



Supplement of

Forest and bioenergy expansion amplifies climate warming by accelerating regional cloud loss

Nanjian Liu et al.

Correspondence to: Zhixin Hao (haozx@igsnr.ac.cn)

The copyright of individual parts of the supplement might differ from the article licence.

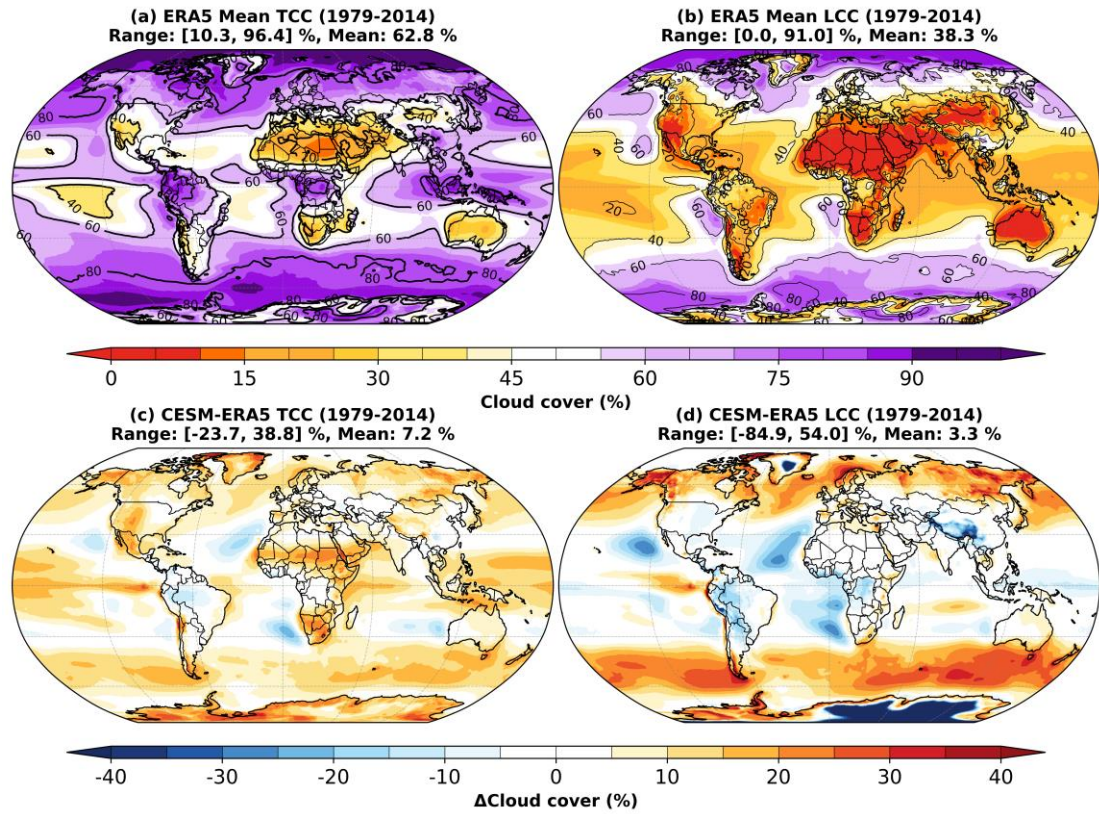


Figure S1. Validation of total cloud cover (TCC) and low-level cloud cover (LCC) by the Community Earth System Model (CESM) against ERA5 (ECMWF Reanalysis v5) during the historical period (1979-2014). The global average is obtained through area-weighted statistics, and the same applies as below.

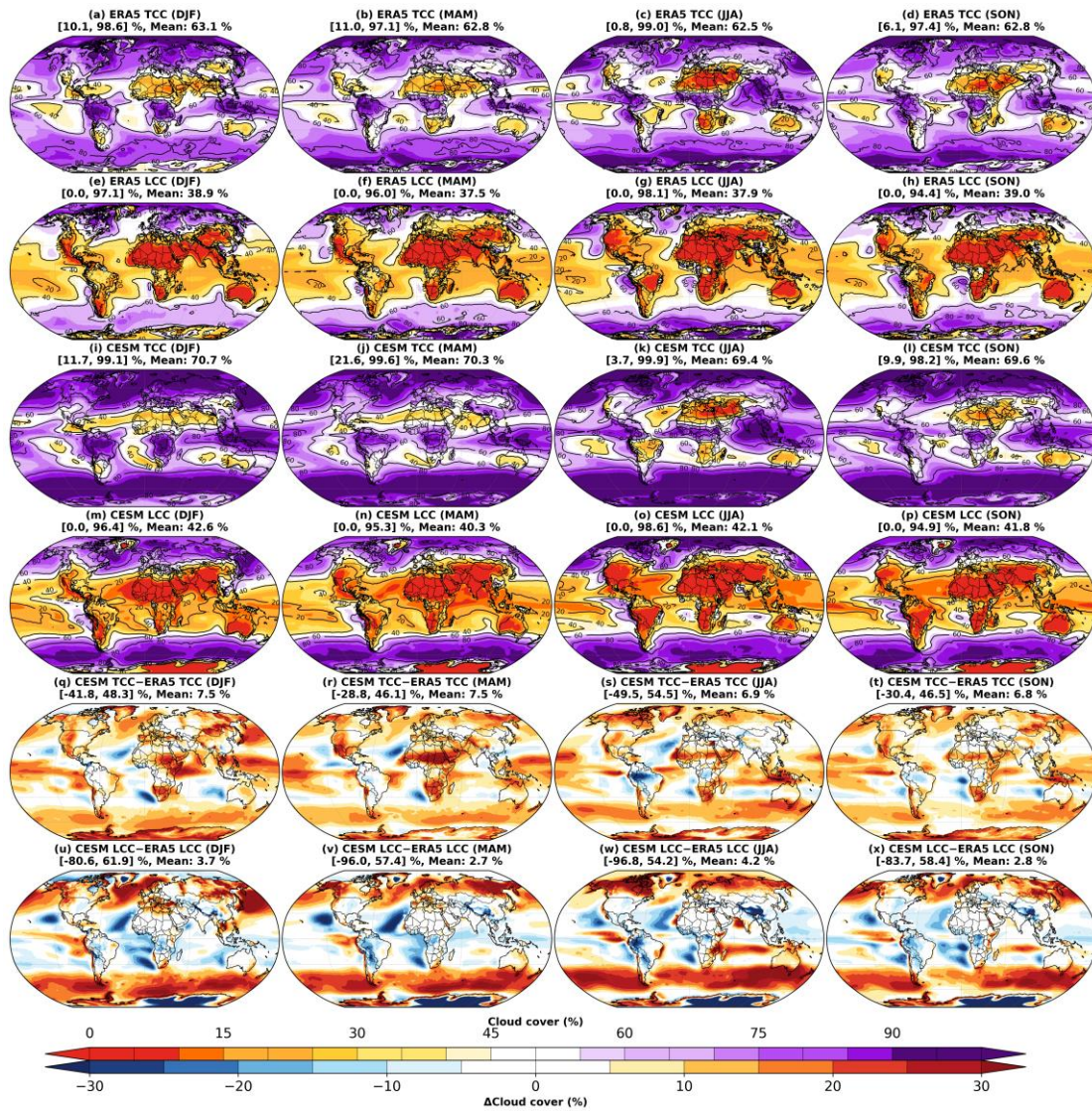


Figure S2. Seasonal validation of total cloud cover (TCC) and low-level cloud cover (LCC) by the CESM during the historical period (1979-2014).

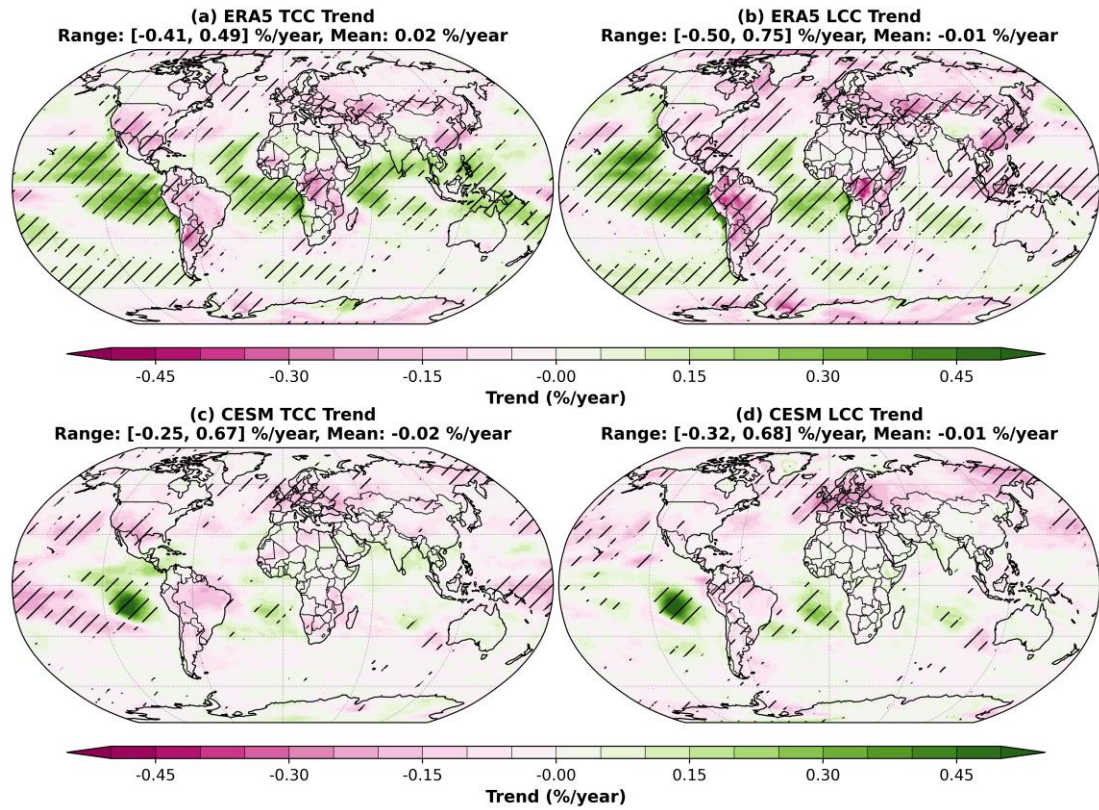


Figure S3. Validation of total cloud cover (TCC) and low-level cloud cover (LCC) trend by the Community Earth System Model (CESM) against ERA5 (ECMWF Reanalysis v5) during the historical period (1979-2014). The hatching indicates the grids with a significant trend ($p < 0.05$).

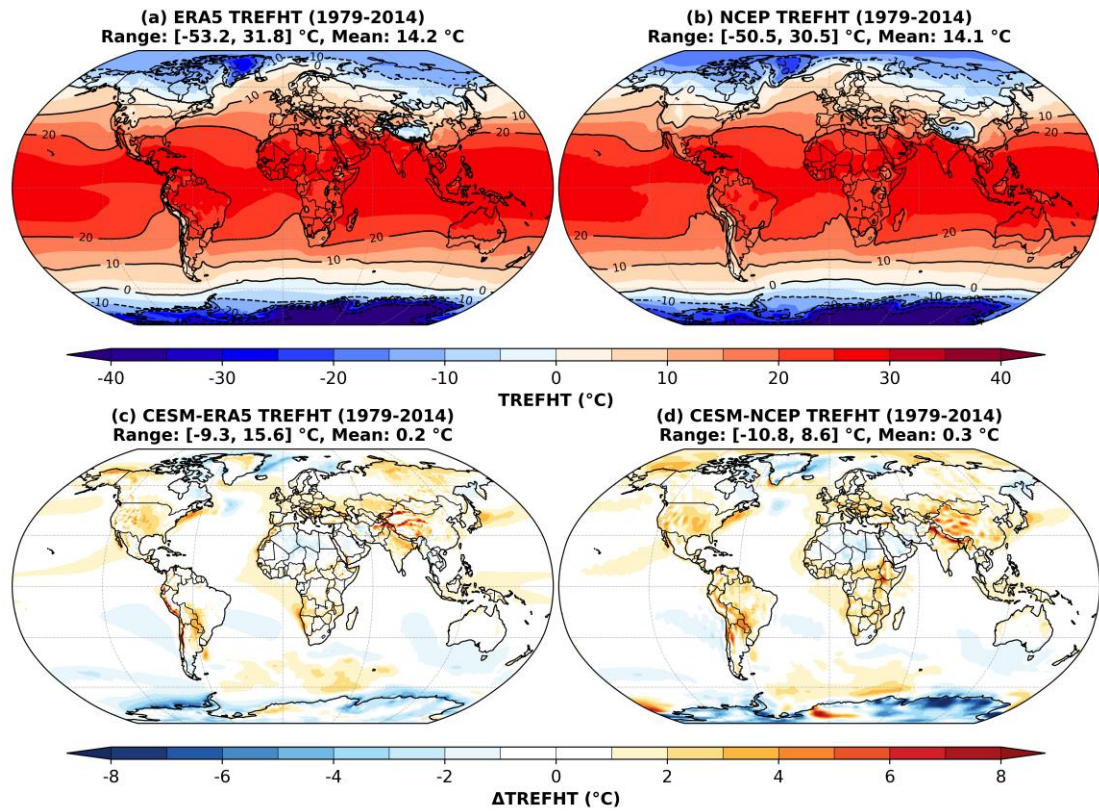


Figure S4. Validation of air temperature at reference height (TREFHT) from ERA5 and National Centers for Environmental Prediction (NCEP) by the CESM during the historical period (1979-2014).

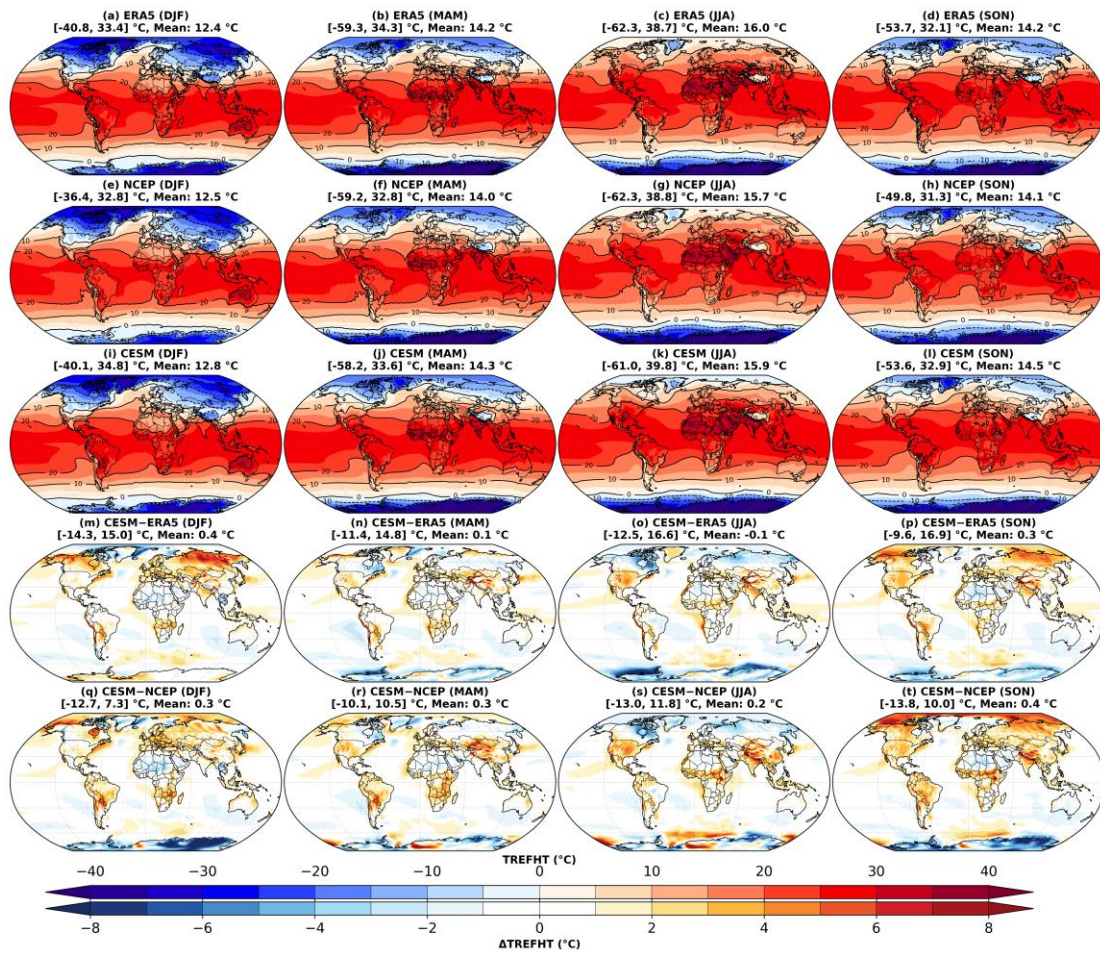


Figure S5. Seasonal validation of air temperature at reference height (TREFHT) from ERA5 and National Centers for Environmental Prediction (NCEP) by the CESM during the historical period (1979-2014).

