Key Points:
- Version 4 of the SMAP Level-4 Soil Moisture product features several model and analysis improvements compared to previous versions
- SMAP Level-4 surface soil moisture, soil moisture uncertainty, and runoff estimates have lower bias in Version 4 than in previous versions
- SMAP Level-4 analysis adjustments are smaller in Version 4 than in previous versions while achieving similar overall product skill

Supporting Information:
- Supporting Information S1

Correspondence to:
R. H. Reichle,
rolf.reichle@nasa.gov

Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product

Rolf H. Reichle[1] , Qing Liu[1,2], Randal D. Koster[1] , Wade T. Crow[3] , Gabrielle J. M. De Lannoy[4] , John S. Kimball[5], Joseph V. Ardizzone[1,2], David Bosch[6] , Andreas Colliander[7] , Michael Cosh[3] , Jana Kolassa[1,8], Sarith P. Mahanama[1,2], John Prueger[9], Patrick Starks[10] , and Jeffrey P. Walker[11]

[1]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA, [2]Science Systems and Applications, Inc., Lanham, MD, USA, [3]Hydrology and Remote Sensing Laboratory, Agricultural Research Service, USDA, Beltsville, MD, USA, [4]Department of Earth and Environmental Sciences, KU Leuven, Heverlee, Belgium, [5]NTSG, University of Montana, Missoula, MT, USA, [6]Southeast Watershed Research, Agricultural Research Service, USDA, Tifton, GA, USA, [7]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [8]GESTAR, Universities Space Research Association, Columbia, MD, USA, [9]National Laboratory for Agriculture and the Environment, Agricultural Research Service, USDA, Ames, IA, USA, [10]Grazinglands Research Laboratory, Agricultural Research Service, USDA, El Reno, OK, USA, [11]Department of Civil Engineering, Monash University, Clayton, Victoria, Australia

Abstract The NASA Soil Moisture Active Passive (SMAP) mission Level-4 Soil Moisture (L4_SM) product provides global, 3-hourly, 9-km resolution estimates of surface (0–5 cm) and root zone (0–100 cm) soil moisture with a mean latency of ~2.5 days. The underlying L4_SM algorithm assimilates SMAP radiometer brightness temperature (Tb) observations into the NASA Catchment land surface model using a spatially distributed ensemble Kalman filter. In Version 4 of the L4_SM modeling system the upward recharge of surface soil moisture from below under nonequilibrium conditions was reduced, resulting in less bias and improved dynamic range of L4_SM surface soil moisture compared to earlier versions. This change and additional technical modifications to the system reduce the mean and standard deviation of the observation-minus-forecast Tb residuals and overall soil moisture analysis increments while maintaining the skill of the L4_SM soil moisture estimates versus independent in situ measurements; the average, bias-adjusted root-mean-square error in Version 4 is 0.039 m^3/m^3 for surface and 0.026 m^3/m^3 for root zone soil moisture. Moreover, the coverage of assimilated SMAP observations in Version 4 is near global owing to the use of additional satellite Tb records for algorithm calibration. L4_SM soil moisture uncertainty estimates are biased low (by 0.01–0.02 m^3/m^3) against actual errors (computed versus in situ measurements). L4_SM runoff estimates, an additional product of the L4_SM algorithm, are biased low (by 35 mm/year) against streamflow measurements. Compared to Version 3, bias in Version 4 is reduced by 46% for surface soil moisture uncertainty estimates and by 33% for runoff estimates.

Plain Language Summary Soil moisture is important because of its impact on the land surface water, energy, and nutrient cycles. The low-frequency microwave observations collected by the NASA Soil Moisture Active Passive (SMAP) satellite are suitable for estimating soil moisture globally. Their sensitivity, however, is limited to the top few centimeters of the soil, and observations are only available every other day depending on location. The SMAP Level-4 Soil Moisture (L4_SM) data product addresses these limitations by merging the satellite observations into a numerical model of the land surface water and energy balance while considering the uncertainty of the observations and model estimates. The resulting L4_SM data product is publicly disseminated within ~2.5 days from the time of observation and provides global estimates of surface and deeper-layer (or "root zone") soil moisture at 3-hourly temporal and 9-km spatial resolution. This study presents an overview of recent updates in the L4_SM algorithm and an assessment of the quality of the resulting Version-4 soil moisture estimates. The algorithm updates reduce bias in L4_SM surface soil moisture and runoff estimates compared to previous versions while otherwise maintaining the product skill, which meets the accuracy requirement specified prior to SMAP's launch.

1. Introduction

The Soil Moisture Ocean Salinity (SMOS; Kerr et al., 2010; Mecklenburg et al., 2016) and Soil Moisture Active Passive (SMAP; Entekhabi et al., 2010) satellite missions measure Earth's L-band (1.4 GHz) passive microwave brightness temperature (Tb), which is highly sensitive to the water content in the top few centimeters of the soil and in the vegetation (Jackson & Schmugge, 1991; Schmugge et al., 1974). With their global coverage and 2- to 3-day revisit time, the L-band observations have revealed new insights into soil moisture drydown characteristics and basin-scale water balance (Akbar et al., 2018; Koster, Crow, et al., 2018; Shellito et al., 2016, 2018), drought (Liu et al., 2017; Mishra et al., 2017; Pablos et al., 2017; Rajasekaran et al., 2018; Sadri et al., 2018), soil moisture memory and its impact on land-atmosphere coupling (Dong & Crow, 2019; McColl et al., 2017), the relation between surface soil moisture and subsoil characteristics (Dirmeyer & Norton, 2018), and vegetation optical depth and plant activity (Feldman, Akbar, & Entekhabi, 2018; Feldman, Short Gianotti, et al., 2018; Konings et al., 2017). Moreover, L-band observations have been used to estimate precipitation (Brocca et al., 2016; Koster et al., 2016; Román-Cascón et al., 2017), to enhance estimates of evapotranspiration and its partitioning into bare soil evaporation and transpiration (Purdy et al., 2018; Rigden et al., 2018), and to improve the characterization of the runoff ratio for flood forecasting and land surface model calibration and evaluation (Crow et al., 2017, 2018, 2019; Koster, Liu, et al., 2018). Recent results also confirm that L-band soil moisture retrievals, and in particular retrievals from SMAP, are generally more accurate than retrievals from C- and X-band instruments (Chen et al., 2018; Kumar et al., 2018).

Among the global soil moisture data sets partly based on L-band observations is the SMAP Level-4 Surface and Root Zone Soil Moisture (L4_SM) product (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017; Reichle, De Lannoy, Liu, Koster, et al., 2017). The L4_SM product is derived from the assimilation of SMAP Tb observations into a land surface model driven with observed precipitation forcing to the extent possible. The L4_SM algorithm thus merges the information from the L-band Tb observations with the water and energy balance constraints encapsulated in the land surface model and the information in the surface meteorological forcing data, which is derived from gauge-based precipitation observations and a large number of atmospheric measurements. Based on the merged information, the L4_SM product provides global, 3-hourly, 9-km resolution soil moisture estimates for the surface (0–5 cm), root zone (0–100 cm), and profile (0 cm to bedrock) layers with a mean latency of ~2.5 days from the time of observation. As of January 2019, L4_SM data have been downloaded by more than 3,000 individual users (based on unique Internet Protocol host addresses) according to the National Snow and Ice Data Center.

The value of the L4_SM product has been demonstrated in a variety of ways. A case study in Australia showed that assimilating SMAP observations successfully corrected soil moisture for short-term errors in the L4_SM rainfall forcing (Reichle, De Lannoy, Liu, Koster, et al., 2017). Validation with in situ measurements indicated that Version 2 of the L4_SM product met its soil moisture accuracy requirement, with a bias-adjusted (or unbiased) root-mean-square error (ubRMSE, also known as standard deviation of the error) of 0.038 m^3/m^3 (0.030 m^3/m^3) for surface (root zone) soil moisture at the 9-km scale (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017). Root zone soil moisture estimates from the L4_SM product also performed well in a multiproduct comparison versus in situ measurements in Spain (Pablos et al., 2018). Moreover, when compared to in situ measurements across the globe, the L4_SM soil moisture estimates were more skillful than model-only estimates that do not benefit from the assimilation of SMAP Tb observations (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017). L4_SM soil moisture estimates could also better predict the streamflow response to subsequent precipitation than model-only estimates and L-band (or C-band) soil moisture retrievals by themselves (Crow et al., 2017). Finally, L4_SM data have been used successfully to evaluate and calibrate the coupling strength between soil moisture and the subsequent, storm-scale runoff ratio in several different land surface models (Crow et al., 2018, 2019).

The observation-minus-forecast (O-F) Tb residuals of the Version 2 L4_SM product had a small global mean bias of ~0.37 K, modest regional (absolute) biases under ~3 K, and typical instantaneous (absolute) values of ~6 K (Reichle, De Lannoy, Liu, Koster, et al., 2017). The modeled Tb ensemble spread, however, overestimated the actual Tb errors in deserts and densely vegetated regions and underestimated them in agricultural regions and transition zones between dry and wet climates (Reichle, De Lannoy, Liu, Koster, et al., 2017). Furthermore, SMAP observations were not assimilated in the Version 2 algorithm across a large region spanning from Eastern Europe to China (Reichle, De Lannoy, Liu, Koster, et al., 2017; section 2.2.1). Moreover,

Reichle, De Lannoy, Liu, Ardizzone, et al. (2017) reported a general wet bias in Version 2 L4_SM soil moisture estimates, particularly for surface soil moisture. Finally, the calibration and validation of the Version 2 algorithm was necessarily based on shorter data records than are available today.

The more recent L4_SM Versions 3 and 4 include substantial modifications of the land model and Tb analysis that address some of the shortcomings of Version 2 (section 2). Moreover, the L4_SM product also provides estimates of soil moisture uncertainty and runoff that have never been validated. The primary objectives of the present study are thus to (i) validate the current Version 4 soil moisture, soil moisture uncertainty, and runoff estimates alongside the previous Version 3 data (section 4), (ii) evaluate the L4_SM data assimilation diagnostics, including the statistics of the O-F Tb residuals and the soil moisture analysis increments (section 5), and (iii), throughout the analysis, relate the differences in performance and key diagnostics to the algorithm changes. The validation data and methods used here are described in section 3. We focus here on comparing the current Version 4 to Version 3 because we have three full years of overlapping data and because the differences between the Version 2 and 3 data are minimal at all in situ measurement sites used for validation (section 2.2.1). Thus, in effect, we quantify the performance of the Version 4 product relative to the documented performance of the Version 2 data (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017; Reichle, De Lannoy, Liu, Koster, et al., 2017).

2. L4_SM Algorithm and Data Product

2.1. Overview

The L4_SM algorithm consists of the assimilation of SMAP Tb observations into a land surface model (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017, their Figure 1). The essential ingredients of this system are the following:

1. horizontal- and vertical-polarization Tb observations from ascending and descending half-orbits of the 36-km resolution SMAP Level-1C product (Piepmeier et al., 2017),
2. the NASA Catchment land surface model (Ducharne et al., 2000; Koster et al., 2000), supplemented with a tau-omega radiative transfer model for L-band Tb (De Lannoy et al., 2013, 2014),
3. surface meteorological forcing data from the NASA Goddard Earth Observing System (GEOS) Forward-Processing (FP) system (Lucchesi, 2013),
4. the Center for Climate Prediction Unified (CPCU) 0.5°, daily precipitation product (Chen et al., 2008; Xie et al., 2007) rescaled so that its climatology matches that of the Global Precipitation Climatology Project (GPCPv2.2; Huffman et al., 2009) data, and
5. the GEOS land assimilation system consisting of a spatially distributed ensemble Kalman filter (EnKF; Reichle, De Lannoy, Liu, Koster, et al., 2017).

The L4_SM estimates are generated on the global, 9-km resolution Equal Area Scalable Earth, version 2 (EASEv2) grid (Brodzik et al., 2012). In addition to soil moisture, the L4_SM product includes a variety of land model variables and assimilation diagnostics (Reichle, Lucchesi, et al., 2018).

Briefly, the L4_SM algorithm consists of driving an ensemble of Catchment model simulations with surface meteorological forcing data from the GEOS FP system after correcting the GEOS precipitation with the CPCU observational product. The ensemble spread represents model uncertainty and is maintained by adding perturbations to the model forcing and prognostic variables as described in Reichle, De Lannoy, Liu, Ardizzone, et al. (2017). Once every 3 hr at 00:00 UTC, 03:00 UTC, etc., the available SMAP Tb observations within a 3-hr window centered on the analysis time are compared with the corresponding model forecast Tb estimates, separately for horizontal and vertical polarization (after first averaging the corresponding fore- and aft-looking SMAP Tb observations). Based on the O-F Tb differences, and taking into consideration the relative uncertainties of each, the EnKF analysis generates soil moisture and temperature increments. These increments are then added to the model estimates in order to make the modeled soil moisture and temperature fields more consistent with the SMAP Tb observations. While the Tb observations are typically only sensitive to soil moisture in the top few centimeters of the soil, the EnKF analysis also generates increments for root zone soil moisture. This is possible by including unobserved Catchment model variables related to deeper-layer soil moisture in the EnKF state vector and calculating increments based on the simulated (ensemble) error correlations between the surface and root zone layers (Reichle, De Lannoy, Liu, Koster, et al., 2017).

During the assimilation process, the L4_SM algorithm interpolates and extrapolates information from the SMAP Tb observations and the model forecast estimates in time and in space; the resulting L4_SM data product represents the merged information. The degree to which L4_SM soil moisture estimates at a given time, location, and depth are impacted by SMAP Tb observations depends on the distance (in space and in time) from the observations, the sensitivity of the observations to surface soil moisture, and the coupling between the surface and root zone layers. Under dense vegetation, for instance, or when the ground is frozen or snow-covered, SMAP Tb observations have little or no immediate impact on the L4_SM soil moisture. Likewise, if the surface and root zone layers are decoupled, SMAP Tb observations have little immediate impact on L4_SM root zone soil moisture.

An integral part of the calibration of the L4_SM system is the removal of seasonally varying bias in the modeled Tb. To this end, seasonally varying and spatially distributed Tb "scaling parameters" were computed as the multiyear mean difference between L-band Tb observations and corresponding model estimates. Prior to each 3-hourly Tb analysis, the appropriate scaling parameters are subtracted from the O-F Tb residuals such that only Tb anomaly information is retained in the EnKF analysis (Reichle, De Lannoy, Liu, Koster, et al., 2017). That is, the system is designed to correct only for errors in short-term and interannual variations and not for errors in the climatological seasonal cycles of the modeled soil moisture or other land surface fields.

2.2. Changes From Version 2 to Version 4

Version 2 of the L4_SM algorithm and product is discussed in great detail by Reichle, De Lannoy, Liu, Ardizzone, et al. (2017) and Reichle, De Lannoy, Liu, Koster, et al. (2017). Using Version 2 as a baseline, Reichle, Liu, et al. (2018) provide a comprehensive project report of the system changes leading to Version 4. The present section summarizes only the most important of these changes. Unless noted otherwise here or by Reichle, Liu, et al. (2018), the configuration of the Version 4 system matches that of the Version 2 system. For example, the ensemble size of 24 members and the model and observation error settings are preserved. Likewise, the overall approach for the Catchment model and ensemble spin-up is unchanged (except for the changes in the model version and forcing inputs discussed below).

