Annual river dataset in China: a new product with a 10 m spatial resolution from 2016 to 2023

Kaifeng Peng[1,2], Beibei Si[1], Weiguo Jiang[2], Meihong Ma[1], Xuejun Wang[1,3,*]

[1]Faculty of Geography, Tianjin Normal University, Tianjin, 300387, China

5   [2]State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University，Beijing, 100875, China

[3] Academy of Eco-civilization Development for Jing-Jin-Ji Megalopolis, Tianjin Normal University, Tianjin, 300387, China

Correspondence to: Xuejun Wang (xuejunw@tjnu.edu.cn)

Abstract: Rivers play important roles in ecological biodiversity, shipping trade and the carbon cycle. Owing to human

10   disturbances and extreme climates in recent decades, river extents have altered frequently and dramatically. The development of sequential and fine-scale river extent datasets, which could offer strong data support for river protection, management and sustainable use, is urgently needed. A literature review revealed that annual river extent datasets with fine spatial resolutions are generally unavailable for China. To address this issue, the first Sentinel-derived annual China river extent dataset (CRED) from 2016 to 2023 was produced in our study. We first produced annual water maps by combining the dynamic world (DW),

15   ESRI global land cover (EGLC) data and the multiple index water detection rule (MIWDR). For the DW and MIWDR water time series, the mode algorithm, which calculates the most common values, was used to generate yearly water maps. Then, an object-based hierarchical decision tree based on geometric features and auxiliary datasets was developed to extract rivers from the water data. The results indicated that the overall accuracies (OAs) of the CRED were greater than 96.0% % from 2016 to 2023. The user accuracies (UAs), producer accuracies (PAs) and F1 scores of the rivers exceeded 95.3%, 91.3% and 93.7%,

20   respectively. A further data intercomparison indicated that our CRED shared similar patterns with the wetland map of East Asia (EA_Wetlands), China land use/cover change (CNLUCC) and China water covers (CWaC) datasets, with correlation coefficients (R) greater than 0.75. Moreover, our CRED outperformed the three datasets in terms of small river mapping and misclassification reduction. The area statistics indicated that the river area in China was 44,948.78 km^2 in 2023, which was mostly distributed in coastal provinces of China. From 2016 to 2023, the river areas were characterized by an initial increase,

25   followed by a decrease and then a slight increase. Spatially, the decreased rivers were located mainly in Southeast China, whereas the increased rivers were distributed mainly in Central China and Northeast China. In general, the CRED explicitly delineated river extents and dynamics in China, which could provide a good foundation for improving river ecology and management. The CRED dataset is publicly available at https://doi.org/10.5281/zenodo.13841910 (Peng et al., 2024a).

1 Introduction

30   River ecosystems are fluid water bodies with linear geometric characteristics. Known as the blood vessels of the Earth, rivers support large amounts of fresh water in natural ecosystems and human societies (Wei et al., 2020) and exert important effects on hydrological cycles and ecological evolution (Li et al., 2020). Moreover, rivers also affect the carbon trade-off in terrestrial ecosystems. As reported, global rivers emit approximately 1.8 Pg of carbon (i.e., carbon dioxide) into the atmosphere annually (Raymond et al., 2013), and the rates of these processes are importantly related to river area changes (Caissie, 2006;

35   Hotchkiss et al., 2015). Thus, investigating river extent changes over time, which could strongly support human well-being, dual carbon targets and hydrological modes, is necessary.

Owing to its large imaging range and low acquisition cost, satellite technology has become an important tool for extracting rivers (Valman et al., 2024). To date, a series of river datasets have been produced at the global or national scale, which can be roughly divided into two categories: river line datasets and river polygon datasets. The river line datasets delineate river channel

40   lines with attributes of width, discharge and others, such as the GRRATS (Global River Radar Altimetry Time Series) (Coss et al., 2020), HydroRIVERS (Lehner and Grill, 2013), and MCRW (Multitemporal China River Width) (Yang et al., 2020a). These datasets were produced on the basis of digital elevation model (DEM) data or satellite images, which effectively display the spatial distribution of the river network. However, they cannot delineate river spatial extents and temporal area changes (Fei et al., 2022). In contrast, river polygon datasets record river complete extents as polygons. In these datasets, rivers together

45   with other wetland types are generally mapped on the basis of satellite images or data compilation, such as the GLWD (Global Lakes and Wetlands Database) (Lehner and Doll, 2004), GRWL (Global River Widths from Landsat) (Allen and Pavelsky, 2018), and CAS_Wetlands (Mao et al., 2020). These datasets effectively display river extents and their spatial distributions. However, a literature review revealed that multitemporal river polygon datasets at the global or national scale are still scarce, as are those for China.

50   Compared with limited river datasets, water-related products are largely produced at global and national scales (Xu et al., 2020). These datasets provide good potential for river classification because rivers are contained within water areas. Generally, water-related datasets are roughly divided into two categories: surface water datasets and land use/cover change (LUCC) datasets (Xu et al., 2022). The former directly map water extents on the basis of time series of satellite images, such as the GSW (Global Surface Water) (Pekel et al., 2016) and GSWD (Global Surface Water Dynamics) (Pickens et al., 2020) datasets.

55   LUCC datasets, such as NLCD (Wickham et al., 2023), GLC_FCS30D (Zhang et al., 2024), and dynamic world (DW) (Brown et al., 2022) datasets, mapped waters along with other land use types. These above datasets have good potential for large-scale river extraction.

China has considerable river resources with world-famous rivers, such as the Yangtze River, Yellow River and Pearl River. According to China's First National Census for water, there are approximately 22,200 rivers with basin areas greater

60   than 100 km^2 (Ministry of Water Resources, 2013). However, China is one of the most water-deficient countries because of its

low per capita water resources (Bai and Zhao, 2023). Moreover, human disturbance and climate change have exacerbated river

losses in recent years (Speed et al., 2016). To balance river protection and economic development, China's government has

implemented a series of projects, such as the Yangtze River Economic Belt (She et al., 2019) and the high-quality development

of the Yellow River (Jiang et al., 2021). Considering the ecological and economic importance of rivers, it is necessary to

65   produce multitemporal river datasets for China.

The objective of this study was to produce annual China river extent datasets (CRED), which, to the best of our knowledge,

are the first river datasets at a 10-m spatial resolution from 2016 to 2023 for China. To achieve this goal, we first collected two

LUCC datasets and Sentinel-2 time series to generate a yearly water map. Second, annual river extents were mapped via

geometric features and knowledge rules. The geometric features were calculated via chessboard segmentation, and the

70   knowledge rules were developed via sample analysis of rivers, lakes and reservoirs. Finally, test samples, created via visual

interpretation, were used to assess the river mapping accuracy. Moreover, comparisons with existing river-related datasets were

implemented to confirm the reliability of our river dataset.

2 Study area and data

2.1 Study area

75   In this study, China was chosen as the study area, with longitudes of 73°30'4" E and 135°5'19" E and latitudes of 6°19'24"

N and 53°33'39" N. China covers five major climate zones (e.g., tropical monsoon, subtropical monsoon, temperate monsoon,

temperate continental, and plateau mountain zones), with a temperature difference of ~30 °C and a precipitation difference of

~2,500 mm (Liu et al., 2014). In China, the distribution of rivers exhibits regional imbalance, with dense networks in the

southern and eastern regions and sparse networks in the northern and western regions (Bai and Zhao, 2023). To analyse river

80   characteristics in detail, we divided China into nine river basins (Fig. 1), namely, the Continental Basin (CB), Southeast Basin

(SEB), Huaihe River Basin (HuRb), Haihe River Basin (HaRB), Yellow River Basin (YeRB), Songhua and Liaohe River Basin

(SLRB), Southwest Basin (SWB), Yangtze River Basin (YaRB), and Pearl River Basin (PRB).
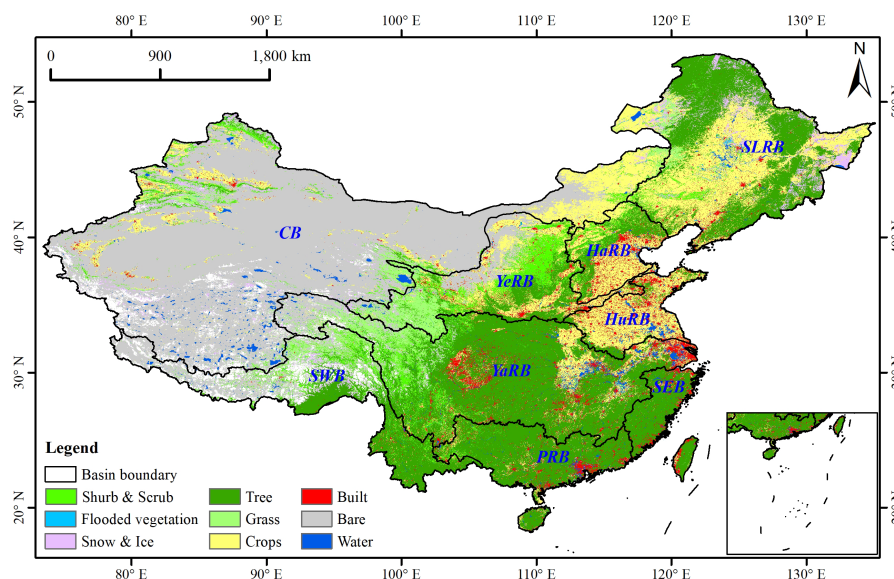
Fig. 1. Land use patterns and river basins of China. The coloured map is a mode-composited image of DW time series images from 2023.

