Supplement

S1 Land-surface representation of LULCC and ecohydrological fluxes in the ESMs

Particularly important for simulating LULCC impacts on precipitation partitioning are the land-surface processes that control (i) how precipitation is converted into runoff (blue water flow) and (ii) how vegetation returns water to the atmosphere via transpiration (green water flow). Across the four ESMs – BCC-CSM2-MR, CMCC-ESM2, MIROC-ES2L, and UKESM1-0-LL – these processes are implemented with differing degrees of process detail in the respective LSMs, which is directly relevant for interpreting model differences in the BGWS response.

S1.1 BCC-CSM2-MR (LSM: BCC_AVIM2.0)

BCC_AVIM2.0 computes transpiration from photosynthesis–stomatal conductance coupling (Farquhar biochemistry with Ball-Berry/Collatz stomata) within a canopy transfer framework; soil-moisture limitation is applied as a root-weighted stress, so green-water fluxes respond directly to both environmental drivers and root-zone moisture (Wu et al., 2019; Li et al., 2019). Relative to earlier versions, BCC_AVIM2.0 includes a prognostic phenology based on plant carbon budgets (leaf-out/senescence and LAI evolution), revised canopy radiation transfer, updated snow/albedo schemes, and a rice paddy parameterisation, all of which affect seasonal transpiration (Li et al., 2019). Runoff follows conventional saturation- and infiltration-excess generation embedded in a CLM-style multilayer soil hydrology: precipitation first partitions to canopy interception and infiltration; surface runoff occurs when rainfall exceeds infiltration capacity or when near-surface layers saturate; vertical percolation and subsurface drainage provide baseflow (Li et al., 2019; Wu et al., 2019). Unlike TOPMODEL/PDM formulations, BCC_AVIM2.0 does not diagnose a subgrid saturated-area fraction explicitly; blue-water generation is therefore governed primarily by vertical soil moisture states and parameterised drainage rather than a prognostic water-table/saturated-area diagnostic (Li et al., 2019). In CMIP6 configurations, transient land-cover/PFT fractions are updated annually from the scenario forcing and mapped to BCC_AVIM vegetation types; LAI and carbon states then evolve prognostically via the phenology and carbon allocation schemes (Wu et al., 2019; Li et al., 2019). Thus, transpiration responds both to imposed PFT redistribution and to climate/CO_2 variability (Li et al., 2019).

S1.2 CMCC-ESM2 (LSM: CLM4.5-BGC)

CLM4.5 in its biogeochemical (BGC) configuration calculates transpiration separately for sunlit and shaded leaves, coupling leaf photosynthesis to Ball-Berry stomata and aggregating to canopy conductance; water-stress effects act through soil water potential, with boundary-layer coupling treated explicitly (Oleson et al., 2013; Koven et al., 2013). Photosynthesis follows Farquhar et al. (1980) biochemistry for C3 plants and Collatz et al. (1992) for C4 plants, with phenology and litterfall simulated prognostically in response to temperature, soil moisture, and photoperiod (Oleson et al., 2013; Lovato et al., 2022). Runoff uses a TOPMODEL-class "saturated fraction" concept (SIMTOP) to produce saturation-excess (Dunne) runoff as a function of water-table depth and a topographic index; an alternative VIC-type option introduces a three-layer soil structure, an explicit

surface-water store, and subgrid microtopography, increasing sensitivity to event intensity and partitioning between infiltration and overland flow (Niu et al., 2005; Oleson et al., 2013). Transient LULCC is implemented by annually updating subgrid PFT fractions in CLM's landunit-column-PFT tiling, conserving mass and energy when cover transitions occur. Vegetation fractions are prescribed by the scenario forcing, whereas LAI within each PFT is prognostic, so transpiration responds to both imposed PFT redistribution and to climate/CO_2 variability. Fire disturbance is simulated via a prognostic fire module, including agricultural, deforestation, and peat fires, which interact with vegetation turnover (Oleson et al., 2013; Lovato et al., 2022).

S1.3 MIROC-ES2L (LSM: MATSIRO (physics); VISIT-e (biogeochemistry))

MATSIRO employs a single-layer canopy energy-water scheme with Farquhar-type photosynthesis and a stomatal resistance formulation ported from SiB2 (Sellers et al., 1996); soil-moisture stress is applied via root-weighted availability (Takata et al., 2003; Guo et al., 2021). Vegetation composition (land-cover and PFT fractions) is prescribed annually from the LUH2 (Hurtt et al., 2020; Ma et al., 2020), representing five LUH2 cover types with no dynamic competition among PFTs (Hajima et al., 2020). These cover fractions are mapped to MATSIRO/VISIT-e vegetation types, while LAI within each type is prognostic and exchanged between MATSIRO and VISIT-e alongside soil water content and runoff (Ito and Inatomi, 2012). Runoff is computed with a simplified TOPMODEL approach in which the saturated fraction is diagnosed from grid-mean water-table depth; surface runoff arises from saturation-excess (Dunne) and infiltration-excess (Hortonian) when near-surface storage is exceeded, while groundwater/baseflow depends on water-table depth and transmissivity, coupling vertical soil moisture to lateral blue-water production (Takata et al., 2003; Guo et al., 2021). LULCC affects hydrology and vegetation by redistributing prescribed cover fractions: clearing events reduce vegetation, while abandonment permits regrowth toward a new equilibrium state according to VISIT-e's growth functions; crop and pasture parameterisations remain fixed apart from fractional area changes, and there is no explicit irrigation or crop management (Hajima et al., 2020).

S1.4 UKESM1-0-LL (LSM: JULES-ES-1.0 with TRIFFID)

JULES derives canopy transpiration from Penman-Monteith using PFT-specific leaf photosynthesis (Farquhar for C_3, Collatz for C_4) coupled to Ball-Berry/Collatz stomata, consistent with the interactive carbon-nitrogen cycle and dynamic vegetation (TRIFFID) (Best et al., 2011; Clark et al., 2011; Sellar et al., 2019). Soil-moisture stress is applied through PFT-dependent availability functions, feeding back on stomatal conductance. Runoff is generated using either a TOPMODEL parameterisation, in which the saturated fraction and baseflow depend on water-table depth and transmissivity, or the Probability Distributed Model (PDM), which uses a subgrid distribution of soil storage capacities to diagnose partial saturation and permits runoff before the gridbox is fully saturated; both schemes include explicit baseflow components (Sellar et al., 2019). Transient LULCC follows the LUH2 forcing: natural PFTs (nine) and agricultural PFTs (four) occupy separate land-units within each gridbox; TRIFFID allows climate-driven competition within the natural and within the agricultural units, but not between them (Burton et al., 2019; Cox, 2001). Fractional cover changes are prescribed annually, while abandonment allows regrowth according to TRIFFID phenology. JULES also includes the INFERNO fire model for disturbance, interacting with both natural and managed vegetation (Burton et al., 2019).

S1.5 Cross-model comparison of land-surface process representation

Among the four, UKESM1-0-LL (JULES with dynamic vegetation, and alternative TOPMODEL/PDM runoff) and CMCC-ESM2 (CLM4.5-BGC with sunlit/shaded canopy, multilayer soil hydrology, and optional VIC runoff) provide the richest process detail for both green- and blue-water branches (Oleson et al., 2013; Sellar et al., 2019). In CMCC-ESM2, total vegetated area is prescribed but PFT composition and LAI evolve prognostically, making it more responsive to LULCC than fully prescribed-LAI models, though less structurally dynamic than UKESM1-0-LL, where natural PFTs compete interactively within land-units. MIROC-ES2L (MATSIRO/VISIT-e) prescribes cover fractions but simulates prognostic LAI dynamics, with explicit saturated-area runoff from a simplified TOPMODEL and multiple surface/groundwater pathways, limiting LULCC feedbacks on transpiration compared to UKESM1-0-LL/CLM (Hajima et al., 2020). BCC-CSM2-MR (BCC_AVIM2.0) includes prognostic phenology and updated canopy, snow, and albedo physics, but retains CLM-style saturation/infiltration-excess runoff without explicit saturated-area diagnostics, making it structurally simpler than the TOPMODEL/PDM/VIC options above (Wu et al., 2019; Li et al., 2019). Notably, higher model complexity is a relative measure, not necessarily predictive of performance – a model with simpler ecohydrology might still match observations better in some regions because of parameter tuning or compensating processes.

S2 LUH2 land-use forcing and scenario contrasts

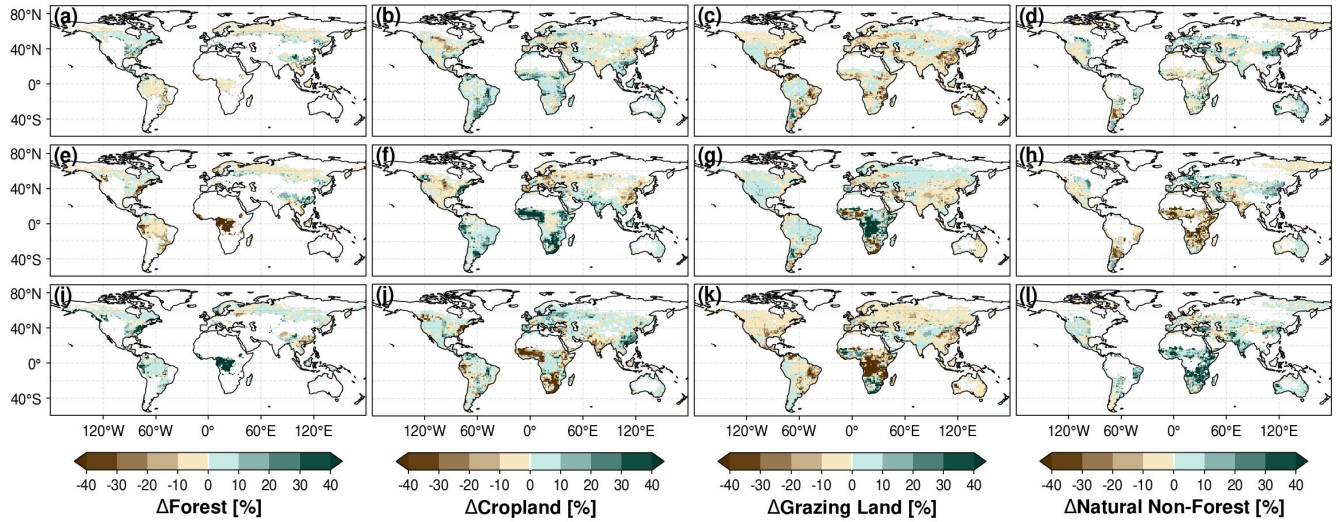

Figure S1. Late-21st-century (2070–2099) fractional changes in land-use state variables [%] under the sustainable (SSP1–2.6) and unsustainable (SSP3–7.0) pathways, based on LUH2 data. Panels **(a)**–**(d)** show changes in forest, cropland, grazing land (pasture and rangeland), and natural non-forest vegetation for SSP1–2.6 relative to the historical period (1985–2014). Panels **(e)**–**(h)** show the corresponding changes for SSP3–7.0 relative to the historical period. Panels **(i)**–**(l)** show differences between the two future pathways (SSP1–2.6 minus SSP3–7.0), highlighting the land-use signal associated with sustainable LULCC. Grid cells with no change are masked (white).

