AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2014WR016109

Key Points:
- River discharge is retrieved solely from satellite imagery and AMHG in 34 rivers
- Successful retrieval is strongly linked to channel morphology and form
- Best practices for retrieval reach selection and method parameterization are identified

Correspondence to:
C. J. Gleason,
cjgleaso@ucla.edu

Citation:
Gleason, C. J., L. C. Smith, and J. Lee (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*, doi:10.1002/2014WR016109.

Received 7 JUL 2014
Accepted 11 NOV 2014
Accepted article online 17 NOV 2014

Retrieval of river discharge solely from satellite imagery and at-many-stations hydraulic geometry: Sensitivity to river form and optimization parameters

Colin J. Gleason[1], Laurence C. Smith[1], and Jinny Lee[1]

[1]Department of Geography, University of California-Los Angeles, Los Angeles, California, USA

Abstract Knowledge of river discharge is critically important for water resource management, climate modeling, and improved understanding of the global water cycle, yet discharge is poorly known in much of the world. Remote sensing holds promise to mitigate this gap, yet current approaches for quantitative retrievals of river discharge require in situ calibration or a priori knowledge of river hydraulics, limiting their utility in unmonitored regions. Recently, Gleason and Smith (2014) demonstrated discharge retrievals within 20–30% of in situ observations solely from Landsat TM satellite images through discovery of a river-specific geomorphic scaling phenomenon termed at-many-stations hydraulic geometry (AMHG). This paper advances the AMHG discharge retrieval approach via additional parameter optimizations and validation on 34 gauged rivers spanning a diverse range of geomorphic and climatic settings. Sensitivity experiments reveal that discharge retrieval accuracy varies with river morphology, reach averaging procedure, and optimization parameters. Quality of remotely sensed river flow widths is also important. Recommended best practices include a proposed global parameter set for use when a priori information is unavailable. Using this global parameterization, AMHG discharge retrievals are successful for most investigated river morphologies (median RRMSE 33% of in situ gauge observations), except braided rivers (median RRMSE 74%), rivers having low at-a-station hydraulic geometry b exponents (reach-averaged $b < 0.1$, median RRMSE 86%), and arid rivers having extreme discharge variability (median RRMSE $> 1000\%$). Excluding such environments, 26–41% RRMSE agreement between AMHG discharge retrievals and in situ gauge observations suggests AMHG can meaningfully address global discharge knowledge gaps solely from repeat satellite imagery.

1. Introduction

Knowledge of the distribution, volume, and temporal availability of river discharge has profound implications for human health and well-being, as the majority of the world's potable and ecosystem water supplies come from rivers [*Barnett et al.*, 2005; *Vörösmarty et al.*, 2010]. In this capacity, quantitative knowledge of river discharge is of critical importance to water management, climate modeling, and protection of riparian and wetland ecosystems. However, in situ gauging stations, the primary way discharge is measured, fail to fully describe its spatial variability and are unavailable for many regions of the globe [*Perry et al.*, 1996; *Dai and Trenberth*, 2002; *Sivapalan et al.*, 2003; *Dai et al.*, 2009]. These knowledge gaps are expected to widen, as the global availability of gauging stations is in decline [*Milliman et al.*, 2008; *Conway et al.*, 2009; *Hannah et al.*, 2011]. In addition, political realities in many regions of the world dictate that some existing gauge data are not shared publicly [e.g., *Conway and Hulme*, 1993; *Dinar*, 2012]. Given these difficulties, supplementary alternatives for quantifying river discharges are needed [*Alsdorf et al.*, 2007].

One such alternative is through use of hydrological water balance models, which assimilate climatological observations and land surface parameterizations to reconstruct the cycling of water between the oceans, atmosphere, and land surface. Water balance models have been used to assess global and regional drought, sustainability, and water use [e.g., *Oki et al.*, 1995, 1999; *Hanasaki et al.*, 2008; *Van Beek et al.*, 2011; *Wada et al.*, 2010]. However, these models typically represent rivers as coarse spatial grids best suited only for very large rivers, and generally estimate total, time-aggregated discharge (i.e., runoff) more accurately than instantaneous discharge. Perhaps most importantly, these models rely heavily on in situ gauge observations for calibration, and are therefore fundamentally limited by the availability of gauges [*Peel and McMahon*, 2006; *Widen-Nilsson et al.*, 2007; *Milly et al.*, 2008; *Hannah et al.*, 2011].

Remotely sensed observations offer a promising opportunity to observe river discharge globally, as images and/or water surface elevations collected from satellites can be obtained in regions where access is limited and ground based discharge measurements are unavailable or proprietary [*Smith*, 1997; *Lettenmaier and Famiglietti*, 2006; *Alsdorf et al*., 2007]. However, river discharge cannot be directly measured from any known satellite or airborne sensor. As such, researchers have sought to combine ground-based information with remotely sensed data to retrieve discharge. These methods include: comparing remotely sensed observations of river inundation area or width with in situ discharge measurements to create rating curves linking the two [e.g., *Smith et al*., 1996; *Gilvear et al*., 2004; *Ashmore and Sauks*, 2006; *Brakenridge et al*., 2007; *Smith and Pavelsky*, 2008; *Nathanson et al*., 2012; *Tarpanelli et al*., 2013; *Pavelsky*, 2014], incorporating remotely sensed observations into hydraulic models within a data assimilation framework [e.g., *Andreadis et al*., 2007; *Durand et al*., 2008, 2010a, 2010b; *Neal et al*., 2009], and using remotely sensed data in conjunction with typical values of river hydraulic parameters like Manning's n and width-depth ratios [e.g., *Bjerklie et al*., 2003, 2005; *Bjerklie*, 2007]. At very broad spatial scales, temporal changes in total water storage alter the Earth's gravity field and can thus be tracked with the GRACE satellites, which allows broad scale runoff to be calculated as the difference between inputs (precipitation), outputs (evaporation), and storage change (GRACE) [e.g., *Syed et al*., 2007, 2009, 2010; *Lee et al*., 2011].

In all of the above remote sensing approaches, some form of in situ hydraulic or climatological data and/or a priori assumptions of river channel properties are required in order to successfully retrieve river discharge. However, *Gleason and Smith* [2014] recently demonstrated a novel method for retrieving discharge solely from repeated imaging of river flow width over space and time, with no required in situ measurements or a priori information whatsoever, using multitemporal Landsat Thematic Mapper (TM) imagery. This method was enabled by the discovery of at-many-stations hydraulic geometry (AMHG), which holds that certain parameters of traditional at-a-station hydraulic geometry (AHG) ($w = aQ^b$, $d = cQ^f$, $v = kQ^m$, where w = width, d = depth, v = velocity, Q = discharge, and a, b, c, f, k, m are empirical parameters, *Leopold and Maddock* [1953]) are not independent but log-linearly related along a river. The work revealed the existence of AMHG by plotting corresponding a–b, c–f, and k–m pairs from 6 to 26 USGS gauging stations positioned along 150 to 3500 km lengths of six U.S. rivers. Given a remotely sensed approximation to AMHG, *Gleason and Smith* [2014] then demonstrated discharge retrievals within 20–30% of in situ gauge estimates for the Mississippi, Athabasca, and Yangtze Rivers derived solely from AMHG-constrained optimization of cross-sectional river width measured from Landsat TM imagery (see section 2.2.3, for further description of the methodology). However, because this discharge retrieval demonstration was based on just three large, morphologically similar rivers, more testing is needed to assess the robustness of AMHG and the associated discharge retrieval method.

To that end, the present paper expands the aforementioned AMHG discharge retrieval method to test its performance and sensitivities, as compared with traditional in situ river gauges, over 34 rivers in different physiographic and climatologic settings (Figure 1). As such, the paper seeks to confirm the presence of AMHG in a larger sample size of rivers, and to confirm if the associated discharge retrieval method functions with similar accuracy in different environments. Of key interest is ascertaining if there are certain rivers or classes of rivers for which AMHG is not observed and/or for which the discharge retrieval method breaks down. This paper also seeks to test if the remotely sensed approximation of AMHG proposed by *Gleason and Smith* [2014] continues to reliably approximate AMHG as calculated from in situ data at the same location. Finally, to evaluate robustness of the discharge retrieval method, 120 different parameterizations for each river are evaluated for sensitivity to key parameter settings, input width accuracy, and reach averaging procedures. The paper concludes with a synthesis of learned strengths and weaknesses of the AMHG discharge retrieval method, and gives some recommended best practices for its application in the future.

2. Data and Methods

2.1. Data

In situ river gauge observations of mean daily discharge were acquired for 34 rivers from the USGS, Environment Canada, the Global Runoff Data Centre, and the Changjiang Wuhan Waterway Bureau for comparison with remotely sensed discharge retrievals (Table 1). These rivers are located in diverse physiographic and climatic settings around the world with at least four gauging stations per continent (Figure 1), and represent six major morphology categories (A, B, C, D, DA, and E) of the Rosgen river classification system [*Rosgen*, 1994].

