Supplementary materials for the article 'Downscaling and projection of multi-CMIP5 precipitation using machine learning methods in the Upper Han river basin

' Ren Xu [1, 2], Nengcheng Chen [2, 3], Yumin Chen [1], Zeqiang Chen [2, 3, *]

[1] School of Resource and Environmental Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
[2] State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
[3] Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China
* Correspondence: ZeqiangChen@whu.edu.cn; Tel.: +86-13871025965

The supplementary materials mainly include the processes and details of Bayesian hyper-parameter optimization for multilayer percptron (MLP), support vector regression (SVR), and random forest (RF) for region mean. In addition, the supplementary materials include the evaluation of 21 NEX-GDDP models and MME model for specific stations in this region, and the changes of projected 21st century precipitation for each station. Each section has been refereed and cited in the manuscript.

The aims of the supplementary materials are to supply the processes and details of Bayesian hyper-parameter optimization, and the models evaluation and future changes of specific station.

The contents of the supplementsary materials include:

1. **Details of Bayesian hyper-parameter optimization for MLP**
2. **Details of Bayesian hyper-parameter optimization for RF**
3. **Details of Bayesian hyper-parameter optimization for SVR**
4. **The evaluation of ML downscaled models for specific stations in this region**
5. **The changes of projected 21st precipitation for each station**

1. Details of Bayesian hyper-parameter optimization for MLP

This section discussed the used software for hyper-parameter optimization, and presented the diagrams of the optimization process and optimization results of MLP methods for the region mean. This study used python's 'hyperopt'package to implement BHO for MLP. The mean squared error (MSE) was used as the optimization objective for MLP. Table S1 represents the optimal results of MLP for region mean. **The elements of the first row represent optimized hyper-parameters.**

Table S1: Optimal results of Bayesian hyper-parameter optimization for MLP for region mean

	activation	alpha	HLZ	LR	Max_iter	Solver	tol	MSE
AC10	tanh	1.024201	20	adaptive	1500	adam	0.00319	1444
AC13	logistic	9.831136	12	adaptive	125	lbfgs	0.008452	1605
GG	tanh	4.26283	16	adaptive	1152	adam	0.002589	1501
GH	tanh	8.535319	14	adaptive	1385	adam	0.000246	1511
GM	logistic	3.838534	14	adaptive	413	adam	0.006199	1528
GR	tanh	9.454777	12	adaptive	1521	adam	0.003523	1755
IN	logistic	7.885871	11	invscaling	145	lbfgs	0.005522	1520
Nor	logistic	9.299373	10	constant	1511	lbfgs	0.000955	1458
MME	tanh	3.534963	6	adaptive	834	adam	0.000142	1340
BMA	logistic	9.045651	9	constant	1360	lbfgs	0.007592	1278

* **HLZ denotes hidden_layer_sizes; LR denotes learning_rate.**

Figs. S1-S10 express the optimal process for region mean. Noting that the X axis of each subplot denote the corresponding hyper-parameters. **Y axis denote the MSE.** The hyper-parameter group which minimizes the MSE are regarded as the optimal parameters.

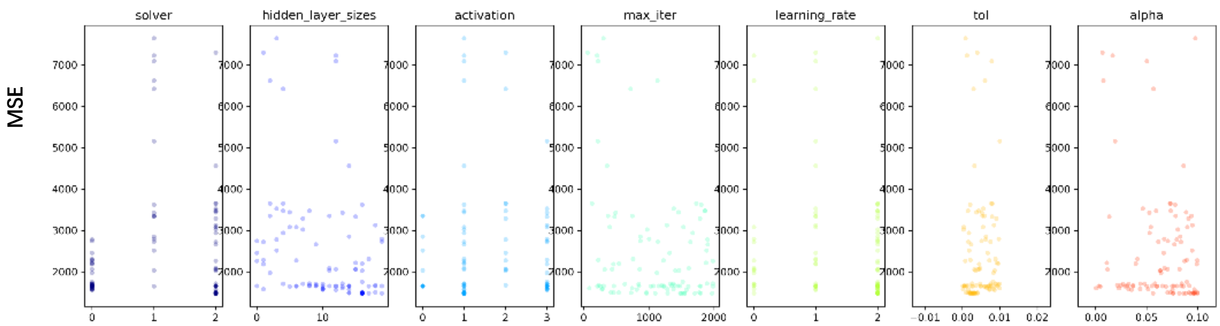


Figure S1: Bayesian hyper-parameter optimization process of MLP downscaling modelling of Ac10 for region mean

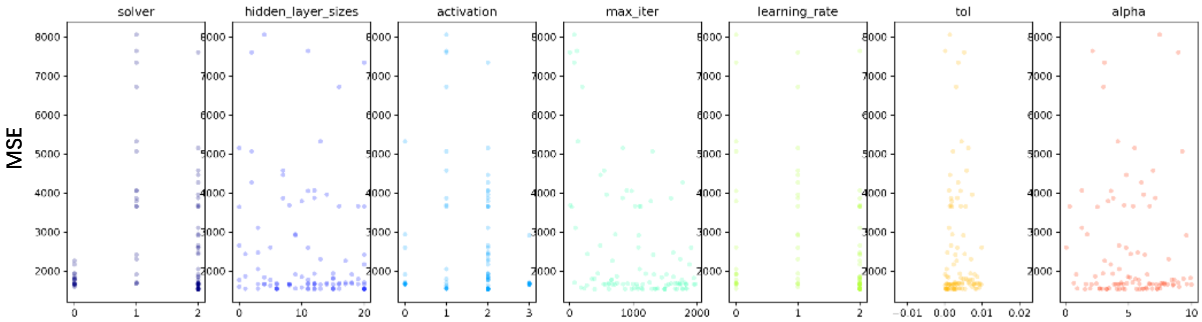


Figure S2: Bayesian hyper-parameter optimization process of MLP downscaling modelling of Ac13 for region mean

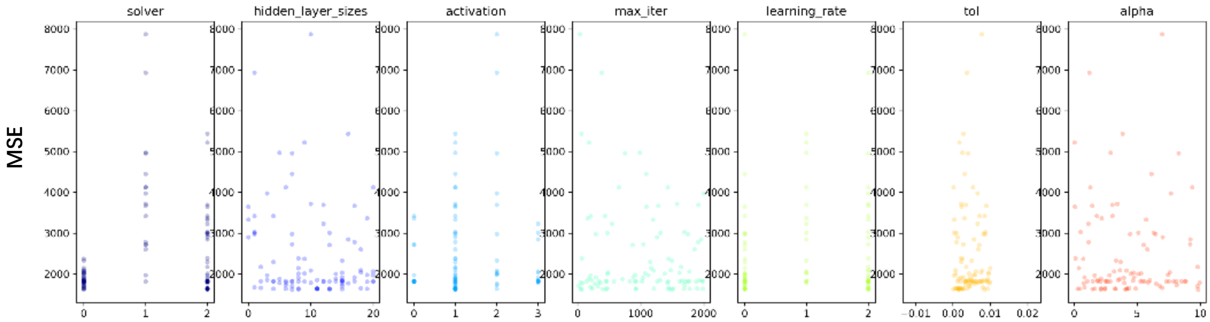


Figure S3: Bayesian hyper-parameter optimization process of MLP downscaling modelling of GG for region mean

