Biome classification influences current and projected future biome distributions

Simon Scheiter[1*], Dushyant Kumar[2], Mirjam Pfeiffer[3], Liam Langan[1]

26th September 2023

[1]Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany
[2]Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany
[3]Öko-Institut e.V., Rheinstrasse 95, 64295 Darmstadt

[*]Author for correspondence:
E-mail: simon.scheiter@senckenberg.de, phone: ++49 69 7542 1864

S1 – aDGVM2 model description

The following paragraphs provide a short description of the aDGVM2. A full description of the original model version is provided by Langan et al. (2017). We used an updated model version that (1) includes both C_3 and C_4 grasses (Kumar et al. 2020, 2021), in contrast to previous versions that only included C_4 grasses, (2) simulates plant-specific leaf-level photosynthetic rates based on leaf temperature and leaf traits such as the specific leaf area (Kumar et al. 2020, 2021), and (3) simulates both single-stemmed trees and multi-stemmed shrubs, where trees are better competitors in light-limited environments whereas shrubs are better competitors in water-limited environments due to a trade-off between height and water uptake capacity (Gaillard et al. 2018).

The aDGVM2 is based on ecophysiological sub-models commonly used in DGVMs (Prentice et al. 2007). These sub-models describe fundamental processes such as photosynthesis (Farquhar et al. 1980; Collatz et al. 1991, 1992), respiration (Thornley & Cannell 2000; Arora 2003; Tjoelker et al. 2001), stomatal conductance (Ball-Berry model, Ball et al. 1987) and evapotranspiration (Penman-Monteith model, Jones 1992; Allen et al. 1998) based on mathematical equations and physical principles. These sub-models were parametrized by using data from observations and field experiments and they allow the simulation of biogeochemical fluxes in response to climate and soil conditions.

The aDGVM2 is individual-based and simulates growth, reproduction, and mortality of individual plants while keeping track of state variables, such as biomass, height, and leaf area. Each plant has leaf, stem, root, storage, reproduction, and bark biomass compartments. Each plant in aDGVM2 is characterized by a plant-specific set of trait values. In the model, traits are inheritable and constant during the lifetime of a plant, whereas state variables change during the lifetime of a plant, dependent on the trait values. The selection of inheritable traits was based on traits or parameters typically used in DGVMs to parameterize PFTs. Traits describe growth form, leaf characteristics, hydraulic characteristics, carbon allocation to each of the six biomass compartments, architecture, reproduction, mortality, and response to disturbance. Most plant traits are linked by trade-offs to constrain possible trait combinations. Selection and trait inheritance assemble plant communities that are adapted to biotic and abiotic conditions. Plants with trait combinations that allow sufficient growth and reproduction rates, and that allow plants to cope with competition and disturbances can contribute their trait values to the community trait pool. Trait mutation and recombination may alter trait values in the community trait pool. Randomly drawn seeds from the trait pool are added to the plant population as seedlings. Plants that are not adapted to the prevailing disturbance regimes, biotic and abiotic conditions, or that do not allocate enough carbon to reproduction disappear from the population. Therefore, successful ecological strategies emerge dynamically from this community assembly process.

The aDGVM2 simulates four different phenological types: evergreen light-triggered, evergreen water-triggered, deciduous light-triggered and deciduous water-triggered (Langan et al.

2017). Woody plants can adopt all four types whereas we assume that grasses are evergreen. Whether a plant is deciduous or evergreen and whether it is light- or water-triggered are two inheritable traits that are constant during the lifespan of a plant, but that can change between generations due to trait inheritance and community assembly processes in aDGVM2. Deciduous vegetation switches between a dormant and a metabolically active state once moving averages of soil matric potential (water-triggered) or solar radiation (light-triggered) exceed or fall below threshold values. Evergreen woody plants remain metabolically active during their entire life time. However, leaf flushing of evergreen plants is stimulated by water and light triggers, i.e., leaf flushing occurs once moving averages of soil matric potential (water-triggered) or solar radiation (light-triggered) exceed threshold values. The threshold values are plant-specific inheritable traits. They are constant during the lifespan of a plant but they can change between generations due to trait inheritance and the community assembly processes in aDGVM2. Leaf flushing of deciduous plants and leaf regrowth of evergreen plants is supported and enhanced by carbon allocation from the storage compartment to the leaves. The amount of storage-to-leaf allocation is an inheritable trait in aDGVM2. While our approach does not allow plants to switch between phenological strategies during their lifetime, growing season length can adjust to inter-annual variation of the climate, because the thresholds used to trigger phenology can be crossed earlier or later in the year.

The aDGVM2 simulates natural surface fire regimes (Langan et al. 2017). Grass biomass and leaf litter on the ground contribute to fuel biomass. Fuel biomass and fuel moisture are used to calculate the potential fire intensity (Higgins et al. 2008). Fire spreads in a grid cell when fuel biomass is dry enough to carry fire and when the fire intensity exceeds a threshold of 300 kJ s^{-1} m^{-1}. To approximate the moisture status of fuel biomass, we use the matric potential of the upper soil layer. Fuel moisture is then used to calculate a fire probability, and fire spreads if a random number is lower than the fire probability. Fire removes the total aboveground grass biomass and aboveground biomass of trees that are not well protected against fire. Specifically, we use a 'topkill' function characterized by tree height and bark thickness (Langan et al. 2017; Higgins et al. 2008) to simulate if fire damages a tree or not. Both grasses and trees can recover from fire disturbance by re-allocating biomass from the storage pool to leaf and stem biomass. The size of the storage pool is thereby influenced by a inheritable allocation trait, that defines the amount of carbon allocated to the storage biomass compartment. The amounts of carbon allocated from storage to leaf and stem biomass are also described by inheritable traits. We assume that the storage pool is belowground where it is protected against damages by fire or herbivores. The community assembly processes and trait inheritance allow aDGVM2 to select for plants and communities that are well adapted to fire. Pathways are high allocation to bark for protection, high allocation to storage for reproduction, or low wood density and an architecture that allows rapid height growth and escaping from the flame zone. While natural fire was included in our simulations, anthropogenic fire with prescribed frequency and fire season was not considered.

While we used aDGVM2 to study grazing impacts and management in mixed cropland-rangeland systems in South Africa (Pfeiffer et al. 2019, 2022), as well as the predecessor model aDGVM to study the impacts of grazing, fuel wood harvesting and fire management (Scheiter et al. 2015; Scheiter & Savadogo 2016; Scheiter et al. 2019), we ignored direct human impacts in this study. Rather, we focused on potential natural vegetation. Previous studies showed that when ignoring areas highly affected by land use, the agreement between simulated and observed model features increases (Kumar et al. 2021; Scheiter et al. 2020). We used a similar approach to exclude areas affected by land use (see section 2.5 in the main text).

