Global hydrological models continue to overestimate river discharge

Stefanie Heinicke *, Jan Volkholz *, Jacob Schewe, Simon N Gosling, Hannes Müller Schmied, Sandra Zimmermann, Matthias Mengel, Inga J Sauer, Peter Burek, Jinfeng Chang, Sian Kou-Giesbrecht, Manoli Grillakis, Luca Guillaumot, Naota Hanasaki, Aristeidis Koutroulis, Kedar Otta, Wei Qi, Yusuke Satoh, Tobias Stacke, Tokuta Yokohata, Katja Frieler

Supplementary methods

Supplementary tables

Supplementary figures

Text S1. Details on the CaMa-Flood simulation

The CaMa-Flood simulations were done with CaMa-Flood, version `v400_20210119`. We have used the standard settings that are provided with the sample runs. In particular this means

- floodplain routing is enabled (`LFLDOUT=.TRUE.`)
- we used the diffuse wave equation instead of kinematic wave equation (`LKINE=.FALSE.`)
- the bifurcation channel scheme is turned on (`LPTHOUT=.TRUE.`)
- the channel width is “GWD-LR + filled with empirical” and the channel depth is set to “empirical power law”
- the Manning coefficients in the river channels are provided by CaMa-Flood
- the Manning coefficient in the floodplain is set to 0.10

We analyzed the variable `outflw`, which is the sum of the discharge in the riverbed and the floodplain, i.e., it represents the total flow of water.

Table S1. Global water models from ISIMIP3a evaluated in this study.

Model	Model version	Modelling group (names of modellers and their institutions)	Main reference	river routing
CLASSIC	CLASSICv1.4	Sian Kou-Giesbrecht: sian.kougiesbrecht@gmail.com, 0000-0002-4086-0561, Canadian Centre for Climate Modelling and Analysis (Canada)	Melton et al., 2020	
CWatM	v1.06	Peter Burek: burek@iiasa.ac.at, 0000-0001-6390-8487, International Institute for Applied Systems Analysis (Austria) Luca Guillaumot: luca.guillaumot@live.fr, International Institute for Applied Systems Analysis (IIASA) (Austria)	Burek et al., 2020	DDM30
H08		Naota Hanasaki: hanasaki@nies.go.jp, 0000-0002-5092-7563, National Institute for Environmental Studies, Japan (Japan) Kedar Otta: otta.kedar@nies.go.jp, 0000-0002-2540-9879, NIES (Japan)	Hanasaki et al., 2018	DDM30
HydroPy	1.0.1 - ISIMIP	Tobias Stacke: tobias.stacke@mpimet.mpg.de, 0000-0003-4637-5337, Max Planck Institute for Meteorology (Germany)	Stacke & Hagemann, 2021	DDM30
JULES-W2	JULESvn6.2	Emmanouil Grillakis: egrillakis@tuc.gr, 0000-0002-4228-1803, Technical University of Crete (Greece) Aristeidis Koutroulis: akoutroulis@tuc.gr, 0000-0002-2999-7575, Technical University of Crete (Greece)	Best et al., 2011	(1) built-in TRIP model (Total Runoff Integrating Pathways), (2) DDM30
MIROC-INTEG-LAND		Yusuke Satoh: yusuke.satoh@kaist.ac.kr, 0000-0001-6419-7330, Korea Advanced Institute of Science and Technology (KAIST) (Japan)	Yokohata et al., 2020	
ORCHIDEE-MICT		Jinfeng Chang: changjf@zju.edu.cn, 0000-0003-4463-7778, College of Environmental and Resource Sciences, Zhejiang University (China) Philippe Ciais: philippe.ciais@lsce.ipsl.fr, 0000-0001-8560-4943, Laboratoire des Sciences du Climat et de l'Environnement (LSCE) (France)	Guimberteau et al., 2014, 2018	STN-30p
WaterGAP2-2e	WaterGAP2.2e	Hannes Müller Schmied: hannes.mueller.schmied@em.uni-frankfurt.de, 0000-0001-5330-9923, Institute of Physical Geography (IPG), Goethe-University Frankfurt (Germany)	Müller Schmied et al., 2021, 2023	DDM30
WEB-DHM-SG	1.0	Wei Qi: QiWei_WaterResources@hotmail.com, Southern University of Science and Technology (China)	Qi et al., 2022	CaMa-Flood model modified based on the DDM30 scheme

Table S2. Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for daily discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
KGE																	
10^{th} percentile	-3.09	-4.98	-4.62	-1.18	-1.14	-2.20	-2.32	-4.06	-4.11	-3.55	-1.34	-3.77	-3.58	-0.14	-1.36	-4.20	-3.38
first quartile	-1.03	-0.90	-1.11	-0.22	-0.24	-0.43	-0.82	-0.69	-0.63	-0.48	-0.27	-0.81	-0.94	0.20	-0.14	-1.45	-1.26
median	-0.25	0.18	0.08	0.25	0.26	0.19	-0.11	0.21	0.28	0.28	0.01	-0.06	0.03	0.46	0.42	-0.43	-0.29
third quartile	0.04	0.48	0.48	0.50	0.51	0.54	0.41	0.48	0.50	0.52	0.16	0.24	0.42	0.63	0.62	0.03	0.08
90^{th} percentile	0.33	0.63	0.65	0.65	0.66	0.70	0.66	0.62	0.64	0.64	0.34	0.46	0.58	0.75	0.75	0.30	0.37
correlation																	
10^{th} percentile	-0.05	0.29	0.23	0.30	0.30	0.21	0.14	0.27	0.28	0.28	0.15	0.15	0.26	0.24	0.33	0.08	0.13
first quartile	0.03	0.46	0.44	0.47	0.49	0.46	0.36	0.43	0.46	0.46	0.29	0.29	0.47	0.45	0.53	0.28	0.29
median	0.17	0.60	0.62	0.61	0.63	0.65	0.60	0.59	0.62	0.60	0.43	0.46	0.63	0.62	0.68	0.50	0.54
third quartile	0.38	0.73	0.72	0.72	0.74	0.79	0.74	0.72	0.74	0.74	0.57	0.61	0.74	0.77	0.78	0.64	0.69
90^{th} percentile	0.62	0.81	0.80	0.80	0.81	0.86	0.82	0.82	0.84	0.83	0.74	0.75	0.82	0.84	0.85	0.78	0.79
bias ratio																	
10^{th} percentile	0.72	0.62	0.62	0.63	0.66	0.61	0.60	0.51	0.63	0.64	0.21	0.08	0.67	0.89	0.82	0.97	0.93
first quartile	0.99	0.83	0.84	0.80	0.80	0.81	0.77	0.74	0.88	0.87	0.36	0.43	0.97	0.95	0.95	1.29	1.28
median	1.39	1.18	1.16	1.06	1.05	1.08	1.01	1.16	1.23	1.21	0.81	1.05	1.49	1.01	1.08	1.88	1.84
third quartile	2.35	2.49	2.33	1.64	1.64	1.64	1.51	2.16	2.17	2.14	1.37	1.94	2.60	1.09	1.59	2.80	2.72
90^{th} percentile	4.79	6.96	6.60	2.82	2.75	3.26	3.18	5.94	5.97	5.51	3.01	5.44	5.51	1.15	2.86	5.24	5.04
variability ratio																	
10^{th} percentile	0.25	0.53	0.58	0.37	0.38	0.63	0.80	0.51	0.45	0.50	0.23	0.36	0.51	0.39	0.60	0.63	0.62
first quartile	0.48	0.68	0.79	0.57	0.60	0.82	1.08	0.66	0.61	0.67	0.34	0.54	0.73	0.61	0.78	0.95	0.91
median	0.94	0.91	1.07	0.86	0.91	1.14	1.52	0.87	0.82	0.89	0.49	0.82	1.01	0.84	1.04	1.31	1.25
third quartile	1.59	1.23	1.49	1.19	1.27	1.65	2.17	1.16	1.10	1.16	0.78	1.25	1.33	1.12	1.32	1.83	1.71
90^{th} percentile	2.07	1.61	2.00	1.69	1.76	2.28	2.85	1.60	1.49	1.52	1.09	1.92	1.78	1.65	1.78	2.55	2.29

Table S3. Correlation between station or catchment properties and model performance for the Kling–Gupta efficiency (KGE) for daily discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	0.25	0.20	0.22	0.22	0.26	0.25	0.27	0.28	0.25	0.28	0.23	0.12	0.28	0.26	0.23	0.23	0.23
Catchment size	0.07	0.04	0.05	0.05	0.07	0.06	0.11	0.03	0.04	0.05	0.05	0.04	0.05	0.14	0.04	0.06	0.06
Clay content	0.04	0.05	0.04	0.06	0.04	0.07	0.07	0.00	0.04	0.02	0.00	-0.03	0.03	-0.03	0.06	-0.01	-0.02
Drainage density	0.07	0.07	0.07	0.09	0.07	0.08	0.04	0.08	0.10	0.09	0.05	-0.12	0.09	-0.07	0.04	0.02	0.00
Elevation	-0.12	-0.09	-0.11	-0.10	-0.14	-0.14	-0.20	-0.12	-0.11	-0.14	-0.12	-0.03	-0.07	-0.26	-0.12	-0.20	-0.20
Irrigated area	-0.06	-0.03	-0.05	-0.02	-0.05	-0.02	-0.05	-0.09	-0.09	-0.10	-0.13	0.00	-0.02	-0.04	-0.01	-0.11	-0.11
Nightlight	0.06	0.06	0.06	0.07	0.05	0.11	0.11	0.17	0.09	0.08	-0.08	0.03	0.04	-0.03	0.00	0.05	0.05
Number of dams	0.04	0.03	0.03	0.04	0.06	0.06	0.08	0.04	0.04	0.04	0.03	0.02	0.03	0.08	0.04	0.03	0.03
Population count	0.06	0.05	0.05	0.05	0.06	0.08	0.11	0.07	0.06	0.07	0.03	0.03	0.05	0.15	0.05	0.05	0.05
Population density	0.03	0.02	0.03	0.03	0.04	0.04	0.05	0.03	0.03	0.03	0.04	0.01	0.04	0.09	0.03	0.03	0.03
Sand content	-0.14	-0.11	-0.13	-0.12	-0.14	-0.10	-0.12	-0.13	-0.13	-0.15	-0.17	-0.19	-0.16	-0.20	-0.15	-0.07	-0.08
Silt content	0.18	0.15	0.15	0.19	0.20	0.12	0.09	0.16	0.16	0.15	0.17	0.07	0.21	0.16	0.16	0.04	0.03
Slope	0.10	0.08	0.08	0.07	0.06	0.08	0.05	0.09	0.09	0.10	0.08	0.03	0.13	-0.09	0.06	0.07	0.07
Storage volume	0.05	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.04	0.05	0.04	0.02	0.05	0.05	0.04	0.04	0.04
Topographic index	-0.07	-0.04	-0.05	-0.03	-0.05	-0.04	-0.05	-0.09	-0.07	-0.09	-0.09	-0.13	-0.08	-0.04	-0.05	-0.08	-0.09

Table S4. Correlation between station or catchment properties and model performance for the first component of the Kling–Gupta efficiency (correlation r) for daily discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	0.15	0.38	0.39	0.39	0.37	0.38	0.32	0.39	0.37	0.43	0.09	0.34	0.39	0.31	0.37	0.31	0.31
Catchment size	0.27	0.13	0.17	0.09	0.13	0.09	0.20	-0.05	-0.08	0.04	0.16	-0.12	0.12	0.23	0.19	0.19	0.19
Clay content	-0.17	0.18	0.13	0.09	0.08	0.17	0.25	0.12	0.09	0.09	0.18	0.17	0.16	0.04	0.10	0.25	0.28
Drainage density	-0.18	0.06	0.02	0.11	0.06	0.06	0.01	0.16	0.16	0.11	0.05	0.15	0.05	-0.01	0.02	-0.01	-0.02
Elevation	-0.18	-0.25	-0.37	-0.30	-0.34	-0.34	-0.38	-0.17	-0.18	-0.27	-0.15	-0.18	-0.29	-0.29	-0.43	-0.33	-0.34
Irrigated area	-0.07	-0.06	-0.05	-0.05	0.01	-0.03	0.01	0.00	0.00	0.00	0.05	0.02	-0.01	-0.06	-0.07	0.02	0.01
Nightlight	-0.21	0.15	0.13	0.13	0.06	0.13	0.11	0.08	0.08	0.03	0.05	0.10	0.00	0.07	0.06	0.10	0.10
Number of dams	-0.01	-0.03	0.04	-0.04	0.06	0.03	0.11	-0.03	-0.05	-0.03	-0.04	-0.10	-0.02	0.09	0.04	0.14	0.15
Population count	0.07	0.14	0.23	0.07	0.19	0.14	0.25	0.07	0.06	0.10	0.11	-0.09	0.10	0.21	0.20	0.24	0.24
Population density	-0.03	0.09	0.11	0.06	0.09	0.08	0.10	0.08	0.09	0.10	0.07	-0.04	0.09	0.10	0.09	0.11	0.10
Sand content	-0.29	-0.05	-0.07	-0.11	-0.11	-0.10	-0.08	-0.03	0.01	-0.01	0.03	-0.09	-0.07	-0.17	-0.10	-0.05	-0.07
Silt content	0.32	-0.04	-0.05	0.11	0.05	0.03	-0.07	-0.07	-0.07	-0.09	-0.16	0.02	-0.06	0.21	0.08	-0.11	-0.10
Slope	-0.11	-0.02	-0.12	-0.10	-0.17	-0.11	-0.16	0.04	0.05	-0.01	-0.06	0.05	0.00	-0.05	-0.16	-0.18	-0.17
Storage volume	0.18	0.07	0.05	-0.02	-0.05	0.01	0.03	-0.17	-0.17	-0.12	-0.03	-0.07	0.01	0.07	0.01	0.06	0.05
Topographic index	-0.10	0.01	0.02	0.02	0.02	0.03	0.05	-0.02	-0.01	-0.01	0.03	-0.03	-0.03	-0.05	0.04	0.08	0.07

Table S5. Correlation between station or catchment properties and model performance for the second component of the Kling–Gupta efficiency (bias ratio β) for daily discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	-0.26	-0.20	-0.22	-0.21	-0.25	-0.24	-0.27	-0.28	-0.25	-0.27	-0.17	-0.12	-0.26	-0.05	-0.24	-0.21	-0.21
Catchment size	-0.05	-0.04	-0.04	-0.03	-0.03	-0.04	-0.05	-0.03	-0.04	-0.05	-0.05	-0.04	-0.04	-0.01	-0.03	-0.04	-0.04
Clay content	-0.03	-0.04	-0.03	-0.08	-0.07	-0.03	-0.01	0.01	-0.03	-0.02	0.09	0.03	-0.02	0.03	-0.05	0.01	0.02
Drainage density	-0.09	-0.08	-0.07	-0.12	-0.12	-0.09	-0.09	-0.08	-0.10	-0.09	-0.03	0.12	-0.10	-0.01	-0.05	-0.03	-0.02
Elevation	0.11	0.09	0.10	0.07	0.09	0.12	0.14	0.11	0.11	0.13	0.06	0.02	0.05	0.01	0.11	0.16	0.16
Irrigated area	0.06	0.03	0.05	0.04	0.07	0.05	0.09	0.10	0.09	0.11	0.16	0.00	0.02	-0.03	0.02	0.11	0.11
Nightlight	-0.07	-0.05	-0.05	-0.07	-0.07	-0.07	-0.07	-0.16	-0.09	-0.08	0.08	-0.03	-0.03	-0.03	-0.03	-0.05	-0.05
Number of dams	-0.03	-0.03	-0.03	-0.03	-0.03	-0.02	-0.02	-0.03	-0.03	-0.03	0.00	-0.02	-0.02	0.07	-0.03	-0.02	-0.02
Population count	-0.05	-0.04	-0.05	-0.03	-0.04	-0.04	-0.05	-0.06	-0.06	-0.06	0.00	-0.03	-0.04	0.09	-0.03	-0.03	-0.03
Population density	-0.03	-0.02	-0.03	-0.03	-0.04	-0.03	-0.04	-0.03	-0.03	-0.03	-0.02	-0.01	-0.03	-0.04	-0.03	-0.03	-0.03
Sand content	0.14	0.12	0.13	0.08	0.09	0.09	0.08	0.14	0.13	0.15	0.20	0.19	0.15	-0.03	0.14	0.06	0.06
Silt content	-0.18	-0.15	-0.16	-0.17	-0.17	-0.13	-0.11	-0.18	-0.17	-0.17	-0.28	-0.07	-0.21	-0.09	-0.17	-0.03	-0.02
Slope	-0.12	-0.08	-0.09	-0.11	-0.12	-0.10	-0.11	-0.10	-0.10	-0.11	-0.13	-0.04	-0.15	-0.02	-0.07	-0.11	-0.11
Storage volume	-0.04	-0.04	-0.04	-0.03	-0.03	-0.04	-0.05	-0.05	-0.05	-0.05	-0.05	-0.03	-0.05	0.06	-0.03	-0.03	-0.03
Topographic index	0.08	0.05	0.06	0.02	0.04	0.06	0.07	0.09	0.07	0.09	0.12	0.13	0.09	-0.03	0.05	0.08	0.09

Table S6. Correlation between station or catchment properties and model performance for the third component of the Kling–Gupta efficiency (variability ratio γ) for daily discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	0.03	-0.13	-0.13	0.14	0.14	0.00	0.01	-0.02	-0.03	-0.05	-0.27	-0.08	0.09	-0.13	0.03	-0.14	-0.12
Catchment size	-0.07	-0.04	-0.04	-0.06	-0.09	-0.02	-0.12	0.04	0.00	-0.07	-0.04	0.04	-0.07	-0.06	-0.13	-0.13	-0.14
Clay content	-0.08	-0.03	-0.11	0.16	0.17	-0.19	-0.15	0.01	0.03	0.06	0.06	-0.01	0.11	0.16	-0.04	0.14	0.16
Drainage density	0.16	0.09	0.07	0.17	0.21	0.04	0.14	0.06	0.12	0.19	0.03	0.02	0.24	0.11	0.15	0.12	0.20
Elevation	0.05	0.21	0.28	0.03	0.10	0.03	0.15	0.09	0.17	0.29	0.34	0.19	0.17	0.28	0.19	0.34	0.32
Irrigated area	-0.10	-0.01	-0.06	0.02	-0.01	-0.07	-0.09	-0.04	-0.03	-0.03	0.05	0.01	-0.02	0.07	-0.03	0.05	0.04
Nightlight	-0.02	0.04	-0.06	0.10	0.07	-0.07	-0.06	-0.01	0.05	0.06	0.03	-0.04	0.03	0.15	0.06	0.12	0.16
Number of dams	-0.12	-0.08	-0.10	0.02	-0.04	-0.05	-0.13	0.01	0.01	-0.07	-0.07	0.06	-0.07	0.06	-0.04	0.01	0.02
Population count	-0.12	-0.06	-0.11	0.03	0.00	-0.03	-0.13	0.00	0.02	-0.03	-0.06	0.02	-0.03	0.05	-0.07	-0.02	0.00
Population density	-0.04	-0.01	0.00	0.02	0.03	-0.04	-0.04	0.00	0.02	0.01	0.00	0.06	0.00	0.01	0.00	0.01	0.02
Sand content	0.05	0.02	-0.05	0.07	0.06	0.03	0.07	0.02	0.00	0.05	0.07	-0.01	0.08	0.12	0.05	0.16	0.22
Silt content	0.05	0.16	0.24	-0.13	-0.11	0.06	0.10	0.01	0.04	0.04	-0.05	-0.06	-0.06	-0.09	0.08	-0.17	-0.19
Slope	0.06	0.24	0.28	0.17	0.22	0.02	0.13	0.12	0.22	0.28	0.25	0.14	0.17	0.29	0.21	0.32	0.29
Storage volume	0.02	0.06	0.07	0.00	0.02	0.09	0.00	0.16	0.04	0.03	0.04	0.20	0.03	-0.03	-0.04	-0.01	-0.04
Topographic index	-0.05	-0.04	-0.13	-0.04	-0.06	-0.06	-0.06	-0.06	-0.09	-0.08	-0.04	-0.13	-0.04	-0.08	-0.10	-0.04	0.03

Table S7. Correlation between station or catchment properties and difference in model performance for the third component of the Kling–Gupta efficiency (variability ratio γ) for daily discharge simulated with models' internal routing scheme (Mir) in comparison to CaMa-Flood (CaMa). Positive values imply that variability is more strongly overestimated for CaMa simulations relative to models' internal routing scheme.

