Biases in active land water storage capacity and water limitation of evapotranspiration in CMIP6 models

Francesco Giardina[1], Ryan S. Padrón[1,2], Benjamin D. Stocker[3,4], Dominik L. Schumacher[1], Sonia I. Seneviratne[1]

[1]Institute for Atmospheric and Climate Science, Department of Environmental Systems Science, ETH Zurich, CH-8092 Zürich, Switzerland
[2]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, CH-8903 Birmensdorf, Switzerland
[3]Institute of Geography, University of Bern, Hallerstrasse 12, 3012 Bern, Switzerland
[4]Oeschger Centre for Climate Change Research, University of Bern, Falkenplatz 16, 3012 Bern, Switzerland

Author for correspondence:
Francesco Giardina
Email: fgiardina@ethz.ch

Abstract

Accurate representation of water available for evapotranspiration on land, i.e. active land water storage capacity (ALWSC), is crucial for climate modeling, due to its significant role in land-atmosphere interactions. Our study analyzes how ALWSC and water limitation of evapotranspiration are represented in Earth System Model (ESM) simulations of the Coupled Model Intercomparison Project phase 6 (CMIP6). First, we compare the models' long-term maximum annual depletion in total-column soil moisture and cumulative water deficit – i.e. proxies for ALWSC – with estimates based on satellite observations of terrestrial water storage from the Gravity Recovery and Climate Experiment (GRACE), as well as remotely-sensed precipitation and evapotranspiration. Our analysis shows that CMIP6 models mostly underestimate ALWSC, especially in the Amazon region. Second, we assess the frequency of water limitation of evapotranspiration in CMIP6 simulations against observations from solar-induced fluorescence (SIF) and GRACE, as well as measurements at 128 sites from the FLUXNET2015 dataset. We find that ESMs overestimate the time under water limitation by 9% over land, and by 18% in the tropics. Our study highlights the importance of improving the representation of ALWSC and its effects on land-atmosphere interactions in ESMs. Model development in this area could have large implications in projections of future climate on land.

Introduction

Climate projections are based on a variety of Earth System Model (ESM) simulations compiled in model intercomparison projects[1]. The accuracy of these simulations is key for progress in climate science and eventually affects the implementation of climate policies globally. The sixth phase of the Coupled Model Intercomparison Project (CMIP6) substantially contributed to the physical science basis of the Sixth Assessment Report (AR6) by the Intergovernmental Panel on Climate Change (IPCC)[1,2]. This phase includes the most advanced Earth System Models (ESMs), simulating historical and future climates based on greenhouse gas and aerosol concentration scenarios outlined in the Shared Socioeconomic Pathways (SSP)[3]. Nonetheless, continuous efforts are needed to keep improving multi-model ensembles, as some models have been shown to not fully align with observational evidence or theoretical understanding[4–8].

The land water available for evapotranspiration (ET) – here referred to as active land water (ALW) – links the global energy, water and carbon cycles, and is hence key for accurate climate projections[9–11]. In particular, ALW denotes the combined water available in the soil profile, including surface water, soil moisture from both upper and deeper layers, accessible groundwater and rock moisture, that collectively contributes to plant transpiration and surface evaporation. By influencing the distribution of available energy (net radiation) at the land surface, ALW primarily directs this energy towards the evaporation of water[9,12]. This process not only affects the water cycle but also modulates the turbulent heat fluxes, thereby impacting the climate system overall[13]. ALW acts as a reservoir for precipitation and radiation anomalies, maintaining stability in the climate system[12,14]. In addition, ALW influences the global carbon cycle given that plants regulate photosynthesis and transpiration in response to water availability[9].

Although there are no direct measurements of active land water storage capacity (ALWSC), i.e. the maximum amount of land water accessible for ET, its spatial variability has been estimated based on cumulative water deficit (CWD) and vegetation activity[15]. Similarly, here we assume that the maximum range of variability over multiple years in the amount of water stored below the surface is indicative of ALWSC. The wettest conditions correspond to saturation, whereas the driest level is reached after all the available water for ET has been consumed. The long-term maximum annual water depletion can be accurately estimated in ESMs given that they represent the water stored in all layers below the surface. On the other hand, some of the most suitable observational proxies for ALWSC include terrestrial water storage (TWS) from the Gravity Recovery and Climate Experiment (GRACE)[10,11,16,17] and the cumulative seasonal difference in evapotranspiration minus precipitation[15,18].

When ALW drops below a critical threshold, the evaporative fraction of net radiation decreases, leading to an increase in the sensible heat fraction and ultimately in air temperature[12,13]. Water limitation is estimated to affect evapotranspiration in 30% to 60% of the Earth's land surface for most of the year[19], and it is an important factor in the exacerbation of heat extremes. However, our understanding of ALWSC and water limitation, as well as their interaction and modulation by climate change, remains limited[12].

This is reflected in the representation of land-atmosphere interactions in climate models[20,21]. Recent research has suggested an overestimation of future warming across CMIP6 models[5], plausibly connected to other documented biases in soil moisture[6] through land-atmosphere interactions. This uncertainty affects our ability to project our success in limiting global warming below the 1.5°C target outlined in the Paris Agreement. It is therefore crucial to constrain these model ensembles with other evidence, such as historical trends and current climate observations.

Using simulations from the land surface model component of seven ESMs within CMIP6[22] driven with observed atmospheric forcing, we show that ALWSC is largely underestimated compared to observations, particularly in the Amazon. Consistently, we find that the frequency of water-limited conditions for evapotranspiration is generally overestimated across models. We discuss the reasons behind these biases and the potential implications for climate projections and policy.

Biases in active land water storage capacity

We focus on the *Land-Hist* experiment within the Land Model Intercomparison Project (LMIP), which consists of simulations from the land component of ESMs participating in CMIP6. These simulations are specifically designed to identify systematic biases in the representations of land processes in current ESMs[22] (see Methods).

We use '*total soil moisture*' from seven ESMs within LMIP[2,22]. This LMIP variable includes moisture from all soil layers in the model (see Methods). In each grid cell, we estimated the maximum depletion of total-column soil moisture (ΔSM_{max}) by first calculating the difference between the highest and lowest total-column soil moisture monthly values in every year. We then identify the greatest annual difference across all analyzed years:

$$\Delta SM_{max} = max(max(SM)_{year} - min(SM)_{year})_{all\,years} \qquad (1)$$

We repeated the calculation using terrestrial water storage (TWS) observations from the Gravity Recovery and Climate Experiment (GRACE)[16] and estimate ΔTWS_{max} from Eq. 1. We assume that changes in total soil moisture and TWS are comparable, as in previous studies[10,11,17].

