nature water

Article

https://doi.org/10.1038/s44221-026-00699-6

A global classification of hydrologic functional diversity in gauged and ungauged catchments

Received: 22 September 2025

Accepted: 23 July 2026

Published online: 11 September 2026

Check for updates

Ali A. Ameli [1] ✉, Hamed Sharif [1] & Jeffrey J. McDonnell [2,3,4]

Most of the world's streams remain ungauged—a longstanding hydrologic challenge. Here we develop a globally scalable, seasonally resolved framework that classifies catchments according to the complexity of how rainfall is converted into runoff across events, yielding three functional types: simple, intermediate and complex. Applied to over 80,000 gauged and ungauged catchments, it reveals that functional complexity is the global norm: catchments draining 87% of the evaluated ungauged land area are classified as complex in the dormant season, with the number of complex catchments increasing by 63% in the growing season. Africa and much of Asia (except Japan) remain predominantly complex year round, Europe reaches 94% complex in the growing season, the US Pacific Northwest and Canada's British Columbia retain simple clusters, and southeastern South America and southeastern Australia show notable transitions from complex to intermediate, demonstrating that catchment functional type is not fixed. Greater urban and agricultural land cover, together with less persistent rainfall, are associated with greater functional complexity. These findings provide a blueprint for prioritizing streamflow gauging in under-represented complex catchments and for guiding model structural selection, transferability and prediction in ungauged catchments, ensuring that models are neither simpler than necessary nor more complex than required.

Rainfall-to-runoff transformation is at the heart of today's water-security challenges, from riverine floods and infrastructure failure to contaminant pulses that jeopardize drinking water and aquatic ecosystems[1–4]. A streamflow hydrograph is best viewed as a sequence of distinct runoff pulses separated by smoother baseflow[5]. These short, event-scale pulses most directly capture the dynamic nature of hydrologic response in rain-dominated catchments and are often composed of a mix of overland flow, subsurface stormflow and groundwater contributions across many landscape types[6–8]. They encode information on how rainfall characteristics interact with antecedent wetness, vegetation and catchment physical controls to shape landscape–stream hydrologic connectivity and runoff production[9–14]. The timing and magnitude of these pulses, governed by the rainfall–runoff relationship, determine the severity and onset of floods[15,16], the mobilization of weathering-derived and biogeogenic solutes, contaminants, and sediments into stream networks[17–20], and the timing and magnitude of groundwater recharge[21,22]. Yet despite their central importance, the functional mapping from rainfall to runoff—and how it varies across catchments and over time—remains poorly characterized across much of the globe[13,23–26].

Recent syntheses have advanced understanding by demonstrating rainfall–runoff linkages in a few hundred gauged catchments[27–32]. Yet

[1]Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, British Columbia, Canada. [2]Global Institute for Water Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada. [3]North China University of Water Resources and Electric Power, Zhengzhou, China. [4]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK. ✉e-mail: aameli@eoas.ubc.ca

the greatest challenge lies in ungauged catchments. Gauged catchments represent only a small fraction of global streams and rivers[33], meaning that most of the world's flowing waters remain unmeasured[23]. This data scarcity perpetuates the 'ungauged catchment problem' as one of hydrology's grand challenges[23,24,34]: the need to forecast floods, protect water quality and plan resilient supplies is often most urgent precisely where hydrometric observations are absent. Identifying the rainfall–runoff functional relationship in ungauged catchments is therefore critical for developing models that anticipate floods[35], track contaminant mobilization[20], safeguard ecosystems[36] and guide climate-robust management[37]. It is also critical for advancing hydrologic theory[38,39] and designing effective monitoring networks[40,41].

Across disciplines, classification provides a powerful framework to synthesize diverse knowledge, reveal hidden patterns and generalize beyond individual case studies[42]. Though early discussions in the hydrologic sciences called for functional classification of rainfall–runoff transformation[43–45], progress has been limited, particularly when applied to ungauged catchments. Such functional classification, if it could be done, could inform every stage of analysis in ungauged catchments: guiding model structure, supporting regionalization, optimizing observation network design and strengthening theoretical foundations.

Here, we develop a globally scalable functional classification that categorizes (un)gauged rain-dominated catchments along a gradient of rainfall–runoff response complexity into three functional types (simple, intermediate and complex), evaluated separately for dormant and growing seasons. The seasonal separation is motivated by experimental evidence that rainfall–runoff relationships can differ substantially between dormant and growing periods owing to changes in antecedent moisture, evapotranspiration demand and vegetation activity[46,47]. Building on a new pilot method developed by Sharif and Ameli[30], we first apply the classification to rain-dominated gauged catchments across 62 countries, which are not regulated by large dams (see the Methods for details of study catchments). Catchment areas range from 0.25 km^2 to 232,850 km^2 (median 299 km^2), and the majority of catchments have at least 30 years of daily rainfall and streamflow records (with minimal missing data), typically spanning 1979–2020. We extract rainfall–runoff events from the daily time series (Methods), yielding 2,030,569 events across seasons.

For each catchment and season, we quantify the rainfall–runoff relationship by regressing total event rainfall volume against runoff volume, where strong linear fits indicate simple behaviour, and strong segmented-linear fits capture a threshold-like pattern in the rainfall–runoff relationship (Methods). Following Sharif and Ameli[30], functional classes are assigned to each catchment separately for dormant and growing seasons using goodness-of-fit (R^2) criteria: 'simple' where rainfall volume linearly explains more than 75% of the variance of runoff volume, 'intermediate' where either the linear fit explains 50–75% of variance or the segmented-linear fit explains more than 75%, and 'complex' where neither fit explains more than 50% (linear) or 75% (segmented-linear) of variance. Extended Data Fig. 1 shows examples of these three functional classes. The conceptual underpinning of these R^2 thresholds is grounded in Dooge's[48] theory of catchment functional behaviour, which posits that any catchment's rainfall–runoff response reflects two competing signals: a repeatable (deterministic) component that produces consistent responses to similar rainfall inputs, and a variable (stochastic) component arising from changing antecedent conditions and storm characteristics. The R^2 of the linear model measures the balance between these two components—high R^2 indicates that the repeatable signal dominates (simple behaviour), while low R^2 indicates that event-to-event variability overwhelms the repeatable signal (complex behaviour) (see the Methods for more detail). A jackknife resampling analysis confirms that the classification of gauged catchments is robust to subsampling, demonstrating that the assigned labels are not driven by outliers and are not overly sensitive to the selected R^2 thresholds (Methods; Supplementary Table 4).

Next, we extend the classification to 77,585 ungauged rain-dominated catchments, spanning 97 countries (Methods). In doing so, two season-specific XGBoost machine learning classifiers are trained and tested on the gauged catchments, using their known functional labels as targets. Predictor variables include climatic metrics (aridity index (AI), seasonality index, rainfall persistence indices and effective precipitation (P-AET)), topographic indices (slope, elevation patterns and geomorphic form/shape factors), geologic attributes (lithology type and depth to bedrock) and land cover fractions (cropland, urban, forest and bare earth) (Methods; Supplementary Table 1 and Supplementary Figs. 2–9). Complementary analysis confirms that gauged and ungauged catchments occupy the same densely sampled region of predictor space, and that the XGBoost model predominantly interpolates within its training domain when applied to ungauged catchments (Methods; Supplementary Text 6 and Supplementary Fig. 14).

XGBoost model performance is assessed using k-fold cross-validation, with precision and specificity metrics measuring the recovery of the three functional classes (Methods). Two season-specific models reproduce each catchment's true class with acceptable precision and specificity across seasons, continents and test catchments held out during model training (Methods; Supplementary Tables 2 and 5). Following validation, each season's model is applied to the ungauged catchments: for each catchment, we compute the predictor variables described above using its boundary polygon and predict its functional class using the trained and validated XGBoost model. This exercise provides a proof of concept: given only a boundary polygon, users can infer whether their catchment behaves as (non)complex in dormant and growing seasons, without requiring local streamflow data.

Prevalence of functional complexity across gauged and ungauged catchments globally

Our results show that across gauged catchments, 62% are complex in the dormant season and 75% in the growing season; intermediate catchments account for 34% (dormant) and 23.5% (growing); and simple catchments for just 4% and 1.5%, respectively (Fig. 1 and Extended Data Fig. 2). In North America (dormant), simple catchments cluster in the Pacific Northwest and western Canada, intermediate behaviour dominates the eastern USA and complex systems span the central USA (Fig. 1d). In Europe (dormant), simplicity is largely confined to western and northern UK, intermediate behaviour appears in Ireland, as well as western Germany, and complexity is widespread across France, and central and eastern Germany and Denmark, and southeastern UK (Fig. 1c). Outside the Northern Hemisphere (dormant), Africa, Latin America and the Caribbean and southern Asia are predominantly complex, with small pockets of intermediate or simple systems (for example, in Tasmania, New Zealand, southern South Africa and southeastern South America) (Fig. 1). These patterns persist into the growing season but shift towards greater complexity (Extended Data Fig. 2): many intermediate catchments transition to complex—402 in Europe and 314 in the USA (Extended Data Fig. 3). Conversely, while complexity remains the dominant mode, portions of Latin America and the Caribbean (143 catchments) and Oceania (66 catchments) shift from complex to intermediate (Extended Data Fig. 3). Our first key finding, therefore, is that complex runoff generation is the global norm for gauged rain-dominated catchments, with limited regions exhibiting persistently simple behaviour and some clear seasonal (dormant-to-growing) modulation in transition zones.

Our second key finding is the overwhelming dominance of functional complexity in ungauged catchments. In the dormant season, streams draining 87% of the evaluated ungauged land area (121 million km^2 out of 139 million km^2) are classified as complex, with only 11.5% intermediate and 1.5% simple (Fig. 2). During the growing season,

