Global water consumption impacts on riverine fish species richness in Life Cycle Assessment

Supplementary Material 1

Eleonore Pierrat[a], Valerio Barbarossa[b,c], Montserrat Núñez[d], Laura Scherer[b], Andreas Link[e], Mattia Damiani[f], Francesca Verones[g], Martin Dorber[g]

a Quantitative Sustainability Assessment division, Department of Environmental and Resource Engineering, Technical University of Denmark (DTU), 2800 Kgs. Lyngby, Denmark

b Institute of Environmental Sciences (CML), Leiden University, Leiden, The Netherlands

c PBL Netherlands Environmental Assessment Agency, The Hague, The Netherlands

d Sustainability in Biosystems, Institute of Agrifood Research and Technology (IRTA), Caldes de Montbui, Barcelona, Spain

e Chair of Sustainable Engineering, Technical University of Berlin, 10623 Berlin, Germany

f European Commission, Joint Research Centre, Via Enrico Fermi 2749, 21027 Ispra, VA, Italy

g Industrial Ecology Programme, Department of Energy and Process Engineering, NTNU, Høgskoleringen 5, 7491 Trondheim, Norway

* Corresponding author email: easpi@dtu.dk

Please cite this work as: Pierrat E., Barbarossa V., Núñez M., Scherer L., Link A., Damiani M, Verones F., Dorber M. *Global characterization factors for quantifying water consumption impacts on freshwater biodiversity in life cycle assessment - Support Material1*. Science of the Total Environment (2022): https://doi.org/10.1016/j.scitotenv.2022.158702.

<u>This document includes:</u>

- Supplementary Figures
- Supplementary Tables
- Supplementary Methods: Marginal and average effect factors calculation
- Supplementary References

The characterization factors are in the Supplementary Material 2 (Microsoft Excel file) and the river basin delineation in basins_5min_pcrglobwb.gpkg (Geopackage file).

1 Supplementary figures

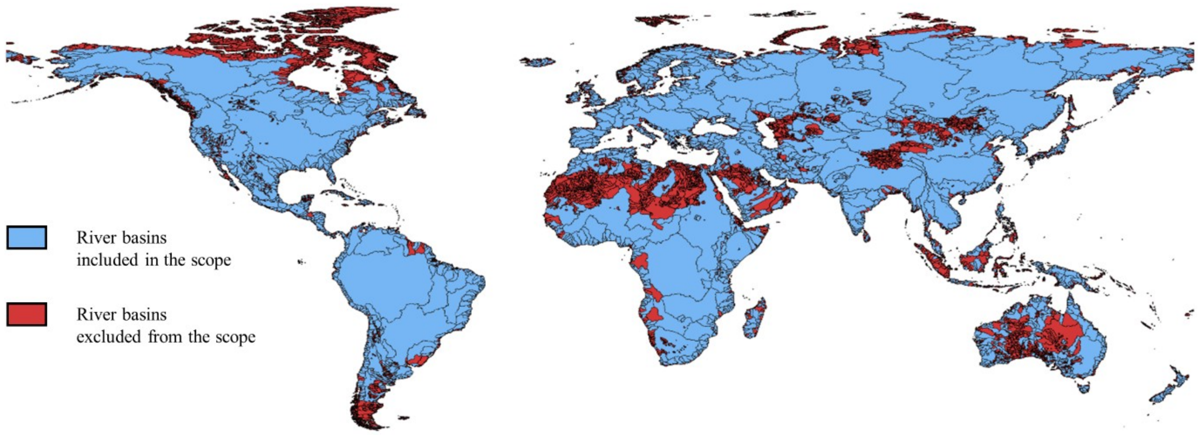

Fig S1: Map of the geographical scope of application of the SDR and resulting effect factors. Landmass covered: 88% (excluding Antarctica and Greenland).

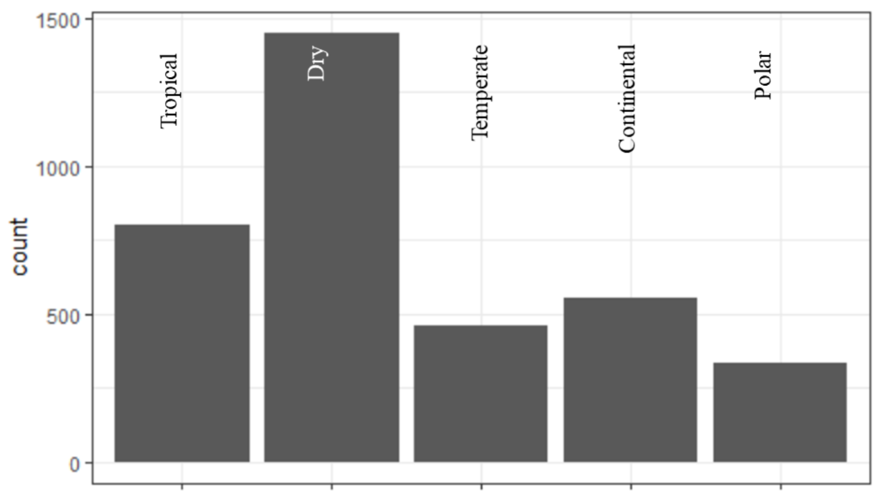

Fig S2 histogram of climate distribution (x-axis) at the basin scale (n=3,592).