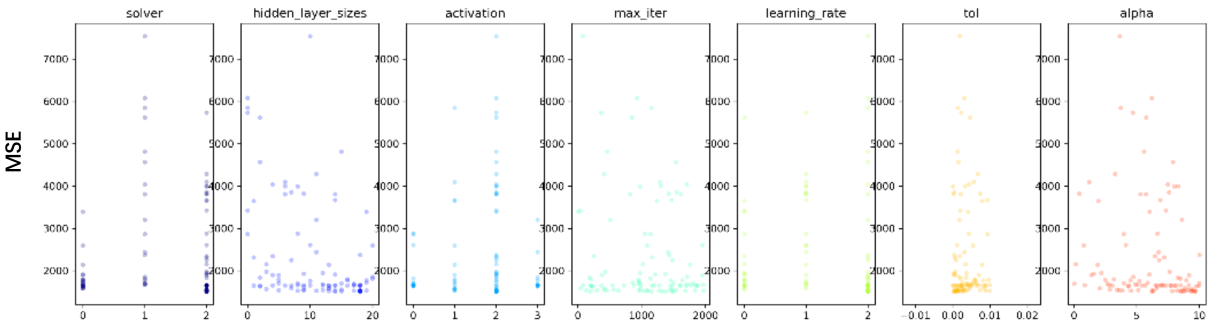


Figure S4: Bayesian hyper-parameter optimization process of MLP downscaling modelling of GH for region mean

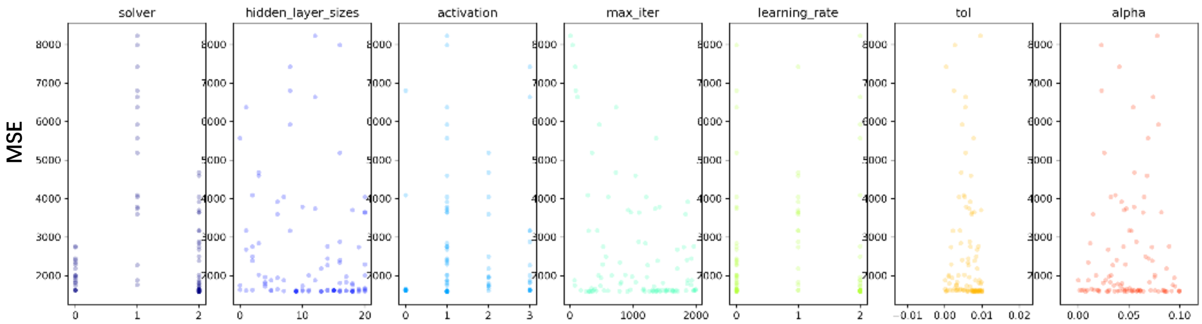


Figure S5: Bayesian hyper-parameter optimization process of MLP downscaling modelling of GM for region mean

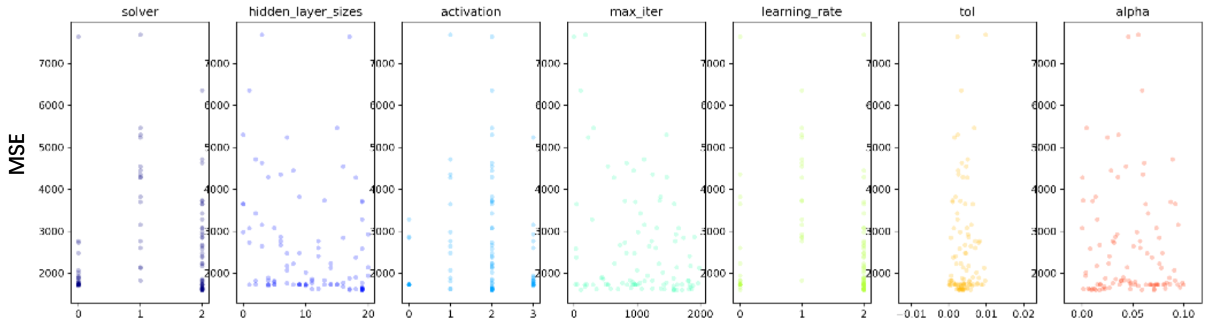


Figure S6: Bayesian hyper-parameter optimization process of MLP downscaling modelling of GR for region mean

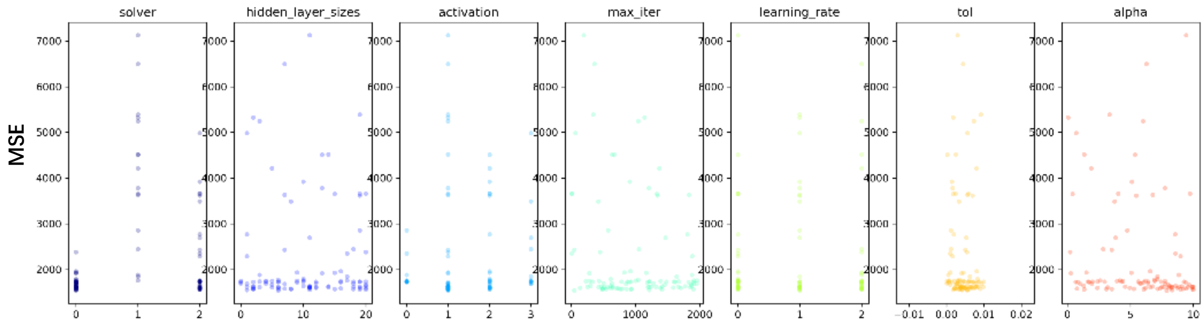


Figure S7: Bayesian hyper-parameter optimization process of MLP downscaling modelling of IN for region mean

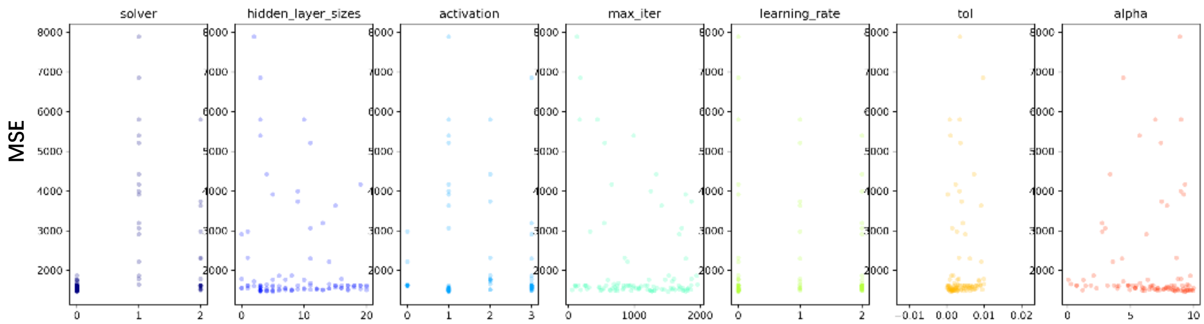


Figure S8: Bayesian hyper-parameter optimization process of MLP downscaling modelling of Nor for region mean

