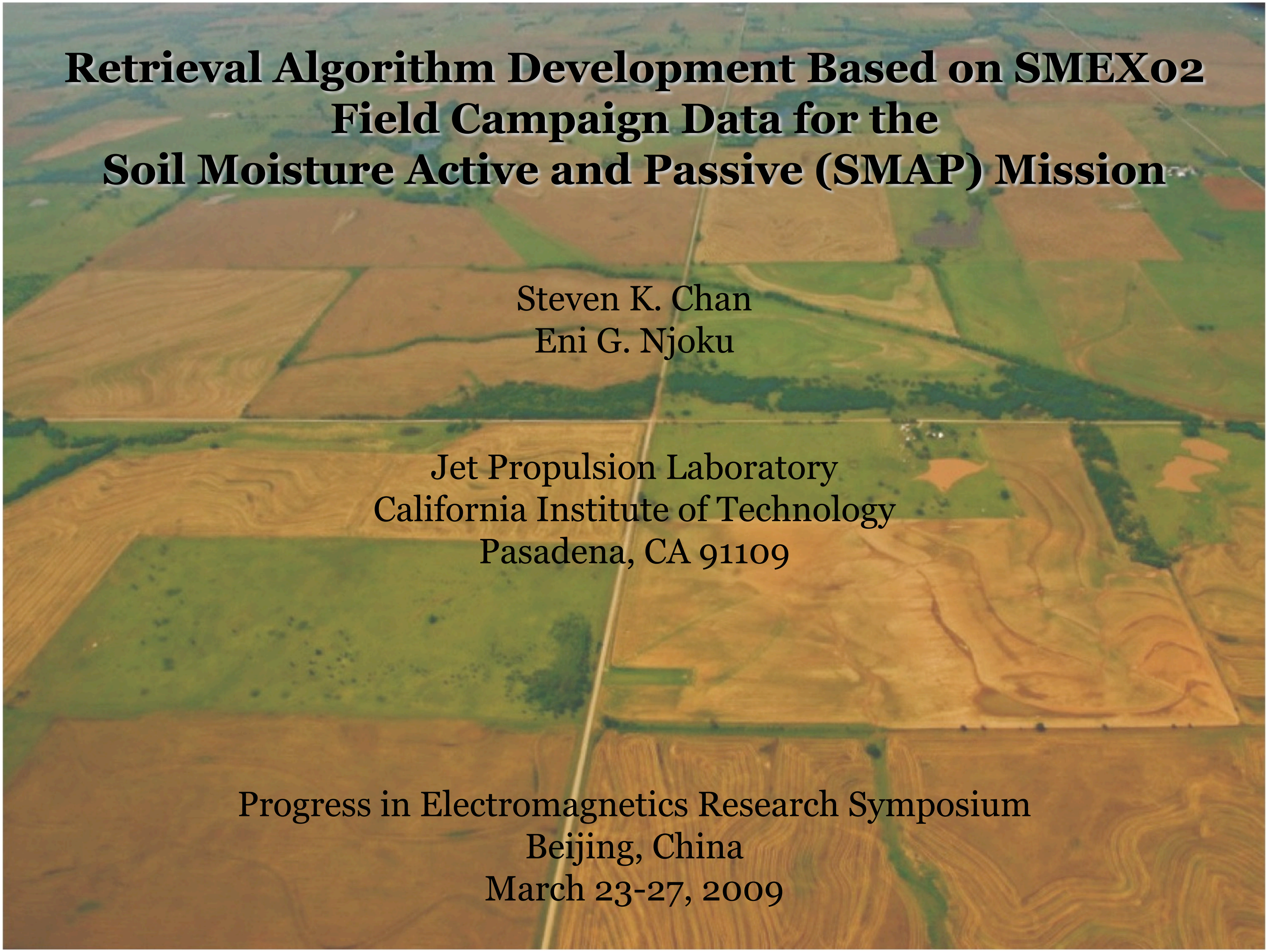


Retrieval Algorithm Development Based on SMEX02 Field Campaign Data for the Soil Moisture Active and Passive (SMAP) Mission

