Critical Soil Moisture Derived From Satellite Observations Over Europe

Jasper M.C. Denissen[1,2] (iD), **Adriaan J. Teuling[2]** (iD), **Markus Reichstein[1]** (iD), and **René Orth[1]** (iD)

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany, [2]Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, Netherlands

Abstract Evapotranspiration (ET) is a crucial quantity through which land surface conditions can impact near-surface weather and vice versa. ET can be limited by energy or water availability. The transition between water- and energy-limited regimes is marked by the critical soil moisture (CSM), which is traditionally derived from small-sample laboratory analyses. Here, we aim to determine the CSM at a larger spatial scale relevant for climate modeling, using state-of-the-art gridded data sets. For this purpose, we introduce a new correlation-difference metric with which the CSM can be accurately inferred using multiple data streams. We perform such an analysis at the continental scale and determine a large-scale CSM as an emergent property. In addition, we determine small-scale CSMs at the grid cell scale and find substantial spatial variability. Consistently from both analyses we find that soil texture, climate conditions, and vegetation characteristics are influencing the CSM, with similar respective importance. In contrast, comparable CSMs are found when applying alternative large-scale energy and vegetation data sets, highlighting the robustness of our results. Based on our findings, the state of the vegetation and corresponding land-atmosphere coupling can be inferred, to first order, from easily accessible satellite observations of surface soil moisture.

Plain Language Summary Water evaporates at the Earth's surface when there is enough soil moisture and incoming radiation. Most water evaporates from plants during photosynthesis. In water limitation, the soil is too dry to support evaporation. When the soil gets wetter, evaporation will increase. At a certain soil moisture, the maximum evaporation has been reached, so that further increases in soil moisture do not further increase the evaporation: this is the critical soil moisture (CSM). When the soil gets wetter than the CSM, evaporation is controlled by the amount of incoming radiation (energy limitation). Incoming radiation is used to i) heat up the air above the surface and ii) to evaporate water. During water limitation, all available energy is used to heat the air. This mechanism potentially intensifies heat waves. Knowing the soil moisture respective to the CSM can tell us something about the intensity of future heat waves. In this study we determine this CSM in Europe using satellite data. With a new method we can see at which soil moistures evaporation is controlled by i) soil moisture, ii) incoming radiation or iii) neither, which denotes the CSM. The CSM helps to improve climate models and consequently heat wave forecasts.

1. Introduction

Evapotranspiration (hereafter referred to as ET) is a crucial variable in land-atmosphere interactions, since it affects the carbon, energy, and water balances. Therefore, ET can potentially impact weather and climate, especially during extreme events such as droughts and heat waves (Guillod et al., 2015; Hirschi et al., 2011; Koster et al., 2016, 2004; Taylor et al., 2012). Conceptually, we distinguish two evaporative regimes: (i) the water-limited regime, where ET is mainly controlled by soil moisture availability, and (ii) the energy-limited regime, where ET is mostly governed by energy (temperature and radiation) supply (Budyko, 1974; Seneviratne et al., 2010). Consequently, regime shifts potentially induce changes in the causality of energy and water availability for ET. This could dampen or amplify land-atmosphere interactions, like evaporative cooling (Seneviratne et al., 2010). Therefore, the critical soil moisture (hereafter referred to as CSM) associated with this regime shift in the conceptual framework is a crucial parameter.

Traditionally, specific CSM values have been associated with characteristic matric potentials for particular vegetation and soil types (Van Genuchten, 1987). By doing so, the determination of CSMs is straightforward in areas of homogenous soils and vegetation (Homaee et al., 2002). However, the range of CSMs across soil

and vegetation types is substantial: Teuling, Uijlenhoet, et al. (2009) report CSMs within one model ranging from 16.7 Vol-% for sandy soils to 42.4 Vol-% for clayey soils based on pedotransfer functions. Additionally, there is considerable variation between CSMs of different models. Novák and Havrila (2006) determine somewhat different CSMs ranging from 2.7 Vol-% in a sandy soil to 13 Vol-% in a loamy soil, which they refer to as the critical soil water content. These different values illustrate that even at smaller scales, there are discrepancies between CSMs determined with various methods. Further, the uncertainty of the CSM, among other soil hydraulic parameters, is enhanced by different pedotransfer functions and different soil texture data sets (Van Looy et al., 2017). In addition, the dependency of the CSM on local soil and vegetation conditions renders it difficult to derive large-scale estimates from previous analyses and literature.

Besides its conceptual relevance, the CSM is an important parameter and/or emergent property in land models, which are embedded into weather and climate (forecasting) models that serve society. Land models have an inherent assumption of the above-mentioned water- and energy-controlled evaporative regimes and of the transition between them, as marked by the CSM (Arora, 2002; Pitman, 2003; Sellers et al., 1997). As the determination of a large-scale CSM is lacking, considerable difference exists between current model estimations of the CSM (Teuling, Uijlenhoet, et al., 2009), leading to inconsistent simulation results. Additionally, comparison of CSMs between models is not straightforward, as absolute values of soil moisture are model dependent and do not necessarily correspond with observed soil moisture (Koster et al., 2009). Further, as these models operate at relatively large spatial scales, the simulation of the evaporative regimes is hard to validate and a source of considerable uncertainty (Guillod et al., 2013).

