Dear editors and reviewers,

We sincerely appreciate your constructive comments and suggestions to improve this manuscript. We have revised the manuscript and addressed these comments point by point. We hope that this revised manuscript fulfills the editor's and reviewers' high standards for the *Earth System Science Data*.

The reviewers' comments are shown in **black**, our responses are highlighted in **blue**, and the revised text in the manuscript is highlighted in **orange**.

We look forward to your feedback.

Yours sincerely,

Yao Li

Reviewer 1

Comments to the Author

This paper presents a river discharge dataset for Chinese rivers. Discharge is estimated by exploiting the correlation between river width as an observable variable from satellite EO and discharge. The predictive relationships (rating curves) are estimated using in-situ river discharge observations from a national database. These approaches are well known and have been used in many previous studies, both at regional and global scale. Authors argue that in earlier studies, only few estimation points (or gauges) have been included from China and that their dataset increases data availability for China significantly. However, I believe this is mainly a consequence of access restrictions for in-situ river gauging data that China imposes on both international and national researchers. For this reason, I see the in-situ dataset as the main contribution of this paper. Authors must ensure that this dataset is accessible for everyone in the public domain before this paper can be published. This is the only way to ensure that other people can benefit from this work and that the authors' work can be benchmarked and compared with alternative approaches.

Response: We sincerely appreciate the reviewer's valuable comments and constructive suggestions. We fully understand the reviewer's concern regarding the accessibility of in situ discharge observations and recognize their critical importance to the hydrologic community. We also appreciate the reviewer's acknowledgment of our work's contribution to improving the availability of runoff observations in China. Here, we would like to clarify the core scientific contributions of this study and their broader implications for the hydrologic community, while providing a formal explanation of the objective limitations on data sharing.

We would like to emphasize that the primary scientific contribution of this study is not merely the compilation of field-observed data, but the construction of the Chinese Daily River Discharge Record (CDR^2) dataset through multi-source satellite-based reconstruction and physically constrained modeling. Although the compiled in situ dataset surpasses existing public databases (e.g., GRDC) in gauge density and temporal resolution, its temporal coverage spans only 2014–2024 and includes non-negligible data gaps (Fig. 2). The key advance lies in extending discharge records to 1990 for 301 gauges by combining multi-satellite-derived river width time series with gauge-specific hydraulic geometry relationships, under strict statistical testing (Spearman rank correlation ≥ 0.25) and physical constraints (hydraulic geometry width exponent between 0 and 0.59). Sacrificing a portion of potentially reconstructable gauges ensures the reliability of the reconstructed discharge records. CDR^2 achieves a median Kling–Gupta Efficiency (KGE) of 0.66 in validation, outperforming existing global satellite remote sensing products, while increasing the number of reconstructed and extended gauges in China by at least a factor of five compared to current global products (Elmi et al., 2024; Lin et al., 2023; Riggs et al., 2023; Saemian et al., 2024). The CDR^2 dataset, along with its complete quality assessment information, has been publicly released on Zenodo at https://doi.org/10.5281/zenodo.20152832 (Wang and Li, 2026). In addition to reconstructed discharge records for 310 gauges, the dataset includes metadata for 1,196 gauges, river width time series, hydraulic geometry parameters, and reconstruction accuracy metrics for each gauge. These data are freely accessible, independently verifiable, and reusable by the research community, ensuring transparency in data sources, reconstruction procedures, and the quality of results. Based on these openly available data, CDR^2 not only provides essential baseline data for cross-validation of satellite-derived discharge products and calibration of hydrological models over China, but also supports research on long-term discharge changes and their hydrological response mechanisms in the region.

Regarding the accessibility of in situ discharge data, we provide the following explanation. The observational data underlying this study were sourced from the National Hydrological and Rainfall Information System of the Ministry of Water Resources of China (http://xxfb.mwr.cn/). This platform provides a dynamic web interface for real-time discharge visualization but offers no download portal for historical archives. Consequently, our dataset was not extracted from an existing database; rather, it was compiled through continuous, long-term web scraping and rigorous cleaning of publicly displayed daily data via a custom automated retrieval program. Due to national data confidentiality policies and regulations, we regret that we are unable to directly release the original gauge records as an open-source dataset. This situation echoes the challenges faced by similar international datasets. For instance, the China Hydrological Project (CHP) dataset (Henck et al., 2010; Schmidt et al., 2011), although used in global river discharge compilations (Riggs et al., 2023) and publicly accessible via the RivRetrieve software package (Riggs et al., 2024), explicitly states that it does not include gauges within China due to the same data policy constraints. Nevertheless, we fully recognize the value of these fundamental observation data in advancing scientific progress and firmly uphold the principle of reproducibility in scientific research. To this end, we warmly welcome researchers who need access to these data to contact our team through collaborative research arrangements. At the same time, we will actively evaluate the possibility of publicly releasing the original observation data within the scope permitted by future policies.

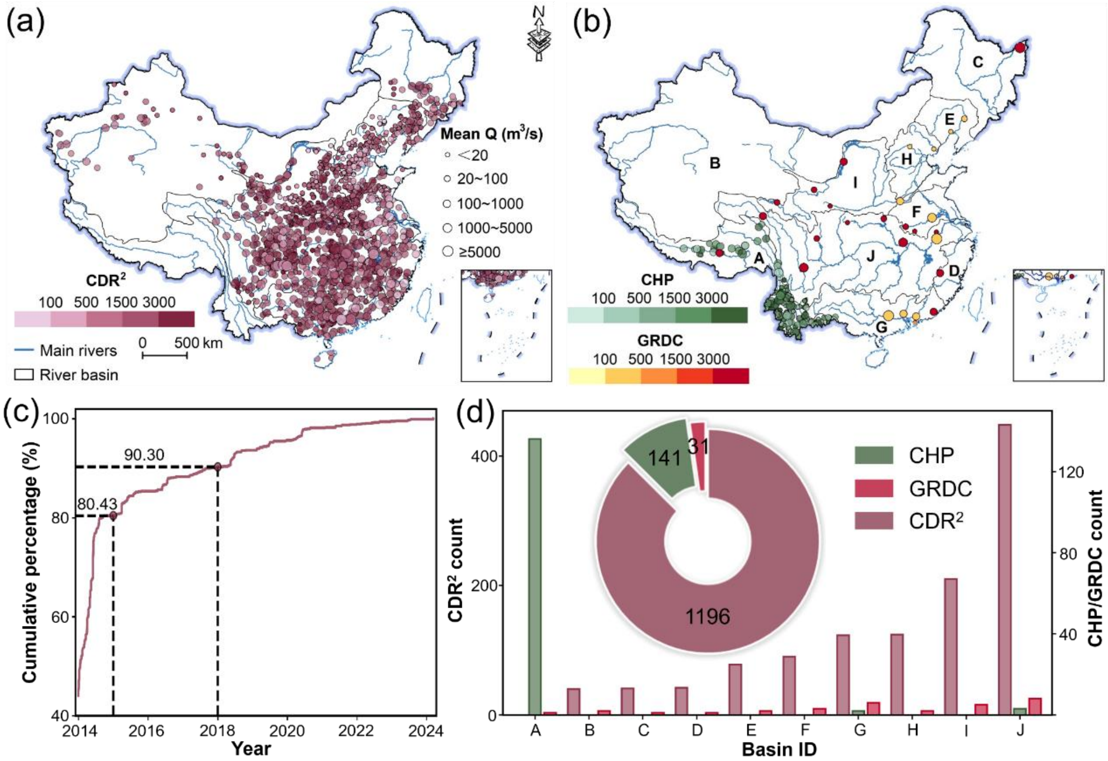

Figure 2. Spatial distribution and temporal coverage of in situ river discharge observations across China (2014–2024). **(a)** Spatial distribution of the 1,196 gauges included in the China Daily River Discharge Records (CDR²) dataset, with symbol size representing mean discharge (Q, m³/s) and color indicating the number of discharge records at each gauge. **(b)** Spatial distribution of river gauges available from the Chinese Hydrology Project (CHP) and the Global Runoff Data Centre (GRDC) within China, with symbol size representing mean discharge (Q, m³/s) and color indicating the number of discharge records at each gauge. The labeled regions (A–J) denote the ten major basins: Southwest Rivers (A), Northwest Rivers (B), Songhua River (C), Southeast Rivers (D), Liao River (E), Hai River (F), Pearl River (G), Huai River (H), Yellow River (I), and Yangtze River (J). **(c)** Cumulative percentage distribution of gauge observation start years in the CDR², indicating that over 80% of gauges contain records beginning in or before 2015, while more than 90% begin in or before 2018. **(d)** Comparison of the number of gauges included in the CDR², CHP, and GRDC datasets across China. (Page 7, Lines 154–166)

References

Elmi, O., Tourian, M. J., Saemian, P., and Sneeuw, N.: Remote sensing-based extension of GRDC Discharge time series – a monthly product with uncertainty estimates, Sci. Data, 11, 240, https://doi.org/10.1038/s41597-024-03078-6, 2024.

Henck, A. C., Montgomery, D. R., Huntington, K. W., and Liang, C.: Monsoon control of effective discharge, Yunnan and Tibet, Geology, 38, 975–978, https://doi.org/10.1130/G31444.1, 2010.

Lin, P., Feng, D., Gleason, C. J., Pan, M., Brinkerhoff, C. B., Yang, X., Beck, H. E., and Frasson, R. P. d. M.: Inversion of river discharge from remotely sensed river widths: A critical assessment at three-thousand global river gauges, Remote Sens. Environ., 287, 113489, https://doi.org/10.1016/j.rse.2023.113489, 2023.

Riggs, R. M., Allen, G. H., Wang, J., Pavelsky, T. M., Gleason, C. J., David, C. H., and Durand, M.: Extending global river gauge records using satellite observations, Environ. Res. Lett., 18, 064027, https://doi.org/10.1088/1748-9326/acd407, 2023.

Riggs, R., Moulds, S., Wortmann, M., Slater, L., and Allen, G.: RivRetrieve: Retrieve Global River Gauge Data, CRAN, https://doi.org/10.32614/CRAN.package. RivRetrieve, 2024.

Saemian, P., Elmi, O., Stroud, M., Riggs, R., Kitambo, B. M., Papa, F., Allen, G. H., and Tourian, M. J.: Satellite Altimetry-based Extension of global-scale in situ river discharge Measurements (SAEM), Earth Syst. Sci. Data, 17, 2063–2081, https://doi.org/10.5194/essd-17-2063-2025, 2025.

Schmidt, A. H., Montgomery, D. R., Huntington, K. W., and Liang, C.: The question of communist land degradation: New evidence from local erosion and basin-wide sediment yield in Southwest China and Southeast Tibet, Ann. Assoc. Am. Geogr., 101, 477–496, https://doi.org/10.1080/00045608.2011.560059, 2011.

Wang, Y. and Li, Y.: China Daily River Discharge Records (CDR²) dataset [Data set], Zenodo, https://doi.org/10.5281/zenodo.20152832, 2026.

