SMAP Level-2 Soil Moisture Cal/Val Requirements

03 May 2011

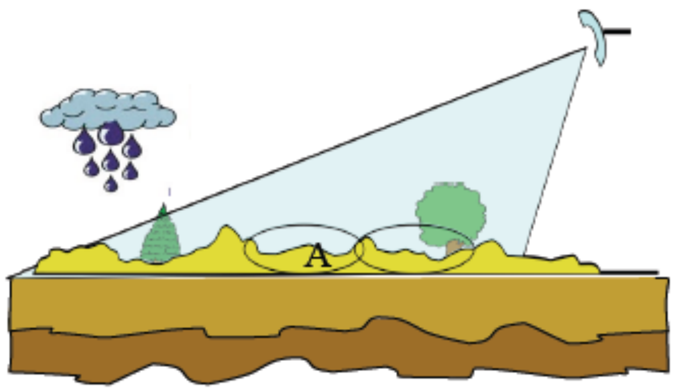
Mahta Moghaddam (University of Michigan)
and SMAP Soil Moisture Algorithms Team

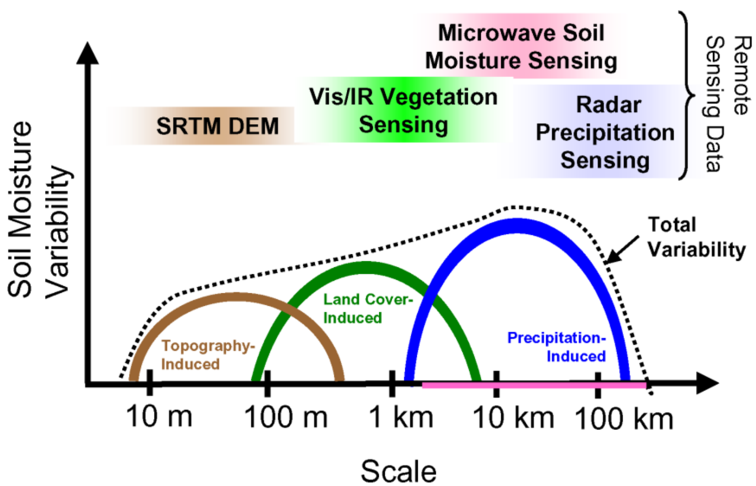
Surface Soil Moisture Products

- Level 2 products are from half orbit data acquisitions
- Level 3 products are daily composites

Data Product	Description	Grid Spacing
L2_SM_P	Soil Moisture (radiometer, half-orbit)	36 km
L2_SM_A/P	Soil Moisture (radar/radiometer, half orbit)	9 km
L2_SM_A	Soil Moisture (radar, half orbit) (Internal Product)	3 km
L3_SM_P	Soil Moisture (radiometer, daily composite)	36 km
L3_SM_A/P	Soil Moisture (radar/radiometer, daily composite)	9 km

The Issue of Scale





Soil moisture is a key state variable in the hydrologic cycle that varies on spatial scales of meters to tens of kilometers, and temporal scales of minutes to days.

SMAP's radar and radiometer instruments will each measure soil moisture, but with different spatial resolutions and different sources of uncertainty. Both capture significant landscape heterogeneity due to large pixel sizes. Each heterogeneous pixel can be thought of as containing many homogeneous sub-pixels (areas "A" in the top figure).

