Supplement of

Historical Trends and Controlling Factors of Isoprene Emissions in CMIP6 Earth System Models

Thi Nhu Ngoc Do[1], Kengo Sudo[1,3], Akihiko Ito[2,3,4], Louisa Emmons[5], Vaishali Naik[6], Kostas Tsigaridis[7,8], Øyvind Seland[9], Gerd A Folberth[10], Douglas I Kelley[11]

[1]Graduate School of Environment Studies, Nagoya University, Nagoya, 464-8601, Japan
[2]Graduate School of Life and Agricultural Sciences, The University of Tokyo, Tokyo, 113-8657, Japan
[3]Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama, 237-0061, Japan
[4]National Institute for Environmental Studies (NIES), Tsukuba, Ibaraki, 305-8506, Japan
[5]Atmospheric Chemistry Observations and Modeling Lab., National Center for Atmospheric Research, Boulder, CO, USA
[6]NOAA/OAR Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
[7]Center for Climate Systems Research, Columbia University, 2880 Broadway, New York, 10025 NY, USA
[8] NASA Goddard Institute for Space Studies, 2880 Broadway, New York, 10025 NY, USA
[9] Norwegian Meteorological Institute, Oslo, Norway
[10]UK Met Office Hadley Centre, Exeter, UK
[11] UK Centre for Ecology & Hydrology, Wallingford, Oxfordshire, OX10 8BB, UK

Correspondence to: Thi Nhu Ngoc Do (ngocdtn19@gmail.com)

Table S1. Descriptions of symbols and parameters in isoprene emission schemes

Model (Scheme)	γ_{TMP}	γ_{PPFD}	γ_{A}	γ_{SM}	$\gamma_{CO_2\_inhi}$	γ_{CE}
CESM2-WACCM/ NorESM2-LM (G2012)				$\gamma_{SM} = 1\ (\theta > \theta_1)$ $\gamma_{SM} = \frac{\theta - \theta_w}{\Delta\theta_1} (\theta_w < \theta < \theta_1)$ $\gamma_{SM} = 0\ (\theta < \theta_w)$ *Note:* $\gamma_{SM} = 1$ for CESM2-WACCM(G2012) in this study	$C_{CO_2} < 365\ ppm$: $\gamma_{CO_2\_inhi} = 1$ $C_{CO_2} \geq 365\ ppm$: $\gamma_{CO_2\_inhi} = I_{Smax} - \frac{I_{Smax} \cdot (C_i)^h}{(C^*)^h + (C_i)^h}$	0.57
GFDL-ESM4 (G2006)	$E_{opt} \cdot \frac{C_{T2} \cdot exp\ exp\ (C_{T1})\ x}{C_{T2} - C_{T1}(1 - exp\ exp\ (C_{T2}x)\ )}$	$C_p \cdot \frac{\alpha PPFD}{\sqrt{1 + \alpha^2 PPFD^2}}$	$F_{new}A_{new} + F_{grow}A_{grow} + F_{mat}A_{mat} + F_{old}A_{old}$	-	-	$\frac{0.49\ LAI_C}{[(1 + 0.2LAI_C^2)^{0.5}]}$
VISIT(G1997)	$\frac{exp\ (\frac{c_{T1}(T_l - T_s)}{RT_sT_l})}{0.961 + exp\ (\frac{c_{T2}(T_l - T_M)}{RT_sT_l}))}$	$\frac{\alpha_1 c_L Q}{\sqrt{1 + \alpha_1^2 Q^2}}$	1) Evergreen: γ_A = 0.05 (leaf age < 1 month) ~ 1.2 (3 <leaf age < 24 months) 2) Deciduous γ_A = 0.05 (leaf age < 1 month) ~ 1.2 (2 <leaf age < 10 months)	-	-	0.5
GISS-E2.1-G (G1995)	$\frac{exp\ (\frac{c_{T1}(T_l - T_s)}{RT_sT_l})}{1 + exp\ (\frac{c_{T2}(T_l - T_M)}{RT_sT_l}))}$	$\frac{\alpha_1 c_L Q}{\sqrt{1 + \alpha_1^2 Q^2}}$	-	-	-	-
UKESM1-0-LL (P2011)	$(e^{0.1(T_a - T_{st})}; 2.3)$	$\frac{A_J + R_D}{(A_J)_{st} + (R_D)_{st}}$	-	-	$\frac{C_{i_{st}}}{C_i}$	-

E_{opt}: the maximum normalized emission capacity (mol km^{-2} h^{-1})

x : Function of leaf temperature

C_{T1}, C_{T2}: empirical coefficient (95, 230)

c_{T1}, c_{T2} : empirical coefficient; 95000, 230000 J mol^{-1})

T_M : empirical coefficient (314 K)

R: Constant (8.314 J K^{-1} mol^{-1})

T_a : air temperature

T_{st} : temperature at a standard condition (303.15 K)

C_p: function related to past PPFD

α: empirical coefficients associated with past PPFD

$PPFD$: instantaneous photosynthesis photo flux density

α_1, c_L: empirical coefficient (0.0027, 1.066)

Q: flux of PAR (μmol m^{-2} s^{-1})

A_J: leaf level net photosynthesis when RuBP is limiting

R_D : Leaf level dark respiration

"st" indicates that the variables are measured at standard conditions

Fraction for four growth stages: new foliage (F_{new}), growing foliage (F_{grow}), mature foliage (F_{mat}), and old foliage (F_{old}).

A_{new}, A_{grow}, A_{mat}, and A_{old} are the relative emission rates assigned to each canopy fraction depending on PFTs.

The only update of equation parameters from G2006 is the relative emission rates assigned to each compound class in G2012.

θ: soil moisture (m^3 m^{-3})

θ_w : soil moisture threshold below which plants cannot extract water from soil (wilting point, m^3 m^{-3})

$\Delta\theta_1$(=0.06): parameter from Pegoraro et al. (2004) for G2006 and $\Delta\theta_1$(=0.04) for G2012.

