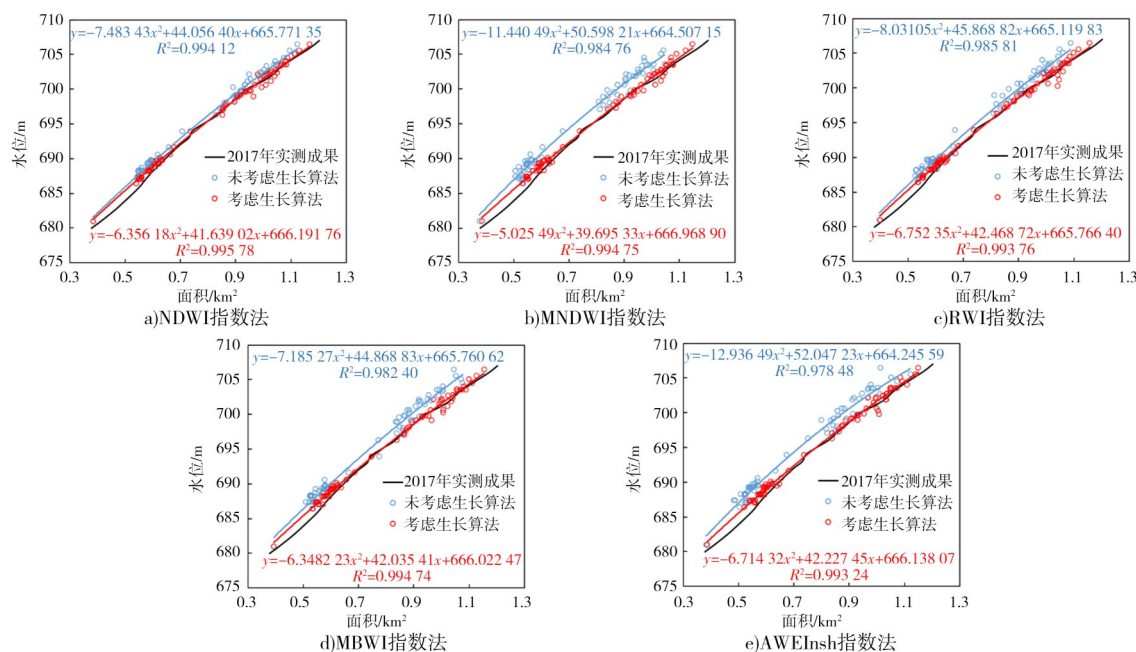


图8 车谷水库水位面积拟合成果

Fig. 8 Results of water level-area fitting for Chegu Reservoir

在选取73景影像中,车谷水库水位变动范围为681.01~706.47 m。因此,在681~707 m取等米水位,从水位-面积拟合曲线上提取面积,并和2017年测量面积进行对比可知:随着水位的抬高,拟合曲线提取的面积相对误差绝对值逐渐减少,特别是应用区域种子生长算法对水体边缘进行优化后,水位690 m以上提取水体面积和实测面积基本一致。各水体指数水位面积拟合误差见表4。

2.5 水库库容曲线构建

取681~707 m等米水位,在考虑水体边缘优化得到的水位-面积拟合曲线上,逐米查取相应的水面面积,采用台柱体积计算区间容积,经累加得到车谷水库681 m水位以上的水位-库容关系曲线(图9),并和2017年实测成果进行对比,分析生成库容曲线的精度,见表5。

表4 水位面积拟合效果

Tab. 4 Effect of water level-area fitting

水体指数	相对/%	
	累积频率曲线法	应用边缘生长算法
NDWI	-1.5~-10.8	1.8~-9.6
MNDWI	-6.7~-15.0	1.4~-10.7
RWI	-2.9~-10.7	2.1~-9.5
MBWI	-4.3~-12.5	1.6~-9.4
AWEI	-6.2~-16.0	1.1~-10.5
水体指数	平均/%	
	累积频率曲线法	应用边缘生长算法
NDWI	5.2	2.9
MNDWI	10.0	3.4
RWI	6.5	2.9
MBWI	8.0	3.0
AWEI	9.9	3.5

