Soil Moisture Active Passive (SMAP) Project: Calibration and Validation for the L2/3_SM_P Version 8 and L2/3_SM_P_E Version 5 Data Products

Citation:

O'Neill, P.[2], S. Chan[3], R. Bindlish[2], M. Chaubell[3], A. Colliander[3], F. Chen[1], S. Dunbar[3], T. Jackson[1], J. Peng[2], M. Mousavi[3], M. Cosh[1], T. Bongiovanni[4], J. Walker[5], X. Wu[5], A. Berg[6], H. McNairn[7], M. Thibeault[8], J. Martínez-Fernández[9], Á. González-Zamora[9], E. Lopez-Baeza[10], K. Jensen[11], M. Seyfried[12], D. Bosch[13], P. Starks[14], C. Holifield Collins[15], J. Prueger[16], Z. Su[17], R. van der Velde[17], J. Asanuma[18], M. Palecki[19], E. Small[20], M. Zreda[21], J. Calvet[22], W. Crow[1], Y. Kerr[23], S. Yueh[3], and D. Entekhabi[24], October 12, 2021. *Calibration and Validation for the L2/3_SM_P Version 8 and L2/3_SM_P_E Version 5 Data Products*, SMAP Project, JPL D-56297, Jet Propulsion Laboratory, Pasadena, CA.

October 12, 2021

JPL D-56297

National Aeronautics and Space Administration

Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91109-8099
California Institute of Technology

Contributors to this report:

O'Neill, P.[2], S. Chan[3], R. Bindlish[2], M. Chaubell[3], A. Colliander[3], F. Chen[1], S. Dunbar[3], T. Jackson[1], J. Peng[2], M. Mousavi[3], M. Cosh[1], T. Bongiovanni[4], J. Walker[5], X. Wu[5], A. Berg[6], H. McNairn[7], M. Thibeault[8], J. Martínez-Fernández[9], Á. González-Zamora[9], E. Lopez-Baeza[10], K. Jensen[11], M. Seyfried[12], D. Bosch[13], P. Starks[14], C. Holifield Collins[15], J. Prueger[16], Z. Su[17], R. van der Velde[17], J. Asanuma[18], M. Palecki[19], E. Small[20], M. Zreda[21], J. Calvet[22], W. Crow[1], Y. Kerr[23], S. Yueh[3], and D. Entekhabi[24].

[1]USDA ARS Hydrology and Remote Sensing Lab, Beltsville, MD 20705 USA
[2]NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA
[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA
[4]University of Texas, Austin, TX 78713 USA
[5]Monash University, Clayton, Australia
[6]University of Guelph, Guelph, Canada
[7]Agriculture and Agri-Food Canada, Ottawa, Canada
[8]Comisión Nacional de Actividades Espaciales (CONAE), Buenos Aires, Argentina
[9]University of Salamanca, Salamanca, Spain
[10]University of Valencia, Valencia, Spain
[11]Center for Hydrology, Technical University of Denmark, Copenhagen, Denmark
[12]USDA ARS Northwest Watershed Research Center, Boise, ID 83712 USA
[13]USDA ARS Southeast Watershed Research Center, Tifton, GA 31793 USA
[14]USDA ARS Grazinglands Research Laboratory, El Reno, OK 73036 USA
[15]USDA ARS Southwest Watershed Research Center, Tucson, AZ 85719 USA
[16]USDA ARS National Laboratory for Agriculture and the Environment, Ames, IA 50011 USA
[17]University of Twente, Enschede, Netherlands
[18]University of Tsukuba, Tsukuba, Japan
[19]NOAA National Climatic Data Center, Asheville, NC 28801 USA
[20]University of Colorado, Boulder, CO 80309 USA
[21]University of Arizona, Tucson, AZ 85751 USA
[22]CNRM-GAME, UMR 3589 (Météo-France, CNRS), Toulouse, France
[23]CESBIO-CNES, Toulouse, France
[24]Massachusetts Institute of Technology, Cambridge, MA 02139 USA

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

During the post-launch Cal/Val Phase of SMAP there are two objectives for each science product team: 1) calibrate, verify, and improve the performance of the science algorithms, and 2) validate accuracies of the current release of the science data products as specified in the L1 science requirements according to the Cal/Val timeline. This report provides analysis and assessment of the SMAP Level 2 Soil Moisture Passive (L2SMP) Version 8 and the L2SMP Enhanced (L2SMP_E) Version 5 data products (also known as the R18 data release in October, 2021). The L2SMP product is provided on a 36-km EASE2 grid and the L2SMP_E on a 9-km EASE2 grid. The SMAP Level 3 Soil Moisture Passive (L3SMP, L3SMP_E) products are simply a daily composite of the L2 half-orbit files. Hence, analysis and assessment of the L2SMP and L2SMP_E products can also be considered to cover the L3SMP and L3SMP_E products. The assessment period now covers the first six years of the SMAP mission from April 1, 2015 to March 31, 2021.

The R18 release in October 2021 incorporates a number of changes since the R17 data release in August 2020 which impact SMAP data products. A point release (R17.1) on April 29, 2021 involved a new Antenna Scan Angle (ASA) correction for the L1B_TB brightness temperature product. The new code examines the initial scan angle in each scan and flags any scan angle that falls out of the normal range as bad (L1B_TB version 5). For the SMAP Level 2 and 3 passive soil moisture products, the R18 data release includes the following:

- L2/3_SM_P_E soil moisture products are now output on both the global EASE2 and the North Polar EASE2 grids. The L2/3_SM_P_E data fields on global and polar grids are organized in two separate HDF5 data groups called 'Soil_Moisture_Retrieval_Data' and 'Soil_Moisture_Retrieval_Data_Polar', respectively, in the same granule. Each group contains the same data fields and their associated definitions. The polar grid projection offers a more uniform spatial sampling at high latitudes (less distortion than the global grid at these high latitudes), which will facilitate studies with soil moisture in the boreal and arctic regions.

- Organic content and sand fraction fields have now been added to the L2 output products.

- A correction has been applied to unmix surface-corrected T_B in the L1C_TB data used as input to the L2/L3_SM_P/P_E soil moisture retrieval algorithms, mainly affecting areas around coastal areas and surrounding inland open water bodies (see Appendix 2). 93% of all global land pixels show a retrieval difference of less than ±0.001 m^3/m^3 after unmixing, confirming the limited scope of this procedure on the majority of global land pixels.

- SCA-V was the original postlaunch baseline algorithm for the L2/L3_SM_P/P_E products from 2015-2021. **Based on improved soil moisture retrieval performance in some agricultural areas, the DCA algorithm is now the SMAP baseline soil moisture retrieval algorithm, although DCA and SCA-V continue to have similar overall performance.** Default pointers in the output products have been changed accordingly (Appendix 3). Note that soil moisture retrievals from all three algorithms – SCA-H, SCA-V, and DCA – are still provided in the L2/3_SM_P/P_E output products.

- Assessment period has been extended to cover April 1, 2015 through March 31, 2021.

Assessment methodologies utilized in this report include comparisons of SMAP soil moisture retrievals with *in situ* soil moisture observations from core validation sites (CVS) and sparse networks, and intercomparison with products from ESA's Soil Moisture Ocean Salinity (SMOS) mission. The primary

assessment methodology is the CVS comparisons using established metrics and time series plots. These metrics include unbiased root mean square error (ubRMSE), bias, mean absolute bias (MAB), and correlation. The ubRMSE captures time-random errors, bias and MAB capture the mean differences or offsets, and correlation captures phase compatibility between data series. It should be noted that some changes have been made in the calibration and upscaling of select CVS based upon follow-up investigations by Cal/Val Partners. In addition, the assessment period is now 72 months (as opposed to 60 months in the L2SMP Version 7 and L2SMP_E Version 4 assessments [27]).

SMAP L2SMP supports a total of three alternative retrieval algorithms: Single Channel Algorithm–H polarization (SCA-H), Single Channel Algorithm–V polarization (SCA-V), and Dual Channel Algorithm (DCA) are the focus of this assessment. These same retrieval algorithms were also used in evaluating the performance of L2SMP_E.

In all previous SMAP data releases through the R16 release in August 2019, the accuracy of retrieved soil moisture from the SMAP L2SMP 36-km standard product was assessed using a validation grid (VG) approach (Sec. 6.1). In this approach, the 36-km EASE 2.0 grid was shifted at each CVS by 3-km increments in order to cover the maximum number of *in situ* measurement points possible at the given CVS before performance statistics were generated comparing SMAP soil moisture to *in situ* soil moisture. The VG shifting was also done at each CVS to minimize the presence of non-representative areas and to maintain homogeneity across the grid cell to the extent possible. However, in the R17 data release of August, 2020 and all data releases going forward, the SMAP project decided to proceed with the L2SMP assessment using the same procedure as used for the L2SMP_E assessment – i.e., without using the shifted VG grid for the 36-km product. Although this non-VG assessment could introduce some additional uncertainty, past assessment reports [1, 20, 25] have documented the close agreement in performance metrics between L2SMP at 36-km posting with VG processing and L2SMP_E at 9-km posting. With the non-VG L2SMP comparisons with *in situ* data from both CVS and sparse networks given in Sec. 8.1 and the L2SMP_E assessments given in Sec. 8.2, it is felt that the community has sufficient information to judge the quality and accuracy of the L2SMP retrieved soil moisture.

The first step in assessment was the comparison of the L2SMP AM (Descending) and PM (Ascending) Version 8 products to the CVS and sparse network observations. CVS AM results indicated that the SCA-V and DCA provided approximately the same best *overall* performance with an ubRMSE of 0.036 m^3/m^3, bias of 0.010 m^3/m^3 and correlation of 0.819 for SCA-V and an ubRMSE of 0.036 m^3/m^3, bias of 0.012 m^3/m^3 and correlation of 0.816 for DCA. DCA shows significantly better ubRMSE for two agricultural sites, Carmen and Monte Buey. The CVS PM results are similar, with an overall ubRMSE of 0.037 m^3/m^3, bias of 0.007 m^3/m^3 and correlation of 0.822 for SCA-V and an ubRMSE of 0.035 m^3/m^3, bias of 0.005 m^3/m^3 and correlation of 0.801 for DCA, with DCA again having much lower ubRMSE at Carmen and Monte Buey. These metrics exceed the SMAP mission requirements and those of the SMOS products. A portion of the change in the metrics from the previous data release may be associated with the longer period of record (72 vs 60 months) since longer records may include a wider range of anomalous conditions or a more typical set of conditions. Sparse network results confirmed the trends seen in the CVS comparisons. The overall conclusion is that the L2SMP AM and PM products have continued to have a sustained performance that exceeds the mission accuracy requirements for L2 passive retrieved soil moisture (*the L2SMP soil moisture shall meet or exceed an accuracy of 0.040 m^3/m^3 ubRMSE over land in the absence of frozen ground, permanent snow/ice, or dense vegetation;* this requirement is actually written for retrieved 6 AM soil moisture at 10 km spatial resolution, but has been applied by the SMAP team to all L2 passive soil moisture products). The combination of analyses using the CVS and sparse networks, intercomparison with products from the SMOS mission, and recent triple colocation analyses have contributed to a better understanding of the performance uncertainties. The assessment now includes 72 months of intercomparisons, and several papers have been published in peer-reviewed journals [2-5, 21-24] as well as numerous investigations listed in the SMAP Bibliography posted at NSIDC at

https://nsidc.org/data/smap/research.html. These analyses satisfy the criteria established by the Committee on Earth Observing Satellites (CEOS) for Stage 4 validation.

The L2SMP_E are posted at 9 km but the contributing domain (i.e. primary spatial area contributing to the radiometer brightness temperature response) is approximately 33 km. A different set of CVS domains than those used for the L2SMP were identified in order to assess the performance of the L2SMP_E product; all ground measurements of soil moisture within the 33-km domain were used and compared to the SMAP retrieved soil moisture at each CVS. Additional information on the L2SMP_E product can be found in [1, 25].

Version 5 of the L2SMP_E, for both AM and PM orbits, was assessed using the CVS and sparse networks. Both the AM and PM products meet the mission requirements and reflect the patterns for L2SMP. Among the algorithms, SCA-V and DCA had similarly good overall metrics with an ubRMSE of 0.038 m^3/m^3, bias of -0.006 m^3/m^3 and correlation of 0.816 for SCA-V AM orbits, and an ubRMSE of 0.037 m^3/m^3, bias of -0.008 m^3/m^3 and correlation of 0.810 for the SCA-V PM orbits. DCA performance shows an ubRMSE of 0.036 m^3/m^3, bias of -0.009 m^3/m^3 and correlation of 0.818 for AM orbits, and an ubRMSE of 0.035 m^3/m^3, bias of -0.013 m^3/m^3 and correlation of 0.792 for the PM orbits. DCA continues to show improved performance over SCA-V at a few specific agricultural CVS. New to the R18 data release, L2SMP_E soil moisture is also output on the North Polar EASE2 grid, with results very consistent with those from the global grid. Both SCA-V and DCA satisfy the SMAP mission accuracy requirements. For the same reasons noted for the L2SMP, the maturity of the L2SMP_E product remains at CEOS Stage 4.

Overall conclusions in this assessment:

- L2SMP and L2SMP_E performances continue to meet project target accuracy (0.040 m^3/m^3 ubRMSE or better).
- The DCA algorithm now has at least comparable performance to SCA-V in meeting project target accuracy (0.040 m^3/m^3 ubRMSE or better).
- Based on improved soil moisture retrieval performance in some agricultural areas and consistent overall performance, the DCA algorithm is now the SMAP baseline soil moisture retrieval algorithm.
- Results from analysis of retrieved soil moisture on the polar grid are consistent with global grid results.
- Both the L2SMP and L2SMP_E products remain at CEOS Validation Stage 4.

2 OBJECTIVES OF CAL/VAL

During the post-launch Cal/Val (Calibration/Validation) Phase of SMAP there are two objectives for each science product team:

- Calibrate, verify, and improve the performance of the science algorithms, and
- Validate accuracies of the science data products as specified in Level 1 science requirements according to the Cal/Val timeline.

The process is illustrated in Figure 2.1. In this Assessment Report the progress of the Level 2 Soil Moisture Passive Team in addressing these objectives is described. The approaches and procedures utilized follow those described in the SMAP Cal/Val Plan [6] and Algorithm Theoretical Basis Document for the Level 2 & 3 Soil Moisture (Passive) Data Products [30].

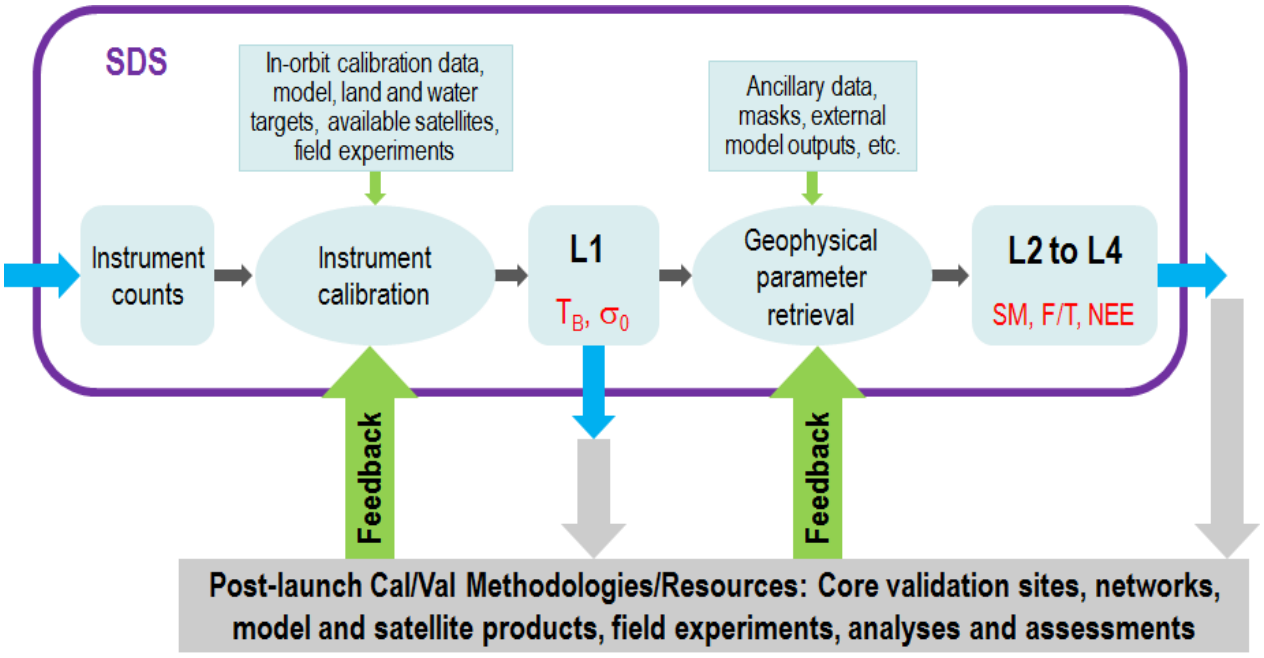


Figure 2.1. Overview of the SMAP Cal/Val Process.

SMAP established a unified definition base in order to effectively address the mission requirements. These are documented in the SMAP Handbook/ Science Terms and Definitions [8], where Calibration and Validation are defined as follows:

- *Calibration*: The set of operations that establish, under specified conditions, the relationship between sets of values or quantities indicated by a measuring instrument or measuring system and the corresponding values realized by standards.
- *Validation:* The process of assessing by independent means the quality of the data products derived from the system outputs.

The L2SMP Team adopted the same soil moisture retrieval accuracy requirement for the fully validated L2SMP data (0.040 m^3/m^3) that is listed in the L1 Mission Requirements Document [9] for the active/ passive soil moisture product.

In assessing the maturity of the L2SMP (and L2SMP_E) products, the team considered the guidance provided by the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) [10]:

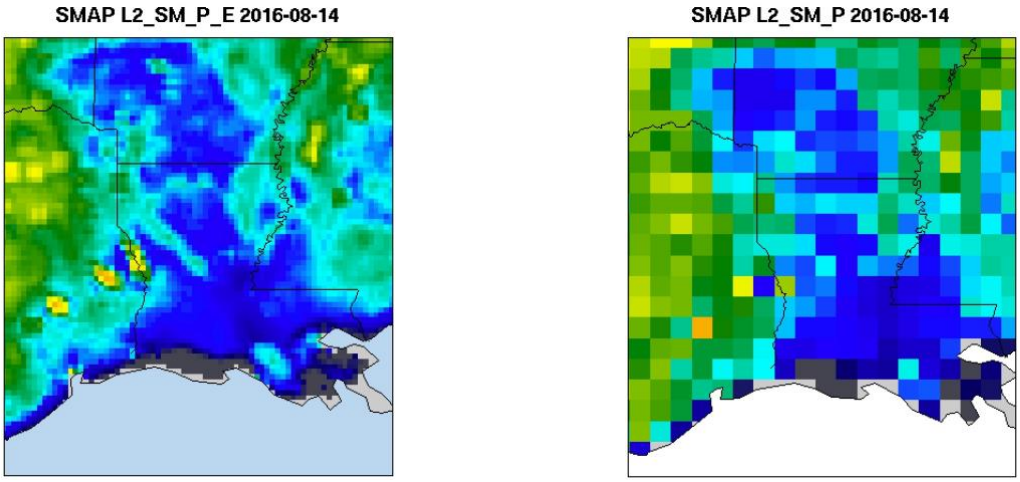
- Stage 1: Product accuracy is assessed from a small (typically < 30) set of locations and time periods by comparison with *in situ* or other suitable reference data.
- Stage 2: Product accuracy is estimated over a significant set of locations and time periods by comparison with reference *in situ* or other suitable reference data. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and time periods. Results are published in the peer-reviewed literature.
- Stage 3: Uncertainties in the product and its associated structure are well quantified from comparison with reference *in situ* or other suitable reference data. Uncertainties are characterized in a statistically robust way over multiple locations and time periods representing global conditions. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and periods. Results are published in the peer-reviewed literature.
- Stage 4: Validation results for stage 3 are systematically updated when new product versions are released and as the time series expands.

Based on the extensive validation analyses to date, the number of peer reviewed publications as well as numerous investigations listed in the Bibliography section of the report, the length of the SMAP period of record, and the utilization of feedback of validation in a systematic update, with this version of L2SMP and L2SMP_E the team has completed Stage 4. The Cal/Val program will continue with the goals of increasing the robustness of the soil moisture products and addressing specific site issues.

3 BRIEF DESCRIPTION OF THE L2SMP AND L2SMP_E

The L2SMP product is derived using SMAP L-band radiometer time-ordered observations (L1CTB product) as the primary input [7] along with other ancillary data on finer grid resolutions, to retrieve soil moisture (and other geophysical parameters as applicable) from a forward model. The resulting soil moisture retrieval output fields, along with others carrying supplementary geolocation information, brightness temperatures, quality flags, and ancillary data, are posted on a 36-km Earth-fixed grid using the global cylindrical Equal-Area Scalable Earth Grid projection, Version 2 (EASE Grid 2.0). The 36-km grid resolution is close to the 3-dB native spatial resolution of the instrument observations. The use of the fixed grid facilitates temporal analyses and ingestion of the products into some user applications. However, it presents challenges to validation given that many core validation sites (CVS) are not centered or contained in a single 36-km EASE grid cell. As a result, a shifted variation of the L2SMP grid has been used in all previous SMAP data releases for validation and assessment purposes (Validation Grid-VG). As noted in Sections 1 (Executive Summary) and 6.1, however, non-VG-based L2SMP assessment was performed in the R17 release in 2020 and similarly in the R18 data release of October 2021.

Following the SMAP launch, methodologies for improving the spatial information of the SMAP radiometer products were explored that resulted in the L1CTB_E (Enhanced) product. Backus-Gilbert (BG) optimal interpolation methodology is used that takes advantage of the radiometer oversampling on orbit. The processing results in data at a higher spatial density by virtue of TB interpolation at a 9-km grid resolution in L1BTB_E. It is important to note that the L1CTB_E processing does not improve the native resolution (~36 km) of the original TB measurements acquired by the SMAP radiometer. It is a posting of data interpolated to a 9-km grid, which can enhance spatial information (see Figure 3.1). The fine grid resolution (9 km) of L1CTB_E provides a convenient basis to produce passive soil moisture retrieval at the same fine grid resolution. Operationally, this is achieved by applying the same soil moisture inversion algorithms used for the standard 36-km L2SMP product to the enhanced 9-km L2SMP_E product. The 9-km posting provides more flexibility in co-locating the grids and CVS data and therefore does not require the use of the VG in validation and assessment of the L2SMP_E product.



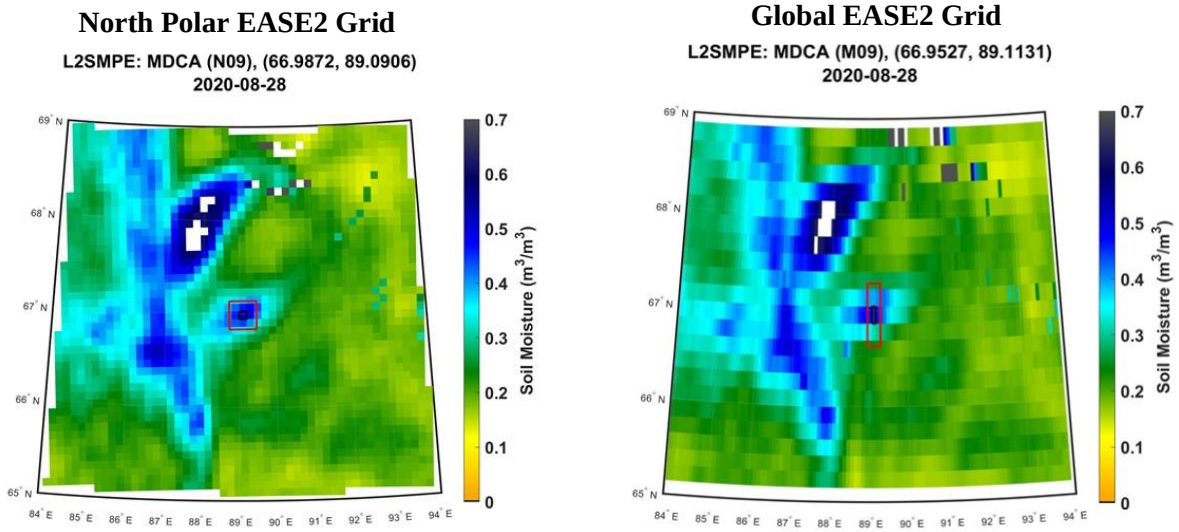
(a) Enhanced Passive Soil Moisture Product (b) Standard Passive Soil Moisture Product

Figure 3.1. Compared with the current standard L2SMP soil moisture product in (b), the enhanced L2SMP_E soil moisture product in (a) demonstrates a more detailed distribution of surface soil moisture and shows spatial features more clearly than does the standard product.

