Remote Sensing of Environment 287 (2023) 113489

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Inversion of river discharge from remotely sensed river widths: A critical assessment at three-thousand global river gauges

Peirong Lin [a,b,*], Dongmei Feng [c], Colin J. Gleason [d], Ming Pan [e], Craig B. Brinkerhoff [d], Xiao Yang [f,g], Hylke E. Beck [h], Renato Prata de Moraes Frasson [i]

[a] Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing 100871, China
[b] International Research Center for Big Data for Sustainable Development Goals, Beijing 100094, China
[c] Department of Chemical and Environmental Engineering, University of Cincinnati, Cincinnati, OH 45211, USA
[d] Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA 01003, USA
[e] Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093, USA
[f] Department of Earth, Marine and Environmental Sciences, University of North Carolina at Chapel Hill, NC 27599, USA
[g] Now at Department of Earth Sciences, Southern Methodist University, Dallas, TX 75205, USA
[h] Physical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[i] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

ARTICLE INFO

Edited by Menghua Wang

Keywords:
Remote sensing of discharge
BAM
geoBAM
Landsat
River width
Global assessment

ABSTRACT

Accurately estimating river discharge from satellite-derived river hydraulic variables (e.g., width, height, and slope) is the overarching goal of the remote sensing of discharge (RSQ) community. Numerous past studies have developed and intercompared different RSQ algorithms to demonstrate their feasibility, yet relatively few have focused on assessing how the RSQ algorithms are adapted to a wide range of rivers globally. As the community is now ready to expand to the global scale given advances in computing power, sensors, and the launch of the Surface Water and Ocean Topography (SWOT) satellite mission, a much broader geographic view of RSQ accuracy should be prioritized toward "better generalizability" instead of "higher accuracy at limited places". To help close this gap, we extracted multi-temporal river widths from >350 K Landsat scenes at >3 K river reaches globally, and used them to estimate discharge using the Bayesian AMHG-Manning (BAM) algorithm and the geomorphologically-enhanced variant (geoBAM). We use this framework to demonstrate how to apply an 'off the shelf' RSQ algorithm and test it globally without methodological intervention and answer: does it live up to its promise? Our daily discharge inversions (1984–2019) showed positive Kling-Gupta Efficiency (KGE) at 27% of the gauges for BAM and 39% for geoBAM, and this percentage increased to 46–65% after feeding richer priors on flow seasonality and monthly variability, amounting to 1400–2000 successful inversions. Exploratory analyses showed that the inversion is the most sensitive to a channel geomorphological parameter *b* and climate aridity, where the optimal conditions are high-*b*, humid environments, as well as moderate width variability, leaf area index (*LAI*), and river width. Although specific to BAM/geoBAM, constraining the factors to their optimal ranges led to a median KGE of 0.33 for >600 gauges, which highlights the promising potential for global RSQ. We further discussed the optimal configuration of the RS/priors by analyzing results derived from different information content in priors. Overall, our critical assessment of BAM/geoBAM reveals a successful global implementation of an existing algorithm that SWOT will improve. We suggest similar large-scale assessments for other RSQs be prioritized to identify the emerging challenges as we move into a new era with global river monitoring capability from space.

Confidential manuscript in revision at *Remote Sensing of Environment*, Jan. 20th, 2023.

1. Introduction

River water is the most accessible and critical water resource for life

* Corresponding author at: Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing 100871, China.
E-mail address: peironglinlin@pku.edu.cn (P. Lin).

https://doi.org/10.1016/j.rse.2023.113489
Received 18 July 2022; Received in revised form 20 January 2023; Accepted 27 January 2023
Available online 4 February 2023

Table 1
Summary of relevant studies focusing on assessing the RSQ accuracy with a global scope. Summarized information includes the assessment sample size, hydraulic variables of interest, data used, requirement for calibration, and their key results. Note that this summary does not incorporate those with a regional focus, those completely within one continent/basin, or one single climatic regime.

Reference	Sample Size	Hydraulic Variables	Data Used	Requirement for calibration	Key Results
Revilla-Romero et al. (2014)	322 stations	River surface extent	GFDS passive microwave observations	Calibrated against discharge measurements	48% stations show NSE > 0
Sichangi et al. (2016)	8 rivers	River width/ stage	MODIS + satellite altimetry	Calibrated against discharge measurements	NSE: 0.6–0.97
Gleason et al. (2014)	34 rivers	River width	Landsat	No	Median RRMSE 33% for non-braided rivers
Van Dijk et al. (2016)	8000+ stations	River surface extent	MODIS + passive microwave data	Trained against discharge measurements	86% stations show correlation of <0.7
Durand et al. (2016)	19 rivers	River width/ stage/slope	Hydraulic model outputs	No	Median standard deviation of relative residuals at 12.5% for 16 non-braided rivers
Hou et al. (2020)	10,000+ stations	River surface extent	MODIS data	Trained against model-based discharge	25% stations show correlation ≥0.4
Frasson et al. (2021)	31 rivers	River stage/ width/slope	Hydraulic model outputs	No	Median normalized standard deviation of residuals: 19%–29%
This study	3000+ rivers	River width	Landsat	No	~1/3 stations show KGE > 0; 46%–65% stations show KGE > 0 if feeding improved priors; 69% stations show daily KGE > 0.2 if confining to optimal conditions

despite its low volumetric percentage among the Earth's freshwater storages. An earth system science view to study the spatio-temporal dynamics of river discharge (or interchangeably referred as discharge or streamflow) is crucial to our understanding of global rivers as a subsystem of the Earth instead of disconnected segments or watersheds. Yet this understanding is largely compromised by spatial gaps (Do et al., 2018), a downward trend (Hannah et al., 2011), and placement biases (Krabbenhoft et al., 2022) of river gauge networks, urging researchers to actively seek for other means, such as remote sensing and global hydrologic modeling approaches, to fill in these gauging gaps.

As a relatively young field, remote sensing of river discharge (hereafter RSQ, where Q denotes discharge) has been burgeoning in recent years along with the increasing availability of satellite data, computational resources, and increased spatial/temporal/spectral resolutions of satellite sensors (see Gleason and Durand (2020) and literatures therein). Since the early demonstration of useful signals of surface water variability in spaceborne measurements (Alsdorf and Lettenmaier, 2003; Brakenridge et al., 2005; Smith, 1997), studies have been devoted to retrieving essential hydraulic variables (e.g., river surface extent/stage/slope, termed as "RS river observables") from space, and then propagating useful RS signals into Q by establishing relationships between RS observables and Q. A plethora of research has been conducted to assess the usability of different optical or active/passive microwave satellite sensors (e.g., Alsdorf et al., 2007; Feng et al., 2019; Huang et al., 2018, 2020a, 2022; Marcus and Fonstad, 2008; Sichangi et al., 2016; Tarpanelli et al., 2021; Tourian et al., 2013), signal/image processing techniques to overcome contaminations from clouds/vegetation/sand bars/wet banks (e.g., Brakenridge et al., 2007; Huang et al., 2018; Tarpanelli et al., 2013), and flow laws/physics (e.g., Bjerklie et al., 2003; Gleason and Smith, 2014), which have collectively contributed to the flourishment of RSQ.

The rationale behind estimating Q with RS observables is simple – Q variability responds to changes in river width (W) or surface inundation area (A), as well as river stage (H) or slope (S) variations. If one can obtain in-situ Q measurements or river cross sections & velocity profile surveys, transferring remotely sensed river W/H into Q is straightforward as it only requires establishing empirical relationships between W/H and Q, the same practice taken when installing automatic gauging stations. Many studies followed this logic to develop RSQ algorithms, which relied on W/H-discharge rating curves (e.g., Pavelsky, 2014) and/or statistical regression models (e.g., Bjerklie et al., 2003, 2018) to estimate Q. However, the need for *in-situ* Q precludes RSQ's applications in truly ungauged scenarios and is thought to have limited gains for Q prediction in ungauged basins (i.e., the PUB challenge, Sivapalan et al., 2003; Hrachowitz et al., 2013). This recognition caused researchers to look to another RSQ group that combines mechanistic river hydraulics/flow laws with Bayesian inference (Durand et al., 2014; Hagemann et al., 2017) or data assimilation (DA) frameworks (Andreadis et al., 2020; Gejadze et al., 2022; Larnier et al., 2020; Oubanas et al., 2018) to estimate Q. The emergence of this group of RSQ reflected the conceptual updates from earlier studies, and as they do not require *in-situ* information, they were selected for the composition of the ensemble algorithms by the Surface Water and Ocean Topography (SWOT) satellite mission to estimate Q globally (Durand et al., 2021). Among them, the Bayesian inference sub-group termed the Mass-conserved Flow Law Inversion (McFLI; more details in Gleason et al. (2017) and Frasson et al. (2021)) is particularly promising as it requires much less computation for large-scale applications compared to DA while retaining the probabilistic way of Q estimation. Given our interest in global-scale RSQ, this study is specifically concerned with the McFLI sub-group of RSQ unless otherwise cautioned.

Despite the prosperous development of RSQ that focused on showing feasibility, relatively few have challenged RSQ's generalizability across a wide range of rivers. Often times researchers tend to make explicit or implicit assumptions on RSQ being readily extensible to the global scale by showing good Nash-Sutcliffe Efficiency (NSE) skill at limited sites (e.g., Gleason and Smith, 2014; Huang et al., 2020b, 2022; Smith and Pavelsky, 2008). Rarely do RSQ algorithmic studies go beyond limited evaluation sites. This confinement comes not only from the significant amount of work with site-specific tuning of physics/parameters for RSQ development, but also from the difficulty in obtaining large-scale RS observations for Q estimation. For example, the SWOT family of RSQ used a few dozen sites when constructing and evaluating the algorithms – the constraints mainly came from the limited hydraulic model outputs to mimic the as-yet-to-exist SWOT data (Durand et al., 2016; Frasson et al., 2021). Outside of the SWOT context, studies have assessed various combination of optical, microwave sensor, or altimetry data for Q estimation at a few to thousands of gauges, but these studies tended to rely heavily on *in situ*/model outputs for calibration/training, or they have not been evaluated by stricter skill metrics beyond correlation (see Table 1 for a summary of studies with similar scopes). As a result, thus far, a clear understanding of when and where state-of-the-art RSQ yields good or limited accuracy is still very much constrained with respect to the geography of global rivers.

Now as SWOT has been launched in December 2022 which is expected to revolutionize the monitoring capability of global inland waters

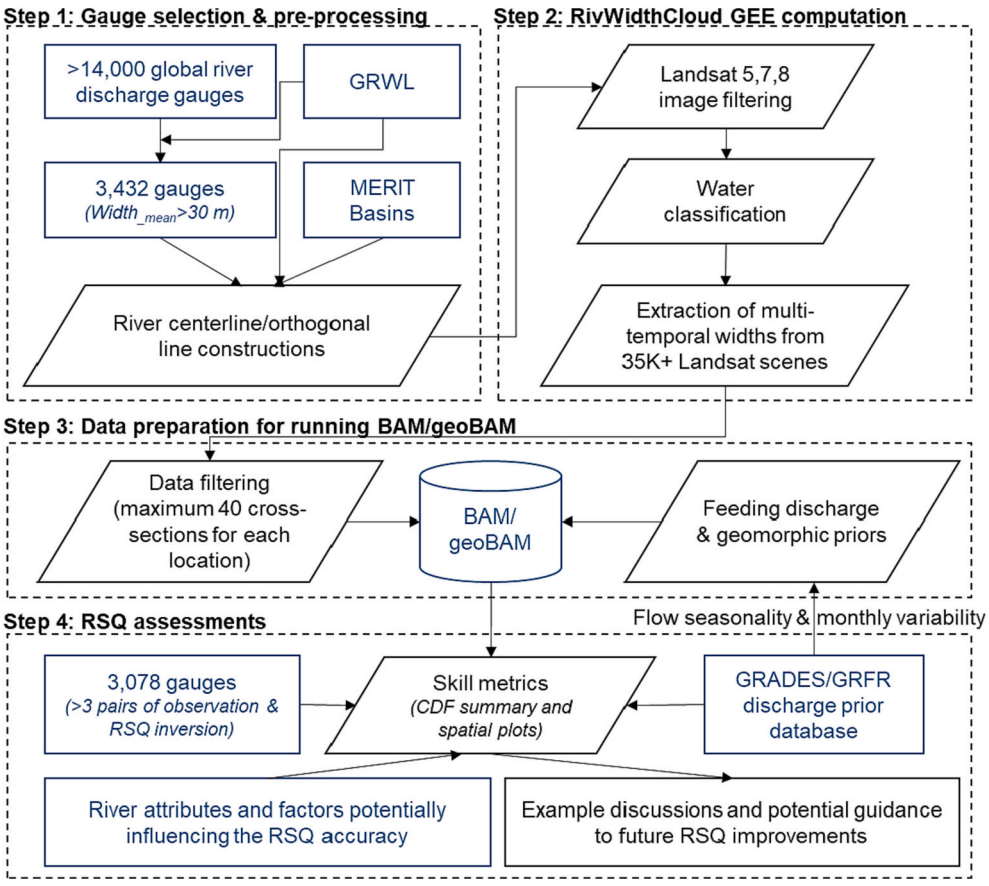

Fig. 1. Technical flowchart of this study. GRWL (Allen and Pavelsky, 2018) refers to the static river width estimates from Landsat; MERIT Basins (Lin et al., 2019) refers to a vector-based river network shapefile; GRADES/GRFR (Lin et al., 2019; Yang et al., 2021) refers to the naturalized river discharge simulated by a global hydrologic model; and see Acknowledgement for the discharge gauge database. Major tools used: RivWidthCloud (Yang et al., 2020b) refers to the GEE tool for river width extraction; BAM (Hagemann et al., 2017) and geoBAM (Brinkerhoff et al., 2020) refer to the discharge inversion algorithms based off the AMHG (Gleason and Smith, 2014; Brinkerhoff et al., 2019) flow law.

(Biancamaria et al., 2016), we are particularly motivated to generate more challenging testing cases, and help prioritize the assessment of how state-of-the-art RSQs are adapted to the wide variety of global rivers, particularly in ungauged basins. Our motivation is in line with the recent petabyte-scale satellite datasets and tools made available via cloud computing platforms such as the Google Earth Engine (GEE) (Gorelick et al., 2017; Pekel et al., 2016; Riggs et al., 2022; Yang et al., 2020b), which has improved our mapping and modeling capabilities for global rivers (Allen and Pavelsky, 2018; Feng et al., 2021; Lin et al., 2019; 2020; Yamazaki et al., 2019; Yang et al., 2020a). Although the available RS data from GEE are less accurate and partial compared to SWOT, the wealth of such data are already valuable to facilitate large-scale assessment, potentially guiding RSQ developers to consider places needing more attention for improvements.

We aim to shed light on the following three questions: *i)* How well does state-of-the-art McFLI perform at daily Q inversion globally? *ii)* What factors may contribute to or limit the RSQ inversion skill and how sensitive are they? *iii)* How to optimize the Q inversion strategies by jointly considering the improved RS data and prior knowledge on global rivers? To achieve the goals, we extracted multi-temporal river widths from >350 K Landsat scenes using GEE, and performed McFLI evaluation at >3000 gauging sites from 1984 to 2019. Two members of the SWOT family RSQ, i.e., BAM (Hagemann et al., 2017) and geoBAM (Brinkerhoff et al., 2020) were used to frame our algorithmic examples, as both of them rely on the at-many-station hydraulic geometry (AMHG) physics (Gleason and Smith, 2014), making them the sole McFLI capable of Q inversion with width-only observations and relatively easy-to-comprehend computations. We use these algorithms as an example to show how to approach generalizing RSQ worldwide as they become testable at global scales. The upcoming data (e.g., SWOT) will improve the accuracy of these results.

