Relevance of near-surface soil moisture vs. terrestrial water storage for global vegetation functioning

Prajwal Khanal[1,2,3], Anne J. Hoek Van Dijke[1], Timo Schaffhauser[2], Wantong Li[1], Sinikka J. Paulus[1,3,4,5], Chunhui Zhan[1,56], René Orth[1,45]

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Hans-Knöll-Straße 10, 07745 Jena, Germany

[2]Chair of Hydrology and River Basin Management, Technical, University of Munich, Arcisstraße 21, 80333 Munich, Germany

[3]Chair[3]Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Langezijds, 7500 AE, Enschede, The Netherlands

[4]Chair of Terrestrial Ecohydrology, University of Jena, Burgweg 11, 07749 Jena, Germany

[4]Chair[5]Chair of Modeling of Biogeochemical Systems, Faculty of Environment and Natural Resources, University of Freiburg, Tennenbacher Straße 4, 79106 Freiburg, Germany

[5]Chair[6]Chair of Land Surface-Atmosphere Interactions, Technical University of Munich, TUM School of Life Sciences Weihenstephan, 85354 Freising, Germany

Correspondence to: Prajwal Khanal (ktm.prajwalkhanal@gmail.com)

Abstract. Soil water availability is an essential prerequisite for vegetation functioning. Vegetation takes up water from varying soil depths depending on the characteristics of their rooting system and soil moisture availability across depth. The depth of vegetation water uptake is largely unknown across large spatial scales as a consequence of sparse ground measurements. At the same time, emerging satellite-derived observations of vegetation functioning, surface soil moisture and terrestrial water storage, present an opportunity to assess the depth of vegetation water uptake globally. In this study, we characterise vegetation functioning through the Near-Infrared Reflectance of Vegetation (NIRv),) and compare its relation to (i) near-surface soil moisture from ESA-CCI and (ii) total water storage from GRACE at the monthly time scale during the growing season. The relationships are quantified through partial correlations to mitigate the influence of confounding factors such as energy and other water-related variables. We find that vegetation functioning is generally more strongly related to near-surface soil moisture, particularly in semi-arid regions and areas with low tree cover. In contrast, in regions with high tree cover and in arid regions, the correlation with terrestrial water storage is comparable to or even higher than with near-surface soil moisture, indicating that trees can and do make use of their deeper rooting systems to access deeper soil moisture, similar to vegetation in arid regions. At the same time we note that this comparison is hampered by different noise levels in these satellite data streams. In line with this, an attribution analysis that examines the relative importance of these soil water storages for vegetation reveals that they are controlled by (i) water availability influenced by the climate and (ii) vegetation type reflecting adaptation of ecosystems to local water resources. Next to variations in space, the vegetation water uptake depth also varies in time. During dry periods, the relative importance of terrestrial water storage increases, highlighting the relevance of deeper water

33 resources during rain-scarce periods. Overall, the synergistic exploitation of state-of-the-art satellite data products to
34 disentangle the relevance of near-surface vs. terrestrial water storage for vegetation functioning can inform the representation
35 of vegetation-water interactions in land surface models to support more accurate climate change projections.

1. Introduction

37 The regulation of water, energy, and biogeochemical cycling between land and atmosphere is primarily dependent on
38 vegetation. In addition, global vegetation provides essential ecosystem services such as food production and uptake of some
39 of the anthropogenic carbon dioxide emissions (Keenan and Williams, 2018). Vegetation growth depends on nutrient, water
40 and energy availability. As a result, on a global scale, there are regions with energy or water limited vegetation functioning
41 (Orth, 2021). In energy-limited regions, the functioning of vegetation is controlled by radiation and temperature, as they often
42 lack sunny and warm conditions but have ample soil moisture. In contrast, soil moisture becomes critical for vegetation growth
43 in water-limited regions. Plant photosynthesis involves opening the stomata for the uptake of CO_2, while at the same time
44 water is lost through transpiration. However, in water-limited conditions, plants can reduce the stomatal opening to avoid water
45 loss, leading to a decrease in photosynthesis. Hence, variations in soil moisture are likely to affect vegetation functioning in
46 water-limited conditions. Moreover, climate change has led to an expanded water limitation on vegetation (Denissen et al.,
47 2022) and increased vegetation sensitivity to soil moisture (Li et al., 2022). For these reasons, it is essential to better understand
48 the dependence of vegetation functioning on soil moisture to comprehend their coping mechanisms during drought to predict
49 the future of global water, energy, and carbon cycles.

51 Plants extract water from varying soil depths based on the positioning of their roots and the availability of soil moisture and
52 nutrients. In general, the plant water uptake depth further differs spatially across different climate regimes and vegetation
53 types, and temporally between seasons. Vegetation in arid regions is more susceptible to fluctuations in near-surface soil
54 moisture compared to vegetation in humid regions (Xie et al., 2019). Grasses, which generally have shorter roots than trees
55 and shrubs, are more reliant on near-surface moisture than deeper moisture (Schenk and Jackson, 2002). Further, root water
56 uptake profiles vary within individual plant types according to above-ground biomass and age, with larger and older trees
57 having deeper roots capable of extracting water from deeper soil layers (Schenk and Jackson, 2002; Tao et al., 2021).
58 Additionally, within similar climate regimes, plant water uptake varies across topographic positions. Upland and lowland roots
59 tend to be shallower, making vegetation more reliant on near-surface soil moisture, while roots go deeper in steep terrain
60 between these landscapes to access both surface and deep moisture (Fan et al., 2017).

62 Though spatial variations of plant water uptake depths across vegetation types and climate regimes, and temporal shift during
63 dry-months, are widely studied at point scale, inadequate deep soil moisture records pose a major obstacle to study vegetation
64 root water uptake at a global scale. Microwave remote sensing allows to infer near-surface soil moisture dynamics globally.

While microwaves penetrate only the top few centimeters and do not cover the entire soil moisture profile, they represent larger depths of moisture variation, providing valuable insights into at least some of the root zone soil moisture (Feldman et al., 2023). Land surface models provide an alternative source of global soil moisture data across depths, but they are subject to uncertainties arising from meteorological data, inaccurate knowledge of soil and vegetation characteristics, and the representation of complex processes such as photosynthesis, infiltration, and evaporation (Koster et al., 2009; Seneviratne et al., 2010). Hence, some studies have employed reanalysis-based soil moisture estimates, to investigate the relationship between vegetation and soil moisture at the global scale ((Li et al., 2021; Miguez- Macho and Fan, 2021)(Li et al., 2021; Miguez-Macho and Fan, 2021); but those are likely to be impacted by model assumptions affecting soil moisture dynamics, particularly for deeper layers where less observational constraints are available. Thus, studying vegetation interactions with the entire water column, including near-surface and deep soil moisture, at a global scale using exclusively observation-based dataset is imperative to enhance the understanding of relevance of near-surface and deep soil moisture for vegetation functioning.

The Gravity Recovery and Climate Experiment (GRACE) satellite mission, launched in 2002, provides total water storage (TWS) anomalies observations at the global scale. The TWS captures not only soil water but also snow and ice, canopy water, surface water and groundwater. Its depth of representation is therefore difficult to physically quantify, and that is why we studiesstudy TWS anomalies. Nevertheless, they seem to be related to variations of overall water availability (near-surface + deep soil moisture) for vegetation (Yang et al., 2014). The inter-annual carbon dioxide growth rate in the atmosphere, for example, has been found to be well correlated with the total water storage anomalies on a global scale, indicating the relevance of total water column for vegetation functioning (Humphrey et al., 2018). In this study, we assume that TWS anomalies can be used to estimate the variation of overall water availability (near-surface + deep soil moisture) for vegetation under (i) snow-free conditions, and assuming that (ii) water storage variations in lakes or groundwater are negligible at the monthly time scale, (iii) and canopy water storage is much smaller than soil water storage and hence also negligible. (Zheng and Jia, 2020; Stocker et al., 2023). While soil moisture fluctuations represent the largest variation of TWS (Rodell and Famiglietti, 2001), it is essential to note that certain regions exhibit notable short term fluctuations in lake and groundwater due to human management (Strassberg et al., 2007; Cooley et al., 2021).

