

## 1                    Revisiting global vegetation controls using multi-layer soil moisture

2    Wantong Li<sup>1\*</sup>, Mirco Migliavacca<sup>1</sup>, Matthias Forkel<sup>2</sup>, Sophia Walther<sup>1</sup>, Markus  
3    Reichstein<sup>1</sup> and René Orth<sup>1</sup>

4    <sup>1</sup>Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, D-  
5    07745 Jena, Germany.

6    <sup>2</sup>Technische Universität Dresden, Institute of Photogrammetry and Remote Sensing,  
7    Helmholtzstr. 10, D-01069 Dresden, Germany.

8    \*Corresponding author: Wantong Li (wantong@bgc-jena.mpg.de)

### 9    **Key Points:**

- 10    •        Vertically resolved soil moisture improves the understanding of large-scale vegetation  
11    productivity.
- 12    •        Extended water-related control on vegetation productivity emerges when considering  
13    multi-layer soil moisture versus total soil moisture.
- 14    •        Sub-surface soil moisture is particularly important for vegetation productivity in semi-  
15    arid regions.

16

## 17 **Abstract:**

18 The productivity of terrestrial vegetation is determined by a multitude of drivers between the  
19 land surface and atmosphere. Water availability is critical for vegetation productivity, but the  
20 vertical dimension of soil moisture has been largely overlooked. Here, we analyze dominant  
21 controls of global vegetation productivity represented by sun-induced fluorescence and spectral  
22 vegetation indices at the half-monthly time scale. We apply random forests to predict anomalies  
23 of vegetation productivity from a comprehensive set of hydro-meteorological variables  
24 including multi-layer soil moisture and quantify the variable importance. Dominant hydro-  
25 meteorological controls generally vary with latitudes: temperature in higher latitudes, solar  
26 radiation in lower latitudes, and soil moisture from sub-surface layers in between. We find that  
27 including vertically resolved soil moisture allows a better understanding of vegetation  
28 productivity and reveals a broader water-related control. This is found especially for semiarid  
29 regions, illustrating the global relevance of deep(er) rooting systems as an adaptation to water  
30 limitation.

## 31 **1. Introduction**

32 Terrestrial vegetation is a key component coupling the global water and carbon cycles  
33 between the atmosphere and the land surface. Its productivity is determined by a multitude of  
34 hydro-meteorological variables (Monteith and Unsworth 1990; Nemani et al., 2003; Piao et al.,  
35 2020). While the underlying relationships are complex in time and space (Pearson et al., 2013;  
36 Cox et al., 2013; Garonna et al., 2018), the hydro-meteorological controls of anomalies in  
37 vegetation productivity are still not fully understood at a global scale. This knowledge gap  
38 contributes to uncertainties in assessing the sensitivity and resilience of ecosystems to different  
39 climate drivers (Seddon et al., 2016; Sakschewski et al., 2016), and in future climate projections  
40 (Feng et al., 2014; Novick et al., 2016; Duveiller et al., 2018).

41 Previous studies investigated dominant hydro-meteorological controls of vegetation  
42 productivity at a global scale and across different ecosystems (Nemani et al., 2003; Beer et al.  
43 2010; Jung et al. 2011; Seddon et al., 2016; Madani et al., 2017; Jung et al., 2017; Walther et al.,  
44 2019; Li & Xiao, 2020). While these studies and recent gross primary production (GPP)  
45 estimates agree that vegetation in (semi-)arid area is significantly impacted by soil moisture (SM)  
46 (Stocker et al., 2018; Stocker et al., 2020), a corresponding global analysis including the impact

47 of SM from multiple depths is lacking. Several studies have already highlighted the local  
48 relevance of multi-layer SM to ecosystems: root water uptake from deeper soil layers can help  
49 mitigate water stress and maintain plant transpiration (Schulze et al., 1996; Migliavacca et al.,  
50 2009); A et al., 2019 demonstrated varying relative importance of surface SM versus deeper SM  
51 depending on land cover types; and Schlaepfer et al., 2017 simulated an increased dryness of  
52 sub-surface SM compared to surface SM which largely impacted vegetation dynamics in  
53 temperate drylands. This way, distinguishing shallow and deep SM is expected to allow for a  
54 more accurate identification of global vegetation controls as the accessibility and availability of  
55 water for plants varies in space and time. For this purpose, the state-of-the-art ERA5 reanalysis  
56 provides SM estimates from multiple layers (Hersbach et al. 2019; Jing et al., 2018), and has  
57 been successfully applied in hydro-meteorological studies (Jing et al., 2018; Tarek et al., 2020;  
58 Li et al., 2020).

59 When considering multiple hydro-meteorological variables, the identification of global  
60 vegetation controls is challenged by potential high collinearity (Dormann et al., 2013) between  
61 some of the variables. Most previous studies did not consider more than three variables, thereby  
62 somewhat circumventing this problem while ignoring potentially important variables (Seddon et  
63 al., 2016; Garonna et al., 2018; Claessen et al., 2019; Li & Xiao, 2020). Machine learning  
64 methods such as random forests have no assumptions on the input data characteristics, and are  
65 designed to process large amounts of diverse input data (Breiman 2001; Forkel et al. 2019; Jiao  
66 et al. 2019). Though they are also challenged by the collinearity in the input data, they are better  
67 placed to deal with this than traditional statistical methods. Further, a pre-processing of the data  
68 can mitigate collinearity by removing potential confounding signals such as long-term trends or  
69 the seasonality (Jung et al., 2017).

70 Aside from model-based estimates (e.g. Jung et al. 2020), reliable observation-based  
71 global photosynthesis proxies are only available for recent years through satellite-derived sun-  
72 induced fluorescence (SIF, Baker et al., 2008; Frankenberg et al., 2011; Joiner et al., 2013). SIF  
73 data is increasingly used to study the relationships between global vegetation productivity and  
74 hydro-meteorological drivers (Yang et al., 2015; Ying et al., 2015; Wagle et al., 2016; Zuromski  
75 et al., 2018; Jiao et al., 2019; Walther et al., 2019; Li & Xiao, 2020). Besides, spectral  
76 vegetation indices and biophysical parameters from multi-spectral satellite instruments such as  
77 the Moderate Resolution Imaging Spectroradiometer (MODIS) are widely used to study drivers

78 of vegetation phenology and productivity (Forkel et al. 2015; Seddon et al., 2016; Buermann et  
79 al., 2018). In this study, we consider SIF alongside two spectral indices (the normalized  
80 difference vegetation index, NDVI; and near-infrared reflectance of terrestrial vegetation, NIR<sub>v</sub>),  
81 and a comprehensive set of explanatory variables representing energy (temperature; radiation;  
82 vapor pressure deficit, VPD) and water availability (precipitation; multi-layer SM) to revisit  
83 global photosynthesis and greenness controls.

## 84 **2. Data and Methods**

### 85 2.1. Vegetation Target Data

#### 86 2.1.1. Sun-Induced Fluorescence (SIF)

87 SIF is a proxy for photosynthesis as it captures radiation emitted by chlorophyll  
88 molecules and is related to photosynthetic activity. We use one of the longest available satellite-  
89 derived SIF retrieval which is based on the Global Ozone Monitoring Experiment-2 (GOME-2)  
90 instrument and ranges from 2007 to 2018 (Köhler et al., 2015). The raw global SIF observations  
91 are filtered to remove data based on (i) high solar zenith angles ( $>70^\circ$ ), (ii) large differences to  
92 the normal local overpass time (2 p.m-8 a.m in the next day), and (iii) large cloud cover ( $>50\%$ ),  
93 as done by Köhler et al., 2015.

#### 94 2.1.2. Vegetation Indices

95 To complement the photosynthesis analysis we use NDVI and NIR<sub>v</sub> as spectral  
96 vegetation indices (Huete et al., 2002; Badgley et al., 2017). We obtain red and near-infrared  
97 reflectances from MOD13C1 v006 product (<https://lpdaac.usgs.gov/products/mod13c1v006/>) in  
98 an original 16-day and  $0.05^\circ$  resolution. NDVI and NIR<sub>v</sub> are computed from data with quality  
99 flags 0 and 1.

### 100 2.2. Hydro-meteorological predictor data

101 We consider a comprehensive selection of energy and water-related variables from the  
102 ERA5 reanalysis (Hersbach et al., 2019). Energy-related variables include air temperature at 2-m  
103 height (hereafter referred to as temperature), surface downward solar radiation (solar radiation)  
104 and VPD, and the water-related variables are total precipitation (precipitation), SM layer 1 (0-7

105 cm), layer 2 (7-28 cm), layer 3 (28-100 cm) and layer 4 (100-289 cm). For comparison, we  
106 compute total SM by averaging values across the individual layers weighted by their thickness.  
107 It is to note that VPD is related to the relative humidity and temperature, and hence we treat it as  
108 an energy-related variable, while it represents the demand of the water in the atmosphere.

