https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Importance of plant functional type, dynamic vegetation, and fire interactions for process-based modeling of gross carbon uptake across the drylands of western North America

Rubaya Pervin[1*], Scott M. Robeson[1], Mallory Barnes[2], Stephen Sitch[3], Anthony P. Walker[4], Ben Poulter[5], Fabienne Maignan[6], Qing Sun[7,8,9], Thomas Colligan[5,10], Sönke Zaehle[11], Kashif Mahmud[12], Peter Anthoni[13], Almut Arneth[13], Vivek K. Arora[14], Vladislav Bastrikov[15], Liam Bogucki[16], Bertrand Decharme[17], Christine Delire[17], Stefanie Falk[18,19], Akihiko Ito[20], Etsushi Kato[21], Daniel Kennedy[22], Jürgen Knauer[23,24], Michael O'Sullivan[3], Wenping Yuan[25] and Natasha MacBean[26,27]

[1]Department of Geography, Indiana University, Bloomington IN, USA.

[2]O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington IN, USA.

[3]Faculty of Environment, Science and Economy, University of Exeter, UK.

[4]Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge TN, USA

[5]NASA Goddard Space Flight Center, Biospheric Sciences Lab., Greenbelt, MD, USA.

[6]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France.

[7]Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland.

[8]Wyss Academy for Nature, University of Bern, Bern, Switzerland.

[9]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland.

[10]Earth System Science Interdisciplinary Center, University of Maryland, College Park MD, USA.

[11]Max Planck Institute for Biogeochemistry, Jena, Germany.

[12]Kimbell School of Geosciences, Midwestern State University, Wichita Falls TX, USA.

[13]KIT, Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

[14]Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria BC, Canada.

[15]Science Partners, Paris, France.

[16]Department of Biology, Western University, London ON, Canada

[17]Centre National de Recherches Météorologiques, Toulouse, France.

[18]Institute of Meteorology and Climate Research Atmospheric Trace Gases and Remote Sensing, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany.

[19]Department of Geography, Ludwig-Maximilians University Munich, Munich, Germany.

[20]Graduate School of Life and Agricultural Sciences, University of Tokyo, Tokyo, Japan.

[21]Institute of Applied Energy, Minato, Tokyo 105-0003, Japan.

[22]National Center for Atmospheric Research, Boulder CO, USA.

[23]School of Life Sciences, Faculty of Science, University of Technology Sydney, Ultimo, NSW 2007, Australia.

[24]Hawkesbury Institute for the Environment, Western Sydney University, Penrith, NSW 2751, Australia.

[25]Institute of Carbon Neutrality, Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing 100091, China.

[26]Department of Geography and Environment, Western University, London ON, Canada.

[27]Department of Biology, Western University, London ON, Canada.

*Corresponding author email: rpervin@iu.edu.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Abstract

Drylands cover ~41% of the Earth's land surface and contribute more than one third of the global net primary productivity. Several studies have demonstrated that drylands play a crucial role in global carbon cycle interannual variability. However, drylands are vulnerable to the impacts of climate change. To predict changes in dryland productivity under climate change we depend on dynamic global vegetation models (DGVMs). Compared to more mesic ecosystems, DGVM carbon cycle dynamics have not been widely evaluated against data. Existing studies are mostly focused at site scale; rarely have these models been assessed or benchmarked against dryland carbon flux products at regional to global scales. Global gross primary productivity (GPP) products have poor performance in dryland regions. Only recently upscaled in situ flux products have been developed specifically for drylands. Here, we evaluated GPP inter-annual variability (IAV) simulated by 15 DGVMs from the TRENDY v11 model intercomparison project against theDryFlux GPP, which is newly developed upscaled GPP product that considers dryland-specific ecohydrological responses. Comparing model simulated GPP IAV to DryFlux, we identified two groups of models: a one group of models with generally lower GPP IAV than DryFlux (e.g., lower standard deviation in annual GPP than DryFlux and slope values of the linear regression between each model and the DryFlux product that are less than 1.0) and a second group of models with generally higher GPP IAV than DryFlux. We examined if including a representation of dynamic vegetation (i.e., changes in the spatial distribution of plant functional type (PFT) fractional cover) or fire in the models can explain the inter-model spread and model performance in comparison to DryFlux. Models that do not include a representation of fire and/or dynamic changes in plant functional type distribution over time generally have lower annual GPP variability compared to DryFlux (1st group of models), except for the eastern and southeastern region of the study area with high rainfall variability. We also found that models with dynamic vegetation exhibit high variability in grass fractional cover (that was higher than two independent reference fractional cover datasets), which was strongly correlated with high GPP IAV. Only some models that included fire simulated burnt area annual variability that correlated well with GPP IAV.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Other models that included fire simulated low burnt area variability and therefore we did not find any strong relation between burnt area and GPP IAV. Finally, we examined the relationship between the dominant PFT and GPP IAV. We did not find a strong correlation between the spatial mean of the slope of the linear regression between each model and DryFlux annual GPP and their spatial mean woody, grass, or C3 grass fractional cover (although many models with generally low GPP IAV had higher woody plant cover). However, we did find a high correlation between the slope of the linear regression between each model and DryFlux annual GPP and spatial mean C4 grass cover. Therefore, our findings suggest that DGVMs inability to accurately represent the spatial distribution of herbaceous (specifically C4 grass) cover as well as processes controlling dynamically changing vegetation distributions over time (including fire) contribute to poor model performance in capturing annual variability in dryland productivity. Our findings can provide a roadmap for DGVM teams seeking to improve vegetation representation in sparsely vegetated dynamic dryland ecosystems.

1. Introduction

Dryland ecosystems are commonly defined as regions where water demand in the form of potential evapotranspiration (PET) is much higher than precipitation (P) (Wang et al., 2022) and encompass grasslands, shrublands, scrublands and savannas. They cover ~41% of the Earth's terrestrial surface and are home to over a third of the world's population (Bastos et al., 2022; Wang et al., 2022) . Dryland regions are hotspots of land-atmospheric coupling (Koster et al., 2004) and are thought to play a dominant role in global carbon cycle variability (Ahlström et al., 2015; Poulter et al., 2014; Zhang et al., 2018). Dryland ecosystem functioning is expected to be extremely sensitive to future changes in water availability (Bastos et al., 2022; Scholes, 2020; Wang et al., 2022). Given the complexity of dryland ecosystem dynamics and their feedbacks to major components of the earth system, it is essential that dynamic global vegetation models (DGVMs), many of which form the land component of earth system

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

models used for climate change projections, can accurately simulate carbon and water fluxes in dryland regions.

DGVMs are process-based models that simulate physical and biogeochemical processes to model exchange of water, energy, carbon (C), and nitrogen (N) fluxes across the land-atmosphere boundary. All DGVMs simulate vegetation dynamics such as leaf phenology and changes in vertical structure or rooting depth; however, only some explicitly simulate spatial distribution of plant functional types (PFTs) and dynamic changes in PFT spatial distribution (fractional cover) over time (henceforth referred to as "dynamic vegetation"). When the spatial distribution of a models' PFTs is not explicitly simulated it needs to be prescribed. While many of these models serve as the land component of their respective ESMs, these models can also be driven offline with meteorological data (either from *in situ* weather stations or from gridded climate reanalyses (Bonan, 2019). The TRENDY ("Trends and drivers of the regional scale terrestrial sources and sinks of carbon dioxide") model intercomparison project began in 2009 with the goal of comparing a suite of DGVM estimates for global atmosphere-land CO_2 flux (Sitch et al., 2015). The TRENDY model ensemble simulations provide important estimates of the natural land carbon sink (cumulative net biome production) in addition to the impact of land use changes on land carbon cycling for the annual Global Carbon Budget (Friedlingstein et al., 2024); however, outputs from TRENDY have been used in a host of other studies exploring land surface and dynamic vegetation process responses to global change drivers beyond the original TRENDY remit (Pan et al., 2020; Yuan et al., 2019; Zhu et al., 2016).

Compared to mesic ecosystems, DGVMs have rarely been evaluated or benchmarked against dryland carbon and water flux data at either site or global scale(MacBean et al., 2021; Renwick et al., 2019; Whitley et al., 2016). However, several recent studies have documented large uncertainties in DGVM simulations of dryland carbon fluxes and stocks(Fawcett et al., 2022; MacBean et al., 2021; Metz et al., 2023; Teckentrup et al., 2021; Traore et al., 2014; Whitley et al., 2016) – although the mean across

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

models tends to replicate benchmark data well (Fawcett et al., 2022). Most of the DGVM evaluations in the aforementioned studies have been conducted at site scale using *in situ* carbon and water fluxes; only a few are focused on regional(Metz et al., 2023, 2025; Teckentrup et al., 2021) or global scale (Fawcett et al., 2022). These spatially continuous large-scale studies used gridded upscaled flux tower data or satellite derived GPP, aboveground biomass, fire emissions and burnt area products to evaluate multiple components of dryland C cycling. Interestingly, there are mismatches between the findings for how well models can capture C fluxes at site versus global scale. Studies evaluating a suite of DGVMs from the TRENDY model intercomparison project across Australian and southwestern US dryland sites found that most models underestimate both the mean gross primary productivity and net ecosystem exchange (NEE) and their interannual variability (IAV)(MacBean et al., 2021; Metz et al., 2023; Wang et al., 2022; Whitley et al., 2016), which the studies suggest may be due to inadequate representations of vegetation dynamics such as tree-rooting depths and leaf phenology, which are essential for capturing plant responses to rainfall variability(Renwick et al., 2019; Whitley et al., 2016, 2017). Global scale flux site data analysis also revealed a similar story that DGVMs underestimate GPP IAV across dryland savanna, grass and shrubland sites(Lin et al., 2023). However, studies comparing TRENDY models with regional to global gridded datasets found a large spread across the models with some models overestimating and some underestimating mean gross and net carbon fluxes and their IAV(Fawcett et al., 2022; Teckentrup et al., 2021). Furthermore, model performance varied regionally, such as a general overestimation of modeled GPP in the Sahel but an underestimation in South Africa and South America(Fawcett et al., 2022). These contrasting findings from site scale studies highlight the need for more regional to global studies that can examine processes and patterns that become important across broader spatial scales.

Many processes have been postulated to cause poor model performance in capturing dryland C fluxes, but the exact causes are not yet known (Fawcett et al., 2022; MacBean et al., 2021; Paschalis et al., 2020; Teckentrup et al., 2021; Traore et al., 2014; Whitley et al., 2016). DGVMs are at various stages of development, with only some including key dryland processes such as fire, grazing, or dynamic

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

vegetation (Sitch et al., 2024). Productivity in drylands is affected by rainfall changes, with herbaceous vegetation responding quickly to rainfall variations, while woody vegetation has slower, longer-term response (Verbruggen et al., 2021, 2024). These responses vary depending on the dryland aridity and the type of vegetation cover; thus, it is important to have correct representations of dryland vegetation type and fractional cover (fCover) (either prescribed or simulated). Inadequate representation of prescribed PFTs, PFT fCover and its variability over time in DGVMs have been shown to cause large spread in model estimates of dryland productivity, especially in sparsely vegetated, mixed shrub-grass ecosystems (Hartley et al., 2017; Wilcox et al., 2023). This uncertainty arises from challenges in separating grass versus woody PFT fractions via classification of remotely sensed imagery (Hartley et al., 2017; Pervin et al., 2022). Different climate dependencies of dryland PFTs, such as differentiating between C3 and C4 grasses based on temperature thresholds, also cause discrepancies across models (Still et al., 2019; Wilcox et al., 2023). In addition, most PFTs are not well adapted for dryland ecosystems. Herbaceous PFTs, and especially C4 grasses, are often oversimplified in models, leading to underestimation of productivity, especially in ecosystems where annual grasses and variations in fire regimes play a major role in ecosystem dynamics (Wilcox et al., 2023).

Models that do include dynamic vegetation have a variety of different methods for simulating changes over time, from relatively simple bioclimatic envelopes to more complex representations based on competition for resources, forest gap dynamics and cohorts of different ages and sizes (Harper et al., 2018; Prentice et al., 1992; Sitch et al., 2003; Smith et al., 2001). The treatment of fires in models also adds to uncertainties in representing dryland vegetation dynamics and productivity in DGVMs (Teckentrup et al., 2021). Fire models vary from relatively simple representations of fuel availability and moisture availability such as the GlobFIRM model (Thonicke et al., 2001) to more complex representations that also account for different sources of ignition and the effect of wind on rates of fire spread (Li et al., 2012; Mangeon et al., 2016; Melton and Arora, 2016). Fires in dryland regions can significantly alter vegetation structure (Bastos et al., 2022; Wilcox et al., 2018), yet models often fail to

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

capture the impact of fire on dryland vegetation, particularly in areas with low fuel availability where fire is less frequent (Baudena et al., 2015; Fawcett et al., 2022; Teckentrup et al., 2021; Verbruggen et al., 2021; Wilcox et al., 2023). Thus, even when these key dryland processes are included in models it remains unclear to what extent they are correctly represented. Failure to accurately model these biogeographic processes and their interplay with climate variability likely affects DGVMs' ability to predict GPP IAV. More in depth model-data comparison studies that focus on the spread across models are needed to help discern which model processes need further evaluation or improvement. These studies can also inform targeted data collection and analysis efforts to support more specific process-based model evaluations and model developments. Regional model-data comparisons fill a unique niche in that they can focus more on larger-scale biogeographical causes of model-data mismatch compared to site based evaluations. Regional scale studies can also focus more on specific dryland regions and utilize localized knowledge compared to global studies (e.g., Australia in Teckentrup et al., 2022).

In this study, we address these research needs by evaluating the ability of 15 DGVMs from TRENDY v11 (Sitch et al., 2024) to capture spatiotemporal patterns of dryland GPP in western North America. We focus specifically on how well models capture GPP IAV as most previous regional dryland model evaluation studies have focused more on mean model biases or trends in productivity (Fawcett et al., 2022; Teckentrup et al., 2021). DGVMs are also being used in studies investigating the ongoing debate around which biome – tropical forests or dryland ecosystems – plays a dominant role in global C cycle IAV (Ahlström et al., 2015); therefore, it is critical that we assess whether these models can indeed capture variability in dryland C fluxes. To evaluate TRENDY model GPP we use the newly developed 'DryFlux' v1.0 GPP product (Barnes et al., 2021). DryFlux is an upscaled flux product developed specifically for dryland ecosystems using flux tower GPP from sites across western North America (Barnes et al., 2021). The random forest model used for flux tower data upscaling includes metrics designed to characterize ecohydrological controls on dryland carbon dynamics (see Section 2.3.1). DryFlux GPP better captures dryland spatiotemporal GPP patterns across western North America

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

compared to the MODIS GPP and FLUXCOM products that have been used in past regional dryland model evaluation studies (Barnes et al., 2021).

The two main objectives of this study are to: 1) identify if TRENDY v11 models have higher or lower variability in annual GPP compared to the DryFlux product across dryland regions of western North America; and 2) examine potential causes of inter-model spread and in differences in model performances with respect to DryFlux, with a specific focus on the role of processes related to biogeography such as vegetation type, fractional cover, dynamic vegetation, and fire. We test 3 main hypotheses.

First, we hypothesized that DGVMs with dynamic vegetation enabled will have a higher GPP variability compared to those with prescribed static vegetation. This is due to the higher grass disturbance rates (Harper et al., 2018) and the ability of most dynamic vegetation enabled DGVMs to establish new grass growth or to spread into bare soil patches following disturbance (Harper et al., 2018; Sitch et al., 2003). Therefore, we predicted that dynamic vegetation-enabled DGVMs will better capture DryFlux GPP IAV, while DGVMs with prescribed static vegetation fCover will underestimate DryFlux GPP IAV.