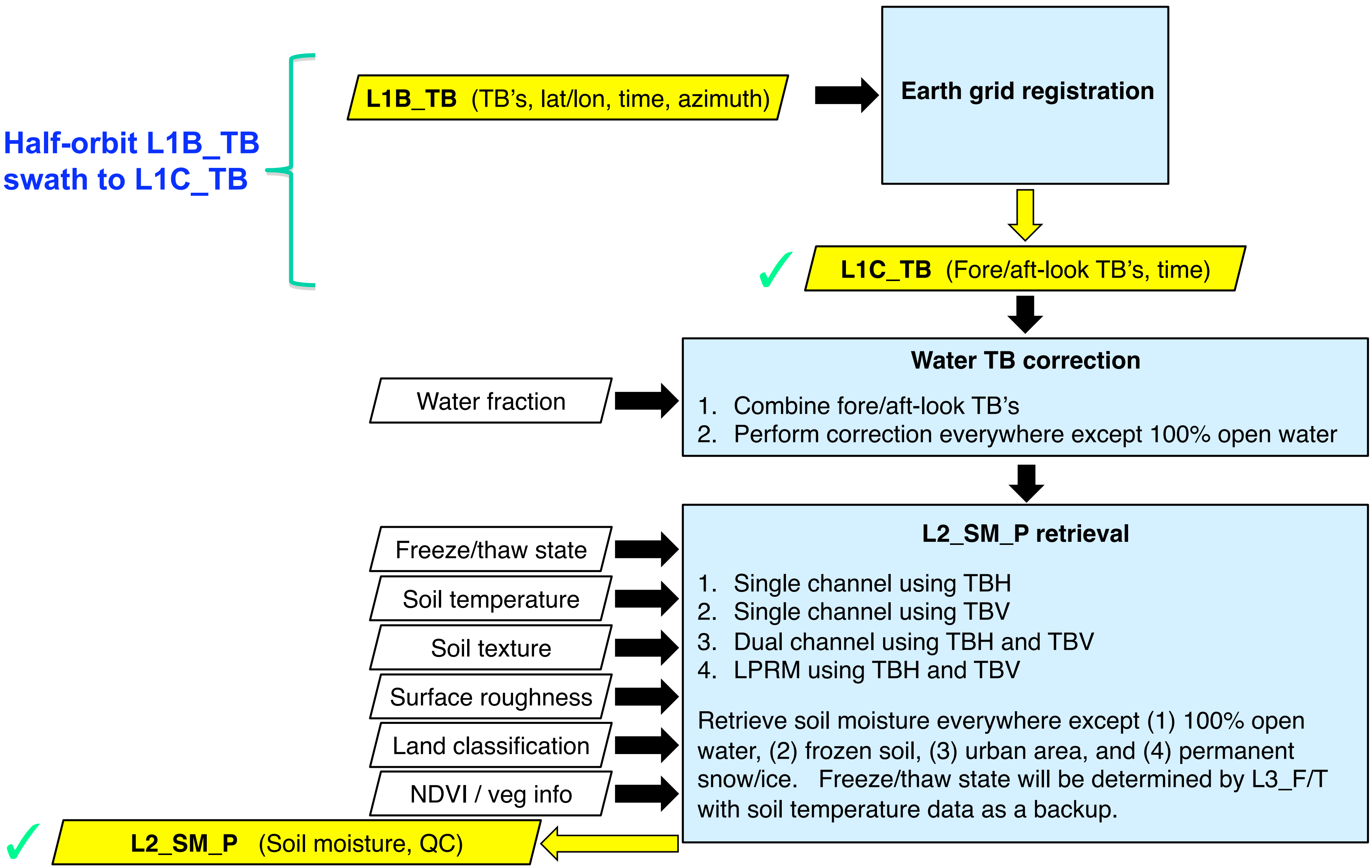
Validation of these measurements requires sophisticated adaptive temporal and spatial sampling strategies.

Land-cover Classes to Validate

- Both active and passive soil moisture algorithms rely on having accurate land cover class information
- All algorithms will use IGBP land-cover units for consistency
- Algorithm development and validations will therefore be done for these classes
- Additionally, crop types will be resolved where available

IGBP class	Definition of IGBP class
Evergreen needle	Tree canopy cover > 60%
Evergreen broadleaf	Same as above
Deciduous needle	Same as above
Deciduous broadleaf	Same as above
Mixed forest	Tree canopy cover > 60% but no type exceeds 60%
Closed shrub	Shrub cover > 60% (evergreen or deciduous)
Open shrub	60% > shrub cover > 10%
Woody Savanna	Herbaceous system & 60% > forest > 30%
Savanna	Herbaceous system & 30% > forest > 10%
Grassland	Herbaceous system & 10% > forest
Wetland	Water + Herbaceous system + tree
Cropland	Temporary crops
Crop/pasture mix	No class exceeds 60%
Bare	Vegetation cover < 10%

L2_SM_P: Processing Flow



L2_SM_P: Algorithm Inputs / Outputs

<u>DATA INPUT</u>:
Grid cell location on fixed Earth grid (lat, lon)
Time tag (date and time of day)
Calibrated water-corrected L1C_TB
Static ancillary data: -- Permanent masks (land, water, urban, etc.) -- Soil type, DEM, % land cover types
Dynamic ancillary data :
-- Soil temperature
-- Vegetation water content
-- Vegetation parameters (*b*, τ, ω)
% open water in pixel [from HiRes radar] -- temperature of open water from Ts at 6 am
Frozen ground flag [from L3_F/T]
Precipitation flag (if set)
Snow/ice flag (if set)
RFI flag [from L1_TB]
Quality flag [from L1_TB & internally set]

<u>DATA OUTPUT</u>:
Grid cell location on fixed Earth grid (lat, lon)
Time tag (date and time of day)
Calibrated water-corrected L1C_TB
Retrieved soil moisture for 6 am overpass
Dynamic ancillary data :
-- Soil temperature
-- Vegetation water content
-- Vegetation parameters (*b*, τ, ω)
% open water in pixel -- temperature of open water
Frozen ground flag
Precipitation flag (if set)
Snow/ice flag (if set)
RFI flag
Quality flag

L2_SM_P: Algorithm Options

- Several candidate radiometer retrieval approaches based on the tau-omega model are being evaluated, with varying requirements for ancillary data:
  - Single-Channel (SCA): uses H-pol brightness temperature which is corrected sequentially for surface temperature, vegetation water content, and surface roughness using ancillary data
  - Iterative (2CA): adjusts soil moisture and vegetation water content iteratively to minimize the difference between computed and observed T_{BV} and T_{BH}; both SM and another parameter (such as VWC) can be retrieved
  - Land Parameter Retrieval Model (LPRM): 2-channel iterative approach which uses a microwave polarization difference index and emissivity to parameterize τ_c; assumes τ_c and ω are the same for H and V polarization; assumes a constant ω
  - Reflectivity Ratio (RR): uses both T_{BV} and T_{BH} and vegetation & roughness correction factors for SM retrieval; algorithm proposes to use SMAP radar data to determine vegetation correction factor needed in the passive retrieval