This study focuses on understanding the relevance of near-surface soil moisture vs. total water storage for vegetation functioning on a global scale using observation-based datasets, thereby inferring vegetation's large-scale water uptake depth from observation-based datasets. For this purpose, we utilise TWS and near-surface soil moisture and correlate them with vegetation functioning, represented by Near-Infrared Reflectance of Vegetation (NIRv). In particular, we analyse (1) what is the relevance of near-surface soil moisture vs. the terrestrial water storage for vegetation functioning?, (2) how does the importance of near-surface soil moisture vs. terrestrial water storage change during dry months? and (3) how do climatic, vegetation, and topographic characteristics explain the variability in the relevance of near-surface vs. terrestrial water storage for vegetation functioning?

99   **2. Data and Methodology**

100   **Table 1: Table summarising all the datasets.**

Datasets	Variables	Source	Spatial Resolution	Temporal Resolution	Temporal Coverage	References
Vegetation Functioning	Near Infrared Reflectance of Vegetation (NIRv)	MODIS/MOD13C1 v061	0.05 degree	16 daily	2000 - present	(Badgley et al., 2017)
	Solar Induced Chlorophyll Fluorescence (SIF)	GOME-2	0.5 degree	16 daily	2007 - 2018	(Köhler et al., 2015)
Soil Water Storage	Near-surface soil moisture (SSM)	ESA-CCI v04.4	0.25 degree	Daily	1978 - 2022	(Dorigo et al., 2017)
	Total Water Storage (TWS) Anomalies	GRACE	0.5 degree	Monthly	2002 - present	(Landerer and Swenson, 2012)
Meteorological	Air Temperature (T_a)	ERA-5	0.25degree	Hourly	1940 - present	(Hersbach et al., 2020)
	Precipitation (P)					
	Net Radiation (R_n)					
	Dew point Temperature (T_d)					
Climatological	Aridity Index	Global Aridity Index and Potential Evapotransp	30 arc seconds	Static	1970-2000	(Zomer et al., 2022)

		iration Database - Version 3				
Vegetation and Land cover class	Tree cover fraction	VFC5KYR	0.05 degree		1982 - 2016	(Hansen, Matthew and Song, Xiao-Peng, 2018)
	Land cover data	ESA-CCI	300 m	Yearly	1992 - 2018	ESA. Land Cover CCI Product User Guide Version 2. Tech. Rep. (2017)
Topographical data	Elevation	Earthenv	1 km	Static		(Amatulli et al., 2018)
	Slope					
Soil data	Fraction of sand	FAO	0.05 degree	Static		(Reynolds et al., 2000)(Reynolds et al., 2000)
	Fraction of clay					
Irrigation	Percentage of Irrigated area	HID	5 arcmin	Yearly	1990 - 2005	(Siebert et al., 2015)

101

102 **2.1 Data**

103 **2.1.1 Vegetation Functioning:**

104 In our study, vegetation functioning is characterised by satellite measurements of Near-Infrared Reflectance of vegetation
105 (NIRv) and Solar Induced Fluorescence (SIF) (**Table 1**). NIRv is the product of near-infrared reflectance and the normalised
106 difference vegetation index (NDVI) and represents the vegetation structure and vegetation greenness (Badgley et al., 2017).
107 The NIRv data is available at a high spatial resolution of 0.05°, and the original 16-day data was aggregated to the monthly
108 NIRv data. SIF is directly related to the photosynthetic activity of plants because the excess energy from sunlight, that triggers
109 the light reaction during photosynthesis, is dissipated by leaf as chlorophyll fluorescence (Mohammed et al., 2019). SIF data
110 is derived from the Global Ozone Monitoring Experiment (GOME-2), because GOME-2 provides relatively reliable data over

111 a long period (2007-2018). The 0.5° spatial and 16-day temporal resolution SIF data is processed into monthly data as described
112 by (Köhler et al., 2015).

113

114 The high spatial resolution of NIRv allows for a detailed study of the correlation of vegetation functioning with soil water
115 availability. Therefore, we performed the main analyses using NIRv data. However, SIF is more sensitive to drought stress
116 than NIRv (Qiu et al., 2022). Therefore, we perform additional analyses with SIF to show that the relationships hold for a
117 different and more direct indicator of vegetation functioning.

118 **2.1.2 Soil Water Storage**

119 This study includes two different measures of soil water availability. The near-surface soil moisture (SSM) provides an
120 estimate of water availability in the top layer of the soil, while the Terrestrial Water Storage (TWS) Anomaly provides an
121 estimate of the overall water column of the soil. The SSM data is derived from the European Space Agency (ESA) Climate
122 Change Initiative Program (CCI), which combines active and passive satellite microwave measurements to provide reliable
123 estimates of SSM (Dorigo et al., 2017). The ESA CCI soil moisture data, at a daily temporal resolution, was aggregated to
124 monthly temporal resolution. The TWS Anomaly data is derived from the GRACE mission, which measures changes in the
125 Earth's gravity field (Landerer and Swenson, 2012). Here, we use the JPL-Mascons product of TWS Anomalies which is
126 available at a 0.5° spatial and monthly temporal resolution (Watkins et al., 2015).

127 **2.1.3 Meteorological Data**

128 Employed climate variables include monthly air temperature(T_a), 2m dew point temperature (T_d), precipitation (P), and net
129 radiation (R_n) from the ERA5 reanalysis products at a 0.25° spatial resolution. The vapor pressure deficit (vpd) is calculated
130 from T_a and T_d. Further, the aridity index is calculated from the ratio between the long-term mean R_n (mm y^{-1}) (1 MJ/sq.m/day
131 = 0.408 mm/day) and P (mm y^{-1}) for each grid cell (Budyko, 1974). We opted for this formulation as it offers a direct estimation
132 of aridity and water (energy) constraints on vegetation. This eliminates the necessity to navigate through various formulations
133 utilized for calculating potential evapotranspiration. However, we conducted additional validations of our results using the
134 Global Aridity Index dataset (Zomer et al., 2022) based upon the FAO Penman-Monteith Reference Evapotranspiration
135 equation. The use of the Global Aridity Index did not change the results of our study (**Section 3.4**). In addition, the mean and
136 standard deviation of the climate variables are calculated and incorporated in the attribution analysis (**Section 2.2.3**).

137 **2.1.4 Vegetation, soil, and topography data**

138 To evaluate the resulting correlation of vegetation functioning and water storages with respect to vegetation characteristics,
139 we employ the tree cover fraction data from the AVHRR vegetation continuous fields products (VCF5KYR,
140 https://lpdaac.usgs.gov/products/vcf5kyrv001/) (Hansen, Matthew and Song, Xiao-Peng, 2018). For this purpose, the mean
141 of tree cover fraction for the years between 2007 and 2016 is calculated.

142 Topographical variables such as elevation and slope are incorporated along with other meteorological variables to determine

143 the relative contribution of different variables to the correlation between vegetation functioning and water storage. Topographic

144 data at a 5 km resolution were downloaded from the EarthEnv. These data are calculated based on the 250 m GMTED dataset

145 and compared against the 90 m SRTM 4.1 dev dataset. The data were resampled to a coarser resolution of 5 km using various

146 aggregation techniques, details of which are in (Amatulli et al., 2018). Furthermore, for each grid cell, the fraction of sand and

147 clay in soil (Reynolds et al., 2000)(Reynolds et al., 2000) along with the percentage of irrigated area (Siebert et al., 2015) were

148 considered in attribution analysis.

149

150 **2.2 Methodology**

151 **2.2.1 Data pre-processing**

152 A flowchart of the data pre-processing and analyses is presented in **Figure S1**. The time period of analysis is from 2007 to

153 2018 constrained by the concurrent availability of all involved datasets. All the analyses were performed in monthly temporal

154 resolution and at 0.05° spatial resolution (for NIRv) and 0.5° spatial resolution (for SIF). The SSM and TWS data were initially

155 available at 0.25° and 0.5° resolution, but were disaggregated or aggregated to 0.05° or 0.5° degrees, depending on the spatial

156 resolution of the analysis performed, based on the assumption that the soil water storage anomalies are representative over

157 larger areas. Also, the meteorological data and vegetation, soil, and topographic data were resampled into the same resolution.