https://doi.org/10.1038/s44221-026-00699-6

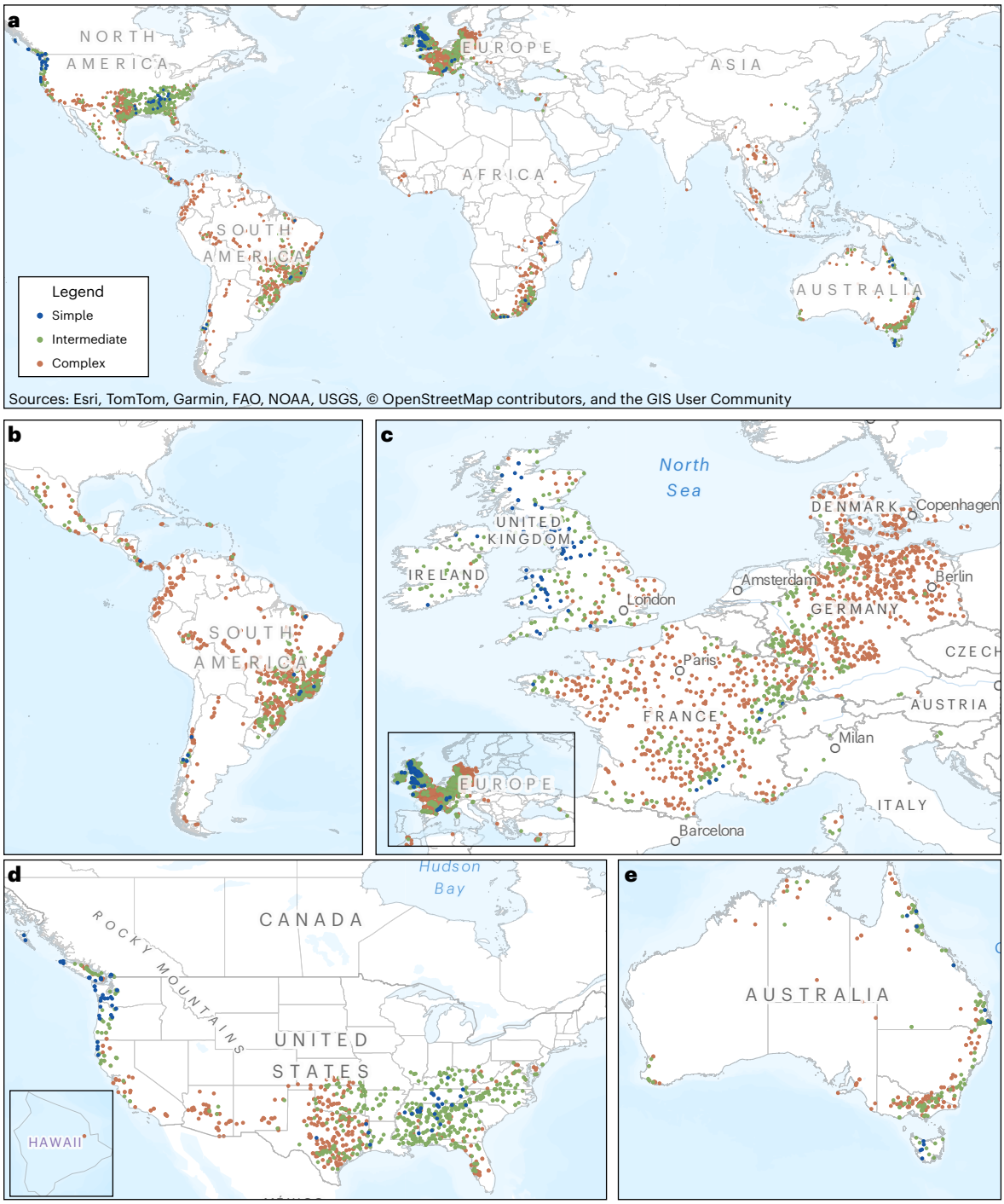


Fig. 1 | Functional complexity prevails in gauged catchments during the dormant season. a, The global distribution of simple (blue), intermediate (green) and complex (red) behavioural classes among gauged catchments. **b**–**e**, Continental insets of functional classes covering Latin America and the Caribbean (**b**), Europe (**c**), North America (**d**) and Oceania (**e**). North America: simple catchments concentrate in the Pacific Northwest and western Canada, intermediate catchments in the eastern USA and complex catchments across the central USA. Europe: simple behaviour is limited to the western and northern UK, intermediate behaviour dominates Ireland and western Germany, and complexity is widespread across France, central and eastern Germany and Denmark, and the southeastern UK. Southern Hemisphere: complex catchments dominate, with pockets of intermediate or simple systems in Tasmania, New Zealand and southeastern South America. Credit: basemaps, Esri, TomTom, Garmin, FAO, NOAA, USGS, © OpenStreetMap contributors and the GIS User Community.

complexity intensifies: at the global scale, the number of catchments labelled as complex increases by 63%, while the number of catchments labelled as simple decreases by 73% (Fig. 3 and Extended Data Fig. 4). Spatial patterns largely mirror those observed in gauged systems. For example, in North America, the Pacific Northwest and western British Columbia (Canada) contain the highest concentration of simple catchments during the dormant season (Fig. 2b). Yet even here, intermediate catchments expand and new complex catchments appear across the states of Washington and Oregon during the growing season (Extended Data Fig. 4b). The eastern USA is predominantly intermediate in the dormant season (Fig. 2a), where 940 catchments transition to complex behaviour during the growing season (Fig. 3 and

https://doi.org/10.1038/s44221-026-00699-6

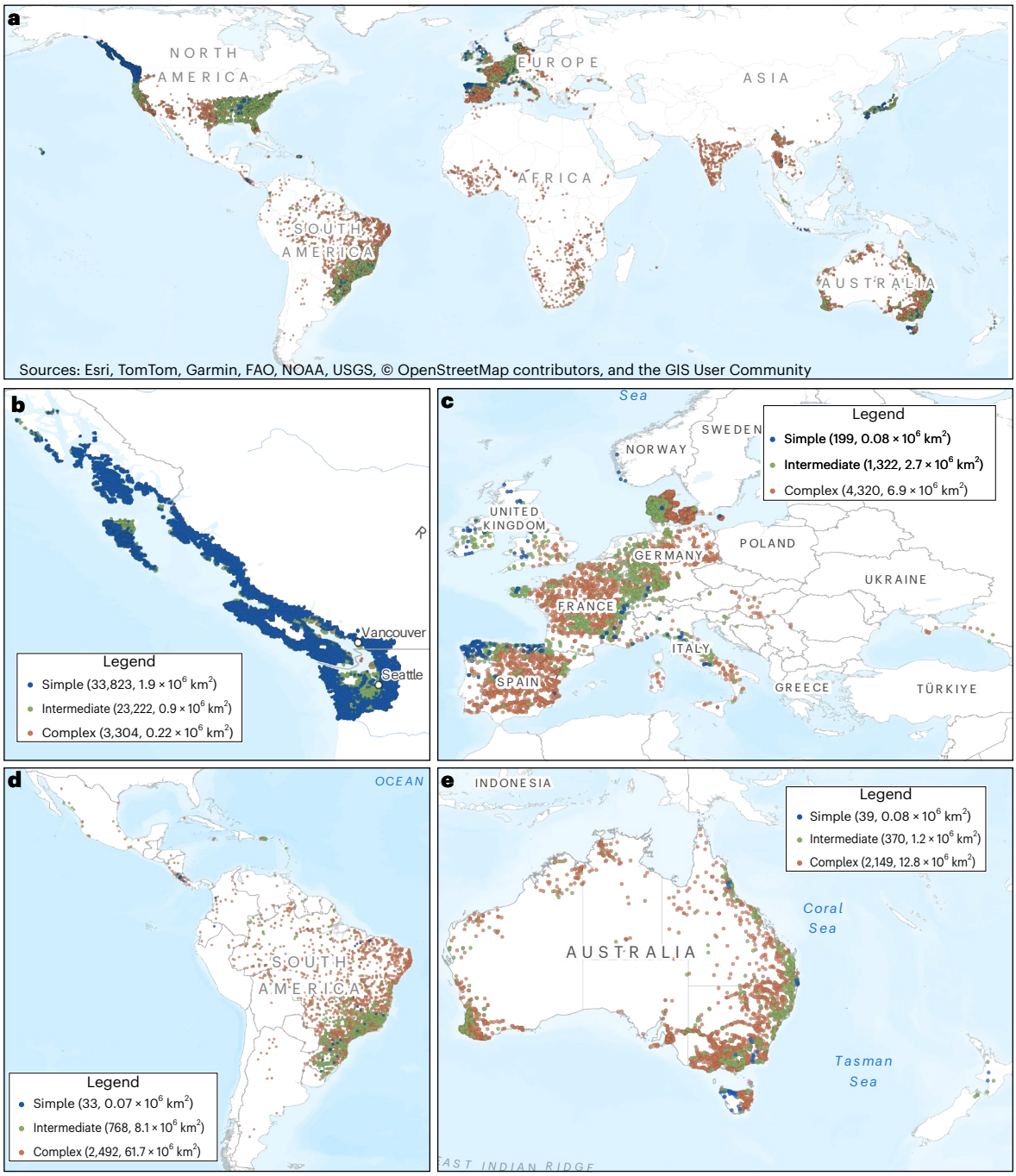


Fig. 2 | Complex catchments drain 121 million km² of ungauged rain-dominated land globally in the dormant season. a, A global map of simple (blue), intermediate (green) and complex (red) classifications for 77,585 ungauged rain-dominated catchments across 97 countries; the legends report the count and total drainage area by class within each region. Globally, complex catchments drain 121 million km² (87%), intermediate 16 million km² (11.5%) and simple 2 million km² (1.5%) of the 139 million km² of ungauged land evaluated in this study. Africa and southern Asia are predominantly complex, while in Japan intermediate behaviour dominates. **b**, The US Pacific Northwest and British Columbia inset highlights a broad band of simple catchments along the steep, humid Coast Mountains. **c**, In Europe, complex catchments prevail in France, Spain, Belgium, eastern Denmark, eastern Germany, and much of Eastern Europe, intermediate catchments cluster in Ireland, northern and western UK, northeastern France, and the western portions of Denmark, Germany, and Italy, and simple catchments are confined to parts of northern and western UK and Ireland, and to select areas of northern Spain. **d**,**e**, Latin America and the Caribbean (**d**) and Oceania (**e**) are dominated by complex catchments, with intermediate patches in southeastern South America and in eastern and southeastern Australia. An interactive version of this global map is available via https://h-sharif.github.io/stormflow-behavior/. Credit: basemaps, Esri, TomTom, Garmin, FAO, NOAA, USGS, © OpenStreetMap contributors and the GIS User Community.

Extended Data Fig. 4a). The central USA remains functionally complex in both seasons (~1,013 complex catchments) (Fig. 3).

In Europe, 74% of ungauged catchments (4,320) are already complex in the dormant season, especially across most of France, Belgium, Spain, and eastern Germany and Denmark. Intermediate behaviour (22.5%) is concentrated in Ireland, the northern and western UK, northeastern France and the western portions of Denmark, Germany and Italy. Simple catchments (3.5%) appear mostly in northern Spain and in parts of Ireland and the northern and western UK (Fig. 2c). By the growing season, Europe becomes overwhelmingly complex

https://doi.org/10.1038/s44221-026-00699-6

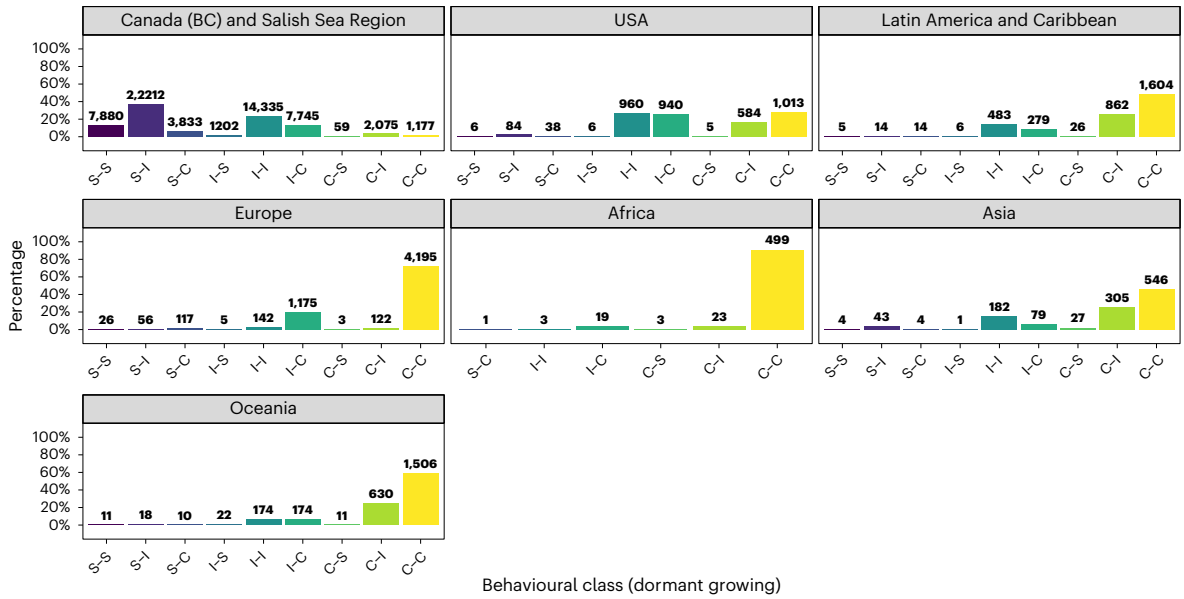


Fig. 3 | Increased functional complexity from dormant to growing seasons is the prevailing global trend in ungauged catchments, though functional simplification emerges in some regions. The bars show the transition of ungauged catchments between functional classes from dormant (left letter on the *x* axis label) to growing (right letter) seasons, with classes denoted as S (simple), I (intermediate) and C (complex). Numbers above each bar give the count of catchments following each transition pathway, and bar heights show their percentage within each region. Globally, from dormant to growing season, simple catchments decline by 73%, intermediate catchments increase by 55% and complex catchments increase by 63% (from 15,277 to 24,968). The dominant global pattern is the intermediate-to-complex transition (I–C), from dormant to growing season, most pronounced in Europe (1,175 catchments) and the USA (940). Conversely, Latin America and the Caribbean (862 catchments) and Oceania (630) show an opposite trend, with catchments shifting from complex to intermediate (C–I). Across all continents, the majority of catchments classified as complex in the dormant season remain complex in the growing season (C–C), underscoring the prevalence and persistence of functional complexity globally. BC denotes the Canadian province of British Columbia; the Salish Sea region includes transboundary catchments spanning the western USA and Canada.