L2_SM_P: Product Highlights

- Assumes TB accuracy of 1.3 K (error budget is quantified)
- Radiometer resolution is ~ 40 km
- L2_SM_P product output is on 36-km EASE grid
- Soil moisture retrieval accuracy requirement is 0.04 (cm^3/cm^3)
  - This requirement only applies to those areas whose vegetation water content (VWC) is ≤ 5 kg / m^2
  - Soil moisture is retrieved but not validated for areas of higher VWC

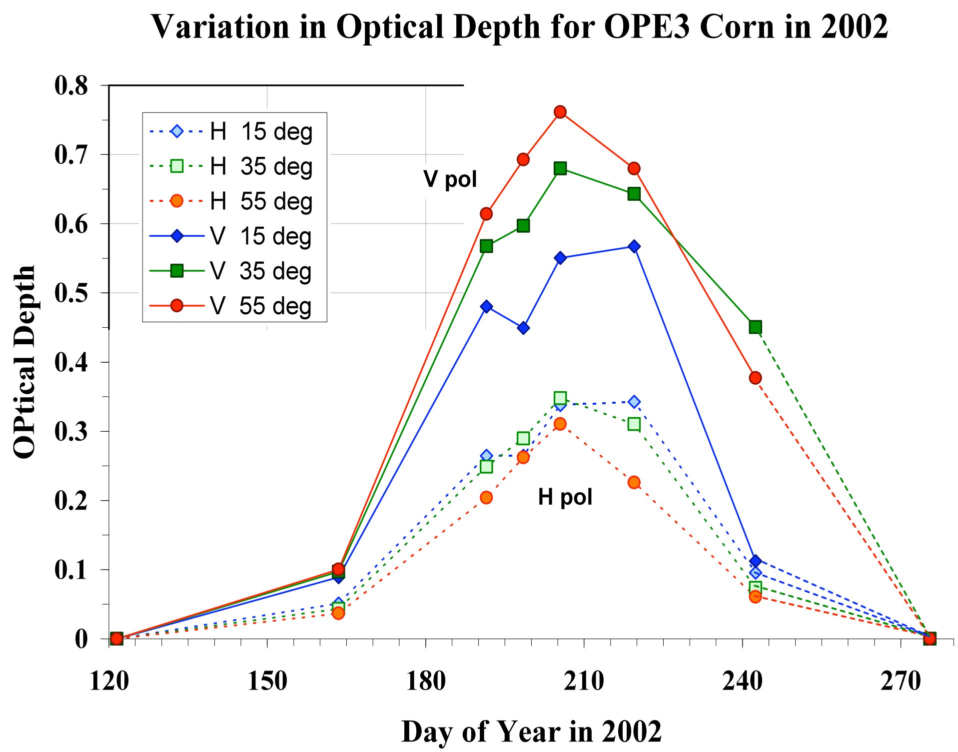
L2_SM_P: Key Open Issues

1. How well do simulations reproduce conditions observed with real data? Test algorithms using SMOS and Aquarius data when available and compare results to simulations.

2. Do the actual accuracies of ancillary data parameters such as T_S and VWC match what we are carrying in our error budgets? Compare to results from ancillary data studies and adjust error budgets accordingly if warranted.

3. Knowledge of the variation of vegetation parameters at H and V pol throughout the year is limited (see plot at right) – our simulations do not currently include such variations. Continue to compile data on temporal variations for each polarization (ComRAD 2012; SMAPEx; others?); analyze SMOS tau climatology when available.

4. How well do current algorithms do globally? How effective are algorithm parameterizations globally? Assess whether current parameterizations based on land cover type are sufficient using SMOS data.



Down-selection of the baseline algorithm will be based on a combination of higher accuracy in retrieved soil moisture, lower bias, better overall performance across land cover classes globally, and operational considerations, given the experimental and simulated data sets used.

L2_SM_P: Cal/Val Requirements

- For Cal / Val, would need to have:
  - In-situ measurements of 0-5 cm soil moisture
  - Effective temperature
  - Vegetation water content
  - All of the above at spatial sampling intervals and spatial coverage applicable to the 36-km SMAP grid
  - All of the above at sites covering IGBP land cover types as well as different crop types
  - Sufficient subsets of above acquired throughout the year to assess seasonality of algorithm parameterizations

