

黄河源区河宽的高效提取算法及年内变化特征

谢朝帅^{1,2,3}, 刘苏峡^{1,2,3}, 周浩伟^{1,3}, 莫兴国^{1,2,3}

(1. 中国科学院地理科学与资源研究所 中国科学院陆地水循环及地表过程重点实验室, 北京 100101;

2. 中国科学院大学中丹学院, 北京 100049; 3. 中国科学院大学, 北京 100049)

摘要: 河流宽度是河流动态过程研究和水文、水动力模型的关键参数之一。传统水文测站无法观测河宽的空间分布, 而卫星遥感影像可以实现河宽的大范围观测。本文选取黄河上游唐乃亥水文站以上流域作为研究区, 基于 Sentinel-2 遥感影像, 采用 Google Earth Engine 大数据云计算平台和 RivWidth 方法提取河宽, 并开发了一种新的河流连通性的构建方法, 提高了 RivWidth 的运算效率, 最终生成了 2019—2021 年黄河源区 10 m 空间分辨率的逐月河流宽度数据 (Monthly River Width Data, MRWD)。结果表明, 本文开发的河流连通性的构建方法使得 RivWidth 的运算效率提高了 24.7%。逐月河流宽度数据包括 6.1×10^6 个河流宽度估算值, 覆盖了黄河源区 2035 km 的河流。与水文站点实测河宽和全球河流宽度数据河宽 (Global River Widths from Landsat, GRWL) 对比验证结果显示, 本文新算法反演的逐月河流宽度数据的平均绝对值百分比误差、均方根误差和相对均方根误差分别为 7.82%、11.62 m、8.84%。基于逐月河流宽度数据, 本文发现黄河源区范围内大约有 41% 的河道窄于 100 m, 85% 以上的河道宽度小于 500 m, 仅 2% 的河道宽度超过 1000 m。本文生成的逐月河流宽度数据可为地表水资源管理提供重要参考依据。

关键词: Sentinel-2; GEE; 黄河源区; 河宽; RivWidth 河宽提取算法

DOI: 10.11821/dlxb202405004

1 引言

河流宽度是重要的水力学参数之一, 可以用于河流流量估算^[1-2], 河岸演变模拟研究^[3]、河流的生态质量评价^[4]、水文模型建模^[5-8]、河流与大气的碳通量交换^[9]、洪水前后河流最小形态空间需求研究^[10]等。因此, 构建准确的河流宽度数据至关重要。

水文站的地面观测是获取河宽数据的传统手段之一^[11-13]。然而, 由于维护资金的不足, 自 1980 年以来全球测站网络不断缩小, 加上国际数据共享政策的限制^[14-15], 通过水文站点获取河宽数据非常困难, 为了满足区域尺度上的河流宽度数据需求, 很多研究通过野外调查、航空照片或卫星图像以目视解译的方式获取河宽^[16-20]。然而, 目视解译费时费力, 且由于受到人为主观意识的影响, 目视解译得到的河宽也与实际河道情况之间存在着误差^[21]。对于区域尺度或者大型河流, 大多通过河流宽度与河流流量或者流域面积之间的经验关系, 得到河流宽度数据^[22-26]。经验关系无法精确捕捉河宽的局部特征和变

收稿日期: 2023-04-12; 修订日期: 2023-12-08

基金项目: 国家重点研发计划(2018YFE0106500); 第二次青藏高原综合科学考察研究(2019QZKK0903, 2019QZKK0403); 第三次新疆综合科学考察项目(2021xjkk0803) [Foundation: National Key R&D Program of China, No.2018YFE0106500; The Second Tibetan Plateau Scientific Expedition and Research, No.2019QZKK0903, No.2019QZKK0403; The Third Xinjiang Scientific Expedition, No.2021xjkk0803]

作者简介: 谢朝帅(2000-), 男, 河南信阳人, 硕士生, 主要从事水文水资源研究。E-mail: xiechaoshuai20@mails.ucas.ac.cn

通讯作者: 刘苏峡(1965-), 女, 湖北黄陂人, 研究员, 博士生导师, 主要从事水文水资源研究。E-mail: liusx@igsnrr.ac.cn

1146-1160 页

化,且由于气候地貌条件差异,不同流域的河流宽度和流量或流域面积之间的经验关系是不同的^[27]。

作为一个集成了大量卫星图像数据并提供了快捷高效的计算能力的云计算平台,谷歌地球引擎(Google Earth Engine, GEE)逐渐成为处理和分析大范围长历时河宽动态的非常合适的方法。此外,MODIS、Landsat和Sentinel等遥感卫星的快速发展为时间尺度更长、空间范围更大的河宽提取提供了重要的基础观测数据。借助卫星图像数据,在没有观测数据的流域中提取河宽的替代方法逐渐发展起来。例如,Yamazaki等^[27]利用航天飞机雷达地形测绘任务(Shuttle Radar Topography Mission, SRTM)水体数据和HydroSHEDS水流方向数据,开发了大型河流全球宽度数据库(Global Width Database for Large Rivers, GWD-LR)。然而,由于使用了SRTM水体数据源,GWD-LR对于河宽小于300 m的河段覆盖范围有限。针对这一问题,Allen等^[9, 28]应用Landsat数据和河宽自动提取工具Rivwidth开发了更大规模更细分辨率的北美河流宽度数据(North American River Width, NARWidth)和全球河流宽度数据(Global River Widths from Landsat, GRWL)。另外,Isikdogan等^[29]使用河宽自动提取工具RivaMap为北美河流开发了大陆尺度的河宽数据,此河宽数据在与NARWidth对比的结果中显示出了较高的准确性。由于GWD-LR只能获取SRTM数据共有时段的河宽数据,NARWidth和GRWL数据计算的是平均流量状态下的单时相河宽,因此以上数据都缺少河宽的多时段动态分析。为了解决这一问题,Yang等^[30]基于全球地表水数据(Global Surface Water, GSW)数据制作了1990—2015年的中国多时相河流宽度数据(Multi-temporal China River Width, MCRW),并分析了河宽的季节波动和动态淹没频率。然而,以上研究大多使用Landsat卫星作为数据源,提取得到的高精度有效河宽仅为90 m,同时这些研究大多是在特定的气象水文条件下得到的单时相河流形态,虽然MCRW分析了多年融合河宽数据的季节变化,但是有关河宽数据的年内动态变化特征分析较少。此外,河道上的水坝和桥梁等建设会阻碍河流的连通性^[31],现有的河宽数据和河网提取方法中还没有明确的解决措施,Yamazaki等^[27]也指出河道连通性的缺失会影响有效河流宽度的提取。因此,目前对于流域尺度(如黄河源区)的高时空分辨率的河流宽度和动态研究不足,考虑河道连通性缺失的相关研究也较少。

为此,本文的目标是:①开发一种河流连通性的构建方法;②建立黄河源区的高时空分辨率河流宽度数据;③分析2019—2021年黄河源区的河流宽度时空动态。具体而言:①针对河道连通性的缺失问题,提出了一种基于水体区域分组和最佳水体区域连接的补全河道连通性的方法;②借助Sentinel-2卫星数据和GEE云处理平台,构建了黄河源区第一个空间分辨率为10 m的逐月河流宽度数据(Monthly River Width Dataset, MRWD);③通过提取黄河源区干流并结合统计分析数据,探究了黄河源区不同区域河流宽度的时空格局和年内动态变化特征;④使用水文年鉴原位测量数据验证了MRWD的准确性,并与现有数据进行了比较。本文提取得到的逐月河流宽度数据MRWD可作为河流形态演变、水文和水动力模型、极端气候事件等研究的基础资料。

