

UTILIZATION OF ANCILLARY DATA SETS FOR SMAP ALGORITHM DEVELOPMENT AND PRODUCT GENERATION

Peggy E. O'Neill¹, Erika Podest², and Eni G. Njoku²

¹Hydrological Sciences Branch / Code 614.3, NASA Goddard Space Flight Center,
Greenbelt, MD 20771 USA tel: 1-301-614-5773, fax: 1-301-614-5808,
Peggy.E.ONeill@nasa.gov

²Jet Propulsion Laboratory, 4800 Oak Grove Drive,
Pasadena, CA 91109 USA, tel: 1-818-354-3693
Erika.Podest@jpl.nasa.gov, Eni.G.Njoku@jpl.nasa.gov,

ABSTRACT

Once launched in late 2014/early 2015, the Soil Moisture Active Passive (SMAP) mission will provide high resolution global mapping of soil moisture and landscape freeze/thaw state every 2-3 days. These measurements are valuable to improved understanding of the Earth's water, energy, and carbon cycles, and to applications of societal benefit. In order for soil moisture and freeze/thaw to be retrieved accurately from SMAP microwave data, a variety of global static and dynamic ancillary data are required. The choice of which ancillary datasets to use for SMAP products will be based on a number of factors including availability and ease of use, their inherent error and resulting impact on SMAP retrieval accuracies, and compatibility with similar choices made by ESA's SMOS mission.

Keywords (Index Terms) -- soil moisture, microwave, geophysical retrievals, ancillary data.

1. INTRODUCTION

The Soil Moisture Active Passive (SMAP) mission* is one of the first Earth observation satellites being developed by NASA in response to the National Research Council's Decadal Survey, *Earth Science and Applications from Space*:

*The SMAP mission has not been formally approved by NASA. The decision to proceed with the mission will not occur until the completion of the National Environmental Policy Act (NEPA) process. Discussions in this paper related to SMAP are being made available for information purposes only.

National Imperatives for the Next Decade and Beyond [1]. Scheduled to launch in late 2014/early 2015, SMAP will provide high resolution and frequent revisit global mapping of soil moisture and freeze/thaw state, utilizing enhanced Radio Frequency Interference (RFI) mitigation approaches to collect new measurements of the hydrological condition of the Earth's surface. The SMAP instrument design incorporates an L-band radar (3 km) and an L band radiometer (40 km) sharing a single 6-meter rotating mesh antenna to provide measurements of soil moisture and landscape freeze / thaw state (Figure 1) [2]. These observations will (1) improve our understanding of linkages between the Earth's water, energy, and carbon cycles, (2) benefit many application areas including numerical weather and climate prediction, flood and drought monitoring, agricultural productivity, human health, and national security, (3) help to address priority questions on climate change, and (4) potentially provide continuity with brightness temperature and soil moisture measurements from ESA's SMOS (Soil Moisture Ocean Salinity) and NASA's Aquarius missions.

In the SMAP prelaunch time frame, baseline algorithms are being developed for generating (1) soil moisture products both from radiometer measurements on a 36 km grid and from combined

radar/radiometer measurements on a 9 km grid, and (2) freeze/thaw products from radar measurements on a 3 km grid. These Level 2 retrieval algorithms need a variety of global ancillary data, both static and dynamic, to run the retrieval models, constrain the retrievals, and provide flags for indicating retrieval quality. The choice of which ancillary dataset to use for a particular SMAP product will be based on a number of factors, including its availability and ease of use, its inherent error and resulting impact on the overall soil moisture or freeze/thaw retrieval accuracy, and its compatibility with similar choices made by the SMOS mission. All decisions regarding SMAP ancillary data sources will be fully documented by the SMAP Project and made available to the user community.

2. SMAP ANCILLARY DATA

Ancillary data needed by the SMAP mission fall into two categories -- static ancillary data are data that do not change during the mission while dynamic ancillary data require periodic updates in time frames ranging from seasonally to daily. Static data include parameters such as permanent masks (land / water / forest / urban / mountain), the grid cell average elevation and slope derived from a DEM, permanent open water fraction, and soils information (primarily sand and clay fraction). The SMAP Project plans to create a master file of all of the static ancillary data resampled to the same grids as the output products, and this master file would be available to any algorithm or end user who needs it. The dynamic ancillary data include land cover, surface roughness, precipitation, vegetation parameters, and effective soil temperatures. Measurements from the SMAP radar will be used to provide information primarily on transient water and frozen ground. Table 1 lists the fourteen ancillary data parameters identified as required by one or more of the SMAP product algorithms.