Second, given dryland grasses, especially those with the C4 photosynthetic pathway, respond strongly to rainfall pulses (Verbruggen et al., 2021, 2024), we hypothesize that models with high grass fCover (whether prescribed or simulated) will show greater GPP variability compared to models with low grass fCover or a higher bare soil or shrub fCover. We predicted that models with a more accurate representation of grass fCover will better match DryFlux GPP IAV compared to those that either over- or underestimate grass cover.

Finally, we hypothesized that the impact of fire on simulated GPP IAV in fire-enabled DGVMs depends on vegetation composition, specifically, the dominance of grasses versus bare soil or woody plant cover. Because grasses typically have a much lower resistance to burning or a higher combustion completeness

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

in models (Sitch et al., 2003), we expect fire to contribute more strongly to GPP IAV in grass-dominated systems than in shrub- or tree- dominated ones. All DGVMs with dynamic vegetation also have their fire module enabled, which we expect will contribute to both increased mean grass cover and greater variability in grass fCover, ultimately resulting in higher GPP IAV. Fire enabled DGVMs without dynamic vegetation should also predict higher GPP IAV than DGVMs with neither dynamic vegetation nor representation of fire, especially if their prescribed PFT map has high vegetation fractional cover.

To assess model performance in capturing GPP IAV, we examined the standard deviation in annual GPP for both the models and DryFlux and the slope values of the linear regression between each model and DryFlux annual GPP. Based on these values we identified models with low GPP IAV and high GPP IAV compared to DryFlux and investigated whether dynamic vegetation and/or fire could explain differences between the two groups. For dynamic vegetation enabled models, we did a pairwise gridcell comparison to explore how the temporal variability in PFT fractional cover was related to GPP IAV metrics (standard deviation and slope). To further disentangle the role that dynamic vegetation and fire are playing in modeled GPP IAV, we examined differences across models in terms of their spatial patterns in standard deviation in annual fractional cover for major plant functional types (and the same for burnt area for fire enabled models). We also explored whether the spatial distribution of dominant PFTs can be linked to spatial patterns in GPP variability. The findings of this study will help clarify which vegetation and disturbance related biogeographical processes most strongly influence interannual variability in dryland GPP across current DGVMs. By identifying where and why models diverge from the observation-driven DryFlux product, we provide a roadmap for dryland process-specific developments that the DGVM community can make to better estimate dryland carbon-vegetation interactions under a changing climate.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

2. Methods

2.1 Study Area: North American Drylands

Our study area is focused on western North American drylands (longitude and latitude bounds: 93° W to 125° W and 15°N to 53°N – Fig. 1). We delineated the dryland region across western North America using the aridity index (AI), which is defined as the ratio of precipitation to potential evapotranspiration. Unlike other studies that typically define drylands as regions with an aridity index between 0 and 0.65 (e.g., Wang et al., 2022), here we limit the range of our study area to aridity index values between 0.05 and 0.5, encompassing both arid (AI 0.05 to 0.2) and semi-arid (0.2 to 0.5) regions, to match the extent where DryFlux v1.0 is tested and can be used as a reference product. We excluded the semi-arid cropland belt in the north and eastern part of the study area (see Section 2.4) to focus on natural vegetation ecosystem productivity in North American drylands. The region exhibits large spatial gradients in key climatic variables. Mean annual precipitation (MAP) ranges broadly from 40 to 1200 mm, while mean annual temperature (MAT) varies from -4°C to +26°C (Anderson-Teixeira et al., 2011; Biederman et al., 2017). The western part near the Pacific Ocean experiences a Mediterranean climate with the majority of annual precipitation falling in winter (November–April). In contrast, the central, eastern, and southeastern portions are heavily influenced by the North American Monsoon, which brings most precipitation during the summer (July–October. Lower elevation areas (≤1610 m) are mainly dominated by C3 shrubs and C4 summer grasses (but with C3 winter grasses) are mostly affected by summer monsoon rainfall, with relatively small amounts of winter and spring rains. Higher elevation forested areas (≥1930 m) are dominated by evergreen shrubs and trees, though there is a transition in vegetation type from grasslands and mixed shrub-grasslands to forests as elevation increases. The higher elevation areas have cooler mean annual temperatures (<10°C) and a more distinct bi-model growing season, receiving moisture from both

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

winter precipitation/spring snowmelt and summer monsoon rainfall (Anderson-Teixeira et al., 2011; Biederman et al., 2017).

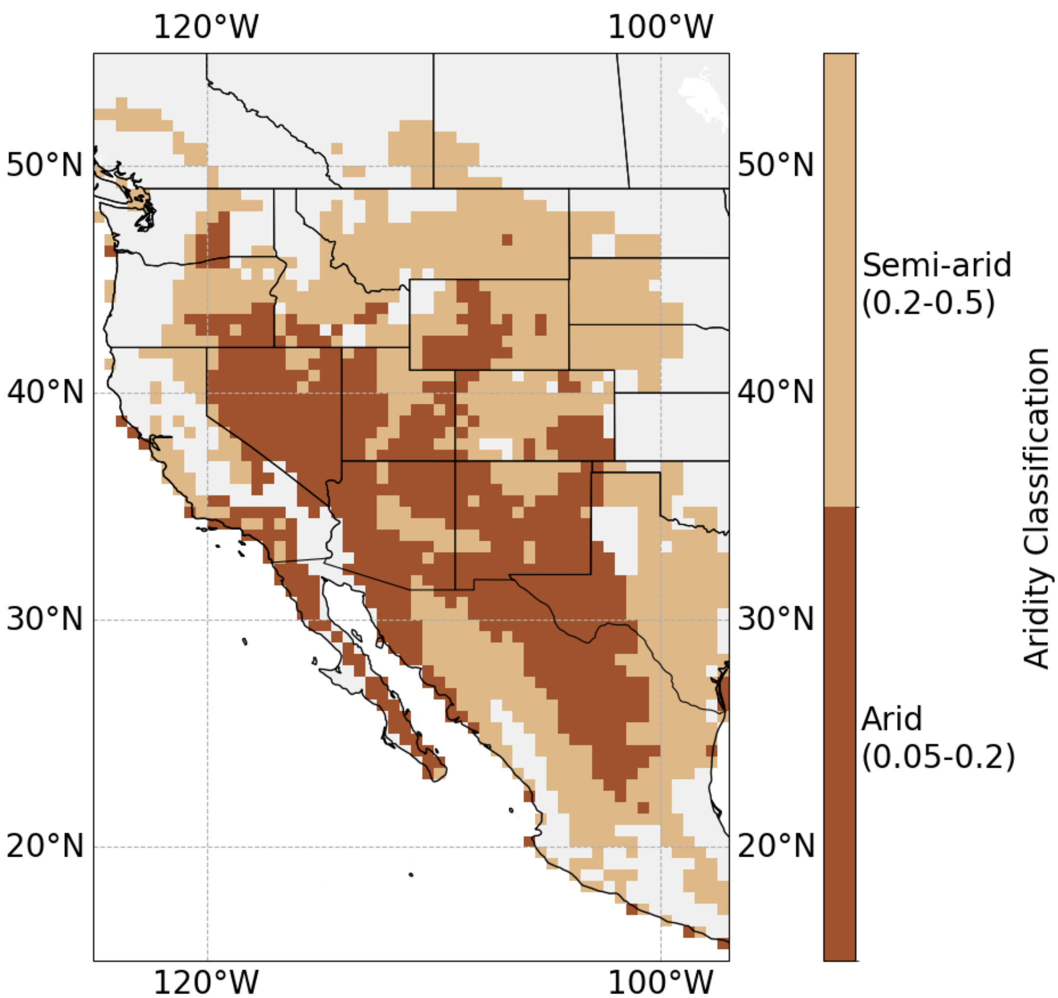

Figure 1: Drylands (excluding modeled cropland PFTs) in western North America. Arid regions (aridity index 0.05 - 0.2) are shown in dark brown, semi-arid regions (aridity index 0.2 - 0.5) in light brown, and all other regions in light grey).

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

2.2 Models

2.2.1 TRENDY v11 GPP

We used the annual GPP (kgCm^{-2}yr^{-1}) simulated by 15 of the 18 DGVMs that contributed to the TRENDY v11 models ensemble (Friedlingstein et al., 2022; Sitch et al., 2024). Three of the TRENDY v11 models were excluded because there were no PFT maps available for the analysis (see Section 2.2.2). Model simulation results were available from the preindustrial period (1700 for some models and 1860 for others) to 2021, but we elected to use 2001 to 2016 only to match the temporal availability of the DryFlux data. We used TRENDY S3 simulations specifically due to their inclusion of time-varying atmospheric CO_2 concentration, climate (using the CRUJRA reanalysis), nitrogen fertilization, and land-use change as forcings. This is in contrast to what the S1 (only time-varying atmospheric CO_2 concentration) and S2 simulations (only time-varying atmospheric CO_2 concentration and climate) included. Additionally, all of these models followed a common experimental protocol, as is further outlined in Sitch et al. (2024). The model outputs were regridded to a common 0.5x0.5° resolution grid. Table 1 lists the models included in this study in addition to whether they included two key biogeographic processes that are important controls on GPP IAV: dynamic vegetation and fire.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

300 Table 1: Information on biogeography processes (dynamic vegetation and fire) included in all TRENDY
301 v11 models used in this study that are relevant to vegetation distribution (fractional cover) and its
302 variability and which in turn affect GPP variability.

Model	**Dynamic Vegetation**	**Fire module**
CABLE POP	No	No
CLASSIC	No	Yes (Melton et al., 2020)
CLM5	No	Yes (Li et al., 2012)
IBIS	No	No
ISBA CTRIP	No	Yes (Thonicke et al., 2001)
JSBACH	No	Yes (Thonicke et al., 2010)
JULES	Yes (Burton et al., 2019; Harper et al., 2018)	Yes (Mangeon et al., 2016) (Burton et al., 2019)
LPJ-GUESS	Yes (Smith et al., 2014) (Smith et al., 2001)	Yes (Thonicke et al., 2001)
LPJwsl	Yes (Sitch et al., 2003)	Yes (Thonicke et al., 2001)
LPX-Bern	Yes (Prentice et al., 1992; Sitch et al., 2003; Wolf et al., 2008)	Yes (Thonicke et al., 2001)
ORCHIDEE	No	No
OCN	No	No
SDGVM	No	Yes (Simple empirical function of litter moisture and/or amount and a fire return interval map).
VISIT	No	Yes (Thonicke et al., 2001)
YIBs	No	No

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

2.2.2 PFT maps

Each of the TRENDY models have their own PFT fCover maps (either prescribed or predicted) based on the PFT types represented in their model (Supplementary Table S1). The common land use change forcing is imposed on the models' own PFT maps (or dynamic PFT simulations) (Sitch et al., 2024). Each model has a different number of PFTs and spatial resolution. Some models have shrubs as separate PFT such as CABLE-POP, CLM5, IBIS, ISBA-CTRIP, JSBACH3.2, VISIT, YIBs but some have combine shrubs with woody tree vegetation such as CLASSIC, LPJ-GUESS, LPJwsl, LPX-Bern, OCN, ORCHIDEE, SDVGM. All models have separate C3 and C4 PFTs, except VISIT and OCN (although in OCN these were defined as tropical/temperate grasses). Some models do not have a separate bare soil such as LPJwsl, LPX-Bern (except for an 'Urban Bare' class), VISIT, and YIBs. To compare the models' vegetation cover fractions to each other and to independent datasets (see Section 2.3.2) we grouped each models' PFTs into 4 main groups: total vegetation, woody vegetation, non-woody vascular vegetation, and bare ground/soil (Table S2). If a model had a mix of grasses and trees like "open shrubland" or a savanna vegetation class it was added within the category of woody vegetation (JULES, JSBACH, CLM5, VISIT, YIBs). For certain analyses we split the non-woody vegetation group further to compare grasses/herbs versus crops and C3 versus C4 types. PFT types that fell outside of these groups (such as 'Peat graminoid" for LPX-Bern) were ignored in this study. All PFT maps were available at annual time scale except for ISBA-CTRIP and LPX-Bern which provided monthly PFT fractional coverages. However, ISBA-CTRIP does not model spatial distribution of vegetation dynamically and therefore we expect to have no monthly variability in fractional coverage of its PFTs. Although LPX-Bern models spatial distribution of vegetation dynamically, from visually inspecting its monthly data using Panoply we did not identify any monthly changes in the PFT distributions (only annual changes were identified). Therefore, we choose the month of July data for both ISBA CTRIP and LPX-Bern PFT distributions in place of annual values. For models that do not have a bare ground or soil PFT, we calculated barren fCover using 1 minus the total fCover of all other vegetation. We note that no model contains a biological

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

soil crust (biocrust) PFT (Table S2) despite the fact that biocrusts are common in dryland regions (REF). The PFT maps were provided by the modelers and mean fCover maps for each of the major groups are provided in Fig. S1. We used nearest neighbor resampling to re-grid all the PFT maps to 0.5° spatial resolution to match the GPP data resolution.

2.3 Data products

2.3.1 Dryflux GPP v1.0

We used the DryFlux v1.0 GPP product (Barnes et al., 2021) to compare and evaluate TRENDY model GPP. DryFlux v1.0 is a dryland specific ecohydrologically informed upscaled flux tower GPP product specifically developed for western North American drylands using remote sensing and gridded meteorological inputs. DryFlux upscaling used a machine learning (random forest) model to identify relationships between predictor variables (such as vegetation greenness, precipitation, temperature, elevation) and flux tower GPP for randomly selected 19 sites out of the 24 sites selected across the US Southwest and Northwestern Mexico. In addition to these common predictor variables, DryFlux used previous months precipitation and the Standardized Precipitation Evapotranspiration Index, SPEI, at different timescales (e.g., a lag of 1, 3, 6, 12 months) to account for antecedents effects that are characteristic of dryland ecosystems (Barnes et al., 2016; Cranko Page et al., 2022; Liu et al., 2019). DryFlux was able to capture the pulse behavior and two growing reasons for the dryland regions. The random forest model was applied to generate an upscaled, gridded global dryland GPP product for 16 years from Feb, 2000 to 2016 with a spatial resolution of 0.5°x0.5°. DryFlux v1.0 data were downloaded from https://github.com/marthageb/DryFlux.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

2.3.2 Reference Fractional Cover Data

2.3.2.1 Rangeland Analysis Platform (RAP) fCover Data

The rangeland analysis platform (RAP) provides fCover estimates of annual and perennial forbs and grasses, shrubs, trees, and bare ground—from 1986 to present at 30-meter spatial resolution for the western US (Allred et al., 2021). Crops were not included in this product. The percentage estimates of these six cover types are model predictions from a temporal convolutional network using vegetation field plots, the Landsat reflectance data and vegetation indices- normalized difference vegetation index (NDVI) and normalized burn ratio two (NBR2) which is sensitive to water content changes within vegetation. Model mean absolute error (MAE) was calculated for fCover against vegetation field data. Overall model mean absolute error was about 6%. Mean absolute error was highest for perennial forbs and grasses which is about 10% and lowest for tree cover which is about 3% (Allred et al., 2021). To better match the model PFTs we added the annual and perennial grasses and forbs layers together to create a herbaceous layer and shrubs and trees to create a woody vegetation layer. RAP fCover data was downloaded from http://rangeland.ntsg.umt.edu/data/rap/rap-vegetation-cover/v3/. The data were aggregated to 0.5°x0.5° resolution to match TRENDY model PFTs and DryFlux and model GPP.

