Extreme flooding controlled by stream network organization and flow regime

Stefano Basso [1,2] ✉, Ralf Merz [1,3], Larisa Tarasova [1,4] & Arianna Miniussi [1,5]

River floods are among the most common natural disasters worldwide, with substantial economic and humanitarian costs. Despite enormous efforts, gauging the risk of extreme floods with unprecedented magnitude is an outstanding challenge. Limited observational data from very high-magnitude flood events hinders prediction efforts and the identification of discharge thresholds marking the rise of progressively larger floods, termed flood divides. Combining long hydroclimatic records and a process-based model for flood hazard assessment, here we demonstrate that the spatial organization of stream networks and the river flow regime control the appearance of flood divides and extreme floods. In contrast with their ubiquitous attribution to extreme rainfall and anomalous antecedent conditions, we show that the propensity to generate extreme floods is well predicted by intrinsic properties of river basins. Most importantly, it can be assessed prior to the occurrence of catastrophes through measurable metrics of these properties derived from commonly available discharge data, namely the hydrograph recession exponent and the coefficient of variation of daily flows. These results highlight the propensity of certain rivers for generating extreme floods and demonstrate the importance of using hazard mapping tools that, rather than solely relying on past flood records, identify regions susceptible to the occurrence of extreme floods from ordinary discharge dynamics.

River floods are primary natural disasters, steadily accounting for several billion dollar losses every year and most of the affected population[1,2]. Assessment of the flood hazard is complicated by runoff generation processes which might be more variable than observed records suggest, let alone ongoing global change[3,4]. A reliable evaluation of the propensity of rivers to undergo extreme floods with magnitudes not previously experienced (here quantified by the values of the river discharge) is therefore crucial for urban planning, designing engineering structures, pricing insurances, and laying out mitigation and adaptation strategies.

Flood hazard assessment is particularly difficult when the magnitude of the rarer floods strongly increases[5,6]. Whenever flood magnitudes rise gradually with diminishing chance of their occurrence (Fig. 1a), they are indeed also characterized by high predictability[7]. In some cases (Fig. 1b), however, a clear growth of the magnitude of the rarer floods points to the possible occurrence of very large events that arise unexpectedly[8], often causing catastrophic socio-economic outcomes, as in the recent case of the July 2021 floods in Germany.

Several studies signalled the pervasiveness of the latter phenomena[9,10]. A few works tried to link these behaviours to the catchment water balance[11,12] and suggest, on the basis of extensive field surveys in two small basins and a synthetic experiment, that they may occur when the catchment storage capacity is exceeded[13,14]. Nonetheless,

[1]Department of Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle (Saale), Germany. [2]Norwegian Institute for Water Research (NIVA), Oslo, Norway. [3]Institute of Geosciences and Geography, Martin-Luther University Halle-Wittenberg, Halle (Saale), Germany. [4]Department of Earth and Environmental Engineering, Columbia University, New York, NY, USA. [5]P&C International, General Reinsurance, Cologne, Germany. ✉e-mail: stefano.basso@niva.no
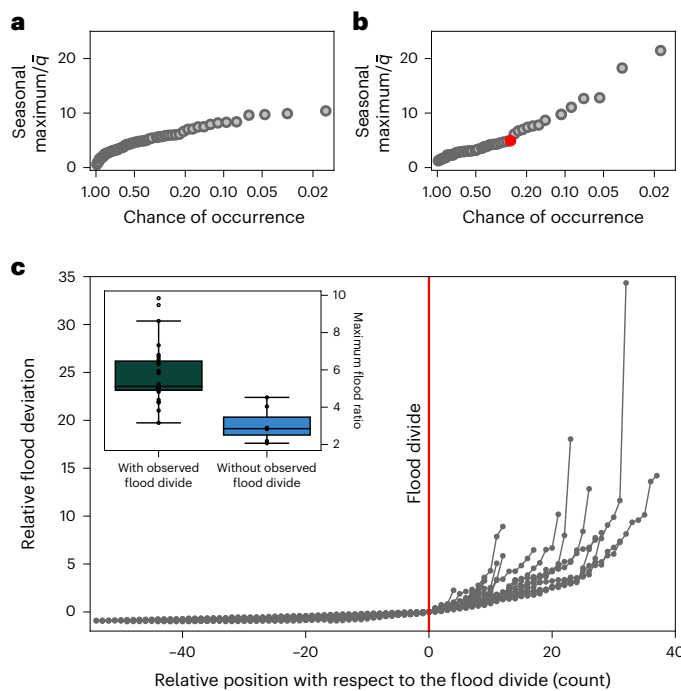
Fig. 1 | Observed flood divides and extreme floods. a,b, Normalized seasonal maxima versus chance of occurrence (Methods) for case studies lacking (ID 11402001, autumn; **a**) and exhibiting (ID 11946000, summer; **b**) a flood divide (red dot). **c**, Relative deviation from the flood divide of observed seasonal floods (grey dots linked by solid lines) sorted by their chance of occurrence. Inset shows ratios between highest observed and mean maximum seasonal flood for case studies (black dots) with ($n = 27$) and without ($n = 7$) flood divides. Centre line: median; box limits: 25th and 75th percentiles; whiskers: minimum and maximum values that are not considered outliers, that is, 1.5 × interquartile range.

data constraints[7,15] and spare knowledge on their causes limit our skills to diagnose the possible occurrence of extreme events based on precursory signals, as done for other natural and societal phenomena[16].

Here, we combine long hydroclimatic records and a mechanistic–stochastic approach to flood hazard assessment to reveal that the spatial organization of stream networks and the river flow regime jointly control the emergence of pronounced increases of the magnitude of the rarer floods. We further demonstrate that the identified controls enable predicting the propensity of rivers to generate extreme floods in an additional set of several thousand case studies.

Sharp rise of flood magnitudes

We identified rivers that exhibit a strong increase in the magnitude of rarer floods from a set of 101 case studies from mid-sized (drainage area: 43–9,052 km^2, median: 865 km^2) unregulated basins in the United States and Germany, denoted as the study dataset (Extended Data Fig. 1, Extended Data Table 1 and Methods). We further pinpointed the case-specific flood magnitude whose exceedance marks the rise of progressively larger floods[14] (Fig. 1b), which we term 'flood divide' as it discriminates between common and increasingly extreme floods.

To identify flood divides we applied a protocol (Methods) to the characteristic relation between river flood magnitudes and their chance of occurrence, which was both empirically derived from available observations and inferred from a well-established mathematical description of precipitation, soil moisture and runoff dynamics in river basins (Methods), referred to as the physically based extreme value (PHEV) distribution of river flows[12,17].

In all instances where the empirical and theoretical methods consistently identified a flood divide (27 out of 101 case studies),

this feature neatly partitions contrasting characters in terms of the increase of flood magnitudes (Fig. 1c). On the left-hand side of the flood divide, magnitudes gently rise within a narrow range of common values, whereas they substantially increase with a remarkable nonlinear growth on the right-hand side of it. The flood divide is thus an effective attribute to distinguish common from increasingly extreme floods that may occur in river basins.

Moreover, the existence of a flood divide indicates whether much larger floods than those observed on average shall be expected in a river[18]. The ratios between the highest observed flood and the mean maximum seasonal flood indeed significantly differ (Kolmogorov–Smirnov test, $P < 0.01$; Methods) for basins with or without flood divides (inset of Fig. 1c). When a flood divide exists, the highest observed floods are much larger, with deviations from their mean values twice as big on average than for basins with no flood divide. In particular, floods can reach exceptionally high magnitudes of up to ten times the mean maximum seasonal flood in these cases, a prospect also evoked by research on historical and palaeofloods[18–20]. The existence and magnitude of the flood divide therefore represent pivotal features to characterize the propensity of rivers to extreme hydrological events and raise awareness of the intrinsic peril of floods in these contexts.

The magnitude of the flood divide varies between 2.5 and 35 (median: 8.4) times the long-term mean discharge of rivers in the study dataset. Comparison of the empirical (that is, derived from the available data records) and theoretical (that is, inferred through PHEV) positions of the flood divide (Extended Data Fig. 2) indicates good agreement (distance and Spearman correlation coefficients of 0.64 and 0.44, respectively, $P < 0.01$; Methods) in this varied set of case studies. The adopted mechanistic–stochastic description of hydrological dynamics is thus suitable to guide investigation on what are the physical controls of the emergence of extreme floods in river basins.

