GLOBAL ABUNDANCE AND MORPHOLOGY OF RIVERS AND STREAMS

George Henry Allen

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctorate of Philosophy in the Department of Geological Sciences.

Chapel Hill
2017

Approved by:

Tamlin M. Pavelsky

Lawrence E. Band

Jonathan M. Lees

Laura J. Moore

Karl W. Wegmann

ABSTRACT

George Henry Allen: The global abundance and morphology of rivers and streams
(Under the direction of Tamlin Muir Pavelsky)

The abundance and morphology of rivers control the rates of hydraulic and biogeochemical exchange between rivers, groundwater and the atmosphere. However, current knowledge of the abundance and morphology of Earth's rivers and streams is based on a series of highly unconstrained hydrologic, geomorphic, climatic, and fractal river network scaling extrapolations. These extrapolations are the source of significant uncertainty in many large-scale hydrologic, geomorphic, and biogeochemical applications. In this dissertation, I characterize the global abundance and morphology of rivers and streams using fieldwork and global-scale satellite remote sensing observations. In Chapter 1, I use field surveys to characterize the distribution of stream widths in thirteen small headwater stream networks across North American and New Zealand. I show a strikingly consistent lognormal statistical distribution of stream width in all surveys, including a characteristic most abundant stream width of 32±7 cm independent of physiographic or hydrologic conditions. I propose a framework showing that, as stream networks expand and contract within the geomorphic channel network in response to changes in streamflow, the most abundant stream width remains approximately static. In Chapter 2, I present the Landsat-derived North American River Width (NARWidth) dataset, the first fine-resolution, continental-scale river centerline and width database. NARWidth contains measurements of >240,000 km of rivers wider than 30 m at mean annual discharge. I find that conventional digital elevation model-derived river width datasets underestimate the abundance of

wide rivers. In Chapter 3, I present the Global River Widths from Landsat (GRWL) Database, the first global survey of river planform geometry at mean discharge. GRWL contains measurements of river geometry of >2.1 x 10^6 km of rivers. GRWL is being used by other researchers to improve the representation of river water resources, hazards, and hydrological processes at large scale. I use GRWL, and the results presented in Chapter 1, to estimate the global distribution of rivers and streams. Using geographic information science and a novel statistical method, I constrain the total surface area of all rivers and streams to 745,000 km^2, or 0.55% of Earth's unglaciated land surface.

to Fiona
who keeps me strong

and to Caroline Rose
who brings me joy for the future

ACKNOWLEDGEMENTS

I thank my advisor, Dr. Tamlin M. Pavelsky for his stellar guidance and inspirational enthusiasm for science and science communication. Tamlin's dedicated support and mentorship have made my graduate school experience tremendously rewarding and enjoyable. I thank the members of my dissertation committee: Dr. Jonathan Lees, Dr. Lawrence Band, Dr. Karl Wegmann, Dr. Laura Moore, and Dr. Jason Barnes, who provided useful advice and encouragement. I also thank my coauthors Elizabeth Altenau, Dr. Xiao Yang, Dr. Colin Gleason, Dr. David Butman, Dr. Mike Lamb, Dr. Mike Durand, Dr. Christine Lion, Zach Miller, Arik Tashie, and Eric Barefoot for fruitful scientific collaborations. I could not have completed my dissertation without the monumental effort made by the talented UNC undergraduate research assistants who contributed to the Global River Widths from Landsat Database: Kyle Hinson, Christine Destefano, Sam Dawson, Claire Nelson, Melony Eimer, Will Rudisill, Shannon Steele, Jonathan Witten, Natalie Burk, Colin Peterson, Mai Dvorak and Wood Robinson. I thank Arik Tashie, Eric Barefoot, Yike Xu, Margaret Zimmer, Dr. Brian McGlynn, Emily Beckham, and Elizabeth Henry for their assistance in the field. I appreciate the statistical assistance I received from Dr. Chris Wiesen at the Odum Institute. This work was funded by NASA NIP grant #NNX12AQ77G, NASA THP grant #NNX14AD82G, the UNC Geological Sciences Martin Research Endowment, the Paul C. Hardin Dissertation Year Fellowship, the NC Space Grant Graduate Research Fellowship, a GSA Graduate Student Research Grant, and the UNC Future Faculty Fellowship. Last but not least I thank my family for their support. I thank my parents, Robert and Charla, for encouraging me to aspire to the highest level of education, and I thank my

other parents, Jef and Katie, for keeping my eyes on the prize: graduating with a job. I thank my sister, Sarah, for pioneering the way through graduate school and I thank Brother Tom for being a friend throughout. Finally, I thank my wife Fiona for her unyielding love and support through graduate school and my new daughter, Caroline Rose, who seems to bring me good luck with every passing day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

$\%SA$	River area to basin area (%)
a	Stream width-discharge coefficient
A	Upstream drainage area (L^2)
$\bar{A}_1$	Mean area contributing to first-order stream segment (L^2)
ADN	Active Drainage Network
A_g	Drainage area at gauge (L^2)
AI	Aridity Index
A_s	Cross sectional area of the stream (L^2)
$\bar{A}_\omega$	Mean drainage area of stream segment of order ω (L^2)
α	Bankfull width to drainage area coefficient
b	Stream width-discharge exponent
BA	Basin Area (L^2)
β	Bankfull width to drainage area exponent
c	Basin-averaged runoff exponent
C	Constant
X	Standard normal variable
DEM	Digital Elevation Model
DHG	Downstream Hydraulic Geometry
g	Gravitational acceleration (L/T^2)
GIS	Geographic Information Systems

GPS	Global Positioning System
GRDC	Global Runoff Data Centre
GRWL	Global River Widths from Landsat Database
h	Wetted stream depth (L)
h_{bf}	*B*ankfull depth (L)
i	Observation location
k	Parker roughness length scale (L)
K_L	Stream length constant (L)
K_r	Basin-averaged runoff coefficient
K_w	Stream width constant
$\bar{L}_1$	Mean length first-order stream segment (L)
L_ω	Total length of all streams of order ω (L)
$\bar{L}_\omega$	Mean length of a stream segment of order ω (L)
MNDWI	Modified Normalized Difference Water Index
n	Gauckler–Manning friction coefficient
N	Number of observations
NARWidth	North American River Width Dataset
N_ω	Stream segment number of order ω
ω	Strahler stream order
Ω	Highest order of stream network
Q	Stream discharge (L^3/T)

Q_g	Stream discharge at gauge (L^3/T)
Q_m	Mean monthly discharge (L^3/T)
Q_μ	Mean annual discharge (L^3/T)
$\bar{Q}_\omega$	Mean stream discharge of segment of order ω (L^3/T)
R_A	Stream area ratio
R_B	Stream bifurcation ratio
R_L	Stream length ratio
s	Channel shape parameter
S	Channel bed slope
SA	Stream surface area (L^2)
SWOT	Surface Water and Ocean Topography Mission
t	Time step
u	Streamflow velocity (L/T)
USGS	United States Geological Survey
w	Wetted stream width (L)
$\bar{w}_\omega$	Mean stream width of segment of order ω (L)
w_{bf}	Bankfull width (L)
WSC	Water Survey of Canada

CHAPTER 1: SIMILARITY OF STREAM WIDTH DISTRIBUTIONS ACROSS HEADWATER SYSTEMS[1]

1.1 INTRODUCTION

Headwater streams (stream order 1-3) (Strahler, 1957; Vannote et al., 1980) comprise an estimated 89% of the global fluvial network length (Downing et al., 2012; Raymond et al., 2013) and are the source of water, sediment, nutrients and organic matter for downstream systems (Gomi et al., 2002). They exhibit highly variable physical, chemical and biotic attributes; as a result, they contribute to significant biodiversity within watersheds (Meyer et al., 2007). They are also more hydraulically coupled to hillslope and groundwater processes compared to larger streams and thus are hotspots for biogeochemical activity (Bencala et al., 2011; Butman et al., 2016; Gomi et al., 2002; Hotchkiss et al., 2015; Raymond et al., 2013). High rates of hyporheic exchange expose transported solutes to unique biogeochemical environments, with subsequent impacts on whole stream metabolism (Findlay, 1995), nutrient cycling (Triska et al., 1989) and contaminant uptake and export (Fuller and Harvey, 2000). Small streams are also a significant source of greenhouse gas to the atmosphere. In fact, over half of the greenhouse gas emitted from the fluvial network originates from small headwater streams (Butman et al., 2016; Raymond et al., 2013). This biogeochemical activity is, in part, a function of stream surface water geometry.

Stream width, defined as the wetted width of flowing water within a channel, reflects natural heterogeneities along a stream such as channel margins, eddies behind large woody

[1] This chapter is in review as a research letter in the journal *Nature*. Some of the supplementary figures and text have been integrated into the main text and the formatting has been changed. The publication citation is as follows: Allen, G.H., Pavelsky, T.M., Barefoot, E.A., Lamb, M.P., Butman, D., Tashie, A., Gleason, C.J., (in review), Similarity of stream width distributions across headwater systems. *Nature*.

debris and hyporheic exchange flow paths (Bencala et al., 2011; Harvey and Bencala, 1993; Kasahara and Wondzell, 2003). These heterogeneities are important because they serve as micro-environmental patches that impact temporary solute storage, material erosion and deposition, biological and ecosystem processes and ultimately large-scale biodiversity (Allen et al., 2013; Vannote et al., 1980; Wondzell, 2011). Where it has been studied, planform stream hydromorphology is often quantified by measuring width at uniformly-spaced intervals along stream centerlines (Allen et al., 2013; Allen and Pavelsky, 2015; Gleason and Smith, 2014; Hankin and Reeves, 1988; Kiel and Cardenas, 2014). Stream width data are used in a broad array of applications including studies of hyporheic flow (Kasahara and Wondzell, 2003; Kiel and Cardenas, 2014), open-channel hydraulics (Gleason and Smith, 2014), material transport and erosion (Allen et al., 2013), lotic habitat (Hankin and Reeves, 1988) and stream-atmosphere gas exchange rates (Butman et al., 2016; Hotchkiss et al., 2015; Raymond et al., 2013). Stream width is also a core variable in the River Continuum Concept, an important conceptual framework that relates lotic ecosystems to stream size (Vannote et al., 1980).

Despite their wide-ranging importance, there has been no published characterization of the distribution of stream widths within an entire headwater catchment. Instead, static, topographically-derived flowlines are typically used to represent stream networks and infer their geometry (Butman et al., 2016; Kiel and Cardenas, 2014; Raymond et al., 2013). However, active drainage networks (ADNs) typically expand and contract with changing streamflow conditions, causing temporal fluctuations in catchment drainage density and stream surface area (Godsey and Kirchner, 2014). Temporal change in drainage density has been studied (Godsey and Kirchner, 2014; Gregory and Walling, 1968), but the simultaneous spatial dynamics of headwater stream widths remains undocumented. Instead, studies requiring stream geometry in

headwater catchments usually estimate stream width distributions using hydraulic scaling principles developed for larger river systems (Allen and Pavelsky, 2015; Butman et al., 2016; Raymond et al., 2013). These scaling principles produce a Pareto (or power-law) distribution of stream width that may be inappropriate for smaller stream networks (Anderson et al., 2004; Gomi et al., 2002).

1.2 METHODS

1.2.1 Field Methods

To characterize stream width distributions in a range of headwater systems, we conducted the most comprehensive field survey to date of stream hydromorphology (wetted width and length) within seven headwater catchments (Figure 1.1). The study catchments spanned a wide range of sizes, environments, geomorphologies and streamflow conditions (see Figure 1.2 for field photos and Appendix Table A1.1 for site-specific attributes). In each of the seven physiographically contrasting study catchments, we paced along all streams within the stream network and measured wetted stream width every 5 m (after Schiller et al., 2011). Additionally, in a 48 hectare subcatchment of Stony Creek Research Watershed in North Carolina, we repeatedly surveyed stream width over a range of hydrologic conditions (Appendix Table A1.2). For the repeat surveys in the Stony Creek subcatchment, we flagged streams every 5 m and surveyed wetted stream widths at each flag over a range of hydrologic conditions. For the repeat surveys in Stony Creek subcatchment, we only analyzed surveys that were collected while streamflow was below the 90th percentile of the streamflow record in order to remove the potential influence of overbank flooding.

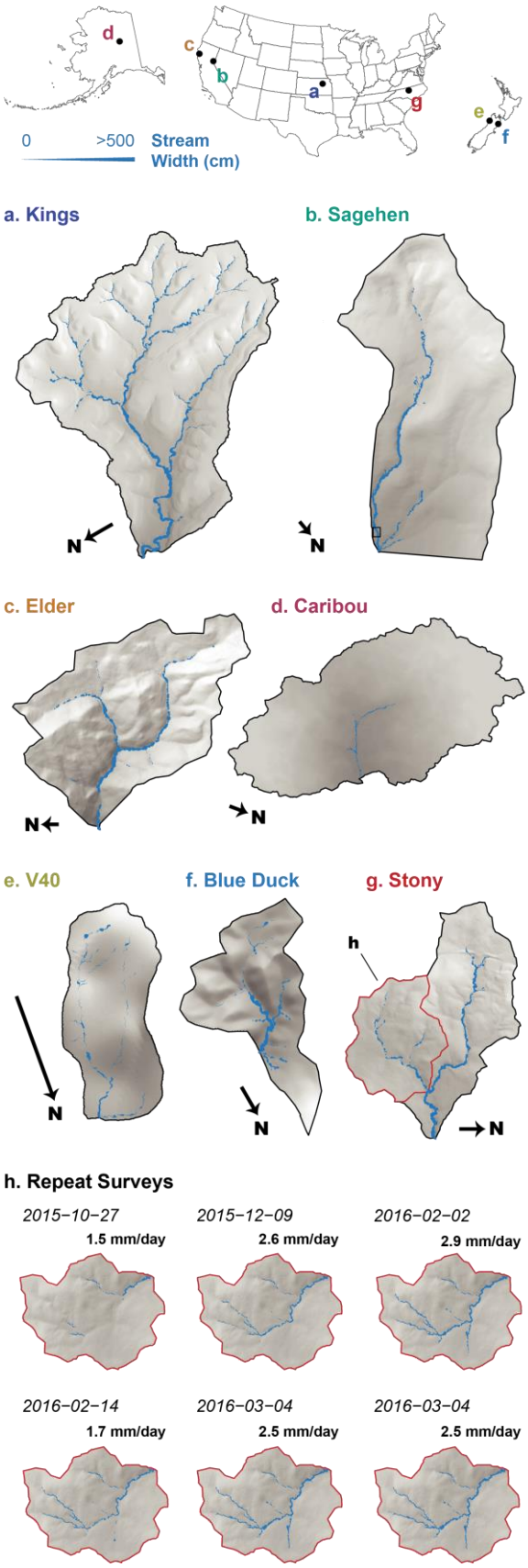

Figure 1.1 - Stream width maps in study catchments

a, K1B tributary, North Branch of Kings Creek, KS; **b,** Sagehen Creek subcatchment, CA; **c,** Upper Elder Creek, CA; **d,** C1 tributary of Caribou Creek, AK; **e,** V40 Stream subcatchment, NZ; **f,** Blue Duck Creek subcatchment, NZ; **g,** Stony Creek Research Watershed, NC. **h,** Stony Creek subcatchment repeat surveys with the basin-averaged runoff values in bold. Lengths of north arrows represent 200 m.

We defined a stream as flowing water within a channel (after Shaw, 2016), including transient channels formed in leaf litter and other debris. We measured wetted stream width with a standard tape measure or, where a tape measure was not practical, with a laser range finder. In multichannel streams, we added the stream widths from all channels together or we visually approximated the percentage of the total width that was dry. We quantified measurement error by repeatedly surveying stream width along a 175 m long stream segment located at the lowermost segment of the Stony Creek Subcatchment. We surveyed the segment five times within 1.5 hours and then compared the width measurements to calculate standard error of 3 cm.

In each catchment, we collected between 160 and 1,797 (mean N=672) stream width measurements. We mapped ADNs with a continuous tracking GPS device, or where necessary, by hand on a topographic map or on optical remotely sensed imagery. We removed 36 survey points (2.5%) from our analysis of Sagehen Creek where snow completely or partially obscured the stream surface We approximated relative hydrological conditions in each physiographically contrasting study catchment by calculating the streamflow percentile and catchment-averaged runoff during the day(s) we surveyed streams relative to the entire gage record (Appendix Table A1.1). Stream gages were often located nearby or downstream from the study catchments and thus the runoff and discharge percentiles presented in Appendix Table A1.1 indicate a general characterization of catchment wetness.

Figure 1.2 - Photographs of study catchments and streams

a-h We surveyed supply- and transport-limited bedrock streams, streams with alluvial substrates (clay, sand, glacial till), and streams impacted by vegetation and woody debris. Study sites also included perennial spring-fed streams, intermittent or ephemeral streams, and streams flowing over permafrost.

1.2.2 Statistical Methods

We described the statistical distribution of stream widths within each study basin by fitting lognormal, gamma, Weibull, and Pareto distributions to the stream width data using maximum likelihood estimation. We quantified model goodness of fit using the Pearson's chi-square goodness of fit test (Venables and Ripley, 2002) and the nonparametric two sided one sample Kolmogorov–Smirnov goodness of fit test (Venables and Ripley, 2002) (Appendix Figure A1.1, Appendix Table A1.3). Using a Gaussian density kernel with a bandwidth of 10 cm, we calculated the mode stream width in each basin. Kernel bandwidth was determined using the Normal Reference Distribution optimal bandwidth selection technique (Venables and Ripley, 2002). We correlated the mode width to physical conditions between physiographically contrasting catchments using the coefficient of determination (R^2) and p-value with significance level, α=0.05. We calculated total stream surface area by summing the product of each stream width and length measurement within a catchment.

1.2.3 Stream Width Model

To understand the origin of the observed distributions of stream widths, we model stream width in each catchment using relationships between stream channel shape, hydraulic resistance, drainage area and discharge. While this model shares some conceptual similarities to preceding studies (Dingman, 2007; Ferguson, 1986; Neal et al., 2015; Parker et al., 2007), it is, to our knowledge, a novel synthesis of downstream hydraulic geometry, at-a-station hydraulic geometry, and natural variability in stream channel cross sectional geometry. Our model begins with the analytical relationship for at-a-station hydraulic geometry presented in by Dingman (2007). The cross sectional shape of stream channels has been modeled as a variety of geometries including triangular, parabolic, trapezoidal, and rectangular (Chow, 1959). Here we

simulate these channel shapes by varying a single shape parameter, s, such that for any wetted depth less than bankfull depth,

$$h = h_{bf}\left(\frac{w}{w_{bf}}\right)^{s}, \tag{1.1}$$

where h_{bf} is bankfull depth, w is wetted stream width, and w_{bf} is bankfull channel width (Figure 1.3a; Dingman, 2007; Neal et al., 2015). Setting s = 1 yields a triangular cross section, and increasing its value beyond 1 yields an increasingly concave (or flat-bottomed) parabolic channel shape (Figure 1.3b).

