



# Algorithm Theoretical Basis Document (ATBD)

## NISAR L3 Soil Moisture Data Product

Release 3.3

JPL D-107679

April 28, 2023

Author: NISAR Soil Moisture Science Team

Paper copies of this document may not be current and should not be relied on for official purposes. The current official version is in JPL EPDM <https://epdm.jpl.nasa.gov>

This document is a working version for the SDS R3.3 delivery and available at DocuShare: <https://charlie-lib.jpl.nasa.gov/docushare/dsweb/View/Collection-349513>

National Aeronautics and  
Space Administration



Jet Propulsion Laboratory  
California Institute of Technology  
Pasadena, California

## SIGNATURE PAGE

### Prepared by:

<u>Email Approval on File</u> Rowena Lohman, NISAR Soil Moisture Science Team Lead	<u>07-Dec-2023</u> Date
<u>Email Approval on File</u> Seungbum Kim, NISAR Soil Moisture Science Team	<u>05-Dec-2023</u> Date
<u>Email Approval on File</u> Narendra Das, NISAR Soil Moisture Science Team	<u>06-Jan-2024</u> Date
<u>Email Approval on File</u> Rajat Bindlish, NISAR Soil Moisture Science Team	<u>05-Jan-2024</u> Date
<u>Electronic Signature on File</u> Steven K. Chan, NISAR Soil Moisture ST/ADT Liaison	<u>05-Dec-2023</u> Date
<u>Email Approval on File</u> Xiaodong Huang, NISAR Soil Moisture ADT Product Developer	<u>05-Dec-2023</u> Date

### Approved by:

<u>Electronic Signature on File</u> Ana Maria Guerrero, NISAR Mission System Manager	<u>05-Dec-2023</u> Date
<u>Electronic Signature on File</u> Chuck Baker, NISAR Mission System Engineer	<u>05-Dec-2023</u> Date
<u>Electronic Signature on File</u> Paul Rosen, NISAR Project Scientist	<u>08-Dec-2023</u> Date
<u>Electronic Signature on File</u> Cecilia Cheng, NISAR SDS Manager	<u>06-Dec-2023</u> Date
<u>Electronic Signature on File</u> Heresh Fattahi, NISAR ADT Lead	<u>06-Dec-2023</u> Date
<u>Electronic Signature on File</u> Helen Mortensen, NISAR SDS System Engineer	<u>05-Dec-2023</u> Date

## EPDM ELECTRONIC SIGNATURES

User-Group/Role	...	Decision	Comments	Date
Cheng, Cecilia S (cecilia)-JPL Consumer/Pr...	...	Approve		06-Dec-2023 09:43
Chan, Steven K (chan)-JPL Consumer/Proje...	...	Approve		05-Dec-2023 14:15
Fattahi, Heresh (fattahi)-JPL Author/JPL Au...	...	Approve		06-Dec-2023 06:40
Guerrero, Ana Maria P (ana)-JPL Author/JP...	...	Approve		05-Dec-2023 15:09
Mortensen, Helen B (hbmorten)-JPL Consu...	...	Approve		05-Dec-2023 22:21
Baker, Charles J (cjbaker)-Engineering/Engi...	...	Approve		05-Dec-2023 14:39
Rosen, Paul A (parosen)-JPL Consumer/Pro...	...	Approve		08-Dec-2023 10:57

## DOCUMENT CHANGE LOG

Revision	Cover Date	Sections Changed	ECR #	Reason, ECR Title, LRS #*
Draft	Oct 25, 2022	All	N/A	Cleared for ATBD review LRS 068407
Initial Release	Apr 28, 2023	All	N/A	Cleared for Unlimited Release CL#23-2417

\* Include the JPL Limited Release System (LRS) clearance number for each revision to be shared with foreign partners.

## TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION</b>	<b>6</b>
1.1	Purpose of Description	6
1.2	Document Organization	6
1.3	Applicable and Reference Documents	6
<b>2</b>	<b>Overview</b>	<b>8</b>
2.1	NISAR mission	8
2.2	NISAR instrument characteristics	8
2.3	NISAR observation strategy	9
2.4	NISAR soil moisture goals and requirements	10
<b>3</b>	<b>Physics of Radar Response to Soil Moisture</b>	<b>12</b>
3.1	Bare soil scattering models	13
3.2	Vegetation scattering models	14
<b>4</b>	<b>Algorithm theoretical basis</b>	<b>15</b>
4.1	Physical Model Inversion Algorithm	15
4.1.1	Adapting and Assessing Forward Models for Inversion	15
4.1.2	Retrieval algorithm	18
4.1.3	Performance assessment using observations	20
4.1.4	Sensitivity analysis	24
4.1.5	Uncertainty characterization	26
4.2	Time Series Algorithm	27
4.2.1	Application to NISAR measurements	33
4.2.2	Performance of the Time Series Algorithm	34
4.3	Multiscale Fusion Algorithm	36
4.3.1	Testing of Multiscale Fusion Algorithm	42
4.3.2	Algorithm Implementation with ALOS PALSAR2 and UAVSAR Data	43
4.3.3	Validation of High Resolution Soil Moisture	44
4.3.4	Uncertainty in Soil Moisture from the Multiscale Fusion Algorithm	47
<b>5</b>	<b>Algorithm implementation and workflow</b>	<b>50</b>
5.1	Algorithm Flow	50
5.2	Preprocessing	50
5.2.1	EASEGRID2 description	50
5.3	GCOV to EASEGRID2 (Statistical filtering)	50

5.4	Output Variables	51
5.5	Quality Flags	52
<b>6</b>	<b>Ancillary data</b>	<b>55</b>
6.1	ECMWF Soil Moisture	55
6.2	Landcover	55
6.3	Cropcover	56
6.4	Vegetation Water Content	56
6.5	Snow	57
6.6	Frozen Ground	58
6.7	Precipitation	58
6.8	Water Bodies	58
6.9	Urban Areas	58
6.10	Soil Texture	59
6.11	Topography	59
<b>7</b>	<b>Calibration and Validation: Pre-launch</b>	<b>60</b>
7.1	SMAPVEX12 experiment	60
7.1.1	Estimation of sample statistics from SMAPVEX12 exercise	61
7.1.2	Errors on each field based on sample statistics	61
7.1.3	Errors on each crop type	62
7.1.4	Results	62
<b>8</b>	<b>Calibration and Validation: Post-launch</b>	<b>65</b>
8.1	In situ Validation	66
8.1.1	Configuration of <i>in situ</i> measurements at each super site	67
8.1.2	Location of Potential Super Sites	68
8.2	Satellite Products	68
8.3	Model Products	68
8.4	Field Experiments	69
<b>9</b>	<b>References</b>	<b>70</b>
	<b>APPENDIX A: Active-Passive Synergy</b>	<b>78</b>
	<b>APPENDIX B: NISAR Water body detection</b>	<b>79</b>
	<b>APPENDIX C: Acronyms</b>	<b>81</b>

# 1 INTRODUCTION

## 1.1 Purpose of Description

This document provides the algorithm theoretical basis document (ATBD) of the NASA-ISRO Synthetic Aperture Radar (NISAR) L-SAR Level-3 Soil Moisture product to be generated by the NASA Science Data System (SDS) and provided to the NASA Alaska Satellite Facility Distributed Active Archive Center (ASF DAAC)].

## 1.2 Document Organization

Section 2 provides an overview of the NISAR mission and the goal of the soil moisture and requirements.

Section 3 provides the physics of the radar response to soil moisture including the bare soil and vegetation scattering models.

Section 4 provides a detailed description of the algorithm theoretical basis including the DSG, TSR, and PMI algorithms.

Section 5 provides a detailed description of the algorithm implementation and workflow.

Section 6 provides a detailed description of the ancillary data used by the algorithms.

Section 7 provides the information of the algorithm calibration and validation before launch.

Section 8 provides the description of the algorithm calibration and validation after launch.

Section 9 provides the references.

Appendix A provides the description of the active-passive synergy.

Appendix B provides the description of the NISAR water body detection.

Appendix C provides a listing of the acronyms used in this document.

## 1.3 Applicable and Reference Documents

Applicable documents levy requirements on areas addressed in this document. Reference documents are cited to provide additional information to readers. In case of conflict between the applicable documents and this document, the Project shall review the conflict to find the most effective resolution.

### Applicable Documents

- [AD1] NISAR NASA SDS Level 4 Requirements, JPL D-95655, Initial, Sep. 13, 2019
- [AD2] NISAR NASA SDS Algorithm Development Plan, JPL D-95678, Initial, Sep. 12, 2019
- [AD3] NISAR Science Data Management and Archive Plan, JPL D-80828, June 1, 2016
- [AD4] NISAR Science Management Plan, JPL D-76340, Rev A, Aug. 14, 2018
- [AD5] NISAR Calibration and Validation Plan, JPL D-102256, September. 2019
- [AD6] NISAR NASA SDS L4 Software Management Plan (SMP), JPL D-95656, Rev A, Sep. 19, 2019
- [AD7] ISO-19115-2, <https://www.iso.org/obp/ui/#iso:std:iso:19115:-2:ed-2:v1:en>

## Reference Documents

- [RD1] EOSDIS Handbook, July 2016, retrieved from <https://cdn.earthdata.nasa.gov/conduit/upload/5980/EOSDISHandbookWebFinalL2.pdf>
- [RD2] NISAR SDS File Naming Conventions, JPL D-102255, Initial, Mar. 2, 2023
- [RD3] NISAR L1\_RSLC Product Specification Document, JPL D-102268, R3.1, August 05, 2022
- [RD4] HDF5 documentation at <https://portal.hdfgroup.org/display/HDF5/HDF5>
- [RD5] Eineder, M. (2003), Efficient simulation of SAR interferograms of large areas and of rugged terrain, *IEEE Transactions on Geoscience and Remote Sensing*, 41(6), 1415-1427.
- [RD6] Brodzik, M. J., B. Billingsley, T. Haran, B. Raup, M. H. Savoie. 2012. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. *ISPRS International Journal of Geo-Information*, 1(1):32-45, [doi:10.3390/ijgi1010032](https://doi.org/10.3390/ijgi1010032).
- [RD7] Brodzik, M. J., B. Billingsley, T. Haran, B. Raup, M. H. Savoie. 2014. Correction: Brodzik, M. J. et al. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. *ISPRS International Journal of Geo-Information* 2012, 1, 32-45. *ISPRS International Journal of Geo-Information*, 3(3):1154-1156, [doi:10.3390/ijgi3031154](https://doi.org/10.3390/ijgi3031154).

As noted on the cover of this document, the latest official versions of NISAR documents should be obtained from JPL Engineering Production Data Management (EPDM) <https://epdm.jpl.nasa.gov>. This document is a working version for the SDS R3.3 delivery and available at DocuShare: <https://charlie-lib.jpl.nasa.gov/docushare/dsweb/View/Collection-349513>.

The NISAR Level 1 science requirements are translated into requirements on the various spacecraft and instrument systems, including the requirements related to the processing system producing the L0-L3 products. These SDS requirements [AD1] fall into three general categories: resolution requirements, radiometric and spatial location accuracy requirements, and latency and throughput requirements. Note that there is no geophysical retrieval accuracy imposed on the NISAR L3\_SME2 product per agreement with the product sponsor – the Satellite Needs Working Group (SNWG); retrieval performance will be attempted on a best-effort basis.

## 2 Overview

Soil moisture is a key variable that impacts the exchange of water and heat energy between the land surface and the atmosphere. It plays an important role in the development of weather patterns and precipitation. Soil moisture also strongly affects the amount of precipitation that runs off into nearby streams and rivers. Soil moisture information can be used for reservoir management, early warning of droughts, irrigation scheduling, and crop yield forecasting. Soil moisture data has the potential to significantly improve the accuracy of short-term weather forecasts and reduce the uncertainty of long-term projections of how climate change will impact Earth's water cycle [e.g., Entekhabi et al., 2010].

Satellite remote sensing of soil moisture has advanced significantly over the last decade due to the success of the Soil Moisture and Ocean Salinity (SMOS) [Kerr et al. 2010] and Soil Moisture Active Passive (SMAP) [Entekhabi et al., 2010] missions, both of which provide global soil moisture retrievals on an approximate 3-day revisit interval at an accuracy of approximately  $0.04 \text{ m}^3/\text{m}^3$ . A key limiting factor of SMAP and SMOS soil moisture measurements is their coarse spatial resolution ( $\sim 40 \text{ km}$ ), which limits their utility for field-scale (i.e.,  $\sim 200 \text{ m}$ ) agricultural monitoring.

The unique capabilities of the NISAR mission [Rosen et al, 2015; Rosen et al., 2017] clearly motivate the production of field-scale ( $\sim 200 \text{ m}$ ) land surface soil moisture products.

### 2.1 NISAR mission

The NASA-ISRO Synthetic Aperture Radar Mission (NISAR), a collaboration between the National Aeronautics and Space Administration (NASA) and the Indian Space Research Organization (ISRO), is a Synthetic Aperture Radar (SAR) satellite that will make global measurements over most of the Earth's land surfaces. The NISAR instrument will provide a means of resolving complex processes ranging from ecosystem disturbances and dynamics, to behavior of the solid earth and cryosphere.

The ability of radar to penetrate clouds and operate day and night allows generation of a time-series of images with consistent temporal spacing that will extend throughout the mission lifetime. Among the important features of the mission characteristics are its wide swath (240 km), high resolution ( $\sim 15\text{m}$ ), 12-day repeat orbit cycle and dual-frequency (L- and S-band) capability.

In addition to measuring radar reflectivity, the NISAR mission has a number of other capabilities that will be useful for soil moisture estimation. Among these features is the capability of performing repeat-pass interferometric coherence analyses, and the collection of data with a range of polarizations. While the core payload is the L-band SAR, a secondary S-band SAR, built by ISRO, will provide opportunities to collect dual-frequency observations over key sites in India and others distributed globally.

### 2.2 NISAR instrument characteristics

With the SweepSAR technology used by NISAR, the entire incidence angle range is imaged as a single strip-map swath, at spatial resolution that depends on the mode, and with the potential for full polarization observations, if desired. The azimuth resolution ( $\sim 8 \text{ m}$ ) is determined by the 12-m reflector diameter, while the range resolution is determined by the bandwidth. The two frequencies of the NISAR, L- and S- band, are being designed such that they share a common clock and frequency references, allowing them to be operated simultaneously. A list of key features of the NISAR system are provided in Table 2.2.1.

**Table 2.2.1.** *NISAR Key Measurement System Characteristics*

System Characteristics	Value
Altitude	747 km
Repeat Period	12 days
Ground-Track Swath	240 km
Mission Duration	3 years
Orbit Inclination	98.5 degrees
Nominal Look direction	Right
Nodal Crossing Time	6AM/6 PM
Antenna diameter	12 m
L-band radar Center Frequency	1260 MHz
L-band Bandwidths	5, 20+5, 25, 37.5, 40+5, 80 MHz
L-band Realizable polarizations	Single through quad-pol, including split-band dual pol and compact pol
Incidence Angle Range	34 – 48 deg

The mission itself includes a large diameter (12 m) deployable reflector and a dual frequency antenna feed that supports implementation of the SweepSAR wide-swath mapping technology. The polarimetric capability of the NISAR system is expected to allow collection of dual-polarized (dual-pol) global observations over most regions for every cycle and has the potential to allow for quad-pol observations in selected areas within India and the U.S. Over land surfaces, the transmit polarization for dual-pol observations will principally be horizontally polarized, with both vertical and horizontal polarizations received, resulting in polarization combinations known as HH and HV to describe the configuration.

For a limited set of targets, the NISAR mission will make fully polarimetric measurements (i.e. fully-polarimetric, or quad-pol) by alternating between transmitting H-, and V-polarized waveforms and receiving both H and V (resulting in HH-, HV-, VH-, VV-polarized imagery). Polarization combinations such as dual- and quad-pol, allow for a fuller characterization of ground-target’s response. Variations in the polarimetric responses of targets to different combinations of polarization can be related to the physical characteristics of the target reflecting energy back to the radar and, hence, can be used for classifying target type and performing quantitative estimates of the target state.

## 2.3 NISAR observation strategy

The 240 km swath and short revisit time of NISAR will allow virtually all of the Earth’s land surfaces to be collected from the same viewing geometry every 12-days. L-band radar will map the global land, coastal and ice regions (Fig. 2.3.1). In addition to the NASA science requirements, ISRO scientists have specified targets of interest in India and its surrounding coastal waters. The ISRO requirements include both L-band and S-band observations. At selected cal/val sites, quad-pol data will be acquired over one track-frame containing the site location.

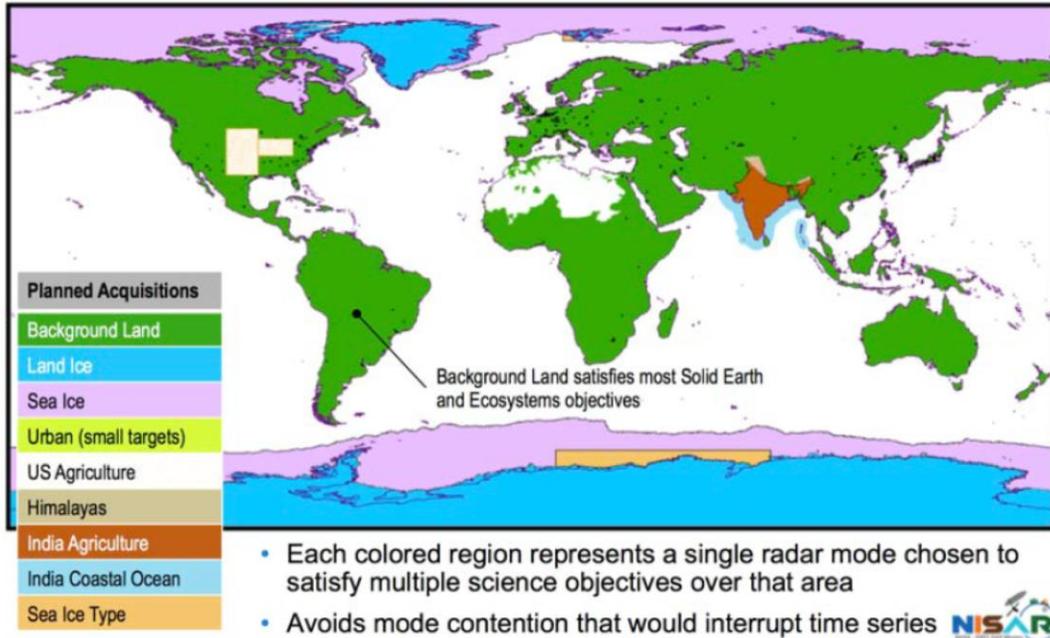


Fig. 2.3.1 NISAR observation strategy

## 2.4 NISAR soil moisture goals and requirements

NISAR backscatter observations will be used to estimate a global high resolution soil moisture product (200 m). This product will be provided on average twice every 12 days. The NISAR soil moisture product is expected to have a data latency of 72 hours (3 days). The NISAR level 2 backscatter product has a data latency of 48 hours. Soil moisture estimates will be provided over areas with dense vegetation (VWC greater than 5 kg/m<sup>2</sup>) but will be flagged during the retrieval process (Fig. 2.4.1). Areas with urban build-up, permanent snow and ice cover, and permanent inland waterbodies (shown in red) will be flagged and no soil moisture retrieval will be performed. No soil moisture retrieval will be performed over areas with excessive precipitation, frozen ground, or areas with snow on ground. The NISAR soil moisture will have an accuracy goal of 0.06 m<sup>3</sup>/m<sup>3</sup> over unflagged areas with vegetation water content below 5 kg/m<sup>2</sup>.

The salient characteristics of the NISAR product are:

- NISAR will produce a global soil moisture product with a repeat interval of 12 days (6 days when using both ascending/descending modes)
- Latency: 72 hours
- Resolution: 200 m
- Accuracy goal: 0.06 m<sup>3</sup>/m<sup>3</sup> (over non-urban areas with VWC below 5 kg/m<sup>2</sup>, with no permanent snow and ice cover, and no permanent inland waterbodies (Fig. 2.4.1))

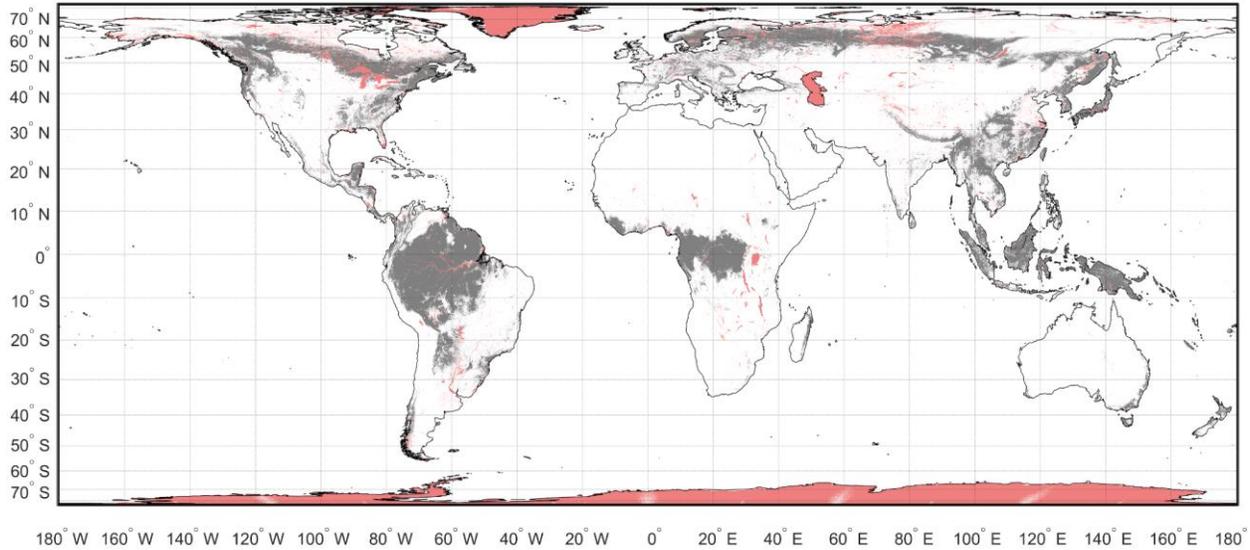


Figure 2.4.1 Soil moisture retrieval coverage. White land areas: retrievals performed with quality flag raised only in case of time-varying properties (e.g., heavy precipitation). Light red areas (urban build-up, permanent snow and ice cover, or permanent inland waterbodies) will be flagged and no NISAR soil moisture retrievals will be performed. Gray areas ( $VWC > 5 \text{ kg/m}^2$ ) will be flagged but soil moisture retrievals will be performed.

### 3 Physics of Radar Response to Soil Moisture

The normalized backscattered radar cross section (NRCS) for a vegetation-covered soil layer can be expressed as a sum of three components based on the scattering theory of distorted Born approximation on a single-scattering case:

$$\sigma_{pq}^t = \sigma_{pq}^s(\varepsilon, h, l, *) \exp(-\tau_{pq}(W, n)) + \sigma_{pq}^{sv}(W, n, \varepsilon, h, l, *) + \sigma_{pq}^v(W, n) \quad (3.1)$$

In this expression,  $\sigma_{pq}^t$  represents the total NRCS in polarization combination  $pq$ ,  $\sigma_{pq}^s$  represents the NRCS of the soil surface that is also multiplied by the two-way vegetation attenuation,  $\sigma_{pq}^v$  is the NRCS of the vegetation volume and  $\sigma_{pq}^{sv}$  represents the scattering interaction between the soil and vegetation. The medium parameters involved are the vegetation water content (VWC, represented by  $W$ ), the vegetation type (i.e. more randomly oriented, more vertical, etc.) indexed by  $n$ , the real part of the soil dielectric constant  $\varepsilon$ , which itself is a function of the volumetric soil moisture  $m$  and has been found sufficient to capture soil moisture effects on radar backscatter, and the surface rms height, correlation length, and other parameters ( $h, l, *$ ). Soil texture, temperature, etc. used for mapping from permittivity to moisture are also implicit parameters. A dependence on the azimuth angle is neglected here because it is assumed that any row oriented features will occupy only a portion of the eventual 200 m product resolution. All the individual cross section quantities and the vegetation attenuation vary with polarization in the most general case. The above formulation neglects any information in polarimetric correlations that may be achievable in the fully-polarimetric case to assist with separating surface and volume contributions. In general, no robust empirical forms for any of the terms in Eqn. 3.1 exists for direct incorporation into a global retrieval strategy.

The confounding factors of vegetation and surface roughness make soil moisture retrieval challenging, particularly when only single- or dual-polarization, single frequency measurements are available. Ancillary information, including vegetation properties or surface roughness, is used to improve the performance of soil moisture retrievals for the algorithms described below.

When NISAR operates in a mode having multiple polarization measurements, additional information can be obtained. Dual-pol measurements will provide both co- and cross-polarized backscatter. The heightened sensitivity of cross-polarization to vegetation contributions provides opportunities for its use in compensating for vegetation effects (particularly vegetation volume scatter contributions  $\sigma_{pq}^v$ ) in the retrieval process. Quad-pol modes offer extensive additional information in the form of HH, VV, and HV powers as well as their correlations, enabling a polarimetric decomposition of scattered fields into volume and surface contributions. Methods for improving soil moisture retrievals in such cases are presented in the following sections and are described in the literature [e.g., DiMartino et al., 2016]. However, the coverage expected for quad-pol mode is expected to be limited. S-band measurements, when available, could also provide benefits in soil moisture sensing. The algorithms reported here are focused on L-band only since this mode will provide the widest global coverage.

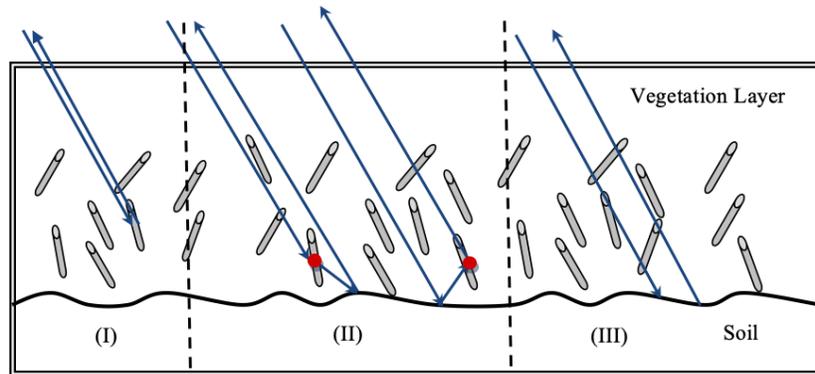


Figure. 3.1 Three scattering mechanisms: I) Direct volume scattering, II) Double bounce effect as exhibited by rough surface effect on the interface of the vegetation layer and soil, and III) Rough surface scattering of the soil

### 3.1 Bare soil scattering models

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 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 obtain 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]. 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 3.1.1. 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.

Results of the NMM3D approach are shown in Figure 3.1.1. and 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. The cross polarization results of NMM3D are also in good agreement with experimental data.

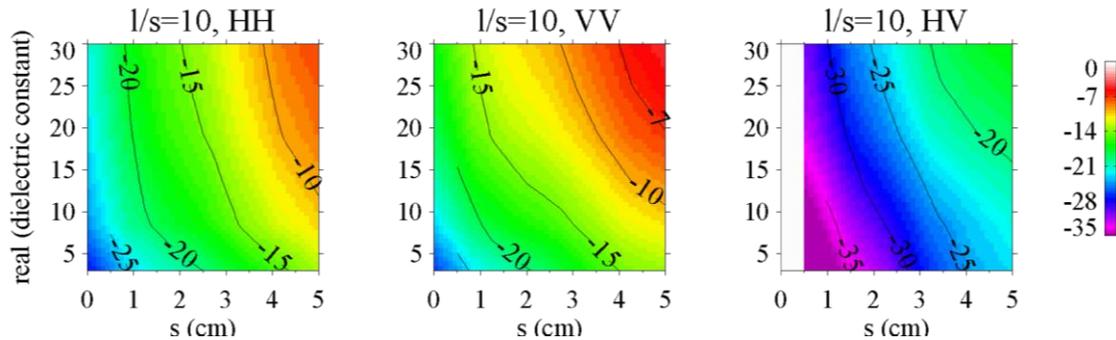


Figure 3.1.1. Color:  $\sigma^0$ (dB) generated by NMM3D bare surface simulations.  $l$  and  $s$  denote the surface correlation length and RMS height, respectively. Simulations were not performed for very smooth ( $s < 0.5$  cm for HV) surfaces. From [Kim et al., 2012].

## 3.2 Vegetation scattering models

The primary approach utilized is a “discrete scatterer” approximation [Lang and Sidhu, 1983; Tsang et al., 2000], in which each vegetation object is assumed to scatter independently. When multiple scattering is needed, a radiative transfer method is employed instead [Liao et al., 2016]. Fields from each vegetation component are summed and averaged over a range of size and orientation distributions. Several variations of the discrete scatterer model exist [Arii et al., 2010; Chauhan et al., 1994], varying in terms of 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. Vegetation components can be 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 the cylinder is larger, numerical solutions of Maxwell equations are solved for the object through the Body of Revolution (BOR) approach [Mautz and Harrington, 1979].

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 are 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 3.1). The model obeys reciprocity of backscattering in Maxwell’s equations and physically accounts for constructive interference in the backscattering direction [Tsang et al., 1985]. The coherent reflectivities of the random rough surface due to 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. In Eqn. 3.1,  $\tau$  is the vegetation opacity along the slant path of a radar beam and is often parameterized as a function of vegetation water content (VWC).

## 4 Algorithm theoretical basis

### 4.1 Physical Model Inversion Algorithm

In the physical model inversion algorithm, the forward scattering model is inverted to estimate soil moisture. The theoretical foundation was developed for the SMAP’s multiple-pol backscattering observations (HH, VV, and HV), which will be presented first. The adaptation to the NISAR’s dual-pol (HH and HV) and quad-pol mode situation will follow.

#### 4.1.1 Adapting and Assessing Forward Models for Inversion

The forward scattering models described in Section 3 are simplified by allowing three independent variables that co- or cross-pol backscattering is most sensitive to at L-band. When there are three independent measurement channels (HH, VV, HV), the retrieval will be able to estimate at most three independent parameters. These are rms height ( $s$ ), real part of dielectric constant ( $\epsilon_r$ ), and vegetation water content (VWC). Furthermore, because the scattering model is computationally intensive, to allow near-real-time estimates, the retrieval algorithm inverts the lookup table representation of the physical forward models [Kim *et al.*, 2014]. Since there are three dimensions, the lookup table is referred to as a ‘data cube’ [Kim *et al.*, 2014] (Fig. 4.1.1.1). The additional benefit of the data cube is to conveniently replace and update a forward model while retaining the same retrieval formulae and product generation system. The data cubes were generated at every 3° of incidence angle and interpolated at the angle of interest.

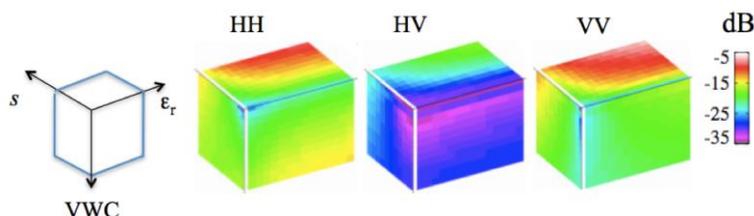


Figure 4.1.1.1 An example of L-band data cubes developed.  $s$  is surface roughness.  $\epsilon_r$  is dielectric constant. vegetation water content is the amount of water within a plant. From [Kim *et al.*, 2014].

