Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR027949

Constraining Remote River Discharge Estimation Using Reach-Scale Geomorphology

C. B. Brinkerhoff[1], C. J. Gleason[1], D. Feng[1], and P. Lin[2]

[1]Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA, USA, [2]Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA

Key Points:
- We introduce geoBAM, an improved version of a remote sensing of river discharge algorithm (BAM)
- geoBAM is tested successfully on over 7,500 river reaches in the Canadian Arctic and on a suite of simulated SWOT observations
- Discharge estimation accuracy is improved by using river-specific prior information informed by classification of a large field data set

Supporting Information:
- Supporting Information S1

Correspondence to:
C. B. Brinkerhoff,
cbrinkerhoff@umass.edu

Citation:
Brinkerhoff, C. B., Gleason, C. J., Feng, D., & Lin, P. (2020). Constraining remote river discharge estimation using reach-scale geomorphology. *Water Resources Research*, *56*, e2020WR027949. https://doi.org/10.1029/2020WR027949

Received 14 MAY 2020
Accepted 14 OCT 2020
Accepted article online 24 OCT 2020

Abstract Recent advances in remote sensing and the upcoming launch of the joint NASA/CNES/CSA/UKSA Surface Water and Ocean Topography (SWOT) satellite point toward improved river discharge estimates in ungauged basins. Existing discharge methods rely on "prior river knowledge" to infer parameters not directly measured from space. Here, we show that discharge estimation is improved by classifying and parameterizing rivers based on their unique geomorphology and hydraulics. Using over 370,000 in situ hydraulic observations as training data, we test unsupervised learning and an "expert" method to assign these hydraulics and geomorphology to rivers via remote sensing. This intervention, along with updates to model physics, constitutes a new method we term "geoBAM," an update of the Bayesian At-many-stations hydraulic geometry-Manning's (BAM) algorithm. We tested geoBAM on Landsat imagery over more than 7,500 rivers (108 are gauged) in Canada's Mackenzie River basin and on simulated hydraulic data for 19 rivers that mimic SWOT observations without measurement error. geoBAM yielded considerable improvement over BAM, improving the median Nash-Sutcliffe efficiency (NSE) for the Mackenzie River from −0.05 to 0.26 and from 0.16 to 0.46 for the SWOT rivers. Further, NSE improved by at least 0.10 in 78/108 gauged Mackenzie rivers and 8/19 SWOT rivers. We attribute geoBAM improvement to parameterizing rivers by type rather than globally, but prediction accuracy worsens if parameters are misassigned. This method is easily mapped to rivers at the global scale and paves the way for improving future discharge estimates, especially when coupled with hydrologic models.

1. Introduction

In recent decades, remote sensing (RS) of rivers has flourished as a subfield within fluvial geomorphology and hydrology. At the global scale, RS of rivers is changing current perceptions of rivers and their role in the Earth system: Globally modeled hydrography at fine-spatial scales (Lehner et al., 2008; Yamazaki et al., 2019), daily runoff routed through almost 3 million river reaches over 30 years (Lin et al., 2019), assessments of rivers and climate (Yang et al., 2020), water quality (Ross et al., 2019), surface area (Allen & Pavelsky, 2018), and hydrological connectivity (Grill et al., 2019) have all debuted recently. These manuscripts extend a continuation of RS for hydrology going back several decades (see Lettenmaier et al., 2015, and Gleason & Durand, 2020, for thorough reviews). These examples, along with similar recent work quantifying global fluvial geomorphic patterns (e.g., Chen et al., 2019; Frasson, Pavelsky, et al., 2019; Lin et al., 2020), suggest that RS is coming of age in its ability to provide global-scale data that honors local differences in rivers. These ideas will be further explored with the launch of the joint NASA/CNES/CSA/UKSA Surface Water and Ocean Topography (SWOT) satellite in 2022, which is expected to provide measurements of water surface elevation and extent at unprecedented spatial and temporal resolutions (Biancamaria et al., 2016).

A particular subset of this literature is showing that accurate global RS of river discharge (RSQ) is presently possible with some gauging information in hand and should be globally possible in ungauged basins in the near future (Gleason & Durand, 2020). In basins with stream gauges or extensive field measurements, RSQ approaches frequently calibrate RS to local channel hydraulics (e.g., Brakenridge et al., 2007; LeFavour & Alsdorf, 2005; Pavelsky, 2014; Pavelsky & Smith, 2009; Tarpanelli et al., 2013) or introduce RS data into hydrologic or hydraulic models (e.g., Bjerklie et al., 2005; Chandanpurkar et al., 2017; King et al., 2018; Lin et al., 2019; Neal et al., 2009; Silvestro et al., 2015; Siqueira et al., 2018; Zhang et al., 2016). These approaches (i.e., merging in situ and RS data) yield good predictive accuracy and can extend existing gauge records in space and time. In ungauged settings however, there are no gauge records to extend. Ground-based knowledge would improve RSQ accuracy in these scenarios, but in lieu of such information, these methods must

produce reasonably accurate results without relying on in situ knowledge (Gleason & Durand, 2020). In ungauged settings, standard practice is again to introduce RS data into hydrologic models (e.g., Emery et al., 2018; Sun et al., 2015) or hydraulic models (e.g., Andreadis et al., 2007; Biancamaria et al., 2011; Durand et al., 2008; Yoon et al., 2012). The most recent and sophisticated methods for assimilating RS into hydraulic models (e.g., Oubanas, Gejadze, Malaterre, Durand, et al., 2018; Oubanas, Gejadze, Malaterre, & Mercier, 2018) are highly accurate in ungauged settings but computationally burdensome for global application (Gleason & Durand, 2020).

A recent branch of RSQ has emerged with global application, SWOT, and ungauged basins in mind. This approach is termed mass conserved flow-law inversion or McFLI (Gleason et al., 2017). McFLIs assume that a river reach is mass conserved and then inversely solve for the unknown parameters in a flow law given some set of RS observations. This means that no hydrologic or hydraulic model is necessary, and discharge is exclusively estimated from RS by inverting basic geomorphic theories. McFLIs are therefore defined by their flow laws. To date, all published McFLIs have used either Manning's equation (Andreadis et al., 2020; Bjerklie et al., 2018; Durand et al., 2014; Garambois & Monnier, 2015; Hagemann et al., 2017; Sichangi et al., 2018) or at-many-stations hydraulic geometry (AMHG; Feng et al., 2019; Gleason et al., 2014; Gleason & Wang, 2015; Hagemann et al., 2017) as a flow law, where AMHG reflects relationships between at-a-station hydraulic geometry (AHG) parameters along a river's course (Brinkerhoff et al., 2019; Gleason & Smith, 2014). Regardless of the geomorphology driving McFLI, all McFLIs suffer from equifinality issues, as multiple sets of flow law variables can solve the inversion in this ill-posed estimation problem (Garambois & Monnier, 2015).

At the core of McFLI inversion is a reliance on initial guesses for parameters not observable from RS, termed "priors" in Bayesian parlance. For example, to invert Manning's equation, priors are generally needed for discharge, channel roughness, and channel cross-sectional area. These priors have previously been estimated from global hydrologic model output (Bonnema et al., 2016; Durand et al., 2016; Feng et al., 2019) and/or from external training data of geomorphic and hydraulic variables (e.g., Canova et al., 2016; Hagemann et al., 2017). Priors take the form of a probability distribution of these RS-unobservable parameters. If in situ data are available, then priors have extremely low variance, and the less certain we are about a parameter a priori, the wider the distribution. Durand et al. (2016) found that McFLIs are sensitive to their priors in a test of five McFLIs on simulated SWOT observations (as SWOT has not launched, McFLIs are tested on "SWOT-like" simulated data—section 2.3). In a similar comparison of algorithms using simulated SWOT-like data, Bonnema et al. (2016) found that AMHG inversion is particularly sensitive to its priors, and Tuozzolo, Lind, et al. (2019) found that McFLI estimation bias is sensitive to the discharge prior bias in the first test of McFLIs from airborne Ka band interferometry measurements of rivers. Finally, Andreadis et al. (2020) found that an expert classification of river planform morphology used to define channel shapes a priori yielded improved discharge prediction. This is logical—the more we know about a river, the better we can invert discharge.

These findings indicate that priors play a pivotal role in McFLI discharge accuracy, yet despite the geomorphic foundations of the McFLI paradigm, present McFLIs have used the same set of geomorphic priors for every river on Earth, regardless of differences in planform geometry, hydraulics, and river size (Bjerklie et al., 2003, 2005; Bonnema et al., 2016; Durand et al., 2016; Hagemann et al., 2017). This means that McFLIs use the same expectations for, for example, Manning's n, width/depth ratios, and AHG exponents in a braided river and a canal. No study to date has explicitly explored the sensitivity of McFLIs to the *quality* of their priors, where "quality" refers to the hydraulic and geomorphic representativeness of a prior for a given river. High-quality priors would give accurate, river-specific knowledge that closely approximates field measurements and contextualizes discharge inversion to the specific hydraulics of the river. While Andreadis et al. (2020) did briefly address prior quality, their method was only applied to the channel shape prior and, more importantly, was not scalable globally. Further, Lin et al. (2019) recently developed a global modeling framework to provide global reach-scale priors on river discharge. However, their work was limited to discharge, and globally implementable priors on channel hydraulics are still underdeveloped.

There are also troves of existing in situ data that can be mapped onto these global products to inform McFLIs. The United States Geological Survey (USGS) makes periodic field measurements of discharge and other hydraulic variables that are freely available, and these measurements are easily joined to existing

hydrographic data sets. For example, Brinkerhoff et al. (2019) joined over 730,000 USGS in situ measurements to the National Hydrography Dataset (NHD), building on earlier work (e.g., HYDRoSWOT; Bjerklie et al., 2020; Canova et al., 2016).

This proliferation of in situ measurements is welcomed but is largely useless for ungauged RSQ if we cannot map these measurements to rivers using RS and have methods capable of ingesting them as priors. Given this context, we hypothesize that McFLI performance will be improved by acknowledging geomorphic differences between rivers by assigning different priors to different rivers, building on recent global RS of rivers and decades of detailed in situ work mapped to global rivers via RS. Further, we hypothesize that this intervention alone should be sufficient to improve accuracies, and no new RS-observations or updates to McFLI algorithm logic is needed to make better predictions of discharge. We use the Bayesian-AMHG-Manning's (BAM) algorithm (Hagemann et al., 2017) as a case study for McFLIs. We provide BAM with improved prior river knowledge by (1) obtaining priors from the largest known repository of in situ data joined to hydrography (Brinkerhoff et al., 2019) and (2) constructing a river classification framework to reduce hydraulic variation to geomorphically distinct river types. We test these interventions (hereafter termed the "geoBAM" algorithm) to produce river discharge from Landsat observations on 7,858 rivers in the Mackenzie River basin (validated at 108 gauges) and from SWOT-simulated data representing 19 rivers from Durand et al. (2016). Ultimately, we provide a method for improving discharge estimation that is globally scalable using only RS observations and could theoretically be applied to any river on Earth with sufficient RS data.

2. Data

The goal of this study was to improve McFLI accuracy by improving the quality of its priors. This required creating three distinct data sets: (1) in situ measured training data for generating new prior river knowledge (section 2.1), (2) RS observations for RSQ in the Mackenzie River basin (section 2.2), and (3) modeled RS observations for RSQ using SWOT-simulated rivers (section 2.3).

2.1. Hydraulic Data Set Used for Prior River Knowledge

Our first task was to gather a comprehensive data set of measured river hydraulics to generate training data. We started with Brinkerhoff et al.'s (2019) data set. This data set merges USGS surface water measurements (NWIS; https://waterdata.usgs.gov/nwis/measurements) of channel discharge and geometry with the NHD (https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus) and filters the data to include only those rivers that have at least six stations of 20 measurements each to derive that river's AMHG. For the present study, we further filtered their data to exclude impossible measurements (i.e., $Q < 0$) and measurements identified by the USGS as "poor," yielding 372,109 unique in situ hydraulic measurements at 1,409 cross sections in 190 rivers in the continental United States (Figure S1 in the supporting information). We added to these data by calculating river and landscape geomorphic variables from the NHD for each observation in the training data. We then reduced this data set to "representative hydraulics" for each cross section such that they did not vary with river stage. These variables were calculated by taking the median and sample variance of the observed values at each cross section. This amounted to the 24 features in Table 1, which were the variables available to our ultimate training data set used to differentiate rivers. All features were continuous, save "stream order" and "USGS waterbody type," which were discrete.

This data set is built exclusively for the United States, and we acknowledge that the continental United States is not reflective of all global landscapes. Further, we have also limited our data set to those rivers where there are six or more stations to parameterize AMHG. However, this is to our knowledge the largest possible freely available fluvial geomorphology data set that covers a wide range of geographies, and these data represent a best-case scenario for our analysis. Notably, our training data are missing observations in equatorial and Arctic/subarctic regions. With the aim of improving global RSQ at the center of this study, we chose to use one of these poorly represented regions as a case study to assess the generalizability of our approach: the Mackenzie River basin in the Canadian Arctic.