S3 Modelled land-cover responses to LUH2 forcing

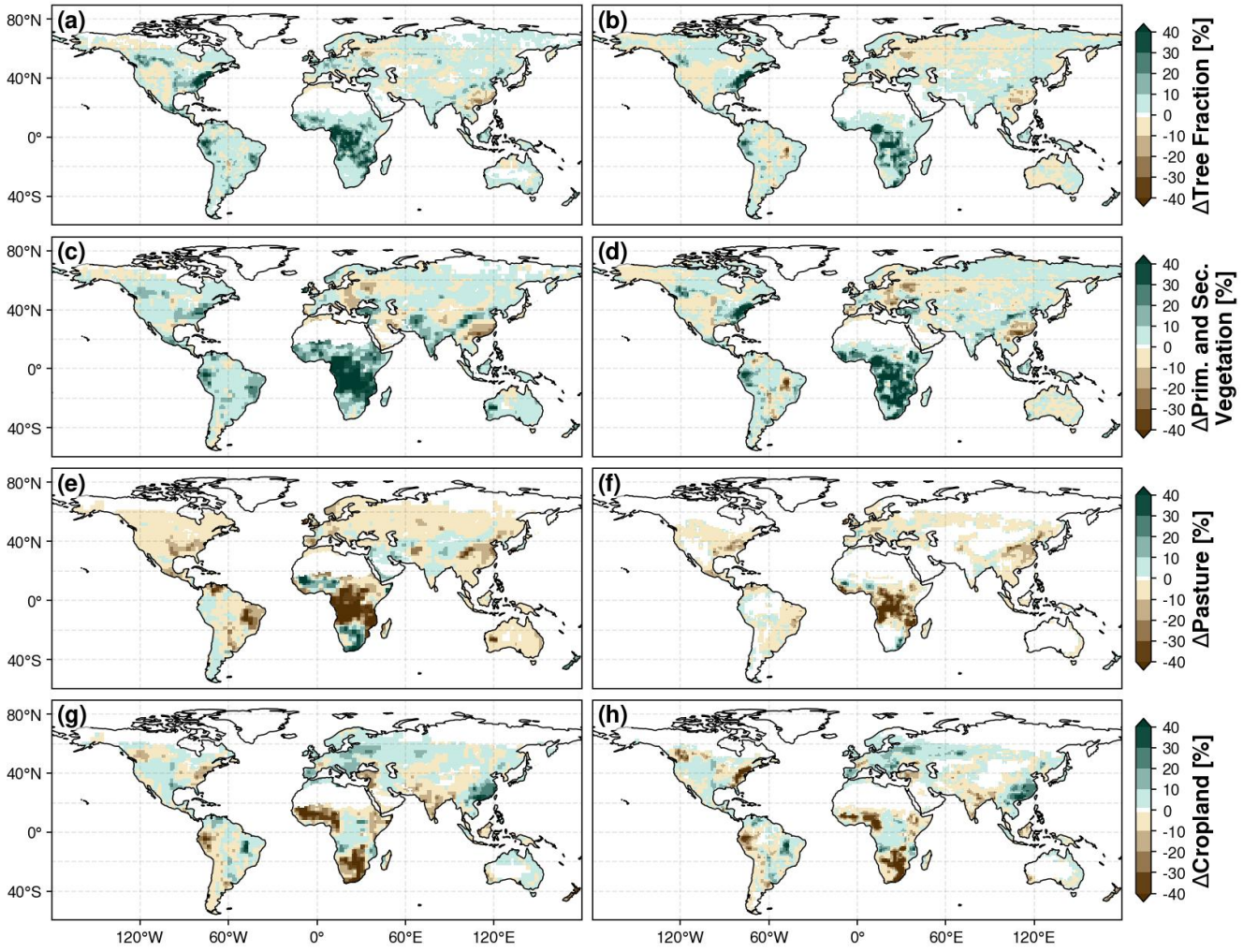

Figure S2. Late-21st-century (2070–2099) differences in simulated land-cover fractions [%] associated with sustainable LULCC. Panels **(a)** and **(b)** show differences in tree cover fraction for CMCC-ESM2 and UKESM1-0-LL (ensemble mean). Panels **(c)** and **(d)** show differences in primary and secondary vegetation fraction for MIROC-ES2L and UKESM1-0-LL (r4i1p1f2). Panels **(e)** and **(f)** show differences in pasture fraction, and panels **(g)** and **(h)** show differences in cropland fraction for the same two models. Model output for BCC-CSM2-MR is not available for these variables. Grid cells with no change are masked (white).

S4 Historical attribution: sensitivity of ΔBGWS to vegetation changes and BGWS response

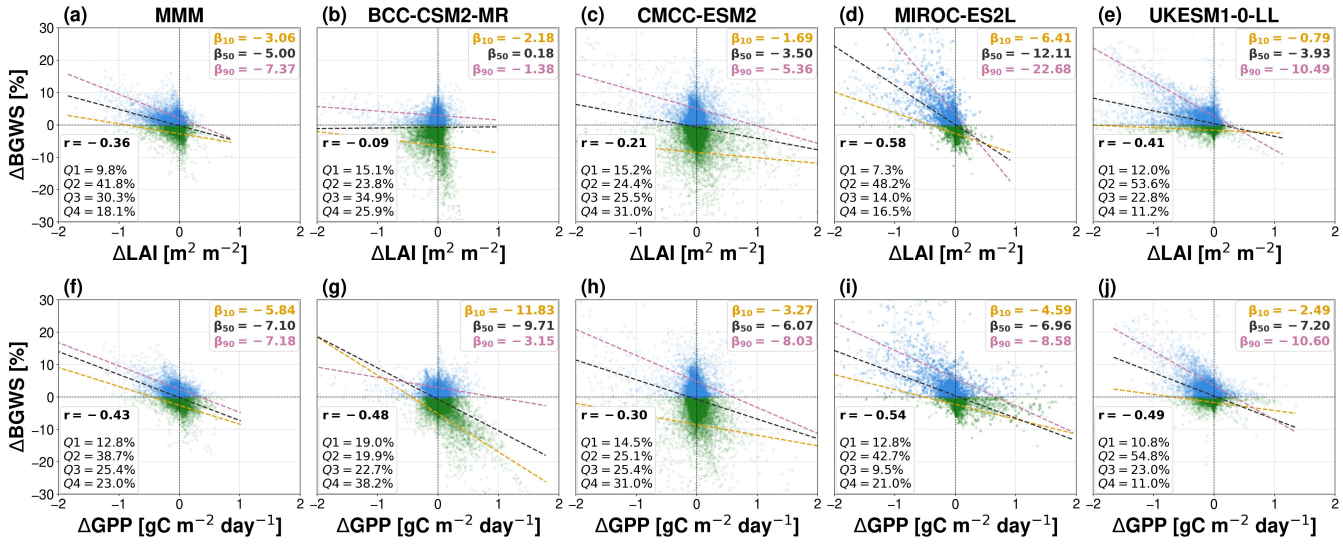

Figure S3. Quantile regression between historical-period (1985–2014) differences in BGWS [%] and differences in LAI [$\mathrm{m^2\,m^{-2}}$] and GPP [$\mathrm{gC\,m^{-2}\,day^{-1}}$] associated with land-use change, defined as the difference between the historical and hist-noLu simulations. Panels **(a)**–**(e)** show quantile regressions between ΔBGWS and ΔLAI for the four-model MMM (**a**) and the individual models BCC-CSM2-MR (**b**), CMCC-ESM2 (**c**), MIROC-ES2L (**d**), and UKESM1-0-LL (**e**). Panels **(f)**–**(j)** show the corresponding relationships between ΔBGWS and ΔGPP. Quantile regressions are shown for the 10th, 50th, and 90th percentiles. The slope coefficient β_1 denotes the change in ΔBGWS per unit LAI or GPP difference at each quantile. Pearson's correlation coefficient (r) is reported for the full sample. The fraction of grid cells in each quadrant is indicated: Q1, positive ΔBGWS and LAI/GPP difference; Q2, positive ΔBGWS and negative LAI/GPP difference; Q3, negative ΔBGWS and LAI/GPP difference; and Q4, negative ΔBGWS and positive LAI/GPP difference.

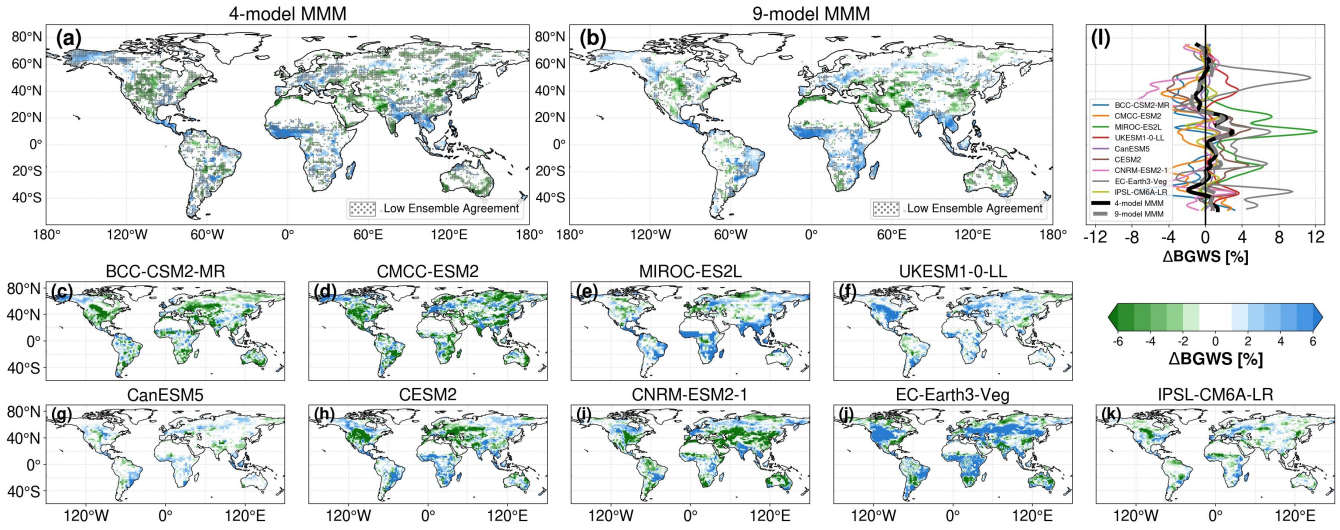

Figure S4. ΔBGWS difference between historical minus hist-noLu simulation (1985-2014) [%]. Spatial distribution of the difference in BGWS between the historical and hist-noLu simulations over the period 1985-2014. **(a)**, MMM of the four-model core ensemble, **(b)**, MMM of the nine-model ensemble, the individual models of the core ensemble: BCC-CSM2-MR ((**c**)), CMCC-ESM2 ((**d**)), MIROC-ES2L **(e)**, UKESM1-0-LL **(f)**; and the additional models of the nine-model ensemble: CanESM5 **(g)**, CESM2 **(h)**, CNRM-ESM2-1 **(i)**, EC-Earth3-Veg **(j)**, and IPSL-CM6A-LR **(k)**. Small differences between -1 and 1 are not coloured. Low ensemble agreement (fewer than three out of four models agree on the sign) is indicated. Panel **(l)** shows the zonal ΔBGWS means.