**Rationale for fixing correlation lengths.** The fourth most influential independent variable in characterizing  $\sigma^0$  is the correlation length. For bare soil, the correlation length introduces relatively constant bias to backscattering with respect to soil moisture (Fig. 4.1.1.2). Then correlation length does not affect the soil moisture retrieval solution as the retrieval searches for the relative minimum difference between model and data. The impact of correlation length on soil moisture retrieval was smaller than 0.005 m<sup>3</sup>/m<sup>3</sup> (Table 3, [Kim *et al.*, 2012]). However, this robustness in the bare soil case could be complicated when scattering from vegetation layers (volume or double bounce) become less dominant over soil surface scattering as soil becomes wetter, demonstrated in a shrubland study (Figures 9 and 10, [Kim *et al.*, 2017a]). It is fixed to 10 (ratio of the length to roughness), after observing that the value allows matching of  $\sigma^0$  between model and field campaign data. An exponential function is known to describe empirical measurements well [Mattia *et al.*, 1997; Shi *et al.*, 1997].

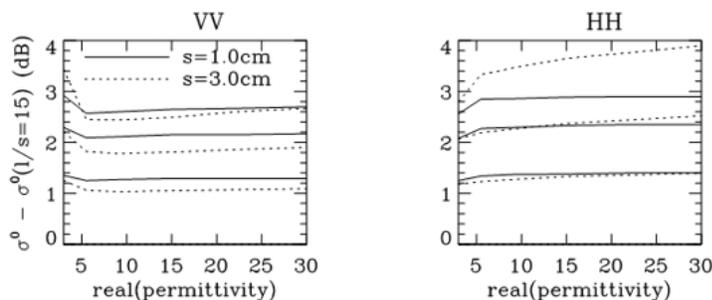


Fig. 4.1.1.2. Difference in  $\sigma^0$  is computed between three ratios (4, 7, and 10; top, middle, and bottom, respectively) and largest ratio (15) for smooth and rough surfaces (roughness 's' of 1 cm and 3 cm, respectively). The ratio is correlation length to roughness height ( $l/s$ ). Simulated by a bare soil scattering model [Kim et al., 2012].

**Developing data cubes for each vegetation type.** Detailed characterization of vegetation properties is needed to construct the forward scattering model as described in Eq. 3.1. The characterization of vegetation geometry includes length, thickness, diameter, and number of branches. The plant samples were weighed before and after drying to obtain the water fraction within a plant. All these data were collected in various field campaigns over the past 10 years funded mainly by the SMAP project as listed in Table 4.1.1.1.

While modeling  $\sigma^0$ , VWC may vary in multiple ways. For crops in an agricultural field, VWC varies predominantly via plant growth [Liao et al., 2016] and soybean [Huang et al., 2016]. For forest, on the time-scale of soil moisture changes (< 20 days), VWC remains static. However, for one data cube to represent forests at different locations, VWC are expected to vary via the changes in trunk density [Tabatabaenejad et al., 2012] or tree height [Kurum et al., 2020]. Shrubs in arid land grow very slowly and the dielectric properties of the plants were found to control VWC [Kim et al., 2017a]. Except for shrubs, the VWC axis of the data cubes reflects the plant growth (Table 4.1.1.1).

**Representing vegetation types globally:** the data cubes are developed for the 9 landcover types, limited by the available field campaign data sets. Although these 9 sets represent all the IGBP's global landcover classes and 4 major crops in the world, there are sub-species (e.g., within the forest category) and crops of smaller presence. Based on the geometry of stems and leaf foliage, the 9 sets are mapped to the more detailed plant types (Table 4.1.1.1).

Table 4.1.1.1 List of data cubes developed for the PMI algorithm, field campaigns that provide the data to train the data cubes, the additional landcover types associated with the data cubes, and how the values of the VWC axis change.

data cube	Field campaign	Additional landcover types to represent	VWC change mechanism
forest	SMAPVEX12 [Kurum et al., 2020]	All forest	Geometry growth
Savanna	Australian [Burgin et al., 2011]	Woody savanna	Geometry growth
Shrub	California [Kim et al., 2017a]	Cotton, pecan, berry	Dielectric change
Corn	SMAPVEX12 [Liao et al., 2016]	Sunflower	Geometry growth
Grass	Eel river [Liao et al., 2021] SMAPVEX12[Kim and Liao, 2021]	Pasture, rangeland, hey, sod	Geometry growth
Wheat	SMAPVEX12[Kim and Liao, 2021]	Forage, oat, barley, rye	Geometry growth
Soybean	SMAPVEX12 [Huang et al., 2016]	Peanut, potato	Geometry growth
Canola	SMAPVEX12[Huang et al., 2021]	rapeseed	Geometry growth

Bare soil	Michigan [Huang et al., 2010]	Fallow	n/a
-----------	-------------------------------	--------	-----

**Periodic structures and terrain slopes:** Periodic structures of soil and vegetation may produce resonance in  $\sigma^0$ . Resonance occurs only at preferred spacing of furrows and plants: such condition was not established in the SMAPVEX12 case [Kim et al., 2018]. With the SMAPVEX12 and SMEX02 analyses, the forward modeling without the special consideration of the periodic structure was accurate to 1-2dB RMSE for corn and soybean [Huang et al., 2016; Liao and Kim, 2022; Liao et al., 2016]. The periodic structure is expected to cause bias in forward model and retrieval (not the unbiased rmse), when the effect of the structure is temporally persistent.

Terrain slopes change  $\sigma^0$  via surface and double-bounce scattering mechanisms. The forward models do not currently incorporate these effects. However, when the surface scattering is dominant as in grassland, the existing data cubes were demonstrated to reliably predict  $\sigma^0$  data by correctly modeling the local incidence angle effect (1-2 dB rmse, [Liao et al., 2021]). Such success may not be guaranteed when the double-bounce process further complicates the effect of the terrain slope [Burgin et al., 2016], which will pose limitations of the algorithm.

**Assessing forward model fidelity with field campaign data:** the forward model predictions were assessed by comparing the SAR or scatterometer data (Table 4.1.1.2). The rms difference ranges from 1 to 2 dB for HH and VV, which is sufficient to produce soil moisture retrievals with an accuracy better than 0.06 m<sup>3</sup>/m<sup>3</sup> in most of the cases. The sufficiency is feasible thanks to the strong sensitivity of L-band  $\sigma^0$  on soil moisture (6 to 8 dB, Kim et al. 2012) and the optimization during retrieval. For grassland, corn and soybean, the results represent independent assessments ([Liao and Kim, 2022; Liao et al., 2021]), in that the models were developed using the data collected from the SMAPVEX12 campaign but the model predictions were evaluated with independently collected data at different times and locations.

Table 4.1.1.2 Performance of the physical model on SAR or scatterometer data. The references are (0) [Kim et al., 2012], (1) [Huang et al., 2016], (2) [Liao et al., 2016], (3) [Liao et al., 2021], (4) [Kim et al., 2017a], (5)[Kurum et al., 2020], (6) [Kim and Liao, 2021] (7) [Huang et al., 2021] (8) [Liao and Kim, 2022].

	bare soil <sup>0</sup>	wheat <sup>6</sup>	bean <sup>1,6,8</sup>	corn <sup>2,8</sup>	canola <sup>7</sup>	pasture, grass <sup>3,6</sup>	shrub <sup>4</sup>	forest <sup>5</sup>
Forward model uncertainty for HH and VV in <u>rmse</u> , dB	~1	~1	~1	~1.7	~2	~1.8	~1.5	~1.5
Soil moisture retrieval unbiased <u>rmse</u> , m <sup>3</sup> /m <sup>3</sup>	0.044	0.050	0.070	0.071	0.08	0.054	0.055	0.044
Soil moisture retrieval correlation	0.89	0.92	0.56	0.50	0.73	0.60	0.95	0.84
# time-series data used for retrieval	11	10	6 to 10	6 to 10	10	6 to 8	14	10

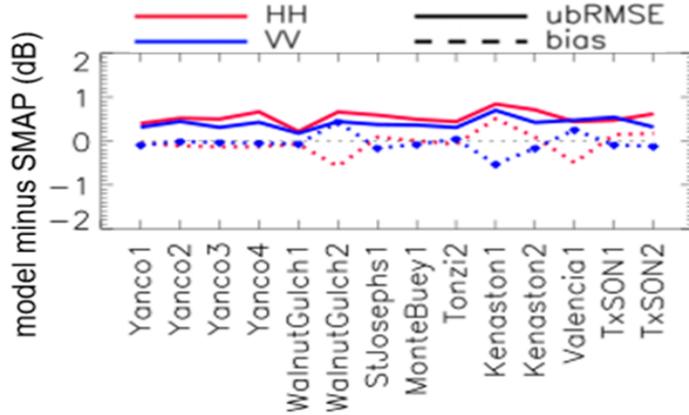


Figure 4.1.1.3. Forward model fidelity (model minus SMAP 3-km  $\sigma^0$ ), assessed over a 2.5-month period of SMAP radar operation [Kim *et al.*, 2017]. Abscissa refers to the names of cal/val site over the world.

**Testing the forward model on a global domain.** SMAP’s radar data offered an opportunity to test the forward model’s fidelity independently at global locations. Despite the coarse resolution of 3 km, the ground validation network was constructed to sample the 3-km pixel with at least 6 points to meet statistical significance. Consistent with the field campaign comparison in Table 4.1.1.2, the bias and unbiased rmse are within 2 dB (Figure 4.1.1.3).

## 4.1.2 Retrieval algorithm

The algorithm searches for a soil moisture solution such that the difference between computed and observed backscatter is minimized in the least squares sense [Kim *et al.*, 2012]. Eq. 4.1.2.1 formulates the cost function to minimize when HH and VV are available in NISAR’s quad-pol mode (VV contribution to the cost vanishes when only HH is available). The algorithm estimates  $s$  (soil roughness) first and then retrieves  $\epsilon_r$  (soil dielectric constant) using the estimated  $s$ . Vegetation effects are corrected by selecting the forward model’s  $\sigma^0$  at the VWC level given by an ancillary source or NISAR HV measurements. This scheme is implemented by slicing the 2-dimensional LUT in the VWC-soil moisture space at the given VWC level (Figure 4.1.2.1). The key components of the algorithm are (1) inverting the physical forward model and (2) reduce retrieval ambiguity using time-series input of backscattering. The end-to-end flow of the retrieval is shown in Figure 4.1.2.2.

$$C(s, VWC, \epsilon_{r1}, \epsilon_{r2}, \dots, \epsilon_{rN}) \tag{4.1.2.1}$$

$$\begin{aligned} &= w_{1,HH} \left( \sigma_{HH}^0(t_1) - \sigma_{HH,fwd}^0(s, VWC, \epsilon_{r1}) \right)^2 + w_{1,VV} \left( \sigma_{VV}^0(t_1) - \sigma_{VV,fwd}^0(s, VWC, \epsilon_{r1}) \right)^2 \\ &+ w_{2,HH} \left( \sigma_{HH}^0(t_2) - \sigma_{HH,fwd}^0(s, VWC, \epsilon_{r2}) \right)^2 + w_{2,VV} \left( \sigma_{VV}^0(t_2) - \sigma_{VV,fwd}^0(s, VWC, \epsilon_{r2}) \right)^2 + \dots \\ &+ w_{N,HH} \left( \sigma_{HH}^0(t_N) - \sigma_{HH,fwd}^0(s, VWC, \epsilon_{rN}) \right)^2 + w_{N,VV} \left( \sigma_{VV}^0(t_N) - \sigma_{VV,fwd}^0(s, VWC, \epsilon_{rN}) \right)^2 \end{aligned}$$

where values from observations and from the forward model are denoted as  $\sigma^0$  and  $\sigma_{fwd}^0$  (both in dB), respectively. Note that the above formulation can accommodate the temporal change in VWC, because  $\sigma_{fwd}^0$  is chosen by the VWC value available at each time. Because  $\sigma^0$  is a monotonic function of  $s$ , the minimum is unique with respect to  $s$ .  $\sigma^0$  is also a monotonic function with respect to  $\epsilon_r$ . Therefore, the minimum associated with  $\epsilon_{ri}$  is unique for a given  $s$ .

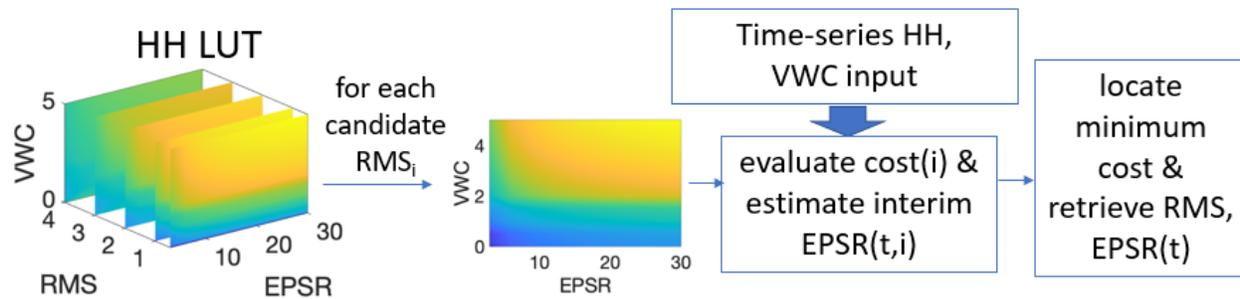


Figure 4.1.2.1. Time-series inversion using the LUT form of the physical model for HH per each vegetation class.  $i$  refers to the candidate of  $s$  (soil roughness) estimate.

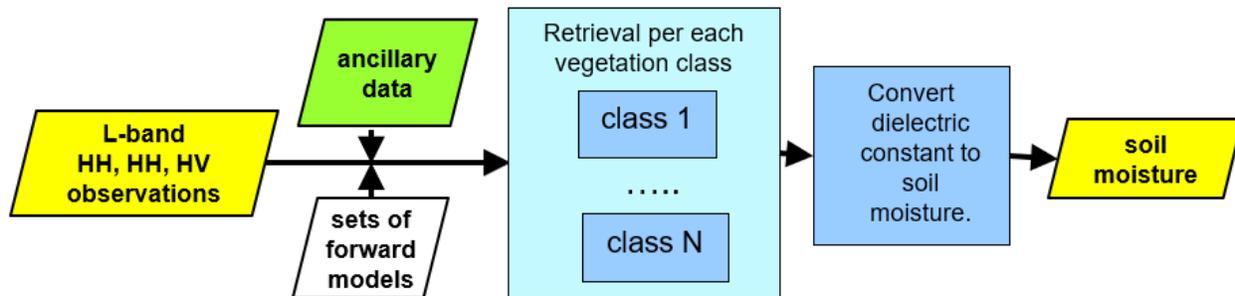


Figure 4.1.2.2. Overall flow of the physical model inversion algorithm.

**Advantage of the time-series approach:** One co-pol measurement (HH or VV) is generally not sufficient to determine  $s$  and  $\epsilon$  [Kim et al., 2012]. 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 1, Kim et al., 2012). Very often the time-scale of the change in  $s$  is longer than that of  $\epsilon$  [Jackson et al., 1997]. Then  $s$  may be constrained to be a constant in time, thus resolving the ambiguity [Kim et al., 2012]. These concepts are presented in terms of how well-conditioned the retrieval is in Table 4.1.2.1.

- The dual-copol time-series inputs facilitate a well-conditioned case with  $2N$  independent input observations, and  $N+1$  unknowns consisting of  $N$   $\epsilon_r$  values and one  $s$  value
- In the presence of speckle noise, the snapshot approach becomes prone to retrieval errors (2 independent inputs vs. 2 unknowns). In comparison, the redundancy of the time-series method provides reliable estimates of soil moisture and roughness.
- When time-series data consist of one co-pol, the retrieval becomes ill-posed but can be resolved by constraining the retrieval with residual and saturation soil moisture values associated with the soil texture (will be demonstrated later in Section 4.1.4).

The simulation in Fig. 4.1.2.3 shows the superior performance of the time-series approach. With experimental data collected over the bare soil, the time-series retrievals were more accurate than the snapshot method by  $0.02 \text{ m}^3/\text{m}^3$  rmse (Table 2 of [Kim et al., 2012]).

Table 4.1.2.1. Comparison of ill-posed condition among retrieval strategies. Degree of orthogonality between HH and VV is assumed to be 100% in this table, but would be smaller in reality.  $N$  is the number of input time-series observations.

	dual co-pol time-series	single co-pol time-series	dual co-pol snapshot
# unknowns ( $N \epsilon_r, 1 s$ )	$N+1$	$N+1$	$2N$

# independent input (HH, VV)

$\angle IN$

$IN$

$\angle IN$

**Number of input time-series ( $M$ ):** To achieve a well-constrained system, the number of time-series should be at least 2 for dual-copol time-series approach (Table 4.1.2.1), assuming HH and VV are orthogonal to each other. Since the orthogonality is not 100%,  $N$  should be larger. Experimental results shown in Table 4.1.1.2 demonstrated that  $N$  of 6 or larger allowed reliable retrieval. Based on these findings,  $N$  for the retrieval with the 3-day repeat SMAP data was determined to be 6. Considering that NISAR's revisit is ~6 and 12 days with and without combining ascending and descending observations, the baseline for  $N$  is 3 for NISAR.

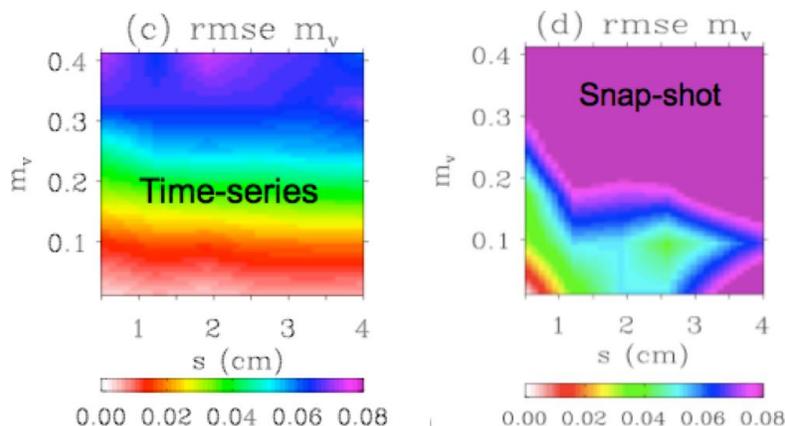


Fig. 4.1.2.3. Soil moisture retrieval errors from Monte-Carlo simulations. Time-series (left) vs snapshot (right). From [Kim et al., 2012].

### 4.1.3 Performance assessment using observations

Across the range of the NISAR incidence angle (33 to 47°), the retrievals were successful using airborne and satellite SAR data over diverse conditions of vegetation, soil moisture, and terrain slope. The validation results are summarized as follows and were presented in Table 4.1.1.2.

**SMAPVEX12 crops:** Over agricultural fields, the soil moisture estimates are accurate to unbiased rmse of 0.041, 0.059 and 0.060  $m^3/m^3$  and correlation of 0.71, 0.83, and 0.69 for grass, wheat, and soybean fields, respectively over the soil moisture dynamic range up to 0.5  $m^3/m^3$ , and soybean's full growth cycle [Kim and Liao, 2021].

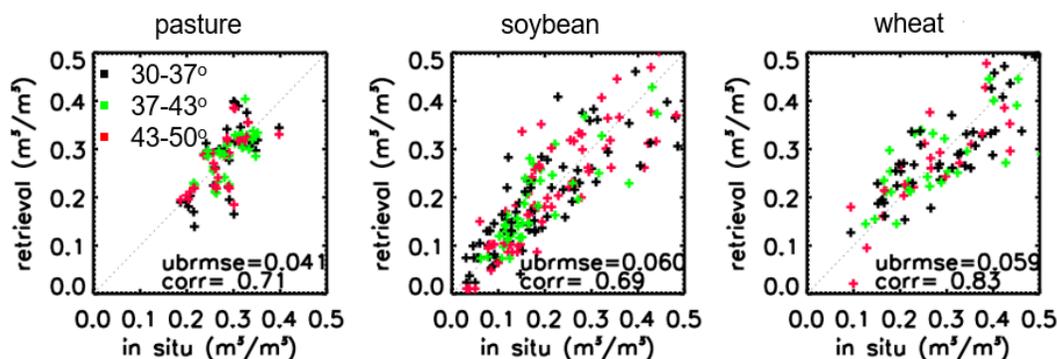


Figure 4.1.3.1. Soil moisture retrievals: grass, wheat, and soybean. To obtain unbiased rmse, a bias was removed per each field and the statistics were averaged. From [Kim and Liao, 2021].

**SMAPVEX12 forests:** For aspen forests, the retrievals over the full dynamic ranges of wetness are accurate to  $\sim 0.04 \text{ m}^3/\text{m}^3$  unbiased rmse with correlation of  $\sim 0.90$ , which is very encouraging for retrieval under the forest canopy (Figure 4.1.3.2) [Kurum et al., 2020]. In these SMAPVEX12 agricultural and forest cases, the physics of the forward models were developed using the same field campaign data. In contrast, the next cases present independent validation.

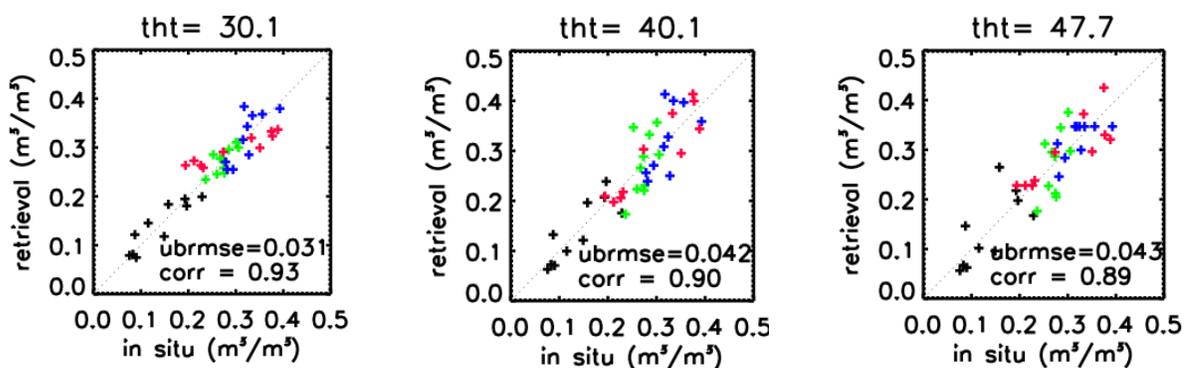


Figure 4.1.3.2. Retrievals over forest observations at different incidence angles ( $tht$ ). Colors correspond to site 1 (black), 2 (blue), 3 (green), and 5 (red). The bias error between retrieval and in situ data is removed individually for each forest site. The statistics are computed per each field, and those from four fields are averaged. From [Kurum et al., 2020].

**SMEX02:** The SMEX02 campaign offered L- and S-band scatterometer observations. NISAR will operate at L- and S-bands over India. Our goal is to explore benefits of the dual-frequency retrieval. The dual-frequency results show the unbiased rmse of  $0.031$  and  $0.057 \text{ m}^3/\text{m}^3$  for corn and soybean, respectively, which are improvements over the single-frequency cases by up to  $0.010$  and  $0.004 \text{ m}^3/\text{m}^3$  (Figure 4.1.3.3). Forward scattering models for corn and soybean fields were previously generated and validated at L-band from the SMAPVEX12 campaign: they are inverted by applying to the SMEX02 data. Also, it is demonstrated that the L-band modeling of forward scattering processes is scalable at S-band in that the physics and parameters behind modeling the vegetation effects remain the same between L- and S- bands [Liao and Kim, 2022].

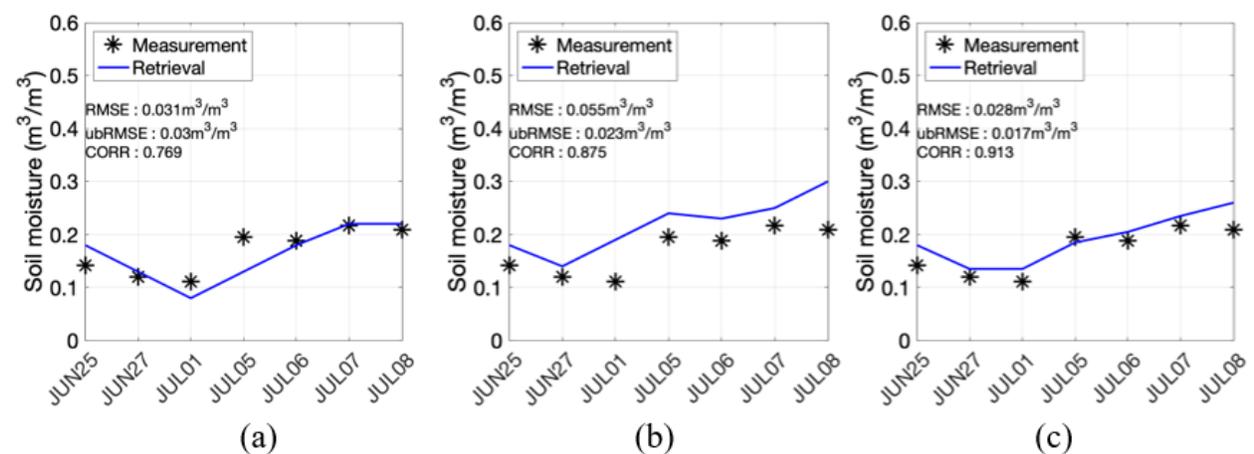


Figure 4.1.3.3. Example of L & S-band retrievals on one corn field from the SMEX02 campaign: (a) L-band (b) S-band (c) combined L- and S- band. From [Liao and Kim, 2022].

Table 4.1.3.1. Retrieval statistics with SMEX02 L- and S-band observations. The units for unbiased rmse and bias are  $m^3/m^3$ .

	L-band		S-band		L- and S-band	
	ub-rmse	R	ub-rmse	R	ub-rmse	R
Corn (17 fields, 7 dates)	0.041	0.44	0.039	0.52	0.031	0.56
Soybean (10 fields, 7 dates)	0.059	0.12	0.061	0.05	0.057	0.09

**AMPM UAVSAR:** ‘AM’ simulates the 6am NISAR acquisition during the ascending pass, and ‘PM’ for descending. We applied the soil moisture retrievals to 9 temporal UAVSAR acquisitions made over a 80-day period. The two objectives are (1) examine if the forward models developed using the SMAPVEX12 data are applicable to these independent sites and (2) whether the ascending and descending observations at given location with different incidence angles (ranging from 30 to 50°) can be combined to form time-series inputs during the retrieval. The second objective will be discussed in more detail in Section 4.1.4. The retrievals agree well with temporal rainfall events (Fig. 4.1.3.4). The validation with in situ measurements shows unbiased rmse (correlation) of 0.066  $m^3/m^3$  (0.73) and 0.079  $m^3/m^3$  (0.48) for corn and soybean crop fields, respectively [Ganesan *et al.*, 2022] (Table 4.1.3.2). Even when the incidence angles change during the time-series inputs (asc+des), the retrieval algorithm is capable of handling the change with ub-rmse degrading by only  $\sim 0.01 m^3/m^3$  (Table 4.1.3.2).

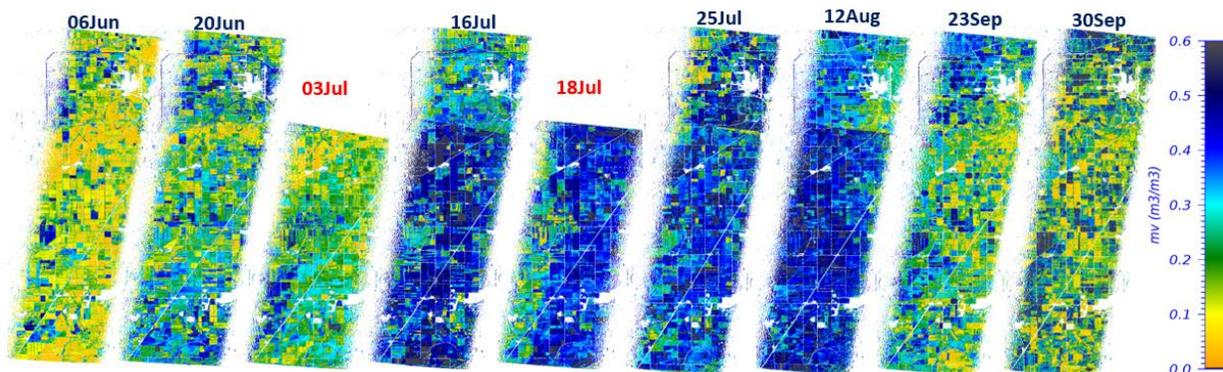


Figure 4.1.3.4. Retrievals over the AMPM UAVSAR domain (20 by 40 km) in 2019 by combining 7 ascending and 2 descending (July 3 and 18) observations. From [Ganesan *et al.*, 2022].

Table 4.1.3.2. Retrieval statistics for AMPM19 UAVSAR observations. The units for rmse and bias are  $m^3/m^3$ . ‘Asc’, ‘Des’, and ‘Asc+Des’ refer to the tracks used as inputs to the time-series algorithm. Descending soybean cases has too few samples to compute correlation.

	corn				soybean			
	rmse	ub-rmse	bias	R	rmse	ub-rmse	bias	R
Asc	0.116	0.053	0.087	0.82	0.114	0.079	0.073	0.48
Des	0.113	0.044	0.087	0.33	0.119	0.070	0.083	-
Asc + Des	0.108	0.066	0.077	0.73	0.119	0.079	0.082	0.48

**AMPM19 PALSAR:** the PALSAR data offer very similar properties as the NISAR data. The goals in this exercise are (1) perform SAR-based retrievals over cotton fields for the first time, and (2) evaluate the algorithm performance using these independent data sets: the forward models trained using the SMAPVEX12 UAVSAR data were applied to the AMPM domain. The retrieved soil moisture was validated over corn, soybean and cotton crops covering different growth stages and a wide range of moisture conditions (Figure 4.1.3.5). The estimated soil moisture is accurate to an unbiased rmse of 0.057, 0.067 and 0.070  $m^3/m^3$  for corn, soybean and cotton crops, respectively (Table 4.1.3.3). We used the shrub

forward model for cotton, noting the geometric similarity of the two plants. These results confirm the general applicability of physical scattering models to an independent test site.