2.2 Data collection

The dynamic world (DW), a near-real-time (NRT) land use/land cover (LULC) dataset, was selected as the primary dataset for producing water maps (Brown et al., 2022). It was produced on the basis of a Sentinel-2 time series, with a revisit interval of 2–5 days and a spatial resolution of 10 m. In the DW dataset, nine land use types were delineated, and the water areas had high mapping accuracy. Statistical analysis revealed that many regions in China had no or few good observations (Fig. S1). This occurred because the DW images were produced from Sentinel-2 images with less than 35% cloud coverage.

For areas where DW observations were missing, we used the ESRI global land cover (EGLC) dataset to supplement the data (Karra et al., 2021). The EGLC dataset was produced on the basis of Sentinel-2 images via deep learning algorithms. It recorded annual global land use types from 2017 to 2023 with high accuracy for water bodies. In addition, we also collected Sentinel-2 images to produce water maps in areas missing data from the DW dataset in 2016 because of the lack of 10-m LUCC or water datasets. All Sentinel-2 images from 2015 to 2016 were used for water classification.

To facilitate river extraction and data intercomparison, some auxiliary datasets, which are listed in Table 1, were used. The China land use/cover change (CNLUCC) dataset is a LUCC dataset that maps 23 land use types at a 30 m spatial resolution (Liu et al., 2018b). It was produced by visual interpretation, with an overall accuracy exceeding 90%. The CNLUCC of 2020 was used to produce training samples of rivers, lakes and reservoirs and make comparisons with our river maps. The Wetland map of East Asia (EA_wetlands) is a detailed wetland dataset (including rivers) of East Asia from 2021, with an overall accuracy of 88% Wang et al. (2023a). The China water cover map (CWaC) is a Chinese water cover dataset (including rivers) from 2020 with a 10 m spatial resolution and an overall accuracy of 86% (Li and Niu, 2022). EA_Wetlands and CWaC were

also used to compare the data to validate the reliability of our river maps. In addition, the dam points, including the reservoir

105 point of interest (POI), OSM_DEM, and GOODD (Mulligan et al., 2020), were used to distinguish reservoirs from natural

water bodies.

Table 1. Study data and information in our study

Type	Name	Spatial Resolution	Time period	Data description	Data source
Land use	DW	10 m	2015–2023	Water extraction	Brown et al. (2022)
	EGLC	10 m	2017–2023		Karra et al. (2021).
	CNLUCC	30 m	2020	Data comparison Sample produce	Liu et al. (2018b)
Satellite image	Sentinel-2	10 m	2015–2016	Water extraction	Google Earth Engine (GEE)
Wetland dataset	EA_Wetlands	10 m	2021	data comparison	Wang et al. (2023a)
	CWaC	10 m	2020		Li and Niu (2022)
Auxiliary dataset	POI	---	2023	Reservoir extraction	https://lbs.amap.com
	GOODD	---	---		Mulligan et al. (2020)
	OSM_Dam	---	---		

3 Methodology

This study aimed to produce a sequential and fine-resolution river extent dataset for China, and the processing chain

110 included yearly water map generation, river classification, and accuracy validation (Fig. 2). First, the mode algorithm was used

to composite yearly water map-based DW datasets. For areas with missing DW data, the EGLC and Sentinel-2 images were

chosen as supplementary datasets, which were utilized to create annual water maps. Second, an object-based hierarchical

decision tree algorithm was developed on the basis of geometric features and auxiliary datasets that can extract rivers well

from water polygons. Finally, test samples were manually produced to assess the classification accuracy. Additionally, river-

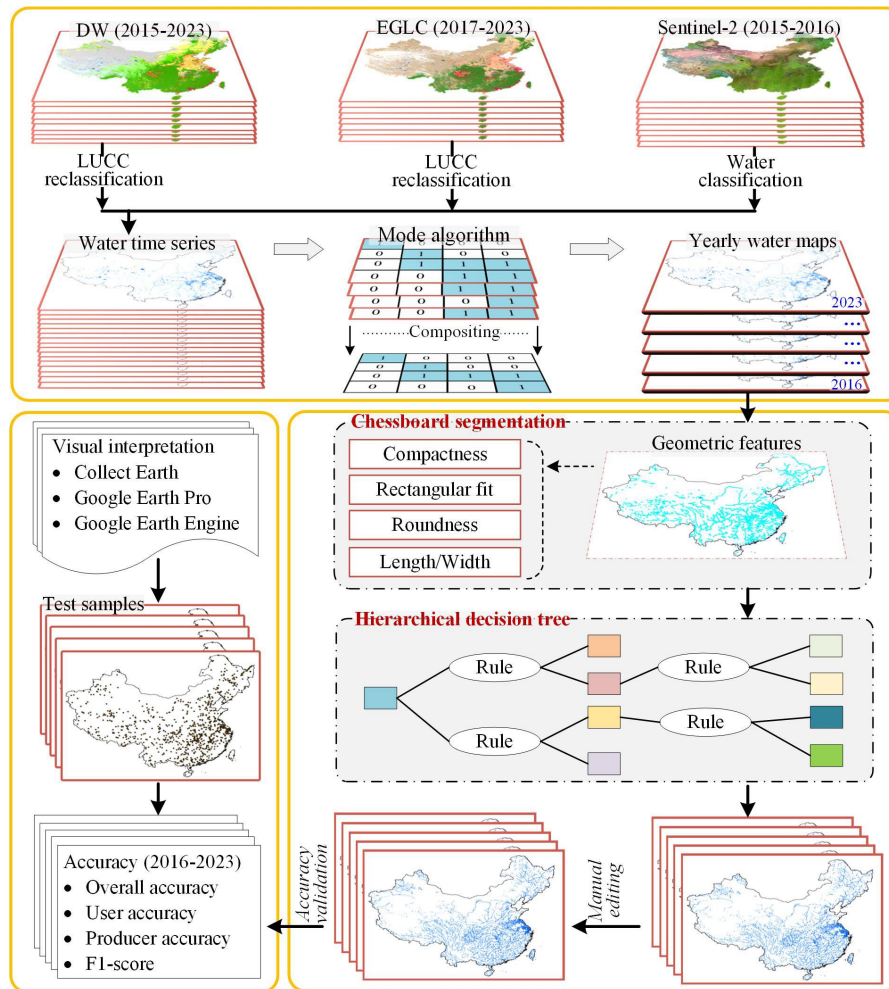115 related datasets were also used to compare data to test the reliability of our river maps.

Fig. 2. Workflow of annual river extraction

3.1 Annual water generation

Rivers are distributed within the extent of surface water. Thus, water extraction can offer a good data foundation for river

120   mapping. To facilitate computation, we divided China into 52 regions using 5°×5° tiles. Through the statistics of valid DW observations from 2016 to 2023, we found that many areas, which are mostly distributed in Southwest China, have a small number or no valid land use pixels (Fig. S1). To address this issue, three scenarios were designed for water extraction (Fig. 3). Specifically, if all the valid DW pixels in a tile exceeded three in a year, the DW dataset was selected to produce annual water maps. Otherwise, the EGLC dataset was used to create annual maps from 2017-2023, or the Sentinel-2 images from 2015 to

125   2016 were used to generate water maps for 2016. In addition, the DW dataset for tiles 49, 51 and 52 had poor water accuracies. For the three tiles, in our study, annual water maps from 2017-2023 were created via the EGLC dataset, whereas those for 2016 were produced via Sentinel-2 images.
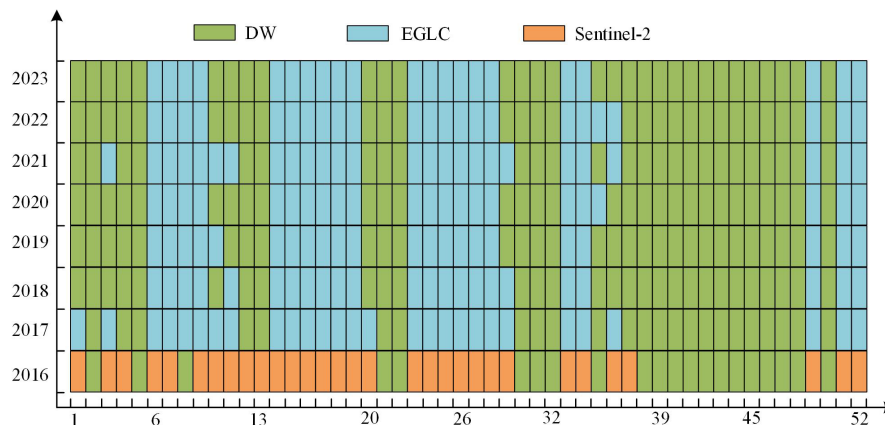
Fig. 3. Three scenarios for water extraction from different datasets