3.1 Addition of North Polar ease2 grid

Starting with SMAP data release R18 in October 2021, the L2/3_SM_P_E Version 5 Soil Moisture Data Product is also provided on the North Polar 9-km EASE2 grid projection in addition to the original Global 9-km EASE2 grid projection. A feature of the global grid is that its grid cells become significantly distorted (from the nominal square shape) at higher latitudes. This distortion in the global grid as compared to the North Polar grid is easily seen in the images in Figure 3.2. The polar grid projection offers a more uniform spatial sampling at high latitudes, which will facilitate studies with soil moisture in the boreal and arctic regions. Better spatial representation of the grid with respect to the SMAP radiometer footprint will also help in the improvement of the retrieval algorithms to account for the unique soil and vegetation conditions of the boreal and arctic landscapes.

The algorithms used to generate soil moisture on the polar grid (SCA-H, SCA-V and DCA) are exactly the same as the ones used to generate soil moisture on the global grid. However, the algorithms use the polar grid TB from the L1C_TB_E product (which has always been available in the L1C_TB products) instead of the TB on the global grid [28]. The ancillary data sources are the same as those used by the global grid soil moisture, but the ancillary data are gridded on the polar grid instead of the global grid before being input to the soil moisture retrieval algorithms for soil moisture on the polar grid. The approach to assign quality and surface flags and the threshold values used in the flag assignments are exactly the same as the ones used for the global grid soil moisture. Any difference between the soil moisture values found on the global grid and on the polar grid for a similar location is, therefore, due to the differences in the gridding. The differences are likely to be more significant at higher latitudes compared to lower latitudes because of the larger distortion of the global grid at the higher latitudes.



(a) (b)

Figure 3.2. Comparison of SMAP retrieved soil moisture on August 28, 2020 from the same northern location (lat/lon 66.99°N, 89.11°E) as plotted on the (a) North Polar EASE2 grid and (b) global EASE2 grid. Soil moisture shown was obtained using the Dual Channel Algorithm (DCA).

4 L1 RADIOMETER PRODUCT UPDATES

In the R18 data release, SMAP L2SMP soil moisture retrievals are based on the 2021 Version 5 of the radiometer Level 1B and 1C brightness temperature data [https://nsidc.org/data/smap/smap-data.html], which includes the new Antenna Scan Angle (ASA) correction code implemented in the R17.1 point release of April 2021 – otherwise, the calibration of the brightness temperature data is the same as reported for the 2020 Version 5 TB data in the R17 release. An assessment of the SMAP brightness temperature data quality and calibration is available in the relevant assessment reports posted at NSIDC [https://nsidc.org/data/smap/technical-references]. A more detailed discussion of the radiometer calibration and products can be found in [11,13,26]. In particular, information about calibration changes in Version 5 of the L1C_TB can be found in [29], which is also posted at NSIDC. The Version 5 TB data meet the noise equivalent delta temperature (NEDT) and geolocation requirements with margin as they did in Versions 3 and 4 (see Table 4.1) [14].

Briefly, the SMAP L1 radiometer calibration algorithm for the Version 5 release in 2020 was significantly revised compared to the previous Version 4 dataset. The calibration changes can be summarized as follows: (1) the radiometer internal reference load was used as the hot calibration target instead of the global ocean used in the Version 4 dataset due to unknown bias in the ocean emissivity model, and (2) the use of concurrent reflector emissivity correction, antenna pattern correction (APC), and noise-diode calibration. The 2020 release radiometer calibration is being used for all L2 passive soil moisture products; the incorporation of the new ASA code is the only brightness temperature product code change for the current R18 data release.

The SMAP radiometer team implemented in 2020 an optimized concurrent calibration scheme that utilizes cold-sky looks with ocean back-lobes and cold-sky looks with ocean/land transition back-lobes. The retrieved noise-diode temperature, reflector emissivity and APC value calibration coefficients along with that of the adjustment of the V and H-pol reflector emissivity values are forced to be the same. The wind speed-dependent sky maps are still used to correct for any galaxy contribution into the antenna temperature for ocean regions. Any previously observed fore-aft differences in L1C_TB radio frequency interference (RFI) still remain. The RFI behavior is similar as before: conditions in the Americas and Europe are good with poorer conditions in Asia.

Table 4.1. Version 3* Characteristics of SMAP L1 Radiometer Data
(*now superseded by Version 5)

Parameter		**Mission Requirement**
NEDT[1]	1.1 K	< 1.6 K[1]
Geolocation accuracy	2.7 km	< 4 km
Land SMAP/SMOS bias (H pol)	-0.90 K	n/a
Land SMAP/SMOS bias (V pol)	-2.37 K	n/a

[1]An NEDT of 1.6 K for a single-look L1B_TB footprint is equivalent to an NEDT of 0.51 K on a 30 x 30 km grid (Table 2.1 in SMAP Radiometer Error Budget, JPL D-61632 [14]). When combined with other error terms in the L1 radiometer error budget, the current single-look footprint NEDT of 1.1 K should result in an NEDT of less than 0.51 K on a 30 x 30 km grid. If all other error sources are within the requirements, this level of NEDT (< 0.51 K) should result in a total radiometric uncertainty of less than 1.3 K as required in the L2SMP error budget.

The changes in SMAP calibration in 2020 resulted in a change in metrics when comparing SMAP TB to SMOS TB. For the assessment period of 2015-2020, SMAP global average brightness temperature comparisons over land areas are approximately -0.90 K (for H pol) and -2.37 K (for V pol) cooler than SMOS (mean bias at top of the atmosphere after Faraday rotation correction was applied). Figure 4.1 shows the SMAP and SMOS observations over land for five years of SMAP data (2015-2020). Statistical analysis results are summarized in Table 4.2. The SMAP brightness temperatures show a very strong correlation with the SMOS observations and a strong linear relationship.

For SMAP/SMOS TB comparisons, it should be noted that while SMOS and SMAP each have an equatorial overpass time of 6 AM, the 6 AM orbits for SMOS are ascending orbits and the 6 AM orbits for SMAP are descending orbits. In order to minimize intercomparison errors associated with temporal changes in soil moisture and temperature, a maximum time window between the two satellite observations of 30 minutes was allowed. Both SMAP and SMOS have an average 3-db footprint size of ~40 km. Spatial variations in the contributing area were minimized by only using observations when the footprint distance was less than 1 km between SMAP and SMOS. Brightness temperatures at the top of the atmosphere (TOA) were used in the intercomparison. This analysis was done for both the horizontal (H) and vertical (V) polarizations. Microwave observations from the SMOS mission were reprocessed to approximate SMAP microwave radiometer observations made at a constant incidence angle of 40°. Only the alias-free portions of the SMOS field-of-view were used in the comparison. Additionally, the alias-free portions of the swath provide brightness temperatures with the lowest NEΔT. SMOS data version v620 was used for the analysis.

Recently, SMOS released Version 7 data (TB data ID is 724 and SM data ID is 700) in May 2021 which resulted in slight changes to the SMOS L1 TB calibration. SMOS-SMAP comparisons for the most recent versions of the respective mission data are ongoing – the numbers in Table 4.2 will be updated when available. The outcome of this new comparison may show that SMOS and SMAP data are in closer agreement.

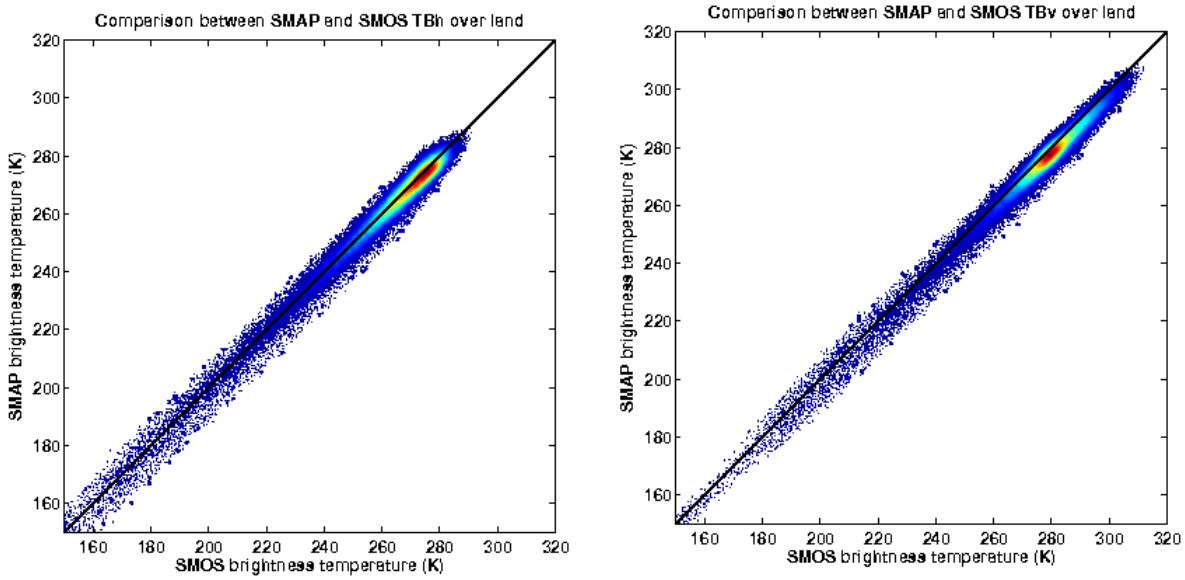


Figure 4.1. Density plot of the L1 brightness temperature comparison (top of the atmosphere) between SMAP (T16516) and SMOS (version 620) observations over land targets for V-pol (right) and H-pol (left) for 2015-2020.

Table 4.2. Summary statistics of the brightness temperature comparison between SMOS (version 620) and SMAP (T16516) for May 5, 2015-March 31, 2020.

		RMSD (K)	R	Bias [SMAP-SMOS] (K)	ubRMSD (K)
H pol	Land	3.56	0.9902	**-0.90**	3.45
	Ocean	2.18	0.8522	-0.46	2.14
	Overall	2.60	0.9994	-0.57	2.53
V pol	Land	3.81	0.9903	**-2.37**	2.98
	Ocean	2.55	0.6967	-1.59	2.00
	Overall	2.92	0.9994	-0.46	2.31

4.1 Water Body Correction

Prior to implementing the actual soil moisture retrieval, a preliminary step in the processing is to perform a water body correction to the brightness temperature data for cases where a significant percentage of the grid cell contains open water. For the Version 7 L2SMP and Version 4 L2SMP_E and all versions going forward, water correction is performed at the footprint level using the SMAP radiometer antenna gain pattern. This correction procedure is performed in the Version 5 SMAP L1B Radiometer Half-Orbit Time-Ordered Brightness Temperatures (L1BTB) product. Both the horizontally and vertically polarized L1B brightness temperatures over land are corrected for the presence of water within the antenna field of view (FOV). The resulting L1B brightness temperatures are then interpolated on the 36-km EASE Grid 2.0 projections using the inverse-distance squared interpolation method and on the 9-km EASE Grid 2.0 projections using the Backus-Gilbert optimal interpolation method. Overall it is expected that over land, the resulting brightness temperatures will become warmer upon the removal of the contribution of water to the original uncorrected observations. As stated in the product page of the Version 5 SMAP L1BTB product, water correction is performed as long as the antenna-gain-weighted water fraction within the antenna FOV is less than or equal to 0.9 and when the antenna boresight falls on a land location as indicated by a static high-resolution land/water mask. Further details of this procedure can be found in the User Guide, ATBD [13], or Assessment Report of the Version 5 SMAP L1B Radiometer Half-Orbit Time-Ordered Brightness Temperatures (L1BTB) product [29].

For Version 8 L2SMP and Version 5 L2SMP_E in the R18 data release of October 2021, an additional correction has been applied to unmix surface-corrected T_B in the L1C_TB data used as input to the L2/L3_SM_P/P_E soil moisture retrieval algorithms, mainly affecting areas around coastal areas and surrounding inland open water bodies (see Appendix 2 for more detailed information). 93% of all global land pixels show a retrieval difference of less than ±0.001 m^3/m^3 after unmixing, confirming the limited scope of this procedure on the majority of global land pixels.

5 ALTERNATIVE L2SMP/L2SMP_E ALGORITHMS

The current L2SMP/L2SMP_E products contain soil moisture retrieval fields produced by the baseline and two optional algorithms. Inside an L2SMP/L2SMP_E granule the *soil_moisture* field is the one that links to the retrieval result produced by the currently-designated baseline algorithm. At present, the operational L2SMP/L2SMP_E Science Production Software (SPS) produces and stores soil moisture retrieval results from the following three algorithms:

1. Single Channel Algorithm V-pol (SCA-V)
2. Single Channel Algorithm H-pol (SCA-H)
3. Dual Channel Algorithm (DCA)

All algorithms operate on the same zeroth-order microwave emission model commonly known as the *tau-omega* model. In essence, this model relates brightness temperatures (SMAP L1 observations) to soil moisture (SMAP L2 retrievals) through ancillary information (e.g. soil texture, soil temperature, surface roughness, and vegetation water content) and a soil dielectric model. The algorithms differ in their approaches to solve for soil moisture from the model under different constraints and assumptions. Below is a concise description of these three algorithms. Further details are provided in [30].

Given the results to date from the L2SMP/L2SMP_E Cal/Val analyses, the SCA-V algorithm and the regularized version of the DCA algorithm deliver comparable overall performance in retrieved soil moisture accuracy, while DCA shows improved retrievals in some problematic agricultural areas. SCA-V was the original postlaunch baseline algorithm for the L2/L3_SM_P/P_E products from 2015-2021. **Based on improved soil moisture retrieval performance in some agricultural areas and consistent overall performance, the DCA algorithm is now the SMAP baseline soil moisture retrieval algorithm.** Default pointers in the output products have been changed accordingly (Appendix 3). Note that soil moisture retrievals from all three algorithms – SCA-H, SCA-V, and DCA – are still provided in the L2/3_SM_P/P_E output products. Throughout the rest of the SMAP mission, the choice of the operational algorithm of the product will be evaluated on a regular basis as analyses of new observations and Cal/Val data become available or if significant improvements can be achieved by algorithm modifications.

5.1 Single Channel Algorithm V-pol (SCA-V)

In the SCA-V algorithm, the vertically polarized TB observations are converted to emissivity using a surrogate for the physical temperature of the emitting layer. The derived emissivity is corrected for vegetation and surface roughness to obtain the soil emissivity. The Fresnel equation is then used to determine the dielectric constant from the soil emissivity. Finally, a dielectric mixing model is used to solve for the soil moisture given knowledge of the soil texture. [Note: The software code includes the option of using different dielectric models. Currently, the software switch is set to the Mironov model [15]]. Analytically, SCA-V attempts to solve for one unknown variable (soil moisture) from one equation that relates the vertically polarized TB to soil moisture. Vegetation information is provided by an 11-year climatological data base of global NDVI and a table of *tau-omega* parameters based on land cover.

5.2 Single Channel Algorithm H-pol (SCA-H)

The SCA-H is similar to SCA-V in that the horizontally polarized TB observations are converted to emissivity using a surrogate for the physical temperature of the emitting layer. The derived emissivity is corrected for vegetation and surface roughness to obtain the soil emissivity. The Fresnel equation is then used to determine the dielectric constant. Finally, a dielectric mixing model is used to obtain the soil

moisture given knowledge of the soil texture. Analytically, SCA-H attempts to solve for one unknown variable (soil moisture) from one equation that relates the horizontally polarized TB to soil moisture. Vegetation information is provided by an 11-year climatological data base of global NDVI and a table of *tau-omega* parameters based on land cover.

5.3 Dual Channel Algorithm (DCA)

The Dual Channel Algorithm (DCA) uses both H-polarized and V-polarized T_B observations to simultaneously retrieve soil moisture and vegetation optical depth by minimizing the cost function *F*

$$F(sm,\tau) = \left(TB_V^{obs} - TB_V^{modeled}\right)^2 + \left(TB_H^{obs} - TB_H^{modeled}\right)^2 + \lambda^2(\tau - \tau^*)^2,$$

where $TB_V^{modeled}$ and $TB_H^{modeled}$ are the brightness temperatures from the *tau-omega* model. The regularization term $\lambda^2(\tau - \tau^*)^2$ was applied to reduce the temporal and spatial noise caused by the nature of the cost function. τ^* is the initial guess for the unknown vegetation optical depth derived from the same 11-year NDVI vegetation climatology used by SCA-V (Section 5.1) and SCA-H (Section 5.2).

Similar to the SCA, some estimates of model parameters (e.g., surface temperature, surface roughness, and vegetation single scattering albedo) must be provided using ancillary datasets in the inversion process. In contrast with the SCA, the polarization mixing factor is assumed to be linearly related to the roughness parameter *h* as in Q = 0.1771 *h* [19]. *h* is provided to the algorithm through a pre-computed static ancillary file with global values of h over the 3 km EASE 2.0 grid. In addition to these differences, DCA uses different values than SCA for the vegetation single scattering albedo.

5.4 Effective Temperature Methodology

Postlaunch, dynamic surface temperature forecast information is routinely ingested by SMAP from the GMAO GEOS-FP model and processed as an ancillary data input as part of the operational processing of the SMAP passive soil moisture product. The effective surface temperature (T_{eff}) is a critical parameter in passive soil moisture retrieval but is not to be confused with an actual physical temperature measured at a single depth. New to the End-of-Prime-Mission data release in 2018 was an improved depth correction scheme for the effective soil temperature, with the parameterization in this scheme reexamined for the R17 release in 2020. This effective temperature methodology developed for the R17 data release in August 2020 is still in use for the R18 data release of October 2021 (Appendix 1).

At L-band frequency, the contributing soil depth of microwave emission may be different from the pre-defined discrete soil depths at which the soil temperatures are available from a land surface model. The resulting discrepancy can contribute to a dry bias of retrieved soil moisture (i.e., retrieval lower than *in situ* soil moisture) if the model-based effective soil temperature is colder than the soil temperature "seen" by the radiometer. Conversely, wet bias of retrieved soil moisture will occur if the model-based effective soil temperature is warmer than the soil temperature "seen" by the radiometer. Since the contributing soil depth of microwave emission varies with soil moisture, the corresponding depth correction scheme for the effective soil temperature must account for soil moisture variability for brightness temperature observations acquired between AM (descending overpasses) and PM (ascending passes). To achieve this objective, the following modified formulation of the Choudhury model [16] has been found to result in good agreement between the *in situ* soil moisture data and the retrieved L2SMP and L2SMP_E soil moisture:

$$T_{eff} = K \times [\, T_{soil2} + C\,(T_{soil1} - T_{soil2})\,]$$

where C = 0.246 for AM soil moisture retrieval and 1.000 for PM soil moisture retrieval, and K = 1.007. K is a factor included to address an observed bias between ancillary modeled soil temperature and measured *in situ* temperature at core validation sites and sparse network stations. T_{soil1} refers to the average soil temperature for the first soil layer (5-15 cm) and T_{soil2} refers to the average soil temperature for the second soil layer (15-35 cm) of the GMAO GEOS-FP land surface model. Additional information on T_{eff} can be found in Appendix 1 of this report.

5.5 General Flag Usage

As with any satellite retrieval data product, proper data usage is encouraged. The following two simple practices are recommended for using SMAP soil moisture retrievals with maximum scientific benefits:

- Use the **retrieval_qual_flag** field to identify retrievals in the **soil_moisture** field estimated to be of recommended quality. A **retrieval_qual_flag** value of either 0 or 8 indicates high-quality retrievals. Proper use of the **retrieval_qual_flag** field is an effective way to ensure that only retrievals of recommended quality will be used in data analyses.

- For further investigation, use the **surface_flag** field and the associated definition described in the User Guide to determine why the **retrieval_qual_flag** field did not report recommended quality at a given grid cell.

5.6 Frozen Soil Flag

At the start of the SMAP mission, the intention was to set the SMAP frozen soil flag during internal SDS processing based on either the radar ground flag (see L3_FT_A ATBD) or on the GMAO-based Teff. After the failure of the SMAP radar on July 7, 2015, a procedure was developed to replace the radar-based frozen soil flag with one generated from the SMAP radiometer data using normalized polarization ratios (see L3_FT_P ATBD). However, due to uncertainties in reference freeze/thaw conditions, subgrid scale heterogeneity, and polarization changes due to changing vegetation and soil moisture conditions during the growing season at low-to-mid latitudes, the L3_FT_P frozen flag can result in false F/T indications. For these reasons, currently (at the time of the R18 data release in 2021) the frozen soil area fraction is still based on the temperature information from the GMAO GEOS-FP model used by the SMAP operational processor (byte value of 0 or 8 in the Retrieval Quality Flag indicates a good non-frozen retrieval).

6 METHODOLOGIES USED FOR L2SMP/L2SMP_E CAL/VAL

Validation is critical for accurate and credible product usage, and must be based on quantitative estimates of uncertainty. For satellite-based retrievals, validation should include direct comparison with independent correlative measurements. The assessment of uncertainty must also be conducted and presented to the community in normally used metrics in order to facilitate acceptance and implementation.

During mission definition and development, the SMAP Science Team and Cal/Val Working Group identified the metrics and methodologies that would be used for L2-L4 product assessment. These metrics and methodologies were vetted in community Cal/Val workshops and tested in SMAP pre-launch Cal/Val rehearsal campaigns. The methodologies identified and their general roles are:

- Core Validation Sites (CVS): Accurate estimates of products at matching scales for a limited set of conditions
- Sparse Networks: One point in the grid cell for a wide range of conditions
- Satellite Products: Estimates over a very wide range of conditions at matching scales
- Model Products: Estimates over a very wide range of conditions at matching scales
- Field Campaigns: Detailed estimates for a very limited set of conditions

In the case of the L2SMP/L2SMP_E data products, all of these methodologies can contribute to product assessment and improvement.

6.1 Validation Grid (VG)

In all previous SMAP data releases prior to 2020, the accuracy of retrieved soil moisture from the SMAP L2SMP 36-km standard product was assessed using a validation grid (VG) approach. In this approach, the 36-km EASE 2.0 grid was shifted at each CVS by 3-km increments in order to cover the maximum number of *in situ* measurement points possible at the given CVS before performance statistics were generated comparing SMAP soil moisture to *in situ* soil moisture (see the VG description below). The VG shifting was also done at each CVS to minimize the presence of non-representative areas and to maintain homogeneity across the grid cell to the extent possible. However, in the R17 data release of August, 2020, the SMAP project decided to proceed with the L2SMP assessment going forward using the same procedure as used for the L2SMP_E assessment – i.e., without using the shifted VG grid for the 36-km product. Although this non-VG assessment could introduce some additional uncertainty, past assessment reports [1, 20, 25, 27] have documented the close agreement in performance metrics between L2SMP at 36-km posting with VG processing and L2SMP_E at 9-km posting. The non-VG assessment approach was also implemented in the R18 data release of October, 2021. With the non-VG L2SMP comparisons with *in situ* data from both CVS and sparse networks given in Sec. 8.1 and the L2SMP_E assessments given in Sec. 8.2, it is felt that the community has sufficient information to judge the quality and accuracy of the L2SMP retrieved soil moisture.

Description of Validation Grid approach used for L2SMP assessments in previous data releases (retained here for information only; VG approach not implemented in the R18 release):

The scanning radiometer on SMAP provides elliptical footprint observations across the scan. The orientation of this ellipse varies across the swath, and on successive passes a point on the ground might be observed with very different azimuth angles. A standard assumption in using radiometer observations is that the signal is dominated by the energy originating within the 3 dB (half-power) footprint (ellipse). The validity of this contributing area assumption will depend upon the heterogeneity of the landscape.

A major decision was made for SMAP to resample the radiometer data to an Earth-fixed grid at a resolution of 36 km. This was to facilitate temporal analyses and the disaggregation algorithm planned for

the AP (active/passive) product. It ignores azimuth orientation and some contribution beyond the 3 dB footprints mentioned above, although the SMAP L1B_TB data do include a sidelobe correction. An important point is that TBs on the Earth-fixed 36-km grid are used in the retrieval of soil moisture, and it is the soil moisture for these 36-km grid cells that must be validated and improved.

The standard SMAP processor provides L2 surface (0-5 cm) soil moisture using only the radiometer (passive) data posted on a 36-km EASE2 Grid. The standard SMAP grid was established without any consideration of where the CVS might be located. In addition, the CVS were established in most cases to satisfy other (non-SMAP) objectives of the Cal/Val Partners. One of the criteria for categorizing a site as a CVS is that the number of individual *in situ* stations (N) within the site is large (target is $N \geq 9$ for 36 km). It was observed when examining the distribution of points at a site that in many cases only a few points fell in any specific standard 36-km grid cell. Therefore, it was decided that special SMAP validation grids (VGs) would be established for validation assessment that would be tied to the existing SMAP 3-km standard grid but would allow the shifting of the 36-km grids at a site to fully exploit N being as large as possible (i.e. the validation grid would be centered over the collection of *in situ* points at a given CVS to the extent possible). The approach used for validation grid processing is illustrated in Figure 6.1.