I_{Smax} : empirically coefficient (1.344)

C_i: Leaf internal CO_2 concentration, which is estimated as 70% of the ambient CO_2 concentration (C_{CO_2})

C^*: empirically coefficient (585)

h: empirically coefficient (1.4614)

$C_{i_{st}}$: Leaf internal CO_2 concentration at standard conditions

19

Table S2. Summary of emulated sensitivity simulations using random forest regressors for CESM2-WACCM(G2012), NorESM2-LM(G2012) and UKESM1-0-LL(P2011)

Simulation No.	CO_2 conc.	LULCC	Climate		
			Temperature	Shortwave radiation	Precipitation
S0	Fixed in 1850	Fixed in 1850		Climate fixed in 1850	
S1	-	Fixed in 1850			
S2	-	Fixed in 1850		-	
S3	-	-		-	
S4	-	-	Fixed in 1850	-	-
S5	-	-	-	Fixed in 1850	-
S6	-	-	-	-	Fixed in 1850

“-” denotes the variable that varied annually during the simulation period.

Table S3. Summary of emulated sensitivity simulations using random forest regressors for GFDL-ESM4(G2006) and GISS-E2.1-G(G1995)

Simulation No.	LULCC	Climate		
		Temperature	Shortwave radiation	Precipitation
S1′	Fixed in 1850		Climate fixed in 1850	
S2′	Fixed in 1850		-	
S3′	-		-	
S4	-	Fixed in 1850	-	-
S5	-	-	Fixed in 1850	-
S6	-	-	-	Fixed in 1850

“-” denotes the variable that varied annually during the simulation period.

Table S4. Mean annual global and regional isoprene emissions (TgC yr^{-1}) in the present day (2000–2014)

Region	Abbr.	CESM2-WACCM (G2012)	NorESM2-LM (G2012)	GFDL-ESM4 (G2006)	GISS-E2.1-G (G1995)	UKESM1-0-LL (P2011)	VISIT-S3 (G1997)	Ensemble mean	Inter-model spreads	Relative inter-model spreads (%)
Alaska/N.W.Canada	ALA	0.44	0.29	0.38	1.54	1.10	1.39	0.85	0.51	59%
Canada/ Greenl./Icel.	CGI	0.61	0.62	1.12	2.73	2.55	2.20	1.64	0.88	54%
W. North America	WNA	1.83	1.85	2.17	8.34	9.34	3.32	4.48	3.14	70%
C. North America	CNA	3.04	3.57	9.46	18.07	10.83	5.72	8.45	5.16	61%
E. North America	ENA	5.41	5.40	5.05	6.66	5.82	5.26	5.60	0.53	9%
Central America/ Mexico	CAM	18.49	22.84	22.39	12.15	13.94	26.59	19.40	5.09	26%
Amazon	AMZ	175.13	175.74	107.26	98.98	116.55	132.90	134.43	30.77	23%
N.E. Brazil	NEB	13.22	8.65	19.94	19.61	27.67	30.86	19.99	7.66	38%
Coast South America	WSA	4.97	6.41	2.91	2.45	7.34	5.33	4.90	1.75	36%
S.E. South America	SSA	12.29	10.97	18.22	23.51	22.25	16.82	17.34	4.64	27%
N. Europe	NEU	0.51	0.39	0.77	1.32	1.04	1.80	0.97	0.48	50%
C. Europe	CEU	1.59	1.55	5.03	6.16	2.63	3.23	3.36	1.71	51%
S. Europe/ Mediterranean	MED	1.40	1.07	5.12	2.53	3.66	4.62	3.07	1.53	50%
Sahara	SAH	0.10	0.08	0.34	24.89	0.16	0.47	4.34	9.19	212%
W. Africa	WAF	58.40	58.86	45.53	56.15	59.83	68.61	57.90	6.78	12%
E. Africa	EAF	17.04	14.57	20.01	37.82	32.61	32.24	25.71	8.84	34%
S. Africa	SAF	14.69	13.14	14.45	23.55	36.18	23.88	20.98	8.06	38%
N. Asia	NAS	3.13	2.75	5.22	9.92	7.64	6.98	5.94	2.53	43%
W. Asia	WAS	0.87	0.43	2.76	6.12	3.53	1.66	2.56	1.91	75%
C. Asia	CAS	0.71	0.47	2.68	3.36	4.50	1.86	2.26	1.42	63%
Tibetan Plateau	TIB	1.16	1.12	1.29	3.06	5.31	1.83	2.30	1.51	66%
E. Asia	EAS	15.15	16.79	15.17	19.85	21.65	17.50	17.69	2.38	13%
S. Asia	SAS	15.20	14.84	29.41	23.55	15.33	17.82	19.36	5.40	28%
S.E. Asia	SEA	71.68	87.83	79.53	14.69	43.16	64.78	60.28	24.67	41%
N. Australia	NAU	9.71	6.41	12.95	49.75	20.86	27.53	21.20	14.57	69%
S. Australia	SAU	5.27	5.18	5.24	5.55	2.12	4.39	4.63	1.18	25%
	Global	452	462	434	482	478	510	470	24	5%

Note: The ensemble mean is calculated by averaging the isoprene emissions values from the *historical* simulation of CMIP6 models and VISIT-S3(G1997). For each region, inter-model spread is defined as the standard deviation of the values across these models. The relative inter-model spread is then calculated by dividing the standard deviation by the ensemble mean and multiplying by 100%, expressed as a percentage.