S5 Seasonal and latitudinal structure of LULCC-driven differences

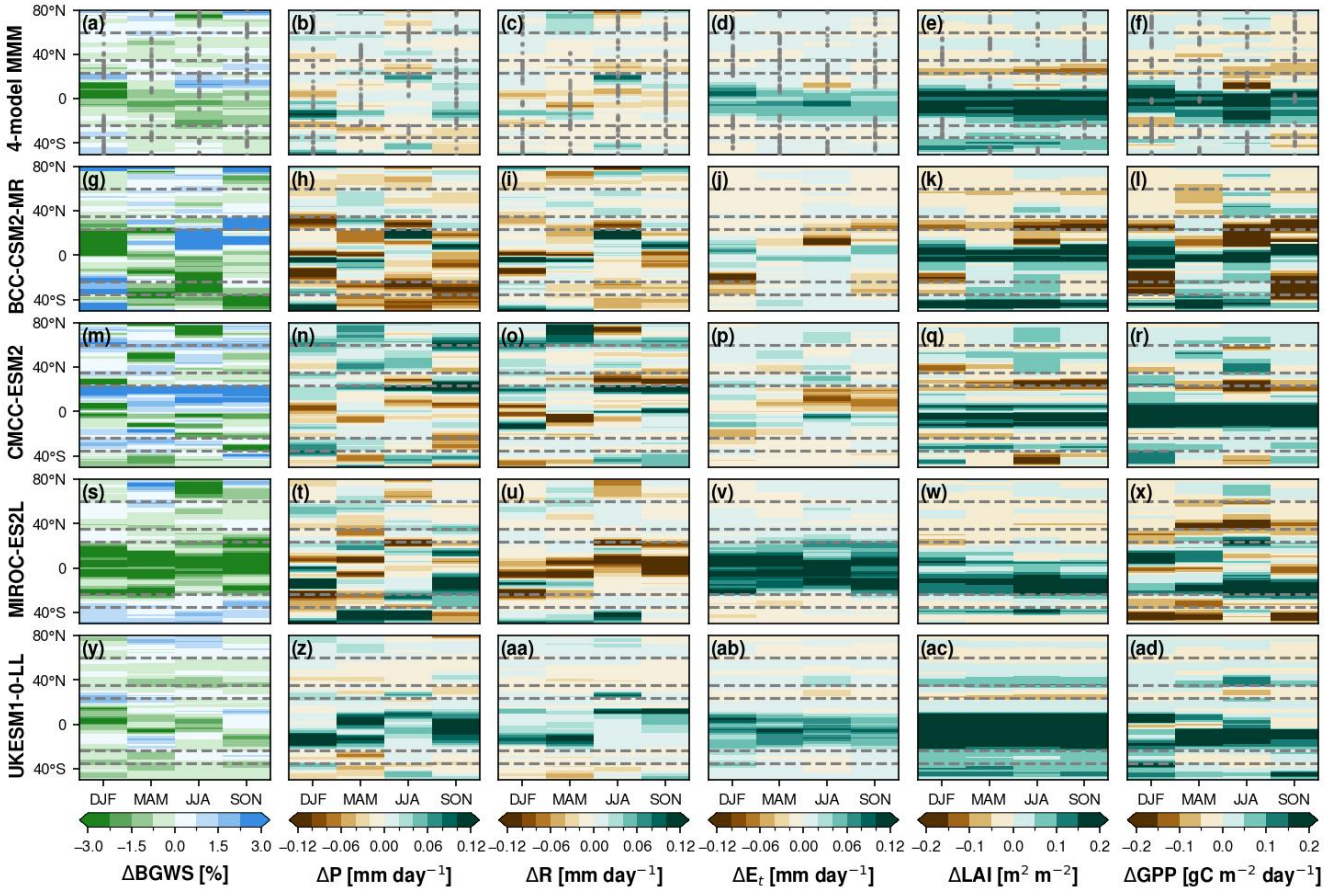

Figure S5. Seasonal and latitudinal structure of late-21st-century (2070–2099) ecohydrological responses associated with sustainable LULCC over land grid cells for the four-model MMM and the individual ESMs. Hovmöller diagrams show seasonal–latitudinal differences in December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON). Columns show ΔBGWS, ΔP, ΔR, ΔE_t, ΔLAI, and ΔGPP. Horizontal dashed lines indicate latitude bands: tropics (23.5°S–23.5°N), subtropics (23.5–35° S/N), mid-latitudes (35–60°N; 35–50°S in our domain), and high latitudes ($\geq$60°N). Grey markers indicate season–latitude cells (zonal means) where fewer than three out of four ESMs agree on the sign of the response.

S6 Spatial patterns of key ecohydrological responses

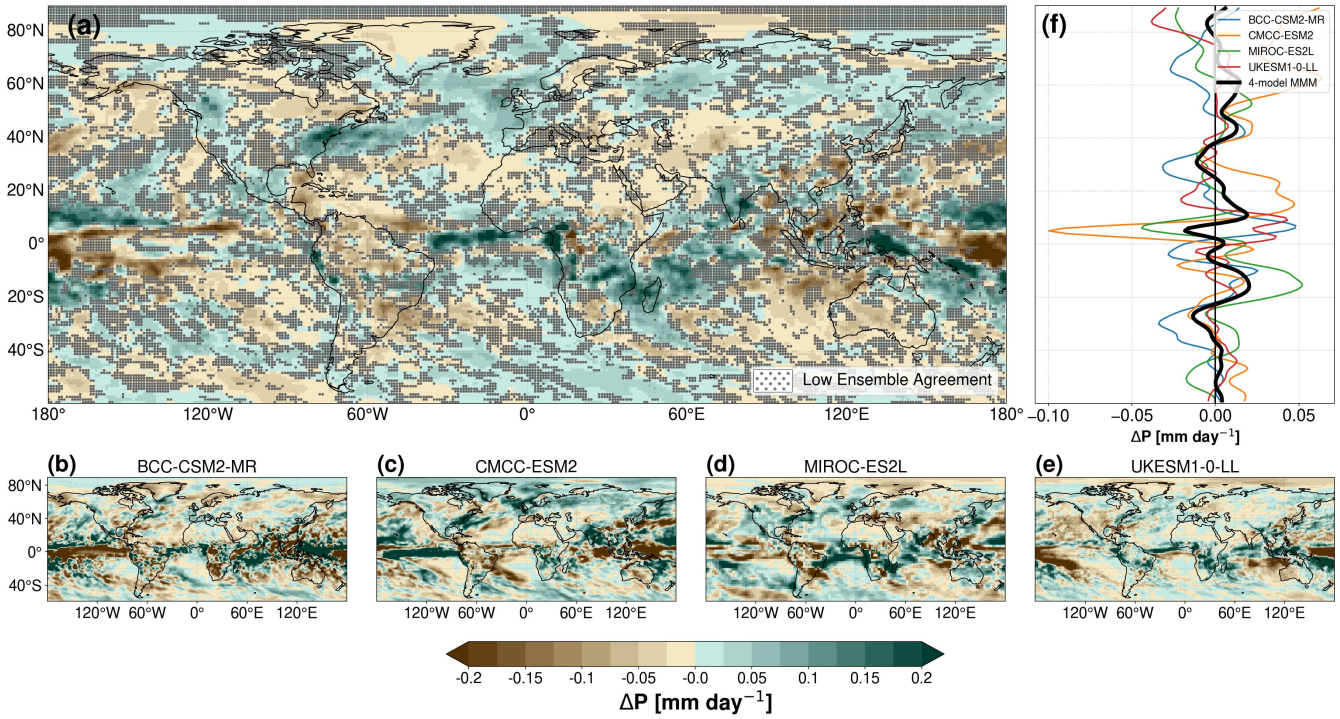

Figure S6. Late-21st-century (2070–2099) difference in precipitation associated with sustainable LULCC, ΔP [mm day^{-1}]. Panels **(a)**–**(e)** show the spatial distribution for the four-model MMM (**a**) and the individual models BCC-CSM2-MR (**b**), CMCC-ESM2 (**c**), MIROC-ES2L (**d**), and UKESM1-0-LL (**e**). In the MMM, regions with low ensemble agreement (fewer than three out of four models agreeing on the sign of ΔP) are indicated. Panel **(f)** shows zonal-mean ΔP.

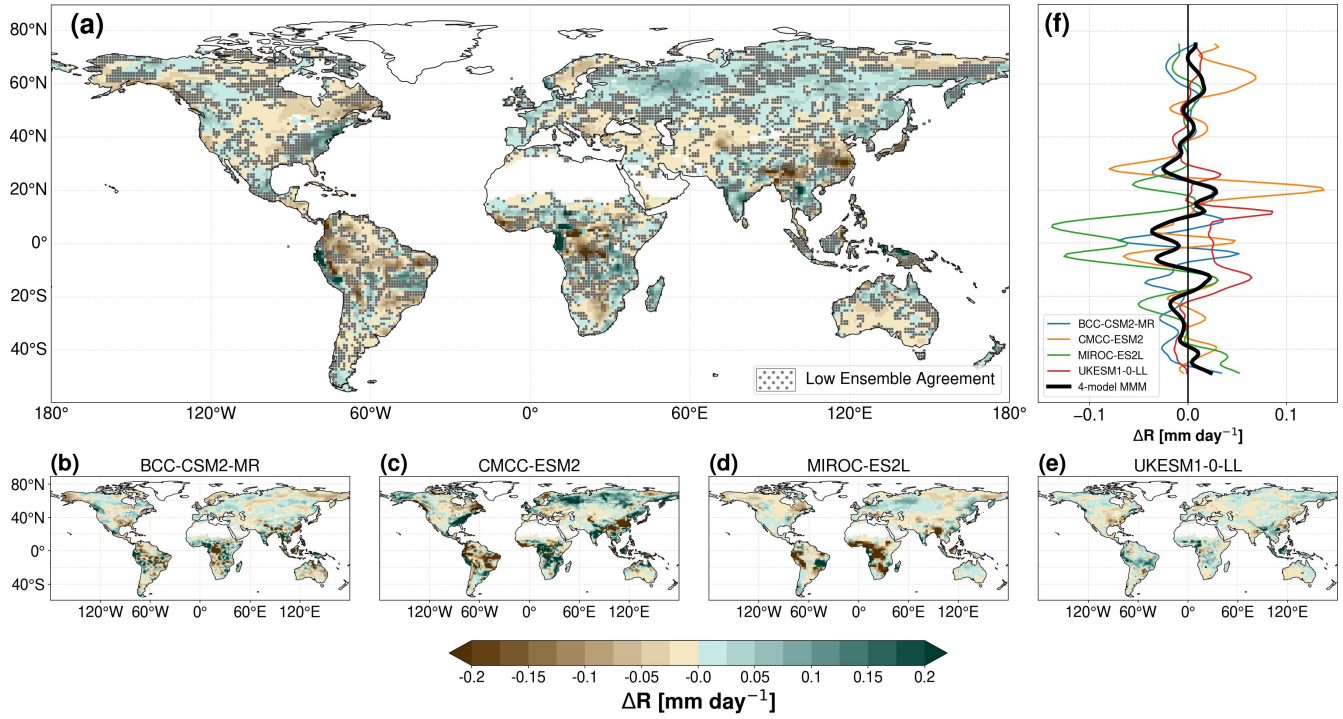

Figure S7. Same as Fig. S6, but for ΔR [mm day^{-1}].

