Supplementary Information for

"Stronger-than-expected CO2 fertilization of soybean inferred from paired C3/C4 yield records"

Contents of this file

Table S1. Model performance metrics and CO_2 sensitivity estimates across scenarios.

R^2 and MAPE (Mean Absolute Percentage Error, %) are reported as mean ± SD across 20 random seeds. CO_2 Sensitivity is reported only for models that include CO_2 as a predictor; dashes (—) indicate not applicable.

Crop	Model	Scenario	Train				Test				Train + Test				CO_2 Sensitivity (% ppm^{-1})	
			R^2 mean	R^2 SD	MAPE mean	MAPE SD	R^2 mean	R^2 SD	MAPE mean	MAPE SD	R^2 mean	R^2 SD	MAPE mean	MAPE SD	Median	SD
Maize	Linear	Climate Only	0.551	0.002	19.79	0.08	0.492	0.009	19.83	0.36	0.541	0.000	19.80	0.02	—	—
		Climate + Time	0.787	0.001	13.87	0.07	0.726	0.005	14.79	0.27	0.776	0.000	14.06	0.01	—	—
	Neural Network	Climate Only	0.569	0.013	18.97	0.51	0.457	0.011	20.32	0.55	0.549	0.011	19.25	0.48	—	—
		Climate + Time	0.876	0.005	10.04	0.31	0.764	0.007	13.45	0.32	0.855	0.005	10.74	0.28	—	—
Soybean	Linear	Climate Only	0.598	0.002	16.20	0.05	0.549	0.008	16.44	0.21	0.589	0.000	16.25	0.01	—	—
		Climate + CO_2	0.785	0.001	12.20	0.06	0.729	0.006	13.05	0.23	0.775	0.000	12.38	0.01	0.649	0.003
		Climate + Agronomy Proxy	0.804	0.001	11.65	0.06	0.748	0.006	12.55	0.23	0.793	0.000	11.84	0.01	—	—
		Climate + Agronomy Proxy + CO_2	0.804	0.001	11.62	0.06	0.750	0.006	12.50	0.23	0.795	0.000	11.80	0.01	0.046	0.006
	Neural Network	Climate Only	0.664	0.021	14.39	0.59	0.537	0.010	16.37	0.29	0.640	0.017	14.80	0.47	—	—
		Climate + CO_2	0.887	0.011	8.39	0.56	0.766	0.005	11.89	0.19	0.864	0.009	9.11	0.45	0.615	0.011
		Climate + Agronomy Proxy	0.887	0.009	8.38	0.43	0.770	0.006	11.75	0.25	0.865	0.008	9.07	0.36	—	—
		Climate + Agronomy Proxy + CO_2	0.890	0.011	8.26	0.54	0.772	0.007	11.73	0.23	0.868	0.009	8.98	0.43	0.291	0.019

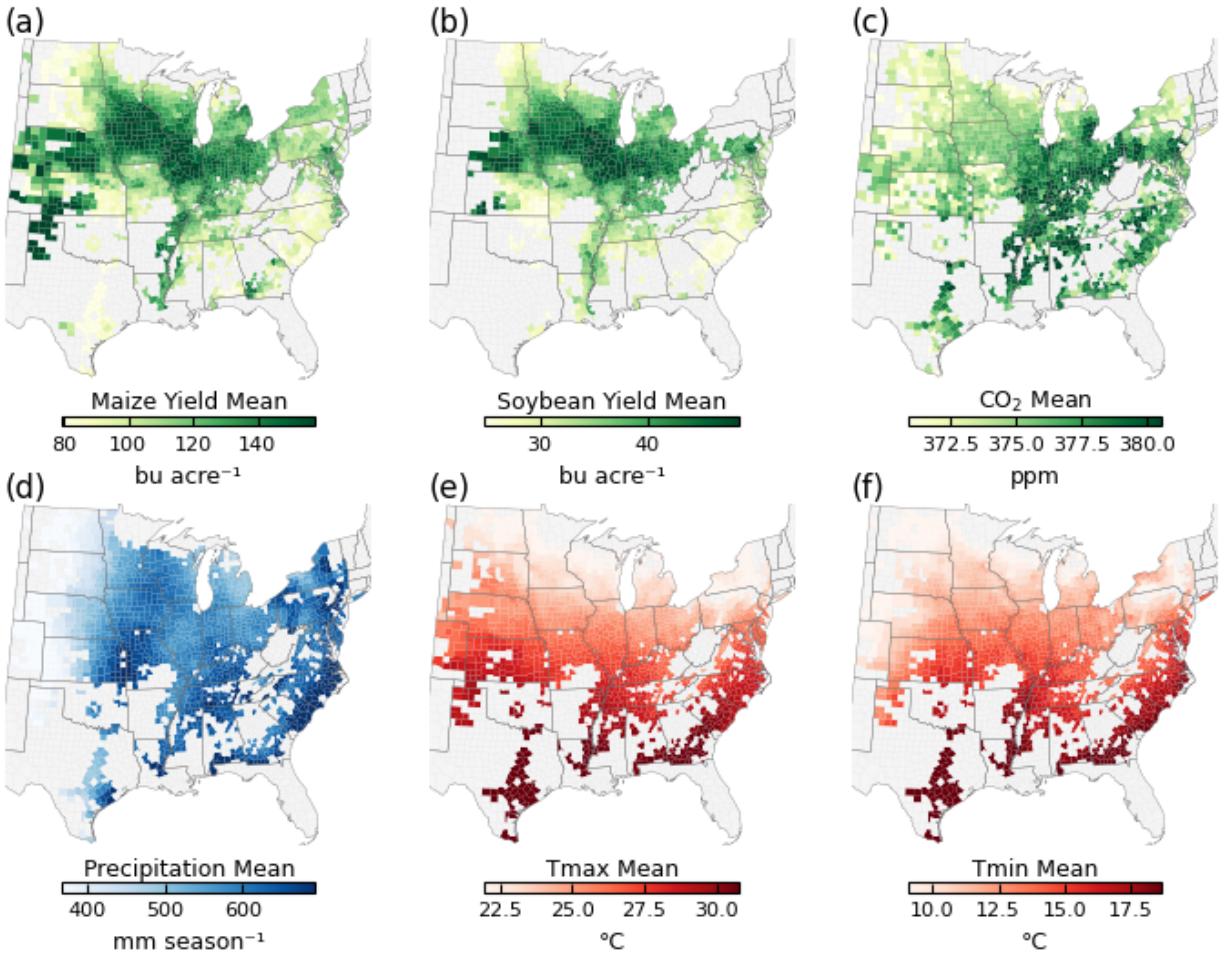

Figure S1: Spatial distribution of county-level temporal means for key variables major agricultural regions in the U.S. (a) Maize yield (bu/acre), (b) soybean yield (bu/acre), growing-season (c) atmospheric CO_2 concentration (ppm), (d) total precipitation (mm), (e) daily maximum temperature (Tmax, °C), and (f) daily minimum temperature (Tmin, °C). All variables are averaged over 1979-2013. Grey counties indicate missing or excluded data.

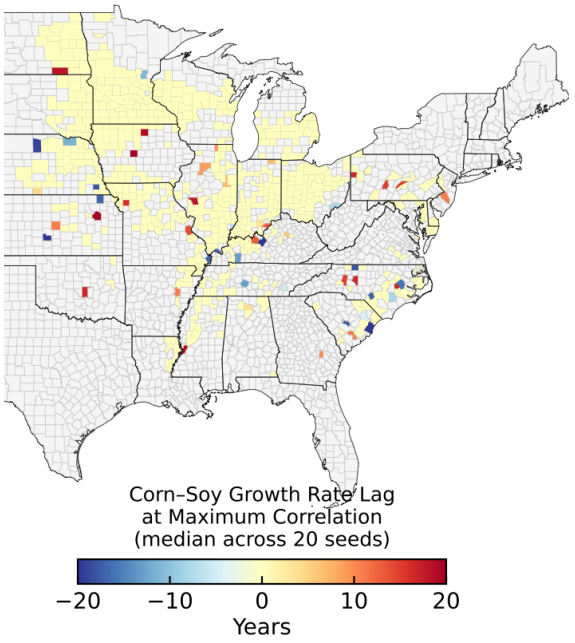

Figure S2. Time lag (years) at maximum cross-correlation between maize and soybean climate residual yield growth rates (median across 20 random seeds). Among 705 reliable counties, 649 (92%) exhibit a median lag of zero, indicating that maize and soybean yield trends are contemporaneous rather than one crop leading the other.

