Soil Moisture Active Passive (SMAP)

Ancillary Data Report

Urban Area

Preliminary, v.1.1
SMAP Science Document no. 046

Narendra Das
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA

January 31, 2013
JPL D-53060

Jet Propulsion Laboratory
California Institute of Technology

Preface

The SMAP Ancillary Data Reports provide descriptions of ancillary data sets used with the science algorithm software in generation of the SMAP science data products. The Ancillary Data Reports may undergo additional updates as new ancillary data sets or processing methods become available. The most recent versions of the ancillary data reports will be made available, along with the Algorithm Theoretical Basis Documents (ATBDs), at the SMAP web site http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

Table of Contents

1 Overview

1.1 Purpose

The purpose of this report is to describe an urban area data set for use in developing quality flags and masks for the SMAP science data products. The urban area dataset is one of a suite of ancillary datasets used by the SMAP science processing algorithms. The algorithms and ancillary data are described in SMAP algorithm theoretical basis documents (ATBDs) and ancillary data reports. The ATBDs and ancillary data reports are listed in Appendices A and B and are available at the SMAP web site:
http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

1.2 Requirement

The SMAP radiometer and radar instruments provide global measurements of the brightness temperature and backscatter cross-section of the land surface, from which global maps of soil moisture and freeze-thaw state are derived. The presence of man-made structures within the sensor footprint or pixel, such as buildings, roads and other artificial or impervious surfaces, contaminates the received brightness temperature and backscatter signals and can lead to errors in the geophysical retrievals. Thus, some means for identifying urban areas where these structures occur predominantly is necessary. Urban areas are also areas of potential radio-frequency-interference (RFI) from man-made sources such as radio transmitters, terrestrial communications links, airport radars, cell-phones, etc., that can contaminate the radar and radiometer measurements. For these reasons, a global dataset indicating urban area is required as a means for generating quality flags for the SMAP data products.

2 Dataset Description and Selection

2.1 Source Datasets

A number of regional and global datasets that include urban area classification are available. These datasets are created using input data from satellite sensors (LANDSAT, SPOT, MODIS, DMSP-OLS), census data, and urban area polygon maps to create digital urban area extent maps at spatial resolution of ~0.5 km to ~10 km. The following table (Table 1) provides a list of the primary urban datasets, their attributes, and the input data used to create them.

After evaluating the characteristics of the datasets in Table 1, the MOD500 and GRUMP datasets were considered the most suitable for developing of an urban mask for SMAP based on the following criteria:

a) High spatial resolution (<= 1 km)
b) Clear binary distinction between rural and urban pixels
c) Better delineation of urban extent

The other datasets from Table 1 were excluded from further consideration since they do not satisfy one or more of the above criteria. The MOD500 and GRUMP data were evaluated further to select one of them as the primary source for the SMAP urban area dataset.

Table 1: List of available urban related datasets.

S. no	Name of the dataset	Inputs to create the urban extent data	Spatial Resolution	Map type
1	Vector Map Level Zero (VMAP0) [Danko 1992]	Navigational charts and maps	1:1 million	Thematic vector maps
2	Global Land Cover 2000 v1.1 (GLC00) [Bartholome et al. 2005]	One year of SPOT data	~1 km	Thematic (22 classes)
3	History Database of the Global Envt. V3 (HYDE3) [Goldewijk 2005]	LSCAN based population density, UN population estimates, and city gazetteers	~10 km	Continuous (% urban)
4	Global Impervious Surface Area (IMPSA) [Elvidge et al. 2007]	LSCAN, LITES, and LANDSAT data for training	~1 km	Continuous (%impervious surface)
5	MODIS Urban Land Cover 500m (MOD500) [Schneider et al. 2009]	One year of MODIS data from Terra and Aqua, and LANDSAT data for training and validation	~0.5 km	Binary (urban / rural)
6	MODIS Urban Land Cover 1000m (MOD1K) [Schneider et al. 2003]	One year of MODIS data from Terra and Aqua, LANDSAT data for training and validation, and LITES	~1 km	Binary (urban / rural)
7	Global Rural-Urban Mapping project (GRUMP) [CIESIN 2004]	VMAP0, LITES, and Census	~1 km	Binary (urban / rural)
8	Nighttime Lights v2 (LITES) [Elvidge et al. 2001]	One year of DMSP-OLS dataset	~1 km	Continuous (nighttime illumination intensity)
9	LandScan 2005 (LSCAN) [Bhaduri et al. 2002]	MODIS1K, VMAP, Census data and LANDSAT	~1 km	Continuous (ambient human population density)

