Water Resources Research

RESEARCH ARTICLE
10.1029/2018WR023166

Global Dynamics of Stored Precipitation Water in the Topsoil Layer From Satellite and Reanalysis Data

Hyunglok Kim[1] and Venkat Lakshmi[1]

[1]The Department of Engineering Systems and Environment, University of Virginia, Charlottesville, Virginia, USA

Key Points:
- The drier the topsoil is, the greater the fraction of precipitated water stored in the topsoil layer
- Of the precipitated water, 37% ± 11% remains in the topsoil layer after 3 days of rainfall event where the aridity index is greater than 5
- Of the precipitated water, 21% ± 2% remains in the topsoil layer after three days of rainfall event where the aridity index is lower than 5

Supporting Information:
- Supporting Information S1

Correspondence to:
H. Kim,
hk5kp@virginia.edu

Citation:
Kim, H., & Lakshmi, V. (2019). Global dynamics of stored precipitation water in the topsoil layer from satellite and reanalysis data. *Water Resources Research*, *54*. https://doi.org/10.1029/2018WR023166

Received 19 APR 2018
Accepted 17 FEB 2019
Accepted article online 19 FEB 2019

Abstract The amount of soil water in the topsoil layer (from 0 to 10 cm) has been regarded as a key factor in controlling land-atmosphere interaction by determining the fraction of net radiation. In the present study, we investigate spatial trends of the stored precipitation fraction in the topsoil layer for varying vegetation and aridity indices by utilizing four satellites and two reanalysis data sets on a global scale. Using the Budyko framework, we relate climate regimes to the stored precipitation fraction on a global scale. A positive relation between the stored precipitation fraction with aridity index and a negative relation between the stored precipitation fraction and free parameter, vegetation optical depth, and isohydric slope are discovered. Even though the stored precipitation fraction values were calculated from different soil moisture and precipitation sources, they share an similar spatial trend: the drier and less vegetated the soil is, the more precipitation is retained in the top layer of the soil. Specifically, the topsoil retains 37% ± 11% of precipitated water three days after a rainfall event where the aridity index was greater than 5. Over wet and forest areas, due to large runoff fluxes and plants intercepting water before the precipitated water reached the ground, the topsoil retains 21% ± 2% of precipitated water three days after a rainfall event. Furthermore, by using the modeled data sets in the calculation of the stored precipitation fraction metric, we are able to conduct a sensitivity analysis of $F_P(f)$ metrics with respect to different sampling frequency values.

1. Introduction

Soil moisture plays a critical role in the global water cycle since the interaction between land and atmosphere is governed by the amount of moisture in the topsoil layer (Famiglietti et al., 1995; Hirschi et al., 2011; Lakshmi et al., 2004; Seneviratne et al., 2010; Tuttle & Salvucci, 2016). The variability of the near-surface atmospheric conditions (e.g., humidity, dust particles, and temperature profiles) have demonstrated high correlation with extreme climate events such as droughts, dust outbreaks, floods, and wildfires. In addition, soil moisture plays an indirect but important role is that latent and sensible heat fluxes are controlled by surface soil moisture, which affects boundary layer stability and low-altitude atmospheric conditions (Brocca et al., 2017; Crow et al., 2018; Delworth & Manabe, 1989; Haarsma et al., 2009; Kim et al., 2017). Even though surface soil moisture can be decoupled from root-zone soil moisture over dry-environment conditions (Hirschi et al., 2014), surface soil moisture can be a good indicator of the variability of deeper soil moisture in many cases (Choi & Jacobs, 2007; Dong & Crow, 2019; Qiu et al., 2016; Zohaib et al., 2017). Furthermore, the forecasting of extreme climate events can be significantly improved by investigating the residence time of soil moisture since soil moisture has a distinctively longer memory than other land surface variables (e.g., evapotranspiration and temperature) and atmospheric anomalies (e.g., dust storms and cyclones; Orth & Seneviratne, 2012); thus, knowledge of the dynamics of surface soil moisture is of importance in climate forecasting on seasonal time scales.

The residence time of surface soil moisture has been actively analyzed at scales ranging from the point to regional scales (Entin et al., 2000; Seneviratne et al., 2006). Soil moisture residence, also known as soil moisture memory, is often quantified by the implicit calculation of lag correlation of soil moisture or explicit estimation of a soil moisture memory time scale over specific time frames using long-term observations. Many previous researchers have conducted studies using a model or point-scale observation. However, in situ data-based estimations of soil moisture memory are limited in their ability to represent large-scale characteristics due to sparse and uneven locations of observations and heterogeneity of study regions. On the other hand, model-based estimates, which can fill in the missing pieces of in situ observations, have a different drawback: the calculated surface soil moisture memory can vary significantly from model to

model because of very different soil moisture dynamics, even when the models are driven by the same meteorological forcing (Dirmeyer et al., 2006; Koster et al., 2009). In addition, previous large-scale estimations of soil moisture memory studies have been limited to providing a proper understanding of residence time of soil moisture due to use of low temporal resolution data sets. In order to overcome the above limitations, and to track an increase in soil moisture corresponding to precipitation events, McColl et al. (2017) proposed a new metric called the stored precipitation fraction ($F_P(f)$) (McColl et al., 2017), which can be related with soil moisture residence time. This new measure of soil moisture dynamics is based on the sum of positive soil moisture increments, total precipitation, sampling depth, and data sampling frequency ($f\,\mathrm{day}^{-1}$) during the specified period. In contrast to other soil moisture memory metrics, the $F_P(f)$ metric can provide the actual fraction of moisture remaining in the topsoil layer after $1/f$ day(s) following a rainfall event. That is, the $F_P(f)$ metric can provide a clue to the complex hydrological processes taking place during sampling gaps. In addition, the $F_P(f)$ metric can provide even more unique information about the remaining amount of precipitated water if a high-frequency sampling data set is employed. For example, if we have a data set which samples instantaneous soil moisture values at a frequency f, this metric can provide the fraction of remaining precipitated water in the soil layer after n (integer) $\times\, 1/f$ day(s) of a rainfall event (detailed information can be found in section 3.2). In addition, the $F_P(f)$ equation does not require monthly or yearly observations to calculate time series autocorrelation, and it does not depend on data types—it can be calculated from short-term data set with in situ, satellite-based, and modeled data sets. However, with an irregular temporal sampling data set, the data set should be refined to include a fixed data sampling frequency before the $F_P(f)$ calculation is performed. For example, soil moisture retrieval satellite systems in Sun-synchronous orbit can provide soil moisture estimates with relatively fixed data sampling frequency, while soil moisture retrievals from a non-Sun-synchronous orbit which is based on the *signals of opportunity* produce an irregular data sampling frequency (Kim & Lakshmi, 2018); this kind of irregular data should be refined to produce a fixed data sampling frequency.

Significant strides have been made in estimating surface soil moisture from space at global and regional scales with relatively high spatial and temporal resolution (Burgin et al., 2017; Karthikeyan et al., 2017; Kim & Lakshmi, 2018). However, as with land surface models, microwave satellite-based soil moisture retrievals from different observation systems bear innate systemic and algorithmic limitations. For example, (1) low microwave frequency bands are susceptible to radio frequency interference (de Nijs et al., 2015); (2) the assumption of the soil moisture retrieval algorithms, such as thermal equilibrium and uniformity at 6 a.m., are violated in certain land cover conditions; and (3) signals from deeper layers cause unpredictable volume scattering over arid areas, and microwaves cannot penetrate high vegetation mass over densely forested regions. These limitations can cause erroneously high or low surface soil moisture retrievals (Kim et al., 2018). Therefore, it is plausible to assume that the $F_P(f)$ metric would show different behavior with respect to its input variables: different soil moisture and precipitation products. As $F_P(f)$ is a new metric, no study has yet been conducted to investigate the sensitivity of $F_P(f)$ metrics with respect to data sources with different data sampling frequencies and characteristics of $F_P(f)$ value with respect to different environmental conditions and land cover types.

The present research seeks to elucidate the priorities of the following questions: (1) do different soil moisture inputs from different microwave frequency-based satellite soil moisture retrievals (e.g., L-, C-, and X-bands) and off-line and coupled land surface models show different global trend in $F_P(f)$?; (2) how does the $F_P(f)$ value change with respect to different f values?; (3) what fraction of precipitated water can be retained in the topsoil layer after a series of rainfall events over a certain period of time, and what factors can affect the amount of stored precipitated water in the ground over arid and wet regions?; and (4) how do vegetation conditions control the variability of $F_P(f)$ on a global scale? Answering these questions will provide us new insights into the implementation of satellite-based $F_P(f)$ as well as model-derived $F_P(f)$ metrics; thus, it can aid in the future application of the $F_P(f)$ metric in climate and hydrology applications.

2. Data Sets

$F_P(f)$ metrics were calculated using four remotely sensed surface soil moisture data sets, half-hourly Global Precipitation Measurement (GPM) data, and two reanalysis data sets over the time period of January 2016 to

Table 1
Product Descriptions

Product Name	Variables	Spatial Resolution	Projection Information	Latitude, longitude	Temporal Resolution	Data Information
GLDAS NOAH025_3H (V1)	SM, ST, P	0.25° × 0.25°	Equiangular	720 × 1440	3-hourly	GLDAS Noah Land Surface Model L4 3-Hourly 0.25 × 0.25° v2.1 (https://search.earthdata.nasa.gov/search)
ERA-Interim land	SM, ST, P	0.25° × 0.25°	Equiangular	720 × 1440	6-hourly	ERA Interim, Daily (http://apps.ecmwf.int/datasets/data/interim-land/type=an/)
SMAP SM	SM	36 km × 36 km	EASE2	406 × 964	Varies	SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, version 4 (https://search.earthdata.nasa.gov/search)
ASCAT SM	SM	25 km resampled to GLDAS	Equiangular	720 × 1440	Varies	H-SAF H102 MetOp-A ASCAT NRT SSM orbit geometry 25-km sampling (WARP NRT 3.1.0)
AMSR2 SM	SM	0.25° × 0.25°	Equiangular	480 × 1440	Varies	Land parameter retrieval model (Parinussa, de Jeu, et al., 2016; Parinussa, Lakshmi, et al., 2016)
SMOS SM	SM	25 km × 25 km	EASE Grids version 1/2 of 25-km spatial resolution	584 × 1388	Varies	The SMOS INRA-CESBIO (SMOS-IC) from https://www.catds.fr/Products/Available-products-from-CEC-SM/SMOS-IC
GPM P	P	0.1° × 0.1°	Equiangular	1800 × 3600	Half-hourly	GPM_3IMERGHH: GPM IMERG Final Precipitation L3 Half-Hourly 0.1°×0.1° v04 (https://search.earthdata.nasa.gov/search)

Note. SM = Soil moisture; ST = Surface temprature; P = Precipitation.

December 2016 (Table 1). All $F_P(f)$ metrics from different data sources were calculated with their original data projections, and the final $F_P(f)$ metrics were reprojected into a regular 0.25° using a nearest-neighbor approach. Detailed information regarding the vegetation, and isohydricity slope data sets (Text S1), is included in the supporting information.