2 研究区和数据

2.1 研究区

黄河源区(32°06'N~36°10'N, 95°47'E~103°27'E)位于中国青藏高原东北部(图1),横跨青海、四川和甘肃3个省份,总流域面积约为12.2万km²,占整个黄河流域面积的16.2%。唐乃亥水文站控制断面以上河长1553 km,占黄河全长的28.4%。多年平均径流

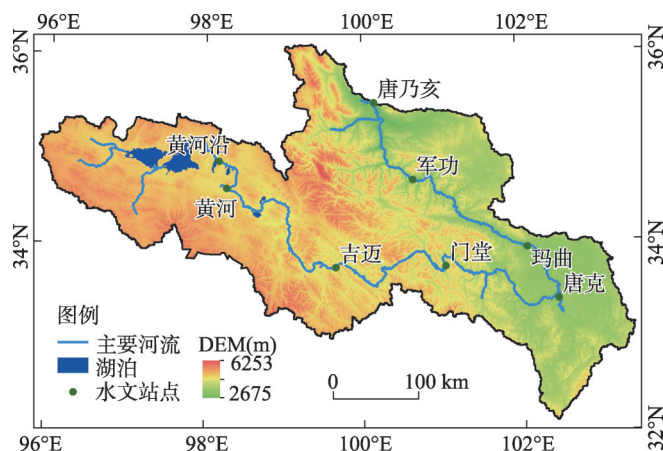


图1 黄河源区水文站点、河流和湖泊分布

Fig. 1 Distribution of the gauging stations, rivers, and lakes in the source region of the Yellow River (SRYR)

量 200 亿 m^3 ，占黄河流域多年平均径流量的 34.5%。黄河源区的地形以山地、丘陵和盆地为主，自西向东，海拔逐渐降低，平均海拔在 3500 m 以上。流域年平均降水量为 522 mm，多集中在夏季，空间上有东南向西北递减的趋势。年平均气温在西北地区 $-4\text{ }^{\circ}\text{C}$ 至东南地区 $2\text{ }^{\circ}\text{C}$ 之间的范围内波动。黄河源区河流地貌和侵蚀类型多样，河流峡谷众多，水资源较为丰富^[32]。位于黄河源区西北部的扎陵湖和鄂陵湖是青海省最大的两个淡水湖。流域中部的阿尼玛卿山主峰海拔达到 6282 m，其上的哈龙冰川是黄河流域最大的冰川^[33-34]。

2.2 数据

2.2.1 Sentinel-2 卫星数据 本文基于地表反射率 (Surface Reflectance, SR) 的 Sentinel-2 卫星数据产品提取黄河源区的地表水体和河宽数据。Sentinel-2 卫星包括 Sentinel-2A 和 Sentinel-2B 两颗同轨道光学遥感卫星，由欧洲空间局 (European Space Agency, ESA) 先后于 2015 年 6 月和 2017 年 3 月发射升空，卫星在轨高度为 786 km，重访周期为 5 d，空间分辨率为 10 m。SR 卫星产品是使用 SNAP 软件进行大气校正后的高精度下垫面反射率卫星数据，并已全部集成于 GEE 平台中。文中使用的 Sentinel-2 卫星图像的空间分布和数量如图 2 所示，卫星产品的有效数据时段范围为 2019 年 1 月—2021 年 12 月，所有月份的卫星图像数量共计 10967 幅。

2.2.2 其他辅助数据 研究中使用到的辅助数据包括：① 航天飞机雷达地形任务数字高程数据 (Shuttle Radar Topography Mission, SRTM)^[35]，空间分辨率为 30 m，SRTM 数据生成的斜坡掩膜，用于去除容易被误标记为水像素的山体阴影像素；② 1975—1990 年、1990—2000 年和 2000—2014 年 3 个时期空间分辨率为 30 m 的全球人类定居区和建成区网格数据 (Global Human Settlement Layers and Built-Up Grid Data, GHSL, <https://ghsl.jrc.ec.europa.eu/download.php>)，用以去除容易被误标记为水像素的建筑网格像素；③ 空间分辨率为 30 m 的全球地表水数据 (GSW)^[36]，利用其 max_extent 波段和研究区多年水体频率图像去除月融合水体图像中的椒盐噪声；④ 全球水库和大坝数据库 (Global Reservoir and Dam Database, GRanD)^[37] 和全球湖泊和湿地数据库 (Global Lakes and Wetlands Database, GLWD)^[38]，用于去除湖泊和水库；⑤ 全球河流宽度数据 GRWL^[28]，用于和 MRWD 数据进行河宽频数统计对比；⑥ 黄河源区主要水文站点的原位测量河宽数据，用于河流宽度数据的验证。

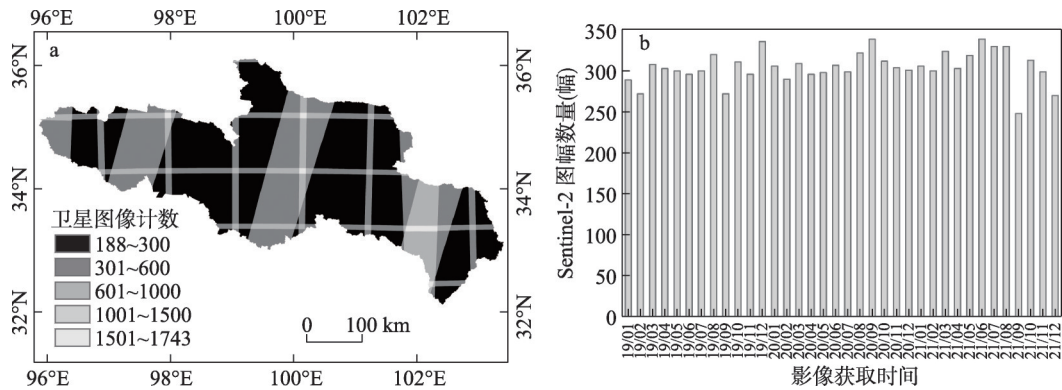


图2 黄河源区卫星图像分布和图幅数量统计

Fig. 2 Spatial distribution and monthly number of all collected Sentinel-2 images in the SRYR

3 研究方法

3.1 水体提取

水体指数法通过构建水体指数模型分析水体与光谱值之间的关系，并设置阈值实现地表水体信息的提取，此方法突出了水体特征，减少了植被、土壤、建筑和阴影的影响，逐渐成为应用频次最高的地表水体提取方法。常用的水体指数包括改进的归一化差异植被指数MNDWI、归一化植被指数NDVI、增强水体指数EVI等。基于以上水体指数，2017年Zou等^[39]提出了一种新的提取地表水体的方法（(MNDWI > NDVI or MNDWI < EVI) and EVI < 0.1），并被广泛应用于水文研究中^[40-44]。然而，Zhou等^[45]发现该标准获得的地表水体结果在黄河源区中是不可靠的，无论是湖泊还是河流，提取的水体信息均不完整，当标准变更为（MNDWI > NDVI or MNDWI < EVI）水体信息的提取效果得到明显改善。此外，Chen等^[46]和Deng等^[47]也发现Zou等^[39]地表水体的方法在塔里木河流域和长江流域中东部地区表现不佳。因此，本文将使用Zhou等^[45]水体指数标准提取地表水体信息，各水体指数的计算公式如下：

$$MNDWI = \frac{\rho_{green} - \rho_{SWIR1}}{\rho_{green} + \rho_{SWIR1}} \quad (1)$$