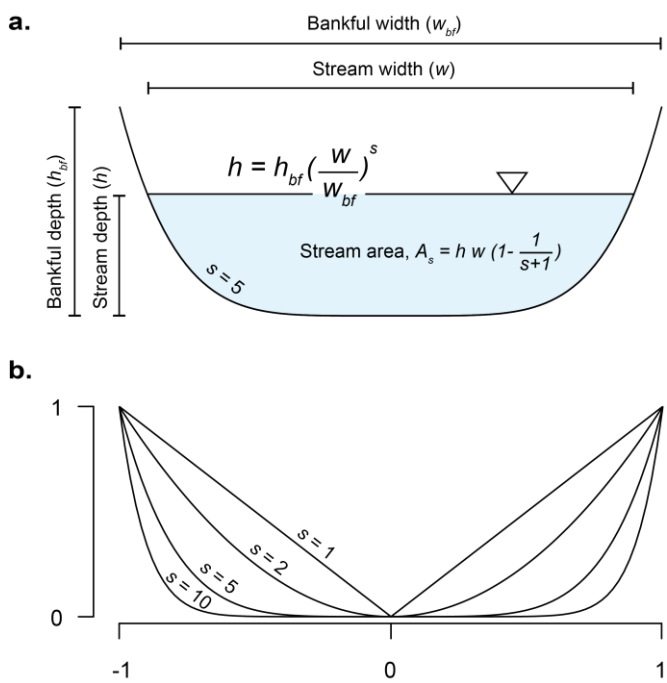

Figure 1.3 - Stream width model schematic
a, Schematic channel cross section; **b**, schematic relationship between the shape parameter, s, and channel cross sectional shape (modified from Neal et al., 2015).

Within a channel, the law of conservation of mass relates stream discharge to channel shape,

$$Q = uA_s, \tag{1.2}$$

where u is mean streamflow velocity and A_s is the cross sectional area of the stream and is calculated as,

$$A_s = hw(1 - \frac{1}{s+1}). \tag{1.3}$$

Flow velocity is modeled using a form of the Manning-Strickler relation for flow resistance (Parker, 1991),

$$u = 8.1(ghS)^{\frac{1}{2}}\left(\frac{h}{k}\right)^{\frac{1}{6}}, \tag{1.4}$$

where g is gravitational acceleration, S is channel bed slope, and k is a roughness length scale equivalent to,

$$k = \left(8.1g^{\frac{1}{2}}n\right)^{6}, \tag{1.5}$$

where n is the Gauckler–Manning friction coefficient (Manning, 1891). Combining Equations (1.1-1.4) yields a relationship, similar to a generalized relationship presented in Dingman (2007), between stream width and other hydraulic and geomorphic parameters:

$$w = Q^{\frac{3}{5s+3}}\, w_{bf}{}^{\frac{s-1}{s+3/5}}\left(8.1(gS)^{\frac{1}{2}}k^{-\frac{1}{6}}\left(\frac{w_{bf}}{h_{bf}}\right)^{-\frac{3}{5}}(1 - \frac{1}{s+1})\right)^{-\frac{3}{5s+3}}. \tag{1.6}$$

Bankfull stream width scales with upstream drainage area (A) as a power law,

$$w_{bf} = \alpha A^{\beta}, \tag{1.7}$$

where α and β are empirical constants (Golden and Springer, 2006; Trampush et al., 2014; Wilkerson et al., 2014; Williams, 1978). We calculate the values of α and β to be 0.008 and 0.42, respectively, using least squares regression on a global compilation of 307 paired log-

transformed w_{bf} and A measurements (R^2 = 0.68, p < 0.001; Trampush et al., 2014), values similar to previous work (Golden and Springer, 2006; Wilkerson et al., 2014; Williams, 1978). Note that using the empirical β value of 0.42, rather than the dimensionally consistent β value of 0.5, propagates a minor imbalance in the dimensions of Equation (1.6). We use the same global database from Trampush et al. (2014) to find that the median bankfull width to depth ratio h_{bf}/w_{bf} is 14 in Equation (1.5), for streams with upstream drainage areas within the range we surveyed in this study. Previous studies have reported h_{bf}/w_{bf} ratios varying from 1.5 to 35 (Petersen, 1992; Wohl and Merritt, 2008; Wohl and Wilcox, 2005; Zimmerman et al., 1967), and the value we use falls within this range.

At each stream width observation site, i, we calculate stream width by combining Equations (1.6) and (1.7),

$$w_i = Q_i^{\frac{3}{5s_i+3}} \left(\alpha A_i^{\beta}\right)^{\frac{s_i-1}{s_i+3/5}} \left(8.1(gS_i)^{\frac{1}{2}} D^{-\frac{1}{6}} \left(\frac{w_{bf}}{h_{bf}}\right)^{-\frac{5}{3}} (1-\frac{1}{s_i+1})\right)^{-\frac{3}{5s_i+3}}. \tag{1.8}$$

We compute A_i and S_i from DEMs listed in Appendix Table A1.1 (O'Callaghan and Mark, 1984). Slope is calculated over the length scale of the resolution of the DEM used. To estimate Q_i at each survey location, we use conservation of mass within a drainage basin,

$$Q_i = Q_g \frac{A_i^c}{A_g^c}, \tag{1.9}$$

where Q_g and A_g are the discharge and drainage area at the stream gage (Appendix Table A1.1) and set c = 1 (Hack, 1957). We set the Gauckler–Manning friction coefficient, n = 0.04, as commonly assumed for mountain streams with gravel/cobble beds (Allen et al., 2013; Chow, 1959). To represent natural diversity in stream channel shape, we randomly vary the value of s_i in Equation (1.8) between 1 and 10, and thus capture channel geometries ranging from a triangle to a function approaching a rectangle (Figure 1.3b). In a separate analysis, we drew values of s_i

from a normal distribution (μ = 5, σ = 2), rather than a uniform distribution, which yielded similar results.

Classic hydraulic geometry relationships are reflected in the exponents of Equation (1.8). The first factor in Equation (1.8), Q_i, represents the at-a-station hydraulic geometry (Leopold and Maddock, 1953) component of the model, the second, aA_i^{β}, represents the downstream changes in channel width, rather than wetted stream width, and the remaining factors encapsulate the influences of flow resistance and channel geometry. If s = 2, then $w_i \propto Q_i^{\frac{3}{13}} = Q_i^{0.23}$, which is similar to typical power-law relations for at-a-station hydraulic geometry (Leopold and Maddock, 1953). Similarly, for downstream hydraulic geometry, if s = 2, then $w \propto Q_i^{0.23} A_i^{\beta 5/13}$, and if $A_i \propto Q_i$ and β = 0.42, then $w_i \propto Q_i^{0.39}$.

1.3 RESULTS

1.3.1 Field Results

The stream widths of all surveys are well characterized by lognormal distributions and exhibit a mode of 32±7 cm (all confidence intervals 1σ, Figure 1.4). The mode width, determined using a Gaussian density kernel, is strikingly similar across all basins and does not significantly correlate with hydrologic conditions (R^2=0.15, p=0.39), basin relief (R^2=0.23, p=0.28), catchment area (R^2=0.04, p=0.69), or drainage density (R^2=0.19, p=0.33). Gamma and Weibull distributions also effectively describe the spread of stream width data (Appendix Figure A1.1, Appendix Table A1.3). The median first-order stream width is 32±8 cm. No instances of overbank flooding were observed during the surveys but disconnections in the ADN were common, particularly in first-order streams.

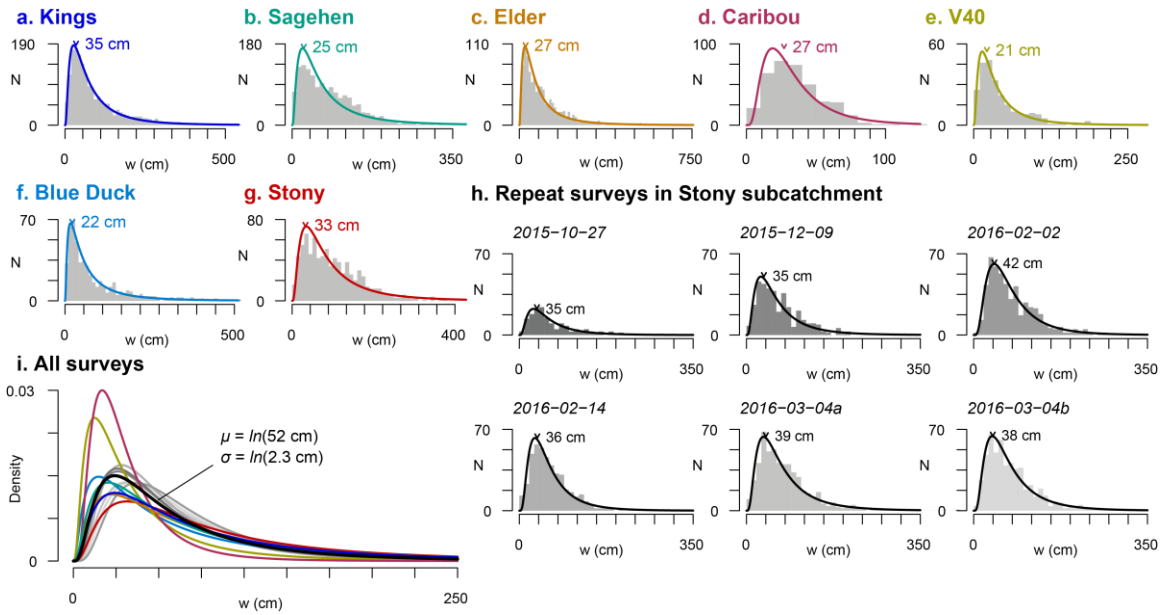

Figure 1.4 - Headwater stream width distributions.
a-h, Stream width histograms with lognormal fits for each survey. Length values indicate the mode stream width in each catchment. **i**, Composite probability density functions of lognormal fits with mean fit in black. μ and σ are the lognormal location and scale parameters in Equation (1.10). N: measurement frequency; w: stream width.

1.3.2 Stream Width Model Results

Caribou and V40 catchments were excluded from the analysis because their available digital elevation models (DEMs) were of insufficient quality to produce accurate stream networks at the scales observed. The model produces stream widths that are spatially realistic (Figure 1.5a) and are distributed similarly to the observations (Figure 1.5c-f). The model, which closely matches observations in four of the five catchments examined, indicates that stream widths are primarily set by discharge and random variability in the channel geometry, from V- to U-shaped, represented by s in Figure 1.3b. The differences between the model outputs and the observations likely stem from simplifying assumptions regarding runoff yield, bankfull channel widths, hyporheic zone transmissivity (Godsey and Kirchner, 2014) and hydraulic resistance (Parker, 1991), none of which were measured in the field.

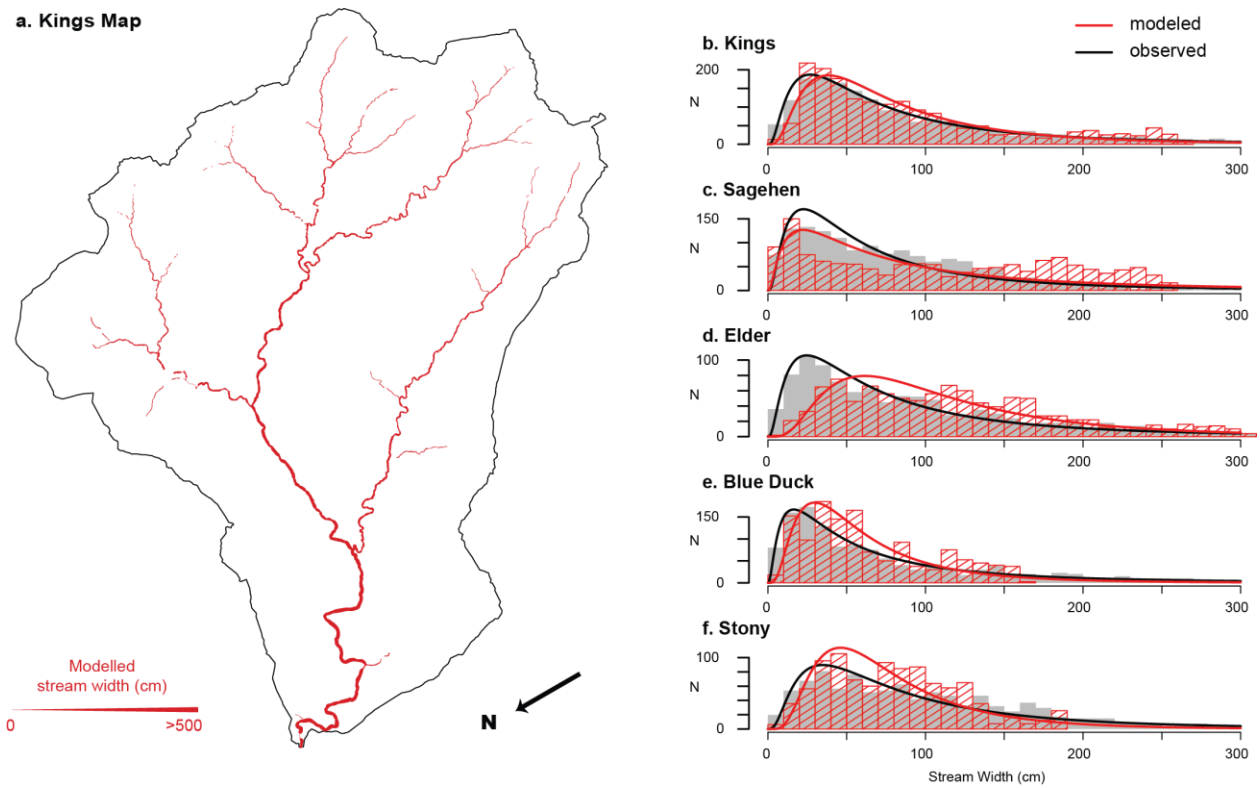

Figure 1.5 - Origin of the lognormal stream width distribution
a, map of modeled stream widths in example study catchment. Length of north arrow is 200 m. **b-f**, the distribution of modeled stream widths compared to field observations. N: measurement frequency.

1.4 DISCUSSION

The insensitivity of stream width distributions to changes in catchment runoff stems from the counteracting processes of stream widening and ADN expansion (see Section 1.4.1). With increasing discharge, streams will simultaneously widen in place and extend upstream such that individual stream segments will increase in stream order as tributary channels are reactivated (Godsey and Kirchner, 2014). As a result, the proportional abundance of narrow streams remains roughly constant (Figure 1.6). Thus, if the cumulative length (L) of an ADN is known, the total stream surface area (SA) may be approximated using the mean lognormal fit in Figure 1.4i,

$$SA = \sum \frac{L}{N} e^{\mu+\sigma X}, \tag{1.10}$$

where N is the number of observations, X is a standard normal variable of length N, μ = ln(32 cm) and σ = ln(2.3 cm). We anticipate that this model may break down with the initiation of overbank flooding or when the channel network is completely occupied by the ADN.

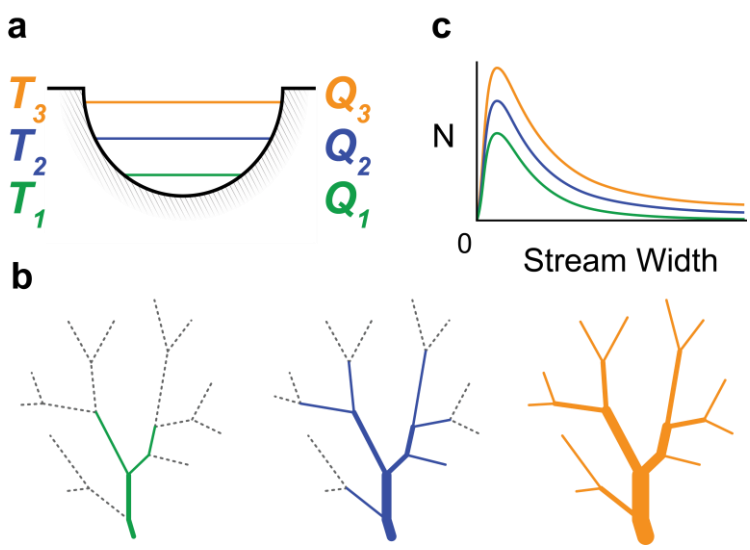

Figure 1.6 - Conceptual model of relationship between changing streamflow conditions (Q1<Q2<Q3) and the stream width distributions
As discharge increases: **a**, wetted width increases at a representative channel cross section; **b**, the ADN extends into narrower lower-order channels, increasing the total length of narrow streams (dotted lines represent channels without flow); and **c**, the phenomena described by panels a and b manifest in the similar width distributions independent of streamflow conditions.

1.4.1 Theoretical Derivation of Observed Stream Width Similarities

The Horton-Strahler ordering system (Strahler, 1957) is commonly used to organize stream network structure. In context of the ADN, stream segments that originate at flowheads are assigned a stream order, ω = 1 and where two segments of the same order join, the resulting segment is of order $\omega = \omega + 1$. If two stream segments of a different order join, the higher order segment continues through the junction and the lower-order segment terminates at the junction.

Horton's law of stream numbers states that with increasing stream order, the number of stream segments of that order, N_ω, decreases geometrically,

$$\frac{N_\omega}{N_{\omega+1}} = R_B \text{ or } N_\omega = {R_B}^{\Omega-\omega}, \tag{1.11}$$

where Ω is the highest stream order of the stream network and R_B is the bifurcation ratio, typically equal to 4, but can vary between 3 and 5 (Horton, 1945; Kirchner, 1993; Strahler, 1957). Assuming that streams occupy the entire length of their channel, Horton's law of stream lengths states that the mean length of a stream segment varies by order such that,

$$\frac{\bar{L}_\omega}{\bar{L}_{\omega-1}} = R_L \text{ or } \bar{L}_\omega = \bar{L}_1 {R_L}^{\omega-1}, \tag{1.12}$$

where $\bar{L}_1$ is the mean length of a first-order stream segment and R_L is the length ratio that can range from 1.5 to 3, with a typical value of 2 (Horton, 1945; Kirchner, 1993; Strahler, 1957). The product of Equation (1.11) and (1.12) shows that the total length of all streams of order ω geometrically decreases with stream order,

$$L_\omega = N_\omega \bar{L}_\omega = \bar{L}_1 {R_B}^{\Omega-\omega} {R_L}^{\omega-1}. \tag{1.13}$$

Setting R_B = 4 and R_L = 2, Equation (1.13) becomes,

$$L_\omega = \bar{L}_1 2^{2\Omega-\omega-1}, \tag{1.14}$$

which can be written in terms of base e,

$$L_\omega = \bar{L}_1 (e^{2\Omega-\omega-1})^{\frac{\ln(2)}{\ln(e)}}, \tag{1.15}$$

or

$$L_\omega = K_L e^{-\ln(2)\omega}, \tag{1.16}$$

where

$$K_L = \bar{L}_1 e^{\ln(2)(2\Omega-1)}. \tag{1.17}$$

Equation (1.16) indicates that total stream length of order ω decreases geometrically with stream order at rate of -ln(2), or -0.69 and thus first-order streams occupy the most length within the stream network. However, due in part to stream disconnections, first-order streams typically occupy only a portion of their channel lengths and thus have decreased abundance relative the rest of the stream network length. Regardless, together, first- and second-order streams comprise the greatest length of stream networks.

