引文格式：王泽雨,李建章. 结合先验概率估计的SAR影像水体提取[J]. 地理空间信息, 2024, 22(8): 81-84. doi:10. 3969/j.issn. 1672-4623.2024.08.018

结合先验概率估计的SAR影像水体提取

王泽雨[1,2]，李建章[1,2*]
（1. 兰州交通大学 测绘与地理信息学院，甘肃 兰州 730030；
2. 兰州交通大学 地理国情监测技术应用国家地方联合工程研究中心 甘肃省地理国情监测工程实验室，甘肃 兰州 730070）

摘 要：提出了一种将水体概率分布图与阈值相结合的方法。该方法首先以DSWE水体概率数据集作为先验数据集，其次用格网将数据集按水体占比分为高、中、低三类，并将其作为指导阈值。然后将Sentinel-1数据经过SNAP预处理得到的后向散射系数按照先验数据集的水体占比分类，采用双峰直方图法、面向对象法和形态学膨胀对影像处理，进行水体提取。最后用同期高分辨率影像作为底图，进行目视解译。实验结果表明，该方法可对鄱阳湖区域的洪水淹没范围进行提取，并具有较高的精度，进一步证明了该方法在水体提取中的有效性和可靠性。

关键词：洪涝灾害；Sentinel-1；DSWE；水体分布数据集
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）08-0081-04

Water Extraction from SAR Images Combined with Prior Probability Estimation

WANG Zeyu[1,2], LI Jianzhang[1,2]
(1. School of Surveying and Mapping and Geographic Information, Lanzhou Jiaotong University, Lanzhou 730030, China;
2. Lanzhou Jiaotong University Geographic Condition Monitoring Technology Application National and Local Joint Engineering Research Center/Gansu Geographic Condition Monitoring Engineering Laboratory, Lanzhou 730070, China)

Abstract: Based on the SAR image and DSWE algorithm, we improved the existing threshold method, and proposed a method combining probabilistic distribution map with threshold value. We took DSWE water probability data set as the prior data set at first. And then, we divided the data set into three categories by water body ratio, such as high, medium and low, and took it as the guiding threshold. We preprocessed the Sentinel-1 data by SNAP software, classified the obtained backscatter coefficients according to the proportion of water in the prior data set, and used the bimodal histogram method, object-oriented method and morphological expansion to process the image and extract water. Finally, we used the simultaneous high-resolution image as the base image for visual interpretation. The experimental result shows that this method can extract the flood inundated area of Poyang Lake with high accuracy, which further proves the effectiveness and reliability of this method in water extraction.

Key words: flood disaster, Sentinel-1, DSWE, water distribution data set

SAR影像的阈值分割法主要是将影像处理为灰度图后选取合适的阈值进行筛选以提取水体对象[1]。由于区域物体特征表现具有差异性，根据区域和选取方式的不同，常用的阈值分割法为双峰直方图法、熵阈值法和Otsu算法[2-3]。

目前采用的影像分类方法中，面向对象法是建立影像与实际世界联系的理想方法。周婷[4]等提出了结合光谱主成分分析及支持向量机的水体提取算法，其识别精度显著高于NDWI单指标法和SVM算法，可为水系结构复杂、光谱特性空间差异大的河流提供理论参考；徐涛[5]等指出，将物体的光谱特征与其空间拓扑特征结合能有效区分城市阴影和水体；王俊海[6]等提出了可区分水体和阴影的ISTD指数，其提取精度较传统分类方法有显著提高，证实了该方法的水体提取的可靠性；崔齐[7]等以矢量数据作为约束条件，提出了一种可实现高空间分辨率遥感影像提取方法，该方法不仅能有效抑制建筑物及其他地物的阴影影响，还能获得具有高精度的细小水体信息；毛莹莹[8]等基于Landsat影像，并结合DEM数据和面向对象方法提取山区水体。研究结果表明该方法能够有效减少水体漏提现象，提升对山区细小水体的提取精度。