3.1.1 Water extraction based on the DW dataset

The DW dataset provides mapped water areas along with eight additional land cover types from 2015 to the present. Its high-frequency revision offers a good foundation for generating annual water maps. Considering that water extent changes over time, a single-temporal DW image cannot adequately represent the annual extent of rivers. Thus, all DW time series images within one year were used to produce annual water maps. As reported by Venter et al. (2022), the mode algorithm can generate the most common land cover for each pixel during the year, which was adopted in this study. The specific principle is illustrated in Fig. 4.
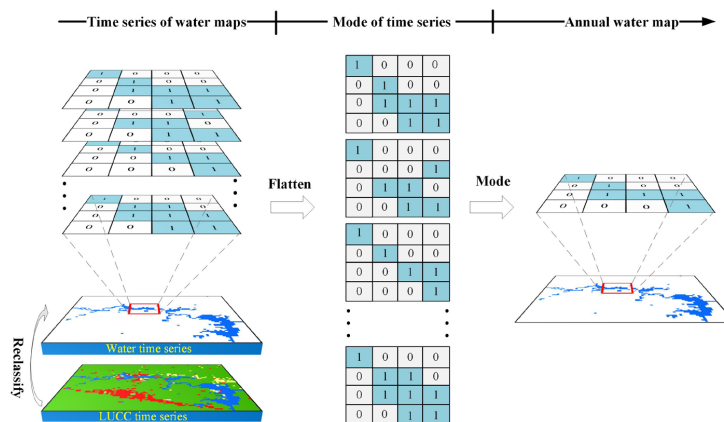

Fig. 4. Illustration of annual water generation

First, water images were generated by reclassifying the DW images. For an individual pixel of a water image, 1 represents water, and 0 represents non-water. Then, water time series images were generated. Finally, the mode algorithm was used to aggregate the water time series images within one year into annual water maps. For example, if a pixel has time series values of (1, 0, 0, 1, 1), then the aggregated value of this pixel is 1. If a pixel has time series values of (1, 0, 0, 0, 1), then the aggregated

7

value of this pixel is 0. In this way, annual water maps were created via the mode algorithm. Additionally, considering the scarcity of DW images before 2016, yearly water maps for 2016 were produced via DW time series images from 2015-2016.

145 The above processing was implemented in the green tiles of Fig. 3.

3.1.2 Water extraction based on the EGLC dataset

For the data gaps in the DW images from 2017 to 2023, we used the EGLC dataset as a supplemental dataset. The EGLC dataset is an annual-scale global land use dataset that records water and eight other land use types. It is produced annually on the basis of Sentinel-2 time series images from 2017 to 2023. In this study, yearly water maps were directly produced via

150 reclassification of the EGLC dataset. The above processing was implemented in the blue tiles of Fig. 3.

3.1.3 Water extraction from Sentinel-2 images

Owing to the lack of 10-m LUCC or water datasets in 2016, we employed Sentinel-2 images to extract water data. Considering the scarcity of Sentinel-2 images before 2016, we collected all the images from June 2015 to December 2016. We classified water areas via the multiple index water detection rule (MIWDR) in the GEE platform, which was developed by

155 Deng et al. (2019). In the MIWDR, five spectral indices were calculated, including the modified normalized difference water index (MNDWI), normalized difference vegetation index (NDWI), enhanced vegetation index (EVI), and automated water extraction index for shadows and no shadows ($AWEI_{sh}$ and $AWEI_{nsh}$, respectively). Water areas were extracted via the criteria "$AWEI_{sh} - AWEI_{nsh} > -0.1$" and "MNDWI>EVI or MNDWI>NDVI". In this way, water time series from 2015 to 2016 were produced, which were then aggregated into annual water maps for 2016 via the mode algorithm. The above processing was

160 implemented in the light red tiles of Fig. 3.

3.2 Object-based hierarchical decision tree algorithm

To extract rivers from water areas, an object-based hierarchical decision tree algorithm was developed on the basis of geometric features and auxiliary datasets. To calculate the geometric features, chessboard segmentation was implemented in eCognition software, with scales larger than those of the columns and rows of the input water image. Five geometric features—

165 compactness, length/width, roundness, rectangular fit, and area—were calculated for the water objects. To analyse the geometric differences in water cover (e.g., rivers, lakes, and reservoirs), training samples were created on the basis of the 2020 CNLUCC map. Specifically, polygons of rivers, lakes and reservoirs were generated on the basis of the CNLUCC map. Random samples of each water cover type were created via a random sampling method. Then, the segmented water objects were overlaid with the training objects, and training objects of corresponding water types were obtained. Fig. 5 displays the

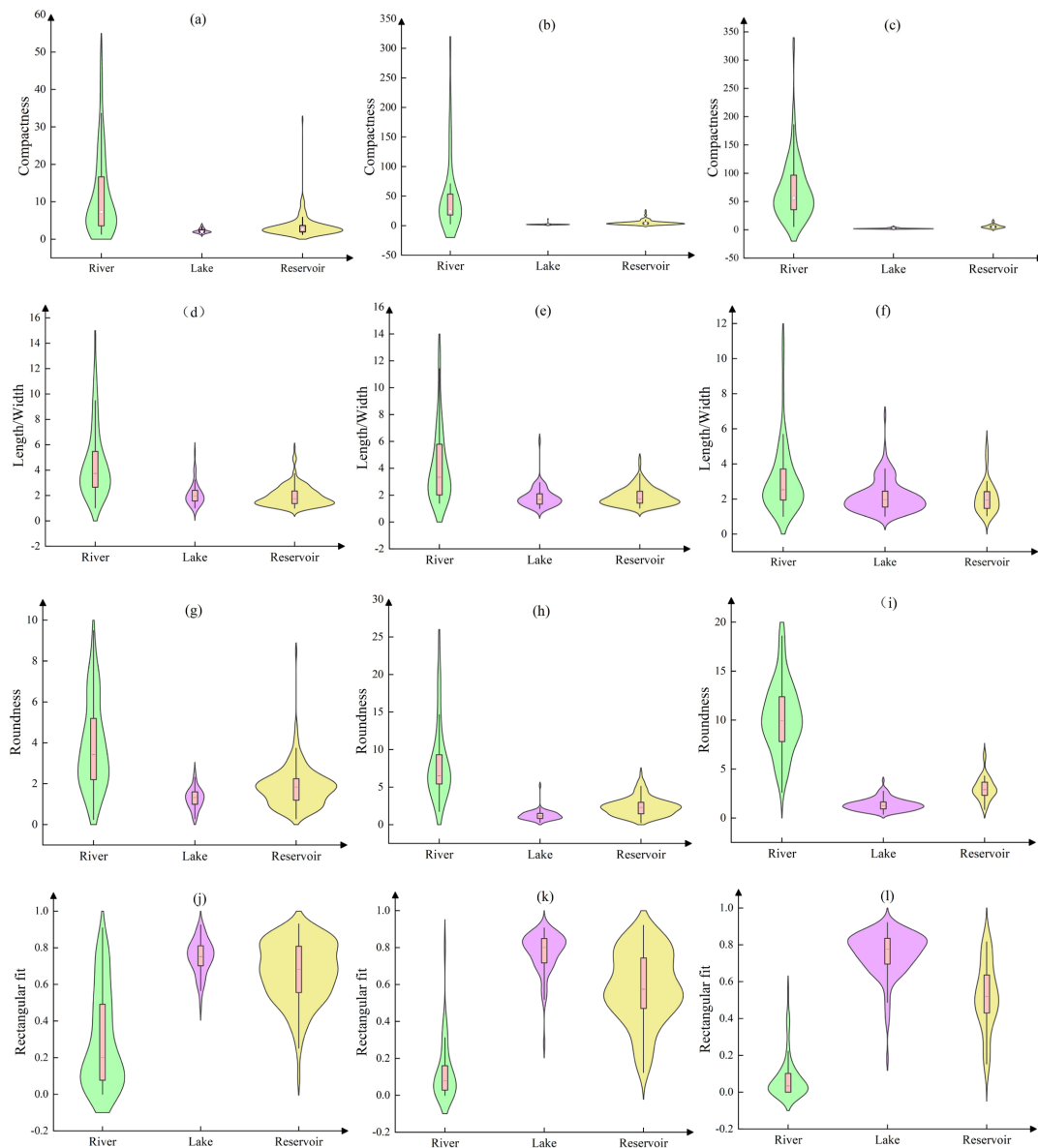170 geometric features of rivers, lakes and reservoirs.

Fig. 5. Violin plots and quartile box plots of geometric features. The columns represent three area range groups of 0–1000 ha, 1000–5000 ha, and >5000 ha, respectively. The rows represent the statistics of compactness, length/width, roundness and rectangular fit, respectively.

The geometric statistics indicated that the geometry of the lakes and reservoirs clearly differed from the rivers. Specifically, the compactness, roundness and length/width of rivers were greater than those of lakes and reservoirs, whereas the rectangular fit of rivers was less than that of lakes and reservoirs. Moreover, the above characteristics for the three water body types were more pronounced for water objects with larger areas. To reduce omission errors, we constructed five weak rules, namely, "compactness >2.3", "rectangular fit <0.5", "length/width >1.8" and "compactness >5.0". To reduce commission

9

errors, we used the five weak rules to construct strong rules. On the basis of our trials and errors, the rule set for river extraction was "compactness >2.3 & rectangular fit <0.5 & length/width >1.8" or "compactness >5.0".

