Soil Moisture Active Passive (SMAP)

Algorithm Theoretical Basis Document
Level 2 & 3 Soil Moisture (Passive) Data Products

Revision C
December 15, 2016

Peggy O’Neill
NASA Goddard Space Flight Center
Greenbelt, MD

Steven Chan, Eni Njoku
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA

Tom Jackson, Rajat Bindlish
USDA Agricultural Research Service
Beltsville, MD

JPL D-66480

Jet Propulsion Laboratory
California Institute of Technology

The SMAP Algorithm Theoretical Basis Documents (ATBDs) provide the physical and mathematical descriptions of algorithms used in the generation of SMAP science data products. The ATBDs include descriptions of variance and uncertainty estimates and considerations of calibration and validation, exception control and diagnostics. Internal and external data flows are also described.

The SMAP ATBDs were reviewed by a NASA Headquarters review panel in January 2012 with initial public release later in 2012. The current version of this ATBD is Revision C. The ATBDs may undergo additional version updates during the mission.

Revision C dated December 15, 2016 contains the following updates from Revision B dated September 14, 2015:

1. Added new ATBDs to list on p. 6.
2. Added updated SMAP Data Products Table on p. 13.
3. Added statement pointing to Sec. 8.1 regarding 6 pm soil moisture retrievals on p. 23.
4. Added statements about 6 pm soil moistures and the L2_SM_P Data Release Version 4 assessment report on p. 25.
5. Added statement that 6 pm compositing uses the same compositing procedure as the 6 am data on p. 29.
6. Added references to the L2_SM_P Data Release Version 3 and L2_SM_P Version 4 /L2_SM_P_E Version 1 assessment reports and to the L2_SM_P journal article on p. 66 and p. 77.
7. Added Section 8.1 on 6 pm soil moisture retrievals and Section 8.2 on the L2_SM_P_E soil moisture retrievals on p. 72.

TABLE OF CONTENTS

SMAP Reference Documents

Requirements:

- 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.
- SMAP Level 2 Science Requirements. SMAP Project, JPL D-45955, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Level 3 Science Algorithms and Validation Requirements. SMAP Project, JPL D-45993, Jet Propulsion Laboratory, Pasadena, CA.

Plans:

- SMAP Science Data Management and Archive Plan. SMAP Project, JPL D-45973, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Science Data Calibration and Validation Plan. SMAP Project, JPL D-52544, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Applications Plan. SMAP Project, JPL D-53082, Jet Propulsion Laboratory, Pasadena, CA.

ATBDs:

- SMAP Algorithm Theoretical Basis Document: L1B and L1C Radar Products. SMAP Project, JPL D-53052, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: L1B Radiometer Product. SMAP Project, GSFC-SMAP-006, NASA Goddard Space Flight Center, Greenbelt, MD.
- SMAP Algorithm Theoretical Basis Document: L1C Radiometer Product. SMAP Project, JPL D-53053, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: L2 & L3 Radar Soil Moisture (Active) Products. SMAP Project, JPL D-66479, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: L2 & L3 Radiometer Soil Moisture (Passive) Products. SMAP Project, JPL D-66480, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: L2 & L3 Radar/Radiometer Soil Moisture (Active/Passive) Products. SMAP Project, JPL D-66481, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: L3 Radar Freeze/Thaw (Active) Product. SMAP Project, JPL D-66482, Jet Propulsion Laboratory, Pasadena, CA.

- SMAP Algorithm Theoretical Basis Document: L4 Surface and Root-Zone Soil Moisture Product. SMAP Project, JPL D-66483, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: L4 Carbon Product. SMAP Project, JPL D-66484, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: Enhanced L1B_TB_E Radiometer Brightness Temperature Data Product. SMAP Project, JPL D-56287, Jet Propulsion Laboratory, Pasadena, CA.
- SMAP Algorithm Theoretical Basis Document: L3 Radiometer Freeze/Thaw (Passive) Product. SMAP Project, JPL D-56288, Jet Propulsion Laboratory, Pasadena, CA.

<u>Ancillary Data Reports</u>:

- Ancillary Data Report: Crop Type. SMAP Project, JPL D-53054, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Digital Elevation Model. SMAP Project, JPL D-53056, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Land Cover Classification. SMAP Project, JPL D-53057, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Soil Attributes. SMAP Project, JPL D-53058, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Static Water Fraction. SMAP Project, JPL D-53059, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Urban Area. SMAP Project, JPL D-53060, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Vegetation Water Content. SMAP Project, JPL D-53061, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Permanent Ice. SMAP Project, JPL D-53062, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Precipitation. SMAP Project, JPL D-53063, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Snow. SMAP Project, GSFC-SMAP-007, NASA Goddard Space Flight Center, Greenbelt, MD.
- Ancillary Data Report: Surface Temperature. SMAP Project, JPL D-53064, Jet Propulsion Laboratory, Pasadena, CA.
- Ancillary Data Report: Vegetation and Roughness Parameters. SMAP Project, JPL D-53065, Jet Propulsion Laboratory, Pasadena, CA.

ACRONYMS AND ABBREVIATIONS

AMSR	Advanced Microwave Scanning Radiometer
ATBD	Algorithm Theoretical Basis Document
CONUS	Continental United States
CMIS	Conical-scanning Microwave Imager Sounder
CV	Calibration / Validation
DAAC	Distributed Active Archive Center
DCA	Dual Channel Algorithm
DEM	Digital Elevation Model
ECMWF	European Center for Medium-Range Weather Forecasting
EOS	Earth Observing System
ESA	European Space Agency
GEOS	Goddard Earth Observing System (model)
GMAO	Goddard Modeling and Assimilation Office
GSFC	Goddard Space Flight Center
IFOV	Instantaneous Field Of View
JAXA	Japan Aerospace Exploration Agency
JPL	Jet Propulsion Laboratory
JPSS	Joint Polar Satellite System
LPRM	Land Parameter Retrieval Model
LSM	Land Surface Model
LTAN	Local Time Ascending Node
LTDN	Local Time Descending Node
MODIS	MODerate-resolution Imaging Spectroradiometer
NCEP	National Centers for Environmental Prediction
NDVI	Normalized Difference Vegetation Index
NPOESS	National Polar-Orbiting Environmental Satellite System
NPP	NPOESS Preparatory Project
NWP	Numerical Weather Prediction
OSSE	Observing System Simulation Experiment
PDF	Probability Density Function
PGE	Product Generation Executable
QC	Quality-Control
RFI	Radio Frequency Interference

RMSD	Root Mean Square Difference
RMSE	Root Mean Square Error
SCA	Single Channel Algorithm
SGP	Southern Great Plains (field campaigns)
SMAPVEX	SMAP Validation EXperiment
SMDPC	Soil Moisture Data Processing Center
SMEX	Soil Moisture EXperiments (field campaigns)
SMOS	Soil Moisture Ocean Salinity (ESA space mission)
T_B	Brightness Temperature
TBC	To Be Confirmed
TBD	To Be Determined
USDA ARS	U.S. Department of Agriculture, Agricultural Research Service
VWC	Vegetation water content (in units of kg/m^2)

1. INTRODUCTION

1.1 Background

The Soil Moisture Active Passive (SMAP) mission is the first of the Earth observation satellites being developed by NASA in response to the National Research Council's Earth Science Decadal Survey, *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond* [1]. The Decadal Survey was released in 2007 after a two-year study commissioned by NASA, NOAA, and USGS to provide them with prioritized recommendations for space-based Earth observation programs. Factors including scientific value, societal benefit, and technical maturity of mission concepts were considered as criteria. In 2008 NASA announced the formation of the SMAP project as a joint effort of NASA's Jet Propulsion Laboratory (JPL) and Goddard Space Flight Center (GSFC), with project management responsibilities at JPL. Launched on January 31, 2015, SMAP is providing high resolution global mapping of soil moisture and freeze/thaw state every 2-3 days on nested 3, 9, and 36-km Earth grids [2]. Its major science objectives are to:

- Understand processes that link the terrestrial water, energy and carbon cycles;
- Estimate global water and energy fluxes at the land surface;
- Quantify net carbon flux in boreal landscapes;
- Enhance weather and climate forecast skill;
- Develop improved flood prediction and drought monitoring capability.

1.2 Measurement Approach

Table 1 is a summary of the SMAP instrument functional requirements derived from its science measurement needs. The goal is to combine the attributes of the radar and radiometer observations (in terms of their spatial resolution and sensitivity to soil moisture, surface roughness, and vegetation) to estimate soil moisture at a resolution of 10 km and freeze-thaw state at a resolution of 1-3 km.

The SMAP instrument incorporates an L-band radar and an L-band radiometer that share a single feedhorn and parabolic mesh reflector. As shown in Figure 1, the reflector is offset from nadir and rotates about the nadir axis at 14.6 rpm (nominal), providing a conically scanning antenna beam with a surface incidence angle of approximately 40°. The provision of constant incidence angle across the swath simplifies the data processing and enables accurate repeat-pass estimation of soil moisture and freeze/thaw change. The reflector has a diameter of 6 m, providing a radiometer 3 dB antenna footprint of 40 km (root-ellipsoidal-area). The real-aperture radar footprint is 30 km, defined by the two-way antenna beamwidth. The real-aperture radar and radiometer data will be collected globally during both ascending and descending passes.

To obtain the desired high spatial resolution, the radar employs range and Doppler discrimination. The radar data can be processed to yield resolution enhancement to 1-3 km spatial resolution over the outer 70% of the 1000-km swath. Data volume constraints prohibit the downlinking of the entire radar data acquisition. Radar measurements that allow high-resolution processing will be collected during the morning overpass over all land regions and extending a short distance into the surrounding coastal oceans. During the evening overpass, data poleward of 45° N will be collected and processed as well to support robust detection of landscape freeze/thaw transitions. The SMAP baseline orbit parameters are:

- Orbit Altitude: 685 km (2-3 day average revisit globally and 8-day exact repeat)
- Inclination: 98 degrees, sun-synchronous
- Local Time of Ascending Node: 6 pm (6 am descending local overpass time)

Table 1. SMAP Mission Requirements

Scientific Measurement Requirements	Instrument Functional Requirements
Soil Moisture: ~± 0.04 cm^3/cm^3 volumetric accuracy (1-sigma) in the top 5 cm for vegetation water content ≤ 5 kg/m^2 Hydrometeorology at ~10 km resolution Hydroclimatology at ~40 km resolution	**L-Band Radiometer (1.41 GHz):** Polarization: V, H, T_3, and T_4 Resolution: 40 km Radiometric Uncertainty*: 1.3 K **L-Band Radar (1.26 and 1.29 GHz):** Polarization: VV, HH, HV (or VH) Resolution: 10 km Relative accuracy*: 0.5 dB (VV and HH) Constant incidence angle** between 35° and 50°
Freeze/Thaw State: Capture freeze/thaw state transitions in integrated vegetation-soil continuum with two-day precision at the spatial scale of landscape variability (~3 km)	**L-Band Radar (1.26 GHz & 1.29 GHz):** Polarization: HH Resolution: 3 km Relative accuracy*: 0.7 dB (1 dB per channel if 2 channels are used) Constant incidence angle** between 35° and 50°
Sample diurnal cycle at consistent time of day (6 am/6 pm Equator crossing); Global, ~3 day (or better) revisit; Boreal, ~2 day (or better) revisit	Swath Width: ~1000 km Minimize Faraday rotation (degradation factor at L-band)
Observation over minimum of three annual cycles	Baseline three-year mission life
* Includes precision and calibration stability ** Defined without regard to local topographic variation	

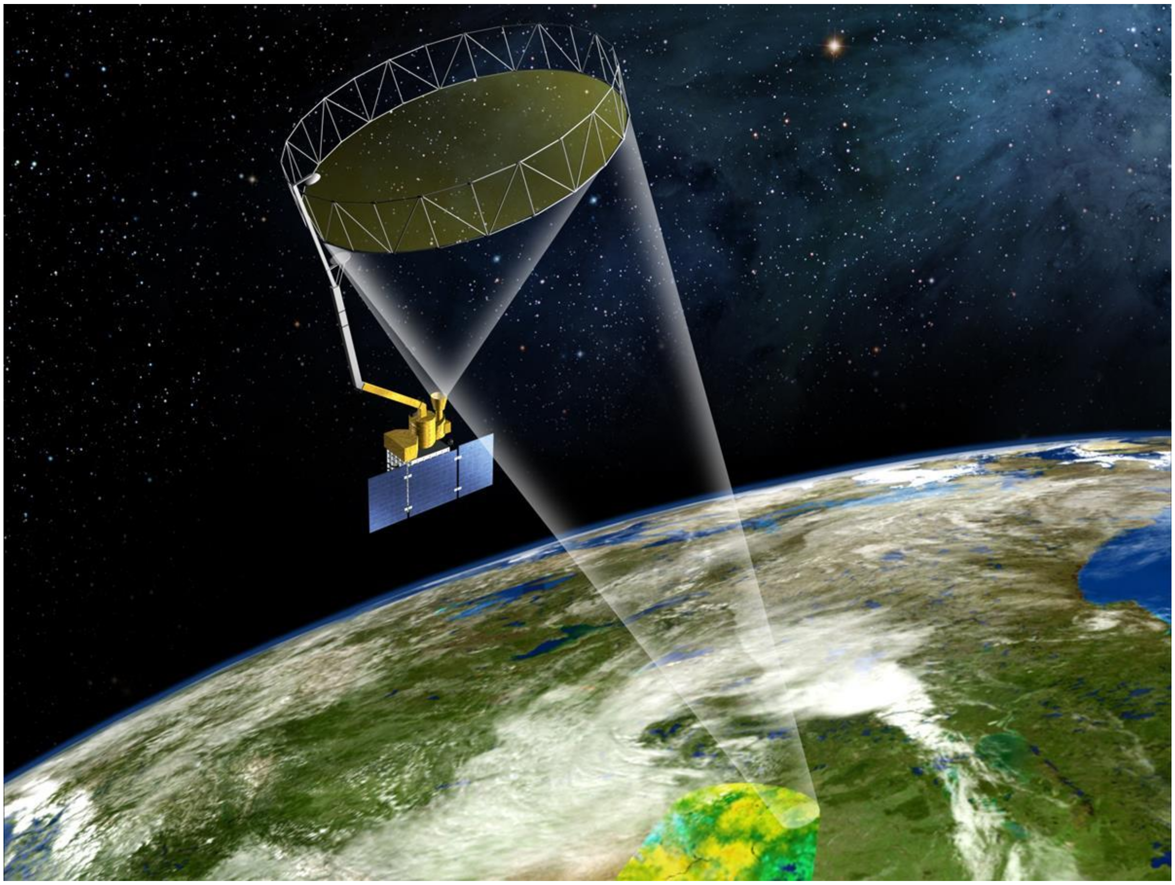


Figure 1. The SMAP mission concept consists of an L-band radar and radiometer sharing a single spinning 6-m mesh antenna in a sun-synchronous dawn / dusk orbit.

On July 7, 2015, the High Power Amplifier of the SMAP radar experienced an anomaly which caused the radar to stop transmitting. All subsequent attempts to power up the radar were unsuccessful to date. At this time the SMAP mission continues to produce high-quality science measurements supporting SMAP's objectives with its radiometer instrument.

The SMAP radiometer measures the four Stokes parameters, V, H, T_3, and T_4 at 1.41 GHz. The T_3-channel measurement can be used to correct for possible Faraday rotation caused by the ionosphere, although such Faraday rotation is minimized by the selection of the 6 am/6 pm sun-synchronous SMAP orbit.

Anthropogenic Radio Frequency Interference (RFI), principally from ground-based surveillance radars, can contaminate both radar and radiometer measurements at L-band. Early measurements and results from ESA’s SMOS (Soil Moisture and Ocean Salinity) mission indicate that in some regions RFI is present and detectable. The SMAP radar and radiometer electronics and algorithms have been designed to include features to mitigate the effects of RFI. The SMAP radar utilizes selective filters and an adjustable carrier frequency in order to tune to predetermined RFI-free portions of the spectrum while on orbit. The SMAP radiometer will implement a combination of time and frequency diversity, kurtosis detection, and use of T_4 thresholds to detect and where possible mitigate RFI.

SMAP observations will (1) improve our understanding of linkages between the Earth's water, energy, and carbon cycles, (2) benefit many application areas including numerical weather and climate prediction, flood and drought monitoring, agricultural productivity, human health, and national security, (3) help to address priority questions on climate change, and (4) potentially provide continuity with brightness temperature and soil moisture measurements from ESA's SMOS (Soil Moisture Ocean Salinity) and NASA's Aquarius missions. The current SMAP data products are listed in Table 2 (as of December, 2016). In the SMAP prelaunch time frame, baseline algorithms are being developed for generating (1) Level 1 calibrated, geolocated surface brightness temperature and radar backscatter measurements, (2) Level 2 and Level 3 surface soil moisture products both from radiometer measurements on a 36 km grid and from combined radar/radiometer measurements on a 9 km grid, (3) Level 3 freeze/thaw products from radar measurements on a 3 km grid, and (4) Level 4 surface and root zone soil moisture and Level 4 Net Ecosystem Exchange (NEE) of carbon on a 9 km grid. Level 1 data are the instrument products; Level 2 data are surface soil moisture in half-orbit format; Level 3 data are global daily composites of the Level 2 data; and Level 4 data combine the SMAP satellite observations with modeling to produce value-added products that support key SMAP applications and more directly address the driving science questions.

The details of each SMAP data product will be described in an associated publicly-available Algorithm Theoretical Basis Document (ATBD), which will be updated periodically as warranted. SMAP data products are generated using algorithm software that converts lower level products to higher level products. Each product has a designated baseline algorithm for its generation. One or more algorithm options may be encoded in the software and evaluated along with the baseline algorithm. The ATBDs describe the product algorithms and their implementation, pre-launch testing, and post-launch validation approaches.

1.3 Scope and Rationale

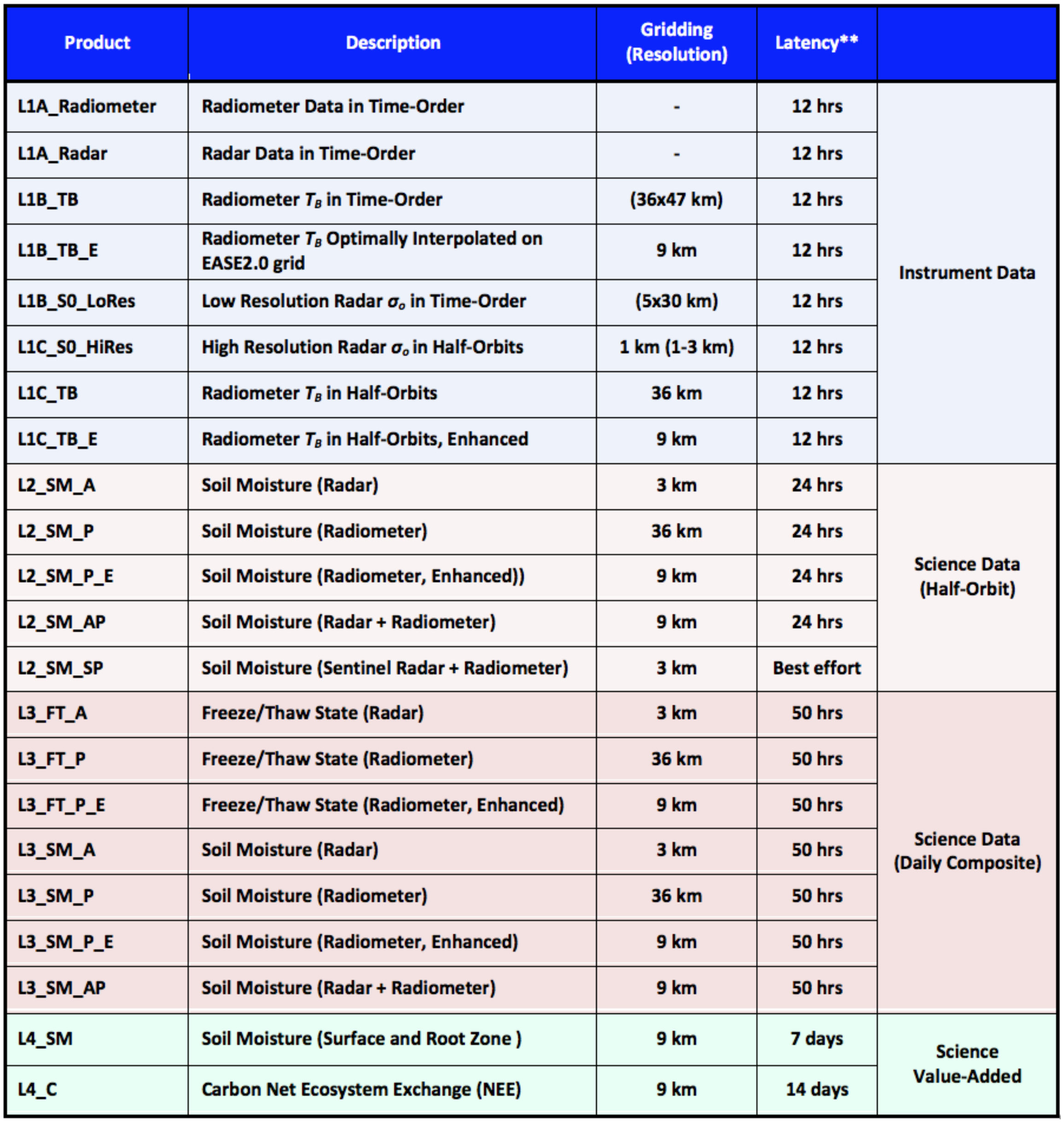
This document is the Algorithm Theoretical Basis Document (ATBD) for the SMAP radiometer-based surface soil moisture products:

1. Level 2 Soil Moisture (L2_SM_P) in half-orbit format.
2. Level 3 Soil Moisture (L3_SM_P) in the form of global daily composites.

The complete list of SMAP data products is provided in Table 2. The L2_SM_P and L3_SM_P products represent the surface soil moisture (0-5 cm layer) derived from the SMAP radiometer as output on a fixed 36-km Earth grid. This grid spacing is close to the approximate spatial resolution of 40 km of the SMAP radiometer footprint and permits nesting with the 3-km grid spacing of the SMAP radar-derived products and the 9-km grid spacing of the L2_SM_A/P combined active/passive product and the L4_SM and

L4_C products. As of December, 2016, L2/3_SM_P includes both AM and PM retrieved soil moisture using the same retrieval algorithms.

Table 2. SMAP Data Products



Product	Description	Gridding (Resolution)	Latency**	
L1A_Radiometer	Radiometer Data in Time-Order	-	12 hrs	Instrument Data
L1A_Radar	Radar Data in Time-Order	-	12 hrs	
L1B_TB	Radiometer T_B in Time-Order	(36x47 km)	12 hrs	
L1B_TB_E	Radiometer T_B Optimally Interpolated on EASE2.0 grid	9 km	12 hrs	
L1B_S0_LoRes	Low Resolution Radar σ_o in Time-Order	(5x30 km)	12 hrs	
L1C_S0_HiRes	High Resolution Radar σ_o in Half-Orbits	1 km (1-3 km)	12 hrs	
L1C_TB	Radiometer T_B in Half-Orbits	36 km	12 hrs	
L1C_TB_E	Radiometer T_B in Half-Orbits, Enhanced	9 km	12 hrs	
L2_SM_A	Soil Moisture (Radar)	3 km	24 hrs	Science Data (Half-Orbit)
L2_SM_P	Soil Moisture (Radiometer)	36 km	24 hrs	
L2_SM_P_E	Soil Moisture (Radiometer, Enhanced))	9 km	24 hrs	
L2_SM_AP	Soil Moisture (Radar + Radiometer)	9 km	24 hrs	
L2_SM_SP	Soil Moisture (Sentinel Radar + Radiometer)	3 km	Best effort	
L3_FT_A	Freeze/Thaw State (Radar)	3 km	50 hrs	Science Data (Daily Composite)
L3_FT_P	Freeze/Thaw State (Radiometer)	36 km	50 hrs	
L3_FT_P_E	Freeze/Thaw State (Radiometer, Enhanced)	9 km	50 hrs	
L3_SM_A	Soil Moisture (Radar)	3 km	50 hrs	
L3_SM_P	Soil Moisture (Radiometer)	36 km	50 hrs	
L3_SM_P_E	Soil Moisture (Radiometer, Enhanced)	9 km	50 hrs	
L3_SM_AP	Soil Moisture (Radar + Radiometer)	9 km	50 hrs	
L4_SM	Soil Moisture (Surface and Root Zone)	9 km	7 days	Science Value-Added
L4_C	Carbon Net Ecosystem Exchange (NEE)	9 km	14 days	

1.4 SMAP Science Objectives and Requirements

As mentioned, the SMAP science objectives are to provide new global data sets that will enable science and applications users to:

- Understand processes that link the terrestrial water, energy and carbon cycles;
- Estimate global water and energy fluxes at the land surface;
- Quantify net carbon flux in boreal landscapes;
- Enhance weather and climate forecast skill;

- Develop improved flood prediction and drought monitoring capability.

To resolve hydrometeorological water and energy flux processes and extend weather and flood forecast skill, a spatial resolution of 10 km and temporal resolution of 3 days are required. To resolve hydroclimatological water and energy flux processes and extend climate and drought forecast skill, a spatial resolution of 40 km and temporal resolution of 3 days are required. To quantify net carbon flux in boreal landscapes, a spatial resolution of 3 km and temporal resolution of 2 days are required. The SMAP mission will also validate a space-based measurement approach that could be used for future systematic hydrosphere state monitoring missions. The SMAP L2/3_SM_P products will meet the needs of the hydroclimatology community.

The SMAP mission Level 1 and Level 2 requirements state that:

"The baseline science mission shall provide estimates of soil moisture in the top 5 cm of soil with an error of no greater than 0.04 cm^3/cm^3 *(one sigma) at 10 km spatial resolution and 3-day average intervals over the global land area excluding regions of snow and ice, frozen ground, mountainous topography, open water, urban areas, and vegetation with water content greater than 5* kg/m^2 *(averaged over the spatial resolution scale)."*

L2-SR-347: "SMAP shall provide a Level 2 data product (L2_SM_P) at 40 km spatial resolution representing the average soil moisture in the top 5 cm of soil."

Although generated at a coarser 40-km spatial resolution, the L2/3_SM_P radiometer-based data products should still meet the 0.04 cm^3/cm^3 volumetric soil moisture retrieval accuracy specified in the mission Level 1 requirements. The SMAP Science Definition Team specified that data will be binned over 6-month time domain periods (April-September, October-March) globally within the SMAP mask when assessing radiometer performance and mission success in terms of soil moisture retrieval accuracies.

1.5 Document Outline

This document contains the following sections: Section 2 describes the basic physics of passive microwave remote sensing of soil moisture; Section 3 provides a description of the SMAP L2_SM_P and L3_SM_P data products; Section 4 introduces the baseline algorithm, along with other algorithm options; Section 5 addresses the use of the SMAP Algorithm Testbed in assessing algorithm performance and estimating error budgets for each candidate algorithm; Section 6 discusses the use of ancillary data and various flags; Section 7 presents procedures for downselecting to a baseline algorithm and for validating the data products; Section 8 provides a list of references; and Appendix 1 contains additional information about correcting observed brightness temperature for the presence of water bodies. This ATBD will be updated as additional work is completed in the pre- and post-launch periods.