The paper is organized as follows. Section 2 presents our technical workflows to obtain long-term river width observations for global RSQ assessments. Section 3 presents the RSQ accuracy with 'off-the-shelf' algorithm setups and the exploratory analyses to assess factors affecting the accuracy. In Section 4, we present the RSQ accuracy after feeding richer priors on Q seasonality and monthly variability, where prominent examples showing the interplay between the river conditions and RS/prior setups are presented. Then, we reach at conclusions and discussions on future work for improvements in Section 5.

2. Data and methods

In this section, we describe the datasets, tools, and methods adopted by this study. Fig. 1 summarizes the technical flowcharts with details provided below.

2.1. Determining gauged rivers observable by Landsat

To provide a stern test for global RSQ, we first identified all gauged rivers observable by Landsat. A global database with >14,000 gauges was used here; it contains daily discharge observations collected from various international and national sources such as the Global Runoff Data Centre (GRDC), the European Water Archive (EWA), the U.S. Geological Survey (USGS), the Brazilian National Water Agency (ANA), Australian Bureau of Meteorology (BoM), and the Chilean Center for Climate and Resilience Research (CR2). This database was spatially matched with the Global River Width from Landsat (GRWL) (Allen and Pavelsky, 2018), resulting in 3432 gauges located on a river wider than 30 m excluding lakes/reservoirs (Fig. 2). These locations were used to search for all available Landsat imagery as they were considered observable by Landsat at 30-m spatial resolution (Section 2.2). These gauges cover a wide range of discharge from 0.06 to 165,770.99 m^3/s (with a median of 96.96 m^3/s and mean of 518.87 m^3/s; Fig. 2a), river

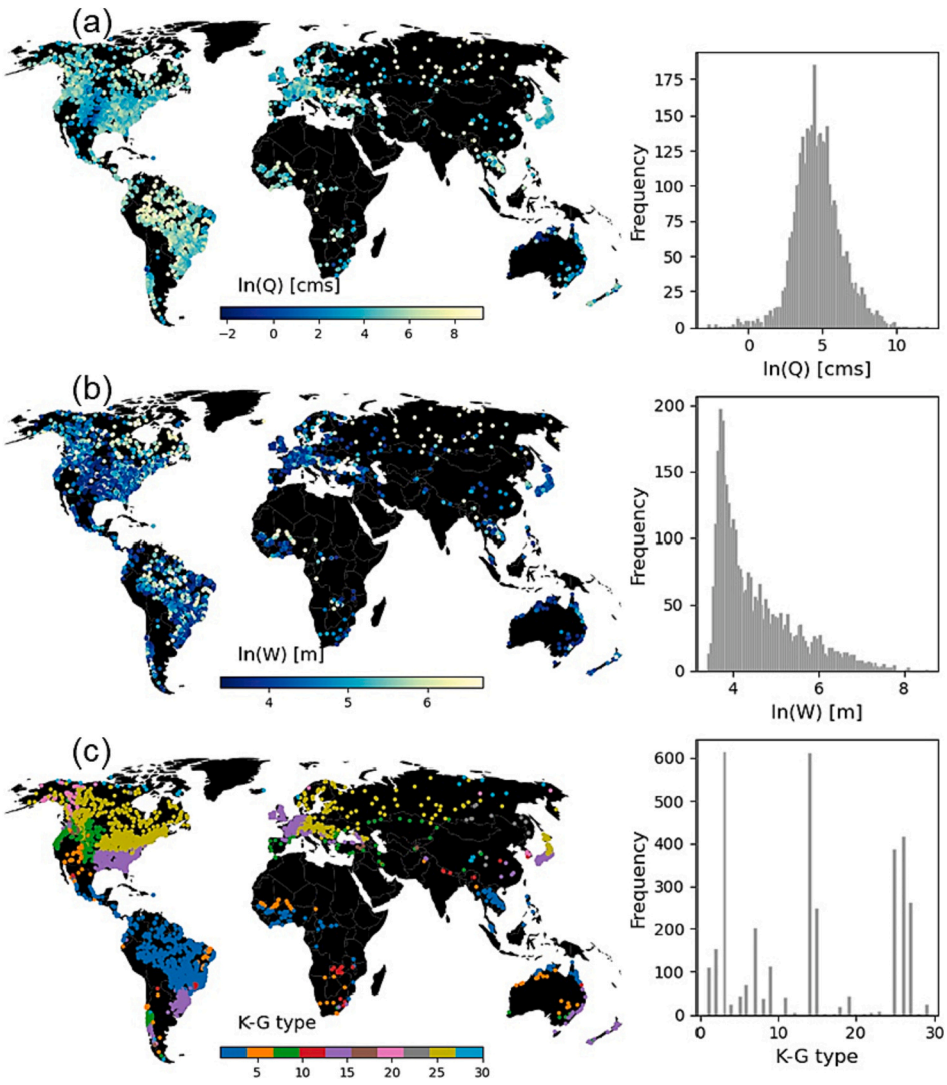

Fig. 2. Spatial pattern and histogram plot of the 3432 gauges used for Landsat width extraction. (a) is for mean annual flow (m^3/s), (b) is for mean width (m), and (c) is for Köppen-Geiger climate classification obtained from Beck et al. (2018); 1–3 for tropical climate, 4–7 for arid climate, 8–16 for temperate climate, 17–28 for tropical climate, and 29–30 for polar climate.

width from 30 m to 5000 m (with a median of 60 m and mean of 150 m; Fig. 2b), and major Köppen-Geiger climate zones (arid, cold, temperate, and tropical; Fig. 2c). By leveraging all possible gauges without a biased view to retain or disregard any gauges a priori, we expect the substantial spatial heterogeneity to challenge any existing RSQ algorithm by nature. The aim is to expose potential problems to the maximum extent possible, and help guide areas to prioritize improvements. This design is different from those starting with idealized conditions and increasingly added errors (Durand et al., 2016; Frasson et al., 2021) – we approach the problem in a conceptually opposite way from the worst-case scenario, and then systematically approach optimal conditions.

2.2. Constructing cross sections/reaches needed by width extraction and the inversion

At each location, multi-temporal river widths were extracted from all available Landsat 5, 7, and 8 imagery from 1984 to 2019 via the Google Earth Engine (GEE) RivWidthCloud tool (Yang et al., 2020b). RivWidthCloud measures wetted river width by overlaying orthogonal lines at selected cross sections with water masks. To avoid the most time-consuming part with river centerline/orthogonal calculation, these lines were pre-defined as inputs for efficient large-scale extraction. More specifically, each gauge was spatially joined with one MERIT river reach and several GRWL segments (Lin et al., 2019) – we used the GRWL segment centroids (a 30–42.4 m interval) as the cross sections, and the MERIT river reach (a median length of 6.7 km) as the unit for imposing mass balance for BAM/geoBAM discharge inversion (see Fig. 3a for an example). This "cross section/node" and "reach" definition is in line with how the SWOT product will be organized (Altenau et al., 2021). Here as the dense GRWL cross sections (interval of 30–42.4 m) have been shown to substantially slowdown the BAM/geoBAM computation, sub-sampling was used to maintain accuracy compared to the full density inversions (Hagemann et al., 2017). We first eliminated the nodes within $2\text{xwidth}_{\text{mean}}$ intervals if the node number exceeded 100 for a given reach – this helped to improve the computational efficiency while obtaining the most complete RS measurements. We then constructed orthogonal lines with a length of $4\text{xwidth}_{\text{mean}}$ (Fig. 3b) and overlaid them with water masks determined by the default RivWidthCloud water classification algorithm (Jones, 2019) to extract widths. Finally, as BAM/geoBAM (Section 2.4) cap the number of input cross-sections at 40 for a given time/river to optimize the computation demand and the performance gain, we further sub-sampled the input cross sections based on the width data availability. This sampling may introduce uncertainties, but in the absence of true dynamic width data as the

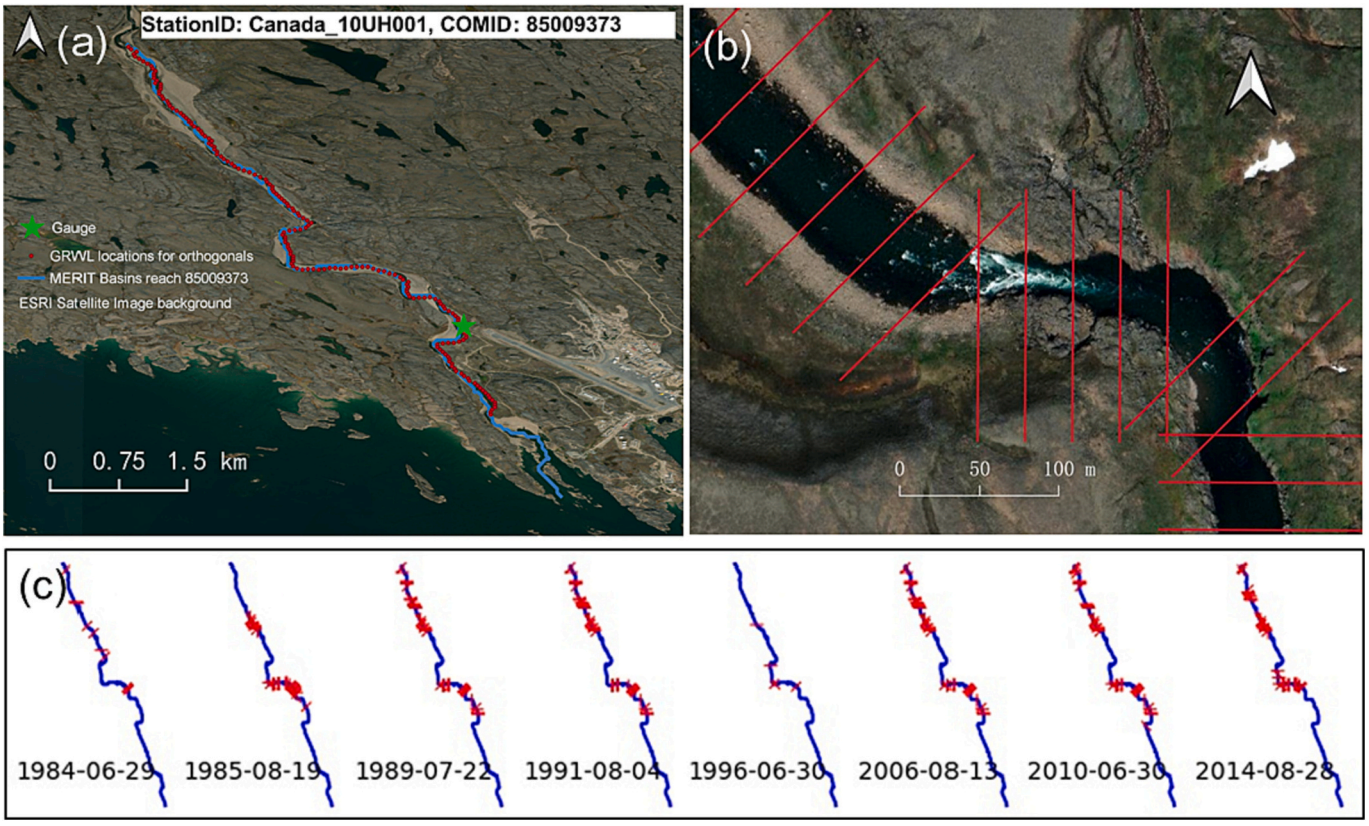

Fig. 3. (a) Examples of the cross sections/nodes (red dots) for orthogonal line construction in the width extraction process, and the river reach (blue segment) for discharge inversion. This example gauge (Canada 10UH001) is randomly selected and located in a river reach of 15.36 km long (COMID: 85009373); the cross-section locations are from the GRWL segment centroids joined to the reach. (b) shows a zoomed-in plot for the constructed river orthogonal lines from which the wetted widths are extracted. (c) shows the sub-sampled cross-sectional widths (ranging between 5 and 40) that feed into BAM/geoBAM algorithms for discharge inversion; for simplicity, only eight dates are shown. Basemap is from the ESRI Satellite image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

reference, this strategy is used to retain the best observable part of a reach (Feng et al., 2021, 2022). Fig. 3c shows the time-varying cross-sectional widths for discharge inversion at different dates, which ranges from 5 to 40.

2.3. Multi-temporal Landsat observations at 3000+ gauged rivers (1984–2019)

Only Landsat measurements not affected by cloud effects, snow cover, and topographical shadows were considered for the water classification. In total, we collected 357,389 Landsat scenes meeting such criteria for the eight continents of MERIT Basins, namely Africa (14,797), Europe (53,605), North Asia (5696), South Asia (27,938), Oceania and South Asian Islands (25,478), South America (77,119), North America (136,428), and Arctic Region (16,328). Fig. 4 summarizes the reach-averaged widths to assess the obtained Landsat river observations.

The revisit time of Landsat is 16 days and with all the qualified images from TM, ETM+, and OLI sensors combined, the temporal interval for width retrievals can be <16 days for many locations. Overall, majority of locations have 150–350 observations across the 36 years due to non-usable scenes with cloud/snow/shade contamination, but some can exceed 600–800 (Fig. 4a). There is also a clear latitudinal pattern showing limited observations for the northern hemisphere tropics and high-latitudes, where clouds or snow covers are high; as a comparison, extra-tropics arrive at the largest number of observations. The mean width (Fig. 4b) derived from the temporally intermittent observations generally match with the static width estimates by GRWL and the gauges' drainage areas (Fig. 2), suggesting the correctness of the multi-temporal width extraction process. The coefficient of variation (CV) (Fig. 4c) reflects the combined effect of the RS observability as well as the geomorphic variability of rivers related to the channel geometry, and it mostly falls between 0.1 and 0.3. Using circular statistics (see SI Text S1 for the equations), we also calculate the mean observation date ($\overline{D}$, Fig. 4d) and how concentrated these dates are (CI index, Fig. 4e) to further summarize the river RS data. It clearly shows that Landsat observations in the Amazon river basin and the high-latitudes are highly concentrated (CI close to 1, with mean observation dates in June to August) to the relatively dry months in the Amazon (little chance of cloud cover) and the warm season in the high-latitude (little chance of snow and river freezing/ice), respectively, which possibly reveals the sampling bias with optical RS data. In comparison, Landsat observation dates are much more dispersed throughout the year for other global regions (CI <0.6 in Fig. 4e, and the mean observation dates vary across regions in Fig. 4d).