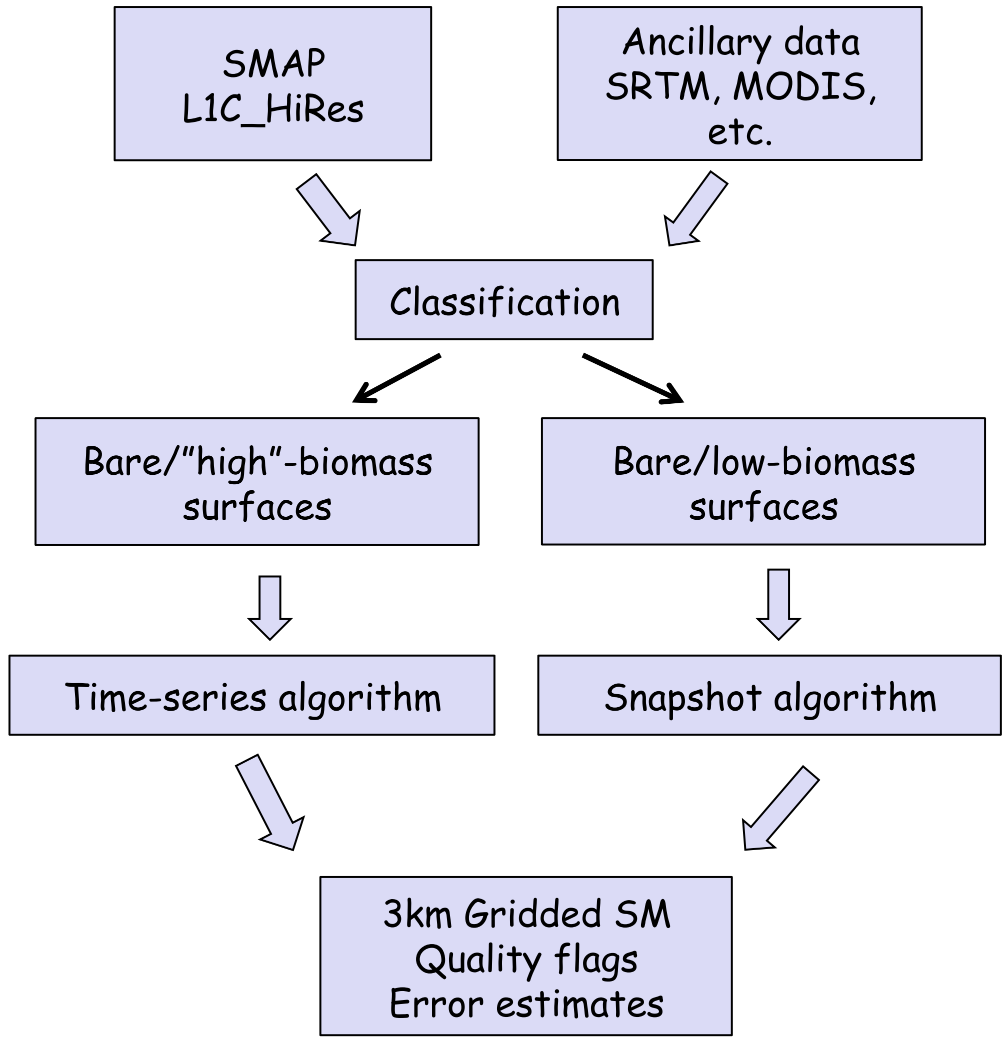
L2_SM_A: Processing Flow

Two algorithms are being carried in parallel:

(1) Snapshot method for bare surface of VWC less than ~0.5 kg/m2. Implementation is feasible over global surfaces but may be applicable only to low-biomass surfaces.

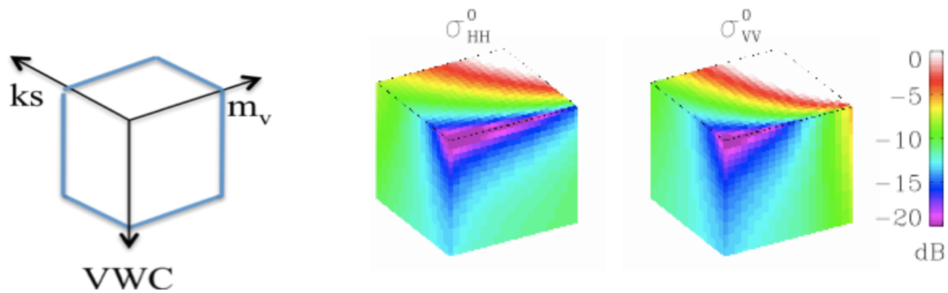
(2) Time-series method over vegetated surfaces up to 3 kg/m2. Either time-series has to be long enough or extreme moisture conditions are needed within the time window

** Both methods may apply to both bare and vegetated surfaces, but so far time-series algorithm has shown better performance for vegetated surfaces.



L2_SM_A: Algorithm Options

- Current baseline algorithm: optimization applied to a physical forward model, pre-computed and represented by data cubes for each radar channel
- Data Cubes are generated according to vegetation types (all IGBP classes will have their specific data cubes)



- Retrieval minimizes the cost *D*, evaluated either at snap shot or over the time-series

$$d(t) = \sum_{ch}^{HH,VV} w_{ch}\left(\sigma^0_{ch,measured} - \sigma^0_{ch,datacube}(m_{v,retrieve}, rough_{retrieve}, vegetation)\right)^2 \ \& \ D = \sum_{t}^{N} d(t)$$

- Other algorithm options have been or are being evaluated:
  - Dubois et al. (snapshot: empirical, no/low veg; likely will not be pursued)
  - Shi dual-pol (snapshot: equivalent roughness parameter; under test)
  - Kim & van Zyl (time-series: linear parameter fit; under test)
  - Wagner index (time-series: under test)