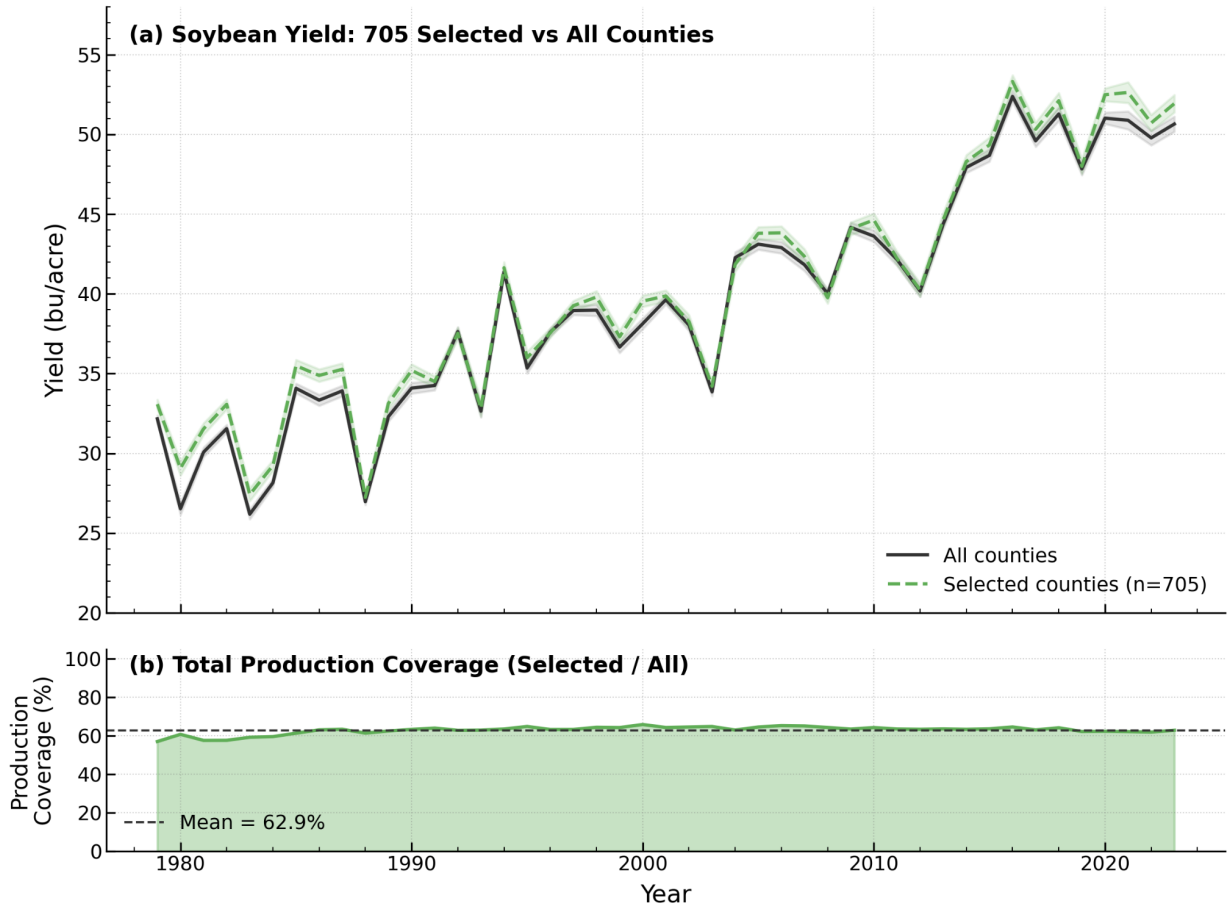

Figure S3. Representativeness of the selected county sample for soybean analysis. (a) Annual mean soybean yield (bu acre^{-1}) for all U.S. counties (black solid line, 1301 counties) and the selected subset (green dashed line, 705 counties) that meet data completeness criteria. (b) Total soybean production coverage of the selected counties relative to all counties, expressed as a percentage. The dashed horizontal line indicates the time-series mean coverage of 62.9%.

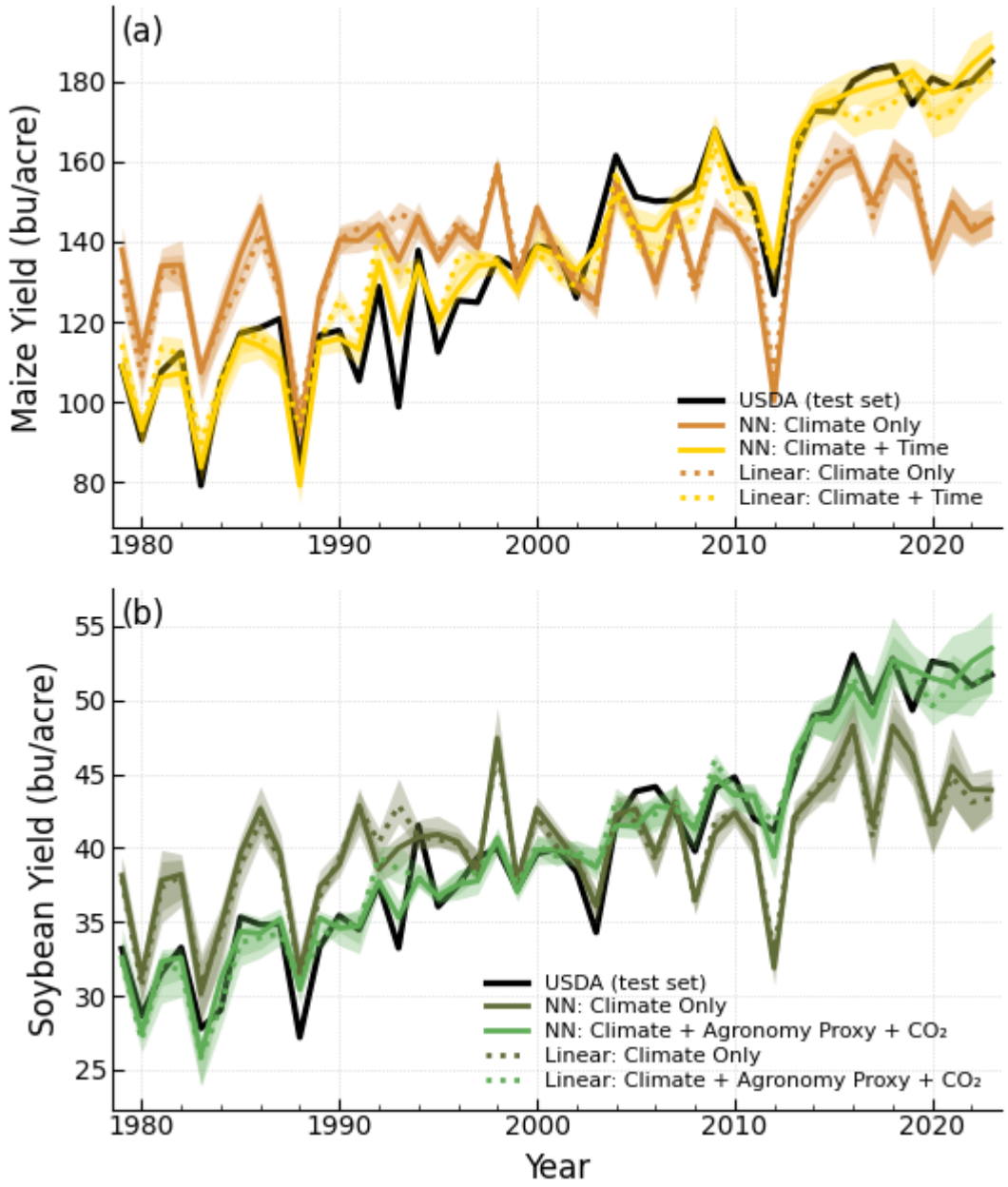

Figure S4. Same as Figure 3a,c but showing out-of-sample (test set) predictions only.