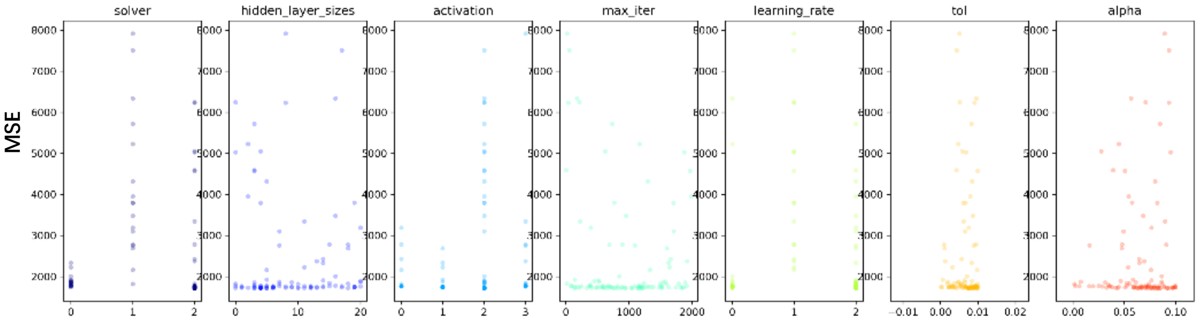


Figure S9: Bayesian hyper-parameter optimization process of MLP downscaling modelling of MME for region mean

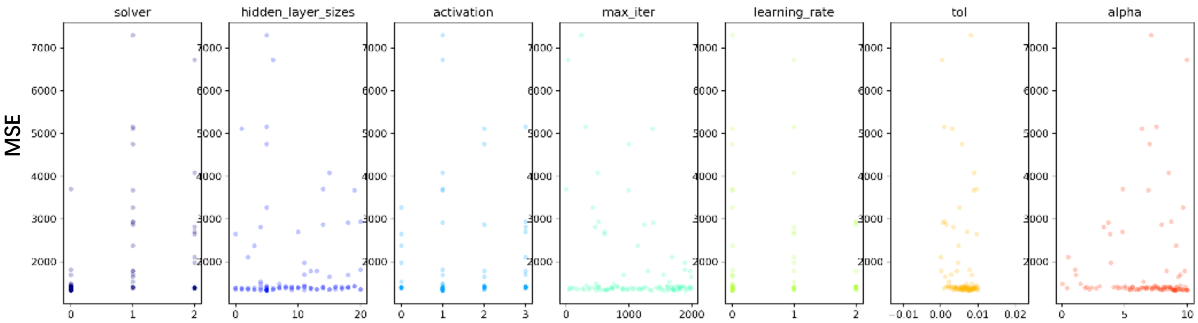


Figure S10: Bayesian hyper-parameter optimization process of MLP downscaling modelling of BMA for region mean

2. Details of Bayesian hyper-parameter optimization for RF

This section discussed the used software for hyper-parameter optimization, and presented the diagrams of the optimization process and optimization results of MLP methods for the region mean. This study used python's 'hyperopt'package to implement BHO for RF. The mean squared error (MSE) was used as the optimization objective for RF. Table S1 represents the optimal results of MLP for region mean. **The elements of the first row represent optimized hyper-parameters.**

Table S2: Optimal results of Bayesian hyper-parameter optimization for RF for region mean

	max_depth	**max_features**	**n_estimators**	**MSE**
AC10	10	5	298	1562.
AC13	18	3	512	1604
GG	5	3	498	1581
GH	4	6	985	1551
GM	6	4	127	1630
GR	12	3	967	1657
IN	6	3	934	1549
Nor	13	2	451	1530
MME	4	2	622	1440
BMA	4	2	190	1401

Figs. S11-S20 express the optimal process for region mean. Noting that the X axis of each subplot denotes the values of corresponding hyper-parameters. **Y axis denotes the MSE.** The hyper-parameter group which minimizes the MSE are the optimal parameters.

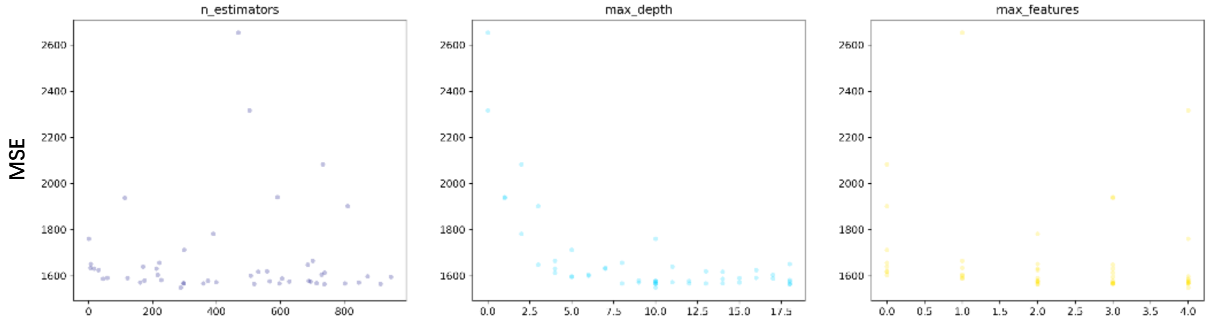


Figure S11: Bayesian hyper-parameter optimization process of RF downscaling modelling of AC10 for region mean

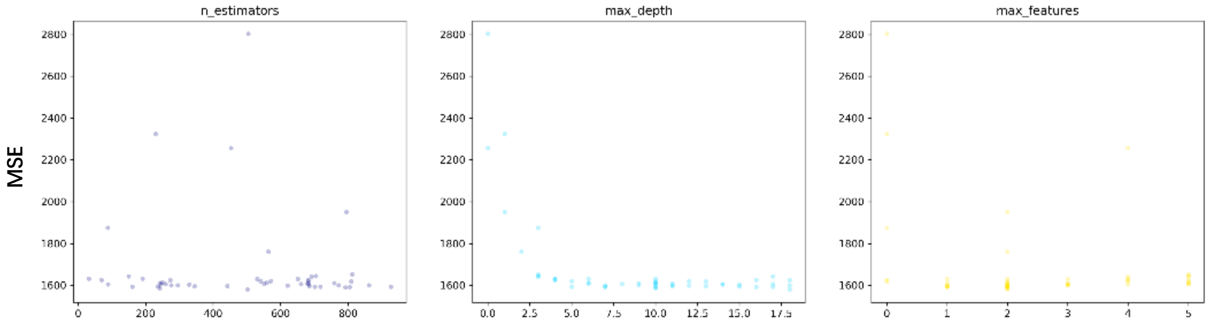


Figure S12: Bayesian hyper-parameter optimization process of RF downscaling modelling of AC13 for region mean

