Article

Rules of river avulsion change downstream

https://doi.org/10.1038/s41586-024-07964-2

Received: 24 November 2023

Accepted: 20 August 2024

Published online: 18 September 2024

Check for updates

James H. Gearon[1]✉, Harrison K. Martin[1,2], Clarke DeLisle[1], Eric A. Barefoot[1,3], David Mohrig[4], Chris Paola[5] & Douglas A. Edmonds[1]

Avulsing rivers create new pathways on the floodplain and the associated flooding can profoundly affect society[1–4]. River avulsions are thought to occur when the water column becomes perched above the floodplain[5] or when the slope down the flanks of the channel provides a steeper descent than the existing river channel[6,7]. We test these classical ideas by quantifying the topography around avulsing rivers and show that these mechanisms, historically invoked separately, work together. Near coasts, rivers avulse when the slope away from the channel is steeper, not because they are perched. The opposite is true near mountain fronts; on fans, the alternative paths are similarly steep to the downstream path, so rivers avulse when they are perched above the surrounding landscape. We reconcile these findings and present a new theoretical framework that identifies which rivers are vulnerable to avulsion and predicts the path of an avulsing river. These first-order rules of avulsion suggest that avulsion risks are underestimated in many coastal environments[8] and that probabilistic predictions of avulsion pathfinding can efficiently map hazards with minimal information. Applying these principles for risk assessment could particularly benefit the Global South, which is disproportionately affected by avulsions.

When a river breaches its banks and shifts course on the floodplain, the resulting avulsion not only reshapes the surrounding landscape but also inflicts devastating floods on nearby communities. Unlike floods caused by precipitation, avulsions strike with little warning, after decades to millennia of accumulating sediment alters the river configuration enough to destabilize the channel[9]. The long intervals between avulsions[10], coupled with their catastrophic nature[1–3], pose substantial challenges in managing avulsion hazards. As climate change intensifies, some rivers may avulse with increasing frequency[8] or in different locations[9], trends that could disproportionately affect the Global South. At present, we cannot predict avulsions because they are incompletely understood; metrics for avulsion likelihood are heuristic and controls on path selection across the floodplain are undefined[5,10,11]. To address these gaps and resolve the fundamental rules of river avulsion, we base our approach on observations.

River avulsions are thought to occur when sediment deposition on the river bed causes superelevation or gradient advantage. Superelevation ($\beta = H_{AR}/H_M$) is the ratio of alluvial ridge height (H_{AR}) to bankfull depth (H_M). Gradient advantage ($\gamma = S_{AR}/S_M$) is the ratio of alluvial ridge slope (S_{AR}) to channel slope (S_M). When the river bed aggrades to the floodplain level, the channel becomes superelevated ($\beta \approx 1$; Fig. 1a). Avulsions become likely as the water surface is perched above the floodplain[10–12]. When the topographic slope perpendicular to the river exceeds the channel slope, the alternative river path has a gradient advantage ($\gamma > 1$); models suggest that $\gamma > 3$–10 makes a river prone to avulsion-inducing overbank flows[6,7,13,14]. Although these avulsion rules are widely used[15], they are largely based on physical intuition and their relative importance is debated[5–7,13,14]. In part, this is because accurately measuring β and γ on global digital elevation models (DEMs) is challenging; topographic artefacts, such as vegetation and human modifications, can obscure subtle topography around riverbanks[16]. Consequently, for a small number of avulsed rivers, β varies from 0.2 to 1.4 (refs. 5,12,17), but it is unknown whether these values are representative.

Here we investigate the rules for river avulsion by measuring the topography of recently avulsed rivers and the factors that influence avulsion pathways. Using a global avulsion dataset (Fig. 1b), ICESat-2, and a bare-earth DEM, we measure β and γ on 58 rivers. Furthermore, for a subset of ten rivers with pre-avulsion topography, we simulate avulsion pathfinding with a random walk model. Our measurements and pathfinding model explain how β and γ govern avulsion mechanics and how their relative importance varies from source to sink.

Avulsion drivers from source to sink

Previous work has shown that avulsions are common near mountain fronts and coastlines because sediment accumulates there[18]. We compiled examples of 174 globally distributed river avulsions throughout the source-to-sink system, observed in Landsat satellite images from 1984 CE to the present (Fig. 1b and Supplementary Table 1). Avulsion is defined as the channel relocating to a new floodplain position outside the active channel belt. We measured the distance (s) along the channel centreline from the mountain front to each avulsion position and divided it by the total length (s_t) of the alluvial portion of the source-to-sink system, giving each avulsion a normalized position from 0 to 1 (X_N; Fig. 2a). s_t was measured from the mountain front to the nearest shoreline, such as an ocean, sea or inland lake. For a subset of avulsions ($n = 58$) with sufficient data quality, we generated cross-sections of

[1]Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, USA. [2]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA. [3]Department of Earth and Planetary Sciences, University of California, Riverside, Riverside, CA, USA. [4]Department of Earth and Planetary Sciences, Jackson School of Geosciences, University of Texas at Austin, Austin, TX, USA. [5]Department of Earth and Environmental Sciences, University of Minnesota, Minneapolis, MN, USA. ✉e-mail: jake.gearon@gmail.com

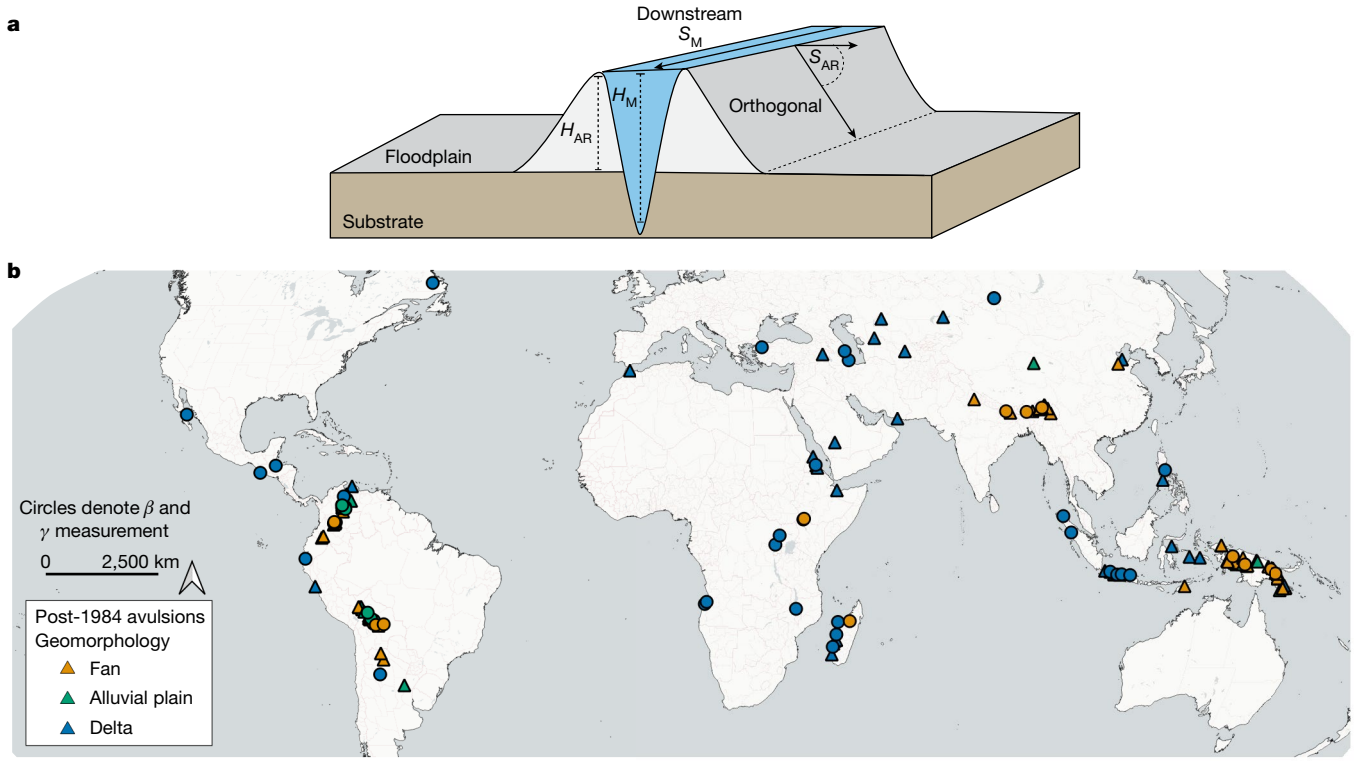


Fig. 1 | Global distribution of river avulsions. a, Schematic of river topography. H_{AR} is the vertical distance from the ridge crest to the floodplain surface. S_{AR} is measured perpendicular to flow direction from the ridge crest to the floodplain and S_M is measured along the levee crest following the channel. **b**, Avulsions (since 1984 CE and coloured by geomorphic landform) used in this study ($n = 174$). Circles show avulsions with measurements of β and γ ($n = 58$).

river topography near the avulsion site. To make cross-sections, we used ICESat-2, a spaceborne lidar platform that can penetrate vegetation canopies[19] (Extended Data Figs. 1 and 2), and FABDEM[20], a bare-earth version of COPDEM30. We validated 38 of the FABDEM measurements of H_{AR} and S_{AR} with ICESat-2 to ensure agreement between data sources (Extended Data Fig. 2). We estimated H_M with a new XGBoost regressor trained on a dataset[21] of hydraulic geometric relationships ($n = 4{,}006$) and measured all other quantities in Fig. 1a with FABDEM or ICESat-2 (see Methods). Using these data, we measured β and γ on the main channel as close to the avulsion site as possible, assuming that the observed topographic state represents the pre-avulsion conditions (Supplementary Table 2).

