Estimation of Landscape Soil Water Losses from Satellite Observations of Soil Moisture

RUZBEH AKBAR AND DANIEL J. SHORT GIANOTTI

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

KAIGHIN A. MCCOLL

Department of Earth and Planetary Sciences, and John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts

ERFAN HAGHIGHI

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

GUIDO D. SALVUCCI

Department of Earth and Environment, Boston University, Boston, Massachusetts

DARA ENTEKHABI

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Manuscript received 20 October 2017, in final form 12 April 2018)

ABSTRACT

This study presents an observation-driven technique to delineate the dominant boundaries and temporal shifts between different hydrologic regimes over the contiguous United States (CONUS). The energy- and water-limited evapotranspiration regimes as well as percolation to the subsurface are hydrologic processes that dominate the loss of stored water in the soil following precipitation events. Surface soil moisture estimates from the NASA Soil Moisture Active Passive (SMAP) mission, over three consecutive summer seasons, are used to estimate the soil water loss function. Based on analysis of the rates of soil moisture dry-downs, the loss function is the conditional expectation of negative increments in the soil moisture series conditioned on soil moisture itself. An unsupervised classification scheme (with cross validation) is then implemented to categorize regions according to their dominant hydrological regimes based on their estimated loss functions. An east–west divide in hydrologic regimes over CONUS is observed with large parts of the western United States exhibiting a strong water-limited evapotranspiration regime during most of the times. The U.S. Midwest and Great Plains show transitional behavior with both water- and energy-limited regimes present. Year-to-year shifts in hydrologic regimes are also observed along with regional anomalies due to moderate drought conditions or above-average precipitation. The approach is based on remotely sensed surface soil moisture (approximately top 5 cm) at a resolution of tens of kilometers in the presence of soil texture and land cover heterogeneity. The classification therefore only applies to landscape-scale effective conditions and does not directly account for deeper soil water storage.

1. Introduction

The profile of soil moisture—from surface to root zone—partially determines water and heat fluxes between the land and atmosphere depending on the dominant evapotranspiration regime (Seneviratne et al. 2010), affects the dynamics of soil respiration (Manzoni et al. 2012), and drives the growth of crops and natural vegetation

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-17-0200.s1.

Corresponding author: Ruzbeh Akbar, rakbar@mit.edu

DOI: 10.1175/JHM-D-17-0200.1

(Rosenzweig et al. 2002). Weather forecasting and climate modeling efforts, especially at large spatial scales, have long been recognized to require knowledge of the soil moisture profile to partition the components of the surface energy balance (Dirmeyer et al. 2000; Koster and Suarez 2003) The short-term evolution of weather and seasonal climate are affected by this partitioning (Vivoni et al. 2008; Seneviratne et al. 2010).

Recent satellite microwave remote sensing 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 have emerged that provide global estimates of soil moisture at about 40-km spatial resolution from daily to 3-day intervals. Operating within the L-band (1.41 GHz) regime of the microwave spectrum, these platforms are capable of sensing water content within the topsoil, and their soil moisture products typically represent an integrated depth of 50 mm, hence only surface soil moisture (SSM; θ). To overcome limitations imposed by the lack of frequent and global observations, a variety of assimilation and land surface modeling efforts have, in the past, been developed (Wagner et al. 1999; Sabater et al. 2007; Reichle et al. 2007; Albergel et al. 2008; Reichle et al. 2008; Kumar et al. 2009) to link and relate surface moisture to deeper soil moisture.

Even though SSM constitutes only a fraction of the total distribution of terrestrial water (Gleeson et al. 2015), it has profound implications on the water, carbon, and energy cycles. It rests at the critical boundary between land and the atmosphere and generally exhibits more temporal variability compared to deeper root-zone moisture. However, the dynamics of surface and root-zone soil moisture are not independent, and indeed share some mutual information. They are physically linked via advection and diffusion processes affected by soil texture, land cover, and vegetation type as well as general climate conditions. In addition, SSM and root-zone moisture are often highly correlated (Ford et al. 2014; Qiu et al. 2014), with some exceptions, such as very dry conditions (Hirschi et al. 2014). Thus, in many applications a strong relationship between surface and root-zone moisture and fluxes exists. In particular, analysis of flux tower data has shown that surface soil moisture's information content for predictive water and energy flux partitioning is equivalent to that of soil moisture integrated over a larger depth on the time scale of remotely sensed observations (2–3 days; Qiu et al. 2016).

Soil moisture dynamics determine landscape hydrological regimes and water/energy limitation through control of surface energy partitioning. Frequent global remotely sensed estimates of SSM allow us to identify these hydrological regimes by characterizing dry-down and water-loss processes.

Over large areal scales, the storage and dynamics of soil moisture within a homogeneous active layer of depth Δz (L) can be described as

$$\Delta z \frac{d\theta}{dt} = I(t) - \mathrm{ET}(\theta) - D(\theta) = I(t) - L(\theta)\,, \qquad (1)$$

where θ is the volumetric water content, I is infiltrating precipitation flux (L T^{-1}) here set equal to actual precipitation (L T^{-1}), ET is evapotranspiration (vegetation transpiration and bare soil evapotranspiration) loss (L T^{-1}), and D is the drainage loss from the volume (L T^{-1}). The total water losses L are defined as the sum of drainage or leakage out of the soil and evapotranspiration (ET + D). Parameter L is a function of the coupled atmospheric and terrestrial processes acting on surface soil moisture; in this study, however, we specifically investigate the primary control of water availability on the loss function under mean climate state, $L = L(\theta)$. Using remotely sensed observations of surface soil moisture, $L(\theta)$ can be estimated to first order at a given location. Therefore, $L(\theta)$ is a function that will be identified based on the entire available record of soil moisture.

Probabilistic modeling of soil water balance shows that the exact form of $L(\theta)$ is a key determinant of water balance dynamics (Dralle and Thompson 2016; Feng et al. 2012). Note that this conceptualization applies to average losses over large scales—on the order of tens of kilometers relevant to weather and climates model as in Dirmeyer (2000)—in which net lateral fluxes are small compared to vertical infiltration, evapotranspiration, and drainage losses.

An idealized soil water loss function is typically characterized as a three-stage piecewise function (Rodriguez-Iturbe and Porporato 2007; Feng et al. 2017) as seen in Figs. 1a and 1b. Soil moisture dynamics following a rain event depend on the prior history of soil moisture and its current state. During and after precipitation the wet soil initially loses water rapidly due to percolation to the subsurface. If under gravity drainage, the loss is initially governed by the hydraulic conductivity and gradient in hydraulic head, both of which are functions of soil volumetric water content. Drainage dominates until the volumetric water content is lowered to the field capacity θ_{fc}, at which point drainage becomes negligible. Post-precipitation, the evapotranspiration from moist soils is considered to be initially in "Stage I," or an energy-limited regime. During this period, evapotranspiration is independent of volumetric water content, and hence the contribution to the loss function from evapotranspiration is determined by the potential evapotranspiration rate such that $d\theta/dt$ is constant under steady weather conditions.

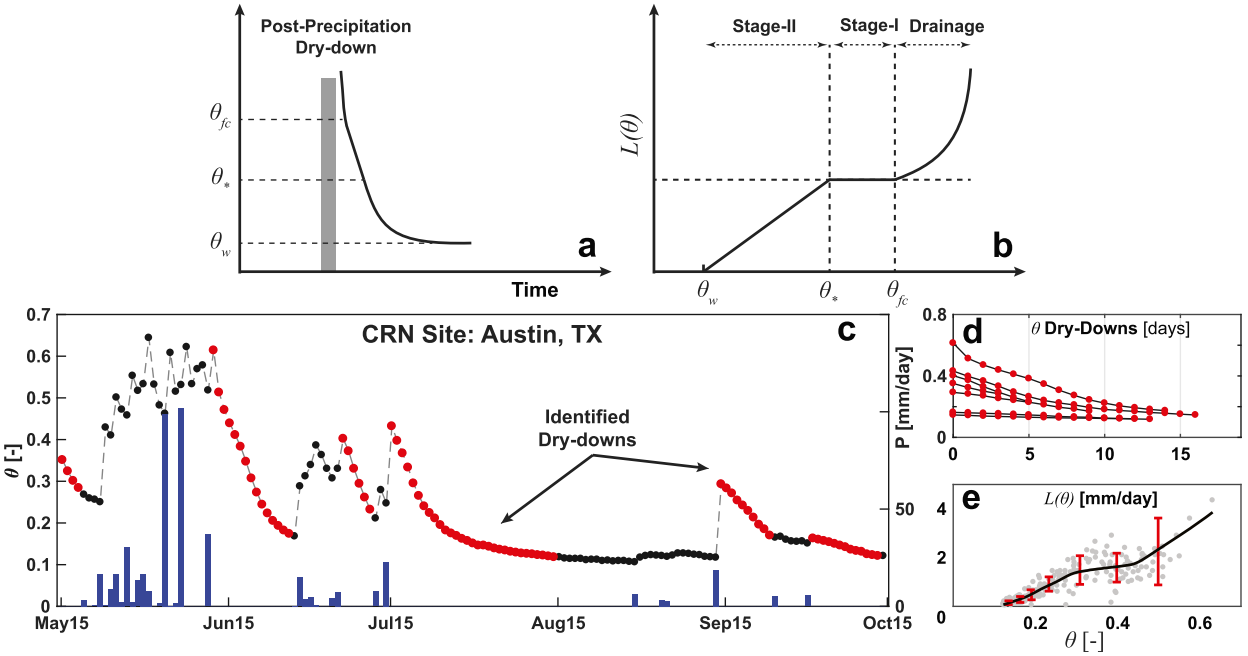

FIG. 1. (a) Canonical three-stage soil moisture dry-down. During interstorms (postprecipitation) the dry-down rate occurs at a rapid rate under drainage until the volumetric soil water content reaches its field capacity. After this time the soil is still moist enough to be under the energy-limited (Stage I) evapotranspiration regime until the volumetric soil water content reaches a critical value θ_*. During this period, the evapotranspiration rate equals the climatic potential evapotranspiration, and the dry-down rate is closer to linear. Below the critical θ_*, the dry-down rate is controlled by soil moisture (Stage II or water-limited evapotranspiration regime). The rate of dry-down is below linear and ultimately diminishes to zero. (b) Loss function: Stage II evapotranspiration for moisture levels below the critical value θ_*, Stage I evapotranspiration and no sensitivity to soil moisture for intermediate wetness levels ($\theta_* \le \theta \le \theta_{fc}$), and percolation or drainage dominant regime above the field-capacity θ_{fc}. In (b) the horizontal dashed line represents the mean climatological potential evapotranspiration rate dependent on available energy and atmospheric evaporative demand (Rodriguez-Iturbe and Porporato 2007). (c) An example in situ USCRN soil moisture time series sampled at 0600 LT daily (Austin, TX). (d) Soil moisture dry-downs, with zero precipitation in between observations are identified. The dry-down cutoff after 15 Aug is due to near-zero increments and is smaller than the detection threshold. (e) The loss function $L(\theta)$ is the negative gradient of these dry-downs, with gray dots indicating an observed $\Delta\theta$ dry-down, i.e., the difference between neighboring red dots in (c). The times series plot in (c) shows 2015 only, while the side plots (d) and (e) include data from 2016 and 2017 as well.