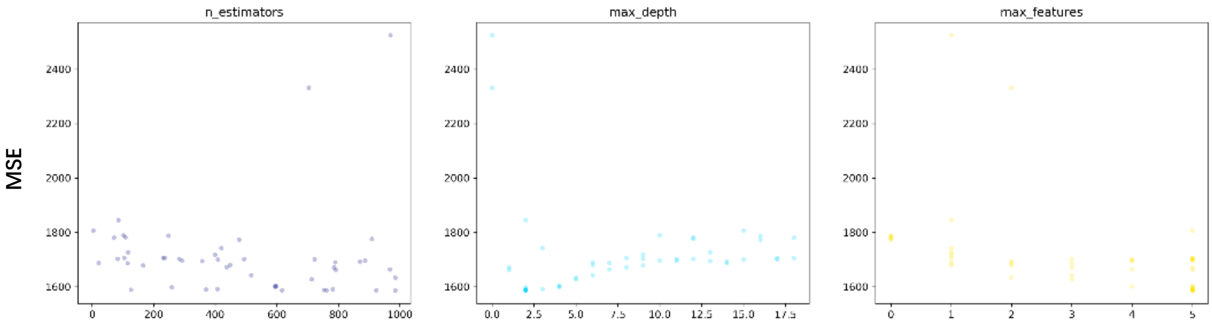


Figure S13: Bayesian hyper-parameter optimization process of RF downscaling modelling of GG for region mean

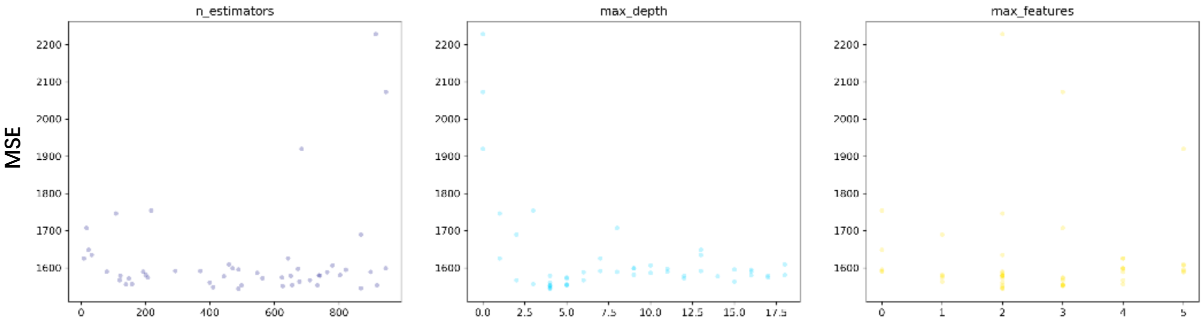


Figure S14: Bayesian hyper-parameter optimization process of RF downscaling modelling of GH for region mean

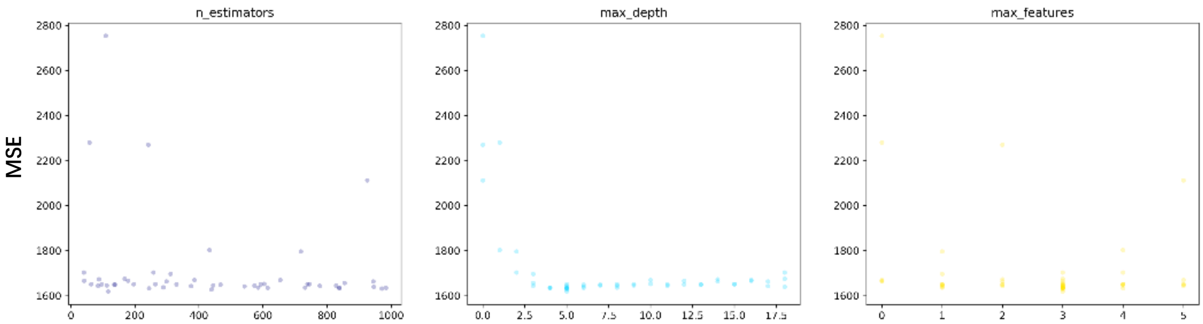


Figure S15: Bayesian hyper-parameter optimization process of RF downscaling modelling of GM for region mean

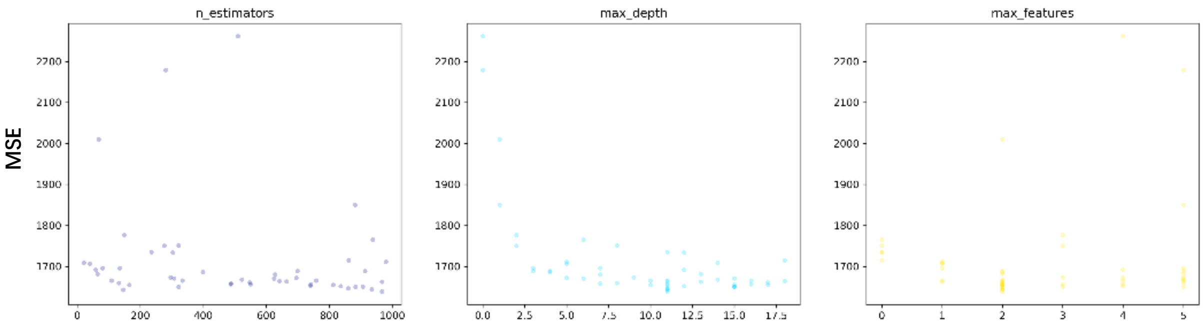


Figure S16: Bayesian hyper-parameter optimization process of RF downscaling modelling of GR for region mean

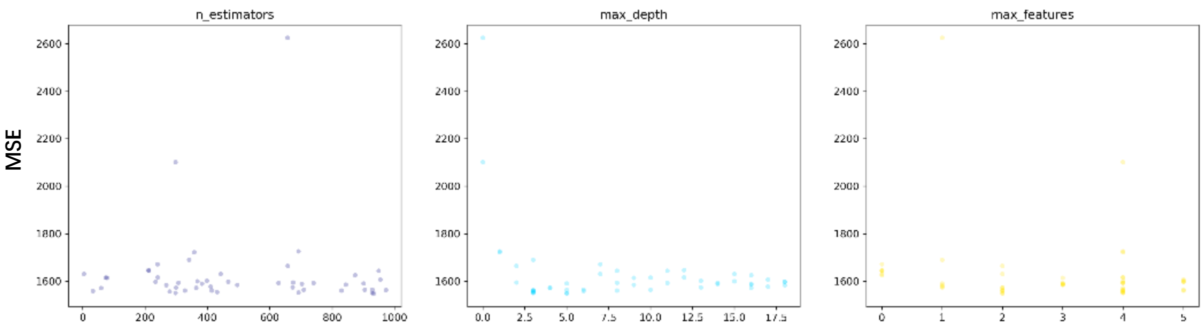


Figure S17: Bayesian hyper-parameter optimization process of RF downscaling modelling of IN for region mean