The full dataset shows that avulsions occur throughout the source-to-sink system but also reveals that the distribution is bimodal. Avulsions are about three times more common at mountain fronts and shorelines than in between the two (Fig. 2a,b). Of the 174 river avulsions we catalogued, 74% ($n = 129$) occur within 15% of the normalized distance to the mountain front ($n = 81$) or shoreline ($n = 48$). For the avulsion sites with sufficient data quality ($n = 58$), the average values of β and γ closely matched predicted theoretical ranges (Fig. 2c,d).

We split our data into three geomorphic groups: (1) rivers on alluvial or fluvial fans ($n = 19$) exit the mountain front with a distributive planform; (2) alluvial plain rivers ($n = 8$) show no radial distribution and have crevasse splays, flood basins or anastomosis; (3) deltaic rivers ($n = 31$) have distributary networks entering a local water body. We then normalized the avulsion position, X_N, by S_M:

$$\mathrm{SPR}^* = X_N / S_M \tag{1}$$

Because S_M scales with upstream basin area[22], this slope–position ratio (SPR*) enables comparisons of avulsions at similar X_N across basins of different sizes and slopes, effectively ordering them from steeply sloping proximal rivers to gently sloping distal rivers (Fig. 2g,h).

Approximately 86% of the avulsions occurred on either alluvial or fluvial fans near the mountain front or on deltaic fans near the shoreline, whereas only 14% occurred on alluvial floodplains in the middle portions of river basins. For avulsed rivers, the values of β and γ vary across geomorphic landform (Fig. 2e,f) and from source to sink (Fig. 2g,h). Near the source, $\beta \geq 1$ consistent with classical expectations for avulsing rivers, before decreasing to <0.1 in more distal settings (Fig. 2g). Notably, as β decreases moving from source to sink, γ increases substantially from 1 to about 32 (Fig. 2h).

Uniting β and γ

β and γ both indicate the likelihood of avulsion, yet on avulsed rivers, we find that they are inversely related (Fig. 3a). Using only one of these metrics to estimate avulsion likelihood is standard practice[7,8,10,12–15,17,18] but leaves half of the problem unconstrained. For example, previous studies[7,23] have interpreted avulsive rivers with high γ and low β as contradictory. Instead, β and γ are fundamentally related when the conditions for avulsion are recast as a shear stress advantage. River avulsions are more likely when water flowing down the alluvial ridge flank exerts greater shear stress (τ_{AR}) than the shear stress on the bed of the main channel (τ_M) by some multiple Λ. A larger τ_{AR} that exceeds the threshold for sediment entrainment should erode the alluvial ridge and the river will avulse if a suitable path exists. The temporal and spatial scales of deposition that set up avulsion are long[10] and thus τ_{AR} and τ_M can be approximated with steady, uniform flow in a rectangular channel. In this framework, we do not account for the shorter-timescale events that trigger avulsions. Then, for an avulsion to occur:

$$\tau_{AR} \geq \Lambda \tau_M \tag{2}$$

$$\rho g H_{AR} S_{AR} \geq \Lambda \rho g H_M S_M \tag{3}$$

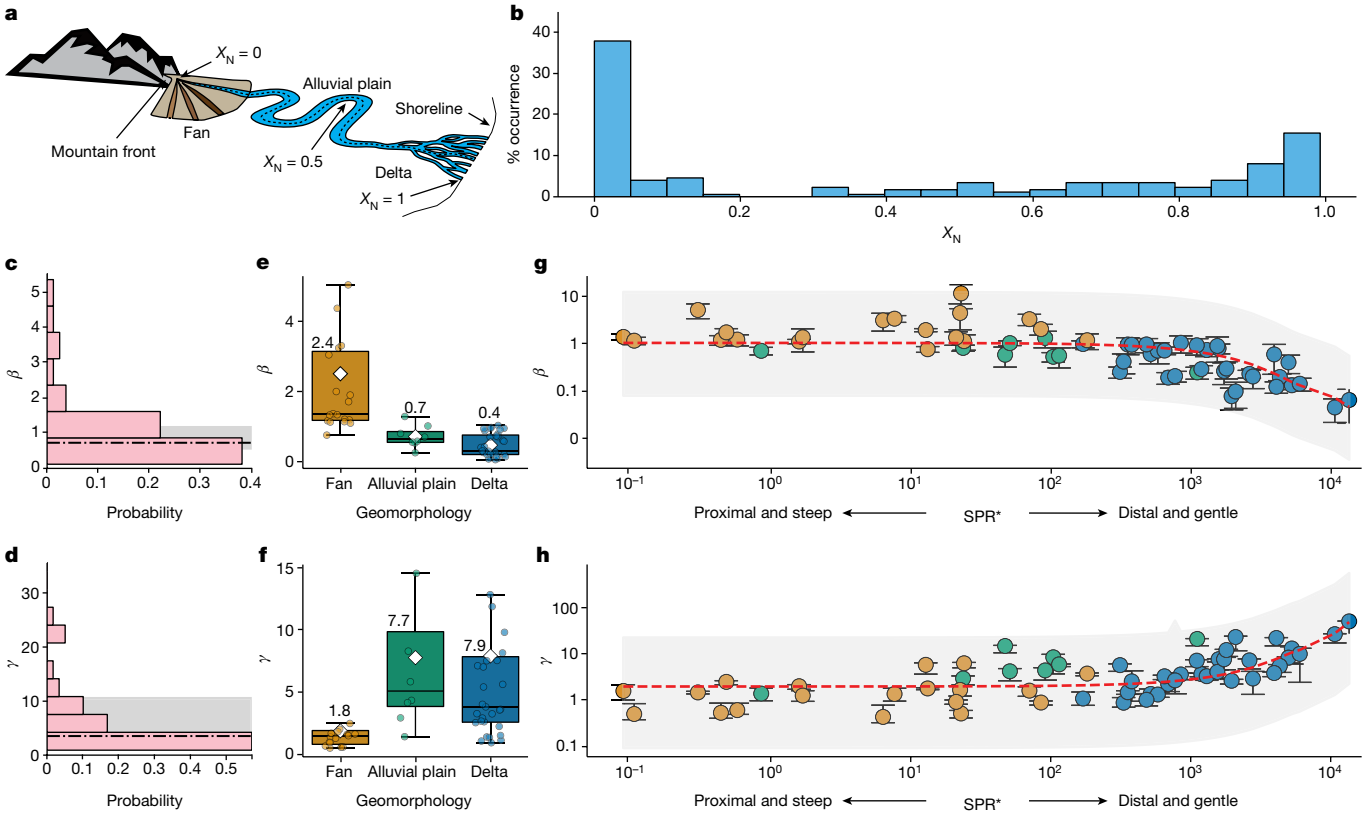


Fig. 2 | Topography of recently avulsed rivers changes from source to sink. **a**, Schematic of X_N (dashed line) within the alluvial source-to-sink system. **b**, Source-to-sink distribution of all avulsions ($n = 174$). **c**,**d**, Histograms of β and γ ($n = 58$). The dash-dot line represents the mean and grey shading shows threshold ranges from ref. 5 in **c** and from ref. 6 in **d**. **e**,**f**, Box-and-whisker plots for each geomorphic landform. White diamonds and numbers correspond to mean values. Whiskers show ±1.5 times the interquartile range. Eight outliers are omitted from the figure for better visualization. **g**,**h**, As SPR* increases, β decreases and γ shows an order of magnitude increase. Points are coloured by geomorphic landform. Shaded regions represent the LOWESS (locally weighted scatterplot smoothing) fit confidence interval. Error bars represent s.d. in **g** and s.e.m. in **h** (see Methods).

in which ρ is fluid density and g is gravitational acceleration. H_{AR} defines the flow depth that could be released onto the floodplain at bankfull flow; this assumption is a useful simplification that could be relaxed in future treatments (for example, ref. 6). We measure S_{AR} orthogonal to the main channel centreline, consistent with crevasse splay take-off angles[24]. The framework provided by equation (3) is simple by design so that it captures the first-order causes of avulsion and is generalizable to a variety of settings.