Our results show that the ALWSC is generally underestimated in CMIP6 models compared to observations, especially in the tropics (Fig. 1a,b and Fig. S1). This observed underestimation could result from GRACE (and thus ΔTWS_{max}) including snow and water bodies in its calculation. This issue could be particularly problematic in regions with large annual variability in snow or water bodies. Therefore, we computed the maximum cumulative water deficit (CWD_{max}), a metric based on the maximum cumulative difference between remotely sensed evapotranspiration (ET) and precipitation (P)[23–25], and compared it to the same quantity calculated with ESM data. We also found an underestimation by the models in this case (Fig. 1c,d and Fig. S2). Assessment with CWD thus reveals patterns of water stress effects resulting from whole-column water depletion (see Methods). For the CWD calculation with CMIP6 data, we directly used ET and P from CMIP6 models. For the observational reference (henceforth referred to as S_{CWDX}), we use the 80-year extreme CWD from Ref.[15], determined using ET data derived from thermal infrared remote sensing via the Atmosphere Land Exchange Inverse (ALEXI) product[23,24] and precipitation reanalysis data from WATCH-WFDEI[25] (see Methods).

As mentioned, there is a potential for overestimating ΔSM_{max} when using GRACE data, particularly in regions prone to regular flooding, such as the Amazon, because the product accounts for all water bodies on land (i.e. rivers, lakes, water in vegetation, etc). However, using the CWD_{max} method with ALEXI and WATCH-WFDEI datasets still indicates that models underestimate ALWSC in the Amazon (Figs. 1c, d and 2). On the other hand, S_{CWDX} tends to underestimate ALWSC because it is derived from ET obtained from infrared thermal remote sensing, which is often inaccurately low at high latitudes and in regions that are not typically water-limited (Fig. 1c). Ultimately, whereas no model or observational dataset provides a flawless comparison, the insights based on GRACE and ALEXI are valuable and generally point towards consistent findings (Figs. 1 and 2). The consistency in patterns of exceptionally higher ALWSC within the wet tropics, as identified through both GRACE and S_{CWDX} observations (Fig. 1a,c), highlights the robustness of our findings (Fig. 1b,d).

When grouping our results by IPCC regions, a clear model underestimation of ALWSC is apparent, especially in the tropics (Fig. 2, Fig. S3 and Fig. S4). Using the S_{CWDX} and the CWD_{max} method reveals further significant regions within the wet tropics where ALWSC is underestimated, extending to Africa, Australia, South Asia, and Central America (Fig. 2b).

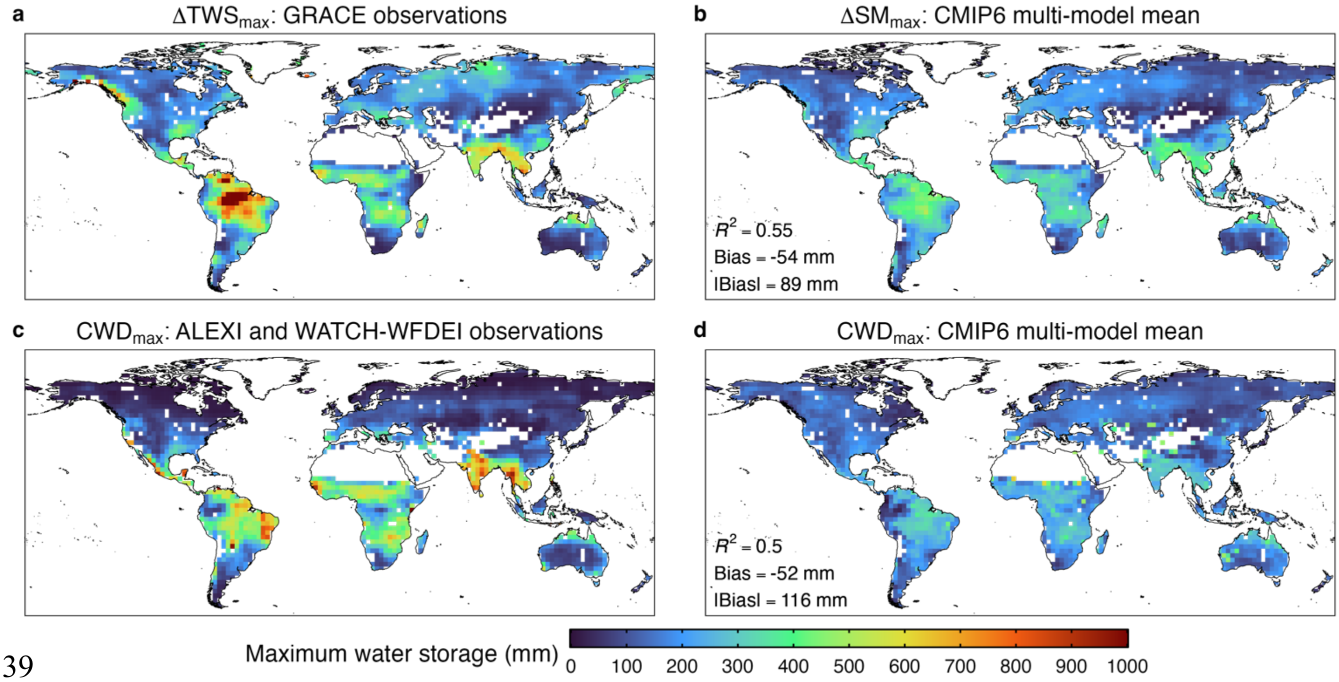

Fig. 1 | Comparison of active land water storage capacity (ALWSC) from LMIP-CMIP6 simulations against observations-based estimates a-b, The maximum soil moisture depletion (ΔSM_{max}), a proxy for ALWSC, was first derived yearly during 2003-2014. For every grid cell, we retained the maximum yearly reduction across those years. (**a**) ΔTWS_{max} from GRACE and (**b**) ΔSM_{max} from LMIP-CMIP6. **c-d,** ALWSC estimated with maximum CWD (**c**), determined from ALEXI and WATCH-WFDEI observations (S_{CWDX}) augmented using an extreme value distribution with an 80-year return period, as detailed in Ref[15] and (**d**) maximum CWD assessed over the 80 years of LMIP-CMIP6 data (1935-2014), to align with the methodology used for S_{CWDX}. The *land-hist* simulation was used across CMIP6 models. Bilinear interpolation was applied to harmonize the GRACE resolution with that of CMIP6. The raw bias was determined by subtracting pixel-by-pixel the observed value from each model and then calculating the mean of these differences across all pixels. For the absolute bias, we computed the mean after taking the absolute value of each pixel-wise subtraction. Biases were weighted to account for the latitudinal variation in grid cell area. Values exceeding 1000 mm are colored as 1000 mm for clarity. The multi-model mean was calculated with models CESM2, CMCC-ESM2, CNRM-ESM2-1, E3SM-1-1, EC-Earth3-Veg, IPSL-CM6A-LR and UKESM1-0-LL.