2.1. Satellite Data Sets

2.1.1. Soil Moisture Active Passive

The Soil Moisture Active Passive (SMAP) mission was launched in January 2015 (Entekhabi et al., 2010) and provides estimates of the moisture content of the top 0–5-cm soil layer in a near-polar and Sun-synchronous orbit with a revisit time of two to three days if the satellite overpasses the Equator at approximately 06:00 and 18:00 local time in descending and ascending orbits, respectively. Soil moisture product is derived from brightness temperature (Tb) observations provided by a low microwave frequency L-band radiometer (1.41 GHz, wavelength = 21 cm; O'Neill et al., 2015). In the present study, we considered the standard global daily 36-km spatial resolution level-3 SMAP soil moisture product from ascending and descending overpass time observations. The SMAP data were masked for soil moisture lower than 0.02 $m^3\ m^{-3}$ and higher than 0.50 $m^3\ m^{-3}$, and the retrieval information in the retrieval quality flag showed recommended quality (0 bit).

2.1.2. Soil Moisture and Ocean Salinity

The Soil Moisture and Ocean Salinity (SMOS) satellite was launched in November 2009. The SMOS accuracy requirements are 0.04 $m^3\ m^{-3}$ at a depth of 3–5-cm surface soil moisture with a revisit every two to three days (Kerr et al., 2001). SMOS provides estimates of the surface soil moisture in a near-polar and Sun-synchronous orbit, ascending in the morning at 06:00 local time and descending at 18:00 local time. The SMOS instrument is an L-band (1.41 GHz, wavelength = 21 cm) 2-D interferometer radiometer. In this study, we used the newest SMOS product, called SMOS-Institut National de la Recherche Agronomique-Centre d'Etudes Spatiales de la BIOsphère (SMOS-IC) product. The SMOS data were masked for soil moisture where the data quality flag was set to 1 (not recommended data quality). This masking process allowed for high-quality information in the soil moisture data considered (Fernandez-Moran et al., 2017).

2.1.3. Advanced Scatterometer

The MetOp-A satellite, which was considered in this study, was launched into a Sun-synchronous orbit in October 2006 and has been operational since May 2007. The Advanced Scatterometer (ASCAT) is an active system microwave remote-sensing instruments operated by the European organization. ASCAT acquires radar backscatter measurements at a frequency of C-band (5.3 GHz, wavelength = 5.7 cm) and has a

spatial resolution of 25 km with a one-to-three-day revisit time (Wagner et al., 2013). The ASCAT overpasses at 09:30 local time in descending orbit, and at 21:30 local in ascending orbit. The surface soil moisture content from ASCAT, so-called degree of saturation, was obtained from EUMETSAT site. This value was calculated from the basic measurement ASCAT—which is the backscattering coefficient ($\sigma°$) measured at a reference angle of 40° and based on the method of Wagner et al. (1999). For details about the change detection algorithm, please refer to Wagner et al. (2013). The porosity values were estimated by applying the equations of Saxton and Rawls (2006), and the texture characteristics were obtained from the Harmonized World Soil Database (Nachtergaele et al., 2009). The ASCAT data were masked, where wetland fractions >15%, topographic complexity >20%, surface soil moisture error >10%, and soil temperature was below the freezing point (Paulik et al., 2014).

2.1.4. Advanced Microwave Scanning Radiometer 2

The Advanced Microwave Scanning Radiometer 2 (AMSR2) sensor onboard the Global Change Observation Mission 1-Water platform was launched in May 2012. It utilizes three different microwave frequency bands at C1 (6.9 GHz, wavelength = 4.3 cm), C2 (7.3 GHz, wavelength = 4.1 cm), and X (10.6 GHz, wavelength = 2.8 cm), for surface soil moisture retrievals. The original ground resolutions depend on frequency channels (C1-band: 24 × 42 km, C2-band: 34 × 58 km, X-band: 35 × 62 km) with a revisit time of one to two days, and the AMSR2 crosses the equator at 01:30 local time and 13:30 local time in descending and ascending orbits, respectively (Maeda & Taniguchi, 2013). In this study, we used the most recently improved Land Parameter Retrieval Model AMSR2 data set with minimum radio frequency interference contamination (Kim et al., 2018; Parinussa, de Jeu, et al., 2016; Parinussa, Lakshmi, et al., 2016).

2.1.5. Global Precipitation Measurement

Launched on 20 February 2014, by NASA the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA), GPM is an international satellite mission intended to provide next-generation observations of precipitation and snowfall with a global scale every 3-hr. The GPM Core Observatory has a Ka + Ku radar and a 13-channel microwave radiometer (Huffman et al., 2014). There are many products available from the GPM Core Observatory; of relevance to present work are those associated with a high-quality, late-run GPM Core Observatory product that anchor a constellation of partner satellite radiometers used in producing the Integrated Multi-satellitE Retrievals (IMERG) at a temporal scale of every 30 min with 0.1° grid boxes. We used GPM IMERG Final precipitation L3 half-hourly data to calculate the satellite-based $F_P(f)$ metric.

2.2. Reanalysis Data Sets

In the present research, we investigated how the $F_P(f)$ value can vary with respect to different f values. Specifically, with regard to the high-frequency sampling data characteristics of reanalysis data, the use of reanalysis data would answer two questions: why the $F_P(f)$ value from high-sampling frequency data should be subsampled in order for comparison with the satellite-based $F_P(f)$ calculation and how the $F_P(f)$ value changes with respect to different f values. It is clear from satellite-based soil moisture data, with current microwave-based soil moisture retrievals satellite systems in Sun-synchronous orbit such as SMAP, SMOS, MetOp-A, and GCOM-W1, it is impossible to have $F_P(f)$ values higher than f ~= 1/3 days. Here we have noted that the model data cannot be regarded as the true value. However, in order to investigate higher sampling frequencies than those possible with satellite-based soil moisture data, using modeled soil moisture data is necessary; thus, in this study, we assumed that reanalysis data would be sufficient as an indicator of soil moisture states based on numerous previous studies (Liu et al., 2011; Spennemann et al., 2015).

2.2.1. Global Land Data Assimilation System

Global Land Data Assimilation System-1 (GLDAS) Noah available from 2000 onward (GLDAS-1) provides numerous atmospheric and land surface variables with a temporal resolution of 3-hr and a spatial resolution from 0.25° × 0.25° to 1° × 1°. GLDAS is a worldwide modeling system based on various satellite- and ground-based observations that produces the best estimation of land surface conditions (Rodell et al., 2004). The GLDAS precipitation value represents the past 3-hr averaged precipitation amount, while soil moisture and surface temperature are instantaneous variables. In the present study, we used the GLDAS data to calculate $F_P(f)$ using soil moisture from the topsoil layer of 0–10 cm and precipitation variables. Soil moisture value was masked out where surface temperature was below 273.15 K.

2.2.2. European Reanalysis Interim

The European Reanalysis Interim (ERA-Interim) product, created by the European Centre for Medium-Range Weather Forecast, has been available since 1979, and its analysis is expected to continue through the end of 2018. ERA-Interim provides numerous atmospheric and land surface variables with a temporal resolution of 3-hr and a spatial resolution from 0.125° × 0.125° to 3° × 3° (Dee et al., 2011). This data set provides instantaneous and accumulated forecast variables. Soil moisture and surface temperature, as instantaneous parameters, represent time scales equal to the model time step, while accumulated parameters, such as precipitation value, are accumulated from the start of the forecast. For detailed information, please refer the following link: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim. In this study, we used the ERA-Interim data set for the calculation of the ERA-Interim-based $F_P(f)$ metric using the top volumetric soil water layer 1 (0–7 cm), topsoil temperature layer 1, and large-scale precipitation variables. Soil moisture value was masked out where soil temperature layer one was below 273.15 K.

3. Methodology

3.1. The Stored Precipitation Fraction $F_P(f)$ From the Satellite Data Sets

The stored precipitation fraction $F_P(f)$ was proposed by McColl et al. (2017) as follows:

$$F_P(f) = \Delta z \frac{\sum_{i=1}^{n} \Delta\theta_{i+}}{\int_0^T P(t)dt}, \tag{1}$$

where

$$\Delta\theta_{i+} = \begin{cases} \Delta\theta_{i+}, & if\ \Delta\theta_i > 0\ and\ \Delta P_i > 0 \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

where Δz is the soil layer depth (m), $\Delta\theta_i$ is $\theta_i - \theta_{i-1}$, n is the number of valid cases of equation (2), P is the precipitation (m), f is a data sampling frequency (day^{-1}), and T (day) is the length of the sampling period. Specifically, the denominator of equation (1) represents the amount of accumulated precipitation events during the study period only when the equation (2) condition was met. Δz values are set to 0.05, 0.1, and 0.07 m for each satellite-based product, GLDAS, and ERA-Interim soil moisture products, respectively. For example, the $F_P(f)$ value for the GLDAS data set (e.g., $f = 8$ and $\Delta z = 0.1$ m) indicates the average proportion of precipitation falling on a soil layer that is still present in the soil layer (0–10 cm) after 1/8 days (or after 3-hr). In addition, we added a $\Delta P_i > 0$ condition in equation (2), rather than simply assuming that $\Delta\theta > 0$ was a positive value only when precipitation occurred. We found that many pixels showed positive $\Delta\theta$ without a precipitation event. Cases where $\Delta P_i = 0$ but $\Delta\theta > 0$ may occur due to irrigation, but they may also occur as a result of errors in soil moisture and precipitation observations. We will discuss this masking impact on the $F_P(f)$ calculation in depth in section 4.1.

3.2. The Calculation of the Stored Precipitation Fraction Using High-Frequency Data Sets

The highest data sampling frequency of global-scale assimilation soil moisture and precipitation data is limited to 8 day^{-1} (i.e., GLDAS and ERA-Interim). Thus, we have shown here the importance of selecting storm event(s) of a particular timing for the $F_P(f)$ calculation when using high sampling frequency data from GLDAS and ERA-Interim data sets.

In McColl et al. (2017), $F_P(f)$ was defined as "the average proportion of precipitation falling on a soil layer that is still present in the soil layer after $1/f$ days." To fully maintain the original significance of the $F_P(f)$ metric from the reanalysis data set-based $F_P(f)$ calculation (i.e., the calculation of $F_P(f)$ with a higher frequency sampling of soil moisture data), storm event(s) was(were) deliberately chosen. Figure 1a shows an example of a soil moisture time series from the GLDAS product with an 8 day^{-1} sampling frequency (black dots) and a soil moisture time series of 1/3 day^{-1} sampling frequency (red dots) with precipitation events. Here we have noted two cases of storm events in order to illustrate how the high sampling frequency data set-based $F_P(f)$ was calculated in this study.