The law of stream areas states that the mean drainage area of a stream segment of order ω, increases geometrically with stream order,

$$\frac{\bar{A}_\omega}{\bar{A}_{\omega-1}} = R_A \text{ or } \bar{A}_\omega = \bar{A}_1 {R_A}^{\omega-1}, \tag{1.18}$$

where $\bar{A}_1$ is the mean area contributing to a first-order stream segment and R_A is the area ratio, which generally ranges between 3 and 6 (Kirchner, 1993; Schumm, 1956). Stream discharge and drainage area scale according to,

$$Q = K_r A^c, \tag{1.19}$$

where $K_r$is a basin-averaged runoff coefficient and c = 1 (Hack, 1957). Combining Equation (1.18) and (1.19) shows that the mean stream discharge of a segment ($\bar{Q}_\omega$) increases geometrically with stream order,

$$\bar{Q}_\omega = K_r {\bar{A}_\omega}^c. \tag{1.20}$$

Stream width scales spatially with discharge in the framework of downstream hydraulic geometry,

$$w = aQ^b, \tag{1.21}$$

where a and b are empirical constants and b is typically equal to 0.5 (Leopold and Maddock, 1953). Combining Equation (1.18-1.21) yields a relationship between mean stream width within a segment of order ω, ($\bar{w}_\omega$) and stream order,

$$\bar{w}_\omega = a\left(K_r\left(\bar{A}_1 {R_A}^{\omega-1}\right)^c\right)^b. \tag{1.22}$$

Setting R_A = 4, Equation (1.22) becomes,

$$\bar{w}_\omega = a{K_r}^b {\bar{A}_1}^{bc} 2^{2bc(\omega-1)} \tag{1.23}$$

which when written in terms of base e becomes,

$$\bar{w}_\omega = a{K_r}^b {\bar{A}_1}^{bc}\left(e^{2bc(\omega-1)}\right)^{\frac{\ln(2)}{\ln(e)}}, \tag{1.24}$$

or

$$\bar{w}_\omega = K_w\, e^{2\,\mathrm{bc}\ln(2)\omega}, \tag{1.25}$$

where

$$K_w = a{K_r}^b {\bar{A}_1}^{bc} e^{-2\ln(2)bc}. \tag{1.26}$$

Setting b = 0.5 and c = 1, Equation (1.25) indicates that the average stream width of order ω increases geometrically with stream order at rate of ln(2) or 0.69, the opposite rate that stream length scales with stream order.

Solving for ω in Equation (1.16) and (1.25) and setting each equation equal to each other yields the relationship between stream width and length by order,

$$L_\omega = \frac{K_L K_\omega}{\bar{w}_\omega}. \tag{1.27}$$

This equation implies that stream width is inversely proportional to stream length with respect to stream order, and thus the total stream surface area of each stream order is approximately equal. Equation (1.27) also implies that the stream width is Pareto distributed with a shape parameter, α = 1, or in other terms, a power-law relationship with an exponent of -1. This theoretical α-value observed from the 13 surveyed streams is α = 1.17±0.29 (Appendix Table A1.3).

As a relatively dry stream network (t=1) expands into previously dry channels (t=2) due to increasing catchment runoff, first-order streams become second-order, second-order streams

become third-order and the highest stream order of the network, Ω, increases by 1 (Godsey and Kirchner, 2014). As the stream network expands, the total length of first-order streams (L_1) increases,

$$L_{1,t=2} = \bar{L}_{1,t=1}\left(\frac{R_B}{R_L}\right), \tag{1.28}$$

and the mean drainage area of first-order streams decreases,

$$\bar{A}_{1,t=2} = \bar{A}_{1,t=1}/R_A \tag{1.29}$$

(Figure 1.6). Equation (1.29) would suggest that with increasing discharge the mode of stream width should decrease, but we observed that the mode width was roughly static, independent of runoff conditions. Thus, for stream networks to maintain the observed fixed distribution of stream widths during changes in discharge, the decrease in first-order drainage area must be balanced by a corresponding increase in runoff yield,

$$K_{r,t=2} = K_{r,t=1}R_A. \tag{1.30}$$

This theoretical framework suggests that increasing catchment runoff causes the total length of each ADN stream order to geometrically increase while simultaneously, the mean stream width of each order remains constant. Thus, the total length of first- and second-order streams, which are the primary source of stream widths remain proportional to each other and their absolute average widths remain equal. This equilibrium results in the static width distribution observed in Figure 1.4h. We expect the natural variability in values of R_B, R_L, R_A, b, and c can cause deviations from a perfect equilibrium real-world stream systems.

1.4.2 Implications for Carbon Efflux Estimates

These results hold significant implications for understanding hydrological, ecological and biogeochemical processes occurring in headwater streams. For example, previous evaluations (Butman et al., 2016; Raymond et al., 2013) of greenhouse gas emissions from rivers and

streams estimated stream surface area using Pareto scaling laws on static USGS and international DEM-derived flowlines, which significantly underestimate the abundance of headwater streams (Benstead and Leigh, 2012). These studies assume that median first-order stream width ranges from 160±110 cm to 315 cm (Butman et al., 2016; Downing et al., 2012; Raymond et al., 2013), an order of magnitude greater than observed in this study.

To evaluate the impact of these differences, we compare our observations against an existing flowline datasets currently used in biogeochemical studies to calculate surface emissions of CO_2 (Butman et al., 2016; Raymond et al., 2013) (see Carbon efflux estimates Appendix Section A1.1). We find that the dynamic expansion and contraction of ADNs causes significant temporal variability in greenhouse gas emissions in headwater stream networks. In the repeat stream width surveys, estimated CO_2 efflux quadruples in response to a doubling in runoff (Appendix Table A1.4). Among the physiographically contrasting catchments, we find that CO_2 efflux calculations based of the USGS flowline datasets differ from our observation-based estimates by as much as 100% (RMSE=17.7 Mg-C/Yr). Using Equation (1.10) to estimate water surface area yields CO_2 efflux values that are more similar to observation-based estimates than the USGS flowline based estimates (RMSE=4.17 Mg-C/Yr). The differences between CO_2 efflux estimates arise from the dissimilarity of stream network length and width distributions between our observations, Equation (1.10), and the USGS flowline datasets.

1.5 CONCLUSIONS

Our observations suggest that stream widths in headwater networks are lognormally distributed, rather than Pareto distributed and that the most common stream width is substantially narrower than previously assumed. This lognormal distribution can be used to more accurately estimate stream surface area in small headwater catchments if the total length of the stream

network is estimated, with implications for stream-atmosphere gas exchange estimates. We find that the dynamic nature of ADNs causes significant variability in greenhouse gas emissions in headwater stream networks. Significant work remains to understand how stream width and network length is controlled by groundwater interactions. Our limited dataset of thirteen surveys likely does not describe the full range of width distributions in headwater stream networks, especially in arid and humid tropical environments. This study's observation of a characteristic modal stream width suggests the existence of a most abundant depth and velocity within ADNs, a new hydrologic framework that may yield greater knowledge of stream generation processes and habitat distributions in headwater stream systems.

CHAPTER 2: PATTERNS OF RIVER WIDTH REVEALED BY THE SATELLITE-DERIVED NORTH AMERICAN RIVER WIDTH (NARWIDTH) DATASET[2]

2.1 INTRODUCTION

Rivers are fundamental to Earth's hydrological and biogeochemical cycles, they are biodiversity hot spots, and they provide vital water supply to human civilization. Despite their widespread importance, relatively limited empirical information on river channel form is available at continental scales to constrain river system models. These models commonly use spatially distributed measurements of river width, centerline location, and/or braiding index to estimate discharge (e.g. Gleason and Smith, 2014), flooding extent (e.g. Neal et al., 2012), landscape evolution (e.g. Lague, 2014), or biogeochemical processes (e.g. Gomez-Velez and Harvey, 2014; Kiel and Cardenas, 2014; Raymond et al., 2013). As models increase in spatial resolution, extent, and sophistication, they require high-resolution, large-scale river width datasets.

A key application of these river width datasets is estimation of the surface area of rivers at different scales. Globally, rivers are significant emitters of greenhouse gas and are estimated to outgas ~1.8 Pg C yr^{-1} of carbon dioxide (Raymond et al., 2013) and ~1.5 Tg CH_4 yr^{-1} of methane (Bastviken et al., 2011). Among other parameters, the surface area of rivers is a primary control on gaseous efflux and is used to estimate global emission rates. Presently, the most sophisticated

[2] This chapter has been published as a research article in the journal *Geophysical Research Letters*. The formatting has been changed to be consistent with the rest of the dissertation. The publication citation is as follows: Allen, G.H., Pavelsky, T.M., (2015), Patterns of river width and surface area newly revealed by the satellite-derived North American River Width (NARWidth) data set. *Geophysical Research Letters*. 42(2), 395-402. doi:10.1002/2014GL062764.

evaluations of global river surface area rely on: 1) calculating river width from digital elevation models (DEMs) by scaling width to upstream drainage area via downstream hydraulic geometry (DHG) relationships (Leopold and Maddock, 1953); 2) extrapolating river width and length from large to small river basins using Horton ratios (Horton, 1945); and 3) extrapolating empirical relationships between climate and percentage water cover from low- to high-latitude basins where high-resolution hydrologically conditioned topographic data does not exist (Raymond et al., 2013). Because this method relies on DHG scaling, which cannot account for anthropogenic modification of riverways, it may not accurately capture the true river surface area (Wehrli, 2013). Further, geographical variability in physical conditions including climate, tectonic deformation, and sediment supply and characteristics can also lead to a breakdown of DHG and climate-percentage water cover scaling (Ferguson, 1986; Park, 1977; Wohl, 2004).

Recent advances in image-processing algorithms have yielded large-scale, high-resolution river width surveys containing hundreds of thousands of measurements (e.g. Allen et al., 2013; Miller et al., 2014; O'Loughlin et al., 2013; Pavelsky et al., 2014a; Yamazaki et al., 2014). These compilations fall within a new class of fluvial geomorphology datasets that directly quantify river width continually downstream. Such datasets are facilitating new approaches for understanding river processes in much the same way that DEMs revolutionized analysis of fluvial systems. Here we present the North American River Width (NARWidth) dataset, the first continental survey of river width at mean annual discharge for rivers wider than 30 m. We analyze the continental-scale frequency distribution of river widths and compare the results to a DEM-derived width distribution. We then use the strong statistical relationship between river width and total river surface area of all rivers at that width to estimate the total surface area of North American rivers at mean annual discharge.

2.2 METHODS

We measured river width at mean discharge from a total of 1,756 Landsat scenes covering North America (see Appendix II for in-depth methods). For each Landsat path-row combination, we calculated the time of year when the observable rivers were most likely to be at mean discharge by analyzing mean monthly discharge records from the Global Runoff Data Center (GRDC) (Appendix Figure 2.1; GRDC, 2017). After acquiring cloud- and river ice-free imagery from the Global Land Cover Facility (glcf.umd.edu) and the USGS (earthexplorer.usgs.gov), we applied the Modified Normalized Difference Water Index formula (Xu, 2006) to Landsat reflectance values and created a binary land-water mask using dynamic thresholding (Li and Sheng, 2012). We visually inspected and corrected the land-water masks and calculated a channel centerline for all river reaches longer than 10 km using RivWidth software (Pavelsky and Smith, 2008). RivWidth computes the river width and braiding index at each centerline pixel and outputs the data as a georeferenced vector (Appendix Figure A2.2). We then flagged measurements of lakes and reservoirs included in the NARWidth dataset using GIS methods and existing water body datasets (Appendix Section A2.1.3).

Landsat-derived river width measurements were validated using 1,049 geographically distributed stream flow and river width records from the United States Geological Survey (USGS) and the Water Survey of Canada (WSC). We included only gauges with records that: (1) span at least 10 complete years of discharge measurement; (2) drain basins larger than 1,000 km^2; (3) are located within 1 km of a RivWidth centerline; (4) are not immediately adjacent to reservoirs or river confluences, and (5) have river width data available. We used daily discharge measurements to calculate the in situ mean annual discharge for each location (Kimbrough et al., 2003) and then compared the corresponding in situ width to the mean of the five spatially closest RivWidth measurements (Appendix Figure A2.3).

To assess conventional width datasets built using DHG, we compared NARWidth to a DEM-derived river width dataset. The DEM-derived dataset was produced by Pavelsky et al. (2014b) to evaluate the spatial distribution of rivers observable by the planned Surface Water and Ocean Topography (SWOT) satellite mission. The dataset was created using methods similar to those developed by Andreadis et al. (2013), except the HYDRO1k DEM (USGS, 2014; Verdin and Verdin, 1999) was used to calculate river width rather than the HydroSHEDS DEM (Lehner et al., 2008). This DEM-derived width dataset was built by using drainage area from HYDRO1k and mean annual discharge from the GRDC in combination with a global averaged width-discharge equation (Moody and Troutman, 2002) to estimate mean annual river width along HYDRO1k DEM streamlines (see Pavelsky et al. (2014b) for a detailed methodology description).

For both the Landsat- and DEM-derived datasets, we analyzed the distribution of river length and surface area binned by river width from 100 to 2000 m, excluding measurements of reservoirs, lakes, and Greenland rivers. River length was calculated using the Euclidean distance between each centerline pixel and the next adjacent centerline pixel. River surface area was calculated by summing the product of river width and length at each centerline pixel (Appendix Figure A2.2). We established a minimum width threshold of 100 m because we are not confident that NARWidth includes all rivers with widths below this threshold (Miller et al., 2014). We excluded rivers wider than 2000 m because they only account for 0.6% of all measurements but significantly skew the results of the analysis.

2.3 RESULTS

2.3.1 Dataset and Validation

The North American River Width (NARWidth) dataset contains 6.7×10^{6} georeferenced

measurements of river width ≥30 m and an additional 1.3×10^6 flagged width measurements of reservoirs and lakes that are connected to the fluvial network (Figure 2.1). In total, NARWidth measures 2.39×10^5 km of rivers with widths ≥30 m corresponding to a water surface area of 4.43×10^4 km^2 and 1.1×10^5 km of rivers wider than 100 m (3.64×10^4 km^2 of water surface area). NARWidth includes rivers ranging from ~4th to 10th Strahler stream orders (Downing et al., 2012; Strahler, 1957). The dataset includes measurements of rivers above 60°N where high-quality river centerline and width data is largely unavailable, but excludes very large lakes (e.g. the Great Lakes), ephemeral streams, deltaic systems, and human-made canals. Additionally, NARWidth includes a braiding index field, defined as the number of channels at each river cross section. The braiding index only includes river channels wider than 30 m, a limitation imposed by the spatial resolution of Landsat imagery. NARWidth is the first continental-scale morphometric survey of rivers at mean discharge and is available for download at http://gaia.geosci.unc.edu/widths/na/NARWidth.

NARWidth width measurements show very little mean bias (-0.35 m) relative to in situ width measurements at mean discharge, suggesting the Landsat scenes were sampled at times that, on average, matched mean discharge timing. The RMSE between NARWidth and in situ widths is 38.0 m, a length similar to the minimum theoretical uncertainty of Landsat-derived river widths calculated from a binary water mask (Pavelsky and Smith, 2008). The RMSE value also incorporates several other sources of error, including differences in discharge between the remotely sensed and in situ measurements and error in the in situ width measurements.

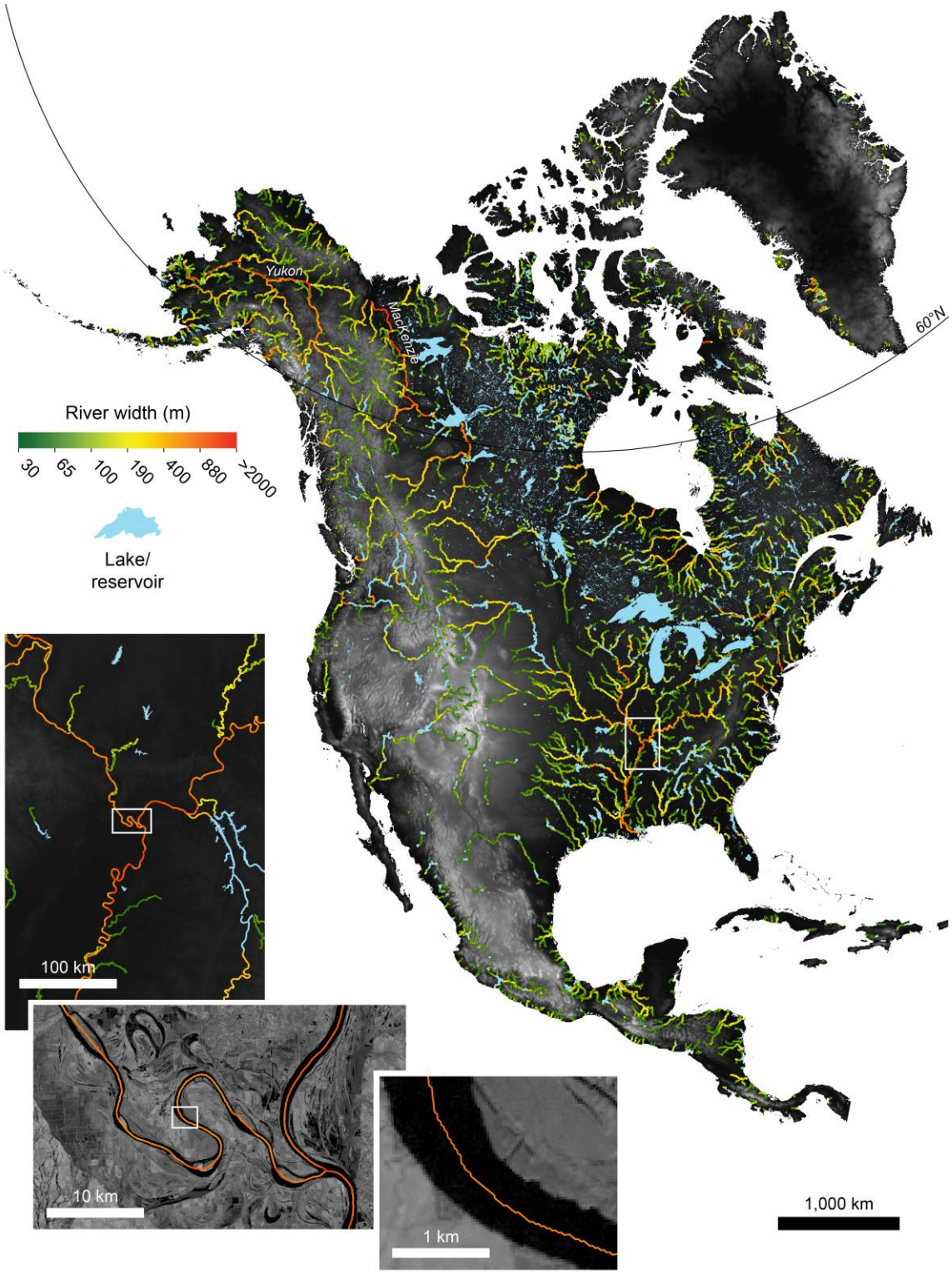

Figure 2.1 - Map of North American River Widths

Inset boxes show levels of detail available at finer spatial resolutions. Note the color bar is stretched geometrically to accent width variability.

To avoid bias from outliers, we used the Theil-Sen median estimator (Sen, 1968) to derive a robust linear regression between NARWidth and in situ width measurements (Figure 2.2). Regression of in situ widths ≥100 m yields a slope that deviates by 3% from unity, but inclusion of all river width data (≥30 m) produces a slope that deviates by 16%. This deviation is expected because NARWidth is more likely to include overestimates of river width compared to underestimates where river width approaches the resolution of the Landsat imagery. For example, NARWidth never includes underestimates of 30 m wide rivers because they are narrower than one Landsat pixel, but it will include overestimates of these rivers. Goodness of fit ($r_s = 0.83$) was characterized using Spearman's nonparametric correlation coefficient (Spearman, 1904). Overall, comparison with in situ measurements suggests that NARWidth provides, on average, an accurate representation of river widths at mean annual discharge to the extent that this is possible from Landsat imagery.