1. Authors have decided to use river width as the observable satellite EO proxy for river discharge. However, several other proxies have been used in the literature, most importantly water surface elevation from satellite altimetry. Can authors discuss the pros and cons of using river width as compared to the other proxies individually or combined in a data fusion approach, and motivate/justify their choice?

Response 1: We sincerely thank the reviewer for this thoughtful comment, which prompts us to provide a more explicit justification for our choice of river width as the primary satellite proxy. We fully agree that the relative merits of different hydraulic observables—particularly river width versus water surface elevation—deserve clear articulation in the manuscript. In response, we have added the following text to the Introduction: "River width and water surface elevation are complementary hydraulic observables for satellite-based river discharge estimation. Unlike satellite altimetry, which samples rivers only along discrete ground tracks and provides one-dimensional profile measurements, optical remote sensing acquires continuous two-dimensional swath observations through wide-swath imaging, enabling much broader spatial coverage of river networks (Alsdorf et al., 2007; Lettenmaier et al., 2015; Tourian et al., 2022). In addition, the long-term continuity of optical satellite archives provides a viable means to reconstruct historical discharge records over multi-decadal periods. Therefore, when aiming to maximize both the spatial density of reconstructible gauges and the temporal coverage of discharge records, river width represents a more suitable remote sensing proxy." (Page 3, Lines 67–76)

Beyond the discussion presented in the manuscript, we further clarify the practical considerations behind this methodological choice. Based on our experience processing satellite altimetry data over inland waters, we find that the effective record length of high-quality water-level observations remains relatively limited. Although earlier altimetry missions, such as Jason-1, provide longer temporal coverage dating back to 2001, their measurement accuracy over inland waters is generally insufficient for reliable discharge reconstruction. More recent missions, such as CryoSat-2, SARAL/AltiKa, and Sentinel-3, have adopted improved inland-dedicated retracking algorithms that substantially enhance measurement quality. However, according to our experience, observations suitable for reliable reconstruction are primarily available from 2016 onward, resulting in a relatively short record length. In contrast, the Landsat archive has provided continuous optical observations since 1984, enabling reconstructions spanning multiple decades. Furthermore, the river sampling frequency of nadir altimeters poses another critical constraint. This study requires a minimum of ten same-day concurrent satellite observations and in situ discharge measurements to establish reliable cross-sectional stage–discharge relationships. However, satellite altimeter revisit intervals substantially exceed those of optical missions. Radar altimeters typically cycle every 10–369 days; similarly, ICESat/ICESat-2 laser altimeters—despite their exceptional along-track spatial resolution (17–70 m footprints and a 0.7 m along-track sampling interval)—operate on an approximate 91-day revisit period. Such infrequent observation severely restricts the acquisition of valid synchronous samples. Consequently, given our objective of reconstructing the maximum number of gauges over the longest possible time span, river width is a more suitable proxy for this study.

References

Alsdorf, D. E., Rodríguez, E., Lettenmaier, D. P.: Measuring surface water from space, Rev. Geophys., 45, RG2002, https://doi.org/10.1029/2006RG000197, 2007.

Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., and Wood, E. F.: Inroads of remote sensing into hydrologic science during the WRR era, Water Resour. Res., 51, 7309–7342, https://doi.org/10.1002/2015WR017616, 2015.

Tourian, M. J., Elmi, O., Shafaghi, Y., Behnia, S., Saemian, P., Schlesinger, R., and Sneeuw, N.: HydroSat: Geometric quantities of the global water cycle from geodetic satellites, Earth Syst. Sci. Data, 14, 2463–2486, https://doi.org/10.5194/essd-14-2463-2022, 2022.

2. The entire workflow hinges on the availability of in-situ discharge at stations. What authors present is not a truly contactless “satellite-only” approach. It is thus essential to share the dataset of in-situ discharge in the public domain to ensure that others can reproduce the work and benchmark it against their own approaches. Given the access limitations for Chinese in-situ discharge, publishing the in-situ data in itself has significant value. Authors refer to a website in the manuscript where the in-situ data supposedly can be downloaded. I made a few checks and tried to navigate this Chinese page, but was unable to actually access in-situ river discharge time series. Authors should either provide detailed guidance as to how the data can be accessed from the website or re-publish it in an open archive using science-friendly data formats (e.g. netcdf).

Response 2: We sincerely thank the reviewer for this constructive comment and for carefully testing the data access pathway described in our manuscript. First, we fully acknowledge the reviewer’s perspective that the in situ dataset used in this study is of significant value. However, as we emphasized in our previous response, the core scientific contribution of this study is the construction of the Chinese Daily River Discharge Record (CDR^2)—a satellite-reconstructed and physically constrained discharge dataset—rather than the original gauge observations alone. The CDR^2 dataset (including reconstructed daily discharge records for 310 gauges, river width time series, hydraulic geometry parameters, and reconstruction accuracy metrics) has been publicly released on Zenodo (https://doi.org/10.5281/zenodo.20152832) to ensure full reproducibility and independent verification. We believe this openly available dataset already serves the scientific community's need to benchmark and compare our results with alternative approaches.

Second, the inability to retrieve historical in situ river discharge time series from the National Hydrological and Rainfall Information System (http://xxfb.mwr.cn/) is inherent to the platform's operational design. Serving strictly as a real-time operational monitoring terminal with a Chinese-only interface, the platform provides dynamic visualization of daily hydrological updates but lacks public interfaces for historical data retrieval, retrospective querying, or bulk downloading (Fig. R1). Notably, the website underwent a redesign in July 2024, after which it ceased providing real-time river discharge data and now displays only water level data. This explains why checking the platform would only reveal water-level dynamics, not discharge. Given these constraints, the long-term historical dataset used in this study was not obtained through direct download but was compiled by running custom web-scraping scripts daily over multiple years, followed by automated quality control and cleaning of the instantaneous real-time data. Regarding data sharing, as we have noted, national data confidentiality policies and regulations preclude the direct release of raw in situ measurements as an open-source dataset. Nevertheless, we fully endorse the principle of scientific reproducibility. Researchers seeking access to these raw data are warmly welcome to contact our group through collaborative research arrangements. We will also actively evaluate the possibility of making the observational data publicly available as future policies permit. In addition, we have explicitly clarified this in the Data Availability section of the manuscript: "The in situ daily river discharge observations were obtained from the National Hydrological and Rainfall Information System of the Ministry of Water Resources of China (http://xxfb.mwr.cn/), a platform accessible only within China." (Page 31, Lines 686–688)

Station 站名	River 河名	Basin 流域	Stage (m) 水位（m）	Time 时间	超警戒水位 (m)
民和	湟水	黄河	1761.76	07-03 08:00	--
泉眼山(二)	清水河	黄河	1204.63	07-03 08:00	--
赵城	汾河	黄河	472.92	07-03 08:00	--
河津	汾河	黄河	371.44	07-03 08:00	--
北道	渭河	黄河	1073.51	07-03 08:00	--
魏家堡	渭河	黄河	486.54	07-03 08:00	--
咸阳	渭河	黄河	373.77	07-03 08:00	--
临潼	渭河	黄河	350.59	07-03 08:00	--
华县	渭河	黄河	338.25	07-03 08:00	--
张家山	泾河	黄河	419.53	07-03 08:00	--
状头	北洛河	黄河	361.19	07-03 08:00	--
戴村坝	大汶河	黄河	40.64	07-03 08:00	--
黑石关	伊洛河	黄河	108.13	07-03 08:00	--
龙门镇	伊河	黄河	146.99	07-03 08:00	--
白马寺	洛河	黄河	112.63	07-03 08:00	--

Figure R1. Interface of the real-time visualization terminal from the National Hydrological and Rainfall Information System. The screenshot demonstrates the platform's core functionality, which is restricted to displaying single-day, transient hydrological dynamics for national key river stations. Red text has been overlaid on the original Chinese-only interface for description.

3. The authors' river discharge reconstruction resulted in a median KGE of 0.66. While this is useful, it is not stellar for a rating-curve based approach using in-situ data, especially for rivers with a highly seasonal flow regime such as in China. It would thus be interesting to benchmark the authors' results using other discharge retrieval techniques developed by other groups.

Response 3: We sincerely thank the reviewer for this constructive suggestion. We fully agree that benchmarking our results against other discharge retrieval techniques is an essential step to contextualize the performance of CDR^2. In fact, this is exactly what we have already done in Section "4.1 Comparison of reconstructed discharge records with existing datasets" of the manuscript (Page 24, Line 531). We compared CDR^2 with four global satellite-derived river discharge products that employ different methodologies—height-based, width-based, and hybrid approaches. These include RSEG by Elmi et al. (2024), SAEM by Saemian et al. (2024), as well as the products by Lin et al. (2023) and Riggs et al. (2023). The comparison shows that CDR^2 achieves the highest median KGE (0.66) among these products, which exhibit a global median range of 0.33–0.48.

4.1 Comparison of reconstructed discharge records with existing datasets

To better contextualize the contribution of the CDR^2 dataset, we conducted a comparative assessment against several representative global remotely sensed river discharge products, including RSEG (Elmi et al., 2024), SAEM (Saemian et al., 2024), and the datasets developed by Lin et al. (2023) and Riggs et al. (2023). Although these datasets have substantially expanded the spatial coverage of global river discharge observations, their representation within China remains extremely limited (Fig. 10 and Table 2). Specifically, the number of gauges in China ranges from only one in Riggs et al. (2023) to 62 in SAEM. In all cases, Chinese gauges account for less than 1% of the total stations included in these global products, reflecting both the scarcity of publicly accessible in situ discharge observations and the challenges of satellite-based discharge retrieval in this region. In contrast, CDR^2 expands the number of gauge-constrained discharge records available for China by approximately 5–300 times relative to existing global datasets, providing discharge estimates at daily temporal resolution for 310 gauges distributed across diverse hydroclimatic and geomorphological settings (Fig. 10). The dataset encompasses rivers spanning a wide range of discharge magnitudes and channel sizes, thereby substantially improving the representation of China within global river discharge archives.

Beyond its improved spatial coverage, CDR^2 also demonstrates clear advantages in temporal resolution and predictive performance. While RSEG and Lin et al. (2023) provide monthly discharge estimates, CDR^2, SAEM, and Riggs et al. (2023) operate at daily time scales. Notably, CDR^2 achieves a median KGE of 0.66, exceeding the reported global median range of 0.33–0.48 among existing products (Table 2), although these comparisons should be interpreted considering differences in gauge selection criteria and river characteristics. Consistent improvements are also evident in other performance metrics, including a lower median NRMSE (12.7%) and a higher median CC (0.80), indicating enhanced capability in reproducing both the magnitude and temporal variability of observed discharge.