(Extended Data Fig. 4), with 94% of catchments in the complex class and 1,175 catchments shifting from intermediate to complex (Fig. 3), across parts of northeastern France, western Germany, Denmark and Italy.

Africa and southern Asia are dominantly complex, and Japan is dominantly intermediate, in both seasons (Fig. 2 and Extended Data Fig. 4). Latin America and Oceania display more balanced distributions of intermediate and complex behaviour, despite the relative dominance of complexity. Importantly, in these regions, the net seasonal shift is opposite to that in Europe and North America: between the dormant and growing seasons, more catchments transition from complex to intermediate (862 in Latin America and 630 in Oceania) than from intermediate to complex (279 and 174, respectively; Fig. 3). Transitions from complex to intermediate generally occur across southeastern South America and eastern and southeastern Australia (Extended Data Fig. 4).

These results underscore the prevalence of rainfall–runoff complexity across continents, covering both the majority of catchments and the majority of land area, and reveal two contrasting seasonal signatures: while most regions intensify towards complexity in the growing season, a few shift the other way, towards intermediate behaviour.

Climatic and physiographic variables associated with functional behaviour

The question stemming from our analysis then becomes: what climate and physiographic variables are associated with the three functional classes? Comparisons of variable distributions across the three functional classes of gauged catchments (Extended Data Fig. 5), together with the results of variable importance rankings from the XGBoost machine learning classification trained on gauged catchments (Supplementary Table 3), highlight the most influential climatic, geological, land cover and topographic controls on functional classification. All lines of evidence converge on the same finding: climate is the dominant control on catchment functional complexity, while land cover, topography and geology fine-tune the response at local scales. The strongest discriminator in both seasons is rainfall persistence (RW_5), a metric that captures the fraction of time in each season with non-negligible rainfall and thereby indicates whether wetness conditions were persistently wet or alternated between wet and dry states. Simple catchments experience frequent non-negligible rainfall, maintaining persistently wet conditions. Complex catchments, by contrast, are characterized by the infrequent occurrence of non-negligible rainfall (Extended Data Fig. 5). Simple catchments are also wetter overall, with a lower AI (median AI of ≈0.3) and higher long-term average P-AET (median P-AET of ≈2 mm day^{-1}) compared with complex catchments (median AI near 1 and median P-AET of ≈0.6 mm day^{-1}).

Land cover, including cropland and urban fractions, exhibits strong importance, particularly for classifying functional behaviour during the growing season (Supplementary Table 3), with increased human modification associated with greater complexity (Extended Data Fig. 5). Simple catchments are generally smaller in area than intermediate and complex catchments (Extended Data Fig. 5), with catchment area playing a moderate role in distinguishing among the three classes (Supplementary Table 3). Complex catchments tend to exhibit greater depth-to-bedrock than intermediate catchments, which may indicate greater subsurface storage capacity that can support a wider range of antecedent wetness states, potentially making the rainfall–runoff response more dependent on prior storage conditions. Simple catchments tend to occur in steeper terrain that promotes rapid runoff response and more efficient landscape-stream connectivity (Extended Data Fig. 5). These findings explain why regional patterns of functional behaviour largely align with broad climatic gradients, while local exceptions emerge where land cover, geology or topography appear to modulate storage, hydrologic connectivity and runoff production.

Seasonal transitions in functional behaviour are closely tied to variability in rainfall persistence, captured here as seasonal variation in RW_5. Supplementary Fig. 10 shows that catchments transitioning from complex to intermediate in the Southern Hemisphere (C–I) experience increased rainfall persistence from the dormant to the growing season, thereby sustaining wet state for longer and producing simpler runoff responses in the growing season. By contrast, catchments shifting from intermediate to complex in the Northern Hemisphere (I–C) experience an overall decline in rainfall persistence during the growing season, leading to more time-variable wetness conditions. These patterns reinforce the finding that rainfall persistence is the primary climatic control on both seasonal classification (Supplementary Table 3) and seasonal transition, with the direction of transition depending on whether rainfall persistence increases or decreases across seasons. Local geology, land cover and topography may further modulate these regional patterns, explaining why not all catchments within a region follow the same trajectory.

Discussion and conclusions

Applicability and uncertainties of the classification

Our classification provides a globally scalable, data-driven diagnostic of rainfall–runoff functional behaviour across rain-dominated catchments. Several boundary conditions define where it can be applied with confidence. First, the classification focuses on event-scale rainfall–runoff behaviour, the component of the hydrograph that captures most streamflow variability and governs floods, hydrologic connectivity and solute transport. The inter-event baseflow component is not classified here, as it is generally smoother and less variable. As catchment hydrologic function is expressed mainly through event-scale responses[5,49,50], the classification captures the most informative component of streamflow behaviour, as confirmed by its ability to explain full-hydrograph model performance (Supplementary Text 1 and Supplementary Fig. 11).

Second, the classification is restricted to rain-dominated catchments (snow fraction <10%). Snow-dominated catchments are excluded because their functional behaviour is governed by snowmelt–runoff processes[51,52], and reliable daily snowmelt observations/simulations are not available at the global scale. Including such catchments would introduce inconsistent data quality and uncertain target labels, which would reduce the robustness of machine learning predictions across all catchments[53,54]. Rain-dominated catchments represent approximately 72% of the global land area below 60° N latitude, where the vast majority of the world's population, agriculture and water infrastructure are concentrated (Supplementary Fig. 1), making the classification broadly applicable at continental to global scales. Third, the classification does not directly apply to catchments influenced by large dams (Methods; Supplementary Text 5), although it can provide a baseline characterization of their underlying non-regulated behaviour.

Finally, while the classification is global, its regional reliability depends on the availability of sufficient gauged catchments within each functional class. Evaluation of model uncertainty (Methods; Supplementary Table 5) confirms that precision and specificity are consistently high, and classification uncertainty is low, across most regions and classes. However, in regions where a given class is sparsely represented—notably the simple class in Africa and in Latin America and the Caribbean—classification uncertainty is higher, and predictions of the simple class in these regions should be interpreted with caution. Reassuringly, when the classification is collapsed to the binary framing of complex versus non-complex, model performance is strong with minimal uncertainty across all regions and both seasons, indicating that the fundamental distinction between complex and non-complex catchments is reliably resolved globally.

Within these conditions and uncertainties, the classification offers actionable guidance for both model structural design and monitoring network prioritization, as discussed in the following sections.

Implications for model structural design

Classifying catchments into functionally simple, intermediate and complex types provides a practical foundation for hydrologic model design. Our classification reflects the degree of time-variance in catchment hydrologic function as demonstrated by Sharif and Ameli[30] for a subset of gauged catchments used in this study: simple catchments exhibit a near time-invariant rainfall–runoff functional relationship, whereas complex catchments show strongly time-variant behaviour, in which similar rainfall inputs can produce different responses. Taherian and Ameli[55] mechanistically showed that this functional time-variance arises because runoff-generation processes are spatially heterogeneous within a catchment: different parts of the catchment operate through different mechanisms, and the antecedent wetness state determines which parts dominate at a given event, so the effective catchment-scale rainfall–runoff function varies between events.

This mechanistic picture translates directly into structural choices for both flexible conceptual modelling platforms (for example, FUSE[56] and SUMMA[57]) and more recent hybrid physics-informed deep learning models[58–60]. In complex catchments, the spatial mosaic of runoff-generation processes motivates laterally distributed model compartments, and antecedent-state-driven shifts in which processes dominate motivate time-variant parameterization in hybrid architectures. Simple catchments, by contrast, can be adequately represented with lumped, parsimonious models, and over-specifying their structure risks equifinality[61].

Our supplementary modelling experiment supports this directly (Supplementary Text 1 and Supplementary Fig. 11). A lumped parsimonious conceptual hydrology model succeeds in simulating simple catchments but fails in complex ones, a structural failure that no amount of calibration can overcome[62]. Model complexity should therefore be matched to functional behaviour, neither simpler than necessary nor more complex than required, and our classification provides a diagnostic to guide this choice. Supplementary Text 2 provides a more detailed discussion of how this classification guides model design across modelling platforms.

Implications for monitoring network design

Our functional classification also guides where to install new gauges or enhance existing networks in ungauged and poorly gauged catchments (Supplementary Text 3), with priority given to complex catchments, whose underlying processes remain the least understood[25,63]. Supplementary Text 4 and Supplementary Figs. 12 and 13 reveal a mismatch between the functional composition of experimental catchment networks—the proportion of simple, intermediate and complex sites they contain—and the composition of the ungauged landscapes they are meant to represent. In Europe, Asia and Oceania, experimental sites are roughly equally distributed between intermediate and complex catchments, while in the USA, intermediate experimental sites outnumber complex ones by nearly two to one. Yet in all four regions, ungauged catchments are overwhelmingly complex, particularly during the growing season. In Africa and Latin America, although the majority of experimental sites are already complex, their small absolute number is insufficient to represent the full diversity of complexity that dominates these regions. Closing the gap between where we observe catchment behaviour and where complexity prevails is essential for building a global hydrologic theory that represents the world's catchments.

Towards a common functional diagnostic for global catchment hydrology

Our framework provides global, observation-based evidence that functional complexity is the norm across rain-dominated catchments and translates this empirical pattern into a catchment-level diagnostic that assigns each gauged and ungauged catchment its functional type in each season. It offers a common language for comparing catchments across regions and modelling traditions, a longstanding aspiration

of comparative hydrology[64,65], and provides actionable guidance for two persistent challenges: identifying where models should account for functional complexity[66–68] and closing global gaps in streamflow observation[1,69]. Because the diagnostic is derived from data rather than from any particular modelling tradition, it can inform the structural design of conventional conceptual models, flexible modelling platforms, hybrid physics-informed deep learning architectures and Earth system models. More broadly, it provides a foundation for matching model structure to catchment functional behaviour and for directing monitoring effort to where it is most needed. This can help ensure that hydrologic models are neither simpler than necessary nor more complex than required, while observation networks prioritize under-represented catchments whose behaviour remains least understood.

Methods

Gauged and ungauged study catchments and their attributes

We assembled overlapping daily streamflow and precipitation records for non-regulated, rain-dominated gauged catchments distributed across six continents (and 62 countries). These 4,306 study gauged catchments spanned catchment areas from 0.25 km^2 to 232,850 km^2. The majority of catchments had at least 30 years of daily records over 1979–2020. Non-regulated catchments were identified by excluding those containing large dams (Supplementary Text 5). We further compiled a global set of 77,585 ungauged, rain-dominated catchments without streamflow observations across 97 countries. All of these gauged and ungauged catchments have catchment-averaged snow fractions below 10%, ensuring minimal snow influence (Supplementary Methods 1).