2.2.1. Changes From Version 2 to Version 3

Version 2 of the L4_SM algorithm, owing to a calibration based solely on SMOS Tb data, was limited to assimilating SMAP Tb observations only in regions where radio frequency interference (RFI) does not adversely impact SMOS data. In Eastern Europe, the Middle East, and East Asia, SMOS observations are corrupted by RFI (Oliva et al., 2012) to such an extent that SMOS cannot provide the L-band climatology necessary to compute the Tb scaling parameters for the L4_SM algorithm. Consequently, SMAP observations were not assimilated in the Version 2 algorithm in these regions (see Figure 4 of Reichle, De Lannoy, Liu, Koster, et al., 2017).

Owing to a variety of hardware and software tools to detect and correct for RFI, SMAP generally provides Tb observations nearly everywhere on the globe (Piepmeier et al., 2014). For the Version 3 system, Tb scaling parameters were computed from 2 years of SMAP Tb observations in those locations where SMOS Tb observations are insufficient due to RFI, resulting in near-global assimilation coverage. Consequently, the Version 3 L4_SM product differs considerably from the Version 2 data in the regions of expanded SMAP assimilation coverage, but the differences between the Version 2 and Version 3 data are minimal elsewhere, including at all in situ measurement sites used for validation here and by Reichle, De Lannoy, Liu, Ardizzone, et al. (2017).

2.2.2. Changes From Version 3 to Version 4

Additional and more pervasive changes to the L4_SM system were implemented with Version 4. The most consequential model change is the calibration of the Catchment model parameter (α) that governs the recharge of soil moisture from the model's root zone excess reservoir into the surface excess reservoir. (The "surface excess" and "root zone excess" are Catchment model prognostic variables that capture the deviation of the modeled soil moisture profile from equilibrium in the 0- to 5-cm surface and 0- to 100-cm root zone layers, respectively.) Specifically, the replenishment of soil moisture near the surface from below under nonequilibrium conditions was reduced substantially in the new model version. This change was motivated by Koster, Liu, et al. (2018), who calibrated α locally by maximizing the correlation with SMAP Level-2 soil moisture retrievals across the continental United States (CONUS). Whereas the original model of the Version 2 and 3 systems uses a spatially constant value of $\alpha = 1$, the calibrated α values of Koster,

Liu, et al. (2018) range from 0.001 to 1, with a mean (median) value of 0.18 (0.004) (their Figure 2). In contrast, the Version 4 L4_SM system uses a spatially constant value of $\alpha = 0.04$. This value is a compromise between high correlation and low ubRMSE of model soil moisture against in situ measurements from the Soil Climate Analysis Network (Schaefer et al., 2007) and the U.S. Climate Reference Network (Bell et al., 2013; Diamond et al., 2013) across CONUS (not shown). Despite their more limited spatial coverage, in situ measurements are used here for model calibration because SMAP data are ultimately assimilated in the L4_SM algorithm. Note that the in situ measurements used for model calibration are distinct from those used for L4_SM validation (section 3.1).

Two important changes in the surface meteorological forcing data were made in Version 4 of the L4_SM system. First, the forcing for the pre-SMAP period in Version 4 is based on the Modern-Era Retrospective Analysis for Research and Applications (MERRA)-2 reanalysis (Gelaro et al., 2017), which enabled the use of a longer period (1980–2014) for L4_SM system spin-up (initialization). Second, and more importantly, the model background precipitation in the Version 4 L4_SM system is rescaled prior to applying the CPCU-based corrections so that its climatology also matches that of the GPCPv2.2 data. In earlier L4_SM versions, only the CPCU data were rescaled to the GPCPv2.2 climatology. Consequently, the precipitation climatology in Version 4 has changed substantially in Africa and the high latitudes, where the CPCU data are not used due to poor gauge coverage (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017).

An important change in the Tb analysis of the Version 4 L4_SM system is the removal from the EnKF state vector of the "catchment deficit" variable, the Catchment model prognostic variable related to the equilibrium soil moisture profile. This change was motivated by the results of Girotto et al. (2019), who similarly eliminated this variable from the EnKF state vector when assimilating SMOS Tb observations into the Catchment model and found good performance for surface and root zone soil moisture estimation without the skill degradation versus groundwater measurements that occurred when the catchment deficit was included in the EnKF state.

The calibration of the Version 4 L4_SM system used longer Tb observation records compared to earlier versions. Specifically, the Tb scaling parameters in Version 4 are computed from 8 years of SMOS (version 620; Kerr et al., 2016) observations (July 2010 to June 2018) and 3 years of SMAP L1C_TB observations (Version 4; Chan, 2018; April 2015 to March 2018) in locations where the SMOS climatology is unavailable due to RFI. The Version 4 SMAP Tb observations over land are warmer by ~3 K on average compared to earlier SMAP versions (Peng et al., 2017, 2018), and neither matches the version 620 SMOS Tb climatology. To address these calibration differences, Tb scaling parameters derived from SMOS observations are further corrected, separately for each grid cell, using the 3-year (April 2015 to March 2018) mean differences between SMOS and SMAP Tb observations.

Additional changes in Version 4 (Reichle, Liu, et al., 2018) include revised values for the surface soil heat capacity, a revised formulation to compute the surface aerodynamic roughness as a function of leaf area index, a change in the model parameter that governs the model's snow accumulation and depletion curve (Reichle, Draper, et al., 2017), and more recent and improved ancillary data sets (Mahanama et al., 2015), including for topography (Slater et al., 2006; Verdin et al., 2007), vegetation height (Simard et al., 2011), and land cover (Bontemps et al., 2011). The latter change resulted in a slightly revised land mask in the Version 4 product compared to earlier versions, notably a better representation of lake areas in East Africa (not shown). Each of these changes should improve some aspect of the L4_SM algorithm, but examining their individual impact is beyond the scope of the present paper.

2.3. L4_SM Data and Processing

SMAP science data collection began on 31 March 2015 at ~16:28 UTC. Throughout this paper, we use the 3-year period from 1 April 2015 to 31 March 2018 for validation and analysis, except for the runoff validation, which is restricted to Northern Hemisphere warm-season months (June–September of 2015–2017) to avoid periods when runoff is dominated by snowmelt. Because production of Version 3 data ended with Data Day 4 June 2018, the validation period is the longest period consisting of full years (or complete warm-season months) that also allows us to compare the skill and characteristics of the Version 4 product directly against those of the Version 3 data.

In the present paper, we use 3-hourly instantaneous surface and root zone soil moisture forecast, analysis, and uncertainty data from the L4_SM "analysis-update" files for Version 3 (Science Version ID Vv3030; Reichle et al., 2017a) and Version 4 (Science Version ID Vv4030; Reichle et al., 2018a). Furthermore, we use 3-hourly time-average total runoff data from the L4_SM "geophysical" files for Vv3030 (Reichle et al., 2017b) and Vv4030 (Reichle et al., 2018b). Besides these L4_SM data, we use soil moisture and runoff output from additional simulations with the L4_SM system in its model-only "Nature Run" configuration, that is, single-member model simulations without ensemble perturbations and without the assimilation of SMAP Tb observations (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017). Since the Nature Run simulations are also used for Catchment model development beyond the SMAP L4_SM system, they follow a different versioning system. Specifically, Nature Run versions NRv4.1 and NRv7.2 correspond to the Catchment model versions and forcing data used in Versions 3 and 4 of the L4_SM algorithm, respectively. To make the discussion below easier to follow, we hereinafter refer to NRv4.1 as NR3030 and to NRv7.2 as NR4030.

3. Validation Data and Approach

3.1. Soil Moisture In Situ Measurements and Validation Metrics

The primary validation reference for L4_SM soil moisture estimates are in situ measurements from SMAP core validation sites (or watersheds; Colliander, Berg, et al., 2017, Colliander, Jackson, et al., 2017). These measurements serve as a fully independent validation data set; again, they were not used in the calibration of the Catchment model parameter α (section 2.2.2). Each core site has a locally dense network of sensor profiles, thereby mitigating the typical upscaling errors encountered in the validation of model output (Crow et al., 2012), which here comprises the L4_SM estimates representing the average soil moisture conditions in the 81-km^2 area of an EASEv2 grid cell. Specifically, in the present paper we use surface (root zone) measurements from 18 (6) different core sites located in a variety of climate regimes and land cover conditions (Table 1). Across these sites, we use surface (root zone) measurements from a total of 31 (12) so-called "reference pixels" (or grid cells) at the 9-km scale. For additional characteristics see supporting information Table S1. Reichle, De Lannoy, Liu, Ardizzone, et al. (2017; their section 3) provide further discussion of processing details. See Chan et al. (2018) for details on the definition of coarser, 33-km reference pixels for which results are referenced below.

In addition to the ubRMSE, correlation (R), and bias metrics computed in Reichle, De Lannoy, Liu, Ardizzone, et al. (2017), here we also compute the *anomaly R* metric. For this, monthly climatological values of in situ soil moisture at a given reference pixel are computed. For a given calendar month, we required at least 80 (3-hourly) measurements in each of the three calendar months falling within the 3-year validation period. Anomaly time series are then computed by subtracting the corresponding monthly climatological seasonal cycle from the raw data. The procedure is repeated for the L4_SM data (after cross-masking against the available in situ measurements to ensure compatible temporal support of the estimated climatology). The anomaly R metric is the correlation coefficient of the L4_SM and in situ anomaly time series. Averages of metrics are computed by first averaging the metrics of all references pixels at a given site and then averaging across all sites.

Statistical uncertainty in the ubRMSE, R, and anomaly R metrics is estimated using 95% confidence intervals, which are computed at each site based on the number of samples in the time series (with a correction for temporal autocorrelation). We refer to changes in metrics as statistically significant whenever the confidence intervals do not overlap. It is important to keep in mind that the confidence intervals are themselves uncertain and only provide practical guidance as to whether the skill differences are meaningful. Statistical uncertainty estimates for the bias are not provided because the upscaling error is considerably larger for the bias than for the second-order metrics (Chen et al., 2019).

As in Reichle, De Lannoy, Liu, Ardizzone, et al. (2017) and Reichle, Liu, et al. (2018), we assume that the upscaled core site observations have much smaller errors than corresponding L4_SM estimates. The bias, ubRMSE, and correlation metrics obtained from directly comparing L4_SM estimates and in situ measurements are thus assumed to be close to the corresponding metrics against the (unknown) true soil moisture. Since it is reasonable to assume that the errors that do exist in the in situ measurements are not correlated with errors in the L4_SM data, measurement uncertainty in the in situ data leads to some overestimation

Table 1

Soil Moisture Core Validation Sites Used in the Present Paper

Site name	Country	Climate regime	Land cover	Number of 9-km reference pixels		Reference
				Surface soil moisture	Root zone soil moisture	
REMEDHUS	Spain	Temperate	Croplands	2	0	Sanchez et al. (2012); González-Zamora et al. (2015)
Reynolds Creek	USA (Idaho)	Arid	Grasslands	2	0	Seyfried et al. (2001)
Yanco	Australia (New South Wales)	Arid	Cropland/natural mosaic	2	0	Panciera et al. (2014)
Carman	Canada (Manitoba)	Cold	Croplands	1	0	McNairn et al. (2015)
Walnut Gulch	USA (Arizona)	Arid	Shrub open	3	0	Keefer et al. (2008)
Little Washita	USA (Oklahoma)	Temperate	Grasslands	3	2	Cosh et al. (2006)
Fort Cobb	USA (Oklahoma)	Temperate	Grasslands	2	2	Cosh et al. (2014)
Little River	USA (Georgia)	Temperate	Cropland/natural mosaic	2	2	Bosch et al. (2007)
Saint Josephs	USA (Indiana)	Temperate	Croplands	1	0	Heathman et al. (2012)
South Fork	USA (Iowa)	Cold	Croplands	3	3	Coopersmith et al. (2015)
Monte Buey	Argentina	Temperate	Croplands	1	0	n/a
Tonzi Ranch	USA (California)	Temperate	Savannas woody	1	0	Clewley et al. (2017); Moghaddam et al. (2016)
Kenaston	Canada (Saskatchewan)	Cold	Croplands	2	1	Rowlandson et al. (2015); Tetlock et al. (2019)
Valencia	Spain	Cold	Savannas woody	1	0	Juglea et al. (2010); Khodayar et al. (2019)
Niger	Niger	Arid	Grassland	1	0	Galle et al. (2018)
Benin	Benin	Tropical	Savannas	1	0	Galle et al. (2018)
TxSON	USA (Texas)	Temperate	Grasslands	2	2	Caldwell et al. (2018)
HOBE	Denmark	Temperate	Croplands	1	0	Bircher et al. (2012); Jensen and Refsgaard (2018)

Note. Further details are provided in the supporting information (Table S1).

of the true ubRMSE and some underestimation of the true correlation (Chen et al., 2017, 2019). Consequently, the skill metrics presented here are conservative estimates of the actual skill.

3.2. Streamflow Measurements

For the validation of the L4_SM runoff estimates (section 4.3) we use streamflow measurements from 239 unregulated hydrological basins of intermediate size (2,000–10,000 km^2) within CONUS. These measurements are published by the U.S. Geological Survey (USGS) and were previously used to assess the skill of model and data assimilation output (Koster, Liu, et al., 2018; Kumar et al., 2014). The streamflow measurements were normalized by basin area to convert their units into millimeters per day. To account for the lack of runoff routing in the L4_SM system, the USGS streamflow measurements were smoothed using a 10-day running mean and directly compared to similarly smoothed runoff estimates from L4_SM and Nature Run data. The model estimates were spatially aggregated with consideration of the fraction of the irregular basin area contained within each 9-km EASEv2 grid cell. As mentioned above, only the warm season (June–September) is included in the runoff validation.

