Supporting Information for

A novel hybrid fine-tuning method for supercharging deep learning model development for hydrological prediction

M.S. Jahangir[1,2], J. Quilty[3], C. Shen[4], A. Scott[5], S. Steinschneider[2], J. Adamowski[1]
[1]Department of Bioresource Engineering, McGill University, Montreal, H9X 3V9, Canada
[2]Department of Biological and Environmental Engineering, Cornell University, Ithaca, 14853, United States
[3]Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, N2L3G1, Canada
[4]Department of Civil and Environmental Engineering, Pennsylvania State University, College town, 16802, United States
[5]Department of Mechanical Engineering, University of Waterloo, Waterloo, N2L3G1, Canada
Correspondence to: Mohammad Sina Jahangir (mohammad.jahangir@mcgill.ca)

Contents of this file

Introduction

This file includes supporting information related to the main article. In S1, the selected catchments are listed. In S2, the error metrics formulation is provided. Afterward, S3 includes additional information about the models' prediction performance. In S4, details are provided for catchment-wise training strategy.

S1 Data

<u>Selected Catchments' ID: CAMESL-US</u>

The catchments listed below in Table S1 were selected from CAMELS-US.

Table S1- Selected gauges and their associated CAMELS IDs.

Index	Catchment ID
1	1022500
2	1031500
3	1047000
4	1052500
5	1054200
6	1055000
7	1057000
8	1078000
9	1134500
10	1137500
11	1139000
12	1144000
13	1169000
14	1170100
15	1181000
16	1333000
17	1350000
18	1365000
19	1413500
20	1423000
21	1435000
22	1440000
23	1440400
24	1451800
25	1485500
26	1487000
27	1491000
28	1518862
29	1532000
30	1539000
31	1543000
32	1543500
33	1544500
34	1545600
35	1547700

36	1548500
37	1550000
38	1552000
39	1557500
40	1568000
41	1580000
42	1583500
43	1596500
44	1605500
45	1606500
46	1632000
47	1632900
48	1634500
49	1638480
50	1639500
51	1644000
52	1664000
53	1666500
54	1667500
55	1669520
56	2011400
57	2011460
58	2013000
59	2014000
60	2015700
61	2016000
62	2017500
63	2018000
64	2027000
65	2027500
66	2028500
67	2046000
68	2051000
69	2051500
70	2053200
71	2053800
72	2056900
73	2059500
74	2064000
75	2065500
76	2069700
77	2070000

78	2074500
79	2077200
80	2081500
81	2082950
82	2092500
83	2108000
84	2111180
85	2111500
86	2112120
87	2112360
88	2118500
89	2128000
90	2137727
91	2143000
92	2149000
93	2152100
94	2177000
95	2212600
96	2216180
97	2221525
98	2231000
99	2245500
100	2246000
101	2296500
102	2297310
103	2298123
104	2298608
105	2299950
106	2312200
107	2324400
108	2327100
109	2349900
110	2361000
111	2369800
112	2371500
113	2372250
114	2374500
115	2408540
116	2422500
117	2450250
118	2464000
119	2469800

120	2472000
121	2472500
122	2479155
123	2479300
124	2479560
125	2481000
126	2481510
127	3010655
128	3011800
129	3015500
130	3021350
131	3028000
132	3049000
133	3069500
134	3070500
135	3076600
136	3078000
137	3144000
138	3165000
139	3170000
140	3173000
141	3180500
142	3182500
143	3186500
144	3237500
145	3238500
146	3241500
147	3280700
148	3281500
149	3285000
150	3340800
151	3346000
152	3364500
153	3366500
154	3384450
155	3439000
156	3456500
157	3460000
158	3463300
159	3471500
160	3473000
161	3479000

162	3488000
163	3498500
164	3500000
165	3500240
166	3504000
167	3574500
168	3604000
169	4015330
170	4024430
171	4027000
172	4040500
173	4045500
174	4057510
175	4057800
176	4059500
177	4063700
178	4074950
179	4105700
180	4122200
181	4122500
182	4127997
183	4185000
184	4196800
185	4197100
186	4213000
187	4216418
188	4221000
189	4224775
190	4256000
191	4296000
192	5057200
193	5120500
194	5291000
195	5362000
196	5393500
197	5399500
198	5408000
199	5413500
200	5414000
201	5444000
202	5458000
203	5466500

