

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2019WR025599

### Key Points:

- We present the first assessment of the suitability of CubeSat satellite data for river discharge estimation
- We estimate discharge using individual and fused satellite data sets on rivers from <20 to >1,000 m wide
- We suggest that the fused width data from both traditional and CubeSat satellites can improve global knowledge of Arctic river discharge

### Supporting Information:

- Supporting Information S1

### Correspondence to:

D. Feng,  
dongmeifeng@umass.edu

### Citation:

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>





Received 20 MAY 2019

Accepted 20 AUG 2019

Accepted article online 24 AUG 2019

Published online 9 SEP 2019

## Comparing Discharge Estimates Made via the BAM Algorithm in High-Order Arctic Rivers Derived Solely From Optical CubeSat, Landsat, and Sentinel-2 Data

Dongmei Feng<sup>1</sup> , Colin J. Gleason<sup>1</sup> , Xiao Yang<sup>2</sup> , and Tamlin M. Pavelsky<sup>2</sup> 

<sup>1</sup>Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA, USA,

<sup>2</sup>Department of Geological Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

**Abstract** Conventional satellite platforms are limited in their ability to monitor rivers at fine spatial and temporal scales: suffering from unavoidable trade-offs between spatial and temporal resolutions. CubeSat constellations, however, can provide global data at high spatial and temporal resolutions, albeit with reduced spectral information. This study provides a first assessment of using CubeSat data for river discharge estimation in both gauged and ungauged settings. Discharge was estimated for 11 Arctic rivers with sizes ranging from 16 to >1,000 m wide using the Bayesian at-many-stations hydraulic geometry-Manning algorithm (BAM). BAM-at-many-stations hydraulic geometry solves for hydraulic geometry parameters to estimate flow and requires only river widths as input. Widths were retrieved from Landsat 8 and Sentinel-2 data sets and a CubeSat (the Planet company) data set, as well as their fusions. Results show satellite data fusion improves discharge estimation for both large (>100 m wide) and medium (40–100 m wide) rivers by increasing the number of days with a discharge estimation by a factor of 2–6 without reducing accuracy. Narrow rivers (<40 m wide) are too small for Landsat and Sentinel-2 data sets, and their discharge is also not well estimated using CubeSat data alone, likely because the four-band sensor cannot resolve water surfaces accurately enough. BAM technique outperforms space-based rating curves when gauge data are available, and its accuracy is acceptable when no gauge data are present (instead relying on global reanalysis for discharge priors). Ultimately, we conclude that the data fusion presented here is a viable approach toward improving discharge estimates in the Arctic, even in ungauged basins.

## 1. Introduction

Rivers and streams are an essential part of the global hydrologic cycle: 90% of water flux transported from continents to oceans (roughly 37% of total terrestrial precipitation) is carried by rivers (Oki & Kanae, 2006). River flows provide valuable freshwater supplies for humans and ecosystems, create habitats for aquatic organisms, provide corridors for human transportation and fish migration, transport nutrients/sediments for estuarine ecosystems, and are also an important interface for mass and energy exchange between water and the atmosphere (Aguilera & Melack, 2018; Allen & Pavelsky, 2018; Barnett et al., 2005; Feng et al., 2016, 2019). Knowledge of river discharge (alternatively streamflow) is thus of great significance for water resource management, civil infrastructure design, water quality control, ecosystem conservation, and for understanding interactions in the Atmosphere-Land-Ocean system.

River flow dynamics respond to other hydrologic components and are subject to substantial spatial and temporal variations, which complicates understanding of streamflow conditions. Traditionally, river discharge is measured continuously in situ at specific locations (i.e., via river gauges). These gauges are reliable, accurate, and necessary for global hydrologic studies. However, they require instrument calibration and long-term maintenance, which limits them to only physically accessible and, largely, economically developed regions (Hagemann et al., 2017). Since 1980, the number of active gauges has dramatically declined globally, rendering data from about 60% of publicly available gauges at least 10 years out of date (Hannah et al., 2011; Vörösmarty et al., 2010, GRDC Database). Thus, in situ measurement of streamflow is limited from both spatial and temporal perspectives. Proprietary realities, which restrict the public availability of gauge data that do exist, make this limitation even worse (Gleason & Hamdan, 2017). Another approach for discharge quantification is hydrologic modeling, which integrates meteorological forcings and watershed hydrogeological

information into a numeric model to estimate discharge. However, as a simplified representation of natural hydrological processes, hydrologic models often do not represent hydrologic reality, even if discharge is modeled correctly. For example, discharge estimations with good fitness to the observations can be achieved by inappropriately simulated runoff quantities, routing, soil parameters, or evaporation rates. Even though discharge is correct, the overall hydrologic understanding could be poor. Further, these models are parameterized to varying extents, and tuning them requires gauge measurements of streamflow, which makes them less applicable and less reliable in ungauged regions (Beven & Cloke, 2012).

Recently, scientists are turning to remotely sensed data sets for alternative solutions to discharge estimation. Satellites or aerial sensors can offer a cost-effective approach to monitor rivers, from which hydraulic variables, including river width, stage, and slope, can be derived for rivers at global scales. For example, the National Aeronautics and Space Administration's upcoming (2022) Surface Water and Ocean Topography (SWOT) Satellite Mission will measure these variables for rivers wider than 100 m, globally (Biancamaria et al., 2016), and has served as a catalyst for remote sensing of discharge. These remotely sensed hydraulic variables have been used as proxies to estimate discharge. Previous studies have developed effective methods to use the remotely sensed data sets for discharge estimation, including establishing empirical relationships between in situ measured discharge and remote observations of channel morphological variables (e.g., width, slope, stage, or water surface area; Bjerklie, 2007; Bjerklie et al., 2003, 2005; Brakenridge et al., 2007; Durand et al., 2010, 2016, 2010; Pavelsky, 2014; Smith et al., 1996; Smith & Pavelsky, 2008; Tarpanelli et al., 2013) and integrating or assimilating remotely sensed data with hydrodynamic/hydraulic models (Brakenridge et al., 2012; Nathanson et al., 2012; Neal et al., 2009; Yoon et al., 2012). More recently, Gleason and Smith (2014) discovered the at-many-stations hydraulic geometry (AMHG), which identified relationships between at-a-station hydraulic geometry (AHG) parameters at multiple cross sections along a river. Compared to AHG, AMHG describes the relationships between hydraulic variables (e.g., width) and discharge in a more constrained way, which makes it an attractive approach to estimate discharge from solely space-based observations. Hagemann et al. (2017) developed this theory into a Bayesian mass conserved flow law inversion (McFLI) discharge algorithm, termed Bayesian AMHG-Manning (BAM), which estimates river discharge using AMHG and/or Manning's equation in a probabilistic manner. BAM integrates prior information of unknown quantities (e.g., hydraulic equation parameters and discharge statistics) from multiple sources including literature, in situ measurements, and reanalysis data sets. Ground-based prior information makes discharge estimation more accurate, yet it is possible to estimate discharge independent of ground-based measurements with BAM (e.g., using literature or reanalysis data), allowing reasonably accurate discharge estimations in completely ungauged settings.

In these previous studies, researchers used remotely sensed imagery from either governmental satellite missions (e.g., Landsat, Moderate Resolution Imaging Spectroradiometer [MODIS], and European Remote Sensing) or conventional commercial satellite platforms (e.g., RapidEye) to estimate discharge. Government-sponsored single-sensor satellite missions (e.g., Landsat 8) can provide high-quality radiometric data via rigorously calibrated and high-performing sensor systems (Irons et al., 2012). However, they generally suffer from the trade-off between spatial and temporal resolutions. For example, Landsat can provide high-quality imagery at moderately fine spatial resolution (i.e., 30 m for visible, near-infrared [NIR], and shortwave infrared bands) but at relatively coarse temporal resolution (i.e., 16-day revisit interval). This coarse temporal resolution limits the ability to capture short-lived phenomena like flood waves, which constrains its hydrologic potential in many applications. Sentinel-2 constitutes an advance toward resolving this limitation by deploying two identical sensors sharing the same orbit but phased 180° to each other, which enables it to monitor Earth surface dynamics at fine spatial resolution (10/20/60 m) globally with a higher repeat frequency (i.e., every 5 days). The integration of conventional commercial satellite platforms (e.g., RapidEye) can improve imagery resolution to 5 m daily, and this can be further enhanced with super high resolution platforms such as WorldView, but these high-resolution commercial sensor data are often prohibitively expensive, especially at global scales (Houborg et al., 2015; Houborg & McCabe, 2018). More recently, CubeSats, a class of small spacecraft with standard dimensions (i.e., 10 × 10 × 10 cm) and weight (i.e., ~1.3 kg; Puig-Suari et al., 2001), are emerging. The low cost of these small sensors makes it possible to launch several together. Planet (Planet Team, 2018), the largest commercial CubeSat operator, has deployed a constellation of over 300 CubeSats monitoring the global land surface at both high spatial (3–5 m) and

temporal (near daily) resolutions. CubeSats have been gaining serious interest from scientists (Aragon et al., 2018; Houborg & McCabe, 2018), especially for studies in Arctic regions (e.g., Cooley et al., 2019). However, the potential of CubeSat imagery for discharge estimation is currently untapped and unknown.

In this study, river flow width data retrieved from both government-sponsored satellites (Landsat and Sentinel-2) and CubeSats (Planet) are used to estimate discharge for 11 Arctic river reaches of various sizes (from <20 to >1 km wide). Sentinel-1 synthetic aperture radar (SAR) can also be used to estimate river widths with a high spatial resolution (i.e., 10 m) and without suffering constraints from clouds and shadows. However, this C-band SAR introduces new challenges to water classification, including salt-and-pepper noise and differences in reflectivity of water surfaces, especially where water interacts with sediment and vegetation. Thus, we have chosen to use exclusively optical satellites, so Sentinel-1 SAR is not included here. The objectives of this study are to answer, for the first time, the following questions: (1) *How does Planet CubeSat imagery perform for discharge estimation in Arctic regions?* (2) *How different are Planet data from traditional optical satellite data (here, Landsat and Sentinel-2) for the purpose of quantifying discharge dynamics?* (3) *Is it possible to improve our ability to estimate discharge for rivers across spatial scales and for ungauged basins by fusing Planet data with conventional optical satellite data?*

To address these questions, we use a developed space-based gauging station (Pavelsky, 2014) and BAM as the methods for discharge estimation. The reason for selecting BAM is that it can estimate discharge using solely remotely sensed river width data, and its capacity to incorporate multiple sources of prior information enables it to estimate discharge independently from in situ measurements, which expands its potential application in ungauged basins. BAM is a relatively new method for discharge quantification using remote sensing; thus, this study also provides the first opportunity to test its performance in quantifying discharge for Arctic rivers across spatial scales. To find potential options for improving our ability for discharge estimation, we test width data from both individual sensors (i.e., Planet, Landsat, and Sentinel-2) and a fusion of them (e.g., Planet + Landsat + Sentinel-2). Finally, multiple surface runoff reanalysis data sets are tested as BAM prior information with an attempt to provide an assessment of this method for potential applications in ungauged basins. This work also provides information about what is possible/necessary for discharge quantification at varying spatial/temporal scales with an eye toward SWOT and global discharge estimation.