Model	CWatM	H08	HydroPy	JULES-W2	ORCHIDEE-MICT	WaterGAP2-2e	WEB-DHM-SG
Aridity index	-0.05	0.02	0.01	-0.02	0.17	0.20	0.07
Catchment size	-0.01	-0.07	-0.19	-0.10	-0.10	-0.08	-0.01
Clay content	-0.17	0.08	-0.01	0.08	0.12	-0.25	0.04
Drainage density	0.00	0.15	0.24	0.19	0.21	0.04	0.15
Elevation	0.23	0.17	0.25	0.28	-0.08	-0.13	-0.08
Irrigated area	-0.10	-0.07	-0.08	0.03	-0.03	-0.12	-0.04
Nightlight	-0.19	-0.03	0.01	0.08	0.09	-0.14	0.06
Number of dams	-0.08	-0.14	-0.17	-0.12	-0.15	-0.11	0.03
Population count	-0.11	-0.06	-0.20	-0.05	-0.06	-0.14	0.04
Population density	0.01	0.04	-0.02	0.03	-0.09	-0.01	0.01
Sand content	-0.11	-0.03	0.08	0.04	0.10	-0.09	0.14
Silt content	0.23	0.03	0.09	0.05	0.03	0.21	-0.03
Slope	0.20	0.16	0.23	0.22	-0.01	-0.11	-0.10
Storage volume	0.05	0.05	-0.14	-0.23	-0.25	-0.01	-0.07
Topographic index	-0.19	-0.07	-0.02	-0.02	0.14	-0.01	0.15

Table S8. Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for maximum annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
KGE																	
10th percentile	-2.93	-4.69	-5.23	-1.35	-1.32	-2.22	-4.03	-2.67	-2.25	-3.17	-0.63	-1.85	-2.68	-0.72	-1.96	-8.13	-6.53
first quartile	-0.93	-0.98	-1.35	-0.32	-0.40	-0.43	-1.48	-0.44	-0.25	-0.57	-0.27	-0.46	-0.93	-0.09	-0.32	-3.15	-2.16
median	-0.18	0.00	-0.13	0.15	0.09	0.14	-0.23	0.10	0.17	0.11	-0.04	-0.02	0.04	0.24	0.23	-0.94	-0.64
third quartile	0.14	0.33	0.26	0.39	0.37	0.41	0.29	0.35	0.38	0.40	0.19	0.25	0.34	0.44	0.47	-0.06	0.01
90th percentile	0.37	0.50	0.48	0.57	0.55	0.57	0.55	0.52	0.55	0.55	0.38	0.41	0.51	0.60	0.62	0.28	0.29
correlation																	
10th percentile	-0.03	0.11	0.05	0.25	0.23	0.18	0.17	0.19	0.20	0.18	0.03	0.09	0.16	0.23	0.23	0.14	0.14
first quartile	0.11	0.28	0.24	0.41	0.40	0.36	0.32	0.34	0.38	0.35	0.19	0.27	0.36	0.41	0.40	0.30	0.29
median	0.31	0.48	0.43	0.57	0.57	0.53	0.51	0.51	0.53	0.51	0.37	0.44	0.55	0.58	0.55	0.46	0.48
third quartile	0.50	0.65	0.60	0.71	0.70	0.70	0.68	0.66	0.67	0.67	0.54	0.59	0.70	0.71	0.70	0.61	0.63
90th percentile	0.65	0.77	0.74	0.78	0.79	0.80	0.79	0.77	0.79	0.79	0.72	0.73	0.80	0.80	0.79	0.74	0.74
bias ratio																	
10th percentile	0.35	0.56	0.76	0.38	0.50	0.45	0.84	0.35	0.40	0.53	0.12	0.09	0.63	0.39	0.79	1.01	0.97
first quartile	0.63	0.78	1.05	0.60	0.73	0.71	1.17	0.61	0.62	0.77	0.26	0.31	0.96	0.65	0.99	1.64	1.47
median	1.17	1.24	1.56	0.91	1.08	1.11	1.81	1.05	1.03	1.21	0.50	0.77	1.50	0.97	1.31	2.71	2.48
third quartile	2.20	2.63	3.16	1.43	1.68	2.10	3.23	2.13	1.90	2.32	1.01	1.65	2.62	1.39	1.99	4.79	4.08
90th percentile	3.90	6.66	7.20	2.49	2.68	4.03	5.93	4.57	4.11	5.06	2.04	3.29	4.54	2.08	3.89	10.08	8.51
variability ratio																	
10th percentile	0.56	0.57	0.49	0.65	0.62	0.60	0.66	0.41	0.39	0.44	0.43	0.52	0.58	0.68	0.61	0.54	0.57
first quartile	0.78	0.76	0.66	0.91	0.91	0.75	0.84	0.56	0.54	0.59	0.61	0.69	0.75	0.87	0.76	0.70	0.72
median	1.12	1.02	0.87	1.27	1.25	0.99	1.08	0.76	0.75	0.80	0.80	0.94	0.99	1.12	0.94	0.88	0.90
third quartile	1.56	1.31	1.15	1.68	1.69	1.28	1.45	1.06	1.06	1.08	1.15	1.39	1.33	1.47	1.20	1.17	1.15
90th percentile	2.40	1.81	1.54	2.28	2.21	1.68	1.87	1.44	1.41	1.41	1.51	1.97	1.75	1.93	1.48	1.57	1.58

Table S9. Correlation between station or catchment properties and model performance for the Kling–Gupta efficiency (KGE) for maximum annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	0.21	0.29	0.29	0.13	0.14	0.21	0.24	0.25	0.26	0.26	0.14	0.18	0.22	0.17	0.26	0.24	0.25
Catchment size	0.09	0.05	0.06	0.04	0.09	0.05	0.11	0.01	0.02	0.05	0.03	0.07	0.05	0.09	0.07	0.07	0.07
Clay content	0.01	0.03	0.02	0.05	0.03	0.11	0.08	-0.02	0.00	0.00	0.04	-0.03	-0.01	-0.04	0.04	-0.01	-0.02
Drainage density	0.05	0.11	0.04	0.11	0.08	0.16	0.06	0.15	0.11	0.04	0.08	-0.04	0.06	-0.07	0.03	0.05	0.02
Elevation	-0.11	-0.12	-0.19	-0.01	-0.08	-0.07	-0.22	-0.08	-0.12	-0.20	-0.08	-0.06	-0.10	-0.19	-0.20	-0.21	-0.24
Irrigated area	-0.11	-0.05	-0.08	0.00	-0.03	-0.07	-0.12	-0.12	-0.08	-0.11	-0.02	-0.12	-0.02	0.01	-0.04	-0.16	-0.14
Nightlight	0.08	0.05	0.08	0.06	0.08	0.15	0.16	0.19	0.15	0.07	-0.08	0.05	0.02	-0.01	0.11	0.09	0.07
Number of dams	0.03	0.03	0.02	0.05	0.07	0.04	0.07	0.03	0.03	0.03	0.03	0.03	0.04	0.05	0.04	0.02	0.02
Population count	0.07	0.05	0.06	0.03	0.07	0.08	0.12	0.06	0.06	0.07	0.03	0.03	0.05	0.05	0.07	0.06	0.06
Population density	0.02	0.03	0.04	0.02	0.03	0.05	0.06	0.03	0.04	0.04	0.02	-0.03	0.04	0.05	0.04	0.03	0.04
Sand content	-0.06	-0.13	-0.13	-0.06	-0.07	-0.02	-0.02	-0.07	-0.10	-0.14	-0.08	-0.11	-0.14	-0.13	-0.15	-0.08	-0.11
Silt content	0.17	0.20	0.15	0.19	0.20	0.15	0.07	0.19	0.18	0.14	0.14	0.11	0.24	0.15	0.18	0.12	0.13
Slope	0.06	0.10	0.05	0.08	0.03	0.11	0.02	0.10	0.07	0.04	0.03	0.06	0.09	-0.12	-0.01	0.05	0.04
Storage volume	0.07	0.05	0.05	0.03	0.04	0.04	0.07	0.04	0.04	0.04	0.02	0.03	0.03	0.04	0.04	0.05	0.04
Topographic index	-0.08	-0.06	-0.07	0.01	0.00	-0.03	-0.03	-0.04	-0.05	-0.08	0.00	-0.11	-0.06	0.02	-0.03	-0.07	-0.07

Table S10. Correlation between station or catchment properties and model performance for the first component of the Kling–Gupta efficiency (correlation r) for maximum annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	0.11	-0.02	-0.01	-0.02	0.01	0.07	0.05	0.05	0.07	0.11	-0.03	0.08	0.02	-0.01	0.02	0.00	0.02
Catchment size	0.10	0.02	0.00	0.00	0.02	0.00	0.01	-0.07	-0.05	-0.02	0.00	0.04	0.06	0.04	0.03	0.02	0.02
Clay content	-0.01	0.18	0.22	0.06	0.05	0.16	0.18	0.02	0.00	0.02	0.17	0.09	0.12	0.03	0.03	0.12	0.14
Drainage density	-0.03	0.00	-0.01	-0.01	-0.03	0.02	-0.01	0.01	-0.04	-0.04	-0.03	0.04	-0.09	-0.08	-0.09	-0.04	-0.03
Elevation	-0.12	-0.01	-0.03	-0.06	-0.08	-0.13	-0.15	-0.02	-0.09	-0.14	-0.07	-0.11	-0.15	-0.12	-0.14	-0.10	-0.10
Irrigated area	0.02	0.11	0.10	0.07	0.07	0.08	0.11	0.07	0.06	0.06	0.08	0.07	0.07	0.07	0.05	0.10	0.11
Nightlight	-0.10	0.09	0.13	0.08	0.04	0.15	0.09	0.07	0.01	0.01	0.09	0.08	0.01	0.01	-0.01	0.13	0.14
Number of dams	-0.01	0.09	0.10	0.07	0.09	0.10	0.11	0.04	0.01	0.06	0.03	0.09	0.12	0.12	0.07	0.10	0.11
Population count	0.02	0.06	0.10	0.05	0.12	0.11	0.13	0.07	0.04	0.09	0.07	0.12	0.13	0.06	0.10	0.10	0.10
Population density	-0.02	0.03	0.03	-0.02	0.01	0.02	0.04	0.05	0.04	0.03	0.04	-0.01	0.05	0.01	0.03	0.01	0.04
Sand content	-0.13	0.06	0.07	0.00	0.01	0.05	0.03	0.08	0.08	0.07	0.12	-0.04	0.01	-0.02	-0.01	0.04	0.01
Silt content	0.06	-0.19	-0.23	-0.07	-0.09	-0.13	-0.17	-0.12	-0.16	-0.15	-0.24	0.03	-0.14	-0.02	-0.05	-0.15	-0.15
Slope	-0.08	0.05	0.04	-0.02	-0.04	-0.03	-0.04	0.05	-0.01	-0.04	-0.07	0.02	-0.07	-0.05	-0.06	-0.06	-0.04
Storage volume	0.13	-0.03	-0.06	-0.10	-0.08	-0.04	-0.07	-0.17	-0.13	-0.16	-0.09	-0.03	-0.06	-0.04	-0.08	-0.16	-0.16
Topographic index	-0.02	0.03	0.07	0.02	0.00	0.07	0.06	-0.01	0.01	0.03	0.10	0.02	0.03	0.04	0.03	0.08	0.06

Table S11. Correlation between station or catchment properties and model performance for the second component of the Kling–Gupta efficiency (bias ratio β) for maximum annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	-0.26	-0.20	-0.22	-0.21	-0.25	-0.24	-0.27	-0.28	-0.25	-0.27	-0.17	-0.12	-0.26	-0.05	-0.24	-0.21	-0.21
Catchment size	-0.05	-0.04	-0.04	-0.03	-0.03	-0.04	-0.05	-0.03	-0.04	-0.05	-0.05	-0.04	-0.04	-0.01	-0.03	-0.04	-0.04
Clay content	-0.03	-0.04	-0.03	-0.08	-0.07	-0.03	-0.01	0.01	-0.03	-0.02	0.09	0.03	-0.02	0.03	-0.05	0.01	0.02
Drainage density	-0.09	-0.08	-0.07	-0.12	-0.12	-0.09	-0.09	-0.08	-0.10	-0.09	-0.03	0.12	-0.10	-0.01	-0.05	-0.03	-0.02
Elevation	0.11	0.09	0.10	0.07	0.09	0.12	0.14	0.11	0.11	0.13	0.06	0.02	0.05	0.01	0.11	0.16	0.16
Irrigated area	0.06	0.03	0.05	0.04	0.07	0.05	0.09	0.10	0.09	0.11	0.16	0.00	0.02	-0.03	0.02	0.11	0.11
Nightlight	-0.07	-0.05	-0.05	-0.07	-0.07	-0.07	-0.07	-0.16	-0.09	-0.08	0.08	-0.03	-0.03	-0.03	-0.03	-0.05	-0.05
Number of dams	-0.03	-0.03	-0.03	-0.03	-0.03	-0.02	-0.02	-0.03	-0.03	-0.03	0.00	-0.02	-0.02	0.07	-0.03	-0.02	-0.02
Population count	-0.05	-0.04	-0.05	-0.03	-0.04	-0.04	-0.05	-0.06	-0.06	-0.06	0.00	-0.03	-0.04	0.09	-0.03	-0.03	-0.03
Population density	-0.03	-0.02	-0.03	-0.03	-0.04	-0.03	-0.04	-0.03	-0.03	-0.03	-0.02	-0.01	-0.03	-0.04	-0.03	-0.03	-0.03
Sand content	0.14	0.12	0.13	0.08	0.09	0.09	0.08	0.14	0.13	0.15	0.20	0.19	0.15	-0.03	0.14	0.06	0.06
Silt content	-0.18	-0.15	-0.16	-0.17	-0.17	-0.13	-0.11	-0.18	-0.17	-0.17	-0.28	-0.07	-0.21	-0.09	-0.17	-0.03	-0.02
Slope	-0.12	-0.08	-0.09	-0.11	-0.12	-0.10	-0.11	-0.10	-0.10	-0.11	-0.13	-0.04	-0.15	-0.02	-0.07	-0.11	-0.11
Storage volume	-0.04	-0.04	-0.04	-0.03	-0.03	-0.04	-0.05	-0.05	-0.05	-0.05	-0.05	-0.03	-0.05	0.06	-0.03	-0.03	-0.03
Topographic index	0.08	0.05	0.06	0.02	0.04	0.06	0.07	0.09	0.07	0.09	0.12	0.13	0.09	-0.03	0.05	0.08	0.09

