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

Nanjian Liu[1,3], Zhixin Hao[*1,3], Siyou Xia[4], Peng Zhao[2,3]

1. Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

2. State Key Laboratory of Mountain Hazards and Engineering Safety, Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, China

3. University of Chinese Academy of Sciences, Beijing 100049, China

4. Department of Geography, University at Buffalo-SUNY, Buffalo 14261, USA

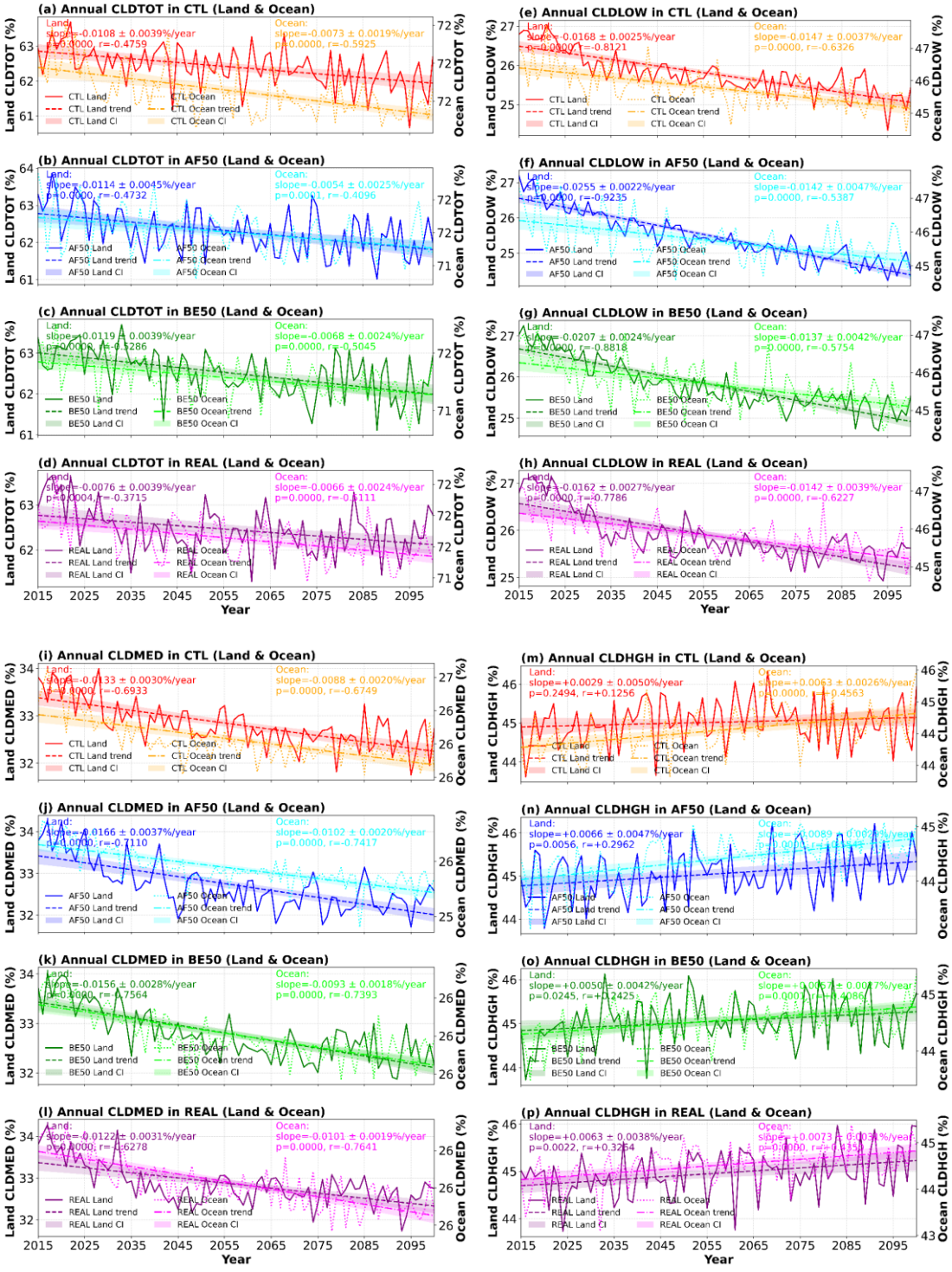

Figure S1. Long-term trends with area weight of cloud cover in land and ocean. The shade indicates 95% confidence interval (CI).

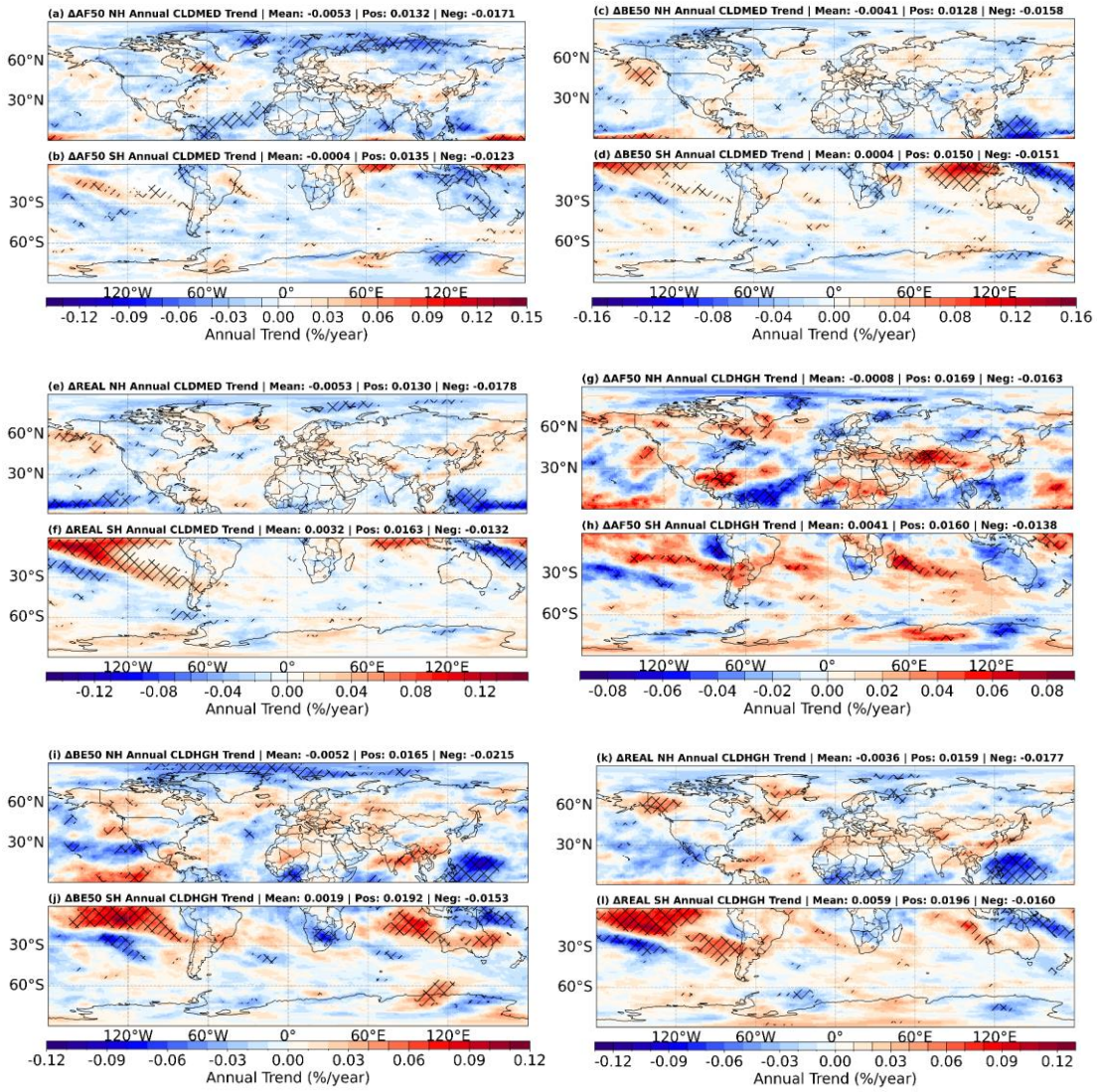

Figure S2. Long-term spatial trends of medium-level cloud cover (CLDMED) and high-level cloud cover (CLDHGH) induced by the expansion of forest and bioenergy in Northern Hemisphere (NH) and Southern Hemisphere (SH). The cross-hatching indicates the grids with a significant trend (p<0.05).

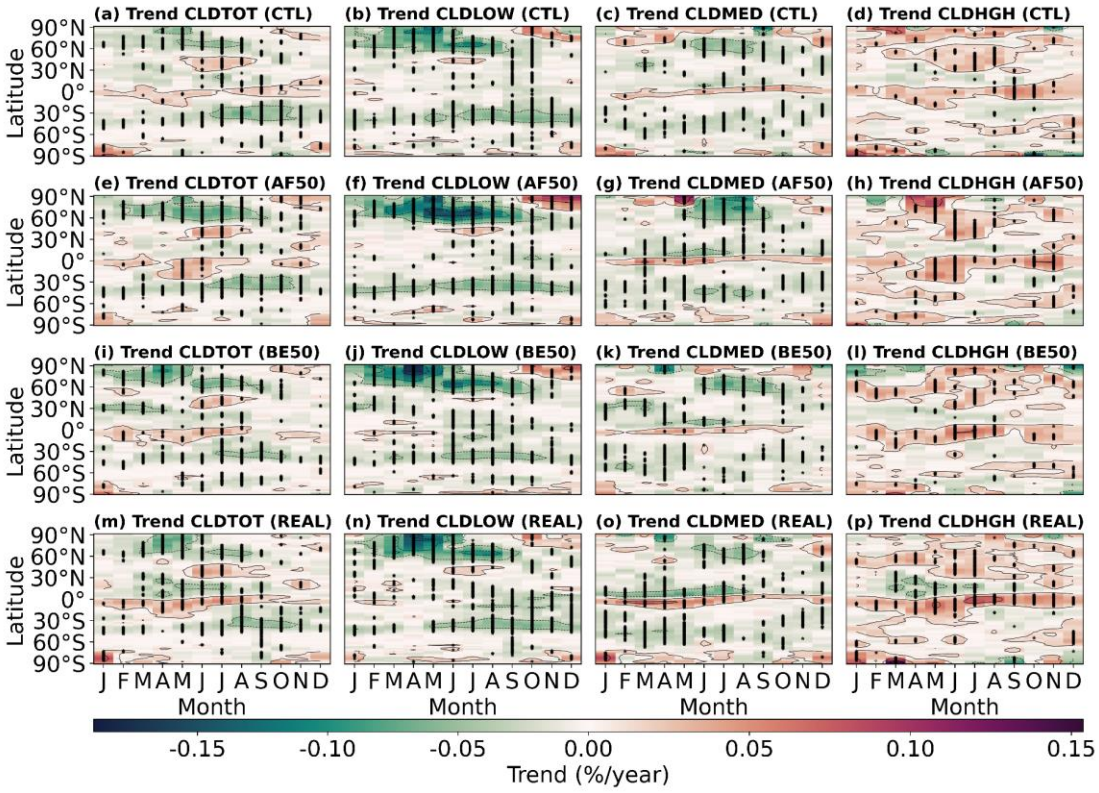

Figure S3. Annual trend in cloud cover across different simulation experiments grouped by month. The black asterisk indicates the grids with a significant trend (p<0.05).