2. PASSIVE REMOTE SENSING OF SOIL MOISTURE

The microwave portion of the electromagnetic spectrum (wavelengths from a few centimeters to a meter) has long held the most promise for estimating surface soil moisture remotely. Passive microwave sensors measure the natural thermal emission emanating from the soil surface. The variation in the intensity of this radiation depends on the dielectric properties and temperature of the target medium, which for the near surface soil layer is a function of the amount of moisture present. Low microwave frequencies (at L band or ~ 1 GHz) offer additional advantages: (1) the atmosphere is almost completely transparent, providing all-weather sensing; (2) transmission of signals from the underlying soil is possible through sparse and moderate vegetation layers (up to at least 5 kg/m^2 of vegetation water content); and (3) measurement is independent of solar illumination which allows for day and night observations.

The microwave soil moisture community has several decades of experience in conducting experiments using ground-based and aircraft microwave sensors [3-6]. These early experiments examined the basic physical relationships between emissivity and soil moisture, determined the optimum frequencies and measurement configurations, and demonstrated the potential accuracies for soil moisture retrievals. From these experiments a number of viable soil moisture retrieval algorithms have evolved, the most promising of which were explored in the Hydros OSSE (Observing System Simulation Experiment) [7] during the risk reduction phase of the project. Hydros was a proposed Earth System Science Pathfinder-class microwave soil moisture mission selected by NASA as a backup mission at the time of the OCO and Aquarius selections in 2002. Funding for Hydros ceased in 2005, but many of its risk reduction activities generated knowledge of direct relevance to SMAP. Additionally, much work was conducted by European and other colleagues prior to the launch of SMOS in 2009 [8-11].

2.1 Physics of the Problem

As mentioned, a microwave radiometer measures the natural thermal emission coming from the surface. At microwave frequencies, the intensity of the observed emission is proportional to the product of the temperature and emissivity of the surface (Rayleigh-Jeans approximation). This product is commonly called the brightness temperature T_B. If the microwave sensor is in orbit above the earth, the observed T_B is a combination of the emitted energy from the soil as attenuated by any overlying vegetation, the emission from the vegetation, the downwelling atmospheric emission and cosmic background emission as reflected by the surface and attenuated by the vegetation, and the upwelling atmospheric emission (Figure 2).

At L band frequencies, the atmosphere is essentially transparent, with the atmospheric transmissivity $\tau_{atm} \approx 1$. The cosmic background Tsky is on the order of 2.7 K. The atmospheric emission is also very small. These small atmospheric contributions will be accounted for in the L1B_TB ATBD, since the primary inputs to the radiometer-derived soil moisture retrieval process described in this L2_SM_P ATBD are atmospherically-corrected surface brightness temperatures as described in Section 3.

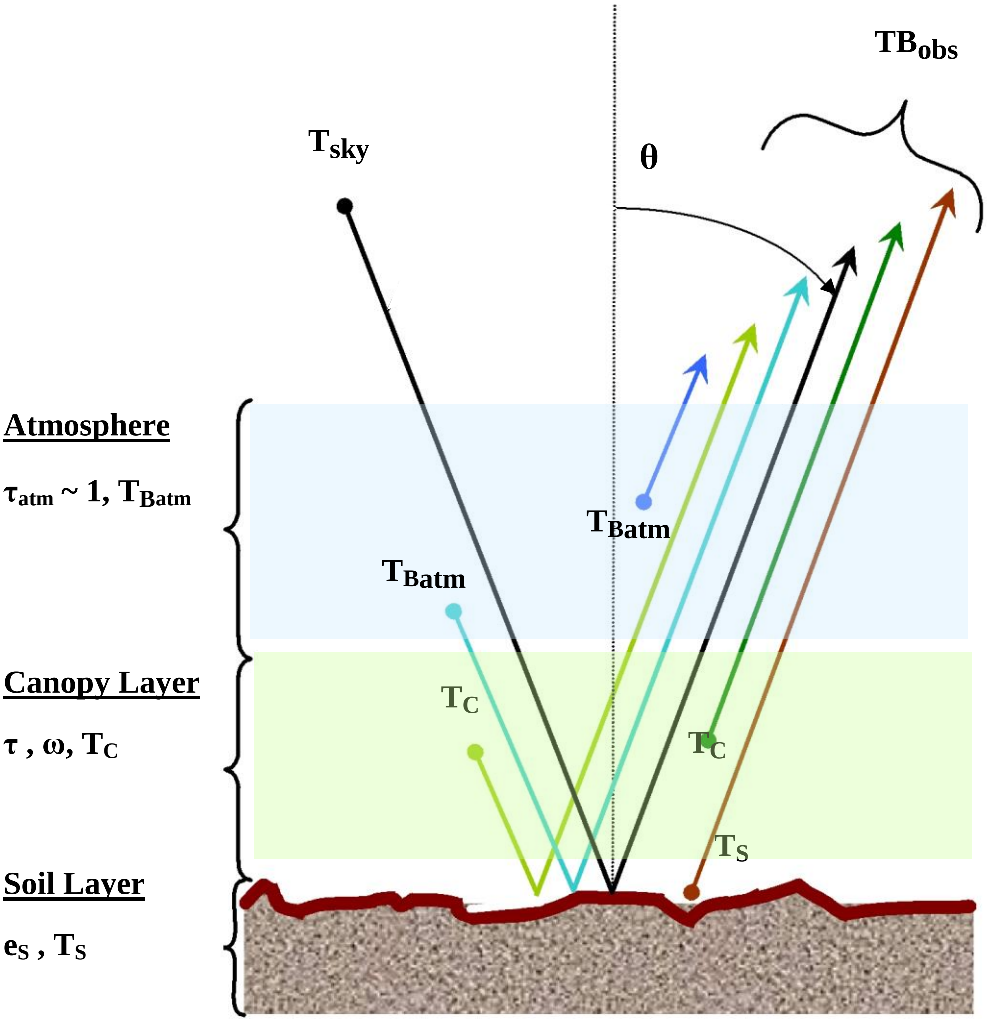


Figure 2. Contributions to the observed brightness temperature T_B from orbit [from SMOS ATBD, ref. 12] .

Retrieval of soil moisture from SMAP surface T_B observations is based on a well-known approximation to the radiative transfer equation, commonly known in the passive microwave soil moisture community as the *tau-omega* model. A layer of vegetation over a soil attenuates the emission of the soil and adds to the total radiative flux with its own emission. Assuming that scattering within the vegetation is negligible at L band frequencies, the vegetation may be treated mainly as an absorbing layer. A model following this approach to describe the brightness temperature of a weakly scattering layer above a semi-infinite medium was developed by [13] and described in [14]. The equation includes emission components from the soil and the overlying vegetation canopy [15]:

$$T_{Bp} = T_s e_p exp(-\tau_p \sec\theta) + T_c(1-\omega_p)[1-exp(-\tau_p \sec\theta)][1+r_p exp(-\tau_p \sec\theta)] \qquad (1)$$

where the subscript p refers to polarization (V or H), T_s is the soil effective temperature, T_c is the vegetation temperature, τ_p is the nadir vegetation opacity, ω_p is the vegetation single scattering albedo, and r_p is the rough soil reflectivity. The reflectivity is related to the emissivity (e_p) by $e_p = (1 - r_p)$, and ω_p, r_p and e_p are values at the SMAP look angle of $\theta = 40°$. The transmissivity γ of the overlying canopy layer is $\gamma = \exp(-\tau_p \sec \theta)$. Equation (1) assumes that vegetation multiple scattering and reflection at the vegetation-air interface are negligible. Surface roughness is modeled as $r_{p\ rough} = r_{p\ smooth} \exp(-h\cos^x \theta)$ where the parameter h is assumed linearly related to the root-mean-square surface height [16-17] and x =0, 1, or 2. Nadir vegetation opacity is related to the total columnar vegetation water content (VWC, in kg/m^2) by $\tau_p = b_p{*}VWC$ with the coefficient b_p dependent on vegetation type and microwave frequency (and polarization) [15].

If the air, vegetation, and near-surface soil are in thermal equilibrium, as is approximately the case near 6:00 am local time (the time of the SMAP descending pass), then T_c is approximately equal to T_s and the two temperatures can be replaced by a single effective temperature (T_{eff}). Soil moisture can then be estimated from $r_{p\ smooth}$ using the Fresnel and dielectric-soil moisture relationships.

The surface reflectance r_p is defined by the Fresnel equations, which describe the behavior of an electromagnetic wave at a smooth dielectric boundary. At horizontal polarization the electric field of the wave is oriented parallel to the reflecting surface and perpendicular to the direction of propagation. At vertical polarization the electric field of the wave has a component perpendicular to the surface. In the Fresnel equations below, θ is the SMAP incidence angle of 40^o and ε is the complex dielectric constant of the soil layer:

$$r_H(\theta) = \left| \frac{\cos\theta - \sqrt{\varepsilon - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon - \sin^2\theta}} \right|^2 \tag{2}$$

$$r_V(\theta) = \left| \frac{\varepsilon\cos\theta - \sqrt{\varepsilon - \sin^2\theta}}{\varepsilon\cos\theta + \sqrt{\varepsilon - \sin^2\theta}} \right|^2 \tag{3}$$

In terms of dielectric properties, there is a large contrast between liquid water ($\varepsilon_r \sim 80$) and dry soil ($\varepsilon_r \sim 5$). As soil moisture increases, soil dielectric constant increases. This leads to an increase in soil reflectivity or a decrease in soil emissivity ($1 - r_p$). Note that low dielectric constant is not uniquely associated with dry soil. Frozen soil, independent of water content, has a similar dielectric constant to dry soil. Thus, a freeze/thaw flag is needed to resolve this ambiguity. As T_B is proportional to emissivity for a given surface soil temperature, T_B decreases in response to an increase in soil moisture. It is this relationship between soil moisture and soil dielectric constant (and hence microwave emissivity and brightness temperature) that forms the physical basis of passive remote sensing of soil moisture. Given SMAP observations of T_B and information on T_{eff}, h, τ_p, and ω_p from ancillary sources (Section 6) or multichannel algorithm approaches (Section 4), Equation (1) can be solved for the soil reflectivity r_p, and equation (2) or (3) can be solved for the soil dielectric ε. Soil moisture can then be estimated using one of several dielectric models and ancillary knowledge of soil texture.

2.2 Rationale for L-Band

Within the microwave portion of the electromagnetic spectrum, emission from soil at L-band frequencies can penetrate through greater amounts of vegetation than at higher frequencies. Figure 3 shows microwave transmissivity as a function of increasing biomass at L-band (1.4 GHz), C-band (6 GHz), and X-band (10 GHz) frequencies, based upon modeling. The results clearly show that L-band frequencies have a significant advantage over the C- and X-band frequencies (and higher) provided by current satellite instruments such as AMSR-E and WindSat, and help explain why both SMOS and SMAP are utilizing L band sensors in estimating soil moisture globally over the widest possible vegetation conditions. Another advantage of measuring soil moisture at L-band is that the microwave emission originates from deeper in the soil (typically 5 cm or so), whereas C- and X-band emissions originate mainly from the top 1 cm or less of the soil (Figure 4).

Although the above arguments support the use of low frequencies, there is, however, a lower frequency limit for optimal T_B measurements for soil moisture. At frequencies lower than L-band, radiometric measurements are significantly degraded by manmade and galactic noise. Since there is a protected band at L band at 1.400−1.427 GHz that is allocated exclusively for radiometric use, the SMAP radiometer operates in this band.

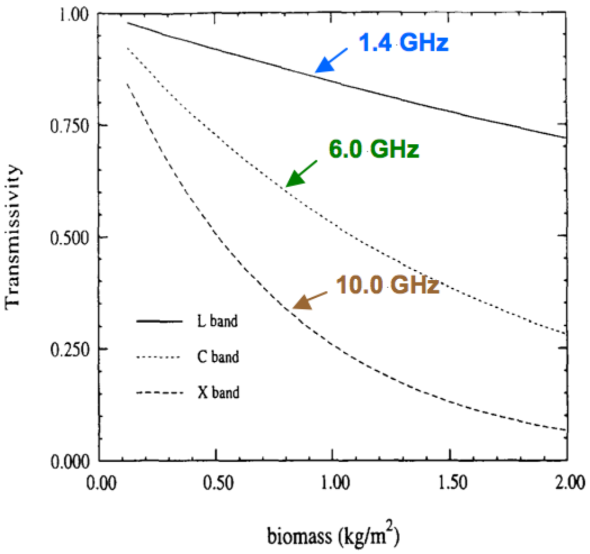


Figure 3. Vegetation transmissivity to soil emission at L-band frequencies (1.4 GHz) is much higher than at C- (6 GHz) or X-band (10 GHz) frequencies [adapted from 22].

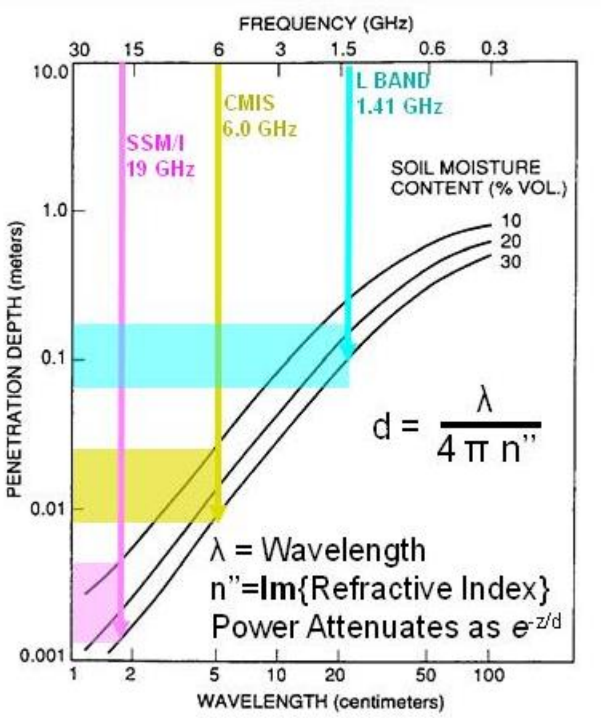


Figure 4. L-band T_B observations are sensitive to emission from deeper in the soil than at higher frequencies [adapted from 23]. Soil moisture curves are given for 10, 20, and 30% (or in absolute units, 100 x cm^3/cm^3).

2.3 Soil Dielectric Models

In the past few decades, a number of soil dielectric models have been developed by the passive microwave remote sensing community. Although they differ in analytical forms, they generally share common dependence on soil moisture, soil texture, and frequency. The details of these models have been described thoroughly in the literature – a good summary can be found in [18, 19]. Currently, the SMAP project is investigating the use of three different soil dielectric models:

(1) Dobson [20] – a semi-empirical mixing model, the Dobson model retains the physical aspects of the dielectric properties of free water of the soil through the Debye equations while also using certain empirical fitting parameters based on the different soil types studied during the model's development; the model requires frequency, soil moisture, soil temperature, sand fraction, clay fraction, and bulk density as input parameters.

(2) Wang & Schmugge [21] – a central point of this empirical mixing model is the use of a transition point of water content beyond which the dielectric constant increases rapidly with soil moisture; the model predicts and illustrates the substantial impact of bound water (as opposed to free water only) on soil dielectric constant.

(3) Mironov [19] – formally known as the Mineralogy-Based Soil Dielectric Model (MBSDM); using a large soil database, Mironov was able to obtain a set of regression equations to derive many of the spectroscopic parameters needed by a model that he

developed earlier; the resulting model not only applies to a wider range of soil types, but also requires fewer input parameters – with clay percentage as the only soil input parameter.

These three models have been widely used due to their simple parameterizations and applicability at L-band frequencies (1.26-1.41 GHz). As part of SMAP pre-launch and post-launch calibration/validation activities, the performance of these dielectric models in terms of bias and accuracy of the retrieved soil moisture will be evaluated and a decision made on which dielectric model to carry forward into the operational production of SMAP data products. The SMAP L2_SM_P processing software has a switch which selects which dielectric model will be used in the soil moisture retrieval. For comparison, ESA's SMOS mission currently uses land cover classification to choose the appropriate dielectric model (Dobson or Mironov). Figure 5 gives an example of the performance of the three dielectric models when used in forward model computations of L band T_B for θ = 40°, assuming smooth bare soils at T_S = 25°C for different soil types.

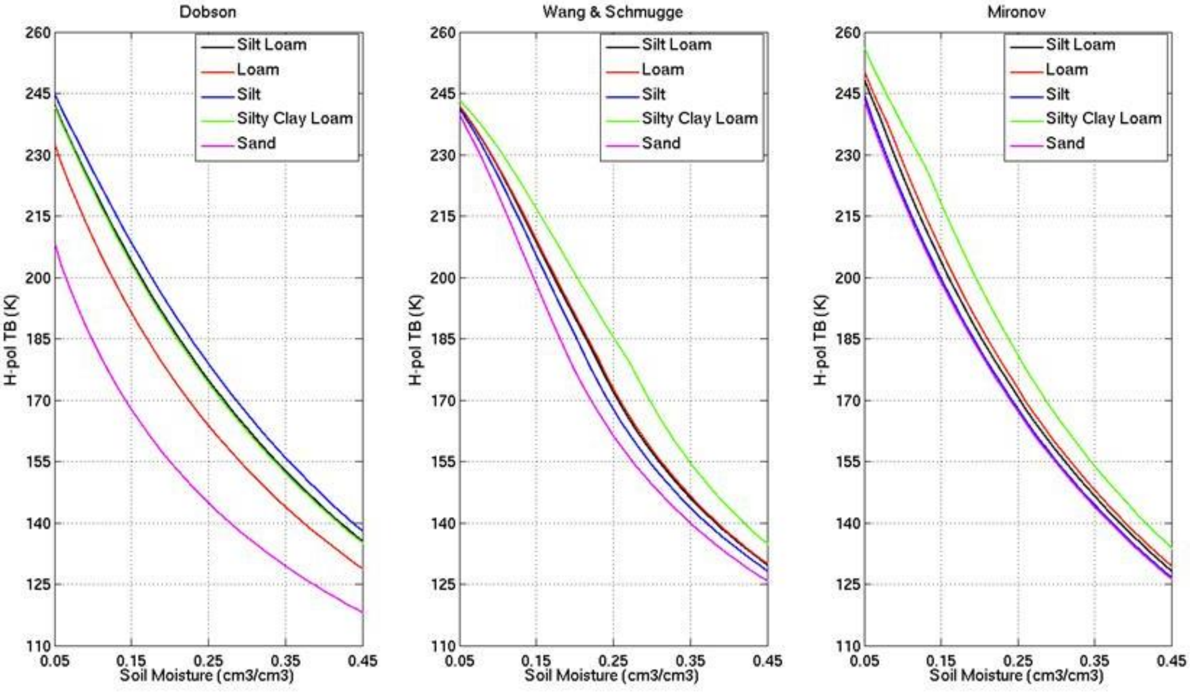


Figure 5. **Bare soil T_B as computed by different soil dielectric models. The selected soil types correspond to the top five most dominant soil texture classes, together accounting for over 80% of the global land area.**

2.4 Use of the 6:00 AM Descending Node Orbit for the Primary Mission Product

The decision to place SMAP into a sun-synchronous 6:00 am / 6:00 pm orbit is based on a number of science issues relevant to the L2_SM_P product [24, 25]. Faraday rotation is a phenomenon in which the polarization vector of an electromagnetic wave rotates as the wave propagates through the ionospheric plasma in the presence of the Earth's static magnetic field. The phenomenon is a concern to SMAP because the polarization rotation increases as the square of wavelength. If uncorrected, the SMAP polarized (H and V) radiometer measurements will contain errors that translate to soil moisture error. Faraday rotation varies greatly during the day, reaching a maximum during the afternoon and a minimum in the pre-dawn hours. By using T_B observations acquired near 6:00 am local solar time as the primary input to the L2_SM_P product, the adverse impacts of Faraday rotation are minimized. Faraday rotation correction to SMAP T_B is described in the L1B_TB ATBD.

At 6:00 am the vertical profiles of soil temperature and soil dielectric properties are likely to be more uniform [13] than at other times of the day (Figure 7). This early morning condition will minimize the difference between canopy and soil temperatures and thermal differences between land cover types within a pixel (Figure 6). These factors help to minimize soil moisture retrieval errors originating from the use of a single effective temperature to represent the near surface soil and canopy temperatures. This same effective temperature can be used as the open water temperature in the water body correction to T_B that will be discussed in Sections 3 and 4.

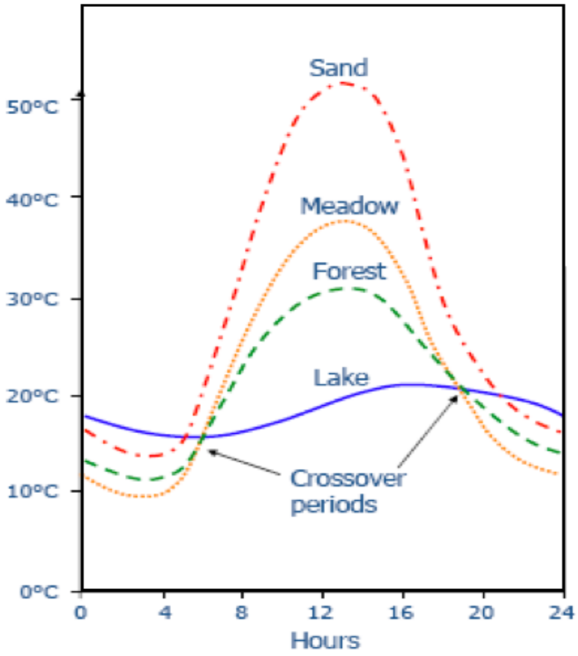


Figure 6. **Schematic showing diurnal variation in temperature and thermal crossover times at approximately 6:00 am / 6:00 pm local time for various broad classes of land surface covers [modified from 24].**

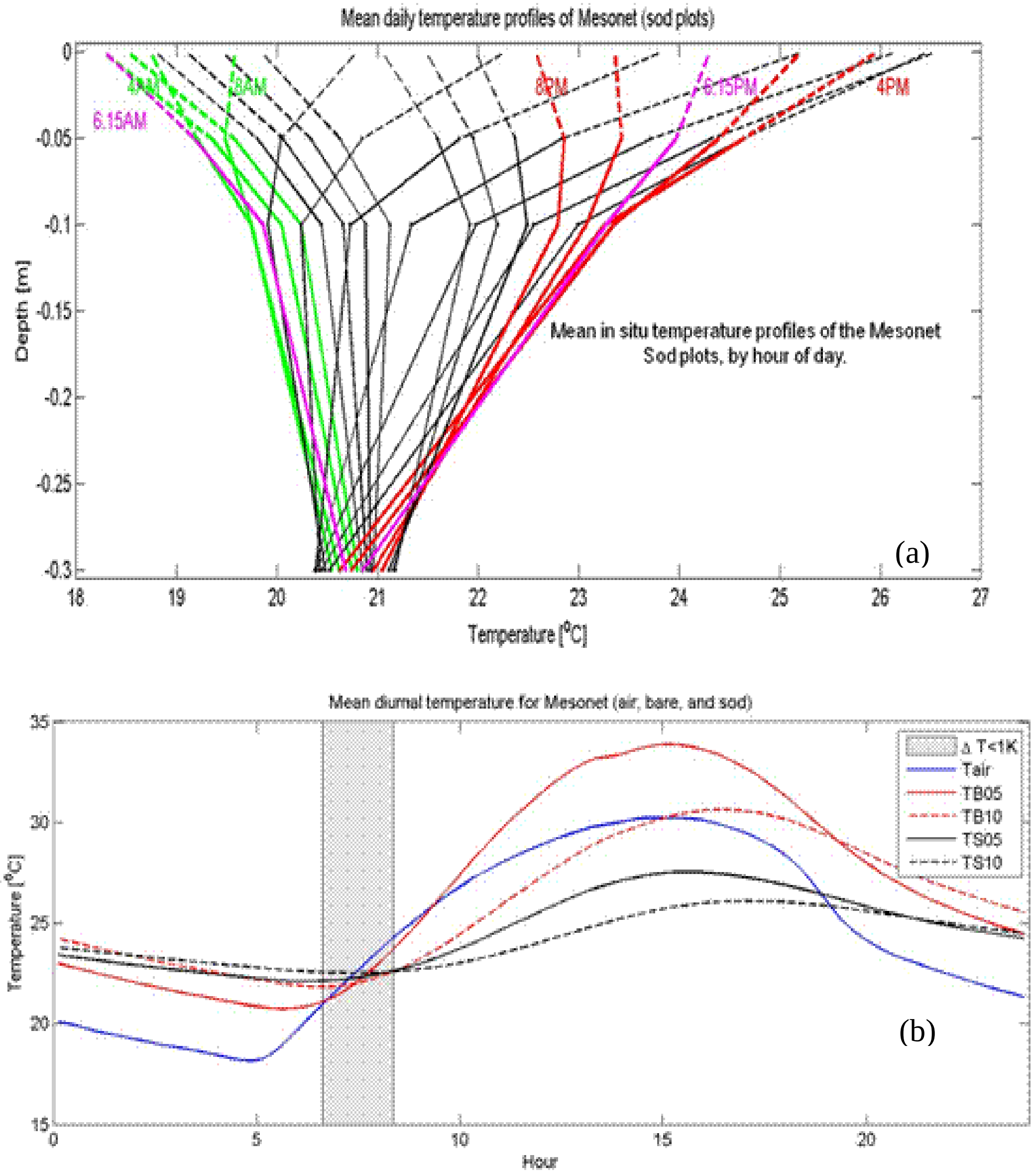


Figure 7. Soil temperature as a function of time based on June 2004 Oklahoma Mesonet data: (a) vertical profiles for a sod covered site and (b) the mean soil temperatures for bare soil (TB05, TB10) and sod (TS05, TS10). The shaded region identifies the period of the day when these effects result in less than 1° C difference among the four temperatures (T. Holmes, personal communication).

Finally, it is desirable to establish a long-term climate data record of L-band brightness temperatures and soil moisture. Such a data record could enable investigations of important trends in emissivity, soil moisture, and other derived variables occurring over annual to decadal periods. Both the SMOS and Aquarius L-band missions will operate in 6 am/6 pm orbits, and SMAP will extend these L-band data records.

As will be discussed in Section 3, the current approach to generation of the baseline L2_SM_P product was originally restricted to input data from the 6:00 am descending passes because of the thermal equilibrium assumption and near-uniform thermal conditions of surface soil layers and overlying vegetation in the early morning hours. Accurate soil moisture retrievals using data from 6:00 pm ascending passes may require use of a land surface model and will be generated as part of the L4_SM product (see ATBD for L4_SM). However, some early results from the SMOS mission suggest that the additional error associated with 6 pm retrievals may not be as large as expected [48]. As described in Section 8.1, the SMAP project in L2_SM_P Data Release Version 4 (Dec., 2016) will produce a 6 pm retrieved soil moisture product using the same retrieval algorithm as the 6 am soil moisture product.

3. PRODUCT OVERVIEW

This ATBD covers the two coarse spatial resolution soil moisture products which are based on the SMAP radiometer brightness temperatures: L2_SM_P, which is derived surface soil moisture in half-orbit format at 40 km resolution output on a fixed 36-km Equal-Area Scalable Earth-2 (EASE2) grid, and L3_SM_P, which is a daily global composite of the L2_SM_P surface soil moisture, also at 40 km resolution output on a fixed 36-km EASE2 grid. Utilizing one or more of the soil moisture retrieval algorithms to be discussed in section 4, SMAP brightness temperatures are converted into an estimate of the 0-5 cm surface soil moisture in units of cm^3/cm^3.