Annual water maps → Chessboard segmentation

Geometric features

| Compactness | Length/width | Roundness | Rectangular fit | Area |

Dam points → Non-Reservoir → River rule set

Reservoir | Rough river | Unidentified waters

Manual editing | Manual editing

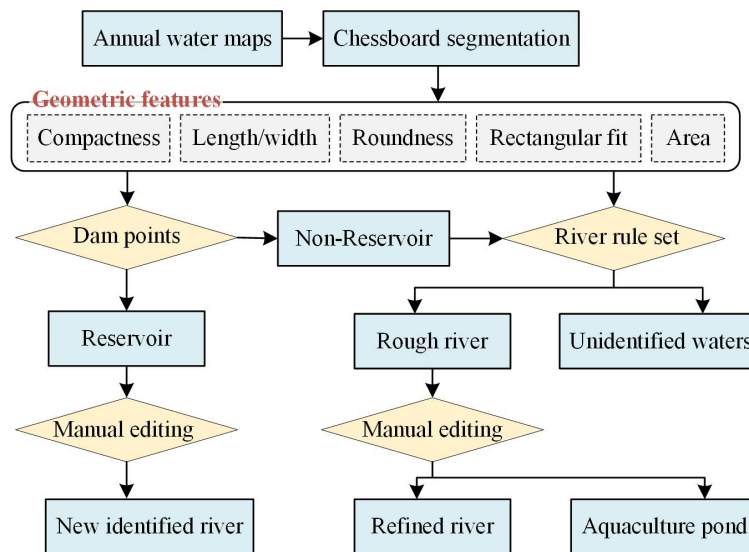New identified river | Refined river | Aquaculture pond

Fig. 6. Workflow of the object-based hierarchical decision tree

On the basis of the designed rule sets and auxiliary datasets, an object-based hierarchical decision tree (HDT) was developed to subsequently separate water cover types (Fig. 6). First, the dam points, including the Gaode reservoir POI, GOODD and OSM_Dam points, were used to identify reservoirs. These points were overlaid with water objects, and the selected water objects were labelled reservoirs. Then, the rivers were extracted via the river rule set. The remaining water objects were labelled as unidentified water. Finally, some manual edits were needed to improve the quality of the river maps, which involved three main situations: (1) Some aquaculture ponds were spatially connected with rivers, which cannot be separated via the river rule set. (2) The identified reservoirs may contain some rivers due to position errors at dam points. (3) Small rivers with narrow channels cannot be well identified via the river rule set because of their spatial disconnection. We manually revised these mapping errors.

3.3 Accuracy validation

To assess the mapping accuracy of rivers, test samples were produced by combining stratified random sampling and visual interpretation. First, river and non-river samples were produced via overlay, buffer and random sampling operations. Specifically, the rivers in the CNLUCC dataset for 2020 were extracted and then overlaid with the rivers from 2020. In the union regions, river samples were created via random sampling. After that, non-river samples were produced within a 300 m outside buffer of the union regions. Second, all random samples were visually interpreted by combining Collect Earth (CE), Google Earth (GE) and the GEE platform (Peng et al., 2024b). The CE software enables user-friendly sample management,

the GE provides high spatial resolution images, and the GEE offers median-composited Sentinel-2 images. Using these three

platforms, river and non-river samples from 2016 to 2023 were produced.

205   Finally, the accuracy indices, including the overall accuracy (OA), production accuracy (PA), user accuracy (UA) and F1

score, were calculated on the basis of the test samples. The F1 score measures the comprehensive accuracy of individual types

that balances the PA and UA, and is calculated as follows:

$$F1 = 2\frac{PA \times UA}{(PA + UA)} \times 100\%$$ (1)

3.4 Dataset intercomparison

210   To further validate the reliability of our river dataset, three existing river-related datasets—EA_Wetlands, CWaC and

CNLUCC—were used for intercomparison. Specifically, the rivers of EA_Wetlands in 2021, the rivers of CWaC in 2020, and

the rivers of the CNLUCC dataset in 2020 were used to test our river maps in the corresponding years. The first two datasets

have the same spatial resolution as our river dataset and were compared in 2021 and 2020. CNLUCC 2020 has a 30 m spatial

resolution with an OA of 91.2%, which was also used for data intercomparison because of its high mapping accuracy. To

215   comprehensively reflect their agreement, we aggregated the river extents within the spatial grid of 0.05°×0.05° and obtained

scatterplots and linear regressions with correlation coefficients (R) and root mean square errors (RMSEs).

3.5 Spatial–temporal analysis of China's areas

On the basis of our river maps from 2016 to 2023, the spatial patterns and areas of rivers at both the national scale and

river basin scale were analysed. Furthermore, the Theil–Sen method was used to calculate the change slope of river areas

220   within a grid of 0.05°×0.05°. This method is a robust nonparametric algorithm that can reduce the impact of outliers on the

computed slope (Qiu et al., 2024). Its formula is as follows:

$$S = Median(\frac{x_j - x_i}{j - i}) \quad \forall j > i$$ (2)

where S denotes the Theil–Sen slope, i and j denote years, and x_i and x_j denote river areas within a grid of 0.05°×0.05°

in different years. If $S < 0$, there was a decreasing trend in the river area of the grid. If $S = 0$, it indicates a stable river area. If

225    $S > 0$, there was an increasing trend in the river area of the grid.

4 Results

4.1 River accuracy assessments

Combining the DW dataset, GLEC dataset and Sentinel-2 images, we produced annual China river extent maps (CRED)

from 2016 to 2023. We validated the CRED via test samples that were manually inspected via visual interpretation. The CRED

230   achieved stable and good accuracy, with all OAs exceeding 96.8% each year (Table 2). For rivers, the UAs and PAs from 2016

to 2023 exceeded 95.3% and 91.3%, respectively. For the non-river areas, the UAs and PAs in individual years exceeded 96.0%

and 97.7%, respectively. The above results indicated that the omission and commission errors of our river maps were very low.

Table 2. Accuracy assessments of our river maps from 2016 to 2023

	River			Non-river			OA
	UA	PA	F1 score	UA	PA	F1 score	
2016	96.8%	92.9%	94.8%	97.9%	99.1%	98.5%	97.6%
2017	97.9%	94.5%	96.2%	97.4%	99.0%	98.2%	97.5%
2018	96.2%	91.3%	93.7%	96.0%	98.3%	97.1%	96.0%
2019	97.2%	94.5%	95.8%	97.2%	98.6%	97.9%	97.2%
2020	96.5%	94.2%	95.4%	96.9%	98.2%	97.5%	96.8%
2021	95.4%	94.4%	94.9%	97.5%	98.0%	97.7%	96.8%
2022	95.3%	94.9%	95.1%	97.5%	97.7%	97.6%	96.8%
2023	98.4%	96.5%	97.5%	98.5%	99.3%	98.9%	98.5%

235   In addition, the F1 score, which measures the comprehensive accuracy of individual types, was calculated. The results

indicated that the F1 scores of rivers and non-rivers were greater than 93.7% and 97.1%, respectively. On the basis of the above

analysis, we believe that the rivers in the CECD from 2016 to 2023 achieved reliable accuracy.

4.2 Intercomparison with river-related datasets

To further test the reliability of the CRED, three river-related river datasets, including CNLUCC, CWaC and

240   EA_Wetlands, were chosen for data intercomparison. To quantify the consistency among datasets, river fractions within a 0.05°

by 0.05° grid were counted (Fig. 7). The results indicated that our CRED exhibited high correlation coefficients (R) with the

three river-related datasets. CNLUCC, considered one of the most reliable land use datasets, had the highest R among the three

datasets, with a value of 0.832. The CNLUCC was produced by experienced interpreters through visual interpretation on the

basis of Landsat images. The high R value between the CNLUCC dataset and the CRED indicated the good reliability of the

245   rivers in the CRED. CWaC and EA_Wetlands were generated via intelligent algorithms at a 10 m spatial resolution, and their

R values with the CRED were 0.775 and 0.750, respectively. Overall, our river maps exhibited good consistency with existing
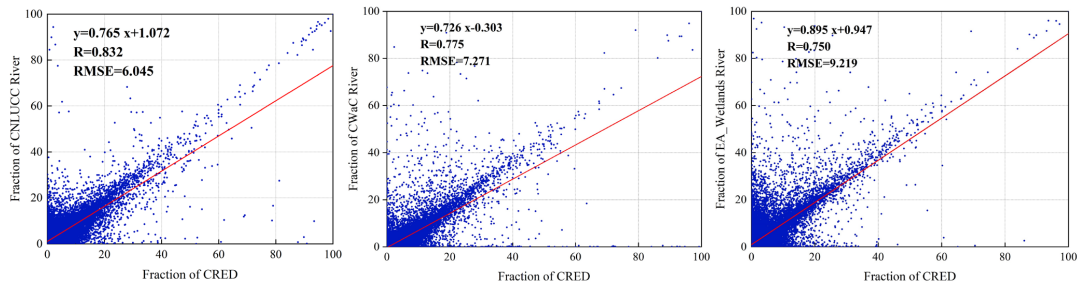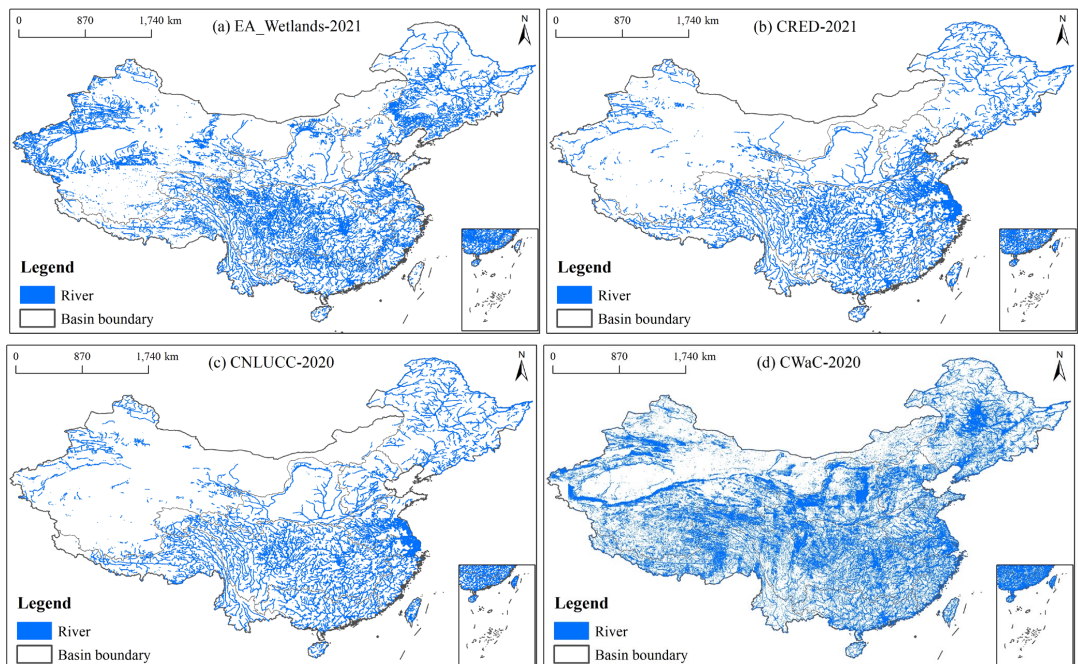
river-related datasets.