3.3. Data Assimilation Diagnostics

By routinely confronting model estimates with observations, a data assimilation system generates valuable internal diagnostics that can be used to assess its proper functioning. In the present paper, we investigate the statistics of the O-F Tb residuals, the observation-minus-analysis (O-A) Tb residuals, and the soil moisture analysis increments. We provide only a brief narrative summary of these diagnostics here and refer the reader to section 2c of Reichle, De Lannoy, Liu, Koster, et al. (2017) for a more detailed description and formal definitions. Because the L4_SM algorithm assimilates Tb observations, the O-F and O-A diagnostics are
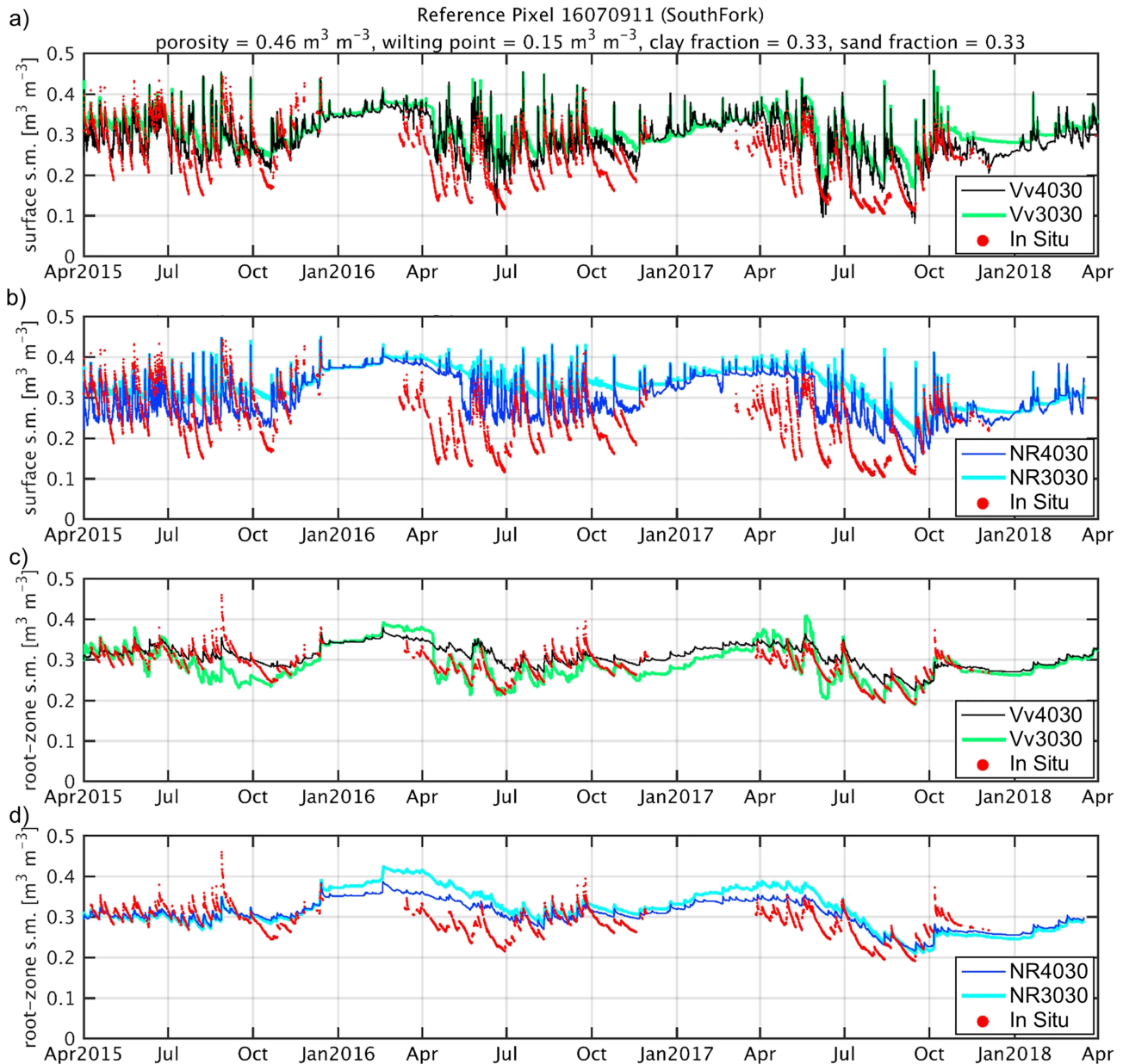
Figure 1. Surface soil moisture at the 9-km South Fork, Iowa, reference pixel #16070911 from (a) L4_SM estimates (Vv4030 and Vv3030) and (b) Nature Run simulations (NR4030 and NR3030). In situ measurements are shown identically in both (a) and (b) as red dots. (c, d) As in (a) and (b), respectively, but for root zone soil moisture. L4_SM = Level-4 Soil Moisture.

in terms of Tb at the 36-km resolution of the assimilated observations, while the analysis increments are at the 9-km resolution of the Catchment model soil moisture. Note that the observations used to compute the O-F and O-A residuals are the climatologically rescaled Tb observations (section 2.1). It is also important to keep in mind that the O-F Tb residuals compare the model forecast Tb with the corresponding Tb observation *before* the observation is assimilated.

In a well-calibrated assimilation system, the long-term mean of the O-F residuals and increments at any given location is 0, indicating that there is no long-term net addition (or subtraction) of water or energy by the analysis. The time series standard deviation of the O-F residuals is a measure of the typical misfit between the model forecast Tb and the (rescaled) SMAP observations. Similarly, the time series standard deviation of the increments is a measure of the typical adjustment of the model forecast soil moisture at any given
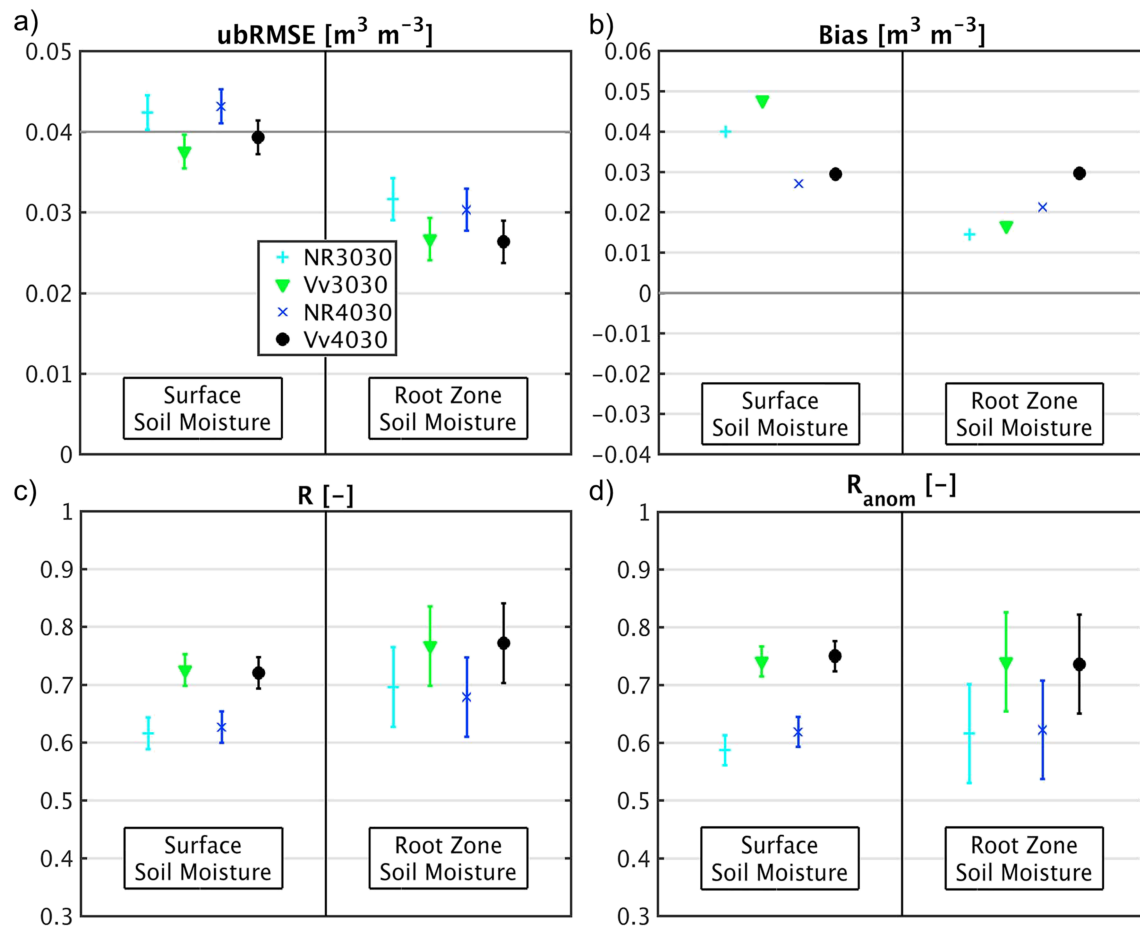
Figure 2. (a) ubRMSE (m^3/m^3), (b) bias (m^3/m^3), (c) R (dimensionless), and (d) anomaly R (dimensionless) for Nature Run and L4_SM estimates of surface and root zone soil moisture versus core validation site measurements, averaged across all 9-km reference pixels. Error bars for ubRMSE, R, and anomaly R indicate 95% confidence intervals (section 3.1). The thick horizontal line in (a) represents the L4_SM accuracy requirement of ubRMSE ≤ 0.04 m^3/m^3. ubRMSE = unbiased root-mean-square error; L4_SM = Level-4 Soil Moisture.

analysis time. A particularly valuable metric is the standard deviation of the O-F residuals after normalization with their simulated (expected) error standard deviation, which includes the prescribed observation error and the model forecast error resulting from the prescribed perturbation characteristics. This metric should be unity. If the metric exceeds unity, the prescribed errors underestimate the actual errors reflected in the O-F residuals. Conversely, if the standard deviation of the normalized O-F residuals is less than unity, the prescribed errors overestimate the actual errors. Finally, all second-order assimilation diagnostics are spatially averaged in quadrature.

4. Validation Versus In Situ Measurements

4.1. Soil Moisture

4.1.1. Performance at South Fork, Iowa

In this subsection we first illustrate the soil moisture skill for Version 4 relative to Version 3 at South Fork, Iowa, before turning to the average performance across all sites in section 4.1.2. The results at South Fork—particularly the changes seen in going from Version 3 to Version 4—are representative of typical results across all sites. Note that at some core validation sites (e.g., Little River, Georgia), the revised model resulted in greater differences (Balsamo et al., 2018, their Figure 8a; Reichle, Liu, et al., 2018, their Figure 6), whereas at other sites (e.g., Little Washita, Oklahoma), the model changes had less impact on the skill of the soil moisture estimates (Reichle, Liu, et al., 2018, their Figure 5). For reference, supporting information Tables S2–S5 provide a complete listing of the performance metrics at all sites.
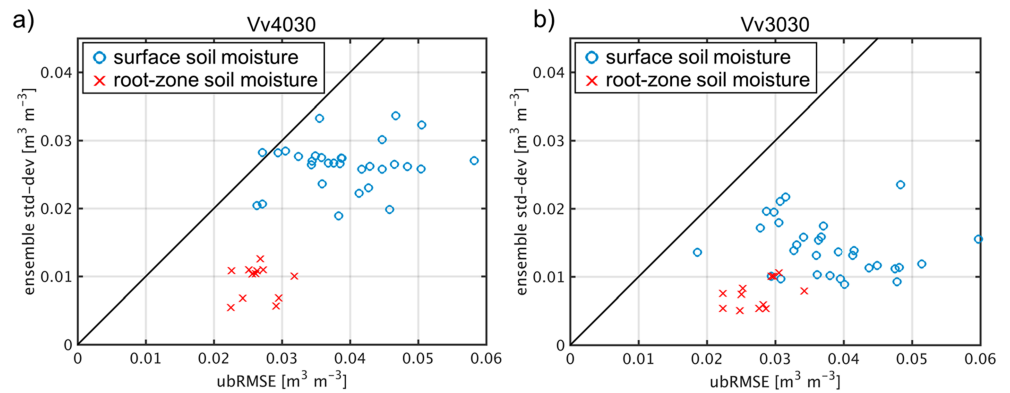
Figure 3. L4_SM estimates of soil moisture uncertainty (ordinate) at 9-km reference pixels versus actual ubRMSE (abscissa) for (a) Vv4030 and (b) Vv3030. Blue (red) circles indicate surface (root zone) soil moisture. The 1:1 line is shown in solid black. L4_SM = Level-4 Soil Moisture; ubRMSE = unbiased root-mean-square error.

Reichle, De Lannoy, Liu, Ardizzone, et al. (2017) provide a detailed description of the South Fork site, which is in a cold climate agricultural region dominated by summer crops of corn and soybeans, with mostly bare fields or stubble in winter. Owing to the high clay content of the soils in this region, agricultural fields are increasingly equipped with tiles to improve drainage, a feature that is not captured in the Catchment model. There are three 9-km reference pixels at South Fork (Table 1), all showing similar relative performance of the Version 3 and 4 Nature Run and L4_SM data (supporting information Tables S2–S5). Here, we discuss pixel #16070911 because it has the most complete time series of measurements. Figure 1a shows surface soil moisture for Vv4030 (thin black lines), Vv3030 (thick light green lines), and the in situ measurements (red dots). Overall, the estimates from both L4_SM product versions track the in situ measurements reasonably well. The bias in surface soil moisture is 0.034 m^3/m^3 for Vv4030, down from 0.053 m^3/m^3 for Vv3030. The respective Nature Run time series, shown in Figure 1b, have bias values of 0.042 m^3/m^3 (NR4030; thin blue lines) and 0.082 m^3/m^3 (NR3030; thick cyan lines). The reduction in the surface soil moisture bias from the Version 3 to the Version 4 system stems primarily from the model change that reduced the upward recharge from the root zone to the surface excess reservoir (section 2.2.2), which results in generally drier (and thus less biased) surface soil moisture. Another consequence of this model change is the increased dynamic range of surface soil moisture in the Version 4 system compared to that of Version 3, which better reflects the wetting and dry-down characteristics of the in situ measurements (Figures 1a and 1b). However, the ubRMSE values for Vv4030 and Vv3030 are the same at ~0.05 m^3/m^3. Similarly, the R values (anomaly R values) are essentially unchanged at 0.70 (0.72) for Vv4030 and 0.69 (0.75) for Vv3030, although it should be noted that the R value of the model-only Nature Run estimates increased significantly from 0.29 for NR3030 to 0.45 for NR4030. In summary, the primary impact of the model changes on the surface soil moisture skill of the L4_SM product is the reduction in bias.

For root zone soil moisture (Figure 1c), Vv4030 exhibits a bias of 0.020 m^3/m^3, whereas Vv3030 was essentially unbiased (−0.003 m^3/m^3). The ubRMSE, however, is better for Vv4030 (0.025 m^3/m^3) than Vv3030 (0.031 m^3/m^3), albeit not significantly. This ubRMSE improvement in the new version is also mirrored in the Nature Run results (Figure 1d), with a root zone soil moisture ubRMSE of 0.032 m^3/m^3 for NR4030 compared to 0.044 m^3/m^3 for NR3030. The smaller ubRMSE values in the Version 4 system are most likely a reflection of the reduced dynamic range in root zone soil moisture caused by the reduction in the upward recharge in the Catchment model. While the root zone soil moisture R value is slightly worse for Vv4030 (0.72) than Vv3030 (0.74), the anomaly R value is better (0.82 versus 0.76).

