Soil Moisture Active Passive (SMAP)

Algorithm Theoretical Basis Document
SMAP L2 & L3 Radar Soil Moisture (Active) Data Products

Revision A
December 9, 2014

Seung-bum Kim, Jakob van Zyl, Scott Dunbar, Eni Njoku
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

Joel Johnson
The Ohio State University, Columbus, OH

Mahta Moghaddam
University of Southern California, Los Angeles, CA

Jiancheng Shi
University of California, Santa Barbara, CA

Leung Tsang
University of Washington, Seattle, WA

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 is Revision A. The ATBDs may undergo additional version updates after SMAP launch.

TABLE OF CONTENTS

1 INTRODUCTION

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

Level 2 Soil Moisture (L2_SM_A) (half-orbit)
Level 3 Soil Moisture (L3_SM_A) (global, daily composite)

1.1 The Soil Moisture Active Passive (SMAP) Mission

1.1.1 Background and Science Objectives

The National Research Council's (NRC) Decadal Survey, *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond*, was released in 2007 after a two year study commissioned by NASA, NOAA, and USGS to provide them with prioritization recommendations for space-based Earth observation programs (National Research Council 2007). Factors including scientific value, societal benefit and technical maturity of mission concepts were considered as criteria. SMAP data products have high science value and provide data towards improving many natural hazards applications. Furthermore SMAP draws on the significant design and risk-reduction heritage of the Hydrosphere State (Hydros) mission (Entekhabi et al. 2004). For these reasons, the NRC report placed SMAP in the first tier of missions in its survey. 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. The target launch date is October 2014 (Entekhabi et al. 2010).

The SMAP science and applications 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.

The Level 1 requirements for soil moisture state "…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/m2 (averaged over the spatial resolution scale)…"

The 10-km soil moisture requirement will be met for SMAP by combining the radar and radiometer measurements in an algorithm that optimizes the high-resolution attribute of the radar with the higher accuracy attribute of the radiometer, as described in the L2_SM_AP ATBD.

The L2_SM_A (radar-only 3 km) product described in this ATBD is considered a research product for SMAP and is not expected to be as accurate as either the L2_SM_P or L2_SM_AP SMAP data products, especially under more densely vegetated conditions (~3-5 kg/m^2). The SMAP mission has targeted 0.06 cm^3/cm^3 as the L2_SM_A accuracy goal, for vegetation water contents up to 5 kg/m^2 (though it is acknowledged that this may not be feasible at the higher vegetation amounts).

1.1.2 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 prohibits the downlink 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 one swath width over the surrounding 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 baseline orbit parameters are:

- Orbit Altitude: 685 km (2-3 days average revisit and 8-days exact repeat)
- Inclination: 98 degrees, sun-synchronous
- Local Time of Ascending Node: 6 pm

Table 1 SMAP Mission Requirements

Scientific Measurement Requirements	Instrument Functional Requirements
Soil Moisture: ~±0.04 m^3m^{-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 land-scape variability (~3 km).	L-Band Radar (1.26 GHz and 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	Swath Width: ~1000 km

(6am/6pm Equator crossing); Global, ~3 day (or better) revisit; Boreal, ~2 day (or better) revisit	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	

The SMAP radiometer measures the four Stokes parameters, V, H and 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 6am/6pm sun-synchronous SMAP orbit.

At L-band anthropogenic Radio Frequency Interference (RFI), principally from ground-based surveillance radars, can contaminate both radar and radiometer measurements. The measurements of the SMOS 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. To combat this, the SMAP radar utilizes selective filters and an adjustable carrier frequency in order to tune to pre-determined 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.

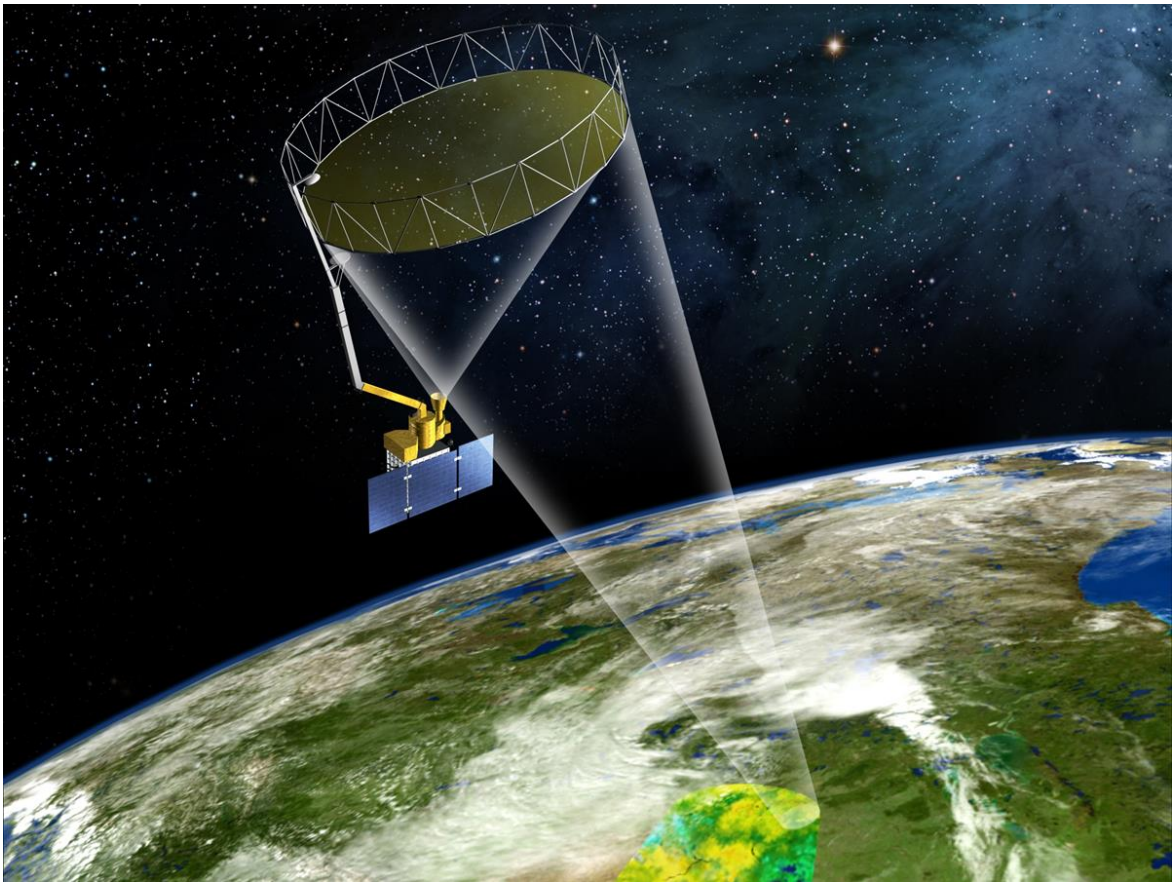

Figure 1 The SMAP observatory is a dedicated spacecraft with a rotating 6-m light-weight deployable mesh reflector. The radar and radiometer share a common feed.

Table 2. SMAP data product table.

Product	Description	Gridding (Resolution)	Latency	
L1A_TB	Radiometer Data in Time-Order	-	12 hrs	Instrument Data
L1A_S0	Radar Data in Time-Order	-	12 hrs	
L1B_TB	Radiometer T_B in Time-Order	(36x47 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	
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_AP	Soil Moisture (Radar + Radiometer)	9 km	24 hrs	
L3_FT_A	Freeze/Thaw State (Radar)	3 km	50 hrs	Science Data (Daily Composite)
L3_SM_A	Soil Moisture (Radar)	3 km	50 hrs	
L3_SM_P	Soil Moisture (Radiometer)	36 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	

1.2 Product Overview

The SMAP planned data products are listed in Table 2. Level 1B and 1C data products are calibrated and geolocated instrument measurements of surface radar backscatter cross-section and brightness temperatures. Level 2 products are geophysical retrievals of soil moisture on a fixed Earth grid based on Level 1 products and ancillary information; the Level 2 products are output on half-orbit basis. Level 3 products are daily composites of Level 2 surface soil moisture and freeze/thaw state data. Level 4 products are model-derived value-added data products that support key SMAP applications and more directly address the driving science questions.

1.3 Instrument Characteristics

This section describes the SMAP radar characteristics related to the L2_SM_A product. The L1C_S0 ATBD offers the full details of the SMAP radar instruments. The SMAP radar operates with VV, HH, and HV transmit-receive polarizations, and uses separate transmit frequencies for the H (1.26 GHz) and V (1.29 GHz) polarizations. The real-aperture ('lo-res') swath width of 1000 km provides global coverage within 3 days or less equatorward of 35°N and 2 days poleward of 55°N. Figure 2 illustrates the instrument and conical scanning configuration.

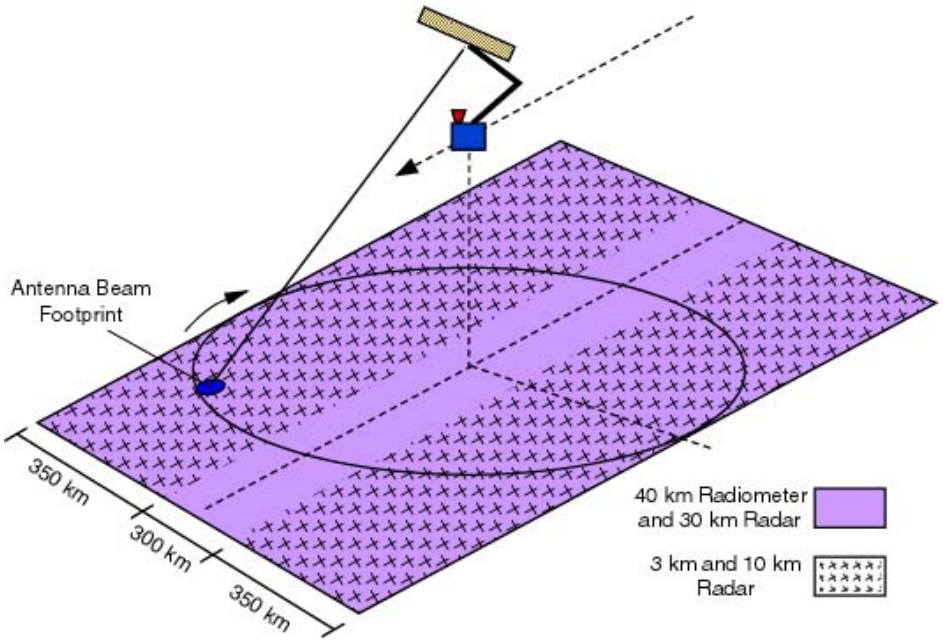

Figure 2. SMAP radar swath and high resolution radar coverage.

The Level 2 and Level 3 radar soil moisture algorithm therefore will re-sample the L1C radar measurements to an Equal-Area Scalable Earth (EASE) grid at a nominal 3 km resolution and estimate soil moisture from the resampled backscatter measurements. The radar's original 1-MHz operating bandwidth yields a ground range resolution of approximately 250 m, so that a minimum of 12 "looks" in range occur within a 3-km range cell. The original azimuth resolution is obtained using an unfocused SAR processing, and is dependent on the total Doppler bandwidth resolved in each footprint. The Doppler diversity has a maximum at scan angles perpendicular to the platform velocity, leading to a single-look azimuth resolution of approximately 450 m. The single-look resolution decreases as the scan angle approaches the nadir track, reaching 1500 m at the inner swath edge (150 km from the nadir track). This varying azimuth resolution across scan causes the number of independent azimuth "looks" within a 3-km cell to vary with scan position. Excluding the nadir track region due to its degraded resolution, the 3-km ('hi-res') resolution portion of the swath is defined as the two outer 350-km-wide segments shown in Figure 2.

The radar relative measurement error (K_p) is impacted by the radar measurement precision (K_{pc}) and any relative calibration errors (K_{pr}) or radio frequency interference contributions (K_{pRFI}). The measurement precision depends on the signal-to-noise ratio (SNR) and the number of independent samples or 'looks' averaged in each measurement. The basic equation for the radar measurement precision, K_{pc} is:

$$K_{pc} = \frac{1}{\sqrt{N_a N_e}}\left(1+\frac{1}{SNR}\right) \qquad (1.3\text{-}1)$$

where K_{pc} is the normalized backscatter standard deviation, $\Delta\sigma^o/\sigma^o$, and the SNR is defined relative to the noise equivalent cross-section σ^o_{ne}, i.e., SNR = σ^o/σ^o_{ne}. The noise equivalent cross-section for SMAP varies with the scan position and ranges from σ^o_{ne} = -31.8 dB (inner edge) to -29.5 dB (outer edge). N_a and N_e are the total number of azimuth and elevation looks, which also are functions of the scan angle as discussed previously. In the outer 350 km (hi-res) portion of the swath the total number of

looks in a 3-km cell varies from about 30 to 100. It is assumed in equation 1.3-1 that the radar's estimation and subtraction of the receiver noise power contribution to the observed σ^o utilizes a very large number of looks so that this process does not impact K_{pc}.

Assuming that in most situations K_{pc} is determined primarily by the number of looks and is not limited by the SNR (this may not be true for smooth and/or dry bare surfaces or for inland water bodies, particularly for HV measurements), the value of K_{pc} varies from about 0.22 to 0.14 from the inner to the outer edge of the hi-res swath for a worst case -25 dB signal. Calibration and RFI contributions are described in detail in L1C_S0_HiRes ATBD and the SMAP radar error budget document (Chan 2011), and the combined estimated values of K_{pr} and K_{pRFI} for SMAP radar products at 3-km resolution are 0.35 and 0.4 dB, respectively. When the three dominant components of the error sources are combined and converted to dB, the range of radar relative error across the swath becomes approximately 1.0 to 0.7 dB. Because the SMAP radar will observe particular locations both in the "fore" and "aft" portions of the scan during an overpass, soil moisture retrievals will average these two measurements to reduce K_{pc} (assuming no azimuthal variations in the true cross section) to 0.73 and 0.55 dB for outer and inner edges, respectively. The nominal K_p levels assumed for 3-km SMAP radar observations for both HH and VV channels therefore range from 0.7 down to 0.5 dB in the following sections. For the cross-pol, K_{pr} and K_{pRFI} are 0.72 and 0.8 dB respectively. The cross-pol's total K_p levels after the fore and aft averaging are estimated to be 1.3 and 1 dB. SMAP provides only one of HV and VH. Therefore it is not expected to further improve the speckle of HV by averaging the two cross-pol observations. Some important characteristics of the SMAP mission and its synthetic aperture radar are provided in Table 3.

Table 3. Characteristics of the SMAP 3-km resolution σ^o.

Swath width	1000 km with 300 km-wide nadir-gap
Temporal revisit period	3 day (35°N) 2 day (55°N) Global coverage within 8 days
Radiometric accuracy (relative, 1σ) [*¶§]	For co-pol, 0.5 dB and 0.7 dB at outer and inner edges of the swath, respectively. For cross-pol, 1 dB and 1.3dB at outer and inner edges of the swath, respectively.
Number of single looks per 3km	200 and 60 at outer and inner edges of the swath, respectively [¶]
Noise equivalent σ^0§	-28.5 dB and -31.5 dB at outer and inner edges of the swath, respectively.

[*] Values include speckle, relative calibration error, and radio frequency interference residual error.
[¶] Averaging of fore-scan and aft-scan images is incorporated.
§ These values are derived based on the required performance of the radar system. The current best estimate (CBE) values are better than the required values.

The use of differing frequencies for horizontal and vertical polarizations implies that the speckle in HH and VV cross sections is uncorrelated. SMAP is therefore not capable of measuring any correlation information in HH and VV returns. SMAP will report only one of either VH or HV returns (reported cross section will average these quantities), and could in theory determine correlations between co- and cross-pol fields. However, whether these correlations will remain for averages over 3-km spatial resolution needs to be examined.

1.4 Historical Perspective

Previous spaceborne L-band synthetic aperture radar (SAR) systems include the SeaSat (1978), SIR-A (1981), SIR-B (1984), SIR-C (1994), JERS-1 (1992-98), and PALSAR (2006-2011). Of these

systems, only SIR-C (Bindlish and Barros 2001; Dubois et al. 1995; Shi et al. 1997) and PALSAR included multi-polarization capabilities. While some success in retrieving soil moisture has been shown (primarily for bare surfaces) with these instruments, none of these systems were specifically designed for soil moisture measurement purposes. In particular the relatively small swath sizes achieved by SAR systems result in temporal revisits typically of thirty days or more, limiting both hydrologic applications and the benefits of any potential time series retrieval approach. Due to these challenges and more than 40 years of investigation, the goal of achieving reliable absolute soil moisture measurements from satellite radar observations remains largely unfulfilled.

Because the use of radar is simplified if only "bare" surfaces are considered (i.e. surfaces with little or no vegetation coverage), most existing radar-only retrieval studies and algorithms are focused on bare surfaces alone. Empirical or semi-empirical algorithms (Dubois et al. 1995; Oh et al. 1992) have been applied to bare surfaces. Functional fits to the Integral Equation Model (IEM) (Shi et al. 1997) for bare surfaces have also been reported. A numerical inversion of the IEM (Verhoest et al. 2007) was also successful for bare surfaces.

Because bare surfaces account for only ~ 14% of the global land area, an increased global coverage for radar soil moisture estimates can be achieved only if algorithms applicable to vegetated surfaces are utilized. The use of L-band radar as opposed to higher frequencies such as C- and X-band has been shown in both theoretical and in ground-based and airborne field campaigns to be advantageous in reducing the influence of vegetation of measured cross sections (De Roo et al. 2001; Moghaddam et al. 2000), but vegetation remains a significant contributor to observed data. While some references have demonstrated the possibility of empirically correcting vegetation effects (Bindlish and Barros 2001), (Joseph et al. 2008)) and subsequently applying a bare surface retrieval algorithm, a robust empirical correction applicable globally remains to be determined, limiting the applicability of this approach.

Because vegetation, roughness, and soil moisture all impact radar measurements, it is difficult to retrieve accurately absolute soil moisture from a single radar measurement. Retrievals become more feasible if multiple simultaneous radar measurements (for example in multiple polarizations), a time series of radar measurements, additional ancillary data on vegetation or roughness(Joseph et al. 2008), or combinations of such measurements or information are available.

The more frequent revisit of the SMAP radar clearly enables the use of time series approaches to improve retrieval. Several examples of time series retrievals have been reported (Kim et al. 2012b; Mattia et al. 2009) and have shown clear improvements over "snapshot" (i.e. performing a retrieval using only the current observation) methods. Some time-series algorithms (Kim and van Zyl 2009) estimate only the relative change in soil moisture. The estimate of an index of such change (Wagner et al. 1999) has been verified extensively using C-band spaceborne scatterometer data. However, when applied to 1 to 3-km resolution SAR data, radar speckle begins to degrade the performance of this method (Pathe et al. 2009). It is noted that each of the previous time series algorithms has limitations that make it difficult to provide quantitative estimates of soil moisture globally with the required accuracy for SMAP applications. The limitations include an assumption of temporally static vegetation conditions, a failure to report absolute soil moisture values, or a reliance on other analyses such as an ancillary geophysical model estimate of soil moisture.

The incorporation of extensive ancillary data is also desirable to improve retrieval performance. However such detailed information is not available globally, and retrievals using satellite measurements must incorporate coarser and less descriptive, but globally available, ancillary descriptions of vegetation or roughness. Surface roughness information in particular is largely unavailable for global land surfaces; approaches for incorporating a stochastic description of surface roughness properties into the retrieval process have recently been described (Verhoest et al. 2007).

It is also noted that the high spatial resolution of previous SAR observations makes them more susceptible to an additional confounding factor: the "row effects" of agricultural surfaces that cause changes in the surface radar signal with azimuthal angle (Zribi et al. 2002). Although it is expected that the impact of such effects should be reduced at the 3-km spatial resolution of SMAP, the assessment of such effects continues.