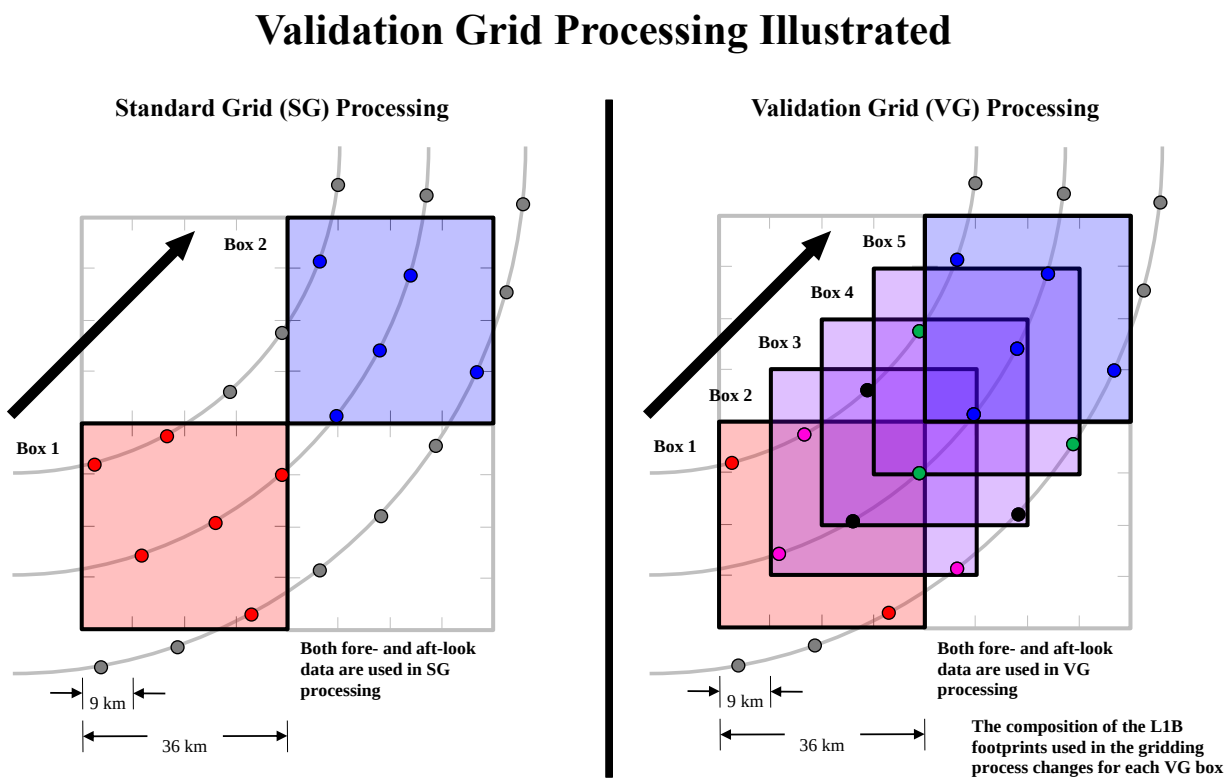


Figure 6.1. Illustration of validation grid processing. The EASE2 grid boxes are shifted by 3-km increments (although 9-km shifts are shown for clarity) to allow a better geographical match with the *in situ* validation sites.

Computationally the L2 and L3 VG products are the same as the standard product. The selection of the VGs for each site was done by members of the SMAP Algorithm Development Team and Science Team prelaunch. As noted, the 3-km grid does not change. The selection of the VGs also considered avoiding or minimizing the effects of land features that were not representative of the sampled domain or were known problems in retrieval (e.g., non-representative terrain, large water bodies, etc.).

7 Summary of Refinements in L2SMP Version 8 and L2SMP_E Version 5

- *Expanded Assessment Period*: For the previous validated data release report, the analysis time period was April 1, 2015 – March 31, 2020 (60 months). The start date in 2015 was based on when the radiometer data were judged to be stable following instrument start-up operations. The end date was based upon the closing date of the Version 7/4 release report. The current assessment report expands the time period from April 1, 2015 through March 31, 2021, which provides a 6-year assessment.

- *Addition of North Polar Grid for L2/3_SM_P_E retrieved soil moisture:* L2/3_SM_P_E soil moisture products are now output on both the global EASE2 and the North Polar EASE2 grids. The L2/3_SM_P_E data fields on global and polar grids are organized in two separate HDF5 data groups called 'Soil_Moisture_Retrieval_Data' and 'Soil_Moisture_Retrieval_Data_Polar', respectively, in the same granule. Each group contains the same data fields and their associated definitions. The polar grid projection offers a more uniform spatial sampling at high latitudes (less distortion than the global grid at these high latitudes), which will facilitate studies with soil moisture in the boreal and arctic regions.

- *New output fields added:* Organic content and sand fraction fields have now been added to the L2 output products.

- *Brightness temperature unmixing near water:* A correction has been applied to unmix surface-corrected T_B in the L1C_TB data used as input to the L2/L3_SM_P/P_E soil moisture retrieval algorithms, mainly affecting areas around coastal areas and surrounding inland open water bodies (see Appendix 2 for more detailed information).

- *New algorithm baseline:* SCA-V was the original postlaunch baseline algorithm for the L2/L3_SM_P/P_E products from 2015-2021. **Based on improved soil moisture retrieval performance in some agricultural areas, the DCA algorithm is now the SMAP baseline soil moisture retrieval algorithm.** Default pointers in the output products have been changed accordingly (Appendix 3). Note that soil moisture retrievals from all three algorithms – SCA-H, SCA-V, and DCA – are still provided in the L2/3_SM_P/P_E output products.

8 ASSESSMENTS

In this section several assessments and intercomparisons are presented. The standard L2SMP AM and PM (Version 8) and L2SMP_E AM and PM (Version 5) are examined for the expanded time period. Changes from the previous assessment (L2SMP Version 7 and L2SMP_E Version 4 [27]) will be noted if they occur. These assessments utilize CVS, sparse network, and SMOS comparisons.

8.1 L2SMP

8.1.1 Core Validation Sites

The primary validation for the L2SMP soil moisture is a comparison of retrievals at 36 km with ground-based observations that have been verified as providing a spatial average of soil moisture at the same scale, referred to as core validation sites (CVS) in the SMAP Calibration/Validation Plan [6].

In situ data are critical in the assessment of the SMAP products. These comparisons provide error estimates and a basis for modifying algorithms and/or parameters. A robust analysis will require many sites representing diverse conditions. However, there are relatively few sites that can provide the type and quality of data required. SMAP established a Cal/Val Partners Program in order to foster cooperation with these sites and to encourage the enhancement of these resources to better support SMAP Cal/Val. The current set of sites that provide data for L2SMP are listed in Table 8.1.

Not all of the sites in Table 8.1 have reached a level of maturity that would support their use as CVS. Prior to initiating the 2015 beta-release assessments, the L2SMP and Cal/Val Teams reviewed the status of all sites to determine which sites were ready to be designated as CVS. This process is repeated periodically during the mission, with the addition of new screening procedures for *in situ* data as well as changes in upscaling at some CVS. The basic process is as follows:

- Develop and implement the validation grid
- Assess the site for conditions that would introduce uncertainty
- Determine if the number of points is large enough to provide reliable estimates
- Assess the geographic distribution of the *in situ* points
- Determine if the *in situ* instrumentation has been either (1) widely used and known to be well-calibrated or (2) calibrated for the specific site in question
- Perform quality assessment of each point in the network
- Establish a scaling function (default function is a linear average of all stations)
- Conduct pre-launch assessment using surrogate data appropriate for the SMAP L2SMP product (i.e. SMOS soil moisture)
- Review any supplemental studies that have been performed to verify that the network represents the SMAP product over the grid domain

The current CVS for the L2SMP product are marked with an asterisk in Table 8.1. A total of 12 CVS were used in this assessment. Each of these should have at least 9 points (ground *in situ* measurement stations); however, exceptions are made to allow fewer *in situ* stations if the site has a well-established scaling and calibration function. The status of candidate sites will continue to be reviewed periodically to determine if they should be classified as CVS and used in future assessments. Note that the table includes comments on sites that are used for some of the L2SMP_E analyses discussed later.

The *in situ* data downloaded from the Cal/Val Partners is run through an automatic quality control (QC) before determining the upscaled soil moisture values for each pixel (grid cell). The QC is implemented largely following the approach presented in [17]. The procedure includes checks for missing data, out of control values, spikes, sudden drops, and physical temperature limits. Additionally, the physical temperature is checked to be above 4°C because many sensors experience change in behavior at colder temperatures. In several cases the sites include stations that do not perform as expected, or at all, during the comparison period. These stations are removed from consideration altogether, and a new configuration is set for the site accounting for only the stations that produce a reasonable amount of data over the comparison period. Consequently, the upscaling functions for these sites are also based on the remaining set of stations.

The key tool used in L2SMP (and L2SMP_E) CVS analyses is illustrated by Figure 8.1. These charts are updated as changes are made to L1 data, L2 algorithms, or in preparation for periodic reviews with Cal/Val Partners. It includes a time series plot of *in situ* and retrieved soil moisture as well as flags that were triggered on a given day, an XY scatter plot of SMAP retrieved soil moisture compared to the average *in situ* soil moisture, and the quantitative statistical metrics. It also shows the CVS individual station distribution. When the *in situ* values are marked with a magenta color instead of red, it means that the *in situ* quality flag is raised. Several alternative algorithms and the SMOS soil moisture product are also displayed (SMOS L2 v650 was used). These plots are carefully reviewed and discussed by the L2SMP Team on a periodic basis. Systematic differences and anomalies are identified for further investigation.

All sites are then compiled to summarize the metrics and compute the overall performance. Tables 8.2 and 8.3 present the overall results for the current L2SMP Version 8 validated data sets. The combined scatter plots associated with these results are shown in Figure 8.2. These metrics and plots include the removal of questionable-quality and retrieval-flagged data.

The key results for this assessment are summarized in the SMAP Average results row in Table 8.2 for the descending (AM) orbits. First, all algorithms have about the same ubRMSE, differing by 0.007 m^3/m^3, with SCA-V and DCA exceeding the SMAP mission goal of 0.040 m^3/m^3. Second, the correlations for SCA-V and DCA are also very similar. For both of these metrics, the SCA-V and DCA algorithms show the best performance, with both being superior to SCA-H. SCA-V and DCA also have similar bias values which are lower than the SCA-H bias. SCA-V has an ubRMSE of 0.036 m^3/m^3, a bias of 0.010, and an R of 0.819 for AM overpasses. DCA has similar performance with an ubRMSE of 0.036 m^3/m^3, a bias of 0.012, and an R of 0.816. Although the SMAP L1 mission requirements for soil moisture retrieval accuracy (0.040 m^3/m^3 or better) strictly apply only to the descending (6 AM) orbits, Table 8.3 presents equally good results for the ascending (6 PM) orbits: SCA-V has an ubRMSE of 0.037 m^3/m^3, a bias of 0.007, and an R of 0.822, and DCA has an ubRMSE of 0.035 m^3/m^3, a bias of 0.005, and an R of 0.801. Given these results, SMAP users can have confidence in using both AM and PM data to increase the temporal frequency of SMAP soil moisture without sacrificing retrieval accuracy.

Based upon the metrics and considerations discussed, DCA will become the new operational baseline algorithm for this release (Version 8 L2SMP and Version 5 L2SMP_E). As a longer period of observations builds and additional CVS are added, the evaluations will be repeated on a periodic basis.

For guidance in expected performance, the SMOS soil moisture products for each site over the same time period were analyzed. Summary statistics are included in Tables 8.2 and 8.3. For the CVS analyzed here, SMAP algorithms outperform SMOS for all metrics.

Also shown in Tables 8.2 and 8.3 are the metric averages from previous L2SMP version assessments. As noted previously, in addition to the product changes, there is also a longer period of record associated with Version 8. Comparing across the versions shown, DCA shows continual improvement while SCA-V has stable performance in meeting mission requirements.

Table 8.1. SMAP Cal/Val Partner Sites Providing L2SMP Validation Data

Site Name	Site PI	Area	Climate regime	IGBP Land Cover
Walnut Gulch*	M. Cosh	USA (Arizona)	Arid	Shrub open
Reynolds Creek*	M. Cosh	USA (Idaho)	Arid	Grasslands
Fort Cobb*	M. Cosh	USA (Oklahoma)	Temperate	Grasslands
Little Washita*	M. Cosh	USA (Oklahoma)	Temperate	Grasslands
South Fork*	M. Cosh	USA (Iowa)	Cold	Croplands
Little River*	M. Cosh	USA (Georgia)	Temperate	Cropland/natural mosaic
TxSON*	T. Bongiovanni	USA (Texas)	Temperate	Grasslands
Millbrook	M. Temimi	USA (New York)	Cold	Deciduous broadleaf
Kenaston*	A. Berg	Canada	Cold	Croplands
Carman*	H. McNairn	Canada	Cold	Croplands
Monte Buey*	M. Thibeault	Argentina	Arid	Croplands
Bell Ville	M. Thibeault	Argentina	Arid	Croplands
REMEDHUS*	J. Martinez	Spain	Temperate	Croplands
Valencia	E. Lopez-Baeza	Spain	Arid	Woody Savannas
Twente	Z. Su	Netherlands	Cold	Cropland/natural mosaic
HOBE	F. Udall	Denmark	Temperate	Croplands
Kuwait	H. Jassar	Kuwait	Temperate	Barren/sparse
Niger	T. Pellarin	Niger	Arid	Grasslands
Benin	T. Pellarin	Benin	Arid	Savannas
Naqu	Z. Su	Tibet	Polar	Grasslands
Maqu	Z. Su	Tibet	Cold	Grasslands
Ngari	Z. Su	Tibet	Arid	Barren/sparse
MAHASRI	J. Asanuma	Mongolia	Cold	Grasslands
Yanco*	J. Walker	Australia	Arid	Croplands
Kyeamba	J. Walker	Australia	Temperate	Croplands
*=CVS used in both L2SMP and L2SMP_E assessments.				

It should be noted that a small bias should be expected when comparing satellite retrievals to *in situ* soil moisture sensors during drying conditions. Satellite L-band microwave signals respond to a surface layer of a depth that varies with soil moisture (this depth is taken to be ~0-5 cm for average soils under average conditions). The *in situ* measurement is centered at 5 cm and measures a layer from ~ 3 to 7 cm. For some surface conditions and climates, it is expected that the surface will be slightly drier than the layer measured by the *in situ* sensors. For example, Adams et al. [18] reported that a mean difference of 0.018 m^3/m^3 existed between the measurements obtained by inserting a probe vertically from the surface versus horizontally at 5 cm for agricultural fields in Manitoba, Canada. Drier conditions were obtained using the surface measurement and this difference was more pronounced for mid- to dry conditions and minimized during wet conditions.

A review of the individual CVS indicates that several agricultural sites (such as South Fork and Little River) have larger bias values, possibly due to heterogeneous land cover and changing conditions throughout the crop growing season that are not being properly addressed by the current algorithm approach. Efforts are under way to better understand the causes of the errors and to determine if there is anything that can done to mitigate these errors. DCA shows a marked reduction in ubRMSE at the agricultural sites in Carmen and Monte Buey, even though its overall metrics are very similar to SCA-V. SCA-V has slightly better performance in grassland sites compared to DCA.

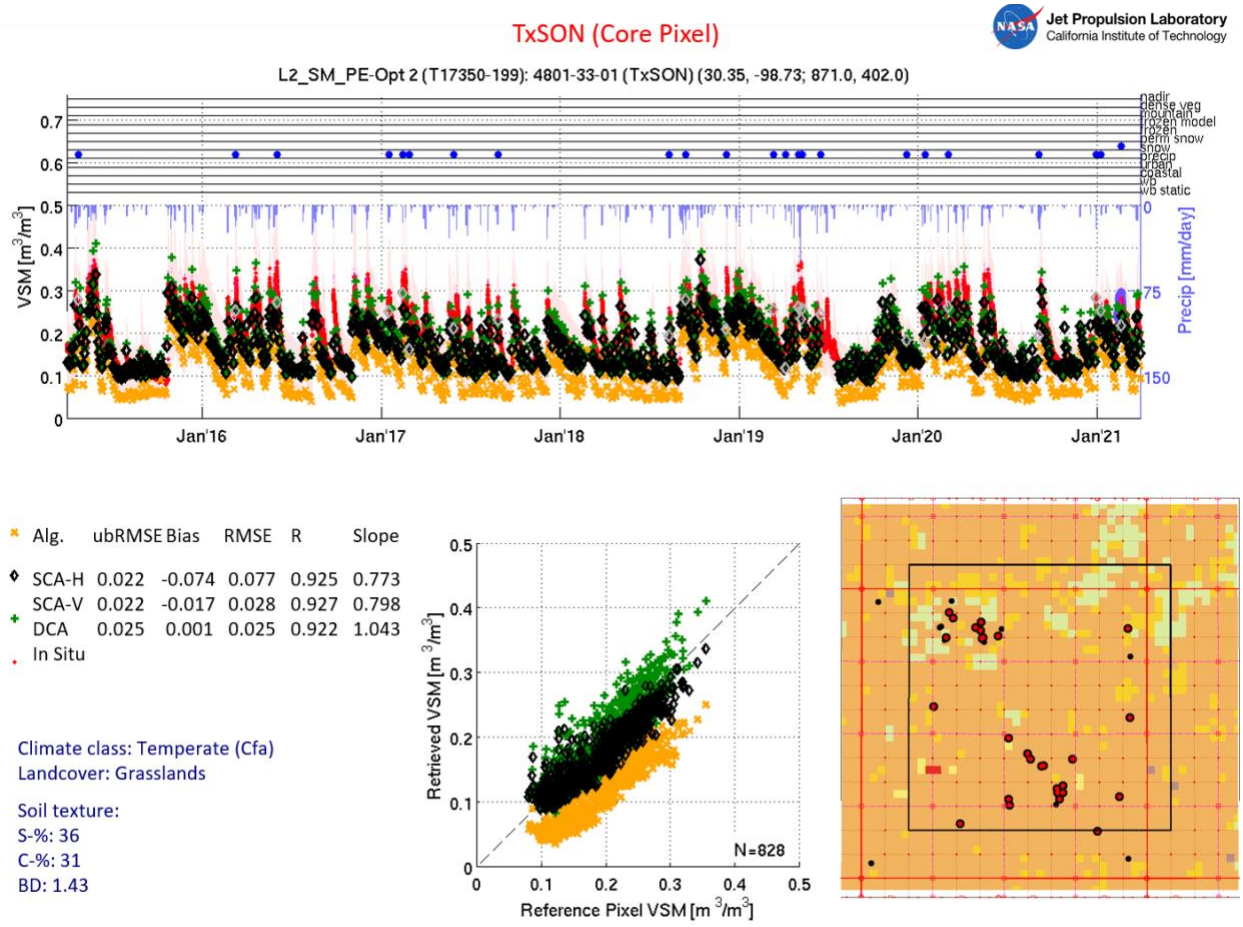


Figure 8.1. L2SMP_E assessment tool report for the TxSON core validation site descending (AM) passes.

Table 8.2. SMAP L2SMP Version 8 CVS Assessment for Descending (AM) Overpasses

CVS	ubRMSE (m^3/m^3)			Bias (m^3/m^3)			RMSE (m^3/m^3)			R			N		
	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA
Reynolds Creek	0.041	0.041	0.043	-0.052	-0.007	0.000	0.067	0.042	0.043	0.638	0.634	0.607	126	126	125
Walnut Gulch	0.024	0.024	0.025	-0.027	0.011	0.016	0.036	0.026	0.030	0.794	0.824	0.788	368	407	384
TxSON	0.021	0.020	0.024	-0.073	-0.016	0.003	0.076	0.026	0.025	0.925	0.929	0.927	619	619	619
Fort Cobb	0.027	0.024	0.029	-0.056	-0.012	-0.004	0.062	0.027	0.029	0.888	0.904	0.901	675	675	675
Little Washita	0.026	0.025	0.031	-0.045	0.000	0.010	0.052	0.025	0.033	0.890	0.900	0.892	694	694	694
South Fork	0.054	0.045	0.043	-0.079	-0.061	-0.052	0.096	0.076	0.067	0.716	0.770	0.758	447	458	372
Little River	0.046	0.036	0.040	0.046	0.096	0.074	0.065	0.102	0.084	0.741	0.753	0.773	640	634	642
Kenaston	0.036	0.031	0.032	-0.009	0.027	0.035	0.037	0.041	0.048	0.787	0.844	0.831	349	349	349
Carman	0.084	0.061	0.047	-0.017	-0.007	-0.013	0.086	0.062	0.049	0.619	0.695	0.747	295	299	301
Monte Buey	0.067	0.048	0.031	-0.021	0.001	-0.011	0.071	0.048	0.033	0.773	0.869	0.856	213	228	230
REMEDHUS	0.038	0.038	0.038	0.014	0.044	0.050	0.041	0.058	0.063	0.759	0.785	0.793	430	442	415
Yanco	0.047	0.043	0.047	0.013	0.044	0.035	0.049	0.061	0.059	0.927	0.926	0.918	196	199	204
Mean Absolute Bias				**0.038**	**0.027**	**0.025**									
SMAP L2SMP Average V8	**0.043**	**0.036**	**0.036**	**-0.026**	**0.010**	**0.012**	**0.061**	**0.050**	**0.047**	**0.788**	**0.819**	**0.816**			
SMAP L2SMP Average V7	**0.043**	**0.037**	**0.036**	**-0.025**	**0.010**	**0.012**	**0.062**	**0.050**	**0.047**	**0.787**	**0.820**	**0.815**			
SMOS L2SM Average V6	**0.052**			**-0.026 (MAB=0.036)**			**0.067**			**0.685**					
SMAP L2SMP Average V5	**0.046**	**0.037**	**0.047**	**-0.028**	**-0.001**	**0.038**	**0.062**	**0.044**	**0.070**	**0.788**	**0.821**	**0.737**			
SMOS L2SM Average V5	**0.053**			**-0.024 (MAB=0.035)**			**0.065**			**0.671**					
SMAP L2SMP Average V4	**0.046**	**0.039**	**0.047**	**-0.037**	**-0.028**	**-0.015**	**0.071**	**0.061**	**0.066**	**0.772**	**0.795**	**0.700**			
SMOS L2SM Average V4	**0.053**			**-0.028**			**0.072**			**0.710**					

Table 8.3. SMAP L2SMP Version 8 CVS Assessment for Ascending (PM) Overpasses

CVS	ubRMSE (m^3/m^3)			Bias (m^3/m^3)			RMSE (m^3/m^3)			R			N		
	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA
Reynolds Creek	0.053	0.053	0.055	-0.059	-0.015	-0.006	0.079	0.055	0.055	0.576	0.575	0.559	188	188	186
Walnut Gulch	0.027	0.027	0.026	-0.040	-0.003	0.000	0.048	0.027	0.026	0.701	0.728	0.668	584	662	602
TxSON	0.019	0.018	0.021	-0.068	-0.015	0.001	0.070	0.024	0.021	0.933	0.938	0.935	697	697	696
Fort Cobb	0.031	0.027	0.030	-0.051	-0.013	-0.009	0.059	0.030	0.031	0.875	0.881	0.867	760	760	757
Little Washita	0.030	0.026	0.028	-0.035	0.002	0.007	0.046	0.026	0.029	0.882	0.899	0.891	745	745	742
South Fork	0.055	0.043	0.039	-0.076	-0.068	-0.064	0.094	0.081	0.076	0.722	0.798	0.767	469	482	359
Little River	0.044	0.036	0.037	0.057	0.101	0.077	0.072	0.107	0.086	0.773	0.766	0.778	505	505	508
Kenaston	0.036	0.031	0.034	-0.002	0.028	0.037	0.036	0.042	0.050	0.800	0.850	0.811	472	472	471
Carman	0.069	0.053	0.042	-0.026	-0.013	-0.016	0.074	0.054	0.045	0.662	0.775	0.752	372	373	360
Monte Buey	0.065	0.047	0.030	0.013	0.017	-0.014	0.066	0.050	0.033	0.823	0.872	0.800	214	234	225
REMEDHUS	0.038	0.039	0.038	0.001	0.025	0.027	0.038	0.046	0.047	0.800	0.840	0.849	566	639	587
Yanco	0.050	0.042	0.040	0.016	0.042	0.026	0.053	0.059	0.048	0.942	0.942	0.940	218	230	216
Mean Absolute Bias				**0.037**	**0.029**	**0.024**									
SMAP L2SMP Average V8	**0.043**	**0.037**	**0.035**	**-0.022**	**0.007**	**0.005**	**0.061**	**0.050**	**0.046**	**0.791**	**0.822**	**0.801**			
SMAP L2SMP Average V7	**0.043**	**0.037**	**0.036**	**-0.025**	**0.010**	**0.012**	**0.062**	**0.050**	**0.047**	**0.787**	**0.820**	**0.815**			
SMOS L2SM Average V6	**0.057**			**-0.028 (MAB=0.036)**			**0.071**			**0.660**					
SMAP L2SMP Average V5	**0.046**	**0.037**	**0.047**	**-0.028**	**-0.001**	**0.038**	**0.062**	**0.044**	**0.070**	**0.788**	**0.821**	**0.737**			
SMOS L2SM Average V5	**0.053**			**-0.024 (MAB=0.035)**			**0.065**			**0.671**					
SMAP L2SMP Average V4	**0.046**	**0.039**	**0.047**	**-0.037**	**-0.028**	**-0.015**	**0.071**	**0.061**	**0.066**	**0.772**	**0.795**	**0.700**			
SMOS L2SM Average V4	**0.053**			**-0.028**			**0.072**			**0.710**					