These improvements are partly attributed to the rigorous screening strategy adopted in this study, whereby only gauges exhibiting stable width–discharge relationships and physically consistent hydraulic behavior were retained (Fig. 5). In contrast, global-scale products often face an inherent trade-off between maximizing spatial coverage and maintaining retrieval quality, frequently incorporating gauges with weaker hydraulic relationships or unstable channel responses, thereby introducing greater uncertainty into discharge reconstruction. Moreover, by integrating observations from both Landsat and Sentinel-2 imagery, CDR^2 achieves a substantially higher effective sampling frequency than single-sensor approaches. This enhanced temporal sampling improves the detection of short-term discharge fluctuations and mitigates the smoothing effects associated with monthly aggregation, as observed in RSEG and Lin et al. (2023), thereby enabling a more realistic reconstruction of discharge dynamics at daily temporal resolution. Overall, CDR^2 provides an expanded and higher-resolution characterization of river discharge across China, a hydrologically critical yet globally underrepresented region. Through its greatly expanded gauge density, discharge records at daily resolution, and enhanced reconstruction performance, CDR^2 fills a major geographic gap in existing global river discharge products.

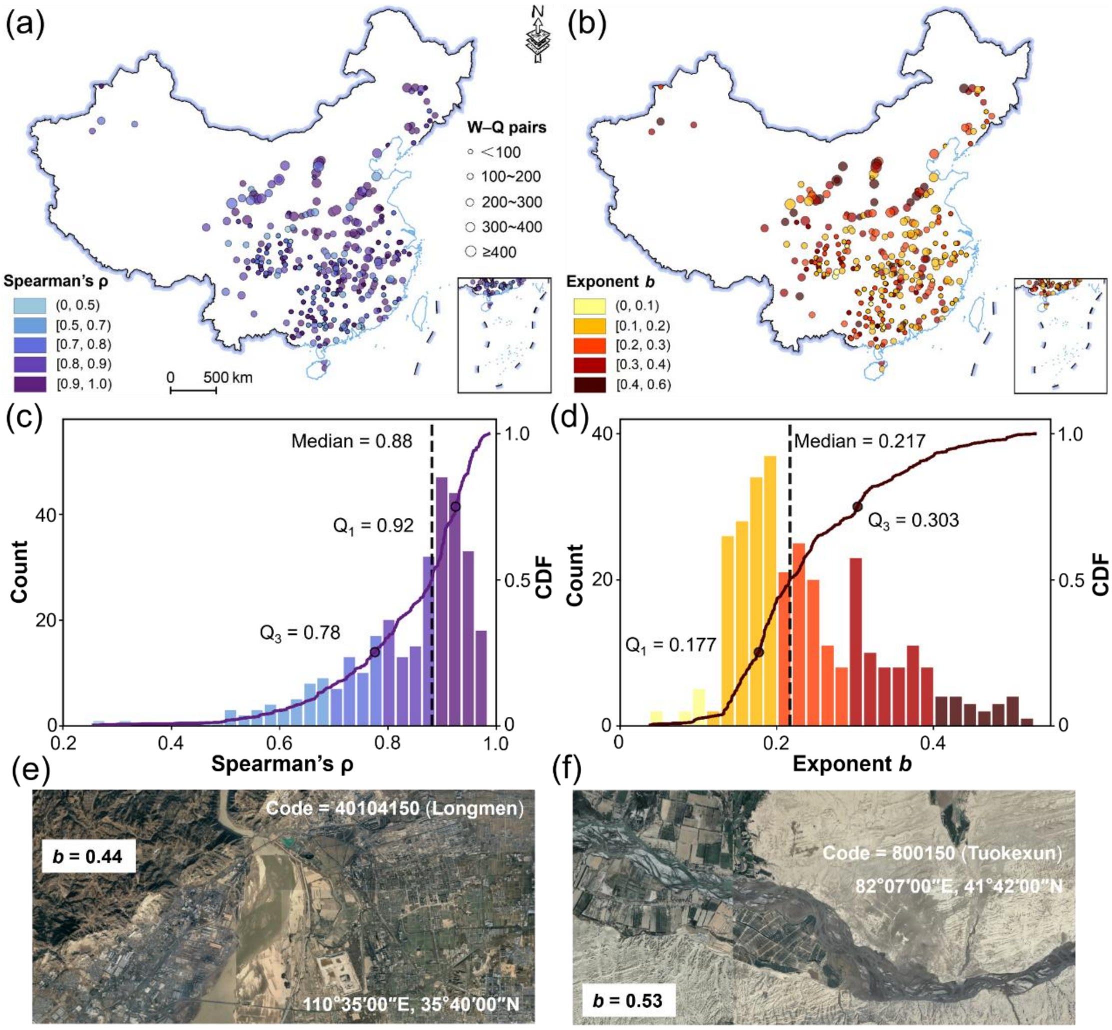

Figure 5. Spatial distribution and statistical characteristics of width–discharge relationships across the 310 selected gauges. (**a**) Spatial distribution of the Spearman rank correlation coefficient (ρ) between river width and discharge, with symbol size representing the number of width–discharge (W–Q) observation pairs at each gauge. (**b**) Spatial distribution of the power-law exponent (*b*) derived from the at-a-station

hydraulic geometry (AHG) relationship. (**c–d**) Frequency histograms and cumulative distribution functions (CDF) of ρ and b, respectively, with median values and interquartile ranges indicated. (**e–f**) High-resolution satellite imagery of representative river reaches exhibiting high b values and morphologically dynamic channel patterns. Source of the underlying image: © Google Earth. (Page 16, Lines 354–363)

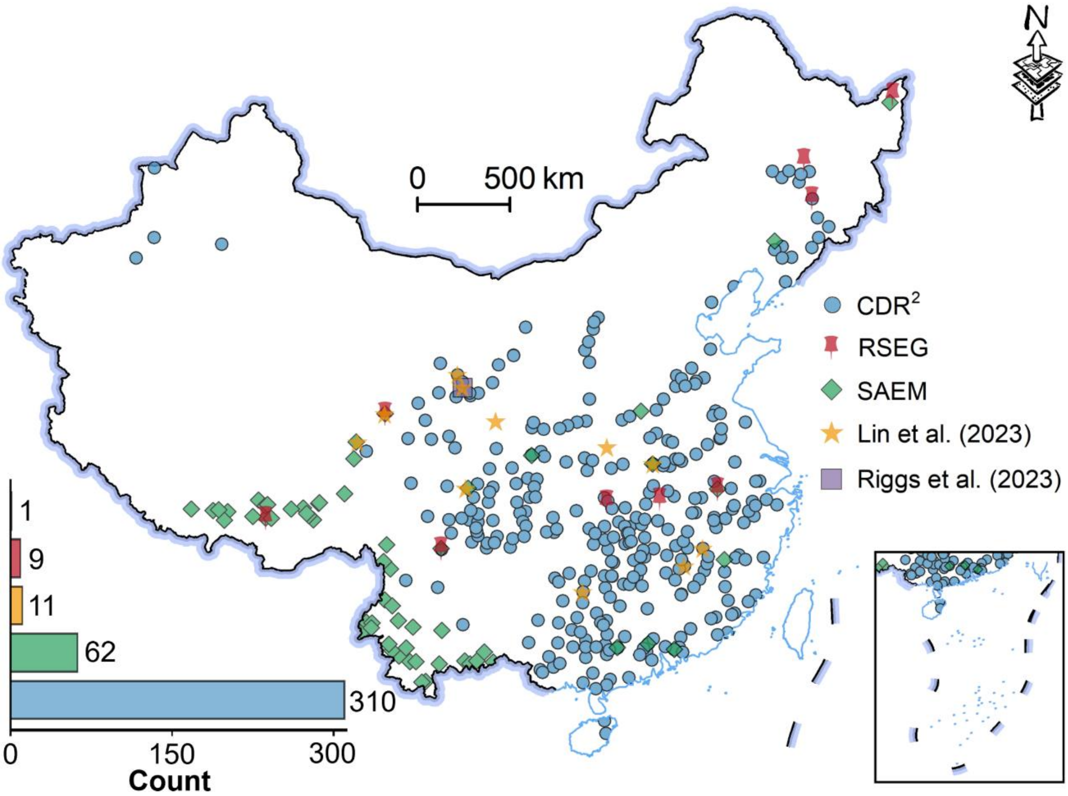

Figure 10. Spatial distribution of river discharge gauges included in CDR^2, compared with gauges from RSEG (Elmi et al., 2024), SAEM (Saemian et al., 2024), Lin et al. (2023), and Riggs et al. (2023). The inset summarizes the number of gauges contained in each dataset within China. (Page 26, Lines 571–574)

Table 2. Comparison of CDR^2 with existing satellite-extended river discharge datasets. (Page 26, Line 575)

Dataset	Hydraulic variables	Data source	Global gauge	China gauge	Temporal resolution	Validation results
Elmi et al., 2024 (RSEG)	River width/ stage	Landsat/Satellite altimetry	3377	9	Monthly	All gauges (mean): KGE = 0.35, rRMSE = 21.8%, CC = 0.43
Saemian et al., 2024 (SAEM)	River stage	Satellite altimetry	8730	62	Daily	All gauges (median): KGE = 0.48, nRMSE = 18%, CC = 0.64
Lin et al., 2023	River width	Landsat	3078	11	Monthly	>600 gauges (median): KGE = 0.33
Riggs et al., 2023	River width	Landsat/Sentinel	2168	1	Daily	All gauges (median): KGE = 0.46, rRMSE = 83%, rBias = 1.4%

CDR[2]	River width	Landsat/Sentinel	310	310	Daily	All gauges (median): KGE = 0.66, nRMSE = 12.7%, rBias = 6.6%, CC = 0.8

References

Elmi, O., Tourian, M. J., Saemian, P., and Sneeuw, N.: Remote sensing-based extension of GRDC Discharge time series – a monthly product with uncertainty estimates, Sci. Data, 11, 240, https://doi.org/10.1038/s41597-024-03078-6, 2024.

Lin, P., Feng, D., Gleason, C. J., Pan, M., Brinkerhoff, C. B., Yang, X., Beck, H. E., and Frasson, R. P. d. M.: Inversion of river discharge from remotely sensed river widths: A critical assessment at three-thousand global river gauges, Remote Sens. Environ., 287, 113489, https://doi.org/10.1016/j.rse.2023.113489, 2023.

Riggs, R. M., Allen, G. H., Wang, J., Pavelsky, T. M., Gleason, C. J., David, C. H., and Durand, M.: Extending global river gauge records using satellite observations, Environ. Res. Lett., 18, 064027, https://doi.org/10.1088/1748-9326/acd407, 2023.

Saemian, P., Elmi, O., Stroud, M., Riggs, R., Kitambo, B. M., Papa, F., Allen, G. H., and Tourian, M. J.: Satellite Altimetry-based Extension of global-scale in situ river discharge Measurements (SAEM), Earth Syst. Sci. Data, 17, 2063–2081, https://doi.org/10.5194/essd-17-2063-2025, 2025.

Reviewer 2

Comments to the Author

Overall, this dataset fills an important gap in publicly accessible river discharge observations across China and will likely be of broad interest to the hydrological and remote sensing communities. My main concerns relate to the clarity and reproducibility of the methodology, the validation strategy, and the dataset uncertainty.