Fig. 7. Scatterplots of the river fraction between the CRED and CNLUCC, CWaC and EA_Wetlands. The river fraction was aggregated within 0.5° by 0.5° grids.

To inspect the spatial consistency between the CRED and the three river-related datasets, the spatial distributions of Chinese rivers in the corresponding datasets were mapped (Fig. 8). These river maps generally shared similar distributions. Among the three water-related datasets, CNLUCC had the most similar patterns with the CRED (Fig. 8 (c) & (e)). For the EW_Wetlands and CWaC, their rivers had considerable differences from the CRED. The EA_Wetlands dataset delineated more rivers west of the Continental Basin (CB) and southwest of the Songhua and Liaohe River Basin (SLRB) (Fig. 8 (a)). CWaC misclassified many small-area patches as rivers, such as ice, snow, and shadows (Fig. 8 (d)). In addition, river areas of individual river maps were calculated, and six typical regions were selected to check spatial details among river datasets (Fig. 8 (f)).
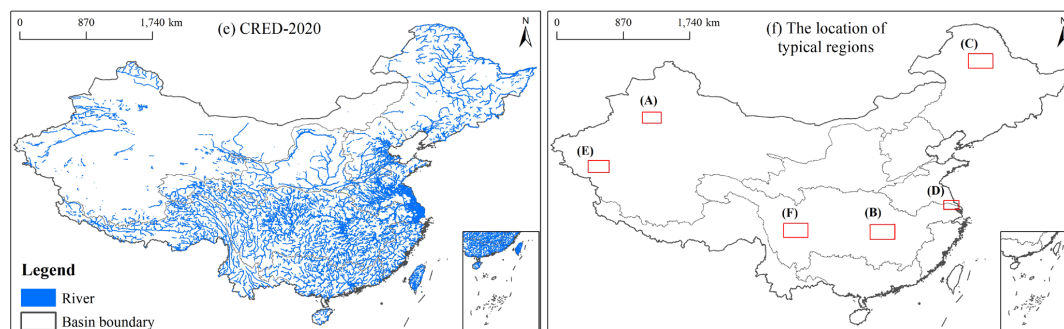
Fig. 8. River patterns of EA_wetlands (2021), CLUCC (2020), CWaC (2020) and the CRED (2020/2021). (f) displays six typical regions, and their rivers are shown in the supplementary materials.

For EA_Wetlands and the CRED in 2021, the river areas were 64886.84 km^2 and 47397.75 km^2, respectively. The river area in EA_Wetlands was larger than that in the CRED, mainly because of various factors. On the one hand, the EA_Wetlands dataset better represented rivers in complex mountainous areas and seasonal rivers than the CRED (Fig. S2(A)). On the other hand, the EA_Wetlands dataset misclassified some linear water areas as rivers, such as Dongting Lake (Fig. S2(B)).

The river areas of the CNLUCC dataset and the CRED in 2020 were 47563.20 km^2 and 44694.64 km^2, respectively. The river areas of the two datasets were nearly equal. However, due to spatial resolution differences in satellite images, some differences were still identified. CNLUCC did not effectively extract narrow rivers in mountainous areas (Fig. S3(C)) or identify canals/channels along coastal regions in Southeast China (Fig. S3(D)). However, the CRED was able to effectively recognize these rivers.

The river areas of CWaC and the CRED in 2020 were 33877.90 km^2 and 44694.64 km^2, respectively. The river area in CWaC was smaller than that in the CRED, likely due to an underestimation of river extents by CWaC. CWaC was produced by combining Sentinel-1 SAR (synthetic aperture radar) images and Sentinel-2 MSI (multispectral image) images. As reported by Li and Niu (2022), water bodies extracted via SAR images are often underestimated. Consequently, the river extents mapped by CWaC may also be underestimated. In addition, considerable small patches, such as ice, snow and shadows, were misclassified as rivers in CWaC (Fig. S4 (E)). Moreover, the CRED was superior to CWaC in terms of data postprocessing. For example, CWaC misclassified some reservoirs as rivers, whereas the CRED accurately excluded them (Figure S4 (F)).

4.3 Spatial pattern and area changes of China's rivers

In this study, the annual river extents of Chinese rivers from 2016 to 2023 were mapped, and their spatial distributions are displayed in Fig. S5. China's rivers are characterized by more southern rivers and fewer northern rivers, more eastern rivers and fewer western rivers. Generally, most rivers are distributed in Southeast China and Southwest China. Northeast China also has many rivers.

In 2023, the river areas of China were 44,948.78 km^2. Among the nine river basins, the YaRB had the largest river area

14

and densest river network, accounting for 42.55% of China's river area in 2023 (Fig. 9). The others were the SLRB, PRB, HuRB, YeRB, SWB, SEB, CB, and HaRB in descending order, with river areas accounting for 14.10%, 11.81%, 7.40%, 7.21%, 6.33%, 5.24%, 3.17% and 2.00%, respectively. The CB is the largest river basin in China but has the sparest river network. The HaRB is the smallest river basin in China and has the smallest river areas; however, it has a relatively dense river network.
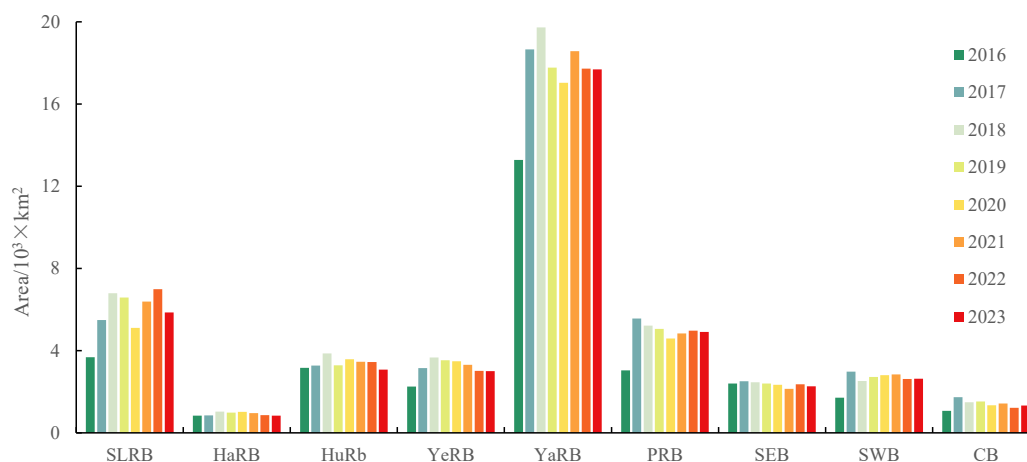

290

Fig. 9. River areas from 2016 to 2023 in nine river basins

Temporally, the river areas of China from 2016 to 2023 were characterized by an initial increase, followed by a decrease and then a slight increase. This change trend was also observed for the SLRB, HuRB, YeRB, PRB, SEB, SWB, and CB. For the HaRB and YeRB, the river area first increased but then decreased. In the past eight years, the river areas of 2018 in the

295   SLRB, HaRB, HuRB, YeRB and YaRB were the largest, whereas the river areas of 2017 in the PRB, SEB, SWB and CB were the largest.