A comparison of Figures 1a and 1b illustrates that the assimilation of SMAP Tb observations improves the surface soil moisture skill considerably in both L4_SM versions. For example, the ubRMSE improves from 0.066 m^3/m^3 for NR4030 to 0.050 m^3/m^3 for Vv4030. Moreover, the R value (anomaly R value) improves from 0.45 (0.41) for NR4030 to 0.70 (0.72) for Vv4030. The assimilation of SMAP Tb observations similarly improves the ubRMSE and R skill for root zone soil moisture at the South Fork site (Figures 1c and 1d).
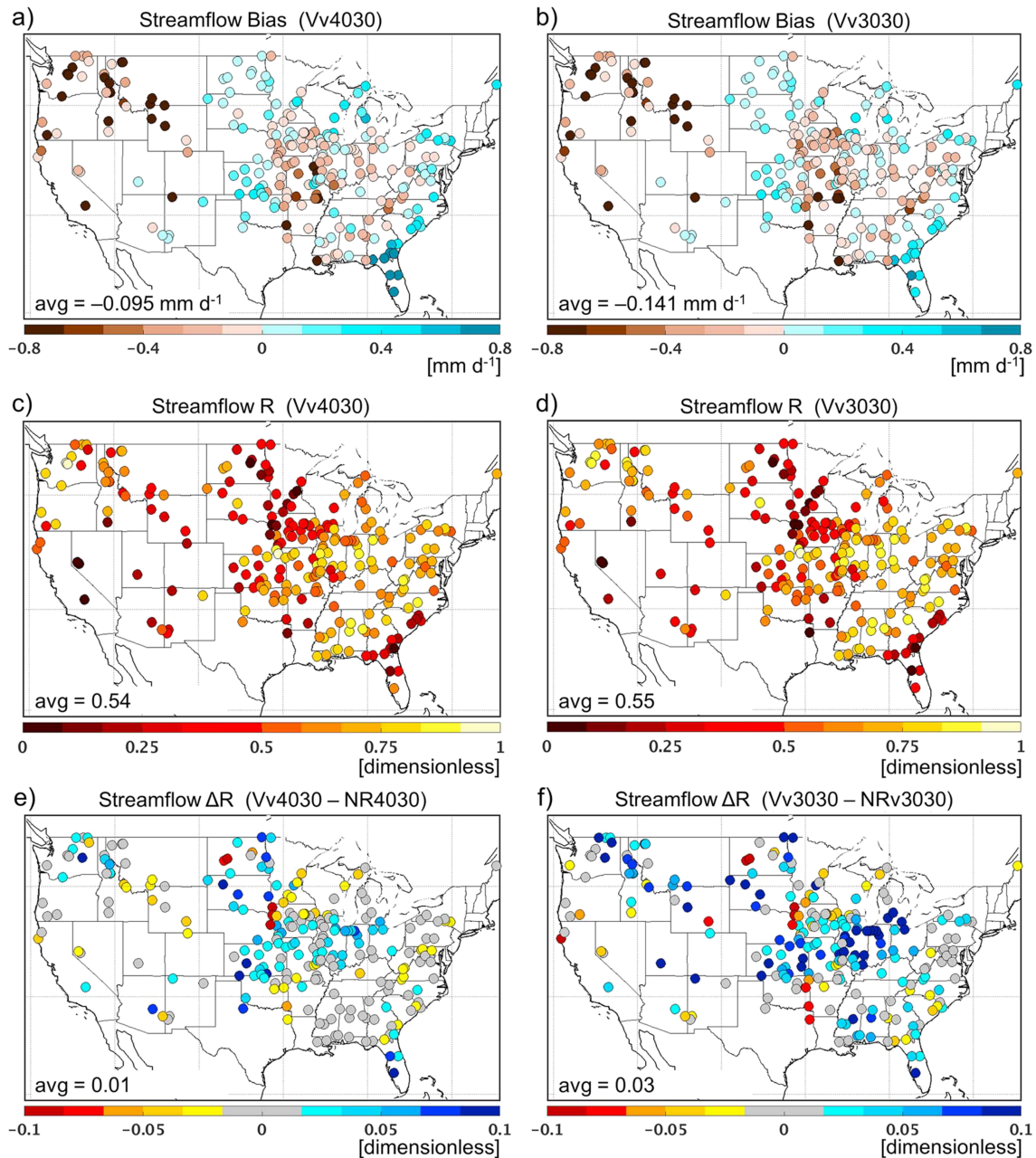
Figure 4. June–September 2015–2017 runoff skill versus U.S. Geological Survey streamflow measurements in 239 U.S. basins for (a, c, e) Vv4030 and (b, d, f) Vv3030. (a, b) Bias and (c, d) *R* values of the L4_SM product. (e, f) Difference in *R* values between the L4_SM product and the corresponding Nature Run data. Blue (red) colors in (e) and (f) indicate that the L4_SM estimates are better (worse) than Nature Run estimates. L4_SM = Level-4 Soil Moisture.

4.1.2. Average Performance Across All Reference Pixels

Figure 2 summarizes the average performance metrics across all 9-km reference pixels from the 18 core validation sites (Table 1). Perhaps the most important validation result is that the Vv4030 surface and root zone soil moisture data meet the L4_SM accuracy requirement (ubRMSE $\leq$ 0.04 m^3/m^3) established prior to the launch of SMAP (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017; their section 3a). This result is now based on three full years of core site measurements, compared to just 2 years as used by Reichle, De Lannoy, Liu, Ardizzone, et al. (2017). Specifically, as shown by the black circles in Figure 2a, the ubRMSE is 0.039 m^3/m^3 for Vv4030 surface soil moisture and 0.026 m^3/m^3 for Vv4030 root zone soil moisture.
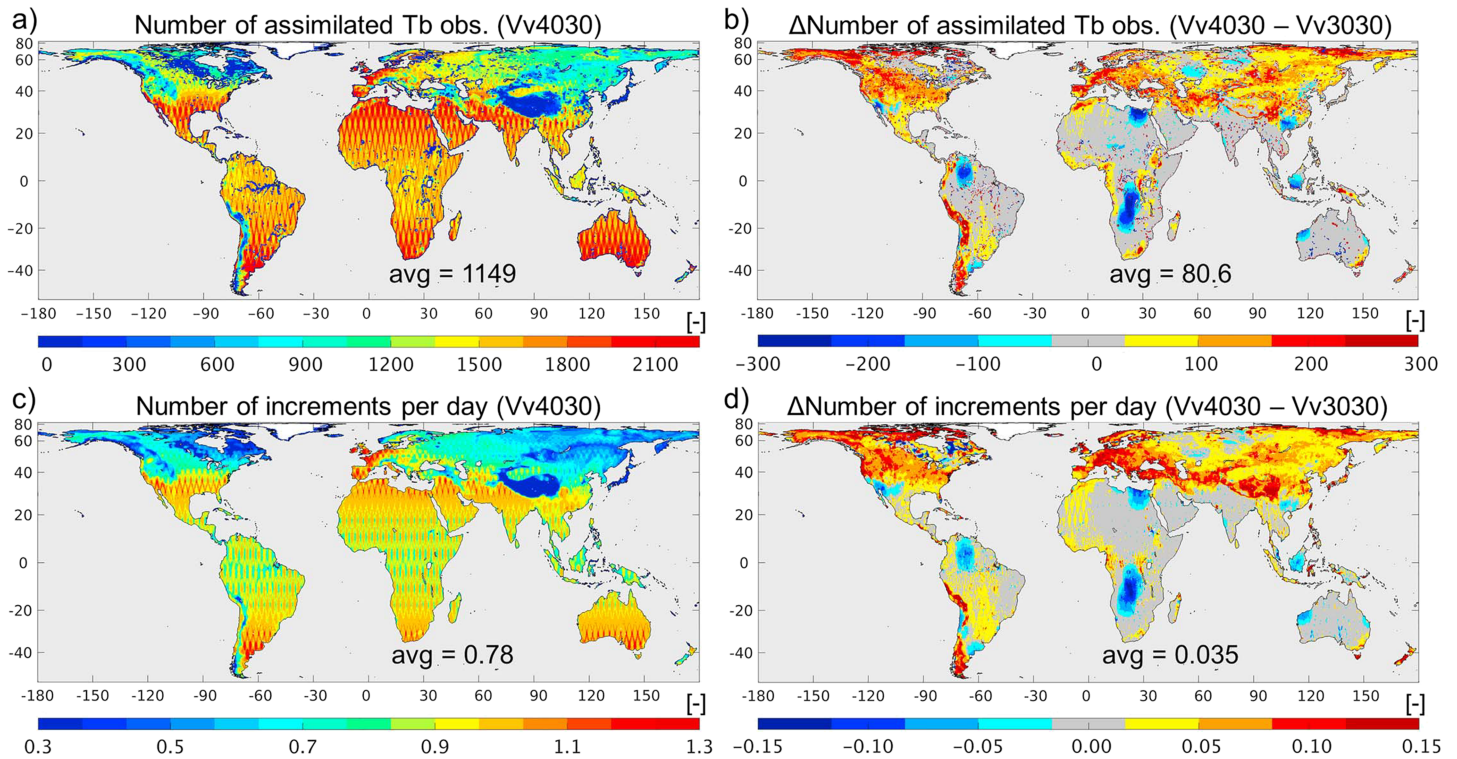
Figure 5. (a) Number of Soil Moisture Active Passive Tb observations assimilated into Vv4030 from April 2015 to March 2018. Data counts include horizontal- and vertical-polarization data from ascending and descending half-orbits. (b) Difference in number of assimilated observations (Vv4030 minus Vv3030). (c) Average number of increments per day generated in Vv4030 from April 2015 to March 2018 (applies equally to surface and root zone soil moisture increments). (d) Difference in number of increments (Vv4030 minus Vv3030). Global averages computed over all land excluding glaciated and open water surfaces. Tb = brightness temperature.

There is a slight (but not statistically significant) increase in the surface soil moisture ubRMSE from 0.038 m^3/m^3 for Vv3030 (light green triangles) to 0.039 m^3/m^3 for Vv4030 (Figure 2a). A similar increase is seen in the underlying modeling systems from 0.042 m^3/m^3 for NR3030 (cyan plus signs) to 0.043 m^3/m^3 for NR4030 (blue crosses; Figure 2a). As illustrated in the South Fork time series (Figure 1), the slight degradation in skill is related to the generally larger dynamic range in surface soil moisture in the Version 4 modeling system following the reduction in the upward recharge of the surface excess reservoir (section 2.2.2). In fact, the surface soil moisture time series standard deviation averaged across all reference pixels, which is 0.053 m^3/m^3 for the in situ measurements, improves from 0.048 m^3/m^3 for Vv3030 to 0.055 m^3/m^3 for Vv4030 (not shown). The same model change results in a reduction of the dynamic range of root zone soil moisture in the Version 4 system, causing the slight decrease in root zone soil moisture ubRMSE from 0.032 m^3/m^3 for NR3030 to 0.030 m^3/m^3 for NR4030, and from 0.027 m^3/m^3 for Vv3030 to 0.026 m^3/m^3 for Vv4030 (Figure 2a). Very minor differences (likewise not statistically significant) between the two versions are also seen for the correlation metrics (Figures 2c and 2d).

Among the metrics considered here, the bias exhibits the largest change between the two product versions. Specifically, the average bias in surface soil moisture decreased considerably from 0.048 m^3/m^3 for Vv3030 to 0.029 m^3/m^3 for Vv4030, while the average bias in root zone soil moisture increased from 0.017 m^3/m^3 for Vv3030 to 0.030 m^3/m^3 for Vv4030 (Figure 2b). A different perspective is gained by looking at the average *absolute* bias, which decreased from 0.052 m^3/m^3 in Vv3030 to 0.047 m^3/m^3 in Vv4030 for surface soil moisture and remained unchanged at 0.048 m^3/m^3 for root zone soil moisture (not shown). These relative changes in bias are mirrored in the underlying model estimates (Figure 2b), suggesting again that the changes in the bias of the L4_SM product are driven by the changes in the Catchment model, and more specifically, by the change in the upward recharge into the surface excess reservoir (section 2.2.2).
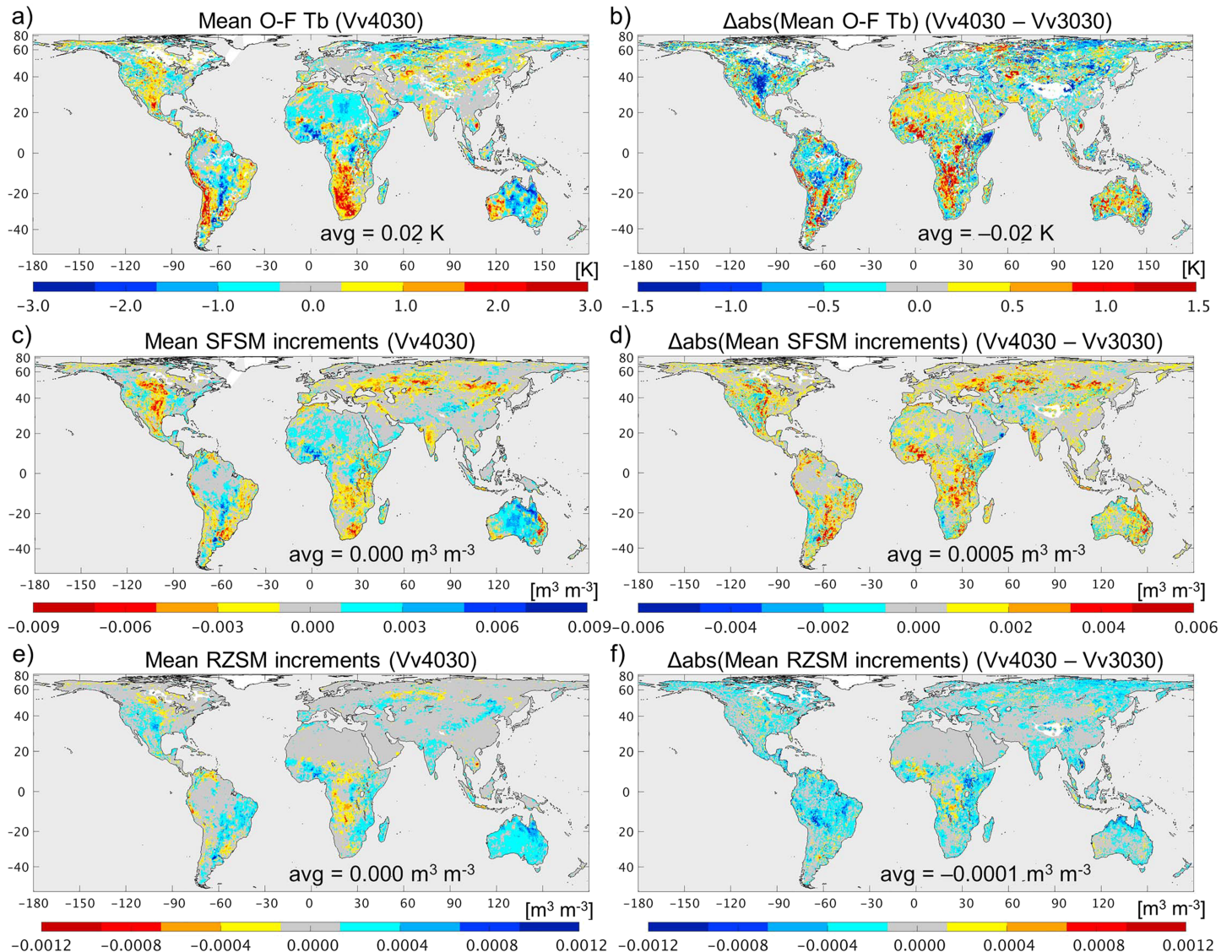
Figure 6. (a, c, e) Mean of Vv4030 assimilation diagnostics. (b, d, f) "Closeness" of Vv4030 and Vv3030 mean assimilation diagnostics to the ideal value of 0, where "closeness" is measured as the Vv4030 minus Vv3030 difference in the absolute value of the mean assimilation diagnostic. Blue (red) colors in (b), (d), and (f) indicate that the Vv4030 (Vv3030) mean diagnostic is better. Assimilation diagnostics include (a, b) O-F Tb residuals, (c, d) surface soil moisture (SFSM) increments, and (e, f) root zone soil moisture (RZSM) increments for April 2015 to March 2018. Regions with insufficient Tb assimilation coverage for the computation of statistics are shown in white and excluded from the global average calculation. O-F = observation-minus-forecast; Tb = brightness temperature.

