



# JGR Earth Surface

## RESEARCH ARTICLE

10.1029/2019JF005033

### Key Points:

- A Calibration set of 134,697 *in situ* measurements was used to estimate suspended-sediment concentration from Landsat satellite imagery
- A Base calibration model (97% uncertainty) is improved by automated river clustering (74% uncertainty) and added *in situ* data (50% uncertainty)
- Cosmopolitan suspended-sampling campaigns may best refine satellite imagery approaches

### Supporting Information:

- Supporting Information S1

### Correspondence to:

E. N. Dethier,  
evan.n.dethier.gr@dartmouth.edu

### Citation:

Dethier, E. N., Renshaw, C. E., & Magilligan, F. J. (2020). Toward improved accuracy of remote sensing approaches for quantifying suspended sediment: Implications for suspended-sediment monitoring. *Journal of Geophysical Research: Earth Surface*, 125, e2019JF005033. <https://doi.org/10.1029/2019JF005033>

Received 5 FEB 2019

Accepted 15 MAY 2020

Accepted article online 9 JUN 2020

## Toward Improved Accuracy of Remote Sensing Approaches for Quantifying Suspended Sediment: Implications for Suspended-Sediment Monitoring

E. N. Dethier<sup>1</sup> , C. E. Renshaw<sup>1</sup> , and F. J. Magilligan<sup>2</sup>

<sup>1</sup>Department of Earth Sciences, Dartmouth College, Hanover, NH, USA, <sup>2</sup>Department of Geography, Dartmouth College, Hanover, NH, USA

**Abstract** Because of the high logistical and financial costs of direct measurements of riverine suspended sediment, remote sensing is increasingly used to supplement the direct-observation record. The accuracy of this method is poorly constrained, and its potential as a tool for understanding river sediment transport is thus limited. We introduce and apply global-scale methods for estimating depth-integrated suspended-sediment concentrations (SSCs) using Landsat 5 and 7 satellite imagery calibrated with 134,697 *in situ* SSC measurements. We account for river-to-river variability in the relationship between water optical properties and SSC by (a) categorizing rivers using unsupervised K-means clustering and/or (b) correcting calibration estimates for individual rivers using local *in situ* measurements of SSC, suspended-sediment grain size, and percent organic carbon (POC). In the absence of site-specific *in situ* SSC measurements, clustering rivers reduces the average relative error of SSC estimates from 97% to 73%. We show that as few as five site-specific *in situ* measurements combined with our algorithm further reduces average relative error to 49% and average relative at-a-station bias in SSC to 7%. Little additional improvement in accuracy or bias is gained by including measurements of percent sand or POC of the suspended sediment. Since only modest additional accuracy is gained after ~5–10 paired *in situ* SSC measurements and satellite observations, sampling campaigns should prioritize limited sampling at diverse locations rather than intensive sampling at a limited number of sites. In addition, we publish standalone calibrations for 151 rivers made solely with *in situ* SSC measurements local to those sites.

**Plain Language Summary** We present methods for using satellite imagery from the Landsat 5 and 7 missions to detect suspended-sediment concentrations in large rivers. We establish best practices for determining the accuracy of these methods and identify specific ways for improving that accuracy. We show that little additional accuracy in the method is gained after about five paired measurements of suspended-sediment concentration and satellite observation, implying that it is better to collect a limited number of paired samples at a large number of sites rather than to intensively sample just a few sites.

## 1. Introduction

Quantitative understanding of large-scale patterns and trends in suspended-sediment transport is limited by the paucity of spatially and temporally robust suspended-sediment measurements. Detecting changes to suspended-sediment transport in the world's rivers, particularly those due to rapid anthropogenic changes (Best, 2019), requires high-quality monitoring and historical context. However, suspended-sediment data sets are of variable quality and coverage, as the monitoring required to quantify suspended-sediment concentration (SSC) is logically and financially expensive. Given these constraints, remote sensing approaches are increasingly used to supplement the observational record (Dethier et al., 2019; Doxaran et al., 2003; Heege et al., 2014; Latrubesse et al., 2017; Long & Pavelsky, 2013; Mertes et al., 1993; Montanher et al., 2014; Overeem et al., 2017; Pavelsky & Smith, 2009; Ritchie & Cooper, 1988; Stumpf & Pennock, 1989; Umar et al., 2018; Villar et al., 2012; Yepez et al., 2017). Implementing best practices for developing, testing, improving, and quantifying uncertainty for these remote sensing methods is critical to their widespread successful use.

Satellite remote sensing of suspended sediment relies on the strong absorption of incoming electromagnetic radiation by water. Sediment suspended in water reflects incoming radiation more strongly than pure water, particularly in the visible and near-infrared (NIR) wavelengths (Bhargava & Mariam, 1991; Chen et al., 1991;

Novo et al., 1989), allowing for development of empirical relationships between surface reflectance (SR) and SSC. Developing a SR-SSC relationship requires linking in situ SSC measurements to contemporaneous satellite reflectance data of high-quality river pixels (no cloud/haze, shadow, snow/ice, or river bank/bar pixel contamination) at the in situ sampling location. This data set of image-in situ SSC pairs can then be used to calibrate physical and semiempirical models (Heege et al., 2014; Stumpf & Pennock, 1989) or empirical models (Dethier et al., 2019; Doxaran et al., 2003; Latrubblesse et al., 2017; Montanher et al., 2014; Overeem et al., 2017; Umar et al., 2018; Yepez et al., 2017). These models are then applied to predict SSC beyond the spatial and/or temporal scope of in situ samples, either local to the in situ sampling location(s) or extended regionally or globally (Dethier et al., 2019; Montanher et al., 2014).

Despite this straightforward premise, numerous challenges complicate the successful estimation of SSC. Prior approaches have varied widely in methodology, and calibration data sets have ranged from  $<10$  in situ measurements made at a single location to  $>1,000$  measurements from tens of rivers. Varying in situ sampling techniques, atmospheric properties, satellite image processing methods, and suspended-sediment characteristics all contribute to uncertainty in estimating SSC from satellite observations. The objectives of this paper are to systematically address these research challenges, quantify uncertainty where it cannot be reduced, and increase the size and representativeness of calibration data sets, seeking to improve and standardize techniques for satellite-derived SSC estimation. The key question addressed by this study is not whether remotely sensed estimates of SSC are as accurate as direct measurements of SSC (in general, they are not), but rather how can researchers maximize the accuracy of remotely sensed estimates of depth-averaged SSC while understanding their inherent uncertainties. Accordingly, we capitalize on the 36-year combined Landsat 5 and 7 satellite record (1984 to the present) and the Google Earth Engine platform to develop uniform, globally scalable methods for estimating SSC along rivers greater than  $\sim 90$  m wide.

## 1.1. Background

Before presenting our methodology, in this section we briefly review the key challenges to developing methods for satellite-derived estimates of SSC, which have been widely discussed in the literature. This review provides the necessary context for understanding our approach, its possible limitations, and the potential it offers for geomorphic applications.

### 1.1.1. Suspended-Sediment Sampling Methods

Successful calibration for satellite-derived estimates of SSC relies on consistent, accurate in situ measurements of SSC. Since the penetration of light through water is at most several meters, and substantially less for turbid water (Stefan et al., 1983), the spectral reflectance captured in remote sensing imagery necessarily best correlates with the surface SSC (SSSC) (Overeem et al., 2017; Umar et al., 2018). Some studies have sought to eliminate inconsistencies in river sediment mixing profiles by developing calibrations for SSSC (Martinez et al., 2015; Park & Latrubblesse, 2014; Umar et al., 2018; Villar et al., 2012), using only measurements taken at the river surface. Others have assumed uniform water column mixing (Overeem et al., 2017) or developed an empirical calibration between SR and depth-averaged SSC, implicitly assuming a consistent vertical-mixing profile (Dethier et al., 2019; Yepez et al., 2017). In each case, where rivers deviate from the assumed mixing profile due to disparities in grain size or flow dynamics, model predictions will necessarily be less accurate.

By extension, variations in in situ SSC sampling techniques used in creating calibration data sets are a primary source of uncertainty for SSC calibrations. Depth-integrated sampling approaches are generally the most accurate in situ methods (Lee & Glysson, 2013; Lurry & Kolbe, 2000) but are more expensive and labor-intensive than samples collected from surface water, from a single depth, and/or using an automatic pump. As a result, many SSC sampling data sets are composed of surface or middepth measurements, which may or may not be adjusted to approximate depth-integrated methods by use of, for example, a site-specific surface SSC-depth-integrated SSC calibration (Filizola & Guyot, 2009; Lee & Glysson, 2013; Lurry & Kolbe, 2000; Martinez et al., 2009).

### 1.1.2. Satellite Imagery Processing

In addition to high-quality in situ data, calibrations for satellite detection of SSC require high-quality imagery of the river reach(es) in question, as well as successful identification and sampling of river-water pixels. Cloud cover, shadows, haze, and ice/snow compromise or preclude satellite techniques for detecting SSC that incorporate visible and NIR electromagnetic spectra. In addition, rivers with high SSC may be

difficult to distinguish from land, clouds, or ice (Hudson et al., 2014). These factors can hinder both the initial automated identification of river water, as well as subsequent calibration of relationships between river-water reflectance and SSC. Some investigators have developed or used nonagency empirical or physical algorithms for quality control and atmospheric corrections (Doxaran et al., 2003; Heege et al., 2014; Kilham et al., 2012; Montanher et al., 2014; Wang & Lu, 2010), while others have relied on the SR and quality control products provided by National Aeronautics and Space Administration (NASA) for Landsat and MODIS satellites (Chu et al., 2012; Dethier et al., 2019; Miller & McKee, 2004; Park & Latrubesse, 2014; Ross et al., 2019). Though few SSC studies have conducted rigorous validation of these agency products, particularly those that are in development, Yepez et al. (2017) and Kuhn et al. (2019), respectively, showed that the U.S. Geological Survey (USGS) SR product was most successful in replicating in situ spectral measurements on the Amazon River and Orinoco River. Theoretical improvement in SSC model calibration resulting from location-specific atmospheric calibrations must be balanced by the benefits of global uniformity, which is currently offered primarily by agency calibrations.

After images with high-quality river pixels are identified and processed, pairing them with contemporaneous in situ measurements of SSC introduces another potential source of uncertainty. Particularly, in systems with frequent SSC fluctuations relative to in situ sampling frequency, lags between in situ sampling and satellite image acquisition may result in imagery of SSCs that do not correspond to the associated in situ sample. Most studies have sought to match in situ measurements with satellite images acquired on the same day. However, limited sampling campaigns have led others to accept lags of  $\pm 1$  day (Stumpf & Pennock, 1989) to as much as  $\pm 9$  days (Dethier et al., 2019; Montanher et al., 2014). The trade-off between additional calibration measurements and reduced calibration accuracy is unknown and warrants systematic investigation.

### 1.1.3. Variability Due to Sediment and River Characteristics

Suspended-sediment and river-water characteristics that influence river optical properties can also impact calibration success (Bhargava & Mariam, 1991; Chen et al., 1991; Dethier et al., 2019; Kilham et al., 2012; Martinez et al., 2009, 2015; Novo et al., 1989). Lower per-mass reflectance of coarser grains can lead to model underestimation of SSC when suspended sediment has a low mass fraction of fine-grained particles (Novo et al., 1989). Sediments with different mineralogy have varying absorption profiles, leading to interriver and intrariver variation in the spectra most strongly correlated with SSC (Bhargava & Mariam, 1991; Chen et al., 1991). Organic material also introduces variability into river reflectance profiles. Dissolved organic matter absorbs blue and green wavelengths, while chlorophyll-*a* generally absorbs blue and red light and reflects NIR light (Martinez et al., 2015). Because high reflectance in these spectra is generally associated with high-SSC water, low-SSC water with significant dissolved organic matter (Martinez et al., 2015; Montanher et al., 2014) or chlorophyll-*a* (Martinez et al., 2015; Mertes et al., 1993; Ritchie & Cooper, 1988; Stumpf & Pennock, 1989) may incorrectly lead to model inference of high SSC.

### 1.1.4. Addressing Interriver and Intrariver Variations in River Properties

Interriver variations in sediment characteristics lead to uncertainty in SSC model calibrations, even when those calibrations are made and applied at a single river location. Some studies have addressed interriver variation by developing calibrations that included additional in situ data about organic content or grain size (Kilham et al., 2012) or developed calibrations for different seasons (Park & Latrubesse, 2014). Reliance on in situ data or a posteriori categorization often requires intensive local sampling and/or knowledge, which may limit the widespread development and application of these methods. In addition, the transferability of the resulting local calibrations is likely limited due to site-specific sediment characteristics and, particularly when in situ sampling campaigns are limited, a narrow range of SSCs incorporated in model calibration data sets (Long & Pavelsky, 2013).

