

## **Response to The Comments from Reviewer #2**

This paper aims to produce a dataset of Chinese rivers spanning the period from 2016 to 2023 at an annual scale with a resolution of 10 m. However, the dataset lacks originality and has gaps in sufficient quality, and is limited in its potential for broader application, which I detail below:

### **Response:**

Thank you very much for your precious time to review our manuscript. We sincerely appreciate your diligent efforts to provide professional comments. Referring to your comments, we have revised our manuscript and provided detailed explanations for each comment. We acknowledge that our paper has limitations, but we believe that it still is a meaningful and interesting study. Our river algorithm is accurate, robust and effective, which can achieve lower misclassification and omission errors compared to using only the length-to-width ratio. Meanwhile, considerable manual editions were implemented for our river maps, which further improve their data quality. The characteristics of being national-scale, annually continuous, and having a 10-m spatial resolution make our river maps valuable for practical applications. Our detailed responses are as follows.

### **Originality:**

1) The classification of water body is more easily achievable compared to other land cover types in the field of remote sensing. It exhibits a significant spectral difference from other land cover types and has a relatively simple texture. Furthermore, it would be easy to screen rivers by simply using the length-to-width ratio of water bodies. However, the authors utilized publicly available 10 m land cover data and did not use an innovative scheme to extract rivers. They also fail to consider the network of rivers and the topographical features that influence the formation of rivers. The originality of the technical solution is limited.

### **Response:**

Thanks for your thorough and professional comments.

As mentioned in the review comments, we did not develop a new algorithm for water classification. We did so mainly consider three aspects: (1) As the comment pointed out, the spectral features of water bodies are simple and differ significantly with other land use types. To date, there are a lot of studies on remote sensing classification of water bodies, and the related algorithms for classifying waters are mature, accurate and effective (Zou et al., 2018; Pickens et al., 2020; Yang et al., 2020). Developing a new algorithm would hardly lead to improvements in water classification, mainly due to the limitations of remote sensing image classification capabilities and the complexity of geographic environments. (2) The Dynamic World (DW) and ESRI global land cover (EGLC) used in this study are land use datasets generated using deep learning algorithm. The broad land use

types (e.g. water, forest, farmland) in these two datasets could achieve high accuracy using deep learning algorithm, especially for the waters. For DW tiles with considerable gaps in invalid observations in 2016, we used multiple index water detection rule (MIWDR) to extract water bodies. This method was proposed by [Deng et al. \(2019\)](#), which exhibited good feasibility for long-term water classification at large scale. Meanwhile, existing research has demonstrated that deep learning, machine learning and water index methods all can obtain reliable results ([Li et al.,2022](#)). Thus, we believe that all three datasets are reliable water-related products, with high accuracy of waters. (3) Considering the availability of mature water-relate datasets, which have been validated for accuracy and reviewed by experts, we directly used these datasets to produce water maps. This approach is not only to avoid duplicating work, but also to save a significant amount of workload.

### References:

- [1] Zou Z, Xiao X, Dong J, et al. Divergent trends of open-surface water body area in the contiguous United States from 1984 to 2016[J]. *Proceedings of the National Academy of Sciences*, 2018, 115(15): 3810-3815.
- [2] Pickens, A. H., Hansen, M. C., Hancher, M., Stehman, S. V., Tyukavina, A., Potapov, P., Marroquin, B., & Sherani, Z. (2020). Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment*, 243(March), 111792.
- [3] Yang, X., Qin, Q., Yésou, H., Ledauphin, T., Koehl, M., Grussenmeyer, P., & Zhu, Z. (2020). Monthly estimation of the surface water extent in France at a 10-m resolution using Sentinel-2 data. *Remote Sensing of Environment*, 244(October 2019), 111803.
- [4] Deng, Y., Jiang, W., Tang, Z., Ling, Z., & Wu, Z. (2019). Long-term changes of open-surface water bodies in the Yangtze River Basin based on the google earth engine cloud platform. *Remote Sensing*, 11(19).
- [5] Li J, Ma R, Cao Z, et al. Satellite detection of surface water extent: A review of methodology[J]. *Water*, 2022, 14(7): 1148.

For our river extraction algorithm, it was accurate and effective for large-scale river mapping. On the one hand, it can accurately distinguish rivers from other water covers. On the other hand, the rules and thresholds used in our algorithm do not change over regions and time.