109 To validate our findings we also use alternative SM products: (i) MERRA-2 surface and  
110 root-zone SM (Gelaro et al., 2017), (ii) GLEAM v3.3 surface and root-zone SM (Martens et al.,  
111 2017), and (iii) SoMo.ml with three layers (O and Orth, 2020). Table S1 shows the information  
112 of depths for all SM products that we use and classify into surface SM, shallow and deep root-  
113 zone SM.

### 114 2.3. Additional data

115 To evaluate the results of our analyses, we compute the aridity index for each grid cell as  
116 the ratio between the long-term averages of net radiation (expressed as mm potential evaporation)  
117 and precipitation from the respective ERA5 data. We distinguish climate regimes using long-  
118 term mean temperatures and aridity index. In addition, we use fractional vegetation coverage  
119 (FVC) data from the AVHRR vegetation continuous fields products (VCF5KYR,  
120 <https://lpdaac.usgs.gov/products/vcf5kyrv001/>) from 2007 to 2016 to classify the percentages of  
121 tree canopy, short vegetation and bare ground (Song et al., 2018). We distinguish vegetation  
122 characteristics using the fraction of vegetation cover (the sum of the fractions of tree canopy and  
123 short vegetation), and the fraction of tree cover in vegetation cover.

### 124 2.4. Methods

#### 125 2.4.1. Data Pre-processing

126 The data pre-processing is illustrated in Figure S1. All vegetation indices and hydro-  
127 meteorological data are aggregated to  $0.5^\circ$  spatial and half-monthly temporal resolution where  
128 SIF is available, and 16-day original NDVI and NIRv are linearly interpolated to half-monthly  
129 resolution. The study time period is 2007-2018, limited by the availability of SIF. In all SIF-  
130 based analyses we focus on data with  $SIF > 0.5 \text{ mW/m}^2/\text{sr/nm}$  to filter out sparse or dormant  
131 vegetation. This filtering is also applied in the NDVI and NIRv analyses, where additionally  
132 negative NDVI and NIRv values are filtered out. Grid cells are only considered in the analysis if  
133 more than 15 data points are left after filtering, and if the vegetation cover from the FVC data  
134 exceeds 5%. For all target and predictor variables, we obtain half-monthly anomalies by

135 subtracting the mean seasonal cycles. We remove long-term trends for each grid cell which are  
136 determined by a locally weighted smoothing filter (Cleveland et al., 1979) with a smoothing  
137 span of 0.4.

#### 138 2.4.2. Identification of main controls

139 Random forests (RF) is a non-parametric regression-based method requiring no  
140 statistical assumptions on predictor and target variables (Breiman 2001). In this study, all hydro-  
141 meteorological anomalies are used as predictor variables, and anomalies of SIF and vegetation  
142 indices are employed as target variables per each grid cell, respectively (Figure S1). RF training  
143 is done using information from each grid cell together with the surrounding grid cells (forming  
144 3x3 grid cell matrices) to yield robust model performance while including data with similar  
145 climatic and landscape characteristics. After training, the performance of the RF model is  
146 evaluated at each grid cell by computing the  $R^2$  between the modeled and observed target  
147 variable for out-of-bag (OOB) data that was not used for training (hereafter referred as  $R^2$ ). Grid  
148 cells with  $R^2$  lower than or equal to 0 are filtered out.

149 The relative importance of each predictor variable is inferred from the decrease in  $R^2$   
150 related to a temporal permutation applied to the particular variable (Cutler et al., 2012; Gómez-  
151 Ramírez et al., 2019). To validate our findings we additionally employ two more methods in this  
152 context: (i) Spearman correlation between each predictor variable and SIF, NDVI or NIRv  
153 (Zwillinger & Kokoska, 2000) and (ii) SHapley Additive exPlanations (SHAP) feature  
154 importance which is based on the average marginal contribution of each predictor to the  
155 modeled target variable (Lundberg et al. 2017; Sundararajan et al., 2019).

156 In addition to the determination of the most relevant hydro-meteorological controls we  
157 study the sensitivity of the vegetation response to each predictor variable. The sensitivity is  
158 determined by the slope from fitted linear quantile (median) regression between the SHAP  
159 dependence of a target variable and a predictor variable, as SHAP dependence enables to  
160 measure the marginal effect each predictor variable has on the target variable for individual and  
161 global explanations (Lundberg et al. 2017; Forkel et al., 2019). While the magnitude of the  
162 sensitivity is usually similar to the identification of feature importance, the sign of sensitivity  
163 complements the information in importance identification. All data-processing and analyses are  
164 done with Python 3.7 by using the NumPy 1.16.1 (Oliphant 2006), Statsmodels 0.11.1 (Skipper

165 & Perktold, 2010), Scikit-learn 0.22.1 (Pedregosa et al., 2011), Matplotlib (Hunter 2007) and  
166 shap 0.35.0 packages (Lundberg et al. 2017).

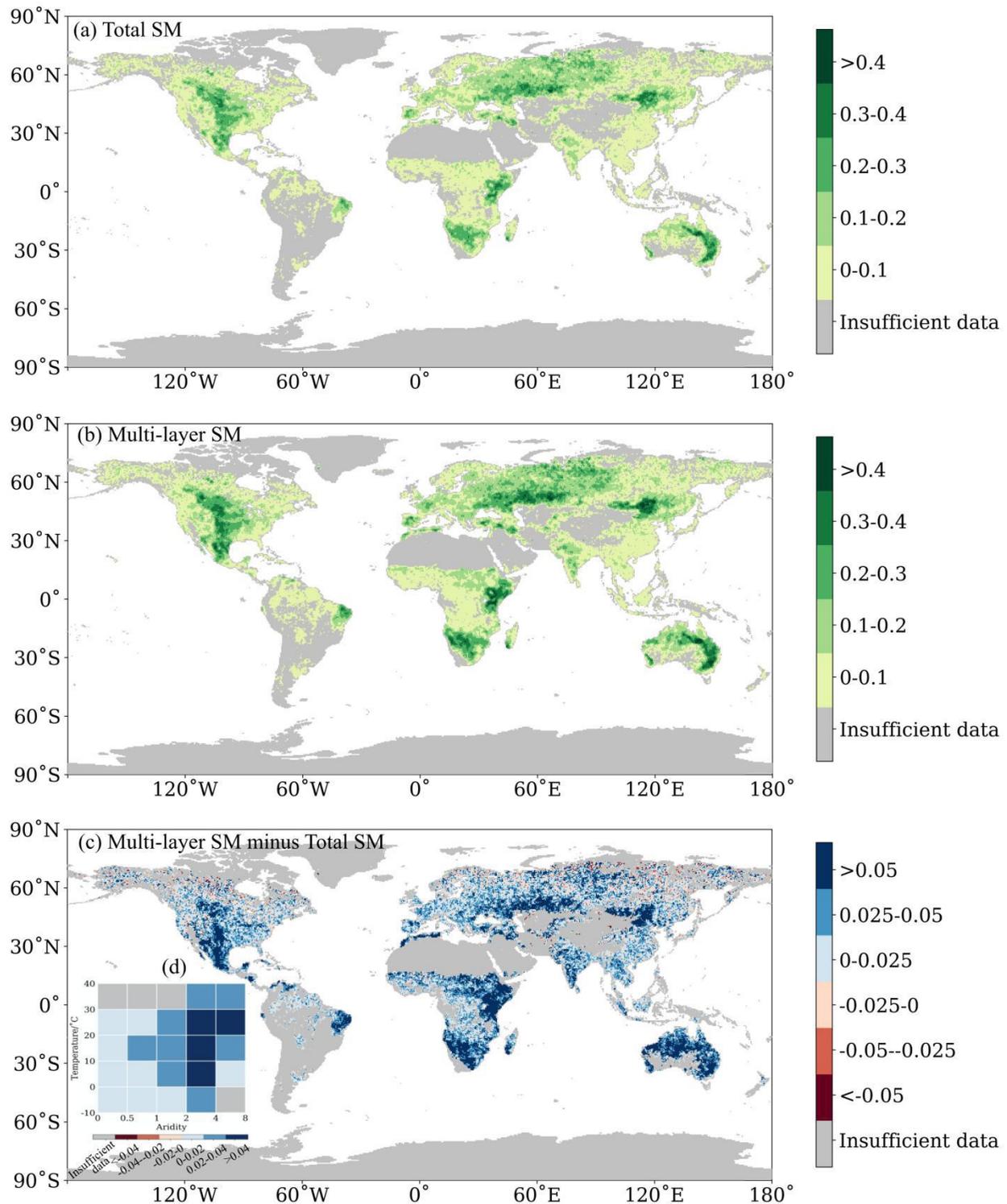
167

168 **3. Results and Discussion**

169 3.1. Model performance

170 Two experiments are performed with RF models differing in how SM is accounted for  
171 (i.e. total versus multi-layer SM), while precipitation, VPD, solar radiation, and temperature are  
172 used consistently in both experiments. Results show that the performance of the RF model in  
173 predicting SIF anomalies is higher using multi-layer SM than that with total SM (Figure 1).

