Published in final edited form as:
Nat Geosci. 2023 February 9; 16: 250–256. doi:10.1038/s41561-023-01125-2.

Global patterns of water storage in the rooting zones of vegetation

Benjamin D. Stocker[1,2,3,4,5,*], **Shersingh Joseph Tumber-Dávila**[1,6], **Alexandra G. Konings**[1], **Martha C. Anderson**[7], **Christopher Hain**[8], **Robert B. Jackson**[1,9,10]

[1]Department of Earth System Science, Stanford University, Stanford, CA 94305-4216, USA

[2]Department of Environmental Systems Science, ETH, Universitatsstrasse 2, 8092 Zurich, Switzerland

[3]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

[4]Institute of Geography, University of Bern, Hallerstrasse 12, 3012 Bern, Switzerland

[5]Oeschger Centre for Climate Change Research, University of Bern, Falkenplatz 16, 3012 Bern, Switzerland

[6]Harvard Forest, Harvard University, Petersham, MA 01366, USA

[7]USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD 20705 USA

[8]NASA Marshall Space Flight Center, AL 35808, USA

[9]Woods Institute for the Environment, Stanford University, Stanford, CA 94305-4216, USA

[10]Precourt Institute for Energy, Stanford University, Stanford, CA 94305-4216, USA

Abstract

The rooting zone water storage capacity – the amount of water accessible to plants – controls the sensitivity of land-atmosphere exchange of water and carbon during dry periods. How the rooting zone water storage capacity varies spatially is largely unknown and not directly observable. Here we estimate rooting zone water storage capacity globally from the relationship between remotely-sensed vegetation activity, measured by combining evapotranspiration, sun-induced fluorescence, and radiation estimates, and the cumulative water deficit calculated from daily time series of precipitation and evapotranspiration. Our findings indicate plant-available water stores that exceed the storage capacity of 2 m deep soils across 37% of the Earth's vegetated surface. We find that

*corresponding author: Benjamin D. Stocker (benjamin.stocker@giub.unibe.ch).

Author contributions statement
B.D.S developed the methods, conducted the analysis and wrote the paper. S.J.T.D. compiled the rooting depth dataset and guided its analysis. A.G.K. helped design the analysis and manuscript. M.B.A. developed the algorithm for thermal infrared remote sensing of evapotranspiration. C.H. generated the evapotranspiration dataset. R.B.J. initiated the study and guided the analysis. All authors contributed to writing the manuscript.

Competing Interests Statement
Authors declare no competing interests.

Europe PMC Funders Author Manuscripts

biome-level variations of rooting zone water storage capacities correlate with observed rooting zone depth distributions and reflect the influence of hydroclimate, as measured by the magnitude of annual cumulative water deficit extremes. Smaller-scale variations are linked to topography and land use. Our findings document large spatial variations in the effective root zone water storage capacity and illustrate a tight link between the climatology of water deficits, rooting depth of vegetation and its sensitivity to water stress.

Introduction

To sustain activity during dry periods and resist impacts of droughts, plants rely on water stored below the surface. The larger the rooting zone water storage capacity (S_0), the longer plants can withstand soil moisture limitation[1]. S_0 is therefore a key factor determining drought impacts, land-atmosphere exchanges, and runoff regimes, particularly in climates with a seasonal asynchrony in radiation and precipitation (P)[2–4]. In models, S_0 is commonly conceived as a function of the soil texture and the plants' rooting depth (z_r), limited to the depth of the soil[3, 5]. Recent research has revealed a substantial component of S_0 and contributions to evapotranspiration (ET) by water stored beneath the soil, in weathered and fractured bedrock and groundwater[6–11]. Plant access to such deep moisture plays an important role in controlling near-surface climate[12–14], runoff regimes[4], global patterns of vegetation cover[15], and mitigating impacts of droughts[16].

However, S_0 is impossible to observe directly across large scales and its spatial variations are poorly understood[17]. Global compilations of local plant rooting depth (z_r) measurements[18, 19] yield information related to S_0, but have resolved this observational challenge only partly because of their limited size, and large documented variations in z_r across multiple scales[7, 18, 18–21]. Empirical approaches for estimating the global z_r distribution made use of relationships between in-situ observations and climatic factors[22]. Modelling approaches for predicting z_r have conceived their spatial variations as the result of optimal adaptation to the prevailing hydro-climate[23–25], or as being adapted to just buffer water demand to sustain ET during dry periods[2, 26]. Such mass-balance approaches make use of the maximum cumulative water deficit (CWD) during dry periods as an indication of the effective S_0. An additional hypothesis posits that it would not be beneficial for plants to root even deeper and thus size their S_0 even larger[26]. However, a link between the magnitude of CWD extremes, the sensitivity of vegetation activity to an increasing CWD, and local z_r observations remains to be shown, and the prevalence of plant access to water stored at depth (here taken as >2 m) across the globe remains to be quantified.

Despite its crucial role in controlling water and carbon fluxes and the scarcity of observations, virtually all models simulating water and carbon exchange between the land surface and the atmosphere rely on a specification of S_0 either directly as the depth of a "water bucket", or indirectly through prescribed rooting depths z_r and soil texture across the profile. Typically, water stored at depth and along the entire Critical Zone (including weathered bedrock) is not fully represented in models[8, 9], and the evident plasticity of z_r and variations S_0 within plant types and along climatic and topographic gradients are

Europe PMC Funders Author Manuscripts

often ignored. Implications of this simplification may be substantial for the simulation of land-atmosphere coupling and drought impacts[8, 12, 13].

Here, we present a method for diagnosing S_0 from the relationship between vegetation activity and CWD. By fusing multiple time series of Earth observation data streams with global coverage, we estimate the global distribution of S_0 at a resolution of 0.05° (~5 km). Using a mass-balance approach[2, 26] and field observations of z_r from a globally distributed dataset, we then show that the sensitivity of vegetation to water stress across the globe is strongly related to the magnitude of CWD extremes and reflects the rooting depth of plants.

Estimating S_0 from Earth observations

We started by estimating S_0 as the CWD at which vegetation "activity" ceases. Our approach accounts for the constraint of the rooting zone water availability on ET and photosynthesis and relates S_0 to the sensitivity of vegetation activity to water stress. The parallel information of ET, P, and the modelled snow mass balance enables a quantification of CWD over time. Vegetation activity was estimated from two alternative observations: from the evaporative fraction (EF, defined as ET divided by net radiation), and from sun-induced fluorescence (SIF, normalised by incident shortwave radiation, see also Methods).

Fig. 1 reveals large global variations in S_0. Estimates based on EF and SIF correlate closely and agree in magnitude (R^2 = 0.78, Supplementary Fig. S1). The lowest sensitivity of vegetation activity to an increasing CWD, and thus the largest apparent S_0, is found in regions with a strong seasonality in radiation and water availability and substantial vegetation cover - particularly in monsoonal climates. In contrast, the lowest S_0 values appear not only in regions where seasonal water deficits are limited due to short inter-storm duration (e.g., western Amazon and Congo basin) and/or low levels of potential ET (e.g., high latitudes), but also in deserts and arid grasslands. This likely reflects the limited water storage accumulating during rain events from which vegetation can draw during dry periods. In these regions, a rapid decline of ET and SIF with an increasing CWD is related to vegetation cover dynamics, governed by greening after rain pulses and browning during dry periods[27].

Clear patterns emerge also at smaller scales (Fig. 2, Extended Data Fig. 1-3). S_{dSIF} and S_{dEF} consistently (see Supplementary Fig. S1) reveal how the sensitivity of photosynthesis and transpiration to drought stress varies across different topographical settings, indicating generally larger S_0 in mountain regions ('M' in Fig. 2) and along rivers ('R') and deltas ('D'). We note however, that ET estimates from the product used here (ALEXI[28, 29]) may be biased high over mountainous terrain where low incident net radiation and surface temperatures are caused not by high evaporative fractions but rather by topography effects and local shading. The maps of S_{dSIF} and S_{dEF} also bear strong imprints of human land use. Major irrigated cropland areas are congruent with some of the highest apparent S_0 values. In these areas, our analysis yields particularly high CWD values and a low sensitivity of SIF and EF to CWD - without using information about the location and magnitude of irrigation. Other major irrigated areas appear as blank cells in Fig. 2 because the algorithm used to calculate CWD (see Methods) fails due to a long-term imbalance between P and ET and a

"runaway CWD". This indicates a sustained over-use of water resources, caused by lateral water redistribution at scales beyond ~5 km via streamflow diversion or groundwater flow and extraction (or bias in P and ET estimates).

Regressing vegetation activity against CWD also identifies locations where a decoupling of the two variables appears, i.e., where the sensitivity of EF or SIF significantly decreases beyond a certain CWD threshold ("Flattening" in Fig. 2, see Methods). Such areas are particularly common in the vicinity of mountain regions, in areas with irrigated croplands, and in savannahs (Supplementary Fig. S2). Related mechanisms may be at play. A flattening of the EF (SIF) vs. CWD relationship is likely due to different portions of the vegetation having access to distinct water resources and respective storage capacities. In areas with large topographic gradients, this may be due to within-gridcell heterogeneity in plant access to the saturated zone. Although relevant for land-atmosphere coupling[12], land surface models typically do not account for such effects. This has potential implications for simulations of ET during prolonged dry periods in these regions. In savannahs, a shift in ET contributions from grasses and trees and a related shift in transpiration occurs as grasses, which are often more shallow-rooted than trees[30], senesce. In irrigated cropland areas, the flattening likely reflects land use heterogeneity within ~5 km grid cells and the persistent water access on irrigated fields while EF and SIF are reduced more rapidly in surrounding vegetation.

What controls spatial variations in S_0 and z_r and the sensitivity of vegetation activity to water stress? Following ref.[2], we hypothesized that annual CWD maxima reflect the total amount of plant-accessible water. That is, z_r and S_0 are sized to just maintain transpiration and photosynthesis under extreme water deficits, commonly experienced over the course of a plant's lifetime (recurring with a return period of T years). Hence, a correlation between the magnitude of CWD extremes and the sensitivity of vegetation activity to an increasing CWD should emerge. For estimating CWD extremes, we started by using T = 80 years and assessed other choices as described in Supplementary Text S1, and see Extended Data Fig. 4.