Watershed features promoting extreme floods

In contrast with the ubiquitous attribution of extreme flood instances to intense rainfall and anomalous antecedent conditions[21–23], we show here that intrinsic attributes of river basins explain the penchant of rivers for generating extreme floods.

We applied a dimensional analysis tool (Methods) to the description of hydrologic dynamics provided by PHEV to set research hypotheses on the key factors promoting the occurrence of flood divides in river basins. We then validated the hypotheses with observations and tested whether the identified controls accurately predict the occurrence of extreme floods. The analysis indicates that two specific watershed properties, namely the hydrograph recession exponent[24] and the coefficient of variation of daily flows[25], control the emergence and magnitude of the flood divide.

The hydrograph recession exponent is a compelling descriptor of the geomorphological structure of the contributing river basin[24,26], which determines how watersheds funnel runoff towards their outlets[27]. Specifically, it stems from the spatial organization of the stream network, which defines how the geometry of saturated areas[26] and drainage of the riparian unconfined aquifer[24] vary in time and contribute to discharge. The coefficient of variation of daily flows instead arises from distinctive interactions among precipitation inputs, evapotranspiration rates, soil moisture dynamics and response times of river basins[25]. It summarizes in a single index how watersheds filter the incoming climate signal[28] and thus recaps the chance of precipitation falling on dry or saturated basins, which in turn controls the mix of small and large runoff events[29]. Although typically estimated from streamflow observations, both these properties and their descriptors can be likewise evaluated from commonly available hydroclimatic data series[30] and geomorphological data only[24,26].

A distance correlation coefficient (Methods) of the multivariate relation between observed flood divides and the two physioclimatic controls equal to 0.47 ($P < 0.05$) confirms significant dependence
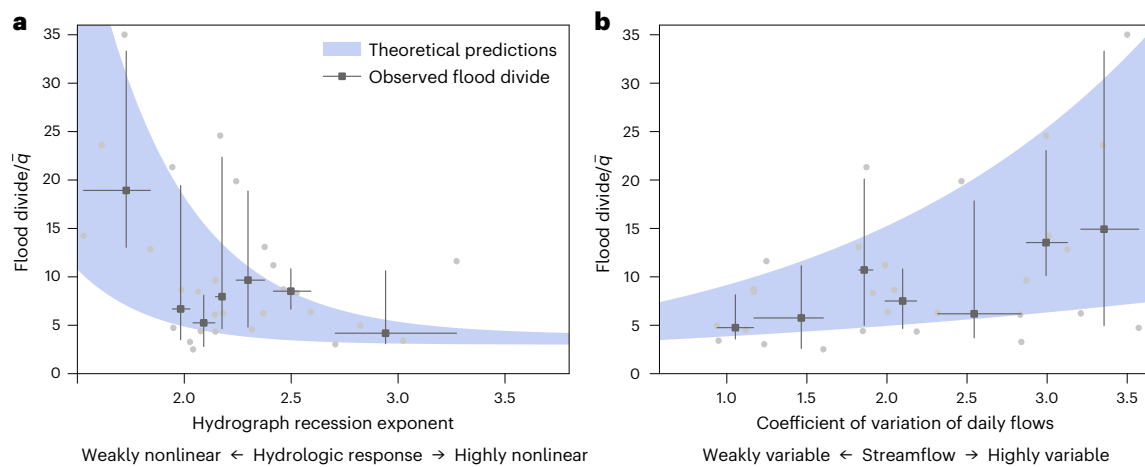
Fig. 2 | Magnitude of the flood divide as a function of its physioclimatic controls. a,b, Normalized magnitude (that is, divided by the long-term mean river discharge $\bar{q}$) of the flood divide in the study dataset as a function of the hydrograph recession exponent (**a**) and the coefficient of variation of daily flows (**b**). Shaded areas span the 95% variability range of theoretical predictions and provide an estimate of their uncertainties. Grey markers display the median (squares), minimum and maximum values of the binning (horizontal bars), and 5th and 95th percentile range (vertical bars) of the observations (dots; equal number of $n = 4$ case studies for each bin), here used for validation.

of the magnitude of the flood divide from the hydrograph recession exponent and the coefficient of variation of daily flows. We inferred the form of the bivariate relations as well as their uncertainties through PHEV (Methods) and validated the theoretical patterns by overlying the available observations, which mostly fall within the anticipated ranges (Fig. 2). The distance and Spearman correlation coefficients of the bivariate relations are respectively 0.45 ($P < 0.05$) and −0.3 ($P = 0.12$) for the hydrograph recession exponent (Fig. 2a), and 0.44 ($P < 0.05$) and 0.40 ($P < 0.05$) for the coefficient of variation of daily flows (Fig. 2b).

When the hydrologic response is highly nonlinear, the flood divide appears for relatively small magnitudes (Fig. 2a). Heterogeneous drainage density typically enhances the nonlinearity of the hydrologic response[26]. In these cases, a given increase of the overall length of the stream network actively draining runoff during events determines a superlinear growth of the connected riparian aquifers[24] and saturated areas[26], causing sharp increments of streamflow and the emergence of flood divides. The areas contributing runoff instead add up gradually with more linear hydrologic responses[26], preventing the appearance of flood divides, which shift to increasingly larger magnitudes (Fig. 2a). This evidence corroborates findings of theoretical[31] and modelling studies[14,32] that suggest a role of nonlinear hydrological responses in the occurrence of extreme runoff events.

The magnitude of the flood divide also increases with the streamflow variability (Fig. 2b). The coefficient of variation of daily flows stems from the ratio between interarrivals of runoff-producing precipitation events and response times of river basins[12,25] (Methods). When the interarrival between events is larger than the time required for draining them (because of sporadic precipitation, intense evapotranspiration or fast hydrologic response), watersheds can dry substantially before new precipitation occurs. Events are likely to be filtered by the available basin storage, decreasing the chance of marked growths of the flow magnitudes and shifting the flood divide to larger values. Conversely, when streamflow weakly varies, watersheds experience sustained wet conditions that are likely to cause marked streamflow increments and the emergence of flood divides for relatively small magnitudes.

The coefficient of variation of daily flows hence recapitulates in a single metric the characteristic water storage dynamics of river basins[25]. Its identification as a key control of the emergence of flood divides and extreme floods agrees with studies pointing at a relation between the predisposition of rivers to flooding and the long-term

wetness conditions of their basins[33,34], which largely affect streamflow variability[35]. Here, we confirm with data the key importance of typical water storage dynamics for the emergence of increasingly extreme floods, and provide general explanations of the underlying mechanisms by means of the PHEV framework.

Foreseeing the chance of extreme floods

A question that naturally arises is whether we can label river basins as hazardous (that is, they may exhibit flood divides and extreme floods) by leveraging the hydrograph recession exponent and the coefficient of variation of daily flows as indicators. We show here that we can indeed provide accurate predictions by means of binary logistic regression (Methods), using the two properties as explanatory variables of the likely occurrence of extreme floods in river basins. We first evaluated reliability and robustness of the predictions over the study dataset in a cross-validation fashion (Methods). The large majority of results are true cases (Extended Data Fig. 3), which indicate good ability to identify either the emergence of flood divides (true positives) or their absence (true negatives) from the two physioclimatic properties. Median balanced accuracy and the Matthews correlation coefficient (MCC; Methods) of 0.87 (interquartile range: 0.80–0.94) and 0.63 (0.44–0.77), respectively, denote overall high prediction accuracy. In particular, hydrograph recession exponent and streamflow variability successfully categorize river basins as either having flood divides or not in 83% of the cases on average (interquartile range: 75%–92%), and outclass a random classifier (Methods) in 97% of the cases.