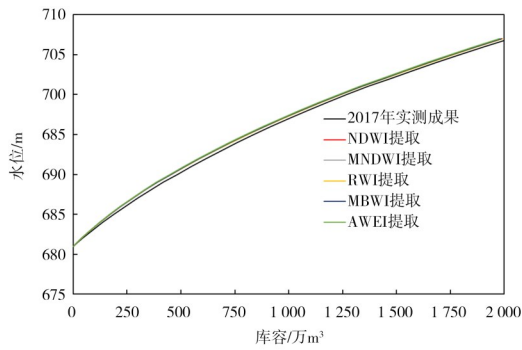


图 9 车谷水库库容曲线提取成果

Fig. 9 Results of Chegou Reservoir capacity curves

选取 73 景影像对应的最低水位为 681.01 m,且 681 ~ 686 m 无对应的影像数据,因此该水位下遥感影像提取的水体面积对低水位的库容影响较大。虽然人工目视检查各水体指数均将水体完全提取,但由于水位选取误差,导致提取面积仍较实测小。因此构建的水库库容随水位抬高,误差逐渐减少。

水位 707 m 时各水体指数法构建的库容相对误差为 -1.7% ~ -2.5%,和实测库容非常接近。

3 验证分析

武安市车谷水库位于山区,周边建筑物较少,应用本文提出的方法提取水库水体面积取得了比较满意的结果。为进一步验证本文提出方法的普适性,对武安市大洛远水库进行验证分析。大洛远水库位于武安市城区南侧,周边建筑物较多,属湖泊型水库。

武安市大洛远水库和车谷水库位于同一景遥感影像中,因此同样选用 73 景进行分析。基于 NDWI 和 RWI 指数法提取水体。分析时将累积频率曲线等分为 500 份;DP 算法处理时,至少保留 200 个点;考虑大洛远水库自库尾至坝前有明显的色差,生长准则的阈值相比车谷水库较大,取 20。经分析,NDWI 和 RWI 指数法分别有效提取 71 和 70 景。

表 5 水位库容提取效果

Tab. 5 Effect of extracting storage capacity from water level

2017年测量成果		NDWI		MNDWI		RWI		MBWI		AWEInsh	
水位/m	相对库容/万 m ³	相对库容/万 m ³	相对误差/%	相对库容/万 m ³	相对误差/%	相对库容/万 m ³	相对误差/%	相对库容/万 m ³	相对误差/%	相对库容/万 m ³	相对误差/%
681	0	0		0		0		0		0	
683	89	80	-9.4	79	-10.6	81	-9.3	81	-9.3	80	-10.3
685	189	172	-9.0	170	-10.0	173	-8.8	172	-8.9	171	-9.9
687	300	275	-8.1	273	-9.0	276	-7.9	276	-8.1	273	-9.0
689	419	390	-6.9	387	-7.6	392	-6.6	390	-6.8	387	-7.7
691	548	517	-5.7	514	-6.3	519	-5.4	517	-5.7	513	-6.5
693	688	656	-4.7	652	-5.2	658	-4.3	656	-4.7	651	-5.4
695	838	808	-3.7	803	-4.2	810	-3.3	807	-3.7	801	-4.4
697	1 002	972	-3.1	967	-3.6	975	-2.7	971	-3.2	964	-3.8
699	1 179	1 149	-2.6	1 143	-3.1	1 152	-2.3	1 147	-2.7	1 140	-3.3
701	1 368	1 339	-2.1	1 332	-2.6	1 342	-1.9	1 337	-2.3	1 329	-2.8
703	1 573	1 542	-2.0	1 534	-2.5	1 546	-1.8	1 539	-2.2	1 531	-2.7
705	1 792	1 759	-1.8	1 750	-2.4	1 762	-1.7	1 756	-2.0	1 747	-2.5
707	2 026	1 990	-1.8	1 979	-2.3	1 993	-1.7	1 986	-2.0	1 976	-2.5

重点对典型时期(2016-06-22、2018-12-14、2019-08-16 和 2020-04-27)进行分析,结果见图 10、11。通过人工目视检查,2种方法均将库区水体栅格提取出来。2020年4月27日2种指数对库区外左上角的水体提取存在明显差异。主要原因是 RWI 将该区域提取为较小的孤立水体,并且该区域