2.4. The Bayesian AMHG-Manning (BAM) algorithm and its recent variant geoBAM

We used two algorithms, i.e., BAM and geoBAM, for Q inversion with river width data obtained from RS ($W_{i,t}$). The two algorithms utilize the same flow law and Bayesian underpinning to probabilistically estimate Q_t – one states the likelihood (a sampling model for $W_{i,t}$ conditional on the non-RS parameters) and the priors (the non-RS terms of equations in Appendix) as probability distributions. Then these two terms are sufficient to describe the joint posterior distribution for inferring Q. To approximate the distribution, a Hamiltonian Monte Carlo sampler is

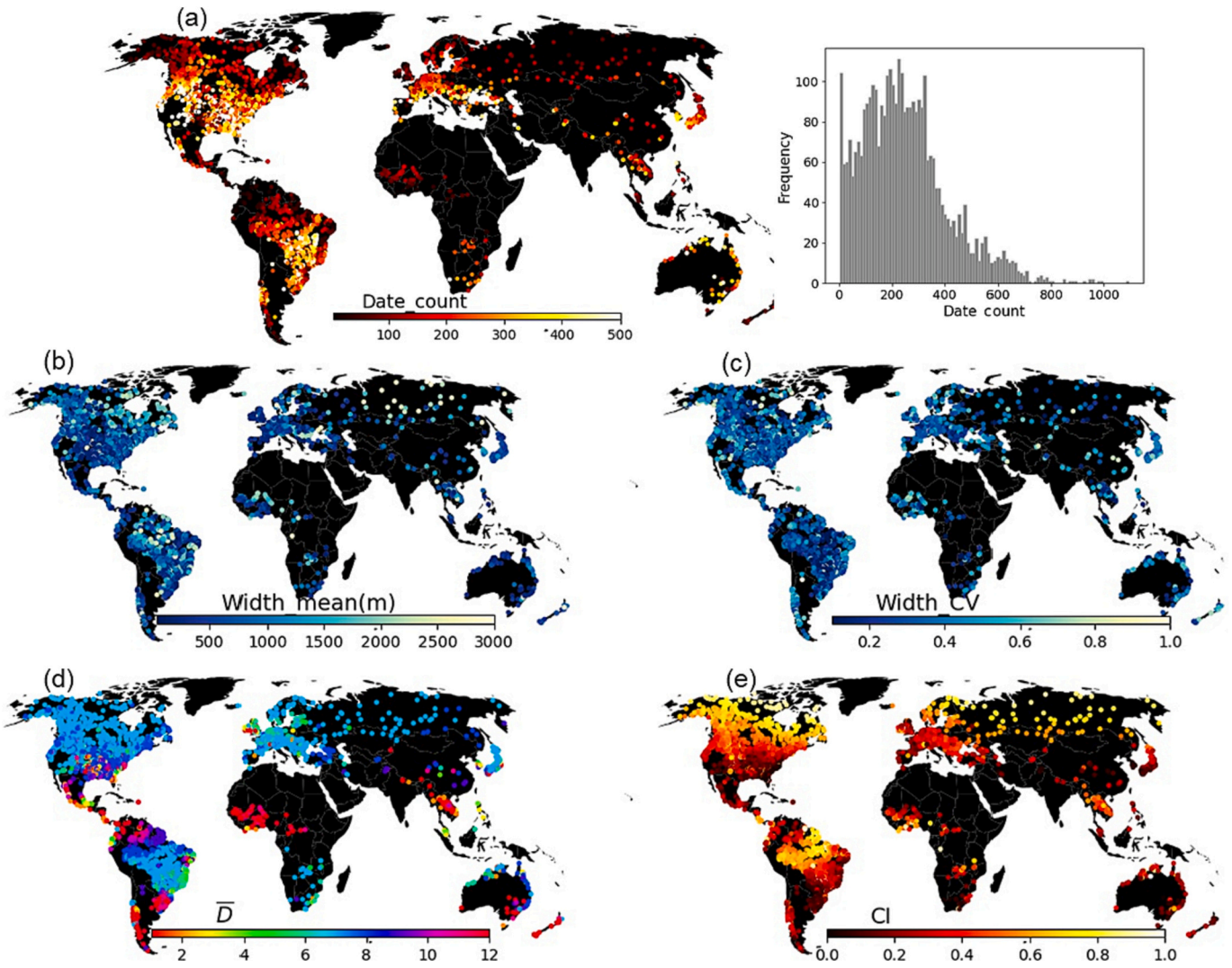

Fig. 4. Reach-averaged spatio-temporal characteristics of the 357,389 Landsat scenes. (a) shows the number of scenes at each gauge (*Date_count*) and the histogram, (b) and (c) show the mean (*Width_mean*) and the coefficient of variation (*Width_CV*) calculated from all observations. (d) and (e) show the mean observational date (*D_bar*, expressed as month number) and its concentration index (*CI*).

adopted for efficient high-dimensional sampling. Beyond the sampling, BAM/geoBAM are largely similar to the other Bayesian McFLIs such as MetroMan (Durand et al., 2014) and GaMo (Garambois and Monnier, 2015), except that it adds AMHG (Gleason and Smith, 2014) as an alternative of the widely used Manning's equation, eliminating the need for stage/slope observations, making it possible to invert Q with width-only data. BAM/geoBAM code packages are openly available (see Acknowledgement). The physics of BAM/geoBAM are described in detail in Appendix.

Priors formalize a priori estimates and uncertainties for non-observable parameters, and they represent the initial understanding of the distribution of these terms. Two types of priors are fed into BAM/geoBAM, namely Q priors and hydraulic geometry priors; the former can be derived from global hydrologic modeling (see below), and the latter can be derived from river geomorphic attributes. To mimic ungauged scenarios, we assumed no gauge data existed for Q_t, different from strategies by Brinkerhoff et al. (2020) and Feng et al. (2021) where all gauged data were used to constrain the inversion when applicable. Following this logic, we used the Global Reach-level A priori Discharge Estimates for SWOT (GRADES) (Lin et al., 2019) and the updated Global Reach-level Flood Reanalysis (GRFR) (Yang et al., 2021) to formalize Q_t priors. GRADES/GRFR leveraged recent advancements in global hydrologic modeling to estimate Q worldwide, and it provides one of the best-verified and openly accessible Q priors based on updated hydroclimatic knowledge for 2.94 million reaches. Although GRADES/GRFR were calibrated and bias-corrected, no direct gauged observations were included when formulating the model; thus, they are appropriate to use as the priors for ungauged rivers. To state the truncated log-normal distribution needed for Q_t and Q_c, the priors include Q_{mean}, Q_{min}, Q_{max}, and Q_{CV}, where the 'off-the-shelf' BAM/geoBAM takes Q_{mean} as the Q climatology of the prior model and Q_{CV} as 1.

The geomorphic priors came from the hydraulic training datasets used when constructing BAM/geoBAM (Canova et al., 2016), which we did not make changes to. In geoBAM, rivers were classified with an unsupervised and a bespoke "expert" scheme, where different geomorphic priors were assigned to the river classes. The unsupervised scheme used DBSCAN to cluster the rivers into 8 classes, and the expert scheme subjectively chose 17 classes to make river width a strong predictor of the river types (ref. to Fig. 2 by Brinkerhoff et al. (2020)). In our initial experiments, we found the unsupervised scheme failed to represent river variabilities worldwide (>90% of rivers classified as Type 1; not shown), suggesting its limitation as also cautioned by Brinkerhoff et al. (2020) for its pure statistical realization. Thus, we used the expert scheme for the ensuing results and interpretations of this scheme for the Q inversions are provided in Section 3.1.

To infer Q, mass conservation is a necessary assumption as lateral

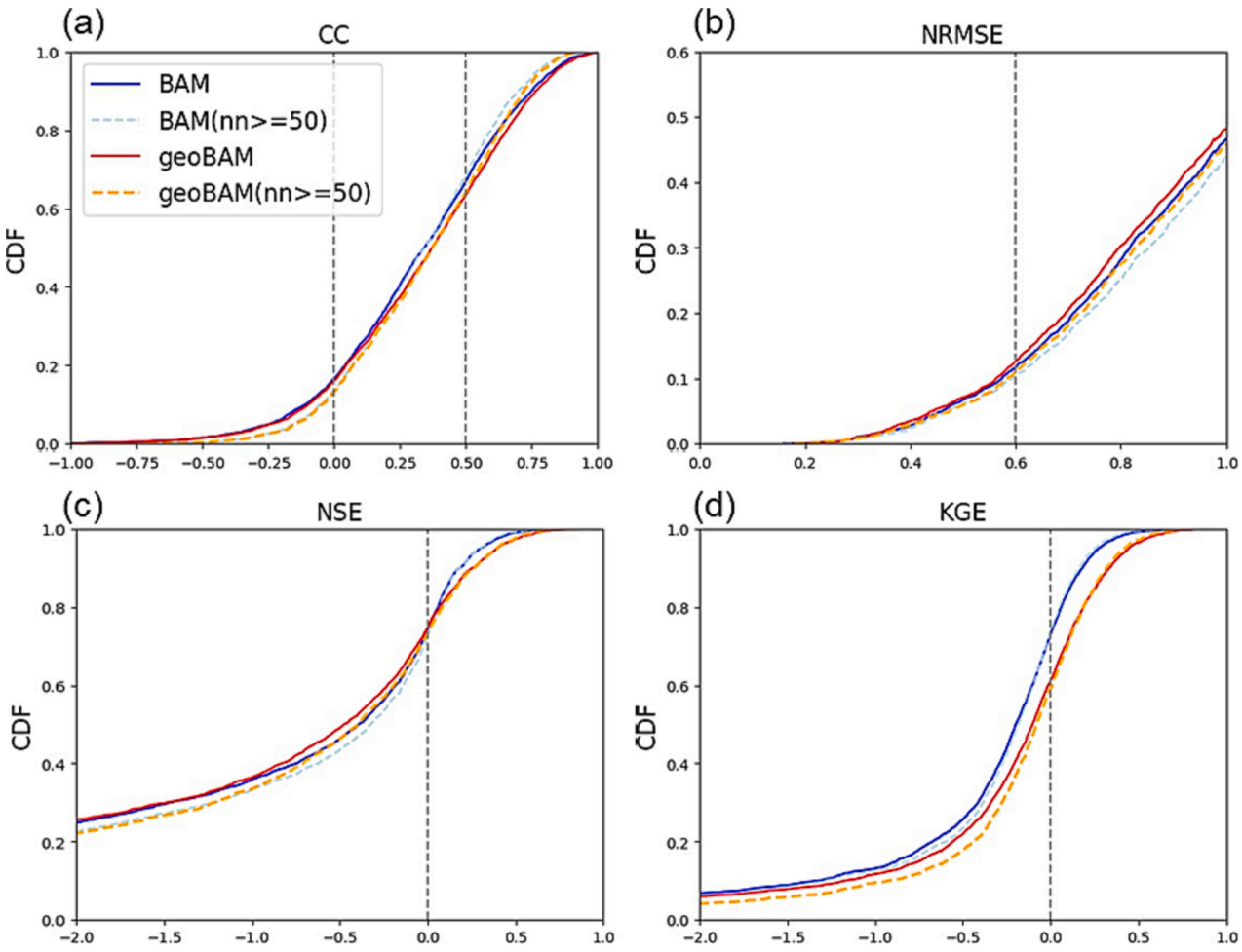

Fig. 5. Cumulative density function (CDF) of the skill metrics for BAM/geoBAM. (a) for correlation coefficient (CC), (b) for normalized root-mean-squared-error (NRMSE), (c) for Nash-Sutcliffe Efficiency (NSE), and (d) for Kling-Gupta Efficiency (KGE). We used at least 3 pairs of the inverted and observed discharge for calculating the skill metrics (3078 locations), while "nn ≥50" used at least 50 pairs for calculating the skill metrics (1691 locations). We note KGE = 0 is used as a stricter threshold to delineate good/bad inversion skills despite studies suggesting KGE > –0.41 can already denote useful skills.

inflows can interfere with the results (Nickles et al., 2020). This assumption is valid here as we segmented global rivers into reaches of ~6–7 km long with a 25 km^2 channelization threshold. This means only tributaries draining ≤25 km^2 are ignored, which better conserves mass compared to Durand et al. (2016) that has a reach length of 11 to 223 km. For meaningful *Q* inversion with BAM/geoBAM, one needs at least five cross sections at a given location/date to derive AMHG. For fair statistical evaluation, one needs at least three pairs of overlapping Landsat widths and observed *Q* (sensitivity to this criterion is assessed in Section 3.1). The screening led to 3078 gauges for final evaluation.

3. Discharge inversion using 'off-the-shelf' BAM/geoBAM configurations

3.1. Summary skill metrics across all locations

We ran BAM/geoBAM at 3078 gauges, and summarized the skill metrics of *Q* inversion using cumulative density function (CDF) plots in Fig. 5; the skill metrics calculated against daily *Q* observations included correlation coefficient (CC), normalized root-mean-squared-error (NRMSE), NSE, and Kling-Gupta Efficiency (KGE), as they offer insights into the RSQ inversion from different angles, e.g., CC for temporal variability, NRMSE for biases and standard deviation errors, and NSE/KGE for the overall skill. We followed Frasson et al. (2021) in choosing these metrics, and they were defined therein.

Fig. 5 shows the baseline RSQ performance for all locations, where the CC distribution is largely similar to Van Dijk et al. (2016) who assessed monthly *Q* estimates with MODIS measurements. BAM/geoBAM performance is seemingly unsatisfactory due to the many negative NSE/KGEs, but these match with our expectation as we provided the sternest criteria to include all possible river types worldwide and did not tailor the algorithm in any way between rivers. Despite this worst-case scenario, we still see ~1000 successful cases, e.g., reasonably captured daily discharge at 33–37% of locations with CC > 0.5 (Fig. 5a) and 27–39% of locations with KGE > 0 (Fig. 5b), which shows promise for large-scale RSQ. Note that we used zero-KGE to delineate good/bad performances, but studies have suggested that KGE > –0.41 can denote informative performance than the mean particularly when NSE alone (Fig. 5c) can be limited in interpreting the results if the observation variability is low (Knoben et al., 2019). This means the criteria on KGE > 0 may be relaxed especially considering the challenging situations in *Q* estimation in ungauged basins (Sivapalan et al., 2003), which shows promise for global RSQ.