2.2 MOD500

The MOD500 dataset is created from Moderate Resolution Imaging Spectroradiometer (MODIS) data from the MODIS instrument aboard NASA's Terra and Aqua satellites [Schneider et al. 2009]. The dataset uses the 8-day Nadir BRDF-Adjusted Reflectance (NBAR) values for the seven land bands (~500 m resolution) for a one year period (BRDF is the Bidirectional Reflectance Distribution Function). MOD500 employs manual interpretation of medium-resolution (~30 m) LANDSAT imagery and other ancillary data sets to construct training sites for supervised classifications. From the supervised classification, the urban areas are locations that are dominated by a 'built environment'. A 'built environment' includes all non-vegetative, human-constructed elements, such as buildings, roads, runways, etc. (i.e. a mix of human-made surfaces and materials). 'Dominated' implies coverage greater than or equal to 50% of a given landscape unit (here, the pixel). Pixels that are predominantly vegetated (e.g. a park) are not considered urban, even though in terms of land use, they may function as urban space. Minimum mapping units for urban areas are contiguous patches of built-up land greater than 1 km^2.

2.3 GRUMP

The GRUMP dataset is a merged product of NOAA's night-time lights dataset (LITES) [Elvidge et al., 2001], digital maps and United Nations census information. LITES is a composite dataset of stable "city" lights using time series data from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) "night-time lights" dataset for the period from 1 October 1994 to 30 April 1995. The pixel values are measurements of the frequency with which lights were observed, normalized by the total number of cloud-free observations. The digital maps are the global-scale dataset: Digital Chart of the World's Populated Places (DCW), an ESRI product originally developed for the US Defense Mapping Agency (DMA) using DMA data and currently available at 1:1,000,000 scale (1993 version). The "populated places" (from UN census information) coverage is available for most countries and contains depictions of the urbanized areas (built-up areas) of the world that are represented as polygons at 1:1,000,000 scale. Thus, the GRUMP dataset merges anthropogenic presence and night-time lights to delineate urban extent.

2.4 Dataset Selection

Figure 1 illustrates clearly the difference between the MOD500 and GRUMP datasets. The input data and the approaches used to create MOD500 and GRUMP are entirely different. The 50% built-up threshold in MOD500 to define urban extent explains why vegetated regions appear within a city's perimeters. However, the GRUMP data for these metropolitan urban extents (Guanzhou, London, and Washington) do not discriminate between vegetation and built-up areas within the city perimeters because of the night-time light and human presence. The GRUMP data encompass the whole metropolitan urban area and thus provides a more conservative urban extent than the MOD500 dataset (Figure 1). The urban area extents for the MOD500 and GRUMP datasets are 0.5% and 2.5% of the total global landmass.

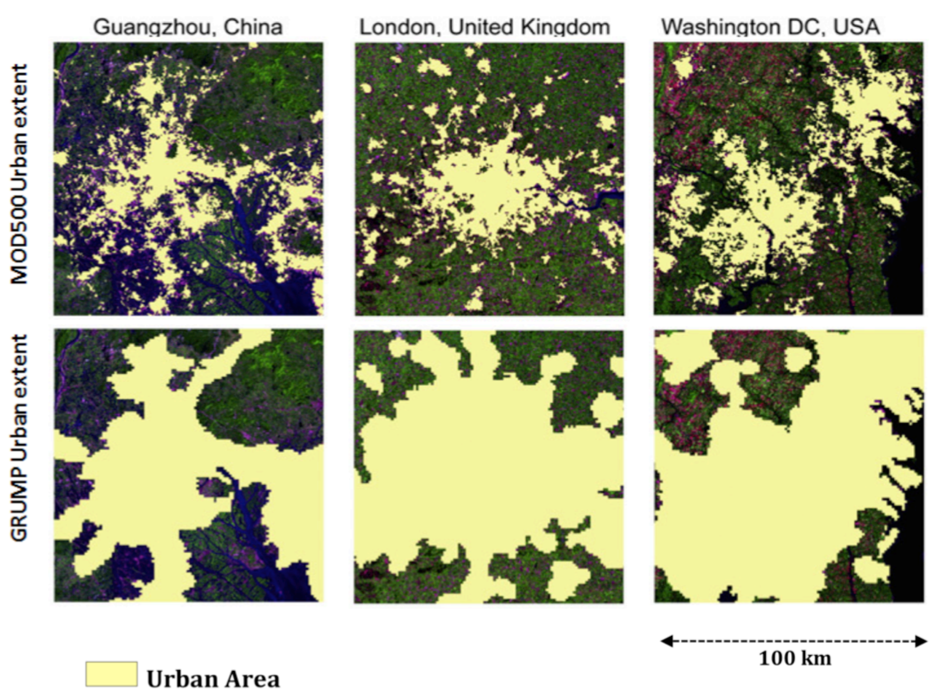


Figure 1. Comparison of MOD500 and GRUMP masks for three metropolitan areas: Guangzhou, China; London, United Kingdom; and Washington D.C.-Baltimore, U.S.A. [Schneider et al. 2010].