The soil may dry further until it reaches a threshold value θ_*, a transition point between water- and energy-limited hydrologic regimes. Any further loss, thereafter, is due to evapotranspiration in the water-limited regime, termed "Stage II." Below θ_*, evaporative loss is a monotonically decreasing function with decreasing soil moisture. The drainage and Stage I regimes can coexist leading to a distinct condition, and it is addressed in this study.

The ability to reconstruct the soil moisture loss function and to identify different hydrologic regimes solely based on observations shows promise. The approach complements prior methods that rely on model outputs. For example, using atmospheric general circulation models (AGCMs) and placing prior transition criteria on soil moisture and evaporations rates, Koster et al. (2009) identified different hydrological regimes globally for the Northern Hemisphere summer (JJA) over a 10-yr record. More recently, Schwingshackl et al. (2017) examined spatiotemporal variations in land–atmospheric coupling and hydrologic regimes using different reanalysis data [i.e., Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), and Global Land Evaporation Amsterdam Model (GLEAM)]. They proposed a model-selection process to arrive at piecewise-linear functions relating soil moisture to model-based estimates of evaporative fraction (EF) and similar to Koster et al. (2009) generated seasonal global maps of different hydrologic regimes.

In contrast, the analysis outlined in this manuscript is purely data driven and uses only remotely sensed soil moisture information. First, the soil water loss function $L(\theta)$ is reconstructed via analysis of soil moisture dry-downs over the continental United States under spring and summertime climate conditions (May–September). Then, the loss function and its transition points (θ_w, θ_*, θ_{fc}) are used to partition SMAP-observed soil moisture into different dominant hydrological regimes, that is, water or energy limited.

Naturally, when leveraging data-only techniques, certain simplifying assumptions must be made, and some limitations are inherent. SMAP reports surface soil moisture on the order of tens of kilometers. This is due to the effective field of view of SMAP constrained by the instruments' half-power, or −3 dB, sampling area on the surface of Earth (approximately 30–40 km). Therefore, subpixel landscape heterogeneity—both in vegetation and soil texture types—will be averaged into an effective landscape surface soil moisture value. This feature may also mask, or diminish, finer-scale processes, for example, subpixel presaturation runoff.

In the water balance equation of (1), all soil water losses are jointly considered. In this context, individual contributing fluxes, that is, evaporation, transpiration, or even runoff, cannot be separated. These processes are not "seen" by the soil moisture observations and require external modeling or ancillary data to capture or approximate. This is because SMAP only senses surface soil moisture (approximately 5 cm but variable depending on the soil moisture and temperature profile) dynamics. Furthermore, soil water losses are, in general, affected by a range of soil texture, vegetation, and climate conditions. In fact, at the larger spatial scales under consideration, these factors are all inherent within the soil moisture dry-down process. Individual dry-down characteristics for a fixed location will be different and will reflect varying atmospheric and vegetation conditions. From a remotely sensed soil moisture perspective, these factors are unobserved, considered latent processes, but are all encoded within the dry-down process. Therefore, similar to Koster et al. (2009), we consider the average, or effective, behavior over the period of study and at the observed SMAP spatial resolution. The expectation to address each individual driving factor is beyond the current scope of the manuscript.

We aim to present a simple methodology to maximally extract information from the soil moisture time series and then identify different hydrologic regimes. The use of ancillary data or models is avoided. This is to prevent "signatures" of ancillary data or models biasing outputs or overshadowing information inherent in the soil moisture time series. Additionally, we place no prior constraints or criteria on transition points as they will be inferred from observations. Currently, the use of these parameters as inputs into land surface models (LSMs) or global climate models (GCMs) is not recommended. By extension of the method globally, and for longer periods of time, this feature can be examined more carefully.

Section 2 describes the source soil moisture and supporting data used in this analysis. Section 3 first discusses the loss function estimation process—via analysis of soil moisture dry-downs—and then presents an unsupervised classification approach, with cross validation, which labels different regions as either being in a water-limited, energy-limited, or transitional regime determined by its overall loss function and a "best-fitting" canonical model. Results and analysis are presented in section 4, where we examine temporal dynamics of different hydrologic regimes. The time frame of the study includes three consecutive Northern Hemisphere spring and summer seasons (May–September 2015, 2016, and 2017) at about 36-km resolution.

2. Data sources

a. Soil moisture and precipitation data

Global surface soil moisture estimates are available from the polar-orbiting NASA SMAP satellite mission (Entekhabi et al. 2010) launched in early 2015. To date, SMAP provides soil moisture estimates with a nominal 1–3-day global coverage depending on latitude. SMAP-derived soil moisture data products undergo frequent and extensive calibration and validation (Chan et al. 2016) in the form of large-scale analysis with respect to many in situ and ground-based soil moisture observing networks. Since SMAP operates at L band (1.41 GHz), soil moisture estimates typically represent an integrated depth of moisture 50 mm within the topsoil. This study uses the descending overpass or 0600 local time SMAP 9-km enhanced radiometer-only soil moisture product (O'Neill et al. 2016). Data from the first three years of operation of SMAP, spanning the Northern Hemisphere's spring and summer months, that is, May–September 2015, 2016, and 2017, are combined and used to derive estimates of the loss function.

To minimize the impact of retrieval uncertainty, data quality control is performed prior to the analysis. Pixels affected by radio frequency interference (RFI), frozen soils, or snow cover, as well as those with larger than 1% water fraction, are removed. In addition, regions with high vegetation density and water content ($>7\,\mathrm{kg\,m^{-2}}$) are also excluded. The latter is due to the high uncertainty in retrieving soil moisture under these conditions. Water fraction and vegetation water content information are readily available from SMAP and are extracted from the soil moisture product, that is, L3SMP_E. Note that the accuracy of soil moisture data is typically inversely proportional to attenuation due to vegetation, that is, vegetation water content, and SMAP places a $5\,\mathrm{kg\,m^{-2}}$ threshold on soil moisture products. In this study, we increased the vegetation threshold to $7\,\mathrm{kg\,m^{-2}}$ to expand the spatial domain of the study and also because the exact threshold level is application dependent.

Daily surface precipitation estimates (mm $\mathrm{day^{-1}}$) are obtained from the Climate Prediction Center's Unified (CPCU) gauged-based daily precipitation dataset

(NCAR 2017). CPCU precipitation is reprocessed from a 0.5° grid (approximately 27 km at the equator) to the SMAP 9-km grid and is utilized as an indicator of the presence, or absence, of precipitation in between SMAP SSM estimates. Only soil moisture dry-downs with less than 1 mm of precipitation in between measurements are used to estimate the loss function.

Additionally, as a means of comparison and context, we use a few soil moisture time series from the U.S. Climate Reference Network (USCRN) climate monitoring stations (Bell et al. 2013) and apply the same loss function classification technique.

b. Climate and soil data

Ancillary climate data, independent from SMAP, are used to analyze and interpret loss function properties. They are not used in the estimation of the loss functions, but rather as diagnostics. Climate data are used to define an aridity index (AI; dimensionless) at 36-km resolution over the contiguous United States (CONUS) for the duration of this study. The AI follows the United Nations Environment Programme (UNEP) definition (Middleton and Thomas 1992) as the ratio between average precipitation P (mm day^{-1}) and potential evapotranspiration (PET; mm day^{-1}), that is, AI = P/PET. Based on this ratio, AI is then labeled as hyperarid (AI $<$ 0.05), arid (0.05 $<$ AI $<$ 0.2), semiarid (0.2 $<$ AI $<$ 0.5), dry/subhumid (0.5 $<$ AI $<$ 0.65), and humid (AI $>$ 0.65). Potential evapotranspiration and precipitation variables are extracted from the Global Land Data Assimilation (GLDAS) Noah 2.1 data products (Beaudoing and Rodell 2016) and reprocessed to the SMAP 36-km grid.

SMAP also provides soil texture data (Das 2013) in the form of sand and clay fractions along with bulk density information at the same spatial resolution as the corresponding soil moisture product. This same soil texture partitioning is used within the SMAP soil moisture retrieval algorithm. In this study, we extract the 36-km sand fraction information in order to examine the dependency of loss functions and their shapes.

3. Estimation of the soil moisture loss function

a. Loss function and dry-downs

Within the land water budget, increases and decreases in soil moisture encode information about different hydrological fluxes and processes. Therefore, separately analyzing positive changes, or increments, and negative increments can assist in isolating the effects of individual fluxes. With the exception of very large scale irrigation, we attribute increases in observed soil moisture to input precipitation, and decreases to evaporative losses and drainage. McColl et al. (2017a) exploited the fact that, at ecosystem scales, positive soil moisture increments are usually due to precipitation and thus quantified the percentage of precipitation that is captured in the dynamics of SMAP-observed surface soil moisture. Under stationary conditions, the positive and negative increments have nearly identical mean values, but their higher-order moments may differ considerably, since they encode information on different fluxes.

This study focuses on the losses or negative changes in surface soil moisture based on the method presented in Akbar et al. (2018). In the absence of or after precipitation, the loss of moisture from the surface layer due to evapotranspiration and drainage, that is, $L(\theta)$, is encoded in the dry-down behavior of soils. These dry-downs are manifested as consecutive and incremental decreases in the amount of soil moisture, that is, the negative temporal differences of soil moisture dry-downs. Figure 1c presents an example of this process where a time series of in situ USCRN surface soil moisture is shown along with consecutive negative ($\Delta\theta < 0$) increments [$\Delta\theta = \theta(t+1) - \theta(t)$] identified from dry-downs (shown in red). The rate of loss can be estimated from negative-only increments, defined as

$$\Delta\theta^{-} = \begin{cases} \Delta\theta(t), & \text{if} \quad \Delta\theta < 0 \\ 0, & \text{otherwise} \end{cases},$$

and the loss function $L(\theta)$ is approximated by conditioning the rate of negative increments with respect to soil moisture itself (Akbar et al. 2018):

$$L(\theta) \approx E\left[-\Delta z \frac{\Delta\theta^{-}}{\Delta t^{\text{obs}}}\middle|\theta\right]\left(\frac{\text{mm}}{\text{day}}\right), \tag{2}$$

where $E[\,]$ is the expectation operator, Δt^{obs} is the observation time interval between consecutive negative increments, and Δz is the depth scale of SMAP volumetric soil moisture estimates and is set to 50 mm. The depth scale Δz scales the loss function in (2), such that $L(\theta)$ is in flux units (L T^{-1}). Its specification does not affect the shape of the loss function nor its inflection points. As long as the surface moisture characterizes the hydrological state of the landscape, that is, temporal dynamics at the same location, then the dynamics of evaporative regimes—the goal of this paper—are consistently estimated. In landscape hydrology drainage losses may ultimately manifest in transpiration or stream outflow. However, in this study we only consider instantaneous soil water losses, calculated using (2). The method does not perform integrated water balance over time. Therefore, in this context, we consider drainage as the rapid loss of moisture from the control volume when the soil is wet. The approach toward obtaining loss functions is as follows:

1) Using the available soil moisture time series, for each 9-km SMAP pixel, dry-downs with less than 1 mm day^{-1} of precipitation in between observation, that is, $P(t) \leq 1\,\text{mm day}^{-1}$ between Δt^{obs}, are extracted (e.g., Fig. 1d). Similar to McColl et al. (2017b), dry-downs and increments less than 1% of the range of observed soil moisture are excluded.
2) Incremental and instantaneous soil water losses, $-\Delta z(\Delta\theta^{-}/\Delta t^{\text{obs}})$, are then calculated.
3) For each 36-km SMAP grid, individual losses from all valid 9-km pixels within the larger grid are aggregated and collected together. In other words, we simply collect all data within a 36-km grid cell. The aggregation to coarse resolution will increase subpixel heterogeneity in terms of weather, vegetation, and texture conditions and has the potential to diminish and mask finer soil or climate-induced inflection points and features (Saleem and Salvucci 2002). However, for a range of conditions this approach is suitable for climate scale analysis, given that GCM grid resolutions are on the order of 50–100 km.
4) Finally, the overall shape of the loss function is reconstructing by conditioning negative increments with respect to soil moisture itself, $E[-\Delta z(\Delta\theta^{-}/\Delta t^{\text{obs}})|\theta]$. The result is a function of soil moisture (the loss function) that is based on the entire observed soil moisture record. Results and analysis (section 4) will be presented at a 36-km spatial resolution

Individual soil moisture dry-downs, in general, are affected by a wide range of soil texture, vegetation, and general climate conditions. Such driving factors are, in fact, inherent and captured within the dry-down process. This feature is apparent in Fig. 1d, where the collection of dry-downs are dissimilar, indicating that atmospheric boundary conditions are implicitly encoded. The methodology outlined here aims at capturing an overall, or effective, first-order data-driven estimate of the soil moisture loss function over a period of time. While naturally other ancillary data or modeling efforts can be introduced, we intentionally avoid doing so. In this manner, we avoid biasing or contaminating the final loss function characteristics by ''signatures'' of the ancillary data and maximally extract information solely from soil moisture observations.