3.1 Inputs to Soil Moisture Retrieval

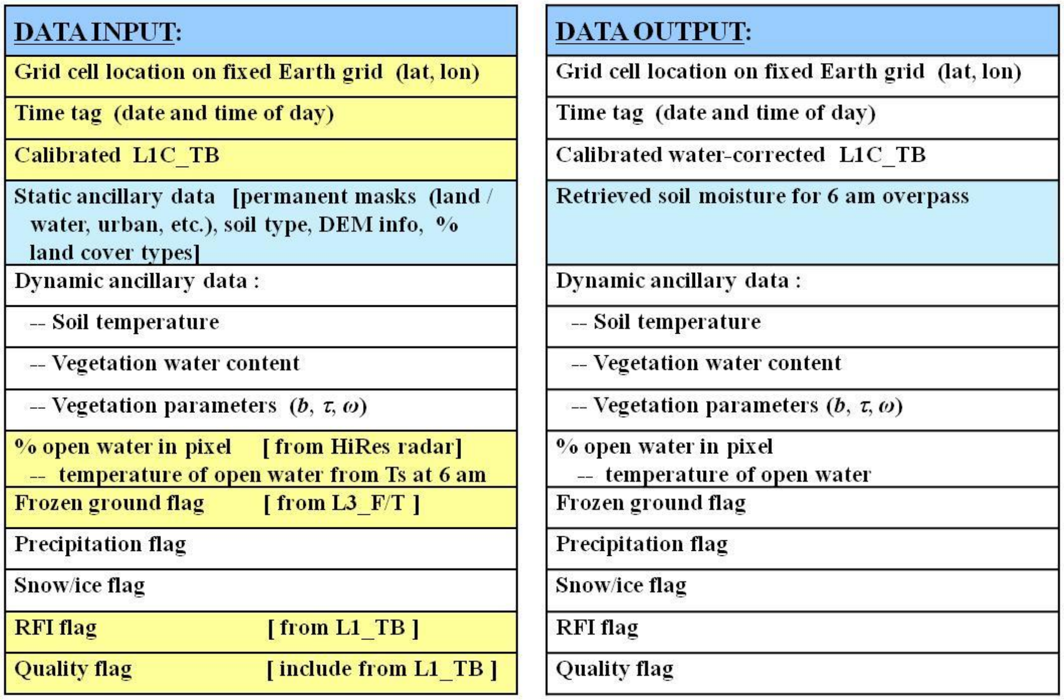
The main input to the L2_SM_P processing algorithm is the SMAP L1C_TB product that contains the time-ordered, geolocated, calibrated L1B_TB brightness temperatures which have been resampled to the fixed 36-km EASE2 grid. In addition to general geolocation and calibration, the L1B_TB data have also been corrected for atmospheric effects, Faraday rotation, and low-level RFI effects prior to regridding. If the RFI encountered is too large to be corrected, the T_B data are flagged accordingly and no soil moisture retrieval is attempted. See the L1B_TB and L1C_TB ATBDs for additional details.

In addition to T_B observations, the L2_SM_P algorithm also requires ancillary datasets for the soil moisture retrieval. These include:

- Surface temperature
- Vegetation opacity (or vegetation water content and vegetation opacity coefficient)
- Vegetation single scattering albedo
- Surface roughness information
- Land cover type classification
- Soil texture (sand, silt, and clay fraction)
- Data flags for identification of land, water, precipitation, RFI, urban areas, mountainous terrain, permanent ice/snow, and dense vegetation

The specific parameters and sources of these and other externally provided ancillary data are listed in Section 6. Other parameters used by the L2_SM_P algorithm are provided internally to the processing chain. These include a freeze/thaw flag, an open water fraction, and a vegetation index, provided by the SMAP Hi-Res radar L2_SM_A product (see L2_SM_A ATBD) or other ancillary sources.

All input T_B and ancillary datasets used in the retrievals are mapped to the 36-km EASE2 grid prior to entering the L2_SM_P processor. All input data, retrieved soil moisture data, and flags utilize the same grid.



DATA INPUT:	
Grid cell location on fixed Earth grid (lat, lon)	
Time tag (date and time of day)	
Calibrated L1C_TB	
Static ancillary data [permanent masks (land / water, urban, etc.), soil type, DEM info, % land cover types]	
Dynamic ancillary data :	
-- Soil temperature	
-- Vegetation water content	
-- Vegetation parameters (b, τ, ω)	
% open water in pixel -- temperature of open water from Ts at 6 am	[from HiRes radar]
Frozen ground flag	[from L3_F/T]
Precipitation flag	
Snow/ice flag	
RFI flag	[from L1_TB]
Quality flag	[include from L1_TB]

DATA OUTPUT:
Grid cell location on fixed Earth grid (lat, lon)
Time tag (date and time of day)
Calibrated water-corrected L1C_TB
Retrieved soil moisture for 6 am overpass
Dynamic ancillary data :
-- Soil temperature
-- Vegetation water content
-- Vegetation parameters (b, τ, ω)
% open water in pixel -- temperature of open water
Frozen ground flag
Precipitation flag
Snow/ice flag
RFI flag
Quality flag

Figure 8. Conceptual list of input and output information for the L2_SM_P soil moisture product.

3.2 Algorithm Outputs

Figure 8 lists in a conceptual way the variety of input and output data associated with the SMAP L2_SM_P soil moisture product. Many of these parameters will be discussed in Section 4 and Section 6. The primary contents of the output L2_SM_P and L3_SM_P products are the retrieved soil moisture and associated quality control (QC) flags, as well as the values of the ancillary parameters needed to retrieve the output soil moisture for that grid cell. The exact Data Product Description for the L2_SM_P and L3_SM_P products was generated in consultation with SMAP Science Data System (SDS) personnel, and will be available to the public through the DAACs (see also Appendix 1).

3.3 Product Granularity

The L2_SM_P product is a half-orbit product. SMAP ascending (6 pm) half-orbits are defined starting at the South Pole and ending at the North Pole, while descending (6 am) half-orbits start at the North Pole and end at the South Pole. Input T_B observations from a given half-orbit are processed to generate output soil moisture retrievals for the same half orbit.

The L3_SM_P product is a daily product generated by compositing one day's worth of L2_SM_P half-orbit granules, separately for ascending and descending half-orbits, onto a global array. T_B observations from descending (6 am) passes will be used to retrieve soil moisture for the L2_SM_P and L3_SM_P standard products as mentioned in Section 2.4. In the L2_SM_P Data Release Version 4 in December, 2016, 6 pm soil moistures will also be produced by applying the baseline 6 am retrieval algorithm to T_B data from the 6 pm ascending passes. The 6 pm soil moisture data will be done on a best effort basis and will not be included in assessments of whether the L2_SM_P product meets the mission Level 1 requirements. However, the 6 pm retrievals will also be compared against observations of soil moisture to assess their accuracy. Currently, the data volume estimate for the L2_SM_P product is 15 MB/day and the data volume estimate for the L3_SM_P product is 41 MB/day; these values are based on products from the 6 am descending pass only.

3.4 SMAP Product Suite

The L2_SM_P and L3_SM_P products are part of the suite of SMAP products shown previously in Table 2. The SMAP L1-L3 products will be generated by the SMAP Science Data Processing System (SDS) at JPL, while the SMAP L4 products will be produced by the Global Modeling and Assimilation Office (GMAO) at NASA GSFC. All SMAP data products approved for release will be archived and made available to the public through a NASA-designated Earth Science Data Center. NASA HQ has designated that the National Snow and Ice Data Center (NSIDC) in Boulder, CO will be the primary SMAP DAAC, although SMAP HiRes radar data will be archived separately at the Alaska Satellite Facility (ASF) in Fairbanks, AK.

3.5 EASE Grid

The grid selected for the SMAP geophysical (L2-L4) products is the updated Equal-Area Scalable Earth-2 (EASE2) grid [26]. This grid was originally conceived at the NSIDC and has been used to archive several satellite instrument data sets including SMMR, SSM/I, and AMSR-E [27]. Using this same grid system for SMAP provides user convenience, facilitates continuity of historical data grid formats, and enables re-use of heritage gridding and extraction software tools developed for EASE grid.

The EASE grid has a flexible formulation. By adjusting one scaling parameter it is possible to generate a family of multi-resolution EASE grids that “nest” within one another. The nesting can be made “perfect” in that smaller grid cells can be tessellated to form larger grid cells, as shown in Figure. 9a. This feature provides SMAP data products with a convenient common projection for both high-resolution radar observations and low-resolution radiometer observations. Figure 9b illustrates the different resolutions for the 3-, 9-, and 36-km EASE grids.

A nominal EASE grid dimension of 36 km has been selected for the L2/3_SM_P products. This is close to the 40-km resolution of the radiometer footprint and scales conveniently with the 3 km and 9 km grid dimensions that have been selected for the radar-only (L2/3_SM_A) and combined radar/radiometer (L2/3_SM_A/P) soil moisture products, respectively. A global 36-km EASE grid can be constructed having an integer number of rows and columns (408 and 963), with northernmost/southernmost latitudes of ±86.6225°, using a scaling parameter[1] that is almost exactly 36 km.

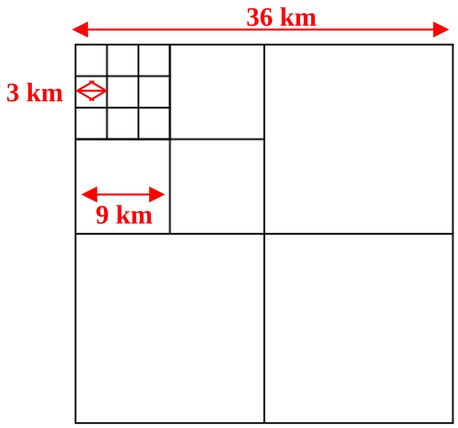


Figure 9a. Perfect nesting in EASE grid – smaller grid cells can be tessellated to form larger grid cells.

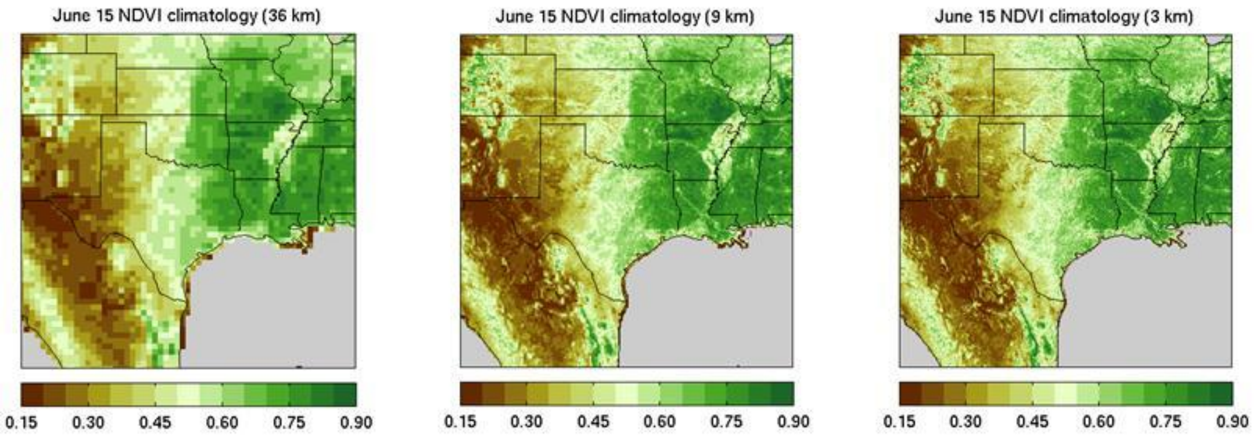


Figure 9b. Example of ancillary NDVI climatology data displayed on the SMAP 36-km, 9-km, and 3-km EASE grids.

---

[1] The precise value of the scaling parameter is 36.00040003 km at ±30° latitudes.

3.6 Soil Moisture Retrieval Process

Figure 10 illustrates the conceptual process used in retrieving soil moisture from SMAP radiometer brightness temperature measurements. In order for soil moisture to be retrieved accurately, a variety of global static and dynamic ancillary data are required (Section 6). Static ancillary data are data which do not change during the mission, while dynamic ancillary data require periodic updates in time frames ranging from seasonally to daily. Static data include parameters such as permanent masks (land/water/forest/urban/ mountain), the grid cell average elevation and slope derived from a DEM, permanent open water fraction, and soils information (primarily sand and clay fraction). The dynamic ancillary data include land cover, surface roughness, precipitation, vegetation parameters, and effective soil temperatures. Measurements from the SMAP radar will be one source of information on open water fraction and frozen ground, in addition to water information from a MODIS-derived surface water data base and temperature information from the GMAO model used in L4_SM. Ancillary data will also be employed to set flags which help to determine either specific aspects of the processing

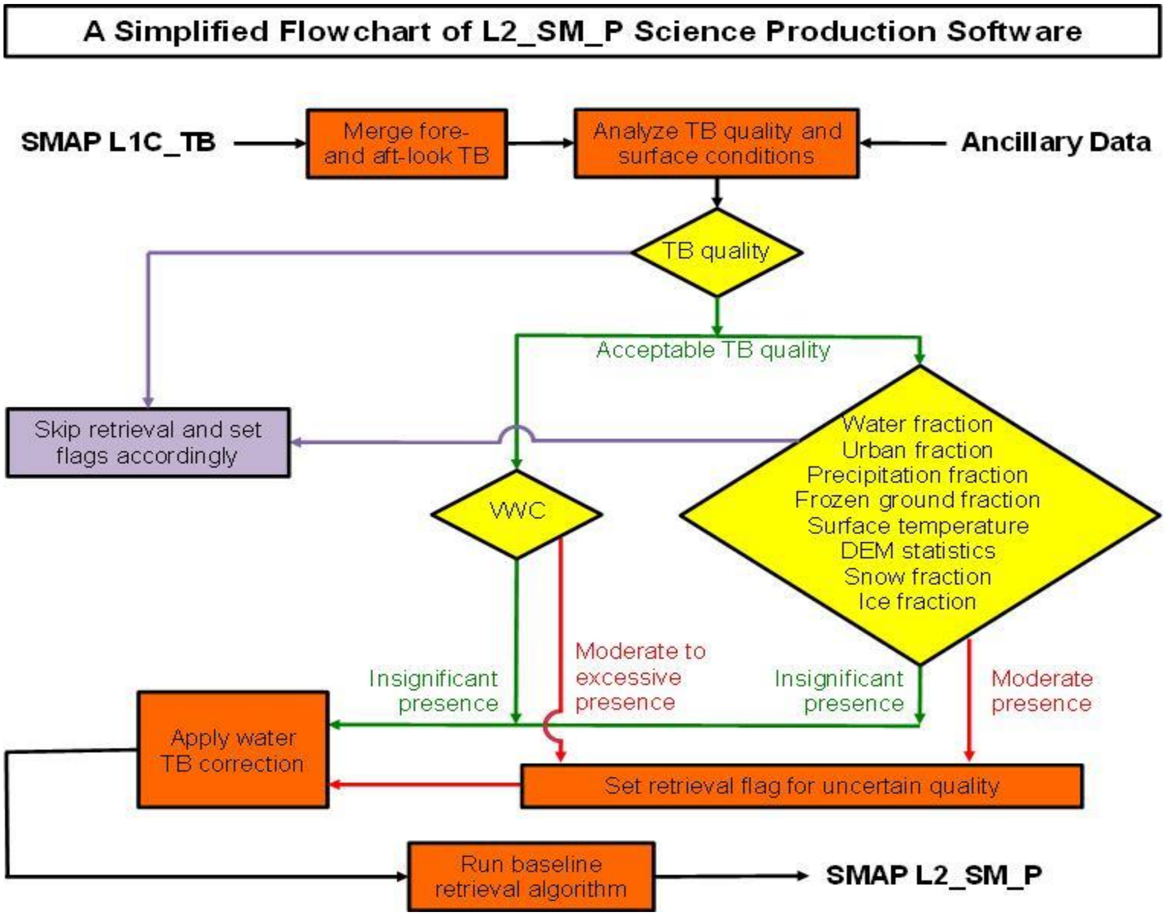


Figure 10. Conceptual flow of L2_SM_P process from input of T_B to output of retrieved soil moisture.

(such as corrections for open water to be discussed in Section 4) or the quality of the retrievals (e.g. precipitation flag). Basically, these flags would provide information as to whether the ground is frozen, snow-covered, or flooded, or whether it is expected to be actively precipitating at the time of the satellite overpass. Other flags will indicate whether masks for steeply sloped topography, or for urban, heavily forested, or permanent snow/ice areas are in effect. All input data to the L2_SM_P processor are pre-mapped to the 36-km EASE2 grid.

Consistent with the SMAP Level 2 mission requirements [28], the L2_SM_P product is a half-orbit product — T_B observations from a given half orbit go through the retrieval algorithm to produce retrieved soil moisture for the same half orbit. An example of the L2_SM_P soil moisture from part of a single half orbit over the United States as simulated on the SMAP Algorithm Testbed (Section 5) is shown in Figure 11. This example is based on a single-channel algorithm operating on H-polarized T_B observations simulated using geophysical data from a land surface model.

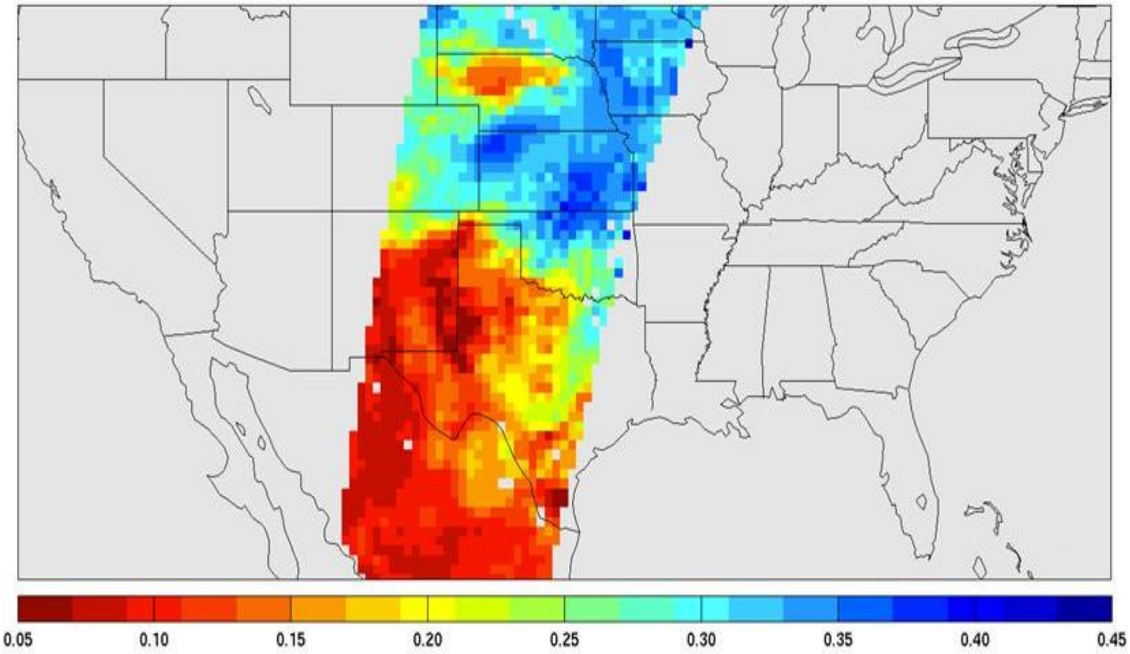


Figure 11. Example of SMAP retrieved soil moisture in cm^3/cm^3. The half-orbit swath pattern is simulated using the orbital sampling module on the SMAP Algorithm Development Testbed.

3.7 Level 3 Radiometer-Based Soil Moisture Product (L3_SM_P)

The L3_SM_P product is a daily global product. To generate the product, individual L2_SM_P half-orbit granules acquired over one day are composited to produce a daily multi-orbit global map of retrieved soil moisture.

The L2_SM_P swaths overlap poleward of approximately ±65° latitude. Where overlap occurs, three options are considered for compositing multiple data points at a given grid cell:

1. Use the most recent (or “last-in”) data point
2. Take the average of all data points within the grid cell
3. Choose the data point observed closest to 6:00 am local solar time

The current approach for the L3_SM_P product is to use the nearest 6:00 am local solar time (LST) criterion to perform Level 3 compositing (a similar procedure is used for 6 pm in Data Release Version 4). According to this criterion, for a given grid cell, an L2 data point acquired closest to 6:00 am local solar time will make its way to the final Level 3 granule; other 'late-coming' L2 data points falling into the same grid cell will be ignored. For a given granule whose time stamp (yyyy-mm-ddThh:mm:ss) is expressed in UTC, only the hh:mm:ss part is converted into local solar time. For example,

UTC Time Stamp	**Longitude**	**Local Solar Time**
2011-05-01T23:19:59	60E	23:19:59 + (60/15) hrs = 03:19:59

The local solar time 03:19:59 is then compared with 06:00:00 in Level 3 processing for 2011-05-01 to determine if the swath is acquired closest to 6:00 am local solar time. If so, that data point (and only that data point) will go to the final Level 3 granule. Under this convention, an L3 composite for 2011-05-01 has all Level 2 granules acquired within 24 hours of 2011-05-01 UTC and Level 2 granules appearing at 2011-05-02 6:00 am local solar time at the equator. Note that this is also the conventional way to produce Level 3 products in similar missions and is convenient to users interested in global applications. Figure 12 shows an example of the L3_SM_P soil moisture output for one day’s worth of simulated SMAP descending orbits globally (Fig. 12a) and over just the continental U.S. (CONUS) (Fig. 12b).

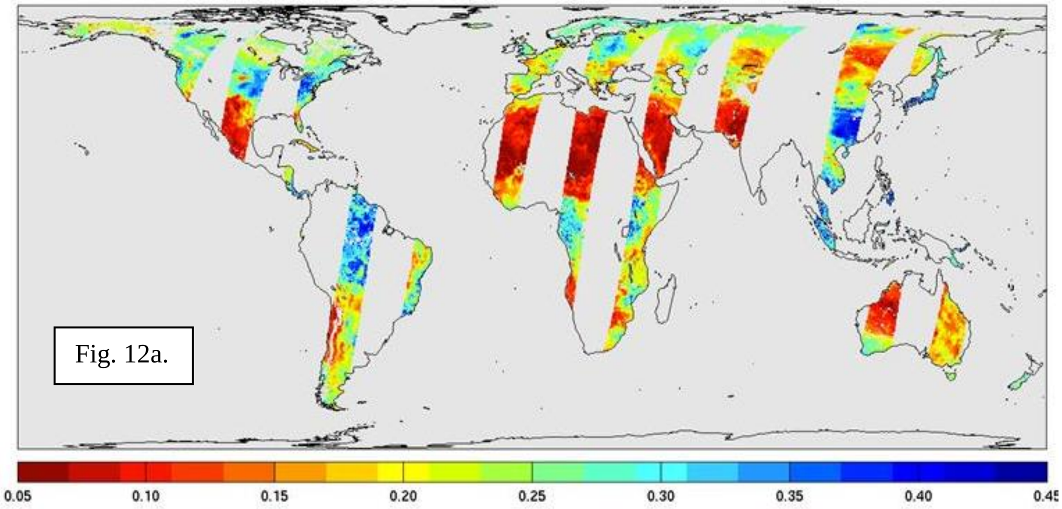


Fig. 12a.

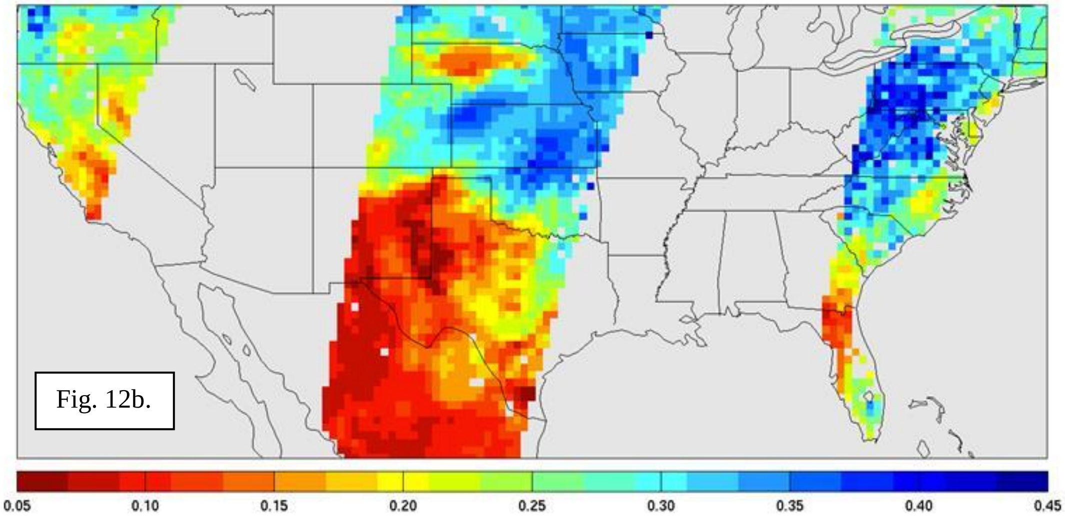


Fig. 12b.

Figure 12. Simulation of L3_SM_P retrieved soil moisture in cm^3/cm^3. This example is based on the single channel algorithm operating on H-polarized T_B observations simulated using geophysical data from a land surface model.

4. RETRIEVAL ALGORITHMS

Decades of research by the passive microwave soil moisture community has resulted in a number of viable soil moisture retrieval algorithms that can be used with SMAP T_B data. ESA's SMOS mission currently flies an aperture synthesis L-band radiometer which produces T_B data at multiple incidence angles over the same ground location. The baseline SMOS retrieval algorithm is based on the *tau-omega* model described in Section 2.1, but utilizes the SMOS multiple incidence angle capability to retrieve soil moisture. SMAP retrievals will also be based on the *tau-omega* model, but will use the constant incidence angle T_B data produced by the SMAP conically-scanning radiometer. Other needed parameters in the retrieval will be obtained as ancillary data.

SMAP baseline and optional algorithms will be evaluated for their soil moisture retrieval performance during the pre- and post-launch time frames. The optional algorithms will be compared against the baseline algorithm using theoretical simulations and observational data. Upon periodic assessment and review by the SMAP science team, a retrieval algorithm option with better performance than the baseline algorithm may replace the earlier baseline and become the new baseline.

For the SMAP L2_SM_P product, five soil moisture retrieval algorithms are currently being evaluated:

- Single Channel Algorithm at H polarization (baseline) (SCA-H)
- Single Channel Algorithm at V polarization (SCA-V)
- Dual-Channel Algorithm (DCA)
- Microwave Polarization Ratio Algorithm (MPRA)
- Extended Dual Channel Algorithm (E-DCA)

Evaluations are done using simulations, testing with observational data from the PALS airborne and ComRAD ground-based instruments (SMAP simulators), other field campaign data, and *in situ* cal/val (CV) site data, as well as by applying candidate SMAP algorithms to SMOS and Aquarius satellite T_B data. The four algorithms (one baseline and three options) are described in this section. 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.

4.1 Water TB Correction

At the 40-km footprint resolution scale of the SMAP radiometer, a significant percentage of footprints within the SMAP land mask will contain some amount of open fresh water due to the presence of lakes, rivers, wetlands, and transient flooding. It is assumed that all ocean pixels will be masked out using the SMAP ocean/land mask. For soil moisture retrieval purposes, the presence of open water within the radiometer footprint (IFOV) is undesirable since it dramatically lowers the brightness temperature and results in anomalously high retrieved soil moisture for that grid cell if soil moisture is retrieved without knowledge of the presence of open water. This results in a bias which degrades the overall soil moisture retrieval accuracy. It is therefore important to correct the SMAP Level 1 T_B observations for the presence of water, to the extent feasible, prior to using them as inputs to the L2_SM_P soil moisture retrieval. Fortunately, this bias can be corrected, especially when it occurs at dawn near inland water/land boundaries where the temperature of water can be reasonably approximated as the temperature of land (as shown in Figure 6).