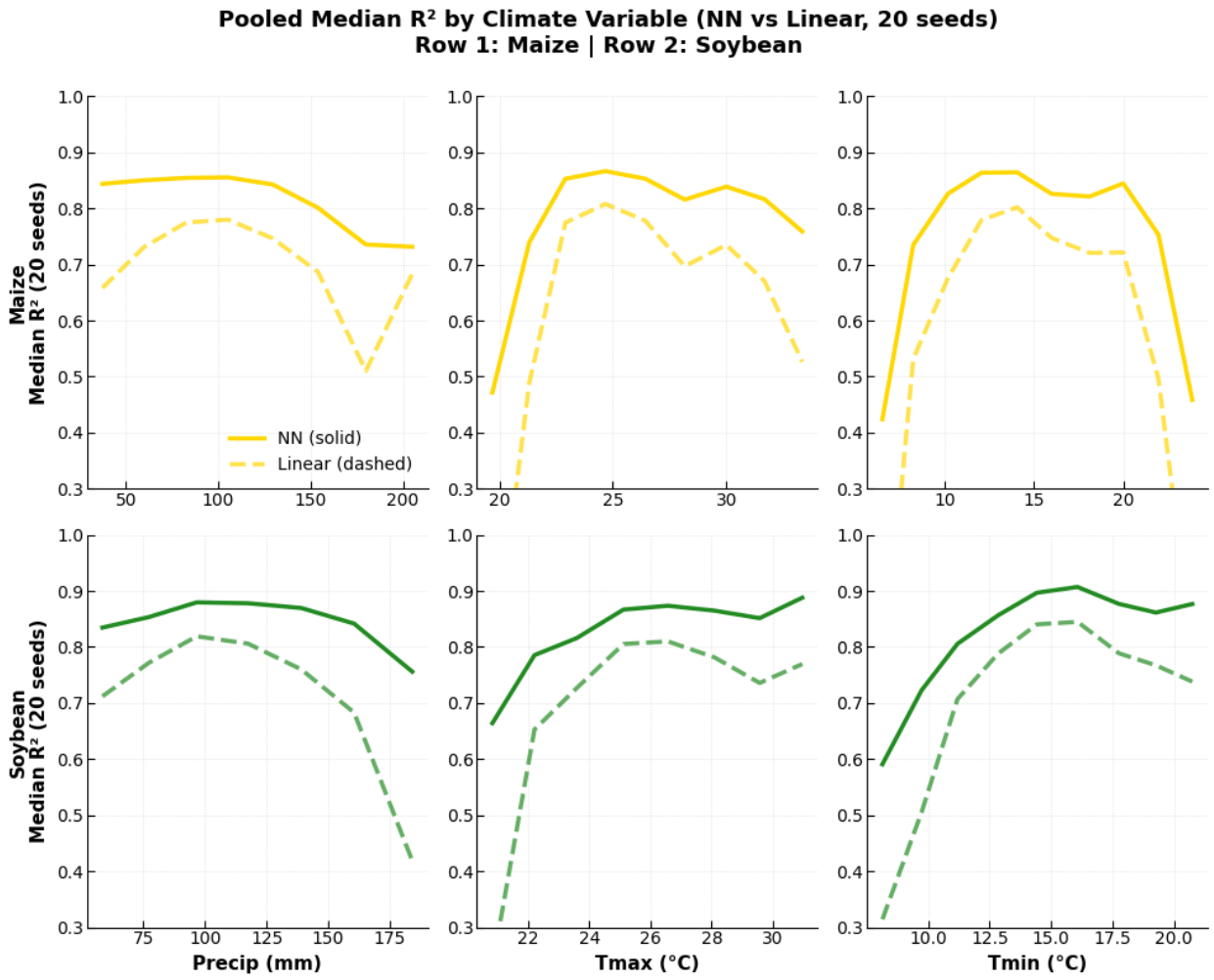

Figure S5. Model R^2 stratified by growing-season climate conditions for maize (top, Scenario 2 with year terms) and soybean (bottom, Scenario 2 with agronomy proxy and CO_2), comparing the linear (dashed) and neural network (solid) models. Counties are binned by growing-season mean precipitation (left), maximum temperature (Tmax, middle), and minimum temperature (Tmin, right); each point shows the median R^2 across all counties and random seeds within that bin.

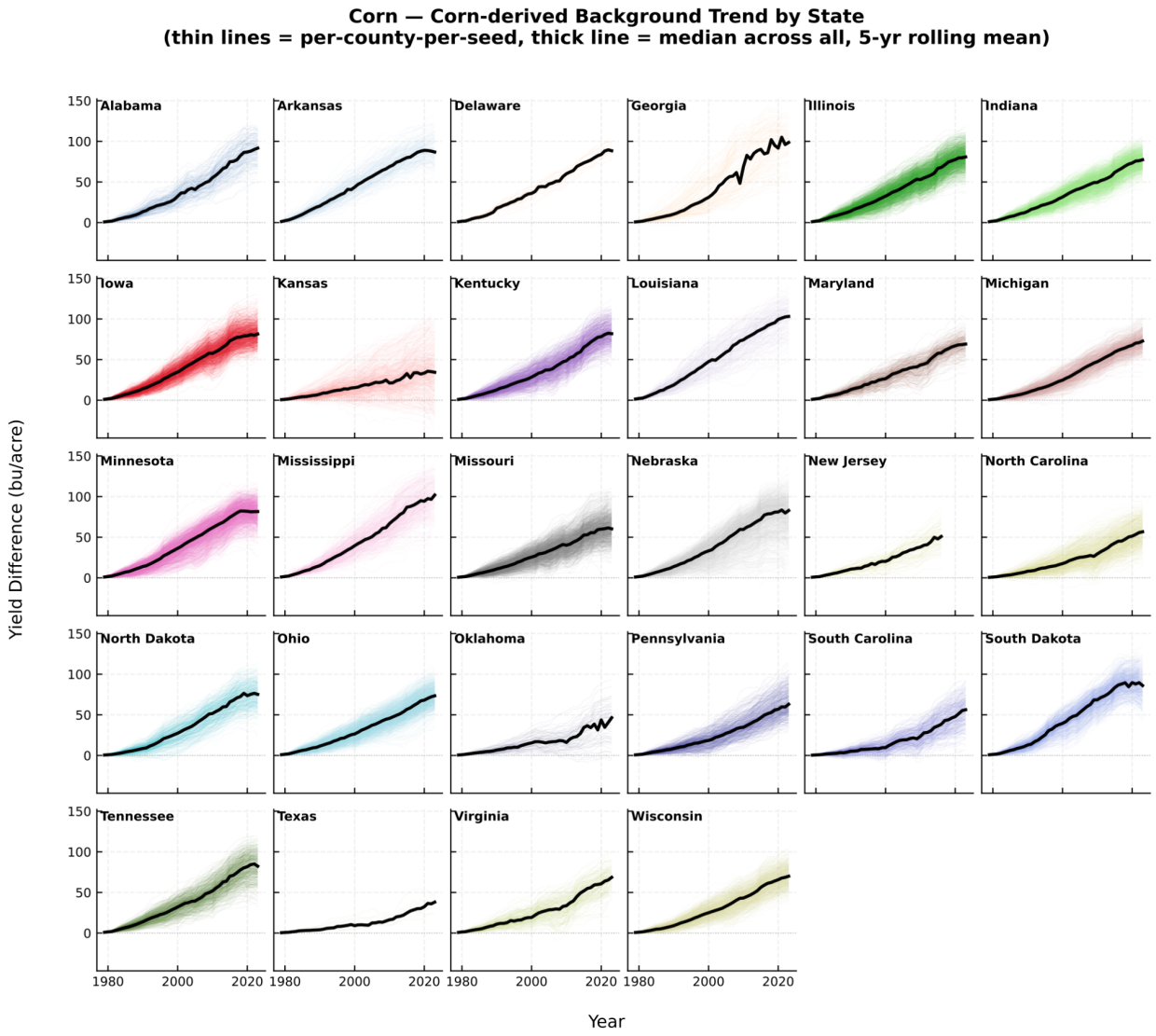

Figure S6 Maize-derived agronomy proxy by state (1979–2023). Each panel shows county-level background trends across 20 random seeds smoothed with a 5-year rolling mean (thin lines, colored by state consistent with Fig. 3b) and the state median (thick black line). The background trend is estimated as the difference between the Scenario 2 neural network predicted yield and the predicted yield with year fixed at 1979, holding climate inputs constant.