174



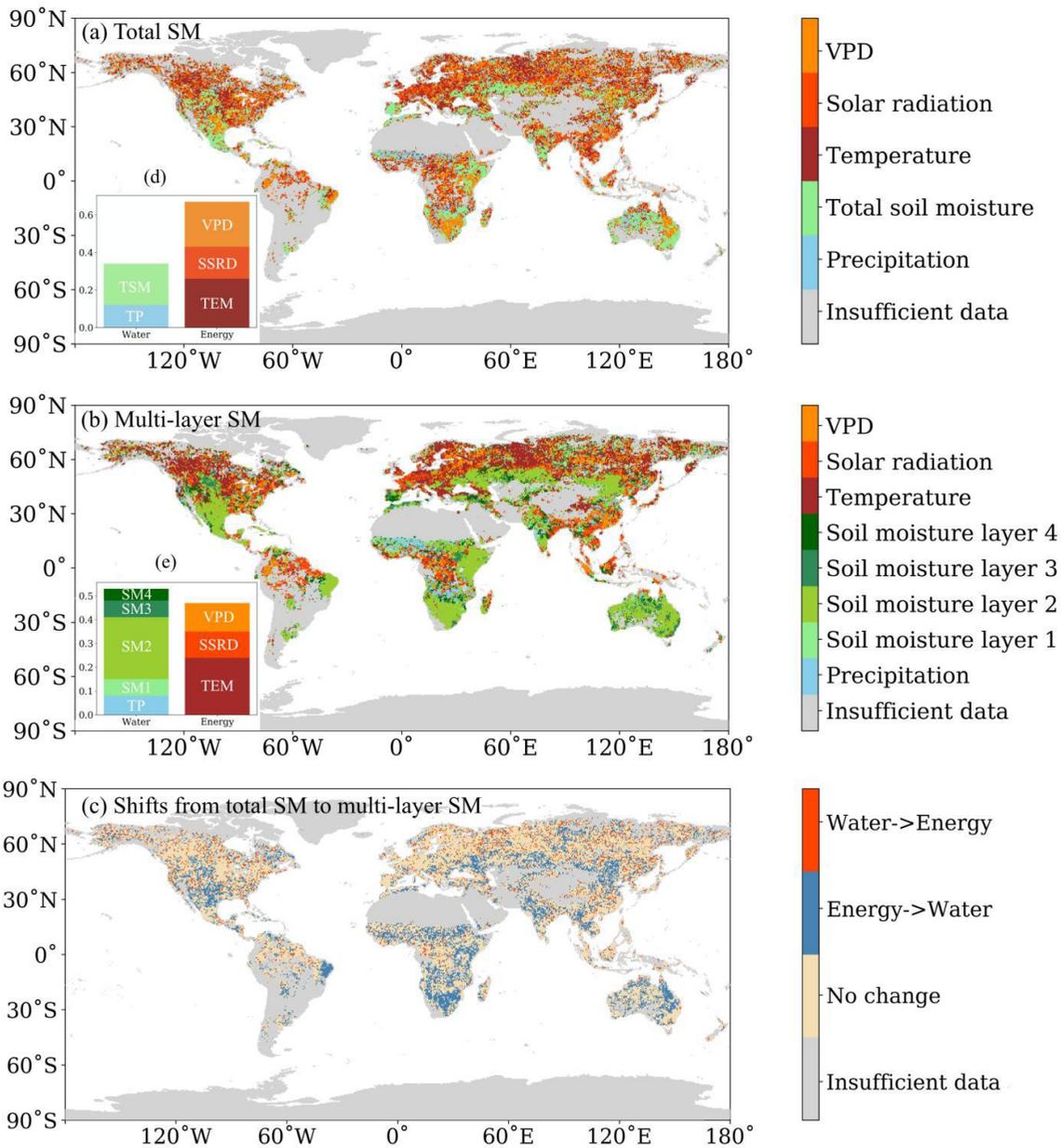
175  
 176 **Figure 1. Model performance ( $R^2$ ) in predicting Sun-Induced Fluorescence (SIF) in (a) the**  
 177 **total soil moisture (SM) experiment and (b) the multi-layer SM experiment (SM layers 1-4).**  
 178 **The panel (c) is the difference between (b) and (a), and (d) summarizes their differences**  
 179 **across climate regimes (i.e. Temperature and Aridity).**

180 The spatial patterns of model performance are similar between both experiments with  
181 higher  $R^2$  ( $> 0.3$ ) in the central North America, central Eurasia, southern and eastern Africa,  
182 central Asia, and eastern Australia. The predictive performance is improved in most regions  
183 across the globe when using multi-layer instead of total SM. Improvements are particularly  
184 found in semi-arid regions such as Australia, central North America and central Asia (Figure 1c,  
185 d). Since multi-layer SM may experience different dynamics across time and space (Schlaepfer  
186 et al., 2017; Berg et al., 2016; Zhang et al., 2016; Lian et al., 2020), plant rooting systems can  
187 develop to adapt for localized water deficits (Fan et al., 2017), such that vertical SM information  
188 can be especially useful in semi-arid regions to predict the vegetation productivity.

189 Though the performance of SIF prediction is improved with multi-layer SM, the  $R^2$   
190 values are still relatively low in many regions. There are even some regions that show  $R^2$  lower  
191 than 0 in South America and central Australia, indicating a worse model performance than a  
192 constant mean value prediction. Such limited reliability of SIF predictions may relate to the  
193 noise of satellite-derived SIF, for example, large regions in South America are located near to  
194 the known South Atlantic Anomaly, which disturbs the satellite-based SIF retrievals (Joiner et  
195 al., 2013; Köhler et al., 2015). This disturbance is less relevant for the NDVI and NIRv  
196 retrievals such that RF model performance is better (Figure S2). Despite the weak model  
197 performance in the case of SIF we believe that our methodology is robust to infer main hydro-  
198 meteorological controls of vegetation productivity, because (i) the employed  $R^2$  of out-of-bag  
199 anomaly data is a challenging metric where information cannot be derived from e.g. seasonal  
200 variations or trends, and also other studies found similarly low values (Kraft et al., 2019); and (ii)  
201 main hydro-meteorological controls on SIF anomalies identified by RF model resemble global  
202 patterns reported in previous studies about main climatic drivers to absolute variations of  
203 vegetation productivity (Figure 2) (Nemani et al., 2003; Seddon et al., 2016; Madani et al.,  
204 2017).

205

206



207

208 **Figure 2. Main hydro-meteorological controls on sun-induced fluorescence (SIF) by applying (a)**  
 209 **total soil moisture (SM) alongside all other predictor variables, and (b) multi-layer SM alongside**  
 210 **all other predictor variables. (c) Shifts between the energy and water controls from (a) to (b).**  
 211 **Proportions of global land area where each variable is the most important controlling factor are**  
 212 **shown in (d) and (e). In (d) and (e), TP denotes precipitation; TSM denotes total soil moisture;**  
 213 **SM1, 2, 3, 4 denote soil moisture in layers 1, 2, 3, 4 respectively; TEM denotes temperature; SSRD**  
 214 **denotes solar radiation; And VPD denotes vapor pressure deficit. As shown in Table S1, SM layer**  
 215 **1 in ERA5 belongs to surface SM, SM layer 2 and 3 belong to shallow root-zone SM, and SM layer**  
 216 **4 belongs to deep root-zone SM.**

217 We perform further RF model experiments to investigate if the added skill in the case of  
218 the multi-layer SM is related to the increased number of predictor variables, and therefore an  
219 increased flexibility of the model, or to the additional information contained in the individual  
220 layers compared with the total SM. First, the experiment of multi-layer RF (4 variables)  
221 preforms better than the experiment of 5 SM variables, showing that the enhanced performance  
222 is not exclusively due to the increased number of variables and related to increased flexibility of  
223 the RF model (Figure S3). Second, regionally enhanced performance can be found when  
224 replacing total SM with individual layers (Figure S4), indicating that additional information can  
225 be explored by the RF model from SM from individual layers.

### 226 3.2. Main hydro-meteorological controls on global vegetation productivity

227 The global patterns of main SIF controls are clearly different between the analyses with  
228 total SM and with multi-layer SM (Figure 2); Total SM does not provide sufficient information  
229 to the RF model to detect all water-controlled regions while these regions are actually covering  
230 the majority of the Earth's land in the analysis with multi-layer SM. Overall, temperature is  
231 identified as the main driver of SIF in the higher northern latitudes, solar radiation dominantly  
232 controls SIF in most tropical regions, and VPD emerges as a main control on SIF in parts of the  
233 western Amazon forests, eastern North America, northern Eurasia and eastern Asia. In between  
234 the tropics and the higher latitudes, where mostly semi-arid climate regimes are prevailing,  
235 water-related variables play the dominant role in controlling SIF. Precipitation and surface SM  
236 control SIF in central India, western Sahel and transition regions between central and southern  
237 Africa. Root-zone SM mainly controls SIF in southern North America, southern Europe, and  
238 many parts of Eurasia, India and Australia. In general, shallow-root zone SM emerges as the  
239 most relevant SM reservoir for vegetation productivity, while deeper SM is particularly  
240 important in the transitional zones and temperate dry regions, such as central North America and  
241 southern South Europe.