1.5 Product/Algorithm Objectives

The foremost objective of the SMAP L2_SM_A soil moisture product is to provide reliable soil moisture retrieval at 3-km resolution. This product will provide finer resolution of the spatial heterogeneity of soil moisture fields as compared to the 9 km L2_SM_AP and 36 km L2_SM_P SMAP products. Both the L2_SM_P and the L2_SM_AP data products are required to meet the 0.04 cm^3/cm^3 soil moisture accuracy specified in the Level 1 baseline mission requirements (Section 6.1). No formal mission requirements are placed on the L2_SM_A (radar-only 3 km) product since this is considered a research product and is not expected to perform as accurately under more densely vegetated conditions. However, this product will be made available publicly as a standard product if its quality, as assessed during the post-launch Cal/Val phase, can be shown to merit general distribution. The SMAP mission has targeted the threshold mission soil moisture accuracy of 0.06 cm^3/cm^3 as the L2_SM_A accuracy goal, for vegetation water contents up to 5 kg/m^2 (it is acknowledged that this may not be feasible at the higher vegetation amounts).

The capabilities of the SMAP radar (multiple polarizations, ~ 3 day revisit, 3-km resolution, use of L-band) have not been previously available simultaneously on a global scale. Thus, SMAP will provide a unique opportunity for understanding the impact of soil moisture on SAR backscatter, leading to a retrieval approach that can combine multiple polarization and time series strategies with available ancillary information. This will enable SMAP to map soil moisture with high temporal frequency (every 2-3 days) with benefit to a wide range of applications at 3-km or better spatial scale in hydrology, agriculture, natural hazards, and human health.

1.6 Outline

This document contains the following sections: Section 2 describes forward models and associated physics governing the radar scattering from bare and vegetated surfaces, and includes example predictions from these models; Section 3 presents the baseline and optional retrieval algorithms (Sections 3.2 and 3.3). Section 3.4 discusses the input and output products, along with the use of ancillary data and various flags. The error budget of the baseline algorithm and the pre-launch algorithm testing on which this error budget is based, are provided in Section 3.5, along with the tests and procedures that will be employed for downselecting the baseline algorithm The L3_SM_A process is discussed in Section 3.6. Section 3.7 presents pre-launch algorithm testing plans. Important practical issues in the operational implementation of the baseline algorithm are discussed in Section 3.8. Section 4 presents procedures for validating the data products after launch and approaches for modifying the algorithms as needed.

2 PHYSICS OF THE SCATTERING PROBLEM

Forward models provide predictions of average co-polarization and cross-polarization backscattering in terms of physical parameters that are used to describe the vegetation-covered soil layer. These are the soil dielectric constant, surface roughness statistics, and vegetation properties, as described below:

1) The soil is most commonly described as a homogeneous medium having a complex dielectric constant $\varepsilon = \varepsilon_r - j\varepsilon_i$ that is a function of the volumetric soil moisture m_v as well as the soil texture, temperature, and bulk density; several empirical models exist for this relationship (Dobson and Ulaby 1986; Hallikainen et al. 1985; Mironov et al. 2004; Peplinski et al. 1995). Thus, forward models can be computed as a function of the soil complex permittivity with a choice of dielectric constant model and associated parameters to be selected in inverting the dielectric constant into soil moisture. It is to be noted that the radar response may be impacted by sub-surface dielectric discontinuities in the near surface region; the depth to which such effects can be significant is increased for drier soils.

2) The interface of the soil layer is usually described as a stationary Gaussian random process having a specified correlation function. Fractal surface descriptions have also been used in some studies but are not considered here. Studies of soil surface profiles and soil moisture remote sensing have shown varied results with regard to the form of the surface correlation function. Controlled field measurements of soil surface profiles (Oh et al. 1992) favor a directionally isotropic exponential correlation function, described by the RMS height s and correlation length l parameters alone. A directionally isotropic description neglects any directional row or tillage features. The degree to which such features would impact radar cross sections at 3 km spatial resolution remains a subject of investigation. Other correlation functions for soil surfaces have also been investigated (Li et al. 2002). Empirical models on the other hand neglect even the correlation length parameter and describe surfaces in terms of an RMS roughness alone. An intermediate approach involves an assumption that the surface correlation length is a fixed multiple of the surface RMS height, e.g., (Fung et al. 1992).

3) A variety of approaches exist for describing vegetation media, including characterizations of vegetation structures such as stalks, trunks, and leaves in terms of canonical cylindrical or disk shapes with specified size and orientation distributions in a set of vegetation layers (P.de Matthaeis 2005; Stiles et al. 2000; Yueh et al. 1992). Simpler approaches only use the vegetation water content (VWC) to provide analytical forms for attenuation and scattering effects. Because ancillary information on vegetation properties again are limited, forward models using detailed vegetation descriptions must use *a priori* information that is a function only of the vegetation classification and/or VWC. A description of the dielectric constant of the vegetation as a function of the water content is also implicit in such approaches (Matzler 1994; Ulaby and El-rayes 1987). Because even small amounts of vegetation can significantly impact radar returns, particularly in cross polarization, "bare surface" models that neglect vegetation effects will have a limited range of applicability to the SMAP project.

Given these parameters, it is typical to describe the radar backscattering for a vegetation-covered soil layer as a sum of three dominant contributions:

$$\sigma_{pq}^{t} = \sigma_{pq}^{s}(\varepsilon, s, l) exp\left(-\tau_{pq}(W)\right) + \sigma_{pq}^{v}(W) + \sigma_{pq}^{sv}(W, \varepsilon, s, l) \qquad (2\text{-}1)$$

In this expression, σ_{pq}^{s}represents the total radar scattering cross-section in polarization pq (=HH, VV, or HV for the SMAP radar), $\sigma_{pq}^{s} exp\left(-\tau_{pq}\right)$ represents the scattering cross-section of the soil surface modified by the two-way vegetation attenuation, σ_{pq}^{v} is the scattering cross-section of the vegetation volume, and σ_{pq}^{sv} represents the scattering interaction between the soil and vegetation. In addition, τ is the vegetation opacity along the slant path of a radar beam, and W is the vegetation water content.

Because SMAP is required to estimate soil moisture over global land regions, forward backscatter models must be adopted or developed for SMAP that are applicable over the full range of land conditions anticipated (but excluding regions that will be masked out of the retrievals, i.e., urban areas, mountainous regions, water bodies, dense vegetation, snow and ice). In this section we describe the forward models that will be used for SMAP, on which the retrieval algorithms are based. As a precursor to this discussion Table 4 lists the landcover classes that will be used by SMAP retrieval algorithms to provide a self-consistent suite of products. This table is based on the classification schemes proposed by the International Geosphere-Biosphere Programme (IGBP) and the crop statistics provided by the United Nations Food and Agriculture Office (UN FAO, http://faostat.fao.org).

Table 4 Land surface classes that form the basis for SMAP soil moisture retrievals: 12 IGBP classes and 4 crop classes (see the references to the SMAP ancillary report for land cover class, in Section 6.1)

1 Evergreen Needleleaf Forest
2 Evergreen Broadleaf Forest
3 Deciduous Needleleaf Forest
4 Deciduous Broadleaf Forest
5 Mixed Forests
6 Closed Shrublands
7 Open Shrublands
8 Woody Savannas
9 Savannas
10 Grasslands
11-14 Croplands : Wheat, Rice, Corn, Soybean
15 Cropland/Natural Vegetation Mosaic
16 Barren or Sparsely Vegetated

For each landcover class shown in Table 4 a forward model must exist on which the retrieval is based. The process for developing such models is described below. The process includes development of a soil model component (bare soil component) and a vegetation layer model above the soil component that can have various forms needed to describe the different types of vegetation shown of Table 4.

2.1 Bare Soil Component Forward Models and Comparisons

Forward modeling studies of scattering from bare surfaces have been based either on approximate or numerical models for the solution of Maxwell's equations. Approximate models include, for example, the Small Perturbation Method (SPM) (Tsang and Kong 2001), the Integral Equation Model (IEM) (Fung et al. 1992), the Advanced Integral Equation Model (AIEM) (Chen et al. 2003), and the Small Slope Approximation (SSA) (Voronovich 1994). Such methods are advantageous because they provide predictions of the expected value of radar returns with minimal to moderate computational requirements. However their inherent approximations can cause errors in predicting "true" radar returns in some situations.

Numerical methods in contrast avoid such approximations, but require Monte Carlo simulations and much greater computational costs to obtained predictions. Numerical methods include, for example, the Method of Moments (MoM) (Tsang et al. 2001), the Extended Boundary Condition Method (EBCM) (Kuo and Moghaddam 2007), the finite element method (Lawrence and co-authors 2010; Lou et al. 1991) and the finite difference time domain method (Chan et al. 1991). "Fast" methods to further improve computational efficiency have also been developed, including the Sparse Matrix Canonical Grid (SMCG) method (Johnson et al. 1996), the Physical Based Two Grid (PBTG) method (Li and Tsang 2001), and the multilevel UV method (Tsang et al. 2004). Fully 3D simulations of Maxwell equations (where the height function $z=f(x,y)$ of the rough surface varies in both horizontal directions)

are required to predict realistic surface behaviors. 3D full wave method of moments simulations based on the "Numerical Maxwell Model in 3 Dimensions" simulations (NMM3D) began in the mid-1990's (Tsang et al. 1994; Tsang et al. 2001).

NMM3D of the Univ. Washington (UW) was selected as the benchmark forward model for the bare surface class. The UV/PBTG/SMCG NMM3D method was used to compute L-band 40 degree 3-D surface backscattering for 200 cases including varying surface RMS heights, correlation lengths, and soil permittivities for co-polarization (Huang et al. 2010) and cross-polarization (Huang and Tsang 2012). Sample results from these simulations (averages over a minimum of 30 Monte Carlo realizations for each case) are shown in Figure 3. Based on these cases, interpolation tables (or "data cubes") were created (interpolated values are within 0.2 dB of the original data values). Since the maximum RMS height considered is 0.21 wavelength (ks=1.32 which is about 5 cm at L band, where k and s are the wavenumber and RMS height), the cases simulated and the interpolations used can be applied to cover a wide range of interests for SMAP. These results were compared with the Dubois empirical model (Dubois et al. 1995), SPM, KA, and AIEM. In parallel, a group in the Univ. Michigan Group used a stabilized EBCM method (Duan and Moghaddam 2011) and computed results up to ks=1.

Results of the UW NMM3D look-up tables were compared with field measurements of co- and cross-polarized backscattering (Oh et al. 1992). The field data includes measurements of soil permittivity, RMS heights, and correlation lengths. The soil surface correlation functions were also measured and found best matched by exponential correlation functions. These ground truth parameters were simulated with NMM3D using the exponential correlation function description. Figure 4 shows the comparison of NMM3D predictions and measured data, and shows that good agreement is achieved. The cross polarization results of NMM3D are also in good agreement with experimental data.

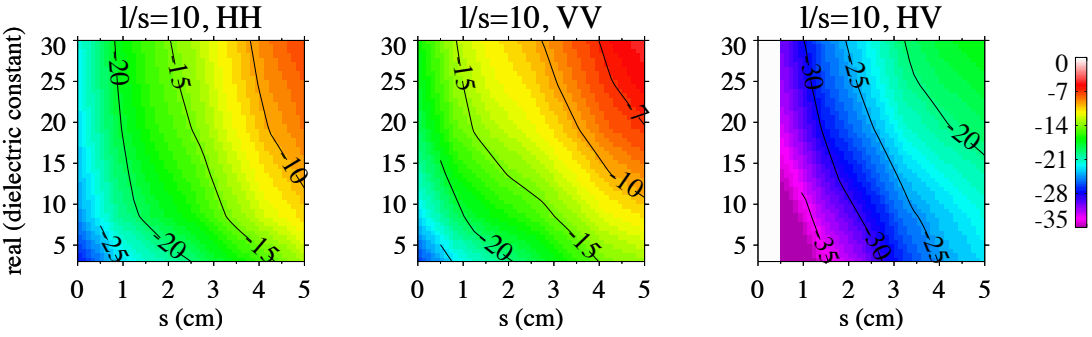

Figure 3. σ^0 (dB) generated by NMM3D bare surface simulations. l and s denote the surface correlation length and RMS height, respectively. Simulations were not performed for very smooth (s < 0.5 cm for HV) surfaces.

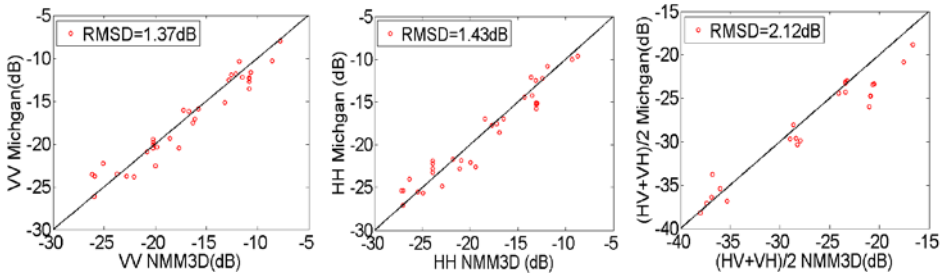

Figure 4. Comparison of backscattering coefficients in dB between NMM3D and Michigan measurement data.

2.2 Models for non-Woody Vegetation and Comparisons

2.2.1 Vegetated Surface Model

The primary approach utilized is a "discrete scatterer" approximation (Lang and Sighu 1983; Tsang et al. 2000), in which each vegetation object, such as a cylinder or disk, is assumed to scatter independently. Fields from each vegetation component are summed and averaged over size and orientation distributions. A first order iterative solution of the radiative transfer equation (Karam et al. 1992; Tsang et al. 1985) yields similar predictions but neglects a coherent double bounce effect. Several variations of the discrete scatterer model exist (Arii et al. 2010; Chauhan et al. 1994), depending on the fidelity with which the vegetation-ground interaction term is treated, the number of vegetation layers included, the method utilized to compute scattering from vegetation and surface components, and the approach used for estimating attenuation.

The level of detail used in describing the vegetation also varies in such models, but typically involves *a priori* assumptions regarding canonical shapes of vegetation components and their size and orientation distributions in an assumed set of vegetation layers (Arii et al. 2010). The required computations for the analytical approaches can vary widely depending on the assumptions utilized, but in most cases remain significant so that pretabling of model results into a "data cube" rather than "on-the-fly" evaluations are necessary.

The UW data cube (Xu et al. 2010), which uses the discrete scatter approximation, is applied for four crop land types that were selected in the order of the harvest area of the globe (grassland, corn, soybean, and wheat). Experimental data are also available for these types. Depending on the basic structure of the crop type, vegetation components are represented by cylinders, disks and spheroids. For the case of small radius cylinders, and for thin disks, analytic approximations are used (Tsang et al. 2000) to calculate the scattering by these objects. When the radius of cylinder is larger, numerical solutions of Maxwell equations are solved for the object through the Body of Revolution (BOR) approach (Mautz and Harrington 1979). The use of Maxwell's equations maintains the smoothness of the results when the cylinder radius increases from small to moderate values.

For a layer of vegetation, the distorted Born approximation is applied for the mean field calculation. The scattering cross section of the vegetation layer and its interaction with the soil surface were derived using a half space Green's Function. The results are expressed as three scattering mechanisms. I) The direct volume scattering, II) The double bounce effect as exhibited by rough surface effect on the interface of the vegetation layer and soil, and III) The rough surface scattering of the soil (Figure 5). A factor of 2 is incorporated in the wave model, which accounts for the the double-bounce forward scattering that the radiative transfer model does not simulate. The wave model obeys reciprocity of backscattering in Maxwell equations and physically accounts for constructive interference in the backscattering direction (Tsang et al. 1985). The coherent reflectivities of the random rough surface of double bounce effects are calculated from the numerical solution (NMM3D) of the rough surface as described in the bare soil calculations (Huang et al. 2010). The bare soil scattering is obtained through the NMM3D look-up table and then reduced by the computed attenuation through the vegetation layer.

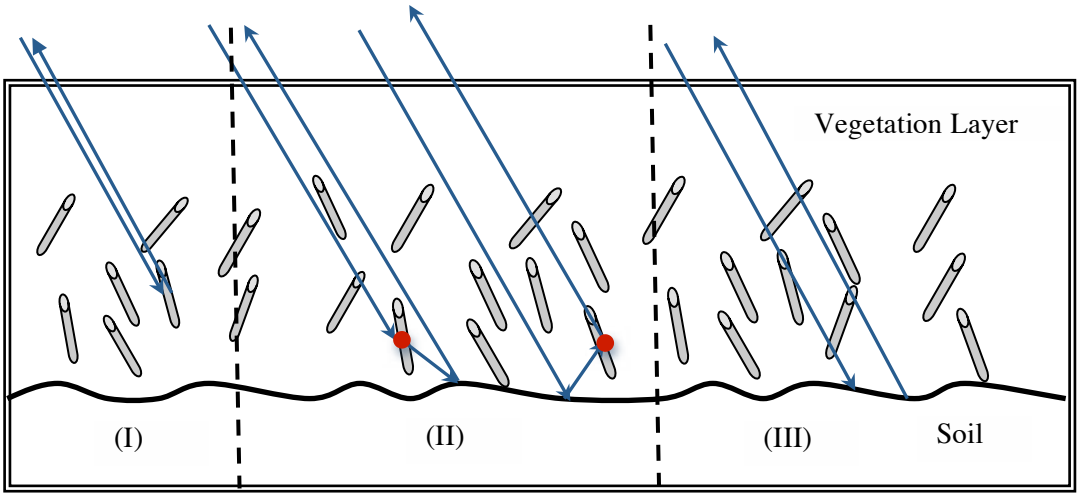

Figure 5. Three scattering mechanisms.

2.2.2 Comparisons with Field Observations

To validate the forward model, comparisons were made between the model predictions and field experiment data from the SGP99 and SMEX02 campaigns. Figure 6 demonstrates the model structures for the two typical crop types. The grassland was represented by thin cylinders uniformly distributed above the rough soil surface. The cylinder size was chosen as radius 2 mm and 50 cm height. The VWC is varied by the radius and grass density. In the grass cube, the radius is a best fit parameter, benchmarked by using the SGP99 measurement data. An empirical relationship between the radius and VWC is formed. All the rest of the grass parameters are kept invariant. According to the measurement from SGP99, most of the grass land has fairly small vegetation water content (VWC less than 2 kg/m^2). In the grassland, when the VWC is small, the vegetation layer serves as an attenuation layer, so that the total backscattering is dominated by surface scattering and is strongly dependent on soil moisture. As the VWC increases, the volume scattering contributes more to VV polarization and the double bounce contributes more to HH polarization. The cross polarization is mainly due to volume scattering from the vegetation layer. The other two grass-like crops are wheat and rice fields before forming grains. Thin cylinder models are used in both these cases also.

Figure 6 also illustrates the corn field model structure. The corn stalk and leaves are modeled as different size cylinders. Larger radius cylinders (radius ~1.7cm) with vertical orientation are used to represent the stalk. Smaller cylinders with a random orientation distribution are used to model the leaves. VWC is varied by the moisture content of the stalk and leaves, whiling keeping the corn structure invariant. Unlike the grass-like classes, the total backscattering is mainly dominated by the double bounce scattering. The soil moisture sensitivity is due to the change of the reflectivity of the coherent wave in the double bounce mechanism.

Soybean's stems are represented by cylinders and the leaves are represented by disks. In the soybean model, the vegetation height and the leaves volume are changing with VWC. The relationship is derived from the SMEX02 in situ measurement empirically. The thickness of the leaves and the ratio of the cylinder's diameter, length and disk's diameter are fixed. In the soybean case, both stems and leaves have significantly contribution to the total backscattering.

The measurement data for comparison in Figure 6 are all from the month of peak biomass. The model structures of each crop are based on the growing phase. For the pre-season and harvest season, the geometry parameters will likely need to be modified.