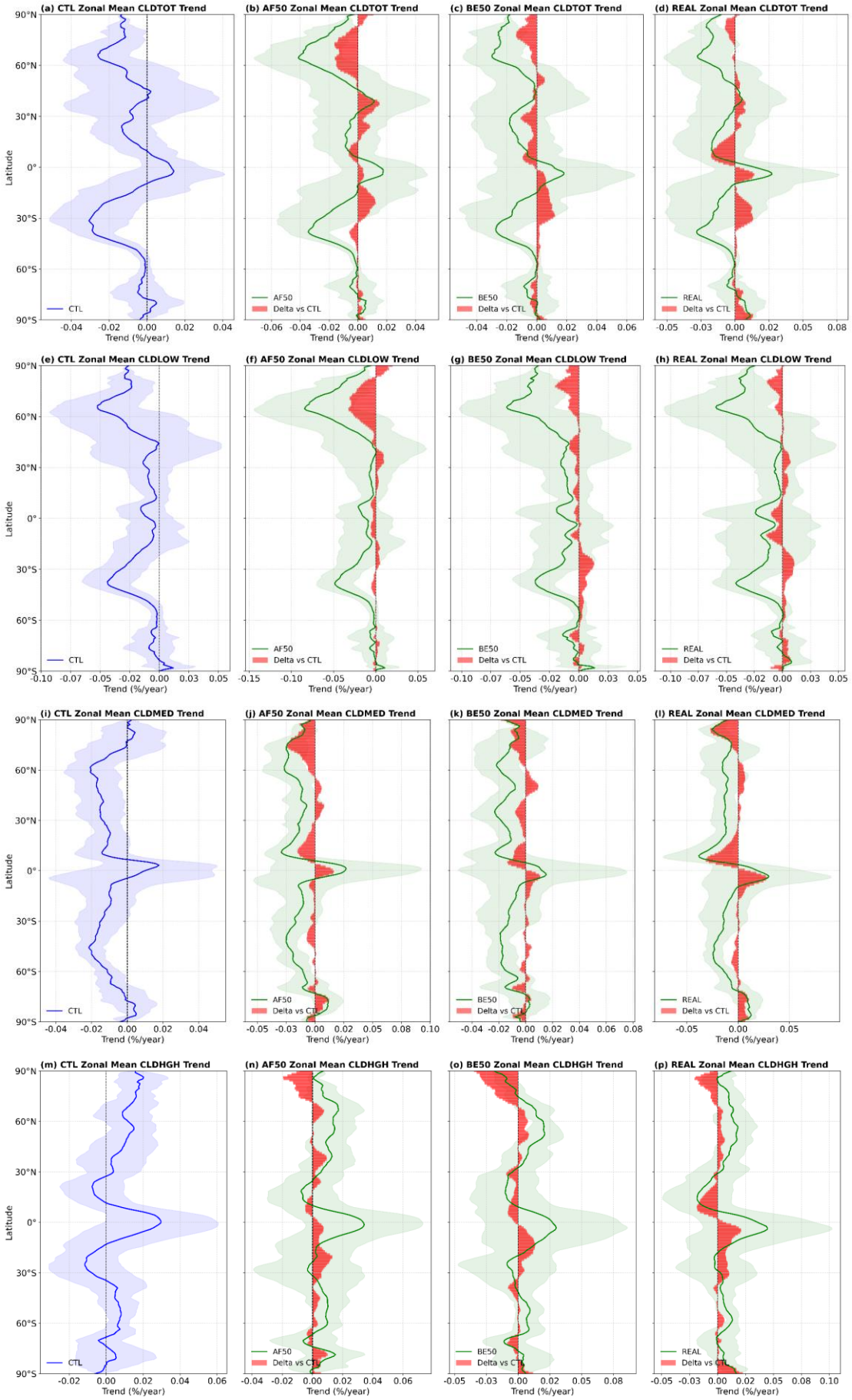
(a) CTL Zonal Mean CLDTOT Trend
(b) AF50 Zonal Mean CLDTOT Trend
(c) BE50 Zonal Mean CLDTOT Trend
(d) REAL Zonal Mean CLDTOT Trend
(e) CTL Zonal Mean CLDLOW Trend
(f) AF50 Zonal Mean CLDLOW Trend
(g) BE50 Zonal Mean CLDLOW Trend
(h) REAL Zonal Mean CLDLOW Trend
(i) CTL Zonal Mean CLDMED Trend
(j) AF50 Zonal Mean CLDMED Trend
(k) BE50 Zonal Mean CLDMED Trend
(l) REAL Zonal Mean CLDMED Trend
(m) CTL Zonal Mean CLDHGH Trend
(n) AF50 Zonal Mean CLDHGH Trend
(o) BE50 Zonal Mean CLDHGH Trend
(p) REAL Zonal Mean CLDHGH Trend
Latitude
90°N
60°N
30°N
0°
30°S
60°S
90°S
CTL
AF50
BE50
REAL
Delta vs CTL
Trend (%/year)

Figure S4. Latitudinal distribution of annual trends in cloud cover across different simulation experiments. The shade indicates standard deviation.

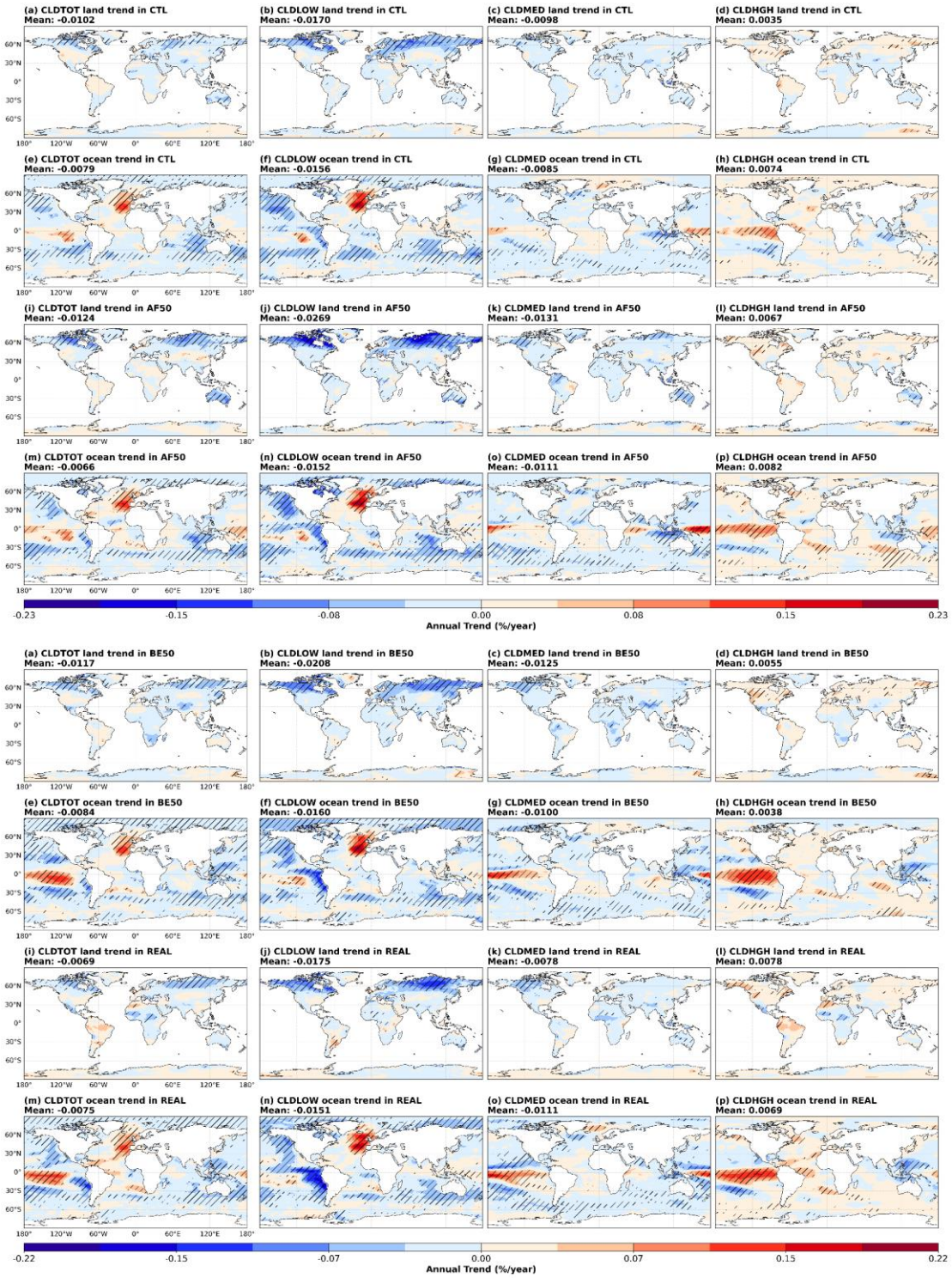

Figure S5. Distribution of cloud cover trend on land and ocean across different simulation experiments. The slash indicates the grids with a significant trend (p<0.05).