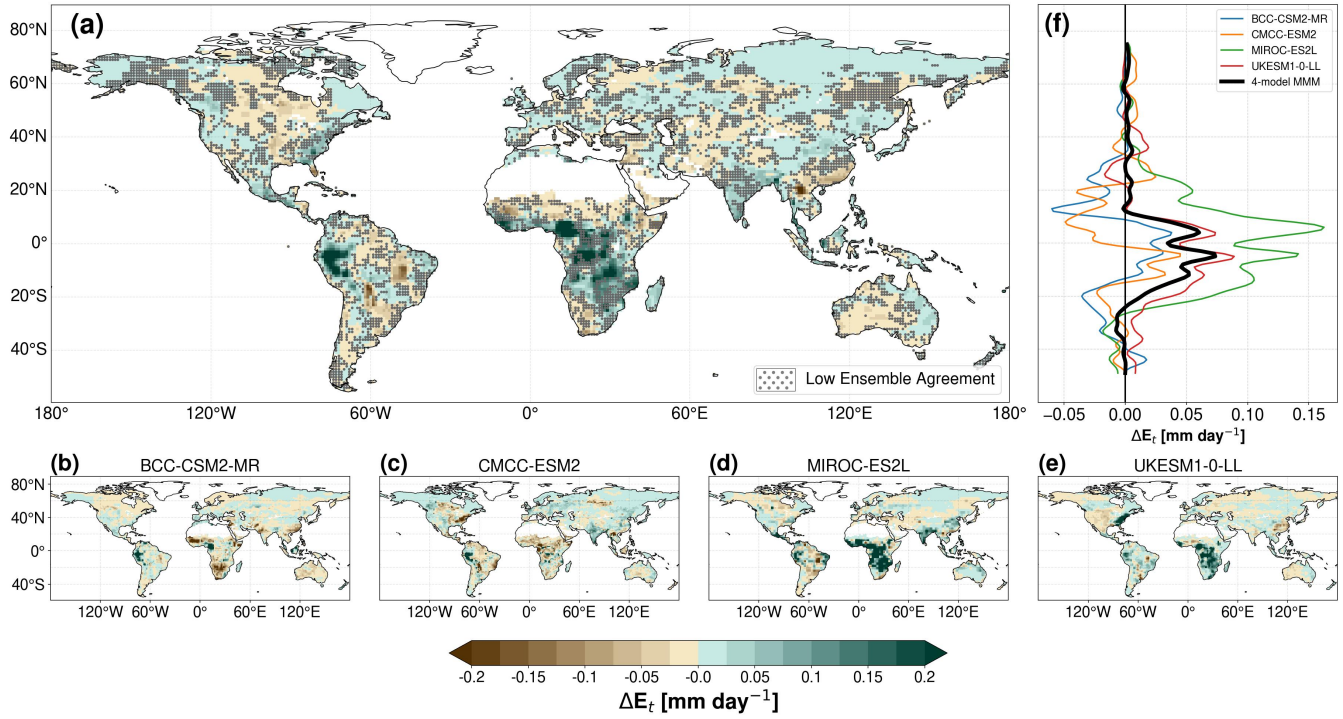

Figure S8. Same as Fig. S6, but for ΔE_t [mm day^{-1}].

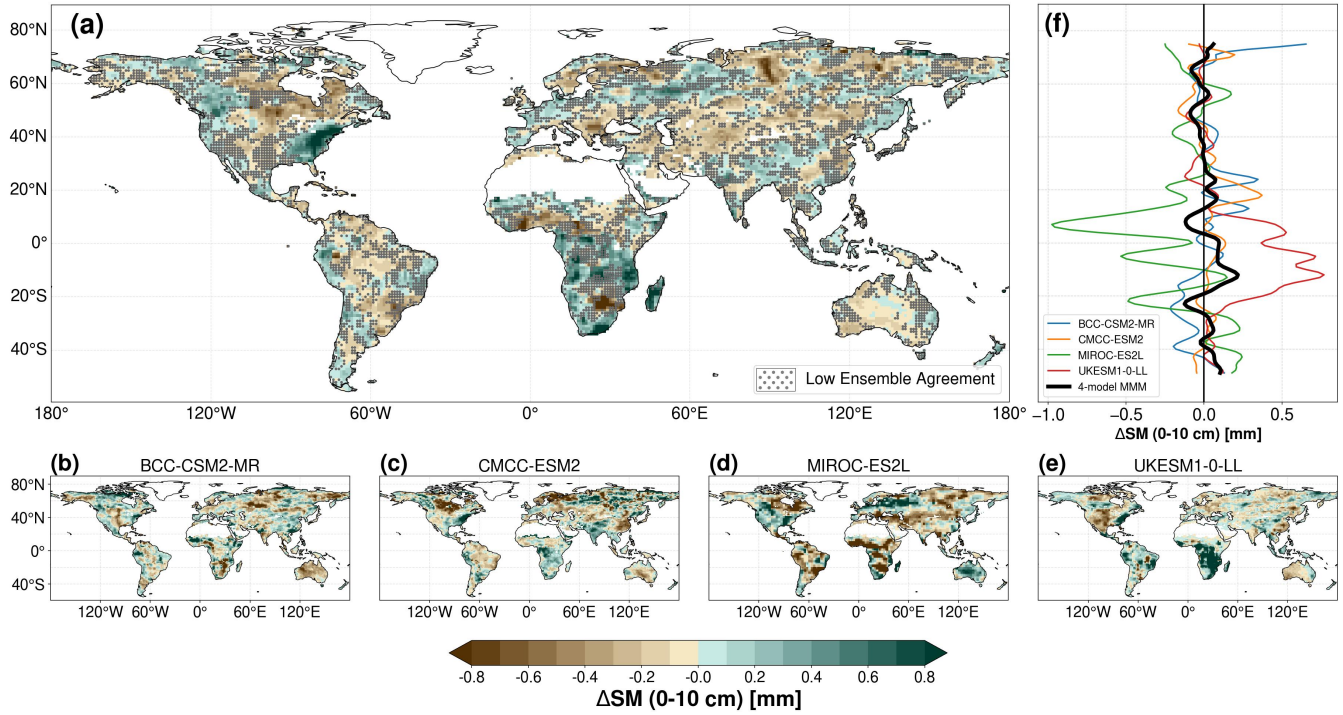

Figure S9. Same as Fig. S6, but for surface soil moisture in the upper 0–10 cm, ΔSM [mm].

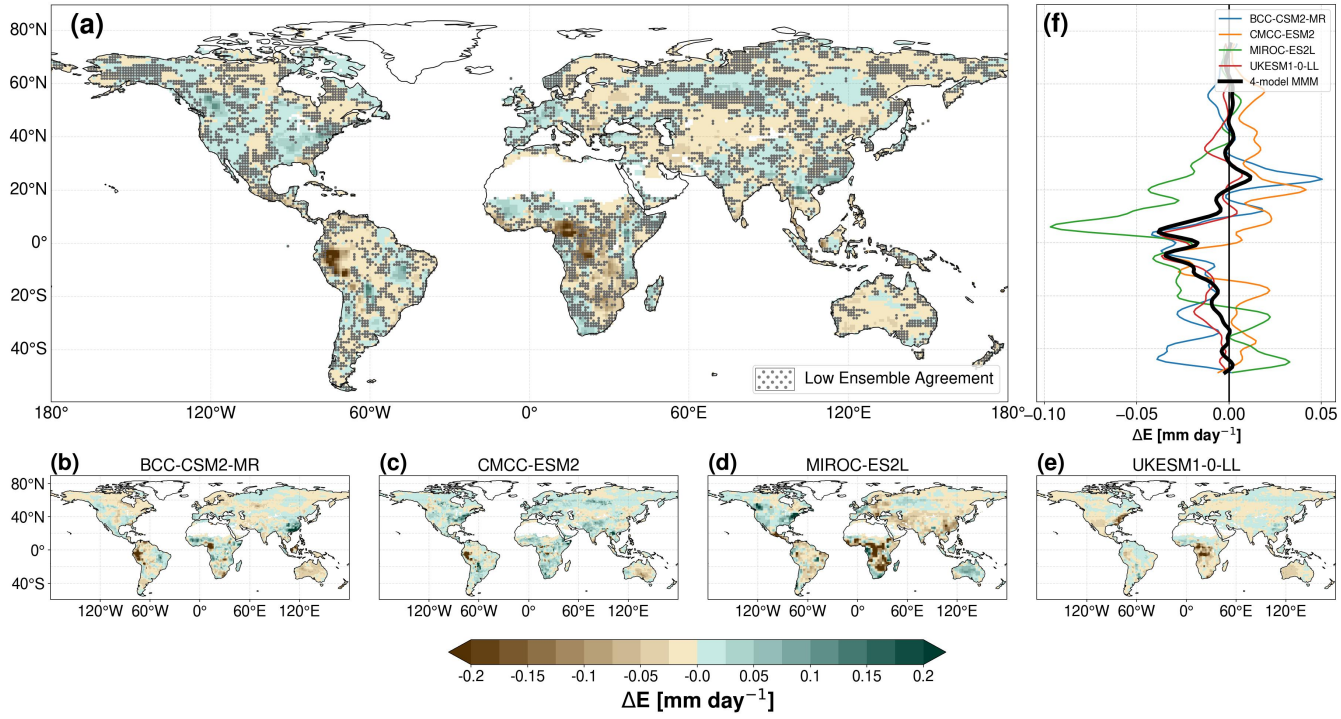

Figure S10. Same as Fig. S6, but for ΔE [mm day^{-1}].