We believe our river mapping algorithm is an innovative algorithm, and our explanations is as follows: It is true that the length-to-width ratio can effective distinguish rivers from other water covers. However, using this single feature would result in considerable misclassification and omission errors. To illustrate this point, we counted length/width of rivers, lakes and reservoirs at different area groups. It was found that a threshold for the length/width set too low misclassifies

many lakes and reservoirs as rivers, whereas a threshold set too high leads to the omission of many rivers. To address this issue, we selected four geometric features—Compactness, Length/width, Roundness and Rectangular fit—based on literature review and experiments. To reduce omission errors, we first developed weak rules to extract as many as rivers as possible. Then, we combined these weak rules into strong rules to distinguish non-rivers such as lakes, reservoirs and aquaculture ponds, thereby reducing misclassification.

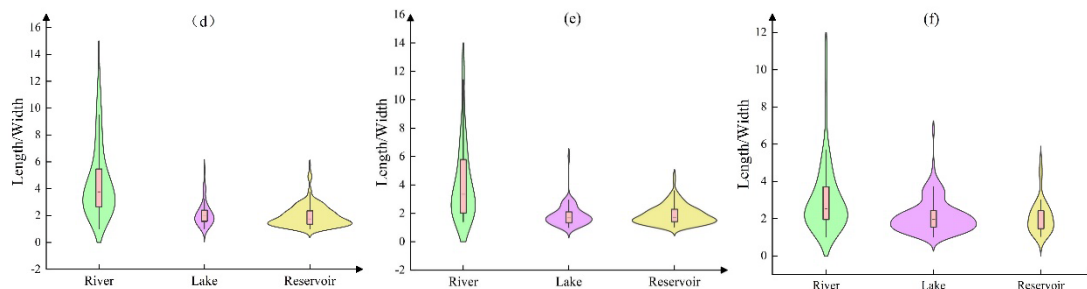


Figure T6. Violin plots and quartile box plots of length/width. The columns represent three area range groups of 0–1000 ha, 1000–5000 ha, and >5000 ha,

In addition to the computation of the river classification algorithm, considerable manual corrections were implemented to improve our river maps. There are three main errors need to be corrected: 1) Some aquaculture ponds were spatially connected with rivers. (2) The identified reservoirs may contain some rivers. (3) Small rivers with narrow channels cannot be well identified via the river rule set. We provided a more detailed description of post-processing improvement in Section 3.3. We also implemented considerable manual corrections. Therefore, our study is not a work that can be realized easily and has good potential in scientific research and practical applications.

2) As shown below, the authors did not acknowledge many relevant river datasets in the text. This makes me seriously concerned about the proper place for this paper.

**Response:**

Thanks for your careful review and offering professional comments.

As the comments mentioned, we indeed missed some typical and representative datasets in our literature review, which is a deficit of our paper. The three recommend datasets are wonderful global river datasets. We have added these datasets to our literature review in Section 5.3. Thanks again for your helpful comments. Additionally, we have re-searched the literatures in Web of Science platform, and added three additional global river datasets in our manuscript.

[1] Lin, P., Pan, M., Wood, E. F., Yamazaki, D., & Allen, G. H. (2021). A new vector-based global river network dataset accounting for variable drainage density. *Scientific data*, 8(1), 28.

[2] Nyberg, B., Sayre, R., & Lujendijk, E. (2024). Increasing seasonal variation in the extent of

- rivers and lakes from 1984 to 2022. *Hydrology and Earth System Sciences*, 28(7), 1653-1663.
- [3] Yan, D., Wang, K., Qin, T., Weng, B., Wang, H., Bi, W., et al. (2019). A data set of global river networks and corresponding water resources zones divisions. *Scientific data*, 6(1), 219.
- [4] Yamazaki D, Ikeshima D, Sosa J, et al. MERIT Hydro: A high - resolution global hydrography map based on latest topography dataset[J]. *Water Resources Research*, 2019, 55(6): 5053-5073.
- [5] Yan D, Li C, Zhang X, et al. A data set of global river networks and corresponding water resources zones divisions v2[J]. *Scientific Data*, 2022, 9(1): 770.
- [6] Altenau E H, Pavelsky T M, Durand M T, et al. The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A global river network for satellite data products[J]. *Water Resources Research*, 2021, 57(7): e2021WR030054.