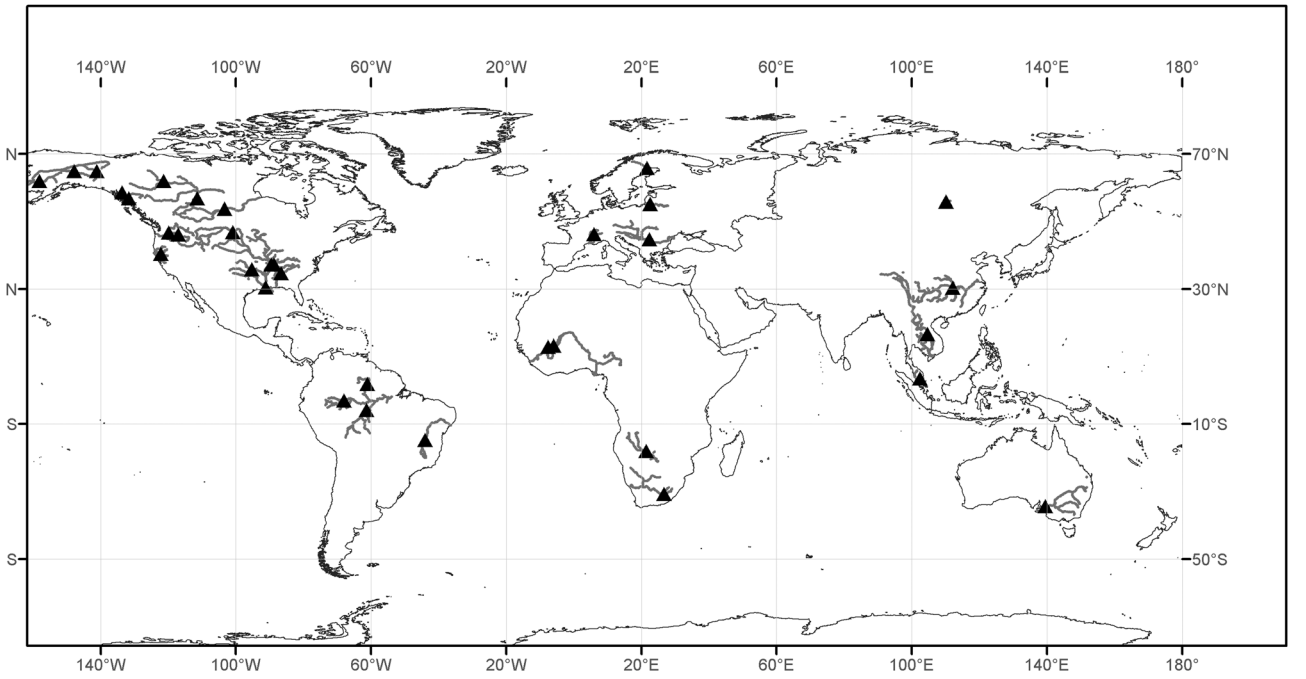

Figure 1. Locations of 34 in situ river gauges used in this study, distributed across a diverse range of geomorphic and climatic settings. A 10 km river reach surrounding or abutting each gauge was tested for AMHG presence, and used to retrieve three different discharge products (Q_{best}, Q_{inv}, Q_{glob}) using 9–16 Landsat TM images. River gauge observations were provided by the USGS, Environment Canada, Changjiang Wuhan Waterway Bureau, and Global Runoff Data Centre. See Table 1 for gauge coordinates and additional information.

At each river gauge, a search for archived, cloud-free Landsat images collected on the same day as reported gauge discharge was conducted using the USGS's Earth Explorer interface (http://earthexplorer.usgs.gov/). The maximum allowable time between the first and last images for a particular river was restricted to 20 years to mitigate impacts of ongoing erosional and depositional changes to channel geometry [*Phillips*, 1990; *Ran et al*., 2014]. Images were acquired by the Landsat MSS (10 images), TM (391 images), and ETM+ (76 images) sensors. TM and ETM+ each have 30 m spatial resolution, while the 10 MSS images collected over the Mekong River have 79 m spatial resolution. Identifying high-quality image-gauge data pairs proved difficult, owing to frequent cloud cover, intermittent gauge data, and winter ice conditions. In total, 477 high-quality image-gauge data pairs were identified, ranging from 9 to 16 cloud-free images per station over the period 1972–2011 (Table 1). In situ gauge discharges were assumed to be representative of flows throughout the day, with no subdaily correction between satellite image acquisition time and mean daily flow.

2.2. Methods

To remotely retrieve river discharge from Landsat imagery, the following workflow was performed: first, study reaches were classified into Rosgen classes and assigned climactic regimes (section 2.2.1). Next, instantaneous cross-sectional flow widths were manually extracted for each satellite image at fixed cross sections (section 2.2.2). Third, these remotely sensed river flow widths were used to approximate AMHG and then passed to an ensemble of genetic algorithms (GA) to retrieve instantaneous river discharge for each satellite image acquisition date (section 2.2.3). Finally, each river was subjected to a series of sensitivity tests to assess method robustness and the impacts of varying input widths and optimization parameter choices on the final discharge result (section 2.2.4). A detailed description of each of these steps follows.

2.2.1. Assignment of Rosgen Classification and Climatic Regime

River Rosgen class was determined for the 10 km reach encompassing each river gauge based on visual inspection of Landsat imagery and of imagery and terrain elevations provided within the Google Earth software package. In accordance with *Rosgen* [1994], these categories were assigned as relatively straight and steep (A stream types), low sinuosity moderately steep (B stream types), meandering (C stream types), braided (D stream types), anastomosed (DA stream types), and tortuously meandering (E stream types). Because this classification system is highly dependent on the scale of analysis, the classes assigned to each

Table 1. Gauge Locations and Descriptions[a]

River	Gauge Location	Latitude/Longitude	Median Flow (m^3/s)	Mean Width (m)	Rosgen Class	Koeppen-Geiger Climate Class	AMHG r^2
Snake	Anatone, WA, USA	46°5′50″N 116°58′36″W	589	148	A	Warm temperate	0.99
Columbia	Grant County, WA, USA	46°37′44″N 119°51′49″W	3172	470	B	Warm temperate	0.98
Mekong	Mukdahan, Thailand	16°32′24″N 104°44′12″E	2445	937	B	Equatorial	0.96
Sacramento	Tehama, CA, USA	40°17′19″N 122°11′08″W	303	89	B	Warm temperate	0.98
Yukon	Eagle, AK, USA	64°47′22″N 141°11′52″W	4757	502	B	Snow	0.99
Rhone	Chancy, Switzerland	46°9′10″N 5°58′14″E	319	85	B	Warm temperate	0.98
Kuskokwim	Crooked Creek, AK, USA	61°52′10″N 158°06′41″W	1846	458	B	Snow	0.93
Athabasca	Fort McMurray, AB, Canada	56°46′49″N 111°24′7″W	591	526	B	Snow	0.97
Nemunas	Smalininkai, Lithuania	55°4′33″N 22°34′48″E	533	223	C	Snow	0.98
Orange	Aliwal Noord, South Africa	30°41′18″S 26°42′59″E	108	115	C	Arid	0.8
Bani	Douna, Mali	13°13′0″N 5°54′0″W	77	238	C	Arid	0.97
Niger	Koulikoro, Mali	12°52′0″N 7°33′0″W	295	718	C	Equatorial	0.97
Pahang	Temerloh, Malaysia	3°27′0″N 102°25′48″E	327	221	C	Equatorial	0.89
Missouri	Bismark, ND, USA	46°48′51″N 100°49′17″W	513	392	C	Warm temperate	0.93
Upper Mississippi	Thebes, IL, USA	37°13′17″N 89°27′46″W	4453	538	C	Warm temperate	0.97
Yangtze	Shahsi, Hubei, China	30°19′12″N 112° 10′12″E	10,196	1100	C	Warm temperate	0.99
Saskatchewan	Tobin Lake, SK, Canada	53°42′10″N 103°17′50″W	526	380	C	Snow	0.92
Iskut	Near Johnson River, BC, Canada	56°44′20″N 131°40′25″W	1095	680	D	Snow	0.87
Tanana	Fairbanks, AK, USA	64°47′34″N 147°50′20″W	1096	1107	D	Snow	0.99
Taku	Juneau, AK, USA	58°32′19″N 133°42′00″W	691	281	D	Snow	0.98
Solimoes	Santo Antonio Do Ica, Brazil	3°7′54″S 67°56′8″W	46,504	2717	DA	Equatorial	0.98
Rio Branco	Caracarai, Brazil	1°49′17″N 61°7′25″W	1963	964	DA	Equatorial	0.97
Okavango	Namibia/Angola border	18°2′0″S 21°25′59″E	180	216	DA	Arid	0.78
Murray	Blanchetown, Australia	34°22′56″S 139°36′56″E	56	329	E	Arid	0.82
Luleaelven	Bodens, Sweden	65°48′27″N 21°40′13″E	593	303	Low-*b* (dam)	Snow	0.88
Danube	Orsova, Romania	44°42′0″N 22°25′12″E	4262	1084	Low-*b* (dam)	Warm remperate	0.53
Tennessee	Whitesburg, AL, USA	34°34′18″N 86°33′29″W	963	352	Low-*b* (dam)	Warm temperate	0.31
Arkansas	Muskogee, OK, USA	35°46′10″N 95°17′49″W	994	460	Low-*b* (dam)	Warm temperate	0.05
Verkhnyaya	Verkhnyaya Zaimka, Russia	55°51′0″N 110°9′0″E	662	257	Low-*b* (lake)	Snow	0.55
Ohio	Metropolis, IL, USA	37°08′51″N 88°44′27″W	4899	1218	Low-*b* (confluence)	Warm temperate	0.58
Lower Mississippi	Baton Rouge, LA, USA	30°26′44″N 91°11′29″W	12,199	1009	Low-*b* (levy)	Warm temperate	0.63
Sao Francisco	Manga, Brazil	14°46′34″S 43°55′56″W	1463	619	Low-*b*	Equatorial	0.33
Rio Madiera	Manicore, Brazil	5°51′0″S 61°18′7″W	22,351	1177	Low-*b*	Equatorial	0.93
McKenzie	Fort Simspon, NT, Canada	61°52′6″N 121°21′32″W	12,200	1497	Low-*b*	Snow	0.35

[a]Locations and summary information for all study reaches. Gauge data in the USA provided by the USGS, gauge data in Canada provided by Environment Canada, gauge data in China provided by the Changjiang Wuhan Waterway Bureau, and all other gauge data provided by the Global Runoff Data Center, Koblenz, Germany. "low-*b*" is not a Rosgen class but rather denotes a river reach where mean AHG *b* exponent is less than 0.1 which was ultimately a more important morphological characteristic than Rosgen class in these rivers. Koeppen-Geiger classifications provided by *Kottek et al*. [2006] based on climate data from 1951 to 2000, and accessed at http://koeppen-geiger.vu-wien.ac.at/present.htm.

river reach examined in this paper are valid only at the study reaches and are not necessarily representative of the rest of the river. River classifications were based primarily on river plan form, valley shape, and floodplain area observable in the Landsat and Google Earth imagery. Limited observation of changes in river stage and bankfull/overbank flows in the available imagery precluded important indicators of channel incision, bankfull width, and flood-prone width usually used to aid Rosgen classification. As such, Classes B and C (seven rivers each) were differentiated mostly by their sinuosity and apparent floodplain area. Classes A and E each only had one river and were classified by their sinuosity. Classification of Class D (braided) and Class DA (anastomosing) rivers was straightforward, with the presence of stable banks and stable, vegetated bars the key discriminant between the two.