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \quad (2)$$

$$EVI = 2.5 \times \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + 6 \times \rho_{red} - 7.5 \times \rho_{blue} + 1} \quad (3)$$

式中： ρ_{green} 、 ρ_{red} 、 ρ_{blue} 、 ρ_{NIR} 和 ρ_{SWIR1} 分别为Sentinel-2卫星图像的绿色波段、红色波段、蓝色波段、近红外波段和短波红外波段。

黄河源区月尺度水体数据的建立主要包括以下几个步骤。① 使用QA波段对2019年1月—2021年12月每个月度时间范围内经过黄河源区的所有Sentinel-2卫星图像进行去除云、云阴影、雪像素的预处理，其余像素表示为有效像素。② 利用SRTM 30 m分辨率的DEM数据计算坡度，并掩膜得到坡度小于10的像素，去除山地阴影。③ 使用GHSL数据来掩膜图像范围内的建筑网格像素，消除错误标记为水体像素。④ 利用GEE云计算平台上的median_composite算法进行月尺度图像融合，使用GSW的最大历史水范围波段max_extent，进一步去除残余的云、云阴影和雪像素引起的椒盐噪声^[47]。⑤ 使用Zhou等^[45]提出的水体指数标准提取水体，并利用基于历史水频率的水体插值方法（Surface Water

Gaps, SWGS)^[45]对最终结果进行水体信息的插补,得到黄河源区2019年1月—2021年12月的逐月水体数据。

3.2 最佳水体区域连接线的构建

2008年Pavelsky等提出一种借助水体信息图自动估算河流宽度的工具RivWidth^[48],在大范围河流宽度数据提取方面具有较高的可靠性,目前已被广泛应用^[9, 28-29, 49]。然而,RivWidth只能计算水体提取结果中的最大连续水体的河流宽度。针对该问题,Yang等^[30]使用迭代的方法(简称迭代法)提取河流宽度,该方法的原理是在给定的水体栅格结果中搜索最大的河流并计算它的宽度,随后删除这条河流栅格数据,并将RivWidth应用于剩余的最大的河流,迭代遍历水体栅格。然而迭代法需要大量计算资源,且迭代停止条件难以确定。因此,本文开发了一种最佳水体区域连接线的构建方法(简称连接法),以达到补全水体信息和单次运算得到研究区全部河流宽度数据的目的。

基于Zhou等^[45]的水体提取方法和SWGS插值得到的研究区域水体栅格数据(图3b),连接法构建水体区域最佳连接线主要分为以下几个步骤:①基于GSW的最大历史水频率范围波段max_extent^[36],使用ArcGIS软件构建30 m缓冲区,掩膜研究区域的月度水体栅格数据;②在GSW的30 m缓冲区掩膜后的水体栅格数据的基础上,将全球水库和大坝数据库GRDD和全球湖泊和湿地数据库GLWD作为参考数据,去除研究区水体栅格数据中的湖泊和水库;③计算湖泊和水库去除后的水体栅格数据中每个河段的水体覆盖面积,去除水体面积小于900 m²的河段,筛选出黄河源区的月度水体栅格数据的主河道水体栅格数据(图3c),因为本文主要关注河流;④使用ArcGIS软件中的区域分组工具对主河道水体栅格数据进行水体区域分组,为每一块连续的水体区域分配唯一的编号(图3d);⑤通过求解区域分组后的各个水体区域之间的距离最短连接线的方式构建水体区域的最佳连接线(图3e);⑥将水体区域最佳连接线通过的主河道水体栅格数据中的非水体区域进行赋值,转换为水体区域,得到研究区域的月度主河道连续水体栅格数据(图3f)。表1展示了连接法的伪代码,其中区域分组处理时采用8相邻像元数方法评估水体栅格区域之间的连通性(评估每个输入像元正交和对角线总计8个最近相邻像元的连通性,相互连接的且具有相同值的栅格像元将构成单个水体区域)。此外,处理得到的最佳水体区域连接线应具备以下两点属性:①贯穿并连接各个水体栅格区域;②连接线的总长度最短。

3.3 河宽数据MRWD的提取

借助水体信息图自动估算河流宽度,需要一种合适的算法。RivWidth方法基于二进制水体信息图(水为1,非水像素为0),使用三角法计算河流中心线中每个像素的宽度,逐渐成为河宽自动提取的常用方法之一。为了从水体信息图中估算河流宽度,RivWidth方法的计算过程主要分为以下几个步骤:①RivWidth方法使用形态距离变换测量水体信息图中每个水像素到最近的非水像素的欧氏距离^[50],生成一个距离矩阵图(图4c);②使用双向拉普拉斯滤波器^[51]对距离图进行卷积,求得各河道的中心线(图4d);③通过将每个正交线和河道相交的宽度相加,求算河流宽度^[48]。本文选择RivWidth方法提取河宽的原因是:①它能够自动从输入的二进制水体信息图中提取河流宽度;②黄河源区有许多河流发源于青藏高原山区,积雪和阴影对MNDWI响应较高,采用RivWidth方法可以减少误差。

3.4 河宽数据MRWD的验证

为了评估MRWD数据的准确性,将处理得到的河流宽度数据MRWD与现场实测河宽数据进行比较。本文汇总了研究区的水文站点的河宽原位记录,将相应的原位宽度与由Allen等^[9]和Pavelsky等^[28]进行的最接近当前站的相应MRWD估计的平均宽度(最邻近