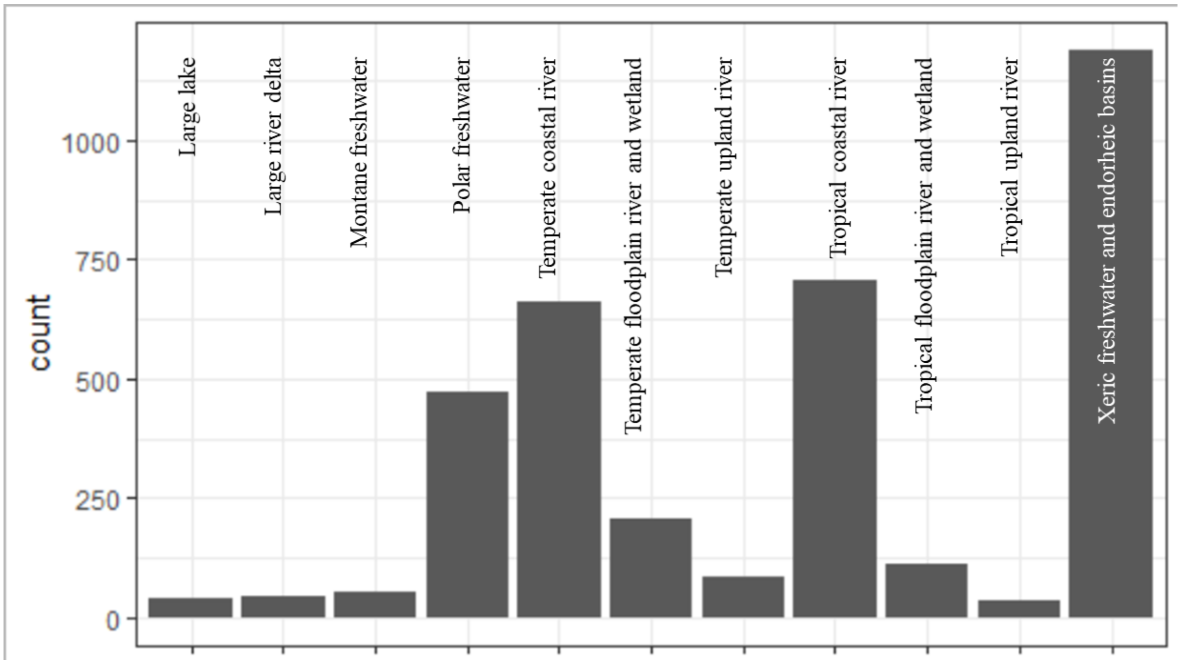

Fig S3 histogram of freshwater habitat type (x-axis) distribution at the basin scale (n=3,592)