Large-scale assessments of evaporative regimes have been performed previously based on various metrics with both observational and modeled data sets; where in some of these studies the determination of the CSM is lacking (Koster et al., 2009; Seneviratne et al., 2006; Teuling, Hirschi, et al., 2009; Zscheischler et al., 2015), other analyses have determined CSMs based on modeled data sets, which reflect implemented relationships between soil moisture (SM) and evaporative fraction (EF; Schwingshackl et al., 2017). There are only a few recent studies that determine observation-based CSMs at the regional-continental scale. Such large-scale analyses have only recently become feasible thanks to the increasing availability of satellite-derived data sets (e.g., Liu et al., 2012; Tramontana et al., 2016). For example, Feldman et al. (2019) estimate the CSM over Africa by assuming a piecewise linear model based on satellite observations of surface soil moisture and diurnal temperature amplitude. Haghighi et al. (2018) determine the CSM in a similar manner but using field observations of SM and EF over semiarid regions outside of the growing season, effectively excluding the effects of plant transpiration from their estimates. Finally, Akbar et al. (2018) determine the CSM over the contiguous United States by assessing the characteristics of dry-downs from satellite surface soil moisture from the National Aeronautics and Space Administration Soil Moisture Active Passive mission during three consecutive summers only.

In this study, we focus on determining regional-continental-scale CSMs from observational data in Europe, as regime transitions are known to occur frequently in this region (Seneviratne et al., 2006; Teuling, Hirschi, et al., 2009). Moving beyond the previous studies, we propose a novel correlation-difference metric to characterize the CSM. This metric uses data on energy and water availability, as well as vegetation functioning, and thereby determines the CSM based on comprehensive Earth observations. We estimate the CSM at different spatial scales: (1) a continental-scale estimate will serve as observational constraint for land surface models (large-scale CSM), while (2) small-scale grid cell estimates will reflect the spatial heterogeneity of the CSM (small-scale CSM). Further, we investigate the sensitivity of the CSM to climate, soil, and vegetation characteristics, and its robustness when determined with different data sets.

2. Data and Methods

We propose a novel metric to evaluate water- versus energy-limited conditions in each grid cell:

$$\Delta corr = corr(A_E, A_V) - corr(A_W, A_V) \qquad (1)$$

where A indicates bimonthly (twice per month, concerning the first and second half of the month) anomalies of particular energy (E), vegetation (V), or water (W) variables, and $corr$ denotes a temporal correlation between anomaly time series. The default $\Delta corr$ metric from equation 1 is calculated using surface soil moisture (from the European Space Agency [ESA] Climate Change Initiative [CCI] program, ESA CCI), surface temperature (from E-OBS), and ET (from FLUXCOM). We note that FLUXCOM ET is not an

Table 1
Data Sets Used in This Study

Data set	Version	Variable(s)	Spatial res.	Temp. res.	Derived from	Reference
ESA CCI	04. 4	Surface soil moisture$_W$	0.5 ° × 0.5°	Daily	Satellite observations	Dorigo et al. (2017), Gruber et al. (2017), and Liu et al. (2012)
E-OBS	17.0	Surface temperature$_E$ and precipitation	0.25 ° × 0.25°	Daily	Station observations	Haylock et al. (2008)
FLUXCOM (RS)	2	Evapotranspiration (ET)$_V$ and gross primary production (GPP)$_V$	0.25 ° × 0.25°	8-daily	Multiple datastreams processed with machine learning techniques, using eddy covariance-derived ET observations as target	Tramontana et al. (2016)
ERA-interim	2	Incoming shortwave radiation$_E$, net shortwave radiation, net longwave radiation, and vapor pressure deficit$_E$	0.5 ° × 0.5°	Daily	Reanalysis model	Dee et al. (2011)
GIMMS3g	4	normalized difference vegetation index (NDVI)$_V$	0.025 ° × 0.025°	Bimonthly	Satellite observations	Zhu et al. (2013)
GOME-2	GFZ	Sun-induced chlorophyll fluorescence (SIF)$_V$	0.5 ° × 0.5°	8-daily	Satellite observations	Köhler et al. (2015)
SoilGrids	0.5.1	Fraction clay, sand, and silt	0.1 ° × 0.1°	—	Satellite observations	Hengl et al. (2017)
MCD12Q1	005	Fraction land cover	500 m × 500 m	—	Satellite observations	Friedl et al. (2010)

Note. Variables are used in equation 1 are denoted with an *E* (energy), *V* (vegetation), and *W* (water). Next to this, the E-OBS data set, SoilGrids data, and the MODIS land cover product MCD12Q1 are used to specify climate, soil, and vegetation types in Europe, respectively.

observational product but derived from multiple data streams using machine learning techniques. This data set is chosen because, unlike process-based models, it does not involve any assumption or implementation of a SM-EF relationship. All time series are linearly detrended, before anomalies are computed by subtracting the mean seasonal cycle. To assure no confounding impacts of nonlinearities, Kendall's rank correlations are computed. $\Delta corr > 0$ indicates that vegetation anomalies correlate stronger with energy than with water anomalies, such that the grid cell would be referred to as energy limited. Correspondingly, $\Delta corr < 0$ indicates that a grid cell is water limited. When $\Delta corr \approx 0$, the magnitudes of energy and water limitations are equal, thus, corresponding to frequent regime shifts and marking the related CSM. Therefore, this metric enables a simple and straightforward determination of the CSM.

All data sets employed in this study are listed in Table 1. All energy, vegetation, and water variables are aggregated to a common 0.5 ° × 0.5° spatial resolution. Thereafter, bimonthly averages are calculated, to mitigate the effect of synoptic weather variability on our analyses and because at this timescale the response of ET to soil moisture is the strongest (Boese et al., 2019; Teuling et al., 2006). A bimonthly average is only calculated when at least 5 days per 2-week period are available, to account for gaps in the ESA CCI SM data set due to, for example, snowy or extremely dry soil. Given the required concurrent availability of data sets, we consider the time period 2007–2015, which meets the requirements for the minimum of 4–6 years of data recommended for calculating land-atmosphere interactions metrics as in equation 1 (Findell et al., 2015).

We focus on the warm season in this study to exclude the impact of ice and snow and to focus on active vegetation functioning. In this context, data will be considered only if the bimonthly temperature exceeds 10 ° C. This can lead to a different number of bimonths filtered in different grid cells. Correlations (equation 1) are calculated per grid cell, and per season. Using all available bimonths from a particular season across all years ensures a meaningful amount of data points. No seasonal correlation is computed if less than six data points are available.

