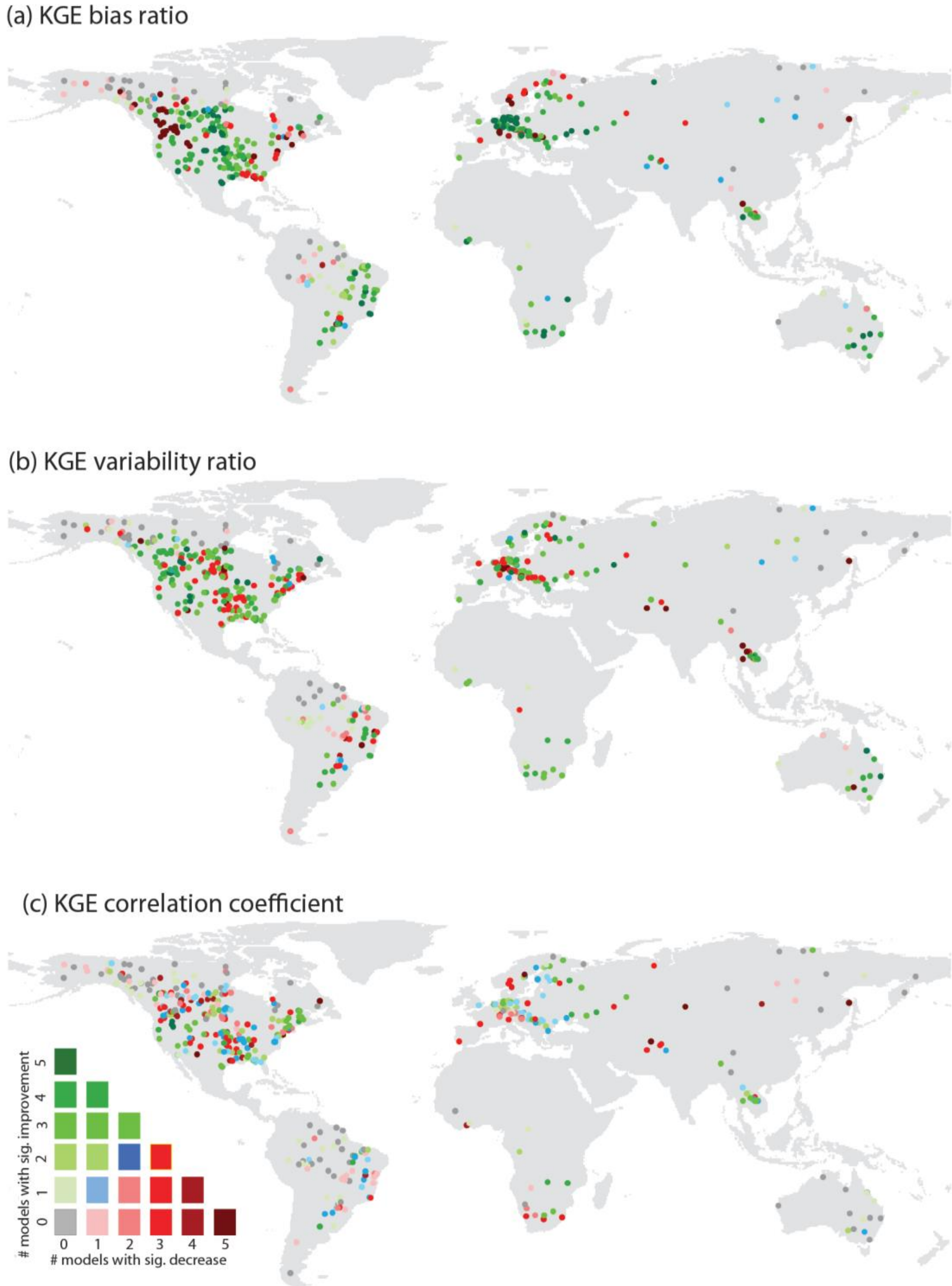
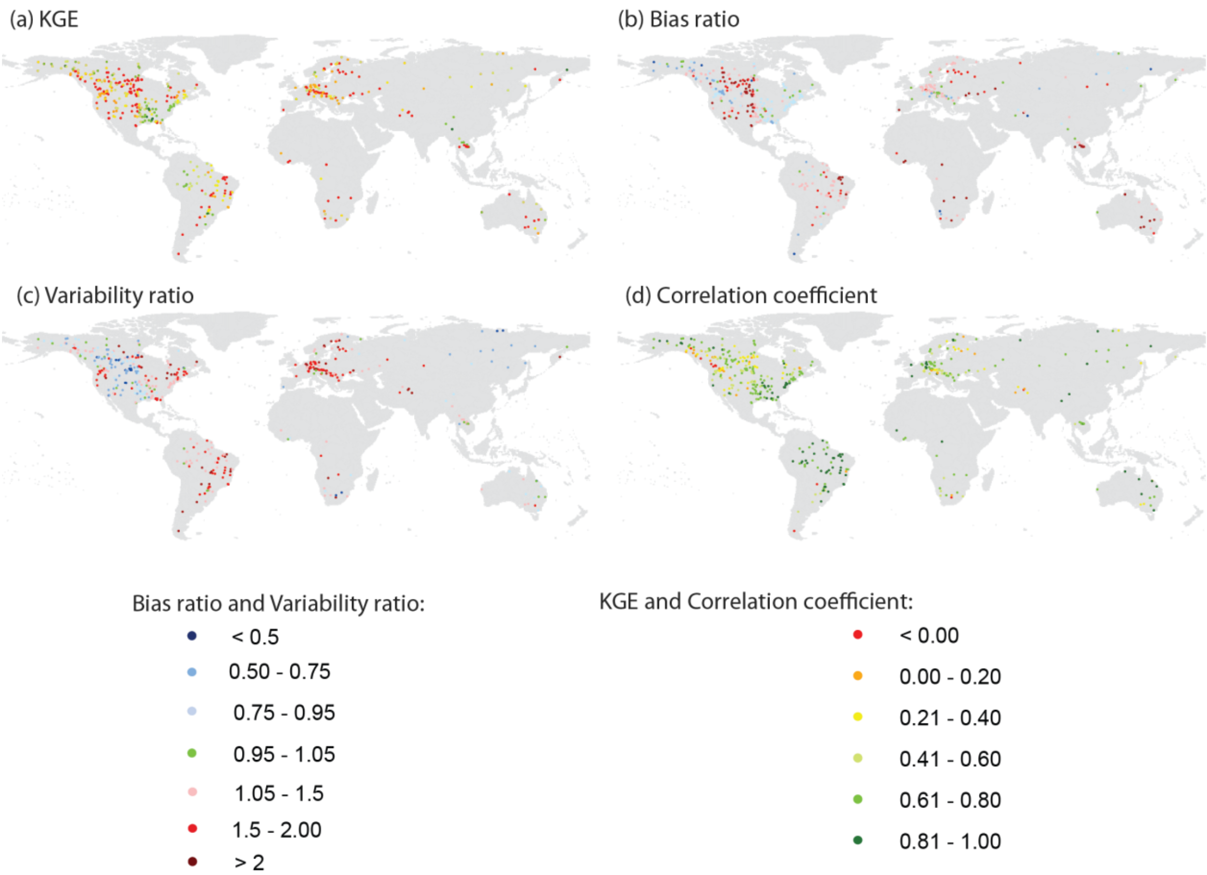
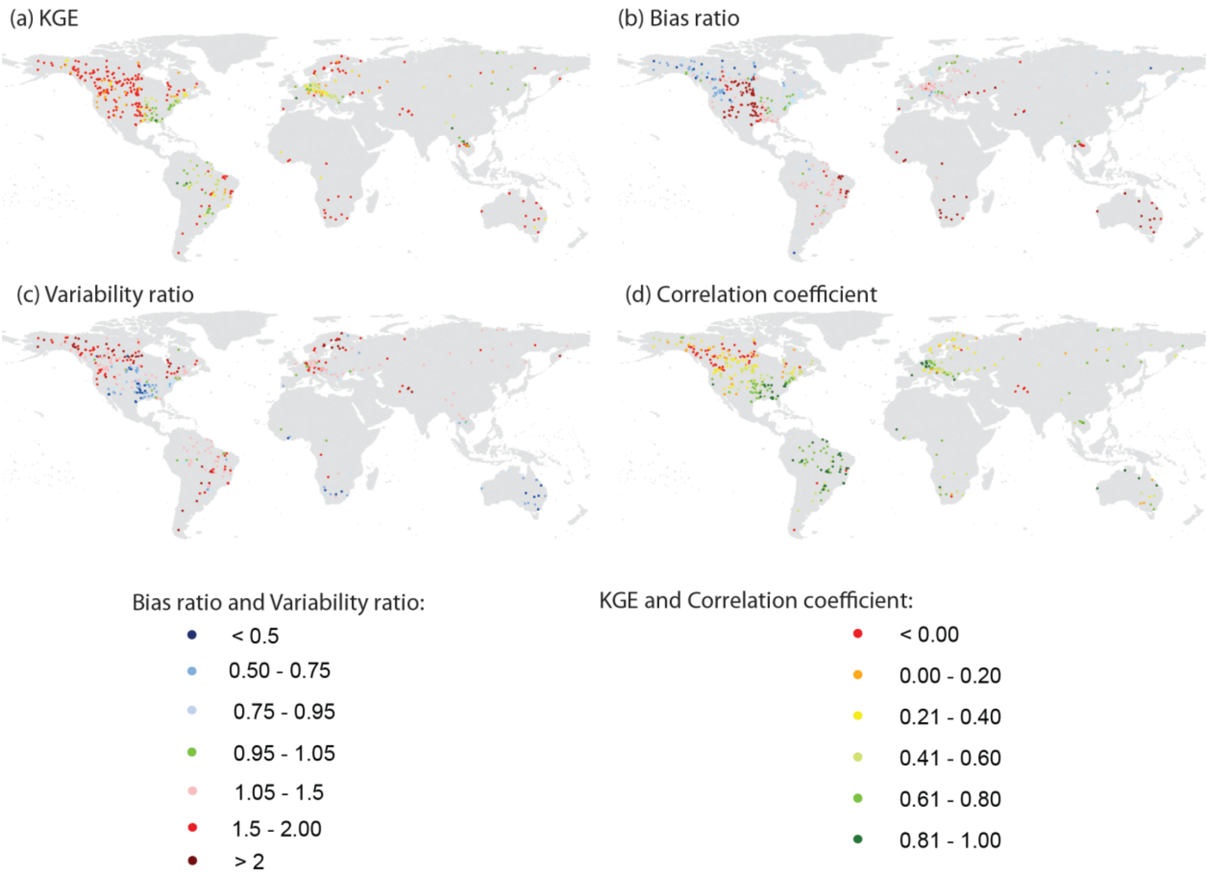
Supplementary Figures