Steven K. Chan
Eni G. Njoku

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

Progress in Electromagnetics Research Symposium
Beijing, China
March 23-27, 2009

SMAP Mission Overview

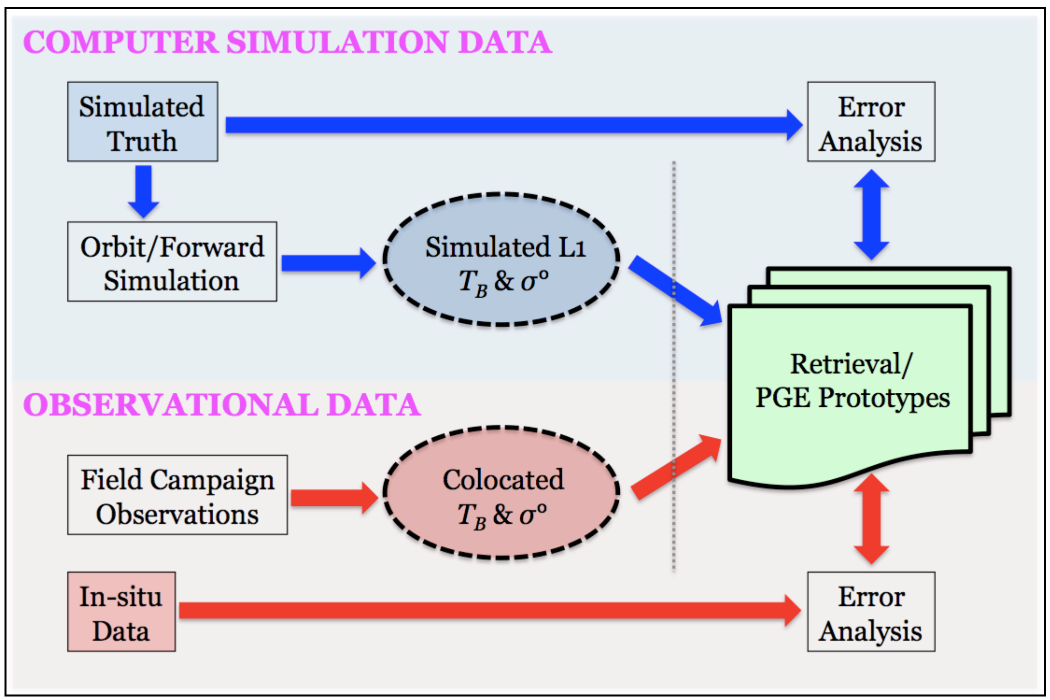
The **S**oil **M**oisture **A**ctive and **P**assive (SMAP) mission is a NASA-directed mission aiming at providing global observations of soil moisture and land surface freeze/thaw state. The resulting hydrosphere state measurements will help advance our understanding of processes that link the water, energy and carbon cycles, as well as enhance existing weather and climate forecast skills. For more information, please visit http://smap.jpl.nasa.gov.

Spaceborne L-band radar and radiometer are used to acquire complete global observations of $\sigma°$ and T_B within 3 days.

	Radar	Radiometer
Antenna	Shared 6-m rotating mesh antenna at 40° incidence	
Frequency	1.26 GHz	1.41 GHz
Orbit	670-km sun-synchronous orbit at 6am/6pm LTAN	
Spatial resolution	1-3 km (SAR)	40 km
Primary data products	$\sigma°$, F/T state, soil moisture	T_B, Soil moisture
Expected launch	2013	

SMAP Testbed

Overview: In support of SMAP mission design, a set of software tools is being developed at JPL. Collectively known as SMAP Science Algorithm Testbed (SAT), it provides a common computing environment where SMAP L1 and L3 algorithms will be coded, tested, and eventually ported to generate operational SMAP data products. Currently, SAT development aims at mimicking the SMAP end-to-end processing flow from observations to data product generation.



Algorithm Testing

Rationale: SMAP geophysical data products are derived from retrieval algorithms. To satisfy mission success criteria, a given algorithm must meet a certain retrieval accuracy requirement. For soil moisture retrieval at 40 km based on SMAP radiometer, the accuracy requirement is **4% vol.**

Approach: In this presentation, we describe the performance of two versions of single-channel retrieval algorithms via Monte Carlo simulations on SAT, and compare their retrieval error budgets against field campaign data collected during the **S**oil **M**oisture **Ex**periment **2002** (SMEX02) in Iowa, USA. The two versions of single-channel retrieval algorithms subject to testing are based on microwave emission modeling of a two-layered land-vegetation medium at horizontal and vertical polarizations:

H-pol: $$T_{B_h} = T_s[(1 - r_{s_h})e^{-\tau} + (1 - \omega_h)(1 - e^{-\tau})(1 + r_{s_h}e^{-\tau})]$$

V-pol: $$T_{B_v} = T_s[(1 - r_{s_v})e^{-\tau} + (1 - \omega_v)(1 - e^{-\tau})(1 + r_{s_v}e^{-\tau})]$$

T_B Uncertainty Analysis

In reality, T_B measurements, model parameters, and ancillary data needed to retrieve soil moisture all come with various sources of noise and uncertainty. The combined impact on T_B measurements by these perturbation sources within a nonlinear system is most conveniently analyzed using Monte Carlo simulations.

Assuming Gaussian (need not be so, though) random error distributions, we perturb each parameter simultaneously and in each run compute T_B based on the perturbed parameters (red and blue below). We then repeat this procedure many times to obtain the probability density function of T_B.