2.2. RS Data for the Mackenzie River Basin

The Mackenzie River basin is representative of Arctic hydrology and is completely unrepresented in our training data. It is also one of the largest watersheds in North America, relatively unregulated, and the authors have previous experience in both field and RS work in the basin. Thus, it is a good test for

Table 1
The 24 Geomorphic Variables Used to Define River Types in This Study

Variable	Symbol	Description	Source
Median channel width	W	Observed channel width	measured—NWIS[a]
Median channel velocity	V	Observed mean channel velocity	measured—NWIS[a]
Median channel depth	D	Observed mean channel depth	measured—NWIS[a]
Slope	S	Observed slope	measured—NHD[b]
Stream order	SO	Strahler stream order	calculated—NHD[b]
Distance downstream	Dd	River kms from headwater	calculated—NHD[b]
Sinuosity	Sn	Deviation from a path of maximum downslope	calculated—Wieczorek et al. (2018)
Hydraulic residence time	HRT	Travel time spent in river reach or waterbody (at mean annual flow)	calculated—NHD[b]
USGS waterbody type	WB	River/waterbody	calculated—NHD[b]
Channel shape	r	Geometrically defined shape parameter (Dingman, 2007)	calculated—NWIS[a]
Drainage area	DA	Upstream catchment area	calculated—NHD[b]
Median Froude number	Fr	Measure of hydraulic flow regime in open-channel flow (median of at-a-station values)	calculated—NWIS[a]
Median shear stress	τ	Force of moving water against channel bed (median of at-a-station values)	calculated—NWIS[a] + NHD[b]
Median unit power	Ω	Energy dissipation against riverbanks, normalized by channel width (median of at-a-station values)	calculated—NWIS[a] + NHD[b]
Maximum grain size entrained	D_e	Largest bed material entrained and transported (Henderson, 1966) (median of at-a-station values)	calculated—NWIS[a] + NHD[b]
Median Manning's n	n	Roughness term for Manning's equation (median of at-a-station values)	calculated—NWIS[a] + NHD[b]
Variance of cross-sectional channel width	Var(W)	Observed channel width (variance of at-a-station values)	measured—NWIS[a]
Variance of cross-sectional channel velocity	Var(V)	Observed mean channel velocity (variance of at-a-station values)	measured—NWIS[a]
Variance of cross-sectional channel depth	Var(D)	Observed mean channel depth (variance of at-a-station values)	measured—NWIS[a]
Variance of cross-sectional Froude number	Var (Fr)	Measure of hydraulic flow regime in open-channel flow (variance of at-a-station values)	calculated—NWIS[a]
Variance of cross-sectional shear stress	Var(τ)	Force of moving water against channel bed (variance of at-a-station values)	calculated—NWIS[a] + NHD[b]
Variance of cross-sectional unit power	Var(Ω)	Energy dissipation against riverbanks, normalized by channel width (variance of at-a-station values)	calculated—NWIS[a] + NHD[b]
Variance of cross-sectional maximum grain size entrained	Var (D_e)	Largest bed material entrained and transported (Henderson, 1966) (variance of at-a-station values)	calculated—NWIS[a] + NHD[b]
Variance of cross-sectional Manning's n	Var(n)	Roughness term for Manning's equation (variance of at-a-station values)	calculated—NWIS[a] + NHD[b]

Note. Note that we require a stage invariant geomorphic classification. Therefore, stage-varying variables are defined using "representative hydraulics," that is, the median and variance of these variables at-a-station.
[a]USGS (https://waterdata.usgs.gov/nwis/measurement). [b]EPA (https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus).

applying our new prior river knowledge at the near-continental scale. We use the Mackenzie as a test basin using real satellite data (as opposed to modeled SWOT data, section 2.3). To do so, we needed RS observations of river widths to drive the BAM/geoBAM algorithms. We extracted multitemporal widths for 220,069 cross sections in the basin following Feng et al. (2019, Text S1). In short, we (1) used the MERIT hydro river network product (Lin et al., 2019), which was vectorized from flow accumulation data by Yamazaki et al. (2019), to define river centerlines and generate width measurement stations at varying intervals along the centerlines based on river widths (Yamazaki et al., 2019, Table S1); (2) we constructed orthogonal cross sections for every station following Yang et al. (2019); and (3) these orthogonal lines were used as inputs to RivWidthCloud, an automated algorithm for river width extraction using the Google Earth Engine (Yang et al., 2019). After filtering for clear-sky images, RivWidthCloud classifies a pixel as water using an algorithm detailed in Yang et al. (2019) and Text S1. The classified water mask was intersected with the orthogonal lines to estimate wetted width at each cross section. We measured widths this way at 220,069 cross sections for all rivers with a mean width >120 m from 7,858 Landsat-visible reaches in the Mackenzie River Basin using Landsat imagery from 1984–2013. Finally,

these cross-sectional river widths were binned by river reach using the MERIT Hydro river network (Lin et al., 2019, median length 6.8 km).

For validation, we collected in situ daily discharge data for the Mackenzie River basin for all time periods that gauges were operational from 1984–2013 from the Water Survey of Canada (WSC; https://wateroffice.ec.gc.ca/mainmenu/real_time_data_index_e.html). These gauge data are analogous to USGS gauging data provided by the NWIS. WSC gauge stations were linked to our hydrography based on their geospatial locations (distance within 500 m of a centerline) and drainage areas (i.e., difference within ±10% of a hydrography reach), resulting in 327 validation gauges. Of these 327 gauges, 108 coincide with Landsat-visible reaches and were ultimately used for validation. Together, these RS widths and completely independent gauge validation data provide a platform to test McFLI improvements in an actual application setting, with all inherent RS errors present in the data. We also use a "perfect" scenario to test our improvements without the influence of real-world errors, described next.

2.3. SWOT-Simulated Rivers

Once SWOT launches, Manning's-based McFLIs can be run globally from SWOT's novel simultaneous observations of river width and height. However, existing SWOT-like data are limited to an airborne Ka band interferometric synthetic aperture radar (InSAR), which is currently available for less than five rivers globally with only a handful of observations each (Tuozzolo, Lind, et al., 2019), or painstaking data fusion of altimetry and imagery (Bjerklie et al., 2018). Therefore, authors typically use simulated SWOT data to test Manning's-based McFLIs. A SWOT simulator has been built by the Jet Propulsion Laboratory (JPL) to simulate the errors that are expected to come with actual SWOT data (namely, radar layover errors and random noise) and has been used to benchmark McFLIs before (e.g., Oubanas, Gejadze, Malaterre, Durand, et al., 2018). However, for this study we are interested purely in algorithm performance and so seek a test scenario that assumes perfect measurement conditions with no introduced errors. Satisfying data to achieve this are simple reach-averaged hydraulic model outputs with water surface heights and river widths labeled as "RS observations" (Durand et al., 2016; Hagemann et al., 2017). For this test, we used 19 test rivers from Frasson, Durand, et al. (2019) which were developed for benchmarking McFLIs by Durand et al. (2016) and outlined in Table S2. These rivers cover the United States, Canada, Great Britain, France, Bangladesh, and Italy. Median discharge ranged from 62–14,199 m^3/s and "observation" windows ranged from 22–365 days over 3–16 reaches. All 19 models solve either 1-D or 2-D St. Vernant equations using field measurements to construct channel bathymetry (see Durand et al., 2016, for a more through overview). Twelve are HEC-RAS models (1-D), one is a MASCARET model (1-D), one is a BreZo model (2-D), one is a ProSe model (1-D), two are LISFLOOD models (1-D and 2-D), and two are H2D2 models (2-D). Simulated rivers mimic perfect measurement conditions and represent better than the best-case scenarios of what SWOT will provide to hydrologists. Specifically, these river models provide "observations" of river width and water surface elevation, which are in turn used to calculate water surface slope for use in a Manning's-based McFLI.

3. Methods

With these data in hand, our methods are as follows (see Figure 1 for a flowchart of our methods):

1. Update BAM to reflect the latest geomorphic understanding of AMHG following Brinkerhoff et al. (2019) and ingest geomorphic priors (section 3.1).

2. Generate new prior river geomorphology knowledge from our field hydraulics data set (section 2.1) and categorize this information following different approaches: an "unsupervised" classification and an "expert" classification (section 3.2).

3. Assign classified prior river knowledge to river reaches using an RS mapping scheme (section 3.3).

4. Compare geoBAM with Hagemann et al.'s (2017) original BAM algorithm for the Landsat and SWOT cases (section 3.4).

3.1. geoBAM and AMHG Physics

We performed McFLI RSQ from both the BAM and geoBAM algorithms. geoBAM does not update the logic or computational aspects of BAM, and therefore, it is important to briefly outline how the original BAM algorithm works and detail what makes BAM and geoBAM different. Readers are referred to Hagemann

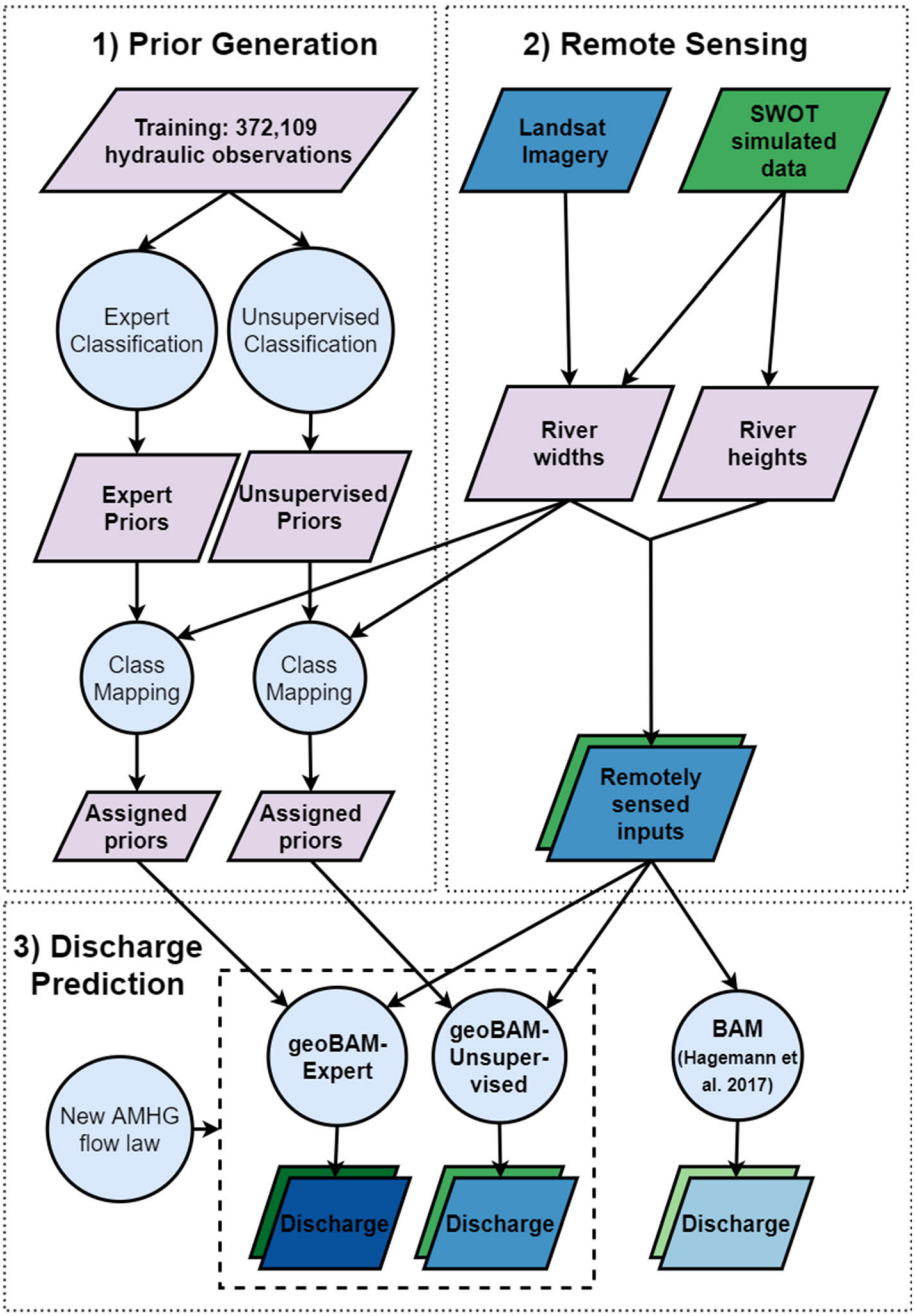

Figure 1. To develop geoBAM, we (1) generate new prior hydraulic data and classify this information, (2) generate remotely sensed inputs, and (3) estimate discharge with BAM/geoBAM. In the final row of discharge products, green outputs correspond to the Mackenzie River basin test (108 samples), and blue outputs correspond to the SWOT-simulated rivers (19 samples). These colors also align with boxplot colors in our results (Figures 3–6 and 8). Note that the "remotely sensed inputs" are different for the SWOT case (widths and heights) versus Landsat case (widths only). See Figure S2 for a flowchart explaining original BAM workflow (Hagemann et al., 2017), and section 3.1 for information on the updated flow law.

et al. (2017) for full details, and Figure S2 gives a flowchart detailing the workflow internal to BAM. In brief, BAM is a McFLI that probabilistically estimates river discharge via Bayesian inference and a Hamiltonian Monte Carlo sampling scheme given RS river widths and/or heights and priors on all non-RS parameters as inputs. BAM yields posterior probability distributions for all non-RS parameters in the flow law (including discharge) as a Bayesian technique and is a formal and intuitive way of achieving this (Hagemann et al., 2017).

BAM relies on two flow laws: Manning's equation and AMHG. Brinkerhoff et al. (2019) recently brought theoretical closure to the previously empirical AMHG, and consequently, AMHG can now be written as a function of known hydraulic variables. Brinkerhoff et al. (2019) found that AMHG is a direct mathematical consequence of a river longitudinal profile's strength of fit in any slope-roughness model. Further, they showed that AMHG's empirical parameters (W_c and Q_{cw} in Equations 1–3, see Brinkerhoff et al., 2019, for definition), coexist with Dingman's (2007) hydraulic geometry model and thus are a valid hydraulic tuple for the given river cross section. Overall, their analysis yielded a new AMHG expression (Equation 1) defined by p (the generalized velocity-depth relation exponent), K (the generalized roughness coefficient), W_b (bankfull width), D_b (bankfull depth), r (a channel shape term defined by Dingman, 2007: $r = 1$ is a triangular channel, $r = 2$ a parabola, and $r =$ infinity a box), and channel slope S. We take this premise and derive a novel flow law (Equation 3) by first substituting Equation 1 into the original AMHG flow law (Equation 2; Gleason & Wang, 2015) and substituting Manning's constants for the generalized terms. In Equation 3, Q is river discharge, b is the width AHG exponent, and W is channel width for cross section i and time step t. n was defined previously as Manning's roughness coefficient. Note that K is equal to $1/n$ when adhering to Manning's relation and consult Dingman (2007) for an explanation of his generalizations of Chezy's and Manning's expressions for hydraulic geometry relations that we follow here.

$$W_c^{1+r+rp}Q_{cw}^{-1} = \left(KS^{\frac{1}{2}}\right)^{-1} W_b^{r+rp} D_b^{-(1+p)} \left(\frac{r}{r+1}\right)^{-(1+p)} \tag{1}$$

$$Q_t = \left(\frac{W_{it}}{W_c}\right)^{\frac{1}{b_i}} Q_{cw} \tag{2}$$

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

It must be stressed that values for p and K (in Equation 1) have been shown to vary widely and not always adhere to Manning's (or Chezy's) constants due to different river morphologies (e.g., Bjerklie et al., 2005; Dingman, 2007; Dingman & Afshari, 2018; Dingman & Sharma, 1997; Ferguson, 2010; Knighton, 1975). With that said, this new formulation of the AMHG flow law is flexible and in theory allows for future generalized implementations. In summary, the AMHG flow law (Equation 3) is now defined explicitly by fluvial geomorphology and when introduced to BAM, requires priors on river channel hydraulics just like Manning's equation does (specifically, Q, W_b, D_b, r, n, and W_c in Equation 3).

