AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR021508

Hydrological Storage Length Scales Represented by Remote Sensing Estimates of Soil Moisture and Precipitation

Ruzbeh Akbar[1], Daniel Short Gianotti[1], Kaighin A. McColl[2], Erfan Haghighi[1], Guido D. Salvucci[3], and Dara Entekhabi[1]

[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, [2]Department of Earth and Planetary Science, Harvard University, Cambridge, Massachusetts, USA, [3]Department of Earth and Environment, Boston University, Boston, Massachusetts, USA

Key Points:
- Surface soil moisture dynamic often reflects deeper soil water content characteristics
- By performing water balance closure using only soil moisture and precipitation a length scale can be obtained
- Estimates of the length scale show an east-west gradient across US with larger values in wetter regions

Correspondence to:
R. Akbar,
rakbar@mit.edu

Citation:
Akbar, R., Gianotti, D. S., McColl, K. A., Haghighi, E., Salvucci, G. D., & Entekhabi, D. (2018). Hydrological storage length scales represented by remote sensing estimates of soil moisture and precipitation. *Water Resources Research*, *54*. https://doi.org/10.1002/2017WR021508

Received 12 JUL 2017
Accepted 8 FEB 2018
Accepted article online 16 FEB 2018

Abstract The soil water content profile is often well correlated with the soil moisture state near the surface. They share mutual information such that analysis of surface-only soil moisture is, at times and in conjunction with precipitation information, reflective of deeper soil fluxes and dynamics. This study examines the characteristic length scale, or effective depth Δz, of a simple active hydrological control volume. The volume is described only by precipitation inputs and soil water dynamics evident in surface-only soil moisture observations. To proceed, first an observation-based technique is presented to estimate the soil moisture loss function based on analysis of soil moisture dry-downs and its successive negative increments. Then, the length scale Δz is obtained via an optimization process wherein the root-mean-squared (RMS) differences between surface soil moisture observations and its predictions based on water balance are minimized. The process is entirely observation-driven. The surface soil moisture estimates are obtained from the NASA Soil Moisture Active Passive (SMAP) mission and precipitation from the gauge-corrected Climate Prediction Center daily global precipitation product. The length scale Δz exhibits a clear east-west gradient across the contiguous United States (CONUS), such that large Δz depths ($>$200 mm) are estimated in wetter regions with larger mean precipitation. The median Δz across CONUS is 135 mm. The spatial variance of Δz is predominantly explained and influenced by precipitation characteristics. Soil properties, especially texture in the form of sand fraction, as well as the mean soil moisture state have a lesser influence on the length scale.

1. Introduction

While surface soil moisture constitutes only a fraction of the total storage of terrestrial water (Gleeson et al., 2015), it rests at a critical boundary between the atmosphere and land. Hence, it plays a disproportionately large role in the water cycle (McColl et al., 2017a). Soil moisture affects the land-atmospheric coupling through partial control on the evaporation of precipitation water accumulated over time (Schwingshackl et al., 2017; Seneviratne et al., 2010) and therefore influencing the probability of subsequent precipitation (Tuttle & Salvucci, 2016). Soil moisture also influences crop and plant growth, dynamics of soil respiration (Rosenzweig et al., 2002), and modulates drought conditions (Koster, 2004). Most significantly, it affects evapotranspiration (ET), which serve as the link between the global water and energy cycles (Fatichi et al., 2016).

In vegetated regions, and more specifically water-limited areas, the amount and rate of plant transpiration are predominantly affected by root-zone soil moisture availability (Vivoni et al., 2008)—that is the soil moisture profile down to tens of centimeters within below the land surface. However, the exact depth is dependent on the ecosystem, plant type, and root distribution (Kurc & Small, 2007).

At global scales, remote sensing measurements of root-zone soil moisture are minimal. Low-frequency radar observations, 430 MHz, by the NASA Airborne Microwave Observatory of Subcanopy and Subsurface (AirMOSS) (Tabatabaeenejad et al., 2015) have recently retrieved and reported root-zone moisture at regional scales (100 $\times$ 25 km) over North America. Overall, however, within the past few decades, technological advancements in microwave remote sensing have enabled frequent spaceborne estimates of surface—or top 50 mm—soil moisture. Recent satellite missions such as the European Space Agency's Soil Moisture Ocean Salinity (SMOS) (Kerr et al., 2001) and the NASA Soil Moisture Active Passive (SMAP) (Entekhabi et al.,

2010) mission continue to provide global soil moisture coverage at about 40 km spatial resolution from daily to 3 day intervals.

1.1. Surface and Root-Zone Soil Moisture

Large scale weather forecasting and climate modeling efforts typically rely on deeper soil moisture information to more accurately describe land-atmosphere coupling, evapotranspiration, along with heat and water exchanges (Dirmeyer, 2000; Koster & Suarez, 2003). To overcome limitations imposed by the lack of observation-based deeper soil moisture information, a wide range of land-surface modeling and assimilation techniques have been developed to link surface observations to deep soil moisture. These methods span a variety of techniques including semiempirical low-pass filtering of time series soil moisture (Albergel et al., 2008; Wagner et al., 1999) and data assimilation approaches (Kumar et al., 2009; Reichle et al., 2007, 2008; Sabater et al., 2007). Applicability of these methods rests on accurate model parameterization, soil texture and vegetation cover, as well as atmospheric forcing.

Due to direct exposure to precipitation and evaporation processes, surface soil moisture dynamics—compared to deeper soil moisture—are generally more rapid. However, it is physically linked via diffusive processes to the deeper soil moisture profile and its temporal dynamics do contain some information about the temporal evolution of deeper soil moisture. The exact nature of this link is complex, and a function of climate, soil texture, as well as land cover type and use (Mahmood & Hubbard, 2007). With the exception of very dry soil conditions (Hirschi et al., 2014) surface soil moisture is strongly correlated with deeper layers (Albergel et al., 2008; Ford et al., 2014; Qiu et al., 2014). While root-zone soil moisture is usually the dominant control on surface fluxes, rather than surface soil moisture, the correlation between surface soil moisture and root-zone soil moisture leads to a strong relationship between surface soil moisture and surface fluxes in many cases (Qiu et al., 2016).

To demonstrate the correlation, Figure 1 shows the explained-variance, R^2, between surface (5 cm) and root-zone soil moisture, i.e., sensors placed at 10, 20, 50, and 100 cm depths. Data from the US Climate Reference Network (USCRN) (Bell et al., 2013) in situ networks over CONUS are used to determine the statistic

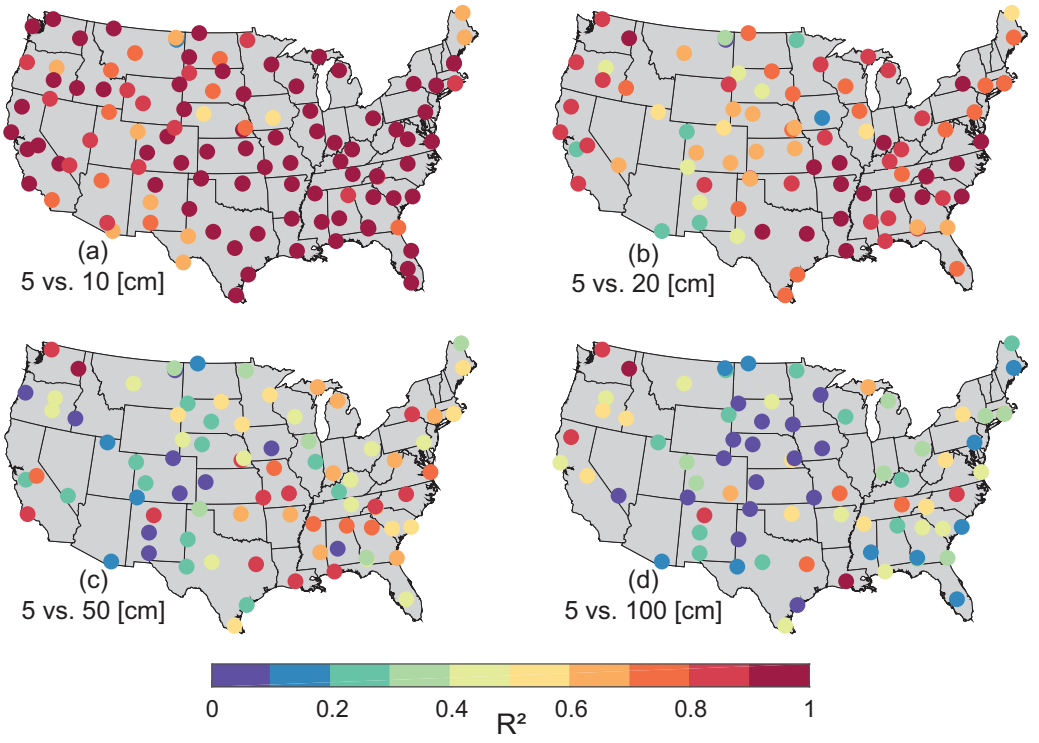

Figure 1. Explained-variance between surface 5 cm and root-zone (plots a through d for 10, 20, 50, and 100 cm) soil moisture from the USCRN in situ networks over a 1 year period. R^2 is highest between 5 and 10 cm soil moisture but degrades, at a variable rate, with respect to deeper depths.

for a 1 year period (2015). Across CONUS, for the 5 and 10 cm depths R^2 is largest, and close to 1. With respect to deeper sensor depths, however, it gradually degrades, but at variable rates due to differences in soil texture and climate. For parts of southeast US, the correlation coefficient is still large $R^2 > 0.5$ down to a depth of 50 cm.

1.2. Hydrological Length Scale

As stated in the previous section, surface and deeper soil moisture dynamics are often correlated and convey mutual information. Thus, analysis of surface-only soil moisture is, at times, reflective of deeper soil fluxes and dynamic. Based on this understanding, a water balance equation can be formulated which describes an effective and active homogenous storage volume characterized by precipitation inputs, $P(t)$ [$L\ T^{-1}$], and water storage dynamics as evident in SMAP surface soil moisture observations, θ_{SMAP} ($m^3 m^{-3}$),

$$\Delta z \frac{d\theta_{SMAP}}{dt} = P(t) - Q(\theta_{SMAP}) \tag{1}$$

$Q(\theta_{SMAP})$ [$L\ T^{-1}$] is the soil moisture hydrologic divergence of water from the control volume. It encapsulates total moisture losses from the control volume due to evapotranspiration, $ET(\theta)$ [$L\ T^{-1}$], and percolation into deeper soil, $D(\theta)$ [$L\ T^{-1}$], as reflected by surface soil moisture dynamics.