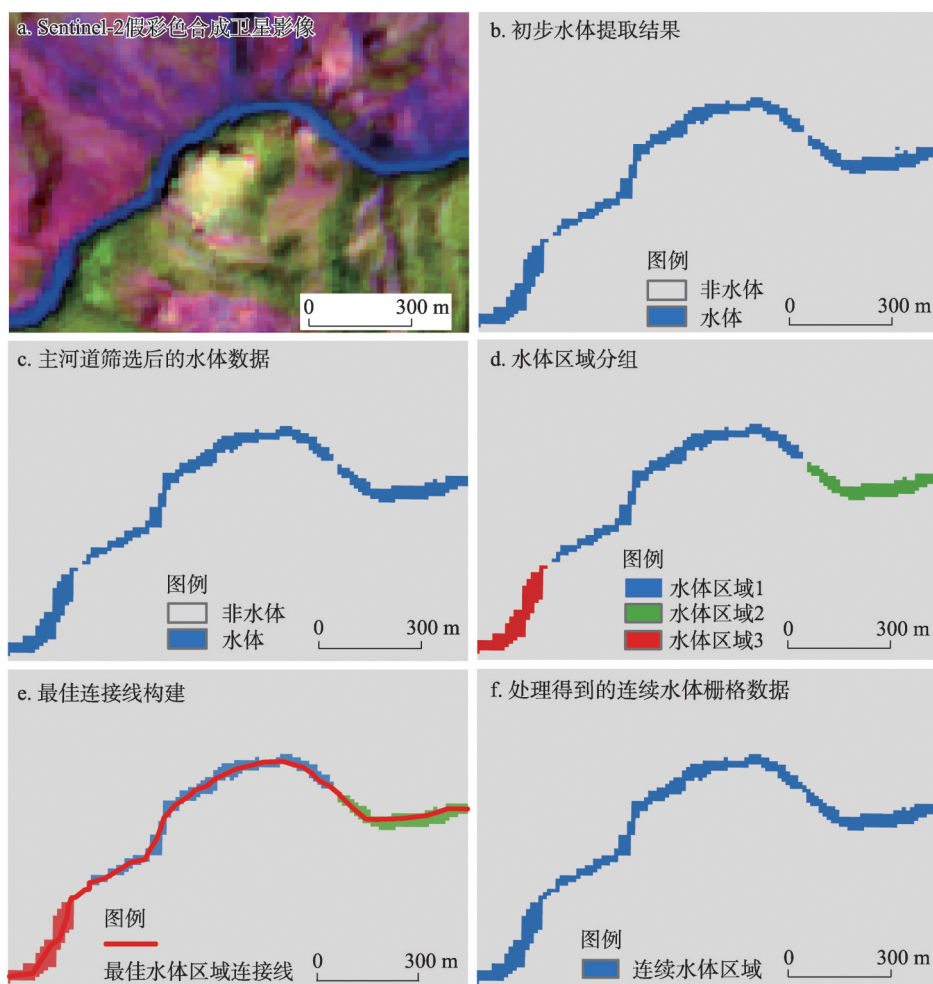


图3 连接法的输入、主要步骤和输出结果

Fig. 3 Input, main steps and output of the connection method

5个河宽的平均值)进行比较。日尺度上河流宽度数据的验证,依照各个水文站点划定的区域边界范围进行河流宽度验证日期的选择,原则如下:①优先考虑Sentinel-2卫星影像与水文年鉴的河宽测定相同天的日期,其次是间隔不超过24 h的其他日期;②当多个日期满足条件时,挑选Sentinel-2卫星影像云层覆盖率最低的日期;③验证日期的卫星影像云层覆盖率需小于20%。使用平均绝对百分误差(MAPE),均方根误差(RMSE)和相对均方根误差(RRMSE)评估数据的准确性。

$$MAPE = \frac{1}{n} \times \sum_{i=1}^n \left| \frac{W_i - \bar{W}_i}{W_i} \right| \times 100\% \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (W_i - \bar{W}_i)^2} \quad (5)$$

$$RRMSE = \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n W_i} \quad (6)$$

式中: W_i 为月平均河宽数据的原位测量数据; \bar{W}_i 为MRWD估算的河流宽度。

表 1 最佳水体区域连接线方法伪代码(连接法)

Tab. 1 Pseudo-code of optimal surface water connection line method (connection method)

算法：连接法	
输入：水体栅格数据(water_data)，GSW 的最大历史水频率范围波段(max_extent)，全球水库和大坝数据库 (GRDD)，全球湖泊和湿地数据库(GLWD)	
输出：主河道连续水体栅格数据(connected_water_data)	
1.	构建 30 m 缓存区掩膜面(buffer_mask)并掩膜 water_data： <ul style="list-style-type: none">1.1 使用 max_extent 创建一个空的 buffer_mask1.2 将 buffer_mask 根据 max_extent 进行缓冲区分析，缓冲范围设置为 30 m1.3 将 buffer_mask 应用于 water_data
2.	去除湖泊和水库： <ul style="list-style-type: none">2.1 以 GRDD 和 GLWD 为参考数据，从 water_data 中去除湖泊和水库
3.	去除小面积水体，得到主河道水体栅格数据(main_water_data)： <ul style="list-style-type: none">3.1 计算 water_data 中每个水体区域的面积，将面积小于 900 m² 的河段从 water_data 中去除
4.	水体区域分组和编号： <ul style="list-style-type: none">4.1 创建一个空的区域编号图层(region_labels)4.2 使用 8 相邻像元数法评估 main_water_data 中水体栅格的连通性，为相互连接的水体分配唯一区域编号4.3 将区域编号数据添加到 region_labels 中
5.	水体区域最佳连接线构建 <ul style="list-style-type: none">5.1 创建一个空的连接线图层(connecting_lines)5.2 计算 region_labels 中各个区域之间的距离最短连接线5.3 将连接线添加到 connecting_lines 中
6.	赋值非水体区域 <ul style="list-style-type: none">6.1 创建一个空的主河道连续水体栅格数据(connected_water_data)6.2 复制 connecting_lines 和 region_labels 数据到 connected_water_data6.3 将 connecting_lines 经过的非水体区域在 connected_water_data 中标记为水体
7.	返回 connected_water_data 作为结果

4 结果和讨论

4.1 改进河宽提取算法的效果

本文对比两种不同方法（迭代法和连接法）处理得到的水体面积和 RivWidth 工具提取河宽的运算时间（表 2）。以 2019 年 7 月的黄河源区的水体栅格数据为例，相较于迭代运算，最佳水体连接后的 Rivwidth 运算效率提高了 24.7%，连接后的水体面积增加了 3.7%。其中，迭代法的统计结果是参考 Yang 等^[30]的算法描述通过代码复现得到的，表 2 中的运算时间为迭代法和连接法处理得到的水体数据输入到 RivWidth 工具后到处理得到河宽数据的时间。出现这种结果的原因是：① 连接法处理后的水体栅格区域为一个整体，只需进行一次运算，而迭代过程需要多次重复运算；② 由于 RivWidth 工具只能计算水体栅格数据中最大连续水体的河流宽度，以往基于水体数据使用 RivWidth 工具提取河宽的研究忽略了河道中的较小水体，如 Yang 等^[30]开发的 MCRW 河宽数据只提取了长度超过 1500 m 的河道的河流宽度，而连接法可以获取到不连续水体中的较小水体的河道数据（图 3），补全了水体信息，因此相较于迭代法，连接法处理后的水体面积有所增加。