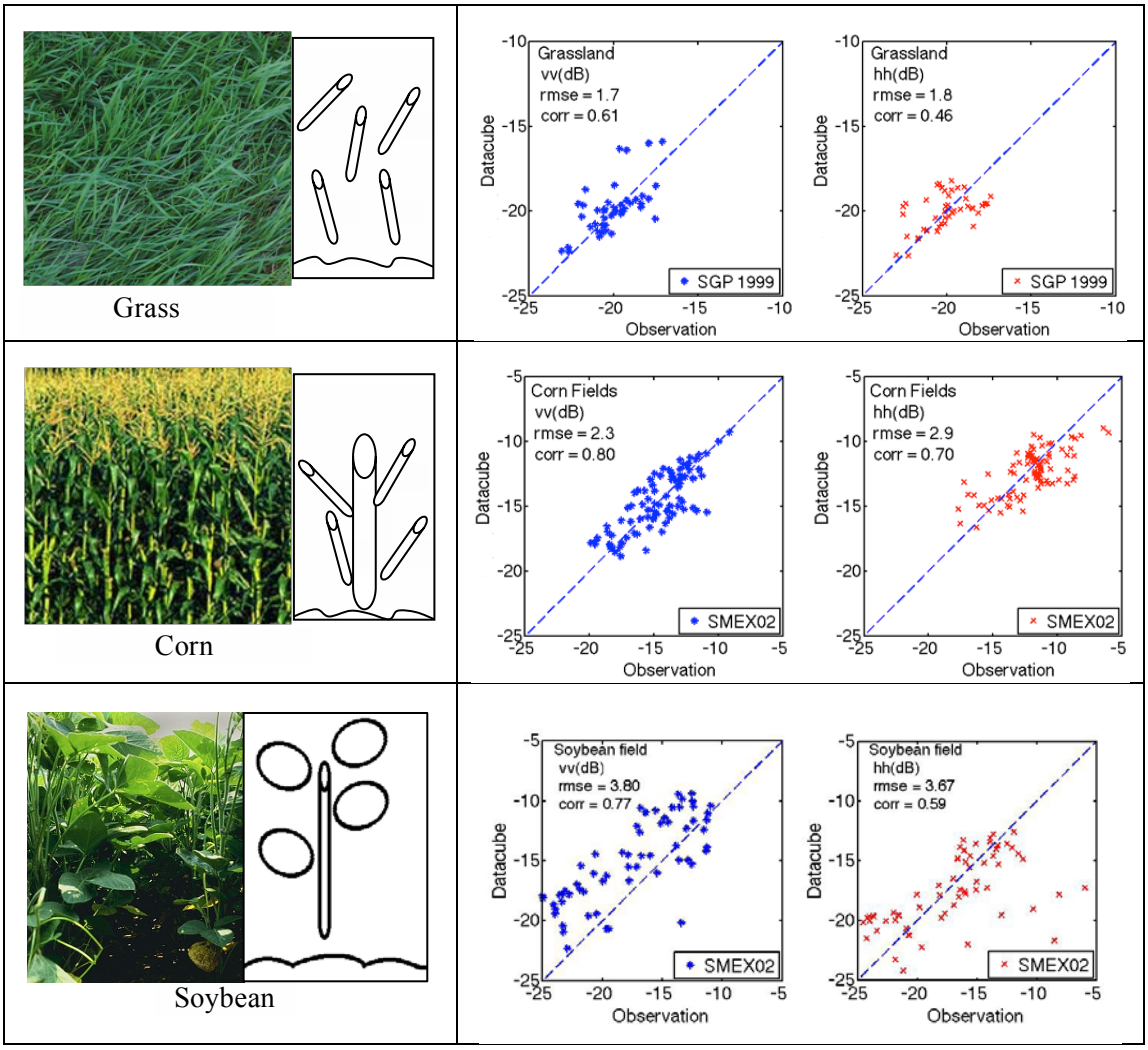

Figure 6. Model structure and evaluation of simulation for grassland, corn, and soybean field.

2.3 Models for Woody Vegetation and Comparisons

2.3.1 Vegetated Surface Model

For woody vegetation, the layered scattering geometry and vegetation model shown in Figure 7 are used. Because coherent effects are prominent in the specular interactions of crown-ground and trunk-ground, it is important that the discrete scatter approach (wave model) is used, in contrast to radiative transfer theory.

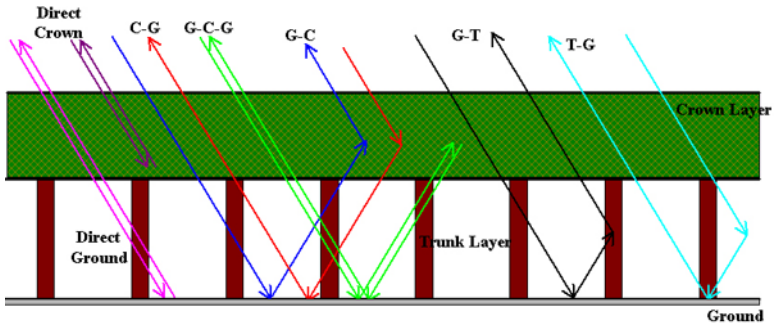

Figure 7. Generalized model of backscattering from a woody (forested) vegetation canopy. The signal can follow a number of paths as it interacts with vegetation and ground. These paths include (1) direct backscatter from the crown layer, (2) direct backscatter from the ground, (3) double-bounce scattering between crown and ground or the opposite path (C-G and G-C), (4) double-bounce scattering between trunks and ground or the opposite path G-T and T-G), and (5) backscattering from the ground to crown to ground (G-C-G), which is often quite small compared to the other 4. The crown layer consists of discrete random scatterers representing branches and leaves.

First, the scattering matrix of a single finite cylinder of arbitrary orientation is derived based on scattering theory. The scattering from the ensemble is then calculated by integrating over an arbitrary, but realistic, probability density function for the size and orientation of scatterers within a given volume of vegetation. The interactions between vegetation layers can be accounted for by cascading the scattering matrices of the relevant layers. The Mueller (or Stokes) matrices M_i of the i^{th} layer is calculated and superimposed to calculate the total scattering from the vegetated surface. Additionally, attenuation due to each layer is calculated from the forward scattering matrix, and is included in the calculation of the backscattered signal from the layer below it. The inclusion of the layer attenuation is also referred to as the distorted Born approximation, since it modifies the single-scattering (Born approximation) solution by allowing the waves to travel in an equivalent, lossy medium, which is more realistic than traveling in free space (Chauhan and Lang 1991; Durden et al. 1989; Israelsson et al. 2000; Lin and Sarabandi 1999; Thirion et al. 2006). The RT and wave-based approaches was compared (Saatchi and McDonald 1997).

Although the two-layer vegetation is the most widely used model, most realistic terrain has vegetation that is of either mixed species or the same species at different stages of growth. A vector RT formulation has been developed to accomplish this task (Liang et al. 2005), as well as a parallel approach using wave theory (Burgin et al. 2011).

The scattering geometry is shown in Figure 8 (upper two panels), where it is seen that an arbitrarily large number of layers can be considered, each with a distinct distribution function for the scatterers. Each layer contains discrete scatterers of several species Figure 8 (bottom panel). These scatterers are a combination of large or small branches and leaves/needles, as well as trunks for softwood trees in branch layers (due to the extension of the stem into the canopy layer) or trunks in trunk layers.

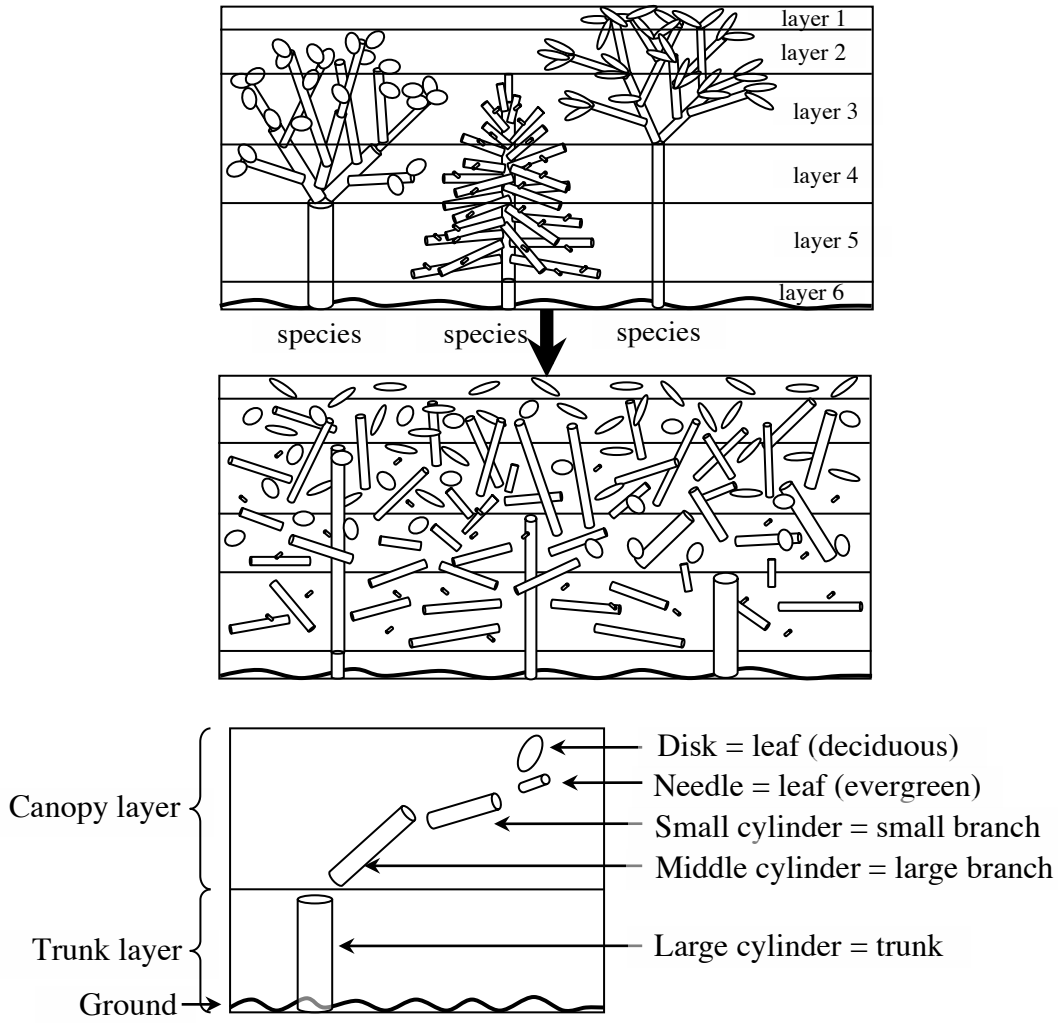

Figure 8. General geometry for a multispecies forest stand, where an arbitrary number of random medium layers can be present. Discrete-scatterers are building blocks of each species type and each random medium layer.

2.3.2 Comparisons with Field Observations

Similar to the case of scattering models of grass and crops, to simplify the retrieval of soil moisture, an optimization approach based on searching within scattering model data cubes is adopted for woody vegetation. The data cubes for each polarization have 3 input parameters. The three polarizations are VV, HH, and HV backscattering and the 3 input parameters (or variables) are soil moisture RMS height (*ks*), soil moisture expressed as the equivalent dielectric constant of soil (ε_r) and vegetation water content (VWC). To construct the data cubes for each woody vegetation type, the forward scattering model described above is parameterized with species-specific input parameters over realistic expected ranges of those parameters. Each axis of the data cube is sampled at a few points (on the order of 10 points), and interpolations are used to construct a dense space of the backscattering coefficients for the entire range of vegetation and soil parameters represented by the data cubes.

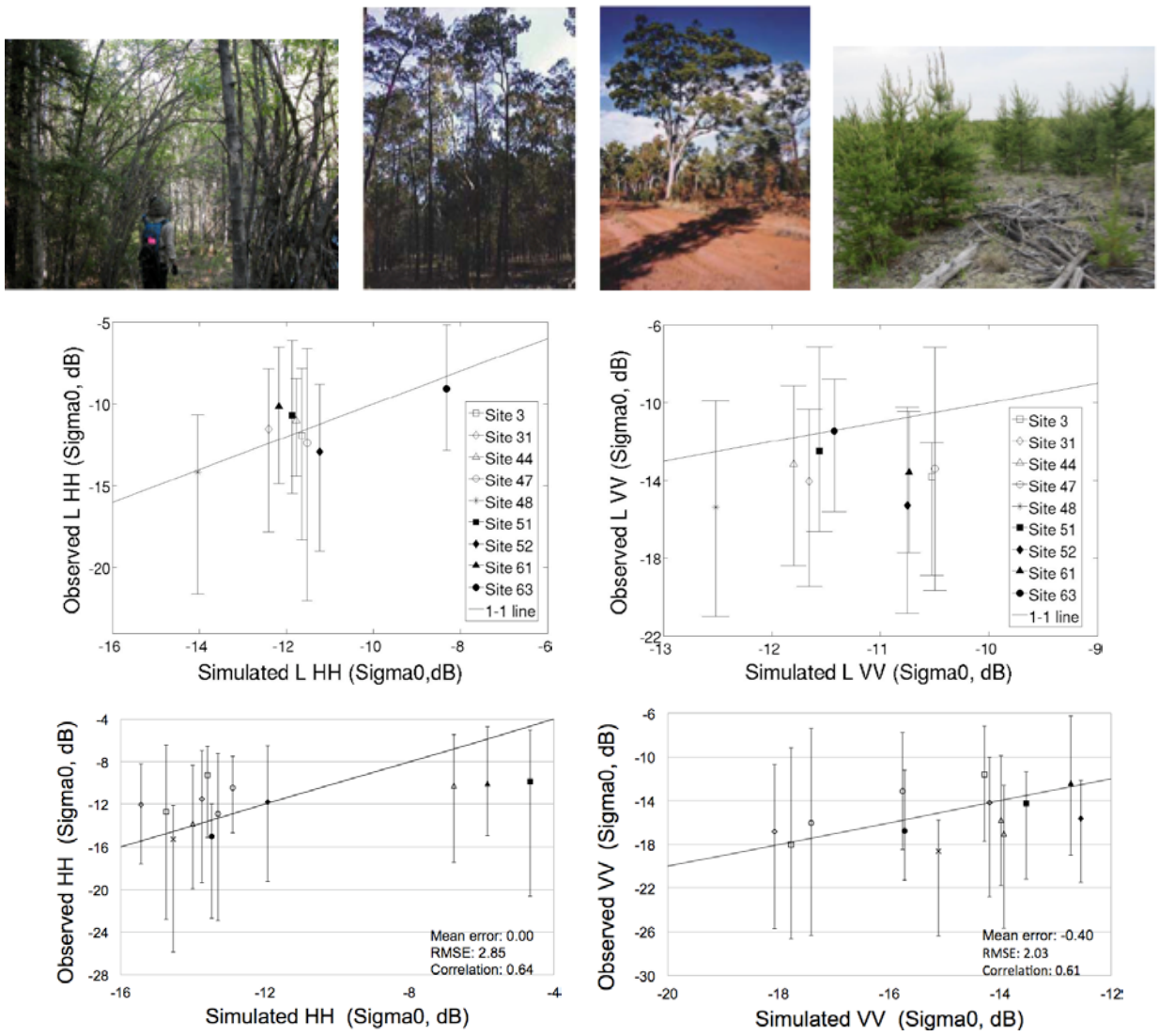

Figure 9. Top row: Examples of woody vegetation species for which field data and SAR data (airborne or spaceborne) were available; results of discrete scatterer model (Burgin et al. 2011; Durden et al. 1989) simulated and validated with SAR data. Middle row: validation using ALOS/PALSAR data for a subset of the plots. HH and VV polarizations are shown. Bottom row: validation using AIRSAR data for a different subset of the plots. HH and VV polarizations are shown. Each point on these plots may correspond to a forest stand with one or more species. The vertical error bars show the speckle distribution within a 5x5 pixel box. Mean error corresponds to model bias (also indicative of calibration bias). If the bias is removed, then the RMSE corresponds to model accuracy.

To validate the date cubes for woody (treed) vegetation, the results of the discrete scatterer model with airborne (AIRSAR) and spaceborne (ALOS PALSAR) data are compared. The ability to simulate radar backscattering coefficients from multiple species allows a combination of calculation results with a wide range of canopies. Validation of the model was undertaken using field data collected from the Tara Downs subregion of the Brigalow Belt Bioregion (BBB) and the Injune Collaborative Landscape Project (ICLP) study area in Queensland, Australia. Field data were available for 82 plots. Nine plots from Tara Downs and 14 plots from ICLP were chosen for model validation. The choices were made to maximize the overlap between field data and SAR data. The plots represented single, double, and multiple species stands. The predicted backscattering coefficients for Tara Downs were compared to fully polarimetric ALOS PALSAR data acquired in 2007, and the ICLP simulations were compared to AIRSAR data from 2000.

The correspondence between field data and simulated SAR data at HH and VV polarizations for an example of woody vegetation at L-band is shown in Figure 9. The errors seen in the figure result from (1) modeling errors, (2) parameterization errors, and (3) radar data absolute and relative calibration errors. The modeling errors, item (1), could be due to the low-order (distorted Born) scattering assumption for scattering within vegetation canopy as well as the first-order surface model used. Both of these approximations will be examined and replaced with higher-order scattering analyses. A full-wave surface scattering solution has already been developed by the team, which will replace the first-order model used in the plots. A full-wave scattering solution for random collection of cylinders has also been developed, which is currently under evaluation and will likely replace the distorted Born solution. Parameterization errors, item (2), are due to errors in ground truth collection and more importantly, due to lack of certain components of ground truth and in particular soil moisture. For most of the points shown in the figure, no in-situ soil moisture measurements were available simultaneously with the radar data acquisitions. Instead, we used gross estimates provided by the ground teams from several days before or after the radar flights. Soil roughness measurements were also not available, and the roughness values were estimated from various photographs. This deficiency highlights the need for well coordinated field campaigns for woody vegetation, on par with those for crops and grasslands. Regarding radar calibration, item (3), AIRSAR co-polarized data were believed to have 1.5 dB absolute calibration and 1 dB relative calibration accuracies. PALSAR likely had better calibration performance, with relative calibration as good as 0.5 dB. UAVSAR has absolute calibration accuracy of 0.5-0.7dB, and is recommended for near-future field campaigns along with accurate and well planned ground data collection.

2.4 Adaptation of Forward Models for SMAP

SMAP has three L-band measurement channels (HH, VV, HV) from which soil moisture can be estimated. The estimation of soil moisture is an inversion of forward models. The estimation becomes challenging because there are many parameters in the forward models that are unknown (or uncertain) a priori, and must therefore be estimated as part of the retrieval, or be provided through ancillary data. With three measurement channels the most free parameters that can be estimated is three, therefore the forward models must be parameterized in terms of at most three dominant parameters in a simplified forward model framework. For SMAP, the three parameters are dielectric constant (equivalently, soil moisture), soil surface roughness, and vegetation water content. The impact of vegetation structure is accounted for by development of separate models for each of the landcover classes shown in Table 4.

The complexity of the non-linear forward models precludes a closed-form analytical solution. Direct numerical inversion of a complicated forward model is also not feasible for global soil moisture retrieval at 3-km resolution. However, a lookup table representation of a complicated forward model was demonstrated to be an accurate and fast approach for retrieval (Kim et al. 2012b). Because the SMAP radar has three measurement channels, and three primary retrievables (dielectric constant, surface roughness, and vegetation water content), the forward models can be pre-computed for a given land surface class, for discrete values covering the full range of the three parameters, and represented as a "data cube" (Kim et al. 2012a; van Zyl 2011). Merits of the data cube search are summarized below:

- to avoid numerical or analytical inversion that is often not feasible for a sophisticated forward model (Functional fits to the forest scattering model may also be used for the fast search of the soil moisture solution (Moghaddam et al. 2000))
- to achieve the same inversion accuracy as the numerical or analytical inversion by adopting a fine interval for the data cube axis, as demonstrated in (van Zyl 2011)
- to conveniently replace and update a forward model while retaining the same retrieval formulae.