158 After aggregating all the datasets to 0.05° resolution, the monthly anomalies were calculated by subtracting the long term mean

159 monthly cycle and by removing linear trends. A SIF threshold was applied in each grid cell to filter out non-growing season

160 data. For this purpose, we filtered out all the months from 2007-2018 when the mean-monthly SIF value was below the

161 threshold of 0.2 mW/m2/sr/nm. We apply an additional temperature threshold ($T_a > 5°C$) to remove the months with frozen

162 soil and snow cover, similar to (Li et al., 2021)(Li et al., 2021). Last, all months with missing soil water storage or vegetation

163 functioning records were excluded.

164 **2.2.2 Calculate the relevance of near-surface (SSM) soil moisture and terrestrial water storage (TWS) for vegetation**
165 **functioning**

166 We calculated the Spearman correlation between vegetation functioning (NIRv) and soil water storages (SSM and TWS) for

167 each grid cell during growing season months when observations for at least 40 months were available. To derive partial

168 correlation estimates between NIRv and the water storages, we employed a bootstrapping approach (resampling with

169 replacement from the original data) within each grid cell, with 1000 repetitions to compute bootstrap means and confidence

170 intervals. The cutoff of 40 months was implemented to guarantee a substantial number of observations for growing-season

171 months in each grid cell. This consideration assumes that the minimum number of growing-season months varies from 3 to 4

172 months per year globally. In addition to soil moisture, also air temperature (T_a) and net radiation (R_n) affect the vegetation

173 functioning. Moreover, SSM (soil moisture) and TWS (total water storage) demonstrate a notable correlation, as illustrated in

174 **Figure S2**, signifying the presence of mutual information. To exclusively examine the individual impacts of each water storage
175 variable on vegetation functioning and disentangle mutual information from other water variables, we accounted for
176 confounding effects. This entailed computing the partial correlation between NIRv and water storages (SSM or TWS), while
177 controlling for Ta, Rn, and the other water storage variable (TWS or SSM). Since we focus on understanding the role of soil
178 moisture on vegetation functioning, which is primarily critical in water-limited conditions, we removed the grids cells with
179 negative partial correlations from our analysis. Such negative partial correlations may hint at vegetation's converse effect on
180 soil moisture (when increasing vegetation activity depletes the soil moisture) and a negative correlation could occur in the grid
181 cells where water limits vegetation productivity through oxygen limitation (Ohta et al., 2014). Also, note that predominant
182 energy limitation of the vegetation prevents the evaluation of the relevance of soil moisture vs. terrestrial water storage as
183 partial correlations will become insignificant when temperature or radiation are mainly controlling vegetation functioning.
184
185 It is important to note that we chose not to apply a significance criterion in analyzing the partial correlation between NIRv and
186 water storages. When controlling for both water storage (TWS or SSM) and energy variables (Ta and Rn) in the partial
187 correlation (NIRv~SSM or TWS), a limited number of grid cells demonstrate significant correlation globally, given the high
188 correlation between SSM and TWS (**Figure S2**). This poses challenges for drawing global inferences on vegetation water
189 uptake. However, our overarching goal is to discern variations in the partial correlation of NIRv with water storages across
190 differing climate-vegetation gradients and how it changes from the growing season to dry months, rather than confirming
191 specific statistical thresholds. For this, we want to maintain a sufficient amount of grid cells necessary for making global
192 inferences. However, to ensure that our results are not affected by the significance criterion, we conducted additional analyses
193 considering only grid cells with a significant partial correlation (though a very small number compared to the total grid cells
194 available for each AI-TC class globally), as described in **section 3.4**.
195
196 The impact of all pre-processing steps on the number of grid cells included in this study is illustrated in **Figure S3**. Generally,
197 our filtering procedures enable us to concentrate primarily on water-limited regions, as they effectively remove a substantial
198 number of grid cells from the wet regions globally.
199
200 To analyse how the importance of SSM and TWS changes during dry months, we specifically selected the months characterized
201 by the lowest 10% SSM for each grid cell, representing the driest conditions within the growing-season months. The partial
202 correlations between NIRv and water storages, r(NIRv~SSM) and r(NIRv~TWS) were calculated separately for dry months.
203 To focus on vegetation response to similar extent of dryness spatially, only grid cells with greater than 100 monthly
204 observations were considered for the dry months analysis. In addition, only the grid cells which had positive partial correlation
205 in growing season months were included for the dry months analysis.
206

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

207 After computing the partial correlations, we grouped the grid cells by aridity and tree cover classes, which allowed us to
208 analyse the evolution of correlations and the difference between the partial correlation across aridity-tree cover classes.
209 Afterwards, we again employed a full bootstrapping methodology with 1000 repetitions to compute the bootstrap means and
210 confidence interval using a full bootstrapping methodology (resampling with replacement from the original data) for each
211 aridity-tree cover class with sufficient number of observations for both growing season and dry months. .

212

213 Moreover, to test the robustness of the results, we did additional partial correlation analyses, for which we correlated the SIF
214 (instead of NIRv) with SSM and TWS. The analyses with SIF were performed at a spatial resolution of 0.5°, at which SIF data
215 was available.

216

2.2.3 Attribution Analysis

218 We used a random forest model to understand the spatial variability in the relevance of SSM versus TWS for NIRv. Random
219 forest is a nonparametric based regression algorithm which does not require any statistical assumptions on the predictor and
220 target variables which makes it particularly useful for detecting the nonlinear relationship (Breiman, 2001). Given potential
221 nonlinear impacts of various factors (climate, soil types, vegetation) on the relationship between moisture storages and
222 vegetation functioning, this study employed the random forest method to assess the relative contributions of these variables.

223

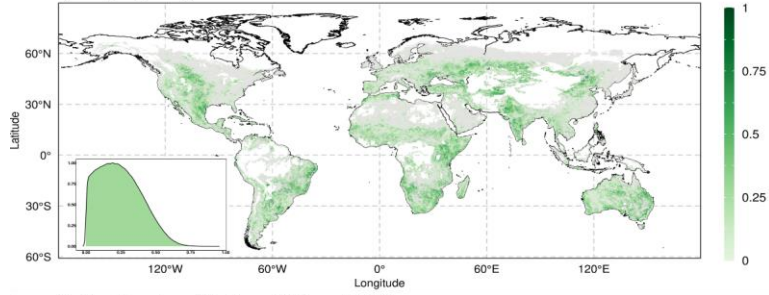
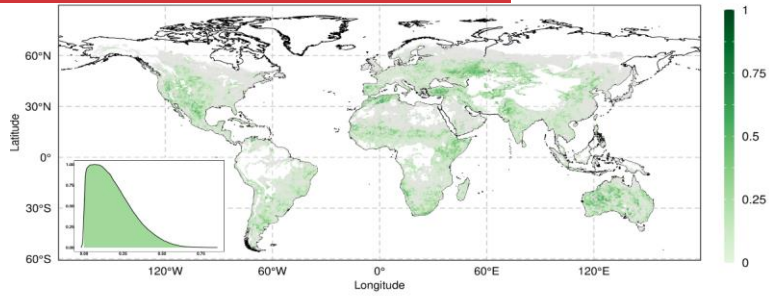
224 In our study, 15 predictors were included in the random forest model based on their potential physical relevance to the target
225 variable, which is the difference in correlation between SSM and TWS with NIRv in growing season months. These predictors
226 included mean and standard deviation of climate variables (T_a, R_n, P and vpd), aridity index, topographical variables (elevation
227 and slope), vegetation variable (tree cover), soil-related variables (fraction of clay and sand), and percentage of irrigated areas
228 for each grid cell. We calculated the mean and standard deviation of the climate variables only during the growing-season
229 months, as determined for the subsequent partial correlation analysis... Furthermore, only the grid cells exhibiting positive
230 partial correlation between NIRv and SSM as well as NIRv and TWS during growing season-months were included in the
231 random forest analysis. For training a random forest model, we used the "xgboost" package in R (Chen and Guestrin, 2016).