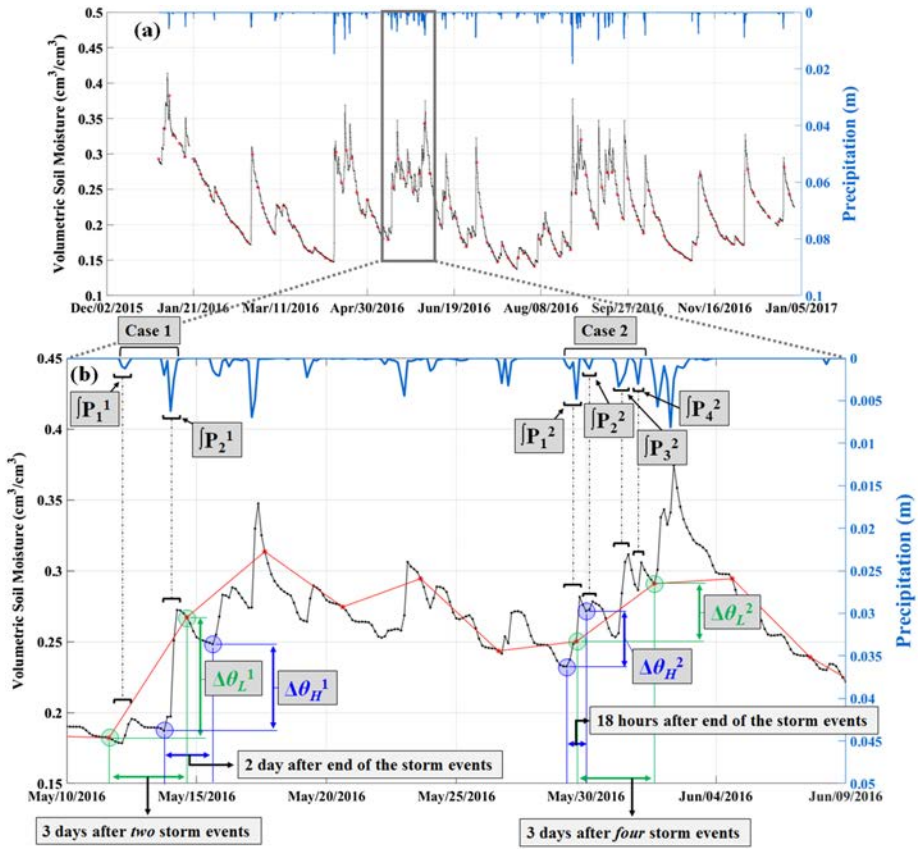

Figure 1. (a) Illustration of soil moisture variability with different sampling frequencies (f day^{-1}) with storm events. The black dots and solid line represent a high sampling frequency data set (i.e., reanalysis data; $f = 8$ day^{-1}), and the red dots and line represent a low sampling frequency data set (i.e., satellite data; $f = 1/3$ day^{-1}). In each series of storm event shown in the top part of Figure 1b, two soil moisture values of low sampling frequency data (green circles) and two soil moisture values of high sampling frequency data (blue circles) are illustrated. In this study, the incremental of soil moisture after end of the storm event(s) ($\Delta\theta_H$) is considered to calculate the $F_P(f)$ value using reanalysis data sets.

In the first case (upper axis marked as Case 1 in Figure 1b), the precipitation fraction with $f = 1/3$ can be calculated thus:

$$F_P\left(\frac{1}{3}\right)_L = \frac{\Delta\theta_L^1}{\int P^1}\Delta z = \frac{\Delta\theta_L^1}{\int P_1^1 + \int P_2^1}\Delta z$$

$F_P(f)_L$ indicates a short-term $F_P(f)$ value for P_n^1 storm events ($n = 1$ and 2) with low sampling frequency soil moisture data. In this case, $\Delta\theta_L^1$ has been calculated following two individual storm events, as shown by the time interval between the two green circles. In the present study, however, we used the $\Delta\theta_H^1$ value, as shown by the time interval between the two blue circles, since that is the soil moisture increment $1/f$ day(s) after end of storm event(s). With storm event Case 1, the various $F_P(f)_H$ $1/f$ day(s) after end of storm events can be calculated. For example, $F_P(1/2)_H$ can be calculated as follows:

$$F_P\left(\frac{1}{2}\right)_H = \frac{\Delta\theta_H^1}{\int P^1}\Delta z = \frac{\Delta\theta_H^1}{\int P_2^1}\Delta z$$

$F_P(f)_H$ indicates a short-term $F_P(f)$ value for a P_n^1 storm event ($n = 2$) with high sampling frequency soil moisture data. For equation (1), we sum up all $\Delta\theta_H$ cases for the numerator and integrate precipitation events that contribute to the $\Delta\theta_H$ value for the denominator.

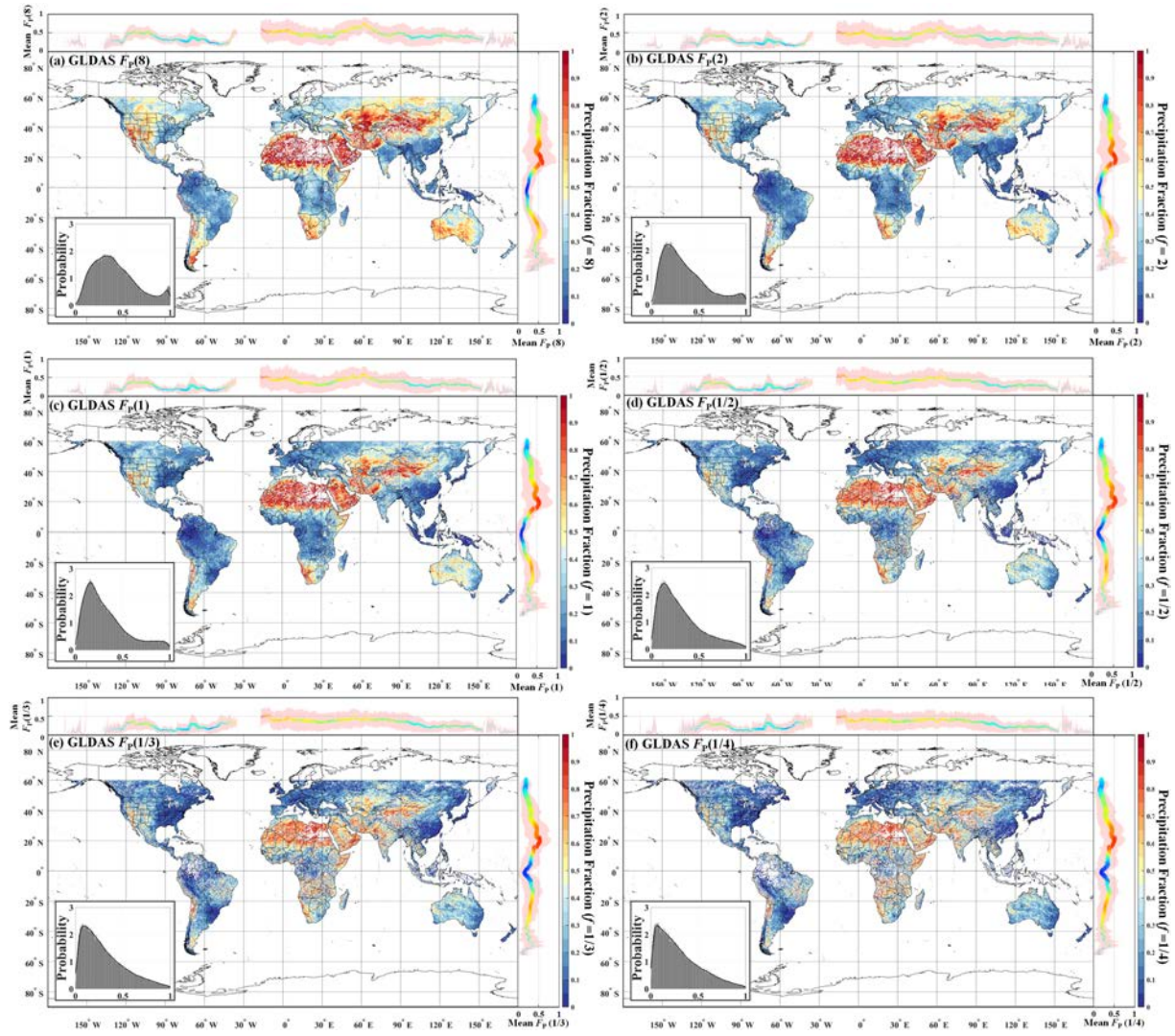

Figure 2. Maps of precipitation fraction based on six different sampling frequencies of GLDAS data: (a) $f = 8\ \text{day}^{-1}$, (b) $f = 2\ \text{day}^{-1}$, (c) $f = 1\ \text{day}^{-1}$, (d) $f = 1/2\ \text{day}^{-1}$, (e) $f = 1/3\ \text{day}^{-1}$, and (f) $f = 1/4\ \text{day}^{-1}$. Top and right panels indicate longitude and latitude zonal means of $F_P(f)$, respectively. Inset graphs represent PDF for each data set, marker sizes in zonal plots illustrate proportional to zonal land area, and the shaded region shows ±1 standard deviation. The opaque circle values in the zonal plot correspond to the zonal mean surface soil moisture.

With this selective process, $F_P(f)$ calculation using reanalysis data represents the average proportion of precipitation falling on a soil layer that is still present in the soil layer $1/f$ day(s) after storm events. Please note that $1/f$ is the time difference between the blue circles. Similarly, for the second case (upper axis marked as Case 2 in Figure 1(b)), there were four precipitation events over the course of three days: $P_1{}^2$, $P_2{}^2$, $P_3{}^2$, and $P_4{}^2$ (the time interval between the two green circles). In this case, we could calculate various $F_P(f)_H$ $1/f$ day(s) after storm event(s), such as $F_P(18/24)_H$ (the time interval between the two blue circles), instead of $F_P(1/3)_L$ (the time interval between the two green circles) as follows:

$$F_P\left(\frac{18}{24}\right)_H = \frac{\Delta\theta_H^2}{\int P^2}\Delta z = \frac{\Delta\theta_H^2}{\int P_1^2}\Delta z$$

Figure 2 shows the maps of $F_P(f)$ based on different sampling frequencies of GLDAS data. We found that $F_P(f)$ from higher f value reanalysis data showed higher values of $F_P(f)$ than lower f value reanalysis-based

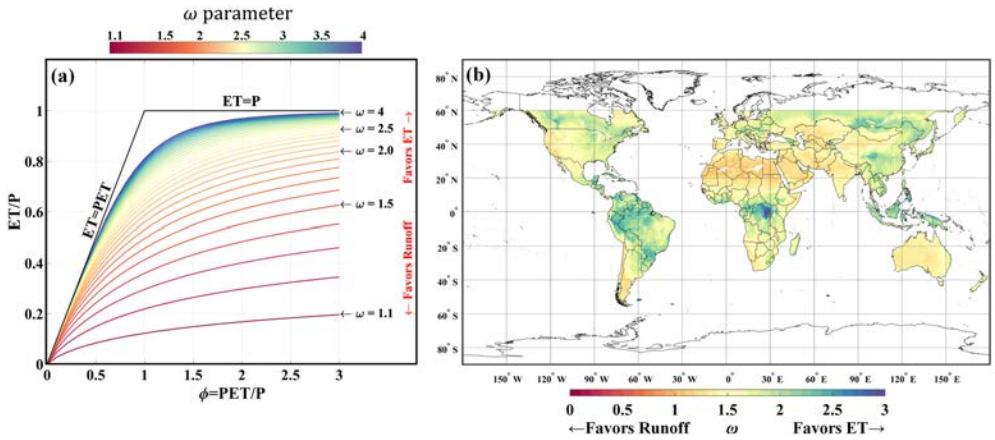

Figure 3. (a) ET/P from equation (3) for different values of the ω parameters are shown with variously colored solid lines. The original Budyko curve is also presented with a solid black line. (b) The global-scale map of the ω parameter.