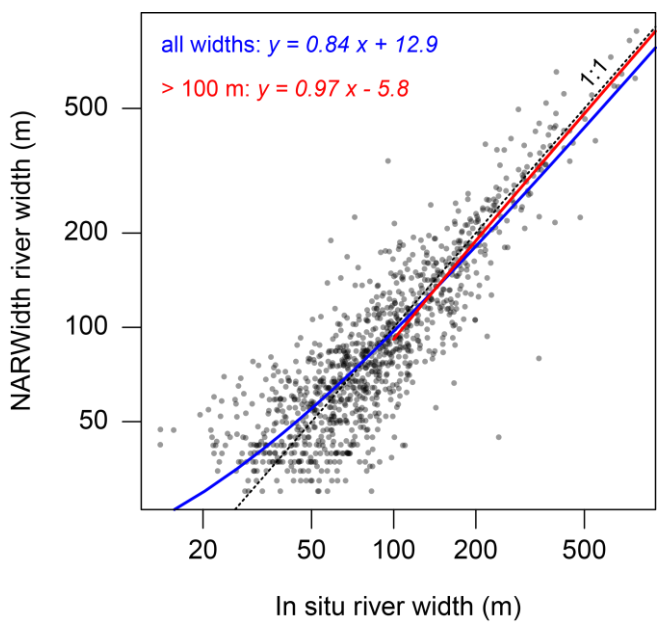

Figure 2.2 - NARWidth river width validation
NARWidth widths were compared to USGS and WSC in situ river width measurements at 1,049 locations.

2.3.2 River width distributions

The NARWidth dataset allows for the first analysis of the frequency distribution of river width measurements on a continental scale. The distribution of average river length binned by river width closely follows a power law of river width such that,

$$Length = C \times Width^{-\alpha} \tag{2.1}$$

where $C = 3.24\times10^{10}$ $m^{\alpha+1}$ and $\alpha = 2.18$ (Figure 2.3a). C and α were calculated using maximum likelihood estimation to avoid assumptions associated with regression analysis of binned data (Clauset et al., 2009; Gillespie, 2014). While other functions may also characterize the distribution of river widths, we use a power function because power laws are regularly used in hydraulic scaling and the curve closely fits the length of rivers from 100 to 2000 m wide ($r^2 >$ 0.996, $p < 0.001$). Outside of this range, the function overestimates the length of rivers, particularly for the widest observed rivers where the distribution of river width appears to deviate from a power-law spectrum. The upper tail, defined here as widths greater than 2000 m, is composed primarily of measurements from large, multichannel rivers. Sixty-three percent of river widths greater than 2000 m are from multichannel rivers compared to only 23% of rivers between 1000 and 2000 m wide. Geographically, multichannel rivers greater than 100 m wide make up 26.2% of all rivers north of 60° N, while only composing 14.9% of rivers south of 60° N.

Although the river width distributions derived from the HYDRO1k DEM and NARWidth both closely fit power law functions, there are two key differences between them. First, the DEM-derived width distribution is characterized by a higher exponent ($\alpha = 2.64$ in equation 1) than the NARWidth distribution ($\alpha = 2.18$), signifying that the DEM-derived dataset contains a lower proportion of wide rivers (Figure 2.3b). Second, the DEM-derived dataset does not include

rivers wider than 894 m, resulting in a frequency distribution with a relatively truncated upper tail (Figure 2.3b inset).

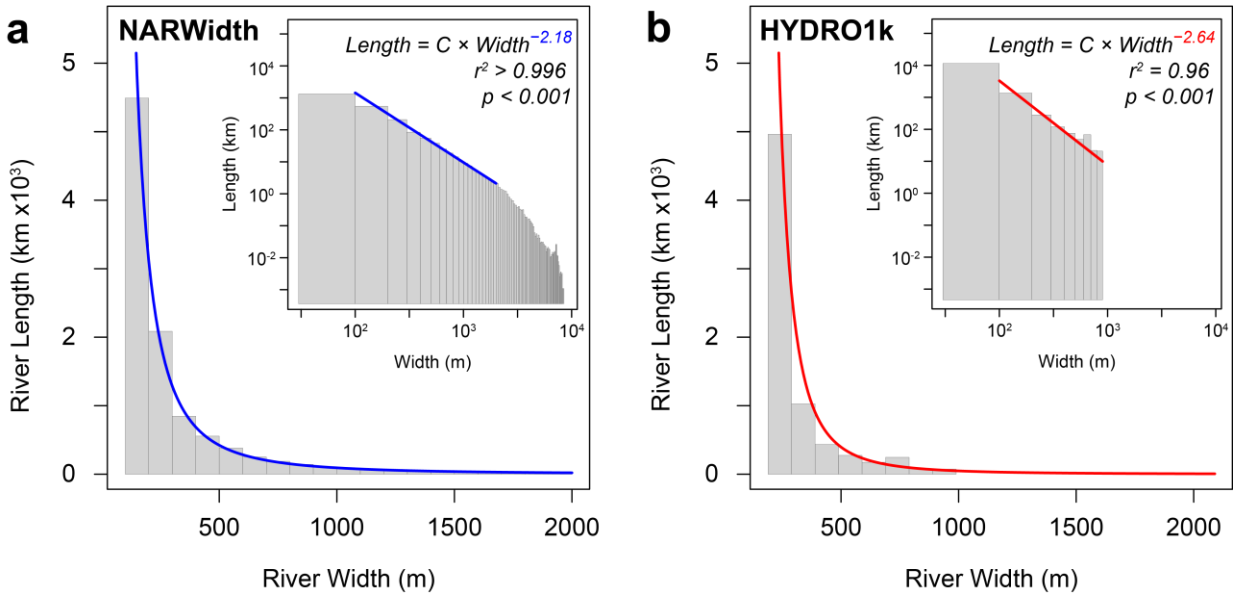

Figure 2.3 - River width distributions from 100 to 2000 m
Insets show the full range of observed data in log space. Width distributions are described by a power function (equation 1) with r^2 values estimated using a linear regression of binned logged data (White et al., 2008). **a,** NARWidth-derived distribution where $C = 3.24\times10^{10}$ $m^{-1.18}$. **c,** HYDRO1k DEM-derived distribution where $C = 3.67\times10^{10}$ $m^{-1.64}$.

2.3.3 Total surface area of North American streams and rivers

We used the distribution of total river surface area binned by width to estimate the overall surface area of all streams and rivers in North America (Figure 2.4). A power function closely describes the distribution of surface area of rivers 100 to 2000 m wide ($r^2 > 0.996$, $p < 0.001$) and is used to extrapolate surface area to rivers narrower than 100 m. The predictable surface area versus width relationship observed here is also likely to persist for these small streams because Horton ratios relating stream order to stream width and total length apply down to first order perennial streams (Downing et al., 2012; Morisawa, 1962). As surface area is the product of stream length and width, it should also scale as a power law down to first order streams.

A considerable unknown is the appropriate lower width boundary of the surface area extrapolation. Studies that use DEMs to extract fluvial networks typically assume a critical threshold drainage area (or support area) ranging from 0.1-1 km^2, which correspond to river widths of ~0.8 m to 2 m (Beighley and Gummadi, 2011; Butman and Raymond, 2011). Instead of a threshold in drainage area, we use a threshold in river width directly. The value of this lower width threshold substantially influences the surface area calculation because of the non-linear relationship between river width and total channel length. The relationship between the lower width threshold (m) and the total surface area of North American rivers (km^2) is described by

$$\textit{Total Surface Area} = 183683 \times \textit{Lower Width Threshold}^{-0.175} - 44468. \qquad (2.2)$$

This relationship only applies for lower width thresholds below 1000 m because above this width, equation (2) begins to deviate from the observed data. Downing et al. (2012) compiled a list of stream order vs. mean width data from 46 perennial first order stream segments worldwide and found the median stream width is 1.6 ± 1.1 m (1σ confidence intervals). Using these widths as the lower width threshold in equation (2), the total surface area of permanently flowing North American rivers and streams is $1.24^{+0.39}_{-0.15} \times 10^5$ km^2 or $0.55^{+0.17}_{-0.07}$% of the terrestrial land surface. This surface area value is likely an underestimation of total river surface area because: 1) it is based on the average width of first order streams rather than the average width at stream heads (streams as narrow as 0.18 m have been observed by Zimmerman et al., 1967); and 2) it excludes the surface area contribution of ephemeral streams, estimated to account for 2-3% of global river surface area (Downing et al., 2012; Raymond et al., 2013).

2.4 DISCUSSION

2.4.1 Distribution of river widths

Differences between NARWidth and the DEM-derived width dataset likely arise from bias in measuring river width at gauge stations and oversimplifications involved in DHG scaling. The width-discharge relationship used to produce the DEM dataset was developed using measurements largely collected at gauging stations (Moody and Troutman, 2002). Stream gauges are typically located at stable, single channel sites, often near bridges or other fixed structures, leading to a possible negative bias of measured river widths relative to the true width distribution (Park, 1977). Because multichannel rivers tend to be wider and because their widths are more sensitive to changes in discharge than are single channel rivers (Smith et al., 1996), average river widths away from gauge stations may be underestimated if only widths at gauge stations are used. Given the non-linear frequency distribution of river widths (equation 1, Figure 2.3), this systematic underestimation of river width may result in an artificially high α value for the DEM-derived width distribution relative to the NARWidth distribution.

Additionally, DHG predicts that the maximum river width within a basin is located wherever discharge is greatest, usually at the basin outlet. Direct observations from Landsat imagery do not fully support this model. The widest width measurements in NARWidth are primarily from large, braided river systems flowing through floodplains (e.g. the Yukon and Mackenzie rivers, Figure 2.1). These locations are examples of river form impacted by unusual physical conditions such as, in the case of braided rivers, abundant sediment supply (Rosgen, 1994). Such physiographic conditions can result in substantial deviation from strict width-discharge relationships that are not captured by DHG. Thus, applying generic DHG scaling over large scales and differing river morphologies should be done with caution.

2.4.2 River surface area estimation

We estimated the total surface area of North American rivers by developing a relationship between river width and total surface area binned by width for rivers 100-2000 m wide, which we then use to extrapolate surface area for rivers narrower than 100 m (Figure 2.4). We based this extrapolation on classic Hortonian analysis which predicts that the distribution of river surface area will display statistical self-similarity at different spatial scales, indicating that similar processes act on river form over a wide range channel sizes (Rodríguez-Iturbe and Rinaldo, 2001). With decreasing basin size, however, hillslope processes begin dominating and the fractal relationships between total stream length and width must inevitably break down (Benda et al., 2005). We propose that the lower width threshold should be the geometric mean width at stream heads, but we are unaware of any robust quantitative information to constrain this value. Headwater stream networks are highly dynamic, largely dependent on changing hydrologic conditions (Godsey and Kirchner, 2014). Further work is needed to quantify the distribution of river length, width, and discharge in headwater catchments to better constrain the frequency distributions of small streams that cannot be measured from remotely sensed datasets.

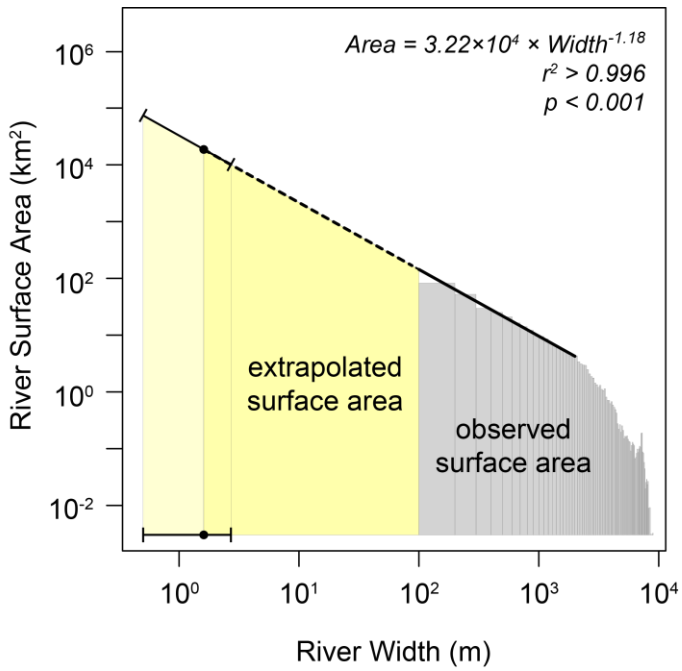

Figure 2.4 - River surface area binned by river width
A power function was fit to data from widths 100 to 2000 m (solid line) and used to extrapolate total surface area of rivers less than 100 m wide (yellow polygon). Error bars denote the upper and lower width threshold used in the surface area extrapolation (1.6±1.1 m). The extrapolated surface area was then added to observed surface area (gray bars) to estimate total river surface area of North America.

Previous estimates of river surface area across a large range of scales and physiographic conditions vary between 0.3% and 1.5% of watershed area, and our estimate of $0.55^{+0.17}_{-0.07}$% falls within this interval (Davidson et al., 2010; Downing et al., 2012; Welcomme, 1976). In the contiguous United States, past estimates range from 0.52 to 0.56%, closely matching our estimate for the entire North American continent (Butman and Raymond, 2011; Downing et al., 2012; Leopold, 1962). Analyses of total global stream and river surface area estimate that rivers cover 0.30–0.56% of land surface (Downing et al., 2012; Raymond et al., 2013). These studies rely on a relatively limited number of width measurements ($N < 1.0\times10^3$) to conduct stream order-scaling analysis. We avoid depending on potentially biased width-order analysis by

building an extensive continental inventory of river width and length and analyzing the frequency distribution of surface area itself.

As part of a pioneering global carbon efflux study, Raymond et al. (2013) estimated that the total surface area of North American rivers is 0.46% of continental surface area. Raymond et al. estimated that, on average, 14% of global surface area is frozen, and this fraction is removed from their analysis. They acknowledge that the impact of river ice on carbon efflux rates is poorly understood, and we do not attempt to account for the proportion of river surface area that is frozen. Noting this limitation of comparability between the surface area estimates, our likely conservative estimate of $0.55^{+0.17}_{-0.07}$%, which excludes streams narrower than 1.6±1.1 m, is 20^{+38}_{-15}% larger than the Raymond et al. value. Further, Raymond et al. extrapolated width-order relationships down to streams with a drainage area of ~0.1 km^2, amounting to a lower width threshold of less than 1 m (Beighley and Gummadi, 2011). The discrepancy between river surface area estimates may arise from the DHG-based extrapolation methods employed by Raymond et al. Their evaluation relies on a global DHG formula that scales river width with regional discharge. Thus, many of the same problems in estimating river width discussed in Section 2.4.1 may also apply to their estimate (e.g. underestimating the abundance of wide rivers). Our estimation of North American river surface area indicates that gaseous emissions from rivers should likely be revised upwards compared to most recent estimates.

CHAPTER 3: THE GLOBAL ABUNDANCE OF RIVERS AND STREAMS DERIVED FROM SATELLITE IMAGERY[3]

3.1 INTRODUCTION

Rivers are the chief source of renewable water to humans and freshwater ecosystems (Vorosmarty et al., 2010) and are also responsible for some of the largest natural disasters in history (Hough, 2014). Through downstream transport of sediment, solutes, and other material, they organize landscapes and play a key role in biogeochemical cycling. Rivers outgas significant quantities of greenhouse gasses to the atmosphere including methane (Bastviken et al., 2011), nitrous oxide (Beaulieu et al., 2011), and carbon dioxide (Raymond et al., 2013). For example, the global river network is estimated to emit ~1.8 PgC yr^{-1} of carbon dioxide to the atmosphere (Raymond et al., 2013), approximately equivalent to one fifth of fossil fuel emissions (Peters et al., 2013).

Despite the importance of rivers, there exists a dearth of observational information on river abundance, morphology, and discharge globally. Global river gage data is highly fragmented and publicly unavailable in most regions, impeding understanding of Earth's river water resources (Pavelsky et al., 2014c). Further, the location, morphology and size of Earth's rivers are commonly simulated using flow routing algorithms based on static digital topography and network scaling principles (Andreadis et al., 2013; O'Callaghan and Mark, 1984). Assumptions involved in these simulations are the source of error in many large-scale hydraulic,

[3] This chapter is in preparation for submission to the journal *Nature*. Some of the supplementary figures and text have been integrated into the main text and the formatting has been updated to conform to the other chapters in this dissertation. The publication citation is as follows: Allen, G.H., Pavelsky, T.M., (*in prep.*), The global abundance of rivers and streams derived from satellite imagery. *Nature*.

hydrologic, and biogeochemical applications (Andreadis et al., 2013; Pavelsky et al., 2014c; Raymond et al., 2013; Wehrli, 2013).

Remote sensing offers a compelling solution to many of the limitations facing large-scale river hydrology. Satellite images can be used to quantify river planform geometry (e.g. spatial variations of river width along a river network), which then can be used in an inverse problem framework to solve for complex processes occurring at the earth's surface. Many modern hydrologic, geomorphic, and biogeochemical models include river geometry as a fundamental predictor variable (e.g. Allen et al., 2013; Gleason and Smith, 2014). While other empirical river width datasets exist, their coverage is not global, or coarse spatial resolution inhibits their usefulness to river system models (Allen and Pavelsky, 2015; Miller et al., 2014; Pavelsky et al., 2014a; Yamazaki et al., 2014).

Global-scale evaluations of fluvial biogeochemical processes rely on the accurate representation of river and stream abundance and morphology (Butman et al., 2016; Downing et al., 2012; Raymond et al., 2013; Richey et al., 2002). Current knowledge of the size distribution of the world's rivers and streams is based on a series of highly unconstrained hydrologic, geomorphic, and fractal river network scaling extrapolations (Downing et al., 2012; Raymond et al., 2013). These approaches have led to estimates of global river area ranging from 485,000 km^2 to 845,000 km^2, a difference in area equivalent to the size of Germany (Downing et al., 2012; Raymond et al., 2013). This unconstrained area is a significant source of uncertainty in global river-atmosphere biogeochemical flux evaluations (Raymond et al., 2013; Wehrli, 2013).

3.2 METHODS

To characterize the global abundance and morphology of rivers, we built the Global River Widths from Landsat (GRWL) Database, the first global database of planform river

geometry at a constant frequency discharge (Figure 3.1). We used a global compellation of 3693 gage station data (GRDC, 2017) to determine during which months rivers were commonly near mean discharge and acquired 7376 Landsat TM, ETM+ and OLI scenes during the most likely months (Appendix Section A2.1.1, A2.1.2). We applied previously published image processing techniques (Allen and Pavelsky, 2015; Pavelsky and Smith, 2008) to classify rivers and measure river morphology (Appendix Section A2.1.3). At a 30-m spatial resolution, GRWL contains width measurements of >2.1M km of rivers 30 m or wider at mean annual discharge. GRWL also contains river centerline location, braiding index, river network topological information, and over 7.6M measurements of flagged lakes/reservoirs and canals connected to the fluvial network.