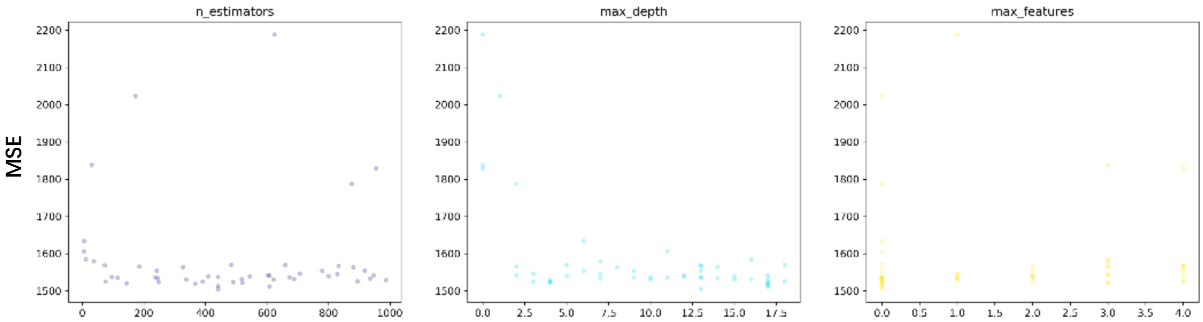


Figure S18: Bayesian hyper-parameter optimization process of RF downscaling modelling of Nor for region mean

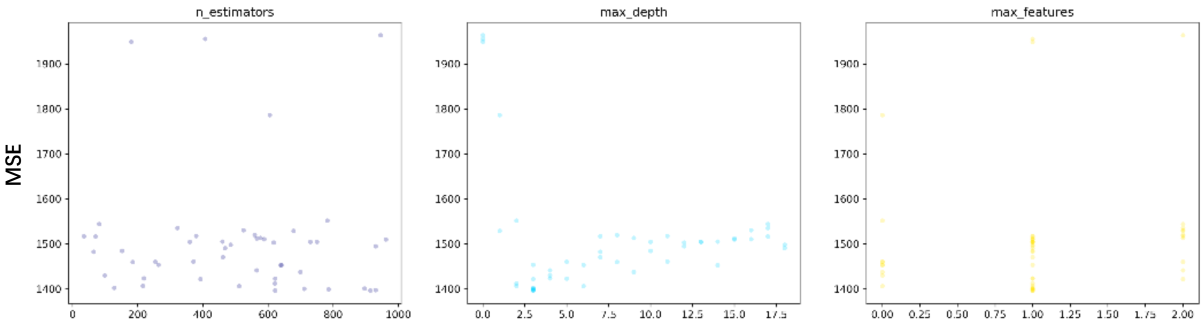


Figure S19: Bayesian hyper-parameter optimization process of RF downscaling modelling of MME for region mean

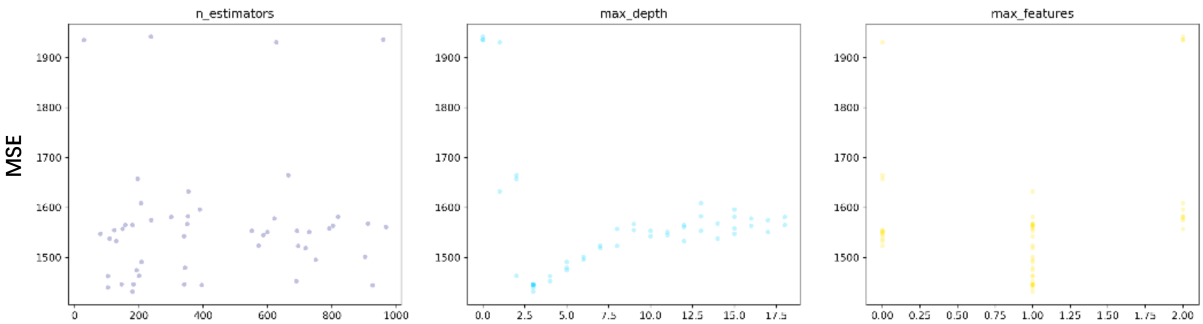


Figure S20: Bayesian hyper-parameter optimization process of RF downscaling modelling of BMA for region mean

3. Details of Bayesian hyper-parameter optimization for SVR

This section discussed the used software for hyper-parameter optimization, and presented the diagrams of the optimization process and optimization results of SVR methods for the region mean. This study used MATLAB's 'fitrsvm' function for SVM because it is a self-contained function in MATLAB 2016b. **The loss function was deemed** as the optimization objective for SVR. Table S3 represents the optimal results of SVR for region mean. **The elements of the row line represent optimized hyper-parameters.**

Table S3: Optimal results of Bayesian hyper-parameter optimization of SVR for region mean

	BoxConstraint	KernelScale	Epsilon	KernelFunction	PolynomialOrder	Standardize	Objective
AC10	976.76	856.48	0.47992	Gaussian	NaN	false	7.3573
AC13	984.83	80.665	1.4186	Gaussian	NaN	false	7.3572
GG	43.612	89.569	0.5045	Gaussian	NaN	false	7.4439
GH	981.77	41.537	15.467	Gaussian	NaN	false	7.4516
GM	949.64	NaN	14.106	Polynomial(rbf)	1	true	7.3604
GR	637.68	61.876	0.071061	gaussian	NaN	false	7.4265
IN	203.68	NaN	0.30124	Polynomial(rbf)	2	true	7.406
Nor	753.41	144.223	0.12571	gaussian	NaN	true	7.4268
MME	12.972	NaN	8.73	Polynomial	2	true	7.2276
BMA	881.15	NaN	0.37308	Polynomial	2	True	7.2081

The following Figs S21-S30 express the optimal process. Noting that the X axis denote numbers of iteration; **Y axis denote the value of loss function.** The hyper-parameter group which minimizes the MSE are the optimal parameters.

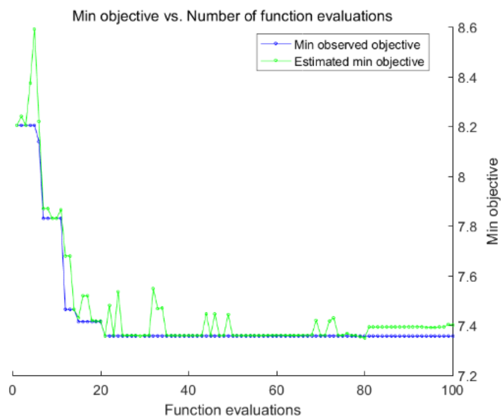
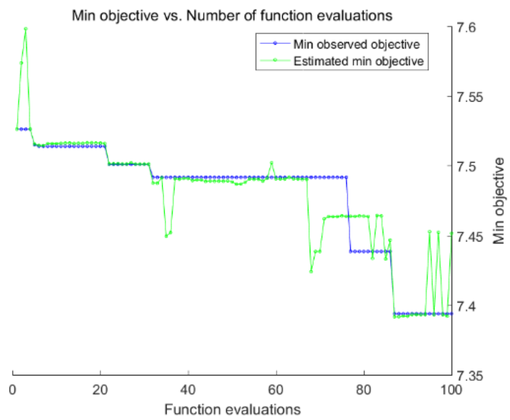


Figure S21: Bayesian hyper-parameter optimization process of SVR downscaling modelling of AC10 region mean



(m-n)

Figure S22: Bayesian hyper-parameter optimization process of SVR downscaling modelling of AC13 for region mean

