Supplementary materials

Table S1 Adjusted R^2 of XGBoost derived from leave-one-out cross validation with samples split by climate regions under rainfed condition in A0 scenario.

GGCMs	Winter wheat	Spring wheat	Maize	Rice
APSIM-UGOE	0.91	0.72	0.66	0.64
CARAIB	0.74	0.82	0.68	0.7
EPIC-IIASA	0.75	0.77	0.81	0.74
EPIC-TAMU	0.81	0.71	0.78	0.68
GEPIC	0.71	0.8	0.8	0.77
LPJ-GUESS	0.93	0.79	-	-
LPJmL	0.8	0.79	0.83	0.74
ORCHIDEE-crop	0.66	-	0.88	0.81
pDSSAT	0.71	0.73	0.71	0.68
PEPIC	0.89	0.85	0.85	0.8

"-": No GGCM simulation.

Table S2 The global median R across all grids and perturbations over current croplands.

GGCM	Winter wheat				Spring wheat				Maize				Rice			
	C	T	W	N	C	T	W	N	C	T	W	N	C	T	W	N
APSIM-UGOE	0.94	0.93	0.94	0.90	0.69	0.79	0.81	0.65	0.92	0.91	0.92	0.88	0.94	0.93	0.94	0.92
CARAIB	0.93	0.93	0.93	-	0.89	0.96	0.95	-	0.84	0.84	0.83	-	0.85	0.85	0.80	-
EPIC-IIASA	0.97	0.97	0.97	0.98	0.72	0.82	0.82	0.98	0.95	0.95	0.95	0.96	0.94	0.94	0.94	0.95
EPIC-TAMU	0.94	0.93	0.93	0.92	0.86	0.92	0.92	0.87	0.94	0.94	0.94	0.89	0.93	0.93	0.93	0.85
GEPIC	0.90	0.91	0.88	0.89	0.88	0.93	0.93	0.87	0.95	0.94	0.94	0.90	0.95	0.94	0.94	0.87
LPJ-GUESS	0.94	0.93	0.94	0.92	0.85	0.91	0.92	0.83	-	-	-	-	-	-	-	-
LPJmL	0.96	0.95	0.95	0.95	0.86	0.95	0.95	0.85	0.93	0.88	0.88	0.90	0.93	0.89	0.89	0.89
ORCHIDEE-crop	0.97	0.97	0.97	0.97	-	-	-	-	0.89	0.89	0.89	0.89	0.92	0.92	0.92	0.92
pDSSAT	0.95	0.94	0.94	0.94	0.81	0.91	0.91	0.83	0.90	0.89	0.88	0.87	0.91	0.88	0.90	0.83
PEPIC	0.94	0.94	0.94	0.93	0.88	0.93	0.93	0.92	0.94	0.94	0.94	0.92	0.93	0.93	0.96	0.91

* "-" denotes the lack of raw GGCM simulation

Table S3 The global median MAE (t/ha) across all grids and all perturbations over current croplands.

GGCM	Winter wheat				Spring wheat				Maize				Rice			
	C	T	W	N	C	T	W	N	C	T	W	N	C	T	W	N
APSIM-UGOE	0.21	0.23	0.21	0.18	0.30	0.28	0.29	0.33	0.36	0.38	0.37	0.37	0.24	0.18	0.19	0.21
CARAIB	1.09	0.25	0.28	-	0.35	0.12	0.19	-	1.52	0.22	0.27	-	0.37	0.29	0.39	-
EPIC-IIASA	0.15	0.14	0.14	0.16	0.3	0.19	0.20	0.10	0.21	0.20	0.20	0.20	0.18	0.15	0.16	0.18
EPIC-TAMU	0.25	0.23	0.24	0.14	0.31	0.15	0.16	0.22	0.35	0.29	0.32	0.26	0.11	0.08	0.10	0.11
GEPIC	0.27	0.19	0.23	0.16	0.24	0.11	0.12	0.17	0.22	0.20	0.23	0.21	0.18	0.13	0.16	0.17
LPJ-GUESS	0.17	0.14	0.14	0.06	0.33	0.10	0.11	0.12	-	-	-	-	-	-	-	-
LPJmL	0.22	0.19	0.20	0.17	0.25	0.20	0.22	0.20	0.23	0.31	0.32	0.26	0.13	0.11	0.14	0.13
ORCHIDEE-crop	0.05	0.05	0.05	0.05	-	-	-	-	0.20	0.20	0.20	0.20	0.10	0.10	0.10	0.10
pDSSAT	0.42	0.40	0.40	0.27	0.40	0.12	0.12	0.22	0.47	0.42	0.50	0.45	0.41	0.38	0.42	0.42
PEPIC	0.23	0.20	0.22	0.15	0.25	0.12	0.13	0.14	0.25	0.23	0.26	0.20	0.23	0.17	0.20	0.18

