How Have Global River Widths Changed Over Time?

Dongmei Feng, Colin J. Gleason, Xiao Yang, George H. Allen, and Tamlin M. Pavelsky

Abstract Changes in a river's width reflect natural and anthropogenic impacts on local and upstream/downstream hydraulic and hydrologic processes. Temporal variation of river width also impacts biogeochemical exchange and reflects geomorphologic evolution. However, while global maps of mean river width and dynamic water surface extent exist, there is currently no standardized global assessment of river widths that documents changes over time. Therefore, we made repeated width measurements from Landsat images for all rivers wider than 90 m collected from 1984 to 2020 (named Global LOng-term river Width, GLOW), which consists of ∼1.2 billion cross-sectional river width measurements, with an average of 3,000 width measurements per 10-km reach. With GLOW, we investigated the temporal variations of global river width, quantified by the interquartile range (IQR) and temporal trend. We found that 85% of global rivers have a width IQR <150 m. We also found that 37% of global river segments show significant temporal trends in width over the past 37 years, and this number is higher (46%) for human-regulated rivers. Further, we leveraged machine learning to identify the most important factors explaining river width variations and revealed that these driving factors are significantly different between free-flowing and non-free-flowing rivers. Specifically, the most important factor driving temporal variations in river width is the climate for free-flowing rivers, and is soil condition for human-impacted rivers. Finally, we anticipate that this study and the public release of GLOW will improve the understanding of river dynamics and catalyze additional interdisciplinary studies.

1. Introduction

Rivers impact humans and the environment by providing water/energy supply, aquatic habitats, transportation corridors, and geohydrological/biochemical connections with the atmosphere, land, and ocean (Feng et al., 2020; Gernaat et al., 2017; Oki & Kanae, 2006; Raymond et al., 2013; Svytski et al., 2005). River width, especially temporal changes in river width, is an essential measure of the dynamics of rivers’ functionality and reflects local and upstream/downstream impacts of hydrological processes and human interventions (Allen & Pavelsky, 2015, 2018; Frasson et al., 2019; Gao et al., 2021). In geomorphology, river width dynamics shed light on geomorphological evolution and river types (e.g., Bertoldi et al., 2010; Brinkerhoff et al., 2020; Dunne & Jerolmack, 2020; Krapesch et al., 2011; Monegaglia et al., 2018). In bioecology, river width variations indicate changes in the water surface area, which partly controls the mass and energy exchange (e.g., carbon emission and heat transfer) at the water-atmosphere interface (e.g., Allen & Pavelsky, 2018; Caissie, 2006; Hotchkiss et al., 2015; Raymond et al., 2013), and other hydraulic variables (e.g., river depth), which reflect habitat conditions in the fluvial ecosystem (Feng et al., 2020; Miller et al., 2014). Knowledge of river width changes is also essential for water resource management, as it provides reference information for flood risk management and human regulations (Chakraborty & Mukhopadhyay, 2015; Krapesch et al., 2011). In hydrology, multitemporal river widths are commonly used as a proxy for river discharge based on a power law usually existing between these two variables (Leopold & Maddock, 1953). This practice serves as the basis for much recent literature on remote sensing of river discharge and hydrologic/hydraulic modeling (e.g., Brinkerhoff et al., 2020; Elmi et al., 2015; Feng et al., 2019; Gleason & Smith, 2014; Hagemann et al., 2017; Harlan et al., 2021; Ishitsuka et al., 2021; Lin et al., 2020; Mengen et al., 2020; Pavelsky, 2014). Therefore, understanding how river width changes over time is essential to both scientific and engineering applications.

While many records of river width dynamics exist, there is little knowledge of the temporal variation of river width globally. River width measurements can be readily obtained from in situ measurements, but are difficult to directly measure automatically: a time-lapse photography setup with good photogrammetric control is
often needed (Leduc et al., 2018). Therefore, the most publicly available in situ data of river width are archived field measurements, and most of these are from North America and cover only a short period of time. Widths are also readily measured from satellite remote sensing data, which provides powerful resources for extracting multitemporal river width at the global scale, and numerous “off-the-shelf” software has been developed to aid in this task (e.g., RivWidth (Pavelsky & Smith, 2008); RivaMap (Isikdogan et al., 2017), RivWidthCloud (Yang et al., 2019)). However, measuring river width from remote sensing images is relatively computationally expensive. Thus, to date, only static mean/bankfull river width data sets are available at the global scale (Allen & Pavelsky, 2018; Lin et al., 2020; Yamazaki et al., 2019), and the investigation of multitemporal river width has only been conducted at regional scales (e.g., Hou et al., 2019; Yang et al., 2020), although Gao et al. (2021) studied the variability of global river extent. Consequently, despite previous studies have shown that river width can be impacted by many factors, including but not limited to climate, human activities, and soil and geology characteristics (Dunne & Jerolmack, 2020; Hou et al., 2019; Pekel et al., 2016; Yang et al., 2020), the investigation of river width variability and associated driving mechanisms at the global scale is still lacking.

Recently developed cloud computing environments, such as Google Earth Engine (GEE), which archive a large number of satellite images and provide highly efficient computing power, make it possible to calculate river widths from multiple satellites over long time spans at the global level. In this study, we developed an efficient width extraction method on GEE built on previous algorithms (Feng et al., 2019, 2021; Yang et al., 2019) to extract river width from every available image from Landsat 5, 7, and 8 for all rivers globally wider than 90 m for 1984–2020 (see Section 2 for details). By applying this method to approximately 1.2 million Landsat images, we obtained a multi-temporal global river width database for the past 37 years (named Global Long-term River Width, GLOW) containing ~1.2 billion river width measurements for global rivers with a total length of ~4 million km. With this data set, we could investigate the temporal variability of global river width and identify significant patterns of river width for the past four decades in a standardized global analysis. With this analysis, we attempted to answer three primary questions: (a) What is the temporal variability of global river width during 1984–2020? (b) Are there any significant temporal trends in river width globally? And (c) what are the major natural and anthropogenic factors that explain these variations? Finally, we hope that our analysis combined with this data set can support applications in water resource management and catalyze interdisciplinary studies related to hydrology/hydraulics, geomorphology, and ecology.