Table S12. Correlation between station or catchment properties and model performance for the third component of the Kling–Gupta efficiency (variability ratio γ) for maximum annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
Aridity index	0.03	-0.13	-0.13	0.14	0.14	0.00	0.01	-0.02	-0.03	-0.05	-0.27	-0.08	0.09	-0.13	0.03	-0.14	-0.12
Catchment size	-0.07	-0.04	-0.04	-0.06	-0.09	-0.02	-0.12	0.04	0.00	-0.07	-0.04	0.04	-0.07	-0.06	-0.13	-0.13	-0.14
Clay content	-0.08	-0.03	-0.11	0.16	0.17	-0.19	-0.15	0.01	0.03	0.06	0.06	-0.01	0.11	0.16	-0.04	0.14	0.16
Drainage density	0.16	0.09	0.07	0.17	0.21	0.04	0.14	0.06	0.12	0.19	0.03	0.02	0.24	0.11	0.15	0.12	0.20
Elevation	0.05	0.21	0.28	0.03	0.10	0.03	0.15	0.09	0.17	0.29	0.34	0.19	0.17	0.28	0.19	0.34	0.32
Irrigated area	-0.10	-0.01	-0.06	0.02	-0.01	-0.07	-0.09	-0.04	-0.03	-0.03	0.05	0.01	-0.02	0.07	-0.03	0.05	0.04
Nightlight	-0.02	0.04	-0.06	0.10	0.07	-0.07	-0.06	-0.01	0.05	0.06	0.03	-0.04	0.03	0.15	0.06	0.12	0.16
Number of dams	-0.12	-0.08	-0.10	0.02	-0.04	-0.05	-0.13	0.01	0.01	-0.07	-0.07	0.06	-0.07	0.06	-0.04	0.01	0.02
Population count	-0.12	-0.06	-0.11	0.03	0.00	-0.03	-0.13	0.00	0.02	-0.03	-0.06	0.02	-0.03	0.05	-0.07	-0.02	0.00
Population density	-0.04	-0.01	0.00	0.02	0.03	-0.04	-0.04	0.00	0.02	0.01	0.00	0.06	0.00	0.01	0.00	0.01	0.02
Sand content	0.05	0.02	-0.05	0.07	0.06	0.03	0.07	0.02	0.00	0.05	0.07	-0.01	0.08	0.12	0.05	0.16	0.22
Silt content	0.05	0.16	0.24	-0.13	-0.11	0.06	0.10	0.01	0.04	0.04	-0.05	-0.06	-0.06	-0.09	0.08	-0.17	-0.19
Slope	0.06	0.24	0.28	0.17	0.22	0.02	0.13	0.12	0.22	0.28	0.25	0.14	0.17	0.29	0.21	0.32	0.29
Storage volume	0.02	0.06	0.07	0.00	0.02	0.09	0.00	0.16	0.04	0.03	0.04	0.20	0.03	-0.03	-0.04	-0.01	-0.04
Topographic index	-0.05	-0.04	-0.13	-0.04	-0.06	-0.06	-0.06	-0.06	-0.09	-0.08	-0.04	-0.13	-0.04	-0.08	-0.10	-0.04	0.03

Table S13. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for daily discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
KGE																	
10th percentile	-7.98	-14.88	-12.31	-3.08	-3.16	-3.82	-4.54	-14.36	-13.71	-11.60	-2.89	-9.43	-8.73	-0.34	-4.31	-10.68	-9.09
first quartile	-2.86	-4.61	-4.67	-0.64	-0.79	-1.40	-1.64	-4.06	-3.48	-3.52	-1.08	-1.79	-3.26	0.03	-1.27	-3.19	-3.04
median	**-0.81**	**-0.65**	**-0.94**	**-0.09**	**-0.12**	**-0.20**	**-0.57**	**-0.32**	**-0.43**	**-0.47**	**-0.19**	**-0.27**	**-0.55**	**0.29**	**-0.11**	**-1.06**	**-0.93**
third quartile	-0.15	0.09	-0.07	0.28	0.34	0.24	0.09	0.30	0.28	0.26	0.08	0.08	0.14	0.52	0.44	-0.22	-0.12
90th percentile	0.09	0.45	0.39	0.56	0.58	0.53	0.47	0.50	0.50	0.48	0.21	0.37	0.44	0.67	0.61	0.17	0.21
correlation																	
10th percentile	-0.03	0.22	0.15	0.18	0.19	0.13	0.11	0.21	0.23	0.21	0.13	0.09	0.21	0.16	0.19	0.04	0.05
first quartile	0.01	0.36	0.29	0.35	0.39	0.34	0.26	0.36	0.36	0.35	0.24	0.22	0.33	0.34	0.38	0.17	0.19
median	0.11	0.51	0.50	0.52	0.53	0.53	0.46	0.50	0.52	0.50	0.41	0.35	0.54	0.49	0.58	0.39	0.42
third quartile	0.30	0.64	0.63	0.65	0.66	0.68	0.65	0.61	0.63	0.60	0.53	0.52	0.67	0.67	0.68	0.54	0.60
90th percentile	0.53	0.73	0.72	0.73	0.74	0.76	0.73	0.71	0.74	0.74	0.70	0.65	0.75	0.79	0.78	0.68	0.70
bias ratio																	
10th percentile	0.71	0.64	0.67	0.53	0.57	0.55	0.58	0.56	0.62	0.64	0.16	0.06	0.57	0.86	0.98	0.74	0.77
first quartile	1.11	1.01	1.07	0.88	0.91	0.75	0.78	0.90	1.00	1.03	0.26	0.41	0.93	0.93	1.13	1.21	1.20
median	2.08	2.13	2.22	1.34	1.39	1.35	1.32	1.88	2.08	1.99	0.81	1.08	2.02	1.01	1.65	2.09	2.06
third quartile	4.64	6.58	6.64	2.34	2.42	2.84	2.72	5.85	5.39	5.47	2.49	3.12	5.05	1.10	2.89	4.42	4.04
90th percentile	9.90	16.87	14.29	4.83	4.85	5.59	6.42	16.35	15.67	13.57	4.73	10.96	10.71	1.19	6.24	12.63	11.08
variability ratio																	
10th percentile	0.13	0.49	0.46	0.29	0.28	0.51	0.66	0.42	0.34	0.41	0.29	0.29	0.36	0.39	0.49	0.64	0.62
first quartile	0.32	0.64	0.69	0.46	0.47	0.73	0.93	0.57	0.49	0.59	0.41	0.50	0.57	0.62	0.71	0.95	0.90
median	0.77	0.94	1.12	0.71	0.77	1.09	1.42	0.86	0.83	0.92	0.62	0.88	0.93	0.88	0.98	1.40	1.36
third quartile	1.60	1.32	1.63	1.00	1.09	1.65	2.19	1.18	1.11	1.26	0.96	1.34	1.39	1.34	1.30	2.23	1.97
90th percentile	2.00	1.83	2.35	1.81	1.91	2.13	2.81	1.61	1.57	1.85	1.48	2.02	1.83	2.04	1.90	3.16	2.82

Table S14. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for mean monthly discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
KGE																	
10^{th} percentile	-8.08	-14.74	-12.23	-3.05	-3.04	-3.83	-4.50	-14.50	-13.71	-11.60	-2.70	-9.57	-8.72	-0.18	-4.40	-10.69	-9.13
first quartile	-2.86	-4.61	-4.64	-0.63	-0.65	-1.36	-1.28	-4.06	-3.48	-3.60	-0.77	-1.72	-3.28	0.15	-1.10	-2.83	-2.38
median	**-0.66**	**-0.49**	**-0.53**	**-0.03**	**-0.04**	**-0.18**	**-0.19**	**-0.23**	**-0.34**	**-0.24**	**-0.10**	**-0.19**	**-0.43**	**0.41**	**0.05**	**-0.72**	**-0.63**
third quartile	-0.05	0.19	0.15	0.32	0.41	0.27	0.28	0.34	0.31	0.37	0.14	0.14	0.23	0.62	0.51	0.01	0.05
90^{th} percentile	0.21	0.48	0.49	0.59	0.63	0.55	0.57	0.54	0.55	0.58	0.32	0.43	0.47	0.76	0.70	0.28	0.33
correlation																	
10^{th} percentile	-0.05	0.30	0.25	0.27	0.26	0.21	0.20	0.31	0.29	0.32	0.20	0.11	0.29	0.24	0.33	0.09	0.12
first quartile	0.02	0.45	0.42	0.40	0.46	0.42	0.40	0.45	0.43	0.47	0.37	0.29	0.42	0.44	0.53	0.32	0.31
median	**0.17**	**0.61**	**0.60**	**0.59**	**0.63**	**0.60**	**0.61**	**0.60**	**0.61**	**0.64**	**0.54**	**0.45**	**0.61**	**0.62**	**0.68**	**0.56**	**0.57**
third quartile	0.39	0.72	0.73	0.71	0.74	0.72	0.73	0.71	0.73	0.75	0.67	0.58	0.73	0.76	0.79	0.70	0.70
90^{th} percentile	0.64	0.80	0.82	0.79	0.80	0.81	0.82	0.78	0.80	0.82	0.80	0.71	0.79	0.85	0.84	0.79	0.80
bias ratio																	
10^{th} percentile	0.71	0.64	0.67	0.52	0.56	0.55	0.58	0.56	0.62	0.65	0.16	0.06	0.58	0.85	0.98	0.74	0.77
first quartile	1.11	1.02	1.08	0.89	0.92	0.75	0.78	0.90	1.00	1.03	0.26	0.40	0.93	0.93	1.13	1.20	1.21
median	**2.07**	**2.13**	**2.20**	**1.33**	**1.38**	**1.36**	**1.32**	**1.88**	**2.07**	**1.99**	**0.81**	**1.07**	**2.02**	**1.00**	**1.64**	**2.09**	**2.06**
third quartile	4.70	6.57	6.63	2.35	2.42	2.85	2.71	5.78	5.34	5.58	2.49	3.10	5.04	1.10	2.88	4.44	4.05
90^{th} percentile	10.01	16.72	14.23	4.83	4.84	5.62	6.44	16.48	15.69	13.59	4.65	11.07	10.71	1.20	6.25	12.64	11.12
variability ratio																	
10^{th} percentile	0.20	0.54	0.51	0.35	0.33	0.70	0.75	0.47	0.45	0.46	0.31	0.39	0.46	0.42	0.52	0.62	0.61
first quartile	0.39	0.69	0.70	0.50	0.50	0.92	0.94	0.60	0.59	0.58	0.41	0.61	0.62	0.63	0.71	0.91	0.88
median	**0.81**	**0.90**	**0.99**	**0.76**	**0.77**	**1.25**	**1.30**	**0.83**	**0.83**	**0.80**	**0.55**	**0.90**	**1.00**	**0.88**	**0.94**	**1.21**	**1.21**
third quartile	1.48	1.25	1.44	1.03	1.01	1.73	1.76	1.16	1.11	1.11	0.83	1.31	1.27	1.18	1.19	1.76	1.73
90^{th} percentile	1.86	1.69	1.82	1.71	1.71	2.20	2.22	1.49	1.43	1.43	1.18	1.82	1.75	1.75	1.49	2.44	2.31

Table S15. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for long-term mean monthly discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
KGE																	
10th percentile	-8.34	-15.37	-12.36	-3.03	-3.03	-5.07	-4.86	-14.89	-15.76	-11.71	-2.83	-9.72	-8.77	-0.49	-4.92	-11.04	-10.31
first quartile	-3.05	-4.78	-5.02	-0.71	-0.86	-1.86	-1.65	-4.67	-4.08	-3.73	-0.92	-2.12	-3.69	0.06	-1.42	-3.22	-2.90
median	**-1.01**	**-0.63**	**-0.92**	**-0.04**	**-0.08**	**-0.24**	**-0.19**	**-0.39**	**-0.50**	**-0.42**	**-0.15**	**-0.21**	**-0.68**	**0.45**	**0.05**	**-0.72**	**-0.69**
third quartile	-0.06	0.19	0.13	0.37	0.44	0.32	0.30	0.36	0.28	0.39	0.17	0.16	0.24	0.68	0.53	0.00	0.06
90th percentile	0.26	0.54	0.50	0.63	0.65	0.61	0.62	0.58	0.59	0.59	0.39	0.44	0.49	0.81	0.76	0.34	0.38
correlation																	
10th percentile	-0.60	0.31	0.22	0.28	0.24	0.14	0.21	0.32	0.22	0.32	0.04	-0.08	0.30	0.04	0.34	0.02	0.09
first quartile	-0.21	0.61	0.57	0.50	0.54	0.49	0.50	0.59	0.59	0.59	0.51	0.34	0.51	0.52	0.69	0.45	0.50
median	**0.30**	**0.80**	**0.82**	**0.76**	**0.77**	**0.77**	**0.76**	**0.79**	**0.78**	**0.82**	**0.76**	**0.59**	**0.73**	**0.79**	**0.85**	**0.76**	**0.75**
third quartile	0.64	0.91	0.91	0.88	0.88	0.87	0.87	0.90	0.90	0.91	0.89	0.78	0.88	0.92	0.94	0.89	0.90
90th percentile	0.83	0.96	0.96	0.93	0.94	0.94	0.95	0.95	0.95	0.96	0.95	0.88	0.94	0.97	0.97	0.95	0.95
bias ratio																	
10th percentile	0.71	0.64	0.67	0.53	0.55	0.57	0.58	0.55	0.63	0.64	0.15	0.06	0.59	0.86	0.97	0.74	0.79
first quartile	1.12	1.02	1.08	0.89	0.92	0.75	0.78	0.89	1.00	1.03	0.26	0.40	0.96	0.93	1.13	1.22	1.21
median	**2.05**	**2.11**	**2.20**	**1.32**	**1.38**	**1.35**	**1.32**	**1.95**	**2.06**	**1.97**	**0.80**	**1.06**	**2.03**	**1.01**	**1.65**	**2.11**	**2.06**
third quartile	4.73	6.71	6.74	2.33	2.42	2.85	2.67	5.72	5.34	5.62	2.48	3.12	4.98	1.10	2.89	4.39	4.02
90th percentile	10.10	17.37	14.34	4.83	4.91	5.73	6.46	16.59	16.55	13.71	4.62	11.27	10.75	1.19	6.25	12.71	11.24
variability ratio																	
10th percentile	0.04	0.59	0.65	0.29	0.30	0.71	0.70	0.58	0.56	0.56	0.23	0.37	0.50	0.29	0.53	0.48	0.48
first quartile	0.15	0.76	0.81	0.43	0.48	0.92	0.92	0.70	0.71	0.69	0.45	0.52	0.71	0.54	0.72	0.89	0.86
median	**0.53**	**1.01**	**1.08**	**0.67**	**0.72**	**1.30**	**1.29**	**0.91**	**0.89**	**0.86**	**0.67**	**0.83**	**1.03**	**0.83**	**1.00**	**1.27**	**1.27**
third quartile	1.37	1.39	1.53	0.97	1.09	1.75	1.75	1.23	1.22	1.22	0.95	1.22	1.42	1.28	1.35	1.77	1.73
90th percentile	1.81	1.98	2.17	1.75	1.85	2.86	2.85	1.77	1.89	1.91	1.43	1.74	2.04	1.86	1.99	2.75	2.75

Table S16. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for mean annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
KGE																	
10th percentile	-7.88	-15.01	-12.69	-3.08	-3.14	-3.82	-4.44	-14.83	-14.17	-11.84	-3.11	-9.49	-8.78	0.02	-4.12	-10.59	-9.16
first quartile	-2.95	-4.82	-4.71	-0.59	-0.63	-1.05	-1.12	-3.93	-3.37	-3.40	-0.85	-1.79	-3.04	0.33	-1.03	-2.78	-2.26
median	**-0.55**	**-0.38**	**-0.36**	**0.09**	**0.10**	**0.02**	**0.09**	**-0.32**	**-0.34**	**-0.31**	**-0.05**	**-0.16**	**-0.45**	**0.59**	**0.10**	**-0.46**	**-0.45**
third quartile	0.16	0.28	0.29	0.48	0.50	0.46	0.47	0.42	0.40	0.42	0.16	0.29	0.26	0.72	0.52	0.19	0.24
90th percentile	0.50	0.59	0.60	0.71	0.70	0.65	0.66	0.64	0.63	0.64	0.36	0.52	0.59	0.83	0.72	0.48	0.58
correlation																	
10th percentile	-0.04	0.38	0.36	0.38	0.39	0.38	0.36	0.36	0.38	0.38	0.20	0.18	0.39	0.42	0.42	0.31	0.30
first quartile	0.15	0.53	0.52	0.59	0.63	0.54	0.55	0.59	0.59	0.60	0.42	0.48	0.60	0.61	0.61	0.55	0.56
median	**0.43**	**0.68**	**0.66**	**0.74**	**0.77**	**0.70**	**0.72**	**0.71**	**0.72**	**0.73**	**0.59**	**0.66**	**0.76**	**0.76**	**0.74**	**0.70**	**0.72**
third quartile	0.66	0.82	0.82	0.85	0.86	0.84	0.84	0.84	0.85	0.85	0.74	0.80	0.86	0.85	0.83	0.81	0.81
90th percentile	0.78	0.88	0.88	0.90	0.90	0.92	0.92	0.91	0.91	0.91	0.83	0.90	0.91	0.91	0.90	0.89	0.89
bias ratio																	
10th percentile	0.69	0.64	0.67	0.51	0.56	0.53	0.58	0.55	0.61	0.64	0.16	0.06	0.59	0.85	0.98	0.74	0.77
first quartile	1.07	1.01	1.08	0.89	0.92	0.75	0.77	0.90	1.00	1.03	0.26	0.40	0.96	0.92	1.12	1.19	1.19
median	**2.05**	**2.14**	**2.16**	**1.36**	**1.39**	**1.35**	**1.33**	**1.84**	**2.02**	**1.98**	**0.81**	**1.05**	**2.03**	**1.01**	**1.65**	**2.10**	**2.09**
third quartile	4.76	6.66	6.64	2.34	2.44	2.81	2.69	5.88	5.28	5.34	2.48	3.10	4.86	1.10	2.94	4.31	4.01
90th percentile	9.71	17.00	14.67	4.89	5.01	5.81	6.37	16.81	16.15	13.82	5.05	11.47	10.75	1.21	6.09	12.55	11.15
variability ratio																	
10th percentile	0.30	0.39	0.37	0.41	0.37	0.61	0.56	0.36	0.36	0.35	0.26	0.51	0.49	0.65	0.51	0.63	0.61
first quartile	0.51	0.53	0.50	0.62	0.60	0.79	0.74	0.49	0.47	0.47	0.36	0.69	0.67	0.79	0.64	0.83	0.81
median	**0.76**	**0.74**	**0.68**	**0.83**	**0.79**	**1.06**	**1.01**	**0.75**	**0.72**	**0.70**	**0.51**	**1.01**	**0.92**	**1.01**	**0.85**	**1.11**	**1.07**
third quartile	1.18	0.96	0.90	1.21	1.12	1.39	1.38	1.06	1.08	1.04	0.78	1.54	1.31	1.28	1.07	1.44	1.39
90th percentile	1.87	1.27	1.19	1.54	1.45	1.97	1.79	1.49	1.48	1.45	0.99	2.49	1.79	1.76	1.36	2.01	1.95