Fig. 3a shows the global distribution of S_{CWDX80} and reveals patterns across multiple scales - in close agreement with S_{dSIF} and S_{dEF} (R^2 = 0.76 and R^2 = 0.83, respectively, Supplementary Fig. S3). This indicates that the sensitivity of vegetation activity to an increasing CWD (measured by S_{dSIF} and S_{dEF}) is strongly controlled by hydroclimate (as measured by S_{CWDX80}). The agreement between S_0 estimates based on water mass balance approaches[2, 26] and vegetation activity suggests that plants tend to size their roots no deeper, and S_0 no larger, than what is suggested by observed CWD extremes. Magnitudes of S_{CWDX80} inferred for 55% (37%) of the Earth's vegetated regions indicate plant access to water stored beyond 1 (2) m soil, assuming texture-dependent water holding capacity[31–33] (Extended Data Figs. 5, 6).

Fine granularity and large spatial heterogeneity of S_{CWDX80} at regional scales reveal the importance of land use and the local topographical setting for determining plant-available water storage capacities (Extended Data Figs. 7, 8). Complex patterns emerge. Mountainous areas feature higher S_{CWDX80} than their surrounding lowlands. In other regions, lowlands

feature some of the highest recorded S_{CWDX80}. In these regions, irrigated agriculture is wide-spread (Fig. 2 and Extended Data Fig. 1). Variations are likely to extend to even smaller scales along the hillslope topography[7] and within individual forest stands[34]. These scales lie beyond the resolution of the satellite remote sensing data used here to calculate CWD.

Evaluation with rooting depth observations

S_0 provides an estimate of the *effective* total plant-available water, independent of assumptions about physical constraint (limited soil depth, shallow bedrock, or groundwater), and independent of uncertain soil texture and water holding capacity (WHC). Due to the absence of direct observational constraints on S_0, we converted S_0 to a corresponding apparent plant rooting depth (z_r), enabling an evaluation of S_0 estimates against fully independent observations. We focused on comparing biome-level distributions of inferred apparent rooting depth (z_{CWDX80}) with a dataset[30] containing 5524 individual field observations of plant rooting depth from 1705 globally distributed sites (Supplementary Fig. S4). We thus tested the link between hydroclimate and belowground vegetation structure across large climatic gradients.

Predicted and observed biome-level maximum rooting depth (90% quantiles) are correlated (Pearson's $r = 0.68$, Fig. 4c), while the lower (10%) quantiles appear to be overestimated by z_{CWDX80} (Fig. 4b). Using a subset of the data where information about the water table depth (WTD) is provided (489 entries from 359 sites), we limited values of z_{CWDX80} to the value of the observed local WTD (53% of all observations). This yields a strongly improved correlation of observed and estimated the biome-level 10% rooting depth quantiles (Pearson's $r = 0.91$, Fig. 4d) compared to estimates that are not capped at the observed WTD (Fig. 4b). This suggests that inferred z_r overestimates values where roots access the groundwater and indicates that groundwater access is relevant across more than half of the globally distributed sites in our dataset. While acting as a constraint on the rooting depth[7], plant access to groundwater or a perched water table implies sustained transpiration during dry periods, correspondingly large CWDs and, by implication of the model design, large S_{CWDX80} and (apparent) z_{CWDX80}.

Influence of biotic and abiotic factors

Using first-principles modelling and integrating multiple data streams, we diagnosed a hydrologically effective ecosystem-level S_0 from the sensitivity of vegetation activity to CWD. We found that large-scale variations in S_0 are driven by the hydroclimate and that global patterns of seasonal water deficits are reflected in the rooting depth of plants. More fine-grained variations in S_0 within regions and biomes are linked to land use and irrigation of agricultural land (Fig. 2), and to topography (Extended Data Fig. 7, 8), and to the water table depth, as indicated by the comparison to plant-level rooting depth observations. The method applied here makes use of the sensitivity of remotely sensed ET to an increasing CWD and thus provides estimates of S_0 even if belowground water stores are never fully depleted during the observational period. Additional analyses, where S_0 was diagnosed from

a simple water balance model with prescribed S_0, confirmed the reliability of the method across a broad range of hydroclimates (Supplementary Text S2 and Supplementary Fig. S5).

S_0 reflects a combination of biotic and abiotic factors. Biotic factors that determine the total plant-available water are, i.a., the rooting depth of the vegetation and plant hydraulic properties. Abiotic factors include the hydroclimate and physical constraints to the rooting depth, related to the texture and depth of the soil and the weathered bedrock[7]. Similarly, human management activities such as irrigation and tile drainage can impact ET, and thus S_0 in agricultural systems. Physical constraints to roots are largely unknown across large scales. Our estimation of S_0 makes no assumptions about such constraints. Instead, the magnitude of the water storage capacity is inferred from mass balance considerations. The CWD we derive from the balances of ET and P imply that the corresponding amount of water is supplied by local storage or supplied from lateral subsurface water convergence - likely a smaller contributor at the ~5 km spatial resolution of the data analysed here[35].

Diagnosed values of S_0 implicitly include water intercepted by leaf and branch surfaces, internal plant water storage, and moisture stored in the topsoil and supplied to soil evaporation. These components are generally smaller in magnitude compared to moisture storage supplied to transpiration[36], and their contribution to ET declines rapidly as CWD increases. Hence, spatial variations in S_0 primarily reflect variations mediated by moisture stored across the root zone.

Particularly in regions with pronounced dry seasons, our estimates of S_0 greatly exceed typical values of the total soil water holding capacity when considering the top 1 or 2 meters of the soil column and texture information from global databases ([31], Extended Data Fig. 5). The discrepancy in magnitude and spatial patterns of total 1 (2) m soil water holding capacity and S_0 diagnosed here hints at a critical role of plant access to deep water and the need to extend the focus beyond moisture in the top 1-2 m of soil for understanding and simulating land-atmosphere exchange[10, 11]. Indications of widespread plant-access to deep water stores are consistent with observations of bedrock-penetrating roots[7, 37] and with evidence for dry-season moisture withdrawal from the weathered bedrock[9, 11]. We note that using the global map of S_{CWDX80} (z_{CWDX80}) for directly parameterizing S_0 (z_r) in models may be misleading in areas with particularly small maximum CWDs and consequently small S_{CWDX80}. Scaling relationships of above and belowground plant architecture[30] and additional effects of how z_r determines access to belowground resources and function (e.g., nutrients, mechanical stability) should be considered.

Underlying the estimates of S_{CWDX80} is the assumption that plant rooting strategies are reflected by CWD extremes with a return period T = 80 years. S_{dSIF} and S_{dEF} provide an independent constraint to test this assumption. Extended Data Fig. 4 suggests that T is not a global constant. A tendency towards higher T emerges with an increasing gridcell average forest cover fraction.

Our analysis identified mountain regions as being characterised by particularly high S_0, in spite of shallow soil and regolith depths[38]. This could be due to hillslope-scale variations in groundwater depth, enabling sustained transpiration during prolonged rain-free periods.

Lateral subsurface flow at scales beyond the resolution of the data used here (~5 km) may additionally supply water for ET and thus contribute to large inferred S_0 in valley bottoms of large drainage basins. Local convergence (divergence) acts to supply (remove) subsurface moisture and sustain (reduce) ET, leading to larger (smaller) CWD values. Without relying on a priori assumptions regarding S_0 or functional dependencies of water stress effect on ET, thermal infrared (TIR)-based remote sensing data (as is used here) offers an opportunity to detect such effects[8]. Our analysis yielded strong contrasts in diagnosed S_0 along topographic gradients (Extended Data Fig. 7, 8). However, further research should assess the accuracy of spatial variations in annual mean ET and potential effects of terrain, where land surface temperature signals on shaded slopes may be mis-interpreted by the ALEXI algorithm as signatures of higher ET.

Our global S_0 estimates are a “snapshot” in time. Regional to continental-scale variations in average tree ages may be associated with changes in rooting depth and S_0. Furthermore, environmental change may trigger changes in vegetation composition and structure[39], with consequences for S_0. Similarly, deforestation implies changes in rooting depth[18], S_0, and the surface energy balance[14]. Such temporal changes are not considered here due to the limited length of available time series of satellite observations (16 years). It remains to be seen whether plasticity in z_r is sufficiently rapid to keep pace with a changing climate with strong and wide-spread increases in rainfall variability[40], and to what degree rising CO_2 alters plant water use and their carbon economy and thereby the costs and benefits of deep roots.

Taken together, constraints available from local z_r observations and from global remote sensing of vegetation activity reveal consistent patterns across multiple spatial scales and suggest widespread plant access to deep water storage, including the weathered bedrock and groundwater, or to other ancillary sources of water, such as irrigation. Our study revealed a tight link of the climatology of water deficits and vegetation sensitivity to drought stress. We demonstrated how land-atmosphere interactions and the Critical Zone water storage capacity are linked with the rooting depth of vegetation and how belowground vegetation structure is influenced by the hydroclimate and topography across the globe.

Methods

Estimating ET

Unbiased estimates of ET during rain-free periods are essential for determining CWD and estimating S_0 and implied z_r. We tested different remote sensing-based ET products and found that the ALEXI-TIR product, which is based on thermal infrared remote sensing[28, 29], exhibits no systematic bias during progressing droughts (Supplementary Text S3 and Supplementary Fig. S6) - in contrast to other ET estimates assessed here. The stability in ET estimates from ALEXI-TIR during drought are enabled by its effective use of information about the surface energy partitioning, allowing inference of ET rates without reliance on *a priori* specified and inherently uncertain surface conductances[44] or shapes of empirical water stress functions[45], and without assumptions of rooting depth or effective S_0. ALEXI-TIR is thus well-suited for estimating actual ET behaviour during drought without introducing circularity in inferring S_0.