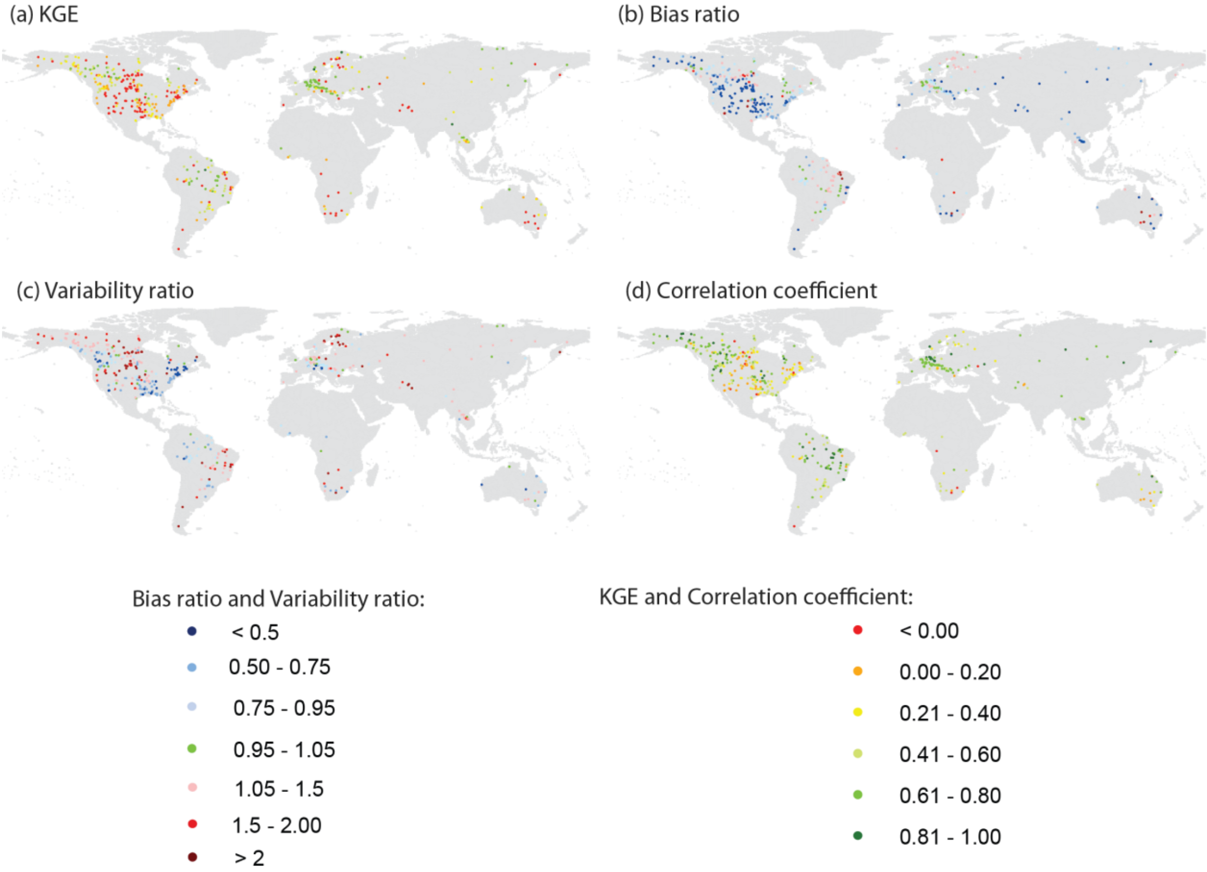
Supplementary figure 1: Number of GHMs with a significant increase or decrease in hydrological performance (KGE) due to HIP. Figures (a), (b), and (c) show for each of the underlying KGE sub-parameters (bias ratio, variability ratio, correlation coefficient) the number of models with a significant increase or decrease in performance.