$F_P(f)$ calculations—more positively skewed distribution from lower sampling data-based $F_P(f)$ values (i.e., the lower median for the lower f sampling data-based $F_P(f)$ values). The reason for this is clear: the total loss of soil moisture due to evapotranspiration, drainage, and runoff is much greater three days after rainfall event(s) (Figure 2e) than 3-hr after rainfall event(s) (Figure 2a). This result indicates that comparing the $F_P(8)$ from default sampling frequencies of reanalysis data with the $F_P(1/3)$ from satellite products would not produce any meaningful result; thus, reanalysis products should be subsampled to have the same sampling frequency as satellite products when reanalysis-based $F_P(f)$ values are compared with satellite-based $F_P(f)$ values. The global average of $F_P(1/3)$ is 0.31, which means that 31% of precipitated water remains in the topsoil layer 3-hr after rainfall event(s). In order to compare satellite- and reanalysis-based $F_P(1/3)$, two reanalysis data sets (i.e., GLDAS and ERA-Interim) were subsampled to estimate $F_P(1/3)$ values. In other words, the time interval between every adjacent sampling of GLDAS soil moisture data was set to 72 hr. These reanalysis-based $F_P(1/3)$ results were then compared with satellite-based $F_P(1/3)$ (please refer Figure 4 in section 4.2).

Furthermore, it is important to recall the advantage of the $F_P(f)$ metric and the reason for using reanalysis data in the $F_P(f)$ calculation. $F_P(f)$ is a measure of soil moisture memory that is dominated by drainage. The use of $F_P(f)$ is advantageous in these circumstances because it can be estimated solely from the time series of soil moisture and precipitation that can in turn be estimated directly from satellite observations; thus, calculating $F_P(f)$ from observation data sets is useful since reanalysis fields of soil moisture can be erroneous. However, in this study, we used reanalysis to investigate the sensitivity of the $F_P(f)$ metric with respect to f values and to understand why different f-based $F_P(f)$ values cannot be used interchangeably.

3.3. Aridity Index and ω Parameter From the Budyko Framework

We considered the ratio between potential evapotranspiration (PET) and precipitation (P) to investigate the impact of atmospheric evaporative demand on the variability of the $F_P(f)$ value. The aridity index commonly denoted as $[\overline{\phi}] = \frac{[\overline{PET}]}{[\overline{P}]}$ has been widely used for the investigation of the Budyko framework (Gentine et al., 2012), where $[\overline{PET}]$ is the climatological annual PET and $[\overline{P}]$ is the climatological annual precipitation (overbars refer to yearly means and brackets indicate long-term climatological means). In addition, the Budyko curve can be represented with evapotranspiration ET/P as a function of ϕ (Figure 3a). In the present research, we used an analytical solution to the Budyko framework following previous research by Padrón et al. (2017) to calculate ω parameter. This analytical equation is derived by combining dimensional analysis and mathematical reasoning as follows:

$$\frac{\left[\overline{ET}\right]}{\left[\overline{P}\right]} = 1 + \left[\phi\right] - \left(1 + \left[\phi\right]^{\omega}\right)^{\frac{1}{\omega}}, \quad (3)$$

where ω refers to a free parameter that integrates the net effect of all controls except ϕ on the partitioning of P into ET and runoff. The ω can be interpreted thus: the higher the ω is, the higher the favorability of the ET condition is. In other words, pixels with the same value of ϕ but with higher ω parameter have land cover conditions that favor ET over runoff (Padrón et al., 2017). The aridity index (ϕ) values and the ω parameters were calculated from GLDAS V2.0 using 63 years of data sets (from 1948 to 2010) on a global scale (Figure 3b).

A more detailed discussion of the methodology and the additional data sets can be found in the supporting information Burgin et al. (2017), Diem et al. (2014), Gruber et al. (2016), Holmes et al. (2006), Maggioni et al. (2016), and Owe et al. (2001).

4. Results and Discussion

4.1. The Probability of $\Delta\theta > 0$ When $\Delta P_i = 0$

It is found that many pixels showed positive $\Delta\theta$ without a precipitation event. Cases where $\Delta P_i = 0$ but $\Delta\theta > 0$ can occur not only due to irrigation but also due to errors in soil moisture and precipitation observations. In the first case, the irrigation effect can be significant over areas where irrigation occurs on an extraordinary scale. However, a few irrigated farms in a 40-km pixel might not substantially alter the 40-km average soil moisture observed by a satellite. In the other case, random, time-varying errors in soil moisture and precipitation observations can also lead to frequent cases where $\Delta P_i = 0$ but $\Delta\theta > 0$. Figure S1 was included to investigate this effect on $F_P(f)$ calculations by utilizing SMAP and GPM data sets. Figure S1a illustrates the probability of $\Delta\theta > 0$ when $\Delta P_i = 0$ during the study period. Figure S1b shows the global irrigation map from the Global Map of Irrigation Areas (Siebert et al., 2005). If we assume that Global Map of Irrigation Areas is the ground-truth data of irrigation areas, it is clear that the case $\Delta\theta > 0$ when $\Delta P_i = 0$ is common even in nonirrigated areas (Figure S1c). However, as shown in Figures S1c, this case is more frequent over irrigated regions than other areas.

4.2. Intercomparison of $F_P(f)$ Metrics From Satellite and Reanalysis Data Sets

Figure 4 illustrates the maps of $F_P(1/3)$ from (a) SMAP-GPM-based, (b) SMOS-GPM-based, (c) ASCAT-GPM-based, (d) AMSR2-GPM-based, (e) GLDAS-based, and (f) ERA-Interim-based calculations. The PDF of both satellite-estimated $F_P(1/3)$ and reanalysis-estimated $F_P(1/3)$ have right-skewed spatial distributions. Over sparsely vegetated regions (Figure S2a) all $F_P(1/3)$ except ERA-Interim show higher $F_P(1/3)$ than densely forested regions; ERA-Interim-based $F_P(1/3)$ values do not show this contrast as strongly as other $F_P(1/3)$ values. Furthermore, the Contiguous United States (CONUS) present a distinctive feature: the western side of the CONUS regions have a relatively higher $F_P(1/3)$ (average $F_P(1/3) = 0.29 \pm 0.06$) than the eastern side of the CONUS regions (average $F_P(1/3) = 0.15 \pm 0.07$; latitude ranges between 25° and 53° and longitude ranges between −120° and −67° in Figure 4). The longer residence of the precipitation falling on the top layer of the soil (i.e., high $F_P(1/3)$) indicates a lower drainage rate from the topsoil layer over Western CONUS. However, this process-level conclusion should be made carefully here since a similar pattern with a higher f value-based $F_P(f)$ calculation could also indicate the presence of significant surface runoff in Eastern CONUS. The previous study showed that groundwater depletion, defined as long-term water-level declines, is a problem in the Western CONUS (Famiglietti et al., 2011; Konikow, 2015). Therefore, high $F_p(f)$ over the Western CONUS areas can be interpreted as groundwater depletion. If more precipitated water remained in the topsoil layer, less water would reach the water table, and this would cause a long-term decline in water levels. In other words, as a result of high $F_p(f)$, Western CONUS could have a serious groundwater depletion due not only to the low amount of precipitation but also to the large fraction of stored precipitation in the topsoil layer. Furthermore, the average aridity indices are 3.93 and 1.08 for the western and eastern parts of CONUS, respectively (Figure 5). Thus, the high $F_P(f)$ value could be due to environmental conditions becoming drier, meaning more precipitated water would be captured in the topsoil layer, consequently leading to groundwater depletion. In addition, it is worth noting that the interpretation of current results is strongly dependent on whether the increased $F_P(f)$ is due to increased infiltration or decreased

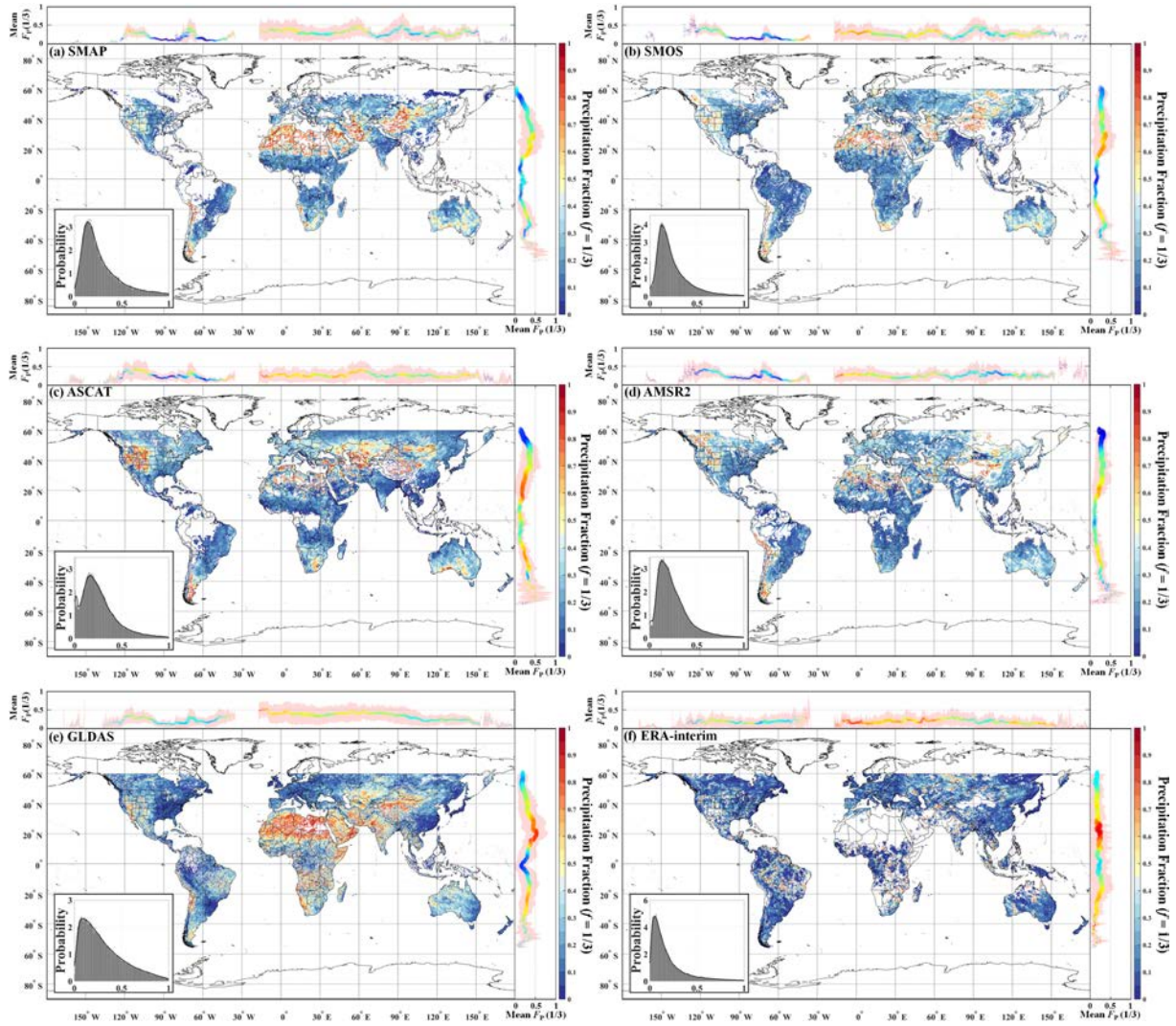

Figure 4. Global maps of $F_P(1/3)$ for (a) SMAP, (b) SMOS, (c) ASCAT, (d) AMSR2, (e) GLDAS, and (f) ERA-Interim. Top and right panels indicate longitude and latitude zonal means of $F_P(f)$, respectively. Inset graphs represent PDF for each data set and marker sizes in zonal plots illustrate proportional to zonal land area, and shaded region shows ±1 standard deviation. The opaque circle values in the zonal plot correspond to the zonal mean surface soil moisture.