Rearranging and incorporating our definitions of β and γ, the avulsion threshold can be expressed as:

$$\beta\gamma \geq \Lambda \quad (4)$$

When formulated as a shear stress advantage, the avulsion threshold depends on the product of superelevation and gradient advantage. This implies that, for a given value of Λ, β and γ are inversely related, which is consistent with our measured field data but opposite to classical ideas (Fig. 3). Previously published ranges of β (0.2–1.4) and γ (3–10) (refs. 5,6,11,12) provide bounds on an initial estimate of Λ of 0.6–14. Our measurements show that Λ ranges from 0.2 to 11 with a median of 2.1, suggesting that an avulsion occurs when the shear stress down the alluvial ridge flank is about twice that in the main channel (Fig. 3b). The range of Λ is relatively narrow (Fig. 3b), even though our framework does not address variations in vegetation, grain size or floodplain conditions such as pre-existing channels[25] or events that trigger avulsion. These uncontrolled factors may account for the variability in Λ.

Avulsed rivers on alluvial or fluvial fans have high β and low γ, whereas the opposite is true for deltas (Figs. 2e–h and 3a), reflecting different avulsion rules near the source compared with the sink. This behaviour is predicted by the Λ framework as avulsion initiation depends on the relative growth rate of both β and γ. As an alluvial ridge grows on fans, β increases faster than γ because channel depths near the source are relatively shallow and surrounding slopes are relatively steep, making superelevation easier to achieve. The opposite is true on deltas; it is easier to achieve gradient advantage because surrounding slopes are relatively flat and deltaic channels are relatively deep. An alternative explanation suggests that low β in deltas results from larger variations in river stage height, which creates more frequent trigger events that lead to avulsion at lower than expected β (ref. 17). In the absence of stage height data, we test this by using the discharge range normalized by the mean discharge for each avulsed river, which also describes variability around the central tendency. We find no correlation between β and this measure of variance ($R^2 = 0.05$, $p = 0.1$; Extended Data Fig. 4).

Modelling the path taken by avulsions

β and γ both influence when avulsions will occur, yet determining the path avulsions will take across the floodplain remains challenging. Roughly 75% of avulsions have $\Lambda > 1$, suggesting that they initiate when there is a shear stress advantage relative to the main channel. To test if path selection is governed by this shear stress advantage, we introduce a random walk algorithm that incorporates the basic properties of a gravity-driven flow by selecting a path based on the slope differences to the surrounding cells (S) and the inertial tendency to resist changes in direction from the previous step (I). We use an algorithm based on the softmax function and cosine similarity to compute transition

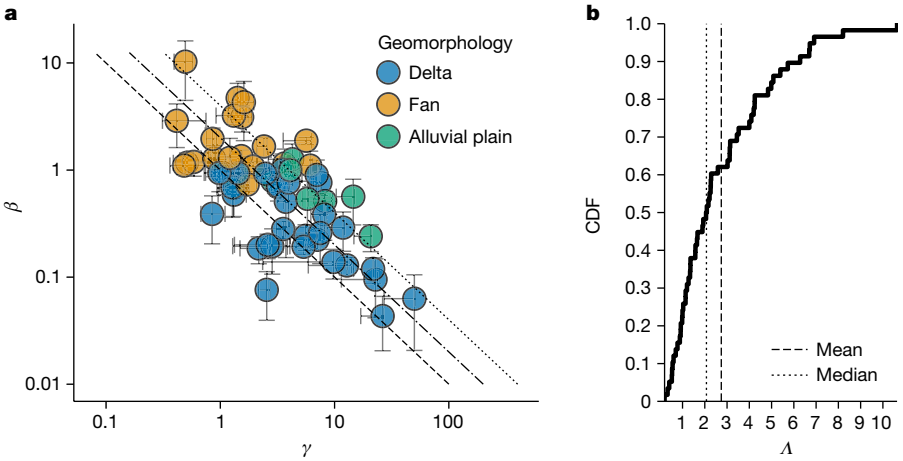


Fig. 3 | Inverse relationship between β and γ for avulsed rivers. a, Lines from equation (4) show the predicted relationship for different values of the avulsion threshold Λ. Points are coloured by geomorphology ($n = 58$), with error bars indicating uncertainty. **b**, Cumulative density function (CDF) of Λ values with mean 2.7 and median 2.1. Standard deviation is 2.2.

probabilities P to neighbouring pixels that depend on exponents for slope (ϕ) and inertia (ψ), and a blending parameter (ω), such that:

$$P((x,y) \to (i,j)) = \mathrm{softmax}_{i,j}(\omega\phi S_{i,j} + (1-\omega)\psi I_{i,j}) \quad (5)$$

in which the softmax function exponentiates the sum of ϕS and ψI weighted by ω (see Methods). We selected the softmax function because it generates a valid probability distribution over a vector of discrete choices. Because avulsions modify their paths, we found six completed avulsions from our original dataset (Fig. 4a–c,g,h and Extended Data Fig. 6a) that occurred after the collection period of a global DEM, including SRTM for unvegetated regions, such that the measured topography represents the pre-avulsion conditions. To increase the number of test cases, we also included four continuing avulsions (Fig. 4d–f and

