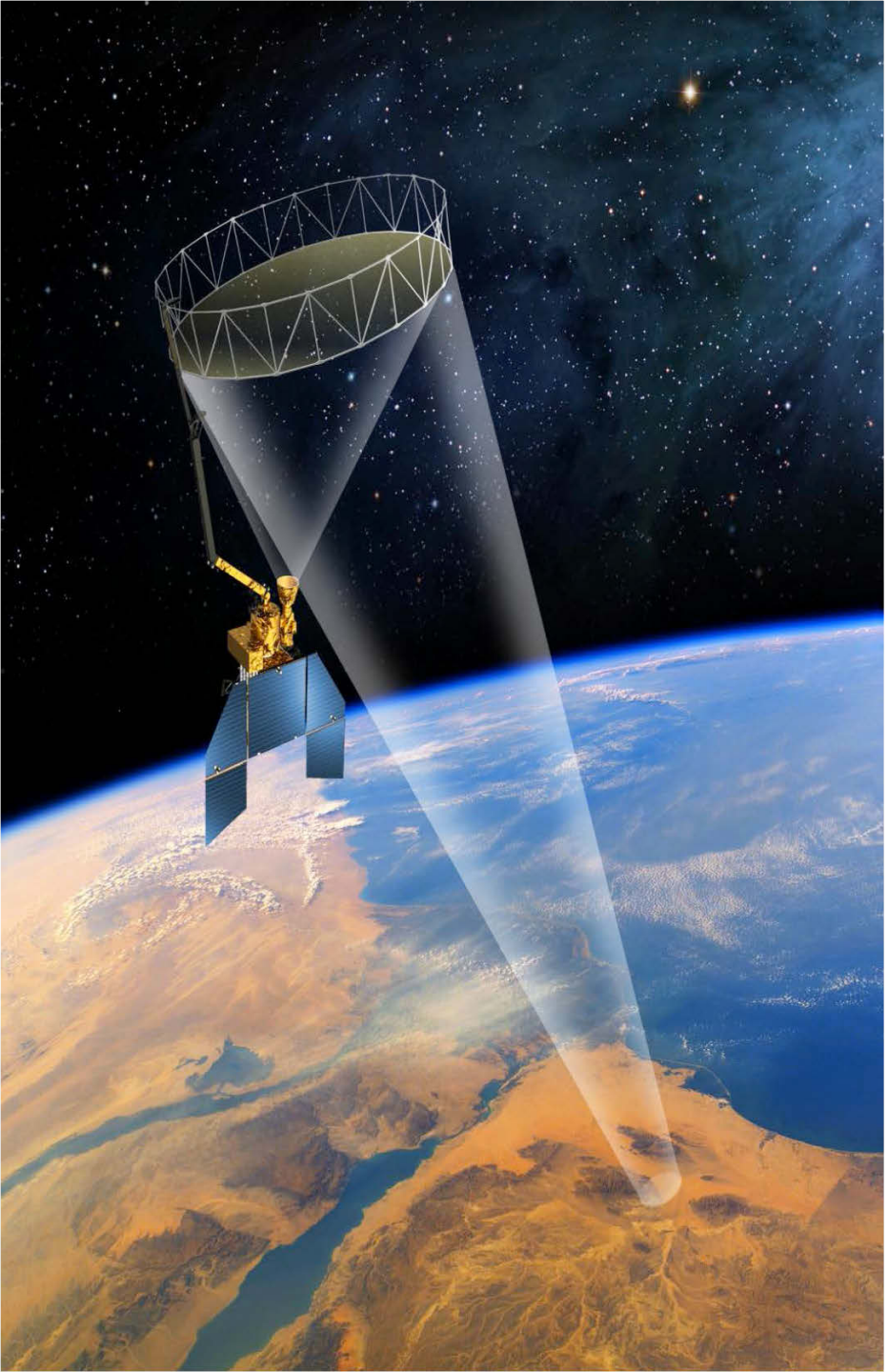


SMAP Data Products

USGS SMAP Joint Mission Tutorial

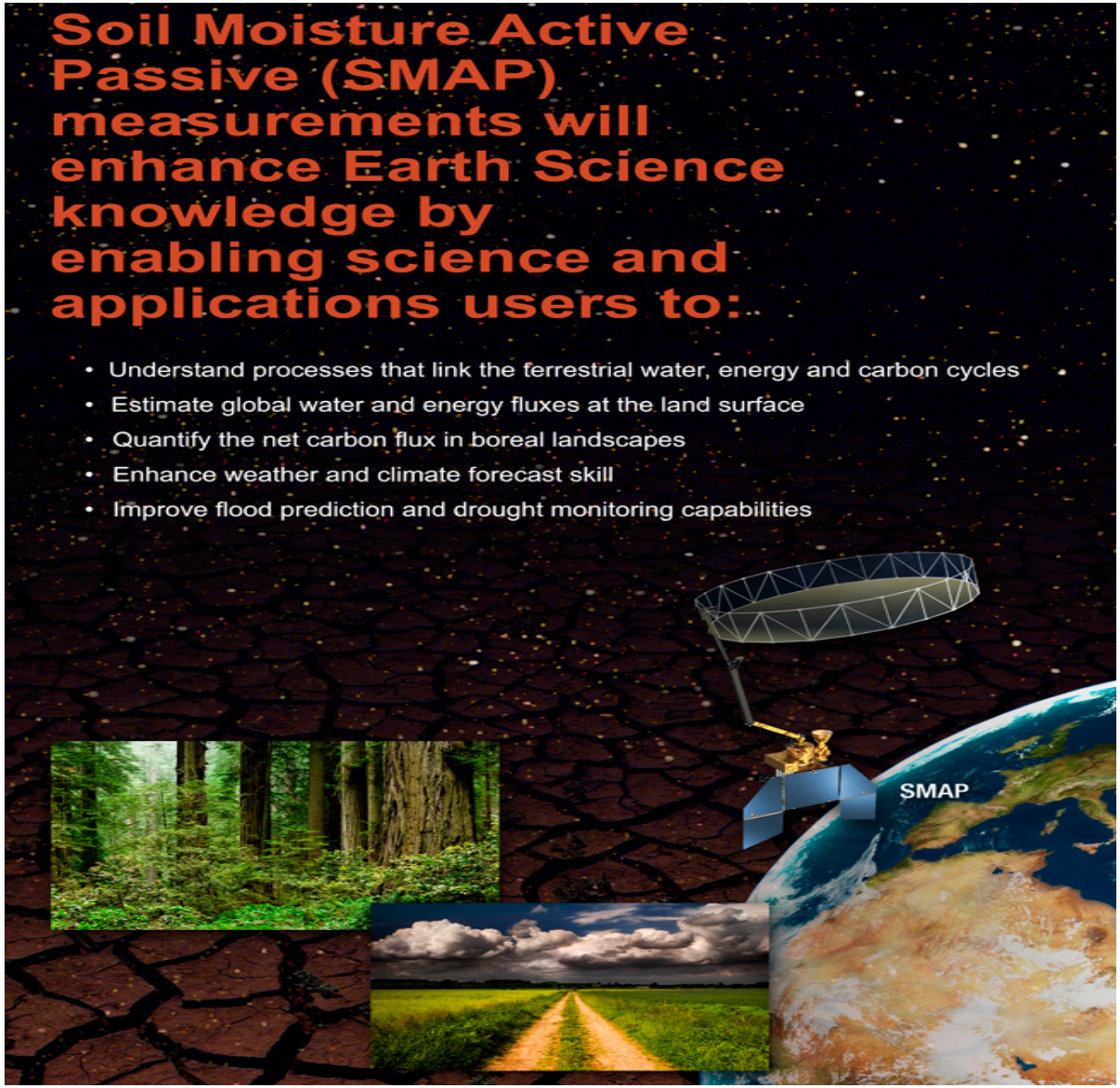
Barry Weiss
Scott Dunbar

SMAP Science Algorithm Development
SMAP Science Data System (SDS)
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA

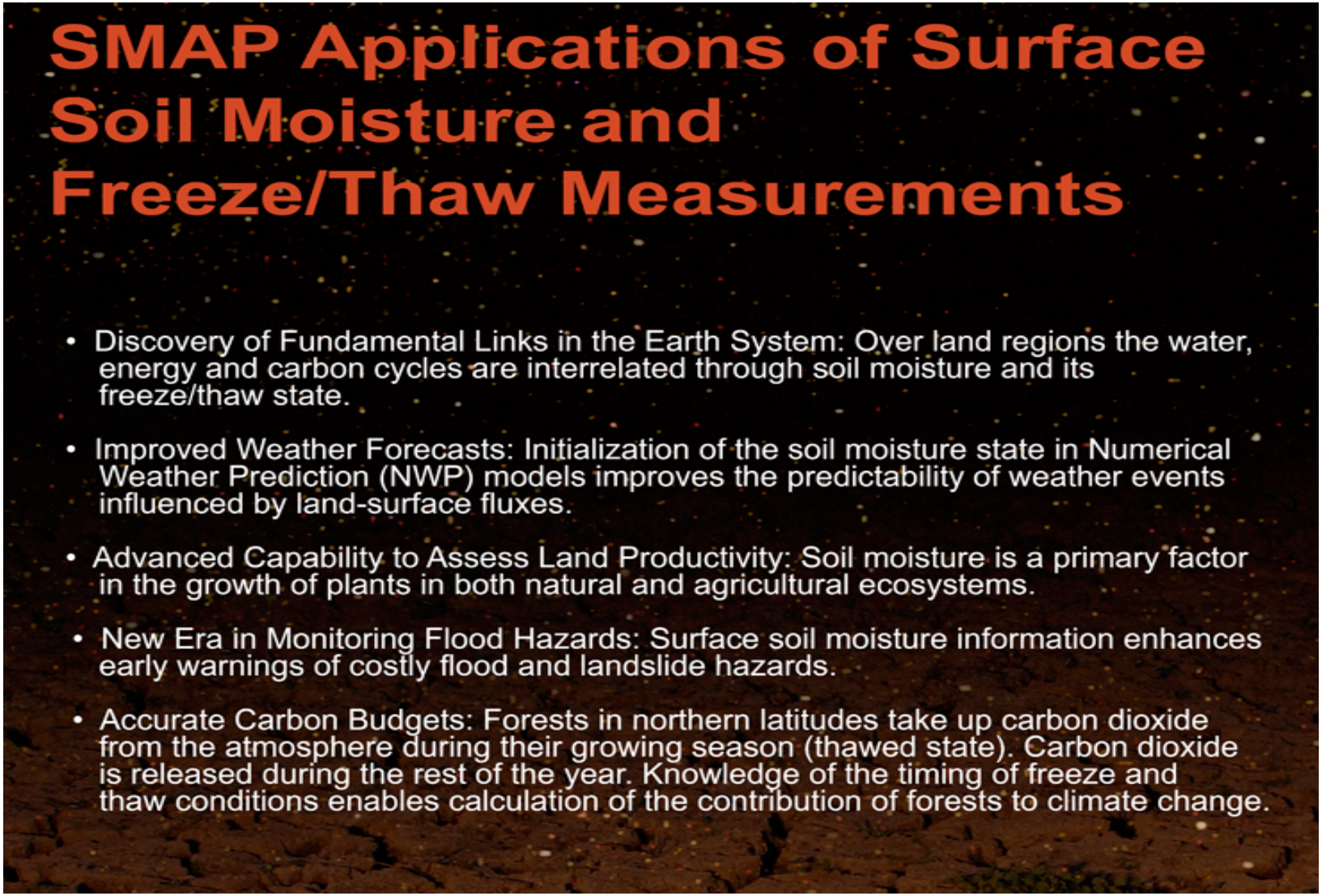
Presentation Outline

- Overview of proposed SMAP Mission
- Data product content
- Data Product format and design
- Provision of simulated data products and other data sets for application development

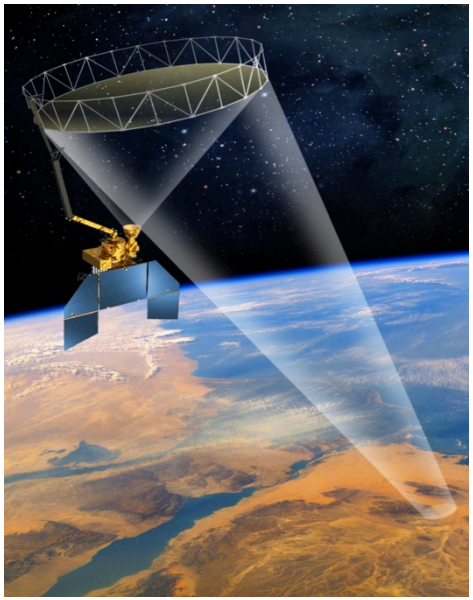
SMAP Science Objectives



SMAP Applications



Mission Overview



The proposed SMAP mission was in the first tier recommended by 2007 NRC Earth Science Decadal Survey

Primary Science Objectives :

Global, high-resolution mapping of soil moisture and its freeze/thaw state to:

- Link terrestrial water, energy and carbon cycle processes
- Estimate global water and energy fluxes at the land surface
- Quantify net carbon flux in boreal landscapes
- Extend weather and climate forecast skill
- Develop improved flood and drought prediction capability

Observatory Features:

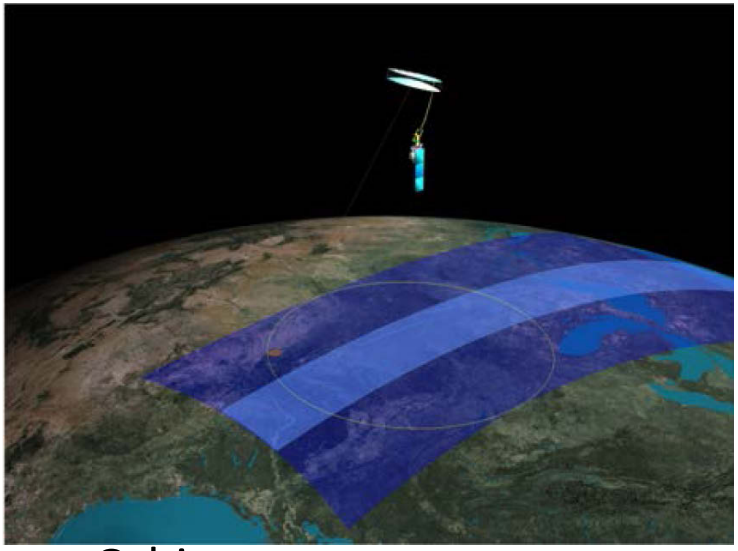
- 3-axis stabilized spacecraft with zero momentum biased attitude control
- Single string avionics and power control/distribution electronics
- Selected redundancy in ACS sensors, actuators, and telecom radios
- Deployable fixed solar array
- Command Telemetry & Doppler via S-band to NEN & SN
- Science data return at 130 Mbps via an X-band link to the NEN
- Hydrazine blow-down propulsion

Proposed Mission Implementation:

Partners	• JPL (project & payload mgmt, science, spacecraft, radar, mission operations, science processing) • GSFC (science, radiometer, science processing)
Risk	• 7120.5D Category 2; 8705.4 Payload Risk Class "C"
Launch	• Oct. 2014, the baseline plan launch vehicle is Delta-2
Orbit	• Polar sun synchronous; 685 km altitude
Duration	• 3 years
Payload	• L-band SAR (JPL) • L-band radiometer (GSFC) • Shared 6m rotating (13 rpm) antenna (JPL)