Climactic regime for each river was assigned a Koeppen-Geiger climactic class according to data products provided by *Kottek et al*. [2006] (accessed at http://koeppen-geiger.vu-wien.ac.at/present.htm). The "Main Classes" data product, corresponding to Polar, Snow, Warm Temperate, Equatorial, and Arid regions, was used to determine the climate zone for each study reach's latitude/longitude river in a GIS analysis (Table 1). As with Rosgen Class, the Koeppen-Geiger class for each river is specific to the study reach and does not represent the entirety of the river.

2.2.2. Extraction of Cross-Sectional Flow Widths in Satellite Imagery

For each Landsat image, instantaneous river flow width (top width) was manually measured at 15 fixed cross sections as per *Gleason and Smith* [2014] within a 10 km study reach either surrounding or immediately abutting the in situ river gauges listed in Table 1. Cross sections were equally spaced within each reach for convenience, as previous work has suggested that the particular locations of cross sections within a reach

do not correlate strongly with discharge retrieval accuracy provided widths are not artificially constrained [*Gleason and Smith* 2014]. A 10 km reach length was determined to be long enough to capture spatial and temporal flow width variations needed for the AMHG discharge retrieval method, yet short enough as to ensure its fundamental assumption of mass conservation along the reach. Reaches were chosen to avoid tributaries and bridges to maintain mass conservation and mitigate influence of control structures that affect river widths.

2.2.3. AMHG Discharge Retrieval Procedure

Several procedures were required to transform cross-sectional width measurements into reach averaged discharge (Figure 2). First, each river's AMHG (equation (1)) was approximated by a previously developed remotely sensed empirical proxy (equation (2)). This approximation is key as AMHG physically defines the relationship between AHG a and b at any station and thus reduces the complexity of AHG and allows for much more efficient discharge retrieval. Given this approximation, the remotely sensed cross-sectional widths extracted above were then passed to an ensemble of GAs that sought to estimate AHG a and b parameters at each cross section by minimizing the difference in discharge between cross sections (cross-sectional discharge calculated from equation (3)) in a pairwise permutation. Output from an ensemble of GAs (the same GA was initialized many times) was then aggregated via reach averaging to arrive at a final discharge for each study reach. These procedures are described in detail below.

Before discharge can be retrieved, a river's AMHG must first be approximated following *Gleason and Smith* [2014]. AMHG holds that the AHG a and b parameters are related to one another along a river (equation (1), adapted from *Gleason and Smith* [2014], equation (S5)). In equation (1), E is a river-specific empirical constant defining the slope of the log-linear relationship between a and b with subscript x referring to any spatially indexed cross-section location along the river, and w_{glob} is the mean of all observed widths in a study reach over space and time:

$$b_x = -\frac{1}{\log(E)} * \log(a_x) + \frac{1}{\log(E)} * \log(w_{glob}) \tag{1}$$

Crucially, *Gleason and Smith* [2014] found a robust empirical proxy, y, for $1/\log(E)$ that can be obtained solely from satellite images (equation (2)). In equation (2), $x_1, x_2, \ldots, x_n$ correspond to the spatial locations of each cross section in a study reach and p and y are empirical regression parameters, where y is a proxy for $1/\log(E)$ in equation (1):

$$max(w_{x1,x2,\ldots,xn}) = p\left(max(w_{x1,x2,\ldots,xn})^2 - min(w_{x1,x2,\ldots,xn})^2\right)^y \tag{2}$$

Given that y approximates $1/\log(E)$, y is also used to calculate the intercept of equation (1) ($y * \log(w_{glob})$) though simple substitution. This remotely sensed AMHG approximation allows for a reduction in the complexity of equation (3) (the classical width AHG equation [$w = aQ^b$, *Leopold and Maddock*, 1953]), as b and $\log(a)$ are explicitly linked under AMHG; thus reducing the number of AHG variables from four to three:

$$\log(Q) = \frac{\log(w) - \log(a)}{b} \tag{3}$$

Discharge may be calculated at any cross section given an a–b pair and observed widths via equation (3). Final, reach averaged discharge in this study was retrieved via unconstrained optimization of these cross-sectional discharges using a GA as in *Gleason and Smith* [2014]. GAs are heuristic algorithms that mimic the processes of natural selection and mutation to optimize parameters for indeterminate optimization systems [*Grefenstette*, 1986; *Deb et al*., 2002]. The GA in this study, which was originally developed to optimize the area of tree crowns in a lidar forestry assessment [*Gleason and Im*, 2012], was modified to minimize the difference in discharge between mass-conserved pairs of cross sections within a reach. This is achieved by testing different parameter combinations (in this case, AHG a and b values) against an objective function, which here is the residual of discharge between any two cross sections. Other optimization methods should be able to estimate discharge in a similar manner, but this specific GA was applied as it is well understood and has been applied to AMHG discharge estimation previously [*Gleason and Smith*, 2014].

The remotely sensed AMHG approximation (equations (1) and (2)) is critical to successful discharge estimation as AMHG provides a physical mechanism whereby the GA can be constrained. Without AMHG, the GA

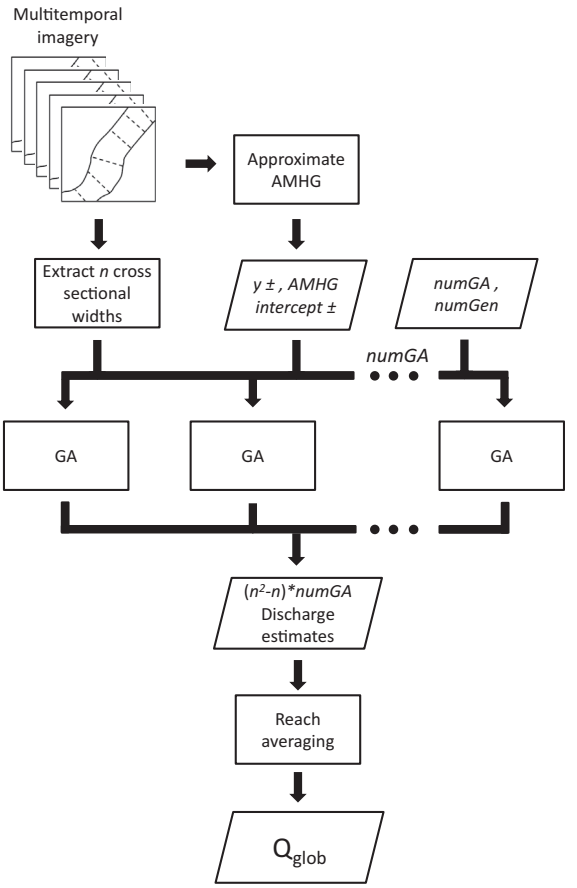

Figure 2. Conceptual diagram of discharge retrieval under the global parameterization recommended in this paper and yielding Q_{glob}. Repeat satellite images over a river gauging station are used to both approximate a river's AMHG (per equation (2)) and to extract *n* cross-sectional river top width measurements. These widths and the AMHG approximation are then passed to an ensemble of genetic algorithms (GAs). Each GA requires values of *y* and the AMHG intercept as approximated from imagery ± a tolerance (±0.10 each), as well as the number of GAs in an ensemble (*numGA*, 10) and how long each of these GAs searches for an optimal solution (*numGen*, 10). Each GA seeks to minimize the difference in discharge between *n* cross sections in a pairwise permutation, and together the ensemble yields a total of $(n^2 - n)*numGA$ discharge estimates. These estimates are then reach averaged by taking the mean of all cross-sectional discharges and yield a final Q_{glob} hydrograph. This process is entirely devoid of in situ information, and only requires multitemporal imagery as an input.

has trouble converging on plausible discharge values. That is, if candidate *a* and b values are allowed to couple randomly, the GA will not be able to estimate discharge accurately either because it takes an inordinate amount of time to initialize the GA or because of false convergence on extreme discharges. Therefore, a river's remotely sensed AMHG approximation is used to narrow the solution space for eligible AHG *a* and *b* parameters from a wide range of possible values (possible *a* values range from 1 to 600 and *b* values range from 0.01 to 0.8 in this study). This physically bounded solution space is then passed to the GA ensemble.

Instead of assuming a fixed AMHG from the remotely sensed empirical proxy *y* (equation (2)) as in *Gleason and Smith* [2014], here, the *y* value passed to the method was given a tolerance that allowed for inaccuracy in *y* rather than accepting it as exactly equal to 1/log(*E*). This produced a range of possible values for both the AMHG slope (*y*) and intercept ($y * \log(w_{glob})$). A number of these tolerances were tested within the sensitivity analyses to determine their impact on discharge retrieval (Table 2, also section 2.2.4), and this had the effect incrementally widening the AHG solution space described above. This change improved the processing speed of the method, and also enabled successful discharge retrieval even when the remotely sensed AMHG approximation was incongruous to observed AMHG.

Finally, the other parameters needed to estimate discharge include the number of GAs used in an ensemble (*numGA*), and the number of generations per GA (*numGen*) that determines how long each GA searches for an optimum solution. These parameters were varied according to the sensitivity analyses described below (Table 2). The final methodological inputs were discharge constraints passed to the GA that inhibit the GA from false convergence on extreme discharge values. The minimum discharge constraint is calculated assuming a minimum 0.5 m/s flow velocity and 0.5 m river depth multiplied by the minimum observed width for a given river, and the maximum discharge constraint is similarly calculated using 5 m/s flow velocity and 10 m river depth. These limits are reasonable considering the observed river widths in this study, and all gauge discharges fall within them.