## 2. Methods

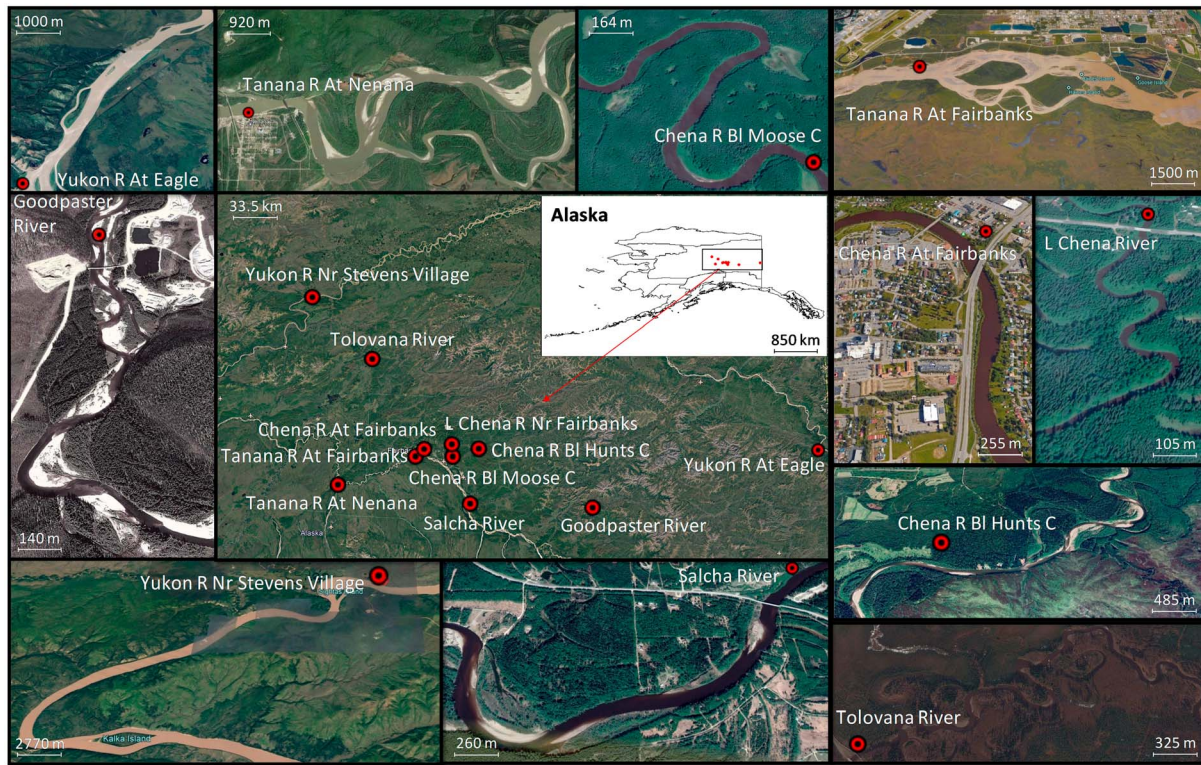
### 2.1. Study Region

Eleven Alaskan river reaches that mostly flow within the same basin were selected (Figure 1) to provide a range of river sizes, minimize differences in climate and hydrologic seasonality, and provide for algorithm use and evaluation. For example, the Yukon is the fifth largest Arctic river, with reach widths approaching 1,000 m, while the Lower Chena reach is as narrow as 16 m. The river flow dynamics of each are similarly impacted by Arctic-specific hydrologic processes such as ice breakup and permafrost dynamics. Each reach has a corresponding stream gauge to allow testing of algorithm performance with and without gauge data and for calibration.

### 2.2. Discharge and Runoff Data

In situ gauge observations of instantaneous discharge were acquired from the U.S. Geological Survey (<https://waterdata.usgs.gov/nwis>) for our 11 river reaches for the summer period (15 May through 15 September) during 2016–2018 (Table 1); 2016–2018 was a hydroclimatologically representative period based on analysis of gauge records (Figure S1 in the supporting information, SI). These subdaily gauge observations were averaged to a daily scale and then used for validation purposes. We also used these data to create a traditional space-based gauging station (e.g., Pavelsky, 2014) in each reach for comparison and to create the Bayesian equivalent of space-based gauging: using gauges to provide Bayesian priors to BAM. These comparisons allow us to assess CubeSat performance across a continuum from completely gauged to completely ungauged scenarios. To create ungauged scenarios, Bayesian priors for BAM available at the global scale are needed. We acquired reanalysis daily surface runoff from National Center for Environmental Prediction (NCEP; data provided by the National Oceanic and Atmospheric Administration-Earth System Research Laboratory Physical Sciences Division, Boulder, Colorado, from their Web site at <https://www.esrl.noaa.gov/psd/>), hourly surface runoff and baseflow from Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2; Global Modeling and Assimilation Office, 2015) and





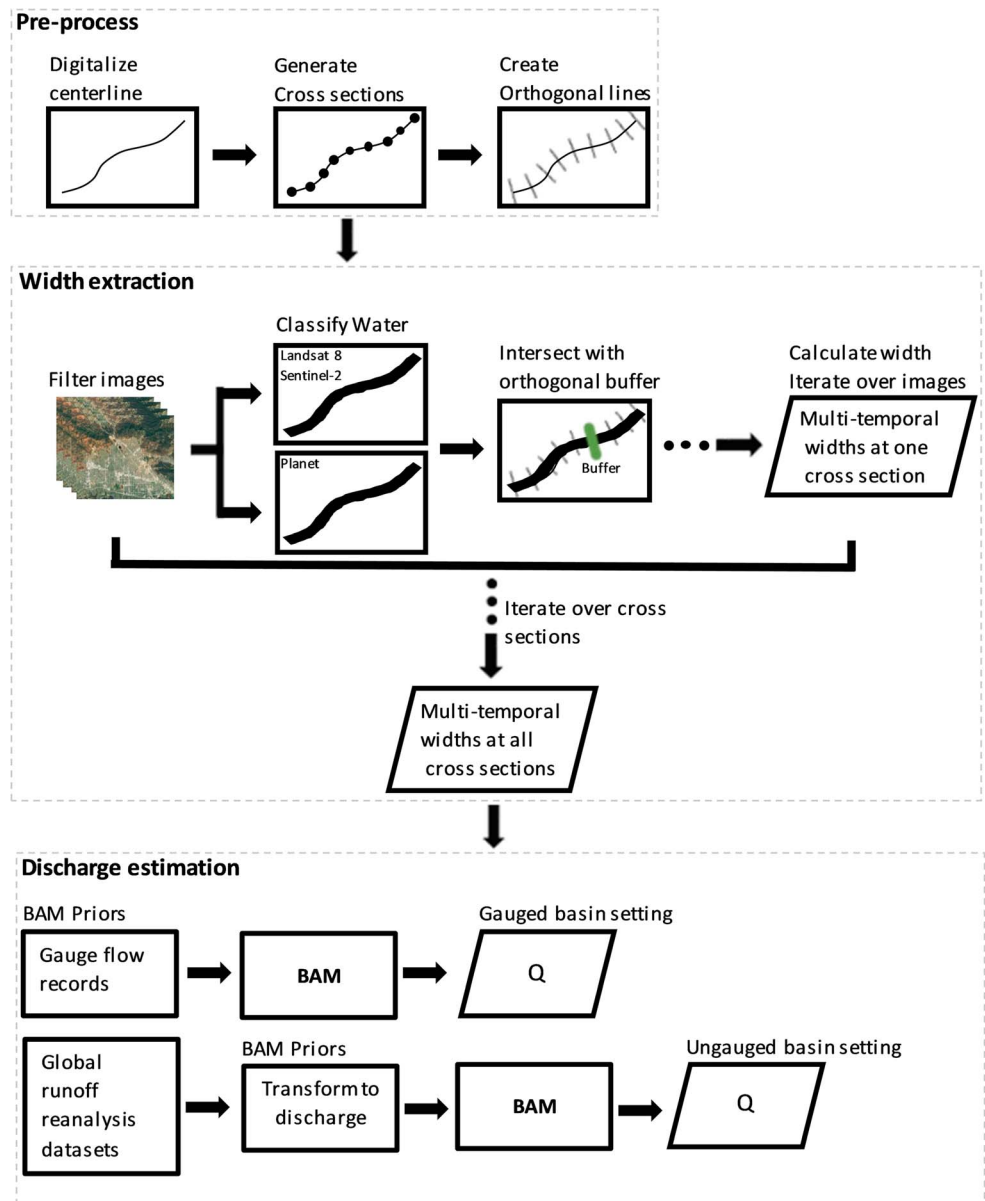
**Figure 1.** Study region: 11 river reaches and gauge locations. The figure in the middle shows an overview of the study region, gauges, and their locations in Alaska; the small figures around the edge show the details of these 11 river reaches (background images from Google Earth, 2018).

hourly surface and subsurface runoff from European Centre for Medium-Range Weather Forecasts Reanalysis Fifth generation (ERA5; Copernicus Climate Change Service, 2017) for our study period for this purpose. The runoff data (surface runoff for NCEP, surface runoff and baseflow for MERRA-2, and surface and subsurface runoff for ERA5 based on data availability) were converted to a discharge prior by aggregating runoff over the upstream drainage catchments for each of our reaches.

**Table 1**  
Gauge Locations and Descriptions

Site No.	Station name	Latitude/longitude	Drainage area (km <sup>2</sup> ) <sup>a</sup>	Mean flow (m <sup>3</sup> /s)	Mean width (m) <sup>b</sup>	Reach length (km)
15356000	Yukon R At Eagle	64°47'22"N/141°11'52"W	285,696	4267	537	8.9
15453500	Yukon R Nr Stevens Village	65°52'32"N/149°43'04"W	496,640	7123	630	22.4
15477740	Goodpaster R Nr Big Delta	64°27'02"N/144°56'32"W	1,731	37	34	1.4
15484000	Salcha R Nr Salchaket	64°28'17.5"N/146°55'41"W	5,632	116	60	2.3
15485500	Tanana R At Fairbanks	64°47'34"N/147°50'20"W	33,106	1236	611	9.6
15493400	Chena R Bl Hunts C Nr Two Rivers	64°51'36"N/146°48'12"W	3,430	75	46	2.9
15493700	Chena R Bl Moose C Dam	64°48'03"N/147°13'40"W	3,738	73	31	2.2
15511000	L Chena R Nr Fairbanks	64°53'10"N/147°14'50"W	952	13	16	0.7
15514000	Chena R At Fairbanks	64°50'45"N/147°42'04"W	5,094	92	52	1.4
15515500	Tanana R At Nenana	64°33'53.8"N/149°05'38.4"W	65,434	1495	399	5.6
15519100	Tolovana R Bl Rosebud C Nr Livengood	65°27'55"N/148°37'43"W	622	7	29	0.7

<sup>a</sup>Drainage area for Gauge 15485500 (not available from USGS) was obtained through watershed delineation, others were obtained from USGS. USGS = U.S. Geological Survey. <sup>b</sup>Widths for 15356000, 15453500, 15485500, and 15515500 are the average channel widths at all cross sections retrieved from usable Landsat 8, Sentinel-2, and Planet imagery during 15 May to 15 September for years 2016–2018; widths for 15484000, 15493400, and 15514000 are from usable Sentinel-2 and Planet imagery; widths for 15477740, 15511000, 15493700, and 15519100 are solely from usable Planet imagery.