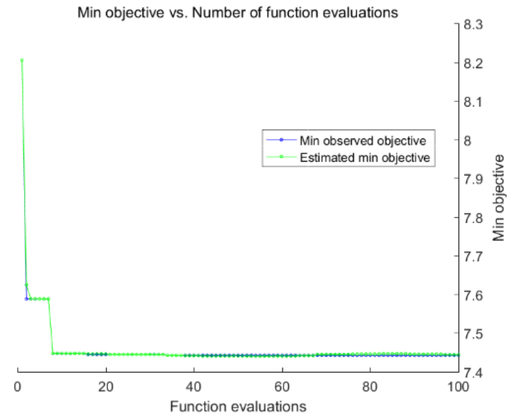


Figure S23: Bayesian hyper-parameter optimization process of SVR downscaling modelling of GG for region mean

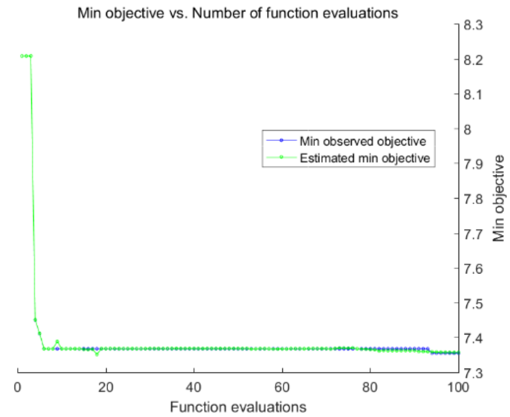


Figure S24: Bayesian hyper-parameter optimization process of SVR downscaling modelling of GH for region mean

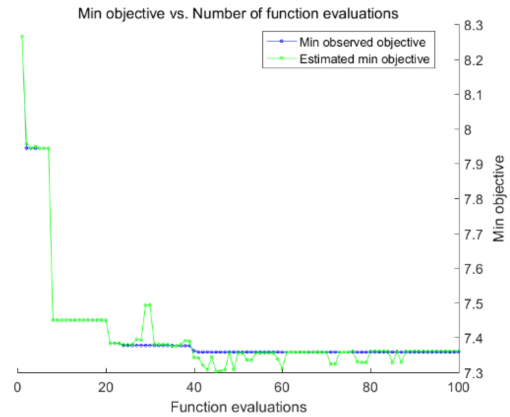


Figure S25: Bayesian hyper-parameter optimization process of SVR downscaling modelling of GM for region mean

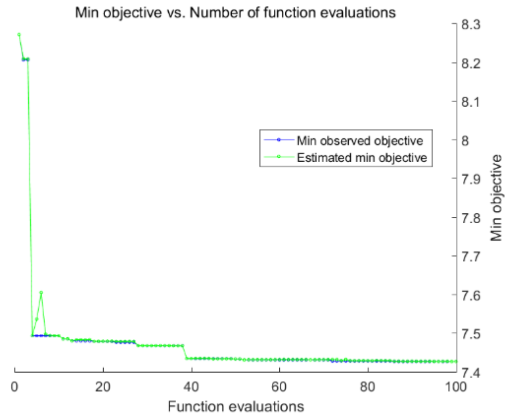


Figure S26: Bayesian hyper-parameter optimization process of SVR downscaling modelling of GR for region mean

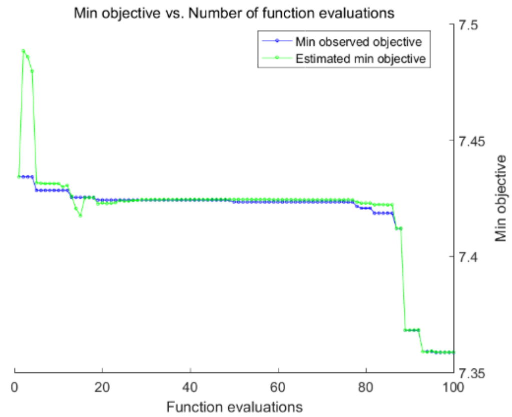


Figure S27: Bayesian hyper-parameter optimization process of SVR downscaling modelling of IN for region mean

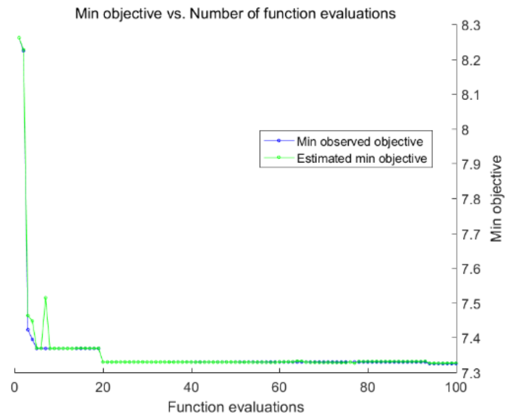


Figure S28: Bayesian hyper-parameter optimization process of SVR downscaling modelling of Nor for region mean

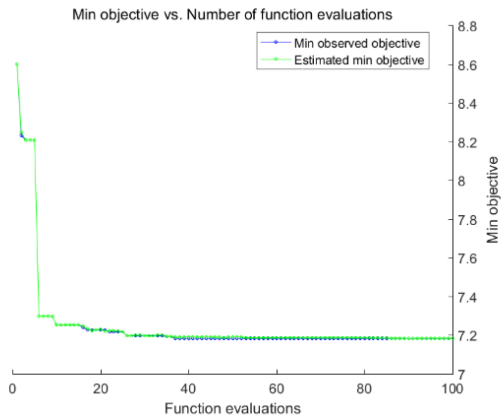


Figure S29: Bayesian hyper-parameter optimization process of SVR downscaling modelling of MME for region mean

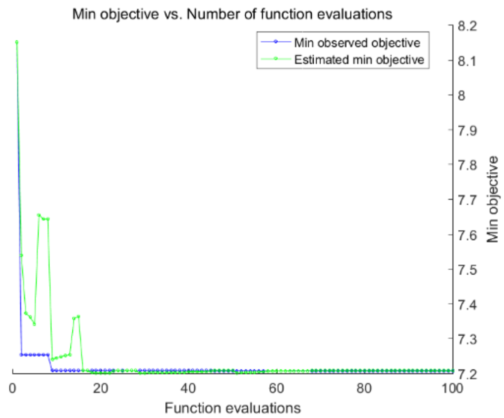


Figure S30: Bayesian hyper-parameter optimization process of SVR downscaling modelling of BMA for region mean

4. The evaluation of ML downscaled models for specific stations in this region

Table content:

Table S4: the validation results of MLP downscaled models for specific stations in this region