Table S5. Plant functional types (PFTs) estimated in CMIP6 models and VISIT(G1997)

Abbreviation	CESM2-WACCM, NorESM2-LM (G2012)	GFDL-ESM4 (G2006)	GISS-E2.1-G (G1995)	UKESM1-0-LL (P2011)	VISIT (G1997)
NeDeBo	Needleleaf deciduous boreal trees	Needleleaf deciduous trees	Needleleaf deciduous trees	Needleleaf deciduous trees	Boreal deciduous forest
NeEvTe	Needleleaf evergreen temperate trees				Temperate needleleaf evergreen forest
NeEvBo	Needleleaf evergreen boreal trees	Needleleaf evergreen trees	Needleleaf evergreen trees	Needleleaf evergreen trees	Boreal evergreen forest
BrEvTr	Broadleaf evergreen tropical trees			Broadleaf evergreen tropical trees	Tropical evergreen forest
BrEvTe	Broadleaf evergreen temperate trees	Broadleaf evergreen trees	Broadleaf evergreen trees	Broadleaf evergreen temperate trees	Temperate broadleaf evergreen forest
BrDeTr	Broadleaf deciduous tropical trees				Tropical deciduous forest
BrDeTe	Broadleaf deciduous temperate trees				Temperate deciduous forest
BrDeBo	Broadleaf deciduous boreal trees	Broadleaf deciduous trees	Cold/Drought broadleaf deciduous trees	Broadleaf deciduous trees	Evergreen/Deciduous mixed forest [a]
EvTeSb	Broadleaf evergreen temperate shrub			Evergreen shrub	Dense shrubland
DeTeSb	Broadleaf deciduous temperate shrub				Open shrubland
DeBoSb	Broadleaf deciduous boreal shrub	Shrubs	Cold/Arid adapted shrubs	Deciduous shrub	Tundra
C4Gr	Warm C4 grass		C4 grass	C4 grass	Desert/Savanna
C3Gr	Cool C3 grass		C3 grass	C3 grass	Grassland/Steppe
ArcticC3Gr	Arctic C3 grass	Grass and others	Arctic C3 grass		Polar Desert/Rock/Ice
Crop	Crops	Crops	C3 crops	C3 crop, C3 pasture, C4 crop, C4 pasture	Crops

[a] temperate/boreal broadleaf and needleleaf trees

Table S6. Emission factors (μgC g_{mass}^{-1} h^{-1}) for each of the plant functional types described in Table S5

Abbreviation	CESM2-WACCM, NorESM2-LM (G2012)	GFDL-ESM4 (G2006)	GISS-E2.1-G (G1995)	UKESM1-0-LL (P2011)	VISIT (G1997)
NeDeBo	0.003	0.0	8.0	8.0	8.0
NeEvTe	1.2	2.0	8.0	8.0	8.0
NeEvBo	4.6	4.0			8.0
BrEvTr	20.6		24.0	24.0	24.0
BrEvTe	29.4			16.0	16.0
BrDeTr	33.3				24.0
BrDeTe	47.6				45.0
BrDeBo	52.4	24.0	24.0/45.0	35.0	8.0
EvTeSb	5.6			20.0	16.0
DeTeSb	19.0				24.0
DeBoSb	19.0	24.0	16.0/24.0	10.0	16.0
C4Gr	1.2		24.0	24.0	24.0
C3Gr	5.0		16.0		16.0
ArcticC3Gr	9.9	0.0	16.0	16.0	16.0
Crop	0.01	0.0	5.0	5.0	5.0

Note:

EFs in CESM2-WACCM(G2012), NorESM2-LM(G2012) are given in units of mass of species per unit area of land surface per unit time (e.g. $\mu g_{isoprene}$ m^{-2} h^{-1}), as opposed to μgC g_{mass}^{-1} h^{-1} used in other models and are denoted hereafter as EF_{area}. Therefore, a conversion must be applied to make these values comparable to the EFs used by other models, which are denoted as EF_{mass} (μgC g_{mass}^{-1} h^{-1}). To convert EF_{area} to EF_{mass}, we applied Eq. (1) of Weber et al. (2023) as shown below:

$$EF_{mass} = EF_{area} \times \frac{1}{LAI_{ref}} \times \frac{1}{SLW} \times \frac{m_{Carbon}}{m_{species}} \times \frac{1}{\gamma_{CE}} \quad \text{(S1)}$$

In that equation, LAI_{ref} is the reference leaf area index used by G2012 scheme (5 $m_{leaf}^{2}\, m_{surface}^{-2}$), SLW is the specific leaf weight ($g_{mass} m_{surface}^{-2}$), the factor $\frac{m_{Carbon}}{m_{species}}$ accounts for the fact that G2012 scheme considers the mass flux of a given species and other land models (e.g., P2011 and VISIT) use the mass flux of carbon, and γ_{CE} is the G2012 canopy environment coefficient (0.57).

As CLM5 land model is incorporated in CESM2-WACCM(G2012), NorESM2-LM(G2012), we use SLW dataset with the CLM5. SLW is inverse from specific leaf area (SLA; $gC\ m_{leaf}^{-2}$) (Ali et al., 2016), and apply a scaling of 2 to convert the mass of carbon to foliar mass.

Emission factors for UKESM1-0-LL(P2011) are derived from Weber et al. (2023), while for G2006 they are derived from the technical description of CLM3 (Oleson et al., 2004). Emission factors in GISS-E2.1-G(G1995) are assigned from Guenther et al. (1995) for corresponding PFTs, while values for VISIT(G1997) were derived from Lathìere et al. (2006).