In the water balance equation of (1), Δz [L] is defined as the unknown characteristic length scale of the control volume. It transforms volumetric water fraction (θ) to a storage volume of water per unit area (depth) in the landscape soil. The distinction is analogous to intensive and extensive variables in thermodynamics. The value of an intensive variable, such as temperature, is independent from the size of the enclosing (balance) control volume. Analogously, volumetric water content is an intensive variable. However, without a length scale, it does not carry information about the amount of water within the system, nor can it be directly applied to balancing input and output fluxes.

Extensive variables, such as mass or enthalpy, depend on the size of the balance control volume and more directly relate to fluxes into and out of the system. In the case of hydrological water balance, the soil water content has to be described in terms of an extensive variable—in volume per unit area or depth units—if it is used in conjunction with water balance and precipitation and hydrologic losses. This is evident in the mass balance relation in (1) with the inclusion of the length scale Δz.

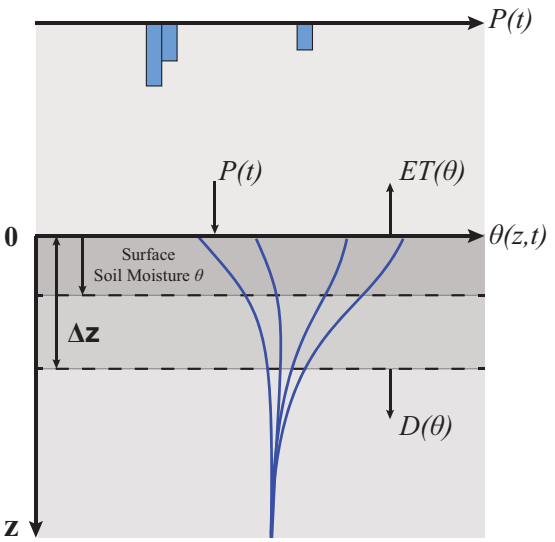

Figure 2. Schematic of conceptual control volume with depth Δz. This length scale is dependent on soil properties (texture, hydraulic conductivity, etc.) as well as precipitation, P(t), characteristics such that surface soil moisture dynamics are also well correlated with total stored water content at this depth.

Using precipitation information, SMAP surface soil moisture observations, and enforcing mass conservation in (1) this unique length scale can be determined. Figure 2 shows a schematic representation of this conceptualization and the control volume with respect to surface soil moisture. Given input precipitation along with estimates of outgoing losses due to percolation into deeper soil and evapotranspiration, a unique length scale Δz exists where surface soil moisture dynamics reflect deeper soil water content—Appendix A also includes a complementary study in support of this feature using in situ profile soil moisture data. In this context, the uniqueness of Δz refers to the fact that, for a given region and over the period of study, a specific length scale Δz can be determined which is in a function of soil and precipitation characteristics.

Conceptually, Δz is analogous to the damping depth of diurnal temperature waves within a soil layer (Dickinson, 1988; Hu & Islam, 1995). This damping depth is a length scale dependent on two sets of properties: first, the soil's thermal characteristics, namely thermal conductivity and heat capacity. The second is the frequency and periodicity of external forcing—diurnal and seasonal input radiation. From a hydrological dynamics perspective in the context of (1), estimates of Δz reflect a length scale dependent on the soil hydraulic conductivity and water storage potential—in general, mean soil moisture and texture—as well as precipitation interstorm periods (forcing intervals)

and potential evaporation rates. Therefore, the degree of hydrological activity, reflected in precipitation and mean soil moisture characteristics, will affect Δz.

Koster and Suarez (2001) demonstrated that, in the context of soil moisture memory and autocorrelation, while temporal soil moisture dynamics can be more accurately described—either from observations or models—an effective water storage or holding capacity is required. In relation to the system's total sorted water content [*L*], the effective holding capacity was estimated via calibration of a monthly water balance equation, similar to (1). By applying a variance matching technique between surface soil moisture dynamics and gravity-based total water content (Crow et al., 2017) also determined a suitable scale factor to evaluate basin level annual water balance closure. Similar to Koster and Suarez (2001) and Crow et al. (2017), a hydrologic length scale is estimated in this study. However, unlike Koster and Suarez (2001), values of fluxes such as evaporation are not required. And unlike Crow et al. (2017) gravity-based measurements of total soil water storage are not required.

In this study, we focus on the estimating this hydrologic length scale using a data-driven approach leveraging SMAP volumetric soil water content and precipitation data. In applying (1) to the estimation process, the functional form of the hydrologic divergence, $Q(\theta_{SMAP})$ must to be known. We seek to estimate this function in a data-driven and nonparametric fashion. In section 2, we use the temporal sequences of remotely sensed surface soil moisture to estimate this function and its state-dependence. An observation-driven approach is presented to partially reconstruct $Q(\theta_{SMAP})$. The hydrological model in (1) considers only vertical input and output of moisture. Over large scales (tens of kilometers) the vertical gains and losses are assumed to dominate over lateral hydrologic exchanges. Additionally, the soil moisture remote sensing community has historically attributed a 50 mm vertical support depth to soil moisture estimates from L-band (1.4 GHz) radiometers. However, discussions surrounding Δz reflect a length scale characterizing a hydrological system with knowledge of precipitation. Therefore, it can be different from the SMAP 50 mm support depth defined for its estimates of volumetric water content.

With the above considerations in mind, we pose the following question: What is the characteristic length scale of a simple hydrological system described only by precipitation inputs and surface soil water dynamics evident in SMAP observations? We focus here on data and observation-driven approaches to enhance utility of satellite remote sensing observation of soil moisture for water balance studies. Naturally, more complete modeling and model-based assimilation techniques can be pursued but are not considered here: only soil moisture and precipitation information are used.

To address the above question, two steps are taken. First, in section 2, we present a simple framework to partially reconstruct the hydrologic divergence function by analyzing consecutive negative increments of soil moisture—specifically during soil moisture dry-downs. Then, an optimization problem is formulated to provide estimates of Δz over the Contiguous United States (CONUS) by minimizing the root-mean-squared (RMS) difference between water balance-based estimates of soil moisture and SMAP soil moisture product. This analysis will focus on May–September 2015 and 2016 (inclusive). In section 3, estimates of the characteristic length scale are presented and its implications are discussed. The influence of precipitation, soil moisture as well as soil properties on this length scale are also examined. Appendix A includes a complementary analysis in support of the relationship between surface soil moisture dynamics and total stored water content at the length scale Δz, but based on multidepth in situ soil moisture observations only.

2. Materials and Methods

2.1. Soil Moisture and Precipitation Measurements

The SMAP Enhanced Radiometer-only surface soil moisture estimates (L3SM_P_E) (O'Neill et al., 2016), at a 9 km posting, are used in this analysis. Each 9 km data granule represents surface soil volumetric water content over a large (approximately 30–40 km) area corresponding to the half-power or -3 (dB) sampling area of the SMAP antenna. Depending on latitude, SMAP soil moisture observations are available from daily to 3 day intervals. The SMAP project ensures the validity of soil moisture estimates, globally, via calibration and validation efforts with respect to multiple in situ sensor networks (Chan et al., 2016). Only the descending, or local "6 am," soil moisture products are considered. This is due to uncertainties associated with model-based predictions of evening soil and vegetation physical temperatures; these in turn will affect soil moisture retrieval. It is important to note that the SMAP retrieval algorithm and soil moisture products

(Entekhabi et al., 2010; O'Neill et al., 2016) assumes a semi-infinite dielectric medium characterized by homogenous and uniform texture within the 50 mm soil layer. We extend the uniform soil texture assumption to this analysis as well.

Surface precipitation estimates are obtained from the Climate Prediction Center's Unified (CPCU) gauged-based global daily precipitation (NCAR, 2017). CPCU precipitation is reprocessed from a 0.5 degree grid to the SMAP 9 and 36 km grids and resampled to 6 hourly accumulations starting from 0600 local time—thus concurrent with SMAP observations, when available.

SMAP soil moisture and CPCU precipitation data spanning May–September 2015 and 2016 over the contiguous United States (CONUS) are used. In all cases, pixels affected by radio frequency interference (RFI), more that 1% water fraction, more than 7 kg m^{-2} vegetation water content as well as those flagged as frozen or snow covered are excluded from the analysis. Soil moisture quality and use-recommendation flags are provided in the SMAP data files and help to minimize the impact of uncertainties associated with soil moisture estimates.

Soil texture, in the form of sand and clay fractions, are also used to approximate the soil's porosity as $\phi = (\text{sand} \times 0.395) + (\text{clay} \times 0.482) + (1 - \text{sand} - \text{clay}) \times 0.451$. We use the same soil texture data set as SMAP (Das, 2013).

2.2. Estimating the Soil Moisture Loss Function

In (1), estimates of volumetric soil moisture are available from SMAP and precipitation from the CPCU gauge-corrected daily products. Note, however, to estimate Δz the water balance model of (1) must be completed with knowledge of the hydrologic divergence function, $Q(\theta_{SMAP})$. Stochastic modeling of soil moisture dynamics have proposed a variety of physics-based analytical forms for $Q(\theta_{SMAP})$ (Feng et al., 2012; Laio et al., 2001; Rodriguez-Iturbe & Porporato, 2007). However, at large scales their applicability diminishes due to difficulties in parameterizing such models. An alternative method for estimating the hydrologic divergence function from precipitation and soil moisture is to conditionally average the daily precipitation amount according to the soil moisture status, as demonstrated by Salvucci (2001) with field data and Tuttle and Salvucci (2014) with AMSR-E estimated of soil moisture.