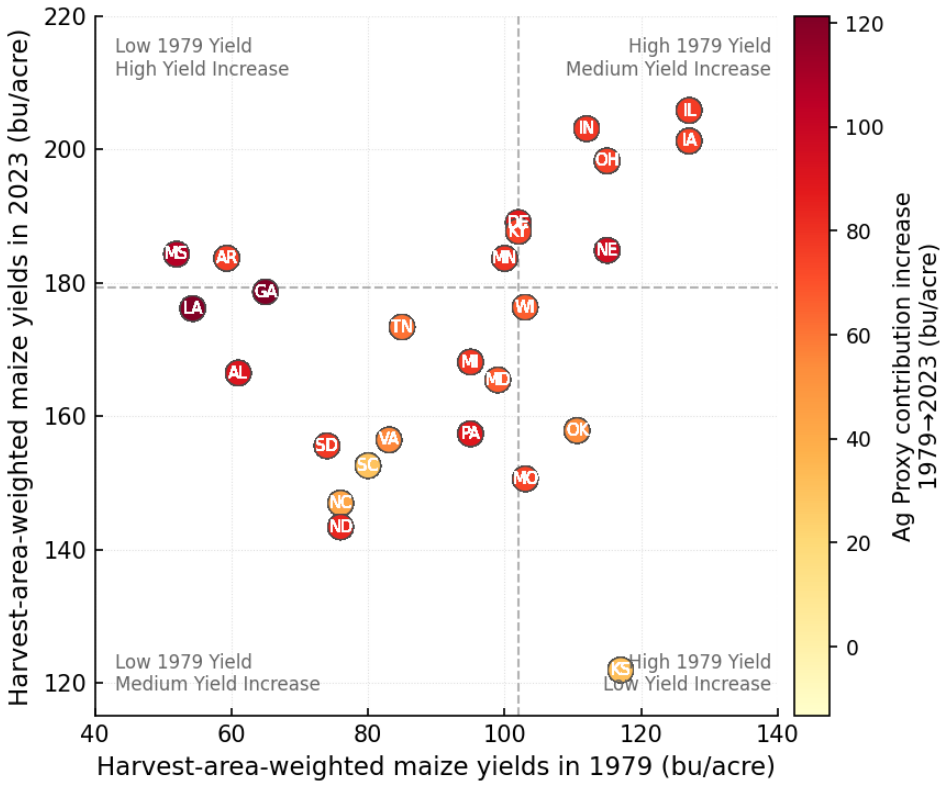

Figure S7. State-level maize yields in 1979 versus 2023, colored by the cumulative agronomic gain (maize-derived agronomy proxy) over the study period. Dashed lines indicate the cross-state median yield in 1979 (vertical) and 2023 (horizontal).

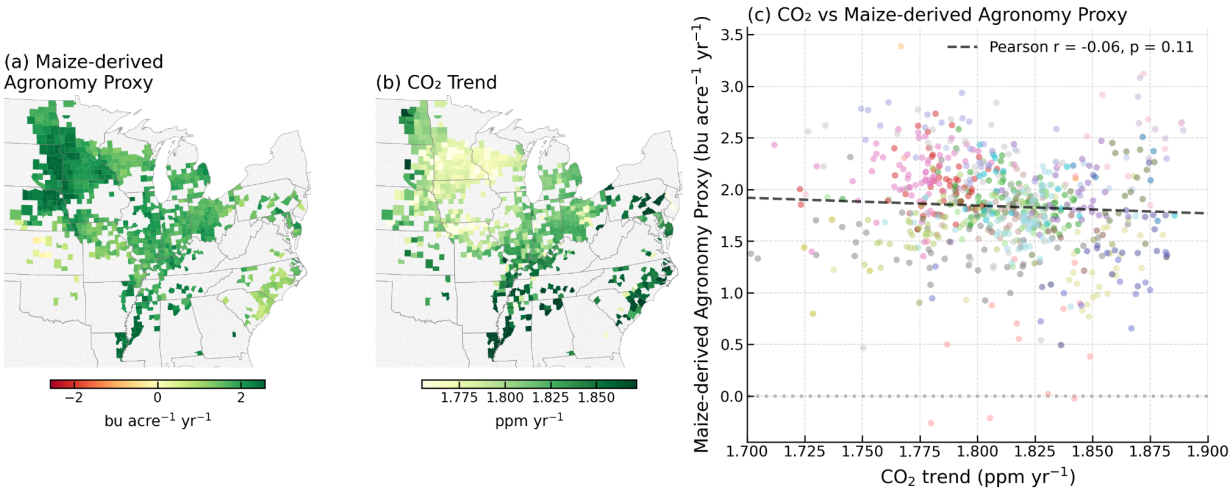

Figure S8. Linear trends county-level (a) maize-derived background trends (bu/acre/yr), (b) growing season mean CO_2 concentration (ppm/yr). (c) Scatter plot of CO_2 trends in (b) versus maize-derived background trends in (a) across all counties, with each point colored by state. The dashed line shows the linear regression fit (Pearson r = −0.06, p = 0.11), indicating that the two signals are not significantly correlated.

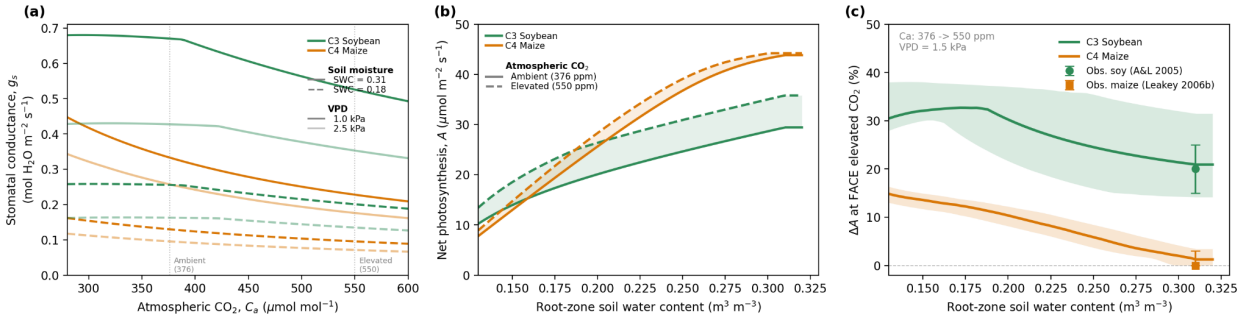

Figure S9. Simulated (a) stomatal conductance as a function of atmospheric CO_2 concentration, (b) net photosynthesis as a function of root-zone soil moisture under ambient (376 ppm) and elevated (550 ppm) CO_2, and (c) the increase in net photosynthesis under elevated CO_2 as a function of root-zone soil moisture, for C3 soybean (green) and C4 maize (orange). In (a), line styles indicate different soil moisture and VPD conditions. Error bars in (c) show FACE-measured photosynthetic responses for soybean (Ainsworth & Long, 2005) and maize (Leakey et al., 2006), confirming consistency with field observations.