2.3.2.2 MODIS Vegetation Continuous Field Data

MODIS Vegetation Continuous Field (VCF) product (MOD44B) contains three layers of global land cover fractions: percent tree cover (tree height > 5m), percent non-tree cover, and percent non-vegetated or bare ground cover plus water. Trees are defined as greater than 5m; therefore, most of the shrubs that are present in this area will be grouped with the non-tree cover class. Crops are also included as part of the non-tree cover layer. These fCover layers were estimated from a fully-automated machine learning algorithm using the MODIS Terra Global 250-meter resolution surface reflectances (bands 1 to 7) and brightness temperature (band 32) (DiMiceli et al., 2021). The root mean square error value against field based fCover data ranges from 9 to 23% (DiMiceli et al., 2021). The MODIS VCF product is available at

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

250-meter pixel spatial resolution and yearly temporal resolution in the Sinusoidal coordinate system. The MODIS VCF product was downloaded using Google Earth Engine. Within Google Earth Engine we reprojected and resampled the data to the WGS84 Geographic Coordinate System (EPSG:4326) at 0.5°x0.5° to match the RAP product, TRENDY model PFTs, and model and DryFlux GPP.

2.4 Model and Data Preprocessing

Following spatial resampling as described in each subsection of Section 2.3, we performed further model and data preprocessing steps to bring them into the same spatial extent and temporal scale. One of our main objectives is assessing model data IAV. Therefore, the monthly DryFlux data were summed to annual GPP to compare to the annual GPP from models participating in the TRENDY intercomparison. All GPP data were converted to the same unit of $kgCm^{-2}y^{-1}$. For the regional total we calculated the weighted sum in $PgCyr^{-1}$. We used aridity index values of between 0.05 to 0.5 (Section 2.1) to create a dryland mask for the GPP and PFT datasets. The aridity index data is based on the method of Zomer et al., 2022 and was downloaded from Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v3 (Trabucco and Zomer, 2022). We note that this is not the same climate data as was used to drive the TRENDY model simulations (Section 2.2). The aridity index data were spatially aggregated to 0.5x0.5° resolution to match the model and data products used in this study. We used the annual average aridity index for the 1970–2000 period to delineate our study area. We masked out croplands in the study area as the main focus of this research is mixed woody non-woody ecosystems that experience natural climate variability (and are not irrigated) and because DryFlux did not include any crop flux tower sites in its upscaling model. We created a crop mask common to all models by masking grid cells for all models that contained crop PFT fCover (Fig. S1g) greater than 0.5 (50%) for any model within the 2001 to 2016 period. That means that we masked out the grid cells that are crop dominated for any model. We masked all grid cells that were not common to both the TRENDY model ensemble and DryFlux.

2.5 Data analysis and evaluation metrics

For a first comparison of whether TRENDY models tend to have higher or lower interannual GPP variability (GPP IAV) compared to DryFlux, we plotted time series of both spatially weighted sum of annual GPP and the annual GPP anomalies (i.e., mean normalized annual GPP over the 2001 to 2016 period for the 15 TRENDY v11 model ensembles and the DryFlux product. We computed Pearson's correlation (R) values between modeled and DryFlux total annual GPP over the study area. We plotted maps of the spatial patterns in how well each of the TRENDY models captures DryFlux GPP IAV by calculating for each grid cell both standard deviation (over the 2001 to 2016 period) of the annual GPP time series for all 15 models and DryFlux as well as the slope of the linear least-squares regression between modeled and DryFlux annual GPP (using the scipy.stats linregress module in Python). To facilitate comparison across models, and between the models and DryFlux, we examined the distribution of GPP standard deviation and slope values across the study region using area-weighted kernel density estimate (KDE) ridgeline plots that summarize probability density of all grid cells.

To investigate our hypotheses related to the role of dynamic vegetation and fire in spatial patterns of modeled annual GPP variability, we first grouped the KDE ridgeline plots by models that used dynamic vegetation and fire (Table 1) to identify if either of these key processes alone or in combination can explain the patterns in model whether models have higher or lower GPP IAV compared to DryFlux. To further explore the reasons behind the differences between models, and between model and DryFlux GPP, we calculated and mapped the standard deviation in annual PFT fCover for each grid cell over the 2001 to 2016 period for each of the 4 main PFT groups: total vegetation, woody vegetation, non-woody vascular vegetation, and bare ground/soil, with a further split of non-woody category into grasses and crops, and then C3 and C4 grasses. We compared the first three categories to determine which major PFT categories (woody versus non-woody) were contributing the most to total PFT annual variability. We then examined whether the spatial patterns in the standard deviation in PFT annual fCover match the spatial patterns in

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

model annual GPP standard deviation and the slope of the linear regression between model and DryFlux annual GPP. For dynamic vegetation enabled models we then performed a gridcell pairwise comparison to determine the nature of the relationship between modeled total vegetation annual fractional cover variability and annual GPP variability. We calculated how much of the variance in annual GPP can be explained by variance in PFT annual fCover by computing least-squares linear regressions (Python module scipy.stats.linregress). For the grid cell pairwise comparison, we calculated the coefficient of variation (CV) which is standard deviation of values over the 2001 to 2016 time period divided by the mean over the same period, because the CV is a more appropriate statistic to compare the variability at locations with different means. We performed the same set of analyses for the variability in modeled burnt area (standard deviation in annual percentage burnt area over the 2001 to 2016 period) to understand fire effects on spatial patterns in both modeled PFT fCover and GPP variability. Finally, we examined how accurately the dynamic vegetation enabled models represent the vegetation variability within the study region by benchmarking the annual PFT fCover standard deviation maps against independent reference fCover datasets from existing remote sensing data (see section 2.3.2).

Besides overall vegetation annual fCover variability due to dynamic vegetation and fire, we also assessed whether higher fCover of one of the major PFT groups can be linked to modeled GPP IAV (specifically, the slope of the linear regression between model and DryFlux annual GPP). For this purpose, for each model we took the spatial mean (± 1 s.d.) fCover of woody, grass, C3 grass, and C4 grass across all grid cells with fCover above a certain threshold (0.1 and 0.5) to assess if a model with higher fCover of a specific PFT tends to over- or underestimate DryFlux annual GPP by comparing with the spatial mean (± 1 s.d.) of the slope of the linear regression We performed this analysis for grid cells with mean fCover above both 0.1 and 0.5 considering that 10% is a typical uncertainty in satellite fCover products while greater than 50% mean fCover represents the grid cells dominated by that cover type. Then we plotted the mean fCover against the slope values to determine if models with higher mean fCover in a specific PFT group can be linked to models' performance in capturing DryFlux GPP IAV. Finally, we compared the

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

model mean PFT fCover with our independent reference datasets (MODIS VCF for woody and nonwoody groups and RAP for woody and grass vegetation types) to assess if the relationships between model vegetation fractional cover and GPP variability are undermined by inaccurate modeled fCover distributions.

3. Results

3.1 Comparison of model and DryFlux annual GPP variability

Figure 2 shows the time series (2001-2016) of annual GPP summed over the entire study area. Although the TRENDY ensemble mean total GPP (thick grey curve in Fig. 2) is lower than the DryFlux GPP (thick black curve in Fig. 2) the sign of the annual anomalies in TRENDY ensemble mean GPP match those of DryFlux well (Figs. 2a and b; R: = 0.96). However, individual TRENDY models simulate both higher and lower annual total GPP (Fig. 2a) and GPP IAV (Fig. 2b) compared to tDryFlux(Fig. 2a). Most models underestimate both DryFlux mean GPP and IAV (brown curves in Figs. 2a and b), while a selected few models (OCN, SDGVM and the LPJ family of models) tend to overestimate both the mean DryFlux GPP magnitude and the IAV (blue curves in Figs. 2a and b).

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

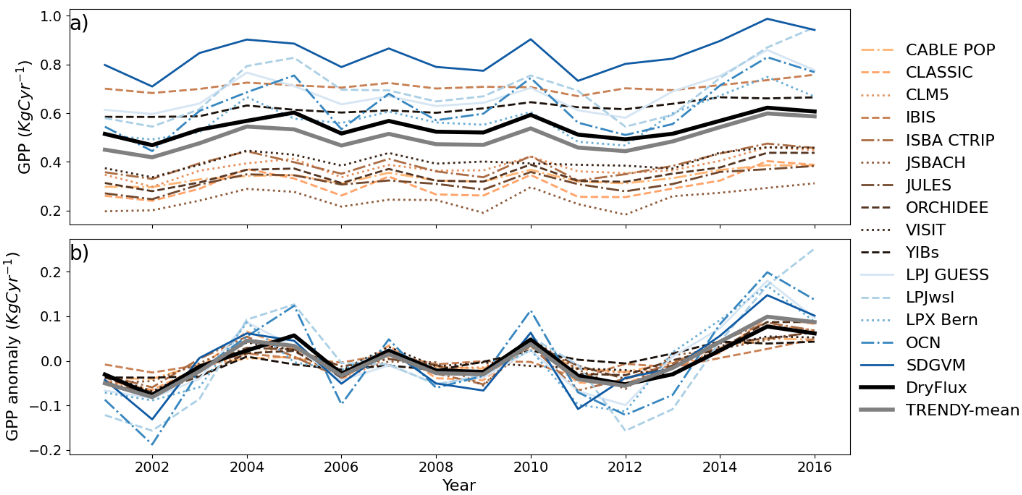

Figure 2: Time series of (a) total annual GPP ($PgCyr^{-1}$) over the study area; and (b)annual GPP anomalies (i.e., annual values minus the mean to focus on GPP IAV) between 2001 to 2016 from TRENDY v11 models (blue and brown curves) compared to DryFlux v1.0 (thick black curve). The mean of the TRENDY models is shown in the thick grey curve. Note that models with the blue curves are those that tend to have higher annual GPP variability (as seen in the anomalies) than DryFlux.

The spatial distribution of standard deviation in annual GPP (Fig. 3a) and slope of the linear regression between model and DryFlux annual GPP across the study area are shown in Figs. 3a and 3c. These are two metrics that represent GPP IAV. As seen in the total annual GPP time series (Fig. 2), most models generally have a lower standard deviation in annual GPP (top two rows of Fig. 3a) compared to DryFlux across much of the study area except the east and southeastern region (see below) and slope values less than 1 (top two rows of Fig. 3c), while several models (LPJ-GUESS, LPJwsl, LPX-Bern, OCN, and SDGVM) generally have higher standard deviation in annual GPP and slope values greater than 1 (bottom row of Figs. 3a and c – although SDGVM and LPX-Bern have lower annual GPP variability compared to s DryFlux in the north-eastern and western regions, respectively).

The spatial pattern of high standard deviation of annual GPP in the east and southeastern part of the study area in the DryFlux product is captured well by most models, although the magnitude is higher than

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

DryFlux for many models (Fig. 3a). In terms of magnitude there is variability among models too e.g. the slope values are often quite a bit higher (up to 3) or lower (down to 0) across most models for this SE region (Fig. 3c). This region corresponds to an area with high annual rainfall variability (data not shown). High values of standard deviation in annual GPP in this region do not necessarily correspond to high mean annual GPP values (Fig. S2). IBIS has a high mean GPP in the southeast region but has very low standard deviation in annual GPP and CLASSIC, CLM5, JULES, ORCHIDEE do not have as high a mean annual GPP as IBIS in that region but those models do have higher magnitude of annual GPP standard deviation (cf. Fig. 3a and Fig. S3).

The Kernel density estimation (KDE) ridgeline plots, which are the probability density functions of temporal standard deviation of annual GPP across space, corroborate the spatial patterns: the majority of models (in the top two rows of Fig. 3a and light purple KDE plots in Fig. 3b) tend to be skewed with peaks towards lower values and longer tails towards higher values of standard deviation in annual GPP compared to DryFlux that looks more like normal distribution although with longer tails towards higher values. For the same models the majority of grid cells have slope values less than 1 (brown KDE plots in Fig. 3d). In contrast, the KDE plots for the LPJ models, OCN and SDGVM generally have a much wider distribution of annual GPP standard deviation and slope values (dark and light blue KDE plots in Figs. 3b and d). While the mode(s) of the standard deviation KDE plots generally match the mode of the DryFlux KDE plot for these models, a greater number of grid cells have higher standard deviation in annual GPP compared to DryFlux (Fig. 3b). The mode of the slope KDE plots is ~1.0 for LPJ-GUESS, LPX-Bern and SDGVM, while LPJwsl and OCN have distributions that are skewed more towards higher slope values (>1.0 with modes ~2.0). Among all the models, LPJ-GUESS is the model that best captures the variability of the DryFlux GPP as evidenced by comparing the spatial pattern of annual GPP standard deviation (Fig. 3a), the standard deviation KDE plot that best matches DryFlux (Fig. 3b), the majority of slope values close to 1 (Fig. 3c) and the most constrained slope KDE plot with a mode of ~1 (Fig. 3d). LPJwsl, LPX-

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Bern, SDGVM and OCN all have a much higher annual GPP standard deviation and high slope values (often exceeding 2) in the eastern or southeastern region of the study area (Figs. 3a and c).

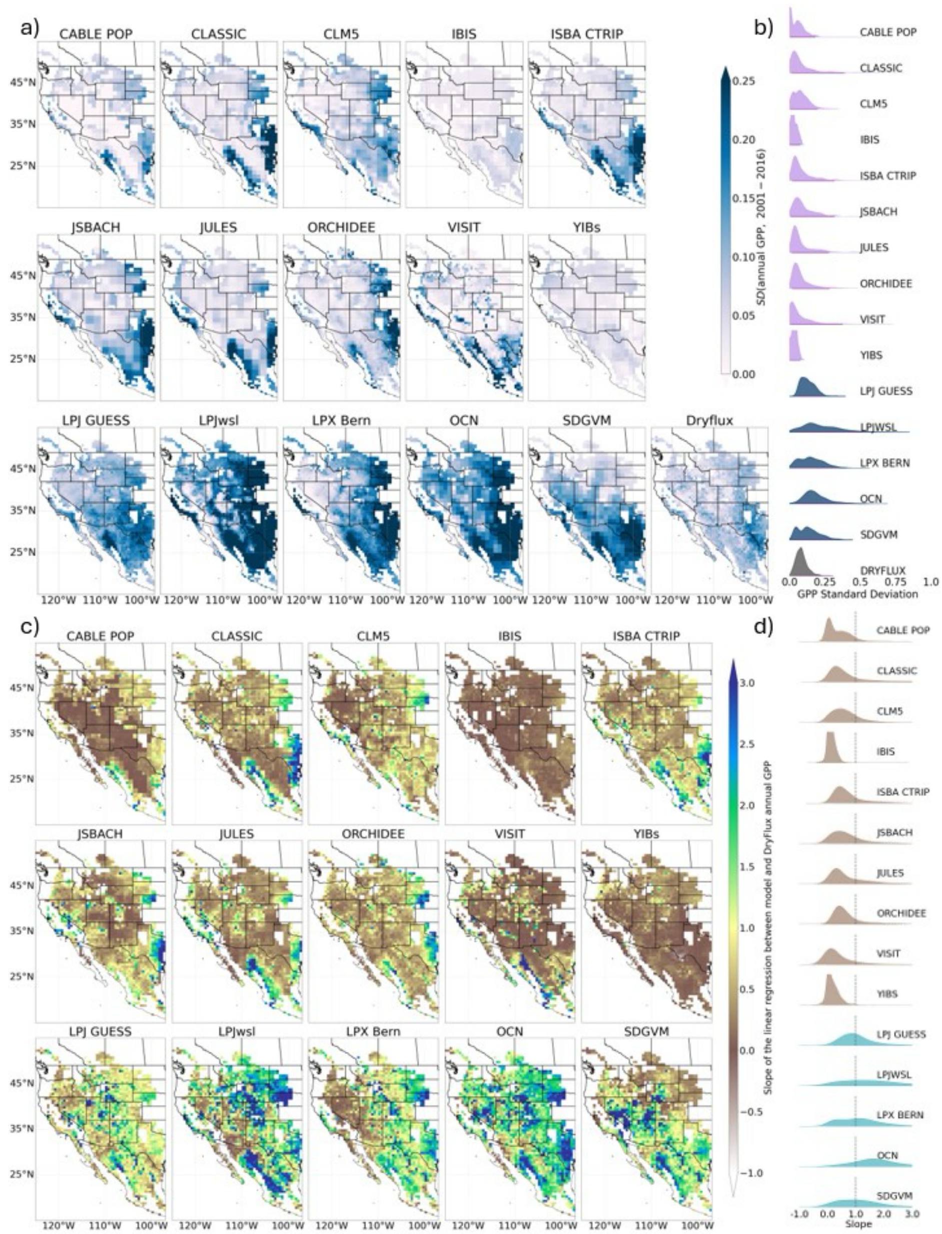

Figure 3: Spatial distribution across the western North American study area for: a) standard deviation of annual GPP (2001-2016) for all TRENDY v11 models used in this study and the DryFlux data; b) a ridgeline plot of the Kernel

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Density Estimates (KDE) representing the area weighted spatial distribution of GPP standard deviation for each of the TRENDY models and DryFlux; c) and d) same as a) and b) but for the slope of the linear regression between model and DryFlux annual GPP. The vertical dashed line in the slope ridgeline KDE plots shows a slope value of 1.0. The top two rows of (a) and (c) and the light purple and brown KDE plots in (b) and (d) correspond to models that tend towards lower standard deviation values than DryFlux or which have slope values of the linear regression between each model and DryFlux that are much less than 1. The bottom row of maps in (a) and (c) and the dark and light blue KDE plots in (b) and (d) correspond to models that tend to have higher standard deviations and a wider range of slope values.