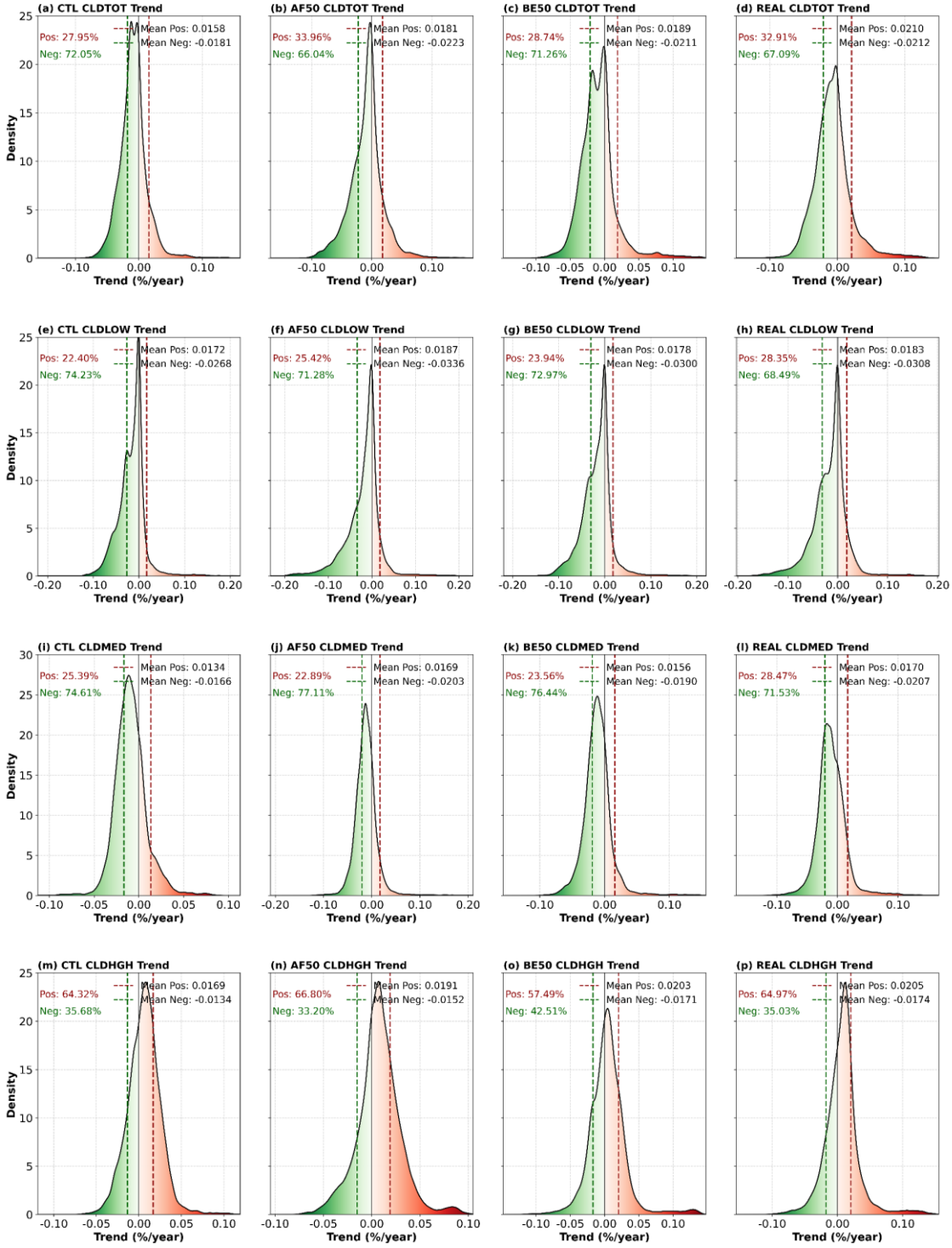

Figure S6. Proportions and average values of cloud cover trends greater than 0 and less than 0 across different simulation experiments.

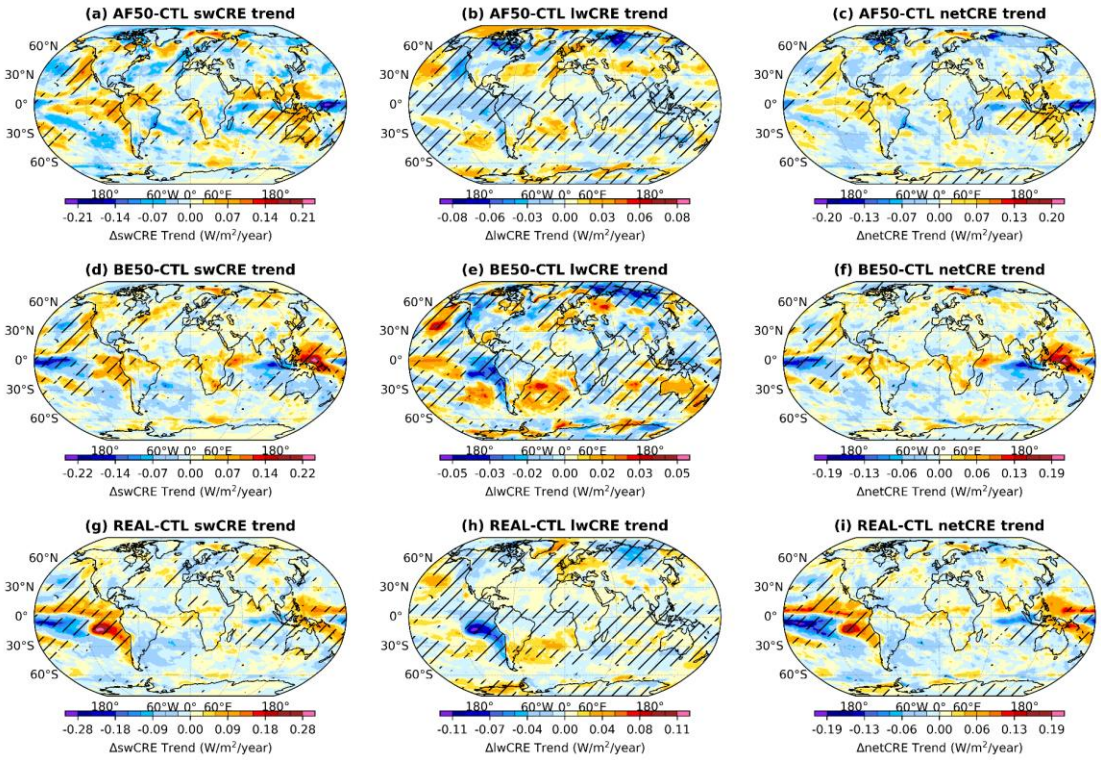

Figure S7. Spatial distribution of annual trends in shortwave cloud radiative effect (CRE), longwave CRE and netCRE induced by different land cover changes. The slash indicates the grids with a significant trend (p<0.05).

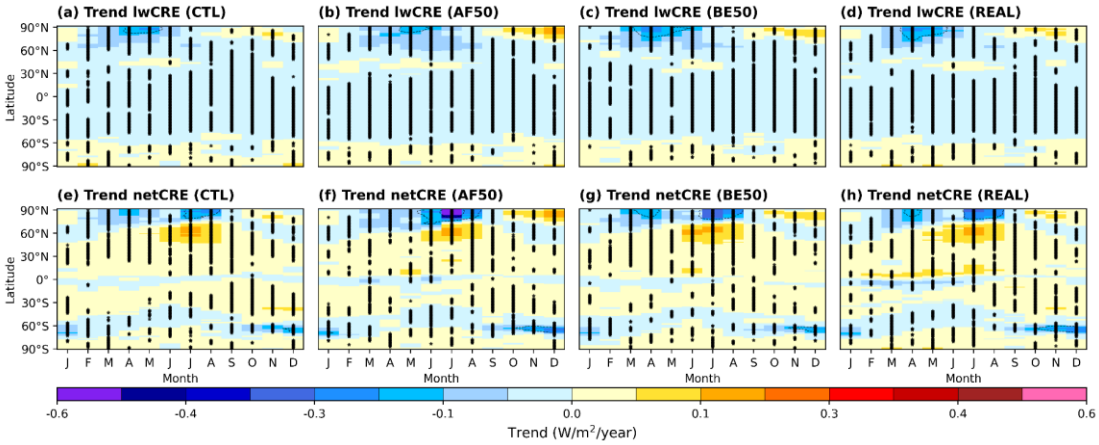

Figure S8. Annual trend in longwave CRE and net CRE across different simulation experiments grouped by month. The black asterisk indicates the grids with a significant trend (p<0.05).

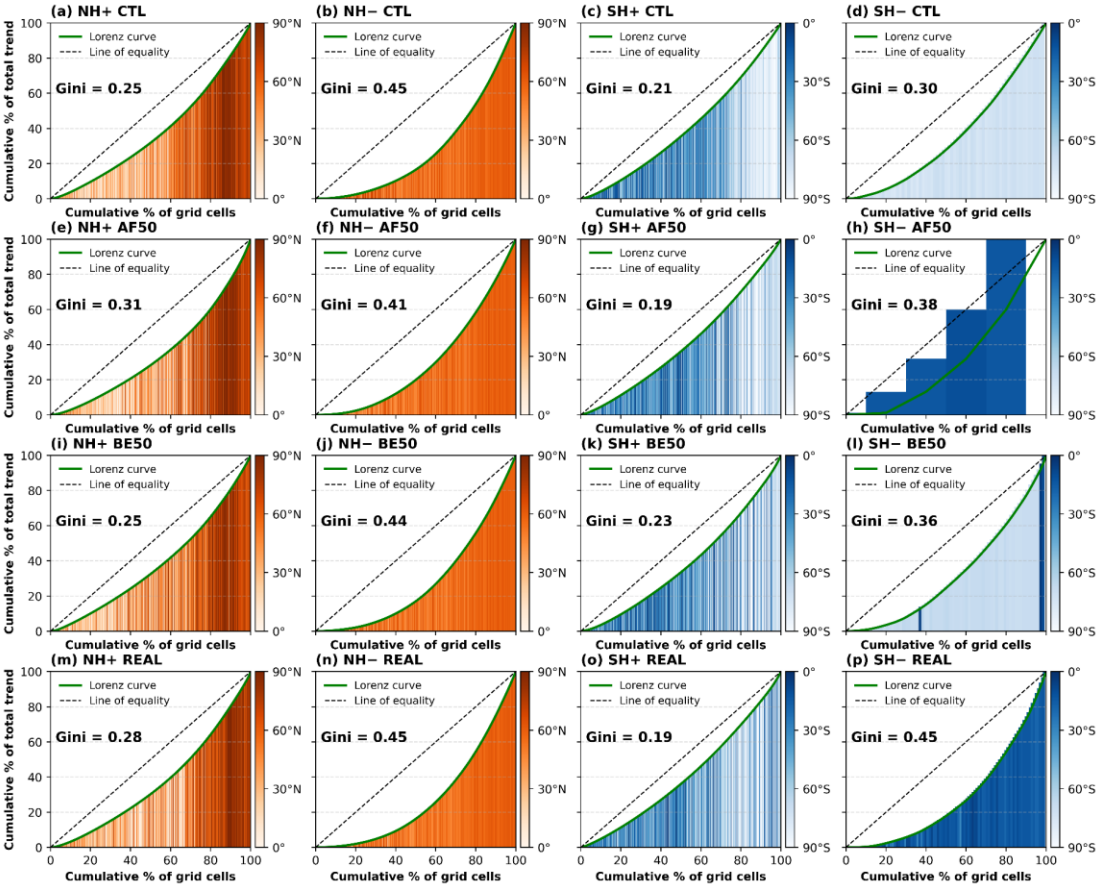

Figure S9. Gini coefficient of temperature at reference height trends in the different simulations