In lieu of model-based approaches, we present an observation-driven method to estimate the loss function. At larger spatial scales, increases in soil moisture are attributed to infiltrating precipitation and decreases due to evapotranspiration and percolation into deeper soil. Thus, soil water losses are partially encoded in the gradients of volumetric soil moisture content increments, especially dry-downs with zero infiltrating precipitation in-between successive observations. For example, McColl et al. (2017b) estimated the time scale of stage-II ET—water-limited regime—from SMAP-observed dry-downs. Here dry-downs are extracted from the SMAP soil moisture time series by analyzing successive negative increments of volumetric soil water content. A loss function and its dependence on the state variable (i.e., volumetric soil water content) is then approximated by conditioning these negative increments on soil moisture itself. With SMAP alone, the loss function is in units of time-increment decrease in θ_{SMAP}, i.e., (m^3 m^{-3} d^{-1}). Therefore, with no precipitation between successive observations the loss function becomes

$$L(\theta_{SMAP}) = E\left[-\frac{\Delta\theta^{-}_{SMAP}}{\Delta t^{obs}}\middle|\theta_{SMAP}\right] \tag{2}$$

where the increments are $\Delta\theta_{SMAP} = \theta_{SMAP}(t+1) - \theta_{SMAP}(t)$ and

$$\Delta\theta^{-}_{SMAP} = \begin{cases} \Delta\theta_{SMAP}\ \Delta\theta_{SMAP} < 0 \ \textit{and}\ P(t) = 0 \\ 0\ \textit{otherwise} \end{cases}$$

where E[] is the expectation operator. Δt^{obs} is the time difference, in days, between successive SMAP soil moisture observations that yield a negative change, or loss, in moisture. Note that the loss function, $L(\theta_{SMAP})$, defined in (2) is in units of volumetric water content per day [T^{-1}]. The total amount of loss in (mm) can be determined once the new length scale, Δz, is determined, i.e., $Q(\theta_{SMAP}) = L(\theta_{SMAP}) \cdot \Delta z$. Noise and uncertainties in the soil moisture retrieval process affect loss function estimates derived from soil moisture increments. An error analysis by McColl et al. (2017a) showed that larger increments are less

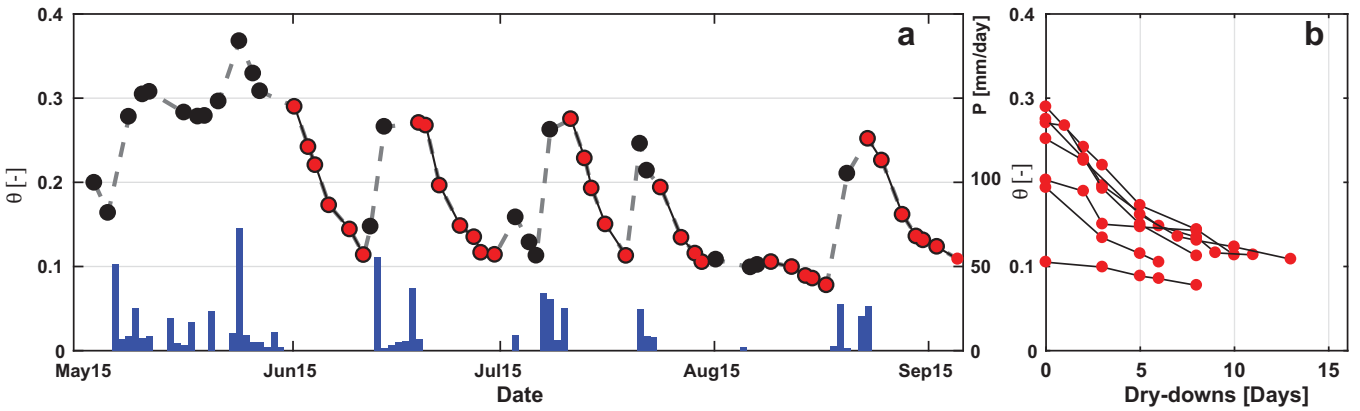

Figure 3. (a) Example SMAP soil moisture time series for a pixel in Oklahoma US for May–September 2015. The red circles indicate dry-downs with zero precipitation (shown as the blue bars) in between SMAP observations. (b) The collection of dry-downs over time. The loss function $L(\theta)$ is the negative gradient of these dry-downs.

susceptible to noise effect. Similar to McColl et al. (2017b), in this study, increments less than 1% of the range of observed soil moisture—along with those with intermeasurement precipitation—are excluded.

In Figure 3a, an example soil moisture time series is shown to highlight how the loss function is constructed. Here dry-downs with a minimum of three soil moisture observations and two consecutive negative increments, $\Delta\theta^-$, are identified. Figure 3b shows the collection of corresponding dry-downs over time, with the longest dry-down spanning 13 consecutive days. For each individual dry-down the incremental soil moisture loss, between time t and $t+1$ is $\frac{-\Delta\theta^-}{\Delta t^{obs}}$. These are shown as blue dots in Figure 4a. The loss function $L(\theta)$ is then obtained by applying a locally weighted linear smoother, or LOWESS, across all the data from all dry-downs. The smoother used as span of 65% of the data when performing local regressions. In general, the loss function is a monotonically increasing function with respect to increasing soil moisture; that is larger losses for larger soil moisture values. The values $\theta_{\min}$ and θ_{up} denote the minimum and upper observed soil moisture values used to define $L(\theta)$ and determine the "observation-based" domain, or segment "**B**" in Figure 3.

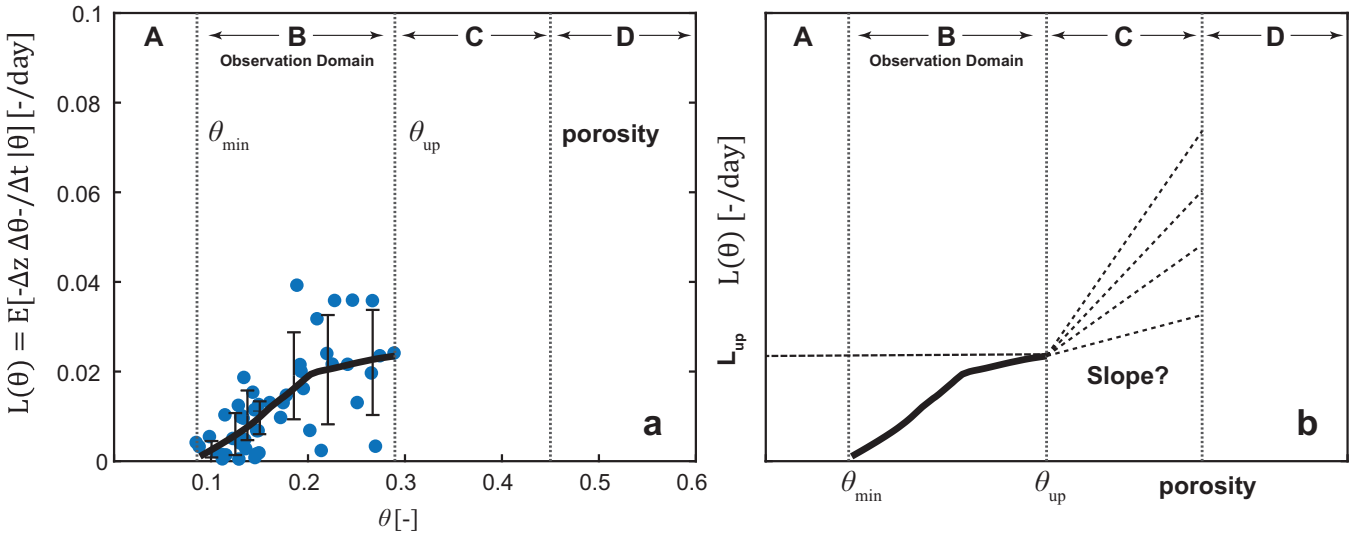

Figure 4. (a) Reconstructed soil moisture loss function based on the SMAP soil moisture time series of Figure 3a. The small blue markers are individual losses, $\frac{-\Delta\theta^-}{\Delta t^{obs}}$. The conditioned expectation of these losses conditioned on soil moisture, after application of a locally weighted linear smoother, is the final loss function: $L(\theta)=E\left[\frac{-\Delta\theta^-}{\Delta t^{obs}}|\theta\right]$ (thick black line). The error bars show the vertical standard deviation of losses based on seven equal-count soil moisture bins. The loss function, based on the dynamic range of observations, is limited between $\theta_{\min}$ and θ_{up}. (b) Schematic representation of the loss function along with four different loss Segments A: dry-end interpolation; B: observation domain; C: slope-estimation domain; and D: losses due to saturation and beyond. In Segment C, the slope of the loss function α is unknown and imbedded in the optimization process.

Given the nominal 3 day SMAP revisit time and variable soil moisture dynamics, certain portions of the loss function are not observed. Specifically, outside the observation-based domain (between $\theta_{\min}$ and θ_{up}), the loss function must be defined in three additional segments:

a. For prolonged dry intervals, with zero precipitation, the amount of loss must also decrease with decreasing soil moisture. However, the dynamic range of observations are limited and do not always span the dry-end. To overcome this limitation, when needed, the loss function linearly interpolates between the minimum observed soil moisture and the minimum possible SMAP value of 0.02 (m^3 m^{-3}). The linear interpolation behavior is consistent with the linear Stage-II evaporation loss process in water-limited areas. We label this segment as "**A**" seen in Figure 3b.
b. Percolation or leakage into the soil during and immediately after precipitation events is a rapidly occurring process. Therefore, if a rain event occurs in-between two consecutive SMAP observations, the percolation process may have finished and is entirely missed. To accommodate this, a linear loss component, below the porosity level and above θ_{up}, with an unknown slope of α [T^{-1}] in the form of $L(\theta_{up})+\alpha\cdot(\theta(t)-\theta_{up})$ is included. This slope-estimation segment is labeled as "**C**" in Figure 3b.
c. Saturation losses are also considered for cases where the soil moisture level approaches and exceeds the soil porosity value. These losses can be considered as runoff or spillage, for example, and are defined as $\frac{P(t)}{\Delta z}-(\phi-\theta(t))$, where ϕ is porosity. This segment is labeled as "**D**."

Therefore, taking in to account (a)–(c), the loss function in (2) is amended as

$$L(\theta)=\begin{cases}(\theta-0.02)\cdot\dfrac{L(\theta_{\min})}{(\theta_{\min}-0.02)} & 0.02\le\theta\le\theta_{\min}\ (\text{Segment A})\\ E\left[\dfrac{-\Delta\theta^{-}}{\Delta t^{obs}}|\theta\right] & \theta_{\min}\le\theta\le\theta_{up}\ (\text{Segment B})\\ L(\theta_{up})+\alpha\cdot(\theta-\theta_{up}) & \theta_{up}<\theta<\phi\ (\text{Segment C})\\ \dfrac{P}{\Delta z}-(\phi-\theta) & \theta\ge\phi\ (\text{Segment D})\end{cases} \tag{3}$$

Hereafter, the subscript SMAP for θ_{SMAP} is dropped for presentation clarity.

Changes in the volumetric soil moisture content in (3) are principally due to evapotranspiration and interstorm percolation from surface to subsurface within the soil column. The last component of (3), Segment D, additionally captures the saturation-excess runoff if θ_{SMAP} reaches porosity. There are, however, important runoff processes that are not adequately captured by (3). The SMAP product (θ_{SMAP}) represents an estimate of near-surface volumetric soil water content over a large area corresponding to the half-power scale of the SMAP instrument antenna—approximately 40 km. Saturation-excess runoff may occur over such a large domain even if the entire domain is not saturated. In humid or adequately wetted domains, θ_{SMAP} may be below porosity but, within the domain, there may be areas where the soil is locally saturated, e.g., downslope, near-channel, floodplain, or local areas with high flow accumulation. Precipitation incident over such domains will generate runoff and not all the precipitation will add to the storage in the control volume. Even though Segment D captures the saturation-excess runoff process when the entire domain is saturated, there is likely an underestimation of this hydrologic process with (3) and the use of coarse scale (~40 km) SMAP estimates of soil moisture. Additionally (3) does not include infiltration excess runoff and lateral flows among adjacent pixels. At the relative coarse scale, systematic exchange between pixels is likely to be far less than vertical evaporation and percolation exchanges.