4.2 反演的河宽数据的验证

河宽数据 MRWD 验证的首要步骤是确定验证日期。按照各个站点的区域范围，筛选

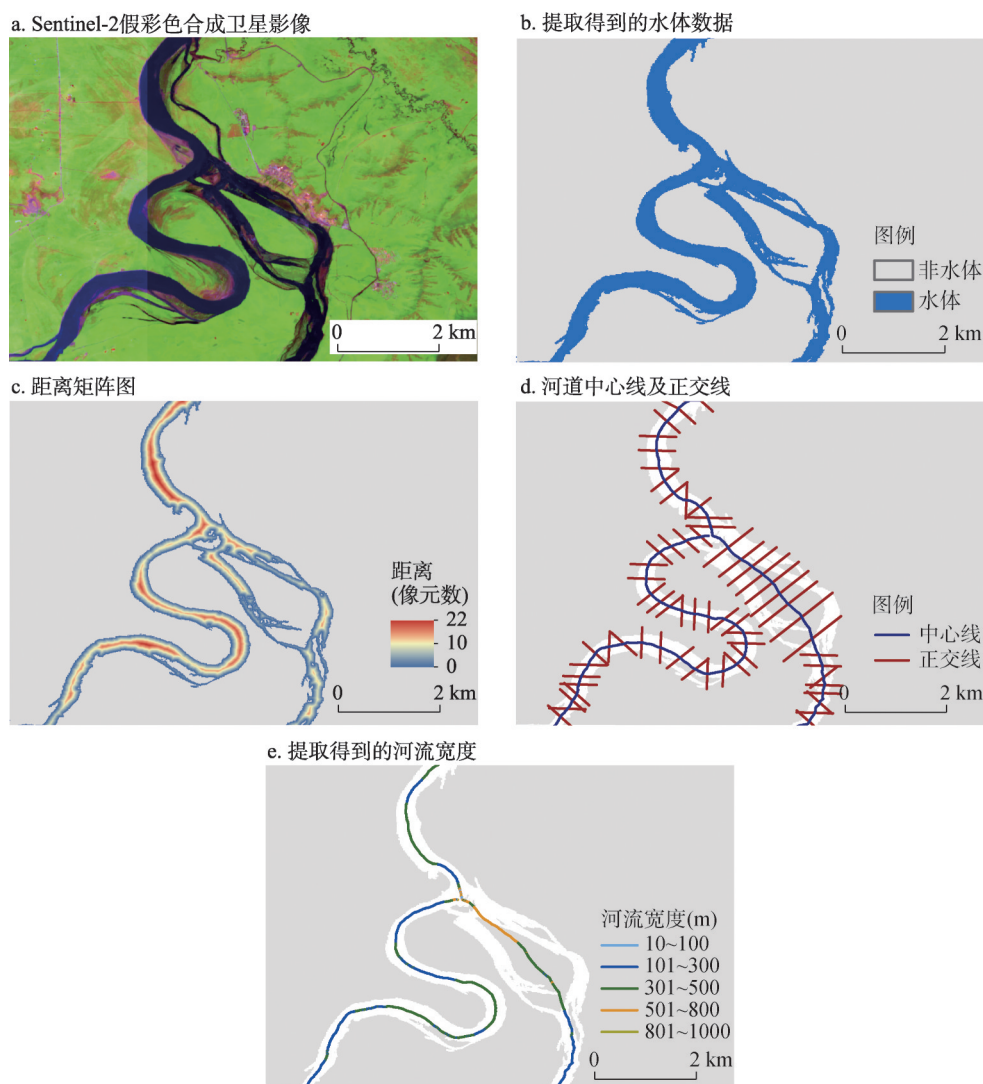


图4 河宽提取过程的输入、主要步骤和输出结果

Fig. 4 Input, main steps and output of river width extraction process

覆盖研究区各个水文站点区域的 Sentinel-2 卫星影像数据, 通过与水文年鉴上各个站点河流宽度数据的实际测定时间进行对比, 决定各个水文站点进行河流宽度日尺度验证的日期 (表3)。

使用原位测定河宽和 MRWD 数据进行验证 (表4), 分别统计河流宽度的平均绝对值百分比误差 (MAPE)、均方根误差 (RMSE) 以及相对均方根误差 (RRMSE)。此外, 将所有用于日尺度验证的推算河流宽度及其对应的原位测定河宽进行了对比 (图5)。由表4和图5可以看出, 河宽验证的 R^2 达到0.96, 最大 MAPE 约为15%, 说明 MRWD 可以较好的表征河流宽度的实际变化。

Allen 等^[9]使用了美国地质调查局 (USGS) 和加拿大水调查局 (WSC) 在北美运营的测量站的原位测量的河流宽度验证了 GRWL。Yang 等^[30]使用了中国水文站点的原位测定数据验证了 MCRW。由于 GRWL (平均宽度为 109.2 m, 中位数为 77.8 m)、MCRW

表 2 迭代法和连接法处理后的水体面积和 RivWidth 的运算时间

Tab. 2 The surface water area processed by the iterative method and the connection method and the corresponding running time of RivWidth tool

RivWidth优化方法	运算时间(s)	水体面积(km ²)	数据来源
迭代法	544.2	437.4	[30]
连接法	409.7	453.5	本文

表 3 黄河源区河流宽度数据验证日期

Tab. 3 The date of the river width dataset for validation in the SRYR

水文站点	纬度(°N)	经度(°E)	验证日期 (YYYYMMDD)
黄河沿	34.88	98.17	20201111
吉迈	33.77	99.65	20190103、20200225、20190418、20200726、20200825、20200904、20191030、20201116
门堂	33.77	101.05	20200502、20190811、20200929
玛曲	33.97	102.08	20200209、20200618、20190727、20200820、20190905、20191017、20201110、20201228
军功	34.70	100.65	20190513、20190612、20200820、20200929、20201118
唐乃亥	33.50	100.15	20190423、20190518、20200624、20190725、20200820、20200929、20191020、20191114、20201226
黄河	34.60	98.27	20190730
唐克	33.42	102.47	20190612、20200726、20200820、20190927、20201110

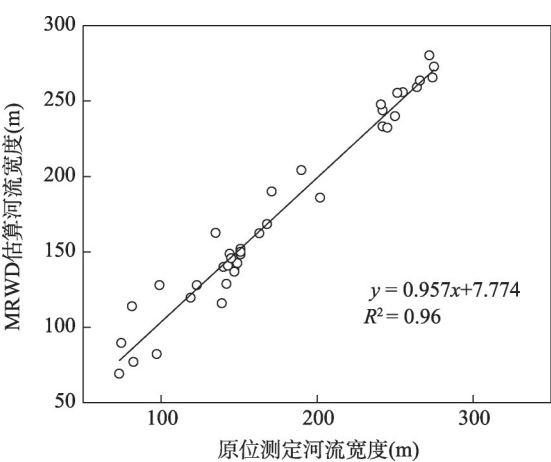
表 4 日尺度河流宽度的验证结果

Tab. 4 Validation results of daily river width

水文站点	黄河沿	吉迈	门堂	玛曲	军功	唐乃亥	黄河	唐克
MAPE(%)	15.52	10.37	12.32	1.54	8.02	3.66	6.55	4.61
RMSE(m)	15.10	13.82	17.58	4.92	16.61	8.74	5.40	10.81
RRMSE(%)	15.52	12.32	13.56	1.67	9.90	6.55	6.55	4.63

(平均宽度为 440.7 m，中位数为 272 m) 和 MRWD 在尺度上存在差异，因此，除了平均绝对值百分比误差 MAPE、和均方根误差 RMSE，本文引入相对均方根误差 (RRMSE)，即均方根误差除以平均实测数据得到的均方根误差归一化，以进行不同实测数据验证的河宽数据之间的比较。结果显示，MRWD 河流宽度数据的平均 MAPE、RMSE、RRMSE 分别为 7.82%、11.62 m、8.84%。优于 GRWL 的 18.9%、121.5 m、25.5%和 WF 为 1%的 MCRW 的 15.0%、73.7 m、15.2%。

尽管 GRWL、MCRW、MRWD 都是基于 RivWidth 方法估算的河流宽度数据，但是 GRWL 调查的是年平均流量下的河流宽度，MCRW 是多年多个月份条件下融



注：Pearson 相关系数为 0.98, $p < 0.05$ 。

图 5 MRWD 河宽和原位测定河宽对比验证
Fig. 5 Comparison and verification of the MRWD and in-situ river width

合的季节河流动态数据。而MRWD是在月度融合条件下的高时空分辨率数据,并在日尺度的条件下进行的验证。因此,MRWD河宽数据可以在更精细的时空尺度上表征黄河源区范围内河流宽度的变化。