420

Table 4. 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)
	Vector-based global river network	Global	90 m	---	(Lin et al., 2021)
	Global River Network	Global	90 m	---	(Yan et al., 2019)
	MERIT Hydro	Global	90 m	---	(Yamazaki et al., 2019)
	SWOT River Database	Global	100	---	(Altenau et al., 2021)
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	---	---	<a href="https://download.geofabrik.de/asia/china.html">https://download.geofabrik.de/asia/china.html</a>
	Global River and Lake Extent	Global	30 m	1984-2022	(Nyberg et al., 2024)

3) Besides, the work of Allen & Pavelsky (2018) has been cited, but the differences from the data of this article have not been explained.

**Response:**

Thanks for your useful comments. To demonstrate the differences between GRWL and our CRED, we conducted additional experiments to perform a comparative analysis of the two datasets.

For GRWL and the CRED in 2018, the river areas were 25583.15 km<sup>2</sup> and 5186223 km<sup>2</sup>, respectively. The river area in GRWL was significantly smaller than that in the CRED, which was also observed in Figure T7. The area difference of these two datasets was mainly due to the difference of river mapping methods. The GRWL only mapped rivers with width greater than 90 m, as it considered river extracted from Landsat imagery to be reliable only when their width exceeded 90 m. In contrast, the CRED mapped more narrow rivers, due to its high spatial resolution.

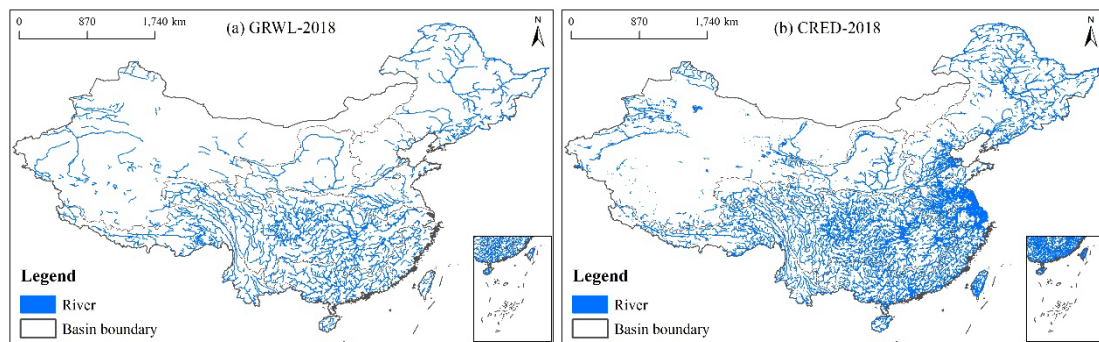


Figure T7. River patterns of GRWL and CRED in 2018

To quantify the consistency between GRWL and CRED, river fractions within a 0.05° by 0.05° grid were counted. It was found that the two datasets had moderate correlation coefficient, with R value of 0.577. This indicated that two datasets shared considerable consistency but still exhibited significant differences. Due to the significant difference in areas and spatial distribution between two datasets, we will no longer added the data inter-comparison between GRWL and CRED in the revised manuscript.

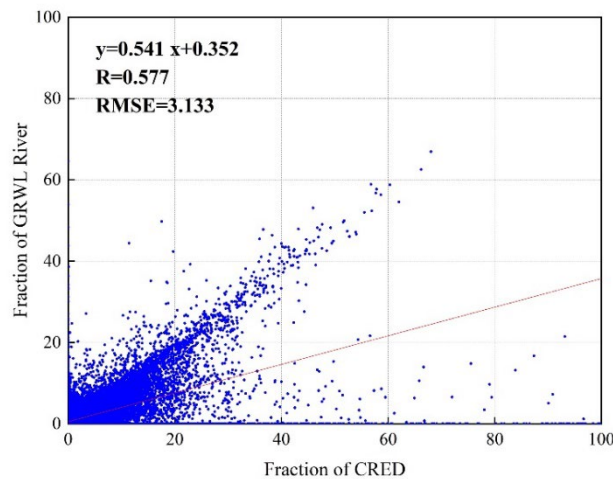


Fig. T8. Scatterplots of the river fraction between the CRED and GRWL

### Scientific quality:

1) The scheme of data validation has considerable uncertainty. The river is typically characterized by the property of network morphology. However, the generated river data display a substantial number of river discontinuities. There is a conspicuous phenomenon of rivers in adjacent years either "disappearing" or "breaking off" noticeably. Although the accuracy is about 95% by visual interpretation, this is based on a pixel-by-pixel basis and does not take into account the connectivity of rivers. Additionally, the visual interpretation is also a highly subjective process. If only the center of the river was selected, the accuracy of the river would be overestimated.