The previous summary is based on the supplementary materials of Scheiter et al. (2020) with several adjustments.

S2 – aDGVM2 model simulation protocol

We conducted aDGVM2 simulations for Africa, tropical and subtropical Asia and Australia using climate forcing compiled for the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP, Warszawski et al. 2014). These data include daily time series of bias-corrected and statistically downscaled climate variables at 0.5° spatial resolution between 1950 and 2099. We used minimum, maximum, and average near-surface air temperature, precipitation, near-surface relative humidity, near-surface wind speed, and downwelling long- and short-wave radiation. Simulations were conducted for RCP8.5 derived from GFDL-ESM2M simulations (Warszawski et al. 2014), because RCP8.5 represents an extreme case scenario within the RCPs, with high carbon emissions, high energy consumption, and low climate mitigation until 2099 (van Vuuren et al. 2011). Substantial vegetation changes simulated under this scenario are therefore most suitable to illustrate differences in future biome patterns and biome changes derived by different classification methods. Climate forcings from GFDL-ESM2M were used because they showed good performance for large parts of the study region (McSweeney & Jones 2016). Soil data was derived from the Harmonized World Soil Database (Nachtergaele et al. 2009) and we used elevation from the Shuttle Radar Topography Mission (SRTM, Jarvis et al. 2008). Atmospheric CO_2 concentrations were derived from van Vuuren et al. (2011) for RCP8.5.

We conducted a 450-year model spin-up to allow modeled state variables, traits, and plant communities to reach a dynamic equilibrium with environmental conditions. For the model spin-up, we used a random sequence of years from the period between 1950 and 1980. After the spin-up, transient simulations were conducted using the time series between 1950 and 2099. Simulations were conducted in the presence of a natural surface fire regime, as represented by the aDGVM2 fire routines (Langan et al. 2017). We simulated potential natural vegetation in the absence of anthropogenic impacts. Model simulations were conducted at 1° spatial resolution. Therefore, climate, soil and elevation data were re-sampled to the required 1×1° spatial resolution using the nearest neighbor method. Replicate runs were not conducted.

The previous description is based on Scheiter et al. (2020) with several adjustments.

S3 – Biome classification

We applied a biome classification approach based on previous aDGVM2 studies (Kumar et al. 2021; Scheiter et al. 2020). Vegetation was classified into seven biomes based on the cover of different PFTs. The approach distinguishes seven biome types: desert, C_4 grassland, C_3 grassland, shrubland, woodland, evergreen forest and deciduous forest. The PFTs used for the classification were C_4 grasses, C_3 grasses, deciduous shrubs, evergreen shrubs, deciduous trees and evergreen trees. To calculate the cover of different PFTs, we first assigned one of the six PFTs to each simulated plant individual and then calculated the sum of the crown area of all individuals of each PFT within a grid cell. Shrubs and trees were distinguished based on the stem count of individual woody plants, a model trait describing whether a plant is single-stemmed or multi-stemmed (Gaillard et al. 2018). Specifically, we assumed trees have 1 or 2 stems and shrubs have 3 or more stems. Deciduous and evergreen vegetation were distinguished based on the phenology trait in aDGVM2, which defines if a plant is deciduous or evergreen (Langan et al. 2017). C_4 and C_3 grasses are explicitly represented as different PFTs in aDGVM2.

If both grass and woody cover were lower than pre-defined cover thresholds ($300m^2$ per simulated 1ha stand, Table S1), vegetation was classified as desert. If woody cover was below $300m^2$/ha and grass cover exceeded $300m^2$/ha, vegetation was classified as C_3 or C_4 grassland, according to the dominance (i.e., higher cover) of the C_3 or C_4 grass type. If woody cover was between 300 and $30.000m^2$/ha, vegetation was classified as shrubland if shrub cover exceeded tree cover, or as woodland if tree cover exceeded shrub cover. If woody cover exceeded $30.000m^2$/ha, vegetation was classified as deciduous or evergreen forest if the cover of deciduous woody plants exceeded the cover of evergreen woody plants or vice versa. For shrubland, woodland and both forest types, grass cover was not considered for the classification. This definition of biomes uses the crown cover of each simulated plant in 1ha stands. As crowns may overlap, their sum can exceed the size of the simulated stand. The threshold values were selected such that observed and simulated biome patterns agree.

References

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: Guidelines for computing crop water requirements. Irrigation & Drainage, Paper 56, FAO, Rome, Italy.

Arora VK (2003) Simulating energy and carbon fluxes over winter wheat using coupled land surface and terrestrial ecosystem models. Agricultural and Forest Meteorology, **118**, 21–47.

Ball JT, Woodrow IE, Berry JA (1987) Progress in Photosynthesis Research, chap. A model

predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, pp. 221–224. Nijhoff, Dordrecht.

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data, **5**, 180214.

Collatz GJ, Ball JT, Grivet C, Berry JA (1991) Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. Agriculture and Forest Meteorology, **54**, 107–136.

Collatz GJ, Ribas-Carbo M, Berry JA (1992) Coupled photosynthesis-stomatal conductance model for leaves of C_4 plants. Australian Journal of Plant Physiology, **19**, 519–538.

Farquhar GD, Caemmerer SV, Berry JA (1980) A biochemical-model of photosynthetic CO_2 assimilation in leaves of C_3 species. Planta, **149**, 78–90.

Gaillard C, Langan L, Pfeiffer M, Kumar D, Martens C, Higgins SI, Scheiter S (2018) African shrub distribution emerges via height - sapwood conductivity trade-off. Journal of Biogeography, **45**, 2815–2826.

Higgins SI, Bond WJ, Trollope WSW, Williams RJ (2008) Physically motivated empirical models for the spread and intensity of grass fires. International Journal of Wildland Fire, **17**, 595–601.

Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled srtm for the globe version 4, available from the cgiar-csi srtm 90m database. http://srtm.csi.cgiar.org.

Jones HG (1992) Plants and microclimate: a quantitative approach to environmental plant physiology. 2nd edn. Cambridge University Press.

Kumar D, Pfeiffer M, Gaillard C, Langan L, Martens C, Scheiter S (2020) Misinterpretation of Asian savannas as degraded forest can mislead management and conservation policy under climate change. Biological Conservation, **241**, 108293.

Kumar D, Pfeiffer M, Gaillard C, Langan L, Scheiter S (2021) Climate change and elevated CO_2 favor forest over savanna under different future scenarios in South Asia. Biogeoscience, **18**, 2957–2979.