Looking at the component statistics, we found the inversion accuracy is mainly compromised by model biases (i.e., only <20% of gauges show an NRMSE <0.6, Fig. 5b), and this behavior of RSQ error dominated by bias rather than temporal variability error has also been highlighted by recent SWOT studies (Frasson et al., 2021, 2022; Tuozzolo et al., 2019). This behavior is different from hydrologic modeling, where temporal variability error is often considered more difficult to cope with - the

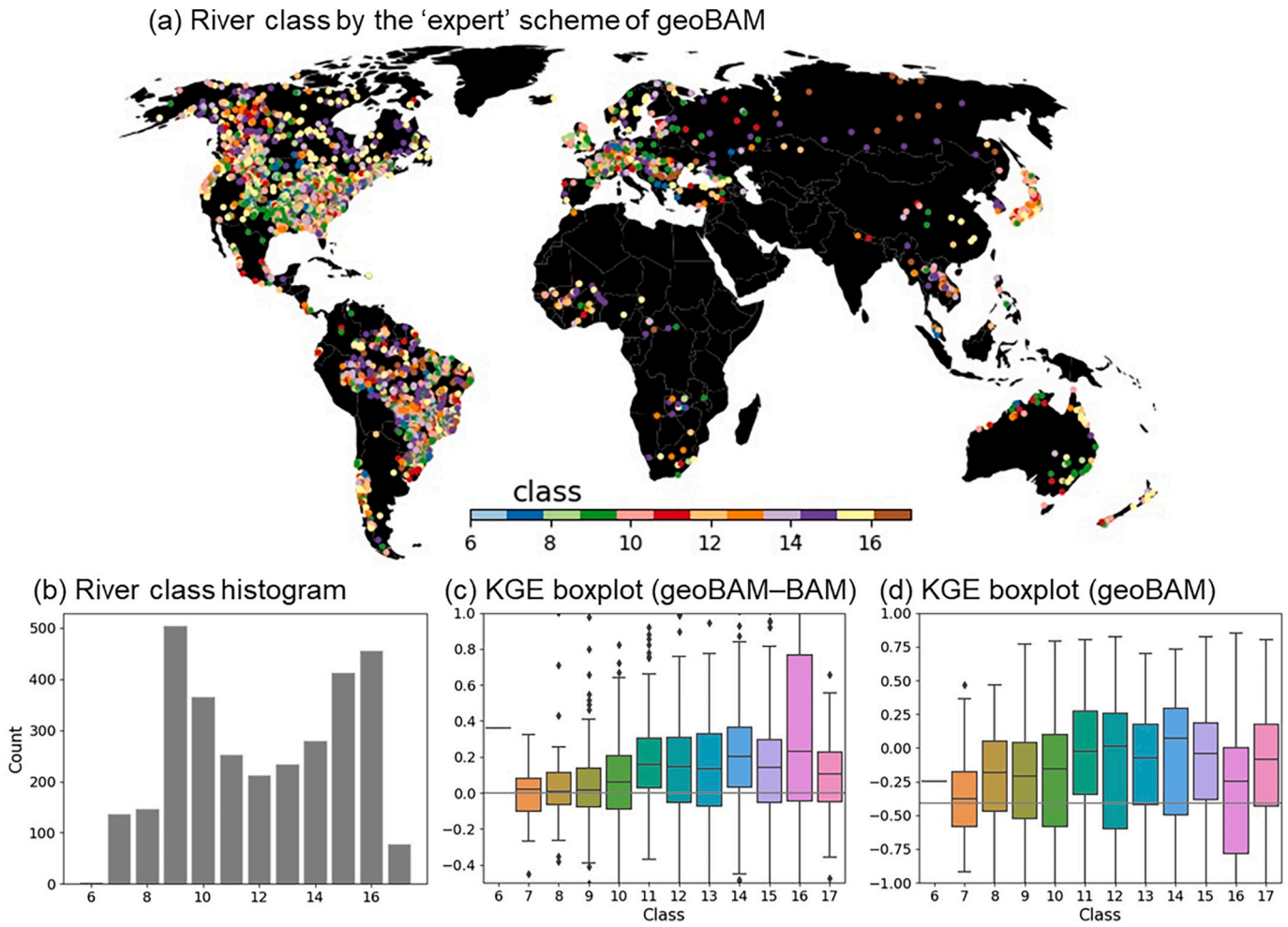

Fig. 6. River classes determined by geoBAM and its inversion accuracy. (a) and (b) show the river class spatial map and the histogram; (c) shows the KGE boxplot calculated as KGE_{geoBAM} minus KGE_{BAM} to understand geoBAM's improvements upon BAM (horizontal line shows $KGE_{diff} = 0$; positive values mean geoBAM is better than BAM); (d) is the geoBAM KGE boxplot against the river class (horizontal line shows KGE = –0.41). Class 6 has only three cases, thus not sufficient to shown on (c) and (d).

uncertainty cascades in hydrologic modeling are dynamic and cannot be addressed using effective bias correction techniques (Lin et al., 2019). Such complementary strengths of RSQ and hydrologic modeling imply that merging of the two will become increasingly useful in the future (Feng et al., 2021; Ishitsuka et al., 2021). We also tested the sensitivity of the skill metrics to the threshold of a minimum of 3 pairs of inverted and observed discharges by increasing it to 50 pairs (1691 out of 3073 gauges in Fig. 5). The latter shows slightly better KGE/NSE when biases are negative (likely due to the high sensitivity of bias to small sample sizes) and when CC is low (likely due to the difficulty in capturing temporal variability with large sample sizes), but the overall sensitivity is small. Thus, in the later sections we consistently used locations with >3 pairs of inverted and observed discharge to increase the spatial coverage of our assessment. We also found geoBAM is mainly superior to BAM in the bias component rather than the correlation component – recall geoBAM is only different in its geomorphic priors, and this suggests that better stated geomorphic priors are key to alleviating the bias error.

In geoBAM expert scheme, rivers were pre-classified to different types mainly by their widths – Classes 1–15 refer to small- to medium-sized rivers, Class 16 are the "highly width-variable" rivers, and Class 17 are big rivers (mean width > 665 m) which also lack sufficient samples in its geomorphic training data (more details in Brinkerhoff et al., 2020). According to this scheme, the ~3000 rivers are classified as Classes 6–17 covering a wide geographic extent (Fig. 6a), and Classes 9 & 16 are the dominant types (Fig. 6b). The KGE_{diff} between geoBAM and BAM against river classes (Fig. 6c) shows that geoBAM improves over BAM in all river classes, particularly for Class 16 (i.e., the highly width-variable rivers), but limited gains were found for Class 7–9 (i.e., smaller rivers). There is a tendency for geoBAM to bring more improvements as rivers get larger, but this tendency drops for Class 17 (i.e., big rivers) likely due to the limited training data in this class. Despite the better skill of geoBAM, Fig. 6d shows it still struggles to accurately estimate *Q* for

Table 2
Six selected factors potentially influencing the RSQ inversion accuracy.

Factor	Description
b	At-a-station hydraulic geometry (AHG) *b* exponent, calculated by fitting reach-averaged width and prior discharge in a power law model
AI	Aridity Index (AI) calculated as mean precipitation divided by potential evapotranspiration (Zomer et al., 2008)
Width_CV	Coefficient of variation (CV) calculated as the standard deviation of width divided by mean width
LAI	Leaf Area Index (LAI) of the 1981–2015 annual mean from the Advanced Very High Resolution Radiometer (AVHRR) Global Inventory Modeling and Mapping Studies (GIMMS) LAI3g version 2 (Mao and Yan, 2019)
CI	Concentration index (CI) of the Landsat observation dates across all years
Width_mean	Mean width in the unit of meters

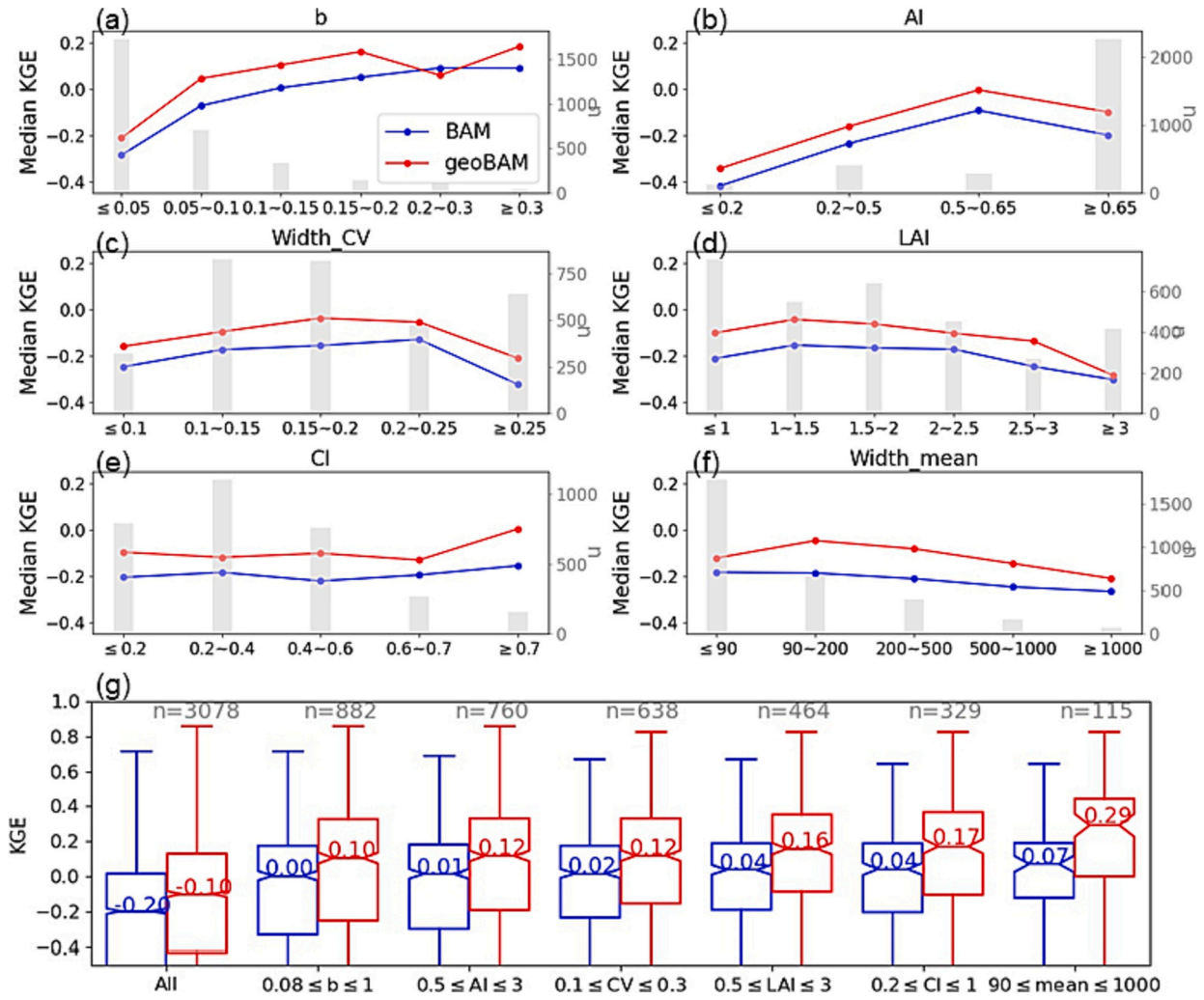

Fig. 7. Exploratory analyses of the BAM/geoBAM inversion accuracy against six selected factors. (a) to (f) show the analyses for *b*, *AI*, *Width_CV*, *LAI*, *CI*, and *Width_mean* (Table 2); the x-axis shows the assessed range, the left y-axis shows the median KGE, and the right y-axis shows the sample size (shown as gray bars). (g) shows the KGE boxplot by incrementally applying filters for one factor at a time to obtain increased inversion accuracy, and the sample size is denoted in n. The KGE ranges on the left y-axes from (a) to (f) are kept the same such that they can better show the sensitivity to each factor.

Classes 16 and 17, and both good (KGE > 0.5) and bad (KGE < 0) performances can be seen in each river class. This implies other factors are still compounding the inversion and better understanding of each river's uniqueness is needed. Based on this result, we assessed a range of other factors possibly influencing RSQ to better understand this question.

3.2. Exploratory analyses on factors influencing RSQ

We selected six factors to assess how they affect the RSQ accuracy (Table 2). Among them, at-a-station hydraulic geometry (AHG) exponent *b* and aridity index (*AI*) have been previously assessed at 34 rivers by Gleason et al. (2014), and *Width_CV* has been discussed to indicate AMHG strengths by Feng et al. (2019, 2021), while we further added the assessment of leaf area index (*LAI*), mean river width, and the *CI* index. The first three factors better reflect the physical properties of a river (*b*, *AI*, *Width_CV*) while others are expected to affect the RS retrieval accuracy more (*LAI*, *CI*, *Width_mean*). We note this assessment is not to be exhaustive but to probe into "what factor can influence the RSQ inversion and how sensitive is the RSQ to them?" Although some factors may be algorithm-specific, this analysis framework can be useful for other RSQs too. In the following section, we elaborate on their expected control on accuracy, and in Fig. 7, the sensitivity of RSQ to these factors is assessed with the median KGE derived from each factor.

Overall, the RSQ exhibits the greatest sensitivity to *b* followed by *AI*, as shown by the greatest slope change (Fig. 7a). Generally, close-to-zero *b* denotes channels of nearly rectangular shapes that lack width variability, thus, these channels show the most unsatisfactory inversion, as also anticipated by Gleason et al. (2014). We show that KGE increases rapidly for increasing *b*, but after $b > 0.15$, the increase in KGE_{median} gradually leveled off. *AI* is a climate regime factor where arid environments tend to show low AMHG strengths (Gleason et al., 2014); here we found AI>0.5 (a divide for sub-humid climate) shows the best KGE, but the overall sensitivity is smaller (Fig. 7b). For *Width_CV*, modest values show the best RSQ; low values can correspond to low-*b* situations and KGE increases as *Width_CV* increases to 0.25, yet we found KGE drops quickly for *Width_CV* > 0.25 (Fig. 7c). This behavior may be attributable to overbank conditions that can introduce compound behavior to the AMHG flow laws, which won over the AMHG strengths to some degree captured by *Width_CV*.

Shifting to factors affecting the RS retrievals that can occur in all river forms, we found that the best KGE is for *LAI* < 1.5–2.5 (Fig. 7d), where too large *LAI* can interfere with width retrievals for optical sensors, and too small *LAI* can indicate dry conditions. High *CI* is associated with seasons of the best RS observational quality with low cloud/snow/ice contaminations, but we do not see prominent KGE changes with increased *CI* (Fig. 7e). Lastly, as the width extraction process was known

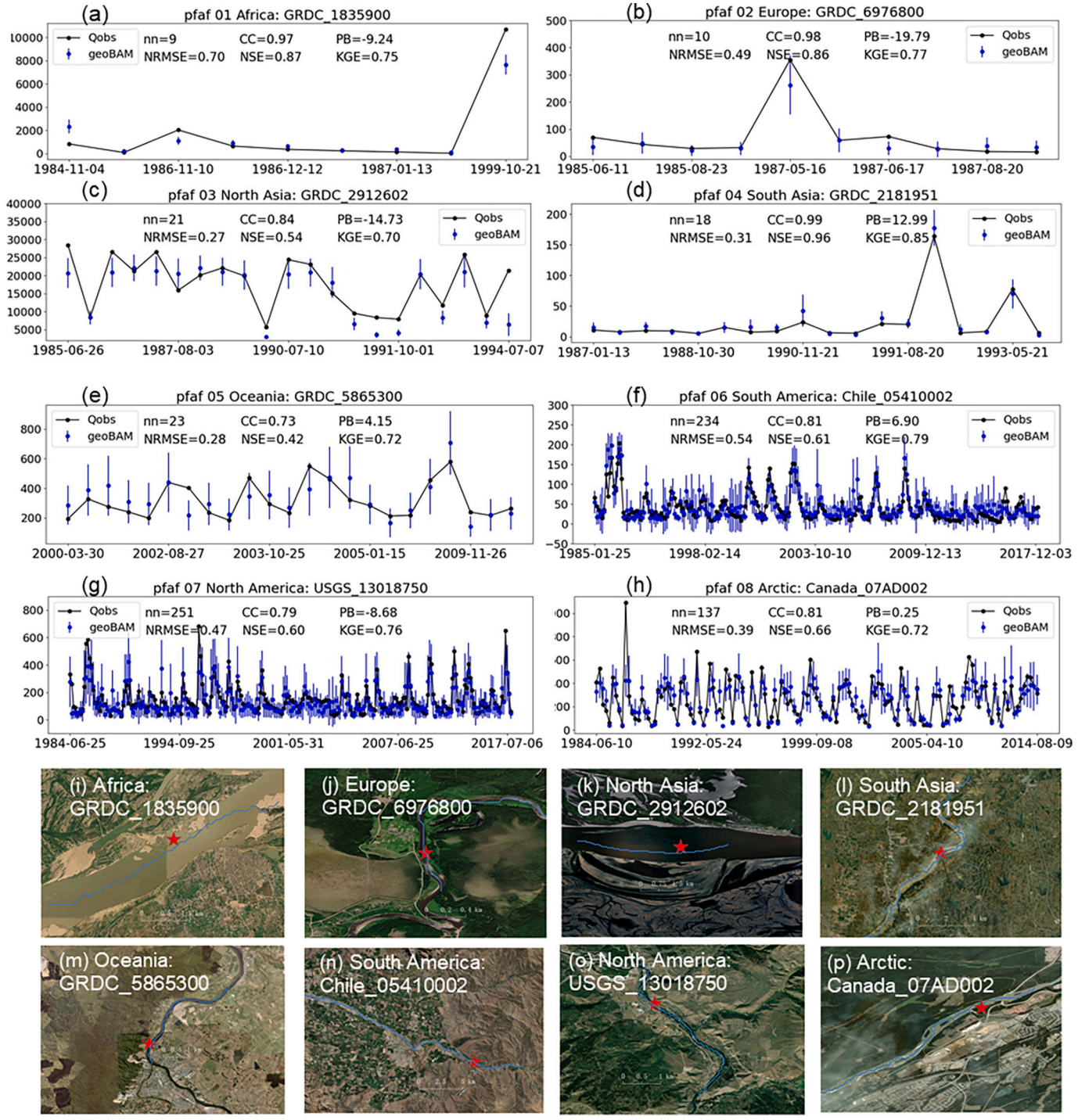

Fig. 8. Example locations exhibiting very good discharge inversion accuracy. For each continent of MERIT Basins, we randomly selected one gauge with KGE > 0.7 in geoBAM for this demonstration. (a) to (h) show the hydrographs, where the texts show the number of inversion dates (nn) and the skill metrics including CC, percentage bias (PB), NRMSE, NSE, and KGE. Error bars show the 95% confidence interval, where uncertainty is estimated by the Bayesian inference. (i) to (h) show the ESRI Satellite Image backgrounds of the gauged river segments.