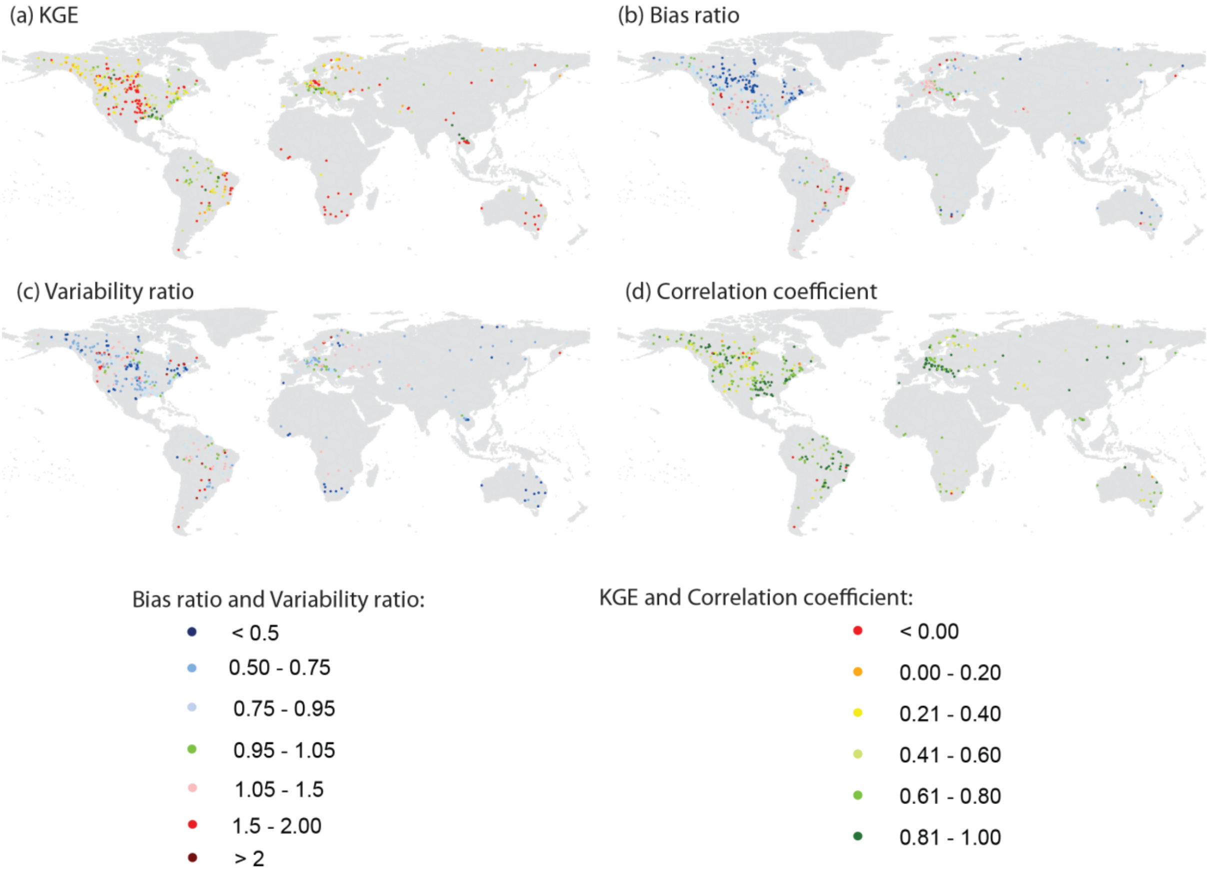
Supplementary figure 2a: KGE performance of H08 under HIP conditions.



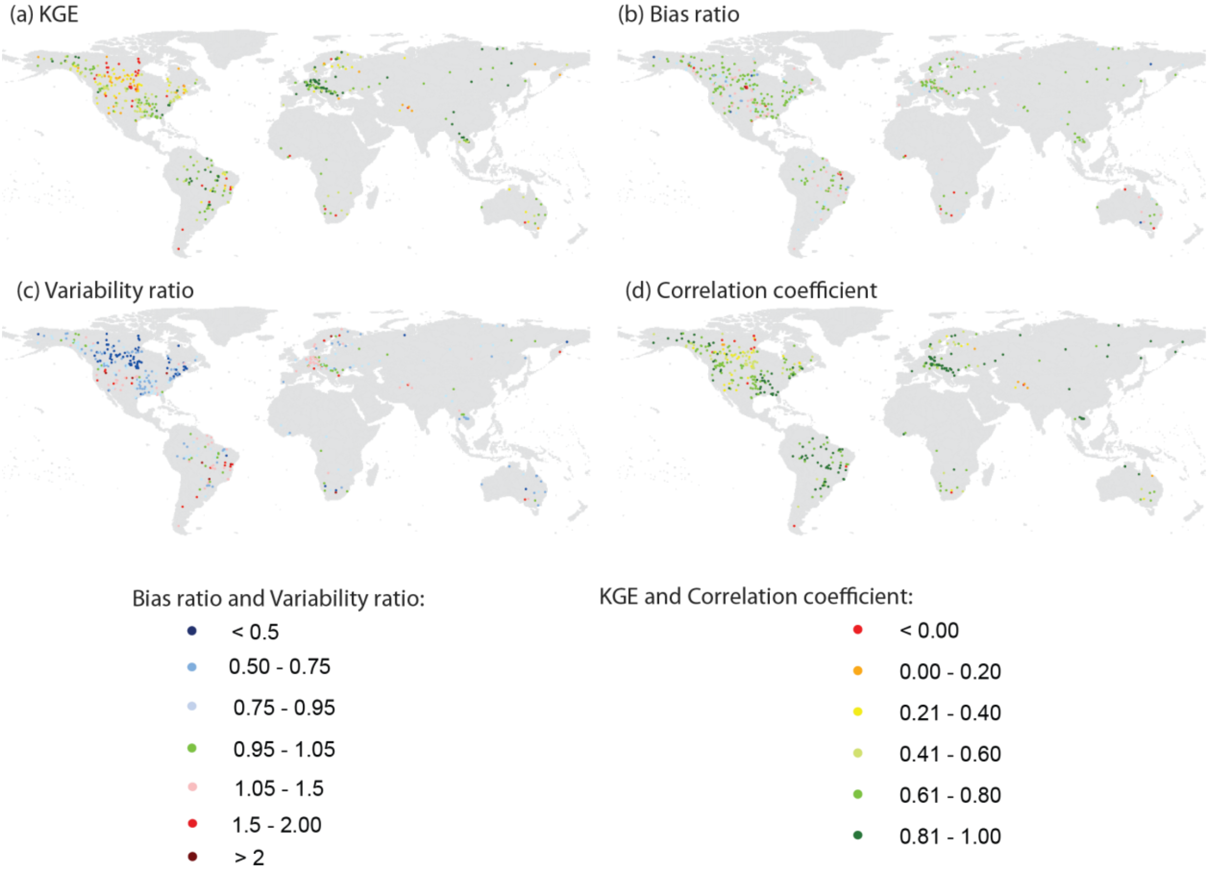
Supplementary figure 2b: KGE performance of LPJmL under HIP conditions.