伪色域空间和库区均值的色差大于给定的阈值,被判断为非水体,从而未实现边缘生长。

整体来看,本文提出采用累积频率曲线拐点法和区域种子生长算法相结合的方法提取水体对位于城区的水库同样适用,具有较好的普适性。

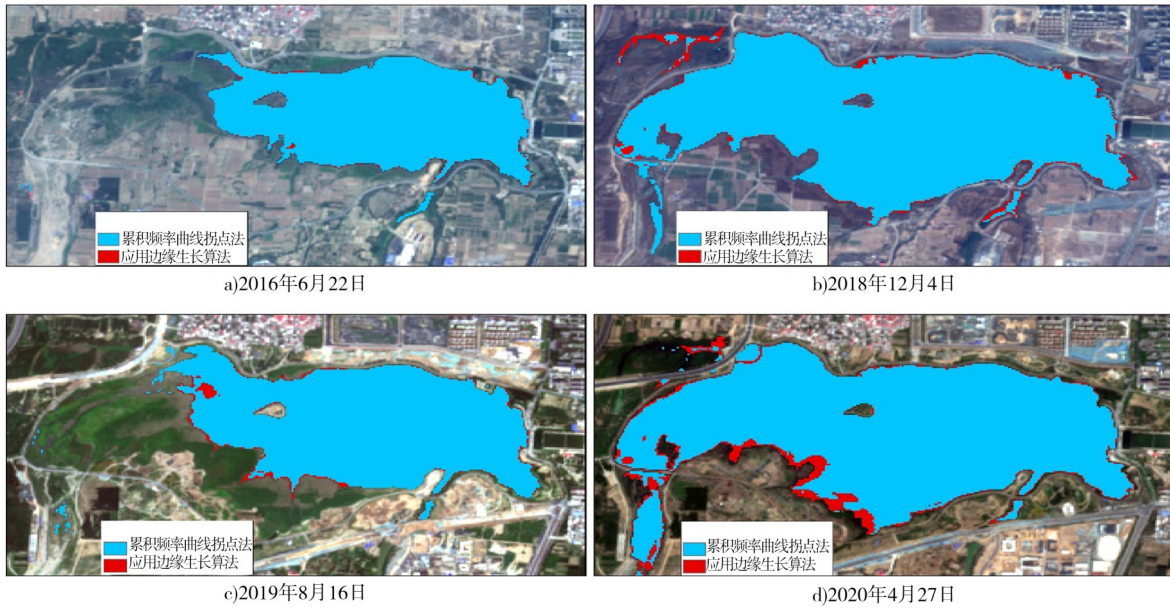


图 10 NDWI 水体指数提取成果
Fig. 10 Extraction results of NDWI

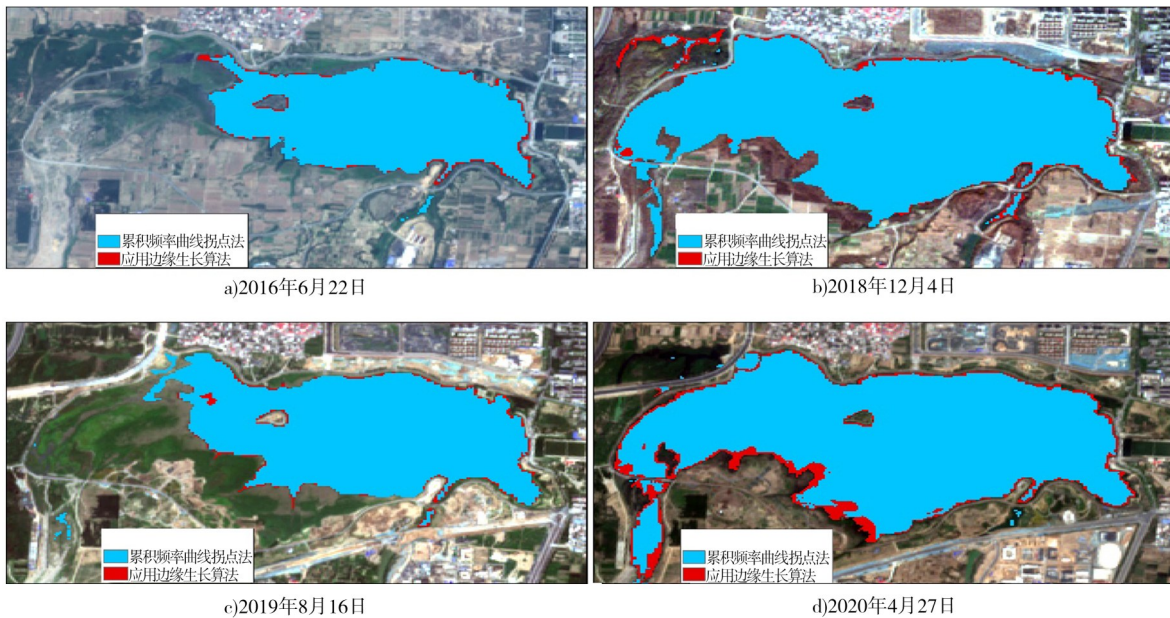


图 11 RWI 水体指数提取成果
Fig. 11 Extraction results of RWI

4 结论

中国中小型水库众多,传统的库容曲线复核工作通常需要投入大量的人力、物力和财力。随着中小型水库水雨情测报系统完善,积累了大量的水位观测资料,为使用高分遥感影像提取水库库容曲线提供了前提条件。本文基于 Sentinel-2 遥感影像,针对水体指数法,提出利用累积频率曲线拐点法自动识别最优阈值,采用区域种子生长算法对水体边缘进行优化,进而构建了一套快速构建水库库容曲线的完整自动化流程。以武安市车谷水库为研究对象,并在大洺远水库进行验证。

a) 水体指数法的原理是通过波段运算,经归一化处理后,将影像二值化,以区分水体和其他地表覆盖物。从统计学角度,基于聚类分类的思路,通过分析归一化处理后水体指数图的累积频率曲线,提出最优阈值位于该曲线上由凹变凸的拐点处,进而提出采用累积频率曲线拐点法确定最优阈值。该方法的提出为遥感水体的提取提供了新思路。

b) 针对基于水体指数法,采用累积频率曲线拐点法提取水体边缘不连续,存在明显系统误差的问题,提出基于伪色域(NIRGB)的生长准则,并利用区域种子生长算法优化水体边缘,有效提高了提取精度。