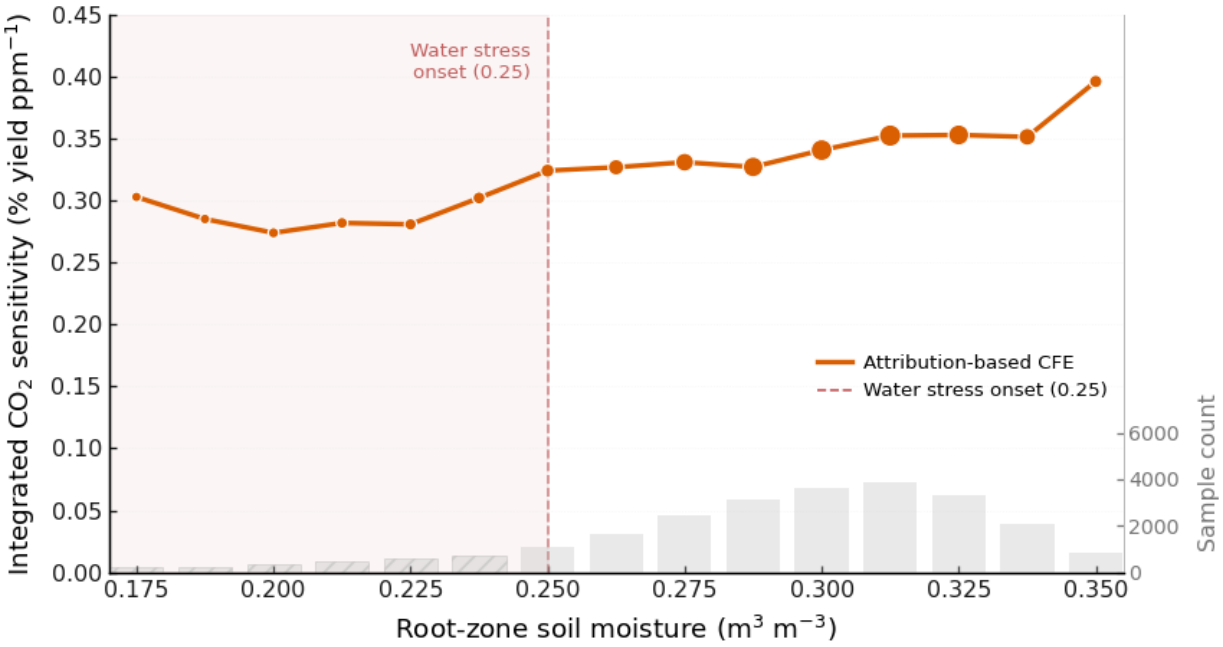

Figure S10 Integrated CFE of soybean as a function of root-zone soil moisture (SMrz). Each county's integrated CFE (CO_2 contribution to yield gain ÷ mean yield ÷ CO_2 increase, in % yield ppm^{-1}) is paired with the growing-season mean root-zone soil moisture of every county-year observation. Points show the bin-median CO_2 sensitivity (bin width = 0.0125 m^3 m^{-3}); point size is proportional to sample count per bin (grey bars, right axis). The dashed red line marks the water stress onset threshold (0.25 m^3 m^{-3})

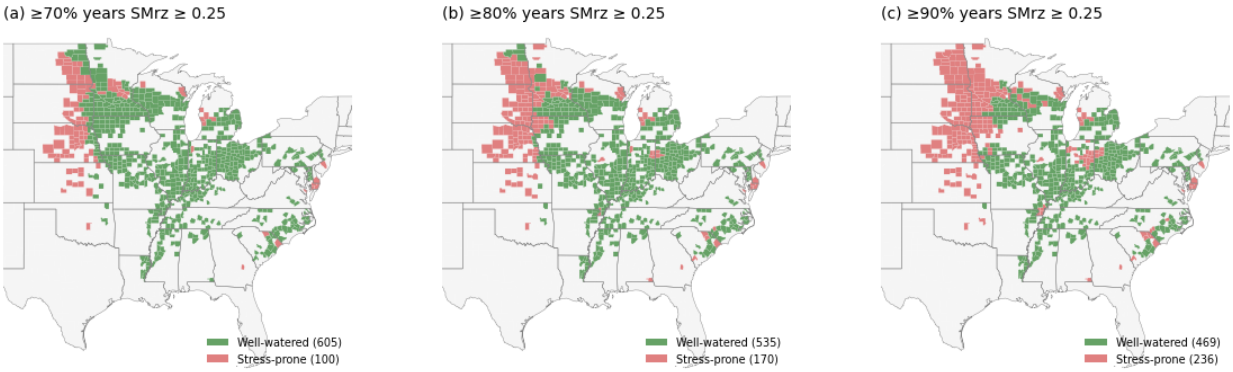

Figure S11 Water stress classification of 705 reliable counties under three thresholds. Counties are classified as well-watered (green) if root-zone soil moisture (SMrz) exceeds 0.25 $m^3\ m^{-3}$ in at least (a) 70%, (b) 80%, or (c) 90% of growing seasons during 1979–2023, and as stress-prone (red) otherwise. The 0.25 $m^3\ m^{-3}$ threshold corresponds to the onset of plant water stress from Figure S10.

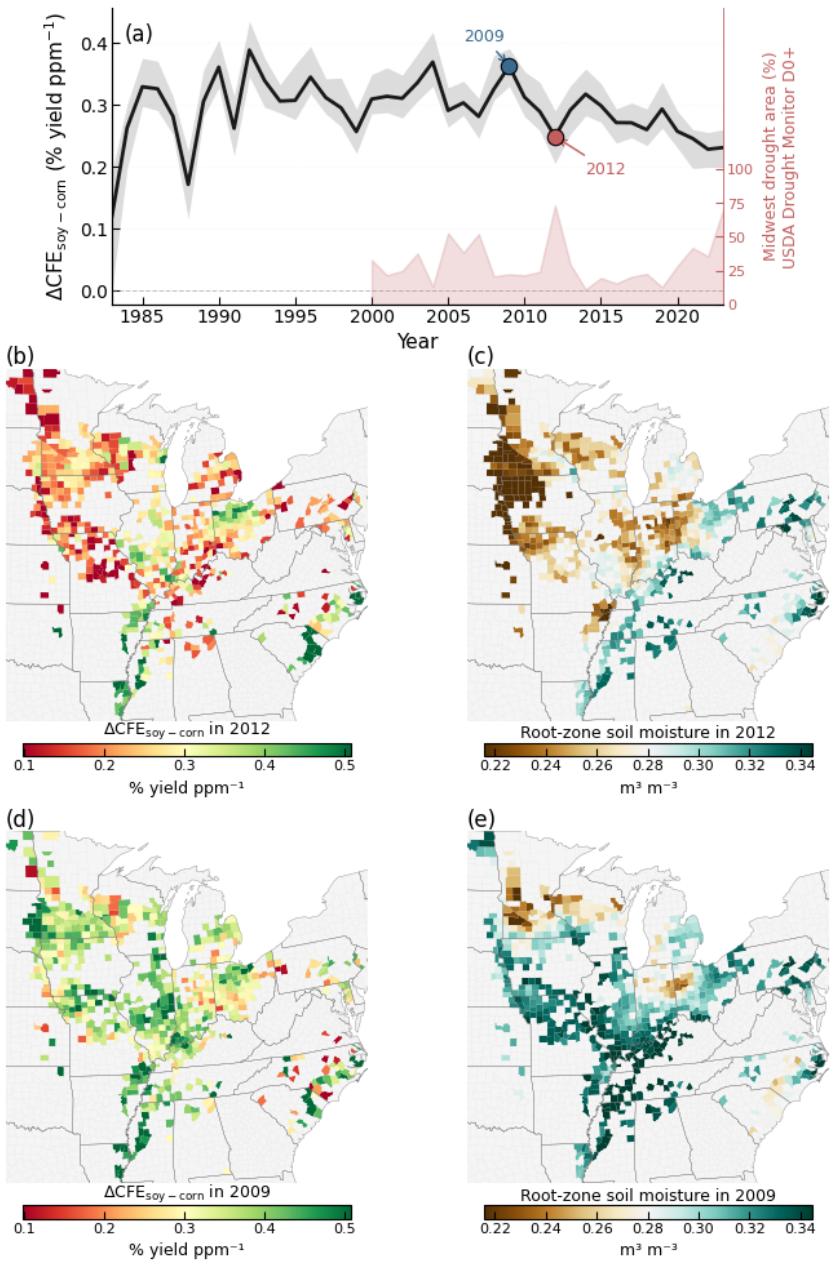

Figure S12 Drought suppresses the differential CO_2 fertilization effect (ΔCFE). (a) Annual median ΔCFE (soybean − maize) across 605 well-watered counties (black line; shading = 5–95th percentile across 20 model seeds), with Midwest drought percent area for D0-D4 category from the USDA Drought Monitor (red fill, right axis). The 2012 drought year shows a pronounced dip in ΔCFE coinciding with peak drought extent. (b–c) County-level ΔCFE and root-zone soil moisture in 2012; (d–e) the same for 2009, a wet reference year. The shared color scales highlight that drier conditions in 2012 (brown tones in c) correspond to lower ΔCFE values (red/orange in b), while the wetter 2009 (teal in e) sustains higher ΔCFE (green in d).