目前针对水体提取应用中，阈值法在小面积水体区域中精度较低，而机器学习方法的复杂性难以满足快速监测的需求。此外，现有的水体分布数据集多是

收稿日期：2023-02-01。
项目来源：国家自然科学基金资助项目（41661091）；兰州交通大学优秀平台资助项目（201806）。
第一作者简介：王泽雨（1997—），硕士，研究方向为遥感信息提取，E-mail：846030371@qq.com。
通信作者：李建章（1974—），博士，副教授，研究方向为全球定位系统与测量数据处理，E-mail：343529953@qq.com。

年度、月度合成数据产品。针对当前技术难点，本文在Sentinel-1数据和GEE（google earth engine）平台支持下，基于SAR影像与动态地表水范围DSWE（dynamic Surface water extent）算法，对现有阈值法进行改进，将生成的水体概率分布图与阈值法结合，并对2020年鄱阳湖特大洪灾范围进行水文提取，分析洪水发生前后鄱阳湖水面面积变化并绘制洪灾淹没区域图。结果表明，该方法实现了SAR影像高精度水体提取，可为未来水体提取算法优化、水资源精细化管理效率提高提供支撑和参考。

1 研究区域概况

鄱阳湖位于江西省北部，横跨九江，南昌和上饶三市，其常年平均降水量为1 500 mm左右，是亚热带湿润气候的典型区域之一。每年4-6月份，随着鄱阳湖流域内的洪水涌入，湖面水位不断上升；至7-9月时，湖水受长江洪水顶托或倒灌影响依旧保持较高水位，直至10月该高水位现象才逐渐下降。湖水的年水位变幅最高可达9.59~14.85 m，最小为3.54~9.59 m。2020年7月6日，受省内的强降雨和长江中下游洪水影响，鄱阳湖水位持续上涨，其湖口水位超过警戒线19.50 m。

2 数据与预处理

2.1 遥感影像数据来源

本文使用的卫星数据主要由两部分组成，第一部分是从阿拉斯加设施（ASF）网站下载了22幅Sentinel-1（IW）模式的GRD图像，该影像数据开始于洪灾发生前2020-06-02，并截止于到洪灾发生后2020-08-01；另一部分数据是在Google Earth Engine（GEE）云平台上调用的Landsat5和Landsat7数据，数据时间范围是从2000-01-01—2019-12-31，该部分数据可用于建立先验掩码。

2.2 数据的预处理

为更好地提取水体信息，本文在使用Sentinel-1（IW）模式的GRD图像进行分析之前先通过SNAP软件进行预处理，包含轨道校正、热噪声去除、辐射定标、多视、相干斑滤波、地形校正和分贝化7个步骤[9]。

3 研究方法

3.1 动态地表水范围（DSWE）算法

动态地表水范围（DSWE）算法生成栅格文件，表明在陆地卫星场景中的每个像素是否存在地表水。它的设计是通过限制对地球资源卫星和数字高程模型生成的地表反射率（SR）数据的必要输入，从而在全球范围内适用，DSWE算法涉及以下步骤。

1）对输入的高程数据进行预处理，生成阴影。

2）编写一个计算水体概率的函数，包含：MNDWI、MBSRV、MBSRN、AWESH、NDVI。对每个像素使用5种方法进行检测，对输出的结果进行条件判断（如果条件满足则为1；如果条件不满足则为0）。0是陆地的概率，1是水体的概率，这5个判断结果被组合成一个5位的二进制输出值（范围00000到11111），其中每个数字代表一个特定的测试结果，根据1的数量来确定最终是不是水体。计算公式如下：

$$\mathrm{MNDWI}=\frac{(\mathrm{Green}-\mathrm{SWIR1})}{(\mathrm{Green}+\mathrm{SWIR1})} \tag{1}$$

$$\mathrm{MBSRV}=\mathrm{Green}+\mathrm{Red} \tag{2}$$

$$\mathrm{MBSRN}=\mathrm{NIR}+\mathrm{SWIR1} \tag{3}$$

$$\mathrm{AWESH}=\mathrm{Blue}+(2.5\times\mathrm{Green})-(1.5\times\mathrm{MBSRN})-(0.25\times\mathrm{SWIR}_2) \tag{4}$$

$$\mathrm{NDVI}=\frac{(\mathrm{NIR}-\mathrm{Red})}{(\mathrm{NIR}+\mathrm{Red})} \tag{5}$$