Hagemann et al.'s (2017) BAM also uses a channel roughness prior (n) that is constant in space and time across all sections of a river, which is known to be both physically inaccurate in many scenarios and poorly reflective of the variation in roughness experienced in space and time (Ferguson, 2010; Tuozzolo, Langhorst, et al., 2019). Thus, geoBAM should update this n prior to vary in space and time. We found this space-and-time-varying implementation significantly slowed geoBAM computations, and further, there was almost no change in performance between space-varying and space-and-time-varying, but both improved on the default prior in BAM. We deemed the lack of improvement in performance when instituting space-and-time varying n not worth the significant computational burden it imposed. Thus, we implemented only the space-varying n prior into the version of geoBAM used in this study. Figure S3 shows the difference between these results.

3.2. River Classification

Recall our hypothesis that defining priors for specific river types will improve McFLI performance. To test this, we constrained prior river knowledge using expert and unsupervised statistical classifications to extract geomorphically distinct river types.

Statistical learning is generally binned into unsupervised and supervised approaches (Hastie et al., 2009; James et al., 2013). Both use a suite of variables extracted from a training data set to define a feature space and then identify patterns in that space. Unsupervised learning identifies these patterns and then clusters

observations solely from the training data given user guidance only on algorithm parameters, while supervised learning uses user-defined predictors to model "target variables" that are known a priori. The familiar concepts of regression (if the target variable is continuous), or classification (if it is discrete) are forms of supervised learning. In this study, we lack a priori river types in the training data to take a supervised approach to learning, and so instead, we implement both unsupervised learning and an expert classification framework that forces specific knowledge on river geomorphology to guide statistical methods of classification.

3.2.1. Unsupervised

As a representative unsupervised clustering approach, we used the "density-based spatial clustering of applications with noise" (DBSCAN: Ester et al., 1996) algorithm. DBSCAN is a density-based clustering algorithm that groups observations in the multidimensional feature space using proximity. Distance between points is determined using Euclidean distance. Unlike simpler unsupervised clustering algorithms, DBSCAN does not assume that all clusters have a convex shape in the feature space and instead uses density to group observations. This means clusters can be arbitrarily shaped or completely surround other clusters. This also permits DBSCAN to identify "noise" points which are outside of the dense areas of the feature space, differing in practice from other simple unsupervised learning methods (e.g., *k*-means clustering will assign every observation to a cluster). The user must provide a minimum number of points for a cluster and a maximum cluster radius, and DBSCAN automatically determines the number of clusters, unlike simpler unsupervised algorithms. We used a standard, "elbow"-based approach to choose a maximum cluster radius of 0.5 (Text S2 and Figure S4) and a minimum cluster size of five river cross sections as a balance of interpretability, within-cluster variation, and number of resulting clusters. We ran DBSCAN on nondimensional forms of the 24 geomorphic variables in Table 1, yielding seven clusters in approximately 95% of the cross sections, with the remaining 5% classified as "noise."

3.2.2. Expert

We also developed a bespoke expert classification framework for extracting river types, built specifically such that river width is a predictor of these types. By using principal component analysis (PCA) as a guiding tool, the approach follows similar methods used to extract global hydroclimatic river types (Dallaire et al., 2019) and hydrologic flow regimes (Olden et al., 2012). Here, a PCA was used to dimensionally reduce our data set and create multivariate, nondimensional principal components (PCs) responsible for some amount of geomorphic variation across the feature space (and ultimately the United States). We ran a PCA on nondimensional forms of the 24 geomorphic variables in Table 1 at all cross sections and selected the three most influential PCs, cumulatively responsible for 54% of the variance in the feature space, as a subjective balance between variance explained and interpretability. Selecting additional PCs would explain more variance but limit our ability to map rivers from RS data and interpret the PCs. In order of most variance explained, these PCs qualitatively represented (1) sediment transport, (2) river size, and (3) variation in velocity/Froude number (Table S3).

There are two parts to the expert classification. In the first, we used PC scores to assist in classifying cross sections. PC scores were calculated for every cross section in the training data to locate each cross section in the PCA subspace (or more formally, linear combinations of the normalized feature values multiplied by their "loadings"). Each cross section has 24 PC scores associated with it. Because the loading vectors, and thus the PC scores, are all normalized to the same global values (Hastie et al., 2009; James et al., 2013), we simply summed the three PC scores for each cross section corresponding to PCs 1–3. This provides a single value per cross section, where cross sections with similar aggregate PC values have similar geomorphology (i.e., have similar values for the three most influential PCs and therefore similar geomorphology in the PCA subspace). We then use this metric to classify all cross sections into river types by segmenting rivers into 15 classes using quantiles of the aggregate PC metric as class thresholds. Since this is an "expert" system, we subjectively chose 15 classes to explicitly maintain river width as a predictor of river types while allowing for diversity of river form in our class system. Adding more classes obscured differences between classes and broke down the predictive relationship between river type and river width (section 3.3). This relationship is essential, as width is easily obtained from RS and thus allows this expert classification to be mapped to all global rivers.

The second part of the expert classification framework was designed to differentiate "big" rivers and "highly width-variable" rivers from the 15 classes. For some very large rivers (e.g., the Mississippi or St. Lawrence

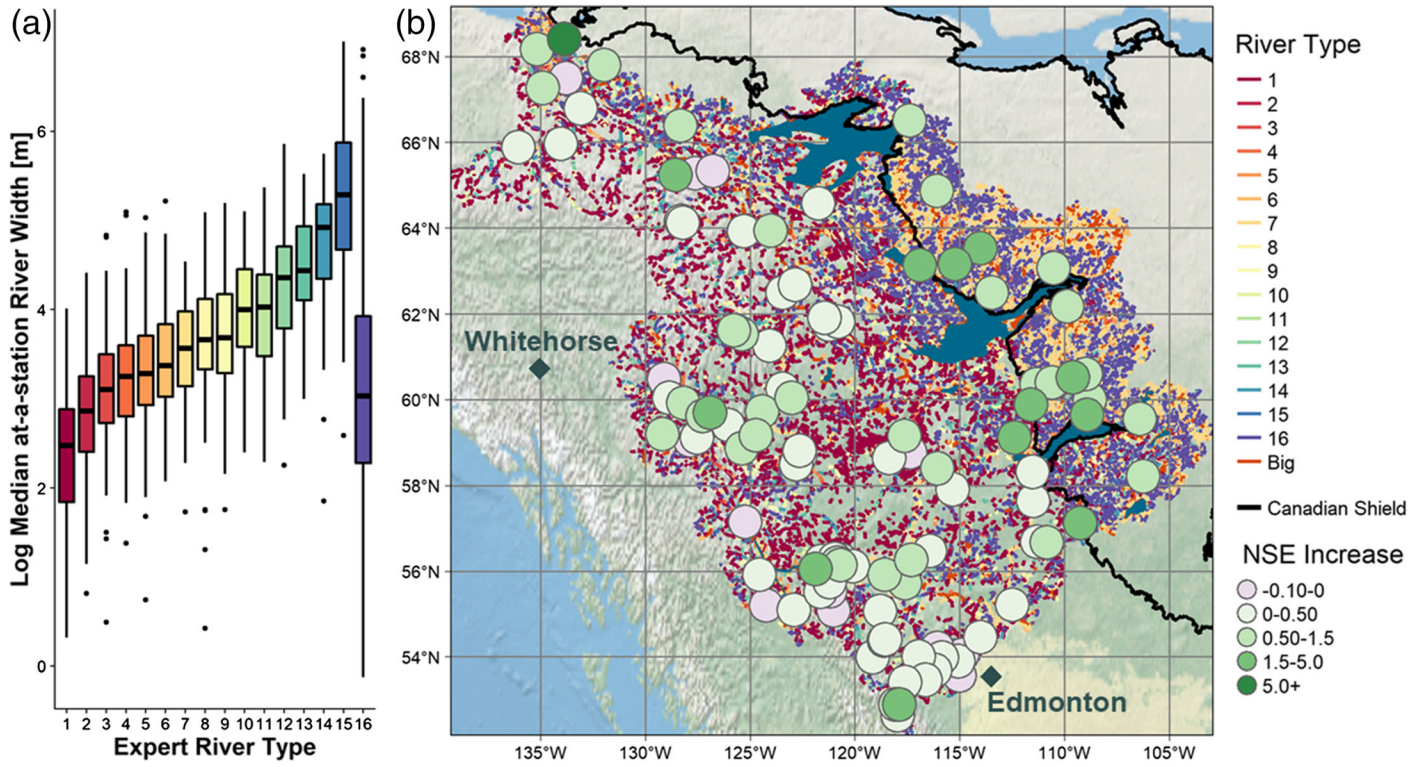

Figure 2. (a) Median at-a-station river widths by expert river type for the training data. Median river width is a strong predictor of most "expert"-classified rivers; however, significant variation of widths within each river type is still observed. (b) Expert classification for over 7,500 reaches and discharge validation at 108 gauges. Increase in NSE from BAM to geoBAM-expert is given as colored circles at the gauge locations. The Canadian Shield (black outline) and major lakes (dark blue) are overlaid upon the classified hydrography (Lin et al., 2019; Yamazaki et al., 2019) to show clear class differences in the shield versus the rest of the basin. Basemap made with Natural Earth and cities added to show the size of the basin.

rivers), the training data had very few cross sections in rivers of similar size, and thus, hydraulics are ill defined. We defined "big" rivers as those with a mean width greater than 665 m. This threshold was subjectively chosen using the distribution of river widths for the widest river type (Type 15). We then defined "highly width-variable" rivers (Type 16 in Figure 2a) as those with a channel shape parameter $r < 1$ (like Andreadis et al.,'s 2020, definition for braided rivers), which guarantees significant variability in river width for both single channel and multichannel rivers (Dingman, 2007).

3.3. Mapping River Type From RS

The goal of this study is to improve discharge prediction at the global scale. Thus, river types must be assignable to rivers using only RS observations or other readily available geospatial data sets. However, our classifications above are defined using in situ measurements that are not remotely sensible. This presents a unique supervised classification challenge necessitating width and water surface height (i.e., current synoptically available fluvial parameters) as the sole predictors of the river types. Here, we use "river type mapping" to refer to a final supervised classification where the target variables are our river types and the predictors are river width and/or slope. This is a second classification step applied to both the unsupervised and expert classifications already performed on the in situ data.

For the DBSCAN framework, neither width nor slope were strong predictors of river types. Therefore, we needed a way to map RS observations at the scale of the Mackenzie basin (220,069 cross sections) to the DBSCAN-generated classes (which use no RS data). To do so, we turned to basic supervised statistical learning to assign river types, which has seen some success at the regional scale (Guillon et al., 2020). Using a classic validation-set approach to model training, we trained a multiclass logistic regression classifier on 80% of the training data using the median of cross-sectional widths as the sole predictor. Logistic regression models yield the probability of an observation being assigned some class, conditional on the observations in the data

set (Hastie et al., 2009). While generally used to predict binary classes, multiclass logistic regression is possible and implemented here using a "one-versus-rest" classifier (Bishop, 2006). When tested on the remaining 20% of the data, the classification accuracy was 87%. Note that this classifier can classify a river as "unclassified," that is, a noisy point outside of any clusters in the feature space (section 3.2.1). Thus, we can reasonably reproduce our classification, which carries rich in situ prior information, solely from RS.

As previously stressed, the expert framework was developed explicitly such that river width was a strong predictor of river type as produced by the PCA. This is confirmed in section 4.1, where the median widths for each river type exhibit a strong relationship with the river types. We acknowledge that within the river types there is significant variation in river width. We ultimately assigned river class by selecting the class which had a median characteristic width closest to the MERIT hydro width. We mapped the "highly width-variable" type to rivers using a minimum standard deviation of at-a-station width >1.87 m as the threshold. Recall that "highly width-variable" rivers were previously defined as having an r less than 1 in the training data (section 3.2.2). We iteratively tested thresholds on the SWOT rivers and chose the minimum value that preserved a large cluster of low variability (Figure S5).

3.4. Discharge Estimation

Ultimately, we produced two forms of geoBAM: geoBAM-Unsupervised and geoBAM-Expert, that compare how pure statistical learning (geoBAM-Unsupervised) compares with imposed prior knowledge of rivers (geoBAM-Expert). The only difference between these methods is their priors, and the RS inputs and discharge physics are identical between them.