For each gauged and ungauged catchment, we derived a common set of long-term catchment-scale predictor attributes (Supplementary Table 1) spanning four domains: (1) long-term climatic indices—AI, seasonality index, asynchronicity index, P-AET and rainfall persistence indices computed separately for dormant and growing seasons, (2) topographic attributes—median topographic wetness index (TWI), mean slope, topographic ruggedness index, and the shape and scale parameters of gamma distributions fitted to pixel-scale topographic attributes, capturing within-catchment topographic heterogeneity, (3) geologic attributes—median depth to bedrock and the two most prevalent lithological classes and (4) land cover fractions of bare earth, urban, forest and cropland. Dormant and growing seasons were identified from satellite-derived phenological transition dates, and rainfall–runoff events were tagged accordingly. Full details of data sources and derivation of climatic, topographic, geologic and land cover attributes are provided in Supplementary Methods 1.

Functional classification of gauged catchments

Classification of gauged catchments into simple, intermediate and complex types was performed in two steps. First, we extracted stormflow (or runoff) events separately for the dormant and growing seasons. Second, we applied linear and segmented-linear regression models to the event data, assigning each catchment to a behavioural class for each season.

Event identification of gauged catchments. We followed the event-detection approach of the ref. 30, which employs detrended moving-average cross-correlation analysis[70] on daily rainfall and streamflow time series. Events were defined by two parameters—minimum rainfall intensity (R_{min}) and maximum moving-average window length (L_{max})—for which we used R_{min} = 1.2 mm day^{-1} and L_{max} = 16 days. Sensitivity tests conducted in ref. 71 indicate that behavioural classification is largely insensitive to these settings. After initial detection, we filtered out: (1) events with runoff ratios >1, (2) events shorter than 1 day or longer than 15 days and (3) catchments with fewer than 15 events in a season. This yielded 947,991 dormant-season and 1,082,578 growing-season events across gauged catchments.

Classification criteria of gauged catchments. Our classification builds on the framework proposed by Dooge[48], who suggested that catchment functional behaviour can be viewed as the superposition of deterministic and stochastic components. When the deterministic part dominates, the catchment behaves with a functionality proximate to the domain of deterministic organized simplicity; when the stochastic part dominates, the system follows unorganized complex functional behaviour. Following his lead, and building upward from the simplest functional form as recommended by Dooge[72], we use the R^2 of the linear regression between event rainfall volume and runoff volume as the metric that reflects the competition between deterministic (or simplicity) dominance and stochastic (or complexity) dominance in a catchment's rainfall–runoff relationship.

For each catchment and season, we regressed event-scale rainfall volume against runoff volume. Catchments with a (single segment) linear fit of $R^2 \geq 0.75$ were labelled as simple, reflecting a consistent rainfall–runoff relationship in which the deterministic component clearly dominates. Those with R^2 of linear fit between 0.50 and 0.75 or those exhibiting strong threshold behaviour were classified as intermediate. The strong threshold behaviour was identified via segmented-linear regression analysis: statistically significant breakpoint(s) were detected via a score test[73], and up to three segments (or two breakpoints) were iteratively fit. If the segmented model achieved $R^2 \geq 0.75$, the catchment was classified as intermediate, indicating threshold-like behaviour. Note that we also evaluated the validity of the nonlinear sigmoid function[74] in explaining the rainfall–runoff relationship. However, no catchment showed strong validity ($R^2 > 0.75$) of such a functional form beyond what was identified using a segmented-linear model. Catchments failing both criteria (linear $R^2 < 0.50$ and segmented $R^2 < 0.75$) were deemed complex, indicating that stochastic dynamics overwhelm the deterministic signal, suggesting temporally inconsistent rainfall-runoff behaviour.

The $R^2 \geq 0.75$ threshold for simplicity is conservative—substantially higher than thresholds used in previous works to characterize simplicity of runoff generation (for example, Ross et al.[29] used an R^2 threshold of 0.45). This conservatism is deliberate: a stricter threshold ensures that only catchments with clearly dominant deterministic behaviour are labelled as simple, reducing the risk of misclassification that could propagate through the machine learning model and compromise the model's predictions in ungauged catchments. Median R^2 values confirm clear separation among the three classes: in the dormant (growing) season, simple catchments exhibit a median linear model R^2 = 0.80 (0.79), intermediate catchments exhibit a median segmented-linear model R^2 = 0.80 (0.80), while complex catchments exhibit weak linear or segmented-linear relationships (median $R^2 \ll 0.5$).

Jackknife resampling analysis for evaluating classification robustness of gauged catchments. To evaluate the robustness and certainty of the functional classification of gauged catchments, we performed a jackknife resampling analysis for each gauged catchment and season. In each round, 80% of the identified runoff events were randomly selected without replacement, and the regression-based classification procedure described in the previous section was repeated on this reduced sample—that is, the linear and segmented-linear models were refit, R^2 values were recomputed and the functional class (simple, intermediate or complex) was re-assigned using the same thresholds. This process was repeated 100 times per catchment per season, yielding a distribution of classifications across 100 resamples. The certainty of the original classification was then quantified as the fraction of the 100 rounds in which the assigned functional class matched the class obtained using the original full samples of events. A high certainty percentage indicates that the classification is stable, not driven by specific event subsets or outliers, and that the R^2 thresholds used for classification are robust given the underlying data. A low percentage would suggest that the catchment's R^2 values fluctuate across the threshold boundaries

under resampling, indicating that the classification is highly sensitive to selected thresholds for that particular catchment.

The results demonstrate high classification robustness across all three functional classes in both seasons (Supplementary Table 4). In the dormant season, the median certainty was 98% for simple, 96% for intermediate and 100% for complex catchments. In the growing season, median certainty was 86% for simple, 92% for intermediate and 100% for complex catchments. The 25th percentile (Q25) values remained above 84% (78%) in all classes of dormant (growing) seasons, indicating that even the least stable catchments retained their classification in the large majority of resampling rounds. Complex catchments exhibited the highest stability, with a Q25 value of 96% (dormant) and 98% (growing), reflecting the strong separation of these catchments from the R^2 thresholds. Overall, these results confirm that the three-class classification is robust to subsampling, that assigned labels are not artifacts of particular event subsets and that the R^2 thresholds delineate stable, meaningful boundaries in the data.

Machine learning-based extrapolation framework for ungauged catchments classification

Predictor-space transferability assessment. Before training and applying a machine learning classifier to extrapolate functional classes to ungauged catchments, it is essential to verify that the ungauged catchments reside within the same region of predictor space as the gauged training catchments. Without such verification, the classifier may be applied to catchments whose attribute combinations fall outside the training domain, leading to unreliable extrapolation. To assess this, we performed a density-based clustering analysis using the DBSCAN algorithm[75], which identifies dense regions in multidimensional predictor space and labels points in sparse areas as noise, without requiring a predefined number of groups. All gauged and ungauged catchments were projected into the same predictor space defined by the predictor variables used in the XGBoost model (Supplementary Table 1), and DBSCAN was applied separately for dormant and growing seasons across 12 trials, each with a different DBSCAN parameter combination to ensure that the conclusions are not sensitive to parameter choice (see Supplementary Text 6 for details).

The results demonstrate strong predictor-space overlap between gauged and ungauged catchments in both seasons (Supplementary Fig. 14). The primary transfer domain—the well-sampled region where both gauged and ungauged catchments are densely represented—consistently contained the majority of both gauged (median ~67%) and ungauged (median ~83%) catchments across all 12 trials. Notably, gauged catchments were more likely to be classified as noise than ungauged ones, indicating that the training data covers a slightly wider territory in predictor space than the prediction domain. No sparse training domain was identified in any dormant-season trial, and it appeared in only 1 of 12 growing-season trials. These results confirm that the XGBoost model will predominantly be interpolating within its well-sampled training domain when applied to ungauged catchments (see Supplementary Text 6 for detailed results).

XGBoost model algorithm. We used gradient-boosted decision trees (XGBoost) to predict the catchment's functional class for each season from physiographic and climatic predictor variables. XGBoost constructs an ensemble of shallow decision trees by sequentially fitting each new learner to the residuals of the preceding model, utilizing a learning rate and subsampling to manage the bias–variance trade-off[76]. Model development followed a multistage process. First, we conducted 200 trials for each season, randomly partitioning the gauged catchments into training and testing sets at 70/30 (100 trials) and 80/20 (100 trials) ratios to assess sensitivity to the training sample size. Within each trial, fivefold stratified cross-validation was used to monitor multiclass log-loss and identify the optimal number of boosting rounds. To prevent overfitting, we implemented early stopping with a patience of 20 rounds of stagnant validation loss, with a maximum cap of 100 boosting iterations.

Key hyperparameters were identified through a preliminary hyperparameter search by minimizing cross-validated log-loss: an eta of 0.3, a maximum tree depth of 8 and a subsample ratio of 0.6 yielded the best cross-validated performance for both dormant and growing-season models. These parameters were then fixed across all 200 trials for each season. For each trial, model performance was evaluated on the held-out test set, and the model's performance metrics were quantified across the global dataset and within each region (see below).

XGBoost model predictor variables (inputs). Predictor variables comprised long-term climatic, topographic, geological and land cover attributes (Supplementary Methods 1 and Supplementary Table 1), each of which is known to influence water balance and catchment storage–release processes. Climatic inputs included aridity, asynchronicity index, P-AET and rainfall persistence indices. Topographic metrics captured average slope, TWI and distributional heterogeneity via gamma-fit parameters for slope, TWI, elevation and topographic ruggedness index. Geological descriptors comprised the two dominant lithology classes and median depth to bedrock. Land cover variables included the percentages of the catchment area as cropland, urban, forest and bare-earth cover. Once the models were developed for the gauged catchments, we used the same set of predictor variables for the ungauged catchments to classify each ungauged catchment during both dormant and growing seasons.

Evaluation of the XGBoost model performance and uncertainty. We evaluated how well the machine learning-based seasonal classifiers recover the ground-truth functional labels of gauged catchments using two complementary metrics: precision—the share of catchments predicted as a given class that truly belong to that class—and specificity—the share of catchments not in a given class that are correctly recognized as such. For each class i

$$\text{Precision}_i = \frac{\text{TruePositives}_i}{\text{TruePositives}_i + \text{FalsePositives}_i}$$

while specificity is

$$\text{Specificity}_i = \frac{\text{True Negatives}_i}{\text{True Negatives}_i + \text{False Positives}_i}.$$

In these equations, true positives and false positives refer to correctly and incorrectly predicted behavioural classes based on ground-truth data, respectively. True negatives are catchments truly not belonging to class i that are correctly predicted as not belonging to it. This pair of metrics provides an intuitive readout of errors on both sides—false inclusions (precision) and false exclusions (specificity)—which is particularly helpful in a multiclass setting where adjacent classes may exhibit similar hydrologic characteristics (for example, simple versus intermediate). For illustration, a precision of 80% for the simple class means that 80% of catchments predicted as simple were truly simple based on ground truth. Specificity, in turn, measures how well non-targets are rejected: a specificity of 80% for the complex class indicates that 80% of truly non-complex catchments (simple or intermediate) were correctly identified as non-complex; similarly, a specificity of 95% for the simple class means that 95% of truly non-simple catchments (intermediate or complex) were correctly identified as non-simple, with only 5% misclassified as simple.

We reported both metrics for the three-class problem (simple, intermediate and complex) and for a binary framing that separates catchments with the dominance of stochasticity (complex) from more deterministic catchments (non-complex: simple and intermediate).

To aid interpretation, we adopted problem-agnostic benchmarks from the classifier-evaluation literature. For three-class classification, each class precision and specificity values ≥60% are acceptable, ≥70% are strong and ≥85% are excellent—reflecting the added difficulty of multiclass separation when adjacent classes are not sharply separated (for example, simple versus intermediate). For the binary complex versus non-complex task, ≥70% is acceptable, ≥80% strong and ≥90% excellent. These ranges accord with general guidance for classifier assessment under class imbalance and overlapping class characteristics, which can inflate misclassification rates[77,78].