Ancillary data will also be employed to set flags which help to determine either specific aspects of the processing (such as corrections for

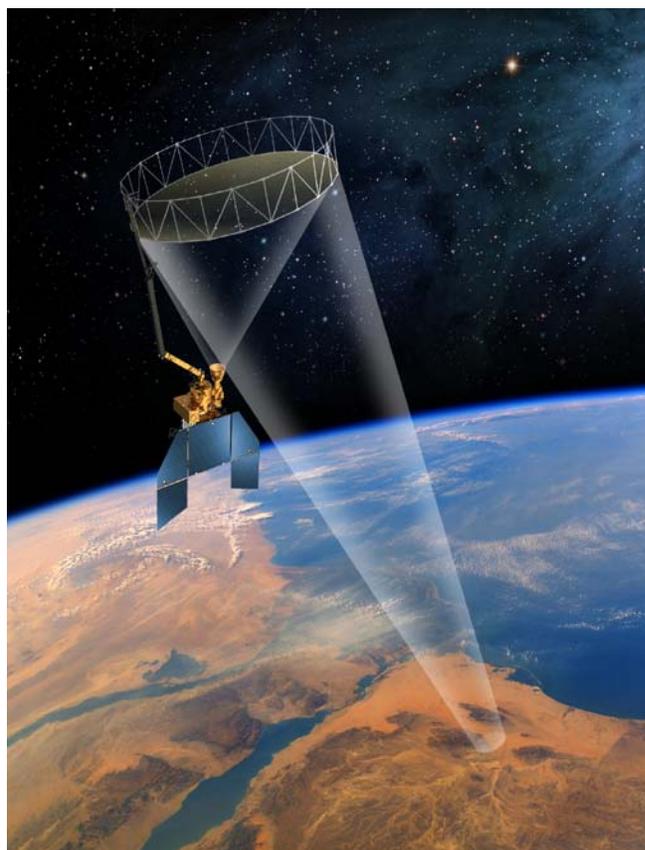


Figure 1. The SMAP mission concept consists of an L-band radar and radiometer sharing a single spinning 6-m mesh antenna in a sun-synchronous dawn/dusk orbit.

transient water) or the quality of the retrievals (e.g. precipitation flag). Basically, these flags would provide information as to whether the ground is frozen, snow-covered, or flooded, or whether it is actively precipitating at the time of the satellite overpass. Other flags will indicate whether masks for steeply sloped topography, or for urban, heavily forested, or permanent snow/ice areas are in effect.

3. ANCILLARY DATA SELECTION PROCESS

A thorough ancillary dataset selection process has been undertaken in order to consider all datasets available for each parameter and to ensure consistency in the datasets used throughout the different retrieval algorithms for the different SMAP baseline products. Consistency with SMOS products is also a consideration.

1	Soil Temperature
2	Surface Air Temperature
3	Vegetation Water Content (VWC)
4	Sand & Clay Fraction
5	Urban Area
6	% Permanent Open Water
7	Crop Type
8	Land Cover Class
9	Precipitation
10	Snow
11	Mountainous Area
12	Permanent Ice
13	<i>b</i>, <i>ω</i>, & <i>τ</i> Vegetation Parameters
14	<i>h</i> Roughness Parameter

Key factors considered in the ancillary data selection are latency, spatial resolution, temporal resolution, global coverage, accessibility/ease of use, and quality checks/internal error. Once a dataset has been selected as the primary source for a given parameter, it is transferred onto the SMAP testbed for regridding in a consistent manner to the 1 km, 3 km, 9 km, and 36 km SMAP EASE grids. Different SMAP algorithms operate at different spatial resolutions, but all SMAP grids are nested. Flags and masks are also generated for the appropriate parameters. Documentation of the selection rationale for a given dataset will include specification of a primary and secondary dataset as well as anticipated new and improved datasets that will potentially replace the selected one by the time SMAP is in orbit.

An example of a selected ancillary dataset is shown in Figure 2. Soil moisture retrieval algorithms require information about soil texture, specifically sand and clay fraction. A global dataset was assembled from an optimized combination of the FAO (Food & Agriculture Organization), HWSO (Harmonized World Soil Database), STATSGO (State Soil Geographic—US), NSDC (National Soil Database Canada), and ASRIS (Australian Soil Resources Information System) soil databases. This composite dataset

uses the best available source for a given region [3], which should improve the accuracy of SMAP products in that region as well as providing consistency with the work of local scientists and end users in that region. A negative consequence of this decision is the potential for discontinuities at international boundaries, such as between the United States and Canada. The SMOS mission, in contrast, has chosen to use the FAO soils database in generating their global products.

Similar analyses have been conducted for selection of some of the other ancillary parameter data bases. The area fraction of permanent open water is required both to discriminate between land and water pixels and to provide ancillary information for a correction to the SMAP brightness temperatures due to the presence of open water (a transient water fraction will be generated from the SMAP radar data). The MODIS44W global static product at 250 m resolution, which is endorsed by the MODIS Land Science Team, has been chosen for use by the SMAP mission [4]. Another MODIS product, the MODIS IGPB global product at 250 m resolution, is the current baseline dataset to provide land cover class information.

Selection of other datasets is currently in progress. (a) Preliminary analysis of ECMWF (European Centre for Medium-Range Weather Forecasts) and GMAO (Goddard Global Modeling & Assimilation Office) soil and air temperature products indicates that both sources may meet the 2°C accuracy desired for soil moisture retrieval [5]. (b) Urban areas are problematic since manmade construction masks the signal from the underlying soil, and cities are also often a source of radio frequency interference. The Global Rural-Urban Mapping Project (GRUMP) dataset, a merged product of NOAA's night-time lights dataset with United Nations census information, will be used to produce an urban/rural mask [6]. (c) A JPL merged DEM combining Shuttle radar and other topography data will be baselined until the new USGS GMTED2010 250 m – 1000 m resolution DEM is released and evaluated.