Table S17. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for maximum annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

Model	CLASSIC	CWatM		H08		HydroPy		JULES-W2			MIROC-INTEG-LAND	ORCHIDEE-MICT		WaterGAP2-2e		WEB-DHM-SG	
River routing	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa	Mir	Mir 2	CaMa	CaMa	Mir	CaMa	Mir	CaMa	Mir	CaMa
<u>KGE</u>																	
10^{th} percentile	-4.34	-11.62	-12.49	-2.51	-2.66	-3.51	-7.98	-7.25	-5.40	-9.35	-1.45	-3.90	-4.86	-1.54	-4.90	-15.37	-15.90
first quartile	-2.01	-4.68	-5.21	-0.49	-0.64	-1.15	-3.16	-2.20	-2.03	-3.22	-0.40	-0.94	-1.80	-0.28	-1.83	-6.67	-5.55
median	**-0.43**	**-0.89**	**-1.64**	**0.02**	**-0.06**	**-0.18**	**-0.77**	**-0.16**	**-0.09**	**-0.58**	**-0.13**	**-0.13**	**-0.44**	**0.16**	**-0.24**	**-1.99**	**-1.76**
third quartile	-0.02	0.14	-0.12	0.33	0.28	0.27	-0.04	0.22	0.26	0.16	0.13	0.14	0.23	0.39	0.31	-0.19	-0.06
90^{th} percentile	0.26	0.36	0.28	0.56	0.52	0.49	0.39	0.42	0.43	0.39	0.29	0.34	0.45	0.57	0.54	0.26	0.29
<u>correlation</u>																	
10^{th} percentile	-0.06	0.17	0.12	0.26	0.23	0.17	0.15	0.14	0.20	0.15	0.01	0.03	0.20	0.26	0.20	0.14	0.15
first quartile	0.08	0.29	0.24	0.42	0.40	0.32	0.31	0.32	0.35	0.31	0.22	0.22	0.33	0.40	0.39	0.30	0.30
median	**0.26**	**0.45**	**0.44**	**0.56**	**0.57**	**0.51**	**0.48**	**0.49**	**0.49**	**0.46**	**0.39**	**0.39**	**0.50**	**0.58**	**0.55**	**0.43**	**0.45**
third quartile	0.45	0.65	0.62	0.72	0.69	0.70	0.69	0.64	0.64	0.62	0.56	0.55	0.68	0.71	0.69	0.62	0.63
90^{th} percentile	0.62	0.77	0.75	0.79	0.80	0.80	0.79	0.75	0.78	0.76	0.73	0.71	0.80	0.80	0.79	0.77	0.75
<u>bias ratio</u>																	
10^{th} percentile	0.23	0.72	1.03	0.27	0.28	0.28	0.60	0.49	0.55	0.66	0.13	0.06	0.62	0.34	0.91	0.93	0.91
first quartile	0.65	1.12	1.61	0.51	0.67	0.70	1.21	0.83	0.87	1.18	0.26	0.30	0.98	0.63	1.22	1.78	1.60
median	**1.44**	**2.55**	**3.58**	**0.87**	**1.12**	**1.34**	**2.39**	**1.60**	**1.59**	**2.35**	**0.77**	**0.95**	**1.84**	**1.01**	**1.97**	**3.60**	**3.31**
third quartile	3.36	6.65	7.18	1.59	2.07	2.78	5.00	4.11	3.95	5.11	1.89	2.23	3.69	1.55	3.69	8.64	7.39
90^{th} percentile	6.03	13.61	14.48	3.49	3.94	5.40	9.96	9.21	7.35	11.33	3.25	5.33	6.81	2.85	6.81	17.34	17.89
<u>variability ratio</u>																	
10^{th} percentile	0.48	0.51	0.44	0.62	0.57	0.56	0.61	0.34	0.34	0.40	0.40	0.46	0.52	0.67	0.56	0.50	0.55
first quartile	0.69	0.65	0.56	0.81	0.80	0.71	0.72	0.46	0.45	0.49	0.54	0.62	0.67	0.83	0.71	0.70	0.71
median	**0.98**	**0.92**	**0.75**	**1.19**	**1.17**	**0.98**	**1.04**	**0.62**	**0.61**	**0.66**	**0.73**	**0.84**	**0.89**	**1.08**	**0.89**	**0.92**	**0.95**
third quartile	1.55	1.27	1.05	1.79	1.68	1.36	1.52	0.95	0.95	0.99	1.03	1.33	1.24	1.47	1.16	1.29	1.29
90^{th} percentile	2.76	1.89	1.48	2.45	2.37	2.01	2.02	1.42	1.42	1.41	1.48	2.10	1.83	2.05	1.42	1.87	1.82

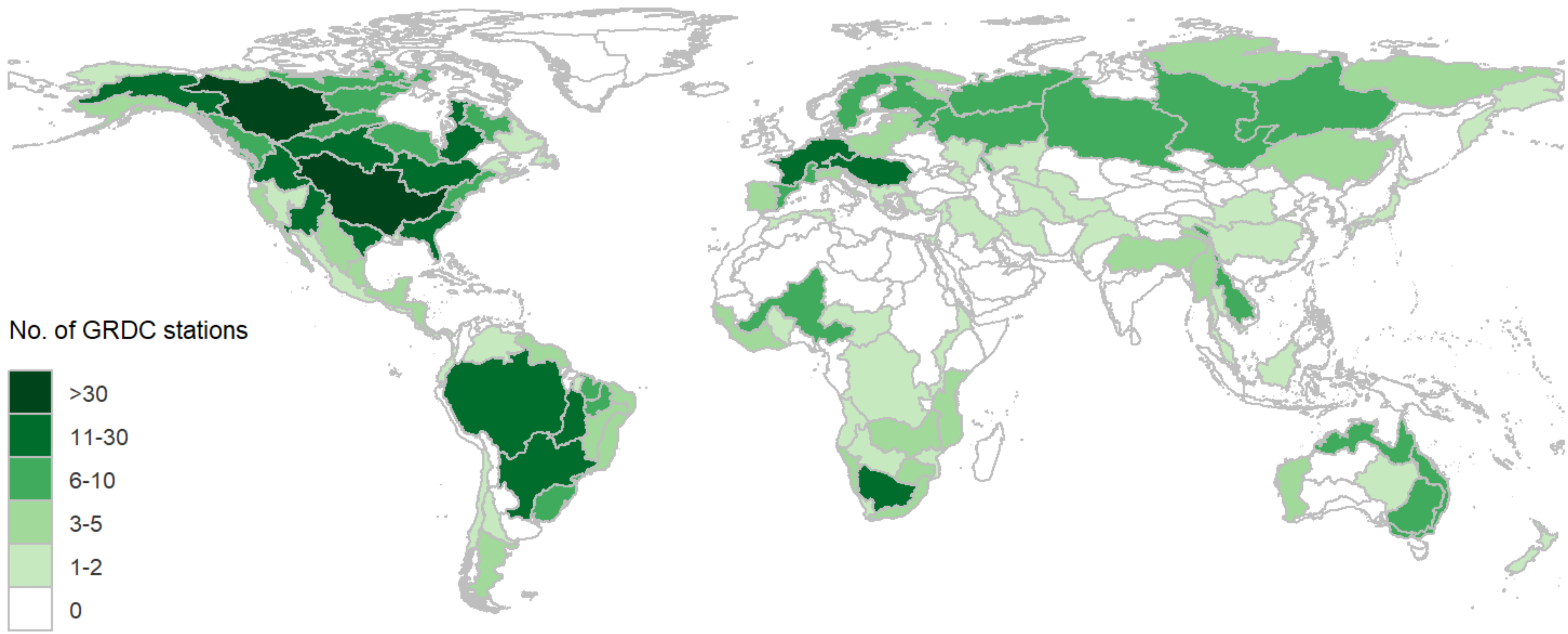

Figure S1. Map with number of stations per basin evaluated for this study. Basins correspond to level 3 of HydroBASINS from the HydroSHEDS database (Lehner & Grill, 2013).

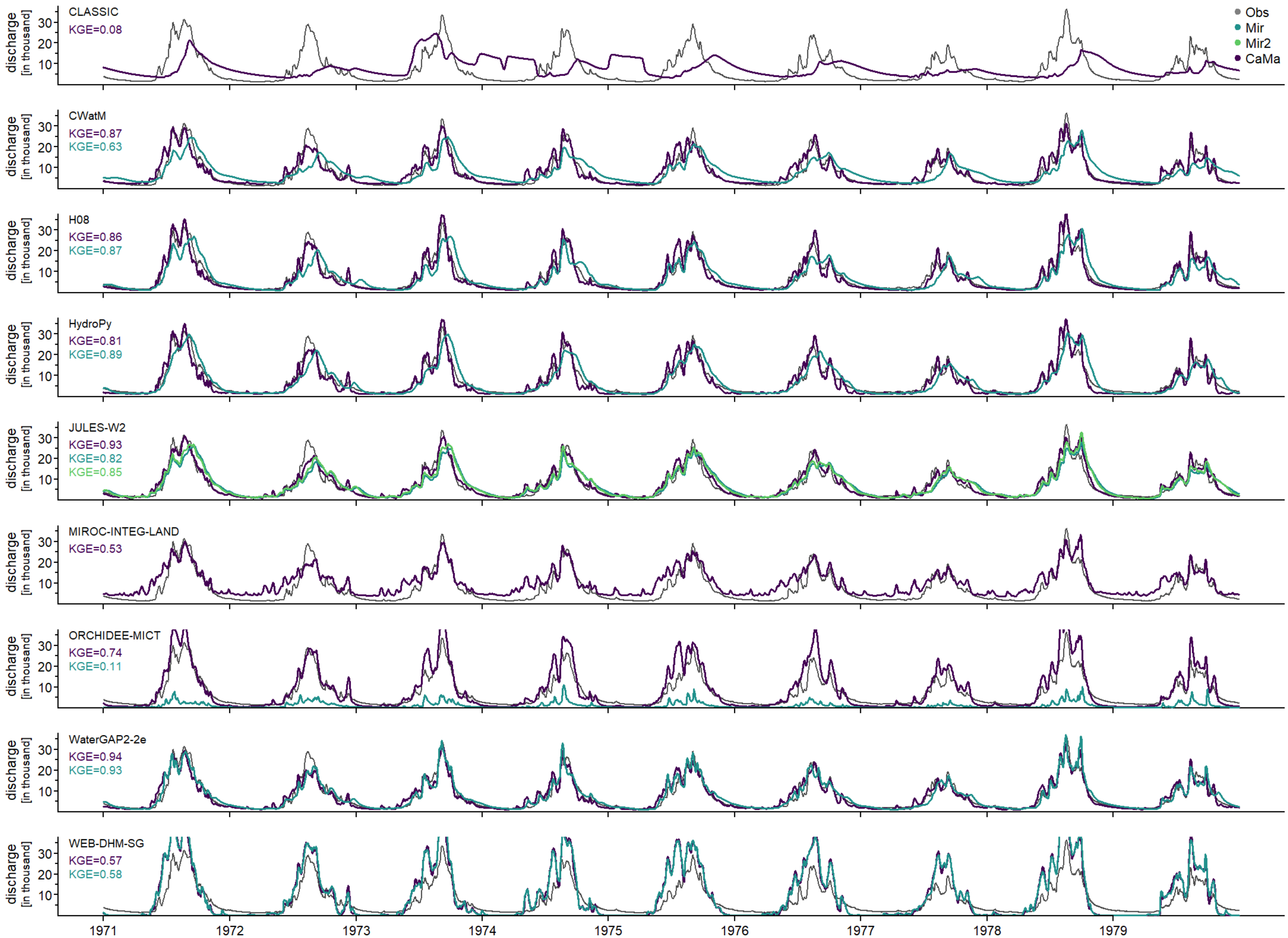

Figure S2. Time-series of daily discharge [m^3/s]for observation (Obs), and simulations by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) for a station for which all models had a high correlation, Mukdahan station, Mekong river, Thailand. JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

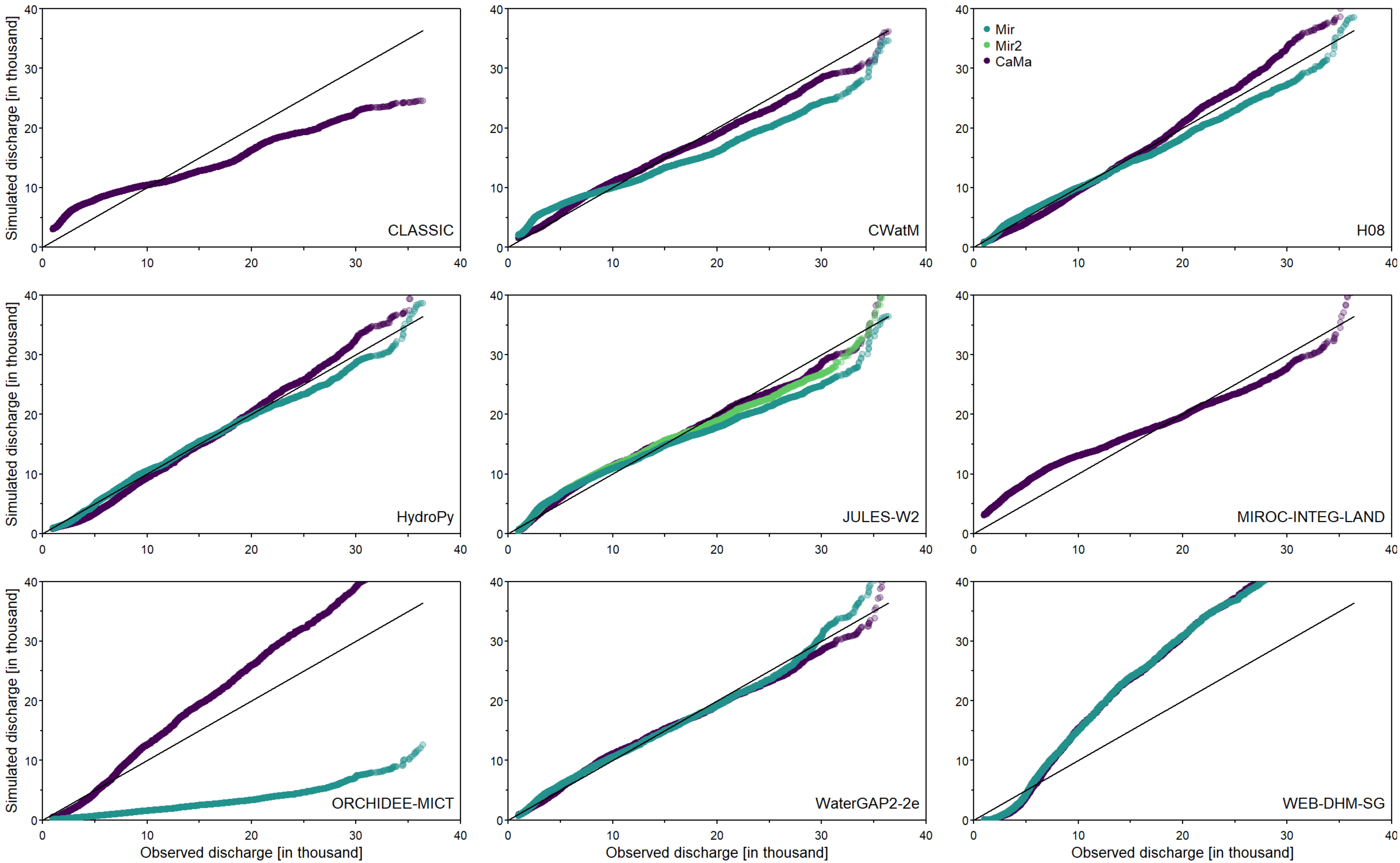

Figure S3. QQ-plot of daily discharge [m^3/s] for the same station as in Fig 2 comparing the distribution of observed values (x-axis) and simulated values (y-axis). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

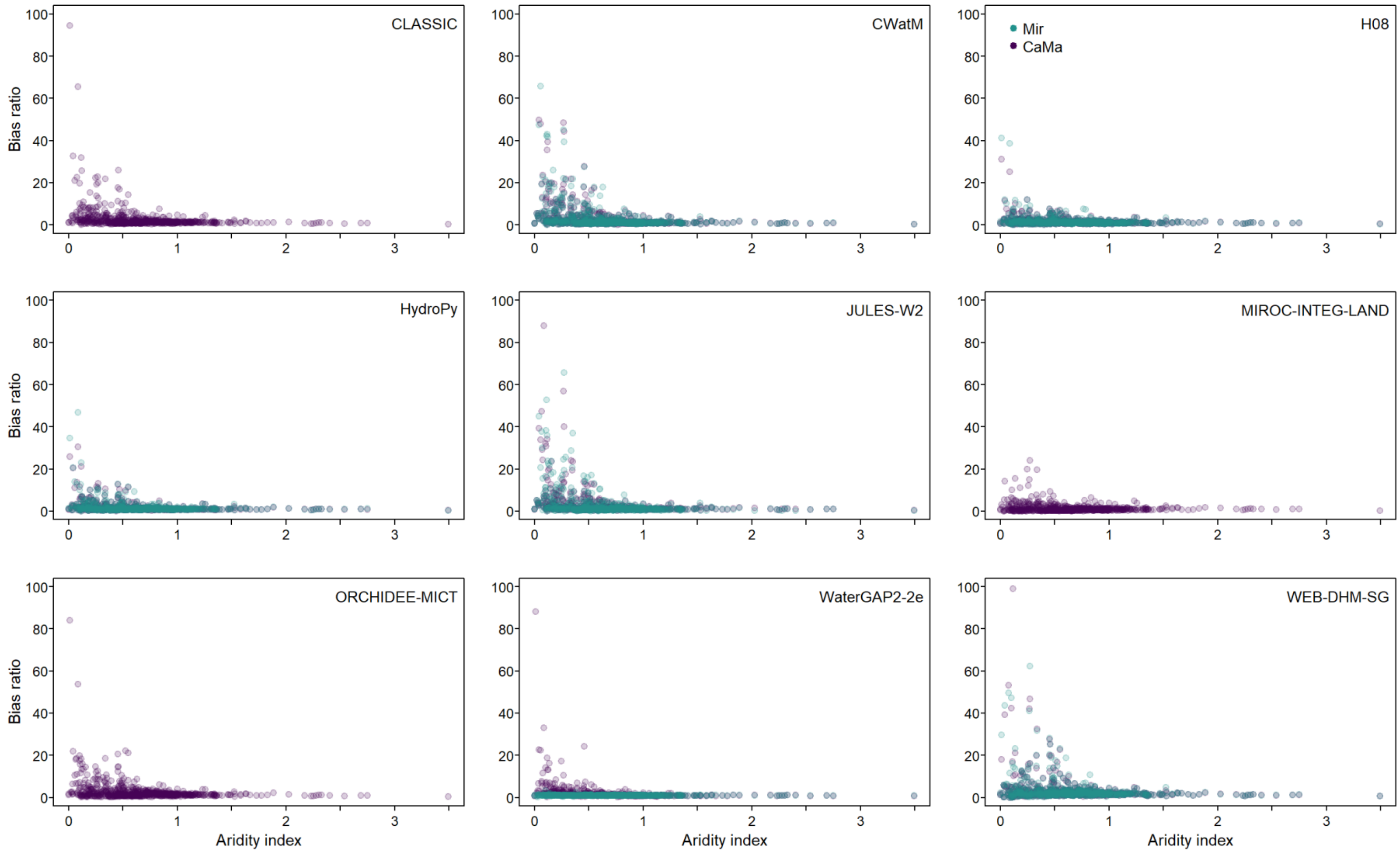

Figure S4. Relationship between aridity and bias ratio (2nd component of KGE) for daily discharge simulated by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) across 644 stations. Aridity index is the ratio of precipitation to potential evapotranspiration (unitless, lower values correspond to more arid conditions, details in Zomer et al., 2022).