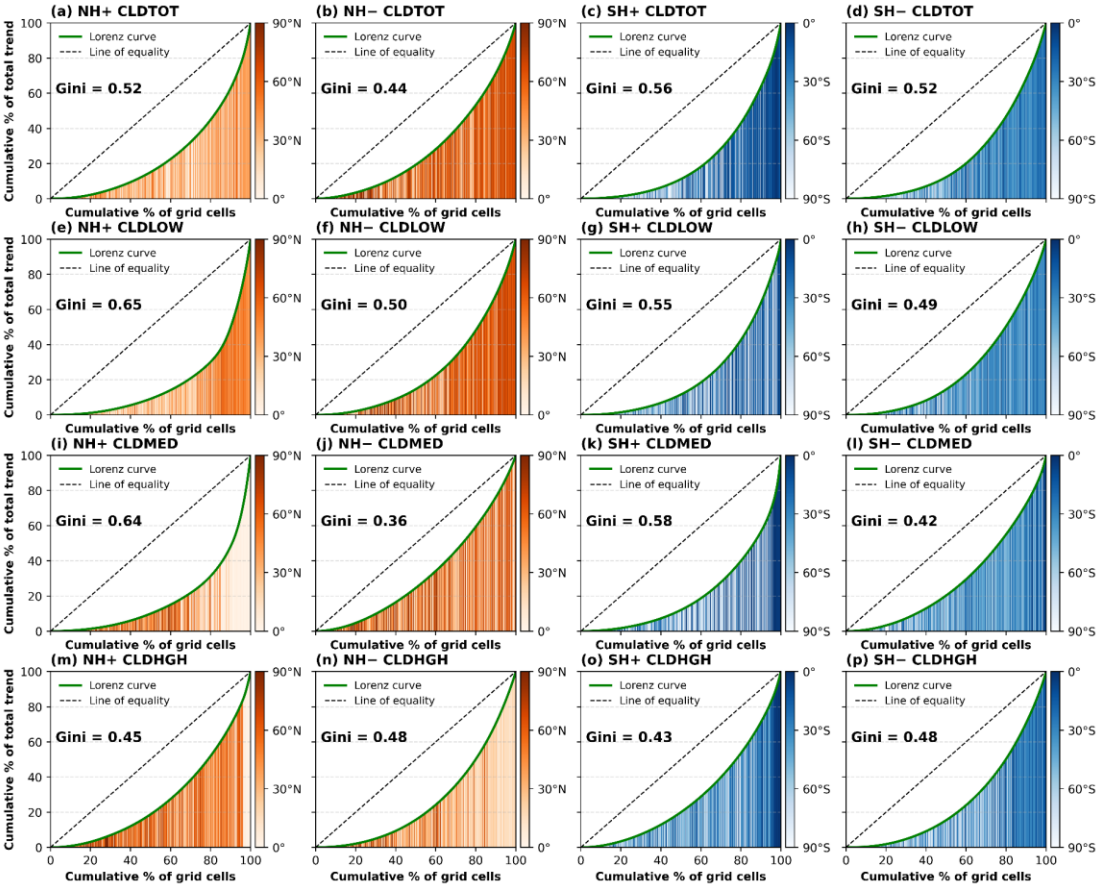

Figure S10. Gini coefficient of cloud cover trends in the AF50 simulations.

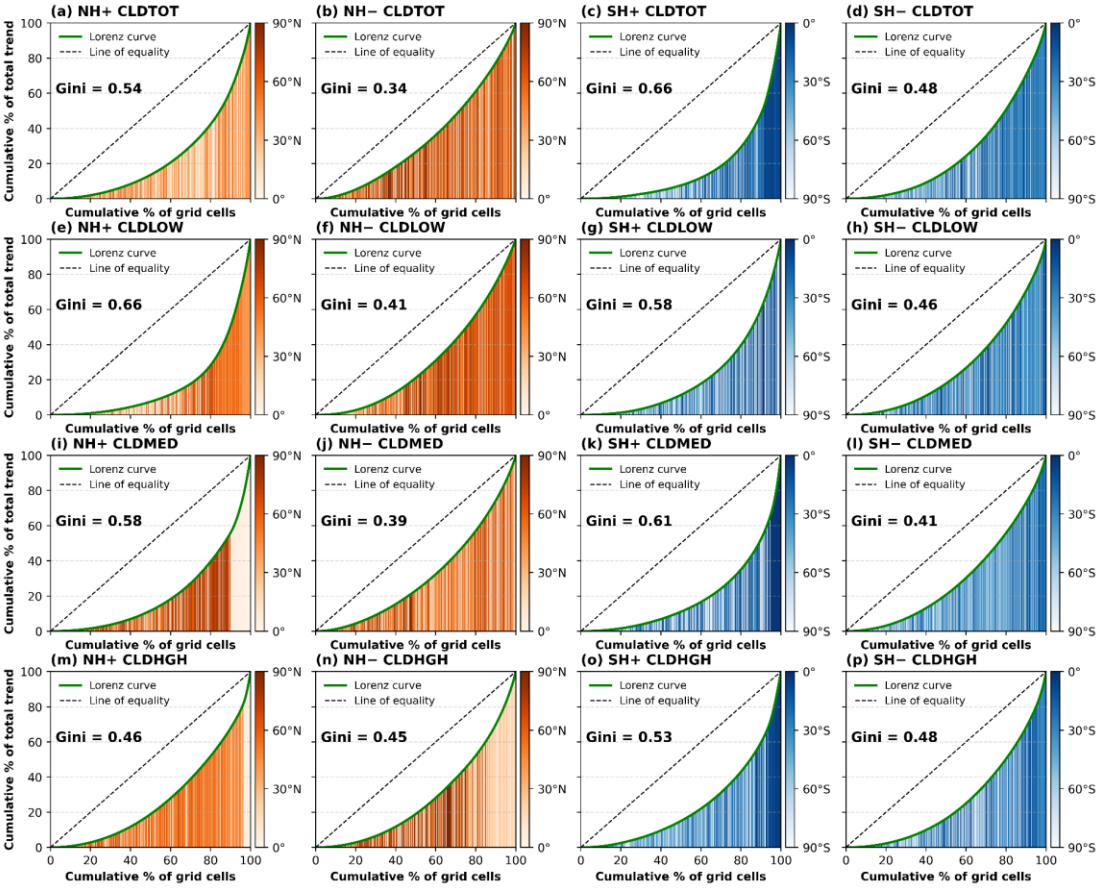

Figure S11. Gini coefficient of cloud cover trends in the BE50 simulations.

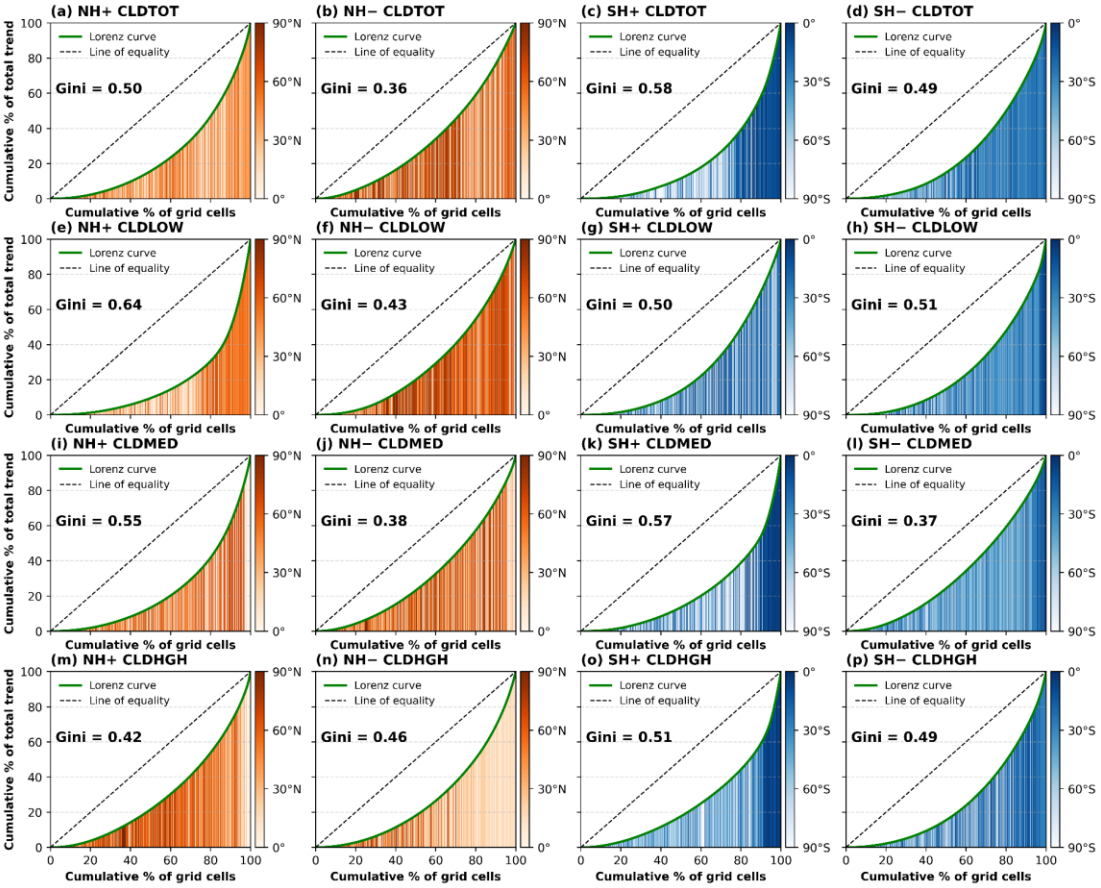

Figure S12. Gini coefficient of cloud cover trends in the REAL simulations.