4.4 Changing characteristics of China's rivers

To explicitly examine the change characteristics of China's rivers, the Theil-Sen slope of river areas within 0.05°×0.05° grids was calculated (Fig. 10). More than half of China's rivers had a negative slope, indicating a decreasing trend in river

300   area. The decreased areas were distributed mainly in coastal regions, such as the Yangtze River Delta, HaRB, HuRB, SEB and PRB. Most of these regions had slope values ranging from 0 to 1 (Fig. 10 (b)). These trends and distribution characteristics were related mainly to human activities. Regions with decreased river areas experienced high-level economic development. Correspondingly, their human disturbances were more frequent. Over the past eight years, the expansion of construction land, industrial and agricultural development, and other areas in these regions have contributed to the degradation of river areas.
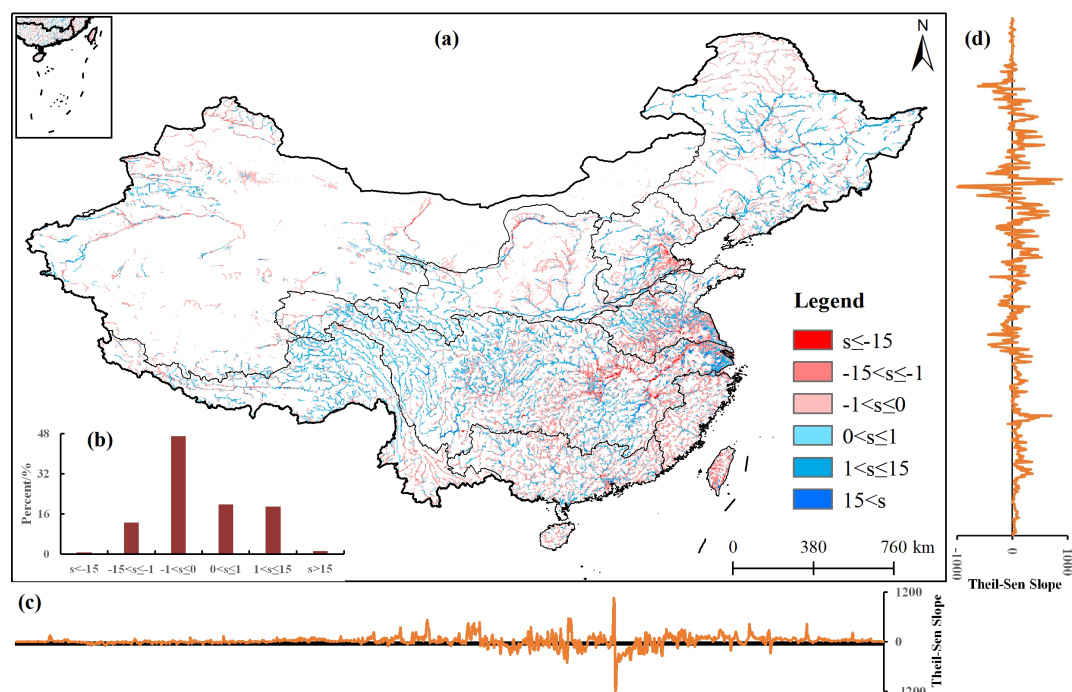
305

Fig. 10 Theil-Sen slope of rivers in the CRED. (a) Theil–Sen slope for 0.05°×0.05° grids in China. (b) Histogram of Theil–Sen slopes for the river grids. (c) and (d) Slope histograms with 0.05° increments in longitude and latitude.

For grids with increasing river areas, most are located in the southern SLRB, western YaRB, and southeastern CB. Most of the increased river grids had slope values ranging from 0 to 15, which accounted for 38.70% of the total river grids (Fig. 10

310 (a&b)). For increased river grids, their change trends and distributions could be linked with ecological restoration and water conservancy projects. On the one hand, local governments have implemented a series of policies to protect natural ecosystems. Large encroachments on river ecosystems are strictly prohibited in China. On the other hand, some water conservancy projects, such as the Yellow River Diversion Irrigation project and the South-to-North Water Diversion project, have greatly promoted the increase in river area.

315 In addition, we also calculated the river change slope by longitude and latitude at a 0.05° interval (Fig. 10 (c&d)). From the perspective of longitude, rivers with significant changes were mainly concentrated in central and eastern China, covering the southern coastal regions. These changes were most related to high-speed economic development. From the perspective of latitude, rivers with significant changes were distributed mainly in northern China. These change characteristics were linked mainly to the new constructions of cannels and conservancy facilities.

16

5 Discussion

5.1 Reliability of river mapping

In this study, we successfully produced annual river maps of China from 2016 to 2023. To our knowledge, this is the first comprehensive effort to map annual river extents across China at a 10 m spatial resolution. The successful implementation of our study can be attributed to three factors: the availability of multisource LUCC datasets, the reliability of existing water extraction algorithms, and the effectiveness of our river mapping algorithm.

First, multisource LUCC datasets at a spatial resolution of 10 m provide a good foundation for river mapping, such as ESRI's world cover (Zanaga et al., 2021), FROM-GLC10 (Gong et al., 2019), DW, and EGLC. The DW dataset, with a revisit frequency of 2–5 days, allowed us to capture seasonal variations in water extent (Brown et al., 2022). However, it is often unavailable in Southwest China and the Tibetan Plateau because the DW dataset was produced using only Sentinel-2 images with less than 35% cloud coverage. To address this limitation, the EGLC was used as supplementary data in areas where the DW dataset was missing, ensuring more comprehensive river mapping.

Second, existing water extraction algorithms also support river mapping. Owing to the scarcity of Sentinel-2 images from 2015 to 2016, the EGLC dataset was unavailable in 2016, and significant portions of the DW dataset were also missing for China during this period (Karra et al., 2021; Wang et al., 2023c). To address this issue, the MIWDR algorithm was employed to extract water areas in 2016, which served as supplementary data for the DW dataset. The MIWDR, which is designed with multiple water indices, has been shown to be effective at extracting water from mountainous and urban areas, making it highly suitable for large-scale water classification (Deng et al., 2019).

Third, we developed an object-based hierarchical decision tree algorithm for river extraction. Rivers, characterized by their linear and narrow shapes, have distinct geometries compared with other water covers (e.g., lakes and reservoirs) (Peng et al., 2023). To effectively extract rivers, five geometric features, which were calculated via chessboard segmentation, were used to develop a hierarchical decision tree. Compared with pixel-based methods, our algorithm can handle the spectral similarity of water cover types. Compared with general object-based methods, our algorithm has good robustness, as its rules and thresholds do not change over time or across regions.

An accuracy validation indicated that our annual river maps achieved high accuracy, with OAs exceeding 96%. For the individual accuracy metrics, the UAs, PAs and F1 scores of the river areas were greater than 95.3%, 91.3% and 93.7%, respectively. To further check the reliability of our CRED, three existing river datasets, including EA_Wetlands, CNLUCC, and CWaC, were selected for intercomparison. The results showed that our CRED exhibited similar spatial distributions to those of the existing datasets, with R values exceeding 0.75. Moreover, the CRED was superior in mapping small and narrow rivers and reducing river misclassifications. Overall, the above analysis confirmed that our river maps are both accurate and reliable, effectively delineating river extents and dynamics.

5.2 Improvements and potential contributions of the CRED

Owing to their ecological and economic importance, river datasets have been produced to support river management and economic development. These datasets can be divided into river line datasets and river polygon datasets (Table 3). River line datasets record channel lines and reflect river spatial distributions well. However, these datasets cannot delineate river extents and area changes. River polygon datasets delineate river extents with polygon patches. However, most of these datasets are

355   limited by low temporal and spatial resolutions. On the one hand, long-term river polygon datasets are scarce and fail to reflect river dynamics. On the other hand, the spatial resolutions of these datasets are intermediate or low and are thus insufficient to delineate river extents in detail. In contrast, the CRED has better spatial and temporal resolution. It maps annual river extents in China at a spatial resolution of 10 m, which can monitor river extents and their dynamics well.

360   **Table 3. List of existing river datasets covering the globe or China. The dashed lines indicate that the corresponding information is unavailable.**

Type	Name	Spatial extent	Spatial resolution	Time period	Resource
River lines	GRRATS	Global	---	---	(Coss et al., 2020)
	HydroRIVERS	Global	500 m	---	(Lehner and Grill, 2013)
	MCRW	China	30 m	1990–2015	(Yang et al., 2020a)
	Reach-level Bankfull river	Global	30 m	---	(Lin et al., 2020)
River polygons	GLWD	Global	1 km	1980s	(Lehner and Doll, 2004)
	GRWL	Global	30 m	---	(Allen and Pavelsky, 2018)
	EA_Wetlands	East Asia	10 m	2021	(Wang et al., 2023a)
	CAS_Wetlands	China	30 m	2015	(Mao et al., 2020)
	CWaC	China	10 m	2020	(Li and Niu, 2022)
	River_OSM	Global	---	---	https://download.geofabrik.de/asia/china.html

Long-term and accurate river extent datasets significantly benefit biodiversity conservation and sustainable river ecosystem management (Gonzalez-Ferreras and Barquin, 2017). For example, river ecosystems not only provide habits and

365   food for fish, amphibians, birds and aquatic plants but also serve as curial natural corridors for species migration, connecting

different ecological zones (Mar Sanchez-Montoya et al., 2022). The CRED maps, with detailed spatial information and dynamic changes, can be used as baseline data for managing, protecting and improving river ecosystems. On the basis of the CRED, many river-related indices, such as hydrological connectivity, patch area, and land degradation, can be calculated, providing valuable insights into the environmental health of river ecosystems (Yan and Niu, 2019).

370   Moreover, the CRED can also effectively serve economic development. A river equipped with good shipping capacity can greatly facilitate commodity transportation and trade ((Spear et al., 2024)). In southeastern China, for example, the dense river network provides good access to shipping, which largely contributes to local economic development ((Liu et al., 2018a)). The CRED records annual river extents over the past eight years and has good potential for enhancing shipping capacity. Shipping conditions, including the river width, area and spatial connection, can be easily derived from the CRED. In agriculture,

375   rivers provide irrigation water for arable land, ensuring crop growth (Lu et al., 2023). Regular and dynamic monitoring of river distribution and area changes is beneficial for agricultural production and management.

5.3 Uncertainties, limitations and improvements

This study focused on river extraction from China and successfully produced high-quality river maps from 2016 to 2023. However, there are still some uncertainties and limitations. First, the multisource LUCC datasets and water extraction

380   algorithm used for CRED production may result in uncertainties. As previously mentioned, the DW dataset cannot cover all of China. To address this issue, the EGLC dataset and water classification by the MIWDR algorithm were used as supplementary datasets. The three datasets shared different characteristics in terms of water mapping. The DW well captured water seasonal variation but missed many water areas in mountain areas (Li and Wang, 2024). The EGLC delineates only annual water extents without seasonal information (Venter et al., 2022). The MIWDR algorithm misclassified some ice, snow

385   and shadows as water. The above characteristics of the three datasets may lead to errors in river extraction.