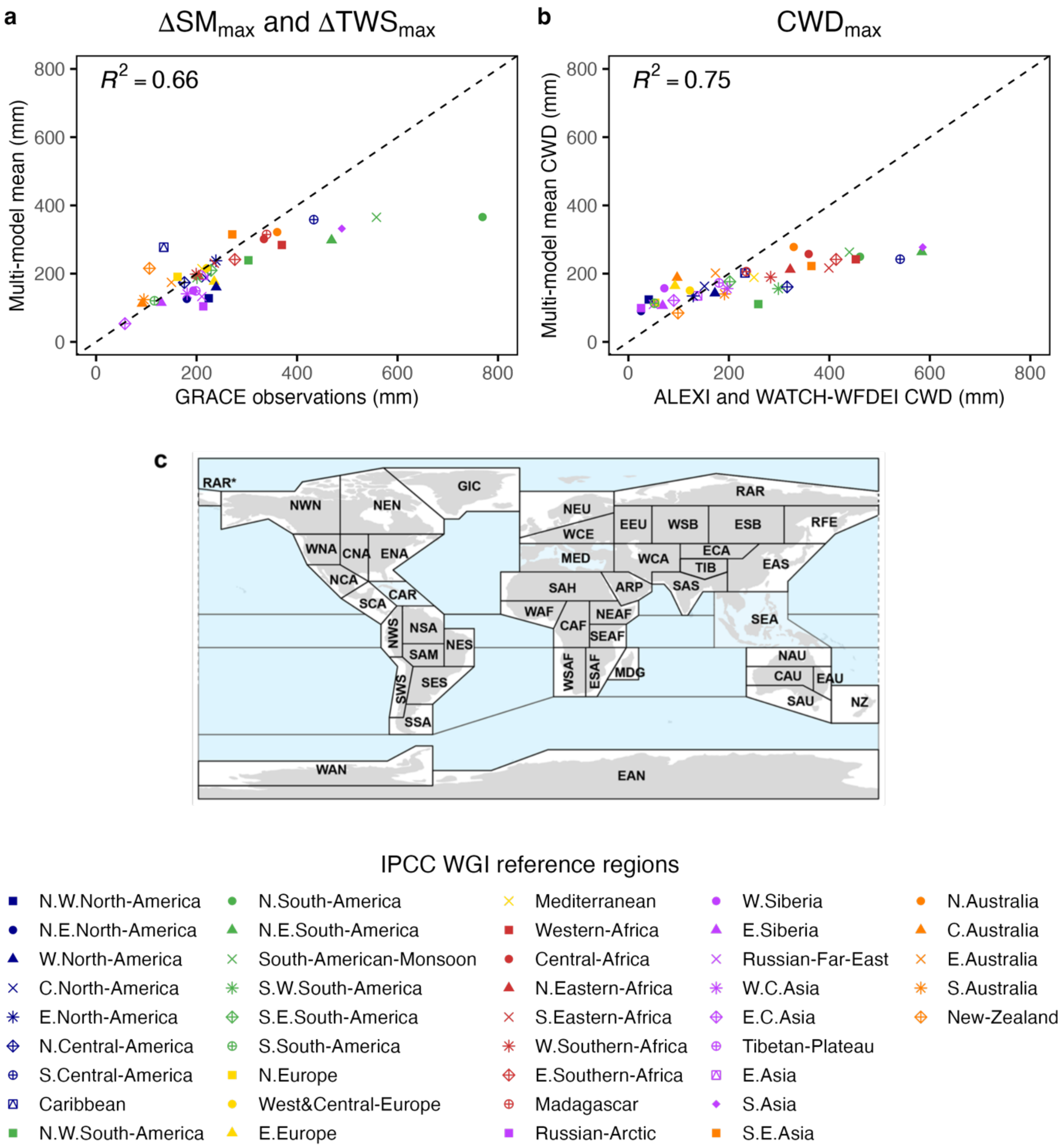

Fig. 2 | Comparison of active land water storage capacity (ALWSC) across different IPCC regions based on LMIP-CMIP6 model simulations and observations. We present a region-wise evaluation of ALWSC, using observations-based estimates and corresponding LMIP-CMIP6 model simulations. **a**, ΔTWS_{max} from GRACE versus ΔSM_{max} from LMIP-CMIP6. **b**, Maximum CWD from ALEXI and WATCH-WFDEI versus maximum CWD from LMIP-CMIP6. **c,** IPCC reference regions. We first determined the ALWSC in every pixel and then we calculated the mean of this value within each IPCC region for model and observational data. The multi-model mean was calculated with models CESM2, CMCC-ESM2, CNRM-ESM2-1, E3SM-1-1, EC-Earth3-Veg, IPSL-CM6A-LR and UKESM1-0-LL.

Biases in water limitation

To quantify water constraints on evapotranspiration and photosynthesis, we study the evaporative fraction of net radiation (EF) as a function of the total-column soil moisture in the LMIP-CMIP6 models (see Methods). Given the complexities of directly observing EF on a global scale, we use normalized remotely sensed solar-induced fluorescence (SIF) as a proxy for EF and TWS from GRACE as a proxy for total water storage, as in Ref[11]. We study the relation of EF as a function of ALW to assess the frequency of water limitation in models and observations. With this aim, we first derive the water limitation threshold θ_{crit} at every grid cell and subsequently compute the proportion of time each grid cell remains beneath θ_{crit} (see Methods).

We find that the frequency of water limitation is on average overestimated in CMIP6 models compared to observations by 9% over land, and up to 18% in the tropics (Fig. 3 and Fig. S5). Conversely, models often underrepresent the occurrence of water limitation in central and northern Europe and generally at high latitudes in the Northern Hemisphere.

The finding that models significantly overestimate the time under water limitation in the Amazon when compared to GRACE and GOME-2 data is further confirmed when performing the same analysis with FLUXNET2015 observations (Fig. 4 and Fig. S6). Our analysis indicates that for many ESMs, even a typically humid biome as the Amazon can experience water limitation for 30% or more of the days in a year (Fig. 4d,g,h,i,j). This pattern of water limitation is similar to that observed in a dry Mediterranean savanna (US-Ton, Fig. 4k-o). ESMs also overestimate the frequency of water limitation at multiple FLUXNET2015 stations in Australia, China, Italy and the USA (Fig. S6). We use daily data and directly calculate EF when comparing the co-located ESM output to the FLUXNET2015 observations. Nevertheless, it is important to acknowledge the potential mismatch in spatial scale between the footprint of the eddy-covariance flux measurements and the size of the ESM grid cell.