Cumulative water deficit estimation

The cumulative water deficit (CWD) is determined here from the cumulative difference of actual ET and the liquid water infiltration to the soil (P_{in}). ET is based on thermal infrared remote sensing, provided by the global ALEXI data product at daily and 0.05° resolution, covering years 2003-2018. Values in energy units of the latent heat flux are converted to mass units accounting for the temperature and air pressure dependence of the latent heat of vaporization following ref.[46]. P_{in} is based on daily reanalysis data of P in the form of rain and snow from WATCH-WFDEI[47]. A simple snow accumulation and melt model[48] is applied to account for the effect of snowpack as a temporary water storage that supplies P_{in} during spring and early summer. Snow melt is assumed to occur above 1°C and with a rate of 1 mm d^{-1} $(°C)^{-1}$. The CWD is derived by applying a running sum of (ET - P_{in}), initiating on the first day when (ET - P_{in}) is positive (net water loss from the soil), and terminating the summation after rain has reduced the running sum to zero (Supplementary Fig. S7). This yields a continuous CWD time series of daily values. In general, $P >$ ET for annual totals. This implies that the CWD summation is initiated at zero each year. In very rare cases, the CWD accumulates over more than one year, and data were discarded if the accumulation extends over five years ("runaway CWD"). All P and snowmelt (P_{in}) is assumed to contribute to reducing the CWD. This implicitly assumes that no runoff occurs while the CWD is above zero. The period between the start and end of accumulation is referred to as a CWD *event*. Within each event, co-varying data, used for analysis, are removed after rain has reduced the CWD to below 90% of its maximum value within the same event. This concerns the analysis of SIF and EF (see below) and avoids effects of relieved water stress by re-wetting topsoil layers before the CWD is fully compensated. The algorithm to determine daily CWD values and events is implemented by the R package *cwd*[49].

Diagnosing S_0 from vegetation activity

By employing first principles for the constraint of the rooting zone water availability on vegetation activity[1], we developed a method to derive how the sensitivity of these fluxes to water stress relates to S_0 and how this sensitivity can be used to reveal effects of access to extensive deep water stores. Two methodologically independent sources of information on vegetation activity were used: the evaporative fraction (EF, defined as ET divided by net radiation), and sun-induced fluorescence (SIF, normalised by incident shortwave radiation). SIF is a proxy for ecosystem photosynthesis[50] and is taken here from a spatially downscaled data product[51] based on GOME-2 data[52, 53]. Since net radiation and shortwave radiation are first-order controls on ET and SIF, respectively, and to avoid effects by seasonally varying radiation inputs, we used EF instead of ET, and considered the ecosystem-level fluorescence yield, quantified as SIF divided by shortwave radiation (henceforth referred to as 'SIF') for all analyses. The resulting estimates for S_0 are referred to as S_{dEF} and S_{dSIF}, respectively.

The principles for relating vegetation activity to the rooting zone water availability were considered as follows. As the ecosystem-level CWD increases, both gross primary production (GPP, ecosystem-level photosynthesis) and ET are limited by the availability of water to plants. Below, we refer to GPP and ET as a generic "vegetation activity" variable $X(t)$. This principle can be formulated, in its simplest form, as a model of $X(t)$ being a linear

function of the remaining water stored along the rooting zone $S(t)$, expressed as a fraction of the total rooting zone water storage capacity S_0:

$$X(t) = X_0 \cdot S(t)/S_0 \quad (1)$$

Following Eq. 1, S_0 can be interpreted as the total rooting zone water storage capacity, or the depth of a "water bucket" that supplies moisture for ET. Following ref.[1] and with $X(t)$ representing ET, the temporal dynamics during rain-free periods (where runoff can be neglected) are described by the differential equation

$$dS/dt = -X(t) \Rightarrow dS/dt = -X_0 \cdot S(t)/S_0 \quad (2)$$

and solved by an exponential function with a characteristic decay time scale λ:

$$X(t) = X_0 \cdot \exp(-[t - t_0]/\lambda) \quad (3)$$

λ is related to S_0 as $S_0 = \lambda X_0$, where X_0 is the initial ET at $S(t_0) = S_0$. In other words, the apparent observed exponential ET decay time scale λ, together with X_0, reflects the total rooting zone water storage capacity S_0.

Fitting exponentials from observational data is subject to assumptions regarding stomatal responses to declines in $S(t)$ and is relatively sensitive to data scatter. Hence, resulting estimates of S_0 may not be robust. With $\mathrm{CWD}(t) = S_0 - S(t)$ and Eq. 1, the relationship of $X(t)$ and $\mathrm{CWD}(t)$ can be expressed as a linear function

$$X(t) = X_0 - X_0/S_0 \cdot \mathrm{CWD}(t) \quad (4)$$

and observational data for $X(t)$ can be used to fit a linear regression model. Its intercept a and slope b can then be used as an alternative, and potentially more robust estimate for S_0:

$$S_0 = -a/b \quad (5)$$

This has the further advantage that estimates for S_0 can be derived using any observable quantity of vegetation activity $X(t)$ (not just ET as in ref.[1]) under the assumption that activity attains zero at the point when the CWD reaches the total rooting zone water storage capacity; i.e., $X(t^*) = 0$ for $\mathrm{CWD}(t^*) = S_0$.

Here, we use a spatially downscaled product of sun-induced fluorescence (SIF[51]), normalised by incident shortwave radiation (WATCH-WFDEI data[47]), and the evaporative fraction (EF), defined as the ratio of ET (ALEXI-ET data[29]) over net radiation (GLASS data[54]), as two alternative, normalised proxies for water-constrained vegetation activity, termed X'. Normalisation by net radiation and incident shortwave radiation respectively removes effects by seasonally varying energy available for vegetation activity. X'_0 is thus assumed to be stationary over time and the relationship of $X'(t)$ and $\mathrm{CWD}(t)$ is interpreted here as a reflection of effects by belowground water availability and used to derive S_{dSIF} and S_{dEF}. All data used for X'_0 are provided at 0.05° and daily resolution.

S_{dSIF} and S_{dEF} were then derived based on the relationship of EF and normalised SIF versus CWD, guided by Eq. 5. The relationship was analysed for each pixel with pooled data belonging to the single largest CWD event of each year, and using the 90% quantile of EF and normalised SIF within 50 evenly spaced bins along the CWD axis. Binning and considering percentiles were chosen to reduce effects of vegetation activity reduction due to factors other than water stress (CWD). We then tested for each pixel whether the data support the model of a single linear decline of SIF (EF) with increasing CWD (Eq. 5), or, alternatively, a segmented regression model with one or two change points, using the R package *segmented*[55]. The model with the lowest Bayesian Information Criterion (BIC) was chosen and S_{dSIF} and S_{dEF} were quantified only for pixels where no significant change point was detected and where the regression of EF (SIF) vs. CWD had a significantly negative slope. "Flattening" EF (SIF) vs. CWD relationships were identified where a significant change point was detected and where the slope of the second regression segment was significantly less negative ($p = 0.05$ of t-test) compared to the slope of the first segment. Examples, visualising the diagnosing of S_0 from EF, are given in Supplementary Fig. S8. We performed additional tests of the method's reliability in estimating S_0 by deriving S_{dEF} from simulations of the ecosystem water balance and ET, where S_0 was prescribed, using the SPLASH model[46]. This demonstrates that the method applied for S_{dSIF} and S_{dEF} yields accurate estimates of S_0 across all climatic conditions and independent of the size of S_0 (Supplementary Text S2 and Supplementary Fig. S5).

Diagnosing S_0 from cumulative water deficits

Following ref.[2], the rooting zone water storage capacity S_0 is estimated based on CWD extremes occurring with a return period of T years. Magnitudes of extremes with a given return period T ($S_{CWDX\,T}$) are estimated by fitting an extreme value distribution (Gumbel) to the annual maximum CWD values for each pixel separately, using the *extRemes* R package[56]. Values $S_{CWDX\,T}$. $S_{CWDX\,T}$ are translated into an effective depth $z_{CWDX\,T}$ using estimates of the plant-available soil water holding capacity, based on soil texture data from a gridded version of the Harmonized World Soil Database[31, 32] and pedo-transfer functions derived by ref.[33]. Associations of $S_{CWDX\,T}$ and topography were analysed considering the Compound Topography Index[57] and elevation from ETOPO1[58]. The Compound Topography Index is a measure for subsurface flow convergence and the water table depth based on the topographical setting[59].

Estimating return periods

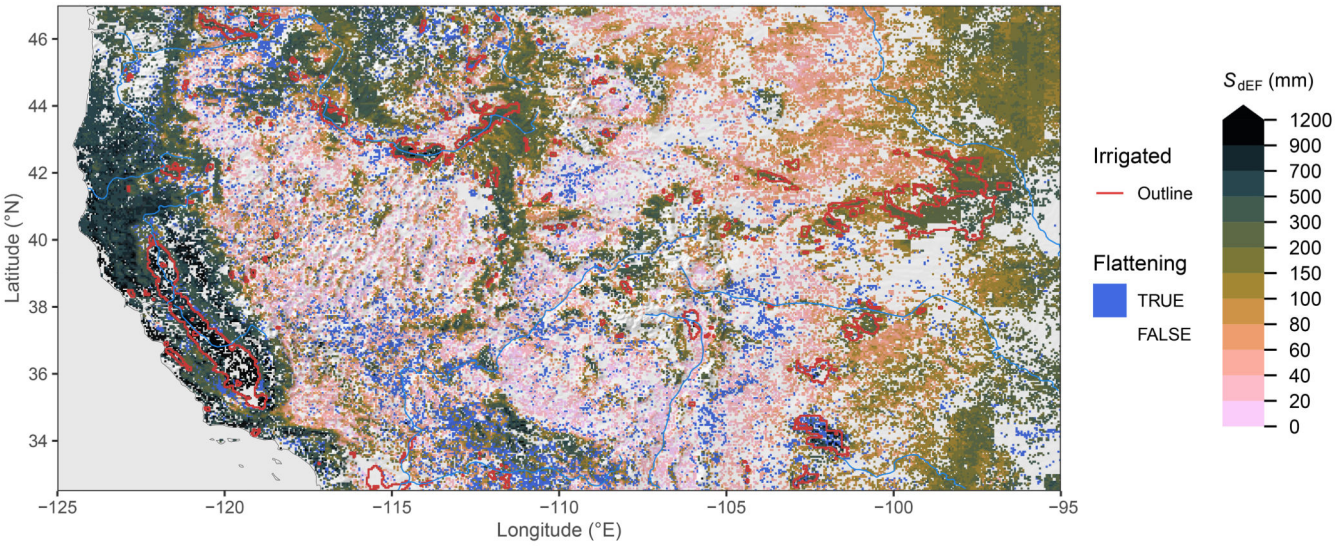
Diagnosed values of S_{dSIF} and S_{dEF} provide a constraint on the return period T. To yield stable estimates of T and avoid effects of the strong non-linearity of the function to derive T from the fitted extreme value distributions and magnitudes estimated by S_{dSIF} and S_{dEF}, we pooled estimates S_{dSIF} (S_{dEF}) and $S_{CWDX\,T}$ values within 1° pixels ($\leq$400 values). A range of discrete values T was screened (10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500 years) and the best estimate T was chosen based on comparison to S_{dSIF} (T_{SIF}) and to S_{dEF} (T_{EF}), i.e., where the absolute value of the median of the logarithm of the bias was minimal. Relationships of best matching T with topography (measured by the Compound Topography Index[57]) and with the forest cover fraction (MODIS MOD44B[60]) were analysed.