Three issues complicate the interpretation of the bias estimates. First, the upscaling uncertainty in the bias estimates is much larger than that in the ubRMSE and correlation metrics (Chen et al., 2019), which makes it difficult to determine the significance of the bias changes between Versions 3 and 4. Second, estimates of model (Nature Run or L4_SM) bias are more sensitive than ubRMSE and correlation metrics to the vertical placement of the sensors. Specifically, the 0–5 cm modeled surface soil moisture is validated using sensors typically placed at 5 cm depth, and the 0–100 cm modeled root zone soil moisture is validated using sensors placed within 50 cm from the surface (see supporting information Table S1 for the depth of the deepest sensor used at each reference pixel; for further details see Reichle, De Lannoy, Liu, Ardizzone, et al., 2017). Considering this mismatch in vertical support, perfect model estimates of surface (root zone) soil moisture should have a dry (wet) diagnosed bias versus the in situ measurements. For surface soil moisture, the reduction in the diagnosed wet bias from Version 3 to Version 4 is thus definitely an improvement. For root zone soil moisture, the increase in the diagnosed wet bias does not necessarily imply a degradation. Finally, the
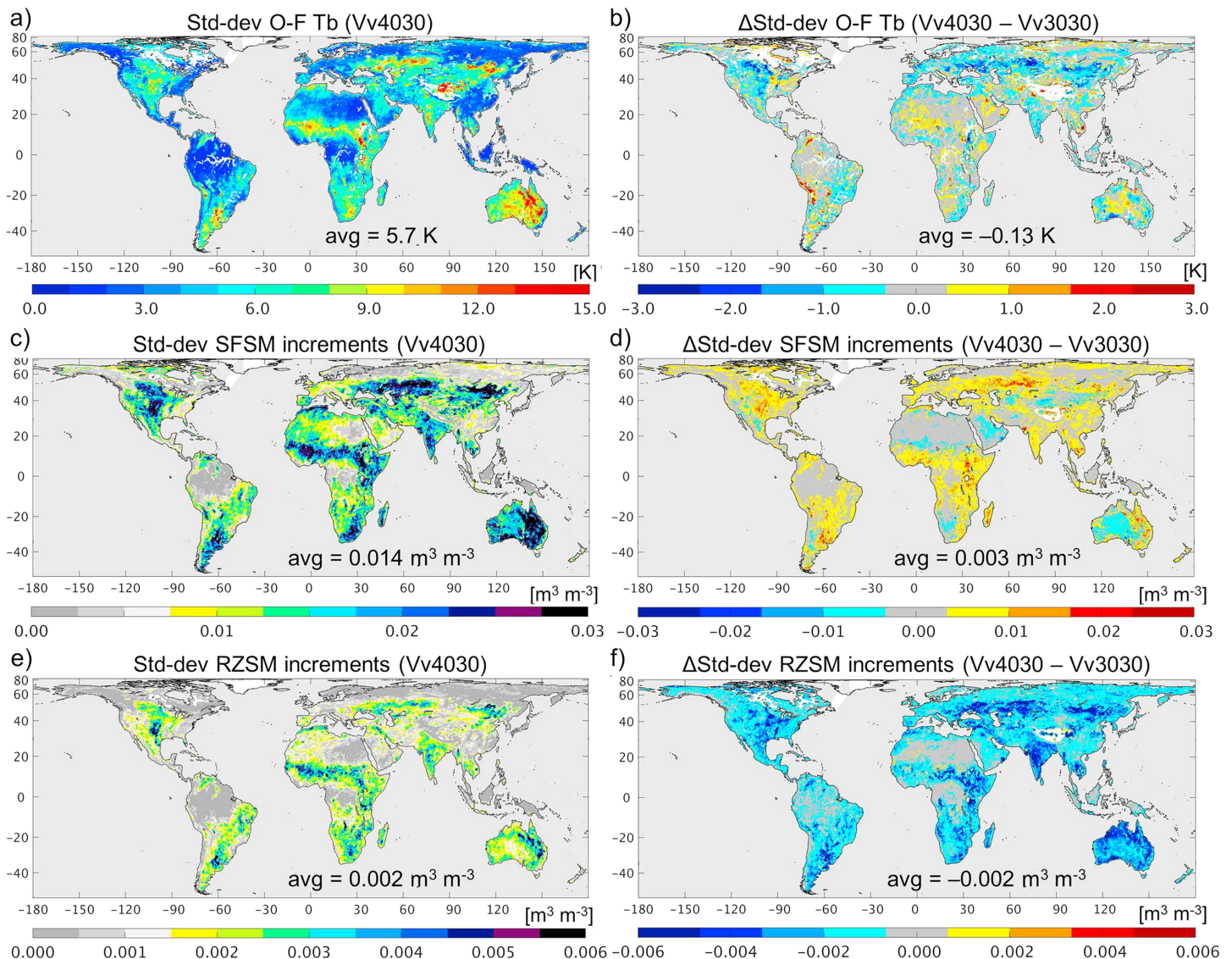
Figure 7. (a, c, e) Standard deviation of Vv4030 assimilation diagnostics and (b, d, f) difference in standard deviation of assimilation diagnostics (Vv4030 minus Vv3030). Assimilation diagnostics include (a, b) O-F Tb residuals, (c, d) surface soil moisture (SFSM) increments, and (e, f) root zone soil moisture (RZSM) increments for April 2015 to March 2018. Regions with insufficient Tb assimilation coverage for the computation of statistics are shown in white and excluded from the global average calculation. O-F = observation-minus-forecast; Tb = brightness temperature.

assumption of a constant Tb sensing depth in the L4_SM analysis likely causes bias in the L4_SM soil moisture estimates. Estimating the actual sensing depth and using this information in soil moisture assimilation systems are subjects of ongoing research (Lv et al., 2018, 2019).

Figure 2 further demonstrates that the Version 4 L4_SM estimates are better than estimates from the corresponding (model-only) Nature Run simulation, which again highlights the beneficial impact of the SMAP Tb data. For example, the surface soil moisture ubRMSE is 0.039 m^3/m^3 for Vv4030, down from 0.043 m^3/m^3 for NR4030, and the root zone soil moisture ubRMSE is 0.026 m^3/m^3 for Vv4030, down from 0.030 m^3/m^3 for NR3030 (Figure 2a). These ubRMSE improvements, however, are not statistically significant.

Statistically significant improvements from assimilating SMAP data are seen for the surface soil moisture R (Figure 2c) and anomaly R (Figure 2d) values, which improve from 0.63 and 0.62, respectively, for NR4030 to 0.72 and 0.75, respectively, for Vv4030. The R and anomaly R metrics for Vv4030 root zone soil moisture are increased by similar margins over NR4030. However, the root zone improvements are not statistically
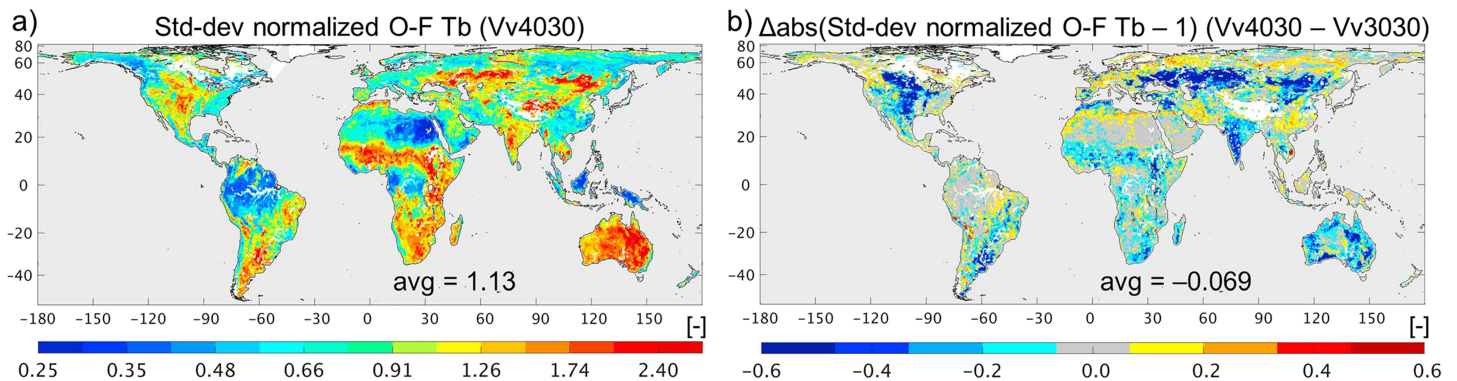
Figure 8. (a) Standard deviation of normalized O-F Tb residuals (ν) for Vv4030 from April 2015 to March 2018. (b) Difference in the absolute departure of ν from unity (Vv4030 minus Vv3030). Blue (red) shading in (b) indicates that ν is closer to (farther away from) unity in Vv4030 than Vv3030; that is, blue (red) shading indicates that Vv4030 (Vv3030) is better. Regions with insufficient Tb assimilation coverage for the computation of statistics are shown in white and excluded from the global average calculation. O-F = observation-minus-forecast; Tb = brightness temperature.

significant, primarily because there are fewer reference pixels with root zone measurements and because root zone soil moisture is more strongly autocorrelated, resulting in larger confidence intervals (section 3.1). The average bias for surface soil moisture is only slightly worse for Vv4030 than for NR4030 (Figure 2b). In comparison, the root zone bias increased from 0.021 m^3/m^3 for NR4030 to 0.030 m^3/m^3 for Vv4030 (Figure 2b). As noted above, this does not necessarily imply degraded performance. Overall, the impact of SMAP data assimilation on the skill versus in situ measurements is very similar in the Version 3 and 4 systems (Figure 2).

Finally, the skill of the published Vv4030 data (Figure 2) closely matches that of the preliminary Version 4 data assessed by Reichle, Liu, et al. (2018; their Figures 8 and 9). This is true across all metrics investigated by Reichle, Liu, et al. (2018), who also assessed the skill of surface soil temperature data, metrics computed at the 33-km reference pixel scale, and the performance versus measurements from more than 400 "sparse" network sites (not shown). We therefore refer the reader to Reichle, Liu, et al. (2018) for additional Version 4 validation results. Taken together, the above results suggest that, apart from the changes in bias, there is no net improvement or degradation in the performance versus in situ measurements between the Version 3 (and thus, effectively, Version 2) and Version 4 soil moisture estimates.

4.2. Soil Moisture Uncertainty Estimates

The L4_SM surface and root zone soil moisture estimates are computed as the mean across the 24-member ensemble of Catchment model replicates (section 2). The L4_SM product also provides the ensemble standard deviation (or ensemble spread) of the analyzed surface and root zone soil moisture at the same 3-hourly, 9-km resolution. This ensemble spread is a measure of the uncertainty in the L4_SM soil moisture estimates. In this section, we assess the quality of these uncertainty estimates by comparing them to the actual errors as computed versus in situ measurements. A related metric is the standard deviation of the normalized O-F Tb residuals, which is available (nearly) globally and will be examined in section 5.4.

Figure 3 shows scatter plots of the (time-average) uncertainty estimates at the 9-km core site reference pixels (Table 1) versus the corresponding (time series) ubRMSE for surface and root zone soil moisture. In a perfectly tuned system, the two metrics should match; that is, the symbols in the scatter plot should be close to the 1:1 line. In the L4_SM system, however, nearly all of the symbols fall well below the 1:1 line, indicating a severe underestimation of the actual errors in the L4_SM product. For the surface soil moisture uncertainty estimates, this bias is −0.013 m^3/m^3 for Vv4030 (Figure 3a), which is an improvement over the bias of −0.024 m^3/m^3 for Vv3030 (Figure 3b). For the root zone soil moisture uncertainty estimates, the bias is −0.017 m^3/m^3 for Vv4030 compared to −0.020 m^3/m^3 for Vv3030. Since the ubRMSE values are very similar for both product versions (section 4.1), the bias in the uncertainty estimates is reduced from Version 3 to Version 4 mostly because the ensemble spread increased considerably in Version 4. The increase in spread is particularly pronounced for surface soil moisture because the reduction in the modeled upward recharge (section 2.2.2) also suppresses the relaxation of the perturbed surface soil moisture to equilibrium conditions.

The overall magnitude of the actual errors is therefore better captured by the Vv4030 uncertainty estimates than by those from Vv3030. The R values for the scatter plots of Figure 3, however, are not significantly different from zero (not shown). This means that—for surface and root zone soil moisture in both product versions—the (time-average) uncertainty estimate for any given grid cell cannot usefully predict the actual error (i.e., the time series ubRMSE) at that grid cell relative to the error at other grid cells.

Finally, we also examined, separately for each 9-km reference pixel, the spread in the distribution of the actual instantaneous errors (computed versus the in situ measurements) as a function of the associated instantaneous uncertainty estimates (not shown). No relationship emerged, however, in either product version for either surface or root zone soil moisture. That is, for a given location, the time series of the uncertainty estimates do not contain useful information about the evolution of the typical magnitude of the actual errors.

Corresponding results for core site reference pixels at the 33-km scale and for sparse networks are qualitatively similar (not shown). It is encouraging that the revised Catchment model in the Version 4 system is better able to represent average actual errors in surface soil moisture. Some of the remaining low bias in the L4_SM uncertainty estimates may also be explained by the likely overestimation of the ubRMSE values due to errors in the in situ measurements themselves (section 3.1). Nevertheless, the above results and further evidence to follow in section 5 indicate that more work is needed to refine the error characterization in the L4_SM system.

4.3. Runoff

To date, the validation of the L4_SM product versus in situ measurements has focused on soil moisture and temperature (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017). Here, we investigate for the first time the warm season (June–September) skill of the L4_SM runoff estimates versus USGS streamflow measurements in 239 unregulated hydrological basins (section 3.2).

Key performance metrics for both product versions are shown in Figure 4. For Vv4030 runoff estimates, the bias versus observed streamflow is −0.095 mm/day (or −35 mm/year) on average across the 239 basins (Figure 4a). The underestimation is modest across much of the Midwest and more severe in many northwestern basins and in a few basins in Missouri and Arkansas. There are also, however, many basins in the Great Plains, along the East Coast, and particularly in Florida, where Vv4030 runoff overestimates the observed streamflow. Nevertheless, the mean runoff in Vv4030 is less biased than that of Vv3030, for which the bias is −0.141 mm/day (or −52 mm/year; Figure 4b).

The overall underestimation of streamflow in the L4_SM system is consistent with that in the MERRA-2 reanalysis (Reichle, Liu, et al., 2017), which also uses the Catchment model to describe the land surface water and energy balance. During the development of the revised Catchment model for Version 4 of the L4_SM system, we tested the strategy of Koster and Mahanama (2012) to improve the Catchment model's representation of the soil moisture-evapotranspiration and soil moisture-runoff relationships. While this indeed improved the model's mean runoff, these modifications also degraded the soil moisture skill versus in situ measurements (even with soil moisture assimilation) and were therefore deemed unsuitable for use in the L4_SM system.

The runoff correlation skill for Vv4030 is 0.54 (Figure 4c). The highest skill values are found in the far northwestern basins and in most of the central and eastern United States. The lowest skill values occur primarily in the western mountains, the north central United States, and along the coast of Georgia and the Carolinas. There are only minor differences in the runoff correlation skill between Vv4030 and Vv3030, for which the average skill is minimally higher at 0.55 (compare Figures 4c and 4d). While both L4_SM product versions have the same runoff correlation skill, the corresponding Nature Run skills are somewhat different. Specifically, the improvement in Vv4030 over the corresponding Nature Run simulation is minimal on average ($\Delta R = 0.01$), with larger changes seen only in a handful of basins along the Texas-Oklahoma border, along the Nebraska-Iowa border, and in Florida (Figure 4e). The average improvement in Vv3030 over NR3030 is slightly bigger ($\Delta R = 0.03$), with larger and predominantly positive changes seen in about two dozen basins in the central United States (Figure 4f). That is, the revised Catchment model better captures temporal variations in runoff than did the old model version, while the assimilation of SMAP Tb observations contributes less to the runoff skill in Vv4030 than in Vv3030, suggesting that the model revisions and

the Tb assimilation correct similar deficiencies. This is consistent with the fact that in the Version 4 system, the root zone soil moisture analysis increments are generally smaller than in Version 3 (section 5), owing to the removal of the catchment deficit variable from the EnKF state vector in Version 4 (section 2.2.2). The overall relatively small runoff improvements from data assimilation seen here are consistent with the findings of Mao et al. (2019) that soil moisture assimilation alone is not sufficient to substantially improve streamflow estimates.