Each GA is then applied in a pairwise permutation of all cross sections, producing $(n^2 - n) * (num\ GA)$ AHG parameter estimates and $(n^2 - n) * (numGA)$ direct discharge estimates per river, where *n* is the number of cross sections. Discharge at each of *n* cross sections is estimated by either taking summary statistics of GA produced *a* and *b* parameters to estimate cross-sectional AHG (and thus instantaneous *Q*, by combining with observed width per equation (3)), or by taking summary statistics of the final discharge estimates produced by GA optimization directly. Either way, the resultant *n* cross-sectional discharge estimates are then reach averaged to yield a single remotely sensed discharge retrieval for each image acquisition time. In the

Table 2. Method Parameterizations[a]

Test Number	*y* Tolerance	AMHG Intercept Tolerance	*numGA*	*numGen*
1	±0.05	±0.10	10	1
2	±0.05	±0.10	10	5
3	±0.05	±0.10	10	10
4	±0.05	±0.10	10	20
5	±0.05	±0.10	10	50
6	±0.05	±0.10	1	10
7	±0.05	±0.10	5	10
8	±0.05	±0.10	10	10
9	±0.05	±0.10	20	10
10	±0.05	±0.10	50	10
11	±0.05	±0.01	10	10
12	±0.05	±0.05	10	10
13	±0.05	±0.10	10	10
14	±0.05	±0.15	10	10
15	±0.05	±0.20	10	10
16	±0.01	±0.10	10	10
17	±0.05	±0.10	10	10
18	±0.10	±0.10	10	10
19	±0.15	±0.10	10	10
20	±0.20	±0.10	10	10

[a]Parameterizations used in the sensitivity analysis. The recommended global parameterization is row 18, and test numbers correspond to the *x* axes in Figure 5.

present paper, both discharge aggregation methods were evaluated to determine their impact on discharge retrieval accuracy.

2.2.4. Sensitivity Analyses

To assess the sensitivity of the discharge retrieval method across different river forms and method parameter settings, the method was applied to all 34 test sites using 20 different parameter combinations. This exercise tested each of the key inputs to the AMHG method, namely: the remotely sensed proxy *y*, the corresponding calculated AMHG intercept ($y * \log(w_{glob})$), the number of GAs used to retrieve discharge, and the number of generations per GA, providing a wealth of output from which to assess method performance (Table 2).

The sensitivity of discharge retrieval to reach averaging procedure was also tested using six different methods of aggregating GA output: calculating reach averaged hydrographs as the mean and median of discharge generated at each cross section by (1) the mean of all GA produced *a* and *b* values per cross section, (2) the median of all GA produced *a* and *b* values per cross section, and (3) direct GA discharge estimates. The aforementioned 20 parameterizations, coupled with these six different reach averaging procedures, yielded 120 different hydrographs for each river in this study. Note that the term "hydrograph" as used in this paper refers to intermittent flow estimates corresponding to dates of image acquisition, often spanning a decade or more. This differs from traditional daily or hourly temporal sampling of in situ hydrographs, with greatly reduced temporal representation of river flows. However, because the method does yield discharge estimates over time, the term hydrograph is retained.

In addition to studying the sensitivity of discharge retrieval from the full 120 hydrographs for every river, four specific hydrographs were extracted from among the suite of 120 for further interriver comparison and termed Q_{inv}, Q_{best}, Q_{glob}, and $Q_{glob+error}$. Q_{inv} represents an AHG "inversion" scenario, in which in situ gauge observations and satellite derived widths were used to calculate classical AHG *a* and *b* parameters at each cross section [see *Smith and Pavelsky*, 2008; *Pavelsky*, 2014, for other examples of this practice]. Remotely sensed widths were then input to these derived rating curves to estimate discharge at each cross section in every available satellite image. These discharges were then reach averaged by taking the mean of all cross-sectional flows to arrive at a final flow retrieval per image acquisition time, similar to *Pavelsky* [2014], in which the median of all cross-sectional flows was taken to produce a final hydrograph. As such, this inversion provides a current standard of satellite + in situ data discharge retrieval for comparison with the satellite-only AMHG discharge retrieval method.

The second specifically extracted hydrograph for each river, Q_{best}, was chosen simply as the most accurate discharge retrieval for each river from among all 120 hydrographs. Since this requires a priori knowledge of river discharge, choosing a hydrograph in this manner is equivalent to using this methodology to fill gaps in gauge records or to extend historical observations from a discontinued gauge. Without any reference flow information, it is impossible to calculate Q_{best}.

Finally, the sensitivity analyses were mined for an optimal parameterization that produced the highest quality output for the greatest number of rivers and was termed as the "global" parameterization (Q_{glob}). This global parameterization was determined after all sensitivity analyses had been run and forms a principal result of this paper as a recommended parameter set for future use in absence of in situ data. The robustness of Q_{glob} was further tested by adding random error to observed river widths to assess the effect of sensor resolution and width measurement errors on discharge retrieval accuracy ($Q_{glob+error}$). These random

Table 3. Remotely Sensed AMHG Proxy *y*

River	In Situ AMHG Slope	Proxy *y*	River	In Situ AMHG Slope	Proxy *y*
Snake	0.30	0.19	Iskut	0.35	0.45
Columbia	0.29	0.30	Tanana	0.32	0.47
Mekong	0.28	0.18	Taku	0.33	0.45
Sacramento	0.33	0.40	Solimoes	0.20	0.18
Yukon	0.30	0.19	Rio Branco	0.28	0.27
Rhone	0.34	0.04	Okavango	0.29	0.14
Kuskokwim	0.30	0.15	Murray	0.60	0.50
Athabasca	0.34	0.34	Luleaelven	0.31	0.04
Nemunas	0.37	0.06	Danube	0.13	0.00
Orange	0.47	0.39	Tennessee	0.05	0.50
Bani	0.42	0.24	Arkansas	0.01	0.30
Niger	0.34	0.43	Verkhynaya	0.20	0.23
Pahang	0.29	0.06	Ohio	0.07	0.18
Missouri	0.35	0.31	Sao Francisco	0.14	0.09
Upper Mississippi	0.27	0.25	Rio Madiera	0.13	0.11
Yangtze	0.25	0.24	Lower Mississippi	0.16	0.09
Saskatchewan	0.38	0.29	Mackenzie	0.04	0.60

errors were normally distributed about a mean of 0 m with standard deviation of 20 m for each width observation, yielding a 3σ error distribution of ~60 m, equivalent to a 2 pixel error on a width measurement from a 30 m resolution Landsat TM image. Each width observation in every river data set was subjected to a random error from this distribution, and the resulting data sets are designed to evaluate how well the method performs with poorly mapped width inputs.

All tests were assessed using the relative root mean square error (RRMSE) statistic, with relative residuals $(\frac{gauge-estimated}{gauge})$ used for each observation rather than direct residuals. This metric allows direct comparison of performance across the varied spatial scales of rivers in this study, and also places emphasis on the ability of the method to match gauge flows at each image acquisition time, rather than aggregating flow retrievals across time (which is inappropriate given the ~10 year span of flow retrievals). The reference discharge used to calculate the RRMSE for each test was the reported gauge discharge for each river, and all discharge retrieval performance is measured against this standard. Gauge discharge itself is subject to error both from overbank flows and from the error associated with the fit of the rating curve at each station. While in situ measurements of reference flows would provide a more accurate description of method performance than gauge flows, the global nature of this study rendered this alternative impossible and gauge data is used as the reference standard. The above sensitivity analyses and width error tests provide a robust means of assessing the performance of the AMHG discharge retrieval method, and increase confidence in the results and recommended parameterizations presented next.

3. Results

3.1. AMHG and Its Remotely Sensed Proxy

The presence of AMHG was observed in nearly all 34 rivers in this study, with r^2 values for the goodness of fit of equation (1) as built from in situ gauge data and remotely sensed widths ranging from 0.05 to 0.99 and a mean of 0.82 (Table 1). This confirms that AMHG is a persistent phenomenon in natural rivers [*Gleason and Smith*, 2014], with notable exceptions being the Arkansas, Tennessee, and Sao Francisco Rivers ($r^2 = 0.05$, 0.31, and 0.33, respectively). AMHG was strongly present in all morphological classes of rivers, but weak in rivers where mean cross-sectional AHG *b* exponent was less than 0.1 (hereafter termed "low-*b*" rivers). AMHG r^2 in these low-*b* rivers averaged 0.51, while mean AMHG r^2 in all remaining rivers was 0.91. Theoretically, discharge retrievals in rivers with weak AMHG are unlikely to be well estimated by the methodology, and discharge retrieval results bear this out, as will be discussed in section 3.2.

A key goal of this study was to assess the robustness of the remotely sensed proxy *y* and the resulting calculation of the AMHG intercept ($y * \log(w_{glob})$) from this value (equations (1) and (2)). The proxy *y* was a less robust approximation of $1/\log(E)$ in this study than was reported by *Gleason and Smith* [2014], as may be seen in Table 3. RMSE between observed $1/\log(E)$ and the remotely sensed approximation *y* for all rivers was 0.18, and RMSE between observed AMHG intercept and the remotely sensed intercept approximation was 0.49, which is considerably higher than the previously reported values by *Gleason and Smith* [2014].

Table 4. Discharge Retrieval Accuracy[a]

River	Q_{inv} RRMSE (%)	Q_{best} RRMSE (%)	Q_{glob} RRMSE (%)	Q_{glob} + width Error RRMSE (%)
Snake	15	18	79	47
Columbia	16	16	58	61
Mekong	43	27	42	58
Sacramento	14	12	27	82
Yukon	14	18	52	60
Rhone[b]	30	32	33	45
Kuskokwim	19	22	22	43
Athabasca	27	25	25	29
Nemunas	13	13	13	179
Orange[c]	166	83	1044	1270
Bani[c]	36	77	6416	3752
Niger	24	17	21	38
Pahang	16	25	27	74
Missouri	15	15	19	22
Upper Mississippi	20	20	60	63
Yangtze	18	39	74	67
Saskatchewan	37	33	266	92
Iskut[d]	88	50	53	57
Tanana[d]	20	72	79	81
Taku[d]	20	52	74	89
Solimoes	20	18	77	78
Rio Branco	20	25	33	39
Okavango[c]	20	25	1668	2097
Murray[c]	43	49	1935	1935
Luleaelven[e]	10	21	23	65
Danube[e]	42	27	39	47
Tennessee[e]	100	89	92	94
Arkansas[e]	64	83	557	83
Verkhynaya[e]	21	24	343	94
Ohio[e]	138	73	134	82
Sao Fransisco[e]	33	29	42	49
Rio Madiera[e]	40	55	81	82
Lower Mississippi[e]	16	28	80	79
McKenzie[e]	18	95	98	98

[a]Discharge estimation for study reaches following four different retrieval scenarios. The reduction of RRMSE when width error is added to some low-*b* data sets reflects a change from overestimation to underestimation of flows. The remotely sensed proxy *y* for the Rhone river was severely underestimated, and as such the slope of AMHG as built from gage data was used to retrieve discharge in this river.
[b]True AMHG used.
[c]Arid region river.
[d]Braided River.
[e]Low-*b* River.