3. Results and Discussion

Analyzing in a first step the summer (June–August) mean surface soil moisture we find a north-south gradient across Europe (Figure 1a). Apart from this general pattern, soil moisture in panel (a) tends to be higher
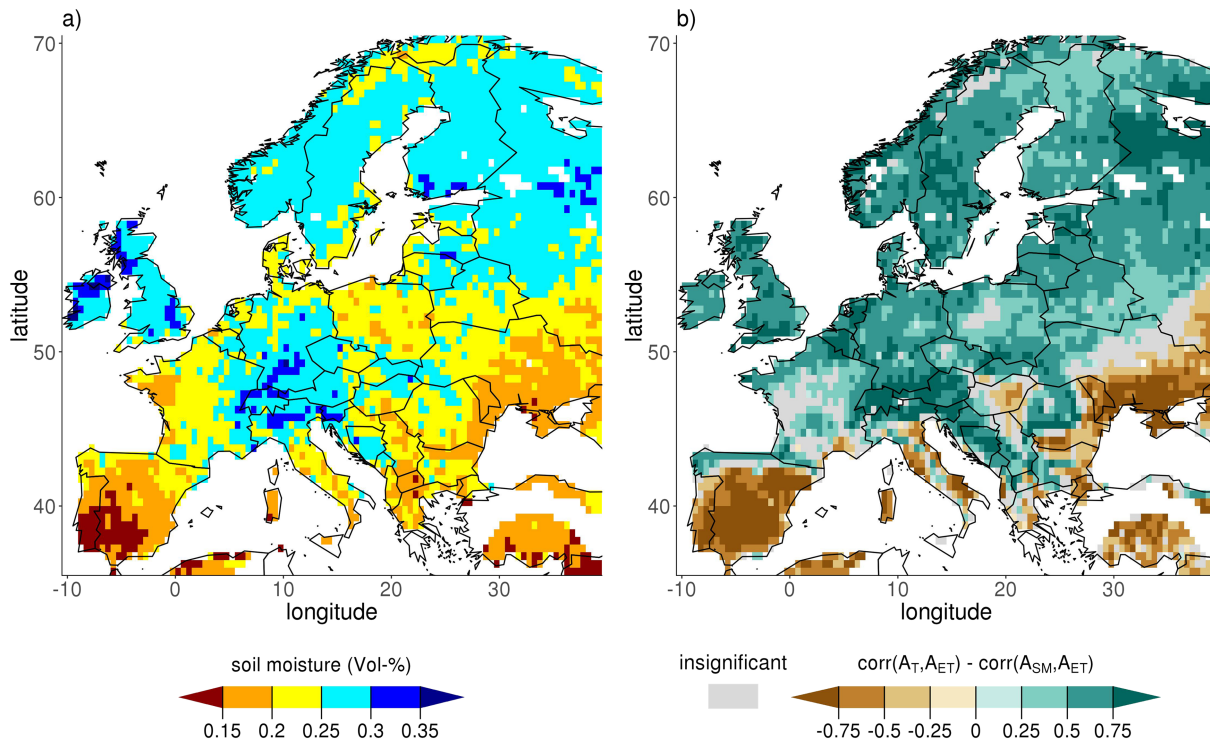
Figure 1. Distribution of surface soil moisture and evapotranspiration regimes in Europe. (a) Spatial variability of average surface soil moisture from the ESA CCI data set and (b) Δ*corr* in the summer (JJA), computed according to equation 1 with soil moisture from ESA CCI, FLUXCOM ET, and E-OBS surface temperature (see also Table 1). The Δ*corr* is only calculated if it is significant, that is, within the 90% confidence interval.
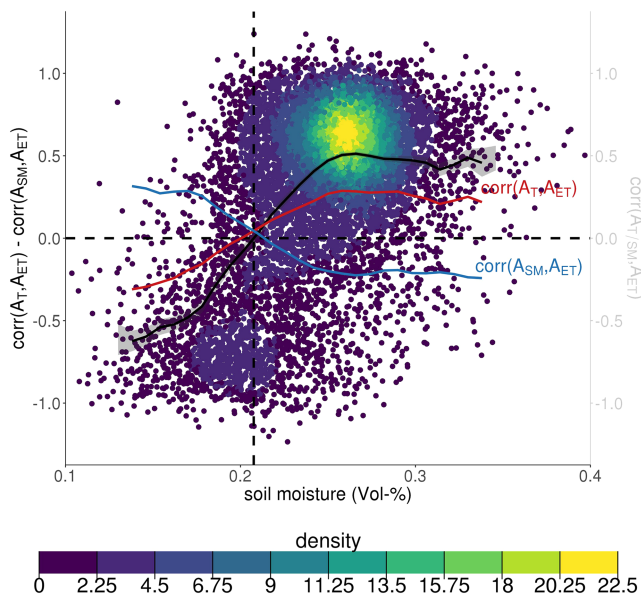

Figure 2. Sensitivity of Δ*corr* to soil moisture across all grid cells in Europe. Each dot represents a particular grid cell and season. The black line is a moving average through the points, while the gray ribbon indicates the 5% and 95% percentiles of moving averages through 1,000 bootstrapped samples of the data points. In addition, moving averages are also displayed for individual correlations from equation 1 where the individual points are not shown. The color imposed on the data points reflect the density of data points.