We further performed a stress test (Methods) to evaluate the skills of our indicators in foretelling the possibility of extreme floods in a broader set of 2,519 case studies from mid-sized (drainage area: 36–23,843 km², median: 966 km²) unregulated basins, denoted as the test dataset (Extended Data Fig. 1, Extended Data Table 1 and Methods). This is an especially severe trial as, contrary to common practice, the training set is here more than 70 times smaller than the validation set. Median balanced accuracy and MCC are in this case equal to 0.60 (interquartile range: 0.54–0.65) and 0.18 (0.07–0.28), respectively. The onset of several false positives (that is, cases where we predict a flood divide that is not confirmed by observations; Extended Data Fig. 4) mainly causes the decrease of accuracy. These false positive instances may be partly owing to the inclusion in the test dataset of case studies undergoing hydrological processes (for example, snowmelt, strongly
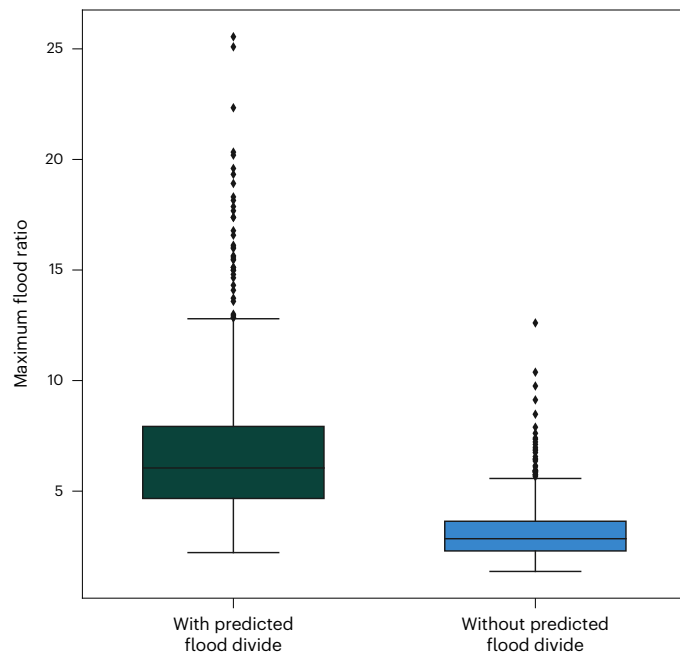
Fig. 3 | Prediction of flood divides and extreme floods from their identified physioclimatic controls. Ratios between the highest observed flood and the mean maximum seasonal flood for case studies in the test dataset for which we predict the presence (true positives, $n = 531$) and the absence (true negatives, $n = 728$) of a flood divide. Centre line: median; box limits: 25th and 75th percentiles; whiskers: minimum and maximum values that are not considered outliers, that is, 1.5 × interquartile range; dots: outliers.

variable recession properties across events) that are not explicitly characterized by the adopted theoretical framework, and for which the identified physioclimatic controls might hence be only partially telling. However, past studies also show that marked growths of the magnitude of the rarer floods are systematically more often detected with longer data records[7], and argue that extreme floods that would allow us to confirm these predictions may not be included in the available observations because of their limited lengths[36]. This is probably the case here, as the fraction of false positive cases in the test dataset consistently decreases with increasing data length (Extended Data Fig. 5a). Moreover, previous studies also highlight that marked rises of the magnitude of the rarer floods are less clearly identified from observations in humid regions[7] characterized by reduced streamflow variability[25], as for our set of false positives (Extended Data Fig. 5b). The lower likelihood of observing extreme events in these contexts[29] hence suggests caution in considering these basins as at low-risk. Here we simply note that, notwithstanding the false positive labels, also in the stress test the descriptors of stream network organization and river flow variability outdo a random classifier in 87% of the cases.

Most importantly, the predicted existence of a flood divide based on these two physioclimatic features of watersheds provides indications on whether much larger floods than those observed on average shall be expected in a river basin (Fig. 3), analogously to that previously shown for observed flood divides (inset of Fig. 1c). In fact, our predictions successfully distinguish river basins in the test dataset where extreme deviations of the highest observed flood from the mean maximum seasonal flood occur. The ratios between the latter variables are indeed significantly larger (Kolmogorov–Smirnov test, $P < 0.01$) for case studies where we predicted the existence of flood divides, regardless of whether we benchmark our predictions against the available observations (Fig. 3) or not (Extended Data Fig. 4b).

Although only tested in mid-sized unregulated river basins, the knowledge gained on the intrinsic attributes of watersheds that control

the emergence of flood divides offers a chance to raise awareness of the propensity of certain rivers to generate extreme floods[18]. The foreseen existence of a flood divide may provide guidance on the choice of alternative statistical tools (for example, light- versus heavy-tailed distributions)[37] widely employed in the practice of flood hazard assessment. Estimates of its expected position empower evaluations of the reliability of discharge records for unveiling the peril of extreme events exceeding the flood divide in river basins subject to varied geomorphological and hydroclimatic settings[38]. Furthermore, the attested feasibility of inferring flood divides from measurable metrics of ordinary discharge dynamics (that is, the hydrograph recession exponent and the coefficient of variation of daily flows), rather than records of streamflow maxima, enables the inception of hazard mapping tools that do not merely rely on past flood records[39], but actively identify hazardous regions that are susceptible to the occurrence of flood divides and extreme floods, thus informing concerned communities of possibly overlooked hazards[40].

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/s41561-023-01155-w.

References

1. *Human Cost of Disasters: An Overview of the Last 20 Years: 2000–2019* (CRED UNDRR, 2020); https://reliefweb.int/report/world/human-cost-disasters-overview-last-20-years-2000-2019
2. Bevere, L. & Remondi, F. *Natural Catastrophes in 2021: The Floodgates Are Open* (Swiss Re Institute, 2022); https://www.swissre.com/institute/research/sigma-research/sigma-2022-01.html
3. IPCC *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (eds Field, C. B. et al.) (Cambridge Univ. Press, 2012).
4. Sharma, A., Wasko, C. & Lettenmaier, D. P. If precipitation extremes are increasing, why aren't floods? *Water Resour. Res.* **54**, 8545–8551 (2018).
5. Merz, B. et al. Causes, impacts and patterns of disastrous river floods. *Nat. Rev. Earth Environ.* **2**, 592–609 (2021).
6. Kreibich, H. et al. The challenge of unprecedented floods and droughts in risk management. *Nature* **608**, 80–86 (2022).
7. Smith, J. A., Cox, A. A., Baeck, M. L., Yang, L. & Bates, P. D. Strange floods: the upper tail of flood peaks in the United States. *Water Resour. Res.* **54**, 6510–6542 (2018).
8. Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable* (Random House, 2007).
9. Bernardara, P., Schertzer, D., Eric, S., Tchiguirinskaia, I. & Lang, M. The flood probability distribution tail: how heavy is it? *Stoch. Environ. Res.* **22**, 5638–5661 (2008).
10. Villarini, G. & Smith, J. Flood peak distributions for the eastern United States. *Water Resour. Res.* **46**, W06504 (2010).
11. Guo, J. et al. Links between flood frequency and annual water balance behaviors: a basis for similarity and regionalization. *Water Resour. Res.* **50**, 937–953 (2014).
12. Basso, S., Schirmer, M. & Botter, G. A physically based analytical model of flood frequency curves. *Geophys. Res. Lett.* **43**, 9070–9076 (2016).
13. Rogger, M. et al. Step changes in the flood frequency curve: process controls. *Water Resour. Res.* **48**, W05544 (2012).
14. Rogger, M., Viglione, A., Derx, J. & Blöschl, G. Quantifying effects of catchments storage thresholds on step changes in the flood frequency curve. *Water Resour. Res.* **49**, 6946–6958 (2013).

15. Rao, M. P. et al. Seven centuries of reconstructed Brahmaputra River discharge demonstrate underestimated high discharge and flood hazard frequency. *Nat. Commun.* **11**, 6017 (2020).

16. Sornette, D. & Ouillon, G. Dragon-kings: mechanisms, statistical methods and empirical evidence. *Eur. Phys. J. Spec. Top.* **205**, 1–26 (2012).

17. Basso, S., Botter, G., Merz, R. & Miniussi, A. PHEV! The physically-based extreme value distribution of river flows. *Environ. Res. Lett.* **16**, 124065 (2021).