### 1.1.5. Previous Studies

Few studies have fully addressed these fundamental challenges to making globally applicable, satellite-derived SSC predictions. Long and Pavelsky (2013) found that a majority of published calibration models poorly predicted in situ measurements made during a field campaign, highlighting the limited transferability of most calibrations. However, several important efforts toward global calibration methods have occurred. Long and Pavelsky (2013) showed that models were more successful and transferrable when developed using a combination of visible and infrared bands, and in situ measurements spanning a wide range of SSCs. In addition, they provided a method for correcting existing calibration models using in situ data made at a location of interest. Montanher et al. (2014) accounted for the difference between “white” and “black”

rivers in Brazil, manually dividing them using satellite image inspection and regional knowledge. Ritchie et al. (2003) suggested different calibration models should be developed for water bodies with low ( $<50 \text{ mg/L}$ ) and high ( $\geq 50 \text{ mg/L}$ ) SSC. More automated, and thus potentially transferrable, grouping of optically similar waterbodies has been used to group lake types using multispectral data (Spyrakos et al., 2018) and some rivers in South America (Dethier et al., 2019; Martinez et al., 2015) for subsequent separate calibration. Though these approaches show promise, they have not yet been widely adopted for river SSC remote sensing.

### 1.2. Toward Globally Applicable Methods for Satellite Detection of SSC

The above challenges to successfully utilizing satellite-derived estimates of SSC must be balanced against the potential value of global methods for estimates of SSC. In particular, the global coverage of satellite measurements holds the potential for scalable measurement of both spatial and temporal variability in SSC, which is difficult, if not logistically unachievable, using at-a-station measurements. Satellite-based estimates allow for investigation of both monitored and otherwise-unmonitored rivers, and extant satellite records provide the means for historical analysis even absent historical in situ data.

Seeking to improve the accuracy and transferability of these increasingly implemented techniques, we address the relative contributions of different sources of error to overall calibration uncertainty for different calibration strategies, made by (1) using K-means clustering to group optically similar rivers and (2) making river-specific corrections to the global calibration models based on local in situ measurements. We develop protocols for minimizing uncertainty and bias in applications of these techniques.

## 2. Methods

We describe below the development of a satellite SR-in situ SSC calibration model. We accessed publicly available data to generate data sets of in situ SSC measurements from 727 stations and water-only Landsat data from in situ station locations. These data sets were linked by date and station, providing the basis for calibration models relating in situ SSC to Landsat-measured SR. As reviewed above, calibrations depended on, and were sensitive to, methods of satellite image processing, filtering, and masking of non-water pixels; automated grouping of optically alike river types; in situ sampling methodology; river and sediment transport characteristics; and Landsat image/sample metadata. We address calibration sensitivity to these variables and present several methods for calibration model development, quantifying model prediction uncertainty using multiple holdout data sets. We describe these methods below, graphically using a flow chart in the supporting information (Figure S1), and provide the Google Earth Engine and R scripts used for processing at <https://github.com/evandethier/satellite-ssc> website.

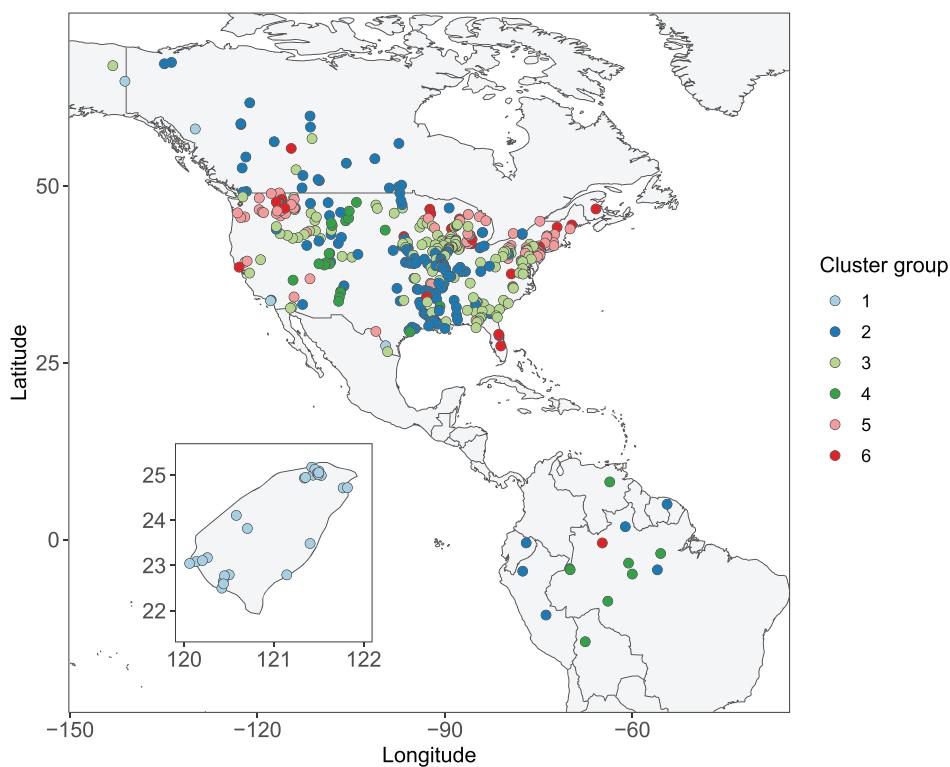
### 2.1. Data Sources: Remote Imagery and In Situ Measurements

#### 2.1.1. In Situ SSC Measurements and Metadata

The in situ measurements used for calibration were made from 1984–2019 in the United States by the USGS (USGS, 2018), from 1984–2001 in Canada by the Water Survey of Canada (WSC) (HYDAT, 2019), from 2000–2019 in South America by Hydro-geochemistry of the Amazonian Basin (HYBAM), and from 1999–2014 in Taiwan by the Taiwan Water Resources Agency (WRA) (WRA, 2018) (Figure 1).

The USGS (De Cicco et al., 2018) and WSC (Albers, 2017) provide bulk download functionality through agency-developed R programming language packages. We used this functionality to generate data sets of in situ SSC measurements. SSC in situ measurements from Taiwan and South America each required manual retrieval from agency websites. The USGS and WSC provided consistent, reliable station metadata information, while station metadata were not consistently reported by WRA and HYBAM. SSC sample metadata relevant to calibration methods were extensive in some cases, but reporting was inconsistent, as summarized in Table 1.

To filter out rivers too narrow for reliable water-only pixel retrieval from Landsat images, the in situ data set was limited to stations with drainage area  $>3,000 \text{ km}^2$  (except in Taiwan, where the threshold was  $800 \text{ km}^2$  due to typically wide, unvegetated channels). We manually ensured that each remaining in situ river-reach site was wide enough for Landsat sampling, eliminating many sites with large drainages but narrow channels in the American West.



**Figure 1.** Map of in situ SSC sampling stations that contributed to the calibration data set, with inset map showing stations in Taiwan. Site color indicates unsupervised K-means cluster group.

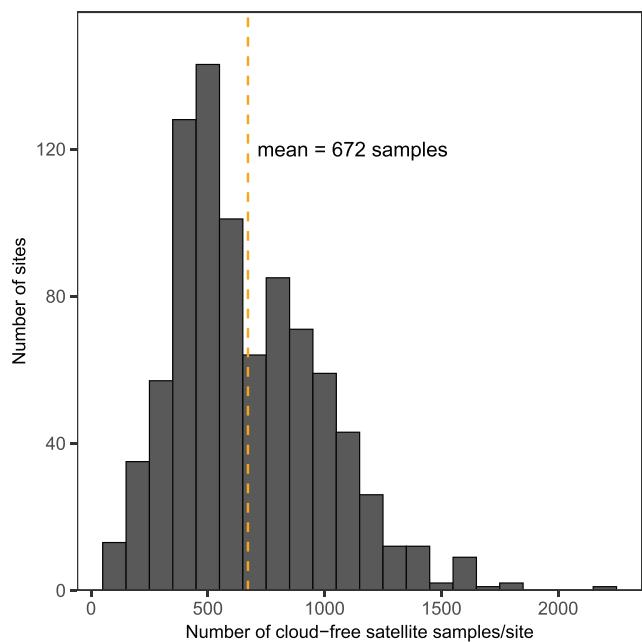
Because metadata regarding in situ SSC measurement technique varied in detail among and within agencies, we could not ensure that every sample was depth averaged. Though this uncertainty likely contributed to total model uncertainty, we believe that the benefits of incorporating these additional measurements justify their use. In order to evaluate error due to variations in sampling techniques, we categorized in situ sampling methodology as depth-integrated, single-depth (nonsurface or unspecified depth), surface (top ~1 m of water column), and unknown/miscellaneous. We converted measurements made at some South American stations from surface concentrations to depth-averaged concentration estimates, following Martinez et al. (2009) and Filizola and Guyot (2009). For each methodology category, we analyzed model calibration error and bias. In addition, where samples had adequate metadata, we assessed whether calibration error was related to sample depth for single-depth measurements or average depth for depth-integrated measurements.

Additional station and sample metadata were used to evaluate calibration model performance. Drainage area was available for all stations. Grain-size information was reported in the USGS and WSC in situ data sets, but only for some samples (Table 1). The most common grain-size measurement ( $n = 10,386$ ) was “P63,” the fraction of suspended sediment finer-than-sand sized. Particulate organic carbon (POC,  $n = 3,864$ ) measurements were also included in SSC sampling metadata for some USGS stations. For each

**Table 1**  
Number of In Situ Samples Matched With Landsat 5 or 7 SR Images

Sampling agency	Stations	Stations ( $\geq 10$ samples)	In situ samples	Landsat 5	Landsat 7	P63 stations	P63	POC stations	POC
USGS	655	423	121,379	73,077	48,302	424	9,740	180	3,864
WCS	28	28	11,042	10,963	79	24	646	—	—
HYBAM	16	15	1,244	473	771	—	—	—	—
WRA	28	27	1,032	494	538	—	—	15 <sup>a</sup>	561 <sup>a</sup>
Total	727	493	134,697	85,007	49,690	448	10,386	180	3,864

<sup>a</sup>In situ samples from Hilton et al. (2010). Not matched with Landsat satellite images.



**Figure 2.** Histogram of the number of Landsat 5 and 7 satellite samples that were extracted for each in situ station. To be counted, a sample of cloud, haze, and ice/snow-free water pixels must have been obtained from the image. Bins are 100 satellite samples wide.

Infrared 1, 1,550–1750 nm), and B7 (Shortwave Infrared 2, 2080–2,350 nm) (Survey, U. G, 2019). The Landsat SR product also includes basic landcover classification and quality assurance codes for each pixel in the “pixel\_qa” band. Some of the automated classification routines are not optimized for water detection, particularly for rivers with high SSC (Hudson et al., 2014). However, in general, the USGS classifications successfully identify clouds, cloud shadows, ice, and snow, and we thus eliminated pixels with these classifications.

We did not use the USGS classification to distinguish water pixels from land pixels. Instead, we used a suite of empirically derived thresholds best suited for this application. The first was a modified normalized difference water index (MNDWI) using B2 and B5. Water pixels were automatically classified using a threshold of 0, as dynamic threshold calculated for each Landsat image (Allen & Pavelsky, 2018; Li & Sheng, 2012) did not lead to significant improvement in water pixel sampling. To eliminate pixels affected by haze, glint/glare, and/or appearing to incorporate information from banks or bars, we only selected pixels with B7 reflectance  $<0.05$ . Cloud edges are sometimes incorrectly classified by the USGS quality control algorithm, so we adopted a more conservative approach, adding an additional threshold of B1 reflectance + B2 reflectance  $<0.5$ . Ice and snow in or proximal to river water often contaminate water pixels by increasing their perceived reflectance. Though we flagged images with pixels within 3 km of the in situ station classified as snow or clouds as low quality in the calibration data set, we nonetheless used these pixels in the analysis to avoid eliminating high-SSC pixels classified as snow (Hudson et al., 2014). We automated sampling of the water pixels within 200 m of each in situ sampling station for each image in the Landsat 5 and 7 archive, calculating the daily median surface reflectance of those pixels in each band (Figure 2). This 200-m radius in general allows for water pixel sampling even in the presence of Landsat 7 scanline errors. This protocol was packaged together in an automated function in Google Earth Engine, which for each processed image generates an image comprising only uncontaminated water pixels, and is available at <https://github.com/evandethier/satellite-ssc> (Figure S1).