Therefore, the capacity to reconstruct loss functions rests on the ability to detect a sufficient number of soil moisture dry-downs. The spatial resolution of the soil moisture product has a lesser effect. We also note that because of SMAP's orbit and measurement scheme, certain shortcomings do exist. As discussed earlier, the vertical support of SMAP soil moisture observations is limited to the topsoil (50 mm); therefore, soil moisture dry-downs may be predominantly evaporation driven. In addition, drainage, or leakage, into the soil is a rapidly occurring process. Given the nominal 3-day sampling frequency of SMAP, the likelihood of detecting a complete leakage process, that is, the power-law shape in Figs. 1a and 1b, is diminished. Furthermore, the distribution of observed soil moisture is climatologically determined and rarely spans the range from wilting point to porosity; loss functions can only be estimated over the range of observed SSM states. Nevertheless, this technique provides a first-order approach to estimate $L(\theta)$ at scales relevant to climate studies. The following sections will discuss the shape and form of these loss functions.

b. *Loss function shape classification*

Loss functions take on different shapes over different hydroclimates. At global and continental scales, examining the shape and form of individual loss functions is impractical. Therefore, to address this, a shape-based categorical classification scheme is implemented to identify and label different pixels based on their underlying soil moisture loss function. We first define a set of piecewise-linear models $M(\boldsymbol{\Theta})$ labeled according to Fig. 2, such that each model is a representation of different segments of the idealized loss function seen in Fig. 1b. Our focus is on identifying the dominant hydrologic regimes based on the governing soil moisture at any time relative to the identified local soil moisture transition values. The term $\boldsymbol{\Theta}$ is the vector of model breakpoints, in volumetric soil moisture units ($m^3\,m^{-3}$), which determine the soil moisture transition points θ_w, θ_*, and θ_{fc} of Figs. 1a and 1b. These breakpoints will vary based on the dynamic range of observed soil moisture, rates of dry-down, soil texture, and overall climate conditions, and in general define the water- and energy-limited regimes. No prior constraints or conditions are placed on these parameters, and they are derived entirely from the classification approach. The prime ($'$) notation in Fig. 2 is used to distinguish between the transition points derived in Fig. 1 and our classification-based estimates. Detailed mathematical descriptions of each of the six canonical models are given in Table 1.

The classification procedure is as follows: a least squares optimization between each of the canonical models and observed loss data $L(\theta)$ is performed to obtain estimates of the mean squared error (MSE) and associated model parameters $\hat{\boldsymbol{\Theta}}$. This least squares model fitting is embedded within a 10-fold cross validation (CV) routine (Friedman et al. 2001; James et al. 2014) with 30 Monte Carlo replicates, that is, 10×30 model–data fittings for each canonical shape. Cross

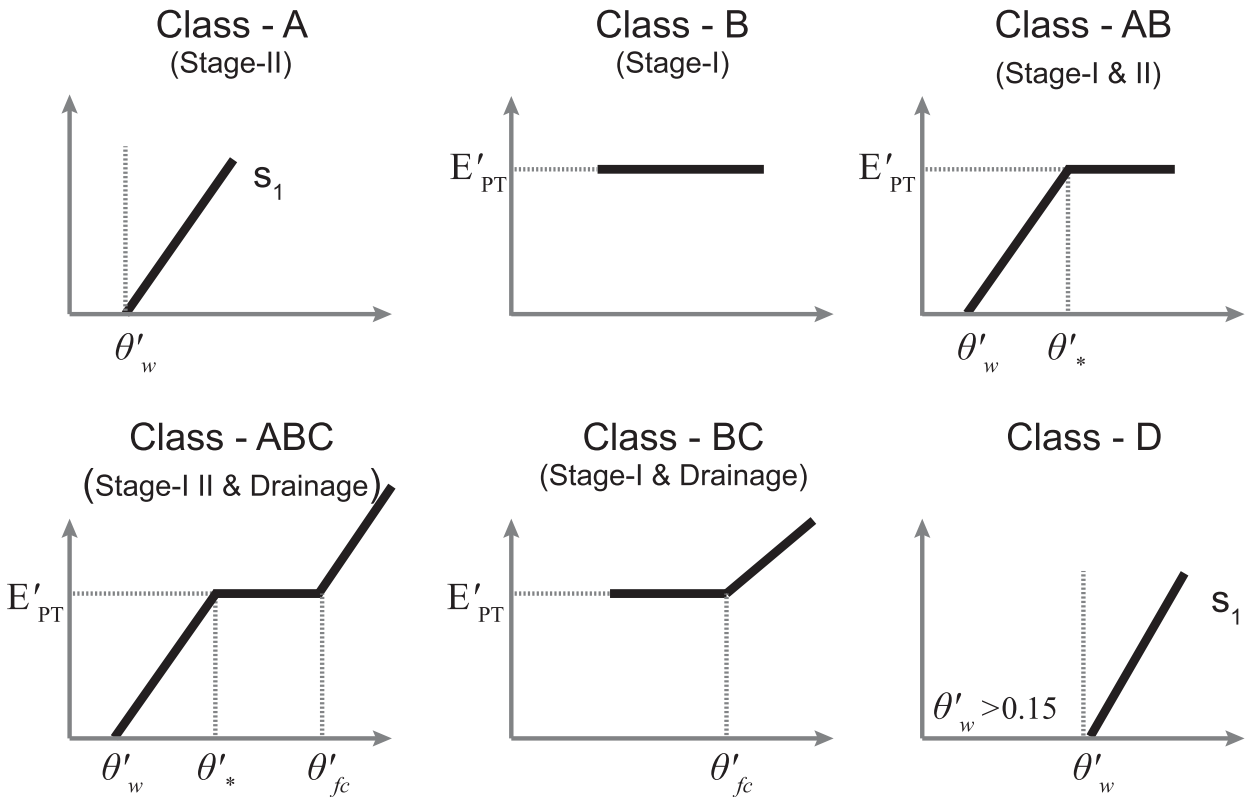

FIG. 2. Six canonical forms used to identify and classify SMAP-derived loss functions $L(\theta)$. Note the primed distinction of the x- and y-axis transition points, for example, θ'_w, θ'_*, and the actual parameters in Fig. 1. These parameters, in general, will vary based on soil texture, land cover, and climate conditions, and each are only determinable given a suitable dynamic range of soil moisture. Classes A and D are differentiated based on the value of the x-axis intercept point, such that Class A is assigned if $\theta'_w < 0.15$; otherwise, Class D is assigned.

validation partitions and splits the data into different segments, such that 90% of the data are used within the least squares routine and 10% are used to assess the performance of each model. For a single canonical model, the mean MSE score and one standard error (1-SE) across all 300 iterations, that is, permutation of data splits, is reported. The 1-SE is calculated as $\mathrm{SE} = \sqrt{\mathrm{var}(\mathrm{MSE}_1, \ldots, \mathrm{MSE}_N)}/N$, where $N = 10$ (James et al. 2014) and is later used to discriminate between models with similar mean MSE scores.

For each of the canonical models in Fig. 2, the CV process is performed multiple times across different partitions of the data to arrive at a set of mean MSE and 1-SE scores for each model, that is, $\mathrm{MSE}^{\mathrm{A}}_{\mathrm{mean}}$, $\mathrm{MSE}^{\mathrm{B}}_{\mathrm{mean}}$, etc. The procedure to determine the most representative, or optimum, model is as follows:

TABLE 1. Mathematical description of the loss function canonical forms.

Canonical form	Physical process	Equation
Class A	Stage II evapotranspiration	$L(\theta) = s_1(\theta - \theta'_w)$
Class B	Stage I potential evapotranspiration E_{PT}	$L(\theta) = E_{\mathrm{PT}}$ for all θ
Class AB	Stages I and II	$L(\theta) = \begin{cases} s_1(\theta - \theta'_w) & \theta \le \theta'_* \\ E_{\mathrm{PT}} & \theta \ge \theta'_* \end{cases}$
Class ABC	Stage I, II, and percolation	$L(\theta) = \begin{cases} s_1(\theta - \theta'_w) & \theta \le \theta'_* \\ E_{\mathrm{PT}} & \theta'_* \le \theta \le \theta'_{\mathrm{fc}} \\ E_{\mathrm{PT}} + s_2(\theta - \theta'_{\mathrm{fc}}) & \theta \ge \theta'_{\mathrm{fc}} \end{cases}$
Class BC	Stage II and percolation	$L(\theta) = \begin{cases} E_{\mathrm{PT}} & \theta \le \theta'_{\mathrm{fc}} \\ E_{\mathrm{PT}} + s_2(\theta - \theta'_{\mathrm{fc}}) & \theta \ge \theta'_{\mathrm{fc}} \end{cases}$
Class D	Same as Class A $\theta'_w > 0.15$	$L(\theta) = s_1(\theta - \theta'_w)$ $\theta'_w > 0.15$