c) 本文构建了一套完整地自动化流程来快速构建水库库容曲线,并对细节问题进行深入分析。水库库容曲线提取的核心为快速构建水位-面积关系曲线,重点在于保证曲线整体的精度,当影像景数足够多时,不必苛求每景影像都具有较高的提取精度。通常由于受区域自然条件和成像时大气条件等环境条件影响,以及水体指数本身的局限性,有时并不存在最优阈值。本文在对车谷水库水体面积进行提取时,共选取 2016 年 6 月至 2020 年 12 月期间共 73 景影像。NDWI 水体指数的影像利用率达到 100%,但 MNDWI 水体指数利用率为 94.5%。

d) 选取比值型水体指数 NDWI、MNDWI 和 RWI,差值型水体指数 MBWI 和 AWEInsh,共 5 种指数进行分析。结果表明:频率曲线拐点法对各类型

指数最优阈值的确定均有较好的适用性和稳定性,且算法简单,可以充分利用先验条件,以提高搜寻效率。比值类指数水面提取效果整体优于差值型指数。在车谷水库水体面积提取中,NDWI 和 RWI 水体指数水体面积提取效果较好,平均误差均为 2.5%,相对误差分别为 -7.3% ~ 5.8% 和 -6.2% ~ 4.7%, R^2 分别为 0.993 36 和 0.990 49。误差主要受水位观测、水体面积提取效果和水位面积关系拟合精度等因素影响。

e) 在对车谷水库库容曲线提取中,采用 NDWI 和 RWI 水体指数提取精度较高,库容相对误差范围分别在 -9.4% ~ -1.8% 和 -9.3% ~ -1.7%,且随水位抬高误差越小,高水位提取的库容值和实测值基本一致。说明基于 Sentinel-2 遥感影像提取的水库库容曲线具有一定的精度,对大批量构建全国中小型水库库容曲线具有一定的参考意义。

f) 选取紧邻武安市城区的大洺远水库进行了验证分析,表明本文基于水体指数法提出的累积频率曲线拐点法和区域种子生长算法相结合的方法提取水体具有较好的普适性。

参考文献:

- [1] 2023 年全国水利发展统计公报[M]. 北京:中国水利水电出版社,2024.
- [2] 王国兴,李士鸿. 应用遥感资料获取库区水下地形的研究方法研究[J]. 河海大学学报(自然科学版),1998,26(6):91-94.
- [3] PHILIP A T, STEPHEN J W. Modeling flood plain inundation using an integrated GIS with radar and optical remote sensing[J]. *Geomorphology*, 1998, 21(3):295-312.
- [4] 陈曦,裴毅,姚帮松,等. 水位库容曲线的卫星影像测定方法研究[J]. 人民长江,2013,44(20):25-28.
- [5] 张莉芳,潘华海,单定军,等. 基于遥感技术的柘林水库库容曲线复核[J]. 水利水电技术,2017,48(6):1-6.
- [6] 王金鑫,桑学锋,常家轩,等. 多源时相遥感影像数据耦合的水库库容曲线重构[J]. 测绘通报,2021(11):42-47.
- [7] 吴庆双,汪明秀,申茜,等. Sentinel-2 遥感图像的细小水体提取[J]. 遥感学报,2022,26(4):781-794.
- [8] 徐涵秋. 水体遥感指数研究进展[J]. 福州大学学报(自然科学版),2021,49(5):613-625.
- [9] MCFEETERS S K. The use of the normalized difference water

- index (NDWI) in the delineation of open water features [J]. *International Journal of Remote Sensing*, 1996, 17 (7) : 1425-1432.
- [10] 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J]. *遥感学报*, 2005, 9(5):589-595.
- [11] WANG X B, XIE S P, ZHANG X L, et al. A robust multi-band water index (MBWI) for automated extraction of surface water from Landsat 8 OLI imagery [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2018, 68:73-91.
- [12] FEYISA G L, MEILBY H, FENSHOLT R, et al. Automated water extraction index: a new technique for surface water mapping using Landsat imagery [J]. *Remote Sensing of Environment*, 2014, 140: 23-35.
- [13] 王茹月,周航宇. 直方图双峰法在遥感影像水体提取中的研究与改进[J]. *测绘通报*, 2023(9):77-81.
- [14] 张珂,吴星宇,吴南,等. 基于高分一号遥感影像的水体提取方法对比分析与改进[J]. *水资源保护*, 2024, 40(4):9-16.
- [15] SEKERTEKIN A. A survey on global thresholding methods for mapping open water body using Sentinel-2 Satellite imagery and normalized difference water index [J]. *Archives of Computational Methods in Engineering*, 2021, 28:1335-1347.
- [16] MUKHERJEE A, KUMAR A A, RAMACHANDRAN P. Development of new index-based methodology for extraction of built-up area from Landsat7 imagery: comparison of performance with SVM, ANN, and existing indices [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(2): 1592-1603.
- [17] LI C M, SHAO Z F, ZHANG L, et al. A comparative analysis of index-based methods for impervious surface mapping using multi-Seasonal Sentinel-2 satellite data [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 3682-3694.
- [18] DOUGLAS DH, PEUCKER TK. 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.
- [19] 李涛,田晓君. 离散型函数拐点算法及应用[J]. *微计算机信息*, 2007, 23(6):248-249.
- [20] ADAMS R, BISCHOF L. Seeded region growing [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(6):641-647.
- [21] 邓孺孺,何颖清,秦雁,等. 近红外波段(900—2500nm)水吸收系数测量[J]. *遥感学报*, 2012, 16(1):192-206.
- [22] 李贵俊,刘正熙,游志胜,等. 一种基于色差和彩色归一化的车身颜色识别算法[J]. *计算机应用*, 2004(9):47-49.

(责任编辑:向 飞)