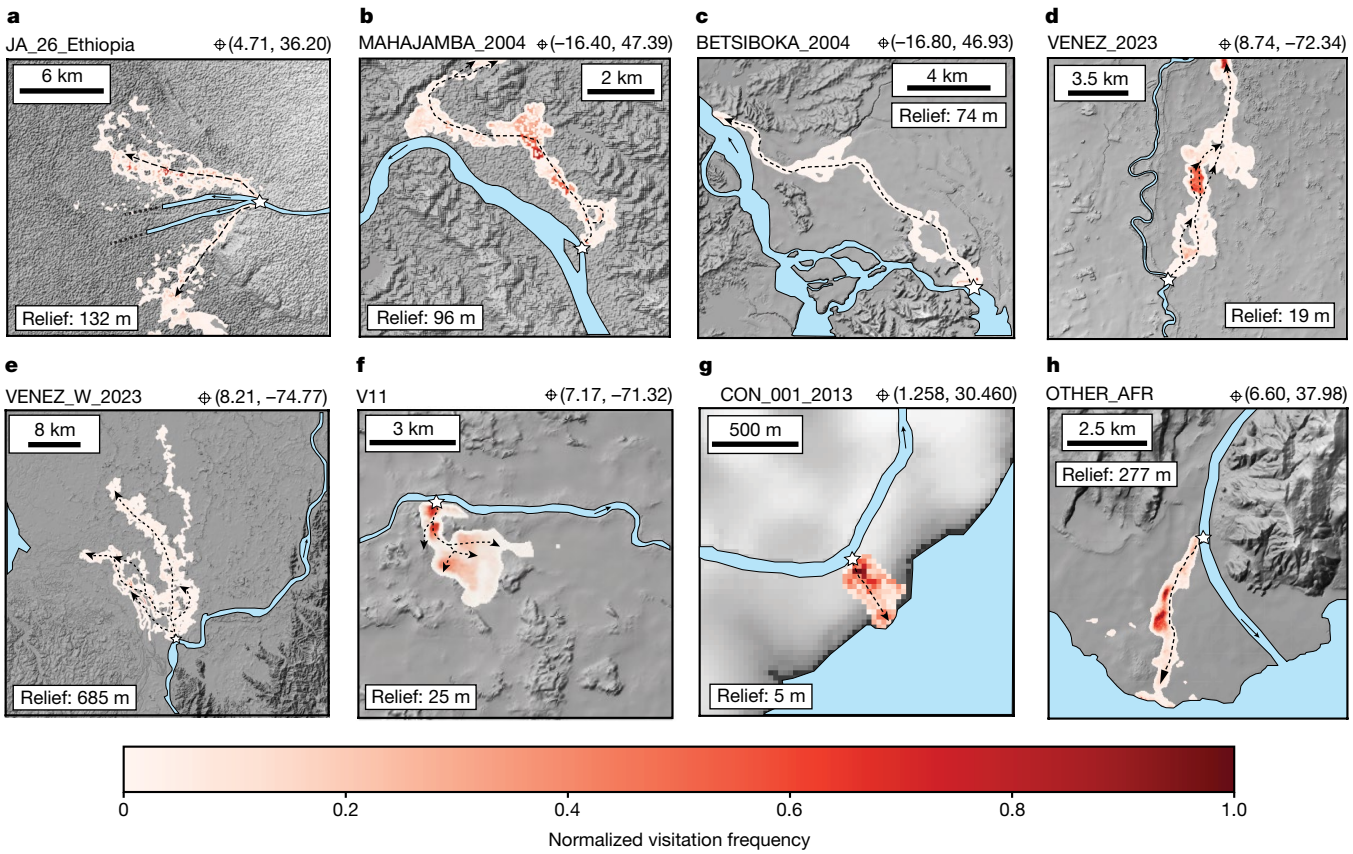


Fig. 4 | Avulsion pathfinding experiments. Red colours represent the normalized probability of occupation after 100 random walk trials and only cells >0.075 are shown. Dashed black lines show the observed avulsion paths. **a**–**c**, Completed fan avulsions. Panel **a** contains two events from 2002 (northern channel) and 2007 (southern channel). **d**–**f**, Continuing alluvial plain avulsions. **g**,**h**, Completed delta avulsions. Light-blue mask shows the parent channel and standing water. Hillshade maps are derived from SRTM[16] (**a**,**b**) and FABDEM[20] (**c**–**h**). Coordinates specify avulsion location. The top of each panel is north. Recent imagery may be different (for example, **g**) owing to continuing changes.

Extended Data Fig. 6b). We conducted 100 random walk model trials starting from the riverbank of the known avulsion location and using $\omega = 0.5$, $\phi = 10$ and $\psi = 1$ (see Methods for details). In all cases, the predicted avulsion path overlaps nearly the entire observed path (Fig. 4 and Extended Data Fig. 6). The model works on all landforms, including proximal fans (Fig. 4a–c), alluvial plains (Fig. 4d–f and Extended Data Fig. 6b) and deltas (Fig. 4g,h and Extended Data Fig. 6a) and in cases in which the pathway is inherited (Fig. 4c,h) or constructed by the advancing river (Fig. 4a,b,d–g and Extended Data Fig. 6a,b). In two cases, the model correctly identifies several pathways from the same event (Fig. 4a,e). The success of these pathfinding experiments implies that avulsions tend to seek steeper paths that minimize directional changes on the floodplain. If these tuning parameters are universal, as our initial results suggest, then equation (5) could be used to predict future avulsion pathways.

Implications for hazard management

These new rules can help prepare for avulsion-related hazards. Previous work commonly estimates the time between avulsions as the time to aggrade a channel depth and achieve $\beta = 0.5$–1.0 (refs. 5–15,17,18,26), but β only accounts for half of Λ and depends on the source-to-sink position. In deltaic environments, avulsion might occur sooner than estimated, as roughly 60% of deltas in our dataset have $\beta < 0.5$ (Fig. 2g,h), which is notable because deltas are heavily populated and are evolving rapidly in response to climate change[27,28].

Our new rules are simple and could be used to generate catchment-scale maps of Λ, a first step in identifying avulsion hazards. Improved global bare-earth DEMs will enable mapping of potential avulsion corridors using topographic data and our pathfinding algorithm. Complete risk assessments will require expanding our framework to account for time-dependent flooding, bank erosion or river engineering.

Avulsions pose invisible flood hazards unaddressed by current flood models focused on predicting surging water levels. In fact, avulsion flooding does not always follow unprecedented rain[1], making traditional warning signs unreliable[29]. Flood models cannot identify avulsion sites because they do not include the erosion and deposition that precede avulsion[30–32]. This limitation perpetuates disparities in flood risk management, especially in the Global South (Fig. 1b and Extended Data Fig. 5), in which avulsions are more frequent, possibly because of geologic and climatic factors or differences in river engineering practices.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07964-2.

1. Syvitski, J. & Brakenridge, R. Causation and avoidance of catastrophic flooding along the Indus River, Pakistan. *GSA Today* **23**, 4–10 (2013).
2. Macklin, M. G. & Lewin, J. The rivers of civilization. *Quat. Sci. Rev.* **114**, 228–244 (2015).
3. Sinha, R. The great avulsion of Kosi on 18 August 2008. *Curr. Sci.* **97**, 429–433 (2009).
4. Todd, O. J. & Eliassen, S. The Yellow River problem. *Trans. Am. Soc. Civil Eng.* **105**, 346–416 (1940).
5. Mohrig, D., Heller, P. L., Paola, C. & Lyons, W. J. Interpreting avulsion process from ancient alluvial sequences: Guadalope-Matarranya system (northern Spain) and Wasatch Formation (western Colorado). *GSA Bull.* **112**, 1787–1803 (2000).
6. Slingerland, R. & Smith, N. D. Necessary conditions for a meandering-river avulsion. *Geology* **26**, 435–438 (1998).
7. Tornqvist, T. E. & Bridge, J. S. Spatial variation of overbank aggradation rate and its influence on avulsion frequency. *Sedimentology* **49**, 891–905 (2002).
8. Chadwick, A. J., Lamb, M. P. & Ganti, V. Accelerated river avulsion frequency on lowland deltas due to sea-level rise. *Proc. Natl Acad. Sci.* **117**, 17584–17590 (2020).
9. Slingerland, R. & Smith, N. D. River avulsions and their deposits. *Annu. Rev. Earth Planet. Sci.* **32**, 257–285 (2004).
10. Jerolmack, D. J. & Mohrig, D. Conditions for branching in depositional rivers. *Geology* **35**, 463–466 (2007).
11. Ganti, V., Chadwick, A. J., Hassenruck-Gudipati, H. J. & Lamb, M. P. Avulsion cycles and their stratigraphic signature on an experimental backwater-controlled delta. *J. Geophys. Res. Earth Surf.* **121**, 1651–1675 (2016).
12. Ganti, V., Lamb, M. P. & Chadwick, A. J. Autogenic erosional surfaces in fluvio-deltaic stratigraphy from floods, avulsions, and backwater hydrodynamics. *J. Sediment. Res.* **89**, 815–832 (2019).
13. Allen, J. R. L. Studies in fluviatile sedimentation: an exploratory quantitative model for the architecture of avulsion-controlled alluvial suites. *Sediment. Geol.* **21**, 129–147 (1978).
14. Bridge, J. S. & Leeder, M. R. A simulation model of alluvial stratigraphy. *Sedimentology* **26**, 617–644 (1979).
15. Hajek, E. A. & Wolinsky, M. A. Simplified process modeling of river avulsion and alluvial architecture: connecting models and field data. *Sediment. Geol.* **257–260**, 1–30 (2012).
16. Farr, T. G. et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **45**, RG2004 (2007).
17. Ganti, V., Chu, Z., Lamb, M. P., Nittrouer, J. A. & Parker, G. Testing morphodynamic controls on the location and frequency of river avulsions on fans versus deltas: Huanghe (Yellow River), China. *Geophys. Res. Lett.* **41**, 7882–7890 (2014).
18. Brooke, S. et al. Where rivers jump course. *Science* **376**, 987–990 (2022).
19. Neuenschwander, A., Guenther, E., White, J. C., Duncanson, L. & Montesano, P. Validation of ICESat-2 terrain and canopy heights in boreal forests. *Remote Sens. Environ.* **251**, 112110 (2020).
20. Hawker, L. et al. A 30 m global map of elevation with forests and buildings removed. *Environ. Res. Lett.* **17**, 024016 (2022).
21. Deal, E. Downstream hydraulic geometry data compilation. *HydroShare* http://www.hydroshare.org/resource/0629ffb81fdb40aa9e6be42cc11918ca (2021).
22. Anderson, R. S. & Anderson, S. P. *Geomorphology: The Mechanics and Chemistry of Landscapes* (Cambridge Univ. Press, 2010).
23. Aslan, A., Autin, W. J. & Blum, M. D. Causes of river avulsion: insights from the late Holocene avulsion history of the Mississippi River, U.S.A. *J. Sediment. Res.* **75**, 650–664 (2005).
24. Rahman, M. M., Howell, J. A. & MacDonald, D. I. M. Quantitative analysis of crevasse-splay systems from modern fluvial settings. *J. Sediment. Res.* **92**, 751–774 (2022).
25. Martin, H. K. & Edmonds, D. A. The push and pull of abandoned channels: how floodplain processes and healing affect avulsion dynamics and alluvial landscape evolution in foreland basins. *Earth Surf. Dyn.* **10**, 555–579 (2022).
26. Jobe, Z. R. et al. Comparing aggradation, superelevation, and avulsion frequency of submarine and fluvial channels. *Front. Earth Sci.* **8**, 53 (2020).
27. Nienhuis, J. H. et al. Global-scale human impact on delta morphology has led to net land area gain. *Nature* **577**, 514–518 (2020).
28. Edmonds, D. A., Caldwell, R. L., Brondizio, E. S. & Siani, S. M. O. Coastal flooding will disproportionately impact people on river deltas. *Nat. Commun.* **11**, 4741 (2020).
29. Best, J., Ashmore, P. & Darby, S. E. Beyond just floodwater. *Nat. Sustain.* **5**, 811–813 (2022).
30. Hirabayashi, Y. et al. Global flood risk under climate change. *Nat. Clim. Change* **3**, 816–821 (2013).
31. Alifu, H., Hirabayashi, Y., Imada, Y. & Shiogama, H. Enhancement of river flooding due to global warming. *Sci. Rep.* **12**, 20687 (2022).
32. Bates, P. Fundamental limits to flood inundation modelling. *Nat. Water* **1**, 566–567 (2023).