Recall that BAM/geoBAM can use two flow laws: The user chooses to invert Manning's equation, AMHG, or a combination of both as its flow law. The user also chooses whether to run the algorithms at every cross section or to use reach-averaged observations in line with how SWOT will observe rivers. For this study, we ran AMHG in the Mackenzie at the cross-section scale and ran a "switch" flow law on the reach-averaged SWOT-simulated data. This "switch" always inverts Manning's equation and inverts Manning's and AMHG when AMHG it is deemed suitably strong (Hagemann et al., 2017). We defined "suitably strong" AMHG differently than Hagemann et al. (2017), following Brinkerhoff et al.'s (2019) finding that when observed slopes strongly fit a river-wide slope model, AMHG is strong. Here, we use the regime theory model defined by Henderson (1966) and reprinted as Equation 4 where Q_b is bankfull discharge. Degree of fit is defined by the coefficient of determination (r^2—note that r^2 is standard notation and not the square of the channel shape parameter r), and strong fit was defined as an $r^2 > 0.90$.

$$S = 0.44 D_e^{1.15} Q_b^{0.46}. \tag{4}$$

To run geoBAM for Manning's and/or AMHG, the user provides water surface heights and/or widths from RS observations, as well as priors on 34 parameters. These parameters define distributions for six hydraulic variables (as well as discharge, flow law errors, and AMHG's W_c term) and are formalized within geoBAM as truncated, lognormal distributions where $\ln(X) \sim N(\mu,\sigma^2)$ for $\lambda < X < \varepsilon$, using mean ($\mu$), standard deviation ($\sigma$), and upper ($\varepsilon$) and lower bounds ($\lambda$) as parameters. The six hydraulic variables, following the geomorphic updates made in section 3.2, are median cross-sectional area (A_0), bankfull width (W_b), bankfull depth (D_b), Manning's roughness term (n), a channel shape parameter (r), and an AHG exponent (b). For our tests (see Figure 1 for a flowchart), we used geomorphic river types to redefine the prior river knowledge on the six hydraulic terms in Figure 7 using the river mapping in section 3.4 to assign a river type to each BAM reach.

Prior parameters were extracted directly from the training data's distribution of each hydraulic variable once a river type had been assigned. W_b and D_b were calculated at each cross section using a return period of 2 years, acknowledging that Q_b has been associated with a range of return periods given local geomorphology (Petit & Pauquet, 1997; Williams, 1978). However, a 1.5- to 2-year return period is a standard statistical definition for bankfull flow and was used here to align with existing literature. b was defined empirically by fitting, at each cross section, Equation 5. r was defined empirically following Dingman (2007) as f/b, where f is the exponent calculated when fitting, at each cross section, Equation 6. Manning's n was calculated using observed flow velocities and depths and following Equation 7.

$$w = aQ^b \tag{5}$$

$$d = cQ^f \tag{6}$$

$$n = \frac{d^{\frac{2}{3}} S^{\frac{1}{2}}}{v} \tag{7}$$

Recall that geoBAM-Expert includes a "big" river type and geoBAM-Unsupervised can classify a river as "unclassified" (section 3.4). For "big" and "unclassified" rivers with poor representation in our training data, we followed the method outlined in Hagemann et al. (2017)—by using a global prior set that represents all rivers in a single summary from our training data set (Figure S6). For the "AMHG switch," Q_b and D_e were also predicted using this global method. For the expert framework, we further accounted for "big" rivers by setting the bounds on some priors to larger values than produced by the global prior (Table S4). The prior estimate on discharge was defined differently for our two tests: for the Mackenzie River basin, we used a stream gauge when available and otherwise used the mean daily discharge estimate from GRADES (Lin et al., 2019). For the SWOT-simulated rivers, output from a water-balance model (Wisser et al., 2010) was used in line with Durand et al. (2016).

3.5. Validation Metrics

Error metrics (Table S5) to quantify RSQ accuracy followed Hagemann et al. (2017). rBIAS and RRMSE define the range and central tendency of prediction errors, respectively. NRMSE is a normalized variant of RRMSE to account for RRMSE's high sensitivity to errors in low flow estimation, and Nash-Sutcliffe efficiency (NSE) represents the amount of variance in the observed data that the model explains. An NSE greater than 0 indicates that our model estimates better than guessing mean flow every time. Scores are reported at the river scale for the SWOT-simulated rivers and at the reach scale for the Mackenzie River basin.

In order to independently validate our interventions on the American SWOT-simulated rivers, we manually removed all training data that corresponds to these SWOT-simulated river reaches before building the classifications and training the logistic classifier for geoBAM-Unsupervised. However, we use the entire training data set for the Mackenzie basin as these reaches are all independent of the training data. Similarly, the classification results presented in section 4.3 use the entire training data set.

4. Results

We compare geoBAM validation scores against BAM scores across several tests. We first present the river classification and discharge prediction for the Mackenzie River basin (section 4.1), followed by discharge prediction accuracy for the SWOT-simulated rivers (section 4.2). Then, we orient these results by exploring the prior distributions (section 4.3) and additional RSQ error metrics (section 4.4).

4.1. Mackenzie River Basin

First, we visually confirm that river width was a strong predictor of Expert river type (Figure 2a—recall section 3.2.2). Then, we map over 7,500 rivers in the Mackenzie River basin classified via the expert framework, along with the increase in NSE from BAM to geoBAM-Expert for all 108 validation reaches (Figure 2b). Rivers mostly fall into two regimes composed of Type 1 and 2 rivers in the western half of the watershed, and Types 16 and 7 in the eastern. These two regimes align nearly perfectly with the boundary of the Canadian Shield, a particularly old, hard, exposed igneous rock that underlies thin soil, large lakes, and wetlands in the eastern portion of the basin (e.g., Spence & Woo, 2008). Meanwhile, the western portion drains the Rocky Mountains—a fundamentally different landscape. Thus, the Expert classification qualitatively reflects observed geomorphology patterns in the landscape of the Mackenzie River basin through its different river types.

The clear manifestation of the Canadian Shield in our classification influences RSQ accuracy as well. All validation reaches on the Shield showed an NSE improvement of at least 0.50, while any reaches with less improvement, or even degradation, in NSE are not located on the Shield. This suggests that geoBAM is more able to hydraulically represent the types of rivers present in "Shield-like" landscapes than BAM. The only

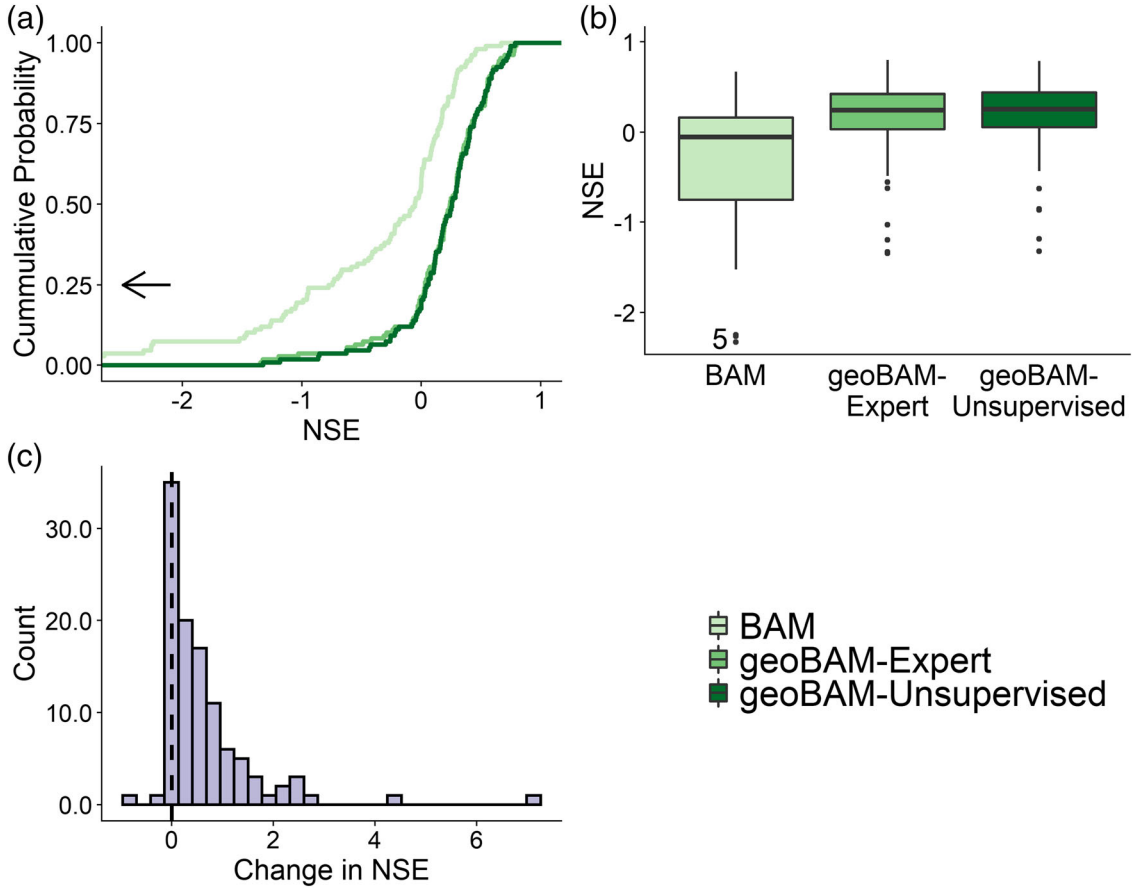

Figure 3. NSE improvement for 108 validation reaches in the Mackenzie River basin: (a) empirical cumulative density functions (CDFs) of the NSE scores for every reach, (b) boxplots of the same results, and (c) histogram of the change in NSE from BAM to geoBAM-Unsupervised. Axes are truncated for visualization's sake. We have provided arrows and number of rivers not plotted when necessary.

river reach with an NSE increase >5.0 is near the outlet of the basin. Finally, there is no clear relationship between NSE improvement and stream order (Figure S7), suggesting that improvement is not a function of river size (although slight decrease in the NSE improvement is seen from Order 1 to Order 5 streams, Figure S7).

Next, we explore accuracy of predicted discharge using geoBAM. geoBAM predicted discharges are much improved over BAM in the Mackenzie, regardless of the Expert or Unsupervised variant (Figure 3). Despite no new RS observations, we have yielded an increase in the median NSE of 0.31 (Figure 3c) across the 108 validation gauges when using geoBAM-Unsupervised. Most reaches (78/108) improved NSE by at least 0.1 (Figure 3c). Almost the entire improvement distribution is greater than 0 when we take the difference in NSE between geoBAM-Unsupervised and BAM, and improvement is often greater than 1.0 (Figure 3c). The largest improvement occurs in reaches that were poorly estimated (approximately NSE < 0) by BAM, and this area of largest improvement is the only region (Figure 3a) with notable differences between unsupervised and expert variants of geoBAM, with geoBAM-Unsupervised performing slightly better. BAM had positive NSE for less than half of the reaches (42%). geoBAM, however, has a positive NSE in most reaches (80% Unsupervised and 79% Expert). Median NSE across all reaches improved from −0.05 to 0.24 with geoBAM-Expert and to 0.26 with geoBAM-Unsupervised. Finally, the entire distribution of NSE scores is higher and more consistent with geoBAM: The middle 50% of scores increased and the interquartile range (IQR) narrowed substantially (from 0.91 to 0.38 for geoBAM-Unsupervised and 0.39 geoBAM-Expert). The minimum geoBAM NSE is also −1.35, which is well within the whiskers for BAM (compared to BAM's minimum NSE of −6.63). In sum, geoBAM's two primary interventions (a geomorphic update to AMHG and differential prior distributions on geomorphic variables—sections 3.2–3.4)

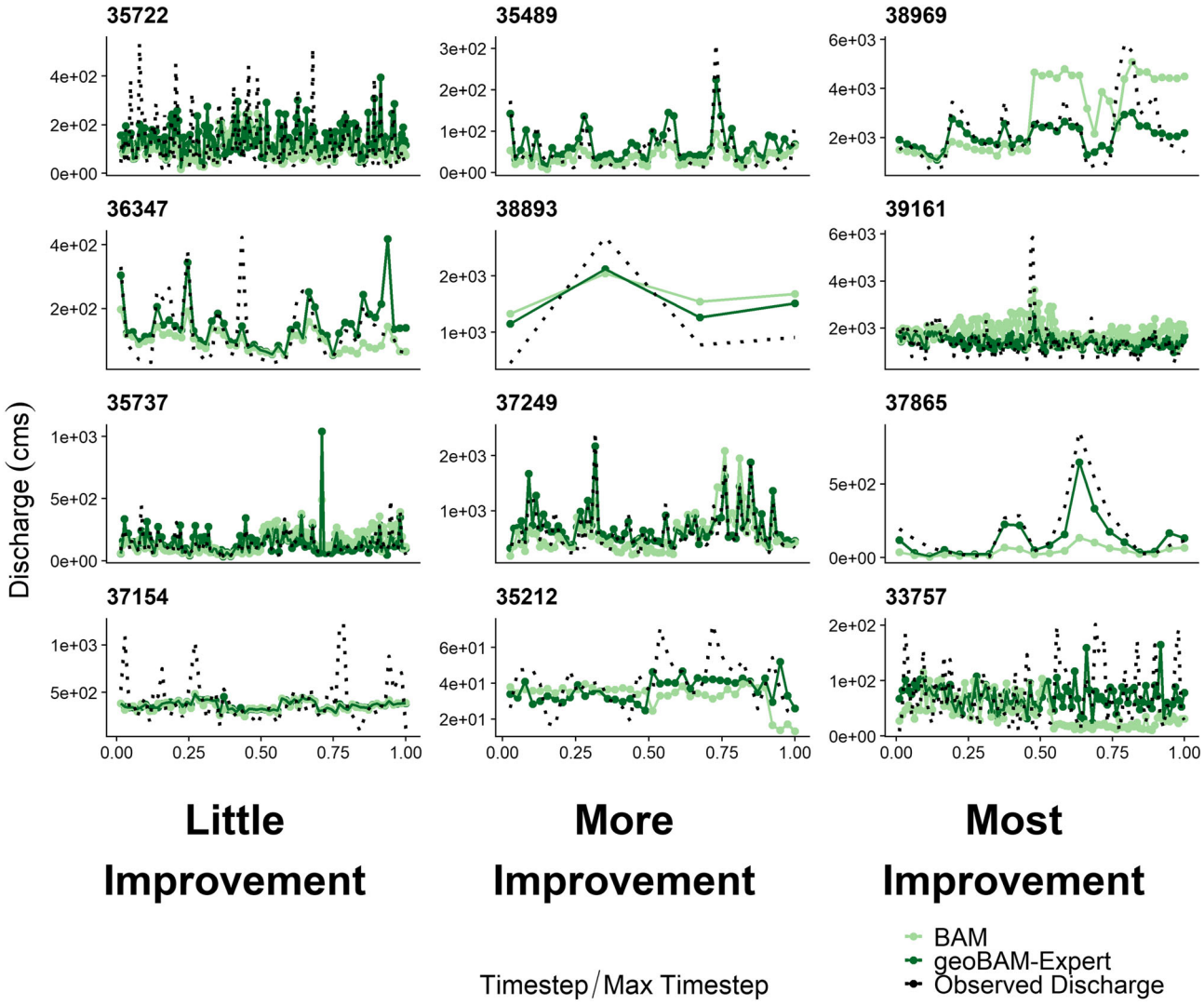

Figure 4. Twelve hydrographs from the Mackenzie River basin: observed discharge (dashed black) is plotted alongside BAM (light green) and geoBAM-Expert (dark green). Each column features four reaches randomly sampled from bins of 33% ("tertiles," here called "little," "more," and "most" improvement) for NSE improvement from BAM to geoBAM-Expert.

substantially outperform BAM and its global-scope priors (section 1 and Figure S6) in both the magnitude and consistency of NSE scores over all 108 validation reaches.