With respect to the SMAP instrument footprint and its sensitivity to urban structures and RFI, the GRUMP dataset seems more appropriate for use by SMAP because it clearly delineates urban and rural boundaries and has potential to provide a conservative urban extent mask/flag. The urban mask provided by the GRUMP dataset will also be of benefit in masking other coarse resolution ancillary datasets such as surface temperature, surface albedo and soil texture within urban areas.

3 Processing

3.1 Source Data Characteristics

The Global Rural-Urban Mapping Project (GRUMP) dataset is an urban/rural extent digital map available at the Center for International Earth Science Information Network (CIESIN). Documentation is available at http://sedac.ciesin.columbia.edu/gpw/. The data file is in ASCII format. The first lines of the file specify the data resolution, the number of rows and columns, and the null data value. The data are stored in geographic coordinates of decimal degrees based on the World Geodetic System spheroid of 1984 (WGS84) with 30 arc-second (~1-km) resolution. In the ASCII data file the integer values for rural, urban, and null pixel are 1, 2 and -9999, respectively, the null value represents water bodies.

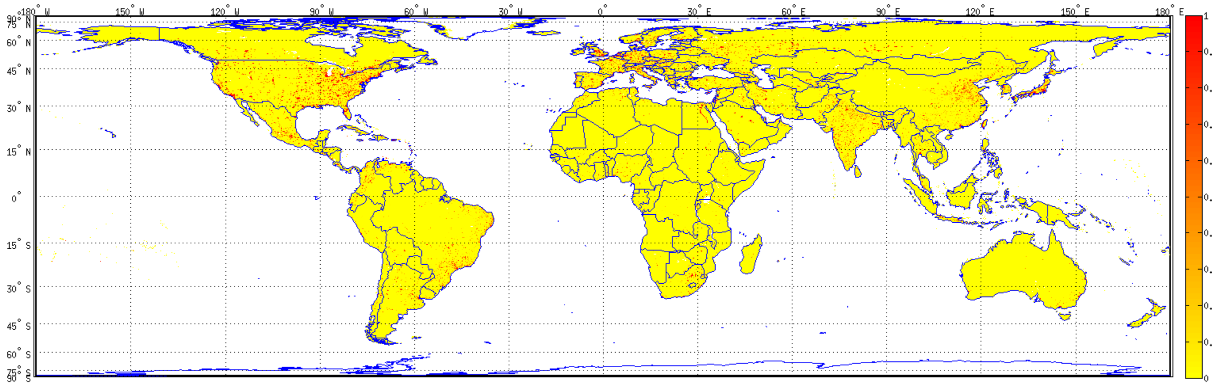
The GRUMP data suffer from a blooming effect due to the use of night-time light. The blooming effect creates an overestimation of the actual urban extent, which is most likely dependent on the intrinsic characteristics of the sensor [Elvidge et al., 2004]. A more detailed comparison of lighted area with built-area estimates from Landsat imagery of 17 cities worldwide [Small et al., 2005] shows that lighted areas are consistently larger than the estimates of built areas, and thresholds >90% of urban area reconcile lighted area with built area.