4.3 河宽的分布特征

河流宽度数据MRWD由2019年、2020年和2021年逐月的河流宽度数据总计 6.1×10^6 个平面测量值组成,涵盖了黄河源区2035 km的河流。MRWD数据中的河流宽度为10 m(即Sentinel-2图像的最小像素大小)到3202 m(在上游鄂陵湖出水口附近,图7),河宽数据MRWD中的每个测量值由河流中心线像素的经纬度坐标及其对应的河流宽度数据组成。在各个月份的河流宽度数据中,2%的河流宽度超过1000 m,85%以上的河流宽度小于500 m,大约有41%的河流宽度小于100 m。研究发现,MRWD河流宽度的分布(图6)非常接近幂律函数(100~500 m河宽, $R^2 = 0.80$, $p < 0.05$):

$$N_{MRWD} = 3.55 \times 10^5 \times W_{MRWD}^{-0.49} \quad (7)$$

式中: W_{MRWD} 为各个像素的河流宽度; N_{MRWD} 为对应河流宽度范围内的像素数。继Allen等^[28]和Yang等^[30]之后,本文计算了GRWL河流宽度的幂律函数(100~500 m河宽, $R^2 = 0.68$, $p < 0.05$):

$$N_{GRWL} = 1.97 \times 10^5 \times W_{GRWL}^{-0.52} \quad (8)$$

式中: W_{GRWL} 为GRWL河流宽度数据中各个像素的河流宽度; N_{GRWL} 为对应河流宽度范围内的像素数。

图6中可以看出,MRWD和GRWL的拟合曲线大致相似,由于MRWD河流宽度数据的空间分辨率为10 m,而GRWL河流宽度数据的空间分辨率为30 m,因此本文MRWD河流宽度测量值高于GRWL河流宽度数据。此外,GRWL得到的宽度分布指数(0.52)大于MRWD(0.49),表明MRWD河流宽度数据具有更高比例的较宽河流^[30]。其原因是:① MRWD具有更高的时空分辨率,对于河流宽度的刻画更加精细,同时求算的是时段内的最大河宽,而GRWL的空间分辨率仅为30 m,求算的是平均流量下的河流宽度;② 两个数据的水文条件不同。MRWD数据是2019—2021年黄河源区的首个10 m空间分辨率的逐月河流宽度测量数据,提供了关于河流中心线和宽度的栅格和矢量格式的历史信息。栅格文件包含具有相应河流宽度值的中心线像素,而矢量文件由连接每个中心线像素与相邻像素的折线组成。

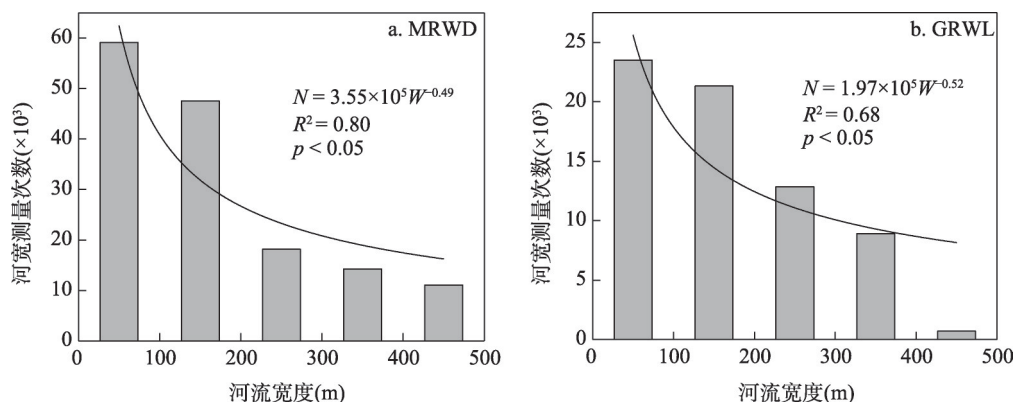


图6 MRWD和GRWL数据0~500 m的河流宽度分布

Fig. 6 Distribution of the river width from 0 to 500 m of MRWD and GRWL

4.4 河宽的时空变化动态

基于MRWD河宽数据（图7），本文提取了黄河源区干流并计算各个河段距离上游河口的累计长度，采用LOWESS方法进行河宽的平滑处理，得到河宽沿干流距离增加的变化（图8a~8c）。研究区的有效河宽在0~1100 m范围内动态变化，从上游到下游，河道高程逐渐降低，河流宽度整体呈现出波动性下降趋势，河源位于扎陵湖和鄂陵湖区附近，河流出水口在唐乃亥水文站点附近。此外，在河道的中后段（400~1250 km），夏季7—9月的河流宽度显著高于其他月份的河宽，月平均河流宽度分别为255 m、257 m、249 m，冬季12—2月的河宽最小，月平均河流宽度分别为152 m、149 m、141 m（图8d）。MRWD中小于100 m的河道主要分布于吉迈—门堂水文站河段（550~700 km）和军功—唐乃亥水文站河段（1200~1450

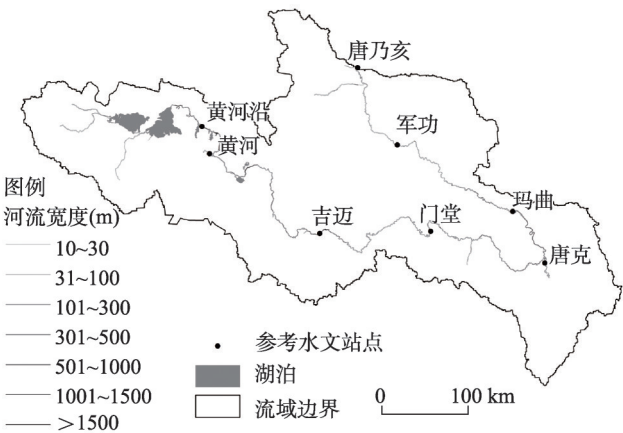


图7 2019年7月黄河源区河流宽度分布
Fig. 7 Distribution of the river width of the SRYR in July 2019

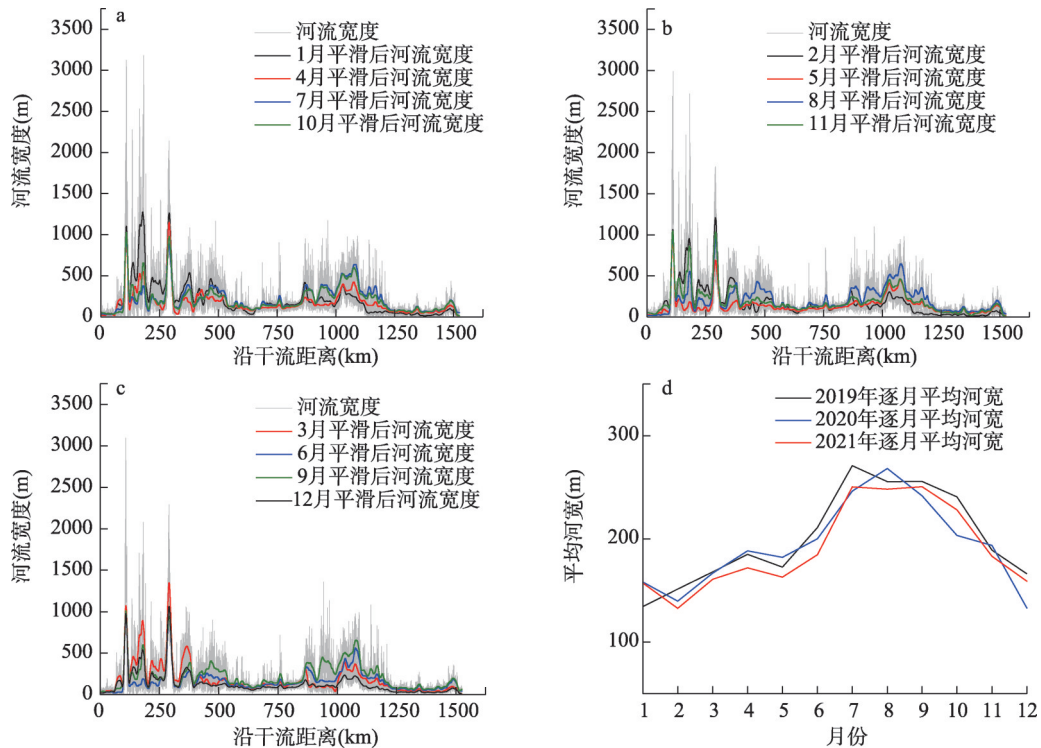


图8 黄河源区干流沿程的原始和平滑河宽及2019—2021年黄河源区逐月平均河宽
Fig. 8 Original and smoothed river width along the main stream of the SRYR and monthly averaged river width from 2019 to 2021 of the SRYR

km)。河道前段(0~400 km)分布有扎陵湖、鄂陵湖两个大型湖泊和众多小型湖泊,唐克水文站附近的黄河九曲第一弯到玛曲水文站附近的河段(1000~1100 km)分布有较多连通的大型辫状河道。