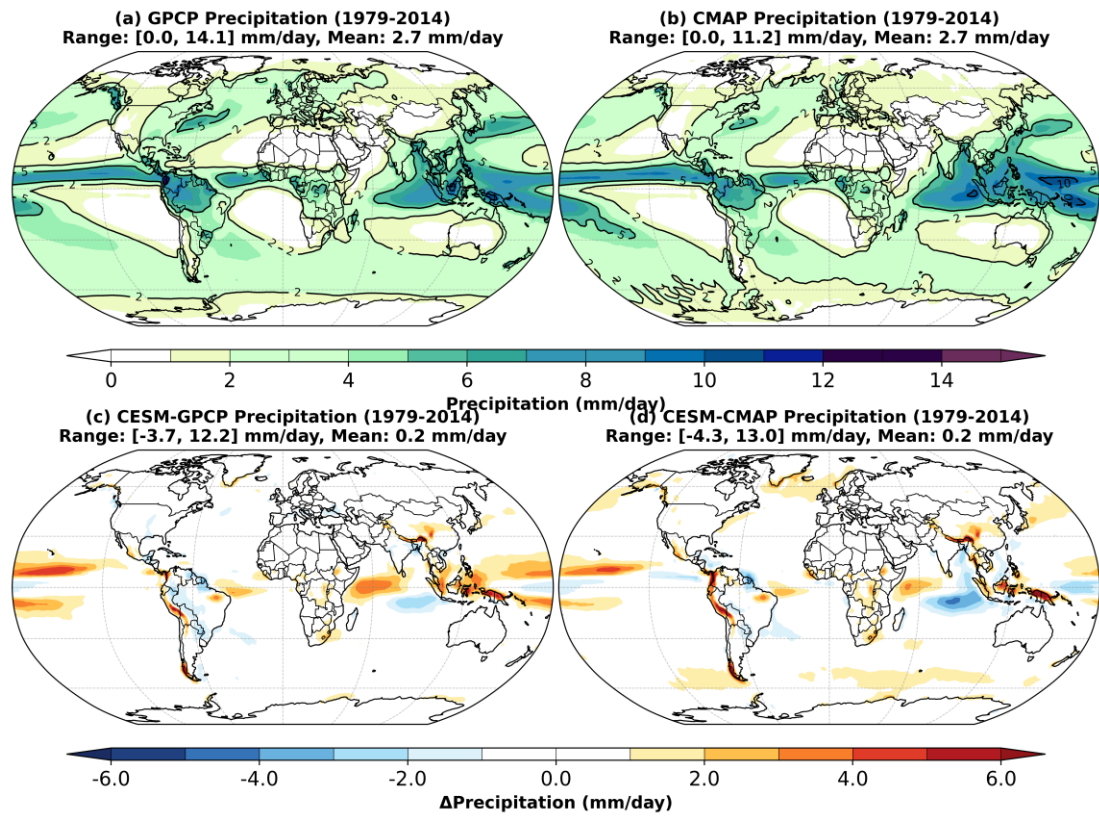


Figure S6. Validation of precipitation by the Community Earth System Model (CESM) against Global Precipitation Climatology Project (GPCP) and CPC Merged Analysis of Precipitation (CMAP) during the historical period (1979-2014).

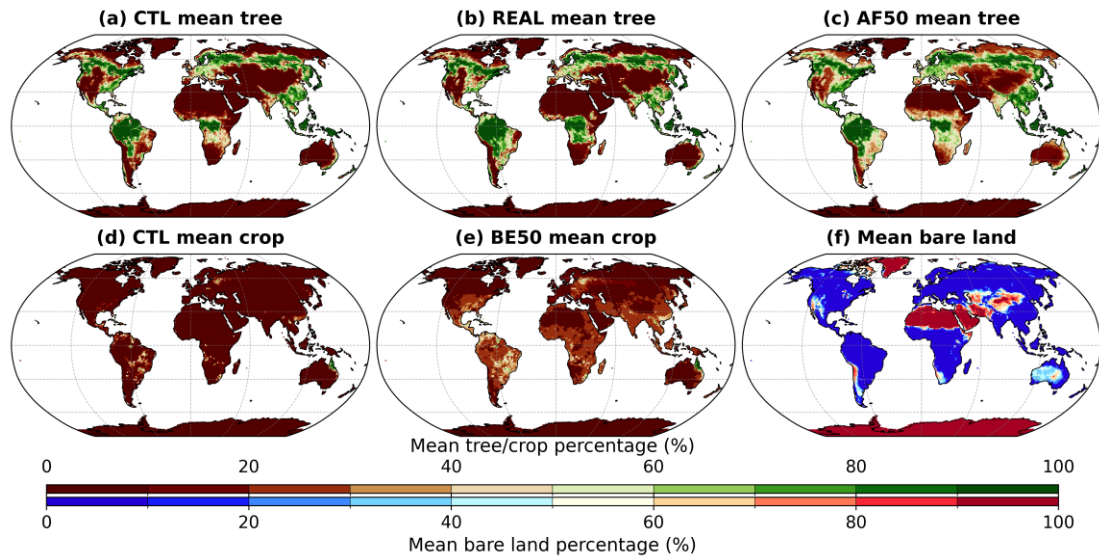


Figure S7. Spatial distribution of forest and bioenergy crop under different land use and land cover scenarios. (a and d) Spatial distribution of forests and bioenergy crops in the original land-use data (CTL run). (b) Spatial distribution of forests in the realistic afforestation scenario (REAL). (c and e) Spatial distribution of forests and bioenergy crops in the 50% idealized linear afforestation (AF50) and bioenergy crop expansion (BE50) scenarios. (f) Spatial distribution of bare land.

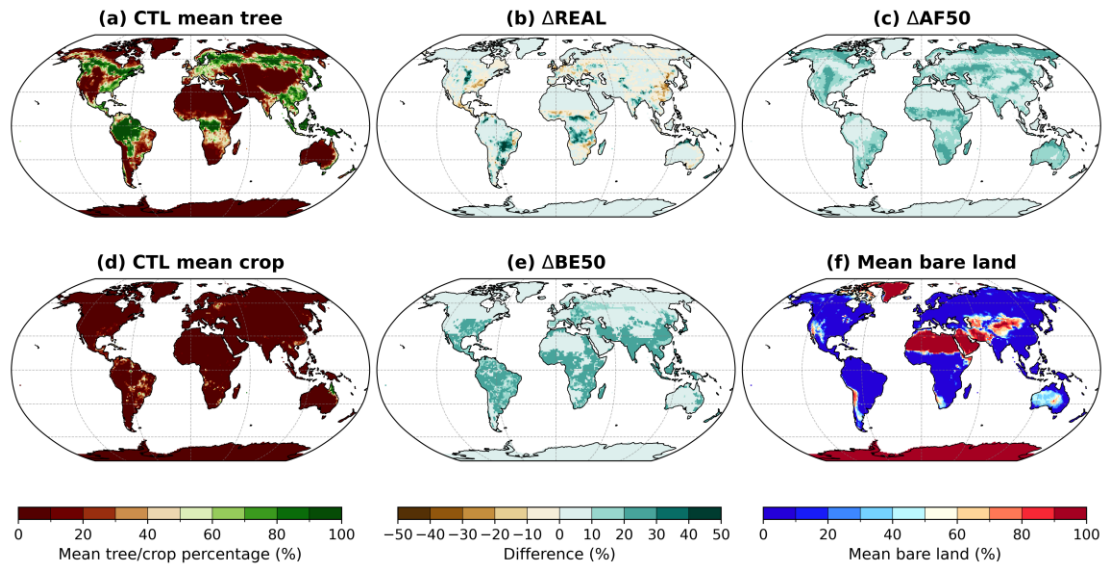


Figure S8. Spatial distribution of mean tree, crop, and bare land percentages and their differences under different land-use scenarios averaged over 2015-2100. (a and d) Mean tree and crop percentage in the CTL scenario. (b, c) Differences in tree percentage for the REAL and AF50 scenarios relative to CTL (Δ REAL and Δ AF50, respectively). (e) Difference in crop percentage for the BE50 scenario relative to CTL (Δ BE50). (f) Mean bare land percentage. The left and right color bars indicate the absolute percentages for vegetation (a, d) and bare land (f), respectively, whereas the middle diverging color bar represents the percentage differences between the forest and bioenergy expansion experiments and CTL scenario.

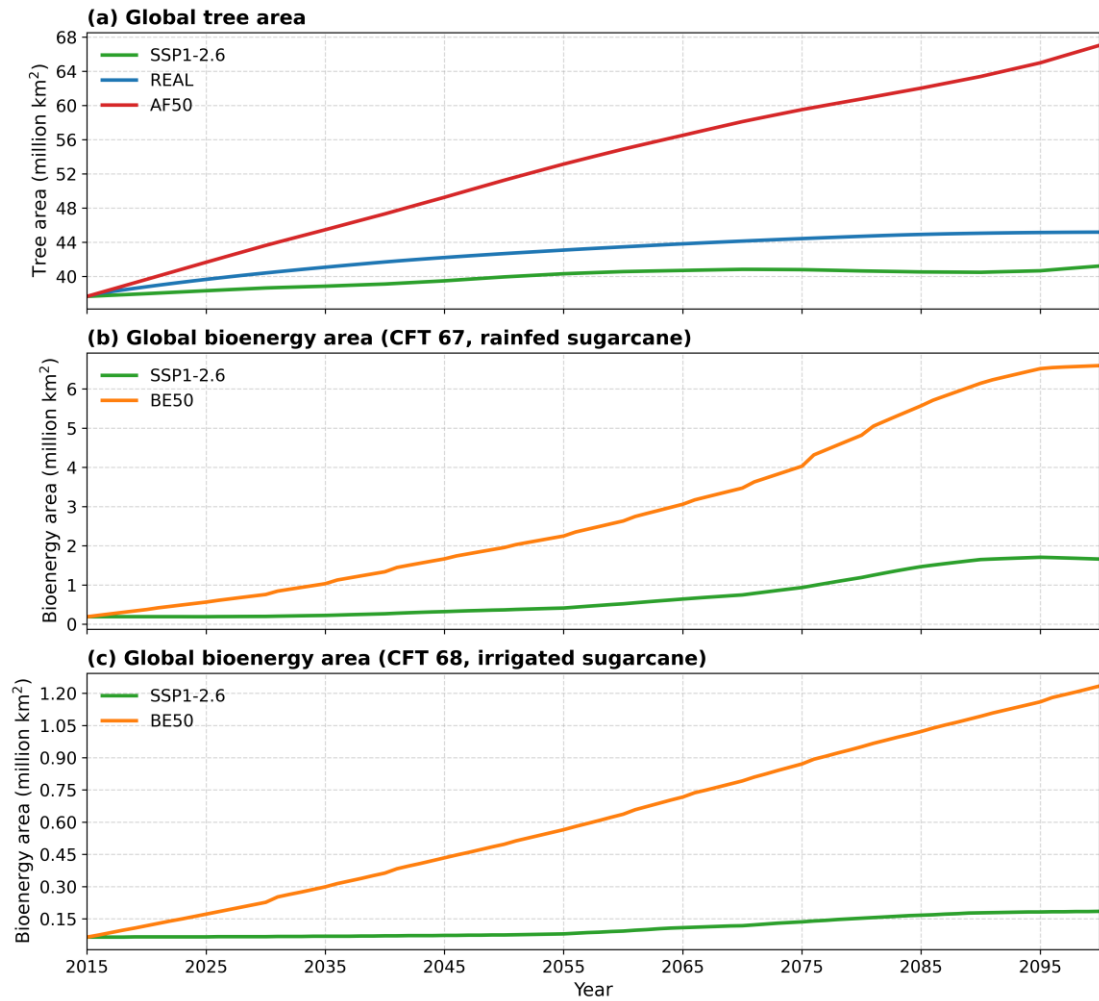


Figure S9. Area change of forest and bioenergy crops under different land-use scenarios (unit: million km²). In the Community Land Model Version 5 (CLM5), indices 67 and 68 correspond to rainfed and irrigated sugarcane, respectively, among bioenergy crops.

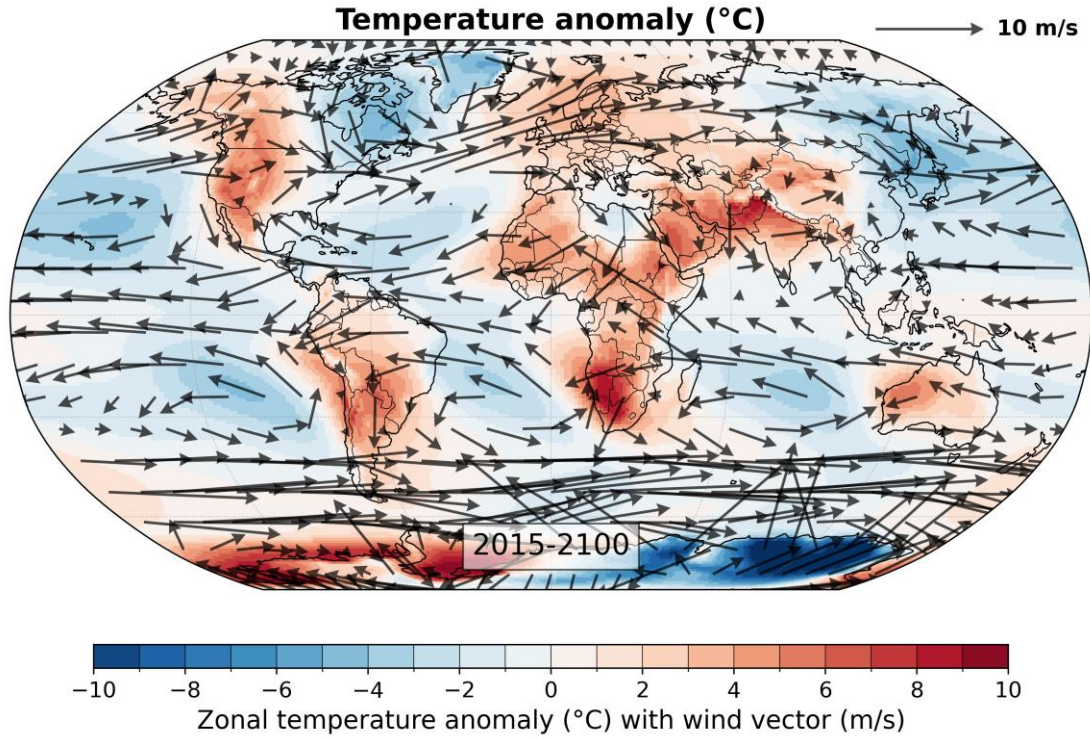


Figure S10. Spatial distribution of 850 hPa zonal temperature anomalies and horizontal wind vectors under the CTL scenario (2015–2100). Color shading illustrates the temperature anomaly of each grid cell relative to its zonal mean (°C). Specifically, warm colors (reds) indicate regions where the lower-tropospheric temperature is higher than the zonal average, predominantly clustered over strongly heated landmasses. Conversely, cool colors (blues) represent temperatures lower than the zonal average, mostly corresponding to vast oceanic regions with higher heat capacity. The black arrows denote the mean wind velocity and direction at 850 hPa (reference vector: 10 m/s). The superposition of wind fields and thermal anomalies highlights the low-level stationary wave structures and temperature advection driven by land-sea thermal contrasts. Notably, relative to the zonal mean, the positive temperature anomalies over land are primarily associated with the westerly circulation, whereas the negative anomalies are largely driven by northerly flows.

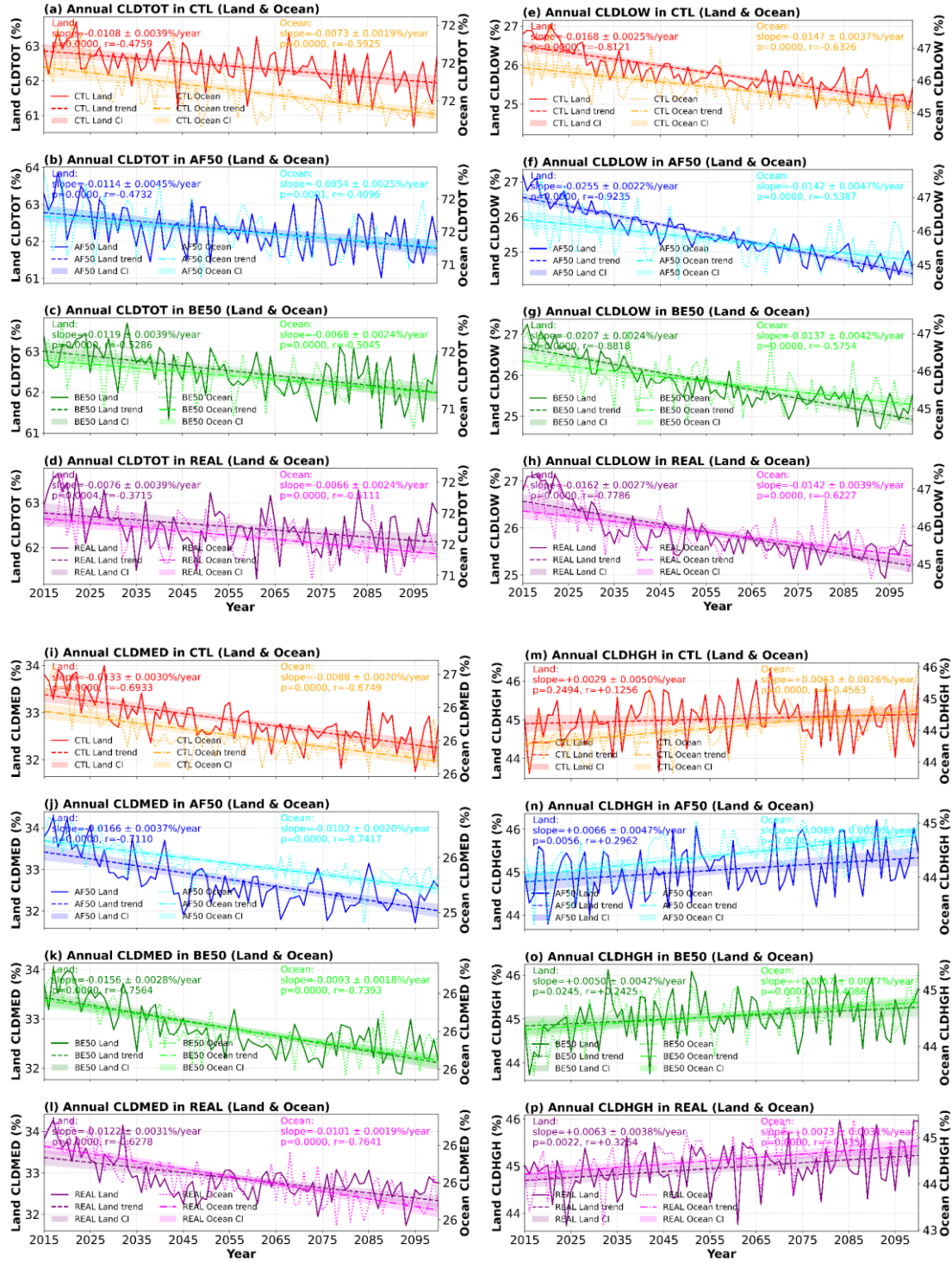


Figure S11. Long-term trends with area weight of cloud cover in land and ocean.