Rooting depth estimation and observations

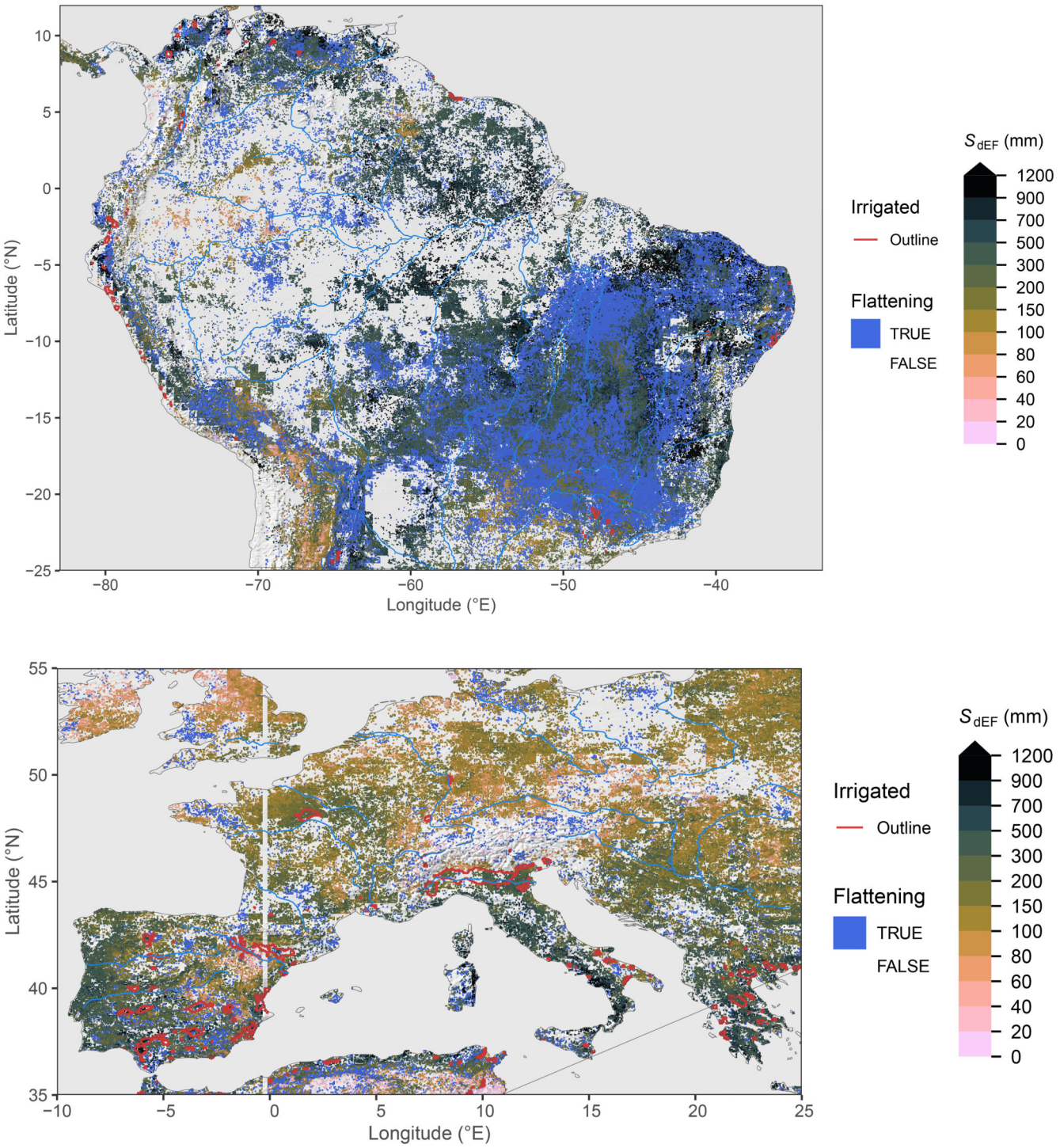
We converted root zone water storage capacity estimates, S_{CWDX80}, to a corresponding apparent rooting depth (z_{CWDX80}) using a global soil texture map[31, 32]. The conversion of S_{CWDX80} into a corresponding depth z_{CWDX80} accounts for topsoil and subsoil texture and water holding capacity (WHC) along the rooting profile (see Methods, Fig. 3b), and - in view of lacking information with global coverage about the WHC of the weathered bedrock - assuming uniform subsoil texture extending below 30 cm depth. The comparison of biome-level quantities (instead of a direct point-by-point comparison) avoids the inevitable scale mismatch between *in situ* plant-level observations and global remote sensing data.

The observational rooting depth data set (N = 5524) was compiled by[30] by combining and complementing published datasets from refs.[22] and[7]. The data include observations of the maximum rooting depth of plants taken from 361 published studies plus additional environmental and climate data. z_r was taken as the plant's maximum rooting depth. Data were aggregated by sites (N = 1705) based on longitude and latitude information. Sites were classified into biomes using maps of terrestrial ecoregions[43]. Quantiles (10%, 90%) were determined for each biome. For a subset of the data (359 sites) where parallel measurements of the water table depth (WTD) was available, we conducted the same analysis, but took the minimum of WTD and z_r.