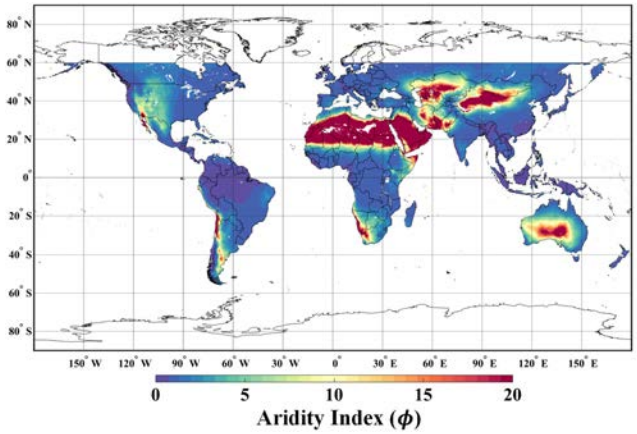

Figure 5. Map of the aridity index (ϕ).

recharge. Specifically, if the increased storage in Δz is due to reduced recharge (i.e., drainage out of the bottom of the root zone), a high $F_P(f)$ result can be interpreted to mean that the drier the environmental conditions are, the more precipitated water is captured in topsoil layers, consequently leading to groundwater depletion. Furthermore, if the increased storage in Δz is due to more infiltration (i.e., into the top of the root zone), a high $F_P(f)$ would likely reduce the occurrence of drought: as environmental conditions become drier, a greater amount of precipitated water infiltrates into the top of the root zone. To understand this result in greater detail, further investigation of $F_P(f)$ value using various f value will be considered in future research.

The longitude and latitude mean $F_P(1/3)$ comparison results are shown in Figure 6, and Table 2 shows the Spearman correlation (ρ) result between each data set with respect to longitude and latitude zonal means $F_P(1/3)$. ERA-based $F_P(1/3)$ showed the lowest agreement (i.e., the lowest ρ values) with other satellite-based $F_P(1/3)$ metrics for both longitude and latitude

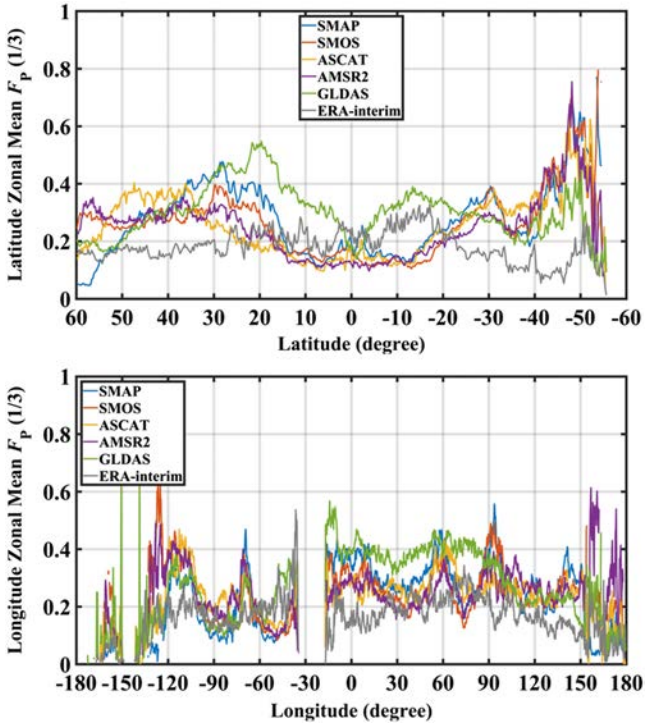

Figure 6. Graph showing the comparisons of (top) latitude zonal mean of $F_P(1/3)$ to (bottom) longitude zonal mean of $F_P(1/3)$. Blue lines indicate SMAP, orange lines indicate SMOS, yellow lines indicate ASCAT, purple lines indicate AMSR2, green lines indicate GLDAS, and gray lines indicate ERA-Interim-based $F_P(1/3)$ calculation.

zonal comparisons. However, two of the reanalysis-based $F_P(1/3)$ metrics showed high agreement (ρ value of latitude mean comparison = 0.47 and ρ value of longitude mean comparison = 0.55). Overall, the SMAP-based $F_P(1/3)$ metric shows good agreement with other satellite-based $F_P(1/3)$ metrics for both longitude and latitude zonal comparisons.

The values of $F_P(1/3)$ from satellite-based calculations have a similar standard deviation. This is because the satellite-based $F_P(1/3)$ metrics share the same precipitation sources (i.e., GPM data) as their input variable for the calculation of $F_P(1/3)$ values and have same Δz value (i.e., 0.05 m). In comparisons of the absolute values of satellite-based $F_P(1/3)$ and reanalysis-based $F_P(1/3)$, GLDAS-based $F_P(1/3)$ is generally higher than satellite-based $F_P(1/3)$; but ERA-Interim-based $F_P(1/3)$ is lower than other satellite-based $F_P(1/3)$ and GLDAS-based $F_P(1/3)$ (average $F_P(1/3)$ values are 0.245, 0.251, 0.233, 0.251, 0.269, and 0.164 for SMAP-, SMOS-, ASCAT-, AMSR2, GLDAS-, and ERA-based, respectively).

In the case of reanalysis-based $F_P(1/3)$, two reanalysis data sets were subsampled to have the same data frequency with the satellite-based $F_P(1/3)$, but they still showed different absolute $F_P(1/3)$ values. We found that the values of $\sum\Delta\theta_i$ from the ERA data were lower than those from the GLDAS product. This is because the soil moisture from ERA represents a thickness of 0–7 cm of the soil layer, while the GLDAS soil moisture represents a thickness of 0–10 cm from the soil layer. This point also indicates that the difference of the parametrization of drainage between the two reanalysis products creates different values of $F_P(1/3)$ in the two products. Additionally, the lesser thickness constant of Δz (0.07 m) in the ERA-based $F_P(1/3)$ calculation also contributes to lower values of $F_P(1/3)$ than GLDAS $F_P(1/3)$.

In addition, satellite-based and model-based $F_P(f)$ values cannot be used interchangeably for several reasons. This is because the sampling depth of soil moisture varies, causing different inputs of Δz in equation (1). The sampling depth of microwave-based soil moisture can change with respect to wavelength, soil wetness conditions, soil properties (i.e., bulk density), etc., while model-based $F_P(f)$ values are based on fixed sampling depths: soil moisture depth from 0 to 10 cm for GLDAS and 0–7 cm for ERA-Interim.

Table 2
Pairwise Spearman Correlation Coefficient Between Each Pair of (Top) Latitude and (Bottom) Longitude Zonal Mean $F_P(f)$ Metrics From SMAP-, SMOS-, ASCAT-, AMSR2-, GLDAS-, and ERA-Interim-Based Calculations, Respectively

	SMAP	SMOS	ASCAT	AMSR2	GLDAS	ERA-Interim	Average
Latitude							
SMAP	1.00	0.79**	0.66**	0.63**	0.29**	−0.24**	0.43
SMOS	0.79**	1.00	0.70**	0.85**	0.00	−0.55**	0.36
ASCAT	0.66**	0.7**	1.00	0.8**	−0.24**	−0.56**	0.27
AMSR2	0.63**	0.85**	0.80**	1.00	−0.11*	−0.57**	0.32
GLDAS	0.29**	0.00	−0.24**	−0.11*	1.00	0.47**	0.08
ERA-Interim	−0.24**	−0.55**	−0.56**	−0.57*	0.47*	1.00	−0.29
Longitude							
SMAP	1.00	0.69**	0.67**	0.40**	0.68**	0.12**	0.51
SMOS	0.69**	1.00	0.54**	0.68**	0.38**	−0.03	0.45
ASCAT	0.67**	0.54**	1.00	0.49**	0.56**	0.38**	0.53
AMSR2	0.40**	0.68**	0.49**	1.00	0.16**	−0.02	0.34
GLDAS	0.68**	0.38**	0.56**	0.16**	1.00	0.55**	0.47
ERA-Interim	0.12**	−0.03	0.38**	−0.02	0.55**	1.00	0.20

**p value<0.01 *p value<0.05

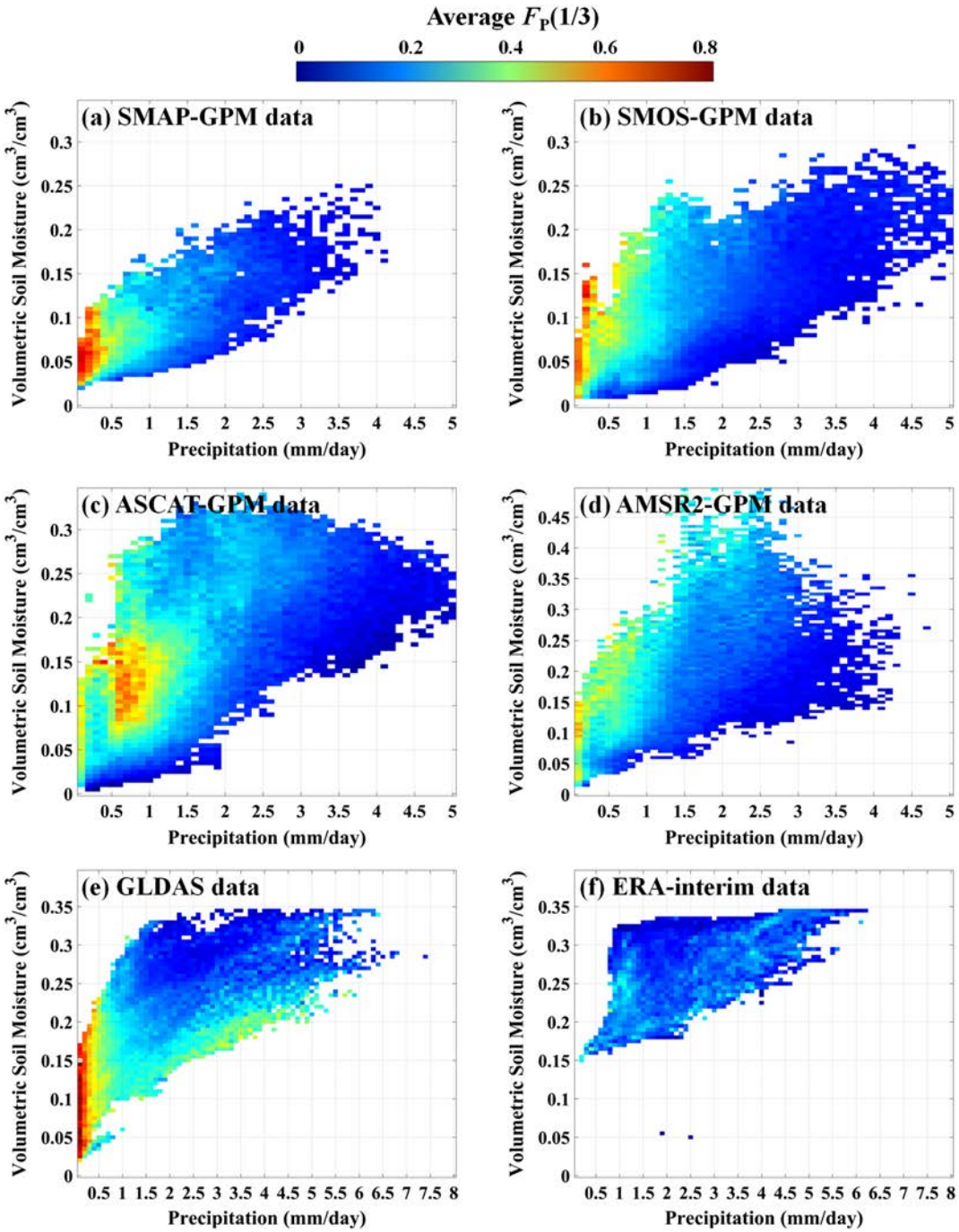

Figure 7. Heatmap shows the results of (a) SMAP-, (b) SMOS-, (c) ASCAT-, (d) AMSR2-, (e) GLDAS-, and (f) ERA-Interim-based F_P (1/3) metrics with mean precipitation and mean surface soil moisture. The pixel colors in the heatmap indicate average value of $F_P(1/3)$ within specific ranges of soil moisture and precipitation.