204	5487980
205	5489000
206	5495000
207	5495500
208	5501000
209	5503800
210	5507600
211	5508805
212	5525500
213	5556500
214	5584500
215	5592050
216	5593575
217	5593900
218	5595730
219	6224000
220	6280300
221	6289000
222	6332515
223	6339100
224	6344600
225	6350000
226	6352000
227	6404000
228	6406000
229	6409000
230	6431500
231	6447500
232	6470800
233	6477500
234	6479215
235	6479438
236	6601000
237	6622700
238	6623800
239	6632400
240	6803510
241	6803530
242	6814000
243	6847900
244	6853800
245	6876700

246	6878000
247	6885500
248	6888500
249	6889200
250	6889500
251	6892000
252	6903400
253	6910800
254	6917000
255	6918460
256	6921070
257	6921200
258	7057500
259	7060710
260	7066000
261	7142300
262	7145700
263	7167500
264	7180500
265	7184000
266	7196900
267	7197000
268	7208500
269	7261000
270	7291000
271	7299670
272	7301410
273	7315200
274	7315700
275	7335700
276	7340300
277	7346045
278	7362100
279	7375000
280	8013000
281	8014500
282	8023080
283	8029500
284	8066200
285	8066300
286	8070000
287	8070200

288	8082700
289	8086212
290	8086290
291	8101000
292	8104900
293	8109700
294	8150800
295	8158700
296	8164300
297	8164600
298	8165300
299	8171300
300	8175000
301	8178880
302	8189500
303	8190000
304	8190500
305	8194200
306	8195000
307	8196000
308	8200000
309	8202700
310	8269000
311	8324000
312	8377900
313	8378500
314	8380500
315	9081600
316	9210500
317	9223000
318	9312600
319	9352900
320	9386900
321	9404450
322	9430600
323	9447800
324	9480000
325	9492400
326	9494000
327	9497980
328	9505350
329	9505800

330	9510200
331	9512280
332	9513780
333	10172700
334	10205030
335	10234500
336	10242000
337	10249300
338	10258500
339	10259200
340	10263500
341	10329500
342	10396000
343	11124500
344	11143000
345	11148900
346	11151300
347	11162500
348	11176400
349	11224500
350	11230500
351	11253310
352	11264500
353	11266500
354	11274500
355	11274630
356	11381500
357	11383500
358	11451100
359	11468500
360	11473900
361	11478500
362	11480390
363	11481200
364	11482500
365	11522500
366	11523200
367	11528700
368	11532500
369	12010000
370	12013500
371	12020000

372	12025700
373	12035000
374	12040500
375	12041200
376	12043000
377	12054000
378	12056500
379	12082500
380	12092000
381	12115000
382	12143600
383	12144000
384	12147500
385	12167000
386	12175500
387	12186000
388	12189500
389	12381400
390	12390700
391	12411000
392	12451000
393	12488500
394	13011500
395	13011900
396	13023000
397	13083000
398	13161500
399	13235000
400	13240000
401	13313000
402	13331500
403	14020000
404	14096850
405	14137000
406	14154500
407	14158500
408	14166500
409	14182500
410	14185000
411	14185900
412	14216500
413	14222500

414	14236200
415	14301000
416	14305500
417	14306500
418	14309500
419	14316700
420	14325000
421	14400000

Selected Catchments’ ID: CAMESL-DE

The catchments listed below in Table S2 were selected from CAMELS-DE.

Table S2- Selected gauges and their associated CAMELS IDs.