Proposed SMAP Measurement Approach

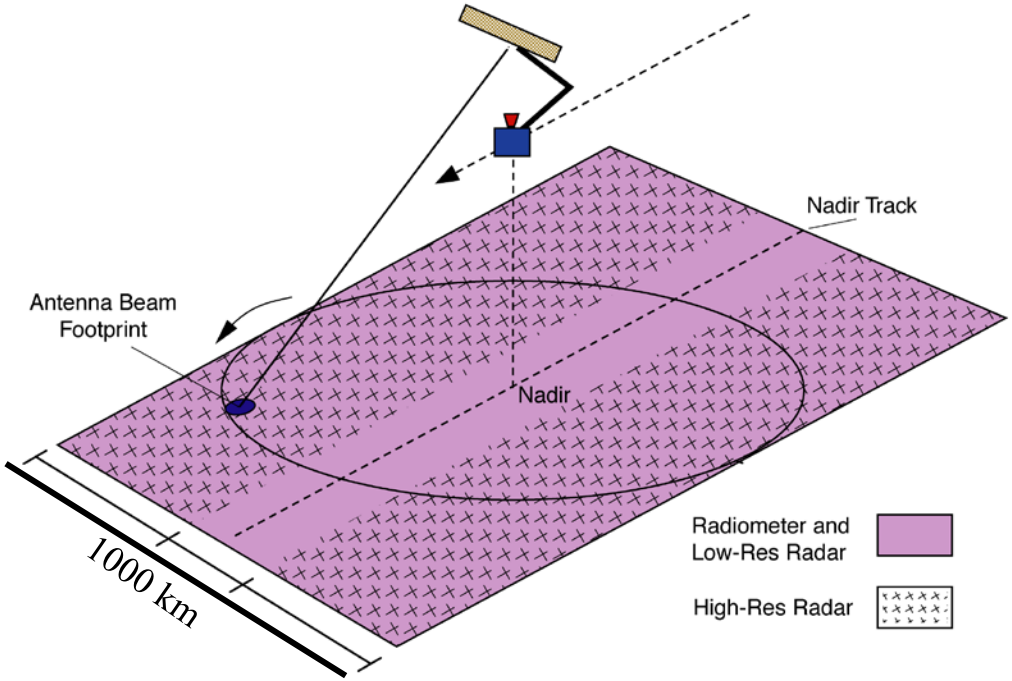
- Instruments:
  - Radar: L-band (1.26 GHz)
    - High resolution, moderate accuracy soil moisture
    - Freeze/thaw state detection
    - SAR mode: 3 km resolution
    - Real-aperture mode: 30 x 6 km resolution
  - Radiometer: L-band (1.4 GHz)
    - Moderate resolution, high accuracy soil moisture
    - 40 km resolution
  - Shared Antenna
    - 6-m diameter deployable mesh antenna
    - Conical scan at 13 rpm
    - Constant incidence angle: 40 degrees
      - 1000 km-wide swath



- Orbit:
  - Sun-synchronous orbit
  - 6 am local time descending
  - 6 pm local time ascending
  - 685 km altitude
  - Global coverage once every three days
- Mission Operations:
  - 3-year baseline mission

Instrument Key Features

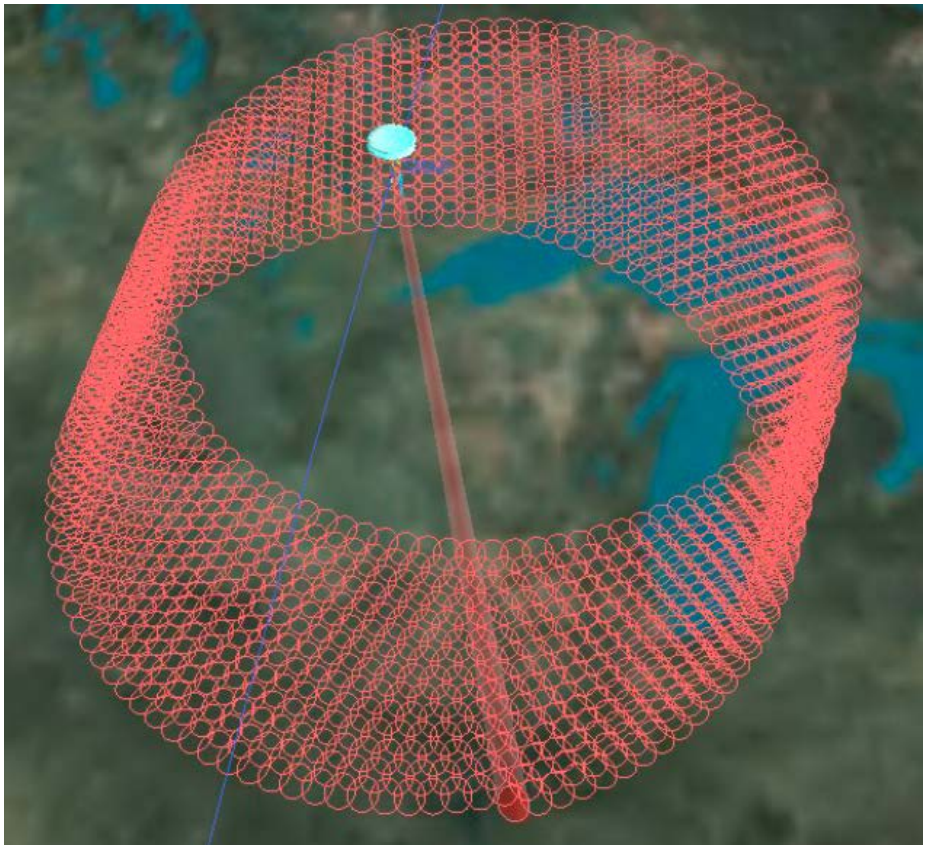
- To meet requirement for 3-day revisit time at AM local time…
  - ⇒ *1000 km swath at 670 - 685 km dawn/dusk sun-synchronous orbit.*
- For wide measurement swath of combined L-Band active and passive measurements…
  - ⇒ *Conically scanning reflector antenna.*



- To achieve L-Band passive resolution of 40 km and active resolution of 3 km …
  - ⇒ *6 meter aperture antenna*
  - ⇒ *13 rpm rotation rate*
  - ⇒ *Real-aperture radiometer*
  - ⇒ *Synthetic-aperture radar processing*
- Incidence angle
  - ⇒ *Near-constant 40 deg incidence angle*

Data Acquisition Plan

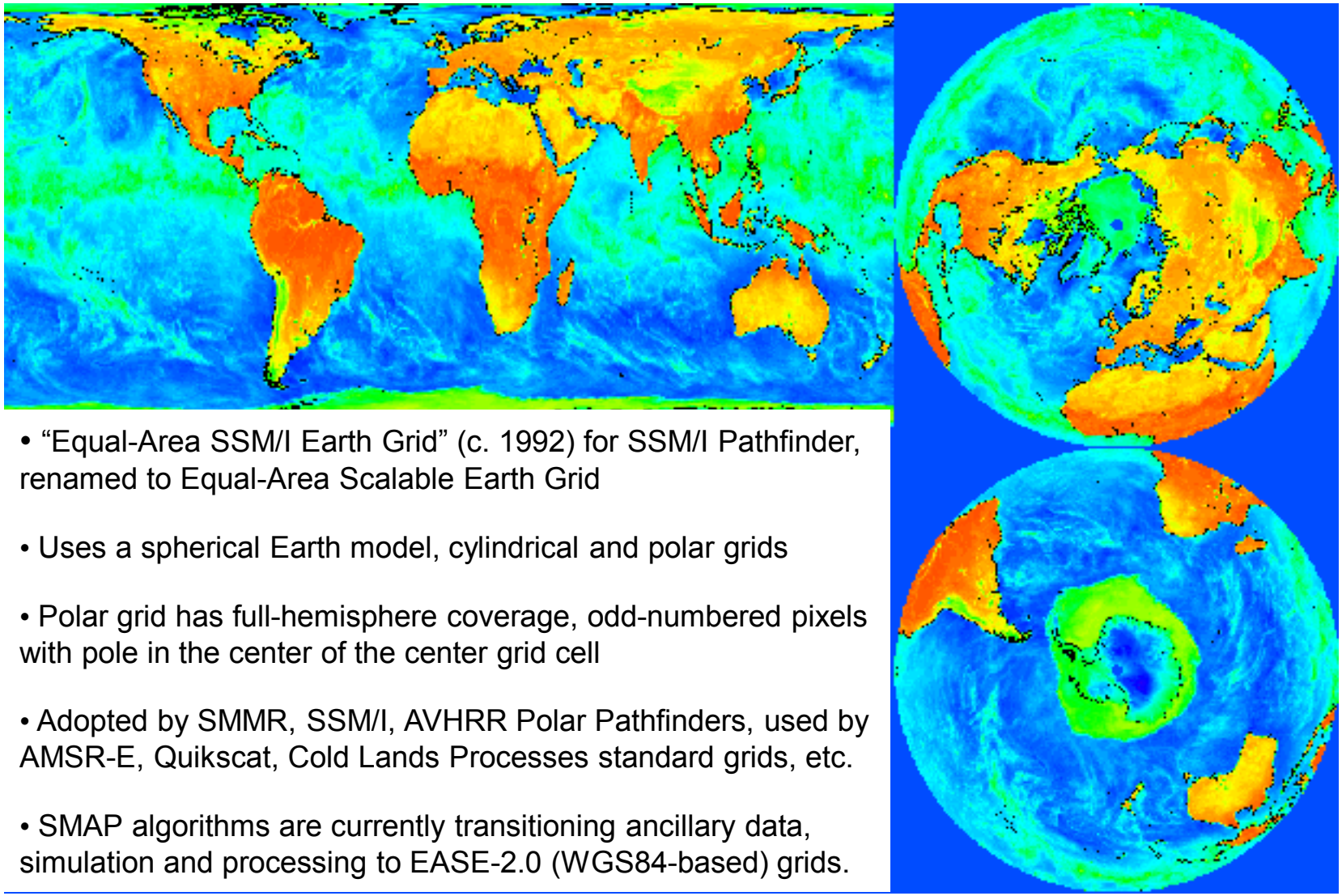
- Radiometer data:
  - Continuous collection over the entire orbit and entire 360 degree antenna scan
  - Capable of periodic “cold sky” looks
- Synthetic Aperture Radar (SAR) data:
  - Collection over the entire 360 degree antenna scan over land and coastal water during the AM orbit
  - Collection over the forward portion of the antenna scan over land and coastal water North of 45 degrees North latitude during the PM orbit.
- Low-resolution, real aperture radar data
  - Continuous collection over entire orbit and entire antenna scan



Proposed SMAP Data Products

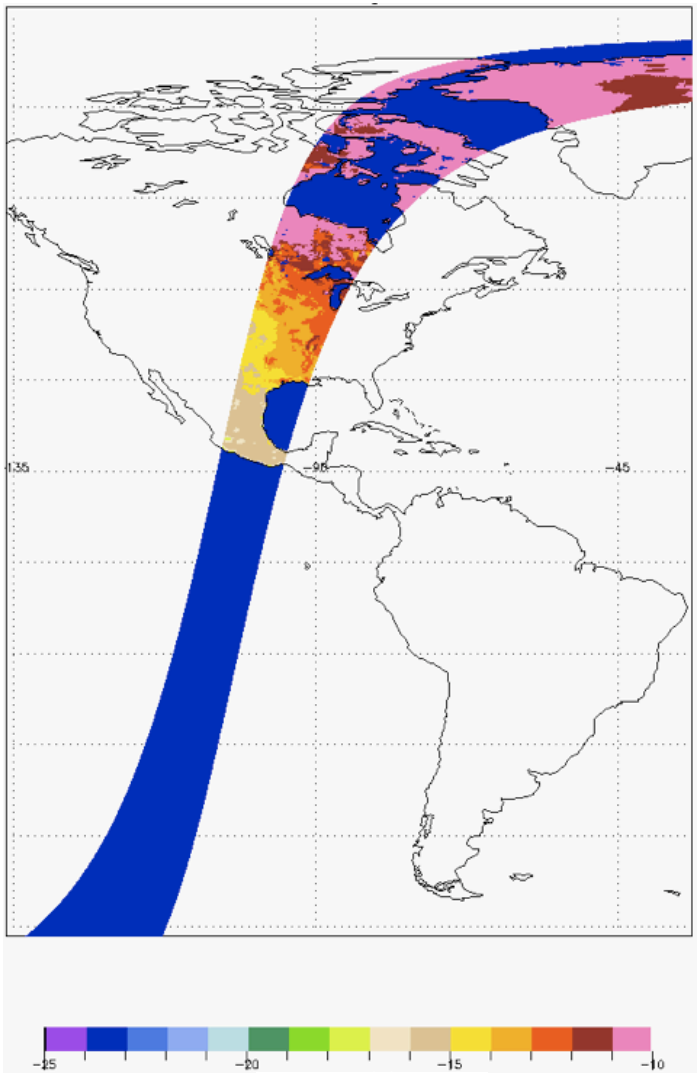
Data Product Short Name	Description	Grid Resolution	Granule Extent
L1A_Radar	Parsed Radar Instrument Telemetry		Half Orbit
L1A_Radiometer	Parsed Radiometer Instrument Telemetry		Half Orbit
L1B_S0_LoRes	Low Resolution Radar σ_o in Time Order	5x30 km (10 slices)	Half Orbit
L1C_S0_HiRes	High Resolution Radar σ_o on Swath Grid	1 km	Half Orbit
L1B_TB	Radiometer T_B in Time Order	39x47 km	Half Orbit
L1C_TB	Radiometer T_B	36 km	Half Orbit
L2_SM_A	Radar Soil Moisture (includes Freeze-Thaw)	3 km	Half Orbit
L2_SM_P	Radiometer Soil Moisture	36 km	Half Orbit
L2_SM_AP	Active-Passive Soil Moisture	9 km	Half Orbit
L3_FT_A	Daily Global Composite Freeze/Thaw State	3 km	North of 45° N
L3_SM_A	Daily Global Composite Radar Soil Moisture	3 km	Global
L3_SM_P	Daily Global Composite Radiometer Soil Moisture	36 km	Global
L3_SM_AP	Daily Global Composite Active-Passive Soil Moisture	9 km	Global
L4_SM	Surface & Root Zone Soil Moisture	9 km	Global
L4_C	Carbon Net Ecosystem Exchange	9 km	North of 45° N

SMAP Level 2, 3 and 4 Products Use the EASE Grid



- "Equal-Area SSM/I Earth Grid" (c. 1992) for SSM/I Pathfinder, renamed to Equal-Area Scalable Earth Grid
- Uses a spherical Earth model, cylindrical and polar grids
- Polar grid has full-hemisphere coverage, odd-numbered pixels with pole in the center of the center grid cell
- Adopted by SMMR, SSM/I, AVHRR Polar Pathfinders, used by AMSR-E, Quikscat, Cold Lands Processes standard grids, etc.
- SMAP algorithms are currently transitioning ancillary data, simulation and processing to EASE-2.0 (WGS84-based) grids.