18. St. George, S., Hefner, A. M. & Avila, J. Paleofloods stage a comeback. *Nat. Geosci.* **13**, 766–768 (2020).

19. Blöschl, G. et al. Current European flood-rich period exceptional compared with past 500 years. *Nature* **583**, 560–566 (2020).

20. Wilhelm, B. et al. Impact of warmer climate periods on flood hazard in the European Alps. *Nat. Geosci.* **15**, 118–123 (2022).

21. Huntingford, C. et al. Potential influences on the United Kingdom's floods of winter 2013/14. *Nat. Clim. Change* **4**, 769–777 (2014).

22. Yin, J. et al. Large increase in global storm runoff extremes driven by climate and anthropogenic changes. *Nat. Commun.* **9**, 4389 (2018).

23. Blöschl, G. et al. Changing climate both increases and decreases European river floods. *Nature* **573**, 108–111 (2019).

24. Biswal, B. & Marani, M. Geomorphological origin of recession curves. *Geophys. Res. Lett.* **37**, L24403 (2010).

25. Botter, G., Basso, S., Rodriguez-Iturbe, I. & Rinaldo, A. Resilience of river flow regimes. *Proc. Natl Acad. Sci. USA* **110**, 12925–12930 (2013).

26. Mutzner, R. et al. Geomorphic signatures on Brutsaert base flow recession analysis. *Water Resour. Res.* **49**, 5462–5472 (2013).

27. Rodriguez-Iturbe, I. & Rinaldo, A. *Fractal River Basins: Chance and Self-Organization* (Cambridge Univ. Press, 1997).

28. Müller, M. F., Roche, K. R. & Dralle, D. N. Catchment processes can amplify the effect of increasing rainfall variability. *Environ. Res. Lett.* **16**, 084032 (2021).

29. Basso, S., Frascati, A., Marani, M., Schirmer, M. & Botter, G. Climatic and landscape controls on effective discharge. *Geophys. Res. Lett.* **42**, 8441–8447 (2015).

30. Doulatyari, B. et al. Predicting streamflow distributions and flow duration curves from landscape and climate. *Adv. Water Resour.* **83**, 285–298 (2015).

31. Basso, S., Schirmer, M. & Botter, G. On the emergence of heavy-tailed streamflow distributions. *Adv. Water Resour.* **82**, 98–105 (2015).

32. Kusumastuti, D. I., Struthers, I., Sivapalan, M. & Reynolds, D. A. Threshold effects in catchment storm response and the occurrence and magnitude of flood events: implications for flood frequency. *Hydrol. Earth Syst. Sci.* **11**, 1515–1528 (2007).

33. Reager, J., Thomas, B. & Famiglietti, J. River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nat. Geosci.* **7**, 588–592 (2014).

34. Slater, L. J. & Villarini, G. Recent trends in U.S. flood risk. *Geophys. Res. Lett.* **43**, 428–436 (2016).

35. Basso, S., Ghazanchaei, Z. & Tarasova, L. Characterizing hydrograph recessions from satellite-derived soil moisture. *Sci. Total Environ.* **756**, 143469 (2021).

36. Brunner, M. I. et al. An extremeness threshold determines the regional response of floods to changes in rainfall extremes. *Commun. Earth Environ.* **2**, 173 (2021).

37. El Adlouni, S., Bobée, B. & Ouarda, T. B. M. J. On the tails of extreme event distributions in hydrology. *J. Hydrol.* **355**, 16–33 (2008).

38. Dethier, E. N., Sartain, S. L., Renshaw, C. E. & Magilligan, F. J. Spatially coherent regional changes in seasonal extreme streamflow events in the United States and Canada since 1950. *Sci. Adv.* **6**, eaba5939 (2020).

39. Wing, O. E. J. et al. Inequitable patterns of US flood risk in the Anthropocene. *Nat. Clim. Change* **12**, 156–162 (2022).

40. Cornwall, W. Europe's deadly floods leave scientists stunned. *Science* **373**, 372–373 (2021).

Methods

PHEV distribution of river flows

PHEV[12,17] is a mechanistic–stochastic characterization of the magnitude and probability of streamflow maxima occurring in a given reference period (for example, a season). It results from a well-established mathematical description of catchment-scale daily precipitation, soil moisture and runoff dynamics[41–45], which has been proved suitable for a wide array of physioclimatic conditions[46–54]. This framework describes precipitation as a marked Poisson process with frequency λ_p $(1/T)$ and exponentially distributed depth with average α (L), where T stands for time and L for length. Soil moisture increases due to precipitation infiltration and decreases as a result of evapotranspiration, which is a linear function of soil moisture between the wilting point and a critical upper threshold. Exceedance of this threshold triggers runoff pulses with frequency $\lambda < \lambda_p$ $(1/T)$ and exponentially distributed magnitude with average α (L). These pulses recharge a single catchment storage, which finally drains into the stream network. A nonlinear storage–discharge relation mimics the hydrological response and the related hydrograph recessions, which are described through the coefficient K (L^{1-a}/T^{2-a}) and exponent a of a power law function[55]. The summarized mechanistic–stochastic description of runoff generation processes enables expressing the probability distributions of daily flows[45], peak flows[12] (that is, local flow peaks occurring as a result of runoff-producing rainfall events) and flow maxima[12] (that is, maximum values in a specified timespan) as a function of a few physically meaningful parameters (α, λ, a, K).

We directly computed three parameters of PHEV (α, λ, a) from daily rainfall and streamflow series: α is the mean precipitation during rainy days; λ is the ratio between the mean specific river discharge $\bar{q}$ (L/T) and α; and a is the median value of the exponents of power law functions fitted to observed $dq/dt - q$ pairs of single hydrograph recessions[24], where dq/dt are the first derivatives in time (t) of the river discharge q. K is instead obtained via maximum likelihood estimation on the observed seasonal maxima[12].

Data

Two datasets are used in this work with distinct objectives. Both of them were analysed on a seasonal basis (spring: March to May; summer: June to August; autumn: September to November; winter: December to February). A case study represents a given catchment during one season. The first set of data, named study dataset (Extended Data Fig. 1 and Extended Data Table 1), includes 101 case studies[56] across the United States (from the MOPEX dataset[57,58]) and Germany[59]. These case studies were selected as they are characterized by observational records at least 30 years long, limited anthropogenic streamflow disturbance caused by reservoirs and human water uses[58,60], and modest snowfall (that is, the average daily temperature is above 0 °C for the majority of instances in each season and for most years) precluding intense snow accumulation and melting processes[50,53]. They also exhibit hydrograph recession coefficients (that is, the coefficients of power law functions with exponent set equal to a fitted to observed $dq/dt - q$ pairs of single hydrograph recessions) that do not consistently decrease with increasing flow magnitudes[17]. These case studies comply with key hypotheses of the adopted theoretical framework[12], thus enabling a rigorous investigation of physical controls on the emergence of flood divides. The second set of data, termed test dataset (2,519 additional case studies; Extended Data Fig. 1 and Extended Data Table 1), consists instead of watersheds from the MOPEX and Germany that do not necessarily fulfil the above requirements. The only two criteria used for selecting them are the limited anthropogenic disturbance on streamflow[58,60] and a minimum length of the observational series equal to 10 years. The test dataset constitutes a separated set of case studies to stress test the capability of the physical variables identified as explanatory of the magnitude of flood divides to predict the emergence of these features in the test catchments.

Identification of flood divides

We applied a robust methodology[56] to detect flood divides in the study dataset, the steps of which are summarized in the following. We identified the point of maximum curvature[14] of the semi-logarithmic relation between the inverse of the exceedance cumulative probability of flow maxima and its normalized magnitude (that is, magnitude divided by the long-term mean river discharge $\bar{q}$). We estimated this relation, which is commonly known as the flood magnitude–frequency curve, both empirically via Weibull plotting position[61,62] of the observations and by means of PHEV. In the former case, the curvature fluctuates, as it is computed on a discrete set of unevenly distributed points. We thus applied a heuristic rule to remove noise and further consider as potential flood divides only observations on the right-hand side of the last point whose second derivative exceeds the range of twice the standard deviation of the curvature itself[63]. We then used the Mann–Whitney U test[64] to evaluate statistical difference (at the 0.05 significance level) of the distributions of first derivatives before and after each potential flood divide. Additionally, we assessed whether this difference is substantial by computing an effect size by means of the Cohen's d[65,66] and the relative increase of the slope of PHEV within the observational range. Increments of the flood magnitude beyond the flood divide are finally considered relevant if the Cohen's d for the point with minimum P value of the Mann–Whitney U test is higher than 0.4, a value that indicates a moderate effect size[67,68], and the slope increment exceeds 1%. The red circle in Fig. 1b provides an example of flood divide identified through this procedure.