One vegetation axis represents the vegetation structure and dielectric properties by an allometric

equation relating volume and weight of the vegetation. The representation of vegetation effects is clearly a simplification, considering that numerous vegetation parameters produce different backscatter coefficients. However, the axis dimension cannot be too large; and for "snapshot" retrieval options the dimension cannot exceed the number of independent observations (e.g., HH, VV, HV). This simplification will result in errors in the soil moisture retrieval (discussed subsequently).

Only the real part of the complex relative permittivity (ε_r) is used, to keep the number of unknowns to 3. It is assumed that the imaginary part of the dielectric constant can be directly related to the real part for specific soil texture information. The results remain dependent on both real and imaginary parts of the permittivity, particularly for cases when the real part of the soil dielectric constant is less than 5.

The data cube for the grass surface developed by the Univ. Washington is shown below as an example. The data cube has the NMM3D simulation for the surface condition the Body of Revolution model for vegetated surfaces. Figure 10 shows the data cubes for the co-pol σ^0. Observed values of *ks* (*k* is the wave number, 26.38) range up to ~1 (equivalently, 3.8 cm RMS height, (Oh et al. 1992)). Values of m_v reach 0.4 cm^3/cm^3 (Dubois et al. 1995); the corresponding value of the dielectric constant is about 30 (real part). The pasture vegetation has the VWC smaller than 3 kg/m^2 in most cases. Accordingly the range of each axis is determined. In general, the co-pol σ^0 increases with *s*, ε_r, and VWC. Also σ^0_{HH} is typically smaller than σ^0_{VV} for low-vegetation surfaces. These are clearly shown by the data cubes in Figure 10.

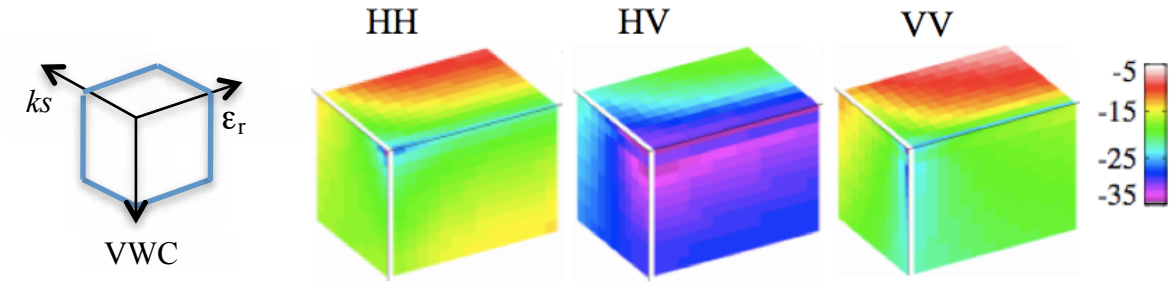

Figure 10. Data cubes of σ^0 simulated for grass surface in dB unit. The ranges of the axes are 0.03 to 1.1, 3 to 30, and 0 to 3 kg/m^2 for *ks*, ε_r, and VWC, respectively.

To address the challenging task of the applicability to global surfaces, the baseline time-series algorithm will utilize data cubes specific to the 16 classes of Table 4. Because crop classes cover 11% of the global land, and the scattering from various crops is not the same, the original 12 IGBP classes are expanded with four data cubes a to represent dominant crops in terms of their spatial extent (http://**faostat**.fao.org). Initial versions of all of the 16 data cubes are currently available, many of which have been validated using existing in situ observations. Because "forward model" error is a key concern for SMAP radar soil moisture retrievals, validation of the data cubes continues and will continue throughout the cal/val period. The backscatter coefficients of datacubes do not include the Kp noise. Application of the noise-free datacubes to the noisy SMAP data will result in retrieval errors and the errors are included in the retrieval error budget.

3 RETRIEVAL ALGORITHMS

The baseline retrieval algorithm inverts a forward scattering model. The equation for forward scattering (Eq. 2-1) presents four unknowns: RMS height (s), dielectric constant (ε), correlation length (l), and vegetation water content (VWC). It will be shown that m_v retrievals are insensitive to errors in knowledge of the correlation length over a wide range of m_v, roughness, and correlation lengths (Appendix 5.2) (Kim et al. 2012b). An exponential function is known to describe empirical measurements well (Mattia et al. 1997; Shi et al. 1997), which was used in the forward model and was not considered as an unknown during the retrieval.

As an introduction to the discussion of the retrieval algorithm, and to place the discussion of each algorithm component in the proper context, the overall algorithm processing flow is described in the following section. Each portion of the algorithm flow, and the details of the retrieval methods and options, will then be described in detail in the subsequent sections.

3.1 Algorithm Flow

The algorithm flow of the L2_SM_A is presented in Figure 11. The portion of the flowchart through the land surface classification is the initial processing done as a precursor to the actual retrievals. The L2_SM_A processor reads in 1-km resolution σ^0 from the SMAP L1C_S0. The 1-km data in natural unit are aggregated onto the 3-km EASE grid, during which various quality flags are applied. Three quality flags are derived using the 3-km σ^0: freeze-thaw state (see L3_FT_A ATBD), radar vegetation index (Section 3.4.2), and transient water body (Section 3.4.3). Static and dynamic ancillary data are sampled for each pixel of the L2_SM_A product. σ^0 values from the past time are sampled and used by the time-series algorithm. For each 3-km pixel, land cover class information is obtained from the mostly annual ancillary data. The land cover information allows to choose an appropriate data cube for each pixel. The retrieval over the different land cover classes is spatially assembled to create a half-orbit output, followed by the conversion from the dielectric constant to soil moisture.

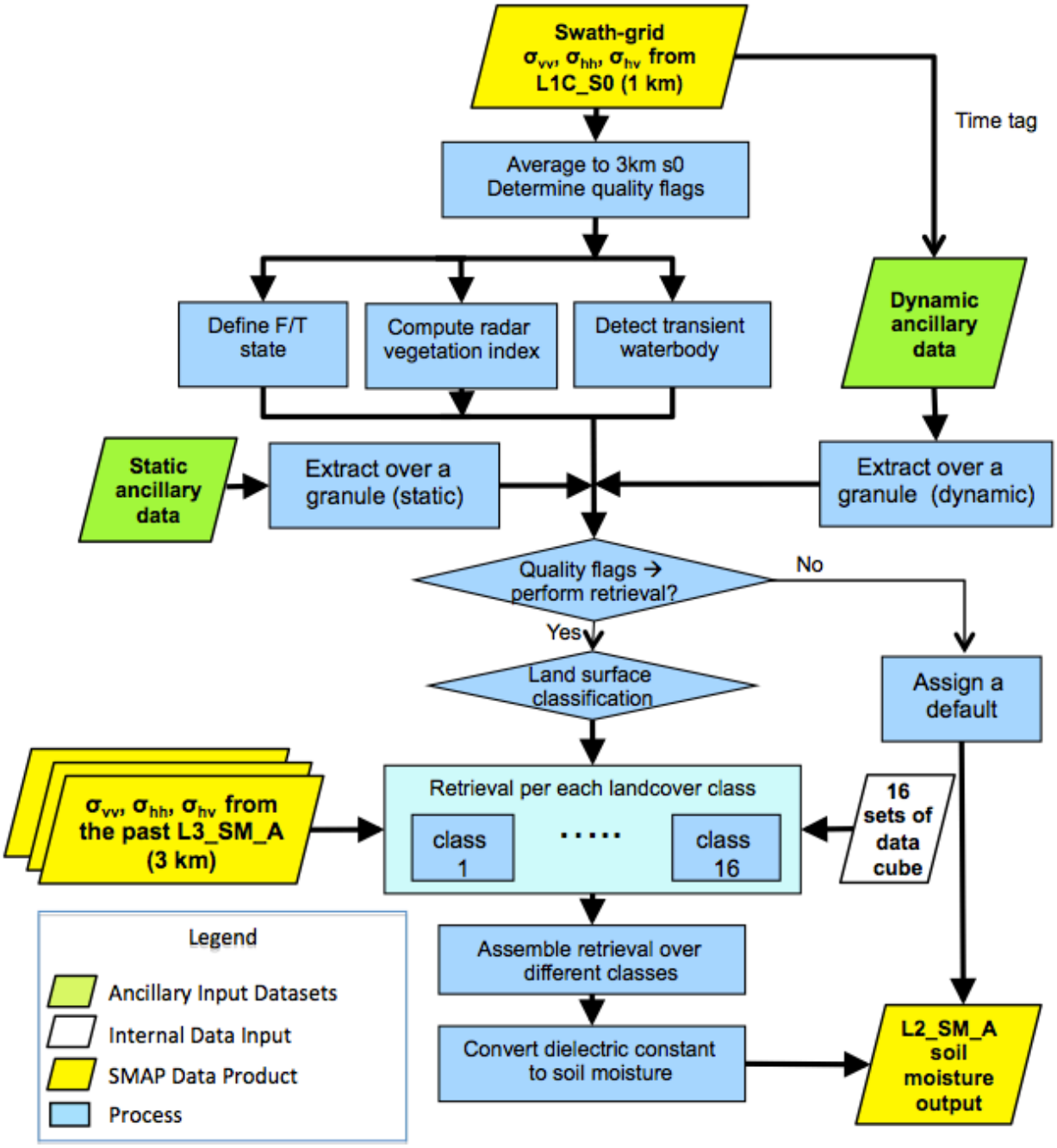

Figure 11. L2_SM_A algorithm flow.

3.2 Baseline Algorithm: Time-series Data-cube Approach

3.2.1 Time-series Retrieval

The SMAP radar provides three independent channels (HH, VV, and HV). HV-channel measurements are reserved for possible use in correcting vegetation effects. The remaining two co-pol measurements (HH and VV) are not always sufficient to determine *s* and ε (Section 3.2.4) (Kim et al. 2012b). One of the main causes is the ambiguity in bare surface scattering: a wet and smooth surface may have the same backscatter as a dry and moderately rough surface has (Figure 3). Very often the time-scale of the change in *s* is longer than that of ε (Jackson et al. 1997). Then *s* may be constrained to be a constant in time, thus resolving the ambiguity (Kim et al. 2012b). The concept of a time-invariant *s* has also been utilized in other studies (Joseph et al. 2008; Mattia et al. 2009; Verhoest et al. 2007). The SMAP baseline algorithm differs from these studies in that no ancillary or ground measurements or statistical assumptions are required to constrain *s*.

The SMAP baseline approach (Kim et al. 2012a; Kim et al. 2012b) is a multichannel retrieval algorithm that searches for a soil moisture solution such that the difference between computed and observed backscatter is minimized in the least squares sense. The algorithm estimates *s* first and then retrieves ε_r using the estimated *s*. Vegetation effects are included by selecting the forward model's σ^0 at the VWC level given by an ancillary source or the SMAP HV measurements. The algorithm retrieves *s* and real part of dielectric constant (ε_r) using a time series of *N* co-pol backscatter measurements: $\sigma^0_{HH}(t_1)$, $\sigma^0_{VV}(t_1)$, $\sigma^0_{HH}(t_2)$, $\sigma^0_{VV}(t_2)$,…, $\sigma^0_{HH}(t_N)$, and $\sigma^0_{VV}(t_N)$. There are thus $2N$ independent input observations and N+1 unknowns consisting of N ε_r values and one *s* value. Note that the VWC provided by ancillary information is allowed to be varying throughout the time series.

Radar backscattering coefficients before conversion to decibels can be modeled as Gaussian random variables (Ulaby et al. 1986) to account for speckle and thermal noise effects. Assuming sufficient integration following power detection, the backscattering coefficient after conversion to decibels can also be modeled as a Gaussian random variable. Because SMAP will observe HH and VV returns at slightly different center frequencies, the effects of speckle and thermal noise on these measurements are statistically independent. Statistical independence of speckle in measurements at differing time steps is also expected. These facts and a maximum likelihood formulation motivate least-squares retrieval approaches based on the average of individual error terms. It is noted that calibration, radio frequency interference, and other error sources may produce correlated error terms. The systematic and correlated components from these sources will be corrected. Any residuals may impact overall algorithm performance and are modeled as uncorrelated Gaussian noise in this ATBD (although they may still contain correlated noise).

The retrieval algorithm therefore minimizes the cost function (*C*):

$$
\begin{aligned}
&C\left(s,n,\varepsilon_{r1},\varepsilon_{r2},..,\varepsilon_{rN}\right)\\
&=w_{1,HH}\left(\sigma^0_{HH}\left(t_1\right)-\sigma^0_{HH,fwd}\left(s,n,\varepsilon_{r1}\right)\right)^2+w_{1,VV}\left(\sigma^0_{VV}\left(t_1\right)-\sigma^0_{VV,fwd}\left(s,n,\varepsilon_{r1}\right)\right)^2\\
&+w_{1,HH}\left(\sigma^0_{HH}\left(t_2\right)-\sigma^0_{HH,fwd}\left(s,n,\varepsilon_{r2}\right)\right)^2+w_{1,VV}\left(\sigma^0_{VV}\left(t_2\right)-\sigma^0_{VV,fwd}\left(s,n,\varepsilon_{r2}\right)\right)^2\\
&+..\\
&+w_{1,HH}\left(\sigma^0_{HH}\left(t_N\right)-\sigma^0_{HH,fwd}\left(s,n,\varepsilon_{rN}\right)\right)^2+w_{1,VV}\left(\sigma^0_{VV}\left(t_N\right)-\sigma^0_{VV,fwd}\left(s,n,\varepsilon_{rN}\right)\right)^2\\
&=\frac{1}{N}\begin{bmatrix}E_1\left(\sigma^0_{HH}\left(t_1\right),\sigma^0_{VV}\left(t_1\right),s,n,\varepsilon_{r1}\right)+E_2\left(\sigma^0_{HH}\left(t_2\right),\sigma^0_{VV}\left(t_2\right),s,n,\varepsilon_{r2}\right)+..\\ E_N\left(\sigma^0_{HH}\left(t_N\right),\sigma^0_{VV}\left(t_N\right),s,n,\varepsilon_{rN}\right)\end{bmatrix}
\end{aligned}
\qquad (3.2.1\text{-}1)
$$

where values from observations and from the forward model are denoted as σ^0 and σ^0_{fwd} (both in dB), respectively. An additional parameter n is included above (so that a total of N+2 parameters are now involved) to represent the number density of vegetation scatterers. This parameter can be viewed as accounting for bias discrepancies between data cube simulations and actual vegetation properties. Again note that the above formulation can accommodate the temporal change in VWC, because σ^0_{fwd} is chosen by the VWC value available at each time. The weights w_i in the cost function is associated with differing errors in the cross section measurement as a function of time or polarization, and the subscript i indexes the time sequence. Uniform weights are utilized in the results to be shown. The cost function then is an average of terms E_i that depend individually on only three (e.g., s, n, and ε_{r1}) of the N+2 parameters. This allows the minimum of the cost function to be located without a search over the complete N+2 dimensional space.

Because s and n are the only parameters in common among the terms of the cost function, the retrieval algorithm considers all possible values of s and n. For each s value, values of ε_{ri} are found that minimize each E_i term individually (i.e., N one dimensional searches are performed), and the resulting average of E_i values is the minimum value of the cost function for the assumed value of s, notated $C_{min}(s)$. The retrieved estimate of s_{rtr} is determined as the value of s that minimizes $C_{min}(s)$, and retrieved estimates of soil moisture are the corresponding ε_{ri} values determined when constructing $C_{min}(s_{rtr})$. Because σ^0 is a monotonic function of s, the minimum is unique with respect to s. σ^0 is also a monotonic function with respect to ε_r. Therefore the minimum associated with ε_{ri} is unique for a given s. In summary, the proposed method is a two-step approach. In the first step, the temporal average of the individual error term, E_i, reduces the Gaussian random noise and improves the retrieval of s. Subsequently the ε_r retrieval benefits from the estimate of s. The vegetation density parameter n is not a direct retrieval parameter, but rather is included to represent potential physical sources of any biases that may remain (i.e. non-zero values of the Cost function) at the retrieved rms height and permittivity values.

3.2.2 Data-cube Search

The least-square minimization of Section 3.2.1 is implemented using the forward model lookup table (generated using the models described in Section 2). The lookup table has three input parameters (a 'data cube') required to predict SMAP's HH, VV, and HV radar observations. The data-cube input parameters are the RMS height of bare soil roughness (s), real part of dielectric constant (ε_r), and the ancillary or HV determined vegetation water content (VWC). Either dielectric constant or volumetric soil moisture (m_v) may characterize the bare soil scattering, because they are interchangeable through a dielectric model. The dielectric constant is a more practical choice because otherwise soil parameters other than m_v, such as texture, force additional dimensions of the data cubes. The current baseline

retrieval is fairly independent of the correlation length (Appendix 5.2) that therefore is not a part of the data cube.

3.2.3 Algorithm Evaluation

Monte-Carlo simulation of 0.5 dB radar measurement noise was performed to evaluate the performance of the time-series algorithm and the results are presented in Figure 12. According to the bare surface analysis in Figure 12a, the retrieval error is mostly smaller than 0.06 cm^3/cm^3 and smaller than 0.04 cm^3/cm^3 in more than half of the cases examined. A challenging problem in the bare surface retrieval is to resolve the ambiguities between soil moisture and roughness effects: the backscatter of the dry-rough surface is the same as that of the wet-smooth surface (Figure 3). The ambiguity is resolved by constraining the retrieval using the estimate of surface roughness.

The Monte-Carlo simulation of the K_p noise in the presence of vegetation is also performed (in Figure 12b). The retrieval error is better than 0.06 cm^3/cm^3 mostly for VWC up to 2 kg/m^2 and m_v up to 0.4 cm^3/cm^3. Generally the m_v retrieval error increases with VWC, as a result of the vegetation effect. Both of the Monte-Carlo simulation results were made with 6 time-series records. This time-length corresponds to 18 days considering SMAP's 3-day repeat cycle. The VWC error is set to 20% (1σ, zero-mean) based on a preliminary analysis of 1.5 dB (1σ, zero-mean) error in the SMAP HV observation (see Section 5.3).

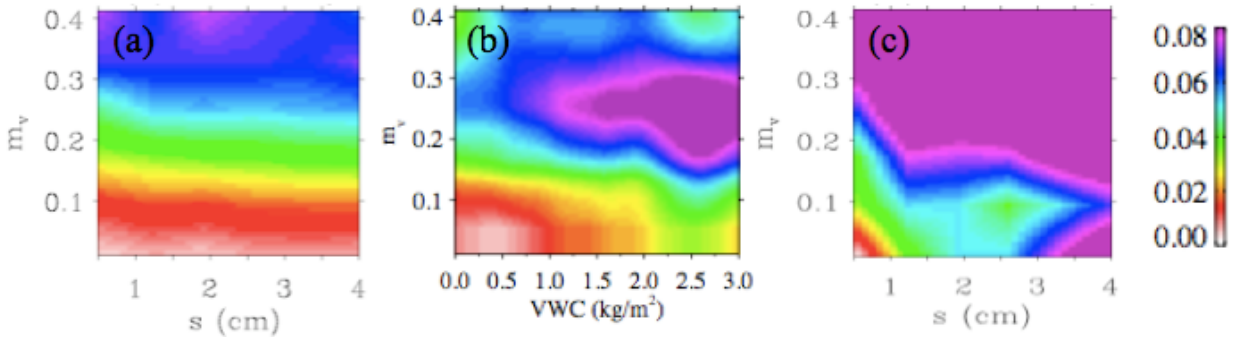

Figure 12. RMS errors in cm^3/cm^3 from the Monte-Carlo simulation of the K_p effect on the data-cube m_v retrieval. (a) time-series retrieval for bare surface, (b) time-series retrieval for pasture surface, (c) snap-shot retrieval for bare surface: all three cases represent *s* of ~1 cm and K_p of 13% for the co-pol signal (0.5dB). The correlation length to the RMS height has the ratio of 10. The number of time series records used during the retrieval is 6 (=18 days).