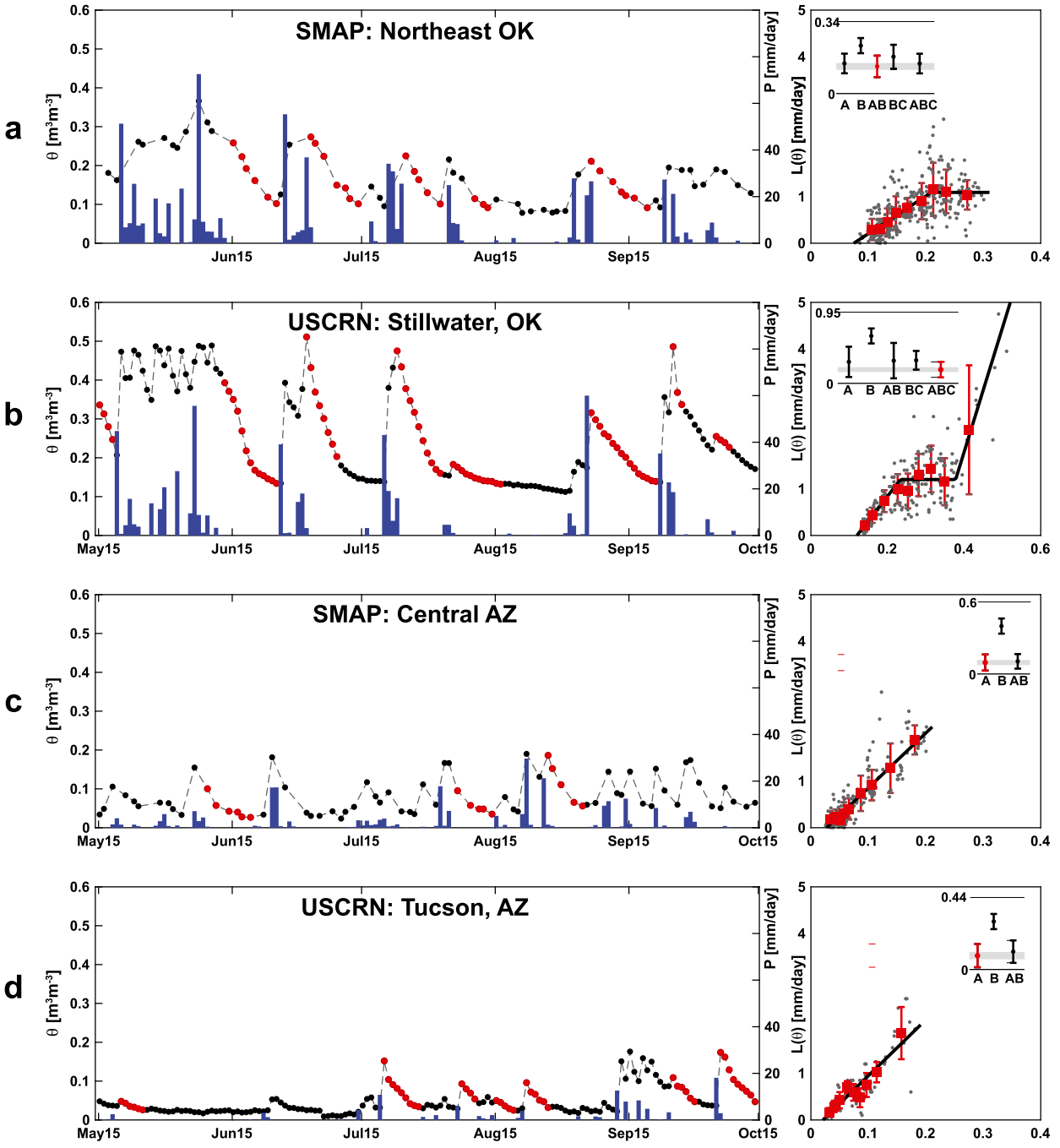

FIG. 3. (left) Pixel-level soil moisture time series and (right) loss function classifications for (a),(c) SMAP and (b),(d) USCRN in situ sensor sites sampled at 0600 LT daily. The red markers in the soil moisture time series plots represent the identified dry-downs used for loss function classification. In all right panels, small gray markers refer to the individual losses, $E[-\Delta z(\Delta\theta^{-}/\Delta t^{\text{obs}})|\theta]$, which are then grouped into nine equal-count bins (red markers with standard deviation error bars). The thick black line is the best-fitting canonical model as determined via the cross-validation process. The inset figure shows the relative magnitude of MSEs, wherein the selected model is colored red. The gray shaded area is the 1-SE interval (approximately 1/3 of standard deviation error bars).

1) The model with the smallest mean MSE score across all six canonical forms is initially selected as the optimum model.

2) If, within 1-SE of the optimum model, an alternative model exists that (i) is simpler (i.e., has fewer parameters) and (ii) has a smaller 1-SE, then this alternative

model replaces the optimum model. Over the CONUS, for approximately 11% of pixels the alternative model was chosen. No spatially coherent patterns were observed for regions where this occurred.

Once the optimum model is chosen, the corresponding pixel is assigned a label (Class A, Class B, etc.) denoting the selected loss function shape and its underlying physical process (Stage II, Stage I, etc.). The associated model parameters are also extracted and later used to determine the various hydrological regimes. For example, Stage II evapotranspiration is reported with a zero-loss intercept point θ'_w and slope s_1 [mm day^{-1} (m^3 m^{-3})$^{-1}$]. The model selection process outlined above is similar to the recent efforts by Schwingshackl et al. (2017) to examine and quantify evaporative fraction versus soil moisture relationships.

The parameter optimization and model assessment process is demonstrated in Fig. 3 for two different regions within the United States: Arizona and northern Oklahoma. The figures include examples from both SMAP (Figs. 3a,c) and in situ sensor data from the USCRN (Figs. 3b,d). In all plots, the 2015-only soil moisture time series is shown along with identified dry-downs (red markers) and precipitation (blue bars). The right panels show the corresponding individual losses $L(\theta)$ (red marker for all three years), as determined by (2). The thick black lines are based on the CV process outcome and denote the optimum and selected loss functions. The smaller inset plots in all figures show the relative magnitude of the mean MSE score for the different canonical model fits. The error bars denote the standard deviations of the MSE scores across all CV folds, that is, 300 iterations, and the thinner gray bars are the 1-SE of the optimum model (marked in red). Figures 3a and 3c are SMAP-derived loss function for regions in Oklahoma and Arizona, respectively. Figures 3b and 3d are similar, but for the closest USCRN in situ soil moisture sensors sites sampled at 0600 LT daily.

Figure 3 also highlights the importance of temporal sampling when classifying loss functions. Depending on latitude, SMAP has a 1–3-day temporal sampling rate. When more frequent sampling can be achieved, the loss function classification technique is able to reconstruct a more complete canonical form. This is evident in Fig. 3b, where the drainage, or percolation process, is reconstructed given the daily USCRN measurements, while in Fig. 3a this was not possible; the process was not observed. Note, however, that both plots yield the energy- and water-limited segments of the loss functions. On the other hand, examples in Figs. 3c and 3d (both in Arizona) only reconstruct the Stage II water-limited regimes. In general, across the CONUS, this can be attributed to either the absence of other processes, such as Figs. 3c and 3d, or very rapid short-lived events. In these examples, the transition points and plateau (Stage I) are not comparable because of the obvious scale disparity; SMAP is 36 km while the in situ station is a point measurement. Furthermore, given the dynamic nature of soil moisture, and its dry-downs, it is not equally distributed within each hydrologic regime. In latter sections, we focus only on dominant hydrologic segments where the time series of soil moisture persists more frequently.

This model selection process is applied over the CONUS and, where applicable, the x-axis transition points (θ'_w,θ'_*, and θ'_{fc}), the linear segment's slopes, and constant plateau values are also extracted. The soil moisture transition points are then used to identify the different hydrological regimes by partitioning the SMAP-observed soil moisture time series.

c. *Sensitivity of classification to measurement noise*

The loss function classification process is susceptible to noise and uncertainty, especially noise inherent in SMAP soil moisture products that in turn affect soil moisture dry-downs and the ability to estimate moisture losses. To examine the robustness of our classification algorithm to this noise, a Monte Carlo simulation study is performed using real precipitation data and two prescribed loss function shapes as follows:

1) Noisy time series of soil moisture are simulated using a simple hydrological model based on (1):

$$\theta_n(t+1) = \theta(t) + \frac{P(t)}{\Delta z} - L(\theta) + \delta\,, \tag{3}$$

   where δ is zero-mean Gaussian noise with standard deviation σ, that is, $\delta \sim N(0, \sigma^2)$, and is applied at every instance of time or measurement to mimic the SMAP instrument. The model is forced with observed daily precipitation $P(t)$ (mm day^{-1}). The subscript n denotes noisy soil moisture. In this experiment, the soil moisture time series from (3) is subsampled to every 3 days to be consistent with the SMAP temporal sampling. We manually define two loss functions in the form of (i) a single Stage II evapotranspiration (Class A) and (ii) a combined transitional Stage II to Stage I form (Class AB).
2) Equation (2) is used to estimate individual losses from the noisy time series $L(\theta_n)$.
3) The loss function classification and CV process are performed on data points obtained in (2). Naturally, because of the influence of noise, the classification can erroneously differ from the "true" prescribed loss function.

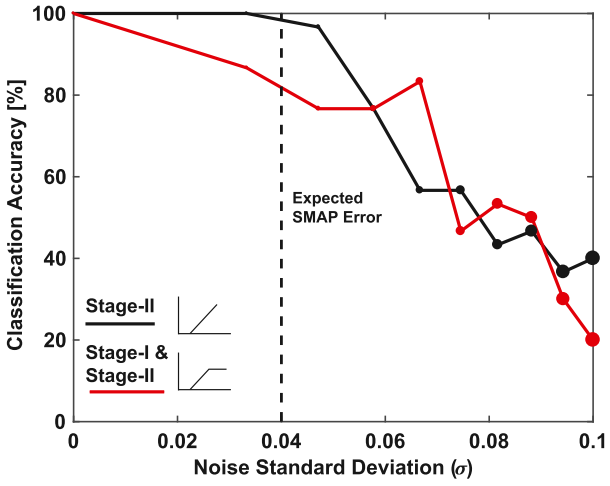

FIG. 4. Loss function classification accuracy with respect to noise. Two loss functions are manually prescribed, and noisy soil moisture time series are simulated. The vertical dashed line is the expected uncertainty associated with SMAP soil moisture data. The increasing marker sizes indicate a large MSE score and thus poor model–data fits with respect to increasing noise. A misclassification error less than 20% is expected.

4) Steps 1–3 are repeated 50 times in a Monte Carlo fashion for six different noise levels. Last, for each noise level and across all 50 iterations, the classification accuracy is calculated and reported as

$$\text{Accuracy} = \frac{\text{Number of times } L(\theta) \text{ is correctly identified}}{\text{Number of iterations}} \times 100. \tag{4}$$

Steps 1–4 are applied to the two most common loss functions observed over the CONUS: Stage II evapotranspiration and combined Stage I and Stage II. Figure 4 shows the classification accuracy as a function of the applied noise standard deviation. As expected, the accuracy drops with increasing noise, and the mean MSE scores of the optimum models increase. This trend is shown as the increasing marker size (MSE) with increasing noise. In Fig. 4, the vertical dashed line is the expected median noise standard deviation for an SMAP observation (Entekhabi et al. 2010) and is set to $0.04\,m^3\,m^{-3}$. This expected error is an SMAP mission requirement, and SMAP products have been shown to meet this criterion with respect to in situ ground truth calibration and validation sites (Chan et al. 2016). Therefore, given SMAP soil moisture data and the loss function classification scheme outlined in the previous section, we would expect somewhat less than a 20% error in our reported classification efforts for Class AB loss functions, and substantially less for Class A.

