Quantifying geomorphically effective floods using satellite observations of river mobility

Anya S Leenman[1], Louise J. Slater[1], Simon J Dadson[1], Michel Wortmann[1], and Richard Boothroyd[2]

[1]University of Oxford
[2]University of Glasgow

April 20, 2023

Abstract

Geomorphologists have long debated the relative importance of disturbance magnitude, duration and frequency in shaping landscapes. For river-channel adjustment during floods, some argue that cumulative flood 'power', rather than magnitude or duration, matters most. However, studies of flood-induced river-channel change often draw upon small datasets. Here, we combine Sentinel-2 imagery with flow data from laterally-active rivers to address this question using a larger dataset. We apply automated algorithms in Google Earth Engine to map rivers and detect their lateral shifting; we generate a large dataset to quantify channel change during 160 floods across New Zealand, Russia, and South America. Widening during these floods is best explained by their duration and cumulative hydrograph. We use a random forest regression model to predict flood-induced channel widening, with potential applications for hazard management. Ultimately, better global data on sediment supply and caliber would help us to understand flood-driven change to river planforms.

Quantifying geomorphically effective floods using satellite observations of river mobility

A. S. Leenman[1], L. J. Slater[1], S. J. Dadson[1,2], M. Wortmann[1,3] and R. Boothroyd[4]

[1]School of Geography and the Environment, University of Oxford
[2]UK Centre for Ecology and Hydrology
[3]European Centre for Medium-Range Weather Forecasts
[4]School of Geographical and Earth Sciences, University of Glasgow

Key Points:

- We develop a method to quantify river planform change during flood events, using Google Earth Engine
- We do so for a dataset of 160 floods that exceeded the 80th percentile stage, at 41 flow gauging sites on laterally active rivers
- Erosion during these high-flow events was most correlated with the event duration and summed hydrograph

Corresponding author: Anya Leenman, `anya.leenman@chch.ox.ac.uk`

Abstract

Geomorphologists have long debated the relative importance of disturbance magnitude, duration and frequency in shaping landscapes. For river-channel adjustment during floods, some argue that cumulative flood 'power', rather than magnitude or duration, matters most. However, studies of flood-induced river-channel change often draw upon small datasets. Here, we combine Sentinel-2 imagery with flow data from laterally-active rivers to address this question using a larger dataset. We apply automated algorithms in Google Earth Engine to map rivers and detect their lateral shifting; we generate a large dataset to quantify channel change during 160 floods across New Zealand, Russia, and South America. Widening during these floods is best explained by their duration and cumulative hydrograph. We use a random forest regression model to predict flood-induced channel widening, with potential applications for hazard management. Ultimately, better global data on sediment supply and caliber would help us to understand flood-driven change to river planforms.

Plain Language Summary

Some rivers change their shape over time. In this paper, we explore how high-flow events drive these river channels to reshape themselves. We use Google Earth Engine to automatically map the shapes of rivers in satellite images. We apply this method to pairs of satellite images before and after high-flow events, to understand how the river shape is changed by the event. We compare the amount of channel-widening measured to aspects of the high-flow event, including its peak, duration and total flow. We do so for 160 high-flow events, and find that the duration and total flow are most important for explaining how much a channel widens during the event. Finally, we build a statistical model to predict the average amount of channel widening for a given high-flow event. This method has potential applications for hazard management in rivers that are known to change their shape.

1 Introduction

The relative importance of disturbance magnitude, duration and frequency for shaping landscapes is a crucial question in geomorphology. Many studies have considered the effects of high-magnitude versus high-frequency events: for cumulative sediment transport (Wolman & Miller, 1960; Webb & Walling, 1982), for generating and reworking landforms (Wolman & Gerson, 1978; Kale, 2002, 2003; Surian et al., 2015), and for the resulting sedimentology (Magilligan et al., 1998; Marren, 2005). Others have considered the duration and total energy expenditure of individual disturbances and how this relates to their ability to transport sediment and reshape river channels (Costa & O'Connor, 1995; Magilligan et al., 2015). In rivers, understanding which disturbances perform the most geomorphic work — both instantaneously, and cumulatively over time — has important implications for sediment budgeting, flood conveyance, depositional records, and natural hazard management.

In rivers, the major disturbances are flood events, which have the power to reshape the channels that convey them. Such reshaping ranges from bar deposition and bank erosion (Bryndal et al., 2017) or aggradation (Morche et al., 2007; Hooke, 2016) through to widening (Fuller, 2008; Yousefi et al., 2018), reoccupation of abandoned channels (Arnaud-Fassetta et al., 2005) and large-scale reworking of floodplains (Miller, 1990). The latter can have severe impacts for society, including erosion of agricultural or residential land (Yousefi et al., 2018) or the destruction of transport and river management infrastructure (Arnaud-Fassetta et al., 2005). Conversely, aggradation during floods can raise riverbeds by several meters (Morche et al., 2007; Tunnicliffe et al., 2018), reducing a channel's conveyance capacity and the freeboard below bridges (Johnson et al., 2001). Quantitative methods are needed to understand, model, and predict how river channels can be reshaped by individual flood events.

The geomorphic effectiveness of a flood is thought to be a function of its duration and magnitude. Here, we define geomorphic effectiveness as the extent to which a flood alters the channel form by eroding or depositing sediment. We use the term 'flood' to mean any temporary rise in the water level (in our analysis, one that exceeds the 80th percentile of the water surface elevation measurements). Previous studies have suggested that the cumulative stream power (defined by Bagnold (1966) as the product of water density, acceleration due to gravity, discharge and slope) beneath a flood hydrograph must be high for the event to be geomorphically effective; the implication is that high-magnitude but brief floods, and low-magnitude but long floods, are not likely to be effective (Costa & O'Connor, 1995). However, others have suggested that additional factors (not just the cumulative power) make a flood geomorphically effective. For instance, Middleton et al. (2019) demonstrated that flood magnitude does influence geomorphic effectiveness: in the proglacial braided river they studied, planform change during floods increased with their peak discharges. Others propose that a flood's geomorphic effectiveness is not determined by the hydrograph alone, but also by the sediment supply (Church, 2014; Hooke, 2016; Bennett et al., 2017; Pfeiffer et al., 2019) or the time since the previous flood, which can influence both sediment availability and the looseness of the riverbed (Gintz et al., 1996; Hooke, 2015). These studies have advanced our understanding of geomorphic effectiveness, but almost all were small-sample case studies of 1-10 flood events or river reaches, often in similar regional or climatic contexts. Larger samples of flood events from a more geomorphically and geographically diverse set of rivers are required to produce a robust empirical assessment of what makes a geomorphically effective flood.

Google Earth Engine (GEE) has recently emerged as a key tool facilitating large-sample analyses of landscape characteristics — through both its computational platform and archive of quality controlled satellite data. The 'large-sample' approach, which addresses environmental questions using data from tens to thousands of sites, is popular in hydrology (Addor et al., 2017; Klingler et al., 2021) and has begun to be applied in geomorphology (Slater et al., 2015; Slater, 2016; Pfeiffer et al., 2019; Sylvester et al., 2019; Valenza et al., 2020; Ahrendt et al., 2022; Brooke et al., 2022; Clubb et al., 2022; Edmonds et al., 2022). A large-sample approach to studying planimetric river adjustments can be readily deployed in GEE, drawing on automated methods to map river planform (Allen & Pavelsky, 2015; Pekel et al., 2016; Zou et al., 2018; Isikdogan et al., 2019; Pickens et al., 2020; Boothroyd et al., 2021) and to track planform deformation (Wickert et al., 2013; Rowland et al., 2016; Schwenk et al., 2017; Jarriel et al., 2021; Chadwick et al., 2022; Langhorst & Pavelsky, 2022). By automating river planform tracking in GEE, the geomorphic effectiveness of a large sample of flood events can be assessed.

In this paper, we investigate the streamflow drivers of geomorphically effective floods using Sentinel-2 satellite imagery in GEE. We pursue two research questions:

1. Which hydrograph metrics best explain a flood's 2D geomorphic effectiveness?
2. How well can a flood's 2D geomorphic effectiveness be predicted from hydrologic and environmental variables?

We measure geomorphic effectiveness as the reach-averaged channel widening during a flood. We compute this planimetric erosion in GEE for flood events in Brazil, Colombia, New Zealand and Russia. We use 160 flood events at 41 flow gauging sites on laterally active rivers to evaluate our research questions (see Figure S1, Supplementary Material (SM), for gauge locations). We ascertain the influence of hydrograph shape on geomorphic effectiveness in our dataset. Finally, we develop an empirical model to predict flood-induced erosion. When coupled with streamflow forecasts, the model may be useful for hazard management in sites that are known to be laterally active.

2 Methods

Our method can be summarized as follows. First, we identified sites with historical daily stage (water level) measurements and a laterally active channel. For those rivers, we identified peaks in the stage records. Second, for each flood peak we extracted the pre- and post-flood channel planform from Sentinel-2 data in GEE, and conducted a change detection between the two planforms to quantify erosion during the flood. Ultimately, we compared the lateral erosion detected to parameters of the flood hydrograph. Figure 1 illustrates these steps with an example of one flood in Colombia. Our code is available at `https://github.com/a-leenman/2dFloodsPublic`; GEE processing was performed via the 'rgee' r package (Aybar, 2022).

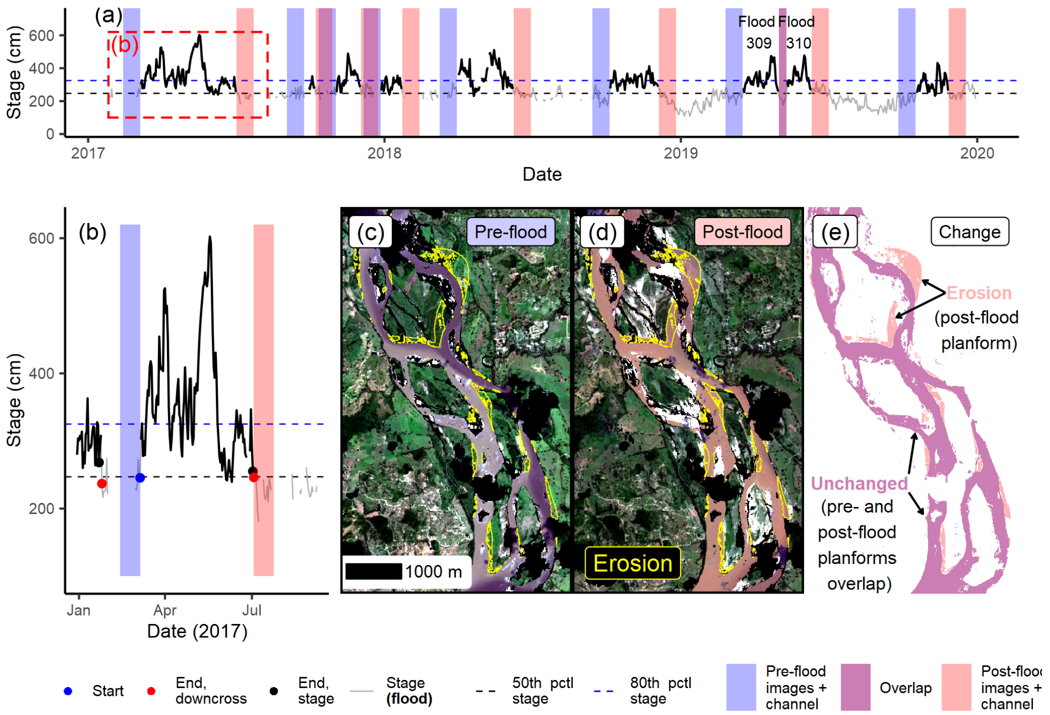

Figure 1. Methods used to define floods and detect planform change. (a) The pre-flood (blue) and post-flood (red) search windows for a sequence of floods (bold lines), showing how the windows can overlap (purple). (b) Example flood from Colombian gauge 23097040, with the flood start date (blue circle), two options for flood end date (black and red circles; the 'downcross' (red circle) method was most appropriate) and the pre- and post-flood search windows. (c) Pre-flood channel morphology, mosaicked from all cloud-free pixels in the six satellite images covering part of the AOI within the pre-flood search window. Erosion during the following flood is outlined in yellow. Black patches have no data due to cloud. (d) Corresponding post-flood mosaic (10 source images within the time and space filter). (e) The pre- and post-flood channel planforms are overlaid, highlighting the erosion (red) detected.

2.1 Site selection and area of interest

Hydrologic records are crucial to our analysis, providing flood occurrence and hydrograph shape data. We obtained publicly available stage records and gauging locations for Brazil, Colombia, New Zealand and Russia. These countries were chosen for their laterally active rivers and availability of recent daily stage records.

Other authors used discharge or stream power records to pursue this problem. However, we chose to use stage data so that differences in stage could provide a proxy for depth fluctuations when estimating the time series of shear stress. Ultimately, we aimed to approximate the sediment transport capacity of each hydrograph.

We filtered the stage records to include only those gauges that:

1. Were located on a river with a mean annual discharge above 100 cm^3 s^{-1} (data from Grill et al. (2019)), to ensure these rivers were large enough to be visible in our 10 m satellite imagery.
2. Were located on a laterally active river whose dynamics could be measured from satellite data. Laterally active rivers were identified by filtering the 'water permanence' layer from Pekel et al. (2016). After computing planform change during floods, a site was removed if the eroded area never exceeded 1% of the water surface area or if the flood-induced widening never exceeded 3 m. These thresholds enabled the largest possible dataset while excluding channels that were not laterally active.
3. Were not adjacent to large lakes or dams.
4. Overlapped with the Sentinel-2 record (June 2015 - present) by at least one year.

This filtering isolated a sample of 41 gauges. River widths ranged from 60 to 1000 m; their gradients ranged from 0.00001 to 0.002. Their mean long-term discharge ranged from 100 to 7000 $cm^3 s^{-1}$, and upstream catchment area ranged from 3800 to 430000 km^2. Values of the Richards-Baker index (Baker et al., 2004) ranged from 0.005 (very seasonal) to 0.33 (moderately flashy). Gauge altitudes ranged from 3 to 500 m. Forest cover at the gauges ranged from 0 to 100%, and mean annual rainfall from 440 to 4100 mm. The range of rivers (including braided, wandering and meandering forms) encompassed by these values highlights the geographic and geomorphic diversity of the rivers we incorporate.

For each gauge, we defined an 'Area of Interest' (AOI) in which we extract the river planform and monitor its deformation. The 'HydroSHEDS Free Flowing Rivers' vector network (Lehner et al., 2008; Grill et al., 2019) was used to select all river segments within 40 km of each gauge. We kept only the segments on the same branch as the gauge, and also removed segments that were past a jump in average discharge of >20%, implying that a 'major' tributary had been passed; we computed such jumps using the average discharge data for each segment in Grill et al. (2019). If two gauges were nearby on the same river, we divided the intervening segments between them. This left a remaining 'linked reach' (comprising one or more HydroSHEDS segments) assigned to each gauge. We extracted water masks along each reach from Allen and Pavelsky (2018a, 2018b), as a first approximation of the channel area. However, these masks do not always encompass the entire channel in our study reaches (which are extremely laterally mobile: some shift by more than 30 m in a single flood) and so we buffered these masks by 500 m to create the AOI. Finally, lakes in the HydroLAKES (Messager et al., 2016) dataset were subtracted from the AOI, to avoid spurious change detection from varying lake levels. We thus assigned to each gauge a unique AOI within which we extracted the river planform before and after each flood.