Langan L, Higgins SI, Scheiter S (2017) Climate-biomes, pedo-biomes or pyro-biomes: which world view explains the tropical forest – savanna boundary in South America? Journal of Biogeography, **44**, 2319–2330.

McSweeney CF, Jones RG (2016) How representative is the spread of climate projections from the 5 CMIP5 GCMs used in ISI-MIP? Climate Services, **1**, 24–29.

Nachtergaele F, van Velthuizen H, Verelst L, et al. (2009) Harmonized world soil database (version 1.1). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Pfeiffer M, Hoffmann MP, Scheiter S, et al. (2022) Effects of alternative crop-livestock management scenarios on selected ecosystem services in smallholder farming – a landscape perspective. Biogeosciences Discussions, **2022**, 1–33.

Pfeiffer M, Langan L, Linstädter A, et al. (2019) Grazing and aridity reduce perennial grass abundance in semi-arid rangelands – insights from a trait-based dynamic vegetation model. Ecological Modelling, **395**, 11 – 22.

Potapov P, Hansen MC, Pickens A, et al. (2022) The Global 2000-2020 Land Cover and Land Use Change Dataset Derived From the Landsat Archive: First Results. Frontiers in Remote Sensing, **3**, 856903.

Prentice I, Bondeau A, Cramer W, et al. (2007) Terrestrial Ecosystems in a Changing World, chap. Dynamic Global Vegetation Modeling: Quantifying Terrestrial Ecosystem Responses to Large-Scale Environmental Change, pp. 175–192. Springer.

Scheiter S, Higgins SI, Beringer J, Hutley LB (2015) Climate change and long-term fire management impacts on Australian savannas. New Phytologist, **205**, 1211–1226.

Scheiter S, Kumar D, Corlett RT, et al. (2020) Climate change promotes transitions to tall evergreen vegetation in tropical Asia. Global Change Biology, **26**, 5106–5124.

Scheiter S, Savadogo P (2016) Ecosystem management can mitigate vegetation shifts induced by climate change in West Africa. Ecological Modelling, **332**, 19–27.

Scheiter S, Schulte J, Pfeiffer M, Martens C, Erasmus BFN, Twine WC (2019) How does climate change influence the economic value of ecosystem services in savanna rangelands? Ecological Economics, **157**, 342 – 356.

Thornley JHM, Cannell MGR (2000) Modelling the components of plant respiration: Representation and realism. Annals of Botany, **85**, 55–67.

Tjoelker MG, Oleksyn J, Reich PB (2001) Modelling respiration of vegetation: evidence for a general temperature-dependent Q_{10}. Global Change Biology, **7**, 223–230.

van Vuuren DP, Edmonds J, Kainuma M, et al. (2011) The representative concentration pathways: an overview. Climatic Change, **109**, 5–31.

Warszawski L, Frieler K, Huber V, Piontek F, Serdeczny O, Schewe J (2014) The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework. Proceedings of the National Academy of Sciences, **111**, 3228–3232.

Table S1: Thresholds used for biome classification. A_{GR} - canopy area of all grasses; A_{WD} - canopy area of all woody plants; A_{C4} - canopy area of C4 grasses; A_{C3} - canopy area of C3 grasses; A_{TR} - canopy area of trees; A_{SH} - canopy area of shrubs; A_{DE} - canopy area of deciduous woody plants; A_{EV} - canopy area of evergreen woody plants. The aDGVM2 simulates vegetation within 1 ha plots and the canopy areas are calculated as the sum of the crown areas of all individuals of the respective functional type. Values are given in m^2.

Biome	A_{GR}	A_{WD}	A_{C4}	A_{C3}	A_{SH}	A_{TR}	A_{DE}	A_{EV}
Baren/Desert	< 300	< 300						
C4 grassland	≥ 300	< 300	$\geq A_{C3}$	$< A_{C4}$				
C3 grassland	≥ 300	< 300	$< A_{C3}$	$\geq A_{C4}$				
Shrubland	≥ 0	$300 \leq \cdot < 30000$			$\geq A_{TR}$	$< A_{SH}$		
Woodland	≥ 0	$300 \leq \cdot < 30000$			$< A_{TR}$	$\geq A_{SH}$		
Deciduous forest	≥ 0	≥ 30000					$\geq A_{EV}$	$< A_{DE}$
Evergreen forest	≥ 0	≥ 30000					$< A_{EV}$	$\geq A_{DE}$

Table S2: Traits and trait assemblages used for the biome classification based on cluster analysis and traits. Numbers 1 to 11 are the trait clusters TC1 to TC11, WC provides traits included the importance analysis where only weakly correlated traits (correlation coefficient < 0.9) were included.

Trait	1	2	3	4	5	6	7	8	9	10	11	WC
Phenology traits												
Phenology	X	X		X				X	X		X	X
Rain-light trigger	X	X		X				X	X		X	X
Light trigger off	X	X		X				X	X		X	X
Light trigger on	X	X		X				X	X		X	
Rain trigger off	X	X		X				X	X		X	
Rain trigger on	X	X		X				X	X		X	X
Carbon allocation traits												
Allocation to bark	X	X			X		X		X		X	X
Allocation to leaf	X	X			X		X		X		X	
Allocation to reproduction	X	X			X		X		X		X	X
Allocation to roots	X	X			X		X		X		X	
Allocation to storage	X	X			X		X		X		X	X
Allocation to wood	X	X			X		X		X		X	
Storage to leaf allocation	X	X			X		X		X		X	
Storage to wood allocation	X	X			X		X		X		X	X
Plant architectural traits												
Allometry parameter 1	X	X				X	X	X			X	
Allometry parameter 2	X	X				X	X	X			X	
Canopy shape parameter 2	X	X				X	X	X			X	
Maximum rooting depth	X	X				X	X	X			X	X
Root shape parameter 1	X	X				X	X	X			X	X
Root shape parameter 2	X	X				X	X	X			X	X
Stem count	X	X				X	X	X			X	X
Leaf economic and reproduction traits												
SLA	X	X					X	X	X	X		X
Wood density	X	X					X	X	X	X		
Leaf Nitrogen	X	X					X	X	X	X		
Seed weight	X	X					X	X	X		X	X
Size-related variable traits (model state variables)												
Crown area	X		X				X	X	X		X	
Height	X		X				X	X	X	X		X
Stem diameter	X		X				X	X	X		X	
Bark thickness	X		X				X	X	X		X	X
Age	X		X				X	X	X		X	X

Table S3: Traits included in the 10 best models in the randomized importance analysis using all traits (TC1). All models had κ >0.54. 'Number' represents the number of traits included in the different models, 'Models' represents the number of the selected models that include a trait.