The impact of these shortcomings can be tested by examining what percent-of-time, and what fraction of the soil moisture dynamics, is due to Segment B of (3). Segment B is constructed based on SMAP observations alone. Segments A, C, and D are extrapolations. In section 3.2, diagnostics on the percent-of-time and fraction of water balance in the B and the A + C + D segments are presented. The results are indicative of the degree to which the problem posed in this study is solved using observations only, i.e., observation-driven.

The procedure to estimate the loss function is in contrast to Koster et al. (2017) where the entire form of $L(\theta)$ is first parameterized as a monotonically increasing piecewise linear function. Then, for each pixel, by minimizing the norm difference between model predictions and SMAP observations, the parameters of $L(\theta)$ were estimated. Here knowledge of the loss function is partially obtained from soil moisture dry-downs and analysis of successive negative increments.

We emphasize that the form of the loss function given in (3) encapsulates all moisture losses as evident in surface soil moisture observations. Determination of the individual components of $L(\theta)$, e.g., evaporation, runoff, etc., requires additional modeling or ancillary data. The method outlined in this section is a first-order attempt at determining $L(\theta)$ based solely on soil moisture observations. It is sufficient for the purpose of estimating Δz, the aim of this paper.

2.3. Estimating the Effective Hydrological Length Scale

The forward temporal evolution of soil moisture, $\theta(t+1)$ (m^{-3} m^{-3}), based on water balance can be written as

$$\theta(t+1) = \theta(t) + \frac{P(t)}{\Delta z} \cdot \Delta t - L(\theta) \cdot \Delta t \tag{4}$$

where $L(\theta)$ is the loss function in (3) and Δt is the model time step—different than Δt^{obs}—and is set to 6 h, i.e., 608 data points between May and September per year. By means of (4), the next time step soil moisture, $\theta(t+1)$, can be predicted using the current estimate of soil moisture $\theta(t)$, while accounting for total precipitation inputs and losses between t and $t+1$. This approach neglects diurnal soil moisture fluctuations. It is also important to note that during and immediately after rain events, precipitation inputs and losses from the control volume are simultaneous and continuous processes. However, the implementation of the loss function and hydrological model in (3) and (4) is a two-step process. First, soil moisture $\theta(t)$ is added to the accumulated precipitation over the time-interval Δt. Then, total losses are subtracted from the combined soil moisture and precipitation quantity. The piecewise implementation of the loss function in (3) supports this approach.

Note that within the current water balance framework two parameters are unknown: the length scale Δz and the extrapolation slope α. These parameters are estimated by formulating an objective function $J(\Delta z, \alpha)$ and then minimizing the root-mean-squared (RMS) difference between a time series of N SMAP observations, $\theta^{obs}(t)$, $t \in \left[t_1^{obs}, t_N^{obs}\right]$, and the corresponding predictions from (4). With a nominal 3 day SMAP revisit rate between May and September 2015 and 2016, N is approximately 130 data points. To estimate Δz and α, (4) is use to forward generate a time series of 6 hourly soil moisture using observations of precipitation and losses. Then, using Simulated Annealing (Ingber, 2012), as the global optimization routine, the overall sum of squared-differences between this time series and SMAP soil moisture is minimized. For each year, the time series in (4) is initialized by the corresponding first SMAP soil moisture observation.

Estimates of Δz can be different from the 50 mm support depth of SMAP soil moisture. This is partly due to the variable degrees of correlation between surface and subsurface soil moisture dynamics. Second, precipitation forcing as well as soil texture characteristic will affect the closure of (1) and minimization of the objective function. The surface soil moisture state must be conceptually redistributed within a control volume of length Δz.

Accurate estimation of Δz and α rest on satisfactory reconstruction of the loss function when considering observation-based approaches. In turn, this relies on detecting a sufficient number of soil moisture dry-downs. With this factor in mind, we proceed by first determining dry-downs from 9 km SMAP soil moisture observations at a given location. Then, using (2) individual observation-based losses are calculated. A collection of valid 9 km SMAP pixels, within a 36 km grid cell, are aggregated and combined together to increase the sample size and to include more weather-induced soil moisture dynamics and dry-downs. The final step involves minimizing the objective function using 36 km mean SMAP soil moisture and CPCU precipitation. Pixels with more that 1% water fraction, vegetation water contents of more than 7 kg m^{-2}, and frozen or snow covered are excluded. Moreover, note that estimation of Δz requires analysis of both wetting and dry-down periods, and thus requiring precipitation data. If dry-downs are only analyzed (P(t) = 0) in (4), Δz is indeterminate.

3. Results

3.1. Hydrological Length Scale

Figure 5 shows three pixel-level examples describing the length scale estimation process. The examples are (a) Grasslands in Western US, (b) Croplands in Southeast US (Northern Florida), and (c) Shrublands in Southwest US. Each example covers a range of soil moisture, texture, and precipitation characteristics. For each location, the figures include the 2015 and 2016 May–September SMAP time series soil moisture at 36 km in black. Concurrent 6 hourly soil moisture predictions from (4) based on estimates of the length scale Δz and slope α are shown in red. For each time series, the corresponding 6 hourly precipitation, $P(t)$, is also given. The annotated plots on the left present the loss functions, as described in Figure 3. In all cases, time series predictions from (4), exhibit high correlations with SMAP soil moisture ($R^2 > 0.6$) and appropriately timely "jumps" in soil moisture with respect to input precipitation. For all cases, the error standard deviations with respect to SMAP are also comparable and approximately 0.035 m^3 m^{-3}. Additionally, the corresponding length scale are 46, 296, and 91 mm, for figures a–c, respectively. Each is uniquely determined.

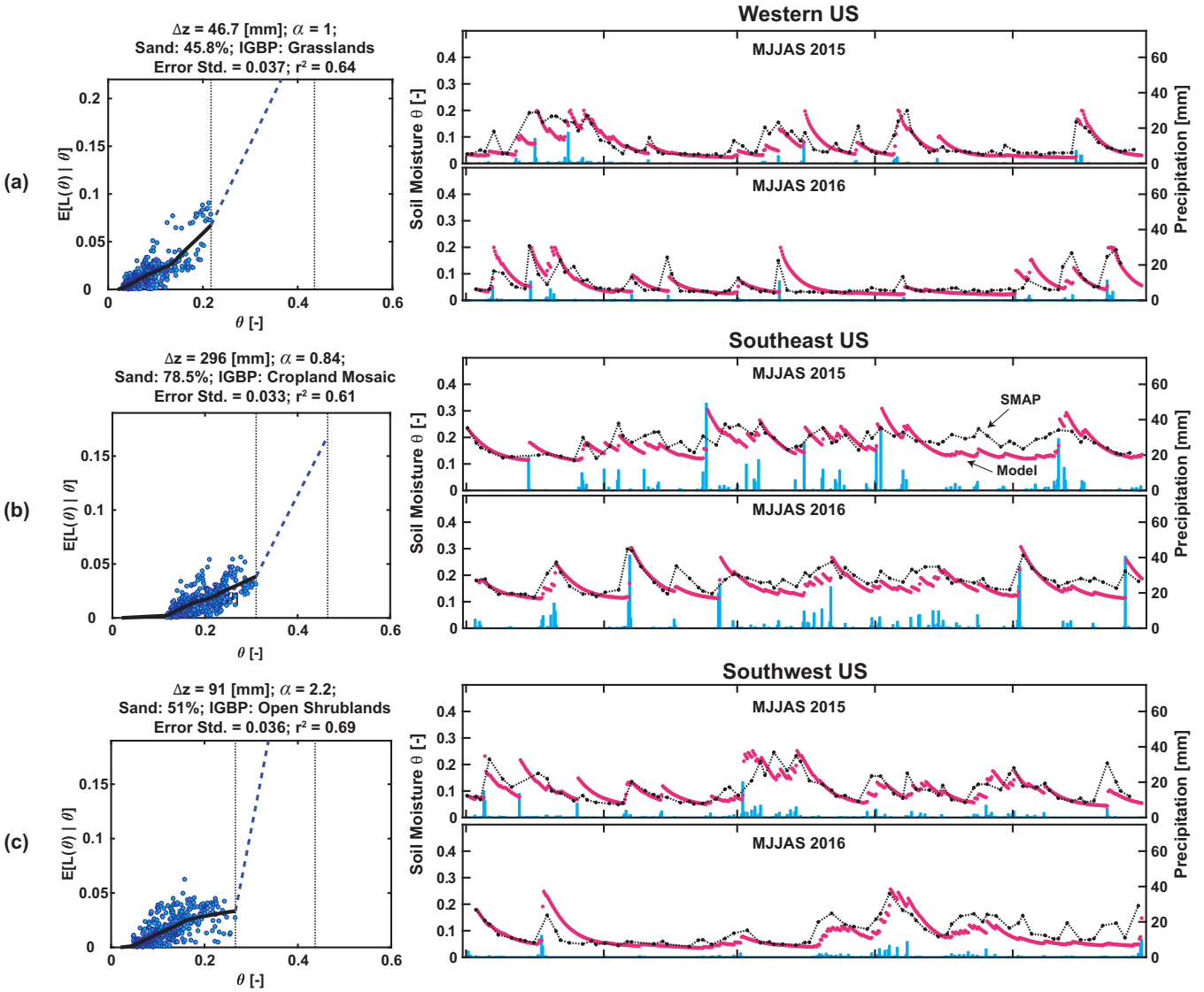

Figure 5. Three representative examples of SMAP (black) and 6 hourly predictions from (4) (magenta) soil moisture times series across CONUS: (a) Western US, (b) Southeast US, and (c) Southwest US. (left) The observation-based loss functions in (2) and the black line is after application of a local smoother. The dashed blue line is the linear loss component between θ_{up} and porosity with slope $\hat{\alpha}$. The two dashed vertical lines are from left to right, θ_{up} and porosity ϕ.

Section 1 states that the length scale is dependent on different soil and precipitation properties. This feature is evident in Figure 5. Compared to Figures 5a and 5c, the example in Figure 5b has larger mean precipitation and mean soil moisture—about 3 mm d^{-1} and 0.18 m^3 m^{-3}, respectively. Therefore, the length scale required to provide closure to (1) also becomes larger, $\Delta z = 296$ mm. Similarly, the example pixel in Figure 5c yields a larger length scale (91 mm) compared to Figure 5a (50 mm) due to slightly larger mean precipitation and soil moisture—1.1 mm d^{-1} and 0.1 m^3 m^{-3} compared to 0.7 mm d^{-1} and 0.06 m^3 m^{-3}. Thus, inherent soil properties and forcing patterns affect this length scale. Naturally, we can expect different length scales when performing the same analysis over winter months, for example, where precipitation patters may be different.