$$T_{B_v} = T_s[(1 - r_{s_v})e^{-\tau} + (1 - \omega_v)(1 - e^{-\tau})(1 + r_{s_v}e^{-\tau})]$$

where

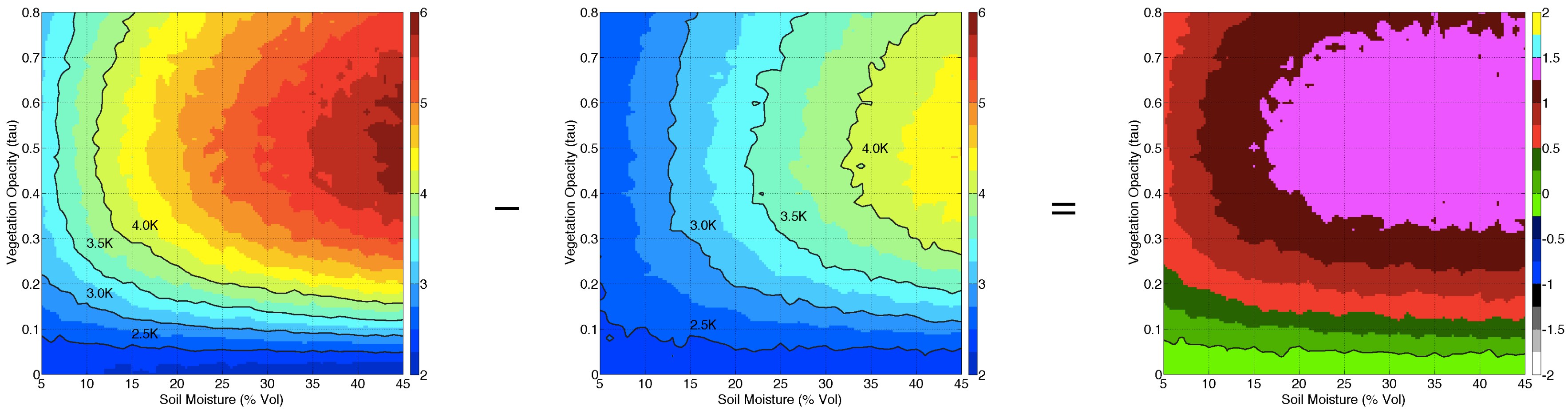
$$r_{s_v} = [(1 - Q)r_{\circ_v} + Q r_{\circ_h}]e^{-h} \;\; ; \quad \tau = b \; \mathrm{VWC} / \cos\theta$$

$$r_{\circ_v} = \left|\frac{\varepsilon_r \cos\theta - \sqrt{\varepsilon_r - \sin^2\theta}}{\varepsilon_r \cos\theta + \sqrt{\varepsilon_r - \sin^2\theta}}\right|^2 \; ; \quad r_{\circ_h} = \left|\frac{\cos\theta - \sqrt{\varepsilon_r - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon_r - \sin^2\theta}}\right|^2$$

T_B Uncertainty Analysis

Using $h \in N(0.1, 5\%)$, $\omega \in N(0.05, 5\%)$, $Q \in N(0.05, 5\%)$, $\tau \in N(\tau^*, 10\%)$, $T_s \in N(25°\mathrm{C}, 2°\mathrm{C})$, and $T_B \in N(T_B^*, 1.5\mathrm{K})$

RMSE (H) RMSE (V) RMSE (H-V)

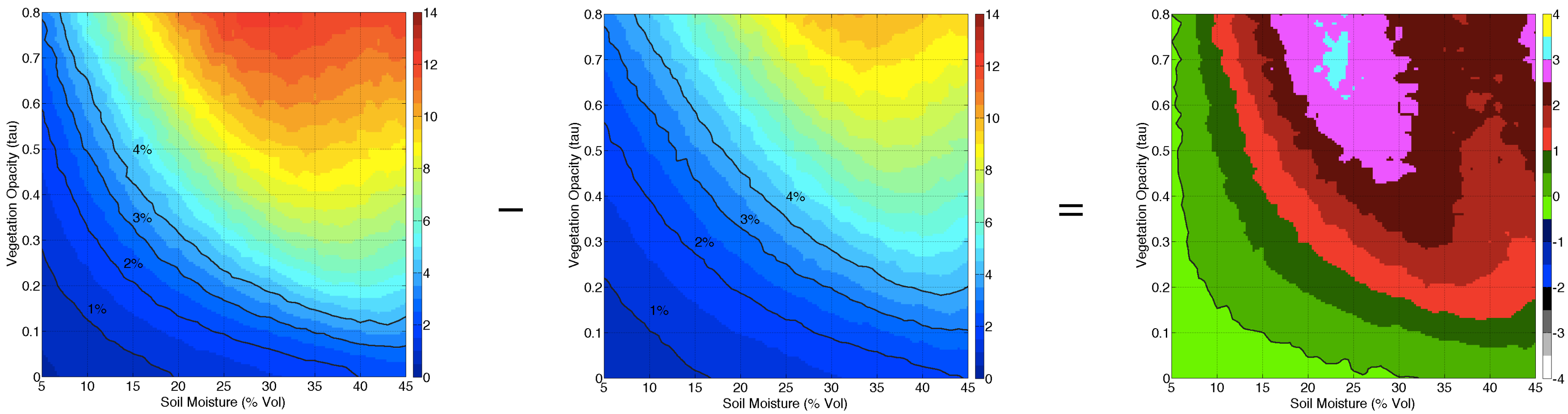


(1) At low vegetation (e.g. τ < 0.1), H is more optimal than V in terms of lower RMSE across the whole range of soil moisture (the green region below the solid line in the H-V plot). **(2)** At moderate-to-high vegetation (e.g. τ > 0.1), V becomes more optimal than H. T_B uncertainty grows as soil moisture increases. **(3)** At any given T_B uncertainty level, V covers a larger area (hence wider range of m_v and τ variability) than H does.