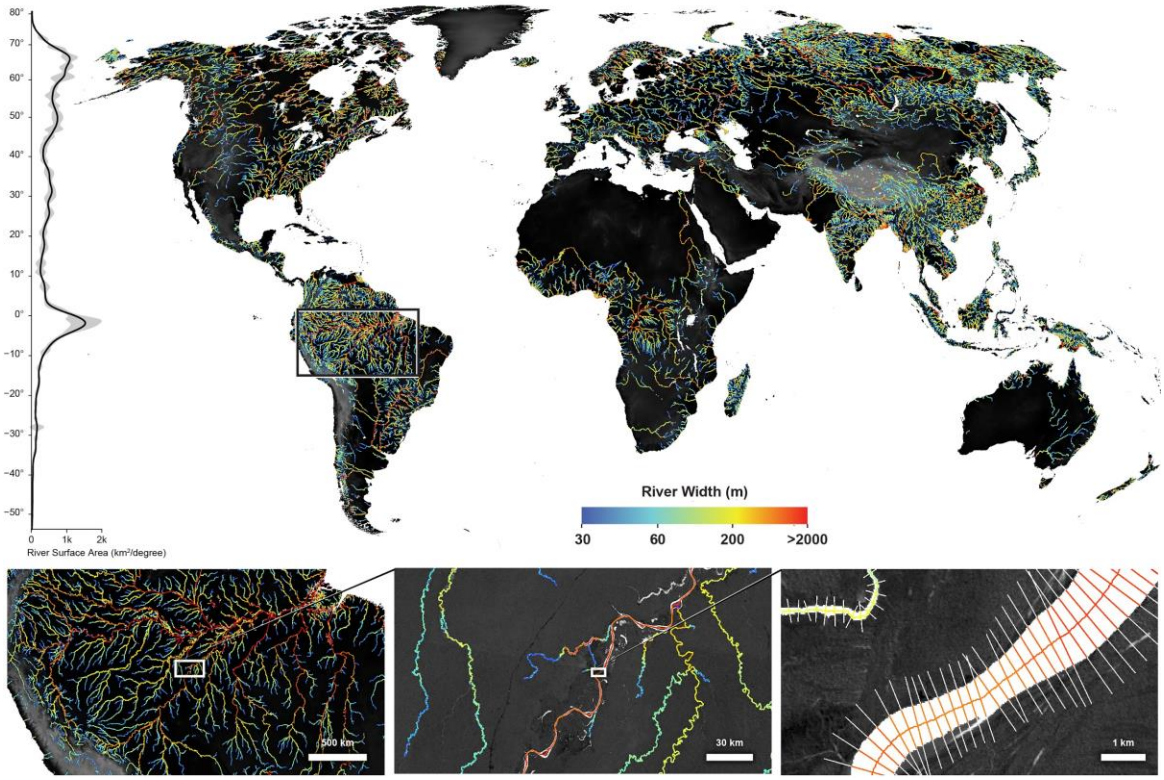

Figure 3.1 - The Global River Widths from Landsat Database (GRWL) contains over 58M measurements of river planform geometry.

We validated GRWL width data using river width measurements from the United States Geological Survey and the Water Survey of Canada at 1146 gaging stations (Figure 3.2, Appendix Section A2.2). GRWL width measurements show little mean bias (1.71 m) compared to in situ width measurements at mean discharge, suggesting the Landsat scenes were sampled at times that, on average, matched mean discharge timing. The RMSE between GRWL and in situ widths is 38.2 m, a length similar to the minimum theoretical uncertainty of Landsat-derived river widths calculated from a binary water mask (Pavelsky and Smith, 2008). The RMSE value also incorporates several other sources of error, including differences in discharge between the remotely sensed and in situ measurements and error in the in situ width measurements. Regression of in situ widths ≥100 m using the Theil-Sen median estimator (Sen, 1968) yields a slope that deviates by 2% from unity, but inclusion of all river width data (≥30 m) produces a slope that deviates by 16% (Figure 3.2). This deviation is expected because GRWL is more likely to include overestimates of river width compared to underestimates where river width approaches the resolution of the Landsat imagery. We characterized goodness of fit using Spearman's nonparametric correlation coefficient, $r_s = 0.82$ (Spearman, 1904). Overall, comparison with in situ measurements suggests that GRWL provides, on average, an accurate representation of river widths at mean annual discharge to the extent that this is possible from Landsat imagery.

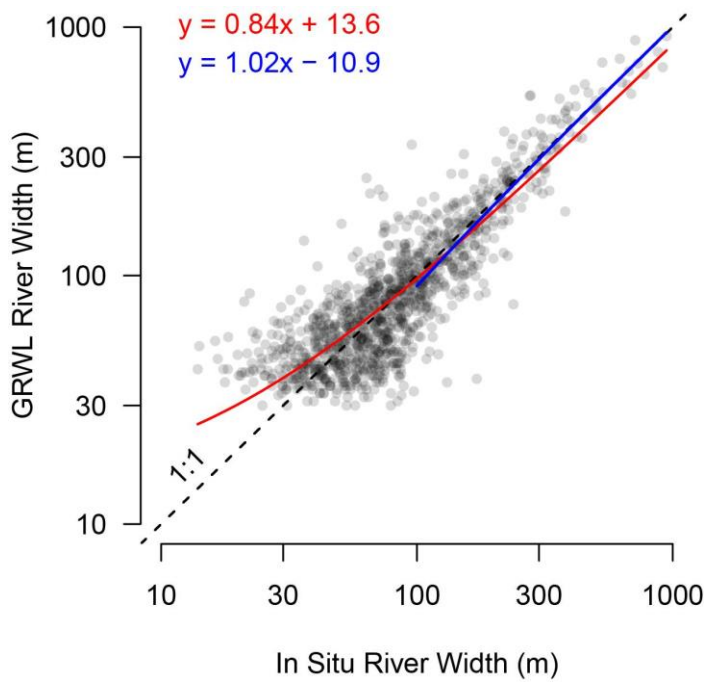

Figure 3.2 - GRWL validation with 1146 in situ width measurements from the United States Geological Survey and the Water Survey of Canada.

3.3 RESULTS

GRWL provides the most direct and comprehensive quantification of river area worldwide to date. By summing the product of each river width measurement and the downstream distance between each width measurement (Figure 3.3a), the total area of rivers measured by GRWL is 415,000 km^2, or 0.31% of Earth's land surface (Antarctica was removed in this analysis). We exclude from our river area analysis lakes and canals, we and remove rivers with elevations less than 1 m above sea level to reduce tidal influence on river morphology. The sum of surface area of rivers wider than 100 m, where GRWL data is most complete and accurate, is 344,000 km^2, a value consistent with a previous aggregate estimate of 360,000 km^2 for rivers wider than 90 m (Lehner and Döll, 2004). However, these river area estimates exclude small rivers and streams that occupy a significant portion of the fluvial network.

To account for rivers and streams <100 m wide, too narrow to be accurately measured

with Landsat imagery, we use a statistical extrapolation technique. In all global river basins (Lehner and Grill, 2013) containing >5000 river measurements >100 m wide (N=415), we fit a Pareto distribution to the raw width data using maximum likelihood estimation (Figure 3.3b). On average, the Pareto distribution fits the data well (mean Kolmogorov-Smirnov D=0.14±0.07, p<0.001, all confidence intervals 1σ), as predicted by theory (Allen et al., in review; Horton, 1945; Kirchner, 1993; Strahler, 1957) and other empirical evidence (Downing et al., 2012; Morisawa, 1962; Trampush et al., 2014). While in Chapter 1 we found that stream widths in headwater networks are lognormally distributed, we find that in large river systems, river area, and hence river width, is Pareto distributed. The departure from the Pareto distribution at the upper tail, seen in Figure 3.3b, is an expected result of the finite-size effect, or, in other words, the limiting consequence of basin size on maximum river width (Rodríguez-Iturbe et al., 1992). In Chapter 1, we find that first-order streams exhibit a highly consistent median wetted width of 0.32±0.08 m (mean of the medians), largely independent of physiographic and hydrological conditions. By extending the GRWL-derived Pareto distribution to a river area corresponding to this median first-order stream width, we calculate total river area in each basin by adding the estimated stream and river area to the observed river area (Figure 3.3b, Figure 3.4a)

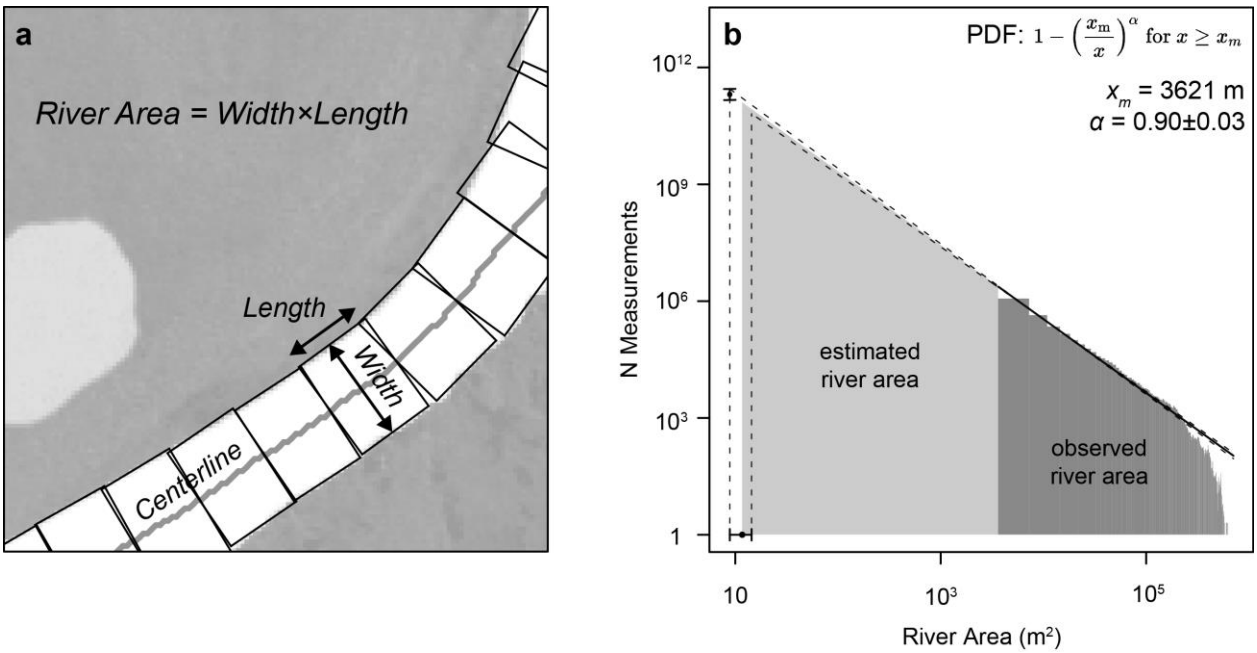

Figure 3.3 - Estimating surface area of rivers and streams within a basin
a, Discretized river area measurements. **b,** River area extrapolation from satellite-measured wide rivers to narrower rivers and streams in an example basin (Amazon shown). Pareto scale (χ_m) and shape (α) parameters.

The addition of the total stream and small river area estimate to the observed river area sums to a global total of 672,000 km^2. This value is similar to 687,000 km^2, or the river area calculated when all the basins are combined and a single Pareto curve is fit to the collective data, as has been done in previous studies (Allen and Pavelsky, 2015; Downing et al., 2012). However, many basins do not contain >5000 river width measurements >100 m (basins with hatched lines in Figure 3.4a), because their climates are arid or because their accumulated drainage areas are insufficient to generate wide rivers. To represent rivers and streams located within these basins, we use basins with areas >100,000 km^2 (N=146) to develop an empirical relationship between average basin aridity (Zomer et al., 2008), basin size, and percent basin occupied by rivers (R^2=0.71, p<0.001, Figure 3.5). Larger basins tend to have a larger river area because they contain higher-order channels. Greenland is excluded from this analysis. Including

these basins in the total increases the global river area at mean annual discharge to 745,000 km^2, or 0.55% of Earth's land surface (Figure 3.4a).

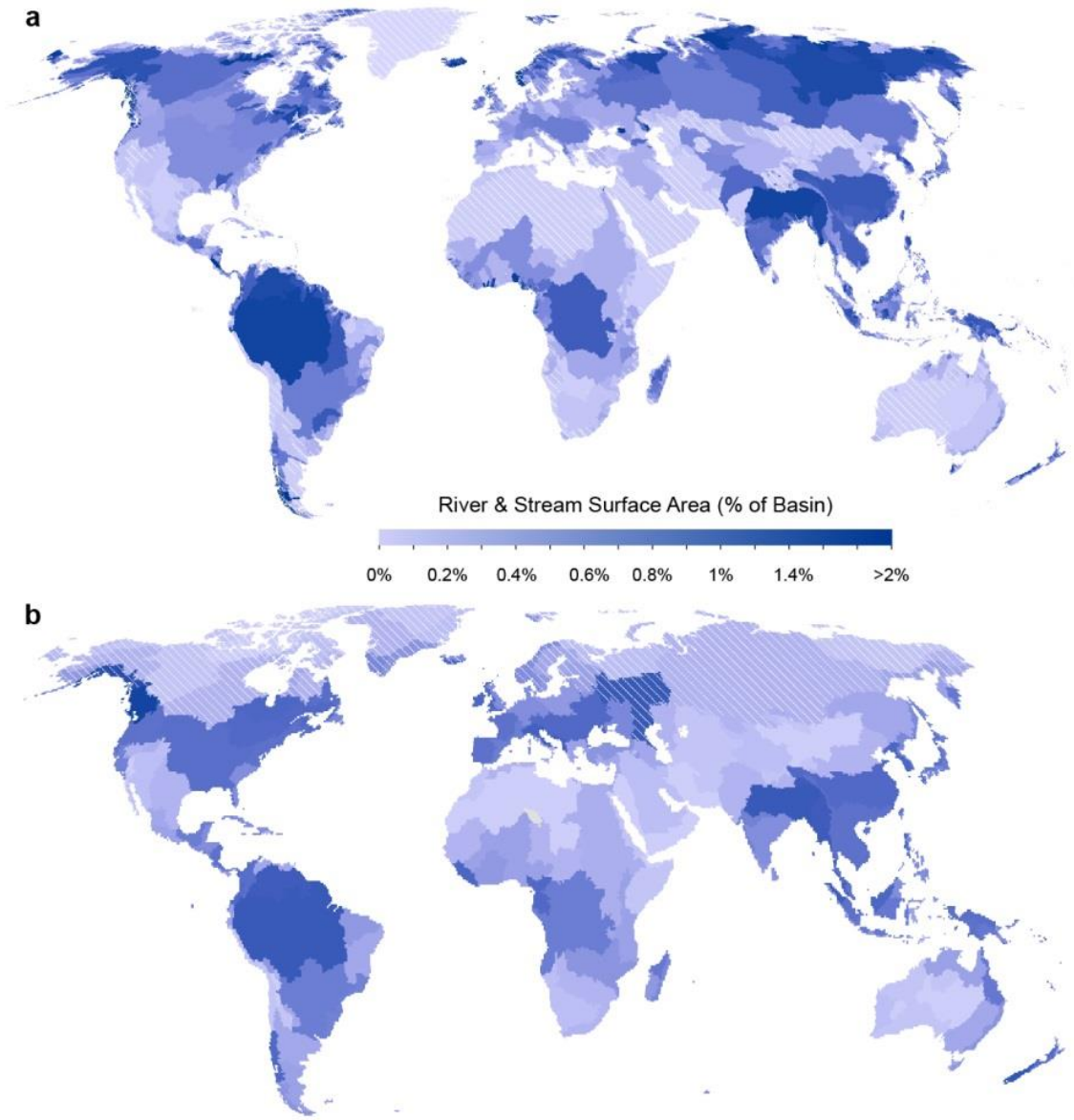

Figure 3.4 - Percentage of global basins occupied by rivers and streams
a, GRWL estimate; **b,** Previous best estimate (Raymond et al., 2013). Rivers are significantly more abundant in tropical regions than previously thought. Large discrepancies in high latitude river area likely stem from the incorporation of the effect of river ice on effective river area in Raymond et al. Hatch lines show basins with climate-based surface area estimates in both studies (Figure 3.5).

3.4 DISCUSSION AND CONCLUSIONS

The global river area presented here is greater than previous estimates, but falls within the range of their uncertainty. Downing et al. (2012) estimates that the global river surface area is between 485,000 and 662,000 km^2 by assuming globally invariant scaling relationships between stream order and river length and width. Raymond et al. (2013) uses DEM-derived flow paths, simulated discharges, stream order scaling extrapolations, and climate extrapolations to estimate that river area is 540,000 km^2. A significant source of unknown in these previous studies originates from estimating the area of intermittent and ephemeral streams. We significantly reduce this uncertainty by relying on the observed consistency of headwater stream widths and the observed length-width relationships in the mean discharge of large rivers.

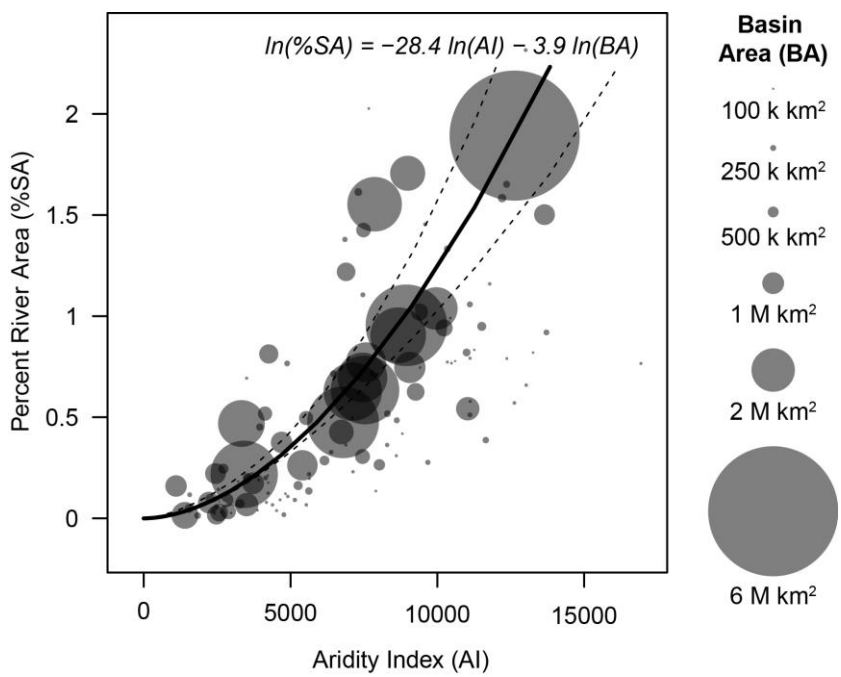

Figure 3.5 - River area-climate relationship between global basins
Multiple linear regressions of log-transformed basin percent river area (%SA) against Aridity Index (AI) and basin area (BA). Regression weighted by basin area. Dashed lines are 1σ error bounds.

We find substantial basin-to-basin spatial variability in our river area estimates (Figure 3.4a). For example, rivers occupy 1.9% of the Amazon Basin at mean annual discharge, an area ~48-86% greater than previous estimates (Beighley and Gummadi, 2011; Raymond et al., 2013) but less than total inundation area estimates of the wet-season flood (Coe et al., 2008; Hess et al., 2003). Compared to the only other existing region-by-region global estimate of river area (Raymond et al., 2013), we estimate a greater abundance of river area in the tropics and less river area in temperate and desert regions (Figure 3.4). Tropical rivers emit a greater amount of greenhouse gas to the atmosphere per unit river area, and thus our results indicate that estimates of total river outgassing rates should likely be revised upwards.

Our new satellite based estimate of river area can be used to improve the accuracy of current estimates of outgassing from the fluvial network. Further, the GRWL Database has significant potential to improve representation of fluvial processes and understanding of river water resources globally (Clark et al., 2015). GRWL is being used to improve hydrologic models, and organize remote sensing surface water observations. GRWL will be used to identify rivers segments observable by the Surface Water and Ocean Topography Mission (SWOT), scheduled to launch in 2021 (Biancamaria et al., 2016). GRWL can also be used to organize large multitemporal datasets of surface water dynamics (Pekel et al., 2016) to study fluvial geomorphology and changes in river discharge through time and space at the global scale (Gleason and Smith, 2014).