While nearly all validation reaches showed improvement in NSE with respect to BAM (Figure 3c), there was a range in the magnitude of this improvement, visualized by the hydrographs of predicted discharge (Figure 4). To visualize, we binned all 108 validation reaches into three groups using the tertiles of the distribution of improvement in NSE from BAM to geoBAM-Expert. We termed these bins "little improvement," "more improvement," and "most improvement" and then randomly selected four reaches from each of these bins for Figure 4. For the reaches with the most improvement (38969, 39161, 37865, and 33757) BAM could not replicate the observed dynamics of streamflow, while geoBAM-Expert was largely able to reproduce the peak events and baseflow periods present in the observed hydrographs. For Reach 37865, both BAM variants modeled baseflow accurately, but only geoBAM-Expert successfully modeled peak flow events. For Reaches 38969 and 33757, BAM's modeled discharge was marked by periods of no discernable streamflow dynamics and sudden jumps to other flow values. geoBAM-Expert was able to more accurately reproduce the differences between baseflow and peak flow events in these reaches. For the "more improvement" reaches, geoBAM-Expert hydrographs also visually resemble the observed hydrographs, though the magnitude of improvement was smaller than for the "most improvement" reaches. Sometimes (Reaches 37249 and

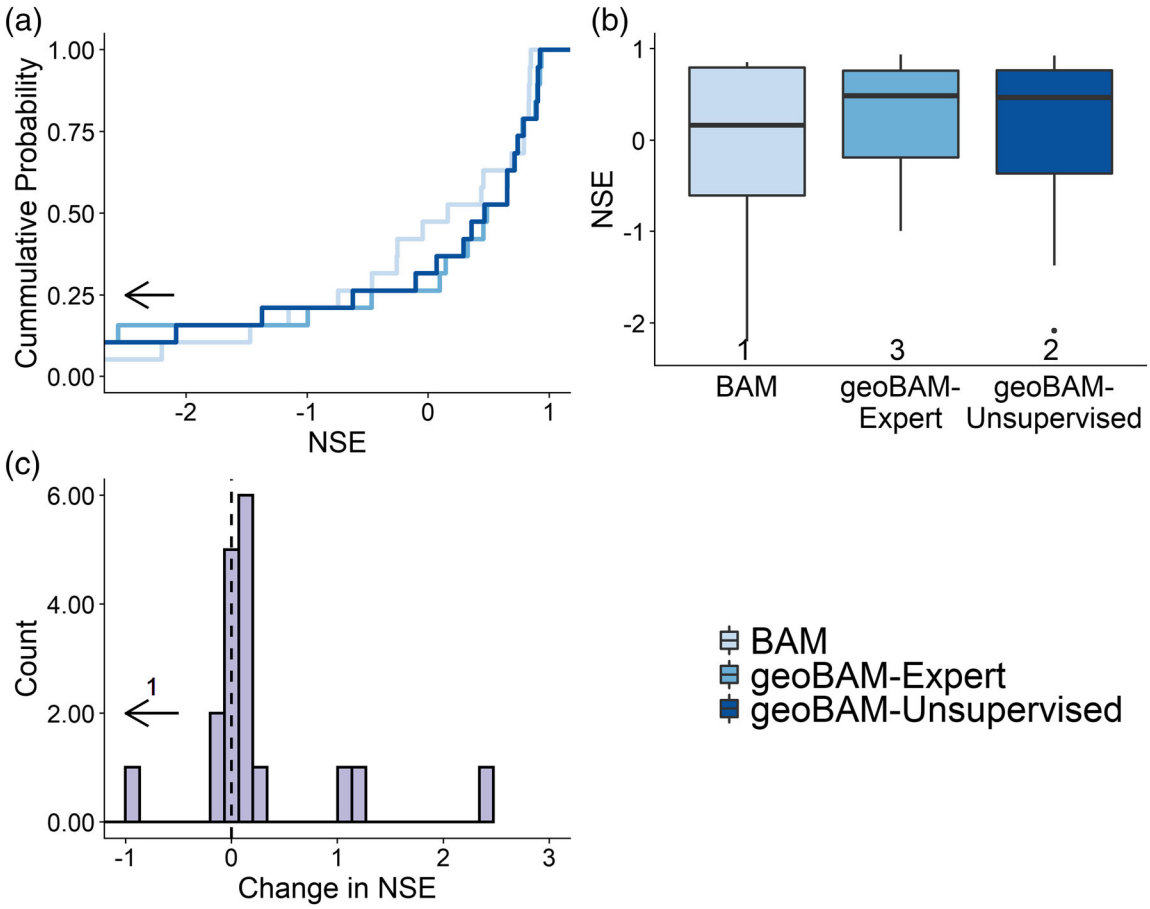

Figure 5. NSE improvement for 19 SWOT-simulated rivers: (a) empirical cumulative density functions (CDFs) of the NSE scores across rivers, (b) boxplots of the same results, (c) histogram of the change in NSE using geoBAM-Unsupervised. Axes are truncated for visualization's sake. We have provided arrows and number of rivers not plotted when necessary.

35489), BAM already modeled streamflow quite well. Here, geoBAM-Expert largely produced the same hydrograph as BAM, but with better accuracy in simulating the magnitude of the high flow events that BAM was missing. While some of the "little improvement" reaches were poorly modeled by BAM and continue to be poorly modeled by geoBAM-Expert (Reach 37154), some were modeled well by both BAM variants (Reach 35737). Finally, there is no discernable relationship between NSE improvement and the size of the river, as a full spectrum of observed discharge, ranging from tens of m^3/s to thousands of m^3/s, is evident across all improvement cases.

4.2. SWOT-Simulated Rivers

There is continued substantial improvement in NSE for the 19 SWOT-simulated rivers (Figure 5), though there is less overall improvement than in the Mackenzie River test (Figure 3). geoBAM-Unsupervised yielded an increase in median NSE for all rivers of 0.27 (Figure 5c), as well as consistently better performance: The IQR shrinks from 1.40 to 0.94 for geoBAM-Expert and 1.13 for geoBAM-Unsupervised (Figure 5b). Most rivers (8/19) improved NSE by at least 0.1. While this improvement is lower than in the Mackenzie River test, the median NSE across all 19 rivers significantly improved from 0.16 to 0.48 with geoBAM-Expert (Figure 5b). Most improvement in the SWOT rivers came from those with a middling NSE (approximately −1 to 0.80—Figure 5a). Those rivers with very high NSE scores did not change at all with our interventions, and some of the very poorly performing rivers got worse. geoBAM-Expert slightly outperformed geoBAM-Unsupervised with respect to median NSE and the IQR of NSE scores; however, geoBAM-Expert produced more poorly performing outliers than geoBAM-Unsupervised or BAM

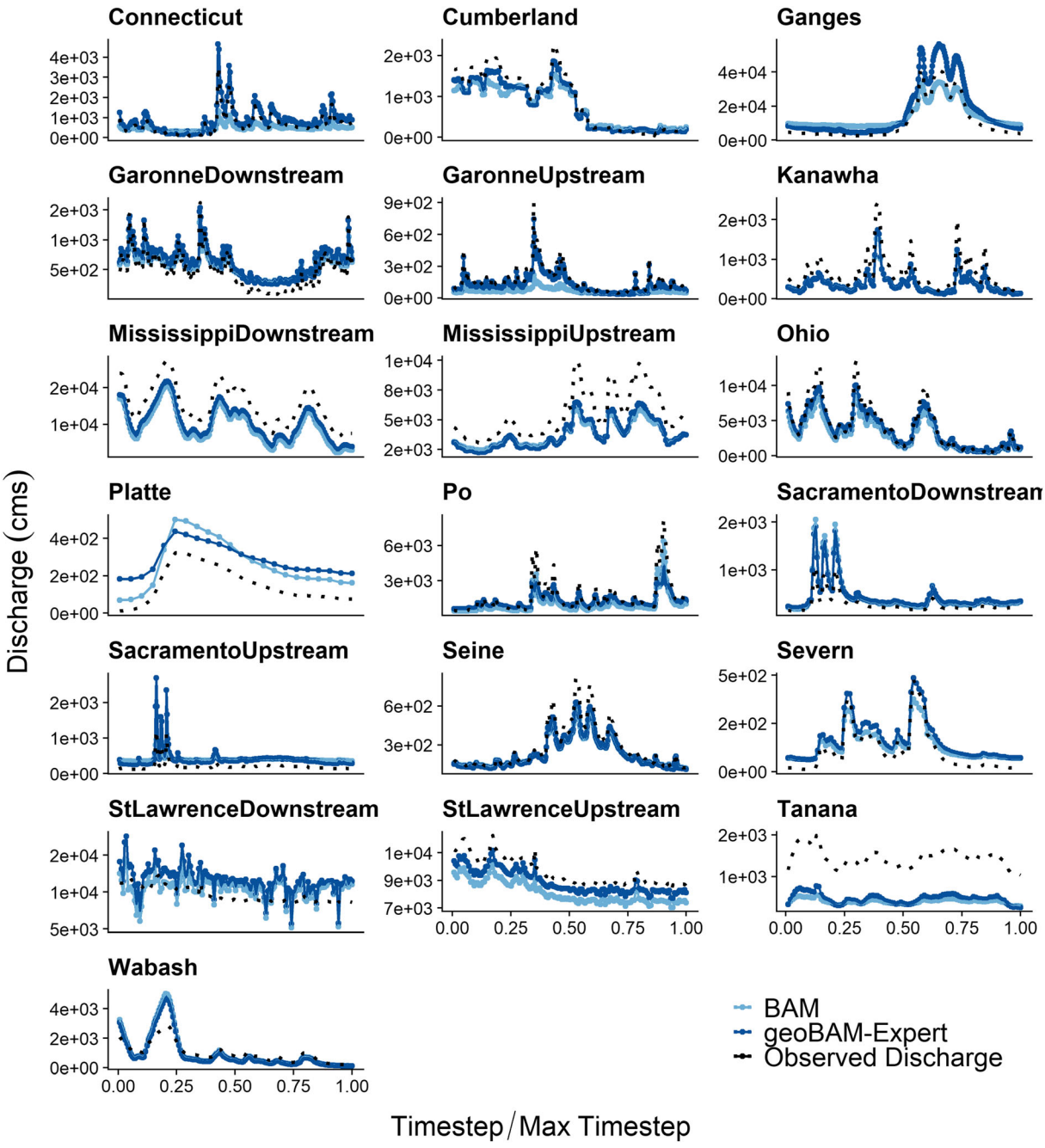

Figure 6. Hydrographs for the SWOT-simulated rivers: Observed discharge (dotted black) is plotted alongside BAM (light blue) and geoBAM-Expert (dark blue).

(Figure 5b). The cumulative density functions (CDFs) (Figure 5a) highlight this, where rivers with NSE < -1.5 had degraded performance with geoBAM-Expert.

Considering hydrographs, BAM performed quite well in low/baseflow periods across rivers, but much worse in high flow events (e.g., Severn, Cumberland, and Seine). In some rivers, geoBAM-Expert hydrographs (Figure 6—dark blue lines) significantly improved from those estimated using BAM (light blue lines—for example, St. Lawrence Upstream, Garonne Upstream, Cumberland, Connecticut, and Severn). In other rivers, very little changed (e.g., Wabash, Mississippi Downstream, and Tanana). Both BAM and geoBAM produced hydrographs that visually resemble observed flow dynamics, but, similar to the Mackenzie River test (Figure 4), geoBAM has "filled in" many errors in predicting the magnitude of peak events with varying