If we identified a flood divide from both empirical and PHEV estimates, we labelled it a true positive (TP) case. Conversely, if both observations and PHEV suggested the absence of a flood divide, we labelled the case as a true negative (TN). When we detected a flood divide from PHEV but not from the observations, we termed it a false positive (FP). If instead PHEV did not signal the existence of a flood divide, which we however identified from the observations, we labelled it a false negative (FN) case. Application of this whole procedure to the study dataset yields true flood divides for 27 case studies, which are displayed in Extended Data Fig. 2.

Dimensional analysis

Starting from a physically meaningful law, the Pi theorem[69] enables reducing the variables of a problem by arranging them into dimensionless groups that help reveal the actual physical controls of the problem[70]. In particular, the Pi theorem states that if we can hypothesize a mechanistic relation involving n physical variables with k independent fundamental dimensions, we can rewrite it in terms of $p = n - k$ dimensionless groups. We postulated that PHEV outlines the pivotal relations among physioclimatic variables that control the hydrological response of river basins and the occurrence of floods. We then leveraged the Pi theorem to unveil physical controls on the shape of the flood magnitude–frequency curve and hence on the magnitude of the flood divide, which are embodied by the dimensionless groups identified through the Pi theorem. The latter groups may differ depending on the hypothesized mechanistic relation and the variables considered in the analysis. Therefore, we validated the above hypothesis and the relevance of the resulting dimensionless groups against observations. We finally tested the predictive power of the identified physical controls in a large set of case studies.

The variables in this case are the normalized magnitude of the flood divide $q/\bar{q}$ (that is, flow magnitude q divided by the long-term mean river discharge $\bar{q}$), the effective rainfall frequency λ $(1/T)$, the average rainfall magnitude α (L), and the hydrograph recession exponent a and coefficient K (L^{1-a}/T^{2-a}). The overall number of variables is $n = 5$. The number of fundamental dimensions is instead $k = 2$ (that is, (L) and (T)). We rearranged the five variables into $p = 3$ dimensionless groups, which are the two dimensionless variables themselves $(q/\bar{q}$ and $a)$, and a combination of the variables that encompasses all the remaining ones and suitably yields a dimensionless group (that is,

$K\lambda^{a-2}\alpha^{a-1}$). We thus identified an expression for the normalized magnitude of the flood divide that reads as: $q/\bar{q} = f(a, K\lambda^{a-2}\alpha^{a-1})$. The second dimensionless group on the right-hand side of the equation is the squared coefficient of variation of daily flows[12,25]. To stress its physical origin, we also express it as the ratio between the mean interarrival of effective rainfall events, $1/\lambda$ (*T*), and the characteristic response time of the basin, $1/K(\alpha\lambda)^{a-1}$ (*T*) (refs. [52,71]). The Pi theorem thus indicates the hydrograph recession exponent and the coefficient of variation of daily flows as the physical controls of the magnitude of the flood divide.

Relations between magnitude of the flood divide and its physioclimatic controls

We proceeded as follows to determine the theoretical relations between the normalized magnitude of the flood divide and its geomorphological and hydroclimatic controls (blue envelopes in Fig. 2). We fitted an exponential function in the form $y = \alpha_i \cdot \exp(\beta_i x_i) + \gamma_i$ to results from PHEV, where the dependent variable y is the normalized magnitude of the flood divide $q/\bar{q}$ estimated by means of PHEV, x_i is either the hydrograph recession exponent (Fig. 2a) or the observed coefficient of variation of daily flows (Fig. 2b) and i labels either of these two cases. We thus obtained the optimal parameters (that is, those for which the sum of the residuals is minimized) and their standard deviations. We finally determined the theoretical relations and their uncertainties (blue envelopes in Fig. 2) by plotting the exponential functions with the sets of parameters that encompass the 95% variability of theoretical predictions for the set of case studies.

Statistics

We used the non-parametric two-sample and two-sided Kolmogorov–Smirnov test[72] to determine whether the ratios between the highest observed flood and the mean maximum seasonal flood for basins with or without flood divides are drawn from the same probability distribution. We applied the test to compare both cases in the study dataset for which empirical and PHEV estimates of the flood divide provided consistent results (inset of Fig. 1c), and cases in the test dataset for which we predicted either the presence or the absence of a flood divide by means of binomial logistic regression of its two identified physioclimatic controls (Fig. 3 and Extended Data Fig. 4b).

We used distance correlation[73] to quantify the strength of the observed relations between normalized magnitude of the flood divide and its physioclimatic controls, as well as between observed and theoretical magnitudes of the flood divide (Extended Data Fig. 2). Distance correlation is a measure of multivariate dependence between random vectors, which is defined, analogously to the Pearson correlation coefficient, as the ratio between their distance covariance and the product of their distance standard deviations. It varies between 1 and 0, with the latter value indicating that the variables are independent.

Binary logistic regression

We used logistic regression to predict whether a flood divide may arise or not in a river basin, by considering the hydrograph recession exponent and the coefficient of variation of daily flows as explanatory variables. Logistic regression is a statistical tool that uses a logistic function to model a binary outcome[74,75], which in this study is the occurrence/non-occurrence of a flood divide. In mathematical terms, let us consider a linear relationship (with coefficients β_0, β_1, β_2) between two predictors (x_1, x_2) and the log-odds *l* (logit) of the event $Y = 1$, with Y being a Bernoulli distributed variable: $l = \log(P/(1-P)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. The probability that $Y = 1$ is thus: $P = \frac{1}{1+\exp^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}} = S_e(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$ where S_e is the sigmoid function with base *e*. We set the cutoff threshold for assigning values of *P* to either class 0 (no flood divide expected) or 1 (flood divide expected) at 0.75, meaning that if we predict a probability lower than 0.75 the case study is allocated to class 0 and vice versa. We determined this value by means of the Youden's statistic computed for the study dataset[76]. We also controlled for collinearity of the

explanatory variables by computing the variable inflation factor[77], which is equal to 1.62. Provided that removal of correlated variables is typically recommended[77,78] for a variable inflation factor >5–10, we retained both the hydrograph recession exponent and the coefficient of variation of daily flows as explanatory variables.

We trained the binary logistic regression by using the true cases in the study dataset, that is, those for which both PHEV and the observations indicate either the presence (true positives, 27 cases) or the absence (true negatives, 7 cases) of a flood divide[56]. We then evaluated the predictions in a twofold fashion. We first applied a cross-validation procedure, randomly extracting for 100 times two-thirds of the 34 true cases (true positives plus true negatives) for fitting the parameters of the logistic model and using the remaining one-third of cases to evaluate the accuracy of the predictions (Extended Data Fig. 3). We later adopted a separated extended dataset (the test dataset; Extended Data Fig. 1) to evaluate the prediction performance under broader conditions. In this case we fitted the parameters of the logistic model on the whole set of 34 true cases identified in the study dataset, randomly extracted for 1,000 times 34 case studies from the test dataset to match the number of case studies used for training the binary logistic regression, and evaluated the performance each time (Extended Data Fig. 4a).