in mountainous regions such as the Alps or the Carpathian Mountains. As expected, negative Δ*corr* in panel (b), indicating water-limited conditions, generally coincide with lower soil moisture. Correspondingly, energy-limited conditions (positive Δ*corr*) occur in regions where ample soil moisture is available. Insignificant Δ*corr* values occur in northern Scandinavia due to a lack of available soil moisture data related to low surface temperatures and in between water- and energy-limited regions across central Europe, marking the transitional regions. Next to these spatial variations, Figure S1 in the supporting information shows seasonal variation of soil moisture in panel (a)–(c) and of Δ*corr* in panels (d)–(f). Winter is not shown, because there are hardly any significant Δ*corr* values. Generally, water-limited conditions coincide with dry soils and energy-limited conditions coincide with wet soils. From springtime to summertime, conditions shift from energy limited to transitional in central Europe and parts of the Mediterranean, likely due to a decrease in soil moisture content. In autumn, water-limited conditions persist in the Mediterranean. This possibly reflects that, while surface soil moisture in the Mediterranean is already replenished (panel c), the root zone, where vegetation extracts the majority of its moisture, is still dry. This derived spatial pattern of Δ*corr* is an important result as it is based solely on (satellite-)observable variables and can hence serve as a benchmark for models, which mostly simulate these variables.

In a next step, we analyze the relation between soil moisture and Δ*corr*, as depicted in Figure 2. Each point in the scatterplot depicts soil moisture and Δ*corr* at a given grid cell and in a given season (basically Figure S1 soil moisture in (a)–(c) plotted against the Δ*corr* in (d)–(f)) and the coloring

indicates the density of the data points. The red and blue moving average lines indicate the governing processes in the respective evaporative regimes: When the soil moisture content is low, ET is water limited, resulting in $corr(A_T, A_{ET}) < 0$ and $corr(A_{SM}, A_{ET}) > 0$. At wetter soil moisture contents, ET is governed by energy supply, resulting in $corr(A_T, A_{ET}) > 0$ and $corr(A_{SM}, A_{ET}) < 0$. The negative $corr(A_{SM}, A_{ET})$ at higher soil moisture contents might be related to a confounding, negative correlation between surface temperature and soil moisture: A wet soil moisture anomaly might result from a precipitation surplus, which tends to occur jointly with a negative surface temperature anomaly.

The difference between the moving averages of the individual correlations yields the moving average of $\Delta corr$ (thick black line), of which negative values indicate water-limited conditions and positive values indicate energy-limited conditions. These fitted lines are likely not representing the behavior at every single grid cell but depict the general relationship. But this illustrates the key advantage of the $\Delta corr$ metric over CSM estimation using actual and potential ET (Seneviratne et al., 2010) or the relationship between SM-EF (Haghighi et al., 2018; Schwingshackl et al., 2017): The CSM can be simply inferred from where the moving average switches sign, without applying piecewise linear models with potentially poor fits. In Figure 2 we derive a large-scale CSM at approximately 21 Vol-%. This value entails temporal and spatial variability, from different seasons and grid cells, respectively. The ribbon around the moving average line reflects the uncertainty in the moving average and illustrates the 5% and 95% percentiles of moving averages based on 1,000 bootstrapped samples from the data points. This uncertainty is relatively small thanks to the large amount of data used. The ribbon is narrowest between approximately 20 and 30 Vol-%, as the majority of the data points have soil moisture contents in that range, as can be seen by the density of the points. Further, we test the potential role of confounding effects for our analysis using partial correlations in Figure S2. Accounting for the confounding effect of soil moisture on the correlation between temperature and ET, as well as the confounding effects of temperature on the correlation between soil moisture and ET, we find very similar results as in Figure 2. This suggests that confounding effects do not significantly influence the individual correlations that form $\Delta corr$.

Note that even though we employ observation-based soil moisture in our analysis, the derived large-scale CSM of 21 Vol-% is somewhat model based. This is because the values of the ESA CCI soil moisture are derived by scaling the satellite-observed temporal dynamics against modeled data (Dorigo et al., 2017). Therefore, only analyses using the same soil moisture product can make use of our absolute derived CSM values, while all other studies should rather use it in a relative sense: 21 Vol-% is drier than 85% of the European grid cell seasonal mean soil moistures. Another study based on satellite-derived surface soil moisture from National Aeronautics and Space Administration's Soil Moisture Active Passive mission reports a median of CSMs of 18 Vol-% over the contiguous United States (Akbar et al., 2018). This result, as well as our estimate, is more to the dry side than currently assumed in models, which often become water limited just below or at the field capacity (Teuling, Uijlenhoet, et al., 2009). Therefore, the large-scale CSM is an emergent property of the European land climate system and thus can be used as a continental reference CSM.