### 2.1.3. Satellite Image-In Situ Measurement Calibration Data Set

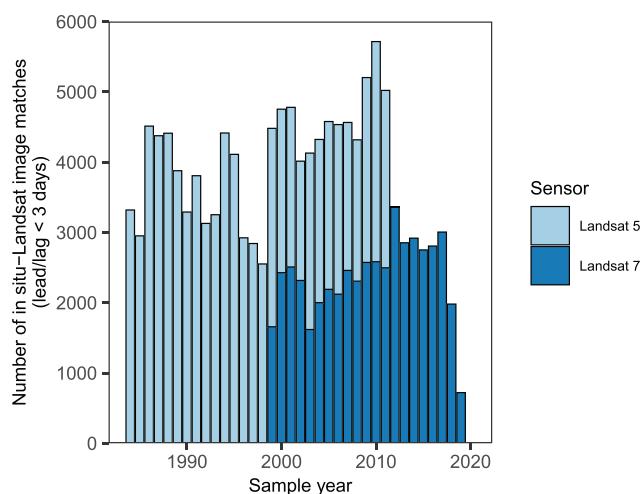
We used the programming language R for all postprocessing and analysis. All the R scripts we used can be found online (<https://github.com/evandethier/satellite-ssc>). We began by joining the in situ and satellite image data sets, using a maximum lead/lag range of  $\pm 8$  days from the in situ sampling date.

station with  $>5$  P63 or POC measurements during the satellite period, we calculated a station-averaged P63 ( $n$  samples = 111,015) and POC ( $n = 48,832$ ), allowing the inclusion of station-averaged POC derived from measurements made by Hilton et al. (2010) for some Taiwanese stations. Each of these metadata parameters was incorporated into calibration model development and/or uncertainty analyses (see section 2.2.5).

### 2.1.2. Image Selection and Sampling Methods

To maximize spatial and temporal coverage of calibration data and application, we restricted our attention to the 36-year Landsat 5 and 7 satellite record. Though the Landsat satellites acquire images less frequently than some other sensors (e.g., MODIS, Sentinel-2, and SPOT), the combined spatiotemporal resolution, consistency, and record length of Landsat imagery is unparalleled. The 30-m multispectral image resolution of Landsat imagery allows for reliable data acquisition from rivers with widths greater than  $\sim 90$  m (Allen & Pavelsky, 2018).

Landsat image analysis was conducted using the Google Earth Engine platform, which hosts the entire Landsat 5 (1984–2012) and 7 (1999–present) imagery archive and allows for rapid, global-scale analyses on spatial and temporal scales that were previously untenable. We used USGS Landsat Tier 1 SR product to generate the calibration data set. The spectral bands available for calibration were B1 (blue, 450–520 nm), B2 (green, 520–600 nm), B3 (red, 630–690 nm), B4 (NIR, 770–900 nm), B5 (Shortwave



**Figure 3.** Number of in situ calibration measurements matched with Landsat images acquired within  $\pm 2$  days for each year in the Landsat 5 (light blue, 1984–2012) and Landsat 7 (dark blue, 1999 to the present) record.

We tested calibration models (described in detail below) using in situ-satellite links with lead/lag of  $\pm 0$ –8 days, evaluating changes in prediction error and bias with increasing lead/lag ranges. As a separate evaluation of sensitivity to lead/lag ranges, we calculated time series autocorrelation at 81 USGS stations with  $>2$  years of consecutive daily SSC measurements (Figures S2 and S3). Based on these tests, we chose an optimal lead/lag range of  $\pm 2$  days from in situ sampling date, producing a data set of 134,697 samples (Figure 3). Using this data set, we generated calibration models by several approaches and tested their varying success and correspondence to each other (see below).

Though the calibration data set does not yet include locations from Europe, Africa, or continental Asia due to the limited availability of public data from these regions, it does include rivers of varying scales (drainage area  $10^3$ – $10^6$  km $^2$ ) and from many climatological and sedimentological settings (Wohl et al., 2015). Variations in river characteristics and sampling methodology in this calibration data set present challenges for generating a globally applicable SSC calibration model. However, the wide range of river characteristics in the calibration data set increases the likelihood of more general prediction model applicability.

## 2.2. Calibration Model Development

### 2.2.1. Base Approach

To accommodate a range of possible relationships between spectral imagery and in situ SSC, we considered multiple possible calibration models relating log-transformed SSC to the surface reflectance of river water. The optimal calibration model was determined by least squares multiple regression, with variables automatically selected from a list of spectral bands and band ratios by cross-validated lasso analysis, a form of multiple regression that eliminates low-power explanatory variables to avoid overfitting and multicollinearity, and is more statistically robust than stepwise regression (Tibshirani, 1996). We used an automated lambda penalty selection, choosing the model with the fewest explanatory variables that was within 1 standard error of the most successful model iteration. We applied this method to develop calibration models for the entire imagery data set described above using several approaches, with each successive approach including additional degrees of spatial or in situ discrimination. In the “base” approach, we developed an optimal global calibration model without any automatic river classification or in situ differentiation. The base approach was used solely as a benchmark for comparison to the four subsequent approaches, respectively, generated after automatic clustering by river optical properties (“cluster” approach), and additional individualization by station in situ SSC data (“station” approach) and/or station grain-size metadata (“grain-size” approach) (see below). We also developed “standalone” models for each station with  $>10$  in situ-Landsat matches, calibrating these with only in situ measurements from a given station.

### 2.2.2. Cluster Approach

To improve our model calibration, the cluster approach accounts for site-specific variability by grouping in situ monitoring stations in the calibration data set with similar optical properties of river water. Because SSC calibration models depend on reflectance in the visible wavelengths, suspended-sediment color affects the relative reflectance in different wavelengths. Methods used by Martinez et al. (2015) and Dethier et al. (2019) were adapted for K-means clustering of all rivers greater than  $\sim 90$  m wide globally. Though the relatively limited Landsat spectral resolution limits the sensitivity of clustering methods relative to those generated using multispectral instruments (Martinez et al., 2015), a combination of bands and band ratios improved cluster differentiation. K-means groups were generated using the multidecadal averages of two spectral bands (B1 and B3) and two band ratios (B2/B1 and B4/B3/B1), computed using the water-only pixels from all images at each station. These bands and band ratios were selected to minimize within-cluster variance by maximizing the Cubic Clustering Criterion (Warren, 1983). Reflectance in the B1 (blue) and B3 (red) wavelengths are an indication of the total reflectance of the pixel. The band ratio B2/B1 (green/blue) indicates the relative “greenness” of the pixel. B4/B3/B1 (NIR/red/blue) indicates the relative reflectance of

B4 and B3 as compared to B1, which may be related to water column stratification (Pavelsky & Smith, 2009). Selecting comparatively few bands minimizes both within-cluster variance and computation expense.

Taken together, these bands and band ratios produce a diagnostic spectral profile at a given location (Figure S7). We tested a range of K-means categories from 1 to 10 for the selected bands and band ratios, and separately generated SSC calibrations for each cluster using all Landsat reflectance bands and band ratios. No in situ SSC or POC data were used for this automated grouping, allowing for application to all global rivers using only the Landsat record. We optimized the number of cluster groups by selecting the calibration model with the lowest total relative error and bias.

After selecting the optimal number of cluster groups, we evaluated typical cluster spectral signatures for different sediment concentrations. Visualization of “true color” (red: B3, green: B2, and blue: B1) and “false color” (red: NIR, green: B3, and blue: B2) color gradients for each cluster and SSC categories of 0–50, 50–100, 100–250, 250–500, 500–750, and >750 mg/L provided reference for the differences among clusters. Analysis of the uncertainty of cluster approach (described further below) indicates the expected error of satellite-derived SSC estimates in the absence of any in situ data from the site where SSC is estimated.

### 2.2.3. Station Approach

Calibration models inevitably result in less accurate SSC estimates for rivers dissimilar to those in the calibration data set, for example, those with unusual vertical SSC profiles or sediment color, and/or high sand concentrations. To evaluate the magnitude of this uncertainty and provide a correction for stations with some in situ SSC samples, we developed a station-specific, or “station” approach. In addition to the spectral band variables used in the Base and Cluster approaches, we added a dummy variable for each station, modifying the cluster approach for lasso regression. The resulting calibration model adjusted the intercept of the calibration model for each station based on the in situ samples from that station.

### 2.2.4. Modified-Station and Standalone Approaches

At all stations with >100 linked in situ sample-Landsat image pairs, we used Monte Carlo simulation to evaluate improvement in calibration resulting from including in model development a specified number,  $N$ , of in situ measurements from the station at which SSC is estimated (from a domain of  $S$  total in situ measurements at the given station). For each of  $N = 1, 2, 3, 4, 5, 10, 20$ , and 50, we developed two satellite calibration models: (1) a “modified-station” model calibrated with a random subsample of  $N$  in situ measurements from the selected station, combined with all the in situ measurements from the other stations in that station’s cluster group and; and (2) a “standalone” model using only the random subsample of  $N$  in situ measurements paired with satellite observations, with no additional in situ data from other stations. The modified-station approach differs from the full station approach in that only a limited subset of the in situ data from the station at which SSC is estimated is used in developing the calibration model. The modified-station approach is designed to assess the additional accuracy gained from incrementally increasing the number of at-the-station in situ data used in developing the calibration model. Comparison of the standalone and modified-station approaches quantifies the relative value of using station data alone to develop a calibration model, versus combining river clustering, in situ data from other rivers, and sparse in situ measurements at a station.

We applied each calibration model (modified-station and standalone) to all  $S$  in situ measurements at the selected station, then calculated the geometric means of the relative error and station bias, respectively. For each station we repeated this Monte Carlo simulation 20 times for each value of  $N$ , then compared the relative error and bias for the standalone and modified-station approaches for all stations. Finally, for all stations with >10 in situ-Landsat matched samples, we used lasso regression to generate a final standalone station model with all available in situ data (with no additional data from other stations). All final standalone station models that met the assumptions of multiple linear regression are published, along with calibration plots (at <https://github.com/evandethier/satellite-ssc>).

### 2.2.5. Additional Variables: Grain-Size and Particulate Organic Carbon Approaches

Several additional river (Dethier et al., 2019; Kilham et al., 2012), SSC/sampling (Bhargava & Mariam, 1991; Chen et al., 1991; Lee & Glysson, 2013; Lurry & Kolbe, 2000; Martinez et al., 2009; Novo et al., 1989), and Landsat-image characteristics (Doxaran et al., 2003; Hudson et al., 2014) may influence the success of calibrations relating water spectral reflectance to SSC. We assembled data sets of these parameters and evaluated their individual impact on SSC calibration success. For samples made contemporaneously with

grain-size and/or POC measurements, we evaluated the impact of including those parameters in both the cluster and station calibration models. Measurements of these parameters were limited, so we also generated calibration models using station-averaged sand fraction and POC fraction and incorporated them as continuous variables in the lasso regression analysis to generate new calibration models for each K-means cluster group.

### 2.3. Estimating SSC and Uncertainty

We assessed the uncertainty in calibration models using holdout sets comprising 25% of the calibration data. That is, we generated the calibration model parameters using 75% of the relevant calibration set data and then used these model parameters to predict SSC for each observation in the remaining 25% of the calibration set data. For the station models, the selection of the in situ holdout data set is random and independent of the selection of the holdout stations. Thus, some (on average 75%) of the in situ data from the station at which SSC is to be predicted is used in developing the calibration model.

Because SSC values span orders of magnitude, reporting error in mg/L suffers from a scaling problem. To address this issue, for each prediction we followed Morley et al. (2018) to calculate the absolute relative error as

$$\text{relative error} = 10^M \left( \log_{10} \left| \frac{\text{SSC}_{\text{predicted}}}{\text{SSC}_{\text{in situ}}} \right| \right) - 1, \quad (1)$$

where  $\text{SSC}_{\text{predicted}}$  and  $\text{SSC}_{\text{in situ}}$  are drawn from the holdout data set not used for calibration and  $M$  denotes the median. This error calculation has the advantage of being readily interpreted (unlike root-mean-square error on log-transformed values) and avoids the problems of asymmetry introduced by using mean absolute error. Because it uses the median, it is insulated against outlier values, which are common in this large data set due to sampling errors, lag between sample time and Landsat image acquisition, image artifacts, and misclassified pixels.