Methods

Avulsion dataset

We assembled a dataset of 174 avulsion locations (Fig. 1b). We ensured the quality and reliability of our data by selecting only rivers with documented avulsion events in the satellite era (1984 CE to the present). Our dataset combines observations from three previous studies[18,33,34] and is supplemented with further observations that we collected. We define avulsion as the relocation of a channel to a new position on the floodplain outside the active channel belt that has reconnected to an existing channel downstream or intersected the shoreline. We do not include small changes that occur within the active channel belt (for example, chute cut-offs, in-channel bifurcation and bar development or crevasse development). Both annexational and progradational styles of avulsion are included. Similarly, both local and regional avulsions are included, as well as partial and full avulsions. An avulsion does not have to be completed within the observation window to be included. We also avoided including any avulsions for which it was reasonably clear that human activities may have rerouted the river.

Topographic datasets

Most global DEMs have inaccurately captured topography surrounding rivers. This is especially true in remote areas in which high-resolution topographic data or bare-earth models are unavailable. In these locations, readily available topography comes from radar-derived DEMs, such as SRTM and ASTER, that have low horizontal resolutions (30–90 m per pixel), speckling effects and cannot separate land elevation from vegetation canopy height in low-relief alluvial settings[16]. Moreover, most avulsions in the satellite era occur in active tectonic settings such as the Andean, Himalayan and Papua New Guinean basins, in which dense tree cover affects radar performance.

To address this issue, we use ICESat-2 and the Forest and Buildings removed Copernicus Digital Elevation Model (FABDEM) to analyse topography around rivers. NASA's ICESat-2 (2018 to the present) provides a lidar platform that balances resolution, coverage and latency for precise measurements of narrow swaths of Earth's surface. ATLAS—ICESat-2's single instrument—is a spaceborne lidar system that uses photon-counting technology and operates at a high repetition rate of about 10 kHz. This enables ATLAS to cover narrow swaths of the Earth's surface with near-continuous data[35]. ICESat-2 can capture the topography around river channels, including the alluvial ridge (Extended Data Fig. 1). Of the products derived from ICESat-2, we use ATL03, the geolocated photon data, and ATL08, land and vegetation heights. ATL03 photons have an along-track spacing of 0.70 m and a circular footprint about 17 m in diameter. ATL08 data are produced by a custom photon classification algorithm that reduces ATL03 data into 100-m segments[36]. These segments capture the ground surface elevation accurately, even in areas with high vegetation cover (up to 90%; root mean square error (RMSE) = 0.73 m (ref. 19)), which is ideal for quantifying gradient advantage and superelevation on recently avulsed rivers in riparian or tropical settings. A key feature of the ATL08 product is that it classifies all ATL03 photons within 100-m segments before reducing them. This allows us to remap ATL08 surface classifications at 100-m resolution back to sub-metre resolution with geolocated photon data (ATL03). Recent work[35] showed that ATL03 photon data with remapped ATL08 classifications can be used reliably in vegetated areas (R^2 = 1.00, RMSE = 0.75 m).

FABDEM was created with a three-step algorithm to remove signals of forests and buildings from the 30-m Copernicus DEM (COPDEM30)[20]. FABDEM outperformed other global DEMs in deltaic flood models owing to its better estimation of delta surface heights and a lack of striping or other artefacts[37,38]. Crucially, FABDEM was originally validated using ICESat-2 ATL08 data, showing good alignment with a mean error of −0.03 m and an error percentage <2 m of 80.42%. The water surfaces in FABDEM are inherited from COPDEM30, which was hydro-flattened, rendering the water surface non-physical[39]. As such, we used ICESat-2 to measure water surface elevations for all 58 rivers. This was possible because water surfaces reflect ICESat-2 photons more readily than vegetated ground cover and orthogonality restraints could be relaxed, leading to more available data on water surface elevations. We further validated FABDEM by comparing measurements of alluvial ridge heights and slopes with those from ICESat-2 (n = 38; Extended Data Fig. 2 and Supplementary Table 3). These validation measurements are denoted in the data files, whereas ten were not co-located with avulsion sites.

Calculating β and γ

We generated three measurements of β and γ for every avulsion site. The measurements were done with a combination of FABDEM and ICESat-2. Repeat measurements were taken within roughly three channel widths to reduce measurement variance around the avulsion node. ICESat-2 was preferred for all topographic measurements apart from S_M, which was exclusively measured using FABDEM. ICESat-2 data were limited in coverage and required transects to be close to the avulsion node (within about ten channel widths), orthogonal (within roughly 20° of the perpendicular to the channel centreline) and to contain high enough signal-to-noise ratio to resolve ground elevations. The data sources for each avulsion site are listed in Supplementary Table 2.

To calculate γ, we measured S_{AR} orthogonal to the main channel at points upstream of the avulsion node on the main channel (Supplementary Table 2). We avoided measurements downstream of the avulsion site because this part of the channel was created during the avulsion and may have different characteristics. Slopes were measured from the ridge crest to the floodplain and calculated as alluvial ridge crest elevation minus floodplain elevation over distance between. For S_M, we measured upstream and downstream of the avulsion node along the main channel within approximately 50 channel widths. Effort was made to maintain local slopes, that is, omitting obvious upstream or downstream slope breaks to ensure a linearly sloping region around the avulsion node. Repeat measurements of γ were taken, one on either side of the channel and a third directly next to one of the previous measurements. Sometimes, only one bank shows a ridge form, in which case all measurements are taken on that side.

Calculating β requires measuring: (1) alluvial ridge elevation; (2) water surface elevation; (3) floodplain elevation; and (4) channel depth. As ICESat-2 can rarely resolve channel bed elevations owing to strong absorption of near-infrared laser pulses by turbid water, the fourth value, channel depth, was estimated with a machine learning model (see the next section, 'Estimating channel depth'). H_{AR} was calculated by subtracting the alluvial ridge elevation from the elevation of the floodplain (Extended Data Fig. 1). The alluvial ridge elevation was determined by selecting the lowest of the two high points on either side of the channel along an orthogonal cross-section (for example, Extended Data Fig. 1).

Determining what constitutes the floodplain is not always straightforward. Following ref. 26, we manually define the floodplain elevation as the location next to the channel for which the ridge flank slope asymptotically approaches the surrounding floodplain. If there is no obvious asymptote but the topography next to the ridge form approaches a constant level of variability at lower scales than the ridge form itself, we select the position on the cross-section that represents the onset of the most consistent low elevation within that range of variability. Sometimes, one bank abuts a valley wall whereas the other side maintains a more typical ridge geometry. In these cases, measurements are taken on the half-ridge form. In edge cases that do not fall into the categories defined above, the floodplain is picked as the lowest point in between the ridge form and the next highest topographic location that is taller than the ridge moving outward along the cross-section.

Estimating channel depth

There are numerous ways to estimate channel depth using empirical equations from hydraulic geometry[40] or semi-theoretical equations from open channel flow[41]. Theoretical equations are difficult to constrain and hydraulic geometry equations can vary from site to site. To overcome these issues and generate a reliable estimate of depth that can be applied globally, we trained a machine learning model on a dataset of 4,006 measurements from two datasets of channel depth, discharge, slope and width[21,42]. Our model, the Boost-Assisted Stream Estimator for Depth (BASED), uses an XGBoost regressor as the prediction engine. XGBoost implements gradient-boosted decision trees and is considered the state-of-the-art regression model for structured data, even outperforming deep learning models in regression and classification tasks[43,44]. Our BASED model uses measured discharge, channel slope and channel width to estimate bankfull depth. The performance of the model was gauged with the mean absolute error (MAE), the RMSE, the coefficient of variation (R^2) and the mean absolute percentage error (MAPE). The final model's metrics are MAE = 33 cm, RMSE = 104 cm, R^2 = 0.89 and MAPE = 20% (Extended Data Fig. 3a).