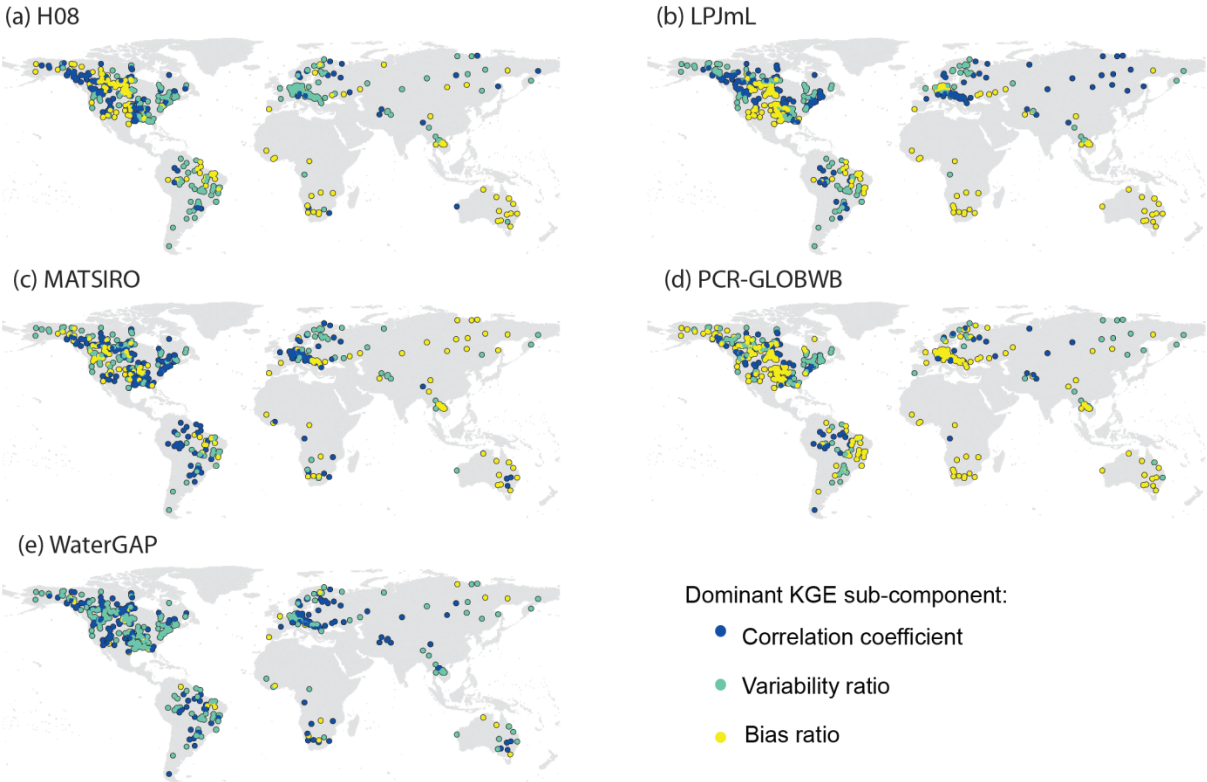
Supplementary figure 2c: KGE performance of MATSIRO under HIP conditions.



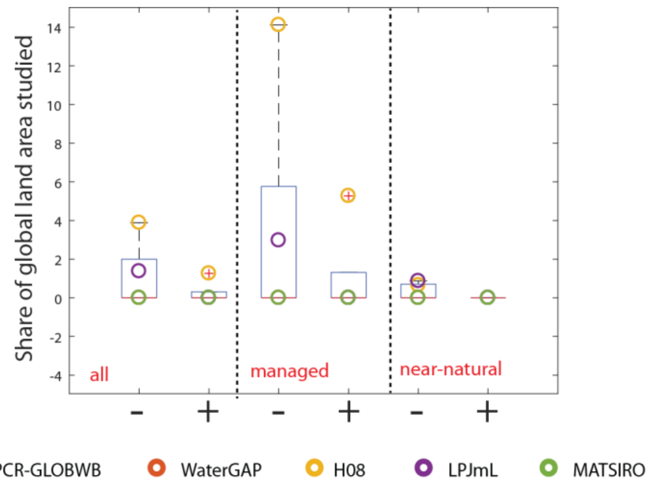
Supplementary figure 2d: KGE performance of PCR-GLOBWB under HIP conditions.