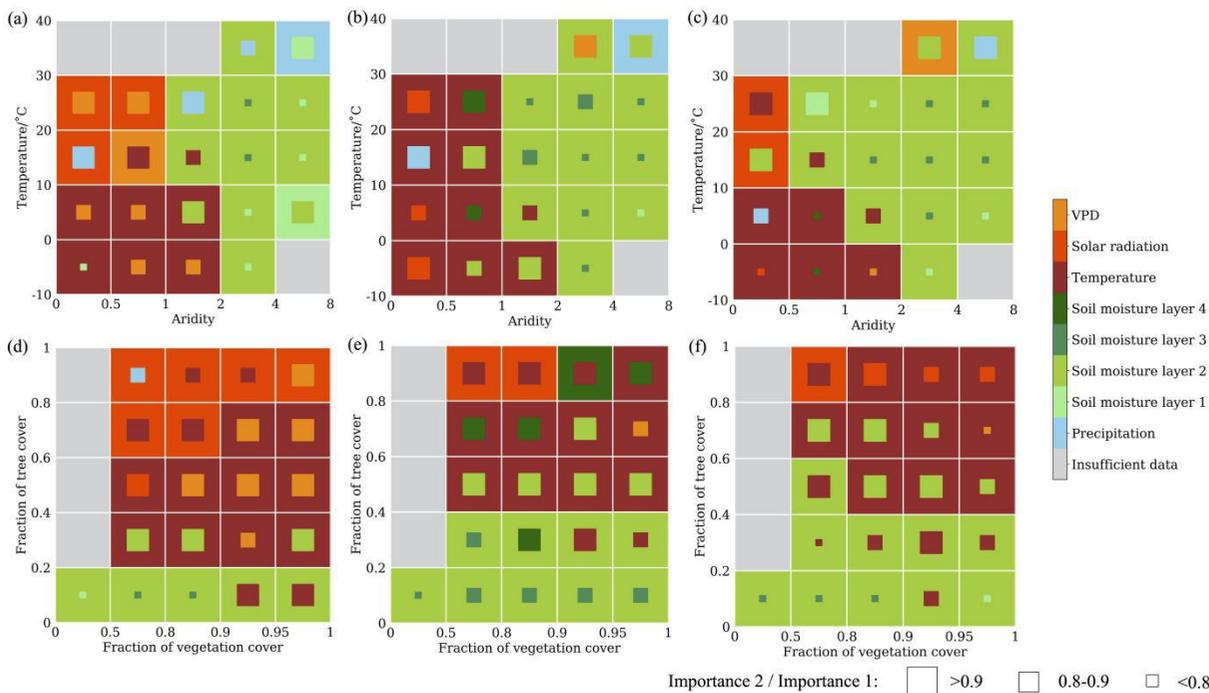
242 Key drivers of NIRv and NDVI present similar global patterns to those of SIF (Figure  
243 S5), while they show extended SM-controlled regions. Walther et al., 2019 also found  
244 inconsistent values of tree cover fraction with shifting relationships between SM and SIF or  
245 vegetation indices, relating to the fact that spectral greenness signals are somewhat influenced  
246 by moisture-related changes in the soil reflectivity or plant water content. Further, for the

247 respective main controlling hydro-meteorological variables identified across space, we typically  
248 find highly positive associated sensitivities of SIF to the respective control, which supports  
249 positive relationships between the identified main controls and SIF (Figure S6).

250         Next, we analyze the main controls with respect to climate regimes. Figure 3a shows that  
251 the SM variables dominantly control SIF in arid regions, energy-related variables dominantly  
252 control SIF in humid regions. In transitional regions water-related variables tend to be more  
253 important at warmer temperatures, while energy-related variables dominate for colder  
254 temperatures. Overall, the pattern is in line with first-order constraints for evapotranspiration  
255 from Seneviratne et al., 2010, and with findings on energy- versus water-dominated vegetation  
256 by Denissen et al. 2020 in Europe. Across all considered hydro-meteorological variables,  
257 shallow-root zone SM is identified as the most important variable in (semi-)arid regions. Among  
258 the energy variables temperature is the most relevant, while solar radiation also plays a role  
259 particularly in warm regions. Similar patterns are found for NDVI and NIRv with SM controls  
260 extending more beyond arid regions (Figure 3b, c).

261

262



263

264 **Figure 3. Main hydro-meteorological controls on (a, d) sun-induced fluorescence SIF, (b, e)**  
 265 **Near-Infrared reflectance vegetation indices (NIR<sub>v</sub>) and (c, f) normalized difference**  
 266 **vegetation indices (NDVI) across climate regimes and vegetation characteristics. Most**  
 267 **important control variables are indicated by the color of the temperature-aridity and tree-**  
 268 **vegetation boxes, respective second most important control variables are denoted by the**  
 269 **color of the inner square, where the size indicates the relative importance compared to the**  
 270 **most important control variable. Temperature-aridity and tree-vegetation boxes**  
 271 **containing less than 10 available data are shown in gray. The aridity index and the fraction**  
 272 **of vegetation cover are visualized by non-linear sequences in terms of skewed distributions**  
 273 **of the data.**

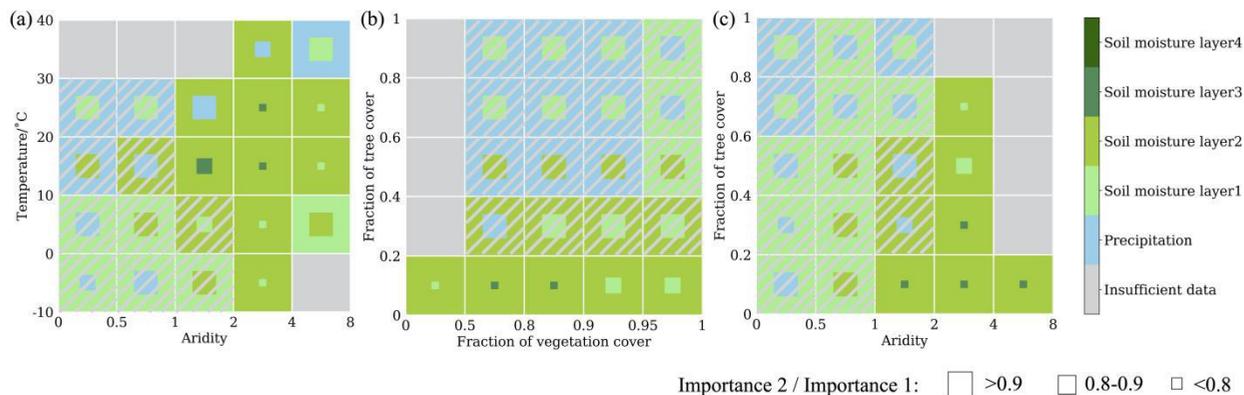
274 Main controls also differ with vegetation types (Figure 3d, e, f), mostly varying along a  
 275 gradient in the fraction of tree cover while they are more similar between different fractions of  
 276 vegetation cover. Regions dominated by grass or shrubs are most water-controlled, regions with  
 277 intermediate tree cover are temperature-controlled, and regions with the highest tree cover and  
 278 presumably wet or temperate climate conditions are mostly radiation-controlled. Such main  
 279 energy controls involve a relatively lower vulnerability of tree ecosystems to droughts than other  
 280 ecosystems (Huang & Xia, 2019), as droughts are typically associated with above-average solar

281 radiation and newly developing leaves that can compensate photosynthesis (Orth & Destouni,  
282 2018; Yan et al., 2019; Hutyra et al., 2007; Wu et al., 2016; Li et al., 2018b). Moreover,  
283 consistent with the previous findings, NDVI and NIRv show extended significant water-related  
284 controls to tree-grass mixed biomes compared with the SIF results (Walther et al., 2019). This is  
285 more pronounced for NDVI, potentially due to larger confounding effects of background  
286 brightness in NDVI, while NIRv contains enhanced information about the proportion of  
287 vegetation in reflectance and partly overcomes this issue (Badgley et al., 2017; Badgley et al.,  
288 2019). Changes in main controls across vegetation characteristics are not simply an artifact of  
289 the correspondingly different climate regimes, as Figure S7 shows that the main hydro-  
290 meteorological controls change in response to both vegetation type and climate.