Index	Gauge ID
1	DE110020
2	DE110030
3	DE110040
4	DE110060
5	DE110130
6	DE110140
7	DE110170
8	DE110250
9	DE110280
10	DE110290
11	DE110300
12	DE110310
13	DE110320
14	DE110350
15	DE110370
16	DE110410
17	DE110460
18	DE110510
19	DE110520
20	DE110530
21	DE110570
22	DE110920
23	DE110970
24	DE111000
25	DE111010
26	DE111070
27	DE111120
28	DE111360

29	DE112070
30	DE210300
31	DE210310
32	DE210320
33	DE210350
34	DE210370
35	DE210430
36	DE210450
37	DE210460
38	DE210470
39	DE210500
40	DE210520
41	DE210540
42	DE210580
43	DE210590
44	DE210610
45	DE210720
46	DE210730
47	DE210830
48	DE210840
49	DE210860
50	DE210870
51	DE210880
52	DE210920
53	DE210940
54	DE211010
55	DE211070
56	DE211080
57	DE211110
58	DE211120
59	DE211150
60	DE211160
61	DE211260
62	DE211310
63	DE211320
64	DE211450
65	DE211460
66	DE211470
67	DE211490
68	DE211500
69	DE211510
70	DE211620

71	DE211720
72	DE211730
73	DE211760
74	DE211800
75	DE212130
76	DE212300
77	DE212310
78	DE212320
79	DE212380
80	DE212390
81	DE212460
82	DE212470
83	DE212500
84	DE212540
85	DE212550
86	DE212570
87	DE212590
88	DE212610
89	DE212640
90	DE212650
91	DE212690
92	DE212710
93	DE212760
94	DE212870
95	DE212890
96	DE212900
97	DE212980
98	DE213010
99	DE213050
100	DE213060
101	DE213140
102	DE213150
103	DE213170
104	DE213180
105	DE213360
106	DE213450
107	DE213460
108	DE213480
109	DE213490
110	DE213510
111	DE213520
112	DE213540

113	DE213560
114	DE213570
115	DE213580
116	DE213700
117	DE213710
118	DE213730
119	DE213780
120	DE213800
121	DE213820
122	DE213840
123	DE213850
124	DE213860
125	DE213880
126	DE213890
127	DE213900
128	DE213910
129	DE213920
130	DE213930
131	DE213940
132	DE213970
133	DE214010
134	DE214020
135	DE214030
136	DE214060
137	DE214090
138	DE214110
139	DE214160
140	DE214200
141	DE214220
142	DE214230
143	DE214250
144	DE214290
145	DE214370
146	DE214450
147	DE214470
148	DE214480
149	DE214520
150	DE214580
151	DE214590
152	DE214610
153	DE214630
154	DE214640

155	DE214690
156	DE214720
157	DE214770
158	DE214860
159	DE214880
160	DE214890
161	DE214960
162	DE215010
163	DE215020
164	DE215030
165	DE215040
166	DE215050
167	DE215090
168	DE215100
169	DE215110
170	DE215120
171	DE215150
172	DE215160
173	DE215190
174	DE215280
175	DE215340
176	DE411220
177	DE411320
178	DE411400
179	DE412090
180	DE710030
181	DE710040
182	DE710090
183	DE710100
184	DE710120
185	DE710170
186	DE710180
187	DE710190
188	DE710200
189	DE710240
190	DE710250
191	DE710270
192	DE710280
193	DE710290
194	DE710300
195	DE710310
196	DE710340

197	DE710380
198	DE710400
199	DE710410
200	DE710420
201	DE710430
202	DE710450
203	DE710460
204	DE710470
205	DE710530
206	DE710550
207	DE710560
208	DE710590
209	DE710600
210	DE710650
211	DE710660
212	DE710690
213	DE710700
214	DE710710
215	DE710780
216	DE710800
217	DE710830
218	DE710900
219	DE710930
220	DE710950
221	DE810300
222	DE810320
223	DE810330
224	DE810470
225	DE810570
226	DE810580
227	DE810920
228	DE811610
229	DE812200
230	DEA10000
231	DEA10040
232	DEA10060
233	DEA10160
234	DEA10230
235	DEA10300
236	DEA10310
237	DEA10320
238	DEA10460