In addition, the DW, EGLC, and MIWDR methods failed to adequately map water areas on both the Tibetan Plateau and in Southwest China. These errors were related mainly to mountain shadows and complex environments. Water areas share spectral characteristics similar to those of mountain shadows, making them prone to misclassification (Wang et al., 2023b). Complex environments, such as cloud contamination and ice/snow, make accurate extraction of water difficult. In addition,

390   the mode algorithm, which is used to map the most common rivers on an annual scale (Calderon-Loor et al., 2021), may also result in uncertainties. Owing to seasonal variability, the boundaries of mapped rivers may be contentious. How to define the extent of a river, such as by using the maximum, minimum or medium water inundated areas, is a debated issue.

In future studies, the data quality and time range of the CRED could be further improved. As previously discussed, many uncertainties in the CRED are largely caused by multisource LUCC datasets or water datasets. Despite the availability of water

395   datasets at 10 m spatial resolution(Li et al., 2023; Vanderhoof et al., 2023; Yang et al., 2020b), long-term global water datasets

are still scarce. Thus, future attempts should focus on developing robust and accurate water classification algorithms for Sentinel-1/2 images, which are aimed at producing long-term and large-scale water data with high consistency. In return, these water classification studies support river extraction well. Additionally, as Sentinel-1/2 imagery accumulates, the CRED should be updated in a timely manner to extend its time range.

400   **6 Data availability**

The China River Extent Dataset (CRED) is publicly available for download from the Zenodo repository https://doi.org/10.5281/zenodo.13841910 (Peng et al., 2024a). It delineates river extents across China from 2016 to 2023. Correspondingly, the CRED contains eight shapefiles. It has an area attribute, with units of hectares (ha). The projection of the CRED is WGS_1984_Albert.

405   **7 Conclusion**

Rivers play crucial roles in terrestrial ecosystems and human well-being, not only by providing freshwater resources and supporting biodiversity but also by serving the functions of navigation, cultural recreation and flood disaster prevention. To support the high-quality development of China, it is vital to produce sequential and fine-resolution river extent maps. On the basis of a literature review, however, fine-resolution annual river extent maps across China have rarely been investigated in

410   past studies. Thus, to better understand the spatial distribution and dynamics of China's rivers, this study generated the first Sentinel-derived annual China river extent dataset (CRED) from 2016 to 2023 on the basis of the DW dataset, EGLC dataset and MIWDR algorithm.

An accuracy evaluation indicated that the OAs of the CRED from 2016 to 2023 exceeded 96.0%. For the individual accuracy indices, the UAs, PAs and F1 scores of the rivers exceeded 95.3%, 91.3% and 93.7%, respectively. To further check

415   the reliability of our CRED, data comparisons with EA_Wetlands, CNLUCC and CWaC were implemented. The results indicated that the CRED shared similar distributions with the three datasets, with R values over 0.75. Moreover, the CRED dataset was superior to existing river-related datasets in terms of spatial details, river misclassifications and omissions. Spatial–temporal analysis revealed that rivers in China are more abundant in East China than in West China and more abundant in South China than in North China. The YaRB had the largest river area among the nine river basins, accounting for 42.55% of

420   China's rivers. From 2016 to 2023, the areas with decreases in rivers were mostly located in Southeast China, whereas the areas with increases in rivers were mostly distributed in Central and Northeast China. Overall, this study produced high-quality annual river maps for China and explored their dynamics from 2016 to 2023, which is important for the protection, management and sustainable use of rivers.

Author Contributions

425   FP conceived this study. The methodology design, manuscript wiring, and format analysis were done by FP and XW. Data processing, software, and programming were done by FP, SB and XW. FP and WJ supervised, acquired and administered the project. Formal analysis and manuscript improvement were done by MM. Quality control and accuracy validation were done by FP, SB, WJ, MM, and XW.

Competing interests

430   The contact author has declared that none of the authors has any competing interests

445   **References**

Allen, G. H. and Pavelsky, T. M.: Global extent of rivers and streams, Science, 361, 585-587, doi:10.1126/science.aat0636, 2018.

Bai, X. and Zhao, W.: Impacts of climate change and anthropogenic stressors on runoff variations in major river basins in China since 1950, Sci. Total Environ., 898, doi:10.1016/j.scitotenv.2023.165349, 2023.

450   Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., and Tait, A. M.: Dynamic World, Near real-time global 10 m land use land cover mapping, Sci. Data, 9, doi:10.1038/s41597-022-01307-4, 2022.

Caissie, D.: The thermal regime of rivers: a review, Freshw. Biol., 51, 1389-1406, doi:10.1111/j.1365-2427.2006.01597.x, 2006.

455   Calderon-Loor, M., Hadjikakou, M., and Bryan, B. A.: High-resolution wall-to-wall land-cover mapping and land change assessment for Australia from 1985 to 2015, Remote Sens. Environ., 252, doi:10.1016/j.rse.2020.112148, 2021.

Coss, S., Durand, M., Yi, Y., Jia, Y., Guo, Q., Tuozzolo, S., Shum, C. K., Allen, G. H., Calmant, S., and Pavelsky, T.: Global River Radar Altimetry Time Series (GRRATS): new river elevation earth science data records for the hydrologic community, Earth Syst. Sci. Data, 12, 137-150, doi:10.5194/essd-12-137-2020, 2020.

460   Deng, Y., Jiang, W., Tang, Z., Ling, Z., and Wu, Z.: Long-Term Changes of Open-Surface Water Bodies in the Yangtze River Basin Based on the Google Earth Engine Cloud Platform, Remote Sens., 11, doi:10.3390/rs11192213, 2019.

Fei, J., Liu, J., Ke, L., Wang, W., Wu, P., and Zhou, Y.: A deep learning-based method for mapping alpine intermittent rivers and ephemeral streams of the Tibetan Plateau from Sentinel-1 time series and DEMs, Remote Sens. Environ., 282, doi:10.1016/j.rse.2022.113271, 2022.

465   Gong, P., Liu, H., Zhang, M., Li, C. C., Wang, J., Huang, H. B., Clinton, N., Ji, L., Li, W. Y., Bai, Y., Chen, B., Xu, B., Zhu, Z., Yuan, C., Suen, H., Guo, J., Xu, N., Li, W., Zhao, Y., Yang, J., Yu, C., Wang, X., Fu, H., Yu, L., Dronova, I., Hui, F., Cheng, X., Shi, X., Xiao, F., Liu, Q., and Song, L.: Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017, Sci. Bull., 64, 370-373, doi:10.1016/j.scib.2019.03.002, 2019.

470   Gonzalez-Ferreras, A. M. and Barquin, J.: Mapping the temporary and perennial character of whole river networks, Water Resour. Res., 53, 6709-6724, doi:10.1002/2017WR020390, 2017.

Hotchkiss, E. R., Hall, R. O., Jr., Sponseller, R. A., Butman, D., Klaminder, J., Laudon, H., Rosvall, M., and Karlsson, J.: Sources of and processes controlling CO_2 emissions change with the size of streams and rivers, Nat. Geosci., 8, 696-699, doi:10.1038/NGEO2507, 2015.

475   Jiang, L., Zuo, Q., Ma, J., and Zhang, Z.: Evaluation and prediction of the level of high-quality development: A case study of the Yellow River Basin, China, Ecol. Indic., 129, doi:10.1016/j.ecolind.2021.107994, 2021.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., and Brumby, S. P.: Global land use/land cover with Sentinel 2 and deep learning, 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, doi:10.1109/IGARSS47720.2021.9553499, 2021.

480   Lehner, B. and Doll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296, 1-

22, doi:10.1016/j.jhydrol.2004.03.028, 2004.

Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrol. Process., 27, 2171-2186, doi:10.1002/hyp.9740, 2013.

Li, D., Wu, B. S., Chen, B. W., Qin, C., Wang, Y. J., Zhang, Y., and Xue, Y.: Open-Surface River Extraction Based on Sentinel-2 MSI Imagery and DEM Data: Case Study of the Upper Yellow River, Remote Sens., 12, doi:10.3390/rs12172737, 2020.

Li, J., Li, L., Song, Y., Chen, J., Wang, Z., Bao, Y., Zhang, W., and Meng, L.: A robust large-scale surface water mapping framework with high spatiotemporal resolution based on the fusion of multi-source remote sensing data, Int. J. Appl. Earth Obs. Geoinf., 118, doi:10.1016/j.jag.2023.103288, 2023.

Li, K. and Wang, J.: A multi-source data fusion method for land cover production: a case study of the East European Plain, Int. J. Digit. Earth, 17, doi:10.1080/17538947.2024.2339360, 2024.

Li, Y. and Niu, Z.: Systematic method for mapping fine-resolution water cover types in China based on time series Sentinel-1 and 2 images, Int. J. Appl. Earth Obs. Geoinf., 106, doi:10.1016/j.jag.2021.102656, 2022.

Lin, P., Pan, M., Allen, G. H., de Frasson, R. P., Zeng, Z., Yamazaki, D., and Wood, E. F.: Global Estimates of Reach-Level Bankfull River Width Leveraging Big Data Geospatial Analysis, Geophys. Res. Lett., 47, doi:10.1029/2019GL086405, 2020.