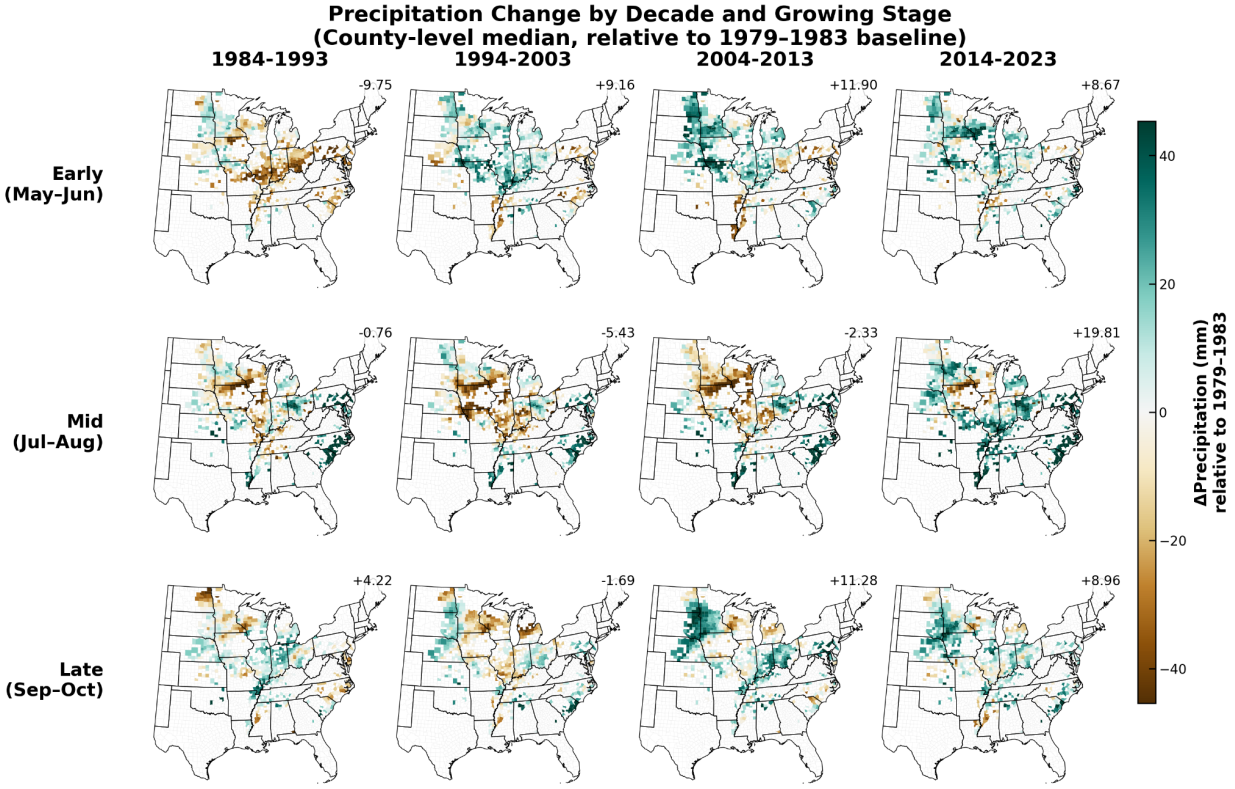

Figure S13. County-level change in growing-season precipitation relative to the 1979–1983 baseline, by decade and stage (median across years within each decade, 705 reliable counties). Positive (blue-green) values indicate wetter conditions; negative (brown) values indicate drier conditions.

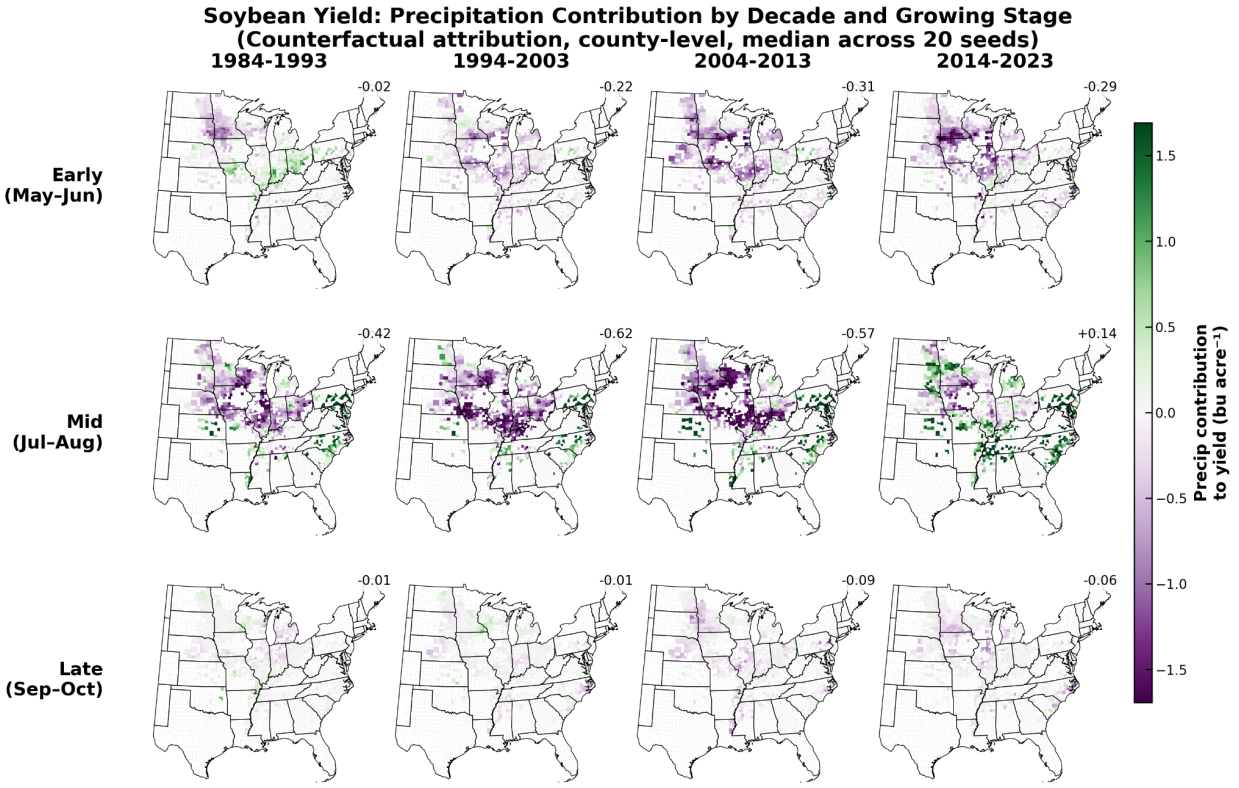

Figure S14. County-level precipitation contribution to soybean yield by growing stage and decade, estimated via counterfactual attribution (median across 20 random seeds). Each panel shows the yield change (bu acre^{-1}) attributable to precipitation deviations from the 1979–1983 baseline for a single growing stage: Early (May–Jun), Mid (Jul–Aug), and Late (Sep–Oct).