2.2 Flood delineation and search window definition

We delineated floods temporally based on the daily stage record for each gauge. Although higher frequency records were available for some countries, we resampled them by taking the daily mean stage. While this process smoothed some maxima and minima, it gave all records the same frequency. We defined a flood as any period exceeding the 80th percentile of the stage record during the Sentinel-2 record (June 2015 onwards; Figure 1a, b). Floods were extracted from the daily stage records using the hydroEvents R package (Wasko & Guo, 2022). To ensure we captured the rising and falling limbs, we defined the flood start date as the first measurement before the peak which was also below the 50th percentile of stage (Figure 1a, blue points). We defined the flood end date in two ways: either as

1. the first measurement following the peak which also fell below the 50th percentile of stage (Figure 1a, red points), or
2. the first measurement following the peak which was within 30 cm of the stage at the start of the flood (Figure 1a, black points). Occasionally, missing data meant that the first method created flood end dates that were unreasonably far after the end of the flood, necessitating the second method.

For each flood, we chose the flood end date with the stage measurement that was closest to the stage on the start date. Following the discussion in Slater et al. (2021), floods separated by less than seven days were counted as one event, and floods lasting more than 5 months were discounted as these were mostly anomalies from missing data. While this approach of using the 50th percentile to give the start and end dates assigns a longer length to floods than some standard approaches, it allows us to capture the geomorphic effects of the rising and falling limbs, and recognizes that geomorphic change and sediment entrainment likely start before the 80th percentile stage is exceeded.

Directly before and after each flood, we defined pre- and post-flood time windows of up to three weeks (Figure 1a, b). We truncated a time window if floods were less than three weeks apart; for example, flood 309 (Figure 1a) finished nine days before the following event, and so its post-flood window was truncated. If sequential events were less than six weeks apart, their pre- and post-flood windows were allowed to overlap; the post-flood window for one flood could even overlap entirely with the pre-flood window of the following event, as with floods 309 and 310 (Figure 1a; this would mean that the post-flood channel mask of flood 309 was reused as the pre-flood mask of flood 310). We used these pre- and post-flood time windows to search the Sentinel-2 archive (Level 1C, harmonized).

2.3 Planform extraction and change detection

Within each pre- and post-flood time window, we extracted the river planform from Sentinel-2 (S2) imagery. First, we mosaicked all cloud-free S2 pixels within the time window and AOI, taking the minimum reflectance in each band if multiple copies of one pixel were available. Figure 1c and d are examples of these mosaics. We proceeded with an event if at least 50% of its AOI was cloud-free; only pixels that were cloud-free in both mosaics were used. For sites in New Zealand and Russia, we also mapped snow using the normalized difference snow index, following Hofmeister et al. (2022). For snow-free scenes that met our cloud threshold, we mapped channel planform from a combination of spectral indices, following Zou et al. (2018) and Boothroyd et al. (2021); these were the normalized difference vegetation index (Rousel et al., 1973), modified normalized difference water index (Xu, 2006), and enhanced vegetation index (Huete et al., 2002). Following Boothroyd et al. (2021), we counted both water and exposed sediment (i.e. non-vegetated bars) as part of the channel, given that a lack of vegetation indicates bars are frequently inundated. While this mapping method is simple, it is generalizable to rivers with different lighting conditions and suspended sediment concentrations.

We conducted change detection between the pre- and post-flood planforms to estimate each flood's geomorphic impact. To isolate areas that were permanently (as opposed to transiently) changed during a flood, we tracked the state (wet or dry) of each pixel at monthly intervals for the following 24 months, loosely following the pixel-by-pixel trend analysis of Nagel et al. (2022). We only considered a pixel to be eroded if it switched from dry-to-wet in the flood and then continued to be wet for the subsequent two years. If cloud cover meant there were <18 months of these after-flood observations for an event, we discounted it; we chose this threshold by checking the change detection for bias due to stage fluctuations. This pixel-tracking method allowed us to eliminate spurious change detection resulting from transient stage fluctuations.

We measured a flood's geomorphic effectiveness as the area that was permanently eroded (i.e. changed from 'dry' to 'wet') during the event. We normalized this eroded area by the reach length to give the reach-averaged channel widening. Because we counted non-vegetated bars as part of the channel, it was difficult to measure deposition following the flood; newly deposited sediment was typically registered as 'channel' by our mapping algorithm. This is why we consider post-flood erosion to be the most appropriate metric of geomorphic change in our data.

Our procedures for gauge selection, cloud- and snow-filtering isolated a dataset of 160 events for which we measured geomorphic effectiveness. Because there were less than 11 floods in some countries, we pooled all floods for our subsequent analyses.

2.4 Regression and prediction

Our first research question considers the influence of hydrograph shape on geomorphic effectiveness. There are numerous metrics to characterize hydrographs, including measures of height, duration, integrated power, volume or transport capacity, and asymmetry (Brunner et al., 2021; Slater et al., 2021). Because these rivers feature a range of hydrographs (for instance, flashy versus seasonal), we use three simple metrics that allow comparison with previous studies. The first is the flood peak height, relative to the mean daily stage. The second is the cumulative value of all daily stage measurements during the flood, measured relative to mean daily stage. This cumulative water level metric is akin to the 'volume' of a hydrograph when using discharge records (e.g. Brunner et al. (2021), Figure 3). Because we use stage records, the metric accounts for the combined influence of changes in flow depth during the flood (exerting stress on the river banks/bed) and of flood duration; we refer to it as the 'summed hydrograph'. The third metric is the flood duration.

As well as exploring how hydrograph metrics correlated with erosion, we built a random forest regression model to rank the predictors' importance (by estimating how much they decreased the model's mean square error, MSE). In addition to these hydrograph metrics, we incorporated the pre-flood channel width, as channel size can positively influence channel mobility (Constantine et al., 2014; Nanson & Hickin, 1986; Langhorst & Pavelsky, 2022). Although sediment supply also increases channel mobility (e.g. Constantine et al. (2014); Ahmed et al. (2019); Donovan et al. (2021)), we do not have sediment supply time-series for our gauging sites. Instead, we used stream gradients and stage records to estimate the sediment transport capacity for each flood (see Section S1, SM for details), and added these estimates to the random forest model. We built the model using the randomForest r package (Liaw & Wiener, 2002) with 500 trees and two variables randomly sampled at each split. We used the model to predict each flood's reach-averaged erosion using leave-one-out cross-validation (LOOCV).

3 Results

In the laterally active rivers we study, floods and their geomorphic impacts vary by orders of magnitude. Peak heights vary from 30 to 700 cm above mean daily stage. The summed hydrographs vary from 40 to 30000 cm above mean daily stage, and flood durations from 1 to 152 days. The geomorphic effects of these floods are diverse, with

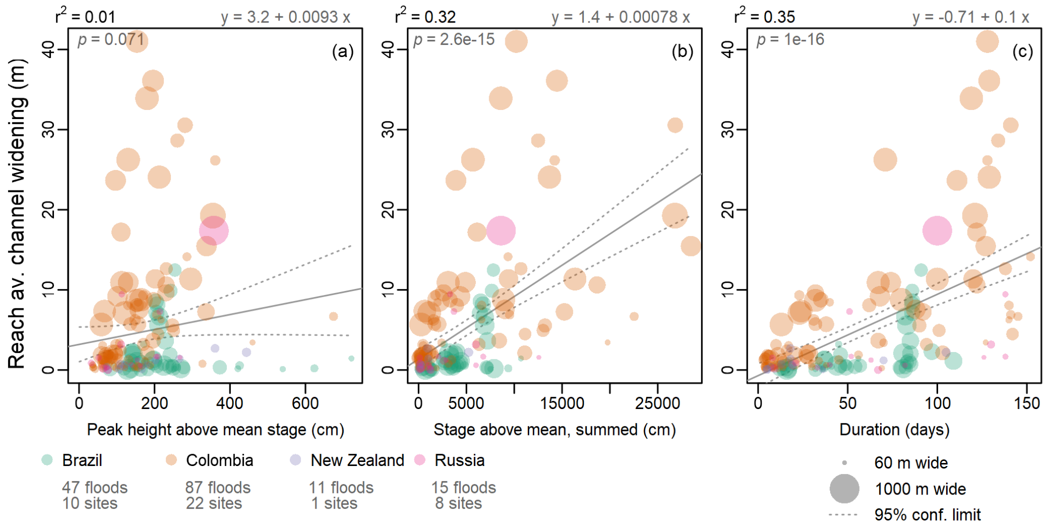

Figure 2. Flood metrics and their relationship to reach-averaged channel widening (i.e. plan-view erosion normalized by reach length) during each flood. (a) Flood peak height above the mean daily stage. (b) Cumulative stage exceeding mean daily stage ('summed hydrograph'). (c) Flood duration. Each point represents one event; colors indicate the four countries; point size is proportional to pre-flood channel width. The solid gray line shows a linear regression and dotted lines show 95% confidence limits; the regression equation is at the top-right. r^2 and p-values are at the top left. r^2 values for individual countries are in Table S1, SM.

reach-averaged widening as low as 0.005 m and as high as 41 m. The least geomorphically active country is New Zealand, with an average flood-induced widening of 0.9 m, while the most active is Colombia, with an average widening of 7 m across all floods.

Our first research question considers the erosional response of river channels to flood hydrographs. Figure 2 demonstrates how reach-averaged erosion varies with three hydrograph metrics in the 160 floods we study. Each point represents one event, with the reach-averaged erosion compared to the flood's peak height (a), summed hydrograph (b), and flood duration (c). Figure 2 therefore shows how hydrograph metrics influence geomorphic effectiveness for 160 floods at 41 sites across Brazil, Colombia, New Zealand and Russia between 2015 and 2021.

Our results indicate that reach-averaged channel widening is only weakly related to flood height in our dataset (Figure 2a). A linear regression of reach-averaged erosion during each flood against the peak height had an r^2 of just 0.01. Erosion scaled more strongly with the summed hydrograph (Figure 2b), with an r^2 of 0.32, and most strongly with flood duration (Figure 2c), with an r^2 of 0.35. See Table S1 (SM) for country-specific relationships. These coefficients of determination are surprisingly high, considering that they represent observations from real systems and are thus confounded by other natural variables in each location. Some of the relationships in Figure 2 appear non-linear (especially panel (c)), but we lack sufficient data to fit non-linear models and so we use linear regression to make a first-order comparison. These metrics are correlated among themselves (see Figure S2, SM); longer floods often had higher peaks, so that the r^2 values shown here indicate *relative* importance and we cannot say that the increase in erosion with flood duration was independent of the concurrent increase in height for many floods. Nevertheless, panels a-c indicate that, at least for our sample of laterally active

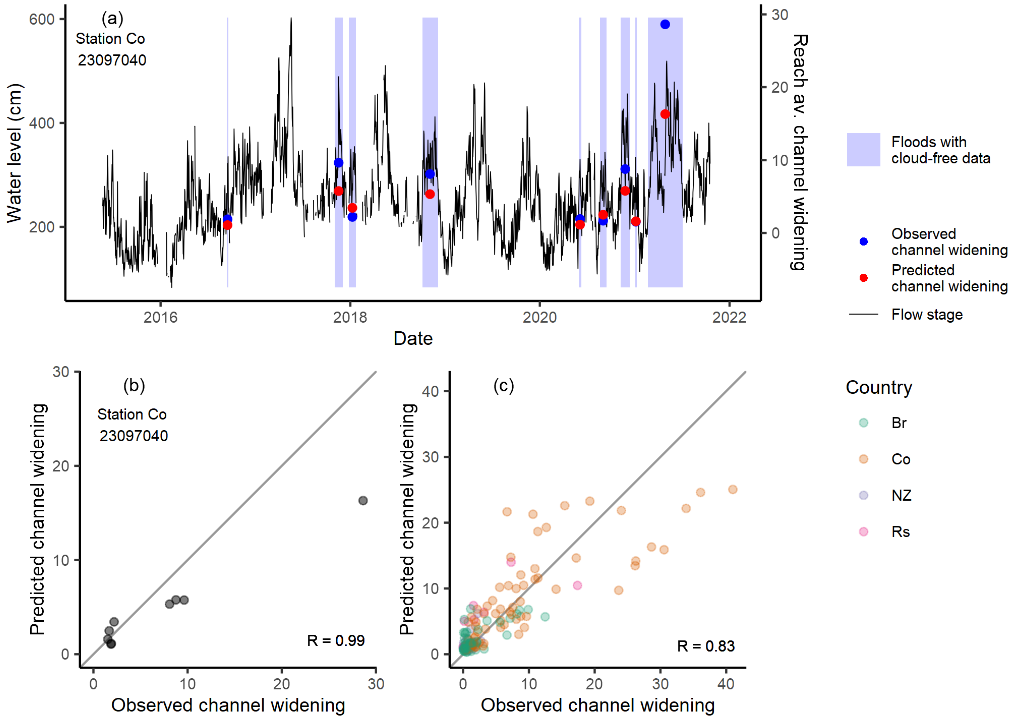

Figure 3. Predictions from our random forest regression model. (a) The stage record for Colombian gauge 23097040; flood events with sufficient cloud-free satellite data are highlighted. The observed and predicted reach-averaged erosion (channel widening) during each flood are overlain and scale with the secondary y-axis. (b) A comparison of observed and predicted channel-widening at this gauge; each point is one flood. (c) A comparison of observed and predicted channel-widening for all floods in our dataset. Grey lines in (b) and (c) show a 1:1 relation.

rivers, flood duration was the most important variable for explaining flood-driven erosion of the vegetated channel boundary.

We built a random forest regression model to rank the importance of the hydrograph metrics, channel width, and estimated sediment transport for explaining flood erosion. The random forest model ranked these variables in the following order: estimated transport, channel width, flood duration, summed hydrograph and peak height; the rankings reflect how much each variable reduced the model's MSE. This ranking is similar to the r^2 values in Figure 2 and Figures S3-S4 (SM). Because the summed hydrograph and flood duration were correlated ($R = 0.79$), we ran two additional model versions, omitting either summed hydrograph or flood duration. Although these omissions altered the variables' MSE reductions, neither altered the remaining variable rankings, implying that the rankings are not affected by this co-linearity in the predictors.

We predicted erosion for all floods in our dataset using the random forest model with LOOCV. We were able to predict erosion with at least 60% accuracy ($R = 0.83$; Figure 3c) using the pooled dataset. The model performed best for sites in Colombia with numerous floods, such as site 23097040 (Figure 3a,b). For Colombian sites with data for > 7 floods, R values were 0.78–0.99. The model tended to under-predict the highest values of reach-averaged erosion.

4 Discussion

Although there is no firm consensus, previous literature has laid the case for a hydrograph's cumulative power as the best explainer of a flood's geomorphic effectiveness. For instance, based on 10 events in Arkansas, California, Colorado, Idaho, Oregon and Washington, Costa and O'Connor (1995) suggested that a flood's geomorphic effectiveness reflected the cumulative unit stream power exceeding the threshold for alluvial erosion. Rose et al. (2020) likewise found that the most geomorphically effective floods in a sample of seven had a high energy expenditure, high peak and long duration. Kale and Hire (2007) observed that sediment transport (a proxy for geomorphic effectiveness) during monsoons rose exponentially with their cumulative stream power. Magilligan et al. (2015) attributed the limited widening during an extreme flood to its low cumulative power, resulting from a high peak but short duration. Our data partly support this hypothesis; the summed hydrograph was positively correlated with erosion during the floods we studied. However, in our dataset flood duration was a slightly better predictor of erosion of the vegetated channel boundary. This result was consistent when we raised the flood definition threshold to the 90th percentile of stage, and the summed hydrograph and flood duration had equal effects when we lowered the threshold to the 70th percentile (Figures S5 and S6, SM).

One reason for the weaker influence of the summed hydrograph in our data may be that these previous studies used the unit stream power hydrograph, whereas we used the stage hydrograph. We used stage so that changes could be used as a proxy for depth fluctuations when estimating shear stress and each hydrograph's sediment transport capacity. Although the unit transport capacity was a weaker predictor than the summed hydrograph or duration, transport became a stronger predictor when multiplied by channel width (see section S1 and Figure S3 (SM) for more detail).