4. Results

a. Loss function classification

The 36-km loss function classification map over the CONUS for the 3-yr duration is shown in Fig. 5. Based on the outcome of the classification–CV effort, each pixel is labeled according to its general water- and/or energy-limited state. Pixels labeled as "no weather" refer to areas with no identifiable soil moisture dry-downs, and pixels labeled as "mask" are excluded because of high water fraction (>1%) and high vegetation water content ($>7\,kg\,m^{-2}$). See Figs. S1 and S2 in the online supplemental material for plot mean estimates of θ_w, θ_*, and θ_{fc}, as well as their associated uncertainty (standard deviations). A gradual transition from west to east is observed. Large regions of the western and southwestern United States are in a primarily Stage II water-limited domain with strong L sensitivity to θ ($dL/d\theta$) up to $20\,mm\,day^{-1}\,(m^3\,m^{-3})^{-1}$. This observation is consistent with field-scale studies in arid and semiarid environments, where high evaporative demand and bare soil evapotranspiration result in rapid loss of soil moisture (Cavanaugh et al. 2011; Vivoni et al. 2008; Kurc and Small 2007) as well as shorter dry-down time intervals (McColl et al. 2017b). Across the CONUS, for different climate zones, the median slope of Stage II evapotranspiration ranges from $20\,mm\,day^{-1}\,(m^3\,m^{-3})^{-1}$ in hyperarid areas to less than $5\,mm\,day^{-1}\,(m^3\,m^{-3})^{-1}$ for humid regions in the east (see Fig. S3h). This trend is consistent with the gradual decrease in the rate of change of evaporative fraction with respect to soil moisture, that is, $\partial EF/\partial\theta$, based on reanalysis studies from Schwingshackl et al. (2017).

Parts of the Great Plains show a transitional behavior, that is, Stage I and II (yellow). The wetter Midwest areas show a combination of Stages I and II, with some regions exhibiting an additional drainage-like behavior (green and light blue, respectively). The eastern and southeastern United States show a mixture of different classifications for neighboring pixels.

Note, however, that behind each pixel is a dynamic soil moisture time series partitioned into the water- and energy-limited segments over time. As will be apparent in the next section, to resolve this pixilation and mixture issue, we use the percentage of time of SMAP observation within each segment (Stage I, Stage II, etc.) to identify the dominant hydrologic regimes. This effort yields more spatially uniform and consistent hydrologic regimes.

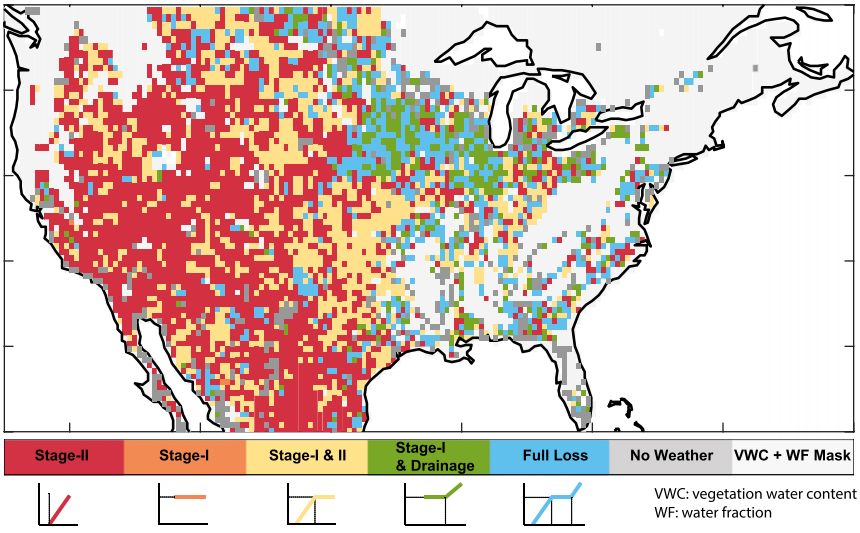

FIG. 5. Soil moisture loss function classification over the CONUS. Masked regions due to high vegetation (>7 kg m^{-2}) and more than 1% water fraction are labeled as "VWC + WF Mask" and shown in light gray. Regions marked as "No Weather" (gray) are areas where no dry-downs were detected.

As a measure of confidence in both the hydroclimatological loss function classification and parameter estimate process, Fig. 6 shows the MSE score and correlation R^2 of the optimum loss function model with respect to the observed losses. The largest errors are concentrated in the U.S. Midwest. In these regions, soil moisture dry-downs are irregular due to rapid crop growth, irregular summertime precipitation patterns, and unknown irrigation inputs. The R^2 between observed losses and the optimum model is highest in the drier regions of the southern and southwestern United States, where derived loss functions are typically linear and monotonically increase with respect to increasing soil moisture, that is, Stage II. An apparent striping pattern is visible in Fig. 6a, especially in southern Texas. In these areas when calculating the MSE score, the variable samples, that is, $1/N$ division, becomes more apparent. This is due to overlapping SMAP half-orbits. These features are not present nor affect the same locations in the classification map of Fig. 5.

The same analysis outlined thus far was also performed at the 9-km spatial resolution (figure not shown). The spatial patterns of the 9-km classification map and the 36-km map (from Fig. 5) are consistent. However, the higher-resolution classification map shows increased pixilation due to increased finer-resolution vegetation and water fraction filtering and is not presented in this study.

b. *Characteristic dry-down time scales*

A characteristic dry-down time scale, in days, from the Stage II slope {s_1 [mm day^{-1} $(m^3 m^{-3})^{-1}$] in Table 1}, can be calculated as $\Delta z/s_1$, shown in Fig. 7, where $\Delta z = 50$ mm. Here, we observe that the median soil moisture characteristic decay time scale gradually increases with increasing wetness, from approximately 3 days in hyperarid regions of the Southwest to 7 days in the more humid areas of the eastern United States. These trends are consistent with McColl et al. (2017b), where characteristic dry-down time scales, at a global scale, were determined via exponential model fitting and span from approximately 11 days for wetter regions to 3 days for highly arid environments. Note that these time-scale estimates are shorter than soil moisture memory and persistence time scales in Katul et al. (2007) and Ghannam et al. (2016), for example, 7 days, or months in the case of Delworth and Manabe (1988). The discrepancies stem from the fact that here, similar to McColl et al. (2017b), we consider only the surface soil moisture content with a nominal vertical support depth, or storage depth, of 50 mm. Prior studies typically evaluate time scales using the entire root-zone soil moisture profiles, up to 1000 mm. Therefore, faster surface soil moisture dissipation from a shallower support depth is expected. Additionally, Shellito et al. (2016) have reported faster SMAP dry-downs with respect to in situ soil sensors. The reported discrepancies stem from differences in the SMAP sensing depth—nominally 50 mm—relative to in situ sensors and also differences in the spatial support of the different measurements. Nevertheless, observed trends are generally consistent such that hyperarid and arid regions have the highest rate of surface moisture loss. While the rate and spatial pattern for the Stage II slope is similar to the evaporative fraction versus soil moisture rates of Schwingshackl et al. (2017), the actual amount of moisture

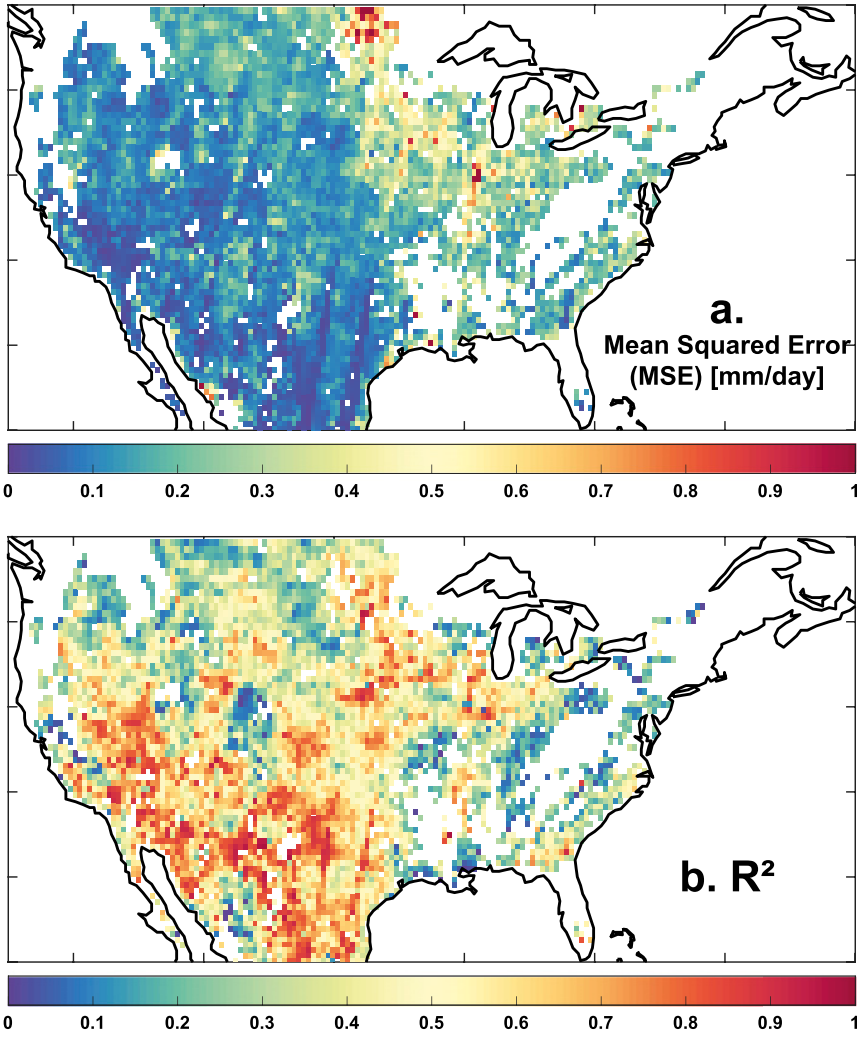

FIG. 6. (a) MSE and (b) R^2 between optimum fitting canonical model and observed losses $L(\theta)$. The largest errors are collocated with croplands in the Midwest.

loss might differ. This is because only surface soil moisture with a nominal depth scale, $\Delta z = 50$ mm, is considered, as discussed in section 2.

c. *Dominant hydrological regimes*

The classification map in Fig. 5 is obtained by the identification of soil moisture dry-downs, calculating their negative increments and then finding the optimum canonical loss function for each pixel. The loss function is the conditional expectation of negative increments conditioned on soil moisture itself. At any location, the estimation of the conditional expectation and its parameters is based on the combined record of 3 years. Because of variable precipitation patterns and atmospheric conditions, the dynamic range and temporal evolution of SMAP-retrieved soil moisture will vary across different hydrological regimes. Studies exploring soil moisture persistence—indicative of wet/dry states—often examine the frequency of occurrence, or spectral components, of soil moisture above or below predefined thresholds. Ghannam et al. (2016), for example, determined wet/dry thresholds by relating measured ET to soil moisture and then investigated the spectral of soil moisture threshold crossing for various high-frequency in situ sites with different climate and biome characteristics. In this section, using the loss function classification outcomes from Fig. 6 and associated local parameters, we quantify the fraction of SMAP-observed soil moisture falling within each of the three hydrological regimes listed in Fig. 1.