It is known that soil, climate, and vegetation characteristics can locally influence the CSM (Feldman et al., 2019; Haghighi et al., 2018; Novák & Havrila, 2006). To investigate this on a large scale, moving average lines based on subselections of data representing particular soil, climate, and vegetation types are shown in Figure 3: (a) soil types have been determined using depth-weighted average soil texture fractions of the top meter from the SoilGrids data set (Hengl et al., 2017). Across all grid cells in Europe, the 75% quantile has been calculated for clay, silt, and sand fractions, respectively. Any grid cell exceeding this respective threshold is classified as clay, sand, or silt, leaving a mixed soil class for the remaining grid cells, in Figure 3; (b) for the climate types, grid cells are classified according to their long-term average surface temperature; and (c) vegetation types have been derived from the Moderate Resolution Imaging Spectroradiometer land cover data set MCD12Q1 (Friedl et al., 2010). The forest class in Figure 3 comprises evergreen/deciduous broadleaf/needleleaf and mixed forest categories. Low vegetation includes closed/open shrublands, (woody) savannas, and grasslands. And crop consists of all cropland land cover types. A grid cell is classified as forest, low vegetation, or crop if the respectively considered land cover fractions exceed the European 75% quantile of the respective vegetation type, leaving a mixed vegetation class for the remaining grid cells. Figure 3 shows that the large-scale CSM varies by few Vol-% in response to different soil textures, climate conditions, and vegetation classes. These tested characteristics seem to have comparable little influence on the CSM. They might be interdependent with, for example, colder surface temperatures predominantly coinciding with
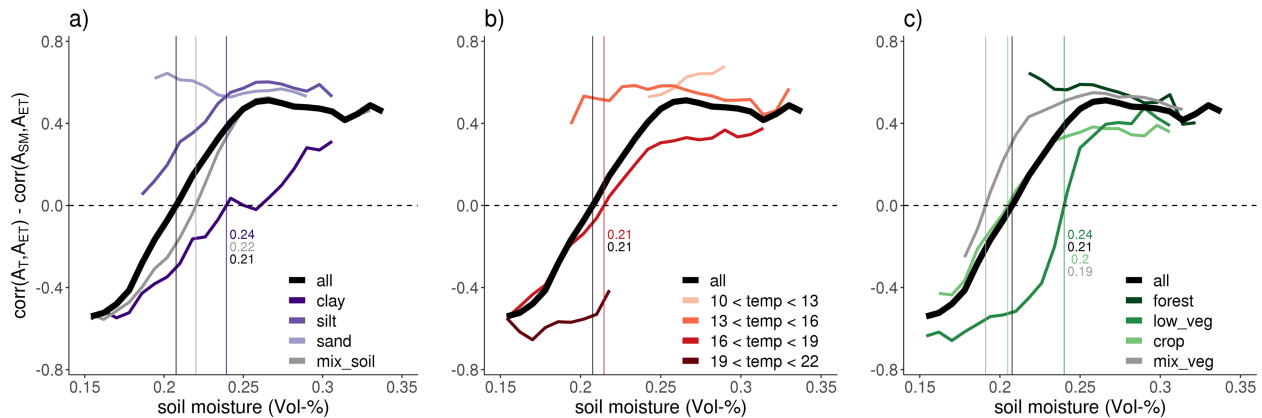
Figure 3. Soil, climate, and vegetation effects on the large-scale mean CSM. Moving average lines are calculated with subgroups of the data, based on, from left to right, different soil textures, climate classifications, and vegetation types. The thick black lines are the same as in Figure 2.

forest. As for the soil types, the CSM for clay is wetter than for all soil textures combined, which is expected because clay has a more negative matric potential than coarser soil textures with dominant sand and silt fractions, and is in line with earlier findings (Akbar et al., 2018; Feldman et al., 2019). The regions within the sand and silt classes appear permanently energy limited. Concerning climate types, interestingly, regime transition, is only observed for the second-warmest class. In contrast, colder climate regions are generally energy-limited and warmer climate regions water limited. Forested regions are not subject to regime transition either, as trees have deep-reaching roots which can access deep(er) water reservoirs to avoid water limitation. Correspondingly, low vegetation and crop with shorter root systems are more water limited, resulting in slightly different CSMs. In summary, in Europe the CSM tends to be slightly wetter for (i) finer soils, (ii) warmer surface temperatures, and (iii) shorter vegetation, which hampers advocation for a single, representative large-scale CSM.

In the previous analyses, $\Delta corr$ was calculated with anomalies from surface soil moisture (water), surface temperature (energy), and ET (vegetation). The largest part of ET is accounted for by plant transpiration (Good et al., 2015; Lawrence et al., 2007; Schlesinger & Jasechko, 2014), associated with photosynthetic activity. Sun-induced fluorescence (SIF), gross primary production, and normalized difference vegetation index (NDVI) are reflections of photosynthetic activity and can therefore be regarded as reasonable proxies for ET. Next to that, plants adjust their stomatal resistance in response to changes in atmospheric energy availability in the form of leaf temperature, incoming shortwave radiation or vapor pressure deficit. Figure 4 shows the corresponding results derived by substituting (a) surface temperature anomalies with
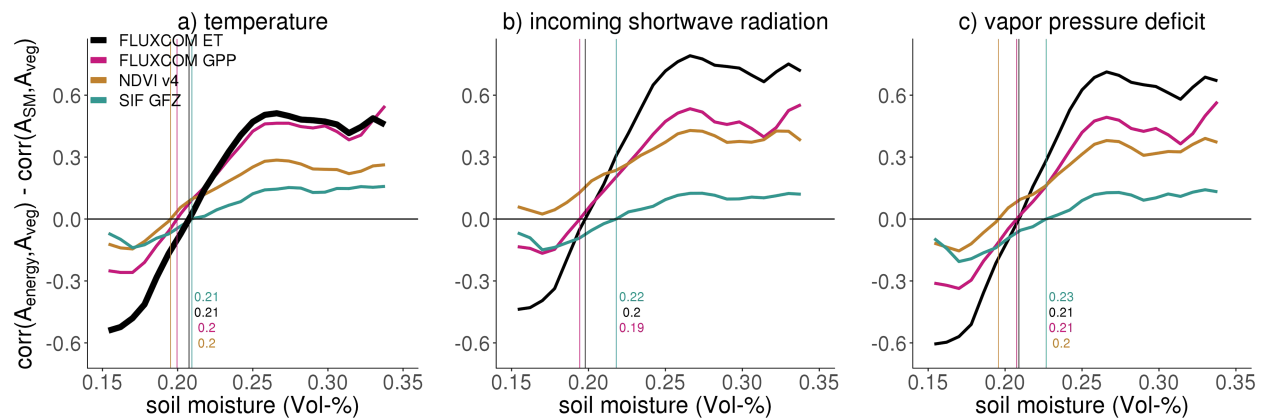