2. Data and Methods

2.1. Width Extraction

2.1.1. Preprocessing

The Landsat program offers multispectral images suitable for water detection dating back to the 1970s. However, with a resolution of 30 m, Landsat imagery cannot reliably measure changing widths on rivers narrower than 90 m (Allen & Pavelsky, 2018). Therefore, before extracting river widths from these images, we filtered river reaches by removing those narrower than 90 m. We used the Multi-Error-Removed-Improved-Terrain (MERIT) Basin hydrography (Lin et al., 2019) as the underlying river network (Figure 1). To identify >90 m reaches, we extracted mean river width from two existing global data sets: $\bar{W}_{GRWL}$ from GRWL (Allen & Pavelsky, 2018) and $\bar{W}_{MERIT}$ from MERIT Hydro (Yamazaki et al., 2019) for each river reach in the MERIT Basin river network. For river reaches without GRWL or MERIT Hydro river width, we used the width-area relationship developed by Frasson et al. (2019) to estimate the mean river width ($\bar{W}_{\text{Area}}$). We then selected reaches satisfying the following criteria: max($\bar{W}_{MERIT}$, $\bar{W}_{GRWL}$) > 90 m when MERIT Hydro or GRWL width is available; otherwise $\bar{W}_{\text{Area}}$ > 90 m. Using these three estimates of mean river width maximized the spatial coverage, as the MERIT Basin river network includes more rivers than either of the previously published static width databases. After determining the river reaches suitable for width extraction, on each reach, we used a dynamic interval technique (i.e., the longitudinal interval between cross-sections is proportional to mean river width; consult Feng et al., 2021 for more details) to locate width measurements due to computational limits at the global scale, which results in an average interval of 170 ± 92 m (1σ). Note that our primary goal was to assess the dynamics of river width, not to maximize the number of cross-sections, and this alignment creates a balance between computational efficiency and data density. To reduce errors in extracted widths, we removed cross-sections close to river confluences, as these cause known errors in the width extraction process (Feng et al., 2019, 2021).
We extracted multitemporal river widths from the Landsat family of satellites for 1984–2020 in GEE (Figure 1). Here, we refer to river width as the distance between the water extent boundaries along a river cross-section, so it is not necessarily the river width within banks. The original GEE-based width extraction method, known as RivWidthCloud, was developed by Yang et al. (2019). Later, Feng et al. (2021) modified RivWidthCloud for large-scale applications. Specifically, Feng et al. (2021) mapped a fixed orthogonal vector onto the river centerline and then calculated river width for each orthogonal per image. This approach is simple and robust and significantly improves the computational efficiency compared to the original RivWidthCloud, which calculates a dynamic orthogonal (angle and length) for each image, but care must be taken in selecting the length of the orthogonal. If orthogonals are too long, surrounding surface water bodies might introduce false positive errors, while if orthogonals are too short, they will not span the full river width and introduce a false maximum width that might look like a terrain constraint. We extracted the GRWL \((W_{GRWL})\) and MERIT Hydro \((W_{MERIT})\) mean width for each cross-section, and then used Equation 1 to determine the half-length of the orthogonal line for each cross-section \((L_{orth})\). We determined \(c\) in Equation 1 by examining river width measurements against in situ river width data and GRWL, and selected the optimal value of 2.0 in this study (i.e., the orthogonal length is four times wider than the maximum cross-sectional mean river width, Equation 1). To ensure the accuracy of the resulting river width measurements, we adopted the screening process from RivWidthCloud to automatically remove width measurements when the end of the orthogonal lines falls in the water extent. This tradeoff to ensure high-quality retrievals is feasible globally but reduces the overall cross-sections used to characterize rivers.

We calculated a fixed orthogonal angle from the MERIT Basin river network, so that the orthogonal angles only need to be calculated once, which significantly improved the computational efficiency. We acknowledge that a fixed centerline and orthogonal cannot account for river meandering, but it makes it computationally feasible to conduct this analysis at a global scale and allows us to compute the temporal variability of river width in a consistent manner. The screening process that detects inappropriate orthogonals ensures the accuracy of our resulting width measurements, as above. We further discuss the impact of our method on the results in Section 4. The alternative approach, that is, a dynamic centerline and orthogonal as developed in RivWidthCloud, has limitations as well (besides the computational cost). For example, when using a dynamic centerline and orthogonal, it can be challenging to differentiate river width changes from the changes resulting from the cross-section shift due to channel meander.

\[
L_{orth} = \begin{cases} 
\max (W_{GRWL}, W_{MERIT}) \times c, & \text{if } W_{GRWL} \text{ or } W_{MERIT} \text{ is available} \\
W_{Area} \times c, & \text{if } W_{GRWL} \text{ and } W_{MERIT} \text{ are unavailable}
\end{cases}
\] (1)
where $L_{orth}$ is the half-length of the orthogonal line; $W_{A,GRWL}$ and $W_{A,MERIT}$ are the cross-sectional mean width extracted from GRWL and MERIT, respectively; $\overline{W}_{Area}$ is the estimated cross-sectional width based on the area-width relationship from (Frasson et al., 2019).

### 2.1.2. Extracting Cross-Sectional River Width

With cross-sections and orthogonals defined, we implemented the following procedures in GEE to extract river widths (illustrated in Figure 1). The major steps for this process include: (a) Upload the cross-section data, including geospatial location, orthogonal angle, and orthogonal length, to GEE. (b) Filter Landsat images from Landsat 5, 7, and 8 satellites for 1984–2020 based on cloud coverage (less than 25% was used in this study). We did not use Landsat 7 images collected after 31 May 2003, due to the Scan Line Corrector (SLC) failure. Here, we only used almost clear sky images (cloud coverage <25%) to reduce errors caused by clouds, especially the thin clouds (e.g., cirrus cloud) and cloud shadows, which are difficult to be identified by the cloud classification algorithm (Foga et al., 2017). We selected a whole scene 25% filter rather than a within-channel cloud filter for computational efficiency: the cloud percentage of a Landsat scene is available as metadata and does not require processing to apply. Subsequent cloud contamination is handled as described later. (c) Classify water pixels of Landsat images using the revised Dynamic Surface Water Extent (DSWE) method (Jones, 2019). Water classification from Landsat has a long history, and numerous methods have been developed, but DSWE is a recent innovation that is accurate and computationally efficient. Compared to other classification methods, DSWE is less sensitive to the selection of threshold values, making it more robust for large-scale applications. In addition, DSWE can detect water extent when pixels are partially covered by water, a unique and useful feature to improve the accuracy of water classification that other algorithms designed to detect open water typically do not have (Jones, 2019). (d) Intersect the classified water mask with the orthogonal line buffer (30 m) at each cross-section to calculate river flow width. The details for the river width calculation can be found in Figure S1 in Supporting Information S1. Steps c and d were repeated automatically over all filtered Landsat images to obtain the time series of cross-sectional river width 1984–2020 globally. To ensure the accuracy of this data set, we conducted a thorough quality control process by identifying and removing width measurements affected by inappropriate orthogonals, clouds, cloud shadow, hill shadow, and snow/ice (Yang et al., 2019). Specifically, we calculate the zonal statistics of clouds, shadows, snow/ice, and the number of orthogonal ends falling in the water extent for each intersection between the water extent and orthogonal buffer, and automatically remove those contaminated by these factors (defined as, for instance, cloud_ zonal_mean >0 for cloud contamination). We further validated our approach by comparing the GEE-based GLOW widths with locally extracted widths from the vectorized water surface area classified from the same Landsat images using the same water classification method (i.e., DSWE) (see Figure S2 in Supporting Information S1 for details). Please note that the method presented in this study is developed upon previous studies (Feng et al, 2019, 2021; Yang et al., 2019); the GEE functions for orthogonal angle calculation, DSWE water classification, width calculation, and the accuracy screening process (for detecting inappropriate orthogonals, clouds, cloud shadow, hill shadow, and snow/ice) are adopted from RivWidthCloud (Yang et al., 2019), and the orthogonal construction method is adopted from Feng et al. (2019, 2021).