The proxy *y* tends to underestimate 1/log(*E*), yet this did not preclude successful discharge retrieval, as will be discussed in section 3.4.

3.2. Sensitivity of AMHG Discharge Retrieval to River Geomorphology, Climate, and AHG b-Exponent

Remotely sensed AMHG discharge retrievals in rivers with Rosgen Class D (braided) morphology (Iskut, Tanana, and Taku Rivers), rivers with Koeppen-Gieger arid climatology and extreme flow variability (Bani, Okavango, Orange, and Murray Rivers), and low-*b* rivers having a mean cross-sectional AHG *b* value less than 0.1 (Luleaelven, Danube, Tennessee, Arkansas, Verkhynaya, Ohio, Sao Francisco, Madeira, Lower Mississippi, and McKenzie Rivers) were all poor. Low-*b* rivers were observed in Rosgen Classes B and C and were all found either upstream of confluences, near receiving waters such as dams and lakes, or where cross sections within the reach had near vertical floodwalls or banks, thus either physically constraining the river or providing backwater effects and resulting in low width-discharge sensitivity.

The availability of gauge data for the rivers in this study allows for the a posteriori selection of the best possible AMHG discharge retrieval parameterization from amongst all of the configurations tested for each individual river (Q_{best}). Q_{best} for each of the three rivers in Rosgen Class D (the Iskut, Tanana, and Taku Rivers) yielded RRMSE of 50, 72, and 52%, respectively (Table 4). This poor performance occurs despite strong AMHG presence in these rivers (AMHG $r^2 = 0.87, 0.99, 0.98$). In addition, Q_{best} tended to be nearly one order of magnitude less than observed flows, yet Q_{best} hydrographs tracked changes in discharge quite well (Figure 3). These rivers were not very sensitive to width error, but nonetheless the method proved inadequate at estimating their flow despite successful AHG inversions and strong AMHG (Table 4).

Discharge was also not well estimated in all four arid region rivers examined (Bani, Okavango, Orange, and Murray Rivers), as the method was unable to recreate the extreme variability of observed flows (in situ discharges for the Bani River, for instance, ranged from 3 to 894 m^3/s). Q_{best} RRMSE was 77, 25, 83, and 49% for the Bani, Okavango, Orange, and Murray Rivers, respectively. Quality of Q_{best} retrievals in these rivers was also sensitive to width accuracy error, yielding RRMSE values >1000% when normally distributed width errors were added to the data set (Table 3). Similar to braided rivers, this poor discharge retrieval occurred despite strong presence of AMHG ($r^2 = 0.97, 0.78, 0.80, 0.82$).

As with braided and arid region rivers, discharge was not accurately retrieved in low-*b* rivers, and average Q_{best} RRMSE for these low-*b* rivers was 52% (Table 4). In addition, low-*b* rivers exhibit less robust AMHG

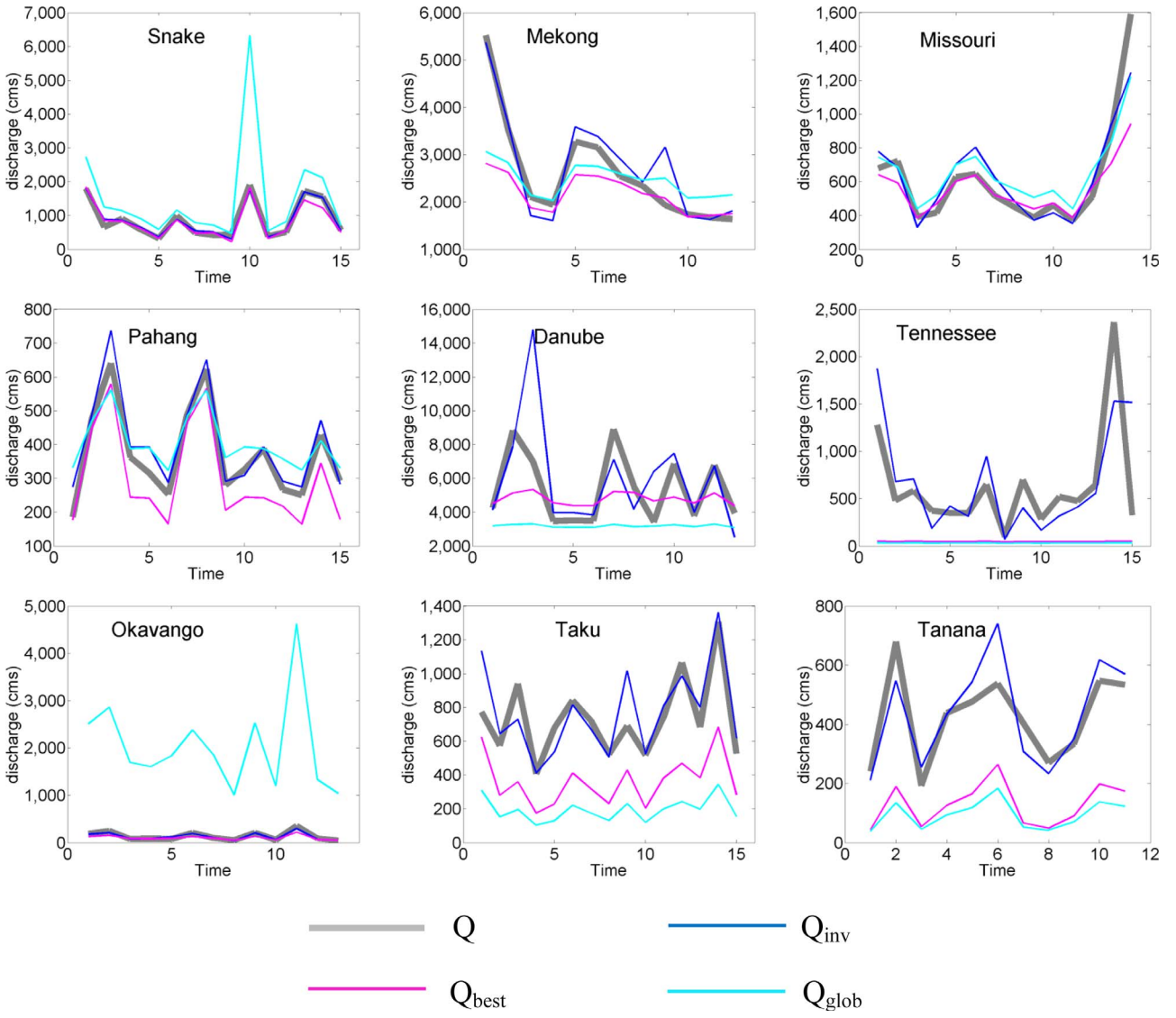

Figure 3. Comparison of in situ (Q) and satellite-derived discharge hydrographs for nine rivers. Satellite retrievals are Q_{best} (using gauge data to tune the retrieval), Q_{inv} (using remotely sensed widths within a rating curve built from gauge data), and Q_{glob} (using no in situ information whatsoever). The Snake, Mekong, Missouri, and Pahang Rivers meet the recommended geomorphic criteria in this study. The other five rivers yield poor discharge retrievals, attributed to braided morphology (Taku, Tanana), arid-climate discharge variability (Okavango), and low AHG b-exponent (Danube, Tennessee).

behavior than any other class of rivers. Mean AMHG r^2 for these low-b rivers was 0.51, and removing these 10 rivers from the collective 34 river data set yields a mean AMHG r^2 of 0.91 for all remaining rivers. Similarly, the remotely sensed proxy functioned poorly over low-b rivers, as y and AMHG intercept had mean RMSE of 0.26 and 0.77, respectively. Removing low-b rivers from the data set yielded y and AMHG intercept RMSE of 0.13 and 0.32, respectively, for all remaining rivers. Most of these low-b rivers tended to be very large (wide), but not all large rivers (e.g., the Mekong, Upper Mississippi, Athabasca, and Solimoes rivers) were low-b.

In addition, Q_{best} was relatively invariant for low-b rivers, and output hydrographs poorly tracked changes in discharge as reported by gauge data. This lack of temporal variability was masked by lower RRMSE values when such static hydrographs occurred near mean flow (e.g., Luleaelven, Danube, Sao Francisco Rivers, see Figure 3). Such invariant output would also be an indicator of a poor retrieval for rivers with moderate b values, because AHG predicts that in these rivers a unit change in width will result in a proportionally larger change in discharge (as b is always less than 1 in equation (3)). Therefore, it is likely that a discharge retrieval with variance (over time) less than or equal to input width variance does not accurately represent flows and should be excluded from analysis.

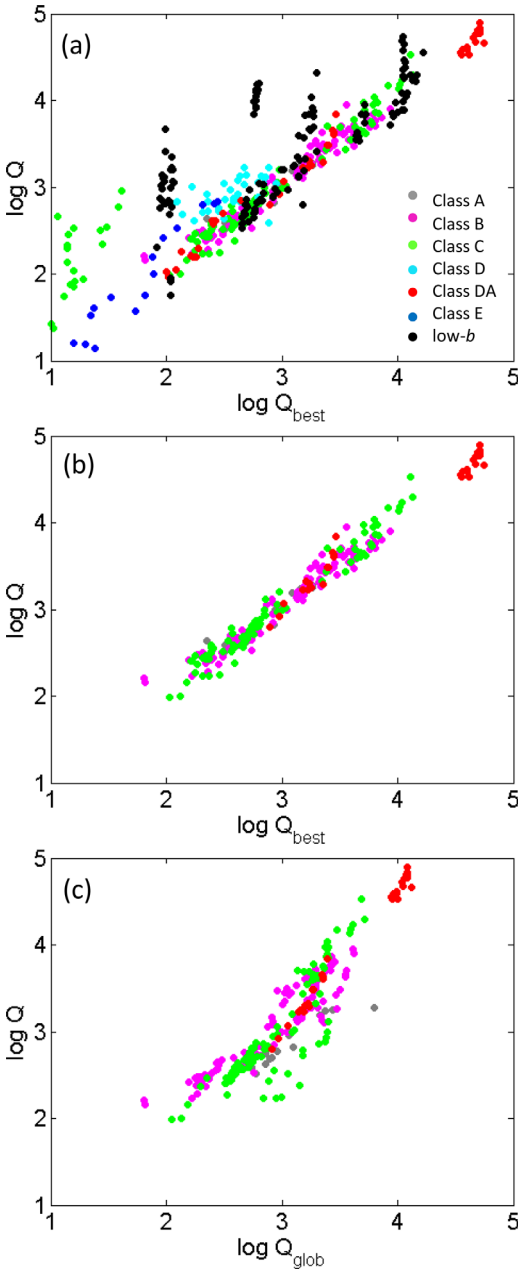

Figure 4. Satellite-derived discharges plotted versus in situ gauge observations for (a) all 34 rivers assessed in this study, using site-specific parameterizations tailored for each river (Q_{best}); (b) a culled data set of 17 operational rivers (i.e., with braided, arid, and low-*b* rivers removed), again using in situ knowledge to identify optimal parameters (Q_{best}); and (c) the same culled data set using no in situ knowledge and a global recommended parameterization (Q_{glob}). With the exception of braided rivers (cyan, Class D) and low-*b* rivers (black), river Rosgen Class imparts little systematic impact on satellite-derived discharge retrievals. Median discharge retrieval RRMSE for (b and c) are 20 and 33%, respectively.