The importance of flood duration in our dataset implies that, once these floods exceed the entrainment threshold, further stage increases have a smaller effect than the duration above the threshold. That is, shear stress exposure duration has a greater effect than the peak stress. This result suggests that the threshold for entrainment was low in the rivers we studied, so that full mobility of all sediment sizes was attained frequently. The regional breakdown of Figure 2 (Table S1, SM) supports this notion, as the influence of duration is strongest for Colombia where some studies have reported sand beds (e.g. Smith (1986); Martínez Silva and Nanny (2020)).

Other studies have used flood peak height, rather than cumulative power, to explain geomorphic effectiveness. For instance, Middleton et al. (2019) mapped planimetric change during floods in a proglacial river and showed that, once an annually-reset threshold discharge had been exceeded, planimetric change increased with peak discharge. Miller (1990) found that, in alluvial rivers wider than 200 m, peak unit stream power during floods was correlated with geomorphic effectiveness. In alluvial fan experiments featuring different hydrographs of the same volume, surface reworking increased with the peak discharge (Leenman et al., 2022). Nevertheless, in our dataset flood height was only weakly related to geomorphic change. It is possible that a threshold above which peak height becomes important can only be extracted by analyzing numerous floods at one location. Such an analysis is difficult in the remote sensing of real rivers, either due to seasonal floods or to persistent cloud cover, both of which limit the number of events that can be assessed.

Our results, and particularly the importance of flood duration, highlight some complexities of investigating flood impacts with a large-sample remote-sensing analysis. First, while we measured the flood-induced erosion of the vegetated channel boundaries, others simply categorized flood-driven change (e.g. (Costa & O'Connor, 1995)) or quantified sedimentological impacts (Magilligan et al., 2015). The importance of duration here is relevant to vegetated channel boundaries, but results may differ if measuring a different aspect of channel morphology — for instance, Magilligan et al. (2015) highlight how a flood event can have large sedimentological effects but a smaller impact on channel shape. Second, our large-sample analysis highlights the difficulty of finding a single parameter explaining flood effectiveness in all rivers. Flood duration was the most important driver of erosion in some rivers in our dataset, but not all; Table S1 shows that peak height was more important in Russia. Third, the relationship between a flood hydrograph and the erosion caused can be compounded by other variables, including the presence and character of vegetation, the caliber and structure of bed and bank sediment, the sediment supplied from upstream, and the time elapsed since the previous flood. In this paper, we make a first attempt at a large-sample analysis of geomorphically effective floods, and our work highlights the need for global datasets on these additional variables in order to fully address this problem.

Others have suggested that the causal relationship between a flood and its geomorphic effectiveness is moderated by sediment supply. For instance, in comparing two events on the Peace River (Canada), Church (2014, Chapter 10) found that their geomorphic effects were best explained by differences in the sediment influx. Pfeiffer et al. (2019) found that bed-level changes in Washington State were not related to high-flow events, but to sediment supply from glaciers upstream. Dean and Schmidt (2013) observed that geomorphic change during a flood in the Rio Grande was highest downstream of sediment-rich tributaries. For longer-term channel mobility, sediment supply positively influences channel migration (Constantine et al., 2014), and some rivers in our dataset (e.g. the Magdalena) have very high sediment loads (Restrepo et al., 2006; Higgins et al., 2016; Dethier et al., 2022). This question is an interesting and important one, and further work to measure sediment transport alongside flow during floods is crucial for understanding how sediment availability modulates a hydrograph's geomorphic effectiveness.

Our methods have some limitations which provide avenues for further research. The first is the suitability of using planform measurements to quantify three-dimensional channel adjustment. For landslides, erosional area scales with volume (Guzzetti et al., 2009; Larsen et al., 2010), but in rivers a 2D for 3D substitution would not be appropriate where channels are laterally confined. We have side-stepped this problem by using only laterally mobile rivers, which are therefore the rivers where a 2D for 3D substitution is most appropriate. Middleton et al. (2019) demonstrated experimentally that sediment transport scaled linearly with planimetric change, providing further justification for 2D change detection. However, further work on the suitability of measuring geomorphic change in planview would be valuable.

Further potential limitations include that of data resolution; the Sentinel-2 imagery we use has a 10 m resolution. Because erosion may occupy a smaller footprint than deposition of the same volume (Lindsay & Ashmore, 2002), finer-scale imagery may better capture erosion and would facilitate equal monitoring of both processes. An investigation of improvements with higher-resolution imagery would be worthwhile. In addition, our method computes change in the vegetated channel boundaries, so that non-vegetated bars moving through these rivers are not counted. Work comparing different algorithms to quantify river dynamics would be a useful contribution. Finally, similarity between the spectral signatures of snow and water in the mNDWI (Huang et al., 2018) meant we had to discard snowy scenes. We thus compromised slightly on our goal of a geomorphically diverse set of rivers. As the S2 record approaches a decade, the main limitation on this work is the availability of flow records, which constrains the range of sites that can be used. Methods to measure or model flow in ungauged basins could extend this work to an even more geographically diverse range of rivers.

5 Conclusions

We used Google Earth Engine and the Sentinel-2 satellite archive to map planform geomorphic change in laterally-mobile rivers during 160 flood events. By tracking each pixel for two years, we were able to separate permanent planform change from transient water extent fluctuations arising from stage variability. We measured each flood's geomorphic effectiveness as the reach-averaged erosion during the flood, and compared this to the flood hydrograph.

In the 41 laterally active rivers studied, we found that the flood peak height was only weakly correlated with erosion. The summed hydrograph was a better predictor, but erosion was most closely correlated with flood duration in our dataset of events exceeding the 80th percentile of stage.

We built a random forest regression model to predict geomorphic change for each flood, using hydrograph metrics, estimated sediment transport and channel size. The model had a prediction accuracy above 60%, which is promising for the predictability of riverbank erosion in mobile reaches.

Our work highlights the need for high-frequency flow monitoring in the world's laterally active rivers, to better understand how a flood's hydrograph controls its erosional impact. Moreover, better data on land cover, bank strength, and sediment caliber at stream gauging sites would elucidate how these characteristics modulate flood-driven erosion. Finally, monitoring sediment transport alongside river flows would help us to understand how sediment availability influences a flood's geomorphic effectiveness.

References

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*(10), 5293–5313.

Ahmed, J., Constantine, J. A., & Dunne, T. (2019). The role of sediment supply in the adjustment of channel sinuosity across the Amazon Basin. *Geology*, *47*(9), 807–810.

Ahrendt, S., Horner-Devine, A. R., Collins, B. D., Morgan, J. A., & Istanbulluoglu, E. (2022). Channel Conveyance Variability can Influence Flood Risk as Much as Streamflow Variability in Western Washington State. *Water Resources Research*, *58*(6), e2021WR031890.

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.

Allen, G. H., & Pavelsky, T. M. (2018a). Global extent of rivers and streams. *Science*, *361*(6402), 585–588.

Allen, G. H., & Pavelsky, T. M. (2018b). *Global River Widths from Landsat (GRWL) Database.* Zenodo. Retrieved from `https://gee-community-catalog.org/projects/grwl/` (Accessed through Google Earth Engine) doi: 10.5281/ZENODO.1297434

Arnaud-Fassetta, G., Cossart, E., & Fort, M. (2005). Hydro-geomorphic hazards and impact of man-made structures during the catastrophic flood of June 2000 in the Upper Guil catchment (Queyras, Southern French Alps). *Geomorphology*, *66*(1-4), 41–67.

Aybar, C. (2022). rgee: R Bindings for Calling the 'Earth Engine' API [Computer software manual]. (https://github.com/r-spatial/rgee/, https://r-spatial.github.io/rgee/, https://github.com/google/earthengine-api/)

Bagnold, R. A. (1966). *An approach to the sediment transport problem from general physics* (Report No. 422I). Retrieved from `http://pubs.er.usgs.gov/publication/pp422I` doi: 10.3133/pp422I

Baker, D. B., Richards, R. P., Loftus, T. T., & Kramer, J. W. (2004). A new flashiness index: Characteristics and applications to midwestern rivers and streams. *JAWRA Journal of the American Water Resources Association*, *40*(2), 503–522.

Bennett, G., Kean, J., Rengers, F., Ryan, S., & Rathburn, S. (2017). Landslide-channel feedbacks amplify flood response and channel erosion. In *EGU General Assembly Conference Abstracts* (p. 14326).

Boothroyd, R. J., Williams, R. D., Hoey, T. B., Barrett, B., & Prasojo, O. A. (2021). Applications of Google Earth Engine in fluvial geomorphology for detecting river channel change. *Wiley Interdisciplinary Reviews: Water*, *8*(1), e21496.

Brooke, S., Chadwick, A. J., Silvestre, J., Lamb, M. P., Edmonds, D. A., & Ganti, V. (2022). Where rivers jump course. *Science*, *376*(6596), 987–990.

Brunner, M. I., Slater, L., Tallaksen, L. M., & Clark, M. (2021). Challenges in modeling and predicting floods and droughts: A review. *WIREs Water*, *8*(3), e1520. doi: https://doi.org/10.1002/wat2.1520

Bryndal, T., Franczak, P., Kroczak, R., Cabaj, W., & Kołodziej, A. (2017). The impact of extreme rainfall and flash floods on the flood risk management process and geomorphological changes in small Carpathian catchments: a case study of the Kasiniczanka river (Outer Carpathians, Poland). *Natural Hazards*, *88*(1), 95–120.

Chadwick, A., Steel, E., Williams-Schaetzel, R., Passalacqua, P., & Paola, C. (2022). Channel migration in experimental river networks mapped by particle image velocimetry. *Journal of Geophysical Research: Earth Surface*, *127*(1), e2021JF006300.

Church, M. (2014). *The Regulation of Peace River: A Case Study for River Management.* Hoboken, UK: John Wiley & Sons.

Clubb, F. J., Weir, E. F., & Mudd, S. M. (2022). Continuous measurements of valley floor width in mountainous landscapes. *Earth Surface Dynamics*, *10*(3), 437–456.

Constantine, J. A., Dunne, T., Ahmed, J., Legleiter, C., & Lazarus, E. D. (2014). Sediment supply as a driver of river meandering and floodplain evolution in the Amazon Basin. *Nature Geoscience*, *7*(12), 899–903.

Copernicus. (n.d.). *Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-1C.* Retrieved from `https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_HARMONIZED#description` (Accessed through Google Earth Engine)

Costa, J. E., & O'Connor, J. E. (1995). Geomorphically Effective Floods. *Natural and Anthropogenic Influences in Fluvial Geomorphology: AGU Geophysical Monograph*, *89*, 45–56.

Dean, D. J., & Schmidt, J. C. (2013). The geomorphic effectiveness of a large flood on the Rio Grande in the Big Bend region: Insights on geomorphic controls and post-flood geomorphic response. *Geomorphology*, *201*, 183–198.

Dethier, E. N., Renshaw, C. E., & Magilligan, F. J. (2022). Rapid changes to global river suspended sediment flux by humans. *Science*, *376*(6600), 1447–1452.

Donovan, M., Belmont, P., & Sylvester, Z. (2021). Evaluating the relationship between meander-bend curvature, sediment supply, and migration rates. *Journal of Geophysical Research: Earth Surface*, *126*(3), e2020JF006058.

Edmonds, D. A., Martin, H. K., Valenza, J. M., Henson, R., Weissmann, G. S., Miltenberger, K., ... Hajek, E. A. (2022, Jan). Rivers in reverse: Upstream-migrating dechannelization and flooding cause avulsions on fluvial fans. *Geology*, *50*(1), 37–41. doi: 10.1130/G49318.1

Fuller, I. C. (2008). Geomorphic impacts of a 100-year flood: Kiwitea Stream, Manawatu catchment, New Zealand. *Geomorphology*, *98*(1-2), 84–95.

Gintz, D., Hassan, M. A., & Schmidt, K.-H. (1996). Frequency and magnitude of bedload transport in a mountain river. *Earth Surface Processes and Landforms*, *21*(5), 433–445.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., ... Zarfl, C. (2019). Mapping the world's free-flowing rivers. *Nature*, *569*(7755), 215–221. Retrieved from `https://developers.google.com/earth-engine/datasets/catalog/WWF_HydroSHEDS_v1_FreeFlowingRivers#description` (Data accessed via Google Earth Engine)

Guzzetti, F., Ardizzone, F., Cardinali, M., Rossi, M., & Valigi, D. (2009). Landslide volumes and landslide mobilization rates in Umbria, central Italy. *Earth and Planetary Science Letters*, *279*(3-4), 222–229.

Higgins, A., Restrepo, J. C., Ortiz, J. C., Pierini, J., & Otero, L. (2016). Suspended sediment transport in the Magdalena River (Colombia, South America): Hydrologic regime, rating parameters and effective discharge variability. *International Journal of Sediment Research*, *31*(1), 25–35.

Hofmeister, F., Arias-Rodriguez, L. F., Premier, V., Marin, C., Notarnicola, C., Disse, M., & Chiogna, G. (2022). Intercomparison of Sentinel-2 and modelled snow cover maps in a high-elevation Alpine catchment. *Journal of Hydrology X*, *15*, 100123.

Hooke, J. (2015). Variations in flood magnitude–effect relations and the implications for flood risk assessment and river management. *Geomorphology*, *251*, 91–107.

Hooke, J. (2016). Geomorphological impacts of an extreme flood in SE Spain. *Geomorphology*, *263*, 19–38.

Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: A review. *Reviews of Geophysics*, *56*(2), 333–360.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, *83*(1-2), 195–213.

Isikdogan, L. F., Bovik, A., & Passalacqua, P. (2019). Seeing through the clouds with deepwatermap. *IEEE Geoscience and Remote Sensing Letters*, *17*(10), 1662–1666.

Jarriel, T., Swartz, J., & Passalacqua, P. (2021). Global rates and patterns of channel migration in river deltas. *Proceedings of the National Academy of Sciences*, *118*(46), e2103178118.

Johnson, P. A., Hey, R. D., Horst, M. W., & Hess, A. J. (2001, February). Aggradation at bridges. *Journal of Hydraulic Engineering*, *127*(2), 154–157. Retrieved from `https://doi.org/10.1061/(asce)0733-9429(2001)127:2(154)` doi: 10.1061/(asce)0733-9429(2001)127:2(154)

Kale, V. S. (2002). Fluvial geomorphology of Indian rivers: an overview. *Progress in physical geography*, *26*(3), 400–433.

Kale, V. S. (2003). Geomorphic Effects of Monsoon Floods on Indian Rivers. In M. M. Q. Mirza, A. Dixit, & A. Nishat (Eds.), *Flood Problem and Management in South Asia* (p. 65–84). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-017-0137-2_3

Kale, V. S., & Hire, P. S. (2007). Temporal variations in the specific stream power and total energy expenditure of a monsoonal river: The Tapi River, India. *Geomorphology*, *92*(3-4), 134–146.

Klingler, C., Schulz, K., & Herrnegger, M. (2021). LamaH-CE: LArge-SaMple DAta for hydrology and environmental sciences for central Europe. *Earth System Science Data*, *13*(9), 4529–4565.

Langhorst, T., & Pavelsky, T. (2022). Global Observations of Riverbank Erosion and Accretion from Landsat Imagery. *Journal of Geophysical Research: Earth Surface*, e2022JF006774.

Larsen, I. J., Montgomery, D. R., & Korup, O. (2010). Landslide erosion controlled by hillslope material. *Nature Geoscience*, *3*(4), 247–251.

Leenman, A., Eaton, B., & MacKenzie, L. G. (2022). Floods on alluvial fans: implications for reworking rates, morphology and fan hazards. *Journal of Geophysical Research: Earth Surface*, *127*(2), e2021JF006367.