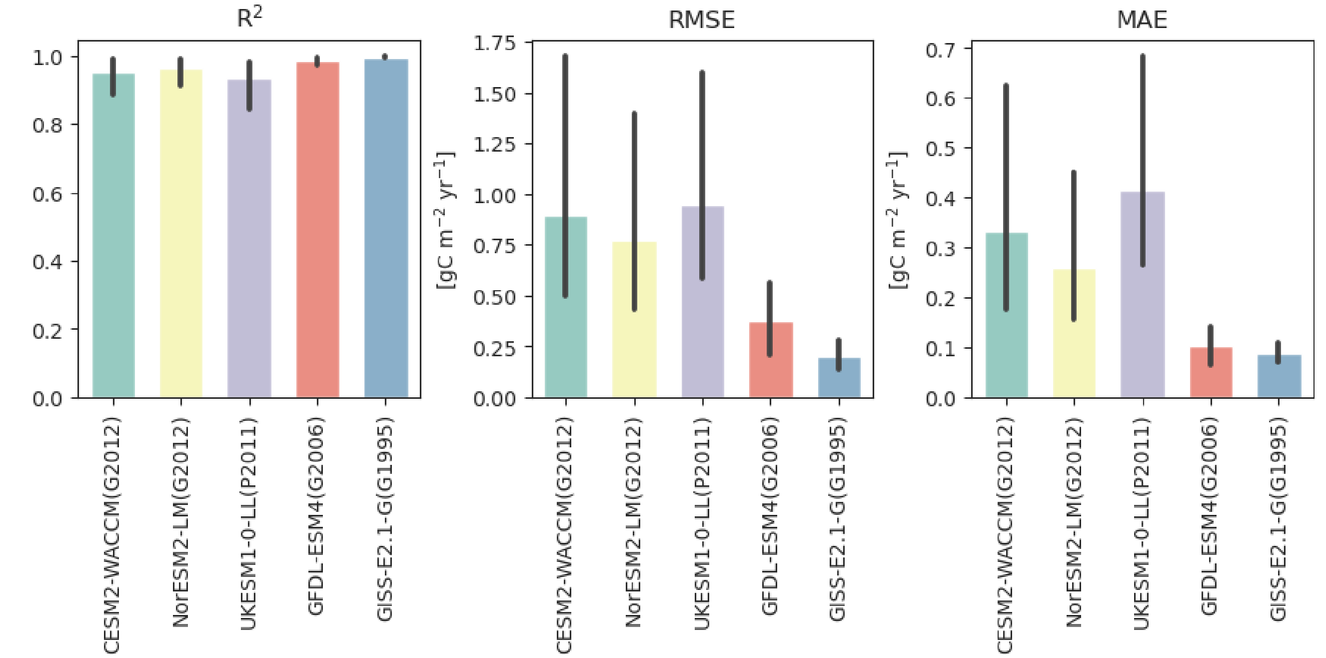


Figure S1. Three-fold cross-validation of random forest regressor applied for each CMIP6 model's data. R^2 is coefficient of determination, RMSE is Root Mean Squared Error, and MAE is Mean Absolute Error.

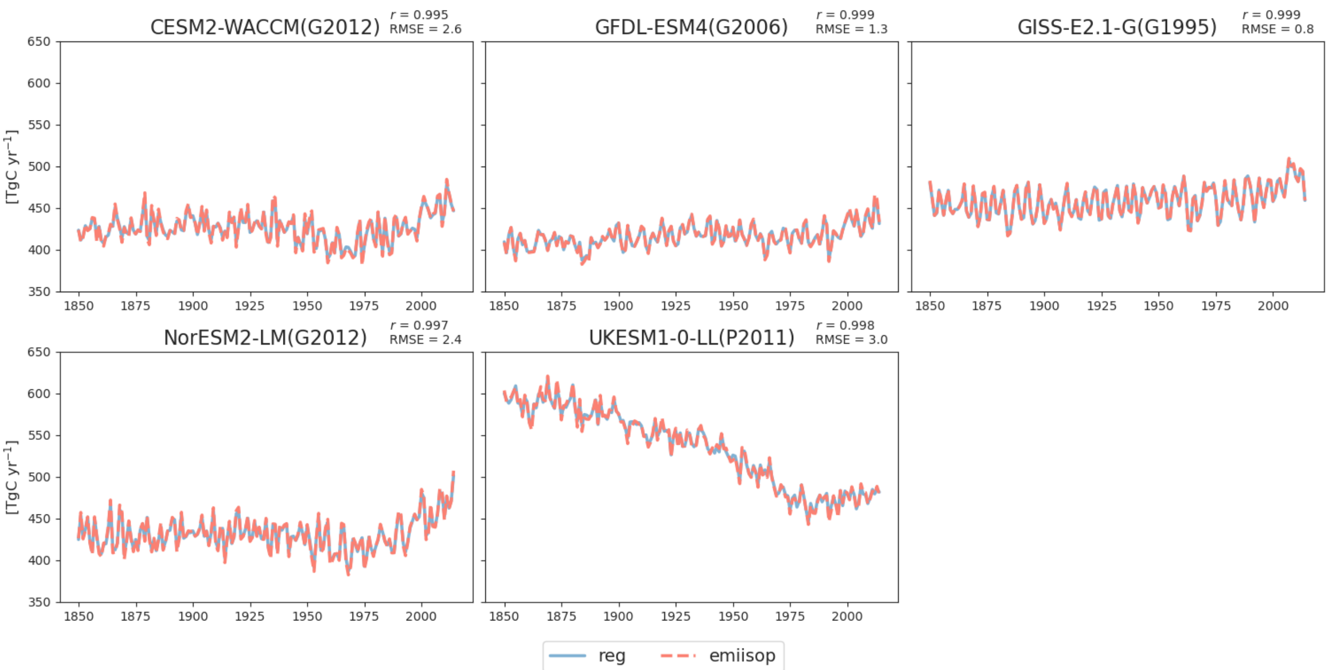
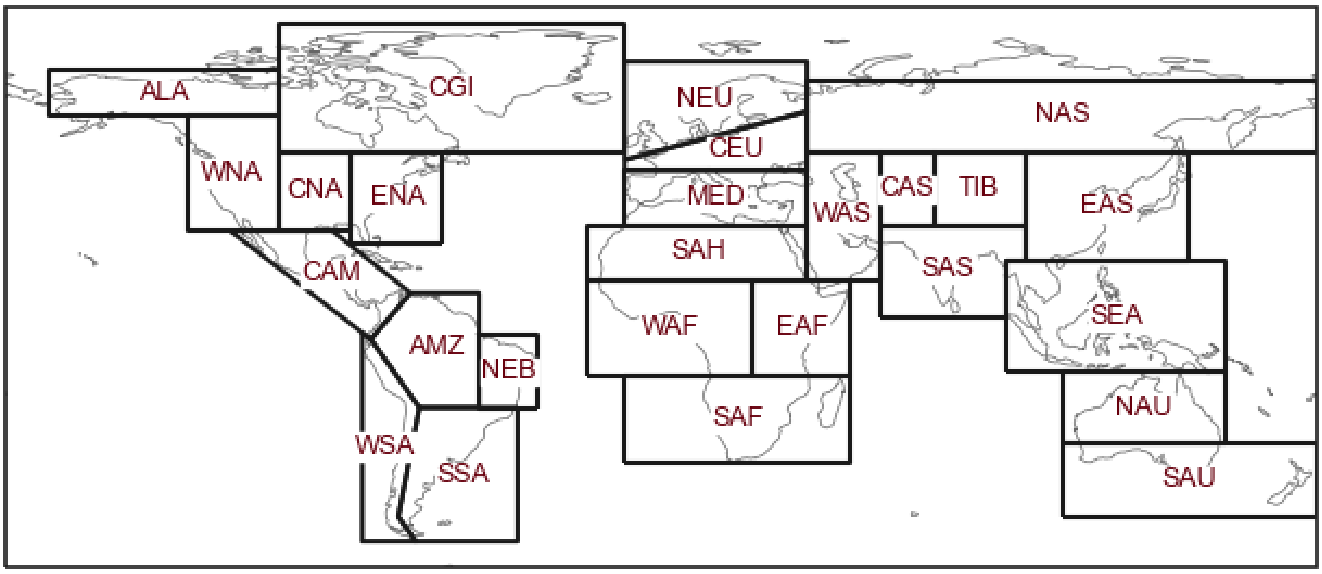