We employed two performance metrics, namely the balanced accuracy and the MCC to evaluate the accuracy of our predictions. The balanced accuracy[79] is a class-wise weighted accuracy rate computed as the arithmetic mean of sensitivity (true positive rate) and specificity (true negative rate). It is recommended when one class (true negatives in this work) is underrepresented in the dataset (that is, imbalanced dataset). The balanced accuracy ranges between 0 and 1, with values lower than 0.5 indicating a worse performance than a random classifier. The MCC[80,81] is a metric unaffected by biases when considering imbalanced datasets[82]. It is defined as $\text{MCC} = \frac{\text{TP}\times\text{TN} - \text{FP}\times\text{FN}}{\sqrt{(\text{TP}+\text{FP})\times(\text{TP}+\text{FN})\times(\text{TN}+\text{FP})\times(\text{TN}+\text{FN})}}$. The MCC ranges between −1 (complete disagreement between predictions and observations) and +1 (perfect prediction), with 0 indicating that the model performs as well as a random classifier. MCC is equivalent to the Pearson correlation coefficient in the special case of two binary variables (that is, predictions and observations). Analyses have been performed with the Python scikit-learn package, version 0.24.2[83].

Data availability

Hydroclimatic data for the United States utilized in this study are available from the MOPEX dataset, https://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data. Daily precipitation for river basins in Germany is obtained from the gridded REGNIE dataset, https://opendata.dwd.de/climate_environment/CDC/grids_germany/daily/regnie/. Daily gridded temperature[84] is available at http://www.ufz.de/index.php?en=41160. Daily streamflow series for river basins in Germany are available from the Global Runoff Data Center of the Federal Institute of Hydrology (https://www.bafg.de/GRDC) and from the federal state water authorities: Baden-Württemberg (https://udo.lubw.baden-wuerttemberg.de), Bavaria (https://www.gkd.bayern.de/de/fluesse/abfluss), Brandenburg (https://pegelportal.brandenburg.de/start.php#loaded), Hesse (https://www.hlnug.de/themen/wasser/wasserstaende-und-abfluesse/pegelmessnetz), Lower Saxony (https://www.nlwkn.niedersachsen.de/startseite/wasserwirtschaft/flusse_bache_seen/fliessgewasserpegel/pegelverzeichnis/pegelverzeichnis-159525.html), Mecklenburg-Vorpommern (http://pegelportal-mv.de/pegel-mv/pegel_mv.html), North Rhine-Westphalia (https://www.lanuv.nrw.de/umwelt/wasser/oberflaechengewaesserfluesse-und-seen), Rhineland-Palantine (https://wasserportal.rlp-umwelt.de/servlet/is/8122/), Saarland (https://www.saarland.de/muv/DE/portale/wasser/informationen/hochwassermeldedienst/wasserstaende_warnlage/wasserstaende_warnlage_node.html), Saxony (https://www.umwelt.sachsen.de/umwelt/infosysteme/hwims/portal/web/download-von-messwerten), Saxony-Anhalt (https://gld.lhw-sachsen-anhalt.de/#), Schleswig-Holstein (https://umweltportal.

schleswig-holstein.de/trefferanzeige?docuuid=1f7866d9-c39c-4e55-823d-3c4faf7d299c) and Thuringia (https://tlubn.thueringen.de/wasser/fluesse-baeche/fluesse-und-baeche-wassermenge). Thirty-year precipitation normals for the United States provided by the PRISM Climate Group, Oregon State University, are available at http://prism.oregonstate.edu (downloaded on 1 June 2021). Thirty-year precipitation normals for Germany provided by the Deutsche Wetter Dienst are available at https://opendata.dwd.de/climate_environment/CDC/grids_germany/multi_annual/precipitation.

Code availability

The code used to perform the analyses and produce the figures for this study is available at https://www.hydroshare.org/resource/a6bcc341413c4fb0b195b25ebe1bb3e6/.

References

41. Rodriguez-Iturbe, I., Porporato, A., Ridolfi, L., Isham, V. & Cox, D. R. Probabilistic modelling of water balance at a point: the role of climate, soil and vegetation. *Proc. R. Soc. A* **455**, 3789–3805 (1999).
42. Laio, F., Porporato, A., Ridolfi, L. & Rodriguez-Iturbe, I. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: II. Probabilistic soil moisture dynamics. *Adv. Water Resour.* **24**, 707–723 (2001).
43. Porporato, A., Daly, E. & Rodriguez-Iturbe, I. Soil water balance and ecosystem response to climate change. *Am. Nat.* **164**, 625–632 (2004).
44. Botter, G., Porporato, A., Rodriguez-Iturbe, I. & Rinaldo, A. Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: slow, leaching-prone components of the hydrologic response. *Water Resour. Res.* **43**, W02417 (2007).
45. Botter, G., Porporato, A., Rodriguez-Iturbe, I. & Rinaldo, A. Nonlinear storage–discharge relations and catchment streamflow regimes. *Water Resour. Res.* **45**, W10427 (2009).
46. Botter, G., Peratoner, F., Porporato, A., Rodriguez-Iturbe, I. & Rinaldo, A. Signatures of large-scale soil moisture dynamics on streamflow statistics across U.S. climate regimes. *Water Resour. Res.* **43**, W11413 (2007).
47. Botter, G., Basso, S., Porporato, A., Rodriguez-Iturbe, I. & Rinaldo, A. Natural streamflow regime alterations: damming of the Piave River basin (Italy). *Water Resour. Res.* **46**, W06522 (2010).
48. Ceola, S. et al. Comparative study of ecohydrological streamflow probability distributions. *Water Resour. Res.* **46**, W09502 (2010).
49. Pumo, D., Viola, F., La Loggia, G. & Noto, L. V. Annual flow duration curves assessment in ephemeral small basins. *J. Hydrol.* **519**, 258–270 (2014).
50. Schaefli, B., Rinaldo, A. & Botter, G. Analytic probability distributions for snow-dominated streamflow. *Water Resour. Res.* **49**, 2701–2713 (2013).
51. Mejía, A., Daly, E., Rossel, F., Jovanovic, T. & Gironás, J. A stochastic model of streamflow for urbanized basins. *Water Resour. Res.* **50**, 1984–2001 (2014).
52. Müller, M. F., Dralle, D. N. & Thompson, S. E. Analytical model for flow duration curves in seasonally dry climates. *Water Resour. Res.* **50**, 5510–5531 (2014).
53. Santos, A. C., Portela, M. M., Rinaldo, A. & Schaefli, B. Analytical flow duration curves for summer streamflow in Switzerland. *Hydrol. Earth Syst. Sci.* **22**, 2377–2389 (2018).
54. Arai, R., Toyoda, Y. & Kazama, S. Runoff recession features in an analytical probabilistic streamflow model. *J. Hydrol.* **597**, 125745 (2021).
55. Brutsaert, W. & Nieber, J. L. Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resour. Res.* **13**, 637–643 (1997).
56. Miniussi, A., Merz, R., Kaule, L. & Basso, S. Identifying discontinuities of flood frequency curves. *J. Hydrol.* **617**, 128989 (2023).
57. Schaake, J. et al. The Model Parameter Estimation Experiment (MOPEX). *J. Hydrol.* **320**, 1–2 (2006).
58. Wang, D. & Hejazi, M. Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States. *Water Resour. Res.* **47**, W00J12 (2011).
59. Tarasova, L., Basso, S., Zink, M. & Merz, R. Exploring controls on rainfall-runoff events: 1. Time series-based event separation and temporal dynamics of event runoff response in Germany. *Water Resour. Res.* **54**, 7711–7732 (2018).
60. Lehner, B. et al. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Front. Ecol. Environ.* **9**, 494–502 (2011).
61. Weibull, W. A statistical theory of strength of materials. *Ing. Vetensk. Akad. Handl.* **151**, 1–45 (1939).
62. Makkonen, L. Plotting positions in extreme value analysis. *J. Appl. Meteorol. Climatol.* **45**, 334–340 (2006).
63. Tchebichef, P. Des valeurs moyennes. *J. Math. Pures Appl.* **2**, 177–184 (1867).
64. Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**, 50–60 (1947).
65. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Erlbaum Associates, 1974).
66. Sullivan, G. & Feinn, R. Using effect size – or why the p-value is not enough. *J. Grad. Med. Educ.* **4**, 279–282 (2012).
67. Gignac, G. E. & Szodorai, E. T. Effect size guidelines for individual differences researchers. *Pers. Individ.* **102**, 74–78 (2016).
68. Lovakov, A. & Agadullina, E. R. Empirically derived guidelines for effect size interpretation insocialpsychology. *Eur. J. Soc. Psychol.* **51**, 485–504 (2021).
69. Buckingham, E. On physically similar systems; illustrations of the use of dimensional equations. *Phys. Rev.* **4**, 345–376 (1914).
70. Porporato, A. Hydrology without dimensions. *Hydrol. Earth Syst. Sci.* **26**, 355–374 (2022).
71. Deal, E., Braun, J. & Botter, G. Understanding the role of rainfall and hydrology in determining fluvial erosion efficiency. *J. Geophys. Res.* **123**, 744–778 (2018).
72. Smirnov, N. V. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bull. Math. Univ. Moscou* **2**, 3–14 (1939).
73. Szekely, G. J., Rizzo, M. L. & Bakirov, N. K. Measuring and testing dependence by correlation of distances. *Ann. Stat.* **35**, 2769–2794 (2007).
74. Cox, D. R. *The Analysis of Binary Data* (Methuen, 1970).
75. Hosmer, D. W. & Lemeshow, S. *Applied Logistic Regression* (Wiley, 1989).
76. Fernández, A. et al. *Learning from Imbalanced Data Sets* (Springer, 2018).
77. Kutner, M. H., Nachtsheim, C. J. & Neter, J. *Applied Linear Regression Models* (McGraw-Hill/Irwin, 2004).
78. Sheather, S. *A Modern Approach to Regression with R* (Springer, 2009).
79. Velez, D. R. et al. A balanced accuracy function for epistasis modeling in imbalanced datasets using multifactor dimensionality reduction. *Genet. Epidemiol.* **31**, 306–315 (2007).
80. Matthews, B. W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta Protein Struct.* **405**, 442–451 (1975).
81. Lever, J., Krzywinski, M. & Altman, N. Classification evaluation. *Nat. Methods* **13**, 603–604 (2016).