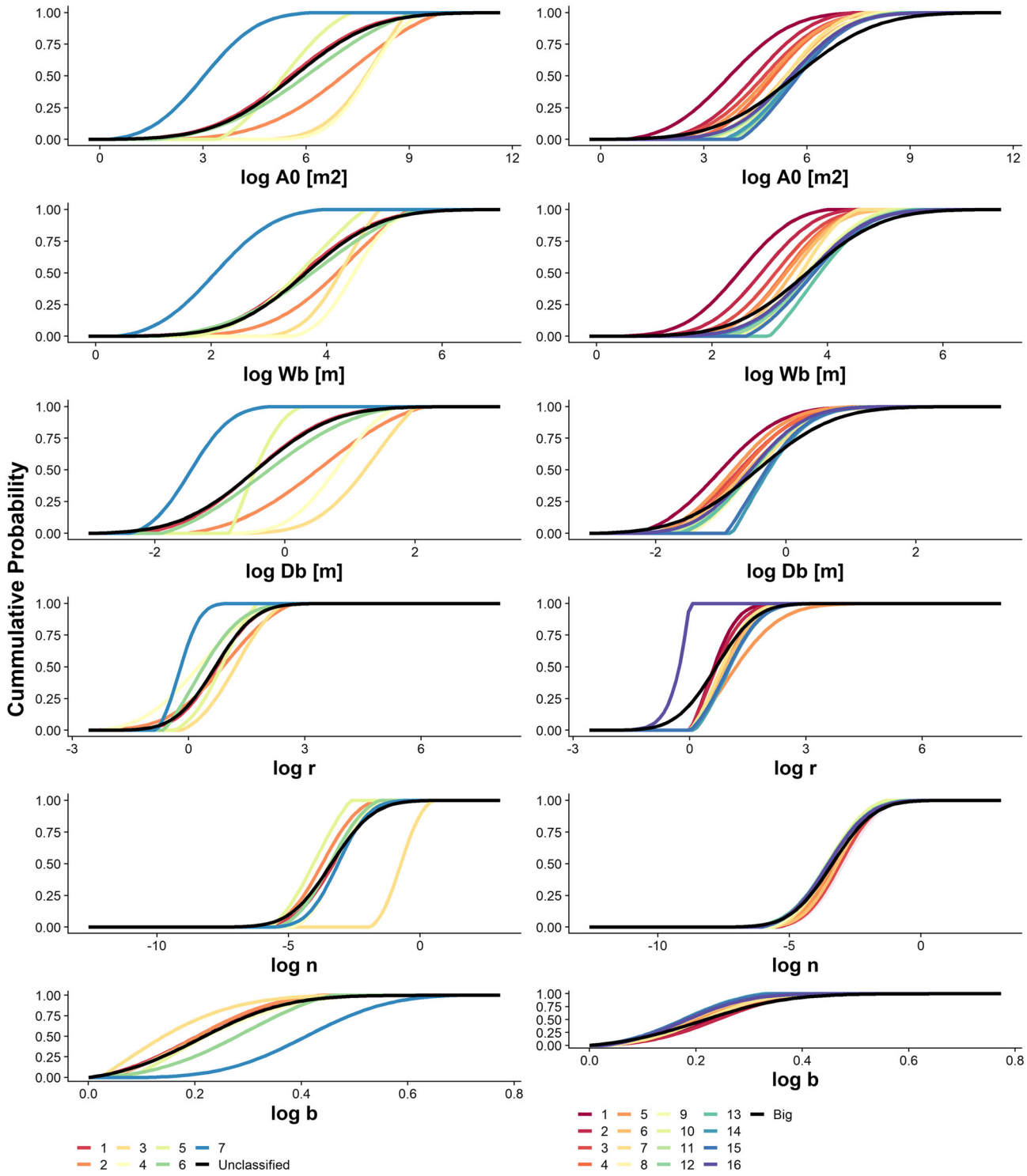

Figure 7. Truncated, lognormal distributions of hydraulic geoBAM priors as defined using the unsupervised classification (left column) or the expert classification (right column). CDFs with classes overlain on one another indicate the variable is not differentiable by class, whereas most variables are differentiable by class.

degrees of success. For example, the Ohio River was mostly modeled correctly by BAM except for errors in the three peak events (Figure 6—"Ohio"). geoBAM-Expert has minimized the error in these peak events, while continuing to accurately model the rest of the hydrograph. Some rivers with little to no change in NSE scores (Figure 5) still appear to have partially "filled in" these discrepancies (e.g., Seine, Severn, and

Cumberland). geoBAM-Expert results for both the Sacramento Upstream and St. Lawrence Downstream show visibly worsened reproduction of observed discharge than BAM.

4.3. Classification Comparison

Sections 4.1 and 4.2 show significant performance improvement in geoBAM discharge estimation when defining prior river knowledge via geomorphic river types. However, geoBAM-Expert and geoBAM-Unsupervised yielded functionally the same NSE performance in both the Mackenzie River basin and the SWOT-simulated rivers, despite being constructed very differently. Thus, we now turn to exploring differences in the river types these classifications produced. Both classification frameworks feature fundamentally different classifications and yield fundamentally different hydraulic priors across river types (Figure 7). Figure 7 plots truncated, parametric distributions of the six hydraulic terms calculated from the training data needed to run geoBAM for both classifications. Distributions for all six hydraulic terms are visually distinct for the unsupervised classification (left column), while A_0, W_b, and D_b appear to increase monotonically by expert river type (right column).

While these plots visually justify both classification frameworks as ways to extract differentiable prior distributions for hydraulic terms, we confirmed the uniqueness of the classes by running two one-way analysis of variance (ANOVA) tests on the medians of the six hydraulic terms' distributions in Figure 7, grouped by river type (Table S6). These ANOVA tests check whether there are statistically significant differences between the centers of the prior distributions per river type and are reported as p values, where a value <0.05 is generally considered statistically significant and a value between 0.05 and 0.1 is considered mildly significant. We note that distribution centers are not the only criteria necessary to define truncated probability distributions and that distributions can have similar centers but drastically different overall shapes. Still, this is a convenient way to quantify one dimension of differentiability. For the expert framework, W_b, D_b, n, and A_0 were significantly different ($p < 0.05$), while other variables were not different by class. For the unsupervised framework, b and r were marginally significantly different (p values between 0.05 and 0.10), while the other hydraulics were not significantly different by class. Overall, Figure 7 and Table S6 confirm that both classification frameworks yield differentiable river types across these six hydraulic terms.

Using the parametric distributions in Figure 7, we qualitatively defined what these river types represent. Based on the significantly different A_0, W_b, and D_b distributions for geoBAM-Expert, we interpret these river types as discrete representations of river size: channel area and bankfull geometry monotonically increase with river type. Per the definition of r and b in the expert system, Type 16 ("highly width-variable") rivers are fundamentally different from the other class distributions. The unsupervised river types are more difficult to qualitatively define, but the r distribution for River Type 7 appears strikingly similar to the expert system's River Type 16, suggesting that the unsupervised method successfully identified the highly width-variable cross sections itself. The unsupervised system also identified a river type with exceptionally high Manning's n values, distinct from the others. The center of this distribution is a Manning's n of 0.48, which is extremely high and generally reserved for very rough, vegetal-lined, artificial channels (Chow, 1959). The remaining river types cover the full spread of channel geometries experienced in our training data, but with varying degrees of certainty in the distribution centers.

4.4. Summary of Performance Metrics

Finally, we analyze RSQ accuracy across all experimental trials in this study using additional error metrics (Figure 8). Again, differences between the two proposed classification frameworks for either test are marginal, but improvement from BAM is stark. In the SWOT rivers, geoBAM-Expert had a slightly better median rBIAS score (−0.07 vs. −0.13), but otherwise, median performance scores were functionally the same across both classification frameworks tested.

While the magnitude of scores is worse across the board for the Mackenzie than the SWOT rivers (a byproduct of using of RS-observed widths only), the improvement in NSE is largely similar for both tests (Figure 8). The one notable difference between tests is the much smaller variation in scores (i.e., boxplot IQRs in Figure 8) for the Mackenzie reaches than the SWOT rivers. The SWOT rivers exhibited large ranges in predictive skill regardless of the BAM/geoBAM implementation.

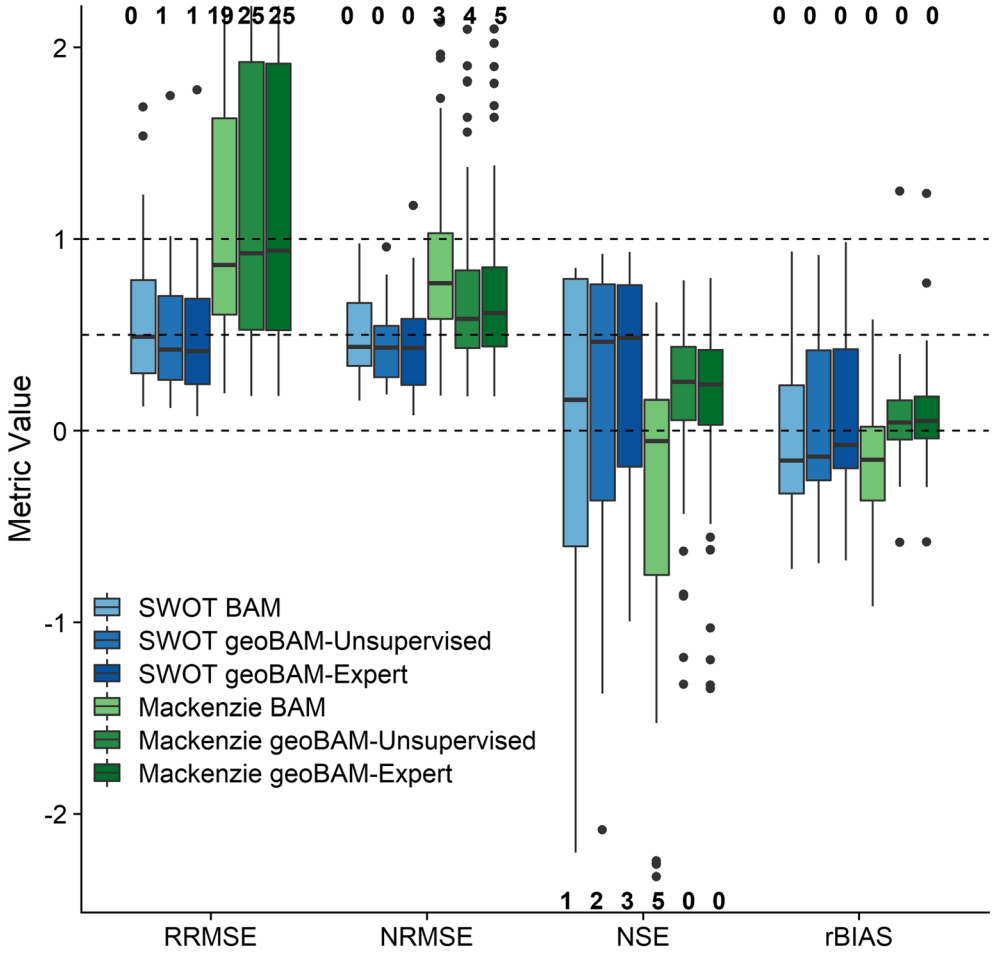

Figure 8. Comparison of geoBAM classifications for both the SWOT-simulated rivers and the Mackenzie River basin. Numbers reflect the number of SWOT rivers or Mackenzie reaches not plotted due to axis truncation.

5. Discussion

5.1. RSQ in Ungauged Basins

We find that the quality of prior knowledge significantly influences RSQ accuracy, as evidenced through large performance improvements from BAM to geoBAM (Figures 3–6 and 8). While algorithm developments also continue to advance our understandings of McFLIs and discharge inversion, we have found that our relatively simple interventions to prior estimation are easy to implement and did not require new inversion logic within the algorithm. We highlight our improvements via real-world RS observations in the Mackenzie River basin over a massive spatial scale for two reasons. First, all reaches used in this study are in the Arctic/Subarctic, where our training data are unrepresentative (i.e., none of the field hydraulic measurements were made there). The success of these interventions in a blind case study like this, with training data only from the continental United States, suggests that this approach is implementable globally. Second, we have satisfyingly replicated Feng et al.'s (2019) BAM results that were run only on 11 Alaskan rivers. We have extrapolated their RSQ workflow to thousands of reaches with geoBAM, relying on the globally available GRADES (Lin et al., 2019) for reach-explicit prior knowledge on discharge. These results corroborate the aggregate use of Feng et al.'s (2019) width mapping method, geoBAM, and GRADES for big-data RSQ and functionally open the door for uncalibrated RSQ across global-scale river networks.

5.2. "Local Relevance" of RSQ and SWOT Observations

Notably, we also achieved skill improvement without relying on new and computationally intensive data assimilation schemes or hydraulic models (e.g., solving full Saint-Venant equations). With that said, these

other approaches may yield much stronger predictive skill than ours, although with additional data and computational requirements. The utility of geoBAM vis a vis its accuracy is therefore interesting in light of other data intensive, higher skill approaches for discharge estimation. This "local relevance" of global hydrology has been explored from multiple perspectives (e.g., Benstead & Leigh, 2012; Bierkens et al., 2015; Brown et al., 2015; Fleischmann et al., 2019; Rajib et al., 2020). Specifically, Fleischmann et al. (2019) suggest that for a global hydrodynamic model to be "locally relevant," discharge estimates need to have an NSE > 0.90. Otherwise, discharge is simply not accurate enough for application in flood forecasting for specific rivers. Others, including Brown et al. (2015), suggest that "local relevance" is not equivalent to high model skill and that for data-poor regions, any improvement in hydrologic knowledge is of local relevance for water resource managers. geoBAM's median NSE in the Mackenize river basin is 0.25, which is far below Fleischmann et al.'s (2019) definition of relevance. Thus, we acknowledge that our results are not sufficiently accurate for flood work and similar applications at the local scale. However, the goal of the McFLI paradigm is as one piece of a broader puzzle in ungauged basins where it is difficult to calibrate hydrodynamic models because of a lack of data (due to physical, logistical, or political realities). McFLI estimates are also sparse in space and time and are not a forward model, so they cannot predict or forecast hydrographs: McFLI by itself can only estimate what happened when the satellite passed overhead. In isolation, geoBAM is therefore not a tool that can or should be used for water resources planning.

However, we show that geoBAM can update understanding of river geomorphology, provide more accurate hydraulic parameters, and provide discharge estimates with a positive NSE for completely ungauged basins. These can then be used to better parameterize forward models needed for water resources, either via calibration or assimilation. Previous work has been performed on the effects of parameterizing more traditional hydrodynamic models by river, rather than using global-scope parameters (e.g., Altenau et al., 2017; Neal et al., 2009). These studies, along with Andreadis et al. (2020) and Tuozzolo, Lind, et al. (2019), support the general notion that river-specific knowledge of a river will notably improve model skill, whether one is explicitly parameterizing a full hydrodynamic model or simply using RS to infer discharge. Thus, while we cannot argue that geoBAM discharges themselves are locally relevant per Fleischmann et al. (2019), we can argue that geoBAM results (discharges and geomorphic parameters) can move global modeling toward this local relevance, especially in ungauged basins. Notably, this capability should improve dramatically with the launch of SWOT, per Figure 8.

Interestingly, for the SWOT rivers significant performance improvement is limited to NSE, with marginal improvement seen for rBIAS and RRMSE (Figure 8). Conversely, the Mackenzie case exhibits noteworthy improvement for NRMSE and rBIAS (on par with NSE) but worsened performance in RRMSE. RRMSE and rBIAS generally track together and are easily inflated due to errors in baseflow prediction. Our interventions alter baseflow predictions very little in both test cases (Figures 4 and 6) and because of this it is possible that these two metrics are relatively insensitive to geoBAM's interventions in those rivers. In the Mackenzie River basin, the switch from a negative to a positive rBIAS is likely the reason for an increasing RRMSE, per these metrics' definitions (Table S5).