The procedure to correct for water T_B is quite simple. Given a mixture of land and water within the antenna IFOV, the observed TB is an areal weighted sum of the T_B contributions from the water and from the land:

$$T_B^{IFOV} = \alpha T_B^{water} + (1-\alpha) T_B^{land} \tag{4}$$

where α denotes the areal fraction of water within the antenna IFOV, and $T_B{}^{water}$ denotes the T_B emission from water computed from a theoretical model (for ex., the Klein-Swift model [29]) with an estimated physical temperature obtained from ancillary sources. At the 6 am SMAP overpass time, the temperature of the water is approximately the same as the temperature of the surface soil layer, so the same ancillary temperature data can be used for both. Once α and $T_B{}^{water}$ are known, $T_B{}^{land}$ can be solved for and then used to retrieve soil moisture in the non-open water part of the IFOV.

In principle, this water T_B correction should be applied to brightness temperatures before gridding for greater retrieval accuracy. However, doing so requires intensive on-the-fly calculations to determine the spatial extent covered by the antenna's 3-dB beamwidth for each 12-ms T_B sample (and there are roughly half a million such samples in a given orbit), while providing only a small gain in retrieval accuracy (less than 0.005 cm^3/cm^3; see Appendix 2) compared with the same correction applied to T_B after gridding. For this reason, the L2_SM_P team decided to implement water T_B correction on the L1C_TB product:

$$T_B^{gridded} = \alpha T_B^{water} + (1-\alpha) T_B^{land} \tag{5}$$

L1B_TB half-orbit T_B data are mapped on a 36-km EASE grid to form $T_B{}^{gridded}$ (the SMAP L1C_TB product). The open water fraction (α) (consisting of both static and transient water) is then calculated for each 36-km grid cell using information from the permanent water data base (Section 6) and the 1-km scale water fraction parameter from the SMAP high-resolution radar (see the L2_SM_A ATBD). $T_B{}^{water}$ is calculated as described above, and once α and $T_B{}^{water}$ are known, $T_B{}^{land}$ can be solved for and then used in the L2-SM_P processor to retrieve soil moisture in the non-open water part of the grid cell.

It is important to recognize that there is a threshold for α above which the correction may generate enough errors that the SMAP's target retrieval accuracy of 0.04 cm^3/cm^3 may not be met. Figure 12 shows how water fraction and land/water classification error affect the retrieval accuracy. Given the uncertainties of T_B observations as well as other model, ancillary, and environmental parameters, a relatively tight margin (0.005 cm^3/cm^3) of retrieval accuracy is plotted as a function of water fraction and classification error for three vegetation water contents (VWC) levels: 0.0, 2.5, and 5.0 kg/m^2. The figure shows that, for a given water fraction, water TB mixed with bare-soil TB is more easily correctable than with densely vegetated TB. In the worst-case scenario (the green curve in Figure 13), a water fraction of 4% with a classification error no greater than 5% is needed to meet a retrieval accuracy of 0.005 cm^3/cm^3 at VWC = 5 kg/m^2.

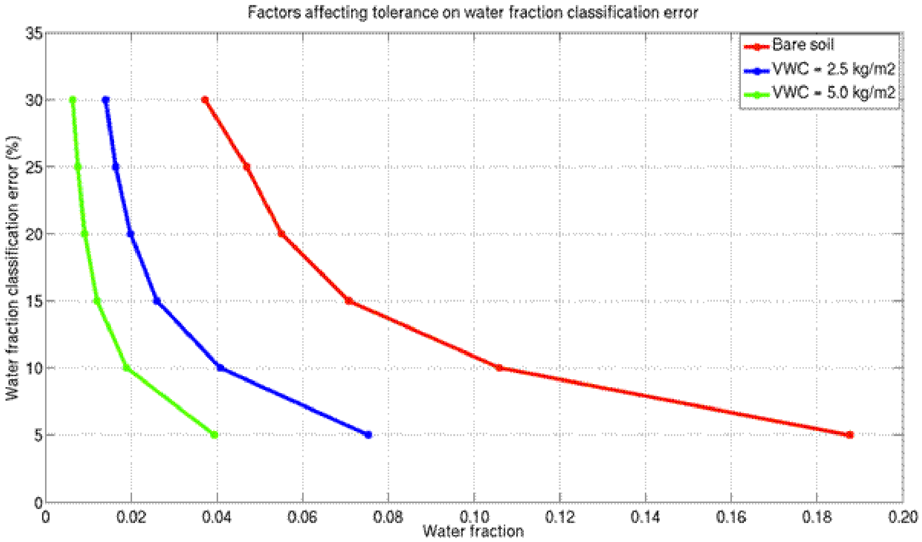


Figure 13. For a given soil moisture retrieval RMSE (0.005 cm^3/cm^3 in this case), more accurate estimation of the water fraction is needed for T_B observations that contain a larger water fraction and/or a larger vegetation water content.

4.2 Single Channel Algorithm (Current Baseline)

From a broad perspective, there are five steps involved in extracting soil moisture using passive microwave remote sensing. These steps are normalizing brightness temperature to emissivity, removing the effects of vegetation, accounting for the effects of soil surface roughness, relating the emissivity measurement to soil dielectric properties, and finally relating the dielectric properties to soil moisture.

In the single channel algorithm (SCA) [4], horizontally polarized T_B are traditionally used due to their sensitivity to soil moisture, but the same algorithm can also be applied to V polarization T_B. The use of H pol T_B with the SCA was the prelaunch SMAP baseline algorithm; SCA-V is the baseline for the postlaunch release of L2_SM_P data. In the SCA approach, brightness temperatures 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. Additional details on these steps follow.

At the L band frequency used by SMAP, the brightness temperature of the land surface is proportional to its emissivity (e) multiplied by its physical temperature (T). It is typically assumed that the temperatures of the soil and the vegetation are the same, especially at the SMAP overpass time of 6 am. The microwave emissivity at the top of the soil surface or vegetation canopy is given by (the polarization subscript p is suppressed in the following equations):

$$e = \frac{T_B}{T} \tag{6}$$

If the physical temperature is estimated independently, emissivity can be determined. In the SMAP formulation, ancillary surface temperature in the form of a Numerical Weather Predication model product is utilized to estimate T (see SMAP Ancillary Data Report: Surface Temperature, JPL D-53064).

The emissivity retrieved above is that of the soil as modified by any overlying vegetation and surface roughness. In the presence of vegetation, the observed emissivity is a composite of the soil and vegetation. To retrieve soil water content, it is necessary to isolate the soil surface emissivity (e^{surf}). Following Jackson and Schmugge [15], the emissivity

$$e = [1-\omega][1-\gamma][1+\left(1-e^{surf}\right)\gamma] + e^{surf}\gamma \tag{7}$$

Both the single scattering albedo (ω) and the one-way transmissivity of the canopy (γ) are dependent upon the vegetation structure, polarization and frequency. The transmissivity is a function of the optical depth (τ) of the vegetation canopy:

$$\gamma = \exp[-\tau \sec \theta] \tag{8}$$

At L-band the single scattering albedo tends to be very small, and sometimes is assumed to be zero in order to reduce dimensionality for computational purposes. For SMAP, the capability for a nonzero ω will be retained. Substituting equation 8 into equation 7 and rearranging yields

$$e^{surf} = \frac{e-1+\gamma^2+\omega-\omega\gamma^2}{\gamma^2+\omega\gamma-\omega\gamma^2} \tag{9}$$

The vegetation optical depth is also dependent upon the vegetation water content (*VWC*). In studies reported in [15], it was found that the following functional relationship between the optical depth and vegetation water content could be applied:

$$\tau = b * VWC \tag{10}$$

where *b* is a proportionality value which depends on both the vegetation structure and the microwave frequency. Since *b* is related to the structure of the overlying vegetation, it is likely that *b* will also vary with microwave polarization for at least certain types of vegetation. The variation of the *b* parameter with polarization is currently being studied by the SMAP team – it is expected that analysis of SMOS data and other field data will resolve when a polarization dependence is evident and is therefore needed to improve soil moisture retrieval accuracy for that type of vegetation.

For SMAP implementation of the SCA, values of *h*, *b*, and ω will be provided by means of a land cover look up table (an example is in Table 3 – note that the most current set of SMAP parameters used in routine processing will be available for download from NSIDC post-launch as the L2_SM_P team works to calibrate the model parameters throughout the intensive cal/val phase of the mission). The vegetation water content can be estimated using several ancillary data sources (Section 6.3). For SMAP, the baseline approach utilizes a set of land cover-based equations to estimate VWC from values of the Normalized Difference Vegetation Index (NDVI) (an index derived from visible-near infrared reflectance data from the EOS MODIS instruments now and the NPP/JPSS VIIRS instrument in the SMAP time frame) (see Equation 18). The τ-ω parameters are derived from information in the refereed literature, from past experiences and analyses conducted by the SMAP team, and from informal discussions with subject matter experts [11, 12, 15, and others]. These values will be updated, and polarization-dependent values added, as new information becomes available.

The emissivity that results from the vegetation correction is that of the soil surface, including any effects of surface roughness. These effects must be removed in order to determine the smooth surface soil emissivity (e^{soil}) which is required for the Fresnel equation inversion. One approach to removing this effect is a model described in [16] that yields the bare smooth soil emissivity:

$$e^{soil} = 1-[1-e^{surf}]\exp[h\cos^2\theta] \tag{11}$$

The $cos^2\theta$ term is sometimes changed to $cos\ \theta$ or dropped completely to avoid overcorrecting for roughness – the specific exponent to use will be determined during the SMAP CV phase. The parameter *h* is dependent on the polarization, frequency, and geometric properties of the soil surface and is related to the surface height standard deviation *s*. *h* values for different land cover types will be included in the SMAP τ-ω parameter look up table.

Emissivity is related to the dielectric properties (ε) of the soil and the viewing or incidence angle (θ). For ease of computational inversion, it can be assumed that the real component (ε_r) of the dielectric constant provides a good approximation of the complex dielectric constant; however, the complex form is also retained in the SMAP L2_SM_P processor. The Fresnel equations link the dielectric constant to emissivity. For horizontal (H) polarization (eq. 2 rewritten for emissivity):

$$e_H(\theta) = 1 - \left| \frac{\cos\theta - \sqrt{\varepsilon_r - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon_r - \sin^2\theta}} \right|^2 \tag{12}$$

and for vertical (V) polarization the relationship is (eq. 3 rewritten for emissivity):

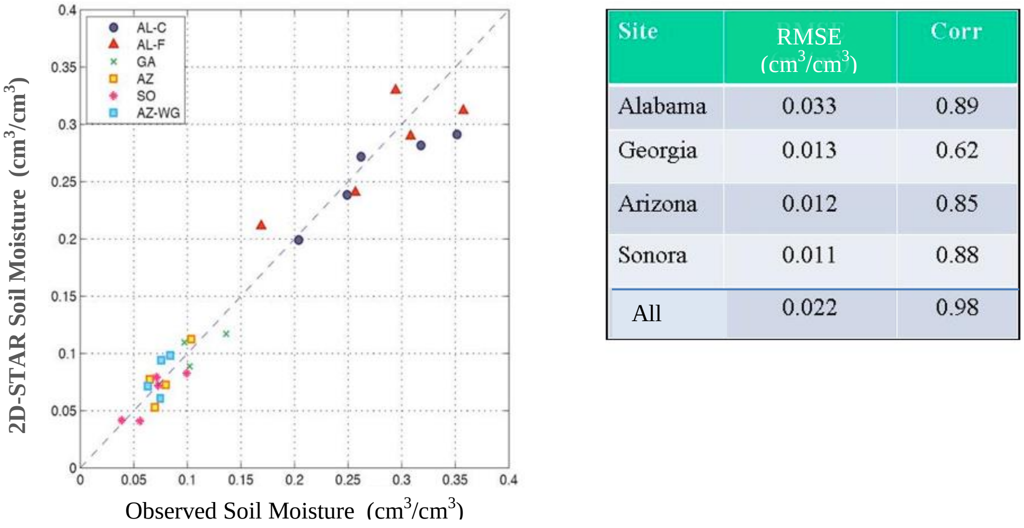
$$e_V(\theta) = 1 - \left| \frac{\varepsilon_r \cos\theta - \sqrt{\varepsilon_r - \sin^2\theta}}{\varepsilon_r \cos\theta + \sqrt{\varepsilon_r - \sin^2\theta}} \right|^2 \tag{13}$$

The dielectric constant of soil is a composite of the values of its components – air, soil, and water – which have greatly different values. A dielectric mixing model is used to relate the estimated dielectric constant to the amount of soil moisture. As described in Section 2.2, there are three dielectric mixing models under consideration that seem to perform differently in different soil moisture ranges (Wang and Schmugge [21], Dobson et al. [20], and Mironov [19]). The SMOS team is evaluating the relative merits of these dielectric models and their impact on overall soil moisture retrieval accuracy, and is currently using mainly the Mironov model. The SMAP project plans to closely monitor and review the SMOS results in making the final selection of a dielectric model for SMAP. The SMAP processor has options for all three of the dielectric models, with the Mironov model also the current SMAP baseline.

Table 3. Example of Look Up Table of Algorithm Parameters by IGBP Class (note: not the SMAP final set)

ID	MODIS IGBP land classification	*s*	*h*	*b*	*ω*	*Stem factor*
0	Water Bodies	--	0	0	0	--
1	Evergreen Needleleaf Forests	1.60	0.160	0.100	0.050	15.96
2	Evergreen Broadleaf Forests	1.60	0.160	0.100	0.050	19.15
3	Deciduous Needleleaf Forests	1.60	0.160	0.120	0.050	7.98
4	Deciduous Broadleaf Forests	1.60	0.160	0.120	0.050	12.77
5	Mixed Forests	1.60	0.160	0.110	0.050	12.77
6	Closed Shrublands	1.00	0.110	0.110	0.050	3.00
7	Open Shrublands	1.10	0.110	0.110	0.050	1.50
8	Woody Savannas	1.00	0.125	0.110	0.050	4.00
9	Savannas	1.00	0.156	0.110	0.080	3.00
10	Grasslands	1.56	0.156	0.130	0.050	1.50
11	Permanent Wetlands	1.00	0	0	0	4.00
12	Croplands - Average	1.08	0.108	0.110	0.050	3.50
	- Wheat	0.83	0.083	TBD	TBD	TBD
	- Mixed (Wheat, Barley, Oats)	1.08	0.108	TBD	TBD	TBD
	- Corn	0.94	0.094	TBD	TBD	TBD
	- Soybean	1.48	0.148	TBD	TBD	TBD
13	Urban and Built-up Lands	--	0	0.100	0.030	6.49
14	Crop-land/Natural Vegetation Mosaics	1.30	0.130	0.110	0.065	3.25
15	Snow and Ice	--	0	0	0	0
16	Barren	1.50	0.150	0	0	0

An example of retrieved soil moisture using the SCA and site-specific correction parameters is shown in Figure 14:



Site	RMSE (cm^3/cm^3)	Corr
Alabama	0.033	0.89
Georgia	0.013	0.62
Arizona	0.012	0.85
Sonora	0.011	0.88
All	0.022	0.98

Figure 14. Soil moisture retrieval error based on L-band H-polarized T_B airborne observations. The vegetation parameter *b* and roughness parameter *h* are optimized using ground measurements of soil moisture from the SMEX03 and SMEX04 field campaigns [[30]; also D. Ryu, personal communication]. The canopy and soil temperatures are assumed to be equal (i.e., $T_c = T_s = T_{eff}$) under the hydraulic equilibrium assumption.

4.2.1 Nonlinear VWC Correction

In terms of soil moisture retrieval performance, the Hydros OSSE [7] revealed that the SCA could produce biased retrievals based on linear VWC correction aggregated from high-resolution vegetation data. However, two relatively simple approaches were developed to create an effective VWC that helps to reduce the bias and overall RMSE in retrieved soil moisture [31, 32]. For example, from [32], the observed $T_B{}^{obs}$ integrated over the IFOV can be written as (assuming uniform soil moisture, soil temperature, surface roughness, and antenna gain):

$$\begin{aligned} T_B{}^{obs} &= \frac{1}{N}\sum_{i=1}^{N} T_{Bi} \\ &= \frac{1}{N}\sum_{i=1}^{N} T_{si}[1 - R_{\circ i}(m_{vi})e^{-h_i}e^{-2\tau_i}] \\ &= T_s\left[1 - R_\circ(m_v)e^{-h}\frac{1}{N}\sum_{i=1}^{N} e^{-2\tau_i}\right] \\ &= T_s\,[1 - R_\circ(m_v)e^{-h}e^{-2\tau^*}] \end{aligned}$$

$$e^{-2\tau^*} = \frac{1}{N}\sum_{i=1}^{N} e^{-2\tau_i}\ \text{, or } \tau^* = -\frac{1}{2}\ln\left[\frac{1}{N}\sum_{i=1}^{N} e^{-2\tau_i}\right] \quad \text{(nonlinear VWC correction)} \tag{14}$$

While these methods have been successfully applied to the SCA (Figure 15), their value to the other candidate soil moisture retrieval algorithms (DCA and MPRA) is currently being investigated.

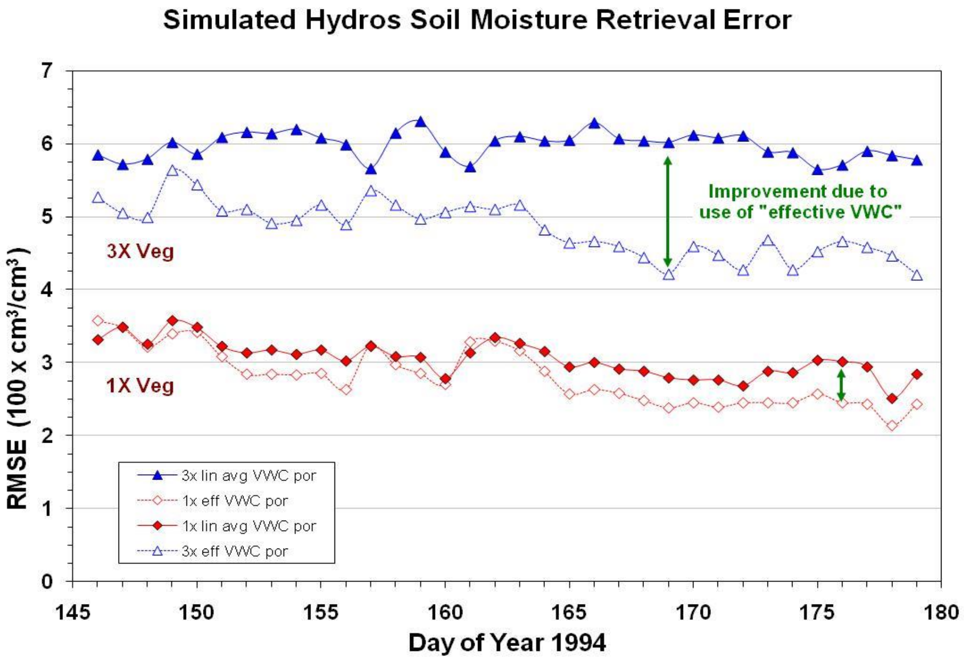


Figure 15. Improvement in simulated Hydros soil moisture retrieval error using a simple effective VWC correction with the SCA algorithm for existing vegetation (1X) and for artificially increased simulated vegetation amounts (3X) [31].

4.3 Dual Channel Algorithm

The Dual Channel Algorithm (DCA or DPOL) is an extension of the SCA described in the previous section — it uses both H-polarized and V-polarized T_B observations to simultaneously retrieve soil moisture and VWC [33]. The inversion mechanism of the DCA starts with feeding the *tau-omega* model (Section 2.1, equation 1) with initial guesses of soil moisture and VWC. The quantities are then adjusted iteratively until the difference between the computed and observed T_B observations is minimized in a least square sense. Similar to the SCA, 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.

The DCA has been used with reasonable success in the 2007 CLASIC field campaign conducted in Oklahoma, USA [34]. Figure 16 shows the variability of the retrieved soil moisture and vegetation opacity.

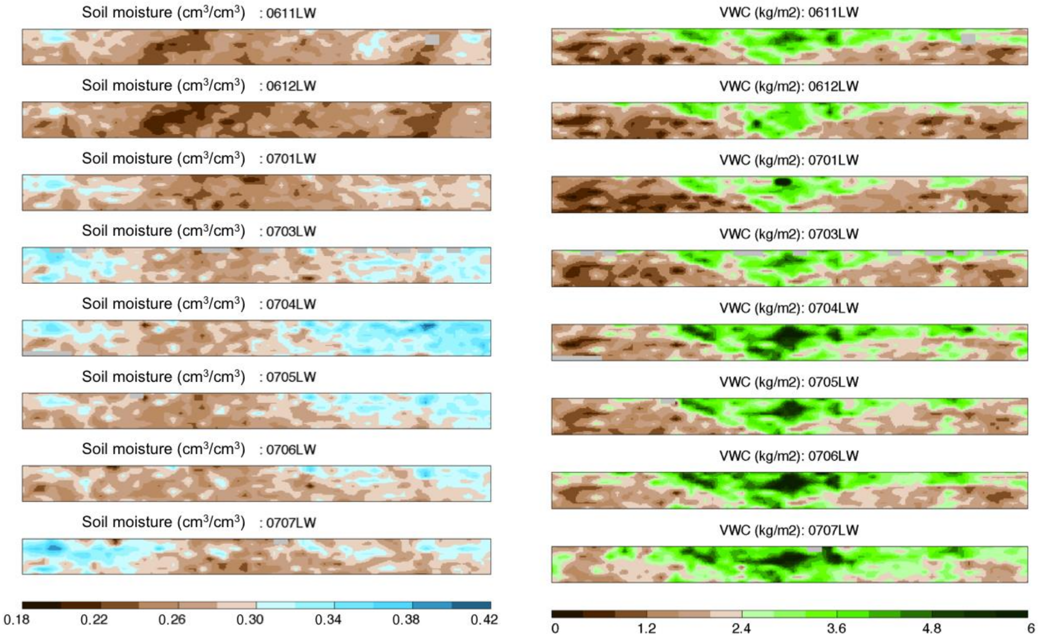


Figure 16. Simultaneous retrieval of soil moisture and VWC using the DCA with airborne PALS data from the 2007 CLASIC field campaign. The spatial and temporal variability of the two parameters retrieved by the DCA agree with the actual wetting-drying pattern observed during the campaign.

The ability of the DCA to simultaneously estimate two geophysical parameters may come with a penalty. While the additional channel allows for estimation of VWC, it also brings in additional T_B errors (uncorrelated between V and H channels) that may adversely affect retrieval accuracy. Also, an assumption implicit in this algorithm is that the optical depth is identical for both polarizations. Exactly which effect outweighs the other is under investigation through simulations using the SMAP SDS Testbed (to be described in Section 5).

4.4 Land Parameter Retrieval Model

The Microwave Polarization Ratio Algorithm (MPRA), based on the Land Parameter Retrieval Model [35], is an index-based retrieval model that uses dual polarization channels at a single microwave frequency (typically C or X-band) to derive soil moisture and vegetation optical depth. As implemented on multi-frequency satellites such as AMSR-E, it also uses the Ka-band V-polarized channel to retrieve physical temperature of the surface. Only a few studies [36] have examined the applicability of this model at L-band frequencies, although analysis of SMOS data with LPRM is currently underway [R. de Jeu, personal communication, 2011]. Because there are no Ka-band V-polarized T_B observations available from the SMAP instruments, surface temperature will be obtained using ancillary data sets as with the other L2_SM_P algorithms.

In the MPRA, the radiative transfer model operates on two assumptions: (1) the soil and canopy temperatures are considered equal (T), and (2) the vegetation transmissivity (γ) and the single-scattering albedo (ω) are the same for both H and V polarizations. If e_s is the soil emissivity, the T_B can be expressed by the *tau-omega* model (Eq. 1) with T_C = $T_S = T$:

$$T_B = e_s \gamma T + (1-\omega)(1-\gamma)T + \gamma(1-e_s)(1-\omega)(1-\gamma)T \tag{15}$$

The single scattering albedo ω represents the loss of energy in the canopy and is assumed by MPRA to be constant globally, in contrast to the other L2_SM_P algorithms where a nonzero ω is assumed to be a function of land cover type and is input as an ancillary parameter (from look up table). In a previous study [36] using L-band T_B from an aircraft experiment in Australia, the global ω was set equal to 0; however, for the ongoing SMOS analyses, $\omega = 0.05$ is being used as the global value [R. de Jeu, personal communications, 2011].

The Microwave Polarization Difference Index (MPDI) and the observed emissivity (e_H and e_V) are used in MPRA to derive the vegetation optical depth [38], which in turn is used to calculate the transmissivity (γ). The MPDI and vegetation optical depth are calculated as follows:

$$MPDI = \frac{T_{BV} - T_{BH}}{T_{BV} + T_{BH}} \tag{16}$$

$$\tau = \cos(\theta)\, ln\left[ad + \sqrt{(ad)^2 + a + 1}\right] \tag{17}$$

where a and d are $a = 0.5\,[(e_V - e_H) / \text{MPDI} - e_V - e_H]$ and $d = 0.5\,\omega / (1 - \omega)$.

The observed emissivity can also be modeled as a function of soil moisture and temperature in three steps. First, the dielectric constant is calculated as a function of soil moisture, temperature, and soil type following the parameterization of Wang and Schmugge [21]. Second, the smooth surface emissivity is calculated by applying the Fresnel equations. Third, this emissivity is corrected for roughness effects according to

Choudhury *et al.* [16]. The roughness parameterization requires an estimate of the parameter *h*, which is dependent on the polarization, frequency and geometric properties of the soil surface. In previous applications of this approach using C- and X- band, it was found acceptable to set this roughness parameter to a constant; with SMAP, a land cover-based roughness will likely be used for consistency with the other algorithms. With this set of equations, soil moisture is retrieved in an optimization routine that minimizes the error between the modeled and observed H-polarized brightness temperatures. The vegetation optical depth at this optimized soil moisture value is an additional retrieval result.

4.5 Extended Dual Channel Algorithm (E-DCA)

The E-DCA is a variant of DCA. Like DCA, E-DCA uses both the vertically and horizontally polarized T_B observations to solve for soil moisture and vegetation optical depth. In E-DCA, however, the cost function (Φ^2) is formulated in a way different from that of DCA. Instead of minimizing the sum of squares of the differences between the observed and estimated T_Bs as in DCA (Equation 1 above), the E-DCA attempts to minimize the sum of squares of the difference between the observed and estimated normalized polarization differences (expressed in natural logarithm) and the difference between the observed and estimated T_Bs (also expressed in natural logarithm) as follows:

$$\min \Phi^2_{\mathrm{E-DCA}} = \left[\log\left(\frac{T^{\mathrm{obs}}_{\mathrm{B,v}} - T^{\mathrm{obs}}_{\mathrm{B,h}}}{T^{\mathrm{obs}}_{\mathrm{B,v}} + T^{\mathrm{obs}}_{\mathrm{B,h}}}\right) - \log\left(\frac{T^{\mathrm{est}}_{\mathrm{B,v}} - T^{\mathrm{est}}_{\mathrm{B,h}}}{T^{\mathrm{est}}_{\mathrm{B,v}} + T^{\mathrm{est}}_{\mathrm{B,h}}}\right)\right]^2 + \left[\log\left(T^{\mathrm{obs}}_{\mathrm{B,h}}\right) - \log\left(T^{\mathrm{est}}_{\mathrm{B,h}}\right)\right]^2 \tag{18}$$

In each iteration step, soil moisture and vegetation opacity are adjusted simultaneously until the cost function attains a minimum in a least square sense. It is clear that when both Φ^2_{DCA} and $\Phi^2_{\mathrm{E-DCA}}$ attain their theoretical minimum value (i.e. zero) in the absence of uncertainties of modeling, observations, and ancillary data, $T^{\mathrm{obs}}_{\mathrm{B,v}} = T^{\mathrm{est}}_{\mathrm{B,v}}$ and $T^{\mathrm{obs}}_{\mathrm{B,h}} = T^{\mathrm{est}}_{\mathrm{B,h}}$, implying that DCA and E-DCA converge to the same solutions. The advantage of E-DCA over DCA, however, is apparent when in reality there is finite uncertainty (e.g., a dry bias associated with the ancillary soil temperature data) -- this uncertainty will be cancelled from the numerator and denominator in the calculation of the normalized polarization difference in $\Phi^2_{\mathrm{E-DCA}}$, leaving such uncertainty affecting only one component of the cost function instead of two components as in Φ^2_{DCA}. This reduction in the impact of soil temperature uncertainty on soil moisture retrieval should make E-DCA more tolerant of soil temperature uncertainty, resulting in fewer instances of retrieval failure than DCA. At present, there are a few caveats associated with E-DCA: (1) its exact performance is being evaluated in the ongoing Cal/Val activities and is not included in the beta assessment report, and (2) the choice of the horizontally polarized T_B in the $\Phi^2_{\mathrm{E-DCA}}$ formulation is subject to further assessment as analyses of new observations and Cal/Val data become available.