### 2.2. Temporal Change Quantification

We were interested in how individual rivers are changing over time and in comparing them with one another. Therefore, we evaluated the temporal changes in river width using two metrics: the interquartile range (IQR) (Equation 2) and the temporal trend during 1984–2020. IQR measures the dispersion of river widths. We chose IQR because it does not require data to have an underlying normal distribution and is less sensitive to outliers. We calculated IQR for each cross-section based on the width measurements at the daily scale (we averaged the measurements when there were multiple data points on the same day at the same cross-section due to the overlap of satellite images) for 1984–2020 as:

$$IQR = Q_3 - Q_1$$

where $Q_3$ is the 75th percentile river width; $Q_1$ is the 25th percentile river width. We reported IQR for cross-sections with at least 30 width measurements. IQR is a non-normalized metric, so we expect it to be sensitive to river widths. We also calculated the coefficient of quartile variation (CQV) (Equation S1 in Supporting Information S1) to represent the relative dispersion of river widths. However, the normalized metric CQV may
introduce biases into the resulting random forest analysis (see details in Section 2.3) by encoding a river width signal into the data. So, we show the CQV results only in the Supporting Information S1 (discussed later in Sections 3 and 4).

To detect the temporal trend signals in river width during 1984–2020, we calculated the annual average river width for each river cross-section and then calculated the trend slope and p-value using the Mann-Kendall test. We reported trends for cross-sections with at least 15 years of river width observations to ensure the reliability of the trend signals. We conducted a sensitivity test on this minimum length of data (10–20 years) for trend analysis and found that our conclusions in this study are insensitive to this value. As before, Landsat river width measurements are more accurate for rivers wider than 90 m (Allen & Pavelsky, 2018), and therefore we quantified the temporal changes of river width only for river reach segments with all measurements larger than 90 m, that is, minimum cross-sectional river width measurement >90 m.

2.3. Random Forest Analysis

To further interpret the temporal variations (i.e., IQR and trends) of global river width, we collected 117 global environmental parameters in six categories: hydrology (e.g., river discharge and flow intermittency), physiography (e.g., elevation), climate (e.g., precipitation and temperature), land cover, anthropogenic impact (e.g., human regulation of rivers), and soil/geology (e.g., soil erosion), based on previously published global data sets (Dallaire et al., 2018; Fan et al., 2013; Grill et al., 2019; Hengl et al., 2017; Huscroft et al., 2018; Lin et al., 2020; Linke et al., 2019; Liu et al., 2018; Messager et al., 2021; Wada et al., 2016; Yamazaki et al., 2017; Zhu et al., 2013), as previous studies have shown that these factors can potentially impact river width (e.g., Dunne & Jerolmack, 2020; Gao et al., 2021; Hou et al., 2019; Lin et al., 2020; Pekel et al., 2016; Yang et al., 2020). These global data sets include information about upstream basin conditions as well as local, reach level conditions, which enables us to account for a broad array of natural and anthropogenic drivers. Moreover, we calculated the temporal variability (IQR) and trend of total precipitation (P), air temperature, evapotranspiration (ET), and snowmelt at the reach level for the period consistent with width data, based on the reanalysis products from ERA5 (https://cds.climate.copernicus.eu/#!/home), which is unique to this study. These statistics are also included in the analysis for identifying the drivers of temporal variations in river width. All environmental parameters used in this study are summarized in Table S1 in Supporting Information S1.

Then we developed Random Forest models to identify the most important environmental parameters for the temporal variations of river width. Random Forest is a machine learning algorithm (Breiman, 2001) capable of dealing with high-dimensional data (numeric, categorical, or mixed) and nonlinear relationships. It is therefore widely used in earth sciences (e.g., Messager et al., 2021). We used the R package “Ranger” (Wright & Ziegler, 2017), which was particularly designed for high-dimensional data, to quickly implement the Random Forest algorithm. With the Random Forest models, we mainly had two goals: (a) identifying the most important parameters explaining the temporal variability and trend of river width and (b) discovering how the temporal variability and trend of river width change with each of these parameters. The 117 environmental parameters are from 59 subcategories (Tables S1 and S2 in Supporting Information S1), and the parameters within each subcategory are correlated to various extents. To avoid the importance dilution problem between correlated parameters (Gregorutti et al., 2017), we first built the Random Forest model to select the most important parameter in each subcategory, and this process resulted in 59 parameters, which were then used as the input for the final Random Forest models. After building these Random Forest models, we calculated the importance of each parameter and identified the most important parameters for both IQR and the trend of river width. Finally, we constructed Accumulated Local Effects (ALE) plots to show how the river width variability and trend change with each of them. Here, the ALE was selected because it works well even for correlated parameters and is computationally efficient (Apley & Zhu, 2020).

Given that most data sets of the global environmental parameters mentioned above are at the reach scale, the Random Forest analysis was also conducted at the reach level. We first calculated the arithmetic mean of the multi-cross-section river widths over each river reach (defined by the MERIT Basin hydrography) to produce a single reach-averaged width for each date, based on which the IQR and trend slopes were calculated. We acknowledge that this approach could bring uncertainty due to the difference in data availability at different cross-sections within one reach. An alternative way is to take the cross-sectional statistics as the reach-scale signal; for example, we can calculate IQR and trends for each cross-section and then take the mean or median to represent that reach.
We compared these two approaches and found that the former represents the reach-scale temporal variations better, as the latter approach tends to underestimate the signals.

To investigate whether the mechanics driving changes in river width are different for free-flowing and non-free-flowing (i.e., human-impacted) rivers, we performed the Random Forest analysis for them separately. To do so, we identified the non-free-flowing river reaches by integrating the previous data set from Grill et al. (2019) with the Global River Obstructions Database (GROD) (Yang et al., 2021). Grill et al. (2019) identified the free-flowing rivers by constructing the connectivity status index (CSI). GROD is a global database of river obstructions along GRWL river centerlines, including dams and nonreservoir-producing structures (e.g., locks, low head dams, and wing dams) that were not included in Grill et al. (2019). Integrating GROD with the data set of Grill et al. (2019) enables us to identify human-impacted river reaches to the maximum extent possible. Here, we defined non-free-flowing river reaches as those with a CSI < 95 (suggested by Grill et al., 2019) or those within 500 m of a GROD structure.

3. Results

Our width extraction process resulted in 1.17 billion width observations, with 101 measurements per cross-section on average, for rivers spanning a total length of ~4 million km, made from 1.19 million Landsat images. The average number of width observations per cross-section per year varies between 2 and 8 in the 36 major basins worldwide (Figure S2 in Supporting Information S1). This large volume of river width data forms the basis for estimating the temporal variations in river width at the global scale for the past ~4 decades. As mentioned before, to ensure the accuracy of the analysis of river width variations, we selected only river segments with all width measurements >90 m (i.e., the minimum width measurement >90 m) (Allen & Pavelsky, 2018), resulting in 2.7 million cross-sections with a median river width of 439 m.