3.2 Processing and Regridding

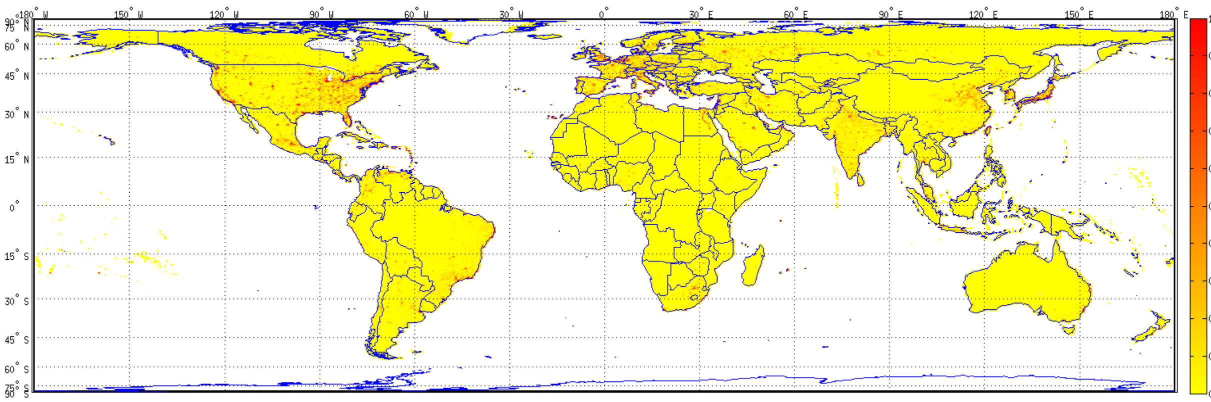
The GRUMP dataset was obtained online from http://sedac.ciesin.columbia.edu/gpw/. The GRUMP data have a native resolution of 30 arc-second (~1 km) and were re-gridded to the Equal-Area Scalable Earth Grid 2.0 (EASE2) projection (reference available at: *nsidc.org/data/ease/versions.html*) at 3, 9, and 36 km. The gridding process used was a simple linear averaging (drop-in-the-bucket technique) where the rural pixels and null value pixels were ignored. This permitted the computation of urban fraction within the grid resolutions of 3, 9, and 36 km. The processed output dataset consists of the urban fraction data values. Examples of urban fraction maps gridded at 9 km and 36 km are shown in Figures 2a and 2b.

Figure 3 shows pixels having an urban fraction greater than 0.5. For SMAP usage, grid cells with an urban fraction greater than 0.25 are assigned a flag in the output processed dataset. However, each geophysical processing algorithm that uses the ancillary dataset can assign a threshold of urban fraction within a pixel, for masking purposes, according to its own requirements.

Histograms of the gridded GRUMP urban fraction data at 3, 9 and 36 km are presented in a tabular form in Table 2. At the native resolution of 30 arc-second (1 km) the GRUMP data indicate urban area covering about 2.5% of the global land mass, with an overestimation of <=10%. In the gridded data, the frequency of an urban area fraction between 0 and 1 is 3.81% for the 3-km EASE2 grid, 7.1% for the 9-km EASE2 grid and 21.4% for the 36-km EASE2 grid. Table 2 shows more details of the distribution of urban area fractions at different resolutions.

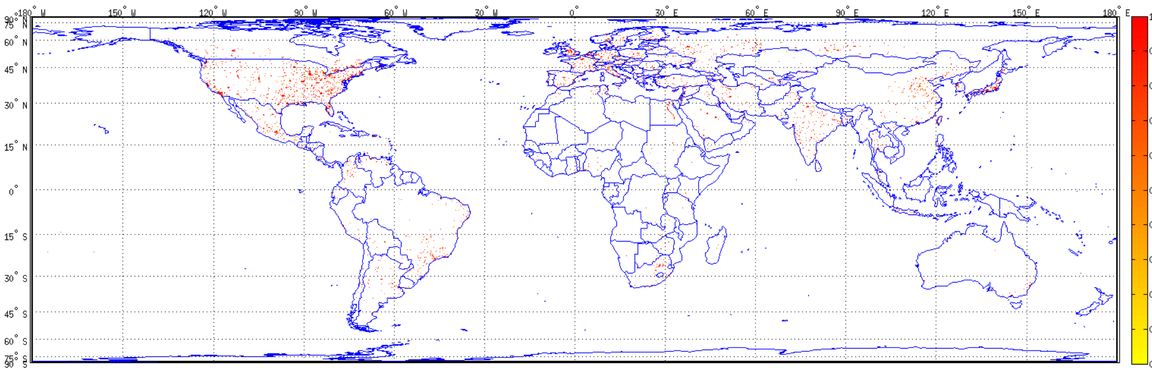


(a)