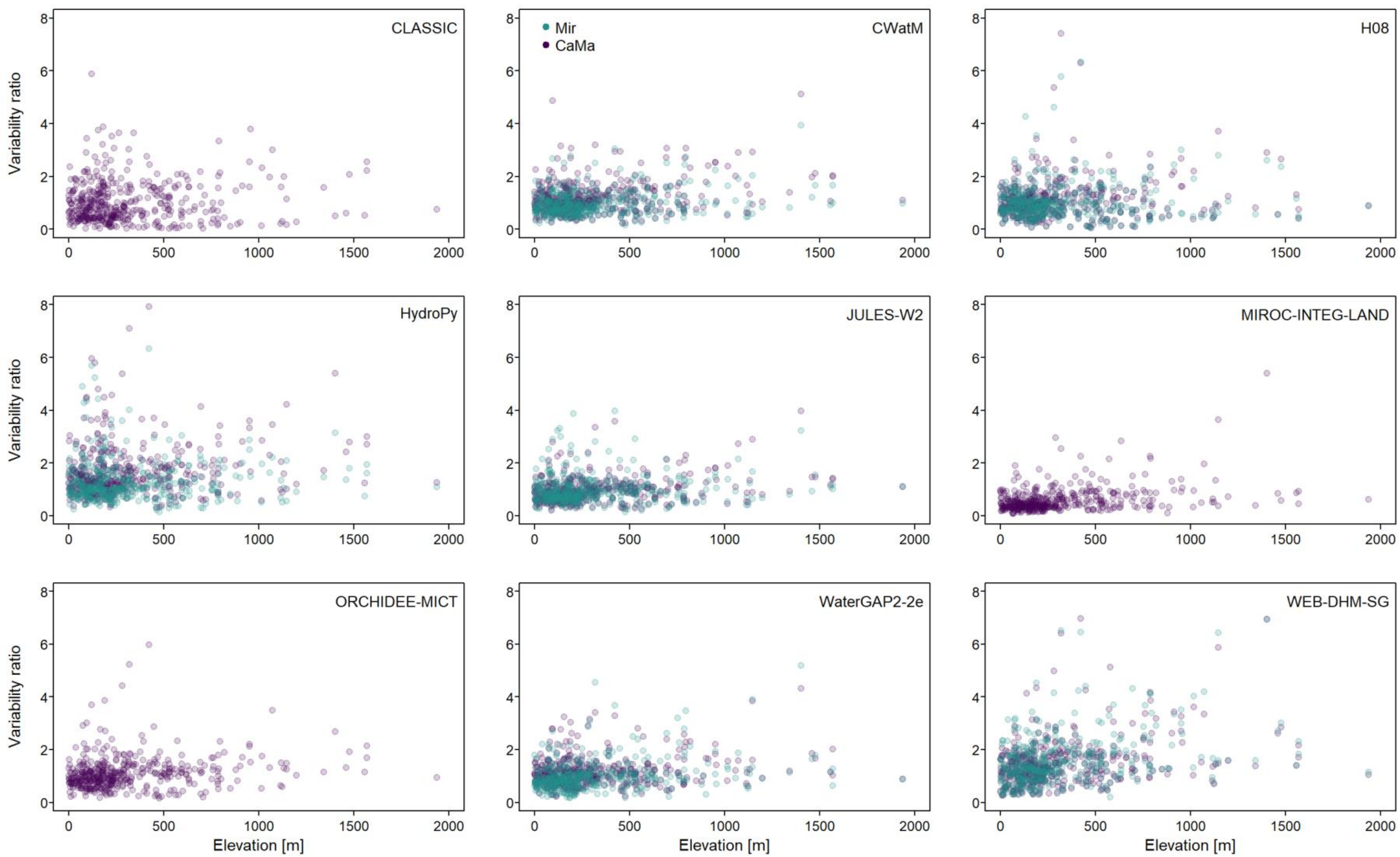

Figure S5. Relationship between mean catchment elevation and variability ratio (3rd component of KGE) for daily discharge simulated by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) across 436 stations.

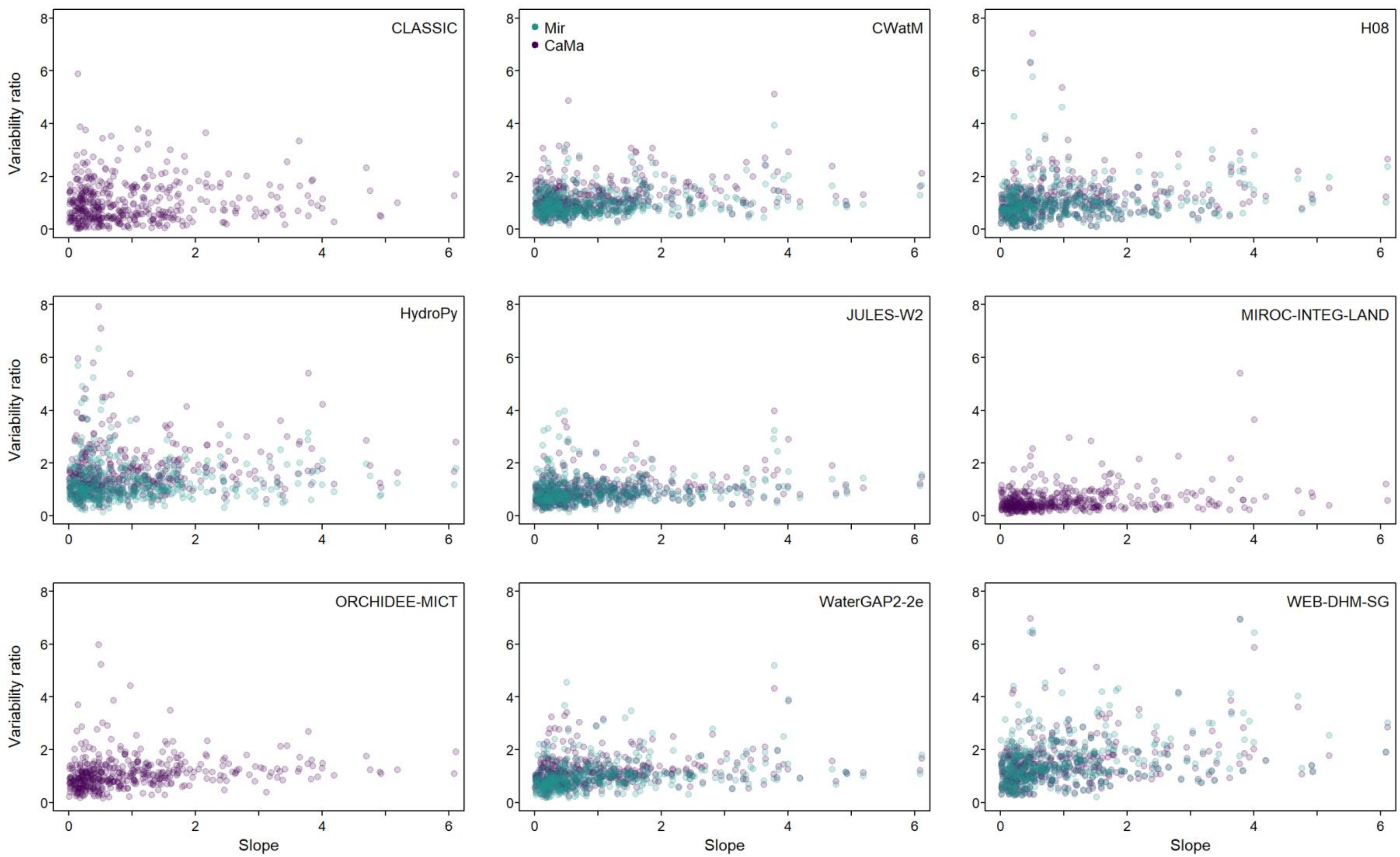

Figure S6. Relationship between mean catchment slope and variability ratio (3rd component of KGE) for daily discharge simulated by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) across 436 stations.

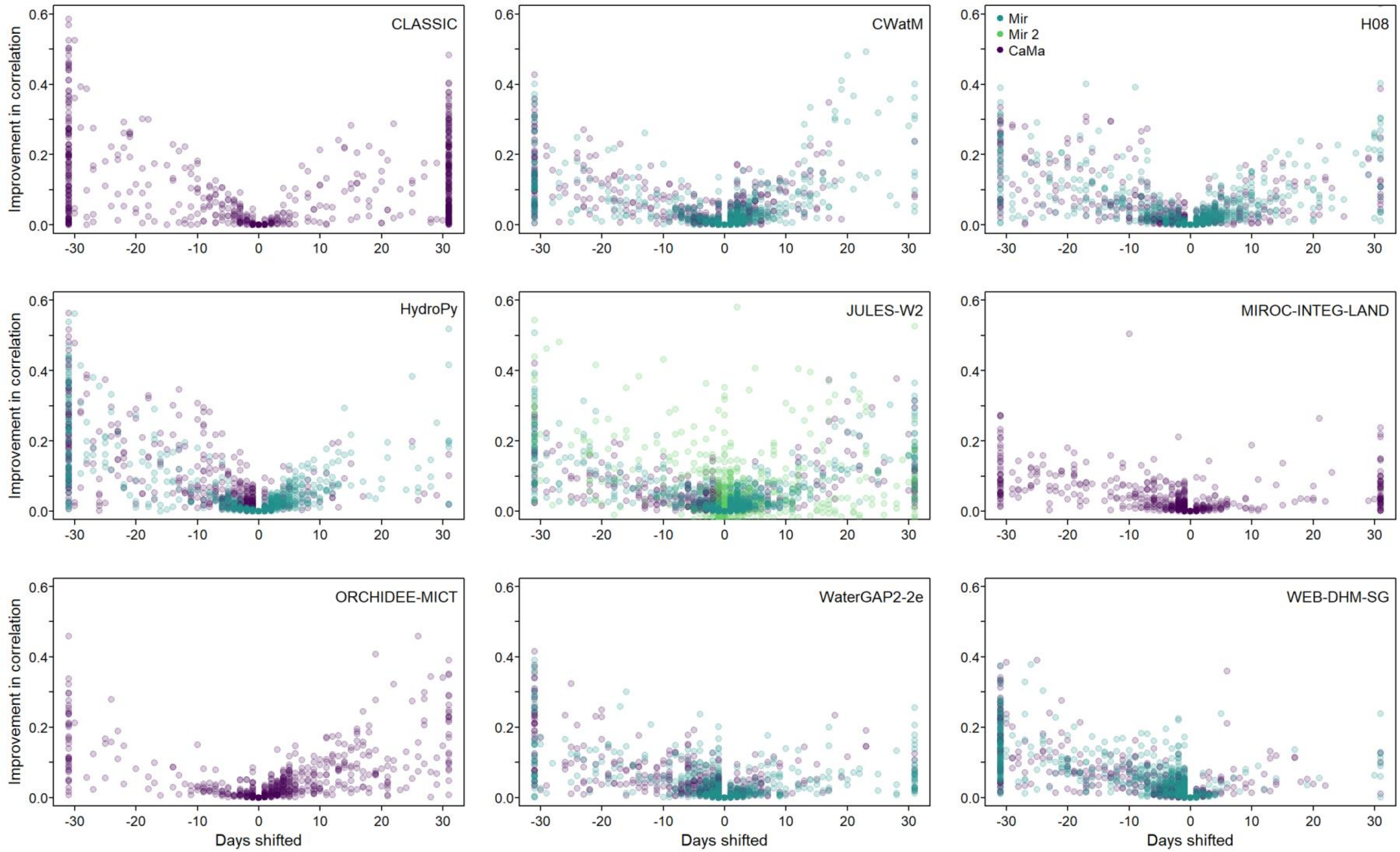

Figure S7. Relationship between days by which simulated time-series were shifted relative to observed time-series of daily discharge, and to which extent this shifting improved correlation (e.g., difference between correlation for shifting 0 days to number of days with highest correlation across 644 stations. Discharge simulated by model's internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

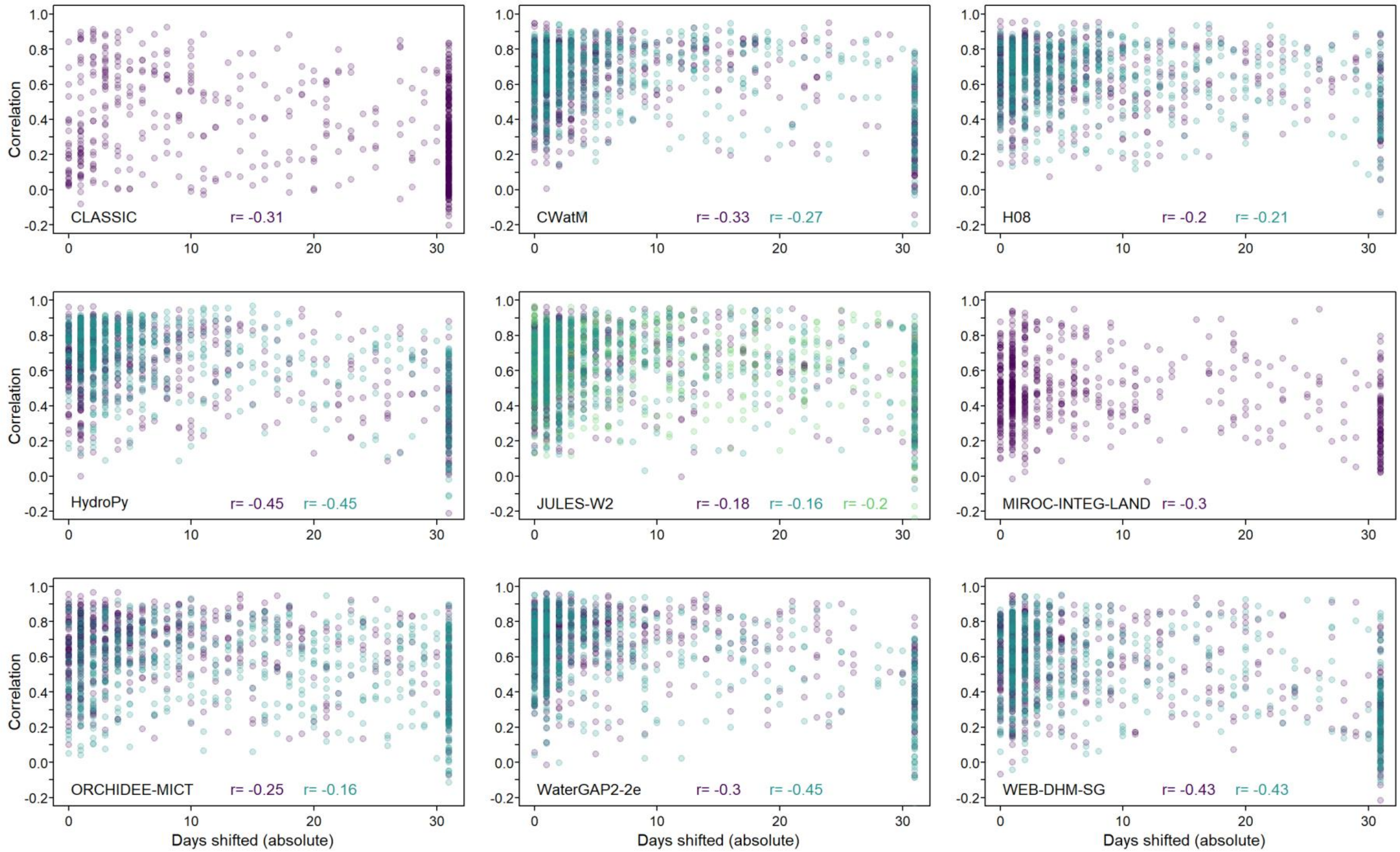

Figure S8. Relationship between days by which simulated time-series were shifted relative to observed time-series of daily discharge, and correlation of shifted time-series (i.e., 1st component of KGE) across 644 stations. Discharge simulated by model's internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

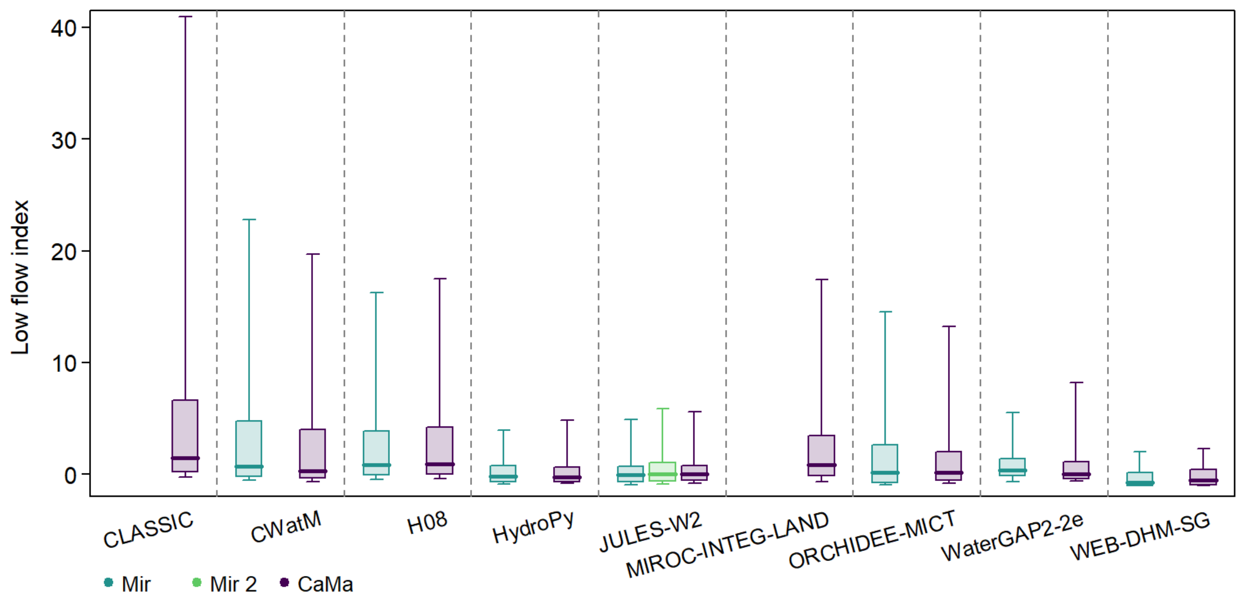

Figure S9. Evaluation of daily discharge simulated by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) across 644 stations for low flow. JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1). Low flow index of zero implies perfect match between observation and simulation, negative values imply that low flow is simulated as too low, and positive values that low flow is simulated too high. Thick lines: median, box: first quartile and third quartile, whiskers: 10th and 90th percentile.