The time-series retrieval algorithm was further tested with in situ field data. The first data were measured by the truck-mounted scatterometer at 1.25 GHz at 40° incidence angle over four bare surface sites near Ypsilanti, Michigan during a two-month period with up to 11 temporal samples per site (Oh et al. 2002). The second data were collected by the airborne Passive/Active L-band Sensor (PALS) instrument at 1.26 GHz at the incidence angle centered at 38° during the 1999 Southern Great Plains (SGP99) experiment over the Little Washita Watershed near Chickasha, Oklahoma (Jackson and Hsu 2001; Njoku et al. 2002). The VWC varies from site to site between 0.16 and 2.5 kg/m^2. Temporally the roughness, correlation length, and VWC remain constant. Both data sets offer additional in situ measurements of soil texture and soil temperature.

Figure 13a shows that the bare surface retrieval has an RMSE of 0.044 cm^3/cm^3 after compiling the retrievals over the 4 in situ locations. The time-series retrieval over the pasture surface has an RMSE of 0.054 cm^3/cm^3 (Figure 13b). The slightly larger error for the pasture surface may reflect the effect of vegetation.

For each field of the SGP99, a constant offset in terms of dB was estimated to minimize the difference between observed backscatter and predicted backscatter using the time series of co-pol backscatter observations. This offset in dB may be understood as an error in the scatterer number density, *n*. The error would appear in the data cube as nearly a constant in dB. A change in *n* results in time-invariant offsets in dB for the volume scattering and the attenuation of the surface scattering (this may not necessarily apply to the double-bounce term if the bare surface condition changes in time.) To the extent that the double-bounce component is small, and also that the dB offsets in the individual scattering components become a dB offset in the total scattering power, the correction for *n* can be achieved by applying the dB offset. The estimated offset values are -2.6, 1.9, 5.2, 0.5, 2.5, and -2.5 dB across the 6 fields of the SGP99 (3 dB offset would roughly mean a doubling of *n*). These values are comparable with the bias difference between the data cube prediction and the backscatter observations: -2.5, 1.1, 1.9, -1.2, 2.5, and -1.2 dB for the 6 fields. The close agreement between the estimated offset and the bias difference suggests that the interpretation of data cube biases in terms of variations in vegetation density is reasonable.

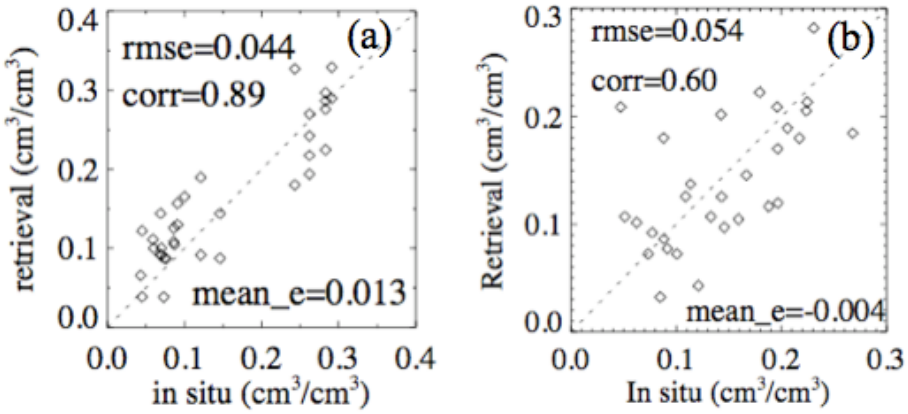

Figure 13. RMS errors in soil moisture retrieval using the data-cube time-series method, assessed with (a) the Yipsilanti bare surface data and (b) the SGP99 pasture surface data.

3.2.4 Snapshot Retrieval

When the number of time-series input is set to 1, the time-series method becomes a conventional snapshot approach. ε_r and *s* are adjusted simultaneously at each time instance so as to minimize the distance between the data cube prediction and the observation. The snapshot retrieval algorithm was also tested with the two sets of field campaign data used in Section 3.2.3. Figure 14a shows that the bare surface retrieval has an RMSE of 0.055 cm^3/cm^3, which is larger than the error of the time-series retrieval. Closer examination however indicates that the retrieval error increases for larger soil moisture. When the performance of the Monte-Carlo simulation of the K_p effect is compared between time-series (Figure 12a) and snapshot (Figure 12c) retrievals, the performance difference is much greater than the in situ validation shows. The radar measurement error of the in situ data (~0.3 dB) is smaller than simulated by the Monte-Carlo analysis, which may explain why the Monte-Carlo analysis of the snapshot retrieval show the poorer performance.

The snapshot retrieval over the pasture surface is poor. Figure 14b shows a retrieval error of 0.105 cm^3/cm^3. It is possible that the structure of the grass was important in characterizing the backscatter to be very different from that of the bare surface. The possibility is supported also by the large retrieval error (not shown) when the retrieval was performed using the bare surface data cube (NMM3D) after treating the grass surface as if it is a bare surface. The evaluation results suggest the limited applicability of the snapshot retrieval, and highlights the need for time-series retrievals.

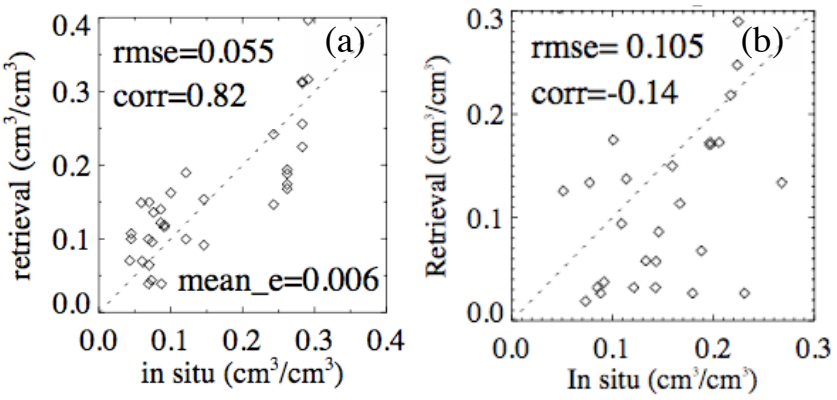

Figure 14. In situ validation of the soil moisture retrieval error (RMS) using the data-cube snapshot method, assessed with (a) the Yipsilanti bare surface data and (b) the SGP99 pasture surface data.

3.3 Optional Algorithms

3.3.1 Change-detection by Kim and van Zyl

The full details of this algorithm are described by (Kim and van Zyl 2009). A brief summary is presented here. After the SMAP radar data are accumulated for a moderate time period (nominally six months), an expression will be derived to relate the backscattering cross section to soil moisture as

$$m_v = f(\sigma_{hh}, \sigma_{vv}) \tag{3.3.1-1}$$

where the normalized radar cross sections are expressed in decibels. Since this expression depends on the biomass level, the cross-polarization will be used to compensate the biomass variation over time: how this will be implemented is being studied. One potential choice for f is given by

$$f(\sigma_{hh}, \sigma_{vv}) = C_0 + C_1 \frac{\sigma_{hh}+\sigma_{vv}}{2} \tag{3.3.1-2}$$

These two coefficients for each pixel will be determined using the expected minimum and maximum values for soil moisture and the time-series backscattering cross section data. Accommodating different vegetation phenology during the implementation of the above relationship remains as a task. The obvious assumption contained in Eq. 3.3.1-2 is that there is a linear relationship between the radar cross-sections and the soil moisture. However, simulated data suggests that this relationship may be nonlinear. Ground radar and airborne data will be used to derive the final expression for the function $f(\sigma_{hh}, \sigma_{vv})$. The slowly varying backscattering cross section component due to the biomass variation will be estimated and compensated using the cross-polarization backscattering cross section. Ground radar measurements will help derive the expression before the SMAP instrument is launched. However, since the ground measurements are limited, with simulated data various vegetation environments will be considered. When the two extreme conditions (completely dry and wet) of the soil moisture and radar backscatter are known exactly, a test with a corn field experiment shows that the soil moisture may be retrieved with an error of 0.026 cm^3/cm^3 (Table 3 of Kim and van Zyl). When the errors in the two extreme values are simulated with the random noise, the retrieval error reaches 0.05 cm^3/cm^3. Alternatively, Kim and van Zyl proposed that C_1 at each radar pixel (e.g., at 3 km resolution) is inverted using the m_v retrieved at a coarse scale (e.g., at 36 km resolution) using the radiometer records. The radar-radiometer method is the subject of the SMAP Level 2&3 Active/Passive ATBD.

3.3.2 Change detection by Wagner et al.

A radar-only time-series algorithm (Wagner et al. 1999) was proposed to retrieve an index of m_v change using the C-band ERS scatterometer data. The index (M_s) is given by

$$M_S = \left(\sigma^0(t) - \sigma^0_{dry}\right) / \left(\sigma^0_{wet} - \sigma^0_{dry}\right), \tag{3.3.2-1}$$

where $\sigma^0(t)$ is the observation at one time. σ^0_{wet} and σ^0_{dry} are two extreme values of σ^0 of a pixel, which may be derived from a multi-year database of radar records. At least one year worth of the SMAP data are needed to derive the two extreme values of σ^0. *Ms* is an index ranging from 0 to 1, and can distinguish 5 different levels of soil moisture states. Later studies retrieved the index varying continuously, e.g., (Naeimi et al. 2009). To obtain the absolute value of soil moisture, two extreme sets of moisture and radar backscatter are needed (Kim and van Zyl 2009), or the index may be assimilated into a land surface numerical model (Scipal et al. 2008). Most of the studies were applied to the low-speckle scatterometer data with the spatial resolution of 25 km or larger. When applied to the 1-km resolution data from the ENVISAT Advanced SAR (ASAR), the speckle began to impact the performance even at 3-km resolution; 3-10 km resolution was recommended for reliable retrieval of the index (Pathe et al. 2009).

3.3.3 Snapshot Algorithm for Bare Surfaces by Sun et al.

A new algorithm to estimate bare surface soil moisture using dual-polarization L-band backscattering measurements is proposed in (Sun et al. 2009). The retrieval over the vegetation surface based on the bare surface algorithm is under development. This algorithm describes the backscattering coefficients as a product of two functions.

$$\sigma_{pq}(\theta) = R_{pq}(\varepsilon_s, \theta) \cdot Sr_{pq}(\theta, ks, kl) \tag{3.3.3-1}$$

One is the dielectric function R_{pq} that reflects the soil moisture information. It can be approximated by the polarization magnitudes as described by the Small Perturbation Model at L-band. The other is the roughness function Sr_{pq} that describes the overall effects of the surface roughness at the different polarization. The subscript *pq* represents for the polarization status. Inversion requires the acknowledgement of the relationship between the surface roughness parameters Sr_{hh} and Sr_{vv} at different polarizations with a good accuracy so that they can be considered as one unknown.

Through analyses on an AIEM surface backscattering database with a wide range of soil moisture and surface roughness conditions, it was found that the relative difference of the overall surface roughness parameters between co-polarizations can be well estimated through a roughness index *Sr*.

$$Sr_{hh}^{C_1} - Sr_{vvhh}^{C_2} = \exp\left[E + F \cdot \log(Sr) + G \cdot Sr\right] \tag{3.3.3-2}$$

where $S_r = A \cdot \dfrac{\sigma^B_{vvhh}}{\sigma_{hh}}$, where $\sigma_{vvhh} = \sqrt{\sigma_{hh}\sigma_{vv}}$, is a roughness index that only related to the surface roughness properties but not the dielectric properties. Equation (3.3.3-2) provides a direct estimation of the differences in the roughness parameters at the certain powers from the radar observations. From this information, we now can directly estimate surface dielectric properties by

$$Sr_{hh}^{C_1} - Sr_{vvhh}^{C_2} = \left(\frac{\sigma_{hh}}{R_{hh}}\right)^{C_1} - \left(\frac{\sigma_{vvhh}}{R_{vvhh}}\right)^{C_2} \tag{3.3.3-3}$$

The left side of Equation (3.3.3-3) is the surface roughness differences and can be estimated directly from the dual co-polarization radar measurements. The right side of Equation (3.3.3-3) has only one unknown: the dielectric properties of soil through the dielectric function in Equation (3.3.3-1). It can be estimated numerically. The advantage of this algorithm is that the estimation of surface soil moisture is based on the roughness relation at the different polarization that can be directly estimated from the radar measurements.

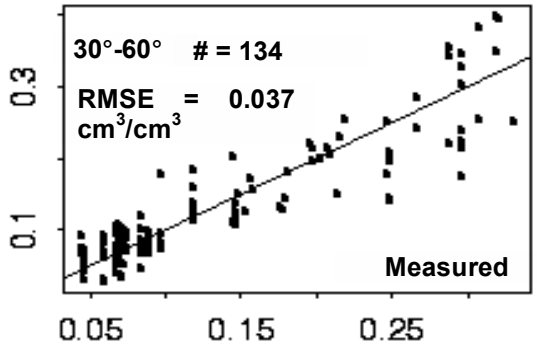

Figure 15. The comparisons between the estimated volumetric soil moisture (y-axis) from two validation data sets and the ground measurements (x-axis).

This algorithm is validated with two sets of ground scatterometer experimental data. Figure 15 shows the comparison of the ground measured soil moisture with that the estimated surface soil moisture using the above two experimental data sets for incidences from 30º to 60º. The inversion results at 20º are very poor with RMSE over 0.06. This is mainly due to the very small polarization differences in the measurements and some measurements with HH larger than VV, which can not be predicted by the AIEM model. The inversion results in the range of 30 º to 60 º incidence, however, have an RMSE of 0.037 over a total of 134 observations. It indicates that bare surface soil moisture may be estimated with SMAP instrument.

Table 5 Input data for the L2_SM_A algorithm.

Data Type	Data Source
Radar backscatter (HH/VV/HV)	SMAP L1C_S0
Elevation and standard deviation of DEM	combination of SRTM, Alaska USGS DEM, Canada Geobase DEM, and GTOPO30
Land cover class	MODIS-IGBP
Crop type	Cropland Data Layer (US) Environment Canada (Canada) ECOCLIMAP (Europe) (Monfreda et al. 2008) (Rest of the world)

Vegetation water content	SMAP L2_SM_A MODIS NDVI
Data cube	SMAP Science Team
Soil texture	Harmonized World Soil Database + regional + FAO
Soil surface temperature	GMAO
Mountain Area	combination of SRTM, Alaska USGS DEM, Canada Geobase DEM, and GTOPO30
Freeze/thaw flag	SMAP L2_SM_A
Permanent ice	MODIS IGBP
Snow	Snow & Ice Mapping System (IMS) - NOAA
Static water body	MODIS MOD44W
Transient water body	SMAP L2_SM_A
Precipitation	ECMWF total precipitation forecasts (or GPM once launched)
Urban area	Global Rural-Urban Mapping Project (GRUMP)
RFI contamination	SMAP L1C S0

3.4 Input Data and Output Parameters

Table 5 lists the required input data sets to the radar-only soil moisture inversion. These includes the primary choices that meet the ancillary data requirements. 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 3-km 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, precipitation, and effective soil temperatures. Although most ancillary data are by definition external to SMAP, the L2_SM_A will provide key pieces of information to the L2_SM_P and L2_SM_AP algorithms including the open water fraction (Section 3.4.3) and the frozen ground flag (see L3_FT_A ATBD).

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, and its inherent error and resulting impact on the overall soil moisture or freeze/thaw retrieval accuracy. Latency, spatial resolution, temporal resolution, and global coverage are also important. The choice of a primary source for each of the ancillary data parameters is fully documented in individual SMAP Ancillary Data Reports (Section 6.1) which will be

made available to the user community. Below brief descriptions are provided for the input radar backscatter, and for the ancillary parameters generated by the L2_SM_A processor.

3.4.1 Radar Backscatter

The primary input to the L2_SM_A algorithm will be the Level 1C (L1C) gridded radar product. This product consists of the calibrated high-resolution (SAR-mode) multiple polarization radar backscatter measurements and associated parameters, Earth-located and mapped to a swath-based, along-track/cross-track, 1 km grid. It is assumed that the backscatter measurements have been fully calibrated and corrected for atmospheric attenuation and ionospheric effects such as Faraday rotation. Information provided in the L1C product or in ancillary calibration files should enable estimates of the σ^0 accuracy (K_p) for each measurement based on instrument characteristics.

3.4.2 Vegetation Water Content derived from Radar Vegetation Index

The SMAP L2_SM_A/L2_SM_P/L2_SM_AP algorithms require information on the VWC. One potential source of vegetation information is a polarimetric parameter derived from SMAP data such as the RVI (Radar Vegetation Index) (Kim and van Zyl 2000) defined by:

$$RVI = \frac{8\sigma_{hv}}{\sigma_{hh} + \sigma_{vv} + 2\sigma_{hv}} \qquad (3.4.2\text{-}1)$$

The RVI nominally ranges between 0 and 1, and is a measure of the randomness of the scattering. For a smooth bare surface, the RVI is near zero, while for surfaces with a dominant volume scattering component the value approaches 1. This parameter is very similar to the polarimetric entropy, but does not require the full covariance matrix to calculate. A linear relationship between RVI and VWC was reported from the 14-day long aircraft measurement over corn and wheat fields in Oklahoma in 2007 (Yueh et al. 2008). On the other hand, the relationship was strongly non-linear during corn's entire growth cycle over a 20-week period (Kim and van Zyl 2009). A further analysis is needed for RVI to be an operational source of VWC.

Alternatively, the vegetation amount may be derived based on low-order polynomial of the normalized difference vegetation index (NDVI) and optical remote-sensing images (see L2_SM_P ATBD). In principle the RVI or any other parameter based on the radar backscatter should correlate with the vegetation amount more strongly than the optical NDVI, because fundamentally optical radiation cannot penetrate vegetation canopies while the L-band microwave signal can. Furthermore the RVI is concurrent with the radar observation. Studies are ongoing with regard to whether the radar backscatter may improve estimates of vegetation effects.

3.4.3 Transient Water Bodies

An algorithm for discriminating transient water bodies from land surfaces is also part of the SMAP L2_SM_A processor. There are many combinations of water surfaces adjoined by land surfaces, (depending on both land and water surface roughness and presence of vegetation) that will affect this discrimination capability. At present, the case of a calm water surface next to lightly vegetated land was studied. It was found that a threshold (3 dB) applied to the ratio of HH to VV effectively separated water from land at SMAP's 40 deg incidence angle. In this study, UAVSAR (Uninhabited Aerial Vehicle Synthetic Aperture Radar) images were analyzed over fairly flat terrain. Despite the apparent differences between UAVSAR and SMAP, the measurement quality for both instruments is determined by the same statistical processes inherent in SAR imaging. Calibration error and the number of single looks of one UAVSAR pixel are similar to those of one pixel of the proposed SMAP SAR. The algorithm was tested over two locations (one of them is shown in Figure 16). The detection error in both

cases is 7% (Kim et al. 2011). Diverse configurations of water and neighboring land exist depending on surface roughness and vegetation cover. A classifier different from the co-pol ratio may be necessary to delineate these types of standing water.

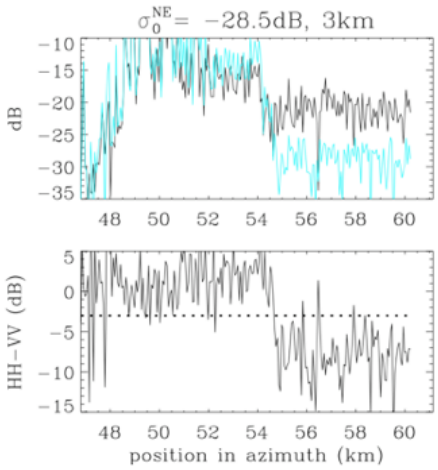

Figure 16 Water body detection tested with a UAVSAR image over the Mono Lake, CA, with the noise level (σ_0^{NE}) of -28.5 dB. The abscissa is the distance along the 40° incidence angle path (for example, the 60-km mark corresponds to the middle of the lake and the 48 km mark is over land). The top panels show the HH (cyan) and VV (black) co-polarized backscatter. The bottom panels show HH-VV, with the threshold of -3dB shown as a dotted line.