5.3. Comparisons and Improvements to McFLI Algorithms

We are aware of one other McFLI that has explicitly tested the influence of prior quality. Andreadis et al.'s (2020) SWOT Assimilated Discharge (SAD) algorithm was developed to address parameter equifinality issues in McFLIs by constraining their parameter space for discharge inversion. To test this, they introduced a suite of interventions to SAD, one of which was an expert geomorphic classification that defined their prior on channel shape. While it is difficult to directly compare results across algorithms and with different interventions, when they isolated this intervention median NSE across 18 SWOT-simulated rivers increased by 0.35. Using geoBAM-Expert, median NSE for our 19 SWOT-simulated rivers increased by 0.27. Such similar results, for this one test, suggest that this degree of improvement occurs across McFLIs, assuming the algorithm is flexible enough to ingest different definitions of prior river knowledge (like BAM and SAD are). Despite the similarities, the SAD geomorphic classification is not globally scalable because it relies on manual interpretation of a river's planform geometry, and classification is limited exclusively to the channel shape parameter r. Further, SAD does not invoke AMHG like BAM does. Conversely, geoBAM is globally scalable, can be used to extract distributions for any desired hydraulic term, and yields equivalent improvement in discharge prediction.

Future work should incorporate geoBAM's classification into SAD, and other McFLIs, to see if performance further improves when using a geomorphic classification for all priors. It is unclear if this approach in improving prior quality will work with all McFLIs. Hagemann et al. (2017) purposefully designed BAM to be flexible in both the priors it ingests and the flow laws it inverts, allowing us to run this study in two disparate settings: (1) using simulated altimetry data to invert Manning's equation and (2) using river widths to invert AMHG. Our similar results in both settings (sections 4.1 and 4.2) suggest that the McFLI paradigm would broadly benefit from improved prior quality, regardless of the flow law or RS observations used. However, we exclusively tested this within BAM, meaning these interventions may not coexist with the mathematical and/or computational setups in other McFLIs. In particular, other McFLI algorithms may have priors that do not correspond to our in situ data or classes. Not all McFLIs are the same even though many use Manning's equation: Each algorithm has its own unique prior set. Whether or not our classifications are useful to all McFLIs or just to BAM is left for future work.

For RSQ, classifying rivers to provide better priors will only yield improvements if the wrong class if assigned. Because we reduce priors to look-up tables with a set of priors assigned to each river type, if the wrong class assigned to a river, then RSQ will considerably worsen. Andreadis et al. (2020) found a similar result when their expert classification misclassified two rivers and yielded poor performance. When they artificially assigned the correct river type, they improved RSQ accuracy. Similar behavior was identified in this study: the St. Lawrence Upstream's NSE score was improved from −1.15 (BAM) to 0.45 (geoBAM-Expert) but degraded to −2.08 when using geoBAM-Unsupervised. Clearly, there is room for improvement, and in this example it appears that geoBAM-Unsupervised is assigning an incorrect class, or the unsupervised classes are unrepresentative of this river. This suggests that RSQ accuracy hinges on how river types are defined and ultimately mapped to rivers, and the massive range of performance accuracy within river types (section 4.2) suggests that there is substantial room to improve how river types are extracted and mapped to RS data. We suggest that because our river type mapping procedure uses only river widths to predict types, we are missing crucial information on predicting “correct” river types which might reduce the variance in performance within river types. Regardless, these errors occurred in only a few rivers (St. Lawrence Upstream, St. Lawrence Downstream, Platte, and Sacramento Downstream). and both of our classification frameworks are globally scalable with much improved predictive accuracy and use just time-varying river widths as predictors.

Finally, the fact that the SWOT rivers are all very large might explain why Manning's-based McFLIs generally perform well in these tests. Manning's equation simulates discharge in deeper, larger rivers quite well, but not in shallower, smaller ones where n is underestimated relative to field measurements (Ferguson, 2010). This is important in the SWOT context, as SWOT will be limited to a relatively coarse spatial resolution and only observe rivers wider than 100 m (with a goal of 50 m wide rivers). Thus, SWOT will miss small streams, making it impossible for global SWOT discharge retrievals via geoBAM (or any other algorithm for that matter) to accurately represent small river hydrology by itself. However, geoBAM via AMHG uses only river widths and can be reasonably run on any river that is “hydraulically visible,” that is, exhibits a hydrological response via RS (Garambois et al., 2017). For instance, Feng et al. (2019) successfully ran BAM (not geoBAM) on rivers as narrow as 20 m using PlanetLabs imagery, which has a spatial resolution less than 3 m.

5.4. Classifying Global Rivers

It is useful to orient our classification workflow in the context of other global-scale river classification frameworks. Interestingly, geoBAM-Unsupervised largely replicated the results using the bespoke geoBAM-Expert, signifying that stark differences in classification (Figure 6) did not manifest in discharge estimation. The success of such simple unsupervised classification suggests that a generalized, global river geomorphology framework is possible. There are currently very few of these frameworks (e.g., Dallaire et al., 2019; Fernandez & Sayama, 2015; Haines et al., 1988; Puckridge et al., 1998), and only one is globally consistent in coverage (Dallaire et al., 2019). Dallaire et al. (2019) clustered features explicitly associated with every observation/river reach within HydroSHEDS (Lehner et al., 2008); however, fluvial geomorphology data do not generally exist in this form at the global scale and so their analysis was largely limited to hydroclimatic river types. Conversely, Guillon et al. (2020) successfully used machine learning models to upscale a priori geomorphic river types for the Sacramento River basin—defined using field geomorphology

campaigns at 290 sites—to over 100,000 reaches. Our study presents a novel amalgamation of these two methods, first using automated clustering of field data to define a geomorphic classification framework and then using supervised learning to upscale river types to anywhere on Earth. section 4.4 shows that this approach is viable in the Mackenzie River basin, where we accurately represent landscape geomorphology through our river types (Figure 2b). However, more sophisticated methods will likely be needed to produce a generalized, global river geomorphology framework that moves beyond RSQ.

Finally, Durand et al. (2016) noted the potential homogeneity of test rivers used for benchmarking McFLIs, but until now there has not been a geomorphically explicit way to quantify this homogeneity. Using the geoBAM-Expert river types, we found that the SWOT-simulated rivers are quite homogenous and large (23/132 reaches were classified as "big" rivers and 52/132 reaches were assigned River Type 15, the widest in the classification). Even more striking, geoBAM-Unsupervised assigned 16/19 rivers the same river type. These are not reflective of the global variation in hydraulics and geomorphology SWOT will encounter. What is thus needed is a wider range of river types for validating and testing McFLIs before SWOT launches, such that we can parse out specific river types that McFLIs model well, and those that McFLIs model poorly.

6. Conclusions

This study presents a first attempt at improving the quality of prior river knowledge for McFLI RSQ in two distinct settings: on thousands of Arctic river reaches using Landsat imagery and on simulated rivers representing NASA/CNES/CSA/UKSA SWOT satellite outputs prior to its launch in 2022. Prior quality was improved via two methods: (1) providing a larger and geomorphically varying training data set and (2) statistical mapping of priors to river types, using both unsupervised and expert methods. We found significant improvement in the accuracy of discharge predictions for both test cases and using both classification methods, with a mean NSE improvement (from BAM to geoBAM-Unsupervised) for the Mackenzie river of 0.64 and of 0.31 for the SWOT-simulated rivers. Both classification methods yielded functionally the same improvement in accuracy, despite vastly different constructions. These findings are significant and highlight the importance of prior knowledge in a Bayesian mathematical setting, where we have shown that starting from a more informed understanding of the river yields more accurate results. These priors may be implementable in any McFLI and will play a pivotal role in both current RSQ efforts in global river networks and future global RSQ from SWOT.

Data Availability Statement

Results and code to reproduce our figures are available online (at http://doi.org/10.5281/zenodo.4073609). geoBAM is available at GitHub (https://github.com/craigbrinkerhoff/geoBAMr). All other data used in this study are freely available online.

Acknowledgments
This work was supported by the following grants, awarded to C. J. Gleason: NSF CAREER grant 1748653 (supporting C. J. Gleason), NASA New Investigator Program grant 80NSSC18K0741 (supporting D. Feng), and NASA SWOT Science Team grant NNX13AD96G (partially supporting C. B. Brinkerhoff). The authors would like to thank Mark Hagemann for building and sharing the original BAM algorithm.

References

Allen, G. H., & Pavelsky, T. M. (2018). Global extent of rivers and streams. *Science*, *361*(6402), 585–588. https://doi.org/10.1126/science.aat0636

Altenau, E. H., Pavelsky, T. M., Bates, P. D., & Neal, J. C. (2017). The effects of spatialresolution and dimensionality on modeling regional-scale hydraulics in a multichannel river. *Water Resources Research*, *53*, 1683–1701. https://doi.org/10.1002/2016WR019396

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

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

Benstead, J. P., & Leigh, D. S. (2012). An expanded role for river networks. *Nature Geoscience*, *5*(10), 678–679. https://doi.org/10.1038/ngeo1593

Biancamaria, S., Durand, M., Andreadis, K. M., Bates, P. D., Boone, A., Mognard, N. M., et al. (2011). Assimilation of virtual wide swath altimetry to improve Arctic river modeling. *Remote Sensing of Environment*, *115*(2), 373–381. https://doi.org/10.1016/j.rse.2010.09.008

Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016). The SWOT mission and its capabilities for land hydrology. In A. Cazenave, N. Champollion, J. Benveniste, J. Chen (Eds.), *Remote sensing and water resources* (pp. 117–147). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-32449-4_6

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., et al. (2015). Hyper-resolution global hydrological modelling: What is next? *Hydrological Processes*, *29*(2), 310–320. https://doi.org/10.1002/hyp.10391

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer. Retrieved April 22, 2020, from https://cds.cern.ch/record/998831

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. *Journal of Hydrology*, *561*, 1000–1018. https://doi.org/10.1016/j.jhydrol.2018.04.005

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

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 Resources Research*, *56*, e2019WR025986. https://doi.org/10.1029/2019WR025986

Bjerklie, D. M., Moller, D., Smith, L. C., & Dingman, S. L. (2005). Estimating discharge in rivers using remotely sensed hydraulic information. *Journal of Hydrology*, *309*(1), 191–209. https://doi.org/10.1016/j.jhydrol.2004.11.022

Bonnema, M. G., Sikder, S., Hossain, F., Durand, M., Gleason, C. J., & Bjerklie, D. M. (2016). Benchmarking wide swath altimetry-based river discharge estimation algorithms for the Ganges river system. *Water Resources Research*, *52*, 2439–2461. https://doi.org/10.1002/2015WR017296

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

Brinkerhoff, C. B., Gleason, C. J., & Ostendorf, D. W. (2019). Reconciling at-a-station and at-many-stations hydraulic geometry through river-wide geomorphology. *Geophysical Research Letters*, *46*, 9637–9647. https://doi.org/10.1029/2019GL084529

Brown, C. M., Lund, J. R., Cai, X., Reed, P. M., Zagona, E. A., Ostfeld, A., et al. (2015). The future of water resources systems analysis: Toward a scientific framework for sustainable water management. *Water Resources Research*, *51*, 6110–6124. https://doi.org/10.1002/2015WR017114

Canova, M. G., Fulton, J. W., & Bjerklie, D. M. (2016). *USGS HYDRoacoustic dataset in support of the Surface Water Oceanographic Topography satellite mission (HYDRoSWOT) [dataset]*. United States Geological Survey. https://doi.org/10.5066/F7D798H6

Chandanpurkar, H. A., Reager, J. T., Famiglietti, J. S., & Syed, T. H. (2017). Satellite- and reanalysis-based mass balance estimates of global continental discharge (1993–2015). *Journal of Climate*, *30*(21), 8481–8495. https://doi.org/10.1175/JCLI-D-16-0708.1

Chen, S.-A., Michaelides, K., Grieve, S. W. D., & Singer, M. B. (2019). Aridity is expressed in river topography globally. *Nature*, *573*(7775), 573–577. https://doi.org/10.1038/s41586-019-1558-8

Chow, V. T. (1959). *Open-channel hydraulics*. New York: McGraw-Hill.