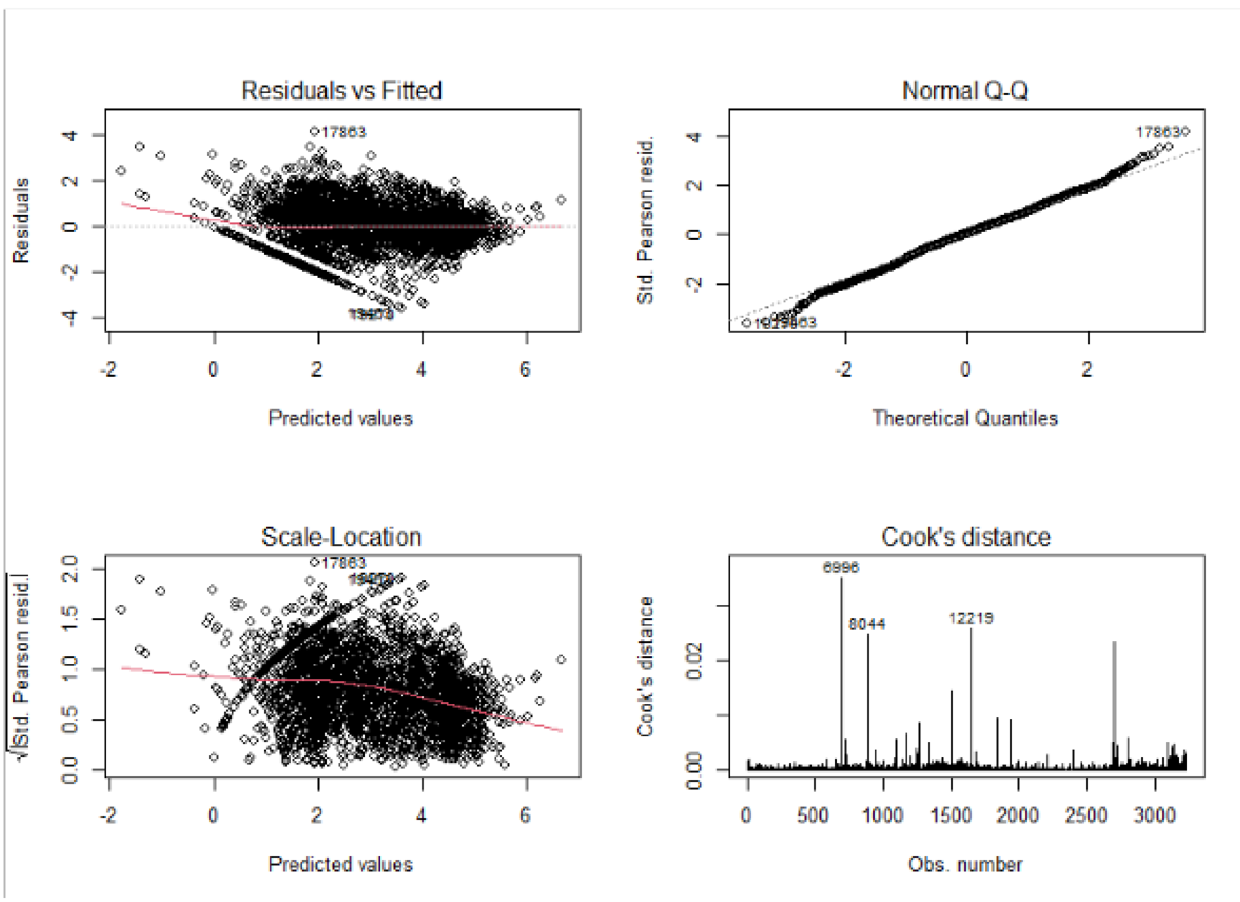

Fig S4: Residual plots of the selected species discharge model M4. The residual plot shows (top-left) shows that for smaller values, the simulated species richness is overestimated. Residual QQ

Plot (top-right) shows that the residuals follows reasonably the normal distribution. Scale-location plot (bottom left) trend is slightly decreasing (between 1 and 0.5) showing that residuals are not completely independent. Cook's distance plot indicates 3 outliers (basins 6996, 7965, 12219) that have a considerable influence on the overall regression.

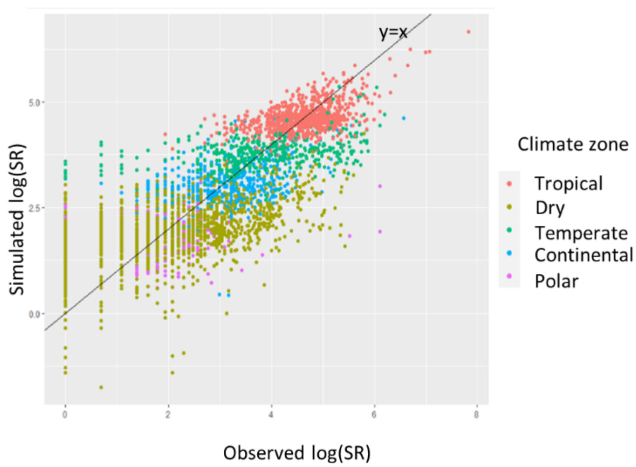

Fig S5: Observed log species richness versus simulated log species richness (y-axis) with the selected species discharge model.