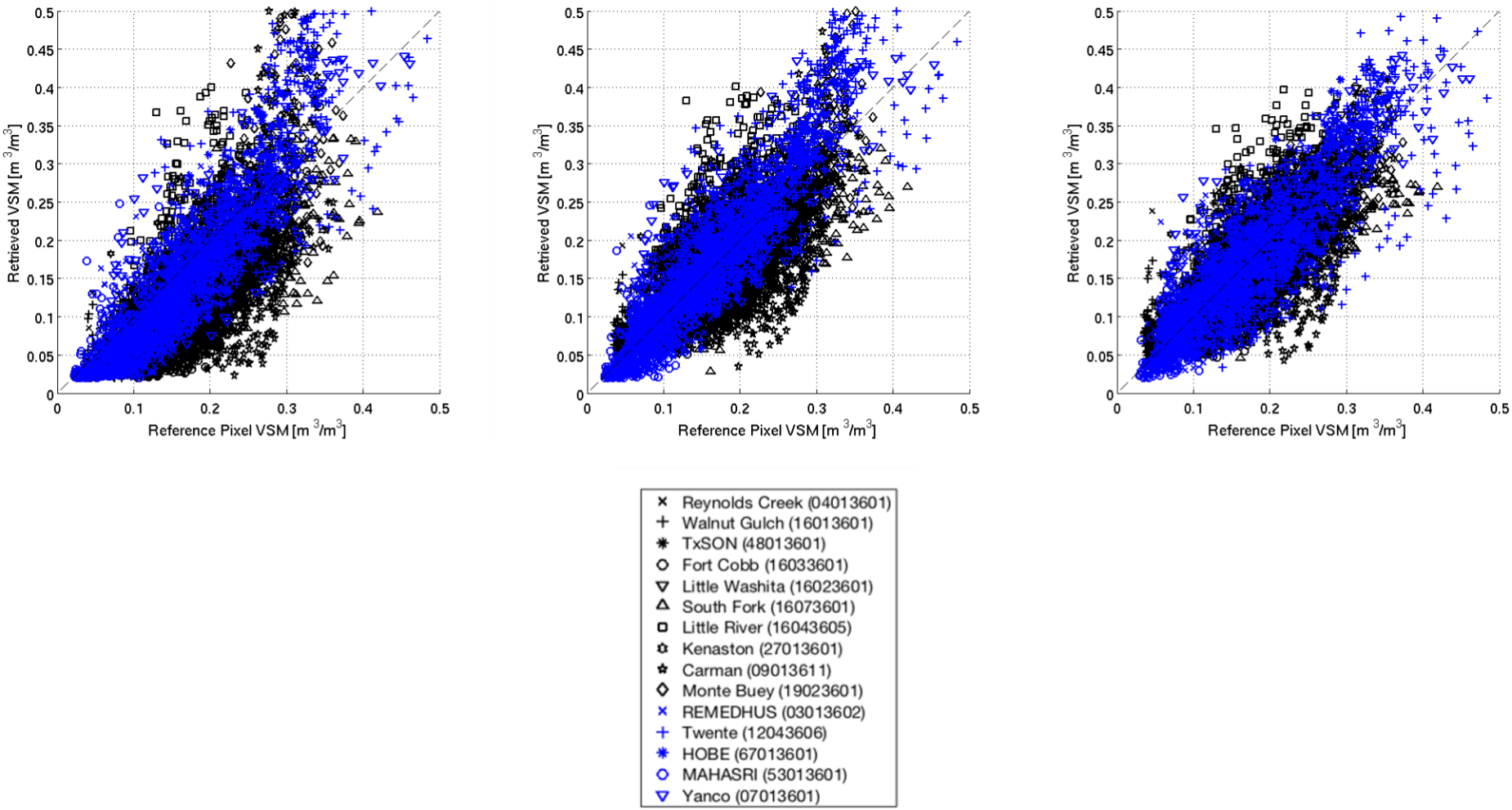


Figure 8.2. Scatterplot of SMAP non-VG L2SMP Version 8 CVS Assessment for Descending (AM) Overpasses (SCA-H left panel, SCA-V middle panel, and DCA right panel).

8.1.2 Sparse Networks

The intensive network CVS validation described above can be complemented by sparse networks as well as by new/emerging types of soil moisture networks. The set of networks being utilized by SMAP are listed in Table 8.4.

The defining feature of these networks is that the measurement density is low, usually resulting in one ground measurement point per SMAP footprint. These observations cannot be used for validation without addressing two issues: verifying that they provide a reliable estimate of the 0-5 cm surface soil moisture layer and that the one measurement point is representative of conditions across the entire SMAP footprint.

SMAP has been evaluating methodologies for upscaling data from these networks to SMAP footprint resolutions. A key element of the upscaling approach is Triple Colocation that combines the *in situ* data and SMAP soil moisture product with another independent source of soil moisture, likely to be a model-based product [5].

Although limited by upscaling, sparse networks do offer many sites in different environments and are typically operational with very low latency. They are very useful as a supplement to the limited number of CVS.

Table 8.4. Sparse Networks Providing L2SMP and L2SMP_E Validation Data

Network Name	PI/Contact	Area	No. of Sites (L2SMP)	No. of Sites (L2SMP_E)
NOAA Climate Reference Network (CRN)	M. Palecki	USA	60	56
USDA NRCS Soil Climate Analysis Network (SCAN)	M. Cosh	USA	101	100
GPS	E. Small	Western USA	80	77
COSMOS	M. Zreda	Mostly USA	30	32
SMOSMania	J. Calvet	Southern France	10	11
Pampas	M. Thibeault	Argentina	16	14
Oklahoma Mesonet	-	Oklahoma, USA	94	96
Mongolian Grasslands (MAHASRI)	J. Asanuma	Mongolia	13	13

The sparse network metrics are summarized in Table 8.5 and 8.6. Because of the larger number of sites, it is possible to also examine the results based upon the IGBP land cover classification used by SMAP. For these comparisons the SMOS metrics are included for each category. The reliability of the analyses based upon these classes will depend upon the number of sites available (N).

Overall, the relative performance of the algorithms is similar to that obtained from the CVS -- SCA-V and DCA again have the best metrics, with SCA-V exhibiting an ubRMSE of 0.051 m^3/m^3, bias of -0.002 m^3/m^3 and correlation of 0.619 for AM orbits, and an ubRMSE of 0.051 m^3/m^3, bias of 0.001 m^3/m^3 and correlation of 0.606 for the PM orbits. DCA shows an ubRMSE of 0.052 m^3/m^3, bias of 0.004 m^3/m^3 and correlation of 0.623 for AM orbits, and an ubRMSE of 0.051 m^3/m^3, bias of 0.005 m^3/m^3 and correlation of 0.611 for the PM orbits. Compared to the CVS results, the sparse network values are higher for ubRMSE and lower for R, which is expected due to the significant change in scale between a point and the grid product. When comparing Version 8 to previous data versions, SCA-V and DCA demonstrate stable or improved performance. Considering the many caveats that must be considered in making sparse network comparisons, the algorithm performance is quite good. This result provides additional confidence in the previous conclusions based on the CVS.

Interpreting the results based on land cover is more complex. There are no clear patterns associated with broader vegetation types, especially since some of the land cover types have a limited number of sites. The ubRMSE values for DCA are all between 0.022 and 0.074 m^3/m^3. Grasslands had larger bias values, which needs to be investigated. Forest results are based on very limited sites and should not be generalized.

SMOS (Level 2 UDP) metrics are also included in Tables 8.5 and 8.6 (in blue) as supporting information. It should be noted that while SMOS retrievals are based on a different land cover classification scheme (ECOCLIMAP), this does not have any impact on the comparisons shown, which compares the soil moisture retrievals to the *in situ* observations for the points that fall into these categories. Overall, the SMOS products are showing a higher bias and ubRMSE and lower correlation than the SMAP retrievals.

Table 8.5. SMAP L2SMP Version 8 Sparse Network Assessment for Descending (AM) Overpasses

IGBP Class	ubRMSD (m^3/m^3)				Bias (m^3/m^3)				RMSD (m^3/m^3)				R				N
	SCA-H	**SCA-V**	**DCA**	**SMOS**	**SCA-H**	**SCA-V**	**DCA**	**SMOS**	**SCA-H**	**SCA-V**	**DCA**	**SMOS**	**SCA-H**	**SCA-V**	**DCA**	**SMOS**	
Evergreen needleleaf forest	0.072	0.071	0.074	0.080	-0.069	-0.017	0.005	-0.016	0.122	0.111	0.117	0.097	0.500	0.530	0.530	0.043	8
Evergreen broadleaf forest																	
Deciduous needleleaf forest																	
Deciduous broadleaf forest				0.119				-0.135				0.180				0.147	1
Mixed forest	0.059	0.057	0.058	0.062	-0.067	-0.009	0.000	-0.065	0.100	0.081	0.071	0.090	0.650	0.650	0.660	0.662	4
Closed shrublands																	
Open shrublands	0.040	0.040	0.041	0.055	-0.042	0.003	0.019	-0.011	0.064	0.055	0.060	0.065	0.490	0.520	0.510	0.509	49
Woody savannas	0.058	0.054	0.054	0.093	-0.018	0.033	0.028	-0.051	0.108	0.101	0.095	0.132	0.660	0.690	0.700	0.476	30
Savannas	0.040	0.037	0.040	0.052	-0.030	0.006	0.005	-0.021	0.066	0.056	0.061	0.063	0.820	0.820	0.820	0.796	3
Grasslands	0.052	0.051	0.052	0.063	-0.075	-0.031	-0.019	-0.043	0.097	0.075	0.073	0.088	0.650	0.670	0.670	0.613	258
Permanent wetlands																	
Croplands	0.078	0.069	0.067	0.085	-0.050	-0.022	-0.018	-0.037	0.127	0.109	0.105	0.124	0.490	0.540	0.550	0.544	83
Urban and built-up																	
Crop/Natural vegetation mosaic	0.070	0.062	0.062	0.095	-0.037	0.010	-0.009	-0.050	0.101	0.088	0.088	0.140	0.610	0.670	0.690	0.425	29
Snow and ice																	
Barren/Sparse	0.020	0.021	0.022	0.029	-0.021	0.012	0.021	-0.011	0.035	0.034	0.039	0.040	0.480	0.480	0.480	0.617	9
SMAP L2SMP Average V8	**0.054**	**0.051**	**0.052**		**-0.045**	**-0.002**	**0.004**		**0.091**	**0.079**	**0.079**		**0.594**	**0.619**	**0.623**		
SMAP L2SMP Average V7	**0.052**	**0.048**	**0.047**		**-0.046**	**0.000**	**0.01**		**0.085**	**0.069**	**0.073**		**0.661**	**0.684**	**0.693**		
SMOS L2SM Average V6				**0.069**				**-0.039**				**0.097**				**0.567**	
SMOS L2SM Average V5				**0.073**				**-0.042**				**0.099**				**0.550**	
SMOS V6 data are used in the land cover rows.																	

Table 8.6. SMAP L2SMP Version 8 Sparse Network Assessment for Ascending (PM) Overpasses

IGBP Class	ubRMSD (m^3/m^3)				Bias (m^3/m^3)				RMSD (m^3/m^3)				R				N
	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	
Evergreen needleleaf forest	0.071	0.070	0.073	0.080	-0.066	-0.014	0.008	-0.005	0.120	0.110	0.115	0.107	0.510	0.530	0.530	0.089	8
Evergreen broadleaf forest																	
Deciduous needleleaf forest																	
Deciduous broadleaf forest				0.124				-0.083				0.149				0.240	1
Mixed forest	0.057	0.054	0.053	0.062	-0.056	-0.003	0.007	-0.050	0.099	0.078	0.065	0.079	0.690	0.710	0.720	0.691	4
Closed shrublands																	
Open shrublands	0.042	0.042	0.043	0.059	-0.047	0.000	0.015	-0.009	0.068	0.057	0.062	0.072	0.450	0.450	0.450	0.426	48
Woody savannas	0.058	0.053	0.052	0.094	-0.004	0.039	0.030	-0.038	0.107	0.102	0.094	0.122	0.660	0.680	0.700	0.518	30
Savannas	0.040	0.038	0.040	0.057	-0.024	0.010	0.008	-0.020	0.064	0.060	0.064	0.076	0.830	0.820	0.820	0.713	3
Grasslands	0.052	0.050	0.051	0.062	-0.072	-0.031	-0.022	-0.040	0.094	0.074	0.074	0.088	0.640	0.650	0.650	0.616	259
Permanent wetlands																	
Croplands	0.077	0.067	0.064	0.083	-0.041	-0.022	-0.024	-0.040	0.123	0.107	0.102	0.117	0.510	0.550	0.560	0.553	82
Urban and built-up																	
Crop/Natural vegetation mosaic	0.069	0.061	0.060	0.093	-0.014	0.021	-0.002	-0.044	0.094	0.089	0.087	0.130	0.610	0.670	0.680	0.451	29
Snow and ice																	
Barren/Sparse	0.020	0.021	0.022	0.037	-0.023	0.012	0.021	-0.012	0.035	0.033	0.037	0.050	0.410	0.390	0.390	0.320	9
SMAP L2SMP Average V8	**0.054**	**0.051**	**0.051**		**-0.039**	**0.001**	**0.005**		**0.089**	**0.079**	**0.078**		**0.590**	**0.606**	**0.611**		
SMAP L2SMP Average V7	**0.052**	**0.049**	**0.049**		**-0.040**	**0.003**	**0.011**		**0.083**	**0.07**	**0.074**		**0.628**	**0.643**	**0.65**		
SMOS L2SM Average V6				**0.069**				**-0.036**				**0.095**				**0.561**	
SMOS L2SM Average V5				**0.073**				**-0.042**				**0.549**				**0.550**	
SMOS V6 data are used in the land cover rows.																	

8.2 L2SMP_E

8.2.1 Core Validation Sites

The new L2SMP_E Version 5 is assessed using the same approach as that employed for L2SMP. The major difference between L2SMP_E and L2SMP is that this product is assessed using a different set of CVS. Because it is possible to now provide a retrieval for every SMAP 9-km grid cell where feasible, the need for using the validation grid (as used in releases for L2SMP prior to 2020) is not expected to be as important an issue in performing validation. It should be noted that the validation grid allowed centering the retrieval on any 3-km grid, whereas the L2SMP_E retrieval process can only be centered on a 9-km grid. Thus, the ability to match the *in situ* network to the grid may be more restrictive for L2SMP_E. Each available CVS was reviewed to identify the 9-km grid cell that satisfied the CVS criteria for the new 33-km contributing domain. Therefore, the mix/weighting of *in situ* stations and grid center will be different between the CVS sets used for the two products.

The CVS results are summarized in Tables 8.7 and 8.8 for the AM and PM overpasses, respectively. SCA-V and DCA show comparable performance, and their ubRMSE meet/exceed the SMAP mission requirements. SCA-V has an ubRMSE of 0.038 m^3/m^3, bias of -0.006 m^3/m^3, and correlation of 0.816 for AM orbits, and an ubRMSE of 0.037 m^3/m^3, bias of -0.008 m^3/m^3 and correlation of 0.810 for the PM orbits. DCA shows an ubRMSE of 0.036 m^3/m^3, bias of -0.009 m^3/m^3 and correlation of 0.818 for AM orbits, and an ubRMSE of 0.036 m^3/m^3, bias of -0.013 m^3/m^3 and correlation of 0.792 for the PM orbits. When compared to the L2SMP retrievals, the differences in the metrics are negligible. These results indicate that the L2SMP_E products can be used in place of L2SMP without loss of accuracy.

New to the SMAP R18 data release of October 2021, L2SMP_E retrieved soil moisture values are also output on the North Polar EASE2 grid as well as the global EASE2 grid. The North Polar grid metrics appear in Tables 8.9 and 8.10. The use of the North Polar grid eliminates the Southern Hemisphere CVS sites of Monte Buey and Yanco from this analysis, but otherwise the performance metrics are very similar to those obtained using the global grid. For both the global and North Polar grids, SMAP overall results are generally better than those obtained from SMOS.

8.2.2 Sparse Networks

The sparse network results are summarized in Tables 8.11 and 8.12 for the AM and PM overpasses, respectively. Comparing the overall metrics for the L2SMP sparse network products to the L2SMP_E sparse network products, the results are very consistent and therefore support the trends observed in the CVS analysis. The sparse network metrics using the North Polar grid are presented in Tables 8.13 and 8.14 and show similar results. Again, overall SMAP metrics are better than SMOS.

Table 8.7. SMAP L2SMP_E Version 5 CVS Assessment for Descending (AM) Overpasses

CVS	ubRMSE (m^3/m^3)			Bias (m^3/m^3)			RMSE (m^3/m^3)			R			N		
	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA
Reynolds Creek	0.041	0.041	0.044	-0.070	-0.022	-0.012	0.081	0.046	0.046	0.639	0.666	0.624	150	174	171
Walnut Gulch	0.025	0.026	0.028	-0.014	0.024	0.029	0.029	0.036	0.041	0.760	0.772	0.686	316	335	301
TxSON	0.022	0.022	0.025	-0.074	-0.017	0.001	0.077	0.028	0.025	0.925	0.927	0.922	828	828	828
Fort Cobb	0.033	0.029	0.032	-0.084	-0.046	-0.042	0.090	0.054	0.053	0.878	0.891	0.886	766	768	752
Little Washita	0.022	0.021	0.027	-0.060	-0.014	-0.002	0.064	0.025	0.027	0.898	0.916	0.915	612	612	612
South Fork	0.052	0.046	0.041	-0.075	-0.054	-0.046	0.091	0.071	0.061	0.702	0.744	0.769	351	358	317
Little River	0.045	0.036	0.040	0.014	0.064	0.039	0.047	0.073	0.056	0.731	0.765	0.785	715	712	715
Kenaston	0.041	0.030	0.031	-0.035	0.000	0.006	0.053	0.030	0.031	0.725	0.784	0.795	265	266	266
Carman	0.085	0.062	0.050	-0.069	-0.062	-0.068	0.109	0.088	0.085	0.614	0.672	0.700	421	427	429
Monte Buey	0.067	0.045	0.033	-0.035	-0.013	-0.021	0.075	0.047	0.039	0.756	0.851	0.849	332	350	354
REMEDHUS	0.039	0.038	0.038	-0.020	0.007	0.010	0.044	0.039	0.039	0.830	0.845	0.843	677	729	690
Twente	0.075	0.062	0.056	0.011	0.034	-0.021	0.076	0.070	0.060	0.828	0.833	0.841	678	705	728
HOBE	0.045	0.036	0.042	0.002	0.001	-0.004	0.045	0.036	0.042	0.750	0.864	0.848	124	124	124
MAHASRI	0.035	0.033	0.025	-0.018	-0.011	-0.023	0.039	0.035	0.034	0.779	0.803	0.886	295	286	132
Yanco	0.042	0.038	0.034	-0.016	0.019	0.018	0.045	0.042	0.039	0.897	0.905	0.923	542	541	529
Mean Absolute Bias				**0.040**	**0.026**	**0.023**									
SMAP L2SMP_E Average V5	**0.045**	**0.038**	**0.036**	**-0.036**	**-0.006**	**-0.009**	**0.064**	**0.048**	**0.045**	**0.781**	**0.816**	**0.818**			
SMOS L2SM Average V6	**0.052**			**-0.026 (MAB=0.036)**			**0.067**			**0.685**					
SMOS L2SM Average V3	**0.053**			**-0.023**			**0.067**			**0.683**					
SMAP L2SMP_E Average V4	**0.044**	**0.037**	**0.036**	**-0.036**	**-0.006**	**-0.009**	**0.064**	**0.048**	**0.045**	**0.782**	**0.817**	**0.817**			
SMOS L2SM Average V2	**0.053**			**-0.022**			**0.067**			**0.665**					

Table 8.8. SMAP L2SMP_E Version 5 CVS Assessment for Ascending (PM) Overpasses

CVS	ubRMSE (m^3/m^3)			Bias (m^3/m^3)			RMSE (m^3/m^3)			R			N		
	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA
Reynolds Creek	0.046	0.044	0.046	-0.074	-0.027	-0.017	0.087	0.051	0.049	0.550	0.629	0.586	207	246	238
Walnut Gulch	0.026	0.025	0.026	-0.026	0.010	0.014	0.036	0.027	0.029	0.702	0.726	0.614	570	668	589
TxSON	0.020	0.020	0.022	-0.068	-0.017	-0.002	0.071	0.026	0.022	0.929	0.934	0.926	897	897	897
Fort Cobb	0.038	0.030	0.030	-0.081	-0.050	-0.051	0.089	0.058	0.059	0.874	0.889	0.868	842	851	792
Little Washita	0.025	0.022	0.026	-0.048	-0.010	-0.003	0.054	0.024	0.026	0.895	0.913	0.901	640	640	637
South Fork	0.052	0.045	0.038	-0.068	-0.057	-0.054	0.086	0.073	0.066	0.732	0.765	0.780	357	367	290
Little River	0.044	0.036	0.038	0.024	0.067	0.042	0.05	0.076	0.057	0.764	0.763	0.779	640	636	640
Kenaston	0.036	0.025	0.030	-0.031	-0.001	0.006	0.047	0.025	0.031	0.834	0.890	0.853	366	367	365
Carman	0.085	0.066	0.050	-0.072	-0.065	-0.073	0.111	0.092	0.088	0.501	0.529	0.586	494	495	486
Monte Buey	0.062	0.040	0.033	-0.005	0.001	-0.022	0.062	0.04	0.039	0.829	0.881	0.801	315	336	330
REMEDHUS	0.039	0.037	0.035	-0.029	-0.007	-0.007	0.049	0.038	0.036	0.819	0.854	0.856	703	858	809
Twente	0.075	0.059	0.052	0.027	0.035	-0.023	0.08	0.069	0.057	0.848	0.853	0.848	847	885	916
HOBE	0.043	0.035	0.042	0.010	0.008	0.001	0.044	0.036	0.042	0.720	0.846	0.836	123	123	123
MAHASRI	0.034	0.031	0.031	-0.026	-0.018	-0.021	0.043	0.036	0.037	0.763	0.773	0.720	351	445	235
Yanco	0.048	0.041	0.033	-0.012	0.016	0.011	0.05	0.044	0.035	0.898	0.906	0.920	571	573	516
Mean Absolute Bias				**0.040**	**0.026**	**0.023**									
SMAP L2SMP_E Average V5	**0.045**	**0.037**	**0.035**	**-0.032**	**-0.008**	**-0.013**	**0.064**	**0.048**	**0.045**	**0.777**	**0.810**	**0.792**			
SMOS L2SM Average V6	**0.057**			**-0.028 (MAB=0.036)**			**0.071**			**0.660**					
SMOS L2SMP_E Average V3	**0.052**			**-0.029**			**0.068**			**0.701**					
SMAP L2SMP_E Average V4	**0.045**	**0.037**	**0.036**	**-0.032**	**-0.008**	**-0.013**	**0.063**	**0.047**	**0.045**	**0.779**	**0.813**	**0.792**			
SMOS L2SMP_E Average V2	**0.055**			**-0.026**			**0.068**			**0.677**					

Table 8.9. SMAP L2SMP_E Version 5 CVS Assessment for Descending (AM) Overpasses over North Polar Grid