232

233 We further incorporate SHAP (SHapley Additive exPlanations) values for interpreting the predictions of the random forest
234 model (Lundberg et al., 2020). The SHAP value for a feature is the average difference in prediction of the model when that
235 feature is included compared to when it is excluded, over all possible combinations of features. By calculating SHAP values
236 for each feature in the model, we identified which features were most important in explaining the spatial variability in the
237 relevance of SSM versus TWS. For calculating the SHAP values, we employed "SHAPforxgboost" package in R.

238 **3. Results and Discussion**

239 **3.1 Coupling of vegetation functioning with surface soil moisture and total water storage in the growing season**

a) Coupling strength between NIRv and SSM

b) Coupling strength between NIRv and TWS

c) Difference between (a) and (b)

240

241

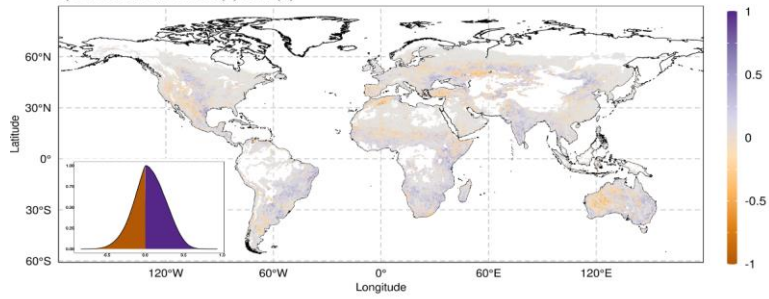
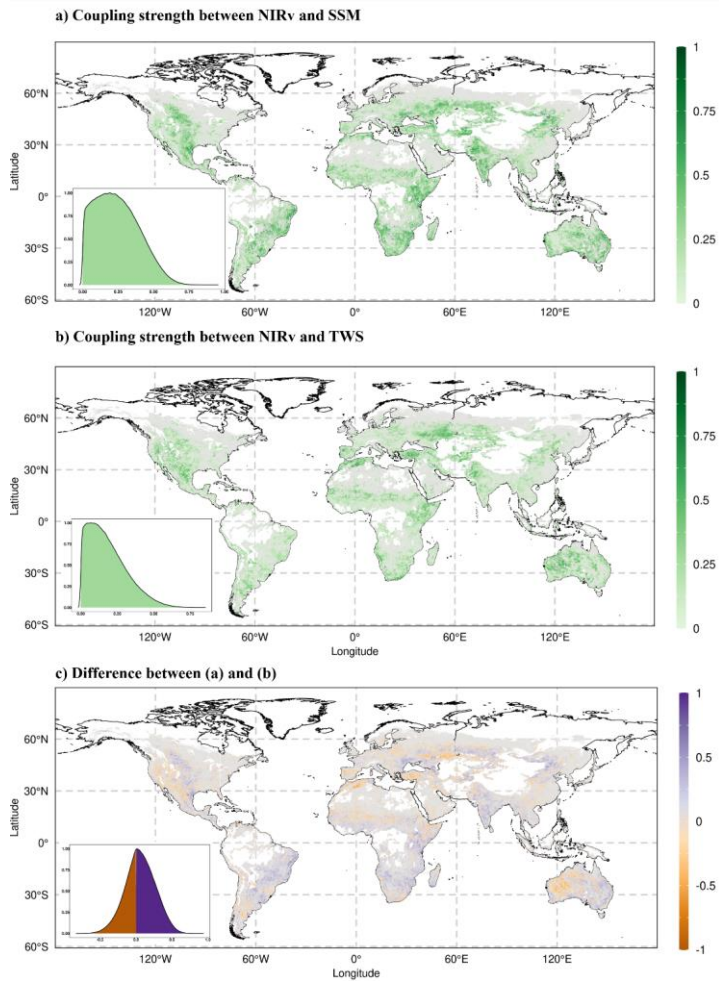



Figure 1: Coupling strength between vegetation functioning (NIRv) and (a) near-surface soil moisture (SSM), and (b) total water storage (TWS) during the growing season months. The color bar denotes the mean partial correlation for each grid cells, computed from the partial correlations across individual bootstrapping samples. Monthly anomalies of all variables are used to calculate the partial correlation. (c) Difference between (a) and (b). The purple colour in (c) indicates the greater partial correlation of NIRv with SSM compared to the partial correlation of NIRv with TWS while orange colour indicates the opposite. Grid cells with positive relationships for both correlations (a) and (b) are shown in (c) with blueish and orange colours. Light grey colour indicates negative partial correlations between NIRv and water storage. The absence of color within the land boundary signifies inadequate

251 The partial correlation of NIRv with near-surface soil moisture varies globally during growing-season months (**Figure 1a**).

252 NIRv demonstrates stronger correlation with near-surface soil moisture within semi-arid climates, Central North America,

253 South America, regions in South Africa and Australia. The correlation is stronger in Southern Europe and the Mediterranean

254 region compared to central and Northern Europe. The correlation gradient from the hot and dry Mediterranean region to wet

255 and cold Northern Europe corresponds to the gradient of water-limited ecosystems to energy-limited ecosystems obtained in

256 other studies (Denissen et al., 2022; Teuling et al., 2009).

257

258 The global correlation of NIRv with TWS follows a similar pattern as with SSM (**Figure 1b**) in growing-season months. The

259 correlation of NIRv with TWS is higher in drier central northern America and Australia compared to other regions. The

260 similarities in the correlation of NIRv with SSM and TWS are expected because the monthly anomalies of SSM and TWS are

261 highly correlated during growing season months in most of our study area **(Figure S2).**

262

263 The difference between the partial correlation of NIRv with SSM and TWS (**Figure 1c**) indicates that the NIRv correlates

264 stronger with TWS in Western America, Southern Europe, and arid regions of Australia compared to other regions globally

265 during growing-season months. In South America and Southern Africa, however, the NIRv shows a stronger correlation with

266 SSM. ~~Although~~To ensure that the observed patterns of difference of partial correlation between SSM and TWS are not the

267 artifacts arising from the computation of differences based on mean partial correlation, we compared the 95% confidence

268 intervals obtained through bootstrapping. Our results indicate that, for the majority of the considered grid cells, the entire

269 confidence intervals of the correlation (NIRv ~ TWS) fall outside the bounds of the correlation (NIRv ~ SSM) which indicates

270 that the correlations differences are significant, thus enhancing the robustness and confidence in our findings (**Figure S4**).

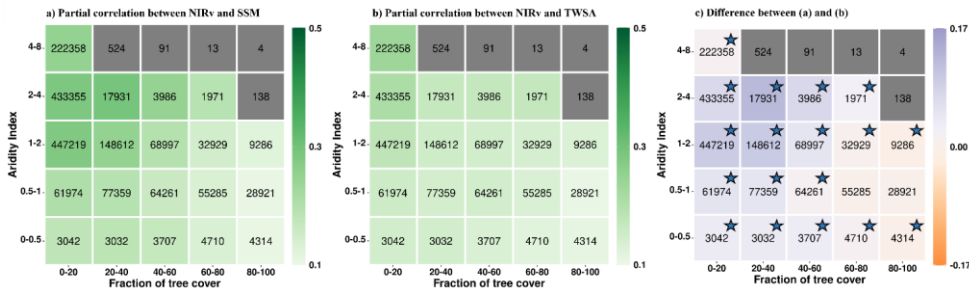
271 Furthermore, even if we control for the effect of soil water storage (SSM or TWS) when computing partial correlation to

272 discern the relative importance for vegetation, it should be noted that the varying noise levels inherent in these datasets might

273 impact our results.