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

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18-22. Retrieved from `https://CRAN.R-project.org/doc/Rnews/`

Lindsay, J. B., & Ashmore, P. E. (2002). The effects of survey frequency on estimates of scour and fill in a braided river model. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *27*(1), 27–43.

Magilligan, F. J., Buraas, E., & Renshaw, C. (2015). The efficacy of stream power and flow duration on geomorphic responses to catastrophic flooding. *Geomorphology*, *228*, 175–188.

Magilligan, F. J., Phillips, J. D., James, L. A., & Gomez, B. (1998). Geomorphic and sedimentological controls on the effectiveness of an extreme flood. *The Journal of geology*, *106*(1), 87–96.

Marren, P. M. (2005). Magnitude and frequency in proglacial rivers: a geomorphological and sedimentological perspective. *Earth-Science Reviews*, *70*(3-4), 203–251.

Martínez Silva, P., & Nanny, M. A. (2020). Impact of microplastic fibers from the degradation of nonwoven synthetic textiles to the Magdalena River water column and river sediments by the City of Neiva, Huila (Colombia). *Water*, *12*(4), 1210.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature communications*, *7*(1), 1–11. Retrieved from https://gee-community-catalog.org/projects/hydrolakes/ (Accessed via Google Earth Engine)

Middleton, L., Ashmore, P., Leduc, P., & Sjogren, D. (2019). Rates of planimetric change in a proglacial gravel-bed braided river: Field measurement and physical modelling. *Earth Surface Processes and Landforms*, *44*(3), 752–765.

Miller, A. J. (1990). Flood hydrology and geomorphic effectiveness in the central Appalachians. *Earth Surface Processes and Landforms*, *15*(2), 119–134.

Morche, D., Schmidt, K.-h., Heckmann, T., & Haas, F. (2007). Hydrology and geomorphic effects of a high-magnitude flood in an alpine river. *Geografiska Annaler: Series A, Physical Geography*, *89*(1), 5–19.

Nagel, G. W., de Moraes Novo, E. M. L., Martins, V. S., Campos-Silva, J. V., Barbosa, C. C. F., & Bonnet, M. P. (2022). Impacts of meander migration on the Amazon riverine communities using Landsat time series and cloud computing. *Science of The Total Environment*, *806*, 150449.

Nanson, G. C., & Hickin, E. J. (1986). A statistical analysis of bank erosion and channel migration in western Canada. *Geological Society of America Bulletin*, *97*(4), 497–504.

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422.

Pfeiffer, A. M., Collins, B. D., Anderson, S. W., Montgomery, D. R., & Istanbulluoglu, E. (2019). River bed elevation variability reflects sediment supply, rather than peak flows, in the uplands of Washington State. *Water Resources Research*, *55*(8), 6795–6810.

Pickens, A. H., Hansen, M. C., Hancher, M., Stehman, S. V., Tyukavina, A., Potapov, P., ... Sherani, Z. (2020). Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment*, *243*, 111792.

Restrepo, J. D., Kjerfve, B., Hermelin, M., & Restrepo, J. C. (2006). Factors controlling sediment yield in a major South American drainage basin: the Magdalena River, Colombia. *Journal of Hydrology*, *316*(1-4), 213–232.

Rose, T., Erskine, W., & Miners, B. (2020). A customised approach to determining the geomorphic effectiveness of small flood events in a regulated river. *River Research and Applications*, *36*(4), 580–594.

Rousel, J., Haas, R., Schell, J., & Deering, D. (1973). Monitoring vegetation systems in the Great Plains with ERTS. In *Proceedings of the Third Earth Resources Technology Satellite—1 Symposium; NASA SP-351* (pp. 309–317).

Rowland, J. C., Shelef, E., Pope, P. A., Muss, J., Gangodagamage, C., Brumby, S. P., & Wilson, C. J. (2016). A morphology independent methodology for quantifying planview river change and characteristics from remotely sensed imagery. *Remote Sensing of Environment*, *184*, 212–228.

Schwenk, J., Khandelwal, A., Fratkin, M., Kumar, V., & Foufoula-Georgiou, E. (2017). High spatiotemporal resolution of river planform dynamics from Landsat: The RivMAP toolbox and results from the Ucayali River. *Earth and Space Science*, *4*(2), 46–75.

Slater, L. J. (2016). To what extent have changes in channel capacity contributed to flood hazard trends in England and Wales? *Earth Surface Processes and Landforms*, *41*(8), 1115–1128.

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., ... Wilby, R. L. (2021). Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management. *Hydrology and*

Earth System Sciences, *25*(7), 3897–3935. doi: 10.5194/hess-25-3897-2021

Slater, L. J., Singer, M. B., & Kirchner, J. W. (2015). Hydrologic versus geomorphic drivers of trends in flood hazard. *Geophysical Research Letters*, *42*(2), 370-376. doi: https://doi.org/10.1002/2014GL062482

Smith, D. G. (1986). Anastomosing river deposits, sedimentation rates and basin subsidence, Magdalena River, northwestern Colombia, South America. *Sedimentary Geology*, *46*(3-4), 177–196.

Surian, N., Barban, M., Ziliani, L., Monegato, G., Bertoldi, W., & Comiti, F. (2015). Vegetation turnover in a braided river: frequency and effectiveness of floods of different magnitude. *Earth Surface Processes and Landforms*, *40*(4), 542–558.

Sylvester, Z., Durkin, P., & Covault, J. A. (2019). High curvatures drive river meandering. *Geology*, *47*(3), 263–266.

Tunnicliffe, J., Brierley, G., Fuller, I. C., Leenman, A., Marden, M., & Peacock, D. (2018). Reaction and relaxation in a coarse-grained fluvial system following catchment-wide disturbance. *Geomorphology*, *307*, 50–64.

Valenza, J., Edmonds, D., Hwang, T., & Roy, S. (2020). Downstream changes in river avulsion style are related to channel morphology. *Nature communications*, *11*(1), 1–8.

Wasko, C., & Guo, D. (2022). Understanding event runoff coefficient variability across Australia using the hydroEvents R package. *Hydrological Processes*, *36*(4), e14563.

Webb, B., & Walling, D. (1982). The magnitude and frequency characteristics of fluvial transport in a Devon drainage basin and some geomorphological implications. *Catena*, *9*(1-2), 9–23.

Wickert, A. D., Martin, J. M., Tal, M., Kim, W., Sheets, B., & Paola, C. (2013). River channel lateral mobility: Metrics, time scales, and controls. *Journal of Geophysical Research: Earth Surface*, *118*(2), 396–412.

Wolman, M. G., & Gerson, R. (1978). Relative scales of time and effectiveness of climate in watershed geomorphology. *Earth surface processes*, *3*(2), 189–208.

Wolman, M. G., & Miller, J. P. (1960). Magnitude and frequency of forces in geomorphic processes. *The Journal of Geology*, *68*(1), 54–74.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025–3033.

Yousefi, S., Mirzaee, S., Keesstra, S., Surian, N., Pourghasemi, H. R., Zakizadeh, H. R., & Tabibian, S. (2018). Effects of an extreme flood on river morphology (case study: Karoon River, Iran). *Geomorphology*, *304*, 30–39.

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

Quantifying geomorphically effective floods using satellite observations of river mobility

A. S. Leenman[1], L. J. Slater[1], S. J. Dadson[1,2], M. Wortmann[1,3] and R. Boothroyd[4]

[1]School of Geography and the Environment, University of Oxford
[2]UK Centre for Ecology and Hydrology
[3]European Centre for Medium-Range Weather Forecasts
[4]School of Geographical and Earth Sciences, University of Glasgow

Key Points:

- We develop a method to quantify river planform change during flood events, using Google Earth Engine
- We do so for a dataset of 160 floods that exceeded the 80th percentile stage, at 41 flow gauging sites on laterally active rivers
- Erosion during these high-flow events was most correlated with the event duration and summed hydrograph

Corresponding author: Anya Leenman, `anya.leenman@chch.ox.ac.uk`

Abstract

Geomorphologists have long debated the relative importance of disturbance magnitude, duration and frequency in shaping landscapes. For river-channel adjustment during floods, some argue that cumulative flood 'power', rather than magnitude or duration, matters most. However, studies of flood-induced river-channel change often draw upon small datasets. Here, we combine Sentinel-2 imagery with flow data from laterally-active rivers to address this question using a larger dataset. We apply automated algorithms in Google Earth Engine to map rivers and detect their lateral shifting; we generate a large dataset to quantify channel change during 160 floods across New Zealand, Russia, and South America. Widening during these floods is best explained by their duration and cumulative hydrograph. We use a random forest regression model to predict flood-induced channel widening, with potential applications for hazard management. Ultimately, better global data on sediment supply and caliber would help us to understand flood-driven change to river planforms.

Plain Language Summary

Some rivers change their shape over time. In this paper, we explore how high-flow events drive these river channels to reshape themselves. We use Google Earth Engine to automatically map the shapes of rivers in satellite images. We apply this method to pairs of satellite images before and after high-flow events, to understand how the river shape is changed by the event. We compare the amount of channel-widening measured to aspects of the high-flow event, including its peak, duration and total flow. We do so for 160 high-flow events, and find that the duration and total flow are most important for explaining how much a channel widens during the event. Finally, we build a statistical model to predict the average amount of channel widening for a given high-flow event. This method has potential applications for hazard management in rivers that are known to change their shape.

1 Introduction

The relative importance of disturbance magnitude, duration and frequency for shaping landscapes is a crucial question in geomorphology. Many studies have considered the effects of high-magnitude versus high-frequency events: for cumulative sediment transport (Wolman & Miller, 1960; Webb & Walling, 1982), for generating and reworking landforms (Wolman & Gerson, 1978; Kale, 2002, 2003; Surian et al., 2015), and for the resulting sedimentology (Magilligan et al., 1998; Marren, 2005). Others have considered the duration and total energy expenditure of individual disturbances and how this relates to their ability to transport sediment and reshape river channels (Costa & O'Connor, 1995; Magilligan et al., 2015). In rivers, understanding which disturbances perform the most geomorphic work — both instantaneously, and cumulatively over time — has important implications for sediment budgeting, flood conveyance, depositional records, and natural hazard management.

In rivers, the major disturbances are flood events, which have the power to reshape the channels that convey them. Such reshaping ranges from bar deposition and bank erosion (Bryndal et al., 2017) or aggradation (Morche et al., 2007; Hooke, 2016) through to widening (Fuller, 2008; Yousefi et al., 2018), reoccupation of abandoned channels (Arnaud-Fassetta et al., 2005) and large-scale reworking of floodplains (Miller, 1990). The latter can have severe impacts for society, including erosion of agricultural or residential land (Yousefi et al., 2018) or the destruction of transport and river management infrastructure (Arnaud-Fassetta et al., 2005). Conversely, aggradation during floods can raise riverbeds by several meters (Morche et al., 2007; Tunnicliffe et al., 2018), reducing a channel's conveyance capacity and the freeboard below bridges (Johnson et al., 2001). Quantitative methods are needed to understand, model, and predict how river channels can be reshaped by individual flood events.

The geomorphic effectiveness of a flood is thought to be a function of its duration and magnitude. Here, we define geomorphic effectiveness as the extent to which a flood alters the channel form by eroding or depositing sediment. We use the term 'flood' to mean any temporary rise in the water level (in our analysis, one that exceeds the 80th percentile of the water surface elevation measurements). Previous studies have suggested that the cumulative stream power (defined by Bagnold (1966) as the product of water density, acceleration due to gravity, discharge and slope) beneath a flood hydrograph must be high for the event to be geomorphically effective; the implication is that high-magnitude but brief floods, and low-magnitude but long floods, are not likely to be effective (Costa & O'Connor, 1995). However, others have suggested that additional factors (not just the cumulative power) make a flood geomorphically effective. For instance, Middleton et al. (2019) demonstrated that flood magnitude does influence geomorphic effectiveness: in the proglacial braided river they studied, planform change during floods increased with their peak discharges. Others propose that a flood's geomorphic effectiveness is not determined by the hydrograph alone, but also by the sediment supply (Church, 2014; Hooke, 2016; Bennett et al., 2017; Pfeiffer et al., 2019) or the time since the previous flood, which can influence both sediment availability and the looseness of the riverbed (Gintz et al., 1996; Hooke, 2015). These studies have advanced our understanding of geomorphic effectiveness, but almost all were small-sample case studies of 1-10 flood events or river reaches, often in similar regional or climatic contexts. Larger samples of flood events from a more geomorphically and geographically diverse set of rivers are required to produce a robust empirical assessment of what makes a geomorphically effective flood.

Google Earth Engine (GEE) has recently emerged as a key tool facilitating large-sample analyses of landscape characteristics — through both its computational platform and archive of quality controlled satellite data. The 'large-sample' approach, which addresses environmental questions using data from tens to thousands of sites, is popular in hydrology (Addor et al., 2017; Klingler et al., 2021) and has begun to be applied in geomorphology (Slater et al., 2015; Slater, 2016; Pfeiffer et al., 2019; Sylvester et al., 2019; Valenza et al., 2020; Ahrendt et al., 2022; Brooke et al., 2022; Clubb et al., 2022; Edmonds et al., 2022). A large-sample approach to studying planimetric river adjustments can be readily deployed in GEE, drawing on automated methods to map river planform (Allen & Pavelsky, 2015; Pekel et al., 2016; Zou et al., 2018; Isikdogan et al., 2019; Pickens et al., 2020; Boothroyd et al., 2021) and to track planform deformation (Wickert et al., 2013; Rowland et al., 2016; Schwenk et al., 2017; Jarriel et al., 2021; Chadwick et al., 2022; Langhorst & Pavelsky, 2022). By automating river planform tracking in GEE, the geomorphic effectiveness of a large sample of flood events can be assessed.

In this paper, we investigate the streamflow drivers of geomorphically effective floods using Sentinel-2 satellite imagery in GEE. We pursue two research questions:

1. Which hydrograph metrics best explain a flood's 2D geomorphic effectiveness?
2. How well can a flood's 2D geomorphic effectiveness be predicted from hydrologic and environmental variables?

We measure geomorphic effectiveness as the reach-averaged channel widening during a flood. We compute this planimetric erosion in GEE for flood events in Brazil, Colombia, New Zealand and Russia. We use 160 flood events at 41 flow gauging sites on laterally active rivers to evaluate our research questions (see Figure S1, Supplementary Material (SM), for gauge locations). We ascertain the influence of hydrograph shape on geomorphic effectiveness in our dataset. Finally, we develop an empirical model to predict flood-induced erosion. When coupled with streamflow forecasts, the model may be useful for hazard management in sites that are known to be laterally active.

2 Methods

Our method can be summarized as follows. First, we identified sites with historical daily stage (water level) measurements and a laterally active channel. For those rivers, we identified peaks in the stage records. Second, for each flood peak we extracted the pre- and post-flood channel planform from Sentinel-2 data in GEE, and conducted a change detection between the two planforms to quantify erosion during the flood. Ultimately, we compared the lateral erosion detected to parameters of the flood hydrograph. Figure 1 illustrates these steps with an example of one flood in Colombia. Our code is available at `https://github.com/a-leenman/2dFloodsPublic`; GEE processing was performed via the 'rgee' r package (Aybar, 2022).