4.6 Algorithm Error Performance

One measure of algorithm performance is determining the accuracy of the retrieved soil moisture in a root square sense. Different algorithms respond differently to uncertainty in a given model / ancillary parameter. One initial test performed by the SMAP team involved retrieving soil moisture from one year of global simulated SMAP brightness temperatures, varying one parameter in turn while keeping the other parameters constant with no error (the SMAP Algorithm Testbed will be described in Section 5). Table 4 lists the error in retrieved volumetric soil moisture (in cm^3/cm^3) for each of the four SMAP L2_SM_P candidate algorithms over the full range of soil and vegetation water content (VWC) conditions encountered in the global simulation. The first column lists the parameter and its assigned error. Across this full range of conditions, with error only in one parameter at a time, all of the algorithms appear to perform to acceptable levels in retrieving soil moisture.

A more stringent simulation is to assign some error to all parameters simultaneously and then assess the accuracy in retrieved soil moisture. Figure 17 shows the results obtained when the indicated errors were applied to the indicated parameters and soil moisture was retrieved for one year and compared to the “true” soil moisture. All algorithms meet the SMAP mission requirements of retrieving soil moisture to an accuracy of 0.04 cm^3/cm^3 for areas within the SMAP mask where VWC is $\leq$ 5 kg/m^2 when averaged across all VWC bins. For this simulation, parameters such as b and h were assumed to be the same for both H and V polarization. This assumption will be re-examined as new information is obtained (through analysis of SMOS and other field data) regarding quantification of any polarization dependence of any of the algorithm parameters.

Additional studies using a new more realistic global simulation (GloSim3) confirmed the initial results of algorithm performance. Although the baseline SCA algorithm performed somewhat better in these simulations, all four algorithms have been coded in the L2_SM_P processor and will be analyzed during the official SMAP CV period. The algorithm which produces the best overall soil moisture retrievals will then be used in the first SMAP bulk reprocessing at Launch + 15 months.

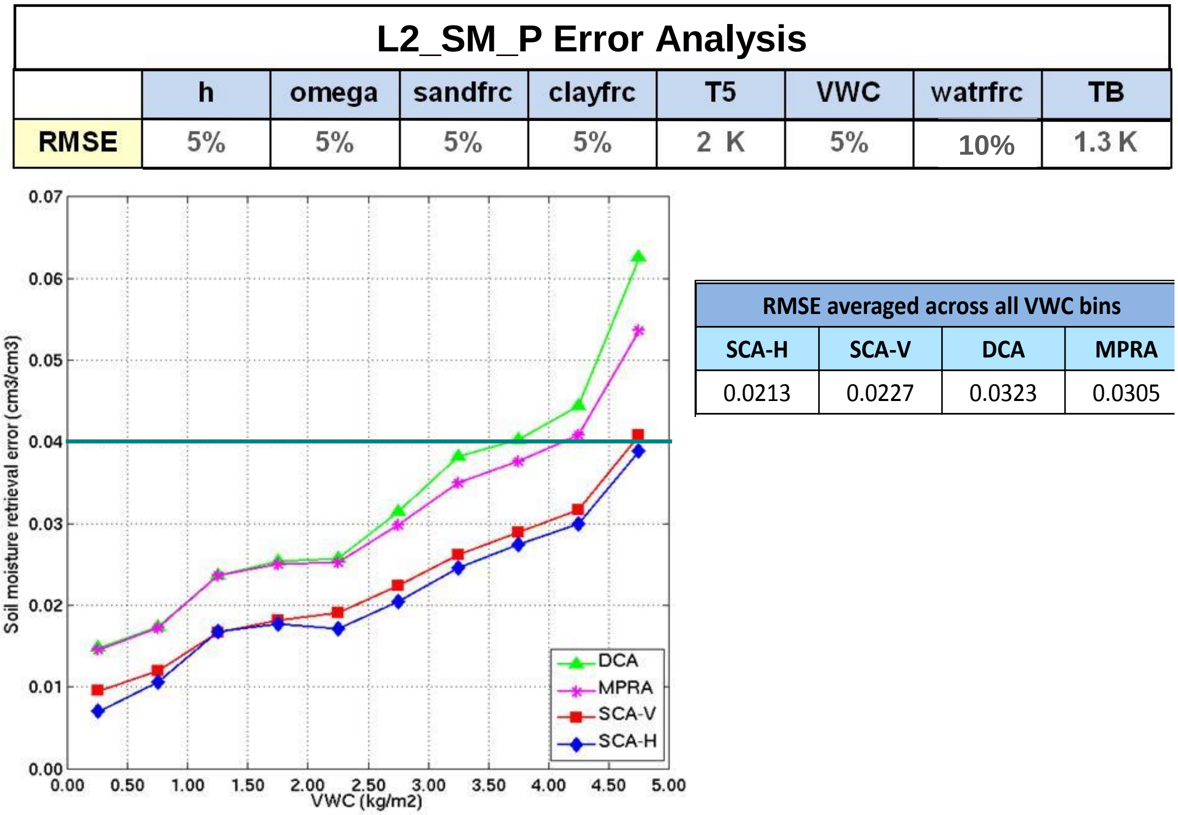
4.7 Algorithm Downselection

Downselection of the baseline SMAP algorithm for the L2_SM_P product will be based on a combination of demonstrated higher accuracy in retrieved soil moisture, lower bias, better overall performance across land cover classes globally, and operational considerations. Performance results will be assessed pre-launch using: (1) simulated data from GloSim3, (2) analyses of past and new field campaign data, especially SMAPVEX-12 and ComRAD (APEX12), and (3) application of SMAP algorithms to SMOS T_B data at the SMAP incidence angle of 40°. Post-launch, performance assessments will be made based on comparisons to *in situ* CV site data and other CV approaches (section 7) during the official SMAP calibration / validation period in the first year after the beginning of routine science operations. Prior to the first bulk reprocessing at Launch + 15 months, a downselection to the best-performing algorithm will be made. However, algorithm performance will continually be assessed throughout the SMAP mission, and one of the other algorithms could be substituted for the baseline algorithm should a problem be detected later. This approach insures that the best possible algorithm will be used in order to deliver the best possible product to the public and NASA archives.

Table 4. Simulated Retrieval Error by Parameter for each Algorithm

	Baseline	Option 1	Option 2	Option 3
Model/Ancillary Uncertainty	SCA (H) RMSE (cm³/cm³)	SCA (V) RMSE (cm³/cm³)	DCA RMSE (cm³/cm³)	MPRA RMSE (cm³/cm³)
Gridding + Aggregation	0.00612	0.00581	0.00591	0.00582
5% h	0.00645	0.00595	0.00595	0.00583
5% omega	0.00629	0.00605	0.00619	0.00611
5% sand fraction	0.00729	0.00699	0.00702	0.00697
5% clay fraction	0.00615	0.00585	0.00594	0.00585
2K T5	0.00871	0.01000	0.01120	0.01200
5% VWC	0.00656	0.00608	–	–
10% VWC	0.00717	0.00647	–	–
5% water fraction	0.00612	0.00582	0.00591	0.00582
10% water fraction	0.00612	0.00582	0.00591	0.00583
20% water fraction	0.00614	0.00584	0.00593	0.00583
1.3 K TB	0.00681	0.00674	0.00828	0.00951
RSS	0.0203	0.0201	0.0205	0.0214

Based on GloSim.



L2_SM_P Error Analysis

	h	omega	sandfrc	clayfrc	T5	VWC	watrfrc	TB
RMSE	5%	5%	5%	5%	2 K	5%	10%	1.3 K

RMSE averaged across all VWC bins			
SCA-H	SCA-V	DCA	MPRA
0.0213	0.0227	0.0323	0.0305

Figure 17. Simulated error performance of all L2_SM_P candidate retrieval algorithms. One year of simulated global SMAP main-beam H- and V-polarized L1C_TB (GloSim3) were used to retrieve soil moisture using perturbed model and ancillary parameters. Static water T_B correction was applied after T_B gridding.

4.7.1 Preliminary Results of Using SMOS Data to Simulate SMAP

Microwave observations from the SMOS mission have been reprocessed to simulate SMAP observations at a constant incidence angle of 40^o (details of the SMOS reprocessing will be explained in Section 7.1.1). This procedure provides a brightness temperature data set that closely matches the observations that will be provided by the SMAP radiometer. SMOS brightness temperatures provide a global real-world, rather than simulated, input for evaluating the different SMAP radiometer-only soil moisture algorithm alternatives. The use of real-world global observations will also help in the development and selection of different land surface parameters (roughness and vegetation) and ancillary observations needed for the L2_SM_P soil moisture algorithms. The ancillary data sets required are dependent on the choice of the soil moisture algorithm. For example, for its needed vegetation information, the single channel algorithm (SCA) might use (a) SMOS-estimated vegetation optical depth, (b) MODIS-based vegetation climatology data, or (c) actual MODIS observations.

Initial results using the SCA-H with a SMOS-based simulated SMAP T_B data set and MODIS-based vegetation (NDVI) climatology data are presented here. For this preliminary analysis, the roughness parameter (h), vegetation parameter (b), and the single scattering albedo (ω) were assumed constant for all land cover classes ($h = 0.1$, $b = 0.08$, $\omega = 0.05$). In subsequent analyses, these parameters will be further refined for different land cover classes as information becomes available and the τ-ω parameter look up table is updated. ECMWF estimates of soil temperature for the top layer (as provided as part of SMOS ancillary data) were used to correct for surface temperature effects and to derive microwave emissivity. ECMWF data were also used for precipitation forecasts and to note the presence of snow and frozen ground.

Figure 18 shows the average soil moisture estimated using the SCA algorithm with the SMOS-simulated SMAP TB data for the ascending orbits (overpass time of 6 AM) for two time periods: (a) June 1-10 and (b) July 1-10, 2011. A MODIS-based NDVI and supporting relationships between NDVI and optical depth were used to correct for vegetation effects. The soil moisture spatial patterns are consistent with geographical features. The estimated soil moisture is very low for desert and arid regions (Africa, Middle East, Central Asia, and Central Australia). High values of soil moisture were observed for forested areas in northern latitudes (Canada and Russia). High soil moisture is also observed over portions of South America.

In June 2011, the northern latitudes of Canada and Siberia had either snow on the ground or the top soil layer was frozen based on the ECMWF forecasts. These areas were flagged during the soil moisture retrieval process. The surface temperature increased in these areas above the freezing mark for the second test period in July 2011. During July, a large part of South-East Asia, Northern South America and Central America is flagged because ECMWF forecasts indicated precipitation at the time of the SMOS overpasses.

Soil moisture retrieved using the SCA-H with SMOS-simulated SMAP T_B for January 2010 – May 2013 was compared to data from *in situ* soil moisture networks in

USDA ARS watersheds that have previously been extensively used in satellite-based soil moisture validation (Jackson et al., 2010, 2011). Figure 19 shows the comparison between observed and estimated soil moisture over the Little River (LR), Little Washita (LW), Walnut Gulch (WG), and Reynolds Creek (RC) watersheds for SMOS ascending orbits (overpass time of 6 AM). Table 5 shows the statistical performance of the SCA-H algorithm over these watersheds. The overall range of soil moisture conditions for the period of record was fairly wide. The SCA retrievals over LR have a low bias and RMSE. For LW, most of the error is because of a dry bias in the soil moisture estimates (-0.028 cm^3/cm^3). The SCA-H soil moisture retrievals for the WG watershed have a good agreement with the *in situ* data with near zero bias. In order to eliminate the effect of snow on SMOS/SMAP retrievals over the RC watershed, only data from July-September were used in the analysis. The SCA results over RC have an underestimation bias that results in a high RMSE. The correlation between the *in situ* observations and estimated soil moisture is good for all the watersheds. As noted earlier, constant values of the roughness parameter, vegetation parameter, and single scattering albedo were used in this analysis, possibly contributing to the observed soil moisture bias.

A second preliminary analysis regridded SMOS 40° T_B data onto the SMAP 36-km EASE2 grid and used SMAP-gridded ancillary data with the SCA to retrieve soil moisture. These retrievals were also compared to in situ soil moisture data from the USDA watersheds without site-specific calibration for *h*, *b*, and *ω* (RMSE = 0.037 cm^3/cm^3, R=0.71) and with site-specific calibration (RMSE = 0.021 cm^3/cm^3, R=0.78). These results are encouraging for the potential of SMAP to meet its required soil moisture accuracy for the L2_SM_P product.

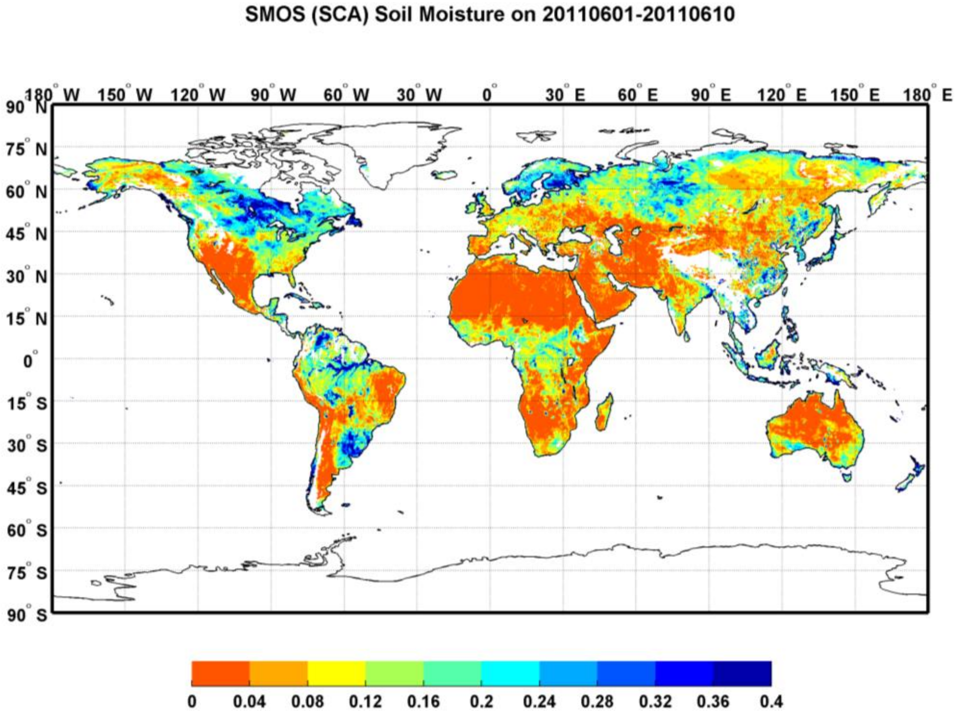


Figure 18a. Average estimated soil moisture using the single channel algorithm (SCA) for SMOS ascending orbits for the period of June 1-10, 2011.

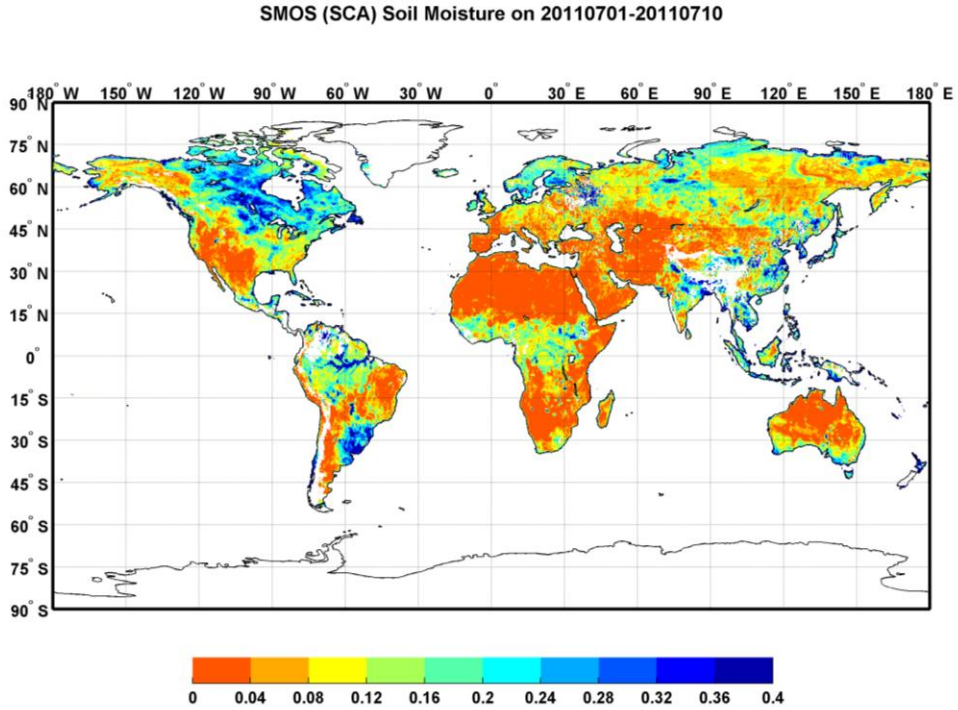


Figure 18b. Average estimated soil moisture using the single channel algorithm (SCA) for SMOS ascending orbits for the period of July 1-10, 2011.

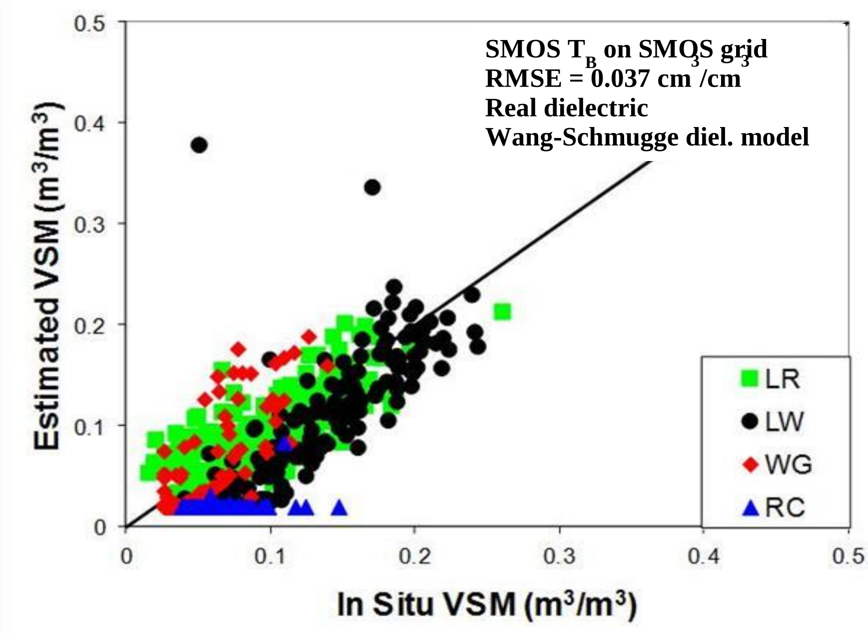


Figure 19. Comparison of estimated soil moisture using SMOS-simulated SMAP T_B with *in situ* observations over USDA ARS watershed sites for ascending orbits (6 AM overpass time) for January 2010-May 2013. One of the two outliers is due to the likely presence of wet snow not predicted by ECMWF; the other outlier is due to an unpredicted active rain event.

Table 5. Statistical summary of the SMOS/SMAP/SCA retrieval algorithm over the USDA watersheds for ascending orbits (6 am overpass time), January 2010 - May 2013.

Watershed	Count	RMSE	R	Bias
Little River, GA	247	0.028	0.767	-0.003
Little Washita, OK	245	0.047	0.841	-0.028
Walnut Gulch, AZ	231	0.025	0.789	-0.008
Reynolds Creek, AZ	74	0.050	0.219	-0.045
Overall	**797**	**0.037**	**0.745**	**-0.016**

RMSE (Root Mean Square Error) and Bias are in cm^3/cm^3 (or m^3/m^3). R = correlation coefficient, N = number of samples. Overall results are essentially the same when binned over 6-month summer and winter seasons (summer RMSE = 0.038, winter RMSE = 0.034).

5. SMAP ALGORITHM DEVELOPMENT TESTBED

The SMAP project has developed and is currently using the Algorithm Development Testbed, a software infrastructure designed to simulate the passive and active microwave observations acquired by SMAP. The Testbed attempts to address the following objectives:

1. To obtain a more rigorous assessment of the soil moisture measurement capability for SMAP relative to that reported in the previous Hydros Risk Mitigation Study,
2. To evaluate how the soil moisture and freeze/thaw measurement capability for SMAP is impacted by different science, instrument, and/or mission trades,
3. To evaluate the relative merits of different microwave models, retrieval algorithms, and ancillary data for meeting the SMAP soil moisture and freeze/thaw science objectives, based on a common set of input and processing conditions, and
4. To provide an end-to-end system that can be used to test the integrated suite of SMAP science product algorithms as a prototype for the SMAP Science Data Processing System (SDPS).

Of relevance to this ATBD, the Testbed can be used to evaluate the performance of different retrieval algorithms and to establish the corresponding error budgets based on a common set of geophysical and instrument conditions. Within the Testbed environment, the following three approaches are adopted: (1) numerical analyses based on land surface model (LSM) outputs, (2) numerical analyses based on Monte Carlo simulations, and (3) algorithm validation based on observations from field campaigns that feature L-band active and passive observations (e.g. SGP99, SMEX, CLASIC, and SMAPVEX). These components and the interrelationships among them are summarized in Figure 20.

With LSM outputs, the Testbed can generate simulated brightness temperature and radar backscatter observations according to SMAP's orbital and instrument sampling pattern. To date one full year of LSM outputs over global and the continental United

States (CONUS) domains are available for simulations. These simulations have been used for initial assessments of algorithm performance as described in the last section. A global map of retrieved soil moisture RMSE from one of these simulations is shown in Figure 21. As evident in the figure, retrieval error varies depending on the amount of vegetation, among other factors. Antenna sidelobe contamination along coastlines and river/lake boundaries also leads to high retrieval errors. Overall, the retrieved soil moisture RMSE stays below 0.04 cm^3/cm^3 over areas with low to moderate amounts of vegetation.

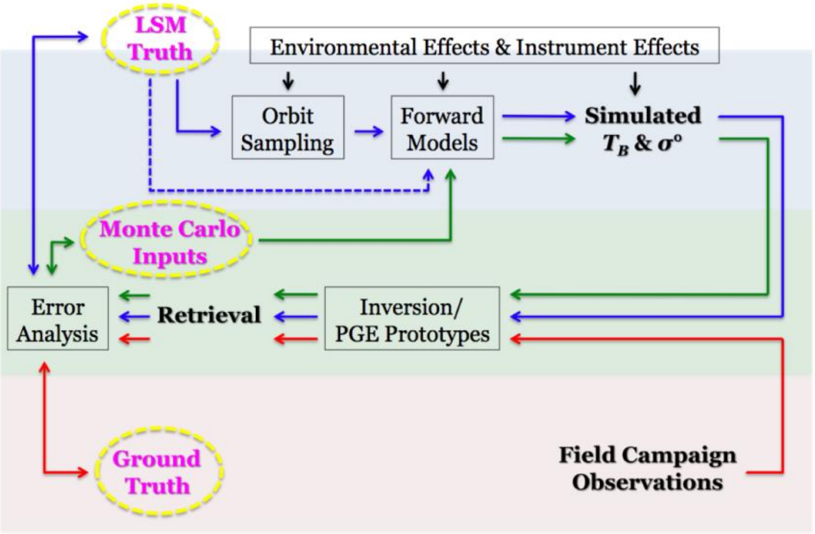


Figure 20. The various simulation modules and input data sources for the SMAP Algorithm Development Testbed.

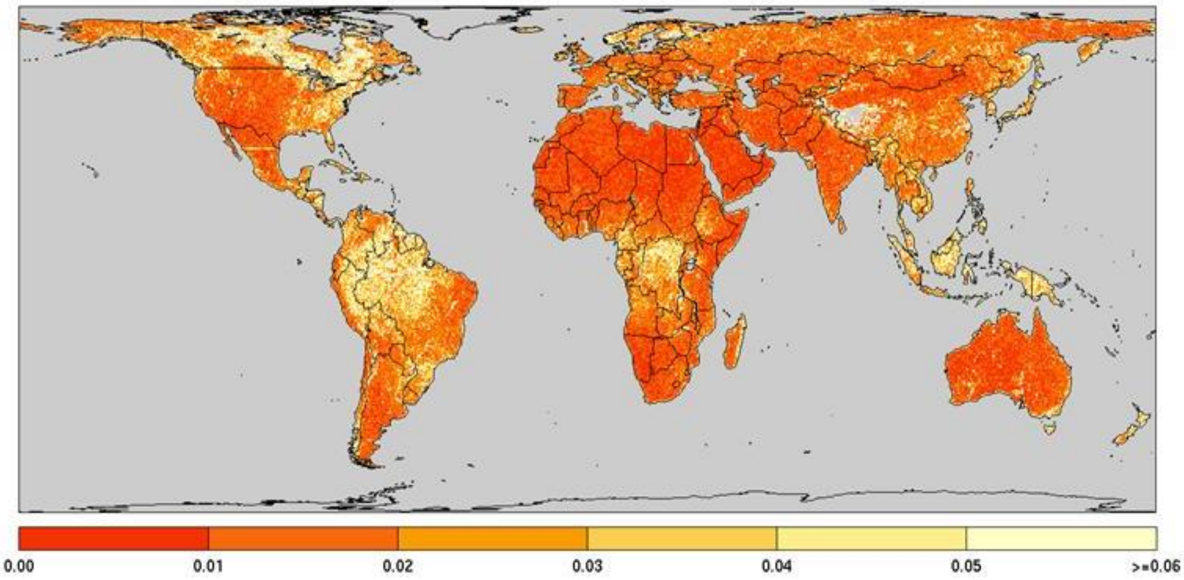


Figure 21a. Global soil moisture retrieval error based on simulations using LSM outputs and a L2_SM_P baseline algorithm.