**Figure 2.** Flowchart of methods for width retrieval and discharge estimation. BAM = Bayesian at-many-stations hydraulic geometry-Manning algorithm.

### 2.3. Remotely Sensed Data and Cross-Sectional Flow Width Extraction

Prior to automated width extraction, a preprocess defining the cross sections from which to extract widths was conducted (Figure 2). We first manually digitized the channel centerline for each river reach based on a water occurrence map (Pekel et al., 2016) and maps on Google Earth Pro (Google Earth, 2018). The channel centerlines could be obtained from other products, for example, the Global River Widths from Landsat Database (Allen & Pavelsky, 2018); however, their accuracy for narrow rivers (i.e., width < 90 m) can be low. Then cross-section points were created along the centerline at 30 meter intervals and the orthogonal angle at each point was calculated (Yang et al., 2019; see Figure 2). Finally, channel orthogonal lines were constructed based on the centerline position, orthogonal angles, and orthogonal line length. In this study, orthogonal line lengths were defined based on a half-width of  $0.75 \times$  maximum channel width from maps on Google Earth Pro (Google Earth, 2018). The ratio 0.75 was selected to ensure the entire channel width is covered and to reduce the impacts of the imperfection of the centerline position (the digitized

**Table 2**  
Summary of Remotely Sensed Data Sets Used in This Study

Satellites	Landsat 8	Sentinel-2	Planet
Spatial resolution	30 m	10/20/60 m	3 m
Temporal resolution	16 days	5 days	daily
Number of bands	11	13	4
Data availability (start time)	4/11/2013	6/23/2015	5/15/2016
Products used in this study	Landsat 8 Surface Reflectance Tier 1	Sentinel-2 MSI: MultiSpectral Instrument, Level-1C	Planet Scope Ortho analytic (PSOrthoTile analytic_dn)
Provider	USGS	European Union/ESA/Copernicus	Image courtesy of Planet Labs, Inc.

Note. Dates are formatted as M/DD/YYYY. USGS = U.S. Geological Survey.

centerline may not be located exactly at the middle of river channel). Further scrutiny of the extracted widths revealed that this orthogonal length was appropriate to cover the full range of temporally varying widths. We used a uniform orthogonal line length within a reach.

After the channel cross-section orthogonal lines were constructed, they were uploaded to Google Earth Engine as the input of a width extraction algorithm (Yang et al., 2019). In this algorithm, the water was classified following a spectral-based approach with reported accuracy of 97% (Zou et al., 2018): [(mNDWI > EVI or mNDWI > NDVI) and (EVI < 0.1)] (defined in equations (1)–(3)); image pixels that meet these criteria were classified as water, and others were classified as nonwater. For Landsat, the surface reflectance (SR) images in the Landsat 8 SR Tier 1 collection were used for width extraction (Table 2). For Sentinel-2, the SR images are unavailable for our study region during 2016–2018, so we used the Sentinel-2 top-of-atmosphere (TOA) Level-1C images for width extraction. To validate the applicability of the water classification method in Zou et al. (2018); developed for SR images) to the TOA images, we compared the widths extracted from Landsat 8 TOA and SR images for River 15515500, and we found a significant ( $R^2 = 0.98$ ) linear relationship with a slope close to unity (1.03) between the widths extracted from these two collections (Figure S2). We argue it is thus reasonable to apply this water classification method to the Sentinel-2 Level-1C images.

We used only clear-sky images, defined as CLOUD\_COVER <25 (%) for Landsat 8 (USGS, 2019) and CLOUDY\_PIXEL\_PERCENTAGE < 25 for Sentinel-2 (ESA, 2015). The cross sections that were affected by clouds were identified based on the cloud classification information and then were excluded. The classified water mask was intersected with the orthogonal line buffer (30 m) at each cross section to calculate river flow width (Figures 2 and 3). The river width at each cross section ( $W_m$ ) can be obtained by rearranging equation (4).

$$\text{mNDWI} = \frac{\text{band}_{\text{green}} - \text{band}_{\text{swir1}}}{\text{band}_{\text{green}} + \text{band}_{\text{swir1}}} \quad (1)$$

$$\text{EVI} = \frac{2.5 \times (\text{band}_{\text{nir}} - \text{band}_{\text{red}})}{1 + \text{band}_{\text{nir}} + 6 \times \text{band}_{\text{red}} - 7.5 \times \text{band}_{\text{blue}}} \quad (2)$$

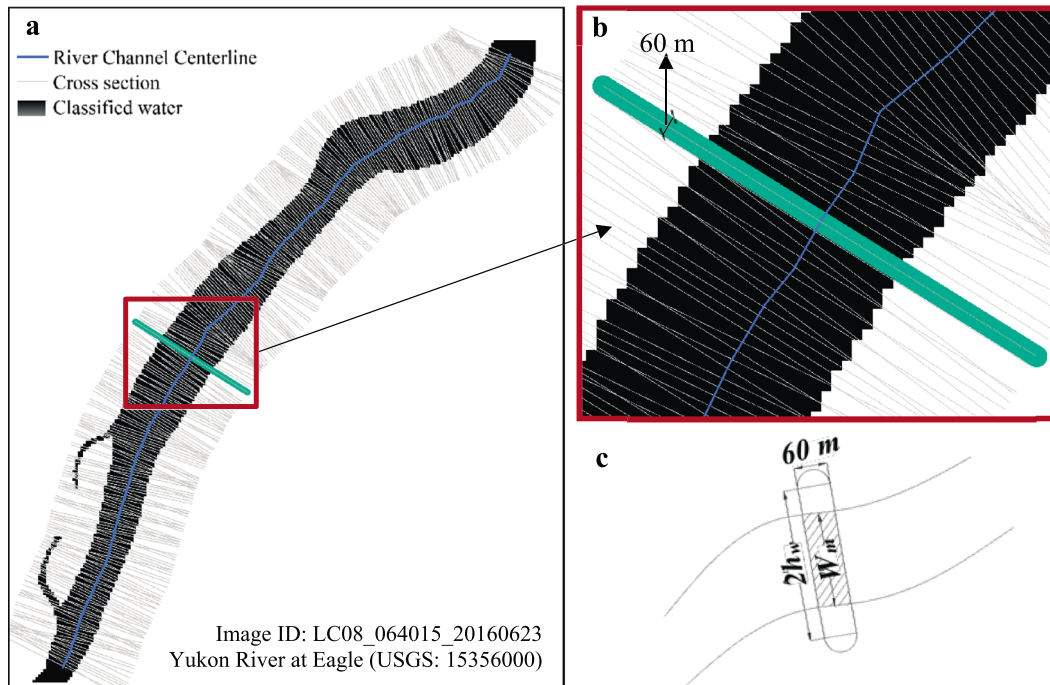
$$\text{NDVI} = \frac{\text{band}_{\text{nir}} - \text{band}_{\text{red}}}{\text{band}_{\text{nir}} + \text{band}_{\text{red}}} \quad (3)$$

where mNDWI is the modified normalized difference water index (Xu, 2006); EVI and NDVI are enhanced vegetation index and normalized difference vegetation index (Huete et al., 1999), respectively;  $\text{band}_{\text{green}}$ ,  $\text{band}_{\text{swir1}}$ ,  $\text{band}_{\text{nir}}$ ,  $\text{band}_{\text{red}}$ , and  $\text{band}_{\text{blue}}$  are the pixel values at bands Green (Band 3 for both Landsat 8 and Sentinel-2), shortwave infrared 1 (Band 6 for Landsat 8 and Band 11 for Sentinel-2), near infrared (NIR; Band 5 for Landsat 8 and Band 8 for Sentinel-2), Red (Band 4 for both Landsat 8 and Sentinel-2), and Blue (Band 2 for both Landsat 8 and Sentinel-2) of satellite images, respectively.

$$\gamma = \frac{W_m \times 60}{2 \times h_w \times 60 + \pi \times 30^2} \quad (4)$$

where  $\gamma$  is the ratio of overlapped water mask area to buffer area;  $W_m$  is the desired quantity, river flow width, m;  $h_w$  is the half length of orthogonal lines, which is predetermined in Step (3) of preprocessing.





**Figure 3.** (a) Cross-section alignment for Yukon River at Eagle, (b) example of cross-section buffer used for width extraction, and (c) diagram for width calculation at each cross section. USGS = U.S. Geological Survey.

Planet imagery is not available in the Google Earth Engine platform. Widths from Planet imagery were extracted following Cooley et al. (2019). Readers are referred to their paper for reference, but in brief, we used the following process. (1) We extracted planet images only in areas of interest that were manually defined to encompass the river and near-bank areas: This limited the total area downloaded from Planet and made classification easier and could be automated in the future using global water masks, for example, Pekel et al. (2016). (2) We calculated NDWI (equation (5)) using Bands 2 and 4. We calculated the Otsu threshold of this NDWI, producing a binary water/nonwater map based on the assumption that water and not-water are the dominant features in the NDWI histogram. (3) We then dilated this Otsu threshold by an 11-pixel kernel and applied an econometrics breakpoint algorithm on this buffered area (using the “Strucchange” R package). This is a more sophisticated automated thresholding algorithm that identifies multiple histogram peaks. (4) Finally, we selected the highest single NDWI peak resulting from this breakpoint analysis in our dilated Otsu polygon as the true water signal; this represented the final water mask. This multistep process has been shown to improve upon simple binary thresholding for Planet images (Cooley et al., 2019). Widths were extracted based on the same manually created cross sections described previously for Landsat and Sentinel-2.

$$NDWI = \frac{band_{green} - band_{nir}}{band_{green} + band_{nir}} \quad (5)$$

where NDWI is normalized difference water index (McFeeters, 1996);  $band_{green}$  and  $band_{nir}$  are Bands 2 (Green) and 4 (NIR) of Planet imagery, respectively.