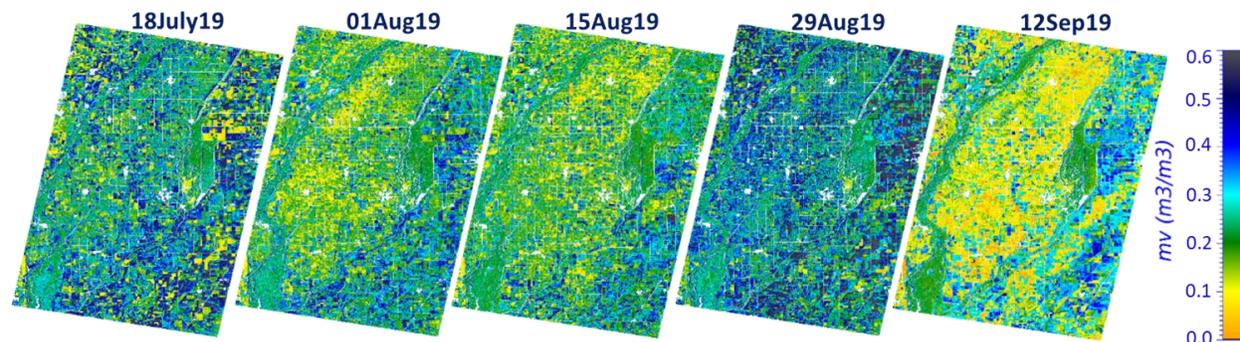


Figure 4.1.3.5. Retrievals using PALSAR2 data over the AMPM domain (40km by 60km) in 2019.

Table 4.1.3.3. Retrieval statistics with PALSAR time-series input. The units for rmse and bias are  $m^3/m^3$ .

	rmse	ub-rmse	bias	R
Cotton (4 fields, 18 data points)	0.065	0.046	0.043	0.68
Corn (2 fields, 6 data points)	0.154	0.046	0.145	0.97
Soybean(4 fields, 17 data points)	0.079	0.042	0.064	0.77

**Land-slide occurring hills:** Over steep hills ( $\sim 15^\circ$  slope) in northern California, 6-m resolution maps of surface soil moisture were generated with a unbiased rmse of  $0.054 m^3/m^3$  [Liao et al., 2021], paving ways for understanding and predicting landslide dynamics (Figure 4.1.3.6).

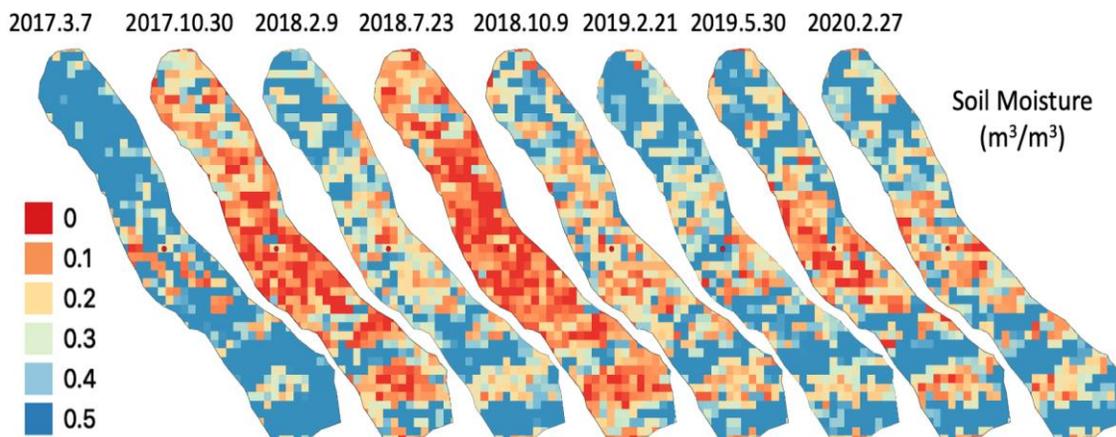


Figure 4.1.3.6. Retrieved soil moisture map for the Eel river landslide domain at the spatial resolution of 6 m. California's annual monsoon cycle is well captured by the retrieval along with the fine spatial details. From [Liao et al., 2021].

**SMAP:** The algorithm has been validated globally using the SMAP radar data at SMAP's  $40^\circ$  incidence angle at 3-km resolution [Kim et al., 2017] (Fig. 4.1.3.7; Table 4.1.3.4). Soil moisture retrievals using HH and VV inputs have an accuracy of  $0.052 m^3/m^3$  ub-rmse,  $-0.015 m^3/m^3$  bias, and a correlation of 0.50, as compared to in-situ measurements. The successful retrieval demonstrates the feasibility of this approach on a global domain over diverse conditions of soil moisture, surface roughness, and vegetation.

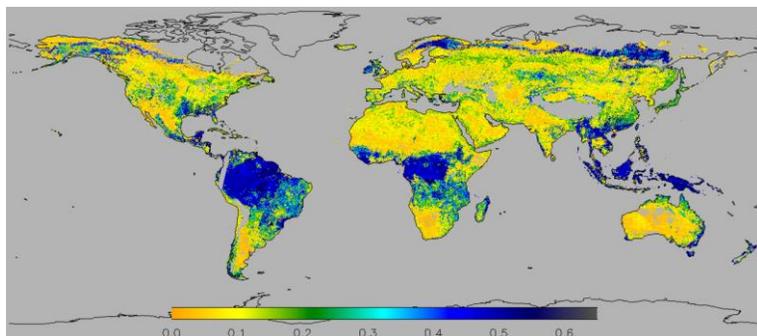


Figure 4.1.3.7. Soil moisture retrieval in  $m^3/m^3$  by the physical model inversion-7 method using 3-km SMAP SAR data over one 8-day cycle in May, 2015.

Table 4.1.3.4. Validation of soil moisture retrievals by the physical model inversion with SMAP SAR data at core validation sites performed over the 2.5-month period in 2015. Units in  $m^3/m^3$  [Kim et al., 2017].

<b>Cropland</b>	ub-rmse	Bias	<b>Non-crop land</b>		ub-rmse	Bias
St. Josephs1	0.051	-0.044	<b>Grassland</b>	Walnut Gulch1	0.014	0.024
Kenaston1	0.104	-0.025		TxSON1	0.047	-0.038
Kenaston2	0.087	-0.037		TxSON2	0.053	-0.029
Monte Buey1	0.08	-0.016		Yanco2	0.058	-0.013
Valencia1	0.032	0.026		Yanco3	0.040	-0.017
Yanco1	0.049	-0.02		Yanco4	0.063	-0.008
			<b>Shrubland</b>	Walnut Gulch2	0.017	-0.013
			<b>Woody Savanna</b>	Tonzi2	0.030	0.002
<b>Average</b>	<b>0.067</b>	<b>-0.019</b>	<b>Average</b>		<b>0.040</b>	<b>-0.012</b>

### 4.1.4 Sensitivity analysis

**Robustness to input polarization:** The majority of NISAR’s observations will be in the dual-pol (HH and HV) mode with limited quad-pol regions. Evaluated with the experimental data, the HH-based retrieval satisfies the accuracy target of  $0.06 m^3/m^3$  ub-rmse in most of the cases. The difference in retrieval ub-rmse between single-copol (HH) and dual-copol inputs is  $\sim 0.01 m^3/m^3$  or smaller, assessed using various campaign data (see below). Analysis in Table 4.1.2.1 and the bare soil experiment in Figure 4.1.2.3 indicated that the dual co-pol inputs offer superior performance. The consequent deficiency in single copol, however, was mitigated by constraining the retrieval with residual and saturation soil moisture values associated with the soil texture. These findings suggest that NISAR’s HH data is capable of delivering reliable estimates. The details of the results are summarized below.

**SMAPVEX12 crop:** The retrieval comparison of the different input channels in ub-rmse using the SMAPVEX12 data show that single-copol and dual-copol inputs produce comparable performance (Section 6.1.4). VV-based retrievals performed better than HH&VV- or HH-input estimates by  $\sim 0.01 m^3/m^3$  for wheat, corn, and soybean crops. The stronger sensitivity of VV to soil moisture changes, than that of HH, explains the result. Most likely the dual-copol’s performance was negatively impacted by the weak sensitivity of HH.

**SMAPVEX12 forest:** This result represents the VV input, because physical modeling of both copols was challenging and resulted in deteriorated retrieval.

AMPM UAVSAR & PALSAR: The results represent HH and VV inputs, with single co-pol cases showing comparable or slight degradations by up to 0.01 m<sup>3</sup>/m<sup>3</sup>.

SMEX02 & land-slide region: We only used HH data to simulate the NISAR operation mode.

**Robustness to the errors in landcover information:** noting that crop database will have latency up to one-year, the sensitivity to the correct choice of data cube types for the vegetation on the ground is assessed by randomly assigning data cubes during the retrieval (Table 4.1.4.1). The choice of data cubes with the wrong crop type increases the ub-rmse by up to 0.002 (when pasture was the target), 0.014 (wheat), 0.021 (bean), 0.024 (corn), and 0.005 (canola) m<sup>3</sup>/m<sup>3</sup>. With these degradation, corn and wheat retrievals exceed the accuracy target and require accurate choice of landcover information.

Table 4.1.4.1. Soil moisture retrieval error in ub-rmse (m<sup>3</sup>/m<sup>3</sup>) and correlation assessed with the SMAPVEX12 data. 'randomly'-chosen data cubes are used to retrieve for the 'target' landcover fields. The differences in ub-rmse or correlation between the correct choice in (d) and the wrong choice on the rest of columns quantify the retrieval error due to the random choice of data cubes.

Random choice of data cube	(a) corn		(b) bean		(c) wheat		(d) retrieval with correct data cube	
	ub-rmse	corr	ub-rmse	corr	ub-rmse	corr	ub-rmse	corr
pasture	0.038	0.30	0.036	0.41	0.036	0.58	0.036	0.58
wheat	0.076	0.69	0.074	0.70	same as in (d)		0.062	0.82
bean	0.049	0.82	same as in (d)		0.072	0.59	0.051	0.80
corn	same as in (d)		0.093	-0.1	0.081	-0.0	0.069	0.27
canola	0.053	0.78	0.052	0.78	0.055	0.74	0.050	0.79

**Robustness to incidence angle:** The NISAR swath has an incidence angle range of 33-47°. Figures 4.1.4.1 and 4.1.4.2 show the retrieval for grass, wheat, soybean, and forest using SMAPVEX12 data at three incidence angle bands (colors correspond to each band). The retrieval performances do not depend on the incidence angles. The robustness benefits from the two properties: forward modeling and retrieval algorithm.

- First, forward models reliably simulate the scattering processes regardless of incidence angle.
- Second, the retrieval algorithm corrects for the vegetation effect. For example, with grown soybeans, volume scattering dominates the total  $\sigma^0$  more severely at high angles than at low angles [Kim and Liao, 2021]. The consequent reduction in sensitivity of  $\sigma^0$  to soil moisture, however, does not adversely affect retrieval. Even if the longer path length produces stronger attenuation at high incidence angles than at low angles, the retrieval performance is satisfactory thanks to reliable forward modeling and correction of the vegetation effect during the retrieval. For forest, the double-bounce components, even at the VV channel, are strong. Since the path length of double-bounce

is the same regardless of incidence angle, the retrieval performances are not affected by the angles.

**Performance over steep terrain:** In a land-slide occurring steep hillslope (~15° slope), the retrieval was accurate with an unbiased rmse of 0.054 m<sup>3</sup>/m<sup>3</sup> [Liao *et al.*, 2021] (Figure 4.1.3.6). In this grassland, all types of scattering (volume, surface, and double-bounce) are found important. The terrain slope alters the local incidence angle. When surface scattering is dominant, the data cubes developed for a flat terrain are expected to perform well through the variation of local incidence angle. The volume scattering is irrelevant to the change in the angle, and the retrieval is expected to correct for the volume scattering effect successfully. These two processes explain the robustness of the retrieval performance to terrain slope.

**Expanding to wider regions:** Independent validations (meaning a forward model developed from one site and applied blindly to the other sites) were performed as one way to test the algorithm’s applicability to the NISAR’s global observations. The validation cases include SMEX02 campaigns (Fig. 4.1.3.3), AMPM UAVSAR domain (4.1.3.4), AMPM PALSAR domains (Fig. 4.1.3.5), land-slide hills (Fig. 4.1.3.6), and SMAP data (Fig. 4.1.3.7). The retrievals were successful in all these independent test cases. The basis of the expandability lies in the idea that the same plants share the similar scattering physics, allowing the application of the data cubes to broader regions in the world.

**Combining ascending and descending data:** NISAR’s exact repeat period is 12 days, which can be too infrequent for many applications. By combining ascending and descending data, 6-day repeat on average can be achieved. However, the two time-series data from the ascending and descending passes will have different incidence angles due to the look angle geometry and local topography. Earlier in this section, the algorithm performance was found robust to the incidence angle variations (where the angle changed by site location, but the time-series input data maintained the same angle). The retrieval performance when the time-series input data have *different* angles was evaluated using the AMPM UAVSAR data (Fig. 4.1.3.4; Table 4.1.3.2). Even when the incidence angles change during the time-series inputs (asc+des configuration), the retrieval algorithm is capable of handling the change with ub-rmse degradation by ~ 0.01 m<sup>3</sup>/m<sup>3</sup> only. Though the experiment’s scope is limited (there were only two descending acquisitions), the results are encouraging to implement the consistent retrieval combining ascending and descending observations.

#### 4.1.5 Uncertainty characterization

Uncertainty information is a key component of retrieval products, which helps end-users quantitatively evaluate the confidence of retrievals at each pixel. The retrieval performance ( $\Delta m_v$ ) is evaluated in a simulated environment (Eq. 4.1.5.1).  $\partial m_v / \partial \epsilon$  is the sensitivity of soil moisture to dielectric constant ( $\epsilon$ ), evaluated using a dielectric model of soil.  $\partial \epsilon / \partial \sigma$  represents the sensitivity of  $\sigma^0$  to soil moisture per each vegetation type and vegetation amount, and is computed using the radar forward models presented in Section 4.1.1.  $\Delta \sigma$  evaluates the speckle noise and calibration error in NISAR HH.  $\Delta \sigma$  also includes the impacts by the uncertainties in ancillary data (roughness  $s$ , vegetation water content  $vwc$ , and landcover  $lc$ ) when the ancillary data are used to sample the forward model  $\sigma^0$  values during the forward model inversion.

$$\Delta m_v = \frac{\partial m_v}{\partial \sigma} \Delta \sigma = \frac{\partial \epsilon}{\partial \sigma} \frac{\partial m_v}{\partial \epsilon} \Delta \sigma \quad (4.1.5.1)$$

$$\Delta \sigma = \sqrt{\Delta \sigma_{inst}^2 + \Delta \sigma_s^2 + \Delta \sigma_{vwc}^2 + \Delta \sigma_{lc}^2}$$

The result of the error model shows that about 70% of the global pixels have the retrieval uncertainty smaller than the target of 0.06 m<sup>3</sup>/m<sup>3</sup> (Figure 4.1.5.1; Table 4.1.5.1) in all four seasons. Considering that 70% is

roughly the size of 1 standard deviation (66%), these results support that the retrievals are accurate to the target of 0.06 m<sup>3</sup>/m<sup>3</sup> at the 1-sigma level

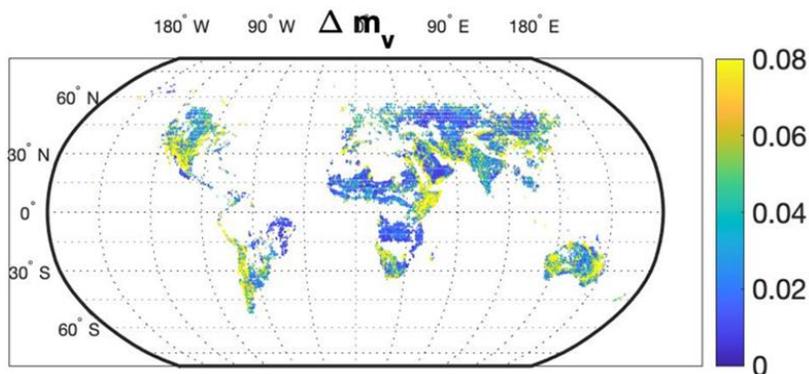


Figure 4.1.5.1. Expected uncertainties in soil moisture retrieval produced by the performance tool. Deserts, dense forests, and frozen lands are excluded.

Table 4.1.5.1. Percentage of pixels with retrieval uncertainty smaller than the target of 0.06 m<sup>3</sup>/m<sup>3</sup>. The numbers are the tabular presentation of the results shown in Figure 4.1.5.1.

	spring	summer	fall	winter
% of pixels	73	67	71	70

## 4.2 Time Series Algorithm

The desire for finer spatial resolution in the retrieval of soil moisture has motivated numerous studies of the use of high resolution synthetic aperture radar including those based on the use of backscatter powers [Burgin and Zyl, 2017; Koyama et al., 2017; Zwieback and Berg, 2018], backscatter time series [Ouellette et al., 2017; Mattia et al., 2009; Mattia et al., 2018; He et al., 2017; Pierdicca et al., 2010; Pierdicca et al., 2013], repeat pass interferometric correlation or phase [Hensley et al., 2011; Zwieback et al., 2015a; Zwieback et al., 2015b; Ouellette et al., 2012], and/or backscatter polarimetric decompositions [Jagdhuber et al., 2015; DiMartino et al., 2016; Ouellette et al., 2014; Trong-Loi et al., 2009], as well as other uses of SAR data to “downscale” other soil moisture products to finer spatial scales [e.g. Das et al., 2017; Peng et al., 2017]. A time-series based retrieval demonstrated for the C-band Sentinel SAR system has shown success in producing 1 km soil moisture products operationally with an RMSE of approximately 0.0675 m<sup>3</sup>/m<sup>3</sup> [Mattia et al., 2018].

Multiple references [Ouellette et al., 2017; Mattia et al., 2009; Mattia et al., 2018; He et al., 2017] describe a “time-series ratio” approach for retrieving soil moisture from radar backscatter time series measurements that attempts to eliminate the confounding influences of vegetation and surface roughness through a “ratio method”. The method is developed for terrain classes in which vegetation volume scatter  $\sigma_{pq}^v$  and soil-vegetation interaction  $\sigma_{pq}^{sv}$  terms can be neglected compared to the vegetation attenuated surface backscatter. Under this assumption, the NRCS has the form  $\sigma_{pq}^s(\varepsilon, h, l, *)e^{-\tau(W,n)}$  which is further assumed to have the form  $|\alpha(\varepsilon)|\sigma_{pq}^{sr}(h, l, *)e^{-\tau(W,n)}$ , i.e. that the NRCS of the bare surface is a multiplication of a functional dependence on permittivity  $|\alpha(\varepsilon)|$  with a functional dependence on roughness. This assumption is consistent with existing “single scattering” theories of backscatter from rough surfaces such as the small perturbation method, physical optics, or the small slope approximation.

The normalized backscattered radar cross section (NRCS) for a vegetated soil layer is a function of parameters related to soil, vegetation, and roughness, making the inverse problem of solving for soil moisture more difficult. The NRCS also includes components caused by multiple scattering of vegetation

and interactions between vegetation and the soil surface. However, short-term changes due to these complex vegetation contributions are negligible. The backscattered NRCS with a first-order small-perturbation model (SPM) can be expressed as

$$\sigma_{PP}^0 = 4\pi k^4 h^2 \cos^4 \theta |\alpha_{PP}|^2 W(2k \sin \theta, 0), PP = HH \text{ or } VV \quad (4.2.1)$$

where  $k = 2\pi/\lambda$  is the wavenumber,  $\theta$  is the incidence angle,  $h$  is rms height, and  $W(\xi_x, \xi_y)$  is the 2D normalized surface roughness spectrum. The alpha coefficient ( $\alpha_{PP}$ ) includes the surface electrical properties for each polarization (either HH or VV), which is a function of the dielectric constants of the soil and the incidence angle.

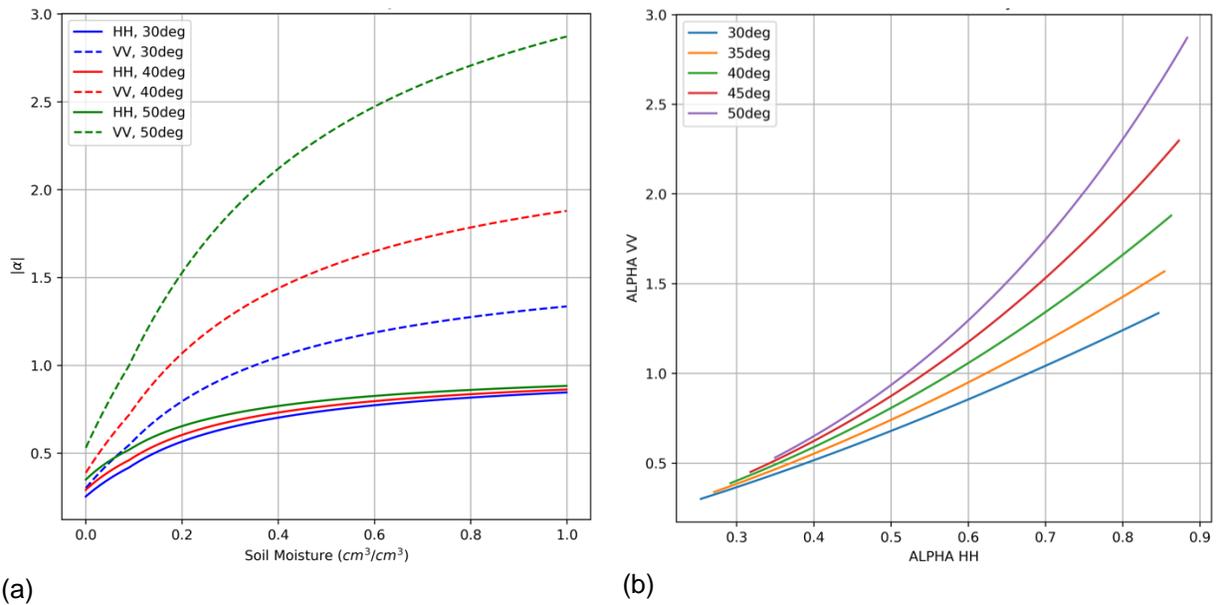


Fig. 4.2.1. The relationship between  $|\alpha_{HH}|$  (a) and  $|\alpha_{VV}|$  (b) with various incidence angles from 30 to 50 degrees. The Mironov model is used in this analysis and clay fraction is fixed at 0.2.

Fig. 4.2.1 shows  $|\alpha_{HH}|$ ,  $|\alpha_{VV}|$  and their relationship when clay fraction is fixed at 0.2. The Mironov model (Mironov et al. 2009) is used in this formulation. Both alpha functions are monotonically increasing with respect to the soil moistures but the dynamic range of VV polarization is much wider than that of HH polarization.

A time-series ratio method assumes that the surface roughness and vegetation properties remain almost constant over two consecutive measurements. If the roughness parameters and incidence angle remain unchanged, the ratio of consecutively measured NRCS values at time  $t_1$  and  $t_2$  can be approximated as

$$\frac{\sigma_{PP}^0(t_2)}{\sigma_{PP}^0(t_1)} \approx \left| \frac{\alpha_{PP}(t_2, \epsilon, \theta)}{\alpha_{PP}(t_1, \epsilon, \theta)} \right|^2, PP = HH \text{ or } VV \quad (4.2.2)$$

With a time-series of N NRCS observations, N-1 ratio values are obtainable, and the matrix equation can be constructed as

$$\begin{bmatrix} 1 & -\sqrt{\frac{\sigma_{PP}^0(t_1)}{\sigma_{PP}^0(t_2)}} & \cdots & 0 & 0 \\ 0 & 1 & \ddots & 0 & 0 \\ & \vdots & \ddots & \vdots & \\ & 0 & 0 & 1 & -\sqrt{\frac{\sigma_{PP}^0(t_{N-1})}{\sigma_{PP}^0(t_N)}} \\ & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} |\alpha_{PP}(t_1)| \\ |\alpha_{PP}(t_2)| \\ \vdots \\ |\alpha_{PP}(t_N)| \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(4.2.3)

where  $PP$  can be either HH or VV polarization. The linear least-squares problem of the combined  $N-1$  by  $N$  matrix equation is solved for the alpha coefficient at each time step. The matrix can also be extended with both HH and VV components for dual polarization scenarios. Also, since this matrix is not fully determined, additional information on the maximum and minimum alpha coefficients is incorporated for solving the problem. Finally, soil moisture is inverted from the alpha coefficients using dielectric mixing model for each polarization configuration.

Considering many samples from longer time series in a matrix equation could obtain more accurate results for soil moisture, but it is more likely to include errors due to vegetation changes including crop growth and so on. In the NISAR mission, the latency between radar acquisitions is 6 to 12 days. Due to this vegetation effect, only several samples up to 3 could be used for the NISAR time-series soil moisture retrieval.

Because the matrix equation is underdetermined, it can be solved only by incorporating additional information into the retrieval. Past works [Ouellette et al., 2017; Mattia et al., 2009; Mattia et al., 2018; He et al., 2017] describe methods for providing this information in terms of maximum and minimum bounds on the retrieved “alpha” coefficients, with these bounds determined through an iterative process, through the incorporation of ancillary information on soil moisture limits potentially from remote sensing observations (e.g., SMAP) or from a climatology of soil moisture behaviors. For the NISAR algorithm, soil moisture bounds from multi-year SMAP radiometer measurements (a SMAP “climatology”) will be used as a baseline in setting bounds for each coarse resolution grid cell.

Figure 4.2.3 compares in-situ soil moistures with values retrieved with this approach using measurements of the L-band SMAP radar, at a spatial resolution of 3 km and at cal/val sites for the SMAP radar. The results show good performance even for varying terrain classes and vegetation contents. The comparison overall showed a correlation between retrieved in-situ soil moistures of  $\sim 0.7$  with unbiased RMS differences of  $\sim 0.06-0.07 \text{ m}^3/\text{m}^3$ . Given the overall  $\sim 0.06 \text{ m}^3/\text{m}^3$  goal typically expected for radar soil moisture retrievals, these results demonstrate the potential of the time-series method. The unbiased RMS were significantly better for VV polarization than HH polarization (HH: ubRMSE= $0.067 \text{ m}^3/\text{m}^3$ ,  $R=0.727$ ; VV: ubRMSE= $0.054 \text{ m}^3/\text{m}^3$ ,  $R=0.809$ ).

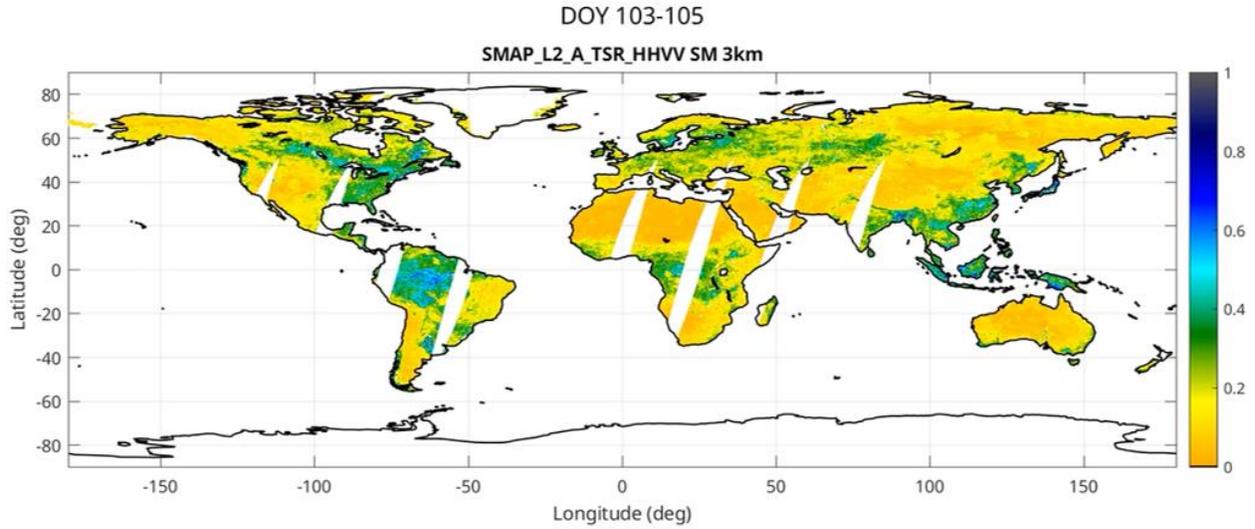


Figure 4.2.2: Example results from the “time series alpha” method with SMAP’s L-band radar measurements from 2015.

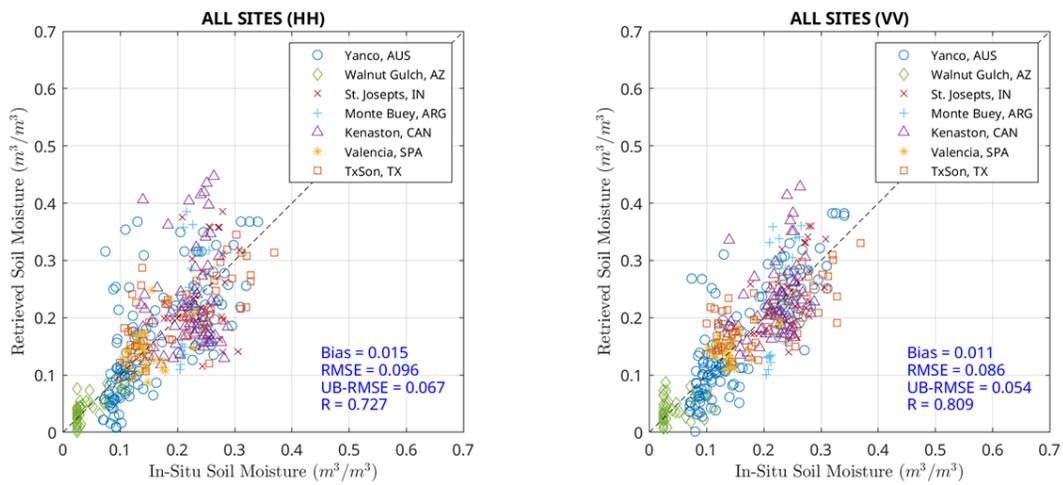


Figure 4.2.3. TSR retrievals compared to in-situ soil moisture measurements at SMAP radar cal/val sites.

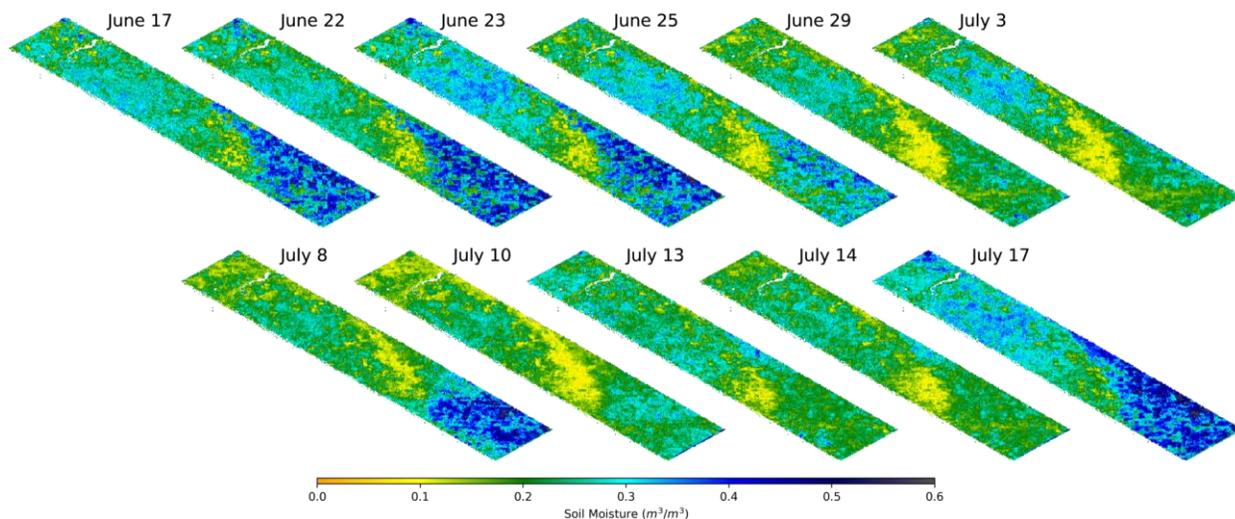


Fig. 4.2.4. Composite retrieved soil moisture images from June 17 to July 17, 2012. 11 common flight lines of 31604 and 31606 were used for the images.