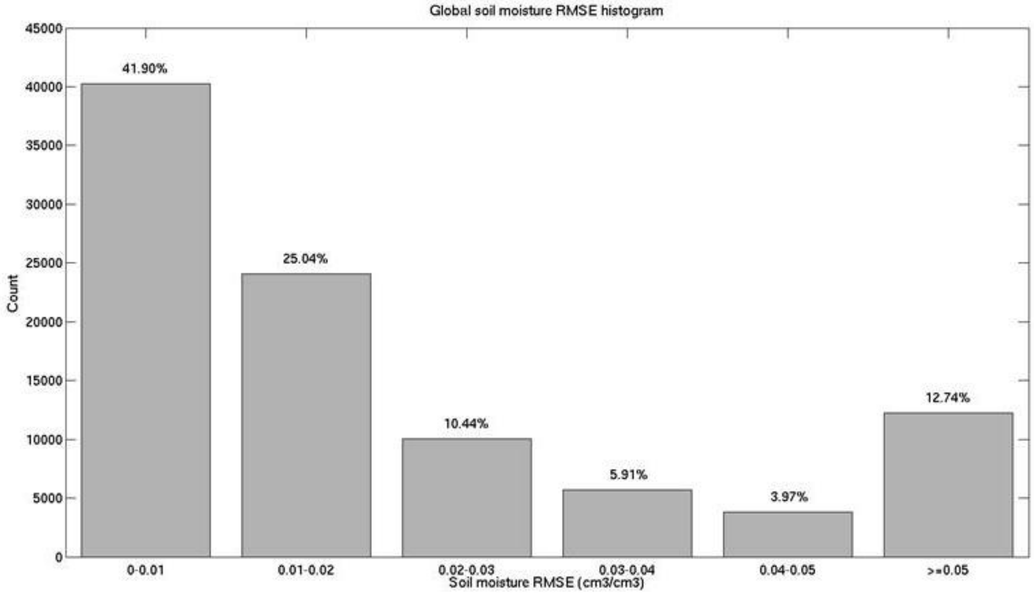


Figure 21b. Histogram of global soil moisture retrieval error based on simulation shown in (a). Approximately 83% of the area in the SMAP land mask has a soil moisture retrieval error ≤ 0.04 cm^3/cm^3.

New improved global simulations (GloSim2 and 3) are essentially complete. They enhance the realism of SMAP forward simulations by adding:

- Consistent global input forcings based on GMAO global nature run data,
- Finer grid resolutions (9-km dynamic fields and 1-km static fields),
- Finer temporal resolution (hourly),
- Improved ancillary datasets (e.g., soil texture, VWC, water fraction, etc),
- Enhanced radar forward modeling (more data cubes),
- More realistic error modeling (consistent spatial scaling of random and non-random perturbations), and
- Data format closer to the SMAP Data Product Specifications.

6. ANCILLARY DATA SETS

6.1 Identification of Needed Parameters

Ancillary data sets are defined as external data sets that are required as inputs to SMAP retrieval algorithms in the generation of the standard L2/3 products. Ancillary data needed by the SMAP mission fall into two categories -- static ancillary data are data that do not change during the mission while dynamic ancillary data require periodic updates in time frames ranging from seasonally to daily. Static data include parameters such as permanent masks (land / water / forest / urban / mountain), the grid cell average elevation and slope derived from a DEM, permanent open water fraction, and soils information (primarily sand and clay fraction). All of the static ancillary data will be resampled to the same 3, 9, and 36-km SMAP EASE grids as the output products, and will be available to any algorithm or end user who needs them. The dynamic ancillary data include land cover, surface roughness, precipitation, vegetation parameters, and effective soil temperatures. Although most ancillary data are by definition external to SMAP, the SMAP HiRes radar will provide key pieces of information to the L2_SM_P algorithms including the open water fraction and the frozen ground flag (see L2_SM_A and L3_FT_A ATBDs). While the exact types of ancillary datasets needed are specific to a given retrieval algorithm, all standard L2/3 products require some ancillary datasets to meet the specified retrieval accuracies.

Table 6 lists the fourteen ancillary data parameters identified as required by one or more of the SMAP product algorithms along with the primary source of information for that parameter (in all cases, there are alternative options for these parameters from climatological data sets, forecast models, or data sets acquired in past or current missions). The choice of which ancillary data set to use for a particular SMAP product is based on a number of factors, including its availability and ease of use, its inherent error and resulting impact on the overall soil moisture or freeze/thaw retrieval accuracy, and its compatibility with similar choices made by the SMOS mission. Latency, spatial resolution, temporal resolution, and global coverage are also important. The choice of a primary source for each of the fourteen ancillary data parameters is fully documented in individual SMAP Ancillary Data Reports which are available to the user community (these data reports are included in the list of SMAP Reference Documents at the front of this ATBD – see also http://smap.jpl.nasa.gov/science/dataproducts/ATBD/).

6.2 Soil Temperature Uncertainty

Errors in ancillary data are factored into the error budgets for each of the SMAP candidate soil moisture retrieval algorithms during SMAP simulations (Section 4.5). A major issue prelaunch is to quantify the expected errors of these ancillary data parameters, especially the error in the effective soil temperature parameter, since it requires the most frequent (daily) updates and is used by all L2_SM_P algorithms. The time resolution of the soil temperature (T_{soil}) is also important – currently, the major global forecast centers (including NCEP, ECMWF, and GMAO) produce T_{soil} forecast products at a time resolution relevant to SMAP (minimum time resolution of 3 hours). For L2/3_SM_P processing, a local 6:00 am T_{soil} will be generated by interpolating in time between the closest available 3-hourly T_{soil} snapshots. A preliminary assessment of the accuracy of ECMWF forecast temperatures was made for an area in the central U.S. encompassing the state of Oklahoma and compared against ground truth temperatures from the Oklahoma Mesonet for every day in 2003 (Figure 22), with an RMSE between forecast and measured surface temperatures of approximately 2.4°C [31].

Table 6. Anticipated Primary Sources of Ancillary Parameters

1	**Soil Temperature**	**GSFC GMAO [consistency ↔ L2-L4]**
2	**Surface Air Temperature**	**GSFC GMAO**
3	**Vegetation Water Content (VWC)**	**MODIS NDVI [T. Jackson/R. Hunt approach]**
4	**Soil Attributes (sand & clay fraction)**	**Combination of HWSD (global), regional data sets (STATSGO-US, ASRIS-Australia, NSD-Canada), FAO**
5	**Urban Area**	**GRUMP data set – Columbia University**
6	**Open Water Fraction**	**a priori static water fraction from MODIS MOD44W to be used in conjunction with open water fraction from SMAP HiRes radar**
7	**Crop Type**	**combination of USDA Cropland Data Layer, AAFC-Canada, Ecoclimap-Europe**
8	**Land Cover Class**	**MODIS IGBP; crop class will be further subdivided into four general crop types**
9	**Precipitation**	**GSFC GMAO**
10	**Snow**	**Snow & Ice Mapping System (IMS) - NOAA**
11	**Mountainous Area [DEM]**	**GMTED-2010**
12	**Permanent Ice**	**MODIS IGBP**
13	***b*, *ω*, and *τ* Vegetation Parameters**	**land cover-driven table lookup**
14	***h* Roughness Parameter**	**land cover-driven table lookup**

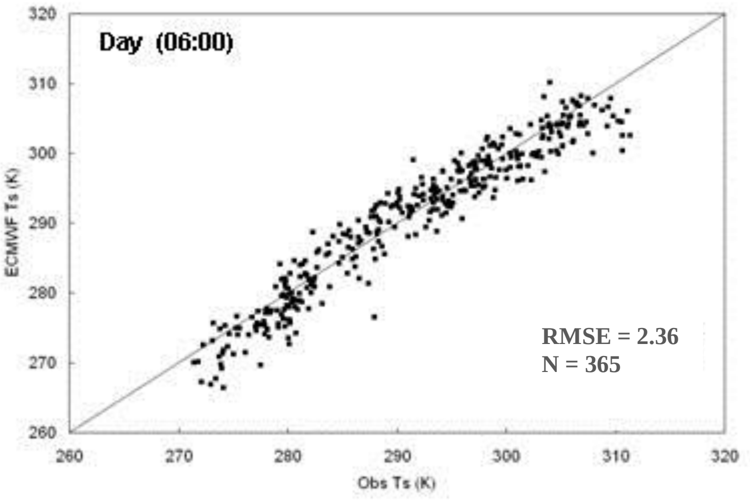
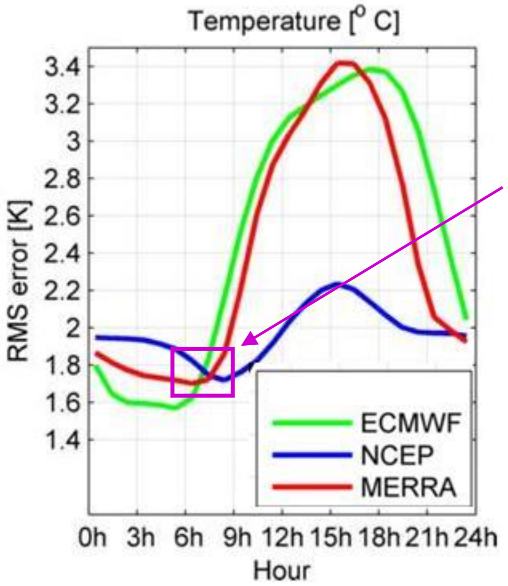


Figure 22. ECMWF forecast surface temperatures and Oklahoma MESONET (2 mm) surface temperatures at the overpass time of 06:00 AM local time for 2003 [31].

More recently, T. Holmes et al. [39] compared the accuracy of 0-5 cm soil temperature derived from the three NWP centers (MERRA is a GMAO data set) against *in situ* soil temperature data from the Oklahoma Mesonet for years 2004 and 2009 [39]. Figure 23 illustrates that at an overpass time of ~ 6 am, all synchronized NWP-derived surface temperature products have errors below 2 K, which is the amount of error budget allocation that is nominally carried for the surface temperature ancillary data parameter.



SMAP 6 am descending orbit

Figure 23. Accuracy of NWP forecast surface soil temperature compared against in situ temperatures for the Oklahoma Mesonet for 2004 and 2009.

6.2.1 Effective Soil Temperature

Postlaunch, dynamic surface temperature forecast information is routinely ingested by SMAP from the GMAO GEOS-5 model and processed as an ancillary data input as part of the operational processing of the SMAP passive soil moisture product. The original baseline computation of the effective surface temperature (T_{eff}) consisted of using the average of the GMAO surface temperature (TSURF) and the GMAO layer 1 soil temperature at 10 cm (TSOIL1). Preliminary analyses showed that a more sophisticated model for computing T_{eff} was required due to non-uniform soil temperature profiles, especially in arid areas, which led to soil moisture retrieval issues. In order to address this problem, several options for T_{eff} were considered and evaluated using SMAP T_B observations along with GMAO soil temperatures for the soil profile.

The SMAP L2_SM_P product currently uses the Choudhury [58] model to compute the effective soil temperature:

$$T_{eff} = T_{soil\_deep} + C\,(T_{soil\_top} - T_{soil\_deep}) \qquad (3)$$

where T_{soil_top} refers to the GMAO layer 1 soil temperature at 0-10 cm (TSOIL1) and T_{soil_deep} refers to the GMAO layer 2 soil temperature at 10-20 cm (TSOIL2). This formulation allows for correct modeling of the deeper sensing depth of emission emanating from deeper in the soil than the surface. C is a coefficient that depends on the observing frequency – for the SMAP L-band data releases Versions 2-4, C= 0.246 as given in [58].

This approach to the calculation of T_{eff} was then applied to all regions in SMAP L2_SM_P soil moisture retrievals, and did minimize the number of non-retrievals due to soil temperature issues.

6.3 Vegetation Water Content

As described in previous sections, a number of retrieval algorithms under investigation rely on vegetation water content (VWC) as an input ancillary parameter. Accurate temporal estimates of VWC, especially at high spatial resolution on a global basis, are very important to achieving accurate soil moisture retrieval using SMAP algorithms. Since VWC is not a parameter that can be measured directly by existing remote sensing techniques, it must be indirectly inferred from other measurable parameters with which it has high correlation. One such parameter is the Normalized Difference Vegetation Index (NDVI).

As described in the SMAP Ancillary Data Report for Global Vegetation Water Content [40], the SMAP team has been collaborating in the development of a more robust and reliable method for estimating VWC from NDVI, taking land cover variability into consideration. This new approach leverages the existing NDVI-based methodology to estimate the foliage water content (leafy part of the vegetation canopy), while using a combination of past field observations and Leaf Area Index (LAI) modeled by NDVI to account for the stem water content (stem and branch part of the vegetation canopy). The result is an estimate of VWC with water content contributions from the foliage and stem components, adjusted for the land cover types in the MODIS IGBP classification scheme.

While the foliage component is expressed in terms of NDVI, the stem component is expressed in terms of Leaf Area Index (LAI), along with annual maximum and minimum NDVI. As LAI exhibits distinct dynamics for different land cover types, this approach makes it possible to use NDVI and land cover classification data sets to construct a global VWC database at high spatial resolution. For croplands and grasslands the current NDVI is used for $NDVI_{ref}$; for all other vegetation types, the annual maximum NDVI is used for $NDVI_{ref}$:

$$VWC = (1.9134 \text{ x } NDVI^2 - 0.3215 \text{ x } NDVI) + StemFactor \text{ x } (NDVI_{ref} - 0.1) / (1 - 0.1) \tag{19}$$

where the stem factor is the product of the average height of a land cover class and the ratio of sapwood area to leaf area. Sapwood area to leaf area ratio [55] is based on the physical requirements for water transport from the soil through the xylem and into the leaves in order to replace water lost by transpiration. When the stem factor is multiplied by leaf area index (derived from canopy water content), the result is the approximate volume of water in the actively-conducting stem xylem per unit ground area (see the last column of Table 3). An example of the VWC distribution using the above formulation over the US for July is shown in Figure 24.

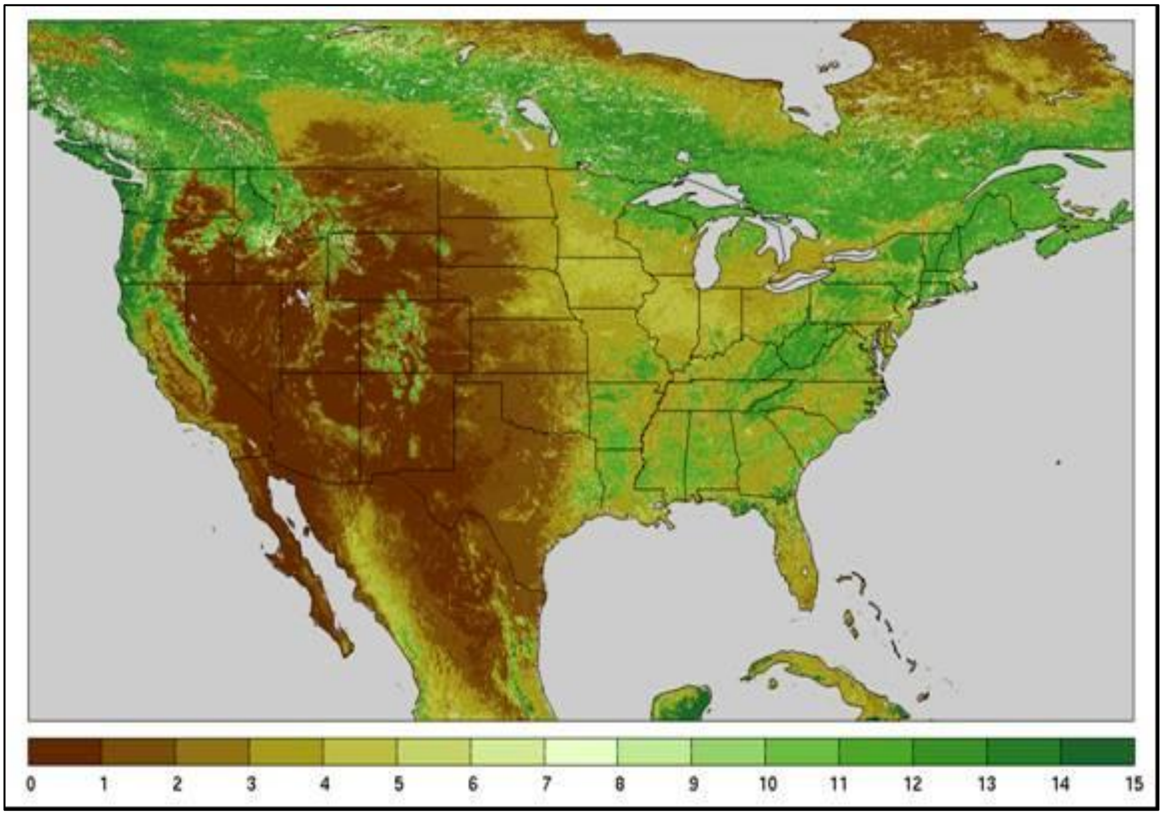


Figure 24. VWC over the continental U.S. for the month of July on a 1-km EASE grid as constructed from a 10-year MODIS NDVI climatology and land cover products.

This new methodology for determining VWC was used in creating a new global 10-year MODIS NDVI climatology at 1 km spatial resolution for use by SMAP [41]. The new climatology was derived from MODIS data from 2000-2010, and is binned over 10-day periods throughout the year; prior to the SMAP launch, this climatology will be updated for the period 2000-2013. In the absence of concurrent NDVI data during the SMAP mission, the historical NDVI for any day of the year for any location can be determined and then used in the VWC calculation described in Equation 17; the annual maximum NDVI is also readily obtained. Figure 25 illustrates the new NDVI climatology for the USDA watershed at Walnut Creek, IA (interpolated where snow is present). Calculation of VWC also serves to set the dense vegetation flag, where the calculated VWC > 5 kg/m^2 for the given grid cell.

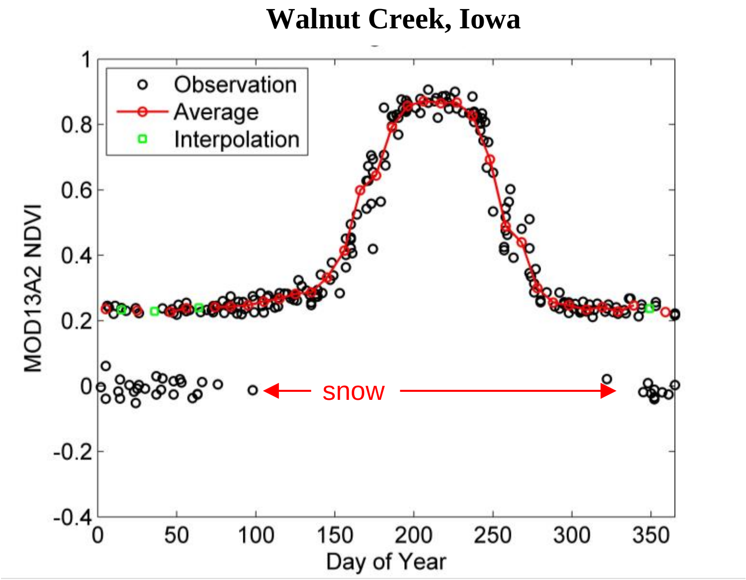


Figure 25. Annual climatology of NDVI for Walnut Creek, IA derived from 2000-2010 MODIS data.

6.4 Soil Texture

Soil moisture retrieval algorithms require information about soil texture, specifically sand and clay fraction. A global dataset was assembled from an optimized combination of the FAO (Food & Agriculture Organization), HWSD (Harmonized World Soil Database), STATSGO (State Soil Geographic—US), NSDC (National Soil Database Canada), and ASRIS (Australian Soil Resources Information System) soil databases (Figure 26). This composite dataset uses the best available source for a given region [54], which should improve the accuracy of SMAP products in that region as well as providing consistency with the work of local scientists and end users in that region. A negative consequence of this decision is the potential for discontinuities at international boundaries, such as between the United States and Canada.

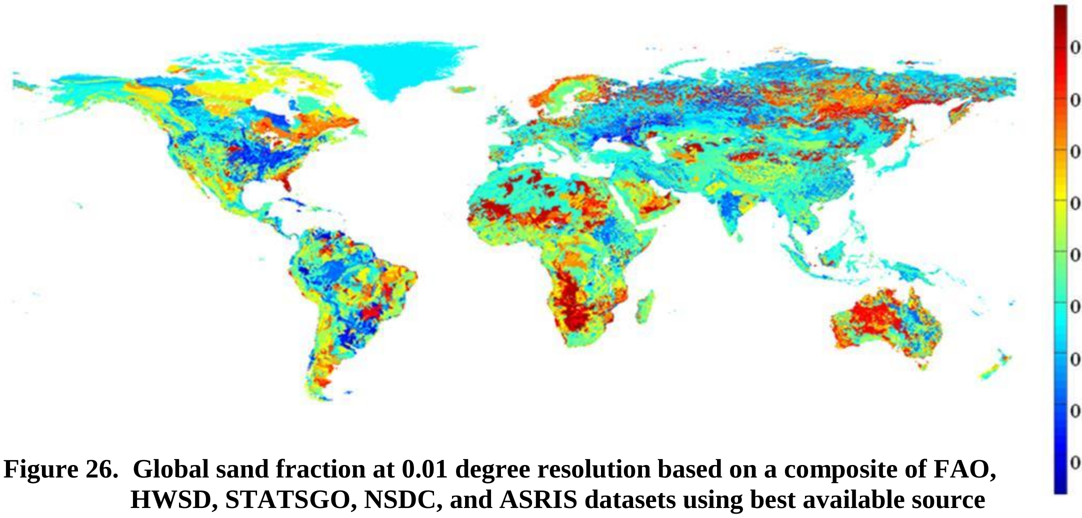


Figure 26. Global sand fraction at 0.01 degree resolution based on a composite of FAO, HWSD, STATSGO, NSDC, and ASRIS datasets using best available source for a given region. All ancillary data will be resampled to the appropriate SMAP EASE grid.

6.5 Data Flags

Ancillary data will sometimes also be employed to set which help to determine either specific aspects of the processing (such as corrections for transient water) or the quality of the retrievals (e.g. precipitation flag). Basically, these flags will provide information as to whether the ground is frozen, snow-covered, or flooded, or whether it is actively precipitating at the time of the satellite overpass. Other flags will indicate whether masks for steeply sloped topography, or for urban, heavily forested, or permanent snow/ice areas are in effect. All flag threshold values are currently under review and may be modified prior to launch and during post-launch calibration/validation activities. All TBDs in flag thresholds will be resolved in final software releases prior to launch.

6.5.1 Open Water Flag

The open water fraction will be produced by the SMAP HiRes radar, *a prioi* information on permanent open fresh water from the MOD44W database, or a combination of the two. It is always reported as part of the L2_SM_P output product (see Appendix 1). This information serves as a flag to affect soil moisture retrieval processing in the following way:

- If water fraction is 0.00–0.05, then flag for recommended quality and retrieve soil moisture
- If water fraction is 0.05–0.50, then flag for uncertain quality and attempt to retrieve soil moisture
- If water fraction is 0.50–1.00, then flag but do not retrieve soil moisture

6.5.2 RFI Flag

The presence of radio frequency interference can markedly impact SMAP T_B, and in turn can adversely affect soil moisture retrieval accuracy or prevent a retrieval from being attempted. RFI is detected and corrected for in the L1B_TB (SPL1BTB) product, which sets an RFI flag which is eventually passed to the L2_SM_P processor (see L1B_TB ATBD). The RFI flag affects soil moisture retrieval processing in the following way:

- If RFI is not detected (bit 2=0 in SPL1BTB's tb_qual_flag), then flag for recommended quality and retrieve soil moisture
- If RFI is detected and can be corrected successfully (bit 3=0 in SPL1BTB's tb_qual_flag), then flag for recommended quality and retrieve soil moisture
- If RFI is detected but can only be corrected partially (bit 14=1 in SPL1BTB's tb_qual_flag), then flag for uncertain quality and attempt to retrieve soil moisture
- If RFI is detected and cannot be corrected (bit 3=1 in in SPL1BTB's tb_qual_flag), then flag but do not retrieve soil moisture.

Note that the RFI information is not embedded in SPL2SMP's surface condition flag. It is mentioned here because it, along with the status of other surface conditions, helps to determine the quality of soil moisture retrieval.

6.5.3 Snow Flag

Although the SMAP L band radiometer can theoretically see through dry snow with its low dielectric to the soil underneath a snowpack, the snow flag is currently envisioned as an area snow fraction based on the NOAA IMS database. The snow flag affects soil moisture retrieval processing in the following way:

- If snow fraction is 0.00–0.05, then flag for recommended quality and retrieve soil moisture
- If snow fraction is 0.05–0.50, then flag for uncertain quality and attempt to retrieve soil moisture
- If snow fraction is 0.50–1.00, then flag but do not retrieve soil moisture

Permanent snow/ice fraction as indicated in the SMAP ancillary land cover map is also treated similarly to snow fraction with the same lower and upper permanent snow/ice thresholds.

6.5.4 Frozen Soil Flag

The SMAP frozen soil flag is set during internal SDS processing based on the radar ground flag (see L3_FT_A ATBD) or on the GMAO T_{eff}. For beta release, the frozen soil area fraction is based on temperature information from the GEOS-5 model. Since the frozen soil flag is generated at high spatial resolution compared to the 36-km grid cell spacing of the L2_SM_P products, a frozen fraction is generated which affects soil moisture retrieval processing in the following way:

- If frozen soil fraction is 0.00–0.05, then flag for recommended quality and retrieve soil moisture

- If frozen soil fraction is 0.05–0.50, then flag for uncertain quality and attempt to retrieve soil moisture
- If frozen soil fraction is 0.50–1.00, then flag but do not retrieve soil moisture

6.5.5 Precipitation Flag

The SMAP precipitation flag is currently set based on forecasts of precipitation from the GEOS-5 model. The use of observational data from the Global Precipitation Mission (GPM) will be evaluated during the intensive cal/val period. The precipitation flag gives the rain rate in mm/hr (or kg/m^2/s), indicating the presence or absence of precipitation in the 36-km grid cell at the time of the SMAP overpass. The presence of liquid in precipitation incident on the ground at the time of the SMAP overpass can adversely bias the retrieved soil moisture due to its large impact on SMAP T_B (precipitation in the atmosphere is part of the atmospheric correction done in SPL1BTB processing). Unlike with other flags, soil moisture retrieval will always be attempted even if precipitation is flagged. However, this flag serves as a warning to the user to view the retrieved soil moisture with some skepticism if precipitation is present.

- If precipitation rate is 0.0–1.0 mm/hr, then flag for recommended quality and retrieve soil moisture
- If precipitation rate is 1.0–25.4 mm/hr, then flag for uncertain quality and attempt to retrieve soil moisture
- If precipitation rate is > 25.4 mm/hr, then flag but do not retrieve soil moisture.

6.5.6 Urban Area Flag

Since the brightness temperature of manmade, impervious, and urban areas cannot be estimated theoretically, the presence of urban areas in the 36-km L2_SM_P grid cell cannot be corrected for during soil moisture retrieval. Thus, the presence of even a small amount of urban area in the radiometer footprint is likely to adversely bias the retrieved soil moisture. The SMAP urban flag will be set based on Columbia University's GRUMP data set [42]. The urban fraction affects soil moisture retrieval processing in the following way:

- If urban fraction is 0.00–0.25, then flag for recommended quality and retrieve soil moisture
- If urban fraction is 0.25–1.00, then flag for uncertain quality and attempt to retrieve soil moisture

6.5.7 Mountainous Area Flag

Large and highly variable slopes present in the radiometer footprint will adversely affect the retrieved soil moisture. The SMAP mountainous area flag will be derived from a statistical threshold based on the slope standard deviation (SD) within each 36-km grid cell. Most likely, soil moisture retrieval will still be attempted in most areas flagged as mountainous.