Figure 4. Robustness of the large-scale mean CSM across different employed energy and vegetation products. The different panels reflect $\Delta corr$ based on anomalies of different energy variables, from left to right: temperature (a), incoming shortwave radiation (b) and vapor pressure deficit (c). The different colored moving averages imply different vegetation data sets. As in Figures 2 and 3, the thick black line in the left panel (a) shows the reference moving average line, which is computed with bimonthly anomalies of ET and surface temperature.
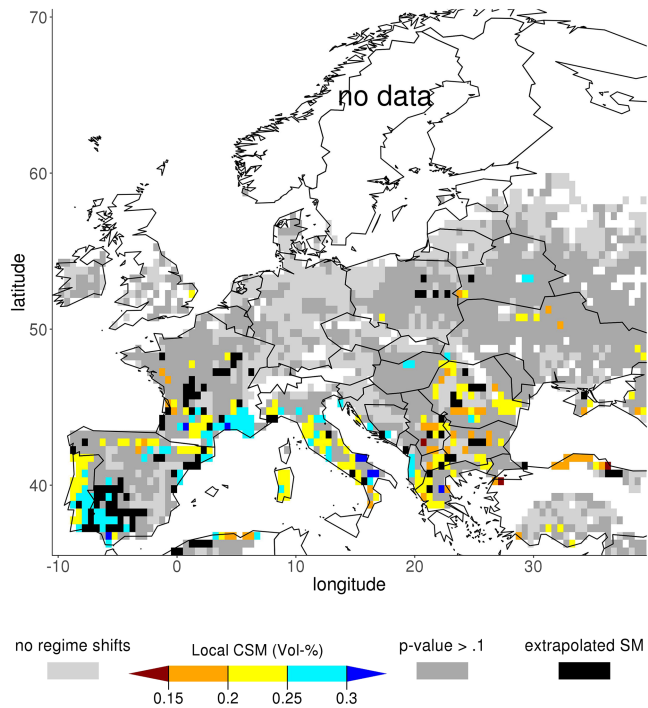
Figure 5. Estimation of the small-scale CSM based on linear regression between bimonth-of-year average soil moisture and Δ*corr*. Light gray grid cells denote regions where no regime shifts occurred. In dark gray regions regime shifts have occurred, but a *p*-value of the linear model exceeds 0.1. Black grid cells indicate that the value of the CSM falls outside of the local measured range of soil moistures. White regions indicate missing data.

(b) incoming shortwave radiation in or (c) vapor pressure deficit anomalies in (c). In addition, ET anomalies are replaced by NDVI, SIF, and gross primary production anomalies. Applying alternative data products yields similar CSMs, generally deviating only a few Vol-% from the previously obtained 21 Vol-%. This highlights the robustness of the Δ*corr* metric across various data products. Further, the choice of data products affects the magnitude of Δ*corr* denoting the strength of the energy or water limitation, respectively. Regardless of the applied energy variable, the strongest Δ*corr* signal is derived with the FLUXCOM ET data set. This can be explained as ET represents a flux, which is expected to respond quicker to changing water availability, yielding stronger Δ*corr* signals than one would expect with state variables. In contrast, NDVI as a state variable responds more slowly, yielding lower Δ*corr* amplitudes. Moreover, changing water use efficiency under dry conditions can affect ET results in Figure 4. The weakest Δ*corr* signals are obtained with SIF. This is surprising as SIF rather represent a flux as, for example, ET, and not a state. A reason for this could be the relatively early equator overpass time of the GOME-2 satellite, 10:00 local solar time (Köhler et al., 2015), which is when radiation and leaf temperature are usually not at their daily maxima such that the vegetation is not yet most active. Table S1 shows a systematic negative bias and less robust values across energy and vegetation products of large-scale CSMs estimated from $corr(A_{energy}, A_{veg})$ in comparison with CSMs estimated from Δ*corr*.

In a next step, we further explore small-scale, grid cell CSMs. Whereas the large-scale CSM is mostly inferred from soil moisture variations in space, the small-scale CSMs are estimated from bimonth-of-year soil moisture variations in time. This allows to study the effect of year-to-year variability of each available bimonth. For this purpose, we focus on grid cells that experience both water- and energy-controlled conditions, that is, where soil moisture crosses the CSM. This is achieved by selecting grid cells where Δ*corr* is negative and positive for at least one bimonth of year, respectively. In each of the grid cells where regime shifts occur, we fit a linear regression on bimonth-of-year soil moistures and corresponding Δ*corr* values. The small-scale CSM is then inferred from the regression line at Δ*corr* = 0. Several steps ensure a meaningful estimation of the small-scale CSM: (1) There should be at least 10 data points (bimonths of year) where soil moisture and Δ*corr* are available. (2) The slope of the linear model should be positive. (3) The *p*value of the linear model should not exceed 0.1 to ensure a reasonably strong linear relationship, given that there only 10–24 data points available per grid cell. (4)2The CSM needs to be within the range of observed soil moistures to ensure a physically possible CSM. An example of the local estimation of the CSM is shown in Figure S3. Figure 5 shows the spatial distribution of the small-scale CSMs. Most of the CSMs are determined in central and southern Europe within the range of 20–25 Vol-%, which is comparable to the previously determined CSMs in Figure 3. Next to that, there are ample grid cells where (i) data availability is insufficient (white grid cells), (ii) no regime shifts are occurring (light gray grid cells), (iii) the linear relationship is too weak (dark gray grid cells), or (iv) the estimated CSM falls outside of the range of observed soil moistures (black grid cells). A closer look at the distribution of the range of CSMs is given in Figure S4. We find similar soil, climate, and vegetation controls for the small-scale CSMs (Figure S5), confirming results from Figure 3.