A map of Δz (mm) over CONUS is shown in Figure 6. Missing regions are due to vegetation water content larger than 7 kg m^{-2}, more than 1% water fraction, frozen or snow-covered soil, or no detectable drydowns. For each pixel, Δz is the unique length scale of the associated hydrological system characterized by precipitation inputs and water storage dynamics evident in SMAP surface soil moisture observations. Estimates of this length scale show an east-west gradient with smaller values in the drier western parts of CONUS, and larger values in the wetter eastern and southeastern parts of the country.

Over the duration of this study, far western parts of the US, specifically California, are dry and hydrologically inactive—with near-zero mean precipitation. Thus, the closure required for (1) yield very small characteristic length scales—similar to the scenario in Figure 5a. This is reflective of larger decoupling and decorrelation between surface and root-zone soil moisture (Hirschi et al., 2014). Surface soil moisture dynamics in these regions represent total water content from only a shallow depth. This behavior does change when the analysis is performed over a different period of time with different precipitation activity (figure not shown). In contrast, larger length scales ($\Delta z > 200$ mm) are evident in south and southeast CONUS—similar to Figure 5c—where on average the regions are wetter.

More humid regions such as West of Mississippi River and the Eastern Seaboard have the highest estimates of Δz in Figure 6. Much of these regions have vegetation water content greater than 7 kg m^{-2} which are masked due to the attenuation of the surface upwelling brightness temperature in dense canopies and associated uncertainty in soil moisture estimates. Over these humid regions the water balance in (3) may not adequately capture all the dominant hydrologic processes (see discussion at end of section 2.2). The Δz estimates in these regions need to be viewed with this uncertainty and source of error in mind. Furthermore, the structure of (4) indicates that if water input into the soil control volume is over-estimated, i.e., local saturation-excess runoff loss is not adequately abstracted from the domain precipitation input, the Δz may be over-estimated to compensate. This feature is more pronounced in the southeastern US where stormflow runoff may occur prior to full saturation.

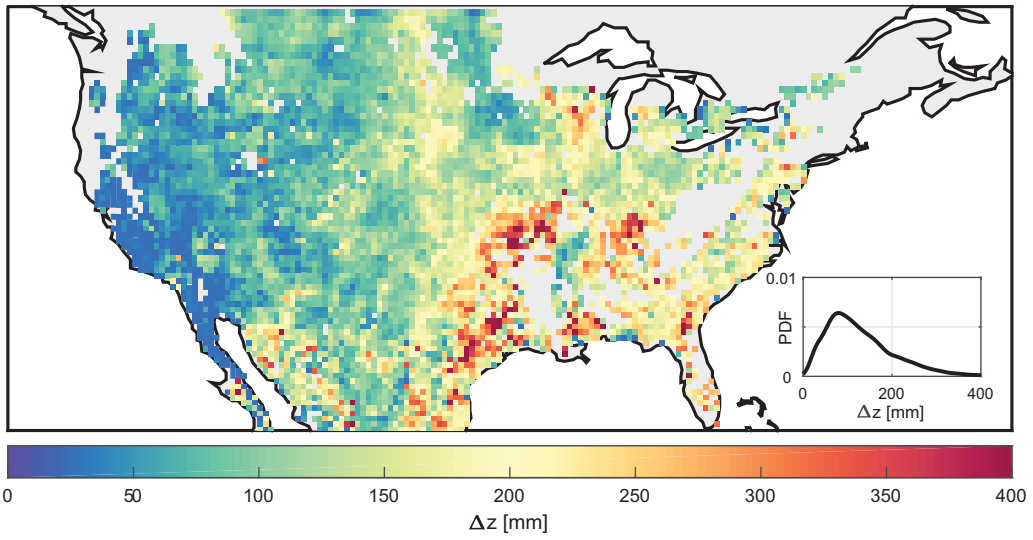

Figure 6. Map of the effective active hydrological length scale Δz (mm) over CONUS. Inset plot shows the probably distribution of the estimated length scale. Less than 12% of pixels have a depth of 50 mm or lower.

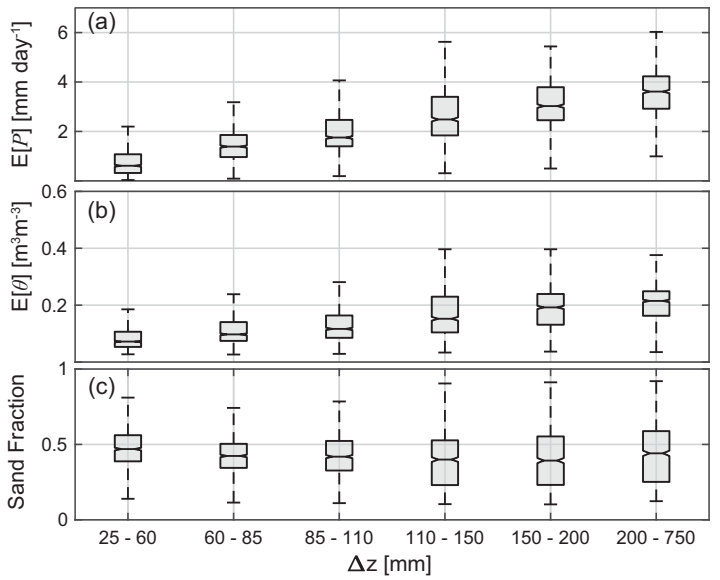

Figure 7. Classification of (a) mean precipitation (mm d^{-1}) (precipitation penetration depth forcing), (b) mean soil moisture (m^3 m^{-3}), and (c) sand fraction with respect to binned Δz (mm). (b) and (c) capture soil's hydraulic conductivity and texture characteristics. Spatially across CONUS, precipitation is the strongest determinant of Δz.

In section 1, Δz was presumed to be dependent on soil properties as well as precipitation forcing characteristics. Overall, we find that precipitation characteristics determine up to 50% of the spatial variance in Δz, while mean soil moisture and texture account for approximately 20%, and potential evaporation rates—from GLDAS—about 10% (figure not shown). In Figure 7 boxplots of (a) mean precipitation, (b) mean soil moisture—a proxy to unsaturated hydraulic conductivity, and (c) sand fraction with respect to binned Δz values are given. Soil texture, in the form of sand fraction, has a lesser influence on Δz. Across all Δz bins, the median sand fraction follows a shallow "U" shape. For every low Δz values, the median sand fraction is 45%, dropping to 40% for intermediate Δz then slightly increasing again.

The estimated landscape water storage ($\Delta z \cdot \theta_{SMAP}$) can now be used in water balance studies with fluxes (e.g., precipitation) balancing change in water storage (depth of water in the soil or volume of soil water per unit surface area). As a test of this statement, the methodology is applied to in situ soil moisture observations where both the surface and the profile soil moisture are measured (Appendix A). The results using in situ measurements show that the surface soil moisture series and precipitation describe a landscape water balance once an applicable Δz is estimated using the methodology introduced in this study.

The spatial pattern of Δz is generally consistent with the correlation map of Figure 1, where R^2 between surface (5 cm) and root-zone soil moisture (at 10, 20, 50, 100 cm) is shown. Observe that, in general, Δz is larger in regions where at depth (up to 50 cm), the surface to root-zone soil moisture correlation is also large. Additionally, the east-west gradient of Δz across CONUS is similar to estimates of the effective column water holding capacity (mm) from Koster and Suarez (2001) but the overall magnitude is naturally smaller since the latter considers total stored water content.

As a measure of the reliability of (4), the error standard deviation between SMAP observations and (4) are shown in Figure 8a. The median error standard deviation across CONUS is 0.06 (–) and the largest errors are concentrated in croplands. The latter is partly due to frequent and sporadic precipitation events undesirably affecting the ability to detect suitable soil moisture dry-down. Since the period of study is concurrent with seasonal crop growth, not only rapid changes in vegetation amount are possible, but high rates of moisture uptake are expected (Konings & Gentine, 2017). The explained-variance, is shown in Figure 8b. R^2 is largest in the Great Plains, especially the southern plains ($>$0.7). It is lowest in the dry western and southwest parts of the US coincident with smaller estimates of Δz. The median R^2 across CONUS is 0.65.

3.2. Loss Function Performance

As discussed in section 2, to complete the water balance in (1), in addition to soil moisture and precipitation, the soil water loss function must be known. In section 2.2, a piecewise loss function was reconstruction based on analysis of consecutive negative increments of soil moisture and defined in (3). As seen in Figure 4b, $L(\theta)$, is partitioned into four segments: A: dry-range extrapolation, B: observation-based domain, C: slope-estimation, and D: saturation and beyond.

Since $L(\theta)$ is partially observation-based, it is important to examine its performance within the proposed hydrological framework of (1), (3), and (4). More specifically, estimates of the length scale Δz may become questionable, if (4) operates outside the observation-based domain (Segment B) too often. If the extreme dry-end or wet-end extrapolation components (Segments A, C, or D) are utilized more than Segment B often, then Δz is a superfluous parameter.