Retrieval Error Budget Analysis

Based on the T_B uncertainty in previous chart, how much retrieval error does it translate to?

RMSE (H) RMSE (V) RMSE (H-V)



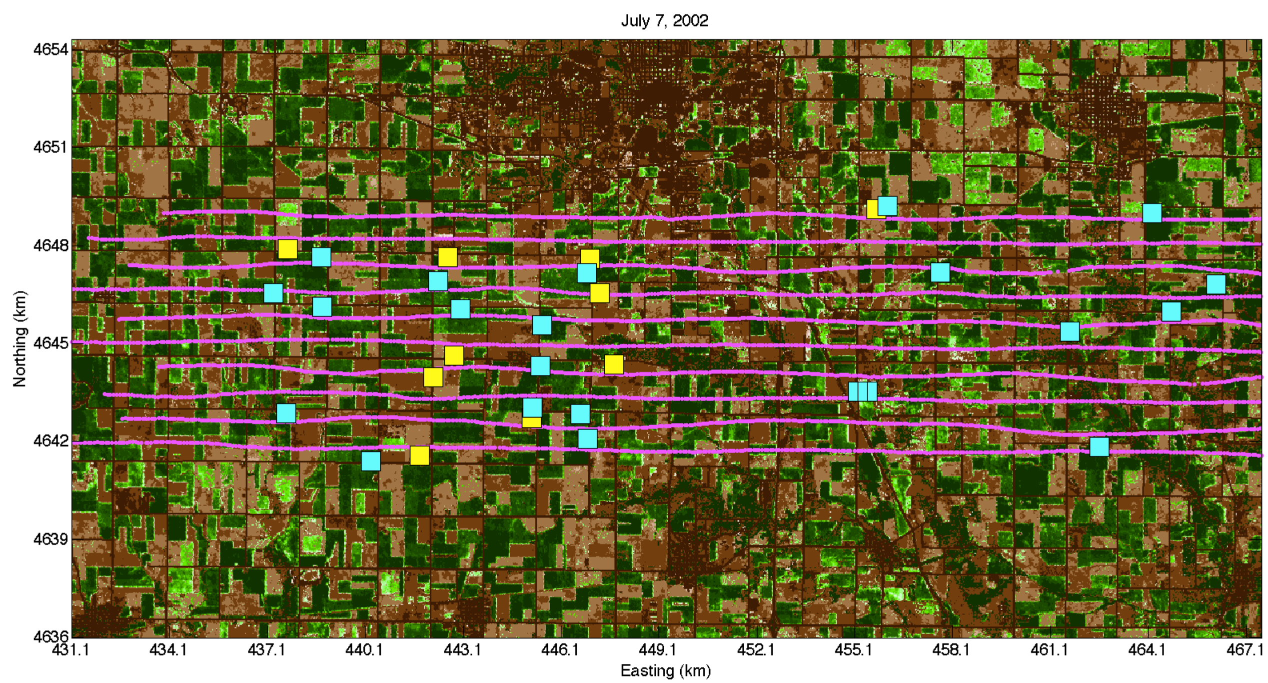
(1) For a given retrieval error budget level (say, 4%), V in general covers a larger area (hence wider range of m_v and τ variability) than H does. **(2)** According to the H-V plot, V has lower RMSE than H except over a small region (the green region below the solid line) defined by limited m_v and τ variability. **(3)** V's optimality over H becomes more pronounced at high m_v and/or τ conditions.

Do Experimental Field Campaign Data Support the Previous Simulation Results?

SMEX02 Field Campaign

Overview: In support of Aqua/AMSR-E validation, measurements were collected throughout the state of Iowa in June-July 2002. In particular, the data acquired over the Walnut Creek Watershed are especially useful for SMAP algorithm development due to the availability of ground measurements and L-band airborne active and passive observations. This study primarily focuses on comparison between ground measurements and PALS T_B observations.