### 291 3.3. Main water-related controls on global vegetation productivity

292 Focusing exclusively on water-related controls reveals that the most important soil layer  
293 varies across climate and vegetation characteristics (Figure 4). Shallow-root zone SM is most  
294 relevant in semi-arid conditions and for grass or shrubs, indicating that plants can adapt to  
295 water-scarce conditions at the surface with deeper-reaching rooting systems (Fan et al., 2017).  
296 This is in line with previous but smaller-scale studies: A et al., 2019 found the strongest  
297 relationship between evapotranspiration and SM between 10-100 cm depth for site-scale  
298 experiments in a transitional zone; further, in dry surface soils in (semi-)arid regions, plants  
299 could easily alter rooting depth distribution and root morphology to utilize water from deeper  
300 soil layers (Schulze et al., 1996), for instance in local Mediterranean grass (Barkaoui et al., 2016)  
301 or savannas ecosystems (Hoekstra et al., 2014; Nippert & Holdo, 2015). For even drier climate  
302 conditions, shallower soil layers become more relevant, probably because the low water supply  
303 does not sustain the development of deep(er) rooting systems such that intermittent vegetation  
304 growth mostly benefits from rainfed surface SM. Interestingly, towards humid climate  
305 conditions our analysis shows a dominant role of surface SM and precipitation, while at the  
306 same time these regions are characterized by high tree cover with deep roots. This could be due  
307 to frequent precipitation keeping surface soil layers wet such that plants can extract significant  
308 fractions of their water demand from there, while the dependence on deeper layers for trees  
309 during short drought periods is not reflected. Furthermore, we note that these regions are  
310 controlled by temperature or solar radiation (see Figure 3) such that the results here could also

311 be an artifact as precipitation and partly also surface SM are expected to co-vary more strongly  
 312 with the dominant energy variables than deeper-layer SM.



313  
 314 **Figure 4. Main water-related controls on sun-induced fluorescence (SIF) across (a) climate**  
 315 **regimes, (b) vegetation characteristics, and (c) classes of fraction of tree covers and aridity.**  
 316 **Similar to Figure 3 but focusing on SIF and water-related controls only. The gray hatching**  
 317 **indicates that temperature, solar radiation or VPD are identified as main controls on SIF**  
 318 **in these boxes in Figure 3.**

319 To illustrate the robustness of our results, we repeat the previous analyses with different  
 320 setups: (i) we use Spearman correlation (Figure S8) and SHAP feature importance (Figure S9)  
 321 as alternative ways to estimate the importance of the considered predictor variables for SIF  
 322 dynamics, and find similar results as for the permutation importance approach, and (ii) we use  
 323 alternative SM products, namely GLEAM, MERRA-2 and SoMo.ml (Figure S10), all of which  
 324 lead to similar results as found with the ERA5 SM.

325 We acknowledge, however, that our analyses do not consider seasonal compensation  
 326 effects, memory effects and irrigation effects when illustrating main hydro-meteorological  
 327 controls on vegetation productivity. Memory effects are found occurring particularly in  
 328 transitional water-driven biomes and sub-tropical regions (Kraft et al., 2019). Precipitation from  
 329 wet seasons can serve as subsurface water storage in subsequent dry seasons (Guan et al., 2015),  
 330 and water transport in roots and stems might be slow or delayed for tree ecosystems, affecting  
 331 energy- or water-control characteristics across biomes. Besides, warm springs benefit  
 332 photosynthesis in the early stage of the growing season, while induce water deficits in the later  
 333 seasons in northern energy-limited ecosystems (Buermann et al., 2018). Finally the main hydro-  
 334 meteorological controls which we determine for the entire growing season may vary between the

335 early, mid and later parts of this period. We further note that our analyses is based on specific  
336 spatial and temporal scales, while vegetation-climate relationships can differ between short-term  
337 and long-term scales (Linscheid et al., 2019), and contrasting signals from nearby regions could  
338 lead to inconclusive results (Jung et al., 2017).

339

#### 340 4. Conclusions

341 This study illustrates that the information of vertically resolved SM improves the  
342 understanding and modeling of anomalies of vegetation productivity. Thereby, vegetation relies  
343 on water from different depths while these characteristic depths vary with climate and vegetation  
344 type. In particular, we show at the global scale that vegetation in semi-arid regions is adapted to  
345 dry conditions through deep(er) rooting systems ensuring more continuous water supply from  
346 deeper soil layers. This complexity was not sufficiently acknowledged in previous studies;  
347 future research should account for vertical SM dynamics by considering multiple layers. The  
348 development of hydrology, land surface, and vegetation models should focus on a reliable  
349 representation of soil layers and vertical soil water transport.

350 Further, we compare the hydro-meteorological controls of vegetation productivity  
351 obtained with different respective proxy metrics. SIF is more strongly related to photosynthesis,  
352 and eventually the carbon cycle, compared with NDVI and NIR<sub>v</sub>, but SIF data is only available  
353 for recent years. Our results show that NDVI and NIR<sub>v</sub>, which are available from the early  
354 1980s, yield similar patterns except for a consistent overestimation of water controls, probably  
355 induced by changes of soil background reflectance as a response to soil moisture changes.

356 Overall, our study contributes to advanced process of understanding within the role of  
357 soil moisture on vegetation productivity by benefiting from the ever-growing suite of global  
358 eco-hydrological data streams.

359

360

361

362

363

364 **Acknowledgments**

365 The authors declare no conflict of interest. We thank Ulrich Weber for processing hydro-  
 366 meteorological data and MODIS data. W. L., R. O., and M. M. acknowledge the financial  
 367 support of the China Scholarship Council that funded the PhD scholarship of W.L. W. L. also  
 368 acknowledges this work under the International Max Planck Research School for Global  
 369 Biogeochemical Cycles. R. O. was funded by the German Research Foundation (Emmy Noether  
 370 grant number 391059971), and S. W. acknowledges funding by an ESA Living Planet  
 371 Fellowship ‘Vad3e mecum’. SIF GFZ data have been retrieved from  
 372 <ftp://fluo.gps.caltech.edu/data/Philipp/GOME-2/ungridded/>. ERA5 data can be downloaded  
 373 from <https://cds.climate.copernicus.eu/>, GLEAM SM from <https://www.gleam.eu/>, MERRA-2  
 374 SM data from <https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/FAQ/>, and SoMo.ml from  
 375 [ftp://ftp.bgc-jena.mpg.de/pub/outgoing/sungmino/somo\\_v1/](ftp://ftp.bgc-jena.mpg.de/pub/outgoing/sungmino/somo_v1/). All the links of the data are  
 376 accessed on 29 September 2020.

377

378 **References**

- 379 A, Y., Wang, G., Liu, T., Xue, B., & Kuczera, G. (2019). Spatial variation of correlations  
 380 between vertical soil water and evapotranspiration and their controlling factors in a semi-arid  
 381 region. *Journal of Hydrology*, 574, 53-63. <https://doi.org/10.1016/j.jhydrol.2019.04.023>
- 382 Albergel, C., Dorigo, W., Reichle, R., Balsamo, G., De Rosnay, P., Muñoz-Sabater, J., Isaksen,  
 383 L., De Jeu, R., & Wagner, W. (2013). Skill and global trend analysis of soil moisture from  
 384 reanalyses and microwave remote sensing. *Journal of Hydrometeorology*, 14(4), 1259-1277.  
 385 <https://doi.org/10.1175/JHM-D-12-0161.1>
- 386 Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial  
 387 photosynthesis. *Science Advances*, 3(3), e1602244. <https://doi.org/10.1126/sciadv.1602244>
- 388 Badgley, G., Anderegg, L. D. L., Berry, J. A., & Field, C. B. (2019). Terrestrial gross primary  
 389 production: Using NIRV to scale from site to globe. *Global Change Biology*, 25(11), 3731-  
 390 3740. <https://doi.org/10.1111/gcb.14729>
- 391 Baker, N. R. (2008). Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annual*  
 392 *Review of Plant Biology*, 59, 89-113.  
 393 <https://doi.org/10.1146/annurev.arplant.59.032607.092759>
- 394 Barkaoui, K., Roumet, C., & Volaire, F. (2016). Mean root trait more than root trait diversity  
 395 determines drought resilience in native and cultivated Mediterranean grass mixtures.  
 396 *Agriculture, Ecosystems & Environment*, 231, 122-132.  
 397 <https://doi.org/10.1016/j.agee.2016.06.035>
- 398 Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., ... & Papale, D.  
 399 (2010). Terrestrial gross carbon dioxide uptake: global distribution and covariation with  
 400 climate. *Science*, 329(5993), 834-838. <https://doi.org/10.1126/science.1184984>
- 401 Berg, A., Sheffield, J., & Milly, P. C. (2017). Divergent surface and total soil moisture  
 402 projections under global warming. *Geophysical Research Letters*, 44(1), 236-244.  
 403 <https://doi.org/10.1002/2016GL071921>
- 404 Breiman, L. (2001). Random forests. *Mach. Learning* 45, 5–32.
- 405 Buermann, W., Forkel, M., O’Sullivan, M., Sitch, S., Friedlingstein, P., Haverd, V., Jain, A. K.,  
 406 Kato, E., Kautz, M., & Lienert, S. (2018). Widespread seasonal compensation effects of  
 407 spring warming on northern plant productivity. *Nature*, 562(7725), 110-114.  
 408 <https://doi.org/10.1038/s41586-018-0555-7>