3.2 Categorising models based on dynamic vegetation and fire

Regrouping the KDE ridgeline plots into dynamic vegetation enabled and fire enabled models, Figs. 4a and b, respectively) provides a first visual evaluation of the impact of each of these processes on model performance. Grouping the models according to whether fire was included did not help to reveal the cause of the differences between the two model groups. Models that did not use dynamic vegetation are mostly underestimating GPP variability compared to DryFlux with the exception of OCN and SDVGM. Of the 4 models that included dynamic vegetation and fire, one model generally had low GPP IAV compared to DryFlux (JULES – Figs. 2 and 3), while the 3 LPJ models tended to have higher GPP IAV compared to DryFlux (Figs. 2 and 3). Thus, our hypotheses related to the inclusion of dynamic vegetation and fire may be partially correct, but the full picture is more complex. Neither of these processes alone, or in combination (the 4 dynamic vegetation enabled models also included fire), is able to explain why models are underestimating or overestimating the DryFlux GPP IAV (Fig. 4). However, with the exception of JULES, including dynamic vegetation and fire may explain why the LPJ models have higher GPP IAV compared to models without either of these processes, as we hypothesised. In contrast, fire alone (in models without dynamic vegetation – Fig. 4b) is not causing models to have higher GPP IAV (possibly with the exception of SDGVM), contrary to what we predicted. In the following sections we examined differences across models in terms of their spatial patterns in temporal variability in fractional cover of

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

the major plant functional types and the burnt area for fire enabled models. We also explored whether the spatial distribution of dominant plant functional types PFTs can be linked to spatial patterns in GPP variability.

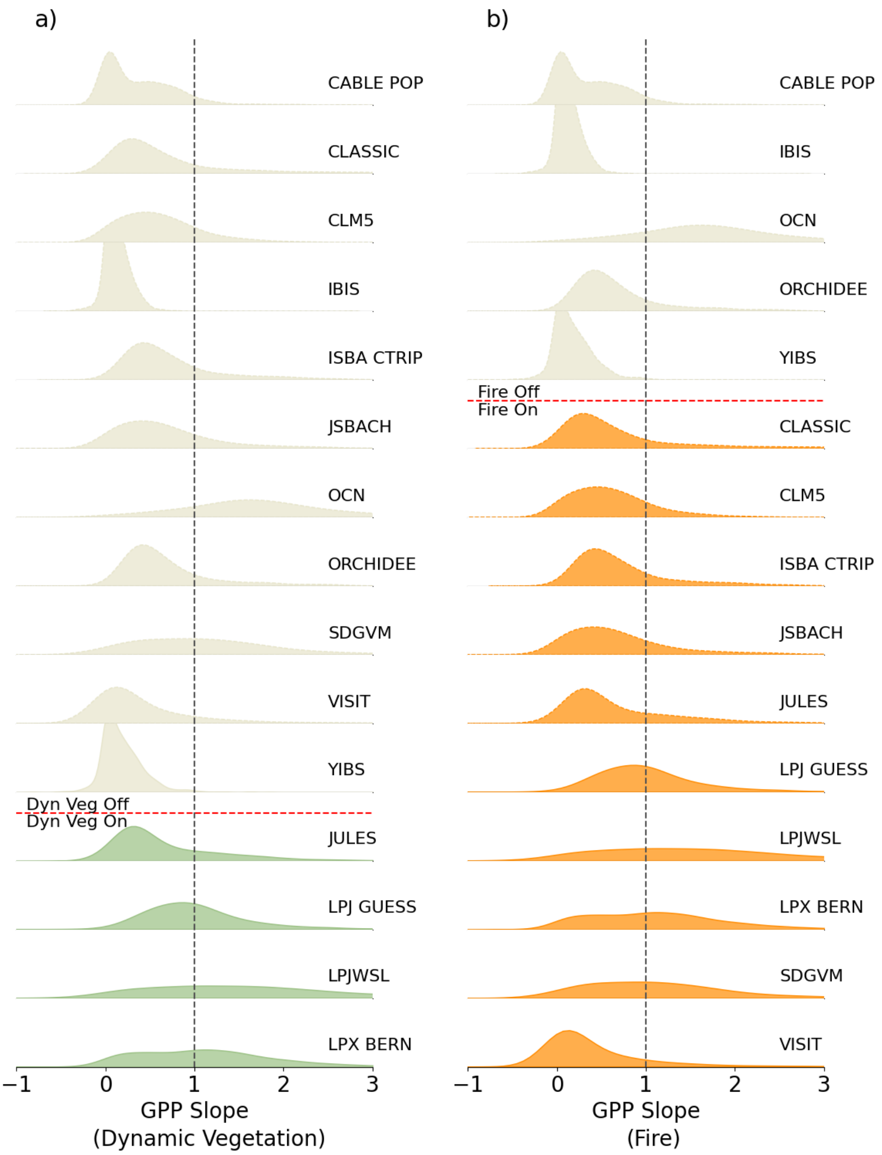

Figure 4: Ridgeline KDE plots summarizing the distribution of slope values of the linear regression between model and DryFlux annual GPP across the entire study area grouped by: a) dynamic vegetation enabled models (green

KDE plots); and b) fire enabled models (orange KDE plots) versus models that do not contain those processes (beige KDE plots) for the models. Note that all models that simulate spatial distribution of fractional vegetation cover dynamically also include fire in their simulations.

3.3 Spatial patterns in PFT fractional cover and burnt area annual variability

As expected based on the results in Section 2.2, models with dynamic vegetation (JULES and the LPJ family of models) have higher standard deviation in PFT annual fCover compared to other models, while all models that included fire but no dynamic vegetation (e.g., CLASSIC, CLM5.0, JSBACH, ISBA-CTRIP, and VISIT) showed no changes in vegetation fCover of vegetation, with the exception of ISBA-CTRIP – Fig. S3a to S3g). For all dynamic vegetation enabled models, which also included fire, we found that temporal variability of annual vegetation fCover – expressed as the coefficient of variation (CV, i.e., standard deviation in annual total PFTcover d divided by the mean) – highly correlated with the temporal variability (CV) of annual GPP (Fig. 5). Among these models, LPJ-GUESS, which captures better the spatial patterns in GPP IAV compared to other models (Fig. 3), has the highest R^2 value between annual total PFT fCover and annual GPP CV, with LPX-Bern a close second. With increasing annual total PFT fCover CV LPJwsl showed an increased spread in GPP IAV. JULES appears to have more of a nonlinear increasing concave down relationship between annual total PFT fCover and annual GPP CV, which implies increases in annual GPP variability (expressed CV) may be limited above an annual total PFT fCover CV of around 0.05.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

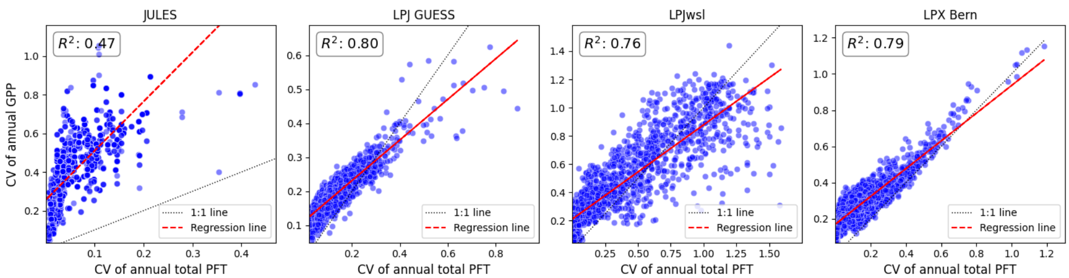

Figure 5: Scatterplots for the four dynamic vegetation enabled TRENDY models (i.e., those that simulate changes in spatial distribution of PFTs over time) showing the pairwise grid cell comparisons of the coefficient of variation (CV; standard deviation divided by the mean) of the total PFT annual cover versus the annual GPP calculated over the 2001 to 2016 time series for each grid cell. CV is used here instead of the standard deviation to directly compare temporal variability of two different quantities.

The standard deviation in annual total PFT fCover in the dynamic vegetation enabled models is mostly from the herbaceous PFTs (cf. 1st 4 columns showing the models in top and bottom rows of Fig. 6), which matches the two independent reference vegetation fCover datasets – MODIS VCF and RAP (last two columns in Fig. 6). In JULES, the spatial patterns of standard deviation in annual PFT fCover match both the MODIS VCF and RAP products (cf. Fig. 6 left column with the rightmost two columns). However, these spatial patterns do not correspond to the spatial patterns seen in JULES standard deviation of annual GPP (Figs. 3a and c). All variants of the LPJ model have higher standard deviation in annual vegetation fCover compared to MODIS VCF and RAP and the spatial patterns also do not match well (Fig. 6). The models appear to be overestimating standard deviation in annual PFT fCover more in the arid central and southwestern part of the study area. These regions typically have sparse dwarf shrubs and higher bare soil cover compared to models (see RAP in Figs. S1a and b). LPJwsl also has a higher standard deviation in annual woody plant fCover in a north central and eastern region of the study area (and high mean woody plant fCover – Fig. S1b) that is not consistent with the two independent data products (Fig. 6). All variants of the LPJ model also predict higher mean and standard deviation in annual grass cover compared to the RAP product (Fig. 6 and S1d). The higher standard deviation in annual PFT

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

fCover in LPJ models when compared to the independent datasets could help explain why those models also overestimate standard deviation in annual GPP (with slope values much greater than 1). However, the spatial patterns in standard deviation in annual PFT fCover (Fig. 6) do not match completely the spatial patterns in standard deviation in LPJ annual GPP (Fig. 3a), with the exception of LPJ-GUESS.

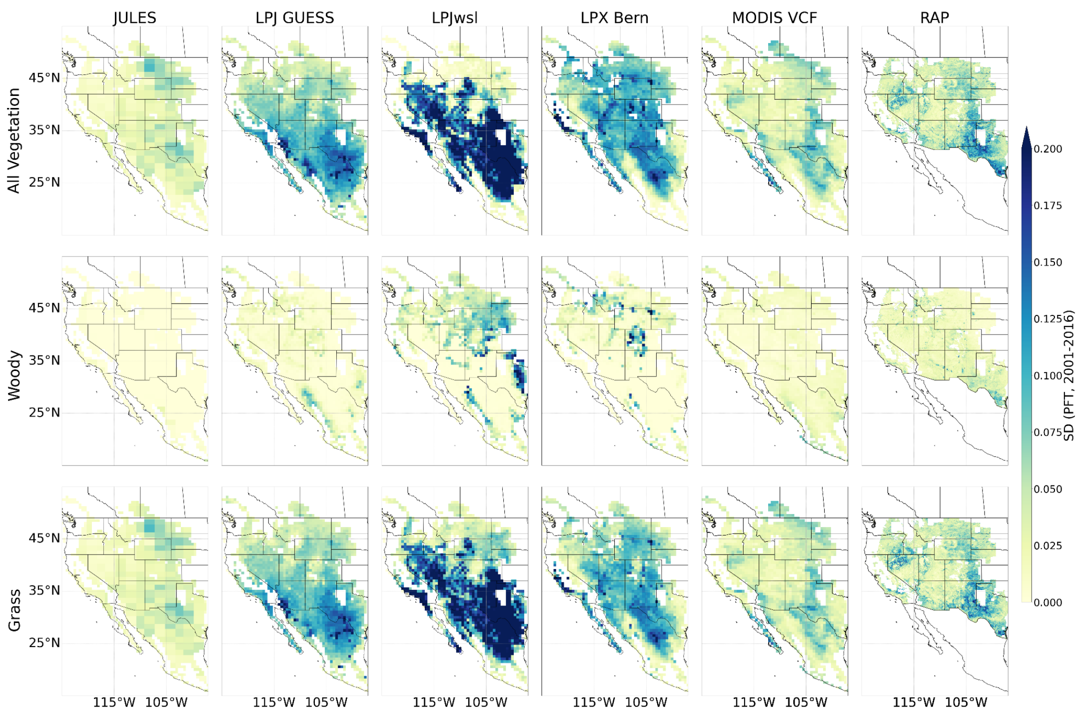

Figure 6: Standard deviation in annual fCover from 2001 to 2016 of all vegetation (all PFTs – top row), woody vegetation (second row), and grasses (bottom row) for dynamic vegetation enabled models compared to independent reference data from MODIS VCF and RAP.