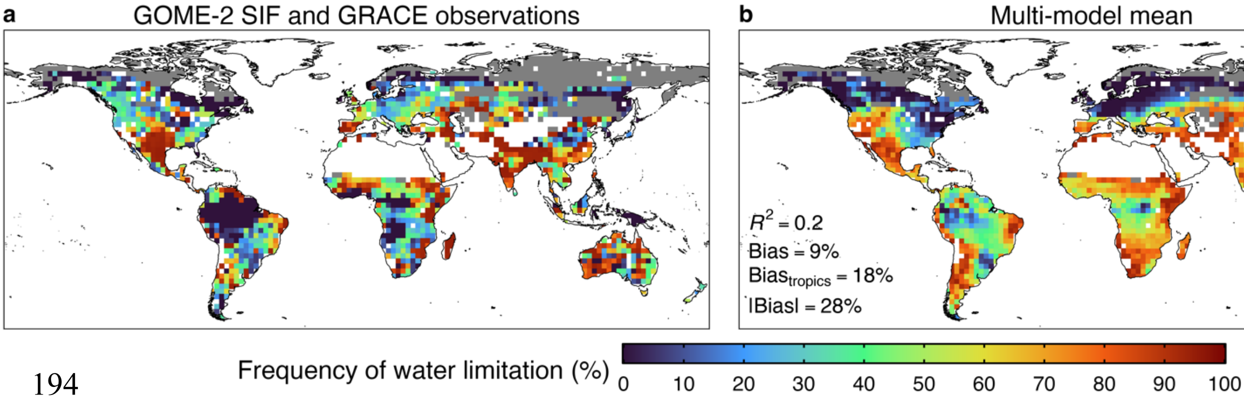

Fig. 3 | Global maps of frequency of water limitation. Pixel-specific critical water limitation thresholds (θ_{crit}) were calculated to determine when plant water stress occurs. We show here the fraction of months with total-column soil moisture below θ_{crit}. **a,** θ_{crit} determined with the normalized SIF versus normalized TWS from GRACE (2007-2014). **b,** θ_{crit} calculated with EF vs normalized total-column soil moisture with LMIP-CMIP6 data (2007-2014). Dark blue pixels represent areas where water is rarely limiting. On the other hand, dark red pixels represent areas that are almost always water-limited. Grey areas illustrate regions where the segmented regression could not be applied due to scarcity of data points. The raw bias was determined by subtracting pixel-by-pixel the observed value from each model and then calculating the mean of these differences across all pixels. For the absolute bias, we computed the mean after taking the absolute value of each pixel-wise subtraction. Biases were weighted to account for the latitudinal variation in grid cell area. Details of all datasets and normalizations can be found in Methods. The multi-model mean was calculated with models CESM2, CMCC-ESM2, CNRM-ESM2-1, E3SM-1-1, EC-Earth3-Veg, IPSL-CM6A-LR and UKESM1-0-LL.

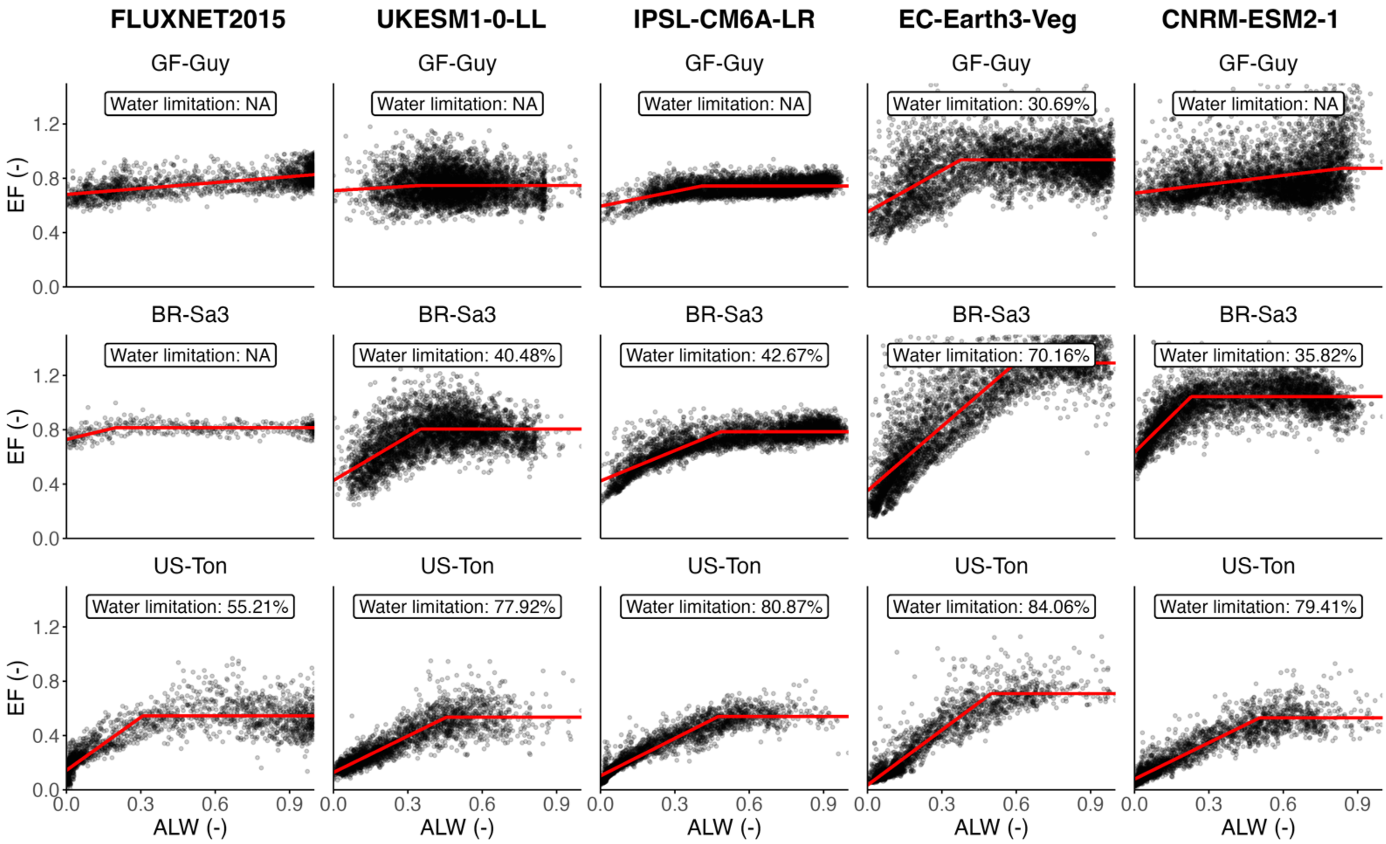

Fig. 4 | Analysis of water limitation at three selected eddy-covariance sites. Evolution of the evaporative fraction (EF) with the active land water (ALW). **a,f,k,** eddy-covariance measurements from the FLUXNET2015 dataset. **b,g,l,** data from UKESM1-0-LL. **c,h,m,** data from IPSL-CM6A-LR. **d,l,n.** data from EC-Earth3-Veg. **e,j,o.** data from CNRM-ESM2-1. When the decrease in the EF was under 30% (as indicated by the y-axis intercept) compared to its maximum value, we assigned a status of no water limitation (NA). The total soil moisture was normalized to ease the comparison between model outputs and the corresponding observed values (see Methods). GF-Guy is an evergreen broadleaf forest site situated in the coastal region of the north-western Amazon, in French Guyana. BR-Sa3, another site characterized by evergreen broadleaf forest, lies further south and inland in Brazil. US-Ton is an oak savanna woodland located near Sacramento, California, United States.