4.5 局限性和不确定性

本文使用 Sentinel-2 卫星影像数据提取了黄河源区的河流宽度,由于源数据空间分辨率的限制,黄河源区部分狭窄河道的河流宽度可能无法提取或者存在一定的误差,例如黄河源区河口区域。随着高空间分辨率遥感图像(如 CubeSat 和 Planet)和 DEM 数据(如 WorldDEM)的发布,未来可以考虑更高精度河宽的提取和监测研究。

本文中获得的河流宽度数据仅为3年的逐月河流宽度数据^[52],无法支撑河流宽度的年际变化规律分析和长期动态演变研究。因此,探讨黄河源区河流宽度的长历时动态变化模式是一个潜在的未来研究课题。

5 结论

本文基于 Sentinel-2 卫星数据和 GEE 大数据云计算平台,结合 RivWidth 方法提取得到了黄河源区 2019—2021 年逐月水体数据,在此基础上提取并分析了黄河源区河宽的空间分布特征和年内动态变化,结果表明:

(1) 连接法通过构建最佳水体区域连接线的方式提高了 RivWidth 的运算效率,补充了河道信息。与迭代法相比,构建最佳水体连接线后的 RivWidth 运算时间缩短了 24.7%,连接后的河道信息增加了 3.7%。

(2) 黄河源区河宽的空间变化差异大,提取得到的河流宽度数据 MRWD 服从幂律函数分布(100 m 到 500 m 河宽)。河流宽度在 10 m (Sentinel-2 空间分辨率)到 3202 m (上游鄂陵湖出水口附近)之间变化,黄河源区范围内大约有 41% 的河道窄于 100 m, 85% 以上的河道宽度小于 500 m, 仅 2% 的河道宽度超过 1000 m。

(3) 黄河源区的河宽年内呈现先增加后降低的趋势,夏季(7—8月)和秋季(9月)的平均河宽最大,分别为 255 m、257 m、249 m,冬季(12—2月)的平均河宽最小,分别为 152 m、149 m、141 m。

参考文献(References)

- [1] Sun W C, Fan J, Wang G Q, et al. Calibrating a hydrological model in a regional river of the Qinghai-Tibet Plateau using river water width determined from high spatial resolution satellite images. *Remote Sensing of Environment*, 2018, 214: 100-114.
- [2] Xu W T, Shen Q, Wang X L, et al. Sensing images for assessing the minimum ecological flux by automatically extracting river surface width. *Remote Sensing*, 2020, 12(18): 2899. DOI: 10.3390/rs12182899.
- [3] Lopez Dubon S, Lanzoni S. Meandering evolution and width variations: A physics-statistics-based modeling approach. *Water Resources Research*, 2019, 55(1): 76-94.
- [4] Paillex A, Schuwirth N, Lorenz A W, et al. Integrating and extending ecological river assessment: Concept and test with two restoration projects. *Ecological Indicators*, 2017, 72: 131-141.
- [5] Possa E M, Maillard P, de Oliveira L M. Discharge estimation for medium-sized river using multi-temporal remote sensing data: A case study in Brazil. *Hydrological Sciences Journal*, 2020, 65(14): 2402-2418.
- [6] Sun W C, Ishidaira H, Bastola S, et al. Estimating daily time series of streamflow using hydrological model calibrated based on satellite observations of river water surface width: Toward real world applications. *Environmental Research*, 2015, 139: 36-45.
- [7] Bates P D, De Roo A P J. A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, 2000, 236 (1-2): 54-77.

- [8] Yamazaki D, Lee H, Alsdorf D E, et al. Analysis of the water level dynamics simulated by a global river model: A case study in the Amazon River. *Water Resources Research*, 2012, 48(9). DOI: 10.1029/2012WR011869.
- [9] Allen G H, Pavelsky T M. Global extent of rivers and streams. *Science*, 2018, 361(6402): 585-588.
- [10] Krapesch G, Hauer C, Habersack H. Scale orientated analysis of river width changes due to extreme flood hazards. *Natural Hazards and Earth System Sciences*, 2011, 11(8): 2137-2147.
- [11] Alsdorf D, Lettenmaier D, Vörösmarty C. The need for global, satellite-based observations of terrestrial surface waters. *Eos, Transactions American Geophysical Union*, 2003, 84(29): 269-276.
- [12] Gleason C J, Smith L C. Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. *PNAS*, 2014, 111(13): 4788-4791.
- [13] Pavelsky T M, Durand M T, Andreadis K M, et al. Assessing the potential global extent of SWOT river discharge observations. *Journal of Hydrology*, 2014, 519: 1516-1525.
- [14] Hrachowitz M, Savenije H H G, Blöschl G, et al. A decade of Predictions in Ungauged Basins (PUB): A review. *Hydrological Sciences Journal*, 2013, 58(6): 1198-1255.
- [15] Hannah D M, Demuth S, van Lanen H A J, et al. Large-scale river flow archives: Importance, current status and future needs. *Hydrological Processes*, 2011, 25(7): 1191-1200.
- [16] Zolezzi G, Luchi R, Tubino M. Modeling morphodynamic processes in meandering rivers with spatial width variations. *Reviews of Geophysics*, 2012, 50(4). DOI: 10.1029/2012RG000392.
- [17] Biancamaria S, Bates P D, Boone A, et al. Large-scale coupled hydrologic and hydraulic modelling of the Ob River in Siberia. *Journal of Hydrology*, 2009, 379(1/2): 136-150.
- [18] Horn J D, Joeckel R M, Fielding C R. Progressive abandonment and planform changes of the central Platte River in Nebraska, central USA, over historical timeframes. *Geomorphology*, 2012, 139: 372-383.
- [19] Xu Jiahui, Wang Shidong, Song Lijuan, et al. Mapping spatio-temporal variation of river width from satellite remote sensing data and hydrometeorological response in the Yarlung Zangbo River. *Acta Geographica Sinica*, 2022, 77(11): 2862-2877. [徐嘉慧, 王世东, 宋利娟, 等. 雅鲁藏布江干流河宽时空变化遥感监测及水文气象响应. *地理学报*, 2022, 77(11): 2862-2877.]
- [20] Zhao Changsen, Pan Xu, Yang Shengtian, et al. Measuring streamflow with low-altitude UAV imagery. *Acta Geographica Sinica*, 2019, 74(7): 1392-1408. [赵长森, 潘旭, 杨胜天, 等. 低空遥感无人机影像反演河道流量. *地理学报*, 2019, 74(7): 1392-1408.]
- [21] Werbylo K L, Farnsworth J M, Baasch D M, et al. Investigating the accuracy of photointerpreted unvegetated channel widths in a braided river system: A Platte River case study. *Geomorphology*, 2017, 278: 163-170.
- [22] Coe M T, Costa M H, Howard E A. Simulating the surface waters of the Amazon River basin: Impacts of new river geomorphic and flow parameterizations. *Hydrological Processes: An International Journal*, 2008, 22(14): 2542-2553.
- [23] De Paiva R C D, Buarque D C, Collischonn W, et al. Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research*, 2013, 49(3): 1226-1243.
- [24] Oki T, Agata Y, Kanae S, et al. Global assessment of current water resources using total runoff integrating pathways. *Hydrological Sciences Journal*, 2001, 46(6): 983-995.
- [25] Hou J W, van Dijk A I J M, Beck H E. Global satellite-based river gauging and the influence of river morphology on its application. *Remote Sensing of Environment*, 2020, 239: 111629. DOI: 10.1016/j.rse.2019.111629.
- [26] Yamazaki D, Kanae S, Kim H, et al. A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, 2011, 47(4). DOI: 10.1029/2010WR009726.
- [27] Yamazaki D, O'Loughlin F, Trigg M A, et al. Development of the global width database for large rivers. *Water Resources Research*, 2014, 50(4): 3467-3480.
- [28] Allen G H, Pavelsky T M. Patterns of river width and surface area revealed by the satellite-derived North American River width data set. *Geophysical Research Letters*, 2015, 42(2): 395-402.
- [29] Isikdogan F, Bovik A, Passalacqua P. RivaMap: An automated river analysis and mapping engine. *Remote Sensing of Environment*, 2017, 202: 88-97.
- [30] Yang J, Huang X, Tang Q H. Satellite-derived river width and its spatiotemporal patterns in China during 1990-2015. *Remote Sensing of Environment*, 2020, 247: 111918. DOI: 10.1016/j.rse.2020.111918.
- [31] Wang Z F, Li J B, Lin Y, et al. GrabRiver: Graph-theory-based river width extraction from remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters*, 2020, 19: 1-5.
- [32] Jin R, Li X. Improving the estimation of hydrothermal state variables in the active layer of frozen ground by