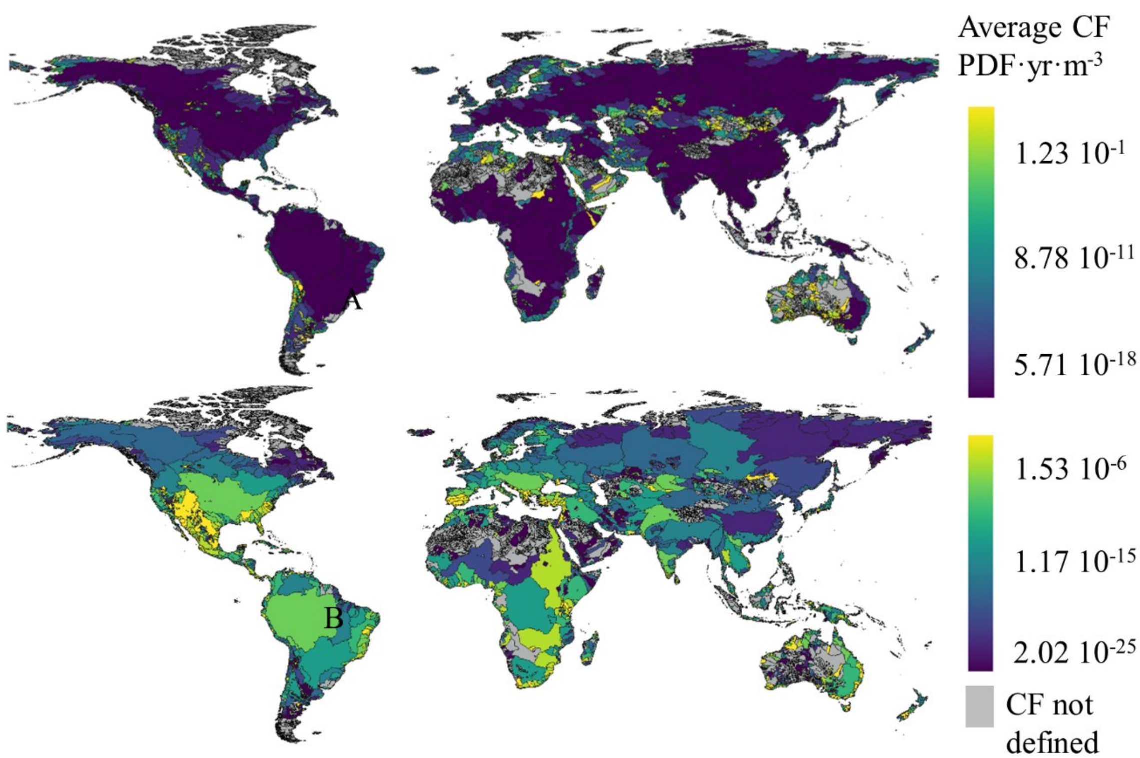

Fig. S6: Average characterization factors of water consumption impacts on local biodiversity (A) and global biodiversity (B).

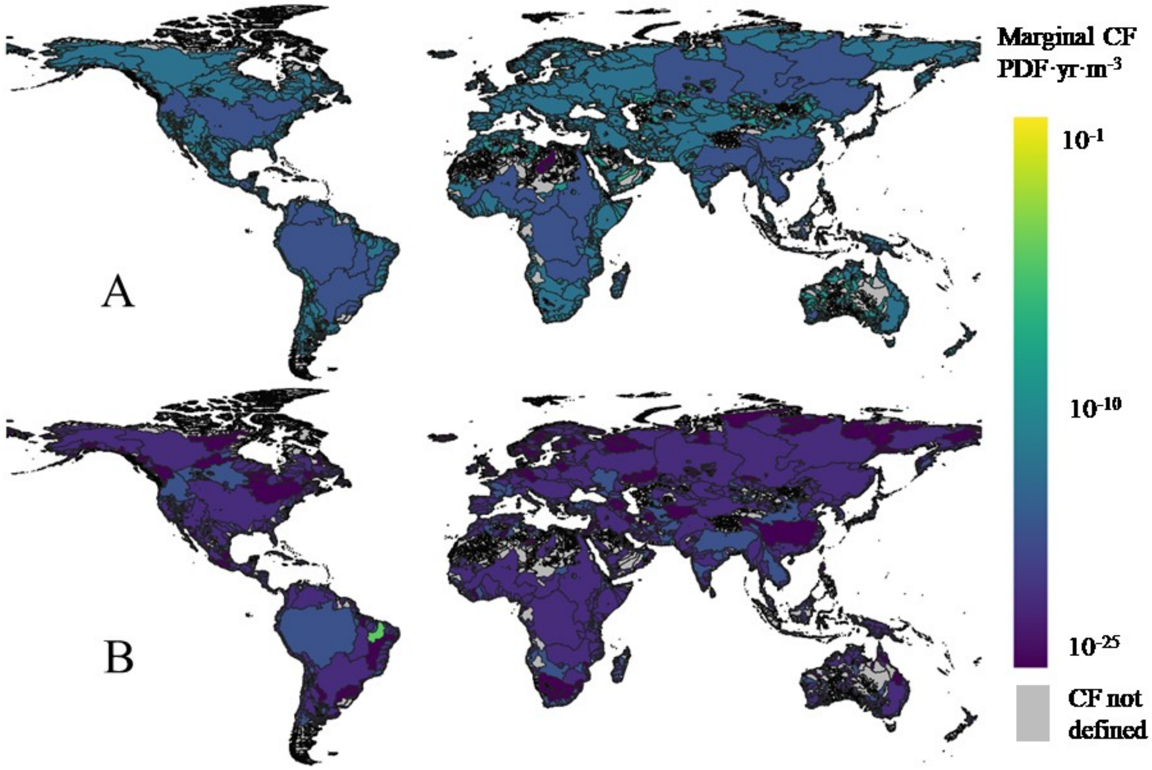

Fig. S7: Marginal characterization factors of water consumption impacts on local biodiversity (A) and global biodiversity (B) in log scale.