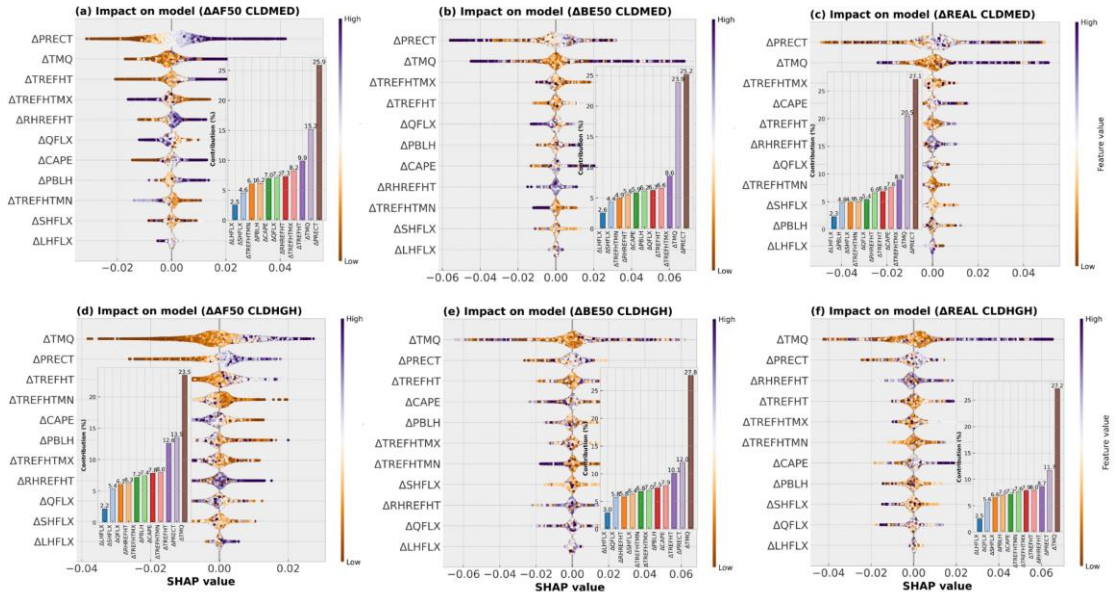

Figure S13. Drivers and contribution of CLDMED and CLDHGH revealed by machine learning response to different land use change. Δ present the difference between sensitivity experiments and control experiment. PRECT, total precipitation (mm/day); TMQ, total precipitable water (kg/m^2); TREFHT, air temperature at reference height (K); TREFHTMX, maximum temperature at reference height (K); RHREFHT, relative humidity at reference height (%); QFLX, evapotranspiration (mm/day); CAPE, convective available potential energy (J/kg); PBLH, planetary boundary layer height (m); TREFHTMN, minimum temperature at reference height (K); SHFLX, sensible heat flux (W/m^2); LHFLX, latent heat flux (W/m^2).

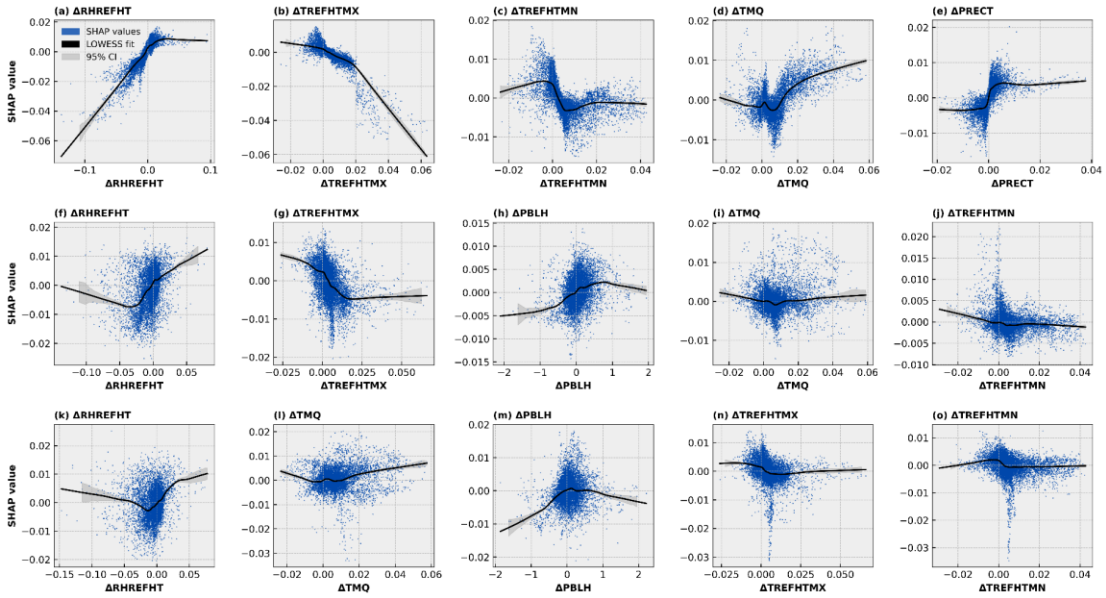

Figure S14. Dependence plots of the top five important variables for low-level cloud cover (CLDLOW), as revealed by the machine learning model. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment. SHAP dependence was visualized using Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland. 1979) with a smoothing fraction of frac=0.3 (using 30% of nearest neighbors for local regression) and it=3 robust iterations to reduce outlier influence, and shaded band indicates 95% confidence interval (CI) of LOWESS regression.

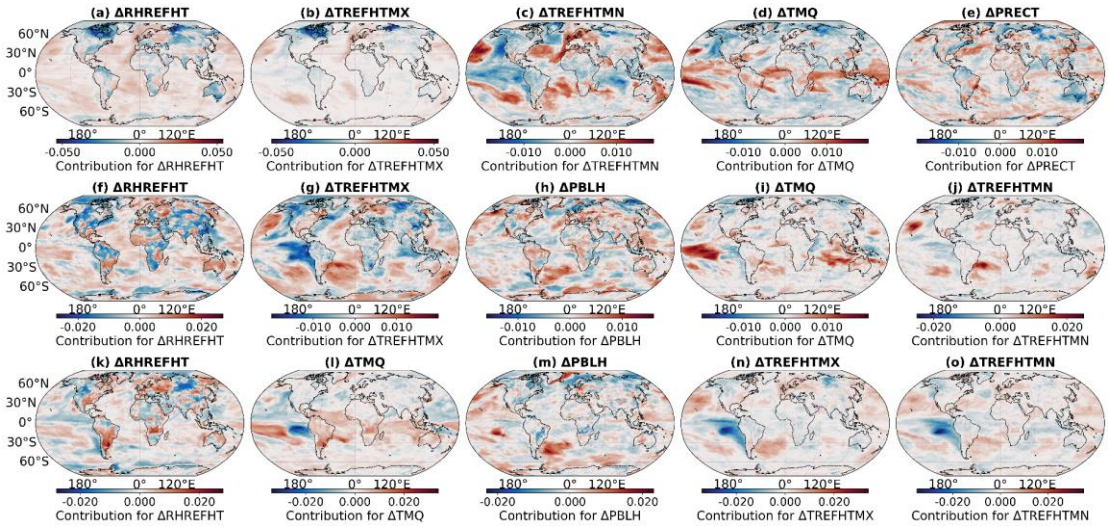

Figure S15. Spatial drivers and contribution of CLDLOW revealed by machine learning response to different land use change. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment.

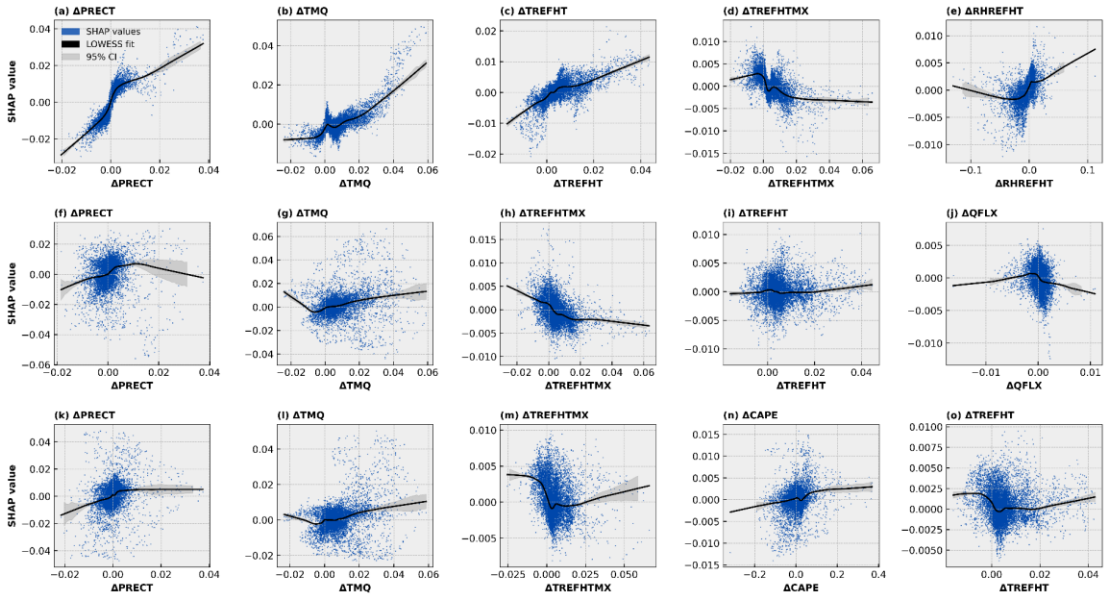

Figure 16. Dependence plots of the top five important variables for CLDMED, as revealed by the machine learning model. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment. SHAP dependence was visualized using Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland. 1979) with a smoothing fraction of frac=0.3 (using 30% of nearest neighbors for local regression) and it=3 robust iterations to reduce outlier influence, and shaded band indicates 95% confidence interval (CI) of LOWESS regression.

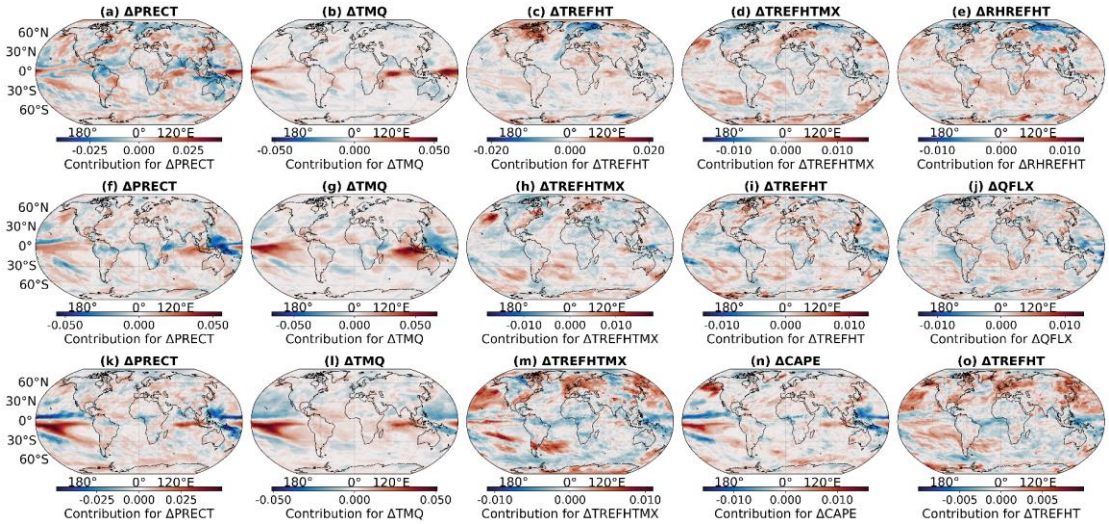

Figure S17. Spatial drivers and contribution of CLDMED revealed by machine learning response to historical land use change. The first row corresponds to the top five drivers of AF50 experiment, the second row to the top five drivers of BE50 experiment, and the third row to the top five drivers of REAL experiment.