Response: We sincerely appreciate the reviewer's constructive feedback and are greatly encouraged by the reviewer's recognition of the value of the Chinese Daily River Discharge Record (CDR²) dataset in filling an important observational gap. In response to the reviewer's concerns, we have revised the manuscript to provide more complete methodological details to ensure reproducibility, clarify our validation strategy, and incorporate a systematic evaluation of dataset uncertainty. Below please find our detailed responses addressing your comments.

Major comments

1. The in situ discharge observations used to establish the width-discharge relationships are primarily from 2014-2024, whereas the reconstructed discharge records span 1990-2024. This reconstruction implicitly assumes that the width-discharge relationship remains stable over the entire reconstruction period. However, channel morphology and hydraulic conditions may have changed substantially over the past three decades due to dam construction, reservoir regulation, channel engineering, sediment erosion/deposition, sand mining, or river course modifications. At a minimum, I recommend explicitly acknowledging this assumption and discussing its implications in the Limitations section. In addition, given that Landsat observations are available before 1990, could the authors clarify why 1990 was selected as the starting year of the reconstruction? Extending the dataset further back in time (e.g., to the 1980s, or earlier if feasible) could potentially provide additional value for long-term trend analyses.

Response 1: We sincerely thank the reviewer for raising these critical points. We fully agree that the implicit assumption of a stable width–discharge relationship over a 35-year period is a significant source of uncertainty, given the potential for morphological and hydraulic changes induced by both natural processes and anthropogenic activities. Accordingly, we have added a dedicated subsection "4.3 Limitations, uncertainties, and future directions" to the revised manuscript (Page 28, Line 623), which provides a thorough discussion of this issue:

4.3 Limitations, uncertainties, and future perspectives

Although the CDR^2 dataset substantially expands the availability of gauge-constrained river discharge records for long-term monitoring, the reconstruction framework remains subject to inherent limitations and uncertainties arising from both hydraulic assumptions and the constraints of remote sensing observations. First, a critical source of uncertainty lies in the implicit assumption that the width–discharge relationship remains temporally stationary. Model calibration relies on in situ observations acquired during 2014–2024 and is retrospectively applied to reconstruct discharge over 1990–2024. Extending hydraulic relationships calibrated over a decade of observations to multi-decadal discharge reconstruction is a common simplification adopted in satellite-based discharge retrieval and implicitly assumes temporal invariance of the AHG relationship (Pavelsky, 2014; Paris et al., 2016). To assess the uncertainty associated with temporal extrapolation, we quantified the extrapolation ratio for each gauge, defined as the proportion of remotely sensed river widths during 1990–2024 that fell outside the calibration width range (Fig. 12). Extrapolation ratios were generally low across the 310 reconstructed gauges, with a median of 0.81% and a mean of 2.62%. Overall, 86.5% of gauges exhibited extrapolation rates below 5% and 55.2% below 1%, suggesting that historical discharge reconstruction at most gauges remained within the calibrated hydraulic regime and that uncertainties associated with extrapolating AHG relationships beyond observed river states were limited. Nevertheless, even when reconstructed widths fall within the calibration range, hydraulic geometry may evolve over multi-decadal timescales in response to both anthropogenic and natural disturbances, including dam construction, reservoir regulation, channel engineering, sediment erosion and deposition, and sand mining (Mansanarez et al., 2019; Zhang et al., 2015). Such changes may alter the underlying width–discharge relationship and introduce additional uncertainty into long-term discharge reconstruction. For example, downstream channel incision caused by sediment retention in upstream reservoirs may substantially increase water depth and flow velocity without corresponding changes in channel width.

Second, discharge reconstruction relies exclusively on river width as the remotely sensed hydraulic observable. Although river width exhibits strong correlations with discharge across most selected gauges (median ρ = 0.88), river discharge is jointly controlled by channel width, flow velocity, bed slope, and hydraulic roughness (Leopold and Maddock, 1953; Singh et al., 2003). These hydraulic controls are not explicitly represented in the current framework and may become particularly important in rivers where channel width responds weakly to discharge variability. This interpretation is consistent with the sensitivity analysis of the hydraulic geometry exponent *b*, which demonstrates progressively improved reconstruction performance with increasing hydraulic sensitivity (Fig. 11). Additionally, the inherent limitations of optical remote sensing introduce further uncertainty. The temporal continuity of CDR^2 is inevitably affected by cloud cover, atmospheric aerosols, and seasonal snow and ice (Feng et al., 2022; Langhorst et al., 2024), resulting in substantial spatial variability in observation frequency among river basins (Fig. 9b). More importantly, extreme rainfall events are frequently accompanied by persistent cloud cover, causing optical sensors to miss peak-flow conditions and hindering the characterization of rapidly evolving flood hydrographs.

Future improvements should focus on integrating multi-source satellite observations with multi-dimensional hydraulic information. Incorporating channel morphology, sediment dynamics, and anthropogenic regulation into the reconstruction framework would facilitate the identification of temporal variations in width–discharge relationships and improve the robustness of long-term discharge reconstruction. Integrating optical imagery with synthetic aperture radar (SAR) observations would substantially improve temporal sampling under persistent cloud cover while enhancing the monitoring of extreme flood events. Furthermore, combining river width with complementary hydraulic observations, including water surface elevation and slope from emerging satellite missions such as the Surface Water and Ocean Topography (SWOT) mission, through nonlinear multivariate inversion frameworks offers a promising pathway toward more physically constrained and accurate satellite-based river discharge reconstruction.

We agree that extending the time series back to the 1980s could theoretically expand the observational window and provide additional insights for long-term trend analysis. However, given the pronounced limitations of the early Landsat archive in spatial coverage, temporal continuity, and data quality, we designated 1990 as the starting year to safeguard the reliability of river width retrieval and subsequent long-term discharge analysis. During its 1980s commercial phase (the EOSAT era), Landsat suffered from limited onboard storage and relied heavily on line-of-sight downlinks to a limited number of international ground stations, resulting in highly fragmented data acquisition outside the United States (Wulder et al., 2016). To clarify this, we have appended the following text to Section 2.2.2 of the manuscript: "Given the limited spatiotemporal coverage, relatively large geolocation errors, and low signal-to-noise ratios of the early Landsat-5 archive (Roy et al., 2014; Vogelmann et al., 2001; Wulder et al., 2016), which constrain frequent and accurate river width retrieval, we selected 1990 as the starting year of the reconstructed time series." (Page 10, Line 221 and Page 11, Lines 222–224)

Statistical analysis of the Global Long-term River Width (GLOW) dataset (Feng et al., 2022) further corroborates the severe scarcity of early data. As the first global Landsat-based river width record spanning decades (1984–2020), GLOW contains more than one billion observations from approximately 11.6 million cross-sections. To quantify the sparseness of early data, we calculated for each cross-section the proportion of observations before 1990 relative to its total observations (Fig. R2). Under the ideal assumption of uniform sampling, the theoretical proportion for 1984–1989 (6 years) relative to the full study period (37 years) is 16.22%. However, the median proportion across cross-sections is only 7.20%, substantially lower than the expected value. Notably, approximately 83% of cross-sections fall below this theoretical benchmark, and 33.15% have no observations prior to 1990 at all. The 1984–1989 period spans only six years, whereas the 1990–2024 period used in this study covers 35 years, which is sufficient to capture long-term hydrological trends. Critically, low-quality and low-frequency early river width observations introduce large errors that undermine the reliability of trend analyses based on reconstructed discharge.

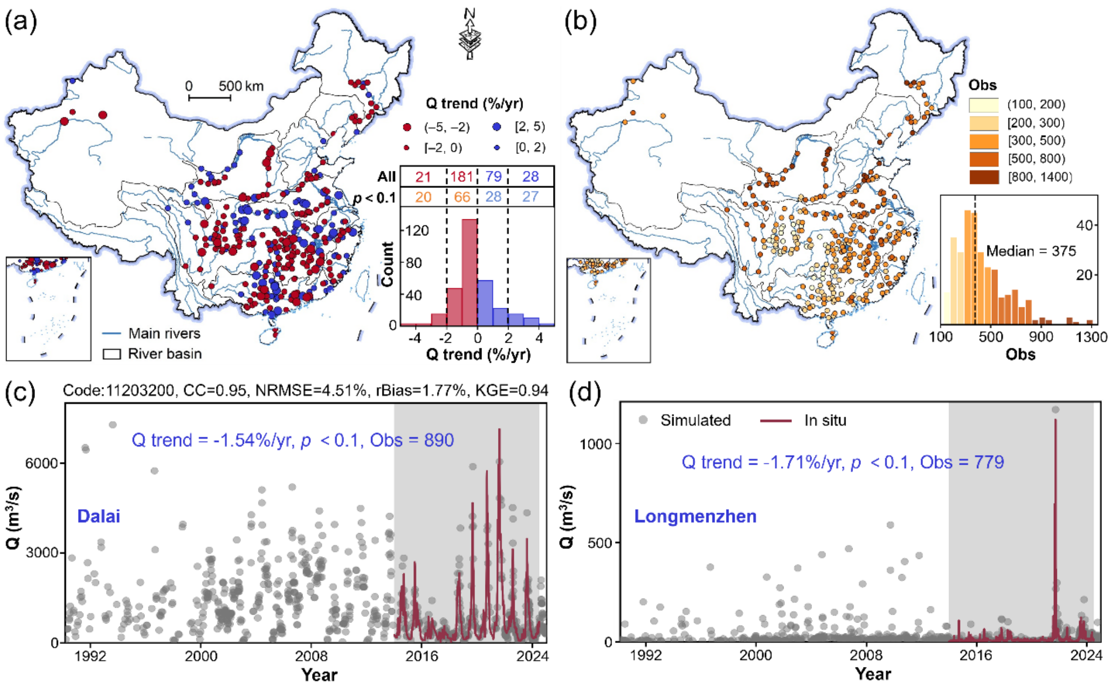

Figure 9. Spatial patterns and long-term trends of satellite-extended river discharge across 309 gauges in China (1990–2024). (**a**) Spatial distribution of relative discharge change rates (Q trend, %/yr), normalized by multi-year mean discharge, with inset

histograms showing the trend distribution and corresponding significance levels (p) across all gauges. (**b**) Spatial distribution of additional observations (Obs) provided by the satellite-extended discharge records at each gauge. (**c–d**) Representative time series of reconstructed daily-scale discharge estimates at Dalai and Longmenzhen gauges, respectively, illustrating long-term discharge variability and trends together with corresponding model performance metrics. (Page 24, Lines 521–529)