3.1. Temporal Variations of Global River Width

After calculating the IQR for all river segments with width measurements >90 m, we found that ~85% of global rivers show a slight or moderate variability in river width (Figure 2). Here, we classified rivers into five categories based on IQR: (a) slightly variable (IQR < 60 m), (b) moderately variable (60 m < IQR < 150 m), (c) variable (150 m < IQR < 300 m), (d) highly variable (300 m < IQR < 500 m), and (e) extremely variable (IQR > 500 m). Figure 2B demonstrates how dispersed the river widths are with different variabilities. The median and mean IQR for global rivers are 34 and 96 m, respectively. At the regional scale, rivers located in Asia (such as the Brahmaputra, Ganges, and Indus Rivers) and West Africa (e.g., the Niger River) show higher river width variabilities (median IQR ranges between 70 and 210 m) compared to other basins (Figures 2A and 2C). GLOW collects river widths for all open-water seasons, and thus its temporal distribution of width measurements reflects the regional seasonality (Figure 2D). For example, in the high-latitude regions, river width measurements are only available during the summer season (May to September); in contrast, for areas in the southern hemisphere, the width data is distributed all year round (Figure 2D). Even though IQR is not normalized, Figure 2A does not follow a map of mean river width, justifying our use of this metric. We repeated the same analysis for the metric CQV (results shown in Figure S3 in Supporting Information S1), and found that the spatial patterns of river width variability are largely the same regardless of the metrics (Figures 2 and S3 in Supporting Information S1).

Moving to the trend of river width, we found that approximately 36.9% of global river segments (~0.6 M out of 1.7 M cross-sections) show a significant temporal trend (p-value < 0.05) in river width during 1984–2020 (Figure 3), with 90% of the trend slopes ranging from ~1.40% to +1.39%/yr. Among these 0.6 M cross-sections, 56.4% show a decreasing trend, while 43.6% show an increasing trend. The temporal trend also presents spatial patterns. Rivers in Asia (especially the High-Mountain-Asia region), eastern Russia (e.g., the Kolyma River), North America (e.g., the Mississippi River), and South America (e.g., the Amazon River) are hotspots of increasing trends in river width (Figure 3A). In contrast, most large rivers in Europe, central Asia, and Africa are dominated by decreasing trends (Figures 3A and 3C). Among the 36 largest river basins in the world, only one third (12/36) of them show more increasing than decreasing trends, with most of these located in Asia (e.g., the Yangtze, Yellow, Mekong, and Kolyma Rivers), South America (e.g., Amazon), and North America (e.g., Mississippi and Nelson Rivers) (Figure 3C). Please note that we are reporting trend signals for flowing water seasons, which are subject to regional differences in seasonality (Figure 2D).
Figure 2.
Figure 2. (A) Temporal variability of global river width (quantified by the interquartile range, or IQR) during 1984–2020. A higher IQR indicates a higher dispersion of river width. Red circles are example cross-sections shown in (B). For clarity, only large rivers are shown in this map. (B) Cross-section examples showing river widths with different variabilities: (a) slightly variable (IQR < 60 m), (b) variable (150 m < IQR < 300 m), and (c) extremely variable (IQR > 500 m). Their major river names are shown in the bottom right of each plot. To show the distribution details consistently, we only present river width measurements for 1990–2015. (C) The distribution of river width IQR across the 36 largest river basins in the world. Numbers above or below the boxplots are the number of cross-sections measured in each basin. Figure 3. (A) Temporal trend (i.e., change rate, %/yr) of global river width during 1984–2020. Positive (negative) change rates indicate increasing (decreasing) trends in river width. An increasing (decreasing) trend means that the river has been getting wider (narrower). Only significant trends (p-value < 0.05) for large rivers are mapped. Green circles are example cross-sections shown in (B). Here we use the relative unit (%/yr, normalized by the average river width of each cross-section) rather than the absolute unit (m/yr) to provide a consistent comparison regardless of river sizes. (B) Three example cross-sections to show an increasing (a), a decreasing (b), and a human-impacted increasing (c) trend of river width. The river name in the bottom right of each plot indicates the basin where the example cross-section is located. (C) The distribution of trend slopes of river width across the 36 largest river basins. Numbers above or below the boxplots are the number of cross-sections showing a significant trend (p-value < 0.05) in river width in each basin. Violin plots above (below) zero suggest a dominant increasing (decreasing) trend in river width within the basin; violin plots spanning a wide range from below to above zero suggest diverging trends in river width in that basin.
Significant trends in river width occur in both free-flowing rivers and non-free-flowing (i.e., human-regulated) rivers (Figure 3B). Due to human interventions (e.g., dam construction and flow diversion), the river width of regulated rivers is likely to show significant changes when alteration began during the study period. For example, the La Grande River in Canada has experienced a sudden rise in river width mainly due to flow diversion from adjacent basins (Figure 3B(c)). Moving to the global scale, we found that non-free-flowing river reaches show a higher fraction (45.8%) of significant trends in river width than free-flowing rivers (33.6%) (Figure 4). This pattern can be found in two-thirds (i.e., 24 out of 36) of the large river basins in the world (Figure 4). In addition, in many basins, the non-free-flowing river reaches show higher fractions of decreasing trends than free-flowing rivers (Figure 4 and Table S2 in Supporting Information S1). Dam/reservoir impacts are dominant in the non-free-flowing rivers: 84% of these river segments are impacted by dams/reservoirs (Figure S4 in Supporting Information S1).

### 3.2. Factors Controlling River Width Variations

#### 3.2.1. Free-Flowing Rivers

For free-flowing rivers, we found that the top nine important parameters for river width temporal variability (i.e., IQR) are from three categories: physiography, climate, and hydrology (Figure 5A). The importance of all 55 parameters can be found in Table S4 in Supporting Information S1. Elevation is the most important parameter influencing river width variability (Figure 5A). Rivers at both extremely high and low elevations show a high IQR, with the lowest IQR occurring at 200–300 m. Rivers at low elevation (<50 m) are mainly located in coastal regions, the Amazon Basin, and some Arctic regions. Among these low-elevation rivers, coastal rivers in southeast China, the Ganges-Brahmaputra River, Arctic rivers (e.g., the Arctic Archipelago, Pechora, Divina, Ob', and Lena river basins), and the Negro River in the Amazon Basin are examples showing a high river width IQR. High-elevation rivers also exhibit high IQR, likely due to the seasonal changes in the cryosphere components (e.g., glacier melt and snowmelt) (Chandel & Ghosh, 2021; Li et al., 2021). In contrast, the lowest IQR is found at elevations of around 200–300 m, and major examples of such rivers include the southern Mackenzie river basin, rivers in the Hudson Bay, and the Xingu River. Sinuosity is another major parameter that affects river width IQR (Figure 5A). We find that rivers with a high sinuosity show a high IQR (Figure 5B). Sinuous rivers are more likely to meander due to the unevenly distributed flow speed in the channel, for example, the “cutbank” effect, leading to changes in channel geometry and thus flow width (Dunne & Jerolmack, 2020). In addition, we
Figure 5.
used a fixed cross-section profile to measure river width; when the river channel meanders, the resulting width measurements will probably change due to the shift in river centerline and orthogonal angle. In this situation, we acknowledge that the high IQR would be an indirect measure of channel meanders instead of river width variability. Since it is very challenging to identify channel meanders at such a fine resolution for global rivers, we did not differentiate river width changes from the effects of channel meanders.