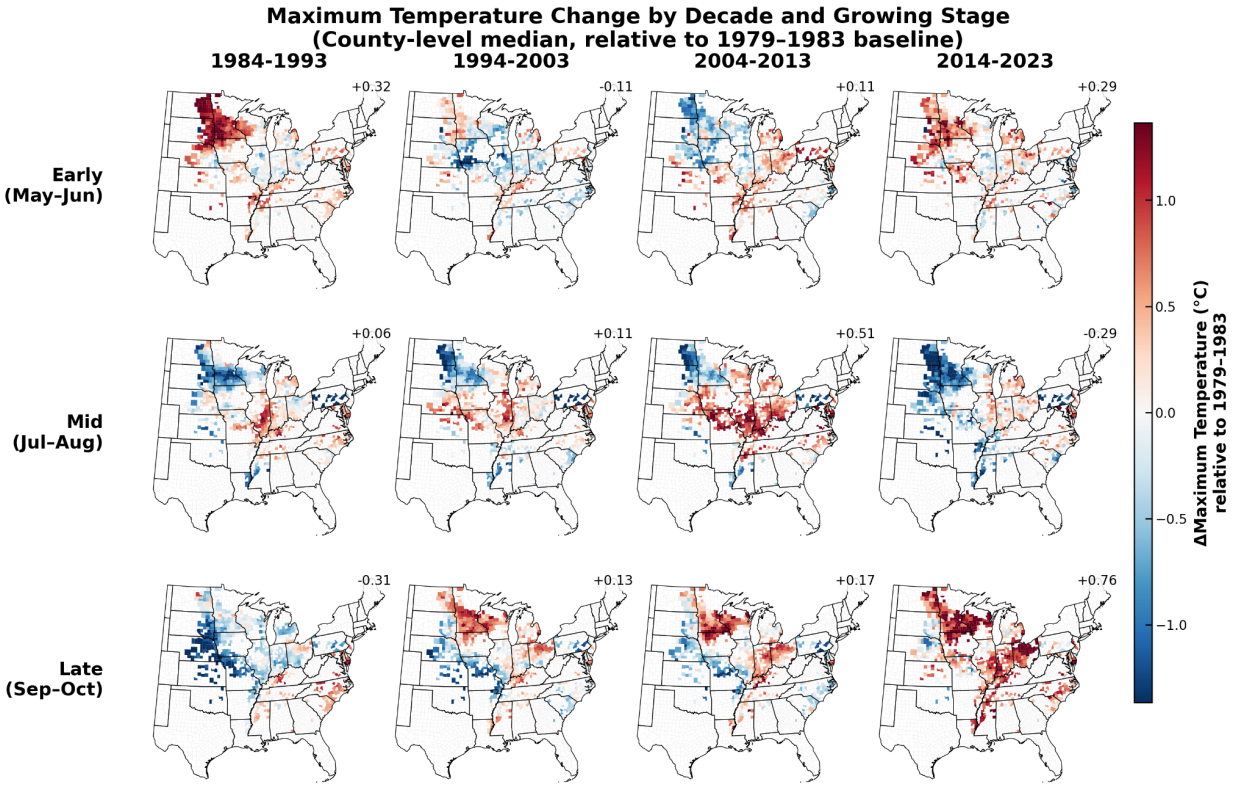

Figure S15. Same as Figure S13 but for maximum temperature (Tmax). Positive (red) values indicate warming; negative (blue) values indicate cooling.

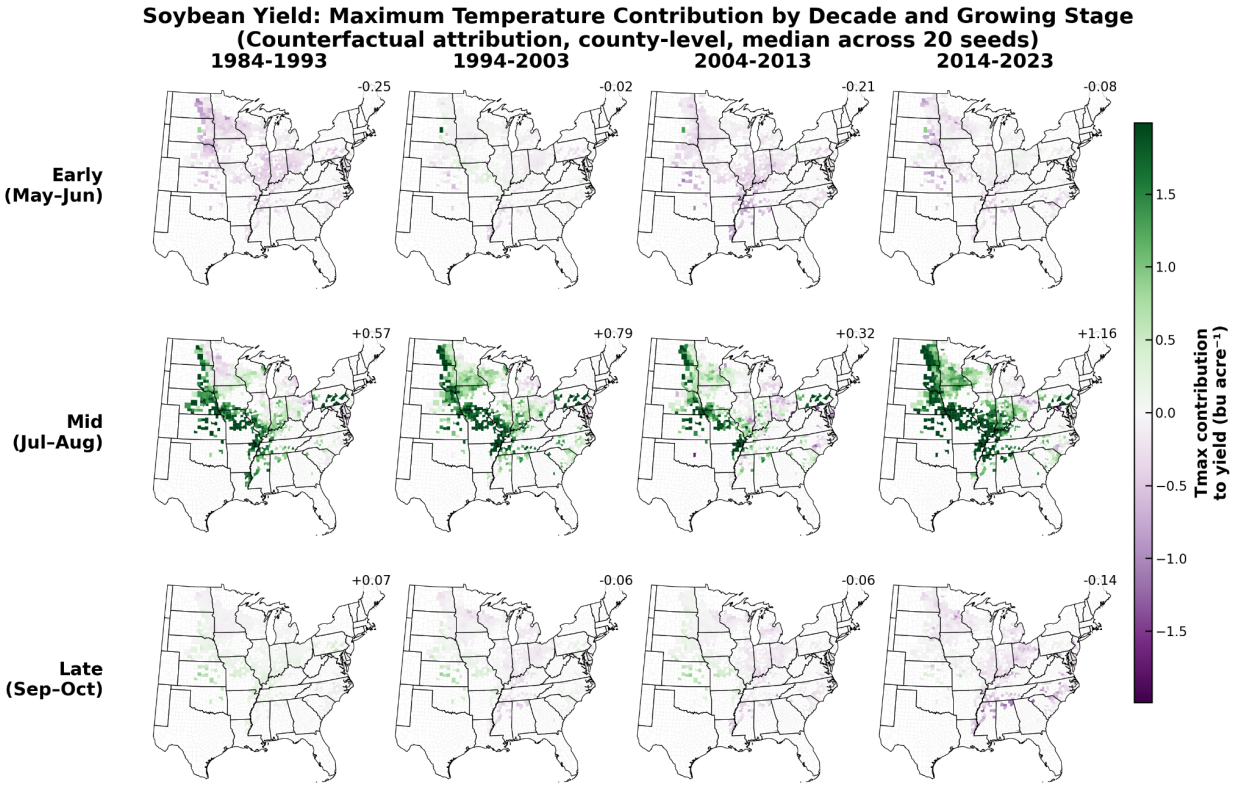

Figure S16. Same as Figure S14 but for maximum temperature (Tmax).