to show higher uncertainties for rivers narrower than 90 m (Allen and Pavelsky, 2018), we also assessed the RSQ accuracy with regard to *Width_mean*. Our results showed that KGE_{median} peaks for *Width_mean* between 90–500 m, but further increases in *Width_mean* can lead to decreased KGE (Fig. 7f). This seemingly counterintuitive behavior indicates that the width uncertainty becomes a less significant error source for wide rivers. We further examined locations with *Width_mean* > 500–1000 m, and identified several other error sources that can explain this behavior. First, we found the worst KGEs were mostly seen within a lake/reservoir. Recall in Section 2.1 we eliminated reaches with lakes/reservoirs, but we found many low KGE locations were incorrectly flagged as rivers by GRWL possibly because of the difficulty to distinguish between a wide channel and a channel-like reservoir; thus, the biased widths and compound hydraulics of a reservoir-like river were propagated into erroneous discharge estimation. Second, in geoBAM, rivers wider than 665 m were classified as a single class that also lacked sufficient geomorphic training sample (see Section 3.1); the lower KGE here indicates the geomorphic priors became less informative for wide

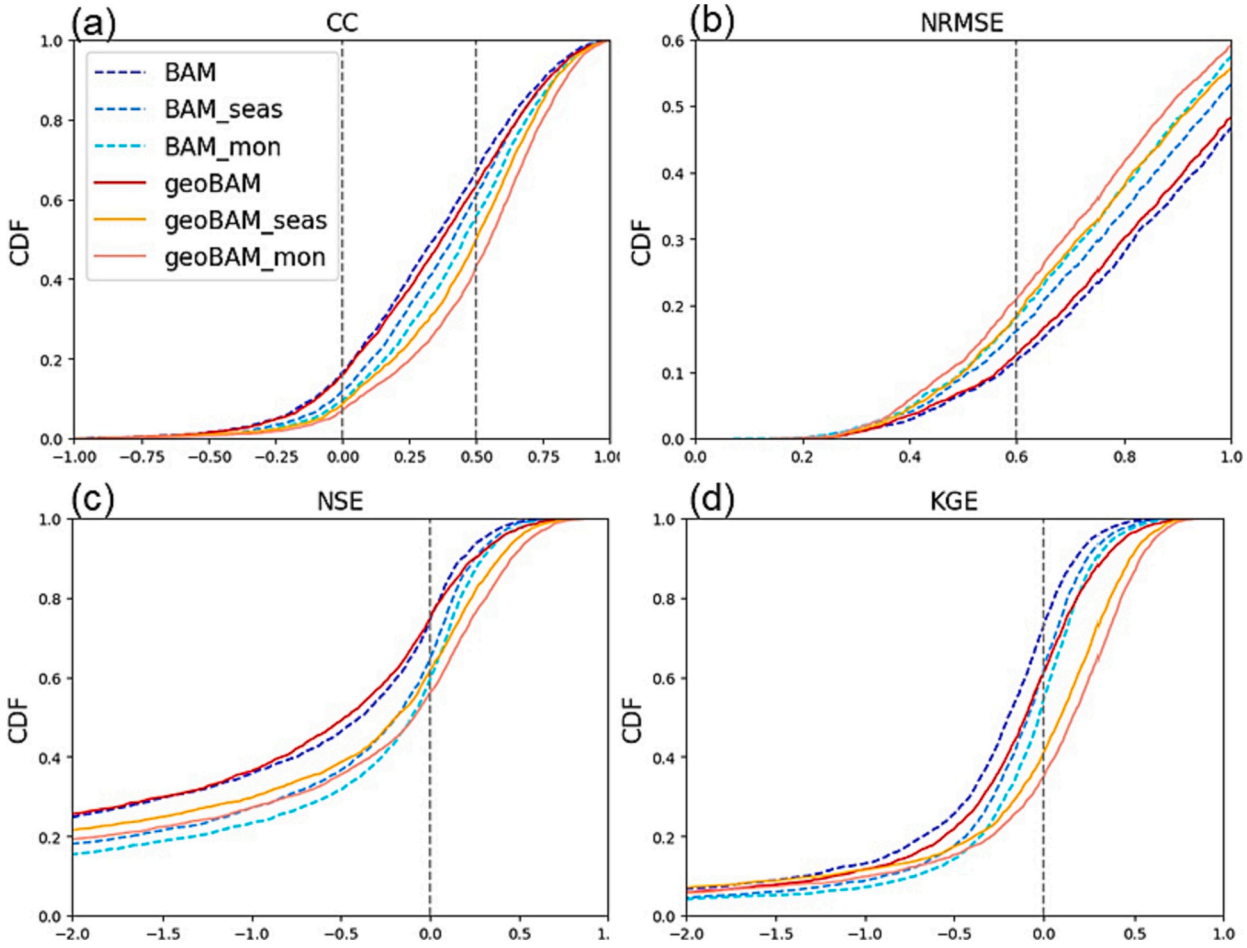

Fig. 9. CDF plots of the skill metrics for BAM/geoBAM by feeding 'seasonal' and 'monthly' Q priors ($N = 3078$). (a) for CC, (b) for NRMSE, (c) for NSE and (d) for KGE, similar to Fig. 4. Cold colors show BAM, BAM_seas, and BAM_mon experiments, respectively; and warm colors show geoBAM, geoBAM_seas, and geoBAM_mon experiments, respectively.

rivers, which is a known but non-prominent issue by Brinkerhoff et al. (2020) and will require further improvement in the future. Third, we found wider rivers also tend to show smaller *b* (Fig. S1). These collectively offer a feasible explanation on why the widest group of rivers does not see further improved RSQ. While it remains difficult to fully derive the mutually exclusive strength as the factor interactions started to show up, here we offer a unique perspective on where to expect better RSQ accuracy based on the prior knowledge of *b*, AI, *Width_CV*, *LAI*, *CI*, and *Width_mean*. The strengths of each factor were further demonstrated in Fig. 7g, where filters were incrementally applied to one factor at a time (x-axis) to gradually increase KGE while accounting for the compromised sample size n; this outlined the optimal conditions that can increase KGE_{median} from −0.10 to 0.29.

3.3. Locations with good inversion accuracy

Fig. 8 shows locations exhibiting very good inversion accuracy; we randomly picked one station with KGE > 0.7 (geoBAM) for this demonstration. These rivers are located in quite different environments, e.g., highly natural (Fig. 8k), highly urbanized (Fig. 8i, m & 8p), limited agricultural activities (Fig. 8j, o), and mountainous regions (Fig. 8n). They are also in different climate zones, but their river discharge has all been well estimated. This confirms the RSQ accuracy based on the Bayesian inversion is not necessarily a function of the physical hydroclimate conditions, but the key to successful inversion is outlined in the factor interactions shown by Fig. 7. We also found that the inversion accuracy is not necessarily lower for smaller mountainous rivers as might be expected (see four locations randomly picked from the Andes, the Qilian Mountain, and the Alps from Fig. S2 for a closer look at RSQ performance in mountainous rivers which has received relatively less attention (Huang et al., 2018)). The fact that both good and bad inversions can be obtained for small mountainous rivers ranging from 55 to 81 m wide suggest that, overall, the successful RSQ requires channel shapes of higher *b* and high *AI* in sub-humid to humid regions, moderate width variability and *LAI* (the order of appearance reflected their relative importance). As long as these conditions are satisfied together within a reasonable way to retrieve the river hydraulic variables, good RSQ inversions can be obtained in different hydroclimatic regions.

4. Improving discharge inversions over the baseline

4.1. The use of enriched discharge priors

In this section, we seek to use openly obtainable geospatial datasets to improve the global RSQ. For challenging locations presented above, one would naturally seek for additional height/slope information from SWOT or other altimetry missions, or from better prior knowledge of river hydroclimatology/geomorphology for optimal inversion. As this study focuses on *Q* inversion with width-only observations, we resort to richer priors to assess how they influence the global RSQ. Here, instead of assuming the known *Q* climatology from the prior model GRADES/GRFR (Section 2.4(2)) with a single set of $\{Q_{mean}$ and $Q_{CV}\}$, we tested the use of two other configurations (i.e., 'seasonal' and 'monthly' priors) to describe the truncated log-normal prior distribution of *Q*. In the

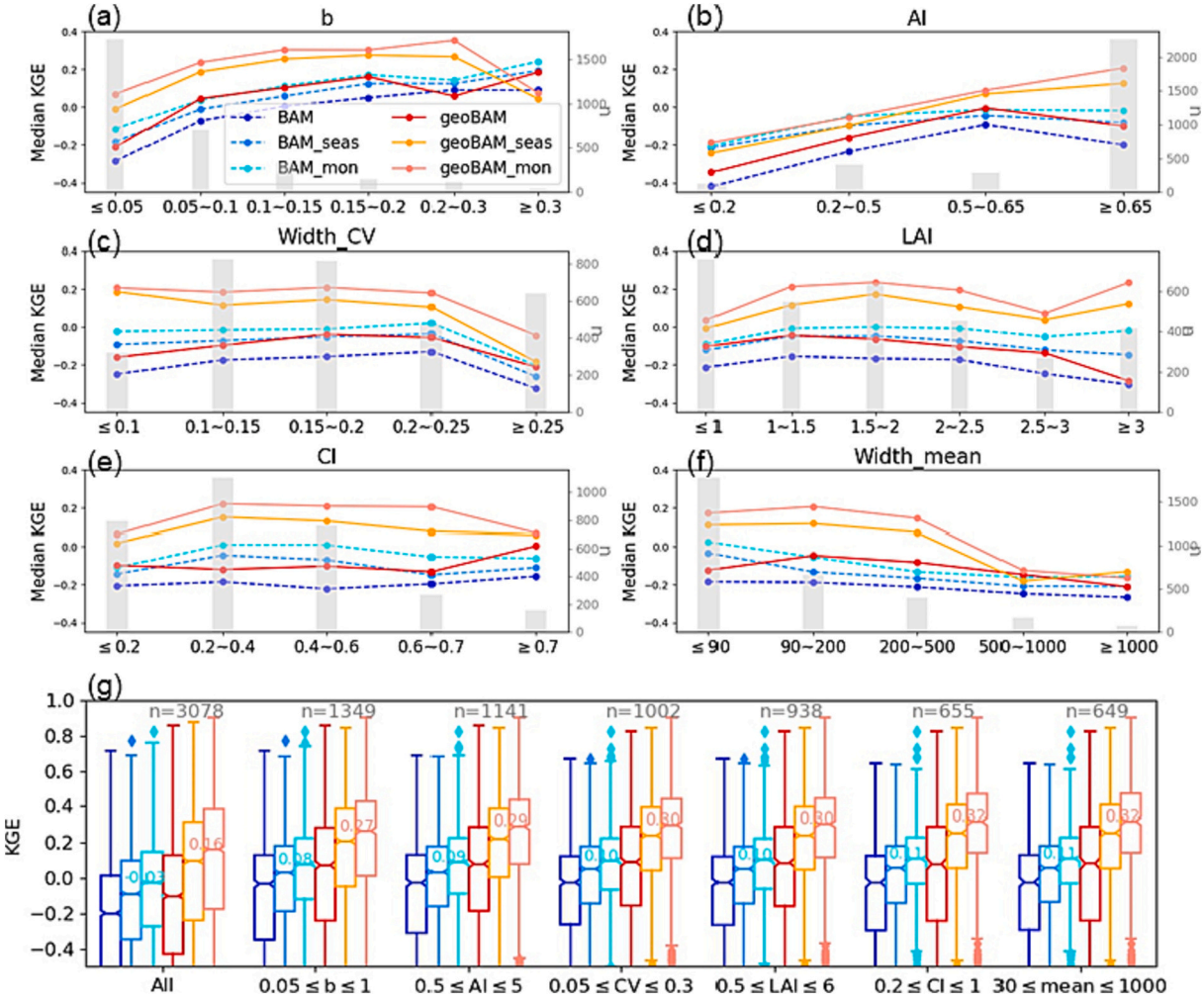

Fig. 10. Exploratory analyses of the BAM/geoBAM discharge inversion accuracy after feeding 'seasonal' and 'monthly' Q priors. The figure is largely similar to Fig. 6. Cold colors show BAM, BAM_seas, and BAM_mon experiments, respectively; and warm colors show geoBAM, geoBAM_seas, and geoBAM_mon experiments, respectively.

'seasonal' configuration, seasonal averages and CV (i.e., four unique values) of the 40-yr GRADES/GRFR *Q* priors (i.e., March–April-May, June–July-August, September–October-November, and December–January-February) were served to {Q_{mean} and Q_{CV}}. Analogously, in the 'monthly' configuration, monthly averages and CV (i.e., 12 unique values) of the 40-yr *Q* priors were served to {Q_{mean} and Q_{CV}}. The rationale behind using these priors is that the flow seasonality and monthly variability from hydrologic modeling provides complementary strengths to RS observations, which is also the philosophy of data assimilation (DA). BAM/geoBAM represent flow as a truncated lognormal distribution, and as a Bayesian process the more accurately we can represent the prior within the distribution, the more accurate the posterior will be. We gradually increase the information content from the 'default', 'seasonal', to 'monthly' configurations with the aim to better understand how to best describe the distribution of the priors versus the observations – the global model is more accurate when aggregated over longer times, but finer time resolution better resolves the distribution. Thus, the configurations below should allow us to take a closer look at this question.

In Fig. 9, the sensitivity of BAM/geoBAM to the enriched priors were revealed. It is seen that the 'monthly' and 'seasonal' configurations consistently outperform the 'default', matching the expectations that when RS data is uninformative, BAM/geoBAM reverts to the prior and ensures valuable dynamic information present in 'seasonal' and 'monthly' *Q* priors preserved. These setups increased the percentage of locations with CC > 0.5 from 33 to 37% (in BAM/geoBAM) to 44–57% (in BAM_mon/geoBAM_mon; the same format is consistently used later). In addition, percentage of locations with KGE > 0 increased from 27 to 39% to 46–65%, suggesting that the successful daily *Q* inversions now increased to 1400–2000 gauges.