For the base, cluster, station, and grain-size approaches, we followed Overeem et al. (2017) to calculate an average SSC,  $\overline{\text{SSC}}$ , for each station:

$$\overline{\text{SSC}} = \frac{1}{n} \sum_{i=1}^n \text{SSC}_i, \quad (2)$$

where  $n$  is the number of individual SSC estimates during the specified timeframe (i.e., the number of cloud- and ice-free Landsat images of the river reach during the timeframe). We quantified the standard error of  $\overline{\text{SSC}}$ ,  $\varepsilon_{\overline{\text{SSC}}}$ , as the sum, in quadrature, of the relative prediction error and the relative standard deviation of the estimated SSC values,  $\sigma_{\text{SSC}}/\overline{\text{SSC}}$

$$\varepsilon_{\overline{\text{SSC}}} = \sqrt{\varepsilon_g^2 + \frac{\sigma_{\text{SSC}}^2}{\overline{\text{SSC}}}} / \sqrt{n}. \quad (3)$$

For each calibration approach, we also evaluated relative station bias at each of  $m$  stations as

$$\text{relative bias}_m = 10^M \left( \log_{10} \left| \frac{\text{SSC}_{\text{measured}}}{\text{SSC}_{\text{predicted}}} \right| \right) - 1. \quad (4)$$

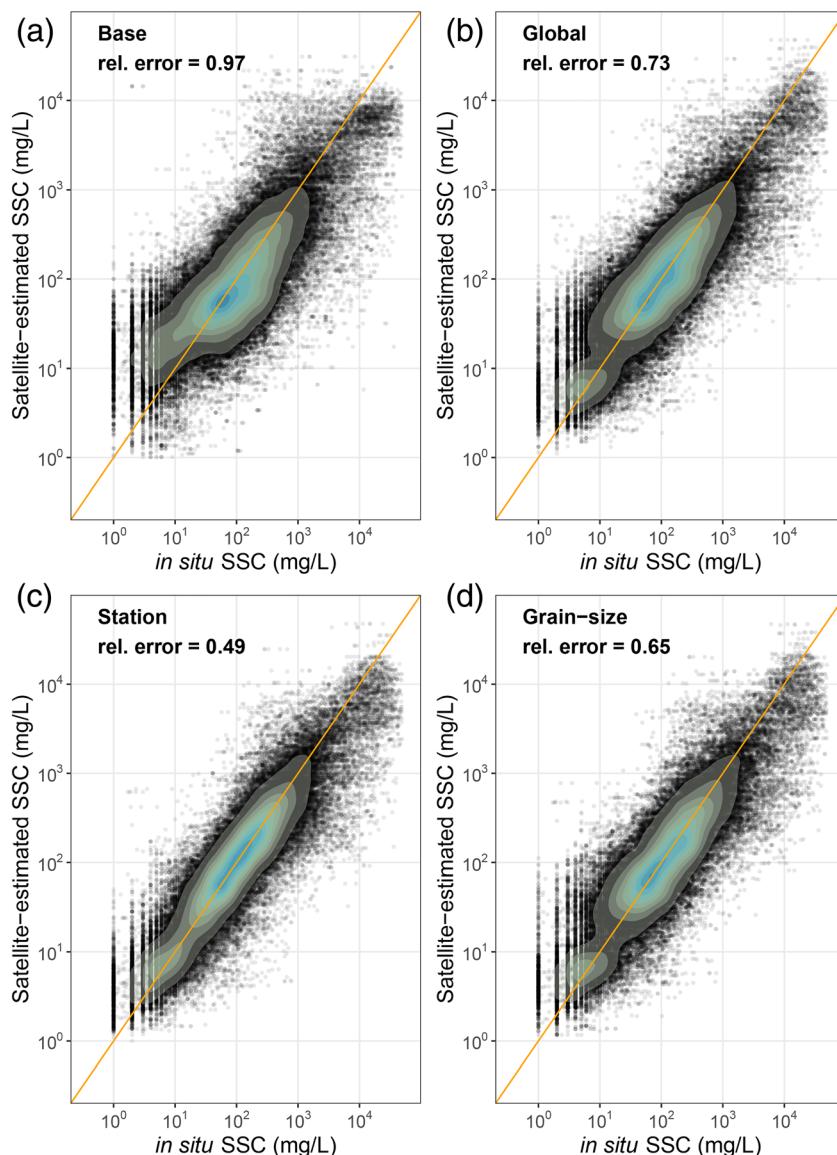
### 2.4. Example Colorado River Application

As a demonstration of our remote sensing methodology, and to explore the potential for spatial extrapolation of the station approach, we applied the cluster and station approaches to the well-monitored Colorado River beginning upstream of Glen Canyon Dam and extending downstream to Grand Canyon, AZ. We used these SSC estimates to track patterns of suspended-sediment transport upstream and downstream of the dam. For purposes of validation, we compared our estimates of SSC to those made in situ by the USGS Grand Canyon Monitoring and Research Center (GCMRC) (USGS, 2019). Measurements from the GCMRC were not incorporated into model calibration for this validation test and thus were a true “test” set. We then evaluated the

**Table 2**

Calibration Uncertainties for the Different Approaches, Given by Relative Error and Bias, for All Data, a Subset of Data From USGS and WCS Stations With Grain-Size Information, and a Subset of USGS and WRA Stations With POC Information

Calibration approach	All data ( <i>n</i> = 134,697)			With sand subset ( <i>n</i> = 111,016)			With POC subset ( <i>n</i> = 48,832)			
	Rel. error	Rel. bias	Rel. error	Rel. error with sand	Rel. station bias	Rel. bias with sand	Rel. error	Rel. error with POC	Rel. station bias	Rel. bias with POC
Base	0.97	0.75	1.04	1.02	0.76	0.68	0.86	0.71	0.64	0.37
Cluster	0.73	0.50	0.71	0.65	0.47	0.38	0.61	0.58	0.40	0.40
Station	0.49	0.08	0.50	0.50	0.07	0.07	0.46	0.46	0.10	0.09



**Figure 4.** Plots of suspended-sediment concentration (SSC) in situ observations (“SSC actual”) versus predictions from multiple regressions for Landsat 5 and 7 sensors (“Estimated SSC”). The orange line is 1:1. The base approach (a) has a relative error = 0.97. The cluster approach (b), with calibration data differentiated by K-means clustering, has a rel. error= 0.73. The station approach (c) includes the individual station name as dummy variable in multiple regression (rel. error= 0.49). The grain-size approach (d) includes station-averaged sand fraction as an additional regressor in the cluster approach (rel. error= 0.65). Contours indicate point density to mitigate information loss due to overplotting.

**Table 3**

Table of Regression Coefficients Determined by Lasso Regression for Each of Six Cluster Groups

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Intercept	-0.4768	1.2156	-1.1399	-0.0441	-0.4352	-0.5023
Variable			Lasso regression coefficients <sup>a</sup>			
B1 <sup>b</sup>	0	0	-2.0 × 10 <sup>-4</sup>	-3 × 10 <sup>-6</sup>	0	-6.0 × 10 <sup>-4</sup>
B2 <sup>b</sup>	0	1.6 × 10 <sup>-4</sup>	0	1.3 × 10 <sup>-4</sup>	0	0
B3 <sup>b</sup>	0	0	1.4 × 10 <sup>-4</sup>	-7.0 × 10 <sup>-5</sup>	4.5 × 10 <sup>-4</sup>	1.3 × 10 <sup>-3</sup>
B4 <sup>b</sup>	3.1 × 10 <sup>-4</sup>	0	0	7.3 × 10 <sup>-4</sup>	6.1 × 10 <sup>-5</sup>	0
B5 <sup>b</sup>	0	0	0	1.0 × 10 <sup>-4</sup>	0	0
B7 <sup>b</sup>	0	2.3 × 10 <sup>-4</sup>	9.0 × 10 <sup>-4</sup>	-1.0 × 10 <sup>-3</sup>	0	0
B1 <sup>2b</sup>	0	-8.0 × 10 <sup>-8</sup>	-1.0 × 10 <sup>-8</sup>	-1.0 × 10 <sup>-7</sup>	-1.0 × 10 <sup>-7</sup>	-5.0 × 10 <sup>-8</sup>
B2 <sup>2b</sup>	0	0	0	0	-7.0 × 10 <sup>-21</sup>	-2.0 × 10 <sup>-19</sup>
B3 <sup>2b</sup>	0	0	0	0	-4.0 × 10 <sup>-21</sup>	-1.0 × 10 <sup>-7</sup>
B4 <sup>2b</sup>	0	0	0	0	-1.0 × 10 <sup>-21</sup>	-1.0 × 10 <sup>-11</sup>
B5 <sup>2b</sup>	0	0	0	0	-8.0 × 10 <sup>-22</sup>	-4.0 × 10 <sup>-20</sup>
B7 <sup>2b</sup>	0	0	0	0	-1.0 × 10 <sup>-8</sup>	-2.0 × 10 <sup>-20</sup>
B2/B1	0	0	0.849	0.457	0	-0.219
B3/B1	0	-0.244	-1.120	-0.193	0.710	1.163
B4/B1	0	0	0.287	-0.695	0	0.116
B5/B1	0	0	0	0	0	0
B7/B1	0	0	0.230	0.280	-0.003	-1.103
B3/B2	2.555	0.928	2.910	1.382	0.616	0
B4/B2	0	1.039	0.421	0.061	0	0
B5/B2	0	0	-0.801	2.693	0	0
B7/B2	0	2.234	0	0.392	0	0
B4/B3	0	-0.567	-0.208	1.136	-0.097	0
B5/B3	-0.157	-0.877	-0.372	-3.413	0	0.575
B7/B3	0	-1.959	-0.098	0	-0.072	0.060
B5/B4	-0.445	0.006	0.271	0	0	0
B7/B4	0	0.079	0	0.131	0	0
B7/B5	0.030	0	0	0.081	0	0.294

Note. The form of the prediction equation for a measurement in Cluster  $i$  is  $\text{Log}_{10}[\text{SSC } (\text{mg/L})] = \text{Intercept } (\text{Cluster}_i) + \text{Variable 1 value} \times \text{Variable 1 Coefficient}_i + \text{Variable 2 value} \times \text{Variable 2 Coefficient}_i + \dots + \text{Variable } N \text{ value} \times \text{Variable } N \text{ Coefficient}_i$ . Lasso regression automatically sets low-power regressors to 0 to minimize overfitting.

<sup>a</sup>Lasso regression automatically sets low-power regressors to 0. <sup>b</sup>Landsat surface reflectance from USGS/NASA LT1 product. Reflectance is scaled by a factor of 10,000.

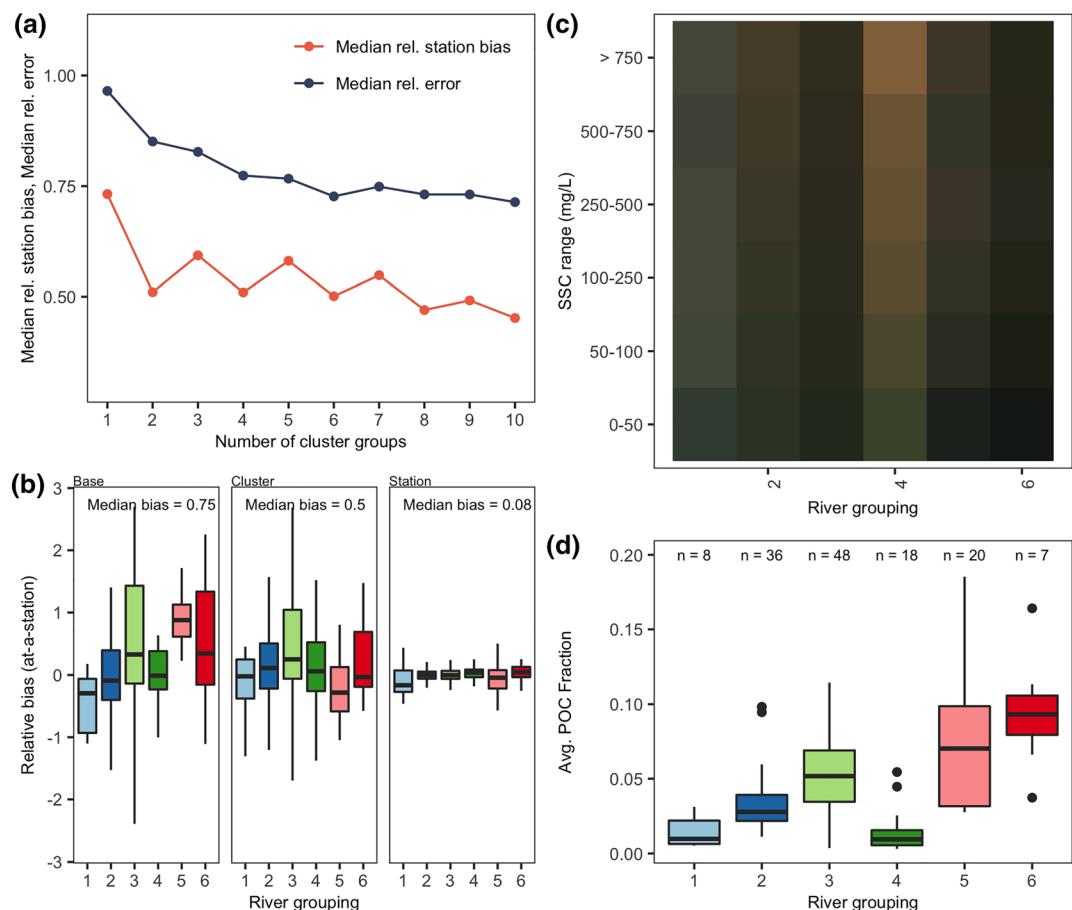
marginal improvement in relative error and relative station bias gained by using station data approaches developed by incorporating in situ data from each station in the GCMRC network.