Liu, H., Jin, X., Wu, L., Wang, X., Fu, M., Lv, Z., Morawska, L., Huang, F., and He, K.: The impact of marine shipping and its DECA control on air quality in the Pearl River Delta, China, Sci. Total Environ., 625, 1476-1485, doi:10.1016/j.scitotenv.2018.01.033, 2018a.

Liu, J., Ning, J., Kuang, W., Xu, X., Zhang, S., Yan, C., Li, R., Wu, S., Hu, Y., Du, G., Chi, W., Pan, T., and Ning, J.: Spatio-temporal patterns and characteristics of land-use change in China during 2010-2015, Acta Geogr. Sin., 73, 789-802, doi:10.11821/dlxb201805001, 2018b.

Liu, J., Kuang, W., Zhang, Z., Xu, X., Qin, Y., Ning, J., Zhou, W., Zhang, S., Li, R., Yan, C., Wu, S., Shi, X., Jiang, N., Yu, D., Pan, X., and Chi, W.: Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s, J. Geogr. Sci., 24, 195-210, doi:10.1007/s11442-014-1082-6, 2014.

Lu, Q., Jing, K., Li, X., Song, X., Zhao, C. P., and Du, S.: Effects of Yellow River Water Management Policies on Annual Irrigation Water Usage from Canals and Groundwater in Yucheng City, China, Sustainability, 15, doi:10.3390/su15042885, 2023.

Mao, D., Wang, Z., Du, B., Li, L., Tian, Y., Jia, M., Zeng, Y., Song, K., Jiang, M., and Wang, Y.: National wetland mapping in China: A new product resulting from object-based and hierarchical classification of Landsat 8 OLI images, ISPRS J. Photogramm. Remote Sens., 164, 11-25, doi:10.1016/j.isprsjprs.2020.03.020, 2020.

Mar Sanchez-Montoya, M., Gomez, R., Calvo, J. F., Bartonicka, T., Datry, T., and Paril, P.: Ecological values of intermittent rivers for terrestrial vertebrate fauna, Sci. Total Environ., 806, doi:10.1016/j.scitotenv.2021.151308, 2022.

Ministry of Water Resources: China census for water: Bulletin of first national water census, in, Beijing, pp. 1-22, 2013.

Mulligan, M., van Soesbergen, A., and Saenz, L.: GOODD, a global dataset of more than 38,000 georeferenced dams, Sci. Data, 7, doi:10.1038/s41597-020-0362-5, 2020.

515 Pekel, J. F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418-422, doi:10.1038/nature20584, 2016.

Peng, K., Si, B., Jiang, W., Ma, M., and Wang, X.: The China river extent maps (CRED) from 2016 to 2023, Zenodo [dataset], https://doi.org/10.5281/zenodo.13841910, 2024a.

Peng, K., Jiang, W., Hou, P., Cui, T., Wu, Z., and Si, B.: Exploring the long-term dynamics of detailed wetland types and their 520 driving forces in coastal metropolitan areas from 1990 to 2020, Int. J. Appl. Earth Obs. Geoinf., 132, doi:10.1016/j.jag.2024.104012, 2024b.

Peng, K., Jiang, W., Hou, P., Wu, Z., Ling, Z., Wang, X., Niu, Z., and Mao, D.: Continental-scale wetland mapping: A novel algorithm for detailed wetland types classification based on time series Sentinel-1/2 images, Ecol. Indic., 148, 110113, doi:10.1016/j.ecolind.2023.110113, 2023.

525 Pickens, A. H., Hansen, M. C., Hancher, M., Stehman, S. V., Tyukavina, A., Potapov, P., Marroquin, B., and Sherani, Z.: Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series, Remote Sens. Environ., 243, doi:10.1016/j.rse.2020.111792, 2020.

Qiu, M., Liu, Y., Tian, F., Wang, S., and Fu, B.: Marsh decrease was much faster than the water increase among the Yellow River Source wetlands during 1986-2022, Sci. Total Environ., 947, doi:10.1016/j.scitotenv.2024.174377, 2024.

530 Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Duerr, H., Meybeck, M., Ciais, P., and Guth, P.: Global carbon dioxide emissions from inland waters, Nature, 503, 355-359, doi:10.1038/nature12760, 2013.

She, Y., Liu, Y., Jiang, L., and Yuan, H.: Is China's River Chief Policy effective? Evidence from a quasi-natural experiment in the Yangtze River Economic Belt, China, J. Clean. Prod., 220, 919-930, doi:10.1016/j.jclepro.2019.02.031, 2019.

535 Spear, M. J., Harris, B. S., Bookout, T. A., Ickes, B. S., Jankowski, K. J., Solomon, L. E., Maxson, K. A., Harris, A. L. W., Mathis, A. T., Schaick, S. J., Williams, J. A., DeBoer, J. A., Lenaerts, A. W., Hine, E. C., Chick, J. H., and Lamer, J. T.: Reduction of large vessel traffic improves water quality and alters fish habitat-use throughout a large river, Sci. Total Environ., 946, doi:10.1016/j.scitotenv.2024.172705, 2024.

Speed, R. A., Li, Y., Tickner, D., Huang, H., Naiman, R. J., Cao, J., Lei, G., Yu, L., Sayers, P., Zhao, Z., and Wei, Y.: A 540 framework for strategic river restoration in China, Water Int., 41, 998-1015, doi:10.1080/02508060.2016.1247311, 2016.

Valman, S. J., Boyd, D. S., Carbonneau, P. E., Johnson, M. F., and Dugdale, S. J.: An AI approach to operationalise global daily PlanetScope satellite imagery for river water masking, Remote Sens. Environ., 301, doi:10.1016/j.rse.2023.113932, 2024.

Vanderhoof, M. K., Alexander, L., Christensen, J., Solvik, K., Nieuwlandt, P., and Sagehorn, M.: High-frequency time series comparison of Sentinel-1 and Sentinel-2 satellites for mapping open and vegetated water across the United States (2017-2021),
545   Remote Sens. Environ., 288, doi:10.1016/j.rse.2023.113498, 2023.

Venter, Z. S., Barton, D. N., Chakraborty, T., Simensen, T., and Singh, G.: Global 10 m Land Use Land Cover Datasets: A Comparison of Dynamic World, World Cover and Esri Land Cover, Remote Sens., 14, doi:10.3390/rs14164101, 2022.

Wang, M., Mao, D., Wang, Y., Xiao, X., Xiang, H., Feng, K., Luo, L., Jia, M., Song, K., and Wang, Z.: Wetland mapping in East Asia by two-stage object-based Random Forest and hierarchical decision tree algorithms on Sentinel-1/2 images, Remote
550   Sens. Environ., 297, doi:10.1016/j.rse.2023.113793, 2023a.

Wang, X., Zhou, G., Lv, X., Zhou, L., Hu, M., He, X., and Tian, Z.: Comparison of Lake Extraction and Classification Methods for the Tibetan Plateau Based on Topographic-Spectral Information, Remote Sens., 15, doi:10.3390/rs15010267, 2023b.

Wang, Y. Z., Sun, Y. H., Cao, X. Y., Wang, Y. H., Zhang, W. K., and Cheng, X. L.: A review of regional and Global scale Land Use/Land Cover (LULC) mapping products generated from satellite remote sensing, ISPRS J. Photogramm. Remote Sens.,
555   206, 311-334, doi:10.1016/j.isprsjprs.2023.11.014, 2023c.

Wei, Z., Jia, K., Jia, X., Khandelwal, A., and Kumar, V.: Global River Monitoring Using Semantic Fusion Networks, Water, 12, doi:10.3390/w12082258, 2020.

Wickham, J., Stehman, S. V., Sorenson, D. G., Gass, L., and Dewitz, J. A.: Thematic accuracy assessment of the NLCD 2019 land cover for the conterminous United States, GIScience Remote Sens., 60, doi:10.1080/15481603.2023.2181143, 2023.

560   Xu, P., Herold, M., Tsendbazar, N. E., and Clevers, J.: Towards a comprehensive and consistent global aquatic land cover characterization framework addressing multiple user needs, Remote Sens. Environ., 250, doi:10.1016/j.rse.2020.112034, 2020.

Xu, P., Tsendbazar, N. E., Herold, M., Clevers, J., and Li, L. L.: Improving the characterization of global aquatic land cover types using multi-source earth observation data, Remote Sens. Environ., 278, doi:10.1016/j.rse.2022.113103, 2022.

Yan, X. and Niu, Z. G.: Preliminary study on wetland connectivity in Baiyangdian basin, Acta Ecol. Sin, 39, 11,
565   10.5846/stxb201809302127, 2019.

Yang, J., Huang, X., and Tang, Q.: Satellite-derived river width and its spatiotemporal patterns in China during 1990-2015, Remote Sens. Environ., 247, doi:10.1016/j.rse.2020.111918, 2020a.

Yang, X., Qin, Q., Yesou, H., Ledauphin, T., Koehl, M., Grussenmeyer, P., and Zhu, Z.: Monthly estimation of the surface water extent in France at a 10-m resolution using Sentinel-2 data, Remote Sens. Environ., 244, doi:10.1016/j.rse.2020.111803,
570   2020b.

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N. E., Ramoino, F., and Arino, O.: ESA WorldCover 10 m 2020 v100 [dataset], doi:10.5281/zenodo.5571936, 2021.

Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X., and Liu, L. Y.: GLC_FCS30D: the first global 30 m land-cover

575   dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-

series Landsat imagery and the continuous change-detection method, Earth Syst. Sci. Data, 16, 1353-1381, doi:10.5194/essd-

16-1353-2024, 2024.