To develop the BASED, we split the 4,006 measurements into training (n = 3,046) and testing (n = 762) sets (Supplementary Table 4). We compared the BASED model to a simple power law function between depth and discharge for the test set (Extended Data Fig. 3a,b). We trained the BASED model using a Bayesian parameter tuning and cross-validation scheme to avoid overfit. The training datasets we use contain a mix of bankfull and non-bankfull discharge measurements, so to validate if BASED can predict bankfull depth, we compared depth predictions to a validation set of bankfull depths (n = 198)[45] (Supplementary Table 5). The BASED can predict bankfull depth measured in the field (blue dots, Extended Data Fig. 3c). After training, testing and validation, the BASED model was retrained on all the available data to generate the most accurate and comprehensive model for deployment.

BASED requires discharge, channel bed slope and channel width to estimate bankfull depth, but we do not have in situ discharge measurements for the avulsed rivers. So, for discharge of the avulsed rivers, we use the maximum modelled values of natural discharge from the RiverATLAS dataset, a global-coverage dataset of yearly modelled river discharge values[46]. Discharges are derived from 30-year (1961–1990) monthly discharge averages calculated by the WaterGAP v2.2 model[46]. River discharge is calculated at 0.5° resolution and downscaled to 15-arcsec resolution[47]. The downscaled discharges are measured to 0.001 $m^3 s^{-1}$, but accuracy at that level is not entirely reliable. To test the reliability of discharges from RiverATLAS, we reran the BASED model predictions for the validation set[45] but used the annual maximum discharge reported by RiverATLAS for each river (red dots, Extended Data Fig. 3c). The BASED model underpredicts depth because the RiverATLAS annual maximum discharge underestimates measured bankfull discharge (Extended Data Fig. 3d). To address this issue, we fit a power law to the relationship between modelled and measured discharge and then inverted it to correct the RiverATLAS discharge. After correcting the RiverATLAS discharges for the validation set, the predicted values are now centred on the 1:1 line with an RMSE of 47 cm (Extended Data Fig. 3e). We predict the depths of avulsed rivers in our dataset applying this correction to the annual maximum discharge values from RiverATLAS.

Superelevation (see ref. 5) defines the channel depth (H_M) as the difference between the elevation of the alluvial ridge crest and the channel bed (Extended Data Fig. 1). We use three operational rules to determine the depth of rivers in our dataset. Let's define A as the distance measured from the alluvial ridge crest to the water surface as measured by ICESat-2 and B as the channel depth predicted from BASED. If: (1) $A \leq B$, we set $H_M = B$ because the predicted depth is below the observed water surface from ICESat-2; (2) $1 < A < 1.5B$, we look for evidence in the ICESat-2 profile that in-channel sediment bars are exposed. If we see these, we set $H_M = A$. If bars are not visible, then we use method 3; (3) $A \geq 1.5B$, then we set $H_M = A + B$. Roughly one-third of the rivers in our dataset have $A \geq 1.5B$. By using $H_M = A + B$, we are effectively assuming that the water surface recorded by ICESat-2 represents some kind of average condition. The depth rule we used for each avulsion is recorded in Supplementary Table 2. In four cases, the BASED predicted a small negative number, a possibility when working with unbounded regressions. In these cases, we substituted a low positive value, 0.1 m.

The cause of why $A > 1.5B$ is not easy to pinpoint, but one possibility is that these rivers are out of equilibrium with their discharge. The ref. 21 dataset represents one of the largest and most complete compilations of river depth and its controlling variables, but the rivers used are from North American and European settings. The avulsing rivers used in this study are in monsoonal climates and strongly aggradational geologic settings in which it is possible that levee aggradation could be decoupled from bed aggradation, leading to larger than expected depths. Avulsing rivers also commonly reoccupy pre-existing channels on the floodplain and the depth they inherit may not be in equilibrium with the discharge.

Error treatment

We used a probabilistic error treatment for β and an analytical one for γ based on the data incorporated into the two ratios. β contains a measurement from FABDEM in the numerator H_{AR} and in the denominator a depth estimate from the BASED, a model output or the elevation datasets, depending on the relationship between A and B (see above). Owing to these nuances in the denominator, the underlying distribution of error is unconstrained, requiring a probabilistic approach. γ consists of two measurements from only ICESat-2 or FABDEM and the standard deviations of both the numerator and the denominator are log-normal. That is, the error distribution is known a priori and, therefore, analytical log-adjusted error propagation methods can be used.

We estimated β ratios and their uncertainties using a Monte Carlo simulation approach. For each data point, we randomly selected one of the three measurements (i = 1, 2 or 3) that correspond to different combinations of alluvial ridge elevation, floodplain elevation and water elevation. We then sampled the ridge height from a triangular distribution with a lower bound of −0.6 m, a mode of 0 m and an upper bound of 0.6 m, based on the MAE values from Extended Data Fig. 2. We sampled the water surface elevation from a triangular distribution with a lower bound of −0.25 m, a mode of 0 m and an upper bound of 0.25 m. We sampled the BASED predicted depth from a triangular distribution with a lower bound of −1.0 m, a mode of 0 m and an upper bound of 1.0 m. The errors represent the uncertainty in the measurements of these variables. We calculated the superelevation ratio using one of the three methods (1, 2 or 3), depending on the method used (see the 'Estimating channel depth' section). We repeated this process for a large number of simulations (n = 10,000).

For the variable γ, we quantified uncertainty by computing the standard errors for the means of S_{AR} and S_M, which were obtained by dividing their respective standard deviations by the square root of the number of measurements involved. The means were log-transformed on the basis of the results of Shapiro–Wilk tests that confirmed the normality of the log-transformed S_{AR} and S_M data. The standard error of the log-transformed means was then calculated, providing a measure of variability in the log domain. Subsequently, we derived the relative error in γ by taking the square root of the sum of the squared log-transformed standard errors, which presumes a log-normal distribution of uncertainty. We back-transformed this relative error to express the absolute uncertainty in γ on its original scale.

Softmax random walk

Our softmax random walk model describes a method for calculating the probability of movement across a grid of elevations as a function of topographic slope and a representation of flow inertia. The slope is

calculated on the basis of the elevation difference between the current position and its neighbours, adjusted for the direction of movement (diagonal or straight). The inertia is determined by comparing the direction of the last move to the potential next move, using cosine similarity to quantify similarity between directions. These two factors are then weighted and combined using specified parameters to represent the influence of slope and inertia on movement. The combined weight is transformed using the softmax function, which converts these weights into probabilities, ensuring they sum to 1 and thus can be interpreted as the likelihood of moving to each neighbouring position. The model avoids returning to the immediate last position to simulate more realistic movement patterns.

This method provides a probabilistic prediction for movement in a terrain-aware context, blending physical geography (slope, ϕ) with movement history (inertia, ψ). The selection of the starting parameters ($\omega = 0.5$, $\phi = 10$ and $\psi = 1$) was informed by repeated trial and error. We opted to use the selected parameters because they provided good fits across a variety of floodplain physiographies. Other combinations that include a larger value of ψ also provided good fits. See the Supplementary Information for an extended derivation.

Data availability

The authors declare that all data supporting the findings of this study are available at https://doi.org/10.5281/zenodo.10338685 (ref. 48).

Code availability

Code for reproduction, including data cleaning, analysis and plotting, is available at https://zenodo.org/records/13693548 (ref. 49).