The shade indicates 95% confidence interval (CI).

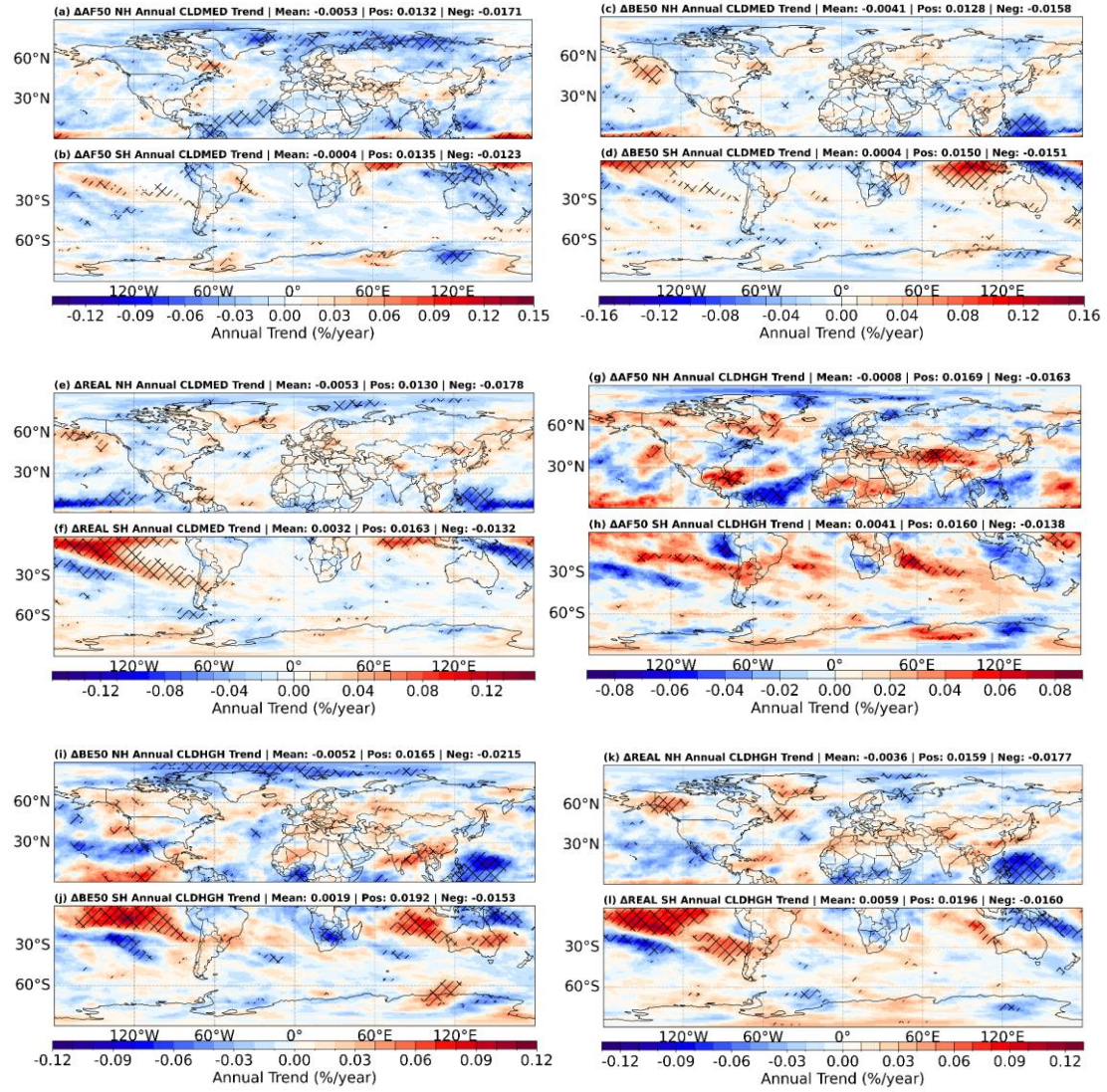


Figure S12. Long-term spatial trends of medium-level cloud cover (CLDMED) and high-level cloud cover (CLDHGH) induced by the expansion of forest and bioenergy in Northern Hemisphere (NH) and Southern Hemisphere (SH). The cross-hatching indicates the grids with a significant trend ($p < 0.05$).

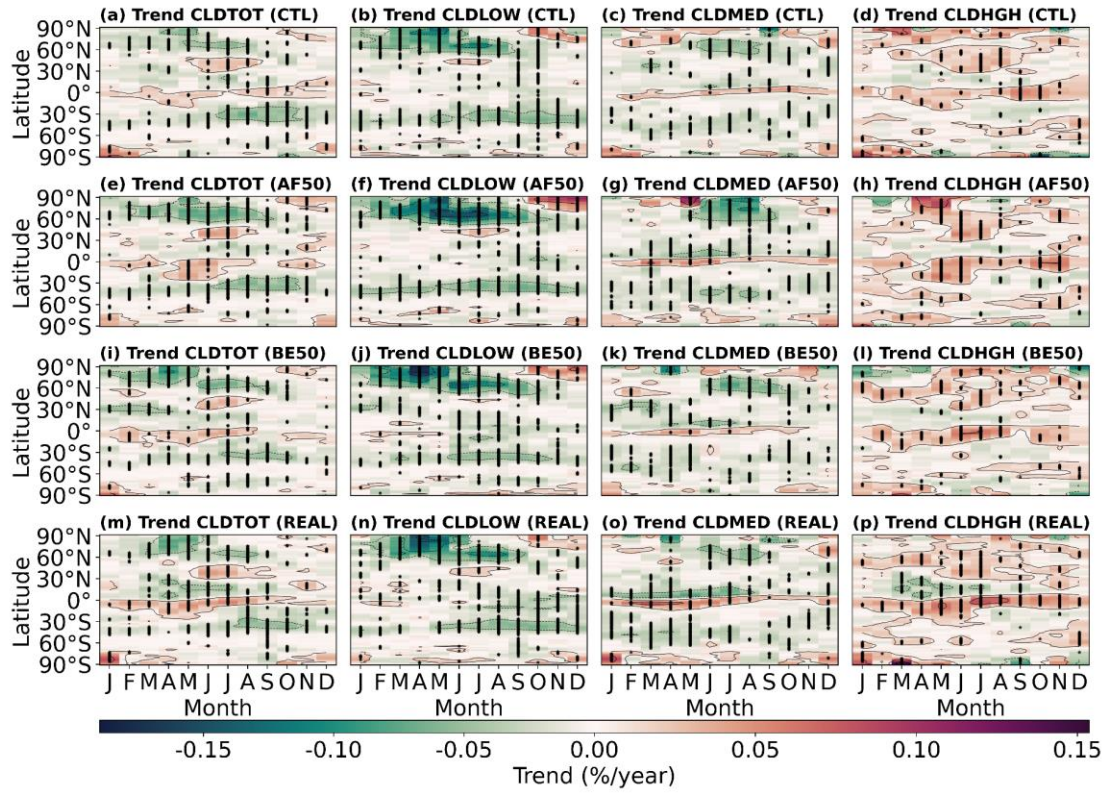


Figure S13. Annual trend in cloud cover across different simulation experiments grouped by month. The black asterisk indicates the grids with a significant trend ($p < 0.05$).

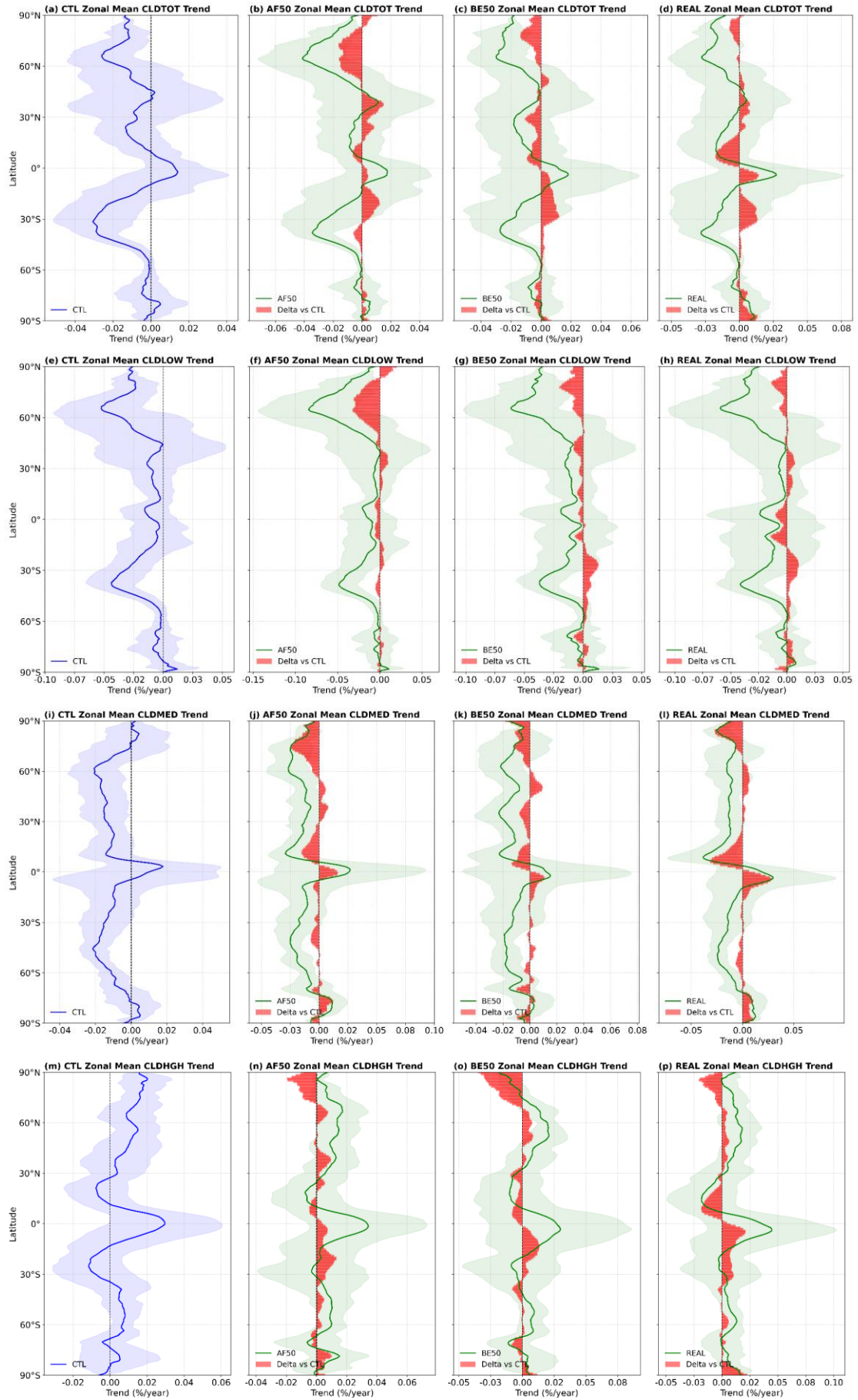


Figure S14. Latitudinal distribution of annual trends in cloud cover across different simulation experiments. The shade indicates standard deviation.

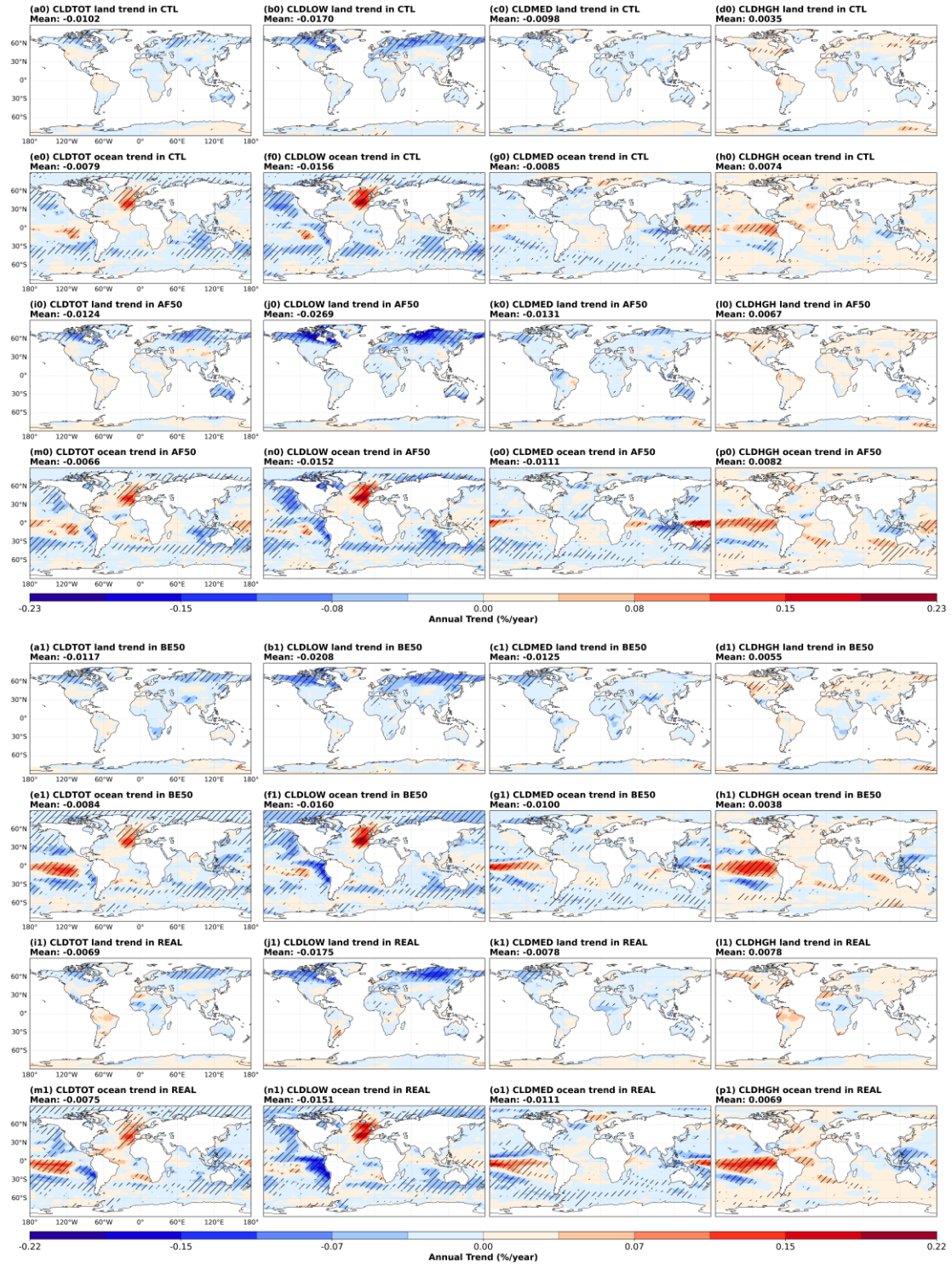


Figure S15. Distribution of cloud cover trend on land and ocean across different simulation experiments. The slash indicates the grids with a significant trend ($p < 0.05$).