We use both percentage-of-time and fraction of the moisture flux within each loss function segment as diagnostics. To examine this behavior, Figures 9a and 9b show plots of the percent-of-time the estimates in (4) are evaluated within the four segments. Figure 9a indicates a median of 80% of time across CONUS the model is within the observation-based domain and approximately 15% of time outside this region

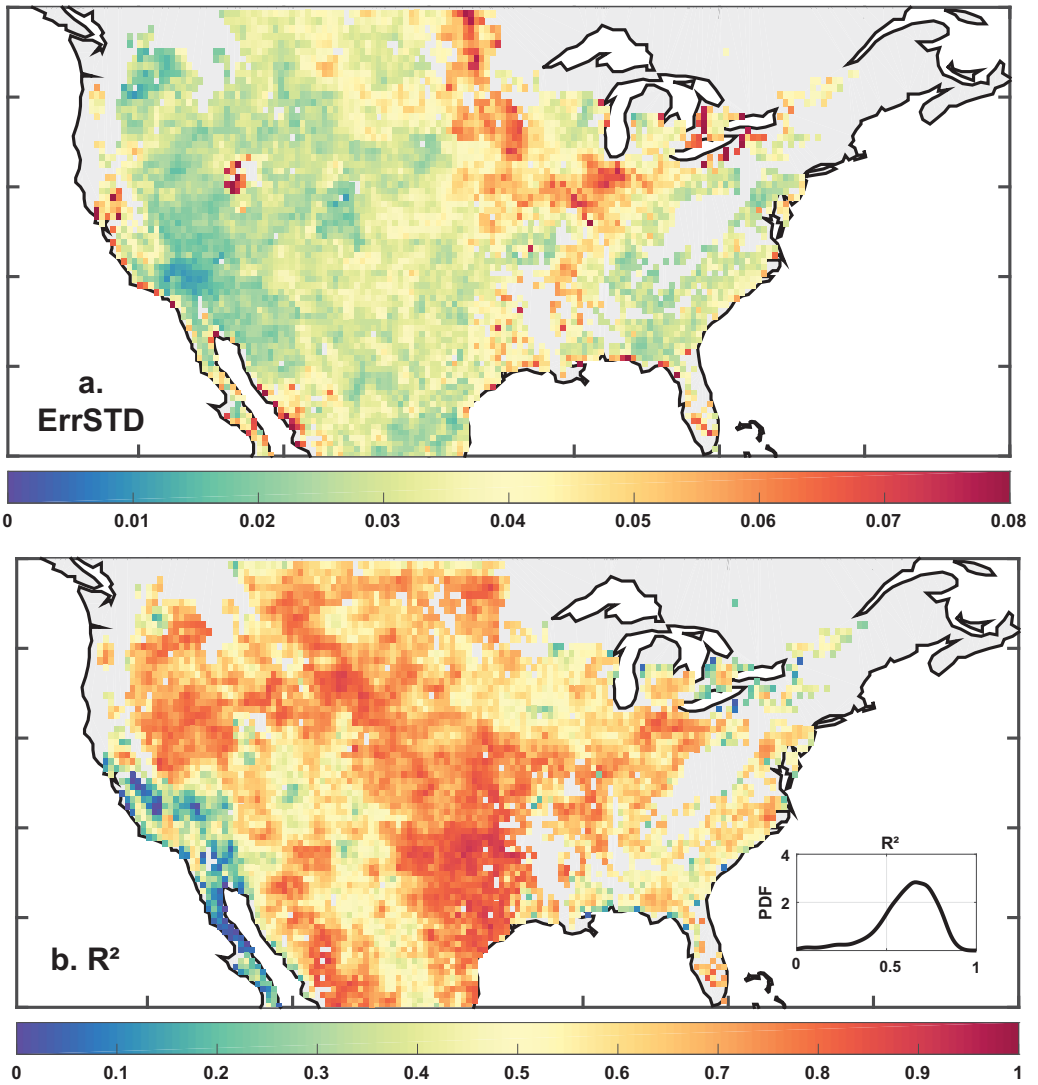

Figure 8. Map of (a) the error standard deviation between SMAP observations and estimates from (4), and (b) the corresponding correlation coefficient, R^2. Inset plot in (b) shows the R^2 distribution with a peak about 0.65. Largest errors are concentrated around croplands in Midwest US.

(extrapolation Segments A, C, and D). Western and Southwest US, especially California, show a large percentage-of-time spent in the dry-end interpolation domain (breakdown of A, C, and D regions is not shown).

Figures 9c and 9d are similar to above, but now show total losses with respect to total precipitation. The percentages are determined by normalizing accumulated losses within each segment by the total precipitation during the 2 years, i.e., $\frac{\sum \Delta z \cdot L(\theta)}{\sum P(t)} * 100$. The figure also highlights an additional insight. Within the dynamic range of SMAP soil moisture, Segment B, approximately 70% of the terrestrial water cycle can be captured. Furthermore, by applying water balance, using precipitation, an active soil layer of depth Δz can be characterized.

4. Discussion

The ability to perform water balance using remotely sensed estimates of surface volumetric soil moisture is the principal motivation behind finding a length scale. This scale factor transforms surface soil moisture estimates—in volumetric units—to the depth of soil water within the landscape—in length units. Mass input

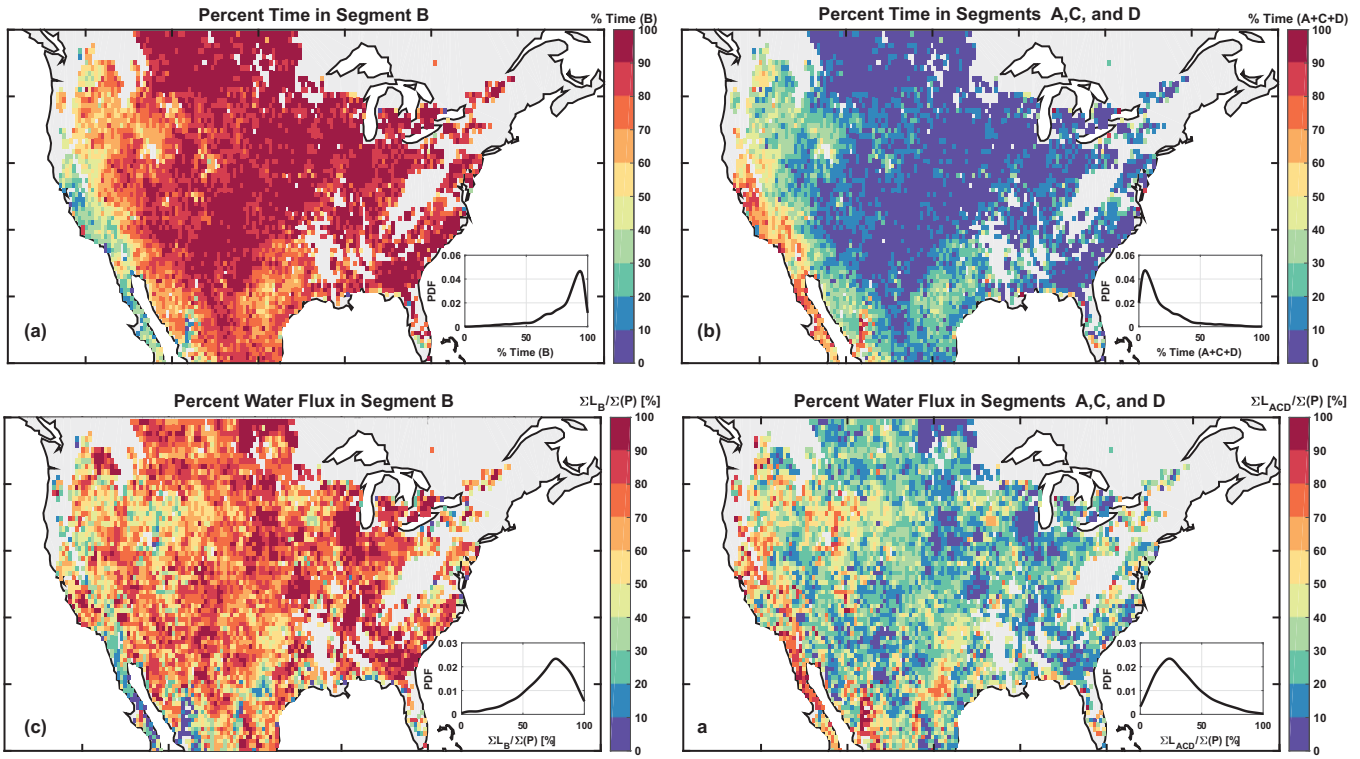

Figure 9. (a) Percent-of-time hydrological model is evaluated in the observation-based loss function domain, i.e., Segment B, with a median of 80%. (b) Percent-of-time hydrological model is either in the dry-end interpolation (e.g., California), slope-estimation domain, or beyond saturation, i.e., Segments A, C, and D, with a median of 15% (c) similar to (a) but for percent-of-total losses with respect to total precipitation $\frac{\sum \Delta z \cdot L(\theta)}{\sum P} \cdot 100$. A median of 70% of the water cycle is within the observation-based domain (d) similar to (b) but showing percent-of-total losses with respect to total input precipitation with a median of 27%. Figures 9a and 9c indicate that more than 70% of the dynamics of the terrestrial water cycle, in the form of surface soil moisture, is captured by SMAP.

into the landscape in the form of precipitation and losses in the form of evapotranspiration and recharge/discharge balance the change in volume of water per unit area (a depth of water). This volume can only be characterized once the soil volumetric water content is combined with a corresponding length scale Δz. Additionally, this length scale should be defined such that it is consistent with the incident precipitation and the observed volumetric soil water content dynamics. Therefore, the identification such a length scale allows usage of surface volumetric soil water content products—such as those produced by SMAP—in water balance studies that begin with local precipitation.

The need for this scaling factor is evident in studies of catchment water balance. Microwave remote sensing observations of surface soil moisture have experimentally been shown to correlate well with gravity-based total water storage estimates (Abelen et al., 2015). At medium scale basins throughout the US ($<$10,000 km^2), Crow et al. (2017) demonstrated that a statistically significant closure in the annual water budget can be obtained when surface-only soil moisture is multiplied by a storage scale factor, e.g., $\kappa \cdot \frac{d\theta}{dt}$. Such a scaling factor is required to transform soil moisture from volumetric units to water content [L] which is the extensive storage variable required for water budget analysis. In Crow et al. (2017), the parameter κ [L] was estimated as the ratio between the variance of change in columnar water volume—from gravity-based observations—to the variance of change in remotely sensed volumetric surface soil water content. Using this storage scale factor κ [L] they obtained high correlations between annual sums of fluxes ($P - Q$) and scaled surface-only soil moisture storage. By application of the methods described in this study over longer periods of time, a long-term estimate of Δz can be obtained, and then similar to Crow et al. (2017), used to examine basin level water balance closure. Moreover, this length scale Δz can be estimated globally (using the methodology introduced in this study) where ground-based gravity anomaly measurements are not available everywhere.

Several studies that have used time series of soil moisture to improve estimates of or fully estimate precipitation have found the need for a scaling factor to apply to the volumetric soil water content series (Brocca et al., 2013, 2014; Crow et al., 2009; Koster et al., 2016). Crow et al. (2009) show improvement in precipitation estimates by application of a multiplicative correction on remotely sensed precipitation accumulations to overcome potential errors in precipitation observations as well as soil water losses within their Antecedent Precipitation Index (API) model. The correction term included a time-constant scaling factor λ (–). This factor was either predetermined or calibrated with respect to independent precipitation data sets and later translated to length units [*L*] via an observation operator within a Kalman Filter. Similarly, Brocca et al. (2014) demonstrates global rainfall estimation, i.e., the SM2RAIN algorithm, using remotely sensed soil moisture within the context of water balance similar to (1). SM2RAIN parameterizes the loss terms in (1) as a parameterized power-law function of soil moisture. The SM2RAIN algorithm also requires a soil depth later Z* [*L*] prior to rainfall estimation. Brocca et al. (2014) estimates this parameter via calibration of the SM2RAIN model with respect to benchmark precipitation products—estimates of Z* were, in general, positively correlated with respect to increasing precipitation. In contrast to these prior studies, estimates of Δz are obtained within a self-consistent water balance framework.