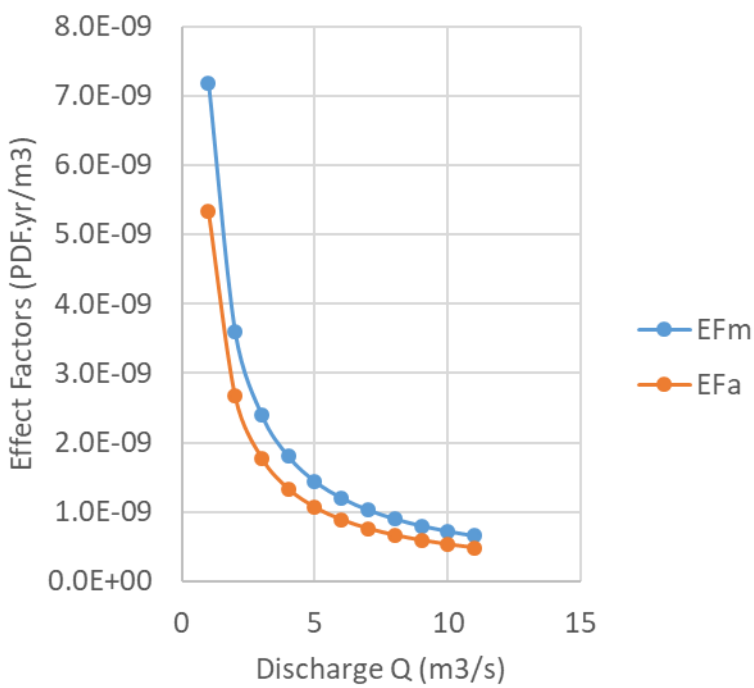

Figure S8: Marginal (EFm) and average (EFa) effect factors dependence on discharge with a fixed level of streamflow depletion (Q/Q0=0.6).

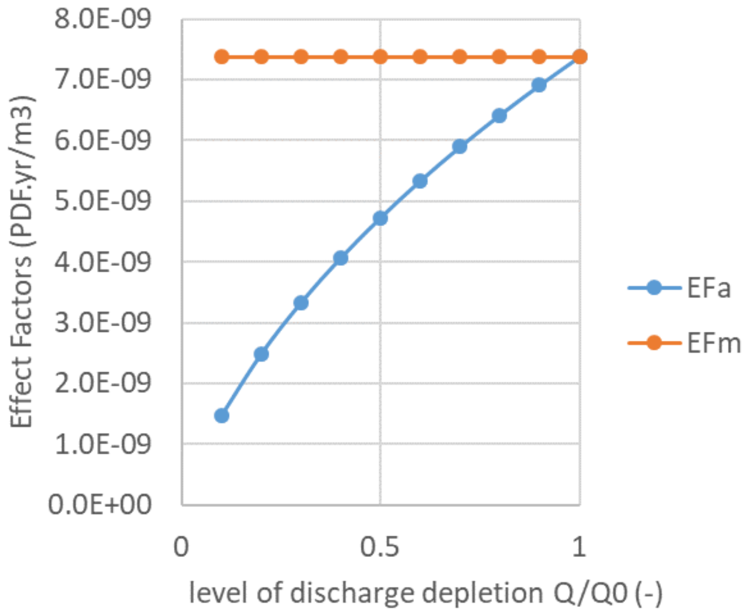

Figure S9: Marginal (EFm) and average (EFa) effect factors in a given basin (Q=1 m^3/s) with different levels of discharge depletion (Q/Q0 varies).

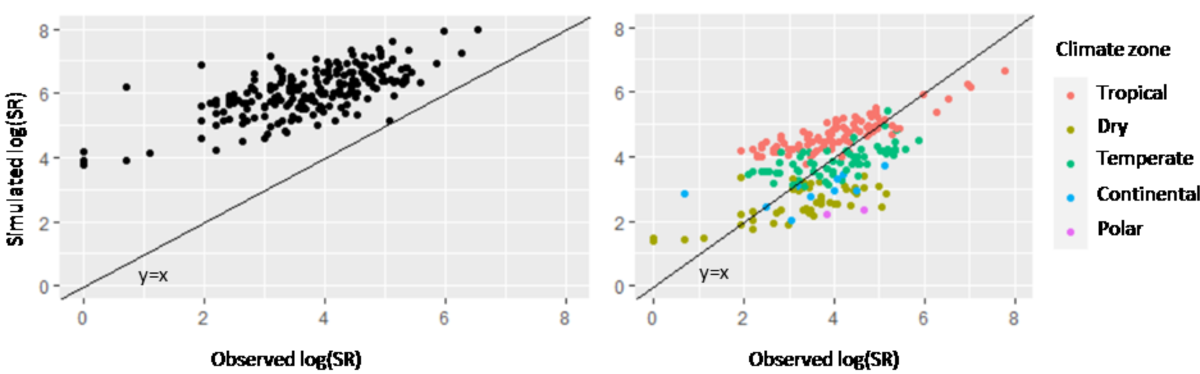

Figure S10: Goodness of fit of Xenopoulos' SDR (left) and the selected model (right) tested with observed discharge data (GSIM).