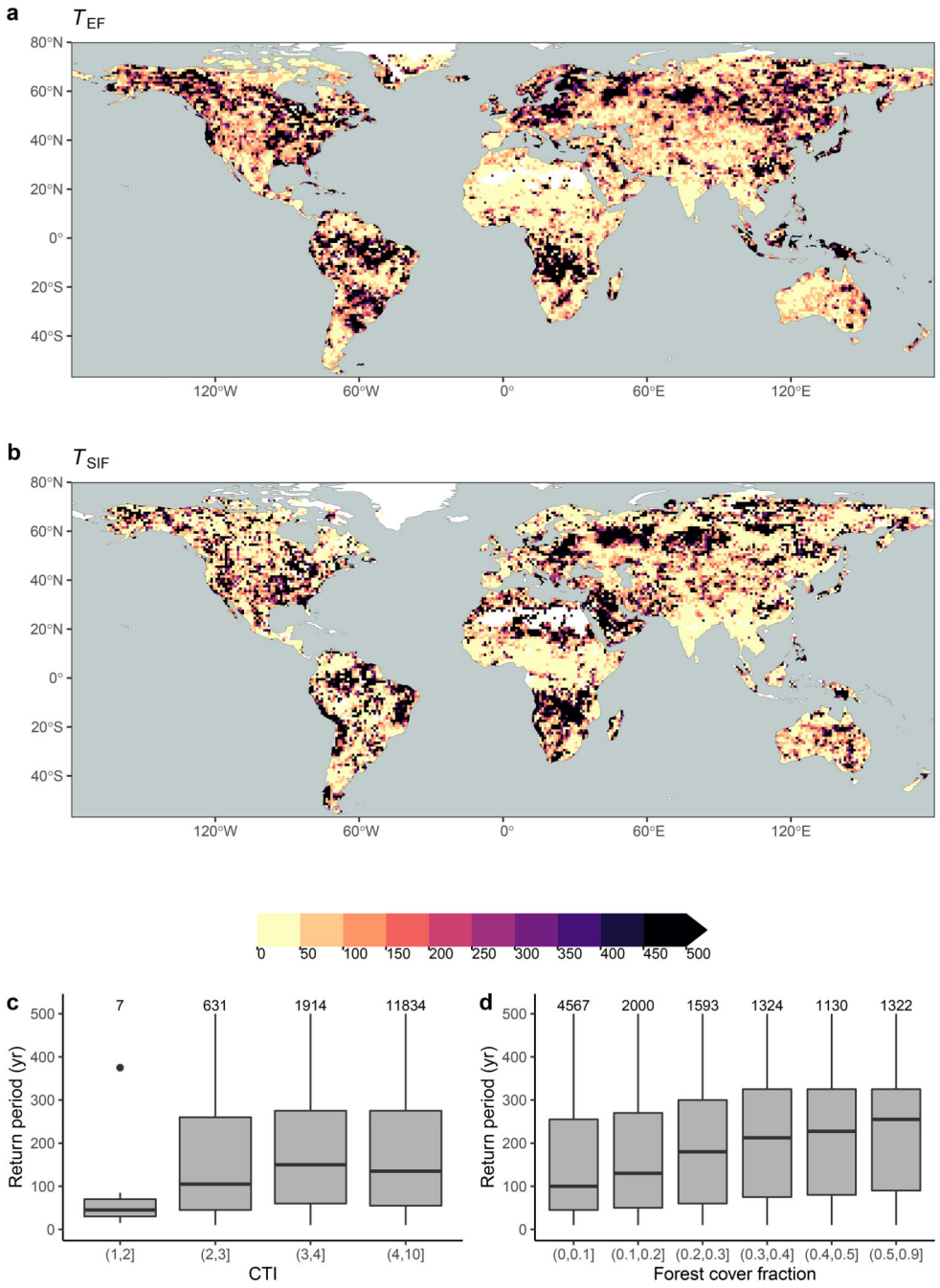
Extended Data



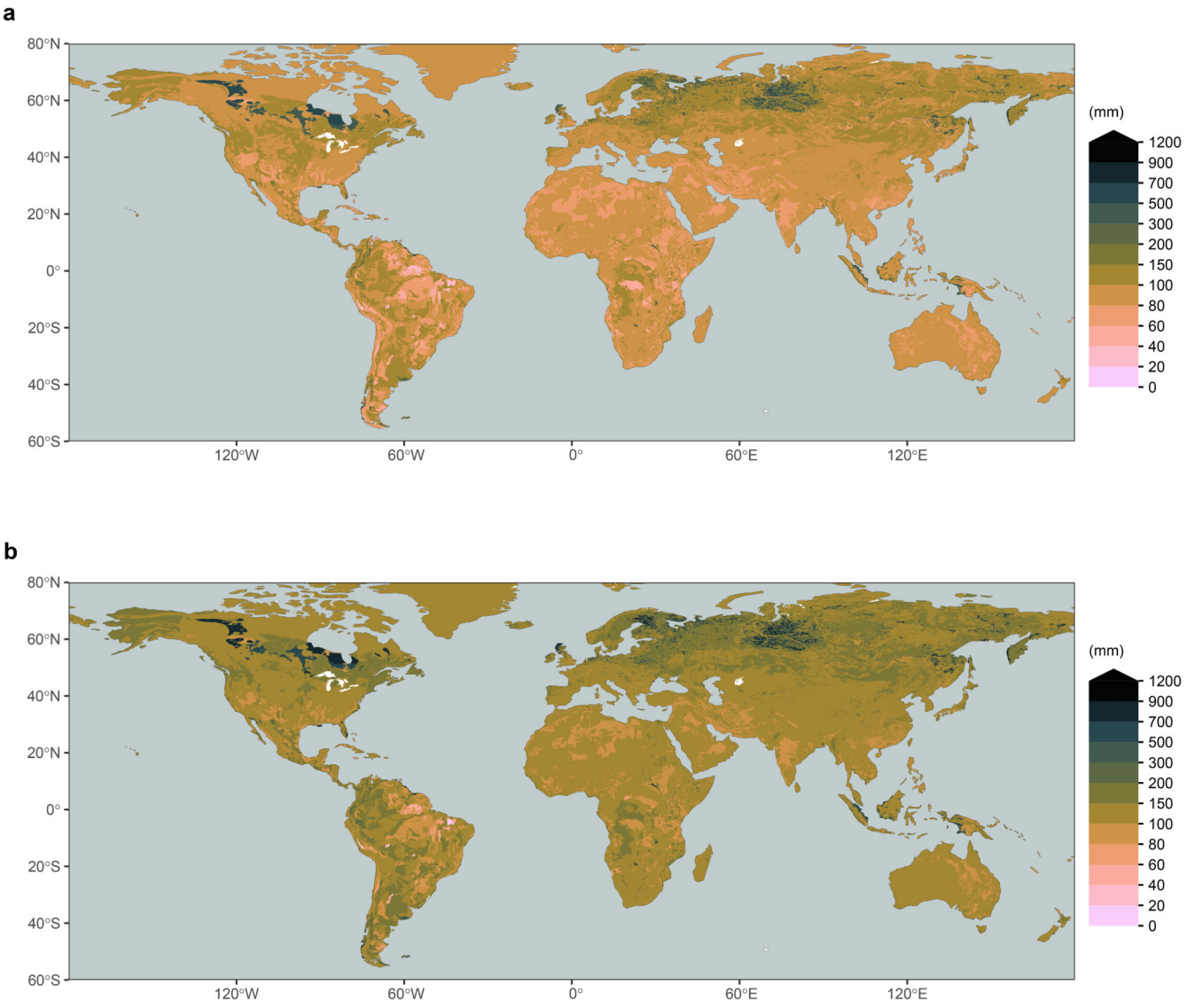
Europe PMC Funders Author Manuscripts



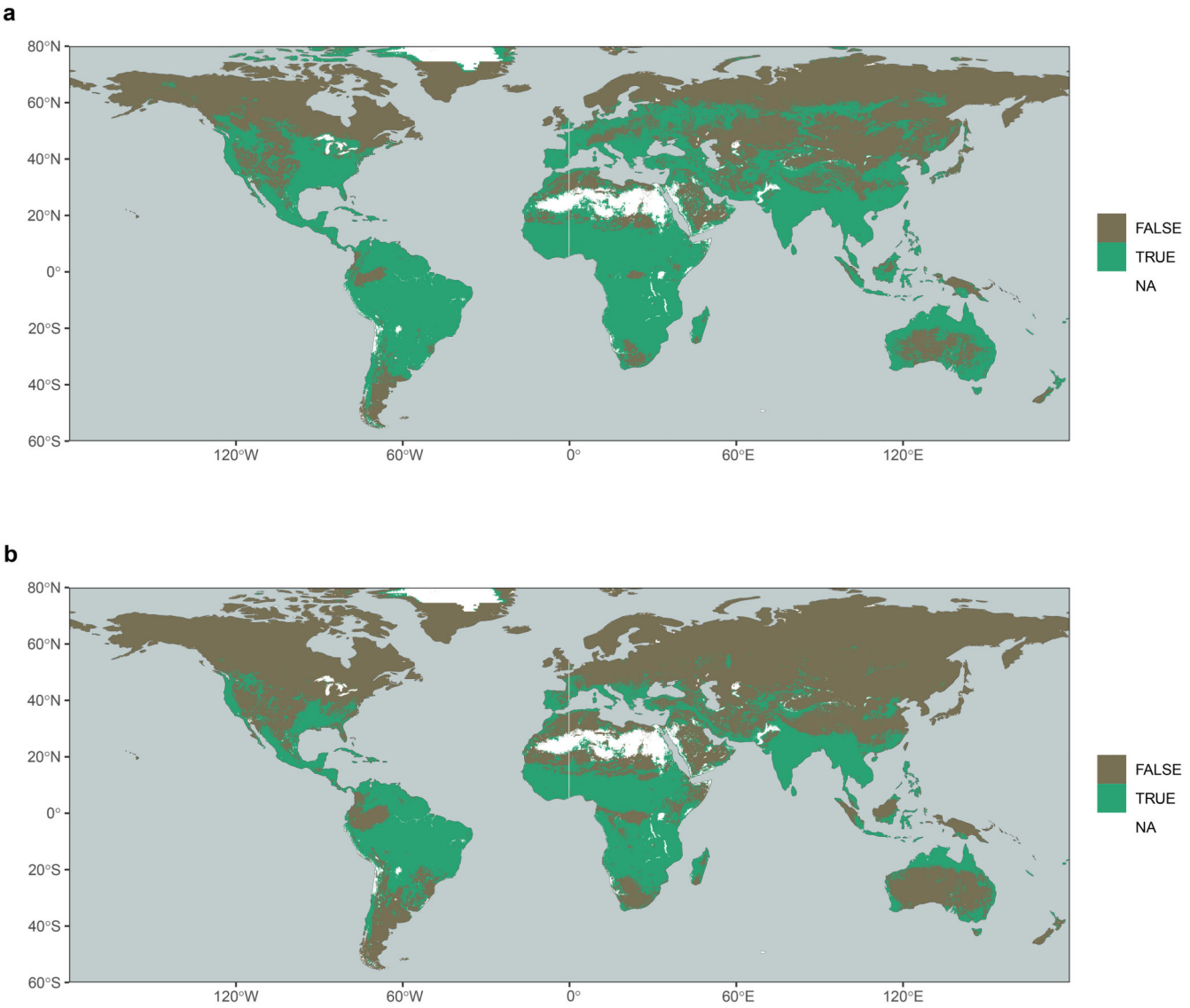
Europe PMC Funders Author Manuscripts

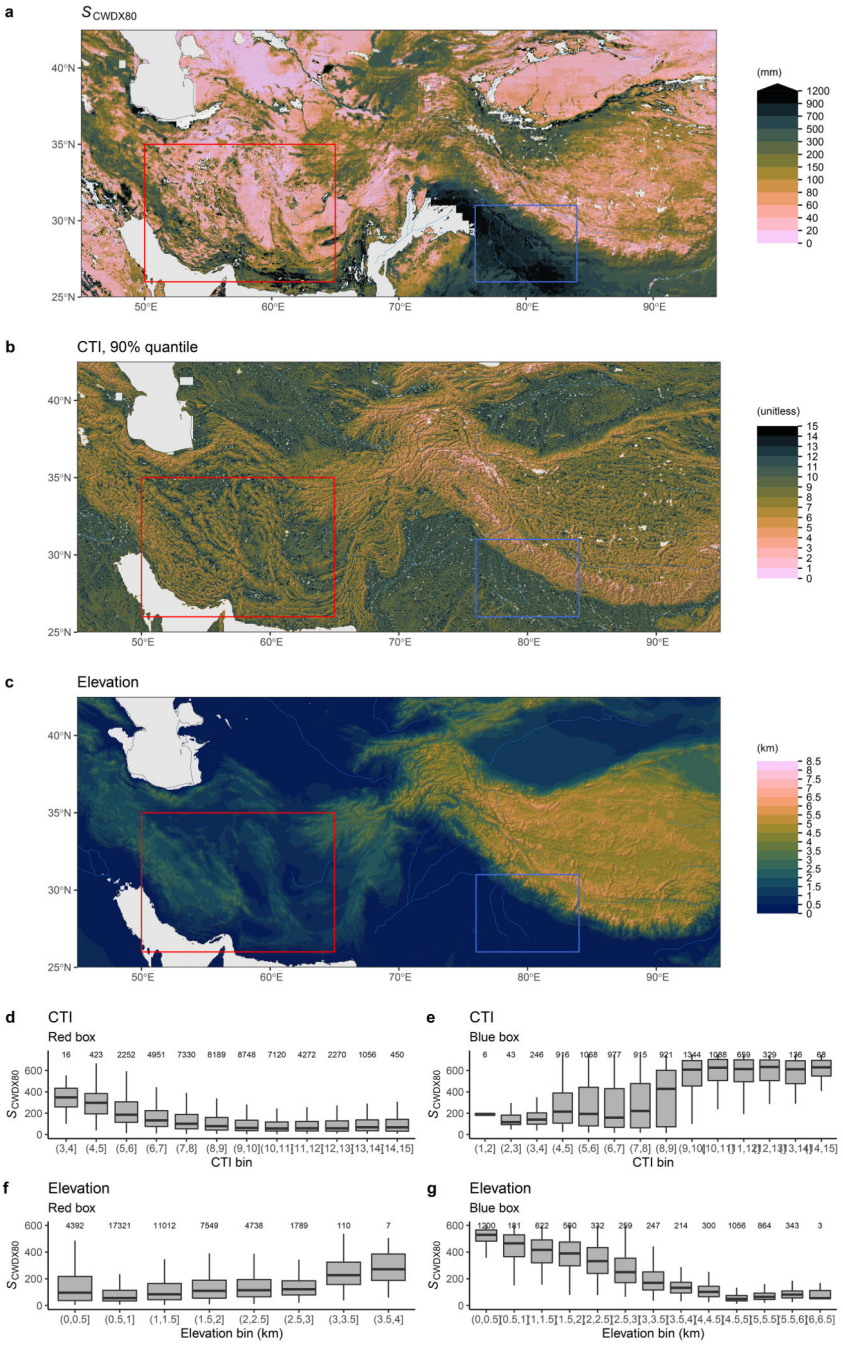


Europe PMC Funders Author Manuscripts

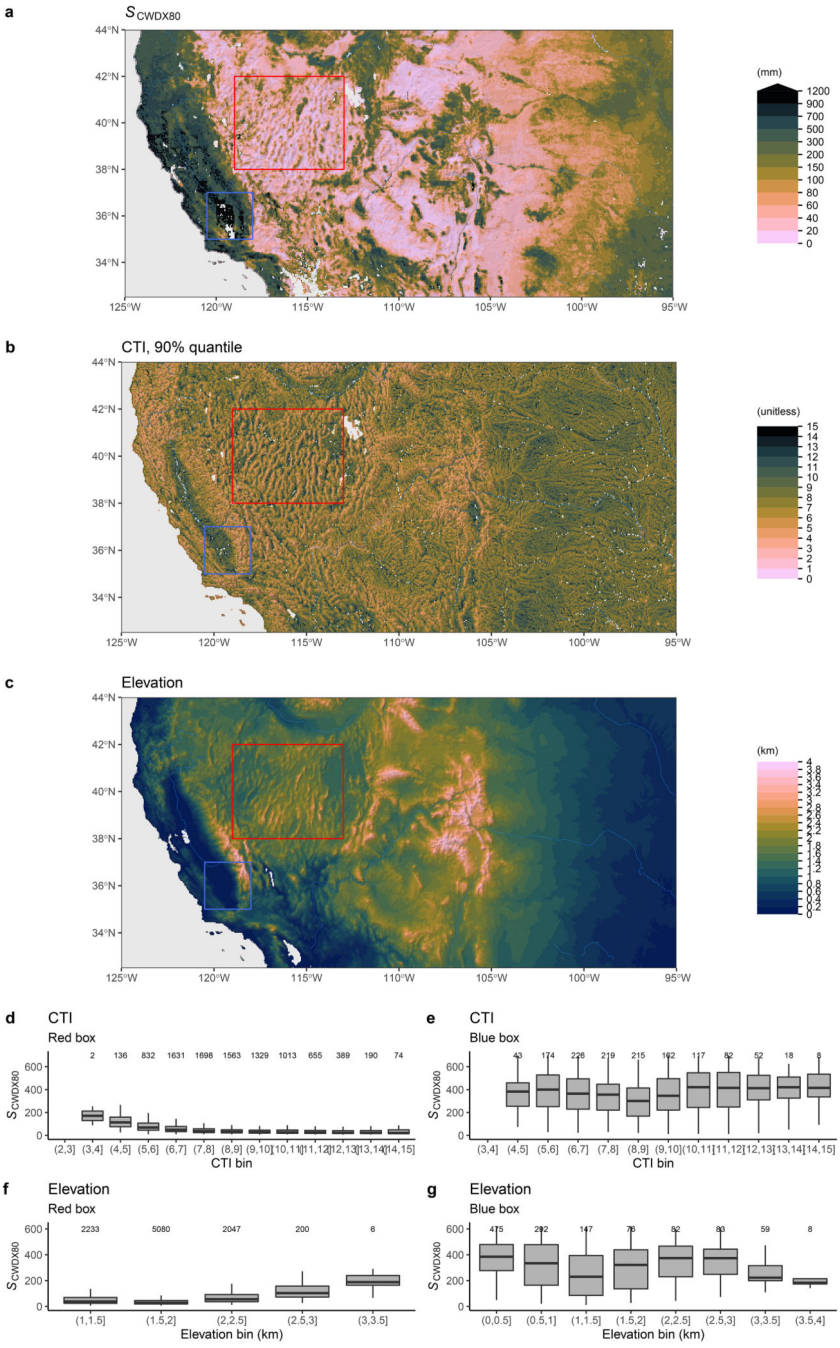


Europe PMC Funders Author Manuscripts





Europe PMC Funders Author Manuscripts



Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

B.D.S was funded by the Swiss National Science Foundation grant no. PCEFP2_181115. This work is a contribution to the LEMONTREE (Land Ecosystem Models based On New Theory, obseRvations and ExperimEnts) project, funded through the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program (BDS). We acknowledge ICDC, CEN, University of Hamburg for data support. S.J.T.D.

Europe PMC Funders Author Manuscripts

was funded by the NSF GRFP and LTER, and the NASEM Ford Foundation Fellowship Program. A.G.K. was supported by NSF DEB-1942133.

Data availability

Global datasets of S_{CWDX80} and z_{CWDX80} are available on *Zenodo*[61]. The rooting depth data are published separately[30].

Code availability

The CWD calculation from ET and P_{in} time series is implemented by the R package *cwd*[49]. All code for this analysis is published on *Zenodo*[62].

References

1. Teuling AJ, Seneviratne SI, Williams C, Troch PA. Observed timescales of evapotranspiration response to soil moisture. Geophys. Res Lett. 2006; 33
2. Gao H, et al. Climate controls how ecosystems size the root zone storage capacity at catchment scale. Geophys Res Lett. 2014; 41: 7916–7923.
3. Milly PCD. Climate, soil water storage, and the average annual water balance. Water Resour Res. 1994; 30: 2143–2156.
4. Hahm WJ, et al. Low subsurface water storage capacity relative to annual rainfall decouples mediterranean plant productivity and water use from rainfall variability. Geophys Res Lett. 2019; 46: 6544–6553.
5. Seneviratne SI, et al. Investigating soil moisture-climate interactions in a changing climate: A review. Earth Science Rev. 2010; 99: 125–161.
6. Thompson SE, et al. Comparative hydrology across AmeriFlux sites: The variable roles of climate, vegetation, and groundwater. Water Resour Res. 2011; 47
7. Fan Y, Miguez-Macho G, Jobbágy EG, Jackson RB, Otero-Casal C. Hydrologic regulation of plant rooting depth. Proc Natl Acad Sci U S A. 2017; 114: 10572–10577. [PubMed: 28923923]
8. Hain CR, Crow WT, Anderson MC, Tugrul Yilmaz M. Diagnosing neglected soil moisture source-sink processes via a thermal infrared-based two-source energy balance model. J Hydrometeorol. 2015; 16: 1070–1086.
9. Rempe DM, Dietrich WE. Direct observations of rock moisture, a hidden component of the hydrologic cycle. Proc Natl Acad Sci U S A. 2018; 115: 2664–2669. [PubMed: 29490920]
10. Dawson TE, Jesse Hahm W, Crutchfield-Peters K. Digging deeper: what the critical zone perspective adds to the study of plant ecophysiology. New Phytol. 2020; 226: 666–671. [PubMed: 31912507]
11. McCormick EL, et al. Widespread woody plant use of water stored in bedrock. Nature. 2021; 597: 225–229. DOI: 10.1038/s41586-021-03761-3 [PubMed: 34497393]
12. Maxwell RM, Condon LE. Connections between groundwater flow and transpiration partitioning. Science. 2016; 353: 377–380. [PubMed: 27463671]
13. Schlemmer L, Schär C, Lüthi D, Strebel L. A groundwater and runoff formulation for weather and climate models. J Adv Model Earth Syst. 2018; 10: 1809–1832.
14. Teuling AJ, et al. Contrasting response of european forest and grassland energy exchange to heatwaves. Nat Geosci. 2010; 3: 722–727.
15. Koirala S, et al. Global distribution of groundwater-vegetation spatial covariation. Geophys Res Lett. 2017; 44: 4134–4142.
16. Esteban EJL, Castilho CV, Melgaço KL, Costa FRC. The other side of droughts: wet extremes and topography as buffers of negative drought effects in an amazonian forest. New Phytol. 2021; 229: 1995–2006. [PubMed: 33048346]