### 3. Results

#### 3.1. Calibration Between In Situ and Estimated SSC

We expanded existing Landsat-in situ SSC calibration sets by more than an order of magnitude, compiling the largest linked satellite image-suspended-sediment-measurement calibration data set to date. Of a data set of more than 1.7 million in situ SSC measurements made at 727 suspended-sediment sampling locations, 134,697 samples encompassing the entire satellite record could be linked to high-quality satellite data of the sampling location within  $\pm 2$  days of the sampling date (Figure 3).

Each primary approach for SSC calibration—base, cluster, and station—shows a statistically significant ( $p < 0.0001$ ) relationship between the spectral properties of river water and SSC (Table 2 and Figure 4). In general, red and NIR wavelengths (B3 and B4), often normalized by B1 or B2, have the strongest positive correlations with SSC, but the influence of individual parameters varies by model and cluster (Figure S10). With the exception of grain-size additions to the cluster model, each successive approach for differentiation in model development improves prediction accuracy. All coefficients incorporated in each model are highly significant, and model residuals are normally distributed. A single river pixel is sufficient for estimating SSC, though prediction model uncertainty decreases slightly with additional sampled pixels (Figure S4).



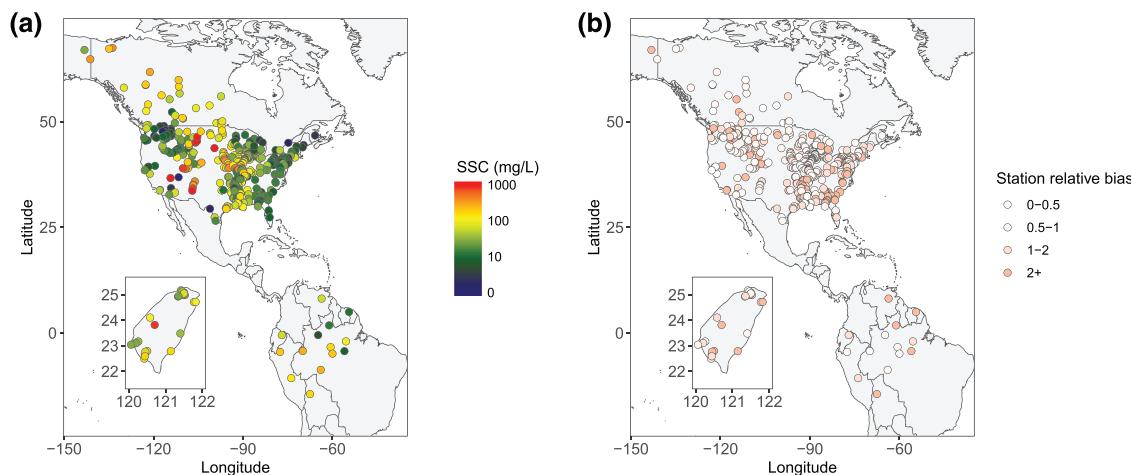
**Figure 5.** (a) Calibration model relative error and station bias for each number of clusters from 1–10. (b) Distribution of station-averaged model relative bias sorted by cluster for the base, cluster, and station prediction approaches. (c) Median “true color” appearance of SSC in six concentration ranges for each cluster group, highlighting the differences among rivers in different groupings. (d) Boxplots of cluster-averaged POC fraction show a relationship with typical river color, with yellow/gray SSC typically associated with low POC fractions, and gray/green/brown associated with high POC fractions. The number of stations with POC measurements is indicated at the top of the plot ( $n = X$ ). Boxplots show the interquartile range and median value, and whiskers extend to the range of the data absent outliers, which are excluded for clarity in (b) and shown in (d).

### 3.1.1. Cluster Approach

Automated K-means clustering resulted in optimal model success for six cluster groups based on calibration relative error and station bias (Table 3 and Figures 5a and 5b). These groups have distinct spectral profiles (Figure S7) and color gradients with increasing SSC (Figures 5c, S8, and S9), which is related to average particulate organic carbon (Figure 5d). The cluster approach results in improved model uncertainty of 25% over the base approach (Table 2 and Figure 4), largely due to a 33% reduction in the relative station bias, particularly for cluster Groups 1, 5, and 6 (Table 2 and Figure 5b). No readily predictable disparity in relative bias exists among cluster groups (Figure 6). We report here and publish on <https://github.com/evandethier/satellite-ssc> the multiple regression models, which are easily deployed in R and readily adapted for Google Earth Engine and other platforms.

### 3.1.2. Station, Modified-Station, and Standalone Approaches

The station approach, which incorporates a dummy variable for each station in addition to using the cluster groups algorithm, reduces uncertainty by 33% and relative bias by 85% in comparison to the cluster approach (Table 2 and Figure 4). This approach is less globally applicable because it requires in situ measurements, which are not available for all rivers. However, the value added is twofold: (1) monitoring at a given station can be expanded temporally to augment the in situ record, if that record has less temporal coverage than the Landsat archive, and (2) monitoring along the river can be extended spatially, provided



**Figure 6.** (a) Average suspended-sediment concentration (SSC) as estimated by the cluster approach. (b) Corresponding relative bias in average SSC estimates with respect to in situ SSC average. Averages are over the entire Landsat database (1984 to the present).

that the river optical properties and/or SSC transport characteristics do not change dramatically along the reaches in question.

Critically, the improvement associated with using both the clustering and paired in situ measurements in the modified-station approach requires only limited in situ data from the station of interest for significant improvement in the calibration accuracy. Indeed, results from the Monte Carlo simulations (Figure 7) show that, on average, the modified-station approach outperforms the standalone station models developed using only paired in situ data from the station when the number of in situ samples at the station  $N < 20$  and performs comparably or better when  $N > 20$ . The modified-station approach with as few as five in situ samples from the station of interest is sufficient to construct a calibration model with similar accuracy to standalone calibration models developed using far more in situ SSC data.

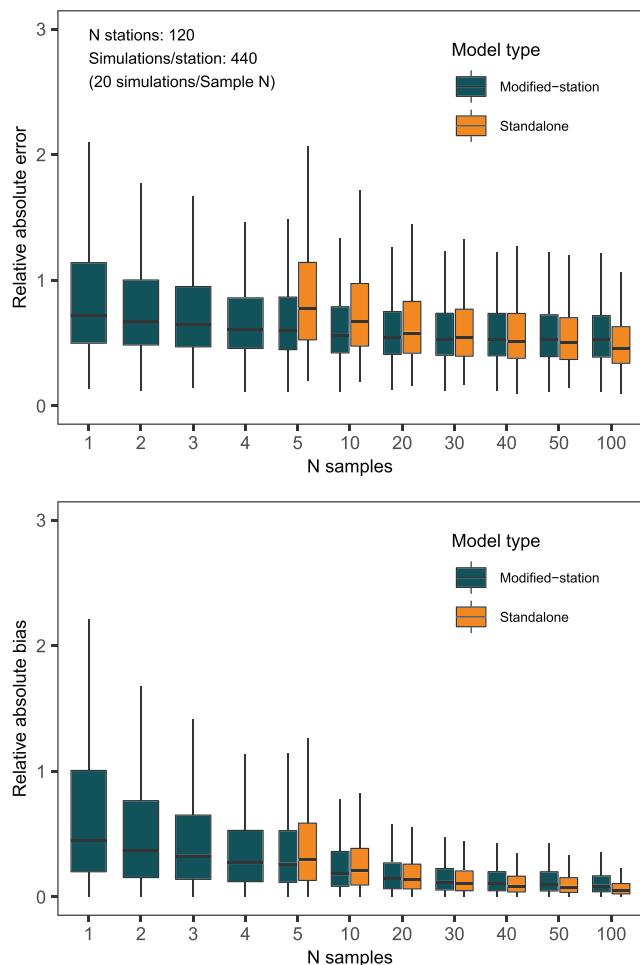
### 3.1.3. Grain-Size and POC Approaches

Grain-size variations likely account for a fraction of the increased success of the station approach that incorporate in situ SSC measurements. Average cluster model prediction bias asymptotically decreases with increased P63 until it stabilizes at  $\sim 50\%$  silt (Figure 8), as the calibration models systematically underpredict SSC in systems with abundant sand. Cluster-to-cluster variations in this relationship are minimal (Figure S5). Average sand fraction is not a critical additional parameter for many systems, likely because most large rivers have average P63  $> 50$  (Figure 9). In the absence of in situ SSC measurements, including the average percent sand improves the cluster model calibration by only 9% (Table 2).

Model accuracy improvement with the inclusion of P63 was better ( $> 15\%$  in some cases) at sites with abundant sand in the Mississippi River, Rio Grande, and Colorado River watersheds. However, in almost all cases, adding in situ measurements of SSC from the station in question—in essence, tuning the cluster calibration to a given station—produced equivalent or greater improvement in uncertainty than adding percent sand data.

Average prediction bias is also correlated with POC fraction, with increasing POC leading to overprediction in SSC (Figure 10). Fewer in situ SSC sampling stations have associated measurements of POC, and the cluster calibration model for in situ samples from those stations performs on average better than for all in situ samples. When station-averaged POC is included as an explanatory variable, model improvement is somewhat less than that due to P63, with a 4% reduction in relative error (Table 2).

Once the dummy variable distinguishing stations is included in the station model regression, the addition of either percent sand or POC yielded no additional improvement in the relative error. Including daily, rather than station-averaged, P63 or POC data also does not on average improve model performance beyond that gained from in situ SSC data alone. Because of the sparseness of daily grain-size and POC data (Table 1), these findings regarding model improvement with additional daily metadata are difficult to assess on a global scale.



**Figure 7.** Results of a simulation of at-a-station relative error and relative station bias in predicted suspended-sediment concentration (SSC) at 120 stations with  $>100$  in situ-Landsat matches. The modified-station approach (dark symbols) outperforms individual station calibrations (orange symbols) when the number of paired in situ samples at the station  $N < 20$ , and performs comparably or better when  $N > 20$ . As few as 5–10 in situ samples is sufficient to construct a calibration model with similar accuracy to the station calibration models developed using all available in situ SSC data.

due to direct human intervention and anthropogenic climate change have accelerated since the end of the twentieth century (Best, 2019), and robust monitoring techniques are key to understanding the likely sources and impacts of these changes, as well as the many fundamental theoretical questions in geomorphology that remain unanswered because of sparse spatiotemporal data. For the many global rivers with sparse or no monitoring, satellite-derived measurements may provide critical insight about patterns and trends in SSC that cannot be obtained without historical and/or spatially extensive data. These techniques offer the potential to supplement with current data monitoring programs that have been recently shuttered, particularly in the United States and Canada (Figure 12) (Warrick & Milliman, 2018), and to provide historical context for incipient monitoring programs elsewhere.