Trait	1	2	3	4	5	6	7	8	9	10	Models
AllocBark	X	X	X	X	X		X	X			7
AllocLeaf	X		X	X	X	X	X	X		X	8
AllocRepr	X		X		X	X			X	X	6
AllocRoot		X		X		X	X	X		X	6
AllocStor	X	X		X	X	X	X	X	X	X	9
AllocWood		X	X	X				X			4
BarkThickness	X	X		X	X		X	X	X	X	8
BiomassPar1		X		X	X	X	X	X	X		7
BiomassPar2	X	X		X	X		X	X	X	X	8
CanfrmPar2	X	X		X	X		X	X	X	X	8
CrownArea	X	X	X	X	X	X	X	X	X	X	10
Evergreen		X					X			X	3
Height	X	X	X	X	X	X	X	X		X	9
LeafNitrogen	X		X	X	X		X	X	X		7
LightThrOff	X	X	X	X	X	X		X	X	X	9
LightThrOn	X	X	X	X	X		X				6
RainLight	X	X	X	X	X	X	X	X	X	X	10
RainThrOff		X		X		X	X		X		5
RainThrOn	X	X	X	X	X	X	X	X	X	X	10
RotfrmMaxD		X				X	X	X	X	X	6
RotfrmPar1	X		X		X						3
RotfrmPar2	X	X	X	X	X		X		X		7
SeedWeight	X	X	X	X	X	X	X	X		X	9
Sla		X				X		X	X		4
StemCount	X	X	X	X	X	X	X		X	X	9
StemDiamTot	X	X	X	X	X	X	X	X	X		9
StorToLeaf	X	X	X		X	X		X	X	X	8
StorToWood		X	X	X			X	X	X	X	7
WoodDensity				X	X	X				X	4
age	X		X		X	X	X	X	X	X	8
Number	21	24	19	23	23	19	23	22	20	20	

Table S4: Traits included in the 10 best models in the randomized importance analysis including only weakly correlated traits (WC). All Models had κ >0.499. 'Number' represents the number of traits included in the different models, 'Models' represents the number of the selected models that include a trait.

Trait	1	2	3	4	5	6	7	8	9	10	Models
AllocBark	X	X	X	X	X	X	X	X	X	X	10
AllocRepr	X	X	X	X	X	X	X	X	X	X	10
AllocStor	X	X	X	X		X	X	X	X	X	9
BarkThickness	X	X	X	X	X	X	X	X	X	X	10
Evergreen				X							1
LightThrOff	X	X	X	X			X	X		X	7
RainLight	X		X	X	X	X	X	X	X	X	9
RainThrOn	X	X		X			X		X		5
RotfrmMaxD	X	X	X	X		X	X	X		X	8
RotfrmPar1				X	X	X				X	4
RotfrmPar2	X		X	X	X	X		X	X		7
SeedWeight		X	X	X	X	X	X		X	X	8
Sla			X	X							2
StemCount	X	X	X		X	X	X	X		X	8
StemDiamTot	X	X	X	X	X	X	X	X	X	X	10
StorToWood	X	X	X	X	X	X	X	X	X	X	10
age	X	X	X	X	X	X	X	X	X	X	10
Number	13	12	14	16	11	13	13	12	11	13	

Table S5: Data-model agreement and percent of the grid cells projected to undergo biome changes until 2099 considering land use. 'κ' quantifies the data-model agreement, 'Δ_{lt}' represents the percent of grid cells projected to undergo biome changes when considering the entire study area as reference, 'Δ_{le}' represents the percent of grid cells projected to undergo biome changes when considering only the area covered by natural vegetation as reference (i.e., excluding areas with land use), 'Cover' represents models where cover of different PFTs was used for the cluster analysis, 'Traits' represents models where trait cluster 8 (TC8) and the Beck et al. (2018) biome map were used for the cluster analysis. For biome classification, the percentage of grid cells projected to undergo biome changes are Δ_{lt} Class=16.3% and Δ_{le} Class=18.7%. This percentage is not related to the F31 biome maps as the maps were not used for developing the classification scheme. Simulations were conducted at 1°spatial resolution, the biome maps were aggregated to the same grid. To account for land use, we used the masked out areas where the fraction of cropland and built-up areas in the Potapov et al. (2022) exceed 15%.

Biome map	Δ_{lt} Cover	Δ_{lt} Traits	Δ_{le} Cover	Δ_{le} Traits
Allen et al., 2020	21.1	18.6	24.7	21.7
Buchhorn et al., 2019	20.9	22.2	24.4	25.9
Beck et al., 2018	16.7	19.5	19.5	22.8
Hengl et al., 2018	20.6	19.9	24.1	23.3
Dinerstein et al., 2017	17.5	13.8	20.5	16.1
Zhang et al., 2017	14.3	19.8	16.7	23.1
Netzel & Stepinski, 2016a	18.5	18.0	21.6	21.1
Netzel & Stepinski, 2016b	19.1	16.4	22.3	19.2
Higgins, Buitenwerf, & Moncrieff, 2016	18.5	14.7	21.6	17.2
Pfadenhauer & Kloetzli, 2014	23.5	22.0	27.5	25.7
Zhang & Yan, 2014	14.2	19.3	16.6	22.6
Metzger et al., 2013	23.0	21.3	26.9	25.0
FAO, 2010	21.2	19.2	24.8	22.4
Tateishi et al., 2011, 2014	23.0	22.2	26.9	26.0
Defries et al., 2010	16.6	19.9	19.5	23.3
Ellis & Ramankutty, 2008	21.2	23.2	24.8	27.2
European Space Agency, 2010	18.1	19.9	21.1	23.3
Friedl et al., 2010	16.6	17.8	19.4	20.9
The Nature Conservancy, 2009	17.7	13.1	20.7	15.4
Peel et al., 2007	17.4	19.1	20.3	22.3
Bartholome & Belward, 2005	18.6	22.6	21.8	26.5
Kaplan et al., 2003	21.1	21.8	24.7	25.5
Olson et al., 2001	21.1	14.8	24.7	17.3
Loveland et al., 2000	19.7	19.5	23	22.8
Ramankutty & Foley, 1999	12.5	18.3	14.6	21.4
Leemans, 1990	25.1	23.6	29.4	27.6
Schultz, 1988, 1995, 2002, 2008, 2016	18.3	17.5	21.5	20.5
Mueller-Hohenstein, 1981	23.1	18.3	27	21.4
Schmithuesen, 1976	23.8	21.4	27.8	25.0
Whittaker, 1975	20.7	20.6	24.2	24.1
Walter, 1964	20.7	16.3	24.2	19.0
Mean	19.5	19.2	22.8	22.4