The emphasis of this work is on observation-driven approaches to provide first-order understanding and estimates of water balance. Note that in section 2, the loss function estimation process uses only soil moisture observations and precipitation information. Importantly, the loss function is estimated from SMAP soil moisture product alone in a nonparametric fashion. No model (e.g., API or parameterized power-law, etc.) are used to estimate the interstorm loss processes. A few obvious caveats exist. First, partitioning the total losses into individual components (evaporation, drainage, possible runoff, etc.) is not possible when using soil moisture and precipitation series alone. This can be addressed by introducing modeling elements as well as additional ancillary data. Furthermore, the applicability and representativeness of the loss function is limited only to the period of study (May–September 2015 and 2016). Implementation and incorporation of this method in to climate models, for example, require additional investigation. In wetter regions, precipitation is a dominant control on the length scale Δz. The influence of seasonal and annual precipitation characteristics on estimated Δz can be examined by application of the methods in section 2 over longer periods of time. We expect, at a given location, Δz will slightly change over time, primarily due to precipitation.

5. Conclusions

The surface soil moisture state often shares mutual information and is well correlated with deep soil water content. By characterizing a water balance control volume with precipitation and surface-only soil moisture dynamics, a unique and representative effective depth, or length scale, can be determined. SMAP surface-only soil moisture observations and precipitation information are used to estimate this effective hydrologic depth. Across CONUS the median Δz is 135 mm. It shows an east-west divide with larger length scales for wetter regions with higher mean precipitation. In dry western parts of the country, especially Californian, the length scale is small due to little hydrological activity—near-zero mean precipitation—and is representative of a higher degree of decorrelation between surface and deeper soil moisture dynamics. Precipitation alone explains up to 50% of the spatial variance in Δz across the US. Based on analysis of soil moisture drydowns and successive negative increments of the soil moisture time series, an observation-based moisture loss function estimation technique was also presented. This method conditions the rate of loss of moisture $\frac{\Delta\theta^-}{\Delta t^{obs}}$ on soil moisture itself, in order to arrive at the loss function, $L(\theta) = E\left[\frac{-\Delta\theta^-}{\Delta t^{obs}} \mid \theta\right]$. Approximately 70% of the terrestrial water cycle dynamics occur within the range of SMAP soil moisture observations such that an active soil layer Δz can be characterized when water balance is applied to precipitation and SMAP soil moisture.

Appendix A: Total Stored Water Dynamics Reflected in the Surface Soil Moisture State

This appendix presents a complementary analysis supporting the linkage between the surface soil moisture state and deeper soil water content and its dynamics. Specifically, we examine the degree to which the total landscape soil water storage [*L*], from surface to depth Δz is correlated with surface-only soil moisture when it is scaled by a factor Δz, i.e., $w_{est}(t) = \Delta z \cdot \theta_{05cm}(t)$. This portion of the study is implemented using

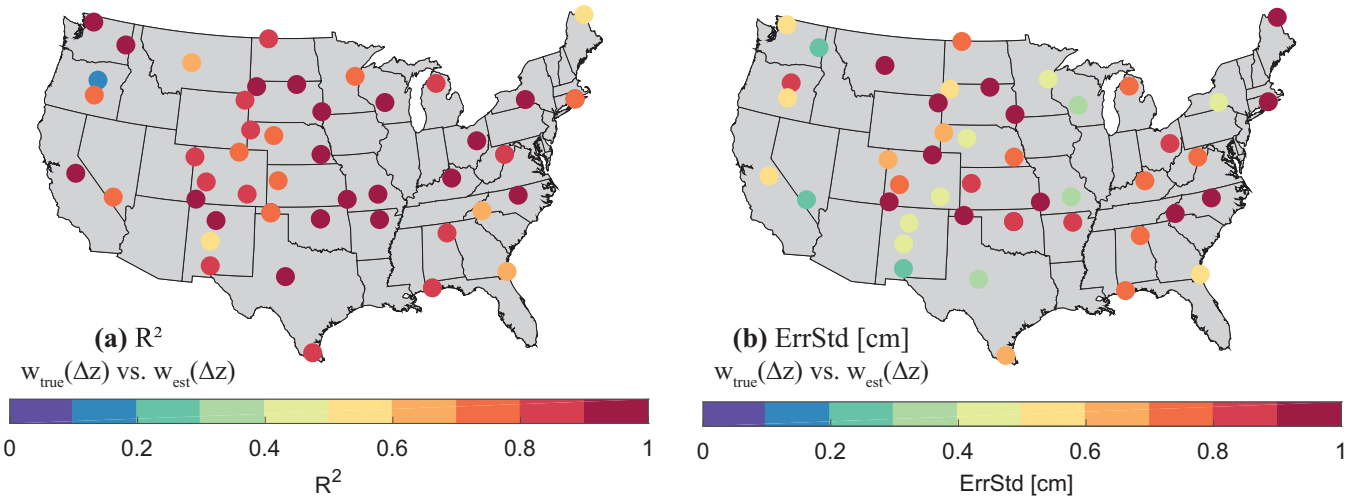

Figure A1. (a) Correlation coefficient, R^2, between w_{true} and w_{est}. (b) Same as Figure A1a but for error standard deviation (cm). Station with missing data, or gaps, in the soil moisture time series at all depths are excluded.

in situ profile soil moisture (5, 10, 20, 50, and 100 cm) from the US Climate Reference Network (USCRN) (Bell et al., 2013).

First, the entire length scale estimation process outlined in section 2 is applied to USCRN in situ surface soil moisture data, θ_{05cm}. That is, first the loss function in (2)–(3) is reconstructed and used to estimate the characteristic length scale, but at a point level. For each location, CPCU precipitation data are used. Over the period of study (May–September 2015 and 2016) only a subset of USCRN station have complete data records at all depths (approximately 50 sites). Locations with missing data, or gaps, in the temporal soil moisture time series are excluded. Additionally, similar to SMAP observations, data are sampled every 3 days at 0600 local time.

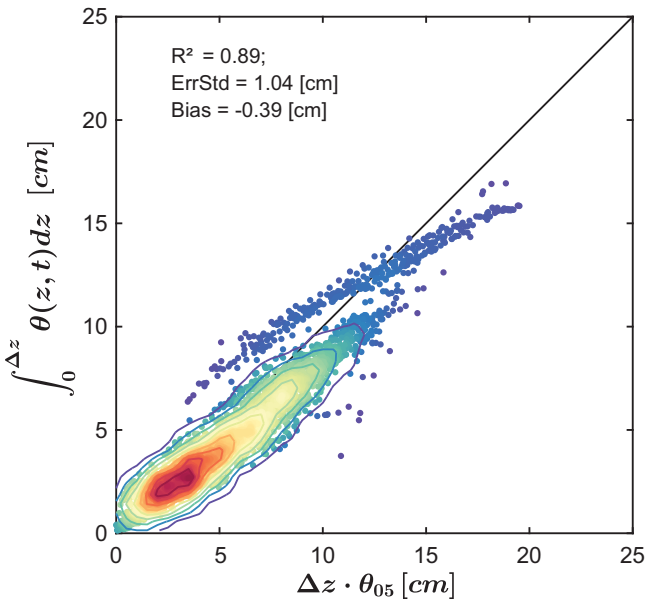

Figure A2. Scatter plots of w_{true} and $\Delta z \cdot \theta_{05cm}$, for all valid USCRN site combined (50 sites, May–September 2015 and 2016). Contours indicate the density of data points. R^2 is 0.89 indicating that the total integrated water content dynamics—from 0 to Δz—are well reflected in the scaled surface soil moisture $\Delta z \cdot \theta_{05cm}$. Each USCRN site has its own associated length scale.

For each valid USCRN site, once estimates of Δz are obtained, we calculate the total columnar stored soil water content from the surface down to a depth Δz, $w_{true}(t, \Delta z)$ (cm). The profile soil moisture state is integrated from 0 to Δz (cm):

$$w_{true}(t) = \int_0^{\Delta z} \theta(t,z)dz \tag{A1}$$

The estimated $w_{est}(t, \Delta z)$ and $w_{true}(t, \Delta z)$ are related through

$$w_{true}(t) = w_{est}(t) + \epsilon \tag{A2}$$

where ϵ reflects measurement errors as well as representativeness errors caused by deviations of $w_{true}(t)$ from $w_{est}(t) = \Delta z \cdot \theta_{05cm}(t)$.

Figure A1 shows the correlation coefficient and error standard deviation statistics between $w_{true}(t)$ and $w_{est}(t,)$. The figures show that $w_{true}(t)$ and $w_{est}(t)$ are well correlated with at most 1 cm error standard deviation. In Figures A1a with the exception of a few sites, $R^2 > 0.65$. Note that each location is associated with a unique Δz (figure not shown).

When all valid USCRN site are combined together the scatter plot in Figure A2 between $w_{true}(t, \Delta z)$ and $w_{est}(t, \Delta z)$ is obtained. The figure demonstrates the strong linkage and very high correlation ($R^2 \approx 0.89$) between the integrated soil moisture profile from (A1), and the corresponding scaled surface-only soil moisture from (2). When considering

all sites, a negative bias exists, −0.4 cm, with an error standard deviation of approximately 1 cm. Figures A1 and A2 demonstrate that within conceptual control volume, a unique length scale exists such that the total integrated soil moisture content (form the surface to this depth) is well reflected in surface-only soil moisture dynamics.

Acknowledgments
The authors acknowledge funding from NASA in the form of a sponsored research grant. K.A.M. is funded by a Ziff Environmental Fellowship from Harvard University's Center for the Environment. E.H. is funded by the Swiss National Science Foundation (SNSF grant P2EZP2–165244). All data used in this study are publicly available. Soil moisture data from SMAP are available for download from the National Snow Ice Data Center (NSIDC) Distributed Active Archive Center (DAAC) at https://nsidc.org/data/smap/smap-data.html. Climate GLDAS data are available for download from https://disc.sci.gsfc.nasa.gov/. Climate Prediction Center Unified Gauge-Based Global daily precipitation can be downloaded from https://climatedataguide.ucar.edu/climate-data/cpc-unified-gauge-based-analysis-global-daily-precipitation Soil Moisture from the U.S. Climate Reference Network (USCRN) are available at https://www.ncdc.noaa.gov/crn/qcdatasets.html.

References

Abelen, S., Seitz, F., Abarca-del-Rio, R., & Güntner, A. (2015). Droughts and floods in the La Plata Basin in soil moisture data and GRACE. *Remote Sensing*, *7*(6), 7324–7349. https://doi.org/10.3390/rs70607324

Albergel, C., Pellarin, T., Fritz, N., Froissard, F., Petitpa, A., Piguet, B., et al. (2008). From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. *Hydrology and Earth System Sciences Discussions*, *12*, 1323–1337.