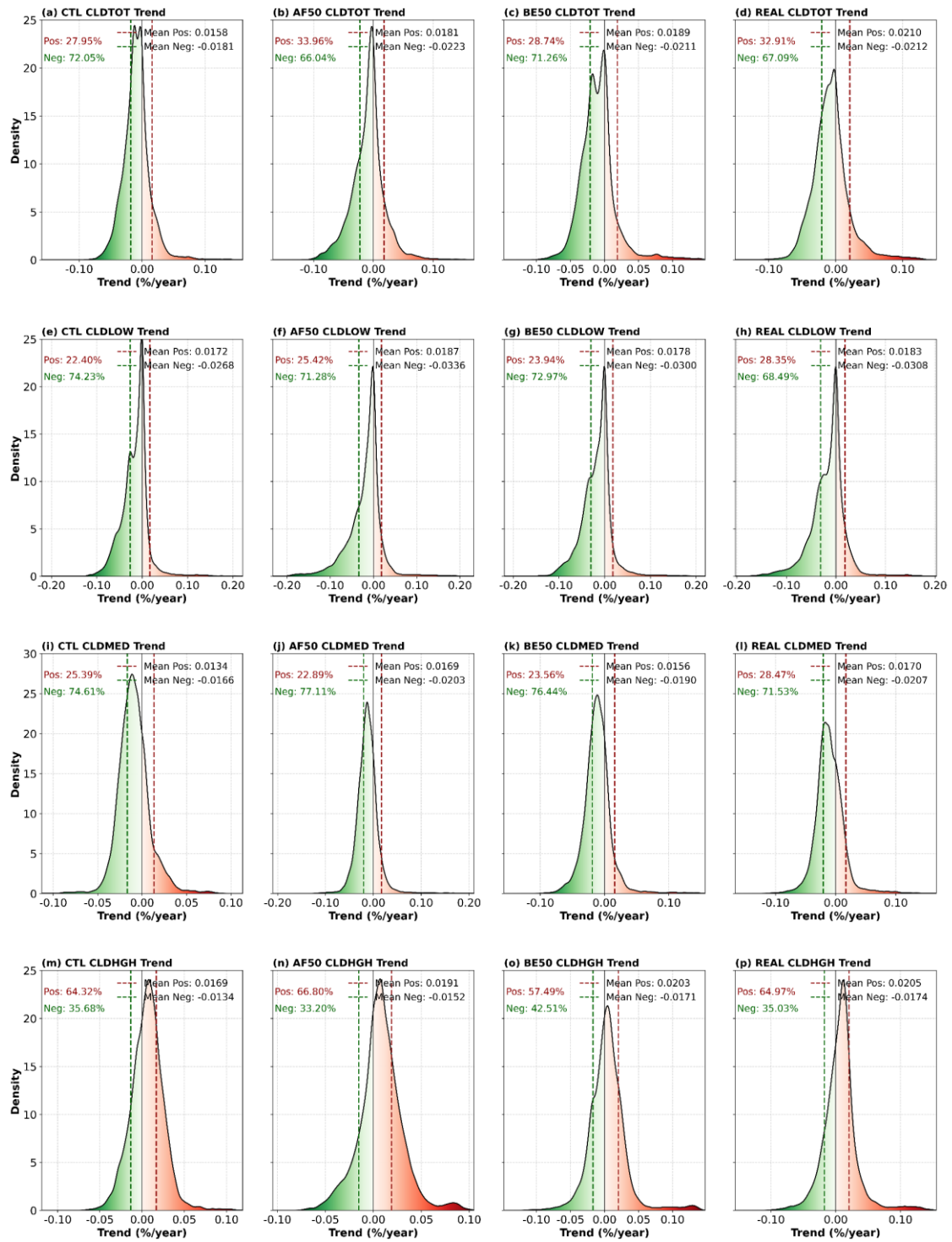


Figure S16. Proportions and average values of cloud cover trends greater than 0 and less than 0 across different simulation experiments.

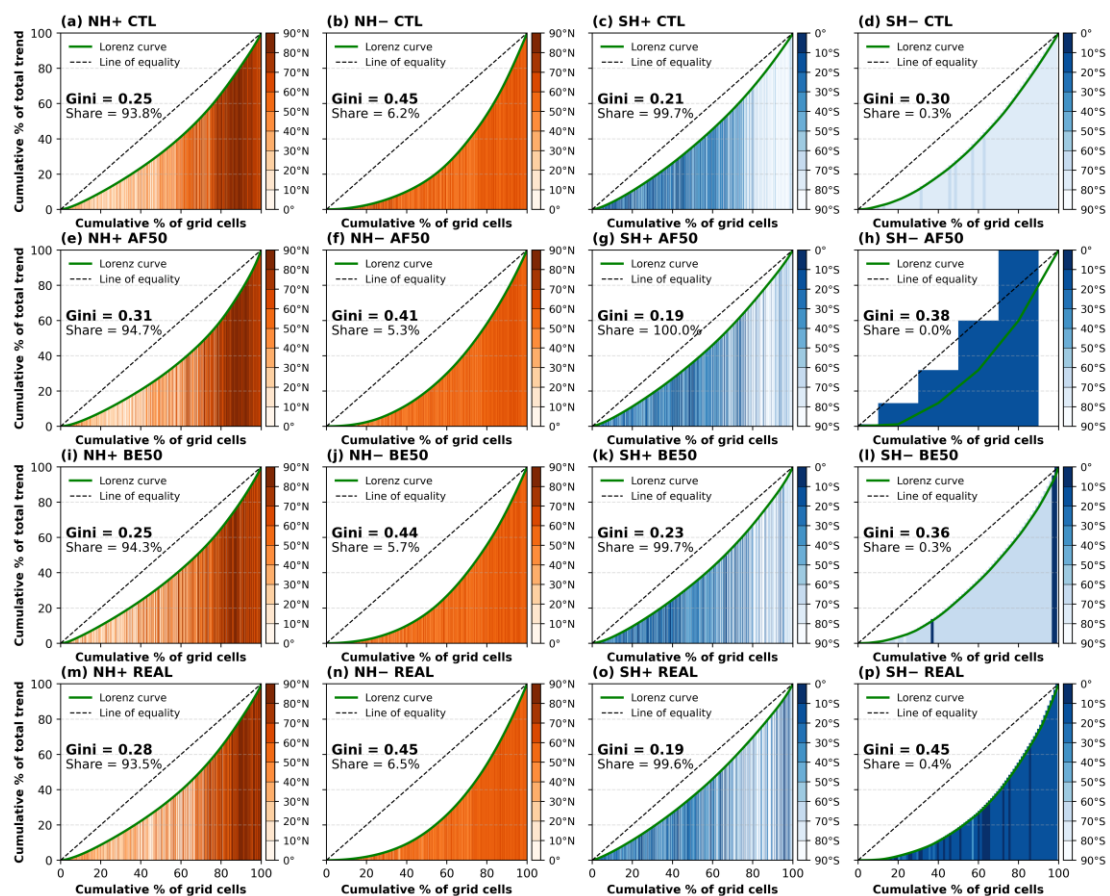


Figure S17. Gini coefficient of temperature at reference height (TREFHT) trends in the different simulations. Share indicates the percentage of an increase or decrease in the trend.

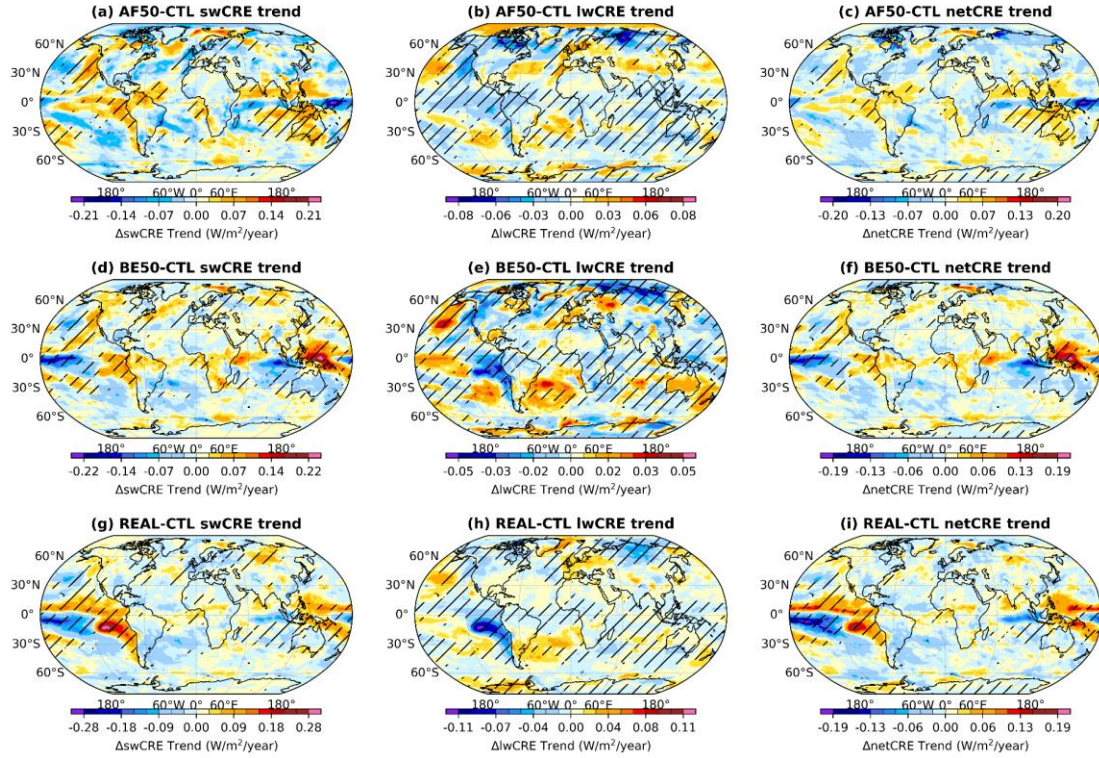


Figure S18. Spatial distribution of annual trends in shortwave cloud radiative effect (CRE), longwave CRE and netCRE induced by different land cover changes. The hatch indicates the grids with a significant trend ($p < 0.05$).

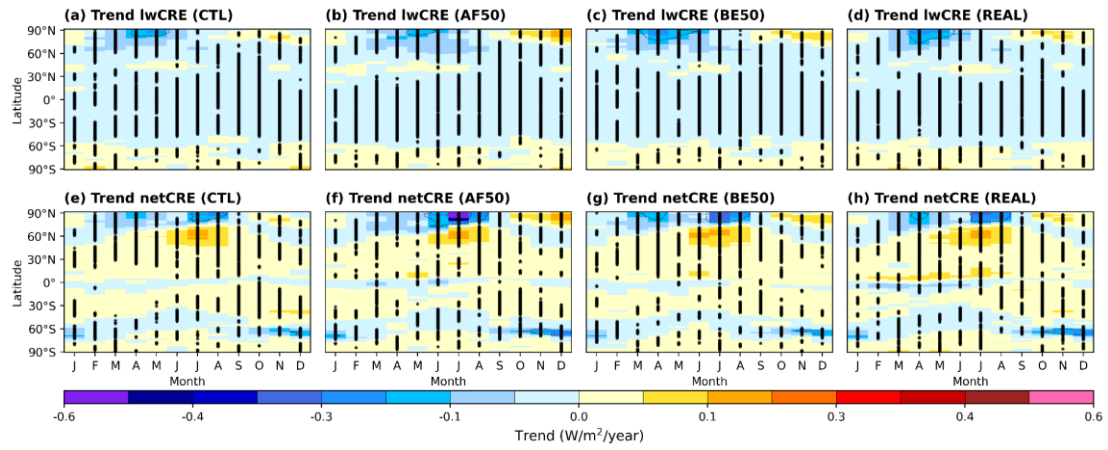


Figure S19. Annual trend in longwave CRE and net CRE across different simulation experiments grouped by month. The black asterisk indicates the grids with a significant trend ($p < 0.05$).

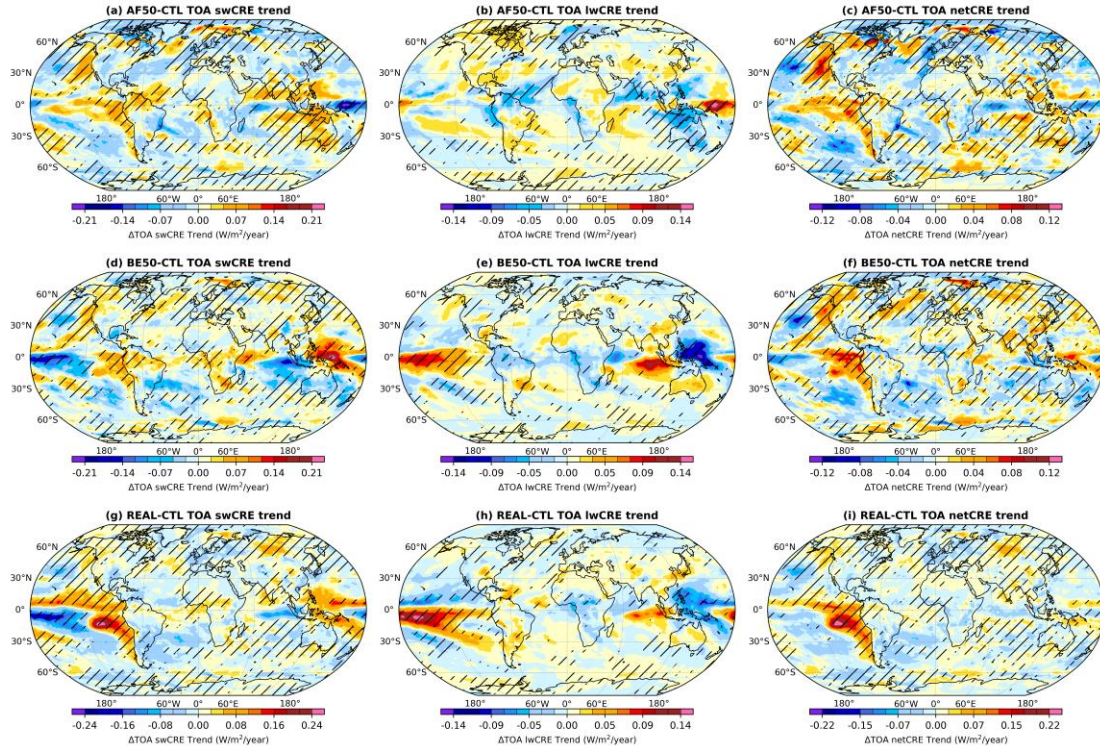


Figure S20. Spatial distribution of annual trends in shortwave cloud radiative effect (CRE), longwave CRE and netCRE at the top of atmosphere (TOA) induced by different land cover changes. The hatch indicates the grids with a significant trend ($p < 0.05$).

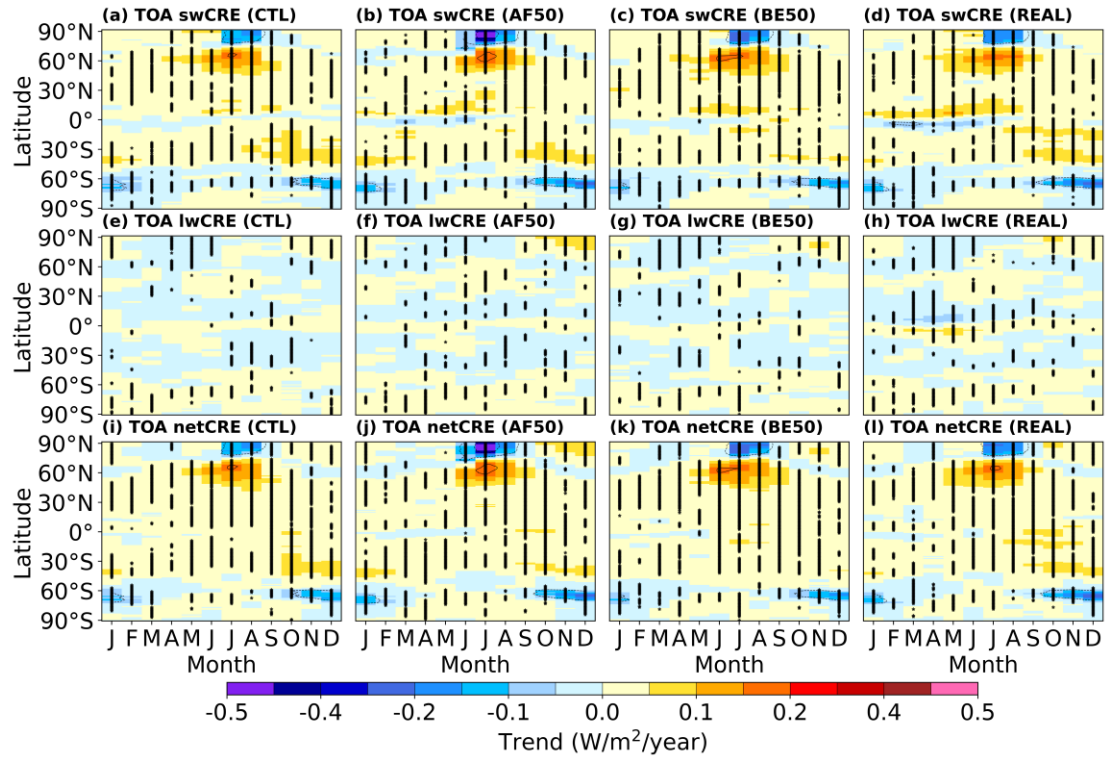


Figure S21. Annual trend in shortwave CRE, longwave CRE and net CRE at the TOA across different simulation experiments grouped by month. The black asterisk indicates the grids with a significant trend ($p < 0.05$).