Only four of the fire-enabled models simulate a standard deviation of annual percentage burnt area greater than 10% for a number of grid cells (CLASSIC, SDGVM, VISIT and LPJwsl); all other models simulate almost no variability in percentage burn area variability (Fig. 7). Note that the spatial patterns in the

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

standard deviation of annual burnt area correspond well with the mean values, with almost all models except SDGVM simulating mean burnt area less than 20% for the majority of grid cells (data not shown). In all fire-enabled models, the spatial patterns in the standard deviation of annual burnt area did not correspond well to spatial patterns in the standard deviation in annual PFT fCover (cf. Fig. 7 with Fig. 6), with the possible exception of the southwestern region of high standard deviation in annual burnt area for LPJwsl. Therefore, while fire may be playing a role in annual variability of PFT fCover and GPP for LPJwsl (Fig. 8), it cannot be the only cause of the much higher standard deviation in annual GPP (and high slope values) compared to DryFlux. The remaining three models (CLASSIC, SDGVM and VISIT) do not model the spatial distribution of vegetation cover dynamically; therefore, despite simulating noticeable variations in annual burnt area this did not affect their year to year changes in PFT cover (Fig. S3). However, these three models did have spatial patterns in standard deviation of annual burnt area that corresponded reasonably well with spatial patterns in standard deviation in annual GPP (cf. Figs. 7 and 3a). This was especially the case for SDGVM, which has the simplest representation of fire (Figs. 7 and 3a). These results suggest that for these three models, fire may be a driver of GPP variability (Fig. 8), as we predicted (albeit that CLASSIC and VISIT still underestimate DryFlux annual GPP variability overall). The remaining fire-enabled models (ISBA-CTRIP, JSBACH, JULES, and LPJ-GUESS) have very low standard deviation in annual burnt area (Fig. 7). (Note LPX-Bern did not output burnt area; however, they use the same fire module,GlobFIRM, as the other LPJ models.) Thus, for these models, fire does not correlate well to standard deviation in annual GPP (Fig. 8), contrary to what we predicted. Furthermore, the burnt area results dispute our hypothesis that fire enabled models with dynamic vegetation would result in higher grass fCover; instead; another process or model input must be driving high standard deviation in grass fCover in the dynamic vegetation enabled models.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

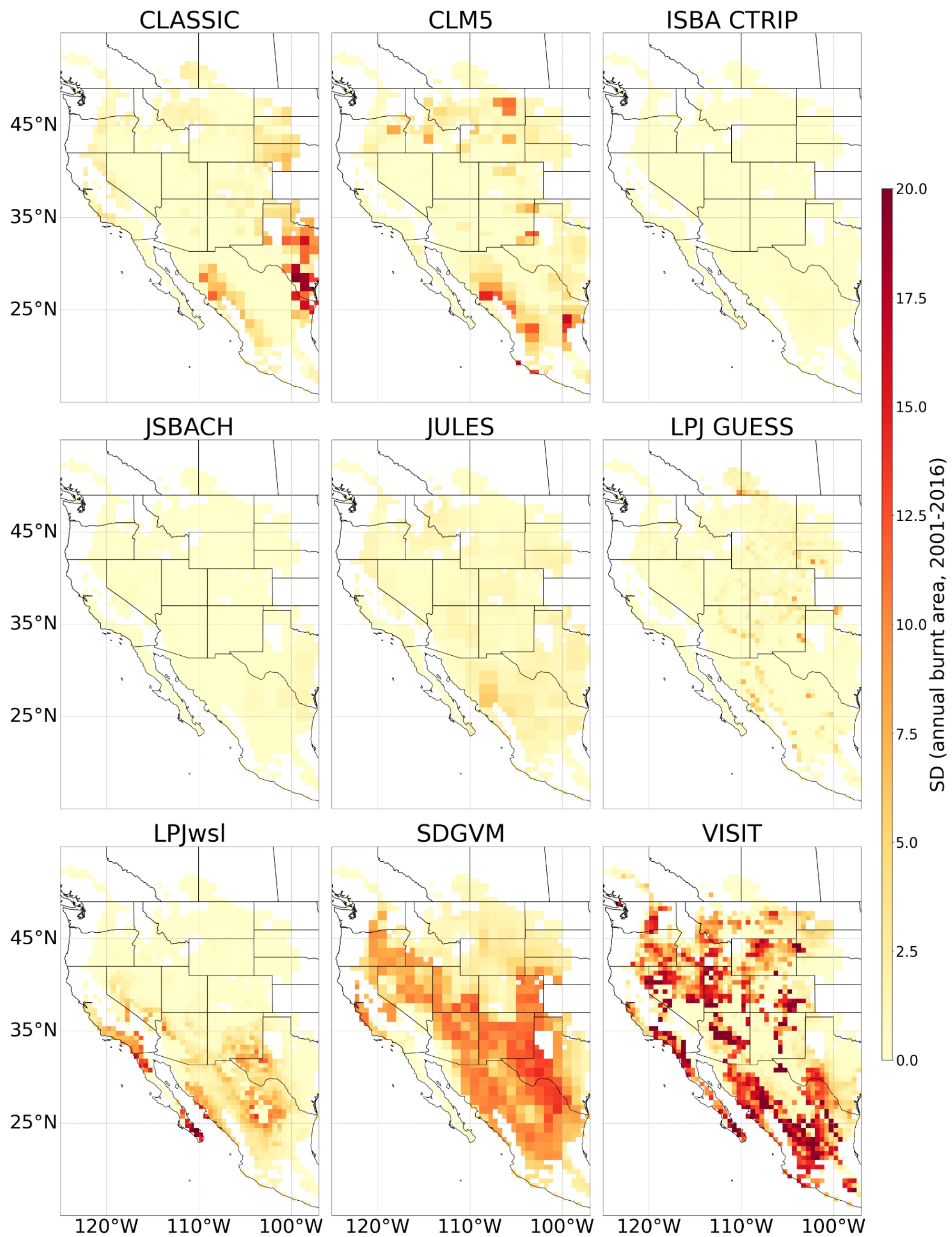

Figure 7: Spatial distribution across the western North American study area of the standard deviation in annual burnt area (%) for each grid cell for the 2001 to 2016 period for fire enabled TRENDY models.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

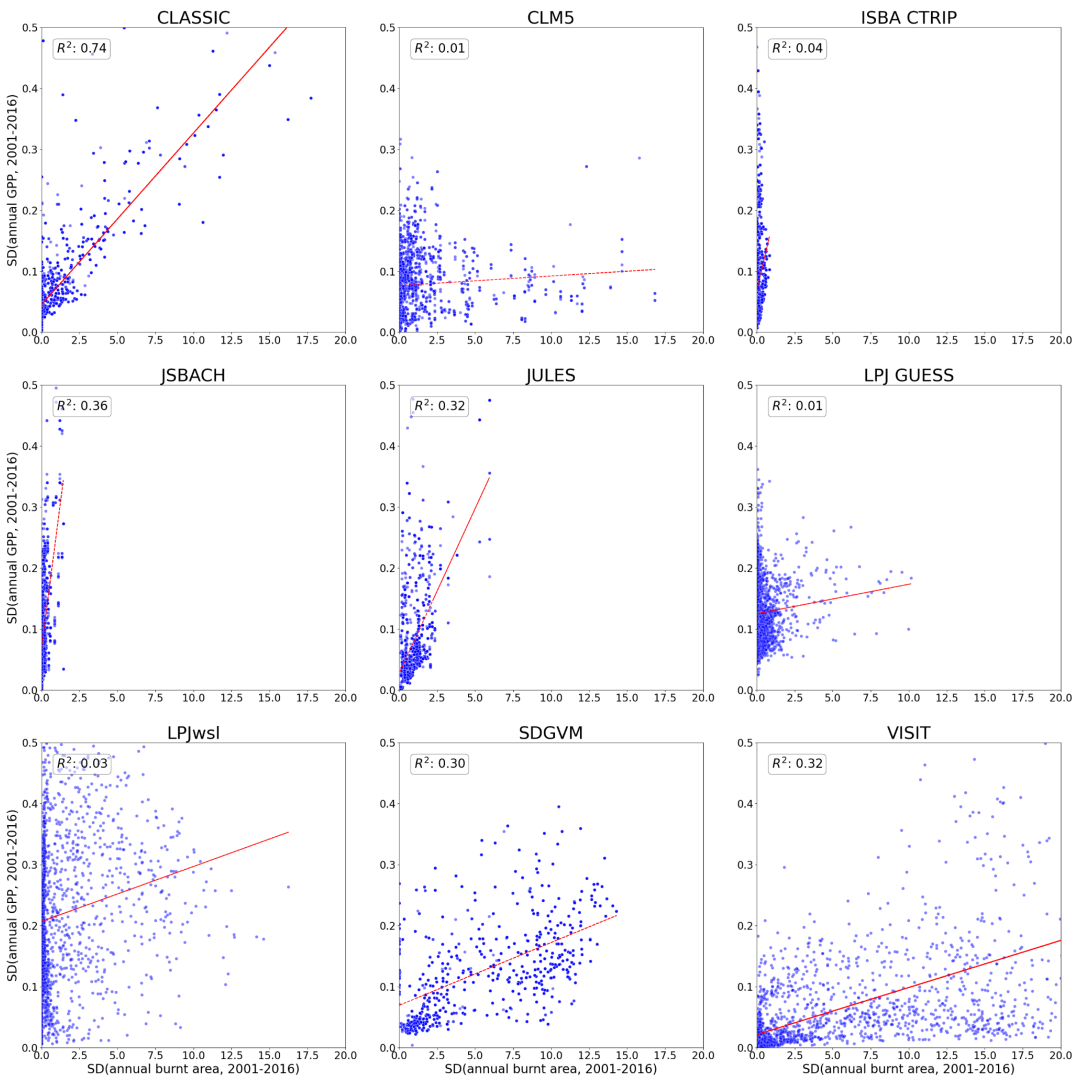

Figure 8: Scatterplots showing for all the fire enabled TRENDY models the pairwise grid cell comparisons of the standard deviation of the total annual burnt area versus the annual GPP calculated over the 2001 to 2016 time series for each grid cell.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

3.4 Relationship between dominant vegetation type and annual GPP variability

In addition to the contribution of dynamically modelled vegetation cover and fire to high annual GPP variability, we hypothesized that model performance in capturing annual GPP variability compared to the reference DryFlux product may be related to the dominant vegetation type. Compatible with our hypothesis that models with high mean grass fCover across the study site will have higher standard deviation in annual GPP compared to models with more shrub or bare ground cover, we found no clear linear relationship between the mean fCover of woody plants for selected grid cells across the study area and the mean slope of the linear regression between each model and DryFlux annual GPP (Fig. 9a). This was the case both when grid cells with a woody fCover greater than 0.1 were selected (Fig. 9a) and grid cells with a woody fCover greater than 0.5 (Fig. 9a inset). Models with both high and low woody fractions underestimate model-DryFlux annual GPP slope values (spatial mean less than 1.0), although it is possible that there is a decreasing concave up relationship in which models with generally low woody fCover have higher GPP variability (slope values greater than 1 – Fig. 9a). Contrary to our hypotheses, the spatial mean grass (C3 plus C4) and C3 grass fCover for different models show no linear relationship with the spatial mean slope of the linear regression between model and DryFlux annual GPP when grid cells with a fCover of greater than 0.1 or 0.5 were selected (R^2 values ≤ 0.06; Figs. 9b and c and insets). However, there is a much stronger positive linear relationship between spatial mean C4 grass fCover and the spatial mean slope of the linear regression between model and DryFlux annual GPP (R^2 0.63 for grid cells with greater than 10% C4 grass fCover and 0.35 for grid cells with greater than 50% C4 grass fCover; Fig. 9d and inset). Models with low C4 grass fCover tend to have low annual GPP model-DryFlux slope values (slopes less than 1; brown points in Fig. 9d) and models with high C4 grass fCover (especially LPX-Bern, OCN and SDGVM) have higher slope values (slopes greater than 1; blue points in Fig. 9d). It is also worth noting that most of the models with higher C4 grass fCover also had a low

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

woody plant cover (low spatial mean and small number of grid cells – Fig. 9a and inset). For these three models (LPX-Bern, OCN and SDGVM), the regions of high mean C4 grass fCover (Fig. S1f) correspond well with regions of high standard deviation in annual GPP s (Fig. 3a) and model-DryFlux slope values much greater than 1 (Fig. 3c). In contrast, regions of high mean C3 fCover do not correspond to regions of high standard deviation in annual GPP for any model (including all the models that tend to overestimate GPP IAV (bottom row of Figs. 3a and c). Therefore, in general our hypothesis that models with higher grass cover would have higher GPP variability was incorrect. However, for some models including LPX-Bern, OCN, and SDGVM, their high C4 grass fCover (and/or a too high response of C4 grasses to climate variability in these models) may be a piece of the puzzle in explaining why they are overestimating DryFlux GPP IAV.

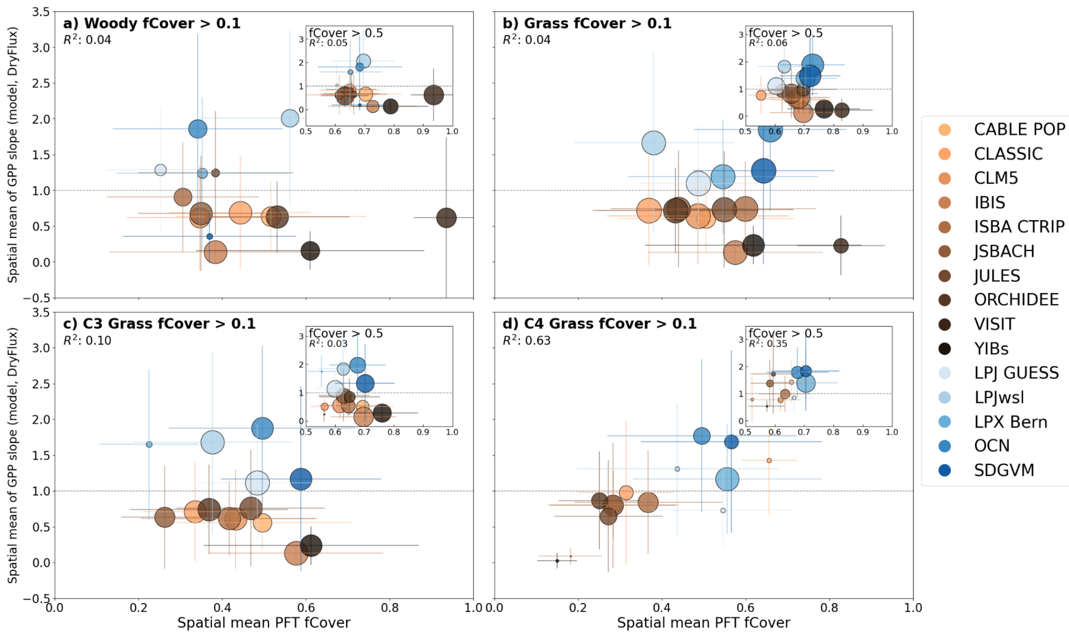

Figure 9: Scatter plots showing the spatial mean (mean across the study area for selected grid cells) fractional cover (fCover) for a) woody, b) grass, c) C3 grass and d) C4 grass cover greater than 0.1 (or 0.5 in the inset plots) versus the spatial mean of GPP slope of the linear regression between model and DryFlux annual GPP for the grid cells with fCover greater than 0.1 (or 0.5). For each scatter plot the point size is proportional to the number of grid cells, the horizontal error bar represents 1 s.d. spread from the mean fCover and the vertical error bar represents 1 s.d.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

spread from the spatial mean GPP slope across selected grid points. R^2 values in each plot represent the proportion of variance in the spatial mean PFT fCover across models explained by the spatial mean in slope between the annual model and DryFlux GPP .