274

275

276

277 **Figure 2: Summarising the coupling strengths of vegetation functioning (NIRv) with (a) near-surface soil moisture (SSM) and (b)**
278 **terrestrial water storage (TWS) in the growing season-months across climate (aridity index) and vegetation regimes (fraction of tree**
279 **cover). (c) shows the difference between (a) and (b). Numbers within the boxes denote the number of grid cells for each aridity-tree**
280 **cover class. Aridity-tree cover classes containing less than 1000 grid cells are shown in grey. The color bar denotes the mean partial**
281 **correlation for each class, computed from bootstrapping. The asterisk in figure (c) signifies that the 95% confidence interval (lower**
282 **and upper) shares the consistent sign (+/-) in the difference of partial correlation. Only grid cells with positive partial correlation**
283 **are considered.**
284

285 Next, we analyse the partial correlation between NIRv and soil water storages across different aridity and tree cover fraction

286 classes during growing season months. For this, we group the grid cells into different aridity and tree cover fraction classes

287 and then do bootstrapping to compute mean partial correlation and the 95 percent confidence intervals for each class with more

288 than 1000 grid cells. We find that the partial correlation of NIRv with SSM (**Figure 2a**) increases with increasing aridity for

289 aridity index (0-4). This can be attributed to the intensification of water stress on vegetation under increasingly arid conditions,

290 resulting in a stronger correlation between NIRv and SSM. However, for a further increase in aridity (4-8), the strength of the

291 correlation of NIRv with SSM declines. This is due to a low soil moisture availability and low temporal variability under

292 extremely arid conditions (**Figure** S4S5). The pattern of increasing correlation along aridity index is also observed in the

293 partial correlation between NIRv and TWS (**Figure 2b**).

294

295 Furthermore, the correlation of NIRv with SSM decreases for higher tree cover fractions (**Figure 2a**). However, such a gradient

296 along tree cover fraction is less pronounced in the partial correlation of the NIRv with TWS (**Figure 2b**). This overall depicts

297 that the coupling of vegetation functioning with SSM is generally higher for non-forested areas compared to forested areas

298 while this gradient is less pronounced in the case of TWS.

299

300 Though the difference in inherent noise levels associated with SSM and TWS impacts partial correlation analysis, we can

301 compare the evolution of the gradient along tree cover or aridity index and assert how the relevance of SSM and TWS changes

302 with varying tree cover or aridity index, assuming that the noise levels are similar across varying AI-TC classes. Taking this

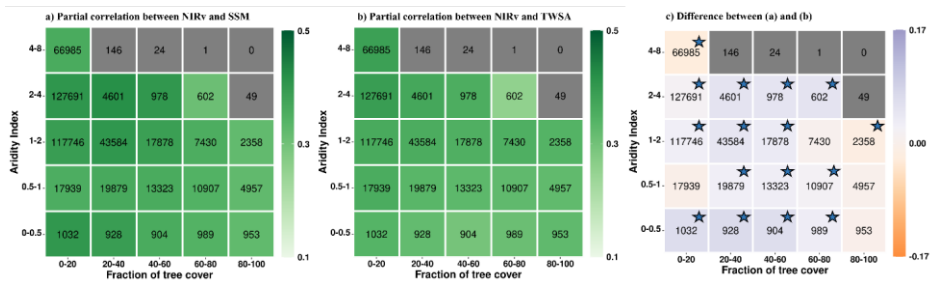
303 into account, we find that NIRv correlates more strongly with near-surface soil moisture compared to terrestrial water storage

304 in semi-arid regions with low tree cover (**Figure 2c**), suggesting that the vegetation preferentially takes up water from SSM

305 whenever available to meet its transpiration demand. This might be due to lower energy expenditure on root water uptake,
306 abundant nutrients and reduced chance of root water logging in the near-surface soil moisture (Feldman et al., 2023; Schenk
307 and Jackson, 2002; Tao et al., 2021). Conversely, the correlation between the NIRv and TWS in arid areas (AI 4-8) and regions
308 with a high fraction of tree cover is equivalent to or greater than that of SSM, suggesting that trees can utilise their extensive
309 root systems to access deeper soil moisture, as also observed in arid vegetation. This is consistent with previous studies
310 reporting that the vegetation dependence on sub-surface soil moisture is higher in arid and seasonal-arid climates (Miguez-
311 Macho and Fan, 2021). However, in certain regions with higher tree cover in humid areas, specifically with AI 0.5-1, such
312 conclusions cannot be confidently drawn statistically. The reason is that the confidence intervals for the difference in partial
313 correlation of NIRv with SSM and TWS fluctuate between positive (indicating greater relevance of SSM) and negative
314 (indicating greater relevance of TWS) values (**Figure 2c**).
315

316 Note that while our analysis focuses on regions with water-controlled vegetation as denoted by positive correlations between
317 NIRv and the considered soil water storages, some of these grid cells are located in comparatively wet climate regimes with
318 aridity index values between 0 and 1 (**Figure 2**). This highlights the relevance of non-climatic factors such as soil and
319 vegetation types or topography in determining vegetation-water relationships in addition to the climate regime. Next to this,
320 in **Figure 2c** it seems that the relevance of terrestrial water storage is comparatively higher in wet climate (aridity 0.5-1) than
321 in transitional climate regimes (aridity 1-2) as shown with the smaller correlation differences. This, however, is probably not
322 the case and might simply be a reflection of reduced variability in surface soil moisture (**Figure S4S5**).
323

324 **3.2 Coupling of vegetation functioning with surface soil moisture and total water storage in dry months**
325 The correlation between NIRv and soil water storage increases during dry months (**Figure 3a,b**) compared to growing season
326 months (**Figure 2a,b**). This increase is consistent for both SSM and TWS and across all tree cover fractions and aridity classes.
327 This is because the water limitation on vegetation increases in dry months and so does the vegetation's sensitivity to the
328 moisture. During the dry months, the correlation with near-surface soil moisture tends to rise, but the correlation with terrestrial
329 water storage increases even more significantly (**Figure 3c**). This indicates the relevance of deeper water resources during
330 periods of scarce rainfall. The partial correlation maps (**Figure S5S6**) also reveal that NIRv's correlation with TWS increases
331 more than its correlation with SSM for most grid cells.
332

333

334 **Figure 3: Summarising the coupling strengths of vegetation functioning (NIRv) with (a) near-surface soil moisture (SSM) and (b)**
335 **terrestrial water storage (TWS) in the 10% driest months in each grid-cell across climate (aridity index) and vegetation regimes**
336 **(fraction of tree cover). (c) shows the difference between (a) and (b). Numbers within the boxes denote the number of grid cells for**
337 **each aridity-tree cover class. Aridity-tree cover classes containing less than 1000 grid cells are shown in grey. The color bar denotes**
338 **the mean partial correlation for each class, computed from bootstrapping. The asterisk in figure (c) signifies that the 95% confidence**
339 **interval (lower and upper) shares the consistent sign (+/-) in the difference of partial correlation. Only grid cells with positive partial**
340 **correlation are considered.**

341 During dry months, the number of analysed grid cells (**Figure 3**) is lower compared to all growing season months (**Figure 2**).

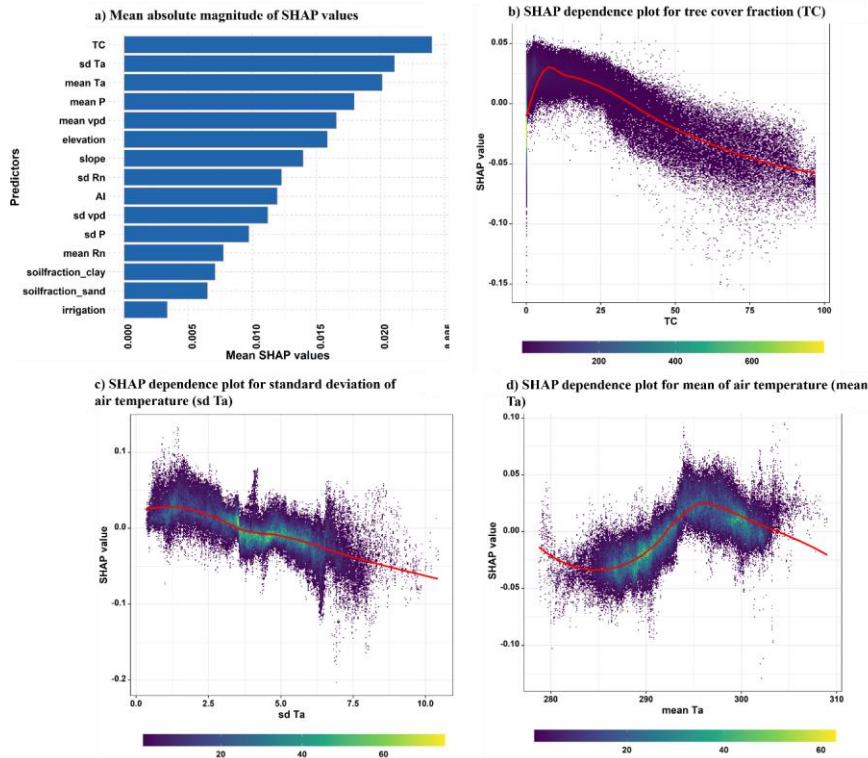
342 We performed a reanalysis of the correlation patterns within aridity-tree cover classes by selecting only those grid cells that