2 Supplementary tables

Table S1: descriptive variables for river basins (3,592) used to build species-discharge relationships. Statistical values are calculated for basins with valid data, area>2550 km^2 (Q80), positive discharge flow and species richness >=1.

Variable	Unit	Description	Min med max	Ref
SR	Unit	Species richness	1;18;2521	Barbarossa et al. 2021
P	Mm/yr	Present mean annual precipitation 1970-2000	0;500;6344	World clim 2 Fick, S.E. and R.J. Hijmans, 2017.

T	K	Present mean annual temperature 1970-2000	201;291;303	World clim 2 Fick, S.E. and R.J. Hijmans, 2017.
Pdelta	%	Precipitation change since the last glacial maximum (avg 3 models CCSM4, MIROC-ESM, MPI-ESM-P)	-0.72;0.12;5.02	World clim 1.4 Karger, et al. (2017)
Tdelta	%	Temperature change since the last glacial maximum (average of 3 models CCSM4, MIROC-ESM, MPI-ESM-P)	-0.016;0.016;0.156	World clim 1.4 Karger, et al. (2017)
Qmean	m^3/s	Mean annual discharge at the river mouth	6e-8; 3.59e1;1.60e5	De Graaf et al. 2019
Q0	m3/s	Simulated mean annual discharge at the river mouth in the natural scenario i.e., without dams and water ocnsumption	0; 3.72e1;1.60e5	De Graaf et al. 2019
Ti	1	Log(area/tan(slope))	6.01;37.53;43.89	De Graaf et al. 2019
Elevation	m	DEM above sea level	0.0 249 5318	De Graaf et al. 2019
Area	km^2	Area	2550 5897 5912646	De Graaf et al. 2019
Climate5	Label	Kopper Geiger climate classification	1 tropical 2 dry 3 temperate 4 continental 5 polar	Beck et al 2018
Realm	Label	Zoogeographic realms	1. Neartic 2. Oceania 3. Indo-Malay 4. Australasia 5. Afrotropic 6. Palartic 7. Neotropic	Abel et al 2008
habitat	Label	Freshwater habitat types	1. Large lake 2. Large river delta 3. Montane freshwater 4. Oceanic island 5. Polar freshwater 6. Temperate coastal river 7. Temperate floodplain river and wetland 8. Temperate upland river 9. Tropical coastal river 10. Tropical floodplain river and wetland 11. Tropical upland river	Abel et al 2008

			12. Xeric freshwater and endorheic basins	
id_basin_pcrglob	Label	Basin identification based on pcrglob wb basin delineation		

Table S2: Spearman-rank coefficients across pairs of continuous variables of the dataset (n=3,592). Autocorrelation between pairs (highlighted in red) were considered too high for coefficients > 0.7.

spearman on centered and scaled data	logQ	prec	temp	prec_delta	temp_delta	ti	slope	elevation	logA
logQ	1.00								
prec	0.75	1.00							
temp	0.03	0.29	1.00						
prec_delta	0.05	-0.08	-0.43	1.00					
temp_delta	-0.01	-0.31	-0.78	0.41	1.00				
ti	-0.06	0.14	0.54	-0.14	-0.39	1.00			
slope	0.15	0.00	-0.30	0.00	0.16	-0.90	1.00		
elevation	-0.16	-0.37	-0.27	-0.02	0.16	-0.57	0.64	1.00	
logA	0.40	0.00	0.03	0.00	0.03	0.04	0.00	0.22	1.00

Table S3-: Cross-validation mean and standard deviation for the goodness of fit indicators (10 folds).

mean	BIC	KGE	CC	BR	RV	RMSE	R2
M5	10113	0.71	0.80	1.00	0.80	0.97	0.63
M4	10252	0.69	0.78	1.00	0.78	1.00	0.61
M3	10424	0.68	0.78	1.00	0.78	1.01	0.60
M2	10723	0.64	0.75	1.00	0.75	1.07	0.56
M1	11750	0.48	0.64	1.00	0.64	1.24	0.41

sd	BIC	KGE	CC	BR	RV	RMSE	R2
M1	0.0000	0.0359	0.0367	0.0224	0.0410	0.0559	0.0459
M2	0.0000	0.0342	0.0336	0.0199	0.0364	0.0681	0.0495
M3	0.0000	0.0236	0.0226	0.0220	0.0269	0.0520	0.0345
M4	0.0000	0.0223	0.0171	0.0211	0.0290	0.0423	0.0264
M5	0.0000	0.0257	0.0260	0.0187	0.0232	0.0606	0.0407

Table S4 : Equation of best model for each model candidate resulting from the cross-validation

Model	Independent variable	Predictors selected	Random component	Link
1	log(SR_tot)	logq	gaussian	identity
2	log(SR_tot)	elevation + logA + logq + prec_delta + temp + 1	gaussian	identity

3	log(SR_tot)	habitat.f+ logA+log1	gaussian	identity
4	log(SR_tot)	climate5.f+elevation +logA+ logq	gaussian	identity
5	log(SR_tot)	habitat.f + logq + prec_delta + temp + 1	gaussian	identity

Table S5: Mean goodness of fit indicators comparing the performance of fitted regressions and Xenopoulos' SDR for observed streamflow taken from GSIM dataset and elevation from the GSGM digital elevation model. A total of 216 basins fulfilled the conditions for all regression models (area>2550 km2, latitude within 42S and 42N, average annual streamflow non null).