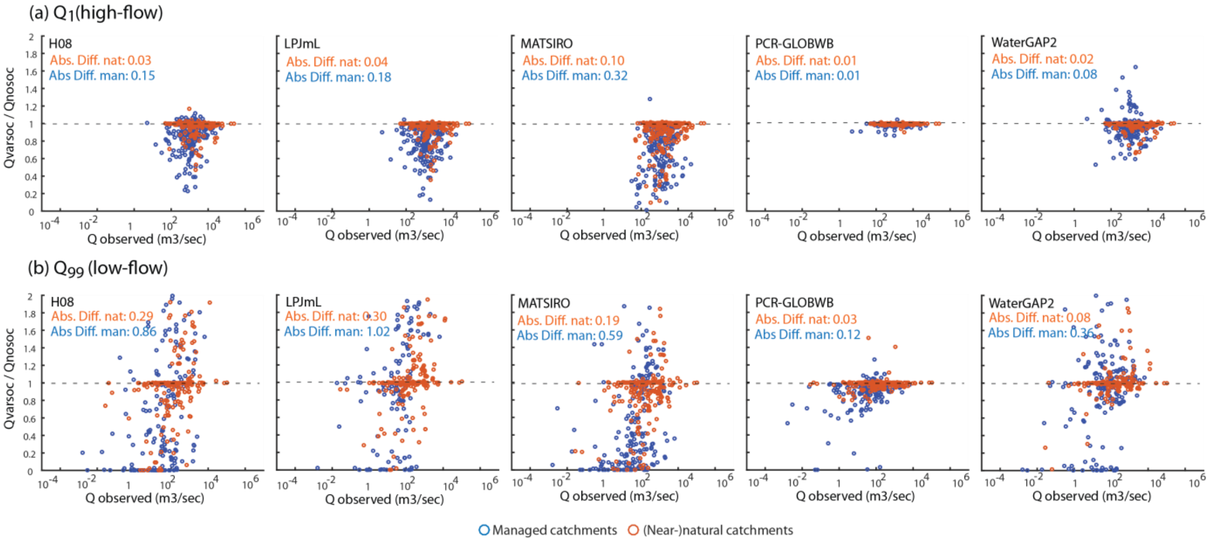
Supplementary figure 2e: KGE performance of WaterGAP2 under HIP conditions.



Supplementary figure 3: Spatial distribution of dominant KGE sub-component limiting optimal hydrological performance.



Supplementary figure 4: Share of land area with a significant change in the representation of the exceedance probability curve, as tested with the Kolmogorov-Smirnov test. WaterGAP2 and PCR-GLOBWB are located underneath the results of MATSIRO.



Supplementary figure 5: Scatterplot showing for the individual global hydrological models (GHMs) the difference in hydrological extremes due to HIP.

The two sets of panels show the difference (factor) in modelled (a) Q_1 high-flow and (b) Q_{99} low-flow discharges under the HIP and NOHIP conditions for the managed (blue) and near-natural catchments (orange). Values >1 imply that discharges are larger under HIP conditions compared to NOHIP conditions, and vice versa. To improve visibility, results are capped at a value of 2. The absolute differences given in each of the panels represent the average absolute deviation from unity for the managed and near-natural catchments. The larger the average absolute deviation, the larger the difference in modelled Q_1 high-flow and Q_{99} low-flow discharges between HIP and NOHIP conditions.

Supplementary Tables

Supplementary table 1: Description of the modelling framework and parameterizations of the models. Table 1.A presents the representation of the hydrological processes in the models, after Zhao et al., 2017. Table 1.B-I present the representation of the human dimensions and its interactions with the hydrological processes for each of the models.

1.A Representation of hydrological processes

Model name	Energy balance	Evaporation scheme	Runoff scheme	Snow scheme	Routing scheme	Flow velocity	Soil Water Layer Depth	References
H08	Yes	Bulk formula	Saturation excess, non-linear	Energy balance	TRIP Linear reservoir (Oki and Sud 1998)	0.5m/s	One soil layer with a depth of 1m	Hanasaki et al., 2008a,b
LPJmL	No	Priestley-Taylor	Saturation excess	Degree-day	Continuity equation derived from linear reservoir model	1 m/s	Five hydrologically active layers of 20, 30, 50, 100 and 100 cm thickness, respectively	Bondeau et al., 2007; Schaphoff et al., 2013
MATSIRO	Yes	Bulk formula	Overland flow, infiltration excess, saturation excess, groundwater.	Energy balance	TRIP Linear reservoir (Oki and Sud 1998)	0.5m/s	12 fully resolved layers (5cm, 20cm, 75cm, and nine next layers of 1m) and a 90m groundwater layer.	Takata et al., 2003; Pokhrel et al., 2012; 2015
PCR-GLOBWB	No	Hamon	Saturation Excess Beta Function	Degree Day	Travel time routing (characteristic distance)	Variable based on Manning's equation	Variable up to 1.5 m soil layers and 50 m groundwater layer	van Beek et al., 2011; Wada et al., 2011
WaterGAP2	No	Priestley Taylor with two alpha factors depending on the aridity of the grid cell	Beta function, saturation excess	Degree Day	Linear reservoir	Variable, based on Manning-Strickler	One soil layer, varying depth in dependence on land cover type (0.1 to 4 m)	Müller Schmied et al., 2014,2016; Verzano et al., 2012

1.B Representation of livestock water use

Model name	Model	Drivers	Parameters	References
H08	-	-	-	
LPJmL	-	-	-	
MATSIRO	-	-	-	
PCR-GLOBWB	Time-series regression by individual countries and regions	Cattle stock	Set from literature review and time-series data	Wada et al., 2014b
WaterGAP2	Time-series regression by individual countries and regions	Cattle stock	Set from literature review and time-series data	Flörke et al., 2013

1.C Representation of domestic water use

Model name	Model	Drivers	Parameters	References
H08	Time-series regression by individual countries and regions	Population	Set from literature review and time-series data	Hanasaki et al., 2008a, b; Yoshikawa et al., 2014
LPJmL	-	-	-	
MATSIRO	Time-series regression by individual countries and regions	Population	Set from literature review and time-series data	Takata et al., 2003; Pokhrel et al., 2012; 2015
PCR-GLOBWB	Time-series regression by individual countries and regions	Population, GDP per capita	Set from literature review and time-series data	Wada et al., 2014b
WaterGAP2	Time-series regression by individual countries and regions	Population, GDP per capita	Calibrated from time-series data	Flörke et al., 2013

1.D Representation of industrial water use

Model name	Model	Sector	Drivers	Parameters	References
H08	Time-series regression by individual countries and regions	Industry	Infrastructure area	Set from literature review and time-series data	Hanasaki et al., 2008a,b, Yoshikawa et al., 2014
LPJmL	-	-	-	-	
MATSIRO	Time-series regression by individual countries and regions	Industry	electricity production	Set from literature review and time-series data	Takata et al., 2003; Pokhrel et al., 2012; 2015
PCR-GLOBWB	Time-series regression by individual countries and regions	Industry	GDP, electricity production, energy consumption, household consumption	Set from literature review and time-series data	Wada et al., 2014b
WaterGAP2	Time-series regression by individual countries and regions	Manufacturing	Manufacturing structural water intensity, Manufacturing gross value added (GVA), Technological change rate for manufacturing	Calibrated from time-series data	Flörke et al., 2013
		Electricity production	Annual Electricity Production, Water use intensity of thermal power plants (by type of plant and cooling system), Technological change rate		