- 409 Claessen, J., Molini, A., Martens, B., Detto, M., Demuzere, M., & Miralles, D. (2019). Global  
 410 biosphere-climate interaction: a causal appraisal of observations and models over multiple  
 411 temporal scales. *Biogeosciences*, *16*(24), 4851-4874. [https://doi.org/10.5194/bg-16-4851-](https://doi.org/10.5194/bg-16-4851-2019)  
 412 [2019](https://doi.org/10.5194/bg-16-4851-2019)
- 413 Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots.  
 414 *Journal of the American Statistical Association*, *74*(368), 829-836.  
 415 <https://doi.org/10.1080/01621459.1979.10481038>
- 416 Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., & Luke,  
 417 C. M. (2013). Sensitivity of tropical carbon to climate change constrained by carbon dioxide  
 418 variability. *Nature*, *494*(7437), 341-344. <https://doi.org/10.1038/nature11882>
- 419 Cutler A., Cutler D.R., & Stevens J.R. (2012) Random Forests. In: Zhang C., & Ma Y. (eds)  
 420 Ensemble Machine Learning. Springer, Boston, MA. [https://doi.org/10.1007/978-1-4419-](https://doi.org/10.1007/978-1-4419-9326-7_5)  
 421 [9326-7\\_5](https://doi.org/10.1007/978-1-4419-9326-7_5)
- 422 Denissen, J. M., Teuling, A. J., Reichstein, M., & Orth, R. (2020). Critical soil moisture derived  
 423 from satellite observations over Europe. *J. Geophys. Res. Atmos.*, e2019JD031672.  
 424 <https://doi.org/10.1029/2019JD031672>
- 425 Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G.,  
 426 Gruber, B., Lafourcade, B., & Leitao, P. J. (2013). Collinearity: a review of methods to deal  
 427 with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27-46.  
 428 <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- 429 Duveiller, G., Hooker, J., & Cescatti, A. (2018). The mark of vegetation change on Earth's  
 430 surface energy balance. *Nature Communications*, *9*, 679. [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-017-02810-8)  
 431 [017-02810-8](https://doi.org/10.1038/s41467-017-02810-8)
- 432 Fan, Y., Miguez-Macho, G., Jobbagy, E. G., Jackson, R. B., & Otero-Casal, C. (2017).  
 433 Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of*  
 434 *Sciences*, *114*(40), 10572-10577. <https://doi.org/10.1073/pnas.1712381114>
- 435 Feng, S., Hu, Q., Huang, W., Ho, C.-H., Li, R., & Tang, Z. (2014). Projected climate regime  
 436 shift under future global warming from multi-model, multi-scenario CMIP5 simulations.  
 437 *Global and Planetary Change*, *112*, 41-52. <https://doi.org/10.1016/j.gloplacha.2013.11.002>
- 438 Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U., &  
 439 Carvalhais, N. (2015). Codominant water control on global interannual variability and trends

- 440 in land surface phenology and greenness. *Global Change Biology*, 21(9), 3414-3435.  
 441 <https://doi.org/10.1111/gcb.12950>
- 442 Forkel, M., Andela, N., Harrison, S. P., Lasslop, G., van Marle, M., Chuvieco, E., Dorigo, W.,  
 443 Forrest, M., Hantson, S., Heil, A., Li, F., Melton, J., Sitch, S., Yue, C., & Arneeth, A. (2019).  
 444 Emergent relationships with respect to burned area in global satellite observations and fire-  
 445 enabled vegetation models. *Biogeosciences*, 16(1), 57-76. [https://doi.org/10.5194/bg-16-57-](https://doi.org/10.5194/bg-16-57-2019)  
 446 [2019](https://doi.org/10.5194/bg-16-57-2019)
- 447 Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. E., Toon, G. C.,  
 448 Butz, A., Jung, M., & Kuze, A. (2011). New global observations of the terrestrial carbon  
 449 cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity.  
 450 *Geophysical Research Letters*, 38, L17706. <https://doi.org/10.1029/2011gl048738>
- 451 Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., Pollock, R., & Taylor,  
 452 T. E. (2014). Prospects for chlorophyll fluorescence remote sensing from the Orbiting  
 453 Carbon Observatory-2. *Remote Sensing of Environment*, 147, 1-12.  
 454 <https://doi.org/10.1016/j.rse.2014.02.007>
- 455 Garonna, I., de Jong, R., Stöckli, R., Schmid, B., Schenkel, D., Schimel, D., & Schaepman, M. E.  
 456 (2018). Shifting relative importance of climatic constraints on land surface phenology.  
 457 *Environmental Research Letters*, 13(2), 024025. <https://doi.org/10.1088/1748-9326/aaa17b>
- 458 Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A.,  
 459 Darmenov, A., Bosilovich, M. G., & Reichle, R. (2017). The modern-era retrospective  
 460 analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14),  
 461 5419-5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>
- 462 Gómez-Ramírez, J., Ávila-Villanueva, M., & Fernández-Blázquez, M. Á. (2019). Selecting the  
 463 most important self-assessed features for predicting conversion to Mild Cognitive  
 464 Impairment with Random Forest and Permutation-based methods. *bioRxiv*, 785519.  
 465 <https://doi.org/10.1101/785519>
- 466 Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., Caylor, K. K., Sheffield, J., Wood, E.  
 467 F., Malhi, Y., Liang, M., Kimball, J. S., Saleska, Scott R., Berry, J., Joiner, J., & Lyapustin,  
 468 A. I. (2015). Photosynthetic seasonality of global tropical forests constrained by  
 469 hydroclimate. *Nature Geoscience*, 8(4), 284-289. <https://doi.org/10.1038/ngeo2382>

- 470 Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Sabater, J. M., Nicolas, J., Radu, R.,  
 471 Schepers, D., Simmons, A., & Soci, C. (2019). Global reanalysis: goodbye ERA-Interim,  
 472 hello ERA5. *ECMWF newsletter*, 159, 17-24. <https://doi.org/10.21957/vf291hehd7>
- 473 Hoekstra, N. J., Finn, J. A., Hofer, D., & Lüscher, A. (2014). The effect of drought and  
 474 interspecific interactions on depth of water uptake in deep-and shallow-rooting grassland  
 475 species as determined by  $\delta^{18}\text{O}$  natural abundance. *Biogeosciences*, 11(16), 4493-4506.  
 476 <https://doi.org/10.5194/bg-11-4493-2014>
- 477 Huang, K., & Xia, J. (2019). High ecosystem stability of evergreen broadleaf forests under  
 478 severe droughts. *Global Change Biology*, 25(10), 3494-3503.  
 479 <https://doi.org/10.1111/gcb.14748>
- 480 Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview  
 481 of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote*  
 482 *Sensing of Environment*, 83(1-2), 195-213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- 483 Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science &*  
 484 *Engineering*, 9(3), 90-95. <https://doi.org/10.1109/MCSE.2007.55>
- 485 Hutya, L. R., Munger, J. W., Saleska, S. R., Gottlieb, E., Daube, B. C., Dunn, A. L., Amaral, D.  
 486 F., De Camargo, P. B., & Wofsy, S. C. (2007). Seasonal controls on the exchange of carbon  
 487 and water in an Amazonian rain forest. *Journal of Geophysical Research: Biogeosciences*,  
 488 112, G3008. <https://doi.org/10.1029/2006JG000365>
- 489 Jeong, S.-J., Schimel, D., Frankenberg, C., Drewry, D. T., Fisher, J. B., Verma, M., Berry, J. A.,  
 490 Lee, J. E., & Joiner, J. (2017). Application of satellite solar-induced chlorophyll fluorescence  
 491 to understanding large-scale variations in vegetation phenology and function over northern  
 492 high latitude forests. *Remote Sensing of Environment*, 190, 178-187.  
 493 <https://doi.org/10.1016/j.rse.2016.11.021>
- 494 Jiao, W., Chang, Q., & Wang, L. (2019). The Sensitivity of Satellite Solar - Induced  
 495 Chlorophyll Fluorescence to Meteorological Drought. *Earth's Future*, 7(5), 558-573.  
 496 <https://doi.org/10.1029/2018ef001087>
- 497 Jing, W., Song, J., & Zhao, X. (2018). Validation of ECMWF multi-layer reanalysis soil  
 498 moisture based on the OzNet hydrology network. *Water*, 10(9), 1123.  
 499 <https://doi.org/10.3390/w10091123>