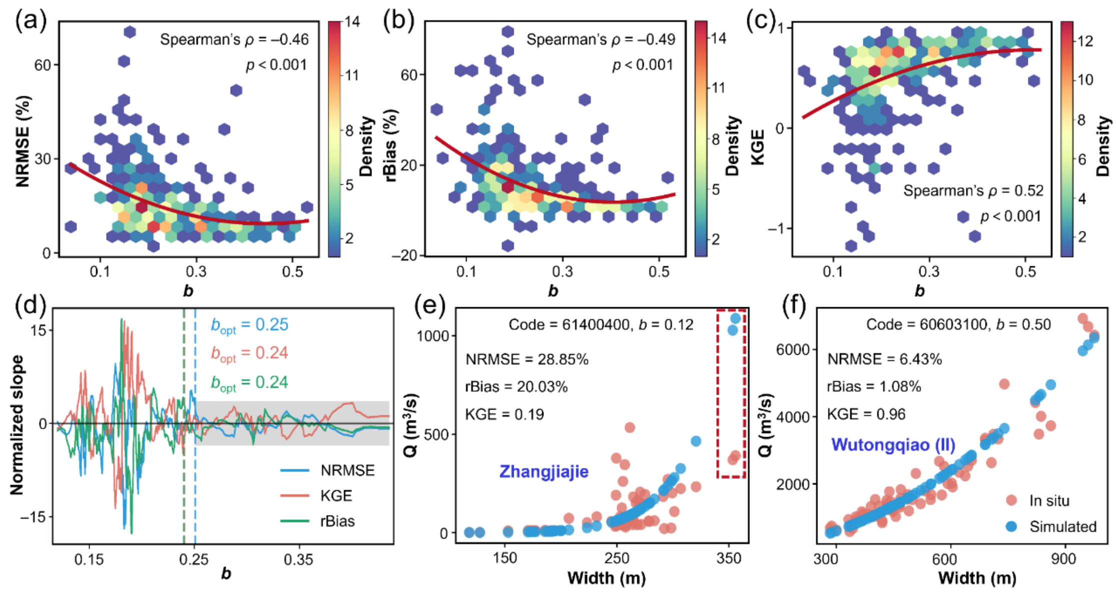

Figure 11. Influence of the width–discharge power-law exponent (b) on the performance of satellite-extended discharge reconstruction. (**a–c**) Relationships between the hydraulic geometry exponent (b) and the performance metrics NRMSE, rBias, and KGE, respectively. Red solid lines indicate fitted relationships, with the corresponding Spearman rank correlation coefficient (ρ) and significance level (p) indicated. (**d**) Variation in the normalized slope of performance metrics with respect to b, used to identify the optimal threshold (b_{opt}) beyond which reconstruction performance becomes stable and less sensitive to changes in hydraulic responsiveness. (**e–f**) Comparison between in situ and reconstructed discharge at the Zhangjiajie and Wutongqiao (II) gauges, representing examples of low-b and high-b hydraulic conditions, respectively. (Page 28, Lines 613–622)

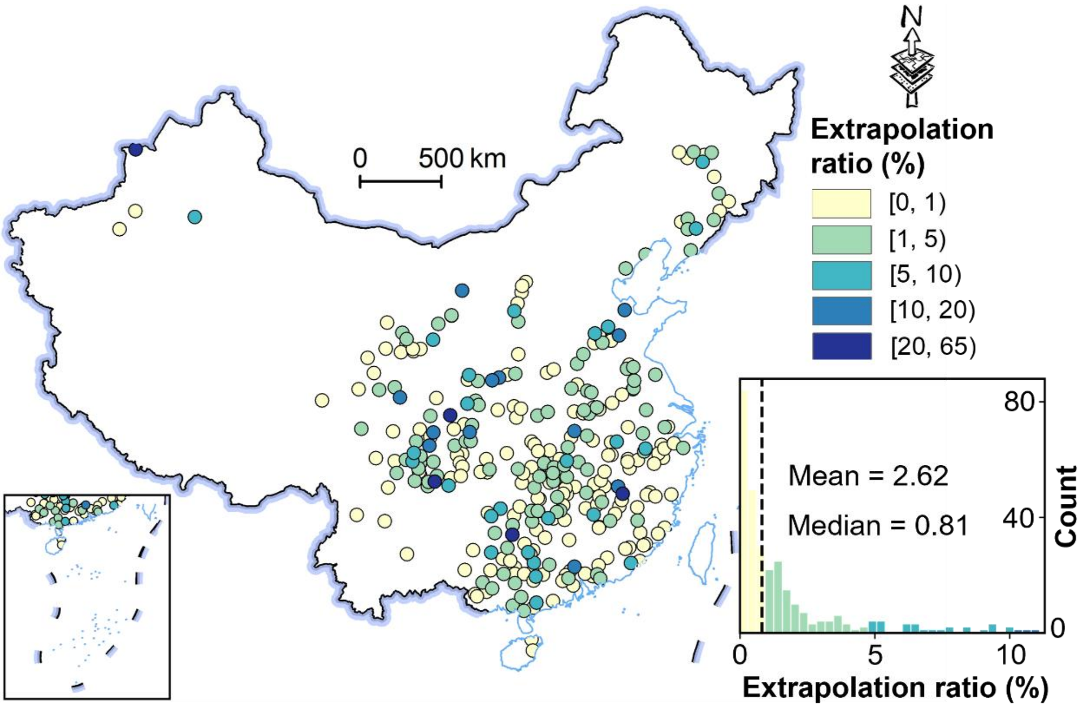

Figure 12. Spatial distribution of extrapolation ratios (%) for the 310 gauges included in the CDR[2] dataset. The inset shows the frequency distribution of extrapolation ratios across all gauges (mean = 2.62, median = 0.81), truncated at the 95th percentile to improve visualization. (Page 31, Lines 675–678)

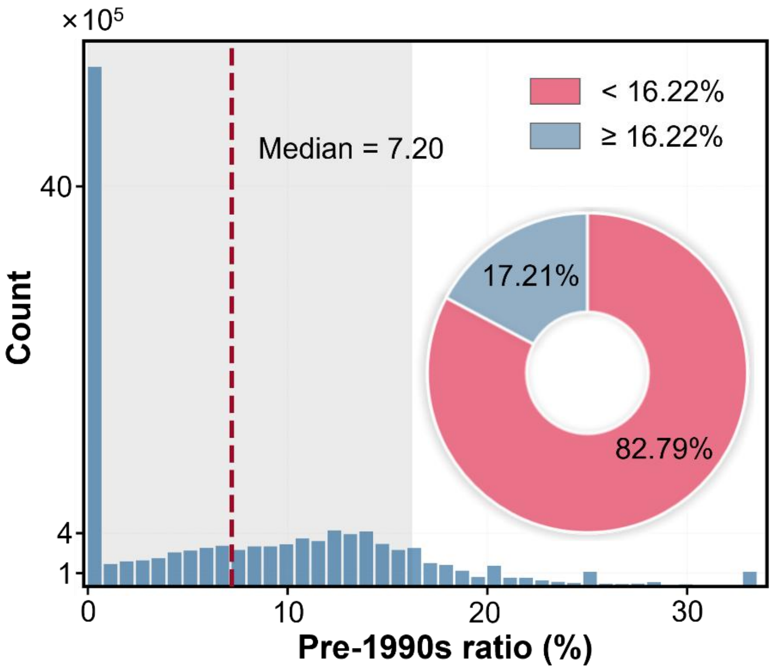

Figure R2. Frequency distribution of pre-1990s observation ratios (%) across river reaches in the Global Long-term River Width (GLOW) dataset (Feng et al., 2022), with a median of 7.20%. The ratio represents the proportion of river width observations collected before 1990 relative to the total observations for each reach. The analysis includes ~11.56 million river reaches. The shaded region denotes reaches with ratios below the theoretical expectation (16.22%), while the inset donut chart summarizes the corresponding proportions. Values above the 95th percentile were truncated for improved visualization.

References

Feng, D., Gleason, C. J., Yang, X., Allen, G. H., and Pavelsky, T. M.: How have global river widths changed over time?, Water Resour. Res., 58, e2021WR031712, https://doi.org/10.1029/2021WR031712, 2022.

Langhorst, T., Andreadis, K. M., and Allen, G. H.: Global cloud biases in optical satellite remote sensing of rivers, Geophys. Res. Lett., 51, e2024GL110085, https://doi.org/10.1029/2024GL110085, 2024.

Leopold, L. B. and Maddock, T.: The hydraulic geometry of stream channels and some physiographic implications, U.S. Government Printing Office, Washington, DC, https://pubs.usgs.gov/pp/0252/report.pdf (last access: 12 March 2026), 1953.

Mansanarez, V., Renard, B., Le Coz, J., Lang, M., and Darienzo, M.: Shift happens! Adjusting stage-discharge rating curves to morphological changes at known times, Water Resour. Res., 55, 2876–2899, https://doi.org/10.1029/2018WR023389, 2019.

Pavelsky, T. M.: Using width-based rating curves from spatially discontinuous satellite imagery to monitor river discharge, Hydrol. Process., 28, 3035–3040, https://doi.org/10.1002/hyp.10157, 2014.

Paris, A., Dias de Paiva, R., Santos da Silva, J., Medeiros Moreira, D., Calmant, S., Garambois, P.-A., Collischonn, W., Bonnet, M.-P., and Seyler, F.: Stage-discharge rating curves based on satellite altimetry and modeled discharge in the Amazon basin, Water Resour. Res., 52, 3787–3814, https://doi.org/10.1002/2014WR016618, 2016.

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., Schott, J. R., Sheng, Y., Vermote, E. F., Belward, A. S., Bindschadler, R., Cohen, W. B., Gao, F., Hipple, J. D., Hostert, P., Huntington, J., Justice, C. O., Kilic, A., Kovalskyy, V., Lee, Z. P., Lymburner, L., Masek, J. G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R. H., and Zhu, Z.: Landsat-8: Science and product vision for terrestrial global change research, Remote Sens. Environ., 145, 154–172, https://doi.org/10.1016/j.rse.2014.02.001, 2014.

Singh, V. P., Yang, C. T., and Deng, Z. Q.: Downstream hydraulic geometry relations: 1. Theoretical development, Water Resour. Res., 39, 1337, https://doi.org/10.1029/2003WR002484, 2003.

Vogelmann, J. E., Helder, D., Morfitt, R., Choate, M. J., Merchant, J. W., and Bulley, H.: Effects of Landsat 5 Thematic Mapper and Landsat 7 Enhanced Thematic Mapper Plus radiometric and geometric calibrations and corrections on landscape characterization, Remote Sens. Environ., 78, 55–70, https://doi.org/10.1016/S0034-4257(01)00249-8, 2001.

Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., Fosnight, E. A., Shaw, J., Masek, J. G., and Roy, D. P.: The global Landsat archive: Status, consolidation, and direction, Remote Sens. Environ., 185, 271–283, https://doi.org/10.1016/j.rse.2015.11.032, 2016.

Zhang, W., Wang, W.-G., Zheng, J.-H., Wang, H.-G., Wang, G., and Zhang, J.-S.: Reconstruction of stage–discharge relationships and analysis of hydraulic geometry variations: The case study of the Pearl River Delta, China, Glob. Planet. Change, 125, 60–70, https://doi.org/10.1016/j.gloplacha.2014.12.004, 2015.

2. The manuscript reports validation statistics based on concurrent width-discharge pairs. However, it is currently unclear whether the calibration and validation datasets are independent. Specifically, readers would benefit from knowing:

- how many observations were used for model calibration versus validation;
- whether an independent testing dataset, cross-validation, or leave-one-year-out validation was employed; or
- whether the same width-discharge pairs were used for both model fitting and performance evaluation.