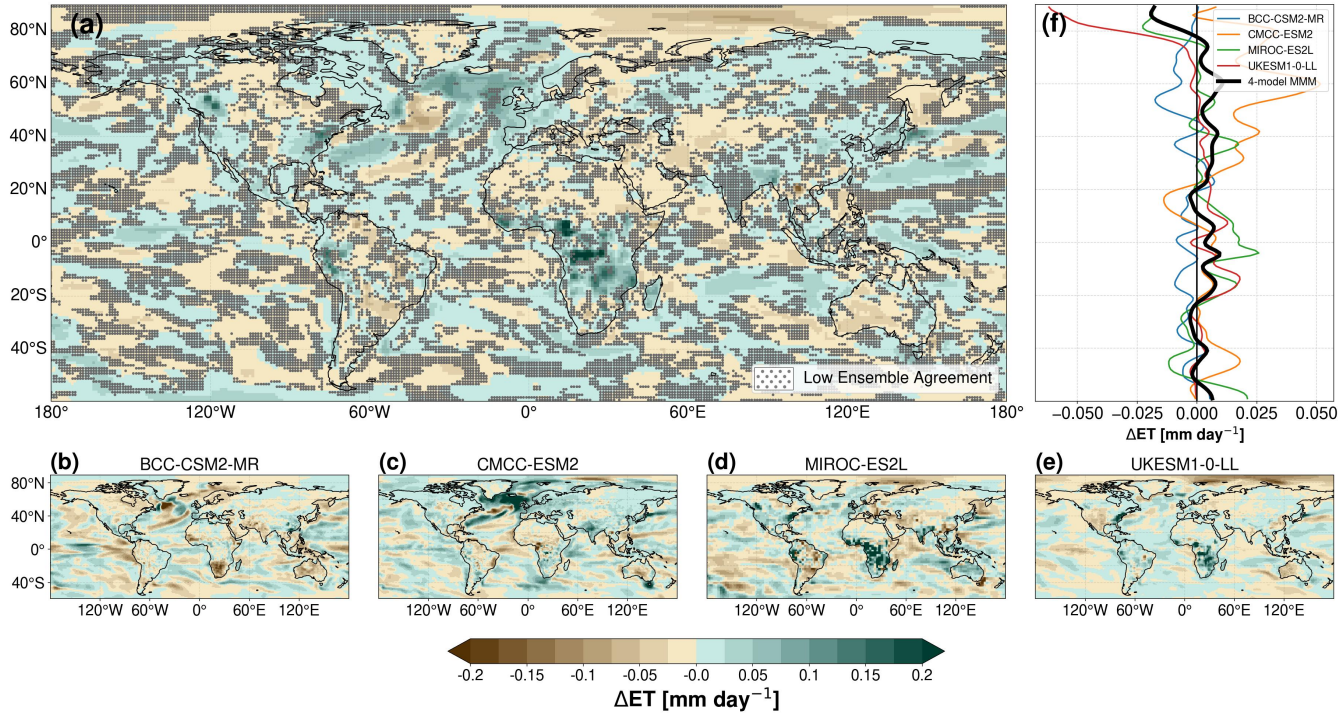

Figure S11. Same as Fig. S6, but for ΔET [$\mathrm{mm\,day^{-1}}$].

S7 Regional detail: Eastern United States

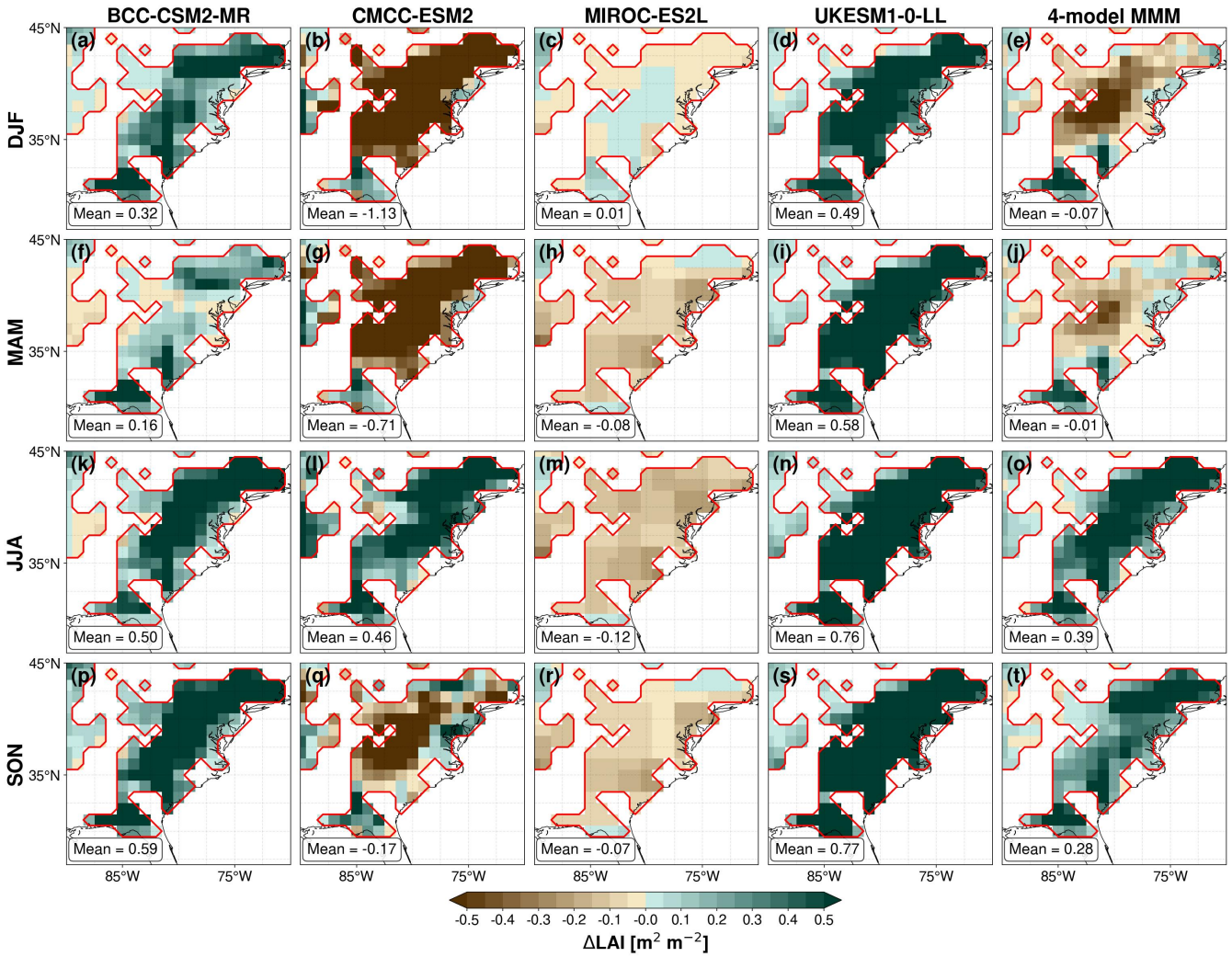

Figure S12. Seasonal late-21st-century (2070–2099) differences in ΔLAI [$\mathrm{m}^2\,\mathrm{m}^{-2}$] associated with sustainable LULCC over the eastern US (30°–45° N, 90°–70° W). Seasons shown are December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON). Panels show the four-model MMM and the individual models. Only grid cells with higher forest and lower cropland fractions in LUH2 under the sustainable pathway are shown; the boundary of this forest-for-cropland mask is outlined in red. For each season, model, and MMM, the area-weighted mean over the masked grid cells is reported.

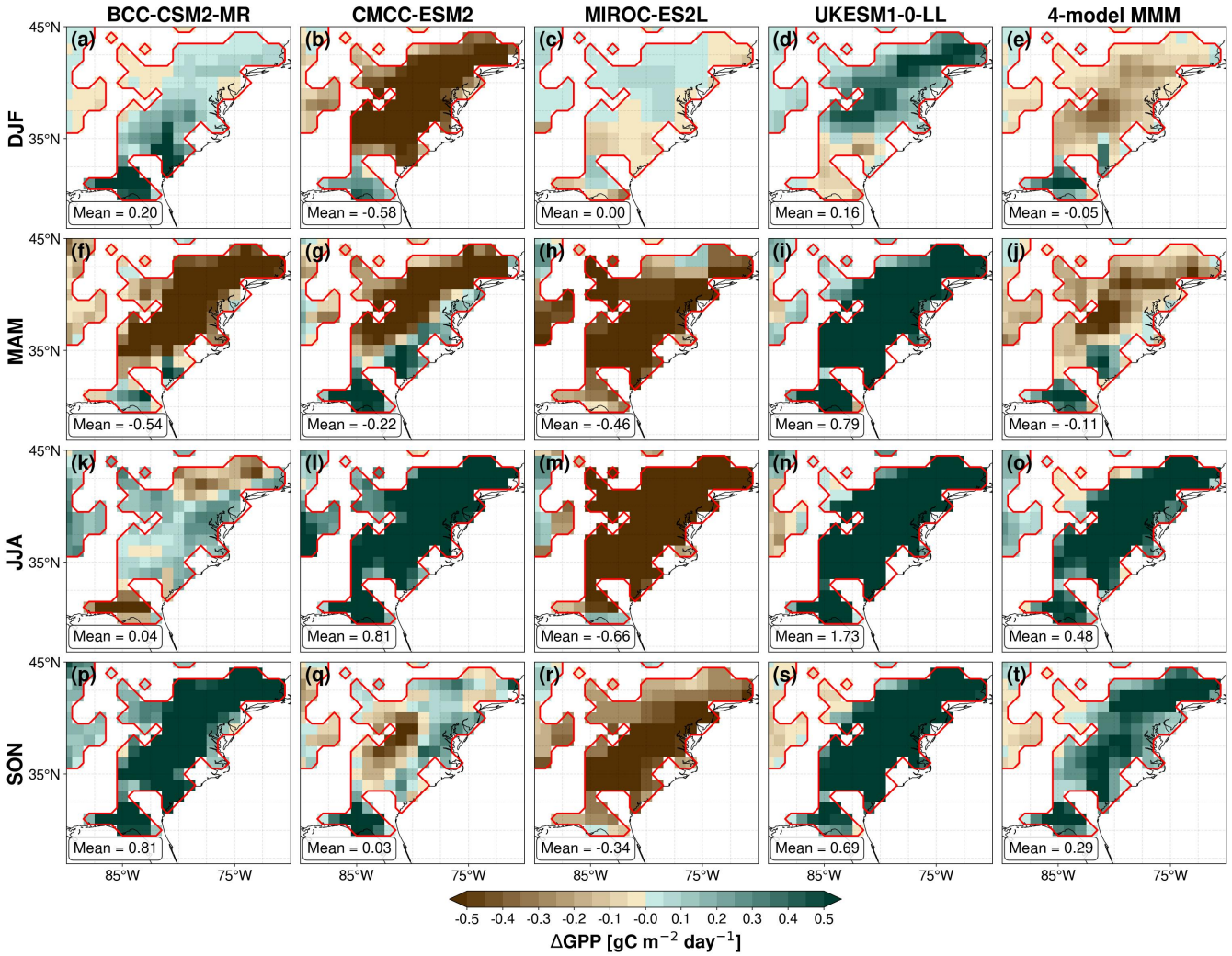

Figure S13. Same as Fig. S12, but for ΔGPP [$\mathrm{gC\,m^{-2}\,day^{-1}}$].