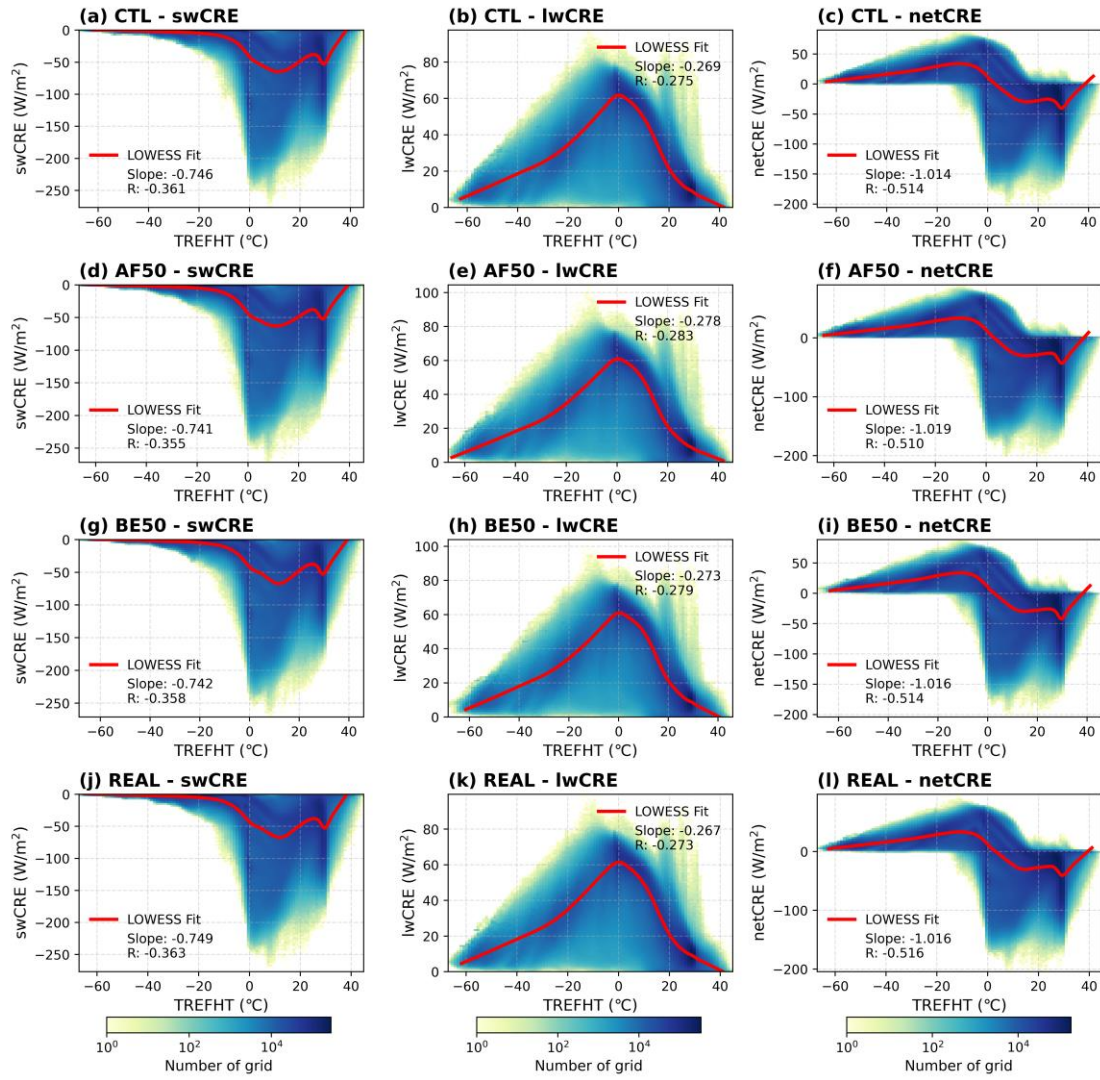


Figure S22. Dependence of cloud radiative effects on temperature. Two-dimensional density histograms illustrate the relationship between shortwave (swCRE), longwave (lwCRE), and net (netCRE) cloud radiative effects and surface reference temperature (TREFHT) across the CTL, AF50, BE50, and REAL scenarios. The color bar indicates the number of grid points on a logarithmic scale. The red solid curves represent the non-linear LOWESS (Locally Weighted Scatterplot Smoothing) fits, while the overall linear regression slopes and regression coefficients (R) are inset in each panel.

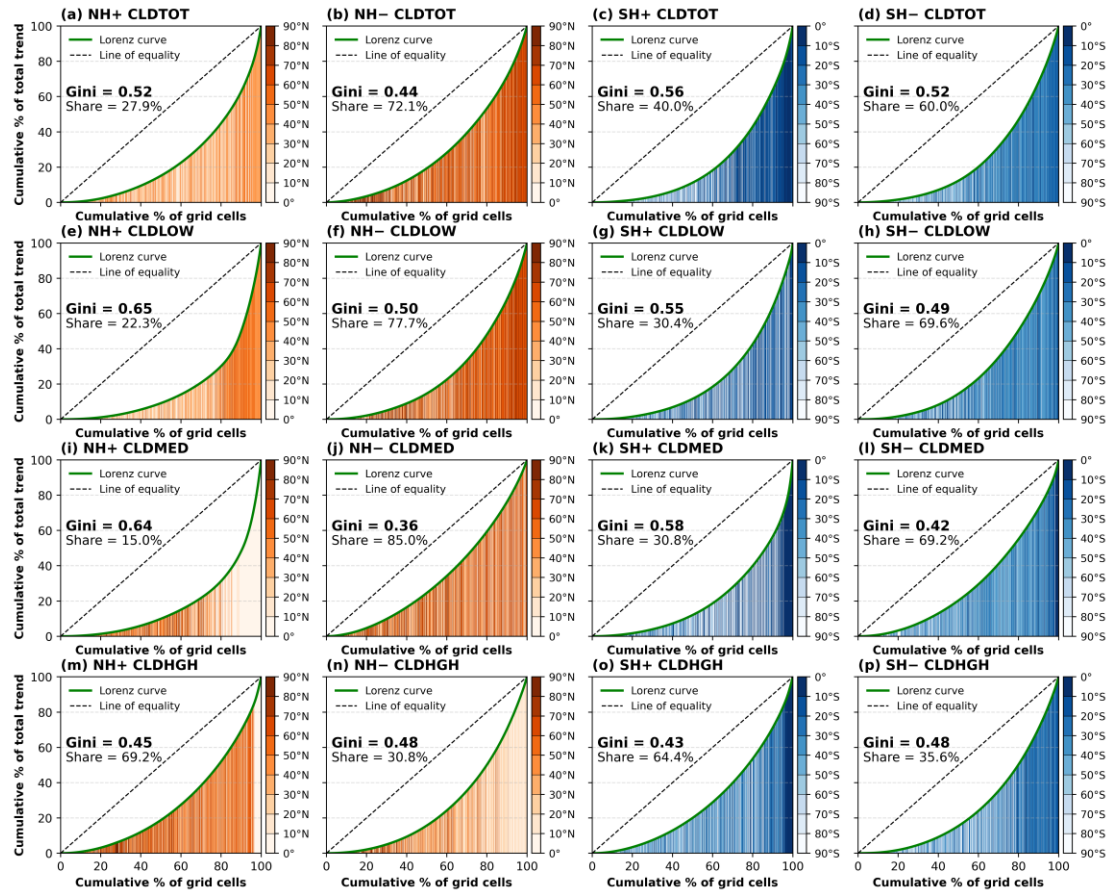


Figure S23. Gini coefficient of cloud cover trends in the AF50 simulation. Share indicates the percentage of an increase or decrease in the trend.

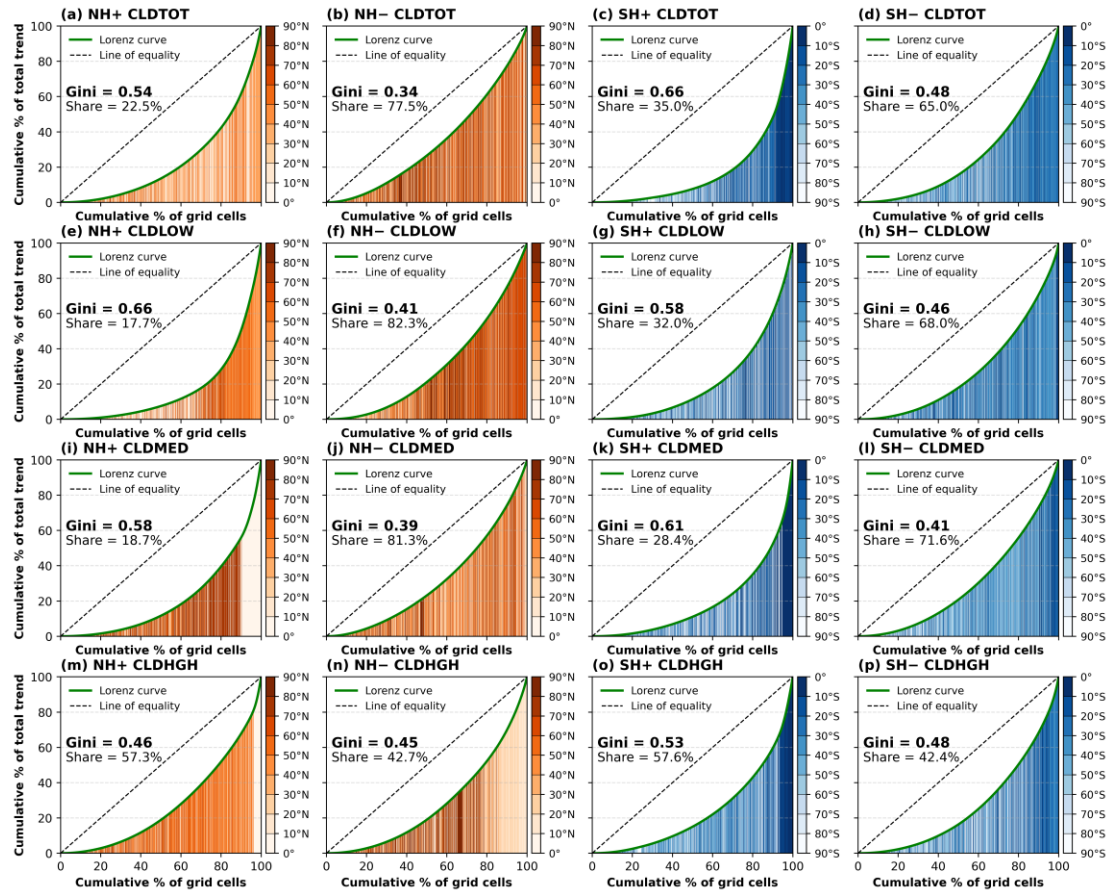


Figure S24. Gini coefficient of cloud cover trends in the BE50 simulation. Share indicates the percentage of an increase or decrease in the trend.

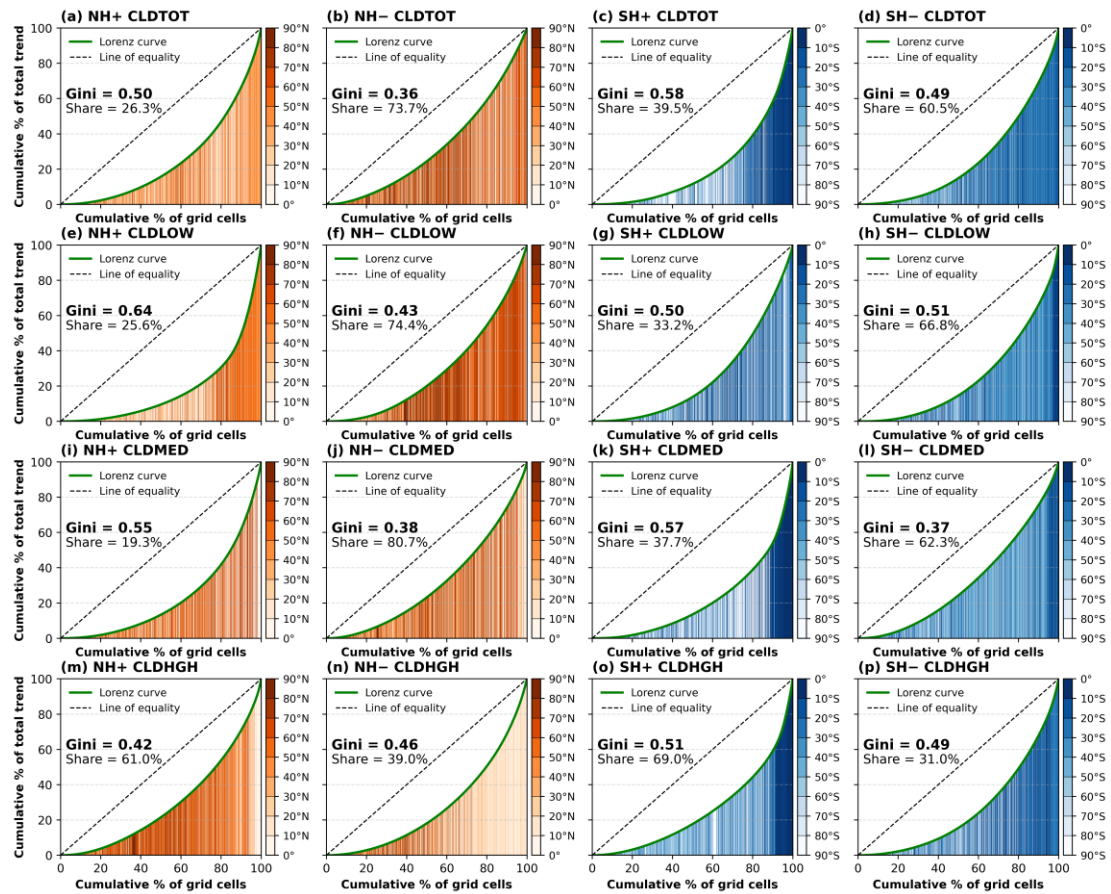


Figure S25. Gini coefficient of cloud cover trends in the REAL simulation. Share indicates the percentage of an increase or decrease in the trend.

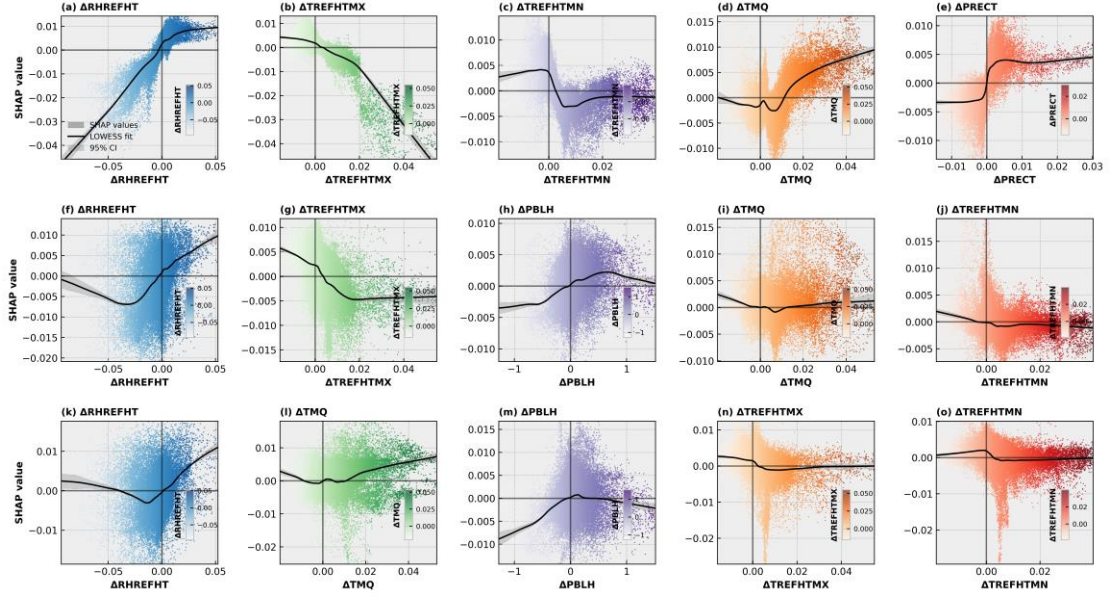


Figure S26. Dependence plots of the top five important variables for low-level cloud cover (CLDLOW), as revealed by the machine learning model. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment. The background scatter displays the full samples colored by feature values. The SHAP dependence trend (black solid line) was visualized using Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland, 1979) with a smoothing fraction of $\text{frac}=0.3$ (using 30% of nearest neighbors for local regression) and $\text{it}=3$ robust iterations to reduce outlier influence, and gray shaded band indicates 95% confidence interval (CI) of LOWESS regression (calculated by a random subsample of 10000 grid cells).

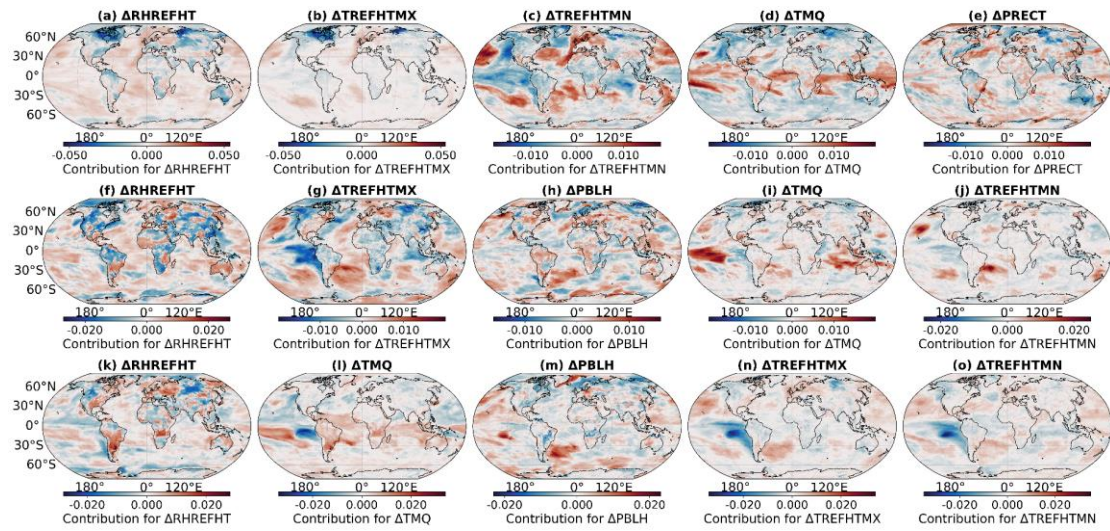


Figure S27. Spatial drivers and contribution of CLDLow revealed by machine learning response to different land use change. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment.