Performing similar exploratory analyses in Fig. 10, we show that the 'monthly' and 'seasonal' configurations improve the inversion for nearly all conditions, particularly for low-*b* channels (see leveled-off lines in Figs. 10a&e), while locations with large *Width_CV* and *Width_mean* remain to show low inversion accuracy (Figs. 10d&f). This confirms that the error sources for *Width_CV* > 0.25 and *Width_mean* > 500–1000 come from the overbank flow conditions and the challenging geomorphic prior setups with geoBAM (as discussed in Section 3.2), which is not expected to be resolvable by enriched *Q* priors. Accordingly, we suggest future geoBAM improvements to be tailored to these rivers.

Interestingly, we find that for high-*b* channels, supplying richer *Q* priors can degrade the geoBAM inversion accuracy (Figs. 10a) – note that *b* is a channel geomorphological indicator where high-*b* values indicate width is a strong predictor of *Q* with no lateral confinement, and in this case less informative priors should be used. Running the filters similar to Fig. 7g shows that constraining factors to their optimal ranges can help improve the daily KGE_{median} of geoBAM to 0.32 while retaining a sample size of 649 (Fig. 10g). These outline the conditions for RSQ to be successful – for the optimal case on the rightmost column of Fig. 10g, gauges with a daily KGE > 0.2 take up 68.9%. We compare this high

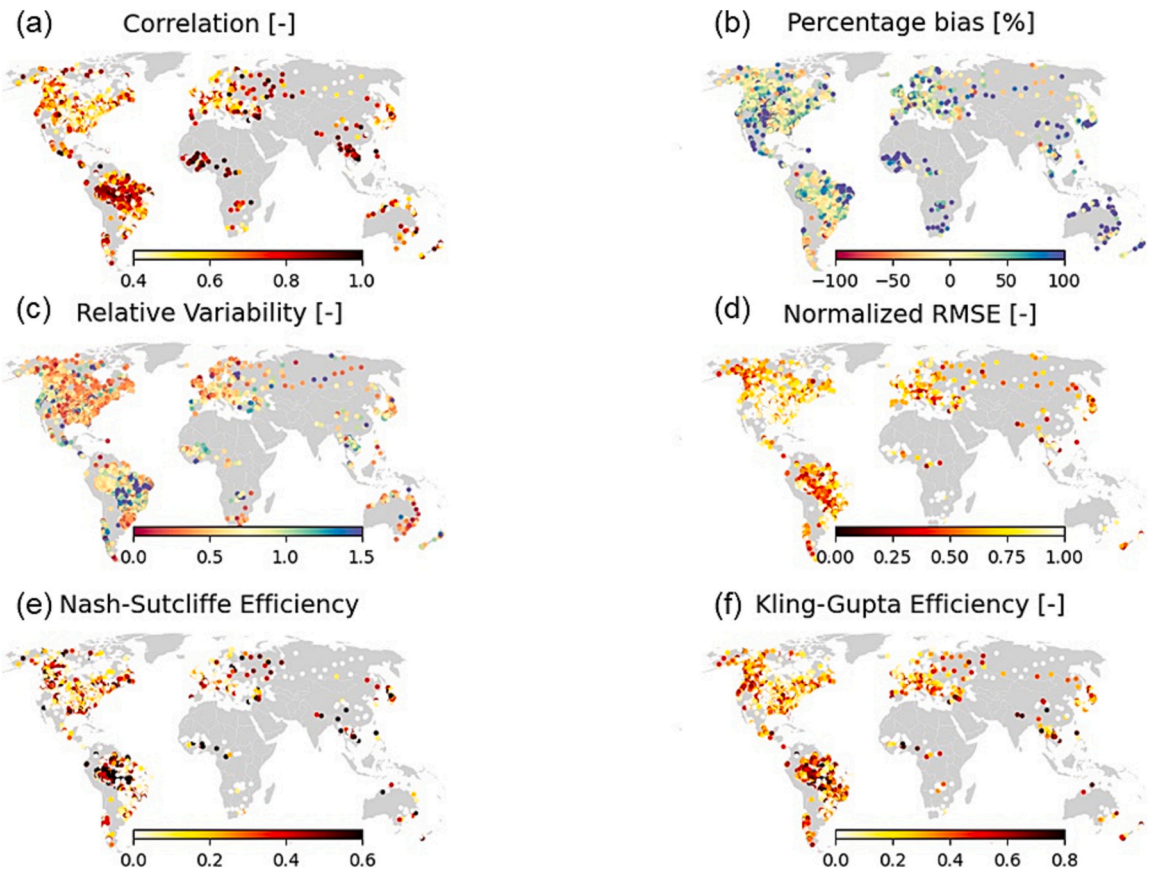

Fig. 11. Spatial distribution of the RSQ inversion accuracy for the *geoBAM_mon* configuration at daily time scales. (a) to (f) show CC, PBIAS, RV, NRMSE, NSE, and KGE, respectively. (a) to (c) show the component statistics for KGE to offer better interpretations of KGE.

percentage with existing large-scale RSQ studies, and find it better than those using MODIS-only (Hou et al., 2020) or combined MODIS-microwave measurements (Van Dijk et al., 2016). The high percentage of daily KGE > 0.2 also complements global hydrologic modeling efforts such as that by Lin et al. (2019), suggesting the combined strengths of RS (i.e., ingesting river width in here) and hydrologic modeling (i.e., ingesting prior knowledge on hydroclimate seasonality/monthly variability here) – this desirable RSQ performance holds further potential for discharge estimation in ungauged rivers.

Fig. 11 shows the RSQ inversion accuracy for the *geoBAM_mon* configuration ($N = 3078$), which offers a geographically explicit understanding of locations with good inversions and those remain challenging. In addition to the integrated metrics such as NSE and KGE (Figs. 11e&f), we also presented the component statistics of KGE (a–c). Fig. 11a shows the highest CC is located in the Amazon, west Africa, Southeast Asia (tropics), eastern Europe, and Canada (high-latitudes), where limited CC in parts of the US may be ascribed to the smaller-than-observed relative variability (Fig. 11c). Gauges that show PBIAS within ±35% take up ~54.5% of all locations, and those with NRMSE <0.6 take up 20.8% for all locations (34.2% for their optimal factor range). High PBIAS and NRMSE are generally seen in arid to semi-arid environments, matching with where prior Q is highly biased (Figs. 11b & d). The *geoBAM_mon* priors significantly improved the inversions for the Amazon River Basin, which consistently showed high biases in the 'off-the-shelf' BAM/geoBAM (Fig. S3) possibly related to the high LAI there. For places where hydrologic modeling struggles to provide good CC such as the high-latitude Arctic and regions influenced by human regulations, RSQ can be more skillful in the temporal dynamics (compare Fig. 11a with Fig. 7a in Lin et al., 2019). This imply that direct RS observations may be better at capturing processes like permafrost freeze/thaw, snow/ice melt, and dam regulations through propagating river widths information to Q – these processes are often difficult to be parameterized well in a physically-based model. The similar complementary strengths of RS and modeling were discussed in two recent studies (Hou et al., 2020; Feng et al., 2021), and here our assessment provides a clearer

Table 3
Summary information of two selected gauges and the KGE values for their RSQ. The best performing experiment in terms of KGE is bolded.

Station ID	*b*	*AI*	$Width_{CV}$	LAI	CI	Width_mean	*CC (lnW, lnQ)*	$Width_{CV}/Q_{CV}$
Brazil_ 14,110,000	0.05	2.29	0.20	4.31	0.28	853.28 m	*0.09*	*0.26*
Brazil_ 46,150,000	0.06	0.46	0.09	1.82	0.29	500.69 m	*0.75*	*0.15*

Station ID	KGE					
	BAM	BAM_seas	BAM_mon	geoBAM	geoBAM_seas	geoBAM_mon
Brazil_ 14,110,000	−0.38	−0.12	−0.09	−0.02	0.39	**0.48**
Brazil_ 46,150,000	−0.13	0.20	0.21	0.00	**0.63**	0.52

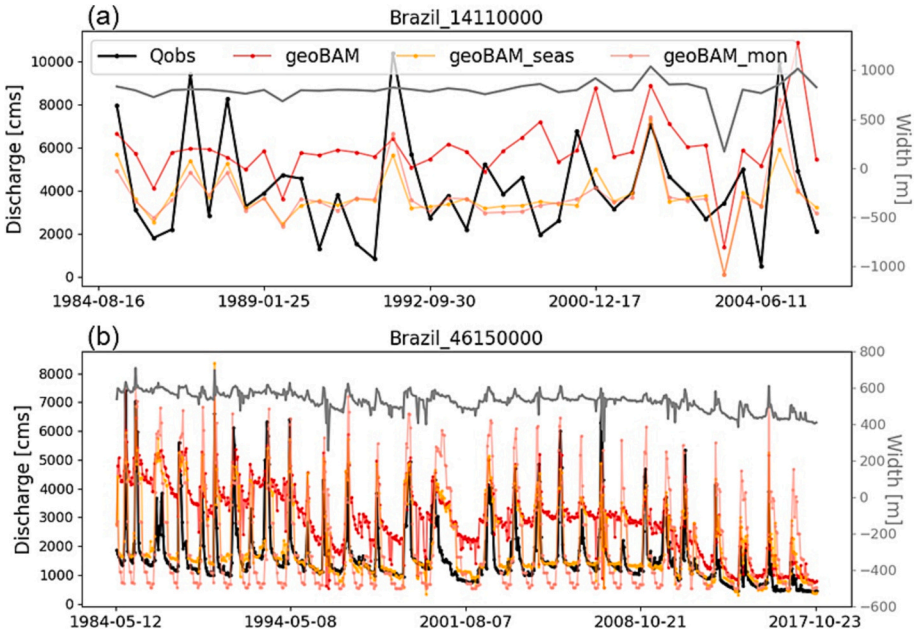

Fig. 12. Hydrographs of the daily discharge inversion for the geoBAM (red), geoBAM_seas (orange), and geoBAM_mon (pink) experiments. Only geoBAM inversion results are presented here for clarity. Width and discharge observations are shown in gray and black lines, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

assessment on this question.

4.2. The optimal combination of Q priors and RS observations

We further show some examples to discuss how the RS data and priors should be optimally combined for RSQ at different locations. Table 3 summarizes the information of two selected gauges in Brazil that consistently showed the worst performance in the original BAM/geoBAM. For both locations, CV ratios of width to discharge (last column of Table 3) are <0.3, suggesting their *W* vary much less than *Q*. But the CC between log-transformed *W* and *Q* for the first gauge is low (~0.09), indicating little useful temporal *Q* information in *W*, while CC for the latter is high (0.75).

For the first location *Brazil_14,110,000* (Rio Negro River, 1.2153°S, 66.8525°W), the limited RSQ accuracy is mostly due to the rectangular-like shape of channel geometry, together with the high *LAI* (~4.3, the Amazon rainforest) that blocks the optical sensors to capture the *W* variability; for this case, consistently improved RSQ was obtained with richer *Q* priors (Fig. 12a). For the contrasting example of *Brazil_46,150,000* (São Francisco River, 12.1825°S, 43.2231°W), richer *Q* priors led to improved inversion accuracy by allowing for more flow variability, but the behavior of *geoBAM_seas* outperforming *geoBAM_mon* also showed up (see the best KGE in Table 3).

We closely examined the hydrographs in Fig. 12, and found that *geoBAM_mon* (pink line) may have provided over-confident *Q* priors on the low/high flows, which resulted in the sub-optimal RSQ when the priors are biased but RS observations are informative. We use the dry months to elaborate on the problem. For example in August, *geoBAM_mon* stated the *Q* prior to have a mean of ~721 m^3/s and a tight CV of 192 m^3/s, while *geoBAM_seas* stated it to have a greater mean of ~1148 m^3/s with less confidence (i.e., a greater CV of 513 m^3/s). Under these setups, in the 1980s when the low flows were relatively high (~1500 m^3/s), *geoBAM_mon* under-estimated *Q* (pink lines) but *geoBAM_seas* estimated it reasonably (orange lines); from 2014 onwards when low flows were observed, both reasonably estimated the discharge to be ~600 m^3/s. The difference is ascribed to the wider *Q* prior distribution of *geoBAM_seas*, allowing for more RS information to be ingested based on the Bayes theorem. Similarly, *geoBAM_mon* almost always overestimated the high flows when it should be relatively low (i.e., peaks of <4000 m^3/s, black lines for the gauge observation). These over-confident prior distributions of *Q* for low/high flows explain why richer priors do not guarantee improved RSQ, suggesting better strategies such as the inflating the prior uncertainty estimates need to be considered for future optimization of the global RSQ.

In light of the above, we further assessed the locations where richer *Q* priors do not lead to improved RSQ, as preliminarily outlined by Fig. S3. Among them, the high-latitude rivers are prominent, and we believe it is related to the challenges with the prior model GRADES/GRFR that struggles to fully capture the flow variability on cryosphere processes such as frozen soil freeze/thaw, and glacier/snow melt. Therefore, the bulk water balance aggregated over longer times yields a better prior as the model cannot finely resolve these processes. We also note that richer *Q* priors almost always degraded the inversion accuracy for locations immediately downstream of a reservoir/lake (see KGE values in Fig. 13; the degradation by *geoBAM_mon* can be more than *geoBAM_seas*). This is because the prior model GRADES/GRFR did not account for dam/reservoir regulations, thus propagating the wrong flow variability to the inversion. In this case, the bulk water balance (or the less temporally resolved priors) also yields better results.