Dallaire, C. O., Lehner, B., Sayre, R., & Thieme, M. (2019). A multidisciplinary framework to derive global river reach classifications at high spatial resolution. *Environmental Research Letters*, *14*(2), 024003. https://doi.org/10.1088/1748-9326/aad8e9

Dingman, S. L. (2007). Analytical derivation of at-a-station hydraulic–geometry relations. *Journal of Hydrology*, *334*(1), 17–27. https://doi.org/10.1016/j.jhydrol.2006.09.021

Dingman, S. L., & Afshari, S. (2018). Field verification of analytical at-a-station hydraulic-geometry relations. *Journal of Hydrology*, *564*, 859–872. https://doi.org/10.1016/j.jhydrol.2018.07.020

Dingman, S. L., & Sharma, K. P. (1997). Statistical development and validation of discharge equations for natural channels. *Journal of Hydrology*, *199*(1), 13–35. https://doi.org/10.1016/S0022-1694(96)03313-6

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

Durand, M., Gleason, C. J., Garambois, P. A., Bjerklie, D., Smith, L. C., Roux, H., et al. (2016). An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. *Water Resources Research*, *52*, 4527–4549. https://doi.org/10.1002/2015WR018434

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. *Journal of Hydrology*, *511*, 92–104. https://doi.org/10.1016/j.jhydrol.2013.12.050

Emery, C. M., Paris, A., Biancamaria, S., Boone, A., Calmant, S., Garambois, P.-A., & Santos da Silva, J. (2018). Large-scale hydrological model river storage and discharge correction using a satellite altimetry-based discharge product. *Hydrology and Earth System Sciences*, *22*(4), 2135–2162. http://doi.org/10.5194/hess-22-2135-2018

Ester, M., Kriegel, H.-P., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, *96*(34), 226–231. https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf?source=post_page

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 Resources Research*, *55*, 7753–7771. https://doi.org/10.1029/2019WR025599

Ferguson, R. (2010). Time to abandon the Manning equation? *Earth Surface Processes and Landforms*, *35*(15), 1873–1876. https://doi.org/10.1002/esp.2091

Fernandez, R., & Sayama, T. (2015). Hydrological recurrence as a measure for large river basin classification and process understanding. *Hydrology and Earth System Sciences*, *19*(4), 1919–1942. https://doi.org/10.5194/hess-19-1919-2015

Fleischmann, A., Paiva, R., & Collischonn, W. (2019). Can regional to continental river hydrodynamic models be locally relevant? A cross-scale comparison. *Journal of Hydrology X*, *3*, 100027. https://doi.org/10.1016/j.hydroa.2019.100027

Frasson, R. P. D. M., Durand, M. T., & Rodríguez, E. (2019). *Compilation of hydraulic models for the study of the spatial averaging on flow laws [data set]*. Zenodo. https://doi.org/10.5281/zenodo.3463541

Frasson, R. P. d. M., Pavelsky, T. M., Fonstad, M. A., Durand, M. T., Allen, G. H., Schumann, G., et al. (2019). Global relationships between river width, slope, catchment area, meander wavelength, sinuosity, and discharge. *Geophysical Research Letters*, *46*, 3252–3262. https://doi.org/10.1029/2019GL082027

Garambois, P.-A., Calmant, S., Roux, H., Paris, A., Monnier, J., Finaud-Guyot, P., et al. (2017). Hydraulic visibility: Using satellite altimetry to parameterize a hydraulic model of an ungauged reach of a braided river. *Hydrological Processes*, *31*(4), 756–767. https://doi.org/10.1002/hyp.11033

Garambois, P.-A., & Monnier, J. (2015). Inference of effective river properties from remotely sensed observations of water surface. *Advances in Water Resources*, *79*, 103–120. https://doi.org/10.1016/j.advwatres.2015.02.007

Gleason, C., Garambois, P.-A., & Durand, M. (2017). Tracking River flows from space. *Eos*. 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 Sensing*, *12*(7), 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. *Proceedings of the National Academy of Sciences*, *111*(13), 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 Resources Research*, *50*, 9604–9619. https://doi.org/10.1002/2014WR016109

Gleason, C. J., & Wang, J. (2015). Theoretical basis for at-many-stations hydraulic geometry. *Geophysical Research Letters*, *42*, 7107–7114. https://doi.org/10.1002/2015GL064935

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., et al. (2019). Mapping the world's free-flowing rivers. *Nature*, *569*(7755), 215–221. https://doi.org/10.1038/s41586-019-1111-9

Guillon, H., Byrne, C. F., Lane, B. A., Solis, S. S., & Pasternack, G. B. (2020). Machine learning predicts reach-scale channel types from coarse-scale geospatial data in a large river basin. *Water Resources Research*, *56*, e2019WR026691. https://doi.org/10.1029/2019WR026691

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 Resources Research*, *53*, 9692–9707. https://doi.org/10.1002/2017WR021626

Haines, A. T., Finlayson, B. L., & McMahon, T. A. (1988). A global classification of river regimes. *Applied Geography*, *8*(4), 255–272. https://doi.org/10.1016/0143-6228(88)90035-5

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer Science & Business Media.

Henderson, F. M. (1966). *Open channel flow*. New York: Macmillian.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). Unsupervised learning. In G. James, D. Witten, T. Hastie, R. Tibshirani (Eds.), *An introduction to statistical learning: With applications in R* (pp. 373–418). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-7138-7_10

King, T. V., Neilson, B. T., & Rasmussen, M. T. (2018). Estimating discharge in low-order rivers with high-resolution aerial imagery. *Water Resources Research*, *54*, 863–878. https://doi.org/10.1002/2017WR021868

Knighton, A. D. (1975). Variations in at-a-station hydraulic geometry. *American Journal of Science*, *275*(2), 186–218. https://doi.org/10.2475/ajs.275.2.186

LeFavour, G., & Alsdorf, D. (2005). Water slope and discharge in the Amazon River estimated using the shuttle radar topography mission digital elevation model. *Geophysical Research Letters*, *32*, L17404. https://doi.org/10.1029/2005GL023836

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94. https://doi.org/10.1029/2008EO100001

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

Lin, P., Pan, M., Allen, G. H., de Frasson, R. P., Zeng, Z., Yamazaki, D., & Wood, E. F. (2020). Global estimates of reach-level bankfull river width leveraging big data geospatial analysis. *Geophysical Research Letters*, *47*, e2019GL086405. https://doi.org/10.1029/2019GL086405

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., et al. (2019). Global reconstruction of Naturalized River flows at 2.94 million reaches. *Water Resources Research*, *55*, 6499–6516. https://doi.org/10.1029/2019WR025287

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

Olden, J. D., Kennard, M. J., & Pusey, B. J. (2012). A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. *Ecohydrology*, *5*(4), 503–518. https://doi.org/10.1002/eco.251

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 Resources Research*, *54*, 2405–2423. https://doi.org/10.1002/2017WR021735

Oubanas, H., Gejadze, I., Malaterre, P.-O., & Mercier, F. (2018). River discharge estimation from synthetic SWOT-type observations using variational data assimilation and the full Saint-Venant hydraulic model. *Journal of Hydrology*, *559*, 638–647. https://doi.org/10.1016/j.jhydrol.2018.02.004

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

Pavelsky, T. M., & Smith, L. C. (2009). Remote sensing of suspended sediment concentration, flow velocity, and lake recharge in the Peace-Athabasca Delta, Canada. *Water Resources Research*, *45*, W11417. https://doi.org/10.1029/2008WR007424

Petit, F., & Pauquet, A. (1997). Bankfull discharge recurrence interval in gravel-bed rivers. *Earth Surface Processes and Landforms*, *22*(7), 685–693. https://doi.org/10.1002/(SICI)1096-9837(199707)22:7<685::AID-ESP744>3.0.CO;2-J

Puckridge, J. T., Sheldon, F., Walker, K. F., & Boulton, A. J. (1998). Flow variability and the ecology of large rivers. *Marine and Freshwater Research*, *49*(1), 55–72. https://doi.org/10.1071/mf94161

Rajib, A., Liu, Z., Merwade, V., Tavakoly, A. A., & Follum, M. L. (2020). Towards a large-scale locally relevant flood inundation modeling framework using SWAT and LISFLOOD-FP. *Journal of Hydrology*, *581*, 124406. https://doi.org/10.1016/j.jhydrol.2019.124406

Ross, M. R. V., Topp, S. N., Appling, A. P., Yang, X., Kuhn, C., Butman, D., et al. (2019). AquaSat: A data set to enable remote sensing of water quality for inland waters. *Water Resources Research*, *55*, 10,012–10,025. https://doi.org/10.1029/2019WR024883

Sichangi, A. W., Wang, L., & Hu, Z. (2018). Estimation of river discharge solely from remote-sensing derived data: An initial study over the Yangtze River. *Remote Sensing*, *10*(9), 1385. https://doi.org/10.3390/rs10091385

Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., & Boni, G. (2015). Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data. *Hydrology and Earth System Sciences*, *19*(4), 1727–1751. https://doi.org/10.5194/hess-19-1727-2015

Siqueira, V. A., Paiva, R. C. D. de, Fleischmann, A. S., Fan, F. M., Ruhoff, A. L., Pontes, P. R. M., et al. (2018). Toward continental hydrologic–hydrodynamic modeling in South America. Retrieved from https://lume.ufrgs.br/handle/10183/184110

Spence, C., & Woo, M. (2008). Hydrology of the northwestern subarctic Canadian Shield. In M. Woo (Ed.), *Cold Region Atmospheric and Hydrologic Studies. The Mackenzie GEWEX Experience*, *Hydrologic Processes* (Vol. 2, pp. 235–256). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-75136-6_13

Sun, W., Ishidaira, H., Bastola, S., & Yu, J. (2015). Estimating daily time series of streamflow using hydrological model calibrated based on satellite observations of river water surface width: Toward real world applications. *Environmental Research*, *139*, 36–45. https://doi.org/10.1016/j.envres.2015.01.002

Tarpanelli, A., Brocca, L., Lacava, T., Melone, F., Moramarco, T., Faruolo, M., et al. (2013). Toward the estimation of river discharge variations using MODIS data in ungauged basins. *Remote Sensing of Environment*, *136*, 47–55. https://doi.org/10.1016/j.rse.2013.04.010

Tuozzolo, S., Langhorst, T., de Moraes Frasson, R. P., Pavelsky, T., Durand, M., & Schobelock, J. J. (2019). The impact of reach averaging Manning's equation for an in-situ dataset of water surface elevation, width, and slope. *Journal of Hydrology*, *578*, 123866. https://doi.org/10.1016/j.jhydrol.2019.06.038

Tuozzolo, S., Lind, G., Overstreet, B., Mangano, J., Fonstad, M., Hagemann, M., et al. (2019). Estimating river discharge with swath altimetry: A proof of concept using AirSWOT observations. *Geophysical Research Letters*, *46*, 1459–1466. https://doi.org/10.1029/2018GL080771

Wieczorek, M. E., Jackson, S. E., & Schwarz, G. E. (2018). *Select attributes for NHDPlus Version 2.1 reach catchments and modified network routed upstream watersheds for the conterminous United States [dataset]*. U.S. Geological Survey. https://doi.org/10.5066/F7765D7V

Williams, G. P. (1978). Bank-full discharge of rivers. *Water Resources Research*, *14*, 1141–1154. https://doi.org/10.1029/WR014i006p01141

Wisser, D., Fekete, B. M., Vorosmarty, C. J., & Schumann, A. H. (2010). Reconstructing 20th century global hydrography: A contribution to the global terrestrial network- hydrology (GTN-H). *Hydrology and Earth System Sciences*, *24*. https://doi.org/10.5194/hess-14-1-2010

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 Resources Research*, *55*, 5053–5073. https://doi.org/10.1029/2019WR024873

Yang, X., Pavelsky, T. M., & Allen, G. H. (2020). The past and future of global river ice. *Nature*, *577*(7788), 69–73. https://doi.org/10.1038/s41586-019-1848-1

Yang, X., Pavelsky, T. M., Allen, G. H., & Donchyts, G. (2019). RivWidthCloud: An automated Google Earth engine algorithm for river width extraction from remotely sensed imagery. *IEEE Geoscience and Remote Sensing Letters*, *17*(2), 217–221. https://doi.org/10.1109/LGRS.2019.2920225

Yoon, Y., Durand, M., Merry, C. J., Clark, E. A., Andreadis, K. M., & Alsdorf, D. E. (2012). Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of Hydrology*, *464–465*, 363–375. https://doi.org/10.1016/j.jhydrol.2012.07.028

Zhang, Y., Pan, M., & Wood, E. F. (2016). On creating global gridded terrestrial water budget estimates from satellite remote sensing. In A. Cazenave, N. Champollion, J. Benveniste, J. Chen (Eds.), *Remote sensing and water resources* (pp. 59–78). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-32449-4_4

References From the Supporting Information

Adams, T., Chen, S., Davis, R., Schade, T., & Lee, D. (2010). The Ohio River community HEC-RAS model. *World Environmental and Water Resources Congress*, *2010*, 1512–1523. https://doi.org/10.1061/41114(371)160

Besnard, A., & Goutal, N. (2011). Comparison between 1D and 2D models for hydraulic modeling of a floodplain: Case of Garonne river. *Houille Blanche-Revue Internationale de l eau*, *3*, 42–47. https://doi.org/10.1051/lhb/2011031

Di Baldassarre, G., Schumann, G., & Bates, P. (2009). Near real time satellite imagery to support and verify timely flood modelling. *Hydrological Processes*, *23*(5), 799–803. https://doi.org/10.1002/hyp.7229

Heniche, M., Secretan, Y., Boudreau, P., & Leclerc, M. (2000). A two-dimensional finite element drying-wetting shallow water model for rivers and estuaries. *Advances in Water Resources*, *23*(4), 359–372. https://doi.org/10.1016/S0309-1708(99)00031-7

Humphries, E., Pavelsky, T., & Bates, P. D. (2014). Two dimensional hydrodynamic modeling of a high latitude braided river. *AGU Fall Meeting Abstracts*, *43*, H43H–H1042H.

Maswood, M., & Hossain, F. (2016). Advancing river modelling in ungauged basins using satellite remote sensing: The case of the Ganges-Brahmaputra-Meghna basin. *International Journal of River Basin Management*, *14*(1), 103–117. https://doi.org/10.1080/15715124.2015.1089250

Rahmah, N., & Sitanggang, I. S. (2016). Determination of optimal epsilon (Eps) value on DBSCAN algorithm to clustering data on peatland hotspots in Sumatra. *IOP Conference Series: Earth and Environmental Science*, *31*, 012012. https://doi.org/10.1088/1755-1315/31/1/012012

Rogers, W. (2014). *Central valley floodplain evaluation and delineation, subtask 5, combined Sacramento River system model Rep*. Sacramento, California: California Department of Water Resources.

Schubert, J. E., Monsen, W. W., & Sanders, B. F. (2015). Metric-resolution 2D river modeling at the macroscale: Computational methods and applications in a Braided River. *Frontiers in Earth Science*, *3*. https://doi.org/10.3389/feart.2015.00074

Siddique-E-Akbor, A. H. M., Hossain, F., Lee, H., & Shum, C. K. (2011). Inter-comparison study of water level estimates derived from hydrodynamic–hydrologic model and satellite altimetry for a complex deltaic environment. *Remote Sensing of Environment*, *115*(6), 1522–1531. https://doi.org/10.1016/j.rse.2011.02.011

Vilmin, L., Flipo, N., de Fouquet, C., & Poulin, M. (2015). Pluri-annual sediment budget in a navigated river system: The Seine River (France). *Science of the Total Environment*, *502*, 48–59. https://doi.org/10.1016/j.scitotenv.2014.08.110

Zou, Z., Xiao, X., Dong, J., Qin, Y., Doughty, R. B., Menarguez, M. A., et al. (2018). Divergent trends of open-surface water body area in the contiguous United States from 1984 to 2016. *Proceedings of the National Academy of Sciences*, *115*(15), 3810–3815. https://doi.org/10.1073/pnas.1719275115