* "-" denotes the lack of raw GGCM simulation

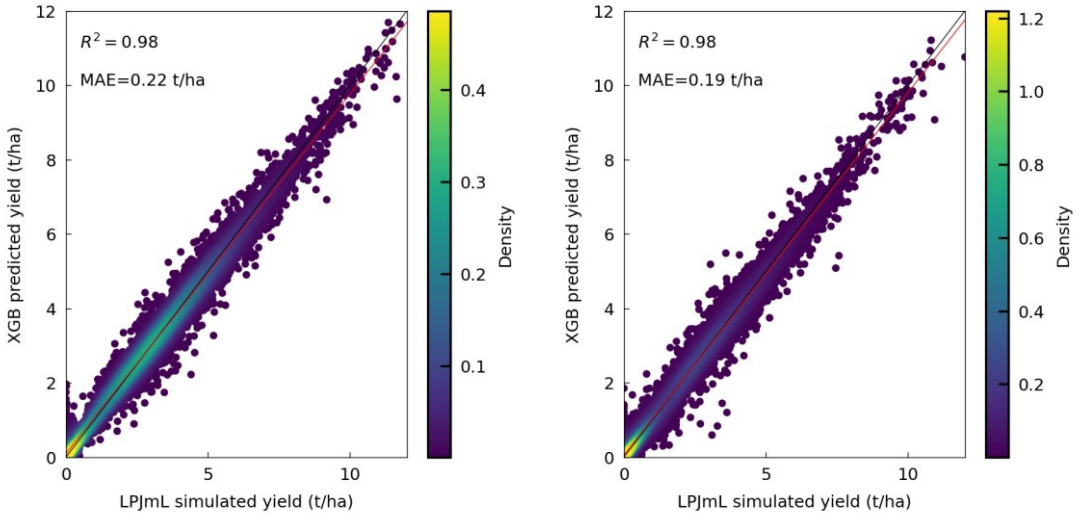

Figure S1 Scatter plots between GGCM simulated yield and emulated yield in testing dataset. Example of LPJmL-A0 for winter wheat. Left and right panel are irrigated and rainfed condition respectively.