The method poorly retrieved discharge in the above three classes of rivers, yet it performed effectively for all remaining rivers (17 of 34 rivers remain after exclusion of braided, arid-climate, and low-*b* rivers, see Figure 4 and Tables 1 and 4). Removal of all arid rivers (4), braided rivers (3) and low-*b* rivers (10) from the collective 34 river data set improves median Q_{best} RRMSE from 27 to 20% (Figures 4a and 4b). The remaining rivers encompass four Rosgen Classes (A, B, C, DA) and a range of nonarid climates. The following results refer to these 17 rivers (Table 1), with braided, arid, and low-*b* rivers herewith excluded from further summary statistics and sensitivity testing.

3.3. Sensitivity of Discharge Retrieval to Reach Averaging Procedure and Parameterization

The AMHG discharge retrieval method is highly sensitive to its reach averaging procedure (Figure 5). Calculating final hydrographs based on GA produced AHG parameters (top four plots, Figure 5) results in much less accurate discharge retrievals than using GA produced flows directly (bottom two plots, Figure 5). In addition, using GA produced AHG parameters exhibits relatively insensitive discharge retrieval accuracy across parameterizations (given in Table 2), while using direct GA discharge estimates show much greater variation in discharge retrieval accuracy as retrieval method parameters are changed. Finally, retrieving discharge by using the mean (left column, Figure 5) or median (Figure 5, right column) of a particular reach averaging procedure exhibits a lack of outlier discharge retrievals, as mean and median retrievals track each other quite well. Only when using direct GA produced discharge estimates do results suggest that the method is susceptible to outliers when following the recommended geomorphic criteria.

The method is also quite sensitive to both the remotely sensed proxy y and the associated calculation of remotely sensed AMHG intercept ($y * \log(w_{glob})$). These parameters exhibit discharge retrieval accuracy sensitivities with distinctive "U" shapes around their central values, suggesting that the AMHG method is less able to resolve discharge when the AMHG tolerance is either too wide or too narrow (Figure 5). In addition, all reach averaging procedures were sensitive to these two parameters, further confirming their importance to discharge retrieval.

Discharge retrieval accuracy is relatively insensitive to the number of GAs and the number generations per GA. Only reach averaging by taking the mean of all GA produced flows (as in the global parameterization,

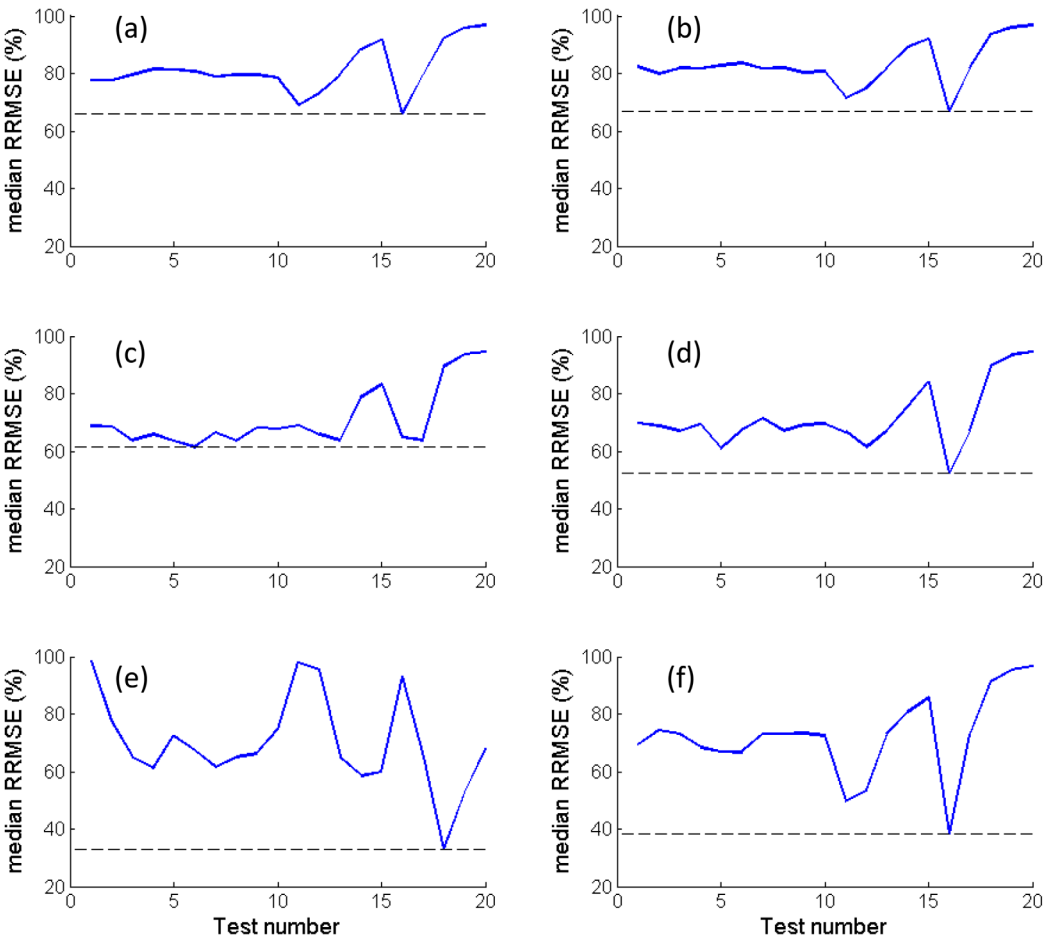

Figure 5. Sensitivity of Q_{glob} to inputs required by the AMHG discharge retrieval method (*y*, AMHG intercept, number of GAs, number of generations per GA, and reach averaging procedure), for 17 operational rivers (nonbraided, nonarid, nonlow-*b*) only. Plots correspond to reach averaging procedure as follows: (a) mean of mean GA produced AHG; (b) median of mean GA produced AHG; (c) mean of median GA produced AHG; (d) median of median GA produced AHG; (e) mean of GA produced discharge; and (f) median of GA produced discharge. Test number refers to the number of the parameterization given in Table 2; blue lines show median discharge retrieval RRMSE of the operational rivers under each parameterization. Black lines show minimum RRMSE for each reach-averaging method. Q_{glob} is most sensitive to tests 11–20 (*y* and AMHG intercept tolerances), and only sensitive to the number of GAs and generations per GA when taking the mean/median of GA produced flows (tests 1–10, bottom two plots). The recommended global parameterization appears in the bottom left plot (Test number 18), yielding a median RRMSE of 33% for all operational rivers.

see section 3.4) yields sharp discharge retrieval improvement with increasing numbers of generations per GA. Collectively, these sensitivity results indicate the method is quite sensitive to parameters *y*, AMHG intercept, and reach averaging procedure, and is much less sensitive to the number of GAs or the number of generations in each GA.

3.4. Discharge Retrieval Accuracy and Global Parameterization

For those 17 rivers meeting the recommended geomorphic criteria ("operational" rivers: nonlow-*b*, nonarid region, nonbraided), the RRMSE of Q_{best} ranged from 12 to 33%, with mean RRMSE of 22% (Table 4 and Figure 5). Separating these results morphologically reveals that Rosgen Classes B and C had mean Q_{best} RRMSE of 22 and 23%, respectively. The two anastomosing Amazonian rivers (Solimoes and Rio Branco, Rosgen Class DA) had Q_{best} RRMSE of 18 and 25%, respectively. These results indicate that river form had little impact on Q_{best}, as Q_{best} was estimated with similar accuracy across Rosgen Classes A, B, C, and DA. Therefore, the methodology may be applied across a wide range of river morphologies and scales, excepting braided rivers where retrievals were poor.

The global parameterization that produced the most accurate output in the greatest number of rivers (Q_{glob}) was found to occur when *y* and AMHG intercept values passed to the GA were given a tolerance of

0.10, when the number of GAs and generations per GA were set to 10, and the mean of all GA produced flows was calculated to produce the final hydrograph. These values represent the central values of each parameter range in the sensitivity analyses. Q_{glob} (which results from these parameters) has median RRMSE of 33% and mean RRMSE of 55% over the operational rivers (Figure 4 and Table 4). The RMSE of the increase in Q_{glob} RRMSE resulting from addition of normally distributed errors to river widths (i.e., the increase in error when widths were randomly degraded: Q_{glob} versus $Q_{glob+error}$) under the global parameterization was 20.6%. The global parameterization is recommended for use in absence of in situ measurements as the default parameter set for the AMHG discharge retrieval method.

4. Discussion

The results of this study indicate that the AMHG discharge retrieval method can successfully retrieve in-channel discharges for a variety of rivers around the world solely from remotely sensed images. Notable exceptions are braided rivers, low-*b* rivers, and rivers displaying extreme variability in discharge as manifested here in arid-climate rivers. Excluding these environments, Q_{best} (selected by examining gauge records a posteriori) retrievals mimic gauge observations quite well (median RRMSE 20%) suggesting that the method best retrieves discharge when some calibrating in situ information is available.

Q_{best} hydrographs also track Q_{inv} hydrographs closely on all 34 rivers, suggesting that the method is able to recreate flows with equivalent accuracy to other width-based satellite methods requiring in situ data [e.g., *Smith and Pavelsky*, 2008; *Pavelsky*, 2014]. The differences between Q_{best}, Q_{inv}, and Q_{glob} (Figure 3) highlight the sensitivity of the method to its parameterization, but the method architecture (as evidenced by Q_{best} versus Q_{inv}), appears to be sound. These results confirm that the AMHG discharge retrieval method should be able to address knowledge gaps in global river runoff at stations where gauges have been discontinued, as the method can be tuned from past observations even if the gauge is no longer functional.