### Response:

Thanks for your valuable comments and careful inspection for our dataset.

The rivers are fluid linear waters with varied extents. Their surfaces area changes are affected by multiple factors, such as seasonal variation, climate changes and human activities. The CRED just mapped the actual extent of rivers in each year based on Sentinel-2 imagery. For the conspicuous phenomenon of rivers in adjacent years, most of them are attributed to natural factors and human activities. Meanwhile, there are still rivers inconsistencies between different years, which could be attributed to two aspects: (1) Some lakes, reservoirs and aquaculture ponds are spatially connected with rivers. For example, in the Poyang Lake Basin, lakes and rivers are spatially connected (Figure T9(a)). It is necessary to implement interrupt operation and delete lakes (Figure T9 (b)). However, since the interruptions are not strictly at the same location, this may lead to some inconsistencies in the rivers between years. (2) Many narrow rivers are spatially discontinuous, due to the limitation of 10 m spatial resolution. In our study, we made our best to manually these rivers. However, the corrections varied between different years, resulting in certain discrepancies in the CRED across years. To illustrate this limitation, we added corresponding descriptions in the Discussion section.

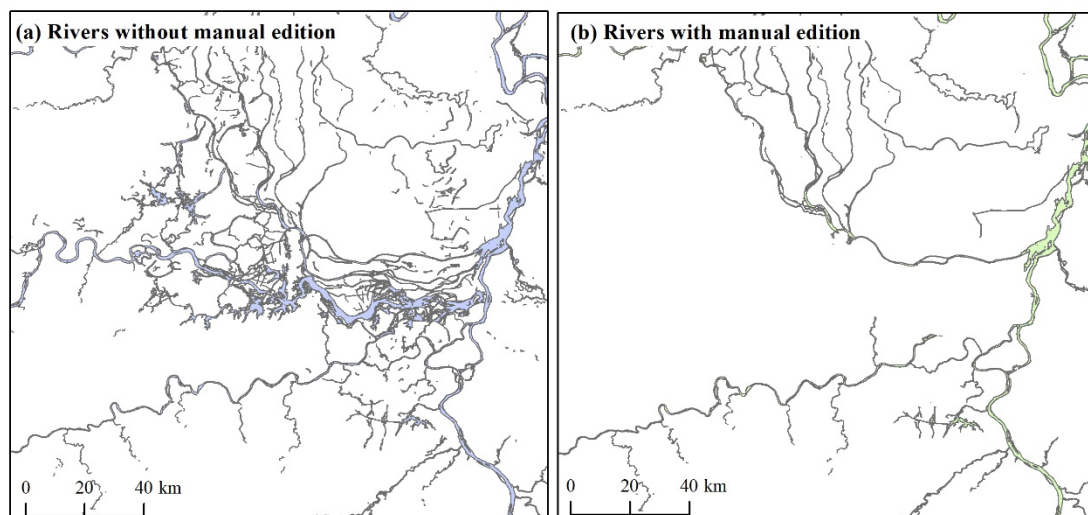




Figure T9. Spatial distribution of rivers without (a) and with (b) manual edition in 2023

smaller than that in other years. In addition, the extent and completeness of manual corrections varied across different years, leading to some inconsistencies of rivers in adjacent years. The above characteristics of the three datasets may lead to errors

In our accuracy validation, test samples were produced based on stratified random sampling and visual interpretation. On the one hand, the locations of test samples were not manually selected but were randomly generated. We did not deliberately produce river samples at the center of rivers. On the other hand, the attributions (river or non-river) were determined based on visual interpretation. This approach ensures both a reasonable sample distribution and accurate sample attributions. We have made more detailed descriptions for the production of test samples.

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 of CRED. 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. We visually labelled these samples based on imagery from different year from 2016 to 2023. Correspondingly, their attributions may vary across different years, although their locations remain unchanged. Using these three platforms, river and non-river samples from 2016 to 2023 were produced.

2) The key data utilized in this paper (i.e., European Space Agency and Dynamic World) are 10 m land-cover classification products. They were not primarily designed for water classification. These products tend to underestimate the area of water body, and consequently, the extent of rivers.