The studies also considered the use of multiple co-pol measurements as well as the incorporation of cross-polarization, but found that no significant gains were achieved on single-polarization (VV) retrievals with this method. Therefore the method is well suited to its application in situations having any polarization measurement mode. In addition, Mattia et al. [2009] have demonstrated methods for using cross-pol measurements as a “pre-screener” for the algorithm to flag regions for which the required assumption of surface-attenuated scatter may be less valid. Generally the approach should be expected to be more applicable for less vegetated scenes, although previous studies have shown that the approach can remain applicable even for vegetation water contents up to  $\sim 5 \text{ kg/m}^2$  for some crop classes.

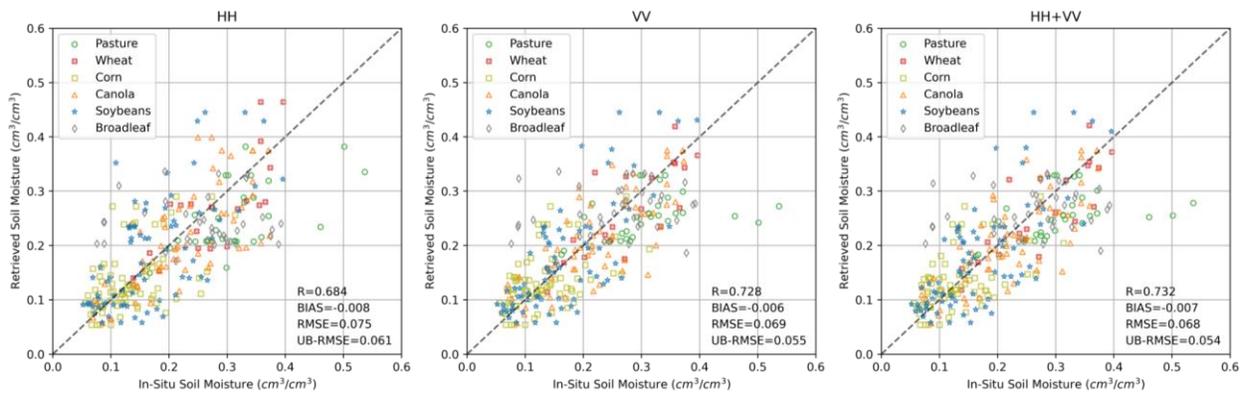
For pre-launch algorithm implementation and validation, datasets from the SMAP Validation Experiment 2012 (SMAPVEX12, Winnipeg in Manitoba, Canada) field campaign acquired with NASA’s airborne L-band UAVSAR instrument were analyzed (Park et al, 2022). During June and July 2012, in-situ observations of soil moisture and vegetation over agricultural and forest areas were collected on the ground. Fig. 4.2.3 shows 11 composite soil moisture images retrieved by time-series ratio algorithm for HH+VV configuration. For the validation of the algorithm, the environmental conditions during the UAVSAR measurements shown in Fig. 4.2.2 could be compared to these global images. First, the UAVSAR experiment started on June 17 just right after the heavy rain. The average in-situ soil moisture was highest at the start of the experiment. After a few days, there were some light rain events, and we could still see very high soil moistures in agricultural areas of the south. Since June 23, there was no precipitation for 10 days and soil moistures have rapidly decreased during the dryer period until July 3. After that, there were few rain events on July 4, July 12, and July 14, the soil moistures (globally or locally) responded to this. From the images, the retrieved soil moisture images are consistent with the in-situ precipitation observations as well as the averaged in-situ soil moistures over the sites. The southern agricultural areas were more affected especially and forest areas were less sensitive to rain events.

For the retrieval performance, the algorithm was assessed using in-situ soil moisture information. Fig. 3.2.4 provides scatter plots comparing retrieved values to those measured at in-situ sites using all 11 days of data for each flight line. In the scatter plot, each crop type is marked with a separate color and symbol for comparison. The statistics of the retrievals including correlation coefficient (R), bias, RMSE, and un-biased RMSE are indicated at the bottom of each figure. For un-biased RMSE, site-based bias correction approach

was used in this analysis. Retrievals show overall reasonable performances for all the methods. In particular, the method using HH+VV combination shows the best performance with the correlation of 0.796 and unbiased RMSE of 0.054. Also, VV performs better than HH because its alpha function shows better sensitivity to the soil moisture.

The crop types of oats, wheat, corn, and canola show overall reasonable performance with unbiased RMSE less than 0.064. However, in the case of soybeans, performance is different depending on the method and unbiased RMSE varies from 0.067 to 0.074. Also, forest sites have good unbiased RMSE, but very low correlations. More studies regarding highly vegetated areas are necessary in the following study. In short, overall performances show unbiased RMSE lower than 0.062 and correlation higher than 0.703.

1. 1604 Flight Line



2. 31606 Flight Line

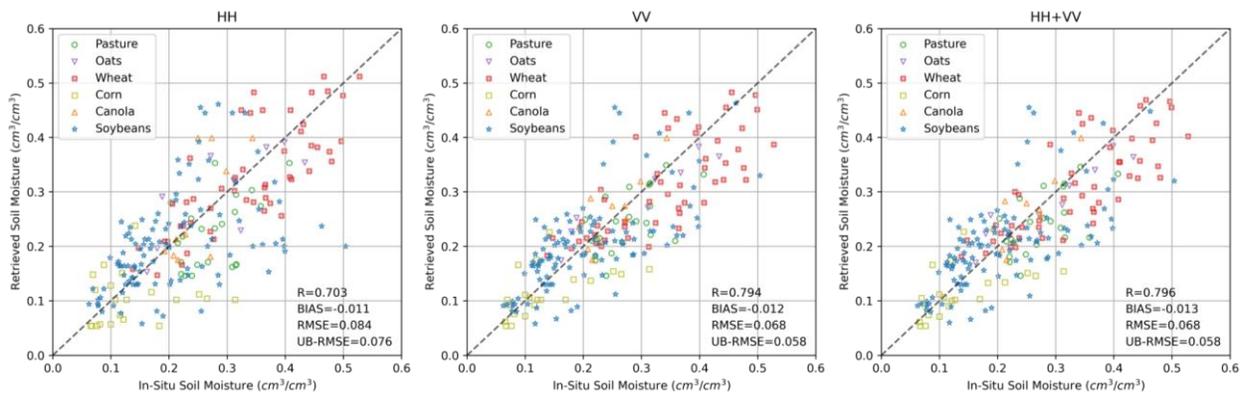


Fig. 4.2.4. Scatter plots between in-situ soil moistures and retrieved soil moistures for HH, VV, and HH+VV combination. (a) 31604 Flight Line and (b) 31606 Flight Line

Due to the limitations of the method to a subset of terrain classes, it should be expected that the algorithm is designed to be combined with other approaches classes for which it is not applicable. The algorithm's focus on "canceling" roughness and vegetation effects however makes it highly desirable for low to moderate vegetated regions.

### 4.2.1 Application to NISAR measurements

The existing time-series ratio algorithm will require some adaptation for use with NISAR data. First, the algorithm should be expected to perform best when time series measurements are acquired at identical incidence angles. This implies a time series step of 12 days when using a times series of NISAR ascending-only or descending-only measurements. The assumption of near-constant vegetation and roughness parameters may be challenged in this situation, but preliminary analyses have shown that the algorithm should still provide acceptable retrievals under most conditions. Time-series retrievals can be conducted both for ascending-only and descending-only datasets, so that retrievals are still available in a 6 day cadence.

An optimal duration for the time series must also be determined. Note that the approach assumes only that vegetation and roughness effects are near constant between two measurements, not throughout the entire time series. Therefore the approach can be applied for an arbitrary time series length, and preliminary results indicate that performance should improve with time series length. Current studies are emphasize ~monthly times series (i.e., 3 measurements over 36 days) but time series up to seasonal scales (i.e., 8 measurements over 96 days) are also under consideration. Note that the ancillary maximum/minimum bound information is required only for the time series length considered.

The time series nature of the algorithm also allows for two types of retrievals: “rapid” and “averaged”. The “rapid” product is obtained using the most recent NISAR measurement as the final point in the time series, so that it is retrieved only from a single time series. This product is desirable because it uses only past measurements and therefore avoids any extended latency required to obtain the required time series. The “averaged” product alternately is obtained after sufficient NISAR measurements have been acquired to allow a point of interest to be retrieved as the first, second, . . . , and final point in the time series. The multiple retrievals so obtained can then be averaged to obtain a second product that is expected to have improved performance (but a latency equal to the time series length). Both products are under development; the “averaged” product is in fact a result of the repeated runs of the “rapid” product necessary for its computation. The “averaged” soil moisture product will be updated in the subsequent NISAR overpass on 12 and 24 days (T-12 days and T-24 days) output using a longer time-series observations.

#### 4.2.2 Performance of the Time Series Algorithm

The soil moisture retrieval error expected from the alpha approximation method can be estimated using the assumption that the minimum bound for  $\alpha_{pq}$  is applied in the matrix equation solution. The successive ratios of a time series of N measurements can be combined to show that the  $\alpha_{pq}$  value at time t is related to that at an earlier time “zero” through

$$\alpha_{pq}^t = \alpha_{pq}^0 \frac{\sigma_{pq}^t}{\sigma_{pq}^0} \quad (4.2.2.1)$$

To apply the minimum bound in the time series solution, we assume in what follows that time “zero” is labeled as the time with the minimum NRCS value, for which it is assumed that the minimum bound  $\alpha_{pq}^0$  is applicable. Note that this time zero may then occur at any point in the time series, with the solution at all other time series points then determined from Eqn. 4.2.2.1.

As previously discussed, the  $\alpha_{pq}$  coefficients obtained from Eqn. 4.2.2.1 are then inverted into soil moisture  $s_m$ :

$$s_m = f_{inverse}(\alpha_{pq}^t) \quad (4.2.2.2)$$

where  $f_{inverse}$  is the inverse of the function relating  $\alpha_{pq}$  to soil moisture for a specified soil texture. The error in soil moisture estimation can be derived from the error in the  $\alpha_{pq}$ . The final expression for the error in soil moisture  $\Delta s_m$  due to speckle contributions can be expressed as a function of the minimum bound  $\alpha_{pq}^0$  and the mean NRCS at time t and at time 0 as:

$$\Delta s_m = df_{inverse}(\alpha_{pq}^t) \alpha_{pq}^t \sqrt{\frac{m_t^2}{m_0^2} \left( \frac{2}{N_{looks}} \right)} \quad (4.2.2.3)$$

Eqn. 4.2.2.3 shows that the error obtained depends on the soil texture and soil moisture values at time  $t$ , the minimum bound  $\alpha_{pq}^0$  as well as the number of looks in the 200 m product. It is noted that the term  $\frac{m_t}{m_0}$  in

Eqn. 4.2.2.3 is equivalent to  $\frac{\alpha_{pq}^t}{\alpha_{pq}^0}$ , so that Eqn. 4.2.2.3 could be further simplified. However, this simplification is avoided at this point in order to enable an inclusion of additional errors that account for uncertainties in the minimum bound applied. These errors in the minimum bound are modeled by describing  $\alpha_{pq}^0$  as its mean value multiplied by a Gaussian random variable with mean equal to 1 and a specified standard deviation. This results in additional error contributions in Eqn. 4.2.2.3 that are included in what follows.

The error model can be simulated over larger regions by incorporating maps of soil texture and applying a climatology of expected soil moisture values. Such information has been previously compiled for the Soil Moisture Active Passive (SMAP) mission, and can be used to develop a large scale error simulation. The required climatology (initially at 40 km spatial scale) has been developed using 6 years of SMAP soil moisture products to determine bi-monthly minimum, maximum, and average soil moisture maps.

The maps in Figure 4.2.2.1 represent the error in soil moisture that is expected in summer for HH (top row) and VV (bottom row) including speckle noise only (left column) and including speckle noise and uncertainty on the minimum bound  $\alpha_{pq}^0$  (right column). The error in soil moisture is found to be larger in HH polarization than VV polarization as expected. Adding uncertainties on the minimum bound  $\alpha_{pq}^0$  also increases the error in soil moisture for both polarizations. However, the performance of the time series algorithms stays in most places below the requirement of the NISAR mission.

Speckle Noise Only	Speckle noise + 2.5% Uncertainty on $\alpha_{pq}^0$
--------------------	---

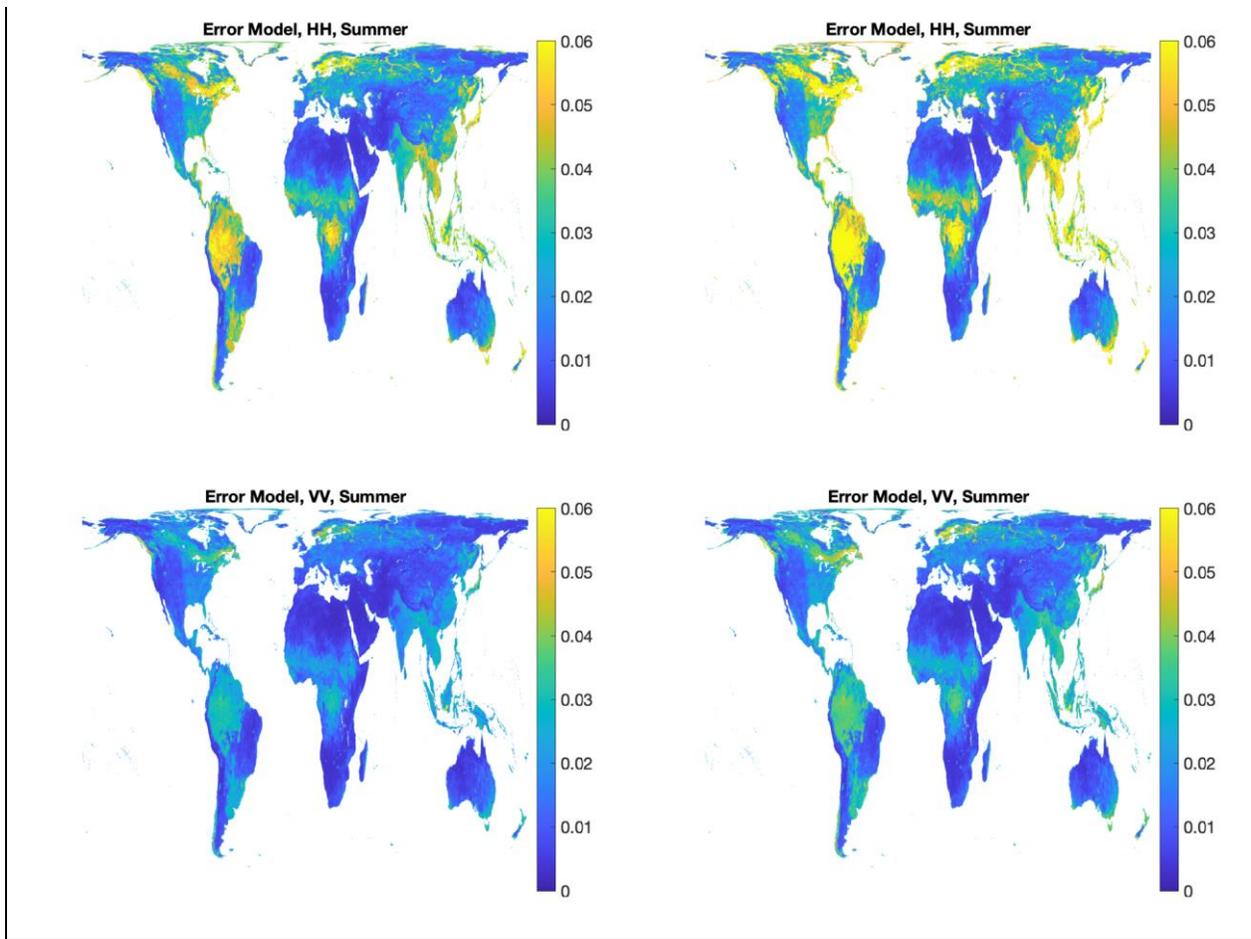


Figure 4.2.2.1 Unbiased RMSE error (m<sup>3</sup>/m<sup>3</sup>) expected during summer for HH (top row) and VV (bottom row) including speckle noise only (left column) and including speckle noise and uncertainty on the minimum bound  $\alpha_{pq}^0$  (right column).

### 4.3 Multiscale Fusion Algorithm

This multiscale fusion algorithm proposed for the NISAR mission takes advantage of existing skills and capabilities of the soil moisture data available at coarser resolution (~9 km). There are many resources available through modeling along with data assimilation and satellite-based retrievals, such as surface soil moisture data from the near real time forecast of the European Center for Medium-Range Weather Forecasting (ECMWF) [Muñoz-Sabater et al., 2021] and soil moisture retrievals from the Soil Moisture Active Passive (SMAP) mission [Chan et al., 2018; Das et al., 2018; Entekhabi et al., 2010], respectively. The accuracy, i.e., unbiased root-mean-square-error (ubRMSE) of the ~9 km soil moisture from these above-mentioned resources is generally (~0.04 m<sup>3</sup>/m<sup>3</sup>) with various biases depending upon landcover [Albergel et al., 2012; Chan et al., 2018, Lal et al., 2022]. At very high resolution (10 m) L-band SAR backscatter observations also carry a distinct signature of soil moisture (10 m). However, the SAR data at its native resolution has speckle noise, thus, aggregating SAR data to 100 m or 200 m does improve the soil moisture-related signature. We plan to use the NISAR co-pol ( $\sigma_{HH}$ ) and cross-pol ( $\sigma_{HV}$ ) observations aggregated to 200 m and expect to reduce the speckle (white) noise nearly ~20 times ( $1/\sqrt{20 * 20}$ ). Thus,

the proposed soil moisture multiscale fusion algorithm takes advantage of high-resolution information of soil moisture in SAR observations and uses them to disaggregate the coarse-resolution soil moisture data.

The proposed multiscale fusion algorithm has a legacy from the SMAP mission algorithm development [Das et al., 2010]. The algorithm presented subsequently is modified according to accommodate the heterogeneity at high-resolution (200 m) within the coarse-resolution (~9 km) as captured by the SAR cross-pol ( $\sigma_{HV}$ ) backscatter observations. The algorithm is elaborated as follows.

For clarity, we first define general grid topologies, mathematical operators, and terms used in the algorithm mathematical formulation. Figure 4.3.1 elaborates the topology of the coarse-resolution (9 km) and high-resolution (200 m) grids. For the sake of convenience to mathematically formulate the algorithm, the naming convention of coarse-resolution (9 km) grid cells as ‘C’ (coarse) and high-resolution grid cells as ‘F’ (fine) are followed subsequently throughout this section. It is obvious from the grid topology (Fig. 4.3.1) that an area grid cell ( $nc = 1$ ) of C encompasses 2025 area grid cells ( $nf = 2025$ ) of F. Two linear operators used frequently in the mathematical formulation are defined as Space-Average operator:  $\langle x \rangle = \frac{1}{A} \int x da$ , where A is the area of larger pixel and a is the area of a smaller pixel within A, and the spatial anomaly operator  $\delta x = x - \langle x \rangle$ .

The relationship between SAR co-pol backscatter ( $\sigma_{HH}$  at L-band and C-band) and volumetric soil moisture are reported by [Kim and van Zyl, 2009; Njoku and Entekhabi, 1996; Ulaby et al., 1986]. At L-band, Kim and van Zyl, [2009] found a nearly linear relationship using truck-mounted L-band radar data collected during the Washita 92 field experiment. A linear relationship between radar backscatter ( $\sigma_{HH}$  at L-band) and volumetric soil moisture is also reported for the airborne L-band remote sensing data of Soil Moisture Experiment, 2002 (SMEX02) by Piles et al., [2009]. The formulation of the algorithm is based on such a linear relationship. However, we use the hypothesis that the volumetric soil moisture and co-polarized backscatter ( $\sigma_{HV}$ ) are linearly related through

$$SM(t) = \alpha + \beta \log[\sigma_{HH}(t)] \tag{4.3.1}$$

Here  $\log$  represents  $10.\log_{10}$ .

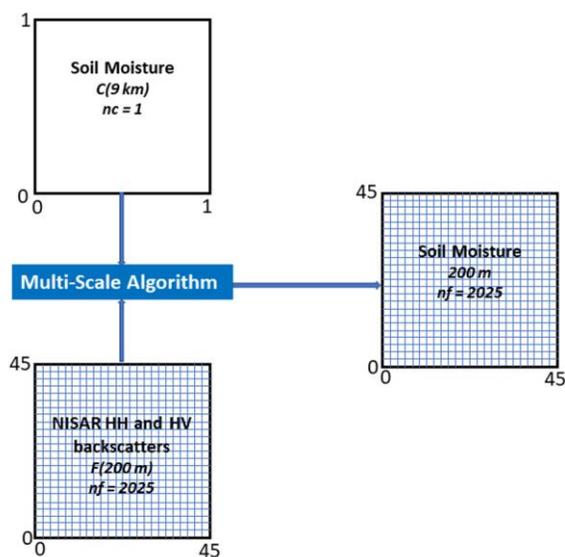


Figure 4.3.1: EASE2 grid topology of the proposed multiscale fusion algorithm.

At a given scale (C or F), in Eqn. 4.3.1,  $\alpha$  and  $\beta$  are parameters that depend on vegetation cover and type, surface roughness, and incidence angle of  $\sigma_{HH}(t)$ . However, using the 12 days exact repeat and not mixing the ascending and descending overpasses makes the time series independent of incidence angle, as all the backscatter observations in the time series over a particular grid cell have the same incidence angle. The parameters  $\alpha$  and  $\beta$ , especially  $\beta$ , vary seasonally and can be estimated at scale C using coarse-resolution soil moisture and aggregated NISAR  $\sigma_{HH}$  time-series as regressant and regressors, respectively, as in

$$SM(C, t) = \alpha(C) + \beta(C) \log < \sigma_{HH}(F_n, t) > \quad (4.3.2)$$

Here  $\sigma_{HH}(F_n, t)$  is co-polarized radar backscatter at a spatial scale of  $F$ ,  $SM(C, t)$  is volumetric soil moisture at a spatial scale of  $C$ . The parameters  $\alpha(C)$  and  $\beta(C)$  can be obtained by temporal sequences regression of  $SM(C, t)$  and  $\log < \sigma_{HH}(F_n, t) >$ . Formulation of the algorithm begins with the hypothesized linear relationship between volumetric soil moisture and co-polarized radar backscatter at spatial scale  $F$ , so that Eqn. 4.3.1 can be written as

$$SM(F_n, t) = \alpha(F_n) + \beta(F_n) \log(\sigma_{HH}(F_n, t)) \quad (4.3.3)$$

where  $\alpha(F_n)$  and  $\beta(F_n)$  are the parameters at spatial scale of  $F$ , and  $SM(F_n, t)$  is a soil moisture value at a spatial scale of  $F$  and at time  $t$  for a particular pixel within  $C$ , and is the target variable to estimate through this algorithm. Subtracting Eqn. 4.3.2 from Eqn. 4.3.3, gives

$$SM(F_n, t) - SM(C, t) = \alpha(F_n) + \beta(F_n) \log(\sigma_{HH}(F_n, t)) - \alpha(C) - \beta(C) \log < \sigma_{HH}(F_n, t) > \quad (4.3.4)$$

Because  $SM(F_n, t)$  is not available, we cannot estimate the parameters  $\alpha(F_n)$  and  $\beta(F_n)$  in the manner that was followed at scale  $C$ . The path forward to incorporate the effects of the variations of these parameters at scale  $F$  with respect to the coarser scale begins with algebraically rewriting Eqn. 4.3.4 by substituting  $\sigma_{HH}(C, t) = \log < \sigma_{HH}(F_n, t) >$  and  $\sigma_{HH}(F_n, t) = \log(\sigma_{HH}(F_n, t))$ ,

$$\begin{aligned} & SM(F_n, t) \\ &= SM(C, t) + \hspace{15em} \text{RHS Term I} \\ & \quad \{\beta(C) \cdot [\sigma_{HH}(F_n, t) - \sigma_{HH}(C)]\} + \hspace{10em} \text{RHS Term II} \\ & \quad \{[\alpha(F_n) - \alpha(C)] + [\beta(F_n) - \beta(C)] \cdot \sigma_{HH}(F_n, t)\} \hspace{5em} \text{RHS Term III} \end{aligned} \quad (4.3.5)$$

The left-hand side of Eq. (4.3.5) is the target variable of the algorithm, i.e., the disaggregated fine resolution soil moisture at 200 m resolution. The first term on the right-hand side (RHS Term I),  $SM(C, t)$ , is the coarse resolution soil moisture from ECMWF at scale  $C$  (9 km).

The RHS Term II,  $\{\beta(C) \cdot [\sigma_{HH}(F_n, t) - \sigma_{HH}(C)]\}$ , can be calculated based on the regression parameter  $\beta(C)$  that is estimated through the time-series of  $SM(C, t)$  from ECMWF and NISAR observation aggregated  $\sigma_{HH}(C)$  to scale  $C$ . The remainder of this second RHS term,  $[\sigma_{HH}(F_n, t) - \sigma_{HH}(C)]$  is also based on the NISAR observations at scales  $F$  and  $C$ .

The RHS Term III accounts for the deviations of the parameters  $\alpha$  and  $\beta$  within the grid cell  $C$ . The term  $\{[\alpha(F_n) - \alpha(C)] + [\beta(F_n) - \beta(C)] \cdot \sigma_{HH}(F_n, t)\}$  is in units of volumetric soil moisture and represents subgrid scale (relative to  $C$ ) heterogeneity effects. The parameters  $\alpha$  and  $\beta$  depend on vegetation and surface roughness. For a perfectly homogeneous region, the parameters  $\alpha(F_n) = \alpha(C)$  and  $\beta(F_n) = \beta(C)$ , and the subgrid heterogeneity term becomes zero. However, in nature homogeneity within  $C$  rarely exists.

NISAR also provides high-resolution cross-polarization radar backscatter measurements at scale  $F$  which are principally sensitive to vegetation and surface characteristics. The subgrid deviation/heterogeneity patterns in vegetation and roughness as captured by the cross-polarization backscatter at scale  $F_n$  is  $[\sigma_{HV}(C) - \sigma_{HV}(F_n)]$ . This indicator can be converted to variations in co-polarization backscatter through

multiplications by a sensitivity parameter  $\left[ \frac{\partial \sigma_{HH}(F_n)}{\partial \sigma_{HV}(F_n)} \right]_C$ . This sensitivity, denoted by the scale  $C$  variable  $\Gamma \equiv \left[ \frac{\partial \sigma_{HH}(F_n)}{\partial \sigma_{HV}(F_n)} \right]_C$ , is specific to the particular grid cell  $C$  and the particular season for grid cell  $C$ .

$\Gamma$  is estimated based on the collection of co-polarization and cross-polarization radar backscatter cross-section within each grid cell  $C$ . The term  $\Gamma \cdot [\sigma_{pq}(C) - \sigma_{pq}(F_j)]$  is the projection of the cross-polarization subgrid deviation onto the co-polarization space. These variations are due to the heterogeneity in parameters  $\alpha$  and  $\beta$  in the radar co-polarization space. It can be converted to soil moisture volumetric units for use in Eqn. 4.3.5 through multiplication by  $\beta(C)$ , the particular radiometer grid scale  $C$  conversion factor relating co-polarization backscatter variations to soil moisture variations. Thus, the product  $\beta(C) \cdot \Gamma \cdot [\sigma_{pq}(C) - \sigma_{pq}(F_j)]$  is the contribution of subgrid (subgrid to scale  $C$ ) variations in  $\alpha$  and  $\beta$  to the soil moisture at scale  $F$ . The NISAR soil moisture disaggregation algorithm is completed by substituting the term  $\beta(C) \cdot \Gamma \cdot [\sigma_{pq}(C) - \sigma_{pq}(F_j)]$  to RHS Term III in Eqn. 4.3.5,

$$SM(F_n, t) = SM(C, t) + \{\beta(C) \cdot [\sigma_{HH}(F_n, t) - \sigma_{HH}(C)]\} + \beta(C) \cdot \Gamma \cdot [\sigma_{HV}(C) - \sigma_{HV}(F_j)] \tag{4.3.6}$$

which can be written more compactly as

$$SM(F_n, t) = SM(C, t) + \beta(C) \cdot \{[\sigma_{HH}(F_n, t) - \sigma_{HH}(C)] + \Gamma \cdot [\sigma_{HV}(C) - \sigma_{HV}(F_j)]\} \tag{4.3.7}$$

Eqn. 4.3.7 is the NISAR soil moisture multiscale fusion algorithm. *The most important aspect of this algorithm is the very low degree of freedom, as no ancillary data is required to operate this algorithm. Its is purely an input data driven algorithm given that  $\beta(C)$  is predetermined.* The details of algorithm implementation is shown in Fig. 4.3.2.

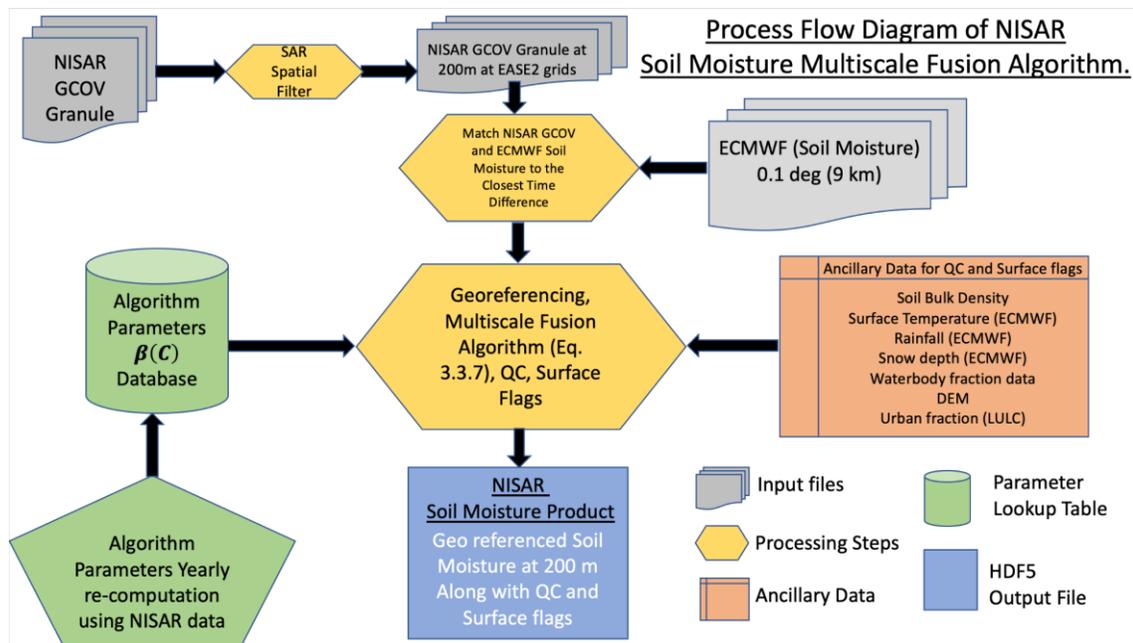


Figure 4.3.2: Process flow diagram of the multiscale fusion algorithm.