Figure S2. Global annual isoprene emission comparison between *historical* simulation from each CMIP6 model (emiisop) and estimation from its random forest regressor (reg). *r* is Pearson correlation and RMSE is Root Mean Square Error.



ALA:	Alaska/Northwest Canada	NEB:	North-East Brazil	WAF:	Western Africa	EAS:	East Asia
CGI:	Eastern Canada/Greenland/Iceland	WSA:	West Coast South America	EAF:	Eastern Africa	SAS:	South Asia
WNA:	Western North America	SSA:	South-eastern South America	SAF:	Southern Africa	SEA:	Southeast Asia
CAN:	Central North America	NEU:	Northern Europe	NAS:	Northern Asia	NAU:	Northern Australia
ENA:	Eastern North America	CEU:	Central Europe	WAS:	Western Asia		
CAM:	Central America/ Mexico	MED:	Southern Europe/ the Mediterranean	CAS:	Central Asia	SAU:	Southern Australia/ New Zealand
AMZ:	Amazon	SAH:	Sahara	TIB:	Tibetan Plateau		

Figure S3. The 26 SREX regions are defined by the IPCC Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (Seneviratne et al., 2012).

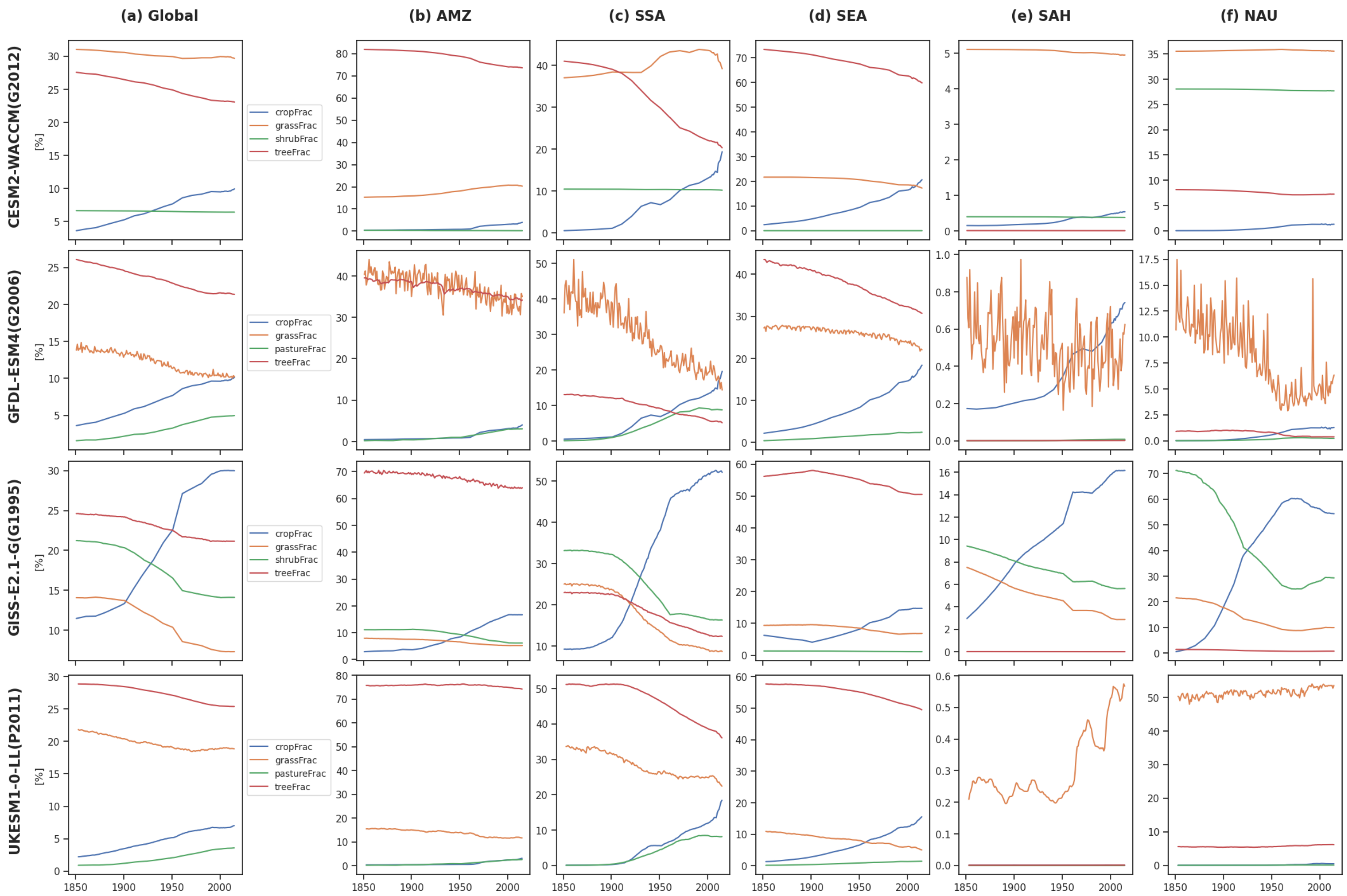


Figure S4. Annual fraction (%) of four plant functional types (PFTs): tree, grass, shrub/pasture, and crop during 1850–2014. Panels (a–f) show changes at (a) the global scale and in regions: (b) Amazon (AMZ), (c) Southeastern South America (SSA), (d) Southeast Asia (SEA), (e) Sahara (SAH), and (f) North Australia (NAU). Data from CESM2-WACCM(G2012), GFDL-ESM4(G2006), GISS-E2.1-G(G1995), and UKESM1-0-LL(P2011). NorESM2-LM(G2012) uses the same land component as CESM2-WACCM(G2012).

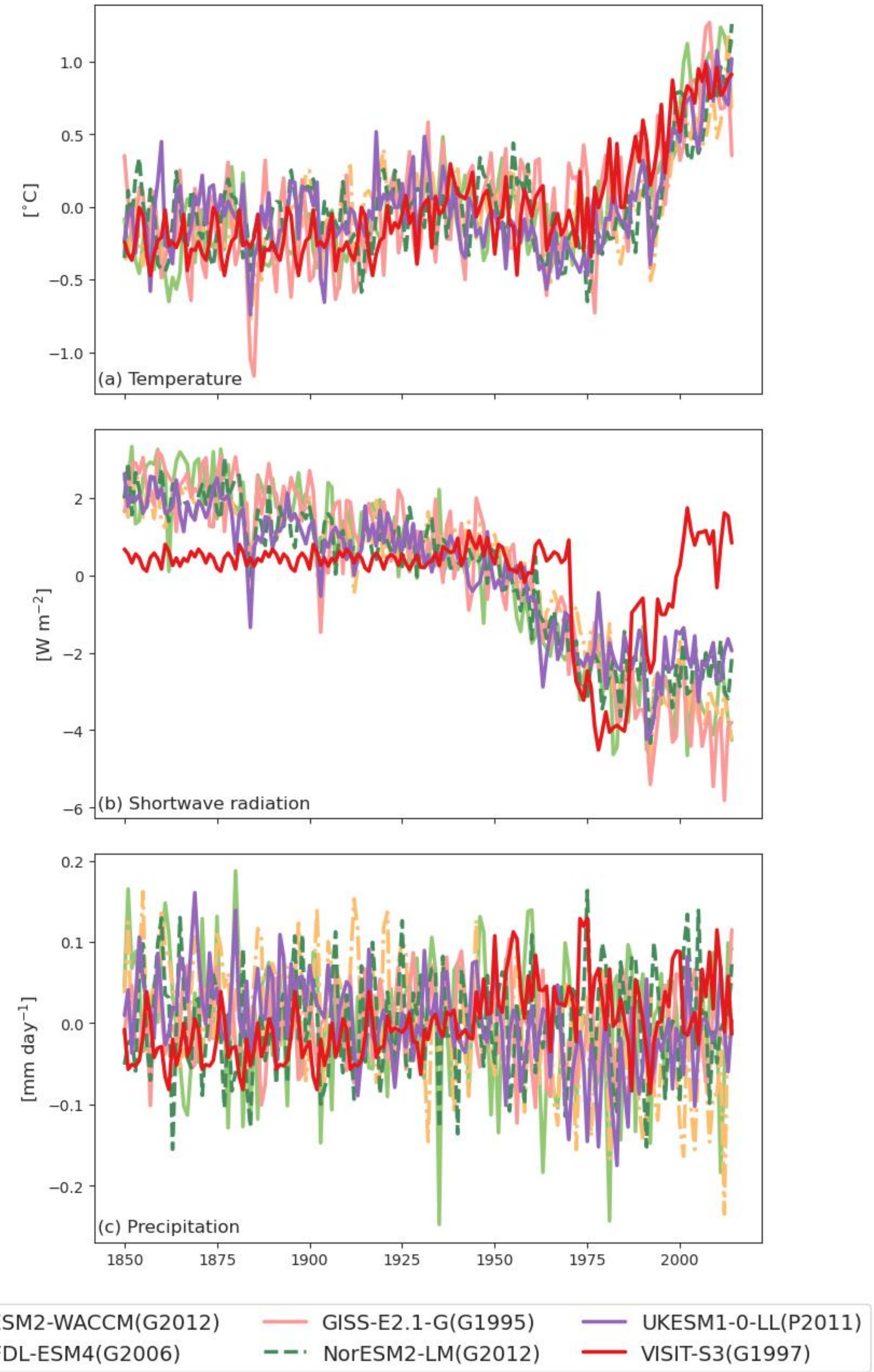


Figure S5. Global annual anomalies of (a) temperature; (b) shortwave radiation and (c) precipitation in CMIP6 models and VISIT-S3 during 1850–2014. Anomalies are deviation from baseline (1850–2014 average).