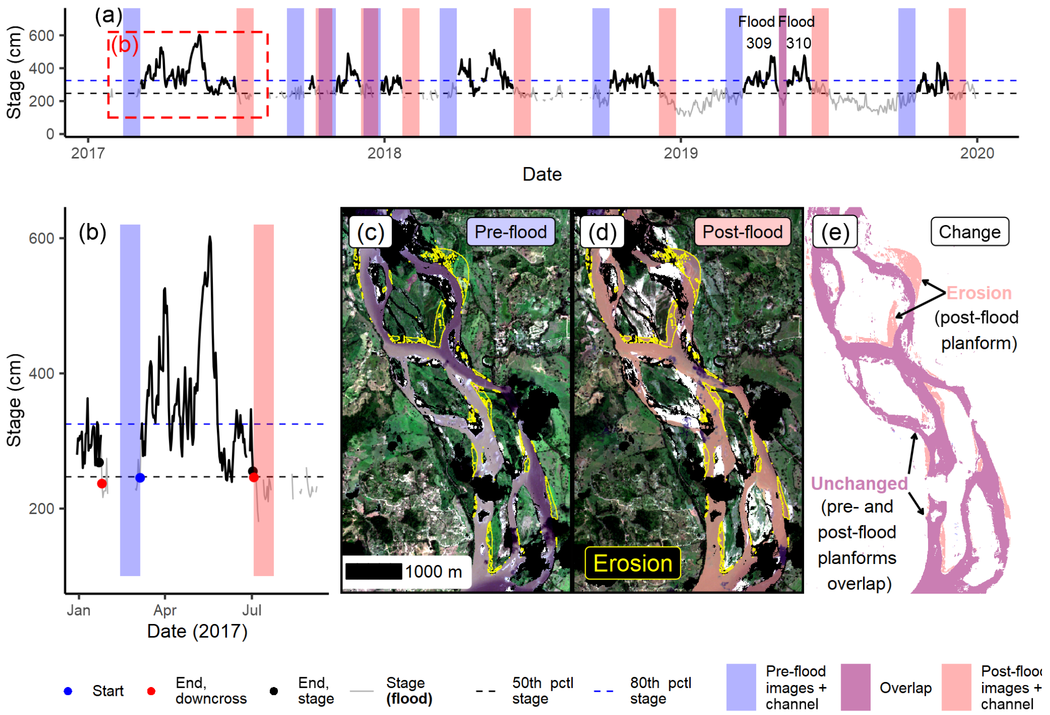

Figure 1. Methods used to define floods and detect planform change. (a) The pre-flood (blue) and post-flood (red) search windows for a sequence of floods (bold lines), showing how the windows can overlap (purple). (b) Example flood from Colombian gauge 23097040, with the flood start date (blue circle), two options for flood end date (black and red circles; the 'downcross' (red circle) method was most appropriate) and the pre- and post-flood search windows. (c) Pre-flood channel morphology, mosaicked from all cloud-free pixels in the six satellite images covering part of the AOI within the pre-flood search window. Erosion during the following flood is outlined in yellow. Black patches have no data due to cloud. (d) Corresponding post-flood mosaic (10 source images within the time and space filter). (e) The pre- and post-flood channel planforms are overlaid, highlighting the erosion (red) detected.

2.1 Site selection and area of interest

Hydrologic records are crucial to our analysis, providing flood occurrence and hydrograph shape data. We obtained publicly available stage records and gauging locations for Brazil, Colombia, New Zealand and Russia. These countries were chosen for their laterally active rivers and availability of recent daily stage records.

Other authors used discharge or stream power records to pursue this problem. However, we chose to use stage data so that differences in stage could provide a proxy for depth fluctuations when estimating the time series of shear stress. Ultimately, we aimed to approximate the sediment transport capacity of each hydrograph.

We filtered the stage records to include only those gauges that:

1. Were located on a river with a mean annual discharge above 100 cm^3 s^{-1} (data from Grill et al. (2019)), to ensure these rivers were large enough to be visible in our 10 m satellite imagery.
2. Were located on a laterally active river whose dynamics could be measured from satellite data. Laterally active rivers were identified by filtering the 'water permanence' layer from Pekel et al. (2016). After computing planform change during floods, a site was removed if the eroded area never exceeded 1% of the water surface area or if the flood-induced widening never exceeded 3 m. These thresholds enabled the largest possible dataset while excluding channels that were not laterally active.
3. Were not adjacent to large lakes or dams.
4. Overlapped with the Sentinel-2 record (June 2015 - present) by at least one year.

This filtering isolated a sample of 41 gauges. River widths ranged from 60 to 1000 m; their gradients ranged from 0.00001 to 0.002. Their mean long-term discharge ranged from 100 to 7000 $cm^3 s^{-1}$, and upstream catchment area ranged from 3800 to 430000 km^2. Values of the Richards-Baker index (Baker et al., 2004) ranged from 0.005 (very seasonal) to 0.33 (moderately flashy). Gauge altitudes ranged from 3 to 500 m. Forest cover at the gauges ranged from 0 to 100%, and mean annual rainfall from 440 to 4100 mm. The range of rivers (including braided, wandering and meandering forms) encompassed by these values highlights the geographic and geomorphic diversity of the rivers we incorporate.

For each gauge, we defined an 'Area of Interest' (AOI) in which we extract the river planform and monitor its deformation. The 'HydroSHEDS Free Flowing Rivers' vector network (Lehner et al., 2008; Grill et al., 2019) was used to select all river segments within 40 km of each gauge. We kept only the segments on the same branch as the gauge, and also removed segments that were past a jump in average discharge of >20%, implying that a 'major' tributary had been passed; we computed such jumps using the average discharge data for each segment in Grill et al. (2019). If two gauges were nearby on the same river, we divided the intervening segments between them. This left a remaining 'linked reach' (comprising one or more HydroSHEDS segments) assigned to each gauge. We extracted water masks along each reach from Allen and Pavelsky (2018a, 2018b), as a first approximation of the channel area. However, these masks do not always encompass the entire channel in our study reaches (which are extremely laterally mobile: some shift by more than 30 m in a single flood) and so we buffered these masks by 500 m to create the AOI. Finally, lakes in the HydroLAKES (Messager et al., 2016) dataset were subtracted from the AOI, to avoid spurious change detection from varying lake levels. We thus assigned to each gauge a unique AOI within which we extracted the river planform before and after each flood.

2.2 Flood delineation and search window definition

We delineated floods temporally based on the daily stage record for each gauge. Although higher frequency records were available for some countries, we resampled them by taking the daily mean stage. While this process smoothed some maxima and minima, it gave all records the same frequency. We defined a flood as any period exceeding the 80th percentile of the stage record during the Sentinel-2 record (June 2015 onwards; Figure 1a, b). Floods were extracted from the daily stage records using the hydroEvents R package (Wasko & Guo, 2022). To ensure we captured the rising and falling limbs, we defined the flood start date as the first measurement before the peak which was also below the 50th percentile of stage (Figure 1a, blue points). We defined the flood end date in two ways: either as

1. the first measurement following the peak which also fell below the 50th percentile of stage (Figure 1a, red points), or
2. the first measurement following the peak which was within 30 cm of the stage at the start of the flood (Figure 1a, black points). Occasionally, missing data meant that the first method created flood end dates that were unreasonably far after the end of the flood, necessitating the second method.

For each flood, we chose the flood end date with the stage measurement that was closest to the stage on the start date. Following the discussion in Slater et al. (2021), floods separated by less than seven days were counted as one event, and floods lasting more than 5 months were discounted as these were mostly anomalies from missing data. While this approach of using the 50th percentile to give the start and end dates assigns a longer length to floods than some standard approaches, it allows us to capture the geomorphic effects of the rising and falling limbs, and recognizes that geomorphic change and sediment entrainment likely start before the 80th percentile stage is exceeded.

Directly before and after each flood, we defined pre- and post-flood time windows of up to three weeks (Figure 1a, b). We truncated a time window if floods were less than three weeks apart; for example, flood 309 (Figure 1a) finished nine days before the following event, and so its post-flood window was truncated. If sequential events were less than six weeks apart, their pre- and post-flood windows were allowed to overlap; the post-flood window for one flood could even overlap entirely with the pre-flood window of the following event, as with floods 309 and 310 (Figure 1a; this would mean that the post-flood channel mask of flood 309 was reused as the pre-flood mask of flood 310). We used these pre- and post-flood time windows to search the Sentinel-2 archive (Level 1C, harmonized).

2.3 Planform extraction and change detection

Within each pre- and post-flood time window, we extracted the river planform from Sentinel-2 (S2) imagery. First, we mosaicked all cloud-free S2 pixels within the time window and AOI, taking the minimum reflectance in each band if multiple copies of one pixel were available. Figure 1c and d are examples of these mosaics. We proceeded with an event if at least 50% of its AOI was cloud-free; only pixels that were cloud-free in both mosaics were used. For sites in New Zealand and Russia, we also mapped snow using the normalized difference snow index, following Hofmeister et al. (2022). For snow-free scenes that met our cloud threshold, we mapped channel planform from a combination of spectral indices, following Zou et al. (2018) and Boothroyd et al. (2021); these were the normalized difference vegetation index (Rousel et al., 1973), modified normalized difference water index (Xu, 2006), and enhanced vegetation index (Huete et al., 2002). Following Boothroyd et al. (2021), we counted both water and exposed sediment (i.e. non-vegetated bars) as part of the channel, given that a lack of vegetation indicates bars are frequently inundated. While this mapping method is simple, it is generalizable to rivers with different lighting conditions and suspended sediment concentrations.

We conducted change detection between the pre- and post-flood planforms to estimate each flood's geomorphic impact. To isolate areas that were permanently (as opposed to transiently) changed during a flood, we tracked the state (wet or dry) of each pixel at monthly intervals for the following 24 months, loosely following the pixel-by-pixel trend analysis of Nagel et al. (2022). We only considered a pixel to be eroded if it switched from dry-to-wet in the flood and then continued to be wet for the subsequent two years. If cloud cover meant there were <18 months of these after-flood observations for an event, we discounted it; we chose this threshold by checking the change detection for bias due to stage fluctuations. This pixel-tracking method allowed us to eliminate spurious change detection resulting from transient stage fluctuations.

We measured a flood's geomorphic effectiveness as the area that was permanently eroded (i.e. changed from 'dry' to 'wet') during the event. We normalized this eroded area by the reach length to give the reach-averaged channel widening. Because we counted non-vegetated bars as part of the channel, it was difficult to measure deposition following the flood; newly deposited sediment was typically registered as 'channel' by our mapping algorithm. This is why we consider post-flood erosion to be the most appropriate metric of geomorphic change in our data.

Our procedures for gauge selection, cloud- and snow-filtering isolated a dataset of 160 events for which we measured geomorphic effectiveness. Because there were less than 11 floods in some countries, we pooled all floods for our subsequent analyses.

2.4 Regression and prediction

Our first research question considers the influence of hydrograph shape on geomorphic effectiveness. There are numerous metrics to characterize hydrographs, including measures of height, duration, integrated power, volume or transport capacity, and asymmetry (Brunner et al., 2021; Slater et al., 2021). Because these rivers feature a range of hydrographs (for instance, flashy versus seasonal), we use three simple metrics that allow comparison with previous studies. The first is the flood peak height, relative to the mean daily stage. The second is the cumulative value of all daily stage measurements during the flood, measured relative to mean daily stage. This cumulative water level metric is akin to the 'volume' of a hydrograph when using discharge records (e.g. Brunner et al. (2021), Figure 3). Because we use stage records, the metric accounts for the combined influence of changes in flow depth during the flood (exerting stress on the river banks/bed) and of flood duration; we refer to it as the 'summed hydrograph'. The third metric is the flood duration.

As well as exploring how hydrograph metrics correlated with erosion, we built a random forest regression model to rank the predictors' importance (by estimating how much they decreased the model's mean square error, MSE). In addition to these hydrograph metrics, we incorporated the pre-flood channel width, as channel size can positively influence channel mobility (Constantine et al., 2014; Nanson & Hickin, 1986; Langhorst & Pavelsky, 2022). Although sediment supply also increases channel mobility (e.g. Constantine et al. (2014); Ahmed et al. (2019); Donovan et al. (2021)), we do not have sediment supply time-series for our gauging sites. Instead, we used stream gradients and stage records to estimate the sediment transport capacity for each flood (see Section S1, SM for details), and added these estimates to the random forest model. We built the model using the randomForest r package (Liaw & Wiener, 2002) with 500 trees and two variables randomly sampled at each split. We used the model to predict each flood's reach-averaged erosion using leave-one-out cross-validation (LOOCV).

3 Results

In the laterally active rivers we study, floods and their geomorphic impacts vary by orders of magnitude. Peak heights vary from 30 to 700 cm above mean daily stage. The summed hydrographs vary from 40 to 30000 cm above mean daily stage, and flood durations from 1 to 152 days. The geomorphic effects of these floods are diverse, with

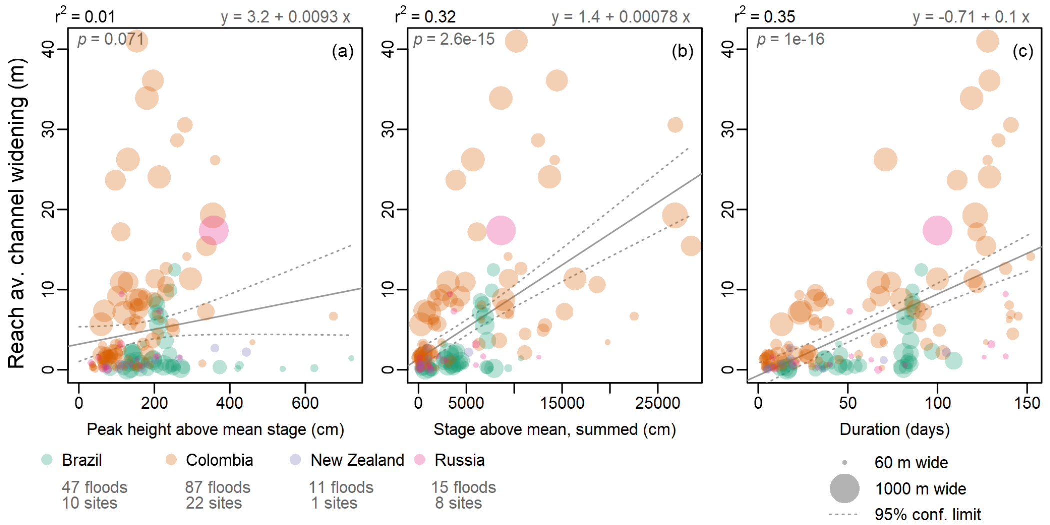

Figure 2. Flood metrics and their relationship to reach-averaged channel widening (i.e. plan-view erosion normalized by reach length) during each flood. (a) Flood peak height above the mean daily stage. (b) Cumulative stage exceeding mean daily stage ('summed hydrograph'). (c) Flood duration. Each point represents one event; colors indicate the four countries; point size is proportional to pre-flood channel width. The solid gray line shows a linear regression and dotted lines show 95% confidence limits; the regression equation is at the top-right. r^2 and p-values are at the top left. r^2 values for individual countries are in Table S1, SM.

reach-averaged widening as low as 0.005 m and as high as 41 m. The least geomorphically active country is New Zealand, with an average flood-induced widening of 0.9 m, while the most active is Colombia, with an average widening of 7 m across all floods.

Our first research question considers the erosional response of river channels to flood hydrographs. Figure 2 demonstrates how reach-averaged erosion varies with three hydrograph metrics in the 160 floods we study. Each point represents one event, with the reach-averaged erosion compared to the flood's peak height (a), summed hydrograph (b), and flood duration (c). Figure 2 therefore shows how hydrograph metrics influence geomorphic effectiveness for 160 floods at 41 sites across Brazil, Colombia, New Zealand and Russia between 2015 and 2021.