**Response:**

Thanks for your valuable comments.

To date, there are many global or national water datasets. However, the long-term and 10 m spatial resolution at national or global scale still remain scarce. The DW dataset is a real-time updated land use dataset since 2015. Its high temporal resolution allows it to effectively capture river dynamics in rivers, while its 10 m spatial resolution enables accurate extraction of rivers, as well as narrow rivers.

The DW datasets were produced using deep learning algorithms, and their reliability has been validated by accuracy assessment and peer-reviewed. Additionally, Li et al (2022) indicated that deep learning algorithms, machine learning algorithms and water index threshold algorithms all achieve good accuracy in water classification. Although the DW were not primarily designed for water classification, rivers in the DW also achieved high accuracy, making them reliable for river

mappings. It is true that the use of DW land use data for river extraction may cause uncertainties. We have illustrated this point in the Discussion section.

465 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 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

**Reference:**

[1] Li J, Ma R, Cao Z, et al. Satellite detection of surface water extent: A review of methodology[J]. Water, 2022, 14(7): 1148.

3) Rivers possess highly pronounced seasonal characteristics. During the summer flood season, rivers become wider, while in winter, they may even disappear. The specific meaning and significance of annual-scale river data remains unclear.

**Response:**

Thanks for your professional comments.

How to define rivers is a challenging task. During the wet season, the extents of river inundation is large, part of which may consist of floodwaters. During the dry season, the extents of river is small, which underestimate the actual extent of rivers. To balance this contradiction, we map river extents using the mode algorithm based on water time series with one year. For a pixel of water, if most of values in its time series are water, the pixel is labeled as yearly water. In our study, the yearly water of rivers is defined as river extents. To explicitly describe the definition of rivers, we made supplements in our manuscript.

value of this pixel is 0. In this way, annual water maps were created via the mode algorithm. For rivers, the composited water extents were defined as yearly river extents. Additionally, considering the scarcity of DW images before 2016, yearly water

4) "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." This strategy is subjective. As depicted in Figure 2, the EGLC data only encompasses the period from 2017 to 2023. In contrast, for the remaining years of 2015 - 2016, classification is carried out using the land cover data that was self-produced. Why use the DW dataset as the primary data of river extraction? How to ensure consistency across datasets? The experimental scheme also has a certain degree of subjectivity.

**Response:**

Thanks for your constructive comments. This is very professional comments.

The Dynamic World (DW) is a 10m near-real-time land use datasets. Normally, the revision of



DW is 3-5 days. This high-frequency is beneficial for capturing river dynamics. Thus, we used the DW as primary data to extract rivers. However, the DW dataset was produced using Sentinel-2 images with cloud coverage less than 35%. This strategy leads to sparse or even absent valid observations in areas severely affected by cloud contaminations.

The following figure counts the number of good observations of 2023 DW in each pixel. Results indicated that there are fewer or missing good observation pixels in Southwest China. To solve this issue, we used EGLC dataset to replace DW dataset in regions with spare or missing observations. However, the time range of EGLC is 2017-2023. To completely produce river maps in 2016, the Sentinel-2 images were used to extract waters in areas with spare or invalid observations of DW. Considering the limited availability of Sentinel-2 images before 2017, all Sentinel-2 images from June 2015 to December 2016 were utilized for water classification.