We compared each model's mean fCover for each PFT category (Fig. S1) with the two independent reference fCover datasets. For this purpose RAP data is more reliable than MODIS VCF for both grasses and woody categories because RAP grass cover is a combination of grass and forbs but not crops; therefore, RAP is a better dataset against which to benchmark grass fCover. MODIS VCF non-woody cover contains all vegetation with height less than 5m; therefore we consider it is suboptimal for evaluating either woody or non-woody cover. Of all the models that tend to have high standard deviation in annual GPP and model-DryFlux slope values much greater than 1 (LPJs, OCN and SDGVM), the models tend to have higher mean grass cover than RAP and the spatial patterns in mean grass cover did not match patterns seen in the RAP product (Fig. S1d). RAP has lower grass cover in the more arid central and southwestern region compared to the models (Fig. S1d) and higher mean woody plant cover (Fig. S1b). (Note that MODIS VCF has a lower woody plant cover that matches the models well; however, MODIS VCF woody cover represents trees greater than 5m). In contrast, many of the models that have lower standard deviation in annual GPP and model-DryFlux slope values much less than 1 in the central and southwestern arid regions tend to have higher mean woody plant cover (or higher bare soil cover) in that part of the study area (Fig. S1b). This could suggest that to simulate higher GPP IAV these models need higher grass cover and lower shrub/bare soil cover. However, LPJwsl has more woody cover in the northern region than seen in the reference datasets or in other models (Fig. S1b), yet still overestimates GPP IAV in this region (Figs. 3a and c). These results lend support for our hypothesis that increased grass cover results in higher GPP variability, with too high grass cover (especially C4) partially resulting in too high GPP IAV.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

4. Discussion

4.1 Implications of uncertainties in DryFlux

The random forest model used to produce the DryFlux product does not include a fCover data nor does it represent different plant functional types (see Section 4.2) and thus cannot be biased by discrepancies in vegetation fractional cover. Instead, DryFlux used satellite derived vegetation greenness indices such as NDVI and EVI as a proxy for vegetation green distribution for photosynthesis. As Smith et al. (2018) showed, these indices do not always capture reductions in GPP in the hot, dry pre-monsoon season because in dryland ecosystems woody plant leaves often remain green during this period; therefore photosynthesis and green leaf area can often be decoupled. This could result in an overestimation of DryFlux GPP during that period, and thus a slight positive bias in DryFlux annual GPP. However, as the DryFlux product performed well in capturing the decline in GPP across southwestern US sites during the pre-monsoon season (Barnes et al., 2021) we are not expecting such a positive bias to occur. Another potential issue with DryFlux is that it may not be as reliable in crop-dominated areas because the random forest model did not include crop sites.. However, this should not affect our analysis as we mask out crop-dominated grid cells from our study area. Finally, a recent study showed that dryland GPP is more sensitive to soil moisture than atmospheric drivers such as vapour pressure deficit or air temperature (Kannenberg et al 2024). DryFlux v1.0 did not consider soil moisture in the upscaling model; however, inclusion of soil moisture at different depths has been tested in a later DryFlux version (Pervin et al., in prep.). As the results of this study show, inclusion of surface soil moisture improved DryFlux GPP accuracy by decreasing RMSE and increasing R2 in both the daily and monthly predictions when compared to in situ GPP data and by helping to capture the bi-modal annual cycle and the magnitude of mean annual GPP across some western North American sites (Pervin et al., in prep). However, inter-

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

annual variability in DryFlux predicted annual GPP is captured well by both versions. Therefore, we are not expecting any drastic change in the comparisons between DryFlux and TRENDY models.

4.2 Model-data uncertainties in PFT distributions

The findings of this study suggest that models without dynamic vegetation tend to have higher shrub versus grass cover, which may contribute to an underestimate in GPP IAV, while models with high prescribed C4 grass cover may overestimate GPP IAV. Models that do not simulate changes in PFT fractional coverage dynamically have their own model-specific methods (including dataset used) for generating spatial distribution of PFTs' fractional coverage. Therefore, mapping remotely sensed vegetation or land cover data to a given model's PFTs is a largely subjective and expert judgement based process, especially in sparsely vegetated dryland regions (Hartley et al., 2017). This can lead to significant variation in PFT fractional coverage maps across models even when they are based on the same underlying remotely sensed data set (Hartley et al., 2017).

Model specific decisions include methods for separating grasses into their C3 and C4 variants (Table S1). Some models define C4 grasses as those in a tropical climate zone based on Köppen Geiger climate zones or ECOCLIMAP (Table S1) and C3 grasses in all other climate zones. Other models use a mean temperature threshold for separating C3 versus C4 grasses; however, the temperature thresholds vary widely between models. For example CABLE POP, LPJwsl, and LPJ-GUESS use a temperature threshold of 15.5°C for the coldest month to separate out C3 and C4 grasses. CLASSIC and SDGVM use the C3/C4 distribution map from Still et al., (2003) and the ESA CCI Land Cover product (Poulter et al., 2011), respectively – both of which follow the study of Collatz et al. (1998) that specified 22°C as a "critical crossover" temperature for C3 versus C4 grasses (Table S1). Note that Still et al. (2003) allows C4 grasses to prevail when temperatures exceed 22°C in any month, while Poulter et al. (2011) specified this temperature threshold for the hottest month. Still et al. (2003) also requires a minimum amount of

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

precipitation (25mm) and assesses whether the temperature and precipitation conditions are met on a seasonal basis, thus permitting mixed C3/C4 grasslands in temperate regions such as our study area. The models with a temperature threshold of 15.5°C for the coldest month have no C4 grass in the study area (Fig. S1f), which is not the case in reality (Section 2.1). In fact, neither a single temperature threshold defined for the hottest or coldest month nor the use of climate biomes to separate out C3 versus C4 grasses is sufficient as both types exist in the southwestern part of our study area: C3 grasses grow in winter, while C4 grasses grow in summer (Ehleringer, 1978).

Large discrepancies in dynamic vegetation enabled modeled C4 grass cover were documented by Still et al. (2019) in comparison to independent proxy data across North America. Their study noted the importance of representing seasonal offsets in both C3 and C4 grasses, in addition to recommending improvements in modeled responses to disturbance, tree-grass competition for water resources, and separation of woody and herbaceous cover in remotely sensed data. The latter is hard to achieve even when LiDAR derived height data are included in the spectral unmixing algorithm (Pervin et al., 2022). Wilcox et al. (2023) further noted the need to differentiate between annual and perennial grasses in DGVMs. However, aside from independent proxy data, there are currently no detailed maps of C3 versus C4 grass distributions for this study region that could help to benchmark models for dynamic vegetation enabled models or to improve PFT input maps for models with prescribed, static PFT fCover. A new approach utilizing the differences in timing of peak growth in C3 versus C4 grasses has been used to map C3 and C4 grass distributions across Australia (Xie et al., 2022). Such an approach should be utilized to create C3/C4 grass distribution maps for other regions such as western North America. There is now a wealth of *in situ* PFT fCover estimates collected at NEON sites that could be used to validate such a product (NEON fCover product: https://data.neonscience.org/data-products/DP1.10058.001). Luo et al. (2024) have applied photosynthetic optimality theory combined with satellite remote sensing data to produce global maps of C4 distributions. Dryland GPP simulations utilising the new product from Luo et al. (2024) should be tested against the default model PFT maps for this study region.

Even when data are available, comparisons with independent reference fCover datasets can be challenging. Vegetation fCover estimates provided in most remote sensing derived classification maps and field estimates is essentially a spatio-temporal average, whereas models need the fCover for the most optimal climate conditions. Models should then limit growth during periods of non optimal climate conditions. One possible option would be to use the year of maximum fCover from products such as RAP to prescribe PFT fCover in models without dynamic vegetation, and compare that to simulations using the standard "spatio-temporal average" PFT fCover approach.

Other issues complicate comparisons of model prescribed or simulated PFT fCover. MODIS VCF is one of the most widely used products for evaluating model PFT fractions (e.g., in Teckentrup et al., 2021); however, as described in Section 2.3.2.2, tree cover in MODIS VCF is defined as trees with a height greater than 5m. In this study region (and many dryland regions) most shrubs at low elevation are less than 5m in height; therefore, we expect MODIS VCF will underestimate true woody cover in dryland regions. Similarly we expect MODIS VCF non-woody cover will overestimate true herbaceous vegetation as it incorporates all vegetation with height less than 5m, which will include small stature woody shrubs present in drylands. Therefore, MODIS VCF is not an optimal product for benchmarking modeled dryland vegetation fractional cover. RAP herbaceous cover is more reliable for providing grass fCover estimates as it incorporates both annual and perennial grasses and forbs. The only drawback of the RAP product is that it does not consider any crop cover types and therefore these have to be removed from a PFT fCover evaluation analysis. RAP also includes both tree and shrub cover layers, which can be summed to provide total woody cover. One additional method for benchmarking dryland woody cover would be to use high resolution datasets of tree crown area (e.g., Brandt et al., 2020) that are becoming more widely available as computational power increases.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

4.3 Implications of inaccurate dynamic vegetation and fire representation

As expected, the lack of dynamic vegetation and fire helps to explain why some models have low GPP IAV compared to DryFlux. However, when analyzed spatially, models that do include a representation of dynamic vegetation and/or fire do not always simulate higher GPP variability across most of the study area, contrary to our hypotheses. Spatial patterns of burnt area variability generally did not match patterns of variability in vegetation fCover for any model (with the slight exception of LPJwsl). This is to be expected for models that do not simulate the spatial distribution of PFTs' fractional coverage dynamically: in these models fires burn the existing biomass of model PFTs and reduce their per unit area biomass density, but do not change the PFTs' fractional cover. However, no change in fCover as a result of fire is not realistic in most cases, thus highlighting the need for models to include both fire and dynamic changes in PFT distributions when modeling dryland ecosystems.

All the dynamic vegetation enabled DGVMs included in this study also included a representation of fire; however, most of these models have low standard deviation in annual percentage burnt area (Fig. 7) which does not represent reality. Fire does occur in western North American dryland ecosystems (McGranahan and Wonkka, 2024; Singleton et al., 2019) though more severe fires typically occur in the higher elevation more shrub to tree dominated ecosystems (Singleton et al., 2019). For models that included dynamic vegetation and fire we predicted that fire would contribute to the prevalence of grass dominated ecosystems (Burton et al., 2019; Verbruggen et al., 2021). However, JULES and all LPJ models simulated very low standard deviation in annual percentage burnt area (Fig. 7) and high grass cover; therefore, for these models it is clearly not fire that is controlling grass dominance. In the S3 simulations pasture covers much of these areas (likely too much based on our knowledge of the study area), but even the LPJ S2 simulations without a pasture land cover class predict high grass cover (data not shown). Further research is needed to ensure that processes controlling grass dominance in the LPJ models, including competition between herbaceous and woody plants, are accurately represented.

The processes that result in high grass cover in the dynamic vegetation enabled DGVMs, including the bioclimatic envelopes defined for each PFT and/or the competition conditions set for each model (Harper et al., 2018; Sitch et al., 2003; Smith et al., 2001), could in turn produce aboveground or litter biomass that is below the minimum fuel load for fire to spread (e.g., 200gCm^{-2} in GlobFIRM – Thonicke et al., 2001; INFERNO – Mangeon et al., 2016; and CLASSIC – (Melton and Arora, 2016)). However, most models simulated litter biomass values greater than this threshold across this region (data not shown); therefore it is not the lack of fuel that is causing low fire occurrence across models. Predicted moisture conditions in the upper soil layers that control flammability may be too high during the hot, dry pre-monsoon period when fires are most likely. Models with a simple bucket layer hydrology scheme (which includes LPJwsl, SDGVM, and LPX-Bern in this study) have been shown to overestimate moisture during the hot, dry period of May and June at sites across the southwestern US compared to more mechanistic soil moisture schemes (MacBean et al., 2020). Even if models are simulating small fires during this period, the impact on annual burnt area, which is a function of length of fire season in GlobFIRM (Thonicke et al., 2001) or prescribed as in INFERNO (Mangeon et al., 2016), may be minimal. A more extensive validation of fire burnt area fraction and its seasonal and inter-annual variability simulated by fire-enabled DGVMs is needed. However, given the complexity of different fire model schemes and the finer spatial and temporal resolutions at which fires occur compared to model grid cell resolution, a more extensive validation is beyond the scope of this study.

For many of the models that simulate higher standard deviation of annual percentage burnt area (CLASSIC, SDGVM and VISIT) there was a reasonable relationship with GPP variability (R^2 values > 0.3; Fig. 8). This finding suggests that fire may play an important role in improving simulations of GPP IAV if it was better represented. Future evaluations of modeled burnt area should consider grouping

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

models according to the different fire parameterisations to try to disentangle the level of complexity needed. . However, even with the higher predicted standard deviation in annual percentage burnt area CLASSIC, CLM5, and VISIT, their GPP IAV was still generally underestimated compared to DryFlux. The role that fire plays in GPP variability in dryland ecosystems could be tested by running factorial simulations with and without the fire module enabled. This should be a priority in the near term for further diagnosing the causes of dryland C flux model-data discrepancies.

Aside from fire, several other processes that control competition between woody and herbaceous species may explain why dynamic vegetation enabled models in this study tended to overestimate grass cover (both C3 and C4 grasses), with spatial patterns that partially corresponded to high GPP IAV compared to DryFlux. Limited water availability shapes the patchy vegetation patterns often seen in drylands (Munson et al., 2016). Differential water uptake by trees and grasses at different depths in the soil as in the Walter hypothesis (Case and Staver, 2018; Munson et al., 2016) in theory are represented in many DGVMs; however, simplistic rooting profiles implemented for more mesic ecosystems may not be representative of dryland plants. More complex vegetation-water interactions are likely at play in drylands that may not be well represented in DGVMs. For example, tree roots are not always deeper than grass roots (Staver et al., 2017). Furthermore, water in deep soil always passes through the shallow soil providing water access to grasses first. Another hypothesis likely not represented in DGVMs considers that shrubs temporarily shift their physiological activity to access water in later periods (Naito et al, 2011). However, in this case grasses will still use up the water first. This leaves the question if grasses are very opportunistic, how do shrubs survive, especially during drought. The answer could be provided by the hypothesis of positive feedback, where shrub clusters support nearby shrub establishment by increasing soil moisture through interception under the canopy (Scanlon et al., 2007). A field study in Kalahari detected wet soils under canopy and dry soils outside the canopy that supports this hypothesis (D'Odorico et al., 2007). However, such vegetation-moisture interactions will not be resolved by DGVMs with no representation of the

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

spatial distribution of mixed tree-grass ecosystems, no understory, and with soil tiling approaches that calculate soil moisture separately for different major PFT groups (such as trees and grasses).

4.4 Potential impact of other processes controlling GPP IAV

The DGVMs investigated in this study all capture the spatial patterns of GPP IAV fairly well in the SE region of the study area with high rainfall variability, with the caveat that for some models the variability is too strong and possibly linked to high C4 grass fractional cover. Larger model-data discrepancies exist in regions that do not have high rainfall variability and can be linked to dominance of shrub versus grass cover (see Section 4.3). The higher variability in annual GPP seen for C4 dominated DGVMs is an interesting finding given that many previous studies have shown models tend to underestimate C flux variability (Keenan et al., 2012), especially in drylands (MacBean et al., 2021; Wang et al., 2022). C4 grass fractional cover and its change over time needs to be better evaluated in models (see Section 4.2); however, it is possible that there is too strong a response of C4 grasses to changing water availability. This could potentially be due to a range of issues in the model, including inaccurate C4 and other dryland plant traits (parameters;(Mahmud et al., 2021)), unrealistic C4 grass phenology (Whitley et al., 2016; De Kauwe et al., 2017), and unrealistic water stress functions (De Kauwe et al., 2015; 2017), with resultant impacts on photosynthesis and the complex interactions between leaf growth and soil water availability (due to partitioning of intercepted rainfall between transpiration and bare soil evaporation). High variability in GPP is likely linked to processes controlling leaf growth (Lin et al., 2023). Indeed, past studies have highlighted leaf area index and GPP in DGVMs can be too tightly coupled, especially in grasslands and arid regions (De Pue et al., 2022; Lin et al., 2023). Adapting moisture limited phenology schemes in DGVMs to better represent dryland PFTs is likely needed (Dahlin et al., 2015; MacBean et al., 2015; Renwick et al., 2019).