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., et al. (2013). U.S. climate reference network soil moisture and temperature observations. *Journal of Hydrometeorology*, *14*(3), 977–988. https://doi.org/10.1175/JHM-D-12-0146.1

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., & Hasenauer, S., et al. (2014). Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *Journal of Geophysical Research: Atmosphere*, *119*, 5128–5141. https://doi.org/10.1002/2014JD021489

Brocca, L., Moramarco, T., Melone, F., & Wagner, W. (2013). A new method for rainfall estimation through soil moisture observations. *Geophysical Research Letters*, *40*, 853–858. https://doi.org/10.1002/grl.50173

Chan, S. K., Bindlish, R., O'neill, P. E., Njoku, E., Jackson, T., & Colliander, A., et al. (2016). Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4994–5007. https://doi.org/10.1109/TGRS.2016.2561938

Crow, W. T., Chen, F., Reichle, R. H., & Liu, Q. (2017). L band microwave remote sensing and land data assimilation improve the representation of prestorm soil moisture conditions for hydrologic forecasting. *Geophysical Research Letters*, *44*, 5495–5503. https://doi.org/10.1002/2017GL073642

Crow, W. T., Han, E., Ryu, D., Hain, C. R., & Anderson, M. C. (2017). Estimating annual water storage variations in medium-scale (2000–10000 km^2) basins using microwave-based soil moisture retrievals. *Hydrology and Earth System Sciences*, *21*(3), 1849–1862. https://doi.org/10.5194/hess-21-1849-2017

Crow, W. T., Huffman, G., Bindlish, R., & Jackson, T. J. (2009). Improving satellite-based rainfall accumulation estimates using spaceborne surface soil moisture retrievals. *Journal of Hydrometeorology*, *10*(1), 199–212. https://doi.org/10.1175/2008JHM986.1

Das, N. (2013). *SMAP ancillary data report: Soil attributes*. Pasadena, CA: Jet Propulsion Laboratory, California Institute of Technology.

Dickinson, R. E. (1988). The force-restore model for surface temperatures and its generalizations. *Journal of Climate*, *1*(11), 1086–1097.

Dirmeyer, P. A. (2000). Using a global soil wetness dataset to improve seasonal climate simulation. *Journal of Climate*, *13*(16), 2900–2922.

Entekhabi, D., Njoku, E. G., O'neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Fatichi, S., Pappas, C., & Ivanov, V. Y. (2016). Modeling plant-water interactions: An ecohydrological overview from the cell to the global scale. *Wiley Interdisciplinary Reviews: Water*, *3*(3), 327–368. https://doi.org/10.1002/wat2.1125

Feng, X., Vico, G., & Porporato, A. (2012). On the effects of seasonality on soil water balance and plant growth. *Water Resources Research*, *48*, W05543. https://doi.org/10.1029/2011WR011263

Ford, T. W., Harris, E., & Quiring, S. M. (2014). Estimating root zone soil moisture using near-surface observations from SMOS. *Hydrology and Earth System Sciences*, *18*(1), 139–154. https://doi.org/10.5194/hess-18-139-2014

Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E., & Cardenas, M. B. (2015). The global volume and distribution of modern groundwater. *Nature Geoscience*, *9*, 161–167. https://doi.org/10.1038/ngeo2590

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. https://doi.org/10.1016/j.rse.2014.08.030

Hu, Z., & Islam, S. (1995). Prediction of ground surface temperature and soil moisture content by the force-restore method. *Water Resources Research*, *31*, 2531–2539.

Ingber, L., Petraglia, A., Petraglia, M. R., & Machado, M. A. S. (2012). Adaptive simulated annealing. In Aguiar e Oliveira Jr., H. J. (Ed.), *Stochastic global optimization & its applications* (pp. 33–62). Berlin, Heidelberg: Springer. Retrieved from https://link.springer.com/content/pdf/10.1007%2F978-3-642-27479-4_4.pdf

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J. M., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1729–1735. https://doi.org/10.1109/36.942551

Konings, A. G., & Gentine, P. (2017). Global variations in ecosystem-scale isohydrocity. *Global Change Biology*, *23*(2), 891–905. https://doi.org/10.1111/gcb.13389

Koster, R. D. (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., Brocca, L., Crow, W. T., Burgin, M. S., & De Lannoy, G. J. M. (2016). Precipitation estimation using L-band and C-band soil moisture retrievals. *Water Resources Research*, *52*, 7213–7225. https://doi.org/10.1002/2016WR019024

Koster, R. D., Reichle, R. H., & Mahanama, S. P. P. (2017). A data-driven approach for daily real-time estimates and forecasts of near-surface soil moisture. *Journal of Hydrometeorology*, *18*(3), 837–843. https://doi.org/10.1175/JHM-D-16-0285.1

Koster, R. D., & Suarez, M. J. (2001). Soil moisture memory in climate models. *Journal of Hydrometeorology*, *2*(6), 558–570. https://doi.org/10.1175/1525-7541(2001)002<0558:SMMICM>2.0.CO;2

Koster, R. D., & Suarez, M. J. (2003). Impact of land surface initialization on seasonal precipitation and temperature prediction. *Journal of Hydrometeorology*, *4*(2), 408–423. https://doi.org/10.1175/1525-7541(2003)

Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., & Peters-Lidard, C. D. (2009). Role of subsurface physics in the assimilation of surface soil moisture observations. *Journal of Hydrometeorology*, *10*(6), 1534–1547. https://doi.org/10.1175/2009JHM1134.1

Kurc, S. A., & Small, E. E. (2007). Soil moisture variations and ecosystem-scale fluxes of water and carbon in semiarid grassland and shrubland. *Water Resources Research*, *43*, W06416. https://doi.org/10.1029/2006WR005011

Laio, F., Porporato, A., Fernandez-Illescas, C. P., & Rodriguez-Iturbe, I. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress IV. Discussion of real cases. *Advances in Water Resources*, *24*(7), 745–762. https://doi.org/10.1016/S0309-1708(01)00007-0

Mahmood, R., & Hubbard, K. G. (2007). Relationship between soil moisture of near surface and multiple depths of the root zone under heterogeneous land uses and varying hydroclimatic conditions. *Hydrological Processes*, *21*(25), 3449–3462. https://doi.org/10.1002/hyp.6578

McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., & Entekhabi, D. (2017). The global distribution and dynamics of surface soil moisture. *Nature Geoscience*, *10*(2), 100–104. https://doi.org/10.1038/ngeo2868

McColl, K. A., Wang, W., Peng, B., Akbar, R., Gianotti, D., Lu, H., et al. (2017). Global characterization of surface soil moisture drydowns. *Geophysical Research Letters*, *44*, 3682–3690. https://doi.org/10.1002/2017GL072819

National Center for Atmospheric Research (NCAR) (2017). *The climate data guide: CPC unified gauge-based analysis of global daily precipitation*. Boulder, CO: National Center for Atmospheric Research. Retrieved from https://climatedataguide.ucar.edu/climate-data/cpc-unified-gauge-based-analysis-global-daily-precipitation

O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., & Bindlish, R. (2016). *SMAP enhanced L3 radiometer global daily 9 km EASE-grid soil moisture, version 1*. Boulder, CO: Natural Snow and Ice Data Center. https://doi.org/10.5067/ZRO7EXJ8O3XI

Qiu, J., Crow, W. T., & Nearing, G. S. (2016). The impact of vertical measurement depth on the information content of soil moisture for latent heat flux estimation. *American Meteorological Society*, *17*(9), 2419–2430. https://doi.org/10.1175/JHM-D-16-0044.1

Qiu, J., Crow, W. T., Nearing, G. S., Mo, X., & Liu, S. (2014). The impact of vertical measurement depth on the information content of soil moisture times series data. *Geophysical Research Letters*, *41*, 4997–5004. https://doi.org/10.1002/2014GL060017

Reichle, R. H., Crow, W. T., Koster, R. D., Sharif, H. O., & Mahanama, S. P. P. (2008). Contribution of soil moisture retrievals to land data assimilation products. *Geophysical Research Letters*, *35*, L01404. https://doi.org/10.1029/2007GL031986

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., & Owe, M. (2007). Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research*, *112*, D09108. https://doi.org/10.1029/2006JD008033

Rodriguez-Iturbe, I., & Porporato, A. (2007). *Ecohydrology of water-controlled ecosystems: Soil moisture and plant dynamics*. Cambridge, UK: Cambridge University Press.

Rosenzweig, C., Tubiello, F. N., Goldberg, R., Mills, E., & Bloomfield, J. (2002). Increased crop damage in the US from excess precipitation under climate change. *Global Environmental Change*, *12*(3), 197–202. https://doi.org/10.1016/S0959-3780(02)00008-0

Sabater, J. M., Jarlan, L., Calvet, J.-C., Bouyssel, F., & De Rosnay, P. (2007). From near-surface to root-zone soil moisture using different assimilation techniques. *Journal of Hydrometeorology*, *8*(2), 194–206. https://doi.org/10.1175/JHM571.1

Salvucci, G. D. (2001). Estimating the moisture dependence of root zone water loss using conditionally averaged precipitation. *Water Resources Research*, *37*, 1357–1365.

Schwingshackl, C., Hirschi, M., & Seneviratne, S. I. (2017). Quantifying spatiotemporal variations of soil moisture control on surface energy balance and near-surface air temperature. *Journal of Climate*, *30*(18), 7105–7124. https://doi.org/10.1175/JCLI-D-16-0727.1

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*(3–4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Tabatabaeenejad, A., Burgin, M., Duan, X., & Moghaddam, M. (2015). P-band radar retrieval of subsurface soil moisture profile as a second-order polynomial: First AirMOSS results. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(2), 645–658. https://doi.org/10.1109/TGRS.2014.2326839

Tuttle, S., & Salvucci, G. (2016). Empirical evidence of contrasting soil moisture-precipitation feedbacks across the United States. *Science*, *352*(6287), 825–828. https://doi.org/10.1126/science.aaa7185

Tuttle, S. E., & Salvucci, G. D. (2014). A new approach for validating satellite estimates of soil moisture using large-scale precipitation: Comparing AMSR-E products. *Remote Sensing of Environment*, *142*, 207–222. https://doi.org/10.1016/j.rse.2013.12.002

Vivoni, E. R., Moreno, H. A., Mascaro, G., Rodriguez, J. C., Watts, C. J., Garatuza-Payan, J., et al. (2008). Observed relation between evapotranspiration and soil moisture in the North American monsoon region. *Geophysical Research Letters*, *35*, L22403. https://doi.org/10.1029/2008GL036001

Wagner, W., Lemoine, G., & Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sensing of Environment*, *70*(2), 191–207. https://doi.org/10.1016/S0034-4257(99)00036-X