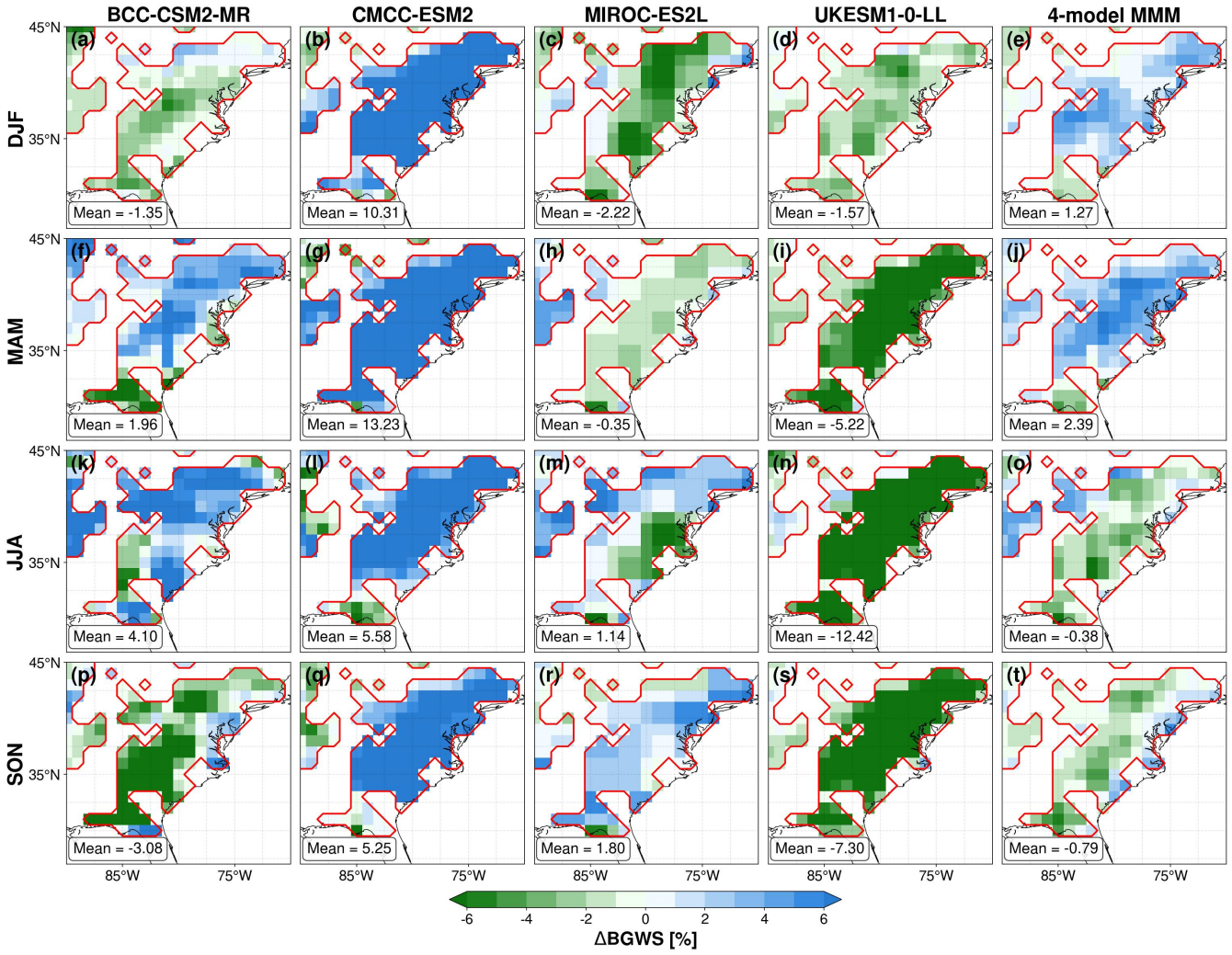

Figure S14. Same as Fig. S12, but for ΔBGWS [%].

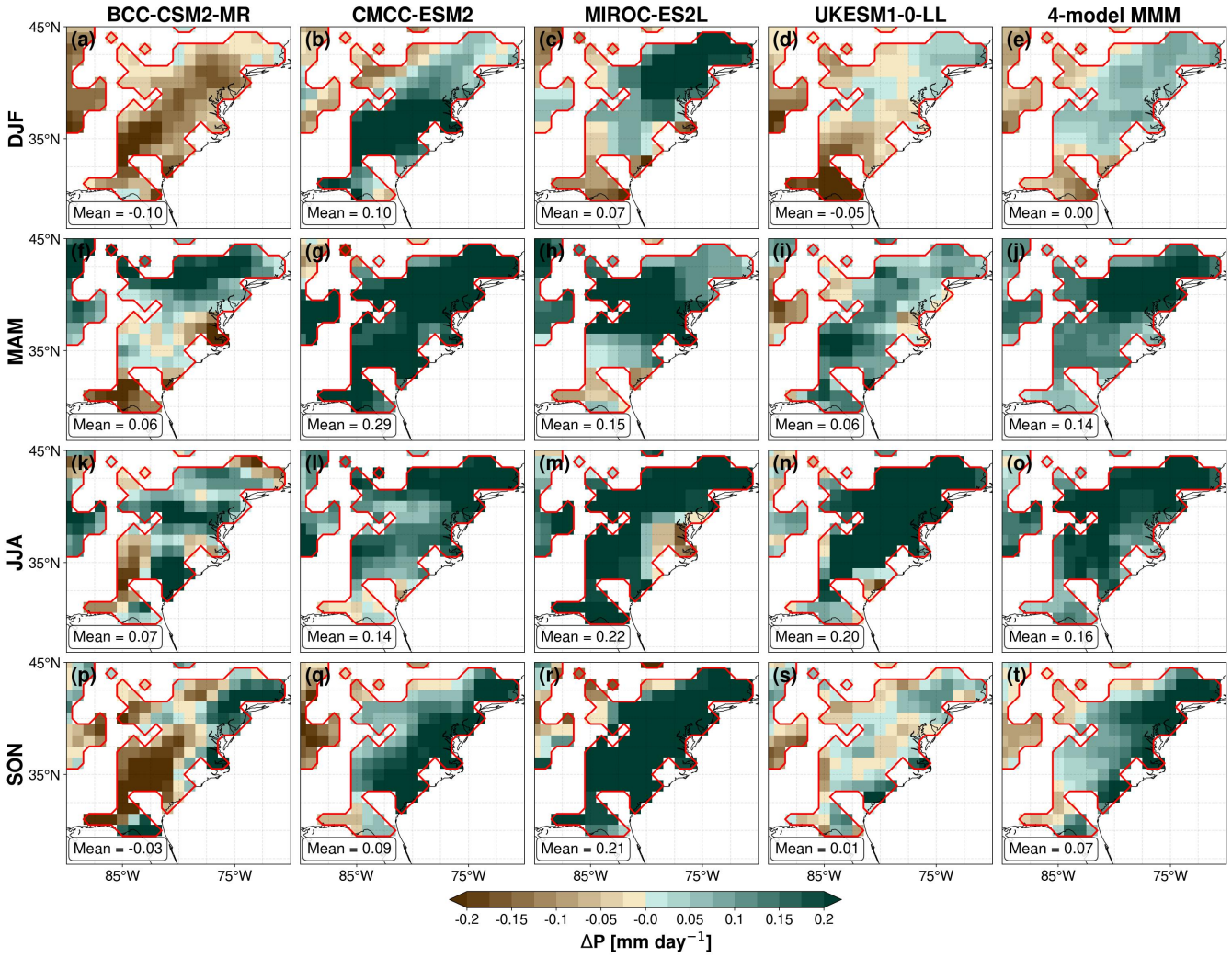

Figure S15. Same as Fig. S12, but for ΔP [mm day^{-1}].

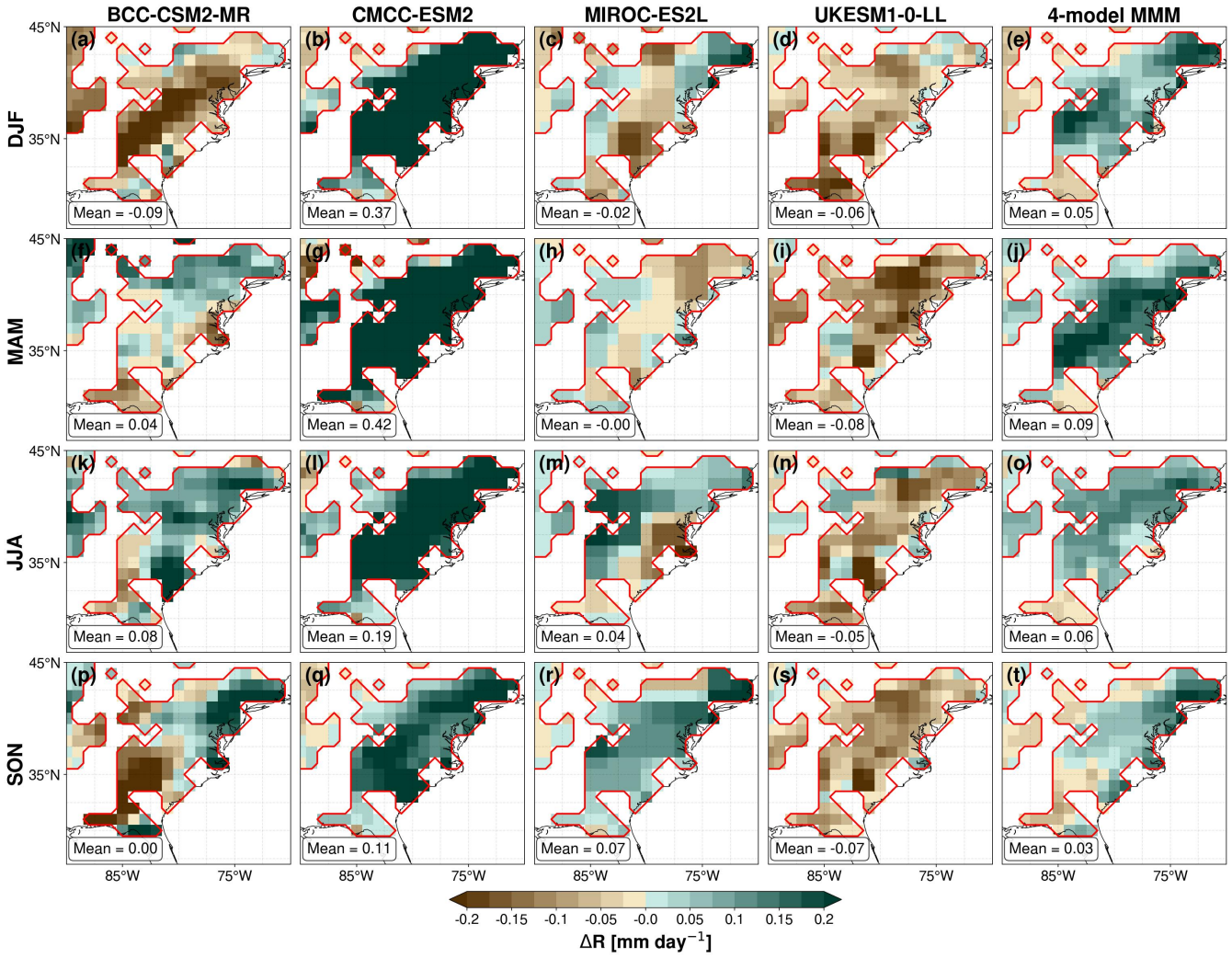

Figure S16. Same as Fig. S12, but for ΔR [mm day^{-1}].