Response 2: We sincerely thank the reviewer for this insightful and highly constructive comment. The reviewer raised a critical point regarding the independence of our model validation. In the original manuscript, the same concurrent width–discharge pairs at each gauge were used for both model fitting and performance evaluation—essentially an in-sample fitting (ISF) approach. We completely agree that clarifying the independent predictive capability of the models is essential for the readers.

To rigorously address your concern and test the predictive independence of our models, we have conducted a leave-one-out cross-validation (LOOCV) analysis for all gauges, and the following description has been added to Section 3.2 "Given the limited sample size available at individual gauges, and because evaluating the predictive capability of the AHG relationships is not the primary objective of this study, the in-sample fitting (ISF) approach was adopted for final parameterization to maximize the calibration sample size and thereby ensure the robustness of the long-term discharge reconstruction.” (Page 17, Lines 365–368) and “To assess the robustness of the performance evaluation, a leave-one-out cross-validation (LOOCV) analysis was performed (Kim et al., 2014). Compared with the ISF results, LOOCV exhibited only marginal reductions in performance, with median decreases in CC and KGE of 0.01–0.04, while differences in NRMSE and rBias remained below 0.7% (Fig. A1). These minor discrepancies indicate that the ISF approach introduces little overfitting and that the reconstructed discharge estimates are robust to the choice of calibration strategy.” (Page 17, Lines 386–391)

In addition, we have added related statements in Section 4.3 to evaluate the uncertainty associated with temporal extrapolation beyond the calibration range: “To assess the uncertainty associated with temporal extrapolation, we quantified the extrapolation ratio for each gauge, defined as the proportion of remotely sensed river widths during 1990–2024 that fell outside the calibration width range (Fig. 12). Extrapolation ratios were generally low across the 310 reconstructed gauges, with a median of 0.81% and a mean of 2.62%. Overall, 86.5% of gauges exhibited extrapolation rates below 5% and 55.2% below 1%, suggesting that historical discharge reconstruction at most gauges remained within the calibrated hydraulic regime and that uncertainties associated with extrapolating AHG relationships beyond observed river states were limited.” (Page 29, Lines 633–640)

Overall, these additional analyses demonstrate that the reported reconstruction performance is robust to the validation strategy and that the long-term discharge reconstruction involves only limited temporal extrapolation beyond the calibrated hydraulic range.

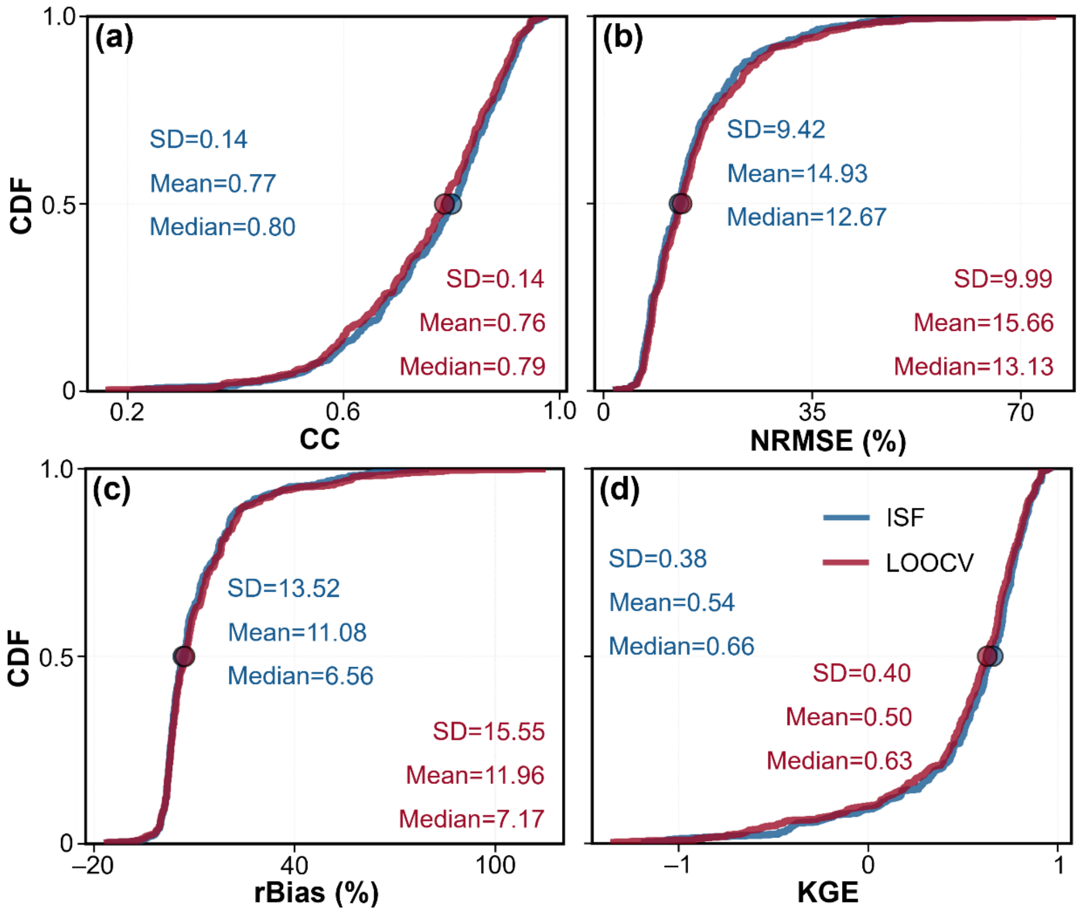

Figure A1. Comparison of reconstruction performance for the satellite-extended river discharge dataset across 310 gauges using in-sample fitting (ISF) and leave-one-out cross-validation (LOOCV) based on concurrent river width–discharge observations. (a–d) Cumulative distribution functions (CDFs) of the Pearson correlation coefficient (CC), normalized root mean square error (NRMSE, %), relative bias (rBias, %), and Kling–Gupta efficiency (KGE), respectively. For each metric, the corresponding mean, median, and standard deviation (SD) are annotated for both ISF and LOOCV. (Page 33, Lines 730–732 and Page 34, Lines 733–737)

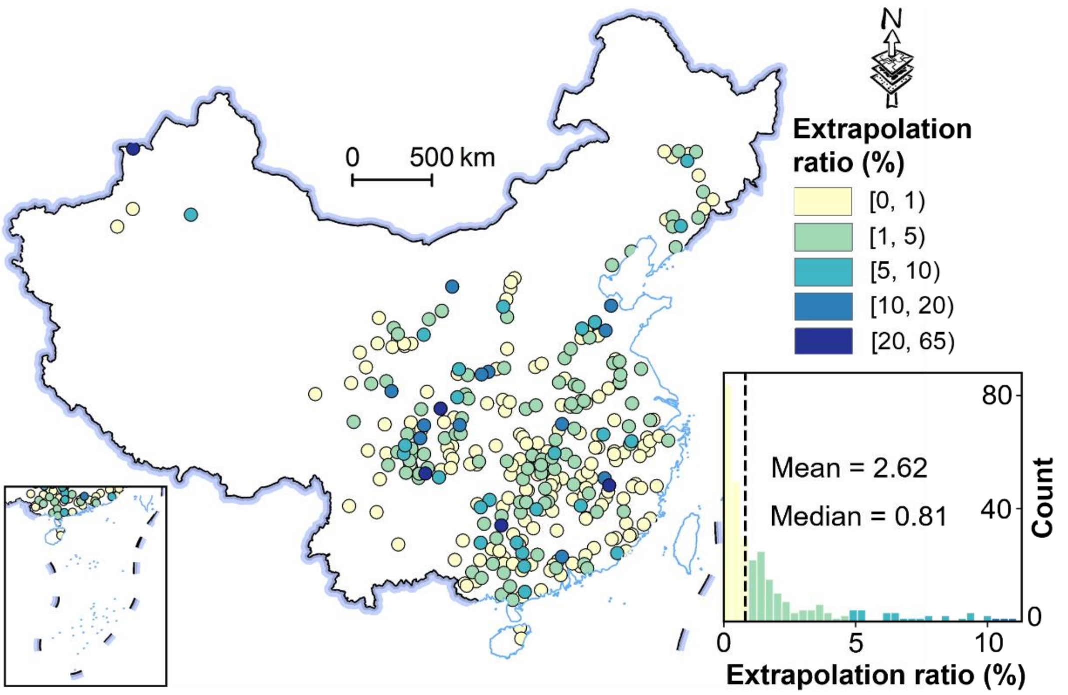

Figure 12. Spatial distribution of extrapolation ratios (%) for the 310 gauges included in the CDR^2 dataset. The inset shows the frequency distribution of extrapolation ratios across all gauges (mean = 2.62, median = 0.81), truncated at the 95th percentile to improve visualization. (Page 31, Lines 675–678)

References

Kim, Y. H., Im, J., Ha, H. K., Choi, J.-K., and Ha, S.: Machine learning approaches to coastal water quality monitoring using GOCI satellite data, GIScience Remote Sens., 51, 158–174, https://doi.org/10.1080/15481603.2014.900983, 2014.

3. I think the dataset would benefit from a more comprehensive assessment of uncertainty. In particular:

◆ Are uncertainty estimates or confidence intervals available for the reconstructed discharge values?

◆ What is the uncertainty associated with the fitted AHG parameters?

Response 3: We sincerely thank the reviewer for this valuable comment. We agree that a comprehensive uncertainty assessment is essential for evaluating the reliability of satellite-based discharge reconstruction. In the original manuscript, uncertainties associated with the fitted AHG parameters and their propagation into discharge estimates were not explicitly quantified. To address this concern, we have added a new uncertainty quantification framework in Section 2.4 (Page 14, Line 300) and corresponding results in Section 3.3 (Page 20, Line 428) based on a nonparametric bootstrap resampling approach.

2.4 Uncertainty quantification

To quantify uncertainty in the fitted AHG parameters and the reconstructed discharge series, a nonparametric bootstrap procedure was applied independently to each gauge. This approach is well suited to hydrological applications because it does not require prior assumptions about the error distribution of the remote sensing observations (Hesterberg et al., 2011; Rodrigues et al., 2015). For each bootstrap realization, N pairs of width–discharge observations were sampled with replacement from the original set of N observations, and the AHG relationship in Eq. (4) was refitted using nonlinear least squares. Repeating this procedure 1,000 times yielded empirical distributions of the coefficient *a* and exponent *b*. Their relative uncertainties, RU_a and RU_b, were defined as the ratio of the half-width of the 95% confidence interval to the median of the corresponding bootstrap distribution:

$$\mathrm{RU_a} = \frac{(a_{97.5} - a_{2.5})}{2a_{50}} \times 100\% \quad (10)$$

$$\mathrm{RU}_b = \frac{(b_{97.5} - b_{2.5})}{2b_{50}} \times 100\% \quad (11)$$

where $a_{2.5}$, a_{50}, and $a_{97.5}$ denote the 2.5th, 50th, and 97.5th percentiles of the bootstrap distribution of coefficient *a*, respectively, with analogous definitions for $b_{2.5}$, b_{50}, and $b_{97.5}$ corresponding to exponent *b*.