CVS	ubRMSE (m^3/m^3)			Bias (m^3/m^3)			RMSE (m^3/m^3)			R			N		
	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA
Reynolds Creek	0.042	0.042	0.045	-0.070	-0.022	-0.012	0.082	0.047	0.047	0.629	0.648	0.594	155	174	170
Walnut Gulch	0.025	0.026	0.029	-0.014	0.024	0.030	0.029	0.035	0.041	0.759	0.772	0.706	317	336	304
TxSON	0.022	0.022	0.025	-0.072	-0.015	0.003	0.076	0.027	0.025	0.923	0.925	0.921	851	851	850
Fort Cobb	0.032	0.028	0.032	-0.086	-0.047	-0.043	0.091	0.055	0.053	0.880	0.892	0.886	766	768	754
Little Washita	0.022	0.021	0.027	-0.059	-0.013	-0.001	0.063	0.025	0.027	0.899	0.917	0.916	611	611	611
South Fork	0.053	0.045	0.041	-0.073	-0.054	-0.046	0.090	0.070	0.061	0.710	0.753	0.764	351	357	316
Little River	0.043	0.035	0.040	0.008	0.059	0.039	0.043	0.068	0.056	0.736	0.768	0.782	718	714	718
Kenaston	0.041	0.030	0.031	-0.033	0.002	0.010	0.052	0.030	0.032	0.725	0.786	0.797	265	266	266
Carman	0.082	0.060	0.049	-0.070	-0.063	-0.070	0.108	0.087	0.085	0.630	0.679	0.707	424	427	429
Monte Buey															
REMEDHUS	0.038	0.037	0.037	-0.022	0.005	0.008	0.044	0.037	0.038	0.826	0.840	0.840	662	717	678
Twente	0.079	0.064	0.056	0.020	0.038	-0.019	0.081	0.074	0.059	0.828	0.833	0.842	654	691	721
HOBE	0.044	0.040	0.044	0.015	0.015	0.008	0.047	0.042	0.044	0.824	0.888	0.866	79	79	79
MAHASRI	0.035	0.033	0.029	-0.019	-0.012	-0.021	0.040	0.035	0.036	0.771	0.801	0.838	285	285	117
Yanco															
Mean Absolute Bias				**0.043**	**0.029**	**0.024**									
SMAP L2SMP_E Average V5	**0.043**	**0.037**	**0.037**	**-0.037**	**-0.006**	**-0.009**	**0.065**	**0.049**	**0.047**	**0.780**	**0.808**	**0.804**			

Table 8.10. SMAP L2SMP_E Version 5 CVS Assessment for Ascending (PM) Overpasses over North Polar Grid

CVS	ubRMSE (m^3/m^3)			Bias (m^3/m^3)			RMSE (m^3/m^3)			R			N		
	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA	SCA-H	SCA-V	DCA
Reynolds Creek	0.049	0.046	0.048	-0.078	-0.031	-0.020	0.092	0.055	0.052	0.575	0.638	0.606	217	252	243
Walnut Gulch	0.026	0.025	0.026	-0.026	0.010	0.014	0.037	0.027	0.029	0.699	0.727	0.624	567	668	594
TxSON	0.020	0.019	0.022	-0.067	-0.015	0.001	0.070	0.024	0.022	0.929	0.935	0.926	932	932	931
Fort Cobb	0.036	0.029	0.030	-0.083	-0.051	-0.051	0.090	0.058	0.059	0.877	0.891	0.873	843	851	808
Little Washita	0.025	0.022	0.026	-0.047	-0.008	-0.001	0.053	0.023	0.026	0.899	0.914	0.900	638	638	634
South Fork	0.053	0.046	0.038	-0.068	-0.058	-0.054	0.086	0.073	0.066	0.737	0.764	0.779	355	366	289
Little River	0.043	0.036	0.038	0.017	0.063	0.042	0.046	0.072	0.056	0.762	0.756	0.774	653	650	652
Kenaston	0.036	0.025	0.030	-0.030	0.000	0.008	0.047	0.025	0.031	0.832	0.890	0.852	356	356	354
Carman	0.083	0.065	0.055	-0.073	-0.066	-0.071	0.110	0.093	0.090	0.510	0.540	0.545	495	497	494
Monte Buey															
REMEDHUS	0.038	0.037	0.035	-0.032	-0.009	-0.008	0.050	0.038	0.036	0.814	0.849	0.852	692	847	793
Twente	0.079	0.061	0.053	0.036	0.039	-0.021	0.087	0.073	0.057	0.846	0.851	0.844	823	879	915
HOBE	0.042	0.039	0.045	0.027	0.024	0.013	0.050	0.046	0.047	0.826	0.892	0.863	83	83	83
MAHASRI	0.034	0.031	0.030	-0.027	-0.019	-0.021	0.043	0.036	0.036	0.764	0.772	0.739	341	434	227
Yanco															
Mean Absolute Bias				**0.047**	**0.030**	**0.025**									
SMAP L2SMP_E Average V5	**0.043**	**0.037**	**0.037**	**-0.034**	**-0.009**	**-0.013**	**0.066**	**0.050**	**0.047**	**0.775**	**0.801**	**0.783**			

Table 8.11. SMAP L2SMP_E Version 5 Sparse Network Assessment for Descending (AM) Overpasses

IGBP Class	ubRMSD (m^3/m^3)				Bias (m^3/m^3)				RMSD (m^3/m^3)				R				N
	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	
Evergreen needleleaf forest	0.042	0.040	0.039	0.080	-0.052	0.006	0.035	-0.016	0.070	0.049	0.059	0.097	0.700	0.710	0.710	0.043	4
Evergreen broadleaf forest																	
Deciduous needleleaf forest																	
Deciduous broadleaf forest	0.071	0.049	0.044	0.119	-0.031	0.016	-0.008	-0.135	0.078	0.051	0.045	0.180	0.330	0.570	0.670	0.147	1
Mixed forest	0.061	0.060	0.060	0.062	-0.065	-0.022	-0.014	-0.065	0.089	0.064	0.061	0.090	0.710	0.720	0.720	0.662	1
Closed shrublands																	
Open shrublands	0.042	0.041	0.042	0.055	-0.050	-0.004	0.013	-0.011	0.070	0.057	0.058	0.065	0.520	0.540	0.540	0.509	44
Woody savannas	0.060	0.057	0.057	0.093	-0.040	0.016	0.009	-0.051	0.101	0.091	0.085	0.132	0.700	0.720	0.730	0.476	27
Savannas	0.038	0.037	0.038	0.052	-0.035	-0.001	-0.004	-0.021	0.066	0.055	0.056	0.063	0.830	0.820	0.830	0.796	3
Grasslands	0.051	0.050	0.051	0.063	-0.076	-0.031	-0.018	-0.043	0.097	0.074	0.072	0.088	0.650	0.660	0.670	0.613	255
Permanent wetlands																	
Croplands	0.077	0.068	0.067	0.085	-0.041	-0.011	-0.007	-0.037	0.120	0.106	0.102	0.124	0.530	0.580	0.580	0.544	74
Urban and built-up																	
Crop/Natural vegetation mosaic	0.071	0.063	0.063	0.095	-0.033	0.014	-0.008	-0.050	0.102	0.090	0.091	0.140	0.610	0.670	0.680	0.425	25
Snow and ice																	
Barren/Sparse	0.022	0.022	0.023	0.029	-0.021	0.008	0.017	-0.011	0.037	0.033	0.037	0.040	0.540	0.530	0.540	0.617	8
SMAP L2SMP_E Average V5	**0.054**	**0.049**	**0.048**		**-0.044**	**-0.001**	**0.002**		**0.083**	**0.067**	**0.067**		**0.612**	**0.652**	**0.667**		
SMAP L2SMP_E Average V4	**0.051**	**0.048**	**0.047**		**-0.043**	**-0.001**	**0.006**		**0.082**	**0.069**	**0.070**		**0.641**	**0.659**	**0.667**		
SMOS L2SM Average V6				**0.069**				**-0.039**				**0.097**				**0.567**	
SMOS L2SM Average V5				**0.069**				**-0.038**				**0.097**				**0.510**	
SMOS V6 data are used in the land cover rows.																	

Table 8.12. SMAP L2SMP_E Version 5 Sparse Network Assessment for Ascending (PM) Overpasses

IGBP Class	ubRMSD (m^3/m^3)				Bias (m^3/m^3)				RMSD (m^3/m^3)				R				N
	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	SCA-H	SCA-V	DCA	SMOS	
Evergreen needleleaf forest	0.039	0.038	0.039	0.080	-0.043	0.016	0.049	-0.005	0.061	0.049	0.064	0.107	0.720	0.710	0.720	0.089	4
Evergreen broadleaf forest																	
Deciduous needleleaf forest																	
Deciduous broadleaf forest	0.073	0.047	0.041	0.124	-0.004	0.030	0.000	-0.083	0.073	0.055	0.041	0.149	0.330	0.580	0.670	0.240	1
Mixed forest	0.055	0.053	0.052	0.062	-0.055	-0.018	-0.013	-0.050	0.078	0.056	0.054	0.079	0.770	0.790	0.800	0.691	1
Closed shrublands																	
Open shrublands	0.042	0.042	0.043	0.059	-0.053	-0.006	0.010	-0.009	0.072	0.057	0.059	0.072	0.470	0.470	0.470	0.426	44
Woody savannas	0.059	0.056	0.055	0.094	-0.025	0.023	0.014	-0.038	0.097	0.091	0.084	0.122	0.700	0.710	0.720	0.518	27
Savannas	0.040	0.038	0.039	0.057	-0.028	0.003	-0.001	-0.020	0.066	0.060	0.059	0.076	0.830	0.820	0.820	0.713	3
Grasslands	0.050	0.049	0.050	0.062	-0.073	-0.031	-0.021	-0.040	0.094	0.073	0.073	0.088	0.650	0.650	0.650	0.616	256
Permanent wetlands																	
Croplands	0.076	0.066	0.064	0.083	-0.034	-0.013	-0.014	-0.040	0.116	0.103	0.099	0.117	0.550	0.590	0.600	0.553	73
Urban and built-up																	
Crop/Natural vegetation mosaic	0.070	0.062	0.061	0.093	-0.011	0.024	0.000	-0.044	0.096	0.090	0.090	0.130	0.620	0.670	0.680	0.451	25
Snow and ice																	
Barren/Sparse	0.022	0.023	0.024	0.037	-0.024	0.009	0.017	-0.012	0.038	0.032	0.036	0.050	0.470	0.450	0.450	0.320	8
SMAP L2SMP_E Average V5	**0.053**	**0.047**	**0.047**		**-0.035**	**0.004**	**0.004**		**0.079**	**0.067**	**0.066**		**0.611**	**0.644**	**0.658**		
SMAP L2SMP_E Average V4	**0.051**	**0.047**	**0.047**		**-0.037**	**0.002**	**0.007**		**0.080**	**0.069**	**0.069**		**0.621**	**0.633**	**0.639**		
SMOS L2SM Average V6				**0.069**				**-0.036**				**0.095**				**0.561**	
SMOS L2SM Average V5				**0.069**				**-0.034**				**0.096**				**0.407**	
SMOS V6 data are used in the land cover rows.																	

Table 8.13. SMAP L2SMP_E Version 5 Sparse Network Assessment for Descending (AM) Overpasses over the North Polar Grid

	ubRMSD (m^3/m^3)			Bias (m^3/m^3)			RMSD (m^3/m^3)			R			N
	SCA-H	**SCA-V**	**DCA**	**SCA-H**	**SCA-V**	**DCA**	**SCA-H**	**SCA-V**	**DCA**	**SCA-H**	**SCA-V**	**DCA**	
Evergreen needleleaf forest	0.051	0.049	0.051	-0.069	-0.007	0.029	0.088	0.057	0.070	0.710	0.720	0.720	6
Evergreen broadleaf forest													
Deciduous needleleaf forest													
Deciduous broadleaf forest	0.067	0.047	0.042	-0.041	0.009	-0.013	0.079	0.048	0.044	0.520	0.690	0.770	1
Mixed forest	0.058	0.058	0.057	-0.071	-0.027	-0.018	0.092	0.064	0.060	0.730	0.730	0.730	1
Closed shrublands													
Open shrublands	0.041	0.041	0.042	-0.043	0.003	0.020	0.065	0.057	0.062	0.540	0.560	0.560	44
Woody savannas	0.060	0.056	0.057	-0.033	0.021	0.010	0.091	0.084	0.080	0.670	0.700	0.710	25
Savannas	0.038	0.036	0.036	-0.054	-0.004	0.012	0.065	0.036	0.038	0.900	0.910	0.920	1
Grasslands	0.051	0.050	0.051	-0.076	-0.030	-0.017	0.096	0.073	0.072	0.660	0.670	0.670	248
Permanent wetlands													
Croplands	0.074	0.067	0.066	-0.072	-0.038	-0.030	0.122	0.104	0.100	0.550	0.600	0.600	60
Urban and built-up													
Crop/Natural vegetation mosaic	0.069	0.060	0.060	-0.035	0.009	-0.009	0.099	0.086	0.087	0.550	0.620	0.640	24
Snow and ice													
Barren/Sparse	0.023	0.024	0.025	-0.024	0.011	0.022	0.039	0.037	0.043	0.470	0.470	0.460	8
SMAP L2SMP_E Average V5	**0.053**	**0.049**	**0.049**	**-0.052**	**-0.005**	**0.001**	**0.083**	**0.067**	**0.067**	**0.612**	**0.652**	**0.667**	

Table 8.14. SMAP L2SMP_E Version 5 Sparse Network Assessment for Ascending (PM) Overpasses over the North Polar Grid

	ubRMSD (m^3/m^3)			Bias (m^3/m^3)			RMSD (m^3/m^3)			R			N
	SCA-H	**SCA-V**	**DCA**	**SCA-H**	**SCA-V**	**DCA**	**SCA-H**	**SCA-V**	**DCA**	**SCA-H**	**SCA-V**	**DCA**	
Evergreen needleleaf forest	0.050	0.049	0.052	-0.061	0.002	0.041	0.082	0.058	0.077	0.710	0.720	0.720	6
Evergreen broadleaf forest													
Deciduous needleleaf forest													
Deciduous broadleaf forest	0.067	0.043	0.039	-0.013	0.022	-0.005	0.069	0.049	0.039	0.500	0.700	0.770	1
Mixed forest	0.053	0.051	0.050	-0.062	-0.022	-0.015	0.082	0.056	0.052	0.780	0.810	0.810	1
Closed shrublands													
Open shrublands	0.043	0.043	0.044	-0.047	0.000	0.016	0.068	0.057	0.061	0.470	0.480	0.470	43
Woody savannas	0.060	0.056	0.057	-0.017	0.027	0.013	0.087	0.085	0.080	0.670	0.680	0.690	25
Savannas	0.039	0.037	0.035	-0.043	0.000	0.011	0.058	0.037	0.037	0.900	0.910	0.920	1
Grasslands	0.051	0.050	0.050	-0.072	-0.029	-0.019	0.093	0.073	0.072	0.640	0.650	0.650	249
Permanent wetlands													
Croplands	0.073	0.066	0.065	-0.062	-0.036	-0.032	0.118	0.103	0.100	0.560	0.610	0.610	60
Urban and built-up													
Crop/Natural vegetation mosaic	0.070	0.060	0.059	-0.013	0.019	-0.004	0.094	0.086	0.086	0.550	0.610	0.620	24
Snow and ice													
Barren/Sparse	0.023	0.024	0.027	-0.028	0.010	0.021	0.039	0.034	0.040	0.520	0.490	0.500	8
SMAP L2SMP_E Average V5	**0.053**	**0.048**	**0.048**	**-0.042**	**-0.001**	**0.003**	**0.079**	**0.064**	**0.064**	**0.630**	**0.666**	**0.676**	

8.3 Summary

Three alternative L2SMP/L2SMP_E retrieval algorithms were evaluated using three methodologies in preparation for this release. The algorithms included the Single Channel Algorithm–H Polarization (SCA-H), Single Channel Algorithm–V Polarization (SCA-V), and Dual Channel Algorithm (DCA). Assessment methodologies were Core Validation Sites (CVS), sparse networks, and intercomparisons with SMOS. **The baseline algorithm has changed from SCA-V to DCA for both L2 and L3 products based on DCA's improved soil moisture retrieval performance in some agricultural areas and consistent overall performance.**

For the current validated release (Version 8 of L2SMP and Version 5 of L2SMP_E), the goal was to update the previous assessment of the retrieval algorithms over a larger period of time and to assess the DCA algorithm in comparison with SCA-V. This assessment was supported by global analyses using sparse networks and SMOS intercomparisons as well as CVS matchups. These analyses indicated that in many areas the SCA-V and DCA had comparable performance statistics, with better unbiased root mean square error and correlation and lower bias than SCA-H. Based on improved soil moisture retrieval performance in some agricultural areas, the DCA algorithm is now the SMAP baseline soil moisture retrieval algorithm. The overall ubRMSE of the DCA and SCA-V retrieved soil moisture from the descending (AM) and ascending PM orbits is on the order of 0.035-0.037 m^3/m^3 for the CVS analyses, which is better than the project target accuracy (0.040 m^3/m^3 ubRMSE or lower).

Sparse network comparisons are more difficult to interpret due to upscaling but provide many more locations than the CVS. The analyses conducted here supported the conclusion reached in the CVS assessment, and contributed to improving validation through Triple Colocation analyses of uncertainties. The sparse network data also allowed the evaluation of performance based on land cover, although the number of sites in a given land cover category was often limited.

SMAP CVS and sparse network retrievals were compared to SMOS. These analyses supported the conclusions of prior assessments that the L2SMP currently has better performance metrics than SMOS.

The analyses described above were repeated for the L2SMP/L2SMP_E PM products. These show comparable performance to the AM results for all metrics. The comparable performance for AM and PM retrievals is attributed to the improved land surface temperature correction approach implemented in the previous version and continued in this release.

The L2SMP_E Version 5 product was assessed using CVS chosen specifically to exploit the L2SMP_E grid posting (9 km) and contributing domain (33 km). Results were essentially the same as those obtained in the L2SMP Version 8 analyses. Results for L2SMP_E using the North Polar grid were very similar to those obtained using the global grid.

Based on the extensive validation analyses to date, the number of peer-reviewed publications (including the numerous independent investigations noted in the SMAP bibliography posted at NSIDC), the length of the SMAP period of record, and the utilization of feedback of validation in a systematic update, with this version of L2SMP and L2SMP_E the team has progressed beyond CEOS Stage 4 validation [10]. The Cal/Val program will continue throughout the mission with the goals of increasing the robustness of the soil moisture products and addressing specific site issues.

9 OUTLOOK AND FUTURE PLANS

Satellite passive microwave retrieval of soil moisture has been the subject of intensive study and assessment for the past several decades. Over this time there have been improvements in the microwave instruments used, primarily in the availability of L-band sensors on orbit. However, sensor resolution has remained roughly the same over this period, which is actually an achievement considering the increase in sensor wavelength from X band to C band to L band over the years. With spatial resolution in the 25-50 km range, there will always be heterogeneity within the satellite footprint that will influence the accuracy of the retrieved soil moisture as well as its validation. Precipitation types and patterns are one of the biggest contributors to this heterogeneity. As a result, one should not expect that the validation metric ubRMSE will ever approach zero except in very homogeneous domains. In contrast, bias tends to be indicative of a systematic error, possibly related to algorithm parameterization and model structure. High quality data are needed to discover and address these systematic errors. Some issues that should be considered during the extended SMAP mission include:

- *Increasing the number of potential validation sites*. Although very useful in SMAP product assessments, the current set of SMAP CVS are somewhat limited in the different environments and land covers they represent, especially at the high resolutions typical of active/passive or disaggregated passive soil moisture products. Additional high resolution core sites in different geographical areas would provide increased confidence in SMAP global performance.

- *Evaluate the impacts of algorithm structure and components on retrieval.* There are some aspects of soil moisture retrieval algorithms that are used because they facilitate operational soil moisture retrieval. One of these simplifying aspects is the use of the Fresnel equations that specify that conditions in the microwave contributing depth are uniform. While there is ample evidence that this is true in most cases, it should be recognized that this assumption is a potential source of error – some effort should be made to evaluate when and where it limits soil moisture retrieval accuracy. Another assumption is that a single dielectric mixing model applies under all conditions globally. All of the commonly-used dielectric models are highly dependent on the robustness of the data set used in their development. The impact of this assumption on retrieval error needs further evaluation. Another consideration in the current DCA is the assumption of equality of the vegetation parameters for the H and V polarizations. This assumption does simplify retrieval but it is not valid for all categories of vegetation.

- *Possible subdivision of crop land cover class into distinct crop subclasses.* Another source of error is SMAP's use of a single IGBP land cover class to cover the great variety of global crops. One area of future work will examine the possibility of subdividing the single crop class into a number of distinct subclasses (e.g., corn, soybeans, wheat, rice) with appropriate parameterization which would better represent the main global crop structural categories. Due to the latency problem in acquiring up-to-date crop maps, this issue is not likely to be addressed until the final bulk reprocessing of SMAP data.

- *Incorporating field campaign results into algorithm assessments and improvements.* Several SMAP field campaigns were conducted postlaunch, and results from these field campaigns will be used in future assessments and algorithm improvements.

- *Improvement of retrievals over forests*. Dense forests (where VWC > 5 kg/m^2) typically exceed the currently accepted threshold for accurate soil moisture retrieval. SMAP provides a flagged retrieval over forests, and the spatial extent of these flagged areas is quite large. At this point there is no supporting validation of the L2SMP soil moisture retrieved for forest areas, and the SMAP forest retrievals are quite different from SMOS. Efforts to improve soil moisture retrievals over forests

should include a careful evaluation of alternative algorithms and improving validation resources through a combination of temporary *in situ* soil moisture networks and field campaigns. The SMAPVEX22 field campaign will collect data in two U.S. temperate forests (Millbrook, NY and Harvard Forest, MA) to begin to address soil moisture retrieval in forested areas. The SMAPVEX22-BERMS campaign in Saskatchewan, Canada will extend these studies to boreal forests. Both field campaigns will deploy temporary *in situ* stations, ground radiometers, and the aircraft PALS SMAP simulator instrument to collect sufficient ground truth data for meaningful comparisons with SMAP data in forested domains.

- *Improvement of retrievals for organic soils.* Organic soils in areas with high vegetation provide a challenge for soil moisture retrievals. SMAP currently uses the Mironov dielectric mixing model for estimating soil moisture. At the present time this model does not account for the presence of organic material in the soil. The impact of organic matter on the dielectric constant is unknown and will be studied in the future along with the forest retrievals. The presence of organic matter also leads to uncertainty in the bulk density value/porosity which is taken to be the upper limit of retrieved soil moisture.

10 ACKNOWLEDGEMENTS

This document resulted from many hours of diligent analyses and constructive discussion among the L2SMP Team, Cal/Val Partners, and other members of the SMAP Project Team.

This research was also supported (in part) by the U.S. Department of Agriculture, Agricultural Research Service. USDA is an equal opportunity provider and employer. *This research was a contribution from the USDA Long-Term Agroecosystem Research (LTAR) network.*

11 REFERENCES

[1] Jackson, T., P. O'Neill, S. Chan, R. Bindlish, A. Colliander, F. Chen, M. Burgin, S. Dunbar, J. Piepmeier, M. Cosh, T. Caldwell, J. Walker, X. Wu, A. Berg, T. Rowlandson, A. Pacheco, H. McNairn, M. Thibeault, J. Martínez-Fernández, Á. González-Zamora, E. Lopez-Baeza, F. Udall, M. Seyfried, D. Bosch, P. Starks, C. Holifield, J. Prueger, Z. Su, R. van der Velde, J. Asanuma, M. Palecki, E. Small, M. Zreda, J. Calvet, W. Crow, Y. Kerr, S. Yueh, and D. Entekhabi, June 6, 2018. Calibration and Validation for the L2/3_SM_P Version 5 and L2/3_SM_P_E Version 2 Data Products, SMAP Project, JPL D-56297, Jet Propulsion Laboratory, Pasadena, CA.

[2] Chan, S., Bindlish, R., O'Neill, P., Njoku, E.., Jackson, T. J., Colliander, A., Chen, F., Burgin, M., Dunbar, R.S., Piepmeier, J, Yueh, S., Entekhabi, D., Cosh, M.H., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martinez-Fernandez, J., Gonzalez-Zamora, A., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Palecki, M., Small, E.E., Zreda, M., Calvet, J.-C., Crow, W.T., and Kerr, Y. Assessment of the SMAP level 2 passive soil moisture product. IEEE Transactions on Geoscience and Remote Sensing, 54 (8): 4994-5007, 2016, doi: 10.1109/TGRS.2016.2561938.

[3] Chan, S., Bindlish, R., O'Neill, P.E., Jackson, T.J., Njoku, E., Dunbar, R.S., Chaubell, J., Peipmeier, J, Yueh, S., Entekhabi, D., Colliander, A., Chen, F., Cosh, M.H., Caldwell, T., Walker, J., Berg, A., McNairn, H., Thibeault, M., Martinez-Fernandez, J., Udall, F., Seyfried, M.S., Bosch, D.D., Starks, P., Holyfield Collins, C., and Prueger, J.H. Development and assessment of the SMAP enhanced passive soil moisture product. Remote Sensing of Environment, 204: 931-941. 2018.

[4] Colliander, A., Jackson, T.J., Chan, S., .O'Neill, P.E., Bindlish, R., Cosh, M.H., Berg, A., Rowlandson, T., Bosch, D., Caldwell, Walker, J.P., Berg, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., Jensen, K.H., Asanuma, J., Seyfried, M., Bosch, D.D., Starks, P., Holifield Collins, C., Prueger, J.H., Su, Z., Lopez-Baeza, E., and Yueh, S.H. An assessment of the differences between spatial resolution and grid size for the SMAP enhanced soil moisture product over homogeneous sites. Remote Sensing of Environment, 207: 65-70. 2018.