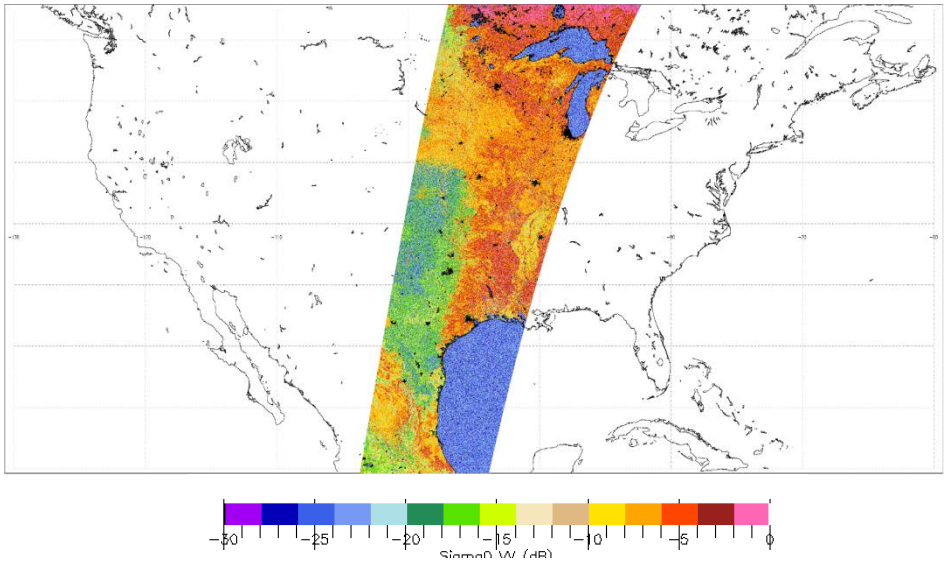
Radar Level 1B Product



Sigma0 VV (dB)

- Each granule contains time ordered data that covers one spacecraft half orbit.
- Coverage is continuous over all surface types.
- Contains Earth-located, calibrated radar backscatter measurements for co-pol and cross-pol data.
- Estimated Kp errors assigned to each measurement.
- Includes spacecraft orbit and attitude information and instrument pointing geometry.
- Includes short term and external calibration data used to generate product output.
- Provides calibrated backscatter measurements for approximately ten range-resolved "slices" of the full radar FOV footprint. (~30 km by 5 km.)

Radar Level 1C Product

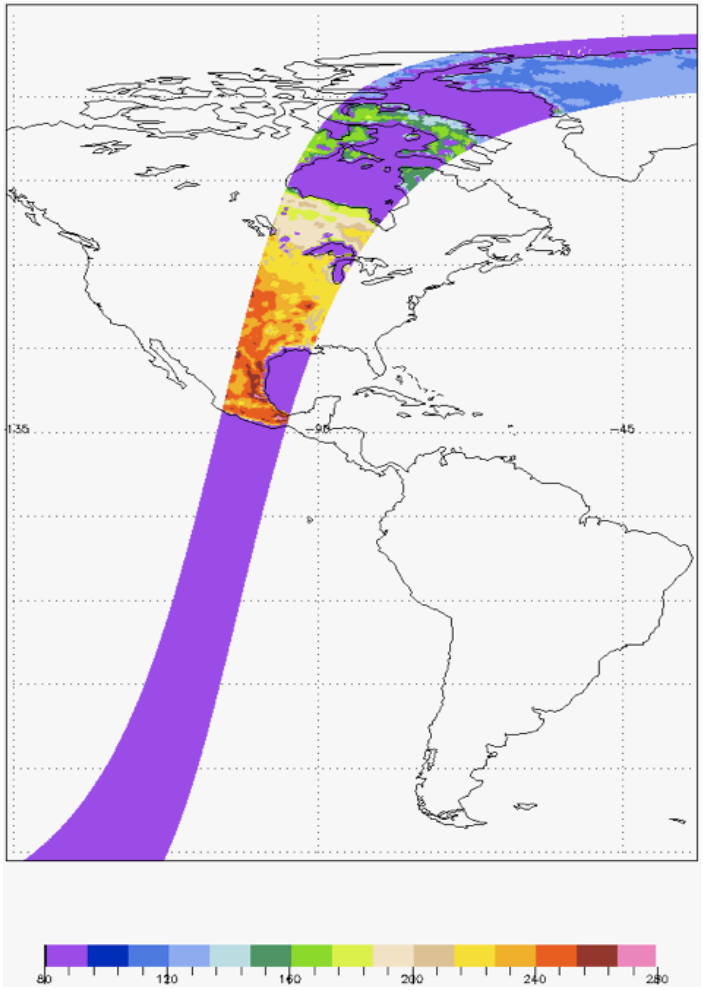


Sigma0 VV (dB)

- Each granule contains geographically ordered data in 1 km grid cells in an along track/cross track swath grid.
- Coverage is restricted to land and coastal water over one spacecraft half orbit.
- SAR provides high-resolution single-look measurements. Resolution varies from ~400 m at the swath edge to about 1.2 km at 150 km from the nadir sub-track. Nadir looks are thin slices as wide as the beam footprint.

- Contains Earth located and calibrated h-pol, v-pol and cross-pol backscatter measurements, each separately multilooked
- Radar measurements achieve 1 km resolution over 70% of the swath. Resolution degrades in the nadir region.
- Forward looking and aft looking measurements stored separately.
- Includes spacecraft orbit and attitude information and instrument pointing geometry.
- Includes short term and external calibration data used to generate product output.
- Provides reference to global and polar 1 km EASE grid coordinates.

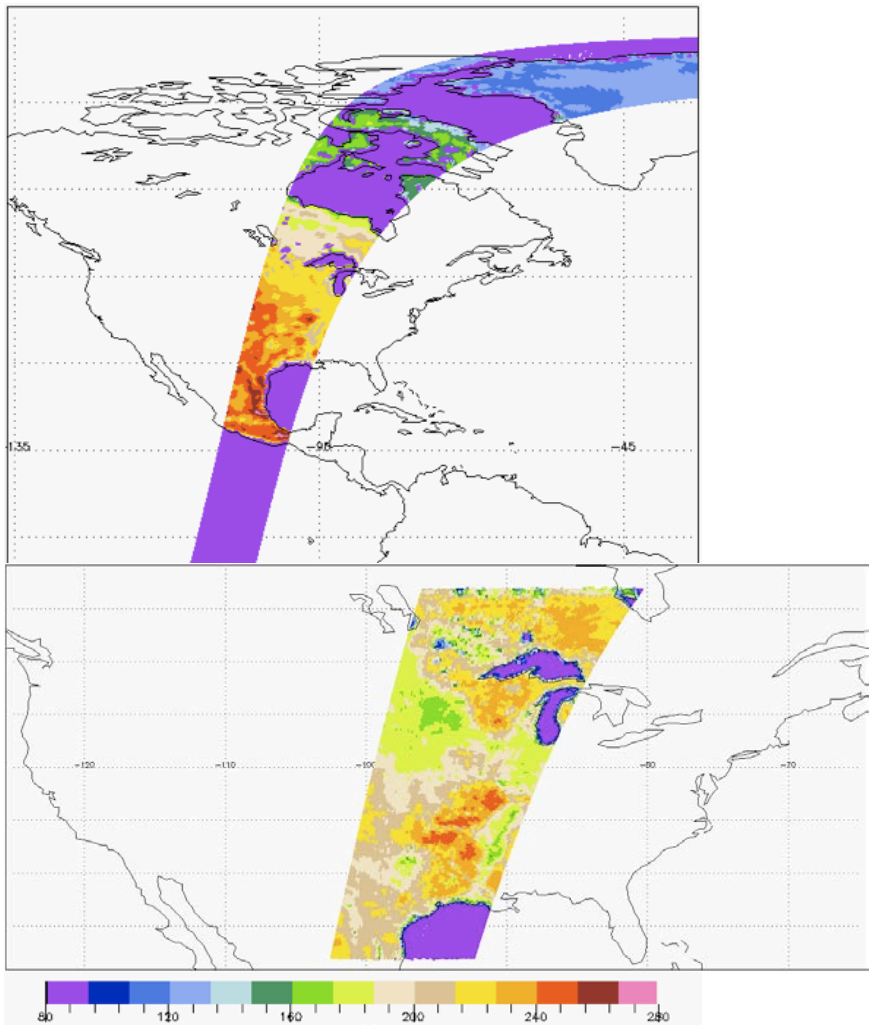
Level 1B Radiometer Product



L1B Time-ordered H-pol TB (K)

- Each granule contains time ordered data that covers one spacecraft half orbit.
- Effective field of view footprint is a 39 km by 47 km ellipse
- Earth-located calibrated data for each EFOV
  - Apparent aperture (antenna) temperatures
  - Top-of-ionosphere (TOI) brightness temperature
  - Surface-referenced brightness temperatures
- Coverage continuous over all surface types.
- All four modified Stokes parameters (V, H, 3 & 4).
- 3rd Stokes used for Faraday rotation correction.
- Time-frequency-polarization diversity used for RFI detection and removal.
- Forward looking and aft looking measurements stored separately.
- Includes spacecraft orbit and attitude information and instrument pointing geometry.

Level 1C Radiometer Product

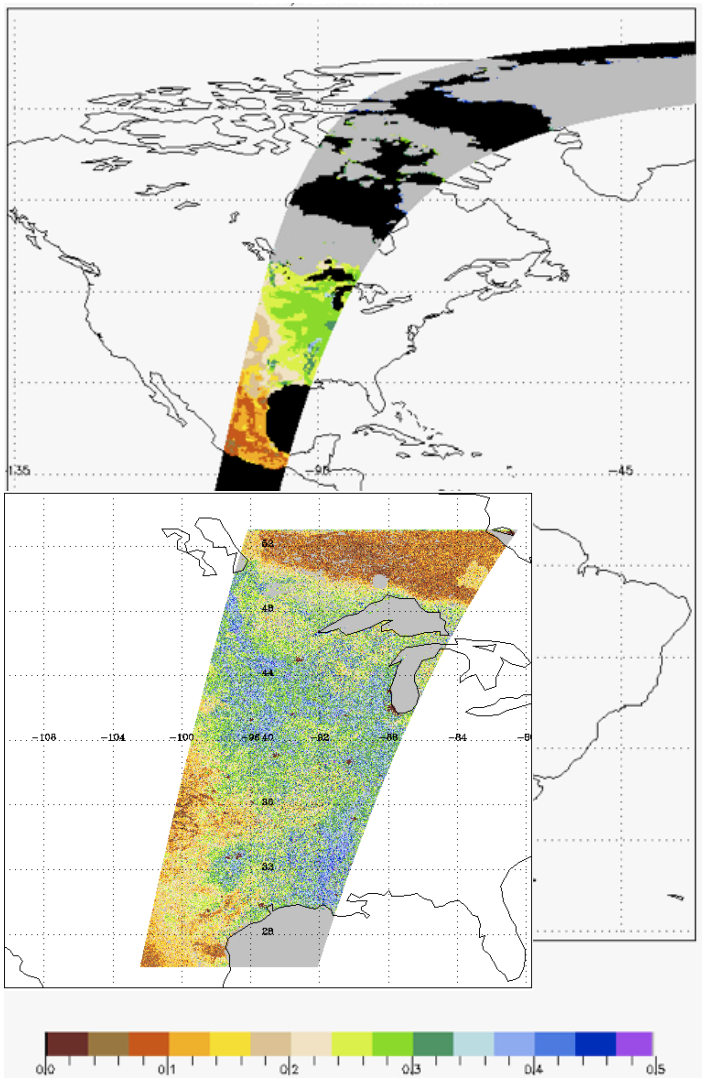


L1C Earth-fixed H-pol TB (K)

- Each granule represents one spacecraft half orbit.
- Contains Level 1B Radiometer data represented on a cylindrical 36 km EASE grid and two polar 36 km EASE grids.
- Data are represented in a one dimensional array.
- Product lists only those EASE grid cells that contain data.
- Latitude and longitude listed for each EASE grid cell.
- Forward looking and aft looking observations stored separately.
- Input to Level 2 Radiometer 40 km soil moisture and Level 2 active/passive soil moisture processing.

National Aeronautics and Space Administration

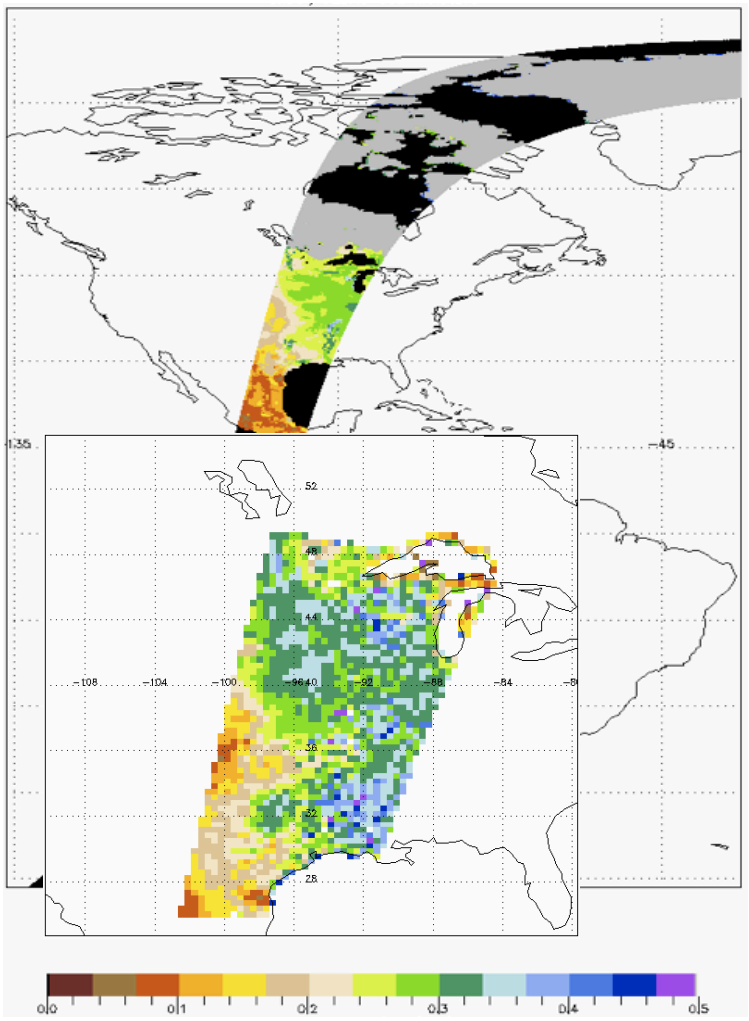
Level 2 High Resolution Radar 3 km Soil Moisture Product



Volumetric Soil Moisture (cm^3/cm^3)

- Each granule contains one half orbit of data posted on 36 km cylindrical EASE grid cells.
- Data are represented in a one dimensional array.
- Product lists only those EASE grid cells that contain data.
- AM Product covers entire Earth land mass, PM product restricted to land north of 45 North longitude
- PM data acquired specifically for freeze-thaw retrievals.
- Employs 1 km high resolution radar L1C data averaged over 3 km cylindrical EASE grid cells to reduce Kp noise.
- Soil moisture retrievals use snapshot and/or time-series algorithms.
- Depending on the terrain classification, multiple optional models/algorithms may be employed for retrieval.
- Provides freeze-thaw state and transient water body information that the other Level 2 soil moisture processes require.
- Includes quality masks for urban areas, mountainous terrain, dense vegetation, snow and ice.