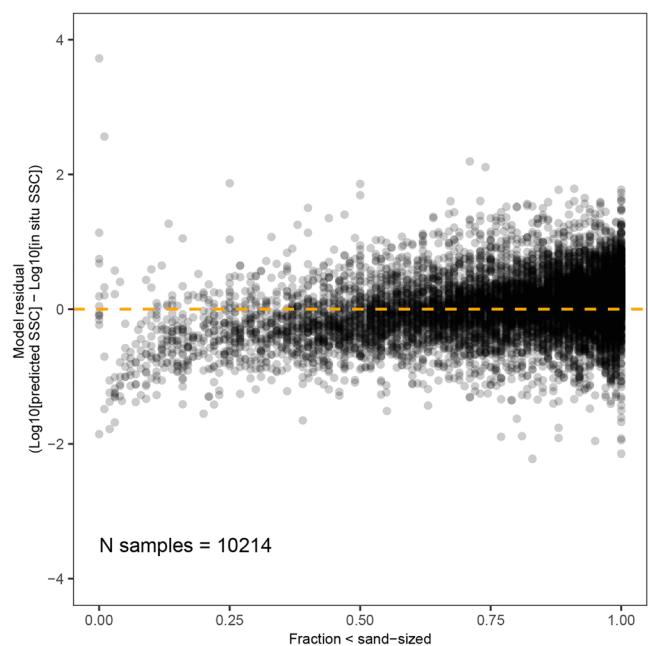
These methods can be applied to all global rivers greater than  $\sim 90$  m wide, provided that caution is taken when interpreting results due to model uncertainty. We note that our assessments of uncertainties are based on the assumption that the in situ data are accurate. As many different investigators and agencies, using different methods, collected the in situ data, it is likely that some amount of the discrepancy in our results is due to unreliability of those measurements. Estimates of the uncertainty of the in situ

### 3.1.4. Example Application on the Colorado River

As an illustrative application, we extracted SSC estimates using our cluster approach (clustering but no in situ data) and station approach (clustering and in situ data) at sediment monitoring stations along the Colorado River downstream of Glen Canyon Dam (Figure 11). Despite challenges presented by narrow canyon reaches at the lower limit of our  $\sim 90$  m width requirement, and varying seasonal shadowing due to changing Sun angles, the accuracy of our SSC estimates (compared to the in situ measurements by the GCMRC) is similar to the average global accuracy of this approach (relative error = 0.77, avg. rel. station bias = 0.20 for the GCMRC sites versus average global accuracy of the cluster approach of relative error = 0.73, rel. station bias = 0.49). Grain size also has a similar effect on model performance as observed with the global data set, with systematic underprediction of SSC when sand fraction exceeds 0.5. Prediction uncertainty is improved by the addition of in situ data from each station in the station approach (relative error = 0.66, avg. station rel. bias = 0.11) for the GCMRC stations versus the average global accuracy of the station approach (relative error = 0.49, rel. station bias = 0.08). Both the cluster and station approach estimates align with field-based research (Hazel et al., 2006; Topping et al., 2000), showing the progressive downstream recovery of SSC due to sediment inputs from the Paria and Little Colorado Rivers (Figure 11). These data additionally show the shifts in seasonal SSC timing associated with disruption of the hydrograph by regulation. Only in the Grand Canyon, some 170 km downstream of Lake Powell, does the Colorado River begin to recover the springtime high-SSC peaks that are apparent 280 km upstream of the dam, yet these peak concentrations are still diminished by  $\sim 50\%$  relative to upstream ( $t$  test,  $p$  value  $< 0.05$ ). Seasonal variations in model prediction accuracy are a function of seasonal shifts in SSC sand fraction (Figure S11) and shadowing due to steep canyon walls (Figure S12).

## 4. Discussion

The remote sensing approaches for determining SSC presented here allow for spatially continuous measurements with sufficient temporal and spatial resolution to detect trends in time and space. Our methods also allow for the quantification of uncertainty and identify clear paths to reduce that uncertainty. Modification of river processes



**Figure 8.** Calibration models show typical underprediction of SSC when sand fraction is high, with an asymptotic approach toward no bias with increasing fraction less than sand sized.

SSC measurements, the median relative error in our SSC estimates using the cluster approach is 74%. For many applications, uncertainty of this scale (less than a factor of 2) may be acceptable, particularly when detecting large spatial or temporal changes. For instance, SSC estimates along the Colorado River derived using the cluster approach successfully identified the seasonal and longitudinal patterns of SSC, including both the reduction due to Glen Canyon Dam and the subsequent downstream recovery (Figure 11, light blue boxes).

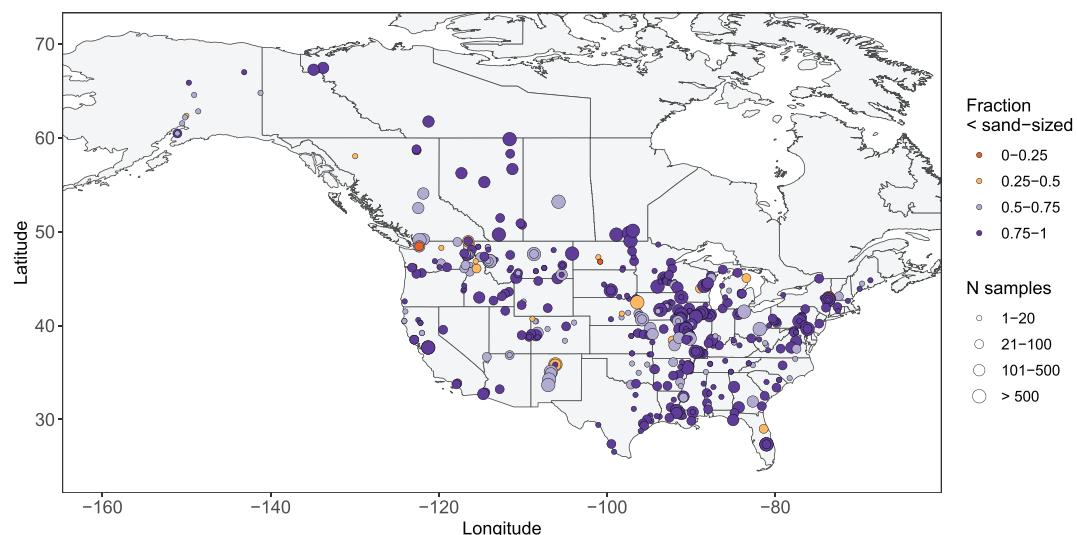
Inclusion of station-averaged sand fraction grain-size or POC does not improve calibration models beyond that gained from using in situ SSC measurements alone. Because we focus on rivers  $>90$  m wide,

measurements are often not reported and, when they are, vary widely (Lee & Glysson, 2013; Syvitski et al., 2003). However, analysis of different sampling methods indicates no systematic bias except for typical overprediction of surface suspended measurements, likely because the models are calibrated mostly with measurements that were not solely obtained at the surface (Figure 13). Similarly, average depth of depth-integrated measurements and point samples with sample depth reported as metadata do not indicate systematic prediction bias (Figure S6).

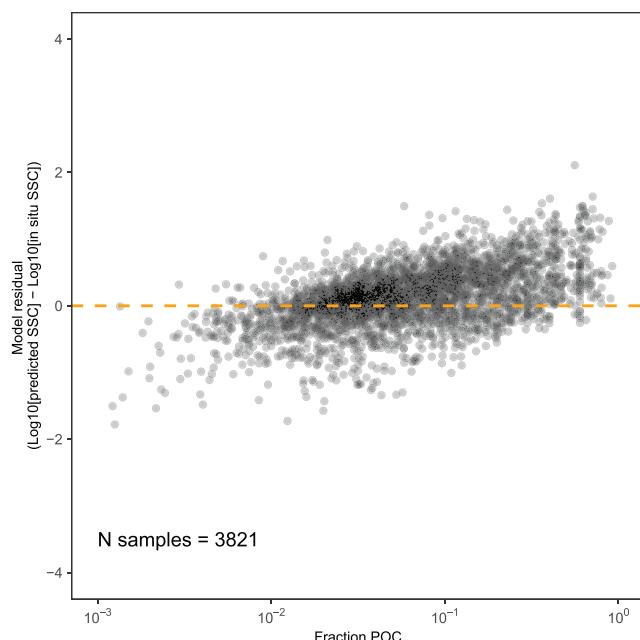
In addition, the temporal resolution of the Landsat record is not suitable for all applications. In cases where high-temporal-resolution data are required, the combination of maximum approximately weekly return period and the inability to detect SSC through cloud cover may not be sufficient. However, these Landsat-based methods offer the longest globally consistent potential for SSC estimation, as they are predicated on two nominally identical satellite sensors operational from 1984 to the present.

#### 4.1. Strengths and Limitations of Calibration Models

It is clear from the reduction in model uncertainty and relative bias achieved by K-means cluster differentiation that incorporating inter-river or intrariver characteristics is a useful step in model development, as has been found for individual rivers by others (Dethier et al., 2019; Martinez et al., 2015). Absent any site-specific in situ



**Figure 9.** Average fraction of suspended sediment that is finer than sand (USGS parameter P63), as measured at USGS and WCS stations used in calibration model development.



**Figure 10.** Calibration models show a relationship between increasing POC fraction and overprediction of SSC.

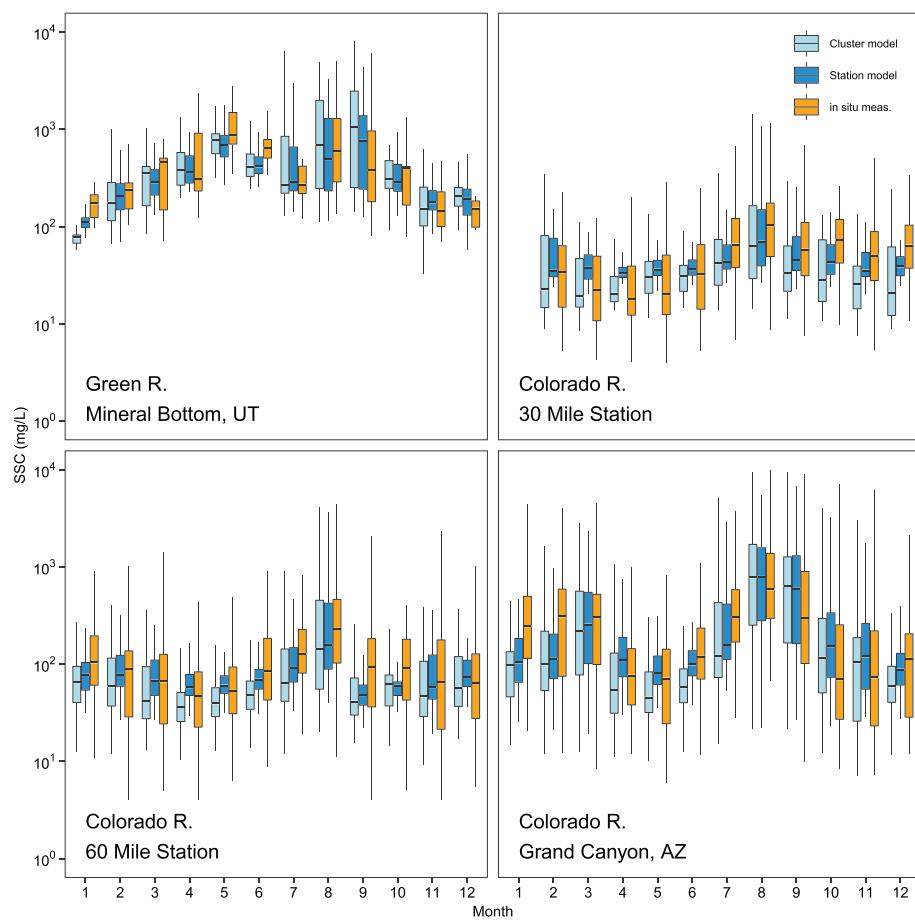
grain-size variations are likely less pronounced than if imagery resolution allowed for detection of SSC on smaller rivers. For example, of those rivers in the calibration data set for which grain-size information is available, in over half of the rivers 80% of the sediment is silt-sized or finer (Figure 9). At stations where incorporating grain-size information does meaningfully reduce station bias relative to the cluster calibration approach, our results indicate that making more straightforward in situ measurements of SSC is an easier means for achieving the same or better result. Inclusion of the station average POC is less impactful than including grain-size measurements, likely because the K-Means clustering that underpins both cluster and station approaches is sensitive to POC (Figure 5d). Incorporating daily, rather than average, values of sand fraction and POC does significantly improve calibration models for some stations. However, the addition of these measurements is of questionable utility in model development, because daily in situ grain-size and/or POC measurements require daily in situ SSC measurements that would render satellite-derived SSC measurements redundant. Instead, daily measurements of these parameters primarily gives context to calibration uncertainty measurements; it is clear that satellite-based studies of rivers with significant sand or organic content are likely to substantially underpredict SSC.

For locations of interest that have in situ data or where investigators intend to collect it, two primary approaches are available: (1) generate a model developed using local in situ data combined with global in situ data from the same cluster group (station approach, reported here for all stations included in the calibration data set) or (2) generate a model using only in situ data local to the station of interest. We discuss below the merits of each approach.

Even sparse in situ measurements can be leveraged in the station approach to dramatically reduce average relative error and relative station bias in average SSC. Similar to Long and Pavelsky (2013), we found that using site-specific corrections substantially improve the success of calibration models applied to other water-bodies. When such a correction is applied, the use of a station model allows for the temporal extension of at-a-station SSC estimates beyond the in situ record. These findings have implications for future sampling campaigns designed to monitor SSC. Since even sparse calibration data can improve the accuracy of remote sensing estimates of SSC by 46% and little additional accuracy is gained beyond 10–20 paired in situ measurements (Table 2 and Figure 7), sampling campaigns should prioritize limited measurements from a greater variety of rivers rather than extensive measurements at fewer sites. In cases where depth integration methods are expensive and/or labor intensive, developing an additional calibration relating surface to depth-integrated SSC may allow for the substitution of surface SSC sampling (Yepez et al., 2017).