3.5 Error Budget

The primary purpose of this section is to verify that the baseline algorithm is able to satisfy the accuracy target of the soil moisture retrieval. The accuracy target for the SMAP's radar-based retrieval is not specified in the science requirement, because this product is a research product. Tentatively the target is set to 0.06 cm^3/cm^3 up to VWC of 5 kg/m^2.

3.5.1 Evaluation with In situ Data

The time-series data-cube retrieval method was tested with the scatterometer data over four bare surface sites near Ypsilanti, Michigan during the two month period (Kim et al. 2012b), and the airborne Passive/Active L-band Sensor (PALS) data collected over pasture fields during the 1999 Southern Great Plains (SGP99) experiment. The number of time-series inputs varies from 6 to 11 for the bare surface data and is fixed to 6 for the pasture data. The retrievals were described in detail in Section 3.2.3, and showed the retrieval errors of 0.044 (bare surface) and 0.054 (pasture) cm^3/cm^3. In both evaluations, the dielectric model error was accounted for: the forward model used the Peplinski model while the retrieval used the Dobson model. Also, included are the errors in the radar measurement by the PALS instrument. Evaluation over different in situ data and for different land cover classes are continuing.

3.5.2 Evaluation with Simulated Data

Monte-Carlo simulations of the K_p radar measurement noise and the VWC error is performed for the time-series algorithm (Kim et al. 2012a). The K_p levels corresponding to the outer edges of the swath (13%, 0.5dB) was used. The VWC error is set to 10% (1σ, zero-mean) based on a preliminary analysis of 1 dB (1σ, zero-mean) error in the SMAP HV observation (after the fore-aft averaging and considering the CBE performance, see Section 5.3). The number of the time-series records is set to 6 (18 days in view of SMAP's 3-day cycle). Each point of the Monte-Carlo simulation results in Figure 17 is an average of the results over the soil moisture truth ranging from 0.05 to 0.45 cm^3/cm^3. The Monte-Carlo study was performed for all available datacubes where m_v retrieval is feasible. For non-woody classes such as grass and soybean, the m_v retrieval error increases with VWC due to the vegetation

attenuation. For woody-vegetation (shrub and woody savanna) the m_v retrieval improves with VWC, as a result of benefiting from the double bounce process. Corn, evergreen needle, and deciduous broadleaf classes also have the dominant vertical stem. Consequently the retrieval benefits from the double bounce mechanism. The two forests classes (evergreen broadleaf and deciduous needle) do not have the vertical stem structure. The vegetation attenuation dominates and m_v retrieval deteriorates with VWC. In summary, the soil moisture retrieval error is found smaller than about 0.05 cm^3/cm^3 except for the two forest classes.

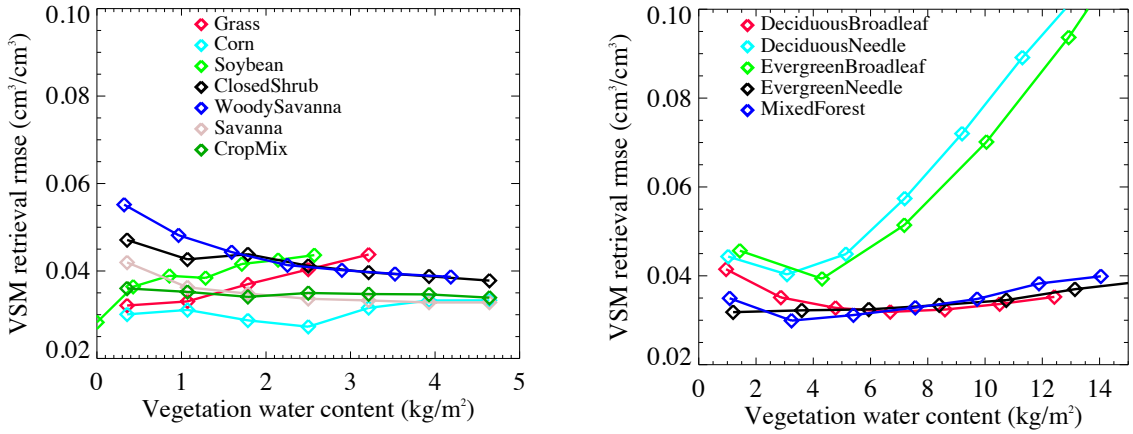

Figure 17. Monte-Carlo simulation of the radar measurement error (13% for the signal, 0.5dB) on the data-cube m_v retrieval. Each point on a curve is an average of m_v RMSE from the Monte-Carlo simulation (e.g., Figure 12), over roughness values of 0.5 cm to 4 cm, and soil moisture truth ranging from 0.05 to 0.45 cm^3/cm^3.

3.5.3 Error Budget

Based on the above results, the error budget is summarized in Table 6. The Monte-Carlo analysis presented in Section 3.5.2 assess the errors due to the K_p noise and the vegetation water content uncertainty. However the analysis does not capture forward model errors. The conversion from dielectric constant to soil moisture is accurate to about 0.02 cm^3/cm^3 (Mironov et al. 2009), which would not affect the total error budget when added to the total error by root-square-sum. The in situ evaluation with the airborne radar data described in Section 3.5.1 accounts for the errors associated with data cube modeling and dielectric conversion (as well as the errors in in situ measurements of soil moisture, both instrumental and scaling). A small degree of heterogeneity exist within the airborne radar footprint, which would also contribute to the in situ evaluation result. The effect of scaling and heterogeneity within the 3-km pixel will be studied further.

Table 6. Error budget for the time-series baseline algorithm. The K_p and VWC errors are modeled with Gaussian random variables with zero mean. The soil moisture retrieval error is obtained from the Monte-Carlo analysis in Figure 17. Number of the pair of HH and VV radar backscatter used as time-series input (6, equivalently 18 days).

Error sources	Budget (cm^3/cm^3)	
	Outer swath edge	Inner swath edge
A) K_p 0.75-1.0 dB 1σ error (1σ, co-pol, fore look)	0.035	0.043
B) Vegetation water content error (1σ, 10%)	0.01	

C) Forward model error (data cubes and heterogeneity)	0.04	
D) Dielectric model uncertainty	0.02	
E) Soil texture: 5% error	0.004	
m_v retrieval error up to VWC of ~ 3 kg/m^2	*0.058*	*0.063*

A,B) the estimate comes from Figure 17, where the effects of the Kp and VWC were combined.
C) the estimate comes from the in situ validation with the SGP99 data (Figure 13b), after subtracting the errors in in situ measurements and the errors in dielectric conversion.
D) The conversion error is about 0.02 cm^3/cm^3 (Mironov et al. 2009)

Refinement of the error budget continues as additional validations of forward models and retrieval performance are conducted.

3.6 Level 3 Processing

The Level 3 product is the daily composite of the Level 2 granules. The current approach for the L3_SM_A product is to use the nearest 6:00 am local solar time criterion to perform Level 3 compositing (see L2_SM_P ATBD for details). All the three Level 2 products (Active, Passive, and Active/Passive) will share the same master processor for the Level 3 products. The L2_SM_A product has a nadir gap where the effective resolution in the crosstrack direction gradually becomes 36 km. A flag will be assigned to identify this region and the region will be excluded during the L3 compositing. Over 3 days on average the gaps will be filled except for the pixels near the equator, where it takes longer to fill the gaps.

3.7 Algorithm Testing & Baseline Selection

The aim of pre-launch activity is to evaluate candidate algorithms, to improve the algorithms, and to select the baseline algorithm. The pre-launch validation of the L2&3_SM_A product retrieval algorithm will be made using (1) global representative simulations on the Algorithm Testbed and (2) field experiment data that offer the radar backscatter and ground truth observations. The implementation of the complete pre-launch Cal/Val process is described in SMAP Calibration and Validation plan (JPL D-52544, Apr. 2011).

3.7.1 Simulated Data in the Algorithm Testbed

Global simulations using ancillary data sources and a spacecraft/instrument simulator will not only reveal performance issues but also processing issues. The ancillary data have to be from the actual source identified by the selected algorithm and meet the specifications of the data that will be used in post-launch processing. The spacecraft/instrument simulator has to provide data compatible with the sampling configuration and performance of the actual spacecraft and instrument.

The simulation data permit testing whether an algorithm can handle with SMAP's sampling that is not well represented by field campaign, such as:

- time gaps
- temporal changes in VWC
- temporal changes in soil surface roughness.

Since the range of surfaces for which measured radar data exist is limited, the output of a reliable radar forward model will be used to establish a database that covers the most likely ranges of the surface soil moisture and roughness properties including RMS height, correlation length, and the forms of the

correlation functions. The various inversion models will then be tested against this database to establish their accuracy, and the ranges of surface parameters over which they are applicable. A similar procedure will be used to simulate vegetation effects by using the data cubes generated for different vegetation classes by radar scattering models.

A realistic global simulation (GloSim2) is currently underway on the SMAP Algorithm Testbed (http://smap-sds-web.jpl.nasa.gov/ confluence/display/algorithm/GloSim+2003). An example of simulated σ_0 is shown in Figure 18 for the HH channel. The simulation covers the world. The simulation is based on the 14 data cubes corresponding to the IGBP scheme and four major crop classes. For the forward simulation, soil moisture and temperature fields were generated by the Goddard Modelling and Assimilation Office (GMAO) Nature Run at 9-km resolution and hourly over 1 year period. VWC was derived from MODIS NDVI data using empirical relationships (see L2_SM_P ATBD). The surface roughness values were generated at 1-km resolution. A mean roughness value was assigned to each of the IGBP and crop classes. The roughness value was then randomly perturbed using a deviation determined from the literature per each IGBP class. The Gaussian random noise with a magnitude of 0.9 dB (1-sigma) is multiplied to both fore or aft look observation of the backscatter, to represent effective K_p noise that includes speckle (0.7 dB), relative calibration error (0.35 dB), and contamination from radio frequency interference (0.4 dB).

Assessments of retrieval performance with the GloSim2 dataset are currently in progress.

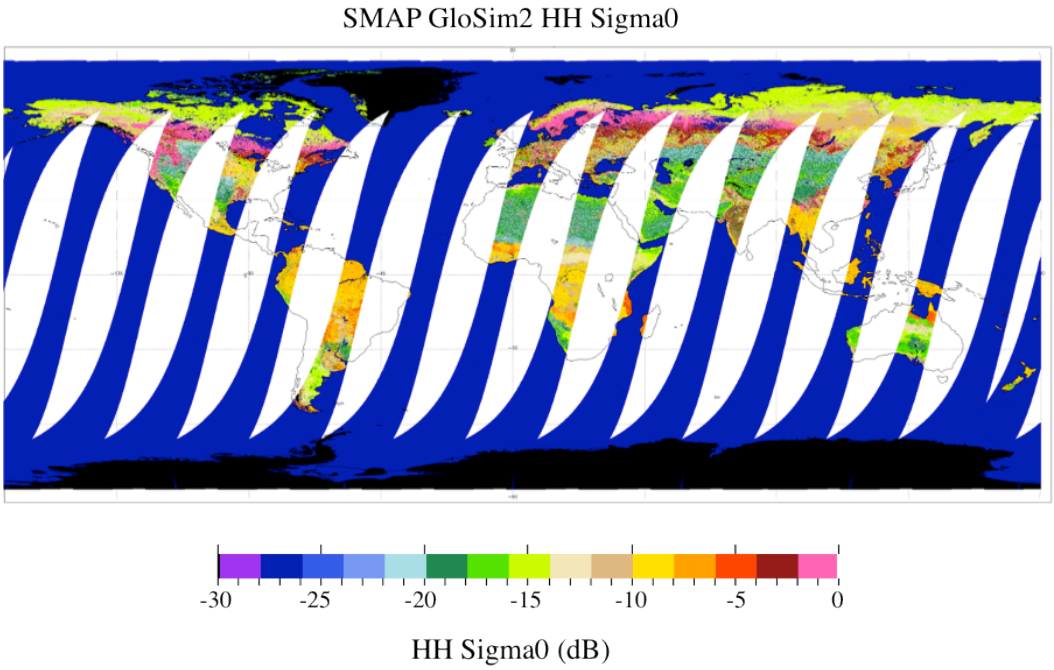

Figure 18. Simulated σ_0^{HH} in dB for a composite of 14 simulated SMAP descending (AM) orbits using radar scattering models.

3.7.2 Field Campaigns

The requirements for field experiment data are two-fold: (1) to provide sufficient knowledge to train the forward model underlying the retrieval algorithm (2) to assess the performance of retrieval algorithms.

To train the forward model, the experimental data have to cover a range of conditions that can be considered to represent the global diversity of surface conditions. Also the time-frame has to be long enough to test the time-series methods (the baseline algorithm so far has been tested with 6 time-series inputs) and to capture the sufficient changes in soil moisture (and desirably vegetation growth). The interval between two consecutive temporal samples is a secondary issue. The following soil surface properties need to be measured:

- Moisture
- Roughness
- Correlation length

Sampling of the following vegetation parameters is desired to characterize the radar scattering model:

- Vegetation class
- Fraction of vegetation cover
- Number of vegetation stem per unit area
- Height of vegetation
- Stem size (diameter)
- Branch, leaves/needles diameter
- Branch, leaves/needles density
- Branch, leaves/needles, trunk volumetric water content
- Dielectric properties for stem, branches, leaves/needles
- VWC from destructive samples and NDVI

Aircraft field experiments or tower-based measurements are suitable for providing the backscatter data to accompany the extensive ground sampling items listed above. Overpass spaceborne L-band SAR measurements (ALOS-2 and SAOCOM) are desired if such can be coordinated.

To test and validate the soil moisture retrieval, the ground sampling requirements for the experimental data are comparatively less demanding. Measurements of soil surface properties, information on vegetation class and VWC would provide sufficient information. Clearly, nonetheless, the full details of the vegetation properties are desired, which will enable further in-depth analysis of the retrieval performance. Overpass spaceborne L-band SAR measurements are desired if such can be coordinated.

The detection of transient water body also require ground truth data. Because the water bodies are transient and often scattered over a large area, the most appropriate source of mapping is satellite or aircraft. Necessary information of ground truths are:

- Vegetation class and descriptive VWC of the adjacent land
- Wind speed over water bodies
- Coordinates of the water body perimeter

As a part of the Cal/Val activities of the SMAP project, the following aircraft/tower campaigns were conducted or are planned that offer the validation capability of the L2_SM_A retrieval:

- SMAPVEX08 (2008, USA).
- CanEx-SM10 (2010, Canada). Extensive vegetation properties were sampled.
- SMAPEx (2010, 2011, Australia). Extensive vegetation properties were sampled. Water body truths were collected.
- San Joaquin Valley Experiment (2010, 2011, USA). Some vegetation properties will be sampled. Water body truths were collected.
- Carbon Arctic Reservoir Vulnerability Experiment (CARVE) flights (2011-2015, USA).
- NASA GSFC ComRAD (Combined Radar/Radiometer System) truck-based instrument (2012, USA). The extensive vegetation properties will be sampled.

- SMAPVEX 12 (2012, Canada). The duration of the campaign will accommodate the testing of the time series approach and the extensive vegetation properties will be sampled.

3.7.3 Baseline Algorithm Selection

The baseline algorithm will be selected through comparison with optional algorithms on an ongoing basis, both before and after launch. The following three criteria will be considered:

- performance in terms of soil moisture retrieval accuracy. The performance of each algorithm will be evaluated by the simulation (GloSim2) and also by tests on observational data.
- feasibility of global soil moisture retrieval with regard to the computing load, the availability of ancillary data, and the applicability to the global surfaces of vegetation and bare soil.
- capability to perform various analysis related to SMAP science, and the design of the mission and the instrument.

3.8 Practical Considerations

This section describes practical issues that would arise during the product generation after launch. The following issues have been identified during the algorithm test with the SMAP Algorithm Testbed. The description pertains to the baseline algorithms, and is subject to change when the baseline algorithm changes.

3.8.1 Numerical Computation Considerations

The numerical computation load is assessed at the time of ATBD writing (September 2012). As the computing hardware advances, the computing load is expected to lessen. As the retrieval algorithms are further refined and economized as the algorithm matures, it will reduce the computing load as well. Preliminary data on processing resources and throughput for the L2_SM_A are acquired by performing the time-series retrieval using Glosim data, and are as follows

Disc storage requirement	60GB for 400 half orbits (about one month worth) of L2_SM_A. The size will increase substantially when all the parameters defined in the data specification document are outputted.
Processing time	25 minutes on average per half orbit

3.8.2 Output Data volume

Currently the 'GloSim2' exercise is being implemented, where the retrieval is performed on the global scale on the basis of the half-orbit granule (http://smap-sds-web.jpl.nasa.gov/confluence/display/algorithm/GloSim+2003.). The global output products occupy about 0.8 terabyte/year and 2 gigabyte/day. The output volume would increase with more output parameters associated with the complete set of parameters related to the ephemeris and geometry; however, the HDF output format employs compression that is expected to significantly reduce the data volume.

3.8.3 Programming/Procedural Considerations

Processing and operational codes for L2_SM_A algorithm are written in Fortran to make it consistent with other algorithms. This will ease inter-algorithm functioning of switches, data and information transactions. Programming of L2_SM_A algorithm will adhere to standard coding specifications to ensure consistent, maintainable, and readable code deliveries within the SMAP SDS. The Fortran program of the L2_SM_A algorithm will also meet the necessary requirements of language compliance, predictable execution, and code clarity. Adherence to these standards by programs will

allows efficient integration of SDS software components (Interfaces) and will simplify the Algorithm-to-PGE development process. The L2_SM_A Fortran code contains comments and version control information to track changes and streamline the development of software. A software specification document will be developed for documentation of all source codes.

Table 7 Soil moisture retrieval flags.

Soil moisture quality
Retrieval recommended flag
Retrieval attempted flag
Retrieval success flag
Radar water body detection success
Freeze-thaw retrieval success
Radar vegetation index retrieval success
Surface quality flag
Static waterbody
Radar water body
Coastal proximity
Urban
Precipitation
Snow ice surface (dynamic)
Permenant snow ice surface
Radar freeze-thaw retrieval
Model frozen ground
Mountain
Dense vegetation (Vegetation water content)
Nadir region

3.8.4 Quality Control and Diagnostics

Quality control (QC) is an integral part of the L2_SM_A processing. The QC steps of the L2_SM_A processing are based on the flags that are provided with the input data streams (L1C_S0), different types of masks, flags, and fractional coverage of other variables provided by ancillary data. The list of the flags is provided in Table 7. 'sigma0 quality flag' is deemed good if all of the sub-flags (fore, aft, RFI) are good. Generally '0' is assigned to the favorable condition for soil moisture retrieval. The L2_SM_A will process all data that have favorable conditions for soil moisture retrieval (VWC <= 5 kg/m^2, no rain, no snow cover, no frozen ground, no RFI, sufficient distance from open water). However, soil moisture retrieval will also be conducted for regions with the high VWC, rain, RFI cleaned data, and places closer to water bodies but appropriate flags will be added to these data points indicating their susceptibility to high errors. Therefore there are three flagging states associated with soil moisture retrieval:

- retrieval recommended flag: soil moisture was retrieved. The retrieval is either recommended or not recommended. For example, even if soil moisture value is retrieved over mountainous areas, the retrieval will be deemed unreliable.
- retrieval success flag: soil moisture retrieval was attempted. The retrieval is either successful or unsuccessful. For example, when sigma0 is out of reasonable range, the retrieval will be attempted but may fail.
- retrieval attempt flag: soil moisture retrieval was either attempted or not attempted. For example on frozen or water surfaces, retrieval will not be attempted.