82. Chicco, D. & Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **21**, 6 (2020).

83. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

84. Zink, M., Kumar, R., Cuntz, M. & Samaniego, L. A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty. *Hydrol. Earth Syst. Sci.* **21**, 1769–1790 (2017).

Author contributions

S.B. and A.M. conceived and designed the study with input from R.M. L.T. and R.M. assembled the German datasets. A.M. performed the investigations with contributions from S.B. L.T. provided support in the choice of the statistical methods. S.B., A.M. and R.M. interpreted the results, with input from L.T. S.B. and A.M. wrote the first draft of the paper. All authors contributed to framing and revising the paper. S.B. and A.M. contributed equally to the work.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41561-023-01155-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41561-023-01155-w.

Correspondence and requests for materials should be addressed to Stefano Basso.

Peer review information *Nature Geoscience* thanks Basudev Biswal, Cedric David and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Tom Richardson, in collaboration with the *Nature Geoscience* team.

Reprints and permissions information is available at www.nature.com/reprints.

**Extended Data Table 1 | Summary of watersheds, case studies and flood divides in the study and test datasets**

		Watersheds	Case studies	Observed flood divides		Predicted flood divides			
				Presence	Absence	TP	TN	FP	FN
Study dataset	*USA*	41	101	27	7	-	-	-	-
	Germany	32							
Test dataset	*USA*	252	2519	-	-	531	728	1133	127
	Germany	403							