L2_SM_A: Key Open Issues

- As with other algorithms, need accurate land-cover classification and other ancillary data
- Soil moisture retrieval is sensitive to surface roughness; can assume unknown and solve for it, but will need cross-pol (HV); is HV high quality?
- Radar backscatter sensitivity to VWC is not unique and depends on vegetation geometry; will the finer definition of land-cover classes (a la IGBP) solve the problem?
- Retrieval results are sensitive to speckle; averaging up will reduce speckle but increase heterogeneity; where is the happy medium?
- Landscape heterogeneity
- Dielectric model

L2_SM_A: Cal/Val Requirements (1)

- Requirements on site representation
  - Soil moisture and roughness must vary over range consistent with mission expectations
  - Vegetation cover should include IGBP classes plus variety of crops
  - Vegetation should be dynamic (such as in crops) to test time-series algorithms
  - Sites should represent mixed vegetation and mixed land cover types to examine the heterogeneity issues

- Requirements on backscatter measurements
  - Sufficient relative calibration accuracy (UAVSAR has < 0.5dB; PALS is same or better)
  - Should provide time-series data
  - Tower-based or airborne

L2_SM_A: Cal/val Requirements (2)

Requirement on ground measurements:

Forward Model Validation
Soil Moisture
Soil Texture
Surface roughness profiles; row structure
Vegetation: woody geometry, water content
Vegetation: leaf geometry, water content
Vegetation: density
Vegetation: dielectric properties
Vegetation class
Fraction of vegetation cover in pixel

Retrieval validation
Soil Moisture
Soil Texture
Surface roughness
Vegetation water content or other vegetation surrogate
Vegetation class
Fraction of vegetation cover in pixel

L2_SM_AP: Algorithm Concept

Temporal changes in T_B and σ_{pp} are related. Parameters α and β are estimated at scale-*C* using successive overpasses.

$$T_{B_p}(C) = \alpha(C) + \beta(C)\cdot\sigma_{pp}(C)$$

Heterogeneity in vegetation and roughness conditions within scale-*C* evaluated by estimating sensitivities in radar cross-pol:

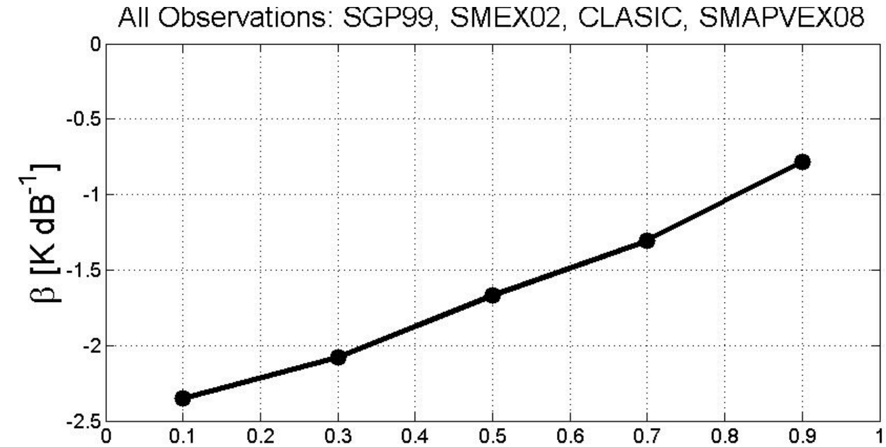
$$\sigma_{pp}(M_j) \approx \sigma^{e}_{pp}(M_j) + \left[\frac{\partial\sigma_{pp}(M_j)}{\partial\sigma_{pq}(M_j)}\right]_C \left(\sigma_{pq}(M_j) - \sigma_{pq}(C)\right)$$

T_B-disaggregation algorithm now becomes:

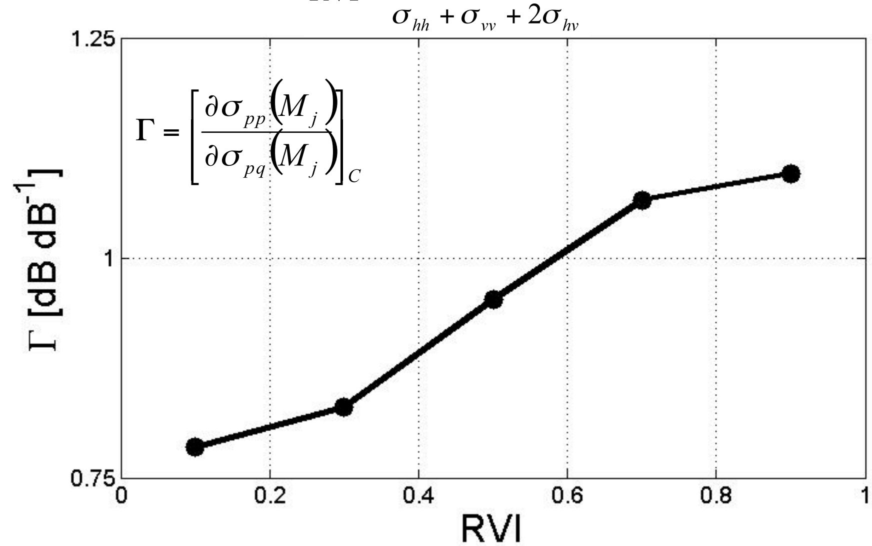
$$T_{B_p}(M_j) = T_{B_p}(C) + \beta(C)\cdot\left\{\left[\sigma_{pp}(M_j) - \sigma_{pp}(C)\right] - \Gamma\cdot\left[\sigma_{pq}(M_j) - \sigma_{pq}(C)\right]\right\}$$