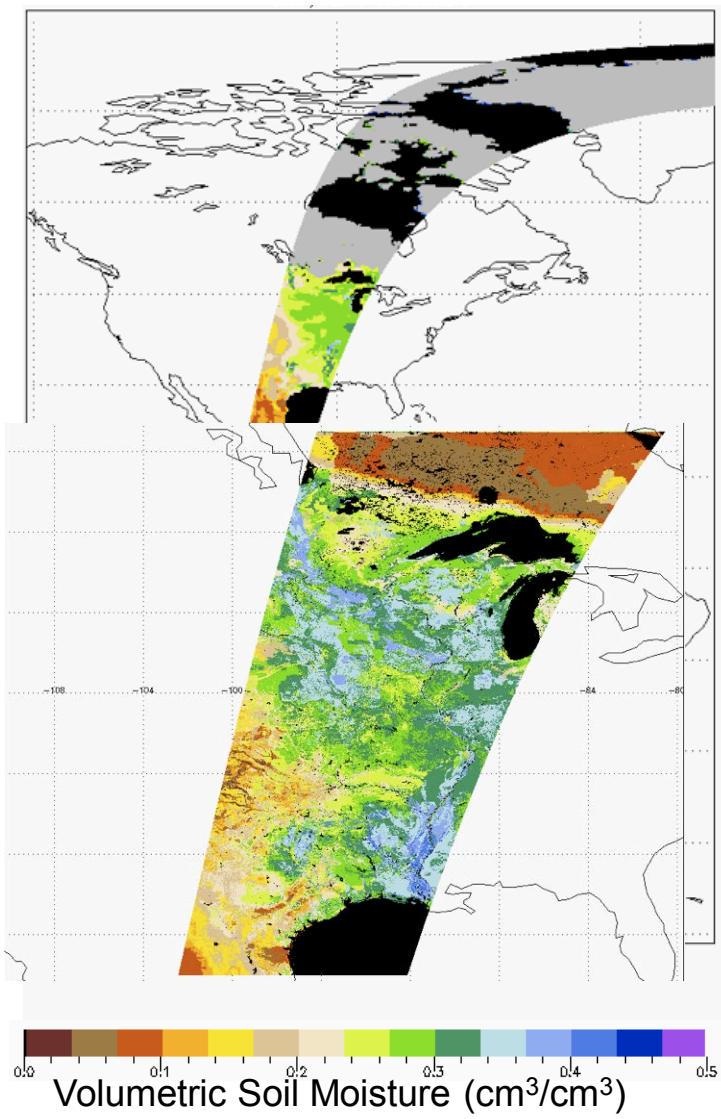
Level 2 Radiometer
36 km Soil Moisture Product



Volumetric Soil Moisture (cm^3/cm^3)

- Each granule contains one half orbit of data posted on 36 km cylindrical EASE grid cells.
- Data are represented in a one dimensional array.
- Product lists only those EASE grid cells that contain data.
- Provides retrieved soil moisture over land with 4% accuracy for low-to-moderately vegetated areas.
  - Low to moderate vegetation defined as vegetation water content <= 5 kg/m^2.
- Requires transient water body and freeze-thaw state retrievals generated with high resolution radar retrievals.
- Estimates soil moisture based on AM observations.
- Includes quality masks for urban areas, mountainous terrain, dense vegetation, precipitation, snow and ice.

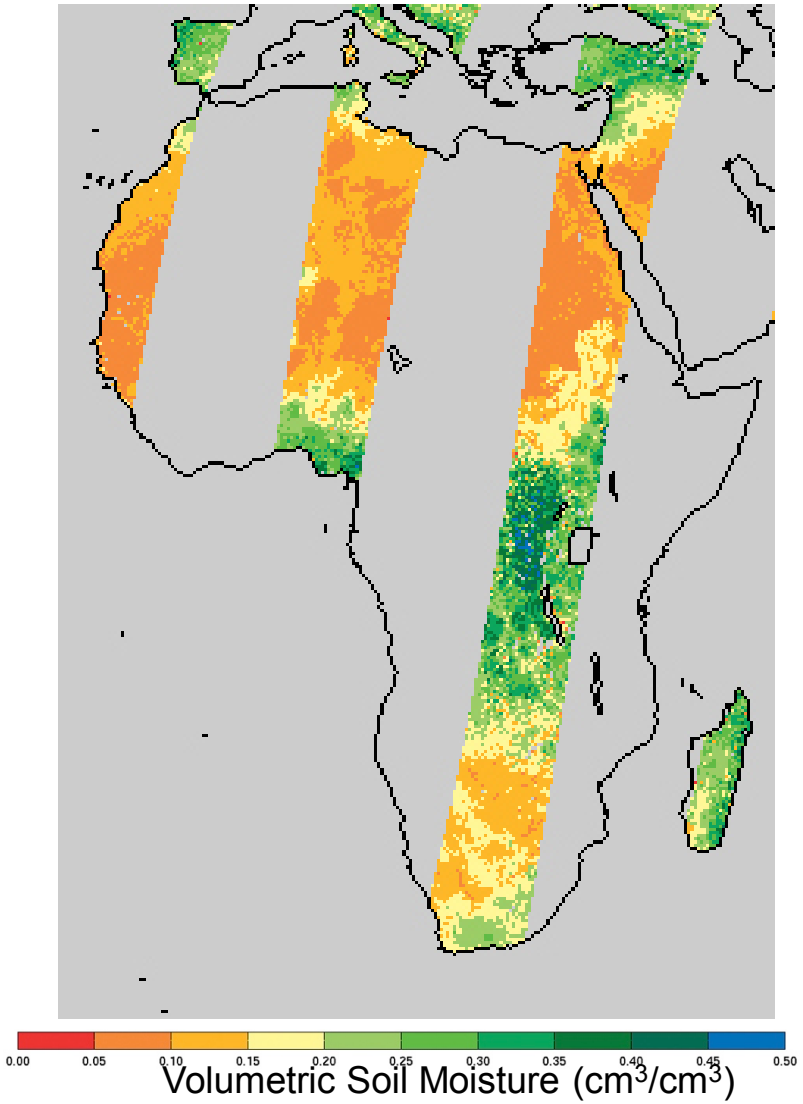
Level 2 Active/Passive 9 km Soil Moisture Product



- Each granule contains one half orbit of data posted on 9 km cylindrical EASE grid cells.
- Data are represented in a one dimensional array.
- Product lists only those EASE grid cells that contain data.
- Merges radar and radiometer channels using a time series algorithm and spatial heterogeneity of L1C radar product.
- Provides dis-aggregated brightness temperatures at 9 km resolution.
- Provides retrieved soil moisture over land with 4% accuracy for low-to-moderately vegetated areas.
  - Low to moderate vegetation defined as vegetation water content <= 5 kg/m^2.
- Employs transient water body and freeze-thaw state generated with high resolution radar retrievals.
- Include quality masks for urban areas, mountainous terrain, dense vegetation, precipitation, snow and ice.

National Aeronautics and Space Administration

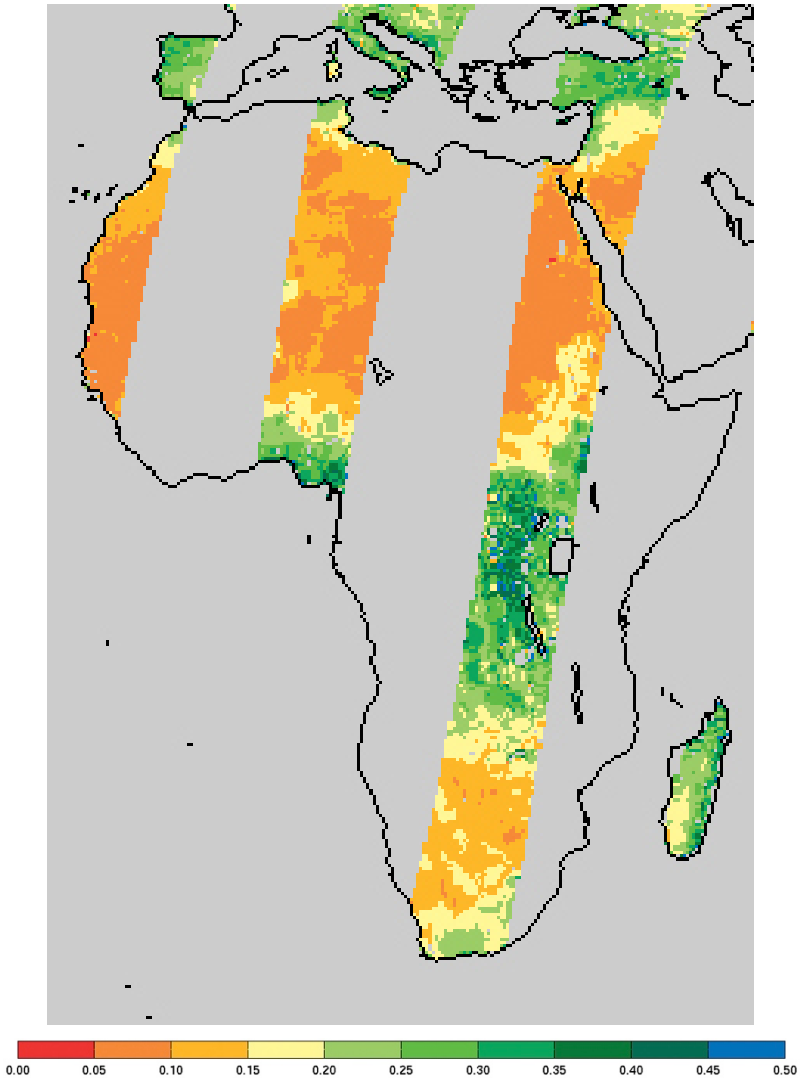
Level 3 High Resolution Radar 3 km Soil Moisture Product



- Composite of all Radar Level 2 half orbit products where the local acquisition time is the same UTC day.
- Multiple measurements may overlap at high latitudes. Algorithm selects those measurements acquired closest to 6 AM solar time.
- Posted on a 3 km cylindrical EASE grid using a two dimensional array.
- Product lists all EASE grid cells, regardless of whether data are available.
- Soil moisture retrievals use snapshot and/or time-series algorithms.
- Depending on the terrain classification, multiple optional models/algorithms may be employed for retrieval.
- Based exclusively on AM data.

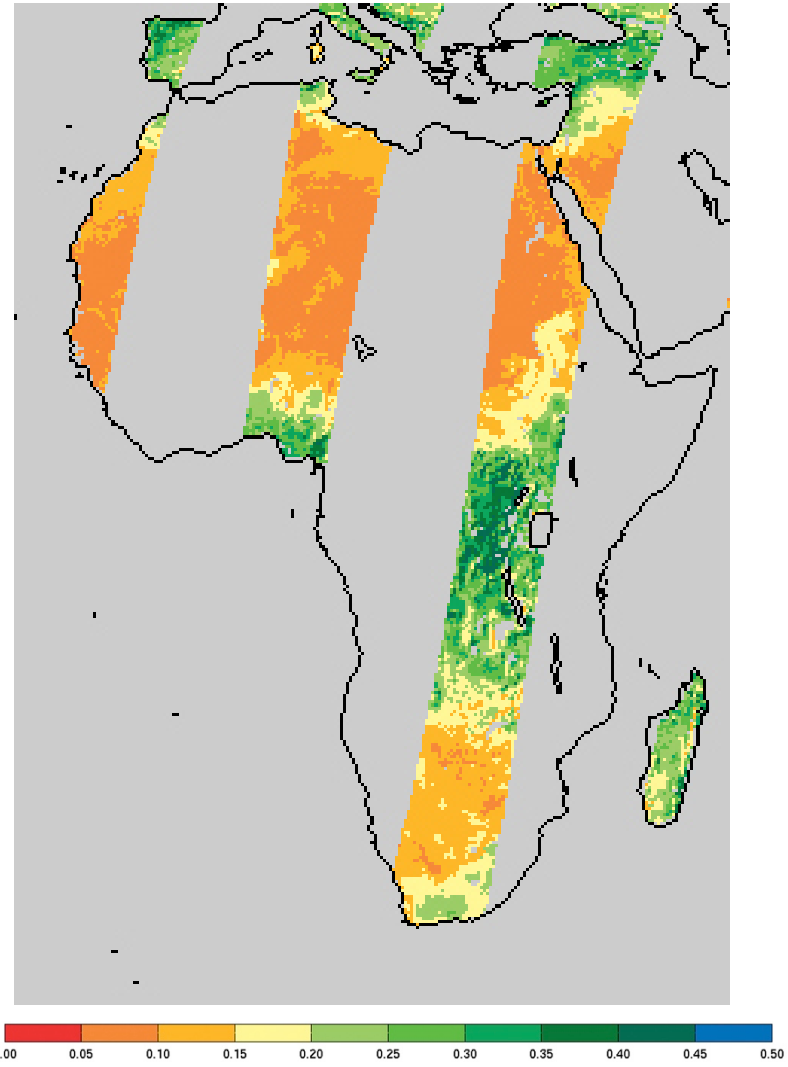
National Aeronautics and Space Administration

Level 3 Radiometer 36 km Soil Moisture Product



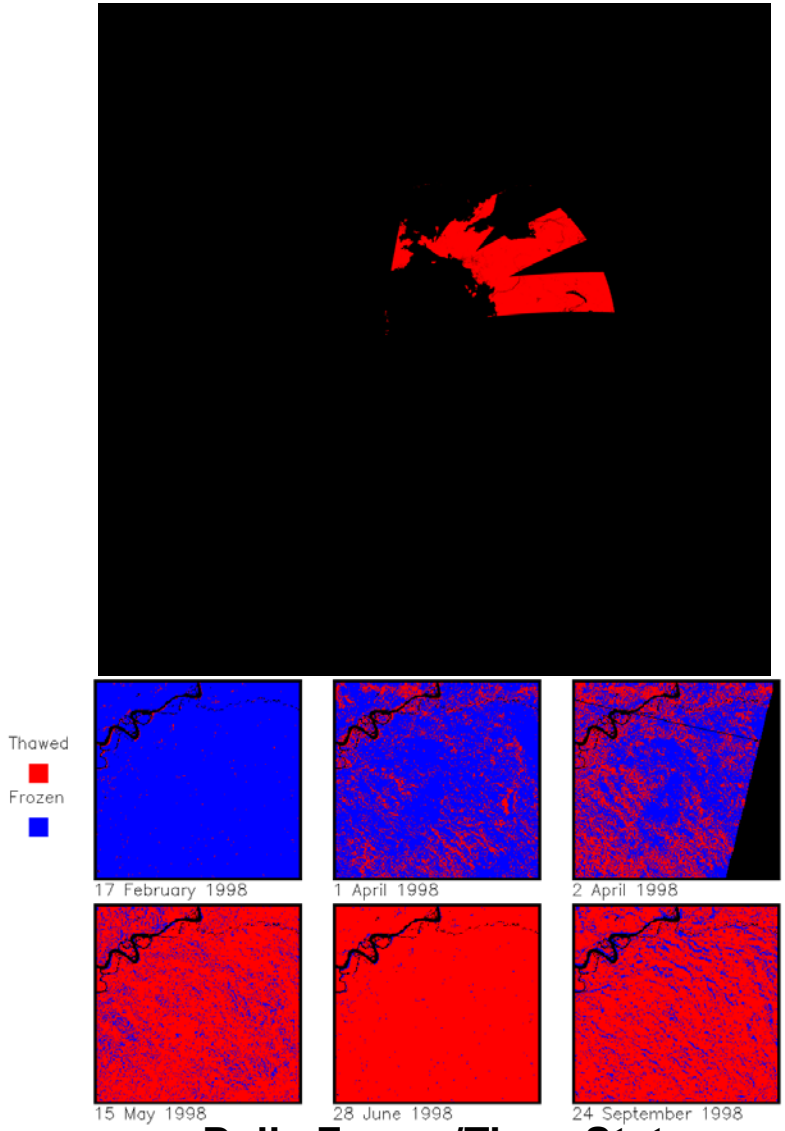
- Composite of all Radiometer Level 2 half orbit products where local acquisition time is the same UTC day.
- Multiple measurements may overlap at high latitudes. Algorithm selects measurements acquired closest to 6 AM solar time.
- Posted on a 36 km cylindrical EASE grid using a two dimensional array.
- Product lists all EASE grid cells, regardless of whether data are available.
- Provides retrieved soil moisture over land with 4% accuracy for low-to-moderately vegetated areas.
  - Low to moderate vegetation defined as vegetation water content <= 5 kg/m^2.
- Based exclusively on AM data.

Level 3 Active/Passive 9 km Soil Moisture Product



- Composite of all Active/Passive Level 2 half orbit products where local acquisition time is the same UTC day.
- Multiple measurements may overlap at high latitudes. Algorithm selects measurements acquired closest to 6 AM solar time.
- Posted on a 9 km cylindrical EASE grid using a two dimensional array.
- Product lists all EASE grid cells, regardless of whether data are available
- Provides retrieved soil moisture over land with 4% accuracy for low-to-moderately vegetated areas.
  - Low to moderate vegetation defined as vegetation water content <= 5 kg/m^2.
- Based exclusively on AM data.