239	DEA10480
240	DEA10620
241	DEA10630
242	DEA10640
243	DEA10760
244	DEA10850
245	DEA10910
246	DEA11090
247	DEA11130
248	DEA11180
249	DEA11200
250	DEA11770
251	DEA11790
252	DEA11980
253	DEA12000
254	DEA12010
255	DEA12100
256	DEA12140
257	DED10300
258	DED10480
259	DED10620
260	DED10690
261	DED10700
262	DED10890
263	DED10970
264	DED11050
265	DED11280
266	DED11420
267	DED11460
268	DED12500
269	DED12510
270	DED12520
271	DED12570
272	DED12680
273	DEE10170
274	DEE10200
275	DEE10430
276	DEE10450
277	DEE10550
278	DEE10570
279	DEE10590
280	DEE10610

281	DEE10630
282	DEE10670
283	DEE10680
284	DEE10710
285	DEG10000
286	DEG10010
287	DEG10020
288	DEG10050
289	DEG10060
290	DEG10090
291	DEG10100
292	DEG10110
293	DEG10140
294	DEG10150
295	DEG10160
296	DEG10170
297	DEG10220
298	DEG10230
299	DEG10240
300	DEG10250
301	DEG10280
302	DEG10300
303	DEG10310
304	DEG10320
305	DEG10330
306	DEG10350
307	DEG10370
308	DEG10380
309	DEG10390
310	DEG10400
311	DEG10420
312	DEG10430
313	DEG10440
314	DEG10450
315	DEG10470
316	DEG10510
317	DEG10530
318	DEG10540
319	DEG10550
320	DEG10560
321	DEG10580
322	DEG10590

323	DEG10600
324	DEG10610

S2 Performance Evaluation

The formulae for the error metrics used to evaluate the different models include modified Kling–Gupta efficiency (Kling et al., 2012), and Nash-Sutcliff efficiency (NSE; McCuen et al., 2006):

$$\mathrm{KGE} = 1 - \sqrt{(\rho - 1)^2 + (\frac{\sigma_{\hat{Q}}}{\sigma_Q} - 1)^2 + (\frac{\mu_{\hat{Q}}}{\mu_Q} - 1)^2} \quad \text{Eq S 1}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}(\hat{Q}_i - Q_i)^2}{\sum_{i=1}^{N}(\mu_Q - Q_i)^2} \quad \text{Eq S 2}$$

where μ and σ are the mean and standard deviation operators, respectively. The parameter N indicates the total number of samples and ρ is the linear correlation (Pearson correlation) between predictions and observations.

S3 Results

<u>Scenario A: CAMELS-US</u>

The 50 randomly selected catchments index numbers are provided in Table below. It is noted that results provided in the main text are associated with pool 1.

Table S3- Selected catchment index numbers to train the base LSTM model.

Pool 1	Pool 2
7	2
8	5
16	12
22	14
33	15
34	16
67	33
78	49
82	53
99	61
119	69
120	108
122	111

133	114
138	118
155	119
190	120
199	131
200	137
203	163
204	177
219	195
241	196
243	200
244	217
268	222
277	225
279	229
282	231
283	242
298	250
303	254
304	271
305	278
310	284
311	292
321	303
326	312
328	333
344	334
366	352
368	357
369	370
373	372
380	390
389	391
398	392
400	393
404	404
411	411

The NSE and KGE values' distribution for 371 catchments that were excluded from training the base model (pool 2) are depicted in Fig S1. $\mathrm{LSTM}^{\mathrm{A}}$ indicates fine-tuning all the weights, while $\mathrm{LSTM}^{\mathrm{B}}$ is associated with fine-tuning only the linear head.