Practitioners seeking to use the AMHG method to retrieve river flows in ungauged basins will not have in situ discharge with which to compare results and optimize parameterizations. Absent in situ knowledge, the results of this study suggest that useful river discharges (Q_{glob}) can still be retrieved. Based on our analyses, we propose that in such situations, a recommended global parameterization consisting of 10 GAs of 10 generations, *y* and AMHG intercept tolerances of ± 0.10, and discharge calculated as the mean of all GA produced flows be applied. Such a parameterization, applied without in situ information to the rivers studied here, yielded Q_{glob} discharge retrievals ranging from 13 to 260% RRMSE with median RRMSE of 33% and RRMSE standard deviation of 7.5%. This lends confidence that the method can provide useful discharge estimates where none currently exist, with accuracies within $\sim$26–41% of what might be expected from in situ river gauge observations.

The purely remotely sensed approximation *y* discovered by *Gleason and Smith* [2014] was less robust in the present study, with a tendency of *y* to underestimate 1/log(*E*) of a river's in situ AMHG (Table 3). However, this did not preclude successful discharge retrieval, as the methodology presented here enabled the GA to tune both a river's AMHG and AHG simultaneously. This change is critical, as any incongruity between *y* and 1/log(*E*) would preclude successful discharge retrieval if *y* were taken as fixed. Further investigation is required to better understand the proxy and its relationship to AMHG, but the proxy proved sufficient for discharge retrieval in this paper.

Without AMHG, the methodology is unable to produce coherent discharge retrievals, yet it is very difficult to perform an uncertainty analysis that numerically describes how error in the approximation of AMHG affects the accuracy of final discharge retrieval given the nature of the procedure. First, any GA tends toward "black box" solutions to a particular problem even though the GA used here is well understood. Second, because discharge is aggregated from thousands of different GA results ($(n^2 - n) * numGA$, discharges, where *n* is the 15 and *numGA* is 10 in the global parameterization), the error propagation of any one GA is somewhat muted. Therefore, while a more accurately approximated AMHG will yield more accurate discharge retrievals, it is difficult to quantify this statement.

Results indicate that the method is sensitive to the accuracy of input river widths. The addition of normally distributed random error ($\simeq$ 60 m at 3σ) to input cross-sectional width data increased corresponding Q_{glob} retrieval error by 20% across all operational rivers (RMSE of increase in RRMSE was 20.6%, calculated from

Table 4). This is logical, as large width errors (especially in narrower rivers) can cause some cross sections to show an increase in width while others show a decrease (violating mass conservation) when such errors are added. This sensitivity requires that input widths be carefully mapped to assure the highest quality discharge retrieval, particularly for smaller rivers.

Despite the successes of the AMHG discharge retrieval method for most river forms, this research identifies three conditions that are inappropriate for its use. First, Q_{glob} in so-called low-*b* reaches is poorly retrieved. This is likely because inverting observed widths through small *b* exponents, especially in wider rivers, leads to enormous flow values that are rejected by the discharge constraints imposed on the GA, thus retaining only smaller values from incorrect *b* exponents. In addition, the observed AMHG for these rivers was much weaker than for all other rivers, and the remotely sensed approximation to this already weak AMHG was also poor in these rivers. Second, the methodology is unable to recreate the extreme flow variability seen in arid region rivers. This is consistent with previously reported poor performance for the Rio Grande River [*Gleason and Smith*, 2014], and may correspond to distinct subbankfull geometry changes that affect the AHG of these rivers or inaccuracies in gauge-derived rating curves [*Ferguson*, 1986; *Mersel et al*., 2013]. Finally, Q_{glob} hydrographs of braided rivers in this study were all underestimated by the methodology, despite tracking relative changes in discharge quite well (Figure 3). It may be that this underestimation is due to the large *b* values often found in braided rivers, but it is not clear at this time why exactly all of these rivers were underestimated by the methodology. Replacing cross-sectional width with effective width W_e [inundation area divided by reach length, *Smith et al*., 1995, 1996] for short (1 km or less) reaches did not improve performance in these rivers, suggesting that braided rivers should be avoided when using the present methodology in ungauged basins.

Critically, the above predictably poor discharge retrieval categories may be identified without in situ information. Braided rivers are easily detected from satellite imagery, and low-*b* rivers can also potentially be identified from imagery by observing changes in width over time relative to mean width. Furthermore, low-*b* rivers tend to be found near confluences, hydraulic controls, and receiving waters, all of which may also be easily identified from satellite images, and their invariant output hydrographs are a sign of a poor retrieval (if less variable over time than width, per equation (3)). The forthcoming NASA/CNES SWOT satellite can also aid detection of low-*b* rivers wider than 100 m [50 m ideally, *Pavelsky et al*., 2014] by providing simultaneous measurements of width and stage, and could show that a river is getting deeper but not wider, a signature of low-*b* rivers. Arid-region rivers are more difficult to determine from imagery alone, but climate maps and the presence of persistent dry (brown) vegetation in imagery can help determine whether or not a study reach is in an arid region.

For reasons that are still unknown, studied reaches on the Yangtze, Saskatchewan, Upper Mississippi, and Columbia Rivers that met the recommended geomorphic criteria had poor Q_{glob} accuracy (RRMSE 74, 266, 60, and 58%, respectively), despite accurate Q_{best} retrievals (39, 34, 20, and 18%). These results suggest that while it is possible to use AMHG calibrated with in situ observations to retrieve discharge (i.e., Q_{best}), using the global parameters does not always yield accurate results. Future work should investigate how particular parameterizations, especially different reach averaging methods, impact discharge retrieval in such cases. More work is required to understand the physical underpinnings of 1/log(*E*), and why the remotely sensed proxy *y* can approximate it. Finally, future work should also consider the nature of AMHG and its relationship to known hydrologic quantities, and employ the AMHG discharge retrieval method in ungauged and poorly-gauged regions to gain further understanding of the global water cycle.

5. Conclusion

The AMHG discharge retrieval method advanced here yields reasonably accurate retrievals of river discharge when derived solely from satellite images. Practitioners seeking to use this method in ungauged basins should adopt the global parameterization described above, and carefully analyze the temporal output, as little variation in discharge output is an indicator of a poorly retrieved hydrograph (provided input widths are sufficiently variable over time). It is of paramount importance that study reaches remain mass conserved and free of tributaries or outflows, and practitioners should avoid choosing reaches near receiving waters, confluences, and hydrologic controls (as these conditions can create low-*b* reaches). Braided rivers, and rivers in arid regions likely to exhibit flow variations of several orders of magnitude, should also be

avoided. Rivers narrower than ~100 m are unlikely to be well retrieved from Landsat TM satellite imagery due to insufficiently precise width measurements, and even very large rivers can problematic if they have small cross-sectional AHG *b* exponents (<0.1). Following these recommendations, our analysis suggests that discharge retrievals having ~26–41% agreement with traditional in situ gauge estimates can likely be made purely from satellite imagery.

Acknowledgments
This research was supported by the NASA Remote Sensing Theory initiative (grant NNX12AB41G), NASA Surface Water Ocean Topography (SWOT) mission (grant NNX13AD88G), and NASA Earth and Space Sciences Fellowship NNX12AN32H. Gauge data were provided by The Global Runoff Data Centre, The USGS, and the Changjiang Wuhan Waterway Bureau. J. Wang provided constructive feedback on the sensitivity analyses herein. All code used to generate the results in this study is available upon request. Comments by three anonymous reviewers greatly improved the clarity and quality of this manuscript.

References

Alsdorf, D. E., E. Rodriguez, and D. P. Lettenmaier (2007), Measuring surface water from space, *Rev. Geophys.*, *45*, RG2002, doi:10.1029/2006RG000197.

Andreadis, K. M., E. A. Clark, D. P. Lettenmaier, and D. E. Alsdorf (2007), Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model, *Geophys. Res. Lett.*, *34*, L10403, doi:10.1029/2007GL029721.

Ashmore, P., and E. Sauks (2006), Prediction of discharge from water surface width in a braided river with implications for at-a-station hydraulic geometry, *Water Resour. Res.*, *42*, W03406, doi:10.1029/2005WR003993.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier (2005), Potential impacts of a warming climate on water availability in snow-dominated regions, *Nature*, *438*(7066), 303–309, doi:10.1038/nature04141.

Bjerklie, D. M. (2007), Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information, *J. Hydrol.*, *341*(3–4), 144–155, doi:10.1016/j.jhydrol.2007.04.011.

Bjerklie, D. M., S. L. Dingman, C. J. Vorosmarty, C. H. Bolster, and R. G. Congalton (2003), Evaluating the potential for measuring river discharge from space, *J. Hydrol.*, *278*(1–4), 17–38, doi:10.1016/s0022-1694(03)00129-x.

Bjerklie, D. M., D. Moller, L. C. Smith, and S. L. Dingman (2005), Estimating discharge in rivers using remotely sensed hydraulic information, *J. Hydrol.*, *309*(1–4), 191–209, doi:10.1016/j.jhydrol.2004.11.022.

Brakenridge, G. R., S. V. Nghiem, E. Anderson, and R. Mic (2007), Orbital microwave measurement of river discharge and ice status, *Water Resour. Res.*, *43*, W04405, doi:10.1029/2006WR005238.

Conway, D., and M. Hulme (1993), Recent fluctuations in precipitation and runoff over the Nile sub-basins and their impact on main Nile discharge, *Clim. Change*, *25*(2), 127–151, doi:10.1007/BF01661202.

Conway, D., A. Persechino, S. Ardoin-Bardin, H. Hamandawana, C. Dieulin, and G. Mahé (2009), Rainfall and water resources variability in Sub-Saharan Africa during the twentieth century, *J. Hydrometeorol.*, *10*(1), 41–59, doi:10.1175/2008JHM1004.1.

Dai, A., and K. E. Trenberth (2002), Estimates of freshwater discharge from continents: Latitudinal and seasonal variations, *J. Hydrometeorol.*, *3*(6), 660–687, doi:10.1175/1525-7541(2002)003<0660:EOFDFC>2.0.CO;2.