Europe PMC Funders Author Manuscripts

17. Liu Y, Konings AG, Kennedy D, Gentine P. Global coordination in plant physiological and rooting strategies in response to water stress. Glob Biogeochem Cycles. 2021; 35 doi: 10.1029/2020GB006758
18. Schenk HJ, Jackson RB. The global biogeography of roots. Ecol Monogr. 2002; 72: 311–328.
19. Canadell J, et al. Maximum rooting depth of vegetation types at the global scale. Oecologia. 1996; 108: 583–595. [PubMed: 28307789]
20. Weaver JE, Darland RW. Soil-Root relationships of certain native grasses in various soil types. Ecol Monogr. 1949; 19: 303–338.
21. Chitra-Tarak R, et al. Hydraulically-vulnerable trees survive on deep-water access during droughts in a tropical forest. New Phytol. 2021; 231: 1798–1813. DOI: 10.1111/nph.17464 [PubMed: 33993520]
22. Schenk HJ, Jackson RB. Mapping the global distribution of deep roots in relation to climate and soil characteristics. Geoderma. 2005; 126: 129–140.
23. Franklin O, et al. Organizing principles for vegetation dynamics. Nat Plants. 2020; 6: 444–453. [PubMed: 32393882]
24. Kleidon A, Heimann M. A method of determining rooting depth from a terrestrial biosphere model and its impacts on the global water and carbon cycle. Glob Chang Biol. 1998; 4: 275–286.
25. Schymanski SJ, Sivapalan M, Roderick ML, Hutley LB, Beringer J. An optimality-based model of the dynamic feedbacks between natural vegetation and the water balance. Water Resour Res. 2009; 45
26. Wang-Erlandsson L, et al. Global root zone storage capacity from satellite-based evaporation. Hydrol Earth Syst Sci. 2016; 20: 1459–1481. DOI: 10.5194/hess-20-1459-2016
27. Knapp AK, Smith MD. Variation among biomes in temporal dynamics of aboveground primary production. Science. 2001; 291: 481–484. [PubMed: 11161201]
28. Anderson M. A Two-Source Time-Integrated model for estimating surface fluxes using thermal infrared remote sensing. Remote Sens Environ. 1997; 60: 195–216.
29. Hain CR, Anderson MC. Estimating morning change in land surface temperature from MODIS Day/Night observations: Applications for surface energy balance modeling. Geophys Res Lett. 2017; 44: 9723–9733. [PubMed: 29403120]
30. Tumber-Dávila SJ, Schenk HJ, Du E, Jackson RB. Plant sizes and shapes above-and belowground and their interactions with climate. New Phytol. 2022; doi: 10.1111/nph.18031
31. Food and Agriculture Organization of the United Nations. Harmonized World Soil Database: Fao/Iiasa/isric/isscas/jrc (Version 10). Food Agriculture Org; 2008.
32. Wieder, W. Regridded harmonized world soil database v1.2. 2014.
33. Balland V, Pollacco JAP, Arp PA. Modeling soil hydraulic properties for a wide range of soil conditions. Ecol Model. 2008; 219: 300–316.
34. Agee E, et al. Root lateral interactions drive water uptake patterns under water limitation. Adv Water Resour. 2021; 151 103896
35. Krakauer NY, Li H, Fan Y. Groundwater flow across spatial scales: importance for climate modeling. Environ Res Lett. 2014; 9 034003 doi: 10.1088/1748-9326/9/3/034003
36. Stoy PC, et al. Reviews and syntheses: Turning the challenges of partitioning ecosystem evaporation and transpiration into opportunities. Biogeosciences. 2019; 16: 3747–3775.
37. Jackson RB, Moore LA, Hoffmann WA, Pockman WT, Linder CR. Ecosystem rooting depth determined with caves and dna. Proc Natl Acad Sci United States Am. 1999; 96: 11387–11392.
38. Pelletier JD, et al. A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. J Adv Model Earth Syst. 2016; 8: 41–65.
39. Parmesan C, Hanley ME. Plants and climate change: complexities and surprises. Annals Bot. 2015; 116: 849–864. DOI: 10.1093/aob/mcv169
40. Pendergrass AG, Knutti R, Lehner F, Deser C, Sanderson BM. Precipitation variability increases in a warmer climate. Sci Reports. 2017; 7
41. Siebert S, et al. Development and validation of the global map of irrigation areas. Hydrol Earth Syst Sci. 2005; 9: 535–547.