## 2.4. Discharge Estimation

### 2.4.1. The BAM Algorithm

BAM estimates river discharge by using either AMHG or Manning's equation in a probabilistic manner (Hagemann et al., 2017). However, observations of water surface elevation and width are needed to drive the Manning variant, and its use is beyond the scope of this study where the focus is optical data alone. These joint observations (of height and width) are also only available where ocean altimeters and optical

**Table 3**  
Priors on Parameters in BAM-AMHG

Parameters	Mean	Variance	Upper bound	Lower bound
$\ln(W_c)$	$\ln\left(\frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n W_{it}\right)$ ; $n$ is number of cross sections within a reach; $T$ is the time steps of width records	$\ln(\text{var}(W_{11,21,\dots,nT}))$	$\ln(\min(W_{11,21,\dots,nT}))$	$\ln(\max(W_{11,21,\dots,nT}))$
$\ln(Q_c)$	$\ln\left(\mu_q / \sqrt{1 + \frac{\sigma_q^2}{\mu_q^2}}\right)$ ; $\mu_q = \text{mean}(Q_r)$ ; $\sigma_q^2 = \text{var}(Q_r)$ ; $Q_r$ is river flow time series obtained from gauge records or reanalysis data sets	$\ln\left(1 + \frac{\sigma_q^2}{\mu_q^2}\right)$	$(1-\alpha) \times \ln\left(\mu_q / \sqrt{1 + \frac{\sigma_q^2}{\mu_q^2}}\right)$ $\alpha$ is 0.1 in this study	$(1+\alpha) \times \ln\left(\mu_q / \sqrt{1 + \frac{\sigma_q^2}{\mu_q^2}}\right)$
$\ln(Q)$	$\ln\left(\mu_q / \sqrt{1 + \frac{\sigma_q^2}{\mu_q^2}}\right)$	$\ln\left(1 + \frac{\sigma_q^2}{\mu_q^2}\right)$	$\min(\ln(Q_r))$	$\max(\ln(Q_r))$
$\ln(b_i)$	$0.02161 + 0.4578 \times \text{SD}(\log W_i)$ $\ln(\text{AHG width-discharge exponent } b)$	$\ln(0.3^2)$	$\ln(0.4)$	$\ln(1.0)$

Note. BAM = Bayesian at-many-stations hydraulic geometry-Manning algorithm; AMHG = at-many-stations hydraulic geometry; AHG = at-a-station hydraulic geometry.

data align in space and time and are difficult to find for small rivers. We therefore used the AMHG physics in this study. The BAM-AMHG is formulated as equation (6) (Hagemann et al., 2017).

$$\log W_i = b_i(\log Q - \log Q_c) + \log W_c + \epsilon \quad (6)$$

where  $W_i$  is river width at cross section  $i$ , (m);  $b_i$  is the AHG width-discharge exponent at cross section  $i$ .  $Q$  is discharge ( $\text{m}^3/\text{s}$ );  $Q_c$  and  $W_c$  are AMHG global parameters for a reach;  $\epsilon$  is an error term that we set to 0.22 as suggested by Hagemann et al. (2017). In BAM-AMHG,  $W_i$  is a known quantity, obtained from satellite imagery in this study;  $b_i$ ,  $Q$ ,  $Q_c$ , and  $W_c$  are unknown parameters that need to be sampled from a predefined a priori distribution. Following suggestions from Hagemann et al. (2017), we assumed these unknown parameters follow a truncated lognormal distribution. Statistics of these a priori distributions were determined (Table 3) based on their original definitions (Gleason & Wang, 2015) and suggestions from Hagemann et al. (2017). The prior information for  $W_c$  was determined by assuming it follows the same distribution as the observed widths extracted from satellite imagery.  $Q_c$  and  $Q$  were assumed to follow the same prior distribution defined by historical flow statistics, but with different boundaries.  $Q$  was assumed to have the same boundary as historical flow, but  $Q_c$  has a narrower boundary based on its definition and suggestions from Gleason and Wang (2015). The parameter  $\alpha$  (in Table 3) used to define the  $Q_c$  boundaries was selected based on an optimization process, during which an ensemble of  $\alpha$  values were tested and the one with the best BAM performance (assessed by normalized root-mean-square error, nRMSE) was selected. In this study, the best BAM performance was achieved when  $\alpha = 0.1$ , although the process was relatively insensitive to this value, and this optimization is not necessary in ungauged settings. Parameter  $b$  was defined using two approaches in this study: The empirical approach suggested by Hagemann et al. (2017) and an AHG- $b$  approach, which transforms the AHG width-discharge exponent at each cross section to the mean of parameter  $b$  in BAM. A standard deviation of 0.3 for  $b$  was used in this study, based on the optimization process described above. Similarly, the BAM results were relatively insensitive to the values of  $b$  standard deviation, and these results should hold in ungauged regions. We used 0.4 and 1.0 for the lower and upper boundaries for  $b$ , respectively, based on the AMHG definition (Gleason & Wang, 2015).

#### 2.4.2. BAM Versus Rating Curves in Gauged Basins

BAM can operate in ungauged cases, but its accuracy improves when prior data are available. We do have gauge data available in our study region, but it is not obvious that BAM should be used when these data are available: why not use space-based gauging in this case? Here, we refer to “spaced-based gauging” as the AHG-based width-discharge rating curve method (Pavelsky, 2014). This method also only needs width (in addition to gauge data) as input, which is comparable to BAM, and it is well developed and more commonly used than BAM. We followed Pavelsky (2014) to combine multiple width-discharge rating curves within a reach for discharge estimation. In brief, we used the in situ gauge discharge and satellite-derived widths at each cross section to create a new rating curve and then merged the results to arrive at a gauging station that could operate if the gauge were removed and if river planform/geometry remained stable. To



optimize the results, cross sections were filtered based on their AHG strength, evaluated by the  $p$  value of the linear regression for  $\log w \sim \log Q$ . Twenty cross sections with the best AHG strength were used for discharge estimation. Twenty was selected as the sample number based on the number of cross sections in the smallest river reach. In the rating curve method, the medians of the estimated discharge at these 20 cross sections were used as the river discharge in each reach. BAM uses the widths from these same 20 cross sections and gauge streamflow data as priors. The performance of BAM will be evaluated by comparing its results with those from the rating curve method.

#### 2.4.3. BAM in Ungauged Basins

To assess the potential of the method proposed in this study for estimating discharge in ungauged basins, we also tested BAM performance when using reanalysis runoff data sets mentioned in section 2.2 as BAM priors. In ungauged basins, in situ measurements are unavailable, so it is impossible to detect the best AHG cross sections as in section 2.4.2. Two experiments were conducted: (1) using width data from 20 cross sections with highest width variabilities (we call this approach *most-variable-width* in the following text) and (2) using an ensemble of 50 combinations of 20 randomly selected cross sections, where the final discharge is the median of these cross sectional discharges (we call this approach *ensemble-median* in the following text). These tests allow us to test the hypothesis that the most-variable-width cross sections will have more accurate discharge estimation than an ensemble-median approach. We removed width outliers to reduce the effects of bias resulting from the width extraction process (e.g., water classification uncertainties) on the BAM results. Width values three standard deviations greater/less than mean width at each cross section were identified as outliers and were excluded for  $Q$  estimation. Consistent with the assumption that the river widths follow a lognormal distribution (in section 2.4.1), this outlier detection process was conducted on the logarithm of width time series. This filtering ensured that commission error in the classification did not introduce overly large/small width values.

### 3. Results

#### 3.1. Satellite Width Measurements

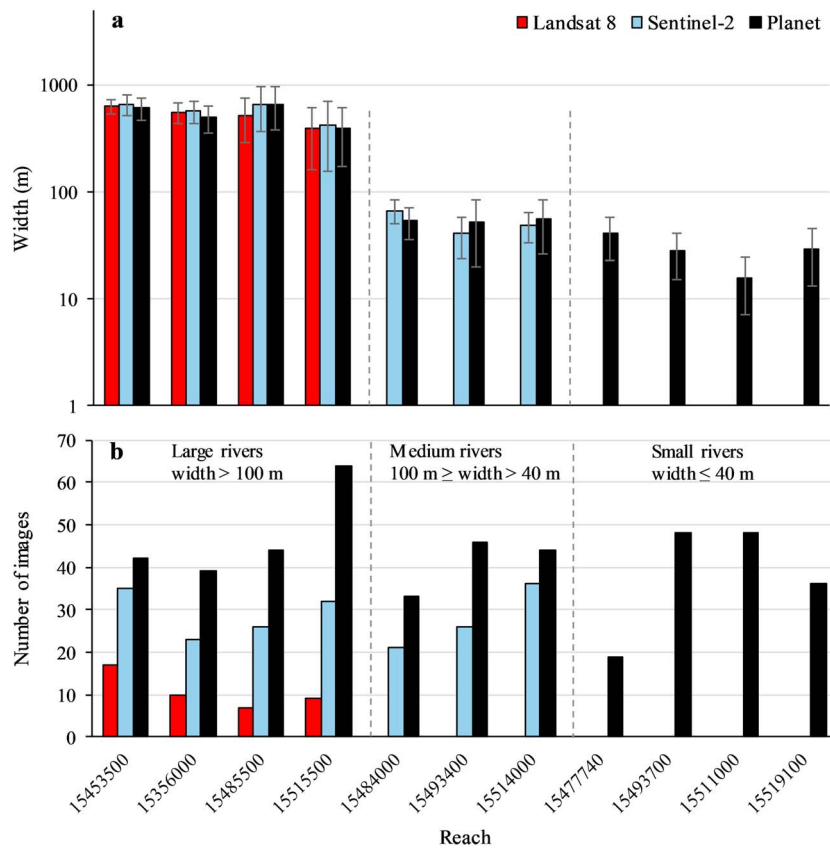
The retrieved widths from three satellites are comparable, with a difference of  $<140$  m for large rivers and  $<15$  m for medium rivers (Figure 4a). Given the spatial resolution of Landsat 8 (i.e., 30 m), only data for large rivers (width  $> 100$  m) are evaluated. With a higher spatial resolution, Sentinel-2 images are used for both large and medium rivers (width  $> 40$  m). Planet images are used for all rivers, considering its fine resolution ( $\sim 3$  m). The number of width measurements is also impacted by the temporal resolution of each satellite (Figure 4b). There are about  $80 \pm 20$  Landsat 8 images available for the large river reaches during the study period, and only 10–20% are cloud free. Sentinel-2 has more usable images due to its higher temporal resolution (i.e., 5 days) as compared to Landsat 8:  $180 \pm 60$  images with 15–25% of them cloud free. Thus, the numbers of widths retrieved from Sentinel-2 for large rivers are greater than those from Landsat 8 by a factor of 3–4 (Figure 4b). With the highest temporal resolution (near daily), Planet has the most usable images for all rivers, about a factor of 4 more than Landsat 8 and twice as many as Sentinel-2 on average, showing its temporal advantages.