Figure 7 shows the heatmaps for the mean surface soil moisture, mean precipitation, and average F_P (1/3) (pixel color). We found higher $F_P(1/3)$ over lower precipitation regions. This is a common trend in all $F_P(1/3)$ from different input data sets. One possible reason for this high $F_P(1/3)$ in low-precipitation areas is that the soil matrix holds water too strongly due to low hydraulic conductivity in dry soil (Hillel, 1998; McColl et al., 2017). By contrast, over higher-precipitation areas, the $F_P(1/3)$ values are much lower than low-precipitation areas. One possible reason for this low $F_P(1/3)$ is that extreme precipitation over saturated or wet land surfaces will result in anomalies of runoff flux rather than soil moisture variability because the

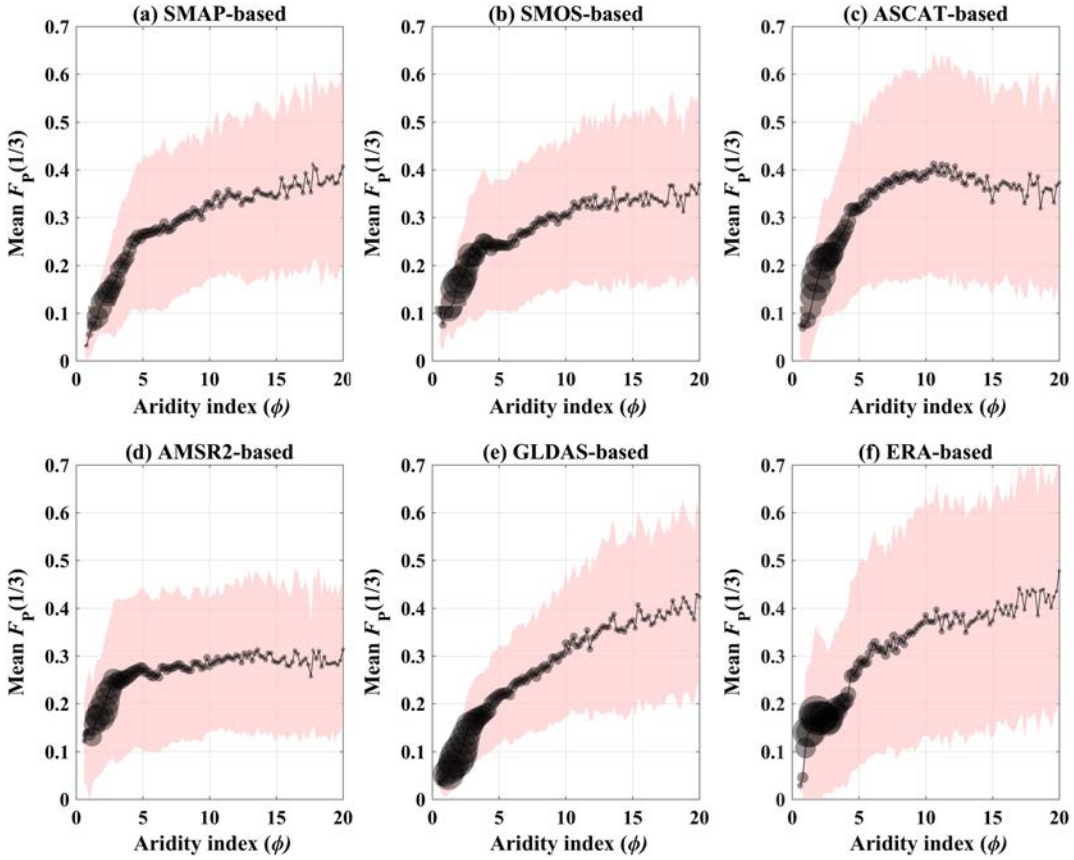

Figure 8. Graphs show the scatterplots of the mean $F_P(1/3)$ from (a) SMAP, (b) SMOS, (c) ASCAT, (d) AMSR2, (e) GLDAS, and (f) ERA-Interim data sets with respect to different values of the aridity index (ϕ). Marker sizes in the plots illustrate proportions to number of data sets, and the shaded region shows ±1 standard deviation.

precipitation rate is higher than the infiltration rate (Horton, 1933). These results will be further discussed in sections 4.3 and 4.4 considering aridity index (ϕ), ω parameter, and vegetation conditions.

4.3. Investigation of Global Spatial Variability of $F_P(f)$

The scatterplots in Figure 8 show the relationships between the $F_P(1/3)$ and the aridity index, and Figure 9 shows the relationship between $F_P(1/3)$ and ω values acquired from four satellites and two reanalysis data sets. It is interesting to note that even when $F_P(1/3)$ values were calculated from different data sources, the global tendency of $F_P(1/3)$ is to increase as the aridity index increases, and to decreases as ω decreases.

First of all, a high $F_P(f)$ represents a high fraction of precipitated water remaining in the ground after a precipitation event. In dry soil, precipitation falling on soil which is held too strongly by the soil matrix due to low hydraulic conductivity will result in a high $F_P(f)$ value (Hillel, 1998; McColl et al., 2017). Consequently, dry soil hampers precipitated water on the ground from undergoing evaporation in arid climate regimes. Higher $F_P(1/3)$ with higher ϕ (i.e., the drier the soil conditions are, the higher the remaining fraction of precipitated water on the ground is) and higher $F_P(1/3)$ with lower ω (i.e., the lower the hydraulic conductivity is, the less the ET-favorable condition is) support this finding (Figures 8 and 9). Furthermore, the response of coupled land-atmosphere processes should be carefully discussed here, especially in cases of negative feedback where the two sides are working in different directions and may have different strengths. This is especially important because the return side of the feedback loop (surface to atmosphere) is generally considered to be weaker than the forward side (atmosphere to surface). If we assume the forward part of the coupling loop is that low soil hydrologic conductivity causes high $F_P(f)$ and high $F_P(f)$ causes high PET, the return

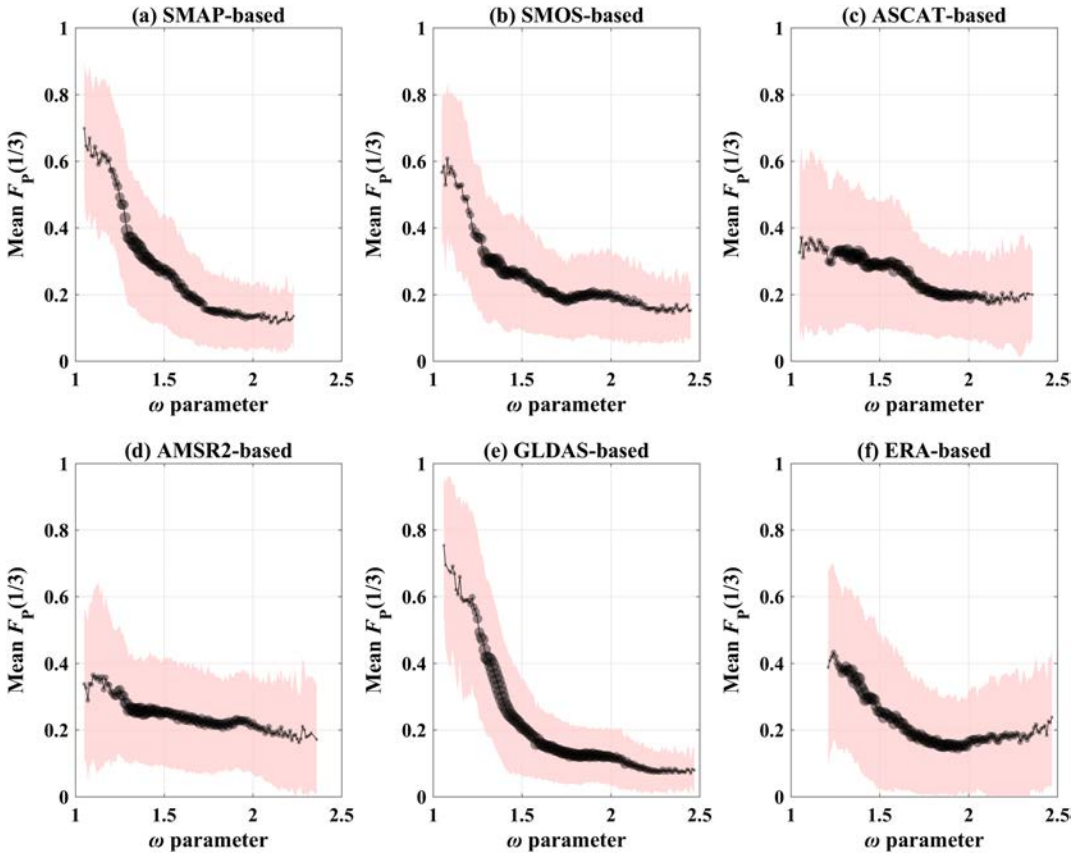

Figure 9. Graphs show the scatterplots of the mean $F_P(1/3)$ from (a) SMAP, (b) SMOS, (c) ASCAT, (d) AMSR2, (e) GLDAS, and (f) ERA-Interim data sets with respect to different values of the ω parameter. Marker sizes in the plots illustrate proportions to number of data sets, and the shaded region shows ± 1 standard deviation.

part of the loop will cause negative feedback. This is due to the fact that high PET causes low $F_P(f)$ because it encourages greater soil moisture loss due to evaporation. In specific, high $F_P(f)$ values in dry climate regimes, low hydraulic conductivity increases PET because soil moisture contents reduce ET cooling so that atmospheric heating from sensible heat flux increases. Consequently, a dry air mass will develop above an arid surface, creating a high vertical vapor pressure gradient (forward part). This high PET will cause low $F_P(f)$ (return part); however, this return part is generally regarded as weaker than the forward part.

Based on our current result, we concluded that these different strengths can be represented by the ω parameter. Over the higher ϕ regions, it is found that the return side of the feedback loop (surface to atmosphere) is generally weaker than the forward part (atmosphere to surface). In other words, we found a positive relationship between $F_P(f)$ and ϕ (Figure 8) and a negative relationship between $F_P(f)$ and ω (Figure 9)—the higher the $F_p(f)$ is (low hydraulic conductivity), the lower the ω is (less ET-favorable) and the higher the ϕ is (drier). Around 37% $\pm$ 11% of precipitated water (averaged from satellite- and reanalysis-based data sets) remains in the topsoil layer after three days of rainfall where the aridity index is greater than 5. It is worth noting that the variability of $F_P(f)$ increases as ϕ increases (Figure 8). This indicates that over arid climate conditions, the remaining fraction of precipitated water on the ground can be greatly affected by other landscape or climate factors such as soil characteristics, average slope, and mean annual temperature rather than by ϕ. In addition, this high variability of $F_P(f)$ over arid climate regions might represent the different strengths of the two feedback sides over arid regions. Some arid areas might be seen as having a weaker forward feedback side, which might be caused by the timing of the soil moisture observations from space. However, this assumption should be further investigated.