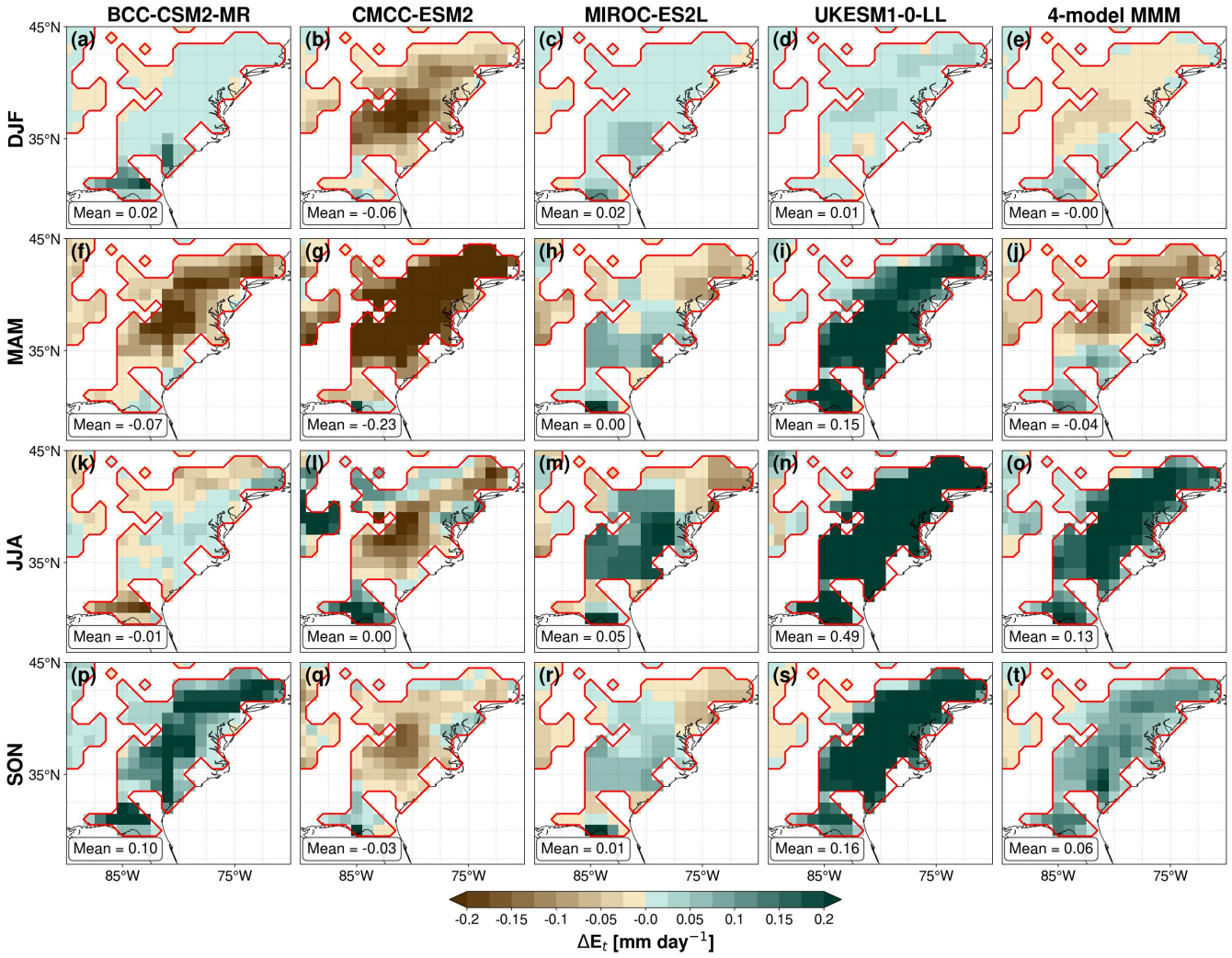

Figure S17. Same as Fig. S12, but for ΔE_t [mm day^{-1}].

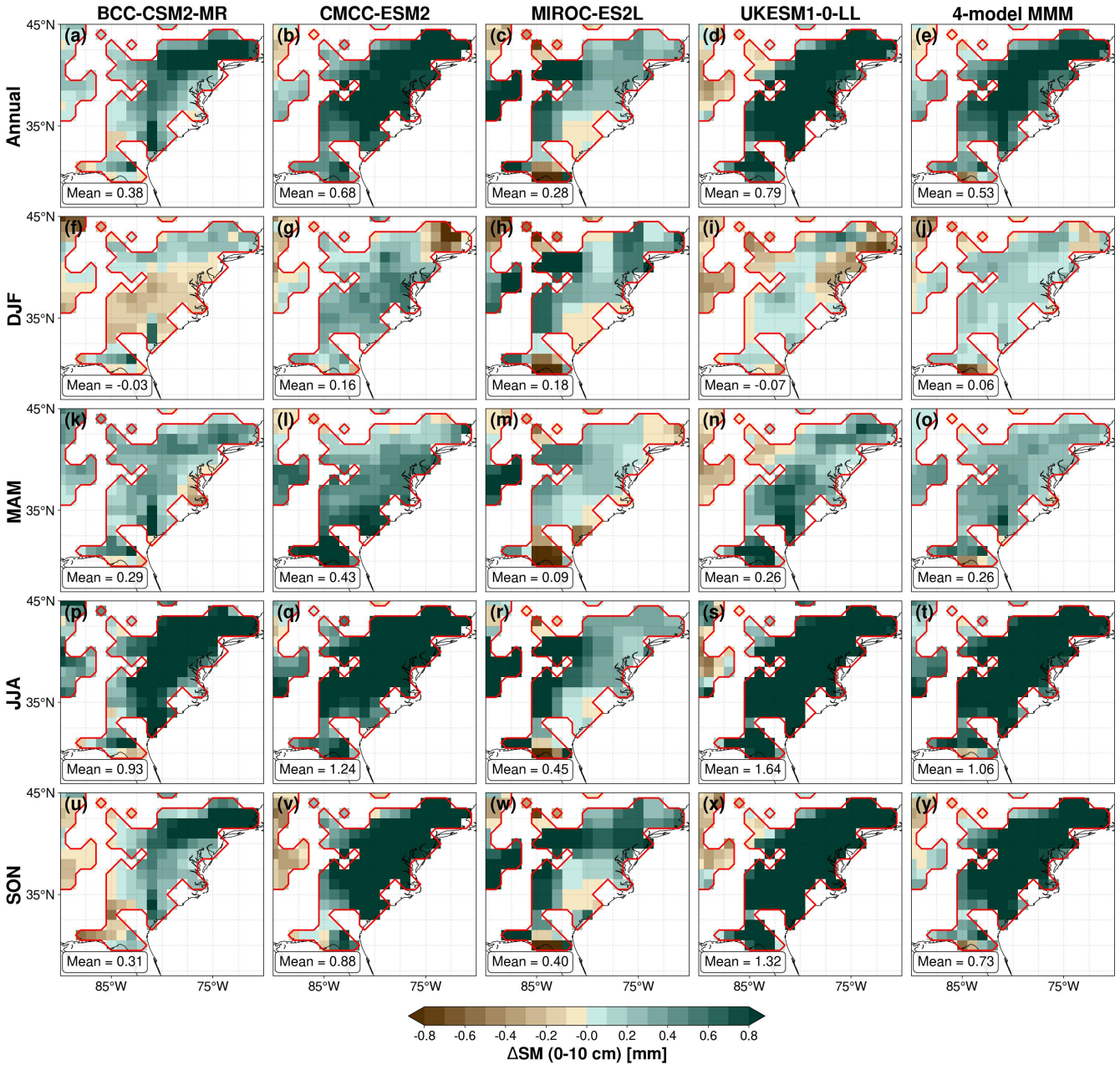

Figure S18. Annual and seasonal late-21st-century (2070–2099) differences in ΔSM (surface soil moisture, mm) associated with sustainable LULCC over the eastern US (30°–45° N, 90°–70° W). Seasonal panels show December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON). Panels show the four-model MMM and the individual models. Only grid cells with higher forest and lower cropland fractions in LUH2 under the sustainable pathway are shown; the boundary of this forest-for-cropland mask is outlined in red. For each time aggregation (annual/seasonal), model, and MMM, the area-weighted mean over the masked grid cells is reported.