- If slope standard deviation is 0-3°, then flag for recommended quality and retrieve soil moisture
- If slope standard deviation is 3-6°, then flag for uncertain quality and attempt to retrieve soil moisture
- If slope standard deviation is > 6°, then flag but do not retrieve soil moisture

6.5.8 Proximity to Water Body Flag

For any given instantaneous measurement, the SMAP radiometer receives a portion of its energy from outside its 3 dB footprint. This becomes an issue if a large water body is just outside the boundaries of a 36-km EASE Grid 2.0 cell but still contributes to the observed signal, since the microwave brightness temperature of standing water is significantly cooler than that of land and would adversely bias the soil moisture retrieved inside the 36-km cell. The proximity to nearby water body flag affects soil moisture retrieval processing in the following way:

- If distance to nearby water body > one 36-km grid cell, then flag for recommended quality and retrieve soil moisture
- If distance to nearby water body < one 36-km grid cell, then flag for uncertain quality and attempt to retrieve soil moisture

6.5.9 Dense Vegetation Flag

The presence of dense vegetation in the grid cell negatively affects the accuracy of retrieved soil moisture. However, at the request of the science community, a soil moisture retrieval will always be attempted regardless of the amount of vegetation present. The dense vegetation flag affects soil moisture retrieval processing in the following way:

- If vegetation water content is 0-5 kg/m^2, then flag for recommended quality and retrieve soil moisture
- If vegetation water content is 5-30 kg/m^2, then flag for uncertain quality and attempt to retrieve soil moisture
- If vegetation water content is > 30 kg/m^2, then flag but do not retrieve soil moisture

6.6 Latency

The SMAP mission requirements impose latency requirements on all SMAP products. L2_SM_P data products have a latency requirement of 24 hours and the L3_SM_P products have a latency of 48-50 hours (to allow for the accumulation of 24 hours of half orbits and their subsequent processing). In operational processing, the SDS is thus responsible for generating the L2_SM_P products within the stated periods from the moment of satellite data acquisition to delivery to the SMAP NSIDC DAAC for distribution to the public.

To meet these requirements, the external ancillary datasets that will be used in L2/3_SM_P processing must be available within the stated periods. The major NWP forecast centers have indicated that most of the needed ancillary data parameters which are highly dynamic and time critical (e.g., surface temperature) will be available to the SMAP SDS for routine product generation within 6 hours of the forecast.

7. CALIBRATION AND VALIDATION

7.1 Algorithm Selection

As discussed in section 4.6, the selection of the algorithm to be used operationally to produce the standard SMAP L2_SM_P surface soil moisture product will be made just prior to the first SMAP bulk reprocessing and continually assessed throughout the mission. Performance evaluations pre- and post-launch will include:

- comparisons of soil moisture estimates using SMOS T_B data processed to the SMAP configuration with *in situ* soil moisture data sets and SMOS algorithm retrievals,
- comparisons of soil moisture estimates based on tower and aircraft field campaign data with ground-based observations,
- sensitivity and uncertainty analyses based upon GloSim3
- comparisons of SMAP retrievals with *in situ* data from CV sites.

L2_SM_P products will satisfy the mission requirement that the retrieved soil moisture will have an ubRMSE (unbiased RMSE) of no more than 0.04 cm^3/cm^3 over areas where the vegetation water content ≤ 5 kg/m^2; this target accuracy was confirmed as achievable in a previous study (Figure 27) for the *Hydros* mission using three candidate algorithms [7]. The *Hydros* study showed that when retrievals were aggregated at the basin level (575,000 km^2), all three algorithms met the target accuracy of 0.04 cm^3/cm^3 volumetric soil moisture, although for individual pixels with high vegetation water content and/or high surface heterogeneity, the soil moisture retrieval accuracy degraded [note: the *Hydros* reflectivity ratio algorithm is not currently a candidate SMAP algorithm]. These general results also apply to SMAP.

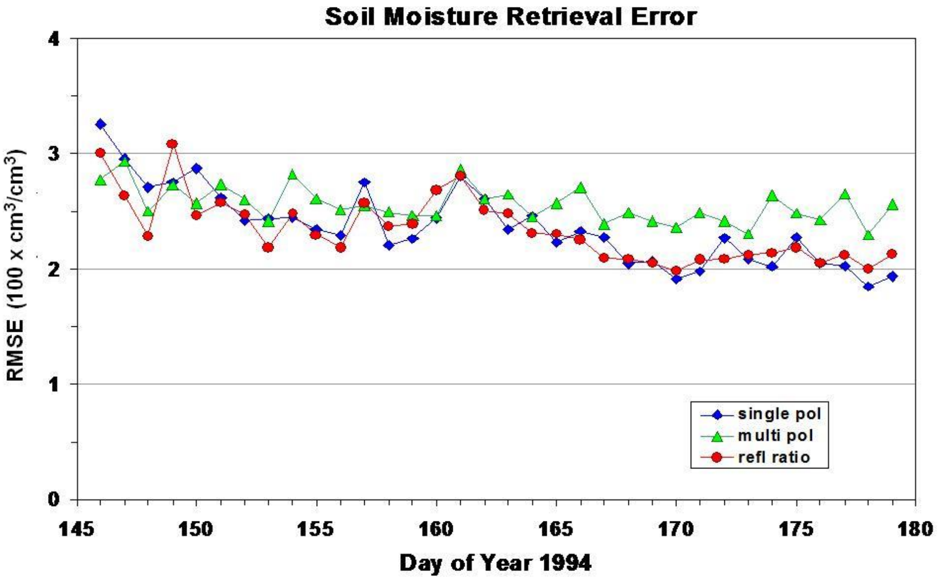


Figure 27. Performance comparison among three candidate retrieval algorithms for the *Hydros* mission based on an OSSE over the Arkansas-Red River basin [7].

7.1.1 SMOS and Aquarius Data Products

The SMAP L2_SM_P team is in a unique position to assess the relative merits of alternative algorithms because data from two currently operating satellites, SMOS and Aquarius, can be used as surrogates for SMAP. SMOS is currently providing L-band brightness temperature as well as a retrieved soil moisture product (since November 2009) [43]. Aquarius began providing brightness temperature data in August, 2011 (launched June 2011) [44] and has also released a soil moisture product [57]. The brightness temperatures from each of these missions require reprocessing in order to simulate the constant 40° incidence angle observations that SMAP will provide.

Initially, the plan was to use the SMOS global gridded L1C browse brightness temperature product as a SMAP surrogate with minimal reprocessing. This SMOS product consists of swath-based dual/full polarization observations resampled to an Earth-fixed grid with a standard incidence angle of 42.5° at the nominal spatial resolution of SMAP. This product provides antenna reference brightness temperatures, but the required parameters for performing the rotation to true surface polarized T_B (including Faraday angle) are not available in the browse product. Upon further consideration of the differences (e.g., grid, incidence angle, etc.) between the two missions and an evaluation of the radiometric quality of the browse product, it was decided that it would be necessary to reprocess the SMOS data using the standard L1C product. Only the unaliased FOV portions of the SMOS orbit are used in the processing. The procedures adopted will result in a higher quality brightness temperature data set at a constant incidence angle of 40° matched to the SMAP grid. Although this product will have a reduced swath width as a consequence of the reprocessing, the loss of some swath width is not critical to the algorithm performance assessment objectives of this analysis.

The following files are acquired for each swath in order to conduct the subsequent analyses: SMOS L2 soil moisture DAP (Data Analysis Product), SMOS L2 soil moisture UDP (User Analysis Product), ECMWF forecast files, and SMOS tau vegetation parameter files for forest and non-forest areas (τ). The first stage of SMOS analysis is generating the constant 40 degree incidence angle brightness temperature data from the SMOS L1C T_B product. This involves the following steps:

- Removing the extended FOV portions of the SMOS orbit
- Filtering to remove anomalous T_B observations and RFI check
- Interpolation to fill in full/dual-pol T_B observations for each snapshot (needed for the next step)
- Transforming from antenna to Earth reference frame (Computing X-Y to H-V T_B)
- RFI check ($0<T_B<320$, $T_{BH}<T_{BV}$)
- For each grid point, brightness temperatures at all available incidence angles are used to develop a prediction equation for T_B as a function of angle, and values at 40 degrees are then predicted.

The next stage of analysis is the retrieval of soil moisture with alternative SMAP algorithms using the SMOS products in their original grid system. All ancillary data are derived from SMOS files. These retrievals will be compared to the SMOS soil moisture products, ground-based soil moisture, and possible model-based products. The first results of this effort were presented in section 4.6.1.

Finally, the new 40 degree SMOS product will be transferred to the SMAP grid. The alternative SMAP algorithms will then be applied with the SMAP ancillary data sets. As described above, evaluation will utilize ground-based observations from well-studied validations sites, the SMOS soil moisture product (re-gridded), and model-based products. Ground-based data sets will include all site data provided by SMAP cal/val partners that meet the product scale and minimum requirements. The validation site analysis will compute standard statistical parameters (RMSE, bias, correlation) for each algorithm at each validation site over at least one annual cycle. Each site will be evaluated individually and then treated as a group to provide a ranking of the algorithms. This process was evaluated in a rehearsal campaign.

In addition to the analysis above that focuses on a number of well-characterized sites, evaluations that incorporate a global synoptic perspective will be conducted. The entire global SMOS soil moisture data set will be compared to the products provided by the candidate SMAP soil moisture algorithms. The performance of each algorithm, relative to SMOS, will be evaluated on an overall basis as well as for categories that include land cover types, NDVI levels, and continents. Statistics will include RMSD, ubRMSD, bias, and correlation. This analysis assumes that the SMOS soil moisture product is accurate and reliable. While it is expected that the SMOS team is doing its best to achieve this goal, it is possible that there may be regions and land covers where the SMOS results are less reliable. It is also possible that some data quality issues may remain, especially the issue of SMOS aliasing, which are likely to be more significant in the algorithms that utilize more than a single polarization. Methods to possibly mitigate this aliasing are currently being investigated.

Aquarius brightness temperature and radar data became available beginning in August, 2011. Aquarius provides L-band data for three beam positions with incidence angles of 28.7, 37.8, and 45.6 degrees. Although methods for normalizing incidence angle will be explored, initially only the middle beam position data will be utilized. It is expected that the radiometric calibration and quality of the Aquarius data will be high, based upon the necessity for high radiometric quality data in order to achieve the mission objectives of measuring sea surface salinity (which has a small dynamic range of T_B). Since its initial data release, there have been two reprocessings by Aquarius as well as SMOS changes that have resulted in Aquarius and SMOS brightness temperatures being closer to each other over the same target. A drawback to the Aquarius data is its spatial resolution, which is several times coarser than SMAP. This coarse resolution (>100 km) might reduce the range of estimated soil moisture and increase the impact of surface heterogeneity. In addition, almost all of the global ground-based soil moisture validation sites have been developed to support products with a spatial resolution of 25-50 km. The coarser scale of Aquarius relative to the ground-based data will have to be carefully considered in the algorithm assessments. The Aquarius program supports the generation of a soil moisture product [57]. Analyses to date indicate that the SMOS and Aquarius products are similar on a global basis.

7.1.2 Tower and Aircraft Field Experiment Data Sets

Because of the natural heterogeneity of landscapes and the inherent coarse scale of satellite radiometers, it can prove challenging to identify the causes of algorithm errors when using satellite-based sensors. Tower and aircraft-based sensors have higher resolutions that allow the control of variability introduced by land cover and soils. Therefore, these instrument platforms can provide additional and valuable insights that are relevant to the algorithm selection decision [56].

Several recent field experiments have provided L-band dual polarization datasets that can be used to evaluate algorithm performance under real-world conditions. These datasets have been compiled and archived for SMAP investigations. These datasets include aircraft observations of SMEX02, CLASIC, SMAPVEX08, CanEx-SM10, SMAPVEX12, and tower-based observations from ComRAD and other instruments. When closely examined, these experiments only cover a limited set of conditions as a result of either design or meteorological circumstances. Therefore, it would be highly desirable for the algorithm selection process to acquire additional data sets. Tower (ComRAD APEX12) and aircraft (SMAPVEX12) experiments were successfully conducted in the Spring/Summer of 2012; data are currently being analyzed. Additional ComRAD data may be collected to aid in radar data cube and radiometer algorithm refinements.

In addition, several other field experiments outside of North America may prove valuable to the algorithm selection process. The airborne soil moisture field campaigns are listed in Table 7.

Table 7. Airborne Soil Moisture Field Campaigns

Campaign	Location	Description
Washita'92	Oklahoma	The first campaign to attempt to collect a time series of spatially distributed hydrologic data, focusing on soil moisture and evaporative fluxes, using both conventional and remotely sensed methods. A NASA C-130 supported the ESTAR L band microwave radiometer and a DC-8 carried Airsar. One of the most successful and scientifically valuable campaigns ever conducted as a result of meteorological conditions and aircraft/instrument performance.
Washita'94	Oklahoma	The primary objective of this experiment was to provide combined ground and aircraft remotely sensed data sets in conjunction with the Space Shuttle Imaging radar missions (SIR-C) in 1994. Each SIR-C mission was to consist of one week of daily observations for the watershed site. ESTAR and AirSAR collected data during portions of the campaign.
SGP97	Oklahoma	SGP97 was a broad multi-disciplinary experiment. One of its main objectives was to establish that the retrieval algorithms for surface soil moisture developed at higher spatial resolution using truck- and aircraft-based sensors can be extended to the coarser resolutions

		expected from satellite platforms. It included the L-band Electronically Scanned Thinned Array Radiometer (ESTAR) and a tower-based L and S-band system. The campaign spanned a longer time period (4 weeks) and covered a domain an order of magnitude larger than prior experiments.
SGP99	Oklahoma	SGP99 returned to the same study region as SGP97 with a new suite of aircraft-based sensors that included AMSR simulators and the recently developed L- and S-band PALS instrument. PALS was flown over the Little Washita Watershed on 5 days over a 6 day period.
SMEX02	Iowa	SMEX02 expanded previous aircraft-based experiments to higher biomass agricultural conditions (corn and soybeans). Both PALS (7 flights over two weeks) and AirSAR (5 flights over 9 days) data were collected.
SMEX03	Georgia, Alabama, Oklahoma	SMEX03 was designed in expand the diversity of land cover conditions that had been examined in previous campaigns. The experiment included the first application of the L-band 2D-STAR instrument and Airsar coverage at one site (Oklahoma).
SMEX04	Arizona, Mexico	SMEX04 continued the expansion of experimental sites conditions that had been examined in previous campaigns. The experiment included the first application of the L-band 2D-STAR instrument and Airsar coverage at one site (Oklahoma).
CLASIC07	Oklahoma	CLASIC included the first flights with new antenna for PALS. Eleven flights were conducted over four weeks for two watersheds.
SMAPVEX08	Maryland	SMAPVEX08First field campaigns dedicated to resolving SMAP algorithm issues. Agricultural sites and PALS were the focus. In addition, the campaign addressed questions related to RFI.
CanEx-SM10	Saskatoon, Canada	CanEx-SM10 was a collaboration between NASA and CSA over agricultural and forest sites. NASA flew the airborne UAVSAR instrument in conjunction the Canadian L-band airborne radiometer and ground sampling observation over one of the SMAP Core Validation Sites.
SMAPEx10-11	New South Wales, Australia	Collaboration led by the University of Melbourne and Monash University in Australia. Three week-long campaigns in 2010 and 2011 designed to specifically address SMAP soil moisture algorithm issues. Two post-launch campaigns will be conducted in 2015. The campaigns will include coincidental aircraft-based radiometer and radar measurements and ground observation over one of the SMAP Core Validation Sites
San Joaquin Valley Experiment	California	SJV involves the UAVSAR instrument deployed t on several days in 2010-2011. Sites are irrigated orchards and vineyards. The primary objective of the experiment is to develop Vegetation Water Content (VWC) retrieval from optical remote sensing instruments, supported by optical instruments. Soil moisture and backscatter relationships will be evaluated. A series of ten flights over 5 months is planned.
SMAPVEX12	Manitoba, Canada (MB)	SMAPVEX12 was conducted in summer 2012 to address the remaining algorithm issues before the launch. This experiment was a collaborative effort between NASA and CSA. The primary L-band observations were collected by the PALS instrument and UAVSAR. A large spatial domain (including agriculture and natural vegetation) over a six-week period was measured.
SMAPEx15 SMAPVEX15 SMAPVEX16	Australia Arizona IA, MB	Post-launch validation of SMAP.

7.1.3 Simulations Using the SMAP Algorithm Development Testbed

As mentioned in sections 4.5 and 5, the SMAP Algorithm Development Team is developing Fortran codes at JPL that enable a set of closed-loop, end-to-end global simulation runs known as GloSim [45]. These simulations will serve several purposes, including providing a mechanism for intercomparison of the relative merits of the four candidate L2_SM_P retrieval algorithms. Additional simulations are now being run with GloSim to examine the performance of all candidate algorithms, and updated error budgets will be generated for each algorithm. These products were incorporated into the 2014 rehearsal campaign.

7.2 Validation

Post-launch validation must provide the information necessary to address whether or not SMAP has achieved its mission requirement to produce an estimate of soil moisture in the 0-5 cm layer with an average ubRMSE of no more than 0.04 cm^3/cm^3 over areas where the vegetation water content ≤ 5 kg/m^2, excluding regions of frozen soil, permanent snow / ice, mountainous terrain, and open water at the footprint measurement scale (40 km for the L2_SM_P). It has been suggested by CEOS (http://lpvs.gsfc.nasa.gov/) that full validation of a satellite product can require a substantial effort in space and time data collection, and that a reasonable approach to the problem is to consider validation as consisting of several stages, which are summarized in Table 8.

Table 8. A Hierarchical Approach to Classifying Land Product Validation Stages as Adopted by CEOS through Consensus of the Land Product Validation Community in 2003 and Revised in 2009

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.

Initial post-launch validation of the L2_SM_P product has been performed and is documented in the SMAP L2_SM_P beta and validated release assessment reports [59, 60] which are available through NSIDC; an L2_SM_P Data Release Version 4 assessment report [61] will be produced in December, 2016. Validation Stages 1 and 2 and beyond will be completed by the end of the official calibration/validation phase of the SMAP mission in 2016 (12 months after the beginning of routine science operations on orbit (IOC)). All validation data and results will be provided through the NSIDC. Given that an initial journal article presenting an assessment of the L2_SM_P soil moisture product was published in August, 2016 [62], it is realistic that within two years after the end of IOC, Stages 1 through 3 will be complete, and work toward Stage 4 will be well underway.

The SMAP Cal/Val plan [46] describes five types of resources that will be utilized as sources of calibration/validation data. These types of data are listed in Table 9, which describes their perceived role and issues that are currently being addressed by the SMAP team. The NSPIRES DCL entry in the table refers to a Dear Colleague Letter request for information that was released by NASA to the science community to solicit members of the SMAP cal/val team and core validation and other validation sites globally. A NASA panel in consultation with the SMAP team selected ~27 investigator sites or supported instrument networks in Summer, 2011. Since then several new sites have been added.

Table 9. Overview of the SMAP Cal/Val Methodologies

Methodology	Role	Issues	Actions
Core Validation Sites	Accurate estimates of products at matching scales for a limited set of conditions	Calibration Up-scaling Limited number	*In Situ* Testbed Scaling methods NSPIRES DCL
Sparse Networks	One point in the grid cell for a wide range of conditions	Calibration Up-scaling Limited number	*In Situ* Testbed Scaling methods NSPIRES DCL
Satellite Products	Estimates over a very wide range of conditions at matching scales	Validation Comparability Continuity	Validation Studies CDF Matching
Model Products	Estimates over a very wide range of conditions at matching scales	Validation Comparability	Validation Studies
Field Experiments	Detailed estimates for a very limited set of conditions	Resources Schedule Conflicts	Simulators Partnerships

The baseline validation for the L2_SM_P soil moisture will be 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. This matches up closely with the Stage 1 validation described in Table 8. Data from core validation sites will be supplemented by field experiments. In order to achieve Stage 2 validation and include a wider range of conditions as well as a synoptic/global assessment, some combination of data from sparse networks, other satellite products, and model-based estimates must be utilized. Each of these data types has caveats associated with it that are described in Table 9. The following sections provide some additional details on how each of the resources listed in Table 10 will be utilized specifically for the L2_SM_P soil moisture product validation.

Table 10. SMAP Cal/Val Methodologies and Their Roles in the L2_SM_P Soil Moisture Product Validation

Methodology	Data Required	Importance	Metrics
Core Validation Sites	Grid Cell averages for each overpass	Primary	RMSE, ubRMSE, Bias, Correlation
Sparse Networks	Spatially scaled grid cell values for each overpass	Secondary: Pending results of scaling analyses	RMSE, ubRMSE, Bias, Correlation
Satellite Products	Orbit-based match-ups Key targets	Primary: Pending assessments and continued operation	RMSD, ubRMSD, Bias, Correlation
Model Products	Orbit-based match-ups Key targets	Secondary	RMSD, ubRMSD, Bias, Correlation
Field Experiments	Detailed estimates for a very limited set of conditions	Primary	RMSE, ubRMSE, Bias, Correlation

7.2.1 Core Validation Sites

As noted previously, the baseline validation (Stage 1) for the L2_SM_P soil moisture will be 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. The CVS have been selected because they satisfied several criteria that included:

- A network of sensors with adequate replication
- For soil moisture, ideally, three nested levels of extent (3, 9, and 36 km)
- For soil moisture, verified against gravimetric samples for the 0-5 cm layer
- Minimal latency in providing data to the SMAP project
- Fully operational well before launch, with infrastructure to support the site through at least 2016

NASA has established agreements with the CVS teams that require the teams to provide the ground-based data in a timely manner to the SMAP project (or the NASA-designated SMAP DAAC at NSIDC). There are expected to be ~15 of these CVS distributed over the globe, and these may increase in number over the next few years. Many of these sites have been used in AMSR-E and SMOS validation [47-50]. Multiple sample points at each site will be averaged to estimate the footprint-scale soil moisture value that will be compared to the SMAP retrieval. The method of averaging will depend upon the amount of information provided by the CVS team. Some of these sites will also be the focus of intensive ground and aircraft field campaigns to further verify the accuracy of the collected data as well as improving scaling.

Having a global distribution of sites will be beneficial to SMAP validation. Based on the launch date of SMAP (January, 2015), the seasonal variations between the northern and southern hemispheres may impact the usefulness of some regions in validation within the official 12-month cal/val period. With a number of core sites in each hemisphere, the official SMAP validation period is less affected by the seasonality of the launch date. The SMAP project is implementing a special product for validation that consists of L1C_TB data centered over the core validation sites to aid in SMAP validation. This is particularly important to the L2_SM_P product because it will allow the full exploitation of the *in situ* data.

7.2.2 Sparse Networks

The intensive network validation described above can be complemented by sparse networks as well as by new/emerging types of networks. Examples of sparse networks include the USDA Soil Climate Analysis Network (SCAN), the NOAA Climate Research Network (CRN), and the Oklahoma Mesonet. The defining feature of these networks is that the measurement density is low, usually resulting in one point per 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 the footprint. SMAP has been evaluating methodologies for upscaling data from these networks to SMAP footprint resolutions. A key element of the upscaling approach will be a method called Triple Co-location (Section 7.2.6) that combines the *in situ* data and SMAP product with another independent source of soil moisture, likely to be a model-based product.

Beyond these operational networks, there are new technologies being evaluated (COSMOS, GPS) that could provide distributed soil moisture information. SMAP is participating in the evaluation of these new technologies as part of its Marena, Oklahoma *In Situ* Sensor Testbed (MOISST) that is assessing both the verification of the relevant depth of measurement of these methods and scaling to SMAP footprints. The upscaling of these sparse networks remains an issue at present, and until this issue is resolved, the sparse networks will likely remain a secondary validation resource for the SMAP L2_SM_P soil moisture products.

7.2.3 Satellite Products

Depending upon mission timing and life, it is possible that SMOS, Aquarius, and JAXA's GCOM-W will be producing global soil moisture products at the same time as SMAP. SMOS and GCOM-W products are at the same nominal spatial resolution as the SMAP L2_SM_P soil moisture and are supported by validation programs, which should be mature by the time of the SMAP launch. As mentioned earlier, Aquarius soil moisture has a coarser resolution than these other satellites.

In a previous section, the use of SMOS data prior to the launch of SMAP was described. Post-launch soil moisture product comparisons with SMOS and GCOM-W are a very efficient means of validation over a wide range of conditions. If confidence in these products is high, they will provide an ideal resource for Stage 2 SMAP validation. The limitations of this type of comparison are the quality of the alternative product, differences in overpass days, and accounting for system differences affecting the soil moisture product. In the case of GCOM-W, which collects data at 01:30 am and 01:30 pm, confusion factors would include both data acquisition at a different time of day from the SMOS/SMAP overpass time of 06:00 am and contributing depth issues associated with GCOM-W's C-band frequency. The SMAP team will actively participate in the validation of these alternative products during the SMAP pre-launch period, which will provide us with knowledge of the quality of both the SMOS and GCOM-W soil moisture.

Post-launch validation will consist of comparisons between the SMAP / SMOS / GCOM-W / Aquarius soil moisture estimates that include:

- Core validation sites (CVS)
- Extended homogeneous regions
- Global maps

For the core validation sites and extended homogeneous regions, statistical comparisons will be conducted (Root Mean Square Difference, RMSD, will be used instead of RMSE because the alternative satellite products are not considered to be "ground truth"). The CVS will likely consist of approximately 15 sites distributed around the globe as described in the SMAP Cal/Val Plan [46]. Comparisons will be initiated as soon as SMAP soil moisture products become available; however, a sufficient period of record that includes multiple seasons will be necessary before any firm conclusions can be reached. It should also be noted that only dates when both satellites cover the same ground target at the same time will be useful. The overlap of the swaths will vary by satellite. The morning (and evening) orbits of SMAP and SMOS cross (the SMOS 6 am overpass is ascending while the SMAP 6 am overpass is descending). Obviously, coverage of a specific site by both satellites will be infrequent. Aquarius and SMAP will have the same overpass times.

Although data collected over the CVS will be of the greatest value, additional sites with concurrent satellite observations are also useful, especially for regions that are relatively homogeneous in terms of land cover/vegetation and soils. One example would be the Sahara region.

Another role for the satellite products is in providing a synoptic perspective. Global image comparisons will be used to identify regions and/or time periods where the soil moisture products from the different satellites diverge.

Assessments will be conducted periodically throughout the SMAP post-launch period to assess, monitor, and possibly correct bias offsets between SMAP products and SMOS/GCOM-W/Aquarius products. In order to fully exploit SMOS/GCOM-W/Aquarius soil moisture products for SMAP validation, it will be necessary for SMAP team members to participate in the assessment and validation of these products and to secure access to the data through ESA and JAXA.

7.2.4 Model-Based Products

In the simplest case, land surface models (either imbedded in a Numerical Weather Prediction (NWP) system or in off-line mode) can be used to generate soil moisture products at larger (basin-wide and continental) scales using land surface and meteorological forcing data sets that are independent of the SMAP remote sensing data. As in the case of satellite products, the resulting soil moisture fields can then be compared with the remotely sensed soil moisture product at validation sites (or synoptically) over diurnal and seasonal cycles. These model-derived soil moisture fields can also be used to extend comparisons to larger space and time domains than available from *in situ* observations, thus supporting Stage 2 validation. The L2_SM_P product matches the typical spatial resolution of currently available NWP products. An advantage of the model-based products is that they produce a synoptic global product every day, which means that more frequent comparisons to SMAP and ground-based observations are possible.