Second, over ET-favorable conditions (i.e., high ω parameter) with low ϕ areas, $F_P(1/3)$ is low—meaning that high ET flux prevents precipitated water in the topsoil layer from remaining longer than three days after

rainfall event(s). However, over wet climate regions with extremely high saturation-excess runoff conditions, much of the incoming precipitation runs off and does not make it into the soil at all because the soil moisture value reaches field capacity quickly after the rainfall event. These areas can be indicated with low $F_P(1/3)$, low ω parameter, and low ϕ. In this case, low $F_P(1/3)$ values do not indicate reduced soil memory or reduced soil residence times rather they simply reflect the fact that a substantial fraction of precipitation never enters the soil column in the first place. Wetter areas (which tend to be more highly vegetated) tend to have more saturated excess runoff and thus lower values of $F_P(f)$. In addition, over forested areas, large amounts of rainfall intercepted by vegetation canopies result in lower $\Delta\theta+$, leading to a smaller fraction of falling precipitation being retained in the surface soil layer. In short, in some areas where high amounts of saturation-excess runoff and vegetation mass, a lower fraction of precipitated water will remain in the topsoil because of high drainage, runoff fluxes, and vegetation canopies. Around $21\% \pm 2\%$ of precipitated water (averaged from satellite- and reanalysis-based data sets) remains in the topsoil layer after three days of rainfall where the aridity index is less than 5.

In addition, it is worth noting that further process-level insight can be fully addressed with other observation data sets such as the leaf area index data from optical sensors for investigating canopy impact on $F_P(f)$ or ΔLST anomalies value which can be observed from various sensors such as microwave and optical sensor (Anderson et al., 2007; Holmes et al., 2009; Price, 1993) for investigating inundation condition impact on $F_P(f)$ (Parinussa, de Jeu, et al., 2016; Parinussa, Lakshmi, et al., 2016). These kind of observation-based data sets that would further support the importance of one process explanation versus another. For example, considering leaf area index data in the calculation of the $F_p(f)$ value will address how much precipitated water is intercepted on canopies; ignoring the amount of water captured on canopies will underestimate the $F_p(f)$ value. Furthermore, considering the inundation conditions in calculation of the $F_p(f)$ will provide a better estimation of the amount of precipitated water contributed to surface runoff; ignoring the amount of surface runoff will result in underestimation of the $F_p(f)$ value.

4.4. Impact of Vegetation on $F_P(f)$ Values

Densely vegetated regions tend to be wetter than sparsely vegetated regions, and relatively high precipitation causes dampened soil, which results in increasing hydraulic conductivity; thus, water on the topsoil layer cannot be retained against gravitational drainage above the field capacity (Hillel, 1998). In these areas, the low-precipitation fraction could be caused not only by high runoff fluxes but also by plants intercepting water before it reaches the ground. The strong negative relationship between VOD and $F_P(1/3)$ supports this statement (Figure 10a). In addition, different types of vegetation have different plant wilting points (Bréda et al., 1995; Rambal et al., 2003) and isohydricity (Konings & Gentine, 2017), which indicates that different vegetation types can have different impacts on soil moisture residence times. Koning and Gentine (2017) found that, in seasonally dry locations, most ecosystems display a more isohydric response during the dry season (low isohydricity slope). In Figure 10b, we found a weak negative relationship between the isohydricity slope and $F_P(1/3)$ values. The isohydricity slope is 0 when the species are perfectly isohydric, and 1 or >1 indicates perfectly anisohydric species. Considering the fact that more isohydric species tend to keep leaf water potential relatively steady under water stress, it seems that more isohydric species contribute to keep a higher fraction of precipitated water on the ground than anisohydric species. By contrast, spatially designed farmlands that grow fast crops are more likely to be anisohydric (high isohydricity slope) because they keep their stomata relatively open to increase carbon uptake and use more water in the topsoil layer; thus, relatively low precipitated water remains in the topsoil layer after three days of rainfall event in comparison with farmlands growing species. This result also indicates that human conversion of bare land or forested regions to farmland could sufficiently change the remaining amount of precipitation falling on soil layers, such as the Corn Belt in United States that account for the greatest proportion of land use. However, there is one important caveat: the observed trend between the isohydric slope and the $F_P(1/3)$ is somehow weak and an extremely large number of (hidden) covariates could cause the observed weak correlation. Therefore, the relationship between the $F_P(f)$ and the isohydric slope need to be further investigated using different f value-based $F_P(f)$ along with detailed land use and land cover data sets.

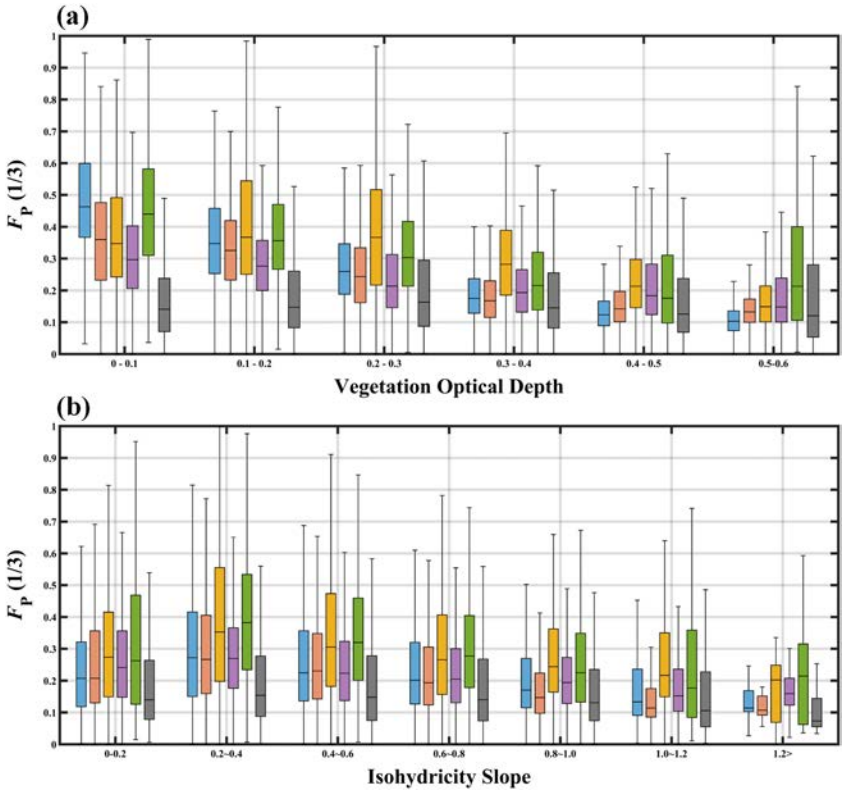

Figure 10. Box plots of $F_P(f)$ at different (a) vegetation optical depth and (b) isohydricity slope ranges for SMAP (blue boxes), SMOS (orange boxes), ASCAT (yellow boxes), AMSR2 (purple boxes), GLDAS (green boxes), and ERA-Interim (gray boxes) based calculations, respectively.

5. Conclusion

The purpose of this study is to analyze the stored precipitation fraction ($F_P(f)$) considering the degree of vegetation, aridity index, and ω parameters using various global-scale satellites and reanalysis products. The $F_P(f)$ metrics from the different data sources were then compared in order to investigate the sensitivity of the $F_P(f)$ to different input products. By studying the spatial variability of $F_P(f)$ metrics with respect to different aridity conditions, we were able to investigate new perspectives into process controls on the spatial variability of $F_P(f)$ over arid and wet regions. Furthermore, we subsampled reanalysis products and estimated the $F_P(1/3)$ from those reanalysis products. The reanalysis-based $F_P(1/3)$ results were then compared with satellite-based $F_P(1/3)$ (i.e., same data sampling frequency). The major findings of the present study are summarized below:

1. Even though $F_P(1/3)$ values were calculated from different soil moisture and precipitation sources, they shared an overall spatial trend: the drier the environment condition is, the more precipitated water remained in the topsoil layer.
2. A positive relation between the $F_P(f)$ and aridity index and a negative relation between the $F_P(f)$ and ω parameter were found. Globally, about 37% $\pm$ 11% of precipitated water (averaged from satellite- and reanalysis-based data sets) remains in the topsoil layer after three days of rainfall where the aridity index is greater than 5 because the soil matrix holds precipitated water on the topsoil layer too strongly due to low hydraulic conductivity. In wet conditions, around 21% $\pm$ 2% of precipitated water remains in the topsoil layer after three days of rainfall event where the aridity index is lower than 5 because much of the incoming precipitation runs off and does not make it into the topsoil layer.
3. Negative relations between $F_P(1/3)$ with VOD and isohydricity slope were found. High $F_P(1/3)$ with high ω parameter (high ET-favorable condition) and low ϕ support the fact that high amounts of rainfall captured by vegetation canopies and high evapotranspiration flux will result in a low fraction of falling

precipitation being retained in the surface soil layer. In addition, the weak negative relationship between the isohydricity slope and $F_P(1/3)$ value led us to conclude that more isohydric species seem to contribute to keeping a higher fraction of precipitated water on the ground than do aniohydric species, because isohydric species tend to keep leaf water potential relatively steady under water stress conditions.

In future studies, calculation of $F_P(f)$ metrics from different soil moisture layers and consideration of observation-based high sampling frequency soil moisture data (Kim & Lakshmi, 2018) and high spatial resolution of soil moisture products (Das et al., 2017) can further help elucidate the $F_P(f)$ metric and many geophysical applications. In addition, the investigation of longer data records of the stored precipitation fraction metric from satellite (Dorigo et al., 2017) and reanalysis data sets would also allow investigation of the impact of climate change on the water cycle.

Acknowledgments
We are particularly grateful to two anonymous reviewers and Wade Crow for their helpful comments. We acknowledge funding from the NASA Terrestrial Hydrology Program (Program Manager Jared Entin, grant NNX12AP75G). The authors thank the teams from NASA and USGS and ESA, Robert Parinussa, and Alexandra Konings for making their data sets publicly available. Data sets used in the present study can be downloaded from the following links: SMAP, GPM, and GLDAS-1 and GLDAS-2 (https://earthdata.nasa.gov/); SMOS (https://www.catds.fr/Products/Available-products-from-CEC-SM/SMOS-IC); ASCAT (https://eoportal.eumetsat.int); ERA-Interim (http://apps.ecmwf.int/datasets/data/interim-full-daily/); and GMIA (http://www.fao.org/nr/water/aquastat/irrigationmap/index.stm). The Isohydricity data are available from the authors on request (Konings & Gentone, 2017).