343 displayed positive partial correlation between NIRv and soil water storages during both the dry months and the growing season

344 months. The results demonstrate that the previously observed patterns remain valid, thereby eliminating the impact of the

345 differing numbers of grid cells analysed. (**Figure S6 S7).**

346

347 **3.3 Climate, vegetation, and topographic controls on the relevance of surface soil moisture vs. total water storage on**
348 **vegetation**

349
350 **Figure 4: (a) Global feature importance based on the mean absolute magnitude of the SHAP values. The higher the mean SHAP**
351 **values, the greater the predictor's relevance. (b-d) Evaluation of SHAP values (=contributions to the correlation difference**
352 **illustrated in Figure 1c) against predictor values for the 3 most relevant predictors tree cover fraction (TC), variability of**
353 **temperature (sd T$_a$) and mean temperature (mean T$_a$) during the growing season months. The colour indicates the density of data**
354 **points. For plotting (b), (c) and (d), only 10 percent random samples of the whole dataset are utilised.**

355 We use a random forest model to understand the spatial variability in the relevance of SSM versus TWS for NIRv. The model

356 was trained with 15 climatic, vegetation, and topographic predictors against the target variable which is the difference of the

357 partial correlations of NIRv with SSM and TWS during growing season-months (R^2 = 0.59, see **methods section 2.2.3**). The

358 mean absolute SHAP value plot shows that the tree cover and the climate variables (mean and standard deviation of T$_a$) are

359 most important variables for explaining the spatial variability in the relative importance of SSM vs. TWS for vegetation

functioning (**Figure 4a**). This overall highlights that the relative importance of SSM vs. TWS for the vegetation is broadly controlled by vegetation type, reflecting the local adaptation of ecosystem and climate, influencing water availability~~.~~ (Stocker et al., 2023).

Tree cover fraction is an important factor in determining the relevance of SSM and TWS for vegetation functioning (**Figure 4c4b**). Regions with a high tree cover are more dependent on TWS, as trees generally have deeper root systems that allow them to adjust water uptake between different depths (Tao et al., 2021). Grasslands on the other hand have shallow roots that are more susceptible to surface soil moisture variations (Yang et al., 2014).

Similarly, the relative importance of SSM and TWS varies non-linearly with the mean growing season temperature (**Figure 4b4d**). TWS tends to be more crucial for vegetation functioning in areas with low (approximately below 20°C) or high (above 27°C) growing season temperatures, while SSM has greater importance in regions with moderate growing season air temperatures. One possible explanation for this trend is that high temperatures induce a strong atmospheric water demand that dries near-surface soil layers, which leads vegetation to increase water extraction from deep soils. This observation is further underscored by the analogous pattern observed in the SHAP dependence plot for vpd, which accentuates atmospheric water demand (**Figure S8S8b**). In contrast, SSM is more available during growing season months in regions characterised by moderate temperatures. We hypothesize that the regions that experience relatively cold growing season temperatures exhibit stronger temperature and weather variability that may contribute to longer dry periods and, thus, emphasises the importance of deeper soil moisture for vegetation functioning. However, it should be noted that our findings regarding the relevance of TWS at high temperatures must be interpreted with caution due to the exclusion of most tropical forest regions from our analysis (**Figure S7S9**). As a result, most warm regions are dry, and there are only a few hot and wet regions included in our training data.

Not only the mean of the growing season temperature, but also its variability is crucial for explaining the significance of SSM and TWS for vegetation functioning (**Figure 4d4c**). A higher temporal variability in temperature increases the importance of TWS for vegetation. This is because atmospheric water demand scales with temperature. Hence, higher variability in temperature implies more peaks in related atmospheric water demand which is a stronger incentive for plants to access deeper water storages which are more often available to meet the vegetation´s transpiration demand.

Figure S8 illustrates the effect of the other six important predictors on the model output. Apart from climatological parameters (mean P, mean vpd, variability in R_n, and aridity index), elevation and slope explain part of the variability in the relevance of SSM vs. TWS for NIRv. Although the reasons for increasing relevance of TWS for vegetation functioning at higher elevation remain unclear, it may be due to elevation's strong correlation with other climatic variables such as T_a and P.

394 Several local studies identified other relevant factors that determine root water uptake depth such as forest stand age and tree
395 height, competition, root hydraulic architecture, and tree species (Zhu et al., 2022; Quijano et al., 2012; Stahl et al., 2013,
396 Gessler et al., 2021; Liu et al., 2021). For example, young trees more easily increase their root activity in the shallow or deep
397 soil dependent on soil moisture than mature trees (Zhu et al., 2022; Drake et al., 2011). These variables were not included in
398 our attribution analysis, because they are not available at global scale.

3.4 Robustness Tests

400 In the aforementioned analysis, we included grid cells exhibiting both positive partial correlations, whether significant or non-
401 significant. Upon further examination, we specifically assessed the evolution of partial correlation between NIRv and water
402 storages, considering only grid cells with significant partial correlation ($p < 0.00505$). The observed patterns along the aridity-
403 tree cover gradient remained similar during growing season months. This suggests the robustness of our results to the choice
404 of the statistical significance criterion, albeit with a substantial reduction in the number of globally available grid cells when
405 considering only significant partial correlation (**Figure S9S10**).

407 Furthermore, to ensure that our results are robust to variations in the threshold for Solar-Induced Fluorescence (SIF) used to
408 define growing season months, we conducted additional analyses with a different SIF threshold. Instead of filtering out all
409 months from 2007-2018 when the mean-monthly SIF value was below the threshold of 0.2 mW/m²/sr/nm, we utilized a
410 threshold of 0.5 mW/m²/sr/nm. Elevating the SIF threshold implies the exclusion of additional months characterized by lower
411 vegetation activity for the partial correlation analysis. However, it is essential to note that this threshold does not seem to affect
412 the number of globally available grid cells during growing season months and hence patterns along AI-TC classes are similar.
413 Instead, it specifically influences the selection of dry months and hence the number of grid cells available for the analysis
414 during dry months. Nevertheless, even with the elevated SIF threshold for defining growing season months, the observed
415 patterns along aridity-tree cover (AI-TC) classes remain largely consistent with the results obtained in our main analyses
416 (**Figure S10S11**).

418 Although NIRv can largely reflect vegetation functioning (Badgley et al., 2017), we repeat our analysis with SIF, which is an
419 alternative and independent indicator for vegetation functioning and shows a near-linear relationship with gross primary
420 productivity at the ecosystem level (Guanter et al., 2012)(Guanter et al., 2012). However, SIF is only available at a coarse
421 resolution of 0.5 degree. The partial correlations, r(SIF~SSM) and r(SIF~TWS) largely agree with the pattern of r(NIRv~SSM)
422 and r(NIRv~TWS) across varying aridity index and tree cover classes (**Figure S11 S12**). This suggests that our overall
423 conclusion on the relevance of SSM or TWS for vegetation functioning is robust across different indicators of vegetation
424 productivity.

Additionally, we tested if our results are robust when the aridity index is calculated based on the FAO Penman-Monteith Reference Evapotranspiration equation, for which we applied aridity classification based on UNEP 1997 guidelines. Our results confirm the findings of **Section 3.1and1 and Figure 2** that as aridity increases, the correlation of NIRv with ~~Soil Surface Moisture~~near-surface soil moisture (SSM) and ~~Total Water Storage~~total water storage (TWS) intensifies. Moreover, in hyper arid regions (AI < 0.03) the correlation with TWS surpasses that with SSM (**Figure ~~S12~~S13**). They also confirm that regions with higher tree Cover (TC) fraction correlates more strongly with TWS compared to SSM. Thus, the choice of aridity index formulation does not alter our main conclusions.