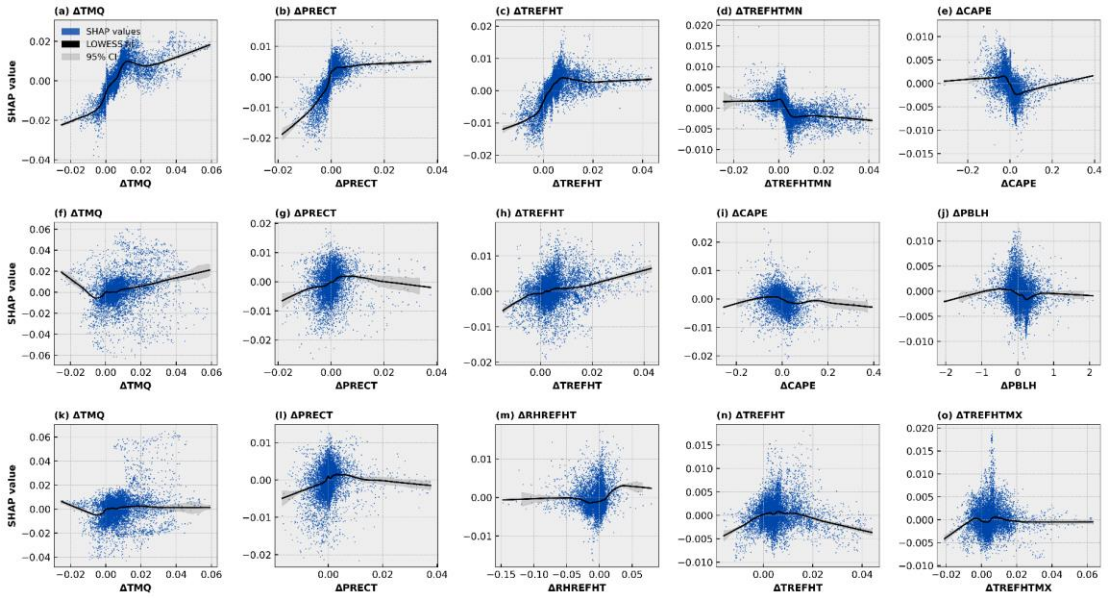

Figure S18. Dependence plots of the top five important variables for CLDHGH, as revealed by the machine learning model. The first row corresponds to the AF50 experiment, the second row to the BE50 experiment, and the third row to the REAL experiment. SHAP dependence was visualized using Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland. 1979) with a smoothing fraction of frac=0.3 (using 30% of nearest neighbors for local regression) and it=3 robust iterations to reduce outlier influence, and shaded band indicates 95% confidence interval (CI) of LOWESS regression.

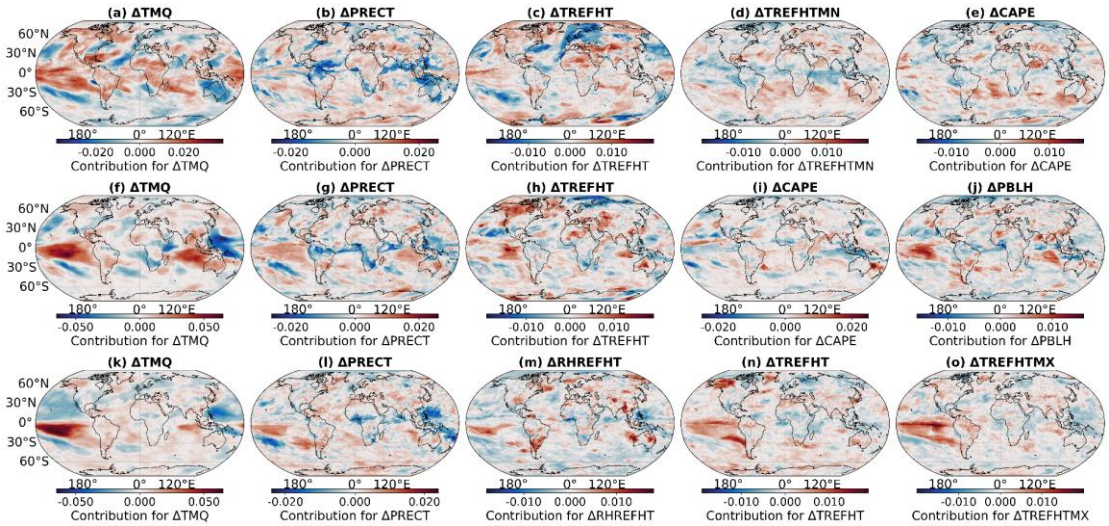

Figure S19. Spatial drivers and contribution of CLDHGH revealed by machine learning response to historical land use change. The first row corresponds to the top five drivers of AF50 experiment, the second row to the top five drivers of BE50 experiment, and the third row to the top five drivers of REAL experiment.

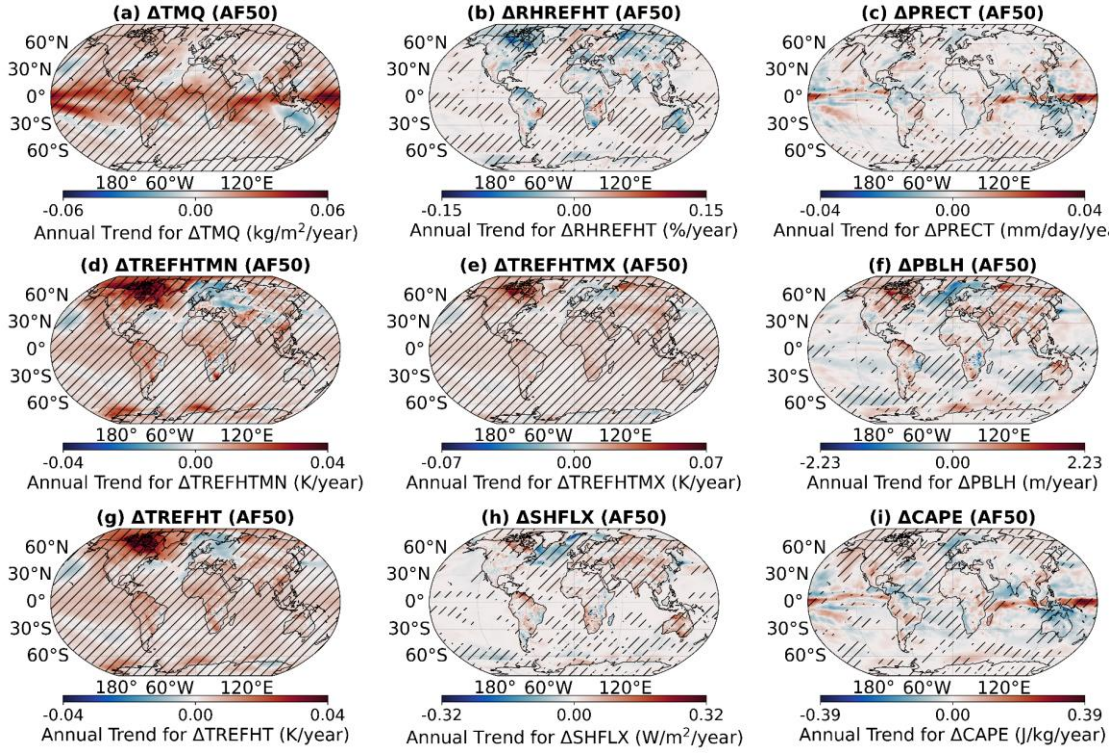

Figure S20. Spatial distribution of annual trends in key atmospheric variables induced by 50% linear afforestation (AF50). The black slash line indicates the grids with a significant trend (p<0.05).

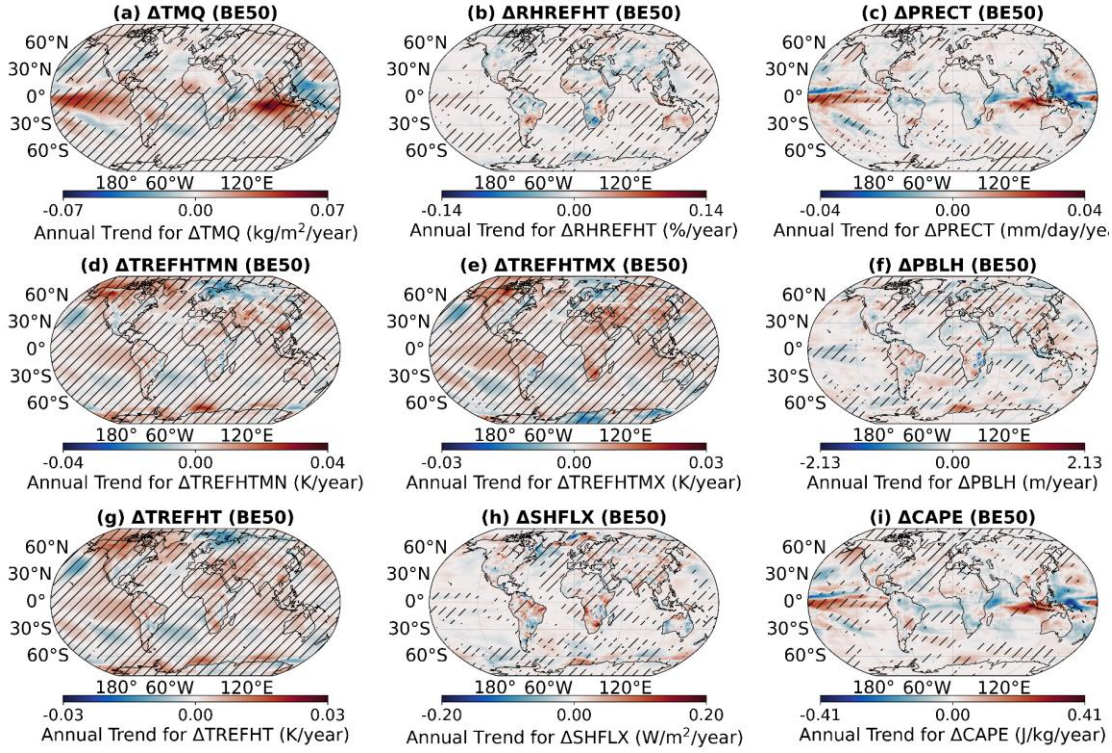

Figure S21. Spatial distribution of annual trend in key atmospheric variables induced by 50% linear bioenergy crop expansion (BE50). The black slash line indicates the grids with a significant trend (p<0.05).