For each 36-km pixel, using estimates of θ'_w, θ'_*, and θ'_{fc}, we calculate the percentage of SMAP observations

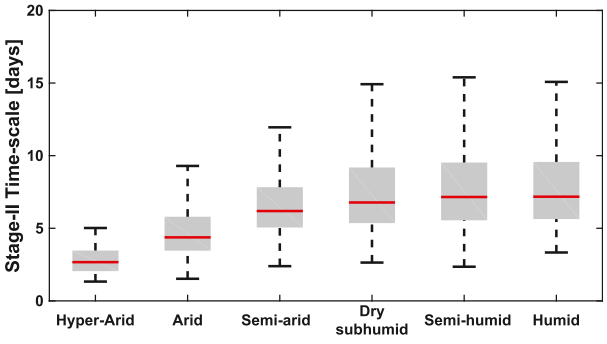

FIG. 7. Estimated soil moisture characteristic dry-down time scale, in days. The time scale is calculated as $\Delta z/s_1$ (day), where the slope of Stage II evapotranspiration s_1 is derived from the loss function classification process and given in Table 1. AI is based on the UNEP definition discussed in section 2b.

falling within 1) Stage II or water-limited evapotranspiration $\theta'_w \le \theta < \theta'_*$, 2) Stage I or energy-limited evapotranspiration $\theta'_* \le \theta < \theta'_{fc}$, and 3) drainage or above field capacity $\theta'_{fc} \le \theta$. For each of the cases, pixels with less than 20% of SMAP observations falling within the defined process are excluded. We further apply a two-dimensional moving-average filter across the CONUS. Partially filled pixels (less than 25%), within the moving-average filter are excluded. The filtering allows further analysis focusing on interannual variations. Hydroclimatic boundaries between large-scale regions are identified and year-to-year movements of these boundaries are tracked. This approach also resolves the pixilation effect seen in Fig. 6, where neighboring pixels at times show a mixture of hydrologic processes. By focusing on dominant and present processes, Fig. 8 yields more spatially coherent and less pixelated hydrologic regimes.

Figures 8a–c show the percentage of SMAP observations falling within the three distinct regions mentioned above. When categorically combined, Figs. 8a–c yield a map of the dominant observed hydrological regions over the CONUS. In general, a west–east gradient is observed with the drier Stage II process dominating the western half of the United States. Large parts of the Great Plains and central United States are observed to be in a transitional state, and the eastern United States is in a combined Stage I and/or drainage regime. The spatial distributions of Fig. 8d are consistent with land surface modeling analyses, even including some relatively finescale features (Koster et al. 2009; Schwingshackl et al. 2017); however, the hydroclimate regions in Fig. 8 are determined using solely observations of soil moisture time series.

Using the procedure above, we can identify regional shifts in the hydrological regimes for each year, that is, 2015, 2016, and 2017. Figures 9a–c are similar to Fig. 8b, but separated by year. In Figs. 9b and 9c the contour lines represent the spatial domain of the 2015 hydrological states. Year to year, we observe small but important weather-induced changes in regional domains that are most noticeable in the western United States. The spatial extent of hydrological process in the Midwest and eastern United States do not vary significantly. Two western regions showing a mix of Stage I and II processes in 2015 show drier conditions in 2016, and a corresponding shift to a dominant Stage II hydrological regime (see the two sizeable green "islands" in the middle of the red region across the western United States in Fig. 9a that disappear in the lower panels). This occurrence is partly due to application of the moving-average filter and the consequent exclusion grid cells. From 2016 to 2017 we observe that large parts of the southern plains (Texas and Oklahoma) revert to a water-limited state (see the large red region advancing north across the hydroclimate boundary in western Texas and Oklahoma). This is attributed to persistent moderate to severe droughts occurring throughout the summer of 2017 (Svoboda et al. 2002; NOAA 2017; Otkin et al. 2018). Occurrence of persistent extreme droughts, especially in eastern Montana throughout the summer of 2017, is also partly captured in Fig. 9c, shown as a large Stage II region. The temperate rain forest of western and northwest Mexico saw above-average precipitation over the period of study (Blunden and Arndt 2017). Consequently, a large Stage II and I transition domain increases from 2015 to 2017. Changes in water availability—and the corresponding change in hydrologic regime—are consequential in these regions. Semiarid transitional areas drive a majority of the interannual variability in the carbon cycle, with fluctuations primarily driven by water availability or the secondary effects of water availability on temperature (Zhang et al. 2016; Jung et al. 2017).

Figure 9c indicates the presence of persistent moderate droughts in parts of Montana, Wyoming, and the Dakotas during the summer season of 2017. As a case study, we examine the temporal change in the dominant hydrologic regimes over this domain. The 2017 "flash" drought across the region expanded rapidly over the season. With the expansion of the region in the Stage II evaporation regime, further drying of soil moisture and increased sensible heat flux due to reduction of evaporation in Stage II can lead to increased surface temperature, greater evaporative demand, and further drying. This represents a positive feedback and a contributor to the intensification of drought. The dynamics of the "flash" drought is accompanied by rapid shifts between Stage I and Stage II evaporation regimes as detected based on the SMAP observations and the hydrologic regimes identification introduced in this study. In Fig. 10, for each pixel and based on the optimum

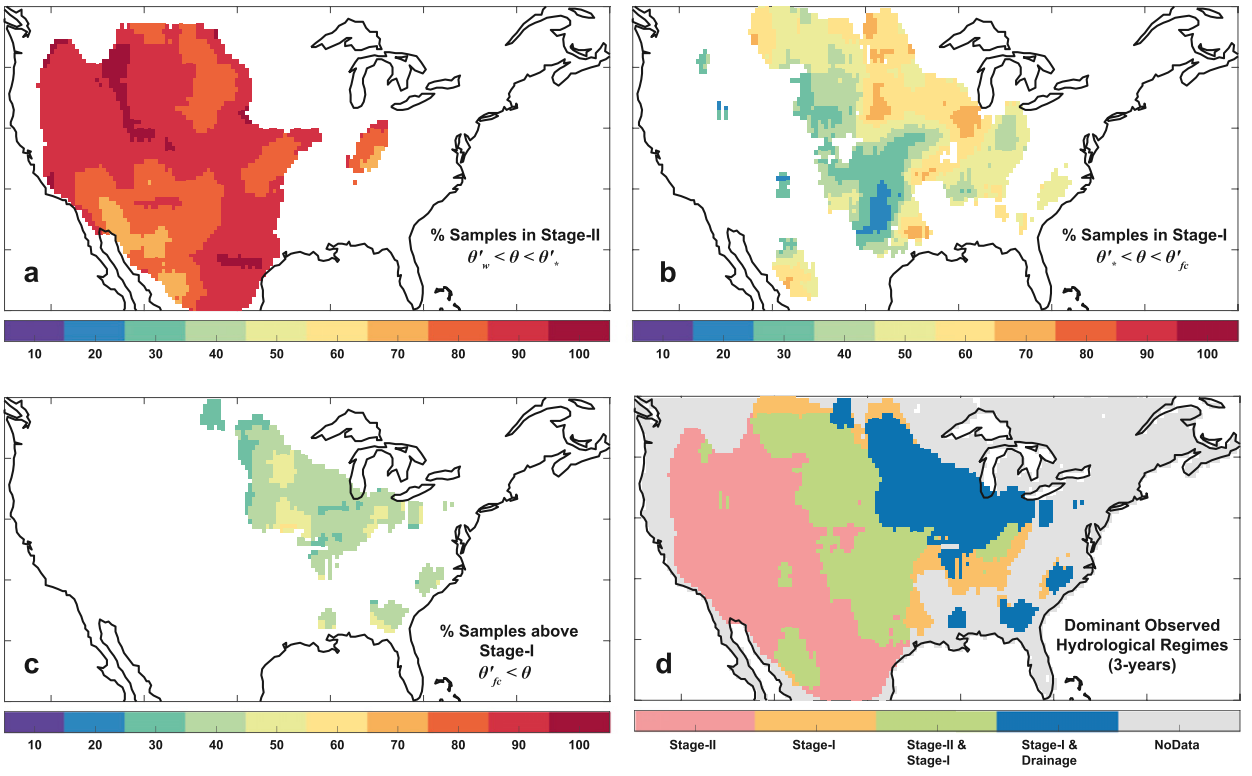

FIG. 8. Percent of SMAP observations (20% or more) falling within different hydrological regimes (a) Stage II evapotranspiration $\theta_w \le \theta < \theta_*$, (b) Stage I evapotranspiration $\theta_* \le \theta < \theta_{fc}$, and (c) above Stage I $\theta_{fc} \le \theta$. (d) The categorical map of these dominant and observed regimes over three consecutive spring and summer periods.

canonical loss function and associated local parameters (θ_w, θ_*, θ_{fc}), we estimate the fraction [0, 1] of SMAP observations falling within the Stage II or Stage I regimes, that is, $\theta_{SMAP} < \theta_*$ and $\theta_{SMAP} > \theta_*$, respectively. Figure 10 shows the time series change in these fractions for the entire study domain along with domain-averaged daily precipitation. Four select snapshots are also shown highlighting the spatial variation of the areas with SMAP observations in Stage II. Over the entire domain, observe the increased spatial extent of Stage II pixels during dry-downs postprecipitation, for example, between 14 May and 7 June. Transitions between Stage II and Stage I due to precipitation events are evident as well as long-term (approximately a few weeks) persistence in Stage II reflective of drought conditions.

5. Discussion

a. Comparison with other methods

Koster et al. (2017) recently proposed a soil moisture forecasting technique where the shape of the loss function is estimated based on a model of soil moisture with precipitation input and a loss function. The function is estimated, piecewise, across the soil moisture range by numerical optimization and matching the dynamics of SMAP-observed soil moisture estimates over the continental United States to model outputs. Other methods include exploiting the statistical properties of stationary processes (Salvucci 2001; Tuttle and Salvucci 2014) to estimate the conditional mean loss function $E[L(\theta)|\theta]$. The expected value of a statistically stationary soil moisture change conditioned on soil moisture itself is zero. Thus, the loss function in (2) can be estimated as the conditional expectation of precipitation. Similar to Koster et al. (2017), this approach also requires precipitation information, while the technique outlined in section 2 only uses precipitation as an indicator of the presence or absence of rain events.

We note that (2) is similar to a recent analysis performed by Shellito et al. (2018), wherein aggregate rates of soil moisture dry-downs over the CONUS were compared to model-based evapotranspiration rates. Faster SMAP drying rates—compared to model-based estimates—were reported for both high surface soil moisture and potential evapotranspiration conditions in addition to regions with sparse to medium amounts of vegetation. Additionally, they report minimal dependency of SMAP soil moisture drying rates with respect to soil texture