3）在代码中编写3个分类函数，分别对应水（像素值=1和2）、土地（像素值=0）和不确定部分（像素值=3）。

4）将最终的分类结果输出到Google Drive。

动态地表水范围流程图DSWE算法总体流程图如图1。

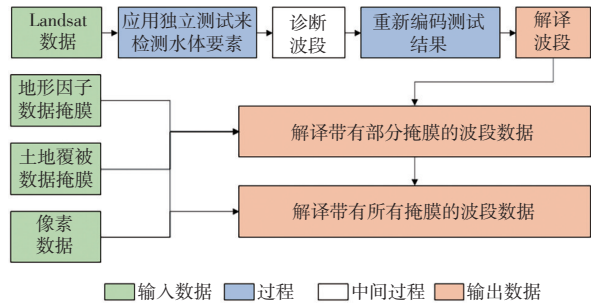


图1 动态地表水范围流量图

3.2 双峰直方图法

双峰直方图法是阈值分割中被广泛应用的方法之一。该方法是基于灰度图像进行处理，当灰度图像中存在2个明显的山峰特性时，可选取山峰间的谷底对应的灰度值作为切割阈值，该方法即PreWitt[10]等提出的双峰法，这是一种具有代表性的全局单阈值分割算法。具体的步骤如下：

1）首先统计图像的灰度信息，得到灰度直方图，设定初始阈值 th 和迭代次数；

2）初始设定第一个最高峰的灰度值 v_1 为0，第二个峰的灰度值 v_2 为255；

3）直方图中灰度值为 v_1 的个数记 P_1，灰度值为 v_1+1 时的个数记为 P_2。如果 $P_2>P_1$，则更新 v_1 的值为 v_1+1.峰值 v_2 则与 v_2-1 相比；

4）第一个峰值从左到右遍历到阈值 th，第二个峰值从右遍历到左边；

5）更新新的阈值为 $T=(v_1+v_2)/2$。

3.3　面向对象法

面向对象分类技术是通过临近像元集合，然后以此为基础作为特征要素去识别感兴趣的光谱要素的方法。该方法能够充分利用高分辨率的全色和多光谱数据的空间、纹理和光谱信息对影像进行分割和分类，具有高精度的分类结果或者矢量输出。该方法主要包括影像分割和影像分类2个步骤。其中影像分割是该技术核心，分割质量的好坏也将直接影响到影像信息提取的准确性，可以认为分割就是局部优化的过程，把图像划为不同的物体，使物体内部保留最大同质性，在对象之间保持最大异质性，从而尽可能减少“同谱异物”和“异谱同物”的干扰。