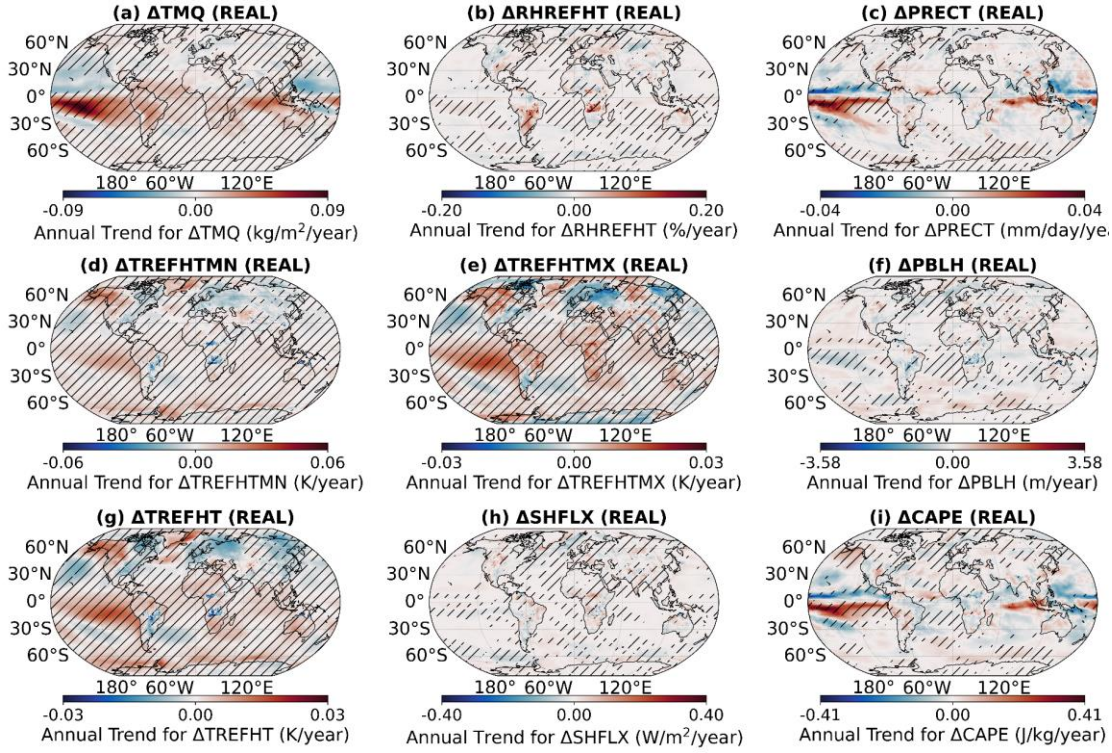

Figure S22. Spatial distribution of annual trend in key atmospheric variables induced by realistic afforestation (REAL). The black slash line indicates the grids with a significant trend (p<0.05).

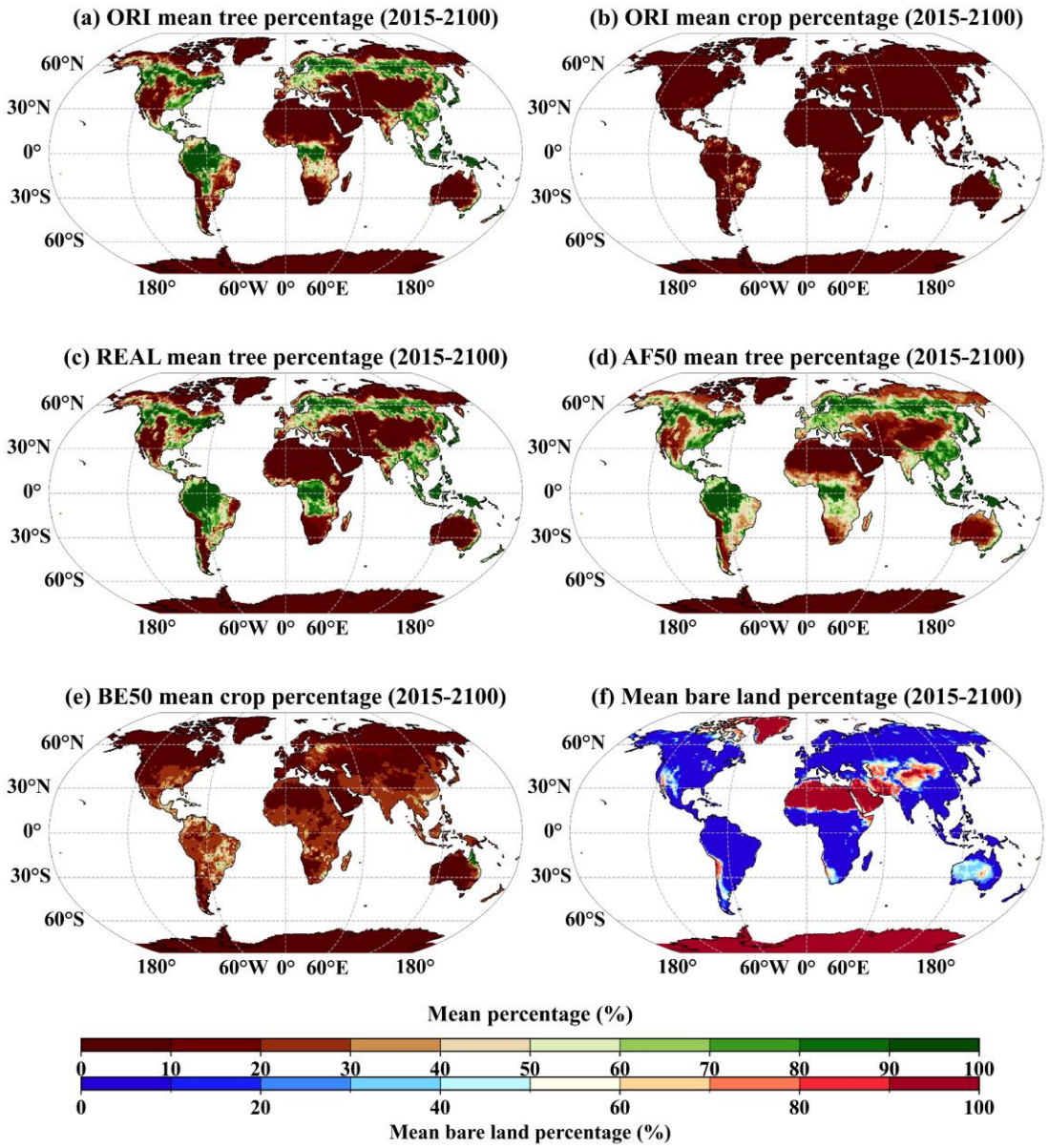

Figure S23. Spatial distribution of forest and bioenergy crops under different land use and land cover scenarios. (a-b) Spatial distribution of forests and bioenergy crops in the original data. (c) Spatial distribution of forests in the realistic afforestation scenario. (d-e) Spatial distribution of forests and bioenergy crops in the 50% linear afforestation and bioenergy crop expansion scenario. (f) Spatial distribution of bare land.

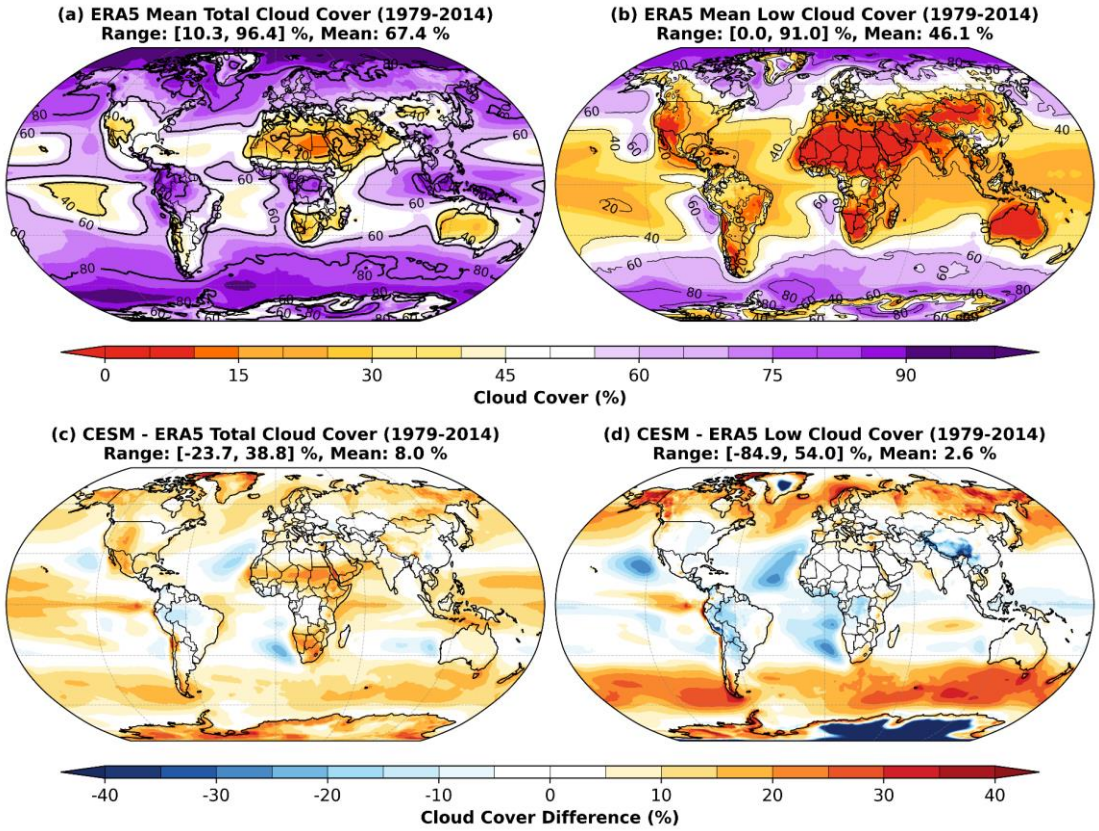

Figure S24. Validation of total cloud cover and low-level cloud cover by the Community Earth System Model (CESM) against ERA5 (ECMWF Reanalysis v5) during the historical period (1979-2014).