For withheld test catchments not used in model training, the dormant-season binary split (complex versus non-complex) yields precision of 0.84 for complex and 0.75 for non-complex, with specificity between 0.74 and 0.84 (Supplementary Table 2). The growing-season binary split yielded a precision of 0.84 for complex and 0.69 for non-complex. For the three-class problem on withheld test catchments, all dormant- and growing-season metrics remain at ≥0.60, with complex class precision equal to 0.84 during both seasons (Supplementary Table 2).

Beyond this global assessment, we examine how model performance and uncertainty vary across regions. We reported precision and specificity for each region and functional class across the 200 cross-validation trials (Supplementary Table 5). We quantified classification uncertainty by the spread between the Q25 and 75th percentile (Q75) of precision and specificity across the 200 trials: narrow spreads combined with high median values for a given class indicate low classification uncertainty, while wide spreads or low median values for a class indicate that the model's ability to reliably identify the given class in that region is uncertain. Low classification uncertainty implies that the model can be confidently used to regionalize functional classes to ungauged catchments in that region, whereas high classification uncertainty indicates that regionalized predictions for the affected class should be interpreted with caution.

In Europe, North America and Oceania, precision and specificity are consistently high with low classification uncertainty across all three functional classes in both seasons (Supplementary Table 5). In Europe, dormant-season median precision ranges from 0.88 (simple) to 0.96 (complex), with Q25 values of precision and specificity remaining above 0.86 and 0.90 across all classes. Growing-season performance was similarly strong, with complex catchments achieving a median precision of 0.98 and simple catchments maintaining a median precision of 0.91. In North America, dormant-season median precision exceeds 0.82 for all classes, with Q25–Q75 differences typically less than 0.05. In the growing season, complex and intermediate classes maintain median precision above 0.87 with narrow Q25–Q75 ranges, and simple catchments achieve a median precision of 0.79 (Q25 of 0.71) with specificity of 1.00. Oceania exhibits comparable performance, with median precision ranging from 0.77 to 0.90 across classes in the dormant season and remaining above 0.79 in the growing season.

By contrast, Africa and Latin America and the Caribbean show strong performance for complex and intermediate classes but higher classification uncertainty for the simple class. In Africa, the dormant-season simple class has a median precision of only 0.50 with a Q25–Q75 range of 0.50–0.63, and in Latin America and the Caribbean, a median of 0.60 with a Q25–Q75 range of 0.50–0.70. These moderately low median values and wide ranges indicate that predictions of simplicity in these regions are not as reliable as in other regions and classes—the model's ability to correctly identify simple catchments varies across trials, indicating higher classification uncertainty and suggesting that regionalization of the simple class to ungauged catchments in these regions should be interpreted with caution. In the growing season, this uncertainty persists for the simple class: in Latin America and the Caribbean, median precision is 0.80 with an interquartile range of 0.20, and in Africa, the median is 0.75 with an interquartile range of 0.31. These wide ranges contrast sharply with the narrow Q25–Q75 spreads of less than 0.05 observed for the same class in Europe and North America, and probably reflect the genuine scarcity of simple catchments in these predominantly complex regions rather than a deficiency in the model. Importantly, when the classification is collapsed to a binary framing (complex versus non-complex), median precision exceeds 0.79 and specificity exceeds 0.75 with minimal interquartile ranges across all regions and both seasons, confirming that the model reliably distinguishes complex from non-complex catchments globally.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

A comprehensive dataset—including daily streamflow and climate time series used to delineate runoff (stormflow) events across gauged catchments, the event catalogue itself and detailed physical and climatic attributes for both gauged and ungauged catchments—is available via Figshare at https://figshare.com/s/24fca5b7d89829035ff1 (ref. 79). The package includes a README that documents each dataset.

Code availability

The code for the trained model used to extrapolate functional classes to ungauged catchments is available via GitHub at https://github.com/h-sharif/stormflow-behavior. It includes an R script for executing the notebook, along with the R Markdown notebook itself. A README in the repository outlines the file structure and provides instructions for reproducing the results. In addition, global interactive maps of functional classes (both seasons) for 77,585 ungauged catchments analysed in this study are available via GitHub at https://h-sharif.github.io/stormflow-behavior. Finally, a user-friendly computer application that classifies catchments from uploaded boundary polygons—returning season-specific labels—has been compiled and is available at https://rainfall-runoff.eoas.ubc.ca/.

References

1. Blöschl, G. et al. Changing climate both increases and decreases European river floods. *Nature* **573**, 108–111 (2019).
2. Talbot, C. J. et al. The impact of flooding on aquatic ecosystem services. *Biogeochemistry* **141**, 439–461 (2018).
3. Wheater, H. S. & Gober, P. Water security and the science agenda. *Water Resour. Res.* **51**, 5406–5424 (2015).
4. Ameli, A. A. & Craig, J. R. Semi-analytical 3D solution for assessing radial collector well pumping impacts on groundwater–surface water interaction. *Hydrol. Res.* **49**, 17–26 (2017).
5. 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).
6. Berghuijs, W. R. & Slater, L. J. Groundwater shapes North American river floods. *Environ. Res. Lett.* **18**, 034043 (2023).
7. Knapp, J. L. A., Berghuijs, W. R., Floriancic, M. G. & Kirchner, J. W. Catchment hydrological response and transport are affected differently by precipitation intensity and antecedent wetness. *Hydrol. Earth Syst. Sci.* **29**, 3673–3685 (2025).
8. Li, H. & Ameli, A. A statistical approach for identifying factors governing streamflow recession behaviour. *Hydrol. Process.* **36**, e14718 (2022).
9. Blöschl, G. *Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales* https://doi.org/10.1017/CBO9781139235761 (2013).
10. McDonnell, J. J. et al. Moving beyond heterogeneity and process complexity: a new vision for watershed hydrology. *Water Resour. Res.* **43**, W07301 (2007).

11. McGlynn, B. L., McDonnell, J. J. & Brammer, D. D. A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. *J. Hydrol.* **257**, 1–26 (2002).
12. Tromp-Van Meerveld, H. J. & McDonnell, J. J. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resour. Res.* **42**, W02411 (2006).
13. Janssen, J. & Ameli, A. A. A hydrologic functional approach for improving large-sample hydrology performance in poorly gauged regions. *Water Resour. Res.* **57**, e2021WR030263 (2021).
14. Ameli, A. A., Laudon, H., Teutschbein, C. & Bishop, K. Where and when to collect tracer data to diagnose hillslope permeability architecture. *Water Resour. Res.* **57**, e2020WR028719 (2021).
15. Saharia, M. et al. On the impact of rainfall spatial variability, geomorphology, and climatology on flash floods. *Water Resour. Res.* **57**, e2020WR029124 (2021).
16. Tarasova, L., Basso, S., Poncelet, C. & Merz, R. Exploring controls on rainfall-runoff events: 2. Regional patterns and spatial controls of event characteristics in Germany. *Water Resour. Res.* **54**, 7688–7710 (2018).
17. Ameli, A. A. et al. Primary weathering rates, water transit times, and concentration-discharge relations: a theoretical analysis for the critical zone. *Water Resour. Res.* **53**, 942–960 (2017).
18. Maher, K. The dependence of chemical weathering rates on fluid residence time. *Earth Planet. Sci. Lett.* **294**, 101–110 (2010).
19. Erlandsson Lampa, M. et al. Catchment export of base cations: improved mineral dissolution kinetics influence the role of water transit time. *SOIL* **6**, 231–244 (2020).
20. Ameli, A. A. Controls on subsurface transport of sorbing contaminant. *Hydrol. Res.* **48**, 1226–1239 (2016).
21. Fan, Y. Groundwater in the Earth's critical zone: relevance to large-scale patterns and processes. *Water Resour. Res.* **51**, 3052–3069 (2015).
22. Wallace, S., Biggs, T., Lai, C.-T. & McMillan, H. Tracing sources of stormflow and groundwater recharge in an urban, semi-arid watershed using stable isotopes. *J. Hydrol. Reg. Stud.* **34**, 100806 (2021).
23. Hrachowitz, M. et al. A decade of predictions in ungauged basins (PUB)—a review. *Hydrol. Sci. J.* **58**, 1198–1255 (2013).
24. Sivapalan, M. et al. IAHS decade on predictions in ungauged basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. *Hydrol. Sci. J.* **48**, 857–880 (2003).
25. Kirchner, J. W. Characterizing nonlinear, nonstationary, and heterogeneous hydrologic behavior using Ensemble Rainfall-Runoff Analysis (ERRA): proof of concept. *Hydrol. Earth Syst. Sci.* **2024**, 1–42 (2024).
26. Kirchner, J. W., Benettin, P. & van Meerveld, I. Instructive surprises in the hydrological functioning of landscapes. *Annu. Rev. Earth Planet. Sci.* **51**, 277–299 (2023).
27. Barthold, F. K. & Woods, R. A. Stormflow generation: a meta-analysis of field evidence from small, forested catchments. *Water Resour. Res.* **51**, 3730–3753 (2015).
28. McMillan, H. et al. Global patterns in observed hydrologic processes. *Nat. Water* **3**, 497–506 (2025).
29. Ross, C. A., Ali, G. A., Spence, C. & Courchesne, F. Evaluating the ubiquity of thresholds in rainfall-runoff response across contrasting environments. *Water Resour. Res.* **57**, e2020WR027498 (2021).
30. Sharif, H. & Ameli, A. A. Searching for functional simplicity of stormflow generation. *Water Resour. Res.* **61**, e2024WR037179 (2025).
31. Wu, S., Zhao, J., Wang, H. & Sivapalan, M. Regional patterns and physical controls of streamflow generation across the conterminous United States. *Water Resour. Res.* **57**, e2020WR028086 (2021).
32. Penna, D. Controls on runoff processes in forested catchments worldwide. *Nat. Water* **4**, 102–114 (2026).
33. Riggs, R. M. et al. Extending global river gauge records using satellite observations. *Environ. Res. Lett.* **18**, 064027 (2023).
34. Blöschl, G. et al. Twenty-three unsolved problems in hydrology (UPH)—a community perspective. *Hydrol. Sci. J.* **64**, 1141–1158 (2019).
35. Yang, W., Yang, H. & Yang, D. Classifying floods by quantifying driver contributions in the Eastern Monsoon Region of China. *J. Hydrol.* **585**, 124767 (2020).
36. Vörösmarty, C. J. et al. Global threats to human water security and river biodiversity. *Nature* **467**, 555–561 (2010).
37. Milly, P. C. D. et al. Stationarity is dead: whither water management?. *Science* **319**, 573–574 (2008).
38. Li, H. & Ameli, A. A. Upland hillslope groundwater subsidy affects low-flow storage–discharge relationship. *Water Resour. Res.* **59**, e2022WR034155 (2023).
39. Sivapalan, M. Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection? *Hydrol. Process.* **17**, 1037–1041 (2003).
40. Gupta, H. V. et al. Large-sample hydrology: a need to balance depth with breadth. *Hydrol. Earth Syst. Sci.* **18**, 463–477 (2014).
41. Seibert, J. & Beven, K. J. Gauging the ungauged basin: how many discharge measurements are needed? *Hydrol. Earth Syst. Sci.* **13**, 883–892 (2009).
42. Milo, R. et al. Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
43. Mcdonnell, J. & Woods, R. On the need for catchment classification. *J. Hydrol.* **299**, 2–3 (2004).
44. Wagener, T., Sivapalan, M., Troch, P. & Woods, R. Catchment classification and hydrologic similarity. *Geogr. Compass* **1**, 901–931 (2007).
45. McGlynn, B. Catchment classification and services—toward a new paradigm for catchment hydrology driven by societal needs. In *Encyclopedai of Hydrological Sciences* (eds Anderson, M. G. & McDonnell, J. J.) https://doi.org/10.1002/0470848944.hsa320 (2008).
46. Scaife, C. I. & Band, L. E. Nonstationarity in threshold response of stormflow in southern Appalachian headwater catchments. *Water Resour. Res.* **53**, 6579–6596 (2017).
47. Scaife, C. I., Singh, N. K., Emanuel, R. E., Miniat, C. F. & Band, L. E. Non-linear quickflow response as indicators of runoff generation mechanisms. *Hydrol. Process.* **34**, 2949–2964 (2020).
48. Dooge, J. C. I. Bringing it all together. *Hydrol. Earth Syst. Sci.* **9**, 3–14 (2005).
49. Tarasova, L. et al. A process-based framework to characterize and classify runoff events: the event typology of Germany. *Water Resour. Res.* **56**, e2019WR026951 (2020).
50. Merz, R. & Blöschl, G. A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resour. Res.* **45**, W01405 (2009).
51. Le, E., Janssen, J., Hammond, J. & Ameli, A. A. The persistence of snow on the ground affects the shape of streamflow hydrographs over space and time: a continental-scale analysis. *Front. Environ. Sci.* **11**, 1207508 (2023).
52. Rahmani, J., Creed, I. F., Badiou, P. & Ameli, A. A. Wetlands set the pace of annual runoff in the northern great plains. *Commun. Earth Environ.* **7**, 368 (2026).
53. Bayati, A., Ameli, A. A. & Razavi, S. Evaluating the functional realism of deep learning rainfall-runoff models using catchment hydrology principles. *Water Resour. Res.* **62**, e2025WR040076 (2026).
54. Janssen, J., Tootchi, A. & Ameli, A. Tackling water table depth modeling via machine learning: from proxy observations to verifiability. *Adv. Water Resour.* **201**, 104955 (2025).
55. Taherian, M. & Ameli, A. Time variance in snowmelt partitioning: a mechanistic modeling approach to explore the role of catchment structure and pre-snow rainfall. *Water Resour. Res.* **62**, e2025WR040679 (2026).