[5] Chen, F., Crow, W. T., Colliander, A., Cosh, M., Jackson, T. J., Bindlish, R., Reichle, R., Chan, S. K., Bosch, D. D., Starks, P. S., Goodrich, D. C., and Seyfried, M. S. 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: 489-502. 2017.

[6] Jackson, T. J., Colliander, A., Kimball, J., Reichle, R., Crow, W., Entekhabi, D. et al. Science data calibration and validation plan, Release A, JPL D-52544, March 14, 2014.

[7] O'Neill, P., Bindlish, R, Chan, S., Njoku, E., Chaubell, J., and Jackson, T. SMAP Algorithm Theoretical Basis Document (ATBD): L2/3_SM_P, Revision F, JPL D-66480, Jet Propulsion Laboratory, Pasadena, CA, August 31, 2020.

[8] Entekhabi, D., Yueh, S., O'Neill, P., Kellogg, K. et al., SMAP Handbook, JPL Publication JPL 400-1567, Jet Propulsion Laboratory, Pasadena, California, 182 pages, July, 2014.

[9] SMAP Level 1 Mission Requirements and Success Criteria. (Appendix O to the Earth Systematic Missions Program Plan: Program-Level Requirements on the Soil Moisture Active Passive Project.), NASA Headquarters/Earth Science Division, Washington, DC, version 5, 2013.

[10] Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV): http://calvalportal.ceos.org/CalValPortal/welcome.do and WWW: Land Products Sub-Group of Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV): http://lpvs.gsfc.nasa.gov.

[11] Peng, J., Misra, S., Piepmeier, J.R., Dinnat, E.P., Hudson, D., Le Vine, D.M., De Amici, G., Mohammed, P.N., Bindlish, R., Yueh, S.H., Meissner, T., and Jackson, T. J. Soil moisture active/passive (SMAP) L-band microwave radiometer post-launch calibration. IEEE Trans. Geoscience and Remote Sensing, 55(4): 1897-1914. 2017.

[12] SMAP Algorithm Theoretical Basis Document: Enhanced L1B_TB_E Radiometer Brightness Temperature Data Product. SMAP Project, JPL D-56287, Jet Propulsion Laboratory, Pasadena, CA.

[13] Piepmeier, J., P. Mohammed, G. De Amici, E. Kim, J. Peng, C. Ruf, and M. Chaubell. SMAP Algorithm Theoretical Basis Document: L1B Radiometer Product. Revision C, June 6, 2018, SMAP Project, GSFC-SMAP-006, NASA Goddard Space Flight Center, Greenbelt, MD.

[14] Dawson, D., SMAP radiometer error budget, SMAP Project, JPL D-61632, Jet Propulsion Laboratory, Pasadena, CA.

[15] Mironov, V. L., Kosolapova, L. G., and Fomin, S. V. Physically and mineralogically based spectroscopic dielectric model for moist soils. IEEE Trans. Geosci. Remote Sens., 47(7): 2059-2070. 2009.

[16] Choudhury, B.J. and Schmugge, T.J. A parameterization of effective soil temperature for microwave emission. Journal of Geophysical Research, 87(C2): 1301-1304.1982.

[17] Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiova, A, Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M. Global automated quality control of in situ soil moisture data from the international soil moisture network. Vadose Zone Journal, 12. doi:10.2136/vzj2012.0097. 2012.

[18] Adams, J. R., McNairn, H., Berg, A. A., and Champagne, C. Evaluation of near-surface soil moisture data from an AAFC monitoring network in Manitoba, Canada: Implications for L-band satellite validation. Journal of Hydrology, 521: 82-592. 2015.

[19] Lawrence, H., Wigneron, J. -P., Demontoux, F., Mialon, A., and Kerr, Y. H. (2013). "Evaluating the Semiempirical H-Q Model Used to Calculate the L-Band Emissivity of a Rough Bare Soil", *IEEE Transactions on Geoscience and Remote Sensing*, 51(7), 4075–4084.

[20] O'Neill, P., S. Chan, R. Bindlish, M. Chaubell, A. Colliander, F. Chen, S. Dunbar, T. Jackson, J. Piepmeier, S. Misra, M. Cosh, T. Caldwell, J. Walker, X. Wu, A. Berg, T. Rowlandson, A. Pacheco, H. McNairn, M. Thibeault, J. Martínez-Fernández, Á. González-Zamora, E. Lopez-Baeza, F. Udall, M. Seyfried, D. Bosch, P. Starks, C. Holifield, J. Prueger, Z. Su, R. van der Velde, J. Asanuma, M. Palecki, E. Small, M. Zreda, J. Calvet, W. Crow, Y. Kerr, S. Yueh, and D. Entekhabi, August 15, 2019. Calibration and Validation for the L2/3_SM_P Version 6 and L2/3_SM_P_E Version 3 Data Products, SMAP Project, JPL D-56297, Jet Propulsion Laboratory, Pasadena, CA.

[21] Chaubell, J., R. Bindlish, S. Chan, F. Chen, A. Colliander, S. Dunbar, D. Entekhabi, P. O'Neill, S. Yueh, J. Asanuma, A. Berg, D. Bosch, T. Caldwell, M. Cosh, C. Holifield Collins, J. Martínez-Fernández , M. Seyfried, P. Starks, Z. Su, M. Thibeault, and J. Walker, "Improved SMAP Dual-Channel Algorithm for the Retrieval of Soil Moisture," *IEEE Trans. on Geoscience and Remote Sensing*, June, 2020, vol. 58, no. 6, pp. 3894-3905, doi: 10.1109/TGRS.2019.2959239.

[22] Colliander, A., T. Jackson, R. Bindlish, S. Chan, S. Kim, M. Cosh, S. Dunbar, L. Dang, L. Pashian, J. Asanuma, A. Berg, T. Rowlandson, D. Bosch, T. Caldwell, K. Caylor, D. Goodrich, H. al Jassar, E. Lopez-Baeza, J. Martinez-Fernandez, A. Gonzalez-Zamora, S. Livingston, H. McNairn, A. Pacheco, M. Moghaddam, C. Montzka, C. Notarnicola, G. Niedrist, T. Pellarin, J. Prueger, J. Pulliainen, K. Rautiainen, J. Ramos, M. Seyfried, P. Starks, Z. Su, Y. Zeng, R. van der Velde, M. Thibeault, W. dorigo, M. Vreugdenhil, J. Walker, X. Wu, A. Monerris, P. O'Neill, D. Entekhabi, E.

Njoku, and S. Yueh, "Validation of SMAP Surface Soil Moisture Products with Core Validation Sites," *Remote Sensing of Environment*, 191: 215-231, March, 2017, http://dx.doi.org/10.1016/j.rse.2017.01.021.

[23] Colliander, A., M. Cosh, S. Misra, T. Jackson, W. Crow, J. Powers, H. McNairn, P. Bullock, A. Berg, R. Magagi, Y. Gao, R. Bindlish, R. Williamson, I. Ramos, B. Latham, P. O'Neill, and S. Yueh, "Comparison of High-Resolution Airborne Soil Moisture Retrievals to SMAP Soil Moisture during the SMAP Validation Experiment 2016 (SMAPVEX16)," *Remote Sensing of Environment*, 227, 2019, pp. 137-150. https://doi.org/10.1016/j.rse.2019.04.004

[24] Forgotson, C., P. O'Neill, M. Carrera, S. Belair, N. Das, I. Mladenova, J. Bolten, J. Jacobs, E. Cho, and V. Escobar, "How Satellite Soil Moisture Can Help to Monitor the Possible Impacts of Climate Change: SMAP Case Studies," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS)*, vol. 13 (1), Dec., 2020, pp. 1590-1596, doi: 10.1109/JSTARS.2020.2982608.

[25] Jackson, T., P. O'Neill, S. Chan, R. Bindlish, A. Colliander, F. Chen, S. Dunbar, J. Piepmeier, M. Cosh, T. Caldwell, J. Walker, X. Wu, A. Berg, T. Rowlandson, A. Pacheco, H. McNairn, M. Thibeault, J. Martínez-Fernández, Á. González-Zamora, E. Lopez-Baeza, F. Udall, M. Seyfried, D. Bosch, P. Starks, C. Holifield, J. Prueger, Z. Su, R. van der Velde, J. Asanuma, M. Palecki, E. Small, M. Zreda, J. Calvet, W. Crow, Y. Kerr, S. Yueh, and D. Entekhabi, December 15, 2016, April 5, 2017. Calibration and Validation for the L2/3_SM_P Version 4 and L2/3_SM_P_E Version 1 Data Products, SMAP Project, JPL D-56297, Jet Propulsion Laboratory, Pasadena, CA.

[26] Peng, J., Misra, S., Piepmeier, J.R., Dinnat, E.P., Yueh, S.H., Meissner, T., Le Vine, D.M., Shelton, K.E., Freedman, A.P., Dumbar, R.S., Chan, S.K., Bindlish, R., De Amici, G., Mohammed, P.N., Hudson, D., and Jackson, T. J., "Soil Moisture Active/Passive (SMAP) L-Band Microwave Radiometer Post-Launch Calibration Upgrade," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 12, no. 6, pp. 1647-1657, June 2019.

[27] O'Neill, P., S. Chan, R. Bindlish, M. Chaubell, A. Colliander, F. Chen, S. Dunbar, T. Jackson, J. Peng, M. Cosh, T. Bongiovanni, J. Walker, X. Wu, A. Berg, H. McNairn, M. Thibeault, J. Martínez-Fernández, Á. González-Zamora, E. Lopez-Baeza, K. Jensen, M. Seyfried, D. Bosch, P. Starks, C. Holifield Collins, J. Prueger, Z. Su, R. van der Velde, J. Asanuma, M. Palecki, E. Small, M. Zreda, J. Calvet, W. Crow, Y. Kerr, S. Yueh, and D. Entekhabi, August 31, 2020. Calibration and Validation for the L2/3_SM_P Version 7 and L2/3_SM_P_E Version 4 Data Products, SMAP Project, JPL D-56297, Jet Propulsion Laboratory, Pasadena, CA.

[28] Chan, S., E. Njoku, and A. Colliander. SMAP Algorithm Theoretical Basis Document: L1C Radiometer Product. Revision B, SMAP Project, JPL D-53053, Jet Propulsion Laboratory, Pasadena, CA, October 14, 2015.

[29] Peng, J., R. Bindlish, A. Bringer, S. Chan, J. Chaubell, A. Colliander, G. De Amici, E. Dinnat, D. Entekhabi, D. Hudson, T. Jackson, J. Johnson, D. Le Vine, T. Meissner, S. Misra, P. Mohammed, J. Piepmeier, and S Yueh. SMAP Radiometer Brightness Temperature Calibration for the L1B_TB (Version 5), L1C_TB (Version 5), and L1C_TB_E (Version 3) Data Products, SMAP Project, JPL-D 56295, Jet Propulsion Laboratory, Pasadena, CA, August 8, 2020.

[30] O'Neill, P., R. Bindlish, S. Chan, J. Chaubell, A. Colliander, E. Njoku, and T. Jackson, 2021. SMAP Algorithm Theoretical Basis Document: Level 2 & 3 Soil Moisture (Passive) Data Products. Revision G, October 12, 2021, SMAP Project, JPL D-66480, Jet Propulsion Laboratory, Pasadena, CA.

12 LINK TO BIBLIOGRAPHY OF PUBLISHED SMAP PAPERS

Please see the SMAP bibliography posted at NSIDC at the following link for SMAP-related papers: https://nsidc.org/data/smap/research.html. The SMAP Science Highlights page on the SMAP website at JPL (https://smap.jpl.nasa.gov/science/highlights/) is another source of information about the use of SMAP data in science and applications.

APPENDIX 1: PARAMETERIZATION OF EFFECTIVE SOIL TEMPERATURE (T_{EFF})

Accurate soil moisture retrievals from SMAP 1.4 GHz brightness temperature data require an estimate of the soil effective temperature (T_{eff}). The Choudhury two-layer approach for T_{eff} [Choudhury et al., 1982] has been used in all previous SMAP data releases and by other missions with good success. This approach combines estimates of a "surface" temperature and a "deep" temperature using a proportional coefficient dependent on the microwave wavelength. In the SMAP 2018 data release, the Choudhury approach was modified by an additional parameter K to address an observed bias between ancillary modeled soil temperatures and measured *in situ* temperatures at core validation sites and sparse network stations. In January, 2020 NASA GMAO modified atmospheric and land modules in the GEOS-FP modeling system, resulting in a change to the soil temperature ancillary data used in SMAP processing and a change in the definition of soil layer depths [Koster et al., 2020]. As a result, the SMAP project conducted a number of different analyses to determine the most appropriate parameterization of the Choudhury effective temperature. This appendix contains two summaries of these efforts. As of the R17 release in August, 2020, the approach outlined in section A3.2 was adopted as the baseline approach for T_{eff} formulation and is used in the R18 release in October 2021.

A1.1. Rationale for a Two-Layer Effective Soil Temperature Model for SMAP Passive Soil Moisture Products

(Lead Investigator: Steven Chan, JPL)

Experimental and theoretical literature (Choudhury 1982 and Lv 2014, 2016) have long established the physics that at L-band frequencies, microwave emission from land primarily comes from the deep soil layer ($\boldsymbol{T_{deep}}$), and to a lesser extent, the shallower soil layer ($\boldsymbol{T_{top}}$). The combined use of $\boldsymbol{T_{deep}}$ and $\boldsymbol{T_{top}}$ in effective soil temperature ($\boldsymbol{T_{eff}}$) modeling has in the past enabled great science advances for L-band satellite missions including SMOS (2009-present) and SMAP (2015-present) as well as enabling other research reported in refereed literature.

The upcoming SMAP R17 release consists of upgrades in/changes to radiometer calibration, GMAO ancillary soil temperatures, GMAO soil layer depths, and the global soil texture database used. As such, the pre-R17 two-layer $\boldsymbol{T_{eff}}$ model was examined to see if parameter retuning was required. Since the impacts of $\boldsymbol{T_{eff}}$ are most directly felt in soil moisture retrieval accuracy, sensitivity analysis was applied to the *in situ* soil moisture observations from core validation sites to determine the new parameters (C, K) for the R17 two-layer $\boldsymbol{T_{eff}}$ model. To ensure validation with minimal partiality in product assessment, observations from only a random half of all CVS stations were used at a time in the sensitivity analysis.

The parameters determined in this way also independently confirmed three outcomes: (1) a larger $\boldsymbol{T_{top}}$ contribution in R17 that is in line with theoretical prediction (e.g., Lv, 2014) and the new GEOS-FP soil depth layer definitions, (2) the validity of a two-layer model to describe $\boldsymbol{T_{eff}}$ for SMAP 6:00 AM data, and (3) the validity of using C=1 to describe $\boldsymbol{T_{eff}}$ for SMAP 6:00 PM data. The resulting retrieval performance metrics and spatial/temporal variability continue to be consistent with those derived before R17, giving additional confidence in SMAP's continued ability to produce high-quality soil moisture data products.

The analysis started with the objective to determine if (a) SMAP $\boldsymbol{T_B}$ observations and (b) footprint-scaled CVS soil moisture observations support a two-layer $\boldsymbol{T_{eff}}$ model or a one-layer $\boldsymbol{T_{eff}}$ model. Baseline soil moisture retrieval time series were obtained at all CVS one (C, K) pair at a time using the following two-layer $\boldsymbol{T_{eff}}$ formulation, which reduces to a one-layer $\boldsymbol{T_{eff}}$ model when $C = 1$.

$$\boldsymbol{T_{eff} = K \times [\, T_{SOIL1} \times C + T_{SOIL2} \times (1 - C)]}$$

where K is a factor that modifies the original Choudhury formulation to address an observed bias between ancillary modeled soil temperature and measured *in situ* temperature at core validation sites and sparse network stations. T_{soil1} refers to the average soil temperature for the first soil layer (5-15 cm) and T_{soil2} refers to the average soil temperature for the second soil layer (15-35 cm) of the GMAO GEOS-FP land surface model. Assuming direct correspondence between (a) and (b) above, favorable metrics are expected along $C = 1$ for the validity of a one-layer $\boldsymbol{T_{eff}}$ model, or along $C \neq 1$ for the validity of a two-layer $\boldsymbol{T_{eff}}$ model. In this sensitivity analysis setup, no implicit *a priori* assumption was made on what (C, K) values should take to dictate the metric pattern one way or the other.

The SMAP CVS data are comprised of *in situ* soil moisture observation time series from 15 stations. To minimize partiality upon assessing the data product with the same dataset, 8 stations were randomly selected (~50% of all available stations) at a time, and their time series were used to determine one realization of (C, K). This process was repeated until all ${}_{15}C_8 = 15! / (7! \times 8!) = 6{,}435$ possible combinations were exhausted. All realizations of (C, K) were then averaged to provide their mean values and standard deviations:

$$(C, K) = (0.43 \pm 0.14, 1.0106 \pm 0.0042)$$

which approach their theoretical limits (C, K) = (0.46,1.0103) had all CVS been used in the first place. However, this random-selection sensitivity analysis approach is not without drawbacks: for example, even though only half of them were used in sensitivity analysis for any given realization, all of them were eventually involved, though not all at the same time. In view of the inherent difficulty in $\boldsymbol{T_{eff}}$'s direct observability other than in soil moisture retrievals, however, this lack of rigor in analysis design appears to be a temporary but reasonable compromise in subsequent validation in order to help answer the following pressing question that initiated this investigation:

> Do SMAP $\boldsymbol{T_B}$ observations and CVS soil moisture observations indicate a two-layer $\boldsymbol{T_{eff}}$ model or a one-layer $\boldsymbol{T_{eff}}$ model?

As an illustration, Figure A1-1 shows the 6:00 AM and 6:00 PM performance metrics over the (C, K) space for the 4-year T16540 OASIS run. Two observations follow from this figure. First, favorable 6:00 PM metrics – defined as low ubRMSE, small bias observed over CVS magnitude, *and* high correlation – all occur simultaneously near the $C = 1$ line. The corresponding K varies within a narrow band of values around 1.01. The fact that $C = 1$ implies that the one-layer model provides the best analytical form for $\boldsymbol{T_{eff}}$ at 6:00 PM. This result independently affirms the choice made by the team in 2017 to simply use $C = 1$ for the SMAP 6:00 PM data. Under this formulation, the effect of $\boldsymbol{T_{SOIL2}}$ is negated and only $\boldsymbol{T_{SOIL1}}$ is used for $\boldsymbol{T_{eff}}$ modeling in Level 2 soil moisture geophysical inversion for 6:00 PM data.

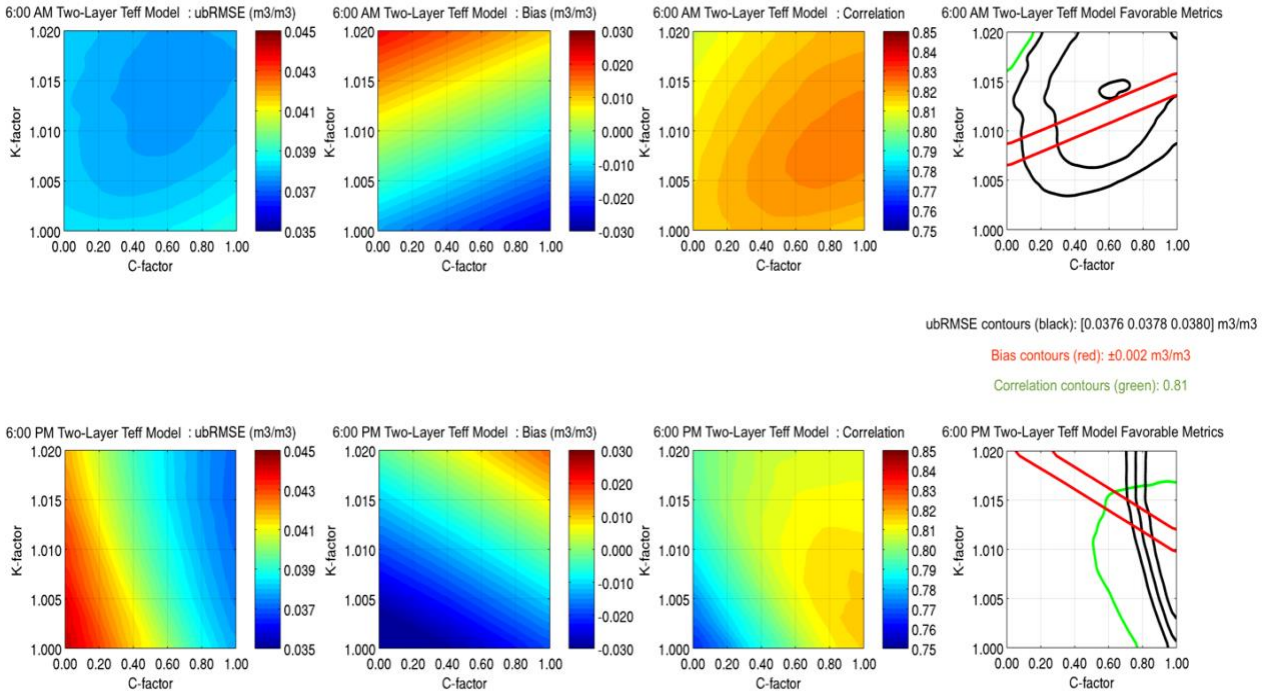


Figure A1-1. 6:00 AM and 6:00 PM performance metrics as a function of *C* and *K*

On the other hand, the favorable 6:00 AM metrics (low ubRMSE, small bias magnitude, *and* high correlation) convey a different structure. Unlike their 6:00 PM counterparts, they do not align well with the $C = 1$ line, implying that the two-layer model is necessary to provide the best analytical form for $\boldsymbol{T_{eff}}$ for the SMAP 6:00 AM data. Since $C \neq 1$ to obtain the best results at 6:00 AM, both $\boldsymbol{T_{SOIL1}}$ and $\boldsymbol{T_{SOIL2}}$ are needed simultaneously to explain the correspondence between the SMAP $\boldsymbol{T_B}$ observations and CVS soil moisture observations for 6:00 AM data.

Although the two-layer model can be reduced to the one-layer model by setting $C = 1$ (as for the 6:00 PM data), the converse is not possible – the one-layer model cannot be modified easily to predict and absorb the contribution of $\boldsymbol{T_{SOIL2}}$ implicitly in the simple analytical form: $\boldsymbol{T_{eff}} = \mathbf{K} \times \boldsymbol{T_{SOIL1}}$. For this reason, the two-layer $\boldsymbol{T_{eff}}$ model is recognized to be the better candidate for its flexibility to take into account the need to use $\boldsymbol{T_{SOIL1}}$ and $\boldsymbol{T_{SOIL2}}$ for the 6:00 AM data and $\boldsymbol{T_{SOIL1}}$ only for the 6:00 PM data.

Note also that the colocation of the favorable 6:00 AM metric patterns in Fig. A1-1 occurs around $C \sim 0.40$. This value of *C* is higher than the traditional Choudhury value ($C = 0.246$) used in previous data releases. In other words, based on the results shown in Figure A3-1, $\boldsymbol{T_{SOIL1}}$ should now assume a 40% weight in $\boldsymbol{T_{eff}}$. This increase is consistent with (1) a change in soil depth definitions in the GEOS-FP ancillary soil temperatures as well as (2) theoretical predictions outlined in Lv, 2014 (Figure A1-2). Because the CVS soil temperatures in the sensitivity analysis do not stay constant everywhere at 20 ºC as assumed in Lv's calculations, a perfect match is not expected between our *C* and Lv's *C*, but they are nonetheless in close agreement.