3.4　实验方法设计

本研究重新对水体提取过程进行完善，阈值水体范围提取方法的具体实验流程图如图2。

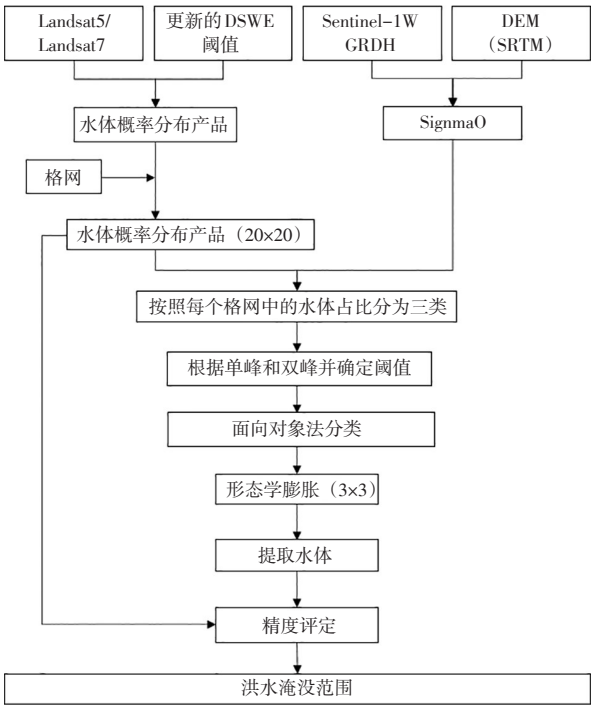


图2　实验流程图

4　实验与分析

4.1　动态地表水范围推导

基于GEE将2000—2019年间的Landsat5和Landsat7数据输入，并通过DSWE算法分别计算水体部分，不确定部分和陆地部分，最终形成动态地表水范围产品。再将得到的产品导出，按照每个格网大小7.5 km×7.5 km，可以将整幅影像分割成20×20的格网，考虑到影像中有部分区域远离鄱阳湖，不是此次研究重点，所以以鄱阳湖流域为中心选取周边12×11的格网作为研究对象，并计算网格内的水体面积比例，按照每个网格中水体在影像中占比分为高（66%~100%）、中（33%~66%）、低（0~33%）三类（图3）。

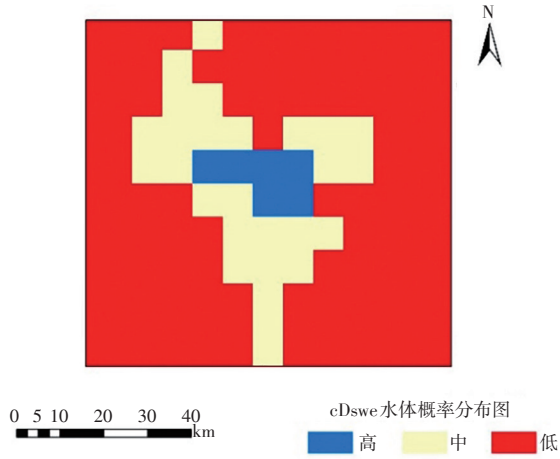


图3　水体在影像中的占比

4.2　水体提取

4.2.1　提取阈值

采用分层抽样，在3个分类中各随机选择3个格网，分别计算阈值，并提取水体。计算研究区2020年6月2日影像的3种分类（高>66%，中33%~66%，低0%~33%）共9幅网格的阈值。为了确定一个普遍的适应性阈值，求其平均值。最后，3个波段阈值确定如下：当水体占比为0 %~ 33%时，小于22.48标记为水；当水体占比为33%~66%时，小于-26.37标记为水；当水体占比为66%~100%时，小于-27.31标记为水。

4.2.2　整体提取结果

根据上一小节确定的阈值，采用面向对象法对研究区2020-06-08—2020-07-30时间范围内（图4）水体进行提取，然后对提取的结果进行形态学变化操作，蓝色部分为水体，从图中可以明显看出从6月8日到7月14日，洪水范围在不断的扩大。

4.3　洪水淹没区域

本文将2020-06-20作为鄱阳湖流域发生洪涝灾害的初期阶段，同年7月30日作为洪涝灾害后期阶段。通过将两阶段时期的水域提取面积进行叠加，计算出淹没区域。结果表明洪涝灾害严重的区域基本出现在鄱阳湖流域边缘地区，其中洪水淹没范围主要分

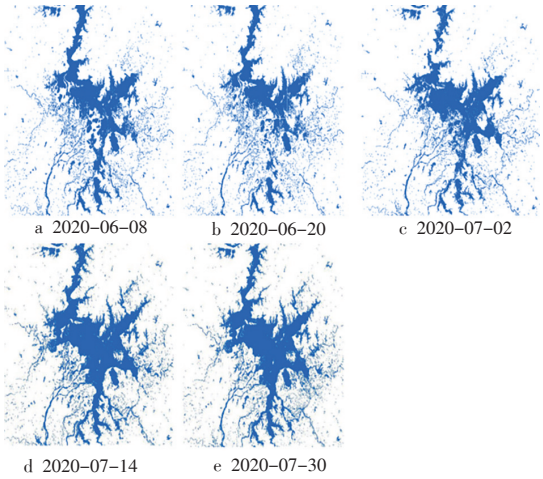


图4 整体提取的水体结果

布在西部、西南部和东北部。根据相关报道，自2020年6月中下旬开始，江西省连续出现大雨及暴雨。截至2020年7月5日凌晨时，鄱阳湖星子站水位已超过警界水位0.6 m（警戒水位19.00 m）；截至7月12日凌晨时，鄱阳湖附近多个水文站已超1998年水位，其水位高度突破历史最高纪录。实验提取结果与实际情况基本相符，说明该实验方法及结果具有较高的可信度。