42. Friedl MA, et al. MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens Environ. 2010; 114: 168–182.
43. Olson DM, et al. Terrestrial ecoregions of the world: A new map of life on earth. BioScience. 2001; 51: 933.
44. Mu Q, Heinsch FA, Zhao M, Running SW. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sens Environ. 2007; 111: 519–536.
45. Fisher JB, et al. ECOSTRESS: NASA's next generation mission to measure evapotranspiration from the international space station. Water Resour Res. 2020; 56 e2019WR026058
46. Davis TW, 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. 2017; 10: 689–708.
47. Weedon GP, et al. The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-Interim reanalysis data. Water Resour Res. 2014; 50: 7505–7514.
48. Orth R, Koster RD, Seneviratne SI. Inferring soil moisture memory from streamflow observations using a simple water balance model. J Hydrometeorol. 2013; 14: 1773–1790.
49. Stocker, B. cwd v1.0: R package for cumulative water deficit calculation. 2021.
50. Zhang Y, et al. Model-based analysis of the relationship between sun-induced chlorophyll fluorescence and gross primary production for remote sensing applications. Remote Sens Environ. 2016; 187: 145–155.
51. Duveiller G, et al. A spatially downscaled sun-induced fluorescence global product for enhanced monitoring of vegetation productivity. Earth Syst Sci Data. 2020; 12: 1101–1116.
52. 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. Atmospheric Meas Tech. 2013; 6: 2803–2823.
53. Köhler P, Guanter L, Joiner J. A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data. Atmospheric Meas Tech. 2015; 8: 2589–2608.
54. Jiang B, et al. Validation of the surface daytime net radiation product from version 4.0 GLASS product suite. IEEE Geosci Remote Sens Lett. 2019; 16: 509–513.
55. Muggeo VMR. Estimating regression models with unknown break-points. Stat Med. 2003; 22: 3055–3071. [PubMed: 12973787]
56. Gilleland E, Katz RW. extremes2.0: An extreme value analysis package inr. J Stat Softw. 2016; 72
57. Marthews TR, Dadson SJ, Lehner B, Abele S, Gedney N. High-resolution global topographic index values for use in large-scale hydrological modelling. Hydrol Earth Syst Sci. 2015; 19: 91–104. DOI: 10.5194/hess-19-91-2015
58. National Geophysical Data Center, NESDIS, NOAA, U.S. Department of Commerce. Etopo1, global 1 arc-minute ocean depth and land elevation from the us national geophysical data center (ngdc). 2011.
59. Beven KJ, Kirkby MJ. A physically based, variable con-tributing area model of basin hydrology. Hydrol Sci J. 1979; 24: 43–69. DOI: 10.1080/02626667909491834
60. Hansen MC, Townshend JRG, DeFries RS, Carroll M. Estimation of tree cover using MODIS data at global, continental and regional/local scales. Int J Remote Sens. 2005; 26: 4359–4380.
61. Stocker, BD. Scwdx80 and zcwdx80. 2021.
62. Stocker, B. stineb/mct: v3.0: re-submission to Nature Geoscience. 2022.

Europe PMC Funders Author Manuscripts

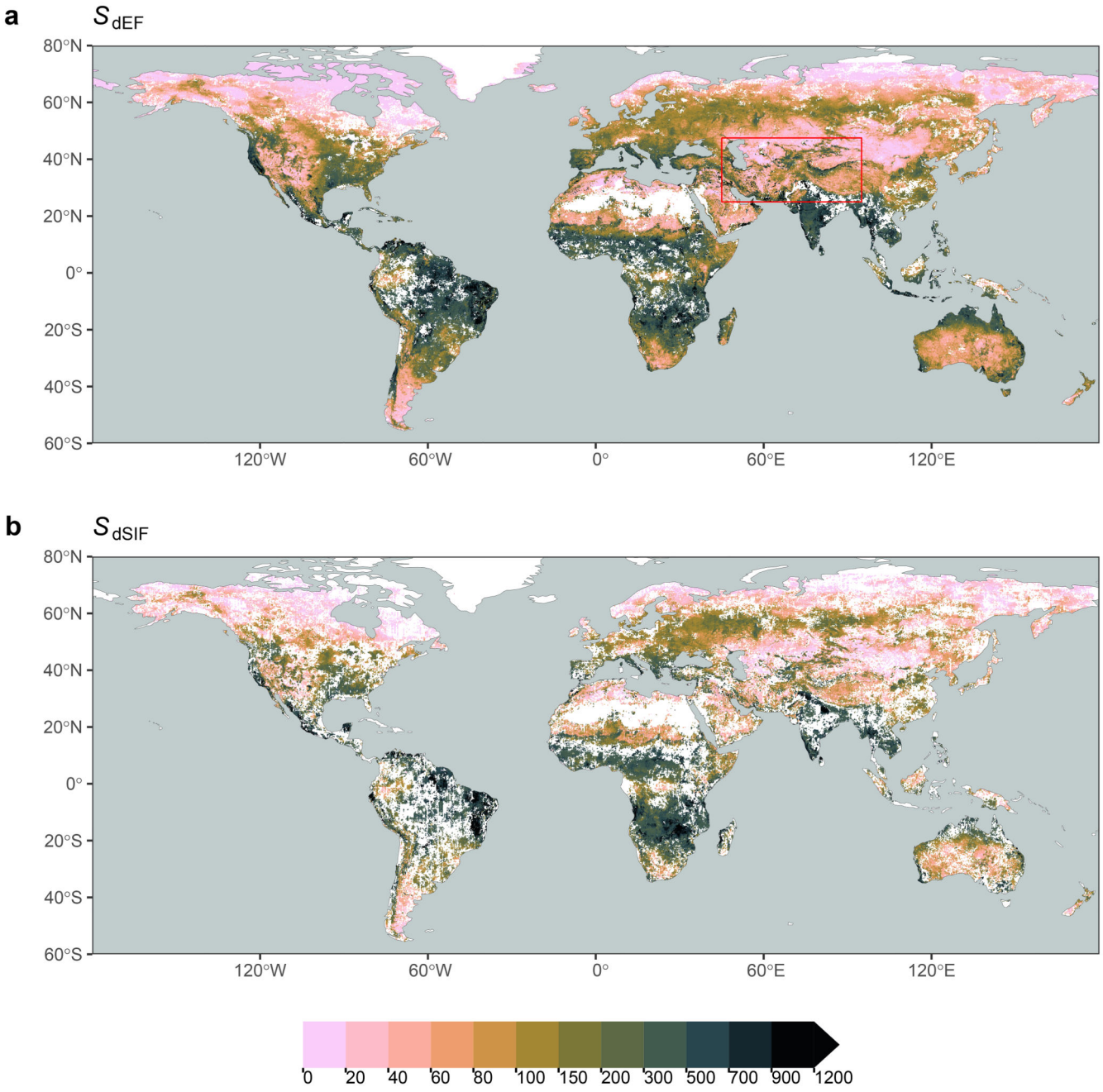


Figure 1.

Rooting zone water storage capacity (mm) estimated from the sensitivity of the evaporative fraction (S_{dEF}, top panel a) and sun-induced fluorescence (S_{dSIF}, bottom panel b) to the cumulative water deficit (CWD). The red box in (a) shows the outline of the magnified map provided in Fig. 2. Data shown is aggregated to 0.1° resolution. Blank cells (white) mark areas where all underlying cells at the original 0.05° resolution did not exhibit a significant and single, linearly declining relationship with increasing CWD.

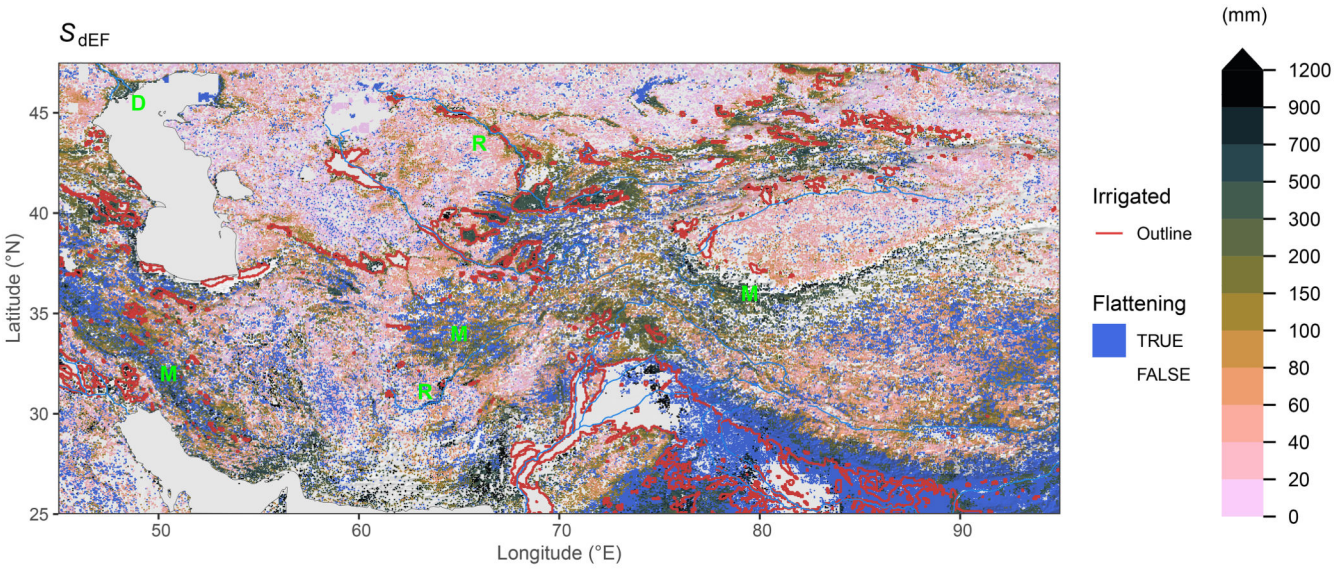


Figure 2.