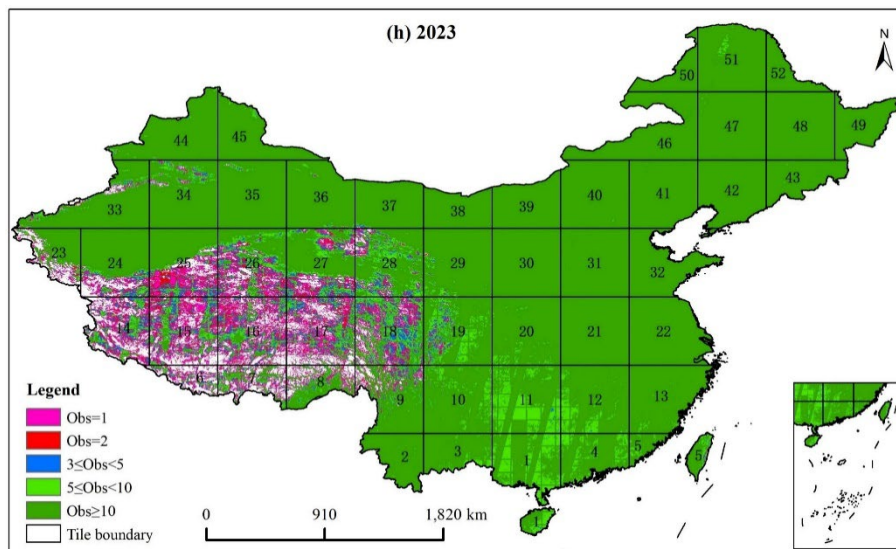


Figure T10. Valid observations for individual pixels in DW image of 2023. This figure was also displayed in Supplementary material

To explicitly illustrate the strategy for the selection of three datasets, we have added supplementary explanations in the manuscript. In addition, we agree that using three different data sources to produce river maps would introduce uncertainties. For instance, river extents may vary in same location and time when derived from different datasets. This is a limitation of our study, which we have described in the Discussion section.

90 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. Sentinel-2 images used in DW dataset may have cloud contaminations, which also resulted in missing observations. Thus, the complete coverage of the entire China to produce a complete China river map cannot be achieved by using DW dataset alone.

To facilitate image computation, we divided China into 52 tiles using a  $5^{\circ} \times 5^{\circ}$  grid (Fig. S1). If a tile contains considerable 95 pixels with fewer than 3 valid observations, it was defined as a data insufficient tile. For these tiles, 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 100 water classification.

465 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 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

## Application

1) As illustrated in Fig. 8, the river data produced in this paper is significantly different from that of other products. In practical applications, it is relatively difficult for users to make a trade-off regarding which one to use.

### Response:

Thanks for your careful and offering professional comments.

In our study, we check the quality of river datasets based on spatial comparison by visual inspection. It was found that, for EA\_Wetlands, it mapped more rivers compared with CRED. However, it contained considerable misclassifications. For example, the Poyang Lake that spatially connected with Yangtze River was labeled as rivers. Aquaculture ponds that connected with rivers were labeled as rivers. For CWaC, many small patches, such as snow/ice, mountain shadows, and other small waters, were misclassified as rivers. For CNLUCC, it failed to identify narrow rivers due to the 30m spatial resolution limitation. In contrast, our CRED not only accurately rivers with fewer misclassifications, but also effectively identified narrow rivers. Based on above analysis, we conclude that CRED not only accurately mapped river extents in China, but also reflect temporal changes of rivers. Therefore, CRED can serve as reliable dataset for river-related applications.

We implemented data inter-comparison for two purposes. On the one hand, it is to show that our data have good agreement with existing river-related datasets. On the other hand, spatial inspections were implemented to analysis the strengths and weakness of different datasets, which

could highlight the superiority of the CRED.

In our study, we did not implement strict accuracy validation for different river datasets, and cannot quantitatively assess the quality of river datasets. This is an insufficient of our work. We illustrated this point in the Discussion section.

narrow streams are extracted in CRED but not in CNLUCC. Meanwhile, we did not use test samples to assess rivers' accuracy of CNLUCC, [CWaC](#) and [EA Wetlands](#), which cannot quantitatively evaluate the accuracy superiority of CRED compared to other river datasets. In addition, our CRED maps were national-scale dataset, and its applicability was not as extensive as

2) This is not a global product. It has relatively limited application potential compared with other global river products.

**Response:**

Thanks for your valuable comments.

It is true that our CRED is not global river dataset, which to some extent limits its widespread application. However, compared with other large-scale river datasets, our CRED has its own advantages. First, our algorithm is accurate and efficient for river extracted, and the identified rivers have small errors in omission and misclassifications. Second, extensive post-processing was conducted on the algorithm-extracted rivers. Aquaculture ponds connected with rivers were manually interrupted and then removed. Lakes and reservoirs connected with rivers were also manually corrected. Narrow rivers with spatial discontinuous patches were manually corrected to map rivers as much as possible. Thus, river boundaries in the CRED are accurate and complete, and the CRED data has good application potentials in China.

Considering the huge workload and our limited time, we did not produce global river dataset. We acknowledge that the river datasets at the national scale has limited applicability compared to global-scale data. This point is illustrated in the discussion section.

other river datasets. In addition, our CRED maps were national-scale dataset, and its applicability was not as extensive as global-scale dataset. ′