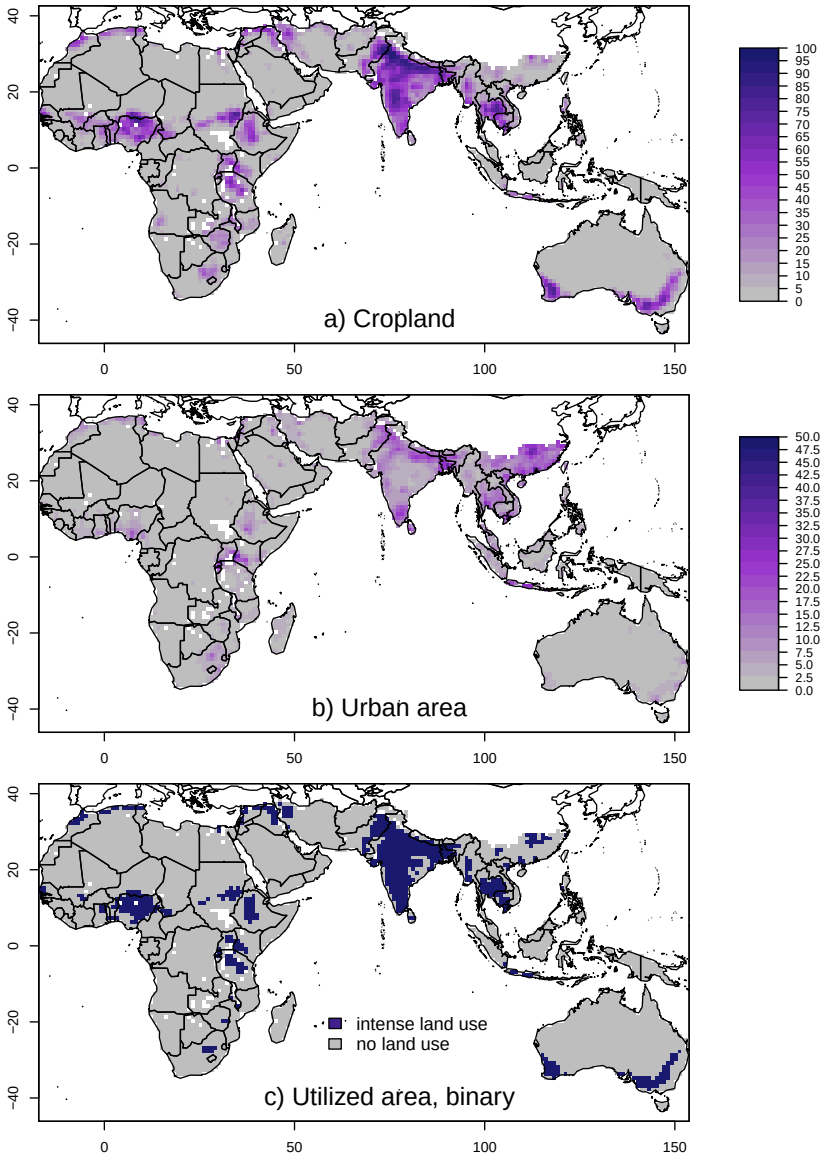

Figure S1: Areas covered by land use. To derive these areas, we used cropland (a) and built-up areas (b) provided by Potapov et al. (2022). Areas where the cover of cropland and built-up areas exceeded 15% cover were classified as intense land use (c). Those areas were masked out for analyses including land use. Maps were aggregated to the 1°spatial resolution used in aDGVM2 simulations by calculating mean values.

Figure S2: Mean trait values of different biome types for the cluster analysis using the Beck et al. (2018) map and all traits (TC1). Mean trait values of selected traits are provided in Table 2 in the main text. Definition of biomes on y-axis: 1 - Tropical rainforest; 2 - Tropical monsoon; 3 - Tropical savanna; 5 - Arid steppe hot; 6 - Temp. dry winter warm summer; 7 - Temp. dry winter hot summer; 9 - Temp. no dry season hot summer; 8 - Arid desert hot; 13 - Temp. no dry season warm summer; 14 - Temp. dry summer hot summer; 15 - Arid desert cold; 17 - Arid steppe cold.