1.E Representation of irrigation water use

Model name	Calculation of irrigation water requirements	Crop Types	Crop Calendar	Additional components	References
H08	Model-based	MIRCA 2000 (Portmann et al., 2010)	Planting date was determined to obtain maximum yield under meteorological conditions for 1960-1999. The planting date was fixed throughout the simulation period. Harvesting date was calculated in the model and changed with years according to meteorological conditions.	Country/Region specific irrigation efficiency	Hanasaki et al., 2008a,b
LPJmL	Model-based	MIRCA 2000 (Portmann et al., 2010)	n.a. (simulated growing season length)	Country-specific irrigation efficiencies	Bondeau et al., 2007; Rost et al., 2008
MATSIRO	Model-based	MIRCA 2000 (Portmann et al., 2010)	Model-based	Country-specific irrigation efficiencies	Takata et al., 2003; Pokhrel et al., 2012; 2015
PCR-GLOBWB	Model-based	MIRCA 2000 (Portmann et al., 2010) with rice/non-rice distinction	Fixed calendar	Dynamically calculated irrigation efficiency	Wada et al., 2014b
WaterGAP2	Model-based	Rice/Non-rice	Model-based	Country specific irrigation efficiency	Döll and Siebert 2002; Döll et al., 2012, 2014; Portmann et al., 2010; Müller Schmied et al., 2014, 2016

1.F Water allocation

Model name	Sources of water included in water use framework	Parameterizations	References
H08	Surface water (including reservoirs)		Hanasaki et al., 2008a,b
LPJmL	River discharge (incl. renewable groundwater), green water in soils		Bondeau et al., 2007; Rost et al., 2008
MATSIRO	Surface water, Groundwater	Pokhrel et al. 2012; 2015	Takata et al., 2003; Pokhrel et al., 2012; 2015
PCR-GLOBWB	Desalinated, Groundwater, Surface water	Wada et al., 2014a,b	Wada et al., 2014b
WaterGAP2	Groundwater, surface water	Sectoral groundwater use fractions	Döll et al., 2012

1.G Representation of return flows

Model name	Irrigation	Industry	Domestic	Livestock	References
H08	Soil via infiltration; Groundwater: via additional recharge.	-	-	-	Hanasaki et al., 2008b
LPJmL	River; 50% returns	-	-	-	Bondeau et al., 2007; Rost et al., 2008
MATSIRO	River	River	River	-	Pokhrel et al., 2012; 2015
PCR-GLOBWB	Soil Layers via infiltration; Groundwater layer: via additional recharge. Amount determined by irrigation efficiency	River: same day. Amount determined by recycling ratios	River: same day. Amount determined by recycling ratios	No return flow	Wada et al., 2014b
WaterGAP2	Returned to groundwater, Fraction of irrigation water abstraction (from groundwater or surface water) returning to groundwater $f_{rgi} = 0.95 – 0.75f_{drain}$ (cell-specific artificial drainage fraction).	Returned to surface water. Difference of water abstraction and consumptive use.	Returned to surface water. Difference of water abstraction and consumptive use.	No return flow, as only consumptive use.	Döll et al., 2012, 2014b

1.H Representation of Reservoirs (after Pokhrel et al., 2016)

Model name	Purposes included	Parameterizations	References
H08	Non-irrigation	Hanasaki et al., 2006	Hanasaki et al., 2006
LPJmL	Irrigation/Non-irrigation	Biemans et al., 2011	Biemans et al., 2011
MATSIRO	Irrigation/Non-irrigation	Hanasaki et al., 2006	Pokhrel et al., 2012
PCR-GLOBWB	Water supply/Flood control/Hydropower/Navigation	Haddeland et al., 2006; Adam et al., 2007	van Beek et al., 2011
WaterGAP2	Irrigation/Non-irrigation	Hanasaki et al., 2006	Döll et al., 2009

1.I Calibration of runoff and discharge

Model name		References
H08	Bias in runoff corrected with modification of two parameters of subsurface flow for four climatic groups: daily maximum subsurface runoff, relation between subsurface runoff and soil moisture.	Hanasaki et al., 2008a,b
LPJmL	-	

MATSIRO	-	
PCR-GLOBWB	-	
WaterGAP2	Biases in long-term mean annual discharge corrected by adjusting exponent of function where part of effective precipitation becomes runoff from land depending on exponent of ratio of actual to maximum soil storage.	Müller Schmied et al., 2014

Supplementary table 2: Abbreviations

Abbreviation	**Full name**
GRDC	Global Runoff Data Centre
KGE	Modified Kling-Gupta Efficiency
rKGE	KGE correlation coefficient (Pearson)
βKGE	KGE bias ratio
γKGE	KGE variability ratio
Q_{99}	Low-flow indicator
Q_1	High-flow indicator
NSE	Nash-Sutcliffe Model Efficiency
KS	Kolmogorov-Smirnov test
NOHIP conditions	‘naturalized’ model run without human influence
HIP conditions	Model run including time-varying human influence
DDM30	Drainage direction map
HIP	Human impact parameterizations
n	Number
P	Probability
s_i	Simulated discharge
o_i	Observed discharge
i	Station
μ_s	Simulated mean monthly discharge
μ_o	Observed mean monthly discharge
σ_s	Standard deviation of simulated discharge
σ_o	Standard deviation of observed discharge
Q_s	Simulated hydrological extreme
Q_o	Observed hydrological extreme
GVA	Gross value added

Supplementary results for selected river basins

Twelve GRDC stations (**supplementary table 3**) were selected to serve as illustrative focus cases in this study. These are: the Amazonas, Amur, Colorado, Congo, Guadiana, Mackenzie, Murray, Ob, Rhine, Tocantins, Volga, and the Zambezi. The stations are selected for their heterogeneous spatial distribution and belong to both the managed and near-natural catchment groups. Moreover, station-specific results for all stations are provided as a **separate supplement**.

