Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Earth's planetary-scale lake storage phenology

Jida Wang

jidaw@illinois.edu

University of Illinois Urbana-Champaign https://orcid.org/0000-0003-3548-8918

Md Sikder

Department of Geography & GIS, University of Illinois https://orcid.org/0000-0002-1910-1800

Mélanie Trudel

Université de Sherbrooke

Sylvain Biancamaria

Université de Toulouse, Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS) (CNES/CNRS/IRD/UT)

Jean-Francois Cretaux

Legos/CNES https://orcid.org/0000-0002-4038-7719

Xinchen He

University of Illinois Urbana-Champaign

Roger Fjørtoft

Centre Spatiales d'Etudes Spatiales (CNES)

Claire Pottier

Centre Spatiales d'Etudes Spatiales (CNES)

Chunqiao Song

State Key Laboratory of Lake and Watershed Science for Water Security, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences https://orcid.org/0000-0003-1060-4636

George Allen

Virginia Tech https://orcid.org/0000-0001-8301-5301

Casey Brown

University of Massachusetts

Huilin Gao

Texas A&M University https://orcid.org/0000-0001-7009-8005

Colin Gleason

University of Massachusetts https://orcid.org/0000-0002-3525-6220

Laurence Smith

Brown https://orcid.org/0000-0001-6866-5904

Eric Levenson

University of North Carolina Chapel Hill

Sarah Cooley
Duke University
Meng Ding
Kansas State University
Etienne Fluet-Chouinard
Pacific Northwest National Laboratory https://orcid.org/0000-0003-4380-2153
Yoshihide Wada
King Abdullah University of Science and Technology https://orcid.org/0000-0003-4770-2539
Cedric David
Jet Propulsion Laboratory, California Institute of Technology https://orcid.org/0000-0002-0924-5907
Hind Oubanas
G-EAU, Univ Montpellier, AgroParisTech, BRGM, CIRAD, INRAE, Institut Agro, IRD
Tamlin Pavelsky
University of North Carolina at Chapel Hill

Physical Sciences - Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11587686/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Earth's planetary-scale lake storage phenology

Jida Wang[1,*], Md Safat Sikder[1], Mélanie Trudel[2], Sylvain Biancamaria[3], Jean-François Crétaux[3], Xinchen He[1], Roger Fjørtoft[4], Claire Pottier[4], Chunqiao Song[5,6], George H. Allen[7], Casey Brown[8], Huilin Gao[9], Colin J. Gleason[8], Laurence C. Smith[10], Eric Levenson[11], Sarah Cooley[12], Meng Ding[13], Etienne Fluet-Chouinard[14], Yoshihide Wada[15], Cédric H. David[16], Hind Oubanas[17], Tamlin M. Pavelsky[11]

Author affiliations

1. Department of Geography and Geographic Information Science, Department of Earth Science and Environmental Change, University of Illinois Urbana-Champaign, Urbana, IL, USA
2. Department of Civil and Building Engineering, Université de Sherbrooke, Sherbrooke, QC, Canada
3. Université de Toulouse, Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS) (CNES/CNRS/IRD/UT), Toulouse, France
4. Centre Spatiales d'Etudes Spatiales (CNES), Toulouse, France
5. State Key Laboratory of Lake and Watershed Science for Water Security, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China.
6. University of Chinese Academy of Sciences, Beijing, China
7. Department of Geosciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
8. Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA, USA
9. Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX, USA
10. Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI, USA
11. Department of Earth, Marine, and Environmental Sciences, University of North Carolina, Chapel Hill, NC, USA
12. Nicholas School of the Environment, Duke University, Durham, NC, USA
13. Department of Geography and Geospatial Sciences, Kansas State University, Manhattan, KS, USA
14. Earth System Science Division, Pacific Northwest National Laboratory, Richland, WA, USA
15. Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
16. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
17. G-EAU, Univ Montpellier, AgroParisTech, BRGM, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

*Corresponding to: jidaw@illinois.edu

Abstract

Lakes, Earth's largest store of liquid surface freshwater, record upstream hydroclimatic variability while mediating downstream flows[1-4]. This dual role is encoded in lake storage phenology – the magnitude, timing, and shape of seasonal storage changes – whose planetary organization and human modification remain poorly understood[4-6]. Here, we use wide-swath satellite altimetry[7] to unveil storage phenology across 5.1 million lakes ≥1 ha worldwide through a four-level framework: individual-lake variability, planetary-scale cycle, recurrent archetypes, and process interpretation via storage–flow coupling. Individual-lake variability, structured by landscape setting, totals 3,047.5–3,571.2 Gt (74.4–119.2% above previous estimates[5,6]), with reservoirs contributing disproportionately and smaller lakes exhibiting greater volume-normalized variability. Phenological asynchrony attenuates these fluctuations to a planetary-scale seasonal cycle of 636.7–928.8 Gt, roughly tripling river-channel variability[8]. Seven archetypes organize this diversity into coherent seasonal patterns, including a high-variability monsoonal mode whose late-season filling opposes global storage recession, dampening planetary variability by 40.1%. While lake storage often retains the seasonal signature of naturalized upstream flow, phase delays of ≥6 weeks characterize nearly half of natural-lake and three-quarters of reservoir storage variability, revealing substantial mediation through hydrologic memory and human regulation. These findings establish lake storage phenology as a measurable dimension of the terrestrial water cycle, linking hydroclimate, storage mediation, and anthropogenic imprint across scales.

Main

Natural lakes and reservoirs, hereafter lakes, dominate Earth's inland surface waters in number, area, and volume[1,9,10]. While rivers convey water across landscapes, lakes retain it, storing nearly 90 times the volume held in river channels[11]. This retention allows lake storage to carry a "memory" of hydroclimatic forcing across seasonal and longer timescales[12,13]. Spanning nearly all climate zones and geographic settings[9,10], seasonal lake storage simultaneously records regional hydroclimatic regimes and mediates freshwater redistribution[2-4,14], while reservoirs additionally bear the imprint of managed filling and drawdown[15,16]. We define the seasonal organization of lake storage – its variability magnitude, timing, and trajectory shape – as lake storage phenology. Resolving this phenology is central to understanding the role of lakes in terrestrial hydrology and how human management reshapes that role.

Characterizing lake storage phenology globally poses challenges of both measurement and scale. Because storage change is three-dimensional, estimating it requires contemporaneous measurements of water surface elevation (WSE) and extent[17,18]. Scaling these estimates globally further requires repeated observations of millions of predominantly small, widely dispersed lakes, many with complex littoral morphology[10]. These demands exceed the spatial coverage and temporal sampling of conventional nadir-pointing satellite altimeters. Recent studies have combined multi-mission radar and lidar altimetry with optical water products to extend observations to approximately 230,000 lakes[5,6] – still only a fraction of the nearly six million lakes larger than 1 ha globally[10]. Observational coverage also remains strongly size-dependent, with most small lakes containing only 2–5 measurements per year[5,19]. Consequently, global assessments of seasonal lake storage have emphasized the magnitude of variability, often relying on extrapolation to account for unobserved smaller lakes[6,20]. Efforts to resolve complete seasonal trajectories remain even rarer, relying on reconstruction from sparse observations and interpretation mainly at basin scales[19]. Such limitations can miss seasonal extrema, bias timing, and obscure phenological heterogeneity among individual lakes, leaving the global magnitude of seasonal lake storage variability uncertain and the phenological organization of lake storage largely unknown.

Launched in December 2022, the Surface Water and Ocean Topography (SWOT) satellite mission introduced global wide-swath radar altimetry, enabling simultaneous, high-resolution measurements of lake WSE and extent on a 21-day orbit cycle[7]. Here we leverage the first 22 months of SWOT science-orbit observations across more than five million lakes down to the hectare scale to operationalize lake storage phenology at near-census detail (Methods). We characterize this phenology using a four-level framework that progresses from individual-lake variability, through planetary-scale lake storage cycle and recurrent archetypes, to process-based interpretation. We first quantify the magnitude of seasonal storage variability and its partitioning by landscape

setting, lake type and size. We then determine how the timing and coherence of millions of individual trajectories aggregate to produce the global lake storage cycle. Next, we inductively identify recurrent phenological archetypes from trajectory timing and shape, quantifying how they reinforce or dampen the planetary signal. Finally, we relate observed phenology to naturalized upstream discharge to interpret how lake storage preserves or modifies hydroclimatic signals through hydrologic memory and human regulation. Collectively, these analyses establish lake storage phenology as a globally observable and hydrologically consequential dimension of the terrestrial water cycle.

Individual-lake storage variability

We analyzed SWOT observations[21] from 21 July 2023 to 5 May 2025 (cycles 1–32) for lakes catalogued in the mission's Prior Lake Database[10], the most comprehensive global inventory of lakes ≥1 ha to date. Quality filtering retained 5.1 million lakes with records spanning at least one year, comprising 1.6 million tier-1 lakes with temporally consistent trajectories and 3.5 million tier-2 lakes with sparser temporal sampling (Extended Data Fig. 1). For each lake, seasonal variability was quantified as the median of interdecile ranges (IDR) across valid one-year windows to reduce sensitivity to residual outliers and interannual variability (Methods).

Natural-lake WSE variability is spatially heterogeneous yet structured by landscape and hydrologic setting (Fig. 1a), represented here by five contextual typologies: permafrost-zone, floodplain, on-river, glacial, and endorheic (later-listed types take precedence when settings overlap; Extended Data Fig. 2a,b; Methods). Permafrost-zone lakes, abundant but predominantly small, show muted fluctuations (median IDR, 0.37–0.69 m across size classes). Floodplain and on-river lakes are more dynamic (0.52–0.90 m and 0.60–1.46 m, respectively), reflecting stronger coupling to seasonal inflows and flood waves. Glacial lakes can span higher values (0.56–2.31 m), reflecting confined basin topography and pulsed meltwater inputs, whereas endorheic lakes show modest variability (0.46–0.78 m), consistent with the aridity of landlocked basins[22]. Reservoirs, by contrast, exhibit median WSE variability of 0.64–2.83 m across size classes, typically 2–4 times that of similarly sized natural lakes (Fig. 1f). This greater variability persists across nearly all typologies and latitude bands (Fig. 1b; Extended Data Fig. 2c,d), particularly in China, India and Brazil, where many high-capacity, multipurpose reservoirs undergo substantial seasonal drawdown and refill[15,23].

When expressed as storage variability, these contrasts are amplified by lake-size disparities (Fig. 1c,d). Although WSE variability scales only weakly with lake size (Fig. 1f), integrating vertical fluctuations over surface areas spanning orders of magnitude steepens the size dependence of storage variability, concentrating absolute variability in large lakes and mega-reservoirs while reducing per-lake contributions of predominantly small systems such as permafrost-zone lakes (Extended Data Fig. 3). Hydroclimatic

extremes can further amplify regional storage excursions: during the 2023–2024 El Niño-associated drought, many Amazon ria lakes varied by >10 m, producing exceptional storage depletion with severe ecological consequences[24]. Human regulation leaves a more pervasive global imprint: although known reservoirs comprise <5% of observed lakes, they account for 48.1–52.5% of summed individual-lake storage variability (both tiers combined versus tier-1 alone).

By observing >20 times as many lakes as previous global studies, our census reveals 3,047.5–3,571.2 Gt of summed individual-lake storage variability (Fig. 1g). This markedly exceeds prior amplitude-based observations (1,390–2,048 Gt) by 74.4–119.2%[5,6], despite our more conservative IDR metric. Converting IDR to an amplitude-equivalent measure raises our estimate to 4,787.6–5,546.1 Gt (Methods), still above previous estimates of 3,685–4,635 Gt that relied partly on statistical extrapolation of small-lake contributions[20].

Direct observation of millions of small lakes also resolves their collective importance to global storage variability. Although individual storage variability declines toward smaller lakes, lake abundance rises steeply (Fig. 1e). Lakes ≤1 km^2 therefore contribute 6.5–10.3% of summed global storage variability (Fig. 1g); the share rises to 17.6–23.6% for lakes ≤10 km^2. Meanwhile, volume-normalized variability tends to increase in smaller lakes (Fig. 1h; Extended Data Fig. 4). This contrast indicates that small lakes are often more seasonally dynamic relative to their storage capacity, extending previous knowledge of greater relative area variability in smaller lakes to storage dynamics[25,26].

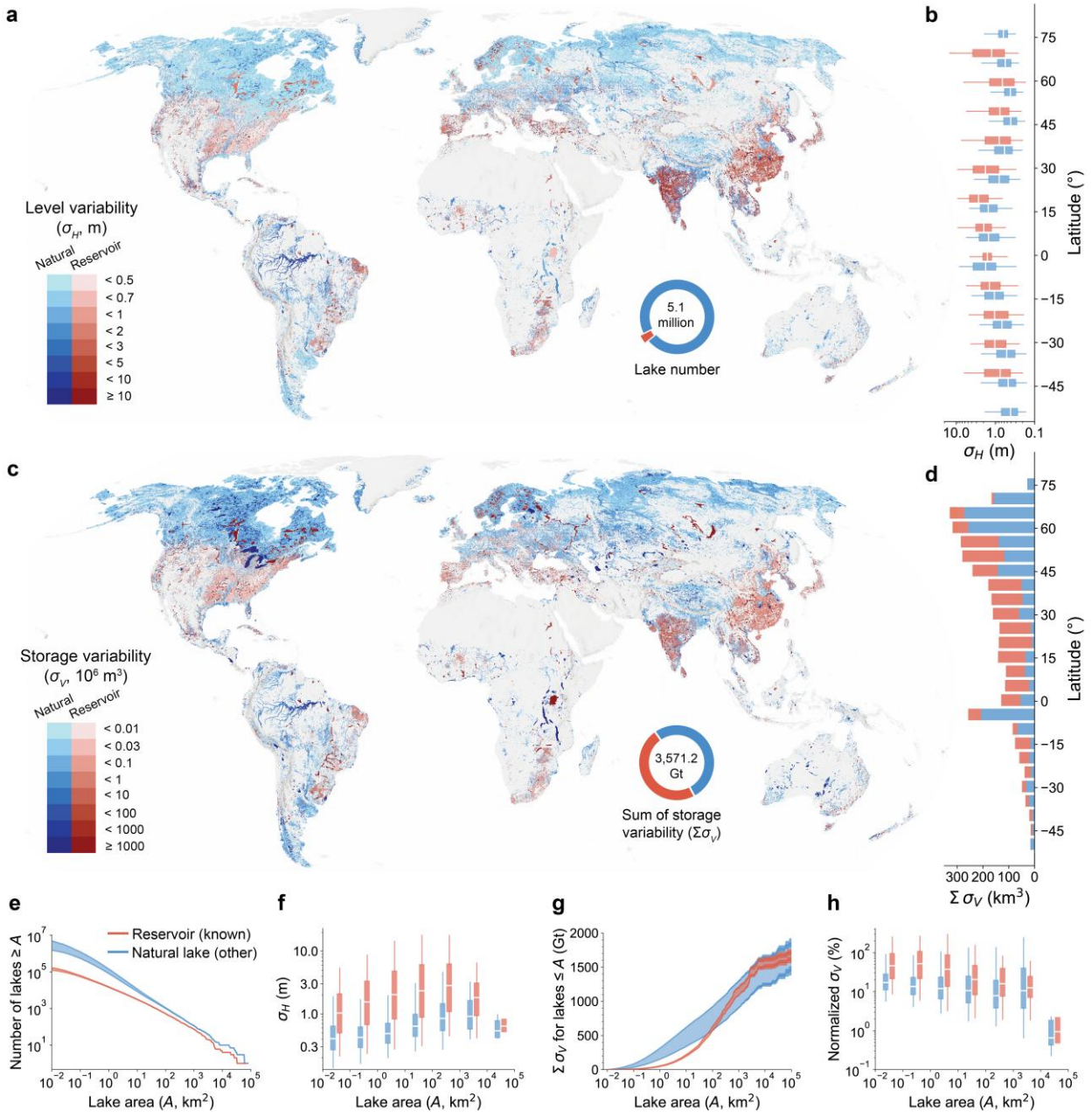


Figure 1. Seasonal water surface elevation (WSE) and storage variability of global lakes. Data comprise 169,600 known reservoirs and 4,911,293 natural lakes larger than 1 ha observed by SWOT from 21 July 2023 to 5 May 2025 (cycles 1–32; Methods). Variability (σ) is quantified by the interdecile range (IDR); blue denotes natural lakes and red denotes reservoirs. **a,** Global individual-lake WSE variability (σ_H). **b,** Latitudinal σ_H in 10° bands. **c,** Global individual-lake storage variability (σ_V). **d,** Latitudinal distribution of summed individual-lake storage variability ($\sum\sigma_V$) in 5° bands. **e,** Cumulative number of observed lakes with area ≥ A. **f,** σ_H across lake-area classes. **g,** Cumulative $\sum\sigma_V$ for lakes with area ≤ A. **h,** Storage variability normalized by estimated lake volume[27] (σ_V / V) across lake-area classes. In **b**, **f**, and **h**, boxes show the 25th–75th percentiles, white

lines indicate medians and whiskers the 5th–95th percentiles. In **e** and **g**, lighter shaded envelopes span estimates based on tier-1 lakes alone (lower bound) to both tiers combined (upper bound); in **g**, darker bands surrounding each bound indicate the corresponding propagated 1-sigma uncertainties, assuming independent lake-level errors (Methods). Panels **f** and **h** include both tiers; tier-1-only distributions are shown in Extended Data Figs. 3 and 4, respectively, and exhibit patterns similar to the corresponding distributions when both tiers are included.

Coherence shapes the global lake cycle

The global lake storage cycle emerges from both the magnitude and temporal alignment of individual lake phenologies. We aggregated tier-1 storage trajectories across the full observation period and calculated statistics over the annual window with the most complete global coverage (cycles 5–22), while using tier-2 lakes to constrain the upper uncertainty bound (Methods). The resulting global cycle exhibits seasonal variability of 719.0 (+209.8/−82.3) Gt, accounting for 12.8% of that in terrestrial water storage (TWS) (Fig. 2a,b; Extended Data Table 1). Integrating absolute storage changes across the annual lake cycle yields a turnover of 3,364.0 Gt, over five times the annual discharge of the Mississippi River[28].

A comparison with rivers underscores the distinction between storage capacity and seasonal dynamics. Although lakes hold nearly two orders of magnitude more water than river channels[11], their seasonal storage variability is only about three times as large. Reservoir variability alone, 297.1 (+68.1/−44.5) Gt, is comparable to global river-channel variability, 265.9 (±85.6) Gt[8], placing regulated storage on a similar seasonal scale to the global river network.

Seasonal phase further reveals a hierarchy of storage mediation across the terrestrial water cycle. The net lake storage cycle reaches a minimum in December–February and peaks in June–July, broadly following TWS seasonality with a 2–3 month lag (Fig. 2a). At the aggregate scale, river storage lags TWS by ~1 month, natural lakes lag rivers by another 1–2 months, and reservoirs lag natural lakes by a further ~3 months, culminating in a nearly half-year phase offset between TWS and reservoirs. This phase progression reflects successive effects of runoff generation and routing, natural-lake retention and hydrologic memory, and managed reservoir storage and release[19,25].

Large variability within individual reservoirs does not translate proportionally into a globally dominant aggregate signal. While reservoirs contribute roughly as much as natural lakes to summed individual-lake variability, their aggregated cycle is less than half as variable as the natural-lake aggregate (Fig. 2b,d; Extended Data Table 1). This attenuation reflects substantial phase cancellation, which we quantify using a coherence ratio, defined as aggregated variability relative to the summed variability of constituent

lakes. Reservoir coherence is only 18.6%, compared with 46.3% for natural lakes, exposing greater diversity among operating regimes. Meanwhile, the aggregate reservoir cycle is anticorrelated with every natural-lake typology (Fig. 2c), indicating that regulated storage collectively follows a seasonal rhythm distinct from that of natural systems[29,30].

Coherence also reshapes aggregate contributions of natural-lake typologies (Fig. 2c,d; Extended Data Table 1). On-river lakes drive 44.6% of global net lake variability, followed by floodplain (26.2%), permafrost-zone (10.6%), endorheic (7.5%) and glacial lakes (0.7%). Coherence reaches ~50% for on-river, floodplain and permafrost-zone lakes but falls to only 12.1% for glacial lakes. Higher coherence elevates the aggregate permafrost-zone signal above that of endorheic lakes despite lower summed individual variability. Conversely, phase cancellation strongly attenuates the glacial-lake signal despite relatively large local fluctuations.

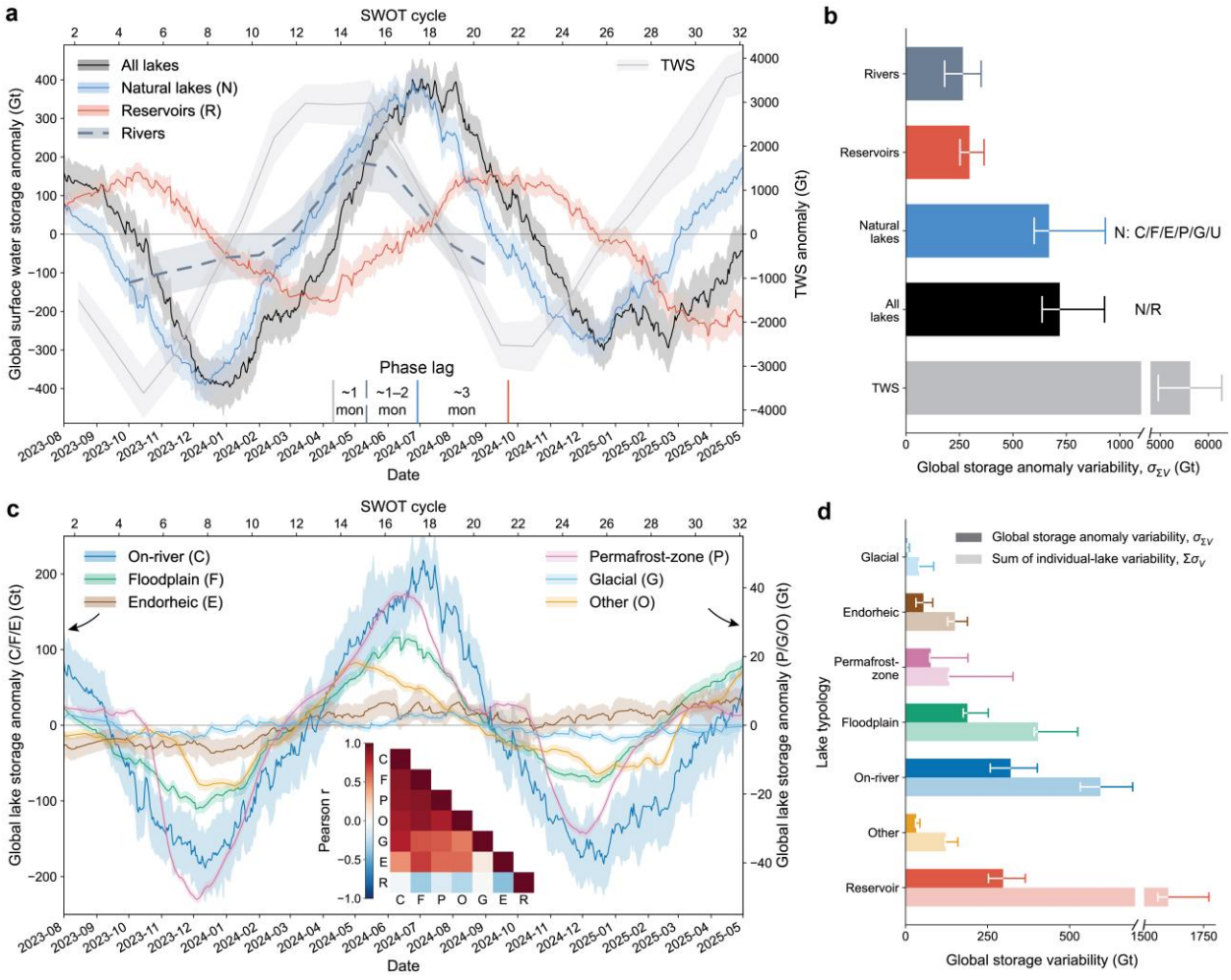


Figure 2. Global seasonal lake storage cycles in the terrestrial water cycle. a, Aggregated storage-anomaly time series of natural lakes, reservoirs, and both,

compared with global river-channel storage and terrestrial water storage (TWS) anomalies from July 2023 to May 2025 (SWOT cycles 1–32). Lake trajectories comprise 1.6 million tier-1 lakes; vertical markers indicate approximate phase offsets among storage cycles. **b,** Magnitudes of seasonal variability of the storage-anomaly series ($\sigma_{\sum V}$) in **a**. **c,** Global natural-lake anomalies disaggregated among on-river, floodplain, endorheic, permafrost-zone, glacial and other unclassified typologies. Inset shows Pearson correlations among lake-typology and reservoir storage-anomaly series. The global map of lake typologies is given in Extended Data Fig. 2. **d,** Magnitudes of seasonal storage variability by lake typology. Dark bars show variability of the globally aggregated storage-anomaly series ($\sigma_{\sum V}$); light bars show the summed individual-lake variability ($\sum\sigma_V$). Throughout this study, seasonal variability (σ) is quantified using the interdecile range (IDR; 90th minus 10th percentile). Variability statistics in **b** and **d** are calculated over cycles 5–22 (12 October 2023 to 21 October 2024), the annual window with the most complete global lake coverage (Methods) and the available water year for the river storage estimates[8]. Shading in **a** and **c** shows propagated time-series uncertainty. Error bars in **b** and **d** propagate tier-1 variability uncertainty, with upper bounds additionally incorporating estimated tier-2 contributions (Methods). Detailed numeric statistics are given in Extended Data Table 1.

Archetypes organize lake phenology

Across individual lakes, the timing of seasonal extrema reveals spatial patterns that mirror major hydroclimatic regimes (Fig. 3a). Storage maxima commonly occur in May–June across boreal regions following thaw and snowmelt[31], January–March in maritime west-coast and Mediterranean climates[32], and September–November across tropical and monsoonal regions[33,34]. Regional gradients are also evident, including an April-to-June downstream progression across the Amazon Basin[35], a winter-to-spring shift from western Europe into Central Asia[32,36], and a south-to-north migration across the East Asian monsoon belt[37]. Annual minima generally resemble these patterns in opposite phase (Extended Data Fig. 5). Extrema timing, however, captures only one dimension of phenology. We therefore clustered standardized annual trajectories from tier-1 lakes to identify phenological archetypes defined by both seasonal phase and trajectory shape – including the rate and persistence of storage rise and decline – independent of absolute variability magnitude (Methods).

From this diversity emerge seven recurrent archetypes of lake phenology (P1–P7; Fig. 3b,d; seasonal descriptors follow Northern Hemisphere convention). P1 (11.7% of classified lakes) features a subdued cold-season high followed by gradual warm-season drawdown, prominent in temperate maritime and northern coastal climates and Northeast Monsoon regions of South and Southeast Asia. P2 (17.0%) exhibits a broad winter–early-spring high, spanning Northern Hemisphere winter–spring recharge

regimes and austral-summer regions. P3 (14.1%) shows a pronounced spring peak and late-autumn–early-winter trough across both northern mid- to high latitudes and tropical wet-season systems, including the middle-lower Amazon and East African lakes.

P4–P6, collectively widespread across boreal and temperate regions, form a depletion–recovery sequence with progressively later winter minima and distinct recovery trajectories. P4 (17.6%) reaches an early-winter minimum followed by a prolonged winter–spring plateau; P5 (17.0%) shows a mid-winter trough followed by broad spring–summer recovery; and P6 (12.7%) continues drawing down into late winter before rising sharply to a midsummer crest. P4 is concentrated among small high-latitude lakes, where its plateau may reflect ice-restricted outflow and subsequent snowmelt-driven filling, alongside cryospheric effects on radar-retrieved elevations (Supplementary Information section 6)[38-41]. P6 is common among larger lakes and reservoirs, including systems with managed winter drawdown before spring freshet (Supplementary Fig. S4)[42,43]. P7 (9.8%) instead exhibits a pronounced monsoonal trajectory, with a pre-monsoon minimum followed by rapid filling to a late-summer–autumn maximum, and is concentrated across South and Southeast Asia, Central America, and the Sahel.

Archetypes differ sharply between numeric abundance and aggregate storage variability (Fig. 3c). P7 is the least abundant yet exhibits the largest aggregate seasonal swing (542.9 Gt), equivalent in magnitude to 75.7% of the global cycle and consistent with its disproportionate reservoir representation (Fig. 4b). By contrast, P4 is the most abundant but carries the weakest aggregate variability (14.2%). P3, P5 and P6 combine broad prevalence with large storage signals, while winter-high P1 and P2 carry weaker variability. Within-archetype coherence ratios mostly exceed 75%, confirming that these archetypes represent recurrent, internally aligned modes of seasonal behavior.

These coherent modes clarify how individual lake phenologies scale up to shape the planetary water storage signal: variability magnitude determines how much seasonal change each archetype carries, whereas phase and trajectory shape determine whether that variability reinforces or cancels the global cycle (Fig. 3e). P3–P6 broadly align with the global spring–summer rise and therefore reinforce net variability, led by P5 (+50.8%) and P6 (+39.2%). P1, P2 and P7 are phase-offset and instead dampen the global signal. Most strikingly, P7 carries the largest seasonal swing yet reduces global net variability by 40.1%, because late-summer–autumn filling across tropical and monsoonal regions opposes the simultaneous global recession.

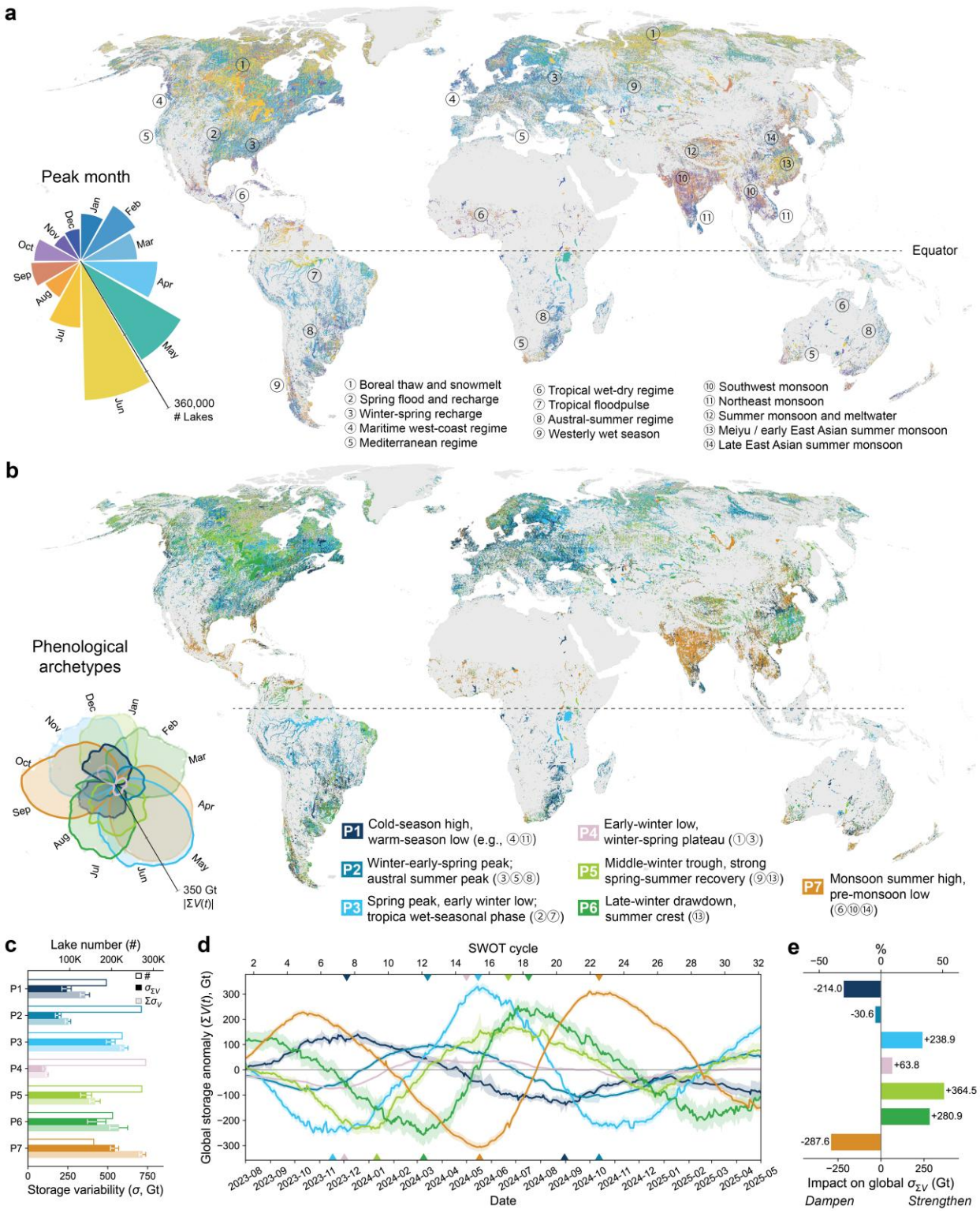


Figure 3. Seasonal timing and phenological archetypes of global lake storage. a, Global distribution of annual storage-maximum (peak) months. The polar histogram shows their global frequency; circled numbers mark representative hydroclimatic regimes associated with regional lake seasonality. Annual storage-minimum months are

shown in Extended Data Fig. 5. **b,** Global distribution of seven phenological archetypes (P1–P7); descriptors use Northern Hemisphere seasonal terminology. The polar plot shows the absolute magnitude of globally aggregated storage-anomaly trajectories ($|\sum V(t)|$) from October 2023 to October 2024, with outline opacity increasing from negative to positive anomalies. **c,** Lake abundance and storage variability by archetype. Outlined bars show lake number (upper axis); solid and light bars show variability of the globally aggregated storage-anomaly trajectory ($\sigma_{\sum V}$) and summed individual-lake variability ($\sum \sigma_V$), respectively (lower axis). **d,** Globally aggregated lake storage-anomaly trajectories ($\sum V(t)$) for P1–P7 from July 2023 to May 2025 (cycles 1–32); shading denotes uncertainty, and colored triangles mark the approximate timing of annual maxima (top) and minima (bottom). **e,** Contribution of each archetype to global net storage variability, quantified as the change in global $\sigma_{\sum V}$ when that archetype is removed. Positive values indicate that removal strengthens the global cycle, whereas negative values indicate that removal dampens it. Analyses include 1.6 million tier-1 lakes; variability statistics are computed over cycles 5–22, as in Fig. 2 (Methods).

Phenology encodes memory and regulation

To interpret the processes underlying lake phenology, we related each tier-1 storage trajectory to naturalized upstream discharge (Q) at the 21-day SWOT-cycle timescale (Methods). Contemporaneous Spearman correlations (ρ) reveal widespread seasonal alignment between lake storage and upstream hydroclimate (Fig. 4a–d). Among natural lakes, 50.5% exhibit at least moderate correlation ($\rho > 0.3$), accounting for 63.4% of summed storage variability ($\sum \sigma_V$); 27.7% show stronger alignment ($\rho > 0.5$), representing 41.0% of $\sum \sigma_V$ (Extended Data Fig. 6; Fig. 4c). Reservoirs show similar count-based proportions, but their storage-variability shares fall to 48.8% and 23.8%. This disparity reflects progressively weaker storage–Q alignment as reservoir size increases (Fig. 4d), suggesting overall greater operational influences in larger systems. Counter-seasonal behavior ($\rho < -0.3$) accounts for 6.1% and 8.7% of natural-lake and reservoir $\sum \sigma_V$, respectively, with hotspots in heavily regulated regions such as southwestern China and the southern Canadian Shield, alongside natural systems like poorly drained Pampean lakes, where groundwater exchange and evaporation exert stronger control on the local water balance[44].

Storage–Q phase offsets (τ) reveal substantial storage-mediated memory: 48.7% of natural-lake storage variability exhibits at least moderate delay ($\tau \geq 2$ SWOT cycles) relative to upstream forcing, rising to 76.0% for reservoirs (Fig. 4e–h). Individual examples, however, reveal diverse responses (Fig. 4a,e). Near-synchronous natural responses include shallow floodplain or on-river systems such as Buabua Lake (Amazon) and Dongting Lake (Yangtze), whereas the larger Laurentian Great Lakes sustain longer hydroclimatic memory ($\tau = 3$–5). Among managed systems, the

multipurpose Three Gorges Reservoir exhibits a nearly counter-seasonal pattern with a pronounced six-cycle delay, but Krishna Raja Sagara Reservoir, primarily operated for water supply and irrigation, shows moderate monsoon-season retention ($\tau = 2$). Along the Tietê cascade, downstream reservoirs become progressively misaligned with naturalized forcing; meanwhile, managed pre-freshet drawdown at Lake Geneva produces a storage lead before Q. Despite this diversity, phase delays generally increase with lake size and are systematically longer in reservoirs, revealing how natural storage memory is further reorganized by human regulation (Fig. 4h).

The joint ρ–τ structure further clarifies how hydroclimatic coupling and storage memory shape phenological archetypes (Fig. 4b,f). Among natural lakes, P6 exhibits the weakest hydroclimatic alignment but the longest delays, consistent with its prevalence among larger systems: 56.5% of natural-lake $\sum\sigma_V$ in P6 shows weak or counter-seasonal correlation ($\rho \leq 0.3$), yet 95.2% of this share recovers to at least moderate alignment after phase adjustment. This recovery ranges from 78.9% to 95.2% across natural-lake archetypes. Reservoirs display even stronger phase reorganization: weak or counter-seasonal correlations account for over 50% of $\sum\sigma_V$ in P1, P6, and P7, yet recovery reaches 92.7–99.0% across archetypes. Globally, this recovery reaches 89.9% for natural lakes and 97.6% for reservoirs, underscoring that apparent storage–Q decoupling often reflects temporal reorganization of hydroclimatic signals through hydrologic memory and human regulation.

Finally, 8.4% of global lake storage variability exhibits storage-leading offsets ($\tau < -1$), concentrated among cryosphere-affected natural lakes, particularly across the permafrost zone and P4 (Fig. 4e,f). Gauge validation and environmental diagnostics suggest that these offsets may combine genuine cryosphere-mediated phenology – including locally retained snowmelt, thaw-related subsurface inflows, seasonal drainage connectivity changes – with lake-storage retrieval and discharge-simulation uncertainties[38-41,45], warranting cautious interpretation (Supplementary Information section 6). Nonetheless, sensitivity analyses show that these negative offsets do not materially alter the broader size dependence of τ or the systematically longer reservoir delays (Extended Data Fig. 7; Supplementary Fig. S28). Removing snowmelt from simulated Q also weakens ρ across high latitudes (Extended Data Fig. 8) and shifts peak-Q timing later relative to lake-storage maxima (Supplementary Fig. S26), suggesting that lake storage phenology in these regions retains the signature of snowmelt-driven freshets central to boreal hydrology.

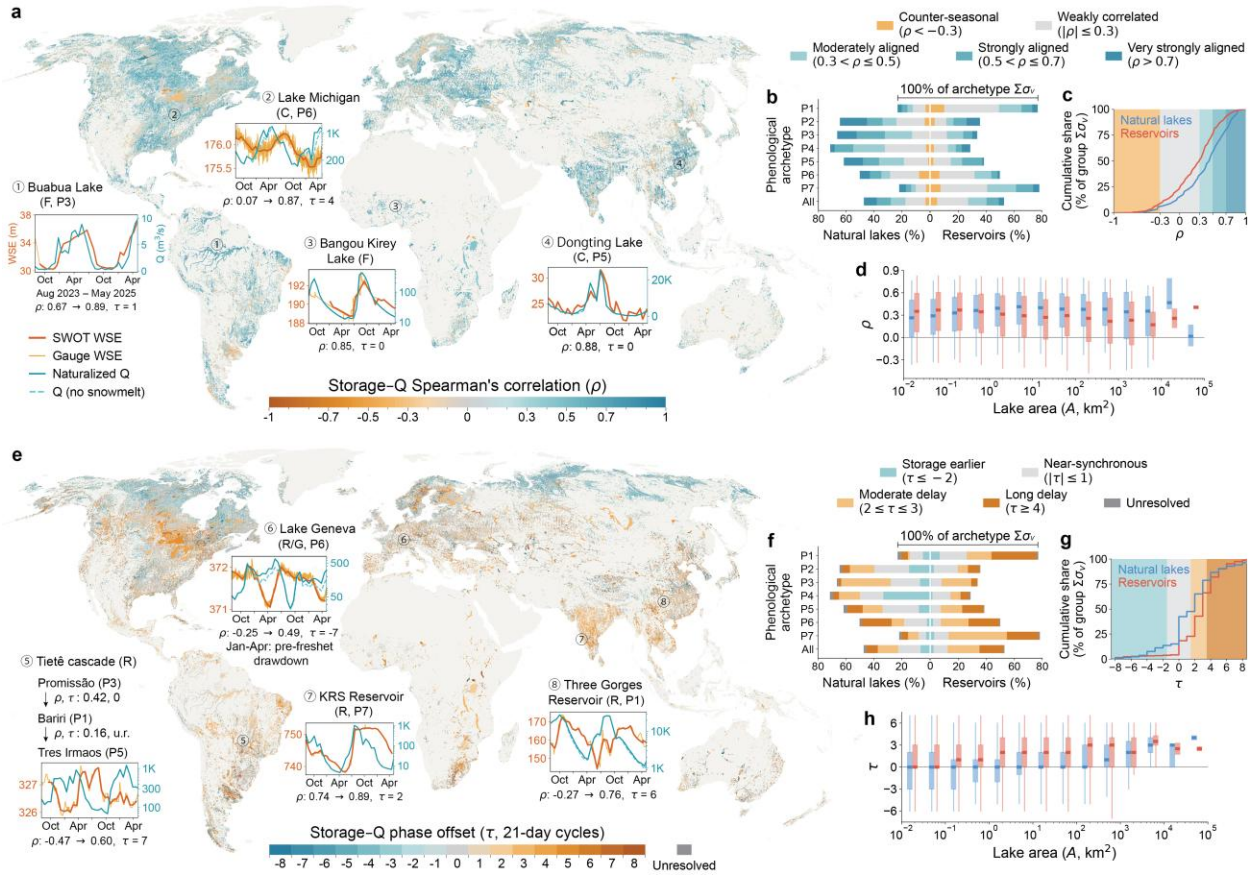
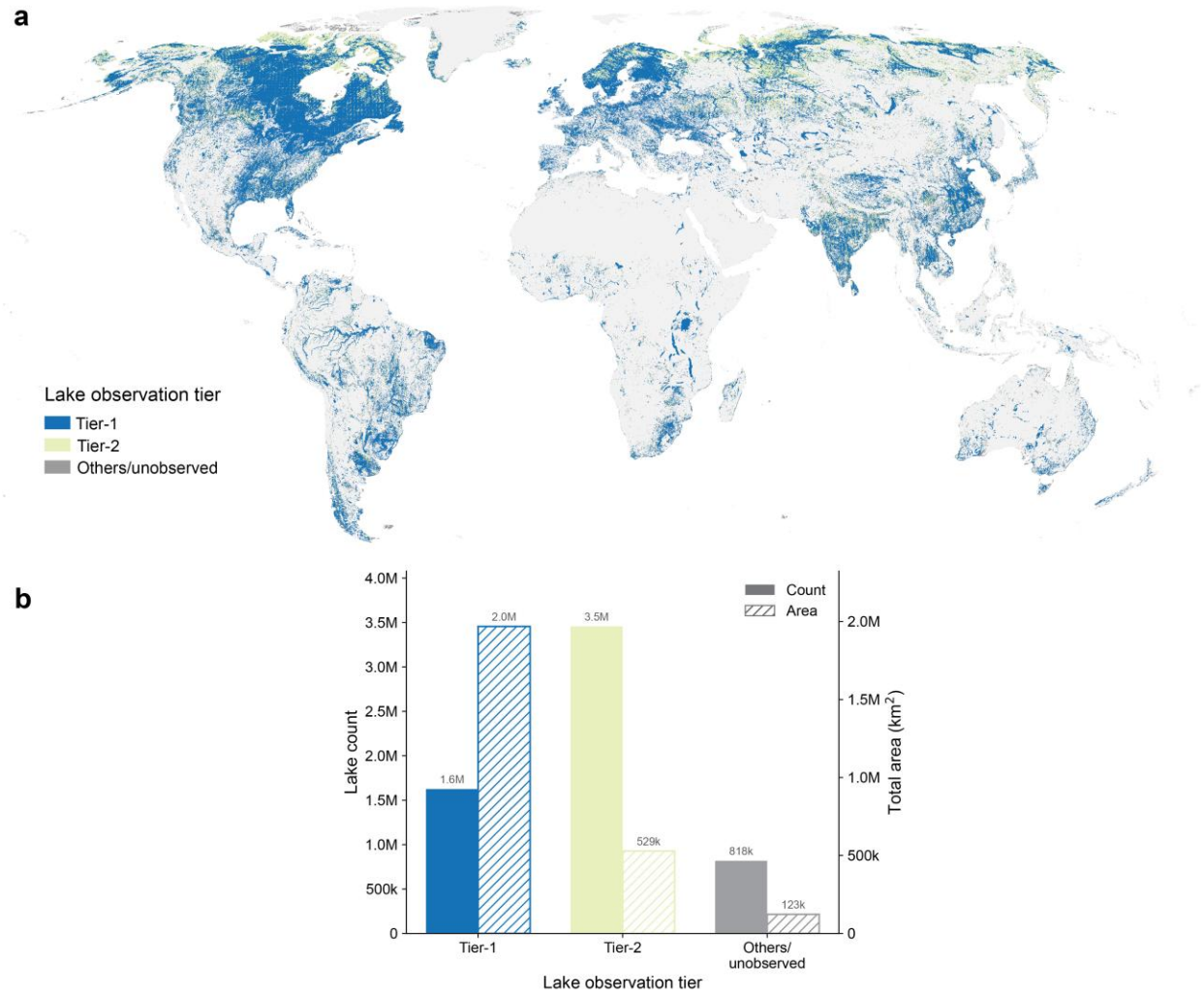


Figure 4. Coupling of lake storage phenology to upstream hydroclimatic forcing. **a,** Global map of contemporaneous Spearman correlation (ρ) between tier-1 lake storage anomalies and modeled naturalized lake-outlet discharge (Q), a surrogate for upstream hydroclimatic forcing, at the 21-day SWOT-cycle scale, July 2023–May 2025. **b–d,** Summaries of ρ by phenological archetype (**b**), cumulative distribution (**c**), and lake size (**d**) for natural lakes and reservoirs. Distributions in **b** and **c** are weighted by individual-lake storage variability (σ_V); "All" includes lakes with and without archetype assignments. **e,** Global map of storage–Q phase offset (τ) in 21-day cycles; positive values mean storage lagging Q. **f–h,** Summaries for τ. Numbered examples illustrate diverse coupling behaviors: (1) minor floodplain buffering in Buabua Lake (Amazon), (2) prolonged hydrologic memory in Lake Michigan, (3) near-synchronous response in shallow Bangou Kirey Lake (Sahel), (4) near-synchronous response in large on-river Dongting Lake (Yangtze), (5) progressive downstream misalignment along the Tietê reservoir cascade, (6) managed pre-freshet drawdown and storage lead in Lake Geneva, (7) monsoon-season retention and delayed release in Krishna Raja Sagara Reservoir, and (8) strong seasonal displacement under multi-purpose regulation at Three Gorges Reservoir. Series show cycle-scale SWOT WSE and Q, with daily gauge WSE where available. Parenthetical labels indicate typologies and assigned archetypes. Horizontal arrows link ρ before and after phase adjustment; "u.r." denotes unresolved τ.

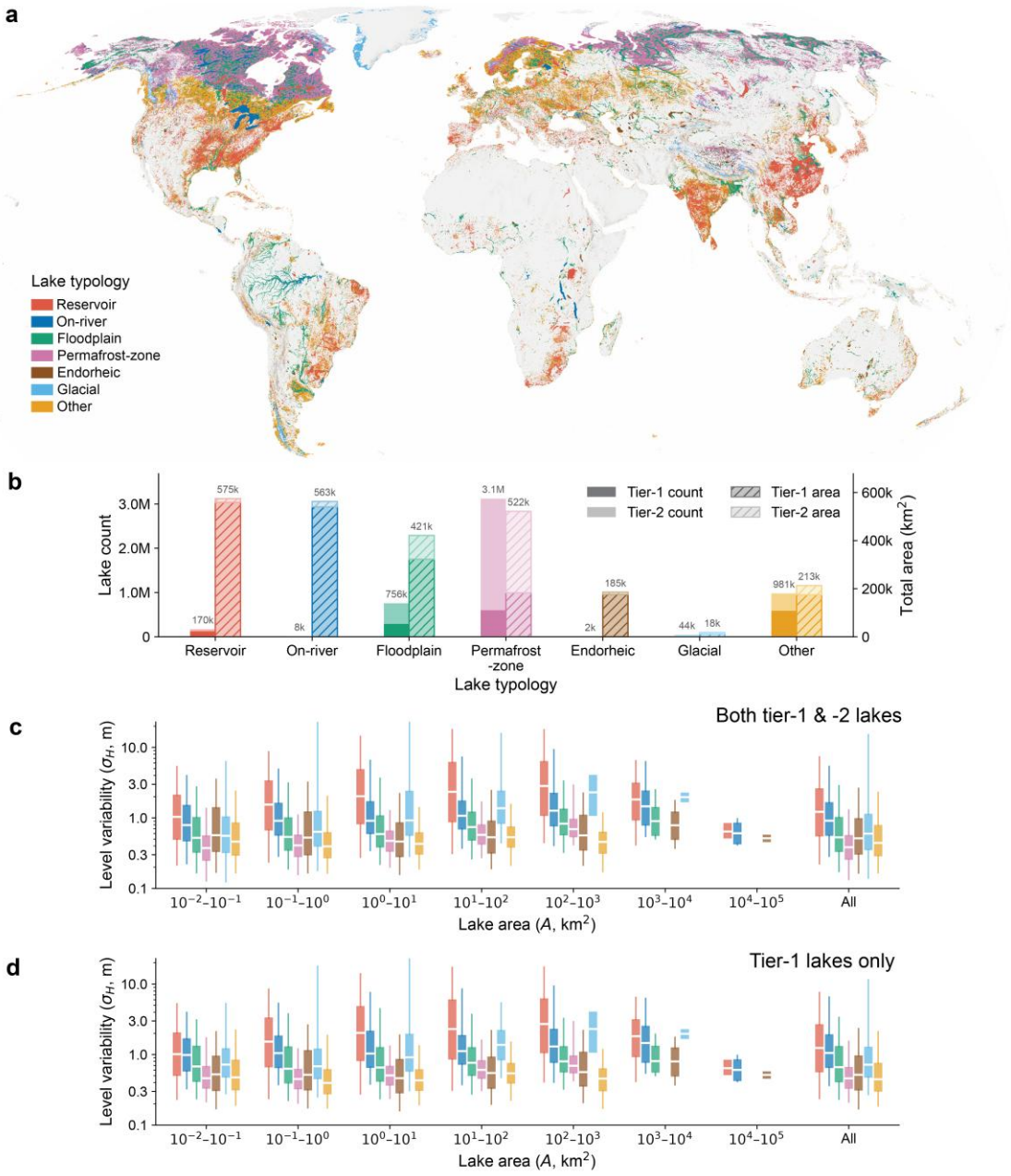
Discussion

We introduce the concept of lake storage phenology and resolve it across millions of lakes globally through a four-level framework linking individual-lake variability, planetary-scale coherence, recurrent archetypes, and coupling to hydroclimatic forcing and regulation. This near-census analysis is enabled by SWOT wide-swath altimetry but constrained by its nominal 21-day orbit cycle and the inevitable sampling irregularities introduced by quality filtering. Our analysis therefore emphasizes seasonal dynamics and may under-resolve storage responses to daily-to-weekly hydrologic events. Likewise, to reduce sensitivity to routing uncertainties in simulated discharge, our storage–Q coupling analyses operate at the SWOT-cycle timescale and primarily diagnose seasonal timing. High correlation or near-synchronous response does not necessarily imply weak hydrologic mediation: lakes can preserve the seasonal rhythm of upstream forcing while still attenuating the amplitude of flow variability through low-pass buffering[46]. Nevertheless, our quality-controlled storage time series retain finer temporal variability wherever sampling permits, offering opportunities for more targeted regional investigations of these buffering processes. More broadly, our four-level framework contributes to a scalable paradigm for integrating the diverse seasonal storage rhythms of millions of lakes into a planetary-scale understanding of hydrology.

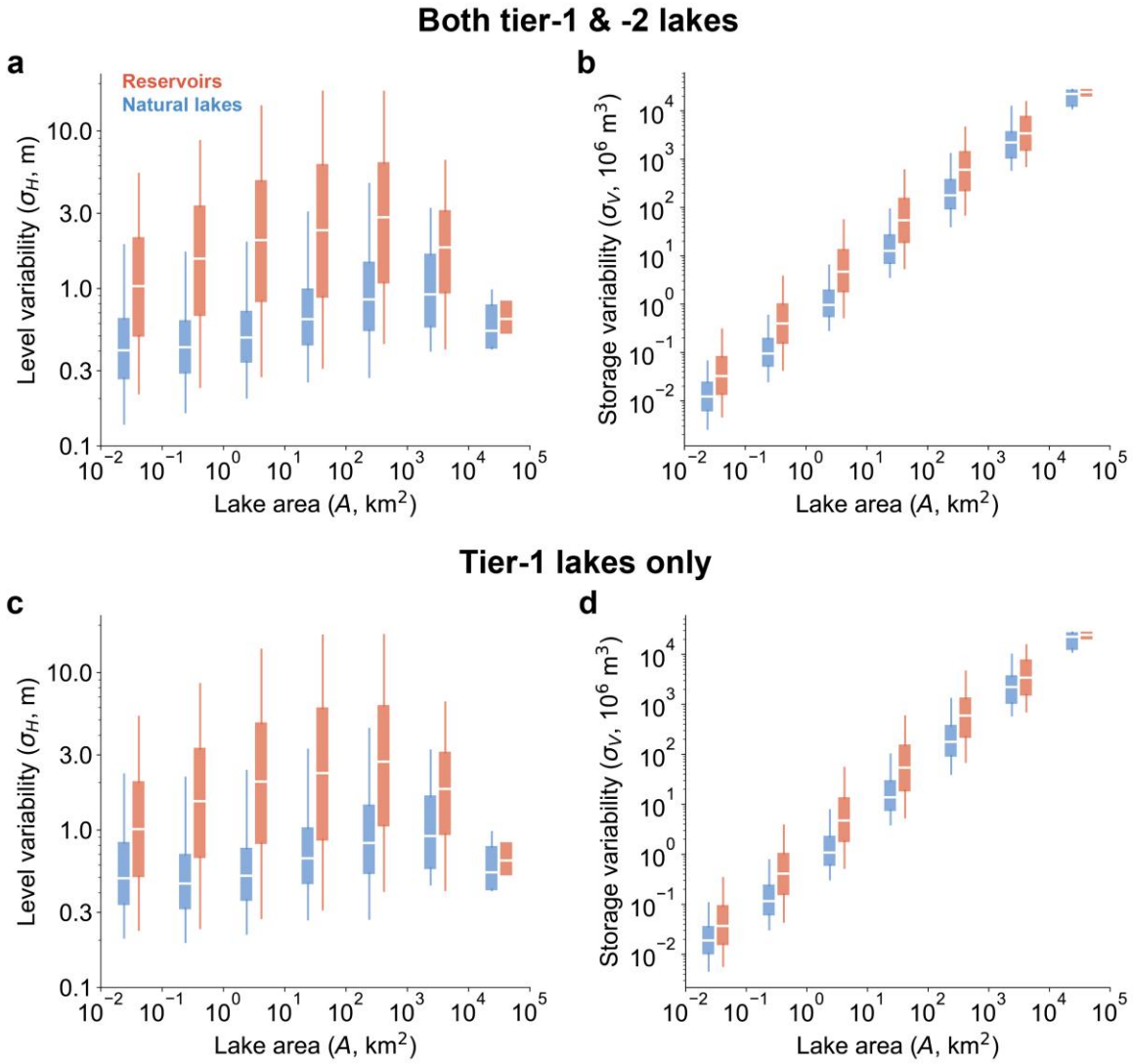
A core implication of this study is that lakes, through their storage phenology, act as both sentinels of regional hydroclimate and mediators of streamflow through storage memory and human regulation. Representing this mediation within drainage networks remains a major challenge for hydrological models[13,47,48] and an essential step towards bridging lake and river sciences. Our global characterization of lake storage phenology provides two complementary pathways: an extensive observational foundation for identifying generalizable patterns in lake hydrologic behavior and refining lake-module formulations and parameterizations; and temporally resolved storage-change constraints for data assimilation[29,49]. Incorporating these observed storage dynamics into modeling frameworks could improve characterization of flow seasonality, flood attenuation and low-flow support, while advancing understanding of fluvial–lacustrine controls on biogeochemical cycling[50]. As SWOT and future Earth-observing systems extend and refine the observational record, these findings establish lake storage phenology as a unique, measurable dimension of the terrestrial water cycle that integrates hydroclimate, storage memory, and anthropogenic impacts across scales.



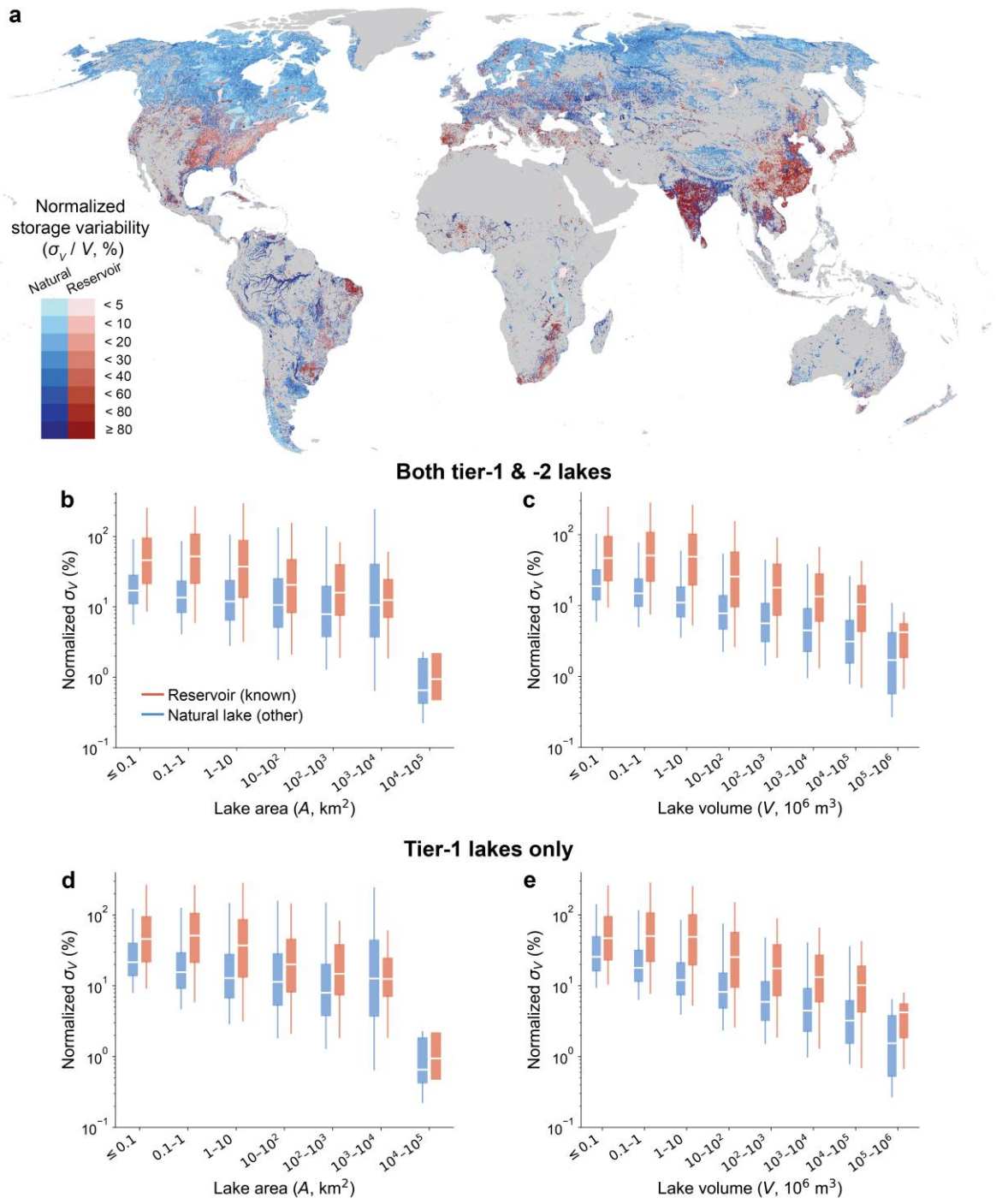
Extended Data Figure 1. Global distribution of lake observation tiers. a, Global map of SWOT prior lakes by observation tier. Tier-1 lakes are used for the primary seasonal storage variability and phenology analyses, whereas tier-2 lakes provide additional observed-lake coverage and inform uncertainty estimates (Methods). Other/unobserved lakes are not included in either tier. Note that unobserved lakes are predominantly small and may therefore be less visible on the map. **b**, Lake count and total lake area summarized by observation tier. Solid bars show lake count, and hatched bars show total area.



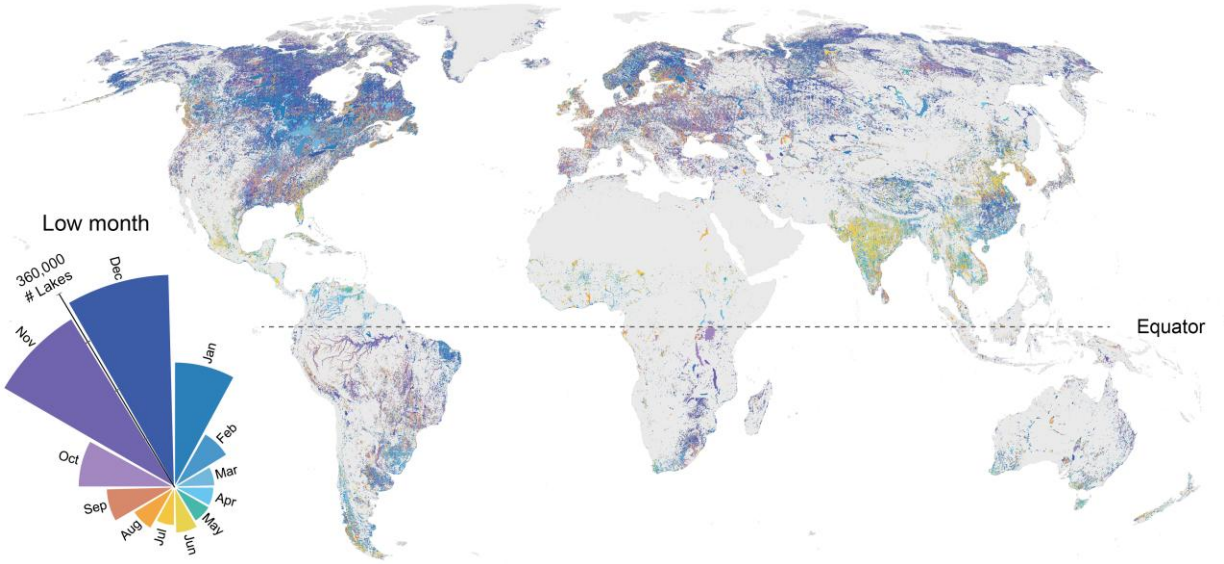
Extended Data Figure 2. Global distribution, abundance, and seasonal water surface elevation (WSE) variability across lake typologies. a, Global map of lakes classified as reservoirs, on-river, floodplain, permafrost-zone, endorheic, glacial, or other lakes. **b**, Lake count and total lake area by typology for tier-1 and tier-2 lakes. Filled bars show lake count and hatched bars show total area; darker and lighter shades denote tier-1 and tier-2 lakes, respectively. Values above bars indicate combined tier-1 and tier-2 totals. **c**, Seasonal WSE variability (σ_H) across lake-area classes for both tier-1 and tier-2 lakes. **d**, Same as c, but for tier-1 lakes only. The "All" category pools lakes across all area classes. In **c** and **d**, boxes denote the 25th–75th percentiles, white lines indicate medians, and whiskers span the 5th–95th percentiles.



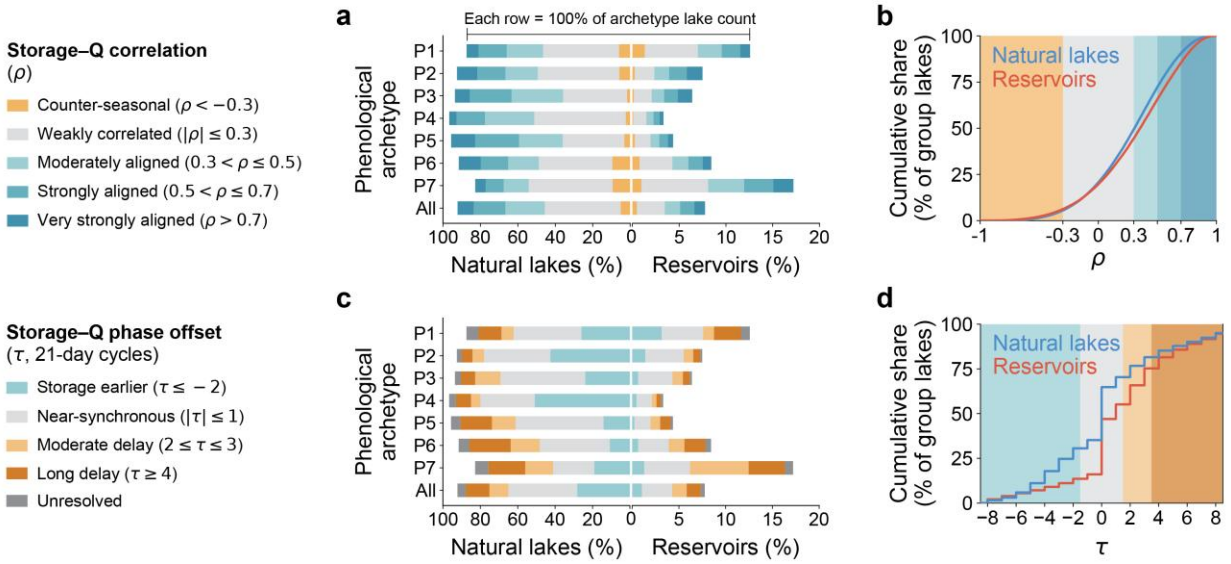
Extended Data Figure 3. Scaling relationships between seasonal lake variability and lake size. a, Boxplots of seasonal WSE variability, and **b**, water storage variability, based on tier-1 and tier-2 lakes combined. **c**–**d**, Same as **a**–**b**, but including tier-1 lakes only. Panel **a** is identical to Fig. 1f and is reproduced here for comparison.



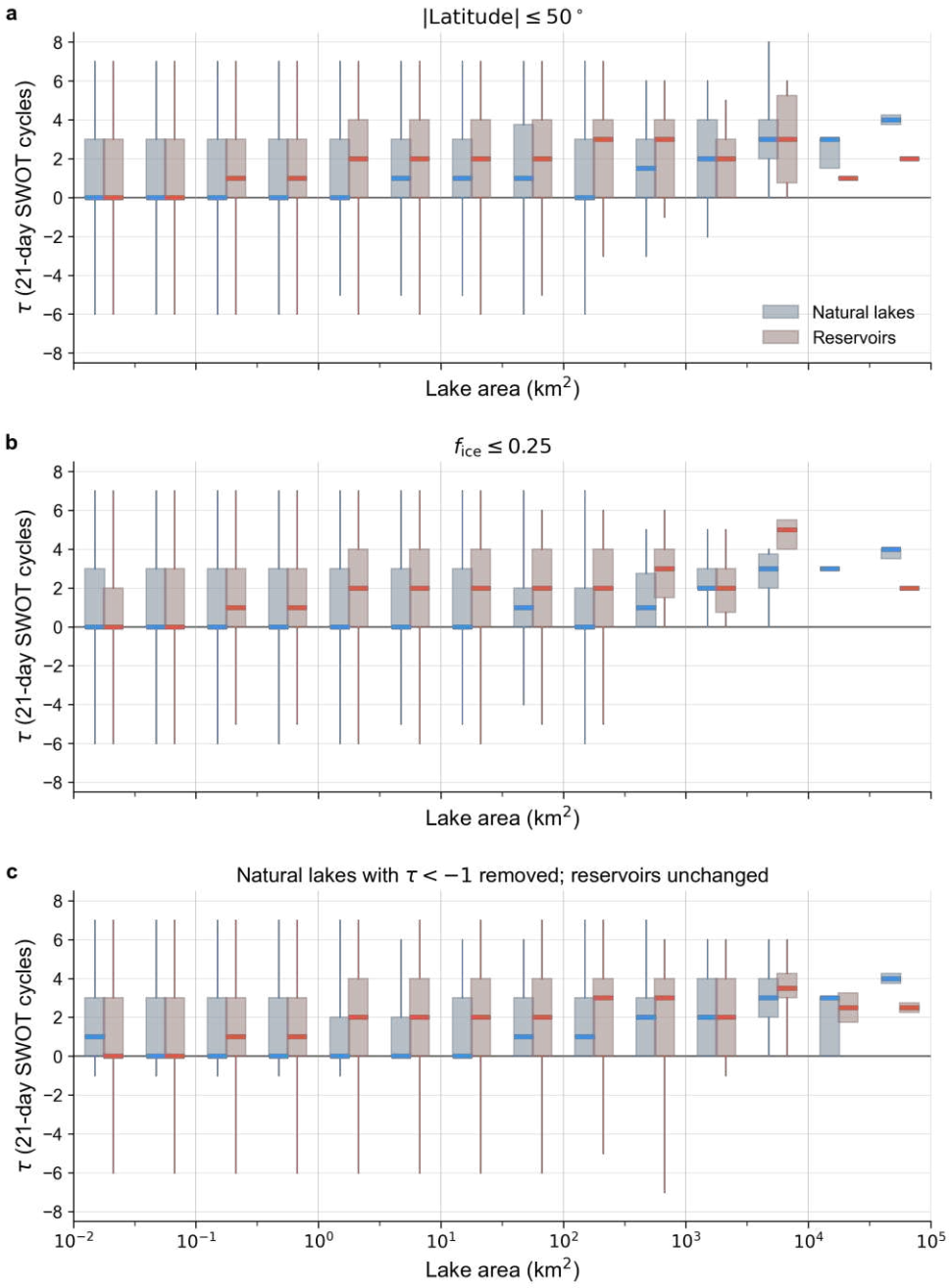
Extended Data Figure 4. Global patterns and size dependence of normalized seasonal lake storage variability. a, Global map of normalized seasonal lake storage variability expressed as a percentage. Normalized variability is calculated as the SWOT-observed seasonal storage variability divided by estimated lake volume[27]. **b,c,** Boxplots showing how normalized storage variability varies with lake area (b) and lake volume (c) for both tier-1 and tier-2 lakes. **d,e,** Same as b,c, but for tier-1 lakes only. Panel **b** is identical to Fig. 1h and is reproduced here for comparison.



Extended Data Figure 5. Annual storage-minimum timing of global lakes. The polar histogram shows the global frequency of storage-minimum (low) months. Spatial patterns of storage-minimum timing broadly mirror those of storage maxima (Fig. 3a) in opposite phase.

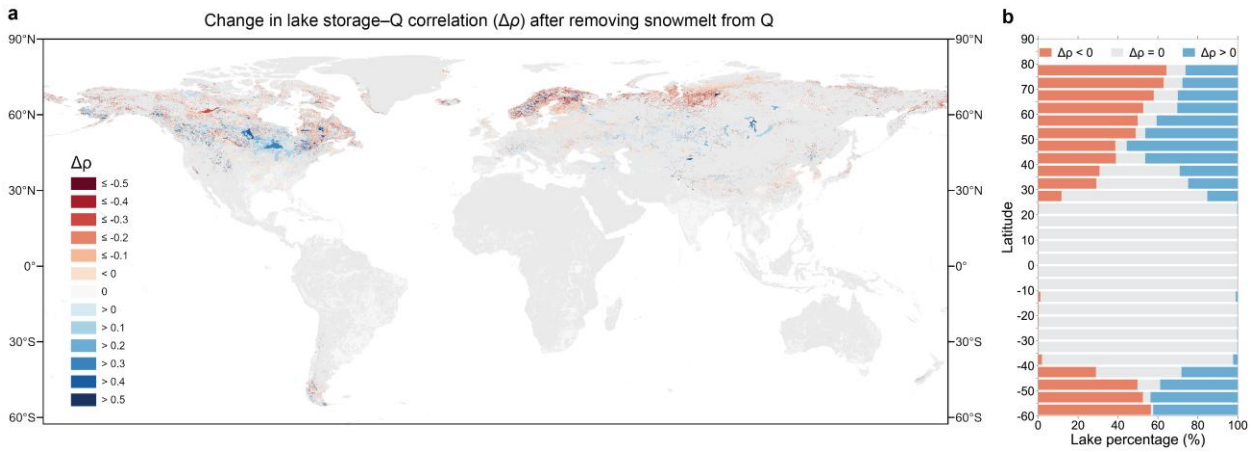


Extended Data Figure 6. Lake-count distributions of storage–Q correlation (*ρ*) and phase offsets (*τ*) across phenological archetypes. This figure provides the count-based counterpart to the storage-variability-weighted analysis in Fig. 4b,c,f,g. **a,** Distribution of contemporaneous storage–Q Spearman correlation (*ρ*) across the seven phenological archetypes (P1–P7) and all lakes. Mirrored bars (both sides of 0) partition each archetype's total lake count between natural lakes (left) and reservoirs (right), with colors indicating correlation classes defined in the legend. **b,** CDF of *ρ* for natural lakes and reservoirs, with global median values of 0.31 and 0.35, respectively. **c,** As in **a**, but for storage–Q phase offset (*τ*, in 21-day SWOT cycles), including cases without resolved lags. **d,** CDF of resolved *τ* for natural lakes and reservoirs. The median *τ* was 0 cycles for natural lakes and +1 cycle for reservoirs, consistent with a systematically later reservoir phase. Unresolved lakes, representing 4.9% of natural lakes and 5.3% of reservoirs, remain in the group but are not assigned a numerical *τ*; accordingly, the cumulative curves terminate at 95.1% and 94.7%, respectively.

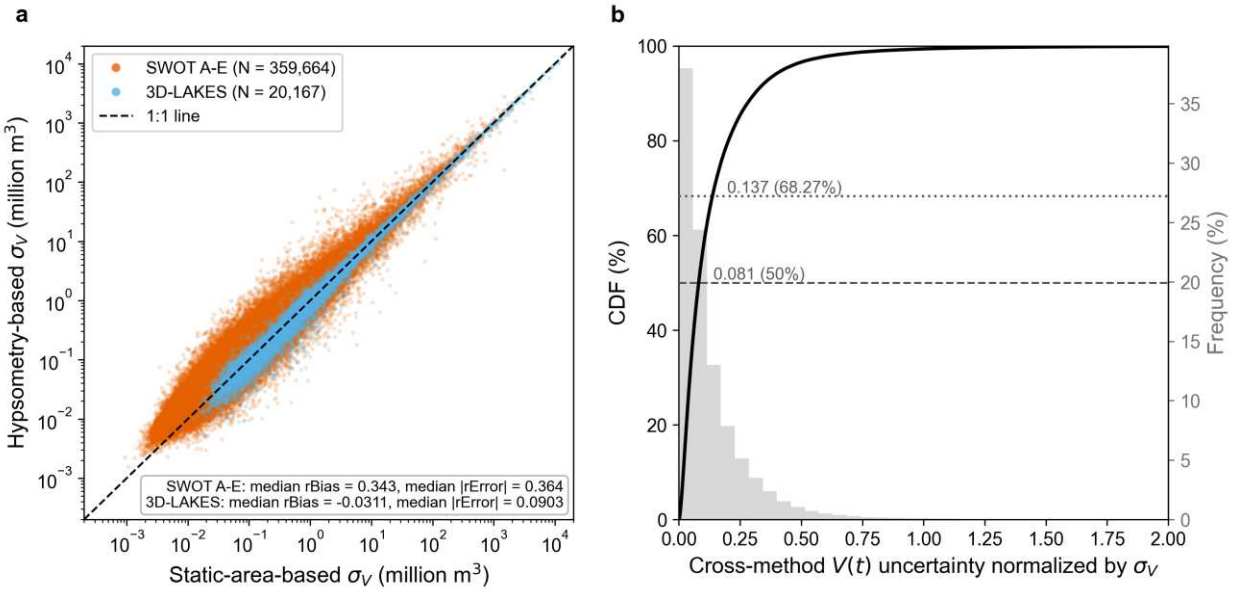


Extended Data Figure 7. Sensitivity of size-stratified storage–Q phase offsets to negative natural-lake lags. a, Natural lakes restricted to |latitude| ≤ 50°. **b,** Natural lakes restricted to low full-ice exposure f_{ice} ≤ 0.25. **c,** Natural lakes restricted to $\tau \geq -1$ SWOT cycles, thereby excluding strongly negative offsets. Reservoir populations are unchanged in all three panels, providing a conservative test of whether cryosphere-associated or strongly negative natural-lake offsets drive the size dependence in Fig. 4h. Across 14 populated size classes, median natural-lake τ remains positively associated with lake size under all treatments (Spearman ρ = 0.893 in **a**, 0.876 in **b**, and 0.829 in **c**, P-value < 0.001 for all three), supporting the overall size dependence

observed in Fig. 4h. Note that the relationship is not monotonic across small and medium lakes, with the clearest increase emerging toward larger size classes. Nevertheless, positive phase offsets remain widespread across smaller size classes, indicating that storage memory is not confined to large lakes and may remain hydrologically important across the lake-size spectrum. Boxes show the interquartile range with median lines, and whiskers span the 5^{th}–95^{th} percentiles.



Extended Data Figure 8: Change in lake storage–Q correlation after removing snowmelt runoff from upstream discharge. a, Global map of changes in the Spearman correlation (ρ) between SWOT-derived lake storage anomalies and mode-simulated naturalized lake-outlet discharge (Q) after excluding snowmelt runoff from Q. Red tones indicate reduced ρ, whereas blue tones indicate increased ρ. Negative changes are more widespread overall, particularly across northern high-latitude regions, suggesting that snowmelt contributes importantly to the seasonal coupling between lake storage and catchment discharge in many snow-influenced systems. **b**, Latitudinal distribution of lake percentages with negative, neutral, and positive correlation changes, showing that correlation decreases are especially common across northern mid- to high-latitude bands.



Extended Data Figure 9. Comparison of lake seasonal storage variability estimated using different storage-construction methods and the resulting cross-method uncertainty. For each lake, $V(t)$ denotes the storage-anomaly time series relative to the temporal mean, and σ_V denotes seasonal storage variability calculated from $V(t)$. $V(t)$ time series were constructed using three WSE-to-storage conversion methods: static PLD area, SWOT-derived area-elevation relationships (SWOT A-E), and 3D-LAKES hypsometry[51] (Methods); the latter two are collectively referred to as hypsometry-based methods. **a,** Comparison of σ_V estimates from the three methods. Each point represents one lake with valid paired estimates. Median relative bias (rBias) and median absolute relative error (|rError|) are reported relative to the static-area-based estimate. **b,** CDF and frequency distribution of relative cross-method uncertainty in $V(t)$. For each lake, cross-method uncertainty was computed as the root-mean-square spread among available method-specific $V(t)$ at retained observation times and normalized by the static-area-based σ_V. Horizontal lines indicate the median and 68.27th-percentile relative uncertainty, corresponding to 0.081 and 0.137, respectively.

Extended Data Table 1. Global seasonal water storage variability by water component. For lake components, central estimates are based on tier-1 lakes. Global storage variability represents the variability of aggregated storage-anomaly time series, whereas the sum of individual-lake variability represents the arithmetic total of variability across individual lakes. Lower bounds account for uncertainty in tier-1 estimates, while upper bounds additionally incorporate the potential contribution of tier-2 lakes. Contribution (%) is calculated relative to the all-lake estimate for the corresponding variability metric. The coherence ratio (%) is defined as the ratio of aggregated storage-anomaly variability to the sum of individual-lake variability; lower values indicate greater phase cancellation among individual lake variations.

Water storage component	Global storage variability ($\sigma_{\sum v}$)				Sum of individual-lake variability ($\Sigma\sigma_v$)				Coherence ratio (%)
	Estimate (Gt)	Lower (Gt)	Upper (Gt)	Contribution (%)	Estimate (Gt)	Lower (Gt)	Upper (Gt)	Contribution (%)	
All lakes	718.97	636.70	928.77	100.00	3047.53	2967.48	3668.14	100.00	23.59
Reservoirs	297.14	252.60	365.22	41.33	1601.45	1557.25	1772.46	52.55	18.55
Natural lakes	668.80	599.52	932.67	93.02	1446.08	1379.34	1933.56	47.45	46.25
On-river	320.28	258.18	402.02	44.55	594.44	532.74	692.60	19.51	53.88
Floodplain	188.47	175.80	252.23	26.21	403.99	392.58	524.91	13.26	46.65
Permafrost-zone	76.38	73.60	189.95	10.62	132.97	131.25	327.58	4.36	57.44
Endorheic	54.17	31.43	82.54	7.53	150.43	127.88	188.59	4.94	36.01
Glacial	5.01	3.53	11.63	0.70	41.49	40.04	85.55	1.36	12.07
Other	32.43	30.40	43.71	4.51	122.76	120.98	159.51	4.03	26.42
River	265.86	180.29	351.44	36.98	---	---	---	---	---
TWS	5619.82	4961.14	6278.50	781.65	---	---	---	---	---

Methods

SWOT lake measurements

The SWOT Level-2 High-Rate Lake Single-Pass Vector Product (L2_HR_LakeSP; hereafter LakeSP) provides polygon-based measurements of lake water surface elevation (WSE), area, and quality attributes across two-50-km swaths per pass on a nominal 21-day repeat orbit[21]. We used the LakeSP_Prior sub-product, which links observations to the SWOT Prior Lake Database (PLD)[10], enabling consistent tracking of individual lakes across repeat passes[21,52]. We retrieved the full nominal-orbit record available in LakeSP Version C (PGC0 and PIC2) via the PO.DAAC Hydrocron API[53], spanning 21 July 2023 to 5 May 2025 (Cycles 1–32, with partial coverage of cycle 32). Version C uses PLD v1.06, which catalogs 5,898,420 prior lakes ≥1 ha covering 2.6 million km^2. Retrieved data yielded time series for 5,629,726 prior lakes, representing 98.5% of global lake area. We used Version C exclusively as newer versions still contained temporal gaps at the time of analysis.

SWOT WSE filtering

LakeSP WSE time series can contain spurious measurements arising from complexities in surface conditions, radar backscatter, and subsequent processing[52]. Because fixed quality criteria can either remove valid hydrological variability or retain residual errors across heterogeneous lakes[54-56], we filtered WSE using the Heuristic Adaptive Lake Filter (HALF)[57], which adapts filtering criteria to lake-specific characteristics. Briefly, HALF (1) identifies a conservative high-quality subset to establish lake-specific thresholds and a physically constrained baseline, (2) iteratively removes outliers relative to a low-pass-filtered baseline while preserving temporal coverage, and (3) corrects cross-pass WSE inconsistencies. Full methods are given in Trudel et al.[57].

For global implementation, we applied HALF with progressively relaxed time-coverage criteria[57] to retain robustly filtered time series for as many lakes as possible. The strict mode required ≥1 year of retained observations with no internal gap >3 months. Lakes failing these criteria were re-evaluated under a relaxed mode requiring >5 retained observations. Lakes failing both modes received only Tukey-fence outlier filtering.

Among prior lakes with LakeSP time series, 1,009,566 passed the strict mode and 4,182,280 passed the relaxed mode, representing 65.0% and 32.2% of global prior lake area, respectively. Among relaxed-mode lakes, 615,253 retained records spanning ≥1 year with no internal gap >4 months. We therefore combined this temporally cohesive subset of relaxed-mode lakes with all strict-mode lakes to define tier-1, comprising 1,624,819 lakes representing 75.1% of global prior lake area (Extended Data Fig. 1). Tier-1 lakes were used for analyses requiring coherent seasonal trajectories, including phenological timing, archetypes, and storage–Q coupling. To improve global

representativeness, we further defined tier-2 as other lakes with at least a 1-year span and at least five retained observations. This tier-2 subset added 3,456,074 lakes, representing 20.2% of global prior lake area. Together, lakes of both tiers represent 95.3%. Tier-2 lakes supplemented coverage-sensitive estimates of individual-lake storage variability and uncertainty estimates for global lake storage variability.

Because lake levels can fluctuate beneath seasonal ice[43] and LakeSP ice flags are climatology-based rather than observation-based, HALF does not categorically exclude ice-flagged observations. Instead, it applies ice-specific quality safeguards to selectively retain them. Filtered LakeSP time series were independently validated against gauge observations from 1,182 lakes worldwide, while ice-removal sensitivity tests assessed the trade-off between preserving seasonal signals and limiting retrieval errors (Supplementary Information section 1). Validation yields 68th-percentile mean absolute errors of 0.147 m and 0.200 m for ice-free and ice-affected lakes, respectively, and a median WSE–gauge correlation of 0.923 (Supplementary Tables S1–S2; Supplementary Figs. S1–S5). Ice-stratified uncertainties were further propagated into estimates of lake storage changes (see “Uncertainty propagation for lake storage changes”). The temporal coverage of filtered time series was also evaluated across global basins, revealing median observation gaps of 0.4 SWOT cycles for tier-1 lakes and 1.4 cycles for tier-2 lakes (Supplementary Information section 2; Supplementary Figs. S6–S9).

Geoid correction for large lakes

LakeSP WSE values are referenced to the Earth Gravitational Model 2008 (EGM08)[58]. EGM08 can exhibit multi-decimeter uncertainties in regions with strong local gravity variations, particularly in mountainous terrain and rift settings[59]. For large lakes sampled by multiple SWOT passes, these errors may produce pass-dependent WSE offsets because different passes observe different portions of the lake surface. Although HALF statistically harmonizes cross-pass offsets, we applied an additional physically based correction for selected large lakes using lake-specific mean water-surface topography.

To remove these biases in a physical way, we applied a geoid correction to the world’s largest lakes using the SWOT Level 2 Water Mask Raster Image 250m Data Product, (Version C)[60], referenced to the WGS84 ellipsoid, together with the mean lake surface topographies developed by Bergé-Nguyen et al[59]. These mean lake surfaces were derived from multi-mission nadir altimetry and served as locally calibrated vertical datums. For each lake and pass, we subtracted the mean lake surface from SWOT raster elevations and computed the mean corrected WSE over the validly observed lake area, defined as pixels with dark-water fraction <0.3 and WSE quality flag = 0.

We initially tested this correction for all 22 large lakes in Bergé-Nguyen et al.[59]. After comparison with HALF-filtered LakeSP time series, we retained the corrected WSE time

series for 15 lakes where this geoid correction outperformed the statistical cross-pass correction: Argentino, Baikal, Balkhash, General Carrera, Issykkul, Ladoga, Malawi, Namco, Nicaragua, Onega, Ontario, Tanganyika, Titicaca, Vänern, and Victoria.

Classifying lake typology

We classified prior lakes into reservoirs and six natural-lake typologies. Reservoirs, including regulated natural lakes, were identified by intersecting PLD polygons[10] with multiple reservoir inventories: the Global Reservoir and Dam database (GRanD v1.3)[61], the Global Dam Watch Database (GDW v1.0)[23], the Georeferenced global Dams And Reservoirs dataset (GeoDAR v1.1 and v2.0)[62], and the Global REservoir Inventory (GREI)[63]. Remaining lakes were classified as permafrost-zone, off-channel floodplain, on-river, glacial, endorheic, or other. Because these categories can overlap, labels were assigned sequentially in the order listed above, with later categories taking precedence.

Permafrost-zone lakes are defined as lakes occurring within continuous, discontinuous, sporadic, or isolated permafrost regions. These lakes are not necessarily thermokarst lakes but may be influenced by permafrost freeze-thaw dynamics and associated hydrogeomorphic and subsurface processes. We identified these lakes as PLD polygons with ≥50% of their area within the Circum-Arctic Map of Permafrost and Ground-Ice Conditions (version 2)[64].

Off-channel floodplain lakes are defined as lakes within floodplain extents and low valley-bottom topography that are not traversed by major river channels. This category is intended to capture backwaters, oxbows, and floodplain depressions with primarily lateral or episodic hydrological connectivity. These lakes were identified as PLD lakes with ≥50% of their area within the SHIFT floodplain dataset (version 3a)[65].

On-river lakes are defined as lakes traversed or intersected by major river channels, indicating direct embedding within primary drainage networks. We identified these lakes using the SWOT River Database (SWORD, version 16)[66], which maps global river centerlines wider than ~30 m. Lakes traversed by smaller rivers not represented in SWORD may therefore be excluded from this classification.

Glacial lakes were identified by intersection with the GIGLak dataset[67], which includes both meltwater-fed lakes and other lakes within glacierized settings. Endorheic lakes were defined as inland terminal lakes without surface outflow and identified using PLD-TopoCat v1.1[68].

Lakes not assigned to any of these typologies were defined as “other”. This category includes, for example, headwater lakes disconnected from major river channels, non-glacier alpine lakes, and other impoundments not inventoried as reservoirs in the above sources. Lake number and area statistics by typology are given in Extended Data Fig. 2.

Constructing lake storage anomalies

Filtered LakeSP time series were converted to storage-anomaly time series, $V(t)$, using a hierarchical hypsometric procedure. For each tier-1 lake, we prioritized the use of the endogenous area-WSE (A-E) relationship derived from LakeSP observations (SWOT-hypsometry method). When this relationship failed quality screening, we used an external A-E relationship from the 3D-LAKES dataset[51] if a reliable PLD-to-HydroLAKES match was available (external-hypsometry method). For the remaining tier-1 lakes and all tier-2 lakes, $V(t)$ was estimated using the PLD lake area (static-area method).

For the SWOT-hypsometry method, we constructed A-E relationships using LakeSP observations with complete area coverage (partial_f = 0) and low dark-water fraction (dark_frac < 0.5). To minimize water-classification and dark-water recovery errors, we accepted a SWOT-derived hypsometry only when the lake had ≥10 valid A-E pairs, a significant increasing A-E trend (Mann–Kendall test, α = 0.05), and a median observed area within a factor of three of the PLD area, allowing substantial area variations while excluding gross water-classification errors[10,69-71]. A robust linear A-E relationship was then fitted using Theil–Sen regression, and $V(t)$ was computed by integrating the fitted relationship from the first valid WSE observation E_0 to each retained WSE. This method was applied to 359,664 lakes, representing 11.1% of the total tier-1 lake area (Supplementary Fig. S10).

When reliable SWOT-derived hypsometry was unavailable, we considered 3D-LAKES, which provides ICESat-2-calibrated A-E relationships for ~0.5 million lakes globally[51]. Because 3D-LAKES uses HydroLAKES[9] to define lake geometries, it was applied only where matched HydroLAKES and PLD polygons had an intersection-over-union ratio ≥0.8. Lakes requiring extrapolation beyond the available 3D-LAKES A-E range were also excluded. To account for non-linear lake hypsometry, we estimated surface areas corresponding to LakeSP WSE observations by piecewise interpolation among available 3D-LAKES A-E pairs. Incremental volume change between adjacent elevations was calculated using the truncated-pyramid formulation[52]:

$$dV_i = \frac{E_i - E_{i-1}}{3}\left(A_{i-1} + A_i + \sqrt{A_{i-1}A_i}\right) \quad \text{eq (1)}$$

where E_i and A_i are elevation and area at level i, respectively; and dV_i is the volume change from level $i-1$ to i. Storage anomaly $V(t)$ at timestep t was derived by integrating dV_i through E_0 to E_t. This external-hypsometry method was applied to 15,694 lakes, representing 41.2% of tier-1 lake area. The remaining 1,249,461 tier-1 lakes, accounting for 47.6% of lake area, used the static-area method.

Although the hierarchy selected one preferred storage-construction method for each lake, all eligible methods were retained to quantify method-related uncertainty. This uncertainty was evaluated from two perspectives: (1) the effect on derived seasonal

storage variability σ_V, quantified by the interdecile range (IDR, P90–P10), and (2) the cross-method spreads in $dV(t)$ time series (Extended Data Fig. 9). Relative to the static-area σ_V, the SWOT-hypsometry method yielded a median relative bias of 0.343 and median absolute relative error of 0.364, whereas the external-hypsometry method yielded −0.031 and 0.090 (Extended Data Fig. 9a). For each lake, method-specific $V(t)$ series were recentered by their period means, and cross-method uncertainty was calculated as the root-mean-square spread among available methods at retained timesteps, normalized by the static-area σ_V. Relative cross-method uncertainty had a median of 0.081 and 68th-percentile of 0.137 (Extended Data Fig. 9b). Both perspectives indicate that method-related differences are generally modest relative to seasonal storage variability.

Uncertainty associated specifically with the static-area storage construction method was further evaluated using independent hypsometric curves for 171 lakes from the HydroWeb dataset[72] (Supplementary Information section 3; Supplementary Fig. S11).

Calculating seasonal lake storage variability

We quantified seasonal lake storage variability at two scales: individual lakes and the globally aggregated lake storage cycle. For individual lakes, the objective was to characterize typical annual-scale storage variability during the approximately two-year study period while reducing sensitivity to interannual variability. Such variability was evident during the study period, which spanned the transition from the strong 2023–2024 El Niño to ENSO-neutral conditions, followed by weak La Niña conditions in late 2024 and early 2025[24,73]. Because lakes also differed in the timing and continuity of retained observations, we used a lake-specific sliding-window approach rather than imposing a common annual window on all lakes.

For each lake, we evaluated all possible one-year windows, each comprising 18 consecutive 21-day SWOT cycles and shifted forward by one cycle at a time. A window was considered valid only when both its first and last cycles contained at least one retained observation, ensuring that annual variability was bracketed by observations without edge extrapolation. Within each valid window, storage-anomaly time series $V(t)$ were interpolated to a regular daily grid to reduce sensitivity to uneven temporal sampling, and seasonal variability was quantified as the IDR of the interpolated time series. The final individual-lake seasonal variability σ_V was defined as the median IDR across all valid sliding windows.

Unlike the individual-lake estimates, construction of the globally aggregated lake storage cycle required a single common annual window so that storage anomalies from different lakes could be summed consistently through time. We evaluated all candidate 18-cycle windows using three basin-level diagnostics: the proportion of lakes requiring edge extrapolation, mean edge-extrapolation length, and mean internal observation gap

(Supplementary Figs. S12–S14). Windows beginning around cycles 5–7 provided the greatest temporal overlap, with cross-basin medians of ~5% of lakes requiring edge extrapolation, a mean extrapolation length of ~0.2 cycles, and internal gaps of ~1.5 cycles. We therefore selected cycles 5–22 (12 October 2023 to 21 October 2024) as the common or reference annual window, which retained these favorable coverage characteristics while starting early enough to reduce ice-related observation loss near the window boundaries (Supplementary Fig. S8). Daily storage anomalies for each lake were re-referenced to their mean over cycles 5–22, and global or class-specific time series, $\sum V(t)$, were obtained by summing across constituent lakes. Seasonal variability of each aggregated series, $\sigma_{\sum V}$, was again quantified using IDR.

Although IDR was used as the primary variability metric, we also evaluated sensitivity of variability to alternative measures, including temporal standard deviation and full amplitude (Supplementary Fig. S15). We first compared these metrics for individual-lake WSE variability σ_H using independent gauge observations, separately for tier-1 lakes and the combined tier-1 + tier-2 population (Supplementary Fig. S15a,b). Standard deviation scaled at 0.37 times IDR, whereas amplitude scaled at 1.24 times IDR, with nearly identical relationships between the two population configurations. Amplitude showed larger scatter, consistent with its greater sensitivity to extreme values.

We then repeated the comparison using SWOT-derived σ_H for individual lakes globally (Supplementary Fig. S15c,d). The resulting relationships were similarly consistent between tier-1 and tier-1 + tier-2 lakes, with standard deviation scaling at 0.39 times IDR and amplitude at 1.31–1.32 times IDR. The close agreement with the gauge-derived relationships supports the fidelity of our LakeSP filtering and variability estimates, while the greater dispersion in the global amplitude relationship further indicates its sensitivity to residual extremes.

Finally, we applied the same comparison to SWOT-derived σ_V globally (Supplementary Fig. S15e,f), for which standard deviation scaled at 0.38 times IDR and amplitude at 1.55–1.57 times IDR, again with similar relationships for the two tier configurations. The scaling factors for σ_V are not expected to be directly comparable to those for σ_H, as storage variability additionally reflects lake area, hypsometry, and storage-construction method. Together, these results support IDR as a robust primary measure while providing empirical scaling relationships for comparison with other studies using standard deviation or amplitude.

Uncertainty propagation for lake storage changes

Uncertainty in each lake's storage-anomaly time series, $V(t)$, comprises two components: (1) within-method uncertainty, arising from WSE observational errors propagated through the selected storage-construction method; and (2) cross-method

uncertainty, representing differences in storage estimates among alternative construction methods. For the within-method component, we combined three WSE-error terms in quadrature: (1) LakeSP-reported WSE uncertainty, wse_u, which accounts for random and systematic errors primarily associated with interferogram quality and processing; (2) within-lake WSE dispersion, wse_std, also reported in LakeSP; and (3) an empirical WSE error term characterized from HALF validation and stratified by ice condition and lake tier (Supplementary Information section 1). The combined WSE uncertainty was propagated through the selected storage-construction method at each timestep to obtain the within-method $V(t)$ uncertainty time series.

For lakes with two or more eligible storage-construction methods, the cross-method uncertainty time series was estimated from the lake-specific spread among method-specific $V(t)$ at each retained timestep. For lakes represented only by the static-area method, we adopted the 68th-percentile relative cross-method uncertainty of 0.137 derived from the multi-method comparison (Extended Data Fig. 9b). Within-method and cross-method uncertainty time series were combined in quadrature to obtain the total $V(t)$ uncertainty time series for each lake.

Uncertainty in each lake's seasonal storage variability, σ_V, was propagated from the $V(t)$ uncertainty time series together with an additional error term accounting for daily interpolation. Expressed as a fraction of σ_V, this interpolation uncertainty was estimated from gauge validation as the increase in normalized variability error introduced by interpolating retained observations to a daily time series. The resulting interpolation terms were $0.029\sigma_V$ for tier-1 lakes and $0.140\sigma_V$ for tier-2 lakes.

For each aggregated storage-anomaly time series, $\sum V(t)$, the $V(t)$ uncertainty time series of constituent lakes were combined at each timestep assuming error independence, to yield a corresponding uncertainty time series for $\sum V(t)$. Then, for the seasonal variability of each aggregated series, $\sigma_{\sum V}$, we calculated a symmetric uncertainty as the root-sum-of-squares of two scalar terms: the root-mean-square of the $\sum V(t)$ uncertainty time series over cycles 5–22, and the interpolation uncertainty, $0.029\sigma_{\sum V}$, derived from tier-1 gauge validation.

Tier-2 lakes, which were not included directly in $\sum V(t)$ time series, were treated as a one-sided coverage uncertainty. To account for phase cancellation among lakes, their summed individual-lake variability, $\sum \sigma_V$, was scaled by the tier-1 coherence ratio, defined as $\sigma_{\sum V} / \sum \sigma_V$ for tier-1 lakes. This scaled tier-2 contribution was added to the upper uncertainty bound of $\sigma_{\sum V}$, whereas the lower bound included only the symmetric uncertainty described above. For tier-1 $\sum \sigma_V$ used for Fig. 2, the lower uncertainty bound was calculated as the root-sum-of-squares of tier-1 per-lake σ_V uncertainties, whereas the upper bound additionally included tier-2 $\sum \sigma_V$.

Timing of lake seasonal extrema

We quantified the timing of seasonal lake storage extrema for each tier-1 lake using a sliding-window consensus approach. To reduce sensitivity to individual outliers and uneven within-cycle sampling, each lake's storage-anomaly time series was first summarized at the SWOT-cycle scale using the median value within each 21-day cycle. Within any annual window, the extrema months were then defined by the calendar months containing the maximum and minimum cycle-median storage anomalies.

Consistent with the sliding-window framework used to estimate individual-lake σ_V, we evaluated all valid one-year windows per lake and recorded the resulting extrema months per window. Final maximum and minimum months were defined as the unique modal months across all valid windows. When no unique mode existed, the corresponding extrema months from the reference annual window, cycles 5–22, were used as fallback estimates.

We evaluated the robustness of the consensus timing using two diagnostics. Agreement was defined as the fraction of valid windows whose maximum or minimum month matched the final consensus month, whereas circular spread was quantified as the circular standard deviation of window-specific extrema months. Maximum and minimum timing could be estimated for 99.8% and 99.8% of tier-1 lakes, respectively, with median agreements of 0.93 and 0.77. Fallback assignment was uncommon, occurring for 2.9% of maximum-month estimates and 5.0% of minimum-month estimates, indicating that most lakes exhibited dominant seasonal extrema timing across annual windows.

Seasonal-extrema timing was further independently validated against gauge-paired tier-1 lakes using the same sliding-window procedure, with additional evaluation of minimum-month robustness under snow- and ice-affected conditions (Supplementary Information section 4; Supplementary Figs. S16–S19).

Identifying phenological archetypes

We inductively classified tier-1 lakes into phenological archetypes using unsupervised K-means clustering of standardized seasonal storage trajectories. The clustering was designed to identify recurrent patterns in seasonal phase and trajectory shape, which we term phenological archetypes. Because trajectories were standardized, the classification emphasized these temporal characteristics rather than differences in absolute storage variability. Consistent with the sliding-window approach used for seasonal extrema timing, we repeated the clustering across all 15 one-year windows from cycles 1–18 through 15–32 and used a sliding-window consensus procedure to determine each lake's final archetype from aligned window-specific classifications.

For each window, we retained lakes whose valid observation span covered the full 18 cycles, interpolated missing interior cycle values, and standardized each cycle-scale

storage-anomaly trajectory using a within-window z-score. We then applied K-means clustering to the standardized time series of all eligible lakes worldwide and evaluated candidate cluster numbers from K = 2 to 20 using the inertia elbow. The elbow represents the parsimonious cluster number beyond which additional clusters yield diminishing reductions in the within-cluster sum of squared distances (i.e., K-means inertia), and we identified it using the maximum distance-to-chord criterion. K = 7 was identified as the elbow in 11 of the 15 windows and was thus adopted as the common cluster number for all windows.

Because K-means labels are arbitrary across independent runs, we aligned the window-specific labels to those from the reference annual window, cycles 5–22. For each window w, lakes classified for both w and the reference window were used to construct a K×K contingency matrix $C^{(w)}$, where $C_{a,r}^{(w)}$ denotes the number of lakes assigned to both cluster a in window w and cluster r in the reference window.

To account for differences in cluster size, we converted $C_{a,r}^{(w)}$ to a Dice overlap score[74]:

$$S_{a,r}^{(w)} = \frac{2C_{a,r}^{(w)}}{\sum_{r'} C_{a,r'}^{(w)} + \sum_{a'} C_{a',r}^{(w)}} \quad \text{eq (2)}$$

where $\sum_{r'} C_{a,r'}^{(w)}$ denotes the total number of lakes assigned to cluster a in window w, and $\sum_{a'} C_{a',r}^{(w)}$ is the total number of lakes assigned to cluster r in the reference window. Large $S_{a,r}^{(w)}$ indicates greater membership overlap between a and r relative to their combined sizes. We then used the Hungarian assignment algorithm[75] to identify the optimal one-to-one mapping, $\pi(a)$, between the window-specific and reference clusters, by maximizing the total overlap score, $\sum_a S_{a,\pi(a)}^{(w)}$. This mapping aligned cluster labels across windows without altering individual lake assignments.

We further evaluated alignment quality using a cluster purity index,

$$p^{(w)}(a) = \frac{\max_r C_{a,r}^{(w)}}{\sum_{r'} C_{a,r'}^{(w)}} \quad \text{eq (3)}$$

where $\max_r$ selects the reference cluster containing the largest number of lakes from cluster a; this way, $p^{(w)}(a)$ gives the fraction (purity) of lakes in cluster a belonging to its dominant reference cluster. Window-level alignment quality was then summarized as the median purity across clusters. We only retained windows containing >80% of tier-1 lakes and having median purity >0.5 in the final consensus analysis.

A final archetype label was assigned to each lake as the modal aligned label across retained windows in which that lake was present. When no unique mode existed, the

reference-window label was used as a fallback. Assignment stability was quantified as the fraction of retained-window labels matching the final archetype label.

The consensus analysis retained six windows spanning cycles 3–25, each containing 87.8–96.7% of tier-1 lakes and having median purity values of 0.54–1.00. Median assignment stability ranged from 0.8 to 1.0 across archetypes, and only 5.8% of lakes classified in the reference window changed archetype after consensus voting. Overall, valid archetype labels were assigned to 1,598,109 lakes (98.4% of tier-1 lakes), indicating that the consensus primarily stabilized phenological class assignment during the study period rather than imposing a substantially different classification.

We further assessed the reliability of the identified phenological archetypes by evaluating how well SWOT-derived lake WSE time series preserved gauge-observed seasonal trajectories and how trajectory fidelity varied across observing environments (ice exposure and latitude), lake types, and phenological archetypes (Supplementary Information section 5; Supplementary Figs. S20–S21).

Simulating naturalized upstream forcing

To characterize the hydroclimatic forcing associated with each tier-1 lake, we simulated naturalized discharge (Q) at lake outlets using the RAPID (Routing Application for Parallel computatIon of Discharge) vector-based river routing model[76] implemented on the Harmonized Prior river-lake (HarP) network[68,77]. HarP represents global river–lake connectivity and links upstream drainage areas to lake outlet nodes. Lakes were treated as non-storing routing elements, with no lake-storage or reservoir-operation schemes activated, so discharge was routed through lake outlets as if lakes behaved as river channels. The resulting Q therefore does not represent lake-mediated outflow, but rather naturalized upstream runoff routed to the lake outlet, including runoff generated over the lake surface. To avoid embedding regulated signals, RAPID was not calibrated against gauge observations and was driven entirely by naturalized hydrological forcings.

Hydrological forcings were obtained from the GLDAS Noah Land Surface Model L4, Version 2.1, 3-hourly 0.25° product[78,79]. We performed two routing scenarios: (1) a baseline simulation using total runoff, including surface runoff, subsurface runoff, and snowmelt, and (2) a no-snow scenario to isolate snowmelt contributions. In the no-snow scenario, runoff R was scaled by the fraction of total water input attributable to rainfall:

$$R_{ns} = \begin{cases} 0, & P = 0, S > 0 \\ R \times \frac{P+\varepsilon}{P+S+\varepsilon}, & \text{otherwise} \end{cases} \quad \text{eq (4)}$$

where R_{ns} is the adjusted runoff used in the no-snow simulation, P is rainfall, S is snowmelt, and ε is a small constant (10^{-6}) to prevent division by zero.

Simulations covered January 2020 through June 2025, with the first two years as model spin-up. RAPID produced discharge at 3-hourly resolution, which was averaged over each 21-day SWOT orbit cycle. This aggregation produced cycle-mean discharge for both routing scenarios, reducing sensitivity to uncertainty in simulated flow travel times.

Lake storage-flow coupling

We quantified coupling between the storage dynamics of tier-1 lakes and their upstream hydroclimatic forcing using two complementary metrics: Contemporaneous Spearman correlation (ρ) to measure how closely seasonal storage variations track naturalized Q without assuming linearity, and phase offset (τ) to characterize temporal displacement of the hydroclimatic signal through lake storage. Both metrics were evaluated at the 21-day SWOT-cycle timescale to reduce sensitivity to short-term routing uncertainties. Consistent with preceding analyses, both metrics were estimated using a one-year (18-cycle) sliding-window approach to emphasize seasonal coupling while limiting the influence of interannual climate variability.

For ρ, an annual window was only retained when it fell within the lake's observation period and contained at least five paired storage–Q observations. ρ was calculated for all valid windows between cycles 1 and 32, and the median across windows was used as the representative coupling for each lake. This yielded valid estimates for 95.64% of tier-1 lakes.

We quantified τ by anchoring each 18-cycle storage window and shifting Q by integer offsets from −8 to +8 SWOT cycles (±5.5 months), with positive τ indicating storage lagging Q. Each shifted comparison required structural support for the full 18-cycle Q window and at least five valid storage–Q pairs. For each candidate τ, lagged and contemporaneous correlations were calculated over the same valid annual windows, and the difference between their median correlations defined the improvement $\Delta\rho$. A non-zero τ was accepted only when the lag-adjusted correlation exceeded 0.3 and $\Delta\rho \geq 0.15$. Among accepted lags, the one maximizing $\Delta\rho$ was selected; for near ties (within 0.03 of the maximum $\Delta\rho$), the smaller absolute lag was preferred. If no non-zero lag met these criteria, lakes with contemporaneous $\rho > 0.3$ were assigned $\tau = 0$, whereas those remaining below this threshold were left unresolved. After phase adjustment, approximately 5.0% of lakes by count, representing ~2.4% of summed storage variability, remained unresolved.

We validated estimated ρ and τ by repeating the coupling analysis with gauge WSE substituted for SWOT-derived trajectories. We also examined the environmental dependence and potential origins of storage-leading (negative) offsets. Sensitivity tests further showed that uncertainty in negative τ does not materially alter the size dependence of τ or the systematically longer reservoir delays. These analyses are detailed in Supplementary Information section 6 and Supplementary Figs. S22–S28.

Use of large language models

Open AI's ChatGPT was used to assist with code development and debugging, discussions of methodological approaches and results, and manuscript refinement. All AI-assisted outputs were critically reviewed and verified by the authors, who retain full responsibility for the study's conceptual framework, scientific originality, methodological and analytical decisions, accuracy, interpretations, and conclusions. AI use was assessed in accordance with Nature's risk-assessment framework (https://www.nature.com/nature/editorial-policies/ai).

Data availability

The Global Lake Storage Phenology dataset (v1.0) produced in this study, including individual-lake attributes such as seasonal water surface elevation (WSE) and storage variability, contextual typologies, and phenological archetype assignments, as well as daily time series of aggregated storage anomalies with propagated uncertainties, is publicly available via Zenodo https://doi.org/10.5281/zenodo.23718039. The input SWOT LakeSP data products (version C) are openly available via NASA PO.DAAC and CNES HydroWeb.next https://doi.org/10.5067/SWOT-LAKESP-2.0. WSE time series from more than 1,200 lake gauges worldwide used for validation are available via Zenodo https://doi.org/10.5281/zenodo.20781336. Time series and hypsometric information for 171 lakes used to evaluate the impact of static lake area on storage-anomaly estimation are available via HydroWeb https://www.theia-land.fr. The SWOT PLD (v1.06) is available through HydroWeb.next https://hydroweb.next.theia-land.fr under the Etalab 2.0 licence. The 3D-LAKES dataset (v2) is available via Zenodo https://doi.org/10.5281/zenodo.13107867. HydroBASINS is available via the HydroSHEDS website https://www.hydrosheds.org/products/hydrobasins. Reservoir inventories are available from the following sources: GRanD (v1.3) and GDW (v1.0) via the Global Dam Watch directory https://www.globaldamwatch.org/directory; GeoDAR (v1.1) via Zenodo https://doi.org/10.5281/zenodo.6163413; and GREI (v1) via Figshare https://doi.org/10.6084/m9.figshare.31067800. SWORD (v16) is available via Zenodo https://doi.org/10.5281/zenodo.10013982, the Circum-Arctic Map of Permafrost and Ground-Ice Conditions (v2) via the National Snow and Ice Data Center https://doi.org/10.7265/skbg-kf16, the SHIFT floodplain dataset (v3a) via Zenodo https://doi.org/10.5281/zenodo.17757266, and GIGLak via Figshare https://doi.org/10.6084/m9.figshare.26310967. The GLDAS Noah Land Surface Model L4 (v2.1), 3-hourly 0.25° product, is available via NASA Earthdata https://doi.org/10.5067/E7TYRXPJKWOQ. The PLD-TopoCat network (v1.1) used for discharge routing is available via Zenodo https://doi.org/10.5281/zenodo.14202301.

Code availability

The Heuristic Adaptive Lake Filter (HALF, v1.0), used to quality-filter SWOT LakeSP observations, is available via GitHub https://github.com/LARIC-jw/HALF/tree/v1.0 and Zenodo https://doi.org/10.5281/zenodo.20782248. The PO.DAAC Hydrocron API used to retrieve LakeSP observations is available via GitHub https://github.com/podaac/hydrocron and Zenodo https://doi.org/10.5281/zenodo.23175535. The Routing Application for Parallel computation of Discharge (RAPID, v1.6.0), used to simulate naturalized upstream discharge, is available via Zenodo https://doi.org/10.5281/zenodo.2560556.

Funding

J.W., M.S.S., X.H., G.H.A., H.G., C.J.G., L.C.M., S.C. and T.M.P. were supported by the National Aeronautics and Space Administration (NASA) SWOT Science Team grants (80NSSC24K1663, 80NSSC20K1143, 80NSSC24K1646, 80NSSC24K1661, 80NSSC25K7230, and 80NSSC25K7715). J.W. and M.S.S. were additionally supported by a Taylor Geospatial Institute (TGI) start-up award to the University of Illinois. M.T. was supported by the Canadian Space Agency (CSA) grant 21SUESAMSW. E.S.L. was supported by NASA's Future Investigators in NASA Earth and Space Science and Technology (FINESST) program (80NSSC23K1556). S.B., J.-F.C. and H.O. were supported by the Centre National d'Études Spatiales (CNES) TOSCA SWOT Science Team projects (SNORKS2, HydroWeb, and PIHYDRO). E.F.-C. was supported by the COMPASS-FME project of the U.S. Department of Energy's Office of Science, Biological and Environmental Research program. C.H.D. was supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA (80NM0018D0004).

Author contributions

Conceptualization: J.W. G.H.A., C.B., H.G., C.J.G., L.C.S., and T.M.P. Methodology: J.W., M.S.S., M.T., S.B., J.-F.C., X.H., C.S., G.H.A., E.L., S.C., E.F.-C., and C.D. Formal analysis and investigation: J.W., M.S.S., M.T., S.B., J.-F.C., X.H., and C.S. Data curation: J.W., M.S.S., M.T., S.B., J.-F.C., R.F., C.P., C.S., H.G., and M.D. Writing – original draft: J.W., M.S.S., and S.B. Writing – review and editing: J.W., S.B., J.-F.C., X.H., R.F., C.S., G.H.A., C.B., H.G., C.J.G., L.C.S., E.L., S.C., M.D., E.F.-C., Y.W., C.D., H.O., and T.M.P. Visualization: J.W. Project administration and financing acquisition: J.W., G.H.A., H.G., H.O., and T.M.P.

Competing interests

The authors declare no competing interests.

References

1 Downing, J. A. *et al*. The global abundance and size distribution of lakes, ponds, and impoundments. *Limnol. Oceanogr.* **51**, 2388-2397 (2006).

2 Schindler, D. W. Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes. *Limnol. Oceanogr.* **54**, 2349-2358 (2009).

3 Woolway, R. I. *et al*. Extreme and compound events in lakes. *Nat. Rev. Earth Environ.* **6**, 593-611 (2025).

4 Li, L., Long, D., Wang, Y. & Woolway, R. I. Global dominance of seasonality in shaping lake-surface-extent dynamics. *Nature* **642**, 361-368 (2025).

5 Cooley, S. W., Ryan, J. C. & Smith, L. C. Human alteration of global surface water storage variability. *Nature* **591**, 78-81 (2021).

6 Chen, T. *et al*. Estimating seasonal water budgets in global lakes by using multi-source remote sensing measurements. *J. Hydrol.* **593**, 125781 (2021).

7 Fu, L. L. *et al*. The Surface Water and Ocean Topography mission: A breakthrough in radar remote sensing of the ocean and land surface water. *Geophys. Res. Lett.* **51**, e2023GL107652 (2024).

8 Cerbelaud, A. *et al*. Wide-swath altimetry maps bank shapes and storage changes in global rivers. *Nature* **651**, 666-671 (2026).

9 Messager, M. L., Lehner, B., Grill, G., Nedeva, I. & Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **7**, 13603 (2016).

10 Wang, J. *et al*. The Surface Water and Ocean Topography Mission (SWOT) Prior Lake Database (PLD): Lake mask and operational auxiliaries. *Water Resour. Res.* **61**, e2023WR036896 (2025).

11 Oki, T. & Kanae, S. Global hydrological cycles and world water resources. *Science* **313**, 1068-1072 (2006).

12 Leader, S., Kettridge, N., Hannah, D. M., Mendoza, C. & Devito, K. J. Multi-year climate memory in shallow lake water levels. *J. Hydrol.* **642**, 131764 (2024).

13 Döll, P., Kaspar, F. & Lehner, B. A global hydrological model for deriving water availability indicators: model tuning and validation. *J. Hydrol.* **270**, 105-134 (2003).

14 Adrian, R. *et al*. Lakes as sentinels of climate change. *Limnol. Oceanogr.* **54**, 2283-2297 (2009).

15 Li, Y., Zhao, G., Allen, G. H. & Gao, H. L. Diminishing storage returns of reservoir construction. *Nat. Commun.* **14**, 3203 (2023).

16 Wang, J., Sheng, Y. & Wada, Y. Little impact of the Three Gorges Dam on recent decadal lake decline across China's Yangtze Plain. *Water Resour. Res.* **53**, 3854-3877 (2017).

17 Crétaux, J.-F. *et al*. Lake volume monitoring from space. *Surv. Geophys.* **37**, 269-305 (2016).

18 Yao, F. *et al*. Satellites reveal widespread decline in global lake water storage. *Science* **380**, 743-749 (2023).

19 Chen, T. *et al*. Satellites reveal contrasting seasonal buffering patterns of global lakes and reservoirs. *Sci. Bull.* **71**, 3753-3764 (2026).

20 Xu, N., Lu, H., Li, W. Y. & Gong, P. Natural lakes dominate global water storage variability. *Sci. Bull.* **69**, 1016-1019 (2024).
21 CNES. Level 2 KaRIn high rate lake single pass vector product (short name: L2_HR_LakeSP), Revision A. SWOT Project Product Description Document, SWOT-TN-CDM-0673-CNES (2022).
22 Wang, J. *et al.* Recent global decline in endorheic basin water storages. *Nat. Geosci.* **11**, 926-932 (2018).
23 Lehner, B. *et al.* The Global Dam Watch database of river barrier and reservoir information for large-scale applications. *Sci. Data* **11**, 1069 (2024).
24 Fleischmann, A. S. *et al.* Extreme warming of Amazon waters in a changing climate. *Science* **390**, 606-611 (2025).
25 Bonnema, M., David, C. H., Frasson, R. P. D., Oaida, C. & Yun, S. H. The global surface area variations of lakes and reservoirs as seen from satellite remote sensing. *Geophys. Res. Lett.* **49**, e2022GL098987 (2022).
26 Pi, X. H. *et al.* Mapping global lake dynamics reveals the emerging roles of small lakes. *Nat. Commun.* **13**, 5777 (2022).
27 Ding, M. An improved geography of surface water abundance in lakes and reservoirs. Master of Arts thesis, Kansas State University, http://hdl.handle.net/2097/40075 (2019).
28 Dai, A. & Trenberth, K. E. Estimates of freshwater discharge from continents: Latitudinal and seasonal variations. *J. Hydrometeorol.* **3**, 660-687 (2002).
29 Riggs, R. M. *et al.* Characterizing operational signatures of reservoirs with the SWOT satellite by comparing natural lake and reservoir dynamics. *Environ. Res. Lett.* **21**, 044008 (2026).
30 Ryan, J. C., Smith, L. C., Cooley, S. W., Pitcher, L. H. & Pavelsky, T. M. Global characterization of inland water reservoirs using ICESat-2 altimetry and climate reanalysis. *Geophys. Res. Lett.* **47**, e2020GL088543 (2020).
31 Spence, C. *et al.* Process synchrony a key control of resilience in a subarctic freshwater system. *Earth's Future* **13**, e2024EF005518 (2025).
32 Seager, R. *et al.* Mechanisms of winter precipitation variability in the European-Mediterranean region associated with the North Atlantic Oscillation. *J. Clim.* **33**, 7179-7196 (2020).
33 Biasutti, M. Rainfall trends in the African Sahel: Characteristics, processes, and causes. *WIREs Clim. Change* **10**, e591 (2019).
34 Maldonado, T., Alfaro, E., Rutgersson, A. & Amador, J. A. The early rainy season in Central America: the role of the tropical North Atlantic SSTs. *Int. J. Climatol.* **37**, 3731-3742 (2017).
35 Richey, J. E. *et al.* Sources and routing of the Amazon River flood wave. *Glob. Biogeochem. Cycles* **3**, 191-204 (1989).
36 Saidaliyeva, Z. *et al.* Precipitation in the mountains of Central Asia: isotopic composition and source regions. *Atmos. Chem. Phys.* **24**, 12203-12224 (2024).
37 Chiang, J. C. H., Kong, W., Wu, C. H. & Battisti, D. S. Origins of East Asian summer monsoon seasonality. *J. Clim.* **33**, 7945-7965 (2020).

38 Alsafi, N. E. *et al*. Icings as sentinels and modifiers of water flow through winter landscapes: An exploration of physico-chemical processes on the lake-dominated, discontinuous permafrost Taiga Shield. *Hydrol. Process.* **38**, e15251 (2024).
39 Luo, S. X., Song, C. Q., Zhan, P. F., Li, J. Y. & Peng, L. A first look at the snow/ice penetration effect of SWOT observations on water level of global glacial lakes. *Geophys. Res. Lett.* **53**, e2026GL122624 (2026).
40 van der Kamp, G., Keir, D. & Evans, M. S. Long-term water level changes in closed-basin lakes of the Canadian Prairies. *Can. Water Resour. J.* **33**, 23-38 (2008).
41 Mielko, C. & Woo, M. K. Snowmelt runoff processes in a headwater lake and its catchment, subarctic Canadian Shield. *Hydrol. Process.* **20**, 987-1000 (2006).
42 Quinn, F. H. Hydraulic residence times for the Laurentian Great Lakes. *J. Great Lakes Res.* **18**, 22-28 (1992).
43 Elchyshyn, L. *et al*. Quantifying the effects of hydrological changes on long-term water quality trends in temperate reservoirs: Insights from a multi-scale, paleolimnological study. *J. Paleolimnol.* **60**, 361-379 (2018).
44 Cirelli, A. F. & Miretzky, P. Ionic relations: a tool for studying hydrogeochemical processes in Pampean shallow lakes (Buenos Aires, Argentina). *Quat. Int.* **114**, 113-121 (2004).
45 Zhou, X. *et al*. Satellite altimetry reveals heterogeneity in global flood peak timing and its shifts. *Commun. Earth Environ*. https://doi.org/10.1038/s43247-026-04115-5 (2026).
46 Götte, J., Astagneau, P. C. & Brunner, M. I. The influence of lakes and reservoirs on estimated flood peaks at hourly vs. daily timescale in the Alps. *Hydrol. Earth Syst. Sci.* **30**, 6235-6248 (2026).
47 Frieler, K. *et al*. Scenario setup and forcing data for impact model evaluation and impact attribution within the third round of the Inter-Sectoral Model Intercomparison Project (ISIMIP3a). *Geosci. Model Dev.* **17,** 1-51 (2024).
48 Gleason, C. J. *et al*. SWOT, empiricism, and river modeling. *Geophys. Res. Lett.* **53**, e2026GL124323 (2026).
49 Ma, X., Wang, J., Gharari, S., Mizukami, N. & Lettenmaier, D. P. Monitoring reservoir storage using SWOT satellite observations and a reservoir operation model. *Water Resour. Res.* **62**, e2025WR041223 (2026).
50 Tranvik, L. J. *et al*. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceanogr.* **54**, 2298-2314 (2009).
51 Huang, C. H. *et al*. 3D-LAKES: Three-dimensional global lake and reservoir bathymetry from ICESat-2 altimetry and landsat imagery. *Sci. Data* **12**, 1625 (2025).
52 CNES. Level 2 KaRIn high rate lake single pass science algorithm software: Level 2 processing (short name: SAS_L2_HR_LakeSP: Level 2 processing), Initial Release. SWOT Project Algorithm Theoretical Basis Document, SWOT-NT-CDM-1753-CNES (2023).
53 Wood, J. *et al*. PODAAC/Hydrocron: 1.11.0a4 (Versions 1.11.0a4) [Computer software]. *Zenodo*, https://doi.org/10.5281/zenodo.23175535 (2026)

54 Girard, F. *et al*. Performance of the Surface Water and Ocean Topography (SWOT) mission for monitoring small lakes in West Africa. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **18**, 14855-14869 (2025).
55 Maubant, L., Dodd, L. & Tregoning, P. Assessing the accuracy of SWOT measurements of water bodies in Australia. *Geophys. Res. Lett.* **52**, e2024GL114084 (2025).
56 Yu, L. P., Zhang, H. W., Gong, W. & Ma, X. Validation of mainland water level elevation products from SWOT satellite. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **17**, 13494-13505 (2024).
57 Trudel, M. *et al*. A heuristic adaptive filter for SWOT lake vector data products. *ESS Open Archive [preprint]*, https://doi.org/10.22541/essoar.15007657/v1 (2026).
58 Pavlis, N. K., Holmes, S. A., Kenyon, S. C. & Factor, J. K. The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). *J. Geophys. Res. Solid Earth* **117**, B04406 (2012).
59 Bergé-Nguyen, M. *et al*. Mapping mean lake surface from satellite altimetry and GPS kinematic surveys. *Adv. Space Res.* **67**, 985-1001 (2021).
60 JPL. Level 2 KaRIn high rate raster product (short name: L2_HR_Raster), Revision C. SWOT Project Product Description Document, JPL D-56416 (2025).
61 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).
62 Wang, J. *et al*. GeoDAR: georeferenced global dams and reservoirs dataset for bridging attributes and geolocations. *Earth Syst. Sci. Data* **14**, 1869-1899 (2022).
63 Liu, K. *et al*. Global patterns of reservoir sedimentation and overlooked risks in small reservoirs. *Nat. Sustain.* **9**, 1325-1337 (2026).
64 Brown, J., Ferrians, O., Heginbottom, J. A. & Melnikov, E. S. Circum-Arctic map of permafrost and ground-ice conditions. *U.S. Geological Survey Circum-Pacific Map* **45**, https://doi.org/10.3133/cp45 (1997).
65 Zheng, K. H., Lin, P. R. & Yin, Z. Y. SHIFT: a spatial-heterogeneity improvement in DEM-based mapping of global geomorphic floodplains. *Earth Syst. Sci. Data* **16**, 3873-3891 (2024).
66 Altenau, E. H. *et al*. The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A global river network for satellite data products. *Water Resour. Res.* **57**, e2021WR030054 (2021).
67 Song, C. Q., Fan, C. Y., Ma, J. S., Zhan, P. F. & Deng, X. Y. A spatially constrained remote sensing-based inventory of glacial lakes worldwide. *Sci. Data* **12**, 464 (2025).
68 Sikder, M. S. *et al*. Lake-TopoCat: A global lake drainage topology and catchment database. *Earth Syst. Sci. Data* **15**, 3483-3511 (2023).
69 Sheng, Y. W. *et al*. Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. *Remote Sens. Environ.* **185**, 129-141 (2016).
70 Wang, J. D., Sheng, Y. W. & Tong, T. S. D. Monitoring decadal lake dynamics across the Yangtze Basin downstream of Three Gorges Dam. *Remote Sens. Environ.* **152**, 251-269 (2014).

71 Pekel, J. F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418-422 (2016).
72 Crétaux, J.-F. *et al*. SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. *Adv. Space Res.* **47**, 1497-1507 (2011).
73 Peng, Q. H. *et al*. Strong 2023-2024 El Niño generated by ocean dynamics. *Nat. Geosci.* **18**, 471-478 (2025).
74 Dice, L. R. Measures of the amount of ecologic association between species. *Ecology* **26**, 297-302 (1945).
75 Kuhn, H. W. The Hungarian method for the assignment problem. *Nav. Res. Logist.* **52**, 7-21 (2005).
76 David, C. H., *et al*. River network routing on the NHDPlus dataset. *J. Hydrometeorol.* **12**, 913-934 (2011).
77 Sikder, M. S. *et al*. HarP: Harmonized Prior river-lake database (Version SWOT-HarP v1.1) [Dataset]. *Zenodo*, https://doi.org/10.5281/zenodo.14205131 (2024).
78 Rodell, M. *et al*. The global land data assimilation system. *Bull. Am. Meteorol. Soc.* **85**, 381-394 (2004).
79 Beaudoing, H. & Rodell, M. GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree, Version 2.1 [Dataset]. *Hydrological Sciences Laboratory (HSL), NASA Goddard Earth Sciences Data and Information Services Center*, https://doi.org/10.5067/E7TYRXPJKWOQ (2020).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExtendedDataFigure1.jpg
- ExtendedDataFigure5.jpg
- ExtendedDataFigure7.jpg
- ExtendedDataFigure3.jpg
- ExtendedDataFigure8.jpg
- ExtendedDataFigure9.png
- ExtendedDataFigure4.jpg
- ExtendedDataFigure6.jpg
- ExtendedDataFigure2.jpg
- WangetalgloballakephenologySI.docx