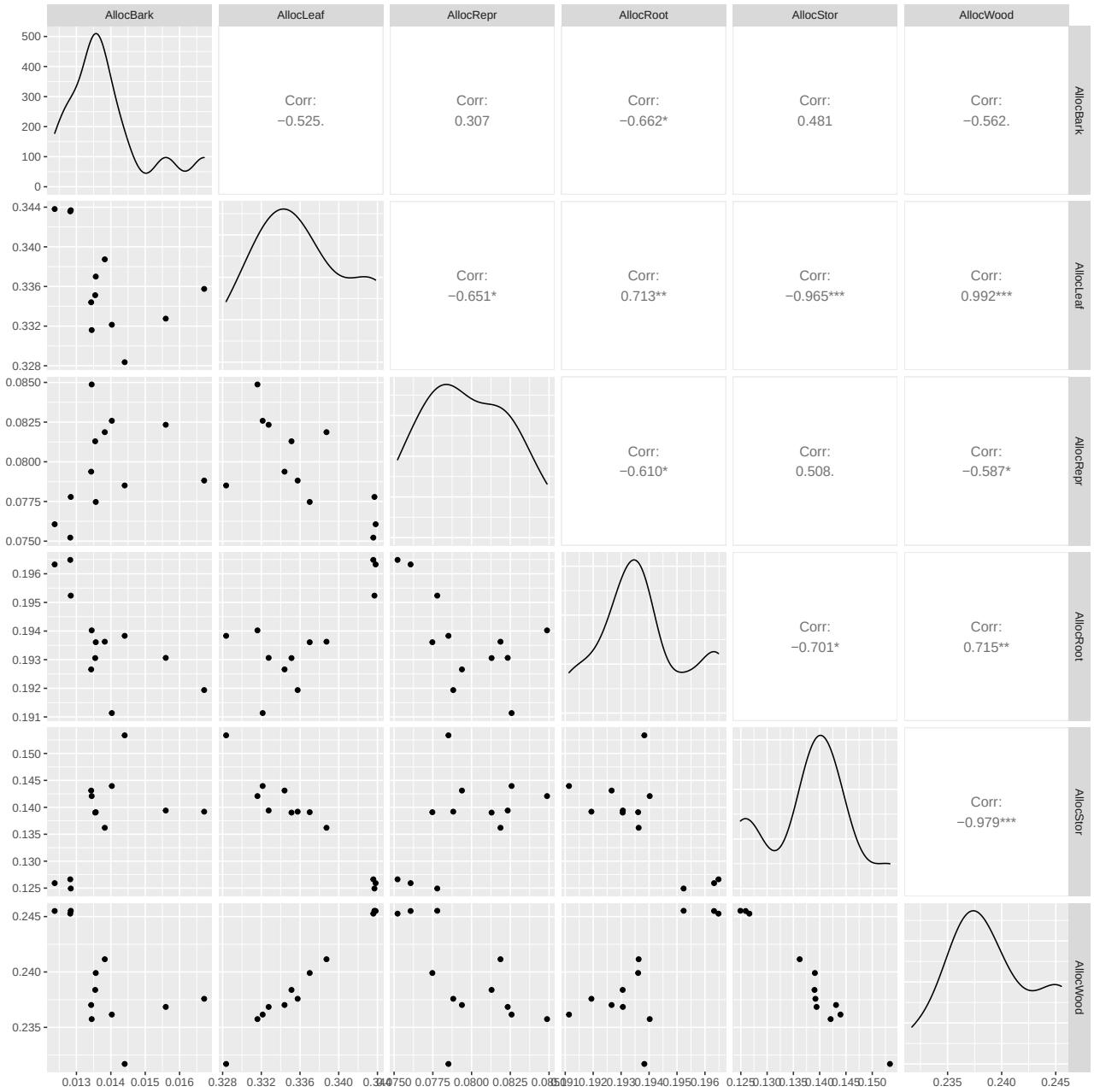

Figure S3: Correlations between mean carbon allocation traits in different biomes. Each point in the plots represents one of the biomes in the Beck et al. (2018) map, classification was conducted with all traits (TC1). Note that the differences in terms of absolute values are small between biomes.

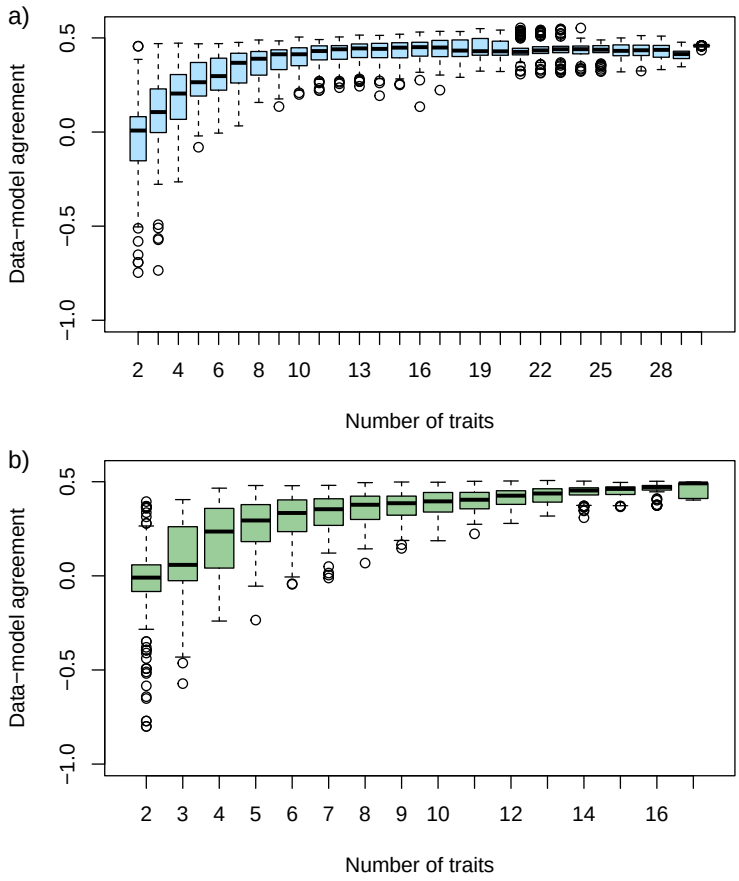

Figure S4: Relation between data-model agreement and the number of traits included in the cluster analysis. For each number of traits, the traits were randomly selected, and the cluster analysis was repeated 150 times. Analyses were conducted including all traits (TC1, panel a) and only weakly correlated traits (WC, panel b). Data-model agreement is represented by the κ statistics.

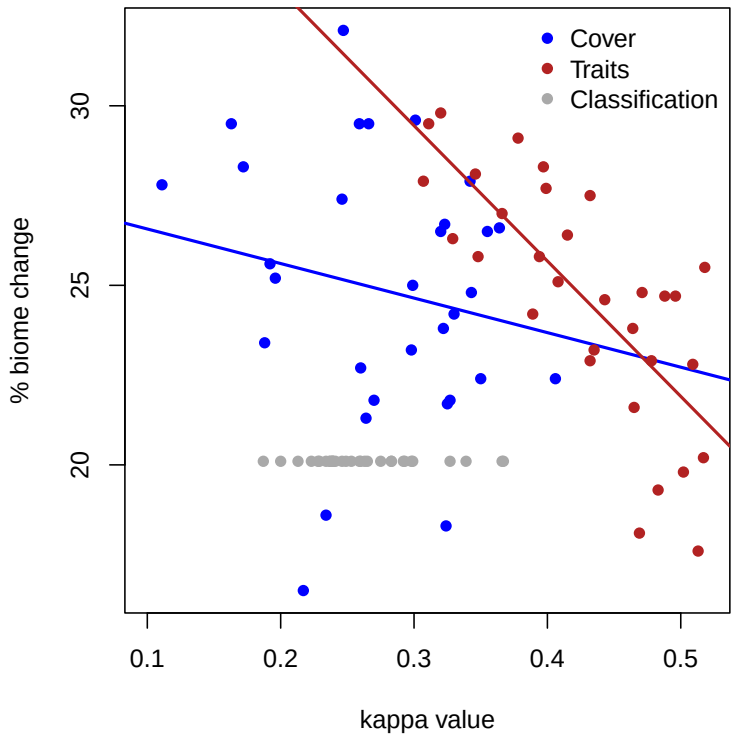

Figure S5: κ values and percent of grid cells affected by biome changes until 2099 for different classification methods and the F31 biome maps used to develop classification schemes. The lines represent linear regression models (adjusted R^2=-0.0006357, p-value=0.3302, correlation coefficient=-0.1808834 for clustering using cover of PFTs; adjusted R^2=0.5536, p-value=9.704e-07, correlation coefficient=-0.7539822 for clustering using traits). The data for figure are provided in Table 1 in the main text.