### Estimation of the $\beta$ and $\Gamma$ Parameters

The accuracy of the sensitivity parameter  $\beta$  and the heterogeneity parameter  $\Gamma$  is important for the performance of the NISAR soil moisture disaggregation algorithm. To estimate the sensitivity parameter  $\beta$  at scale  $C$  time series data is required. With all the available SAR L-band data through satellite missions and airborne platforms (such as ALOS PALSAR2 and UAVSAR) no longer temporal time series of concurrent SAR observations and soil moisture estimates at scale  $C$  over a global extent can be created, except using the SMAP mission ~3 months L-band SAR and the concurrent ECMWF soil moisture. Therefore, the SMAP L-band SAR and ECMWF soil moisture are used to understand the assumption of linearity between  $SM$  and  $\sigma_{HH}$  over various landcover and soil moisture conditions. Figures 4.3.3A and 4.3.3B illustrate a spatial map of the correlation between the SMAP L-band SAR and ECMWF soil moisture and sensitivity parameter  $\beta$  at scale  $C$  over a global extent, respectively. As shown in Fig. 4.3.3A high correlation ( $>0.6$ ) is observed in most parts of the world except over deserts and forests because during the three months duration there was not much temporal variability, leading to low correlation. However, a much longer time series data is required to get desired statistically significant correlation over these landcovers. The spatial evolution sensitivity parameter  $\beta$  looks reasonable in Fig 4.3.3B. For the NISAR mission prelaunch soil moisture multiscale fusion algorithm, we created a look-up table for different landcovers. Figure 4.3.4 shows the median and standard deviation over different landcovers. After the NISAR launch, we plan to keep updating  $\beta$  seasonally and yearly to create a database for each and every  $C$  scale grid cell. This exercise will help improve the quality of  $\beta$  after the NISAR launch.

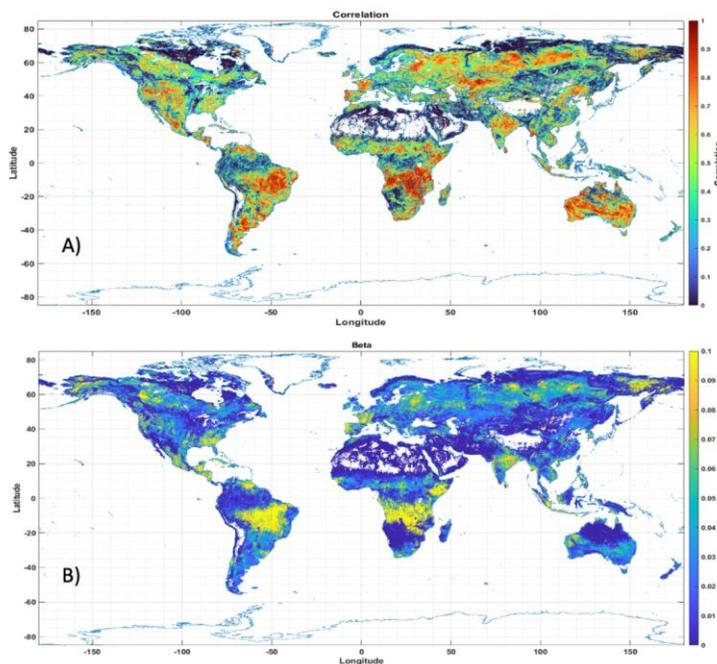


Figure 4.3.3: A) Correlation at  $C$  scale between the SMAP L-band SAR and ECMWF soil moisture using data from March 1<sup>st</sup> 2015 to July 7<sup>th</sup> 2015, and; B)  $\beta$  parameter at scale  $C$  globally, obtained using the SMAP L-band SAR and ECMWF soil moisture using data from March 1<sup>st</sup> 2015 to July 7<sup>th</sup> 2015.

The algorithm heterogeneity parameter  $\Gamma$  exhibits more temporal stability as compared to the  $\beta$  parameter. Figure 4.3.5 shows the global distribution of the  $\Gamma$  parameter computed using SMAP L-band SAR 3 months data. The range of values of  $\Gamma$  parameter corresponds with the parameters derived from the Soil Moisture Field Campaign (SMAPVEX12)

UAVSAR data. To evaluate the stability of the  $\Gamma$  parameter, the coefficient of variation was computed for a one-month period as shown in Fig. 4.3.6. The coefficient of variation is very low for most part of the world

suggesting stability in derived  $\Gamma$  parameter. It is obvious from the algorithm formulation that to estimate  $\Gamma$  parameter no time-series data is required and can be computed on-the-fly using all the  $\sigma_{HH}(F_n)$  and  $\sigma_{HV}(F_n)$  data with the  $C$  scale grid cell. In an ideal condition there will be 2025 combination of  $\sigma_{HH}(F_n)$  and  $\sigma_{HV}(F_n)$  to robustly estimate a statistically significant  $\Gamma$  parameter.

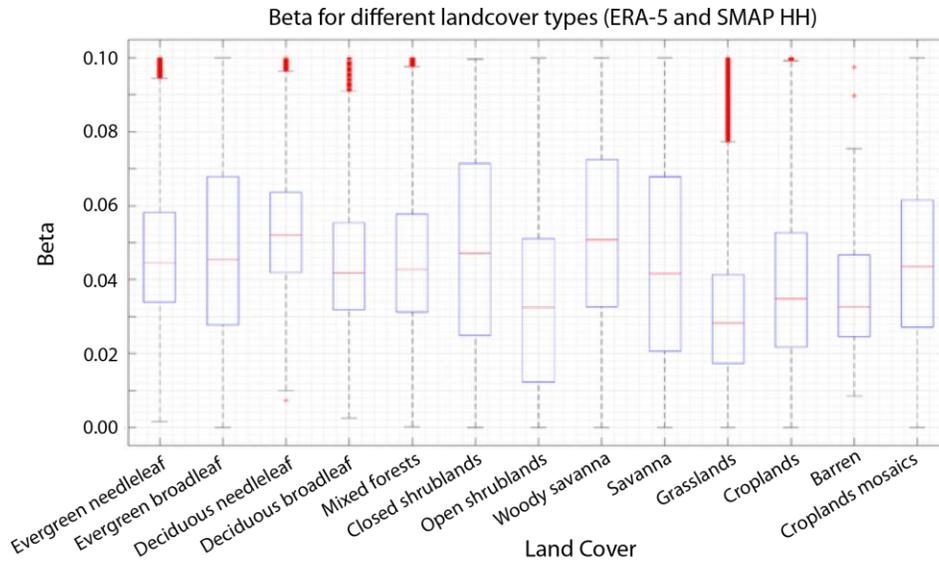


Figure 3.3.5  $\beta(C)$  estimates for each landcover type created from the ECMWF ERA-5 Land Soil Moisture and SMAP L-band SAR  $\sigma_{HH}$ .

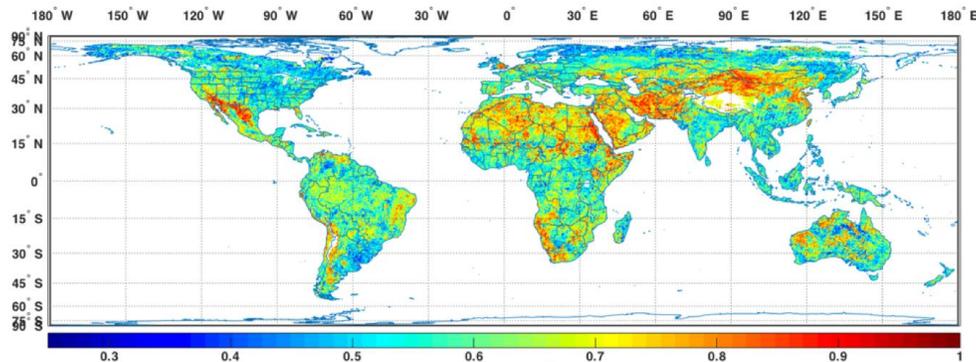


Figure 3.3.6: Map of  $\Gamma$  parameter at global extent averaged for 04-28-2015 to 05-28-2015.

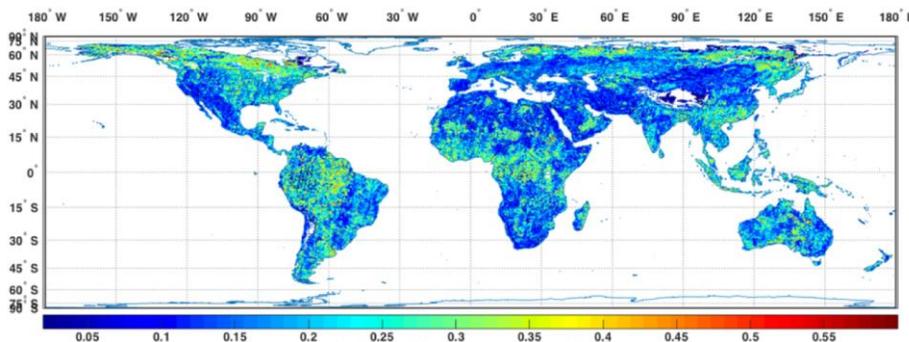


Figure 3.3.7: Coefficient of variation of  $\Gamma$  parameter computed for 04-28-2015 to 05-28-2015.

### 4.3.1 Testing of Multiscale Fusion Algorithm

The NISAR soil moisture multiscale fusion algorithm developed as Eqn. 4.3.7 and the uncertainty estimates (from Eqn. 4.3.12, Eqn. 4.3.19, and Eqn. 4.3.20) are tested and validated in limited conditions over various airborne- and satellite-based SAR L-band observations. Primarily data from UAVSAR and ALOS PALSAR2 are used for testing. Although the SAR data from UAVSAR and PALSAR2 are not identical to data from the NISAR L-band instrument, these datasets do provide L-band SAR data over a large spatial extent and different landcovers. For example, the expected noise floor of NISAR L-band  $\sigma_{HH}(F_n)$  and  $\sigma_{HV}(F_n)$  backscatter data are expected to be inferior to UAVSAR and PALSAR2 observations. Despite such differences, the SAR data from NISAR will be not used as full-resolution single-look-complex (SLC) (~10 m resolution) input but will be aggregated to 200 m spatial resolution before soil moisture retrievals are performed. This step reduces the speckle noise by averaging over 400 full-resolution SLC pixels. Any biases associated with the NISAR instrument will not impact the algorithm (Eq. 4.3.12), as the algorithm relies on the spatial deviation of  $\sigma_{HH}(F_n)$  and  $\sigma_{HV}(F_n)$  from their respective mean values ( $\sigma_{HH}(C)$  and  $\sigma_{HV}(C)$ ) within the coarse resolution 9 km grid cell. Figure 4.3.1.1 shows the study areas used to test the algorithm (Eq. 4.3.7). The study areas are from various parts of North America and India over different landcovers.

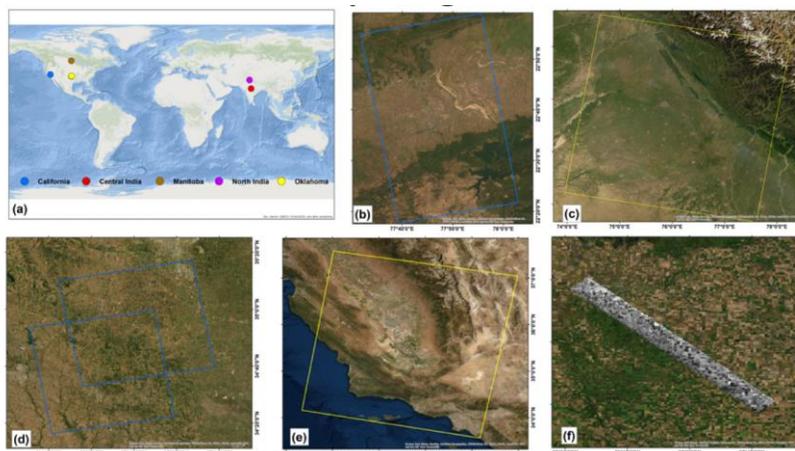


Figure 4.3.1.1: (a) Test sites, ALOS PALSAR 2 coverage for (b) Central India (Fine mode), (c) North India (ScanSAR mode), (d) Oklahoma, USA (Fine mode), (e) California, USA (ScanSAR mode) and, (f) UAVSAR over Manitoba, Canada.

The following describes the characteristics of the study areas:

Central India: The area is in the state of Madhya Pradesh, and has croplands in the middle of the

study area with few rivers and forests and mountains in the northern and southern parts of the study area. The PALSAR2 data is in fine mode.

North India: This large study area covers parts of Punjab, Haryana, Uttarakhand, Uttar Pradesh, and Madhya Pradesh. The study area is primarily dominated by agricultural regions with many urban areas. The Northeastern portion of the study area is mostly forest and foothills of the Himalayan range. The PALSAR data is in SCANSAR mode.

Oklahoma: This study region is mainly grasslands mixed with some agricultural farmlands. The PALSAR2 data is in fine mode.

Southern California: The study region encompasses agricultural landscapes of San Joaquin Valley nestled between the Mountain ranges. The PALSAR data is in SCANSAR mode.

**Manitoba, Canada:** This is the only study area that is covered by UAVSAR data. The study region is totally covered with maize, soybean, and canola farmlands. The UAVSAR data was acquired in the summer of 2012 during the NASA SMAPVEX12 field campaign.

### 4.3.2 Algorithm Implementation with ALOS PALSAR2 and UAVSAR Data

As discussed in the above sections, the input L-band SAR data to the algorithm needs spatial averaging to 200 m. It is obvious from Fig. 4.3.2.1 (top panel) that at the original ~10m resolution the SAR data has significant speckle noise. The undesirable impact of urban areas and waterbodies can also be seen. Therefore, it is essential to conduct spatial filtering before averaging the SAR data. For this purpose, a special hybrid filter [Das et al., 2019] is implemented. The PALSAR2 data after filtering and averaging shows (bottom panel of Fig. 4.3.2.1) a significant reduction of speckles and removal of urban areas and waterbodies.

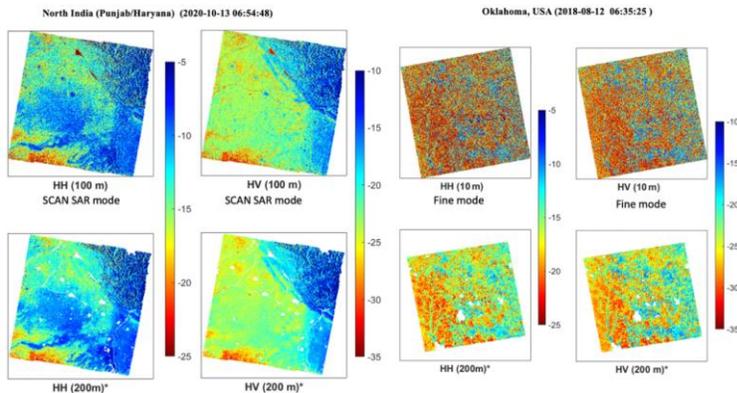


Figure 4.3.2.1: Filtering and spatial averaging of ALOS PALSAR2 data over the study regions.

Figures 4.3.2.2 and 4.3.2.3 illustrate the soil moisture retrievals from the algorithm (Eq. 4.3.7) over the North India, Central India, Southern California, and Oklahoma study regions. The 200 m soil moisture in Figs. 4.3.2.2 and 4.3.2.3 clearly show high-resolution features otherwise not visible in the input coarse resolution 9 km soil moisture (ECMWF) inputs.

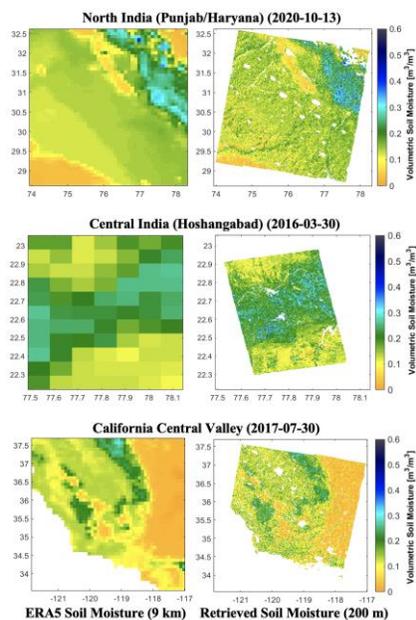


Figure 4.3.2.2: Soil Moisture maps over the study domain (first row for North India and second row for Central India). Left panel: ECMWF ERA-5 land based coarse resolution (9 km) soil moisture; Right panel: high-resolution (200 m) soil moisture estimated using the multiscale fusion algorithm.

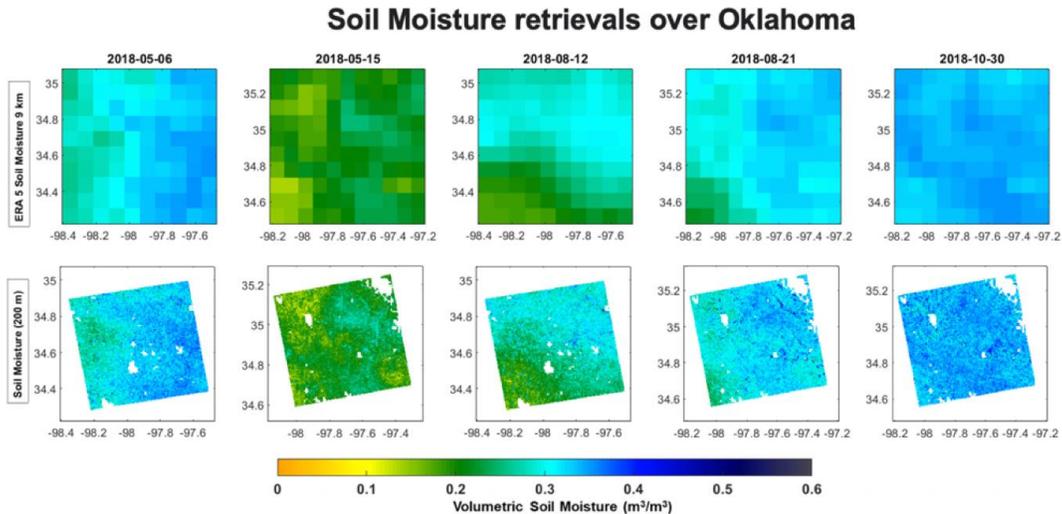


Figure 4.3.2.3: Soil Moisture maps over the Oklahoma region. Upper panel: ECMWF ERA-5 land based coarse resolution (9 km) soil moisture. Middle panel: High-resolution (200 m) soil moisture estimated using the multiscale fusion algorithm.

The implementation of the algorithm over the Manitoba region using UAVSAR data is more challenging than the ALOS PALSAR2 data because the UAVSAR swath is very narrow and does not encompass fully within a 9 km EASE2 grid. There may be some residual errors in the retrieved soil moisture due to this anomaly in the algorithm implementation. However, as shown in Fig. 4.3.2.4, the algorithm performed satisfactorily and the soil moisture pattern over various UAVSAR overpasses successfully captured dry-down and then wetting of the farmland over a month.

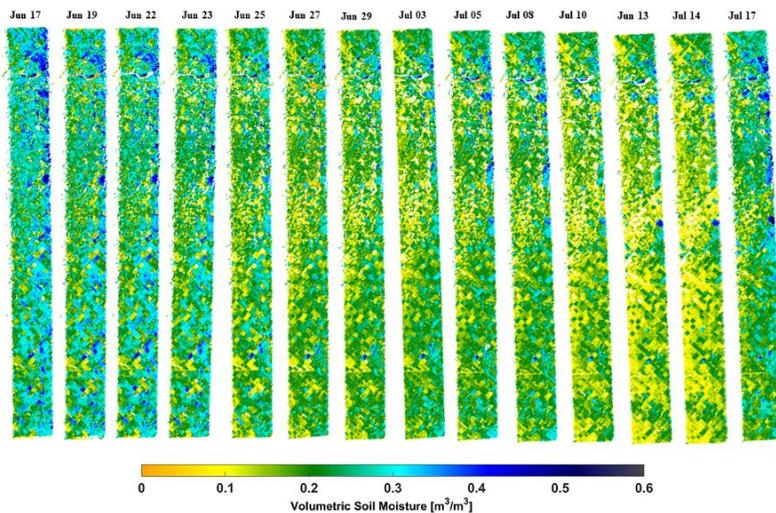


Figure 4.3.2.4: High resolution (200 m) Soil Moisture maps over the region surrounding Southport, Manitoba, Canada.

### 4.3.3 Validation of High Resolution Soil Moisture

A limited validation is conducted for the soil moisture retrievals over the Northern India study region. In situ soil moisture was collected over 13 agricultural fields in the state of Punjab near Ludhiana and coincided with the ALOS PALSAR2 overpass on Oct 13<sup>th</sup>, 2020 (Fig. 4.3.2.2). Here we did an ergodic substitution of

space for time with the assumption that all the agricultural fields are nearby and are almost similar. Thus, the spatially distributed in situ and retrieved soil moisture represent a time series with a range of soil moisture conditions. The comparison in Fig. 4.3.3.1 demonstrated encouraging results with an ubRMSE of  $0.0367 \text{ m}^3/\text{m}^3$  that is within the accuracy goals for the NISAR-based soil moisture product.

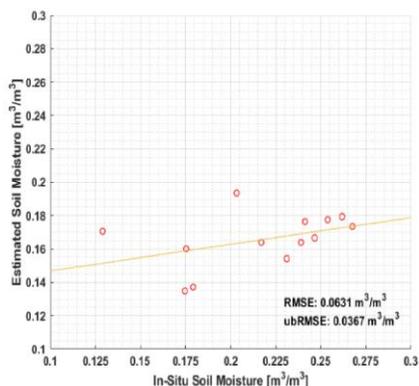


Figure 4.3.3.1: Comparison of in situ soil moisture and PALSAR2-based high-resolution (200 m) soil moisture over 13 agricultural fields in the state of Punjab near Ludhiana.

We also compared the soil moisture retrievals over the California Central Valley as shown in Fig. 4.3.2.2 with the SMAP-Sentinel soil moisture data at 1 km spatial resolution that is available from the same date (07-31-2017) over the study domain. The soil moisture patterns from both data are almost similar. To validate it further, we created the histogram of soil moisture data at 1 km resolution for the overlapping area of both products. The normal distribution and spread of the histograms from both products are the same (almost equal variance), However, the mean of PALSAR2-based soil moisture is slightly higher as expected because of the deeper soil penetration depth of PALSAR2 L-band as compared to the Sentinel1A C-band.

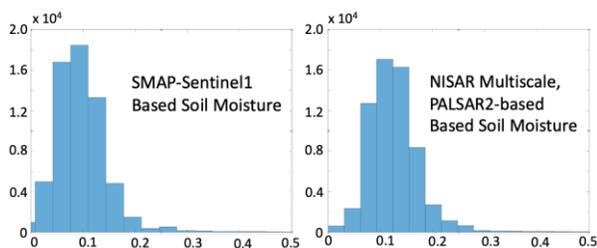


Figure 4.3.3.2: Comparison of soil moisture retrievals from the NISAR multiscale fusion algorithm (with PALSAR2 as input) with the SMAP-Sentinel1A.

Comprehensive validation of the multiscale fusion algorithm was further conducted using the UAVSAR data from the SMAPVEX 2012 experiment conducted in Manitoba, Canada. The test of the algorithm at this site is of particular importance as the UAVSAR observations cover agricultural fields with tremendous heterogeneity in crop types (soybean, maize, sunflower, oats, canola, wheat, and forage). Thus, in the NISAR prelaunch phase, the SMAPVEX 2012 is considered a major site for testing algorithm performance within agricultural landscapes. We extensively validated the retrieved high-resolution soil moisture against the measured in-situ soil moisture over various agricultural fields with different crop types. Detailed soil moisture maps of the UAVSAR flight line 31606 is shown in Fig. 4.3.2.4. The maps show dry and wet conditions and how the multiscale algorithm captures the variability in soil moisture as visible in different fields. The variability in high-resolution surface soil moisture is primarily due to various crop covers, diversity in crop density, variability in soil types, and difference in vegetation-water-content, and these attributes are captured by the L-band SAR and observations. The north-south aligned agricultural fields are clearly seen because of different soil moisture conditions. Such high-resolution soil moisture products will enable the monitoring of individual agricultural fields.

Figure 4.3.3.3 shows the retrieved high-resolution (200 m) soil moisture time series over fields (randomly selected) with different crop covers. The most important aspect to notice in Fig. 4.3.3.3 is that the high-resolution (200 m) soil moisture retrievals closely follow and match the mean of the field in situ measurements for all the days. The retrieved soil moisture even matched closely with the mean of the in-situ soil moisture where the variability is high, such as the agricultural fields that have sunflower, forage, and grassland. Figure 4.3.3.3 also demonstrates another aspect of the algorithm, the retrieved high-resolution soil moisture time series from two different flight lines (31604 and 31606) are almost similar. This

demonstrates that the algorithm is quite capable of retrieving almost the same values of soil moisture irrespective of the difference in incidence angle and azimuth of the SAR observation geometry. Time series plots from various other fields with different landcover were also analyzed (plots not shown in the article) and have similar matching characteristics and results, as shown in Fig. 4.3.3.3.

Figure 4.3.3.4 illustrates the statistics for soil moisture retrievals obtained from UAVSAR flight line 31606 (similar statistics were obtained for flight line 31604, not shown here). The statistics come from all the fields of a certain crop cover. The ubRMSE of the high-resolution (200 [m]) soil moisture retrievals is below the NISAR science goal of 0.06 m<sup>3</sup>/m<sup>3</sup>. As expected, the biases on fields with some crop cover, such as forage and oats, are on the higher side. These high biases can be attributed to wet bias in ERA5-Land coarse resolution soil moisture and a lesser number of observations, as it noted that had fewer than 3 fields were samples and also did not have observations during each UAVSAR overflight. However, the ubRMSE numbers (Fig. 4.3.3.4) over all the fields impart more credence to the multi-scale algorithm as they were calculated under a range of soil conditions (dry to wet) and with crop growth (i.e., increasing vegetation-water-content).

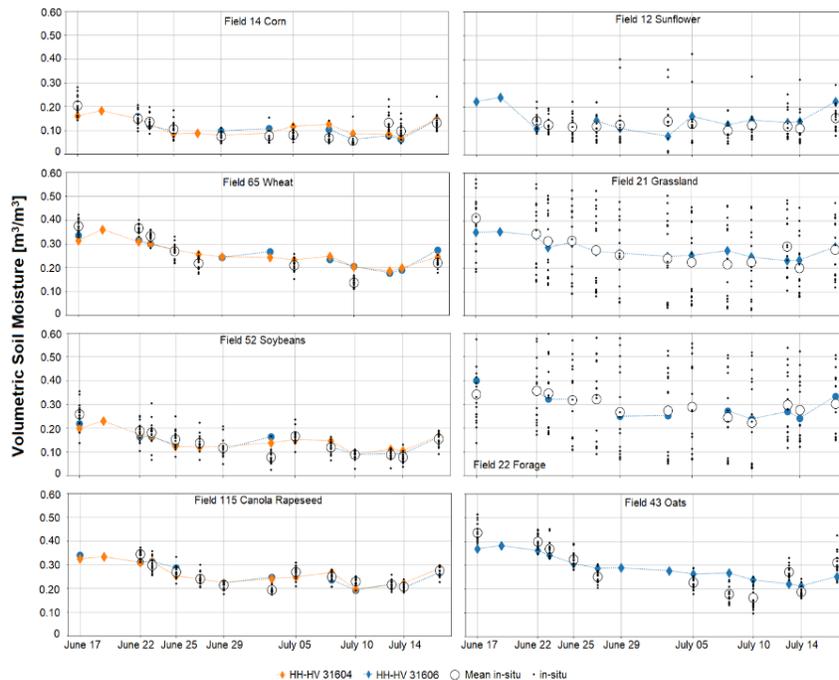


Figure 4.3.3.3: Time series of multiscale fusion algorithm high-resolution (200 m) soil moisture retrievals from two different UAVSAR flight lines (31604 and 31606) over agricultural fields with different crops. Orange and blue symbols show the retrievals from line 31605 and 31606, respectively, and black open circles and small dots show the mean of in situ observations for a given field and the total of all observations for that field.

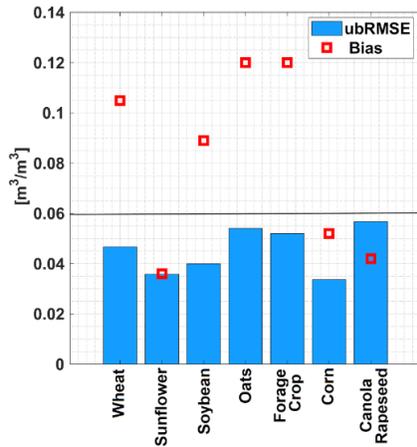


Figure 4.3.3.4: Bias and *ubRMSE* of retrieved high-resolution soil moisture (200 m) for various crop types. Note that some of the crops, such as oats, forage, and broadleaf, had fewer than 3 fields and did not have observations during each UAVSAR overflight. Black solid line indicates the accuracy goals of 0.06 m<sup>3</sup>/m<sup>3</sup> for the NISAR mission.

### 4.3.4 Uncertainty in Soil Moisture from the Multiscale Fusion Algorithm

#### Analytical Approach to Estimate Uncertainty in Algorithm

Uncertainty estimation is a critical part of any algorithm development process. The NISAR soil moisture multiscale fusion algorithm (Eqn. 4.3.7) can be subjected analytically to estimate overall uncertainty in  $SM(F_n, t)$ . The total uncertainty in  $SM(F_n, t)$  is partitioned in two parts, that is due to: i) errors in the input data  $SM(C, t)$ ,  $\sigma_{HH}(F_n)$ , and  $\sigma_{HV}(F_n)$ , and; ii) errors in the algorithm parameters  $\beta$  and  $\Gamma$ . The following analytical formulation quantifies the total uncertainty in  $SM(F_n, t)$ .

#### Uncertainty due to Algorithm Input Data

$$var[SM(F_n)]_{input} = var[SM(C)] + \beta^2 \cdot (10/\ln 10)^2 \{KpC_{HH}^2 + \Gamma^2 \cdot KpC_{HV}^2\} \quad (4.3.12)$$

Eqn. 4.3.12 describes the variance in  $SM(F_n)$  due to errors in the input data.

#### Uncertainty due to Algorithm Parameters

$$var[SM(F_n)]_{param} = \delta_{HH}^2 var[\beta^2] + \delta_{HV}^2 \{ \beta^2 var[\Gamma^2] + \Gamma^2 var[\beta^2] \} \quad (4.3.19)$$

Eqn. 4.3.19 captures the expected variance contribution due to uncertainties in parameters  $\beta$  and  $\Gamma$ .