Climatic parameters also play an important role in explaining river width variability. Temperature variability (quantified by Temperature IQR), shows the highest importance among climate factors (Figure 5A). Rivers in cold regions with a high temperature variability (IQR > 35°C) show a high river width IQR, such as the Lena and Kolyma Rivers in the Arctic. ET variability (quantified by IQR) is another climate parameter influencing river width variability. We found in regions with a low ET variability (IQR < 0.0005 m), rivers show high width variability. Such rivers are mainly distributed in dry (e.g., the Congo and Nile Rivers in Africa) or cold areas (e.g., the Himalaya). In regions where ET IQR is higher than 0.001 m, river width variability increases with ET IQR, likely due to the role of ET in regulating water that contributes to river streamflow. This pattern is also found in precipitation and snowmelt variabilities: when precipitation IQR > 0.001 m or snowmelt IQR > 0.0005 m, river width variability increases with precipitation and snowmelt IQRs (Figure 5B, panels 7 and 9). In contrast, at the lower end (i.e., low precipitation or snowmelt IQRs, which are mainly located in mid-high latitudes), rivers also show high width variability, likely due to the impacts of other factors, such as elevation and temperature. CMI, defined as $P/ET_{potential}^{−1}$ and ranging from −1 to 1, is an integral climate index of precipitation and temperature, reflecting water availability to river discharge. Our results show that in dry or hot regions (CMI < 0 in Figure 5B), river width IQR decreases as CMI increases. Examples of a low CMI and a high river width IQR include the high mountain Asia regions and surrounding rivers (e.g., the Indus and Amu Darya Rivers). On the other hand, in humid or cold areas (CMI > 0 in Figure 5B), such as North Europe, river width variability increases with CMI. The relationships between climate parameters and river width IQR suggest that regions with extreme climate conditions (e.g., extremely dry, cold, or hot) are more likely to show a high river width variability.

We found that river width IQR increases with flow intermittency and discharge variability for most natural rivers (Figure 5B). This suggests that rivers with highly variable discharge tend to show a higher width IQR. Please note that the discharge variability parameter used in this study is defined as the ratio of max monthly discharge to mean discharge (log-scale) (Dallaire et al., 2018), and for some large rivers (e.g., the mainstream Amazon River), the river width IQR is high, but the discharge variability is relatively low (<0.2), leading to the high IQR at the lower end of the discharge variability (Figure 5B panel 8).

Moving to trends in width for free-flowing rivers, the three major categories of important parameters are also climate, hydrology, and physiography (Figure 6), but climate parameters show higher importance than those for river width IQR (comparing Figures 5A and 6A). The most important variable explaining the river width trend is the minimum air temperature (MAT) (Figure 6A). We find that regions with MAT of −30°C~−17°C show the highest increasing trend in river width, and these regions are mainly located in the Arctic (e.g., the Yukon, Mackenzie, Yenisey, and the Hudson Bay), northeast China (i.e., the Amur river basin), and the High-Mountain-Asia region. The significant decreasing trend of river width for MAT lower than −35°C is mainly from the Lena river basin. The trend of river width is also sensitive to temperature change. For most rivers, the increasing temperature tends to weaken the increasing signals of river width or lead to a decrease in river width (Figure 6B). However, when the temperature increase rate is greater than 0.1°/year, it results in a positive trend in river width, and such rivers, although only a few, are mainly located in the most northern part of the Arctic (Figure 6B).

Another important factor explaining the river width trend for free-flowing rivers is the elevation (Figure 6A); the river width trend slope increases with elevation. We can find this pattern in many regions, such as the Himalayas, Andes, and Rocky Mountains (Figure 6A). We also find that these regions have experienced temperature increases at a median rate of 0.02°/year. This also partially explains the peak increasing trend in river width.

Figure 5. (A) Top nine parameters that show the highest importance for the variability of river width (quantified by interquartile range [IQR]) of free-flowing rivers during 1984–2020, based on the Random Forest analysis. Parameter importance is measured in terms of node impurity and calculated as the decrease in the residual sum of squares that results from splitting the response variable (i.e., IQR) based on each parameter. A higher node impurity (or parameter importance) indicates that the parameter is more important. (B) The Accumulated Local Effects (ALE) plots show how each parameter influences the predicted IQR. The y-axis value of the ALE curve suggests the main effect of the parameter at a certain value compared to the average prediction of the entire data set. So, a negative (positive) y-axis value indicates a lower (higher) than the average of IQR predictions, that is, a larger (smaller) y-axis value indicates a higher (lower) predicted IQR at a specific feature value. The rug lines at the bottom of each figure show the distribution of each parameter among global river reaches investigated in this study.
Figure 6.
around the temperature change rate of 0.01–0.02°/year (Figure 6B, panel Temperature Change). As expected, the acceleration of precipitation and discharge facilitates the increasing trend in river width (Figure 6B); in contrast, the increasing ET tends to reduce river width. Other factors, including mean river width, river discharge, and runoff, also impact the trend of river width, but their effects are relatively small (i.e., the relatively flat curves in Figure 6B) compared to other factors.

3.2.2. Non-Free-Flowing Rivers

For non-free-flowing rivers, the makeup of the most important features affecting the river width variations is different from those for free-flowing rivers, with soil/geology and anthropogenic parameters playing more important roles (Figures 5–8, Tables S4–S7 in Supporting Information S1). Particularly, soil conditions like soil types and soil moisture are more important for regulated rivers in explaining their width variability (Figure 7A compared to Figure 5A). The river width IQR decreases with the sand fraction when the sand content is less than 50%, but increases with sand fraction when sand is higher than 50% (Figure 7B). Examples of rivers with a low sand fraction (<30%) and a high width IQR (>100 m) include reaches in the Mississippi River, downstream Yangtze River, and rivers in central/southern India (e.g., the Narmada and Godavari Rivers). Rivers with a high sand content (>65%) and high width variability (IQR > 100 m) are mainly located in South America (e.g., the Rio Colorado and Rio Negro), North Europe (e.g., Sweden), and southern Africa (e.g., the Zambezi River). The silt fraction of the upstream soil is another soil parameter that influences river width variability, although its effect is lower than the mean sand fraction (Figure 7B). When the soil silt content is less than 40%, width IQR decreases with the silt fraction; in contrast, when soil silt is higher than 40%, IQR increases with the silt content (Figure 7B). The relatively higher river width IQR (>100 m) is mainly found in the Rio Colorado and Rio Negro in Argentina, the Zambezi River in Africa (for low silt fraction, <15%), and the Mississippi, Amur, and downstream Yangtze Rivers (for high silt fraction, >50%). We also find that river width variability in non-free-flowing rivers decreases with soil water content (Figure 7B). The soil water content was defined as the fraction of soil water available for evapotranspiration and thus is an indicator of soil stress (Trabucco & Zomer, 2019). Our results show that rivers in regions with severe soil stress, such as the Indus and Niger Rivers, tend to show high temporal variability in river width.

Anthropogenic parameters, including the population density and gross domestic production (GDP), are also important factors impacting river width variability for regulated rivers. River width IQR increases with the population density in populated regions (>50 people/km²). In highly populated areas, such as the Ganges River in India, the median width IQR can be up to 300 m (compared to 70 m for globally regulated rivers and 54 m for globally free-flowing rivers). Other similar rivers include the downstream Yangtze, Yellow, and Nile Rivers. In areas with high GDP, such as the downstream Yangtze, Mississippi, Yellow, Volga, Ganges, and Ob’ Rivers, many of them coinciding with high-population regions, have a high river width variability. The average river width IQR in these regions (GDP > 10¹² US dollars) is 250 m, significantly higher than the global average (70 m). This suggests that human regulation and economic activities play an important role in impacting the river width variability of regulated rivers. Rivers in highly populated areas with intensive industrial production activities are more likely to experience more frequent and intensive human regulations and thus higher river width variabilities.