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Finally, we note that we might expect that inclusion of N cycling in models may result in higher nutrient limitation that reduces annual GPP variability; however, in this study there was no apparent difference in GPP IAV with or without N cycling included in the models (Fig. S4). Nonetheless, Smith et al. (2014) noted that N limitation favours grasses in some regions, which could be contributing to high annual GPP variability for dynamic enabled models. It is likely that nutrient cycling plays a role in dryland productivity dynamics that requires further understanding and implementation. More complex interactions between fire, nutrients, and selective grazing in grasslands should also be considered in future dryland model developments (Wilcox et al., 2023).

Even in reality it is impossible to estimate exact contributions of the myriad of processes that contribute to annual variability in ecosystem scale GPP. Doing so will require novel ways to probe or reinterpret existing *in situ* and gridded datasets. For example, future studies could assess non biogeography related process contributions to GPP IAV in DryFlux (and DGVMs) by considering time periods without fire and when vegetation fractional coverages are not changing considerably.

4.5 Implications for other dryland regions

While our study has elucidated differences between DGVMs in the form of high or low standard deviation of GPP in western North America, we predict that these findings may not be consistent across drylands on an intra-model basis. This prediction is based upon a recent study (Bogucki et al. in prep.) which utilized the same TRENDY DGVMs, but on a global scale, finding that a relatively high standard deviation of NEE in one particular dryland region (such as western North America) does not necessarily equate to a high standard deviation of NEE in other dryland regions. Despite the fact that Bogucki et al. (in prep.) investigate a time period spanning 52 years (1970 – 2020) instead of the 16 years (2001 – 2016) used in this study, and their use of NEE instead of GPP, our findings are generally similar. Both Bogucki et al. (in prep.) and this study identify the same DGVMs as either high (LPJ-GUESS, LPJwsl, LPX-Bern, OCN, and SDGVM) or low (top two rows of Fig. 3a) standard deviation within the bounds of western

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

North America. Additionally, all of the models which we have identified as having high standard deviation in annual GPP also display high standard deviation in annual NEE across multiple dryland regions such as the Caatinga, Sahel, and eastern Australia in Bogucki et. al (in prep.). However, not all of the models which both Bogucki et al. (in prep.) and this study have identified as low standard deviation within the bounds of western North America also have low standard deviation in other dryland regions. While some of the models are consistently low standard deviation across dryland regions (CLM5, ISBA CTRIP, JULES, ORCHIDEE, and YIBS), some models display low standard deviation in western North America but high standard deviation in other dryland regions (CABLE POP, CLASSIC, IBIS, JSBACH, and VISIT). Thus, our explanations for why some DGVMs have higher or lower standard deviations in annual GPP may not be applicable to these five DGVMs dryland regions on a global scale. Further studies focusing on other dryland regions such as the Caatinga, Sahel, and Eastern Australia are required to better understand what may be causing this variation between models and between dryland regions within each of the models.

5. Conclusions

We evaluated the ability of 15 DGVMs from the TRENDY v11 model inter-comparison compared with a dryland focused upscaled flux data product (DryFlux v1.0 GPP) to capture the IAV of dryland GPP in western North America. We also examined potential causes of different performances across models, with a specific focus on the role of processes related to biogeography (e.g., vegetation type and fractional cover, dynamic vegetation, and fire). We tested three hypotheses: that models with dynamic vegetation better capture GPP variability due to their ability to simulate grass regrowth after disturbance; that models with more accurate representation of grass fractional cover leads to better alignment with DryFlux GPP IAV; and that fire effects on GPP IAV are stronger in grass-dominated systems, especially in models with both dynamic vegetation and fire modules. While it is not possible to estimate exact contributions of different processes to poor model performance in capturing annual variability in an emergent quantity

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

such as GPP, this study nevertheless explored how vegetation characteristics and disturbances influence model performance in simulating dryland carbon dynamics. We identified TRENDY v11 models without dynamic vegetation and fire enabled capture low GPP IAV compared to the DryFlux product across the western North America dryland region. Among different PFT types, models with high C4 grass cover have high correlation with GPP IAV. LPJ models have high grass fCover, and high fCover variability, which is likely contributing to their higher GPP IAV. However, the fCover variability is not due to fire, and the spatial patterns in fCover variability do not match spatial patterns in GPP IAV, suggesting that other processes are also contributing to high GPP variability. However our findings provide a roadmap to the modellers in terms of improving vegetation representation in sparse woody shrub grass dynamic dryland ecosystems, specifically to focus on accurate representation of C4 grasses.

Data and Code Availability

The Aridity Index (AI) dataset was downloaded from: https://cgiarcsi.community/2019/01/24/global-aridity-index-and-potential-evapotranspiration-climate-database-v3/. The DryFlux v1.0 GPP product was downloaded from https://github.com/marthageb/DryFlux and is freely available under a MIT License Copyright (c) 2021 marthageb. MODIS VCF was downloaded through google earth engine, April 2024, but also available at: https://lpdaac.usgs (DiMiceli et al., 2015). The RAP fCover data was downloaded from http://rangeland.ntsg.umt.edu/data/rap/rap-vegetation-cover/v3/. TRENDY model data is available from the Global Carbon Budget Data Browser (https://mdosullivan.github.io/GCB/). Scripts produced in this study are deposited in the following github repository: 'https://github.com/Rubaya-Pervin/TRENDY-models-evaluation-using-DryFlux'. Once the manuscript has been accepted for publication after any revisions have been completed, this statement will be updated at that time and a Zenodo doi will be created for this repository so that a permanent record is available for this paper.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Author Contribution

All the work including all data processing, methods and code development, analyses, figure plotting, and most of the writing was carried out by RP. NM primarily wrote Section 4.2, 4.3 and 4.4 and co-author LB primarily wrote Section 4.5 in collaboration with RP. Co-authors SR, KM, and MB provided feedback on preliminary analyses and manuscript drafts. TRENDY modelers contributed information on processes in their models as well as feedback on manuscript drafts.

Competing interests

At least one of the (co-)authors is a member of the editorial board of Biogeosciences

References

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., Jain, A. K., Kato, E., Poulter, B., Sitch, S., Stocker, B. D., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S., and Zeng, N.: Carbon cycle. The dominant role of semi-arid ecosystems in the trend and variability of the land CO_2 sink, Science, 348, 895–899, https://doi.org/10.1126/science.aaa1668, 2015.

Allred, B. W., Bestelmeyer, B. T., Boyd, C. S., Brown, C., Davies, K. W., Duniway, M. C., Ellsworth, L. M., Erickson, T. A., Fuhlendorf, S. D., Griffiths, T. V., Jansen, V., Jones, M. O., Karl, J., Knight, A., Maestas, J. D., Maynard, J. J., McCord, S. E., Naugle, D. E., Starns, H. D., Twidwell, D., and Uden, D. R.: Improving Landsat predictions of rangeland fractional cover with multitask learning and uncertainty, Methods in Ecology and Evolution, 12, 841–849, https://doi.org/10.1111/2041-210X.13564, 2021.

Anderson-Teixeira, K. J., Delong, J. P., Fox, A. M., Brese, D. A., and Litvak, M. E.: Differential responses of production and respiration to temperature and moisture drive the carbon balance across a climatic gradient in New Mexico, Glob. Chang. Biol., 17, 410–424, https://doi.org/10.1111/j.1365-2486.2010.02269.x, 2011.

Barnes, M. L., Susan Moran, M., Scott, R. L., Kolb, T. E., Ponce-Campos, G. E., Moore, D. J. P., Ross, M. A., Mitra, B., and Dore, S.: Vegetation productivity responds to sub-annual climate conditions across semiarid biomes, Ecosphere, 7, e01339, https://doi.org/10.1002/ecs2.1339, 2016.

Barnes, M. L., Farella, M. M., Scott, R. L., Moore, D. J. P., Ponce-Campos, G. E., Biederman, J. A., MacBean, N., Litvak, M. E., and Breshears, D. D.: Improved dryland carbon flux predictions with explicit consideration of water-carbon coupling, Communications Earth & Environment, 2, 1–9, https://doi.org/10.1038/s43247-021-00308-2, 2021.

Bastos, A., Naipal, V., Ahlström, A., MacBean, N., Smith, W. K., and Poulter, B.: Semiarid ecosystems, in: Balancing Greenhouse Gas Budgets, Elsevier, 311–335, https://doi.org/10.1016/b978-0-12-814952-2.00012-5, 2022.

Baudena, M., Dekker, S. C., van Bodegom, P. M., Cuesta, B., Higgins, S. I., Lehsten, V., Reick, C. H., Rietkerk, M., Scheiter, S., Yin, Z., Zavala, M. A., and Brovkin, V.: Forests, savannas, and grasslands: bridging the knowledge gap between ecology and Dynamic Global Vegetation Models, Biogeosciences, 12, 1833–1848, https://doi.org/10.5194/bg-12-1833-2015, 2015.

Biederman, J. A., Scott, R. L., Bell, T. W., Bowling, D. R., Dore, S., Garatuza-Payan, J., Kolb, T. E., Krishnan, P., Krofcheck, D. J., Litvak, M. E., Maurer, G. E., Meyers, T. P., Oechel, W. C., Papuga, S. A., Ponce-Campos, G. E., Rodriguez, J. C., Smith, W. K., Vargas, R., Watts, C. J., Yepez, E. A., and Goulden, M. L.: CO exchange and evapotranspiration across dryland ecosystems of southwestern North America, Glob Chang Biol, 23, 4204–4221, https://doi.org/10.1111/gcb.13686, 2017.

Bonan, G.: Climate change and terrestrial ecosystem modeling, Cambridge University Press, Cambridge, England, 456 pp., https://doi.org/10.1017/9781107339217, 2019.

Burton, C., Betts, R., Cardoso, M., Feldpausch, T. R., Harper, A., Jones, C. D., Kelley, D. I., Robertson, E., and Wiltshire, A.: Representation of fire, land-use change and vegetation dynamics in the Joint UK Land Environment Simulator vn4.9 (JULES), Geosci. Model Dev., 12, 179–193, https://doi.org/10.5194/gmd-12-179-2019, 2019.

Case, M. F. and Staver, A. C.: Soil texture mediates tree responses to rainfall intensity in African savannas, New Phytol, 219, 1363–1372, https://doi.org/10.1111/nph.15254, 2018.

Collatz, G. J., Berry, J. A., and Clark, J. S.: Effects of climate and atmospheric CO partial pressure on the global distribution of C grasses: present, past, and future, Oecologia, 114, 441–454, https://doi.org/10.1007/s004420050468, 1998.

Cranko Page, J., De Kauwe, M. G., Abramowitz, G., Cleverly, J., Hinko-Najera, N., Hovenden, M. J., Liu, Y., Pitman, A. J., and Ogle, K.: Examining the role of environmental memory in the predictability of carbon and water fluxes across Australian ecosystems, Biogeosciences, 19, 1913–1932, https://doi.org/10.5194/bg-19-1913-2022, 2022.

Dannenberg, M. P., Barnes, M. L., Smith, W. K., Johnston, M. R., Meerdink, S. K., Wang, X., Scott, R. L., and Biederman, J. A.: Upscaling dryland carbon and water fluxes with artificial neural networks of optical, thermal, and microwave satellite remote sensing, Biogeosciences, 20, 383–404, https://doi.org/10.5194/bg-20-383-2023, 2023.

De Pue, J., Barrios, J. M., Liu, L., Ciais, P., Arboleda, A., Hamdi, R., Balzarolo, M., Maignan, F., and Gellens-Meulenberghs, F.: Local-scale evaluation of the simulated interactions between energy, water and vegetation in ISBA, ORCHIDEE and a diagnostic model, Biogeosciences, 19, 4361–4386, https://doi.org/10.5194/bg-19-4361-2022, 2022.

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D.-H., Kelly, M., and Townshend, J.: MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006, https://doi.org/10.5067/MODIS/MOD44B.006, 2015.

DiMiceli, C., Townshend, J., Carroll, M., and Sohlberg, R: Evolution of the representation of global

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

vegetation by vegetation continuous fields, Remote Sensing of Environment, 254, 112271, https://doi.org/10.1016/j.rse.2020.112271, 2021.

Ehleringer, J. R.: Implications of quantum yield differences on the distributions of C and C grasses, Oecologia, 31, 255–267, https://doi.org/10.1007/BF00346246, 1978.

Fawcett, D., Cunliffe, A. M., Sitch, S., O'Sullivan, M., Anderson, K., Brazier, R. E., Hill, T. C., Anthoni, P., Arneth, A., Arora, V. K., Briggs, P. R., Goll, D. S., Jain, A. K., Li, X., Lombardozzi, D., Nabel, J. E. M. S., Poulter, B., Séférian, R., Tian, H., Viovy, N., Wigneron, J.-P., Wiltshire, A., and Zaehle, S.: Assessing model predictions of carbon dynamics in global drylands, Front. Environ. Sci., 10, https://doi.org/10.3389/fenvs.2022.790200, 2022.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., Le Quéré, C., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Bates, N. R., Becker, M., Bellouin, N., Bopp, L., Chau, T. T. T., Chevallier, F., Chini, L. P., Cronin, M., Currie, K. I., Decharme, B., Djeutchouang, L. M., Dou, X., Evans, W., Feely, R. A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Luijkx, I. T., Jain, A., Jones, S. D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P. C., Melton, J. R., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Schwingshackl, C., Séférian, R., Sutton, A. J., Sweeney, C., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G. R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, C., Yue, X., Zaehle, S., and Zeng, J.: Global Carbon Budget 2021, Earth System Science Data, 14, 1917–2005, https://doi.org/10.5194/essd-14-1917-2022, 2022.

Harper, A. B., Wiltshire, A. J., Cox, P. M., Friedlingstein, P., Jones, C. D., Mercado, L. M., Sitch, S., Williams, K., and Duran-Rojas, C.: Vegetation distribution and terrestrial carbon cycle in a carbon cycle configuration of JULES4.6 with new plant functional types, Geosci. Model Dev., 11, 2857–2873, https://doi.org/10.5194/gmd-11-2857-2018, 2018.

Hartley, A. J., MacBean, N., Georgievski, G., and Bontemps, S.: Uncertainty in plant functional type distributions and its impact on land surface models, Remote Sens. Environ., 203, 71–89, https://doi.org/10.1016/j.rse.2017.07.037, 2017.

Keenan, T. F., Baker, I., Barr, A., Ciais, P., Davis, K., Dietze, M., Dragoni, D., Gough, C. M., Grant, R., Hollinger, D., Hufkens, K., Poulter, B., McCaughey, H., Raczka, B., Ryu, Y., Schaefer, K., Tian, H., Verbeeck, H., Zhao, M., and Richardson, A. D.: Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO_2 exchange, Glob. Chang. Biol., 18, 1971–1987, https://doi.org/10.1111/j.1365-2486.2012.02678.x, 2012.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y. C., Taylor, C. M., Verseghy, D., Vasic, R., Xue, Y., Yamada, T., and GLACE Team: Regions of strong coupling between soil moisture and precipitation, Science, 305, 1138–1140, https://doi.org/10.1126/science.1100217, 2004.

Li, F., Zeng, X. D., and Levis, S.: A process-based fire parameterization of intermediate complexity in a Dynamic Global Vegetation Model, Biogeosciences, 9, 2761–2780, https://doi.org/10.5194/bg-9-2761-2012, 2012.