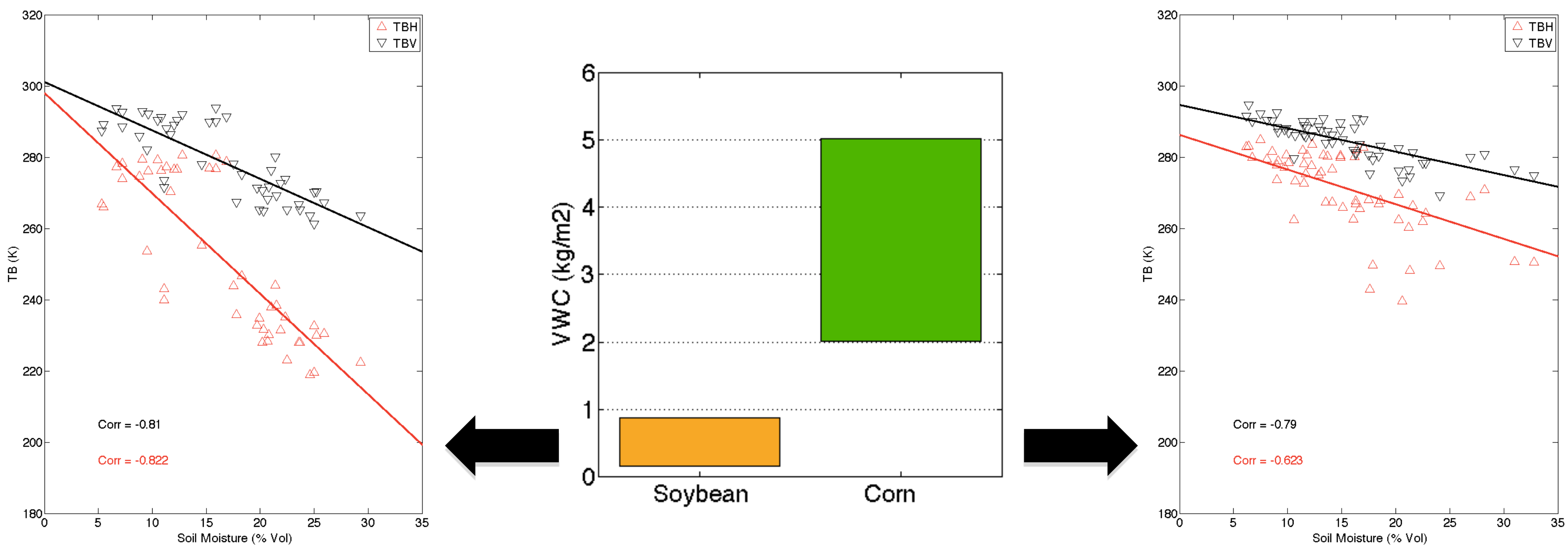
Ground data were collected at corn and soybean fields while PALS flew over the sites simultaneously to acquire L-band passive and active observations. ➔



PALS Passive Observations: Soil Moisture

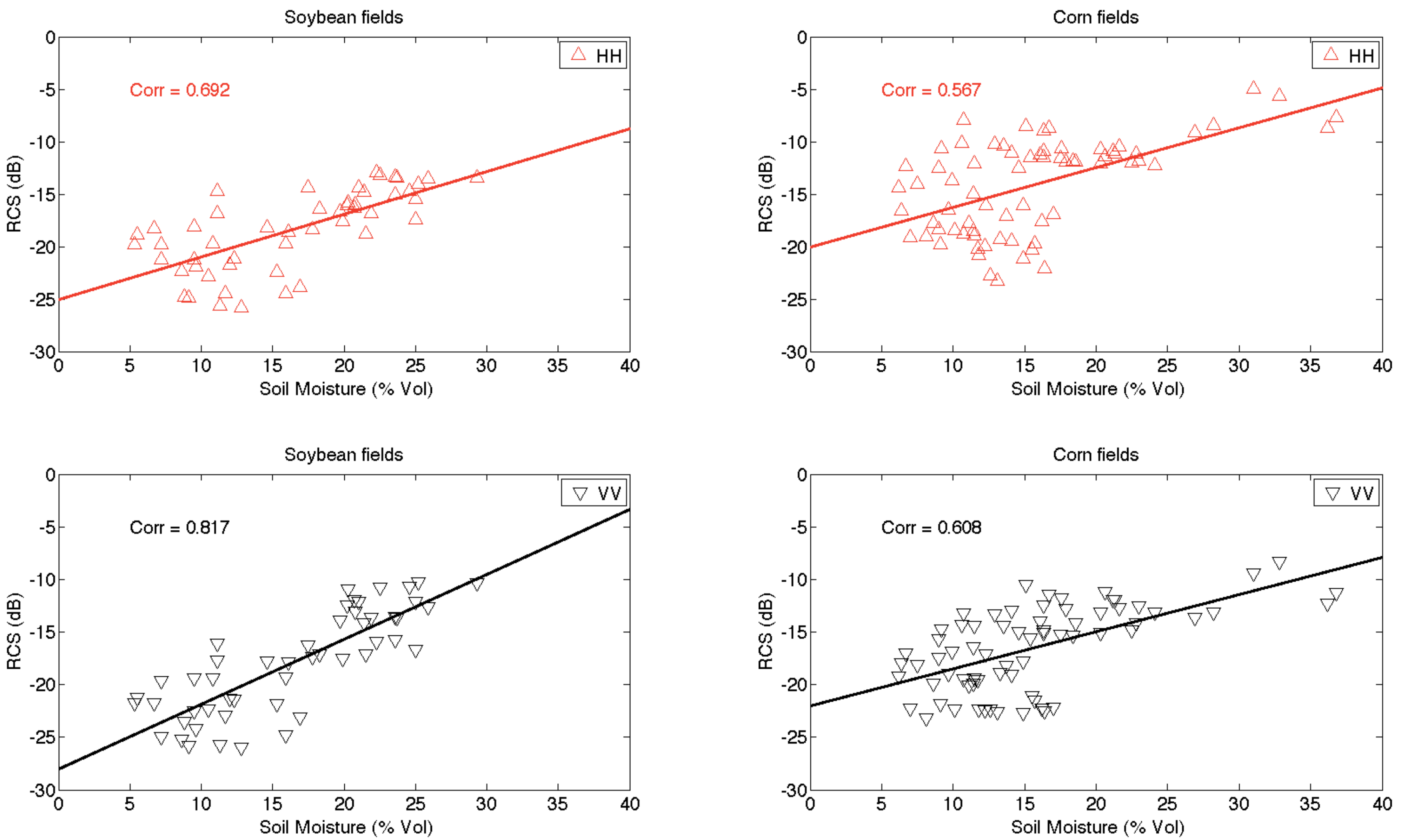
It was relatively dry at the beginning of the campaign. The dry period was followed by a period of sporadic rainfall, resulting in widespread dampness.