Our results indicate that reach-averaged channel widening is only weakly related to flood height in our dataset (Figure 2a). A linear regression of reach-averaged erosion during each flood against the peak height had an r^2 of just 0.01. Erosion scaled more strongly with the summed hydrograph (Figure 2b), with an r^2 of 0.32, and most strongly with flood duration (Figure 2c), with an r^2 of 0.35. See Table S1 (SM) for country-specific relationships. These coefficients of determination are surprisingly high, considering that they represent observations from real systems and are thus confounded by other natural variables in each location. Some of the relationships in Figure 2 appear non-linear (especially panel (c)), but we lack sufficient data to fit non-linear models and so we use linear regression to make a first-order comparison. These metrics are correlated among themselves (see Figure S2, SM); longer floods often had higher peaks, so that the r^2 values shown here indicate *relative* importance and we cannot say that the increase in erosion with flood duration was independent of the concurrent increase in height for many floods. Nevertheless, panels a-c indicate that, at least for our sample of laterally active

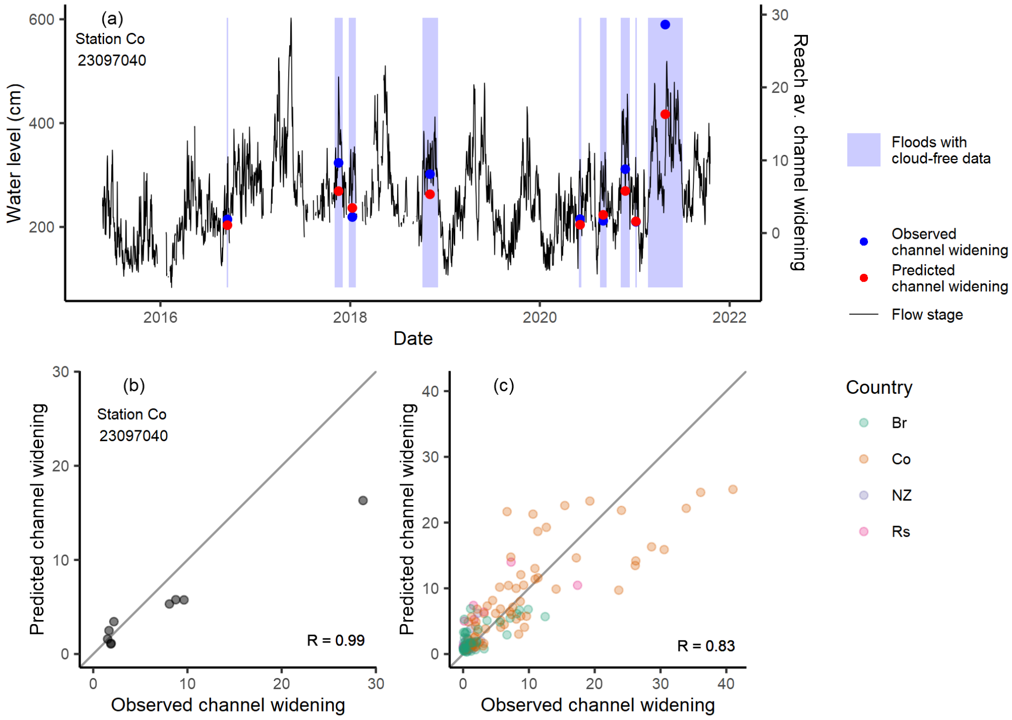

Figure 3. Predictions from our random forest regression model. (a) The stage record for Colombian gauge 23097040; flood events with sufficient cloud-free satellite data are highlighted. The observed and predicted reach-averaged erosion (channel widening) during each flood are overlain and scale with the secondary y-axis. (b) A comparison of observed and predicted channel-widening at this gauge; each point is one flood. (c) A comparison of observed and predicted channel-widening for all floods in our dataset. Grey lines in (b) and (c) show a 1:1 relation.

rivers, flood duration was the most important variable for explaining flood-driven erosion of the vegetated channel boundary.

We built a random forest regression model to rank the importance of the hydrograph metrics, channel width, and estimated sediment transport for explaining flood erosion. The random forest model ranked these variables in the following order: estimated transport, channel width, flood duration, summed hydrograph and peak height; the rankings reflect how much each variable reduced the model's MSE. This ranking is similar to the r^2 values in Figure 2 and Figures S3-S4 (SM). Because the summed hydrograph and flood duration were correlated ($R = 0.79$), we ran two additional model versions, omitting either summed hydrograph or flood duration. Although these omissions altered the variables' MSE reductions, neither altered the remaining variable rankings, implying that the rankings are not affected by this co-linearity in the predictors.

We predicted erosion for all floods in our dataset using the random forest model with LOOCV. We were able to predict erosion with at least 60% accuracy ($R = 0.83$; Figure 3c) using the pooled dataset. The model performed best for sites in Colombia with numerous floods, such as site 23097040 (Figure 3a,b). For Colombian sites with data for > 7 floods, R values were 0.78–0.99. The model tended to under-predict the highest values of reach-averaged erosion.

4 Discussion

Although there is no firm consensus, previous literature has laid the case for a hydrograph's cumulative power as the best explainer of a flood's geomorphic effectiveness. For instance, based on 10 events in Arkansas, California, Colorado, Idaho, Oregon and Washington, Costa and O'Connor (1995) suggested that a flood's geomorphic effectiveness reflected the cumulative unit stream power exceeding the threshold for alluvial erosion. Rose et al. (2020) likewise found that the most geomorphically effective floods in a sample of seven had a high energy expenditure, high peak and long duration. Kale and Hire (2007) observed that sediment transport (a proxy for geomorphic effectiveness) during monsoons rose exponentially with their cumulative stream power. Magilligan et al. (2015) attributed the limited widening during an extreme flood to its low cumulative power, resulting from a high peak but short duration. Our data partly support this hypothesis; the summed hydrograph was positively correlated with erosion during the floods we studied. However, in our dataset flood duration was a slightly better predictor of erosion of the vegetated channel boundary. This result was consistent when we raised the flood definition threshold to the 90th percentile of stage, and the summed hydrograph and flood duration had equal effects when we lowered the threshold to the 70th percentile (Figures S5 and S6, SM).

One reason for the weaker influence of the summed hydrograph in our data may be that these previous studies used the unit stream power hydrograph, whereas we used the stage hydrograph. We used stage so that changes could be used as a proxy for depth fluctuations when estimating shear stress and each hydrograph's sediment transport capacity. Although the unit transport capacity was a weaker predictor than the summed hydrograph or duration, transport became a stronger predictor when multiplied by channel width (see section S1 and Figure S3 (SM) for more detail).

The importance of flood duration in our dataset implies that, once these floods exceed the entrainment threshold, further stage increases have a smaller effect than the duration above the threshold. That is, shear stress exposure duration has a greater effect than the peak stress. This result suggests that the threshold for entrainment was low in the rivers we studied, so that full mobility of all sediment sizes was attained frequently. The regional breakdown of Figure 2 (Table S1, SM) supports this notion, as the influence of duration is strongest for Colombia where some studies have reported sand beds (e.g. Smith (1986); Martínez Silva and Nanny (2020)).

Other studies have used flood peak height, rather than cumulative power, to explain geomorphic effectiveness. For instance, Middleton et al. (2019) mapped planimetric change during floods in a proglacial river and showed that, once an annually-reset threshold discharge had been exceeded, planimetric change increased with peak discharge. Miller (1990) found that, in alluvial rivers wider than 200 m, peak unit stream power during floods was correlated with geomorphic effectiveness. In alluvial fan experiments featuring different hydrographs of the same volume, surface reworking increased with the peak discharge (Leenman et al., 2022). Nevertheless, in our dataset flood height was only weakly related to geomorphic change. It is possible that a threshold above which peak height becomes important can only be extracted by analyzing numerous floods at one location. Such an analysis is difficult in the remote sensing of real rivers, either due to seasonal floods or to persistent cloud cover, both of which limit the number of events that can be assessed.

Our results, and particularly the importance of flood duration, highlight some complexities of investigating flood impacts with a large-sample remote-sensing analysis. First, while we measured the flood-induced erosion of the vegetated channel boundaries, others simply categorized flood-driven change (e.g. (Costa & O'Connor, 1995)) or quantified sedimentological impacts (Magilligan et al., 2015). The importance of duration here is relevant to vegetated channel boundaries, but results may differ if measuring a different aspect of channel morphology — for instance, Magilligan et al. (2015) highlight how a flood event can have large sedimentological effects but a smaller impact on channel shape. Second, our large-sample analysis highlights the difficulty of finding a single parameter explaining flood effectiveness in all rivers. Flood duration was the most important driver of erosion in some rivers in our dataset, but not all; Table S1 shows that peak height was more important in Russia. Third, the relationship between a flood hydrograph and the erosion caused can be compounded by other variables, including the presence and character of vegetation, the caliber and structure of bed and bank sediment, the sediment supplied from upstream, and the time elapsed since the previous flood. In this paper, we make a first attempt at a large-sample analysis of geomorphically effective floods, and our work highlights the need for global datasets on these additional variables in order to fully address this problem.

Others have suggested that the causal relationship between a flood and its geomorphic effectiveness is moderated by sediment supply. For instance, in comparing two events on the Peace River (Canada), Church (2014, Chapter 10) found that their geomorphic effects were best explained by differences in the sediment influx. Pfeiffer et al. (2019) found that bed-level changes in Washington State were not related to high-flow events, but to sediment supply from glaciers upstream. Dean and Schmidt (2013) observed that geomorphic change during a flood in the Rio Grande was highest downstream of sediment-rich tributaries. For longer-term channel mobility, sediment supply positively influences channel migration (Constantine et al., 2014), and some rivers in our dataset (e.g. the Magdalena) have very high sediment loads (Restrepo et al., 2006; Higgins et al., 2016; Dethier et al., 2022). This question is an interesting and important one, and further work to measure sediment transport alongside flow during floods is crucial for understanding how sediment availability modulates a hydrograph's geomorphic effectiveness.

Our methods have some limitations which provide avenues for further research. The first is the suitability of using planform measurements to quantify three-dimensional channel adjustment. For landslides, erosional area scales with volume (Guzzetti et al., 2009; Larsen et al., 2010), but in rivers a 2D for 3D substitution would not be appropriate where channels are laterally confined. We have side-stepped this problem by using only laterally mobile rivers, which are therefore the rivers where a 2D for 3D substitution is most appropriate. Middleton et al. (2019) demonstrated experimentally that sediment transport scaled linearly with planimetric change, providing further justification for 2D change detection. However, further work on the suitability of measuring geomorphic change in planview would be valuable.

Further potential limitations include that of data resolution; the Sentinel-2 imagery we use has a 10 m resolution. Because erosion may occupy a smaller footprint than deposition of the same volume (Lindsay & Ashmore, 2002), finer-scale imagery may better capture erosion and would facilitate equal monitoring of both processes. An investigation of improvements with higher-resolution imagery would be worthwhile. In addition, our method computes change in the vegetated channel boundaries, so that non-vegetated bars moving through these rivers are not counted. Work comparing different algorithms to quantify river dynamics would be a useful contribution. Finally, similarity between the spectral signatures of snow and water in the mNDWI (Huang et al., 2018) meant we had to discard snowy scenes. We thus compromised slightly on our goal of a geomorphically diverse set of rivers. As the S2 record approaches a decade, the main limitation on this work is the availability of flow records, which constrains the range of sites that can be used. Methods to measure or model flow in ungauged basins could extend this work to an even more geographically diverse range of rivers.

5 Conclusions

We used Google Earth Engine and the Sentinel-2 satellite archive to map planform geomorphic change in laterally-mobile rivers during 160 flood events. By tracking each pixel for two years, we were able to separate permanent planform change from transient water extent fluctuations arising from stage variability. We measured each flood's geomorphic effectiveness as the reach-averaged erosion during the flood, and compared this to the flood hydrograph.

In the 41 laterally active rivers studied, we found that the flood peak height was only weakly correlated with erosion. The summed hydrograph was a better predictor, but erosion was most closely correlated with flood duration in our dataset of events exceeding the 80th percentile of stage.

We built a random forest regression model to predict geomorphic change for each flood, using hydrograph metrics, estimated sediment transport and channel size. The model had a prediction accuracy above 60%, which is promising for the predictability of riverbank erosion in mobile reaches.

Our work highlights the need for high-frequency flow monitoring in the world's laterally active rivers, to better understand how a flood's hydrograph controls its erosional impact. Moreover, better data on land cover, bank strength, and sediment caliber at stream gauging sites would elucidate how these characteristics modulate flood-driven erosion. Finally, monitoring sediment transport alongside river flows would help us to understand how sediment availability influences a flood's geomorphic effectiveness.

References

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*(10), 5293–5313.

Ahmed, J., Constantine, J. A., & Dunne, T. (2019). The role of sediment supply in the adjustment of channel sinuosity across the Amazon Basin. *Geology*, *47*(9), 807–810.

Ahrendt, S., Horner-Devine, A. R., Collins, B. D., Morgan, J. A., & Istanbulluoglu, E. (2022). Channel Conveyance Variability can Influence Flood Risk as Much as Streamflow Variability in Western Washington State. *Water Resources Research*, *58*(6), e2021WR031890.

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.

Allen, G. H., & Pavelsky, T. M. (2018a). Global extent of rivers and streams. *Science*, *361*(6402), 585–588.

Allen, G. H., & Pavelsky, T. M. (2018b). *Global River Widths from Landsat (GRWL) Database.* Zenodo. Retrieved from `https://gee-community-catalog.org/projects/grwl/` (Accessed through Google Earth Engine) doi: 10.5281/ZENODO.1297434

Arnaud-Fassetta, G., Cossart, E., & Fort, M. (2005). Hydro-geomorphic hazards and impact of man-made structures during the catastrophic flood of June 2000 in the Upper Guil catchment (Queyras, Southern French Alps). *Geomorphology*, *66*(1-4), 41–67.

Aybar, C. (2022). rgee: R Bindings for Calling the 'Earth Engine' API [Computer software manual]. (https://github.com/r-spatial/rgee/, https://r-spatial.github.io/rgee/, https://github.com/google/earthengine-api/)

Bagnold, R. A. (1966). *An approach to the sediment transport problem from general physics* (Report No. 422I). Retrieved from `http://pubs.er.usgs.gov/publication/pp422I` doi: 10.3133/pp422I

Baker, D. B., Richards, R. P., Loftus, T. T., & Kramer, J. W. (2004). A new flashiness index: Characteristics and applications to midwestern rivers and streams. *JAWRA Journal of the American Water Resources Association*, *40*(2), 503–522.

Bennett, G., Kean, J., Rengers, F., Ryan, S., & Rathburn, S. (2017). Landslide-channel feedbacks amplify flood response and channel erosion. In *EGU General Assembly Conference Abstracts* (p. 14326).

Boothroyd, R. J., Williams, R. D., Hoey, T. B., Barrett, B., & Prasojo, O. A. (2021). Applications of Google Earth Engine in fluvial geomorphology for detecting river channel change. *Wiley Interdisciplinary Reviews: Water*, *8*(1), e21496.

Brooke, S., Chadwick, A. J., Silvestre, J., Lamb, M. P., Edmonds, D. A., & Ganti, V. (2022). Where rivers jump course. *Science*, *376*(6596), 987–990.

Brunner, M. I., Slater, L., Tallaksen, L. M., & Clark, M. (2021). Challenges in modeling and predicting floods and droughts: A review. *WIREs Water*, *8*(3), e1520. doi: https://doi.org/10.1002/wat2.1520

Bryndal, T., Franczak, P., Kroczak, R., Cabaj, W., & Kołodziej, A. (2017). The impact of extreme rainfall and flash floods on the flood risk management process and geomorphological changes in small Carpathian catchments: a case study of the Kasiniczanka river (Outer Carpathians, Poland). *Natural Hazards*, *88*(1), 95–120.

Chadwick, A., Steel, E., Williams-Schaetzel, R., Passalacqua, P., & Paola, C. (2022). Channel migration in experimental river networks mapped by particle image velocimetry. *Journal of Geophysical Research: Earth Surface*, *127*(1), e2021JF006300.

Church, M. (2014). *The Regulation of Peace River: A Case Study for River Management.* Hoboken, UK: John Wiley & Sons.