APPENDIX 1: HEADWATER STREAM WIDTH DISTRIBUTION METHODS AND RESULTS

A1.1 Carbon efflux estimates methods

To estimate carbon efflux within each North American catchment, we used topographically derived USGS flowlines to estimate stream surface area and stream gas transfer velocity, and CO_2 efflux methods described in detail by Butman et al. (2016). We did not conduct this analysis in the New Zealand watersheds because no suitable flowline dataset was available in this region. In the conterminous United States, we used NHDPlus V2 flowlines (http://www.horizon-systems.com/nhdplus/NHDPlusV2_home.php) to calculate carbon efflux, and in the Caribou Creek catchment in Alaska, where NHDPlus data are unavailable, we used EDNA flowlines (http://edna.usgs.gov). NHDplus V2 and EDNA flowlines are derived from merging the Nation Hydrology Dataset (NHD) with the National Elevation Dataset (NED). The NHD contains perennial, intermittent, and ephemeral streamlines that were field mapped by the USGS. Thus, the exact conditions in which the NHDplus V2 flowlines represent are poorly constrained.

In each study catchment, we used the median (5th and 95th percentile ranges) dissolved CO_2 concentrations and temperatures for first-order stream systems in the larger 2-digit USGS hydrologic unit code (HUC) region. For each HUC, we used established hydraulic geometry equations[3] to estimate first-order stream velocity from the stream slope and the median discharge values provided by the USGS flowline datasets. Using these stream velocities, we estimated the median CO_2 gas transfer velocity for first-order streams in each HUC after Raymond et al. (2013). For the Stony Creek Research Watershed where no NHDPlus flowlines exist, we used the discharge measurements taken at the catchment outlet to directly estimate velocity. We calculated the amount of stream surface area using three different methods: 1) field measured

surface area; 2) Equation 1.10-derived surface area; and 3) USGS flowline-derived surface area (Raymond et al., 2013). Then we ran a Monte Carlo simulation to estimate the median and 5th-95th percentile ranges of potential CO_2 efflux (Appendix Table A1.4).

Table A1.1 - Attributes of the seven physiographically contrasting study catchments.

Catchment	North Branch Kings Creek (K1B tributary)	Upper Sagehen Creek sub-catchment	Upper Elder Creek	C1 tributary of Caribou Creek	V40 Stream sub-catchment	Blue Duck Creek sub-catchment	Stony Creek Research Watershed
Location	Konza Prairie Biological Station, KS	Sagehen Creek Field Station, CA	Angelo Coast Range Reserve, CA	Caribou-Poker Crk. Research Watershed, AK	V40 Stream, NZ	Blue Duck Creek, NZ	Duke Forest, NC
Date Surveyed	May 25-26, 2015	May 29-30, 2015	Jun 1-2, 2015	Jun 20, 2015	Jun 28-29, 2015	Jul 12-14, 2015	Sep 27, 2015
Surveyors	G Allen, A Tashie	G Allen, A Tashie	G Allen, A Tashie	G Allen, D Butman	E Barefoot	E Barefoot, E Beckham	G Allen, E Barefoot
Outlet Lat. (DD)	39.0972	39.433	39.718	65.1486	-41.77	-42.278	36.04
Outlet Lon. (DD)	-96.5714	-120.285	-123.604	-147.637	171.777	173.746	-79.067
Drainage Area (ha)	180	449	362	258	6	56	128
Altitude Range (m)	113-135	635-811	175-381	401-657	699-741	136-400	158-226
Stream Network Relief (m)	21.5	131.6	160.2	89	39.5	95.2	36.9
N Width Obs.	1,797	1,422	1,044	363	249	519	805
ADN Length (km)	8.99	7.11	5.22	1.82	1.25	2.6	4.03
ADN Drainage Density (km^{-1})	4.99	1.58	1.44	0.7	20.75	4.63	3.15
% Basin Stream Surface Area	0.46	0.12	0.14	0.02	0.91	0.4	0.31
1st Order Median Width (cm)	30.5	22.6	35.6	20.3	22.9	23.4	30.5
Mode Width (cm)	34.6	25.2	27.4	27	21.5	21.6	32.7
Gage	USGS gage	USGS gage	USGS gage	CPCRW gage	West Coast Regional Council gage	Environment Canterbury gage	Duke Forest Research Watershed
Gage Location	Kings Creek near Manhattan (downstream from catchment)	Sagehen Creek near Truckee (downstream from catchment)	Elder Creek near Branscomb (downstream from catchment)	Caribou Creek near Poker Flat (downstream from catchment)	Buller River at Te Kuha (on a nearby river)	Lyell Creek at Warren Creek Confluence (on a nearby creek)	Stony Creek near Hillsborough (at bottom of study catchment)
Gage Drainage Area (km^2)	10.6	27.2	16.8	104	6,350	64	1.3
Flow Record Length (yrs)	37	62	50	11*	6	12	2
Flow Percentile (1σ)	87.4±8.4	19.4±1.8	32.1±0.9	6.6±0	73.1±0.6	2.3±0.5	34.6±0
Fig. 1 DEM source	FEMA 2006 LIDAR†	NED‡	NED‡	GDEM V2§	LINZ§	LINZ‖	NCFMP LIDAR¶
Fig. 1 DEM resolution (m)	2	10	10	15	8	8	6
Lithology	Interbedded mudstone and limestone	Siliceous intrusives and metamorphics with glacial deposits	Sandstone and mudstone	Greenschist and loess deposits	Interbedded coal and sandstones	Sandstones and mudstone	Intrusives and meta-sediments
Climate	Continental climate with wet summers, cold winters	Continental subarctic with cold winters	Medi-terranean	Continental subarctic	Temperate maritime	Temperate dry	Humid subtropical
Vegetation	Native tallgrass prairie with deciduous forest in valley bottom	White alder bushes, wet montane meadows and conifer forests	Old-growth Douglas-fir forest	Black spruce/feather moss slopes and treeless muskeg in valley bottom	Disturbed temperate rainforest. Native beech and tussock grass.	Mixed pasture grasses and native mixed beech and fern forest	Formerly farmed oak and hickory forest
Notes	Annual controlled burn	Formerly glaciated, spring fed	Spring fed, high rock uplift rates	Permafrost on north-facing slopes	Former coal mine	Cattle grazed	Partially developed, natural piping

*Note: The Caribou Creek gage station is operation during summer months only
†Source: USGS, Federal Emergency Management Agency 2006 LIDAR (2006), see: http://www.kansasgis.org/catalog/_cat_metadata.cfm?meta_id=223
‡Source: USGS, National Elevation Dataset (2009), see: https://lta.cr.usgs.gov/NED
§Source: ASTER Global Digital Elevation Map (2011), see: https://asterweb.jpl.nasa.gov/gdem.asp
‖Source: Land Information New Zealand New Zealand 8m Digital Elevation Model (2012), see: https://data.linz.govt.nz/
¶Source: North Carolina Floodplain Mapping Program 2007 LIDAR (2007), see: http://www.nconemap.com

Table A1.2 - Attributes of the six repeat stream width surveys in the Stony Subcatchment
Stream gage is located at the outlet of the surveyed subcatchment. The Stony Subcatchment has a drainage area of 48 ha and an elevation ranging from 163 to 210 m.

Date Surveyed	2015/10/27	2015/12/09	2016/02/02	2016/02/14	2016/03/04a	2016/03/04b
Surveyors	E Barefoot,	E Barefoot, E Henry	E Barefoot	E Barefoot	E Barefoot, A Tashie	E Barefoot, A Tashie
N Width Obs.	160	368	514	428	531	535
Mode Width (cm)	35.5	35.0	41.7	36.1	39.0	38.5
ADN Length (m)	800	1840	2570	2140	2655	2675
ADN Drainage Density (km^{-1})	1.65	3.80	5.31	4.42	5.49	5.53
1st Order Median Width (cm)	33.0	33.0	43.2	39.4	40.6	40.6
% Basin Stream Surface Area	0.11	0.25	0.42	0.28	0.42	0.43
Discharge (L/s)	7.5	14.5	15.3	8.8	13.9	14.3
Catchment-averaged Runoff (mm/day)	1.5	2.6	2.9	1.7	2.5	2.5
Flow Percentile (%)	53	73	78	57	72	72

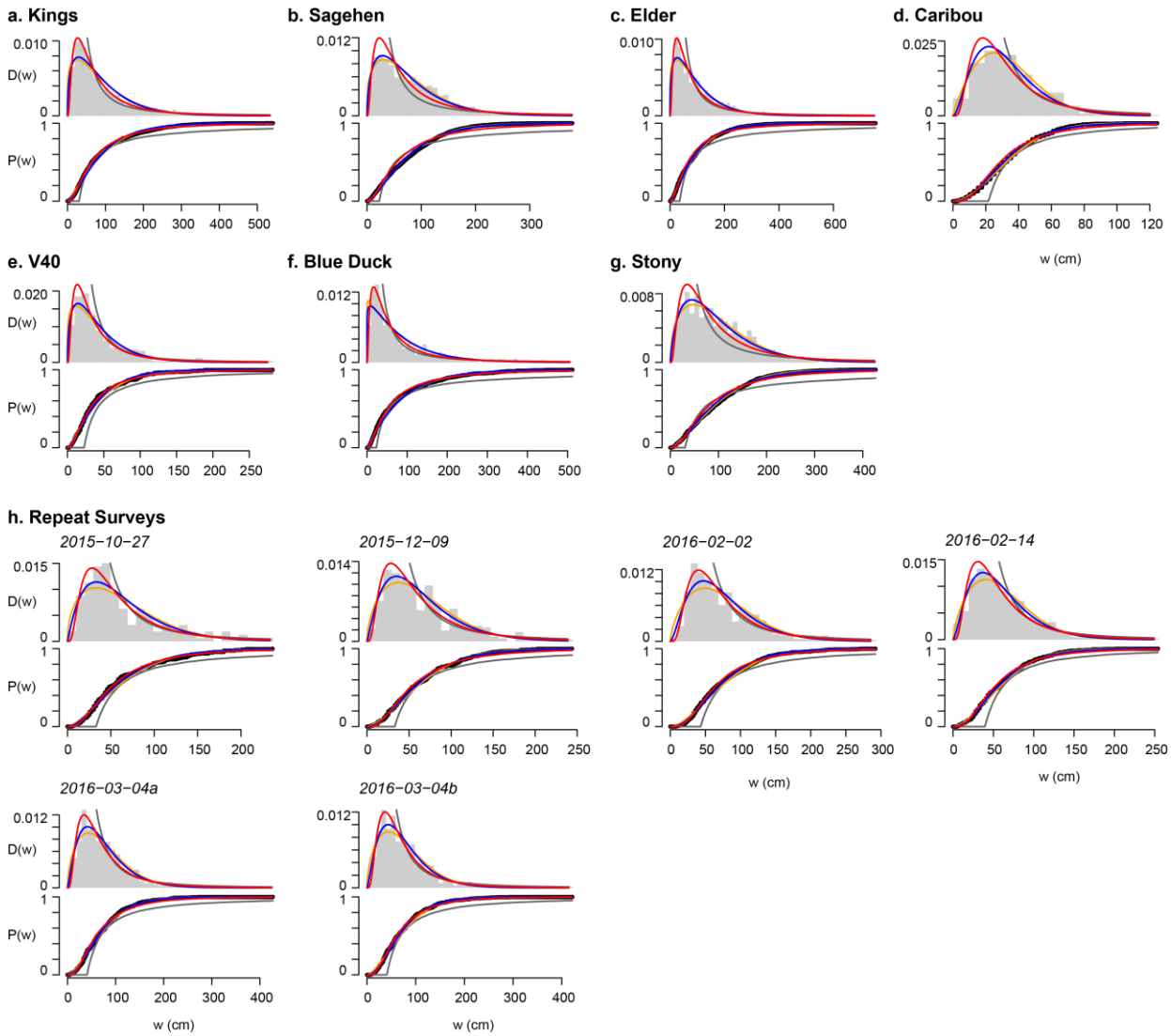

Figure A1.1 - Fitted distributions to stream width data

a-h, The upper panels show histograms with probability density functions and the lower panels show cumulative distribution functions. Unimodal distributions (lognormal, gamma and Weibull) describe stream width data better than the Pareto distribution in all study catchments. See Appendix Table A1.3 for statistics of fits. To improve goodness of fit, Pareto distributions were fit to data greater than the median first-order stream width after Allen and Pavelsky (after Allen and Pavelsky, 2015).

Table A1.3 - Statistics of distribution fits

Lognormal, gamma, Weibull, and Pareto distribution maximum likelihood estimated statistical parameters with Pearson's X^2 statistic and corresponding *p*-value, and with two sided one sample Kolmogorov-Smirnov (K-S) statistic (*D*) and corresponding *p*-value.

Statistical Parameter	Kings	Sagehen	Elder	Caribou	V40	Blue Duck	Stony	2015/ 10/27	2015/ 12/09	2016/ 02/02	2016/ 02/14	2016/ 03/04a	2016/ 03/04b
Lognormal Distribution													
location, μ	4.153	4.01	4.17	3.363	3.417	3.939	4.3	3.912	3.938	4.149	3.916	4.081	4.108
scale, σ	0.927	0.944	0.983	0.659	0.892	1.068	0.867	0.749	0.758	0.656	0.688	0.739	0.697
X^2	81.6	166.3	95.7	37.4	28.1	57.3	103.1	22.3	52	42.8	92.8	70.9	44.2
X^2 *p*	0.018	0.001	0.099	0.001	0.418	0.282	0.001	0.496	0.001	0.086	0.001	0.023	0.333
K-S *D*	0.028	0.081	0.061	0.08	0.047	0.039	0.083	0.041	0.063	0.044	0.061	0.061	0.05
K-S *p*	0.129	<0.001	0.001	0.019	0.628	0.397	<0.001	0.948	0.107	0.266	0.079	0.037	0.139
Gamma Distribution													
shape, *k*	1.463	1.573	1.382	2.888	1.512	1.109	1.809	2.074	2.186	2.639	2.573	2.264	2.378
rate, β	0.016	0.02	0.014	0.083	0.034	0.013	0.018	0.032	0.033	0.034	0.042	0.03	0.031
X^2	131	45.8	211.5	12.2	91.2	101.5	45.5	28.4	40.1	53.3	19.1	129.4	152.1
X^2 *p*	0.001	0.157	0.008	0.434	0.009	0.005	0.313	0.185	0.041	0.01	0.741	0.01	0.007
K-S *D*	0.062	0.041	0.041	0.043	0.077	0.071	0.035	0.078	0.035	0.05	0.034	0.038	0.046
K-S *p*	<0.001	0.017	0.065	0.513	0.101	0.011	0.29	0.278	0.764	0.148	0.703	0.435	0.202
Weibull Distribution													
shape, k	1.215	1.336	1.192	1.866	1.215	1.026	1.44	1.47	1.569	1.671	1.704	1.562	1.571
scale, λ	99.5	85.2	103.1	39.1	47.1	86.9	109.9	72.1	73.3	87.2	69.4	83.8	85.3
X^2	164.6	37.8	376.4	19.6	113.6	96.3	51.6	35	50.1	88.3	51.3	1287.2	1113.1
X^2 *p*	0.001	0.408	0.002	0.105	0.006	0.005	0.163	0.066	0.009	0.002	0.037	0.002	0.001
K-S *D*	0.064	0.034	0.038	0.038	0.083	0.06	0.031	0.089	0.053	0.061	0.049	0.041	0.052
K-S *p*	0	0.069	0.101	0.666	0.064	0.045	0.422	0.158	0.262	0.042	0.26	0.323	0.11
Pareto Distribution													
scale, x_m	30.48	22.60	35.56	21.59	22.86	23.368	30.48	33.02	33.02	43.18	39.37	40.64	40.64
shape, α	0.914	0.823	0.923	1.623	1.211	0.792	0.852	1.256	1.242	1.385	1.596	1.309	1.307
X^2	316	712	237.3	87.6	45.9	135.5	346.1	35.5	124.5	110.1	57.4	109.7	106.6
X^2 *p*	0.001	0.001	0.001	0.001	0.029	0.001	0.001	0.031	0.001	0.001	0.001	0.001	0.001
K-S *D*	0.156	0.209	0.183	0.19	0.106	0.162	0.212	0.151	0.173	0.162	0.176	0.184	0.171
K-S *p*	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table A1.4 - Carbon dioxide efflux calculation attributes

Values of CO_2 efflux are calculated using three metrics of stream surface area: 1) the observed surface area measured in the field; 2) the surface area calculated by applying Equation (1.10) on the observed stream lengths; and 3) the surface area derived from DEM flowline datasets. Numbers in parentheses represent 5th and 95th percentile ranges from Monte Carlo simulations (see Appendix Section A1.1).

Survey	Kings Creek (K1B Tributary)	Sagehen Creek Subcatch-ment	Upper Elder Creek	Caribou Creek (C1 Tributary)	Stony Creek Research Watershed	Stony Creek Subcatch-ment 27/10/2015	Stony Creek Subcatch-ment 09/12/2015	Stony Creek Subcatch-ment 02/02/2016	Stony Creek Subcatch-ment 14/02/2016	Stony Creek Subcatch-ment 04/03/2016a	Stony Creek Subcatch-ment 04/03/2016b
Observed Stream Length (km)	8.98	7.11	5.22	1.82	4.03	0.8	1.84	2.57	2.14	2.66	2.68
Runoff (mm/day)	2.47	0.14	0.46	0.11	0.50	1.34	2.59	2.73	1.57	2.48	2.55
Flowline Dataset	NHDPlus V2	NHDPlus V2	NHDPlus V2	EDNA	NHDPlus V2	NHDPlus V2	NHDPlus V2	NHDPlus V2	NHDPlus V2	NHDPlus V2	NHDPlus V2
HUC Region	HUC 10	HUC 18 Dry	HUC 17	AK HUC 4	HUC 3	HUC 3	HUC 3	HUC 3	HUC 3	HUC 3	HUC 3
Flowline Length (km)	1.9	2.68	3.19	0.878	0	0	0	0	0	0	0
Observed Surface Area (SA) (m^2)	8351	5568	5059	630	4011	519	1207	1991	1321	1992	2037
Eq. (1) Derived SA (m^2)	6609	5226	3840	1335	2963	589	1353	1892	1574	1955	1967
Flowline Derived SA (m^2)	610	8269	9340	940	0	0	0	0	0	0	0
CO_2 Flux from Observed SA (Mg-C/Yr)	26.2 (9.2-45.1)	65.6 (18.3-91.7)	49.6 (13.6-70.4)	2.1 (0.8-3.6)	9.7 (3.3-15.7)	1.3 (0.43-2.1)	3.2 (1.1-5.2)	5.2 (1.7-8.3)	3.33 (1.1-5.2)	5.24 (1.8-8.6)	5.2 (1.8-8.6)
CO_2 Flux from Eq. (1) SA (Mg-C/Yr)	21.4 (7.2-35.1)	62 (17.0-87.2)	37.7 (10.6-53.6)	4.6 (1.8-7.9)	7.0 (2.4-11.6)	1.4 (0.5-2.3)	3.6 (1.2-5.7)	4.9 (1.7-8.1)	3.77 (1.3-6.2)	5.17 (1.7-8.2)	5.3 (17-8.4)
CO_2 Flux from Flowline SA (Mg-C/Yr)	1.9 (0.7-3.2)	97.2 (26.7-137.2)	90.4 (25.6-131)	3.3 (1.2-5.5)	0	0	0	0	0	0	0

APPENDIX 2: DEVELOPMENT AND VALIDATION OF NARWIDTH AND GRWL

A2.1 Methods for developing NARWidth and GRWL

A2.1.1 Determining the time of year when rivers are at mean discharge

NARWidth and GRWL are composed of planform morphometric measurements of rivers at approximately mean discharge. While bankfull discharge usually represents the channel-forming flow (Wolman and Miller, 1960), mean discharge occurs much more often and is thus more consistently observable from limited remotely sensed imagery. Studies of fluvial hydrology suggest that geomorphic relationships relating width to discharge or drainage area are similar regardless of the discharge frequency used (Leopold and Maddock, 1953; Stewardson, 2005). Temporal fluctuations in discharge are a combined result of unpredictable events (e.g. storms and droughts) and more predictable seasonal variability in runoff. Unfortunately, no high-density, global-scale, daily discharge datasets exist to track the specific days that rivers are at their mean discharge. Instead, we approached the problem by determining the time of year that rivers in different parts of the globe are most likely to be at mean discharge.