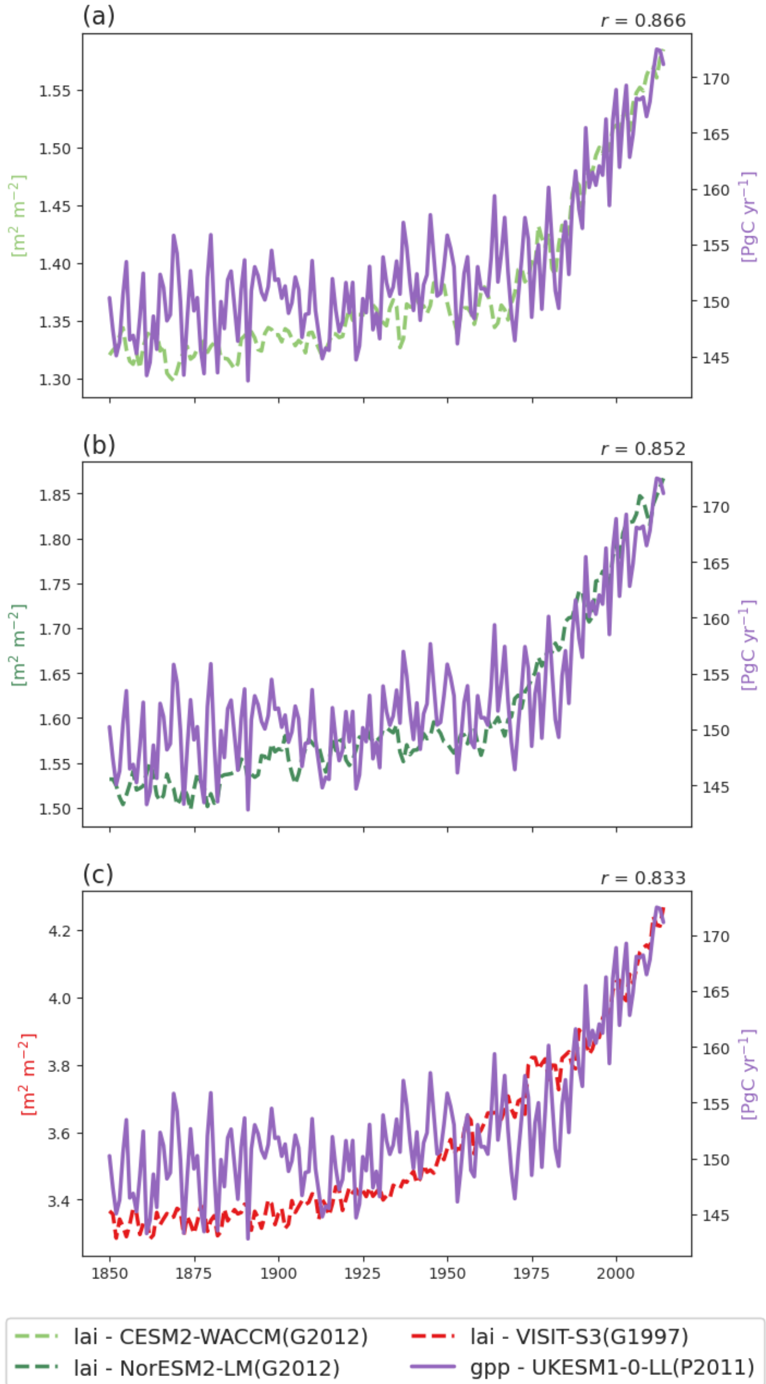


Figure S6. Global annual GPP of UKESM1-0-LL(P2011) and LAI of: (a) CESM2-WACM(G2012), (b) NorESM2-LM(G2012), (c) VISIT-S3(G1997) over land areas during 1850–2014. *r* is Pearson correlation.