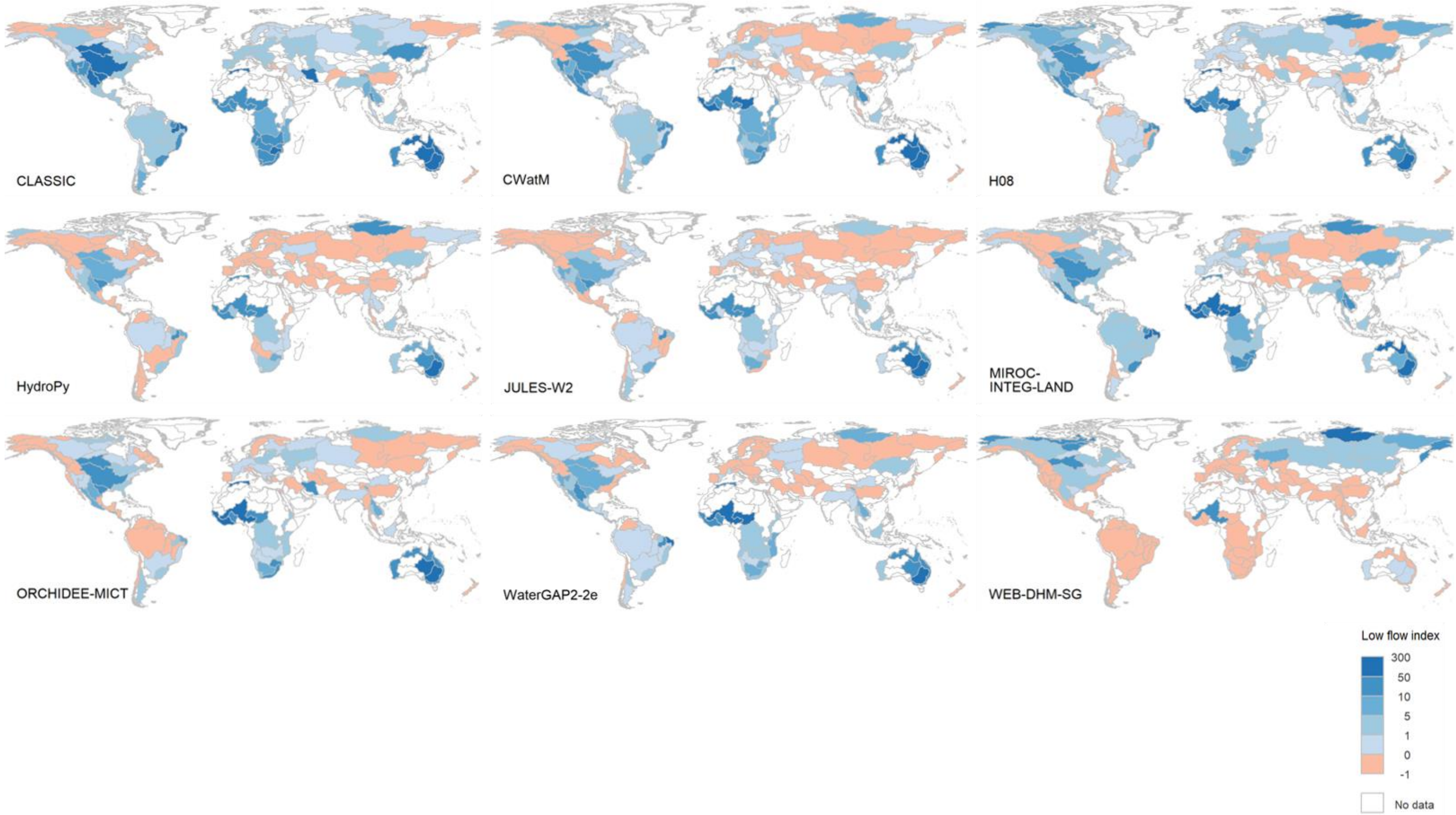

Figure S10. Low flow index (Q90) derived from daily discharge routed by CaMa-Flood, shown as basin average. Low flow index of zero implies perfect match between observation and simulation, negative values imply that low flow is simulated as too low, and positive values that low flow is simulated too high.

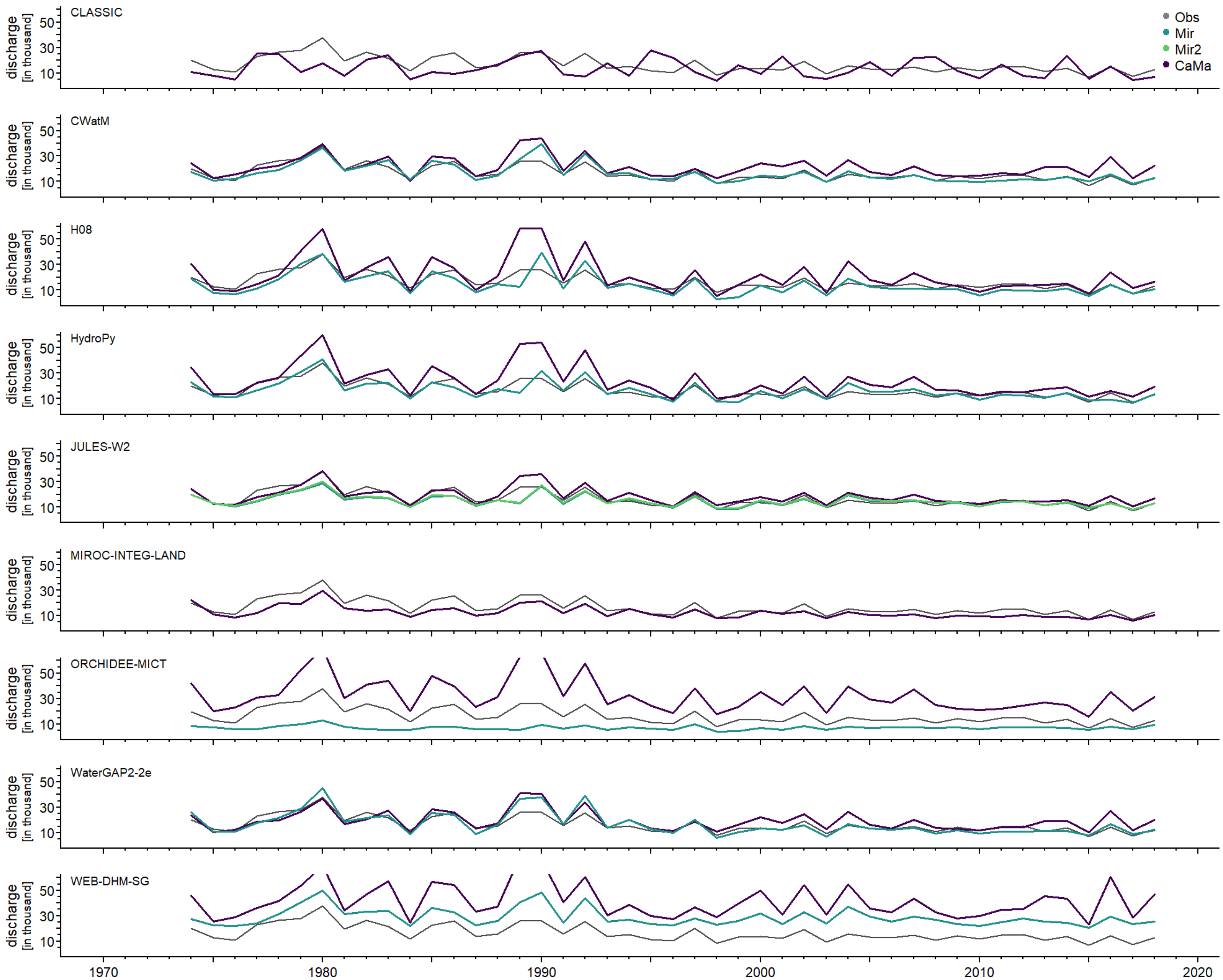

Figure S11. Time-series of maximum annual discharge [m^3/s] for observation (Obs), and simulations by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) for a station for which all models had a high correlation, Descarreto station, Rio Tocantins, Brazil). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

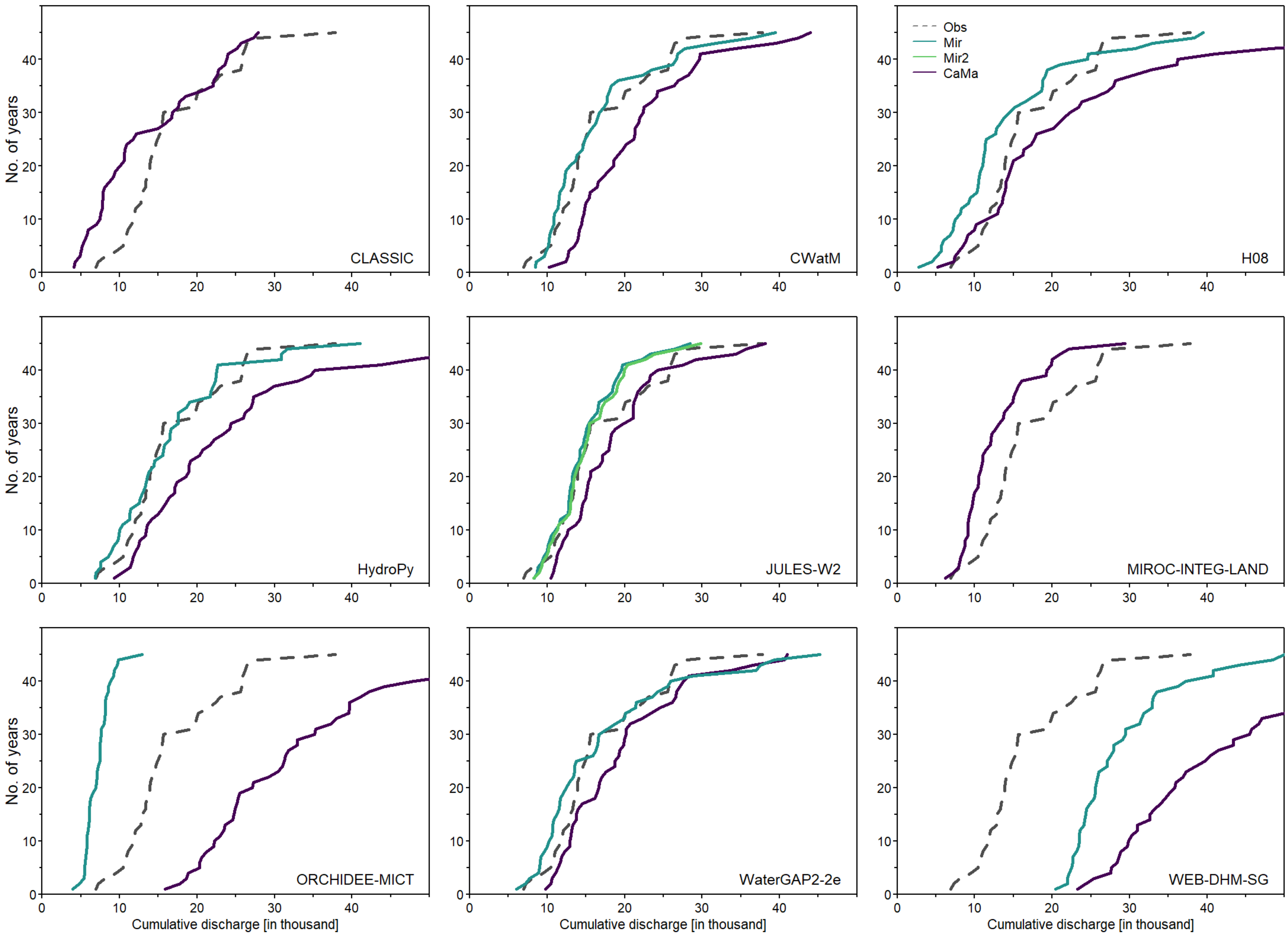

Figure S12. Cumulative plot of maximum annual discharge [m^3/s] for the same station as in figure S8 for observation (Obs), and simulations by model's internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1). When the line for simulated discharge is to the left of the 'Obs' line then discharge is underestimated, and when the line is to the right of the 'Obs' line then discharge is overestimated.

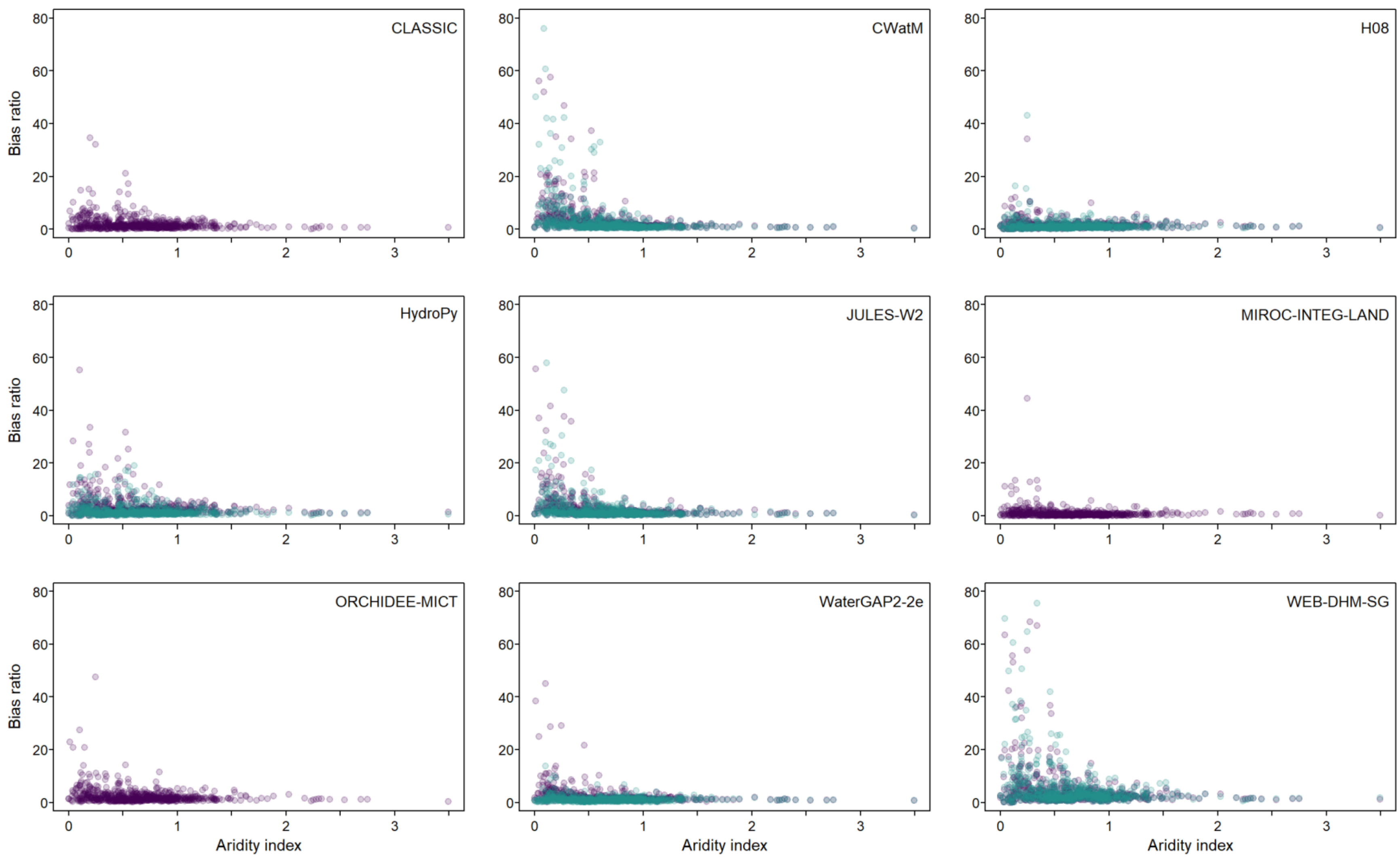

Figure S13. Relationship between aridity and bias ratio (2nd component of KGE) for maximum annual discharge simulated by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) across 644 stations. Aridity index is the ratio of precipitation to potential evapotranspiration (unitless, lower values correspond to more arid conditions, details in Zomer et al., 2022).