	KGE	CC	BR	RV	rmse	R2
M4	0.564	0.59	1.018	0.852	0.99	0.348
M1	0.561	0.71	0.916	0.681	0.882	0.504
Xenopoulos	0.134	0.71	1.591	0.437	2.406	0.504

Table S6: Moran I test results for spatial autocorrelation of residuals. Pairs of residuals are weighted by the inverse of the distance between basin centroids. The residual show significant spatial autocorrelation (p-val =0) however this autocorrelation is small as Moran I coefficient is << 1.

	observed	expected	sd	p,value
Moran_I	2,31E-02	-2,78E-04	9,37E-04	0,00E+00

Table S7: distribution of average and marginal effect factors (n=3592 basins).

	EF_marginal	EF_average
Minimum	3.93E-14	3.93E-14
Percentile 25%	4.14E-11	3.79E-11
Percentile 50%	1.73E-10	1.37E-10
Percentile 75%	1.18E-09	7.61E-10
Maximum	1.00E-01	1.41E-04

Table S8: Comparison of the marginal characterization factors to Hanafiah et al. (2011) for emblematic rivers of the world.

River	river volume (m^3) (Hanafiah et al. 2011)	CF ($PDF.yr.m^3.m^{-3}$) (Hanafiah et al. 2011)	CF ($PDF.yr.m^{-3}$) (Hanafiah et al. 2011)	New EF marginal ($PDF.yr.m^{-3}$)	New/Hanafiah-1
Ganges (id:17313)	2.65E+09	2.41E-03	9.09E-13	1.94E-13	-0.79
Nile (id:22098)	1.60E+09	8.42E-03	5.26E-12	9.05E-13	-0.83
Yangtze (id:16558)	1.24E+10	5.19E-03	4.19E-13	2.20E-13	-0.47
Euphrates (id:14990)	1.84E+09	3.75E-03	2.04E-12	3.22E-12	0.58
Amazon (id:24212)	6.23E+10	3.90E-03	6.26E-14	3.93E-14	-0.37
Niger (id:20745)	2.06E+09	5.59E-03	2.71E-12	5.18E-13	-0.81

Table S9: Dominance analysis matrix of selected model M4. Average contribution of each variable is computed using four different coefficient of determinations.

indicator	climate5	elevation	logA	logq	total
r2m	14%	1%	1%	9%	26%
r2cs	33%	4%	3%	23%	62%
r2n	33%	4%	3%	23%	62%
r2e	36%	4%	3%	25%	67%

Table S10: Dispersion of characterization factors and their components

	Relative interquartile range
FF	0.6
$EF_{marginal}$	5.0
$EF_{average}$	7.8
GEP	7.9
$CF_{reg,marginal}$	4.1
$CF_{reg,average}$	4.2
$CF_{glo,marginal}$	5.3
$CF_{glo,average}$	5.3

Table S11: Spearman rank correlation coefficients between characterization factors and their components

	FF	EF	GEP
$CF_{reg,marginal}$	0.25	0.90	NA
$CF_{reg,average}$	0.24	0.90	NA
$CF_{glo,marginal}$	0.16	-0.07	0.71
$CF_{glo,average}$	0.15	-0.05	0.70

Table S12: Results of the rice case study for the three methodologies tested.

	LCI	CF			Impact		
Basin, Country and identifier of the basin in PCRGLOB-WB2	Water Consumption (m3/kg rice)	New, reg (PDF reg yr/m3)	Hanafiah (PDF reg m3 yr/m3)	New, glob (PDF glob yr/m3)	New, reg (PDF reg yr/kg rice)	Hanafiah (PDF reg yr/kg rice)	New, glob (PDF glob yr/kg rice)
Ganges, India (id:17313)	8.26E-01	3.39E-14	2.67E-03	4.07E-16	6.69E-14	2.21E-03	8.03E-16
Godavari, India (id:19039)	8.26E-01	3.87E-13	1.31E-03	3.10E-16	7.63E-13	1.08E-03	6.11E-16
Yellow River, China (id:13884)	4.87E-01	6.23E-13	6.13E-03	1.55E-16	1.23E-12	2.99E-03	3.06E-16
Pearl River, China (id:17644)	4.87E-01	1.00E-13	2.12E-03	4.83E-16	1.97E-13	1.03E-03	9.54E-16