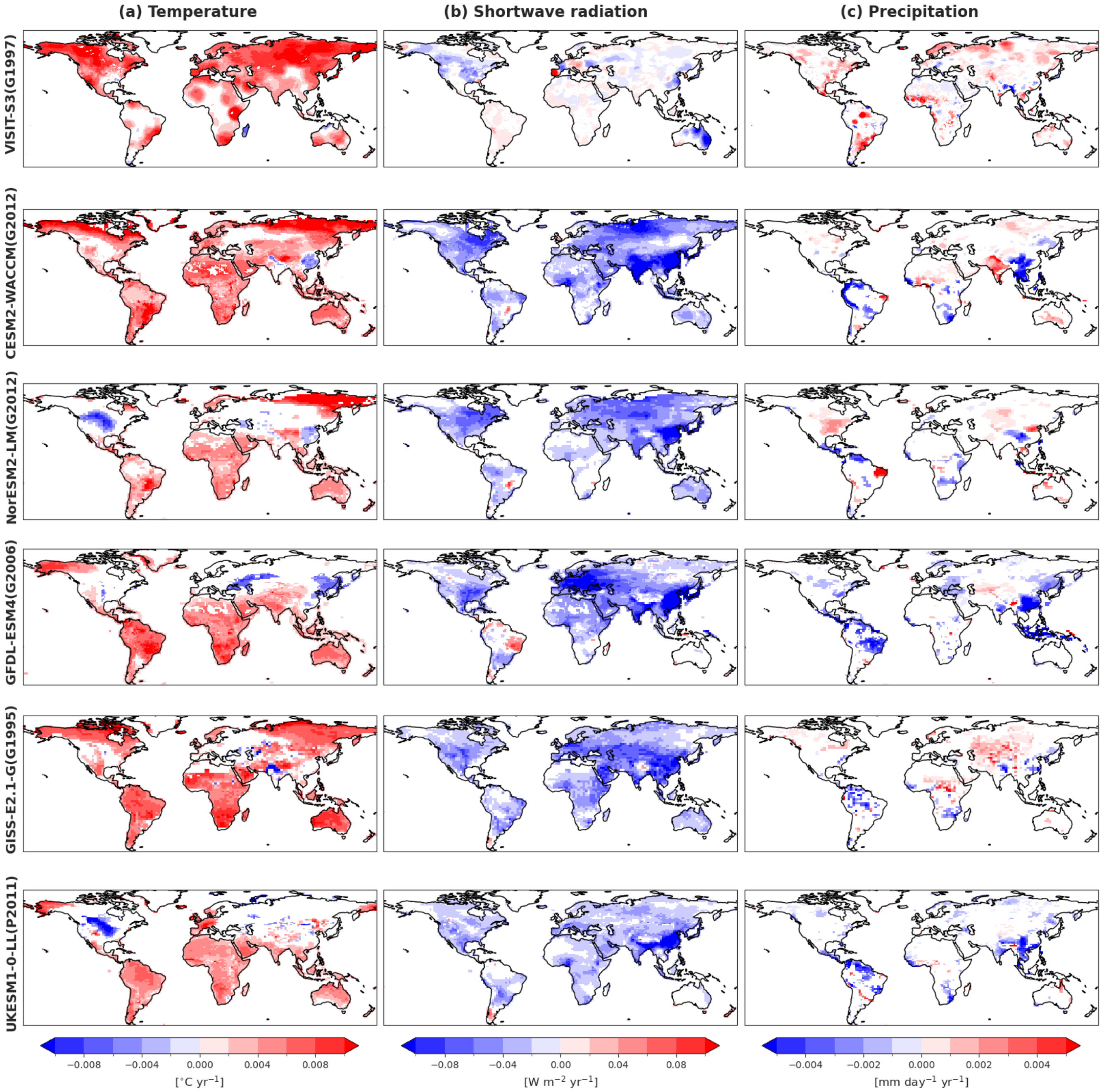


Figure S7. Spatial distribution of annual trends in (a) temperature, (b) shortwave radiation, and (c) precipitation over land areas during 1850–2014. Only significant trends (with $p < 0.05$) are presented.

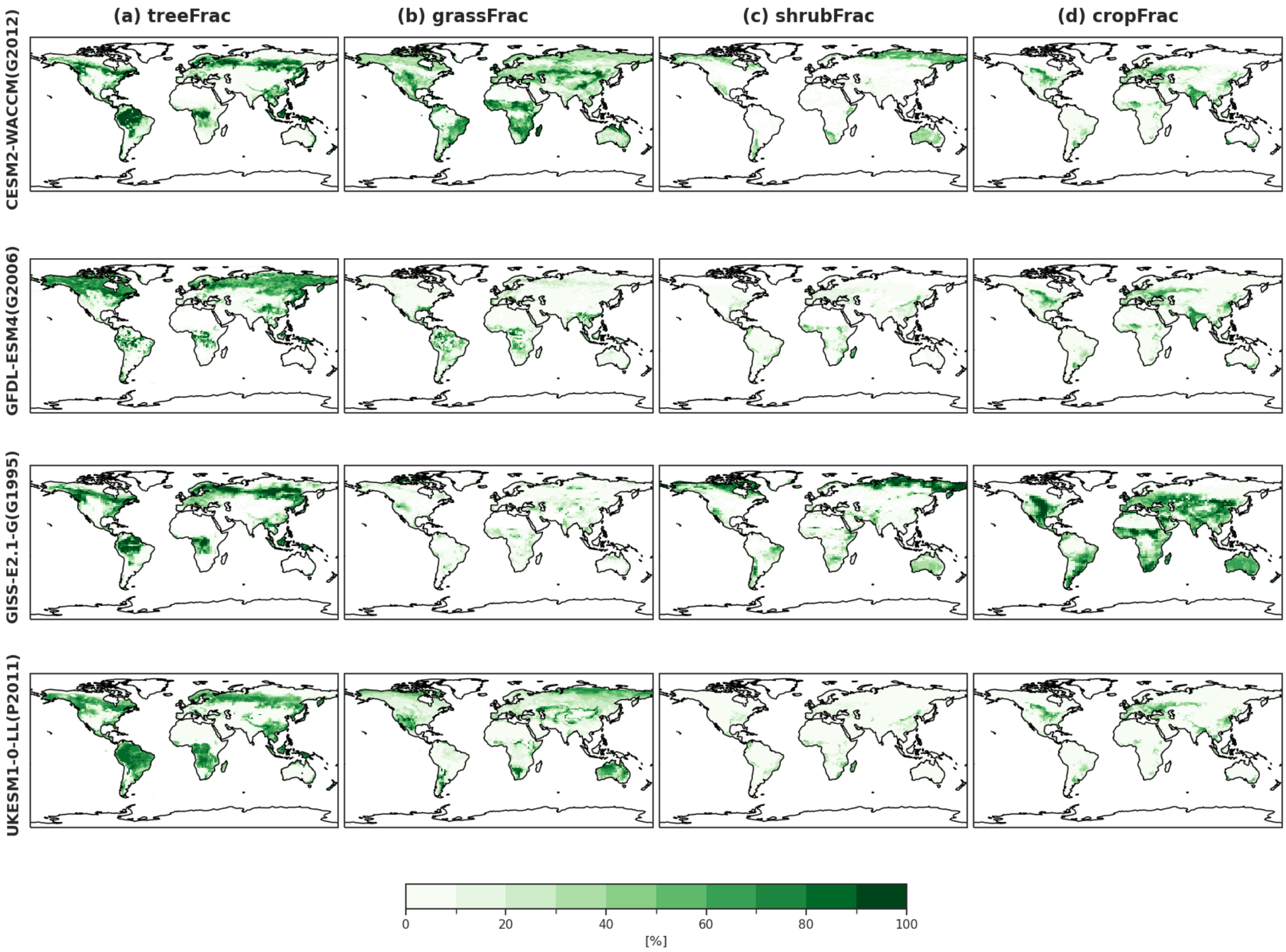


Figure S8. Mean annual fraction (%) in the present day (2000–2014) of four plant functional types (PFTs): (a) tree, (b) grass, (c) shrub, and (d) crop. Data from CESM2-WACCM(G2012), GFDL-ESM4(G2006), GISS-E2.1-G(G1995), and UKESM1-0-LL(P2011). NorESM2-LM(G2012) uses the same land component as CESM2-WACCM (G2012).