Interpretations of these cases helped to outline a clearer picture on how we can optimally combine the RS observations and the priors to improve the generalizability of global RSQ. For example, we suggest using inflated distributions for the monthly priors, which may help better capture the flow extremes. Additionally, we suggest global rivers to be more explicitly labeled such that less informative *Q* priors should be used at places with known problems of a particular prior model. Here, rivers with runoff contributed from cryosphere processes or those immediately downstream of a reservoir should be labeled to use less informative priors, while other prior models may have similar or different error structures, depending on how the prior model was constructed. Lastly, one should also account for the physical properties of rivers, e.g., by nature, rivers of certain geomorphological types (e.g., incised rivers) should have limited *W* variability; for this case, any additional prior knowledge can be key to the *Q* estimation. By contrast, rivers of high-*b* values (e.g., those with high width-to-depth ratio) can have *W* as a strong predictor of *Q*, and for this case uninformative priors

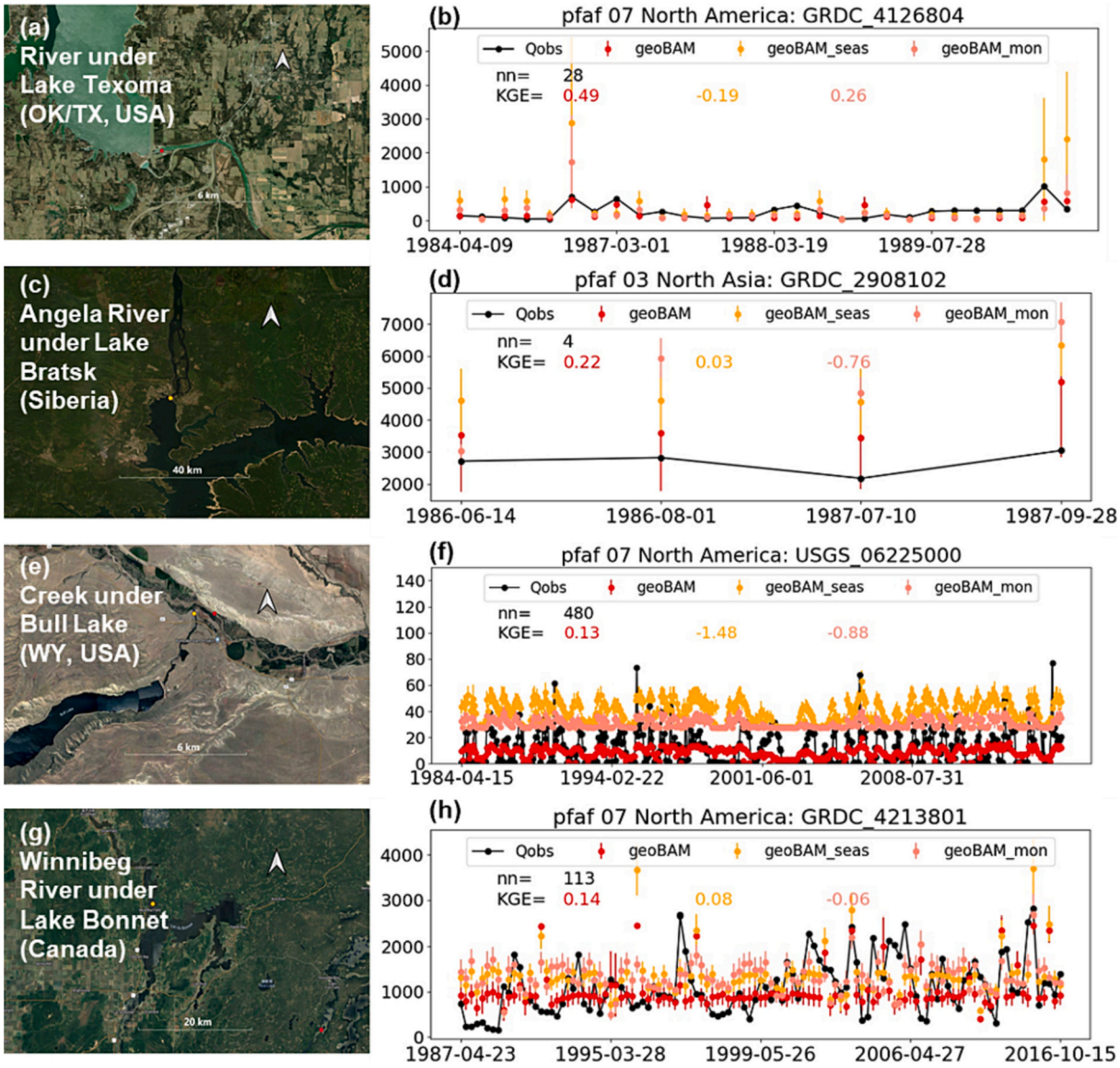

Fig. 13. The RSQ inversion accuracy for four randomly selected locations immediately downstream of a reservoir/dam. The left column shows the ESRI satellite images of these locations; the right column shows the RSQ time series (black: gauge observations; red: geoBAM; orange: geoBAM_seas; pink: geoBAM_mon). The KGE for each prior configuration is printed in the same colour as the hydrograph. Note the strong control of the prior on the bias of the output in panel f. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

should be used, where explicitly labeling of high-*b* channels can lead to optimal inversions. Although the RSQ problem for SWOT will be different from this assessment as one would obtain river height/slope information, we believe the suggestions on how to configure the priors and RS observations for optimal inversion will hold true for the SWOT settings (e.g., using wider distribution, labeling rivers, accounting for the river's intrinsic geomorphologic type). In the future, a clearer understanding of this question will be possible by using increasingly available geospatial datasets, such as hydrologic modeling (Lin et al., 2019), or new soil, vegetation, geology (Beck et al., 2015), or geomorphic datasets (Brinkerhoff et al., 2020) to update our hydroclimatic/geomorphologic/hydraulic understanding of global rivers, and we believe filling up the knowledge gap is the important next steps for the SWOT era.

5. Conclusions and discussions

We extracted multi-temporal river widths from 357,389 Landsat scenes (1984–2019) and used them for daily Q inversion with BAM/geoBAM for evaluation at 3078 sites. The large spatial coverage and diversity of rivers presented a stern test, which promoted the understanding of the promising status of global RSQ and the remaining challenges that require further attention. We suggest this framework is useful for all RSQ, including SWOT, with the following main findings:

- Using 'off-the-shelf' BAM/geoBAM configured with width-only RS observations, we found 33–37% of the three-thousand evaluation sites show daily correlation >0.5 and 27–39% show positive KGEs. This number reaches 46–65% with richer *Q* priors on flow seasonality and monthly variability, which translates to 1400–2000 successful inversions. As our evaluation setup mimics ungauged scenarios and that past studies have suggested KGE > –0.41 already denote useful model skills than the benchmark, the results here with the stricter KGE threshold demonstrate a conservative estimate of the skill, showing great promise for the global scale RSQ.

- By probing into factors influencing the RSQ accuracy for BAM/geoBAM, we found that the inversion is the most sensitive to the channel geomorphological factor *b* and climatic aridity, where optimal conditions for inversion include high-*b*, sub-humid to humid environments, moderate variability in width, *LAI*, and mean width. By applying filters on one factor at a time to gradually increase KGE, favorable RSQ conditions were outlined and the complementary strengths of RS and hydrologic modeling started to appear.
- The results prove the generalizability of BAM/geoBAM, while pointing to the condition of rivers that requires future improvements (e.g., low-*b* channels, arid climate, low/high width variability, high *LAI*, and wide rivers; the wide river group calls for a better differentiation for channel-like reservoirs, as well as improving the geomorphic priors in geoBAM). We suggest other RSQ algorithms that claim global applicability also undergo similar large-scale evaluation as data become obtainable; this can help prioritize algorithm improvements as we move to the global scale RSQ.
- By presenting examples on richer priors did not always lead to improved RSQ, we discuss the optimal ways of configuring the RS observations and the prior model, which includes inflating the distribution in the monthly priors, labeling global rivers with known problems in the prior model, and accounting for the river's geomorphology. These suggestions should also hold for the SWOT setting for global discharge inversion.

Although our work has been largely motivated by the SWOT's launch, it is noted again that the accuracy revealed should not be directly linked with the SWOT discharge product. The additional information on river stage/slope and better data retrievals from SWOT will lead to better RSQ particularly at locations where relying on *W* alone is inadvisable. Additionally, as low-*b* channels take up almost two thirds of our assessment likely because gauges are preferentially located in places with stable hydraulics, one should expect even better RSQ for naturalized and ungauged rivers for SWOT. Overall, our work adds value to the existing RSQ studies by prioritizing "spatial generalizability" over "perfect skill at limited locations". This emphasis not only helps to offer insights into factors compounding the RSQ inversion behavior at scale, but also helps to outline the optimal ways to combine RS data and priors globally. In the future, we suggest more systematic geospatial data collections on river hydraulic variables from altimetry missions (e.g., Coss et al., 2020; Hou et al., 2020; Huang et al., 2020a; Nielsen et al., 2022) and optical sensors (e.g., Feng et al., 2022), together with improved prior knowledge of river hydroclimatology, geomorphology, and hydraulics (Bjerklie et al., 2020; Brinkerhoff et al., 2020; Lin et al., 2019; Linke et al., 2019). Such integrated knowledge and explicit labeling of rivers holds potential to vastly improve our monitoring capability for the discharge variability worldwide in the SWOT era.

Credit authors statement

PL conceived the study, conducted the main analyses, produced the figures, and wrote the initial draft; DF assisted with RivWidthCloud data collection and processing; CJG and MP assisted with the BAM/geoBAM experimental setup; CJG, MP and CB assisted with part of the analyses; XY, HEB, RF provided discussions and technical support. All authors contributed to the editing and revisions of the manuscript.

Declaration of Competing Interest

The authors declare no conflict of interest.

Data availability

The multi-temporal reach-averaged river width data extracted from Landsat (1984-2019) for the 3,078 locations and the BAM/geoBAM discharge estimated from these width RS measurements are publicly shared at Zenodo: https://zenodo.org/record/6655532#.Yqw37XZBz-g. Note that the multi-temporal width data were obtained prior to and independent from Feng et al. (2022) with a slightly different river centerline/orthogonal line definition and cross-section sampling strategy, so it is reasonable to expect different width data accuracies from Feng et al. (2022) at locations. The BAM/geoBAM code packages for discharge inversion are openly available at https://github.com/markwh/bamr and https://github.com/craigbrinkerhoff/geoBAMr. The data processing codes are available upon request to the corresponding author.

Acknowledgement

This study is supported by the Open Research Program of the International Research Center of Big Data for Sustainable Development Goals, Grant No. CBAS2022ORP05, the National Key Research and Development Program (2022YFF0801303), and the startup fund from Peking University on "Numerical modeling and remote sensing of global river discharge" (#7100603564). We thank Prof. Eric F. Wood (deceased in November 2021) who offered valuable input to this work at its early stage. Renato Frasson is supported by NASA SWOT Science Team funding (80NSSC20K1339). Colin Gleason is supported by NASA SWOT Science Team funding (80NSSC20K1141) and NSF CAREER funding (VGJHK59NMPK9). Dongmei Feng is supported by NASA Terrestrial Hydrology Program funding (80NSSC22K0987). Craig Brinkerhoff is supported by a NASA FINESST Fellowship (80NSSC21K1591). We would like to thank the Global Runoff Data Centre (GRDC) for providing river discharge data. A portion of this work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). Three anonymous reviewers and the editors are thanked for offering constructive comments to this manuscript. Author contribution: PL conceived the study, conducted the main analyses, produced the figures, and wrote the initial draft; DF assisted with RivWidthCloud data collection and processing; CJG assisted with the BAM/geoBAM experimental setup; CJG, MP and CB assisted with part of the analyses; XY, HEB, RF provided discussions and technical support; all authors contributed to the editing and revisions of the manuscript. The authors declare no conflict of interest.

Appendix A. Appendix

A.1. Physics of BAM and geoBAM

Both BAM and geoBAM are based off the AMHG theory. The equation adopted by geoBAM is different from BAM in that it has more explicit geomorphic physics based on Dingman's (2007) channel formulation. The original BAM writes the log-transformed AMHG as Eq. (1):

$$logW_{i,t} = b_i(logQ_t - logQ_c) + logW_c + \epsilon_g \quad (1)$$

where $W_{i,t}$ stands for width observation for cross section i at time t; Q_c, W_c are global parameters for AMHG; b_i is the width-discharge exponent; and ϵ_g

is the error term. Later, geoBAM leverages theoretical work (Brinkerhoff et al., 2019) reconciling AMHG with traditional hydraulic geometry via the model of Dingman (2007). This re-writes AMHG as Eq. (2):

$$Q_t = \left(\frac{W_{it}}{W_C}\right)^{\frac{1}{b_i}} \left(W_{b_i}^{\frac{5r_i}{3}} D_{b_i}^{\frac{-5}{3}} \left(\frac{r_i}{r_i+1}\right)^{\frac{-5}{3}} S_{it}^{-\frac{1}{2}} n_i W_c^{-\left(1+\frac{5r_i}{3}\right)} \right)^{-1} \quad (2)$$

where several variables were added into AMHG, including W_{b_i}, D_{b_i}that denote bankfull width/depth, r_i that represents a channel shape parameter by Dingman (2007), n_i as the Manning's roughness, and S_{it} for channel slope. This updated Eq. 2 makes the hydraulic form of AMHG more physics-based and resemble the Manning's (Brinkerhoff et al., 2020).

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2023.113489.

References

Allen, G.H., Pavelsky, T.M., 2018. Global extent of rivers and streams. Science 361, 585–588. https://doi.org/10.1126/science.aat0636.

Alsdorf, D.E., Lettenmaier, D.P., 2003. Tracking fresh water from space. Science 301, 1491–1494. https://doi.org/10.1126/science.1089802.

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

Altenau, E.H., Pavelsky, T.M., Durand, M.T., Yang, X., Frasson, R.P.M., Bendezu, L., 2021. The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A Global River Network for Satellite Data Products. Water Resour. Res. 57, e2021WR030054 https://doi.org/10.1029/2021WR030054.

Andreadis, K.M., Brinkerhoff, C.B., Gleason, C.J., 2020. Constraining the assimilation of SWOT observations with hydraulic geometry relations. Water Resour. Res. 56, e2019WR026611 https://doi.org/10.1029/2019WR026611.

Beck, H.E., de Roo, A., van Dijk, A.I.J.M., 2015. Global maps of streamflow characteristics based on observations from several thousand catchments. J. Hydrometeorol. 16, 1478–1501. https://doi.org/10.1175/JHM-D-14-0155.1.

Beck, H., Zimmermann, N., McVicar, T., et al., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 180214. https://doi.org/10.1038/sdata.2018.214.

Biancamaria, S., Lettenmaier, D.P., Pavelsky, T.M., 2016. The SWOT Mission and its capabilities for land hydrology. Surv. Geophys. 37, 307–337. https://doi.org/10.1007/s10712-015-9346-y.

Bjerklie, D.M., Birkett, C.M., Jones, J.W., Carabajal, C., Rover, J.A., Fulton, J.W., Garambois, P.-A., 2018. Satellite remote sensing estimation of river discharge: application to the Yukon River Alaska. J. Hydrol. 561, 1000–1018. https://doi.org/10.1016/j.jhydrol.2018.04.005.

Bjerklie, D.M., Fulton, J.W., Dingman, S.L., Canova, M.G., Minear, J.T., Moramarco, T., 2020. Fundamental hydraulics of cross sections in natural Rivers: preliminary analysis of a large data set of acoustic doppler flow measurements. Water Resour. Res. 56, e2019WR025986 https://doi.org/10.1029/2019WR025986.

Bjerklie, D.M., Lawrence Dingman, S., Vorosmarty, C.J., Bolster, C.H., Congalton, R.G., 2003. Evaluating the potential for measuring river discharge from space. J. Hydrol. 278, 17–38. https://doi.org/10.1016/S0022-1694(03)00129-X.

Brakenridge, G.R., Nghiem, S.V., Anderson, E., Chien, S., 2005. Space-based measurement of river runoff. EOS Trans. Am. Geophys. Union 86, 185–188. https://doi.org/10.1029/2005EO190001.

Brakenridge, G.R., Nghiem, S.V., Anderson, E., Mic, R., 2007. Orbital microwave measurement of river discharge and ice status. Water Resour. Res. 43 https://doi.org/10.1029/2006WR005238.

Brinkerhoff, C.B., Gleason, C.J., Feng, D., Lin, P., 2020. Constraining Remote River discharge estimation using reach-scale geomorphology. Water Resour. Res. 56, e2020WR027949 https://doi.org/10.1029/2020WR027949.

Brinkerhoff, C.B., Gleason, C.J., Ostendorf, D.W., 2019. Reconciling At-A-Station and at-many-stations hydraulic geometry through river-wide geomorphology. Geophys. Res. Lett. https://doi.org/10.1029/2019GL084529.

Canova, M., Fulton, J., Bjerklie, D., 2016. USGS HYDRoacoustic dataset in support of the Surface Water Oceanographic Topography satellite mission (HYDRoSWOT).

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

Dingman, S.L., 2007. Analytical derivation of at-a-station hydraulic–geometry relations. J. Hydrol. 334, 17–27. https://doi.org/10.1016/j.jhydrol.2006.09.021.