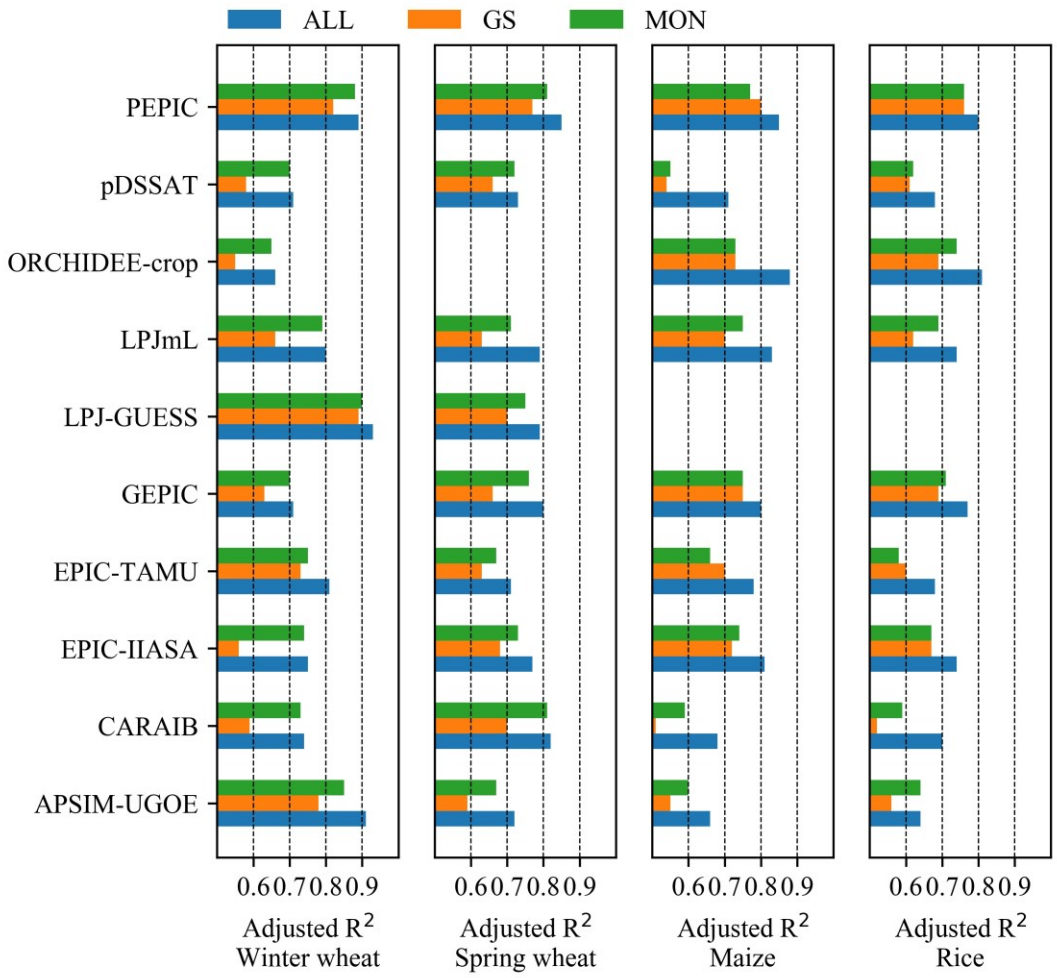

Figure S2 Adjusted R^2 of emulators (leave-one-out cross validation with samples split by Köppen–Geiger climate regions) with different strategery of predictors. All: "Full model", GS: "GS model", Mon: "MON model". Emulators for ORCHIDEE by spring wheat, and LPJ-GUESS by Maize and Rice were not fitted due to the lack of simulation of raw GGCM.