56. Clark, M. P. et al. Framework for Understanding Structural Errors (FUSE): a modular framework to diagnose differences between hydrological models. *Water Resour. Res.* **44**, W00B02 (2008).
57. Clark, M. P. et al. A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.* **51**, 2498–2514 (2015).
58. Wang, Y.-H. & Gupta, H. V. A mass-conserving-perceptron for machine-learning-based modeling of geoscientific systems. *Water Resour. Res.* **60**, e2023WR036461 (2024).
59. Feng, D., Liu, J., Lawson, K. & Shen, C. Differentiable, learnable, regionalized process-based models with multiphysical outputs can approach state-of-the-art hydrologic prediction accuracy. *Water Resour. Res.* **58**, e2022WR032404 (2022).
60. Rahmani, J., Shen, C. & Ameli, A. A. Regionalization of hydrologic behavior and pothole water storage dynamics in the prairie pothole region. *Water Resour. Res.* **62**, e2025WR040280 (2026).
61. Beven, K. & Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **249**, 11–29 (2001).
62. Seibert, J., McDonnell, J. J. & Woodsmith, R. D. Effects of wildfire on catchment runoff response: a modelling approach to detect changes in snow-dominated forested catchments. *Hydrol. Res.* **41**, 378–390 (2010).
63. McDonnell, J. J., Spence, C., Karran, D. J., van Meerveld, H. J. & Harman, C. J. Fill-and-spill: a process description of runoff generation at the scale of the beholder. *Water Resour. Res.* **57**, e2020WR027514 (2021).
64. Sivapalan, M. Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. In *Encyclopedia of Hydrological Sciences* (eds Anderson, M. G. & McDonnell, J. J.) 193–219 https://doi.org/10.1002/0470848944.hsa012 (Wiley, 2006).
65. Thompson, S. E. et al. Developing predictive insight into changing water systems: use-inspired hydrologic science for the Anthropocene. *Hydrol. Earth Syst. Sci.* **17**, 5013–5039 (2013).
66. Fowler, K., Peel, M., Western, A. & Zhang, L. Improved rainfall-runoff calibration for drying climate: choice of objective function. *Water Resour. Res.* **54**, 3392–3408 (2018).
67. Fowler, K. J. A., Peel, M. C., Western, A. W., Zhang, L. & Peterson, T. J. Simulating runoff under changing climatic conditions: revisiting an apparent deficiency of conceptual rainfall-runoff models. *Water Resour. Res.* **52**, 1820–1846 (2016).
68. Peterson, T. J., Saft, M., Peel, M. C. & John, A. Watersheds may not recover from drought. *Science* **372**, 745–749 (2021).
69. Golden, H. E. et al. Advancing the science of headwater streamflow for global water protection. *Nat. Water* **3**, 16–26 (2025).
70. Giani, G., Tarasova, L., Woods, R. A. & Rico-Ramirez, M. A. An objective time-series-analysis method for rainfall-runoff event identification. *Water Resour. Res.* **58**, e2021WR031283 (2022).
71. Sharif, H. & Ameli, A. Searching for Functional Simplicity of Stormflow Generation. *Zenodo* https://doi.org/10.5281/zenodo.14253670 (2024).
72. Dooge, J. C. I. Searching for Simplicity in Hydrology. *Surv. Geophys.* **18**, 511–534 (1997).
73. Muggeo, V. M. R. Testing with a nuisance parameter present only under the alternative: a score-based approach with application to segmented modelling. *J. Stat. Comput. Simul.* **86**, 3059–3067 (2016).
74. Zehe, E., Elsenbeer, H., Lindenmaier, F., Schulz, K. & Blöschl, G. Patterns of predictability in hydrological threshold systems. *Water Resour. Res.* **43**, W07434 (2007).
75. Ester, M., Kriegel, H.-P. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. Second International Conference on Knowledge Discovery and Data Mining (KDD-96)* (eds Simoudis, E., Han, J. & Fayyad, U. M.) 226–231 (AAAI Press, 1996).
76. Chen, T. & Guestrin, C. XGBoost: a scalable tree boosting system. In *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (eds Krishnapuram, B. et al.) https://doi.org/10.1145/2939672.2939785 (Association for Computing Machinery, 2016).
77. Sokolova, M. & Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **45**, 427–437 (2009).
78. Powers, D. Evaluation: From precision, recall and f-measure to roc, informedness, markedness & correlation. *J. Mach. Learn. Technol.* **2**, 37–63 (2011).
79. Data for 'The functional diversity of ungauged catchments: a global hydrological synthesis' *Figshare* https://figshare.com/s/24fca5b7d89829035ff1 (2026).

Acknowledgements

To meet the high computational demands of event identification and functional classification for a large sample of catchments, the majority of analyses were carried out through the network of parallel high-performance computing clusters provided by Compute Canada.

Author contributions

A.A.A. peformed the study design, conceptualization, writing and funding acquisition. H.S. performed the methodology and analyses, data compilation and visualization; J.J.M. performed the conceptualization and writing.

Funding

This study was funded by grants awarded to A.A.A. by the Canadian Statistical Science Institute and by the NSERC Discovery Grants programme. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s44221-026-00699-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44221-026-00699-6.

Correspondence and requests for materials should be addressed to Ali A. Ameli.

Peer review information *Nature Water* thanks Uwe Ehret and Heather Golden for their contribution to the peer review of this work. Peer reviewer reports are available.

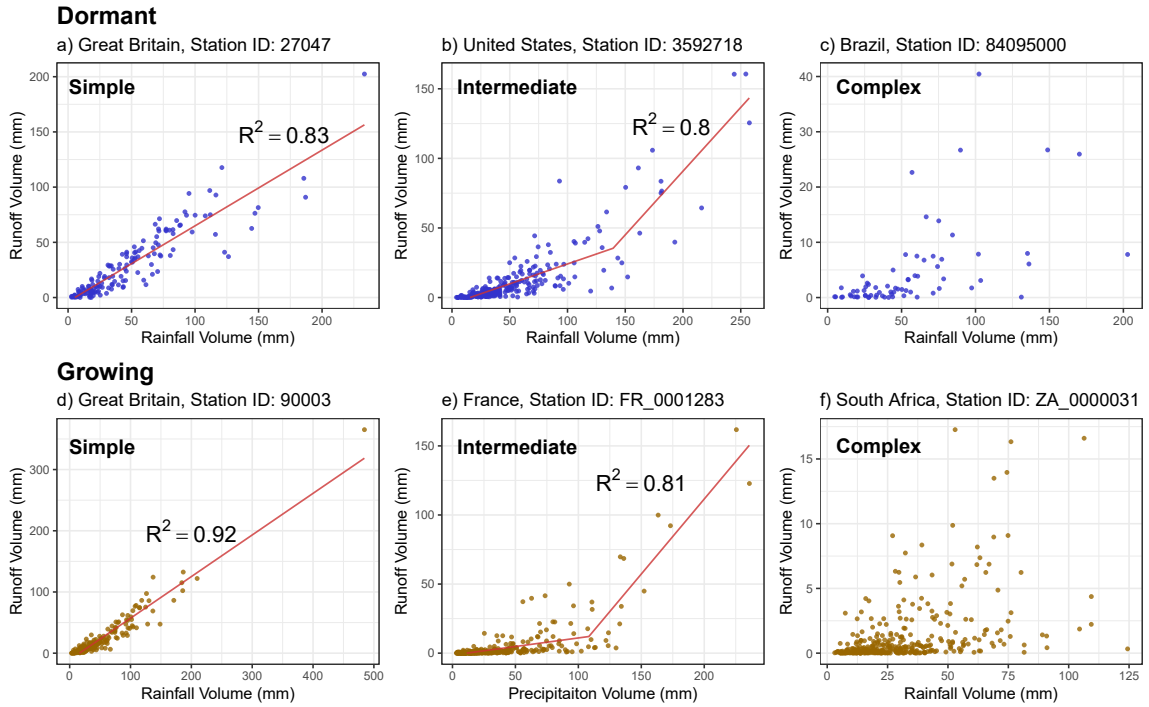
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