#### 3.2. Discharge Estimation

##### 3.2.1. Estimation With Gauge Prior Information

The performance of BAM as compared with the rating curve method is shown in Figure 5, and we show full hydrographs for only one river for each satellite data category for conciseness (Figures 6 and 7). Results for other rivers can be found in the SI (Figures S3–S10). We also report only normalized RMSE (residuals divided by the mean: nRMSE, %) in this paper, and other error metrics, including rRMSE (%), rBias (%), and NSE, can be found in the SI (Tables S1–S3). The definitions of these error metrics can also be found in equations (S1)–(S4) in the SI. For BAM, we tested both  $b$  algorithms as shown in Table 3. The results for these two algorithms are very similar, so only results for the empirical- $b$  algorithm are shown and discussed here. The error metrics for results of BAM with the AHG- $b$  algorithm can be found in Table S1 in the SI.

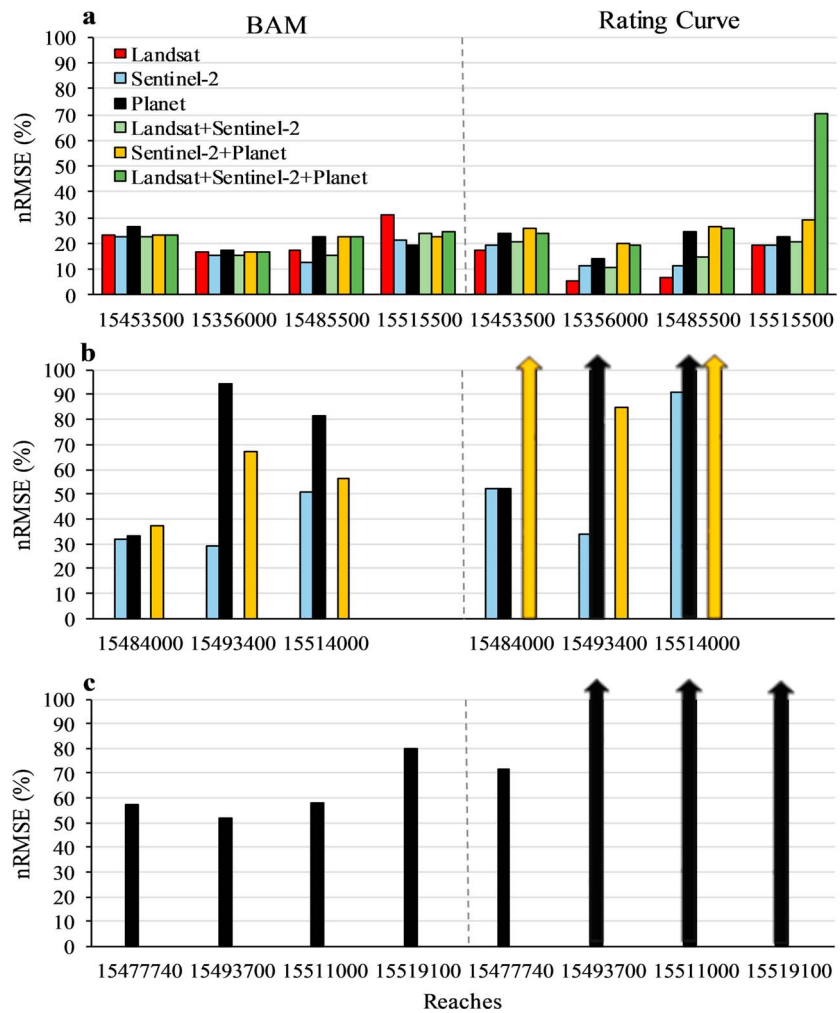
For large rivers (width  $> 100$  m), the rating curve method performs better than BAM when using both individual and fused governmental satellite data (i.e., Landsat, Sentinel-2, and Landsat+Sentinel-2). This is probably because the high-quality satellite imagery (Landsat and Sentinel-2) can produce more accurate



**Figure 4.** Comparison of retrieved river flow width (a) and number of usable images (b) from three data sets for 11 river reaches in the summer period (15 May to 15 September) during 2016–2018. Error bars represent standard deviation of retrieved widths at all cross sections within a reach in all available dates. Considering the spatial resolutions, Landsat 8 imagery were only used for large rivers and Sentinel-2 imagery were only used for large and medium rivers. The top panel (a) indicates that widths are similar, but not identical, from all three satellites, while the temporal advantages of Planet are clear in the bottom panel (b).

width data for large rivers, which help establish strong width-discharge AHG, making the rating curve method more accurate for discharge estimation than BAM. In contrast, BAM performs better than the rating curve method when using fusion of government-based and CubeSat data. When using BAM, data fusion improves discharge estimation for all large rivers, although the magnitude of this effect is small (less than 7%). This is because the rating curve method is more sensitive to noise and outliers in width data than BAM, and BAM produces more “flat” hydrographs that mute dynamics that give better error metrics (e.g., Figures S4 and S5).

For medium rivers, BAM performs better than the rating curve method for both individual and fused data sets. When using BAM, fusion achieves a better performance (average 26% reduction in nRMSE) for most medium rivers (except for River 15484000) with nRMSE in a range of 37–67% (Figures 5 and 7). For small rivers, similarly, BAM is more stable than the rating curve method (Figure 7), with nRMSE ranging between 52% and 80% for BAM and 72% to >100% for rating curve method (Figure 5). Errors are relatively high for small rivers for both BAM and rating curve method compared with those for large and medium rivers. This may suggest that the individual CubeSat data are not favorable for estimating discharge for small rivers, likely due to the lower quality of CubeSat images as compared with those from flagship government-sponsored satellites. In addition, the numbers of cross sections covering small river reaches, especially for Rivers 15519100 (24 cross sections) and 15511000 (23 cross sections), are much smaller than those covering large and medium rivers. Available cross sections for small rivers exhibit weak width-discharge AHG (i.e., only 20–50% AHG curves are significant at a 10% level), which may also contribute to the poor performance of both BAM and rating curve method in these rivers.

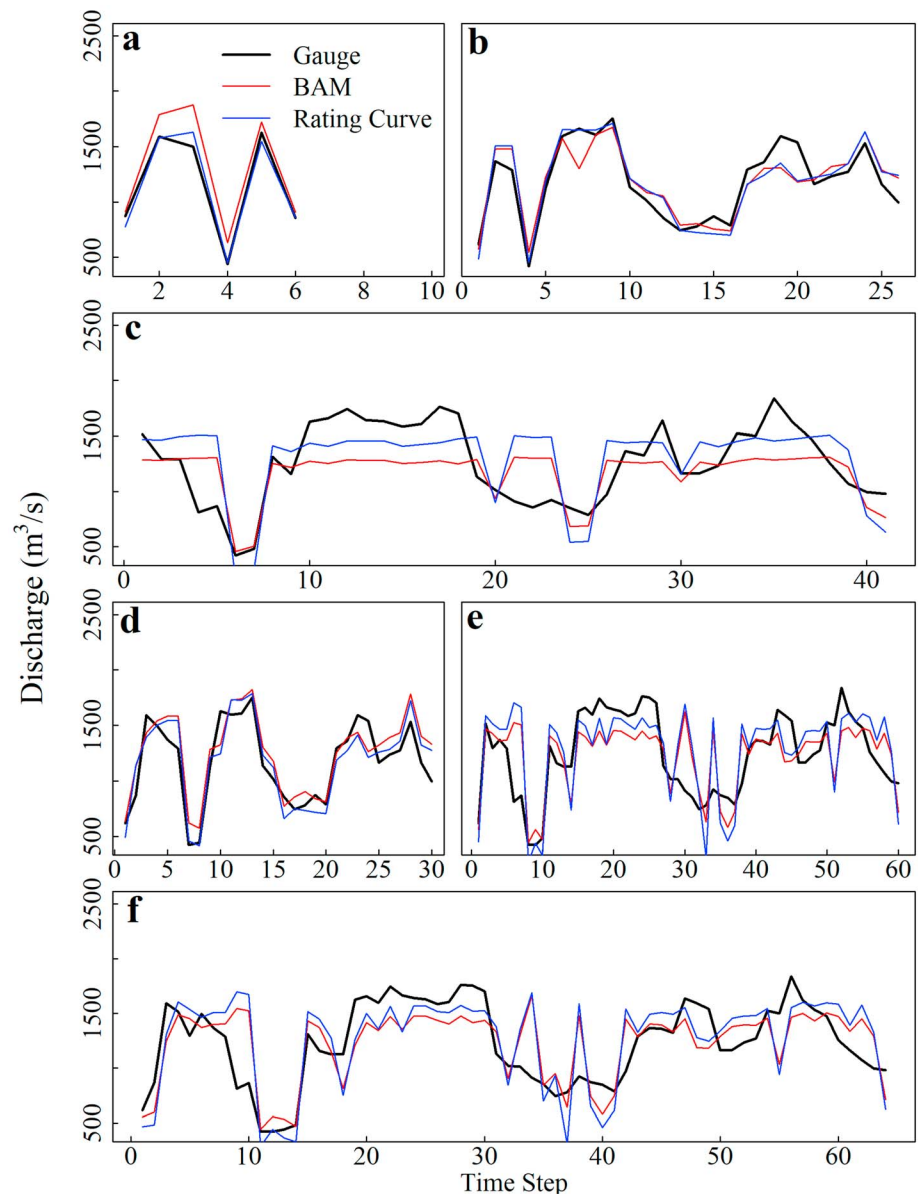


**Figure 5.** Error (nRMSE) of estimated discharge for large rivers (a), medium rivers (b), and small rivers (c) by BAM (left part) and Rating Curve method (right part). Colored bars represent different width data sets, which are a factorial of available data for this study. Per Figure 4, not all data sets are appropriate for all rivers. We see here that BAM outperforms the rating curve method for medium and small rivers and performs well for large rivers. Errors are truncated at 100% for display purposes (arrow points indicate errors higher than 100%), and values for these are given in the supporting information (Table S1). BAM = Bayesian at-many-stations hydraulic geometry-Manning algorithm; nRMSE = normalized root-mean-square error.

Large overestimations in discharge occur when Planet data are used individually or fused as inputs for rating curve method (Figures 7b–7d, S3f, S6b and S6c, S7b and S7c, S8, and S9), suggesting errors in water classification. Although Landsat and Sentinel-2 have coarser spatial resolutions (30/10/20 m) as compared to Planet (3 m), the widths from these governmental satellites have lower uncertainties than those from the CubeSat. This is probably due to the higher quality of Landsat 8 and Sentinel-2 imagery compared to Planet imagery: Planet’s four-band sensors struggle to differentiate turbid water from sand bars or dry channels.