33. Edmonds, D. A., Hajek, E. A., Downton, N. & Bryk, A. B. Avulsion flow-path selection on rivers in foreland basins. *Geology* **44**, 695–698 (2016).
34. Valenza, J. M., Edmonds, D. A., Hwang, T. & Roy, S. Downstream changes in river avulsion style are related to channel morphology. *Nat. Commun.* **11**, 2116 (2020).
35. Xing, Y., Huang, J., Gruen, A. & Qin, L. Assessing the performance of ICESat-2/ATLAS multi-channel photon data for estimating ground topography in forested terrain. *Remote Sen.* **12**, 2084 (2020).
36. Neuenschwander, A. L. et al. ATLAS/ICESat-2 L3a Land and Vegetation Height, Version 6. *National Snow and Ice Data Center* https://doi.org/10.5067/ATLAS/ATL08.006 (2023).
37. Dandabathula, G., Hari, R., Ghosh, K., Bera, A. K. & Srivastav, S. K. Accuracy assessment of digital bare-earth model using ICESat-2 photons: analysis of the FABDEM. *Model. Earth Syst. Environ.* **9**, 2677–2694 (2023).
38. Seeger, K. et al. Assessing land elevation in the Ayeyarwady Delta (Myanmar) and its relevance for studying sea level rise and delta flooding. *Hydrol. Earth Syst. Sci.* **27**, 2257–2281 (2023).
39. Fahrland, E. Copernicus DEM product handbook. Airbus Defence and Space GmbH (2022).
40. Leopold, L. B. & Maddock, T. The hydraulic geometry of stream channels and some physiographic implications. USGS Professional Paper No. 252 (1953).
41. Parker, G. Hydraulic geometry of active gravel rivers. *J. Hydraul. Div.* **105**, 1185–1201 (1979).
42. Dunne, K. B. J. & Jerolmack, D. J. Evidence of, and a proposed explanation for, bimodal transport states in alluvial rivers. *Earth Surf. Dyn.* **6**, 583–594 (2018).
43. Shwartz-Ziv, R. & Armon, A. Tabular data: deep learning is not all you need. *Inf. Fusion* **81**, 84–90 (2022).
44. NVIDIA Data Science Glossary. XGBoost https://www.nvidia.com/en-us/glossary/data-science/xgboost/.
45. Trampush, S. M., Huzurbazar, S. & McElroy, B. Empirical assessment of theory for bankfull characteristics of alluvial channels. *Water Resour. Res.* **50**, 9211–9220 (2014).
46. Linke, S. et al. Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Sci. Data* **6**, 283 (2019).
47. Müller Schmied, H. et al. The global water resources and use model WaterGAP v2.2d: model description and evaluation. *Geosci. Model Dev.* **14**, 1037–1079 (2021).
48. Gearon, J. Rules of river avulsion supplementary data files. *Zenodo* https://doi.org/10.5281/zenodo.10338685 (2024).
49. Gearon, J. jameshgrn/rulesofriveravulsion: publication_release. *Zenodo* https://doi.org/10.5281/zenodo.13693548 (2024).
50. Lehner, B. & Grill, G. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **27**, 2171–2186 (2013).
51. Center for International Earth Science Information Network (CIESIN), Columbia University. Global Gridded Relative Deprivation Index (GRDI), Version 1. NASA Socioeconomic Data and Applications Center (SEDAC) https://doi.org/10.7927/3xxe-ap97 (2022).
52. G20 Background Brief. G20 Presidency https://www.g20.org/en/about-the-g20 (2023).

Acknowledgements D.A.E., J.H.G. and H.K.M. were supported by U.S. National Science Foundation grant EAR-1911321. H.K.M. was also supported by National Aeronautics and Space Administration (NASA) Future Investigators in NASA Earth and Space Science and Technology (FINESST) grant 80NSSC21K1598. E.A.B was supported by US National Science Foundation grant EAR 2052844.

Author contributions J.H.G. and D.A.E. conceived the presented ideas, with help from H.K.M. J.H.G. and D.A.E. wrote the manuscript, with contributions and revisions from E.A.B., H.K.M., C.D., C.P. and D.M. H.K.M. and C.D. contributed to the conception of equation (5). C.P. and D.M. contributed to and reviewed equations (1)–(4). E.A.B., D.M. and C.P. contributed to the design of analyses and interpretation of results. D.A.E. supervised the project.

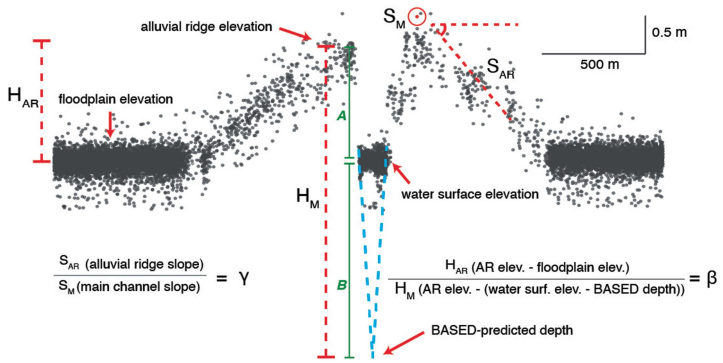
Competing interests The authors declare no competing interests.

Additional information

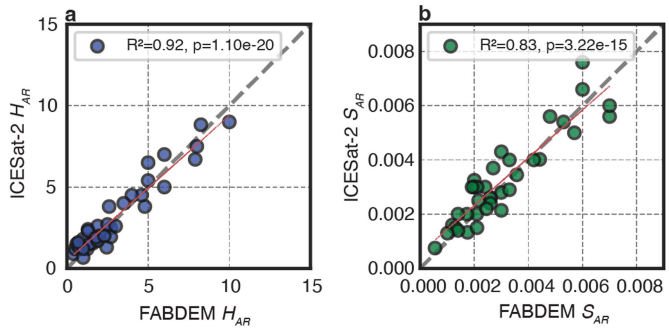
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-07964-2.

Correspondence and requests for materials should be addressed to James H. Gearon.

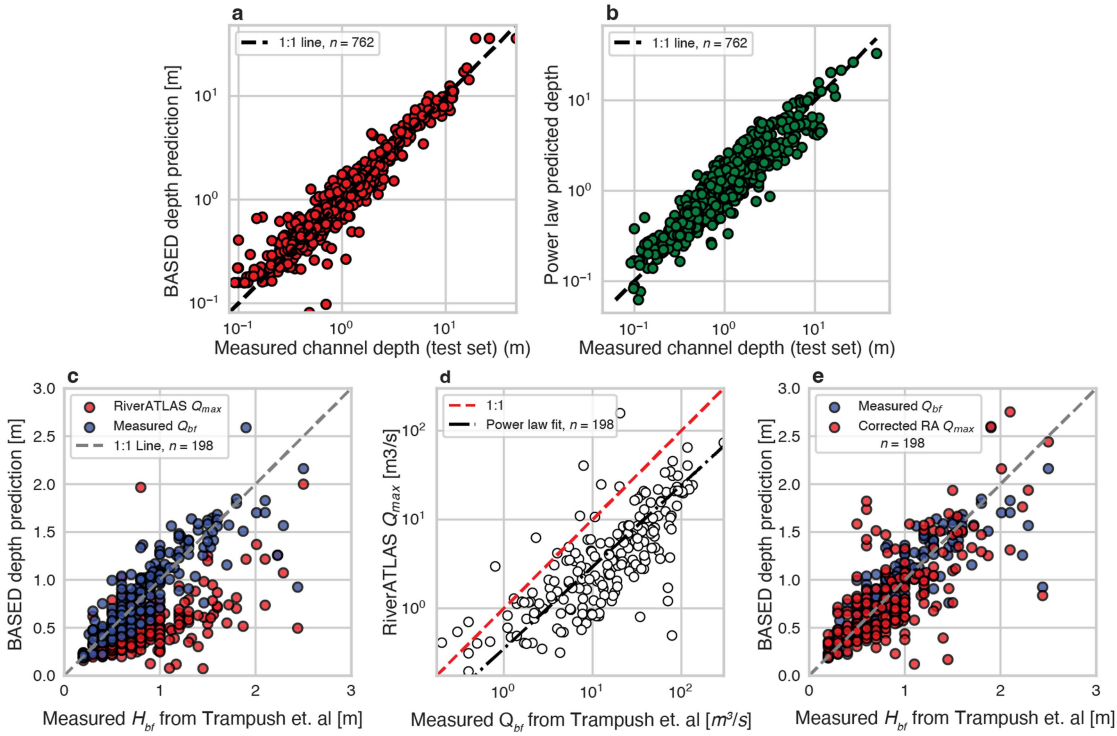
Peer review information *Nature* thanks Eric Lajeunesse and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



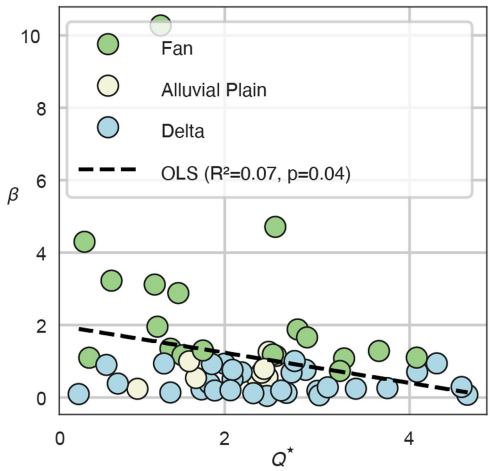
Extended Data Fig. 1 | Computation of β and γ for avulsion TURK_002_1991 (40.143, 28.556). Using both ICESat-2 and FABDEM, these calculations were done three times based on distinct measurements.