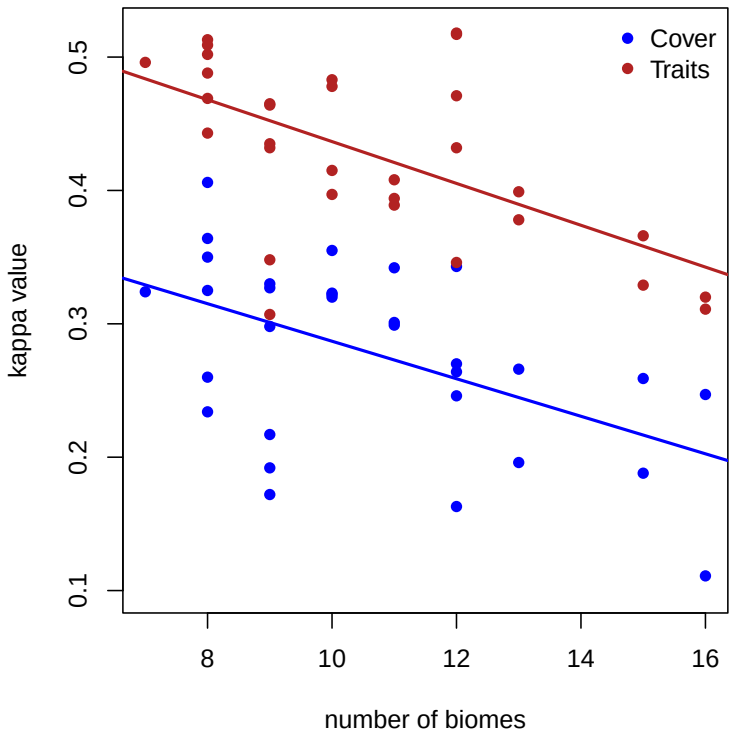

Figure S6: Relation between the κ values and the number of biomes in the different F31 biome maps used to develop classification schemes. The lines represent linear regression models (adjusted R^2=0.2362, p-value=0.003268, correlation coefficient=-0.5115406 for clustering using cover of PFTs; adjusted R^2=0.3364, p-value=0.0003725, correlation coefficient=-0.5987969 for clustering using traits).

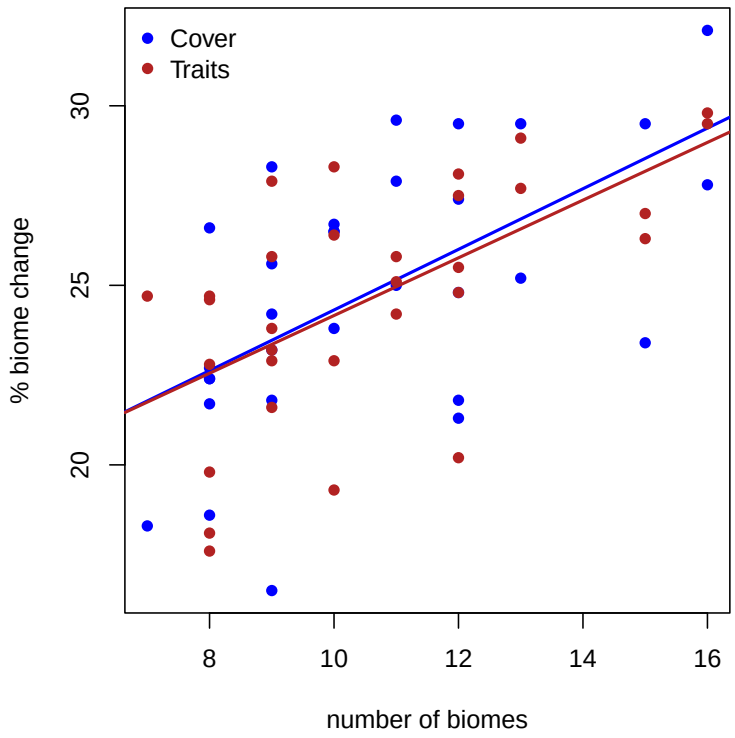

Figure S7: Relation between the percent of grid cells undergoing biome changes until 2099 and the number of biomes in the different F31 biome maps used to develop classification schemes. The lines represent linear regression models (adjusted R^2=0.3099, p-value=0.000679, correlation coefficient=0.5769748 for clustering using cover of PFTs; adjusted R^2=0.3549, p-value=0.0002422, correlation coefficient=0.6135361 for clustering using traits).

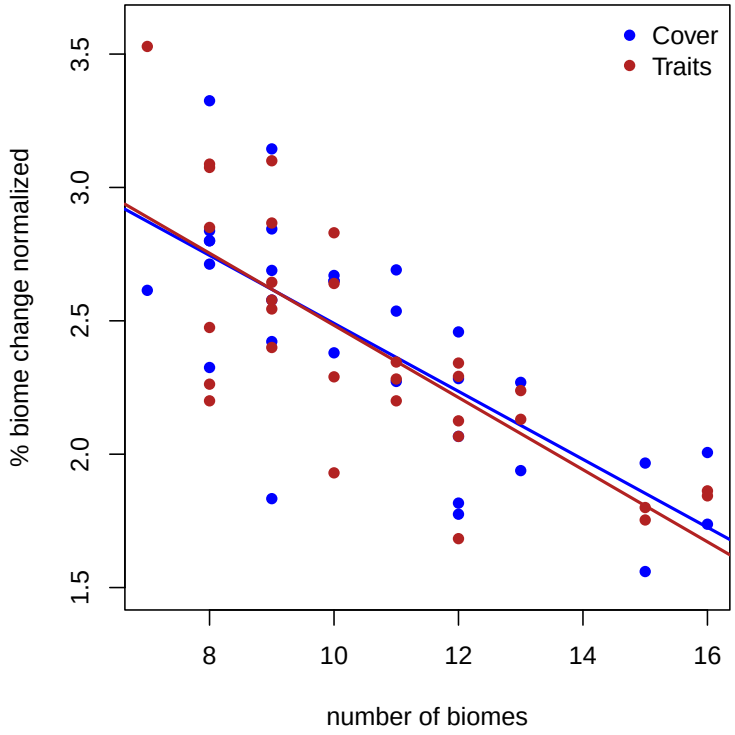

Figure S8: Relation between the average percent of grid cells per biome undergoing biome changes until 2099 and the number of biomes in the different F31 biome maps used to develop classification schemes. The lines represent linear regression models (adjusted R^2=0.5332, p-value=1.887e-06, correlation coefficient=-0.740768 for clustering using cover of PFTs; adjusted R^2=0.5609, p-value=7.612e-07, correlation coefficient=-0.7586113 for clustering using traits).