Exception Handling

To obtain L2_SM_A product from the SMAP observations involves many aspects of product generation. This includes instrument performance, satellite data downlink, data preprocessing activities, quality of data (e.g., data drop-off), preceding algorithm performance, availability of ancillary data, and computing related resources. Due to these activities, exceptions are expected while operating the L2_SM_A algorithm on the SDS Testbed. The development of the L2_SM_A software also includes identification of expected exceptions. The list of expected exceptions is being developed. However, the formulation and computer coding of the L2_SM_A software will be made robust to withstand the expected exceptions.

3.8.5 Interface assumptions

The L2_SM_A product interacts with other products of SMAP as follows. It takes the radar backscatter and RFI flag from the SMAP L1C_S0, and freeze/thaw flag. The flag is generated by the freeze/thaw module of the L3_FT, which is used by the L2_SM_A (L3_FT is a daily composite and its output does not feed directly into the L2_SM_A).

The aggregated 3-km resolution of radar backscatter will be used as an input of the L2_SM_AP. The soil moisture output from the L2_SM_A provides an input to the L3_SM_A and to the optional algorithm of the L2_SM_AP. The water body flag and radar vegetation index from the L2_SM_A is used by the L2_SM_P and L2_SM_AP.

3.8.6 Latency

The L2_SM_A product has to meet the latency requirement of 24 hours from the time of data acquisition by the satellite. The latency requirement for the L1C_S0 product is 12 hour, which leaves 12 hour for the L2_SM_A product. There are 14 half-orbit acquisitions per day. Accordingly the target for the processing speed is ±1 hour per half orbit.

The latency of L2_SM_A product is mainly governed by the latencies involved in the time-series retrieval and the dynamic ancillary data (surface temperature). The time-series retrieval may be accomplished on time (Section 3.8.1). The surface temperature is used by a dielectric model (some models do not require the parameter). Furthermore, the effect of the temperature on the dielectric conversion is secondary. If the surface temperature becomes a major bottleneck, it may be omitted or the source may be replaced with a faster source.

Several outputs of the L2_SM_A processor are used by the L2_SM_P and L2_SM_AP processors. These are the transient water body flag, freeze/thaw flag, RVI information, and 3-km resolution sigma0. The evaluation of these parameters is fast, and will be performed prior to the time-time series retrieval. Subsequently, these will be made available internally for the L2_SM_P and L2_SM_AP so that the two processors do not need to wait until the completion of the L2_SM_A processor.

4 CALIBRATION AND VALIDATION

According to the mission requirements, the post-launch validation needs to establish that the product meets certain criteria to be released as a beta version, a provisional version (not discussed in this version of the document), and finally as a validated product. The implementation of the post-launch Cal/Val process is described in SMAP Calibration and Validation plan (JPL D-52544, Sept. 2012). The following cal/val elements are available:

- Field experiments (intensive ground sampling with airborne measurements): ground sampling of soil moisture with simultaneous airborne radar measurements; the required suite of ground sampling parameters depends on the final forward model; the underflights of airborne radar will be important for validating the scaling aspects of the in situ sampled soil moisture.
- Core sites (dense long term networks, limited spatial domain): calibrated soil moisture estimate established to the 3-km footprint scale (uncertainty assessed) based on several in situ stations within a footprint area supported with sub-footprint scale ancillary information (as specified by the retrieval algorithm). Figure 19 show the current list of the core sites. The sites have been categorized based on the applicability to different scales of the SMAP L2 products.
- Sparse networks (sparse long term networks, large spatial domain): calibrated soil moisture estimate established to the 3-km footprint scale (uncertainty assessed) based on one in situ station within a footprint area supported with sub-footprint scale ancillary information (as specified by the retrieval algorithm).
- Spaceborne remote sensing observations: radar backscatter from the spaceborne SAR instruments onboard, together with in situ ground truth.
- Independent hydrological models: global hydrological models including surface soil moisture as an output.

The soil moisture information alone is sufficient to perform the post-launch validation of the L2_SM_A product. However, whenever available, additional information of soil surface properties (roughness and correlation length), vegetation class, and VWC will be valuable to further analyze the validation results.

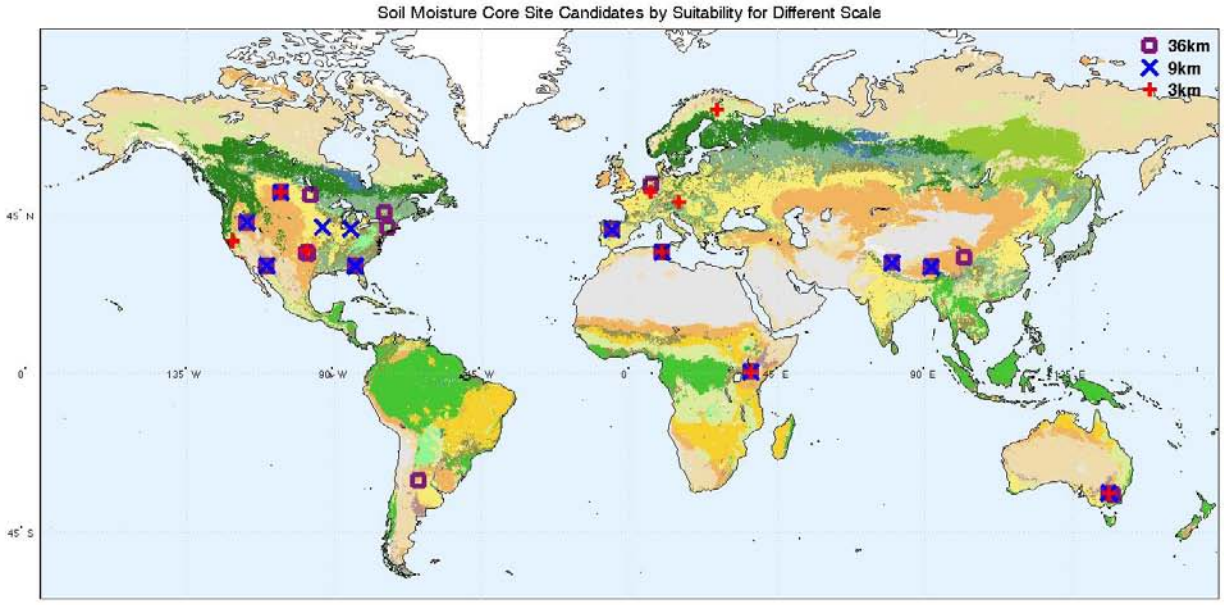

Figure 19. Potential Core Validation Sites categorized by scale.

4.1 Evaluation for Beta Release

For evaluating the L2&3_SM_A product quality for Beta release, a selected subset of the core sites and sparse networks need to be utilized. These resources need to provide almost real time soil moisture estimates in foot-print scale and their comparison with the satellite product should be possible without additional processing, allowing the fast assessment of the product quality and preliminary estimate of the uncertainties.

Subset of these resources will be selected so that they represent the major biomes and high priority areas of the globe. The selection will be performed prior to launch. Furthermore, the inclusion of analyses of the field experiment data is not necessary at this phase (although first field experiment data should be available), which saves time required for doing the evaluation.

The L2&3_SM_A product processing will require the possibility to adjust (calibrate) the basic parameters of the retrieval algorithms for quick tuning of the product. These parameters include adjustment of any bias component in radar backscatter, and other parameters that may reduce the retrieval errors arising from the heterogeneity. This needs to be accounted for in the preparation for the cal/val activities in pre-launch stage.

4.2 Validation

In the first stage of the validation a subset of the core sites, sparse networks and field experiments are needed to establish the L2&3_SM_A product accuracy in confined space and time domain. This refers to the areas directly covered by the selected resources and to the time period of the field experiments. The process is based on the comparison of the mentioned validation resources and the retrieved products. The comparison will be given in terms of mean, standard deviation, and correlation in soil moisture.

In the second stage of the validation, all core sites and sparse networks need to be included and also other satellite products to establish the L2&3_SM_A product accuracy in wider distribution of space and time domain. This refers to the areas covered by the identified validation resources and to the time period, or periods, representing one full season.

The first and second stages are expected to be completed during the geophysical product Cal/Val phase (12 months after launch), which will be followed by the release of the L2_SM_A product to the public. The validation with the other satellite products may have limited importance, depending on the fidelity of those products.

In the third stage of the validation the hydrological models are included as the accuracy of the L2&3_SM_A product is evaluated globally over an extensive time period. The combination of all cal/val elements will allow establishing the uncertainty of the product over global conditions in its lifetime.

The post-launch field experiments should cover high priority biomes and have separation in seasonal cycle. The L2&3_SM_A product will be sensitive to vegetation cover. This means sites with type and density of vegetation that is most critical for the performance of the retrieval (has to be determined in the pre-launch phase).

In the comparison to soil moisture products obtained with other spaceborne L-band SAR instruments (such as ALOS-2 and SAOCOM) the quality of these products, difference in both overpass date and time and differences in the retrieval systems has to be accounted for on case-by-case basis.

The combination of all validation data with the product flow is required to be a clearly defined independent process. This process needs to respond to the need to minimize the retrieval error over selected validation resources. There has to be a possibility not only to adjust parameters, which is needed also for the Beta release, but also to do more extensive updates to the algorithm based on the output of the validation process. The process needs to be defined and implemented before the launch.

5 APPENDIX

5.1 Hydros-era Algorithms

Two retrieval algorithms were considered in previous studies for the Hydros mission. The first is the "bare-soil with vegetation correction". In this algorithm (Dubois et al. 1995) the empirical equations relating co-polarized backscatter for bare soil, σ^0_{hh} and σ^0_{vv}, to soil dielectric constant and roughness are inverted. A correction is made for vegetation attenuation and scattering using the cross-polarized backscatter. Soil moisture is then retrieved from the dielectric constant using the dielectric model. This algorithm is considered by SMAP as one of the snapshot retrieval algorithms for low-vegetation surfaces.

The second Hydros approach used the numerical inversion of a forward model. In this algorithm, soil moisture and vegetation water content estimates are adjusted iteratively in computations of two co-pol σ^0 until the difference between computed and observed backscatter coefficients is minimized in a least squares sense. Both soil moisture and vegetation water content are retrieved by this algorithm. Initial guesses of the vegetation amount and knowledge of the surface roughness were required for successful retrieval. The forward model considered during the iterative minimization incorporates the bare surface σ_0 of Dubois et al. (Dubois et al. 1995) for the co-pol and σ_0 of Oh et al. (Oh et al. 1992) for the cross-pol, and Ulaby et al.'s cloud model for the volume and interaction components (Ulaby et al. 1986). This model is referred to as DOU hereafter.

Below, the performance of the Hydros second algorithm is revisited through examination of soil moisture retrieval errors obtained in a Monte Carlo simulation. In Figure 20 the soil moisture and surface roughness are estimated (VWC is assumed known with 10% standard deviation error with no bias error). Except for low-vegetation surfaces with VWC smaller than 0.5 kg/m^2, the m_v retrieval error exceeds 0.1 cm^3/cm^3. Further analyses suggest that the iterative retrieval from the Hydros era works well when the surface roughness is known with an accuracy better than ~40%. In practice, reliable ancillary information for surface roughness is not readily available (Joseph et al. 2008; Mattia et al. 2009; Shi et al. 1997; Verhoest et al. 2007). Therefore this approach does not meet the specified requirements for the SMAP baseline algorithm.

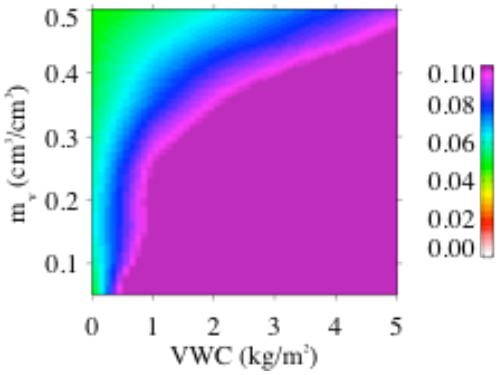

Figure 20. RMS errors from the Monte-Carlo simulation of the mv retrieval using the numerical iteration of the DOU forward model. m_v and surface roughness are estimated using three channels as input (HH, VV, HV), and assuming 10% Gaussian error in VWC and vegetation albedo and 2K error in surface temperature.

5.2 Insensitivity of retrieval to correlation length

The correlation length is not retrieved in the baseline time-series algorithm. Additional retrieval studies in which the correlation length to RMS height ratio was varied in creating the forward model tabulations showed that the time series algorithm retrieval of soil moisture is not sensitive to uncertainty in knowledge of the correlation length (Kim et al. 2012b). Figure 21 illustrates this by experimenting with a ratio of correlation length to RMS height of 4, 7, 10, and 15. Even if the correlation length changes σ^0 by several dB (Figure 21a-b), the changes are nearly constant with respect to the permittivity (or equivalently the soil moisture) and also between the two co-pols. In such a case a different correlation length does not affect the locations in the permittivity solution space where the difference in σ^0 between the forward model and the input measurements reaches its minimum. Therefore the retrieval of soil moisture would be essentially unaffected by errors in the correlation length. The exception to this is for rough terrain where the RMS height is 3 cm or greater and the ratio of correlation length to RMS height is 4 or less (Figure 21c-d). For these cases the retrieval of soil moisture and roughness becomes less accurate. The uniform increase in σ^0 due to the shorter correlation length may be explained by the decoupling of the roughness effect on backscatter from the dielectric effect. The decoupling would occur for a smooth random surface when the single scattering process is dominant (Fung et al. 1992).

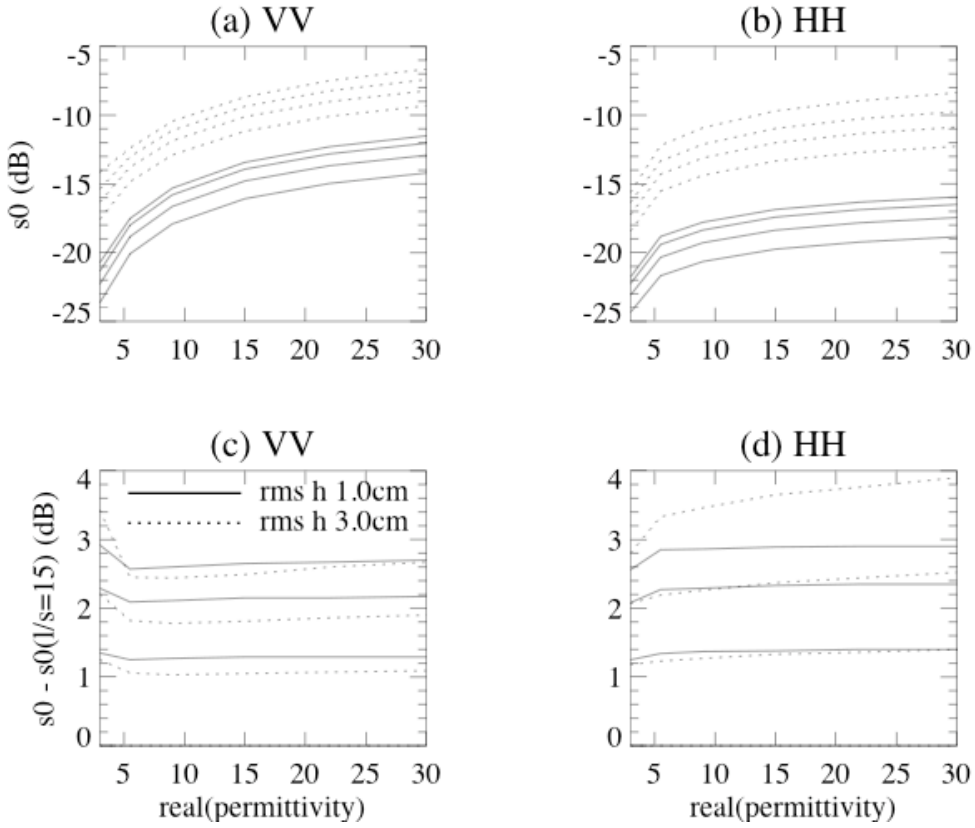

Figure 21. (a-b) NMM3D bare surface σ^0 as a function of polarization, the ratio of correlation length (*l*) to RMS height (*s*: 4, 7, 10, and 15), and roughness height (1cm and 3cm). σ^0 increases with the smaller ratio. (c-d) Difference in σ^0 is computed using the curves in (a-b) between three ratios of correlation length to correlation height (4, 7, 10) and the largest ratio (15), to show the increase in σ^0 with the shorter correlation length. For example, the largest difference in σ^0 is found with ratio 4 and ratio 15, and the smallest difference is with ratio 10 and ratio 15.

To validate the insensitivity of retrieval to correlation length, retrievals were made for four choices of the ratio of correlation length (l) to RMS height for roughness (s), which are available with the Ypsilanti bare surface data set. The m_v retrieval varies among the four retrievals by 0.004 (Site 1), 0.008 (Site 2), 0.012 (Site 3), and 0.015 (Site 4) cm^3/cm^3. As the roughness increases, the m_v retrieval begins to depend on the choice of l/s, consistent with the above analysis. Nevertheless the impact of incorrect knowledge in l/s on the m_v retrieval is less than 0.015 cm^3/cm^3 (Site 4). For the smooth surfaces such as Sites 1 and 2, the impact is negligible, which is consistent with the earlier analysis.

Table 8. Time-series data cube mv retrieval validated with the bare surface data over Ypsilanti, Michigan. The ratio of correlation length (l) to RMS height for roughness (s) is fixed to 10 during the retrieval (denoted by the superscript a); is chosen to match the in situ value (superscript b).

Site	l/s =4	l/s = 7	l/s = 10	l/s = 15
S-1	0.037	0.041	0.038[(a)]	0.038[(b)]
S-2	0.043	0.048 [(b)]	0.051[(a)]	0.048
S-3	0.043[(b)]	0.043	0.047[(a)]	0.055
S-4	0.055[(b)]	0.042	0.040[(a)]	0.042

5.3 Errors in VWC knowledge

During the Monte-Carlo simulation of the retrieval in the presence of the Kp and VWC errors (Sections 3.2.3 and 3.5.2), the VWC error is set to 10% (1σ, zero-mean). One candidate source of the VWC for the SMAP algorithms is the SMAP HV observation. Even if the sensitivity of σ^0 to VWC decreases towards higher VWC, the saturation of σ^0_{HV} occurs at higher VWC values than for co-pol σ^0. After considering the fore and aft averaging and the current-best-estimate of radar performance, the total error on the SMAP HV observation is 1 dB (1σ, zero-mean) at -30 dB signal (Table 3). 1 dB error results in the 10%-15% error up to 5 kg/m^2 VWC according to the following preliminary analysis. Let HV and VWC described by a certain function: VWC = f(HV). An example of $f()$ is given in Figure 22a by the discrete scatterer model (Arii et al. 2010) of the pasture vegetation. The corresponding error in VWC, $\delta VWC = \frac{\partial f}{\partial HV}\,\delta HV$, is shown in Figure 22b (in absolute size) and Figure 22c (in percentage). Figure 22c shows the VWC error of ~10% except when the percentage error becomes large due to the small VWC value. The finding would apply to various cases of surface roughness and soil moisture because HV is determined primarily by the volume scattering. A similar analysis will be performed for different vegetation classes.