Our results provide a cautionary warning to attempts to build calibration models based on only a few station-specific paired in situ and satellite observations. Calibration models based only on fewer than 20–50 paired measurements from a given station do not always perform well (Figure 7), particularly if the distribution of SSC measurements is narrow and/or when applied to river reaches not included in calibration (Long & Pavelsky, 2013). Our claim that only a few in situ measurements are needed to construct an accurate calibration model is only true if the in situ measurements are combined with our cluster algorithm and in situ data from other rivers (our station approach). The cluster algorithm leverages information from other similar rivers, allowing for an accurate calibration even with just a few site-specific measurements.

For some in situ stations, a standalone calibration approach made solely with >20–50 in situ measurements from that station outperforms a station approach calibration that incorporate measurements from other stations. These standalone calibrations are successful when in situ calibration measurements span the likely range of SSC values for that station. We have generated standalone regression models for 463 stations; those calibrations that satisfy the assumptions for multiple regression can be found at <https://github.com/evandethier/satellite-ssc> and thus can be utilized by researchers studying those rivers.

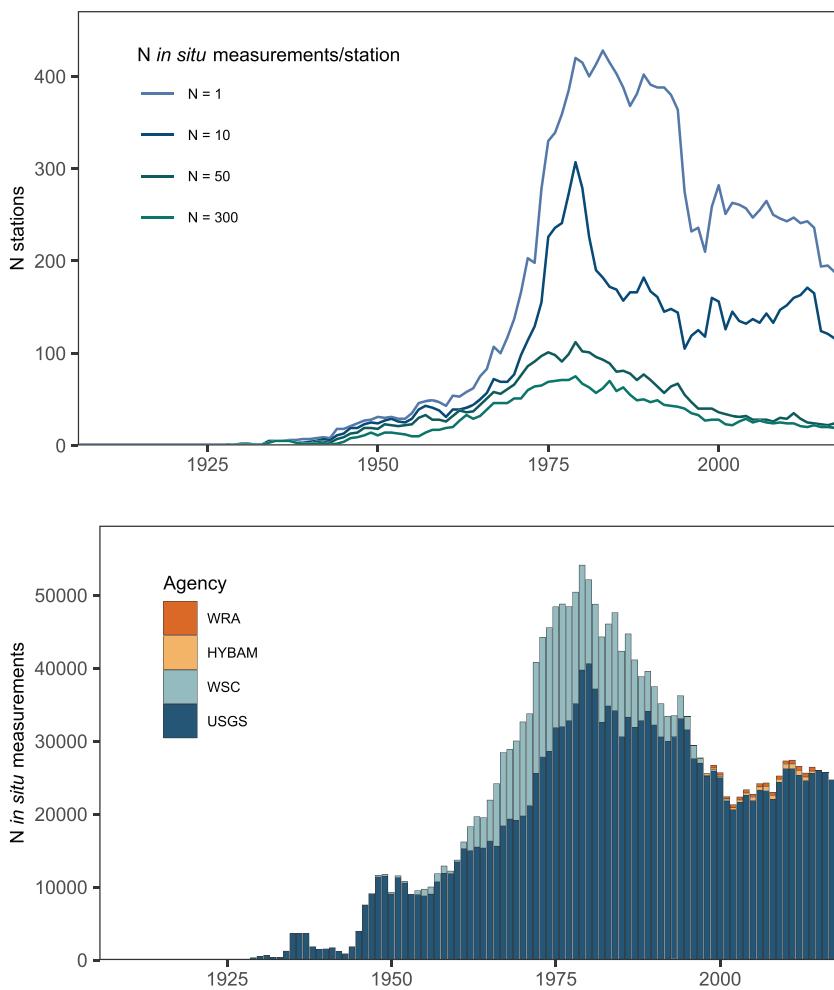


**Figure 11.** Seasonal suspended-sediment concentration (SSC) at the Green River, the best monitored upstream source river to Lake Powell, and on the Colorado River downstream of the dam (30 Mile Station, 60 Mile Station, Grand Canyon, AZ) as estimated by the station approach (light blue symbols), the cluster approach (blue symbols), and as measured in situ by the Grand Canyon Monitoring Research Center (yellow symbols).

As Long and Pavelsky (2013) also reported, we find that caution should be used when transferring a model generated using a station or standalone approach to make predictions of SSC outside the range of in situ calibration measurements, applying the model to Landsat reflectance spectra included in the calibration data set, and to transferring the calibration model to other river reaches. In both the station-specific model approaches, spatial extrapolation of station model predictions is best when limited to reaches lacking significant disruptions in sediment flux due to, for example, major tributary inputs. Because clustering reflects a major shift in sediment optical characteristics, spatial extrapolation of such a model to river reaches in a different cluster should be avoided, particularly in the standalone approach, when the calibration model does not incorporate in situ measurements from other stations. When estimating SSC across multiple cluster groups, using a combination of the station-specific model and cluster approach model(s) may allow for greater spatial extrapolation with only a modest decrease in accuracy.

#### 4.2. Next Steps

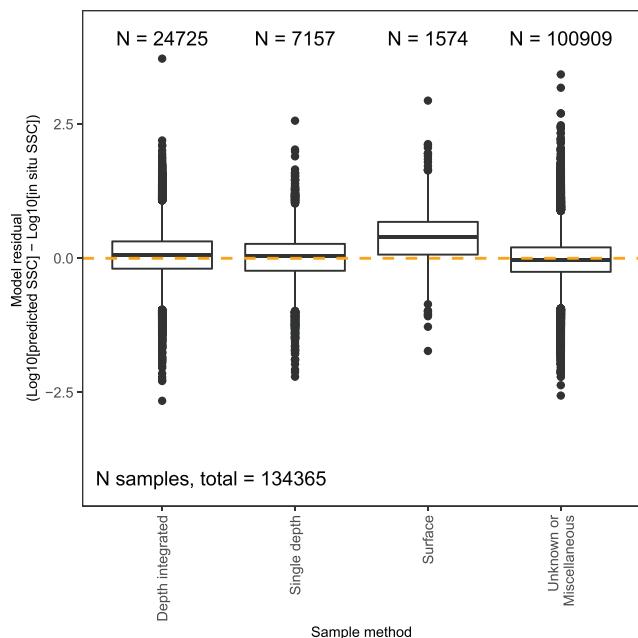
Though our calibration data set includes rivers of varying scales and from many climatological and sedimentological settings, it does not include locations from Europe, Africa, or continental Asia. However, the distribution of cluster types in these regions is similar to those in regions covered by our calibration data set and thus we have no a priori reason to suspect our cluster approach should be any less accurate in these regions. Further refinement of these approaches with additional in situ data is a constantly ongoing process. We continue to seek additional calibration data sets both for increased global coverage and to develop calibrations for newer satellite sensors.



**Figure 12.** Suspended-sediment in situ monitoring at sites suitable for in situ-Landsat SSC calibration (river width greater than  $\sim 90$  m) has declined from a peak in the 1980s–1990s in the United States (USGS) and Canada (WSC), both as reflected in number of stations (upper panel), and total number of samples taken (see also Warrick & Milliman, 2018). In particular, the number of stations making few measurements per year (<50) has declined dramatically, leading to a less cosmopolitan potential calibration data set. Increases in sampling in South America (HYBAM) and Taiwan (WRA) were important for calibrations presented here but only represent a fraction of USGS and WSC measurements.

The Landsat satellites have collected high-quality data continuously since 1984. Other sensors suitable for SSC analysis suffer from substantially shorter records: for example, MODIS has been operational since only 2000, and Sentinel-2 since 2016. Though we have developed calibration models for the Landsat 8 sensor, we do not use results of these models in our current analysis because the Landsat 8 sensor is different than that used on Landsat 5 and 7, and SR products are still in development. In addition, since the Landsat 8 sensor is relatively new, there are an order-of-magnitude fewer in situ calibration measurements than for the other sensors. Sensor fusion has the potential to increase the spatial and temporal coverage of satellite-based SSC estimates, but significant care must be taken to address issues of continuity among sensors.

Both for further development of the Landsat 5 and 7 calibrations, and especially for newer sensor calibrations, adding in situ samples from additional global stations will limit our dependence on matching in situ samples to satellite images not acquired on the same day. Limiting lead/lag in generating a calibration data set of matched in situ and satellite measurements is optimal for reducing errors due to rapid changes in SSC. Our lead/lag limit of  $\pm$  two days is more conservative than some other studies (Dethier et al., 2019; Montanher et al., 2014) and adequate for large rivers that do not change as rapidly as smaller, more flashy



**Figure 13.** Box-and-whisker plots of model residuals plotted for each major suspended-sediment sampling category. Of the different sampling methods, only the surface samples show a systematic bias, with typical overprediction of SSC. Boxplots show the interquartile range and median value, and whiskers extend to the range of the data absent outliers, which are shown as points.

fication, though we do not observe systematic bias in estimating SSC for depth-integrated samples. Using depth-integrated in situ data to correct cluster approach predictions reduces this uncertainty. All code and workflows are available at <https://github.com/evandethier/satellite-ssc> so that other researchers may implement and/or modify them for locations of interest. Continued improvement of these calibrations depends on increased availability of in situ measurements.

## 5. Conclusions

The geospatial focus of our remote sensing estimates of SSC comes at a time of transition for global monitoring of river material transport. Long-standing monitoring programs in the United States have recently been terminated due to budget or logistical constraints (Lee & Glysson, 2013; Warrick & Milliman, 2018), and much of the observational data that are collected outside of the United States are held privately or are decentralized and difficult to access. Further, the synthesis of different data sets is challenging, as the methods by which agencies and private groups collect suspended sediment vary widely around the world. Numerous long-standing geomorphic questions remain unanswered because available data are limited in spatial and/or temporal resolution and extent. Our remote sensing approach, validated by 134,697 ground truth measurements, addresses many of these challenges with a quantified accuracy. This level of uncertainty (73% relative error in the absence of any localized in situ data) may not be sufficient for answering all questions. In some cases, just a few paired in situ and satellite observations may be sufficient to reduce the uncertainty to acceptable levels (50% relative error and 7% relative station bias). These estimates assume the in situ data are accurate; the uncertainty of in situ measurements is often not reported and, when it is, varies widely. While satellite-derived estimates of SSC will likely never be more accurate than careful in situ measurements, we note that the uncertainty of remotely sensed estimates of SSC is offset by their spatial and temporal resolutions that are far greater than can be achieved by in situ measurements alone. In many cases the trade-off of lesser accuracy for greater spatial and temporal coverage may justify the use of remotely sensed estimates of SSC.

systems (Figure S2). In the current form of the Landsat 5 and 7 calibration data set, allowing some lead/lag introduces more calibration data from sparsely sampled sites and regions without adding significantly to model uncertainty. In particular, adding a lead/lag range allowed for increased contribution from South America and Taiwan, where samples are taken approximately monthly, as well as the stations in the United States and Canada that did not measure SSC every day. In allowing imperfect matches, we were able to increase the size of our calibration data sets in these regions by approximately a factor of four. As the number of stations in this data set making daily measurements continues to decline below 50/year (Figure 12), calibration will be increasingly challenging, especially for smaller, flashier rivers (Figure S2). Without continued intensive in situ sampling, new sensors offering increased spatial resolution, and thus the potential for better SSC detection on small rivers, will be undercalibrated and/or underutilized.

These methods are ideally suited for detecting longitudinal and temporal changes in SSC with important geomorphic applications in helping quantify, for example, the impacts of dams and the spatial extent of downstream sediment disconnectivity; important channel-floodplain exchanges; and/or the effects of land use/land cover change. The parallel processing power and scalability of the Google Earth Engine platform allows rapid sampling along transects and/or locations of interest, creating time series spanning the entire 36-year Landsat 5 and 7 record. The accuracy of annual and/or decadal sediment flux is potentially limited by the temporal resolution of Landsat data and uncertainty in water column strati-

## Data Availability Statement

All of the data we use in this work are publicly available. Google Earth Engine and R code used in the methods is included in its entirety online (<https://github.com/evandethier/satellite-ssc>). We are indebted to the NASA Landsat missions (<https://landsat.usgs.gov>), which can be accessed through the Google Earth Engine platform (<http://earthengine.google.com>) as described in the text and github link, or alternatively through <http://ear�explorer.usgs.gov> website. We cite the sources of in situ suspended-sediment data in the text, and these can also be downloaded and analyzed as described in the text using the code online (<https://github.com/evandethier/satellite-ssc>).