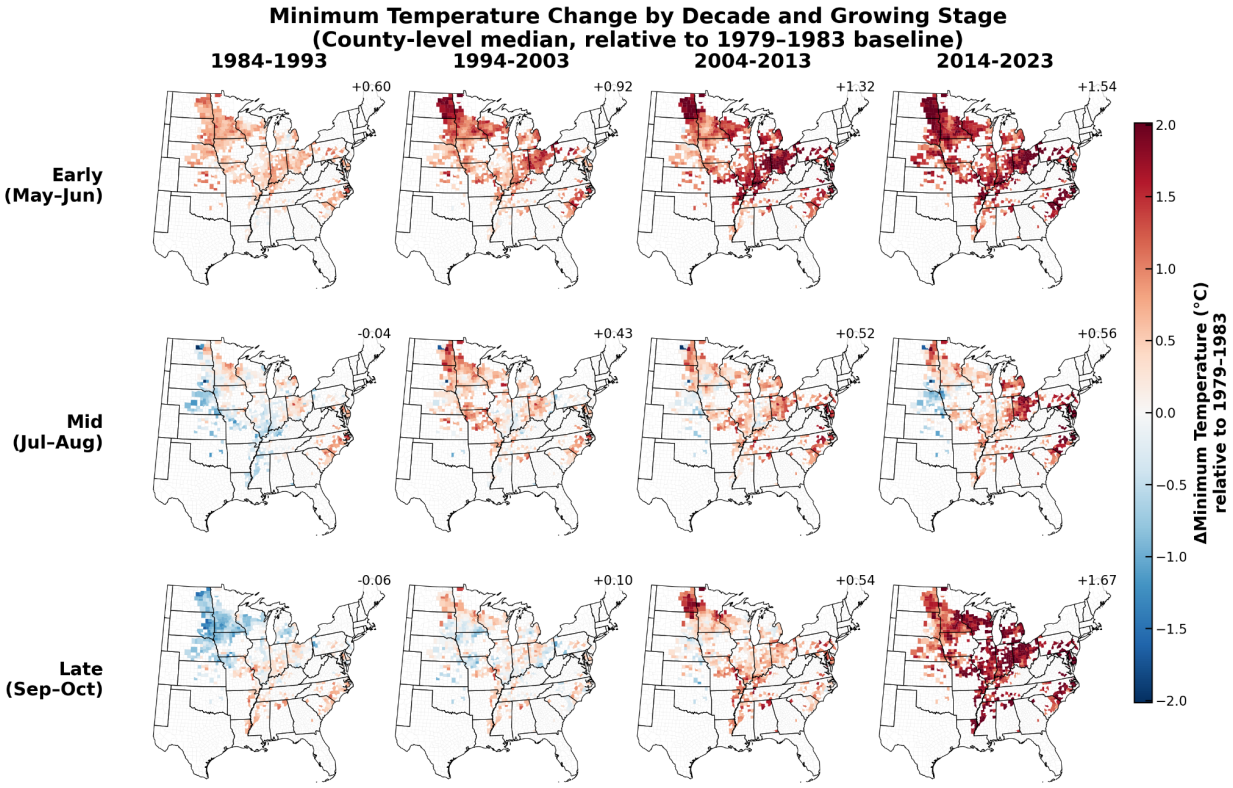

Figure 17. Same as Figure S13 but for minimum temperature (Tmin).

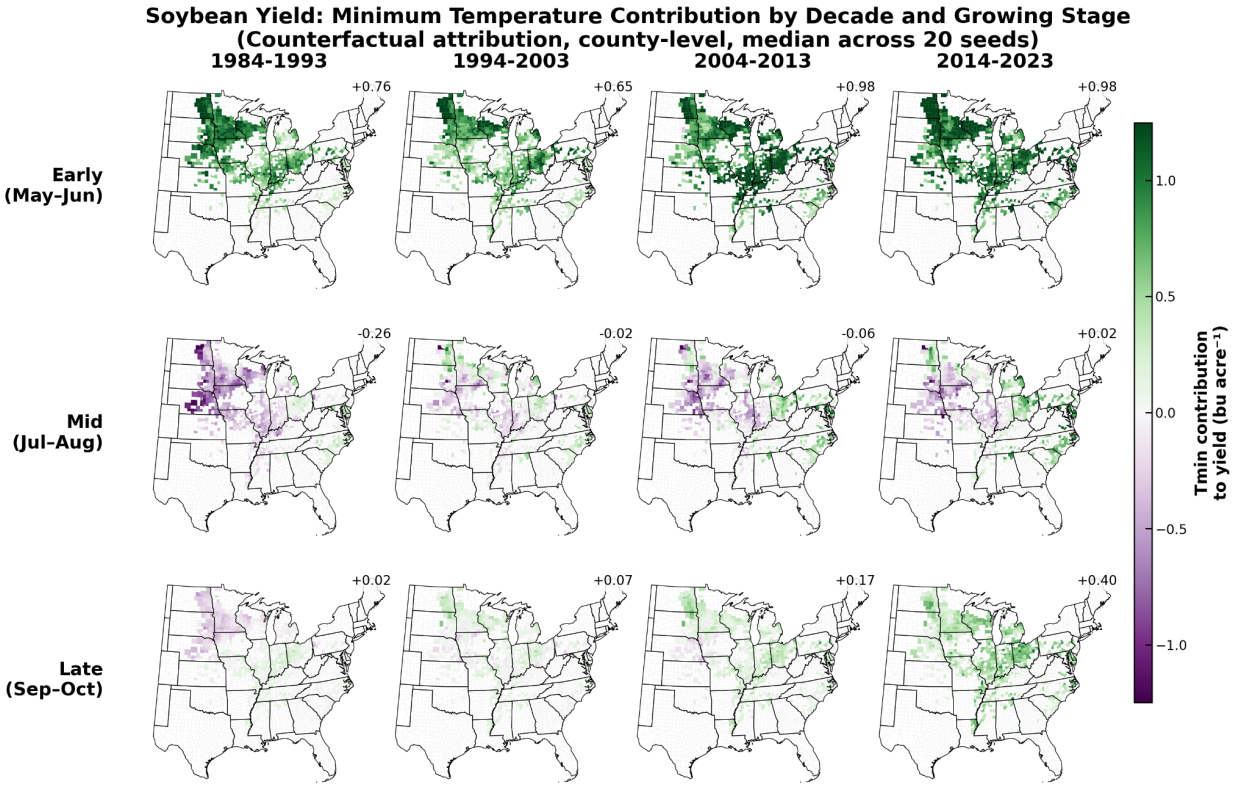

Figure S18. Same as Figure S14 but for minimum temperature (Tmin).

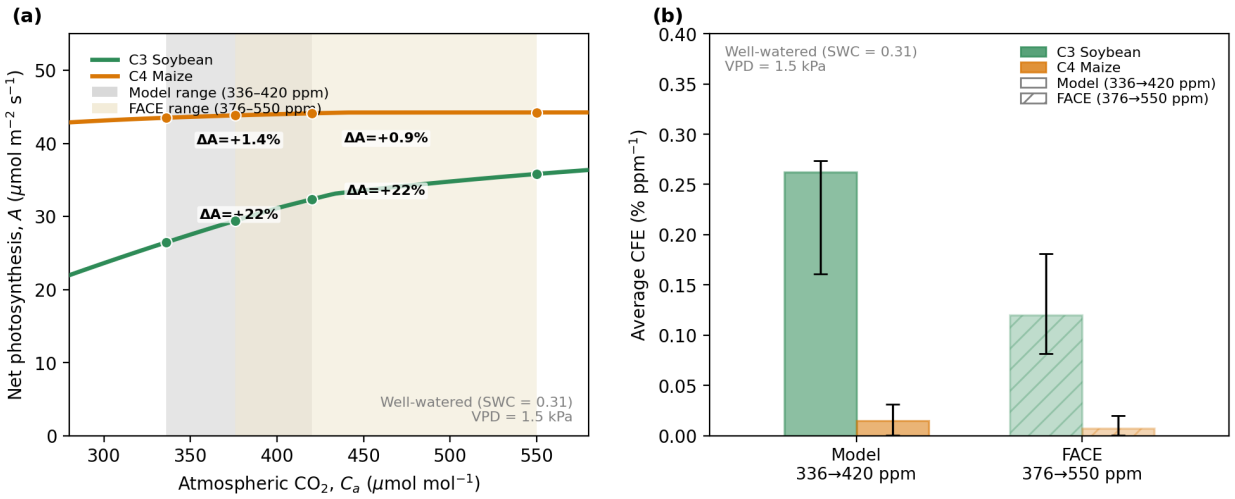

Figure S19 Concavity of the A/C_a curve explains much of the gap between observational and FACE-based CFE estimates. (a) Modeled net photosynthesis vs atmospheric CO_2 under well-watered conditions (SWC = 0.31, VPD = 1.5 kPa). The shaded bands mark the observational range (336→420 ppm, grey) and the FACE range (376→550 ppm, orange). Soybean (C3) gains 22% in photosynthesis over both ranges but from different baseline levels; maize (C4) gains <2%. (b) Average CFE (% yield ppm^{-1}) over the two CO_2 ranges.