Observed flood divides (presence or absence thereof) in the study dataset are those for which empirical and theoretical estimates provided consistent results. Predicted flood divides in the test dataset based on binary logistic regression of their physioclimatic controls. TP: true positives, TN: true negatives, FP: false positives, FN: false negatives.
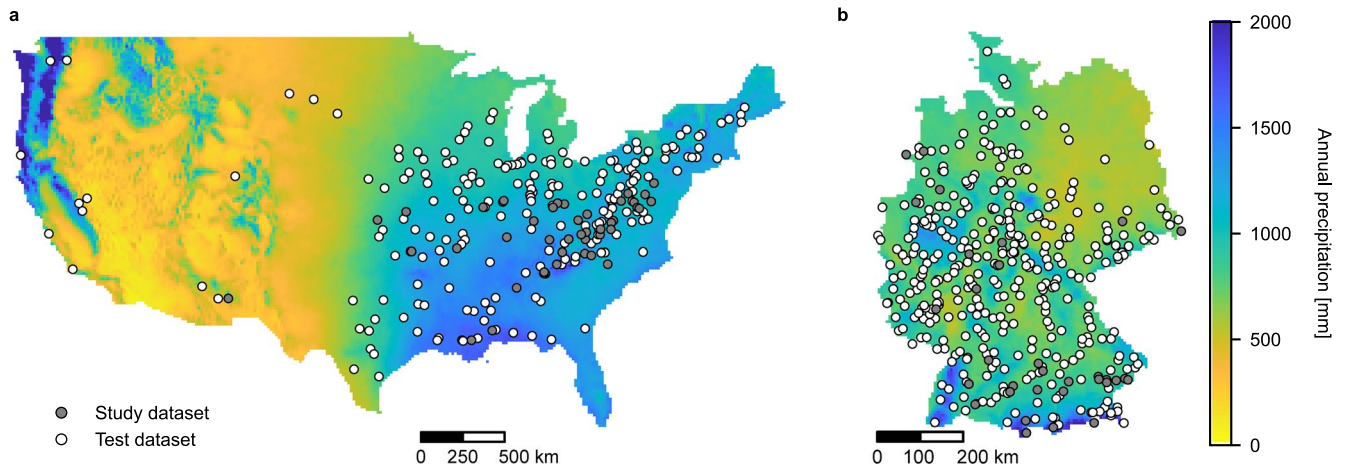
Extended Data Fig. 1 | Datasets. Outlet locations of river basins in **a**, the continental United States and **b**, Germany employed as case studies. Gray dots indicate watersheds in the study dataset (Methods; n = 41 in the United States and n = 32 in Germany). White dots represent watersheds in the test dataset (Methods; n = 252 in the United States and n = 403 in Germany). Colorbar shows 30-years annual precipitation normals (USA: 1981–2010; Germany: 1991–2020).
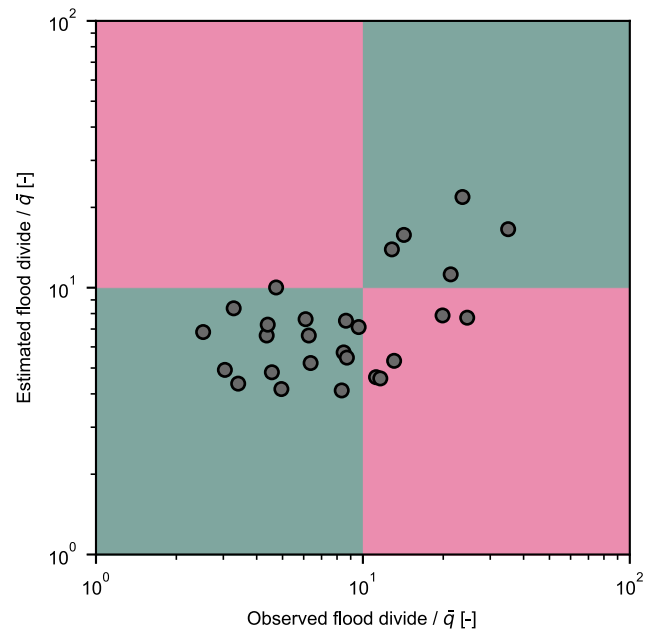
Extended Data Fig. 2 | Estimated versus observed magnitudes of the flood divides. Dots represent case studies (n = 27) for which we identified a flood divide both from observations and through PHEV. Magnitudes are normalized with respect to the long-term mean river discharge $\bar{q}$. Green (magenta) areas encompass theoretical and empirical estimates with same (different) orders of magnitude. The distance and Spearman correlation coefficients between observed and PHEV estimates are 0.64 and 0.44 (p < 0.01).
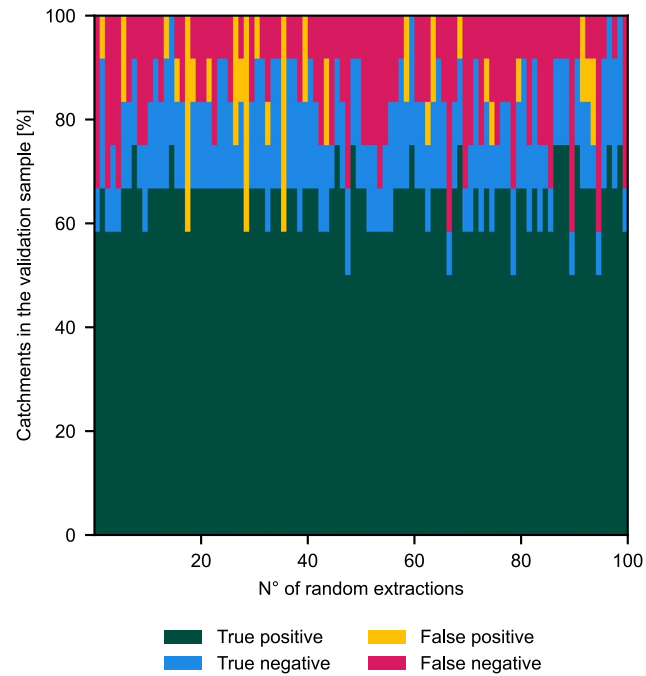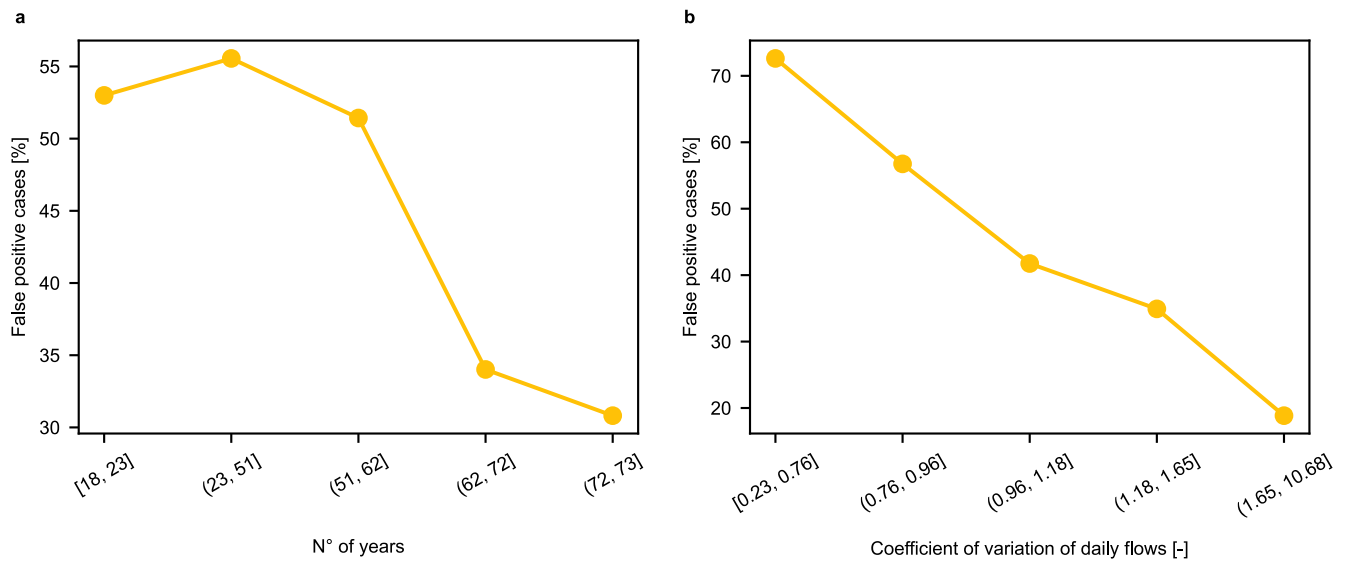
Extended Data Fig. 3 | Prediction of flood divides in the study dataset from their physioclimatic controls. Accuracy in inferring presence or absence of flood divides in a cross-validation experiment performed with the study dataset (Methods). True positives (green) and negatives (blue) are cases for which we respectively anticipate and rule out the occurrence of flood divides, as confirmed by data records. False positives (yellow) are cases for which we signal the existence of flood divides which are not identified from observations. False negatives (magenta) are cases where data records display flood divides which are unforeseen by the identified physical controls.
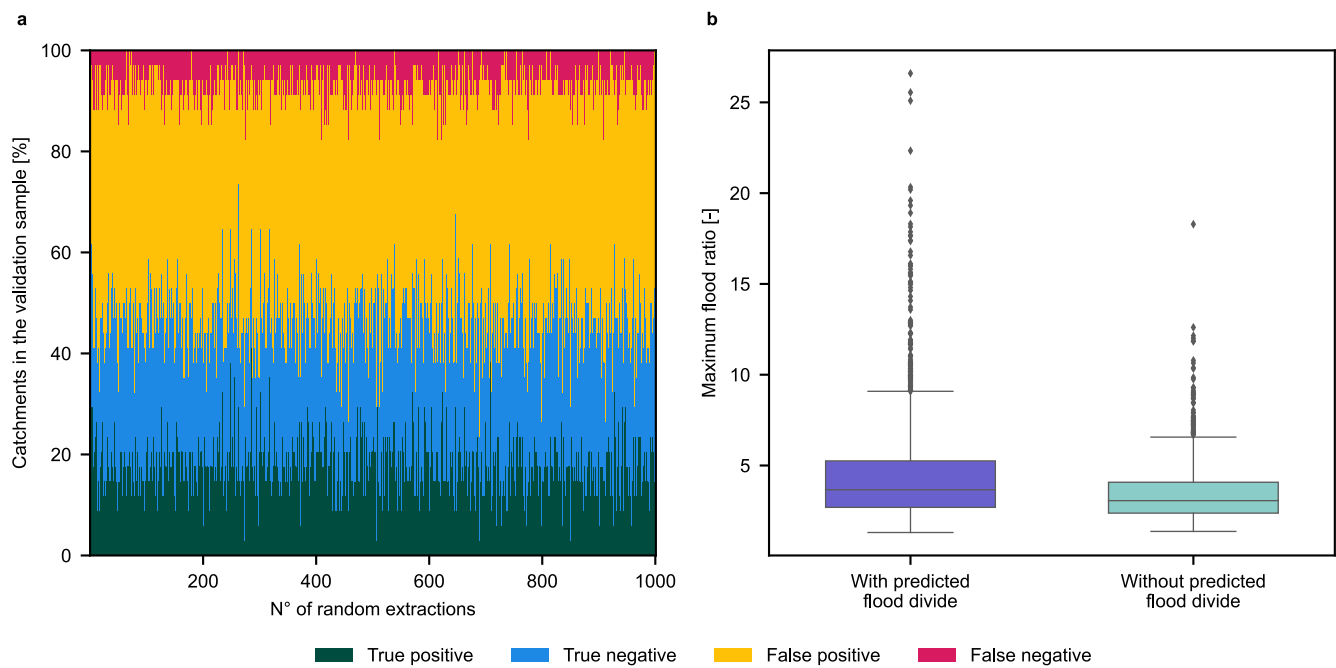
Extended Data Fig. 4 | Prediction of flood divides in the test dataset from their physioclimatic controls. a, Accuracy in inferring presence or absence of flood divides in the test dataset (Methods). **b**, Ratios between highest observed and mean maximum seasonal flood for case studies in the test dataset for which we predict the presence (true and false positives, violet, n = 1664) and the absence (true and false negatives, aquamarine, n = 855) of a flood divide (center line: median; box limits: 25th and 75th percentiles; whiskers: minimum and maximum values that are not considered outliers, that is, 1.5x interquartile range; dots: outliers). The distributions are significantly different (two-sided Kolmogorov-Smirnov test, p = 6*10^{-15}).

a

b

Extended Data Fig. 5 | Detection of flood divides versus data length and streamflow variability. Percentage of false positive cases for increasing **a**, number of years in the data record used for validation and **b**, coefficient of variation of daily flows. Ranges (whose boundaries are reported in the x-axis) were set to have approximately equal numbers of case studies ($n = 504$) for each bin.