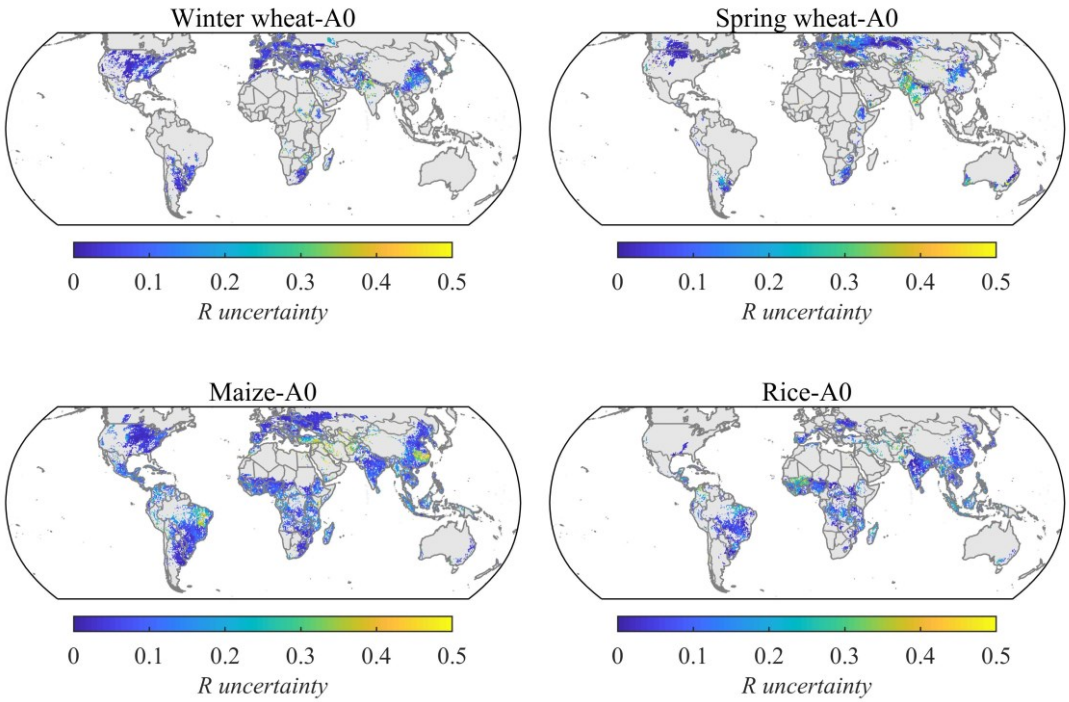

Figure S3 Uncertainty of correlation coefficient (*R*) across multi-model ensemble in the baseline over current cropland. The uncertainty was measured by the standard deviation of *R*s across multi-models.