5. Data Assimilation Diagnostics

5.1. Data Counts

The total number of SMAP L1C_TB observations that were assimilated into Vv4030 during the 3-year assessment period is shown in Figure 5a. This count includes all horizontal- and vertical-polarization observations from ascending and descending orbits. The global average count is 1,149 assimilated SMAP observations (per 36-km grid cell) during the 3-year period, or ~1 per day, across all land areas excluding permanently glaciated and permanent open water surfaces. Outside of the high latitudes, the figure exhibits the typical diamond pattern associated with SMAP's regular 8-day repeat orbit. There are gaps in coverage in the vicinity of lakes (such as in northern Canada) and along major rivers (e.g., the Amazon), where the open water area fraction often exceeds 5% and the observations are excluded from assimilation (Reichle, De Lannoy, Liu, Ardizzone, et al., 2017) because the L4_SM system does not simulate the contribution of open water to the Tb, so that the O-F Tb residuals in these locations are not suitable for use in the L4_SM analysis. Data counts are lower than average in the high latitudes and in high-elevation and mountainous areas (including the Rocky Mountains, the Andes, the Himalayas, and Tibet). In these regions, the L4_SM analysis is limited to a much shorter warm (unfrozen) season (section 2.1). In the high latitudes, the polar orbit of the SMAP spacecraft results in more frequent revisit times, thereby balancing somewhat the lower counts resulting from the shorter warm season there. Note that despite the gaps in assimilation coverage, the L4_SM product provides soil moisture estimates everywhere, even if in some regions the L4_SM estimates rely solely on the information in the land model and forcing data.

Perhaps most importantly, Figure 5a illustrates that SMAP observations were assimilated into Vv4030 throughout most of Eastern Europe, the Middle East, and East Asia, where L-band RFI is common (Oliva et al., 2012). This was not possible in Version 2 of the L4_SM algorithm (Reichle, De Lannoy, Liu, Koster, et al., 2017; their Figure 4). At the time of Version 2, only SMOS Tb observations were available to address the bias in the modeled Tb climatology, and SMOS data cannot be used for this purpose in the RFI-impacted regions (section 2.2.1).

Next, Figure 5b shows that more SMAP observations were assimilated in Version 4 than in Version 3 primarily in the high latitudes and in high-elevation regions (e.g., the Tibetan and Andean plateaus). This increase in coverage in Version 4 is the result of using longer SMOS and SMAP Tb records to compute the (seasonally varying) L-band climatology (section 2), which consequently has fewer gaps in spatial and temporal coverage and thus supports the assimilation of more SMAP Tb observations. Conversely, the count of assimilated SMAP observations in Version 3 exceeds that of Version 4 across several distinct, circular-shaped regions, including southern Colombia and Venezuela, Egypt, southern China, and portions of Zambia, Angola, Namibia, and Botswana (Figure 5b). These elevated data counts reflect an error in the L4_SM algorithm software through Version 3 and comprise ~2% of the global average count. In the affected locations, fore- and aft-looking Tb observations from the same half-orbit were inadvertently assimilated at two different L4_SM analysis times when the fore and aft observation times fell on opposite sides of the boundary between two subsequent 3-hr analysis windows. This error was corrected in the Version 4 algorithm software, which now always averages pairs of fore- and aft-looking Tb observations prior to assimilation. Consequently, the Version 4 data counts (Figure 5a) no longer show the spatial artifacts seen in the earlier versions (Reichle, De Lannoy, Liu, Koster, et al., 2017; their Figure 4). Finally, there are relatively large positive and negative differences in data counts in small, isolated patches across much of the globe (Figure 5b) because the improved land mask in Version 4 (section 2.2.2) changes the exact location and size of inland open water surfaces and glaciers and thereby impacts where Tb observations are assimilated. Furthermore, many isolated patches of higher data counts in Version 4, for example, across the southern United States, are again the result of better coverage of the L-band Tb climatology.

Figure 5c shows the average number of analysis increments that the L4_SM algorithm generated per day during the validation period. The global mean is 0.78 for Version 4, which means that for a given location, there are approximately four increments applied every 5 days on average, either from an ascending or a descending overpass. The overall spatial pattern of the increments count follows that of the count of the assimilated observations (Figure 5a), with the highest counts in the subtropics where overpasses are more frequent than in the tropics and assimilation is not subject to the temporal gaps owing to frozen or snow-covered surface conditions. The spatial coverage of the increments is somewhat greater than that of the assimilated observations, especially near major water bodies, due to the spatial interpolation and extrapolation of the observational information in the distributed analysis update of the L4_SM algorithm. The differences in the number of analysis increments between Vv4030 and Vv3030 (Figure 5d) and the reasons for them match those of the assimilated observations (Figure 5b).

5.2. Mean of O-F Tb Residuals and Soil Moisture Increments

Nonnegligible time series mean values of the O-F Tb residuals and soil moisture increments can reveal shortcomings in the L4_SM algorithm calibration (section 3.3). Figure 6a shows the global distribution of the time series mean of the O-F Tb residuals in Vv4030. The values are typically small and mostly range from −3 to 3 K. The largest regional bias values are found across the Southern Hemisphere. This is because the L4_SM algorithm calibration (that is, the removal of seasonally varying bias in the modeled Tb; section 2.1) relies heavily on the climatological consistency between the various surface meteorological forcing data sets used during the SMAP period and the retrospective (calibration) period, including (i) the SMAP-period forcing data from the 1/4° GEOS FP system (GEOS-5.13.0 through GEOS-5.17 during the 3-year validation period), (ii) the historic forcing data from the 1/2° MERRA-2 system (GEOS-5.12.4), and (iii) the gauge-based CPCU precipitation data (outside of Africa and the high latitudes). In the Southern Hemisphere, the GEOS data are based on far fewer conventional measurements than in the Northern Hemisphere and are thus more prone to the typical discontinuities in global atmospheric analyses that result from changes in the satellite observing system (Robertson et al., 2014). Similarly, the changes over time in the precipitation gauge network that supports the CPCU product are more severe in the Southern Hemisphere (Reichle, Liu, et al., 2017, their Figure 8), resulting in greater discontinuities in the L4_SM precipitation forcing there.

In the global average, the bias (absolute bias) in the O-F Tb residuals is just 0.02 K (0.56 K) for Vv4030 (Figure 6a) and 0.12 K (0.58 K) for Vv3030 (not shown). Regionally, the difference map of the absolute bias (Figure 6b) indicates that the bias is improved (closer to zero) in Vv4030 compared to Vv3030 in central North America, the Amazon, and in the Horn of Africa but worse in southern Africa.

The time mean values of the surface soil moisture analysis increments in Vv4030 vanish in the global average (Figure 6c). Regionally, however, the 3-year mean (net) increments constitute a nonnegligible fraction of the water balance, with maximum (absolute) values reaching as high as 0.01 m^3/m^3 in some regions. Central North America, eastern South America, southern Africa, and portions of central Asia experience net drying increments, whereas West Africa and Australia experience net wetting increments. This pattern roughly matches the 3-year mean bias in the O-F residuals (Figure 6a). The mean increments are much smaller in root zone soil moisture, where maximum (absolute) values of the net increments are within 0.001 m^3/m^3 (Figure 6e).

The global average of the absolute values of the net increments for surface soil moisture in Version 4 is 0.0013 m^3/m^3 (not shown), which is almost twice as large as that in Version 3 (0.0007 m^3/m^3; not shown), with fairly uniform increases seen across the globe (Figure 6d). Conversely, the absolute net increments for root zone soil moisture are smaller in Vv4030 than in Vv3030 by a factor of 2 (Figure 6f). Since the root zone layer is much deeper than the surface layer, the overall water imbalance is therefore much smaller in Version 4 than in Version 3.

In summary, Figure 6 suggests that the L4_SM algorithm is reasonably unbiased and that Version 4 is less biased than Version 3 (and thus the documented Version 2; Reichle, De Lannoy, Liu, Koster, et al., 2017). The system, however, could obviously still benefit from further calibration.

5.3. Standard Deviation of O-F Tb Residuals and Soil Moisture Increments

The time series standard deviation of the O-F Tb residuals and soil moisture increments measures the typical magnitude of the adjustments resulting from the assimilation of the SMAP Tb observations. The time series

standard deviation of the O-F residuals ranges from a few Kelvin to around 15 K in Vv4030 (Figure 7a). The highest values are found in central North America, southern South America, southern Africa, the Sahel, portions of central Asia, India, and (particularly) Australia. These regions have sparse or modest vegetation cover and typically exhibit strong variability in soil moisture conditions. The O-F Tb residuals are generally smallest in more densely vegetated regions, including the eastern United States, the Amazon basin, and tropical Africa. Small values are also found in the high latitudes, including Alaska and Siberia, and in the Sahara Desert.

The global average of the standard deviation of the O-F Tb residuals is about 5.7 K in Vv4030 (Figure 7a), which is slightly reduced from about 5.9 K in Vv3030. The O-F standard deviations are reduced primarily in central North America and across midlatitude Eurasia (Figure 7b), where the Vv4030 (and Vv3030) values are among the largest. The smaller standard deviations in Vv4030 suggest that the Version 4 modeling system is better able to predict the observed Tb just prior to each analysis. This improvement is also reflected in the spatially averaged time series standard deviation of the O-A residuals, which is 3.7 K in Vv4030 and 4.0 K in Vv3030 (not shown). Note that the 3.7 K standard deviation of the Vv4030 O-A residuals relative to the respective 5.7 K value of the O-F residuals reflects the impact of the SMAP observations in the L4_SM system. (Note also that the global average O-F standard deviation of the preliminary Version 4 data used by Reichle, Liu, et al., 2018, is 5.8 K. Owing to a processing error, the 5.1 K number quoted in their Figure 15 is wrong.)

The time series standard deviation of the soil moisture increments is a measure of the typical magnitude of instantaneous increments. This metric is shown in Figures 7c and 7e for Vv4030 surface and root zone soil moisture increments, respectively. The spatial patterns mostly match that of the standard deviation of the O-F Tb residuals (Figure 7a). Typical values for surface soil moisture increments are on the order of 0.02–0.03 m^3/m^3 in central North America, southern South America, southern Africa, the Sahel, portions of central Asia, India, and most of Australia (Figure 7c). In the same regions, root zone soil moisture increments are typically around 0.004 m^3/m^3 (Figure 7e). Compared to Vv3030, the Vv4030 surface soil moisture increments in these regions are ~50% larger (Figure 7d), while the root zone soil moisture increments are ~50% smaller (Figure 7f). The smaller root zone increments are primarily the result of removing the catchment deficit model prognostic variable from the EnKF state vector (section 2.2.2). The larger surface increments are consistent with the increased dynamic range of surface soil moisture in Vv4030 resulting from the reduction in the modeled upward recharge of the surface layer from below (section 2.2.2). In densely vegetated regions, in particular in tropical forests, surface and root zone soil moisture increments are generally negligible (Figures 7c and 7e); in those areas, SMAP Tb observations are mostly sensitive to the dense vegetation and can thus provide only minimal corrections to the model forecast soil moisture.

5.4. Magnitude of Simulated Versus Actual Tb Errors

The standard deviation, v, of the *normalized* O-F Tb residuals measures the consistency between the simulated (modeled) uncertainty and the actual Tb errors (section 3.3). If the perturbation parameters that determine the simulated error standard deviations are chosen such that the simulated errors are statistically consistent with the actual errors, this metric, shown in Figure 8a for Vv4030, should be unity everywhere. The global average of the metric in Version 4030 is 1.13, suggesting that, on average, the simulated errors only slightly underestimate the actual errors. The metric, however, varies greatly across the globe. Typical values are either too low or too high. In the Amazon basin, the eastern United States, tropical Africa, the eastern Sahara Desert, and portions of the high northern latitudes, values range from 0.25 to 0.5, and thus the actual errors there are considerably overestimated. Conversely, in central North America, southern South America, southern Africa, the Sahel, portions of central Asia, India, and most of Australia, values range from 1.5 to 3, meaning that the actual errors in these regions are considerably underestimated. This is consistent with the finding that the L4_SM uncertainty estimates are biased low versus ubRMSE values core validation sites (Figure 4 and section 4.2), which are located within these regions (Table 1).

The global pattern of low and high values of the normalized O-F standard deviation for Vv3030 is very similar to that of Vv4030, but Vv3030 has a larger global average of 1.28 (not shown). The larger average in Vv3030 results primarily from larger values (compared to Vv4030) in central North America and across midlatitude Eurasia, where v is closer to the ideal value of 1 in Vv4030 than in Vv3030 (Figure 8b). Put differently, the underestimation of the actual errors was more pronounced in Vv3030 and is thus improved in Vv4030. This improvement stems primarily from the larger ensemble spread in the Vv4030 model forecast surface

soil moisture (which implies a greater ensemble spread in model forecast Tb) and, to a lesser extent, from the slightly smaller errors in the Vv4030 model forecast Tb. However, more work is clearly needed to further improve the calibration of the input parameters that determine the model and observation errors in the L4_SM system. (Note that the global average of the normalized O-F standard deviation for April 2015 to March 2017 Version 2 data is 1.26 and not 1.0, as erroneously stated by Reichle, Liu, et al., 2018, because of a processing error. The same error also applies to the 0.77 average quoted for the preliminary Version 4 data in Figure 15 of Reichle, Liu, et al., 2018, for which the corrected value is 1.14.)

6. Summary and Conclusions

The SMAP L4_SM algorithm assimilates L-band Tb observations from SMAP into the NASA Catchment land surface model driven with gauge-based precipitation observations. The resulting L4_SM data product provides global, 3-hourly, 9-km resolution estimates of surface and root zone soil moisture and land surface fluxes. This paper describes the changes imposed in Version 4 of the L4_SM algorithm and quantifies their impact on the quality of the L4_SM data product. The most important changes in the L4_SM system are (i) the reduction of the modeled upward recharge into the surface layer from below under nonequilibrium conditions, (ii) the removal from the analysis state vector of the model prognostic variable (catchment deficit) that governs the equilibrium soil moisture profile, and (iii) the use of SMAP Tb records together with longer SMOS Tb records for algorithm calibration.

The model change (i) results in generally drier and more variable surface soil moisture in the most recent product (Vv4030) compared to earlier versions, whereas root zone soil moisture in Vv4030 is generally wetter and less variable. Consequently, the surface and root zone soil moisture time series are more distinct from each other in Vv4030 (Figure 1). Across all 31 of the 9-km reference pixels from 18 SMAP core validation sites (Table 1), the surface soil moisture bias is smaller by 0.018 m^3/m^3 and the root zone soil moisture bias is larger by 0.013 m^3/m^3 in Vv4030 compared to Vv3030 (Figure 2). When discrepancies in vertical support of the model and in situ data are considered, the reduced bias in surface soil moisture is definitely an improvement, while the increased bias in root zone soil moisture may not represent a degradation. In any case, because of these changes in the soil moisture climatology, Version 3 and Version 4 data should not be mixed in applications. The skill in terms of ubRMSE, R, and anomaly R versus in situ measurements is essentially the same in both product versions. Perhaps most importantly, with an ubRMSE of 0.039 m^3/m^3 for surface soil moisture and 0.026 m^3/m^3 for root zone soil moisture, the Vv4030 estimates meet the L4_SM accuracy requirement of 0.04 m^3/m^3 on average across the 9-km core site reference pixels. As in previous versions, the Version 4 L4_SM product has better ubRMSE, R, and anomaly R skill than the corresponding Nature Run simulation, reflecting the added value of assimilating SMAP Tb observations.