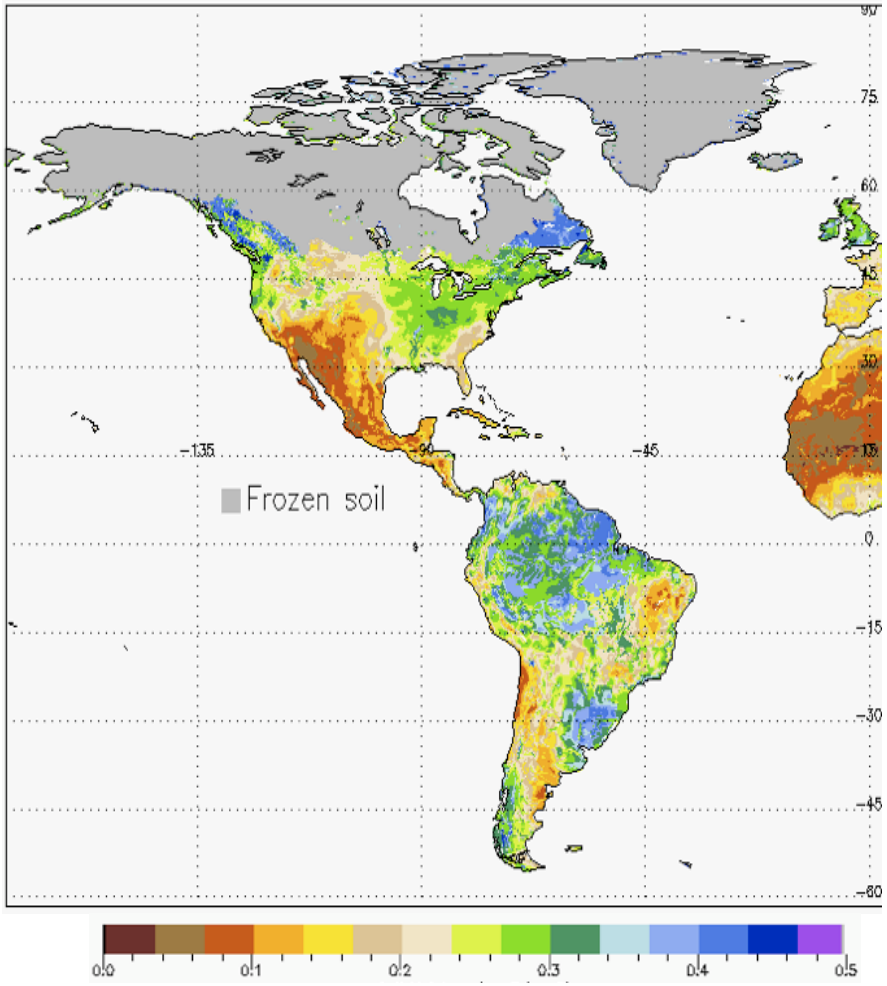
Level 3 Freeze/Thaw Product



Daily Freeze/Thaw State

- Employs the 1 km Level 1C high resolution radar data and a time-series change detection algorithm to infer freeze/thaw state.
- Quantifies daily freeze/thaw state as a binary condition for land surface.
- Includes both AM and PM data, with intra-day state transition flags.
- Posted on a 3 km polar EASE grid with 3 km spatial resolution using a two dimensional array.
- Each product represents a single calendar day UTC.
- Required to achieve 80% freeze/thaw state classification accuracy.

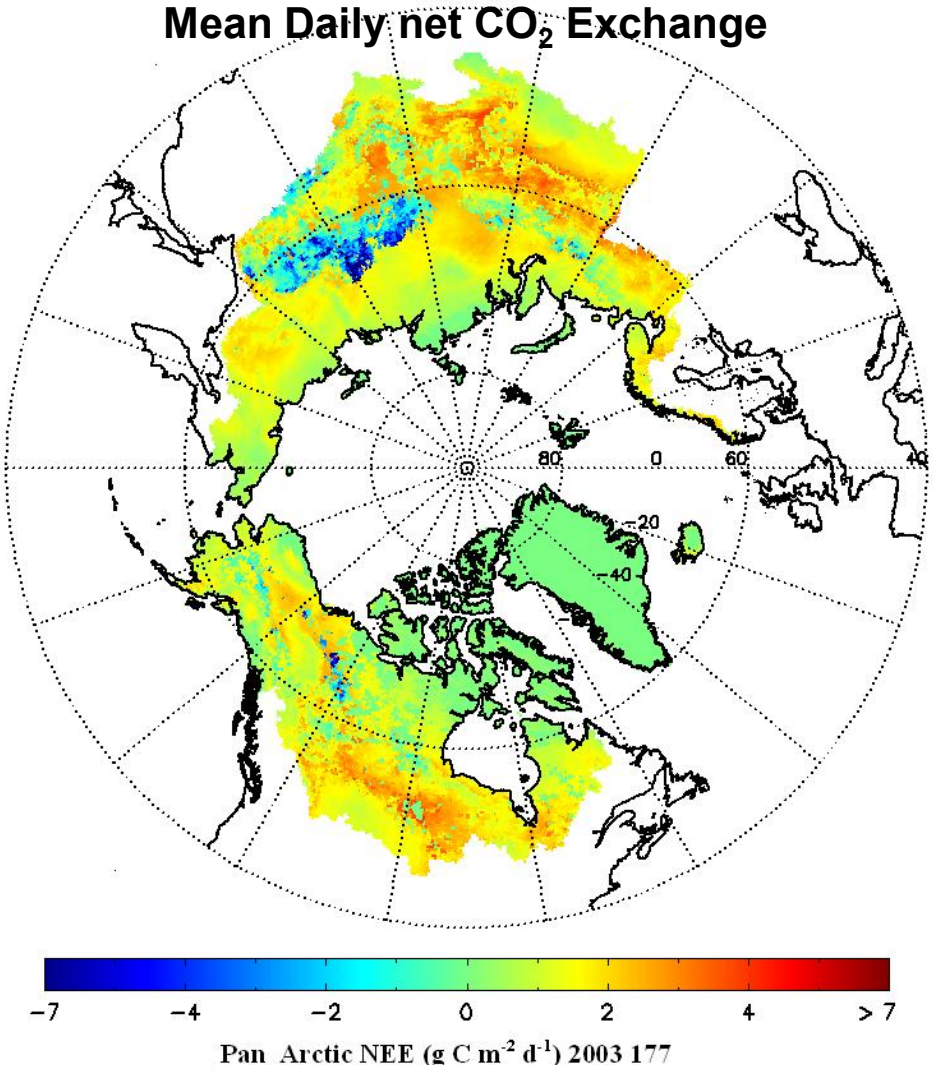
Level 4 Surface and Root-Zone Soil Moisture Product



Volumetric Soil Moisture (%)

- Global output represents 3 hour intervals at 9 km resolution with 7-day latency.
- Product subdivides
  - Instantaneous measures once every 3 hours
  - Time averaged values over a three hour time span.
- Employs SMAP L1C Radiometer, Level 2 AP Disaggregated Brightness Temperatures as well as Level 3 Freeze/Thaw products.
- Assimilates SMAP data into a state-of-the-art land surface model to derive global estimates of root-zone moisture.

Level 4 Carbon Product



- Daily global maps of net ecosystem CO_2 exchange (NEE) at 9 km resolution with 14-day latency.
- Quantifies the net carbon flux in boreal landscapes.
- Reduces uncertainty with regard to existing carbon sinks on land.
- Applies a soil decomposition algorithm driven by SMAP L4_SM and Gross Primary Production (GPP) inputs to compute net land-atmosphere CO_2 exchange (NEE).
- Accuracy commensurate with tower based CO_2 observations. (RMSE ≤ 30 g C m^{-2} yr^{-1} or 1.6 g C m^{-2} d^{-1}).

Data Product Design Drivers

- Data design must balance requirements of the mission Science Team with the needs of the User Community
- All Data Products must:
  - Conform with mission requirements
  - Have a consistent format and architecture
  - Contain adequate self descriptive information
  - Be easy to use

Product Generation

- Within the project - Algorithm Development Team and Science Data System Team collaborate to generate design
  - Develop a complete product design that would include essential data as well as valuable ancillary data
  - Coordinate product content to ensure that data required in subsequent processes are documented and available
  - Coordinate naming conventions so that elements that represent same measures have the same or very similar names
  - Ensure that the products provide the data that the science community needs
- Within the larger community – SMAP Science Data System (SDS) works directly with NASA's Earth Science Data and Information Systems (ESDIS)
  - Ensure that standards are endorsed by the Standard Process Group
  - Ensure that products contain essential metadata so that requisite data could be located on multiple systems including the Data Centers and ECHO
  - Track upcoming technologies, incorporate features that would enable adoption of those technologies as they become more widespread

Product Format
Hierarchical Data Format (HDF5)

- HDF5 combines a robust data model with a file format and a user library
  - Major object components:
    - Datasets – major product content
    - Dataspace – description of product content shapes
    - Datatypes – flexible means to type data
    - Groups – flexible means to group data
    - Attributes – descriptive elements
  - Supports a large variety of platforms and common software languages
    - IDL and Matlab both have well defined and easy to use HDF5 interfaces
  - Has considerable heritage – introduced in 1999-2000
  - Provides a highly flexible architecture to read, extract and write data
  - Operates efficiently with very large data sets
  - Large numbers of new technology data service providers function well with HDF5
- Support available through the HDF Group.
  - Website at http://www.hdfgroup.org/

SMAP Data Product Design
A Consistent and Simple Standard

- **Data product types or collections**
  - Set of files with conformant output generated using the same Science Processing Software (SPS)
  - All data product types should employ a single standard format
- **Data granule or file**
  - An instance of a data product type
- **Product data**
  - Includes both algorithmic output as well as valuable ancillary data
  - Separated into distinct groups of arrays
    - Arrays within each group are conceptually related
    - Within each group, arrays have the same or very similar dimension shapes
    - Within each group, array elements with the same dimension indices apply to the same measurement or retrieval instance
- **Metadata**
  - Product metadata – applies to the entire content of a data granule
  - Local metadata – applies to a particular element in the product

Data Correlation in SMAP Product Design

Three data groups appear in the prototype Radar Level 1B Product design:

Spacecraft Data Group

Each of the arrays in this group contains one representative element for each 0.1 second during the time span the product covers. The following array elements are examples of spacecraft positions and attitudes with a representative time at ***antenna_scan_time_utc[1000]***:

- ***sc_nadir_lon[1000]***
- ***y_pos[1000]***

Sigma0 Data Group

Each of the arrays in this group contains one representative element for each low resolution sigma0. The following array elements are examples of measurements for the low resolution sigma0 with a representative time at ***sigma0_time_utc[20000]*** :

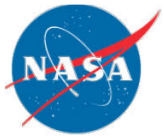
- ***antenna_boresight_azimuth[20000]***
- ***sigma0_hh[20000]***

Sigma0 Slice Data Group

Each of the arrays in this group contains one representative element for a slice in each low resolution sigma0. The following array elements are examples of measurements for the 7th slice of the low resolution sigma0 with a representative time at ***sigma0_time_utc[20000]***:

- ***slice_slant_range[20000][6]***
- ***slice_sigma0_vv[20000][6]***

Metadata Guidelines



- The metadata shall provide users with adequate self descriptive information to enable an assessment of the content, the quality and the algorithmic conditions associated with any SMAP data product.
- The metadata shall enable users to locate specific and appropriate sets of data that they need for their investigation.
- The metadata shall enable users to correlate, interoperate and integrate SMAP data products with those generated by disparate sources, within and outside of NASA.

Metadata Categories and Coverage

- **Product metadata – applies to the entire content of a data granule**
  - Mission specific information
  - Spatial and time boundary information
  - Data version information – algorithm, Science Processing Software (SPS), Science Data System (SDS) release, HDF5 version
  - Granule lineage or pedigree
    - Lists of the input that were used to generate a data granule
  - Technical parameters that apply to the entire data granule
    - Orbit mechanical data
    - Instrument specific information
    - Small tables of calibration and/or algorithmic coefficients
    - Algorithmic parameters and options
  - Data quality and completeness
  - References to related documentation
- **Local metadata – applies to particular arrays in the product.**
  - Maxima, minima, units, dimension definitions, identification of statistical methods

Global Metadata – ISO 19115

- "Geographic Information - Metadata" from the International Organization for Standardization
  - Provides a standardized means to describe Earth data
  - Provides a means to make products "self descriptive and independently understandable"
  - Incorporates all of the major categories required for a complete set of global metadata for each product granule
  - Incorporates all of the major categories required to generate a complete set of collection metadata.
  - Enables fulfillment of the requirement "to correlate, interoperate and integrate SMAP data products with those generated by disparate sources".
  - Uses standardized XML serialization to ease portability to the wider user community. Standard specified in ISO 19139.

SMAP Adaptation of ISO 19115

- Metadata appear in two representations
  - HDF5 group attribute structure within each product
  - ISO 19139 XML, in a separate file as well as in a single metadata attribute
- The ISO 19139 XML rigorously conforms to the ISO model layer
- The HDF group/attribute structure provides a representation layer that is more user friendly
  - All metadata appear under the metadata group
  - Groups within the metadata group represent major ISO classes
  - Attributes in the these groups map to attributes in the ISO classes
- Design may employ modified names of HDF5 groups or attributes to ease user comprehension of the model.
    - For example, the product contains a large number of LI_Lineage/LE_Source classes. HDF5 group names reflect the product described in the group. Group names include:
      - Attitude, Ephemeris, Antenna Azimuth, Spacecraft Clock

ISO Metadata Structure Example

Earth Science Data Model – old

AncillaryInputName = SMAP_L1C_S0_HIRES_SPS_InputConfig_L1A_Radar.xml,

SMAP_L1C_S0_HIRES_SPS_OutputConfig_L1C_S0_HiRes.xml,

SMAP_L1C_S0_HIRES_SPS_MetConfig_L1C_S0_HiRes.xml,

smap_cl_v00001.tsc,

smap_pf_v07.tf,

earth_070425_370426_070425.bpc,

naif0010.tls,

pck00009.tpc,

de421.bsp,

bfpq_mant_decode_array_m_4_v1,

smap_tr_1505291400_1512291400_v01.bsp,

smap_at_1505291459_1605291259_v01.bc,

smap_ar_1505301619_1505301620_v01_t.bc

CollectionDescription = Point Target Simulation Data

CompositeReleaseID = R00301

…….