References

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. A., & Kustas, W. P. (2007). A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *Journal of Geophysical Research*, *112*, D10117. https://doi.org/10.1029/2006JD007506

Bréda, N., Granier, A., Barataud, F., & Moyne, C. (1995). Soil water dynamics in an oak stand. *Plant and Soil*, *172*(1), 17–27. https://doi.org/10.1007/BF00020856

Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. (2017). Soil moisture for hydrological applications: Open questions and new opportunities. *Water*, *9*(2), 140. https://doi.org/10.3390/w9020140

Burgin, M. S., Colliander, A., Njoku, E. G., Chan, S. K., Cabot, F., Kerr, Y. H., et al. (2017). A comparative study of the SMAP passive soil moisture product with existing satellite-based soil moisture products. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(5), 2959–2971. https://doi.org/10.1109/TGRS.2017.2656859

Choi, M., & Jacobs, J. M. (2007). Soil moisture variability of root zone profiles within SMEX02 remote sensing footprints. *Advances in Water Resources*, *30*(4), 883–896. https://doi.org/10.1016/j.advwatres.2006.07.007

Crow, W. T., Chen, F., Reichle, R. H., Xia, Y., & Liu, Q. (2018). Exploiting soil moisture, precipitation, and streamflow observations to evaluate soil moisture/runoff coupling in land surface models. *Geophysical Research Letters*, *45*, 4869–4878. https://doi.org/10.1029/2018GL077193

Das, N. N., Entekhabi, D., Kim, S., Jagdhuber, T., Dunbar, S., Yueh, S., & Colliander, A. (2017). High-resolution enhanced product based on SMAP active-passive approach using Sentinel 1 data and its applications. In *2017 IEEE International on Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 2493–2494). Fort Worth: IEEE. https://doi.org/10.1109/IGARSS.2017.8127500

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Delworth, T., & Manabe, S. (1989). The influence of soil wetness on near-surface atmospheric variability. *Journal of Climate*, *2*(12), 1447–1462. https://doi.org/10.1175/1520-0442(1989)002<1447:TIOSWO>2.0.CO;2

Diem, J. E., Hartter, J., Ryan, S. J., & Palace, M. W. (2014). Validation of satellite rainfall products for western Uganda. *Journal of Hydrometeorology*, *15*(5), 2030–2038. https://doi.org/10.1175/JHM-D-13-0193.1

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., & Hanasaki, N. (2006). GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, *87*(10), 1381–1398. https://doi.org/10.1175/BAMS-87-10-1381

Dong, J., & Crow, W. T. (2019). L-band remote-sensing increases sampled levels of global soil moisture-air temperature coupling strength. *Remote Sensing of Environment*, *220*, 51–58. https://doi.org/10.1016/j.rse.2018.10.024

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI soil moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

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

Entin, J. K., Robock, A., Vinnikov, K. Y., Hollinger, S. E., Liu, S., & Namkhai, A. (2000). Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research*, *105*(D9), 11,865–11,877. https://doi.org/10.1029/2000JD900051

Famiglietti, J. S., Braswell, B. H., & Giorgi, F. (1995). Process controls and similarity in the US continental-scale hydrological cycle from EOF analysis of regional climate model simulations. *Hydrological Processes*, *9*(3–4), 437–444. https://doi.org/10.1002/hyp.3360090314

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., et al. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, *38*, L03403. https://doi.org/10.1029/2010GL046442

Fernandez-Moran, R., al-Yaari, A., Mialon, A., Mahmoodi, A., al Bitar, A., de Lannoy, G., et al. (2017). SMOS-IC: An alternative SMOS soil moisture and vegetation optical depth product. *Remote Sensing*, *9*(5), 457. https://doi.org/10.3390/rs9050457

Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., & Salvucci, G. (2012). Interdependence of climate, soil, and vegetation as constrained by the Budyko curve. *Geophysical Research Letters*, *39*, L19404. https://doi.org/10.1029/2012GL053492

Gruber, A., Su, C. H., Zwieback, S., Crow, W., Dorigo, W., & Wagner, W. (2016). Recent advances in (soil moisture) triple collocation analysis. *International Journal of Applied Earth Observation and Geoinformation*, *45*, 200–211. https://doi.org/10.1016/j.jag.2015.09.002

Haarsma, R. J., Selten, F., Hazeleger, W., & Wang, X. (2009). Drier Mediterranean soils due to greenhouse warming bring easterly winds over summertime central Europe. *Geophysical Research Letters*, *36*, L04705. https://doi.org/10.1029/2008GL036617

Hillel, D. (1998). *Environmental Soil Physics: Fundamentals, Applications, and Environmental Considerations*. Waltham: Elsevier.

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

Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., et al. (2011). Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*, *4*(1), 17–21. https://doi.org/10.1038/ngeo1032

Holmes, T. R. H., De Jeu, R. A. M., Owe, M., & Dolman, A. J. (2009). Land surface temperature from Ka band (37 GHz) passive microwave observations. *Journal of Geophysical Research*, *114*, D04113. https://doi.org/10.1029/2008JD010257

Holmes, T. R. H., de Rosnay, P., de Jeu, R., Wigneron, R. J. P., Kerr, Y., Calvet, J. C., et al. (2006). A new parameterization of the effective temperature for L band radiometry. *Geophysical Research Letters*, *33*, L07405. https://doi.org/10.1029/2006GL025724

Horton, R. E. (1933). The role of infiltration in the hydrologic cycle. *Eos, Transactions American Geophysical Union*, *14*(1), 446–460. https://doi.org/10.1029/TR014i001p00446

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., & Yoo, S. H. (2014). *NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG)*. Greenbelt, MD: Algorithm Theoretical Basis Document (ATBD), NASA/GSFC.

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., & Wood, E. F. (2017). Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons. *Advances in Water Resources*, *109*, 236–252. https://doi.org/10.1016/j.advwatres.2017.09.010

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J. A. M. J., 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

Kim, H., & Lakshmi, V. (2018). Use of Cyclone Global Navigation Satellite System (CYGNSS) observations for estimation of soil moisture. *Geophysical Research Letters*, *45*, 8272–8282. https://doi.org/10.1029/2018GL078923

Kim, H., Parinussa, R., Konings, A. G., Wagner, W., Cosh, M. H., Lakshmi, V., et al. (2018). Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products. *Remote Sensing of Environment*, *204*, 260–275. https://doi.org/10.1016/j.rse.2017.10.026

Kim, H., Zohaib, M., Cho, E., Kerr, Y. H., & Choi, M. (2017). Development and assessment of the sand dust prediction model by utilizing microwave-based satellite soil moisture and reanalysis datasets in East Asian desert areas. *Advances in Meteorology*, *2017*(13), 1–13. https://doi.org/10.1155/2017/1917372

Konikow, L. F. (2015). Long-term groundwater depletion in the United States. *Groundwater*, *53*(1), 2–9. https://doi.org/10.1111/gwat.12306

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

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

Lakshmi, V., Piechota, T., Narayan, U., & Tang, C. (2004). Soil moisture as an indicator of weather extremes. *Geophysical Research Letters*, *31*, L11401. https://doi.org/10.1029/2004GL019930

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., de Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., et al. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrology and Earth System Sciences*, *15*(2), 425–436. https://doi.org/10.5194/hess-15-425-2011

Maeda, T., & Taniguchi, Y. (2013). *Descriptions of GCOM-W1 AMSR2 Level 1R and Level 2 Algorithms*. Ibaraki, Japan: Japan Aerospace Exploration Agency Earth Observation Research Center.

Maggioni, V., Meyers, P. C., & Robinson, M. D. (2016). A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. *Journal of Hydrometeorology*, *17*(4), 1101–1117. https://doi.org/10.1175/JHM-D-15-0190.1

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

Nachtergaele, F., van Velthuizen, H., Verelst, L., Batjes, N., Dijkshoorn, K., van Engelen, V., et al. (2009). Harmonized world soil database. Wageningen: ISRIC.

de Nijs, A. H., Parinussa, R. M., de Jeu, R. A., Schellekens, J., & Holmes, T. R. (2015). A methodology to determine radio-frequency interference in AMSR2 observations. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(9), 5148–5159. https://doi.org/10.1109/tgrs.2015.2417653

O'Neill, P. E., Njoku, E. G., Jackson, T. J., Chan, S., & Bindlish, R. (2015). SMAP algorithm theoretical basis document: Level 2 & 3 soil moisture (passive) data products. Jet Propulsion Lab., California Inst. Technol., Pasadena, CA, USA, JPL D-66480.

Orth, R., & Seneviratne, S. I. (2012). Analysis of soil moisture memory from observations in Europe. *Journal of Geophysical Research*, *117*, D15115. https://doi.org/10.1029/2011JD017366

Owe, M., de Jeu, R., & Walker, J. (2001). A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1643–1654. https://doi.org/10.1109/36.942542

Padrón, R. S., Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Large-scale controls of the surface water balance over land: Insights from a systematic review and meta-analysis. *Water Resources Research*, *53*, 9659–9678. https://doi.org/10.1002/2017WR021215

Parinussa, R. M., de Jeu, R. A., van der Schalie, R., Crow, W. T., Lei, F., & Holmes, T. R. (2016). A quasi-global approach to improve daytime satellite surface soil moisture anomalies through the land surface temperature input. *Climate*, *4*(4), 50. https://doi.org/10.3390/cli4040050

Parinussa, R. M., Lakshmi, V., Johnson, F. M., & Sharma, A. (2016). A new framework for monitoring flood inundation using readily available satellite data. *Geophysical Research Letters*, *43*, 2599–2605. https://doi.org/10.1002/2016GL068192

Paulik, C., Dorigo, W., Wagner, W., & Kidd, R. (2014). Validation of the ASCAT soil water index using in situ data from the International Soil Moisture Network. *International Journal of Applied Earth Observation and Geoinformation*, *30*, 1–8. https://doi.org/10.1016/j.jag.2014.01.007

Price, J. C. (1993). Estimating leaf area index from satellite data. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(3), 727–734. https://doi.org/10.1109/36.225538

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. *Journal of Hydrometeorology*, *17*(9), 2419–2430. https://doi.org/10.1175/JHM-D-16-0044.1

Rambal, S., Ourcival, J. M., Joffre, R., Mouillot, F., Nouvellon, Y., Reichstein, M., & Rocheteau, A. (2003). Drought controls over conductance and assimilation of a Mediterranean evergreen ecosystem: Scaling from leaf to canopy. *Global Change Biology*, *9*(12), 1813–1824. https://doi.org/10.1111/j.1365-2486.2003.00687.x

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Saxton, K. E., & Rawls, W. J. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal*, *70*(5), 1569–1578. https://doi.org/10.2136/sssaj2005.0117

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

Seneviratne, S. I., Koster, R. D., Guo, Z., Dirmeyer, P. A., Kowalczyk, E., Lawrence, D., et al. (2006). Soil moisture memory in AGCM simulations: Analysis of global land–atmosphere coupling experiment (GLACE) data. *Journal of Hydrometeorology*, *7*(5), 1090–1112. https://doi.org/10.1175/JHM533.1

Siebert, S., Döll, P., Hoogeveen, J., Faures, J. M., Frenken, K., & Feick, S. (2005). Development and validation of the global map of irrigation areas. *Hydrology and Earth System Sciences Discussions*, *2*(4), 1299–1327.

Spennemann, P. C., Rivera, J. A., Saulo, A. C., & Penalba, O. C. (2015). A comparison of GLDAS soil moisture anomalies against standardized precipitation index and multisatellite estimations over South America. *Journal of Hydrometeorology*, *16*(1), 158–171. https://doi.org/10.1175/JHM-D-13-0190.1

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

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., et al. (2013). The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorologische Zeitschrift*, *22*(1), 5–33. https://doi.org/10.1127/0941-2948/2013/0399

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

Zohaib, M., Kim, H., & Choi, M. (2017). Evaluating the patterns of spatiotemporal trends of root zone soil moisture in major climate regions in East Asia. *Journal of Geophysical Research: Atmospheres*, *122*, 7705–7722. https://doi.org/10.1002/2016JD026379