River width variability of regulated rivers is also impacted by river discharge variability and flow intermittence. Like free-flowing rivers, most regulated rivers (70%) show an increasing width IQR when discharge variability increases. This discharge variability parameter is normalized by mean discharge and thus sensitive to river size. For some large rivers, such as the Mississippi, Yellow, Yangtze, Indus, and Ob’ Rivers, the discharge variability is relatively low (<0.3), but the river width IQR is high, which results in a high IQR at the lower end of the discharge variability (Figure 7B panel 4), similar to natural rivers (Figure 5B). Like free-flowing rivers, regulated river width IQR increases with flow intermittence (Figure 7B).

Climate parameters, such as temperature annual range and min air temperature (MAT), are also important factors explaining river width variability (Figure 7). In general, river width IQR increases with the temperature range...
Figure 7. (A) Top nine parameters that show the highest importance for the variability of river width (quantified by interquartile range [IQR]) of non-free-flowing rivers during 1984–2020, based on the Random Forest analysis. (B) The Accumulated Local Effects (ALE) plots show how each parameter influences the predicted IQR. The rug lines at the bottom of each figure show the distribution of each parameter among global river reaches investigated in this study. The parameter name followed by “(up)” indicates that the parameter is derived from the whole upstream drainage area of the current river reach; otherwise, the parameter is derived from the local catchment.
Figure 8. (A) Top nine parameters that show the highest importance for the temporal trend of river width (quantified by change rate, %/year) of non-free-flowing rivers during 1984–2020, based on the Random Forest analysis. (B) The Accumulated Local Effects (ALE) plot show how each parameter influences the predicted trend slope. The rug lines at the bottom of each figure show the distribution of each parameter among global river reaches investigated in this study. The parameter name followed by “(up)” indicates that the parameter is derived from the whole upstream drainage area of the current river reach; otherwise, the parameter is derived from the local catchment.
and decreases with MAT, except in some tropical regions where the temperature range is less than 1500°C and MAT is higher than 20°C. Example rivers with high river width IQR, high temperature range (>4,000°C), and low MAT (≤−20°C) include the Mackenzie, Ob’, Lena, Kolyma, and Amur Rivers, with an average river width IQR of >300 m.

As previously mentioned, soil and geological factors play more important roles in explaining the river width trend of non-free-flowing rivers than free-flowing rivers (Figure 8A). Higher increasing rates of river width occur in rivers with higher potential soil loss due to erosion (Figure 8B). The upstream soil types also affect the river width changing rate (Figure 8B). The high-silt and low-sand soils are favorable to the increasing trend of river width; in contrast, the low-silt and especially the high-sand soils lead to a decreasing trend of river width. Highly sandy regulated rivers, for example, the Negro and Colorado Rivers in Argentina, the São Francisco River in Brazil, and the Orange and Niger Rivers in Africa, have shown a significant decrease in river width during the past four decades.

Actual ET is another important factor for the river width changes in non-free-flowing rivers (Figure 8A). In high ET regions (>1,000 mm), such as South America, southeast China, and the southeast US, river width tends to increase over time. In low ET regions (<450 mm), mostly distributed in the northern hemisphere, many rivers show a decrease in river widths, such as the Euphrates and Tigris Rivers in west Asia, the Amu Darya and Syr Darya Rivers in central Asia, the Ob’ River in the Arctic region, the Negro River in Argentina, and the Orange River in South Africa. The river width change rate increases with the temperature range (Figure 8B), except for some tropical regions (e.g., Southeast Asia and the Madeira River in South America) where the temperature range is low (<2,000°C), but the river width has been increasing. Like free-flowing rivers, non-free-flowing rivers also show increasing river width when river discharge increases or ET decreases (Figure 8B). Upstream precipitation is also an important factor impacting the river width trend. In general, river width increases in high-precipitation regions and decreases in low-precipitation areas, especially in the Euphrates-Tigris, Amu Darya, Syr Darya, Ob’, and upper Yenisey river basins. The effect of MAT on the river width trend is relatively small (i.e., the flat curve for the most part). The spike between −40°C and −20°C is due to a small number of river reaches in the Lena and Kolyma Rivers (MAT < −30) showing a significant decrease in river width while some reaches in the upper Yenisey and the Bureya-Amur Rivers (MAT > −30) showing a positive trend in river width.

4. Discussion

This study presented the first assessment of the temporal variations of river width and associated controlling factors at the global scale. We revealed several unprecedented findings that can advance our understanding of global river dynamics. First, the temporal variability and trend signals of global river width show substantial differences across basins, regions, and natural versus regulated rivers (Figures 2–4); we identified the hotspots showing high variabilities (e.g., the Ganges–Brahmaputra Basins, Figure 2) and significant trends of river width (e.g., increasing in the Himalaya region, and decreasing in European rivers, Figure 3). Second, we found that climate variables are among the most important factors for both IQR and trends of river width for both free-flowing and regulated rivers (Figures 5–8). Previous studies (e.g., Gao et al., 2021; Yang et al., 2020) have suggested that climate factors can explain river water dynamics, but few have explicitly quantified their importance for global river width variations. Our study revealed that climate parameters show different impacts on river width variations among natural and regulated rivers. Third, elevation is the most important physiographic parameter explaining river width variations of natural rivers: rivers in high-elevation regions show a higher width IQR (Figure 5) and a faster increase in river width (Figure 6). Previous studies (e.g., Dhimal et al., 2021) have shown that mountainous areas are on the front lines of climate change. Our study implies that river width changes in mountainous regions (e.g., the Himalayan region) are probably more sensitive to climate change, especially considering that the temporal variability and trend in river width are significantly impacted by changes in temperature and precipitation (Figures 5 and 6). Fourth, soil and geological parameters play more important roles in regulated rivers than natural rivers (Figures 5–8). Specifically, soil properties (e.g., the sand and silt fractions) and soil erosion are among the most important factors affecting the temporal variations of regulated rivers (Figures 7 and 8). This finding is essential to enhance our understanding of geomorphological evolution at the global scale (Yang et al., 2015) and also provides valuable information for river management (e.g., river restoration), especially for rivers impacted by dams/reservoirs (as 84% of regulated rivers in this study are affected by dams/reservoirs, Figure S4 in Supporting
Information S1). We acknowledge that our study did not explicitly investigate the driving mechanisms behind the relationships identified by the machine learning analysis. However, by identifying the changing patterns of river width and associate controlling factors, our results chart a course for future studies; for example, further investigations are needed to explain why/how climate and soil erosion impact river width in different regions, and our work provides essential method and data for such studies. Please note that the global environmental parameters used for the RF analysis are mostly model simulations or reanalysis products, and the uncertainties or errors within these data sets are expected to propagate into the RF results.