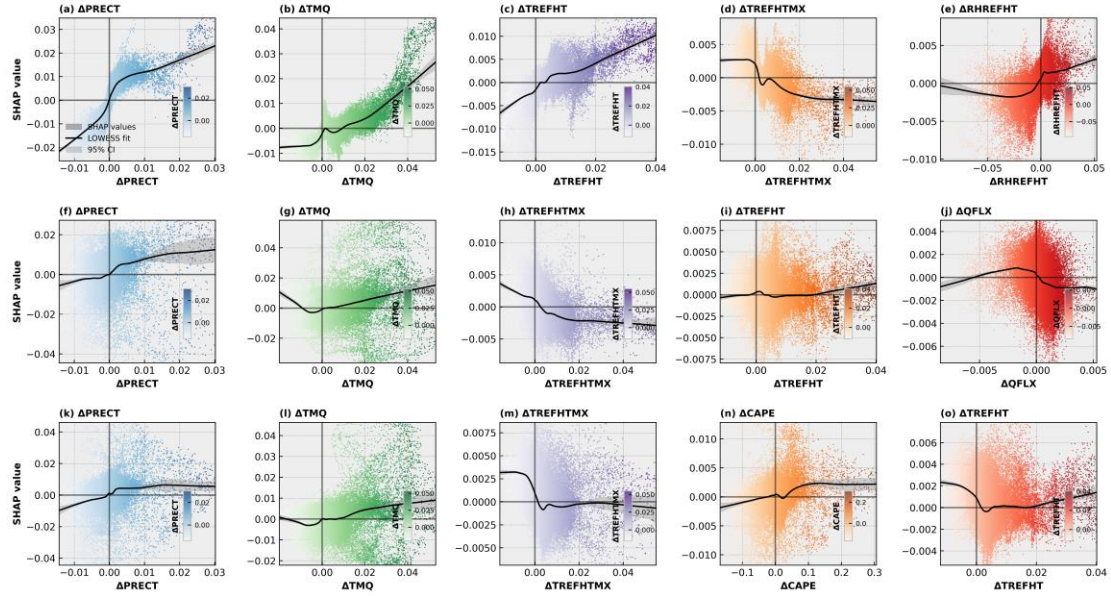


Figure 28. Dependence plots of the top five important variables for CLDMED, as revealed by the machine learning model. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment. The background scatter displays the full samples colored by feature values. The SHAP dependence trend (black solid line) was visualized using Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland, 1979) with a smoothing fraction of $\text{frac}=0.3$ (using 30% of nearest neighbors for local regression) and $\text{it}=3$ robust iterations to reduce outlier influence, and gray shaded band indicates 95% confidence interval (CI) of LOWESS regression (calculated by a random subsample of 10000 grid cells).

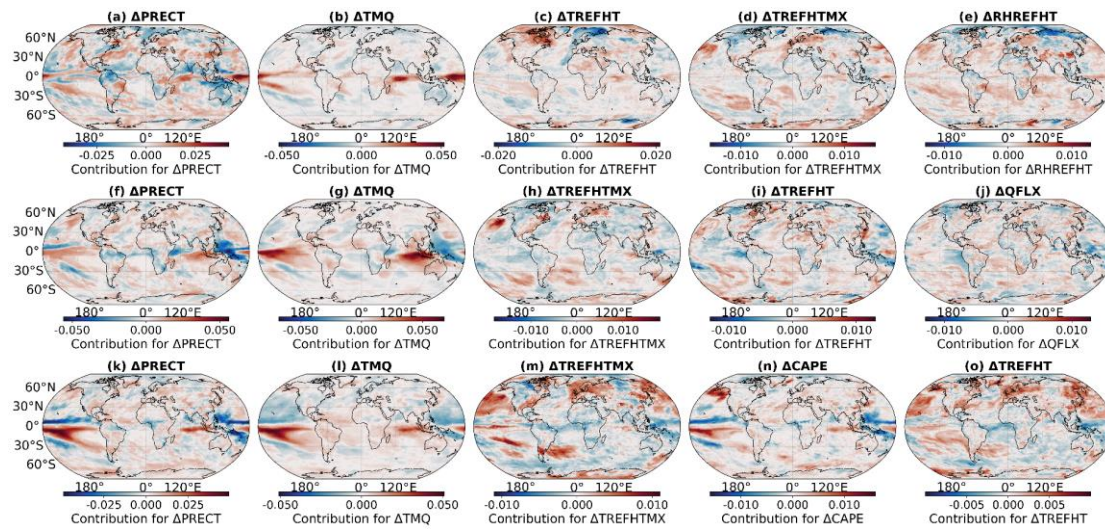


Figure S29. Spatial drivers and contribution of CLDMED revealed by machine learning response to historical land use change. The first row corresponds to the top five drivers of AF50 experiment, the second row to the top five drivers of BE50 experiment, and the third row to the top five drivers of REAL experiment.

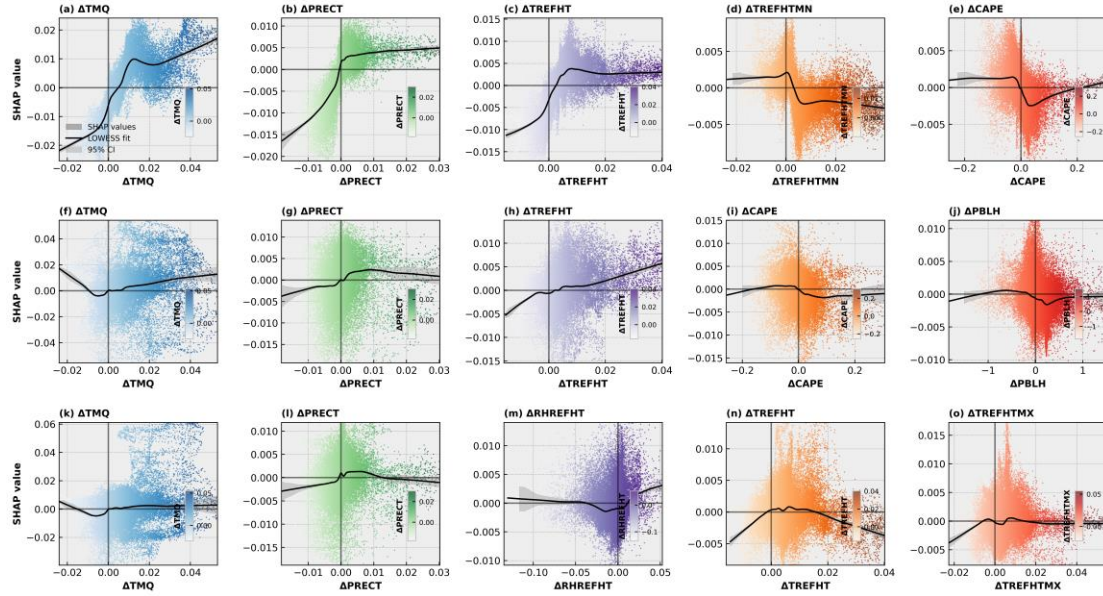


Figure S30. Dependence plots of the top five important variables for CLDHG, as revealed by the machine learning model. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment. The background scatter displays the full samples colored by feature values. The SHAP dependence trend (black solid line) was visualized using Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland, 1979) with a smoothing fraction of $\text{frac}=0.3$ (using 30% of nearest neighbors for local regression) and $\text{it}=3$ robust iterations to reduce outlier influence, and gray shaded band indicates 95% confidence interval (CI) of LOWESS regression (calculated by a random subsample of 10000 grid cells).

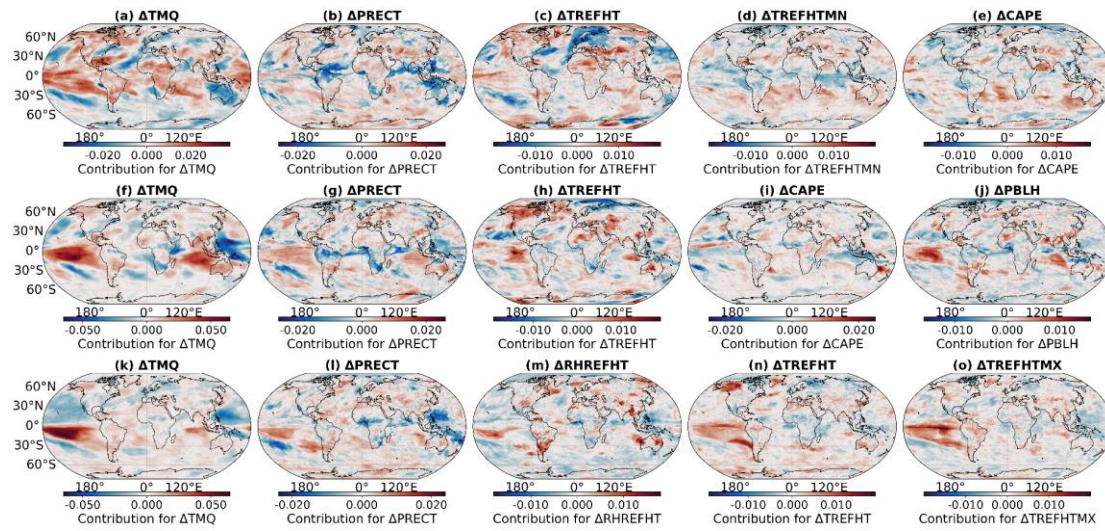


Figure S31. Spatial drivers and contribution of CLDHGH revealed by machine learning response to historical land use change. The first row corresponds to the top five drivers of AF50 experiment, the second row to the top five drivers of BE50 experiment, and the third row to the top five drivers of REAL experiment.

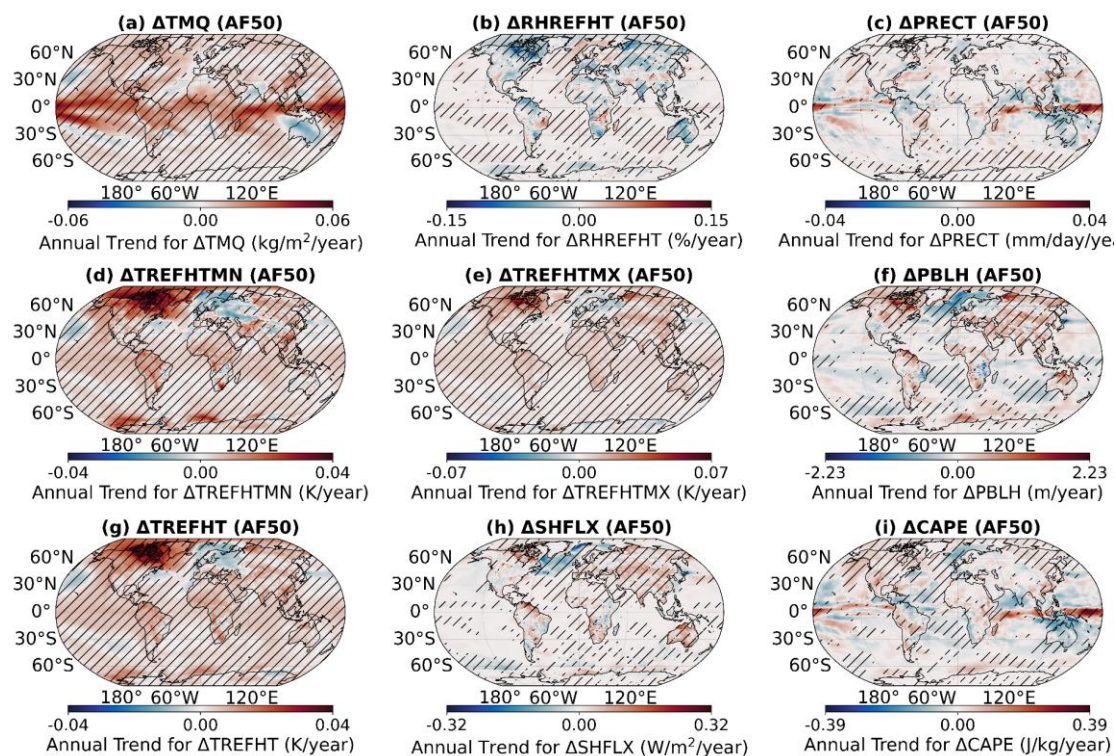


Figure S32. Spatial distribution of annual trends in key atmospheric variables induced by 50% linear afforestation (AF50). The hatch indicates the grids with a significant trend ($p < 0.05$).

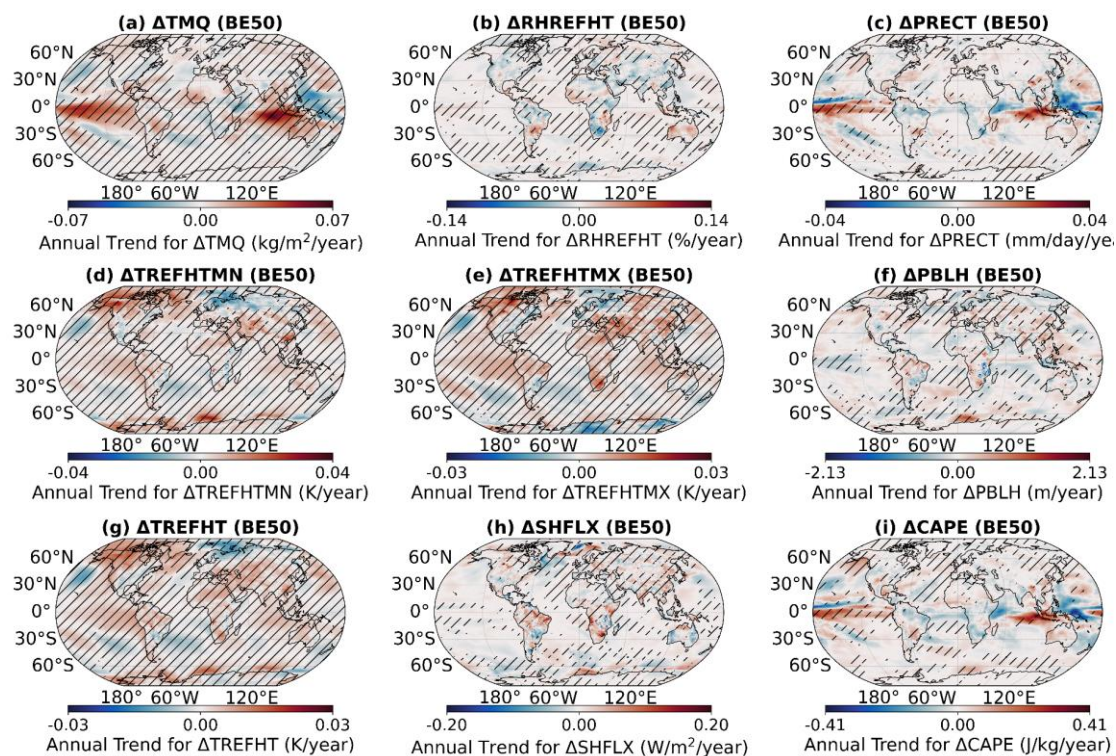


Figure S33. Spatial distribution of annual trend in key atmospheric variables induced by 50% linear bioenergy crop expansion (BE50). The hatch indicates the grids with a significant trend ($p < 0.05$).