ISO 19115 Group/Attribute Model – new

L1A_Radar

DOI = http://dx.doi.org/10.5067/smap/radar/data100

creationDate = 2015-05-30

description = Parsed and reformatted SMAP radar telemetry. The Level 1A Product contains both synthetic aperture radar data and real aperture radar data. The product also includes loopback data as well as health and status data.f

fileName = SMAP_L1A_Radar_00016_A_20150530T160100_R00201_001.h5

identifier = L1A_Radar

version = R00201

Ephemeris

creationDate = 2015-05-29

description = SPICE SP-kernel of SMAP trajectory

fileName = smap_tr_1505291400_1512291400_v01.bsp

version = 01

AntennaAzimuth

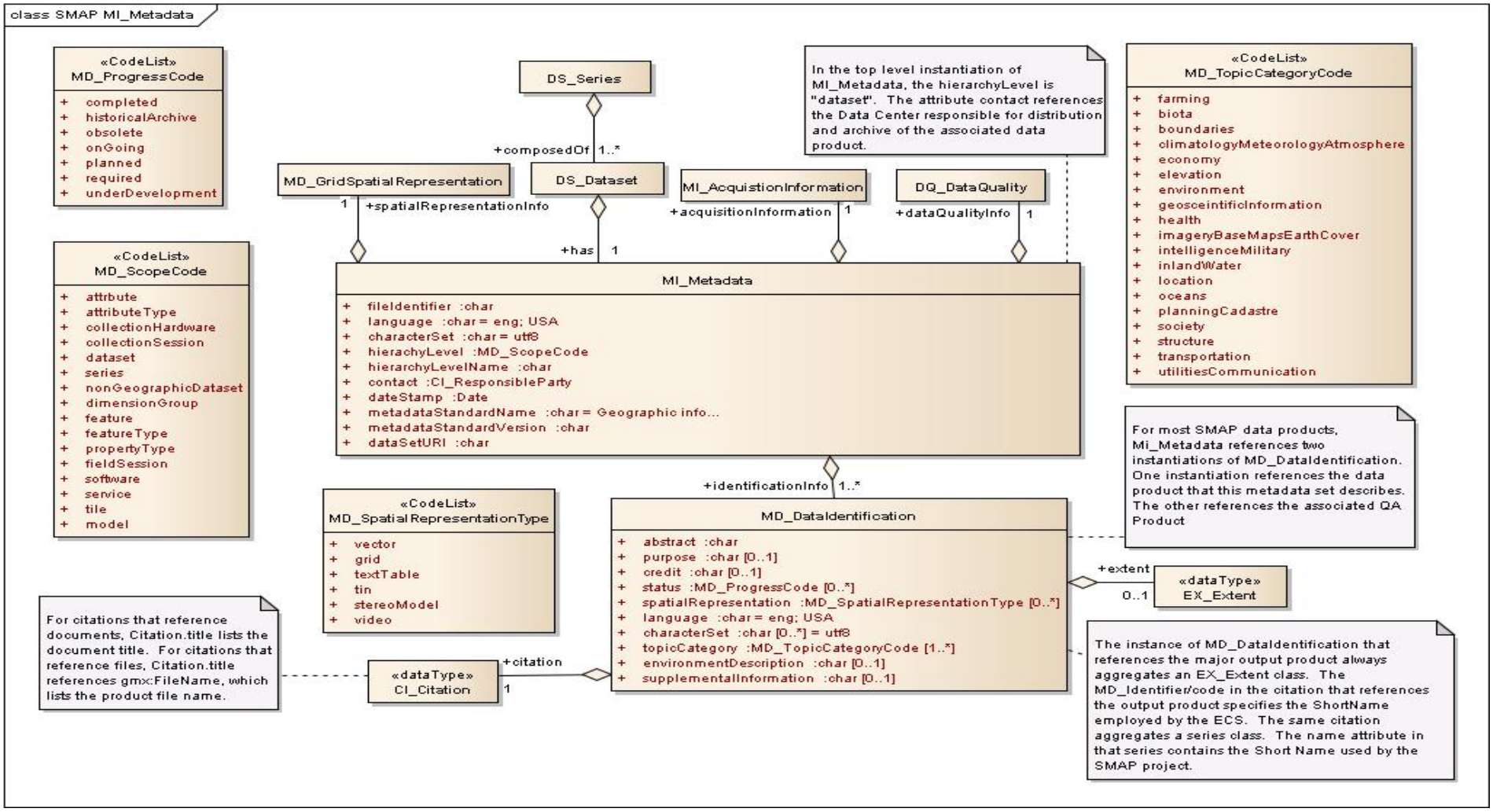
…….

ISO 19139 Compliant XML

Sequences like the following will appear in separate metadata.

```
<gmd:source>
        <gmi:LE_Source>
                <gmd:description>
                        <gco:CharacterString>description of the radar level 1C SPS input configuration file product</gco:CharacterString>
                </gmd:description>
                 <gmd:sourceCitation>
                        <gmd:CI_Citation>
                             <gmd:title>
                                  <gmx:FileName>SMAP_L1C_S0_HIRES_SPS_InputConfig_L1A_Radar.xml</gmx:FileName>
                             </gmd:title>
                            <gmd:date>
                                <gmd:CI_Date>
                                     <gmd:date>
                                          <gco:Date>2015-05-30</gco:Date>
                                     </gmd:date>
                                     <gmd:dateType>
                                          <gmd:CI_DateTypeCode codeList=
http://www.isotc211.org/2005/resources/CodeList/gmx:Codelists.xml#CI_DateTypeCode
codeListValue="creation">creation</gmd:CI_DateTypeCode>
                                     </gmd:dateType>
                                </gmd:CI_Date>
                            </gmd:date>
                            <gmd:edition>
                                 <gco:CharacterString>D00301</gco:CharacterString>
                            </gmd:edition>
                           <gmd:identifier>
                               <gmd:MD_Identifier>
```

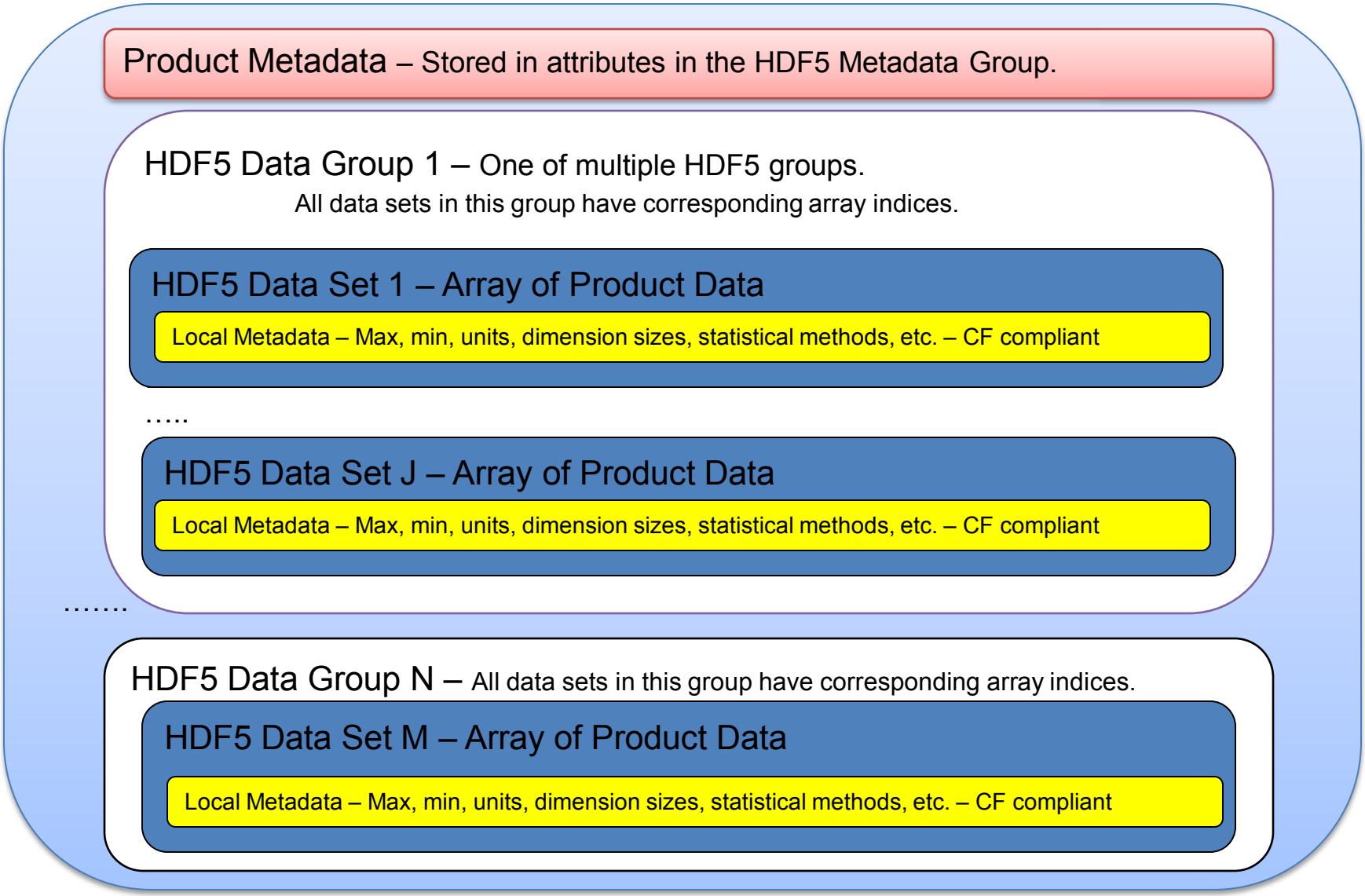
Proposed Top Level ISO Metadata Model



Quality Information

- Product metadata – covers the content of the entire product
  - Overall quality information for entire product content
    - Completeness – fraction of expected data that appears in the product
    - Range Check – fraction of data pixels that fall within an acceptable range
    - Quality Check – fraction of data pixels that are deemed acceptable quality
- Associated with each pixel – enables users to assess quality of each individual element
  - Quantitative error bar whenever possible
  - Bit flags for qualitative information
    - Individual bits flag individual conditions
    - In general, '0' indicates desirable conditions, '1' represents questionable or undesirable conditions
    - When the overall bit flag value is zero, the data are clean

Proposed SMAP Data Product Structure



SMAP Data Product Availability

Data Product Short Name	Description	Initial Availability After Commissioning	Latency to User Community after Acquisition
L1A_Radar	Parsed Radar Instrument Telemetry	3 months	12 hours
L1A_Radiometer	Parsed Radiometer Instrument Telemetry	3 months	12 hours
L1B_S0_LoRes	Low Resolution Radar σ_o in Time Order	3 months	12 hours
L1C_S0_HiRes	High Resolution Radar σ_o on Swath Grid	3 months	12 hours
L1B_TB	Radiometer T_B in Time Order	3 months	12 hours
L1C_TB	Radiometer T_B	3 months	12 hours
L2_SM_A	Radar Soil Moisture	3 months	24 hours
L2_SM_P	Radiometer Soil Moisture	3 months	24 hours
L2_SM_AP	Active-Passive Soil Moisture	3 months	24 hours
L3_FT_A	Daily Global Composite Freeze/Thaw State	6 months	50 hours
L3_SM_A	Daily Global Composite Radar Soil Moisture	6 months	50 hours
L3_SM_P	Daily Global Composite Radiometer Soil Moisture	6 months	50 hours
L3_SM_AP	Daily Global Composite Active-Passive Soil Moisture	6 months	50 hours
L4_SM	Surface & Root Zone Soil Moisture	6 months	7 days
L4_C	Carbon Net Ecosystem Exchange	6 months	14 days

Availability of Validated Data Products

- The initial product delivery to the public will be three months after instrument commissioning, six months after launch. **These will be beta version products**.
  - Initial **validated** Level 1 products appear six months after instrument commissioning, or about nine months after launch.
  - Initial **validated** Level 2, Level 3 and Level 4 products appear twelve months after instrument commissioning, or about fifteen months after launch.

Data Maturity Levels

BetaProducts intended to enable users to gain familiarity with the parameters and the data formats.

ProvisionalProduct was defined to facilitate data exploration and process studies that do not require rigorous validation. These data are partially validated and improvements are continuing; quality may not be optimal since validation and quality assurance are ongoing.

ValidatedProducts are high quality data that have been fully validated and quality checked, and that are deemed suitable for systematic studies such as climate change, as well as for shorter term, process studies. These are publication quality data with well-defined uncertainties, but they are also subject to continuing validation, quality assurance, and further improvements in subsequent versions. Users are expected to be familiar with quality summaries of all data before publication of results; when in doubt, contact the appropriate instrument team.

- **Stage 1 Validation**: Product accuracy is estimated using a small number of independent measurements obtained from selected locations and time periods and ground-truth/field program efforts.
- **Stage 2 Validation**: Product accuracy is estimated over a significant set of locations and time periods by comparison with reference in situ or other suitable reference data. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and time periods. Results are published in the peer-reviewed literature.
- **Stage 3 Validation**: Product accuracy has been assessed. Uncertainties in the product and its associated structure are well quantified from comparison with reference in situ or other suitable reference data. Uncertainties are characterized in a statistically robust way over multiple locations and time periods representing global conditions. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and periods. Results are published in the peer-reviewed literature.
- **Stage 4 Validation**: Validation results for stage 3 are systematically updated when new product versions are released and as the time-series expands.

Extracted from http://science.nasa.gov/earth-science/earth-science-data/data-maturity-levels/

National Aeronautics and Space Administration

Anticipated SMAP Data Product Volumes

Data Product Short Name	Description	Daily Volume (GBytes)	Yearly Volume (TBytes)
L1A_Radar	Parsed Radar Instrument Telemetry	89.135	32.557
L1A_Radiometer	Parsed Radiometer Instrument Telemetry	32.879	12.009
L1B_S0_LoRes	Low Resolution Radar σ_o in Time Order	9.084	3.318
L1C_S0_HiRes	High Resolution Radar σ_o on Swath Grid	44.728	16.356
L1B_TB	Radiometer T_B in Time Order	1.958	0.715
L1C_TB	Radiometer T_B	0.472	0.172
L2_SM_A	Radar Soil Moisture	2.144	0.783
L2_SM_P	Radiometer Soil Moisture	0.015	0.006
L2_SM_AP	Active-Passive Soil Moisture	0.149	0.054
L3_FT_A	Daily Global Composite Freeze/Thaw State	2.349	0.343
L3_SM_A	Daily Global Composite Radar Soil Moisture	10.014	1.218
L3_SM_P	Daily Global Composite Radiometer Soil Moisture	0.046	0.006
L3_SM_AP	Daily Global Composite Active-Passive Soil Moisture	0.842	0.102
L4_SM	Surface & Root Zone Soil Moisture	16.284	1.784
L4_C	Carbon Net Ecosystem Exchange	0.604	0.026

SMAP Data Simulations

- Two major simulated data sets
- GloSim1 Data
  - Mission scientists used these simulated data to generate initial sets of retrieval algorithms
  - Combines 0.01 degree North American dataset with a 0.25 Global Land Data Assimilation System (GLDAS) dataset.
- GloSim2 Data
  - These data enable end-to-end tests into Level 4 modeling, and thus provide more realistic scientific results
  - Forward model generates Radar Level 1C and Radiometer Level 1B Products based on these simulations
  - Level 1 simulations and developing algorithms generate higher level products

Simulated Soil Moisture and Freeze Thaw Data Products

- The simulated products are in HDF5 format
- The simulated products display the planned product design
  - "Metadata" Group contains product metadata
  - In the Level 2 Products, the "Soil_Moisture_Retrieval_Data" Group contains the data arrays
  - Some products may have additional groups
- The simulated data products contain a subset of the data and metadata that will appear in the products the SMAP project will distribute after launch.
- At the very least, all of the Soil Moisture and Freeze/Thaw products contain the following:
  - Latitude
  - Longitude
  - Soil Moisture or Freeze/Thaw
- Additional data elements appear that may contain fill values
  - In most cases, null values are currently NaN
  - The project will generate a standard for null values, and use them in all data products
    - The project will document the specific null value in the product metadata as well as product documentation

Current SMAP Product Simulations

- Simulated data products based on GloSim1 are available to Early Adopters
  - Contents are not complete, but are adequate to test developing algorithms
- The project processes these simulated data sets into HDF5 format that conforms with SMAP product design
  - Project moves data to NSIDC for distribution to Early Adopters
  - Use of HDF5 products enables users to:
    - Generate software now that would function after launch
    - Provide feedback about product format and design
  - The following data sets are currently on the server:
    - One year of SMAP L2_SM_P data
    - Two months of SMAP L2_SM_A data
    - Two months of SMAP L2_SM_AP data
    - Seven files of SMAP_L3_FT_A – representative dates that span from winter through summer
    - One file each of the SMAP_L3_SM_A, SMAP_L3_SM_P and L3_SM_AP

Current SMAP Product Simulations

- Two simulated data products based on GloSim2 are available to Early Adopters
  - Contents are not complete, but are adequate to test developing algorithms
  - The following data sets are currently on the server:
    - One year of SMAP L1B_TB data
    - One year of SMAP L1C_S0_HiRes data

Simulated Data Access

- SMAP simulated data products are available from a restricted FTP site at the National Snow and Ice Data Center (NSIDC)
  - **ftp://smap_data@sidads.colorado.edu/data**
  - Password is required

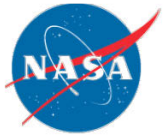
- Data access information has been sent to all Early Adopter members
  - If you need information resent or need assistance accessing data, please contact NSIDC User Services (nsidc@nsidc.org)

- NSIDC will notify Early Adopters as new simulated data products become available

Data Support @ NSIDC

- NSIDC provides general HDF and EASE-Grid documentation and tools
  - Information is included within the /data/00README.txt file on the FTP site
- NSIDC User Services is available for data access questions and general HDF and EASE-Grid support
  - NSIDC User Services currently does not support SMAP data product information
  - To contact User Services email **nsidc@nsidc.org**

Validation Data Access



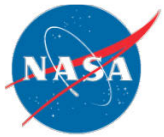
- Validation data are available at the following URL:
  - **ftp://smap_data@sidads.colorado.edu/validation_data**
  - Password is required

- Data access information has been sent to all Early Adopter members
  - If you need information resent or need assistance accessing data, please contact NSIDC User Services (nsidc@nsidc.org)

- NSIDC will notify Early Adopters as new validation data products become available

Current SMAP Validation Data

- Data from two validation campaigns are available.
  - SMAPVEX08: Soil Moisture Active Passive Validation Experiment 2008
  - CLASIC07: Cloud and Land Surface Interaction Campaign 2007
- Several data sets are available for each campaign
  - SMAPVEX08
    - PALS Backscatter, PALS TBs, VWC Map, Land Cover Map, In situ Soil Moisture, Soil Texture, Surface Roughness, In situ Vegetation
  - CLASIC07
    - PALS Backscatter, PALS TBs, VWC Map, Land Cover Map, In situ Soil Moisture, Soil Texture, In situ Vegetation
- README files describing each validation data set are provided in each subdirectory.