In addition to natural factors (e.g., climate, physiography, and soil), we also revealed interesting results about the impacts of anthropogenic factors on river width variations. We have shown that human-impacted rivers are more likely to show significant trends in river width than non-regulated rivers (Figure 4), mainly driven by dam/reservoir reaches (Figure S5 in Supporting Information S1). We found that human population density and total GDP are the most important anthropogenic parameters for river width temporal variabilities (Figure 7, Tables S6–S7 in Supporting Information S1). In fact, other anthropogenic parameters representing human water consumption and degree of regulation are also in the pool of parameters tested for relative importance (Tables S1 and S2 in Supporting Information S1). But the results show that they are less important than the population and economic activities (i.e., GDP) (Tables S6–S7 in Supporting Information S1). This suggests that the static status, such as the water use rate and whether it is being regulated or not, is insufficient to explain the anthropogenic impacts on temporal variations of river width. The influences of human activities on river width are likely more complicated than represented in existing global data sets. For example, the impacts of dams/reservoirs on river width vary significantly across rivers: we found increasing, decreasing, and no trends in river width for rivers with dams (see examples in Figure S8 in Supporting Information S1), which suggests that the numeric/categorical indices (e.g., CSI and with/without dams) are probably unable to sufficiently represent the impacts of human regulations on river widths. In addition, most river infrastructure (e.g., dams and levees) in highly populated regions is not represented in existing global databases. Such infrastructure is likely becoming more common in populated areas, for example, the Amazon, Ganges–Brahmaputra, and Yangtze river basins (Zarfl et al., 2015), which could result in river width changes but was not represented here. Therefore, more data representing anthropogenic influences, when available, should be included in future studies to fully account for human impacts.

GLOW is the first long-term global river width data set and will provide useful information for other studies in hydrology, geomorphology, and fluvial eco-biology (e.g., Dunne & Jerolmack, 2020; Hotchkiss et al., 2015; Monegaglia et al., 2018; Raymond et al., 2013). However, GLOW is derived from the optical spectral reflectance of Landsat imagery and is thus subject to uncertainties induced by cloud impact and the choice of the water classification algorithm. We used only almost clear sky images (cloud coverage <25%) to reduce the effects of cloud errors caused by the imperfection of the cloud mask algorithm (Foga et al., 2017). Without this screening process, we would have 20%–30% more data in GLOW, but the quality of this additional data would be low. With this filter, we attempted to achieve a balance between accuracy and data density. The DSWE method was only validated for sites across the US and Canada (Yang et al., 2019). Thus, applying DSWE to global rivers may cause errors in the resulting width measurements, especially in places where land covers are different from those in the validation sites (Jones, 2019), although the US and Canada do represent a large area with representations from most biomes outside the humid tropics. Another primary source of uncertainty in GLOW is the definition of river centerline and orthogonals. In this study, we used fixed cross-sections and orthogonals. This approach significantly improved the computational efficiency of the width extraction process. However, it also introduces uncertainty in the width measurements by ignoring the effect of channel migration. In addition, the orthogonals with a fixed length may not be able to capture extreme water extents (i.e., width measurements were removed when the orthogonals are too short to cover the extremely wide water surface, although this situation is rare), which likely leads to underestimated temporal variability of river width from GLOW. Determined by the path geometry of Landsat satellites (with a revisit frequency of 16 days) (Allen et al., 2020) and the impact of clouds and seasons (e.g., the frozen season in the boreal and Arctic regions) (Figure 2D), the temporal resolution of GLOW widths is low (i.e., 2–8 measurements per year per cross-section on average, Figure S3 in Supporting Information S1). This may also result in a biased representation of the temporal variability of global river width.

Our selection of a representative metric also necessarily influences our results. We have chosen the IQR, a non-normalized metric, to represent the temporal variability of river widths. Therefore, our results represent absolute rather than relative changes in river widths: a mean 90 m wide river and a mean 1,000 m wide river with...
the same IQR of, for example, 50 m, represent rivers of very different morphologies but with the same absolute variability. We repeated our results using the CQV (i.e., the coefficient of quartile variation), or the IQR divided by the sum of the first and third quartiles (Eq. S1). Figure S3 in Supporting Information S1 repeats Figure 2 with this normalized metric, and the patterns of variability are largely the same. Additionally, the trends in width variability are unaffected by normalizing. Width changes are governed by cross-sectional geometry, for instance, characterized by the channel shape parameter \(r\) introduced by (Dingman, 2007), and discharge. Neither variable is measured globally. In addition, by mapping the actual river widths, we are implicitly controlling for discharge. That is, a river with shallow side slope geometry but low discharge variability has the capacity for width variability but is de facto invariant. Our analysis using observed flows therefore assesses actual variability, which is a result of complex interactions of geometry and discharge and the other factors identified by RF (Figures 5–8). We are encouraged that our results are broadly similar between normalized and non-normalized metrics (Figures 2 and S4 in Supporting Information S1). Note that our trends should accurately represent changes in river widths over time regardless of the metric.

Beyond investigating river width variations, our ultimate goal is to enhance the understanding of global river dynamics. The GLOW data set will aid in achieving this goal. For example, as mentioned in Section 1, river width is one of the most important hydrologic variables for the remote sensing of river discharge. The relationships between river width and discharge, such as the at-a-station-hydraulic-geometry (AHG, \(w = aQ^b\)) (Leopold & Maddock, 1953) or at-many-stations-geometry (AMHG) (Gleason & Smith, 2014), are commonly used for remote sensing of river discharge (e.g., Brinkerhoff et al., 2020; Gleason & Smith, 2014; Hagemann et al., 2017; Harlan et al., 2021; Pavelsky, 2014). GLOW will enable us to build river-specific AHG or AMHG relationships at the global scale (e.g., Figure S9 in Supporting Information S1), thus enhancing our understanding of these algorithms and improving the remote sensing of river discharge. Recent studies (e.g., Feng et al., 2021; Ishitsuka et al., 2021) have shown that assimilating these remote sensing-derived river discharges to hydrologic model simulations can significantly improve the accuracy of estimated river discharge at the global scale, especially for ungauged regions. GLOW may also catalyze interdisciplinary studies related to rivers, such as investigating the dynamics of interactions between the atmosphere and aquatic system (e.g., carbon emission from the river water surface) (e.g., Raymond et al., 2013) and hydro-ecological conditions for freshwater animals (e.g., Feng et al., 2020).

5. Conclusions

This study presents the first investigation of the temporal variations of river width for 1984–2020 at the global scale. We collected 1.17 billion width measurements for global rivers wider than 90 m, made from 1.19 million Landsat images (named Global LOng-term river Width, or GLOW). With this data set, we quantified the temporal variability and trends of river width for the past four decades. We found that the temporal variability of river width shows substantial differences across watersheds and regions. IQR varies between <60 m (slightly variable) and >500 m (extremely variable), with 85% within 0–150 m. Among the 1.7 million river segments investigated for trend detection, 36.9% show a significant temporal trend in river width, with 56.4% decreasing and 43.6% increasing. We also found that regulated rivers are more likely to show a significant trend in river width than free-flowing rivers (46% vs. 34%). By leveraging big data and machine learning techniques, we further investigated the mechanisms driving changes in river width. We found that the most important factors are substantially different between free-flowing and regulated rivers. The most important features for free-flowing rivers are mainly in the categories of climate, hydrology, and physiography; in contrast, for regulated rivers, aside from these three categories, soil and anthropogenic parameters also play important roles in regulating river width variations. Land cover factors are less important for both natural and regulated rivers compared to other categories of parameters.