#### Total Variance in Retrieved Soil Moisture

It is obtained from combining Eqns 4.3.12 and 4.3.19.

$$var[SM(F_n)]_{total} = var[SM(F_n)]_{input} + var[SM(F_n)]_{param} \quad (4.3.20)$$

Various aspects that contribute to the uncertainty of retrieved soil moisture is shown in Fig. 4.3.4.1. The impact due to input errors in the coarse resolution soil moisture based on certain nominal values of other variables and at three different levels of  $\beta(C)$  values are shown in Fig. 4.3.4.1A. It is obvious from the plot that the error in the input dataset  $SM(C)$  has a first order effect on the uncertainty of high-resolution soil moisture, and an increase in the error of  $SM(C)$  leads to an increase in the errors on the estimate of  $SM(F_n)$ .

It is also visible in Figs. 4.3.4.1A that a greater value of  $\beta(C)$  raises the uncertainty in  $SM(F_n)$  estimates. Figure 4.3.4.1B shows the influence of spatial deviation of NISAR co-pol and cross-pol observation of the uncertainty estimate. High spatial deviation ultimately magnifies the errors in algorithm parameters, and therefore increasing spatial deviation in the algorithm increases the uncertainty in the high-resolution soil moisture retrievals.

The uncertainty data cube (Figure 4.3.4.2) shows the expected overall uncertainty in the high-resolution soil moisture retrievals. From the cube, it is visible that more attention is required in better estimation of algorithm sensitivity parameter  $\beta(C)$  and also to have low input error in the coarse resolution soil moisture to keep the overall uncertainty in high-resolution soil moisture under  $\sim 0.06 \text{ m}^3/\text{m}^3$ .

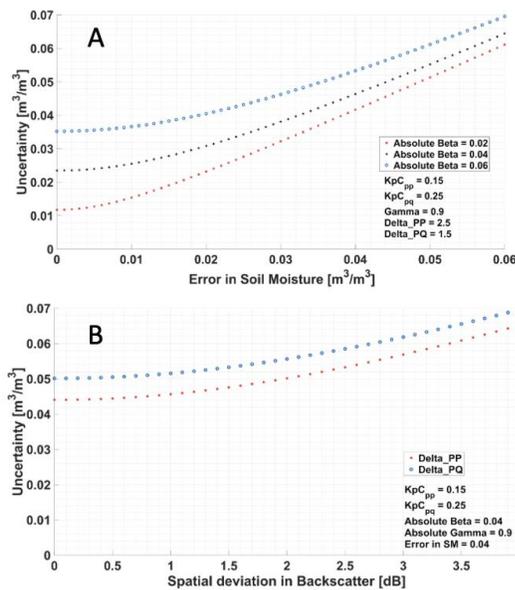


Figure 4.3.4.1: Scenarios of the uncertainty in retrieved high-resolution soil moisture at 200 m from the multiscale fusion algorithm due to errors in input data uncertainties in the algorithm parameters. A) Expected uncertainty in retrieved soil moisture due to errors in the input soil moisture, and; B) Shows the influence of spatial deviation of NISAR co-pol and cross-pol observation on the uncertainty in the retrieved high-resolution soil moisture.

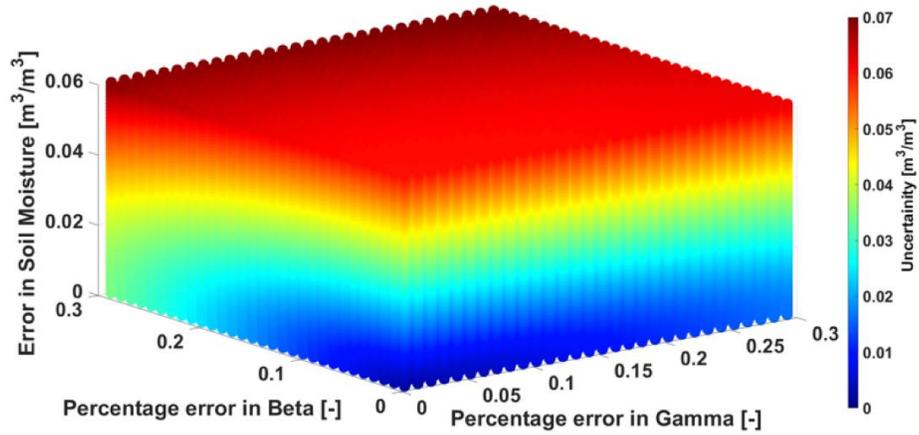


Figure 4.3.4.2: High-resolution soil moisture uncertainty data cube with respect to % error in algorithm parameters and the error in the coarse-resolution soil moisture.

An example of analytically derived uncertainty is shown in Fig. 4.3.4.3 for soil moisture retrievals performed over the Oklahoma region. Such an uncertainty estimate will be provided in the NISAR product for all the valid grid cells at 200 m resolution.

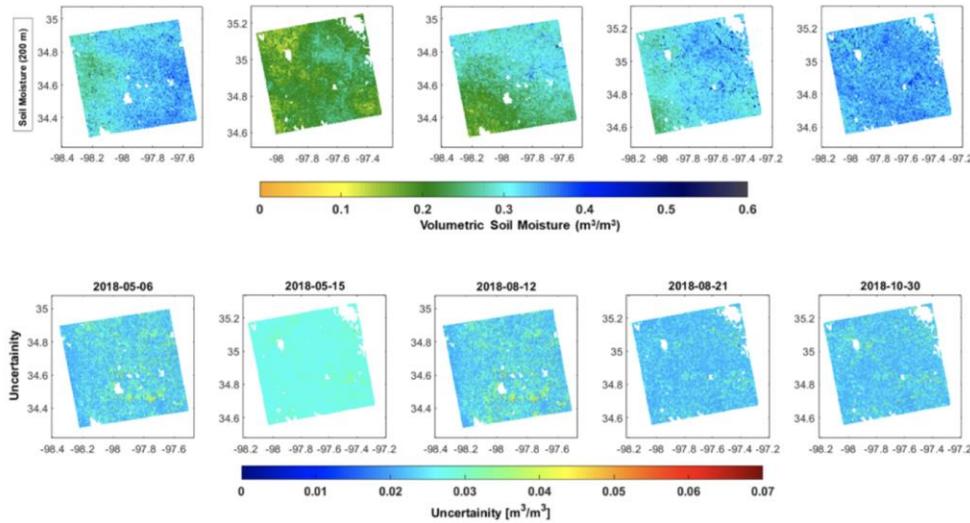
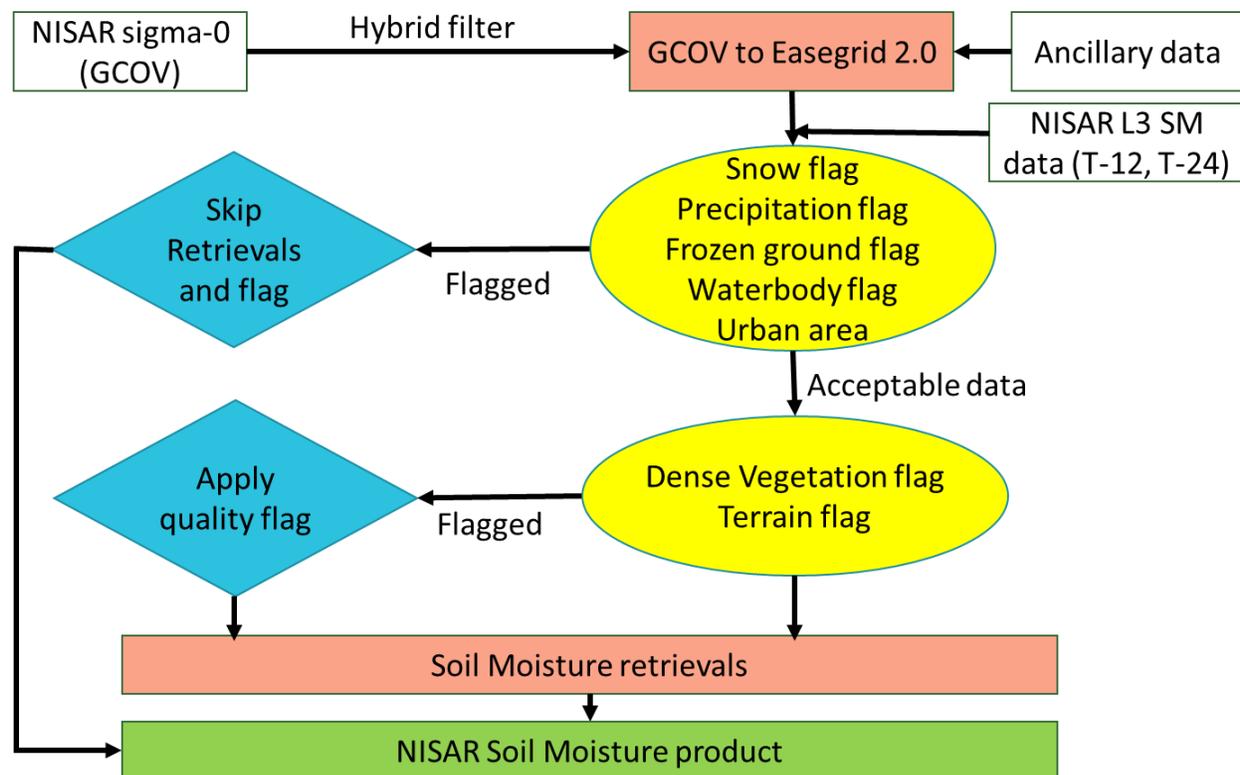


Figure 4.3.4.3: Lower panel: Expected uncertainty in the high-resolution (200 m) soil moisture computed using the analytical approach.

## 5 Algorithm implementation and workflow

### 5.1 Algorithm Flow



### 5.2 Preprocessing

#### 5.2.1 EASEGRID2 description

The grid selected for the NISAR soil moisture products is the updated Equal-Area Scalable Earth-2 (EASE2) grid [Brodzik et al., 2014]. This grid was originally conceived at the NSIDC and has been used to archive several satellite instrument data sets including SMMR, SSM/I, AMSR-E, and SMAP. Using this same grid system for NISAR provides user convenience, facilitates continuity of historical data grid formats, and enables re-use of heritage gridding and extraction software tools developed for EASE2 grid.

#### 5.3 GCOV to EASEGRID2 (Statistical filtering)

The native resolution of GCOV backscatter is ~10 m. The fine-resolution radiometrically terrain corrected (RTC) GCOV data comprises speckle noise and instrument bias. However, the GCOV data granule also includes backscatter observations from water bodies and non-natural scatterers such as the urban area (built up structures) that are undesirable from the soil moisture retrieval algorithms perspective. The NISAR soil moisture algorithms operate at 200 m EASE2 grid resolution. Thus, it is operationally required to average the GCOV data from ~10 m resolution to 200 m EASE2 grid resolution. This makes it essential to eliminate the undesirable backscatter observations going into the averaging process to achieve EASE2 grid

200 m resolution GCOV data. To address the issue of undesirable backscatter outlier observations not going into the averaging process, a hybrid spatial filter is developed similar to the one used for Sentinel-1 data for the SMAP mission [Das et al., 2019].

The spatial hybrid filter is implemented on NISAR GCOV as follows:

1. For each 200 m grid cell within a GCOV granule, the mean ( $m_i$ ) and standard deviation ( $s_i$ ) of  $\sigma_{pp}$  and  $\sigma_{pq}$  are calculated,  $i = 1, \dots, N_c$  where  $N_c$  is the number of 200 m grid cells within the GCOV granule.
2. Then, mean standard deviation ( $MSD$ ) is computed using over all  $s_i, i = 1, \dots, N_c$ .
3. For all the 200 m grid cells with  $s_i > MSD$ , a moving window median filter within a 3 x 3 samples window size is applied.
4. For all the 200 m grid cells with  $s_i \leq MSD$ , all GCOV SAR backscatter ( $\sigma_{pp}$  and  $\sigma_{pq}$ ) values out of range  $[m_i - MSD : m_i + MSD]$  are eliminated.

After the GCOV data are subjected through the hybrid spatial filter, the normal drop-in-bucket technique is used to linearly average the ~10 m data into 200 m EASE2 grid resolution  $\sigma_{pp}$  and  $\sigma_{pq}$  dataset.

## 5.4 Output Variables

The output HDF5 product file (Table 5.4.1) for one NISAR GCOV granule that stores all the information in various data elements at 200 m resolution. The names of all the data elements are self-descriptive.

Table 5.4.1: Description of HDF5 output file.

Fields	Name	Data Type	Units
1	Algorithm/DSG/Algorithm_Param_Beta	32-bit little-endian floating point	m <sup>3</sup> /(dB · m <sup>3</sup> )
2	Algorithm/DSG/Algorithm_Param_Gamma	32-bit little-endian floating point	unitless
3	Algorithm/DSG/Retrieval_Qflag	16-bit little-endian signed integer	unitless
4	Algorithm/DSG/Soil_moisture	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
5	Algorithm/DSG/Soil_moisture_std	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
6	Algorithm/TSR/Alpha1_parameter	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
7	Algorithm/TSR/Alpha1_parameter_uncertainty	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
8	Algorithm/TSR/Alpha2_parameter	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
9	Algorithm/TSR/Alpha2_parameter_uncertainty	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
10	Algorithm/TSR/Retrieval_Qflag	16-bit little-endian signed integer	unitless
11	Algorithm/TSR/Soil_moisture	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
12	Algorithm/TSR/Soil_moisture_uncertainty	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
13	Algorithm/PMI/Dielectric_constant_estimate	32-bit little-endian floating point	unitless
14	Algorithm/PMI/Retrieval_quality_flag	16-bit little-endian integer	unitless
15	Algorithm/PMI/Roughness_estimate	32-bit little-endian floating point	meters
16	Algorithm/PMI/Soil_moisture_estimate	32-bit little-endian floating point	m <sup>3</sup> /m <sup>3</sup>
17	Algorithm/PMI/Soil_moisture_uncertainty	32-bit little-endian floating point	kg/m <sup>2</sup>
18	Algorithm/PMI/Vegetation_water_content_HV	32-bit little-endian floating point	kg/m <sup>2</sup>
19	Algorithm/PMI/Vegetation_water_content_NDVI	32-bit little-endian floating point	kg/m <sup>2</sup>
20	Algorithm/PMI/Vegetation_water_content_estimate	32-bit little-endian floating point	kg/m <sup>2</sup>
21	EASE_column_index	32-bit little-endian signed integer	unitless
22	EASE_row_index	32-bit little-endian signed integer	unitless
23	IncidenceAngle_aggregated	32-bit little-endian floating point	degree

Fields	Name	Data Type	Units
24	IncidenceAngle_aggregated_std	32-bit little-endian floating point	degree
25	NES0_hh	32-bit little-endian floating point	dB
26	NES0_hv	32-bit little-endian floating point	dB
27	NES0_vh	32-bit little-endian floating point	dB
28	NES0_vv	32-bit little-endian floating point	dB
29	Landcover	8-bit little-endian signed integer	unitless
30	Numberoflooks_hh	16-bit little-endian signed integer	unitless
31	Numberoflooks_hv	16-bit little-endian signed integer	unitless
32	Numberoflooks_vh	16-bit little-endian signed integer	unitless
33	Numberoflooks_vv	16-bit little-endian signed integer	unitless
33	Sigma0_hh_aggregated	32-bit little-endian floating point	linear
34	Sigma0_hv_aggregated	32-bit little-endian floating point	linear
35	Sigma0_vh_aggregated	32-bit little-endian floating point	linear
36	Sigma0_vv_aggregated	32-bit little-endian floating point	linear
37	Surface_Qflag	16-bit little-endian signed integer	unitless
38	Waterbody_fraction	32-bit little-endian floating point	unitless
39	latitude	32-bit little-endian floating point	degree
40	longitude	32-bit little-endian floating point	degree
41	identification/boundingPolygon	character string made up of one or more bytes	unitless

## 5.5 Quality Flags

Quality control (QC) is an integral part of creating an operational geophysical product. The QC information flow of the NISAR mission operational soil moisture processing are based on the quality information that are provided with the input GCOV metadata, and different types of masks, flags, and fractional coverages derived from the ancillary data (such as DEM, water fraction, vegetation-water-content, urban fraction, ECMWF-based precipitation, and ECMWF-based soil/ambient temperature). The SDS at JPL processes all data from the input data streams and the ancillary data to store QC information as surface flags in a two-byte integer for each and every valid grid cell (200 m resolution) in the NISAR soil moisture product. A binary bit is allocated for each type of surface status in the two-byte integer. Table 5.5.1 shows the bit position allocation of the surface flags in the two-byte integer. A binary bit is set to 1, if questionable quality is determined for a respective surface condition, otherwise the bit is set to zero that indicates the surface condition is good and is below or equal to a threshold value. The surface flag process provides clear indication in the data that the retrieval has favorable surface conditions or not (such as VWC  $\leq$  5 kg/m<sup>2</sup>, no rain, no frozen ground, no waterbody, and no urban/manmade-structure buildup) to generate the NISAR soil moisture product.

Table 5.5.1: Surface flag bit configuration within the 2-byte integer

Bit Position	Surface flag	Remarks
0	static water body flag	The flag is set if the static water fraction > 10%

1	urban area flag	The flag is set if the urban/man-made structure detected
2	precipitation flag	The flag is set if the rainfall rate (from ECMWF) above a predetermined threshold is reported close to acquisition time (mentioned in Section 5.7)
3	snow or ice flag	The flag is set if the snow is reported in the grid cell.
4	permanent snow or ice flag	The is flag is set if the permanent ice and snow is reported in the grid cell
5	frozen ground flag based on soil surface temperature	The is flag is set if the ground is < 0 deg C, it is based on soil temperature data from ECMWF
6	Terrain flag	The is flag is set if the standard deviation of slope is greater than TBD
7	dense vegetation flag	The is flag is set if the vegetation water content is greater than 5 kg/m <sup>2</sup>

With the current configuration, nine bits are used in the surface flag. The remaining 7 bits are kept idle (set as 0) and that could be used for future inclusion of new surface flags. The surface flag information is also used in the overall soil moisture retrieval quality flag. The soil moisture retrieval quality flag is made of 2-byte integers. Table 5.5.2 shows the bit configuration of the soil moisture retrieval flag data. The retrieval flag is set for each and every valid grid at 200m where soil moisture retrieval is attempted. The retrieval flag 0-bit position is set to 0 or 1 depending upon the surface flag data, for example, if the surface flag overall value for a particular grid cell is 0 then the retrieval flag 0-bit position is set to 0 provided all the other bits in the retrieval flag are set to 0. The remaining 10 bits in the retrieval quality flag are kept idle (set as 0) and that could be used for the future inclusion of new flags.

Table 5.5.2: Retrieval flag bit configuration within the 2-byte integer

Bit Position	Retrieval flag	Remarks

0	Retrieval recommended flag	The flag is set if the surface flag value is greater than 0 or any other bits in the retrieval flag is raised.
1	Retrieval attempted flag	The flag is set if the attempt is not made to retrieve soil moisture
2	Retrieval success flag	The flag is set if the retrieval attempt is not successful
3	Soil moisture thresholds between (0.02 and 0.60)	The flag is set if the soil moisture retrieval is below 0.02 m <sup>3</sup> /m <sup>3</sup> or greater than 0.60 m <sup>3</sup> /m <sup>3</sup> soil moisture
4	NISAR $\sigma_{hh}$ anomaly flag	The is flag is set if <b>TBD</b> anomaly (e.g., RFI) detected in NISAR $\sigma_{hh}$ backscatter
5	NISAR $\sigma_{hv}$ anomaly flag	The is flag is set if <b>TBD</b> anomaly (e.g., RFI) detected in NISAR $\sigma_{hv}$ backscatter
6	NISAR $\sigma_{vv}$ anomaly flag	The is flag is set if <b>TBD</b> anomaly (e.g., RFI) detected in NISAR $\sigma_{vv}$ backscatter
7	NISAR $\sigma_{vh}$ anomaly flag	The is flag is set if <b>TBD</b> anomaly (e.g., RFI) detected in NISAR $\sigma_{vh}$ backscatter

## 6 Ancillary data

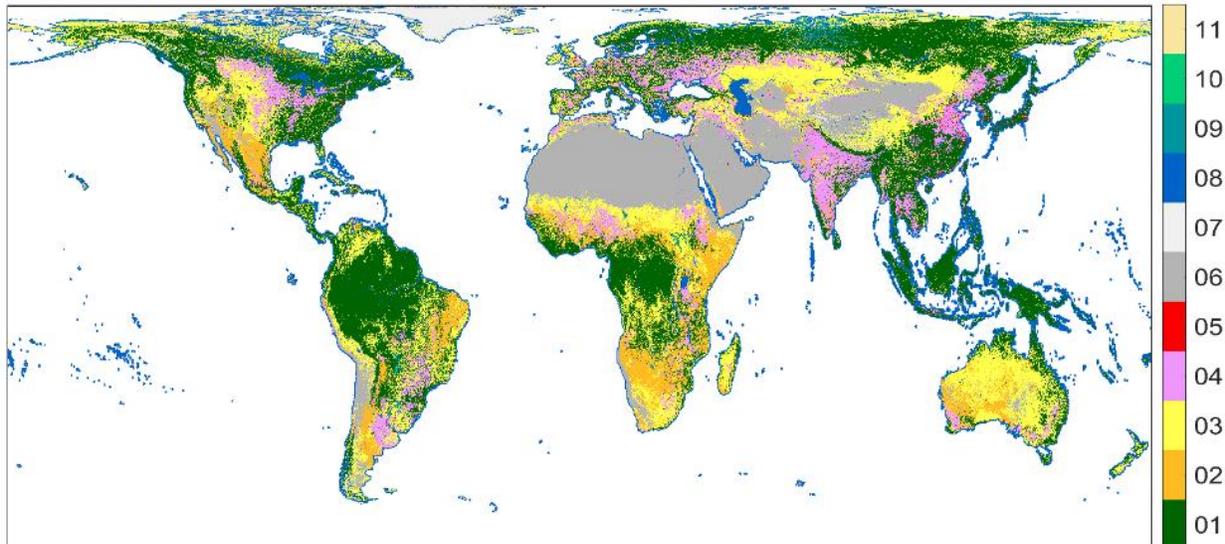
### 6.1 ECMWF Soil Moisture

The ECMWF (European Centre for Medium-Range Weather Forecasts) is a European global forecast seamless model. It is widely regarded as the best and most reliable model currently in existence [Muñoz-Sabater et al., 2021; Roberts et al., 2018]. ECMWF produces operational ensemble-based analyses and predictions that describe the range of possible scenarios and their likelihood of occurrence. ECMWF's forecasts cover time frames ranging from the week (medium-range), to monthly and seasonal (Extended-range forecast), and up to a year ahead (Long-range forecast). The NISAR project is planning to use the near real time ECMWF forecast product at ~10 km resolution. One of the soil moisture algorithms, i.e., the multiscale fusion approach uses the ECMWF coarse resolution ~10 km soil moisture analysis as an input to the algorithm. Lal et al., 2022 conducted a study to evaluate the accuracy and uncertainties of the ECMWF soil moisture at ~10 km resolution. An approach similar to SMAP product validation was implemented to evaluate the ECMWF soil moisture analyses against the SMAP Core Validation Sites (CVS in-situ). They reported that the ECMWF soil moisture has overall low ubRMSE (~0.04 m<sup>3</sup>/m<sup>3</sup>) but has moderately higher bias specifically over cropland in North America and Europe. In addition to evaluation from SMAP CVS in-situ measurement, ECMWF soil moisture was also evaluated against SMAP passive soil moisture 9 km gridded product (L3SMP\_E). ECMWF soil moisture has an overall wet bias, and the least bias observed over tropical and temperate zones with a high R-value (good correlation). The subtropical region has a low bias, but there is a poor correlation between the two datasets, as the desert-like region doesn't have enough soil moisture variability. Overall evaluation with SMAP CVS in-situ measurement and SMAP passive soil moisture indicated that ECMWF soil moisture has a slightly wet bias and is reliable to use in the multiscale fusion algorithm.

### 6.2 Landcover

Landcover classification information is needed as an ancillary input and as a quality flag for the NISAR soil moisture retrievals algorithms. The European Space Agency (ESA) provided WorldCover global land cover products at 10 m resolution version V100 in geographic coordinates. This data is used to produce a global land cover map at a resolution of ~ 200 m EASE 2.0 grid projection. The ESA WorldCover 10 m 2020 product can be downloaded through WorldCover viewer, available at <https://viewer.esa-worldcover.org/worldcover/>. WorldCover is derived using the Sentinel-1 and Sentinel-2 satellite data time-series (<https://sentinels.copernicus.eu/web/sentinel/>). The algorithm used to generate the ESA WorldCover product is based on the approach used to produce the dynamic yearly Copernicus Global Land Service Land Cover (CGLS-LC) map at 100 m resolution (Buchhorn et al., 2020), which is well validated on global scale with an overall mapping accuracy of ~80.3% (Tsendbazar et al., 2021). The overall mapping accuracy of ESA WorldCover 10 m products on a global scale is ~74.4% based on 1.93 million secondary sample units at 21624 primary sample unit locations (Zanaga et al., 2021).

The WorldCover product has a total of 11 discrete land cover classes that are similar to the Land Cover Classification System developed by the United Nations Food and Agriculture Organization. The most recent ESA WorldCover 10 m product available for the year 2020, was processed to develop a global landcover map at EASE2-grid 200 m by selecting the dominant class within each 200 m grid-cell. An overview of the discrete map with 11 land cover classes on a global scale is given in Fig. 6.2.1. A coastal area is classified as permanent water bodies in the ESA WorldCover product. Therefore, all the coastal regions in the Fig 6.2.1 are highlighted (blue color) as permanent water bodies.



**01:** Tree cover; **02:** Shrublands; **03:** Grassland; **04:** Cropland; **05:** Built-up; **06:** Bare/sparse vegetation; **07:** Snow & Ice; **08:** Permanent water bodies; **09:** Herbaceous wetland; **10:** Mangroves; **11:** Moss & lichen

Fig 6.2.1 Global land cover map at EASE2-grid 200 m (*data source: ESA WorldCover 10m 2020*).

## 6.3 Cropcover

Sources of the cropland database are as follows

- Contiguous US. The cropland data layer (CDL) generated by the USDA provides a 30-m resolution of 250+ classes map. CDL is updated annually and published around January for the past year information. The accuracy of the CDL product exceeds 90% for dominant crops such as wheat, corn, soybean, and rice (Johnson and Mueller 2010). For less common crops, the accuracy ranges between 75-80%.
- Canada produces a similar database as CDL (30-m, 250+ class, published in January for the past year) by AAFC (Agriculture Agri-Food Canada).
- For the rest of the world, the static database will be used as a baseline (Monfreda et al. 2008) was derived from a ground survey. The database has a 10-km resolution. The SMAP project statistically downscaled to a 1-km product, which was further interpolated to 200 m for NISAR. Over Europe, there is a legacy database at 250-m (ECOCLIMAP) that is not adopted for NISAR due to its age. More recently, the Dynamic World database was generated by a consortium led by Google (Brown et al. 2022). It creates dynamic (daily) maps of a global crop layer at 10-m Sentinel-2 resolution using machine learning approaches. The expansion of the one crop class to multiple crops can be achieved in the machine learning framework. Exploring the Dynamic World's capacity is on-going and should be regarded as an alternative source.

## 6.4 Vegetation Water Content

Vegetation water content (VWC) ancillary serves two purposes in the soil moisture algorithms: (1) to correct for the effect by vegetation during retrieval by the physical model inversion approach (2) to flag heavily-

vegetated regions where the retrievals are deemed not so reliable such as forest. The daily climatology of VWC has been estimated using optical NDVI data from Landsat or MODIS [Hunt *et al.*, 2011]. This approach was a long heritage in radiometry retrieval of soil moisture and provided successful input to the global soil moisture retrieval using the SMAP SAR data [S B Kim *et al.*, 2017]. Based on the success, the primary source of the VWC for the NISAR retrieval algorithm and the quality flag is chosen to be the daily climatology derived using the MODIS 250-m optical spectral observations (Figure 6.4.1).

The NDVI-based climatology has two deficiencies: synchronization and mismatches in landcover information (to derive NDVI-VWC one landcover database is used; while the application of the NDVI-VWC to NISAR soil moisture relies on a different database). VWC may be derived from the NISAR observations, noting concurrency and collocation. In the literature, the radar vegetation index (RVI, [Y Kim and van Zyl, 2009]) or dual-pol RVI [Bhogapurapu *et al.*, 2022] were promising, we found that RVI is not uniquely associated with in situ VWC when tested using the SMAPVEX12 data for corn, soybean, and canola. Instead HV itself was more representative of VWC. The dependence of the HV-VWC relationship to incidence angle was weak, which is favorable for applying to the NISAR data. If sufficient progress is made, the estimate of VWC from the NISAR polarimetric data may replace the NDVI-derived climatology.

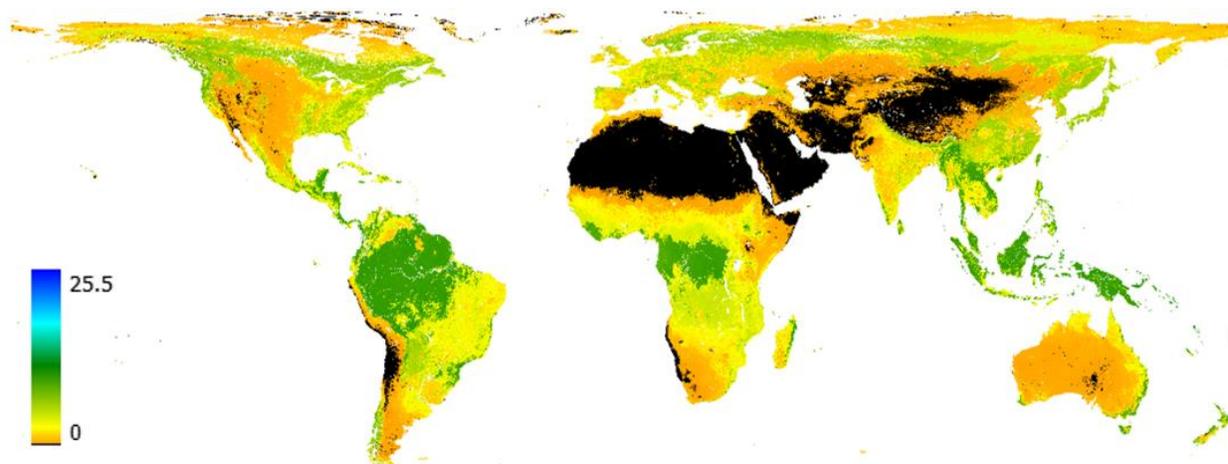


Fig. 6.4.1. VWC daily climatology at 200m resolution derived using MODIS 250m spectral data. Units in kg/m<sup>2</sup>.

## 6.5 Snow

Although the NISAR L-band observations can theoretically see through dry snow with its low dielectric to the soil underneath a snowpack, the snow flag is currently envisioned as an area snow fraction based on the ECMWF ancillary dataset. The snow flag affects soil moisture retrieval processing in the following way:

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

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

## 6.6 Frozen Ground

The retrieval of soil moisture is not meaningful in regions where the ground is frozen. With this rationale, a 0-7 cm soil temperature data field “STL1” from the ECMWF near real-time forecast data will be used to flag frozen ground conditions and soil moisture retrieval will not be attempted. To follow the SMAP heritage, the temperature threshold to determine frozen states is set to 0 °C.