4. Limitations

The Δ*corr* metric is computed using ESA CCI surface soil moisture, which is determined from satellite observations. These are based on microwaves that penetrate only into the upper few centimeters of the soil (Ulaby et al., 1982). The depth of the surface layer in this soil moisture product is not well defined, since this soil moisture product is a composite of multiple microwave sensors with different frequencies and hence slightly different penetration depths (Dorigo et al., 2017). It is not fully representative of the vegetation-accessible soil moisture. However, there is no root zone soil moisture data set with a spatial and temporal coverage

comparable to that of the ESA CCI data set. Assessing the potential effect of this shortcoming, we analyze reanalysis and station-based soil moisture from multiple depths and find that surface soil moisture is a reasonable proxy for root zone soil moisture, albeit with seasonal variations in their relationship (Figure S6). Similar results are reported for example by Hirschi et al. (2014) who find similar surface and root zone soil moisture in mean climatological conditions, or Qiu et al. (2016) who find differences only under extremely dry conditions. Some decoupling between surface soil moisture and root-zone also emerges in Figure S1, where similar surface soil moisture values coincide with different $\Delta corr$ values. This is also reflected in the seasonal variability of $\Delta corr$ results as shown in Figure S7. This pattern can be explained with seasonal discrepancies between surface and root-zone soil moisture: In springtime, root-zone soil moisture is generally still readily available, but the surface soil moisture is generally lower due increased bare soil evaporation. This means that ET could occur at its maximum rate, while lower surface soil moistures are registered by satellites, resulting in a lower springtime CSM. This discrepancy between surface and root-zone soil moisture causes the entire moving average to shift to the dry end, as can be seen in Figure S7. In autumn, the contrary is observed: As precipitation occurs more frequently after summer, first, the surface layer is moistened, but it takes time to replenish the moisture deficit in the root-zone, leading to a higher CSM (shift of the autumn moving average to the wet end). The seasonal variability in representativeness of surface soil moisture for the root-zone can possibly impact the linear model which is used for the estimation of the local CSM in Figure 5, as these linear models are based on maximally 24 data points. Further, surface satellite soil moisture estimates have larger measurement uncertainties on densely vegetated and/or organic soils (Dorigo et al., 2017; Ulaby et al., 1982).

While similar $\Delta corr$ values could in principle be derived with different combinations of the individual correlations, Figure S8 illustrates that $\Delta corr$ usually corresponds to unique combinations, making it unambiguous. Further, there is a significant large scatter across data points in Figure 2 and illustrated in Figure S8. There are several reasons for this underlying uncertainty, next to the limitations related to using satellite surface soil moisture: (1) Soil moisture is known to have profound memory characteristics (Orth & Seneviratne, 2012), such that legacy effects might play a role but are not considered here for simplicity, (2) confounding impacts of soil moisture on $corr(A_T, A_{ET})$, and of surface temperature on $corr(A_{SM}, A_{ET})$, are also not taken into account. The role of these confounding effects is investigated in Figure S2 using partial correlations. We find overall negligible impact of such effects on $\Delta corr$. Finally, (3) human influence on soil moisture and consequently vegetation through, for example, irrigation or land use changes, can introduce non-natural variability into our analysis. Fortunately, none of the limitations listed above affects all grid cells at the same time. Therefore, we are confident that our large-scale analysis with thousands of grid cells employs enough information to derive meaningful results despite the uncertainties introduced by the limitations.

5. Conclusions

In this study we build upon the conceptual frameworks of Budyko (1974) and Seneviratne et al. (2010). We introduce a novel metric to infer energy- or water-limited conditions, which does not rely on prior assumptions on the relationship between soil moisture and EF and uses observation-based data sets, which describe water and energy availability and vegetation functioning. This metric is applied to determine the CSM.

We derive a large-scale CSM representative for the European continent, as well as a range of small-scale CSMs representative locally at particular grid cells. Within the large-scale analysis we obtain spatial patterns of water versus energy limitation in Europe and respective dependency on mean surface soil moisture. At this continental scale, the CSM is determined at 21 Vol-%. This is more toward the dry end of the spatiotemporal European soil moisture distribution and therefore in contrast to land models, which often assume water-limited conditions just below or at the field capacity. This finding can help to improve soil moisture stress representations in models. Application of the determined large-scale CSM directly as a land surface parameter in land models, however, should be avoided due to dependency on the employed surface soil moisture data set. Next to the large-scale CSM, we determine a range of small-scale CSMs and find ample variability according to local soil, climate, and vegetation characteristics. With readily available satellite soil moisture information, the small-scale CSMs allow real-time diagnosis of land-atmosphere interactions and their corresponding role during climate extremes, such as heat waves or droughts.

References

Akbar, R., Short Gianotti, D. J., McColl, K. A., Haghighi, E., Salvucci, G. D., & Entekhabi, D. (2018). Estimation of landscape soil water losses from satellite observations of soil moisture. *Journal of Hydrometeorology*, *19*(5), 871–889.

Arora, V. (2002). Modeling vegetation as a dynamic component in soil-vegetation-atmosphere transfer schemes and hydrological models. *Reviews of Geophysics*, *40*(2), 1006. https://doi.org/10.1029/2002JD002055

Boese, S., Jung, M., Carvalhais, N., Teuling, A. J., & Reichstein, M. (2019). Carbon-water flux coupling under progressive drought. *Biogeosciences*, *16*(13), 2557–2572.

Budyko, M. I. (1974). *Climate and life* (p. 508). New York: Academic Press.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI soil moisture for improved Earth system understanding: State-of-the-art and future directions. *Remote Sensing of Environment*, *203*, 185–215.

Feldman, A. F., Short Gianotti, D. J., Trigo, I. F., Salvucci, G. D., & Entekhabi, D. (2019). Satellite-based assessment of land surface energy partitioning-soil moisture relationships and effect of confounding variables. *Water Resources Research*, *55*.

Findell, K. L., Gentine, P., Lintner, B. R., & Guillod, B. P. (2015). Data length requirements for observational estimates of land-atmosphere coupling strength. *Journal of Hydrometeorology*, *16*(4), 1615–1635.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1), 168–182.

Good, S. P., Noone, D., & Bowen, G. (2015). Hydrologic connectivity constraints partitioning of global terrestrial water fluxes. *Science*, *349*(6344), 175–177.

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple collocation-based merging of satellite soil moisture retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(12), 6780–6792.