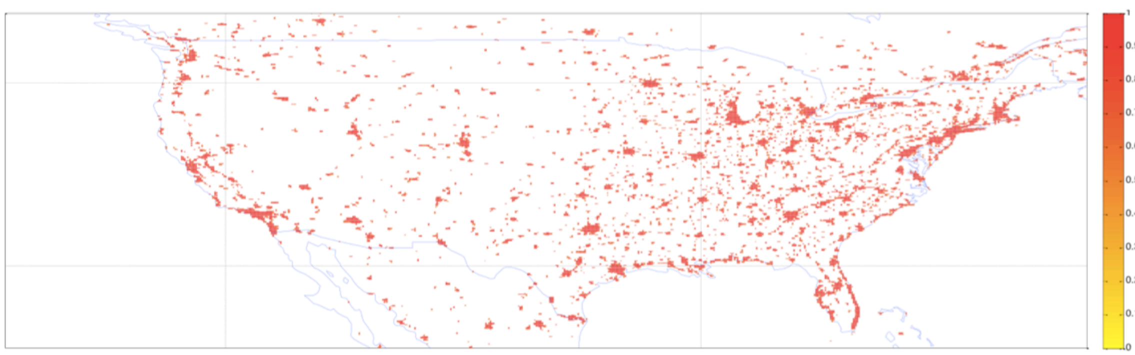


(b)

Figure 2: GRUMP-derived urban fractions at EASE2 grid resolutions: (a) 9 km, and (b) 36 km.



(a)



(b)

Figure 3: Gridded GRUMP data with urban fraction > 0.5 at on the 9-km EASE2 grid; (a) Global, and (b) Continental U.S.

Table 2. Distribution of urban area fraction at different EASE grid resolutions

Fraction of Urban area	**At 3 km** in (%)	**At 9 km** in (%)	**At 36 km** in (%)
0	96.19	92.90	78.60
> 0 and <=0.1	0.22	1.74	12.83
> 0.1 and <=0.2	0.28	1.05	3.92
> 0.2 and <=0.3	0.21	0.79	1.73
> 0.3 and <=0.4	0.23	0.65	1.03
> 0.4 and <=0.5	0.29	0.55	0.63
> 0.5 and <=0.6	0.16	0.43	0.39
> 0.6 and <=0.7	0.21	0.38	0.27
> 0.7 and <=0.8	0.19	0.37	0.26
> 0.8 and <=0.9	0.21	0.36	0.2
> 0.9	0.18	0.17	0.12

4 Data Availability

The processed and gridded urban fraction global datasets at 3, 9 and 36 km will be made available at the National Snow and Ice Data Center (NSIDC). The files are flat binary with data written in column order with little-endian byte ordering. In all the three files the null data/waterbodies pixel value is -9999. The remaining grids cells contain the urban fraction over landmass as described in Table 2. Appendix C provides the dataset description for all three gridded urban fraction datasets at 3, 9, and 36 km..

5 Acknowledgment

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

6 References

Bartholome, E., and A.S. Belward (2005): GLC2000: a new approach to global land cover mapping from Earth observation data, *International Journal of Remote Sensing, 26*, 1959-1977.

Bhaduri, B., E. Bright, P. Coleman, and J. Dobson (2002): LandScan: locating people is what matters, *Geoinfomatics, 5,* 34-37.

CIESIN (2004): *Global Rural-Urban Mapping Project (GRUMP): Urban/Rural Extents*, Center for International Earth Science Information Network (CIESIN), Columbia University Palisades, NY. Available at http://sedac.ciesin.columbia.edu/gpw/.

Danko, D.M. (1992): The Digital Chart of the World Project, *Photogrammetric Engineering and Remote Sensing*, *58,* 1125-1128.

Elvidge, C., B. T. Tuttle, P. C. Sutton, K. E. Baugh, A. T. Howard, C. Milesi, B. L. Bhaduri, and R. Nemani (2007): Global Distribution and Density of Constructed Impervious Surfaces, *Sensor,s 7,* 1962-1979.

Elvidge, C., M. L. Imhoff, K. E. Baugh, V. R. Hobson, I. Nelson, J. Safran, J. B. Dietz, and B. T. Tuttle (2001): Nighttime lights of the world: 1994-95, *ISPRS Journal of Photogrammetry and Remote Sensing, 56,* 81-99.

Goldewijk, K. (2005): Three centuries of global population growth: A spatially referenced population density database for 1700 – 2000, *Population and Environment, 26,* 343-367.

Schneider, A., M. Friedl, and D. Potere (2009): A new map of global urban extent from MODIS data, *Environmental Research Letters*, *4,* article 044003.