To assess retrieval robustness over different amounts of vegetation, data were collected and analyzed over sites with low-to-moderate vegetation (soybean and corn). Observations: **(1)** Higher T_B dynamic range and better correlation over soybean fields. **(2)** V has better correlation, especially at high vegetation.



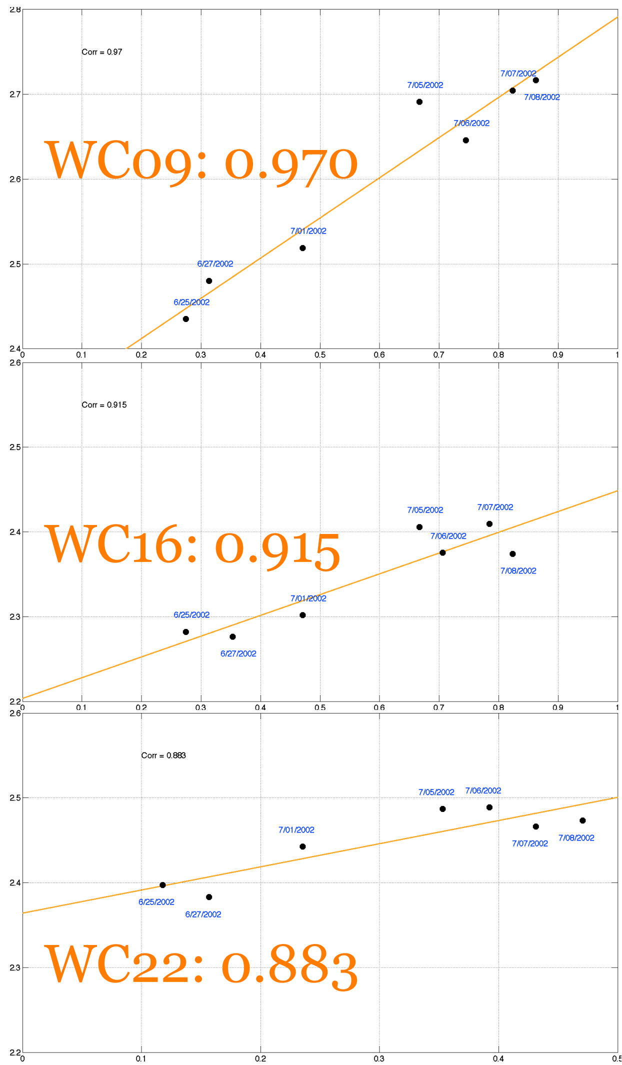
PALS Active Observations: Soil Moisture

Observations: **(1)** Better correlation over soybean fields. **(2)** V has better correlation for both crops. **(3)** RVI shows excellent correlation with VWC.