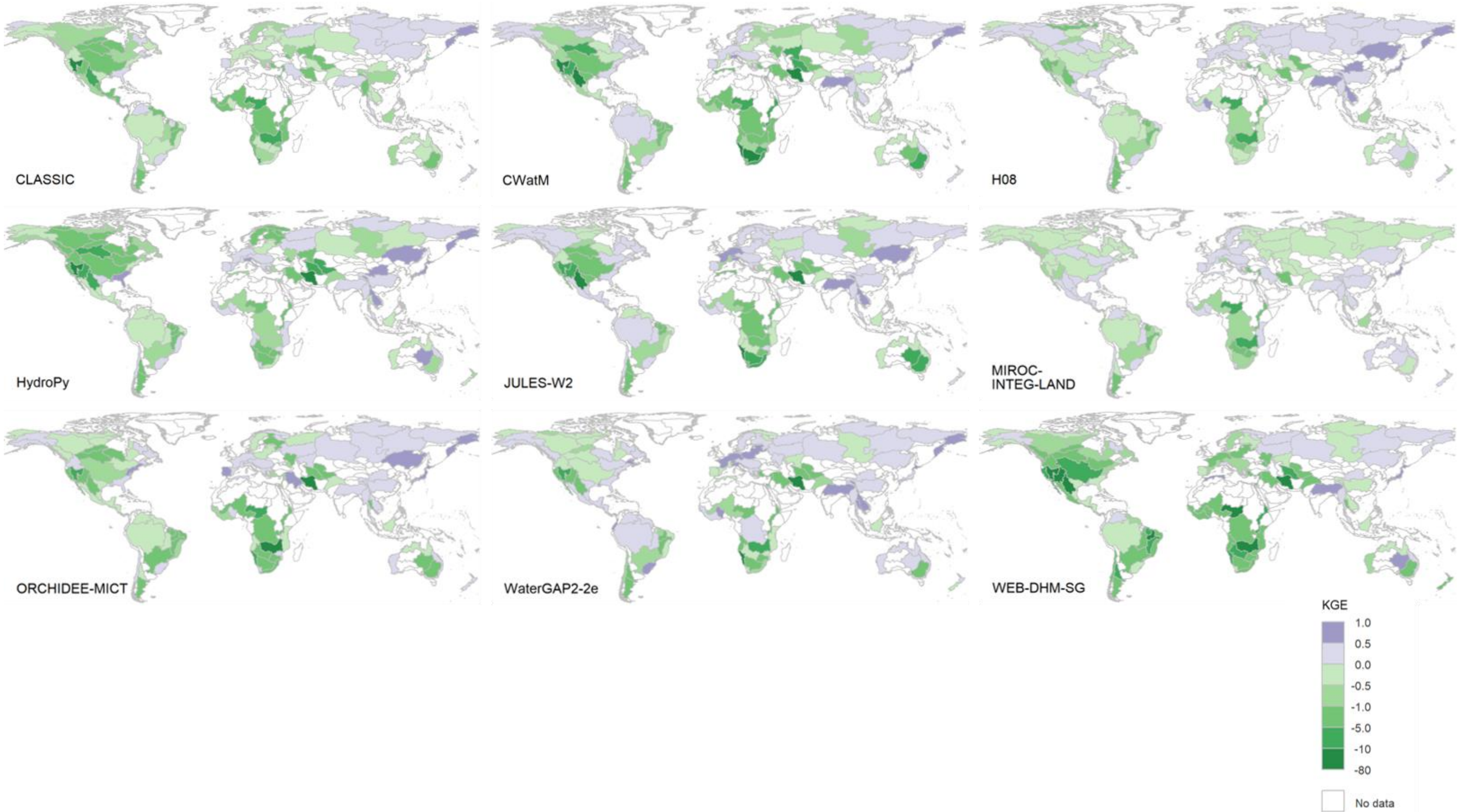

Figure S14. Kling-Gupta efficiency (KGE) for maximum annual discharge routed by CaMa-Flood, shown as basin average.

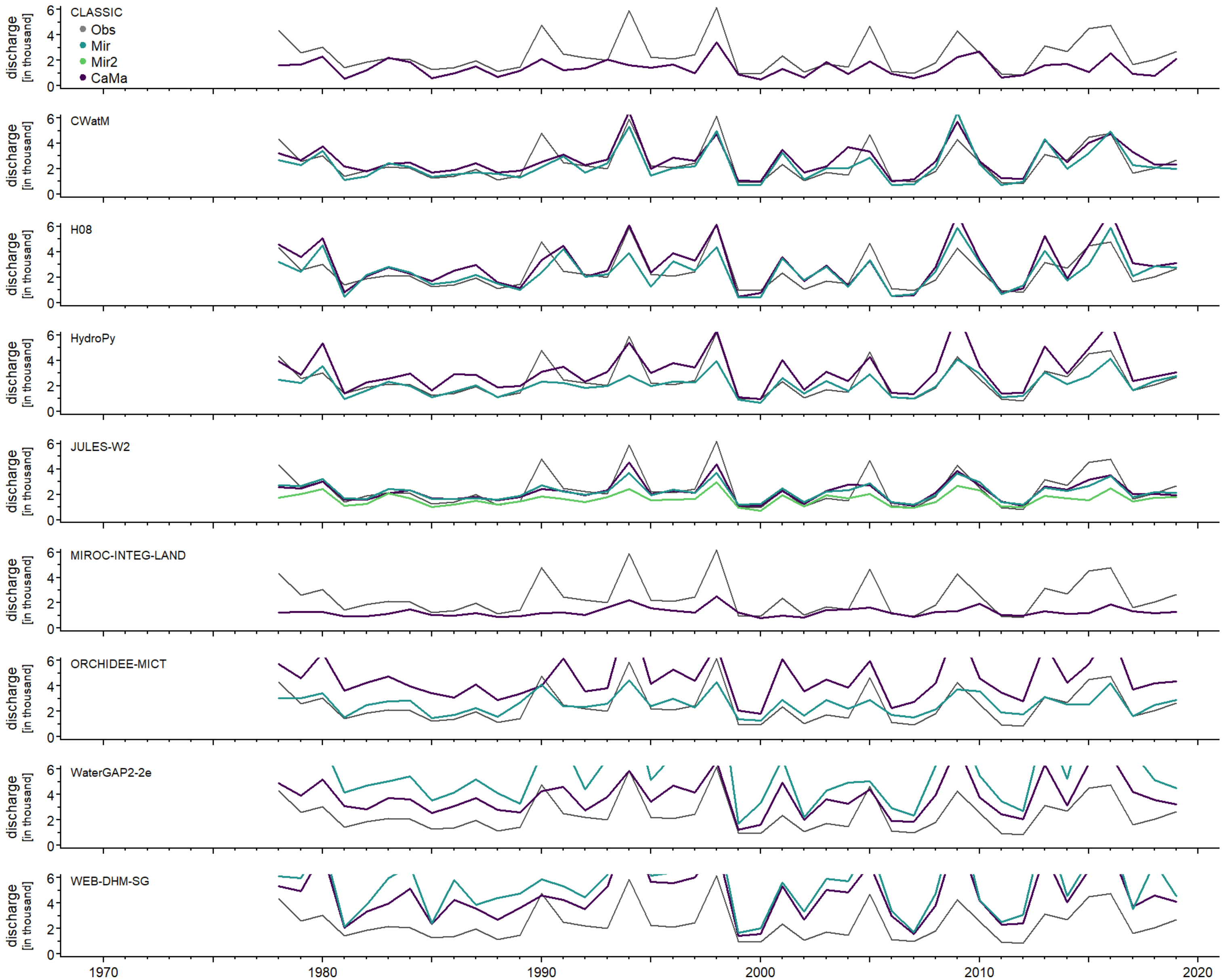

Figure S15. Time-series of maximum annual discharge [m^3/s] for observation (Obs), and simulations by model's internal routing scheme (Mir) and CaMa-Flood (CaMa) for one of the few cases where models (here, JULES-W2, ORCHIDEE) overestimate low values and underestimate high values (Near Sumatra, Florida, Apalachicola River, USA). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

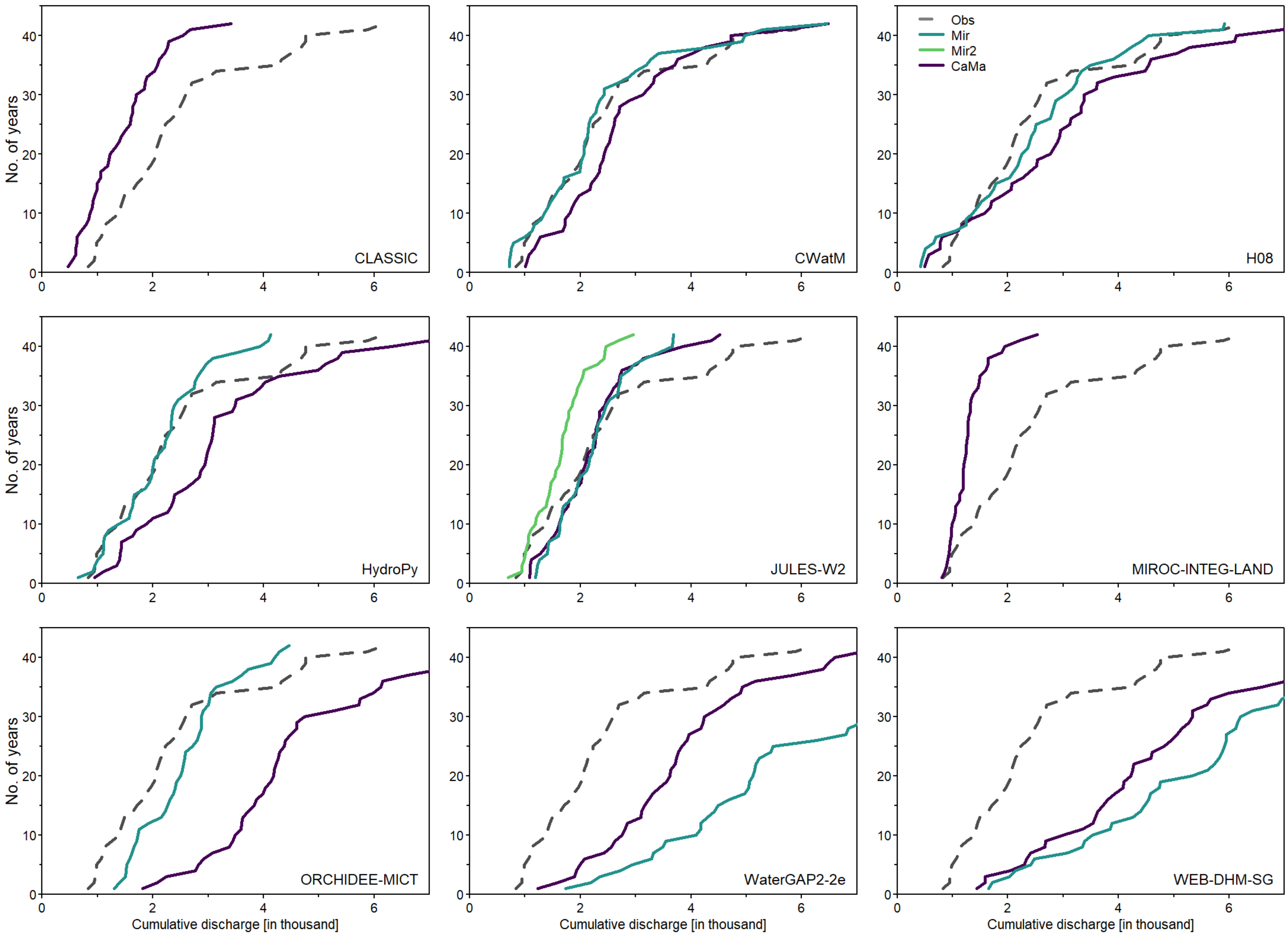

Figure S16. Cumulative plot of maximum annual discharge [m^3/s] for the same station as in figure S11 for observation (Obs), and simulations by model's internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1). Here it can be seen how JULES-W2 and ORCHIDEE overestimate low values and underestimate high values.

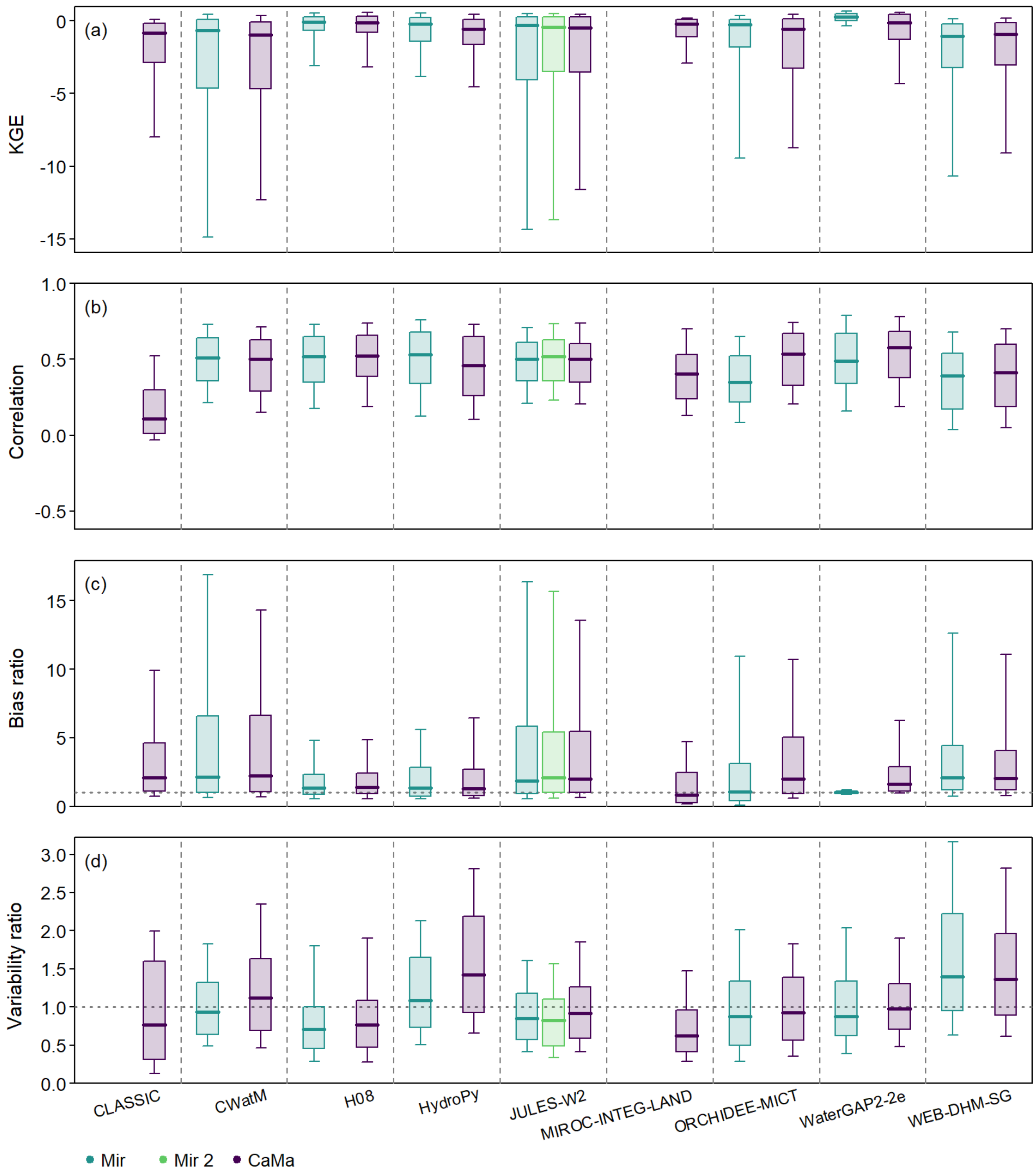

Figure S17. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for daily discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