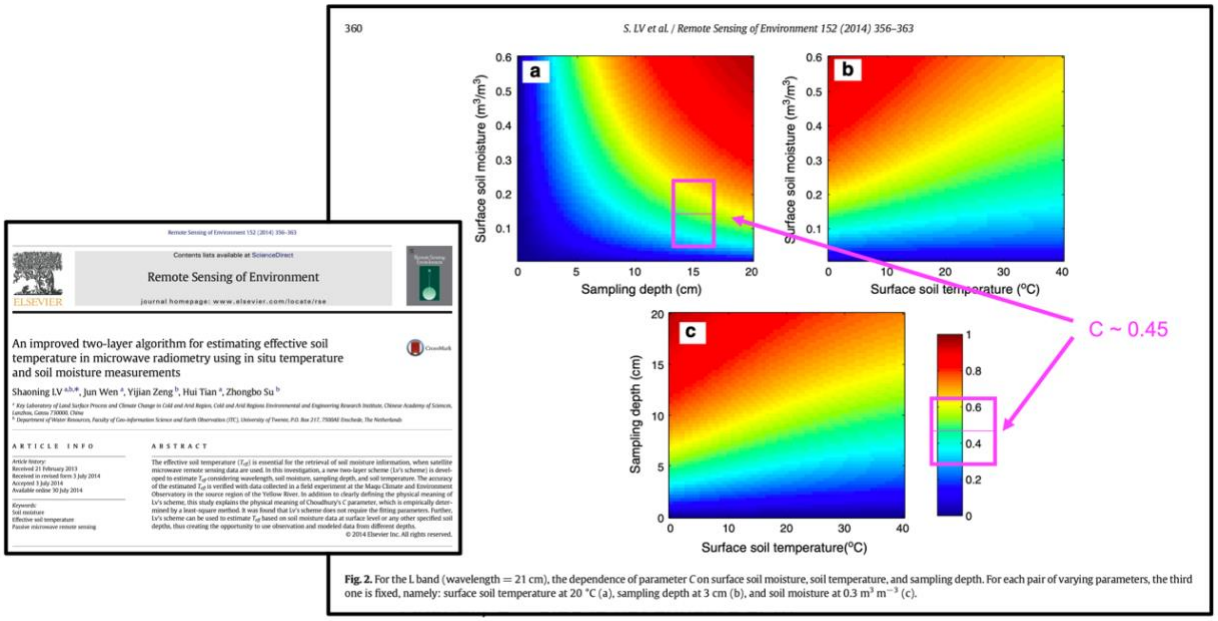


Figure A1-2. Theoretical prediction of $C \sim 0.45$ for the typical dynamic range of soil moisture observations at CVS and the sampling depth of $\boldsymbol{T_{SOIL1}}$ according to the new GMAO GEOS-FP soil depth layer definition.

CONCLUSION

Using the sensitivity analysis outlined above, the two-layer $\boldsymbol{T_{eff}}$ model was determined to provide the best explanation for the correspondence between the SMAP $\boldsymbol{T_B}$ observations and CVS soil moisture observations, with $C = 0.435$ for the SMAP 6:00 AM data and $C = 1$ for the SMAP 6:00 PM data ($K = 1.01$ in both cases) over biomes where SMAP soil moisture retrieval is expected to meet or exceed the mission target accuracy of 0.040 m^3/m^3 (1-sigma).

ADDENDUM

No account for justifying the two-layer $\boldsymbol{T_{eff}}$ model is considered complete without a critical examination into why the other alternative – the one-layer $\boldsymbol{T_{eff}}$ model – is considered a less preferred candidate. From a modeling perspective, the one-layer $\boldsymbol{T_{eff}}$ model is conceptually simple and elegant; it adheres well to a minimalist's approach, requiring no "knobs" to account for the contribution from the deeper soil layers.

It turns out the lack of "knobs" in the one-layer $\boldsymbol{T_{eff}}$ model is also one of its major limitations – there is simply no robust way to account for diverse land surface conditions where the contribution from the deeper soil layers is significant, unless additional corrections are developed on top of this formulation to model the impacts of heat diffusion, soil texture, and soil moisture from the deeper soil layers. None of these are trivial to implement in Level 2 soil moisture geophysical inversion. As a rule of thumb in using and understanding remote sensing data, effective corrections should be applied as close to the first source(s) as possible. Since $\boldsymbol{T_{SOIL2}}$ is already available in the SMAP operational data stream, there is strong incentive to use it to provide correction sooner at $\boldsymbol{T_{eff}}$ rather than later. Furthermore, there is a large body of theoretical and experimental research which indicates that at L-band frequencies, there is significant contribution to the observed $\boldsymbol{T_B}$ from the deeper soil layers. This is especially true in areas where there is not an excessive

amount of vegetation and/or soil moisture. Globally, these areas represent a significant portion of the original "SMAP Retrieval Mask" where mission target accuracy is expected to be satisfied.

For the one-layer T_{eff} model to be able to "ignore" contributions to the emission temperature from T_{SOIL2}, it is necessary that the isothermal condition be met. The isothermal condition is a rather stringent condition because it goes beyond requiring the vertical thermal gradient in the soil to be small or minimum (e.g. at dawn hours) – it actually requires $T_{SOIL1} = T_{SOIL2}$ so that their difference is zero (mathematically, as long as T_{SOIL2} is a constant multiple of T_{SOIL1}, the two-layer T_{eff} model can still be reduced to the one-layer T_{eff} model but nature is not likely to act like this everywhere all the time!). To examine the validity of this isothermal condition, both T_{SOIL1} and T_{SOIL2} were extracted and global local solar time correction was performed using all available hourly files from the GEOS-FP soil temperatures. Figure A1-3 shows the global T_{SOIL1} - T_{SOIL2} differences at local solar time 6:00 AM everywhere.

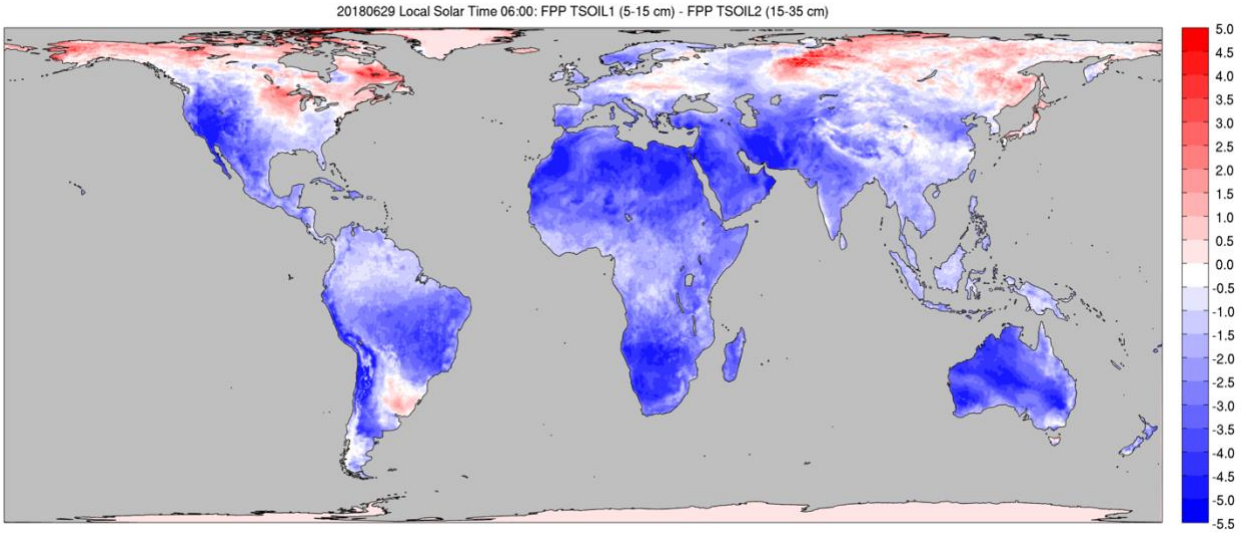


Figure A1-3. Global T_{SOIL1} - T_{SOIL2} difference at local solar time 6:00 AM

White areas indicate where $T_{SOIL1} = T_{SOIL2}$, whereas blue areas indicate where $T_{SOIL1} < T_{SOIL2}$ by about 3-5 K. Clearly, there are non-zero thermal gradients over many places even at dawn hours, thus undermining the validity of the one-layer T_{eff} model. Note that Figure A3-3 refers to only the difference between the 5-15 cm soil layer and the 15-35 cm soil layer as given by the GEOS-FP model. If soil temperature for the 0-5 cm soil layer were available, its gradient from the 5-15 cm soil layer would be even greater (see Fig. 1 in Choudhury et al., 1982). This would further challenge the isothermal condition required by the one-layer T_{eff} model.

A1.2. Emission Temperature and Its Inference from Dynamic Ancillary Data

(Lead Investigator: Dara Entekhabi, MIT)

I. The Effects of Temperature and Moisture Profiles and Their Parameterization

The microwave emission brightness temperature of soils T depends on the soil temperature profile $T_g(z)$ and the attenuation coefficient $\alpha(z)$ profile:

$$T = \int_0^{\infty} T_g(z) \cdot \alpha(z)\, exp\left[-\int_0^{z} \alpha(z')\, dz'\right] dz$$

where

$$\alpha(z) = \frac{4\pi}{\lambda} \frac{\epsilon''(z)}{2\sqrt{\epsilon'(z)}}$$

$\epsilon = \epsilon' + i\,\epsilon''$ is the complex dielectric constant that is dependent on the soil volumetric water content, and λ is the wavelength.

In order to produce estimates of surface soil moisture from brightness temperature observations, the soil temperature T_g has to be treated as a dynamic ancillary data input which can be inferred from atmospheric forecast models that include land surface parameterizations (surface energy balance and heat diffusion into discretized soil media). NASA GMAO's Goddard Earth Observing System-Forward Processing **(**GEOS-FP) modeling system is the source of these soil temperature forecasts for the SMAP project.

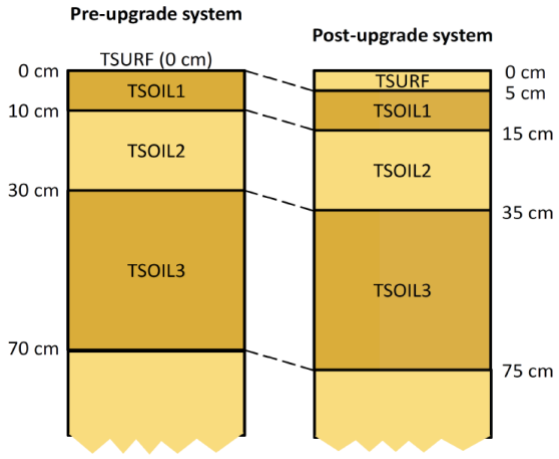
Atmospheric modeling systems discretize the soil column into a finite set of nodes for solving the heat diffusion and advection in the soil vertical profile. The number of nodes is usually limited (typically in the two to five range). Therefore, the emission temperature has to be parameterized in terms of a finite number of soil layers. The Choudhury et al. (1982) model is applicable to estimating an emission temperature based on finite soil temperature estimates within the profile. An emission temperature is defined based on two soil temperature values – one near the surface T_{surf} and another deep in the soil T_{deep} – such that:

$$T = T_{deep} + C \cdot \left(T_{surf} - T_{deep}\right)$$

The factor C primarily depends on wavelength λ but can also be influenced by soil texture, time-of-day, and soil moisture and temperature profile shapes. Choudhury et al. (1982) used the coherent radiative transfer model of Wilheit (1978) in conjunction with observed soil moisture temperature and moisture profiles at a site in Phoenix, Arizona in order to estimate the value of factor C. The soil temperature and moisture observations are at multiple depths, allowing definition of a vertical profile. The site in Arizona experienced large variations in diurnal soil temperature profiles. Since the site is in an arid region and was sporadically irrigated, the soil moisture profile conditions also spanned a large range. Based on the coherent radiative transfer model and the soil profile observations, Choudhury et al. (1982) reported that at L-band $\lambda = 21\ [cm]$, $C = 0.246 \pm 0.009$. SMAP has adopted this value for estimating an emission temperature T based on T_{surf} and T_{deep} ancillary data inputs (current baseline SMAP L2/L3 products have all used $C = 0.246$ for the descending 6 AM overpass time since the start of SMAP science data acquisition). The SMAP Level 1 science requirements for the accuracy of retrieved surface soil moisture products are defined only for descending overpass times (about 6 AM local time). At this time the soil and overlying vegetation temperatures are least divergent compared to the rest of the day (near isothermal condition) which allows simplification of the radiative transfer model at the core of the retrieval algorithm. The ~6 PM ascending overpass data are similarly used to produce estimates of surface soil moisture for 6 PM, but these retrievals are not strictly required to meet the same level of accuracy as the 6 AM retrievals. Not having an accuracy requirement also means that 6 PM soil moisture retrievals are not a driver of algorithm design and independent ground truth science product assessment and validation.

II. Source of Dynamic Soil Temperature Ancillary Data

For SMAP the source of ancillary soil temperature data is the GMAO GEOS-FP system. In January 2020 the GEOS-FP system upgraded to a new land surface model that redefined the surface and soil temperature layer depths as shown below (Koster et al., 2020):



In the upgraded GEOS system, $Tsurf$ is a temperature prognostic that is formed based on bare soil, vegetation, and snowpack temperatures in proportion to their prevalence across the pixel area. This temperature is a flux temperature which is used in the gradient-driven exchanges of heat between the surface and the atmospheric boundary layer. $TSOIL1$ and $TSOIL2$ are consistently soil media temperature prognostic variables that are characterized by these extensive attributes:

	Depth Range	Node Center Position
$TSOIL1$	-5 to -15 [cm]	-10 [cm]
$TSOIL2$	-15 to -35 [cm]	-25 [cm]

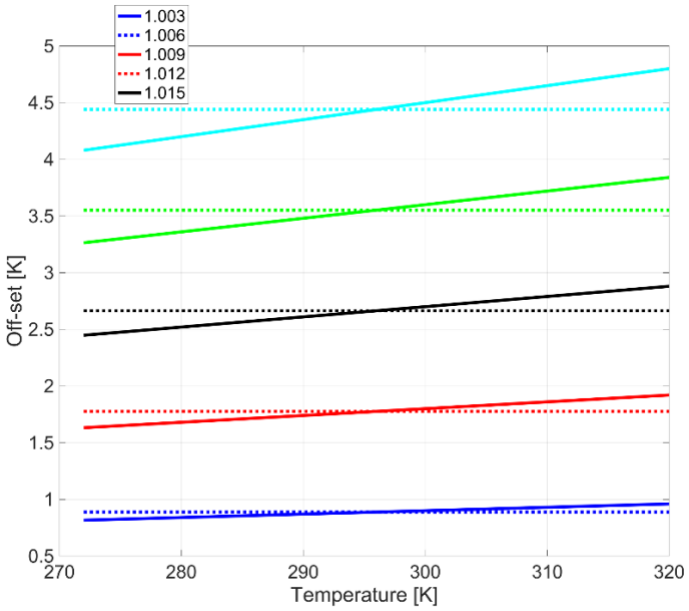
These GEOS temperature variables ($TSOIL1$ and $TSOIL2$) need to be linked to the Choudhury et al. (1982) model variables (T_{surf} and T_{deep}).

III. Offset and Scaling Biases in Model Prognostics

Atmospheric and land surface model forecasts of soil temperatures may be biased (model structural biases) and may need corrections to represent soil temperatures T_{surf} and T_{deep}. Such structural biases are inevitable given the fact that they are parameterized and the result of numerical models as well as because the model grids (in the horizontal) are often tens of kilometers which span large heterogeneities in landscape conditions like slope, aspect, land use, and soil texture.

The model variables can be biased (offset) or they can be modulated (scaled). The offset and scaling may be seasonally variable and depend on land use and soil physical characteristics. Identification of the required scaling (modulation of the seasonal and diurnal amplitude) and its dependence on landscape characteristics is complicated and beyond the immediate priorities of the SMAP Project. The offset, however, is identifiable with the use of *in situ* soil temperature measurements.

To start this analysis, the $TSOIL1$ variable at the time of the descending overpass was compared with the most common measurement of soil temperature at *in situ* stations, i.e., $T_g(z = -5\ [cm])$. The offset between the two variables may be captured by an additive factor or a multiplicative factor. The figure below over the range of typical non-frozen soil temperatures shows that when the temperatures are given in units of [Kelvin], the additive and multiplicative characterization of the offset factor is comparable.



Based on the above figure, the decision was made to proceed with a multiplicative model for the offset so that:

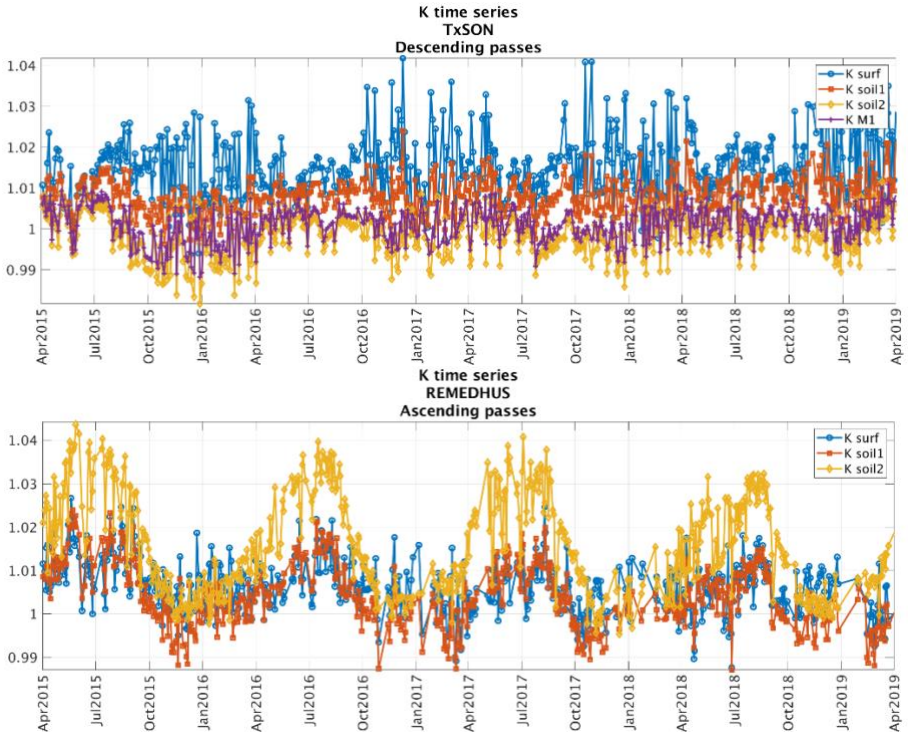
$$T_{surf} = K_{surf} \cdot TSOIL1$$

IV. Descending Overpass Time

Based on ground-based observations of soil temperature $T_g(z = -5\,[cm])$ and GMAO GEOS-FP $TSOIL1$, the values of K_{surf} across CVS and sparse network ground stations are:

Descending Overpasses (~6:00 am Local Time)		
	K_{surf}	Number of Stations
CVS (Chaubell, 2020, personal communication)	1.006 +/- 0.005	15
USCRN (Colliander, 2020, personal communication)	1.0068 +/- 0.0051	299

Colliander (2020) and Chaubell (2020) show that K_{surf} has a distinct seasonal cycle at some stations (peak in the summer) and that drier soil climates have higher K_{surf} values. Examples at the TxSON and REMEDHUS stations (in red) are shown below (Chaubell, 2020):



Repeat of the same offset identification but for the deep soil temperature, i.e.

$$T_{deep} = K_{deep} \cdot TSOIL2$$

has more options to consider. The first option is to estimate K_{deep} in the same manner as for the surface soil but using deeper ground-based observations of soil temperature such as $T_g(z = -20\ [cm])$. The availability of ground-based soil temperature at this depth is more limited. But a more serious consideration is the potential hazard of disturbing the soil temperature gradient provided by the GMAO GEOS-FP system if separate multiplicative offset factors are applied.

As a solution, offsets are applied to the GMAO GEOS-FP soil temperature prognostics that preserve the gradient but shift the profile so that they match the surface ground-based soil temperature measurements. This implies $K_{deep} = K_{surf} = K$. With this approach the GMAO GEOS soil temperature gradient is not disturbed and the shift is a translation on the temperature scale.

Based on the table above, $K = 1.007$ is selected together with the Choudhury et al. (1982) $C = 0.246$ for the 6 AM descending passes according to the combined model:

$$T = K \cdot [C \cdot TSOIL1 + (1 - C) \cdot TSOIL2]$$

V. Ascending Overpass Time

There are no SMAP project requirements on the accuracy of soil moisture retrievals using ascending pass 6 PM data. Nevertheless, the project makes every effort to provide accurate soil moisture retrievals from the 6 PM SMAP observations as well as from 6 AM to essentially double the soil moisture estimates available to the science and applications community. There are known challenges to applying the retrieval

algorithm to ascending data since isothermal conditions between the soil and the overlying vegetation canopy is one of the major requirements of the soil moisture retrieval algorithm as it is formulated now, and this condition is more likely to be violated at 6 PM.

The SMAP project has identified two major use-cases for the ascending soil moisture retrievals: 1) those that are interested in the diurnal cycle of soil moisture (in response to daytime drying and nighttime re-moistening by hydraulic lift of soil water from below to the drier surface) during non-precipitating days, and 2) those who seek to double their data count for more robust time-series analyses. These two use-cases lead to conflicting design criteria: the first use-case requires detection of the subtle small changes associated with the diurnal cycle of soil moisture, while the second use-case requires much consistency between ascending and descending estimates, i.e., no saw-tooth patterns.

Meeting the requirements of the first use-case is very challenging since characterization of the temperature gradient between the soil and the canopy – to the accuracy required to detect subtle diurnal soil moisture changes – is not readily feasible. This use-case community is encouraged to access the SMAP Level 1 brightness temperature data and design their own retrieval algorithms that might better suit their particular application (including the use of alternative ancillary data sources).

The second use-case can be enabled by ensuring an ascending (6 PM) data product that best meets two criteria: 1) best possible performance in terms of matching *in situ* soil moisture at the time of the evening overpass, and 2) as much as possible tracking the descending soil moisture product except for when precipitation events occur in between the two measurements.

The offset factor for the ascending overpass time (6 PM) may be approached in the same manner as for the descending (6 AM) overpasses described above. The results of this estimation process are listed in the table below.

Ascending Overpasses (~6 PM Local Time)		
	K_{surf}	Number of Stations
CVS (Chaubell, 2020, personal communication)	1.002 +/- 0.007	15
USCRN (Colliander, 2020, personal communication)	1.0024 +/- 0.0071	299

Since the ascending overpass time is not part of the SMAP Science Requirements, there is greater latitude in selecting the K and C parameters, with the possibility that the nominal values from the current analysis will be revisited in the future.

For the parameter C for the 6 PM ascending overpasses, the project plans to continue to use $C = 1$, which is the same as the value used to produce the baseline surface soil moisture science data product publicly available now. This value was selected based on the understanding of the evolution of canopy and soil profile temperatures during the course of the solar day. It should be revisited with more evidence based on radiative transfer modeling and *in situ* observations of canopy and multi-layer soil temperatures.

For the parameter K, a lower value than the descending pass estimated value is warranted. At this time it was decided to use the same K during the ascending time as the descending time, i.e., $K = 1.007$, until the fidelity of the forecasts of the diurnal cycle amplitude in the GMAO GEOS-FP system is better understood. This diurnal cycle amplitude is certainly related to the vegetation cover, soil moisture level,

season, and available energy and other environmental factors. It is prudent to not assign a different offset to the ascending overpass time model temperatures until there is greater understanding of the role of the amplitude of the diurnal cycle.

	K	C
Descending	1.007	0.246
Ascending	1.007	1.000

VI. Assessments

In the analysis described here, the soil emission temperature parameter estimation has only used *in situ* soil temperatures and model temperature prognostics. This leaves CVS and sparse network soil moisture information intact for use invalidation of the SMAP retrieved soil moisture values. The table below shows the assessment of the SMAP surface soil moisture product at CVS using Plan 2.2 of the L1 brightness temperature calibration as well as the new OpenLandMap soil texture database. The results of the independent assessment are:

	ubRMSE (m3 m-3)	Bias (m3 m-3)	RMSE (m3 m-3)	Correlation
Descending Orbits	0.037	-0.006	0.048	0.816
Ascending Orbits	0.037	-0.008	0.048	0.811

The random errors (ubRMSE) and correlation are unaffected by the updated parameters of the soil emission model and updated GMAO GEOS-FP soil temperature dynamic soil temperature ancillary data inputs. Although slightly larger than before the updates, the bias is small and within the bounds of uncertainty (~0.01) in determining grid-average values based on finite point samples (Chen et al., 2019).

Two major use-cases were identified for the combined descending and ascending surface soil moisture products. Besides product performance based on 6 AM/6 PM local time *in situ* soil moisture ground truth, efforts were made to minimize spurious 'saw-tooth' variations beyond normal differences in the two overpass soil moisture estimates. This saw-tooth pattern would manifest itself in the mean bias between the ascending and descending soil moisture as well as in other statistics. With the current selection of soil emission parameters, the bias between the descending and ascending surface soil moisture estimates averaged over all CVS is small (around -0.002 m^3 m^{-3}).

REFERENCES

A1. Chen, F. et al., "Uncertainty of Reference Pixel Soil Moisture Averages Sampled at SMAP Core Validation Sites," *J. Hydrometeorology*, 20, 2019.

A2. Choudhury, B., T. J. Schmugge and T. Mo, "A Parameterization of Effective Soil Temperature for Microwave Emission," *Journal of Geophysical Research*, 87(C2): 1301-1304, 1982.

A3. Koster, R. et al., "Land-Focused Changes in the Updated GEOS-FP System (Version 5.25)," *GMAO Research Brief*, 11 pgs, 2020.