Do, H.X., Gudmundsson, L., Leonard, M., Westra, S., 2018. The global streamflow indices and metadata archive (GSIM) – part 1: the production of a daily streamflow archive and metadata. Earth Syst. Sci. Data 10, 765–785. https://doi.org/10.5194/essd-10-765-2018.

Durand, M., Gleason, C.J., Garambois, P.A., Bjerklie, D., Smith, L.C., Roux, H., Rodriguez, E., Bates, P.D., Pavelsky, T.M., Monnier, J., Chen, X., Baldassarre, G.D., Fiset, J.-M., Flipo, N., Frasson, R.P.M., Fulton, J., Goutal, N., Hossain, F., Humphries, E., Minear, J.T., Mukolwe, M.M., Neal, J.C., Ricci, S., Sanders, B.F., Schumann, G., Schubert, J.E., Vilmin, L., 2016. An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. Water Resour. Res. 52, 4527–4549. https://doi.org/10.1002/2015WR018434.

Durand, M., Gleason, C.J., Pavelsky, T.M., Frasson, R.P.M., Turmon, M.J., David, C.H., Altenau, E.H., Tebaldi, N., Larnier, K., Monnier, J., Malaterre, P.-O., Oubanas, H., Allen, G.H., Bates, P.D., Bjerklie, D.M., Coss, S.P., Dudley, R.W., Marc, L.F., Garambois, P.-A., Lin, P., Margulis, S.A., Matte, P., Minear, J.T., Muhebwa, A., Pan, M., Peters, D., Riggs, R.M., Tarpanelli, A., Schulze, K., Tourian, M.J., Wang, J., 2021. A framework for estimating global river discharge from the Surface Water and Ocean Topography satellite mission [WWW Document]. Earth Space Sci. Open Arch. https://doi.org/10.1002/essoar.10508946.1.

Durand, M., Neal, J., Rodríguez, E., Andreadis, K.M., Smith, L.C., Yoon, Y., 2014. Estimating reach-averaged discharge for the river Severn from measurements of river water surface elevation and slope. J. Hydrol. 511, 92–104. https://doi.org/10.1016/j.jhydrol.2013.12.050.

Feng, D., Gleason, C.J., Lin, P., Yang, X., Pan, M., Ishitsuka, Y., 2021. Recent changes to Arctic river discharge. Nat. Commun. 12, 6917. https://doi.org/10.1038/s41467-021-27228-1.

Feng, D., Gleason, C.J., Yang, X., Pavelsky, T.M., 2019. Comparing discharge estimates made via the BAM algorithm in high-order Arctic Rivers derived solely from optical CubeSat, landsat, and Sentinel-2 data. Water Resour. Res. 55, 7753–7771. https://doi.org/10.1029/2019WR025599.

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

Frasson, R.P.M., Durand, M.T., Larnier, K., Gleason, C., Andreadis, K.M., Hagemann, M., Dudley, R., Bjerklie, D., Oubanas, H., Garambois, P.-A., Malaterre, P.-O., Lin, P., Pavelsky, T.M., Monnier, J., Brinkerhoff, C.B., David, C.H., 2021. Exploring the factors controlling the error characteristics of the surface water and ocean topography mission discharge estimates. Water Resour. Res. 57, e2020WR028519 https://doi.org/10.1029/2020WR028519.

Frasson, R.P.M., Turmon, M.J., Durand, M.T., David, C.H., 2022. Estimating the relative impact of measurement, parameter, and flow law errors on discharge from the Surface Water and Ocean Topography mission. J. Hydrometeorol. https://doi.org/10.1175/jhm-d-22-0078.1.

Garambois, P.-A., Monnier, J., 2015. Inference of effective river properties from remotely sensed observations of water surface. Adv. Water Resour. 79, 103–120. https://doi.org/10.1016/j.advwatres.2015.02.007.

Gejadze, I., Malaterre, P.-O., Oubanas, H., Shutyaev, V., 2022. A new robust discharge estimation method applied in the context of SWOT satellite data processing. J. Hydrol. 127909 https://doi.org/10.1016/j.jhydrol.2022.127909.

Gleason, C., Garambois, P.-A., Durand, M., 2017. Tracking river flows from space. EOS Earth Space Sci. News. https://doi.org/10.1029/2017EO078085.

Gleason, C.J., Durand, M.T., 2020. Remote sensing of river discharge: a review and a framing for the discipline. Remote Sens. 12, 1107. https://doi.org/10.3390/rs12071107.

Gleason, C.J., Smith, L.C., 2014. Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. Proc. Natl. Acad. Sci. 111, 4788–4791. https://doi.org/10.1073/pnas.1317606111.

Gleason, C.J., Smith, L.C., Lee, J., 2014. Retrieval of river discharge solely from satellite imagery and at-many-stations hydraulic geometry: sensitivity to river form and optimization parameters. Water Resour. Res. 50, 9604–9619. https://doi.org/10.1002/2014WR016109.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. Big Remote. Sens. Data 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Hagemann, M.W., Gleason, C.J., Durand, M.T., 2017. BAM: bayesian AMHG-manning inference of discharge using remotely sensed stream width, slope, and height: BAM FLOW USING STREAM WIDTH SLOPE HEIGHT. Water Resour. Res. 53, 9692–9707. https://doi.org/10.1002/2017WR021626.

Hannah, D.M., Demuth, S., van Lanen, H.A.J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., Tallaksen, L.M., 2011. Large-scale river flow archives: importance, current status and future needs. Hydrol. Process. 25, 1191–1200. https://doi.org/10.1002/hyp.7794.

Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., Cudennec, C., 2013. A decade of predictions in ungauged basins (PUB)—a review. Hydrol. Sci. J. 58 (6), 1198–1255.

Hou, J., van Dijk, A.I.J.M., Beck, H.E., 2020. Global satellite-based river gauging and the influence of river morphology on its application. Remote Sens. Environ. 239, 111629 https://doi.org/10.1016/j.rse.2019.111629.

Huang, Q., Long, D., Du, M., Han, Z., Han, P., 2020a. Daily Continuous River discharge estimation for ungauged basins using a hydrologic model calibrated by satellite altimetry: implications for the SWOT Mission. Water Resour. Res. 56 https://doi.org/10.1029/2020WR027309.

Huang, Q., Long, D., Du, M., Han, Z., Han, P., 2020b. Daily Continuous River discharge estimation for ungauged basins using a hydrologic model calibrated by satellite altimetry: implications for the SWOT Mission. Water Resour. Res. 56, e2020WR027309 https://doi.org/10.1029/2020WR027309.

Huang, Q., Long, D., Du, M., Zeng, C., Li, X., Hou, A., Hong, Y., 2018. An improved approach to monitoring Brahmaputra River water levels using retracked altimetry data. Remote Sens. Environ. 211, 112–128. https://doi.org/10.1016/j.rse.2018.04.018.

Huang, Q., Long, D., Han, Z., Han, P., 2022. High-resolution satellite images combined with hydrological modeling derive river discharge for headwaters: a step toward discharge estimation in ungauged basins. Remote Sens. Environ. 277, 113030 https://doi.org/10.1016/j.rse.2022.113030.

Ishitsuka, Y., Gleason, C.J., Hagemann, M.W., Beighley, E., Allen, G.H., Feng, D., Lin, P., Pan, M., Andreadis, K., Pavelsky, T.M., 2021. Combining optical remote sensing, McFLI discharge estimation, global hydrologic modelling, and data assimilation to improve daily discharge estimates across an entire large watershed. Water Resour. Res. 57, e2020WR027794 https://doi.org/10.1029/2020WR027794.

Jones, J.W., 2019. Improved automated detection of subpixel-scale Inundation—Revised dynamic surface water extent (DSWE) partial surface water tests. Remote Sens. 11, 374. https://doi.org/10.3390/rs11040374.

Knoben, W.J., Freer, J.E., Woods, R.A., 2019. Inherent Benchmark or Not? Comparing Nash-Sutcliffe, and Kling-Gupta Efficiency Scores. Hydrol. Earth Syst. Sci 23, 4323–4331. https://doi.org/10.5194/hess-23-4323-2019.

Krabbenhoft, C.A., Allen, G.H., Lin, P., Godsey, S.E., Allen, D.C., Burrows, R.M., DelVecchia, A.G., Fritz, K.M., Shanafield, M., Burgin, A.J., Zimmer, M.A., Datry, T., Dodds, W.K., Jones, C.N., Mims, M.C., Franklin, C., Hammond, J.C., Zipper, S., Ward, A.S., Costigan, K.H., Beck, H.E., Olden, J.D., 2022. Assessing placement bias of the global river gauge network. Nat. Sustain. 1–7 https://doi.org/10.1038/s41893-022-00873-0.

Larnier, K., Monnier, J., Garambois, P.-A., Verley, J., 2020. River discharge and bathymetry estimations from SWOT altimetry measurements. Inverse Problems in Science and Engineering 29 (6), 759–789.

Lin, P., Pan, M., Allen, G.H., de Moraes Frasson, Zeng, Z., Yamazaki, D., Wood, E.F., 2020. Global estimates of reach-level bankfull river width leveraging big data geospatial analysis. Geophys. Res. Lett. 47, e2019GL086405. https://doi.org/10.1029/2019GL086405.

Lin, P., Pan, M., Beck, H.E., Yang, Y., Yamazaki, D., Frasson, R., David, C.H., Durand, M., Pavelsky, T.M., Allen, G.H., Gleason, C.J., Wood, E.F., 2019. Global reconstruction of naturalized river flows at 2.94 million reaches. Water Resour. Res. 0 https://doi.org/10.1029/2019WR025287.

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., Thieme, M., 2019. Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. Sci. Data 6, 283. https://doi.org/10.1038/s41597-019-0300-6.

Mao, J., Yan, B., 2019. Global Monthly Mean Leaf Area Index Climatology, 1981-2015. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1653.

Marcus, W.A., Fonstad, M.A., 2008. Optical remote mapping of rivers at sub-meter resolutions and watershed extents. Earth Surf. Process. Landf. 33, 4–24. https://doi.org/10.1002/esp.1637.

Nickles, C., Beighley, E., Durand, M., de Moraes, Prata, Frasson, R., 2020. Integrating lateral inflows into a SWOT Mission River discharge algorithm. Water Resour. Res. 56, e2019WR026589 https://doi.org/10.1029/2019WR026589.

Nielsen, K., Zakharova, E., Tarpanelli, A., Andersen, O.B., Benveniste, J., 2022. River levels from multi mission altimetry, a statistical approach. Remote Sens. Environ. 270, 112876 https://doi.org/10.1016/j.rse.2021.112876.

Oubanas, H., Gejadze, I., Malaterre, P.-O., Durand, M., Wei, R., Frasson, R.P.M., Domeneghetti, A., 2018. Discharge estimation in ungauged basins through variational data assimilation: the potential of the SWOT Mission. Water Resour. Res. 54, 2405–2423. https://doi.org/10.1002/2017WR021735.

Pavelsky, T.M., 2014. 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.

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

Revilla-Romero, B., Thielen, J., Salamon, P., De Groeve, T., Brakenridge, G.R., 2014. Evaluation of the satellite-based global flood detection system for measuring river discharge: influence of local factors. Hydrol. Earth Syst. Sci. 18, 4467–4484. https://doi.org/10.5194/hess-18-4467-2014.

Riggs, R.M., Allen, G.H., David, C.H., Lin, P., Pan, M., Yang, X., Gleason, C., 2022. RODEO: an algorithm and Google earth engine application for river discharge retrieval from landsat. Environ. Model. Softw. 148, 105254 https://doi.org/10.1016/j.envsoft.2021.105254.

Sichangi, A.W., Wang, L., Yang, K., Chen, D., Wang, Z., Li, X., Zhou, J., Liu, W., Kuria, D., 2016. Estimating continental river basin discharges using multiple remote sensing data sets. Remote Sens. Environ. 179, 36–53. https://doi.org/10.1016/j.rse.2016.03.019.

Sivapalan, M., et al., 2003. IAHS decade on predictions in ungauged basins, PUB. 2003–2012: shaping an exciting future for the hydrological sciences. Hydrol. Sci. J. 48 (6), 857–880.

Smith, L.C., 1997. Satellite remote sensing of river inundation area, stage, and discharge: a review. Hydrol. Process. 11, 1427–1439. https://doi.org/10.1002/(SICI)1099-1085(199708)11:10<1427::AID-HYP473>3.0.CO;2-S.

Smith, L.C., Pavelsky, T.M., 2008. Estimation of river discharge, propagation speed, and hydraulic geometry from space: Lena River, Siberia. Water Resour. Res. 44 https://doi.org/10.1029/2007WR006133.

Tarpanelli, A., Brocca, L., Lacava, T., Melone, F., Moramarco, T., Faruolo, M., Pergola, N., Tramutoli, V., 2013. Toward the estimation of river discharge variations using MODIS data in ungauged basins. Remote Sens. Environ. 136, 47–55. https://doi.org/10.1016/j.rse.2013.04.010.

Tarpanelli, A., Camici, S., Nielsen, K., Brocca, L., Moramarco, T., Benveniste, J., 2021. Potentials and limitations of Sentinel-3 for river discharge assessment. Adv. Space Res. 68, 593–606. https://doi.org/10.1016/j.asr.2019.08.005, 25 Years of Progress in Radar Altimetry.

Tourian, M.J., Sneeuw, N., Bárdossy, A., 2013. A quantile function approach to discharge estimation from satellite altimetry (ENVISAT). Water Resour. Res. 49, 4174–4186. https://doi.org/10.1002/wrcr.20348.

Tuozzolo, S., Lind, G., Overstreet, B., Mangano, J., Fonstad, M., Hagemann, M., Frasson, R.P.M., Larnier, K., Garambois, P.-A., Monnier, J., Durand, M., 2019. Estimating river discharge with swath altimetry: a proof of concept using AirSWOT observations. Geophys. Res. Lett. 46, 1459–1466. https://doi.org/10.1029/2018GL080771.

Van Dijk, A.I.J.M., Brakenridge, G.R., Kettner, A.J., Beck, H.E., De Groeve, T., Schellekens, J., 2016. River gauging at global scale using optical and passive microwave remote sensing. Water Resour. Res. 52, 6404–6418. https://doi.org/10.1002/2015WR018545.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G.H., Pavelsky, T.M., 2019. MERIT hydro: a high-resolution global hydrography map based on latest topography dataset. Water Resour. Res. https://doi.org/10.1029/2019WR024873.

Yang, J., Huang, X., Tang, Q., 2020. Satellite-derived river width and its spatiotemporal patterns in China during 1990–2015. Remote Sens. Environ. 247, 111918 https://doi.org/10.1016/j.rse.2020.111918.

Yang, X., Pavelsky, T.M., Allen, G.H., Donchyts, G., 2020. RivWidthCloud: an automated Google earth engine algorithm for river width extraction from remotely sensed imagery. IEEE Geosci. Remote Sens. Lett. 17, 217–221. https://doi.org/10.1109/LGRS.2019.2920225.

Yang, Y., Pan, M., Lin, P., Beck, H.E., Zeng, Z., Yamazaki, D., David, C.H., Lu, H., Yang, K., Hong, Y., Wood, E.F., 2021. Global reach-level 3-Hourly River flood reanalysis (1980–2019). Bull. Am. Meteorol. Soc. 102, E2086–E2105. https://doi.org/10.1175/BAMS-D-20-0057.1.

Zomer, R.J., Trabucco, A., Bossio, D.A., van Straaten, O., Verchot, L.V., 2008. Climate change mitigation: a spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. Agric. Ecosyst. Environ. 126, 67–80.