Potential causes of biases between CMIP6 models and observations

Our analysis, based on different methods and datasets, consistently shows that models underestimate ALWSC compared to observations, particularly in the tropics (Figs. 1 and 2). This suggests that models potentially overestimate the occurrence of water limitation, as confirmed in Figs. 3, 4 and S5 (i.e. the raw bias is positive for all models). Given that ALWSC and ET are closely linked through land-atmosphere interactions[9,12], the underestimation of ALWSC and overestimation of the time under water limitation is consistent with previous studies suggesting that models underestimate ET in the tropics[7], and across most regions during dry periods[18,26–28].

Europe and North America emerge among the least biased regions when compared to observations (Figs. 1-4). This is probably due to the large availability of ground-based observations to constrain ESMs in these areas compared to the rest of the world. At the same time, ALWSC appears to be most biased in the Amazon and in general in the wet tropics (Fig. 2). This supports previous findings that ESMs tend to overestimate water stress in the Amazon and do not adequately capture the positive sensitivity to atmospheric aridity in its most humid regions[29]. This also potentially reflects the inadequate representation of tropical forest root traits found in global models[30]. Given the key role of the Amazon for the global water and carbon cycles[31,32], it is crucial to improve model accuracy when representing this region, also because the response of tropical rainforests to water limitation is one of the main uncertainties in ESMs[33].

Among the seven LMIP-CMIP6 models considered in this study, CESM2 has the most accurate representation of the soil-plant-atmosphere continuum[27,34,35]. It explicitly accounts for plant hydraulics and calculates water potentials in soil, roots, stems, and leaves[34]. This enables plants in CESM2 to draw more water for transpiration from deeper soil layers compared to other ESMs[34]. CNRM-ESM2-1 is also one of the best ESMs in simulating deep soil moisture[6]. This may explain why CESM2 is one of least biased models in the Amazon (Fig. S3) and overall the model with the lowest raw bias (Figs. S1, S2 and S5). UKESM1-0-LL rank among the models with lowest absolute bias and highest linear fit for the spatial pattern (R^2) (Figs. S1, S2). This is perhaps linked to the fact that the model simulates LAI dynamics instead of assuming a constant LAI seasonality and they account for vegetation properties and land use change (Table S1). EC-Earth3-Veg, while being one of the best models at estimating ALWSC (Figs. S1 and S2), ranks as one of the least accurate in terms of water limitation (Fig. S5). This discrepancy likely stems from its limited accuracy in simulating EF (Fig. 4).

Implications for predicting future climate on land

Our global comparison of seven state-of-the-art ESMs to observational estimates has revealed a tendency to underestimate ALWSC and overestimate water limitation. In other words, these models generally (i) simulate less water potentially available for ET on land, and (ii) limit ET more frequently in response to moisture deficits than what is suggested by observational estimates. These biases hamper the model representations of both regional and global water cycles. For example, the Amazon is characterized by a high precipitation recycling ratio, as about one third of the rainfall has previously evaporated from the Amazon itself[36]. In this and other regions strongly reliant on terrestrial ET, exaggerated ET suppression in response to ALWSC deficits in the model could result in excessive drought self-intensification and self-propagation[12,37]. Given that precipitation projections are uncertain due to both internal climate variability and the reliance on parameterizations at subgrid-scales[1], it is crucial to improve model fidelity of ET to prevent an amplification of this uncertainty through unrealistic land–atmosphere interactions. As the global hydrological cycle is expected to intensify in response to our warming climate[38], although the extent to which this already emerges remains debated[39,40], these soil moisture-limitation biases in ET are likely to disproportionately affect the reliability of future projections.

Owing to the fundamental role of ALW in modulating not only water but also energy and carbon fluxes, model biases in water use and limitation propagate beyond the hydrological cycle. Soil moisture–temperature feedbacks are known to amplify hot extremes in most land areas [41], which also emerges clearly in climate projections[42,43]. In fact, it has recently been shown that across much of Europe, air temperature increases are outpaced by even stronger soil temperature trends, suggesting that the heat comes from below[44]. It is challenging to reliably quantify the role of soil moisture–temperature coupling in changing climate, but for certain regions such as the moist tropics, including Amazonia, CMIP6 simulations point to a strong contribution of land feedbacks to extreme heat[45]. Recent work, again based on CMIP6 model experiments, indicates that strong land-atmosphere coupling will become more widespread under increasing atmospheric CO2, suggesting an amplification of future climate sensitivity to such feedbacks[46]. These findings distinctly rely on the ability of the CMIP6 multi-model ensemble to adequately capture the interactions between land and atmosphere, yet our results indicate systematic deficiencies with respect to how the land surface models make use of the available subsurface water, and how they respond to drought conditions. As such, targeted efforts to improve the representation of these processes in climate models would likely enable more accurate projections of hot and dry extremes.

We remark here that in certain regions, e.g. Eastern North America, Northern Europe, and India, the CMIP6 model subset used here underestimates the frequency of water limitation (Fig. 3). Consequently, in those regions, increases in both the occurrence and magnitude of future heatwaves could be underestimated by current state-of-the-art ESMs. Individual hot and dry events can undo several years' worth of net carbon uptake at regional scales[47], and global soil moisture variability has been shown to dictate the strength of the terrestrial carbon sink[10,11,48], which in turn largely governs the fraction of anthropogenic CO2 emissions remaining in the atmosphere. Due to this inherent link between land carbon sequestration and climate extremes, model improvements of both subsurface water utilization and limitation could also decrease the intermodel uncertainty of carbon uptake and hence long-term climate projections.

In this study we highlight the need to improve the representation of ALWSC and water limitation of ET and photosynthesis in climate models. The consistent biases we find across CMIP6 models with respect to observations, i.e. underestimation of ALWSC and overestimation of water limitation frequency, indicate a significant potential for model development. Our analysis illustrates the challenges ESMs face in accurately capturing the specificities of the land water cycle, with significant implications for the simulated land water, energy and carbon fluxes. We note that although the overestimated frequency of water limitation could be interpreted as a direct consequence of the ALWSC underestimation, it is also conceivable in turn that the land component of the ESMs responds too strongly to water stress, reducing ET and thereby limiting further soil moisture depletion, which would decrease our estimate of ALWSC. This is consistent with other studies pointing to a general overreliance on shallow rather than deep soil moisture in models[27,49] and the stronger drying trends in projections of surface compared to deep soil moisture[50]. Future research into land-climate dynamics under climate change can be supported by focusing on land surface model simulations at flux sites and, in general, using higher frequency data outputs[51]. This would offer a more direct comparison between model predictions and empirical observations, avoiding scale mismatches and thus advancing our understanding of terrestrial water and carbon cycles.