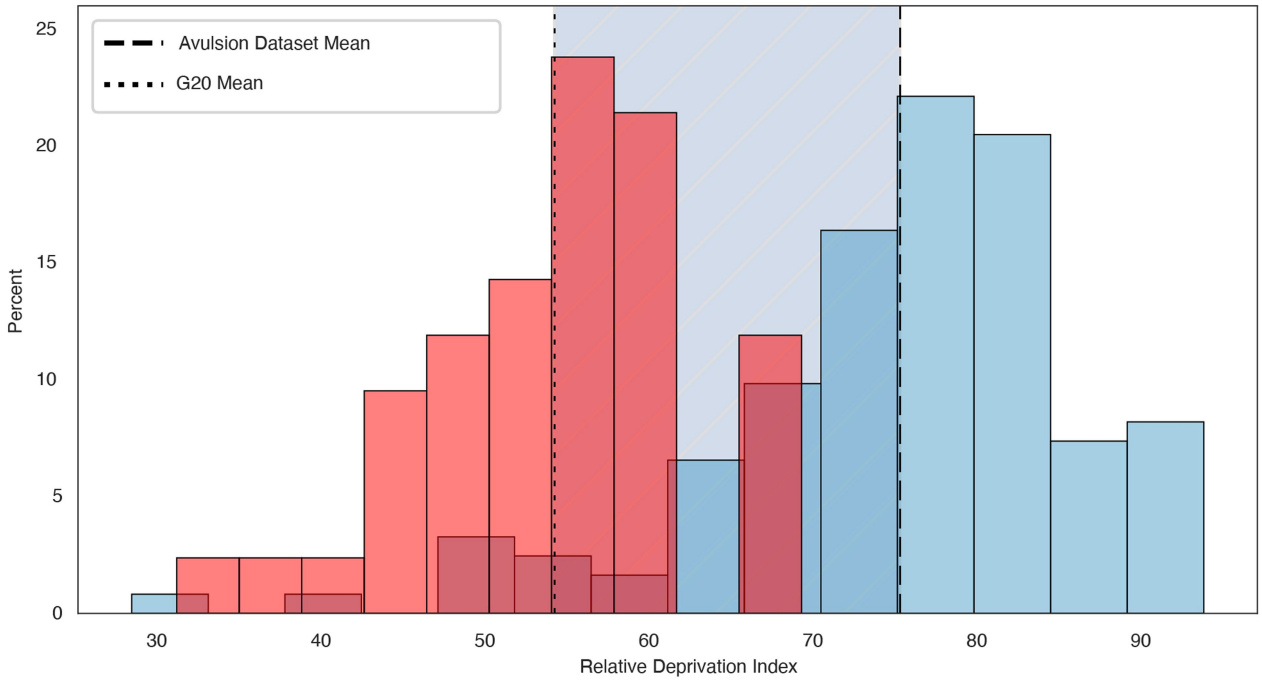
Extended Data Fig. 2 | Validation of alluvial ridge height (H_{AR}) and slope (S_{AR}) measurements between FABDEM and ICESat-2, $n = 38$. a, Alluvial ridge height 1:1 plot, MAE is 0.59 m, RMSE is 0.71 m and R^2 is approximately 0.92 with a p-value of 1.1×10^{-20}. **b**, Alluvial ridge slope 1:1 plot, MAE is 0.0056 m m^{-1}, RMSE is 0.007 m m^{-1} and R^2 is 0.83.



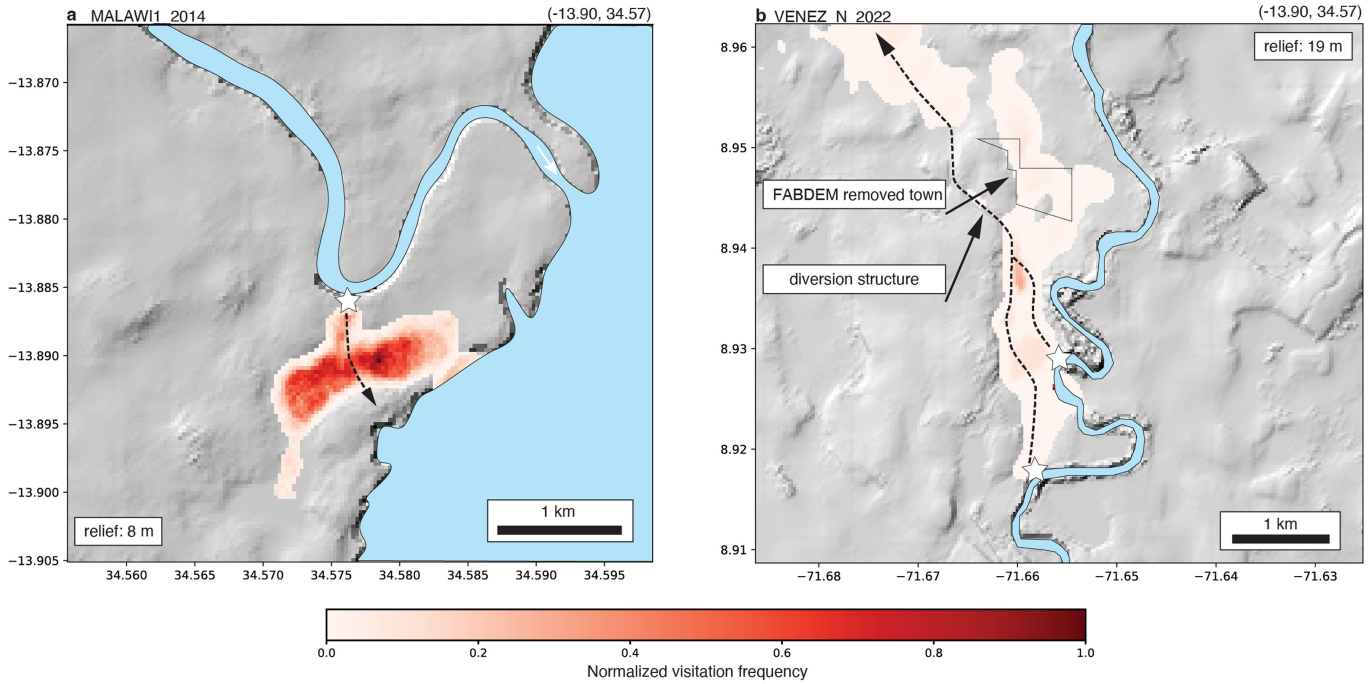
Extended Data Fig. 3 | BASED model validation and RiverATLAS discharge correction. a, BASED-predicted depths against measured depth from the test set ($n = 762$). MAE = 33 cm, RMSE = 102 cm, $R^2 = 0.89$ and MAPE = 20%. The BASED achieves higher accuracy than a power-law fit for estimates of channel depths. **b**, Predicting channel depth is typically done with a power-law fit ($n = 762$). Using the test set from a compilation of hydraulic geometric data[21,42], we fit a relationship between depth and discharge ($y = 0.31x^{0.38}$). In this case, MAE = 64 cm, RMSE = 129 cm, $R^2 = 0.83$ and MAPE is 35%. **c**, BASED-predicted depths against a validation set of measured bankfull depths from ref. 45. Blue points are BASED depth estimates using in situ discharge measurements, red points are BASED depth estimates using RiverATLAS-modelled maximum discharge values ($n = 198$). **d**, Relationship between RiverATLAS maximum discharge and measured discharge from ref. 45. RiverATLAS systematically underestimates discharge and to correct this, we fit a power law ($y = 0.39x^{0.90}$) to the data and invert it to get **e** ($n = 198$). **e**, Relationship between BASED depth estimates using the corrected RiverATLAS maximum discharge and measured bankfull depth from ref. 45. Associated RMSE is 47 cm ($n = 198$).



Extended Data Fig. 4 | No notable correlation between β and discharge variance. We calculated Q^* as the range of corrected annual discharge over corrected mean annual discharge from RiverATLAS[50]. The more typical coefficient of variation requires the standard deviation to compute, which is unavailable in RiverATLAS. As such, we chose Q^* as it also describes variance around the central tendency. An OLS (ordinary least squares) regression on calculated β values and Q^* shows no notable correlation between the two ratios ($n = 58$).



Extended Data Fig. 5 | Disproportionate impact of river avulsion on the Global South. Using data from NASA's Global Gridded Relative Deprivation Index (GRDI), we compiled mean values for watersheds that contain our avulsion nodes, based on the Pfafstetter level 7 (blue). Comparing these with mean values of the G20 nations (red)—which represent a group of industrialized nations accounting for 85% of the global GDP[51,52]—we find that G20 deprivation indices are notably lower than in areas in which recent river avulsions have taken place.



Extended Data Fig. 6 | Further examples of the softmax random walk. **a**, Delta avulsion. Note that the avulsion carved through the downstream floodplain, something the random walk cannot reproduce. **b**, Floodplain continuing avulsions. In this case, the flow was directed around a town with a diversion structure, causing a slight misfit in the results. These avulsions are not included in the avulsion dataset as they are incomplete.