Rooting zone water storage capacity in Central Asia, estimated from the evaporative fraction (S_{dEF}). Blue areas ("flattening") show grid cells where a significant reduction in the slope in EF vs. CWD was identified beyond a certain threshold. S_{dEF} values are not calculated for gridcells classified as "Flattening". Red lines show outlines of major irrigated areas, i.e. where the irrigated land area fraction is above 30%[41]. Information about irrigated areas was used only for mapping here, but is not used for other parts of the analysis. Blank grid cells (white) indicate areas with a sustained imbalance of ET being greater than *P*. Green letters indicate locations of mountains ('M'), rivers ('R'), and delta ('D'), referred to in the main text. Additional regional maps are provided by Extended Data Figs. 1-3.

Europe PMC Funders Author Manuscripts

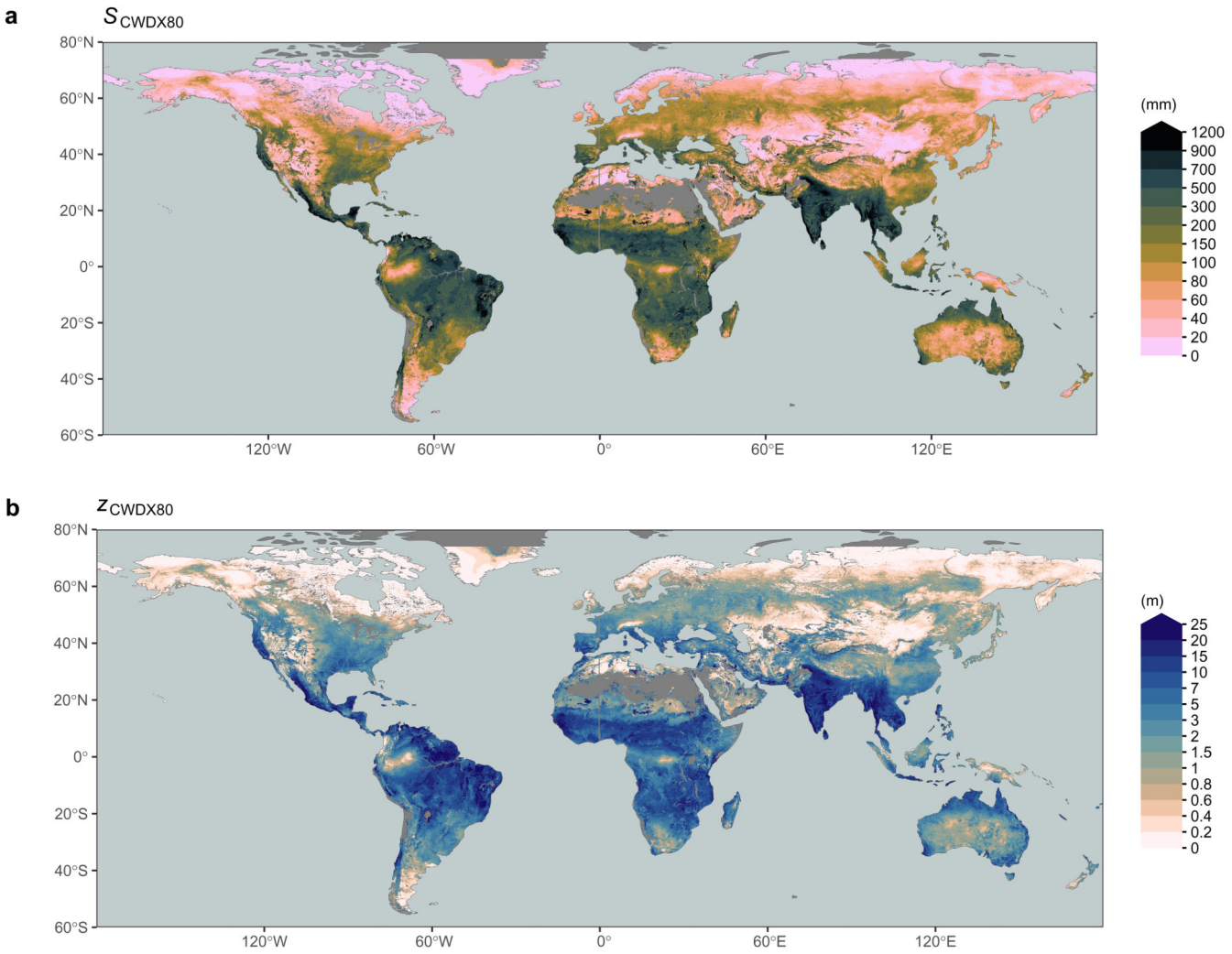


Figure 3.

Spatial variations of (a) the rooting zone water storage capacity, estimated by S_{CWDX80} (mm), and (b) the apparent plant rooting depth z_{CWDX80} (m). Values are remapped to a 0.1° resolution. Blank grid cells (grey) are either permanent inland water bodies and ocean, or locations with long-term accumulation of water deficits. Values are removed in gridcells where more than 99% is non-vegetation surface based on MODIS Landcover[42].

Europe PMC Funders Author Manuscripts

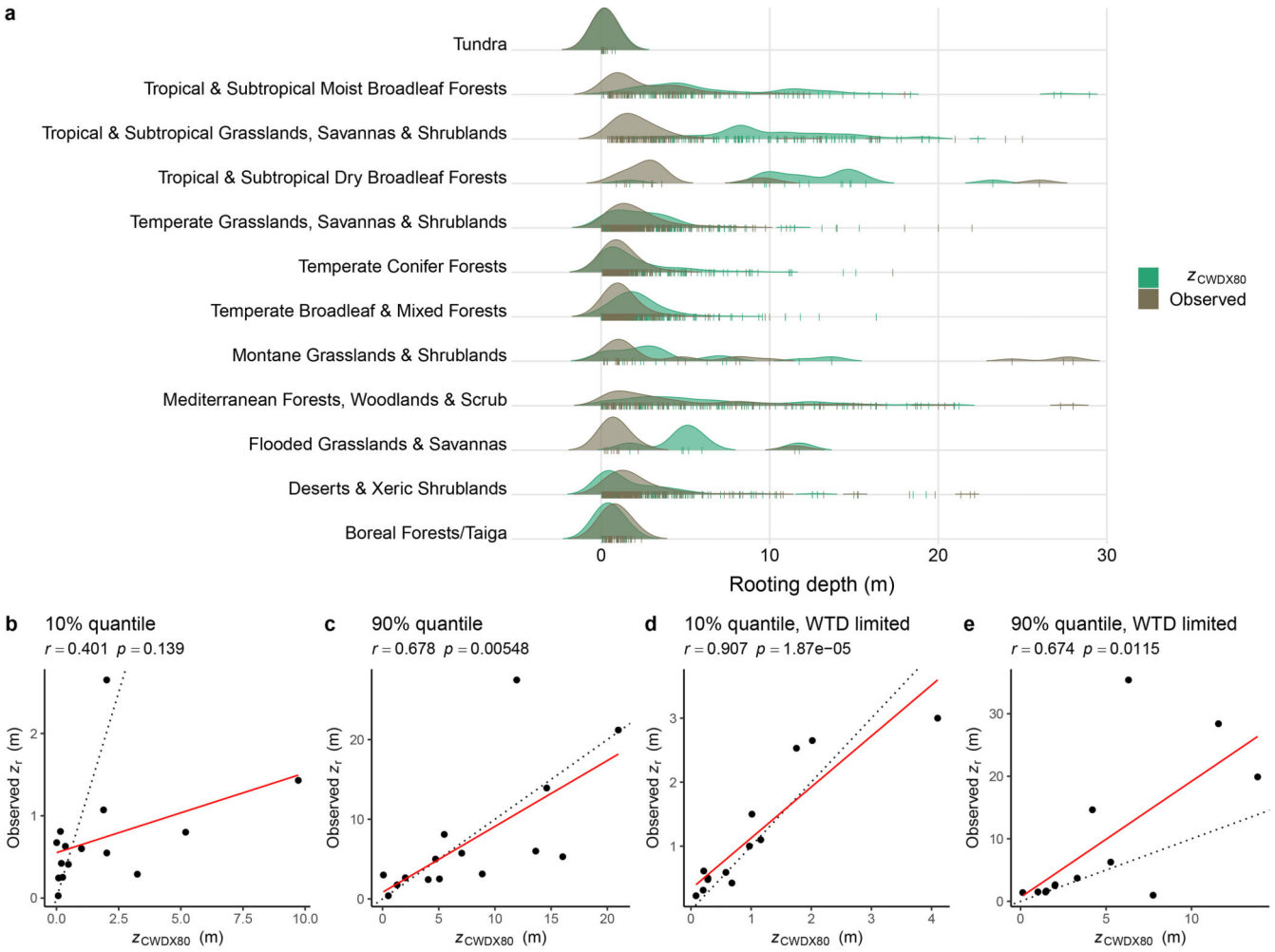


Figure 4. Modelled and observed rooting depth by biomes.

(a) Kernel density estimates of observed and predicted (z_{CWDX80}) rooting depth by biomes, based on data aggregated by sites, shown by vertical colored tick marks. 10% and 90% quantiles of observed vs. predicted (z_{CWDX80}) rooting depth by biome of all data (b,c), and of a subset of the data where the water table depth (WTD) was measured along with rooting depth (d,e). Classification of sites into biomes was done based on[43]. Dotted lines in b-e represent the 1:1 line. In subtitles of panels (b-e), *r* is the Pearson's correlation coefficient, and *p* is the test statistic based on Pearson's product moment correlation coefficient.