Clubb, F. J., Weir, E. F., & Mudd, S. M. (2022). Continuous measurements of valley floor width in mountainous landscapes. *Earth Surface Dynamics*, *10*(3), 437–456.

Constantine, J. A., Dunne, T., Ahmed, J., Legleiter, C., & Lazarus, E. D. (2014). Sediment supply as a driver of river meandering and floodplain evolution in the Amazon Basin. *Nature Geoscience*, *7*(12), 899–903.

Copernicus. (n.d.). *Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-1C.* Retrieved from `https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_HARMONIZED#description` (Accessed through Google Earth Engine)

Costa, J. E., & O'Connor, J. E. (1995). Geomorphically Effective Floods. *Natural and Anthropogenic Influences in Fluvial Geomorphology: AGU Geophysical Monograph*, *89*, 45–56.

Dean, D. J., & Schmidt, J. C. (2013). The geomorphic effectiveness of a large flood on the Rio Grande in the Big Bend region: Insights on geomorphic controls and post-flood geomorphic response. *Geomorphology*, *201*, 183–198.

Dethier, E. N., Renshaw, C. E., & Magilligan, F. J. (2022). Rapid changes to global river suspended sediment flux by humans. *Science*, *376*(6600), 1447–1452.

Donovan, M., Belmont, P., & Sylvester, Z. (2021). Evaluating the relationship between meander-bend curvature, sediment supply, and migration rates. *Journal of Geophysical Research: Earth Surface*, *126*(3), e2020JF006058.

Edmonds, D. A., Martin, H. K., Valenza, J. M., Henson, R., Weissmann, G. S., Miltenberger, K., ... Hajek, E. A. (2022, Jan). Rivers in reverse: Upstream-migrating dechannelization and flooding cause avulsions on fluvial fans. *Geology*, *50*(1), 37–41. doi: 10.1130/G49318.1

Fuller, I. C. (2008). Geomorphic impacts of a 100-year flood: Kiwitea Stream, Manawatu catchment, New Zealand. *Geomorphology*, *98*(1-2), 84–95.

Gintz, D., Hassan, M. A., & Schmidt, K.-H. (1996). Frequency and magnitude of bedload transport in a mountain river. *Earth Surface Processes and Landforms*, *21*(5), 433–445.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., ... Zarfl, C. (2019). Mapping the world's free-flowing rivers. *Nature*, *569*(7755), 215–221. Retrieved from `https://developers.google.com/earth-engine/datasets/catalog/WWF_HydroSHEDS_v1_FreeFlowingRivers#description` (Data accessed via Google Earth Engine)

Guzzetti, F., Ardizzone, F., Cardinali, M., Rossi, M., & Valigi, D. (2009). Landslide volumes and landslide mobilization rates in Umbria, central Italy. *Earth and Planetary Science Letters*, *279*(3-4), 222–229.

Higgins, A., Restrepo, J. C., Ortiz, J. C., Pierini, J., & Otero, L. (2016). Suspended sediment transport in the Magdalena River (Colombia, South America): Hydrologic regime, rating parameters and effective discharge variability. *International Journal of Sediment Research*, *31*(1), 25–35.

Hofmeister, F., Arias-Rodriguez, L. F., Premier, V., Marin, C., Notarnicola, C., Disse, M., & Chiogna, G. (2022). Intercomparison of Sentinel-2 and modelled snow cover maps in a high-elevation Alpine catchment. *Journal of Hydrology X*, *15*, 100123.

Hooke, J. (2015). Variations in flood magnitude–effect relations and the implications for flood risk assessment and river management. *Geomorphology*, *251*, 91–107.

Hooke, J. (2016). Geomorphological impacts of an extreme flood in SE Spain. *Geomorphology*, *263*, 19–38.

Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: A review. *Reviews of Geophysics*, *56*(2), 333–360.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, *83*(1-2), 195–213.

Isikdogan, L. F., Bovik, A., & Passalacqua, P. (2019). Seeing through the clouds with deepwatermap. *IEEE Geoscience and Remote Sensing Letters*, *17*(10), 1662–1666.

Jarriel, T., Swartz, J., & Passalacqua, P. (2021). Global rates and patterns of channel migration in river deltas. *Proceedings of the National Academy of Sciences*, *118*(46), e2103178118.

Johnson, P. A., Hey, R. D., Horst, M. W., & Hess, A. J. (2001, February). Aggradation at bridges. *Journal of Hydraulic Engineering*, *127*(2), 154–157. Retrieved from `https://doi.org/10.1061/(asce)0733-9429(2001)127:2(154)` doi: 10.1061/(asce)0733-9429(2001)127:2(154)

Kale, V. S. (2002). Fluvial geomorphology of Indian rivers: an overview. *Progress in physical geography*, *26*(3), 400–433.

Kale, V. S. (2003). Geomorphic Effects of Monsoon Floods on Indian Rivers. In M. M. Q. Mirza, A. Dixit, & A. Nishat (Eds.), *Flood Problem and Management in South Asia* (p. 65–84). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-017-0137-2_3

Kale, V. S., & Hire, P. S. (2007). Temporal variations in the specific stream power and total energy expenditure of a monsoonal river: The Tapi River, India. *Geomorphology*, *92*(3-4), 134–146.

Klingler, C., Schulz, K., & Herrnegger, M. (2021). LamaH-CE: LArge-SaMple DAta for hydrology and environmental sciences for central Europe. *Earth System Science Data*, *13*(9), 4529–4565.

Langhorst, T., & Pavelsky, T. (2022). Global Observations of Riverbank Erosion and Accretion from Landsat Imagery. *Journal of Geophysical Research: Earth Surface*, e2022JF006774.

Larsen, I. J., Montgomery, D. R., & Korup, O. (2010). Landslide erosion controlled by hillslope material. *Nature Geoscience*, *3*(4), 247–251.

Leenman, A., Eaton, B., & MacKenzie, L. G. (2022). Floods on alluvial fans: implications for reworking rates, morphology and fan hazards. *Journal of Geophysical Research: Earth Surface*, *127*(2), e2021JF006367.

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

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18-22. Retrieved from `https://CRAN.R-project.org/doc/Rnews/`

Lindsay, J. B., & Ashmore, P. E. (2002). The effects of survey frequency on estimates of scour and fill in a braided river model. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *27*(1), 27–43.

Magilligan, F. J., Buraas, E., & Renshaw, C. (2015). The efficacy of stream power and flow duration on geomorphic responses to catastrophic flooding. *Geomorphology*, *228*, 175–188.

Magilligan, F. J., Phillips, J. D., James, L. A., & Gomez, B. (1998). Geomorphic and sedimentological controls on the effectiveness of an extreme flood. *The Journal of geology*, *106*(1), 87–96.

Marren, P. M. (2005). Magnitude and frequency in proglacial rivers: a geomorphological and sedimentological perspective. *Earth-Science Reviews*, *70*(3-4), 203–251.

Martínez Silva, P., & Nanny, M. A. (2020). Impact of microplastic fibers from the degradation of nonwoven synthetic textiles to the Magdalena River water column and river sediments by the City of Neiva, Huila (Colombia). *Water*, *12*(4), 1210.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature communications*, *7*(1), 1–11. Retrieved from `https://gee-community-catalog.org/projects/hydrolakes/` (Accessed via Google Earth Engine)

Middleton, L., Ashmore, P., Leduc, P., & Sjogren, D. (2019). Rates of planimetric change in a proglacial gravel-bed braided river: Field measurement and physical modelling. *Earth Surface Processes and Landforms*, *44*(3), 752–765.

Miller, A. J. (1990). Flood hydrology and geomorphic effectiveness in the central Appalachians. *Earth Surface Processes and Landforms*, *15*(2), 119–134.

Morche, D., Schmidt, K.-h., Heckmann, T., & Haas, F. (2007). Hydrology and geomorphic effects of a high-magnitude flood in an alpine river. *Geografiska Annaler: Series A, Physical Geography*, *89*(1), 5–19.

Nagel, G. W., de Moraes Novo, E. M. L., Martins, V. S., Campos-Silva, J. V., Barbosa, C. C. F., & Bonnet, M. P. (2022). Impacts of meander migration on the Amazon riverine communities using Landsat time series and cloud computing. *Science of The Total Environment*, *806*, 150449.

Nanson, G. C., & Hickin, E. J. (1986). A statistical analysis of bank erosion and channel migration in western Canada. *Geological Society of America Bulletin*, *97*(4), 497–504.

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422.

Pfeiffer, A. M., Collins, B. D., Anderson, S. W., Montgomery, D. R., & Istanbulluoglu, E. (2019). River bed elevation variability reflects sediment supply, rather than peak flows, in the uplands of Washington State. *Water Resources Research*, *55*(8), 6795–6810.

Pickens, A. H., Hansen, M. C., Hancher, M., Stehman, S. V., Tyukavina, A., Potapov, P., … Sherani, Z. (2020). Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment*, *243*, 111792.

Restrepo, J. D., Kjerfve, B., Hermelin, M., & Restrepo, J. C. (2006). Factors controlling sediment yield in a major South American drainage basin: the Magdalena River, Colombia. *Journal of Hydrology*, *316*(1-4), 213–232.

Rose, T., Erskine, W., & Miners, B. (2020). A customised approach to determining the geomorphic effectiveness of small flood events in a regulated river. *River Research and Applications*, *36*(4), 580–594.

Rousel, J., Haas, R., Schell, J., & Deering, D. (1973). Monitoring vegetation systems in the Great Plains with ERTS. In *Proceedings of the Third Earth Resources Technology Satellite—1 Symposium; NASA SP-351* (pp. 309–317).

Rowland, J. C., Shelef, E., Pope, P. A., Muss, J., Gangodagamage, C., Brumby, S. P., & Wilson, C. J. (2016). A morphology independent methodology for quantifying planview river change and characteristics from remotely sensed imagery. *Remote Sensing of Environment*, *184*, 212–228.

Schwenk, J., Khandelwal, A., Fratkin, M., Kumar, V., & Foufoula-Georgiou, E. (2017). High spatiotemporal resolution of river planform dynamics from Landsat: The RivMAP toolbox and results from the Ucayali River. *Earth and Space Science*, *4*(2), 46–75.

Slater, L. J. (2016). To what extent have changes in channel capacity contributed to flood hazard trends in England and Wales? *Earth Surface Processes and Landforms*, *41*(8), 1115–1128.

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., … Wilby, R. L. (2021). Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management. *Hydrology and*

Earth System Sciences, *25*(7), 3897–3935. doi: 10.5194/hess-25-3897-2021
Slater, L. J., Singer, M. B., & Kirchner, J. W. (2015). Hydrologic versus geomorphic drivers of trends in flood hazard. *Geophysical Research Letters*, *42*(2), 370-376. doi: https://doi.org/10.1002/2014GL062482
Smith, D. G. (1986). Anastomosing river deposits, sedimentation rates and basin subsidence, Magdalena River, northwestern Colombia, South America. *Sedimentary Geology*, *46*(3-4), 177–196.
Surian, N., Barban, M., Ziliani, L., Monegato, G., Bertoldi, W., & Comiti, F. (2015). Vegetation turnover in a braided river: frequency and effectiveness of floods of different magnitude. *Earth Surface Processes and Landforms*, *40*(4), 542–558.
Sylvester, Z., Durkin, P., & Covault, J. A. (2019). High curvatures drive river meandering. *Geology*, *47*(3), 263–266.
Tunnicliffe, J., Brierley, G., Fuller, I. C., Leenman, A., Marden, M., & Peacock, D. (2018). Reaction and relaxation in a coarse-grained fluvial system following catchment-wide disturbance. *Geomorphology*, *307*, 50–64.
Valenza, J., Edmonds, D., Hwang, T., & Roy, S. (2020). Downstream changes in river avulsion style are related to channel morphology. *Nature communications*, *11*(1), 1–8.
Wasko, C., & Guo, D. (2022). Understanding event runoff coefficient variability across Australia using the hydroEvents R package. *Hydrological Processes*, *36*(4), e14563.
Webb, B., & Walling, D. (1982). The magnitude and frequency characteristics of fluvial transport in a Devon drainage basin and some geomorphological implications. *Catena*, *9*(1-2), 9–23.
Wickert, A. D., Martin, J. M., Tal, M., Kim, W., Sheets, B., & Paola, C. (2013). River channel lateral mobility: Metrics, time scales, and controls. *Journal of Geophysical Research: Earth Surface*, *118*(2), 396–412.
Wolman, M. G., & Gerson, R. (1978). Relative scales of time and effectiveness of climate in watershed geomorphology. *Earth surface processes*, *3*(2), 189–208.
Wolman, M. G., & Miller, J. P. (1960). Magnitude and frequency of forces in geomorphic processes. *The Journal of Geology*, *68*(1), 54–74.
Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025–3033.
Yousefi, S., Mirzaee, S., Keesstra, S., Surian, N., Pourghasemi, H. R., Zakizadeh, H. R., & Tabibian, S. (2018). Effects of an extreme flood on river morphology (case study: Karoon River, Iran). *Geomorphology*, *304*, 30–39.
Zou, Z., Xiao, X., Dong, J., Qin, Y., Doughty, R. B., Menarguez, M. A., . . . Wang, J. (2018). Divergent trends of open-surface water body area in the contiguous United States from 1984 to 2016. *Proceedings of the National Academy of Sciences*, *115*(15), 3810–3815.

Supplementary Material: Quantifying geomorphically effective floods using satellite observations of river mobility

A. S. Leenman[1], L. J. Slater[1], S. J. Dadson[1,2], M. Wortmann[1,3] and R. Boothroyd[4]

[1]School of Geography and the Environment, University of Oxford, UK
[2]UK Centre for Ecology and Hydrology
[3]European Centre for Medium-Range Weather Forecasts
[4]School of Geographical and Earth Sciences, University of Glasgow

Corresponding author: Anya Leenman, `anya.leenman@chch.ox.ac.uk`

Table 1. The r^2 values for the relationships in Figure 2 (main manuscript), for each individual country.'Av. mag.' is the mean peak height in cm (measured above the gauge's mean daily stage) across all flood events in that country.'Av. total' is the mean (across all floods in a country) of the total water level (in cm) exceeding mean stage.'Av. dur.' is the mean flood duration (in days) in that country. 'Av. widening' is the mean reach-averaged widening (in m) across all floods and sites in that country.

Country	r^2, Peak height above mean stage (cm)	r^2, Stage above mean, summed (cm)	r^2, Duration (days)	N. floods	N. gauge sites	Av. mag.	Av. total	Av. dur.	Av. widening
Brazil	-0.018	0.219	0.293	47	10	230	4100	53	1.90
Colombia	0.110	0.321	0.442	87	22	160	5300	57	7.30
New Zealand	0.796	0.278	0.171	11	1	190	1100	29	0.92
Russia	0.348	0.082	0.081	15	8	150	3700	71	3.20

Table 2. The variables used in the random forest model. Column 1 shows how each variable contributed to reducing MSE. The final column shows the rank assigned to each variable by the random forest regression.