Backup

Planned Ancillary Data Sets

- SMAP processing algorithm developers are considering a large set of ancillary data sets
  - Approximately 20 sources of ancillary data for Level 1
    - Many are provided by the mission and are obvious – spacecraft ephemeris, spacecraft attitude, antenna pointing information, spacecraft clock correlation information
    - Some of the ancillary data sets that are currently under consideration will not be used
    - Some data sets under consideration will be combined or used to generate climatologies
  - Approximately 12 sources of ancillary data for Level 2
    - Ancillary data set selection process is more mature
    - Specific sources for many of these data sets are TBD
- Ancillary data sets will be selected using multiple criteria
  - Applicability – how well they fit into the application
  - Accessibility – how readily available the data are, how easy the data are to access
  - Timeliness – whether availability impacts data product latency requirements
  - Ease of Use
  - Quality

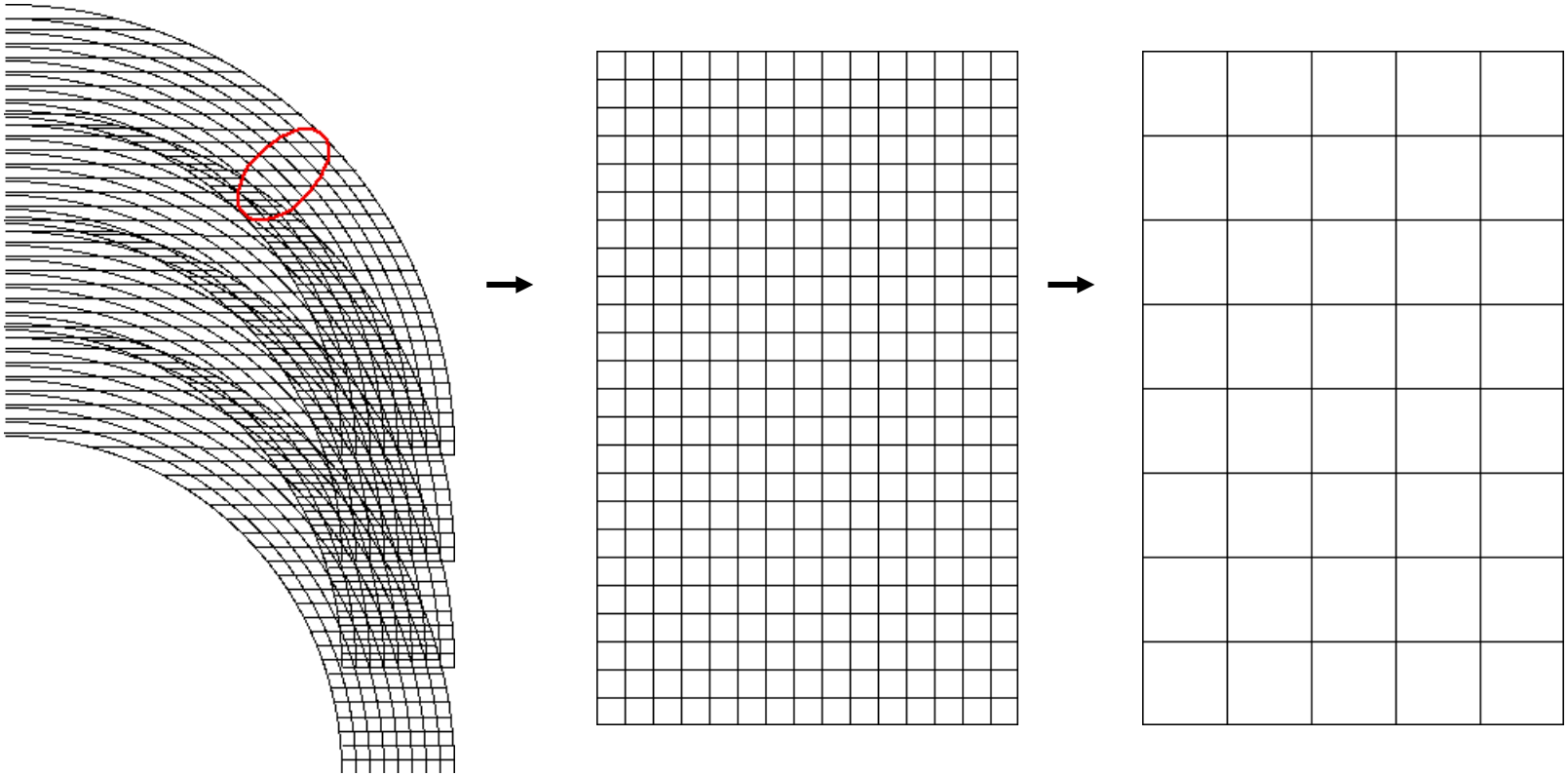
Ancillary Data in SMAP Science Data Products

- Much of the ancillary data appears in SMAP data products
- The proposed SMAP strategy for inclusion of ancillary data in data products:
  - Dynamic ancillary data needed for algorithmic use would appear in SMAP products
  - Static ancillary data needed for algorithmic use may appear in distributable tables rather than in the data products
    - The Data Centers will provide the static ancillary data tables to Early Adopters before launch
    - Data Centers will provide the static ancillary data tables to the general public after launch

Current SMAP Validation Data

- Data from two validation campaigns are available.
  - SMAPVEX08: Soil Moisture Active Passive Validation Experiment 2008
  - CLASIC07: Cloud and Land Surface Interaction Campaign 2007
- Several data sets are available for each campaign
  - SMAPVEX08
    - PALS Backscatter, PALS TBs, VWC Map, Land Cover Map, In situ Soil Moisture, Soil Texture, Surface Roughness, In situ Vegetation
  - CLASIC07
    - PALS Backscatter, PALS TBs, VWC Map, Land Cover Map, In situ Soil Moisture, Soil Texture, In situ Vegetation
- README files describing each validation data set are provided in each subdirectory.

Radar Data Products



Single-Look, Time-Ordered Data

- Native resolution: 250 m in range, 400+ m resolution in azimuth.
- Each resolution element constitutes one independent "look" at the surface.

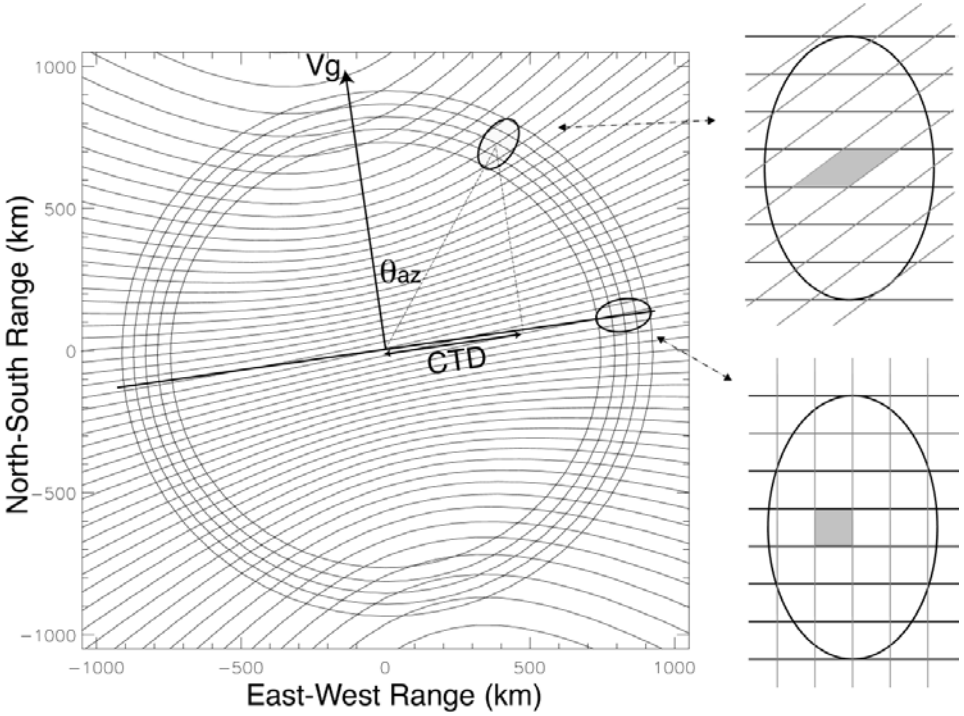
1 km Gridded, Re-Sampled Data

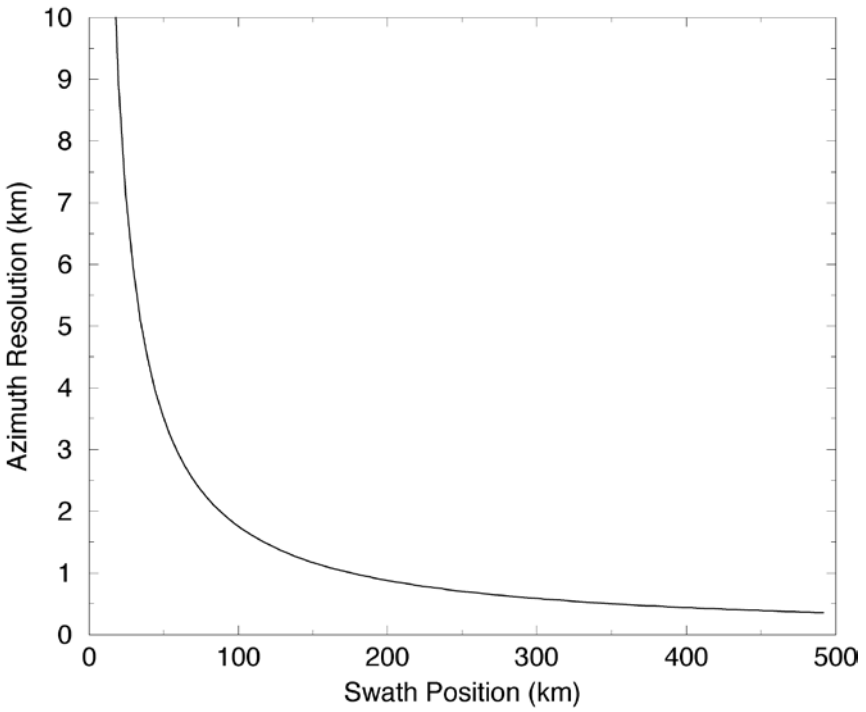
- Data resampled and posted on 1 km grid, resolution may still be > 1 km near nadir.
- Each resolution cell now has multiple "looks" at surface, decreased measurement variance.

3 km Averaged Data

- 1 km posted product can be averaged up to 3 km, 10 km, etc. by investigators.
- Improved number of looks (and hence precision) at expense of spatial resolution.