A4. Lv, S. et al., "An Improved Two-Layer Algorithm for Estimating Effective Soil Temperature in Microwave Radiometry using *In Situ* Temperature and Soil Moisture Measurements," *Remote Sensing of Environment*, vol. 152, pp. 356-363, 2014.

A5. Lv, S. et al., “A Reappraisal of Global Soil Effective Temperature Schemes,” *Remote Sensing of Environment*, vol. 183, pp. 144-153, 2016.

A6. Wilheit, T., “Radiative Transfer in a Plane Stratified Dielectric,” *IEEE Transactions on Geoscience Electronics,* 16(2), 1978.

APPENDIX 2: UNMIXING OF SURFACE-CORRECTED BRIGHTNESS TEMPERATURES IN THE SMAP LEVEL 1C GRIDDED BRIGHTNESS TEMPERATURE PRODUCT

A2.1. INTRODUCTION

Surface-based radiometric correction was a key update introduced in Version 4 of the SMAP Level 1B Time-Ordered Brightness Temperature Product ("SPL1BTB") in 2018. This correction procedure was intended to mitigate the impacts of contaminating radiometric sources on SMAP brightness temperature ($\boldsymbol{T_B}$) observations. Specifically, it dealt with (1) $\boldsymbol{T_B}$ correction over land by removing emission contribution of nearby water and (2) $\boldsymbol{T_B}$ correction over water by removing emission contribution of nearby land. Around coastal areas and areas surrounding inland open water bodies (e.g. lakes, rivers, etc.), this correction procedure is expected to improve the retrieval accuracy of soil moisture over land and ocean salinity over water.

Which type of correction is to be performed within the SPL1BTB processor is determined by the boresight location of a particular time-ordered $\boldsymbol{T_B}$ field-of-view (FOV) sample. When the boresight falls onto a location considered as land according to a high-resolution static land/water mask, $\boldsymbol{T_B}$ correction over land for water contamination will be performed as long as the water fraction within the FOV is below 0.90, and when the boresight falls onto a location considered as water as determined by the same static mask, $\boldsymbol{T_B}$ correction over water for land contamination will be performed as long as the land fraction within the FOV is below 0.90. The resulting water-corrected $\boldsymbol{T_B}$ and land-corrected $\boldsymbol{T_B}$ data fields are available in SPL1BTB, in addition to the original uncorrected $\boldsymbol{T_B}$. **Table A2.1** summarizes the types, triggers, and correction sources considered by these correction schemes as of Version 4 of SPL1BTB:

Table A2.1: Surface-based radiometric correction implemented in SPL1BTB.

Correction Type	Retrieval	Boresight	Trigger	Correction Source
$\boldsymbol{T_B}$ correction over land to remove water contamination	Soil Moisture	Land (surface_status = 0)	FOV water fraction ≤ 0.9	Nearby 3° × 3° area-averaged SMAP $\boldsymbol{T_B}$ over water
$\boldsymbol{T_B}$ correction over water to remove land contamination	Salinity	Water (surface_status = 1)	FOV land fraction ≤ 0.9	Nearby 1° × 1° area-averaged SMAP $\boldsymbol{T_B}$ over land

A more detailed description of this surface-based radiometric correction procedure can be found in:

Chaubell, M.J., S.H. Yueh, R.S. Dunbar, S.K. Chan, F. Chen, J. Piepmeier, R. Bindlish, D. Entekhabi, P.E. O'Neill, "Improving Brightness Temperature Measurements Near Coastal Areas for SMAP," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 4578—4588, 12(11), 2019.

The SMAP Level 1C Gridded Brightness Temperature Product ("SPL1CTB") processor ingests the time-ordered data fields from SPL1BTB and grids them on the EASE Grid 2.0 global (M), north polar (N), and south polar (S) projections. The following discussion applies to how SPL1BTB data fields were and are binned by the SPL1CTB processor under all three projections.

Between Versions 4 (2018) and 5 (2020), the SPL1CTB processor binned water-corrected $\boldsymbol{T_B}$ and land-corrected $\boldsymbol{T_B}$ from SPL1BTB together when they fall into the same EASE Grid 2.0 grid cell. This mixing is undesirable because the resulting gridded $\boldsymbol{T_B}$ would undo the intended purpose of $\boldsymbol{T_B}$ correction over land

for water contamination and T_B correction over water for land contamination implemented in the SPL1BTB processor. Nonetheless, this indiscriminate binning was decided and implemented out of necessity, given the constraints of processing infrastructure available at that time. Note that this indiscriminate binning affected only the SPL1CTB product; its enhanced counterpart ("SPL1CTB_E") was not affected. The rest of this memo deals with our attempt to undo this mixing of water-corrected T_B and land-corrected T_B within the SPL1CTB processor.

A2.2. REMEDY

The SPL1CTB processor has recently gone through a major code overhaul to implement an algorithm-based workaround to circumvent the limitations of the previous indiscriminate binning. The fix involved creation of extra "containers" during the binning process to separate the binning of water-corrected T_B from the binning of land-corrected T_B. The binned data fields from these "containers" remain available until the end of processing where the SPL1CTB processor will decide if a given grid cell is considered land or water based on the value of the nearest-neighbor SPL1BTB's surface_status field. If the nearest-neighbor surface_status field is 0 (**Table A2.1**, 2nd row), the processor will retrieve the "container" for water-corrected T_B and drop the corresponding binned result in the final output field. Conversely, if the nearest-neighbor surface_status field is 1 (**Table A2.1**, 3rd row), SPL1CTB will retrieve the "container" for land-corrected T_B and drop the corresponding binned result in the final output field. In other words, the nearest-neighbor surface_status field of SPL1BTB was used as a transfer platform to enable the SPL1CTB processor to copy the input high-resolution static land/water mask available only to the SPL1BTB processor.

This code overhaul was implemented entirely internally inside the SPL1CTB processor and thus did not require any sophisticated external processing configurations. **Figure A2.1** describes how the indiscriminate binning in SPL1CTB since Version 4 was revised for the next data release of the product.

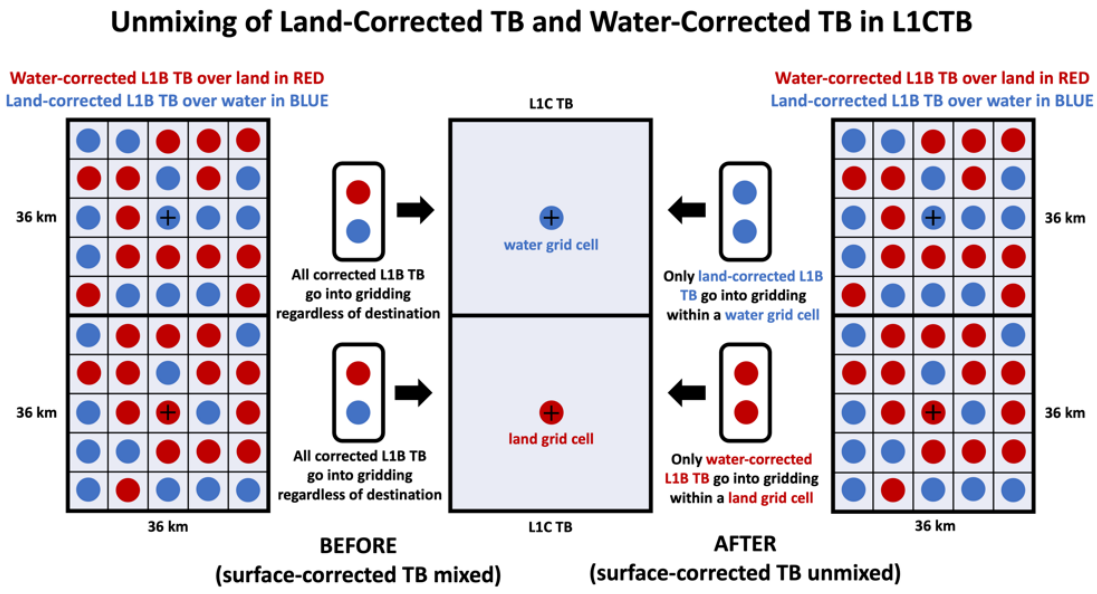


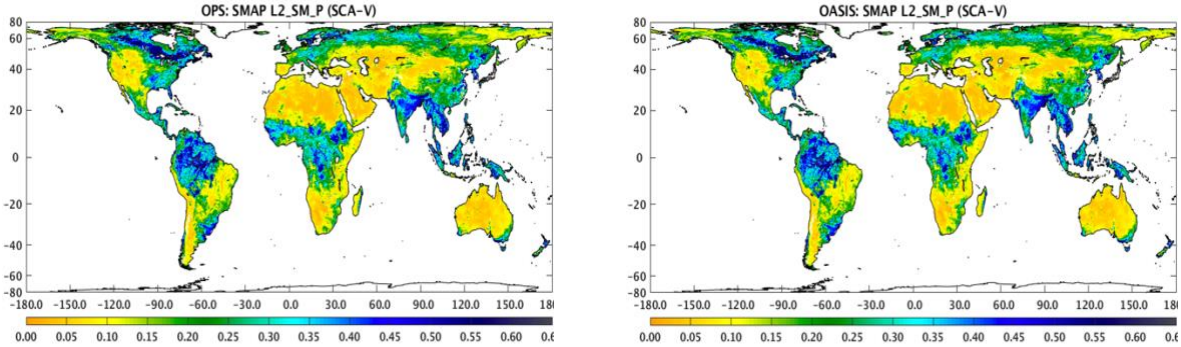
Figure A2.1: Separate binning of water- and land-corrected T_B (right) for the next data release of SPL1CTB. After unmixing, only water-corrected T_B samples go to land grid cells (red) and only land-corrected T_B samples go to water grid cells (blue).

It is evident that the revised binning scheme would preserve the water-corrected T_B and land-corrected T_B from SPL1BTB with a higher fidelity. Given the triggers (**Table A2.1**, 4th column) for this surface correction procedure, the improvements will be most visible along boundaries that separate land from water. In general, they correspond to coastal areas and areas surrounding inland open water bodies.

A few observations on the accuracy of the water-corrected T_B from SPL1BTB follow. From **Table A2.1**, the correction source for T_B correction over land is the average SMAP T_B over water within a 3° × 3° latitude/longitude search box. This coverage is about 9 times greater than what is covered by the main beam (2.5 × HPBW ~ 108 km) of the SMAP antenna. As such, the corresponding correction may involve T_B observations not directly contributing to a single time-ordered T_B field-of-view (FOV) sample, potentially leading to correction errors that translate to subsequent soil moisture retrieval bias. One example of this processing artifact can be seen over Finland in **Figure A2.4b**, where a 3° × 3° latitude/longitude search box at this high latitude would introduce SMAP T_B dominated by the cold water T_B in the vertical dimension. The resulting water-corrected T_B over land after water T_B subtraction may become over-corrected (i.e. higher than reality), leading to very dry soil moisture retrieval. The limitations of this latitude/longitude search approach can be mitigated by a search domain that coincides with the actual FOV footprint projected on the land surface.

A2.3. IMPACTS ON SOIL MOISTURE RETRIEVAL

The SMAP Level 2 Radiometer-only Soil Moisture Product ("SPL2SMP") uses the water-corrected T_B over land from SPL1CTB for soil moisture retrieval. The revised binning described in Section 2 is expected to more thoroughly remove the contribution by water to SMAP T_B, resulting in higher T_B for soil moisture retrieval. A comparison in soil moisture retrieval before and after this unmixing is expected to reveal (1) lower soil moisture around coastal areas and areas surrounding inland open water bodies and (2) little change in retrieval elsewhere where the triggers in **Table A2.1** are not set. **Figures A2.2 and A2.3** confirm these expectations. By and large, **Figure A2.2a** and **A2.2b** are visually identical. Their difference (**Figure A2.3)**, however, reveals dry bias in retrieval around coastal areas and areas surrounding inland open water bodies.



(a) (b)

Figure A2.2: 6:00 am SPL2SMP between June 1-5, 2020 using SPL1CTB gridded vertically polarized T_B (a) before the unmixing of surface-corrected T_B and (b) after unmixing of surface-corrected T_B.

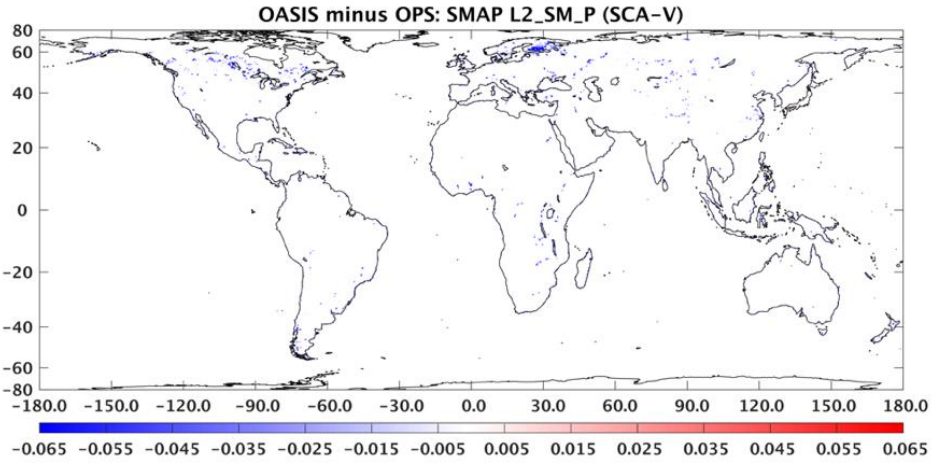
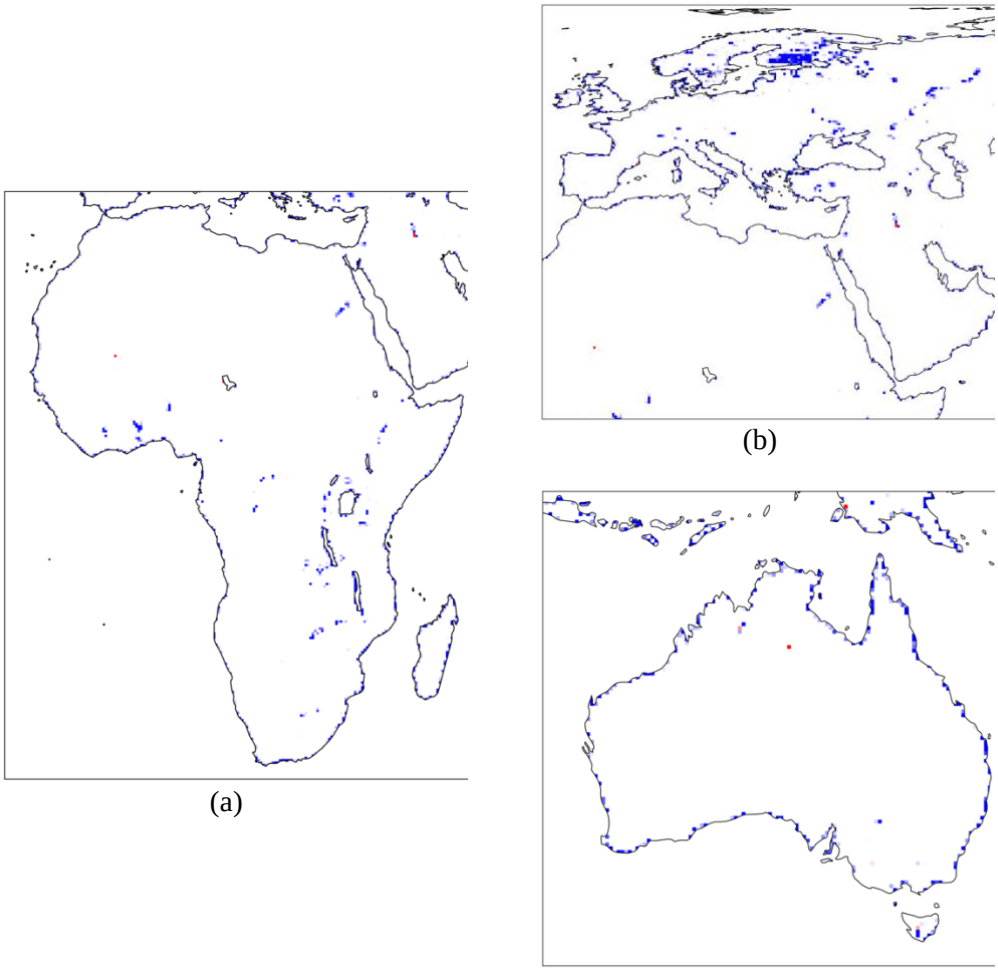


Figure A2.3: Difference in SPL2SMP soil moisture retrieval before and after the unmixing of surface-corrected T_B.

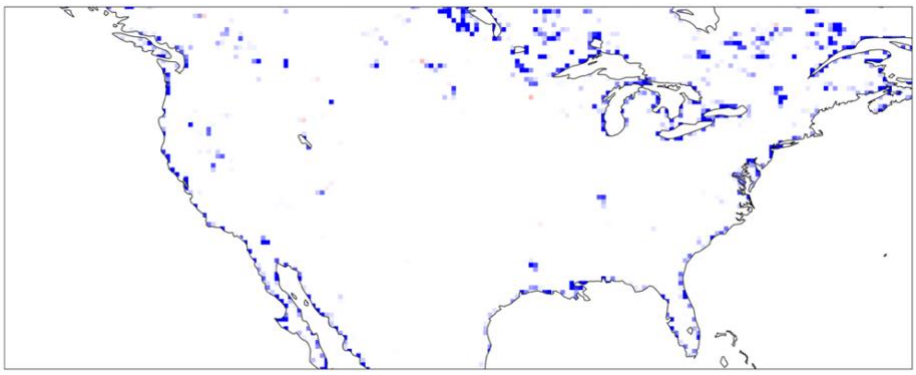
Figure A2.3 can be further cropped over different geographical regions for a better visualization of the regional retrieval difference before and after unmixing. This is shown in **Figure A2.4**.



(a)

(b)

(c)



(d)

Figure A2.4: Difference in SPL2SMP soil moisture retrieval before and after the unmixing of surface-corrected T_B over (a) Africa, (b) Europe, (c) Australia, and (d) CONUS.

Figure A2.3 also indicates the very small percentage of land pixels that see a change in soil moisture retrieval before and after the unmixing of surface-corrected T_B. The histogram in **Figure A2.5** shows the distribution of retrieval difference in percentage upon a consideration of all global land pixels.

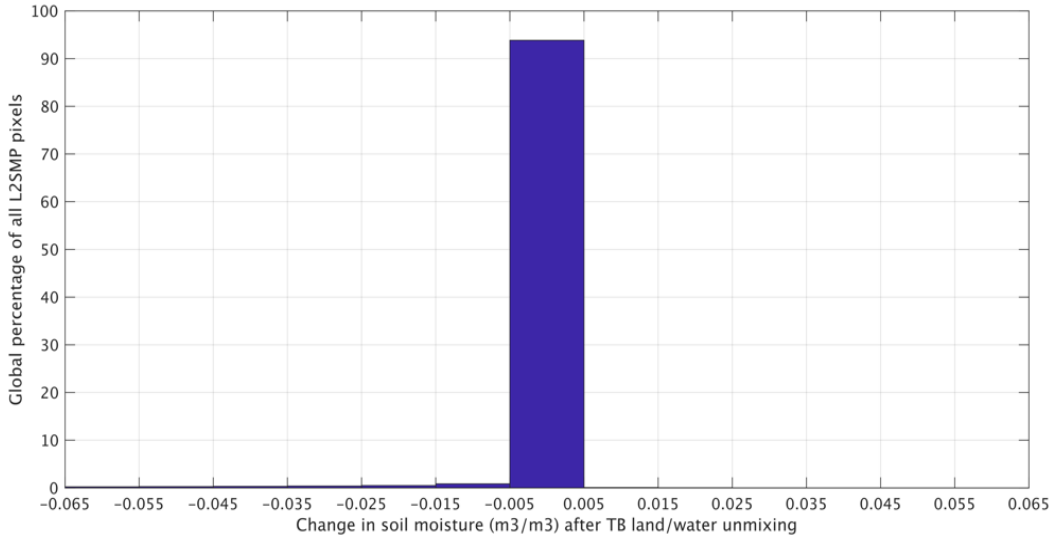


Figure A2.5: Histogram of retrieval difference before and after the unmixing of surface-corrected T_B.

It is clear from **Figure A2.5** that the histogram exhibits a tail that mostly extends to the negative half of the x-axis, implying the lower soil moisture retrieval after the unmixing of surface-corrected T_B, which is consistent with the discussion earlier in this section. Most global land pixels, on the other hand, do not see any change in soil moisture retrieval, as indicated by the overwhelmingly high percentage in the middle

of the histogram. **Table A2.2** presents a numerical breakdown of **Figure A2.5**. Further calculations indicated that 92.9% of all global land pixels show a retrieval difference of less than ±0.001 m^3/m^3.

Table A2.2: Numerical breakdown of soil moisture retrieval difference in Figure A2.5.

Soil Moisture Range (m^3/m^3)		Percent (%)
−0.065	−0.055	0.23
−0.055	−0.045	0.28
−0.045	−0.035	0.33
−0.035	−0.025	0.40
−0.025	−0.015	0.50
−0.015	−0.005	0.87
−0.005	+0.005	93.84
+0.005	+0.015	0.08
+0.015	+0.025	0.04
+0.025	+0.035	0.02
+0.035	+0.045	0.01
+0.045	+0.055	0.01
+0.055	+0.065	0.00

A2.4. CONCLUSION

The unmixing of surface-corrected $\boldsymbol{T_B}$ described in this memo provides an improved mechanism to preserve water-corrected $\boldsymbol{T_B}$ and land-corrected $\boldsymbol{T_B}$ with a higher fidelity. The resulting surface-corrected $\boldsymbol{T_B}$ fields will be available beginning in Version 6 of SPL1CTB and Version 8 of SPL2SMP in the upcoming R18 data release expected in 2021. The impacts of unmixing are observable mostly around coastal areas and areas surrounding inland open water bodies. 92.9% of all global land pixels show a retrieval difference of less than ±0.001 m^3/m^3 after unmixing, confirming the limited scope of this procedure on the majority of global land pixels and thus ensuring the continuing consistency of the upcoming SPL2SMP baseline soil moisture retrieval.

APPENDIX 3: ALGORITHM VARIABLE POINTER TABLES FOR L2/L3 PASSIVE SOIL MOISTURE PRODUCT OUTPUT FIELDS

These tables are provided as a quick reference guide for users of these soil moisture products to associate the correct variable output fields with the correct retrieval algorithm, and to associate the pointer elements to the correct variable for the L2/L3 passive soil moisture products in the SMAP R18 data release of October 2021. More detailed information can be found in the L2 Passive Soil Moisture ATBD [30] and the appropriate SMAP User Guides and Product Specification Documents available at NSIDC (https://nsidc.org/data/smap/technical-references).

For the L2_SM_P/P_E soil moisture products:

List of Variables Associated with Each Soil Moisture Algorithm Option		
DCA	**SCA-V**	**SCA-H**
soil_moisture_option3^	soil_moisture_option2	soil_moisture_option1
vegetation_opacity_option3^	vegetation_opacity_option2	vegetation_opacity_option1
retrieval_quality_flag_option3^	retrieval_quality_flag_option2	retrieval_quality_flag_option1
roughness_coefficient_option3	roughness_coefficient*	roughness_coefficient*
albedo_option3	albedo*	albedo*
* the SCA-H and SCA-V algorithms share the same roughness_coefficient and albedo variables		
^ these variables also correspond to the pointer elements listed in the table below		

List of Pointer Elements and Their Corresponding Variables	
Pointer Element	**Variable**
soil_moisture	soil_moisture_option3
vegetation_opacity	vegetation_opacity_option3
retrieval_quality_flag	retrieval_quality_flag_option3

For the L3_SM_P/P_E soil moisture products:

List of Variables Associated with Each Soil Moisture Algorithm Option		
DCA	**SCA-V**	**SCA-H**
soil_moisture_dca^	soil_moisture_scav	soil_moisture_scah
vegetation_opacity_dca^	vegetation_opacity_scav	vegetation_opacity_scah
retrieval_quality_flag_dca^	retrieval_quality_flag_scav	retrieval_quality_flag_scah
roughness_coefficient_dca^	roughness_coefficient_scav	roughness_coefficient_scah
albedo_dca^	albedo_scav	albedo_scah
^ these variables also correspond to the pointer elements listed in the table below		
NOTE: The variable names in the Soil_Moisture_Retrieval_Data_PM data group have the suffix "_pm" attached.		

List of Pointer Elements and Their Corresponding Variables	
Pointer Element	**Variable**
soil_moisture[_pm]	soil_moisture_dca[_pm]
vegetation_opacity[_pm]	vegetation_opacity_dca[_pm]
retrieval_quality_flag[_pm]	retrieval_quality_flag_dca[_pm]
roughness_coefficient[_pm]	roughness_coefficient_dca[_pm]
albedo[_pm]	albedo_dca[_pm]