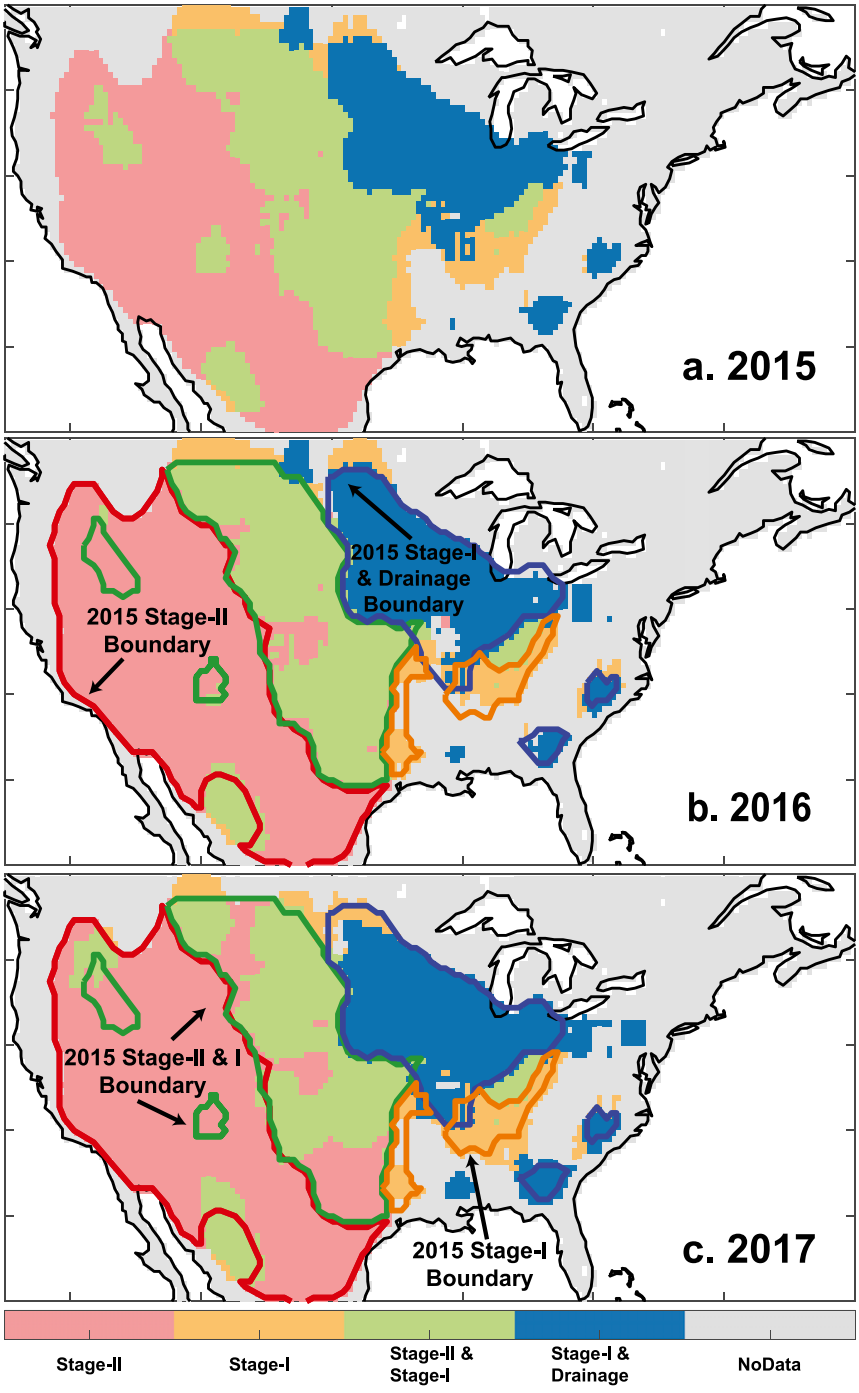

FIG. 9. Dominant observed hydrological regimes partitioned into (a) 2015, (b) 2016, and (c) 2017 May–September months. The thick contour lines in (b) and (c) are the domain boundaries of 2015 overlaid on the other years to highlight changes in the different hydrological regimes.

across the CONUS. Estimated values of θ_w and θ_* derived from the classification scheme exhibit variations with respect to soil texture (see Fig. S3). Median values of θ_w and θ_*, in general, decrease with increasing sand fraction across the CONUS. Median values of θ_{fc} also show slight dependency on sand fraction, but the exact dependency is inconclusive since the spatial domain where θ_{fc} is estimated is limited. Shellito et al. (2018) also

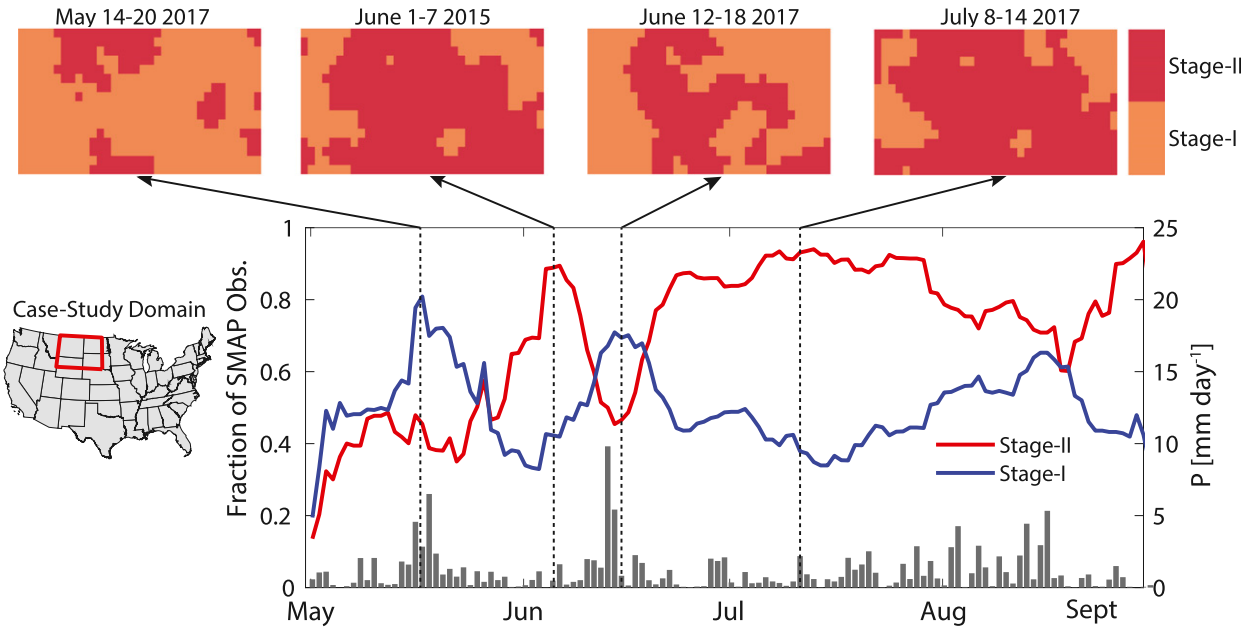

FIG. 10. Dynamics of hydrologic regimes based on 2017 time series of change in the fraction of SMAP observations falling in Stage II or Stage I, that is, $\theta_{SMAP} < \theta_*$ or $\theta_{SMAP} > \theta_*$ across a region including Montana, parts of the Dakotas, and Wyoming; the inset map shows the extent of the study. The time series of regional mean precipitation is also included. Four weekly-averaged snapshot plots show the spatial extent of the Stage II–dominated areas. Colors are consistent with Fig. 5.

identified a spatial-median soil moisture transition point θ_* of $0.15\,m^3\,m^{-3}$ over the CONUS compared to our median values of $0.18\,m^3\,m^{-3}$. Additionally, our mean estimates of θ_* (Fig. S1b) over the United States are consistent with reanalysis studies from Schwingshackl et al. (2017). They report an east–west gradient in θ_* with smaller values in the west (approximately $0.15\,m^3\,m^{-3}$) and up to $0.4\,m^3\,m^{-3}$ in the east. The important difference is that θ_* values from Schwingshackl et al. (2017) are model based, while here the loss function transition points are all estimated from surface soil moisture alone. Furthermore, for the top 50 mm of surface soil moisture storage volume, they report evaporative efficiency and drying rates exhibiting mostly a linear response to changes in soil moisture, thus indicative of a more water-limited behavior. In our analysis, water-limited states cover the majority of the western and central United States; however, wetter and transitional hydrological regimes are also observed in the Midwest and Great Plains.

b. Limitations

The spatial resolution of SMAP is constrained by the half-power, or −3 dB, beamwidth of the instrument's antenna (approximately 30–40 km) on the surface of Earth. Because of this observation geometry and the large spatial scales of SMAP soil moisture, subpixel heterogeneity—in terms of variations in vegetation cover and soil texture—are effectively averaged into a single soil moisture values. This feature may also mask, or diminish, finer subpixel processes that may be dominant at scales on the order of tens of meters, for example, subpixel presaturation runoff in U.S. Southeast. Nevertheless, the focus of this study is on large spatial scales relevant to climate studies where, similar to Koster et al. (2009), we consider only effective, or averaged, processes under mean climate state.

Accurate reconstruction of $L(\theta)$ requires sufficient detection of soil moisture dry-downs. Infrequent observations or rapid dry-downs (less than a day) will not yield adequate negative increments to observe transitions between the various moisture loss stages. SMAP observations vary from daily at high latitudes to 3-day intervals at midlatitudes. As a result, faster processes such as wet soil drainage cannot be adequately detected. This is evident in Fig. 3b, where the loss function estimated based on daily in situ measurements has a clear drainage regime whereas the loss function based on the SMAP temporal sampling (2–3 days) cannot characterize this hydrologic regime.

While care is taken to remove and exclude unreliable SMAP soil moisture retrievals—based on quality flags and surface conditions—they do contain some random noise. The noise can lead to a bias in the estimation of the loss functions. McColl et al. (2017a) provided an error analysis of soil moisture increments and showed that larger (more positive or more negative) soil moisture increments are least susceptible to noise. A means

to further improve the dry-down detection process (i.e., negative increments) is to develop model-based or predetermined thresholds in the future. That is, negative increments must meet the condition $\theta(t+1) - \theta(t) \leq \epsilon$, where, for example, ϵ is derived based on modeling analysis.

Mean estimates and standard deviations of the loss function transition points (θ_w, θ_*, θ_{fc}) are shown in Figs. S1 and S2, respectively. Currently, because of the short time span of this study, the use of these transition points as inputs in LSMs or GCMs is not recommended. They are important soil parameters in physics-based land surface models. Therefore, they must be estimated using observations that are not affected by landscape heterogeneity, that is, in situ data. Furthermore, if the objective is to produce a more robust set of transition point parameters, other data such as soil texture, land use, topography, precipitation, and radiation can be leveraged. Furthermore, model parameters are often specific to the structure of the model and not readily transferrable to another model. A more traditional model calibration approach may thus be a more robust approach to identifying parameters for modeling purposes. In this study, these landscape-scale parameters are intermediate variables that are ultimately used for the purposes of classifying the dominant hydrologic regimes and their dynamics. Models and ancillary data are not used in order to not pass their signatures to the identified map of the dominant hydrologic regimes. To evaluate models and perform model intercomparisons, the spatial and temporal patterns of the dominant hydrologic regimes—rather than parameters themselves—can be effectively used and compared with observation-only estimates.

6. Conclusions

Using soil moisture estimates from the NASA SMAP mission, the topsoil water loss function is characterized over the CONUS for three consecutive spring and summer periods (May–September, 2015–17) at about 36-km spatial resolution. The loss function is the conditional expectation of negative soil moisture time increments where the conditioning is on soil moisture itself. At any soil moisture level, the rate of loss (and hence negative increment) is indicative of the dominant hydrologic regime. A shape-based classification approach is implemented to 1) detect soil moisture dry-downs with negligible precipitation in between measurements, 2) calculate soil water losses from successive negative increments of dry-downs to determine the loss function $L(\theta)$ for the entire soil moisture record, and 3) categorize the observed loss function into known parametric canonical forms through an intensive cross-validation process. Each canonical shape is a representation of the different water- and energy-limited loss regimes, that is, Stage I, Stage II, transitional, etc. The results presented here are observation driven. Only SMAP soil moisture time series are used to identify the dominant hydrologic regimes without reliance on models or ancillary data. Precipitation is used as a quality control in identifying dry-downs, although the positive soil moisture increments disrupting series of sequential negative increments are themselves a good indicator of the end of an interstorm dry-down period.

The observed loss functions and associated parameters, in general, are consistent with field-scale studies of the dynamics of soil moisture loss and evapotranspiration as well as recent model-based analyses. Strong dependency on climate is seen across the United States, notably a clear east–west divide. The western and southwest United States show large rates of moisture loss from the surface layer in a Stage II or water-limited regime with a characteristic decay time scale of about 3 days. For each of the three years of the SMAP record, we produce maps of large-scale boundaries between regions with common dominant hydrologic regimes. Comparisons of the shifts in the boundaries are linked to interannual variability such as the transition of western Texas into and out of drought over the 2015–17 summer seasons. Other shifts in the boundaries related to interannual variability are also evident.