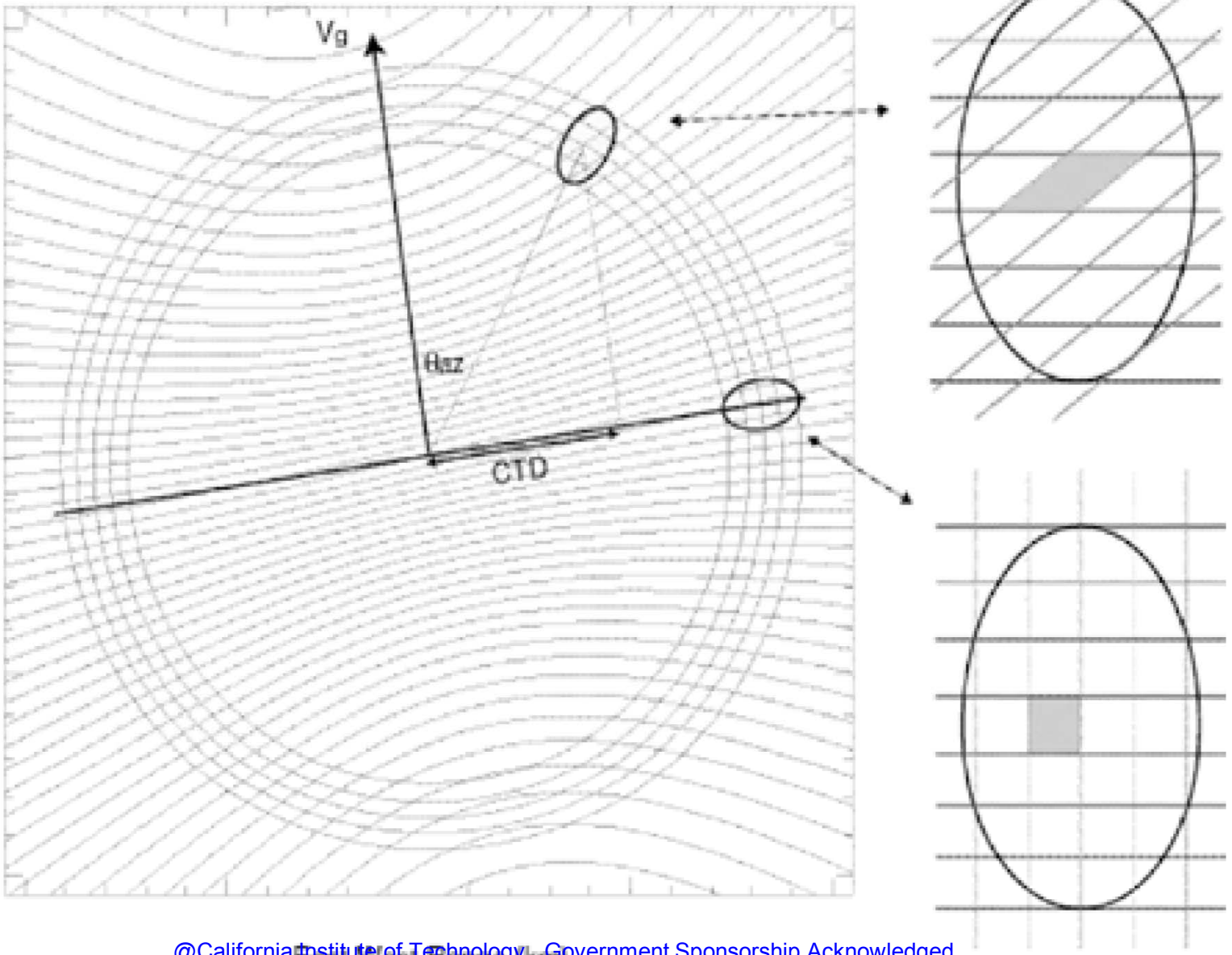
Radar Resolution Approach



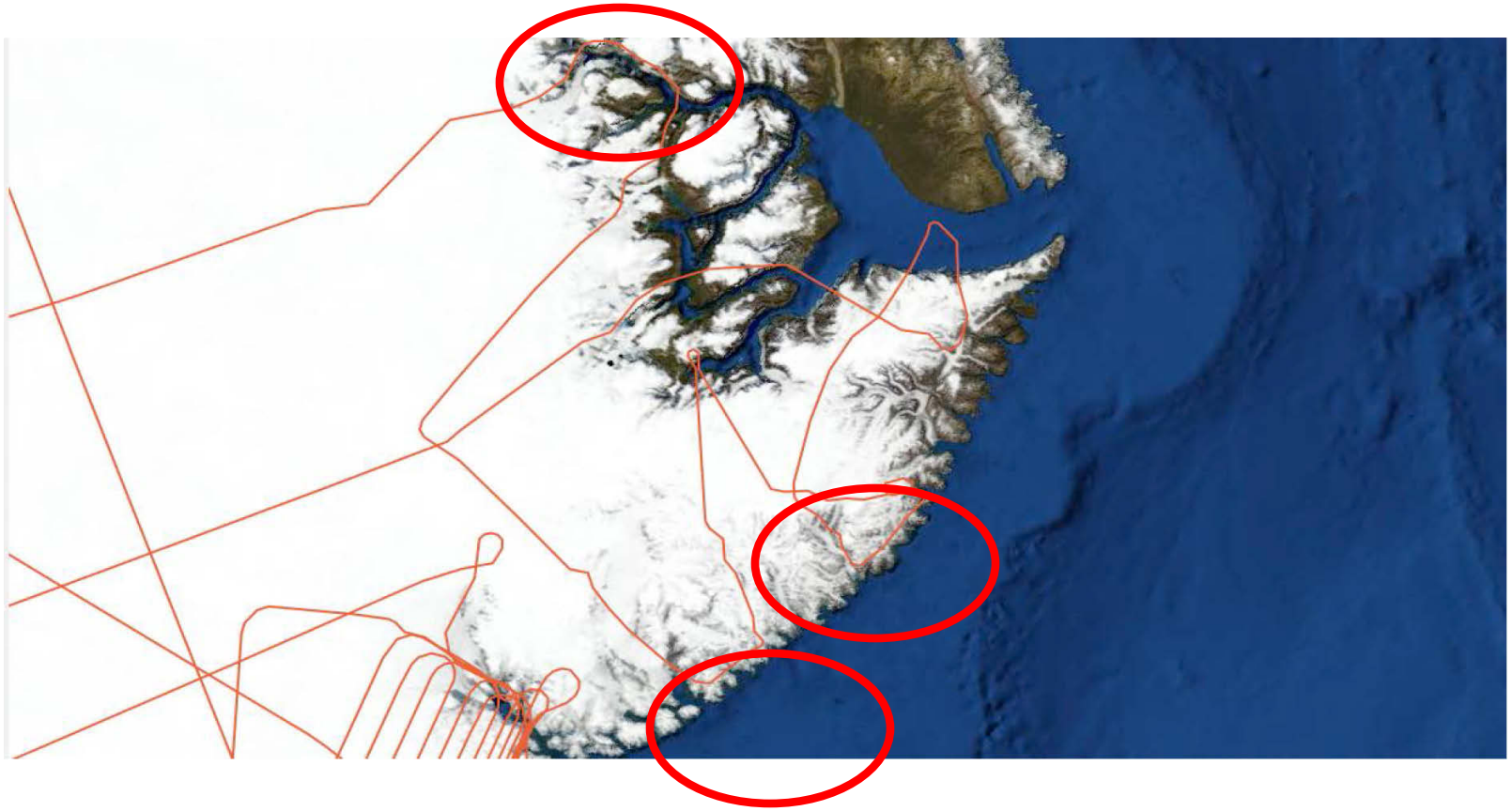


- Unfocused SAR processing.
- Azimuth resolution, and number of azimuth looks, driven by unique scanning geometry.
- High-resolution SAR data that can meet the proposed science requirements for resolution and accuracy over the outer 70% of the measurement swath.

Radar Measurement Geometry



Geolocation Errors with EASE-Grid



Example of geolocation error induced in NASA (500 m) Blue Marble incorrectly re-projected to EASE-Grid. The problem was discovered years later, when IceBridge flight lines were overlaid. The flight manager indicated that the plane flew down the middle of the fjord, not off to the side. Also noted that coastline traverses were over the ocean.

Proposal for EASE-Grid-2.0

- 2010: EASE-Grid has been in use for almost 20 years
- Problems are accumulating, especially datum/projection ellipsoid issues.
  - Haran, T. 2008. HDF-EOS vs. GeoTIFF: GIS problems when Projections and Datum Spheroids are Different. MODIS Science Team Meeting. ftp://sidads.colorado.edu/pub/ppp/conf_ppp/Haran/HDF-EOS_vs._GeoTIFF__GIS_Problems_when_Projection_and_Datum_Spheroids_are_Different.pdf
  - Billingsley, B. 2008. Using GeoTIFFs for Data Sharing: Limitations and Solutions. AGU Fall Meeting ftp://sidads.colorado.edu/pub/ppp/conf_ppp/Billingsley/Using_GeoTIFFs_for_Data_Sharing__Limitations_and_Solutions.pdf
- NSIDC proposes changes that addresses known problems
- EASE-Grid-2.0 Goals:
  - Eliminate datum/ellipsoid issue to ensure that
    - software packages “do the right thing”
    - geoTIFFs are possible without reprojection
  - Make nesting grids simpler
  - Eliminate “undefined” grid cells at corners of azimuthal grids
  - Decouple scales between cylindrical and azimuthal grids
  - Choose new dimensions that will immediately distinguish data from EASE-Grid(-1.0) data

EASE Grid Comparison

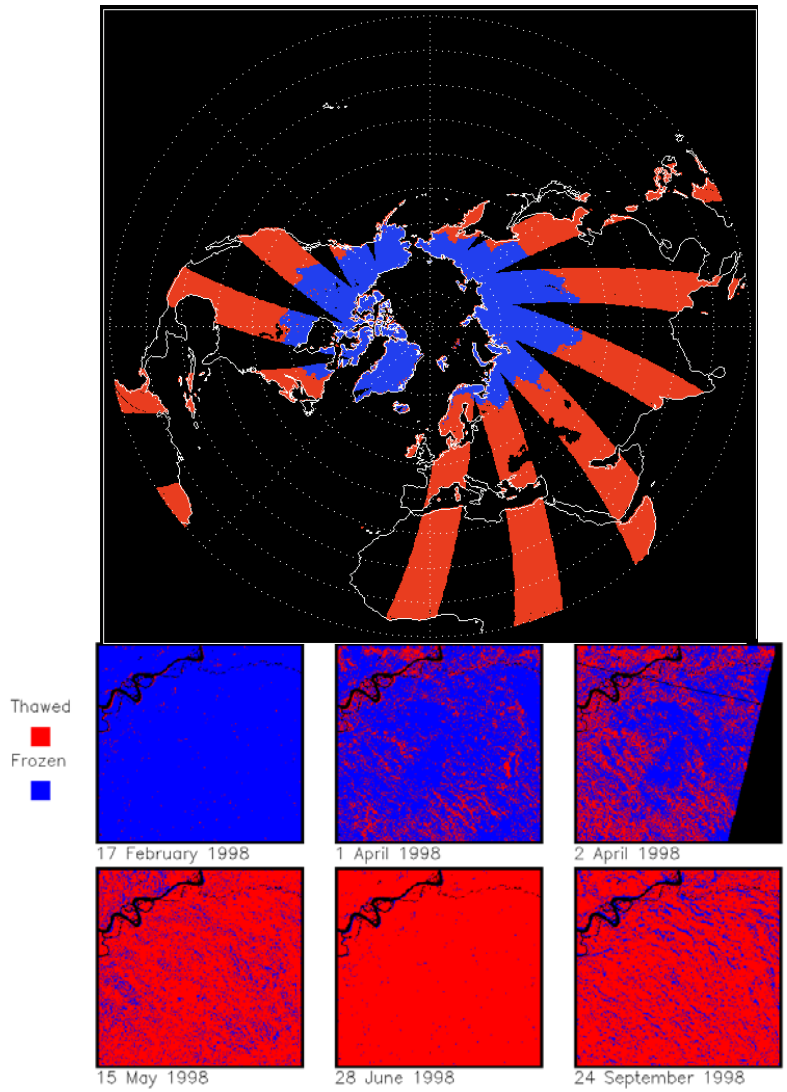
Advantages

- EASE grid comprises a family of three projections that serve a large community: North and South Azimuthal for polar research, and Global cylindrical for low- to mid-latitude applications.
- Equal-area projection makes area calculations “easy”.
- “Scalability” provides for an infinite number of grid resolutions (“graph paper”) to suit any specific application.
- Growing community of adopters make inter-comparison and time-series easy.
- Branding: “EASE-Grid” is recognized, the concept is easy to learn and easy to adapt to new applications.

Disadvantages

- GIS and other software packages ***assume reference datum and projection ellipsoid were the same***: This was hard to explain to users, and harder for them to override in order to teach their software to interpret/re-project it.
- EASE-Grid data ***cannot be formatted into a legitimate geoTIFF*** without re-projecting.
- Original ***scale was defined as 200.5402/8. =~ 25.0675… messy***, hard to represent; exact value was necessary for integer cells in global grid, but shouldn’t have influenced scale in azimuthal grids.
- Equator in azimuthal grids was properly “inside” the span of the grid array, which made ***corner points undefined since they “wrapped beyond” opposite pole.***

Level 3 Freeze/Thaw Product



Daily Freeze/Thaw State

- Each product represents a single calendar day UTC.
- Posted on a 3 km polar EASE grid with 1 to 3 km spatial resolution using a two dimensional array.
- Quantifies daily freeze/thaw state as a binary condition for land surface.
- Includes both AM and PM data
- Employs the 1 km Level 1C high resolution radar data and a time-series change detection algorithm to infer freeze/thaw state.
- Required to achieve 80% freeze/thaw state classification accuracy.

CF Convention – Local Metadata

- The Climate and Forecast (CF) is a highly descriptive metadata convention with a widespread science user community
  - CF designed specifically designed to fit within attributes in netCDF files.
  - CF is based upon the Cooperative Ocean/Atmospheric Data Service (COARDS) standard
- The CF convention includes:
  - A standard to provide descriptive names for each variable in the product
  - Standards for the specification of data units for each variable in the product
    - UDUNITS provides a list of supported unit names
  - Standards for fill values for each variable in the product
  - Standards to express the range of data for each variable in the product
  - Standards to express bit flag definitions and define flag values
  - Standards to specify relationships between spatial and time coordinates for each variable in the product
    - Indicates which particular spatial or temporal coordinates correspond with which dimension axes and indices of a data variable.
  - Standards to specify statistical methods that were used to calculate each variable in the product
    - Clarifies temporal or spatial intervals that were used to provide statistical results.

CF Global Metadata

- Five standard elements for description of the entire product
  - Title
  - Institution
  - Source
  - History
  - References
  - Each of the above five elements are character free-form stings, with no specification about content.

Metadata Use Recommendation

- The ISO 19115 provides ample metadata to cover the proposed SMAP requirements for granule and collection metadata
- The CF conventions provide detail required for local metadata that describe individual arrays in each product
- The small number of global elements that are required in the CF convention presents a minor issue:
  - Could SMAP ignore the small number of CF global elements, and remain “CF compliant”?
  - Is there some means to map the CF global elements into ISO?

Ancillary Data in GloSim2 SMAP Simulations

The table below lists ancillary datasets that are being used for SMAP GloSim2 simulations. They are reflective of the datasets that SMAP may use in Level 2 processing. The table does not reflect the ultimate choice of ancillary data sets for the operational Level 2 algorithms.

Data Type	Resolution	Size	Notes
Soil Moisture (Mv)	9-km EASE grid/hourly	25 MB	GMAO Nature Run
Surface Temperature (Ts)	9-km EASE grid/hourly	25 MB	GMAO Nature Run
NDVI climatology	1 to 36-km EASE grid/10-day (plus max/min grids)	1.02 GB max	Bindlish/Jackson; interpolated in time at the grid cell to use in VWC algorithm
Water Body Fraction	1 to 36-km EASE grid	0.51GB max	MODIS land/water database (generated by S. Chan)
Landcover type	1 to 36-km EASE grid	0.51GB max	Dominant IGBP class at each grid cell (generated by S. Chan)
Sand/Clay fraction, bulk density	1 to 36-km EASE grid	0.51GB max	See N. Das soil texture memo; data are scaled to byte type to save space
Crop type	1-km EASE grid	0.51GB	Based on USDA, Canada, and Europe crop databases; other regions assigned
Surface roughness (s)	1-km EASE grid	2.04GB	Based on Hydros/J. Johnson means, with added variance by IGBP class
Radar Datacubes	3-channel sigma0, function of roughness, dielectric constant, and VWC (280,280,140)	0.125GB each, 2.0 GB total	Currently 16 cubes including 4 crop types, assigned by IGBP/crop indices

GloSim2 Radiometer Simulation Flow

MATLAB LINE INPUT

Specify orbit ID range (1-5,400)

→ **Run Sim Cycle**

Run orbit/instrument sampling

- Read SMAP time in UTC @ 12-ms interval
- Read 3-dB FOV arrays of lat/lon's, antenna gain, and incidence

Read global 1km semistatic data

- Read annual max NDVI
- Read NDVI at two adjacent 10-day periods
- Interpolate NDVI for a given DOY; interpolate again within a single half orbit if necessary

Read global 1-km static data

- Sand Fraction
- Clay Fraction
- Bulk Density
- IGBP Dominant Land Cover
- Surface RMS Roughness

Read global 9-km dynamic data

- Read GMAO hourly 10-cm soil moisture; read two files within a single half orbit if necessary
- Read GMAO hourly 10-cm surface temperature; read two files within a single half orbit if necessary

Static Data

Dynamic Data

→

STEP 1: Sample at FOV lat/lon's

- 246,575-by-338 elements per half orbit

STEP 2: Water pixels within FOV

- Compute dielectric constant using Klein-Swift model
- Compute water TB using Fresnel equations, along with FOV incidence

STEP 3: Land pixels within FOV

- Compute dielectric constant using Mironov model (thawed state)
- Compute dielectric constant using modified Dobson model (frozen state)
- Construct VWC array using (a) interpolated NDVI, (b) annual max NDVI, and (c) IGBP-based stem factor
- Construct 'b' and 'omega' arrays from IGBP-based lookup table
- Construct 'h' array from 's' array using the h = 0.1 * s relationship
- Compute land TB using Fresnel equations, along with FOV incidence and tau-omega model

STEP 4: Antenna synthesis

- Integrate water TB and/or land TB according to FOV antenna gain. Result: one TB per boresight location

TB Calculations

→

Radiometer Level 1B File

Noise-free time-ordered radiometer H- and V-polarized TB observations output at every 12 ms

GloSim2 Radar Simulation Flow