### 3.2.2. Estimation in Ungauged Basins

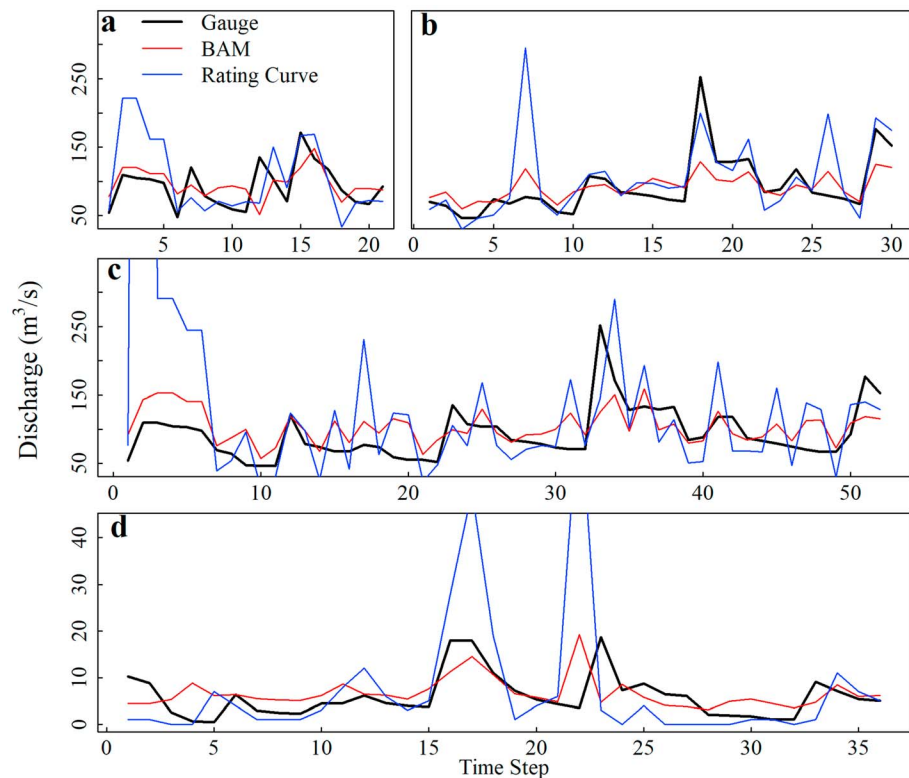
The previous section provides examples of how BAM might perform in cases where gauge data are present. BAM’s great advantage, however, lies in its ability to function in completely ungauged basins without relying on assumptions of hydrologic regionalization or transferability. Here, we use globally available reanalysis runoff data sets as BAM prior information: These estimates are available for all rivers on Earth. As a Bayesian technique, BAM improves with improving prior information, so the estimation performance relies in part on how well the reanalysis data represent the magnitude and dynamics of real river flows. Figure 8



**Figure 6.** Simulated discharge hydrographs for large River 15485500 (Tanana River at Fairbanks) using BAM and rating curve method with width data from individual satellites and their fusions: (a) Landsat 8; (b) Sentinel-2; (c) Planet; (d) Landsat 8 + Sentinel-2; (e) Sentinel-2 + Planet; (f) Landsat 8 + Sentinel-2 + Planet. The different  $x$  axis scales reflect the number of available images, and hydrographs are shown as connected for ease of interpretation. In reality, these time steps cover three summers, so Landsat would appear as seven points across a 3-year time period in panel (a). BAM = Bayesian at-many-stations hydraulic geometry-Manning algorithm.

shows the results, which are generated completely without in situ data and using widths same as in section 3.2.1. Among the three reanalysis data sets tested in this study, NCEP performed best, because it has the most similar statistics (mean and variance) to those of the gauge records (Figure 8). MERRA-2 runoff values are relatively low, leading to underestimated discharge for most rivers (with nRMSE ranging between 50% and 95%; Figure 8b). In contrast, the variation of discharge derived based on ERA5 runoff is much higher than that of gauge measurements, which results in overestimated discharge for most rivers, with nRMSE > 100% for more than half of the studied rivers (Figure 8c). These results also indicate that better performance of BAM discharge estimation can be expected when more realistic runoff reanalysis data are available in the future.





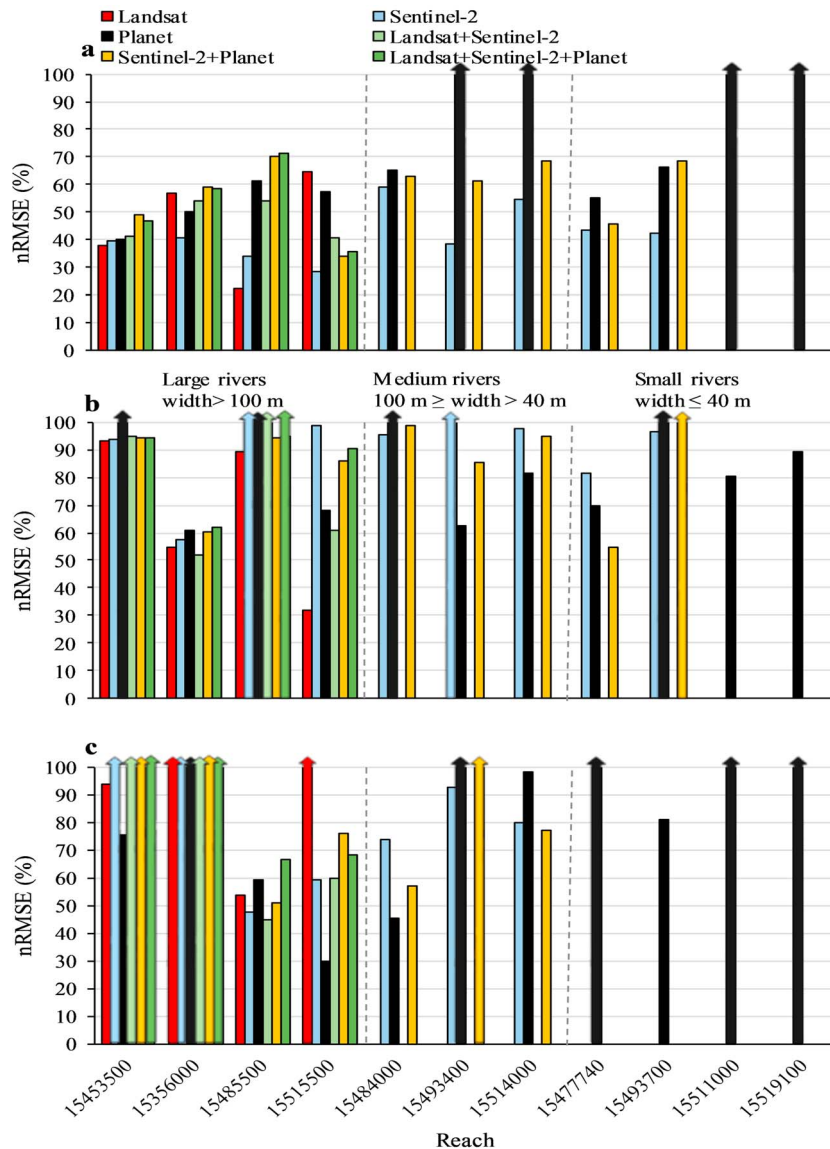
**Figure 7.** (a–c) Simulated discharge for medium River 15484000 (Salcha River near Salchaket) using BAM and rating curve method with width data from individual satellites and their fusions: (a) Sentinel-2; (b) Planet; (c) Sentinel-2 + Planet. (d) Simulated discharge for small River 15519100 (Tolovana River) using BAM and rating curve method with width data from Planet. BAM = Bayesian at-many-stations hydraulic geometry-Manning algorithm.

Using the best (NCEP) forcing, BAM can estimate discharge with an nRMSE in a range of 22–71% in large rivers (Figure 8a). For medium rivers, Sentinel-2 performs better than Planet, with 39–59% nRMSE. However, the fusion of Sentinel-2 and Planet data can achieve a more stable performance across medium rivers. For small rivers, BAM does not perform well, probably due to both the poor width extractions and worsened prior information.

Without gauge data to detect the AHG strength for ungauged basins, we need a different way of choosing which 20 cross sections should be used to run BAM in this case. To provide suggestions on how cross sections should be selected when no gauge data are available, BAM was forced with both width from the 20 most width-variable and an ensemble of randomly selected cross sections, as mentioned in section 2.4.3. The results (only results for NCEP are shown here considering its relatively better performance than MERRA-2 and ERA5; Figure 9) show that the ensemble-median approach performs better than the most-variable approach in regard to the accuracy of discharge estimation, although the difference is not statistically significant ( $p$  value = 0.75 for Mann-Whitney test) probably because of the small number of samples. In addition, the width-most-variable approach is sensitive to the width uncertainties. We found that, without the outlier detection process, the results for the width-most-variable approach are much worse than those shown in Figure 9a. Therefore, the ensemble-median approach is likely a better way to select cross sections for estimating discharge with BAM for ungauged basins.

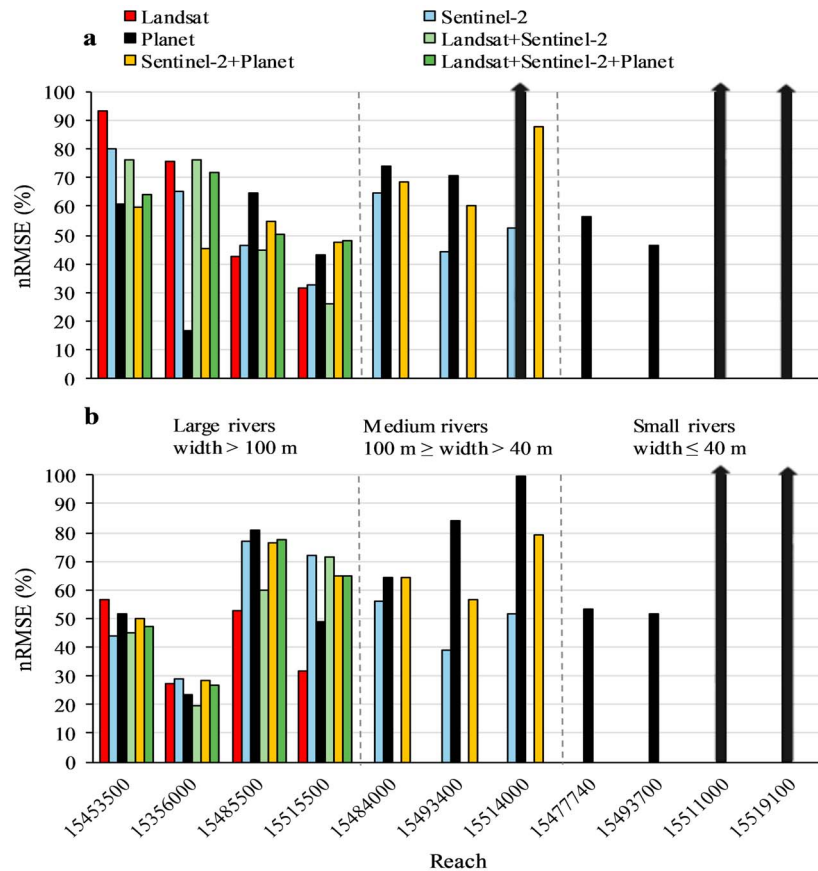
#### 4. Discussion

Since discharge estimates rely on satellite-derived widths, disparities in spatial resolution, band classification techniques, temporal resolution, and spectral resolution are important. Landsat 8 can provide very accurate flow width information for large rivers, but its temporal data density is very low (Figures 4 and 5). Sentinel-2 can provide more usable images at finer scales, but the quality of the extracted widths is not as good as that of



**Figure 8.** Error metrics (nRMSE) of BAM discharge estimates using reanalysis runoff data sets from (a) NCEP, (b) MERRA-2, and (c) ERA5 as priors. These represent completely ungauged scenarios, and we can see how BAM performance changes with changing prior information. NCEP-driven performance is good, suggesting that ungauged Arctic rivers similar to those in the study area might be estimated with acceptable accuracy for these ungauged cases. BAM = Bayesian at-many-stations hydraulic geometry-Manning algorithm; NCEP = National Center for Environmental Prediction; MERRA-2 = Modern-Era Retrospective analysis for Research and Applications version 2; ERA5 = European Centre for Medium-Range Weather Forecasts Reanalysis Fifth generation; nRMSE = normalized root-mean-square error.