Several Numerical Weather Prediction (NWP) centers (including ECMWF, NCEP, and NASA/GMAO) routinely produce operational or quasi-operational soil moisture fields at a scale comparable to the SMAP radiometer product that could be used in SMAP validation. [This is distinct from the GMAO generation of the SMAP L4_SM surface and root zone soil moisture product which uses an ensemble Kalman filter (EnKF) to merge SMAP observations with soil moisture estimates from the NASA Catchment land surface model.] The NWP-derived data products rely on the assimilation of a vast number of atmospheric observations (and select land surface observations) into General Circulation Models (GCMs). Although there are many caveats that need to be considered in using these data, they are readily available and they are consistent with the atmospheric forcings (precipitation and radiation) and land use information that determine the spatial and temporal patterns in soil moisture fields.

There is significant inherent uncertainty in any model-based soil moisture product since this is not one of the NWP primary variables. In addition, the models typically simulate a thicker surface soil layer than the layer that dominates the satellite measurement. Little effort has put so far into validating the soil moisture products of these models. Therefore, while these model products are useful, they must be used very carefully. As a result, they are considered to be a secondary resource for validating L2_SM_P soil moisture.

7.2.5 Field Experiments

Post-launch field experiments will play an important role in a robust validation of the L2_SM_P data product. These experiments provide critical information that can be used to independently assess the contributions of radiometer calibration, algorithm structure and parameterization, and scaling on performance. Field experiments require numerous elements that include ground and aircraft resources, which involve many participants and associated financial support. However, they provide moderate-term intensive measurements of soil moisture and other surface characteristics at SMAP footprint scales.

While it is desirable to acquire such information as soon as possible after launch, the uncertainties of the actual launch date, the relationship of the launch date to the season, and the logistics of allocation of fiscal year resources require that such commitments be conservative. Therefore, the field experiments should be scheduled for some time post-launch and used as part of the more robust validation of the SMAP products. Based on a January, 2015 launch, field campaigns are scheduled in 2015 in Australia and Arizona. Additionally, one major extended post-launch field campaign, which will include one or more core validation sites (such as Manitoba, Canada and the U.S. Midwest/Iowa), is scheduled for Summer 2016.

7.2.6 Combining Techniques

Recent work has extended the application of the "Triple Co-location" (TC) approach to soil moisture validation activities [51, 52]. These approaches are based on cross-averaging three independently-acquired estimates of soil moisture to estimate the magnitude of random error in each product. One viable product-triplet is the use of passive-based remote sensing, active-based remote sensing and a model-based soil moisture product [51, 53]. If successfully applied, TC can correct model versus SMAP soil moisture comparisons for the impact of uncertainty in the model product. However, TC cannot provide viable bias information and, therefore, only assesses the random error contribution to total RMSE. Note that TC can also be applied to reduce the impact of sampling error when upscaling sparse *in situ* measurements during validation against ground-based soil moisture observations.

8. MODIFICATIONS TO ATBD

This ATBD will continue to be modified under configuration control as new information becomes available and as the SMAP team refines its decisions regarding algorithm configuration, ancillary data selection, and the setting of flag thresholds.

The following updates are relevant to L2_SM_P Data Release Version 4 and L2_SM_P_E Version 1 in December, 2016 [61]:

8.1 Soil Moisture Retrievals at 6 PM

From the start of routine SMAP science operations on March 31, 2015, the soil moisture retrievals in the standard L2_SM_P product were generated using 6 am brightness temperatures as described in Section 2.4. Starting with SMAP's L2_SM_P Data Release Version 4 in December, 2016, 6 pm soil moistures will also be produced by applying the baseline 6 am retrieval algorithm to T_B data from the 6 pm ascending passes. It is anticipated that the accuracy of the 6 pm soil moisture values will be somewhat worse than the 6 am soil moistures since some of the assumptions underlying the L2_SM_P retrieval algorithm at 6 am (low Faraday rotation, hydraulic and thermal equilibrium between the upper soil layers and the overlying air/vegetation layer, etc.) are more likely to be violated at 6 pm. However, some early results from the SMOS mission suggest that the additional error associated with 6 pm retrievals may not be as large as expected [48]. Although the 6 pm soil moisture retrieval performance will not be included in the evaluation of whether the L2_SM_P product meets the SMAP mission Level 1 requirements, the 6 pm retrievals will be compared against observations of soil moisture to assess their accuracy as is done with the 6 am soil moisture values. These comparisons will be reported in the Data Release Version 4 assessment report available from NSIDC [61].

8.2 Soil Moisture Retrievals using the Enhanced L1C_TB_E Product

After the demise of the SMAP radar in July, 2015, the SMAP Project focused its attention on generating a new brightness temperature data set by using a Backus-Gilbert interpolation approach to take advantage of the SMAP radiometer oversampling on orbit. The resulting brightness temperatures are posted on a 9 km EASE2 grid. Details of this new algorithm approach can be found in the SMAP Algorithm Theoretical Basis Document: Enhanced L1B_TB_E Radiometer Brightness Temperature Data Product, SMAP Project, JPL D-56287, Jet Propulsion Laboratory, Pasadena, CA. The new L1B_TB_E brightness temperatures are used to produce a L1C_TB_E product which is the starting point of the L2_SM_P_E soil moisture retrievals. These retrievals use the same algorithm as the standard L2_SM_P soil moisture retrievals, but have an aggregation domain of 33 km compared to 36 km for the standard product, and are posted at 9 km. The accuracy of the L2_SM_P_E soil moistures are compared to soil moisture observations in the L2_SM_P_E Data Release Version 1 assessment report [61], which also includes a more complete description of the L2_SM_P_E product. The L2_SM_P_E product will include both 6 am and 6 pm retrieved soil moisture posted on a 9 km grid.

9. REFERENCES

1. National Research Council, "Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond," pp. 400, 2007.
2. Entekhabi, D., E, Njoku, P. O'Neill, K. Kellogg, plus 19 others, "The Soil Moisture Active Passive (SMAP) Mission," *Proceedings of the IEEE*, Vol. 98, No. 5, May, 2010.
3. Y. Kerr and J.-P. Wigneron, "Vegetation models and observations –A review," in: *Passive Microwave Remote Sensing of Land-Atmosphere Interactions (B. Choudhury, Y. Kerr, E. Njoku, P. Pampaloni, Eds.),* VSP, Utrecht, 1995.
4. Jackson, T. J., "Measuring surface soil moisture using passive microwave remote sensing," *Hydrol. Process.*, vol. 7, pp. 139-152, 1993.
5. Jackson, T. J., D. E. LeVine, C. T. Swift, T. J. Schmugge, and F. R. Schiebe, "Large area mapping of soil moisture using the ESTAR passive microwave radiometer in Washita'92," *Remote Sens. Environ*, 53:27-37, 1995.
6. Jackson, T. J. and D. E. LeVine, "Mapping surface soil moisture using an aircraft-based passive microwave instrument: Algorithm and example," *J. Hydrol.*, 184:85-99, 1996.
7. Crow, W., T. Chan, D. Entekhabi, P. Houser, T. Jackson, E. Njoku, P. O'Neill, JC Shi, and X. Zhan, "An observing system simulation experiment for SMAP radiometer soil moisture products," *IEEE Trans. Geosci. Rem. Sens.*, 43(6), pp. 1289-1303, 2005.
8. Kerr, Y., P. Waldteufel, J.-P. Wigneron, J.-M. Martinuzzi, J. Font, and M. Berger, "Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission," *IEEE Trans. Geosci. Rem. Sens.*, 39(8), pp. 1729-1735, 2001.
9. Wigneron, J.-P., Y. Kerr, P. Waldteufel, K. Saleh, M.-J. Escorihuela, P. Richaume, P. Ferrazoli, P. de Rosnay, R. Gurney, J.-C. Calvet, J. Grant, M. Guglielmettih, B. Hornbuckle, C. Matzler, T. Pellarin, and M. Schwank, "L-band microwave emission of the biosphere (L-MEB) model: description and calibration against experimental data sets over crop fields," *Remote Sens. Environ*, 107(4), pp. 639-655, 2007.
10. Saleh, K., J.-P. Wigneron, P. Waldteufel, P. de Rosnay, M. Schwank, J.-C. Calvet, and Y. Kerr, "estimates of surface soil moisture under grass covers using L-band radiometry," *Remote Sens. Environ*, 109(1), pp. 42-53, 2007.
11. Wigneron, J.-P., M. Parde, P. Waldteufel, A. Chanzy, Y. Kerr, S. Schmidl, and N. Skou, "Characterizing the Dependence of Vegetation Model Parameters on Crop Structure,

Incidence Angle, and Polarization at L-Band, ” *IEEE Trans. Geosci. Rem. Sens.*, 42(2), pp. 416-425, 2004.

12. Kerr, Y., P. Waldteufel, P. Richaume, I. Davenport, P. Ferrazoli, and J.-P. Wigneron, *SMOS Level 2 Processor Soil Moisture Algorithm Theoretical Basis Document (ATBD)*, Toulouse, France, CESBIO, SM-ESL (CBSA), vol. SO-TN-ESL-SM-GS-0001, v5.a, 2006.

13. Basharinov, A. and A. Shutko, “Simulation studies of the SHF radiation characteristics of soils under moist conditions,” *NASA Tech. Transl.*, TT F-16, 1975.

14. Ulaby, F., R. Moore, and A. Fung, *Microwave Remote Sensing: Vols. I, II, and III*, Addison-Wesley, Reading, MA, 1982.

15. Jackson, T. J. and T. J. Schmugge, “Vegetation effects on the microwave emission from soils,” *Rem. Sens. Environ.*, vol. 36, pp. 203-212, 1991.

16. Choudhury, B. J., T. J. Schmugge, A. Chang, and R. W. Newton, “Effect of surface roughness on the microwave emission from soil,” *J. Geophys. Res*, 84(C9): 5699-5706, 1979.

17. Wang, J. R., “Passive microwave sensing of soil moisture content: the effects of soil bulk density and surface roughness,” *Remote Sens. Environ.*, vol. 13, pp. 329-344, 1983.

18. Wang, J and B. J. Choudhury, “Passive Microwave Radiation From Soil: Examples of Emission Models and Observations,” in Choudhury, B., Y. Kerr, E. Njoku, and P. Pampaloni, *Passive Microwave Remote Sensing of Land-Atmosphere Interactions*, VSP, Utrecht, 1995.

19. Mironov, V. L., L. G. Kosolapova, and S. V. Fomin, “Physically and mineralogically based spectroscopic dielectric model for moist soils,” *IEEE Trans. Geosci. Remote Sens.*, 47(7), pp. 2059–2070, 2009.

20. Dobson, M. C., F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes, “Microwave dielectric behavior of wet soil – Part II: Dielectric mixing models,” *IEEE Trans. Geosci. Rem. Sens.*, vol. GE-23, pp. 35-46, 1985.

21. Wang, J. R and T. J. Schmugge, “An empirical model for the complex dielectric permittivity of soils as a function of water content,” *IEEE Trans. Geosci. Rem. Sens.*, 18, pp. 288-295, 1980.

22. Ulaby, F., P. Dubois, and J. van Zyl, “Radar mapping of surface soil moisture,” *Journal of Hydrology*, 184(1-2), pp. 57-84, 1996.

23. Njoku, E. G. and J. A. Kong (1977): Theory for passive microwave remote sensing of near-surface soil moisture, *J. Geophys. Res., 82*, 3108-3118.

24. Fagerlund, E., B. Kleman, L. Sellin, and H. Svensson, “Physical Studies of Nature by Thermal Mapping,” *Earth-Science Reviews*, 6, pp. 169-180, 1970.

25. Jackson, T. and J. Kimball, “SMAP Mission Science Issues Associated with Overpass Time,” SMAP Science Document No. 003, JPL, March 31, 2009.

26. NSIDC, “EASE-Grid Data,” [Online] Available: http://nsidc.org/data/ease/index.html [Accessed: Jun 14, 2010].

27. NSIDC, “EASE-Grid Data: Data Summaries,” [Online] Available: http://nsidc.org/data/ease/data_summaries.html [Accessed: Jun 14, 2010].

28. SMAP Level 2 Science Requirements (JPL D-45955), Feb 11, 2009. Available online: https://pdms.jpl.nasa.gov.

29. Klein, L. A. and C. T. Swift, "An Improved Model for the Dielectric Constant of Sea Water at Microwave Frequencies," *IEEE Journal of Oceanic Engineering*, vol. 2, no. 1, January, 1977.

30. Ryu, D., T. Jackson, R. Bindlish, D. LeVine, and M. Haken, "Soil Moisture Retrieval using Two-Dimenional L-Band Synthetic Aperture Radiometer in a Semi-Arid Environment," *IEEE Trans. Geosci. Remote Sens.*, 48 (12), pp. 4273-4284, 2010.

31. O'Neill, P., M. Owe, B. Gouweleeuw, E. Njoku, J. Shi, and E. Wood, "Hydros Soil Moisture Retrieval Algorithms: Status and Relevance to Future Missions," *Proc. 2006 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2006)*, Denver, Colorado, July 31-August 4, 2006.

32. Zhan, X., W. Crow, T. J. Jackson, P. O'Neill, "Improving Spaceborne Radiometer Soil Moisture Retrievals With Alternative Aggregation Rules for Ancillary Parameters in Highly Heterogeneous Vegetated Areas," *IEEE Trans. Geosci. Rem. Sens.*, 5(2), pp. 261-265, 2008.

33. Njoku, E. and L. Li, "Retrieval of Land Surface Parameters Using Passive Microwave Measurements at 6-18 GHz," *IEEE Trans. Geosci. Rem. Sens*., vol. 37, pp. 79–93, 1999.

34. Yueh, S., S. Dinardo, S. K. Chan, E. Njoku, T. Jackson, R. Bindlish, "Passive and Active L-Band System and Observations During the 2007 CLASIC Campaign," *Proc. 2008 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2008)*, Boston, Massachusetts, July 6-11, 2008.

35. Owe, M., R. De Jeu, and J. Walker, "A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index," *IEEE Trans. Geosci. Rem. Sens.*, 39, pp. 1643-1654, 2001.

36. de Jeu, R., T. Holmes, R. Panciera, and J. Walker, "Parameterization of the Land Parameter Retrieval Model for L-Band Observations Using the NAFE'05 Data Set," *IEEE Geoscience and Remote Sensing Letters*, 6 (4), pp. 630-634, October, 2009.

37. Mo, T., B. J. Choudhury, T. J. Schmugge, J. R. Wang, and T. J. Jackson, "A model for microwave emission from vegetation-covered fields," *J. Geophys. Res.*, 87(13), pp. 11229-11237, 1982.

38. Meesters, G. C., R. de Jeu, R. and M. Owe, "Analytical derivation of the vegetation optical depth from the microwave polarization difference index," *IEEE Geosci. and Remote Sensing Letters*, 2, pp. 121-123, 2005.

39. Holmes, T., T. Jackson, R. Reichle, and J. Basara, An Assessment of Surface Soil Temperature Products from Numerical Weather Prediction Models using Ground-based Measurements," *Water Resources Research*, in review, 2011.

40. Chan, S, R. Hunt, R. Bindlish, E. Njoku, J. Kimball, and T. Jackson, "Ancillary Data Report for Vegetation Water Content," *SMAP Project Document #* D-53061, JPL, July, 2011.

41. Jackson, T., R. Bindlish, and T. Zhao, "Justification Memo for Vegetation Index Climatology," *SMAP Science Documents*, JPL, February, 2011.

42. Das, N., "Evaluation of Urban/Rural Datasets for the SMAP Mission," *SMAP Science Document* #D-53060, 1.0, JPL, March, 2011.

43. Kerr, Y. H. , Waldteufel, P., Wigneron, J., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M., Font, J., Reul, N., Gruhier, C., Juglea, S. E., Drinkwater, M. R., Hahne, A., Martín-Neira, M.,and Mecklenburg, S., "The SMOS mission: New tool for monitoring key elements of the global water cycle," *Proceedings of the IEEE*, 98(5), pp. 666- 687, 2010.

44. Le Vine, D. M., Lagerloef, G. S., and Torrusio, S. E., "Aquarius and remote sensing of sea surface salinity from space," *Proceedings of the IEEE*, 98(5), pp. 688-703, 2010.

45. SMAP SDS Web, "GloSim 2003," Available online: http://smap-sds-web.jpl.nasa.gov/confluence/display/algorithm/GloSim+2003.

46. SMAP SDS Web, "SMAP Science Data Calibration and Validation Plan," Available online: http://smap-sds-web.jpl.nasa.gov/confluence/display .

47. Jackson, T. J., Cosh, M. H., Bindlish, R., Starks, P. J., Bosch, D. D., Seyfried, M. S., Goodrich, D. C., and Moran, M. S., "Validation of Advanced Microwave Scanning Radiometer soil moisture products," *IEEE Trans. Geosci. Rem. Sens.*, 48, pp. 4256-4272, 2010.

48. Jackson, T. J., Bindlish, R., Cosh, M. H., Zhao, T., Starks, P. J., Bosch, D. D., Seyfried, M. S., Moran, M. S., Kerr, Y., and Leroux, D., "Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the U.S.," *IEEE Trans. Geosci. Rem. Sens.*, 2012, pp: 1530 - 1543.

49. Draper, C. S., Walker, J. P., Steinle, P. J., de Jeu, R. A, and Holmes, T. R., "An evaluation of AMSR-E derived soil moisture over Australia," *Remote Sens. Environ*, 113(4), pp. 703-710, 2009.

50. Gruhier, C., de Rosnay, P., Hasenauer, S., Holmes, T., de Jeu, R., Kerr, Y., Mougin, E., Njoku, E., Timouk, F., Wagner, W., and Zribi, M., "Soil moisture active and passive microwave products: intercomparison and evaluation over a Sahelian site," *Hydrol. Earth Syst. Sci. Discuss.*, 6, pp. 5303-5339, 2009.

51. Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., and Wagner, W., "A possible solution for the problem of estimating the error structure of global soil moisture data set", *Geophys. Res. Let.*, 35, 2008.

52. Miralles, D. G., Crow, W. T., and Cosh, M. H., "A technique for estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations," *Journal of Hydrometeorology*, 11(6), pp. 1404-1410, 2010.

53. Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V., "Error characterisation of global active and passive microwave soil moisture datasets," *Hydrol. Earth Syst. Sci.*, 14, pp. 2605-2616, doi:10.5194/hess-14-2605-2010, 2010.

54. Das, N. and P. O'Neill, "Selection of Soil Attributes Datasets for the SMAP Mission," SMAP Science Document #D-53058, 1.1, JPL, Dec., 2010.

55. Calvo-Alvarado, J., N. McDowell, and R. Waring, "aalometric Relationships Predicting Foliar Biomass and Leaf Area/Sapwood Area Ratio from Tree Height in Five Costa Rican Rain Forest Species," *Tree Physiology*, 28, pp. 1601-1608, 2008.

56. O'Neill, P., T. Jackson, D. Entekhabi, and E. Njoku, "Survey of Tower and Airborne L-Band Sensor Systems Relevant to Soil Moisture Space Missions," *Geoscience & Remote Sensing Newsletter*, IEEE, 151, June, 2009, pp. 13-16.

57. Bindlish, R., T. Jackson, M. Cosh, T. Zhao, and P. O'Neill, "Global Soil Moisture from the Aquarius Satellite: Description and Initial Assessment," IEEE Geoscience and Remote Sensing Letters, 2014, in press, doi 10.1109/LGRS.2014.2364151.

58. Choudhury, B., Schmugge, T., and Mo, T, "A Parameterization of Effective Soil Temperature for Microwave Emission," J. Geophys. Res., 87: 1301-1304, 1982.

59. Jackson, T., P. O'Neill, E. Njoku, S. Chan, and R. Bindlish, "SMAP Project Calibration and Validation for the L2/3_SM_P Beta-Release Data Products," JPL D-93981, jet Propulsion Laboratory, Pasadena, CA, September, 2015.

60. Jackson, T., P. O'Neill, E. Njoku, 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, M. Seyfried, D. Bosch, P. Starks, D. Goodrich, 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, April 30, 2016. *Calibration and Validation for the L2/3_SM_P Version 3 Data Products*, SMAP Project, JPL D-93720, Jet Propulsion Laboratory, Pasadena, CA.

61. 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, December 15, 2016. *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.

62. Chan, S., R. Bindlish, P. O'Neill, E. Njoku, T. Jackson, A. Colliander, F. Chen, M. Burgin, S. Dunbar, J. Piepmeier, S. Yueh, D. Entekhabi, M. Cosh, T. Caldwell, J. Walker, X. Wu, A. Berg, T. Rowlandson, A. Pacheco, H. McNairn, M. Thibeault, J. Martínez-Fernández, Á. González-Zamora, M. Seyfried, D. Bosch, P. Starks, D. Goodrich, J. Prueger, M. Palecki, E. Small, J. Calvet, W. Crow, and Y. Kerr, "Assessment of the SMAP Level 2 Passive Soil Moisture Product," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 54, no. 8, August 2016, pp. 4994-5007, doi: 10.1109/TGRS.2016.2561938.

APPENDIX 1 – L2_SM_P Output Product Data Fields

The specific details of all of the fields in the L2_SM_P and L3_SM_P output products can be found in the L2/3_SM_P Data Product Specification Document. A summary is given in the table below, which includes some fields which will be retained only during the official CV period ending with the first bulk reprocessing.

Data Fields	Description
Time in J2000 seconds	Average TB sample acquisition time in a grid cell
Time in ASCII text	Average TB sample acquisition time in a grid cell
EASE Grid cell row index	Global 36-km EASE2 Grid cell row index
EASE Grid cell column index	Global 36-km EASE2 Grid cell column index
EASE Grid center latitude	L1C_TB center latitude
EASE Grid center longitude	L1C_TB center longitude
EASE Grid centroid latitude	L1C_TB centroid latitude
EASE Grid centroid longitude	L1C_TB centroid longitude
Incidence angle	L1C_TB incidence angle
TBH	H-polarized TB used as input to retrieval [2]
TBV	V-polarized TB used as input to retrieval [2]
TB3	Third Stokes parameter
TB4	Fourth Stokes parameter
TBH quality flag	Quality flag of TBH
TBV quality flag	Quality flag of TBV
TB3 quality flag	Quality flag of TB3
TB4 quality flag	Quality flag of TB4
Soil moisture (SCA-H)	Retrieved soil moisture using the SCA-H algorithm
Soil moisture (SCA-V)	Retrieved soil moisture using the SCA-V algorithm
Soil moisture (DCA)	Retrieved soil moisture using the DCA algorithm
Soil moisture (MPRA)	Retrieved soil moisture using the MPRA algorithm
Vegetation opacity (SCA-H)	Retrieved ‘tau’ parameter derived from NDVI
Vegetation opacity (SCA-V)	Retrieved ‘tau’ parameter derived from NDVI
Vegetation opacity (DCA)	Retrieved ‘tau’ parameter
Vegetation opacity (MPRA)	Retrieved ‘tau’ parameter
Retrieval quality flag (SCA-H)	Quality flag of retrieved soil moisture using SCA-H

Retrieval quality flag (SCA-V)	Quality flag of retrieved soil moisture using SCA-V
Retrieval quality flag (DCA)	Quality flag of retrieved soil moisture using DCA
Retrieval quality flag (MPRA)	Quality flag of retrieved soil moisture using MPRA
Surface flag	Surface conditions that indicate retrievability
Vegetation water content	Vegetation water content derived from NDVI
Soil temperature	Modeled soil temperature (ancillary forecast temperature)
Static water fraction	Areal fraction of static water in a 36-km grid cell
Radar-based water fraction	Areal fraction of transient water in a 36-km grid cell
Radar-based freeze/thaw fraction	Areal fraction of freeze/thaw state in a 36-km grid cell
Top 3 land cover class types	Top three dominant IGBP land cover classes [1]
Top 3 land cover class fractions	Top three dominant land cover class fractions
Albedo	Single-scattering albedo
Roughness coefficient	‘h’ parameter
Soil moisture error	Error of retrieved soil moisture using the baseline algorithm

[1] The crop class (IGBP = 12) is further delineated into crop type classes based on available crop classification information.

[2] Subject to water body correction wherever water fraction is below a threshold.

APPENDIX 2 – Water Correction Before or After T_B Gridding

Initially, it was proposed that water T_B correction be applied to T_B observations before gridding (on the L1B_TB product), not after gridding (on the L1C_TB product). However, upon further analysis it was found that:

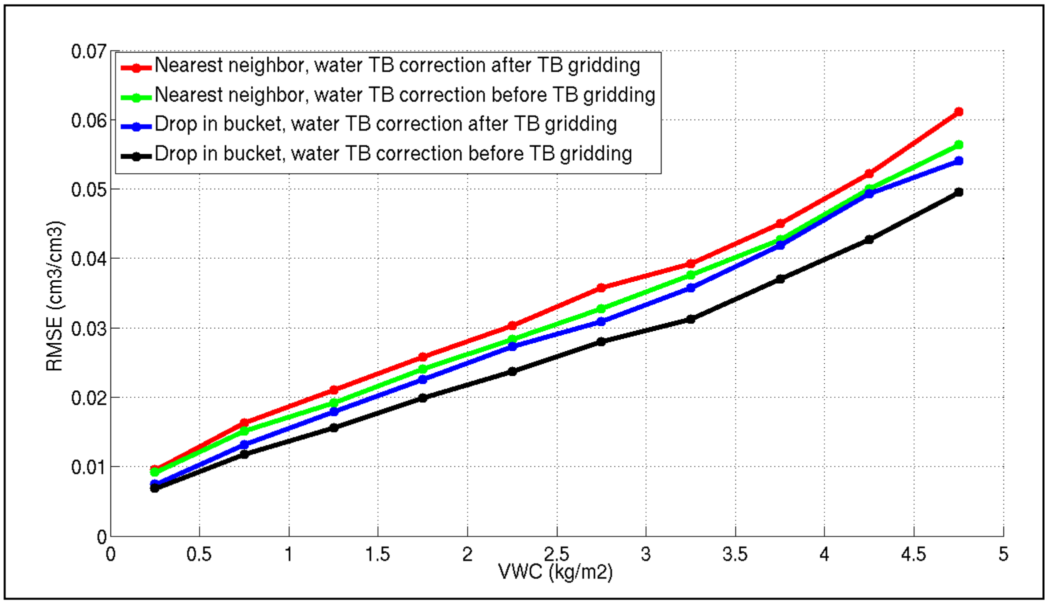
1. Performing water T_B correction before gridding is numerically prohibitive. The process consumes significant CPU resources on determining the half power beamwidth (HPBW) contour and geolocating the radar pixels within it (needed for generation of the water fraction) *for every single T_B sample*. The gain in retrieval accuracy is minimal.
2. In contrast, Level 2 water TB correction is straightforward; it only requires water fraction from gridded sources (e.g., static database or L2_SM_A's water flag).

Scenario	**Gridding Method**	**Water T_B Correction**
1	Drop-in-Bucket	Before T_B gridding
2	Nearest Neighbor	Before T_B gridding
3	Drop-in-Bucket	After T_B gridding
4	Nearest Neighbor	After T_B gridding

Simulations were run comparing soil moisture retrieval accuracy for different scenarios in which the gridding methods and water T_B correction sequence were varied. For all scenarios, retrieval was performed based on global noisy L1B_TB perturbed by $N(0.64, 2.58^2)$, 5% uncertainty in surface roughness, 5% uncertainty in albedo, 5% uncertainty in sand fraction, 5% uncertainty in clay fraction, 2 K uncertainty in soil temperature, and 10% uncertainty in VWC. Retrieval was performed one day per month from Jan 2003 to Dec 2003 over non-frozen areas and areas where VWC ≤ 5 kg/m^2.

As evident from the results, for a given gridding method (e.g., drop-in-bucket), the difference in retrieval accuracy between pre-gridding correction (black lines) and post-gridding correction (blue lines) is minimal. Even at high VWC values, the difference in retrieval accuracy is still less than 0.005 cm^3/cm^3.

H-pol single-channel algorithm (SCA)



V-pol single-channel algorithm (SCA)