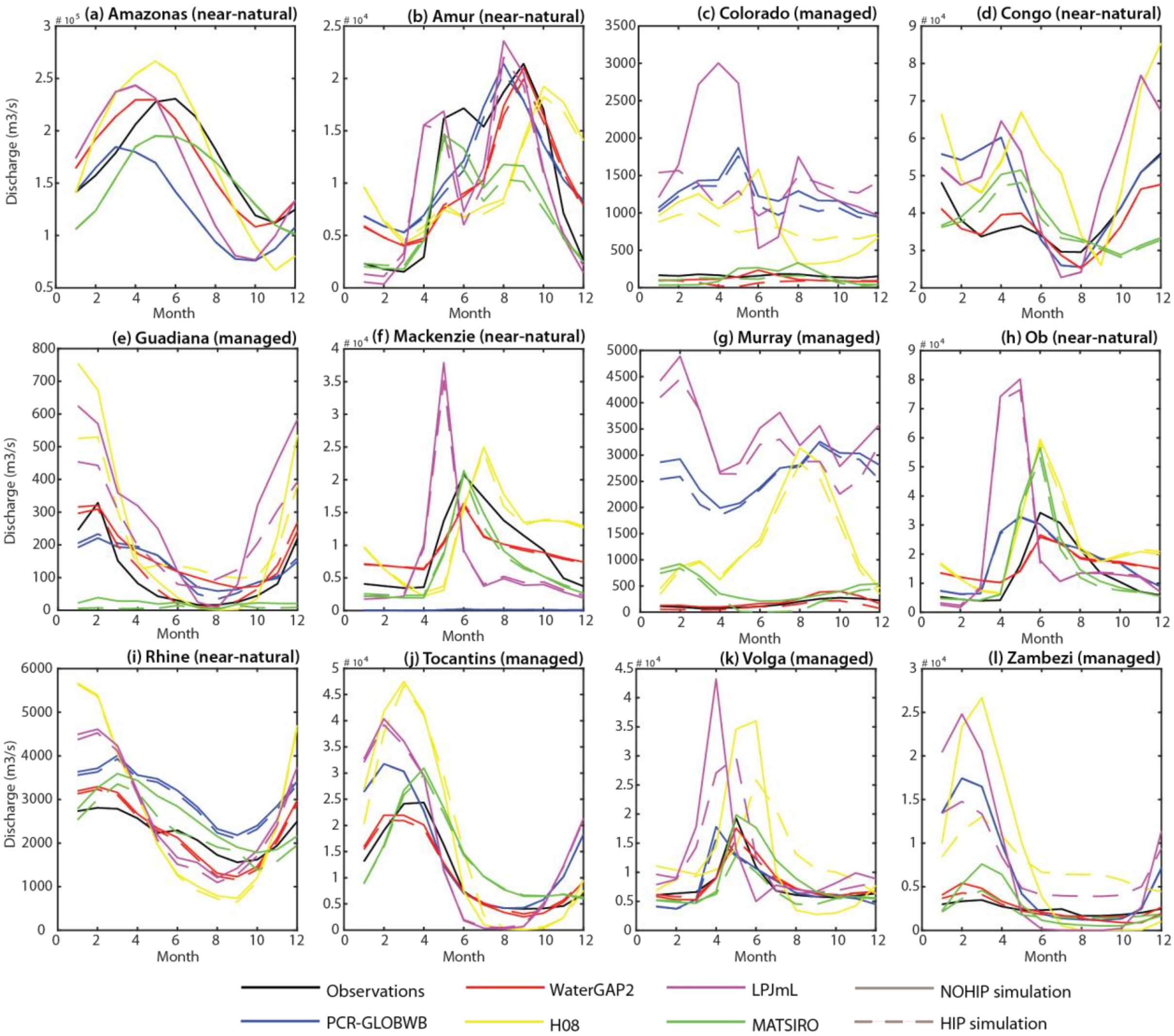
Supplementary table 3: Characteristics of focus stations.

River (station name)	GRDC number	GRDC Catchment area (km^2)	Mean discharge (m^3/sec)	Start- end-year	data-points coverage (%)	Managed/ near-natural
Amazonas (Obidos-Porto)	3629000	4,680,000	170,356	1971-1998	97.9%	near-natural
Amur (Bogorodskoy)	2906901	1,790,000	10,386	1971-1987	99.5%	near-natural
Colorado (Lindero Internacional Norte)	4352100	631,960	160	1971-1995	92%*	managed
Congo (Kinshasa)	1147010	3,475,000	39,127	1971-2010	100%	near-natural
Guadiana (Pulo do Lobo)	6116200	60,883	107	1971-1990	98.8%	managed
Mackenzie (Arctic Red River)	4208025	1,660,000	9,225	1972-2010	98.1%	near-natural
Murray (Overland Corner)	5404270	1,000,001	164	1985-2010	98.1	managed
Ob (Salekhard)	2912600	2,949,998	13,132	1971-2010	97.5%	near-natural
Rhine (Lobith)	6435060	160,800	2,235	1971-2010	100%	near-natural
Tocantins (Tucurui)	3649950	742,300	11,186	1978-2010	94.95%*	managed
Volga (Volgograd Power Plant)	6977100	1,360,000	7,885	1971-2010	97.5%	managed
Zambezi (Matundo-Cais)	1891500	940,000	2,455	1971-1990	96.67%	managed

*stations do not comply to minimum requirement of data-point coverage in time-series (95%) set for full analysis but were included for their representative spatial distribution and representative human/(near-)natural characteristics.

1.Impact of HIP in overall hydrological performance

The change in hydrological performance for 12 managed and near-natural focus stations is shown in **supplementary figure 6**. Little to no change in performance (as calculated with the KGE over the full time-series) is found before and after HIP for the Amazonas, Congo, Mackenzie, Ob, Rhine, and the Tocantins. In contrast, HIP significantly influences the hydrographs of the Amur, Colorado, Guadiana, Murray, Volga, and the Zambezi river. Monthly discharge diminishes significantly for the latter catchments, either in an equal way throughout all months of the year (Murray, Amur) or predominantly during the peak-flow months (Colorado, Guadiana, Volga, Zambezi). For the latter four catchments we observe in most models a significant redistribution of monthly discharge between the peak-flow and low-flow months due to HIP. These cases highlight the influence of dams and reservoirs (and their operation), being part of the HIP framework, on the volume and timing of estimated river discharge. The ability to represent the general hydrologic characteristics varies significantly from catchment to catchment and across models (**supplementary figure 6, supplementary table 4**). Results show that some catchments (Amazonas, Amur, Rhine, or the Tocantins) are generally better represented, whereas the models have relatively more difficulties representing the hydrologic characteristics of others (Colorado, Murray). On the other hand, different models outperform others looking at the various catchments. An overall best performing model does not exist, with WaterGAP2 producing the best results for the Colorado, Congo, Murray, Tocantins, Volga, and Zambezi, compared to LPJmL for the Amur, or MATSIRO for the Amazonas and Ob (**supplementary table 4**).



Supplementary figure 6: Hydrographs for the focus stations under HIP and NOHIP.
Each panel shows the long-term mean monthly discharge for each model. Dashed lines indicate the long-term mean monthly discharges with HIP whilst the solid lines indicate the performance under NOHIP conditions.