4.4 精度评定

精度评定是确认影像分类中遥感解译结果的关键步骤。本研究使用ENVI数据进行精度评价，首先在影像上随机选取了200个验证点，然后通过目视对其进行解译判断，最后选取总体精度和Kappa系数进行精度评定，其具体结果见表1。

表1 精度评价结果

日期	总体精度	Kappa系数
2020-06-08	93.21	0.824 3
2020-06-20	90.75	0.752 6
2020-07-02	92.83	0.789 4
2020-07-14	95.36	0.885 1
2020/07/30	94.54	0.856 3

为了验证精度的可行性，选取了周小莉[11]对于鄱阳湖水体提取的精度，见表2。

表2 精度评价结果

方法	总体精度/%		Kappa系数	
MNDWI	75.57	80.57	0.54	0.6
NDVI	94.10	93.16	0.82	0.80
AWEInsh	87.39	89.04	0.75	0.75

可以看出，文中方法大部分的精度更高。总体精度总体大于90%，满足实验精度要求。

5 结 论

本文基于先验水体数据集对水体覆盖范围进行提取，该方法首先使用GEE云平台上2000—2019年的Landsat数据，利用DSWE算法对数据进行运算得到水体分布概率图；其次对预处理后的Sentinel-1数据按照在水体概率图中水体占比不同划分区域；然后对这些区域分别进行阈值提取；最后采用面向对象法进行水体提取。现有的大多数水体产品使用的是年度综合数据，该方法提取的淹没区范围在一定程度上弥补了光学产品低时间分辨率的问题，此外该方法在小面积水体区域中比其他方法提取水体结果更加准确。从一定程度上改善了水体算法在小区域内分类精度低的问题。

参考文献

[1] 李丹,吴保生,陈博伟,等. 基于卫星遥感的水体信息提取研究进展与展望[J]. 清华大学学报(自然科学版),2020,60(2):147-161

[2] 殷悦,宫辉力,赵文吉. 基于SAR影像的洪水淹没范围信息提取的研究[J]. 测绘与空间地理信息,2007,29(4):50-54

[3] 王航,秦奋. 遥感影像水体提取研究综述[J]. 测绘科学,2018,43(5):23-32

[4] 周婷,邹俊,崔玉环,等. 基于主成分分析和支持向量机的遥感影像水体提取方法及应用[J]. 水资源保护,2022,37(4):1-15

[5] 徐涛,谭宗坤,闫小平. 面向对象的城市水体信息提取方法[J]. 地理空间信息,2010,8(3):64-66

[6] 王俊海,阮仁宗,柴颖,等. 基于高分二号的面向对象城市水体信息提取[J]. 地理空间信息,2018,16(9):34-37

[7] 崔齐,王杰,汪闽,等. 矢量约束的面向对象高分遥感影像水体提取[J]. 遥感信息,2018,33(4):115-121

[8] 毛莹莹,陈友飞. 基于面向对象法的Landsat8影像山区细小水体提取方法[J]. 亚热带资源与环境学报,2015,10(4):86-92

[9] 姚熠,段莹,李光一,等. 卫星遥感资料在山区洪涝灾害监测及决策服务中的应用[J]. 中低纬山地气象,2022,46(2):76-81

[10] 杨修国. 关于直方图双峰法的研究与改进[J]. 电子设计工程,2012,20(12):176-179

[11] 周小莉,郭加伟,刘锟铭.基于陆地成像仪影像和主成分分析的水体信息提取:以鄱阳湖区为例[J].激光与光电子学进展,2016,53(8):83-90