To determine the optimal time of year to measure rivers, we used an international archive of long-term (>10 years) mean monthly discharge measurements (GRDC, 2017) composed of 1920 gauges in North America and 3693 gauges globally. For each station with a complete record (i.e. no missing data values), we constructed a mean monthly hydrograph and calculated the mean and standard deviation (1σ) for all monthly measurements (Appendix Figure A2.2). All months with discharges that fall within one standard deviation of the mean received a download score,

$$score(m) = Q_{\mu} - \tfrac{1}{2}\, Q_m - \tfrac{1}{4}\,(Q_{m-1} + Q_{m+1}), \qquad \text{(A2.1)}$$

where m is a given month being scored, Q_{μ} is the mean of all monthly discharges, Q_m is the mean monthly discharge of the given month, Q_{m-1} is the mean monthly discharge one month before month m, and Q_{m+1} is the mean monthly discharge of the month after month m. River discharge is more likely to be at or near the overall mean value during months with lower scores. Thus at each gauge station we produced a list of months that were ranked by the probability that the river was at mean discharge. To assign the monthly rankings from the gauge stations to each Landsat tile (landsat.usgs.gov), we considered both the proximity to gauge location and the monthly ranking. For example, the highest ranked month from the nearest gauge station has the greatest weight in setting the monthly preference order at a given Landsat tile. Each Landsat footprint is assigned at least one monthly preference and up to five ordered preferences (Appendix Figure A2.1b).

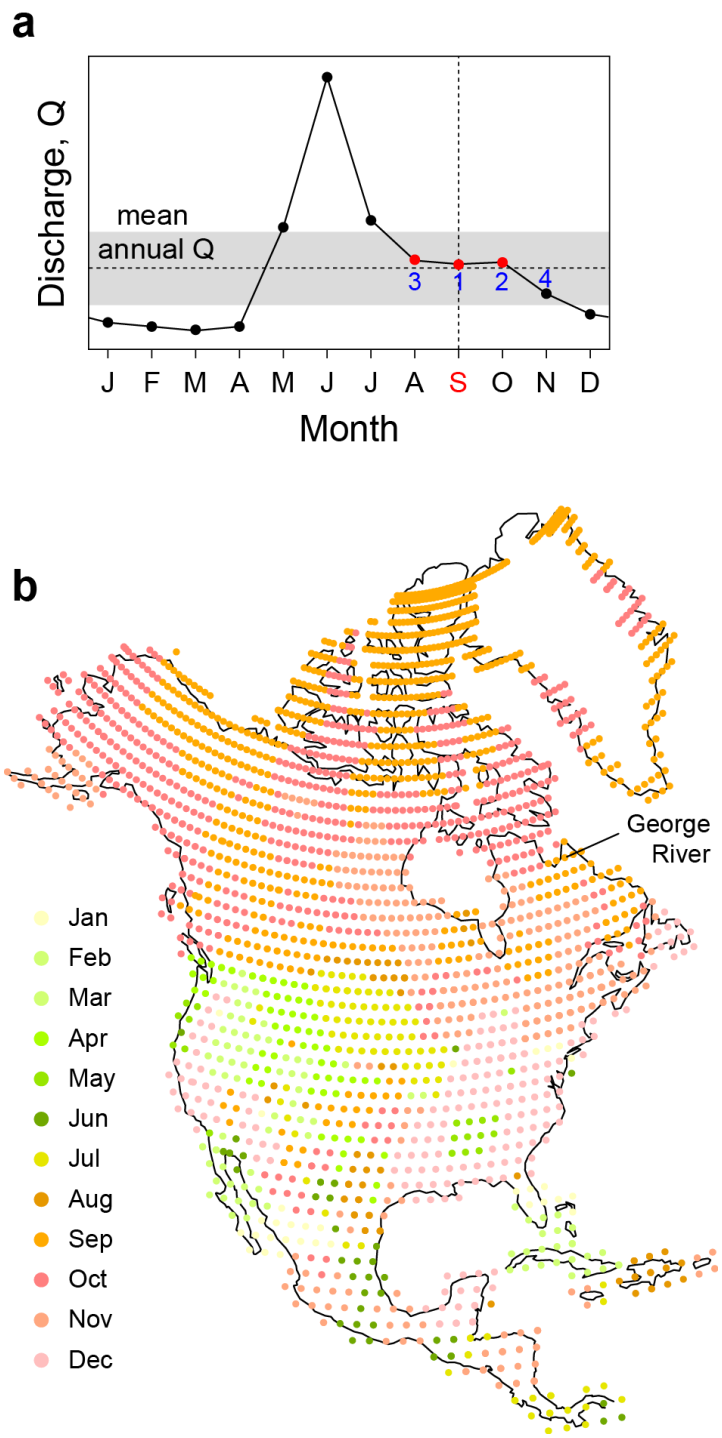

Figure A2.1 - Method for determining the time of year to analyze rivers

a, Example mean monthly hydrograph (George River, CN). Months with discharges within one standard deviation (gray box) of the mean discharge (dashed horizontal line) are ranked (blue numbers) based on their discharge and that of their two neighboring months (Equation A2.1). The best month to measure river width is September. **b,** Month that rivers are most likely to be at mean discharge for each Landsat tile.

To evaluate the validity of this method, we used in situ records to determine the variability of river width within each of the top ranked months. We restricted the analysis to United States Geological Survey (USGS) and the Water Survey of Canada (WSC) stream gauge records that were used to validate NARWidth/GRWL and that contained river width measurements during the top ranked month (N = 1,026). For each stream gauge, we calculated the standard deviation of (a) all recorded width measurements and (b) widths collected only during the top-ranked month. The median standard deviation of all width measurements was 30.0% of the mean annual river width while the standard deviation of width from the top ranked month was 17.3%. Thus, this method reduces the degree of variability associated with measuring rivers from satellite imagery by 42.3% relative to random sampling of rivers year round.

A2.1.2 Landsat imagery acquisition

Landsat TM, ETM+ (SLC-on), and OLI scenes were acquired from two data sources. We automatically downloaded 1,071 of from the Global Land Cover Facility (GLCF, glcf.umd.edu) over North America. The 6,261 remaining scenes were downloaded manually from the USGS website, (https://earthexplorer.usgs.gov/). The highest ranking scene was downloaded first. Upon download, each scene was visually inspected for flaws (e.g. clouds, river ice, shadows, flooding, no rivers) and either kept for use or discarded. If discarded, the next highest ranking scene was automatically downloaded. Four hundred and forty three Landsat tiles, located primarily in the tropics, had no cloud-free scenes available. To address this problem, we developed a program in IDL (version 8.0) that identified clouds based on their spectral signature and splices two or more complementary scenes together to eliminate clouds (Martinuzzi et al., 2007). Fourteen tiles located high in the Canadian Archipelago had no scenes free of cloud and ice available during

any of the monthly preferences listed. These scenes most likely have few if any wide rivers because they are located on relatively small and glacially dominated islands. Apart from these fourteen tiles, we successfully acquired imagery for all observable rivers on Earth.

A2.1.3 Image processing and GIS

We visually inspected several water classification methods including: the normalized difference water index (McFeeters, 1996), the modified normalized difference water index (Xu, 2006), the normalized difference vegetation index (Rouse, 1973), Monitoring the vernal advancement and retrogradation of natural vegetation), and the tasseled cap wetness index (Crist and Cicone, 1984). We found that the best performing classification method was the modified normalized difference water index,

$$MNDWI = (green-MIR)/(green+MIR), \tag{A2.2}$$

where MIR is the middle infrared band (e.g. TM Band 5) and green is the green band (i.e. TM Band 2) (Xu, 2006). We applied the MNDWI formula to all Landsat scenes, mosaicked, and clipped images to 4° Latitude by 6° Longitude tiles. We then created a binary water mask by applying a dynamic threshold (Li and Sheng, 2012) which was visually inspected and corrected for any gaps in continuity and classification errors. These errors stem from sources including river view obstruction by topographic shadows, bridges, or dams, or the erroneous inclusion of swamps, large lakes, or deltas in the river network. RivWidth (version 0.4) calculated a channel centerline for all river reaches longer than 10 km (Appendix Figure A2.2). After RivWidth runs on a mosaic image, we visually inspected the RivWidth output for errors.

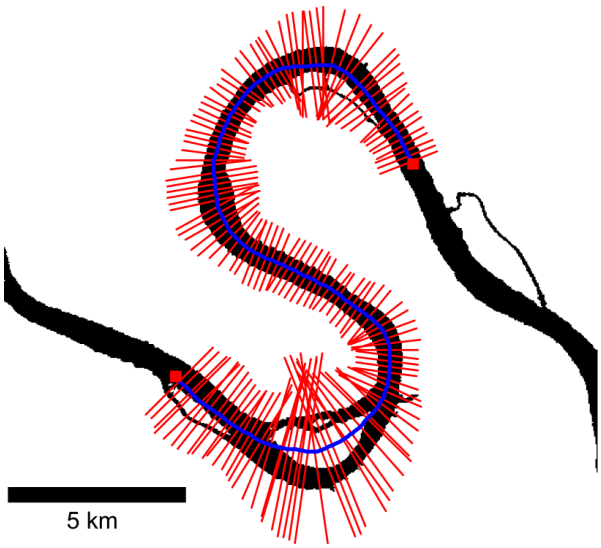

Figure A2.2 - The RivWidth program calculates a river centerline (blue) from a binary river mask (black) derived from Landsat imagery (modified from Miller et al., 2014)
At each centerline pixel, RivWidth computes the river width and braiding index. A river length was computed at each width measurement by calculating the Euclidean distance between each centerline pixel and the next adjacent centerline pixel.

Reservoirs and lakes connected to the fluvial network were labeled using GIS methods and several water body datasets: 1) the Global Lakes and Wet Lands Database (Lehner and Döll, 2004); 2) the Global Reservoir and Dam Database (Lehner et al., 2011); 3) the U.S. and Canada Water Polygons dataset (TomTom North America, 2012); and 4) the Mexico Water Bodies dataset (Sistemas de Información Geográfica, 2008). The locations of lakes and reservoirs were then visually inspected and corrected in ArcGIS.

A2.2 Database validation Image processing and GIS

We validated Landsat-derived river width measurements using 1,049 stream flow and river width records from the USGS (http://nwis.waterdata.usgs.gov/nwis) and the WSC (http://www.ec.gc.ca/rhc-wsc/). At each gauge location, we estimated the river width at mean annual discharge and compared this value to the average of the five spatially closest RivWidth

measurements (Appendix Figure A2.3). We excluded river width measurements that: 1) were taken more than 200 m upstream or downstream from the gauge station; 2) were taken when river ice was observed; or 3) were labelled as "Poor" measurements. We then took the mean of all width measurements that were taken when river discharge was within 10% of the mean annual discharge (red dots, Appendix Figure A2.3) and compared mean in situ width with the mean width of the five nearest NARWidth/GRWL river width measurements. The residuals between in situ and remotely sensed width measurements show the effect of Landsat's 30 m resolution on the validation analysis (Appendix Figure A2.4). The residuals show heteroscedasticity, are uncorrelated, and are unbiased with changing in situ width.

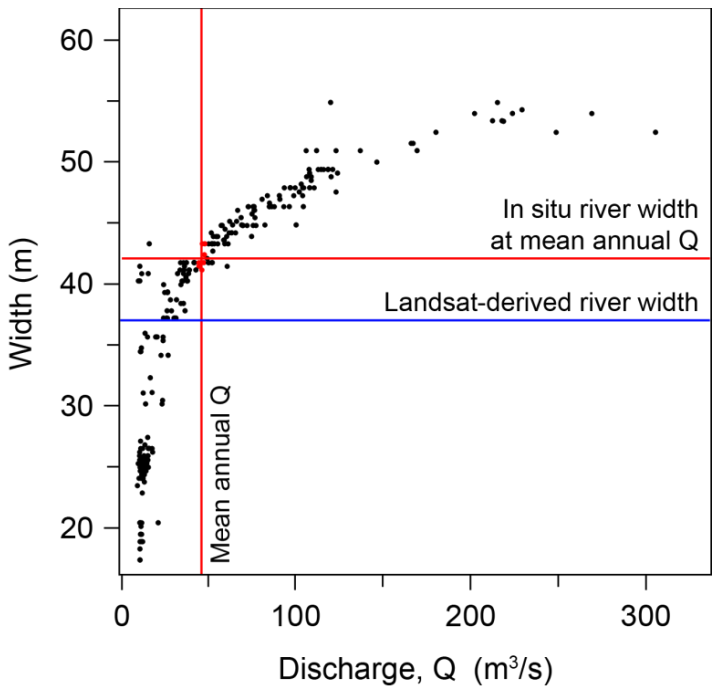

Figure A2.3 - Example in situ river discharge-width rating curve used to validate NARWidth and GRWL

Mean annual discharge was calculated using daily discharge over at least a 10 year period (vertical line). The corresponding river width (red line) was then compared to the mean of the five nearest Landsat-derived NARWidth and GRWL measurements at that location (blue line).

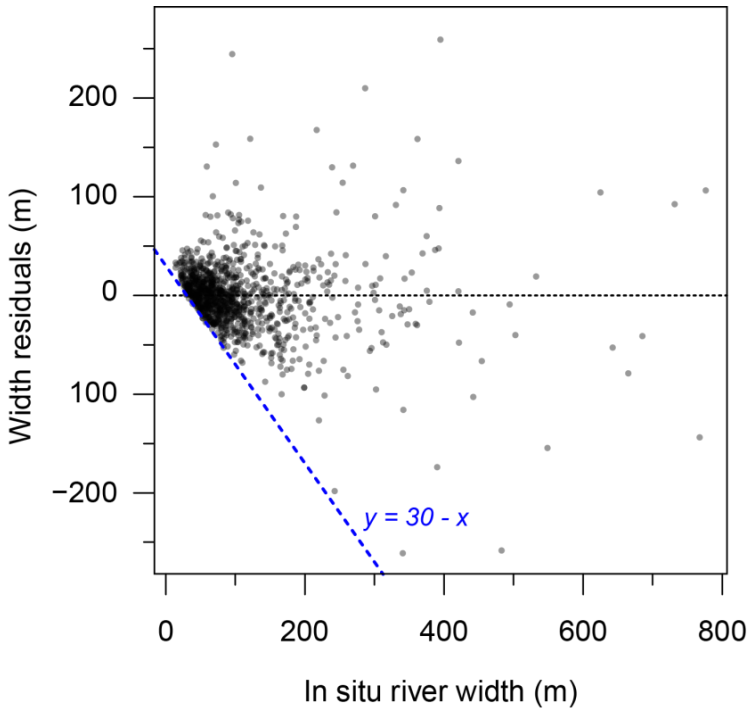

Figure A2.4 -- In situ and remotely sensed river width residuals
Landsat's spatial resolution of 30 m excludes any residuals below the dashed blue line, demonstrating the influence of Landsat's resolution on the regression analysis of the full range of width data (blue line in Figure 2.2).

A2.3 NARWidth river and stream surface area calculation

River surface area was computed at each centerline pixel by multiplying river width and length. To characterize the relationship between river surface area and width, we binned surface area by width using 100 m width intervals (Figure 2.4). We then multiplied the binned surface area data by the bin interval (100 m) and performed a least squares linear regression of the data in log space. We calculated the total surface area of rivers and streams with widths narrower than 100 m by taking the definite integral of this curve from 1.6±1.1 m (estimated by Downing et al., 2012) to 100 m. Finally, we summed this extrapolated surface area and all measured surface areas in rivers wider than 100 m to estimate the total surface area of all rivers and streams in North America.

A2.4 GRWL river and stream surface area calculation

To calculate the surface area of rivers and streams from GRWL, we used a similar strategy as we did with NARWidth, but using a more rigorous statistical method. The statistical extrapolations of river surface area are based on principles of fractal river network scaling theory within basins (Allen et al., in review; Horton, 1945). We isolate individual river networks by clipping GRWL to each highest order basin in the HydroBASINs global hydrography database (Lehner and Grill, 2013). Observed river area is calculated by multiplying GRWL widths by the downstream distance between GRWL measurements (Figure 3.3a). In basins with >5000 river width measurements greater than 100 m, we apply a Pareto distribution using maximum likelihood estimation (MLE) in R. We determine the variance of the Pareto shape parameter, α, using MLE (dashed diagonal lines in Figure 3.3b). We then extend this statistical fit to narrower rivers and streams, down to a surface area of 11.6±2.9 m^2, corresponding with the mean of the medians of first order stream width, 0.32±0.08 m. We combine the river area error associated with fitting α and the error of the extrapolation minimum by assuming that both populations are normally distributed and multiplying the probability densities.

To estimate the surface area in small and/or dry basins, we develop a relationship between basin percent river area (%SA), basin area (BA) (Lehner and Grill, 2013), and aridity index (AI) (Zomer et al., 2008). We use a basin area weighted multiple linear regression of log-transformed data weighted by basin area (Figure 3.5). We weight basin area because larger basins have higher quality river width data. We use Monte-Carlo error propagation (N=10,000) to estimate uncertainty associated with the Pareto extrapolation and the multiple linear regression shown in Figure 3.5.

REFERENCES

Allen, G.H., Barnes, J.B., Pavelsky, T.M., and Kirby, E., 2013, Lithologic and tectonic controls on bedrock channel form at the northwest Himalayan front: Journal of Geophysical Research: Earth Surface, v. 118, p. 1806-1825.

Allen, G.H., and Pavelsky, T.M., 2015, Patterns of river width and surface area revealed by the satellite-derived North American River Width data set: Geophysical Research Letters, v. 42, p. 395-402.

Allen, G.H., Pavelsky, T.M., Barefoot, E.A., Lamb, M.P., Butman, D., Tashie, A., and Gleason, C.J., in review, Similarity of stream width distributions across headwater systems: Nature.

Anderson, R.J., Bledsoe, B.P., and Hession, W.C., 2004, Width of stream and rivers in response to vegetation, bank material, and other factors: JAWRA Journal of the American Water Resources Association, v. 40, p. 1159-1172.

Andreadis, K.M., Schumann, G.J.P., and Pavelsky, T.M., 2013, A simple global river bankfull width and depth database: Water Resources Research.

Bastviken, D., Tranvik, L.J., Downing, J.A., Crill, P.M., and Enrich-Prast, A., 2011, Freshwater Methane Emissions Offset the Continental Carbon Sink: Science, v. 331, p. 50.