% Decr. in MSE	Variables	Rank
22.18	Estimated sediment transport	1
16.77	Channel width	2
15.65	Duration	3
8.54	Total stage exceeding mean	4
7.61	Peak height above mean daily stage	5

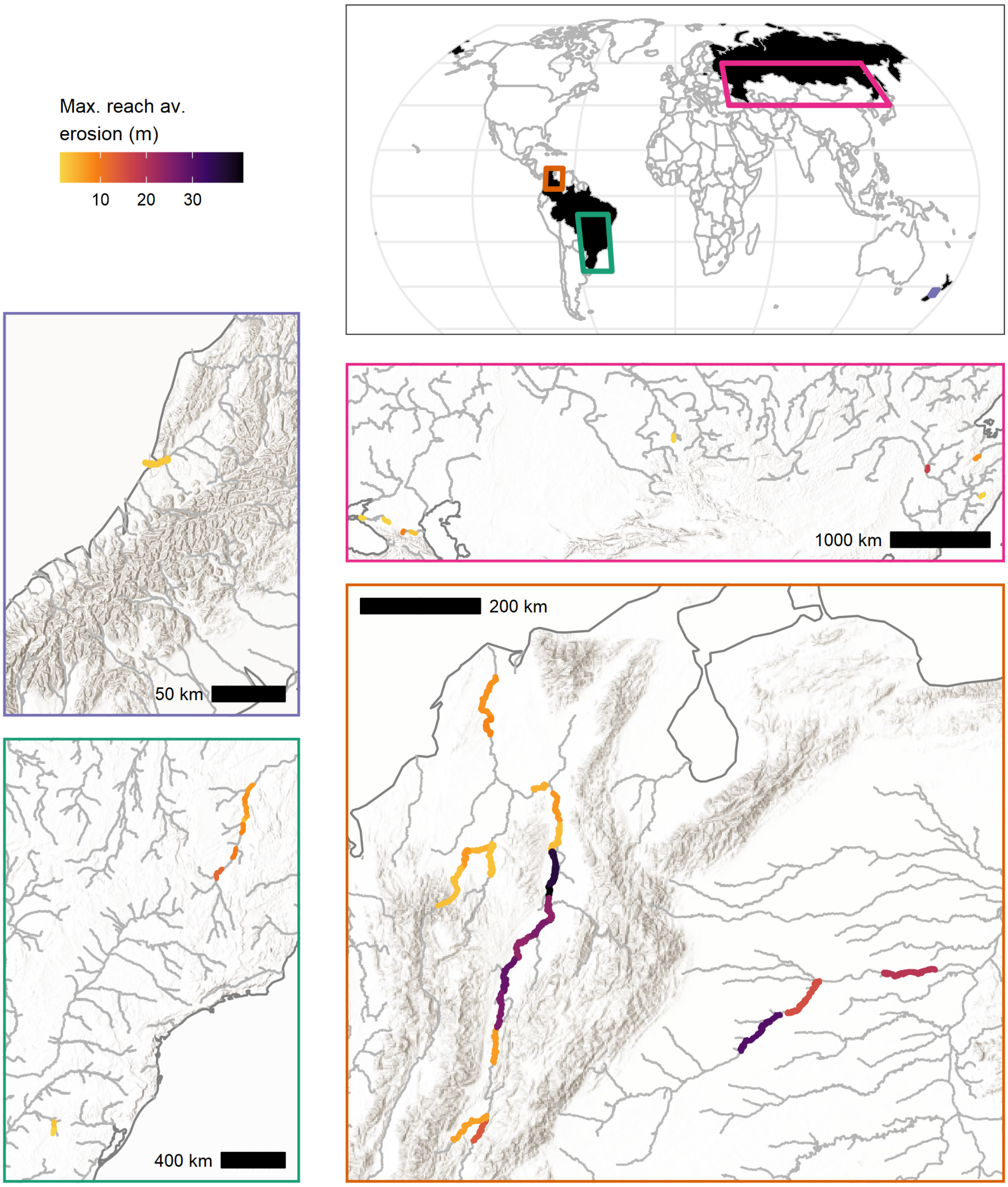

Figure 1. Map showing the areas of interest (AOIs) associated with each gauge. Colors show the magnitude of reach-averaged widening (in metres) during the most effective flood at each site.

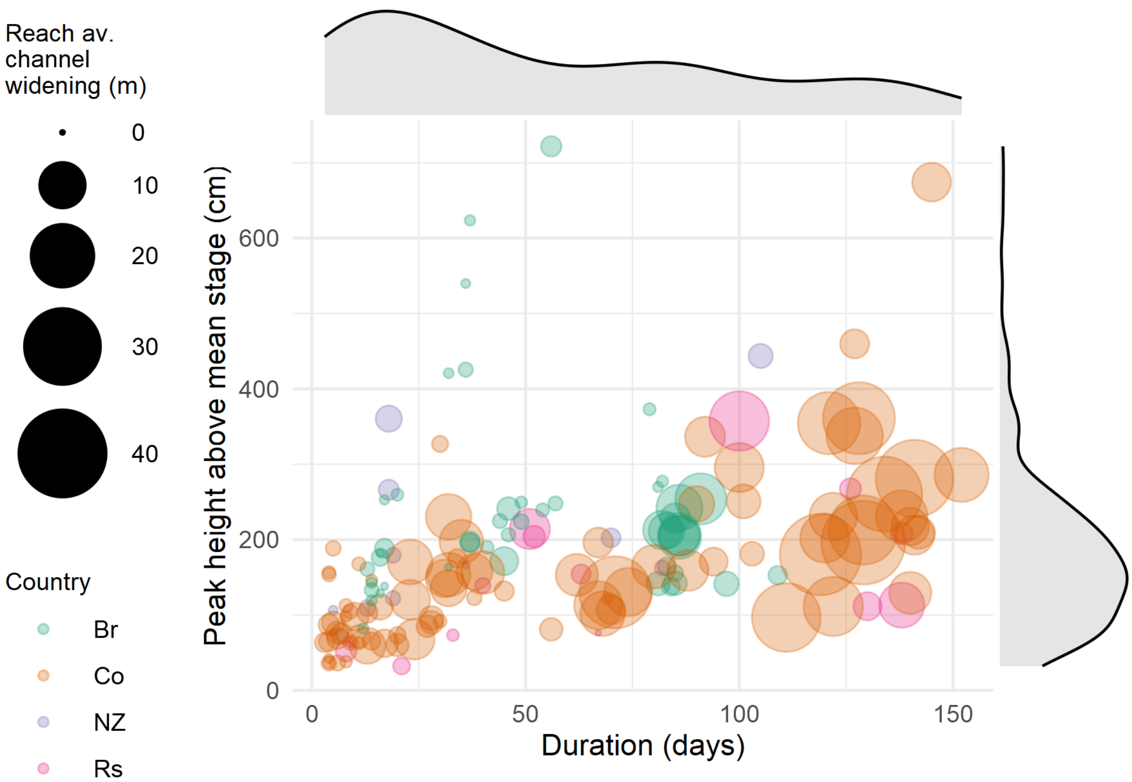

Figure 2. Duration, magnitude (peak height) and geomorphic effectiveness (reach-averaged erosion) for each flood event in our dataset. Each point is one flood event; colours correspond to countries and size corresponds to geomorphic effectiveness of each flood.

1 Sediment transport capacity

We estimated sediment transport capacity based on the stage and slope data available to us. Sediment transport equations often predict transport as the $\frac{3}{2}$ power of some flow property — often that which exceeds a threshold value at which sediment of a given size can be entrained (Church, 2010). Often that flow property is the dimensionless shear stress τ_*, but we have no data on grain size with which to calculate this. Instead, we approximate the dimensional boundary shear stress τ, which scales with the depth-slope product dS. We have no data on flow depth and approximate it with flow stage h instead; our estimates of channel slope S are calculated along the area of interest polygon for each gauging site using elevation data from the MERIT DEM (Yamazaki et al., 2017).

We therefore estimate unit sediment transport q_s as a function of stage and slope. We do not have data on the threshold for motion in our study sites, so we assume that the threshold is 25% of the difference between minimum and maximum stage in each gauge record, during the ˜7 year period for which we have satellite data. While arbitrary, this value of 25% is based on a literature search for reported values of the onset of transport as a percentage of peak discharge, and it also performed better than other thresholds we tried.

We thus estimate a flood's cumulative transport as a function of changes in stage:

$$q_s = \sum_{1}^{n} ((h - h_{r25})S)^{\frac{3}{2}} \tag{1}$$

where n is the total number of days in the flood, h is the stage value for each day, $r25$ is the stage that is 25% of the difference between the minimum and maximum stage during the satellite record, and S refers to the channel slope. We performed this calculation for each day in a flood and summed across the entire event.

Finally, we multiply q_s by channel width to estimate the channel-integrated (total) sediment transport Q_s. While q_s did not scale with erosion as well as the flood duration or summed hydrograph did, the estimated Q_s scaled rather closely ($r^2 = 0.63$) with each flood's geomorphic effectiveness (Figure S3). It is Q_s that we used in our random forest model.

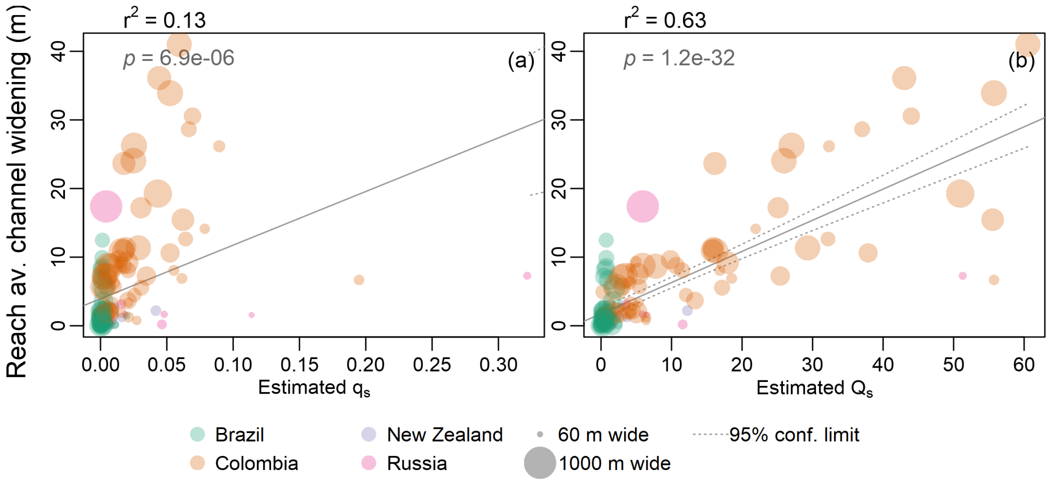

Figure 3. Linear regression of flood-driven erosion (reach-averaged) against our estimates of the cumulative sediment transport capacity of each hydrograph: (a) unit transport q_s (b) integrated (total) transport Q_s.

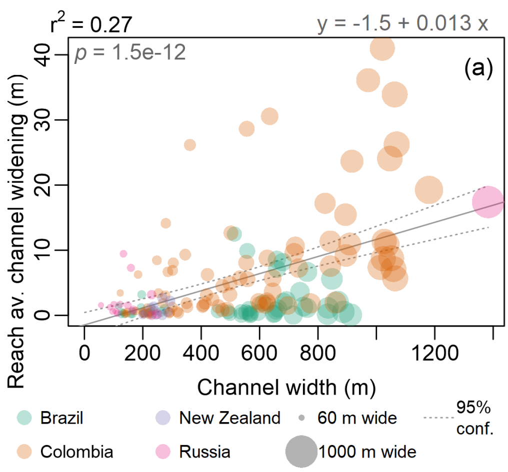

Figure 4. Linear regression of flood-driven erosion (reach-averaged) against mean channel width prior to each flood.

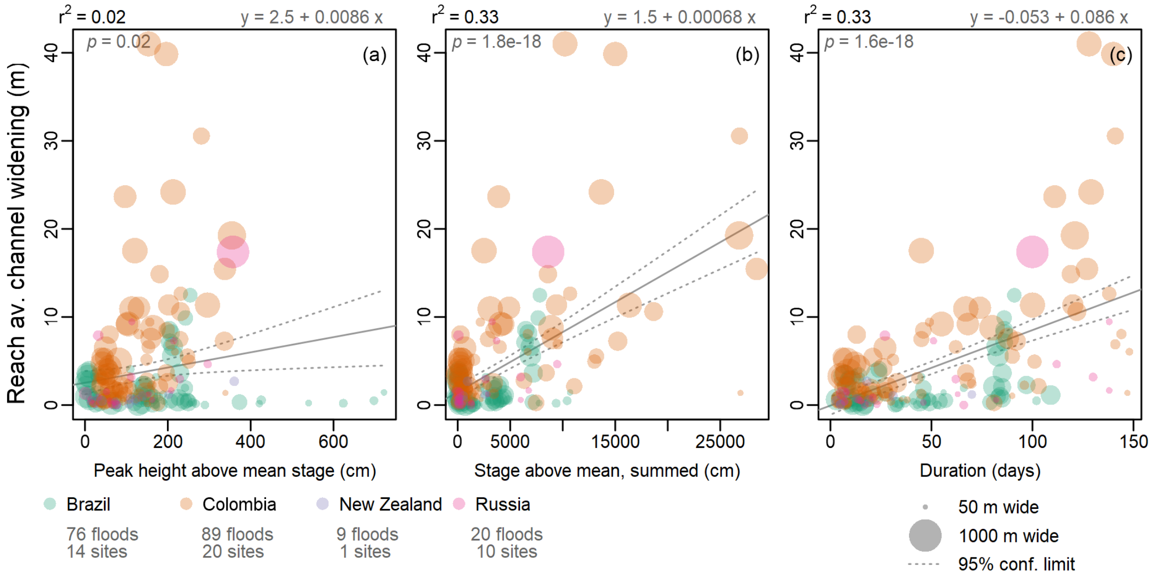

Figure 5. The results in Figure 2 (main manuscript) when the flood-delineation threshold is lowered to the 70th percentile of stage.

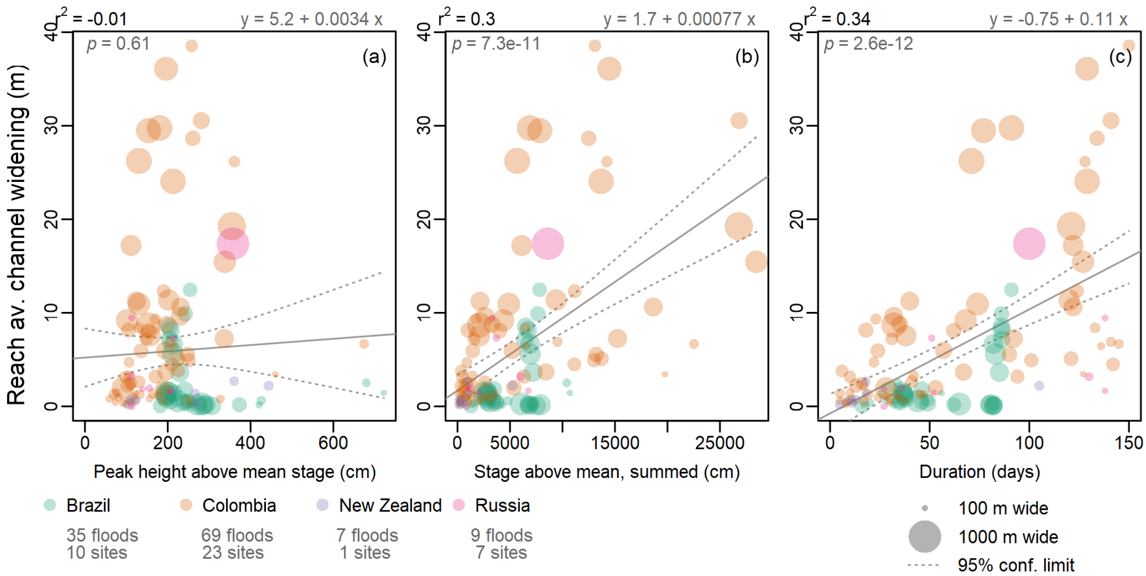

Figure 6. The results in Figure 2 (main manuscript) when the flood-delineation threshold is raised to the 90th percentile of stage.

References

Church, M. (2010). Gravel-Bed Rivers. In T. Burt & R. Allison (Eds.), *Sediment Cascades: An Integrated Approach* (p. 241-269). Chichester: Wiley-Blackwell.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., ... Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, *44*(11), 5844–5853. Retrieved from `https://developers.google.com/earth-engine/datasets/catalog/MERIT_DEM_v1_0_3#description` (Data accessed via Google Earth Engine)