$T_B(M_j)$ is used to retrieve soil moisture at 9 km (consistent algorithm and ancillary data as radiometer algorithm)

Airborne observations from four field experiments with PALS combine to form test database



$$\mathrm{RVI} = \frac{8\sigma_{hv}}{\sigma_{hh} + \sigma_{vv} + 2\sigma_{hv}}$$

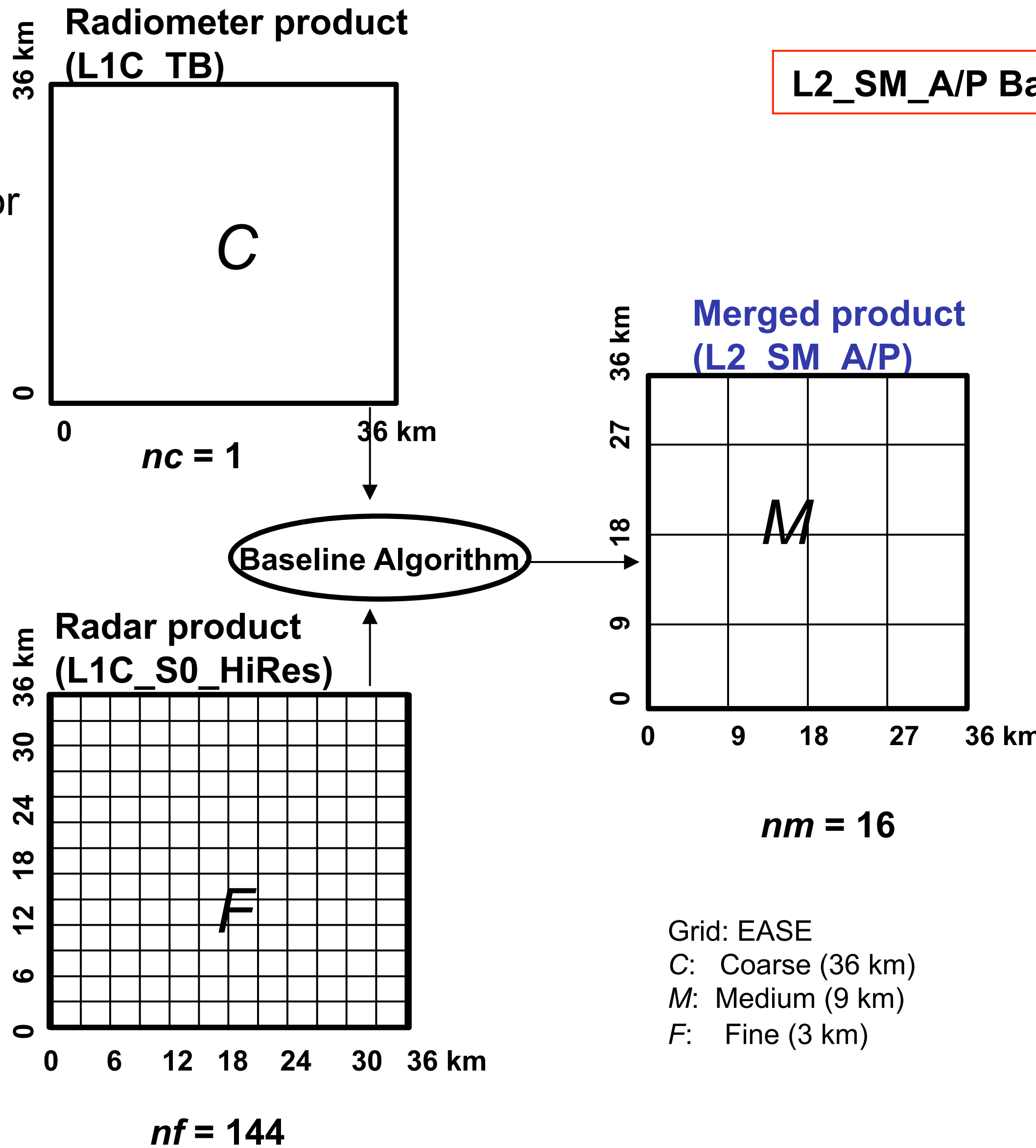


L2_SM_AP: Algorithm Schematic

L2_SM_A/P Baseline Algorithm

T_b (v) corrected for effect of water bodies

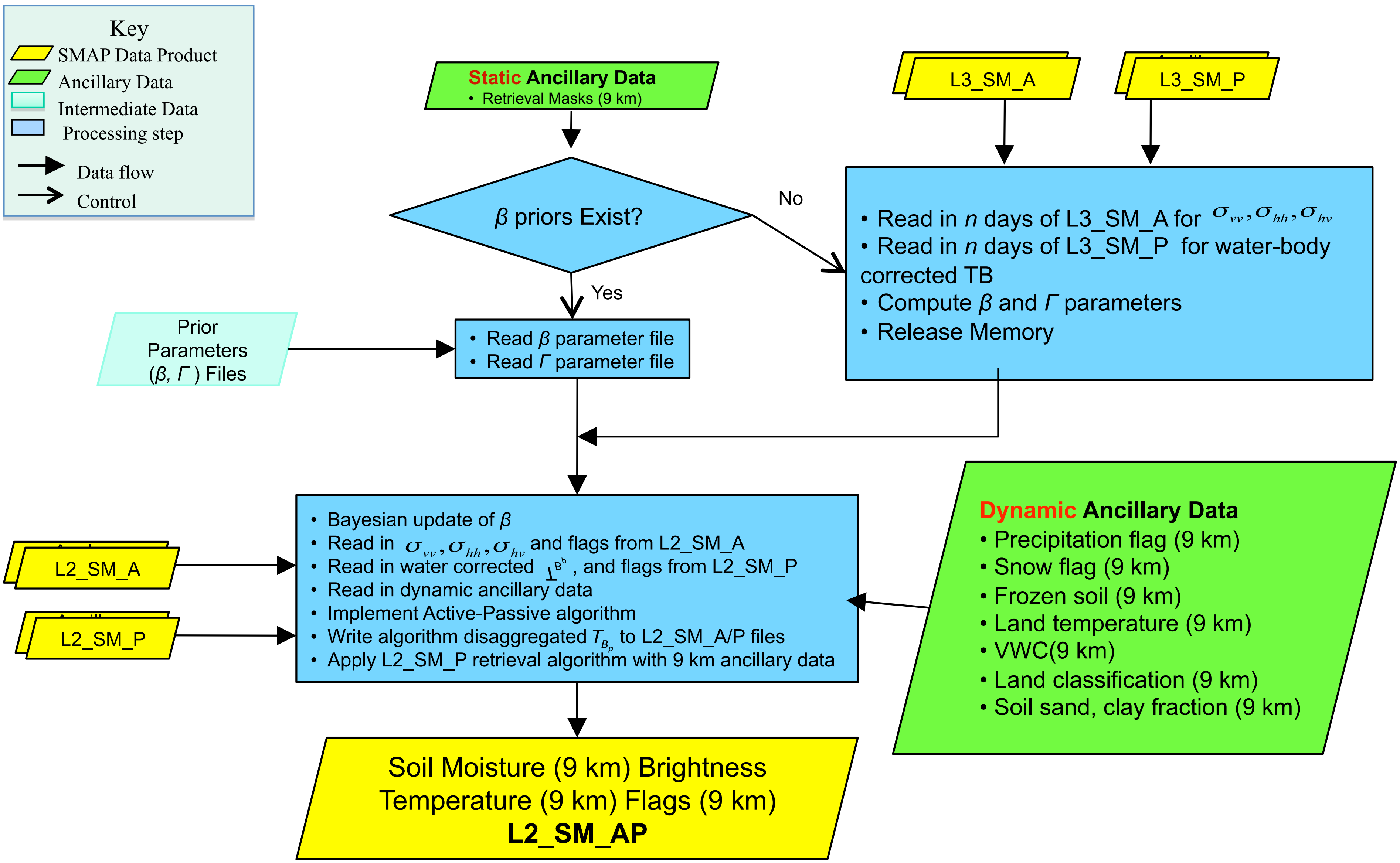
σ(vv) and σ(hv) form land fraction



Soil Moisture:
Volumetric accuracy of 0.04 [cm^3 cm^{-3}]; 9 km resolution ; Vegetation water content < 5 [kg m^{-2}]

Grid: EASE
C: Coarse (36 km)
M: Medium (9 km)
F: Fine (3 km)

L2_SM_AP: Algorithm Flow



L2_SM_AP: Key Open Issues

- Time window to estimate algorithm parameters are TBD
- Development of prior parameter database (airborne data)
- Testing of baseline and option algorithms under diverse vegetation conditions
- Assessing the impact on algorithm flow with complete set of mask and flags
- Rigorous sensitivity analysis to evaluate the impact of various ancillary data inputs on algorithm performance

L2_SM_AP: Cal/Val Requirements

- Need long time series of PALS data from diverse hydroclimatic regions (for algorithm development and parameter validation)

- Focus on regions having IGBP landcover of shrubland, savanna, grassland and mixed forest because they have greater share of global landmass

- Important measurements:

	Type or Property
Airborne	Preferably PALS (T_B, σ_0)
In-situ	Soil moisture, VWC, Surface roughness, Surface temperature
Spatial Sampling	Compatible with 9 km^2 product
Duration	Optimally one month to get time series
Frequency	At least every other day

Summary

- Will need satellite, Airborne, tower-based, in-situ data sets
- Pre-launch cal/val data are required for:

  - Parameterizing forward models for passive-only and active-only algorithms
  - Deriving disaggregation parameters for joint active/passive algorithm
  - Validation of all retrieval algorithms
  - Scaling, aggregation, disaggregation
  - Bottom-line: A common set of validation data serves to select and improve the baseline Level 2 soil moisture algorithms

BACKUP

L2_SM_P: Error Allocation

Error Allocation: Different algorithms respond differently to uncertainty of a given model/ancillary parameter. The following table lists the retrieval errors (RMSE in % VSM) of all algorithms over the full range of soil moisture and VWC conditions encountered in GloSim.