Mode		1	2	3	4	5	6	7	8	9	10	11	12	13
	CC	0.74	0.71	0.65	0.72	0.69	0.71	0.66	0.69	0.67	0.68	0.65	0.70	0.54
AC1.0	RMS	43.46	53.61	55.97	54.58	40.17	45.42	53.71	53.66	57.52	47.96	48.10	43.94	53.06
	Pbias	-15.56	-9.78	13.07	-7.56	-13.32	6.41	11.18	-2.88	1.68	-6.01	-0.71	-1.98	5.64
	CC	0.72	0.70	0.62	0.70	0.64	0.67	0.63	0.69	0.66	0.66	0.61	0.71	0.53
AC1.3	RMS	45.08	54.35	57.76	56.10	42.45	48.24	55.48	53.85	58.80	49.26	50.47	43.56	53.52
	Pbias	4.94	4.24	1.99	5.55	5.76	-6.07	-4.92	-5.70	-10.78	-7.11	5.89	8.50	1.83
	CC	0.73	0.71	0.66	0.72	0.67	0.70	0.62	0.64	0.68	0.69	0.63	0.72	0.56
GG	RMS	44.41	53.22	55.28	54.45	40.95	46.50	55.87	57.02	56.76	47.46	49.51	42.64	52.38
	Pbias	-11.87	-4.05	10.45	-8.78	-7.51	3.79	1.10	1.32	-7.38	1.87	-1.10	1.31	-4.11
	CC	0.75	0.72	0.65	0.73	0.65	0.70	0.63	0.65	0.67	0.67	0.64	0.69	0.54
GM	RMS	43.08	52.37	56.16	53.61	41.80	46.21	55.48	56.27	57.71	48.79	49.11	44.51	53.68
	Pbias	3.85	11.77	11.49	-5.88	-8.44	-6.95	-7.77	6.71	18.76	12.47	5.93	1.39	5.98
	CC	0.73	0.72	0.66	0.71	0.66	0.69	0.62	0.64	0.65	0.68	0.63	0.73	0.52
GH	RMS	44.50	52.30	55.36	55.29	41.39	46.58	56.07	57.35	58.71	47.97	49.30	42.03	54.14
	Pbias	-0.89	-3.33	-3.97	12.97	9.64	14.38	-0.54	-2.86	14.37	-1.11	9.23	2.51	5.15
	CC	0.71	0.70	0.64	0.70	0.64	0.67	0.61	0.69	0.64	0.65	0.61	0.69	0.54
GR	RMS	45.94	53.94	56.96	56.48	42.42	48.23	56.64	53.55	59.46	49.67	50.64	44.64	53.16
	Pbias	-4.33	-0.65	1.45	5.82	4.35	1.05	1.05	4.07	13.71	5.41	-8.52	-2.75	6.93
	CC	0.74	0.71	0.64	0.72	0.66	0.69	0.63	0.64	0.66	0.68	0.62	0.72	0.51
IN	RMS	43.77	53.68	56.62	54.84	41.54	46.93	55.28	57.26	58.19	48.04	49.76	42.99	54.35
	Pbias	2.08	8.45	5.89	-13.28	9.09	-6.29	3.34	-3.90	3.80	-2.45	8.70	11.41	-9.40
	CC	0.73	0.71	0.65	0.73	0.66	0.70	0.65	0.68	0.67	0.68	0.64	0.74	0.58
Nor	RMS	44.55	53.12	55.98	53.72	41.29	46.26	54.07	54.62	57.61	47.91	48.98	41.60	51.62
	Pbias	-9.71	11.13	-4.00	4.92	-12.28	-5.19	4.66	-12.82	-14.56	-9.05	3.66	-1.48	-9.55
	CC	0.75	0.74	0.68	0.74	0.68	0.72	0.66	0.68	0.70	0.69	0.65	0.73	0.57
MME	RMS	41.80	50.13	53.48	51.77	39.95	44.26	53.27	53.69	54.96	47.00	48.33	41.80	52.02
	Pbias	-4.22	-3.15	2.99	-2.13	-6.70	9.55	-8.10	-18.18	-2.49	-3.69	-14.61	-5.77	-2.45
	CC	0.76	0.75	0.68	0.77	0.70	0.73	0.68	0.68	0.71	0.70	0.67	0.74	0.56
BMA	RMS	41.21	48.87	52.88	49.64	38.50	43.41	51.50	53.12	53.57	45.68	46.60	40.43	51.14
	Pbias	-4.11	-3.56	8.69	-6.41	-3.92	12.00	-11.89	9.93	-11.90	-6.15	8.99	-7.18	2.60

Table S5: the validation results of SVR downscaled models for specific stations in this region

Mode		1	2	3	4	5	6	7	8	9	10	11	12	13
	CC	0.75	0.72	0.67	0.73	0.70	0.78	0.71	0.68	0.69	0.74	0.66	0.73	0.56
AC1.0	RMS	43.76	53.03	55.12	54.38	39.81	41.45	51.75	55.36	56.78	45.09	48.49	42.42	53.61
	Pbias	-3.03	-3.50	-4.03	-0.96	5.06	-2.16	-7.17	-5.32	-1.73	-4.27	-4.65	4.30	-8.29
	CC	0.71	0.74	0.62	0.71	0.63	0.64	0.63	0.67	0.64	0.69	0.65	0.72	0.57
AC1.3	RMS	46.12	52.54	58.61	56.18	43.52	50.38	56.28	55.60	60.19	48.40	49.75	43.25	52.92
	Pbias	-2.26	-5.09	-5.08	-1.73	-4.94	-6.63	-7.44	-4.93	-2.55	-5.85	-7.82	2.94	-4.36
	CC	0.72	0.71	0.64	0.72	0.66	0.68	0.63	0.65	0.66	0.70	0.62	0.72	0.56
GG	RMS	45.58	54.33	57.52	54.87	41.64	47.78	56.23	58.04	59.41	47.12	50.68	43.37	54.10
	Pbias	-4.71	-6.34	-4.88	0.80	3.14	1.63	3.11	-7.45	-6.83	-1.10	-6.97	1.36	-10.01
	CC	0.83	0.72	0.67	0.73	0.66	0.72	0.66	0.67	0.71	0.71	0.67	0.72	0.57
GM	RMS	36.44	53.25	55.77	54.46	41.96	45.79	54.62	55.60	55.06	46.87	47.17	43.16	52.60
	Pbias	-1.92	-2.87	-2.44	-3.72	-5.70	-4.40	-7.31	0.77	-3.89	-3.83	7.15	4.55	0.41
	CC	0.73	0.73	0.68	0.73	0.69	0.73	0.69	0.68	0.67	0.71	0.71	0.73	0.55
GH	RMS	44.90	52.54	54.37	54.43	40.82	44.69	52.61	56.01	58.61	46.78	45.77	42.49	53.07
	Pbias	-4.27	-3.72	-1.49	2.53	-3.18	-1.92	-5.78	8.15	-3.68	-2.61	1.20	3.43	-0.13
	CC	0.74	0.72	0.66	0.71	0.65	0.66	0.64	0.68	0.66	0.63	0.94	0.69	0.54
GR	RMS	44.47	52.99	56.09	56.73	42.27	49.12	56.25	55.14	59.77	51.71	23.65	44.85	53.62
	Pbias	-3.76	-3.53	-5.55	-5.63	-3.17	-4.54	-8.28	-4.23	-7.56	-5.05	-0.07	0.89	5.47
	CC	0.75	0.84	0.80	0.76	0.67	0.70	0.63	0.65	0.70	0.68	0.69	0.74	0.58
IN	RMS	43.43	42.26	44.79	52.58	41.75	46.89	56.21	57.54	56.37	48.74	46.97	42.12	52.56
	Pbias	-2.55	-1.51	6.43	-6.30	-4.64	-4.65	-1.27	-4.94	-5.05	-1.95	0.22	0.18	-5.75
	CC	0.78	0.75	0.72	0.76	0.72	0.75	0.69	0.68	0.72	0.72	0.64	0.72	0.58
Nor	RMS	41.24	50.84	51.61	51.60	39.02	43.70	52.84	55.13	54.69	46.12	49.33	43.03	51.56
	Pbias	-1.10	-3.42	-1.42	-3.35	2.10	-2.25	-2.30	0.96	-4.29	-2.18	0.97	1.81	2.42
	CC	0.78	0.75	0.71	0.78	0.74	0.75	0.71	0.70	0.75	0.72	0.68	0.74	0.58
MME	RMS	41.31	50.13	52.70	50.20	37.76	43.32	51.05	54.05	51.89	46.92	47.41	41.65	52.03
	Pbias	0.26	0.63	-1.11	-2.01	-1.37	4.51	-2.29	5.25	-1.21	-4.53	-1.98	6.38	-3.16
	CC	0.77	0.76	0.69	0.76	0.72	0.76	0.69	0.68	0.75	0.74	0.71	0.75	0.63
BMA	RMS	42.14	49.72	54.07	52.40	39.00	42.45	52.29	54.55	52.02	44.57	45.07	40.97	50.03
	Pbias	-1.62	1.23	-2.53	-4.03	-4.62	0.52	-1.43	4.89	-3.18	-0.21	-2.67	3.92	7.72