When analyzing partial correlations between Total Water Storage (TWS) and vegetation metrics (NIRv or SIF) at finer resolutions (0.05 degrees for NIRv or 0.5 degrees for SIF), it is crucial to acknowledge the potential emergence of significant spatial autocorrelation. This is attributed to the fact that the actual spatial resolution of the satellite signal underlying the TWS data is 2-3 degrees.

4. Summary and Conclusions

In this study we compare the relevance of near-surface soil moisture and of terrestrial water storage for vegetation functioning across the globe. We find that in semi-arid regions and regions with low tree cover, vegetation preferentially utilises the water from shallow soil, which is related to continuous availability of near-surface water availability and lack of deep rooting systems respectively. The stronger correlation of NIRv with SSM than TWS is supported by site-level studies that find a higher root water uptake of surface soil moisture (Brinkmann et al., 2019, Gessler et al., 2021, Deseano Diaz et al., 2023; Kulmatiski and Beard, 2013), also when deeper water is available. Some local studies however find a higher root water uptake from deeper layers (Zhu et al., 2022).

By contrast, in mostly forested regions and in relatively dry climate regimes, the correlation with terrestrial water storage is comparable or higher than with near-surface soil moisture, indicating that trees and vegetation in arid regions use their deep root systems to access deeper soil moisture. Point-scale studies also found a different water uptake depth for trees and grasses ~~in~~ for example in savanna ecosystems (Kulmatiski et al., 2010), and a different water uptake depth for tree species (Kahmen et al., 2022). Liu et al. (2021) showed for example that for a karst forest in Southwest China, evergreen species rely mostly on water sources from the 0-30 cm layer, while deciduous species extracted most water from the 30-70 cm layer.

We also find that vegetation's preferential water uptake depth changes over time. During particularly dry months, the relative importance of terrestrial water storage is higher, highlighting the importance of deep water resources during periods of low soil water availability. This is in line with previous studies showing changes in vegetation's water uptake depth during drought

458 periods at small spatial scales where accessing water in deeper soil layers helps plants to alleviate water stress and maintain
459 transpiration (Migliavacca et al., 2009; Tao et al., 2021).
460
461 Our global results are supported by site-scale studies that find that, during drought, the deeper roots play a more active role in
462 water extraction (Stahl et al., 2013, Volkmann et al., 2016; Tao et al., 2021). In some studies however, the increase of deep
463 water uptake is only relative: the absolute uptake of deep water does not increase, but the uptake of shallow water decreases
464 (Brinkmann et al., 2019, Gessler et al., 2021, Rasmussen et al., 2020; Kühnhammer et al., 2023). This means that the uptake
465 of deeper soil layers cannot compensate for the loss of water uptake from the dry topsoil. Contrary to trees, grasses do not shift
466 their uptake depth (Deseano Diaz et al., 2023), or even extract water from the most shallow soils (Prechsl et al., 2015,
467 Kulmatiski and Beard, 2013).
468
469 Furthermore, we show that the spatial variability of the importance of near-surface soil moisture vs. terrestrial water storage
470 for vegetation functioning is influenced by fraction of tree cover and mean and standard deviation of air temperature. This
471 emphasises the role of climate in determining shallow vs. deep soil water resources, and the role of vegetation in adapting to
472 different soil water availability patterns.
473
474 Vegetation functioning and soil water storages are generally coupled in both directions, i.e. while soil moisture availability
475 affects vegetation functioning (positive coupling), this in turn also affects soil moisture through transpiration (negative
476 coupling). As our study focuses on water-controlled vegetation we only consider positive couplings and filter out grid cells
477 with negative correlations. Future research may consider the relevance of soil moisture across depths for the ~~positive~~negative
478 coupling regions.
479
480 Overall, our analysis illustrates that satellite-based data can be used for belowground analysis at large spatial scales thanks to
481 the fact that satellite retrievals can assess soil water storage dynamics across depths and because vegetation in water-controlled
482 areas can be used as an indicator of soil water dynamics. Such novel ways to improve our understanding of belowground water
483 dynamics is necessary and valuable as respective in-situ observations are scarce and of limited representativeness for larger
484 areas, particularly given the typical spatial heterogeneity of soils and vegetation. Our results can further inform a better
485 representation of belowground processes in global models in order to support more accurate projections of future changes in
486 climate, water resources, and ecosystem services.

487 **Data availability**

488 The monthly SIF data is available from https://www.gfz-potsdam.de/sektion/fernerkundungund-

489 geoinformatik/projekte/global-monitoring-of-vegetation-fluorescence-globfluo/daten.The NIRv was calculated from the red
490 and near-infrared reflectance obtained from the MOD13C1 v006 product (https://lpdaac.usgs.gov/products/mod13c1v061/).
491 The ESA-CCI soil moisture can be accessed through https://esa-soilmoisture-cci.org/ and Terrestrial Water Storage Anomaly
492 data can be accessed through https://podaac.jpl.nasa.gov/dataset/TELLUS_GRACGRFO_MASCON_CRI_GRID_RL06_V2.
493 The ERA5 climate variables are available from https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5 . Tree
494 cover fraction data is available from the AVHRR vegetation continuous fields products
495 https://lpdaac.usgs.gov/products/vcf5kyrv001/, land cover data is available from https://www.esa-landcover-cci.org/, and
496 topographic data is available via https://www.earthenv.org/topography. Similarly, the irrigation fraction data could be accessed
497 from https://mygeohub.org/publications/8 .

Competing Interests

499 The contact author has declared that none of the authors has any competing interests.

Acknowledgements

501 The authors thank Ulrich Weber for help with obtaining and processing the data, Sujan Koirala for valuable scientific and
502 technical support and the Hydrology–Biosphere–Climate Interactions group at the Max Planck Institute for Biogeochemistry
503 for fruitful discussions. Prajwal Khanal, Anne Hoek van Dijke and Rene Orth acknowledge funding by the German Research
504 Foundation (Emmy Noether grant no. 391059971).

505
506
507
508
509
510
511
512
513
514
515
516
517
518

519

References

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., and Jetz, W.: A suite of global, cross-scale topographic variables for environmental and biodiversity modeling, Sci Data, 5, 180040, https://doi.org/10.1038/sdata.2018.40, 2018.

Badgley, G., Field, C. B., and Berry, J. A.: Canopy near-infrared reflectance and terrestrial photosynthesis, Sci. Adv., 3, e1602244, https://doi.org/10.1126/sciadv.1602244, 2017.

Breiman, L.: [No title found], Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Budyko, M. I.: Climate and life, Academic press, 1974.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA, 785–794, https://doi.org/10.1145/2939672.2939785, 2016.

Cooley, S. W., Ryan, J. C., and Smith, L. C.: Human alteration of global surface water storage variability, Nature, 591, 78–81, https://doi.org/10.1038/s41586-021-03262-3, 2021.

Denissen, J. M. C., Teuling, A. J., Pitman, A. J., Koirala, S., Migliavacca, M., Li, W., Reichstein, M., Winkler, A. J., Zhan, C., and Orth, R.: Widespread shift from ecosystem energy to water limitation with climate change, Nature Climate Change, 12, 677–684, https://doi.org/10.1038/s41558-022-01403-8, 2022.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215, 2017.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., and Otero-Casal, C.: Hydrologic regulation of plant rooting depth, Proceedings of the National Academy of Sciences of the United States of America, 114, 10572–10577, https://doi.org/10.1073/pnas.1712381114, 2017.