The change in the Tb analysis (ii), in combination with the model change (i), resulted in ~50% larger typical surface soil moisture increments and ~50% smaller typical root zone increments in Vv4030 compared to Vv3030 (Figures 7d and 7f). Similarly, the 3-year mean surface (root zone) soil moisture increments are larger (smaller) in magnitude in Vv4030 (Figures 6d and 6f). Because the 5-cm surface layer is much thinner than the 100-cm root zone layer, the net water imbalance is smaller in Vv4030. Moreover, the typical O-F Tb residuals are slightly smaller in Vv4030 (Figure 7b). Finally, the simulated Tb error variance matches the variance of the actual O-F Tb residuals better in Vv4030 (Figure 8b). Taken together, these results indicate that the Vv4030 algorithm achieves the same product skill as the Vv3030 algorithm but with generally smaller O-F Tb residuals and smaller analysis increments. This is desirable because it means that—without loss of skill—the Version 4 L4_SM estimates are more consistent with the water- and energy-balanced process dynamics of the land surface model and less driven by nonphysical analysis increments.

The use of more extensive Tb records from SMAP and SMOS for L4_SM algorithm calibration (iii) greatly enhanced the locations and times for which SMAP Tb observations are assimilated in Version 4. SMAP data are now assimilated with near-global coverage, including in most of Eastern Europe, the Middle East, and East Asia, where SMAP observations could not be assimilated with the Version 2 algorithm. At the time of Version 2, SMOS data provided the only sufficiently long L-band record to support L4_SM algorithm calibration but could not be used in these regions because of strong and persistent RFI there. The longer SMOS and SMAP records used to calibrate Vv4030 also enhanced the assimilation coverage compared to Vv3030 across the high latitudes and in high-elevation regions (Figure 5).

The present paper also evaluates, for the first time, the soil moisture uncertainty and runoff estimates provided in the L4_SM product. The Vv4030 surface and root zone soil moisture uncertainty estimates are too low compared to actual errors (computed versus in situ measurements), although the underestimation is less severe than in Vv3030 (Figure 3). Moreover, the spatial and temporal patterns in the uncertainty estimates do not reflect those in the actual errors. Across 239 unregulated basins in the contiguous United States, the Vv4030 runoff data underestimate streamflow gauge measurements by 35 mm/year, which is a 33% improvement over the 52 mm/year underestimation in Vv3030 (Figure 4). The long-standing runoff bias in the Catchment model was examined by Koster and Mahanama (2012); their suggested approach to improve the mean runoff would have degraded the model's soil moisture skill and was thus deemed unsuitable for use in the L4_SM algorithm. The runoff correlation skill is slightly improved in the Version 4 modeling system, but the skill of the Vv4030 and Vv3030 runoff estimates after SMAP data assimilation ends up being essentially the same.

In summary, the data assimilation system underlying Version 4 of the L4_SM product includes a number of model and analysis changes. These changes resulted in several small improvements in the Vv4030 product, including more realistic surface soil moisture dynamics and smaller total soil moisture increments, all without compromising the accuracy of the product versus in situ measurements. There are, however, still opportunities for further improvement, both in the skill of the L4_SM estimates and in the measurements and methods used for validating the product. For example, recalibrating the parameters for the model forcing and prognostics perturbations might result in a more optimal analysis and thus further improve the skill of the L4_SM product. Moreover, evapotranspiration estimates should be evaluated against remote sensing products or validated against flux tower measurements once sufficiently long tower records for the SMAP era become publicly available. Compared to the ~1.6 million L4_SM model grid cells, however, there are precious few soil moisture core sites, streamflow gauges in unregulated basins, and flux tower sites. Our validation against in situ measurements should thus be supplemented with novel, spatially distributed evaluation approaches. For example, Dong et al. (2019) used independent satellite soil moisture observations to quantify the impact of assimilating SMAP observations into the L4_SM system and demonstrate that the greatest improvements are in otherwise data-sparse regions.

References

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

Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., et al. (2018). Satellite and in situ observations for advancing global Earth surface modelling: A review. *Remote Sensing*, *10*(12), 2038. https://doi.org/10.3390/rs10122038

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

Bircher, S., Skou, N., Jensen, K. H., Walker, J. P., & Rasmussen, L. (2012). A soil moisture and temperature network for SMOS validation in Western Denmark. *Hydrology and Earth System Sciences*, *16*(5), 1445–1463. https://doi.org/10.5194/hess-16-1445-2012

Bontemps, S., Defourny, P., van Bogaert, E., Arino, O., Kalogirou, V., & Ramos Perez, J. (2011). GLOBCOVER 2009 products description and validation report, ESA technical note. Available at http://due.esrin.esa.int/files/GLOBCOVER2009_Validation_Report_2.2.pdf. Last accessed 2 April 2019.

Bosch, D. D., Sheridan, J. M., & Marshall, L. K. (2007). Precipitation, soil moisture, and climate database, Little River Experimental Watershed, Georgia, United States. *Water Resources Research*, *43*, W09472. https://doi.org/10.1029/2006WR005834

Brocca, L., Pellarin, T., Crow, W. T., Ciabatta, L., Massari, C., Ryu, D., et al. (2016). Rainfall estimation by inverting SMOS soil moisture estimates: A comparison of different methods over Australia. *Journal of Geophysical Research: Atmospheres*, *121*, 12,062–12,079. https://doi.org/10.1002/2016JD025382

Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., & Savoie, M. H. (2012). EASE-Grid 2.0: Incremental but significant improvements for Earth-gridded data sets. *ISPRS International Journal of Geo-Information*, *1*(1), 32–45. https://doi.org/10.3390/ijgi1010032

Caldwell, T. G., Bongiovanni, T., Cosh, M. H., Halley, C., & Young, M. H. (2018). Field and laboratory evaluation of the CS655 soil water content sensor. *Vadose Zone Journal*, *17*(1), 0. https://doi.org/10.2136/vzj2017.12.0214

Chan, S. (2018). SMAP L1C radiometer half-orbit 36 km EASE-grid brightness temperatures, version 4. NASA National Snow and Ice Data Center DAAC, accessed 1 December 2018. https://doi.org/10.5067/zvilg0ps6cti

Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., et al. (2018). Development and assessment of the SMAP enhanced passive soil moisture product. *Remote Sensing of Environment*, *204*, 931–941. https://doi.org/10.1016/j.rse.2017.08.025

Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., & Aida, K. (2018). Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. *Remote Sensing of Environment*, *214*, 1–13. https://doi.org/10.1016/j.rse.2018.05.008

Chen, F., Crow, W. T., Colliander, A., Cosh, M. H., Jackson, T. J., Bindlish, R., et al. (2017). Application of triple collocation in ground-based validation of Soil Moisture Active/Passive (SMAP) level 2 data products. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(2), 489–502. https://doi.org/10.1109/jstars.2016.2569998

Chen, F., Crow, W. T., Cosh, M. H., Colliander, A., Asanuma, J., Berg, A., et al. (2019). Uncertainty of reference pixel soil moisture averages sampled at SMAP core validation sites. *Journal of Hydrometeorology*, *20*(8), 1553–1569. https://doi.org/10.1175/jhm-d-19-0049.1

Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Wayne Higgins, R., & Janowiak, J. E. (2008). Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research*, *113*, D04110. https://doi.org/10.1029/2007JD009132

Clewley, D., Whitcomb, J. B., Akbar, R., Silva, A. R., Berg, A., Adams, J. R., et al. (2017). A method for upscaling in situ soil moisture measurements to satellite footprint scale using random forests. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(6), 2663–2673. https://doi.org/10.1109/jstars.2017.2690220

Colliander, A., Berg, A., Montzka, C., Martinez-Fernandez, J., Seyfried, M., Al Jassar, H., et al. (2017). SMAP/in situ core validation site land surface parameters match-up data, version 1, NASA National Snow and Ice Data Center DAAC, accessed 1 December 2018, https://doi.org/10.5067/dxavixly18km

Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., et al. (2017). Validation of SMAP surface soil moisture products with core validation sites. *Remote Sensing of Environment*, *191*, 215–231. https://doi.org/10.1016/j.rse.2017.01.021

Coopersmith, E. J., Cosh, M. H., Petersen, W. A., Prueger, J., & Niemeier, J. J. (2015). Soil moisture model calibration and validation: An ARS watershed on the South Fork Iowa River. *Journal of Hydrometeorology*, *16*(3), 1087–1101. https://doi.org/10.1175/jhm-d-14-0145.1

Cosh, M. H., Jackson, T. J., Starks, P., & Heathman, G. (2006). Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation. *Journal of Hydrology*, *323*(1-4), 168–177. https://doi.org/10.1016/j.jhydrol.2005.08.020

Cosh, M. H., Starks, P. J., Guzman, J. A., & Moriasi, D. N. (2014). Upper Washita River experimental watersheds: Multiyear stability of soil water content profiles. *Journal of Environmental Quality*, *43*(4), 1328. https://doi.org/10.2134/jeq2013.08.0318

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., et al. (2012). Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, *50*, RG2002. https://doi.org/10.1029/2011RG000372

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

Crow, W. T., Chen, F., Reichle, R. H., & Xia, Y. (2019). Diagnosing bias in modeled soil moisture/runoff coefficient correlation using the SMAP Level 4 soil moisture product. *Water Resources Research*, *55*. https://doi.org/10.1029/2019WR025245

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

De Lannoy, G. J. M., Reichle, R. H., & Pauwels, V. R. N. (2013). Global calibration of the GEOS-5 L-band microwave radiative transfer model over nonfrozen land using SMOS observations. *Journal of Hydrometeorology*, *14*(3), 765–785. https://doi.org/10.1175/jhm-d-12-092.1

De Lannoy, G. J. M., Reichle, R. H., & Vrugt, J. A. (2014). Uncertainty quantification of GEOS-5 L-band radiative transfer model parameters using Bayesian inference and SMOS observations. *Remote Sensing of Environment*, *148*, 146–157. https://doi.org/10.1016/j.rse.2014.03.030

Diamond, H. J., Karl, T. R., Palecki, M. A., Baker, C. B., Bell, J. E., Leeper, R. D., et al. (2013). U.S. Climate Reference Network after one decade of operations: Status and assessment. *Bulletin of the American Meteorological Society*, *94*(4), 485–498. https://doi.org/10.1175/bams-d-12-00170.1

Dirmeyer, P., & Norton, H. (2018). Indications of surface and sub-surface hydrologic properties from SMAP soil moisture retrievals. *Hydrology*, *5*(3), 36. https://doi.org/10.3390/hydrology5030036

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

Dong, J., Crow, W. T., Reichle, R. H., Liu, Q., Lei, F., & Cosh, M. (2019). A global assessment of added value in the SMAP Level-4 Soil Moisture product relative to its baseline land surface model. *Geophysical Research Letters*, *46*, 6604–6613. https://doi.org/10.1029/2019GL083398

Ducharne, A., Koster, R. D., Suarez, M. J., Stieglitz, M., & Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model: 2. Parameter estimation and model demonstration. *Journal of Geophysical Research*, *105*(D20), 24,823–24,838. https://doi.org/10.1029/2000JD900328

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

Feldman, A. F., Akbar, R., & Entekhabi, D. (2018). Characterization of higher-order scattering from vegetation with SMAP measurements. *Remote Sensing of Environment*, *219*, 324–338. https://doi.org/10.1016/j.rse.2018.10.022

Feldman, A. F., Short Gianotti, D. J., Konings, A. G., McColl, K. A., Akbar, R., Salvucci, G. D., & Entekhabi, D. (2018). Moisture pulse-reserve in the soil-plant continuum observed across biomes. *Nature Plants*, *4*(12), 1026–1033. https://doi.org/10.1038/s41477-018-0304-9

Galle, S., Grippa, M., Peugeot, C., Moussa, I. B., Cappelaere, B., Demarty, J., et al. (2018). AMMA-CATCH, a critical zone observatory in West Africa monitoring a region in transition. *Vadose Zone Journal*, *17*(1), 0. https://doi.org/10.2136/vzj2018.03.0062

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/jcli-d-16-0758.1

Girotto, M., Reichle, R. H., Rodell, M., Liu, Q., Mahanama, S., & de Lannoy, G. J. M. (2019). Multi-sensor assimilation of SMOS brightness temperature and GRACE terrestrial water storage observations for soil moisture and shallow groundwater estimation. *Remote Sensing of Environment*, *227*, 12–27. https://doi.org/10.1016/j.rse.2019.04.001

González-Zamora, Á., Sánchez, N., Martínez-Fernández, J., Gumuzzio, Á., Piles, M., & Olmedo, E. (2015). Long-term SMOS soil moisture products: A comprehensive evaluation across scales and methods in the Duero Basin (Spain). *Physics and Chemistry of the Earth, Parts A/B/C*, *83-84*, 123–136. https://doi.org/10.1016/j.pce.2015.05.009

Heathman, G. C., Cosh, M. H., Merwade, V., & Han, E. (2012). Multi-scale temporal stability analysis of surface and subsurface soil moisture within the Upper Cedar Creek Watershed, Indiana. *Catena*, *95*, 91–103. https://doi.org/10.1016/j.catena.2012.03.008

Huffman, G. J., Adler, R. F., Bolvin, D. T., & Gu, G. (2009). Improving the global precipitation record: GPCP Version 2.1. *Geophysical Research Letters*, *36*, L17808. https://doi.org/10.1029/2009GL040000

Jackson, T. J., & Schmugge, T. J. (1991). Vegetation effects on the microwave emission of soils. *Remote Sensing of Environment*, *36*(3), 203–212. https://doi.org/10.1016/0034-4257(91)90057-d

Jensen, K. H., & Refsgaard, J. C. (2018). HOBE: The Danish Hydrological Observatory. *Vadose Zone Journal*, *17*(1), 0. https://doi.org/10.2136/vzj2018.03.0059

Juglea, S., Kerr, Y., Mialon, A., Wigneron, J.-P., Lopez-Baeza, E., Cano, A., et al. (2010). Modelling soil moisture at SMOS scale by use of a SVAT model over the Valencia Anchor Station. *Hydrology and Earth System Sciences*, *14*(5), 831–846. https://doi.org/10.5194/hess-14-831-2010

Keefer, T. O., Moran, M. S., & Paige, G. B. (2008). Long-term meteorological and soil hydrology database, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resources Research*, *44*, W05S07. https://doi.org/10.1029/2006WR005702

Kerr, Y. H., al-Yaari, A., Rodriguez-Fernandez, N., Parrens, M., Molero, B., Leroux, D., et al. (2016). Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation. *Remote Sensing of Environment*, *180*, 40–63. https://doi.org/10.1016/j.rse.2016.02.042

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., et al. (2010). The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, *98*(5), 666–687. https://doi.org/10.1109/jproc.2010.2043032

Khodayar, S., Coll, A., & Lopez-Baeza, E. (2019). An improved perspective in the spatial representation of soil moisture: Potential added value of SMOS disaggregated 1-km resolution "all weather" product. *Hydrology and Earth System Sciences*, *23*(1), 255–275. https://doi.org/10.5194/hess-23-255-2019

Konings, A. G., Piles, M., Das, N., & Entekhabi, D. (2017). L-band vegetation optical depth and effective scattering albedo estimation from SMAP. *Remote Sensing of Environment*, *198*, 460–470. https://doi.org/10.1016/j.rse.2017.06.037

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

Koster, R. D., Crow, W. T., Reichle, R. H., & Mahanama, S. P. (2018). Estimating basin-scale water budgets with SMAP soil moisture data. *Water Resources Research*, *54*, 4228–4244. https://doi.org/10.1029/2018WR022669

Koster, R. D., Liu, Q., Mahanama, S. P. P., & Reichle, R. H. (2018). Improved hydrological simulation using SMAP data: Relative impacts of model calibration and data assimilation. *Journal of Hydrometeorology*, *19*(4), 727–741. https://doi.org/10.1175/jhm-d-17-0228.1

Koster, R. D., & Mahanama, S. P. P. (2012). Land surface controls on hydroclimatic means and variability. *Journal of Hydrometeorology*, *13*(5), 1604–1620. https://doi.org/10.1175/jhm-d-12-050.1

Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., & Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *Journal of Geophysical Research*, *105*(D20), 24,809–24,822. https://doi.org/10.1029/2000JD900327

Kumar, S. V., Dirmeyer, P. A., Peters-Lidard, C. D., Bindlish, R., & Bolten, J. (2018). Information theoretic evaluation of satellite soil moisture retrievals. *Remote Sensing of Environment*, *204*, 392–400. https://doi.org/10.1016/j.rse.2017.10.016

Kumar, S. V., Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., et al. (2014). Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. *Journal of Hydrometeorology*, *15*(6), 2446–2469. https://doi.org/10.1175/jhm-d-13-0132.1

Liu, D., Mishra, A. K., Yu, Z., Yang, C., Konapala, G., & Vu, T. (2017). Performance of SMAP, AMSR-E and LAI for weekly agricultural drought forecasting over continental United States. *Journal of Hydrology*, *553*, 88–104. https://doi.org/10.1016/j.jhydrol.2017.07.049

Lucchesi, R. (2013). File specification for GEOS-5 FP, NASA GMAO Office Note 4 (version 1.0), 63 pp. Available at https://ntrs.nasa.gov. Last accessed 2 April 2019.