Guillod, B. P., Davin, E. L., Kündig, C., Smiatek, G., & Seneviratne, S. I. (2013). Impact of soil map specifications. *Climate Dynamics*, *40*(1–2), 123–141.

Guillod, B. P., Orlowsky, B., Miralles, D. G., & Teuling, A. J. (2015). Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nature Communications*, *6*, 6443. https://doi.org/10.1038/ncomms7443

Haghighi, E., Short Gianotti, D. J., Akbar, R., Salvucci, G. D., & Entekhabi, D. (2018). Soil and atmospheric controls on the land surface energy balance: A generalized framework for distinguishing moisture-limited and energy-limited evaporation regimes. *Water Resources Research*, *54*(3), 1831–1851.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, *113*(D20).

Hengl, T., de Jesus, J. M., Heuvelink, G. B., Gonzalez, M. R., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, *12*(2), e0169748.

Hirschi, M., Mueller, B., Dorigo, W., & Seneviratne, S. I. (2014). Using remotely sensed soil moisture for land-atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability. *Remote Sensing of Environment*, *154*, 246–252.

Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., et al. (2011). Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*, *4*(1), 17–21.

Homaee, M., Feddes, R. A., & Dirksen, C. (2002). Simulation of root water uptake: II. Non-uniform transient water stress using different reduction functions. *Agricultural Water Management*, *57*(2), 111–126.

Köhler, P., Guanter, L., & Joiner, J. (2015). A linear method for retrieval of Sun-Induced Chlorophyll Fluorescence from GOME-2 and SCIAMACHY data. *Atmospheric Measurement Techniques*, *8*(6), 2589–2608.

Koster, R. D., Chang, Y., Wang, H., & Schubert, S. D. (2016). Impacts of local soil moisture anomalies on the atmospheric circulation and on remote surface meteorological fields during boreal summer: A comprehensive analysis over North America, J. *Climacteric*, *29*(20), 7345–7364.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., et al., & GLACE Team (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, *305*(5687), 1138–1140. https://doi.org/10.1126/science.1100217

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., & Puma, M. J. (2009). On the nature of soil moisture in land surface models. *Journal of Climate*, *22*(16), 4322–4335.

Lawrence, D. M., Thornton, P. E., Oleson, K. W., & Bonan, G. B. (2007). The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmosphere interaction. *Journal of Hydrometeorology*, *8*(4), 862–880.

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A., Wagner, W., McCabe, M. F., et al. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment*, *123*, 280–297.

Novák, V., & Havrila, J. (2006). Method to estimate the critical soil water content of limited availability for plants. *Biologia*, *61*(19), S289–S293.

Orth, R., & Seneviratne, S. I. (2012). Analysis of soil moisture memory from observations in Europe. *Journal of Geophysical Research*, *117*(D15).

Pitman, A. J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, *23*, 479–510.

Qiu, J., Crow, W. T., Nearing, G. S., Mo, X., & Liu, S. (2016). The impact of vertical measurement depth on the information content of soil moisture for latent heat flux estimation. *Journal of Hydrometeorology*, *17*(9), 2419–2430.

Schlesinger, W. H., & Jasechko, S. (2014). Transpiration in the global water cycle. *Agric. Forest Meteorol.*, *189*, 115–117.

Schwingshackl, C., Hirschi, M., & Seneviratne, S. (2017). Quantifying spatiotemporal variations of soil moisture control on surface energy balance and near-surface air temperature. *Journal of Climate*, *30*(18), 7105–7124.

Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F., Berry, J., et al. (1997). Modeling the exchanges of energy, water and carbon between continents and the atmosphere. *Science*, *275*(5299), 502–509.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3), 125–161.

Seneviratne, S. I., Lüthi, D., Litschi, M., & Schär, C. (2006). Land-atmosphere coupling and climate change in Europe. *Nature*, *443*(7108), 205–209. https://doi.org/10.1038/nature05095

Taylor, C. M., de Jeu, R. A., Guichard, F., Harris, P. P., & Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *Nature*, *489*(7416), 423–426.

Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., et al. (2009). A regional perspective on trends in continental evaporation. *Geophysical Research Letters*, *36*(2).

Teuling, A. J., Seneviratne, S. I., Williams, C., & Troch, P. A. (2006). Observed timescales of evapotranspiration response to soil moisture. *Geophysical Research Letters*, *33*(23).

Teuling, A. J., Uijlenhoet, R., van den Hurk, B., & Seneviratne, S. I. (2009). Parameter sensitivity in LSMs: An analysis using stochastic soil moisture models and ELDAS soil parameters. *Journal of Hydrometeorology*, *10*(3), 751–765.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., et al., et al. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, *13*, 4291–4313.

Ulaby, F. T., Moore, M. K., & Fung, A. K. (1982). *Microwave remote sensing, active and passive: Radar remote sensing and surface scattering and emission theory* (Vol. 2). Norwood, MA: Artech House.

Van Genuchten, M. T. (1987). *A numerical model for water and solute movement in and below the root zone, Research Report*. Riverside, CA: US Salinity Lab.

Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., et al. (2017). Pedotransfer functions in Earth system science. *Challenges and Perspectives Reviews of Geophysics*, *55*(4), 1199–1256.

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., et al. (2013). Global data sets of vegetation leaf area index (LAI) 3g and fraction of photosynthetically active radiation (FPAR) 3g derived from global inventory modeling and mapping studies (GIMMS) normalized vegetation difference vegetation index (NDVI) for the period 1981 to 2011. *Remote Sensing*, *5*(2), 927–948.

Zscheischler, J., Orth, R., & Seneviratne, S. I. (2015). A submonthly database for detecting changes in vegetation-atmosphere coupling. *Geophysical Research Letters*, *42*(22), 9816–9824.

Erratum

In the originally published version of this article, the Plain Language Summary did not appear. The Plain Language Summary has been added, and this version may be considered the authoritative version of record.