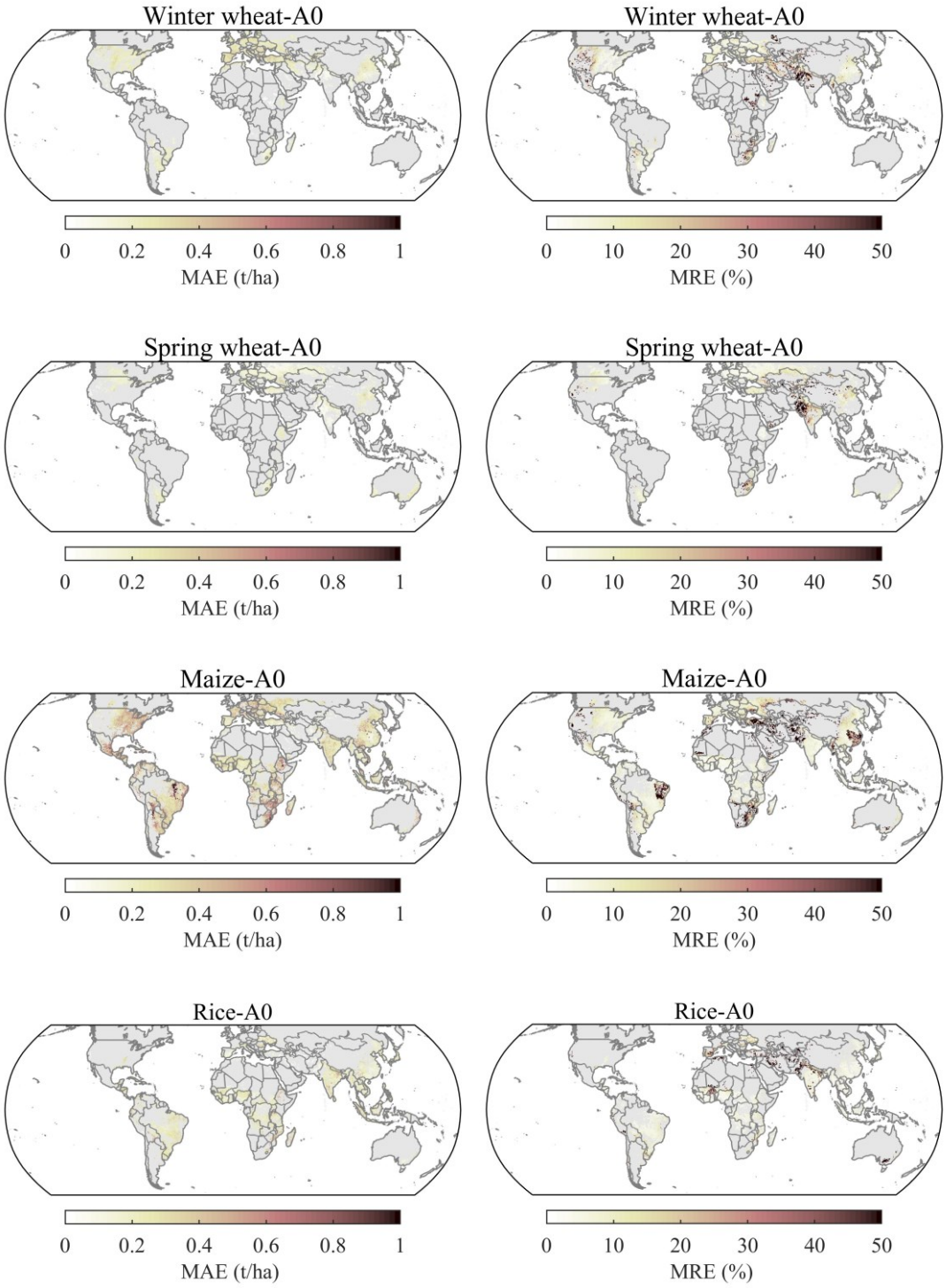

Figure S4 Spatial distribution of MAE and MRE over current croplands in the baseline. MAE: mean absolute error. MRE: mean relative error.

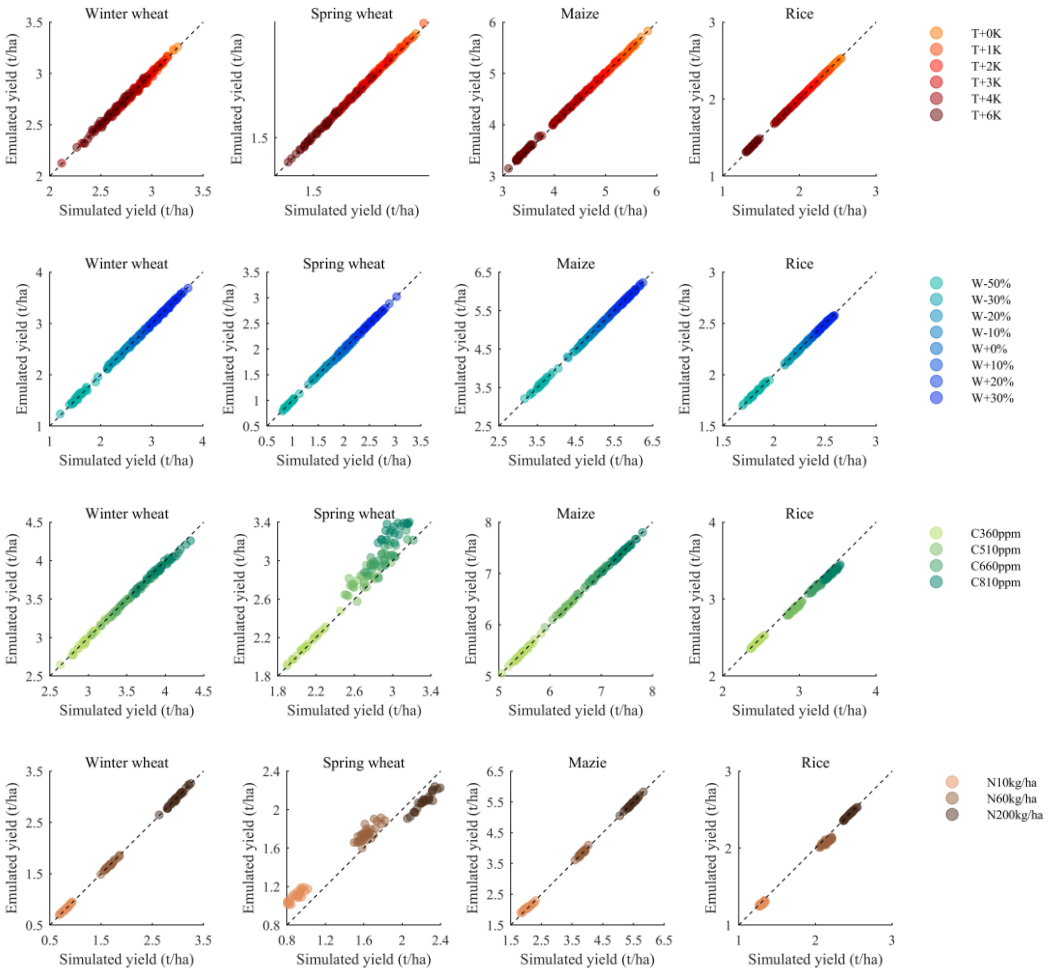

Figure S5 Performance of emulator (EPIC-TAMU) in reproducing the year to year variation of global mean yield from 1981 to 2010 under varied individual CTWN perturbations.