- 500 Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich,  
 501 K. F., Yoshida, Y., & Frankenberg, C. (2013). Global monitoring of terrestrial chlorophyll  
 502 fluorescence from moderate-spectral-resolution near-infrared satellite measurements:  
 503 methodology, simulations, and application to GOME-2. *Atmospheric Measurement*  
 504 *Techniques*, 6(10), 2803-2823. <https://doi.org/10.5194/amt-6-2803-2013>
- 505 Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., ... &  
 506 Williams, C. (2011). Global patterns of land - atmosphere fluxes of carbon dioxide, latent  
 507 heat, and sensible heat derived from eddy covariance, satellite, and meteorological  
 508 observations. *Journal of Geophysical Research: Biogeosciences*, 116, G00J07.  
 509 <https://doi.org/10.1029/2010JG001566>
- 510 Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlstrom, A., Arneeth, A.,  
 511 Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E.,  
 512 Papale, D., Poulter, B., Raduly, B., Rodenbeck, C., Tramontana, G., Viovy, N., Wang, Y. P.,  
 513 Weber, U., Zaehle, S., & Zeng, N. (2017). Compensatory water effects link yearly global  
 514 land CO2 sink changes to temperature. *Nature*, 541(7638), 516-520.  
 515 <https://doi.org/10.1038/nature20780>
- 516 Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P.,  
 517 Besnard, S., Bodesheim, P., & Carvalhais, N. (2020). Scaling carbon fluxes from eddy  
 518 covariance sites to globe: synthesis and evaluation of the FLUXCOM approach.  
 519 *Biogeosciences*, 17(5), 1343-1365. <https://doi.org/10.5194/bg-17-1343-2020>
- 520 Köhler, P., Guanter, L., & Joiner, J. (2015). A linear method for the retrieval of sun-induced  
 521 chlorophyll fluorescence from GOME-2 and SCIAMACHY data. *Atmospheric Measurement*  
 522 *Techniques*, 8(6), 2589-2608. <https://doi.org/10.5194/amt-8-2589-2015>
- 523 Kraft, B., Jung, M., Körner, M., Mesa, C. R., Cortés, J., & Reichstein, M. (2019). Identifying  
 524 dynamic memory effects on vegetation state using recurrent neural networks. *Frontiers in*  
 525 *Big Data*, 2(31). <https://doi.org/10.3389/fdata.2019.00031>
- 526 Li, M., Wu, P., & Ma, Z. (2020). A comprehensive evaluation of soil moisture and soil  
 527 temperature from third - generation atmospheric and land reanalysis data sets. *International*  
 528 *Journal of Climatology*, 1-23. <https://doi.org/10.1002/joc.6549>

- 529 Li, X., & Xiao, J. (2019). A Global, 0.05-Degree Product of Solar-Induced Chlorophyll  
 530 Fluorescence Derived from OCO-2, MODIS, and Reanalysis Data. *Remote Sensing*, 11(5).  
 531 <https://doi.org/10.3390/rs11050517>
- 532 Li, X., & Xiao, J. (2020). Global climatic controls on interannual variability of ecosystem  
 533 productivity: Similarities and differences inferred from solar-induced chlorophyll  
 534 fluorescence and enhanced vegetation index. *Agricultural and Forest Meteorology*, 288-289.  
 535 <https://doi.org/10.1016/j.agrformet.2020.108018>
- 536 Li, X., Xiao, J., He, B., Altaf Arain, M., Beringer, J., Desai, A. R., Emmel, C., Hollinger, D. Y.,  
 537 Krasnova, A., & Mammarella, I. (2018a). Solar - induced chlorophyll fluorescence is  
 538 strongly correlated with terrestrial photosynthesis for a wide variety of biomes: First global  
 539 analysis based on OCO - 2 and flux tower observations. *Global Change Biology*, 24(9),  
 540 3990-4008. <https://doi.org/10.1111/gcb.14297>
- 541 Li, X., Xiao, J., & He, B. (2018b). Higher absorbed solar radiation partly offset the negative  
 542 effects of water stress on the photosynthesis of Amazon forests during the 2015 drought.  
 543 *Environmental Research Letters*, 13, 044005. <https://doi.org/10.1088/1748-9326/aab0b1>
- 544 Linscheid, N., Estupinan-Suarez, L. M., Brenning, A., Carvalhais, N., Cremer, F., Gans, F.,  
 545 Rammig, A., Reichstein, M., Sierra, C., & Mahecha, M. D. (2020). Towards a global  
 546 understanding of vegetation-climate dynamics at multiple timescales. *Biogeosciences*, 17(4),  
 547 945-962. <https://doi.org/10.5194/bg-17-945-2020>
- 548 Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions, Proc.  
 549 Adv. Neural Inf. Process. Syst., pp. 4768-4777.
- 550 Madani, N., Kimball, J. S., Jones, L. A., Parazoo, N. C., & Guan, K. (2017). Global analysis of  
 551 bioclimatic controls on ecosystem productivity using satellite observations of solar-induced  
 552 chlorophyll fluorescence. *Remote Sensing*, 9(6), 530. <https://doi.org/10.3390/rs9060530>
- 553 Martens, B., Gonzalez Miralles, D., Lievens, H., Van Der Schalie, R., De Jeu, R. A., Fernández-  
 554 Prieto, D., Beck, H. E., Dorigo, W., & Verhoest, N. (2017). GLEAM v3: Satellite-based land  
 555 evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5), 1903-  
 556 1925. <https://doi.org/10.5194/gmd-10-1903-2017>
- 557 Matheny, A. M., Bohrer, G., Garrity, S. R., Morin, T. H., Howard, C. J., & Vogel, C. S. (2015).  
 558 Observations of stem water storage in trees of opposing hydraulic strategies. *Ecosphere*, 6(9),  
 559 1-13. <https://doi.org/10.1890/ES15-00170.1>

- 560 Migliavacca, M., Meroni, M., Manca, G., Matteucci, G., Montagnani, L., Grassi, G., ... &  
 561 Seufert, G. (2009). Seasonal and interannual patterns of carbon and water fluxes of a poplar  
 562 plantation under peculiar eco-climatic conditions. *Agricultural and Forest Meteorology*,  
 563 *149*(9), 1460-1476. <https://doi.org/10.1016/j.agrformet.2009.04.003>
- 564 Monteith, J. (1973). Principles of environmental physics. *Edward Arnold, London*.
- 565 Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., Myneni,  
 566 R. B., & Running, S. W. (2003). Climate-driven increases in global terrestrial net primary  
 567 production from 1982 to 1999. *Science*, *300*(5625), 1560-1563.  
 568 <https://doi.org/10.1126/science.1082750>
- 569 Nippert, J. B., & Holdo, R. M. (2015). Challenging the maximum rooting depth paradigm in  
 570 grasslands and savannas. *Functional Ecology*, *29*(6), 739-745. [https://doi.org/10.1111/1365-](https://doi.org/10.1111/1365-2435.12390)  
 571 [2435.12390](https://doi.org/10.1111/1365-2435.12390)
- 572 Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., Papuga, S.  
 573 A., Blanken, P. D., Noormets, A., Sulman, B. N., Scott, R. L., Wang, L., & Phillips, R. P.  
 574 (2016). The increasing importance of atmospheric demand for ecosystem water and carbon  
 575 fluxes. *Nature Climate Change*, *6*(11), 1023-1027. <https://doi.org/10.1038/nclimate3114>
- 576 Oliphant, T. E. (2006). A guide to NumPy. *Trelgol Publishing USA*, Vol. 1.
- 577 O and Orth, Global soil moisture from in situ measurements using machine learning - SoMo.ml,  
 578 Earth Syst. Sci. Data Discuss. (submitted), 2020
- 579 Orth, R., & Destouni, G. (2018). Drought reduces blue-water fluxes more strongly than green-  
 580 water fluxes in Europe. *Nature Communications*, *9*, 3602. [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-018-06013-7)  
 581 [018-06013-7](https://doi.org/10.1038/s41467-018-06013-7)
- 582 Pearson, R. G., Phillips, S. J., Loranty, M. M., Beck, P. S., Damoulas, T., Knight, S. J., & Goetz,  
 583 S. J. (2013). Shifts in Arctic vegetation and associated feedbacks under climate change.  
 584 *Nature Climate Change*, *3*(7), 673-677. <https://doi.org/10.1038/nclimate1858>
- 585 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M.,  
 586 Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python.  
 587 *the Journal of Machine Learning Research*, *12*, 2825-2830.
- 588 Piao, S., Wang, X., Wang, K., Li, X., Bastos, A., Canadell, J. G., Ciais, P., Friedlingstein, P., &  
 589 Sitch, S. (2020). Interannual variation of terrestrial carbon cycle: Issues and perspectives.  
 590 *Global Change Biology*, *26*(1), 300-318. <https://doi.org/10.1111/gcb.14884>