By extension, this method can be applied globally and over longer periods of time to examine seasonal progressions and hydrologic transitions as well as investigating regional and short-lived anomalies. Additionally, analysis over longer periods (multiple years) of time may yield more representative loss functions, enabling a better observation-based approach to examine the role of soil moisture in controlling land–atmospheric processes. Characterization of shifts in the boundaries separating regions with common dominant hydrologic regime can provide important insights into how interannual changes in climate can force a shift in surface hydrologic regimes. These shifts can either reinforce or dissipate the climate anomaly (depending on where the land–atmosphere feedback is positive or negative). The sign of the feedback can also factor into the rate of movement of the boundary. Such diagnostics—based on remote sensing observations alone—can be important tools in advancing the understanding of land–atmosphere interactions.

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/. USCRN in situ soil moisture data can be downloaded from https://www.ncdc.noaa.gov/crn/qcdatasets.html.

REFERENCES

Akbar, R., D. J. Short Gianotti, K. A. McColl, E. Haghighi, G. D. Salvucci, and D. Entekhabi, 2018: Hydrological storage length scales represented by remote sensing estimates of soil moisture and precipitation. *Water Resour. Res.*, **54**, 1476–1492, https://doi.org/10.1002/2017WR021508.

Albergel, C., T. Pellarin, N. Fritz, F. Froissard, A. Petitpa, B. Piguet, and E. Martin, 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. *Hydrol. Earth Syst. Sci.*, **12**, 1323–1337, https://doi.org/10.5194/hess-12-1323-2008.

Beaudoing, H., and M. Rodell, 2016: GLDAS Noah Land Surface Model L4 monthly 0.25 x 0.25 degree V2.1. Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed September 2017, https://doi.org/10.5067/SXAVCZFAQLNO.

Bell, J. E., and Coauthors, 2013: U.S. Climate Reference Network soil moisture and temperature observations. *J. Hydrometeor.*, **14**, 977–988, https://doi.org/10.1175/JHM-D-12-0146.1.

Blunden, J., and D. S. Arndt, 2017: State of the Climate in 2016. *Bull. Amer. Meteor. Soc.*, **98** (Suppl.), https://doi.org/10.1175/2017BAMSStateoftheClimate.1.

Cavanaugh, M. L., S. A. Kurc, and R. L. Scott, 2011: Evapotranspiration partitioning in semiarid shrubland ecosystems: A two-site evaluation of soil moisture control on transpiration. *Ecohydrology*, **4**, 671–681, https://doi.org/10.1002/eco.157.

Chan, S. K., and Coauthors, 2016: Assessment of the SMAP passive soil moisture product. *IEEE Trans. Geosci. Remote Sens.*, **54**, 4994–5007, https://doi.org/10.1109/TGRS.2016.2561938.

Das, N., 2013: SMAP ancillary data report: Soil attributes. SMAP Science Doc. 044, JPL D-53058, 16 pp., http://smap.jpl.nasa.gov/files/smap2/044_soil_attrib.pdf.

Delworth, T. L., and S. Manabe, 1988: The influence of potential evaporation on the variabilities of simulated soil wetness and climate. *J. Climate*, **1**, 523–547, https://doi.org/10.1175/1520-0442(1988)001<0523:TIOPEO>2.0.CO;2.

Dirmeyer, P. A., 2000: Using a global soil wetness dataset to improve seasonal climate simulation. *J. Climate*, **13**, 2900–2922, https://doi.org/10.1175/1520-0442(2000)013<2900:UAGSWD>2.0.CO;2.

——, F. J. Zeng, A. Ducharne, J. C. Morrill, and R. D. Koster, 2000: The sensitivity of surface fluxes to soil water content in three land surface schemes. *J. Hydrometeor.*, **1**, 121–134, https://doi.org/10.1175/1525-7541(2000)001<0121:TSOSFT>2.0.CO;2.

Dralle, D. N., and S. E. Thompson, 2016: A minimal probabilistic model for soil moisture in seasonally dry climates. *Water Resour. Res.*, **52**, 1507–1517, https://doi.org/10.1002/2015WR017813.

Entekhabi, D., and Coauthors, 2010: The Soil Moisture Active Passive (SMAP) mission. *Proc. IEEE*, **98**, 704–716, https://doi.org/10.1109/JPROC.2010.2043918.

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

——, T. E. Dawson, D. D. Ackerly, L. S. Santiago, and S. E. Thompson, 2017: Reconciling seasonal hydraulic risk and plant water use through probabilistic soil-plant dynamics. *Global Change Biol.*, **23**, 3758–3769, https://doi.org/10.1111/gcb.13640.

Ford, T. W., E. Harris, and S. M. Quiring, 2014: Estimating root zone soil moisture using near-surface observations from SMOS. *Hydrol. Earth Syst. Sci.*, **18**, 139–154, https://doi.org/10.5194/hess-18-139-2014.

Friedman, J., T. Hastie, and R. Tibshirani, 2001: *The Elements of Statistical Learning.* Springer, 533 pp.

Ghannam, K., T. Nakai, A. Paschalis, C. A. Oishi, A. Kotani, Y. Igarashi, T. Kumagai, and G. G. Katul, 2016: Persistence and memory timescales in root-zone soil moisture dynamics. *Water Resour. Res.*, **52**, 1427–1445, https://doi.org/10.1002/2015WR017983.

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

Hirschi, M., B. Mueller, W. Dorigo, and S. I. Seneviratne, 2014: Using remotely sensed soil moisture for land–atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability. *Remote Sens. Environ.*, **154**, 246–252, https://doi.org/10.1016/j.rse.2014.08.030.

James, G., D. Witten, T. Hastie, and R. Tibshirani, 2014: *An Introduction to Statistical Learning: With Applications in R.* Springer, 426 pp.

Jung, M., and Coauthors, 2017: Compensatory water effects link yearly global land CO_2 sink changes to temperature. *Nature*, **541**, 516–520, https://doi.org/10.1038/nature20780.

Katul, G. G., A. Porporato, E. Daly, A. C. Oishi, H.-S. Kim, P. C. Stoy, J.-Y. Juang, and M. B. Siqueira, 2007: On the spectrum of soil moisture from hourly to interannual scales. *Water Resour. Res.*, **43**, W05428, https://doi.org/10.1029/2006WR005356.

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

Koster, R. D., and M. J. Suarez, 2003: Impact of land surface initialization on seasonal precipitation and temperature prediction. *J. Hydrometeor.*, **4**, 408–423, https://doi.org/10.1175/1525-7541(2003)4<408:IOLSIO>2.0.CO;2.

——, S. D. Schubert, and M. J. Suarez, 2009: Analyzing the concurrence of meteorological droughts and warm periods, with implications for the determination of evaporative regime. *J. Climate*, **22**, 3331–3341, https://doi.org/10.1175/2008JCLI2718.1.

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

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

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

Manzoni, S., J. P. Schimel, and A. Porporato, 2012: Responses of soil microbial communities to water stress: Results from a meta-analysis. *Ecology*, **93**, 930–938, https://doi.org/10.1890/11-0026.1.

McColl, K. A., S. H. Alemohammad, R. Akbar, A. G. Konings, S. Yueh, and D. Entekhabi, 2017a: The global distribution and

dynamics of surface soil moisture. *Nat. Geosci.*, **10**, 100–104, https://doi.org/10.1038/ngeo2868.

——, W. Wang, B. Peng, R. Akbar, D. Gianotti, H. Lu, M. Pan, and D. Entekhabi, 2017b: Global characterization of surface soil moisture drydowns. *Geophys. Res. Lett.*, **44**, 3682–3690, https://doi.org/10.1002/2017GL072819.

Middleton, N., and D. S. G. Thomas, Eds., 1992: *World Atlas of Desertification*. Arnold, 69 pp.

NCAR, 2017: The Climate Data Guide: CPC Unified gauge-based analysis of global daily precipitation. NCAR, https://climatedataguide.ucar.edu/climate-data/cpc-unified-gauge-based-analysis-global-daily-precipitation.

NOAA, 2017: State of the Climate. NOAA/NCEI, accessed 20 October 2017, https://www.ncdc.noaa.gov/sotc/drought.

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, and R. Bindlish, 2016: SMAP enhanced L3 radiometer global daily 9 km EASE-Grid soil moisture, version 1. NSIDC, accessed September 2017, http://nsidc.org/data/SPL3SMP_E/versions/1.

Otkin, J. A., M. Svoboda, E. D. Hunt, T. W. Ford, M. C. Anderson, C. Hain, and J. B. Basara, 2018: Flash droughts: A review and assessment of the challenges imposed by rapid onset droughts in the United States. *Bull. Amer. Meteor. Soc.*, https://doi.org/10.1175/BAMS-D-17-0149.1, in press.

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

——, ——, and ——, 2016: The impact of vertical measurement depth on the information content of soil moisture for latent heat flux estimation. *J. Hydrometeor.*, **17**, 2419–2430, https://doi.org/10.1175/JHM-D-16-0044.1.

Reichle, R. H., R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe, 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). *J. Geophys. Res.*, **112**, D09108, https://doi.org/10.1029/2006JD008033.

——, W. T. Crow, R. D. Koster, H. O. Sharif, and S. P. P. Mahanama, 2008: Contribution of soil moisture retrievals to land data assimilation products. *Geophys. Res. Lett.*, **35**, L01404, https://doi.org/10.1029/2007GL031986.

Rodriguez-Iturbe, I., and A. Porporato, 2007: *Ecohydrology of Water-Controlled Ecosystems: Soil Moisture and Plant Dynamics*. Cambridge University Press, 464 pp.

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

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

Saleem, J. A., and G. D. Salvucci, 2002: Comparison of soil wetness indices for inducing functional similarity of hydrologic response across sites in Illinois. *J. Hydrometeor.*, **3**, 80–91, https://doi.org/10.1175/1525-7541(2002)003<0080:COSWIF>2.0.CO;2.

Salvucci, G. D., 2001: Estimating the moisture dependence of root zone water loss using conditionally averaged precipitation. *Water Resour. Res.*, **37**, 1357–1365, https://doi.org/10.1029/2000WR900336.

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

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

Shellito, P. J., and Coauthors, 2016: SMAP soil moisture drying more rapid than observed in situ following rainfall events. *Geophys. Res. Lett.*, **43**, 8068–8075, https://doi.org/10.1002/2016GL069946.

——, E. E. Small, and B. Livneh, 2018: Controls on surface soil drying rates observed by SMAP and simulated by the Noah land surface model. *Hydrol. Earth Syst. Sci.*, **22**, 1649–1663, https://doi.org/10.5194/hess-2017-338.

Svoboda, M., and Coauthors, 2002: The Drought Monitor. *Bull. Amer. Meteor. Soc.*, **83**, 1181–1190, https://doi.org/10.1175/1520-0477-83.8.1181.

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

Vivoni, E. R., H. A. Moreno, G. Mascaro, J. C. Rodriguez, C. J. Watts, J. Garatuza-Payan, and R. L. Scott, 2008: Observed relation between evapotranspiration and soil moisture in the North American monsoon region. *Geophys. Res. Lett.*, **35**, L22403, https://doi.org/10.1029/2008GL036001.

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

Zhang, Y., and Coauthors, 2016: Precipitation and carbon-water coupling jointly control the interannual variability of global land gross primary production. *Sci. Rep.*, **6**, 39748, https://doi.org/10.1038/srep39748.