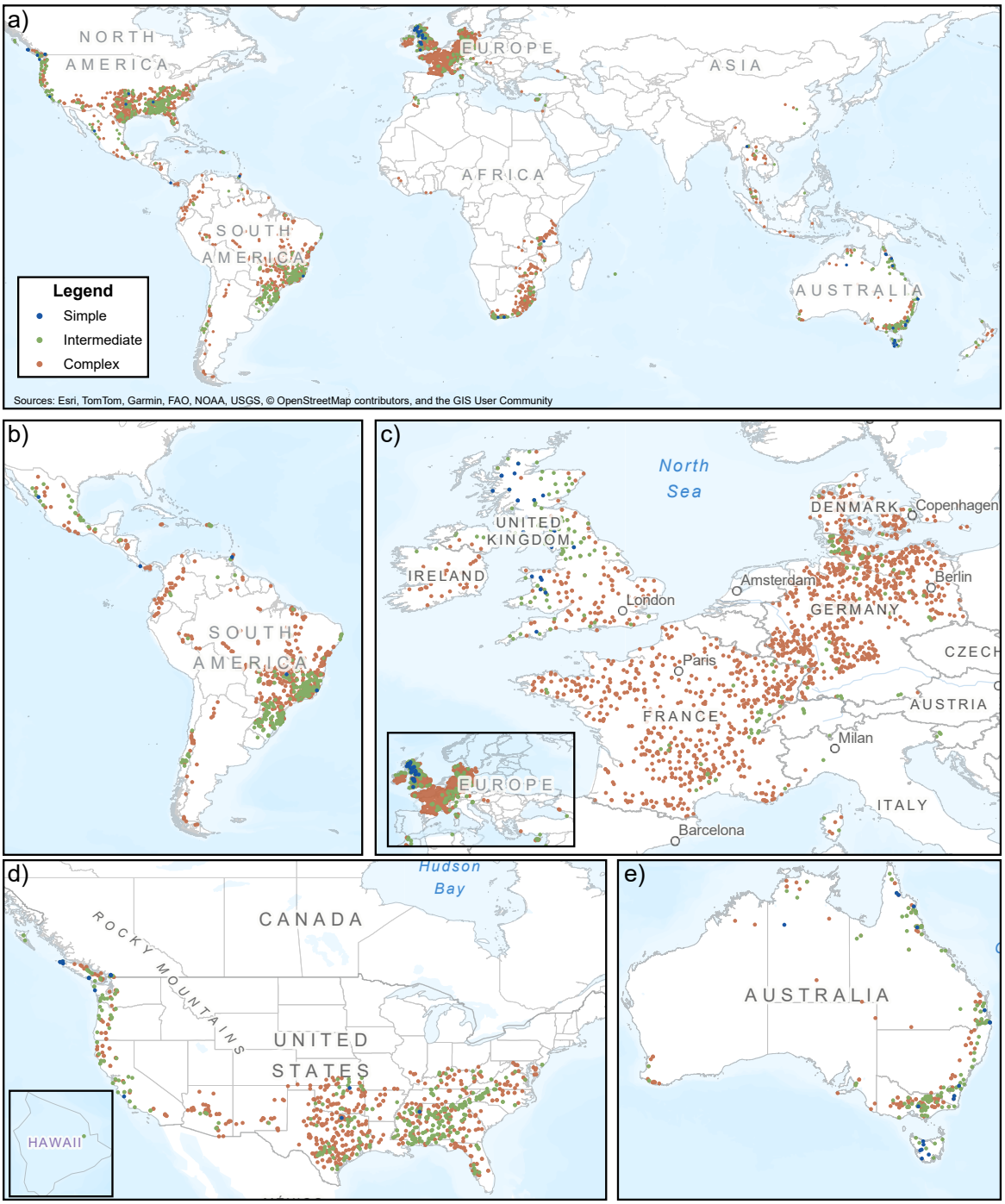
© The Author(s), under exclusive licence to Springer Nature Limited 2026

https://doi.org/10.1038/s44221-026-00699-6



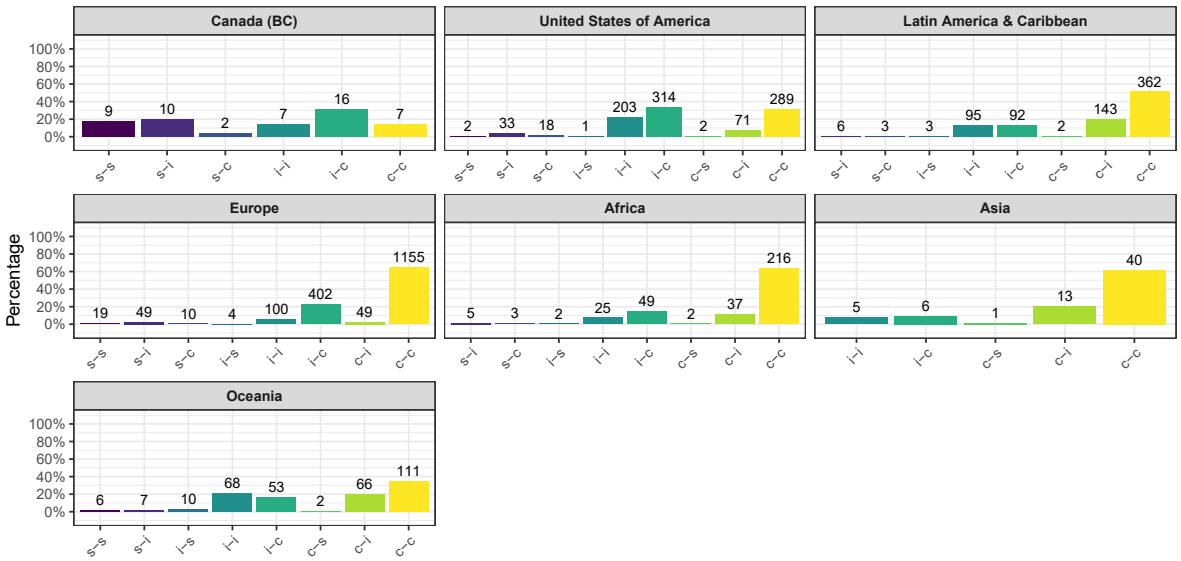
Extended Data Fig. 1 | Rainfall–Runoff relationships vary from linear to highly nonlinear. (**a**–**c**) Dormant-season examples: (**a**) a simple catchment in Great Britain (Station 27047) exhibits a near-linear rainfall–runoff relationship ($R^2 = 0.83$); (**b**) an intermediate U.S. catchment (Station 3592718) shows a segmented-linear response, with a steeper slope beyond ~140 mm of rainfall ($R^2 = 0.80$); (**c**) a complex Brazilian catchment (Station 84095000) displays highly variable rainfall– runoff relationship. (**d**–**f**) Growing-season examples: (**d**) a simple Great Britain catchment (Station 90003) yields a strong linear rainfall– runoff fit ($R^2 = 0.92$); (**e**) an intermediate French catchment (Station FR_0001283) reveals threshold-like behavior with a breakpoint at ~110 mm of rainfall ($R^2 = 0.81$); (**f**) a complex South African catchment (Station ZA_0000031) produces scattered, stochastic runoff responses, highlighting its highly nonlinear functional behavior.

https://doi.org/10.1038/s44221-026-00699-6



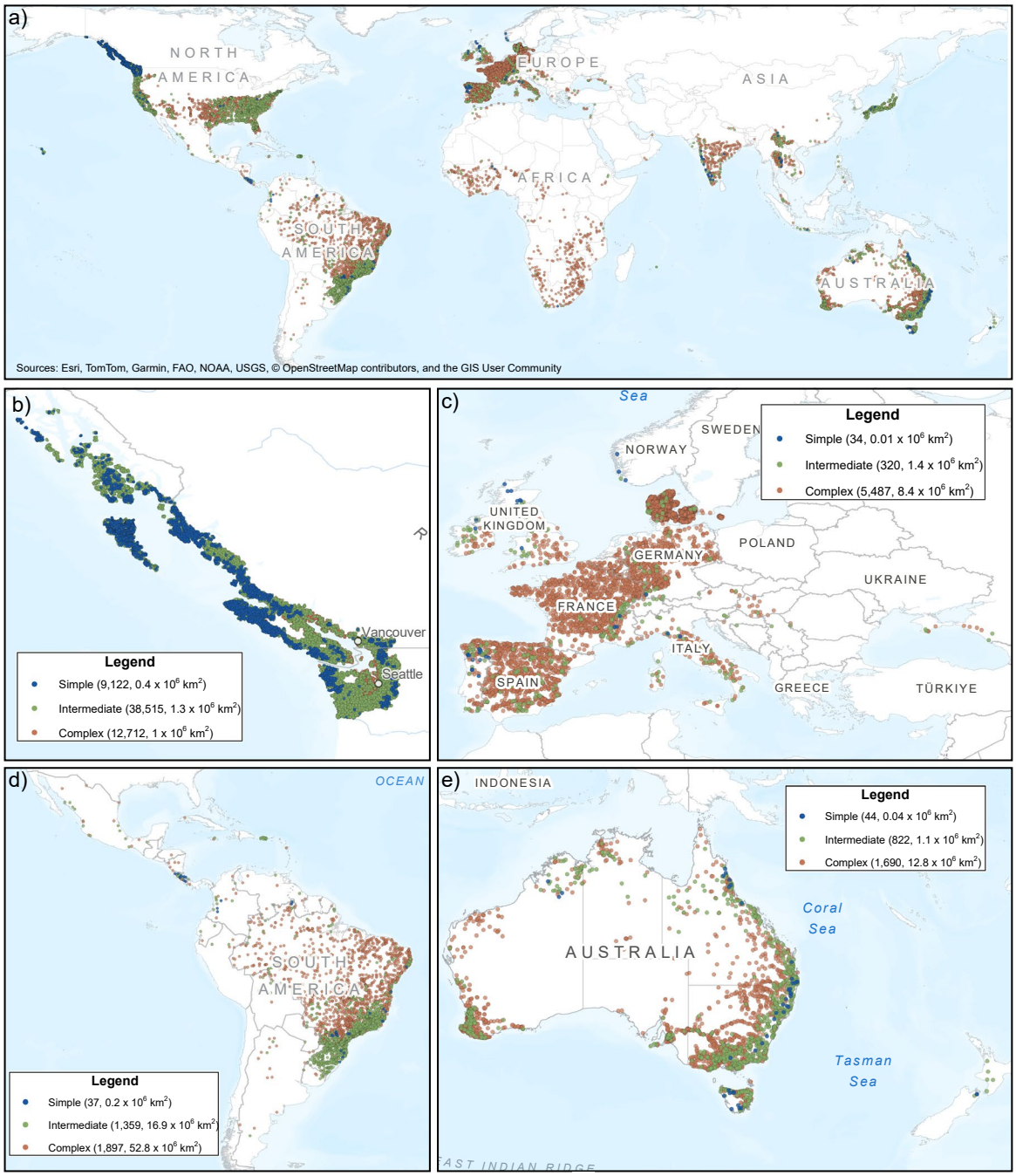
Extended Data Fig. 2 | Functional complexity prevails in gauged catchments during the growing season. (**a**) Global distribution of simple (blue), intermediate (green), and complex (red) behavioral classes among gauged catchments in the growing season. (**b**–**e**) Continental insets of functional classes covering Latin America and the Caribbean, Europe, North America, and Oceania. South and Central America: complexity dominates, with dense clusters of intermediate systems across southeastern South America. Europe: widespread complex behavior across France, Germany, Denmark, Ireland, and the southeastern UK; simple and intermediate catchments remain confined to small parts of the western and northern UK. North America: a pronounced shift toward complexity dominance in the eastern U.S. from dormant to growing season. Oceania: complexity dominates, with intermediate patches in southeastern Australia and simple patches in Tasmania. Credit: basemaps, Esri, TomTom, Garmin, FAO, NOAA, USGS, © OpenStreetMap contributors and the GIS User Community.