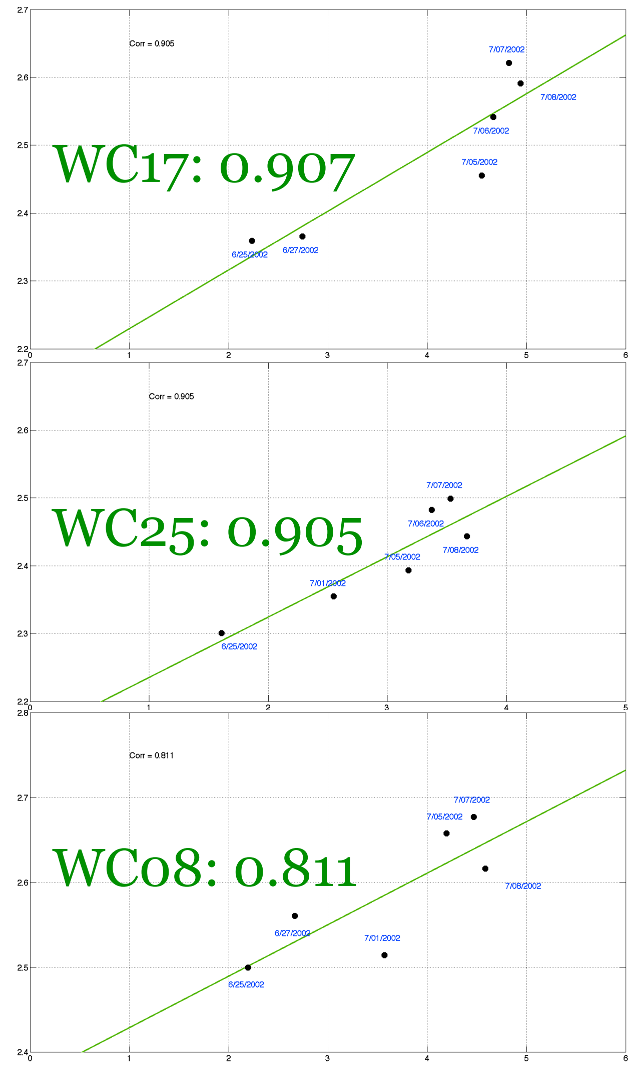
PALS Active Observations: VWC

Soybean fields



VWC

Corn fields



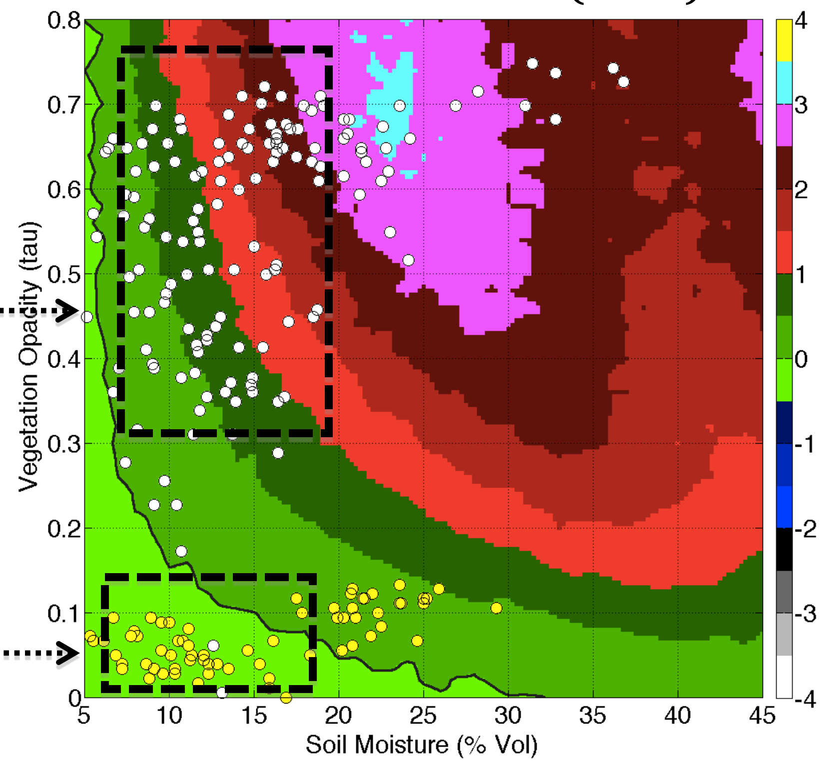
VWC

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

Algorithm Optimality According to Simulation Results

White dots represent (m_v,τ) conditions observed over corn fields during the campaign, yellow dots over soybean fields. The underlying contour plot indicates the difference in retrieval error between the H- and V-pol algorithms. V-pol algorithm produces more accurate retrieval under (m_v,τ) conditions where the difference is positive.

Retrieval RMSE (H-V)



Test 1: Does V-pol algorithm perform better over corn fields?

Test 2: Does H-pol algorithm perform better over soybean fields?

Testing Procedure

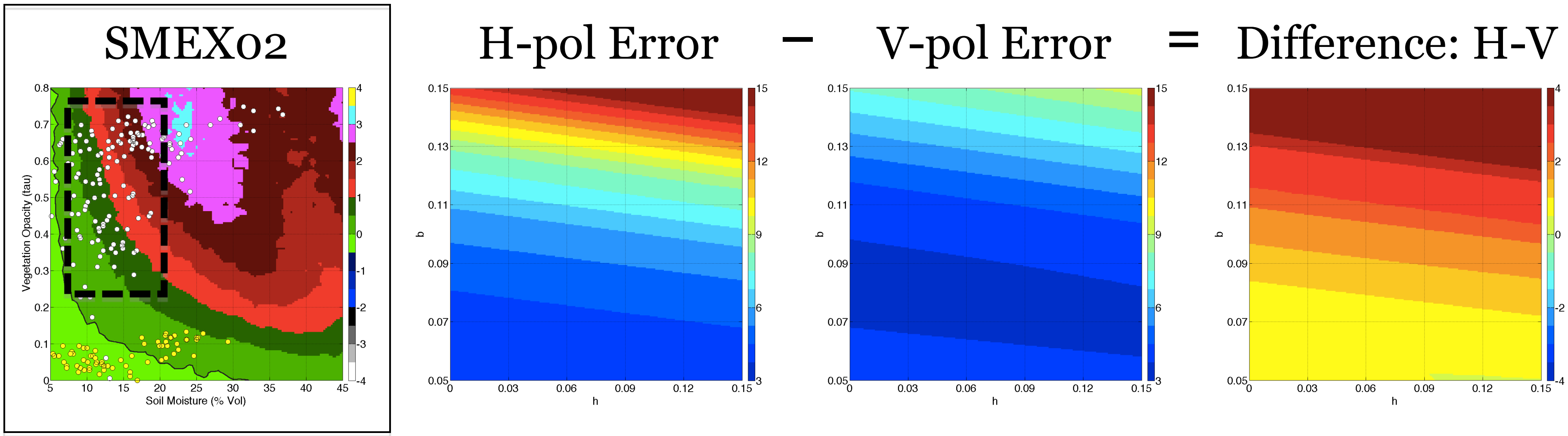
Testing Basis: Under a given set of geophysical conditions (crop types, soil moisture, VWC, surface temperature, surface roughness, ... etc), an algorithm delivers better retrieval than another if it provides a more accurate description of the relationship between the inputs (e.g. ground measurements) and outputs (e.g. PALS T_B observations).