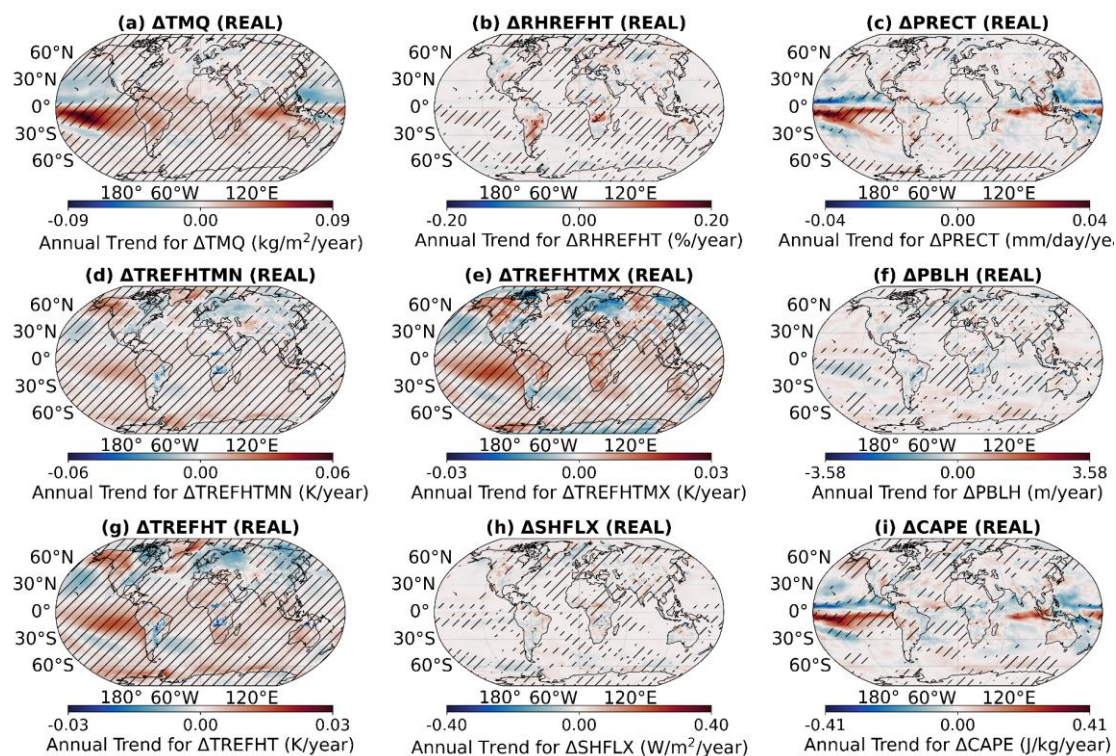


Figure S34. Spatial distribution of annual trend in key atmospheric variables induced by realistic afforestation (REAL). The hatch indicates the grids with a significant trend ($p < 0.05$).

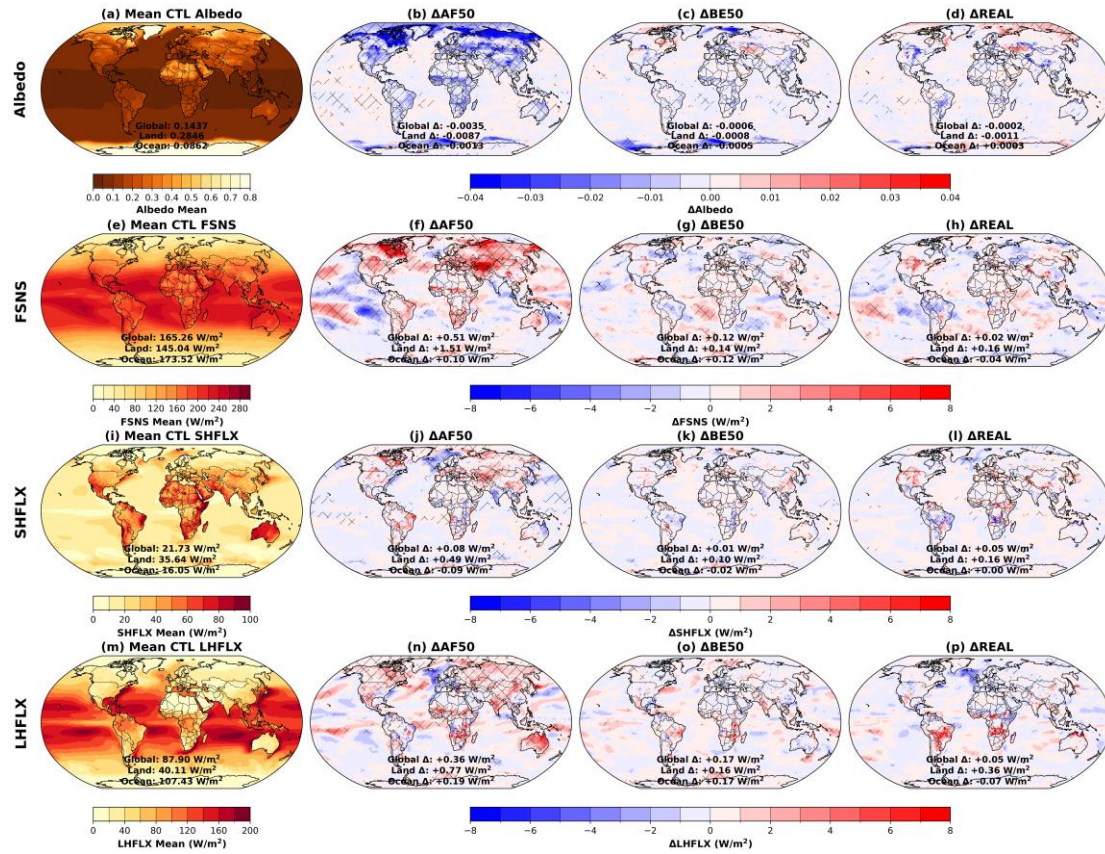


Figure S35. Spatial distributions of multi-year mean (2015–2100) surface albedo and energy fluxes, and their responses to different LUCC scenarios. The rows display four key variables: (a-d) surface albedo, (e-h) net surface shortwave radiation (FSNS), (i-l) sensible heat flux (SHFLX), and (m-p) latent heat flux (LHFLX). The leftmost column presents the climatological mean states under the control simulation (CTL). The subsequent three columns depict the difference for the idealized afforestation (AF50), bioenergy expansion (BE50), and realistic afforestation (REAL) scenarios relative to the CTL. In the difference panels, warm colors (reds) indicate an increase in the respective variable, whereas cool colors (blues) denote a decrease. Cross-hatched regions highlight statistically significant differences at the 95% confidence level, determined by a paired *t*-test ($p < 0.05$). Area-weighted mean values (or mean difference) for the global, land, and ocean domains are inset at the bottom of each panel.

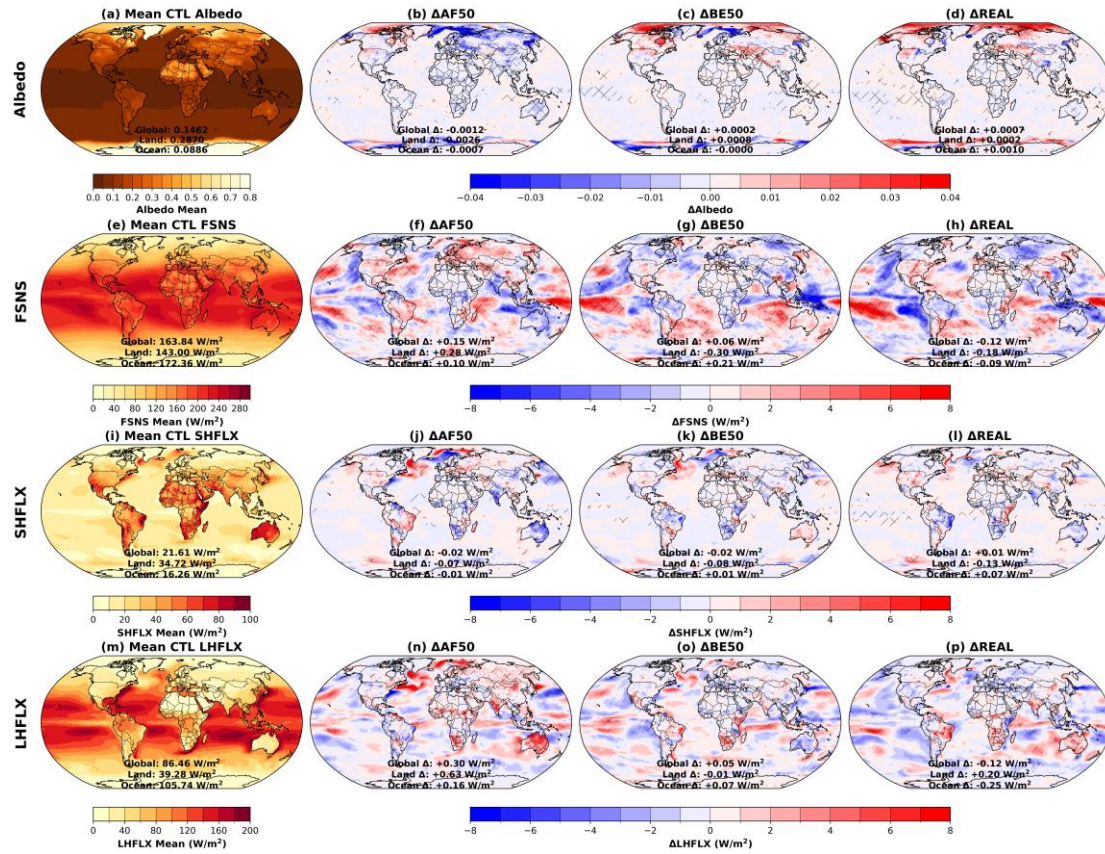


Figure S36. Spatial distributions of multi-year mean (2015–2034) surface albedo and energy fluxes, and their responses to different LUCC scenarios. The rows display four key variables: (a-d) surface albedo, (e-h) net surface shortwave radiation (FSNS), (i-l) sensible heat flux (SHFLX), and (m-p) latent heat flux (LHFLX). The leftmost column presents the climatological mean states under the control simulation (CTL). The subsequent three columns depict the difference for the idealized afforestation (AF50), bioenergy expansion (BE50), and realistic afforestation (REAL) scenarios relative to the CTL. In the difference panels, warm colors (reds) indicate an increase in the respective variable, whereas cool colors (blues) denote a decrease. Cross-hatched regions highlight statistically significant differences at the 95% confidence level, determined by a paired t -test ($p < 0.05$). Area-weighted mean values (or mean difference) for the global, land, and ocean domains are inset at the bottom of each panel.

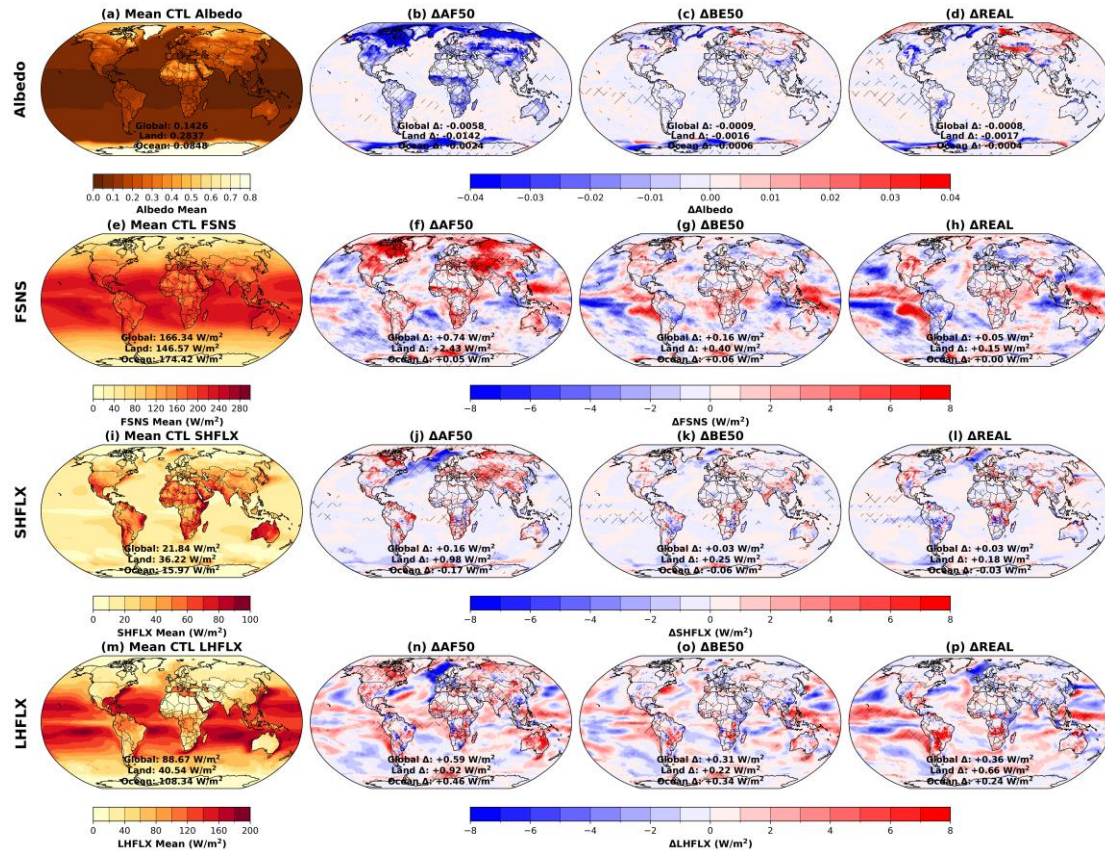


Figure S37. Spatial distributions of multi-year mean (2081–2100) surface albedo and energy fluxes, and their responses to different LUCC scenarios. The rows display four key variables: (a-d) surface albedo, (e-h) net surface shortwave radiation (FSNS), (i-l) sensible heat flux (SHFLX), and (m-p) latent heat flux (LHFLX). The leftmost column presents the climatological mean states under the control simulation (CTL). The subsequent three columns depict the difference for the idealized afforestation (AF50), bioenergy expansion (BE50), and realistic afforestation (REAL) scenarios relative to the CTL. In the difference panels, warm colors (reds) indicate an increase in the respective variable, whereas cool colors (blues) denote a decrease. Cross-hatched regions highlight statistically significant differences at the 95% confidence level, determined by a paired *t*-test ($p < 0.05$). Area-weighted mean values (or mean difference) for the global, land, and ocean domains are inset at the bottom of each panel.

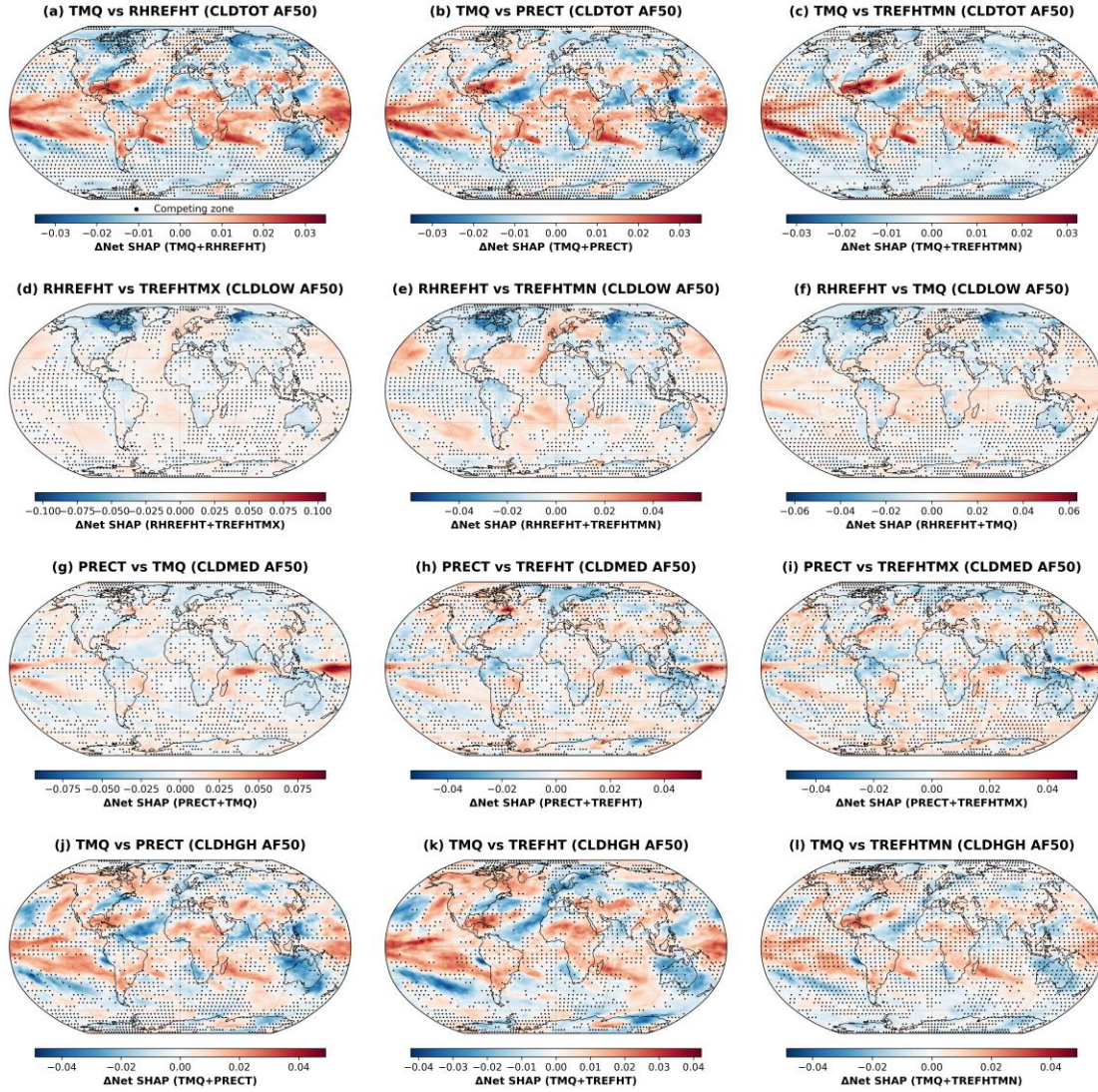


Figure S38. Spatial competitive patterns of key meteorological drivers for cloud cover changes induced by the idealized afforestation (AF50) scenario. Rows from top to bottom correspond to total (CLDTOT), low (CLDLOW), mid (CLDMED), and high (CLDHGH) cloud cover. The background color shading represents the net SHAP contribution (Net SHAP) of the specific paired dominant and secondary drivers. The stippled areas highlight the "competing zones," defined as regions where the two paired drivers have opposite impacts on cloud formation (i.e., one driver promotes cloud cover while the other suppresses it, indicated by opposite signs of their SHAP values).