For each remotely sensed river width observation (W_t) at time step t, the 1,000 bootstrap parameter sets were then used to derive the corresponding discharge distribution:

$$\widehat{\mathrm{Q_t}}^{(i)} = \left(\frac{\mathrm{W_t}}{a^{(i)}}\right)^{1/b^{(i)}}, i = 1, 2, ..., 1000 \quad (12)$$

The 2.5th, 50th, and 97.5th percentiles of this distribution, denoted as $\mathrm{Q_{t,2.5}}$, $\mathrm{Q_{t,2.5}}$, and $\mathrm{Q_{t,97.5}}$, respectively, were used to calculate the relative discharge uncertainty at time step t:

$$\mathrm{RU}_Q(\mathrm{t}) = \frac{(\mathrm{Q_{t,97.5}} - \mathrm{Q_{t,2.5}})}{2\mathrm{Q_{t,50}}} \times 100\% \quad (13)$$

Finally, the median $\mathrm{RU}_Q(\mathrm{t})$ across all available time steps was used to represent the overall uncertainty in discharge reconstruction at each gauge. All uncertainty measures are expressed as dimensionless percentages, enabling direct comparison among gauges with different discharge magnitudes and across geographic regions.

3.3 Uncertainty of river discharge estimates

Although the reconstructed discharge records agree well with in situ observations, uncertainties in the parameters fitted from the AHG relationships may accumulate and propagate through the discharge retrieval process. To quantify this effect, we employed bootstrap resampling to estimate the propagation of AHG parameter uncertainty in the reconstructed discharge series across 310 gauges (Fig. 8). The median relative uncertainties of the coefficient *a* and exponent *b* (RU_a and RU_b) were 14.36% and 13.24%, respectively (Fig. 8a, b). Approximately 68% of gauges exhibited RU_a below 20%, and about 80% showed RU_b below 20%, indicating robust parameter estimation for the majority of gauges, with the half-widths of their 95% confidence intervals all less than one-fifth of the median parameter estimates. The median relative uncertainty in reconstructed discharge attributable to parameter propagation (RU_Q) was 11.89%, and approximately 83% of gauges had RU_Q below 20% (Fig. 8c), suggesting that the influence of parameter estimation errors on discharge reconstruction is generally acceptable. Fewer than 2% of gauges exhibited RU_Q exceeding 40%, all of which had fewer than 25 width–discharge paired samples, implying that limited calibration data are the primary source of elevated discharge uncertainty. Spatially, gauges with higher parameter and discharge uncertainties were mainly located in the Yangtze and Pearl River basins (Fig. 8a–c). Frequent cloud cover in these regions may reduce the availability of effective optical observations and seasonal representativeness, thereby weakening the statistical constraints on AHG parameters and amplifying discharge uncertainty.

Moreover, RU_Q showed a significant positive correlation with the exponent *b* ($R^2 = 0.40$, $p < 0.001$), but no significant correlation with coefficient *a* ($R^2 = 0.01$, $p > 0.05$) (Fig. 8d). This divergence is consistent with the AHG equation, in which uncertainty in the exponent *b* is amplified through its reciprocal exponent. The exponent *b* therefore exerts a stronger control on uncertainty propagation, supporting the necessity of filtering out gauges with physically unreasonable b values, as described in Section 3.1. Overall, the CDR^2 dataset exhibits relatively constrained parameter-induced uncertainty for most gauges, providing a data foundation with quantitative confidence bounds for subsequent analyses of long-term discharge change across China. However, RU_Q only characterizes the propagation of AHG parameter uncertainty under given satellite-derived river widths and does not incorporate other sources of uncertainty, including remote sensing observation errors, model structural errors, or changes in external hydrological processes.

To further reflect the importance of uncertainty characterization for the CDR^2 dataset, we have also incorporated the key uncertainty result into the Abstract and

Conclusions: "Compared with existing global satellite-derived discharge products, CDR^2 expands the number of available gauges in China by more than fivefold while delivering substantially improved performance, achieving a median Kling–Gupta efficiency of 0.66 during validation. Moreover, uncertainties propagated from the fitted hydraulic geometry parameters remain low, with a median relative uncertainty of only 11.89% in the reconstructed discharge estimates." (Page 1, Lines 17–22) and "Notably, 91% of gauges exhibit positive KGE values, approximately 73% achieve CC values greater than 0.7, and 79% maintain NRMSE below 20%. Uncertainty analysis further indicates that uncertainties propagated from model parameters remain limited, with a median relative uncertainty of 11.89% in discharge estimates." (Page 32, Lines 711–714)

These revisions establish a quantitative uncertainty assessment framework for CDR^2, allowing users to better understand the confidence level of the reconstructed discharge records and the influence of AHG parameter variability on discharge estimation.

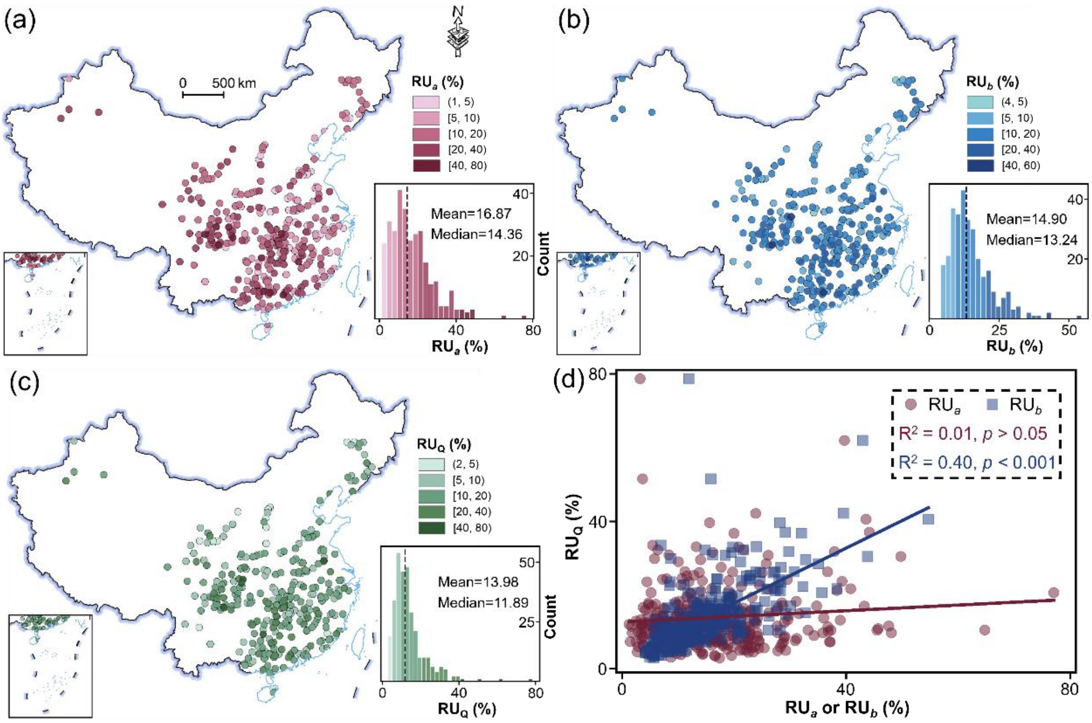

Figure 8. Uncertainty assessment of satellite-extended river discharge at 310 gauges based on the bootstrap method. (**a–c**) Spatial distributions of the relative uncertainties of the width–discharge power-law coefficient *a* (RU_a), exponent *b* (RU_b), and reconstructed discharge (RU_Q), respectively. The relative uncertainty is defined as the relative half-width of the 95% confidence interval, determined by the 2.5th and 97.5th percentiles of the bootstrap distribution. Insets in each panel show frequency

distribution histograms, with mean and median values indicated. (**d**) Scatter plots and linear regressions between RU_Q and RU_a or RU_b, demonstrating that uncertainty in discharge reconstruction is primarily controlled by exponent *b* during error propagation, with the corresponding coefficients of determination (R^2) and statistical significance levels (*p*) indicated. (Page 21, Lines 460–469)

References

Hesterberg, T.: Bootstrap, WIREs Comput. Stat., 3, 497–526, https://doi.org/10.1002/wics.182, 2011.

Rodrigues, D. B. B., Gupta, H. V., Mendiondo, E. M., and Oliveira, P. T. S.: Assessing uncertainties in surface water security: An empirical multimodel approach, Water Resour. Res., 51, 9013–9028, https://doi.org/10.1002/2014WR016691, 2015.

Minor comments

1. Line 143: Regarding Fig. 2d, could the authors explain why CDR^2 contains substantially fewer gages in the Southwest Rivers Basin (Basin A) than the CHP dataset? Is this mainly due to the site-selection criteria adopted in this study?

Response 1: We sincerely thank the reviewer for this insightful comment. The substantially fewer gauges of the CDR^2 dataset in the Southwest Rivers Basin (Basin A) are not due to any basin-specific gauge-selection criteria adopted in this study. In fact, we directly compiled all available daily discharge observations from the National Hydrological and Rainfall Information System of the Ministry of Water Resources of China (http://xxfb.mwr.cn/), with the only processing step being the removal of spatially redundant gauges located within 100 m of each other.

The difference between CDR^2 and the CHP dataset mainly results from their different data sources and construction purposes. The CHP dataset was specifically developed for the Southwest Rivers Basin and therefore contains a relatively dense gauge network in this region. In contrast, no daily discharge records from the Southwest Rivers Basin were available from the source platform used for CDR^2 during the study period, resulting in the absence of gauges from this basin. To avoid potential misunderstanding, we have revised the corresponding description in the manuscript: "The resulting gauge network spans nearly all major river basins in China, with the Southwest Rivers Basin being the only exception because discharge records were unavailable from the source platform." (Page 6, Lines 137–139)