Beyond the significant findings regarding the temporal variations of river width, we also developed GLOW, a ~40-year global river width data set. GLOW will provide data support for studying hydraulics, hydrology, geomorphology, and other river-related questions. Thus, this study enhances the understanding of global river dynamics, and we anticipate that the GLOW data set will catalyze additional work toward advancing river science and interdisciplinary studies.
Data Availability Statement
GLOW and the environmental parameters generated in this study are available at Zenodo: https://doi.org/10.5281/zenodo.6425657.

References
Allen, G. H., & 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, 42(2), 395–402. https://doi.org/10.1002/2014gl062764
Allen, G. H., & Pavelsky, T. M. (2018). River width, slope, catchment area, Meander wavelength, sinuosity, and discharge. Geophysical Research Letters, 45(2), 821–828. https://doi.org/10.1002/2017gl074417
Allen, G. H., Yang, X., Gardner, J., Holliman, J., David, C. H., & Ross, M. (2020). Timing of Landsat overpasses effectively captures flow conditions of large rivers. Remote Sensing, 12(9), 1510. https://doi.org/10.3390/rs12091510
Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. Journal of the Royal Statistical Society: Series B, 82(4), 1059–1086. https://doi.org/10.1111/rssb.12377
Bertoldi, W., Zanoni, L., & Tubino, M. (2010). Assessment of morphological changes induced by flow and flood pulses in a gravel bed braided river: The Tagliamento River (Italy). Geomorphology, 114(3), 348–360. https://doi.org/10.1016/j.geomorph.2009.07.017
Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/a:1010933404324
Brinkerhoff, C. B., Gleason, C. J., Feng, D., & Lin, P. (2020). Constraining remote river discharge estimation using reach-scale geomorphology. Water Resources Research, 56(11), e2020wr027949. https://doi.org/10.1029/2020wr027949
Caisse, D. (2006). The thermal regime of rivers: A review. Water Resources Research, 42(1), 659–678. https://doi.org/10.1029/2005wr004157
Caissie, D. (2006). The thermal regime of rivers: A review. Water Resources Research, 42(1), 659–678. https://doi.org/10.1029/2005wr004157
Caissie, D. (2006). The thermal regime of rivers: A review. Water Resources Research, 42(1), 659–678. https://doi.org/10.1029/2005wr004157
Dallaire, C. O., Lehner, B., Sayre, R., & Thieme, M. (2018). A multidisciplinary framework to derive global river reach classifications at high spatial resolution. Environmental Research Letters, 14(2), 024003. https://doi.org/10.1088/1748-9326/aad8e9
Feng, D., Gleason, C. J., Lin, P., Yang, X., Pan, M., & Ishitsuka, Y. (2021). Recent changes to Arctic river discharge. Nature Communications, 12(1), 6917. https://doi.org/10.1038/s41467-021-27228-1
Feng, D., Yang, X., & Pavelsky, T. M. (2019). Comparing discharge estimates made via the BAM algorithm in high-order Arctic rivers derived solely from optical CubeSat, Landsat, and Sentinel-2 data. Water Resources Research, 55(7), 7753–7771. https://doi.org/10.1029/2019wr025599
Feng, D., Raoufi, R., Beighley, E., Melack, J. M., Goulding, M., Barthem, R. B., et al. (2020). Future climate impacts on the hydrology of headwater streams in the Amazon River Basin: Implications for migratory goliath catfishes. Hydrological Processes, 34(26), 5402–5416. https://doi.org/10.1002/2019hyp13952
Foga, S., Scaramuzza, P. L., Guo, S., Zhu, Z., Dilley, R. D., Beckmann, T., et al. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. Remote Sensing of Environment, 194, 379–390. https://doi.org/10.1016/j.rse.2017.03.026
Frassel, R. P. D., Pavelsky, T. M., Fonstad, M. A., Durand, M. T., Allen, G. H., Schulman, G., et al. (2019). Global relationships between river width, slope, catchment area, Meander wavelength, sinuosity, and discharge. Geophysical Research Letters, 46(6), 3252–3262. https://doi.org/10.1029/2019gl082027
Gao, S., Li, Z., Chen, M., Lin, P., Hong, Z., Allen, D., et al. (2021). Spatiotemporal variability of global river extent and the natural driving factors revealed by decades of Landsat observations. GRACE gravimetry observations, and land surface model simulations. Remote Sensing of Environment, 267, 112725. https://doi.org/10.1016/j.rse.2021.112725
Gernaat, D. E. H. J., Bogaart, P. W., Vuuren, D. P. v., Biemans, H., & Niessink, R. (2017). High-resolution assessment of global technical and operational Landsat data products. Remote Sensing of Environment, 194, 379–390. https://doi.org/10.1016/j.rse.2017.03.026
Gleason, C. J., & Smith, L. C. (2014). Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geomorphology. Proceedings of the National Academy of Sciences of the United States of America, 111(13), 4788–4791. https://doi.org/10.1073/pnas.1317606111
Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. Statistics and Computing, 27(3), 659–678. https://doi.org/10.1007/s11222-016-9464-1
Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., et al. (2019). Mapping the world’s free-flooding rivers. Nature, 569(7755), 215–221. https://doi.org/10.1038/s41586-019-1111-9
Hagemeister, M. W., Gleason, C. J., & Durand, M. T. (2017). BAM: Bayesian AMHG-Manning inference of discharge using remotely sensed stream width, slope, and height. Water Resources Research, 53(11), 9692–9707. https://doi.org/10.1002/2017wr021626
Harlan, M. E., Gleason, C. J., Allen, E. H., Butman, D., Carter, T., Chu, V. W., et al. (2021). Discharge estimation from dense arrays of pressure transducers. Water Resources Research, 57(3), e2020wr028714. https://doi.org/10.1029/2020wr028714
Hengl, T., Vargas, H., Raes, D., Poesen, J., Meitinger, E., Goovaerts, P., et al. (2017). A spatial data infrastructure for earth system science in a changing environment. Data in Brief, 17, 524–533. https://doi.org/10.1016/j.dib.2017.04.013
Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruizperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. PLoS One, 12(2), e0169748. https://doi.org/10.1371/journal.pone.0169748
Hotchkiss, E. R., Hall, R. O., Jr., Sponseller, R. A., Butman, D., Klaminder, J., Laudon, H., et al. (2015). Sources of and processes controlling CO2 emissions change with the size of streams and rivers. Nature Geoscience, 8(9), 696–699. https://doi.org/10.1038/tnes2507
Zarfl, C., Lumsdon, A. E., Berlekamp, J., Tydecks, L., & Tockner, K. (2015). A global boom in hydropower dam construction. *Aquatic Sciences*, 77(1), 161–170. https://doi.org/10.1007/s00027-014-0377-0

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., et al. (2013). Global data sets of vegetation leaf area index (LAI)3g and fraction of photosynthetically active radiation (FPAR)3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (NDVI3g) for the period 1981 to 2011. *Remote Sensing*, 5(2), 927–948. https://doi.org/10.3390/rs5020927

**References From the Supporting Information**

Lehner, B., Reidy Liermann, C., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., et al. (2011). Global Reservoir and Dam Database, Version 1 (GRanDv1): Dams, Revision 01. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY.