Supplementary for

Using satellite observations to validate and improve reservoir storage simulations in global hydrological models

Emmanuel Okiria[1], Naota Hanasaki[1,2], Simon N. Gosling[3], Emmanuel Nyenah[4,5], Peter Burek[6], Yusuke Satoh[7], Sebastian Ostberg[8], Kedar Otta[1], Luca Guillaumot[6,9]

[1]National Institute for Environmental Studies (NIES), Tsukuba, 305-8506, Japan
[2]School of Engineering, The University of Tokyo, Tokyo, 113-8656, Japan
[3]School of Geography, University of Nottingham, Nottingham, NG7 2RD, United Kingdom
[4]Institute of Physical Geography, Goethe University, Frankfurt, 60438 Frankfurt am Main, Germany
[5]Senckenberg Leibniz Biodiversity and Climate Research Centre (SBiK-F), Frankfurt, 60325 Frankfurt am Main, Germany
[6]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[7]Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama, 237-0061, Japan
[8]Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Potsdam, Germany
[9]BRGM, F-45060 Orléans, France

Correspondence to: Emmanuel Okiria (okiria.emmanuel@nies.go.jp)

S1. Derivation of the Global Dam Storage (GDS) dataset (bathymetry approximation and storage retrieval)

To anchor the bathymetric curves, we derived the area at capacity (A_0), depth at capacity (D_0) and the volume at capacity (V_c) for each reservoir. V_c was obtained from the Global Dam Watch (GDW) inventory (Lehner *et al*. 2024). To obtain A_0, we compared the GDW reported area at capacity (A_{GDW}) with the 95th percentile of the Global Reservoir Surface Area Dataset (GRSAD) (Zhao and Gao, 2018) area time series (A_{95}), with the choice based on Eq. (S1). To account for potential unrealistically large areas reported in GDW, A_{GDW} would qualify to be A_0 only if it fell within a specific range relative to GRSAD areas (within 20% of A_{95}).

$$A_0 = \begin{cases} A_{GDW}, & A_{95} < A_{GDW} \leq 1.2 \times A_{95} \\ A_{95}, & \end{cases} \quad \text{(S1)}$$

D_0 was estimated as 95% of the GDW reported dam height (H_{dam}), *i.e*., $D_0 = 0.95 \times H_{dam}$. When dam height was not available, D_0 was approximated using the volume-to-area ration, as $D_0 = V_c/A_0$. To reduce extrapolation errors during periods of high water storage, we augmented the raw GRDL (Hao *et al.* 2024) bathymetry. If the geometric values from GDW at capacity exceeded those from GRDL, the bathymetry was extended to cover the GDW values. For each reservoir, we tested multiple hypsometric models mapping area to area to volume and area to depth. Candidate models included interpolation techniques—piecewise linear and cubic spline interpolation—as well as polynomials fits ranging from degrees one to five. Building on the methodologies established by Yigzaw *et al*. (2018) and Li *et a*l. (2023)—who determined optimal reservoir geometry by minimising the difference between the estimated and reported storage at capacity—the optimal model in our study was selected by minimising the *combined relative error* (CRE). CRE is a composite metric that weights volumetric and depth errors equally (Eq. (S2)). By converting absolute errors into relative percentages, CRE effectively allows us to compare *apples and oranges* (volumetric and linear values). This

unitless scale ensures that the massive numerical values of reservoir volume do not overshadow the much smaller depth values (or vice versa) in the objective function.

$$CRE = 0.5\left|\frac{V_{est}-V_c}{V_c}\right| + 0.5\left|\frac{D_{est}-D_0}{D_0}\right| \quad \text{(S2)}$$

Here, V_{est} and D_{est} are the volume and depth at capacity respectively estimated by the candidate model. If a model failed to converge for one model and or parameter (resulting in a bogus/non-numeric values), the error metric was calculated solely based on the valid model or parameter. The model yielding the lowest CRE was adopted as the bathymetry to extract the storage time series from GRSAD.

S2. Performance evaluation of satellite-derived storage time series against *in situ* observations.

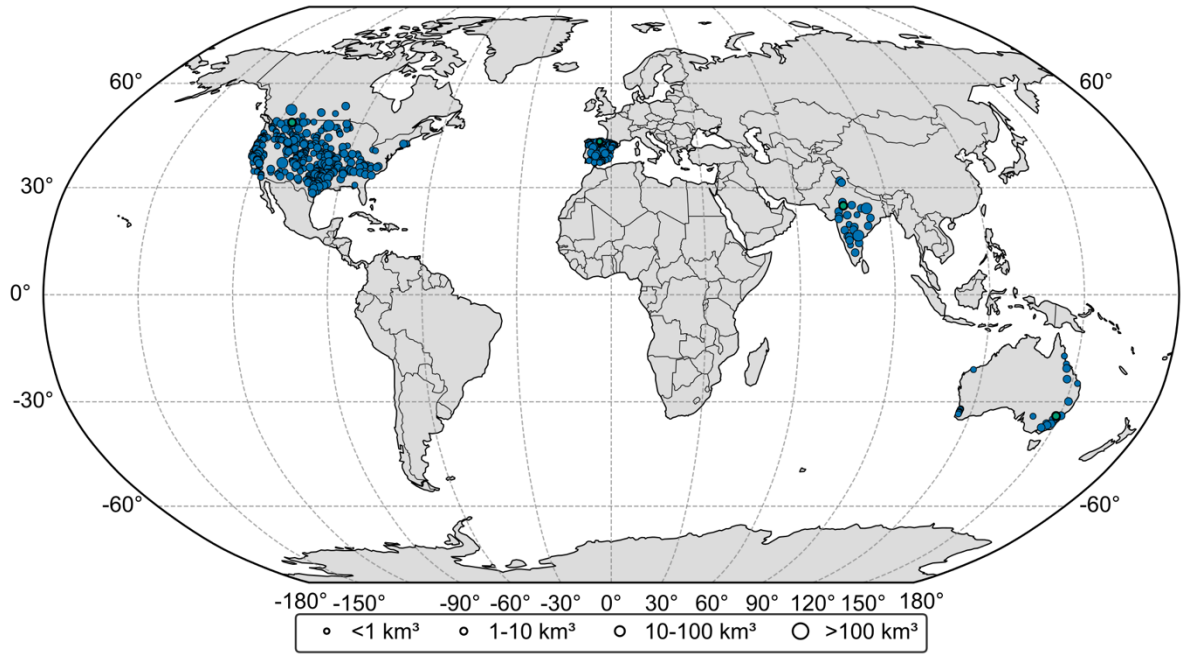


Figure S1: Global distribution of validated reservoirs. Markers denote dams with available *in situ* data and satellite-derived storage time series from GRS, GloLakes and GDS, where each satellite time series overlaps with the *in situ* data for at least 60 months (N = 638 dams). The green markers indicate four selected reservoirs: Libby (the Americas), Ebro (Europe), Rana Pratap Sagar (Asia) and Wyangala (Oceania).

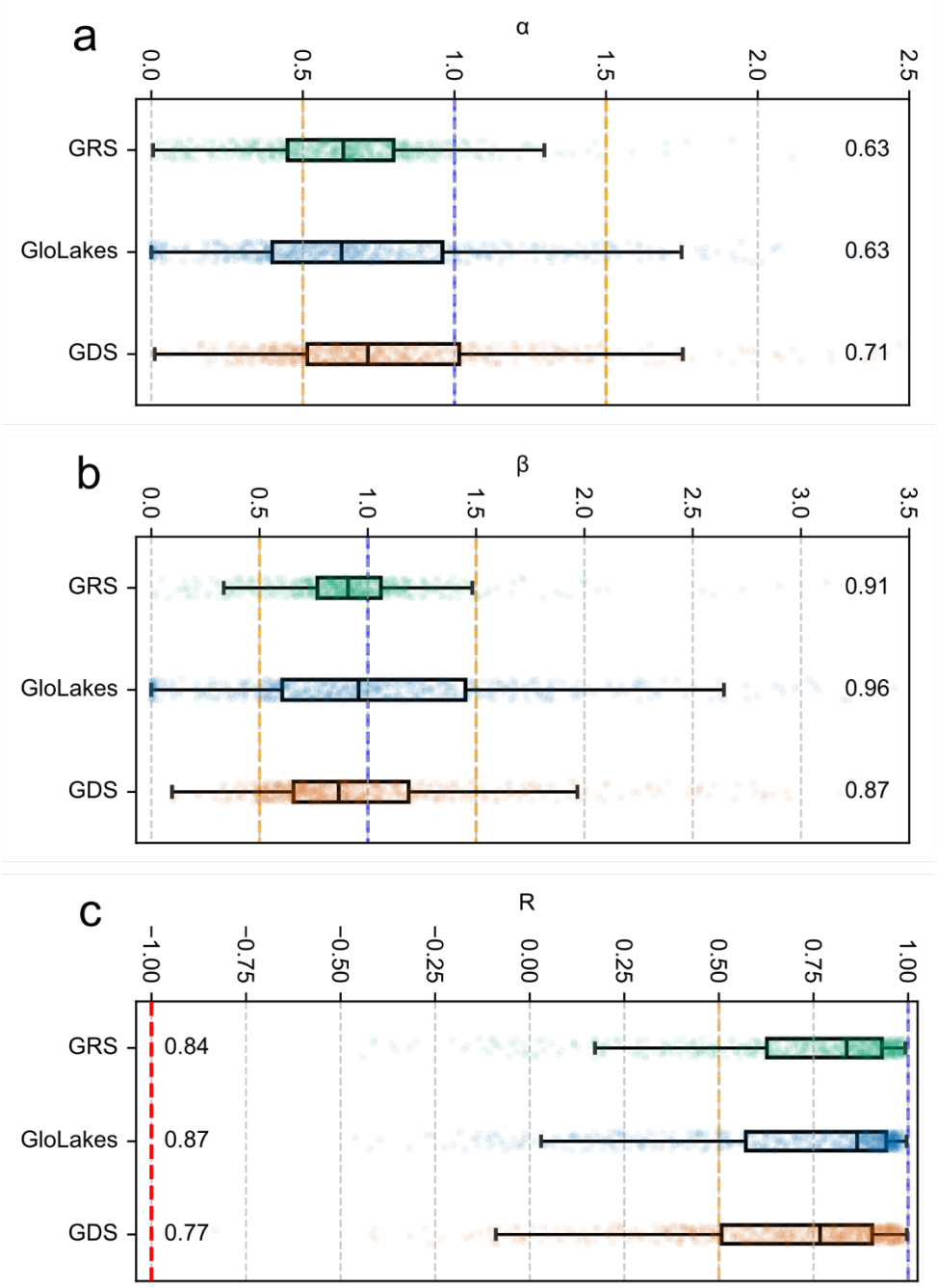


Figure S2: Box plots of KGE components (α, β and R) for GRS, GloLakes and GDS. The dashed orange line marks the threshold for a skilful value while the dashed blue line indicates the optimal value. Median values are shown as vertical black lines within each box, with corresponding numerical labels. Whiskers represent 1.5× IQR, where IQR is the *interquartile range*.

S3. Formulation of the LISFLOOD (LIS) reservoir operation scheme (Burek *et al*. 2013, 2020).

$$\mathrm{Rls} = \begin{cases} \min(\mathrm{Rls_{min}}, \frac{1}{\Delta t}\mathrm{FS}), \ \mathrm{F} \leq 2\mathrm{LC} \\ \mathrm{Rls_{min}} + (\mathrm{Rls_{norm}} - \mathrm{Rls_{min}})\left(\frac{\mathrm{F}-2\mathrm{LC}}{\mathrm{LN}-2\mathrm{LC}}\right), \ 2\mathrm{LC} < \mathrm{F} \leq \mathrm{LN} \\ \mathrm{Rls_{norm}} + \frac{(\mathrm{F}-\mathrm{LN})}{(\mathrm{LF}-\mathrm{LN})}\max\{(\mathrm{inf_m} - \mathrm{Rls_{norm}}), (\mathrm{Rls_{nd}} - \mathrm{Rls_{norm}})\}, \ \mathrm{LN} < \mathrm{F} \leq \mathrm{LF} \\ \max\left\{\frac{(\mathrm{F}-\mathrm{LF})}{\Delta t}\mathrm{S}, \mathrm{Rls_{nd}}\right\}, \mathrm{Rls_{nd}}, \ \mathrm{F} > \mathrm{LF}, \end{cases} \quad \text{(S3)}$$

where, Rls is the release from the reservoir (kg/s), $\mathrm{Rls_{min}}$ is the minimum release from the reservoir (kg/s), $\mathrm{Rls_{norm}}$ is the normal release from the reservoir (kg/s), $\mathrm{Rls_{nd}}$ is the non-damaging release from the reservoir (kg/s) and $\mathrm{inf_m}$ is the instantaneous inflow into the reservoir in month m (kg/s), S is reservoir storage capacity (kg), F is the reservoir fill fraction (one at reservoir storage capacity and zero when empty), LC is the conservation limit (we adopted a value of 0.1 for global simulation), LN is the normal storage limit (we adopted a value of 0.5 for global simulations) and LF is the flood storage limit (we adopted a value of 0.9 for the global simulation).

We set $\mathrm{Rls_{norm}}$ as the long-term average annual discharge, $\mathrm{Rls_{min}} = 0.2 \times \mathrm{Rls_{norm}}$ and $\mathrm{Rls_{nd}} = 4 \times \mathrm{Rls_{norm}}$, following (Burek *et al*. 2013, 2020).

S4. Skill recovery from modified inflow into the reservoir

During preliminary diagnostics, we found a logic error in H08's handling of reservoir inflow: H08 used the previous step's outflow from the reservoir's local grid cell instead of the true upstream inflow entering the cell. This introduced a feedback that recirculated dam release, breaking connectivity to upstream cells. We corrected the code ad recalculated H08 storage simulations in strict compliance with ISMIP3a protocol.

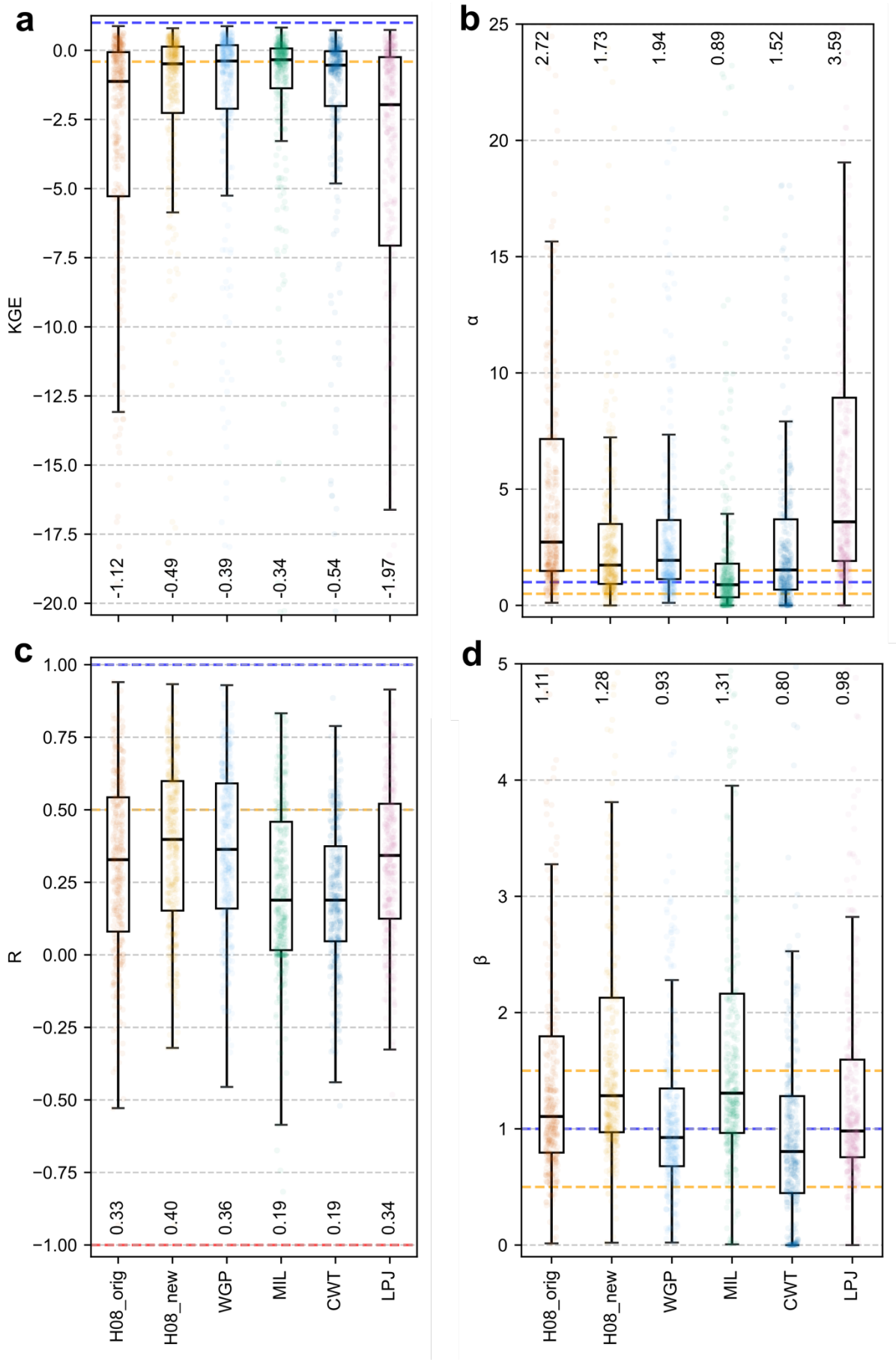

a
KGE
0.0
−2.5
−5.0
−7.5
−10.0
−12.5
−15.0
−17.5
−20.0
-1.12
-0.49
-0.39
-0.34
-0.54
-1.97
b
α
25
20
15
10
5
0
2.72
1.73
1.94
0.89
1.52
3.59
c
R
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0.33
0.40
0.36
0.19
0.19
0.34
d
β
5
4
3
2
1
0
1.11
1.28
0.93
1.31
0.80
0.98
H08_orig
H08_new
WGP
MIL
CWT
LPJ

Figure S3: Comparison of the original ISIMIP3a H08 reservoir storage submission (H08_orig) against the corrected simulation run (H08_new) alongside the other four GHMs (WGP, MIL, CWT and LPJ). The boxplots display the distribution of (a) KGE, (b) Variability ratio (α), (c) Pearson Correlation coefficient (R), and (d) Bias ratio (β) across 371 reservoirs. The bug fix in the H08 code results in a notable recovery of skill, particularly in the KGE score (median improvement from -1.12 to -0.49) and correlation (R), bringing H08 performance into alignment with the highest performing models. Note that the main text utilises 424 reservoirs common to all five GHMs. This is an expansion from the earlier subset of 371 reservoirs, prior to fixing the bug in H08. Specifically, fixing the bug yielded 55 additional useable reservoir time series and the exclusion of two reservoirs due to undefined KGE values.

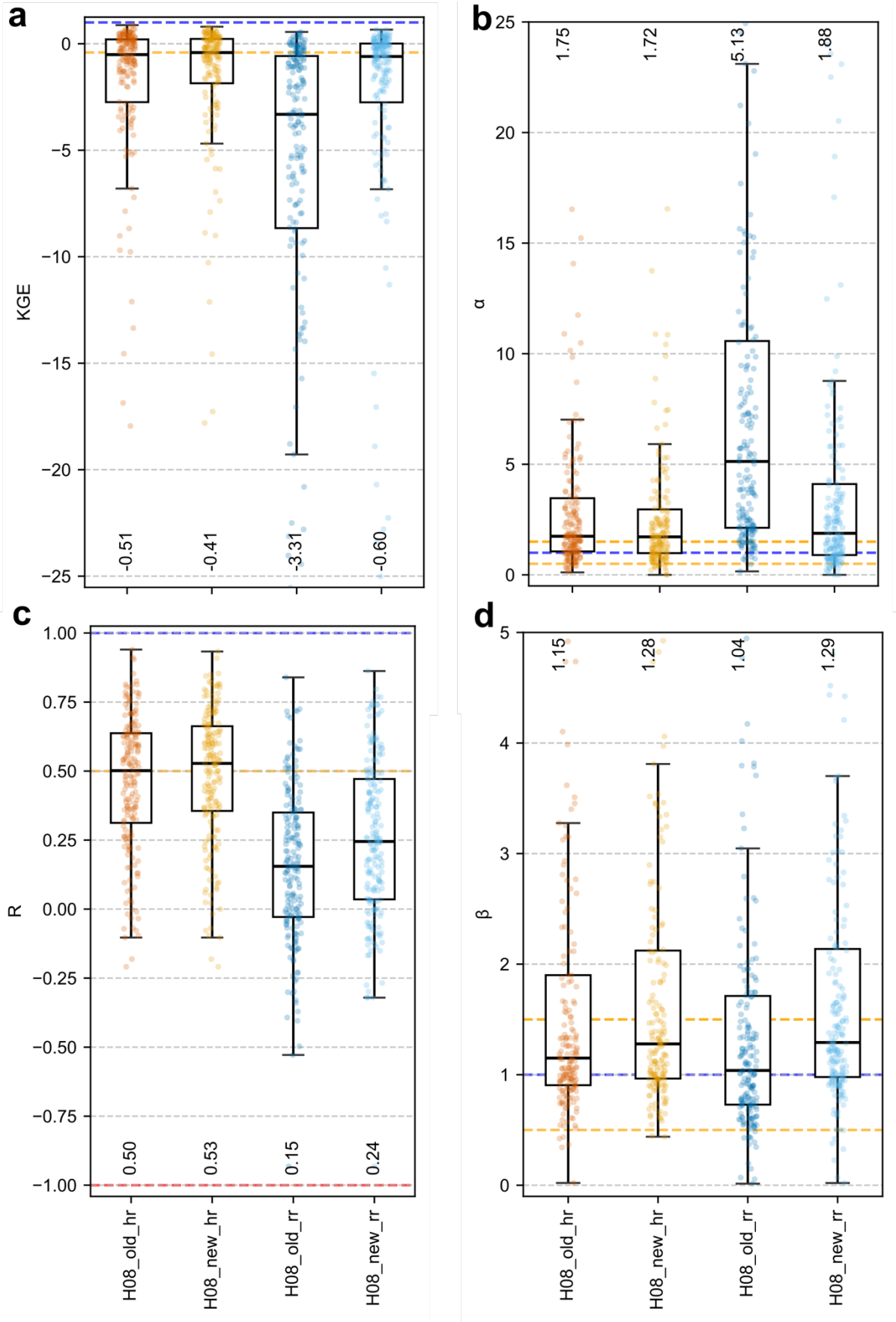


Figure S4: Comparison of H08 reservoir storage simulations pre- and post-bug fix, shown separately for high and low regulation reservoirs. Reservoirs are grouped by *degree of regulation* (DOR): high-regulation (hr) reservoirs, with DOR ≥

0.5 and low regulation or run-of-the river (rr) reservoirs with DOR < 0.5. Again, in the original H08 code, inflow for each time step was incorrectly set using the outflow from the previous time step, an error that disproportionately affected low-regulation (rr) reservoirs (their release is heavily dependent on the instantaneous monthly inflow). After fixing this bug, H08 simulation of low-regulation reservoirs (rr) reservoirs (H08_new_rr) saw a substantial improvement in KGE, with the median increasing from -3.31 to -0.60. In contrast, high-regulation (hr) reservoirs show only minor performance improvements.

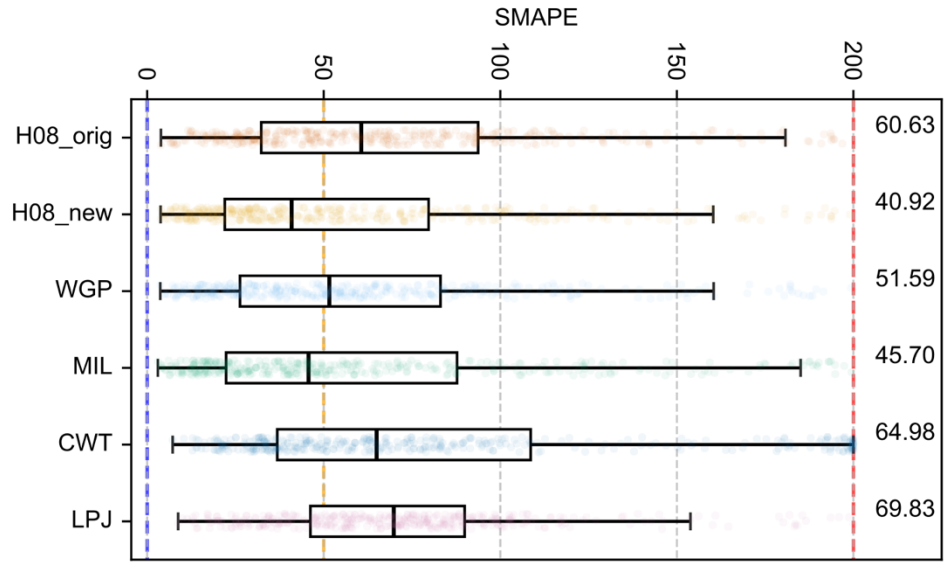


Figure S5: Reduction in Symmetric Mean Absolute Percentage Error (SMAPE) following the bug fix. Distribution of SMAPE for the original (H08_orig) and corrected (H08_new) H08 simulations compared against other GHMs. The correction reduces H08's median SMAPE from 60.63% to 40.92%, making the corrected H08 simulation highly competitive within the ensemble. The solid black line in the box plot denotes the median value (corresponding annotations of median for each GHM are shown), while the dashed blue line shows SMAPE=0 and the dashed red line is the worst possible value, SMAPE=200 %. This is a comparison for 371 dams across the globe.

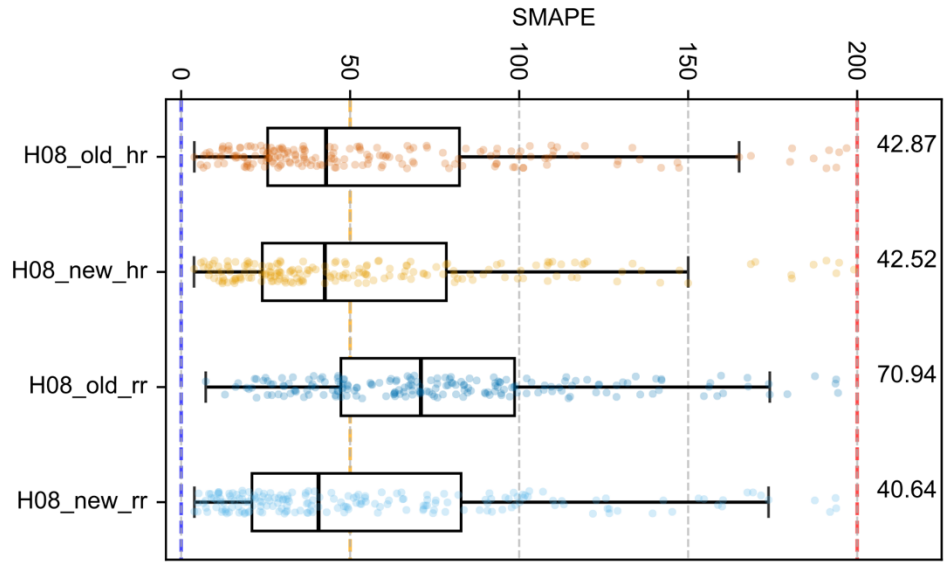


Figure S6: Impact of the code correction on SMAPE stratified by regulation type. Decomposition of SMAPE results for high regulation (hr) and low regulation (rr) reservoirs. Consistent with the KGE results in Fig. S4, the error reduction is

driven almost entirely by the rr subset, where the median SMAPE drops from 70.94% in the original run to 40.64% in the corrected run. The high regulation subset (hr) remains stable (42.87% to 42.52%), confirming that the feedback loop bug primarily degraded the performance of reservoirs with smaller storage-to-inflow ratios.

S5. Intercomparison of H06 and LIS schemes implemented in H08 using default global parameters.

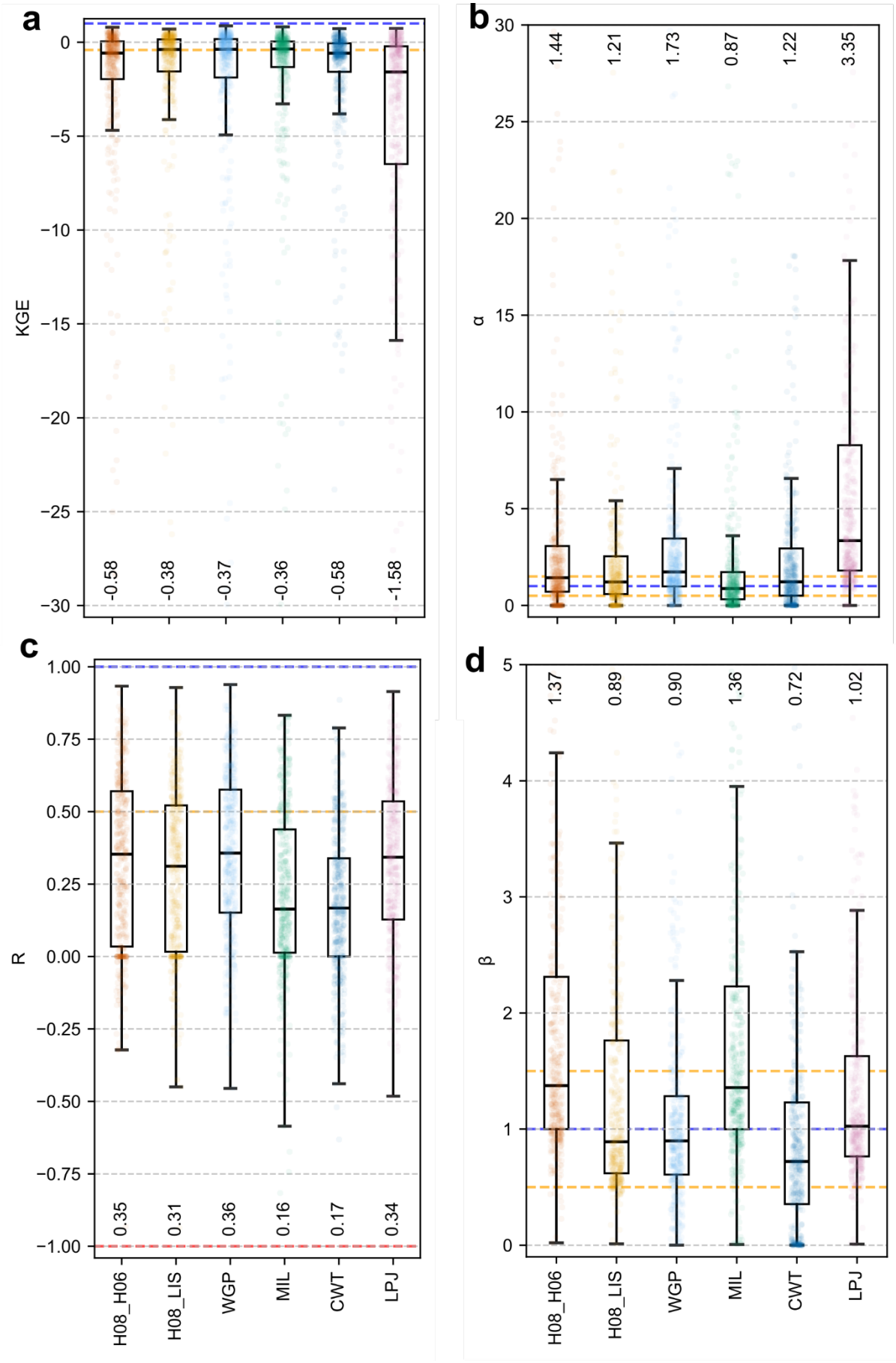

a
KGE
0
−5
−10
−15
−20
−25
−30
-0.58
-0.38
-0.37
-0.36
-0.58
-1.58
b
α
30
25
20
15
10
5
0
1.44
1.21
1.73
0.87
1.22
3.35
c
R
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0.35
0.31
0.36
0.16
0.17
0.34
d
β
5
4
3
2
1
0
1.37
0.89
0.90
1.36
0.72
1.02
H08_H06
H08_LIS
WGP
MIL
CWT
LPJ

Figure S7*:* Performance evaluation of reservoir storage simulations by ISMIP3a GHMs against GRS (N = 424 dams). Box plots of (a) KGE, (b) its variability term (α), the (c) Pearson correlation coefficient (R) and (d) the bias-term (β) scores for five ISIMIP3a GHMs (H08, WGP, MIL, CWT and LPJ). H08_H06 is H08 with its default H06 reservoir operation scheme, while H08_LIS is H08 with the LIS reservoir operation scheme. The vertical, dashed orange line marks the threshold for skilful simulations, while the vertical, dashed blue line indicates the best possible value (KGE = 1, α =1, β=1 and R=1). Median values are denoted by vertical, solid black lines within each box, with corresponding numerical annotations. Whiskers represent 1.5 × IQR, where IQR is the interquartile range.

S6. Performance evaluation of reservoir storage simulations by ISMIP3a GHMs against GRS.

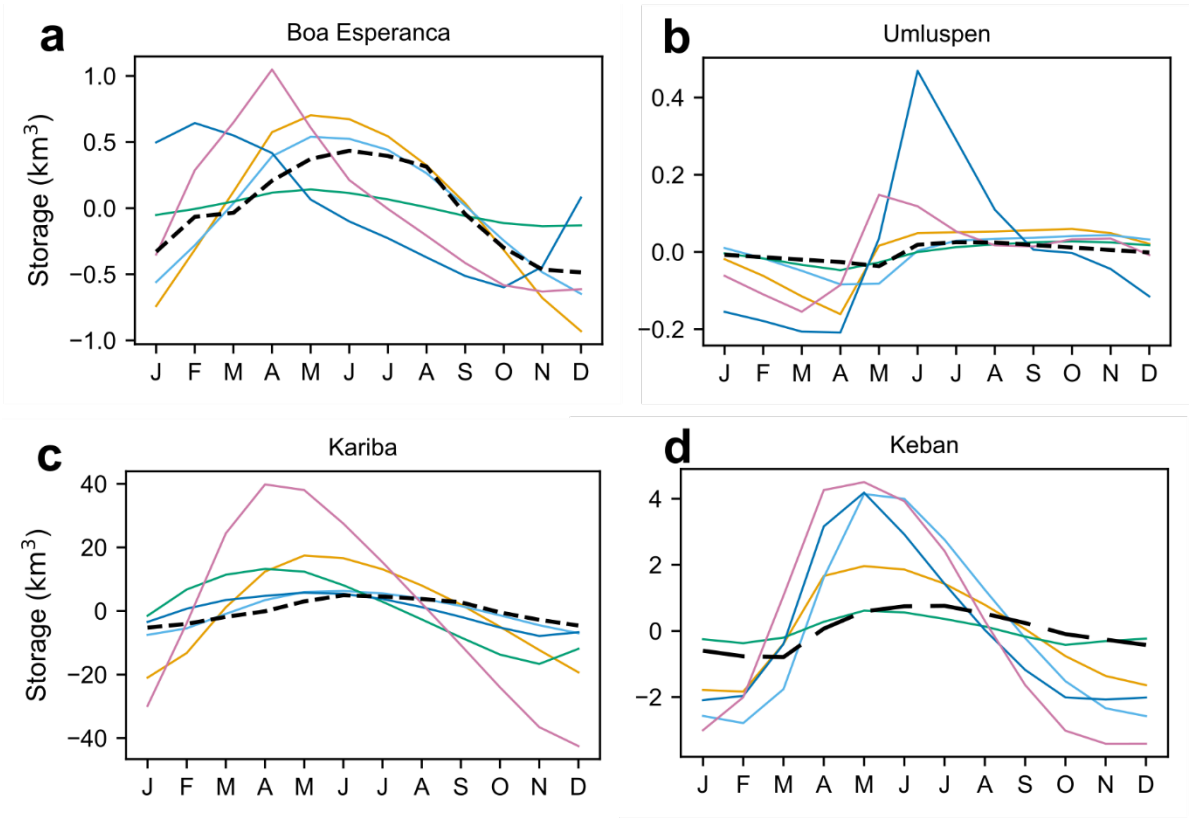


Figure S8: Seasonal time series for select dams. The plots represent the seasonal component extracted from the decomposition of raw (non-bias corrected) long-term monthly time series into their annual, seasonal and residual components. Note that the colour scheme is as in Fig. S11.

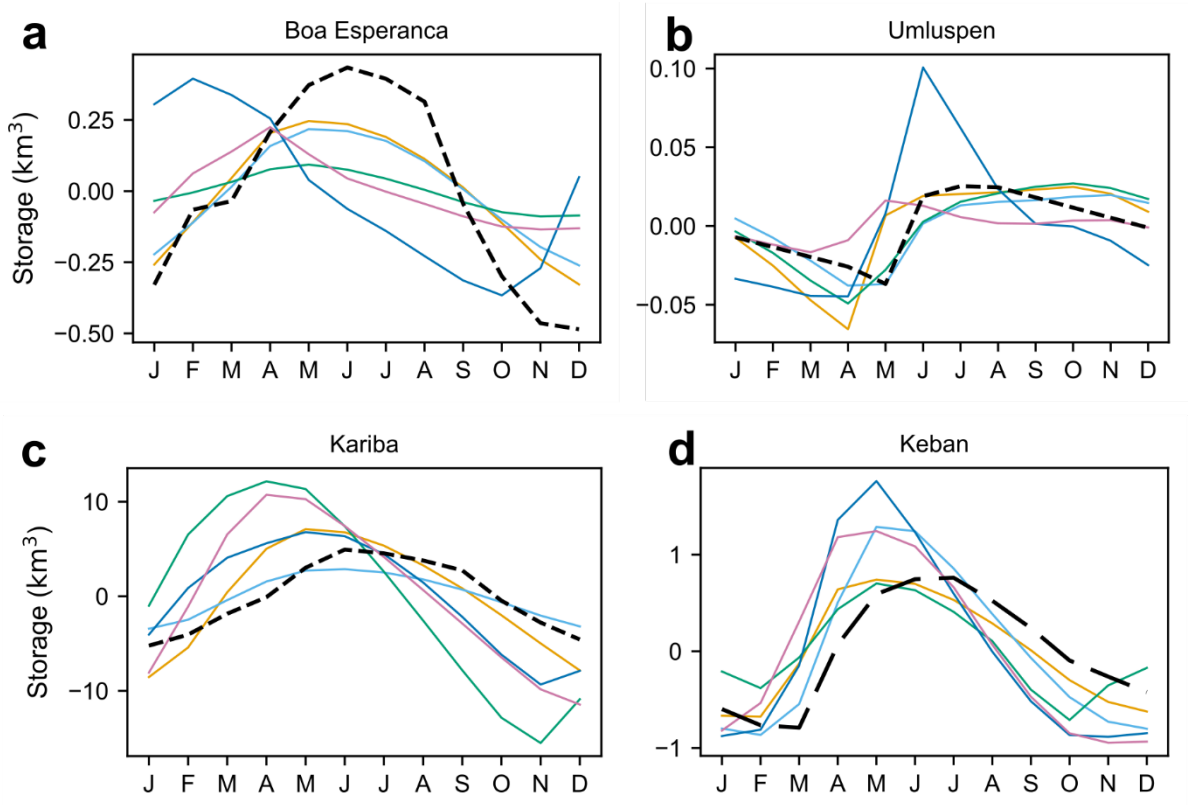


Figure S9: Same as Fig. S8, but for variance-matching bias corrected time series.

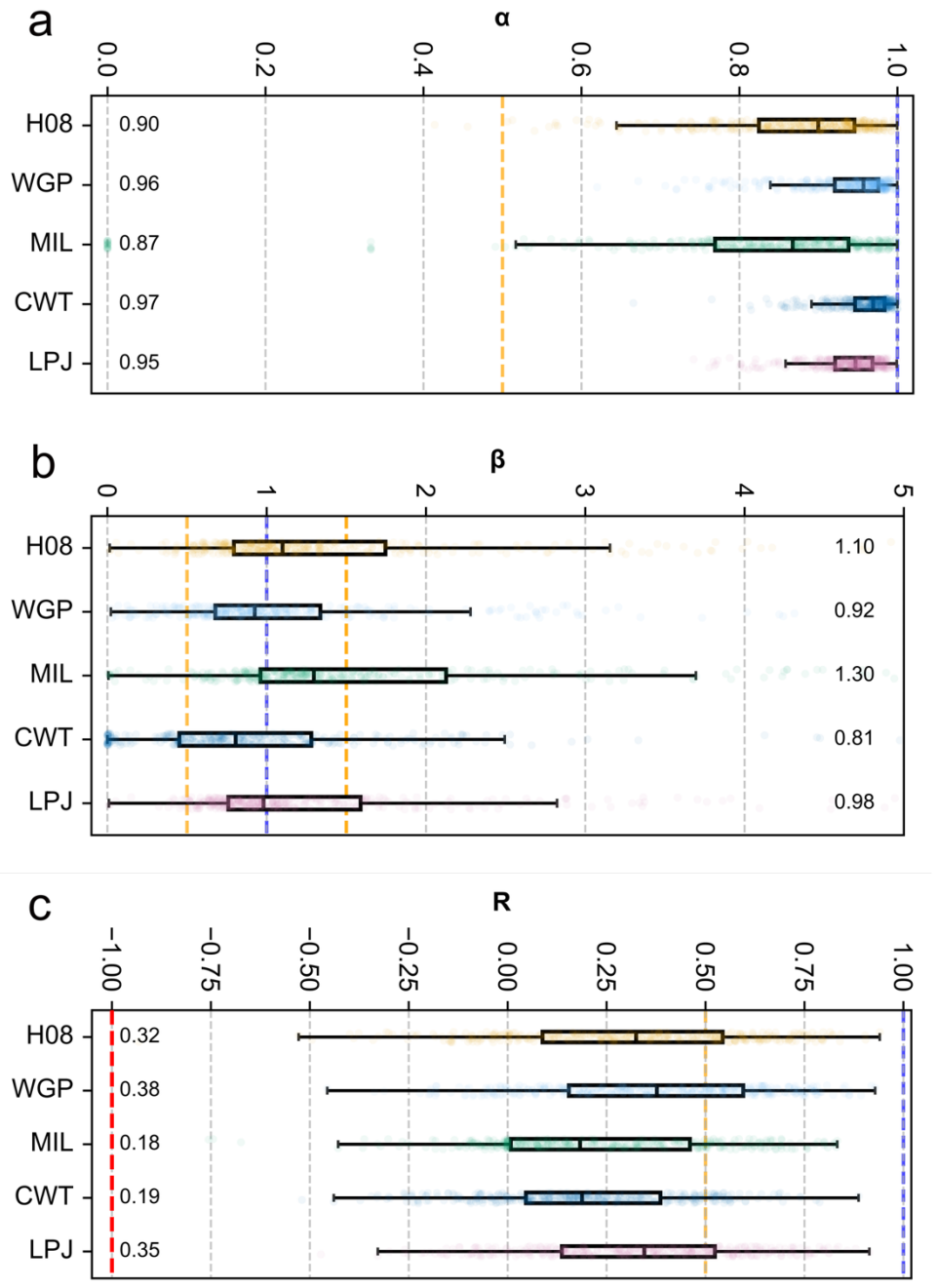


Figure S10: Performance evaluation of variance-matched bias corrected reservoir storage simulations by ISMIP3a GHMs against GRS (N = 409 dams). (a) Box plots of α component (b) β component and (c) R component of KGE for five ISIMIP3a GHMs (H08, WGP, MIL, CWT and LPJ) with overlaid strip plots coloured according to the time series legend. The orange, vertical dashed orange line marks the threshold for skilful simulations (0.5 for α, 1.5 and 0.5 for β and 0.5 for R), while the vertical, dashed blue line indicates the best possible value (KGE = 1). Median values are denoted by vertical, black lines within each box, with corresponding numerical annotations. Whiskers represent 1.5 × IQR, where IQR is the interquartile range.

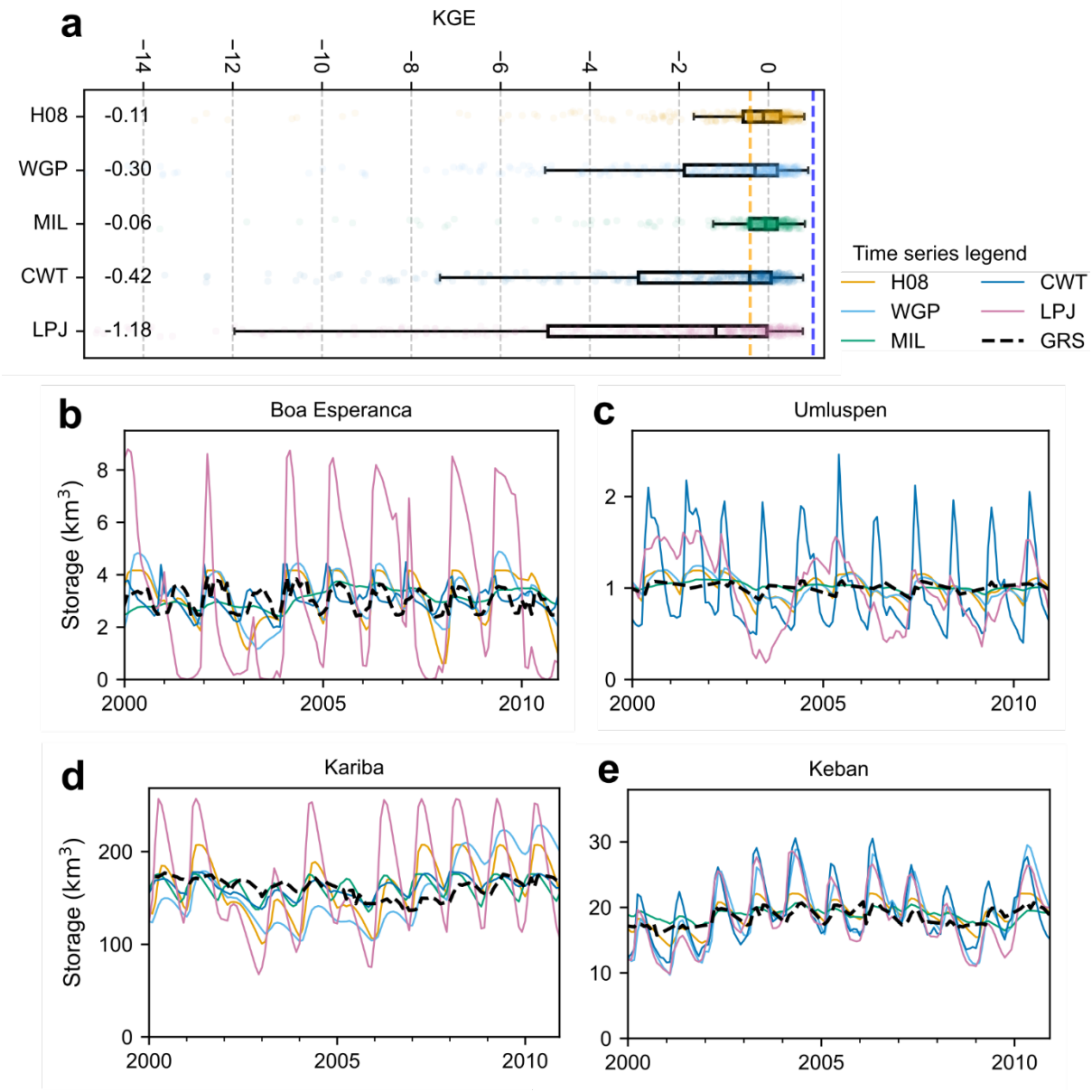


Figure S11: Performance evaluation of reservoir storage simulations by ISIMIP3a GHMs against GRS after applying linear bias correction (N = 431 dams). (a) Box plots of KGE scores for five ISIMIP3a GHMs (H08, WGP, MIL, CWT and LPJ) with overlaid strip plots coloured according to the time series legend. The vertical, dashed orange line marks the threshold for skilful simulations (KGE = -0.41), while the vertical, dashed blue line indicates the best possible value (KGE = 1). Median values are denoted by vertical, black lines within each box plot, with corresponding numerical annotations. Whiskers represent 1.5 × IQR. (b–e) Representative long-term monthly series (2000–2010) for select dams.

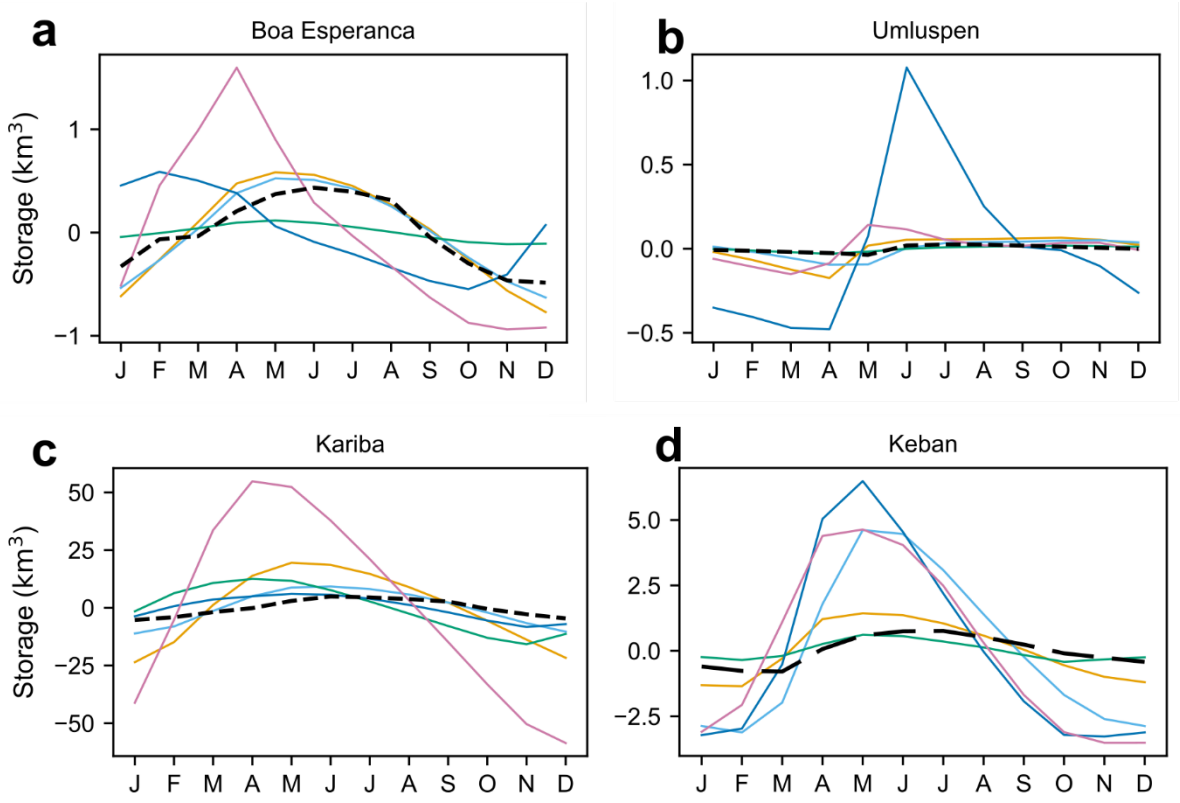


Figure S12: Same as Fig. S8 but for linear bias corrected time series.

Table S1: Performance metrics—KGE scores and their corresponding components (α, β and R terms)—for the five ISIMIP3a GHMs across selected dams. Note that these metrics are derived from long-term monthly time series after linear bias correction.

GRanD ID	Dam Name	Metric	ISIMIP3a GHMs				
			H08	WGP	MIL	CWT	LPJ
2490	Boa Esperança	KGE	-0.28	-0.22	0.09	0.00	-4.51
		R	0.58	0.47	0.09	0.09	0.35
		α	2.21	2.10	0.93	1.40	6.47
		β	1.00	1.00	1.00	1.00	1.00
3720	Umluspen	KGE	-0.70	-0.55	0.19	-8.48	-6.02
		R	0.32	0.34	0.25	0.24	0.19
		α	2.56	2.40	0.69	10.45	7.97
		β	1.00	1.00	1.00	1.00	1.00
4056	Kariba	KGE	-0.65	-1.11	0.22	0.44	-3.00
		R	0.56	0.52	0.22	0.47	0.36
		α	2.59	3.05	0.99	0.85	4.94
		β	1.00	1.00	1.00	1.00	1.00
4422	Keban	KGE	0.13	-1.54	0.18	-1.63	-1.62
		R	0.64	0.71	0.19	0.53	0.58
		α	1.80	3.53	0.89	3.59	3.59
		β	1.00	1.00	1.00	1.00	1.00

S7. Uncertainty in GHM storage simulations using different climate forcings data.

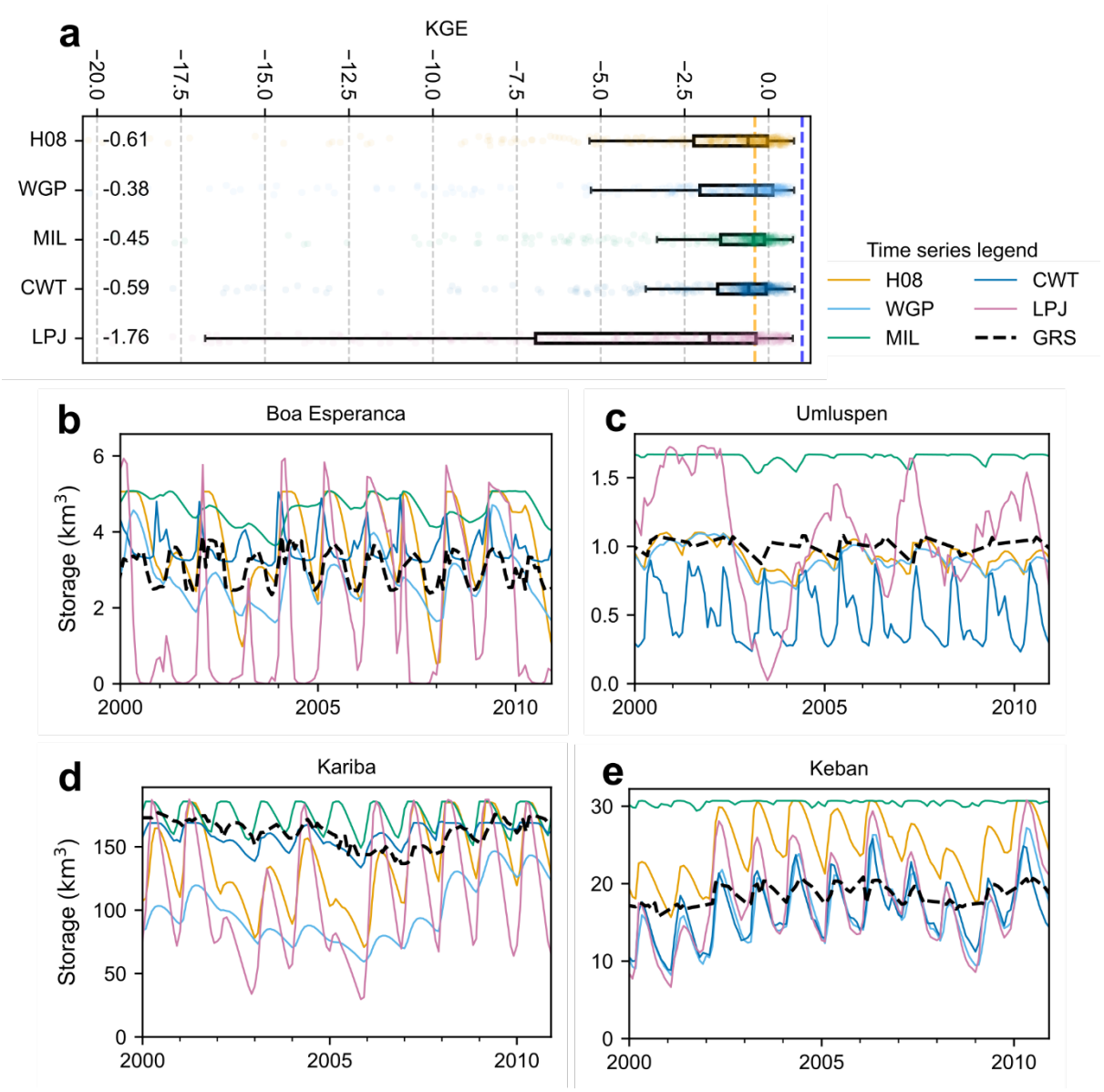


Figure S13: Performance evaluation of reservoir storage simulations by ISMIP3a GHMs against GRS (N = 424 dams). These are simulations using the 20CRv3-ERA5 (CE) climate forcing. (a) Box plots of KGE scores for five ISIMIP3a GHMs (H08, WGP, MIL, CWT and LPJ) with overlaid strip plots coloured according to the time series legend. The vertical, dashed orange line marks the threshold for skilful simulations (KGE = -0.41), while the vertical, dashed blue line indicates the best possible value (KGE = 1). Median values are denoted by vertical, black lines within each box plot, with corresponding numerical annotations. Whiskers represent 1.5 × IQR. (b–e) Representative long-term monthly series (2000–2010) for select dams.

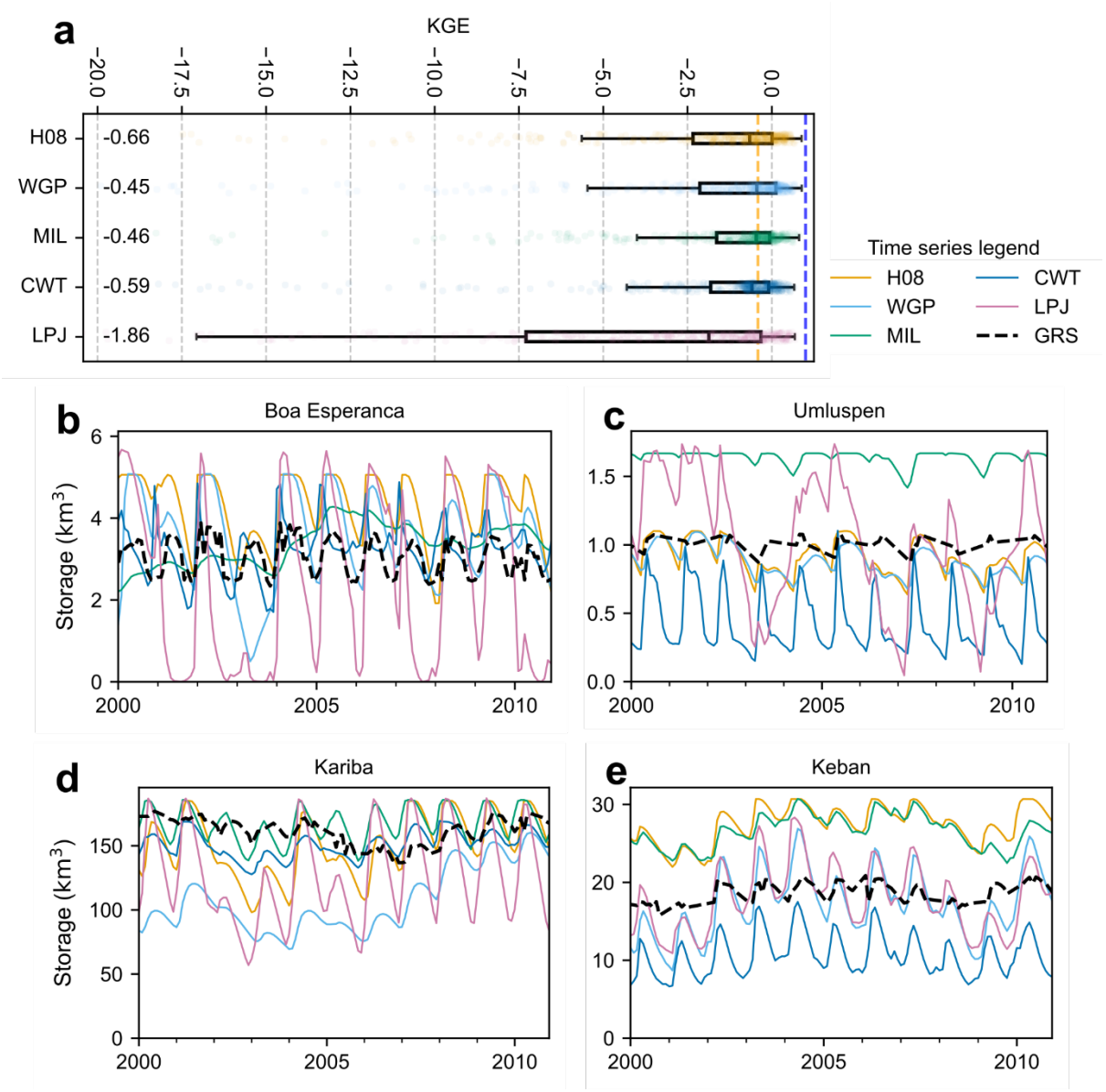


Figure S14: Same as Fig. S13, but for storage simulated using 20CRv3-W5E5 (CW) climate forcing.

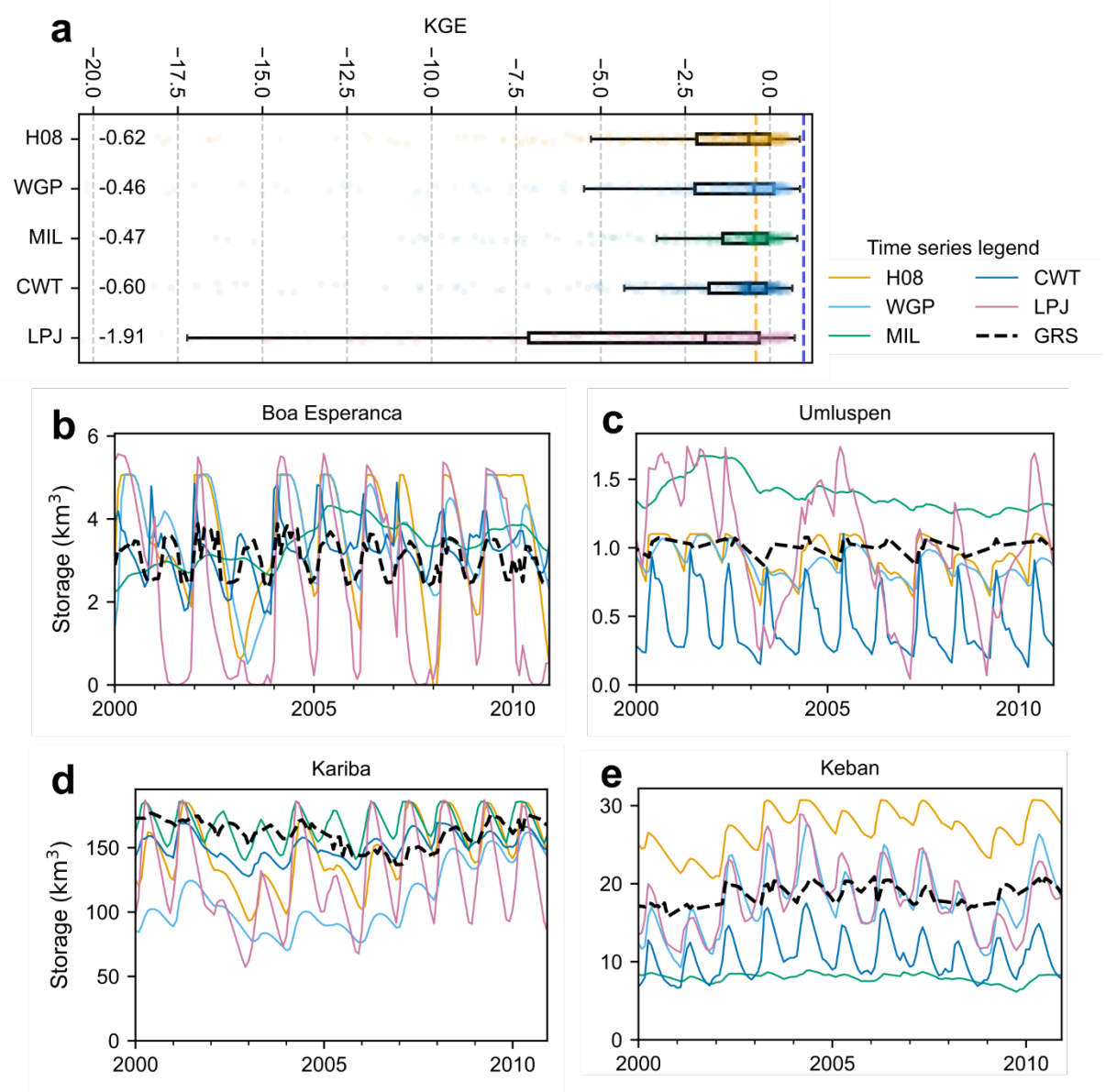
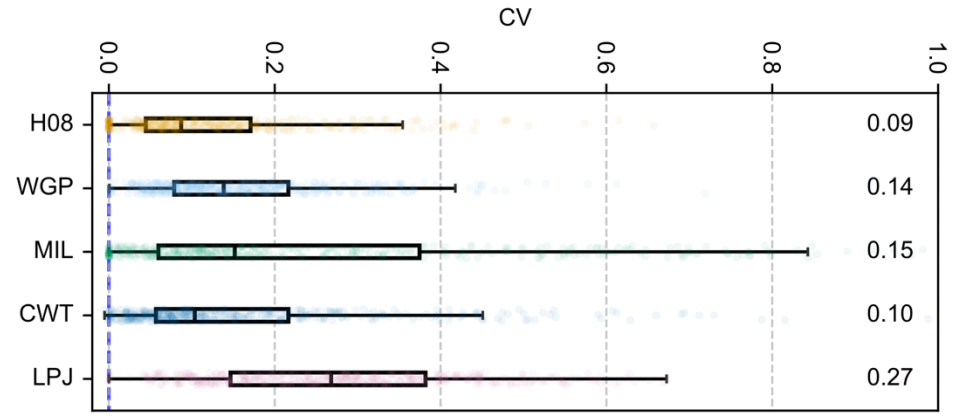


Figure S15: Same as Fig. S13, but for storage simulated using GSWP3-W5E5 (GW) climate forcing.



185

Figure S16: Box plots showing the co-efficient of variation (CV) for five ISIMIP3a GHMs (H08, WGP, MIL, CWT and LPJ) driven by three climate forcings (CE, CW and GW). The vertical, dashed blue line, at CV = 0, indicates perfect consistency in the simulated storage time series across the different climate forcings. Median values are denoted by vertical, black lines within each box plot, with corresponding numerical annotations. Whiskers represent 1.5 × IQR.

S8. Comparison of ISMIP3a GHM simulated storage discharge against GDS

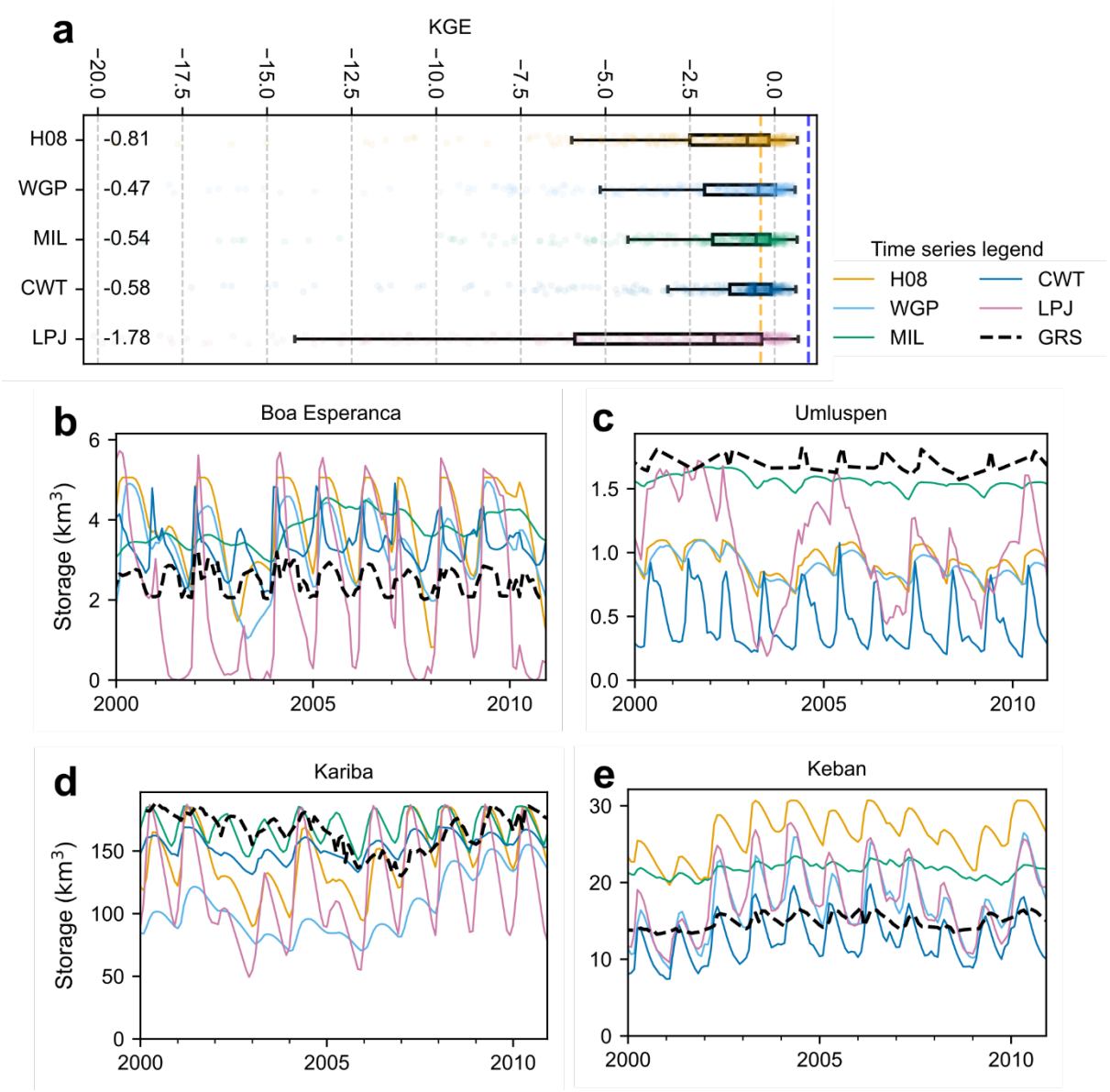


Figure S17: Performance evaluation of reservoir storage simulations by ISIMIP3a GHMs against GDS (N = 424 dams). (a) Box plots of KGE scores for five ISIMIP3a GHMs (H08, WGP, MIL, CWT and LPJ) with overlaid strip plots coloured according to the time series legend. The vertical, dashed orange line marks the threshold for skilful simulations (KGE = -0.41), while the vertical, dashed blue line indicates the best possible value (KGE = 1). Median values are denoted by vertical, black lines within each box plot, with corresponding numerical annotations. Whiskers represent 1.5 × IQR. (b–e) Representative long-term monthly series (2000–2010) for select dams.

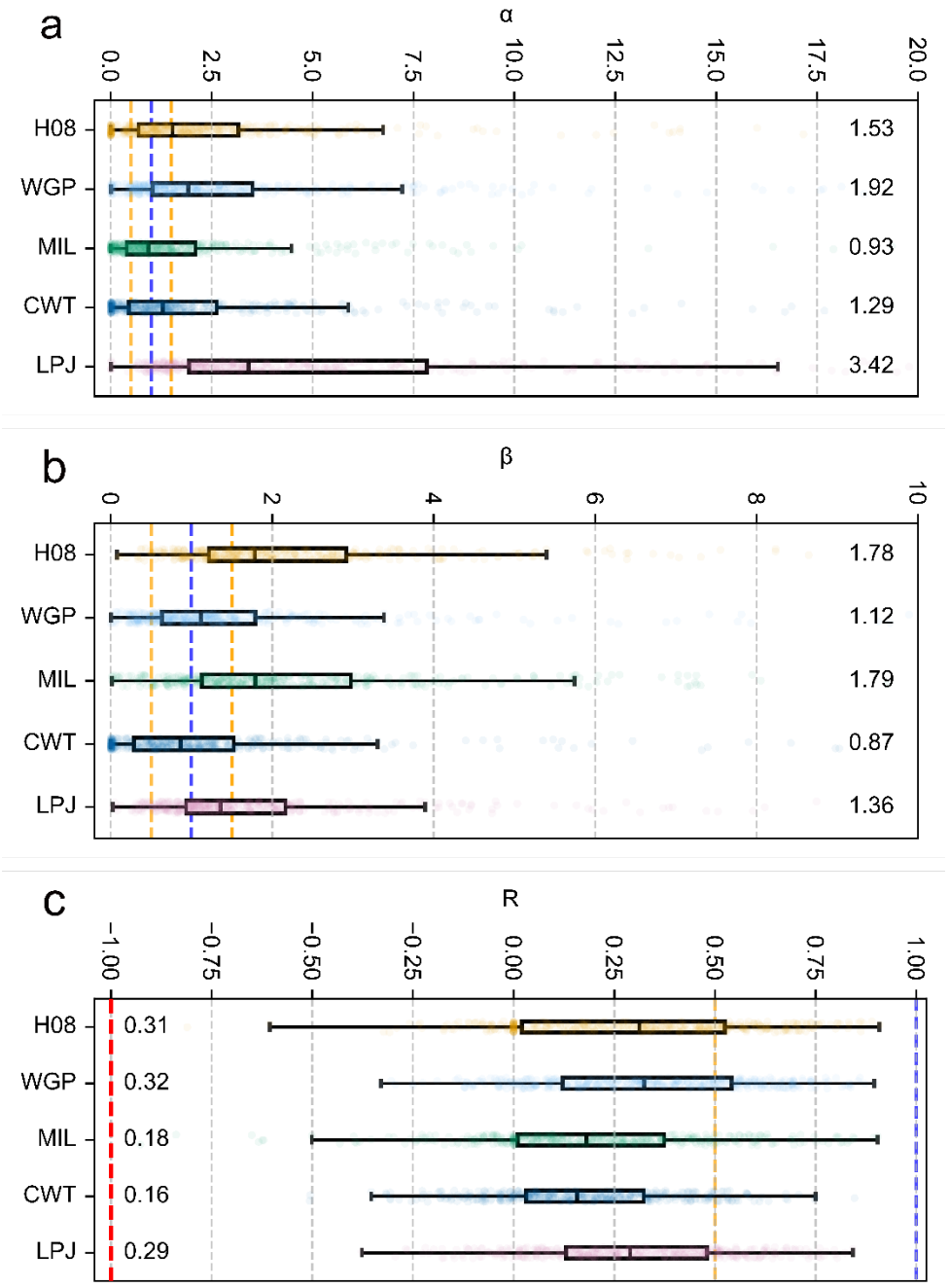


Figure S18: Same as Fig. S17, but for the KGE components: the alpha (α), beta (β) and the correlation (R) terms.

S9. Comparison of ISMIP3a GHM simulated storage discharge against GRS for only *global* dams across all GHMs

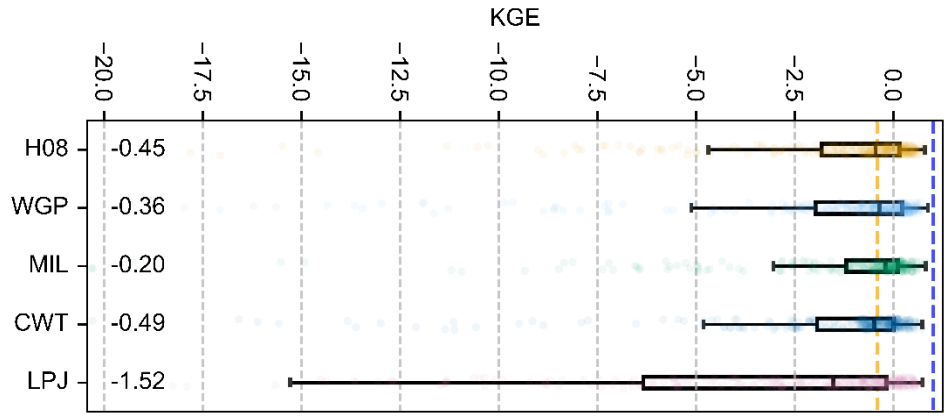


Figure S19: Performance evaluation of reservoir storage simulations by ISIMIP3a GHMs against GRS for only dams classified as *global* by all the GHMs (N = 321 dams). Box plots of KGE scores for five ISIMIP3a GHMs (H08, WGP, MIL, CWT and LPJ). The vertical, dashed orange line marks the threshold for skilful simulations (KGE = -0.41), while the vertical, dashed blue line indicates the best possible value (KGE = 1). Median values are denoted by vertical, black lines within each box plot, with corresponding numerical annotations. Whiskers represent 1.5 × IQR.

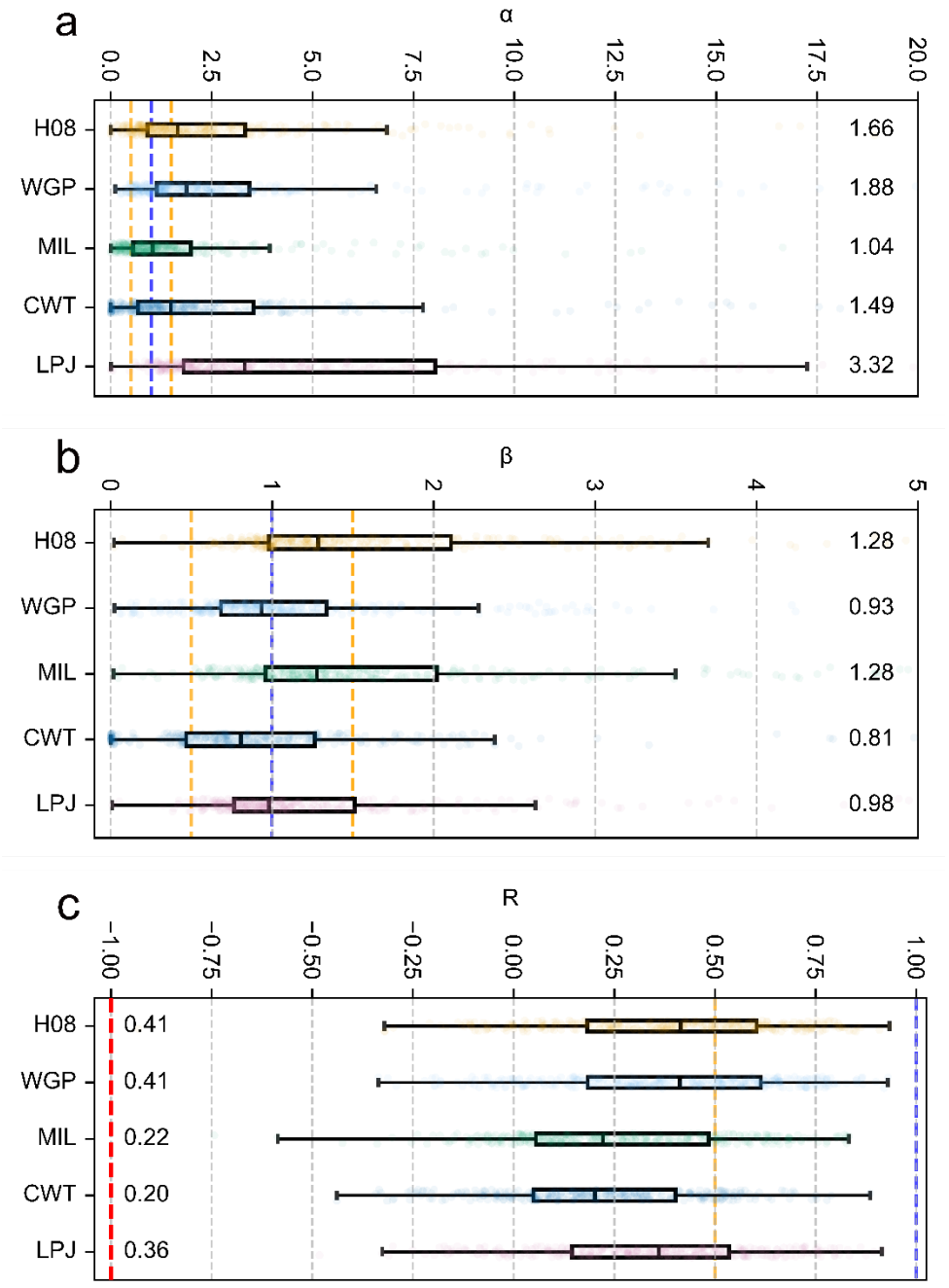


Figure S20: Same as Fig. S19, but for KGE components: the alpha (α), beta (β) and the correlation (R) terms.

S10. Parameter sensitivity test of H06 and LIS

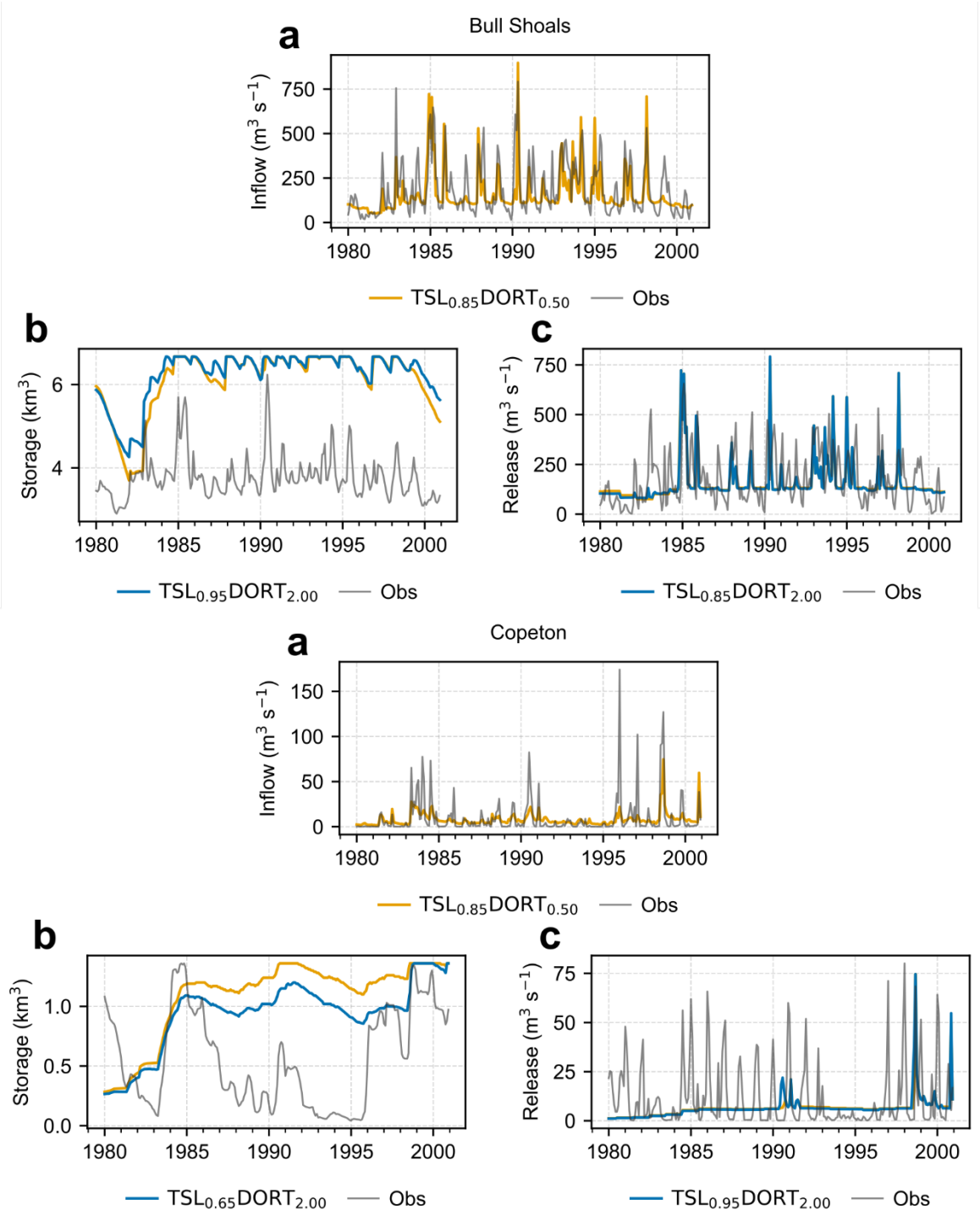


Figure S21: Sensitivity analysis of H06 parameters at Bull Shoals and Copeton dams. (a) Long-term monthly observed (*in situ*) inflow compared to H08 simulations using default global parameters (TSL = 0.85 and DORT = 0.5). (b) Comparison of observed (*in situ*) storage against H08 storage simulated with default global parameters (vermilion curve) and the best-performing tuned parameter set (blue curve). (c) Same as panel (b) but for release time series.

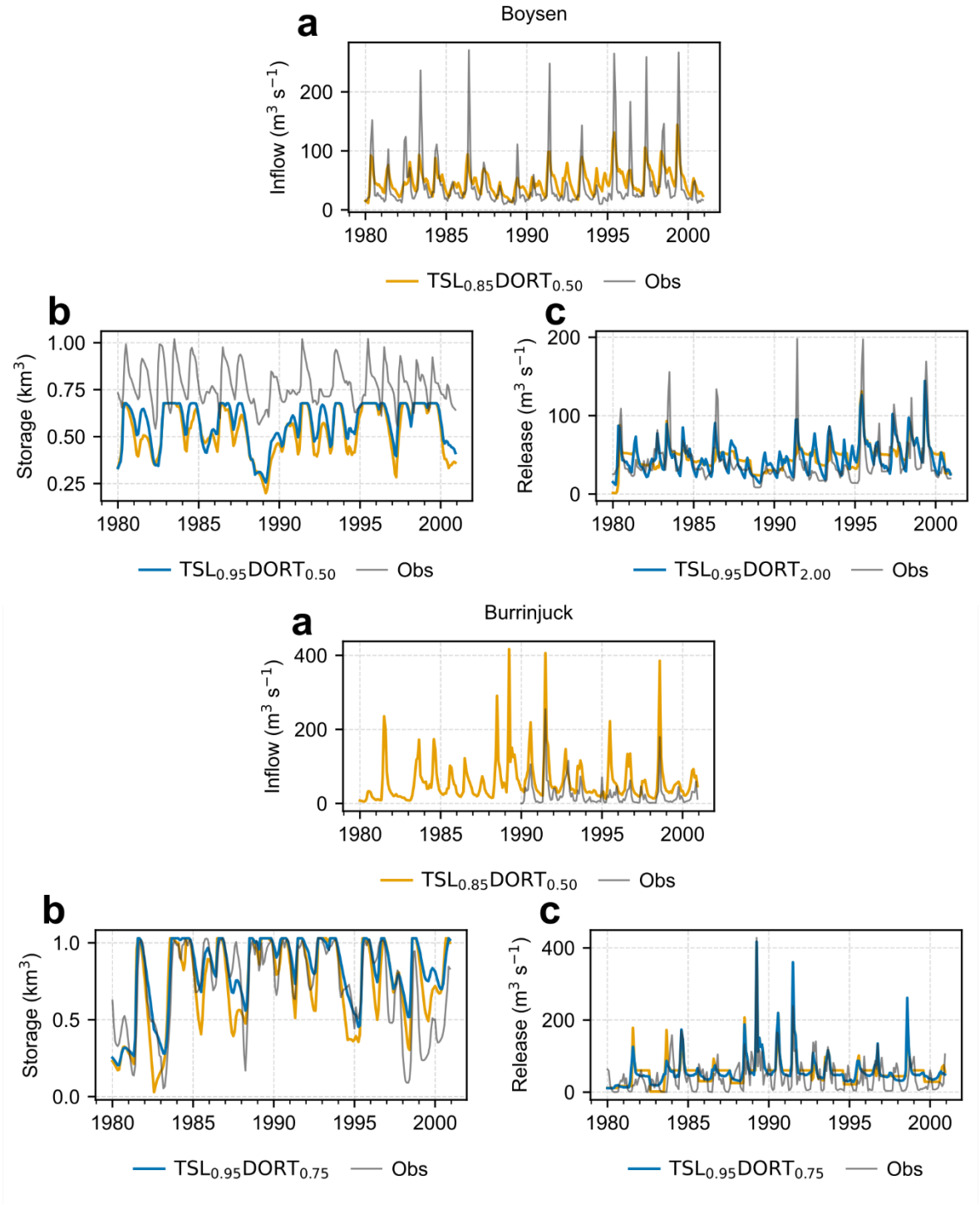
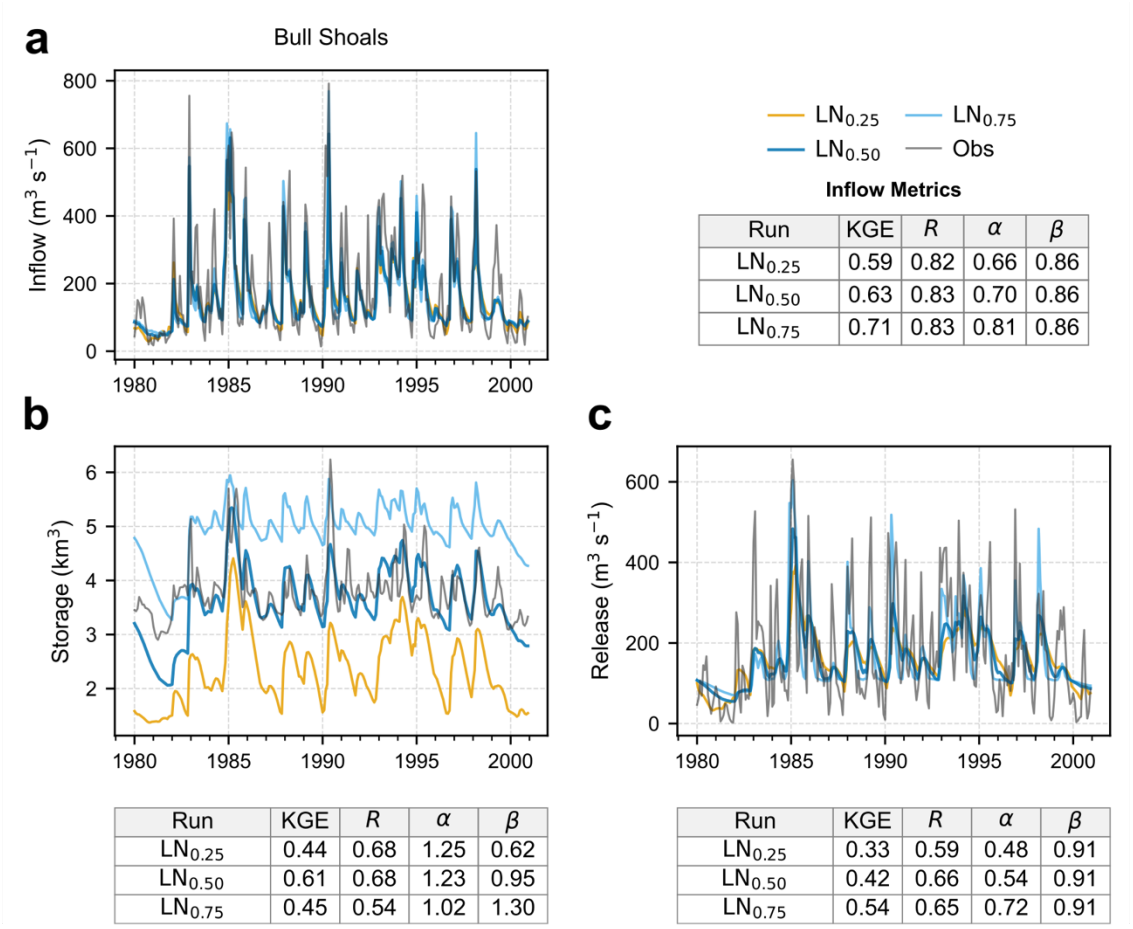


Figure S22: Same as Fig. S21 but for Boysen and Burrinjuck dams.

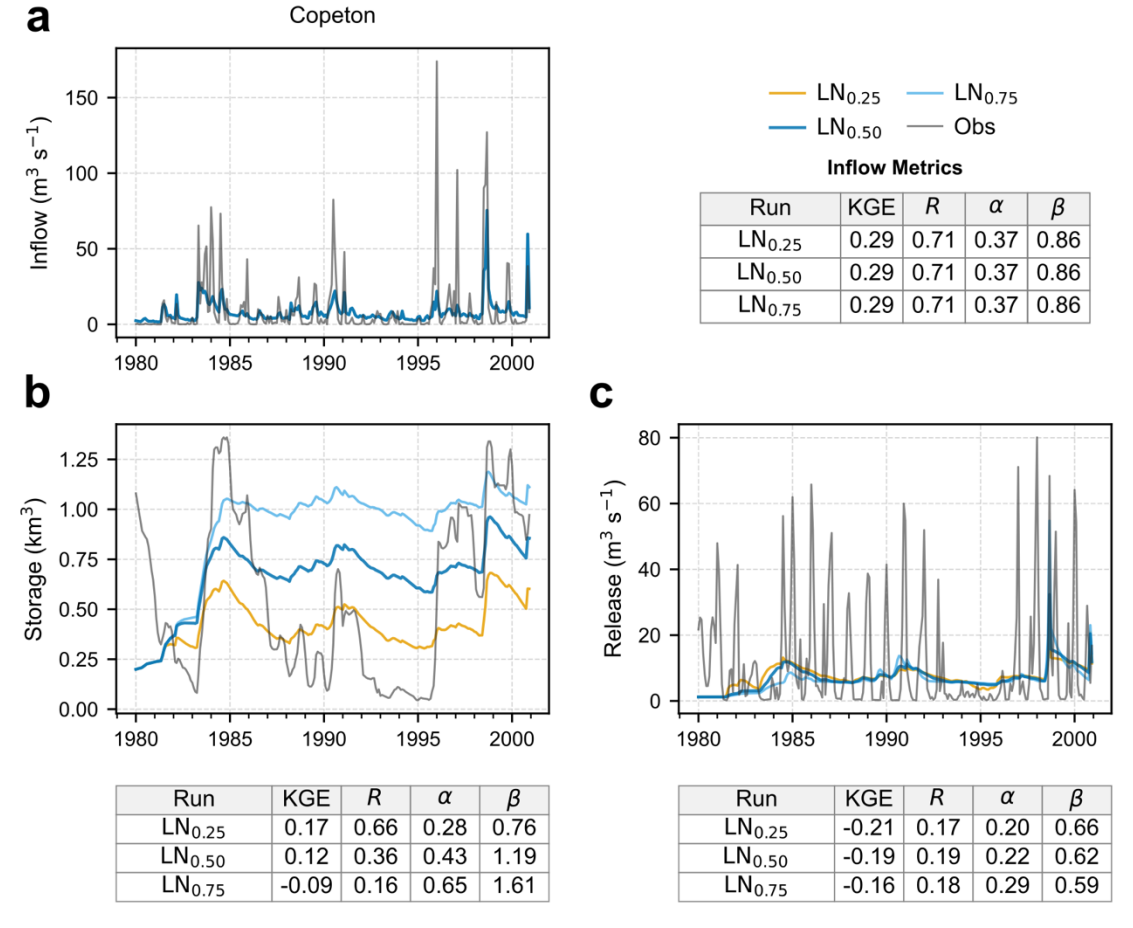


Inflow Metrics

Run	KGE	*R*	*α*	*β*
$LN_{0.25}$	0.59	0.82	0.66	0.86
$LN_{0.50}$	0.63	0.83	0.70	0.86
$LN_{0.75}$	0.71	0.83	0.81	0.86

Run	KGE	*R*	*α*	*β*
$LN_{0.25}$	0.44	0.68	1.25	0.62
$LN_{0.50}$	0.61	0.68	1.23	0.95
$LN_{0.75}$	0.45	0.54	1.02	1.30

Run	KGE	*R*	*α*	*β*
$LN_{0.25}$	0.33	0.59	0.48	0.91
$LN_{0.50}$	0.42	0.66	0.54	0.91
$LN_{0.75}$	0.54	0.65	0.72	0.91



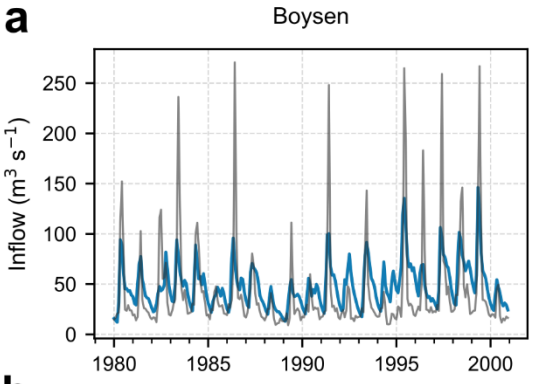
Inflow Metrics

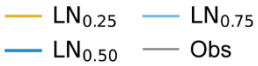
Run	KGE	*R*	*α*	*β*
$LN_{0.25}$	0.29	0.71	0.37	0.86
$LN_{0.50}$	0.29	0.71	0.37	0.86
$LN_{0.75}$	0.29	0.71	0.37	0.86

Run	KGE	*R*	*α*	*β*
$LN_{0.25}$	0.17	0.66	0.28	0.76
$LN_{0.50}$	0.12	0.36	0.43	1.19
$LN_{0.75}$	-0.09	0.16	0.65	1.61

Run	KGE	*R*	*α*	*β*
$LN_{0.25}$	-0.21	0.17	0.20	0.66
$LN_{0.50}$	-0.19	0.19	0.22	0.62
$LN_{0.75}$	-0.16	0.18	0.29	0.59

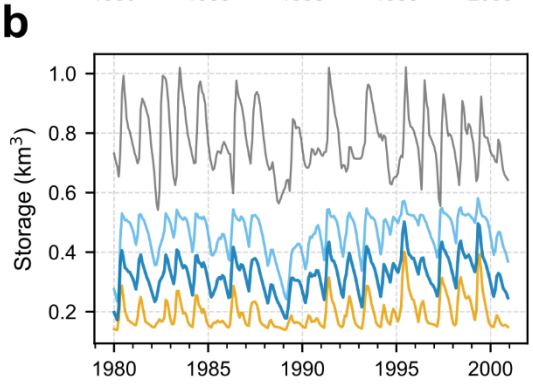
Figure S23 : Sensitivity analysis of LIS' LN parameter at Bull Shoals and Copeton dams. (a) Long-term monthly observed (*in situ*) inflow compared to simulations. (b) Comparison of observed (*in situ*) storage against storage simulated with different parameter sets. (c) Same as panel (b) but for release time series.



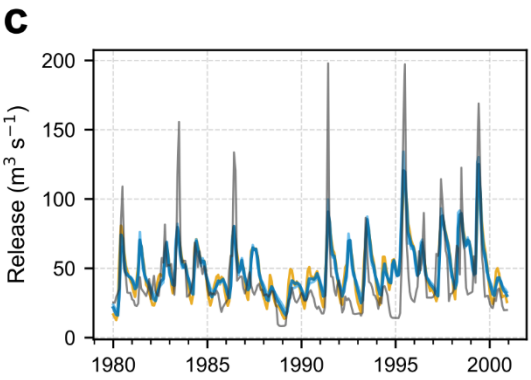


Inflow Metrics

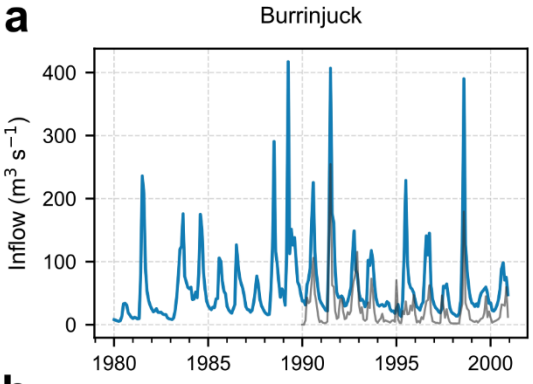
Run	KGE	*R*	*α*	*β*
LN$_{0.25}$	0.38	0.68	0.49	1.17
LN$_{0.50}$	0.38	0.68	0.49	1.17
LN$_{0.75}$	0.38	0.68	0.49	1.17

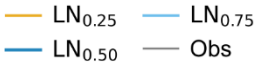


Run	KGE	*R*	*α*	*β*
LN$_{0.25}$	-0.09	0.40	0.49	0.25
LN$_{0.50}$	0.13	0.52	0.57	0.42
LN$_{0.75}$	0.30	0.54	0.65	0.60



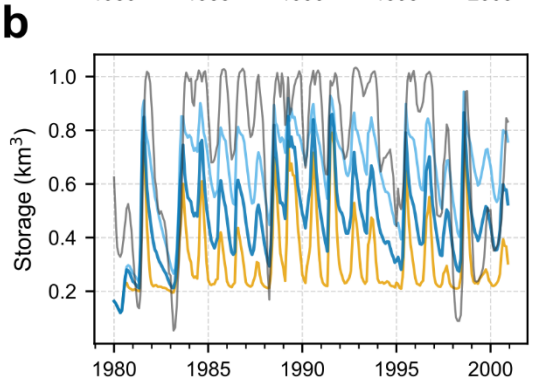
Run	KGE	*R*	*α*	*β*
LN$_{0.25}$	0.53	0.68	0.69	1.17
LN$_{0.50}$	0.53	0.73	0.65	1.17
LN$_{0.75}$	0.53	0.73	0.64	1.17



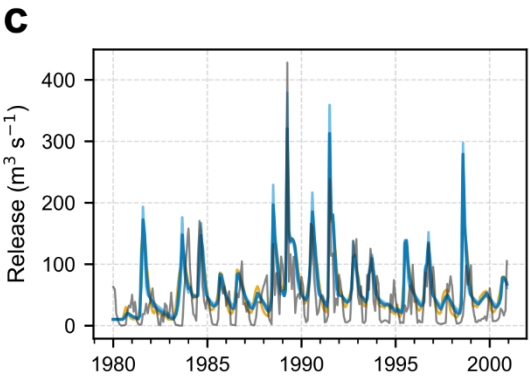


Inflow Metrics

Run	KGE	*R*	*α*	*β*
LN$_{0.25}$	-0.68	0.85	1.72	2.51
LN$_{0.50}$	-0.68	0.85	1.72	2.51
LN$_{0.75}$	-0.68	0.85	1.72	2.51



Run	KGE	*R*	*α*	*β*
LN$_{0.25}$	0.11	0.50	0.51	0.45
LN$_{0.50}$	0.36	0.68	0.56	0.66
LN$_{0.75}$	0.55	0.74	0.65	0.89



Run	KGE	*R*	*α*	*β*
LN$_{0.25}$	0.40	0.57	0.98	1.42
LN$_{0.50}$	0.44	0.62	1.00	1.42
LN$_{0.75}$	0.47	0.66	1.06	1.41

Figure S24: Same as Fig. S23 but for Boysen and Burrinjuck dams.

References.

Burek, Peter, Johan van Der Knifff, and Arie de Roo. 2013. *LISFLOOD - Distributed Water Balance and Flood Simulation Model - Revised User Manual 2013*. https://doi.org/10.2788/24719.

Burek, Peter, Yusuke Satoh, Taher Kahil, et al. 2020. "Development of the Community Water Model (CWatM v1.04) – a High-Resolution Hydrological Model for Global and Regional Assessment of Integrated Water Resources Management." *Geoscientific Model Development* 13 (7): 3267–98. https://doi.org/10.5194/gmd-13-3267-2020.

Hao, Zhen, Fang Chen, Xiaofeng Jia, et al. 2024. "GRDL: A New Global Reservoir Area-Storage-Depth Data Set Derived Through Deep Learning-Based Bathymetry Reconstruction." *Water Resources Research* 60 (1): e2023WR035781. https://doi.org/10.1029/2023WR035781.

Lehner, Bernhard, Penny Beames, Mark Mulligan, et al. 2024. "The Global Dam Watch Database of River Barrier and Reservoir Information for Large-Scale Applications." *Scientific Data* 11 (1): 1069. https://doi.org/10.1038/s41597-024-03752-9.

Li, Yao, Gang Zhao, George H. Allen, and Huilin Gao. 2023. "Diminishing Storage Returns of Reservoir Construction." *Nature Communications* 14 (1): 3203. https://doi.org/10.1038/s41467-023-38843-5.

Yigzaw, Wondmagegn, Hong-Yi Li, Yonas Demissie, et al. 2018. "A New Global Storage-Area-Depth Data Set for Modeling Reservoirs in Land Surface and Earth System Models." *Water Resources Research* 54 (12). https://doi.org/10.1029/2017WR022040.

Zhao, Gang, and Huilin Gao. 2018. "Automatic Correction of Contaminated Images for Assessment of Reservoir Surface Area Dynamics." *Geophysical Research Letters* 45 (12): 6092–99. https://doi.org/10.1029/2018GL078343.