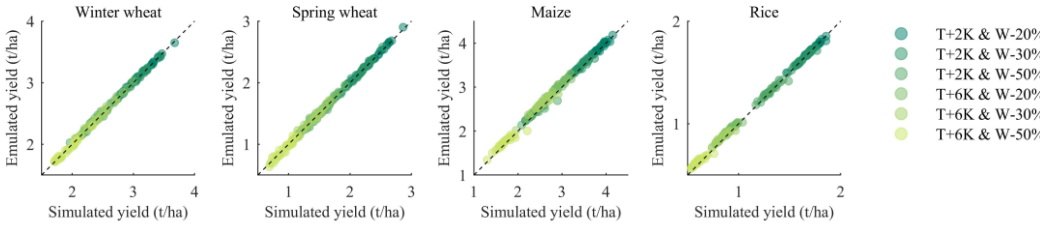

Figure S6 Performance of emulator (LPJmL-A0) in reproducing the year to year variation of global mean yield from 1981 to 2010 under varied T, W perturbations and T+W perturbations.

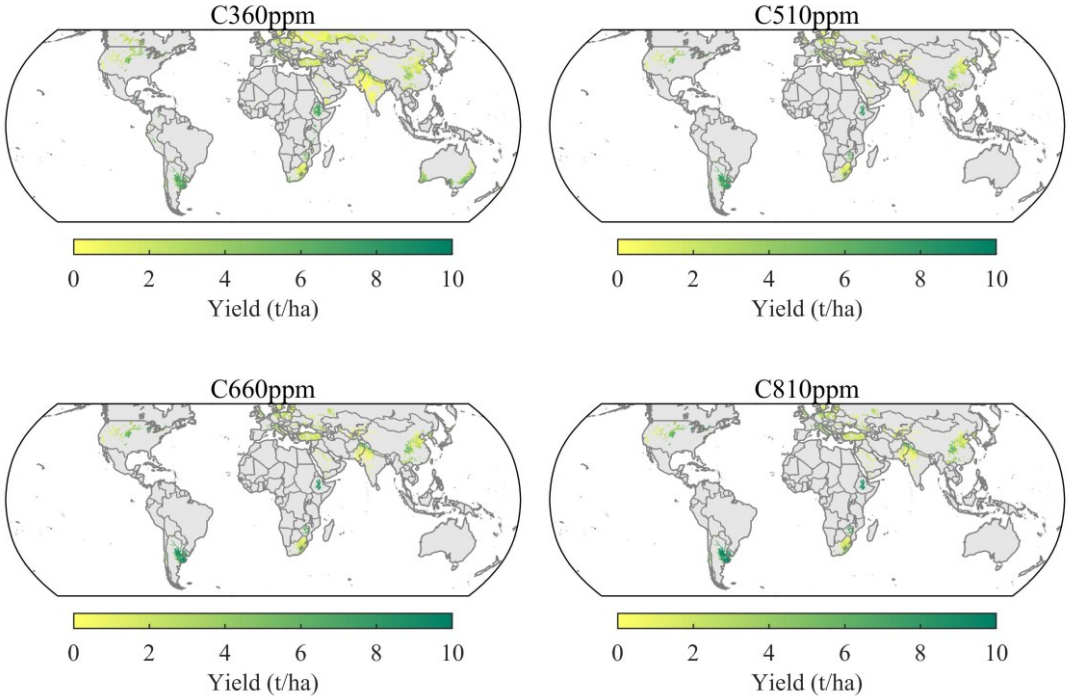

Figure S7 The spatial distribution of simulated yield under CO2 perturbations. The spatial extent has declined sharply under C510ppm, C660ppm and C810ppm relative to the C360ppm. Particularly, in Russia, Canada, India and Australia.