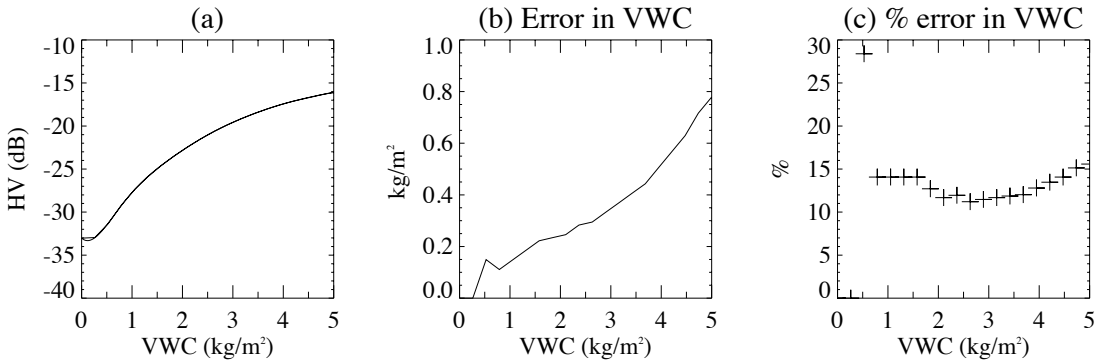

Figure 22. % error in VWC when VWC is derived from HV observation that has 1.5 dB error.

6 REFERENCES

Arii, M., J. J. van Zyl, and Y. Kim, 2010: A general characterization for polarimetric scattering from vegetation canopies. *IEEE Trans. Geosci. Remote Sens.*, **48,** 3349-3357.

Bindlish, R., and A. P. Barros, 2001: Parameterization of vegetation backscatter in radar-based soil moisture estimation. *Remote Sens. Environ.*, **76,** 130-137.

Burgin, M., D. Clewley, R. Lucas, and M. Moghaddam, 2011: A generalized radar backscattering model based on wave theory for multilayer multispecies vegetation. *IEEE Trans. Geosci. Remote Sens.*, **49,** 4832-4845.

Chan, C. H., S. H. Lou, L. Tsang, and J. A. Kong, 1991: Electromagnetic scattering of waves by random rough surface: a finite-diference time-domain approach. *Microwave & Optical Technology Letters*, **4,** 355-359.

Chan, S., 2011: SMAP radar error budget, JPL Report D-61622, Oct. 1, 2011, 67pp.

Chauhan, N. S., and R. H. Lang, 1991: Radar modeling of a boreal forest. *IEEE Trans. Geosci. Remote Sens.*, **29,** 627-638.

Chauhan, N. S., D. M. Le Vine, and R. H. Lang, 1994: Discrete scatter model for microwave radar and radiometer response to corn: comparison of theory and data. *IEEE Trans. Geosci. Remote Sens.*, **32,** 416-426.

Chen, K. S., T. D. Wu, L. Tsang, Q. Li, J. C. Shi, and A. K. Fung, 2003: The emission of rough surfaces calculated by the integral equation method with a comparison to a three-dimensional moment method simulations. *IEEE Trans. Geosci. Remote Sens.*, **41,** 90-101.

De Roo, R. D., Y. Du, and F. T. Ulaby, 2001: A semi-empirical backscattering model at L-band and c_band for a soybean canopy with soil moisture inversion. *IEEE Trans. Geosci. Remote Sens.*, **39,** 864-872.

Dobson, M. C., and F. T. Ulaby, 1986: Preliminary evaluation of the SIR-B response to soil-moisture, surface-roughness, and crop canopy cover. *IEEE Trans. Geosci. Remote Sens.*, **24,** 517-526.

Duan, X., and M. Moghaddam, 2011: Electromagnetic scattering from arbitrary random rough surfaces using stabilized extended boundary condition method (SEBCM) for remote sensing of soil moisture. *IEEE Trans. Geosci. Remote Sens.*, **50,** 87-103.

Dubois, P. C., J. J. van Zyl, and E. T. Engman, 1995: Measuring soil moisture with imaging radars. *IEEE Trans. Geosci. Remote Sens.*, **33,** 915-926.

Durden, S. L., J. J. van Zyl, and H. A. Zebker, 1989: Modeling and observation of the radar polarization signature of forested areas. *IEEE Trans. Geosci. Remote Sens.*, **27,** 290-301.

Entekhabi, D., E. G. Njoku, P. R. Houser, M. W. Spencer, T. Doiron, J. Smith, R. Girard, S. Belair, W. T. Crow, T. J. Jackson, Y. H. Kerr, J. S. Kimball, R. D. Koster, K. C. McDonald, P. E. O'Neill, T. Pultz, S. Running, J. C. Shi, E. Wood, and J. J. van Zyl, 2004: The Hydrosphere State (HYDROS) mission concept: An Earth system pathfinder for global mapping of soil moisture and land freeze/thaw. *IEEE Trans. Geosci. Remote Sens.*, **42,** 2184-2195.

Entekhabi, D., E. G. Njoku, P. E. O'Neill, K. H. Kellog, W. T. Crow, W. J. Edelstein, J. K. Entin, S. D. Goodman, T. J. Jackson, J. T. Johnson, J. S. Kimball, J. R. Piepmeier, R. D. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J. C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, and J. J. van Zyl, 2010: The Soil Moisture Active Passive (SMAP) mission. *Proc. IEEE*, **98,** 704-716.

Fung, A. K., Z. Li, and K. S. Chen, 1992: Backscattering from a randomly rough dielectric surface. *IEEE Trans. Geosci. Remote Sens.*, **30,** 356-369.

Hallikainen, M. T., F. T. Ulaby, M. C. Dobson, H. A. El-rayes, and L. K. Wu, 1985: Microwave dielectric behavior of wet soil. Part I: Empirical models and experimental observation. *IEEE Trans. Geosci. Remote Sens.*, **GE-23,** 25-34.

Huang, S., and L. Tsang, 2012: Electromagnetic scattering of randomly rough soil surfaces based on numerical solutions of Maxwell equations in 3 dimensional simulations using a hybrid UV/PBTG/SMCG method. *IEEE Trans. Geosci. Remote Sens.*, in press.

Huang, S., L. Tsang, E. G. Njoku, and K. S. Chen, 2010: Backscattering coefficients, coherent reflectivities, emissivities of randomly rough soil surfaces at L-band for SMAP applications based on numerical solutions of Maxwell equations in three-dimensional simulations. *IEEE Trans. Geosci. Remote Sens.*, **48,** 2557-2567.

Israelsson, H., L. Ulander, T. Martin, and J. Askne, 2000: A coherent scattering model to determine forest backscattering in the VHF-band. *IEEE Trans. Geosci. Remote Sens.*, **38,** 238-248.

Jackson, T. J., and A. Y. Hsu, 2001: Soil moisture and TRMM microwave imager relationships in the Southern Great Plains 1999 (SGP99) experiment. *IEEE Trans. Geosci. Remote Sens.*, **39,** 1632-1642.

Jackson, T. J., H. McNairn, M. A. Weltz, B. Brisco, and R. Brown, 1997: First order surface roughness correction of active microwave observations for estimating soil moisture. *IEEE Trans. Geosci. Remote Sens.*, **35,** 1065-1069.

Johnson, J. T., T. Leung, R. T. Shin, K. Pak, H. C. Chi, A. Ishimaru, and Y. Kuga, 1996: Backscattering enhancement of electromagnetic waves from two-dimensional perfectly conducting random rough surfaces: a comparison of Monte Carlo simulations with experimental data. *IEEE Transactions on Antenna Propagation*, **44,** 748.

Joseph, A. T., R. van der Velde, P. E. O'Neill, R. H. Lang, and T. Gish, 2008: Soil moisture retrieval during a corn growth cycle using L-band (1.6GHz) radar observations. *IEEE Trans. Geosci. Remote Sens.*, **46,** 2365-2374.

Karam, M. A., A. K. Fung, R. H. Lang, and N. S. Chauhan, 1992: A microwave scattering model for layered vegetation. *IEEE Trans. Geosci. Remote Sens.*, **30,** 767-784.

Kim, S. B., R. West, and E. G. Njoku, 2011: Effect of radar measurement error on the detection of transient inland water bodies. *IGARSS*, Vancouver, Canada.

Kim, S. B., M. Moghaddam, L. Tsang, M. Burgin, X. Xu, and E. G. Njoku, 2012a: Models of L-band radar backscattering coefficients over the global terrain for soil moisture retrieval. *IEEE Trans. Geosci. Remote Sens.*, submitted.

Kim, S. B., L. Tsang, J. T. Johnson, S. Huang, J. J. van Zyl, and E. G. Njoku, 2012b: Soil moisture retrieval using time-series radar observations over bare surfaces. *IEEE Trans. Geosci. Remote Sens.*, **50,** 1853-1863.

Kim, Y., and J. J. van Zyl, 2000: On the relationship between polarimetric parameters. *IEEE Geoscience and Remote Sensing Symposium*, Hawaii, USA.

——, 2009: A time-series approach to estimate soil moisture using polarimetric radar data. *IEEE Trans. Geosci. Remote Sens.*, **47,** 2519-2527.

Kuo, C. H., and M. Moghaddam, 2007: Electromagnetic scattering from multilayer rough surfaces with arbitrary dielectric profiles for remote sensing of subsurface soil moisture. *IEEE Trans. Geosci. Remote Sens.*, **45,** 349-366.

Lang, R. H., and J. S. Sighu, 1983: Electromagnetic Backscattering from a Layer of Vegetation: A Discrete Approach. *IEEE Trans. Antenn. Propag.*, **GE-21,** 62-71.

Lawrence, H., and co-authors, 2010: L-Band emission of a bare soil rough surface and a rough soil surface covered with a grass litter layer: Comparison between experimental data and a numerical modeling approach. *IGARSS*, Hawaii, USA.

Li, Q., and L. Tsang, 2001: Wave scattering from lossy dielectric random rough surfaces using the physics-based two-grid method in conjunction with the multilevel fast multipole method. *Radio Science*, **36,** 571-583.

Li, Q., J. C. Shi, and K. S. Chen, 2002: A Generalized Power Law Spectrum and Its Applications to the Backscattering of Soil Surface on the Integral Equation Model. *IEEE Trans. Geosci. Remote Sens.*, **40,** 271-280.

Liang, P., M. Moghaddam, L. Pierce, and R. Lucau, 2005: Radar backscattering model for multilayer mixed species forests. *IEEE Trans. Geosci. Remote Sens.*, **43,** 2612-2626.

Lin, Y. C., and K. Sarabandi, 1999: A Monte carlo coherent scattering model for forest canopies using fractal-generated trees *IEEE Trans. Geosci. Remote Sens.*, **37,** 440-551.

Lou, S. H., L. Tsang, and C. H. Chan, 1991: Application of the finite element method to Monte Carlo simulations of scattering of waves by random rough surfaces: penetrable case. *Waves in Random Media*, **1,** 287-307.

Mattia, F., G. Satalino, V. R. N. Pauwels, and A. Loew, 2009: Soil moisture retrieval through a merging of multi-temporal L-band SAR data and hydrologic modelling *Hydrol. Earth Syst. Sci.*, **13,** 343-356.

Mattia, F., T. LeToan, J. C. Souyris, G. DeCarolis, N. Floury, F. Posa, and G. Pasquariello, 1997: The effect of surface roughness on multifrequency polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.*, **35,** 954-966.

Matzler, C., 1994: Microwave (1-100 GHz) dielectric model of leaves. *IEEE Trans. Geosci. Remote Sens.*, **32,** 947-949.

Mautz, J. R., and R. F. Harrington, 1979: Electromagnetic scattering from a homogeneous material body of revolution. *Archiv fuer Elektronik und Uebertragungstechnik*, **33,** 71-80.

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

Mironov, V. L., M. C. Dobson, V. H. Kaupp, S. A. Komarov, and V. N. Kleshchenko, 2004: Generalized refractive mixing dielectric model for moist soils. *IEEE Trans. Geosci. Remote Sens.*, **42,** 773-785.

Moghaddam, M., S. Saatchi, and R. H. Cuenca, 2000: Estimating subcanopy soil moisture with radar. *J. Geophys. Res.*, **105,** 14,899-814,911.

Monfreda, C., N. Ramankutty, and J. A. Foley, 2008: Farming the Planet. Part 2: The Geographic Distribution of Crop Areas and Yields in the Year 2000. *Glob. Biogeochem. Cycles***,** doi:10.1029/2007GB002947.

Naeimi, V., K. Scipal, Z. Bartalis, S. Hasenauer, and W. Wagner, 2009: An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. *IEEE Trans. Geosci. Remote Sens.*, **47,** 1999-2013.

National Research Council, 2007: Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, 400 pp.

Njoku, E. G., W. J. Wilson, S. H. Yueh, S. Dinardo, F. K. Li, T. J. Jackson, V. Lakshmi, and J. Bolten, 2002: Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99. *IEEE Trans. Geosci. Remote Sens.*, **40,** 2659-2673.

Oh, Y., K. Sarabandi, and F. T. Ulaby, 1992: An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Trans. Geosci. Remote Sens.*, **30,** 370-382.

——, 2002: Semi-empirical model of the ensemble-averaged differential Mueller matrix for microwave backscattering from bare soil surfaces. *IEEE Trans. Geosci. Remote Sens.*, **40,** 1348-1355.

P.de Matthaeis, R. H. L., 2005: Microwave scattering models for cylinderical vegetation components. *Progress In Electromagnetics Research*, **55,** 307-333.

Pathe, C., W. Wagner, S. Sabel, M. Doubkova, and J. B. Basara, 2009: Using ENVISAT ASAR Global Mode data for surface soil moisture retrieval over Oklahoma, USA. *IEEE Trans. Geosci. Remote Sens.*, **47,** 468-480.

Peplinski, N. R., F. T. Ulaby, and M. C. Dobson, 1995: Dielectric properties of soils in the 0.3-1.3-GHz range. *IEEE Trans. Geosci. Remote Sens.*, **33,** 803-807.

Saatchi, S., and K. C. McDonald, 1997: Coherent effects in microwave backscattering models for forest canopies. *IEEE Trans. Geosci. Remote Sens.*, **35,** 1032-1044.

Scipal, K., M. Drusch, and W. Wagner, 2008: Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weal-her prediction system. *Advances in water resources*, **31,** 1101-1112.

Shi, J. C., J. R. Wang, A. Y. Hsu, P. E. O'Neill, and E. T. Engman, 1997: Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data. *IEEE Trans. Geosci. Remote Sens.*, **35,** 1254-1266.

Stiles, J. M., K. Sarabandi, and F. T. Ulaby, 2000: Electromagnetic scattering from grassland. II Measurments and modeling results. *IEEE Trans. Geosci. Remote Sens.*, **38,** 349-356.

Sun, R., J. C. Shi, T. J. Jackson, K. S. Chan, and Y. Oh, 2009: Improvement of bare surface soil moisture estimation with L-band dual-polarization radar. *IGARSS*, Cape Town, South Africa, IV971-974 pp.

Thirion, L., E. Colin, and C. Dahon, 2006: Capabilities of a forest coherent scattering model applied to radiometry, interferometry, and polarimetry at P-band and L-band. *IEEE Trans. Geosci. Remote Sens.*, **44,** 849-862.

Tsang, L., and J. A. Kong, 2001: *Scattering of electromagnetic waves - vol 3: advanced topics.* John Wiley & Sons.

Tsang, L., J. A. Kong, and R. T. Shin, 1985: *Theory of microwave remote sensing.* Wiley.

Tsang, L., C. H. Chan, and K. Pak, 1994: Backscattering enhancement of a two-dimensional random rough surface (three-dimensional scattering) based on Monte Carlo simulations. *Journal of Optical Society of America*, **11,** 711-715.

Tsang, L., J. A. Kong, and K. H. Ding, 2000: *Scattering of electromagnetic waves - vol 1: theories and application.* John Wiley & Sons.

Tsang, L., J. A. Kong, K. H. Ding, and C. O. Ao, 2001: *Scattering of electromagnetic waves - vol 2: numerical simulations.* John Wiley & Sons.

Tsang, L., D. Chen, P. Xu, Q. Li, and V. Jandhyala, 2004: Wave scattering with the UV multilevel partitioning method: 1. Two-dimensional problem of perfect electric conductor surface scattering. *Radio Science*, **39,** RS5010.

Ulaby, F. T., and H. A. El-rayes, 1987: Microwave dielectric spectrum of vegetation − Part II: Dual-dispersion model. *IEEE Trans. Geosci. Remote Sens.*, **GE-25**.

Ulaby, F. T., A. K. Moore, and A. K. Fung, 1986: Microwave remote sensing: active and passive. Artech House, 1064.

van Zyl, J. J., 2011: Ch. 5 Applications: Measurement of Surface Soil Moisture. *Synthetic Aperture Radar Polarimetry*, J. J. van Zyl, Ed., Wiley, 312.

Verhoest, N. E. C., B. De Baets, F. Mattia, G. Satalino, C. Lucau, and P. Defourny, 2007: A possibilistic approach to soil moisture retrieval from ERS synthetic aperture radar backscattering under soil roughness uncertainty. *Water Resour. Res.*, **43,** W07435.

Voronovich, A., 1994: Small-slope approximation for electromagnetic wave scattering at a rough interface of two dielectric half-spaces. *Waves in Random Media*, **4,** 337-367.

Wagner, W., G. Lemoine, and H. Rott, 1999: A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sens. Environ.*, **70,** 191-207.

Xu, X., L. Tsang, S. Huang, and E. G. Njoku, 2010: Microwave scattering model of vegetated surfaces for applications in SMAP mission. *PIERS*, Xi'an, China.

Yueh, S. H., J. A. Kong, J. K. Jao, R. T. Shin, and T. Le Toan, 1992: Branching model for vegetation. *IEEE Trans. Geosci. Remote Sens.*, **30,** 390-402.

Yueh, S. H., S. Dinardo, S. Chan, E. G. Njoku, T. J. Jackson, and R. Bindlish, 2008: Passive and Active L-Band System and observations during the 2007 CLASIC campaign. *IGARSS*.

Zribi, M., O. Taconet, V. Ciarletti, and D. Vidal-Madjar, 2002: Effect of row structures on radar microwave measurements over soil surface. *Int. J. Remote Sens.*, **23,** 5211-5224.

6.1 List of 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.

Ancillary Data Reports:

- 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.

7 ACRONYMS

BOR	Body of Revolution
CLASIC07	Cloud Land Atmosphere Interaction Campaign 2007
DDA	Discrete Dipole Approximation
DSM	Discrete Scatterer Model
DOU	Dubois/Oh/Ulaby radar scattering model
ECMWF	European Centre for Medium-Range Forecasting
EASE (grid)	Equal-Area Scalable Earth (grid)
GMAO	Goddard Modelling and Assimilation Office
HDF	Hierarchical Data Format
HYDROS	Hydrosphere State (satellite)
IEM	Integral Equation Model
IGBP	International Geosphere-Biosphere Programme
MODIS	Moderate-resolution Imaging Spectroradiometer
NMM3D	Numerical Maxwell Model in 3-Dimension
NDVI	Normalized Difference Vegetation Index
PALS	Passive/Active L-band Sensor
PALSAR	Phased Array type L-band Synthetic Aperture Radar
PGE	Product Generation Executable
RFI	Radio Frequency Interference
RMS	Root Mean Square
RMSE	Root Mean Square Error
RVI	Radar Vegetation Index
SAR	Synthetic Aperture Radar
SAOCOM	SAtélite Argentino de Observación COn Microondas
SDT	(SMAP) Science Definition Team
SDS	(SMAP) Science Data System
SGP99	Southern Great Plains 1999
SMAP	Soil Moisture Active Passive
SMEX02	Soil Moisture Experiment 2002

SMOS	Soil Moisture Ocean Salinity (mission)
SNR	Signal to Noise Ratio
SPM	Small Perturbation Model
SRTM	Shuttle Radar Topography Mission
UAVSAR	Uninhabited Aerial Vehicle Synthetic Aperture Radar
USDA	United States Department of Agriculture
USGS	United States Geological Survey
UW	Univ. Washington
VSM	Volumetric Soil Moisture
VWC	Vegetation Water Content

8 SYMBOLS

k	wave number
m_v	volumentric soil moisture
l	correlation length of bare surface height
s	RMS height of bare surface
ε	soil dielectric constant
ε_r	real part of soil dielectric constant
ε_{ri}	real part of soil dielectric constant at i-th time instance