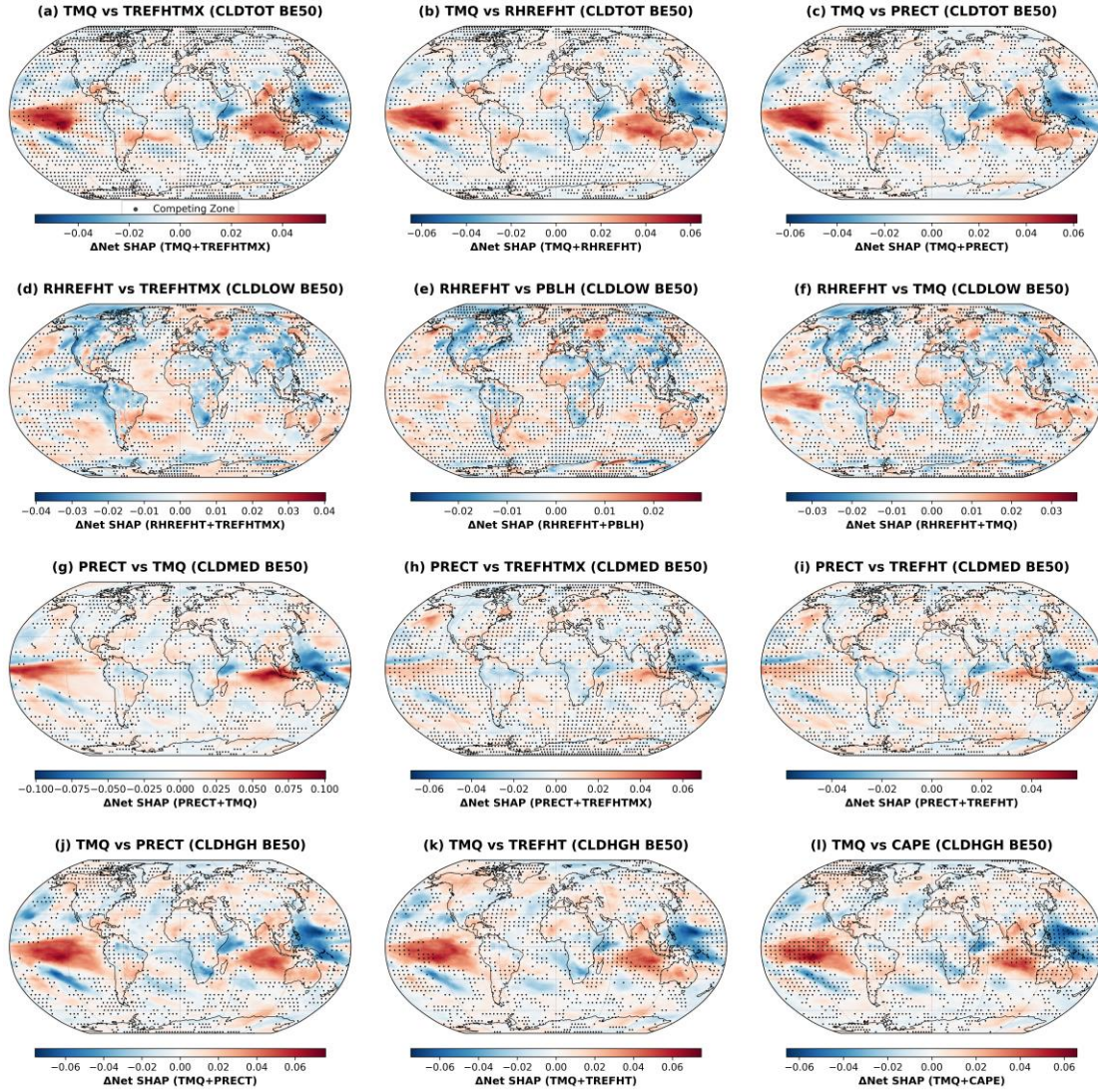


Figure S39. Spatial competitive patterns of key meteorological drivers for cloud cover changes induced by the idealized bioenergy crop expansion (BE50) scenario. Rows from top to bottom correspond to total (CLDTOT), low (CLDLOW), mid (CLDMED), and high (CLDHGH) cloud cover. The background color shading represents the net SHAP contribution (Net SHAP) of the specific paired dominant and secondary drivers. The stippled areas highlight the "competing zones," defined as regions where the two paired drivers have opposite impacts on cloud formation (i.e., one driver promotes cloud cover while the other suppresses it, indicated by opposite signs of their SHAP values).

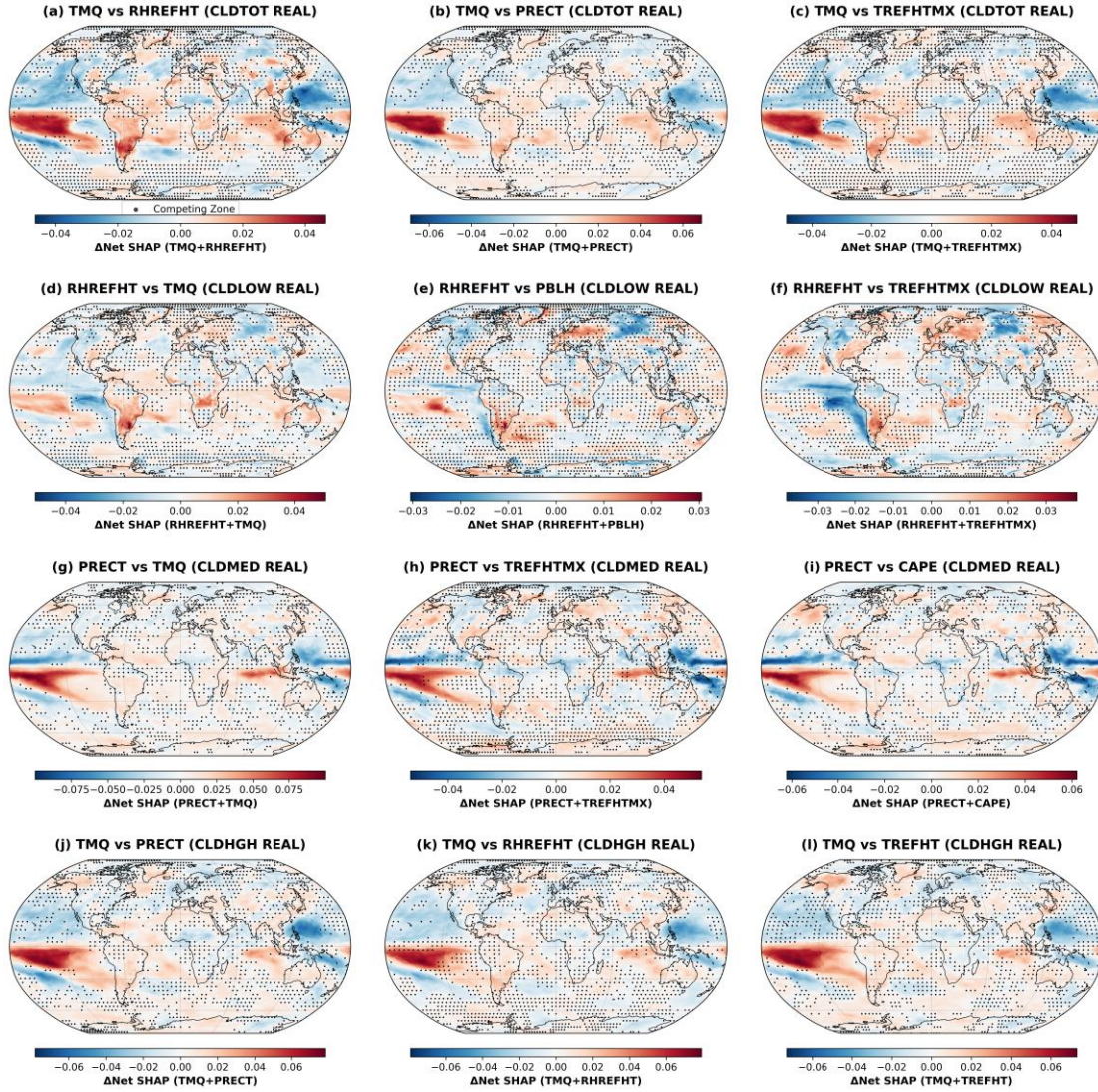


Figure S40. Spatial competitive patterns of key meteorological drivers for cloud cover changes induced by the realistic afforestation (REAL) scenario. Rows from top to bottom correspond to total (CLDTOT), low (CLDLOW), mid (CLDMED), and high (CLDHGH) cloud cover. The background color shading represents the net SHAP contribution (Net SHAP) of the specific paired dominant and secondary drivers. The stippled areas highlight the "competing zones," defined as regions where the two paired drivers have opposite impacts on cloud formation (i.e., one driver promotes cloud cover while the other suppresses it, indicated by opposite signs of their SHAP values).

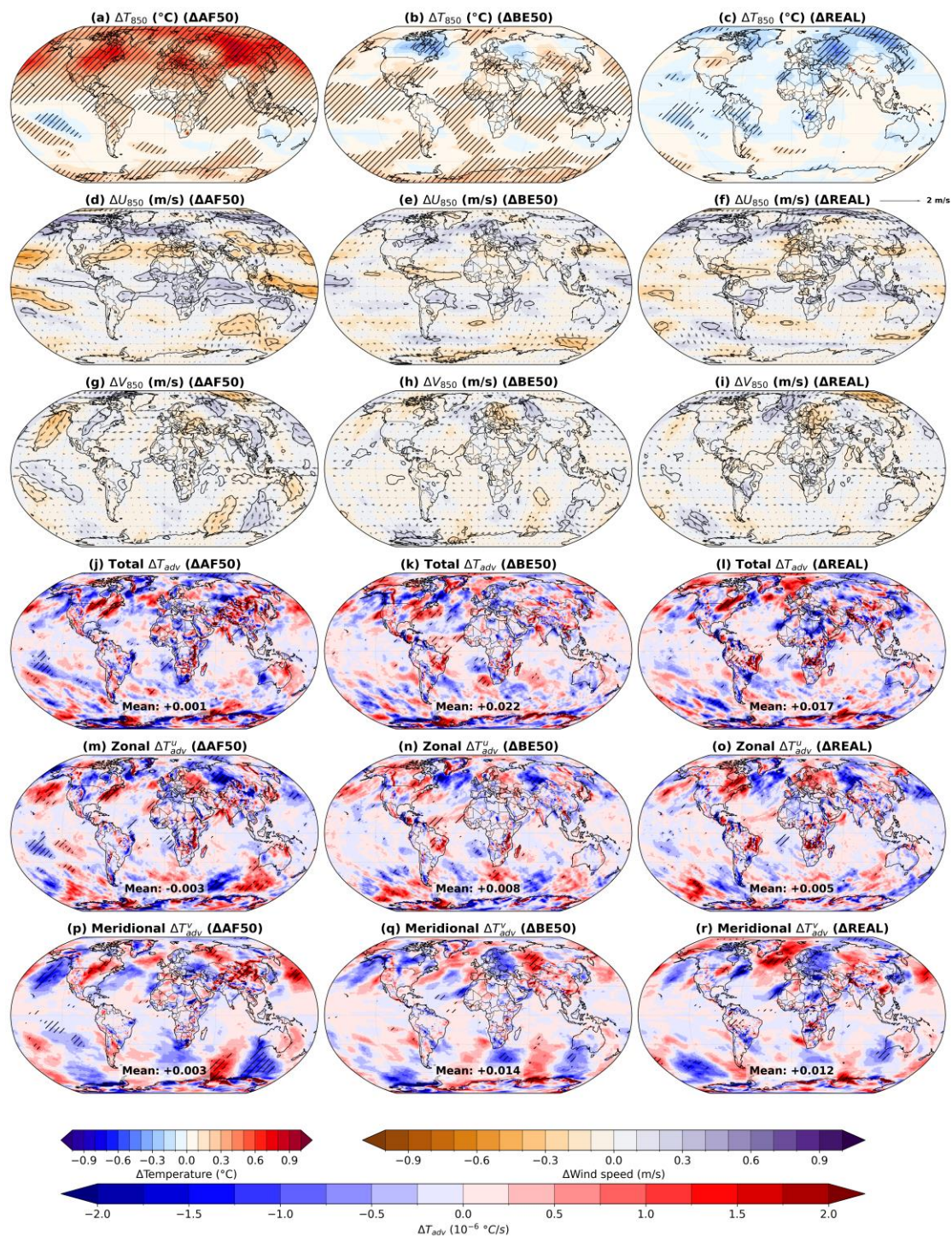


Figure S41. Spatial distributions of lower-tropospheric temperature, horizontal wind, and temperature advection anomalies. Climatological mean differences for the three sensitivity experiments (AF50, BE50, and REAL) relative to the CTL experiment averaged over the 2015–2100 period. Rows from top to bottom display: (1st row) 850 hPa temperature anomalies (ΔT_{850} ; units: $^{\circ}\text{C}$); (2nd row) 850 hPa zonal wind anomalies (ΔU_{850} ; units: m/s); (3rd row) 850 hPa meridional wind anomalies

(ΔV_{850} ; units: m/s); (4th row) total temperature advection anomalies (Total ΔT_{adv} ; units: $^{\circ}\text{C/s}$); (5th row) zonal component of temperature advection anomalies (Zonal ΔT_{adv}^u ; units: $^{\circ}\text{C/s}$); and (6th row) meridional component of temperature advection anomalies (Meridional ΔT_{adv}^v ; units: $^{\circ}\text{C/s}$). In the 2nd and 3rd rows, black vectors represent the combined horizontal wind anomaly field, with a reference vector of 2 m/s shown at the top right of panel (f). Statistical significance at the 95% confidence level ($p < 0.05$) is indicated by hatching in the 1st, 4th, 5th, and 6th rows, and by solid black contours in the 2nd and 3rd rows.

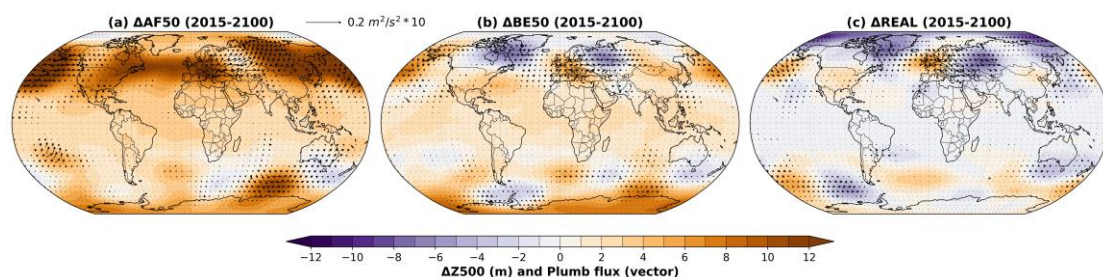


Figure S42. Spatial distributions of the mid-tropospheric geopotential height and horizontal wave activity flux anomalies. Climatological mean differences in the 500 hPa geopotential height (shading; units: m) and the horizontal components of the Plumb wave activity flux (vectors; units: m^2/s^2) for the (a) AF50, (b) BE50, and (c) REAL experiments relative to the CTL experiment averaged over the 2015–2100 period. The brown and purple shadings indicate positive and negative geopotential height anomalies, respectively. The black arrows denote the horizontal propagation direction and intensity of the stationary Rossby wave energy anomalies. The reference vector scale is provided at the top left of panel (a).

Table S1. Optimal parameters of the XGBoost model obtained by Bayesian search

Exps	n_estimato rs	max_dept h	learning_ra te	subsampl e	colsample_bytr ee
AF50- CLDTO T	787	12	0.0484	0.9459	0.6817
AF50- CLDLO W	306	14	0.0498	0.9460	0.9637
AF50- CLDME D	526	12	0.1010	0.8114	0.9082
AF50- CLDHG H	526	13	0.1666	0.8382	1.0
BE50- CLDTO T	494	14	0.0431	0.4538	0.8915
BE50- CLDLO W	305	14	0.0531	0.6602	1.0
BE50- CLDME D	228	13	0.0694	0.8243	0.9313
BE50- CLDHG H	526	12	0.0538	0.6360	1.0
REAL- CLDTO T	787	12	0.0484	0.9459	0.6817
REAL - CLDLO W	526	12	0.0761	0.6833	0.8067
REAL - CLDME D	306	14	0.0446	0.5380	0.8882
REAL - CLDHG H	769	13	0.0564	0.2395	1.0

All hyper-parameters were tuned via Bayesian optimization (Snoek, 2012) within the following ranges: n_estimators 50–800, max_depth 4–14, learning_rate 0.01–0.5, subsample 0.2–1.0 and colsample_bytree 0.2–1.0.

References

- Cleveland, W. S.. Robust Locally Weighted Regression and Smoothing Scatterplots, J. Am. Stat. Assoc., 74(368), 829-836, <https://doi.org/10.1080/01621459.1979.10481038>, 1979.
- Snoek, J., Larochelle, H., Adams, R. P.: Practical Bayesian Optimization of Machine Learning Algorithms, Adv. Neural Inf. Process. Syst., 25, 2960-2968, 2012.