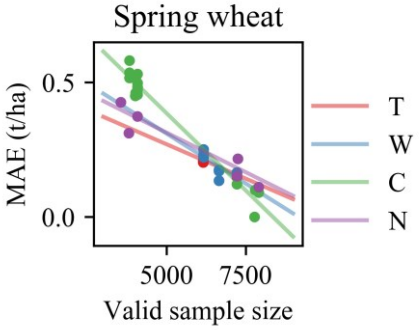

Figure S8 Dependence of median MAE and valid sample size. Dots denote the median MAE of one emulator across all grids under each perturbation. The valid sample size denotes the number of valid gridded yield for each perturbation.

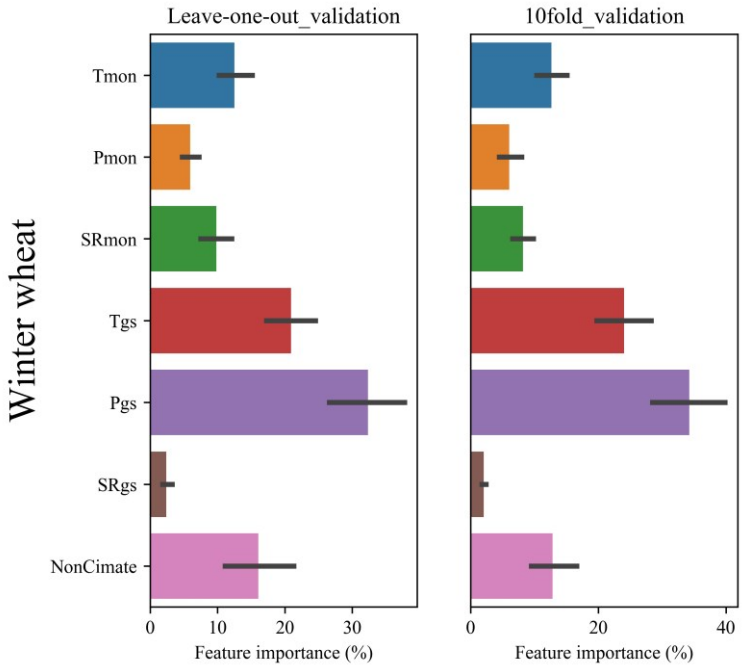

Figure S9 The feature importance of winter wheat emulators given by leave-one-out cross validation with samples split by Köppen–Geiger climate regions (Leave-one-out_validation) and 10-fold cross validation with randomly selected samples (10fold_validation). The bars denote the multi-model ensemble average feature importance. The error bars denote 95% confidence interval estimated from 1000 times bootstrap. Tmon, Pmon, SRmon are temperature, precipitation and solar radiation related variables at monthly scale. Tgs, Pgs, SRgs are temperature, precipitation and solar radiation related variables at growing season scale. NonClimate represent the other variables.