Table S6: the validation results of RF downscaled models for specific stations in this region

Mode		1	2	3	4	5	6	7	8	9	10	11	12	13
	CC	0.72	0.70	0.64	0.71	0.65	0.69	0.63	0.65	0.66	0.66	0.62	0.68	0.53
AC1.0	RMS	44.79	54.31	56.52	55.21	42.02	46.77	55.30	56.87	58.64	49.66	49.90	45.35	53.81
	Pbias	3.62	11.34	14.24	9.62	2.11	12.49	5.16	-8.97	5.13	-4.65	-5.03	6.71	8.31
	CC	0.71	0.69	0.64	0.71	0.63	0.68	0.64	0.65	0.65	0.65	0.61	0.70	0.51
AC.3	RMS	45.97	54.93	56.66	55.72	42.67	47.70	55.12	56.74	58.99	49.89	50.30	44.00	54.53
	Pbias	7.38	6.58	5.42	7.32	13.48	10.63	8.49	9.97	-5.48	7.50	-5.44	9.58	-2.20
	CC	0.72	0.69	0.65	0.71	0.66	0.67	0.61	0.62	0.67	0.67	0.62	0.70	0.55
GG	RMS	45.25	54.64	56.19	55.50	41.43	47.93	56.59	58.54	57.91	48.71	49.90	43.95	52.88
	Pbias	-1.98	3.41	-4.07	-1.92	-6.01	-1.87	1.76	-4.84	-2.72	-3.98	1.95	8.78	2.81
	CC	0.73	0.70	0.65	0.72	0.65	0.69	0.62	0.67	0.67	0.66	0.63	0.70	0.57
GM	RMS	44.21	53.99	56.11	54.61	42.07	46.74	56.19	55.41	57.40	49.36	49.37	43.91	52.07
	Pbias	12.00	4.64	-4.32	-6.69	-4.94	-4.01	-3.60	-6.96	-9.77	-7.39	-8.42	-9.69	-9.93
	CC	0.70	0.71	0.65	0.70	0.64	0.66	0.62	0.59	0.63	0.65	0.62	0.70	0.52
GH	RMS	46.79	53.75	56.28	56.50	42.50	48.74	55.91	59.82	60.24	49.71	49.86	43.88	54.05
	Pbias	122.76	6.35	4.75	2.02	-2.66	16.08	12.51	15.26	-1.82	117.78	7.08	11.19	12.33
	CC	0.72	0.73	0.65	0.70	0.63	0.66	0.60	0.66	0.65	0.64	0.60	0.68	0.52
GR	RMS	44.76	51.70	55.99	56.35	42.60	48.77	57.08	55.83	59.03	50.51	51.10	45.29	54.15
	Pbias	8.46	1.45	13.57	4.04	12.51	15.86	5.72	-3.38	111.63	16.61	101.08	6.68	2.97
	CC	0.74	0.72	0.67	0.74	0.66	0.68	0.60	0.62	0.66	0.68	0.61	0.71	0.52
IN	RMS	43.46	52.85	54.88	53.39	41.63	47.18	57.31	58.60	58.28	48.05	50.17	43.41	53.93
	Pbias	-0.71	9.78	7.48	17.83	4.54	1.87	4.96	10.96	2.67	5.47	3.31	5.13	3.28
	CC	0.72	0.70	0.64	0.72	0.67	0.70	0.63	0.66	0.66	0.67	0.64	0.71	0.57
Nor	RMS	45.23	54.12	56.46	54.57	41.08	46.55	55.41	55.64	58.10	48.88	48.85	43.12	52.33
	Pbias	15.20	-1.27	4.18	11.95	-2.42	-8.80	-3.00	-0.34	-6.68	-4.85	-3.62	-1.09	-3.56
	CC	0.74	0.74	0.66	0.75	0.67	0.71	0.65	0.67	0.68	0.69	0.65	0.73	0.55
MME	RMS	43.54	51.36	55.44	52.51	40.94	45.53	54.03	55.02	56.96	47.32	48.19	41.98	52.86
	Pbias	4.73	-3.96	-1.07	-1.13	-1.78	2.41	1.84	1.46	3.20	2.12	4.75	1.98	-1.37
	CC	0.77	0.74	0.68	0.75	0.69	0.74	0.66	0.71	0.69	0.69	0.67	0.74	0.58
BMA	RMS	42.38	51.90	54.90	53.12	40.54	44.80	54.45	53.64	57.49	48.31	47.96	42.42	52.14
	Pbias	7.46	2.16	-3.02	-1.43	7.62	2.00	-0.61	2.65	8.91	1.51	2.04	2.54	-0.75

5. The changes of projected 21st century precipitation for each station

Table S7: Projected changes of each station under RCP4.5 and RCP8.5 (%)

station	RCP4.5		RCP8.5	
	2040-2069	2070-2099	2040-2069	2070-2099
1	-6.28	-7.20	-6.30	-4.87
2	4.97	5.13	6.08	5.77
3	-0.61	-2.30	0.55	-0.05
4	9.76	10.53	11.99	9.51
5	-8.85	-9.66	-8.02	-7.10
6	-0.37	0.30	3.11	4.47
7	0.80	0.33	-0.56	1.78
8	4.19	6.20	3.90	6.53
9	7.91	6.02	8.96	9.97
10	-3.60	-3.28	-2.04	-0.98
11	-0.87	1.34	-0.78	1.33
12	5.27	8.21	5.12	8.37
13	3.31	4.78	2.96	5.13