Appendix A: SMAP Science Data Products and ATBDs

The SMAP Algorithm Theoretical Basis Documents are available at the SMAP web site http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

Data Product	Description	ATBD
L1A_Radar	Radar raw data in time order	(Joint with L1C_S0_HiRes)
L1A_Radiometer	Radiometer raw data in time order	(Joint with L1B_TB)
L1B_S0_LoRes	Low resolution radar σ_o in time order	(Joint with L1C_S0_HiRes)
L1C_S0_HiRes	High resolution radar σ_o (half orbit, gridded)	West, R., L1B & L1C radar products, JPL D-53052, JPL, Pasadena, CA.
L1B_TB	Radiometer T_B in time order	Piepmeier, J. et al., L1B radiometer product, GSFC SMAP-006, GSFC, Greenbelt, MD.
L1C_TB	Radiometer T_B (half orbit, gridded)	Chan, S. et al., L1C radiometer product, JPL D-53053, JPL, Pasadena, CA.
L2_SM_A	Soil moisture (radar, half orbit)	Kim, S. et al., L2 & L3 radar soil moisture (active) product, JPL D-66479, JPL, Pasadena, CA.
L2_SM_P	Soil moisture (radiometer, half orbit)	O'Neill, P. et al., L2 & L3 radiometer soil moisture (passive) product, JPL D-66480, JPL, Pasadena, CA.
L2_SM_AP	Soil moisture (radar/radiometer, half orbit)	Entekhabi, D. et al., L2 & L3 radar/radiometer soil moisture (active/passive) products, JPL D-66481, JPL, Pasadena, CA.
L3_FT_A	Freeze/thaw state (radar, daily composite)	McDonald, K. et al., L3 radar freeze/thaw (active) product, JPL D-66482, JPL, Pasadena, CA.
L3_SM_A	Soil moisture (radar, daily composite)	(Joint with L2_SM_A)
L3_SM_P	Soil moisture (radiometer, daily composite)	(Joint with L2_SM_P)
L3_SM_AP	Soil moisture (radar/radiometer, daily composite)	(Joint with L2_SM_AP)
L4_SM	Soil moisture (surface & root zone)	Reichle, R. et al., L4 surface and root-zone soil moisture product, JPL D-66483, JPL, Pasadena, CA.
L4_C	Carbon net ecosystem exchange (NEE)	Kimball, J. et al., L4 carbon product, JPL D-66484, JPL, Pasadena, CA.

Appendix B: SMAP Ancillary Data Reports

The SMAP Ancillary Data Reports are available with the ATBDs at the SMAP web site http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

Data/Parameter	Ancillary Data Report
Crop Type	Kim, S., Crop Type, JPL D-53054, Pasadena, CA
Digital Elevation Model	Podest, E. et al., Digital Elevation Model, JPL D-53056, Pasadena, CA
Landcover Classification	Kim, S., Landcover Classification, JPL D-53057, Pasadena, CA
Soil Attributes	Das, N. et al., Soil Attributes, JPL D-53058, Pasadena, CA
Static Water Fraction	Chan, S. et al., Static Water Fraction, JPL D-53059, Pasadena, CA
Urban Area	Das, N., Urban Area, JPL D-53060, Pasadena, CA
Vegetation Water Content	Chan, S. et al., Vegetation Water Content, JPL D-53061, Pasadena, CA
Permanent Ice	McDonald, K., Permanent Ice & Snow, JPL D-53062, Pasadena, CA
Precipitation	Dunbar, S., Precipitation, JPL D-53063, Pasadena, CA
Snow	Kim, E. et al., Snow, GSFC SMAP-007, Greenbelt, MD
Surface Temperature	Fisher, J. et al., Surface Temperature, JPL D-53064 Pasadena, CA
Vegetation and Roughness Parameters	Colliander, A., Vegetation & Roughness Parameters, JPL D-53065, Pasadena, CA

Appendix C: Product Description

The global urban fraction files are available in three grid resolutions i.e., 3 km, 9 km, and 36 km. Mentioned below are the file names and respective characteristics:

File_Name	Format	Row	Col	Projection	Resolution	NoData	Precision
UrbanFraction.03km.4872x11568.float32.EZ2.bin	Binary	4872	11568	EASE2	3 km	-9999	real*4
Urban_Fraction.09km.1624x3856.float32. EZ2.bin	Binary	1624	3856	EASE2	9 km	-9999	real*4
Urban_Fraction.36km.406x964.float32. EZ2.bin	Binary	406	964	EASE2	36 km	-9999	real*4

The binary files are written in column major order. For georeferencing, the above binary files use the EASE2 grid lat-lon files, available via the National Snow and Ice Data Center (NSIDC), Boulder, CO.

Contact email: Narendra.N.Das@jpl.nasa.gov