Lv, S., Zeng, Y., Su, Z., & Wen, J. (2019). A closed-form expression of soil temperature sensing depth at L-band. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(7), 4889–4897. https://doi.org/10.1109/tgrs.2019.2893687

Lv, S., Zeng, Y., Wen, J., Zhao, H., & Su, Z. (2018). Estimation of penetration depth from soil effective temperature in microwave radiometry. *Remote Sensing*, *10*(4), 519. https://doi.org/10.3390/rs10040519

Mahanama, S. P., Koster, R. D., Walker, G. K., Takacs, L. L., Reichle, R. H., de Lannoy, G., et al. (2015). Land boundary conditions for the Goddard Earth Observing System Model Version 5 (GEOS-5) Climate Modeling System—Recent updates and data file descriptions, NASA technical report series on global modeling and data assimilation, NASA/TM-2015-104606, Vol. 39, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA, 55 pp. Available at https://ntrs.nasa.gov. Last accessed 2 April 2019.

Mao, Y., Crow, W. T., & Nijssen, B. (2019). A framework for diagnosing factors degrading the streamflow performance of a soil moisture data assimilation system. *Journal of Hydrometeorology*, *20*(1), 79–97. https://doi.org/10.1175/jhm-d-18-0115.1

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

McNairn, H., Jackson, T. J., Wiseman, G., Belair, S., Berg, A., Bullock, P., et al. (2015). The Soil Moisture Active Passive Validation Experiment 2012 (SMAPVEX12): Prelaunch calibration and validation of the SMAP soil moisture algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(5), 2784–2801. https://doi.org/10.1109/tgrs.2014.2364913

Mecklenburg, S., Drusch, M., Kaleschke, L., Rodriguez-Fernandez, N., Reul, N., Kerr, Y., et al. (2016). ESA's Soil Moisture and Ocean Salinity mission: From science to operational applications. *Remote Sensing of Environment*, *180*, 3–18. https://doi.org/10.1016/j.rse.2015.12.025

Mishra, A., Vu, T., Veettil, A. V., & Entekhabi, D. (2017). Drought monitoring with Soil Moisture Active Passive (SMAP) measurements. *Journal of Hydrology*, *552*, 620–632. https://doi.org/10.1016/j.jhydrol.2017.07.033

Moghaddam, M., Silva, A. R., Clewley, D., Akbar, R., Hussaini, S. A., Whitcomb, J., et al. (2016). Soil moisture profiles and temperature data from SoilSCAPE sites, USA, ORNL Distributed Active Archive Center. https://doi.org/10.3334/ornldaac/1339

Oliva, R., Daganzo, E., Kerr, Y. H., Mecklenburg, S., Nieto, S., Richaume, P., & Gruhier, C. (2012). SMOS radio frequency interference scenario: Status and actions taken to improve the RFI environment in the 1400–1427-MHz passive band. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1427–1439. https://doi.org/10.1109/tgrs.2012.2182775

Pablos, M., González-Zamora, Á., Sánchez, N., & Martínez-Fernández, J. (2018). Assessment of root zone soil moisture estimations from SMAP, SMOS and MODIS observations. *Remote Sensing*, *10*(7), 981. https://doi.org/10.3390/rs10070981

Pablos, M., Martínez-Fernández, J., Sánchez, N., & González-Zamora, Á. (2017). Temporal and spatial comparison of agricultural drought indices from moderate resolution satellite soil moisture data over northwest Spain. *Remote Sensing*, *9*(11), 1168. https://doi.org/10.3390/rs9111168

Panciera, R., Walker, J. P., Jackson, T. J., Gray, D. A., Tanase, M. A., Ryu, D., et al. (2014). The Soil Moisture Active Passive Experiments (SMAPEx): Toward soil moisture retrieval from the SMAP mission. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(1), 490–507. https://doi.org/10.1109/tgrs.2013.2241774

Peng, J., Misra, S., Piepmeier, J. R., Dinnat, E. P., Meissner, T., le Vine, D. M., et al. (2017). Recalibration and validation of the SMAP L-band radiometer, 2017 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, https://doi.org/10.1109/igarss.2017.8127511

Peng, J., Misra, S., Chan, S., Chaubell, J., Bindlish, R., Bringer, A., et al. (2018). SMAP radiometer brightness temperature calibration for the L1B_TB, L1C_TB (Version 4), and L1C_TB_E (Version 2) data products, SMAP project report, Jet Propulsion Laboratory, Pasadena, CA. Available at: https://nsidc.org/sites/nsidc.org/files/technical-references/SMAP_L1_Assessment%2020180601_v9.pdf

Piepmeier, J. R., Focardi, P., Horgan, K. A., Knuble, J., Ehsan, N., Lucey, J., et al. (2017). SMAP L-band microwave radiometer: Instrument design and first year on orbit. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(4), 1954–1966. https://doi.org/10.1109/tgrs.2016.2631978

Piepmeier, J. R., Johnson, J. T., Mohammed, P. N., Bradley, D., Ruf, C., Aksoy, M., et al. (2014). Radio-frequency interference mitigation for the Soil Moisture Active Passive microwave radiometer. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(1), 761–775. https://doi.org/10.1109/tgrs.2013.2281266

Purdy, A. J., Fisher, J. B., Goulden, M. L., Colliander, A., Halverson, G., Tu, K., & Famiglietti, J. S. (2018). SMAP soil moisture improves global evapotranspiration. *Remote Sensing of Environment*, *219*, 1–14. https://doi.org/10.1016/j.rse.2018.09.023

Rajasekaran, E., Das, N., Poulsen, C., Behrangi, A., Swigart, J., Svoboda, M., et al. (2018). SMAP soil moisture change as an indicator of drought conditions. *Remote Sensing*, *10*(5), 788. https://doi.org/10.3390/rs10050788

Reichle, R. H., de Lannoy, G., Koster, R. D., Crow, W. T., & Kimball, J. S. (2017a). SMAP L4 global 3-hourly 9 km surface and rootzone soil moisture analysis update, version 3, NASA National Snow and Ice Data Center DAAC, accessed 1 June 2018, https://doi.org/10.5067/20uljh6ezkfj

Reichle, R. H., de Lannoy, G., Koster, R. D., Crow, W. T., & Kimball, J. S. (2017b). SMAP L4 global 3-hourly 9 km surface and rootzone soil moisture geophysical data, version 3, NASA National Snow and Ice Data Center DAAC, accessed 1 June 2018, https://doi.org/10.5067/b59dt1d5umb4

Reichle, R. H., de Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., & Liu, Q. (2018a). SMAP L4 global 3-hourly 9 km EASE-grid surface and root zone soil moisture analysis update, version 4, NASA National Snow and Ice Data Center DAAC, accessed 1 December 2018, https://doi.org/10.5067/60hb8vip2t8w

Reichle, R. H., de Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., & Liu, Q. (2018b). SMAP L4 global 3-hourly 9 km EASE-grid surface and root zone soil moisture geophysical gata, version 4, NASA National Snow and Ice Data Center DAAC, accessed 1 December 2018, https://doi.org/10.5067/kpjnn2gi1dqr

Reichle, R. H., de Lannoy, G. J. M., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., et al. (2017). Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements. *Journal of Hydrometeorology*, *18*(10), 2621–2645. https://doi.org/10.1175/jhm-d-17-0063.1

Reichle, R. H., de Lannoy, G. J. M., Liu, Q., Koster, R. D., Kimball, J. S., Crow, W. T., et al. (2017). Global assessment of the SMAP Level-4 surface and root-zone soil moisture product using assimilation diagnostics. *Journal of Hydrometeorology*, *18*(12), 3217–3237. https://doi.org/10.1175/jhm-d-17-0130.1

Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P. P., Koster, R. D., & de Lannoy, G. J. M. (2017). Assessment of MERRA-2 land surface hydrology estimates. *Journal of Climate*, *30*(8), 2937–2960. https://doi.org/10.1175/jcli-d-16-0720.1

Reichle, R. H., Liu, Q., Koster, R. D., Ardizzone, J. V., Colliander, A., Crow, W. T., et al. (2018). Soil Moisture Active Passive (SMAP) project assessment report for version 4 of the L4_SM data product, NASA technical report series on global modeling and data assimilation, NASA/TM-2018-104606, vol. 52, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA, 67 pp. Available at https://ntrs.nasa.gov. Last accessed 2 April 2019.

Reichle, R. H., Liu, Q., Koster, R. D., Draper, C. S., Mahanama, S. P. P., & Partyka, G. S. (2017). Land surface precipitation in MERRA-2. *Journal of Climate*, *30*(5), 1643–1664. https://doi.org/10.1175/jcli-d-16-0570.1

Reichle, R. H., Lucchesi, R. A., Ardizzone, J. V., Kim, G.-K., Smith, E. B., & Weiss, B. H. (2018). Soil Moisture Active Passive (SMAP) mission level 4 surface and root zone soil moisture (L4_SM) product specification document, GMAO Office Note 10, version 1.5, 83 pp. Available at https://ntrs.nasa.gov. Last accessed 2 April 2019.

Rigden, A. J., Salvucci, G. D., Entekhabi, D., & Short Gianotti, D. J. (2018). Partitioning evapotranspiration over the continental United States using weather station data. *Geophysical Research Letters*, *45*, 9605–9613. https://doi.org/10.1029/2018GL079121

Robertson, F. R., Bosilovich, M. G., Roberts, J. B., Reichle, R. H., Adler, R., Ricciardulli, L., et al. (2014). Consistency of estimated global water cycle variations over the satellite era. *Journal of Climate*, *27*(16), 6135–6154. https://doi.org/10.1175/jcli-d-13-00384.1

Román-Cascón, C., Pellarin, T., Gibon, F., Brocca, L., Cosme, E., Crow, W., et al. (2017). Correcting satellite-based precipitation products through SMOS soil moisture data assimilation in two land-surface models of different complexity: API and SURFEX. *Remote Sensing of Environment*, *200*, 295–310. https://doi.org/10.1016/j.rse.2017.08.022

Rowlandson, T., Impera, S., Belanger, J., Berg, A. A., Toth, B., & Magagi, R. (2015). Use of in situ soil moisture network for estimating regional-scale soil moisture during high soil moisture conditions. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques*, *40*(4), 343–351. https://doi.org/10.1080/07011784.2015.1061948

Sadri, S., Wood, E. F., & Pan, M. (2018). A SMAP-based drought monitoring index for the United States. *Hydrology and Earth System Sciences Discussions*, *1–19*, 1–19. https://doi.org/10.5194/hess-2018-182

Sanchez, N., Martinez-Fernandez, J., Scaini, A., & Perez-Gutierrez, C. (2012). Validation of the SMOS L2 soil moisture data in the REMEDHUS Network (Spain). *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1602–1611. https://doi.org/10.1109/tgrs.2012.2186971

Schaefer, G. L., Cosh, M. H., & Jackson, T. J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *Journal of Atmospheric and Oceanic Technology*, *24*(12), 2073–2077. https://doi.org/10.1175/2007jtecha930.1

Schmugge, T., Gloersen, P., Wilheit, T., & Geiger, F. (1974). Remote sensing of soil moisture with microwave radiometers. *Journal of Geophysical Research*, *79*(2), 317–323. https://doi.org/10.1029/JB079i002p00317

Seyfried, M. S., Murdock, M. D., Hanson, C. L., Flerchinger, G. N., & van Vactor, S. (2001). Long-term soil water content database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resources Research*, *37*(11), 2847–2851. https://doi.org/10.1029/2001WR000419

Shellito, P. J., Small, E. E., Colliander, A., Bindlish, R., Cosh, M. H., Berg, A. A., et al. (2016). SMAP soil moisture drying more rapid than observed in situ following rainfall events. *Geophysical Research Letters*, *43*, 8068–8075. https://doi.org/10.1002/2016GL069946

Shellito, P. J., Small, E. E., & Livneh, B. (2018). Controls on surface soil drying rates observed by SMAP and simulated by the Noah land surface model. *Hydrology and Earth System Sciences*, *22*(3), 1649–1663. https://doi.org/10.5194/hess-22-1649-2018

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research*, *116*, G04021. https://doi.org/10.1029/2011JG001708

Slater, J. A., Garvey, G., Johnston, C., Haase, J., Heady, B., Kroenung, G., & Little, J. (2006). The SRTM data "finishing" process and products. *Photogrammetric Engineering & Remote Sensing*, *72*(3), 237–247. https://doi.org/10.14358/pers.72.3.237

Tetlock, E., Toth, B., Berg, A., Rowlandson, T., & Ambadan, J. T. (2019). An 11-yr (2007-2017) soil moisture and precipitation dataset from the Kenaston Network in the Brightwater Creek basin, Saskatchewan, Canada. *Earth System Science Data*, *11*(2), 787–796. https://doi.org/10.5194/essd-11-787-2019

Verdin, K. L., Godt, J. W., Funk, C., Pedreros, D., Worstell, B., & Verdin, J. (2007). Development of a global slope dataset for estimation of landslide occurrence resulting from earthquakes, U.S. Geological Survey, Open-File Report 2007-1188, 25 pp. Available at https://pubs.usgs.gov. Last accessed 2 April 2019.

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., & Liu, C. (2007). A gauge-based analysis of daily precipitation over East Asia. *Journal of Hydrometeorology*, *8*(3), 607–626. https://doi.org/10.1175/jhm583.1