- 591 Reichle, R. H., Koster, R. D., De Lannoy, G. J., Forman, B. A., Liu, Q., Mahanama, S. P., &  
592 Touré, A. (2011). Assessment and enhancement of MERRA land surface hydrology  
593 estimates. *Journal of climate*, 24(24), 6322-6338. [https://doi.org/10.1175/JCLI-D-10-](https://doi.org/10.1175/JCLI-D-10-05033.1)  
594 [05033.1](https://doi.org/10.1175/JCLI-D-10-05033.1)
- 595 Sakschewski, B., Von Bloh, W., Boit, A., Poorter, L., Peña-Claros, M., Heinke, J., Joshi, J., &  
596 Thonicke, K. (2016). Resilience of Amazon forests emerges from plant trait diversity.  
597 *Nature Climate Change*, 6(11), 1032-1036. <https://doi.org/10.1038/nclimate3109>
- 598 Schlaepfer, D. R., Bradford, J. B., Lauenroth, W. K., Munson, S. M., Tietjen, B., Hall, S. A.,  
599 Wilson, S. D., Duniway, M. C., Jia, G., & Pyke, D. A. (2017). Climate change reduces  
600 extent of temperate drylands and intensifies drought in deep soils. *Nature Communications*,  
601 8, 14196. <https://doi.org/10.1038/ncomms14196>
- 602 Schulze, E.-D., Mooney, H. A., Sala, O., Jobbagy, E., Buchmann, N., Bauer, G., Canadell, J.,  
603 Jackson, R., Loreti, J., & Oesterheld, M. (1996). Rooting depth, water availability, and  
604 vegetation cover along an aridity gradient in Patagonia. *Oecologia*, 108(3), 503-511.  
605 <https://doi.org/10.1007/BF00333727>
- 606 Seddon, A. W., Macias-Fauria, M., Long, P. R., Benz, D., & Willis, K. J. (2016). Sensitivity of  
607 global terrestrial ecosystems to climate variability. *Nature*, 531(7593), 229-232.  
608 <https://doi.org/10.1038/nature16986>
- 609 Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., &  
610 Teuling, A. J. (2010). Investigating soil moisture-climate interactions in a changing climate:  
611 A review. *Earth Science Reviews*, 99(3-4), 125-161.  
612 <https://doi.org/10.1016/j.earscirev.2010.02.004>
- 613 Skipper, S. and Perktold, J. (2010). Sstatsmodels: Econometric and statistical modeling with  
614 python. Proceedings of the 9th Python in Science Conference.
- 615 Song, X. P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., &  
616 Townshend, J. R. (2018). Global land change from 1982 to 2016. *Nature*, 560(7720), 639-  
617 643. <https://doi.org/10.1038/s41586-018-0411-9>
- 618 Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Peñuelas, J., & Seneviratne, S. I.  
619 (2018). Quantifying soil moisture impacts on light use efficiency across biomes. *New*  
620 *Phytologist*, 218(4), 1430-1449. <https://doi.org/10.1111/nph.15123>

- 621 Stocker, B. D., Wang, H., Smith, N. G., Harrison, S. P., Keenan, T. F., Sandoval, D., Davis, T.,  
 622 & Prentice, I. C. (2020). P-model v1.0: an optimality-based light use efficiency model for  
 623 simulating ecosystem gross primary production. *Geoscientific Model Development*, 13(3),  
 624 1545-1581. <https://doi.org/10.5194/gmd-13-1545-2020>
- 625 Sun, Y., Fu, R., Dickinson, R., Joiner, J., Frankenberg, C., Gu, L., Xia, Y., & Fernando, N.  
 626 (2015). Drought onset mechanisms revealed by satellite solar - induced chlorophyll  
 627 fluorescence: Insights from two contrasting extreme events. *Journal of Geophysical*  
 628 *Research: Biogeosciences*, 120(11), 2427-2440. <https://doi.org/10.1002/2015JG003150>
- 629 Sundararajan, M., & Najmi, A. (2019). The many Shapley values for model explanation. *arXiv*  
 630 *preprint, arXiv:1908.08474*
- 631 Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a  
 632 potential reference dataset for hydrological modelling over North America. *Hydrology and*  
 633 *Earth System Sciences*, 24(5), 2527-2544. <https://doi.org/10.5194/hess-24-2527-2020>
- 634 Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein,  
 635 M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S.,  
 636 & Papale, D. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET  
 637 sites with regression algorithms. *Biogeosciences*, 13(14), 4291-4313.  
 638 <https://doi.org/10.5194/bg-13-4291-2016>
- 639 Wagle, P., Zhang, Y., Jin, C., & Xiao, X. (2016). Comparison of solar - induced chlorophyll  
 640 fluorescence, light - use efficiency, and process - based GPP models in maize. *Ecological*  
 641 *Applications*, 26(4), 1211-1222. <https://doi.org/10.1890/15-1434>
- 642 Walther, S., Duveiller, G., Jung, M., Guanter, L., Cescatti, A., & Camps - Valls, G. (2019).  
 643 Satellite Observations of the Contrasting Response of Trees and Grasses to Variations in  
 644 Water Availability. *Geophysical Research Letters*, 46(3), 1429-1440.  
 645 <https://doi.org/10.1029/2018gl080535>
- 646 Walther, S., Guanter, L., Heim, B., Jung, M., Duveiller, G., Wolanin, A., & Sachs, T. (2018).  
 647 Assessing the dynamics of vegetation productivity in circumpolar regions with different  
 648 satellite indicators of greenness and photosynthesis. *Biogeosciences*, 15(20), 6221-6256.  
 649 <https://doi.org/10.5194/bg-15-6221-2018>
- 650 Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Kohler, P., Jung, M., Varlagin, A.,  
 651 & Guanter, L. (2016). Satellite chlorophyll fluorescence measurements reveal large-scale

- 652 decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. *Global*  
 653 *Change Biology*, 22(9), 2979-2996. <https://doi.org/10.1111/gcb.13200>
- 654 Wu, J., Albert, L. P., Lopes, A. P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K. T., Guan,  
 655 K., Stark, S. C., Christoffersen, B., & Prohaska, N. (2016). Leaf development and  
 656 demography explain photosynthetic seasonality in Amazon evergreen forests. *Science*,  
 657 351(6276), 972-976. <https://doi.org/10.1126/science.aad5068>
- 658 Lian, X., Piao, S., Li, L. Z., Li, Y., Huntingford, C., Ciais, P., Cescatti, A., Janssens, I.A.,  
 659 Peñuelas, J., Buermann, W., Chen, A., Myneni, R. B., Wang, X., Wang, Y., Yang, Y., Zeng,  
 660 Z., Zhang, Y. & Mcvicar, T. R. (2020). Summer soil drying exacerbated by earlier spring  
 661 greening of northern vegetation. *Science Advances*, 6(1), eaax0255.  
 662 <https://doi.org/10.1126/sciadv.aax0255>
- 663 Yan, H., Wang, S. Q., Huete, A., & Shugart, H. H. (2019). Effects of light component and water  
 664 stress on photosynthesis of Amazon rainforests during the 2015/2016 El Niño drought.  
 665 *Journal of Geophysical Research: Biogeosciences*, 124(6), 1574-1590.  
 666 <https://doi.org/10.1029/2018JG004988>
- 667 Yang, X., Tang, J., Mustard, J. F., Lee, J. E., Rossini, M., Joiner, J., Munger, J. W., Kornfeld, A.,  
 668 & Richardson, A. D. (2015). Solar - induced chlorophyll fluorescence that correlates with  
 669 canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest.  
 670 *Geophysical Research Letters*, 42(8), 2977-2987. <https://doi.org/10.1002/2015GL063201>
- 671 Zarco-Tejada, P. J., Morales, A., Testi, L., & Villalobos, F. J. (2013). Spatio-temporal patterns  
 672 of chlorophyll fluorescence and physiological and structural indices acquired from  
 673 hyperspectral imagery as compared with carbon fluxes measured with eddy covariance.  
 674 *Remote Sensing of Environment*, 133, 102-115. <https://doi.org/10.1016/j.rse.2013.02.003>
- 675 Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H., Vaze, J., Liu, C., Lu, X., Zheng,  
 676 H., Wang, Y., & Liu, Y. Y. (2016). Multi-decadal trends in global terrestrial  
 677 evapotranspiration and its components. *Scientific Reports*, 6, 19124.  
 678 <https://doi.org/10.1038/srep19124>
- 679 Zuromski, L. M., Bowling, D. R., Köhler, P., Frankenberg, C., Goulden, M. L., Blanken, P. D.,  
 680 & Lin, J. C. (2018). Solar - induced fluorescence detects interannual variation in gross  
 681 primary production of coniferous forests in the Western United States. *Geophysical*  
 682 *Research Letters*, 45(14), 7184-7193. <https://doi.org/10.1029/2018GL077906>

683 Zwillinger, D. and Kokoska, S. (2000). Standard Probability and Statistics Tables and Formulae.  
684 *Florida: Chapman & Hall/CRC.*  
685