- assimilating in situ observations and SSM/I data. *Science in China Series D: Earth Sciences*, 2009, 52(11): 1732-1745.
- [33] Shi H Y, Li T J, Wei J H, et al. Spatial and temporal characteristics of precipitation over the Three-River Headwaters region during 1961-2014. *Journal of Hydrology: Regional Studies*, 2016, 6: 52-65.
- [34] Qin Y, Yang D W, Gao B, et al. Impacts of climate warming on the frozen ground and eco-hydrology in the Yellow River source region, China. *Science of the Total Environment*, 2017, 605: 830-841.
- [35] Farr T G, Rosen P A, Caro E, et al. The shuttle radar topography mission. *Reviews of Geophysics*, 2007, 45(2). DOI: 10.1029/2005RG000183.
- [36] Pekel J F, Cottam A, Gorelick N, et al. High-resolution mapping of global surface water and its long-term changes. *Nature*, 2016, 540(7633): 418-422.
- [37] Lehner B, Liermann C R, Revenga C, et al. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, 2011, 9(9): 494-502.
- [38] Lehner B, Döll P. Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 2004, 296(1-4): 1-22.
- [39] Zou Z H, Dong J W, Menarguez M A, et al. Continued decrease of open surface water body area in Oklahoma during 1984-2015. *Science of the Total Environment*, 2017, 595: 451-460.
- [40] Chen B Q, Xiao X M, Li X P, et al. A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017, 131: 104-120.
- [41] Zou Z H, Xiao X M, Dong J W, et al. Divergent trends of open-surface water body area in the contiguous United States from 1984 to 2016. *PNAS*, 2018, 115(15): 3810-3815.
- [42] Xia H, Zhao J, Qin Y, et al. Changes in water surface area during 1989-2017 in the Huai River basin using Landsat data and Google earth engine. *Remote Sensing*, 2019, 11(15): 1824. DOI: 10.3390/rs11151824.
- [43] Wang X X, Xiao X M, Zou Z H, et al. Gainers and losers of surface and terrestrial water resources in China during 1989-2016. *Nature Communications*, 2020, 11(1): 3471. DOI: 10.1038/s41467-020-17103-w.
- [44] Basnayaka V, Samarasinghe J T, Gunathilake M B, et al. Planform changes in the lower Mahaweli River, Sri Lanka using landsat satellite data. *Land*, 2022, 11(10): 1716. DOI: 10.3390/land11101716.
- [45] Zhou H W, Liu S X, Hu S, et al. Retrieving dynamics of the surface water extent in the upper reach of Yellow River. *Science of the Total Environment*, 2021, 800: 149348. DOI: 10.1016/j.scitotenv.2021.149348.
- [46] Chen J H, Kang T T, Yang S, et al. Open-surface water bodies dynamics analysis in the Tarim River basin (North-Western China), based on Google Earth Engine cloud platform. *Water*, 2020, 12(10): 2822. DOI: 10.3390/w12102822.
- [47] Deng Y, Jiang W G, Tang Z, et al. Long-term changes of open-surface water bodies in the Yangtze River basin based on the Google Earth Engine cloud platform. *Remote Sensing*, 2019, 11(19): 2213. DOI: 10.3390/rs11192213.
- [48] Pavelsky T M, Smith L C. RivWidth: A software tool for the calculation of river widths from remotely sensed imagery. *IEEE Geoscience and Remote Sensing Letters*, 2008, 5(1): 70-73.
- [49] Miller Z F, Pavelsky T M, Allen G H. Quantifying river form variations in the Mississippi Basin using remotely sensed imagery. *Hydrology and Earth System Sciences*, 2014, 18(12): 4883-4895.
- [50] Rosenfeld A, Pfaltz J L. Distance functions on digital pictures. *Pattern Recognition*, 1968, 1(1): 33-61.
- [51] Marr D, Hildreth E. Theory of edge detection. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 1980, 207(1167): 187-217.
- [52] Drusch M, Del Bello U, Carlier S, et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 2012, 120: 25-36.

An efficient river width extraction algorithm for analyzing the annual variation of the river width in the upper reaches of the Yellow River

XIE Chaoshuai^{1,2,3}, LIU Suxia^{1,2,3}, ZHOU Haowei^{1,3}, MO Xingguo^{1,2,3}

(1. Key Laboratory of Land Water Cycle and Surface Processes, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China; 2. Sino-Danish College, University of Chinese Academy of Sciences, Beijing 100049, China; 3. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: River width is one of the key parameters for studying river dynamic processes and hydrological and hydrodynamic models. Compared with traditional hydrological stations, satellite images can achieve large-scale observations of river width. Based on Sentinel-2 imagery, Google Earth Engine (GEE) and RivWidth tool, a new river connectivity construction method was developed to improve the computational efficiency of RivWidth tool for river width extraction, and the monthly river width dataset (MRWD) with a 10-m spatial resolution from 2019 to 2021 in the upper reaches of the Yellow River was generated by the improved algorithm. The results show that the river connectivity construction method developed in this study (referred to as the connection method) improves the computing efficiency of RivWidth tool by 24.7%. The comparison and verification results with the in-situ river width and GRWL (Global River Widths from Landsat) river width show that the mean absolute percentage error (MAPE), root mean square error (RMSE) and relative root mean square error (RRMSE) of MRWD are respectively 7.82%, 11.62 m, and 8.84%. The MRWD includes 6.1×10^6 river width estimations, covering 2035 km of rivers in the upper reaches of the Yellow River, of which 41% are narrower than 100 m, more than 85% are less than 500 m, and 2% are more than 1000 m. The MRWD generated in this study can provide an important reference for surface water resources management.

Keywords: Sentinel-2; GEE; the upper reaches of Yellow River; river width; RivWidth