Methods

Data sets

This study investigates how ALWSC and water limitation of evapotranspiration are represented across several models from the Coupled Model Intercomparison Project phase 6[2]. Our focus is on the *Land-Hist* experiment within the Land Model Intercomparison Project (LMIP), which consists in global land-only offline simulations driven with observed atmospheric forcing over a historical interval improving snow and soil moisture estimates. Sharing the same configuration of *historical* simulations of the parent model within CMIP6, the *Land-Hist* experiment is conceived for diagnosing systematic biases within the land component of ESMs[22].

To benchmark CMIP6 models, we use several observational datasets. For the estimation of the long-term maximum annual soil moisture depletion in Fig. 1, we use total water storage (TWS) data from the Gravity Recovery and Climate Experiment (GRACE)[16]. TWS accounts for soil moisture, groundwater, surface water, snow and ice and it has been successfully used in previous hydrological studies[52,53]. ΔTWS is used here as a proxy of ALWSC. The key advantage of the GRACE dataset lies in its foundation on mass balance principles, ensuring its water balance aligns with that of CMIP6 models. Both CMIP6 models and GRACE operate on this principle, providing consistency in their approach to water balance, despite the CMIP6 models likely not capturing all physical processes perfectly. We use SIF from version 2.6 of the Global Ozone Monitoring Experiment-2 (GOME-2)[54] as a proxy of photosynthetic activity (Fig. 3), consistent with previous studies[32,54,55]. Monthly means are calculated retaining days when the effective cloud fraction is $<30\%$, as described in Ref[32]. SIF is a complementary process of photosynthesis, and it is thus directly related to the photosynthetic rate[56]. We used eddy-covariance data from the FLUXNET2015 dataset[57] complemented with soil moisture simulated with a bucket-type soil water balance model[58] when observational soil moisture from FLUXNET2015 was unavailable or inconsistent, as documented in Ref[18].

We use CMIP6 data from the ETH Zürich CMIP6 next generation (CMIP6ng) archive[59], which adds extra validation for processed variables and consistency among files from different sources. Even though some observational data sets were available at a greater spatial resolution, all data were bilinearly interpolated to a common $2.5° \times 2.5°$ global grid for comparison. We retained pixels with vegetated land using a global land cover dataset from MODIS[60]. To group the vegetated land pixels of the world in meaningful climatic regions

(Fig. 2), we use the fourth version of the IPCC WGI reference regions[61]. All analyses were performed using R Statistical Software[62]. To access all code and R packages used in this study, please refer to our published repository on GitHub and Zenodo (see 'Data Availability' section).

Estimating active land water storage capacity

We estimate ALWSC as the long-term maximum annual soil moisture depletion in CMIP6 models, by using the total-column soil moisture (variable 'mrso') at the monthly resolution. We first calculated the difference in annual maximum and minimum soil moisture values, and then selected the greatest difference across all years (Fig. 1). Directly calculating the difference as the maximum minus the minimum soil moisture across all years yielded similar results (Fig. S7). The initial calculation was chosen because land cover is subject to change over time, making the subtraction of soil moisture values from widely spaced time points potentially not physically meaningful. For the observational map in Fig. 1, we repeated the calculation of the long-term maximum annual soil moisture depletion with TWS from GRACE (Fig. 1), also used at monthly resolution. The calculation was performed for the years 2003-2014, when both GRACE and CMIP6 data sets were available. GRACE was converted from cm to mm, whereas total-column soil moisture was already available in Kg m^{-2} (equivalent to mm H_2O).

To ensure robustness of our analysis, we also estimated the maximum cumulative water deficit calculated as in Refs[15,18], using ALEXI and WATCH-WFDEI data as observational benchmark, which confirmed our results (Fig. 1).

To visualize regional biases in CMIP6 predictions, we grouped the results of Fig. 1 by IPCC climate reference regions[61]. We determined the mean of the land ALWSC given by ΔSM_{max} across all points within each region, using CMIP6 data. We then compared to the corresponding observational data for the same region (Fig. 2).

Determining water limitation thresholds globally with monthly data

We studied the evaporative fraction (EF) as a function of the total-column soil moisture (SM, variable 'mrso') using monthly data from CMIP6 models. EF was calculated as the ratio of latent heat flux to net radiation:

$$EF = \frac{latent\ heat\ flux}{R_n} = \frac{\mathrm{hfls}}{(\mathrm{rsds} - \mathrm{rsus}) + (\mathrm{rlds} - \mathrm{rlus})} \tag{2}$$

where 'hfls' (W m^{-2}) is latent heat flux from CMIP6 and 'rsds', 'rsus', 'rlds' and 'rlus' were respectively incoming and outgoing shortwave radiation and incoming and outgoing longwave radiation (W m^{-2}), also from CMIP6. We retained pixels with $R_n > 75$ W m^{-2} to focus on the growing season. We then fitted a segmented linear regression with one breakpoint (also known as "linear-plus-plateau model"[63,64]) to the EF vs SM relationship at each pixel, using R package "segmented"[65]. The pixel-specific estimate of the breaking point θ_{crit} was determined by least-square fit; its value represents the SM threshold up to which EF increases linearly as a function of SM (water-limited regime)[9,63,64]. The percentage of time under SM limitation was calculated as the ratio of the number of months with SM < θ_{crit} divided by the total number of months (Fig. 3).
The global observational map shown in Fig. 3 was created with GOME-2 SIF[54] data and TWS data from GRACE[16] as a proxy of ALW. To derive a metric comparable to EF, SIF data were normalized to their maximum value over the entire measurement period, as in Ref.[11]. We focused on monthly data from the growing season by retaining pixels greater than or equal to half of the pixel-specific maximum. Both model and observational analyses were limited to the period from January 2007 to December 2014 (eight years), based on the availability of the observational and modelled datasets.
We do not extend our analysis of water limitation to the intercept and slope from the segmented regression given the different variables used for the models and observations, as well as the high sensitivity of both the intercept and slope to the underlying assumptions of the segmented regression and the quantity of data points included in the analysis.