Test 1: We use ground measurements and PALS T_B observations collected over corn fields to first calibrate the emission model. Then, we apply both H- and V-pol algorithms and compare their resulting retrieval errors.

Test 2: The procedure is similar to that for Test 1, except that the data collected over soybean fields are now used for initial model calibration.

Test 1: V-pol vs. H-pol over Corn Fields

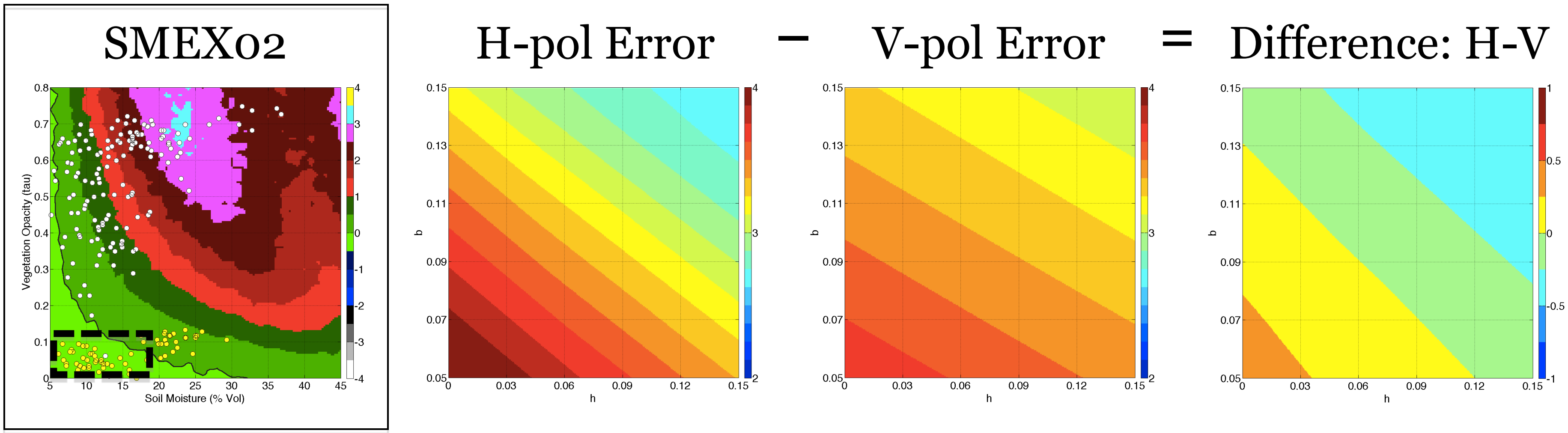
We feed different combinations of (*h*,*b*) into the emission model and for each combination we compare retrieved soil moisture with ground truth. The resulting error is plotted below as a function of *h* and *b* for both H- and V-pol algorithms.



(1) Whatever the final calibrating (*h*,*b*) combination may be, V results in lower retrieval error than H by an average of ~2% VSM across the board. **(2)** H-pol error looks like a down-shifted version of V-pol error, suggesting polarization dependence of *b*? **(3)** V performs better than H over corn fields.

Test 2: V-pol vs. H-pol over Soybean Fields

We feed different combinations of (*h*,*b*) into the emission model and for each combination we compare retrieved soil moisture with ground truth. The resulting error is plotted below as a function of *h* and *b* for both H- and V-pol algorithms.



(1) Conclusion not so clear-cut: some (*h*,*b*)'s make V retrieve better than H but some don't. **(2)** In either case, the difference in retrieval accuracy between V and H is minimal: only ±0.1% VSM. **(3)** Bottom line: both V and H work equally well over soybean fields, neither is superior to the other by any significant margin.

Summary

SMAP Science Algorithm Testbed (SAT) is an important component of the mission. It provides a common computing environment for performing science trades, algorithm testing, and future operational infrastructure development.

This study illustrates a scheme in which algorithm testing can be performed on SAT using computer simulation data (via Monte Carlo simulations) and observational data (SMEX02 dataset).

The unique objectives of each field campaign must be well understood before the experimental data can be used to properly test algorithms for a particular metric (e.g. impacts on retrieval performance by vegetation, pixel-to-footprint scaling, azimuthal dependence, and sub-footprint surface heterogeneity).