Landsat (Figures 5 and 6), likely due to cloud classification issues that introduces bias to water masks (Coluzzi et al., 2018). In addition, the water classification method in Zou et al. (2018) is calibrated based on the performance for Landsat SR images, therefore applying this method to Sentinel-2 TOA images may also result in biases in extracted widths. CubeSat satellites produce the most useable images (about 2–6 times greater than a single government-based satellite data set) and record river flow information for all rivers. This lack of radiometric calibration for CubeSat data is partially overcome by our use of the method in Cooley et al. (2019), where thresholding to detect water is performed on a relative basis, rather than on an absolute threshold as in NDWI. This allows each image to be optimally classified, regardless of its radiometric calibration versus other images. However, the performance of individual CubeSat data for discharge estimation is still not as good as governmental satellites (Figures 5–7), which is probably



**Figure 9.** Error metrics (nRMSE) of BAM with prior information from NCEP reanalysis runoff data set for (a) BAM discharge estimates using widths from 20 most variable cross sections, and (b) ensemble median of BAM discharge estimates using widths from 50 combinations of 20 randomly selected cross sections. BAM = Bayesian at-many-stations hydraulic geometry-Manning algorithm; nRMSE = normalized root-mean-square error; NCEP = National Center for Environmental Prediction.

because of water classification issues (i.e., only four bands are available for water classification; Houborg & McCabe, 2018). For example, for the Tanana River (U.S. Geological Survey No. 15485500), the discharge error for Planet is especially higher than those for Landsat and Sentinel-2 (Figure 5), probably because Planet CubeSats do not have a midinfrared band, which makes it challenging to differentiate the sediment-laden water in Tanana River from the sand bars using visible and near-infrared (VNIR) bands of Planet imagery alone.

When these three satellites are fused, noise and uncertainty can mask the dynamic signals of width variation, which leads to worse discharge estimates from the rating curve method than with single data sets (higher nRMSE for fusions than individuals; Figure 5). This is because the rating curve method is a deterministic approach, and it is more sensitive to outliers and data noise than BAM. In contrast, BAM-AMHG tends to produce more stable discharge estimates across scales after fusion: Hydrographs tend to be more similar to one another. This is mainly because in BAM, the discharge estimates are retrieved based on the statistics of Bayesian inference from a prior distribution, which makes it “statistically defensible” (Hagemann et al., 2017). This suggests that the fusion of government-sponsored and CubeSat satellites data sets, integrated with BAM, is a possible approach to improve discharge estimation for large and medium rivers by increasing data density (about 2–6 times greater than a single governmental satellite data set), which is particularly important to capture short-lived phenomena like flood waves. The upcoming SWOT mission can provide high quality hydraulic variables, which will enhance global discharge estimation significantly. The results in this study may indicate that the fusion of SWOT and other satellites (e.g., CubeSats) can address limitations caused by the lower temporal resolution of SWOT data (i.e., 3–21 days) and thus improve discharge estimation at the global scale.

Of chief interest in this study is the ability to estimate discharge at spatial scales not previously shown using satellite data. However, the BAM-AMHG discharge estimates for small rivers are not as good as for large rivers. In addition to the CubeSat data quality and spatial resolution issues, the sample size is also problematic. The numbers of total cross sections for our 11 reaches are not consistent, varying from 748 for 15453500 to 23 for 15511000. For small rivers, the pool size from which the widths were selected is much smaller, which can reduce the signals required for discharge estimation. Increasing this pool size would require denser cross sections (which may not introduce new information), or lengthening our study reach. Since BAM is a McFLI algorithm, lengthening the reaches would possibly introduce errors in mass conservation: the shorter the reach around the gauge, the more likely that gauge discharge represents the reach. This balance has not been studied intensively and is a worthy subject for future research. Small rivers are also poorly described by AHG from space-borne widths, even at 3-m resolution. Only 20–50% of cross sections exhibit significant AHG ( $p < 0.10$ ) for small rivers. Since both the rating curve method and BAM are fundamentally trying to invert AHG, this explains a large part of the poor performance by BAM as well.

In this study, only 20 cross sections at each reach were used for discharge estimation to allow for consistent comparison across scales. This can be problematic for some large rivers, since the spatial/temporal width variations of the 20 selected cross sections may not be sufficient to drive reach-scale AMHG in the BAM algorithm. For example, for the Yukon River reaches (15453500 and 15356000; Figures S4 and S5), the selected cross sections were located within a reach of 600 m, and the AMHG built on such short reaches is likely weaker than that previously reported by Gleason and Smith (2014). This issue can be avoided in the future by modifying the cross-section filtering technique. For example, the cross sections can be divided into groups based on their spatial locations, and then those with the best AHG in each group passed to BAM. As a Bayesian technique, BAM performance relies on prior information (Figure 8). In this study, a simple approach was used to transform surface runoff to discharge (i.e., discharge = runoff \* drainage area), which ignores water transportation processes and thus likely introduces some bias. Therefore, integrating runoff (surface and/or subsurface runoff) and/or groundwater data sets with some hydrologic routing models to generate BAM priors may be a possible approach to improve discharge estimation for ungauged basins. Furthermore, we have not tested how varying levels of prior information affect discharge retrieval for the gauged cases. These ungauged results indicate that the worse a prior distribution approximates true flow, the worse BAM performs (as expected). We could have also added another test where we subsampled the gauge records to test the role of prior data, but this test lies outside of our scope. Another interesting point revealed by this study is that the NCEP reanalysis runoff data set performs better than the newly released MERRA-2 and ERA5 when using the simple runoff-discharge transformation approach. Therefore, NCEP is recommended as a prior source for discharge estimation, at least for the rivers in this study. However, further investigation about their performance when using different runoff-discharge transformation approaches (e.g., incorporated with routing models) is necessary for future work.

Also for ungauged basins, we found that the error metrics at each cross section within a reach are similar, which implies that there is no obvious clue for detecting “good” cross sections for BAM input. This has long been a goal of McFLI research—trying to determine, without in situ data, where discharge estimation will be most successful. We tested and rejected the hypothesis that choosing the most-variable-width cross sections is more accurate than an ensemble-median in this study. Although the better performance of the ensemble-median approach is not statistically significant, we still recommend it, as it is less sensitive to the outliers in widths (which is particularly important when the width uncertainties are large), and it does not rely on prior knowledge of discharge and is thus readily applied in all basins.

BAM-AMHG is developed based on the AMHG theory, which fundamentally requires cross sectional measurements of width. However, this may not be the best approach to detect the width-discharge signals, especially for small rivers. First, the width-discharge relationship at cross section scales can be susceptible to morphological changes. In this study, we assume the morphology did not change during the study period (2016–2018). Although our study period is short, morphology is still likely to change somewhat, which may weaken the AMHG strength and subsequent discharge simulation. Second, a precise measurement of widths at cross section scales requires high resolution images, which makes most of the current remote sensing imagery with moderate spatial resolutions (e.g., MODIS and Landsat) poor for small rivers in this study. The development of remote sensing techniques (i.e., higher spatial and spectral resolutions), combined with appropriate consideration of cross section allocation and morphology changes, can improve the



performance of the approach proposed in this study. Even for currently available remote sensing techniques, including the optical, SAR and microwave sensors, a reach-averaged width, instead of cross sectional width, may be an option to alleviate the issues mentioned above (e.g., Brakenridge et al., 2007, 2012; Durand et al., 2016; Van Dijk et al., 2016). However, AHG and AMHG theories have not been firmly established at the reach scale. How to allocate measurement units (e.g., reach vs. cross section, length, and interval) to balance width accuracy, morphology change impacts, and AMHG validation requires further investigation.

Finally, we can contextualize this study with other remote sensing of discharge literature. This is the first study to estimate discharge solely from remotely sensed information at such fine spatial scales, and among the first to show such high data density for Arctic rivers. Our results show that BAM can be as accurate as, if not more accurate than, traditional space-based rating curves. This is an interesting point and is strengthened by the diminishing gauge record. As long as past gauges represent current gauge realities (i.e., channel morphology, AHG and AMHG strength, mean, and variability of the hydrograph should be similar), BAM can continue to provide accurate estimates without the need to resample the rating curve distribution. The methods of, for example, Tarpanelli et al. (2019) could also be compared to this gauge-based case, but our goal was not an exhaustive intercomparison. We would expect that any of the many discharge remote sensing techniques that rely on gauge data (e.g., Bjerklie (2007), Bjerklie et al. (2003), Pavelsky (2014), Smith et al. (1996), Smith and Pavelsky (2008), Van Dijk et al. (2016)) would have similar performance in the gauged case.

In an ungauged case, our results are similar to previously published AMHG-based results (Durand et al., 2016; Gleason & Smith, 2014; Hagemann et al., 2017). These results are not as accurate, however, as expected results using the Manning equation within BAM after the launch of SWOT. Durand et al. (2016) show (using hydraulic model output as a stand-in for SWOT) that this extra dimension (observations of width and height, as opposed to just width here) greatly improves performance. While Bjerklie et al. (2018) fused altimeters with optical data to generate such data, this is infeasible for the spatial scales in this study, as altimetry is only available for large rivers. We therefore suggest, following our results, that width-only discharge estimates are useful in both gauged and ungauged settings but that these will have diminished value after the launch of SWOT. SWOT, however, will only observe rivers wider than 100 m with a goal of 50 m (Biancamaria et al., 2016), so even SWOT will not see more than half of the rivers in this study. In addition, CubeSat technology may improve in the future, by, for example, introducing midinfrared spectral bands and improving the techniques for radiometric calibrations between instruments. Seen in this light, the fusion of Landsat/Sentinel-2 and CubeSats holds great promise for future hydrologic application, even with the advent of SWOT.