Determining water limitation thresholds at flux tower locations with daily data

We repeated the EF vs SM analysis outlined in the preceding section at the site-scale, using FLUXNET2015 daily data at selected sites (Fig. 4 and Fig. S6). We calculated EF using FLUXNET2015 data as $EF = \frac{latent\ heat\ flux}{R_n}$. Due to inconsistencies of measured soil moisture at several FLUXNET2015 sites[18], we simulated soil moisture at eddy-covariance locations with SPLASH, a bucket-type soil water balance mode based on a Priestley-Taylor formulation for ET estimation; with water-holding capacity set to 220mm[58,66]. We focused on the growing season by retaining site-days with GPP equal or greater than half of the site-specific maximum. We extracted EF and SM data at FLUXNET2015 locations using daily datasets from 2000 to 2014. We used daily LMIP-CMIP6 data (available only for models UKESM1-0-LL, IPSL-CM6A-LR, EC-Earth3-Veg, and CNRM-ESM2-1) and focused on the grid cells corresponding to the FLUXNET2015 sites for comparison. We determined the critical threshold θ_{crit} and calculated the percentage of days when SM was less than θ_{crit} relative to the total number of days. When the decrease in EF was less than 0.3, as indicated by the y-axis intercept compared to its maximum value, we assigned a status of no water limitation (NA) to avoid interpreting noise as water limitation.

Data availability

All intermediate data and computer code that support this study are available from the Zenodo Digital Repository: https://zenodo.org/doi/10.5281/zenodo.10810324 (Giardina et al. 2024).

All data used in this study are openly available:

- LMIP-CMIP6 data (see Table S1 and Methods for details of the experiments and run IDs): https://esgf-node.llnl.gov/search/cmip6/ or directly from the ETH Zurich CMIP6 next generation archive: https://zenodo.org/records/3734128
- GRACE land data: https://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons/
- GOME-2 SIF: https://avdc.gsfc.nasa.gov/pub/data/satellite/MetOp/GOME_F
- Ecosystem fluxes and meteorological data: https://fluxnet.org/data/fluxnet2015-dataset/
- Global estimates of maximum CWD determined from ALEXI and WATCH-WFDEI data, augmented using an extreme value distribution: https://zenodo.org/records/5515246

Acknowledgements

The authors extend their gratitude to all data providers who have made this study possible. In particular, we want to acknowledge the FLUXNET community, for their role in making the FLUXNET2015 dataset available.

Author Contributions

F.G. wrote the main manuscript in collaboration with R.S.P. and D.L.S. F.G. prepared figures. F.G., S.I.S. and R.S.P. designed the study. F.G. performed the analysis in collaboration with R.S.P. All authors reviewed and edited the manuscript.

Competing Financial Interests

The authors declare no competing financial interests.

References

1. Seneviratne, S. I. *et al.* Weather and Climate Extreme Events in a Changing Climate. in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* 1513–1766 (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021). doi:https://doi.org/10.1017/9781009157896.013.
2. Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci Model Dev* **9**, 1937–1958 (2016).
3. Meinshausen, M. *et al.* The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geosci Model Dev* **13**, 3571–3605 (2020).
4. Tebaldi, C. & Knutti, R. The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* vol. 365 2053–2075 Preprint at https://doi.org/10.1098/rsta.2007.2076 (2007).
5. Tokarska, K. B. *et al.* Past warming trend constrains future warming in CMIP6 models. *Sci Adv* **6**, 9549–9567 (2020).
6. Qiao, L., Zuo, Z. & Xiao, D. Evaluation of Soil Moisture in CMIP6 Simulations. *J Clim* **35**, 779–800 (2022).
7. Wang, Z., Zhan, C., Ning, L. & Guo, H. Evaluation of global terrestrial evapotranspiration in CMIP6 models. *Theor Appl Climatol* **143**, 521–531 (2021).
8. Fu, Z. *et al.* Global critical soil moisture thresholds of plant water stress. *Nat Commun* **15**, 1–13 (2024).
9. Seneviratne, S. I. *et al.* Investigating soil moisture-climate interactions in a changing climate: A review. *Earth Sci Rev* **99**, 125–161 (2010).
10. Humphrey, V. *et al.* Sensitivity of atmospheric CO2 growth rate to observed changes in terrestrial water storage. *Nature* **560**, 628–631 (2018).
11. Green, J. K. *et al.* Large influence of soil moisture on long-term terrestrial carbon uptake. *Nature* **565**, 476–479 (2019).
12. Miralles, D. G., Gentine, P., Seneviratne, S. I. & Teuling, A. J. Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. *Ann N Y Acad Sci* **1436**, 19–35 (2019).
13. Budyko, M. I. *Climate and Life*. (Academic Press, 1974).
14. Humphrey, V., Gudmundsson, L. & Seneviratne, S. I. Assessing Global Water Storage Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes. *Surveys in Geophysics* vol. 37 357–395 Preprint at https://doi.org/10.1007/s10712-016-9367-1 (2016).
15. Stocker, B. D. *et al.* Global patterns of water storage in the rooting zones of vegetation. *Nat Geosci* **16**, 250–256 (2023).
16. Landerer, F. W. & Swenson, S. C. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resour Res* **48**, (2012).
17. Liu, L. *et al.* Increasingly negative tropical water–interannual CO2 growth rate coupling. *Nature* **618**, 755–760 (2023).
18. Giardina, F., Gentine, P., Konings, A. G., Seneviratne, S. I. & Stocker, B. D. Diagnosing evapotranspiration responses to water deficit across biomes using deep learning. *New Phytologist* **240**, 968–983 (2023).
19. Schwingshackl, C., Hirschi, M. & Seneviratne, S. I. Quantifying spatiotemporal variations of soil moisture control on surface energy balance and near-surface air temperature. *J Clim* **30**, 7105–7124 (2017).