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Lin, S., Hu, Z., Wang, Y., Chen, X., He, B., Song, Z., Sun, S., Wu, C., Zheng, Y., Xia, X., Liu, L., Tang, J., Sun, Q., Joos, F., and Yuan, W.: Underestimated interannual variability of terrestrial vegetation production by terrestrial ecosystem models, Global Biogeochem. Cycles, 37, https://doi.org/10.1029/2023gb007696, 2023.

Liu, Y., Schwalm, C. R., Samuels-Crow, K. E., and Ogle, K.: Ecological memory of daily carbon exchange across the globe and its importance in drylands, Ecol Lett, 22, 1806–1816, https://doi.org/10.1111/ele.13363, 2019.

Luo, X., Zhou, H., Satriawan, T. W., Tian, J., Zhao, R., Keenan, T. F., Griffith, D. M., Sitch, S., Smith, N. G., and Still, C. J.: Mapping the global distribution of C vegetation using observations and optimality theory, Nat Commun, 15, 1219, https://doi.org/10.1038/s41467-024-45606-3, 2024.

MacBean, N., Scott, R. L., Biederman, J. A., Peylin, P., Kolb, T., Litvak, M. E., Krishnan, P., Meyers, T. P., Arora, V. K., Bastrikov, V., Goll, D., Lombardozzi, D. L., Nabel, J. E. M. S., Pongratz, J., Sitch, S., Walker, A. P., Zaehle, S., and Moore, D. J. P.: Dynamic global vegetation models underestimate net CO2 flux mean and inter-annual variability in dryland ecosystems, Environ. Res. Lett., 16, 094023, https://doi.org/10.1088/1748-9326/ac1a38, 2021.

Mahmud, K., Scott, R. L., Biederman, J. A., Litvak, M. E., Kolb, T., Meyers, T. P., Krishnan, P., Bastrikov, V., and MacBean, N.: Optimizing carbon cycle parameters drastically improves terrestrial biosphere model underestimates of dryland mean net CO 2 flux and its inter-annual variability, J. Geophys. Res. Biogeosci., 126, https://doi.org/10.1029/2021jg006400, 2021.

Mangeon, S., Voulgarakis, A., Gilham, R., Harper, A., Sitch, S., and Folberth, G.: INFERNO: a fire and emissions scheme for the UK Met Office's Unified Model, Geosci. Model Dev., 9, 2685–2700, https://doi.org/10.5194/gmd-9-2685-2016, 2016.

McGranahan, D. A. and Wonkka, C. L.: Pyrogeography of the Western Great Plains: A 40-year history of fire in semi-arid rangelands, Fire, 7, 32, https://doi.org/10.3390/fire7010032, 2024.

Melton, J. R. and Arora, V. K.: Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0, Geosci. Model Dev., 9, 323–361, https://doi.org/10.5194/gmd-9-323-2016, 2016.

Melton, J. R., Arora, V. K., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., and Teckentrup, L.: CLASSIC v1.0: the open-source community successor to the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) – Part 1: Model framework and site-level performance, Geosci. Model Dev., 13, 2825–2850, https://doi.org/10.5194/gmd-13-2825-2020, 2020.

Metz, E.-M., Vardag, S. N., Basu, S., Jung, M., Ahrens, B., El-Madany, T., Sitch, S., Arora, V. K., Briggs, P. R., Friedlingstein, P., Goll, D. S., Jain, A. K., Kato, E., Lombardozzi, D., Nabel, J. E. M. S., Poulter, B., Séférian, R., Tian, H., Wiltshire, A., Yuan, W., Yue, X., Zaehle, S., Deutscher, N. M., Griffith, D. W. T., and Butz, A.: Soil respiration-driven CO pulses dominate Australia's flux variability, Science, 379, 1332–1335, https://doi.org/10.1126/science.add7833, 2023.

Metz, E.-M., Vardag, S. N., Basu, S., Jung, M., and Butz, A.: Seasonal and interannual variability in CO_2 fluxes in southern Africa seen by GOSAT, Biogeosciences, 22, 555–584, https://doi.org/10.5194/bg-22-555-2025, 2025.

Munson, S. M., Sankey, T. T., Xian, G., Villarreal, M. L., and Homer, C. G.: Decadal shifts in grass and woody plant cover are driven by prolonged drying and modified by topo-edaphic properties, Ecol Appl,

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

26, 2478–2492, https://doi.org/10.1002/eap.1389, 2016.

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Ottlé, C., Poulter, B., Zaehle, S., and Running, S. W.: Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling, Hydrol. Earth Syst. Sci., 24, 1485–1509, https://doi.org/10.5194/hess-24-1485-2020, 2020.

Paschalis, A., Fatichi, S., Zscheischler, J., Ciais, P., Bahn, M., Boysen, L., Chang, J., De Kauwe, M., Estiarte, M., Goll, D., Hanson, P. J., Harper, A. B., Hou, E., Kigel, J., Knapp, A. K., Larsen, K. S., Li, W., Lienert, S., Luo, Y., Meir, P., Nabel, J. E. M. S., Ogaya, R., Parolari, A. J., Peng, C., Peñuelas, J., Pongratz, J., Rambal, S., Schmidt, I. K., Shi, H., Sternberg, M., Tian, H., Tschumi, E., Ukkola, A., Vicca, S., Viovy, N., Wang, Y.-P., Wang, Z., Williams, K., Wu, D., and Zhu, Q.: Rainfall manipulation experiments as simulated by terrestrial biosphere models: Where do we stand?, Glob Chang Biol, 26, 3336–3355, https://doi.org/10.1111/gcb.15024, 2020.

Pervin, R., Robeson, S. M., and MacBean, N.: Fusion of airborne hyperspectral and LiDAR canopy-height data for estimating fractional cover of tall woody plants, herbaceous vegetation, and other soil cover types in a semi-arid savanna ecosystem, Int. J. Remote Sens., 43, 3890–3926, https://doi.org/10.1080/01431161.2022.2105176, 2022.

Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., and Zimmermann, N. E.: Plant functional type mapping for earth system models, Geosci. Model Dev., 4, 993–1010, https://doi.org/10.5194/gmd-4-993-2011, 2011.

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y., Running, S. W., Sitch, S., and van der Werf, G. R.: Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle, Nature, 509, 600–603, https://doi.org/10.1038/nature13376, 2014.

Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., and Solomon, A. M.: Special paper: A global biome model based on plant physiology and dominance, soil properties and climate, J. Biogeogr., 19, 117, https://doi.org/10.2307/2845499, 1992.

Renwick, K. M., Fellows, A., Flerchinger, G. N., Lohse, K. A., Clark, P. E., Smith, W. K., Emmett, K., and Poulter, B.: Modeling phenological controls on carbon dynamics in dryland sagebrush ecosystems, Agric. For. Meteorol., 274, 85–94, https://doi.org/10.1016/j.agrformet.2019.04.003, 2019.

Scholes, R. J.: The future of semi-arid regions: A weak fabric unravels, Climate, 8, 43, https://doi.org/10.3390/cli8030043, 2020.

Singleton, M. P., Thode, A. E., Sánchez Meador, A. J., and Iniguez, J. M.: Increasing trends in high-severity fire in the southwestern USA from 1984 to 2015, For. Ecol. Manage., 433, 709–719, https://doi.org/10.1016/j.foreco.2018.11.039, 2019.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Chang. Biol., 9, 161–185, https://doi.org/10.1046/j.1365-2486.2003.00569.x, 2003.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N.,

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653–679, https://doi.org/10.5194/bg-12-653-2015, 2015.

Sitch, S., O'Sullivan, M., Robertson, E., Friedlingstein, P., Albergel, C., Anthoni, P., Arneth, A., Arora, V. K., Bastos, A., Bastrikov, V., Bellouin, N., Canadell, J. G., Chini, L., Ciais, P., Falk, S., Harris, I., Hurtt, G., Ito, A., Jain, A. K., Jones, M. W., Joos, F., Kato, E., Kennedy, D., Klein Goldewijk, K., Kluzek, E., Knauer, J., Lawrence, P. J., Lombardozzi, D., Melton, J. R., Nabel, J. E. M. S., Pan, N., Peylin, P., Pongratz, J., Poulter, B., Rosan, T. M., Sun, Q., Tian, H., Walker, A. P., Weber, U., Yuan, W., Yue, X., and Zaehle, S.: Trends and drivers of terrestrial sources and sinks of carbon dioxide: An overview of the TRENDY project, Global Biogeochem. Cycles, 38, https://doi.org/10.1029/2024gb008102, 2024.

Smith, B., Prentice, I.C. and Sykes, M.T.: Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, Glob. Ecol. and Biogeogr., 621-637, 2001.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences, 11, 2027–2054, https://doi.org/10.5194/bg-11-2027-2014, 2014.

Still, C. J., Berry, J. A., Collatz, G. J., and DeFries, R. S.: Global distribution of C_3 and C_4 vegetation: Carbon cycle implications, Global Biogeochem. Cycles, 17, 6–1–6–14, https://doi.org/10.1029/2001gb001807, 2003.

Still, C. J., Cotton, J. M., and Griffith, D. M.: Assessing earth system model predictions of C 4 grass cover in North America: From the glacial era to the end of this century, Glob. Ecol. Biogeogr., 28, 145–157, https://doi.org/10.1111/geb.12830, 2019.

Teckentrup, L., De Kauwe, M. G., Pitman, A. J., Goll, D. S., Haverd, V., Jain, A. K., Joetzjer, E., Kato, E., Lienert, S., Lombardozzi, D., McGuire, P. C., Melton, J. R., Nabel, J. E. M. S., Pongratz, J., Sitch, S., Walker, A. P., and Zaehle, S.: Assessing the representation of the Australian carbon cycle in global vegetation models, Biogeosciences, 18, 5639–5668, https://doi.org/10.5194/bg-18-5639-2021, 2021.

Thonicke K, Venevsky S, Sitch S, Cramer W.: The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model, Glob. Ecol. and Biogeogr., 10(6), 661-77, https://doi.org/10.1046/j.1466-822X.2001.00175.x, 2001.

Thonicke, K., Spessa, A., Prentice, I. C., Harrison, S. P., Dong, L., and Carmona-Moreno, C.: The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model, Biogeosciences, 7, 1991–2011, https://doi.org/10.5194/bg-7-1991-2010, 2010.

Trabucco, A. and Zomer, R.: Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v3, https://doi.org/10.6084/m9.figshare.7504448.v4, 2022.

Traore, A., Ciais, P., Vuichard, N., MacBean, N., Dardel, C., Poulter, B., Piao, S., Fisher, J., Viovy, N., Jung, M., and Myneni, R.: 1982–2010 trends of Light Use Efficiency and inherent Water Use Efficiency in African vegetation: Sensitivity to climate and atmospheric CO2 concentrations, Remote Sens. (Basel), 6, 8923–8944, https://doi.org/10.3390/rs6098923, 2014.

Verbruggen, W., Schurgers, G., Horion, S., Ardö, J., Bernardino, P. N., Cappelaere, B., Demarty, J., Fensholt, R., Kergoat, L., Sibret, T., Tagesson, T., and Verbeeck, H.: Contrasting responses of woody and herbaceous vegetation to altered rainfall characteristics in the Sahel, Biogeosciences, 18, 77–93, https://doi.org/10.5194/bg-18-77-2021, 2021.

Verbruggen, W., Schurgers, G., Meunier, F., Verbeeck, H., and Horion, S.: Simulated tree-grass competition in drylands is modulated by CO_2 fertilization, Earths Future, 12, https://doi.org/10.1029/2023ef004096, 2024.

Wang, L., Jiao, W., MacBean, N., Rulli, M. C., Manzoni, S., Vico, G., and D'Odorico, P.: Dryland productivity under a changing climate, Nat. Clim. Chang., 12, 981–994, https://doi.org/10.1038/s41558-022-01499-y, 2022.

Whitley, R., Beringer, J., Hutley, L. B., Abramowitz, G., De Kauwe, M. G., Duursma, R., Evans, B., Haverd, V., Li, L., Ryu, Y., Smith, B., Wang, Y.-P., Williams, M., and Yu, Q.: A model inter-comparison study to examine limiting factors in modelling Australian tropical savannas, Biogeosciences, 13, 3245–3265, https://doi.org/10.5194/bg-13-3245-2016, 2016.

Whitley, R., Beringer, J., Hutley, L. B., Abramowitz, G., De Kauwe, M. G., Evans, B., Haverd, V., Li, L., Moore, C., Ryu, Y., Scheiter, S., Schymanski, S. J., Smith, B., Wang, Y.-P., Williams, M., and Yu, Q.: Challenges and opportunities in land surface modelling of savanna ecosystems, Biogeosciences, 14, 4711–4732, https://doi.org/10.5194/bg-14-4711-2017, 2017.

Wilcox, B. P., Birt, A., Fuhlendorf, S. D., and Archer, S. R.: Emerging frameworks for understanding and mitigating woody plant encroachment in grassy biomes, Curr. Opin. Environ. Sustain., 32, 46–52, https://doi.org/10.1016/j.cosust.2018.04.005, 2018.

Wilcox, K. R., Chen, A., Avolio, M. L., Butler, E. E., Collins, S., Fisher, R., Keenan, T., Kiang, N. Y., Knapp, A. K., Koerner, S. E., Kueppers, L., Liang, G., Lieungh, E., Loik, M., Luo, Y., Poulter, B., Reich, P., Renwick, K., Smith, M. D., Walker, A., Weng, E., and Komatsu, K. J.: Accounting for herbaceous communities in process-based models will advance our understanding of "grassy" ecosystems, Glob Chang Biol, 29, 6453–6477, https://doi.org/10.1111/gcb.16950, 2023.

Wolf, A., Callaghan, T. V., and Larson, K.: Future changes in vegetation and ecosystem function of the Barents Region, Clim. Change, 87, 51–73, https://doi.org/10.1007/s10584-007-9342-4, 2008.

Xie, Q., Huete, A., Hall, C. C., Medlyn, B. E., Power, S. A., Davies, J. M., Medek, D. E., and Beggs, P. J.: Satellite-observed shifts in C3/C4 abundance in Australian grasslands are associated with rainfall patterns, Remote Sens. Environ., 273, 112983, https://doi.org/10.1016/j.rse.2022.112983, 2022.

Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., Ryu, Y., Chen, G., Dong, W., Hu, Z., Jain, A. K., Jiang, C., Kato, E., Li, S., Lienert, S., Liu, S., Nabel, J. E. M. S., Qin, Z., Quine, T., Sitch, S., Smith, W. K., Wang, F., Wu, C., Xiao, Z., and Yang, S.: Increased atmospheric vapor pressure deficit reduces global vegetation growth, Sci Adv, 5, eaax1396, https://doi.org/10.1126/sciadv.aax1396, 2019.

Zhang, X., Wang, Y.-P., Peng, S., Rayner, P. J., Ciais, P., Silver, J. D., Piao, S., Zhu, Z., Lu, X., and Zheng, X.: Dominant regions and drivers of the variability of the global land carbon sink across timescales, Glob Chang Biol, 24, 3954–3968, https://doi.org/10.1111/gcb.14275, 2018.

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., Pan, Y., Peng, S., Peñuelas, J., Poulter, B., Pugh, T. A. M., Stocker, B. D., Viovy, N., Wang, X., Wang, Y., Xiao,

https://doi.org/10.5194/egusphere-2025-2841
Preprint. Discussion started: 31 July 2025

Z., Yang, H., Zaehle, S., and Zeng, N.: Greening of the Earth and its drivers, Nat. Clim. Chang., 6, 791–795, https://doi.org/10.1038/nclimate3004, 2016.