Red River, US (id: 14842)	8.35E-01	1.37E-13	2.55E-03	2.51E-15	2.71E-13	2.13E-03	4.95E-15
Arkansas River, US (id:14842)	8.35E-01	1.37E-13	2.75E-03	2.51E-15	2.71E-13	2.30E-03	4.95E-15

3 Marginal and average effect factors calculation

- Species-discharge relationship

Model M4 equation related species richness to discharge flow at the river mouth:

Equation S1

$$\log(SR) = a \cdot log(Q) - b \cdot E + c \cdot \log(A) + CZ$$

Where:

- log(SR) is the natural logarithm transformation of the fish species richness SR (-) in the river basin,
- log(Q) is the natural logarithm transformation of the mean annual discharge at the river mouth (m3/yr) in the river basin,
- log(A) is the natural logarithm transformation of the river basin area A (m2)
- E (m above sea level) is the average elevation of the basin
- CZ is the climate zone intercept (tropical, temperate, arid, cold, polar) provided in Table 2 of the main manuscript
- a is the regression coefficient of logQ provided in Table 2 of the main manuscript
- b is the regression coefficient of elevation provided in Table 2 of the main manuscript

Therefore:

Equation S2

$$SR = \exp(a \cdot log(Q) - b \cdot E + c \cdot \log(A) + CZ)$$
$$= \exp(a \cdot log(Q)) \cdot \exp(c \cdot \log(A) \, \cancel{CZ}) \cdot \exp(CZ) \cdot \exp(-b \cdot E) = d \cdot Q^a$$

Where: $d = A^c \cdot \frac{exp(CZ)}{\exp(b \cdot E)}$

We define the function SDR(Q) the species-discharge relationship obtained in Equation S2 that allows estimating species richness:

Equation S3

$$SDR(Q) = d \cdot Q^a$$

Therefore, the SDR is regionalized because the coefficient *d* is specific for each river basin and depends on its elevation, area, and climate zone. The loss of species resulting from discharge flow reduction is estimated deriving SDR(Q) in Equation S3:

Equation S4

$$\frac{dSDR(Q)}{dQ} = d \cdot a \cdot Q^{a-1}$$

- Marginal effect factor

Following the definition of the marginal effect factor given in eq. 5:

Equation S5

$$EF_{marginal} = \frac{1}{SR} \cdot \frac{dSR}{dQ} = \frac{1}{SDR(Q)} \cdot \frac{dSDR(Q)}{dQ}$$

Where $EF_{marginal}$ refers to the potentially disappearing fraction of fish species per m^3 water consumed times year ($PDF \cdot yr \cdot m^{-3}$). The unit of this proposed effect factor ($PDF \cdot yr \cdot m^{-3}$) is the same as proposed by Verones et al. (2017). $EF_{marginal}$ can be described through the variables SR, dSR and dQ, where SR refers to the fish species richness (-), dSR (-) to the marginal loss of fish species and dQ to the discharge loss at the river mouth ($m^3 \cdot yr^{-1}$). Furthermore, it is noteworthy that eq.3 is only applicable for $dQ < 0$, i.e. when water consumption causes river discharge depletion.

Substituting in eq.3 in the main manuscript with Equation S3 and Equation S4 we obtain Equation S6:

Equation S6

$$EF_{marginal} = \frac{1}{d \cdot Q^a} \cdot d \cdot a \cdot Q^{a-1} = \frac{a}{Q}$$

- Average effect factor

Following the definition of the average effect factor given in eq. 6 in the main manuscript:

Equation S7

$$EF_{average} = \frac{1}{Q - Q_0} \cdot \frac{SR - SR_0}{SR_0}$$

Where Q_0 ($m^3 \cdot yr^{-1}$) and SR_0 (-) are the annual discharge flow at the river mouth and species richness if there were no water consumption in the basin (natural simulation, see Section 2.2), Q ($m^3 \cdot yr^{-1}$) and SR (-) are their respective counterparts when human consumption is considered.

We estimate SR and SR_0 using SDR(Q) provided in Equation S3, then we substitute in eq. 4 in the main manuscript (Equation S8):

Equation S8

$$EF_{average} = \frac{1}{Q - Q_0} \cdot \frac{d \cdot Q^a - d \cdot Q_0^a}{d \cdot Q_0^a} = \frac{1}{Q - Q_0} \cdot \left(\left(\frac{Q}{Q_0} \right)^a - 1 \right)$$

4 Supplementary references

Fick, S.E. and R.J. Hijmans, 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37 (12): 4302-4315.

de Graaf, I.E.M., Gleeson, T., (Rens) van Beek, L.P.H., Sutanudjaja, E.H., Bierkens, M.F.P., 2019. Environmental flow limits to global groundwater pumping. Nature 574, 90–94. https://doi.org/10.1038/s41586-019-1594-4

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P. & Kessler, M. (2017) Climatologies at high resolution for the earth's land surface areas. Scientific Data 4, 170122. https://www.worldclim.org/data/v1.4/paleo1.4.html