Dai, A., T. Qian, K. E. Trenberth, and J. D. Milliman (2009), Changes in continental freshwater discharge from 1948 to 2004, *J. Clim.*, *22*(10), 2773–2792, doi:10.1175/2008JCLI2592.1.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. Evolut. Comput.*, *6*(2), 182–197, doi:10.1109/4235.996017.

Dinar, S. (2012), The geographical dimensions of hydro-politics: International freshwater in the Middle East, North Africa, and Central Asia, *Eurasian Geogr. Econ.*, *53*(1), 115–142, doi:10.2747/1539-7216.53.1.115.

Durand, M., K. M. Andreadis, D. E. Alsdorf, D. P. Lettenmaier, D. Moller, and M. Wilson (2008), Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model, *Geophys. Res. Lett.*, *35*, L20401, doi:10.1029/2008GL034150.

Durand, M., E. Rodriguez, D. E. Alsdorf, and M. Trigg (2010a), Estimating river depth from remote sensing swath interferometry measurements of river height, slope, and width, *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.*, *3*(1), 20–31, doi:10.1109/jstars.2009.2033453.

Durand, M., L.-L. Fu, D. P. Lettenmaier, D. E. Alsdorf, E. Rodriguez, and D. Esteban-Fernandez (2010b), The surface water and ocean topography mission: Observing terrestrial surface water and oceanic submesoscale eddies, *Proc. IEEE*, *98*(5), 766–779, doi:10.1109/jproc.2010.2043031.

Ferguson, R. I. (1986), Hydraulics and hydraulic geometry, *Prog. Phys. Geogr.*, *10*(1), 1–31, doi:10.1177/030913338601000101.

Gilvear, D. J., C. Davids, and A. N. Tyler (2004), The use of remotely sensed data to detect channel hydromorphology; River Tummel, Scotland, *River Res. Appl.*, *20*(7), 795–811, doi:10.1002/rra.792.

Gleason, C. J., and J. Im (2012), Forest biomass estimation from airborne LiDAR data using machine learning approaches, *Remote Sens. Environ.*, *125*, 80–9191, doi:10.1016/j.rse.2012.07.006.

Gleason, C. J., and L. C. Smith (2014), Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry, *PNAS*, *111*(13), 4788–4791, doi:10.1073/pnas.1317606111.

Grefenstette, J. J. (1986), Optimization of control parameters for genetic algorithms, *IEEE Trans. Syst. Man Cybernet.*, *16*(1), 122–128, doi:10.1109/tsmc.1986.289288.

Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, N. Shirakawa, Y. Shen, and K. Tanaka (2008), An integrated model for the assessment of global water resources—Part 2: Applications and assessments, *Hydrol. Earth Syst. Sci.*, *12*(4), 1027–1037, doi:10.5194/hess-12-1027-2008.

Hannah, D. M., S. Demuth, H. A. J. van Lanen, U. Looser, C. Prudhomme, G. Rees, K. Stahl, and L. M. Tallaksen (2011), Large-scale river flow archives: Importance, current status and future needs, *Hydrol. Processes*, *25*(7), 1191–1200, doi:10.1002/hyp.7794.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel (2006), World Map of the Köppen-Geiger climate classification updated, *Meteorol. Z.*, *15*(3), 259–263, doi:10.1127/0941-2948/2006/0130.

Lammers, R. B., A. I. Shiklomanov, C. J. Vorosmarty, B. M. Fekete, and B. J. Peterson (2001), Assessment of contemporary Arctic river runoff based on observational discharge records, *J. Geophys. Res.*, *106*(D4), 3321–3334, doi:10.1029/2000JD900444.

Lee, H., R. E. Beighley, D. Alsdorf, H. C. Jung, C. K. Shum, J. Duan, J. Guo, D. Yamazaki, and K. Andreadis (2011), Characterization of terrestrial water dynamics in the Congo Basin using GRACE and satellite radar altimetry, *Remote Sens. Environ.*, *115*(12), 3530–3538, doi:10.1016/j.rse.2011.08.015.

Leopold, L. B., and T. Maddock (1953), *The Hydraulic Geometry of Stream Channels and Some Physiographic Implications*, U.S. Geol. Surv., Washington, D. C.

Lettenmaier, D. P., and J. S. Famiglietti (2006), Hydrology: Water from on high, *Nature*, *444*(7119), 562–563, doi:10.1038/444562a.

Mersel, M. K., L. C. Smith, K. M. Andreadis, and M. T. Durand (2013), Estimation of river depth from remotely sensed hydraulic relationships, *Water Resour. Res.*, *49*, 3165–3179, doi:10.1002/wrcr.20176.

Milliman, J. D., K. L. Farnsworth, P. D. Jones, K. H. Xu, and L. C. Smith (2008), Climatic and anthropogenic factors affecting river discharge to the global ocean, 1951–2000, *Global Planet. Change*, *62*(3–4), 187–194, doi:10.1016/j.gloplacha.2008.03.001.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Stationarity is dead: Whither water management?, *Science*, *319*(5863), 573–574, doi:10.1126/science.1151915.

Nathanson, M., J. W. Kean, T. J. Grabs, J. Seibert, H. Laudon, and S. W. Lyon (2012), Modelling rating curves using remotely sensed LiDAR data, *Hydrol. Processes*, *26*(9), 1427–1434, doi:10.1002/hyp.9225.

Neal, J., G. Schumann, P. Bates, W. Buytaert, P. Matgen, and F. Pappenberger (2009), A data assimilation approach to discharge estimation from space, *Hydrol. Processes*, *23*(25), 3641–3649, doi:10.1002/hyp.7518.

Oki, T., K. Musiake, H. Matsuyama, and K. Masuda (1995), Global atmospheric water-balance and runoff from large river basins, *Hydrol. Processes*, *9*(5–6), 655–678, doi:10.1002/hyp.3360090513.

Oki, T., T. Nishimura, and P. Dirmeyer (1999), Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP), *J. Meteorol. Soc. Jpn.*, *77*(1B), 235–255.

Pavelsky, T. M. (2014), Using width-based rating curves from spatially discontinuous satellite imagery to monitor river discharge, *Hydrol. Processes*, *28*(6), 3035–3040, doi:10.1002/hyp.10157.

Pavelsky, T. M., M. T. Durand, K. M. Andreadis, R. E. Beighley, R. C. D. Paiva, G. H. Allen, and Z. F. Miller (2014), Assessing the potential global extent of SWOT river discharge observations, *J. Hydrol.*, *519*(B), 1516–1525, doi:10.1016/j.jhydrol.2014.08.044.

Peel, M. C., and T. A. McMahon (2006), Continental runoff: A quality-controlled global runoff data set, *Nature*, *444*(7120), E14–E14, doi:10.1038/nature05480.

Perry, G. D., P. B. Duffy, and N. L. Miller (1996), An extended data set of river discharges for validation of general circulation models, *J. Geophys. Res.*, *101*(D16), 21,339–21,349, doi:10.1029/96JD00932.

Phillips, J. D. (1990), The instability of hydraulic geometry, *Water Resour. Res.*, *26*(4), 739–744, doi:10.1029/WR026i004p00739.

Ran, L., S. Wang, and X. X. Lu (2012), Hydraulic geometry change of a large river: A case study of the upper Yellow River, *Environ. Earth Sci.*, *66*(4), 1247–1257, doi:10.1007/s12665-011-1336-x.

Rosgen, D. L. (1994), A classification of natural rivers, *Catena*, *22*(3), 169–199, doi:10.1016/0341-8162(94)90001-9.

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, doi:10.1623/hysj.48.6.857.51421.

Smith, L. C. (1997), Satellite remote sensing of river inundation area, stage, and discharge: A review, *Hydrol. Processes*, *11*(10), 1427–1439, doi:10.1002/(SICI)1099-1085(199708)11:10<1427::AID-HYP473>3.3.CO;2-J.

Smith, L. C., and T. M. Pavelsky (2008), Estimation of river discharge, propagation speed, and hydraulic geometry from space: Lena River, Siberia, *Water Resour. Res.*, *44*, W03427, doi:10.1029/2007WR006133.

Smith, L. C., B. L. Isacks, R. R. Forster, A. L. Bloom, and I. Preuss (1995), Estimation of discharge from braided glacial rivers using ERS 1 synthetic aperture radar: First results, *Water Resour. Res.*, *31*(5), 1325–1329, doi:10.1029/95WR00145.

Smith, L. C., B. L. Isacks, A. L. Bloom, and A. B. Murray (1996), Estimation of discharge from three braided rivers using synthetic aperture radar satellite imagery: Potential application to ungaged basins, *Water Resour. Res.*, *32*(7), 2021–2034, doi:10.1029/96WR00752.

Syed, T. H., J. S. Famiglietti, V. Zlotnicki, and M. Rodell (2007), Contemporary estimates of Pan-Arctic freshwater discharge from GRACE and reanalysis, *Geophys. Res. Lett.*, *34*, L19404, doi:10.1029/2007GL031254.

Syed, T. H., J. S. Famiglietti, and D. P. Chambers (2009), GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales, *J. Hydrometeorol.*, *10*(1), 22–40, doi:10.1175/2008jhm993.1.

Syed, T. H., J. S. Famiglietti, D. P. Chambers, J. K. Willis, and K. Hilburn (2010), Satellite-based global-ocean mass balance estimates of interannual variability and emerging trends in continental freshwater discharge, *Proc. Natl. Acad. Sci. U. S. A.*, *107*(42), 17,916–17,921, doi:10.1073/pnas.1003292107.

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

Van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens (2011), Global monthly water stress. 1: Water balance and water availability, *Water Resour. Res.*, *47*, W07517, doi:10.1029/2010WR009791.

Vörösmarty, C. J., et al. (2010), Global threats to human water security and river biodiversity, *Nature*, *467*(7315), 555–561, doi:10.1038/nature09440.

Wada, Y., L. P. H. van Beek, C. M. van Kempen, J. W. T. M. Reckman, S. Vasak, and M. F. P. Bierkens (2010), Global depletion of groundwater resources, *Geophys. Res. Lett.*, *37*, L20402, doi:10.1029/2010GL044571.

Widen-Nilsson, E., S. Halldin, and C. Xu (2007), Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation, *J. Hydrol.*, *340*(1–2), 105–118, doi:10.1016/j.jhydrol.2007.04.002.