https://doi.org/10.1038/s44221-026-00699-6



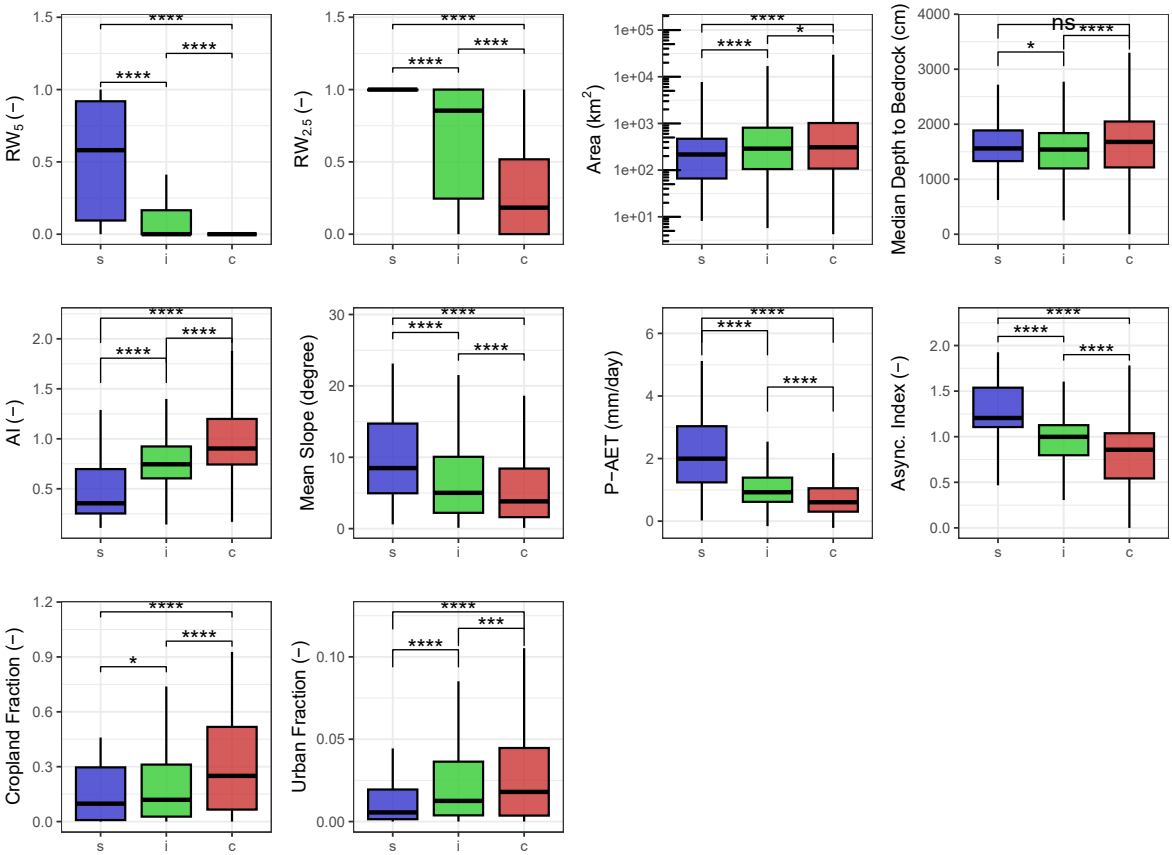
Extended Data Fig. 3 | Seasonal intensification of functional complexity among gauged catchments, with widespread transitions from intermediate to complex behavior between the dormant and growing seasons. Bars show the transition of gauged catchments between functional classes from dormant (left letter on the x-axis label) to growing (right letter) seasons, with classes denoted S (simple), I (intermediate), and C (complex). Numbers above each bar give the count of catchments following each transition pathway, and bar heights show their percentage within each region. A clear pattern emerges: many catchments shift from intermediate to complex (I–C), most notably 402 in Europe and 314 in the United States, reflecting enhanced runoff generation complexity during the growing season. At the same time, the majority of catchments classified as complex in the dormant season remain complex in the growing season (C–C), highlighting the prevalence and persistence of functional complexity globally. In contrast, many catchments in Latin America and the Caribbean (143 catchments) and Oceania (66 catchments) exhibit the opposite trend, transitioning from complex to intermediate (C–I), underscoring the regional heterogeneity of seasonal functional transitions. BC denotes the Canadian province of British Columbia.

https://doi.org/10.1038/s44221-026-00699-6



Extended Data Fig. 4 | Complex catchments expand across ungauged rain-dominated land in the growing season. (**a**) Global map of simple (blue), intermediate (green), and complex (red) classifications for 77,585 ungauged rain-dominated catchments during the growing season; panel legends report the count and total drainage area by class within each region. Africa and southern Asia remain predominantly complex from dormant to growing seasons. (**b**) The U.S. Pacific Northwest and British Columbia inset shows a marked reduction in simple catchments and expansion of intermediate and complex classes relative to the dormant season, most pronounced in Washington and Oregon. (**c**) The European inset shows a widespread shift toward complexity in areas that were simple or intermediate during the dormant season, across catchments in northeastern France, northern Spain, southeastern Ireland, western Germany, Denmark, and Italy. Only small remnants of intermediate behavior remain in the northern and western UK and Ireland. (**d, e**) Latin America and the Caribbean and Oceania are dominated by complex catchments in the growing season, though catchments in southeastern South America and in eastern and southeastern Australia shift from complex to intermediate between the dormant and growing seasons. An interactive version of this global map is available at https://h-sharif.github.io/stormflow-behavior/. Credit: basemaps, Esri, TomTom, Garmin, FAO, NOAA, USGS, © OpenStreetMap contributors and the GIS User Community.

https://doi.org/10.1038/s44221-026-00699-6



Extended Data Fig. 5 | Climate dominates functional behavior, with topography, land cover, and geology fine-tuning regional patterns. Boxplots compare the distributions of key climatic and physiographic variables among simple (S, n = 183), intermediate (I, n = 1,464), and complex (C, n = 2,659) gauged catchments during the dormant season. In each boxplot, the center line indicates the median; box bounds indicate the 25th and 75th percentiles; and whiskers extend to the most extreme data points within 1.5 × the interquartile range from the box; data beyond the whiskers are not shown. Pairwise comparisons were performed using two-sided Wilcoxon rank-sum tests. Because most p-values fall below the numerical precision of the test, significance is reported by threshold rather than as exact values ($p > 0.05$: ns; $p \le 0.05$: *; $p \le 0.01$: **; $p \le 0.001$: ***; $p \le 0.0001$: ****). Rainfall persistence indices (RW_5 and $RW_{2.5}$), which capture the fraction of time in each season with non-negligible rainfall, show the strongest separation, with simple catchments persistently wet, complex catchments more variable, and intermediate catchments falling in between. Aridity Index (AI) and effective precipitation (P-AET) provide measures of overall water availability, again separating wetter simple catchments from drier complex ones. Topographic, geologic, and land cover variables refine this climatic template. Simple catchments are generally smaller than intermediate and complex catchments. Complex catchments have greater depth to bedrock than intermediate catchments, and simple catchments tend to occur in steeper terrain. Land cover attributes further modulate functional behavior: catchments with higher percentages of cropland and urban cover tend to exhibit greater functional complexity. The variables shown here are among the most important predictors of catchment classification identified by the XGBoost machine-learning model (see Supplementary Table 3).

nature portfolio

Corresponding author(s): Ali A Ameli

Last updated by author(s): Jul 13, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect the data
Data analysis	1- For event extraction from time series, we employed the method explained in Giani et al (2022). The publicly available code published along with their paper, and used in our study, is available at "https://github.com/giuliagiani/DMCA-ESR". Giani, G., Tarasova, L., Woods, R. A. & Rico-Ramirez, M. A. An Objective Time-Series-Analysis Method for Rainfall-Runoff Event Identification. Water Resour. Res. 58, e2021WR031283 (2022) 2- For identifying breakpoint in rainfall-runoff relationship, we used the methodology developed by by Muggeo (2016). We used the publicly available code published along with their paper, available at "https://github.com/cran/segmented/blob/master/man/segmented-package.Rd" Muggeo, V. M. R. Testing with a nuisance parameter present only under the alternative: a score-based approach with application to segmented modelling. J. Stat. Comput. Simul. 86, (2016). 3- Machine learning model development for extrapolation of functional classes was conducted using the publicly available and executable code published by us and available at "https://github.com/h-sharif/stormflow-behavior"

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

A comprehensive dataset—including daily streamflow and climate time series used to delineate stormflow events across gauged catchments, the event catalog itself, and detailed physical and climatic attributes for both gauged and ungauged catchments—is available via Figshare: https://figshare.com/s/24fca5b7d89829035ff1. The package includes a README that documents each dataset.
We acknowledge that there is no restriction on our data and all data used in our paper were published in mentioned repository.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	item is not relevant to our study
Reporting on race, ethnicity, or other socially relevant groupings	item is not relevant to our study
Population characteristics	item is not relevant to our study
Recruitment	item is not relevant to our study
Ethics oversight	item is not relevant to our study

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We developed and applied a globally scalable classification framework to characterize rainfall–stormflow functional relationships in both gauged and ungauged natural basins. The framework assigns basins into three behavioral classes (simple, intermediate, complex) using event-scale rainfall–runoff relationships, evaluated separately for dormant and growing seasons. The study comprises gauged catchments (experimental units), each with more than 30 years of continuous daily rainfall and streamflow records, providing millions of stormflow events that serve as replicates.
Research sample	Our research sample comprises 4306 rain-dominated gauged catchments distributed across six continents (62 countries). Because the study aims to diagnose rainfall–stormflow functional behavior, we restricted analyses to rain-dominated basins and excluded non-rain-dominated (snow-dominated) systems using a catchment-average snow fraction ≤10% (computed from the proportion of days with air temperature ≤0 °C). To isolate natural behavior, we further screened for minimal Dam regulation (no major dams). Each included catchment has ≥30 years of daily precipitation and streamflow records (typically ~1979–2020). The sample is intended to represent non-regulated rain-dominated drainage basins globally; it supports model training/validation and subsequent extrapolation to ungauged, rain-dominated catchments. Streamflow came from national and international archives (e.g., GRDC and CAMELS family datasets), and climate drivers from gridded reanalyses/products as detailed in the manuscript's supplementary document.
Sampling strategy	We assembled all available gauged basins that met our pre-set inclusion criteria (rain-dominated, non-regulated) across the global archives described in the supplementary docuemnt. No statistical method was used to predetermine sample size; instead, we maximized coverage subject to these filters to ensure representativeness. The resulting 4306 catchments span six continents and cover wide gradients in climate, geology, and topography (see Table S1). This breadth provides both statistical power (with more than 2 million identified stormflow events across seasons) and geographic and physiographic diversity sufficient to train and validate robust season-specific classifiers and to justify extrapolation to the ungauged, rain-dominated catchments analyzed.
Data collection	Streamflow and meteorological data for gauged catchments were drawn from public domain hydrometric archives (e.g., CAMELS,

Data collection	GRDC datasets and national agencies). Event extraction and attribute derivation followed a standardized algorithm applied across all catchments. Ungauged basin attributes (climatic, topographic, geologic) were derived from globally available datasets (e.g., EM-Earth, global DEMs, lithological maps)
Timing and spatial scale	Compiled data spans 1970–2020, with most gauged records covering 1979–2020. Daily time series of rainfall and streamflow were used to identify more than 2 million stormflow events across gauged basins. Spatially, the study covers six continents (North America, South America, Europe, Africa, Asia, Oceania), with catchment sizes ranging from <10 km² headwaters to >10,000 km² regional basins. Seasonal classification was performed separately for dormant and growing seasons within each hemisphere.
Data exclusions	Catchments were excluded if they (i) were not rain-dominated, (ii) showed evidence of reservoir or dam regulation. These inclusion/exclusion criteria were pre-established to isolate non-regulated, rain-dominated catchment behavior and to ensure data quality—a routine and widely accepted practice in hydrology and water resources.
Reproducibility	Reproducibility was ensured by (i) applying a consistent workflow to all catchments worldwide, (ii) using standardized and publicly available datasets, and (iii) validating machine learning models with cross-validation and independent test sets. Supplementary HBV modeling further independently verified the classification framework's accuracy, showing consistent differences in model adequacy across functional classes. All datasets, code, and interactive classification maps are publicly available, allowing independent replication of our analyses
Randomization	Randomization is not applicable in this study. Our work is based on existing hydrometric and climatic datasets rather than experimental allocation of samples. Catchments were not randomly assigned to groups but instead classified into functional categories (simple, intermediate, complex) according to their observed rainfall–runoff relationships.
Blinding	Blinding is not relevant in this study. All analyses were conducted using publicly available streamflow, climate, and catchment attribute datasets. The classification framework was applied algorithmically without subjective interpretation, and the evaluation metrics (precision, specificity) were computed automatically.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	item is not relevant to our study
Novel plant genotypes	item is not relevant to our study
Authentication	item is not relevant to our study