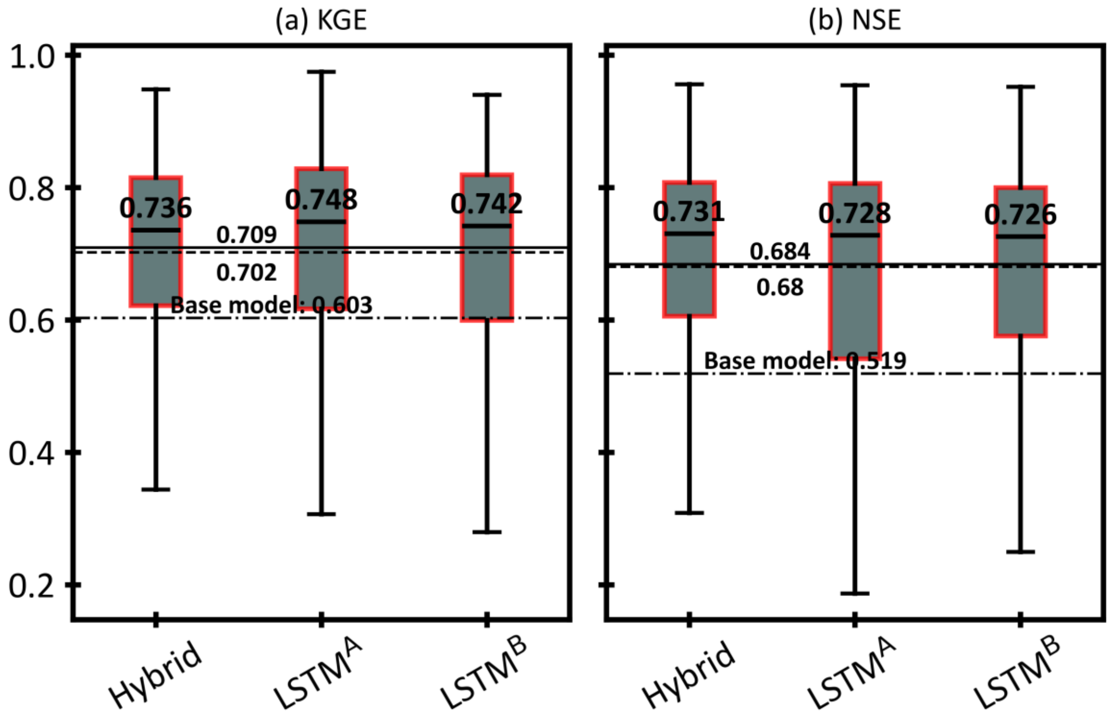

Fig S1- NSE and KGE distribution over the 371 catchments. The solid line indicates the regional model median trained on all 421 catchments, while the dashed line shows the median of the independently trained LSTM models.

<u>Scenario A: CAMELS-DE</u>

The 50 randomly selected catchments index numbers used to train the base model for CAMELS-DE are listed in Table S4.

Table S4- Selected catchment index numbers to train the base LSTM model.

Pool 1
2
5
12
14
15
16
33
49
52
53
61
62
69

96
108
111
113
114
118
119
120
131
137
149
152
163
171
177
195
196
200
214
217
222
225
229
231
236
242
250
254
257
271
278
284
285
292
303
312
323

The empirical cumulative distribution functions (CDFs) for percent bias of the top 2% peak flow range (FHV) and the percent bias of the bottom 30% (baseflow) range (FLV), corresponding to the regional and hybrid models, are presented in Figure S2.