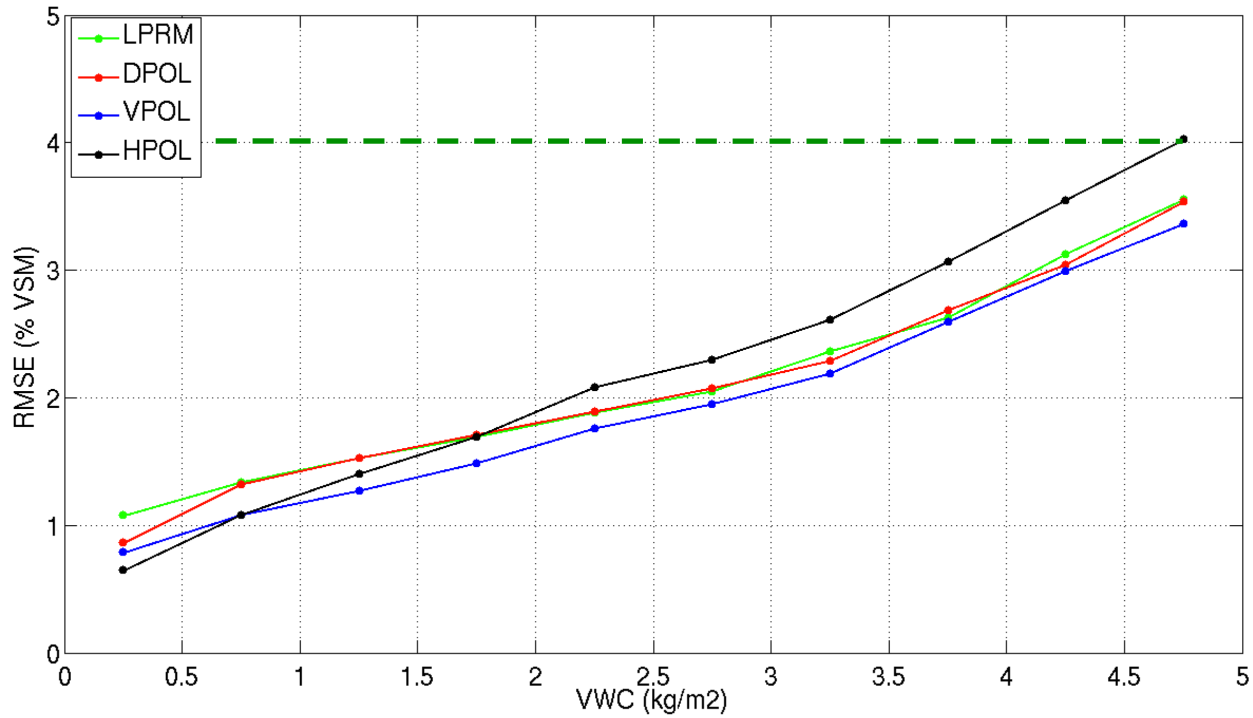
	Baseline	**Option 1**	**Option 2**	**Option 3**
Model/Ancillary Uncertainty	**Single Pol (H)**	**Single Pol (V)**	**Dual Pol**	**LPRM**
Gridding + Aggregation	0.612	0.581	0.591	0.582
5% h	0.645	0.595	0.595	0.583
5% omega	0.629	0.605	0.619	0.611
5% sandfrc	0.729	0.699	0.702	0.697
5% clayfrc	0.615	0.585	0.594	0.585
2K T5	0.871	1.000	1.120	1.200
5% VWC	0.656	0.608	——	——
10% VWC	0.717	0.647	——	——
5% watrfrc	0.612	0.582	0.591	0.582
10% watrfrc	0.612	0.582	0.591	0.583
20% watrfrc	0.614	0.584	0.593	0.583
1.13K TB	0.681	0.674	0.828	0.951

The entries marked in blue were used to produce the RMSE vs. VWC plots on the next page.

L2_SM_P Error Allocation

Simulation Conditions: One year of simulated H- and V-polarized L1B_TB's were used to retrieve soil moisture and/or vegetation opacity, using perturbed model and ancillary parameters. Static water TB correction was applied after TB gridding.

	h	omega	sandfrc	clayfrc	T5	VWC	watrfrc	TB
RMSE	5%	5%	5%	5%	2K	5%	10%	1.13K



L2_SM Single-Channel Error Budget Table *

Error Source	Est. TB Error (K)
Atmospheric Gases & Clouds **	0.15
Soil Temperature (2°C error)	1.7
Vegetation Water Content (10%)	1.6
Model Parameterization (h, ω, b, all at 5% error, classification, etc.)	1.4
Surface Heterogeneity	0.9
Total RSS of Geophysical Errors	2.87
Radiometer Precision & Calibration Stability	1.3
Total RSS Error	3.15

[* Error budget to be generated and updated for each candidate algorithm using SMAP simulations and analysis of SMOS data]

[** This error moved to L1B_TB error budget; L1B_TB are calibrated TBs at the surface]