Feldman, A. F., Short Gianotti, D. J., Dong, J., Akbar, R., Crow, W. T., McColl, K. A., Konings, A. G., Nippert, J. B., Tumber-Dávila, S. J., Holbrook, N. M., Rockwell, F. E., Scott, R. L., Reichle, R. H., Chatterjee, A., Joiner, J., Poulter, B., and Entekhabi, D.: Remotely Sensed Soil Moisture Can Capture Dynamics Relevant to Plant Water Uptake, Water Resources Research, 59, e2022WR033814, https://doi.org/10.1029/2022WR033814, 2023.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., Suto, H., and Grainger, R. G.: Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements, Remote Sensing of Environment, 121, 236–251, https://doi.org/10.1016/j.rse.2012.02.006, 2012.

Hansen, Matthew and Song, Xiao-Peng: Vegetation Continuous Fields (VCF) Yearly Global 0.05 Deg, https://doi.org/10.5067/MEASURES/VCF/VCF5KYR.001, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., and Seneviratne, S. I.: Sensitivity of atmospheric CO2 growth rate to observed changes in terrestrial water storage, Nature, 560, 628–631, https://doi.org/10.1038/s41586-018-0424-4, 2018.

Keenan, T. F. and Williams, C. A.: The Terrestrial Carbon Sink, Annual Review of Environment and Resources, 43, 219–243, https://doi.org/10.1146/annurev-environ-102017-030204, 2018.

Köhler, P., Guanter, L., and Joiner, J.: A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data, Atmos. Meas. Tech., 8, 2589–2608, https://doi.org/10.5194/amt-8-2589-2015, 2015.

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., and Puma, M. J.: On the nature of soil moisture in land surface models, Journal of Climate, 22, 4322–4335, https://doi.org/10.1175/2009JCLI2832.1, 2009.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resources Research, 48, https://doi.org/10.1029/2011WR011453, 2012.

Li, W., Migliavacca, M., Forkel, M., Walther, S., Reichstein, M., and Orth, R.: Revisiting Global Vegetation Controls Using Multi-Layer Soil Moisture, Geophysical Research Letters, 48, https://doi.org/10.1029/2021GL092856, 2021.

Li, W., Migliavacca, M., Forkel, M., Denissen, J. M. C., Reichstein, M., Yang, H., Duveiller, G., Weber, U., and Orth, R.: Widespread increasing vegetation sensitivity to soil moisture, Nature Communications, 13, 3959, https://doi.org/10.1038/s41467-022-31667-9, 2022.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nat Mach Intell, 2, 56–67, https://doi.org/10.1038/s42256-019-0138-9, 2020.

Migliavacca, M., Meroni, M., Manca, G., Matteucci, G., Montagnani, L., Grassi, G., Zenone, T., Teobaldelli, M., Goded, I., Colombo, R., and Seufert, G.: Seasonal and interannual patterns of carbon and water fluxes of a poplar plantation under peculiar eco-climatic conditions, Agricultural and Forest Meteorology, 149, 1460–1476, https://doi.org/10.1016/j.agrformet.2009.04.003, 2009.

Miguez-Macho, G. and Fan, Y.: Spatiotemporal origin of soil water taken up by vegetation, Nature, 598, 624–628, https://doi.org/10.1038/s41586-021-03958-6, 2021.

Mohammed, G. H., Colombo, R., Middleton, E. M., Rascher, U., van der Tol, C., Nedbal, L., Goulas, Y., Pérez-Priego, O., Damm, A., Meroni, M., Joiner, J., Cogliati, S., Verhoef, W., Malenovský, Z., Gastellu-Etchegorry, J.-P., Miller, J. R., Guanter, L., Moreno, J., Moya, I., Berry, J. A., Frankenberg, C., and Zarco-Tejada, P. J.: Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress, Remote Sensing of Environment, 231, 111177, 2019.

Ohta, T., Kotani, A., Iijima, Y., Maximov, T. C., Ito, S., Hanamura, M., Kononov, A. V., and Maximov, A. P.: Effects of waterlogging on water and carbon dioxide fluxes and environmental variables in a Siberian larch forest, 1998–2011, Agricultural and Forest Meteorology, 188, 64–75, https://doi.org/10.1016/j.agrformet.2013.12.012, 2014.

592 Orth, R.: When the Land Surface Shifts Gears, AGU Advances, 2, https://doi.org/10.1029/2021AV000414, 2021.

593 Qiu, R., Li, X., Han, G., Xiao, J., Ma, X., and Gong, W.: Monitoring drought impacts on crop productivity of the U.S. Midwest
594 with solar-induced fluorescence: GOSIF outperforms GOME-2 SIF and MODIS NDVI, EVI, and NIRv, Agricultural and
595 Forest Meteorology, 323, 109038, https://doi.org/10.1016/j.agrformet.2022.109038, 2022.

596 Reynolds, C. A., Jackson, T. J., and Rawls, W. J.: Estimating soil water-holding capacities by linking the Food and Agriculture
597 Organization Soil map of the world with global pedon databases and continuous pedotransfer functions, Water Resour. Res.,
598 36, 3653–3662, https://doi.org/10.1029/2000WR900130, 2000.

599 Rodell, M. and Famiglietti, J. S.: An analysis of terrestrial water storage variations in Illinois with implications for the Gravity
600 Recovery and Climate Experiment (GRACE), Water Resources Research, 37, 1327–1339,
601 https://doi.org/10.1029/2000WR900306, 2001.

602 Schenk, H. J. and Jackson, R. B.: Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants
603 in water-limited ecosystems, Journal of Ecology, 90, 480–494, https://doi.org/10.1046/j.1365-2745.2002.00682.x, 2002.

604 Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating
605 soil moisture-climate interactions in a changing climate: A review, , 99, 125–161,
606 https://doi.org/10.1016/j.earscirev.2010.02.004, 2010.

607 Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., and Scanlon, B.: Historical Irrigation Dataset (HID),
608 https://doi.org/10.13019/M20599, 2015.

609 Stocker, B. D., Tumber-Dávila, S. J., Konings, A. G., Anderson, M. C., Hain, C., and Jackson, R. B.: Global patterns of water
610 storage in the rooting zones of vegetation, Nat. Geosci., 16, 250–256, https://doi.org/10.1038/s41561-023-01125-2, 2023.

611 Strassberg, G., Scanlon, B. R., and Rodell, M.: Comparison of seasonal terrestrial water storage variations from GRACE with
612 groundwater-level measurements from the High Plains Aquifer (USA), Geophysical Research Letters, 34, 2007GL030139,
613 https://doi.org/10.1029/2007GL030139, 2007.

614 Tao, Z., Neil, E., and Si, B.: Determining deep root water uptake patterns with tree age in the Chinese loess area, Agricultural
615 Water Management, 249, 106810, 2021.

616 Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., Buchmann, N., Ammann, C., Montagnani, L.,
617 Richardson, A. D., Wohlfahrt, G., and Seneviratne, S. I.: A regional perspective on trends in continental evaporation,
618 Geophysical Research Letters, 36, https://doi.org/10.1029/2008GL036584, 2009.

619 Watkins, M. M., Wiese, D. N., Yuan, D., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time
620 variable mass distribution with GRACE using spherical cap mascons, JGR Solid Earth, 120, 2648–2671,
621 https://doi.org/10.1002/2014JB011547, 2015.

622 Xie, X., He, B., Guo, L., Miao, C., and Zhang, Y.: Detecting hotspots of interactions between vegetation greenness and
623 terrestrial water storage using satellite observations, Remote Sensing of Environment, 231, 111259, 2019.

624 Yang, Y., Long, D., Guan, H., Scanlon, B. R., Simmons, C. T., Jiang, L., and Xu, X.: GRACE satellite observed hydrological
625 controls on interannual and seasonal variability in surface greenness over mainland Australia, Journal of Geophysical
626 Research: Biogeosciences, 119, 2245–2260, https://doi.org/10.1002/2014JG002670, 2014.

627 Zheng, C. and Jia, L.: Global canopy rainfall interception loss derived from satellite earth observations, Ecohydrology, 13,
628 e2186, https://doi.org/10.1002/eco.2186, 2020.

629 Zomer, R. J., Xu, J., and Trabucco, A.: Version 3 of the Global Aridity Index and Potential Evapotranspiration Database, Sci
630 Data, 9, 409, https://doi.org/10.1038/s41597-022-01493-1, 2022.

631
632
633
634