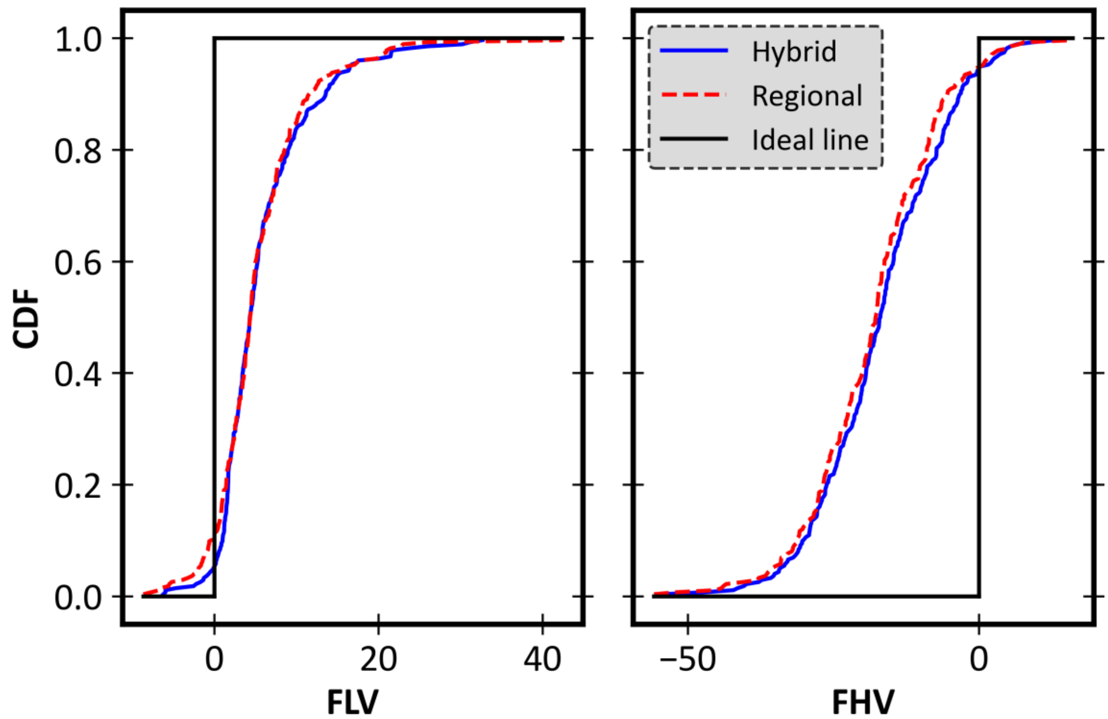

Fig S2- Empirical CDF for FHV and FLV associated with the regional and hybrid model.

<u>Scenario B: CAMELS-US</u>

The significance of NSE change in the catchments was tested using Wilcoxon test (Wilcoxon, 1945). Catchments with significant KGE improvements are marked by green positive markers, whereas those with significant decreases are indicated by red negative markers. Transparent markers represent catchments where the differences were not statistically significant.

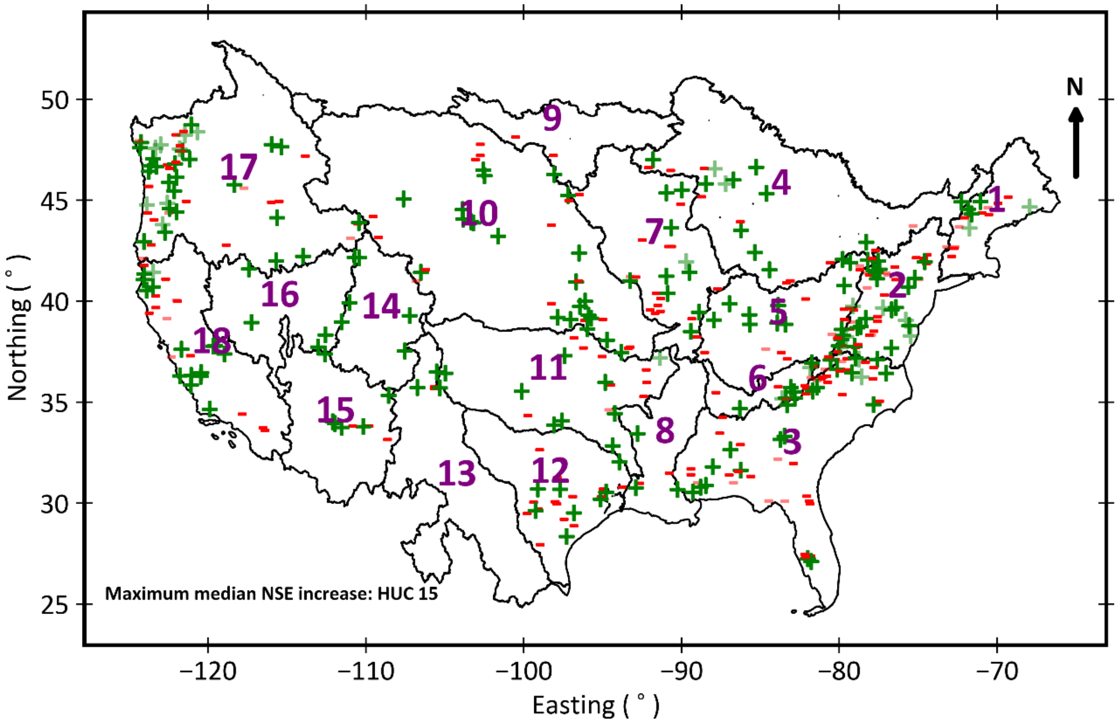

Fig S3- KGE increase (green +) or decrease (red -) across the selected 421 catchments. Only catchments with KGE>-1 were considered for comparison. The transparent signs indicate non-significant difference.