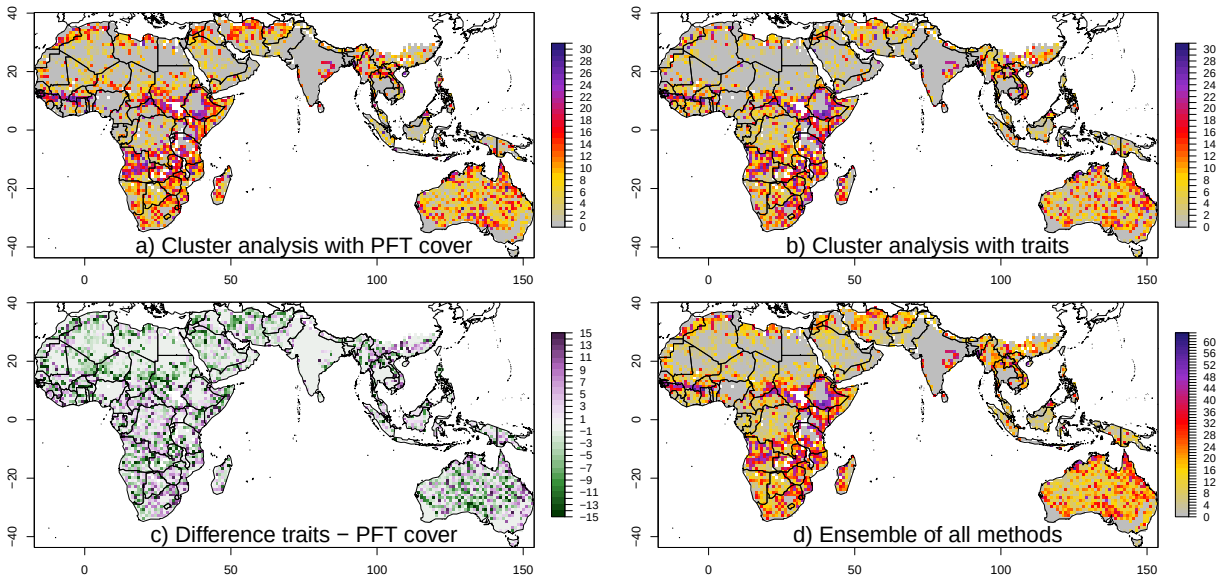

Figure S9: Consensus maps of biome transitions until 2099 considering land use. The consensus maps were derived by counting the number of models that simulate a biome change in each simulated grid cell, (a) 31 models using cover of different PFTs, (b) 31 models using traits, and (d) all 63 models (1 classification using PFTs, 31 cluster analyses using PFTs, 31 cluster analyses using traits). Panel (c) shows the difference between models using traits and models using PFT covers (difference between panels a and b), and illustrates where biome projections from clustering methods with PFTs and traits deviate. We constrained the values to a range between -15 and 15 to improve the visibility of the differences. Areas covered by land use were derived from Potapov et al. (2022), and excluded in this analysis. See Fig. S1 for areas covered by land use, and Fig. 3 in the main text for results ignoring land use. Figures represent results from TC8.

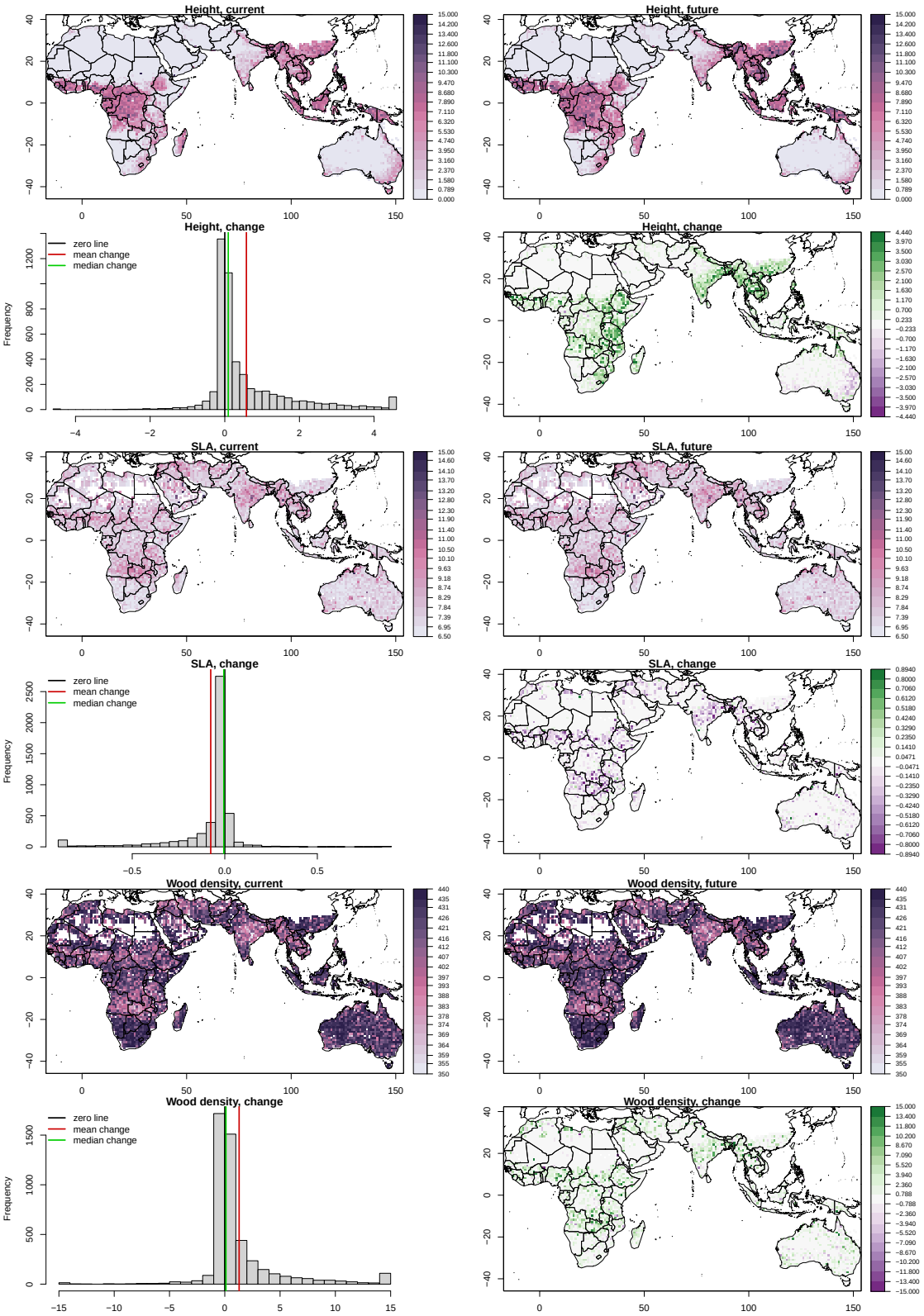

Figure S10: Patterns of height, SLA and wood density simulated with aDGVM2 under current and future conditions, and changes between current and future conditions.