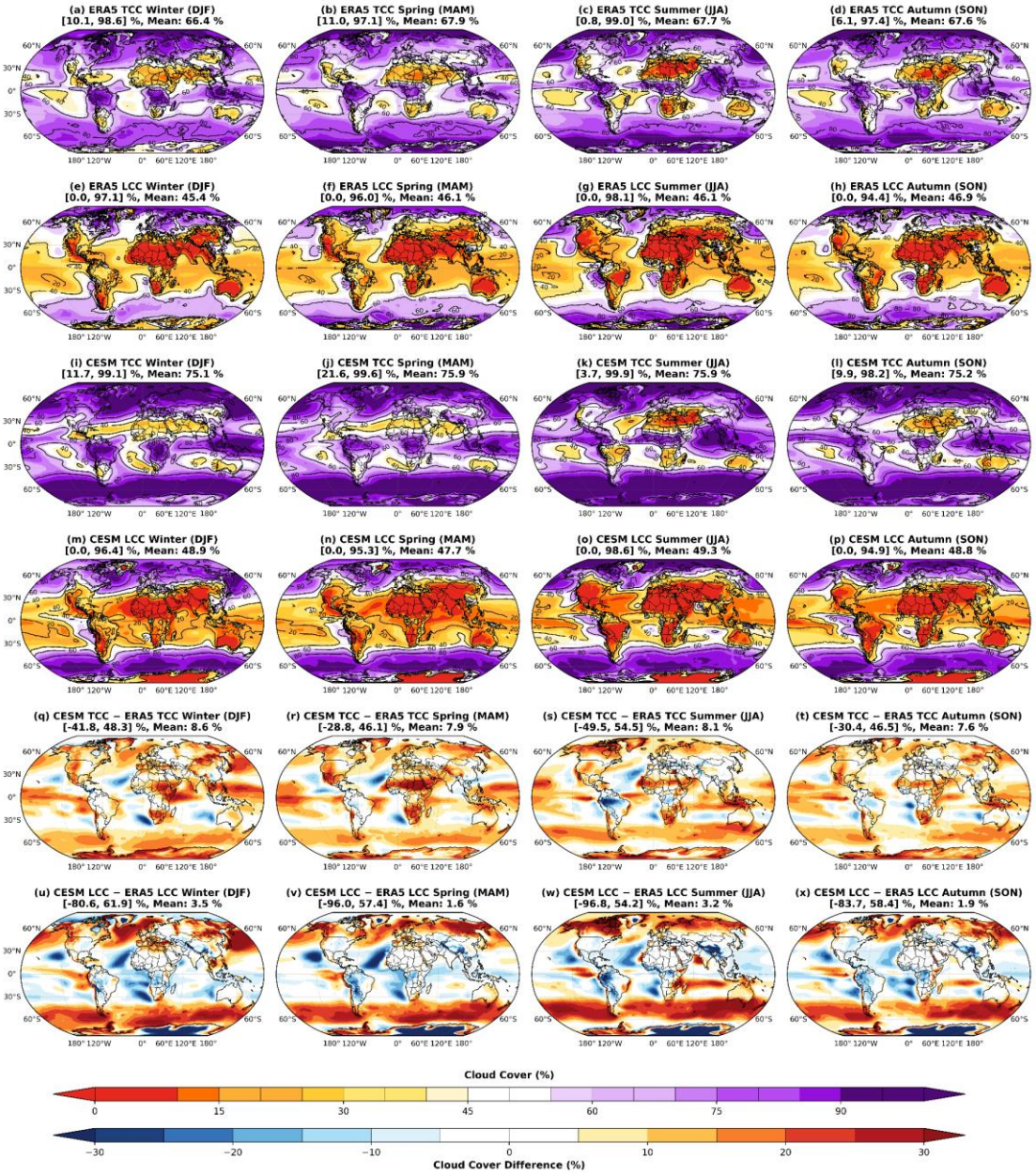

Figure S25. Seasonal validation of total cloud cover and low-level cloud cover by the CESM during the historical period.

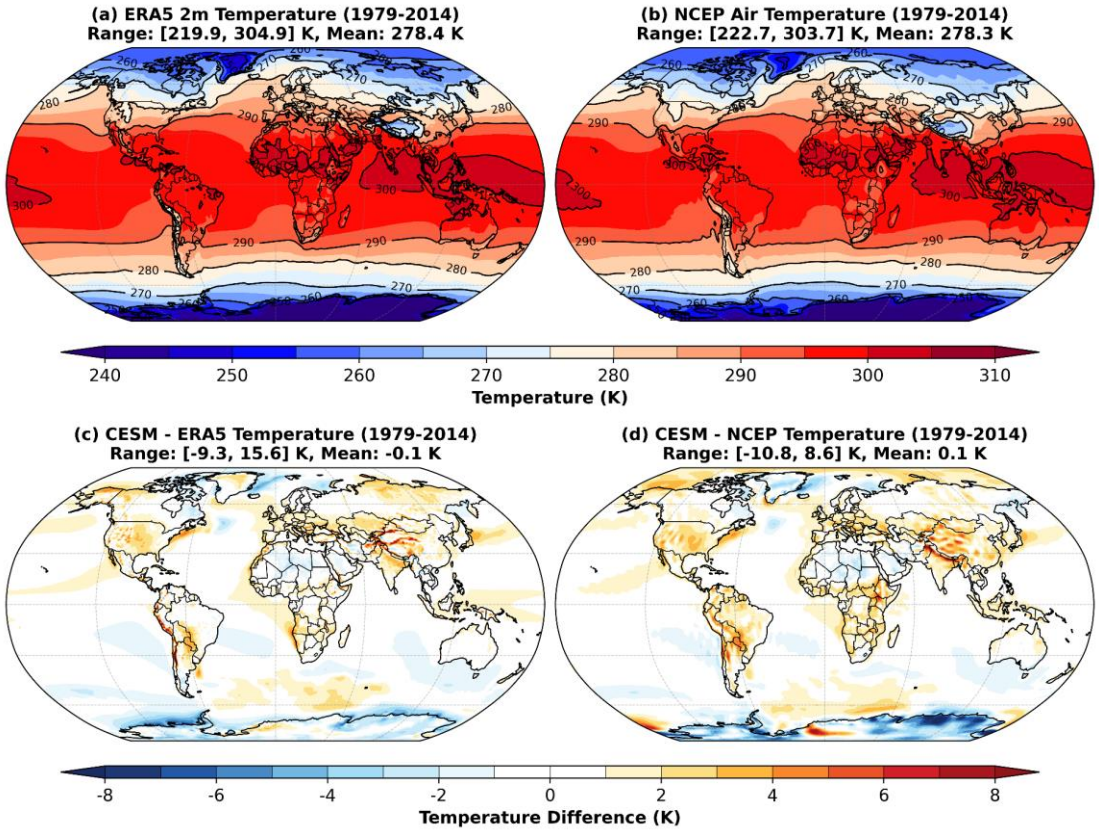

Figure S26. Validation of air temperature from ERA5 and National Centers for Environmental Prediction (NCEP) by the CESM during the historical period (1979-2014).

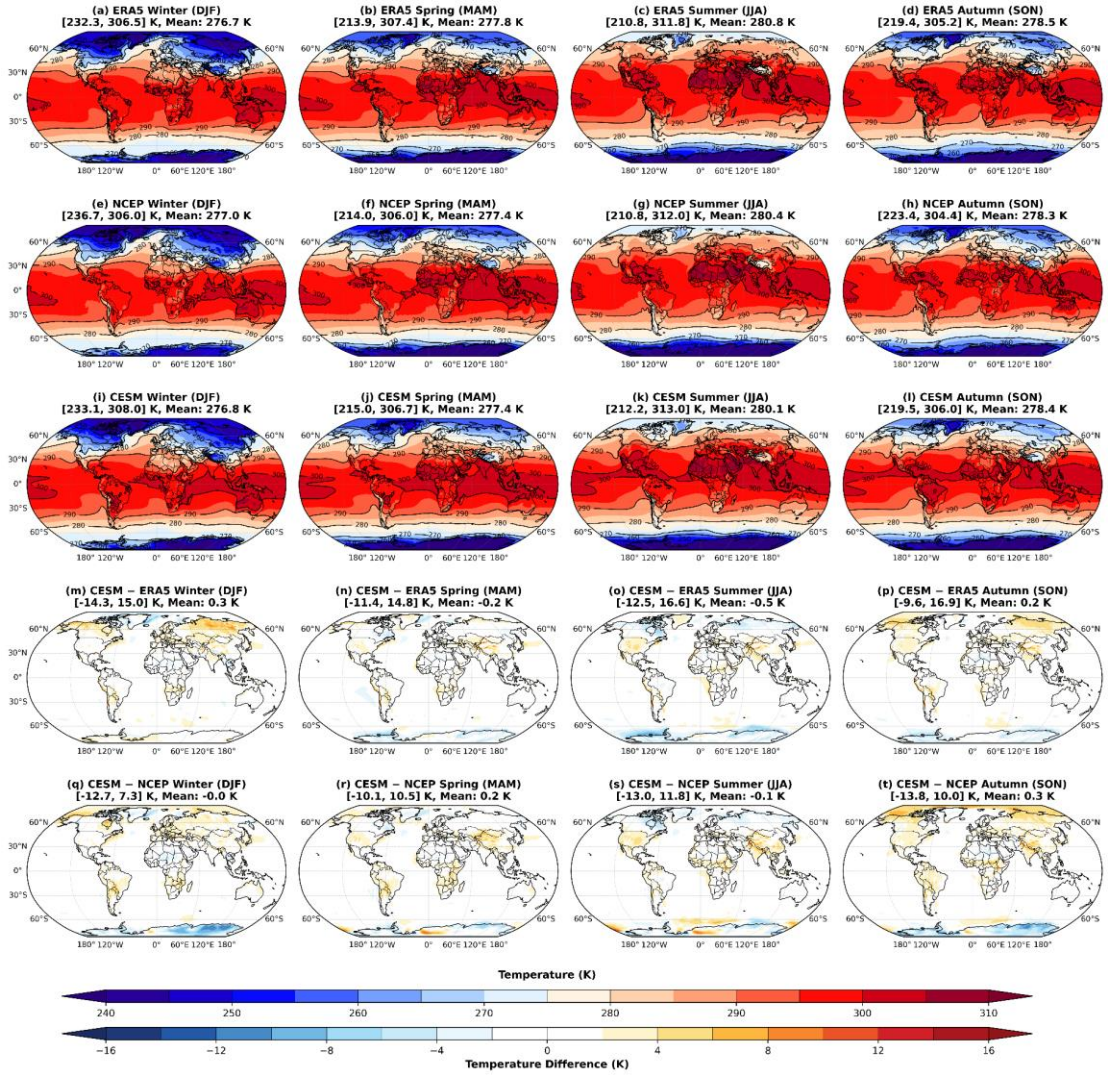

Fig. S27. Seasonal validation of air temperature from ERA5 and National Centers for Environmental Prediction (NCEP) by the CESM during the historical period.

Table 1. Optimal parameters of the XGBoost model obtained by Bayesian search

Exps	n_estimators	max_depth	learning_rate	subsample	colsample_bytree
AF50-CLDTOT	787	12	0.0484	0.9459	0.6817
AF50-CLDLOW	306	14	0.0498	0.9460	0.9637
AF50-CLDMED	526	12	0.1010	0.8114	0.9082
AF50-CLDHGH	526	13	0.1666	0.8382	1.0
BE50-CLDTOT	494	14	0.0431	0.4538	0.8915
BE50-CLDLOW	305	14	0.0531	0.6602	1.0
BE50-CLDMED	228	13	0.0694	0.8243	0.9313
BE50-CLDHGH	526	12	0.0538	0.6360	1.0
REAL-CLDTOT	787	12	0.0484	0.9459	0.6817
REAL - CLDLOW	526	12	0.0761	0.6833	0.8067
REAL - CLDMED	306	14	0.0446	0.5380	0.8882
REAL - CLDHGH	769	13	0.0564	0.2395	1.0

All hyper-parameters were tuned via Bayesian optimisation 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.