The NSE and KGE values for hybrid fine-tuned (epoch 1) and the regional model associated with a different random seed compared to results shown in the main body of the paper are provided in https://github.com/sinajahangir/Hybrid_LSTM_RF/tree/main/Results%20Samples. The boxplots of the metrics are shown in Figure S4.

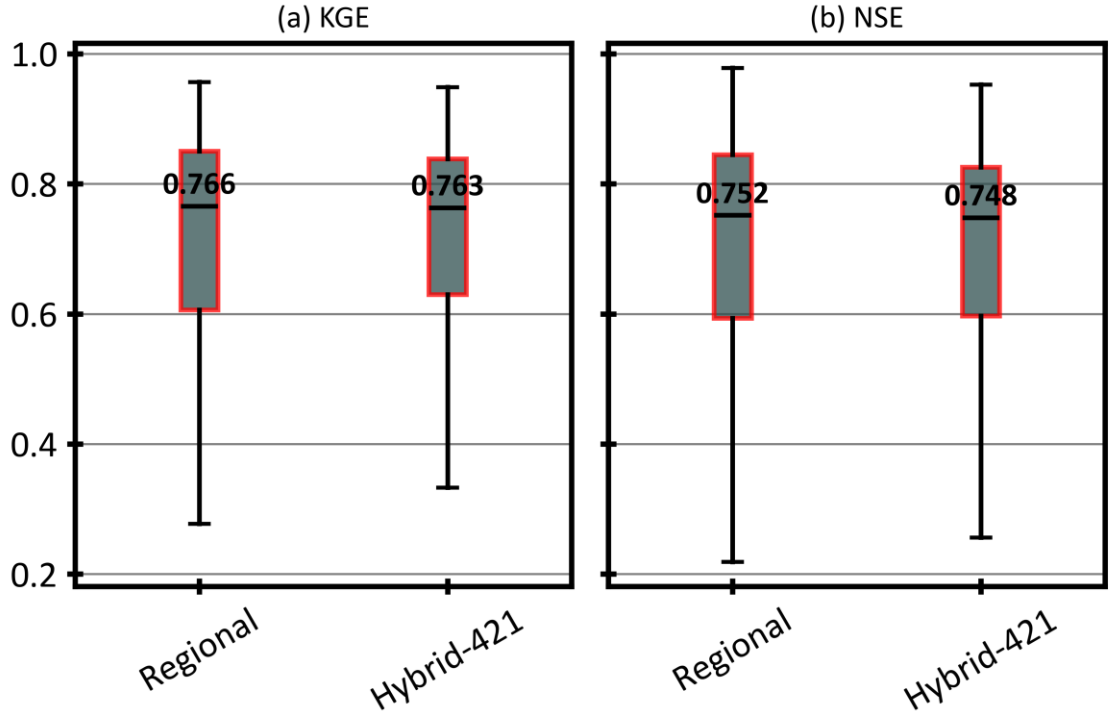

Fig S4- NSE and KGE distribution over the 421 catchments associated with the regional, and hybrid-421 models.

Table S5- NSE and KGE values of the hybrid model and the regional model.

S4 Catchment-Wise Training Strategy

For each catchment, an LSTM model was trained using a 365-day input lag, a cell size of 256, a dropout rate of 0.2, and a linear output layer. A dynamic learning rate schedule was employed, starting at 1e-3 and reduced by a factor of 10 after five epochs if no improvement in validation performance was observed. Prior to training, all input forcings were z-normalized.

References

Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. Journal of Hydrology 424–425, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011

McCuen, R.H., Knight, Z., Cutter, A.G., 2006. Evaluation of the Nash–Sutcliffe Efficiency Index. Journal of Hydrologic Engineering 11, 597–602. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:6(597)

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. Biometrics Bulletin 1, 80–83. https://doi.org/10.2307/3001968