## 6.7 Precipitation

The NISAR precipitation flag is currently set based on forecasts of precipitation from the ECMWF model. The precipitation flag gives the rain rate in mm/hr (or kg/m<sup>2</sup>/s), indicating the presence or absence of precipitation in the 200 m grid cell at the time of the NISAR overpass. The presence of liquid in precipitation incident on the ground at the time of the NISAR overpass can adversely bias the retrieved soil moisture due to its large impact on NISAR. Unlike with other flags, soil moisture retrieval will always be attempted even if precipitation is flagged. However, this flag serves as a warning to the user to view the retrieved soil moisture with some skepticism if precipitation is present.

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

## 6.8 Water Bodies

Water fraction refers to the fraction (by area) of water relative to land in a given spatial extent. This spatial extent of the fraction is a grid cell of a given grid resolution. In all cases, it varies between 0 (no water) and 1 (all water). The static water fraction dataset will be used to discriminate between land and water pixels. The dataset provides information for brightness temperature correction for pixels containing partial land and water, and also to correct for its impact on aggregated radar  $\sigma_0$ . As the NISAR radar-based water detection algorithm matures, information regarding the presence of transient water can be used to augment the static water information and serve the two purposes above more accurately. We will use the static water fraction that is the water occurrence probability map derived using the 30-m resolution Landsat-based dataset by [Pekel *et al.*, 2016].

## 6.9 Urban Areas

In the presence of urban areas, no soil moisture retrieval will be performed and a surface flag will be set for the user. The urban area class at 200 m resolution is present in the WorldCover landuse and landcover map described in Section 6.2. Thus, no other ancillary urban area data is required for the algorithm processing in the Science Data System implementation. In the future, the NISAR project may produce a global urban area mask at a finer resolution.

## 6.10 Soil Texture

The NISAR soil moisture algorithms use the soil's physical attributes such as clay fraction, which is used as an input to the dielectric model to determine the dielectric constant. Soil bulk density data is also used in the NISAR soil moisture SAS to set the possible upper limit soil moisture range. Therefore, better quality soil physical attribute data is essential to perform the above-mentioned steps in the NISAR soil moisture SAS. The best soil data at fine resolution at global extent is available at <https://openlandmap.org> called SoilGrid250m [Hengl et al., 2017]. This database (web repository) has all the high-resolution soil attributes (clay fraction, sand fraction, bulk density, and organic carbon content) required for the NISAR SAS. SoilGrids250m provides data for the following soil properties at global extent (a) Soil organic carbon content in ( $\text{g kg}^{-1}$ ), (b) Soil pH, (c) Sand, silt and clay (weight %), (d) Bulk density ( $\text{kg m}^{-3}$ ), (e) Cation-exchange capacity ( $\text{cmol} + /\text{kg}$ ), (f) Coarse fragments (volumetric %), (g) Depth to bedrock (cm), and (h) USDA Soil Taxonomy classes. NISAR Science team member Narendra Das evaluated the quality and accuracy of the SoilGrid250m data against other available global resources (such as HWSD), and he reported that the SoilGrid250m database has the most accurate estimate of soil properties when compared to the in-situ soil profile data from the World Soil Information Service (WoSIS). More details about the characteristics of this dataset are available in Hengl et al. [2017].

## 6.11 Topography

Historically, soil moisture retrieval algorithms were developed using  $\sigma_0$  rather than  $\gamma_0$ . Therefore the algorithms incorporate local incidence angle (LIA) information. The algorithms plan to flag the pixels with high topography variations, on two considerations. First, topography can alter LIA beyond the range expected by the algorithms. For example, the inversion of the physical model and water-cloud model relies on the models developed at a specific LIA. Though the retrieval tests reported that algorithms were robust to changes in LIA (Fig. 4.1.4.1 and 4.1.4.2) within 30 to 50°, LIA beyond the range is not validated. The scattering mechanism through the vegetation media will also vary according to LIA and may impact the time-series algorithm performance. The algorithm parameters and the fidelity of the multiscale fusion algorithm depend on LIA as well.

The second rationale of the flagging is that the multi-looked  $\sigma_0$  within a 200-m NISAR pixel from the original 20-m single look observations may not represent the spatial mean of underlying soil moisture, because the relationship between  $\sigma_0$  and soil moisture is nonlinear.

To address the above two cases where topography may potentially degrade the retrieval performance, we will flag retrievals based on the mean and standard deviation of LIA within a 200-m pixel computed using the 20-m resolution LIA. The 20-m LIA will be computed dynamically for every single-look pixel based on the beam's look vector and surface normal vector [Shiroma et al. 202]. The 30-m Copernicus DEM will be used for the calculation. The threshold values to the mean and standard deviation are to be determined.

When the surface normal is not on the beam's incidence plane, the cross-pol data will include the depolarized power due to topography, as well as due to vegetation scatterers. The retrieval algorithms assume the cross-pol is due entirely to the vegetation (except for a bare soil). Then the topography may introduce retrieval errors. This 'aspect angle' effect is more difficult to assess and compute. The current strategy is to rely on the LIA flagging described above, noting that steep terrain would have large in-pixel variations due to both LIA and aspect angle variations.

## 7 Calibration and Validation: Pre-launch

During the pre-launch interval, testing and calibration/validation of the retrieval algorithms will be performed using simulated NISAR data based on UAVSAR observations and ALOS PALSAR/PALSAR2 data, where available. Results from the SMAP Validation Experiment in 2012 (SMAPVEX12) experiment (e.g., McNair et al., 2012) suggest that NISAR can meet the goal of ubRMSE of no more than 0.06 m<sup>3</sup>/m<sup>3</sup> for crop types where there were enough *in situ* observations available for assessment.

### 7.1 SMAPVEX12 experiment

SMAPVEX12 covered an agricultural region in Manitoba, Canada (Figure 7.1.1), and included L-band UAVSAR overflights on two overlapping tracks (13604, with 11 observation days and 13606, with 14 observation days, respectively) as well as the collection of soil moisture, surface temperature and other site characteristics at 55 agricultural and 4 forested sites.

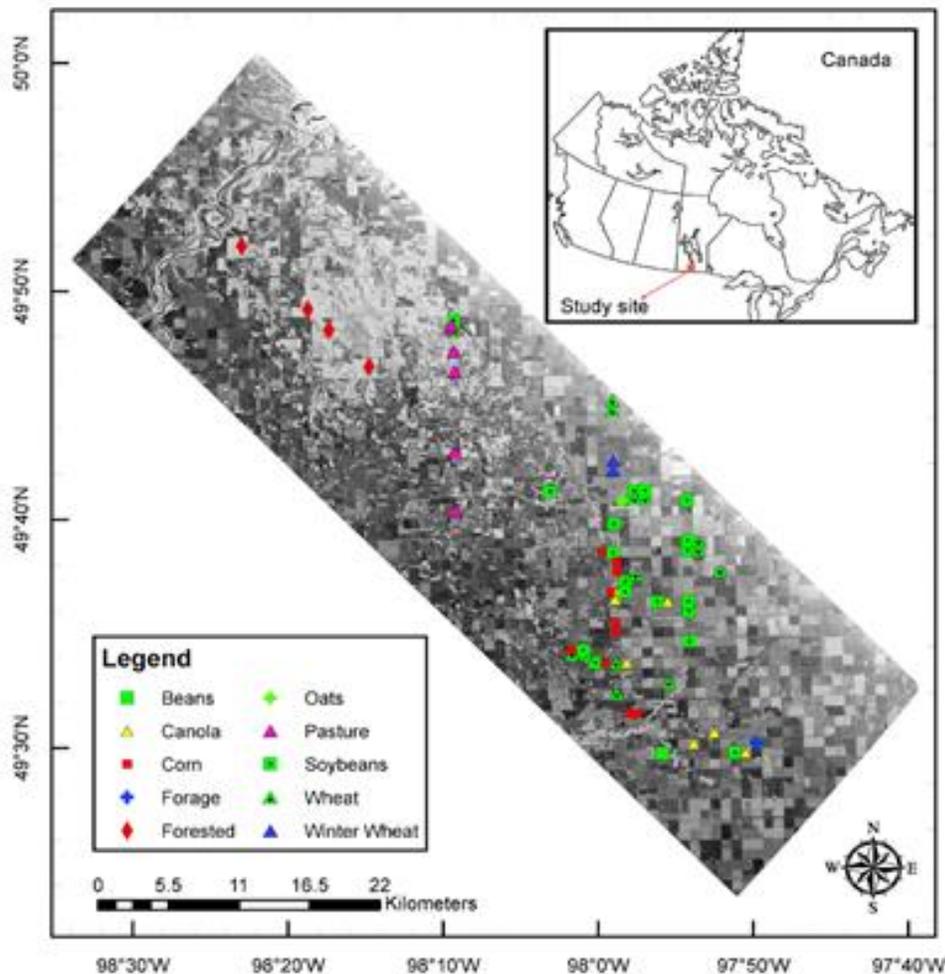


Figure 7.1.1: SMAPVEX12 study area and location of *in situ* field observations (colored points) overlain on SAR backscatter (grayscale).

Each retrieval algorithm was applied to a version of the UAVSAR data that had been simulated to closely match the expected characteristics of NISAR (e.g., spatial resolution, noise), using only the incidence angle range between 30 and 50 degrees.

### 7.1.1 Estimation of sample statistics from SMAPVEX12 exercise

During the SMAPVEX12 exercise, fields had between 6 and 13 days where both *in situ* and UAVSAR observations were made. For each field and for each retrieval algorithm, we find the average bias between the *in situ* observations and retrievals. We also characterize the residual between the *in situ* values and retrievals, and our error on that characterization.

For each retrieval model,  $i$ , at each field site,  $j$ , at each date,  $k$ , we have *in situ* observations,  $d_{jk}^o$ , soil moisture retrievals,  $d_{ijk}^r$ , and residuals,  $d_{jk}^o - d_{ijk}^r$ .

### 7.1.2 Errors on each field based on sample statistics

Except for the forest sites, which each have only one observation on each date, most of the *in situ* observations included observations at 16 locations within each field. At each of those 16 locations, 3 observations are generally made, with a sampling strategy that requires the second observation to have twice the weight of the other two. We use this approach, but find that it has negligible effect on the outcome compared with taking a simple average of all 48 observations at each site.

For each model, at each field, we are interested in how well the modes fit the data on average (i.e., is the bias near zero), but are also very interested in how well the models capture the temporal variability. We calculate the bias and other metrics associated with the fit to the *in situ* observations for  $n$  observations as:

Bias:

$$b_{ij} = \frac{1}{n} \sum_{k=1}^n r_{ijk}$$

Unbiased RMSE:

$$ubRMSE_{ij} = \sqrt{\frac{1}{n} \sum_{k=1}^n (r_{ijk} - b_{ij})^2}$$

Sample standard deviation of residual:

$$\sigma_{ij} = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (r_{ijk} - b_{ij})^2}$$

For our validation efforts, it is useful to not only infer the error on the retrievals, but to evaluate our confidence on that error estimate. This allows us to weight the contribution from fields with differing numbers of observations, and to assess whether differences between the models are significant. For normally distributed noise with an actual standard deviation  $\sigma_{ij}$  and  $n$  samples, the standard deviation of our estimate of  $\sigma_{ij}$ ,  $\hat{\sigma}_{ij}$

$$SD(\widehat{\sigma}_{ij}) = \sigma_{ij} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sqrt{\frac{n-1}{2} - \left(\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right)^2}$$

where  $\Gamma(\cdot)$  is the Gamma function. For more complicated probability distributions that account for the physical bounds on soil moisture, numerical approaches may be required to estimate these errors. We use this inferred error on the sample standard deviation to weight the average standard deviation within a given crop type that is sampled at multiple field sites.

### 7.1.3 Errors on each crop type

The nine land cover types sampled during SMAPVEX12 include between one and thirteen fields each, with varying numbers of observations for each field. Each field is also potentially covered by one or both of tracks 13604 and 13606.

For a given track:

Crop-specific average ubRMSE, for each crop,  $l$ , which is sampled by  $m$  fields, and retrieval model,  $i$ :

$$C_{i,ubRMSE}^l = \frac{1}{m} \sum_{j=1}^m ubRMSE_{ij}$$

### 7.1.4 Results

Here we show ubRMSE and RMSE for several crop types and each of the two tracks. Error bars on RMSE indicate the standard deviation of the fields included for each crop type - so, for crops with few fields, the quality of this estimate is lower. The total number of fields examined is indicated in the title of each figure, but not all fields are covered by each track, and some did not have *in situ* observations acquired on the day of a UAVSAR overflight. We do not show our metrics associated with  $\sigma_{ij}$  since those are not the standard metrics used in the soil moisture literature, but the jupyter notebook that produces these results allows the user to select whichever of these metrics they would like to examine.

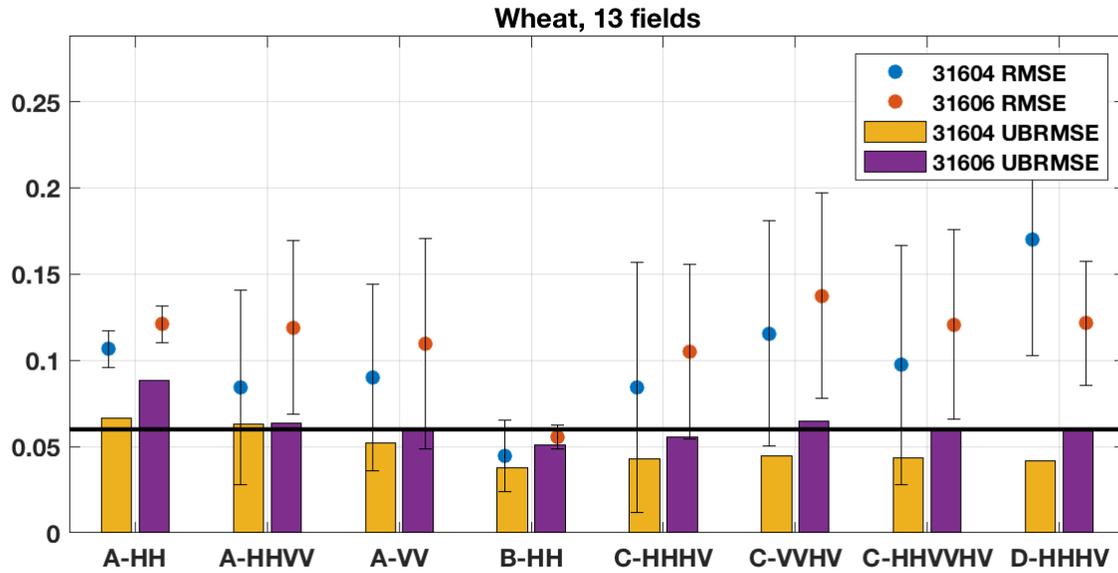


Figure 7.1.4.1: ubRMSE (bars) and RMSE (dots) evaluated over the 13 wheat fields in the SMAPVEX12 study, for each retrieval type and various polarization options. Black heavy line indicates the 0.06 m<sup>3</sup>/m<sup>3</sup> target for NISAR. A: Physical model algorithm, B: Time series algorithm, C: Multiscale Fusion algorithm, D: Semi-empirical algorithm..

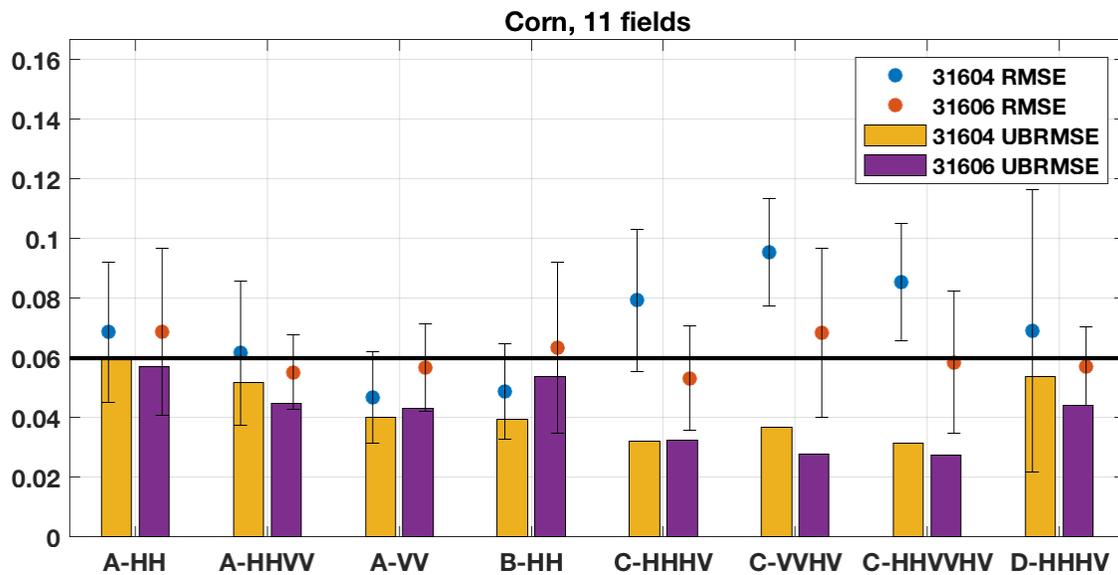


Figure 7.1.4.2: ubRMSE (bars) and RMSE (dots) evaluated over the 11 corn fields in the SMAPVEX12 study, for each retrieval type and various polarization options. Black heavy line indicates the 0.06 m<sup>3</sup>/m<sup>3</sup> target for NISAR. A: Physical model algorithm, B: Time series algorithm, C: Multiscale Fusion algorithm, D: Semi-empirical algorithm.

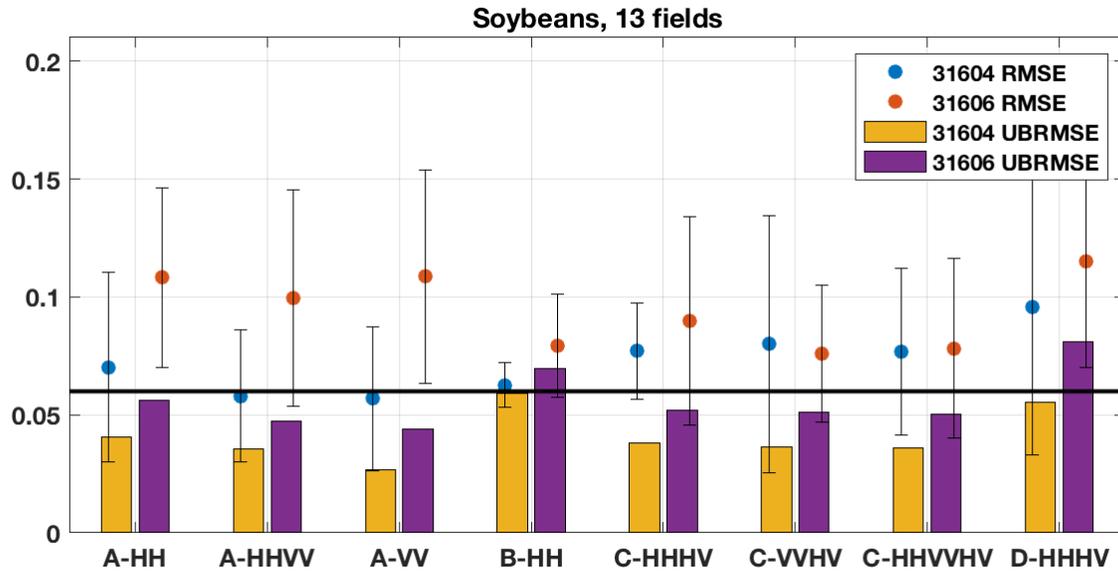


Figure 7.1.4.3: ubRMSE (bars) and RMSE (dots) evaluated over the 13 soybean fields in the SMAPVEX12 study, for each retrieval type and various polarization options. Black heavy line indicates the 0.06 m<sup>3</sup>/m<sup>3</sup> target for NISAR. A: Physical model algorithm, B: Time series algorithm, C: Multiscale Fusion algorithm, D: Semi-empirical algorithm.

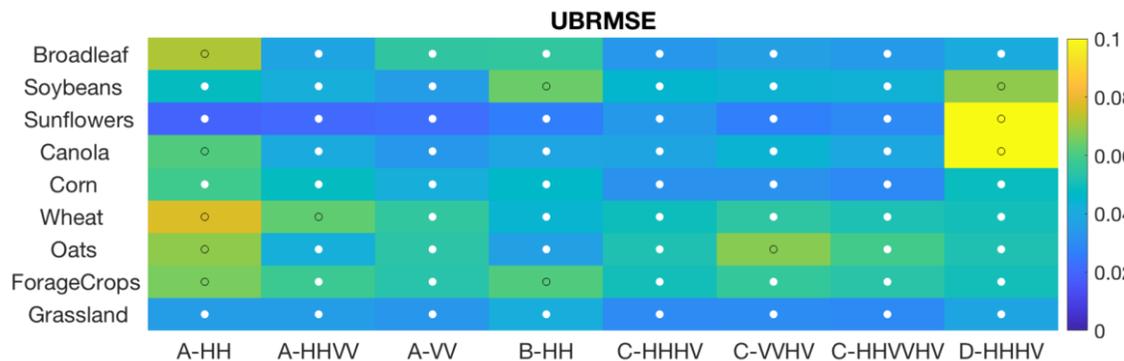


Figure 7.1.4.4: ubRMSE for each group and crop, averaged across both tracks where present. White dots indicate values that are below the 0.06 m<sup>3</sup>/m<sup>3</sup> target, black circles are above the target. A: Physical model algorithm, B: Time series algorithm, C: Multiscale Fusion algorithm, D: Semi-empirical algorithm.

Figure 7.1.4.4 shows that, at least for the data available through the SMAPVEX12 study, the different algorithms are all able to meet the target 0.06 m<sup>3</sup>/m<sup>3</sup> error for most cases. Note that some of the crops, such as oats, forage and broadleaf, had fewer than 3 fields and did not have observations during each UAVSAR overflight.

## 8 Calibration and Validation: Post-launch

Post-launch validation must provide the information necessary to address whether or not NISAR has achieved its mission requirement to produce an estimate of soil moisture in the 0-5 cm layer with an average ubRMSE of no more than 0.06 m<sup>3</sup>/m<sup>3</sup> over areas where the vegetation water content  $\leq 5$  kg/m<sup>2</sup>, excluding regions of frozen soil, permanent snow / ice, mountainous terrain, and open water at the footprint measurement scale (200 m). Different methodologies and data types will be utilized for validating the NISAR soil moisture product. These types of data are listed in Table 8.1, which describes their perceived role and issues that are currently being addressed by the NISAR team.

Table 8.1. Overview of the NISAR Cal/Val Methodologies

Methodology	Role	Issues	Actions
<i>In Situ</i> Validation Sites	Allow assessment of products at matching scales for a limited set of conditions	Calibration Up-scaling Limited number	<i>In Situ</i> Testbed Scaling methods
Satellite Products	Estimates over a very wide range of conditions at matching scales	Validation Comparability Continuity	Validation Studies CDF Matching
Model Products	Estimates over a very wide range of conditions at matching scales	Validation Comparability	Validation Studies
Field Experiments	Detailed estimates for a very limited set of conditions (physical sampling, additional samples observed on overpass dates)	Resources Schedule Conflicts	Simulators Partnerships

The baseline validation for the NISAR soil moisture will be a comparison of retrievals at 200 m with ground-based observations (described below in section 8.1). The period of the initial validation will be 3 to 6 months after the end of the Initial Orbit Checkout, with the number of sites and temporal sampling subject to availability of funding. In order to conduct validation studies that cover a wider range of conditions, as well as a synoptic/global assessment, some combination of data from sparse networks, other satellite products, and model-based estimates must be utilized.

## 8.1 In situ Validation

NISAR validation will rely on all available resources for validating the high-resolution, 200-m soil moisture product. Unlike SMAP, the need for multiple sites within a 200 m pixel is greatly reduced due to the high spatial resolution. Some of the key characteristics needed for *in situ* sites are:

- Accessibility to researchers
- Long term commitment by the host
- Low temporal latency
- Compatible with the validation requirements in terms of depths, etc.

Sparse networks such as USDA Soil Climate Analysis Network (SCAN) and US Climate Reference Network (USCRN) sample a wide range of land cover types and climatic conditions, but generally provide only one site (or possibly a few sites) within a satellite footprint. Data from International Soil Moisture Network (ISMN) will also be used for validation. In situ soil moisture data is also collected at various cal/val networks which can be used for NISAR validation depending on availability (for example: USDA ARS; TxSON; Kenaston, Canada; Carmen, Canada; Monte Buoy, Argentina; Hoal, Austria; Yanco, Australia). ISRO is also developing validation sites that cover a wide range of conditions in India. Data from all potential networks will be used for validation purposes.

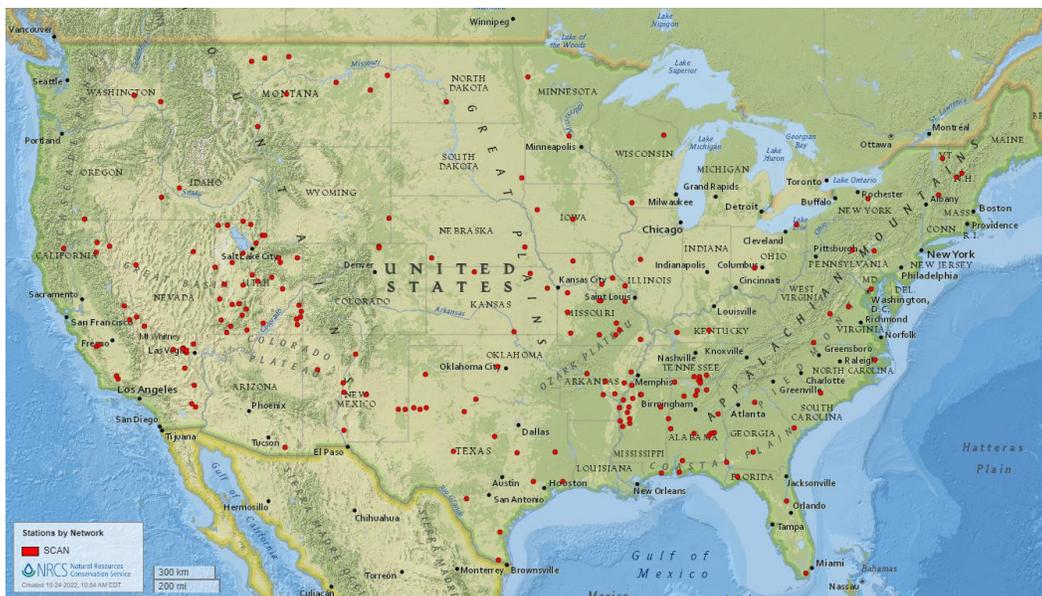


Figure 8.1.1 Location of USDA Natural Resources Conservation Service (NRCS) Soil Climate Analysis Network (SCAN) sites in the United States.

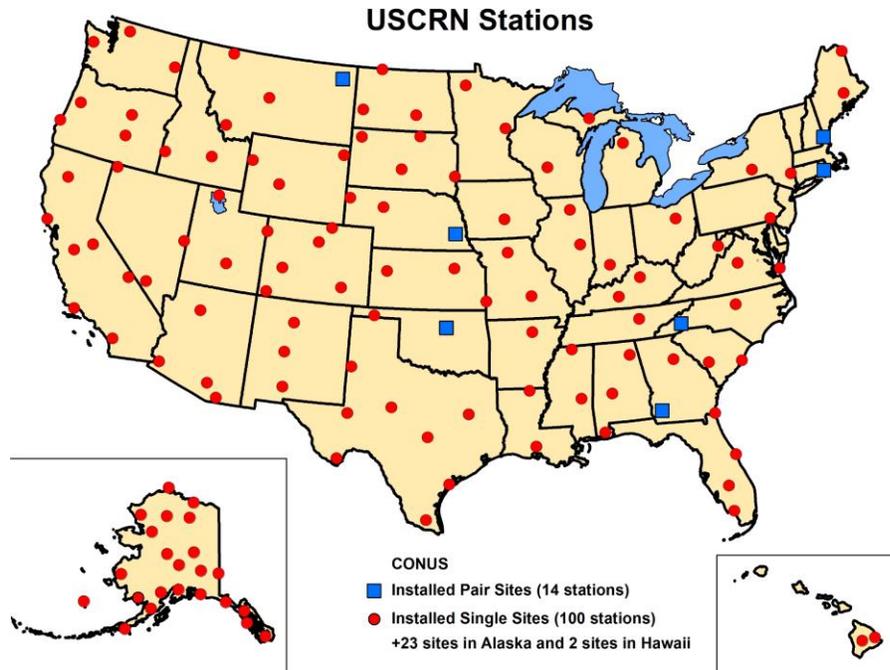


Figure 8.1.2 Location of US Climate Reference Network (CRN).

### 8.1.1 Configuration of *in situ* measurements at each super site

A super site should have a minimum of 3 semi-permanent, instrumented stations distributed across the landscape on the order of the scale of the pixel (200m). If the budget allows, at least five stations within the 200x200m pixel is ideal. Should field experimentation be available for days of overflight, it would be ideal to also collect physical samples of volumetric soil water content for the field scale with a replication of at least 14 locations across the field.



Fig. 8.1.1.1. Example layout of instrumented stations (yellow star) and temporary intensive manual locations (blue circle) for *in situ* soil moisture reading at a super site for post-launch validation.

## 8.1.2 Location of Potential Super Sites

By ~6 months prior to launch, the final set of the validation sites will be determined. Each of the final sites should include at least the minimum number of instrumented stations (as illustrated in Fig. 7.1.1.1) with intensive manual sampling of soil moisture on overflight days if possible.



Fig 8.1.2.1. Locations of candidate validation sites in North America (yellow pins).

## 8.2 Satellite Products

NISAR soil moisture product will be compared with other available satellite products (for example: SMAP, SMOS, Sentinel) at a global scale. Satellite products provide a wide range of soil moisture conditions across all the biomes. Cross-calibration exercises between different satellite instruments have been successfully carried out improving the quality of the time series. The limitations of this type of comparison are the quality of the alternative product, differences in overpass days, and accounting for systematic differences affecting the soil moisture product. Another role for the satellite products is in providing a synoptic perspective. Global comparisons will be used to identify regions and/or time periods where the soil moisture products from the different satellites diverge.

## 8.3 Model Products

Validation based on land surface modeling and data assimilation will be used to complement *in situ* based validation. As discussed in previous sections, validation against *in situ* observations is difficult because the observation sites span limited geographic regions and environmental settings. Hydrological land surface

models and data assimilation approaches can provide continuous (in space and time) soil moisture products that match the spatial support of NISAR soil moisture products.

Several Numerical Weather Prediction (NWP) centers (including ECMWF, NCEP, and NASA/GMAO) routinely produce operational or quasi-operational soil moisture fields. NISAR soil moisture estimates will be aggregated to the coarser resolution provided by these models. These data products rely on the assimilation of a vast number of atmospheric observations (and select land surface observations) into General Circulation Models (GCM's). Although there are many caveats that need to be considered in using these data, they are readily available and they are consistent with the atmospheric forcing (precipitation and radiation) and land use information that determine the spatial and temporal patterns in soil moisture fields. Moreover, surface temperature from at least one NWP system (ECMWF) will be used in the generation of the NISAR data product. High resolution soil moisture data at 1 km resolution over continental United States is available from the Land Information System (LIS). These model estimates provide an independent source of cross-comparison across a wide range of conditions.