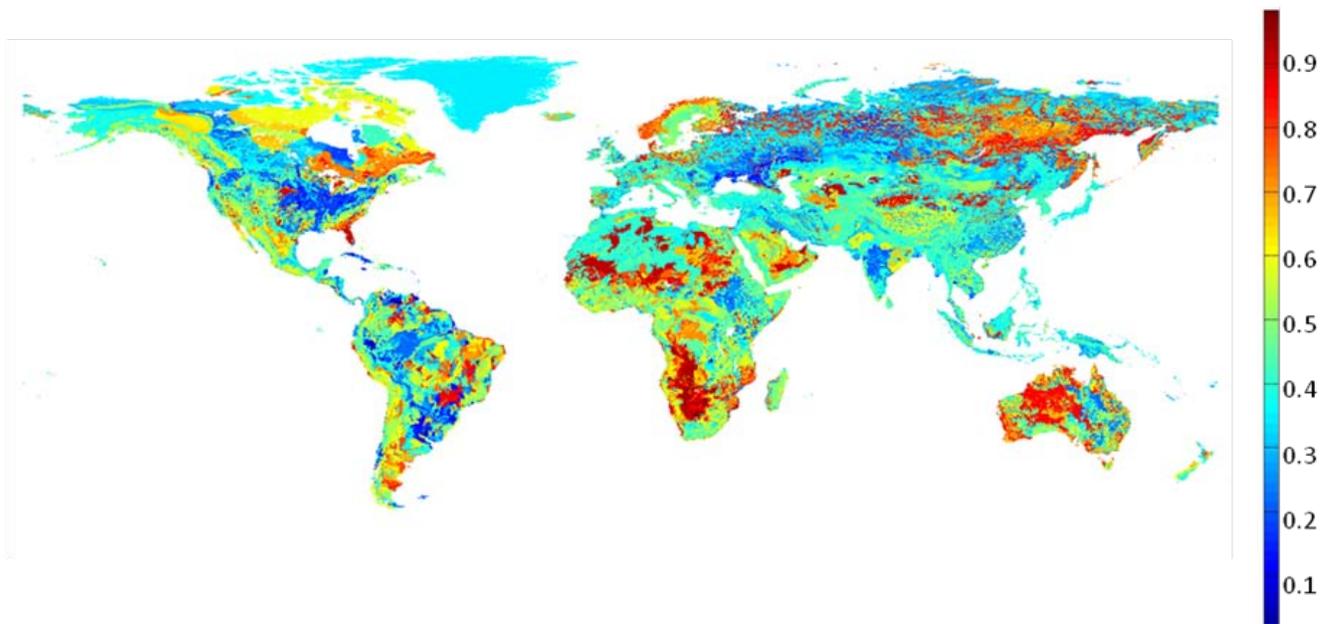


Figure 2. Global sand fraction at 0.01 degree resolution based on a composite of FAO, HWSD, STATSGO, NSDC, and ASRIS datasets using best available source for a given region.

6. SUMMARY

Algorithms being developed for the SMAP mission require a variety of both static and ancillary data on a global scale. Fourteen different ancillary data parameters have been identified by the SMAP team needed for the Level 2 product generation. The selection of the most appropriate source for each ancillary data parameter is driven by a number of considerations, including accuracy, latency, availability, and consistency across all SMAP products and with SMOS. It is anticipated that initial selection of all ancillary datasets, which are needed for ongoing algorithm development activities on the SMAP algorithm testbed at JPL, will be completed within the year. These datasets will be updated as new or improved sources become available, and all selections and changes will be documented for the benefit of the user community. Wise choices in ancillary data will help to enable SMAP to provide new global measurements of soil moisture and freeze/thaw state at the targeted accuracy necessary to tackle hydrologically-relevant societal issues.

7. REFERENCES

- [1] National Research Council, "Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond," pp. 400, 2007.
- [2] Entekhabi, D., E. Njoku, P. O'Neill, K. Kellogg, plus 19 others, "The Soil Moisture Active Passive (SMAP) Mission," *Proceedings of the IEEE*, Vol. 98, No. 5, May, 2010.
- [3] Das, N. and P. O'Neill, "Selection of Soil Attributes Datasets for the SMAP Mission," SMAP Science Document #xz, Draft 1.1, JPL, Dec., 2010.
- [4] Steven Chan, "Selection of a Water-Land Fraction Dataset for the SMAP Mission," SMAP Science Document #xx, Draft 1, JPL, November, 2010.
- [5] Holmes, T., T. Jackson, R. Reichle, and J. Basara, "An Assessment of Surface Soil Temperature Products from Numerical Weather Prediction Models using Ground-based Measurements," submitted to *Water Resources Research*, 2011.
- [6] Das, N., "Evaluation of Urban/Rural Datasets for the SMAP Mission," SMAP Science Document #yy Draft 1.0, Jet Propulsion Laboratory, March, 2011.