20. García-García, A., Cuesta-Valero, F. J., Beltrami, H. & Smerdon, J. E. Characterization of Air and Ground Temperature Relationships within the CMIP5 Historical and Future Climate Simulations. *Journal of Geophysical Research: Atmospheres* **124**, 3903–3929 (2019).
21. Sippel, S. *et al.* Refining multi-model projections of temperature extremes by evaluation against land-Atmosphere coupling diagnostics. *Earth System Dynamics* **8**, 387–403 (2017).
22. Van Den Hurk, B. *et al.* LS3MIP (v1.0) contribution to CMIP6: The Land Surface, Snow and Soil moisture Model Intercomparison Project - Aims, setup and expected outcome. *Geosci Model Dev* **9**, 2809–2832 (2016).
23. Hain, C. R. & Anderson, M. C. Estimating morning change in land surface temperature from MODIS day/night observations: Applications for surface energy balance modeling. *Geophys Res Lett* **44**, 9723–9733 (2017).
24. Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P. & Mecikalski, J. R. A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sens Environ* **60**, 195–216 (1997).
25. Weedon, G. P. *et al.* The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resour Res* **50**, 7505–7514 (2014).
26. Ukkola, A. M. *et al.* Land surface models systematically overestimate the intensity, duration and magnitude of seasonal-scale evaporative droughts. *Environmental Research Letters* **11**, 104012 (2016).
27. Zhao, M., A, G., Liu, Y. & Konings, A. G. Evapotranspiration frequently increases during droughts. *Nat Clim Chang* **12**, 1024–1030 (2022).
28. Teuling, A. J., Seneviratne, S. I., Williams, C. & Troch, P. A. Observed timescales of evapotranspiration response to soil moisture. *Geophys Res Lett* **33**, 0–4 (2006).
29. Green, J. K., Berry, J., Ciais, P., Zhang, Y. & Gentine, P. Amazon rainforest photosynthesis increases in response to atmospheric dryness. *Sci Adv* **6**, (2020).
30. Pagán, B., Maes, W., Gentine, P., Martens, B. & Miralles, D. Exploring the Potential of Satellite Solar-Induced Fluorescence to Constrain Global Transpiration Estimates. *Remote Sens (Basel)* **11**, 413 (2019).
31. Pan, Y. *et al.* A large and persistent carbon sink in the world's forests. *Science (1979)* **333**, 988–993 (2011).
32. Giardina, F. *et al.* Tall Amazonian forests are less sensitive to precipitation variability. *Nat Geosci* **11**, 405–409 (2018).
33. Huntingford, C. *et al.* Simulated resilience of tropical rainforests to CO2-induced climate change. *Nat Geosci* **6**, 268–273 (2013).
34. Kennedy, D. *et al.* Implementing Plant Hydraulics in the Community Land Model, Version 5. *J Adv Model Earth Syst* **11**, 485–513 (2019).
35. Lawrence, D. M. *et al.* The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *J Adv Model Earth Syst* **11**, 4245–4287 (2019).
36. Dominguez, F. *et al.* Amazonian Moisture Recycling Revisited Using WRF With Water Vapor Tracers. *Journal of Geophysical Research: Atmospheres* **127**, e2021JD035259 (2022).
37. Schumacher, D. L., Keune, J., Dirmeyer, P. & Miralles, D. G. Drought self-propagation in drylands due to land–atmosphere feedbacks. *Nat Geosci* **15**, 262–268 (2022).

38. Allen, M. R. & Ingram, W. J. Constraints on future changes in climate and the hydrologic cycle. *Nature* vol. 419 224–232 Preprint at https://doi.org/10.1038/nature01092 (2002).
39. Koutsoyiannis, D. Revisiting the global hydrological cycle: Is it intensifying? *Hydrol Earth Syst Sci* **24**, 3899–3932 (2020).
40. Akhoudas, C. H. *et al.* Isotopic evidence for an intensified hydrological cycle in the Indian sector of the Southern Ocean. *Nat Commun* **14**, 1234567890 (2023).
41. Mueller, B. & Seneviratne, S. I. Hot days induced by precipitation deficits at the global scale. *Proc Natl Acad Sci U S A* **109**, 12398–12403 (2012).
42. Vogel, M. M. *et al.* Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks. *Geophys Res Lett* **44**, 1511–1519 (2017).
43. Vogel, M. M., Zscheischler, J. & Seneviratne, S. I. Varying soil moisture-atmosphere feedbacks explain divergent temperature extremes and precipitation projections in central Europe. *Earth System Dynamics* **9**, 1107–1125 (2018).
44. García-García, A. *et al.* Soil heat extremes can outpace air temperature extremes. *Nat Clim Chang* **13**, 1237–1241 (2023).
45. Dirmeyer, P. A., Sridhar Mantripragada, R. S., Gay, B. A. & Klein, D. K. D. Evolution of land surface feedbacks on extreme heat: Adapting existing coupling metrics to a changing climate. *Front Environ Sci* **10**, 949250 (2022).
46. Hsu, H. & Dirmeyer, P. A. Soil moisture-evaporation coupling shifts into new gears under increasing CO2. *Nat Commun* **14**, 1–9 (2023).
47. Ciais, P. *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* **437**, 529–533 (2005).
48. Padrón, R. S., Gudmundsson, L., Liu, L., Humphrey, V. & Seneviratne, S. I. Drivers of intermodel uncertainty in land carbon sink projections. *Biogeosciences* **19**, 5435–5448 (2022).
49. Dong, J., Lei, F. & Crow, W. T. Land transpiration-evaporation partitioning errors responsible for modeled summertime warm bias in the central United States. *Nat Commun* **13**, 1–8 (2022).
50. Berg, A., Sheffield, J. & Milly, P. C. D. Divergent surface and total soil moisture projections under global warming. *Geophys Res Lett* **44**, 236–244 (2017).
51. Findell, K. L. *et al.* Accurate assessment of land–atmosphere coupling in climate models requires high-frequency data output. *Geosci Model Dev* **17**, 1869–1883 (2024).
52. Jiang, W. *et al.* Annual variations of monsoon and drought detected by GPS: A case study in Yunnan, China. *Sci Rep* **7**, 1–10 (2017).
53. Yang, Y. *et al.* GRACE satellite observed hydrological controls on interannual and seasonal variability in surface greenness over mainland Australia. *J Geophys Res Biogeosci* **119**, 2245–2260 (2014).
54. Joiner, J. *et al.* Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2. *Atmos Meas Tech* **6**, 2803–2823 (2013).
55. Frankenberg, C. *et al.* New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophys Res Lett* **38**, 1–6 (2011).
56. Porcar-Castell, A. *et al.* Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges. *J Exp Bot* **65**, 4065–4095 (2014).
57. Pastorello, G. *et al.* The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci Data* **7**, 225 (2020).

58. Davis, T. W. *et al.* Simple process-led algorithms for simulating habitats (SPLASH v.1.0): Robust indices of radiation, evapotranspiration and plant-available moisture. *Geosci Model Dev* **10**, 689–708 (2017).
59. Brunner, L., Hauser, M. & Beyerle, R. L. and U. The ETH Zurich CMIP6 next generation archive: technical documentation. Preprint at https://doi.org/10.5281/zenodo.3734128 (2020).
60. Friedl, M. A. *et al.* MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens Environ* **114**, 168–182 (2010).
61. Iturbide, M. *et al.* An update of IPCC climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets. *Earth Syst Sci Data* **12**, 2959–2970 (2020).
62. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. (2023).
63. Fu, Z. *et al.* Critical soil moisture thresholds of plant water stress in terrestrial ecosystems. *Sci Adv* **8**, 7827 (2022).
64. Fu, Z. *et al.* Uncovering the critical soil moisture thresholds of plant water stress for European ecosystems. *Glob Chang Biol* **28**, 2111–2123 (2022).
65. Muggeo, V. M. R. Estimating regression models with unknown break-points. *Stat Med* **22**, 3055–3071 (2003).
66. Orth, R., Koster, R. D. & Seneviratne, S. I. Inferring soil moisture memory from streamflow observations using a simple water balance model. *J Hydrometeorol* **14**, 1773–1790 (2013).