## 8.4 Field Experiments

Field experiments serve a valuable role during pre-launch by providing diverse but controlled condition data that can be used for developing algorithms, establishing algorithm parameterization, and defining validation site scaling properties. Field experiments require a significant amount of effort and resources. Field experiments can be done provided these resources are available post-launch. Effort will be made to coordinate the field experiment with other soil moisture missions (eg. SMAP).

## 9 References

Albergel, C., De Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., ... & Wagner, W. (2012). Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sensing of Environment*, 118, 215-226.

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.

Bhogapurapu, N., S. Dey, D. Mandal, A. Bhattacharya, L. Karthikeyan, H. McNairn, and Y. S. Rao (2022), Soil moisture retrieval over croplands using dual-pol L-band GRD SAR data, *Remote Sens. Environ.*, 271, doi:10.1016/j.rse.2022.112900.

Brodzik, MJ, B Billingsley, T Haran, B Raup and MH Savoie (2014) EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets (vol 1, pg 32, 2012) ISPRS Int. Geo-Inf. 3 (3) 1154-1156, , issn: CO1RVdoi: [10.3390/ijgi3031154](https://doi.org/10.3390/ijgi3031154)

Brown, C. F., et al. (2022), Dynamic World, Near real-time global 10 m land use land cover mapping, *Scientific Data*, 9(1), doi:10.1038/s41597-022-01307-4.

Buchhorn, M.; Lesiv, M.; Tsendbazar, N.-E.; Herold, M.; Bertels, L.; Smets, B. Copernicus Global Land Cover Layers—Collection 2. *Remote Sens.* 2020, 12, 1044. <https://doi.org/10.3390/rs12061044>

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 Sensing*, 49, 4832-4845.

Burgin, M., U. K. Khankhoje, X. Duan, and M. Moghaddam (2016), Generalized Terrain Topography in Radar Scattering Models, *IEEE Trans. Geosci. Remote Sensing*, 54(7), 3944-3952.

Burgin, Mariko S., and Jakob J. van Zyl (2017), Regression-based soil moisture estimation from SMAP polarimetric radar data with Aquarius derived coefficients. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE.

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

Chan, S., Bindlish, R., Hunt, R., Jackson, T., & Kimball, J. (2013). Soil moisture active passive (SMAP) ancillary data report: vegetation water content. Jet Propulsion Laboratory: Pasadena, CA, USA.

Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., ... & Kerr, Y. (2018). Development and assessment of the SMAP enhanced passive soil moisture product. *Remote Sensing of Environment*, 204, 931-941.

Chapman, B., McDonald, K., Shimada, M., Rosenqvist, A., Schroeder, R., & Hess, L. (2015). Mapping regional inundation with spaceborne L-Band SAR. *Remote Sensing*, 7(5), 5440-5470.

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(1), 90-101.

Das, N. N., Entekhabi, D., & Njoku, E. G. (2010). An algorithm for merging SMAP radiometer and radar data for high-resolution soil-moisture retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 49(5), 1504-1512.

Das N. N., D. Entekhabi, E. G. Njoku, J. Johnston, J. C. Shi, and A. Colliander, (2014), Tests of the SMAP Combined Radar and Radiometer Brightness Temperature Disaggregation Algorithm Using Airborne Field Campaign Observations, *IEEE Trans. Geosci. Remote Sens.*, vol. 52, pp. 2018–2028.

Das, N. N., Entekhabi, D., Kim, S., Jagdhuber, T., Dunbar, S., Yueh, S., & Colliander, A. (2017), High-resolution enhanced product based on SMAP active-passive approach using sentinel 1A and 1B SAR data. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 2543-2545). IEEE.

Das N. N., D. Entekhabi, R. S. Dunbar, A. Colliander, F. Chenc, W. Crow, T. J. Jackson, A. Berg, D. D. Bosch, T. Caldwell, M. H. Cosh, C. H. Collins, E. Lopez-Baeza, M. Moghaddam, T. Rowlandson, P. J. Starks, M. Thibeault, J. P. Walker, X. Wu, P. E. O'Neill, S. Yueh, and E. G. Njoku, (2018), The SMAP mission combined Active-Passive soil moisture product at 9 km and 3 km spatial resolutions,"*Remote Sens. Environ.*, vol. 211, pp. 204-217.

Das N. N., D. Entekhabi, R. S. Dunbara, M. J. Chaubell, A. Colliander, S. Yueh, T. Jagdhuber, F. Chen, W. Crow, P. E. O'Neill, J. P. Walker, Aaron Berg, D. D. Bosch, T. Caldwell, M. H. Cosh, C. H. Collins, E. Lopez-Baeza, and M. Thibeault., (2019) The SMAP and Copernicus Sentinel-1A/B microwave Active-Passive high-resolution surface soil moisture product, *Remote Sens. Environ.*, vol. 233.

DiMartino, Gerardo, et al. (2016), Polarimetric two-scale two-component model for the retrieval of soil moisture under moderate vegetation via L-band SAR data, *IEEE Transactions on Geoscience and Remote Sensing* 54.4: 2470-2491.

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(1), 87-103.

Elvidge, C.D., Imhoff, M.L., Baugh, K.E., Hobson, V.R., Nelson, I., Safran, J., Dietz, J.B., Tuttle, B.T. (2001), Night-time lights of the world: 1994–1995. *ISPRS J. Photogramm. Remote Sens.* 56, 81–99. [https://doi.org/10.1016/S0924-2716\(01\)00040-5](https://doi.org/10.1016/S0924-2716(01)00040-5).

Entekhabi, Dara, et al., (2010), The soil moisture active passive (SMAP) mission., *Proceedings of the IEEE* 98.5 (2010): 704-716.

Entekhabi, D. and N. N. Das, (2014), Algorithm Theoretical Basis Document (ATBD) JPL D-66481, SMAP L2 and L3 Radar/Radiometer Soil Moisture (Active/Passive) Data Product: L2\_SM\_AP and L3\_SM\_AP. NASA, Jet Propulsion Laboratory, California Institute of Technology. (Peer Reviewed).

Fung, A. K., Z. Li, and K. S. Chen (1992), Backscattering from a randomly rough dielectric surface, *IEEE Trans. Geosci. Remote Sens.*, 30(2), 356-369.

Ganesan, P. G., S. B. Kim, T. H. Liao, M. Reba, and M. Cosh (2022), Field-scale soil moisture estimation under corn and soybean crops using airborne SAR data, Proc. IGARSS 2022, Kuala Lumpur, Malaysia.

He, L., Q. Qin, R. Panciera, M. Tanase, J. P. Walker, and Y. Hong, "An extension of the alpha approximation method for soil moisture estimation using time-series SAR data over bare soil surfaces," IEEE Geosc. Rem. Sens. Lett., vol. 14, pp. 1328-1332, 2017.

Hengl, T., Jesus, J.M. de, Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B. (2017), SoilGrids250m: Global gridded soil information based on machine learning. PLOS ONE 12, e0169748.  
<https://doi.org/10.1371/journal.pone.0169748>.

Hensley, S., Michel, T., Van Zyl, J., Muellerschoen, R., Chapman, B., Oveisgharan, S., ... & Mladenova, I. (2011). Effect of soil moisture on polarimetric-interferometric repeat pass observations by UAVSAR during 2010 Canadian soil moisture campaign. In 2011 IEEE International Geoscience and Remote Sensing Symposium (pp. 1063-1066). IEEE.

Huang, H. T., S. B. Kim, S. Tsang, X. L. Xu, T. H. Liao, T. J. Jackson, and S. H. Yueh (2016), Coherent model of L-band radar scattering by soybean plants: model development, validation and retrieval, *IEEE J. Sel. Topics Appl. Earth Observ.*, 9(1), 272 - 284.

Huang, H. T., T. H. Liao, S. B. Kim, X. L. Xu, S. Tsang, T. J. Jackson, and S. H. Yueh (2021), L-Band Radar Scattering and Soil Moisture Retrieval of Wheat, Canola and Pasture Fields for SMAP Active Algorithm, *Progress In Electromagnetics Research*, su170, 129-152.

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 Sensing*, 50, 4025-4035.

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(6), 2557-2567.

Hunt, E. R., L. Li, M. T. Yilmaz, and T. J. Jackson (2011), Comparison of vegetation water contents derived from shortwave-infrared and passive-microwave sensors over central Iowa, *Remote Sens. Environ.*, 115, 2376-2383.

Jackson, T. J., H. McNairn, M. A. Wetz, 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(4), 1065-1069.

Jagdhuber, T., I. Hajnsek, and K. Papathanassiou, "An iterative generalized hybrid decomposition for soil moisture retrieval under vegetation cover using fully polarimetric SAR," IEEE J. Sel. Top. Appl. Rem. Sens., vol. 8, pp. 3911-22, 2015.

Johnson, D., and R. Mueller, 2010: The Cropland Data Layer. *Photogrammetric Engineering and Remote Sensing*, 76, 1201-1205.

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.

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., ... & Mecklenburg, S. (2010). The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, 98(5), 666-687.

Kim, S. B., M. Aree, and T. J. Jackson (2017a), Modeling L-band synthetic aperture radar observations through dielectric changes in soil moisture and vegetation over shrublands, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, 10.1109/JSTARS.2017.2741497.

Kim, S. B., H. T. Huang, T. H. Liao, and A. Colliander (2018), Estimating vegetation water content and soil surface roughness using physical models of L-band radar scattering for soil moisture retrieval, *Remote Sens.*, 10(4), 556, doi:10.3390/rs10040556.

Kim, S. B., and T. H. Liao (2021), Robust retrieval of soil moisture and roughness at field scale across wide-ranging SAR incidence angles for soybean, wheat, forage and grass, *Remote Sens. Environ.*, <https://doi.org/10.1016/j.rse.2021.112712>, doi:<https://doi.org/10.1016/j.rse.2021.112712>.

Kim, S. B., M. Moghaddam, L. Tsang, M. Burgin, X. Xu, and E. G. Njoku (2014), Models of L-band radar backscattering coefficients over the global terrain for soil moisture retrieval, *IEEE Trans. Geosci. Remote Sens.*, 52(2), 1381-1396.

Kim, S. B., L. Tsang, J. T. Johnson, S. Huang, J. J. van Zyl, and E. G. Njoku (2012), Soil moisture retrieval using time-series radar observations over bare surfaces, *IEEE Trans. Geosci. Remote Sens.*, 50(5), 1853-1863.

Kim, S. B., et al. (2017), Surface soil moisture retrieval using the L-band synthetic aperture radar onboard the Soil Moisture Active Passive (SMAP) satellite and evaluation at core validation sites, *IEEE Trans. Geosci. Remote Sens.*, 55(4), 1897 - 1914.

Kim, Y., T. J. Jackson, R. Bindlish, H. Y. Lee, and S. Y. Hong (2012), Radar Vegetation Index for Estimating the Vegetation Water Content of Rice and Soybean, *IEEE Geoscience and Remote Sensing Letters*, 9(4), 564-568.

Kim, Y., and J. J. van Zyl (2009), A time-series approach to estimate soil moisture using polarimetric radar data, *IEEE Trans. Geosci. Remote Sensing*, 47(8), 2519-2527.

Koyama, Christian N., et al., (2017), The effect of precipitation and soil moisture variations on (partial) polarimetric L-band SAR backscatter in tropical forest regions, 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE.

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.

Kurum, M., S. B. Kim, R. Akbar, and M. Cosh (2020), Surface Soil Moisture Retrievals Under Forest Canopy For L- band SAR Observations Across A Wide Range of Incidence Angles By Inverting A Physical Scattering Model, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, 10.1109/JSTARS.2020.304788, doi:10.1109/JSTARS.2020.304788.

Lal, P., G. Singh, N. N. Das, A. Colliander, and D. Entekhabi (2022), Assessment of ERA5-Land Volumetric Soil Water Layer Product Using In Situ and SMAP Soil Moisture Observations, *IEEE Geoscience and Remote Sensing Letters*, Submitted.

Lang, R. H., and J. S. Sidhu (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, paper presented at 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.

Liao, T. H., and S. B. Kim (2022), Dual-frequency Retrieval of Soil Moisture from L- and S-band Radar Data for Corn and Soybean, *Remote Sens.*, submitted.

Liao, T. H., S. B. Kim, A. Handwerger, E. J. Fielding, M. C. Cosh, and W. S. Schultz (2021), High Resolution Soil Moisture Maps Over Landslide Regions in Northern California Grassland, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, 10.1109/JSTARS.2021.3069010.

Liao, T. H., S. B. Kim, S. Tan, S. Tsang, C. X. Su, and T. J. Jackson (2016), Multiple scattering effects with cyclical terms in active remote sensing of vegetated surface using vector radiative transfer theory, *IEEE J. Sel. Topics Appl. Earth Observ.*, 9(4), 1414 - 1429.

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., 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(4), 954-966.

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, *Hydrology and Earth System Sciences*, vol. 13, pp. 343-356.

Mattia, Francesco, et al, (2018), Sentinel-1 & Sentinel-2 for soil moisture retrieval at field scale. IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE.

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.

McNairn, H., Jackson, T. J., Wiseman, G., Belair, S., Berg, A., Bullock, P., ... & Hosseini, M. (2014). The soil moisture active passive validation experiment 2012 (SMAPVEX12): Prelaunch calibration and validation of the SMAP soil moisture algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5), 2784-2801.

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

- 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.
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., (2021), ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* 13, 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>.
- 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.
- Oh, Y. (2004). Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 42(3), 596-601.
- Ouellette, J., S. Hensley, and J. T. Johnson, "Backscattered field correlations for rough surfaces with varying dielectric properties," IGARSS 2012, 2012.
- Ouellette, J. D., Johnson, J. T., Kim, S., van Zyl, J. J., Moghaddam, M., Spencer, M. W., ... & Entekhabi, D. (2014), A simulation study of compact polarimetry for radar retrieval of soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, 52(9), 5966-5973.
- Ouellette, Jeffrey D., et al., (2017), A time-series approach to estimating soil moisture from vegetated surfaces using L-band radar backscatter." *IEEE transactions on geoscience and remote sensing* 55.6, 3186-3193.
- Park, Jeonghwan, Rajat Bindlish, Alexandra Bringer, Dustin Horton, Joel T. Johnson, 2021. Soil moisture retrievals using a time series ratio algorithm for the NISAR mission. *IEEE Geoscience and Remote Sensing Symposium 2021*. pp. 5873-5876, doi: 10.1109/IGARSS47720.2021.9554619.
- Park, Jeonghwan, Rajat Bindlish, Alexandra Bringer, Dustin Horton, Joel T. Johnson, 2022. Time series ratio algorithm for NISAR soil moisture retrievals. *IEEE Geoscience and Remote Sensing Symposium 2022*.
- Pierdicca, N., Pulvirenti, L., & Bignami, C. (2010), Soil moisture estimation over vegetated terrains using multitemporal remote sensing data. *Remote Sensing of Environment*, 114(2), 440-448.
- Pierdicca, N., Pulvirenti, L., & Pace, G. (2013), A prototype software package to retrieve soil moisture from Sentinel-1 data by using a bayesian multitemporal algorithm. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(1), 153-166.
- Pekel, J. F., A. Cottam, N. Gorelick, and A. S. Belward (2016), High-resolution mapping of global surface water and its long-term changes, *Nature*, 540, 418-422.
- Peng, J., Loew, A., Merlin, O., & Verhoest, N. E. (2017). A review of spatial downscaling of satellite remotely sensed soil moisture. *Reviews of Geophysics*, 55(2), 341-366.
- Roberts, C.D., Senan, R., Molteni, F., Boussetta, S., Mayer, M., Keeley, S.P.E. (2018), Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP. *Geosci. Model Dev.* 11, 3681–3712. <https://doi.org/10.5194/gmd-11-3681-2018>.

Rosen, P. A., Hensley, S., Shaffer, S., Veilleux, L., Chakraborty, M., Misra, T., ... & Satish, R. (2015). The NASA-ISRO SAR mission-An international space partnership for science and societal benefit. In *2015 IEEE Radar Conference (RadarCon)* (pp. 1610-1613). IEEE.

Rosen, P., Hensley, S., Shaffer, S., Edelstein, W., Kim, Y., Kumar, R., ... & Sagi, R. (2017). The NASA-ISRO SAR (NISAR) mission dual-band radar instrument preliminary design. In *2017 IEEE international geoscience and remote sensing symposium (IGARSS)* (pp. 3832-3835). IEEE.

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(5), 1254-1266.

Shiroma, G. H. X., M. Lavalley, and S. M. Buckley (2022), An Area-Based Projection Algorithm for SAR Radiometric Terrain Correction and Geocoding, *IEEE Trans. Geosci. Remote Sensing*, 60, doi:10.1109/tgrs.2022.3147472.

Tabatabaeenejad, A., M. Burgin, and M. Moghaddam (2012), Potential of L-Band Radar for Retrieval of Canopy and Subcanopy Parameters of Boreal Forests, *IEEE Trans. Geosci. Remote Sensing*, 50(6), 2150-2160, doi:10.1109/tgrs.2011.2173349.

Truong-Loi, M., A. Freeman, P. Dubois-Fernandez, and E. Pottier, "Estimation of soil moisture and Faraday rotation from bare surfaces using compact polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3608–3615, Nov. 2009.

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., 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.

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

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

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

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

Tsendbazar, M. H., L. Li, A. Tarko, S. de Bruin, D. Masiliunas, M. Lesiv, S. Fritz, M. Buchhorn, B. Smets, R. Van De Kerchove, M. Duerauer (2021). Towards operational validation of annual global land cover maps *Remote Sens. Environ.*, 266 (2021), Article 112686, 10.1016/J.RSE.2021.112686.

Ulaby, F. T., A. K. Moore, and A. K. Fung (1986), "Microwave remote sensing: active and passive," vol. 3. Norwood, MA: Artech House, pp. 1064.

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.

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, Linlin, Tsendbazar, N.E., Ramoino, F., Arino, O., 2021. ESA WorldCover 10 m 2020 v100. <https://doi.org/10.5281/zenodo.5571936>

Zwieback, S. and A. A. Berg, "Fine-scale SAR soil moisture estimation in the subarctic tundra," EarthArXiv Preprints, doi: 10.31223/osf.io/kp5xd, 2018.

Zwieback, S., Hensley, S., & Hajnsek, I. (2015a). Assessment of soil moisture effects on L-band radar interferometry. *Remote Sensing of Environment*, 164, 77-89.

Zwieback, S., Hensley, S., & Hajnsek, I. (2015b). A polarimetric first-order model of soil moisture effects on the DInSAR coherence. *Remote Sensing*, 7(6), 7571-7596.

## APPENDIX A: Active-Passive Synergy

Reflectivity ( $r$ ) and emissivity ( $e$ ) are the physical attributes of the microwave emissions from the Earth surface. They are physically and fundamentally related through the simple equation  $e = 1 - r$ . The microwave radar uses  $r$  and the radiometer uses  $e$  from the Earth's surface targets to observe the microwave radiometric characteristics. The synergy between radar and radiometer stems from the equation  $e = 1 - r$ . Thus, these observations can be synergistically used to derive many microwave radiometric characteristics of the Earth's surface target. The SMAP mission has demonstrated such active (SAR) and passive (radiometer) synergy to obtain high-resolution and reasonably accurate soil moisture information for the top soil (~5 cm) [Entekhabi and Das, 2014, Das et al., 2014, Das et al., 2018]. The NISAR mission present such an opportunities to merge the L-band SAR backscatter observations with the L-band radiometer observations from the SMAP and the SMOS satellites with careful considerations, such as: a) minimal observation time difference between SAR and radiometer platforms, and; b) taking the effect of SAR and radiometer observation geometries in synergistic computations. Such synergistic formulations to derive Earth's geophysical characteristics is possible; however, a great deal of ancillary information on soil physical properties, soil temperature, and vegetation-water-content data is required to retrieve soil moisture, leading to increase in degree of freedom. The SMAP mission operationally generates an active-passive soil moisture product [Das et al., 2019] and plans to include NISAR data in operation to synergistically generate a high-resolution soil moisture product.

## APPENDIX B: NISAR Water body detection

The baseline algorithm is based on detection thresholds identifying those areas with either inundated vegetation or open water, as this has been the most commonly used and effective algorithm found in the scientific literature. A recent example that describes a procedure similar to that which would satisfy the NISAR requirement can be found in Chapman et al, 2015.

The basis of the algorithm is that, due to double bounce scattering in areas of inundated vegetation, the HH backscatter becomes significantly brighter than for non-inundated vegetation or other landscapes characteristics within areas prone to inundation. In contrast, the HV backscatter does not change appreciably.

Change detection can be a complimentary tool for discovering change in inundation extent. If a pixel is only partially inundated, the backscatter values may not reach threshold values for detecting inundation. However, change detection over a time sequence of images in wetland areas with backscatter values at least as large as those associated with vegetation can be used to detect subtle increases or decreases in brightness as result of a change in flooding extent. Those areas detected as moving from non-inundated vegetation to inundated or vice-versa can be compared with those areas classified using threshold values to refine the classification.

In open water, the backscatter is generally significantly lower than non-inundated landscapes. Depending on the noise equivalent  $\sigma_0$  of NISAR, distinguishing open water from other low backscatter targets such as bare ground may be difficult. Additionally, at small incidence angles, wind roughening can make open water brighter than typical open water values, as can the presence of floating vegetation. Because open water area does not generally change quickly or to a large degree compared to the extent of inundated vegetation, averaging the data over time or over area can be helpful for reducing the noise. Since the incidence angle of a ground target for ascending and descending orbit tracks will, in general, be different, some open water areas not classified as open water at low incidence angles may be observed from imagery at the alternate orbit direction and correctly classified. Another potential approach for distinguishing open water from bare ground may be to examine the interferometric coherence: if the backscatter is above the noise equivalent  $\sigma_0$ , the coherence over bare ground should be higher than that over open water.

The subcategory of floating vegetation is not classified, but may be classified as an inundated area and therefore meet the objectives of the requirement. Alternative data sets could be used to separately categorize these conditions, such as S-band data from NISAR or using data from other spaceborne sensors. The algorithm flow of the L2-SCI-677 is presented in Figure 10.1.

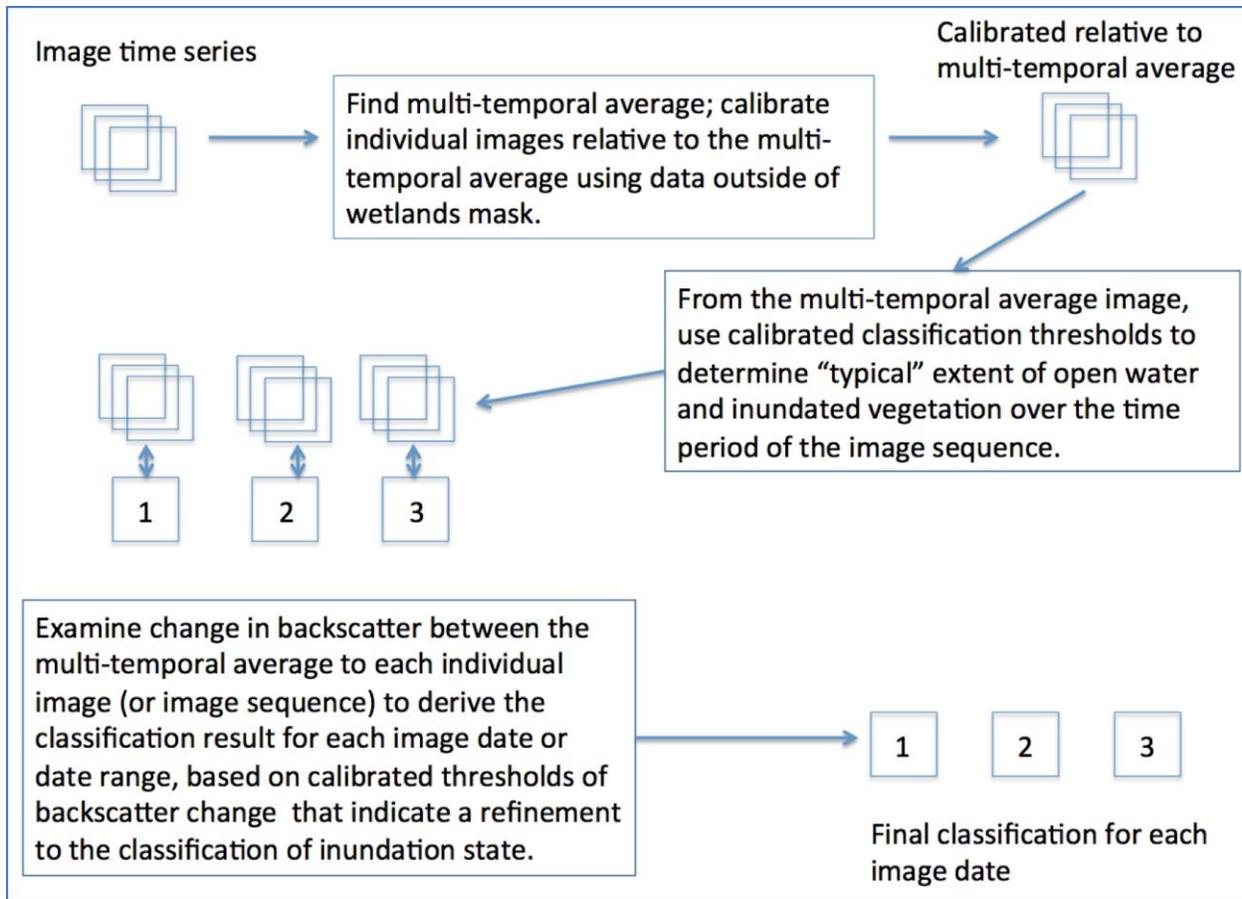


Figure 10.1 Algorithm flow for an example multi-temporal sequence of 3 images

## APPENDIX C: Acronyms

ADT	Algorithm Development Team
ANF	Area Normalization Factor
AT	Along Track
ATBD	Algorithm Theoretical Basis Document
AWS	Amazon Web Services
BFPQ options)	Block (adaptive) Floating-Point Quantization (adaptive may indicate implementation options)
Cal/Val	Calibration and Validation (also sometimes cal/val)
CDR	Critical Design Review
CDL	Cropland Data Layer
CF	Climate and
Forecast	
CPU	Central Processing Unit
CRSD	Calibration Raw Signal Data
CSV	Comma-separated values
DAAC	Distributed Active Archive Center
DBF	Digital Beam Forming
DEM	Digital Elevation Model
DM	Diagnostic Mode
DN	Digital Number
DSG	Disaggregation
EAR	Export Administration Regulations
EASE	Equal-Area Scalable Earth
ECMWF	European Centre for Medium-Range Weather Forecasts
ECEF	Earth Centered Earth Fixed
ER#.#	Engineering Release #.#
ERA5	ECMWF Reanalysis 5th generation
FFT	Fast Fourier Transform
FM	Frequency Modulation
FOE	Forecast Orbit Ephemeris
FOV	Field of View
GCOV	Geocoded Polarimetric Covariance (L2_GCOV)
GCP	Ground Control Point
GDAL	Geospatial Data Abstraction Library
GDS	Ground Data System
GeoTIFF	Geographic Tagged Image File Format
GIS	Geographic Information System
GMTED	Global Multi-resolution Terrain Elevation Data
GNSS	Global Navigation Satellite System
GOFF	Geocoded Pixel Offsets (L2_GOFF)
GPU	Graphics Processing Unit
GSLC	Geocoded Single Look Complex (L2_GSLC)
GUNW	Geocoded Unwrapped Interferogram (L2_GUNW)

HH	Horizontal-transmit, Horizontal-receive polarization
HK, HKTm	Housekeeping Telemetry
HDF5	Hierarchical Data Format version 5
HV	Horizontal-transmit, Vertical-receive polarization
ICU	Integrated Correlation Unit
InSAR	Interferometric Synthetic Aperture Radar
ISCE	InSAR Scientific Computing Environment
ISCE3	InSAR Scientific Computing Environment Enhanced Edition (for NISAR)
ISO	International Organization for Standardization
ISRO	Indian Space Research Organisation (British spelling)
JPL	Jet Propulsion Laboratory
JSON	JavaScript Notation
L0B	Level-0B (data)
L1	Level-1 (data)
L2	Level-2 (data)
L3	Level-3 (data)
LIA	Local Incidence Angle
LRR	[JPL] Limited Release Request
LRS	[JPL] Limited Release System
LUT	Lookup Table
Mbps	Megabits per second
MHz	Megahertz
MOE	Medium-precision Orbit Ephemeris
NASA	National Aeronautics and Space Administration
NETCDF4	Network Common Data Format 4 (also netCDF4)
NISAR	NASA-ISRO Synthetic Aperture Radar
NOE	Near-Realtime Orbit Ephemeris
OpenMP	Open Multi-Processing
PALSAR	Phased Array L-band Synthetic Aperture Radar
PCM	Process Control Management
PDF	Portable Document Format (often pdf)
PDR	Preliminary Design Review
PMI	Physical Model Inversion
POD	Precision Orbit Determination
POE	Precision Orbit Ephemeris
PRF	Pulse Repetition Frequency
QA	Quality Assurance
R#.#	Release #.# (.0 often not used)
REE	Radar Echo Emulator
RFI	Radio Frequency Interference
RIFG	Range-Doppler Interferogram (L1_RIFG)
RMS	Root Mean Square
RMSE	Root Mean Square Error
ROFF	Range-Doppler Pixel Offsets (L1_ROFF)
RRSD	Raw Radar Signal Data
RRST	Raw Radar Signal Telemetry

---

RSLC	Range-Doppler Single Look Complex (L1_RSLC)
RTC	Radiometric Terrain Correction
RUNW	Range-Doppler UnWrapped Interferogram (L1_RUNW)
RV	Right-circular, V-receive compact polarization
SAR	Synthetic Aperture Radar (L-SAR: L-band. S-SAR: S-band)
SAS	Science Algorithm Software
SDS	Science Data System
SDT	Science Definition Team
SIS	Software Interface Specification
SLC	Single Look Complex
SME2	Soil Moisture product based on a 200-meter global EASE Grid projection
SMAP	Soil Moisture Active Passive (Mission)
SMOS	Soil Moisture and Ocean Salinity
SNAPHU	Statistical-cost, Network-flow Algorithm for Phase Unwrapping
SWST	Sampling Window Start Time
SRTM	Shuttle Radar Topography Mission
ST	Science Team
SWST	Sampling Window Start Time
TAI	International Atomic Time (Temps Atomique International)
TCF	Terrain Correction Factor
TEC	Total Electron Content
TFdb	Trackframe Database
TSR	Time Series Ratio
UAVSAR	Uninhabited Aerial Vehicle Synthetic Aperture Radar
ubRMSE	unbiased RMSE
UR	Urgent Response
UTC	Universal Time Coordinated
UTM	Universal Transverse Mercator
VH	Vertical-transmit, Horizontal-receive polarization
VV	Vertical-transmit, Vertical-receive polarization
VWC	Vegetation Water Content
WGS84	World Geodetic System 84
XML	eXtensible Markup Language (xml in code)
YAML	YAML Ain't Markup Language