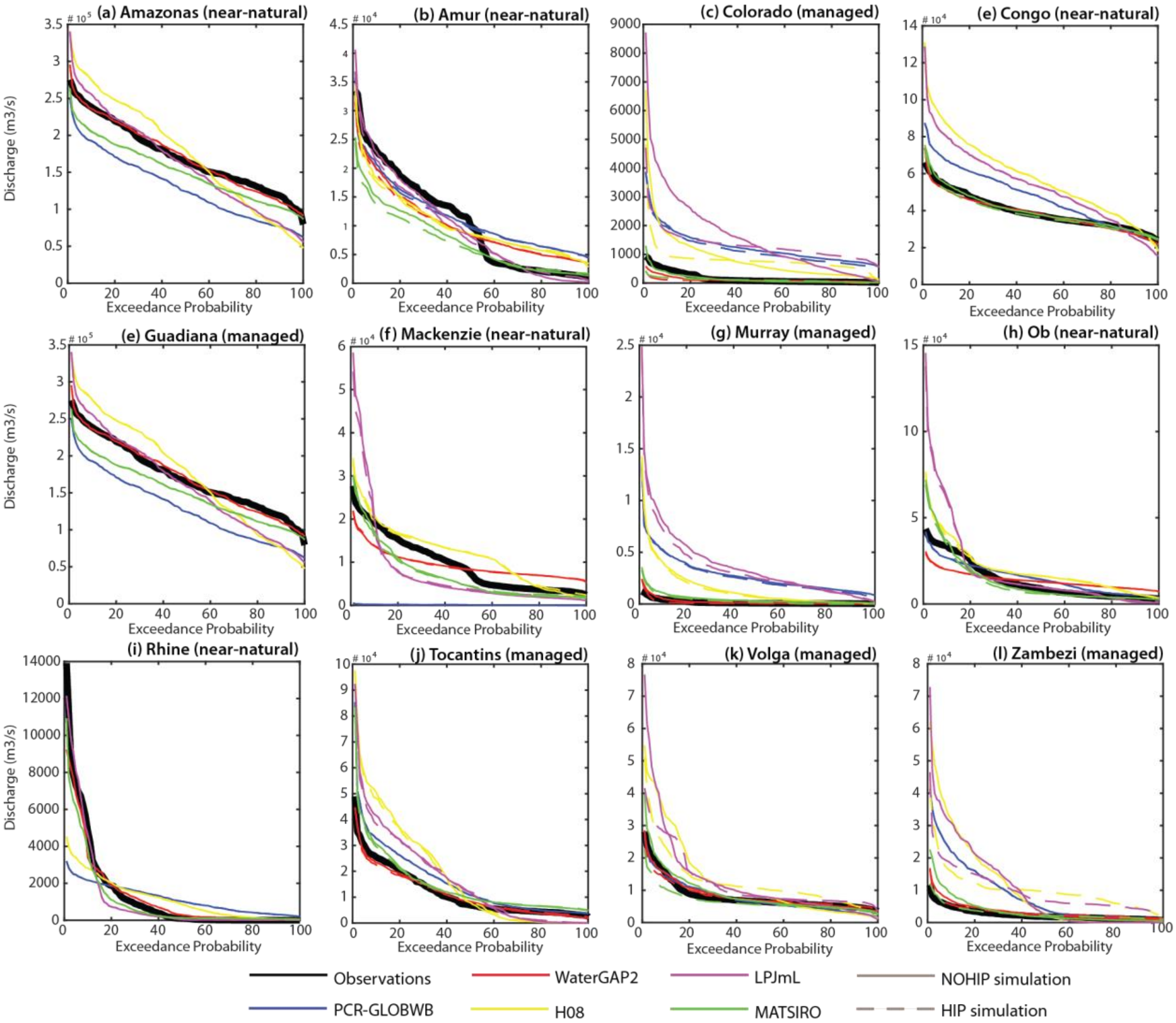
Supplementary table 4: Overall hydrological performance (KGE) for different models for the focus stations.

Performance values with inclusion of human impact parameterizations (HIP) are presented first, performance values without HIP are presented after the slash. Bold values show the best performing models per station. Underscored values show the best performing station per model.

	Amazonas	Amur	Colorado	Congo	Guadiana	Mackenzie	Murray	Ob	Rhine	Tocantins	Volga	Zambezi
PCR-GLOBWB	0.45/ 0.45	0.55/ 0.54	-4.21/ -4.68	0.36/ 0.36	0.32/**0.29**	**0.70/ 0.70**	-13.19/ -14.15	0.51/ 0.50	0.54/0.50	0.58/ 0.58	0.63/ 0.58	-0.92/ -0.92
WaterGAP2	0.81/ 0.82	0.67/ 0.67	**0.13/ 0.23**	**0.66/ 0.66**	**0.49**/0.18	0.48/ 0.47	**-0.11/ 0.64**	0.34/ 0.33	**0.78/0.81**	**0.88/ 0.91**	**0.73/ 0.85**	**0.59/ 0.57**
H08	0.41/ 0.41	0.26/ 0.28	-3.36/ -3.88	0.24/ 0.23	-0.32/- 0.41	0.48/ 0.49	-5.51/ -6.42	0.23/ 0.23	-0.20/- 0.14	0.27/ 0.23	0.15/ 0.04	-1.15/ -1.43
LPJmL	0.41/ 0.40	**0.71/ 0.73**	-7.13/ -8.56	0.23/ 0.23	-0.50/- 1.07	-0.22/ -0.38	-16.96/ -19.44	-0.31/ -0.34	0.34/0.36	0.41/ 0.36	0.36/ -0.31	-1.10/ -1.37
MATSIRO	**0.84/ 0.84**	0.64/ 0.70	-0.21/ 0.14	-0.07/ -0.11	- 0.16/0.09	0.60/ 0.55	-0.33/ -0.95	**0.57/ 0.58**	0.76/0.70	0.71/ 0.69	0.69/ 0.72	0.52/ 0.23

2 Impact of HIP in representation of hydrological extremes

Supplementary figure 7 shows for the focus stations the modelled probability exceedance curves for the different models for both NOHIP and HIP conditions, with as reference the probability exceedance curves based on observational data. Substantial changes in modelled discharge estimates that are the result of HIP can be found under various exceedance probability levels for the Amur, Colorado, Murray, Volga, and Zambezi. For some stations, the most substantial changes are being found at the tails of the exceedance probability curve (at high- and/or low-flows), e.g. the Zambezi or Volga. Other stations show changes for predominantly intermediate exceedance probability levels, e.g. the Amur or Murray. While all models tend to generally overestimate both high- and low-flows in the Murray, Colorado, Tocantins and Zambezi; the results are more varied for the Amazonas, Amur, Guadiana, Mackenzie, and the Ob. For the Congo and Volga most models overestimate high-flows whereas they underestimate low-flow discharges, even with HIP. For the Rhine, this is the other way around, with mostly underestimations for high-flow discharges but overestimations when moving towards low-flows. Both the Amur and Mackenzie see a sudden drop in observed discharges at intermediate levels of exceedance probability. This is likely caused by the strong seasonal characteristics that play an important role in the generation of discharge in these rivers, with a substantial period of the year being snow-dominated. None of the models seems to be able to reflect these particular hydrologic characteristics correctly, both without and with HIP.



Supplementary figure 7: Exceedance probability curves for the focus stations.
Each panel shows the modelled discharge at various levels of exceedance probability with HIP (dashed line) and without HIP (solid line). The thick black lines show the observed discharge at various levels of exceedance probability.