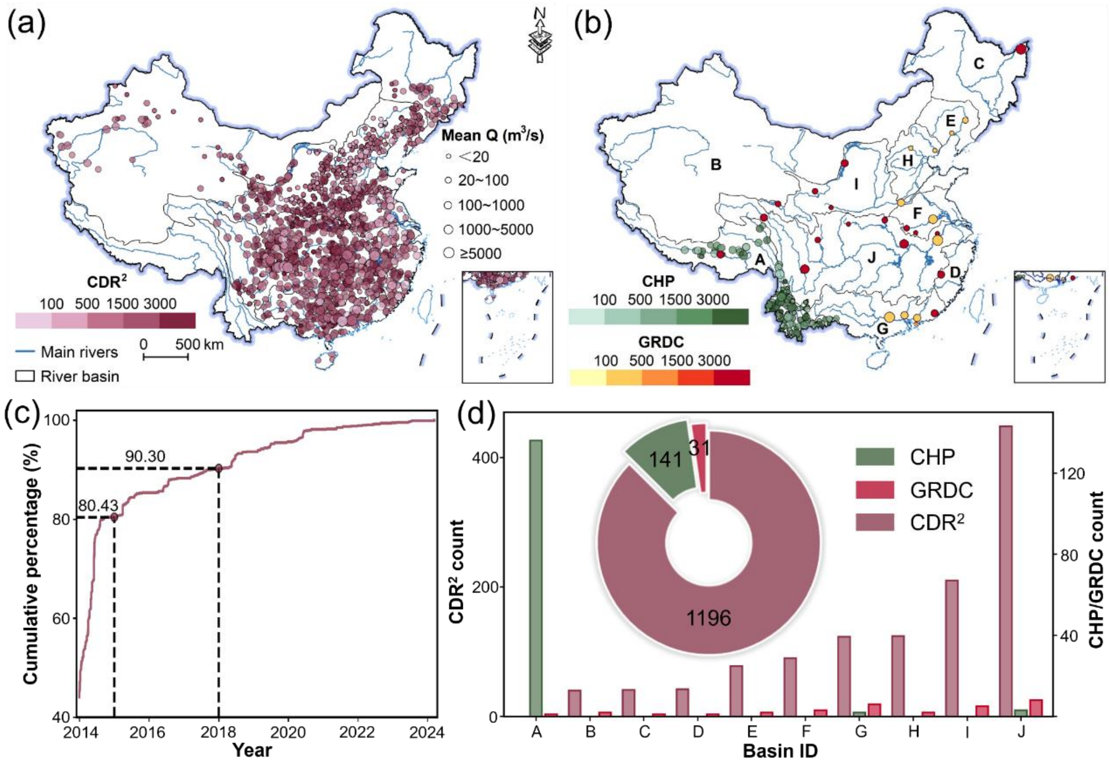

Figure 2. Spatial distribution and temporal coverage of in situ river discharge observations across China (2014–2024). (**a**) Spatial distribution of the 1,196 gauges included in the China Daily River Discharge Records (CDR^2) dataset, with symbol size representing mean discharge (Q, m³/s) and color indicating the number of discharge records at each gauge. (**b**) Spatial distribution of river gauges available from the Chinese Hydrology Project (CHP) and the Global Runoff Data Centre (GRDC) within China, with symbol size representing mean discharge (Q, m³/s) and color indicating the number of discharge records at each gauge. The labeled regions (A–J) denote the ten major basins: Southwest Rivers (A), Northwest Rivers (B), Songhua River (C), Southeast Rivers (D), Liao River (E), Hai River (F), Pearl River (G), Huai River (H), Yellow River (I), and Yangtze River (J). (**c**) Cumulative percentage distribution of gauge observation start years in the CDR^2, indicating that over 80% of gauges contain records beginning in or before 2015, while more than 90% begin in or before 2018. (**d**) Comparison of the number of gauges included in the CDR^2, CHP, and GRDC datasets across China. (Page 7, Lines 154–165)

2. Lines 206–207: Since both Landsat and Sentinel were used, did the authors perform any cross-sensor bias correction or consistency assessment before combining observations from the two sensor systems?

Response 2: We sincerely thank the reviewer for this valuable comment. In this study, we did not perform an additional cross-sensor bias correction before combining observations from Landsat and Sentinel-2. This is because river width was extracted independently from the Level-2 surface reflectance products of each sensor, which have already undergone standard radiometric calibration, atmospheric correction, and geometric correction to ensure a high level of consistency in radiometric properties and geolocation accuracy across sensor systems (Claverie et al., 2018). Moreover, the river width retrieval was based on MNDWI-derived water masks rather than direct spectral value fusion, thereby reducing the potential influence of inter-sensor radiometric differences on the final width estimates.

It should also be clarified that the median compositing procedure was applied only to multiple images acquired by the same sensor on the same day, rather than across different sensors. Therefore, no pixel-level co-registration or cross-sensor image fusion was involved in the processing workflow. When daily river width estimates from different sensors were available for the same date, the independently derived width measurements were averaged to obtain a single daily river width value. To avoid potential misunderstanding, we have revised the corresponding description in the manuscript: "For each sensor, images acquired on the same day were median-composited to generate daily surface reflectance mosaics while minimizing cloud contamination. When multiple river width estimates from different sensors were available for the same date, they were averaged to obtain a single daily width value." (Page 11, Lines 228–231)

References

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., Skakun, S. V., and Justice, C.: The Harmonized Landsat and Sentinel-2 surface reflectance data set, Remote Sens. Environ., 219, 145–161, https://doi.org/10.1016/j.rse.2018.09.002, 2018.

3. Lines 250–251: The site-selection procedure could be described more explicitly. It would be helpful to clearly state here that only 310 of the 401 candidate gages were ultimately retained because they satisfied the valid width-discharge pairing criteria.

Response 3: Thanks for this comment. We agree that the gauge-selection procedure should be described more explicitly. The criterion of at least ten valid width–discharge pairs (N≥10) represents only the initial screening step to ensure sufficient observations for width–discharge relationship assessment, rather than the final selection criterion.

Specifically, among the 401 candidate gauges, 358 gauges satisfied the minimum requirement of N≥10 and were retained for further screening. Additional screening procedures were then applied to remove gauges with weak or physically unrealistic width–discharge relationships, including the exclusion of gauges with low Spearman correlation coefficients ($\rho < 0.25$) and physically implausible hydraulic geometry exponents ($b < 0$ or $b > 0.59$). After these quality-control procedures, a total of 310 gauges were finally retained for discharge reconstruction.

Although the final number of retained gauges (310) was previously reported in Section 3.1, we agree that explicitly stating this information in the Methods section improves the transparency of the gauge-selection procedure. Therefore, we have added the following description after the screening criteria: "Gauges with fitted exponents outside this range were removed to avoid incorporating unstable or physically implausible hydraulic relationships into the discharge reconstruction framework. After applying all screening criteria, a total of 310 gauges were retained for river discharge reconstruction." (Page 13, Lines 281–284) In addition, the description of the initial screening step has been revised as follows: "Accordingly, the instantaneous river widths derived from satellite imagery (Section 2.2) were paired with concurrent in situ daily discharge observations at each gauge (Section 2.1). Among the 401 candidate gauges, only those with at least ten valid width–discharge pairs (N≥10) were retained, yielding 358 gauges for subsequent evaluation of the width–discharge relationship." (Page 12, Lines 265–269)

4. Line 349: "river dynamism" or "river dynamics"?

Response 4: Thanks for this comment. We agree that "river dynamics" is a more commonly used term to describe temporal variations in river systems. Although "river dynamism" was intended to represent the degree of river width variability quantified by the coefficient of variation of river width (CV_w), we have replaced it with "river dynamics" for clarity and consistency with the terminology commonly used in hydrological studies.

Accordingly, we have revised the corresponding description in the manuscript: "To further investigate the influence of river dynamics on reconstruction performance, the selected gauges were stratified into five equal-frequency groups (G1–G5) according to the coefficient of variation of river width (CV_w), with group mean values ranging from 6% to 23%." (Page 18, Lines 398–400)

5. Line 383: The font of "with Sen's slope estimator" differs from the surrounding text. Similar formatting inconsistencies also appear around Lines 393 and 410.

Response 5: Thanks for pointing out this formatting issue. We have re-examined the formatting of the relevant sections and ensured that the font style and formatting of "with Sen's slope estimator" and the corresponding text around Lines 393 and 410 are consistent with the surrounding text in the revised manuscript.

6. Lines 412–413: The manuscript refers to reconstructed "daily discharge" from 1990-2024. However, the reconstructed time series are not temporally continuous on a daily basis. I suggest clarifying this point in the manuscript to avoid potential misunderstanding.

Response 6: We agree that the term "daily discharge" may potentially imply a temporally continuous daily record. In this study, the reconstructed discharge estimates are derived from satellite-observed river widths from Landsat and Sentinel-2 imagery. Therefore, each estimate corresponds to a specific date (daily temporal resolution), but the time series are not continuously available for every day due to satellite observation frequency and data availability. To avoid potential misunderstanding, we have revised the corresponding description in the manuscript: "Using the satellite-extended discharge estimates with daily temporal resolution from 310 gauges spanning 1990–2024, we quantified the long-term trends and observational characteristics of rivers across China." (Page 22, Lines 471–473)

In addition, we have also checked the manuscript and revised other potentially ambiguous uses of "daily discharge" and "continuous" to ensure that the reconstructed products are clearly described as date-specific discharge estimates rather than continuous daily records.

7. Line 439: I recommend adding a dedicated subsection discussing the limitations of the dataset and its appropriate scope of application.

Response 7: Thanks for this comment. We agree that the assumptions, limitations, and appropriate application scope of the CDR^2 dataset should be explicitly discussed. Accordingly, we have added a dedicated subsection to discuss the major limitations and potential uncertainties associated with the dataset, including the temporal stationarity assumption of width–discharge relationships, optical satellite observation constraints, and future improvement directions.

The newly added discussion is presented in Section 4.3 "Limitations, uncertainties, and future directions" of the revised manuscript (Page 28, Line 623), which also addresses the related concerns raised in Major Comment 1.

8. Line 479 (Section 4.1 and Fig. 9): The comparison with existing products should be interpreted with greater caution. CDR^2 is derived from a screened subset of Chinese gages that satisfy the width-discharge relationship requirements, whereas the global datasets include a much broader range of river types, including hydraulically challenging sites. Therefore, statements such as "substantially outperforming" existing products may not represent a fully equivalent comparison and could be moderated accordingly.

Response 8: Thanks for this helpful comment. We agree that comparisons among satellite-based discharge datasets should be interpreted with caution, as different products may involve different gauge selection strategies and represent distinct ranges of river environments. In particular, CDR^2 was developed using a screened subset of Chinese gauges with stable width–discharge relationships, whereas global products aim to maximize spatial coverage and may include a broader range of hydraulically challenging gauges.

To avoid overinterpretation, we have moderated the corresponding statements in the manuscript: "Notably, CDR^2 achieves a median KGE of 0.66, exceeding the reported global median range of 0.33–0.48 among existing products (Table 2), although these comparisons should be interpreted considering differences in gauge selection criteria and river characteristics." (Page 25, Lines 549–552) and "Overall, CDR^2 provides an expanded and higher-resolution characterization of river discharge across China, a hydrologically critical yet globally underrepresented region." (Page 25, Lines 566–568)

In addition, we have revised related statements in the Abstract and Conclusions to avoid implying an absolute performance advantage of CDR^2 over existing global products. The corresponding revisions are as follows: "Compared with existing global satellite-derived discharge products, CDR^2 expands the number of available gauges in China by more than fivefold while delivering substantially improved performance, achieving a median Kling–Gupta efficiency of 0.66 during validation. Moreover, uncertainties propagated from the fitted hydraulic geometry parameters remain low, with a median relative uncertainty of only 11.89% in the reconstructed discharge estimates." (Page 1, Lines 17–22) and "Compared with existing global satellite-based discharge datasets, CDR^2 substantially expands the spatial coverage of discharge records at daily temporal resolution across China by more than fivefold while achieving improved reconstruction accuracy." (Page 32, Lines 715–717).