## Acknowledgments

Thank you to Hannah Rubin, Olivia Lantz, Shannon Sartain, Greta Bolinger, and Sydney Stearns for their contributions to this project. This work was supported by the Dartmouth College Earth Science Department, and National Science Foundation Grants EAR-1545623 and BCS-1636415.

## References

- Albers, S. (2017). Tidyhydat: Extract and tidy Canadian hydrometric data. *Journal of Open Source Software*, 2(20), 511.
- 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>
- Best, J. (2019). Anthropogenic stresses on the world's big rivers. *Nature Geoscience*, 12(1), 7–21. <https://doi.org/10.1038/s41561-018-0262-x>
- Bhargava, D. S., & Mariam, D. W. (1991). Light penetration depth, turbidity and reflectance related relationships and models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46(4), 217–230. [https://doi.org/10.1016/0924-2716\(91\)90055-Z](https://doi.org/10.1016/0924-2716(91)90055-Z)
- Chen, Z. M., Hanson, J. D., & Curran, P. J. (1991). The form of the relationship between suspended sediment concentration and spectral reflectance—Its implications for the use of Daedalus 1268 Data. *International Journal of Remote Sensing*, 12(1), 215–222. <https://doi.org/10.1080/01431169108929647>
- Chu, V. W., Smith, L. C., Rennermalm, A. K., Forster, R. R., & Box, J. E. (2012). Hydrologic controls on coastal suspended sediment plumes around the Greenland Ice Sheet. *The Cryosphere*, 6(1), 1–19. <https://doi.org/10.5194/tc-6-1-2012>
- De Cicco, L., Lorenz, D., Hirsch, R., & Watkins, W. (2018). dataRetrieval: R packages for discovering and retrieving water data available from US federal hydrologic web services, edited.
- Dethier, E. N., Sartain, S. L., & Lutz, D. A. (2019). Heightened levels and seasonal inversion of riverine suspended sediment in a tropical biodiversity hot spot due to artisanal gold mining. *Proceedings of the National Academy of Sciences of the United States of America*, 116(48), 23,936–23,941. <https://doi.org/10.1073/pnas.1907842116>
- Doxaran, D., Froidefond, J. M., & Castaing, P. (2003). Remote-sensing reflectance of turbid sediment-dominated waters. Reduction of sediment type variations and changing illumination conditions effects by use of reflectance ratios. *Applied Optics*, 42(15), 2623–2634. <https://doi.org/10.1364/AO.42.002623>
- Filizola, N., & Guyot, J. L. (2009). Suspended sediment yields in the Amazon basin: An assessment using the Brazilian national data set. *Hydrological Processes*, 23(22), 3207–3215. <https://doi.org/10.1002/hyp.7394>
- Hazel, J. E., Topping, D. J., Schmidt, J. C., & Kaplinski, M. (2006). Influence of a dam on fine-sediment storage in a canyon river. *Journal of Geophysical Research*, 111, F01025. <https://doi.org/10.1029/2004JF000193>
- Heege, T., Kiselev, V., Wettle, M., & Hung, N. N. (2014). Operational multi-sensor monitoring of turbidity for the entire Mekong Delta. *International Journal of Remote Sensing*, 35(8), 2910–2926. <https://doi.org/10.1080/01431161.2014.890300>
- Hilton, R. G., Galy, A., Hovius, N., Horng, M. J., & Chen, H. E. (2010). The isotopic composition of particulate organic carbon in mountain rivers of Taiwan. *Geochimica et Cosmochimica Acta*, 74(11), 3164–3181. <https://doi.org/10.1016/j.gca.2010.03.004>
- Hudson, B., Overeem, I., McGrath, D., Syvitski, J. P. M., Mikkelsen, A., & Hasholt, B. (2014). MODIS observed increase in duration and spatial extent of sediment plumes in Greenland fjords. *The Cryosphere*, 8(4), 1161–1176. <https://doi.org/10.5194/tc-8-1161-2014>
- HYDAT (2019). The water survey of Canada, edited, doi: August, 2019.
- Kilham, N. E., Roberts, D., & Singer, M. B. (2012). Remote sensing of suspended sediment concentration during turbid flood conditions on the Feather River, California—A modeling approach. *Water Resources Research*, 48, W01521. <https://doi.org/10.1029/2011WR010391>
- Kuhn, C., de Matos Valerio, A., Ward, N., Loken, L., Sawakuchi, H. O., Kampel, M., et al. (2019). Performance of Landsat-8 and Sentinel-2 surface reflectance products for river remote sensing retrievals of chlorophyll-a and turbidity. *Remote Sensing of Environment*, 224, 104–118. <https://doi.org/10.1016/j.rse.2019.01.023>
- Latrubesse, E. M., Arima, E. Y., Dunne, T., Park, E., Baker, V. R., d'Horta, F. M., et al. (2017). Damming the rivers of the Amazon basin. *Nature*, 546(7658), 363–369. <https://doi.org/10.1038/nature22333>, <http://www.nature.com/nature/journal/v546/n7658/abs/nature22333.html#supplementary-information>
- Lee, C. J., & Glysson, G. D. (2013). Compilation, quality control, analysis, and summary of discrete suspended-sediment and ancillary data in the United States, 1901–2010, 776, 35p.
- Li, J. L., & Sheng, Y. W. (2012). An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: A case study in the Himalayas. *International Journal of Remote Sensing*, 33(16), 5194–5213. <https://doi.org/10.1080/01431161.2012.657370>
- Long, C. M., & Pavelsky, T. M. (2013). Remote sensing of suspended sediment concentration and hydrologic connectivity in a complex wetland environment. *Remote Sensing of Environment*, 129, 197–209. <https://doi.org/10.1016/j.rse.2012.10.019>
- Lurry, D., & Kolbe, C. (2000). Interagency field manual for the collection of water-quality data (no. OFR 00-213), edited, USGS.
- Martinez, J. M., Espinoza-Villar, R., Armijos, E., & Moreira, L. S. (2015). The optical properties of river and floodplain waters in the Amazon River Basin: Implications for satellite-based measurements of suspended particulate matter. *Journal of Geophysical Research: Earth Surface*, 120, 1274–1287. <https://doi.org/10.1002/2014JF003404>
- Martinez, J. M., Guyot, J. L., Filizola, N., & Sondag, F. (2009). Increase in suspended sediment discharge of the Amazon River assessed by monitoring network and satellite data. *Catena*, 79(3), 257–264. <https://doi.org/10.1016/j.catena.2009.05.011>
- Mertes, L. A. K., Smith, M. O., & Adams, J. B. (1993). Estimating suspended sediment concentrations in surface waters of the Amazon River wetlands from Landsat images. *Remote Sensing of Environment*, 43(3), 281–301. [https://doi.org/10.1016/0034-4257\(93\)90071-5](https://doi.org/10.1016/0034-4257(93)90071-5)
- Miller, R. L., & McKee, B. A. (2004). Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters. *Remote Sensing of Environment*, 93(1–2), 259–266. <https://doi.org/10.1016/j.rse.2004.07.012>

- Montanher, O. C., Novo, E. M. L. M., Barbosa, C. C. F., Renno, C. D., & Silva, T. S. F. (2014). Empirical models for estimating the suspended sediment concentration in Amazonian white water rivers using Landsat 5/TM. *International Journal of Applied Earth Observation*, 29, 67–77. <https://doi.org/10.1016/j.jag.2014.01.001>
- Morley, S. K., Brito, T. V., & Welling, D. T. (2018). Measures of model performance based on the log accuracy ratio. *Space Weather*, 16, 69–88. <https://doi.org/10.1002/2017SW001669>
- Novo, E. M. M., Hansom, J. D., & Curran, P. J. (1989). The effect of sediment type on the relationship between reflectance and suspended sediment concentration. *International Journal of Remote Sensing*, 10(7), 1283–1289. <https://doi.org/10.1080/01431168908903967>
- Overeem, I., Hudson, B. D., Syvitski, J. P. M., Mikkelsen, A. B., Hasholt, B., van den Broeke, M. R., et al. (2017). Substantial export of suspended sediment to the global oceans from glacial erosion in Greenland. *Nature Geoscience*, 10(11), 859. <https://doi.org/10.1038/Ngeo3046>
- Park, E., & Latrubesse, E. M. (2014). Modeling suspended sediment distribution patterns of the Amazon River using MODIS data. *Remote Sensing of Environment*, 147, 232–242. <https://doi.org/10.1016/j.rse.2014.03.013>
- 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>
- Ritchie, J. C., & Cooper, C. M. (1988). Comparison of measured suspended sediment concentrations with suspended sediment concentrations estimated from Landsat MSS data. *International Journal of Remote Sensing*, 9(3), 379–387. <https://doi.org/10.1080/01431168808954861>
- Ritchie, J. C., Zimba, P. V., & Everitt, J. H. (2003). Remote sensing techniques to assess water quality. *Photogrammetric Engineering & Remote Sensing*, 69(6), 695–704. <https://doi.org/10.14358/Pers.69.6.695>
- 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>
- Spyrakos, E., O'Donnell, R., Hunter, P. D., Miller, C., Scott, M., Simis, S. G. H., et al. (2018). Optical types of inland and coastal waters. *Limnology and Oceanography*, 63(2), 846–870. <https://doi.org/10.1002/lno.10674>
- Stefan, H. G., Cardoni, J. J., Schieber, F. R., & Cooper, C. M. (1983). Model of light penetration in a turbid lake. *Water Resources Research*, 19(1), 109–120. <https://doi.org/10.1029/WR019i001p00109>
- Stumpf, R. P., & Pennock, J. R. (1989). Calibration of a general optical equation for remote sensing of suspended sediments in a moderately turbid estuary. *Journal of Geophysical Research*, 94(C10), 14,363–14,371. <https://doi.org/10.1029/JC094iC10p14363>
- Survey, U. G. (2019). Landsat 4–7 surface reflectance LEDAPS product, edited by D. o. t. interior.
- Syvitski, J. P. M., Peckham, S. D., Hilberman, R., & Mulder, T. (2003). Predicting the terrestrial flux of sediment to the global ocean: A planetary perspective. *Sedimentary Geology*, 162(1–2), 5–24. [https://doi.org/10.1016/S0037-0738\(03\)00232-X](https://doi.org/10.1016/S0037-0738(03)00232-X)
- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B*, 58(1), 267–288.
- Topping, D. J., Rubin, D. M., & Vierra, L. E. (2000). Colorado River sediment transport—1. Natural sediment supply limitation and the influence of Glen Canyon Dam. *Water Resources Research*, 36(2), 515–542. <https://doi.org/10.1029/1999WR900285>
- Umar, M., Rhoads, B. L., & Greenberg, J. A. (2018). Use of multispectral satellite remote sensing to assess mixing of suspended sediment downstream of large river confluences. *Journal of Hydrology*, 556, 325–338. <https://doi.org/10.1016/j.jhydrol.2017.11.026>
- USGS (2018). U.S. Geological Survey, edited, doi: August, 2018.
- USGS (2019). Grand canyon monitoring and research center, edited by U. S. G. Survey, March, 2019.
- Villar, R. E., Martinez, J. M., Guyot, J. L., Fraizy, P., Armijos, E., Crave, A., et al. (2012). The integration of field measurements and satellite observations to determine river solid loads in poorly monitored basins. *Journal of Hydrology*, 444–445, 221–228. <https://doi.org/10.1016/j.jhydrol.2012.04.024>
- Wang, J. J., & Lu, X. X. (2010). Estimation of suspended sediment concentrations using Terra MODIS: An example from the lower Yangtze River, China. *Science of the Total Environment*, 408(5), 1131–1138. <https://doi.org/10.1016/j.scitotenv.2009.11.057>
- Warren, S. (1983). Cubic clustering criterion, SAS Institute Incorporated.
- Warrick, J. A., & Milliman, J. D. (2018). Do we know how much fluvial sediment reaches the sea? Decreased river monitoring of US coastal rivers. *Hydrological Processes*, 32(23), 3561–3567. <https://doi.org/10.1002/hyp.13276>
- Wohl, E., Bledsoe, B. P., Jacobson, R. B., Poff, N. L., Rathburn, S. L., Walters, D. M., & Wilcox, A. C. (2015). The natural sediment regime in rivers: Broadening the foundation for ecosystem management. *Bioscience*, 65(4), 358–371. <https://doi.org/10.1093/biosci/biv002>
- WRA (2018). Taiwan water resource agency, edited, doi: August, 2018.
- Yepez, S., Laraque, A., Martinez, J.-M., De Sa, J., Carrera, J. M., Castellanos, B., et al. (2017). Retrieval of suspended sediment concentrations using Landsat-8 OLI satellite images in the Orinoco River (Venezuela). *Comptes Rendus Geoscience*, 350(1–2), 20–30. <https://doi.org/10.1016/j.crte.2017.08.004>