## 5. Conclusions

This study presents a first assessment of CubeSat data for discharge estimation for Arctic rivers. Landsat 8, Sentinel-2, and Planet CubeSat data sets and their fusions were tested for estimating discharge for 11 Arctic river reaches with various sizes spanning from >1,000 to <20 m wide. The relatively low temporal/spatial resolutions (for Landsat and Sentinel-2) or low radiometric data quality (for Planet) limit the applications of these data sets used alone. Their fusion, however, can be an approach to enhance their ability to estimate discharge. The BAM method is typically thought of as a technique for ungauged basins, but we here show that BAM can, in some cases, outperform a traditional space-based rating curve when given the same input data. Space-based rating curves are sensitive to outliers or noise in the width data and the sampling of hydrograph dynamics, which makes them less accurate when using the fused satellite data. In contrast, BAM-AMHG can provide more statistically stable discharge estimates. The results in this study show that BAM-AMHG forced with fused width data from both government-sponsored and CubeSat satellites can improve discharge estimation for both large and medium rivers (>40 m wide) by increasing data density by a factor of 2–6 without hurting accuracy. In ungauged basins, using globally available NCEP surface runoff reanalysis data, BAM forced with fused satellite data can estimate discharge in the study region with acceptable accuracy (34–71% nRMSE for large rivers and 61–69% for medium rivers). SWOT's two-dimensional observations are expected to improve these results substantially (Durand et al., 2016), but this

study suggests that the fusion of CubeSat and government-sponsored satellites (e.g., Landsat and Sentinel 2) represents a possible approach to improve global discharge estimation before the launch of SWOT.

### Acknowledgments

This work was partially supported by NASA SWOT Science Team Grant NNX13AD96G, and NSF CAREER Grant 1748653 to C. J. Gleason. D. Feng was supported by NASA New Investigator Grant 80NSSC18K0741 to C. J. Gleason. T. Pavelsky and X. Yang were supported by a contract from the SWOT Project Office at the NASA/Caltech Jet Propulsion Lab. We gratefully acknowledge Planet for providing access to their satellite imagery archive through their Education and Research Program. All data and source code used in this study are available online (<https://github.com/dongmeifeng-2019/CubeSat>).

### References

- Aguilera, R., & Melack, J. M. (2018). Relationships among nutrient and sediment fluxes, hydrological variability, fire, and land cover in Coastal California catchments. *Journal of Geophysical Research: Biogeosciences*, *123*, 2568–2589. <https://doi.org/10.1029/2017JG004119>
- 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>
- Aragon, B., Houborg, R., Tu, K., Fisher, J. B., & McCabe, M. (2018). CubeSats enable high spatiotemporal retrievals of crop-water use for precision agriculture. *Remote Sensing*, *10*, 1867. <https://doi.org/10.3390/rs10121867>
- Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309. <https://doi.org/10.1038/nature04141>
- Beven, K. J., & Cloke, H. L. (2012). Comment on “Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water” by Eric F. Wood et al. *Water Resources Research*, *48*, W01801. <https://doi.org/10.1029/2011WR010982>
- 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. Space Sciences Series of ISSI* (Vol. 55, pp. 117–147). Cham: Springer. [https://doi.org/10.1007/978-3-319-32449-4\\_6](https://doi.org/10.1007/978-3-319-32449-4_6)
- Bjerklie, D. M. (2007). Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information. *Journal of Hydrology*, *341*(3–4), 144–155. <https://doi.org/10.1016/j.jhydrol.2007.04.011>
- 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](https://doi.org/10.1016/S0022-1694(03)00129-X)
- 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*(1–4), 191–209. <https://doi.org/10.1016/j.jhydrol.2004.11.022>
- 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>
- Brakenridge, R. G., Cohen, S., Kettner, A. J., De Groeve, T., Nghiem, S. V., Syvitski, J. P. M., & Fekete, B. M. (2012). Calibration of satellite measurements of river discharge using a global hydrology model. *Journal of Hydrology*, *475*, 123–136. <https://doi.org/10.1016/j.jhydrol.2012.09.035>
- Coluzzi, R., Imbrenda, V., Lanfredi, M., & Simoniello, T. (2018). A first assessment of the Sentinel-2 Level 1-C cloud mask product to support informed surface analyses. *Remote Sensing of Environment*, *217*, 426–443. <https://doi.org/10.1016/j.rse.2018.08.009>
- Cooley, S. W., Smith, L. C., Ryan, J. C., Pitcher, L. H., & Pavelsky, T. M. (2019). Arctic-boreal lake dynamics revealed using CubeSat imagery. *Geophysical Research Letters*, *46*, 2111–2120. <https://doi.org/10.1029/2018gl081584>
- Copernicus Climate Change Service (C3S) (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), [March 20, 2019].
- Durand, M., Fu, L., Lettenmaier, D. P., Alsdorf, D. E., Rodriguez, E., & Esteban-Fernandez, D. (2010). The surface water and ocean topography mission: Observing terrestrial surface water and oceanic submesoscale eddies. *Proceedings of the IEEE*, *98*(5), 766–779. <https://doi.org/10.1109/JPROC.2010.2043031>
- 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., Rodriguez, E., Alsdorf, D. E., & Trigg, M. (2010). Estimating river depth from remote sensing swath interferometry measurements of river height, slope, and width. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(1), 20–31. <https://doi.org/10.1109/JSTARS.2009.2033453>
- ESA (2015). In B. Hoersch (Ed.), *Sentinel-2 user handbook*, (pp. 1–64).
- Feng, D., Beighley, E., Hughes, R., & Kimbro, D. (2016). Spatial and temporal variations in eastern U.S. Hydrology: Responses to global climate variability. *JAWRA Journal of the American Water Resources Association*, *52*(5), 1089–1108. <https://doi.org/10.1111/1752-1688.12445>
- Feng, D., Beighley, E., Raoufi, R., Melack, J., Zhao, Y., Iacobellis, S., & Cayan, D. (2019). Propagation of future climate conditions into hydrologic response from coastal southern California watersheds. *Climatic Change*, *153*(1–2), 199–218. <https://doi.org/10.1007/s10584-019-02371-3>
- Gleason, C. J., & Hamdan, A. N. (2017). Crossing the (watershed) divide: satellite data and the changing politics of international river basins. *The Geographical Journal*, *183*(1), 2–15. <https://doi.org/10.1111/geoj.12155>
- 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 of the United States of America*, *111*(13), 4788–4791. <https://doi.org/10.1073/pnas.1317606111>
- 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>
- Global Modeling and Assimilation Office (GMAO) (2015). MERRA-2 tavG\_U\_2d\_Ind\_Nx: 2d,diurnal, Time-averaged, single-level, assimilation, land surface diagnostics V5.12.4. Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA.
- Google Earth (2018) V 7.3.2.5776. (June 1, 2017). Alaska, .64° 55' 48.32"N, 145° 52' 30.63"W, Eye alt 277.74 mile. Landsat/Copernicus 2018. <http://www.earth.google.com> [April 06, 2019].
- 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. *Water Resources Research*, *53*, 9692–9707. <https://doi.org/10.1002/2017WR021626>
- Hannah, D. M., Demuth, S., van Lanen, H. A. J., Looser, U., Prudhomme, C., Rees, G., et al. (2011). Large-scale river flow archives: Importance, current status and future needs. *Hydrological Processes*, *25*(7), 1191–1200. <https://doi.org/10.1002/hyp.7794>

- Houborg, R., Fisher, J. B., & Skidmore, A. K. (2015). *Advances in remote sensing of vegetation function and traits*, (Vol. 43, pp. 1–6). Elsevier. <https://doi.org/10.1016/j.jag.2015.06.001>
- Houborg, R., & McCabe, M. F. (2018). A Cubesat enabled Spatio-Temporal Enhancement Method (CESTEM) utilizing planet, Landsat and MODIS data. *Remote Sensing of Environment*, 209, 211–226. <https://doi.org/10.1016/j.rse.2018.02.067>
- Huete, A., Justice, C., & Van Leeuwen, W. (1999). MODIS vegetation index (MOD13), Algorithm theoretical basis document.
- Irons, J. R., Dwyer, J. L., & Barsi, J. A. (2012). The next Landsat satellite: The Landsat Data Continuity Mission. *Remote Sensing of Environment*, 122, 11–21. <https://doi.org/10.1016/j.rse.2011.08.026>
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432. <https://doi.org/10.1080/01431169608948714>
- Nathanson, M., Kean, J. W., Grabs, T. J., Seibert, J., Laudon, H., & Lyon, S. W. (2012). Modelling rating curves using remotely sensed LIDAR data. *Hydrological Processes*, 26(9), 1427–1434. <https://doi.org/10.1002/hyp.9225>
- 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>
- Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *Science*, 313(5790), 1068–1072. <https://doi.org/10.1126/science.1128845>
- 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>
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422. <https://doi.org/10.1038/nature20584>
- Planet Team (2018) Planet application program interface: In space for life on Earth, San Francisco, CA.
- Puig-Suari, J., Turner, C., & Ahlgren, W. (2001). Development of the standard CubeSat deployer and a CubeSat class PicoSatellite, 2001 IEEE Aerospace Conference Proceedings (Cat. No.01TH8542). <https://doi.org/10.1109/AERO.2001.931726>
- Smith, L. C., Isacks, B. L., Bloom, A. L., & Murray, A. B. (1996). Estimation of discharge from three braided rivers using synthetic aperture radar satellite imagery: Potential application to ungauged basins. *Water Resources Research*, 32(7), 2021–2034. <https://doi.org/10.1029/96WR00752>
- Smith, L. C., & Pavelsky, T. M. (2008). Estimation of river discharge, propagation speed, and hydraulic geometry from space: Lena River, Siberia. *Water Resources Research*, 44, W03427. <https://doi.org/10.1029/2007WR006133>
- 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>
- Tarpanelli, A., Santi, E., Tourian, M. J., Filippucci, P., Amarnath, G., & Brocca, L. (2019). Daily river discharge estimates by merging satellite optical sensors and radar altimetry through artificial neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), 329–341. <https://doi.org/10.1109/TGRS.2018.2854625>
- USGS (2019). In K. Zanter (Ed.), *Landsat 8 (L8) Data Users Handbook*.
- Van Dijk, A. I. J. M., Brakenridge, G. R., Kettner, A. J., Beck, H. E., De Groeve, T., & Schellekens, J. (2016). River gauging at global scale using optical and passive microwave remote sensing. *Water Resources Research*, 52, 6404–6418. <https://doi.org/10.1002/2015wr018545>
- Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. *Nature*, 467(7315), 555–561. <https://doi.org/10.1038/nature09440>
- Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. <https://doi.org/10.1080/01431160600589179>
- 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*, 1–5. <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>
- 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>