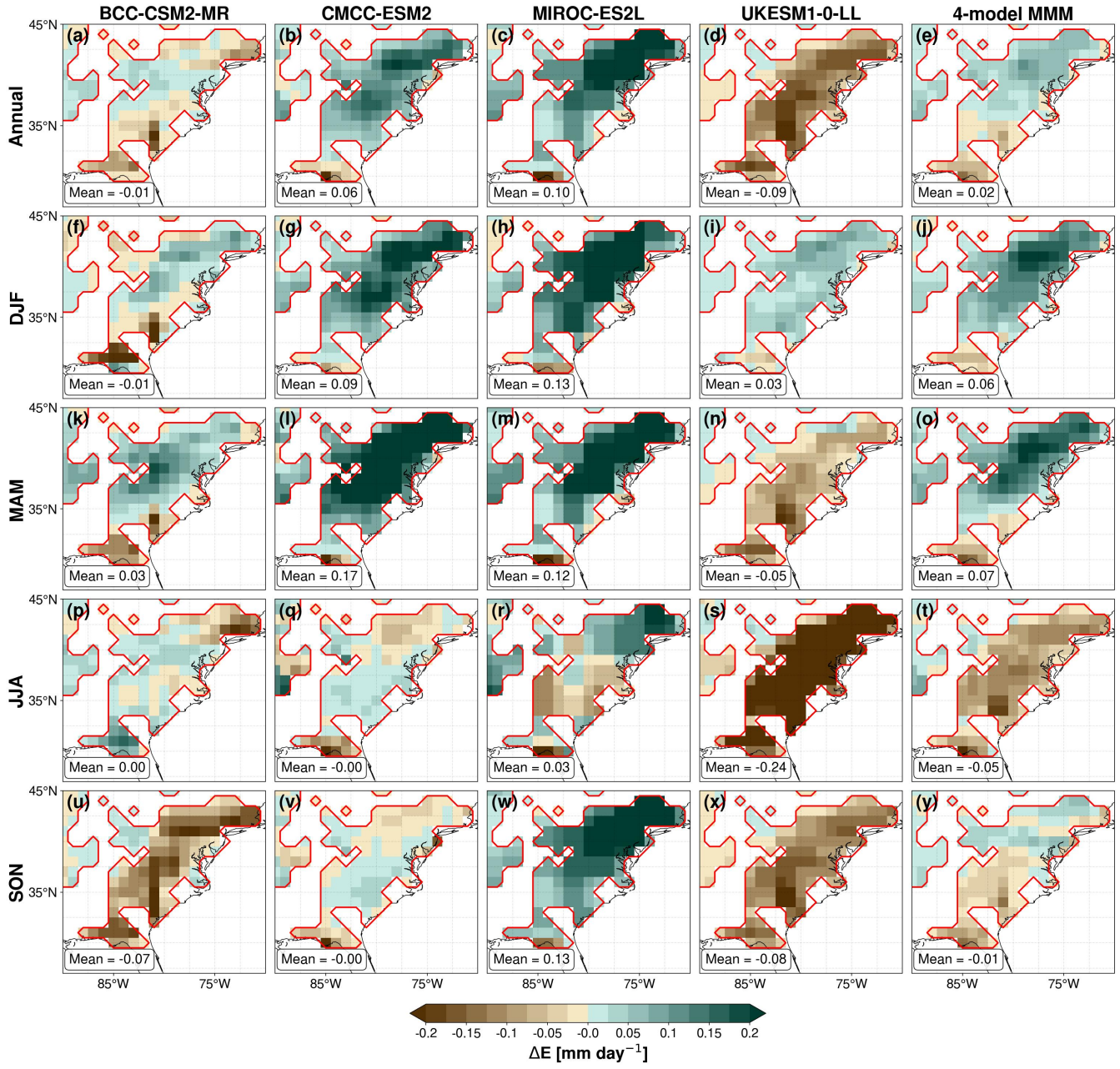

Figure S19. Same as Fig. S18, but for ΔE [mm day^{-1}].

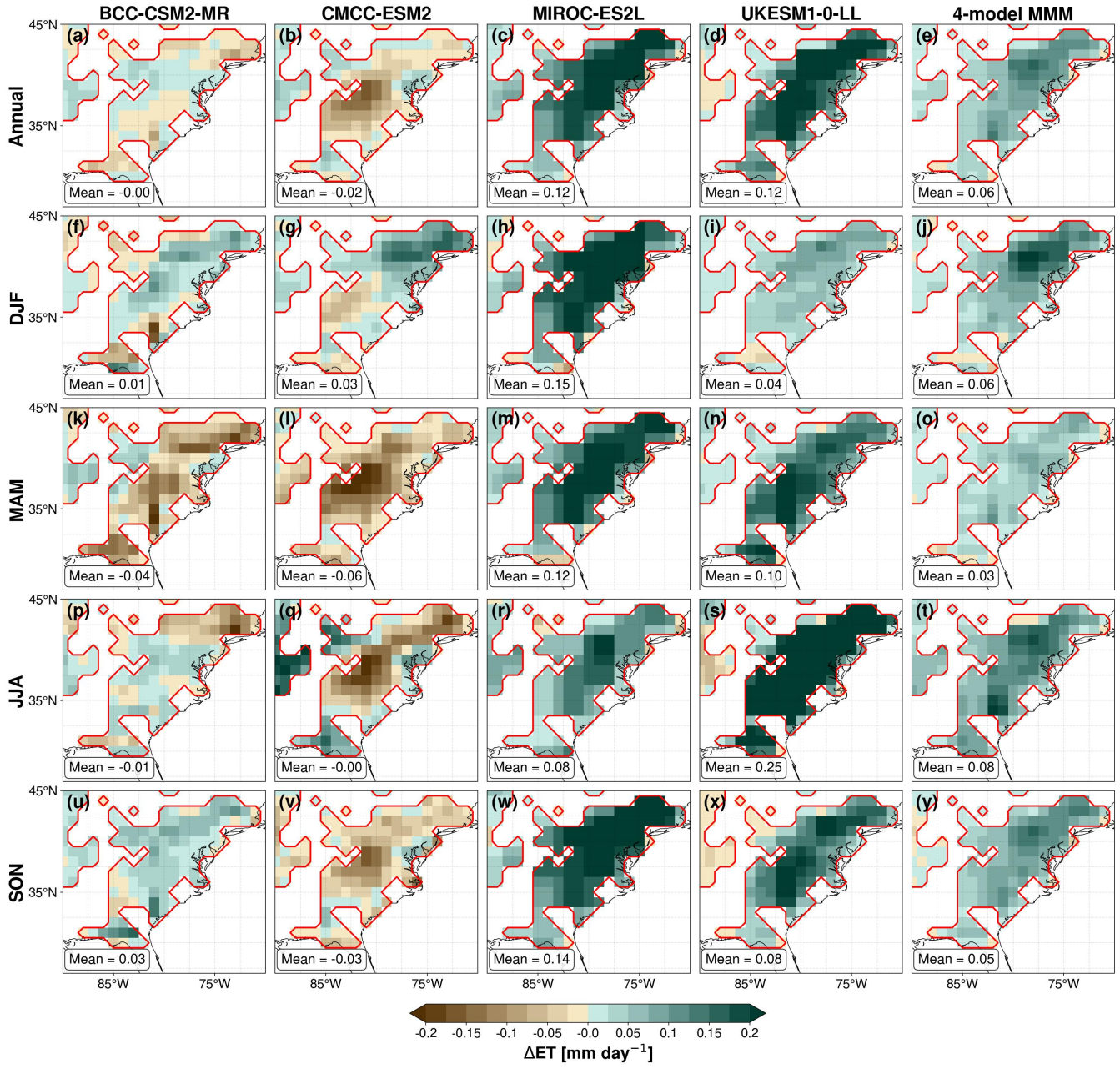

Figure S20. Same as Fig. S18, but for ΔET [mm day^{-1}].

S8 Vegetation diagnostics and carbon allocation responses

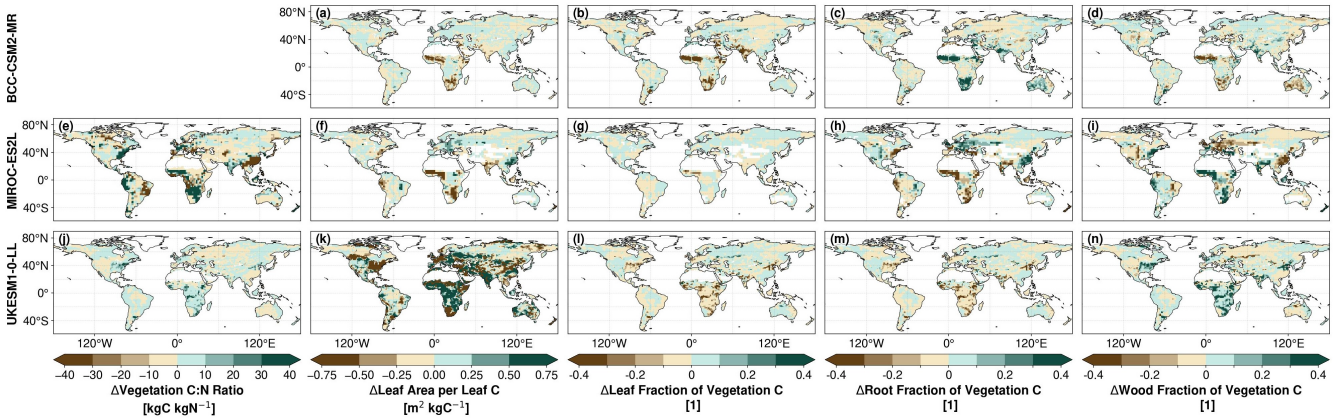

Figure S21. Late-21st-century (2070–2099) differences in vegetation nutrient status and carbon allocation diagnostics associated with sustainable LULCC. Panels show the individual models (rows) for the diagnostics available (columns). Variables shown are LAI/c_{Leaf} [m^2 kgC^{-1}], C:N $= c_{\mathrm{Veg}}/n_{\mathrm{Veg}}$ [–], and the carbon allocation fractions $f_{\mathrm{leaf}} = c_{\mathrm{Leaf}}/c_{\mathrm{Veg}}$, $f_{\mathrm{root}} = c_{\mathrm{Root}}/c_{\mathrm{Veg}}$, and $f_{\mathrm{wood}} = c_{\mathrm{Wood}}/c_{\mathrm{Veg}}$ [–]. In BCC-CSM2-MR, vegetation nitrogen is not archived, so C:N is not available. CMCC-ESM2 is excluded because these diagnostics are not available.

S9 Supplementary tables

Table S1. Overview of the extended CMIP6 ESMs (historical period only) used in this study, including their land-surface component, horizontal resolution, ensemble members, and primary references.

Model	Ensemble member(s)	Land Surface Model	Resolution (lat×lon)	Reference
CanESM5	r1–r10i1p1f1	CLASS-CTEM	∼2.8°× ∼2.8° (T63, Gaussian)	Swart et al., 2019
CESM2	r1–r3i1p1f1	CLM5.0	0.9°× 1.25°	Danabasoglu et al., 2020
CNRM-ESM2-1	r1–r3i1p1f2	ISBA-CTRIP	∼1.4°× ∼1.4° (T127, Gaussian)	Séférian et al., 2019
EC-Earth3-Veg	r1i1p1f1	LPJ-GUESS + HTESSEL	∼0.7°× ∼0.7° (T255, Gaussian)	Döscher et al., 2022
IPSL-CM6A-LR	r1–r4i1p1f1	ORCHIDEE	1.3°× 2.5°	Boucher et al., 2020

Table S2. Spatial agreement metrics between modelled and LUH2 land-use variables for the eastern US region. Pearson correlation coefficient (r), sign agreement (SA, in %), and cosine similarity (CS, in %) are shown. For UKESM1-0-LL, Forest results are based on the ensemble mean across r1–r4i1p1f2, while Primary and Secondary Vegetation, Cropland, and Pasture results are available only for the r4i1p1f2 member.

Model	Variable Pair	r	SA (%)	CS (%)
CMCC-ESM2_r1i1p1f1	Forest vs Forest	0.832	81.94	89.24
UKESM1-0-LL	Forest vs Forest	0.781	79.82	86.19
UKESM1-0-LL	Forest vs Primary and Secondary Vegetation	0.834	81.19	89.21
UKESM1-0-LL	Cropland vs Cropland	0.827	78.85	86.39
UKESM1-0-LL	Grazing Land vs Pasture	0.820	100.00	90.48
MIROC-ES2L	Forest vs Primary and Secondary Vegetation	0.586	79.73	75.10
MIROC-ES2L	Cropland vs Cropland	0.603	75.76	70.13
MIROC-ES2L	Grazing Land vs Pasture	0.789	99.13	88.90