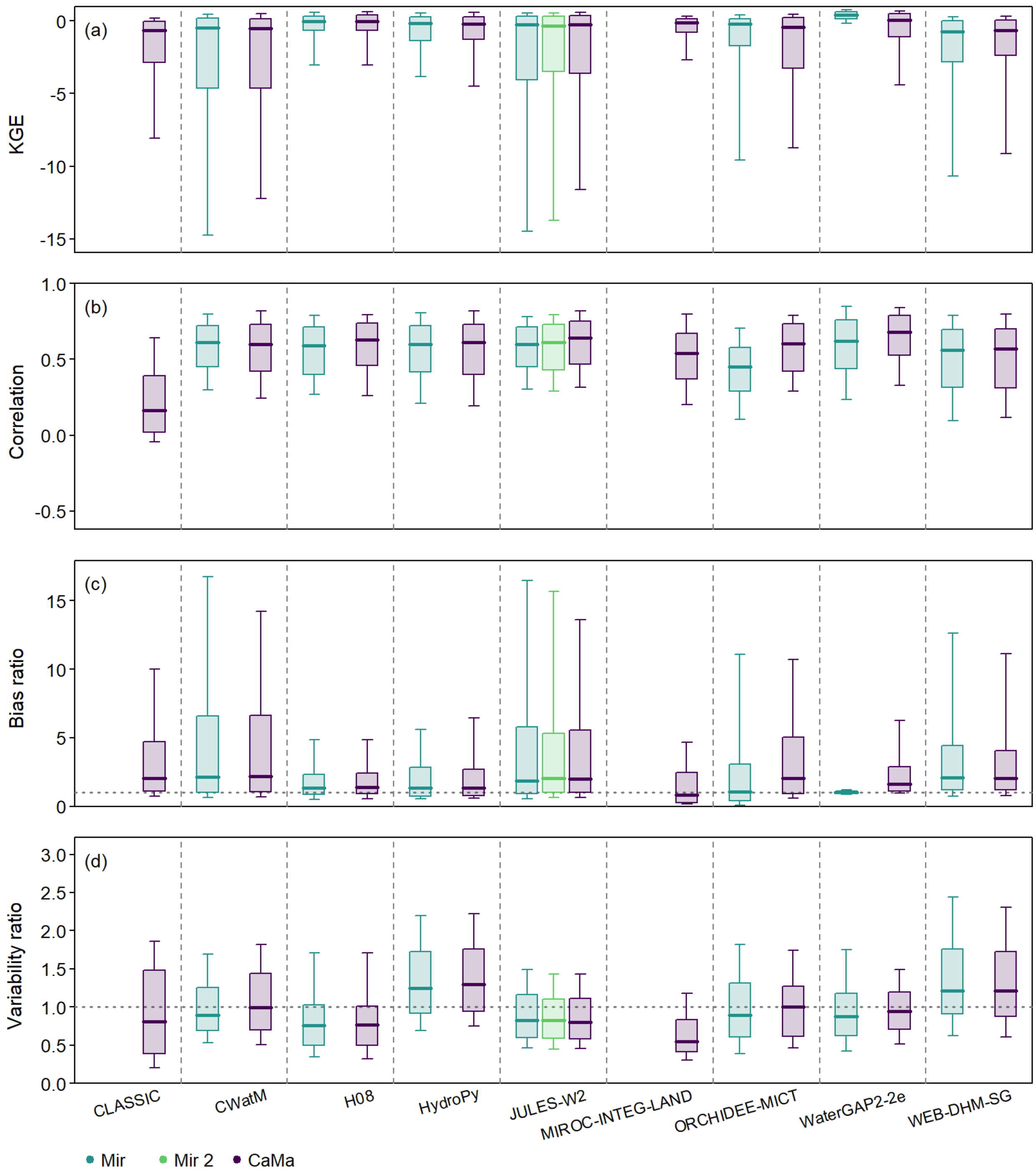

Figure S18. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for mean monthly discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

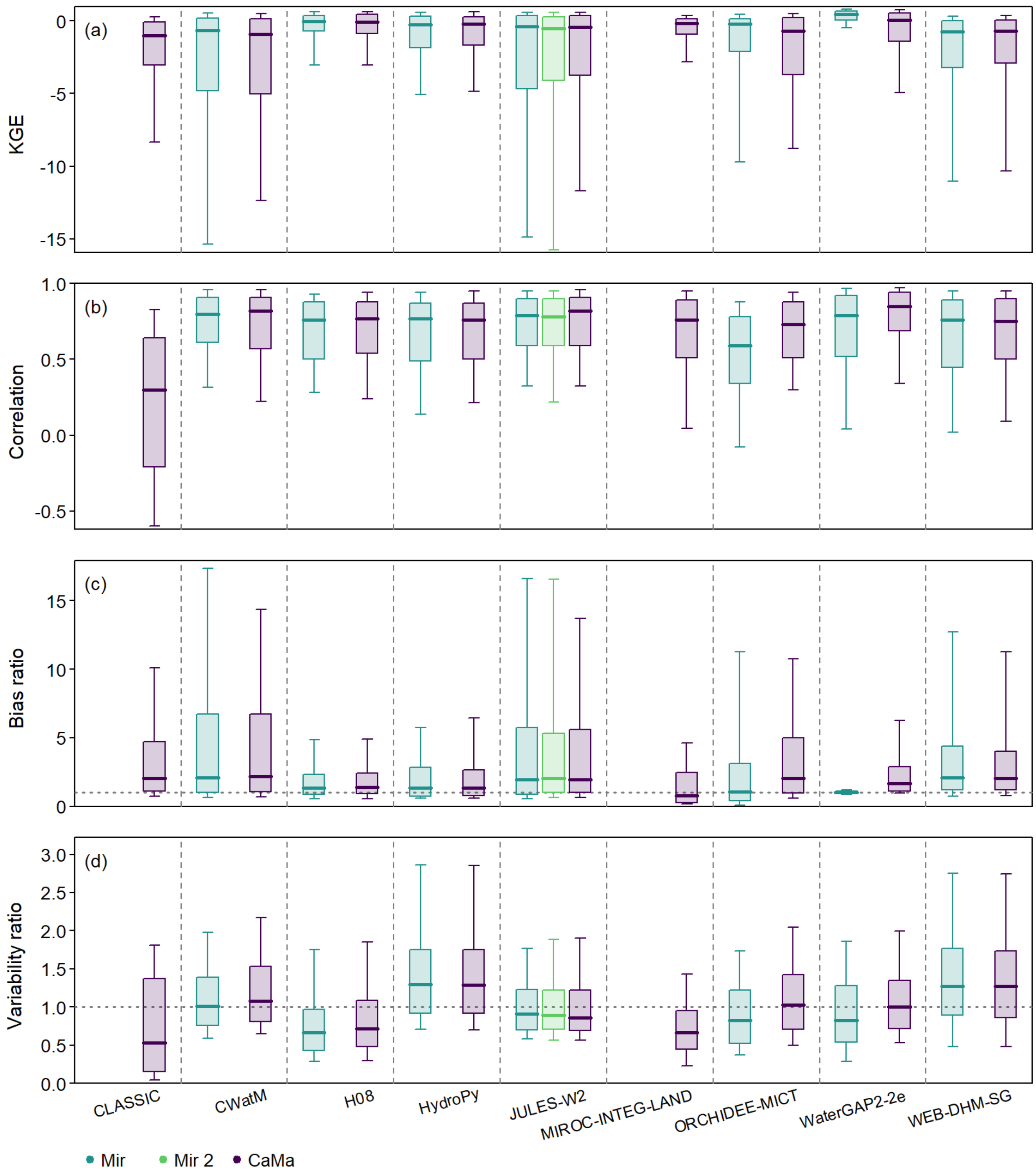

Figure S19. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for long-term mean monthly discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

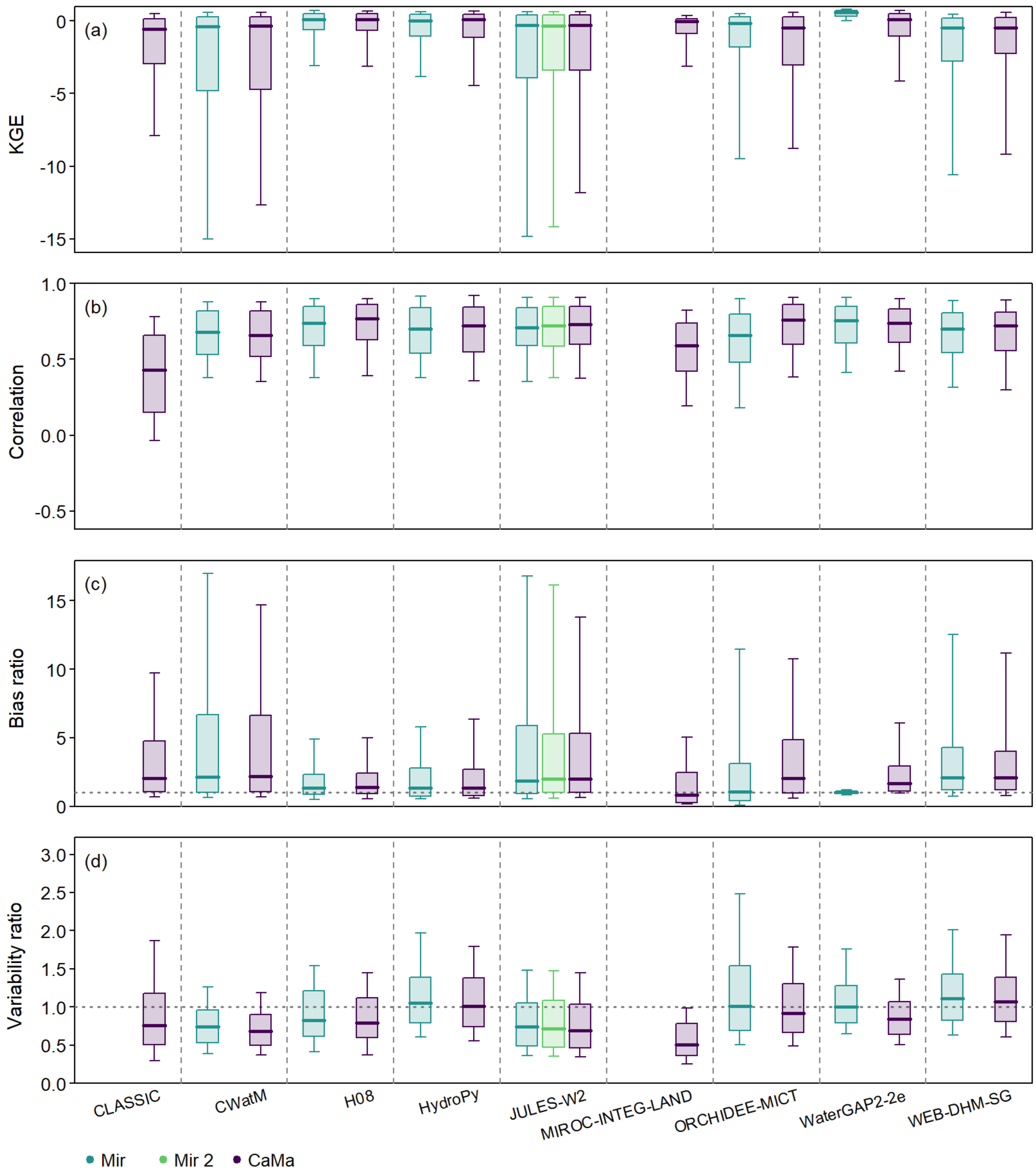

Figure S20. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for mean annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

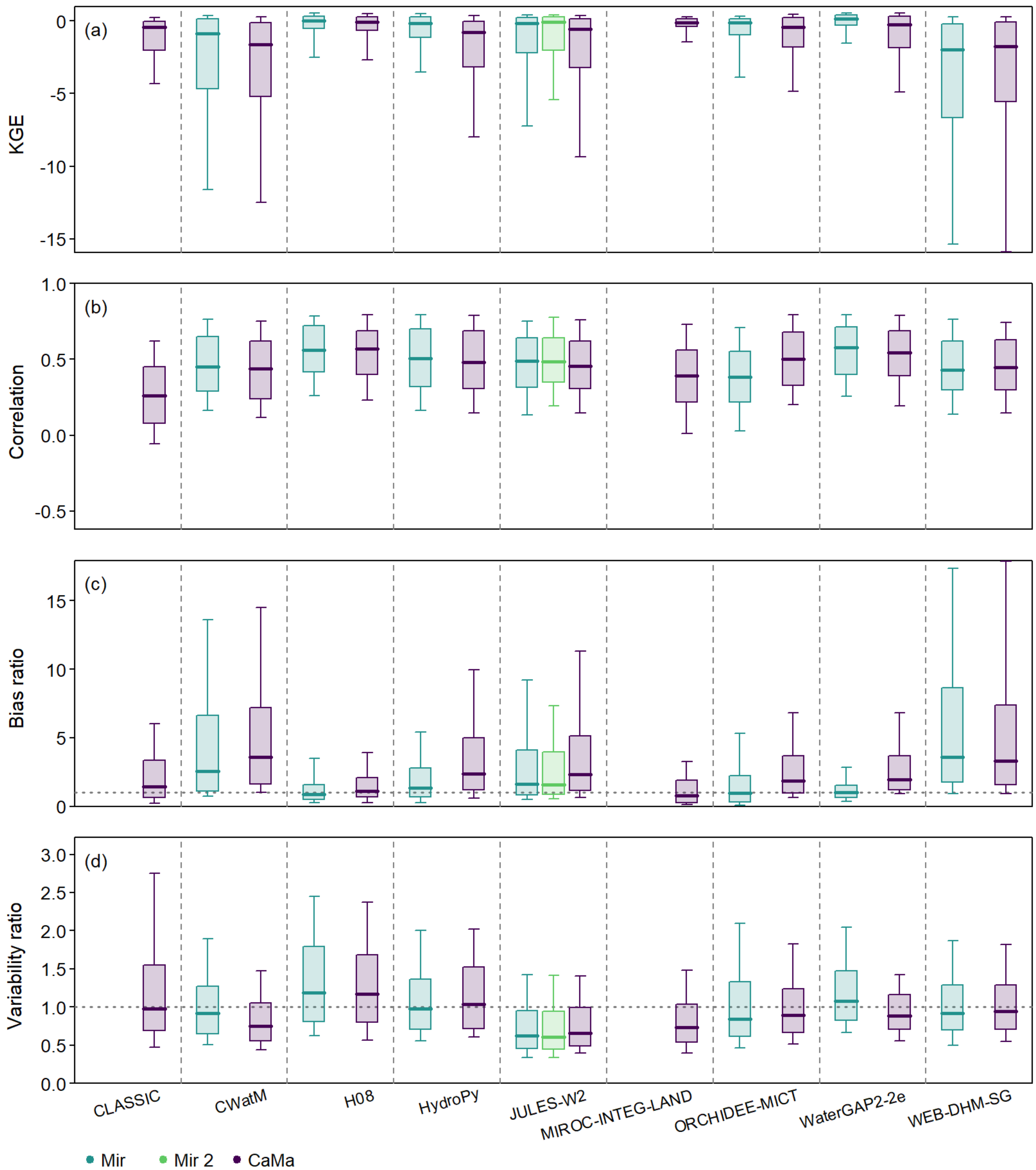

Figure S21. Arid stations (n=236). Kling–Gupta efficiency (KGE) and its three components: correlation (r), bias ratio (β), and variability ratio (γ) for maximum annual discharge simulated with models' internal routing scheme (Mir) and CaMa-Flood (CaMa). JULES-W2 provides discharge routed with two routing schemes (Mir, Mir2, details in table S1).

References

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., & Harding, R. J. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geoscientific Model Development*, *4*(3), 677–699. https://doi.org/10.5194/gmd-4-677-2011

Burek, P., Satoh, Y., Kahil, T., Tang, T., Greve, P., Smilovic, M., Guillaumot, L., Zhao, F., & Wada, Y. (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), 3267–3298. https://doi.org/10.5194/gmd-13-3267-2020

Guimberteau, M., Ducharne, A., Ciais, P., Boisier, J. P., Peng, S., De Weirdt, M., & Verbeeck, H. (2014). Testing conceptual and physically based soil hydrology schemes against observations for the Amazon Basin. *Geoscientific Model Development*, *7*(3), 1115–1136. https://doi.org/10.5194/gmd-7-1115-2014

Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., Ottlé, C., Jornet-Puig, A., Bastos, A., Laurent, P., Goll, D., Bowring, S., Chang, J., Guenet, B., Tifafi, M., Peng, S., Krinner, G., Ducharne, A., Wang, F., … Ciais, P. (2018). ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: Model description and validation. *Geoscientific Model Development*, *11*(1), 121–163. https://doi.org/10.5194/gmd-11-121-2018

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., & Kanae, S. (2018). A global hydrological simulation to specify the sources of water used by humans. *Hydrology and Earth System Sciences*, *22*(1), 789–817. https://doi.org/10.5194/hess-22-789-2018

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, *27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Melton, J. R., Arora, V. K., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., & Teckentrup, L. (2020). CLASSIC v1.0: The open-source community successor to the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) – Part 1: Model framework and site-level performance. *Geoscientific Model Development*, *13*(6), 2825–2850. https://doi.org/10.5194/gmd-13-2825-2020

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T. A., Popat, E., Portmann, F. T., Reinecke, R., Schumacher, M., Shadkam, S., Telteu, C.-E., Trautmann, T., & Döll, P. (2021). The global water resources and use model WaterGAP v2.2d: Model description and evaluation. *Geoscientific Model Development*, *14*(2), 1037–1079. https://doi.org/10.5194/gmd-14-1037-2021

Müller Schmied, H., Trautmann, T., Ackermann, S., Cáceres, D., Flörke, M., Gerdener, H., Kynast, E., Peiris, T. A., Schiebener, L., Schumacher, M., & Döll, P. (2023). The global water resources and use model WaterGAP v2.2e: Description and evaluation of modifications and new features. *Geoscientific Model Development Discussions*, 1–46. https://doi.org/10.5194/gmd-2023-213

Qi, W., Feng, L., Yang, H., Liu, J., Zheng, Y., Shi, H., Wang, L., & Chen, D. (2022). Economic growth dominates rising potential flood risk in the Yangtze River and benefits of raising dikes from 1991 to 2015. *Environmental Research Letters*, *17*(3), 034046. https://doi.org/10.1088/1748-9326/ac5561

Stacke, T., & Hagemann, S. (2021). HydroPy (v1.0): A new global hydrology model written in Python. *Geoscientific Model Development*, *14*(12), 7795–7816. https://doi.org/10.5194/gmd-14-7795-2021

Yokohata, T., Kinoshita, T., Sakurai, G., Pokhrel, Y., Ito, A., Okada, M., Satoh, Y., Kato, E., Nitta, T., Fujimori, S., Felfelani, F., Masaki, Y., Iizumi, T., Nishimori, M., Hanasaki, N., Takahashi, K., Yamagata, Y., & Emori, S. (2020). MIROC-INTEG-LAND version 1: A global biogeochemical land surface model with human water management, crop growth, and land-use change. *Geoscientific Model Development*, *13*(10), 4713–4747. https://doi.org/10.5194/gmd-13-4713-2020

Zomer, R. J., Xu, J., & Trabucco, A. (2022). Version 3 of the Global Aridity Index and Potential Evapotranspiration Database. *Scientific Data*, *9*(1), 1. https://doi.org/10.1038/s41597-022-01493-1