Beaulieu, J.J., Tank, J.L., Hamilton, S.K., Wollheim, W.M., Hall, R.O., Mulholland, P.J., Peterson, B.J., Ashkenas, L.R., Cooper, L.W., Dahm, C.N., Dodds, W.K., Grimm, N.B., Johnson, S.L., McDowell, W.H., Poole, G.C., Valett, H.M., Arango, C.P., Bernot, M.J., Burgin, A.J., Crenshaw, C.L., Helton, A.M., Johnson, L.T., O'Brien, J.M., Potter, J.D., Sheibley, R.W., Sobota, D.J., and Thomas, S.M., 2011, Nitrous oxide emission from denitrification in stream and river networks: Proceedings of the National Academy of Sciences, v. 108, p. 214-219.

Beighley, R.E., and Gummadi, V., 2011, Developing channel and floodplain dimensions with limited data: a case study in the Amazon Basin: Earth Surface Processes and Landforms, v. 36, p. 1059-1071.

Bencala, K.E., Gooseff, M.N., and Kimball, B.A., 2011, Rethinking hyporheic flow and transient storage to advance understanding of stream-catchment connections: Water Resources Research, v. 47.

Benda, L., Hassan, M.A., Church, M., and May, C.L., 2005, Geomorphology of steepland headwaters: the transition from hillslopes to channels: JAWRA Journal of the American Water Resources Association, v. 41, p. 835-851.

Benstead, J.P., and Leigh, D.S., 2012, An expanded role for river networks: Nature Geosci, v. 5, p. 678-679.

Biancamaria, S., Lettenmaier, D.P., and Pavelsky, T.M., 2016, The SWOT Mission and Its Capabilities for Land Hydrology: Surveys in Geophysics, v. 37, p. 307-337.

Butman, D., and Raymond, P.A., 2011, Significant efflux of carbon dioxide from streams and rivers in the United States: Nature Geoscience, v. 4, p. 839-842.

Butman, D., Stackpoole, S., Stets, E., McDonald, C.P., Clow, D.W., and Striegl, R.G., 2016, Aquatic carbon cycling in the conterminous United States and implications for terrestrial carbon accounting: Proceedings of the National Academy of Sciences, v. 113, p. 58-63.

Chow, V.t., 1959, Open - channel hydraulics: Auckland, McGraw-Hill.

Clark, M.P., Fan, Y., Lawrence, D.M., Adam, J.C., Bolster, D., Gochis, D.J., Hooper, R.P., Kumar, M., Leung, L.R., Mackay, D.S., Maxwell, R.M., Shen, C., Swenson, S.C., and Zeng, X., 2015, Improving the representation of hydrologic processes in Earth System Models: Water Resources Research, v. 51, p. 5929-5956.

Clauset, A., Shalizi, C., and Newman, M., 2009, Power-Law Distributions in Empirical Data: SIAM Review, v. 51, p. 661-703.

Coe, M.T., Costa, M.H., and Howard, E.A., 2008, Simulating the surface waters of the Amazon River basin: impacts of new river geomorphic and flow parameterizations: Hydrological Processes, v. 22, p. 2542-2553.

Crist, E.P., and Cicone, R.C., 1984, A Physically-Based Transformation of Thematic Mapper Data---The TM Tasseled Cap: Geoscience and Remote Sensing, IEEE Transactions on, v. GE-22, p. 256-263.

Davidson, E.A., Figueiredo, R.O., Markewitz, D., and Aufdenkampe, A.K., 2010, Dissolved CO2 in small catchment streams of eastern Amazonia: A minor pathway of terrestrial carbon loss: Journal of Geophysical Research: Biogeosciences, v. 115, p. G04005.

Dingman, L.S., 2007, Analytical derivation of at-a-station hydraulic–geometry relations: Journal of Hydrology, v. 334, p. 17-27.

Downing, J.A., Cole, J.J., Duarte, C.M., Middelburg, J.J., Melack, J.M., Prairie, Y.T., Kortelainen, P., Striegl, R.G., McDowell, W.H., and Tranvik, L.J., 2012, Global abundance and size distribution of streams and rivers: Inland Waters, v. 2, p. 229 - 236.

Ferguson, R.I., 1986, Hydraulics and hydraulic geometry: Progress in Physical Geography, v. 10, p. 1-31.

Findlay, S., 1995, Importance of surface-subsurface exchange in stream ecosystems: The hyporheic zone: Limnology and Oceanography, v. 40, p. 159-164.

Fuller, C.C., and Harvey, J.W., 2000, Reactive Uptake of Trace Metals in the Hyporheic Zone of a Mining-Contaminated Stream, Pinal Creek, Arizona: Environmental Science & Technology, v. 34, p. 1150-1155.

Gillespie, C.S., 2014, Fitting heavy tailed distributions: the poweRlaw package. *R package version* 0.20.5: Newcastle, UK.

Gleason, C.J., and 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.

Godsey, S.E., and Kirchner, J.W., 2014, Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order: Hydrological Processes, v. 28, p. 5791-5803.

Golden, L.A., and Springer, G.S., 2006, Channel geometry, median grain size, and stream power in small mountain streams: Geomorphology, v. 78, p. 64-76.

Gomez-Velez, J.D., and Harvey, J.W., 2014, A hydrogeomorphic river network model predicts where and why hyporheic exchange is important in large basins: Geophysical Research Letters, v. 41, p. 2014GL061099.

Gomi, T., Sidle, R.C., and Richardson, J.S., 2002, Understanding processes and downstream linkages of headwater systems: Bioscience, v. 52, p. 905-916.

GRDC, 2017, Long-Term Mean Monthly Discharges and Annual Characteristics of GRDC Station, *in* Centre, G.R.D., ed.: Koblenz, Germany, Federal Institute of Hydrology (BfG).

Gregory, K.J., and Walling, D.E., 1968, The variation of drainage density within a catchment: International Association of Scientific Hydrology. Bulletin, v. 13, p. 61-68.

Hack, J.T., 1957, Studies of longitudinal stream profiles in Virginia and Maryland: US Geol. Surv. Prof. Pap., p. 45-97.

Hankin, D.G., and Reeves, G.H., 1988, Estimating Total Fish Abundance and Total Habitat Area in Small Streams Based on Visual Estimation Methods: Canadian Journal of Fisheries and Aquatic Sciences, v. 45, p. 834-844.

Harvey, J.W., and Bencala, K.E., 1993, The Effect of streambed topography on surface-subsurface water exchange in mountain catchments: Water Resources Research, v. 29, p. 89-98.

Hess, L.L., Melack, J.M., Novo, E.M.L.M., Barbosa, C.C.F., and Gastil, M., 2003, Dual-season mapping of wetland inundation and vegetation for the central Amazon basin: Remote Sensing of Environment, v. 87, p. 404-428.

Horton, R.E., 1945, Erosional Development of streams and their drainage basins; hydrophysical approach to quantitative morphology: Geological Society of America Bulletin, v. 56, p. 275-370.

Hotchkiss, E.R., Hall Jr, R.O., Sponseller, R.A., Butman, D., Klaminder, J., Laudon, H., Rosvall, M., and Karlsson, J., 2015, Sources of and processes controlling CO2 emissions change with the size of streams and rivers: Nature Geosci, v. 8, p. 696-699.

Hough, P., 2014, Understanding global security, Routledge.

Kasahara, T., and Wondzell, S.M., 2003, Geomorphic controls on hyporheic exchange flow in mountain streams: Water Resources Research, v. 39, p. SBH 3-1-SBH 3-14.

Kiel, B.A., and Cardenas, M.B., 2014, Lateral hyporheic exchange throughout the Mississippi River network: Nature Geosci, v. 7, p. 413-417.

Kimbrough, R.A., Smith, R.R., Ruppert, G.P., and Wiggins, W.D., 2003, Explanation of Water-Quality Records *in* USGS, ed., Water Data Report WA-03-1: Vancouver, WA, United State Geological Survey, p. 647.

Kirchner, J.W., 1993, Statistical inevitability of Horton's laws and the apparent randomness of stream channel networks: Geology, v. 21, p. 591-594.

Lague, D., 2014, The stream power river incision model: evidence, theory and beyond: Earth Surface Processes and Landforms, v. 39, p. 38-61.

Lehner, B., and Döll, P., 2004, Development and validation of a global database of lakes, reservoirs and wetlands: Journal of Hydrology, v. 296, p. 1-22.

Lehner, B., and Grill, G., 2013, Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems: Hydrological Processes, v. 27, p. 2171-2186.

Lehner, B., Reidy Liermann, C., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., Doll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rodel, R., Sindorf, N., and Wisser, D., 2011, Global Reservoir and Dam Database, Version 1 (GRanDv1): Reservoirs, Revision 01, Global Water System Project (GWSP).

Lehner, B., Verdin, K., and Jarvis, A., 2008, New Global Hydrography Derived From Spaceborne Elevation Data: Eos Trans. AGU, v. 89.

Leopold, L.B., 1962, Rivers: American Scientist, v. 50, p. 511-537.

Leopold, L.B., and Maddock, T., Jr., 1953, The hydraulic geometry of stream channels and physiographic implications USGS Prof. Paper, p. 57.

Li, J., and Sheng, Y., 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, v. 33, p. 5194-5213.

Manning, R., 1891, On the flow of water in open channels and pipes: Transactions of the Institution of Civil Engineers of Ireland, v. 20, p. 161-207.

Martinuzzi, S.n., Gould, W.A., and Ramos González, O.M., 2007, Creating cloud-free Landsat ETM+ data sets in tropical landscapes cloud and cloud-shadow removal: Rio Piedras, PR, U.S. Dept. of Agriculture, Forest Service, International Institute of Tropical Forestry.

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, v. 17, p. 1425-1432.

Meyer, J.L., Strayer, D.L., Wallace, J.B., Eggert, S.L., Helfman, G.S., and Leonard, N.E., 2007, The Contribution of Headwater Streams to Biodiversity in River Networks1: JAWRA Journal of the American Water Resources Association, v. 43, p. 86-103.

Miller, Z.F., Pavelsky, T.M., and Allen, G.H., 2014, Quantifying river form variations in the Mississippi Basin using remotely sensed imagery: Hydrology and Earth System Sciences, v. 18, p. 4883-4895.

Moody, J.A., and Troutman, B.M., 2002, Characterization of the spatial variability of channel morphology: Earth Surface Processes and Landforms, v. 27, p. 1251-1266.

Morisawa, M.E., 1962, Quantitative Geomorphology of Some Watersheds in the Appalachian Plateau: Geological Society of America Bulletin, v. 73, p. 1025-1046.

Neal, J., Schumann, G., and Bates, P., 2012, A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas: Water Resources Research, v. 48, p. W11506.

Neal, J.C., Odoni, N.A., Trigg, M.A., Freer, J.E., Garcia-Pintado, J., Mason, D.C., Wood, M., and Bates, P.D., 2015, Efficient incorporation of channel cross-section geometry uncertainty into regional and global scale flood inundation models: Journal of Hydrology, v. 529, Part 1, p. 169-183.

O'Callaghan, J.F., and Mark, D.M., 1984, The extraction of drainage networks from digital elevation data: Computer Vision, Graphics, and Image Processing, v. 28, p. 323-344.

O'Loughlin, F., Trigg, M.A., Schumann, G.J.P., and Bates, P.D., 2013, Hydraulic characterization of the middle reach of the Congo River: Water Resources Research, v. 49, p. 5059-5070.

Park, C.C., 1977, World-wide variations in hydraulic geometry exponents of stream channels: An analysis and some observations: Journal of Hydrology, v. 33, p. 133-146.

Parker, G., 1991, Selective Sorting and Abrasion of River Gravel. II: Applications: Journal of Hydraulic Engineering, v. 117.

Parker, G., Wilcock, P.R., Paola, C., Dietrich, W.E., and Pitlick, J., 2007, Physical basis for quasi-universal relations describing bankfull hydraulic geometry of single-thread gravel bed rivers: Journal of Geophysical Research: Earth Surface, v. 112, p. n/a-n/a.

Pavelsky, T.M., Allen, G.H., and Miller, Z.F., 2014a, Spatial Patterns of River Width in the Yukon River Basin, Remote Sensing of the Terrestrial Water Cycle: Hoboken, NJ., John Wiley & Sons, Inc, p. 131-141.

Pavelsky, T.M., Durand, M.T., Andreadis, K.M., Beighley, R.E., Paiva, R.C.D., Allen, G.H., and Miller, Z.F., 2014b, Assessing the potential global extent of SWOT river discharge observations: Journal of Hydrology, v. 519, Part B, p. 1516-1525.

Pavelsky, T.M., Durand, M.T., Andreadis, K.M., Beighley, R.E., Paiva, R.C.D., Allen, G.H., and Miller, Z.F., 2014c, Assessing the potential global extent of SWOT river discharge observations: Journal of Hydrology, v. 519, Part B, p. 1516-1525.

Pavelsky, T.M., and Smith, L.C., 2008, RivWidth: A Software Tool for the Calculation of River Widths From Remotely Sensed Imagery: Geoscience and Remote Sensing Letters, IEEE, v. 5, p. 70-73.

Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A.S., 2016, High-resolution mapping of global surface water and its long-term changes: Nature, v. 540, p. 418-422.

Peters, G.P., Andrew, R.M., Boden, T., Canadell, J.G., Ciais, P., Le Quere, C., Marland, G., Raupach, M.R., and Wilson, C., 2013, The challenge to keep global warming below 2 [deg]C: Nature Clim. Change, v. 3, p. 4-6.

Petersen, R.C., 1992, The RCE: a Riparian, Channel, and Environmental Inventory for small streams in the agricultural landscape: Freshwater Biology, v. 27, p. 295-306.

Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Durr, H., Meybeck, M., Ciais,

P., and Guth, P., 2013, Global carbon dioxide emissions from inland waters: Nature, v. 503, p. 355-359.

Richey, J.E., Melack, J.M., Aufdenkampe, A.K., Ballester, V.M., and Hess, L.L., 2002, Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric CO2: Nature, v. 416, p. 617-620.

Rodríguez-Iturbe, I., Ijjász-Vásquez, E.J., Bras, R.L., and Tarboton, D.G., 1992, Power law distributions of discharge mass and energy in river basins: Water Resources Research, v. 28, p. 1089-1093.

Rodríguez-Iturbe, I., and Rinaldo, A., 2001, Fractal river basins: chance and self-organization: Cambridge Cambridge University Press.

Rosgen, D.L., 1994, A classification of natural rivers: CATENA, v. 22, p. 169-199.

Rouse, J.W., Jr. , 1973, Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation, *in* NASA, ed.: College Station, TX, United States.

Schiller, D., Acuña, V., Graeber, D., Martí, E., Ribot, M., Sabater, S., Timoner, X., and Tockner, K., 2011, Contraction, fragmentation and expansion dynamics determine nutrient availability in a Mediterranean forest stream: Aquatic Sciences, v. 73, p. 485-497.

Schumm, S.A., 1956, Evolution of drainage systems and slopes in badlands at Perth Amboy: Geological Society of America Bulletin, v. 67, p. 597-646.

Sen, P.K., 1968, Estimates of the Regression Coefficient Based on Kendall's Tau: Journal of the American Statistical Association, v. 63, p. 1379-1389.

Shaw, S.B., 2016, Investigating the linkage between streamflow recession rates and channel network contraction in a mesoscale catchment in New York state: Hydrological Processes, v. 30, p. 479-492.

Sistemas de Información Geográfica, S.A., 2008, Mexico Water Bodies [electronic resource]: Redlands, California, USA, Esri.

Smith, L.C., Isacks, B.L., Bloom, A.L., and Murray, A.B., 1996, Estimation of Discharge From Three Braided Rivers Using Synthetic Aperture Radar Satellite Imagery: Potential Application to Ungaged Basins: Water Resources Research, v. 32, p. 2021-2034.

Spearman, C., 1904, The Proof and Measurement of Association between Two Things: The American Journal of Psychology, v. 15, p. 72-101.

Stewardson, M., 2005, Hydraulic geometry of stream reaches: Journal of Hydrology, v. 306, p. 97-111.

Strahler, A.N., 1957, Quantitative analysis of watershed geomorphology: Eos, Transactions American Geophysical Union, v. 38, p. 913-920.

TomTom North America, I., 2012, U.S. and Canada Water Polygons [electronic resource]: Redlands, California, USA, Esri.

Trampush, S.M., Huzurbazar, S., and McElroy, B., 2014, Empirical assessment of theory for bankfull characteristics of alluvial channels: Water Resources Research, v. 50, p. 9211-9220.

Triska, F.J., Kennedy, V.C., Avanzino, R.J., Zellweger, G.W., and Bencala, K.E., 1989, Retention and Transport of Nutrients in a Third-Order Stream in Northwestern California: Hyporheic Processes: Ecology, v. 70, p. 1893-1905.

USGS, U.S.G.S., 2014, HYDRO1K documentation, *in* Survey, U.S.G., ed., Volume 2014.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., and Cushing, C.E., 1980, The river continuum concept: Canadian Journal of Fisheries and Aquatic Sciences, v. 37, p. 130-137.

Venables, W.N., and Ripley, B.D., 2002, Modern Applied Statistics with S: New York, Springer-Verlag

Verdin, K.L., and Verdin, J.P., 1999, A topological system for delineation and codification of the Earth's river basins: Journal of Hydrology, v. 218, p. 1-12.

Vorosmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., and Davies, P.M., 2010, Global threats to human water security and river biodiversity: Nature, v. 467, p. 555-561.

Wehrli, B., 2013, Biogeochemistry: Conduits of the carbon cycle: Nature, v. 503, p. 346-347.

Welcomme, R.L., 1976, Some general and theoretical considerations on the fish yield of African rivers: Journal of Fish Biology, v. 8, p. 351-364.

White, E.P., Enquist, B.J., and Green, J.L., 2008, On Estimating the exponent of power-law frequency distributions: Ecology, v. 89, p. 905-912.

Wilkerson, G.V., Kandel, D.R., Perg, L.A., Dietrich, W.E., Wilcock, P.R., and Whiles, M.R., 2014, Continental-scale relationship between bankfull width and drainage area for single-thread alluvial channels: Water Resources Research, v. 50, p. 919-936.

Williams, G.P., 1978, Bank-full discharge of rivers: Water Resources Research, v. 14, p. 1141-1154.

Wohl, E., 2004, Limits of downstream hydraulic geometry: Geology, v. 32, p. 897-900.

Wohl, E., and Merritt, D.M., 2008, Reach-scale channel geometry of mountain streams: Geomorphology, v. 93, p. 168-185.

Wohl, E.E., and Wilcox, A., 2005, Channel geometry of mountain streams in New Zealand: Journal of Hydrology, v. 300, p. 252-266.

Wolman, M.G., and Miller, J.P., 1960, Magnitude and frequency of forces in geomorphic processes: J. Geol., v. 68, p. 54-74.

Wondzell, S.M., 2011, The role of the hyporheic zone across stream networks: Hydrological Processes, v. 25, p. 3525-3532.

Xu, H., 2006, Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery: International Journal of Remote Sensing, v. 27, p. 3025-3033.

Yamazaki, D., O'Loughlin, F., Trigg, M.A., Miller, Z.F., Pavelsky, T.M., and Bates, P.D., 2014, Development of the Global Width Database for Large Rivers: Water Resources Research.

Zimmerman, R.C., Goodlett, J.C., and Comer, G.H., 1967, The influence of vegetation on channel form of small streams, Proceedings of the Symposium on River Morphology, Volume 75: Wallinford, U. K., International Association of Scientific Hydrology, p. 255–275.

Zomer, R.J., Trabucco, A., Bossio, D.A., and Verchot, L.V., 2008, Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation: Agriculture, Ecosystems & Environment, v. 126, p. 67-80.