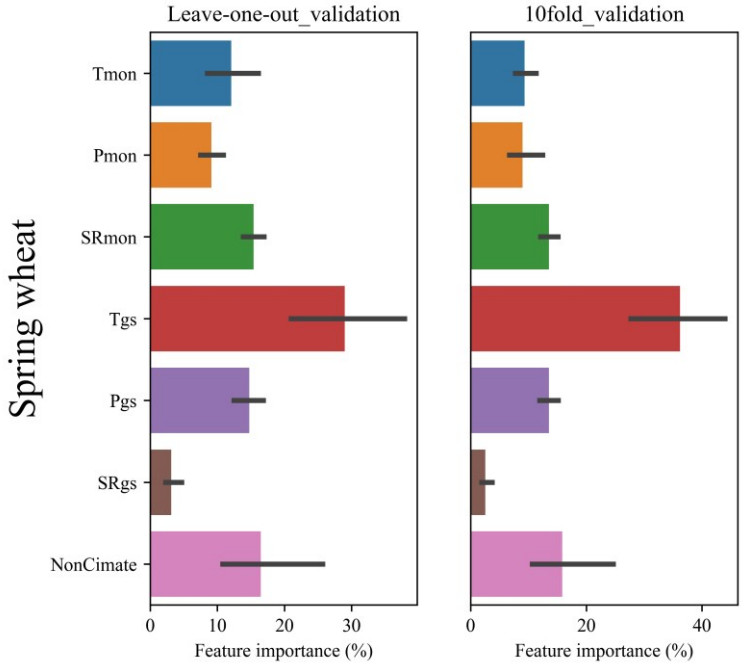

Figure S10 Similar to Figure S9, but for spring wheat.

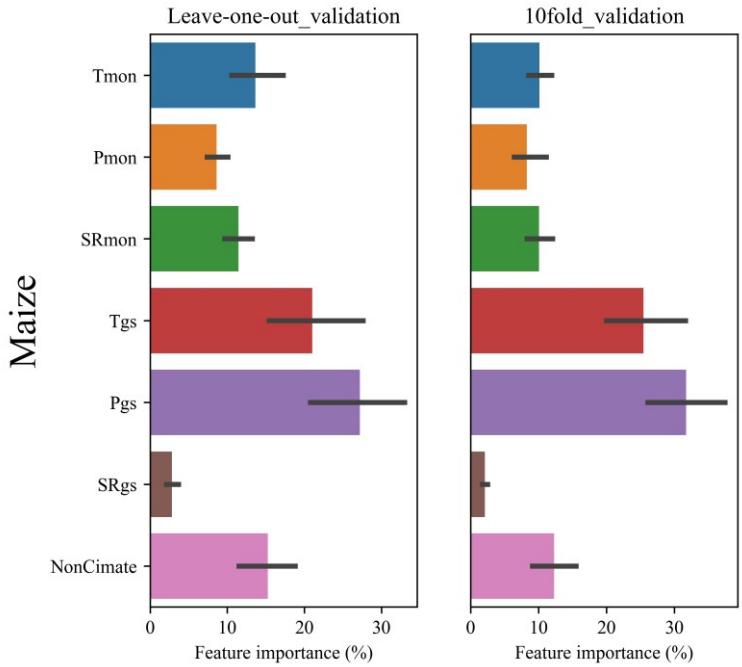

Figure S11 Similar to Figure S9, but for maize

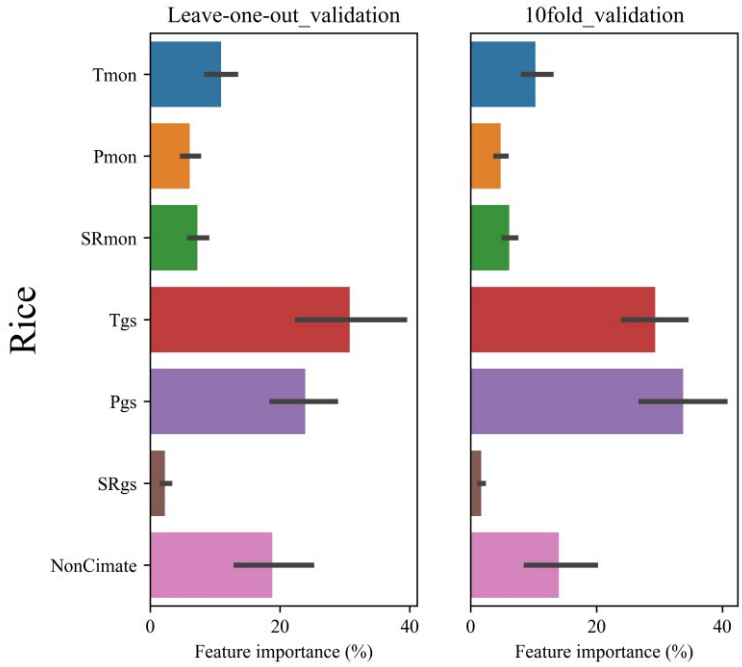

Figure S12 Similar to Figure S9, but for rice