

# **Soil Moisture Active Passive (SMAP) Project Calibration and Validation for the L3\_FT\_P and L3\_FT\_P\_E Data Products (Version 1)**

## **Citation:**

Derksen, Chris<sup>1</sup>, Xiaolan Xu<sup>2</sup>, R. Scott Dunbar<sup>2</sup>, Andreas Colliander<sup>2</sup>, Youngwook Kim<sup>3</sup>, John Kimball<sup>3</sup>

<sup>1</sup>Environment and Climate Change Canada (ECCC), Toronto ON M3H 5T4, CA

<sup>2</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA

<sup>3</sup>University of Montana, Missoula, MT

Paper copies of this document may not be current and should not be relied on for official purposes. The current version is in the Product Data Management System (PDMS): <https://pdms.jpl.nasa.gov/>

**December 5, 2016**

**JPL D-56296**

National Aeronautics and Space Administration



Jet Propulsion Laboratory  
4800 Oak Grove Drive  
Pasadena, California 91109-8099  
California Institute of Technology

Copyright 2016 California Institute of Technology. U.S. Government sponsorship acknowledged.

# TABLE OF CONTENTS

<b>1</b>	<b>EXECUTIVE SUMMARY.....</b>	<b>3</b>
<b>2</b>	<b>OBJECTIVES OF CAL/VAL.....</b>	<b>4</b>
<b>3</b>	<b>L3_FT_P ALGORITHM.....</b>	<b>6</b>
<b>4</b>	<b>L3_FT_P VALIDATION METHODOLOGY .....</b>	<b>8</b>
<b>5</b>	<b>Final RELEASE PROCESS .....</b>	<b>10</b>
5.1	Freeze and Thaw References.....	10
5.2	False Freeze Mitigation .....	12
5.3	Validated Release Testing .....	13
<b>6</b>	<b>ASSESSMENTS .....</b>	<b>14</b>
6.1	Large Scale Patterns and Features .....	14
6.2	Core Validation Sites (CVS).....	16
6.2.1	L3_FT_P .....	18
6.2.2	L3_FT_P_E .....	27
6.2.3	Core Site Summary.....	29
6.3	Sparse Networks.....	31
6.4	Air Temperature Validation.....	37
6.5	Satellite Time Series Inter-comparison .....	41
6.6	Summary .....	43
<b>7</b>	<b>OUTLOOK .....</b>	<b>44</b>
<b>8</b>	<b>ACKNOWLEDGEMENTS.....</b>	<b>45</b>
<b>9</b>	<b>REFERENCES .....</b>	<b>46</b>

# 1 EXECUTIVE SUMMARY

During the post-launch Cal/Val Phase of SMAP there are two objectives for each science product team: 1) calibrate, verify, and improve the performance of the science algorithms, and 2) validate accuracies of the science data products as specified in the L1 science requirements according to the Cal/Val timeline. This report provides analysis and assessment of the validated SMAP Level 3 Landscape Freeze/Thaw (L3\_FT\_P) and enhanced resolution freeze/thaw (L3\_FT\_P\_E) products. The SMAP Level 3 Landscape Freeze/Thaw products are daily composites of half-orbit freeze/thaw retrievals derived from SMAP radiometer (L1C\_TB and L1C\_TB\_E) products.

Assessment methodologies utilized include comparisons of SMAP freeze/thaw retrievals with *in situ* observations from core validation sites (CVS) and sparse networks, and inter-comparison with products from the NASA Aquarius and JAXA Advanced Microwave Scanning Radiometer-2 (AMSR-2) missions. These analyses meet the criteria established by the Committee on Earth Observing Satellites (CEOS) Stage 2 validation.

The SMAP baseline science mission objective for freeze/thaw is to provide binary estimates of landscape freeze/thaw state for the region north of 45° N latitude, which includes the boreal forest zone, with a spatial classification accuracy of 80% and 2-day average intervals separated by AM and PM overpasses. Evaluation of SMAP measurements and FT retrievals from 1 April 2015 indicates the baseline classification accuracy (80%) is met for most metrics and evaluation datasets, and the minimum mission requirement of 70% is met in all cases. The nominal spatial resolution was relaxed from 3 km to 36 km due to the change from SMAP radar to radiometer inputs, but this evaluation shows the enhanced resolution 9 km FT\_P\_E product provides very similar accuracy to what is obtained at 36 km. The primary sources of uncertainty are:

- (1) the SMAP spring thaw signal tends to lead the soil temperature based FT reference flag, which is attributed to the influence of wet snow cover on the radiometer signal and subsequent delay in soil thawing until snow melt is at an advance state.
- (2) long periods in the early winter, particularly at boreal forest sites, when flag agreement with Tsoil measurements are weak because the soil at depth remains unfrozen due to the insulating effect of snow cover. This is despite air temperatures being consistently below zero continuously for weeks and snow lying on the surface. The landscape is effectively 'frozen' even if this is not captured by the Tsoil measurements. It is therefore important for users to understand that the satellite FT retrievals represent an integrated landscape state, not simply the near-surface soil layer.
- (3) obviously false freeze retrievals during summer months, particularly near the southern boundary of the FT domain, caused by small differences in the freeze and thaw reference values. These will largely be mitigated by the application of daily 'never frozen' and 'never thawed' masks from Advanced Microwave Scanning Radiometer (AMSR-E) FT climatologies, which will be applied in the next release of the FT products.

## 2 OBJECTIVES OF CAL/VAL

During the post-launch Cal/Val (Calibration/Validation) Phase of SMAP there are two objectives for each science product team:

- Calibrate, verify, and improve the performance of the science algorithms, and
- Validate accuracies of the science data products as specified in L1 science requirements according to the Cal/Val timeline.

The process is illustrated in **Figure 2.1**. In this Assessment Report, the progress of the L3 Freeze/Thaw Team in addressing these objectives for the release of SMAP radiometer derived freeze/thaw products at two spatial resolutions is described. The L3\_FT\_P product utilizes standard SMAP L1C\_TB inputs at 36 km resolution; the L3\_FT\_P\_E product utilizes L1C\_TB\_E inputs at an enhanced resolution of 9 km. The approaches and procedures follow those described in the SMAP Cal/Val Plan [1]; full details on the Level 3 radiometer derived freeze/thaw algorithm and products are provided in the Algorithm Theoretical Basis Document [2].

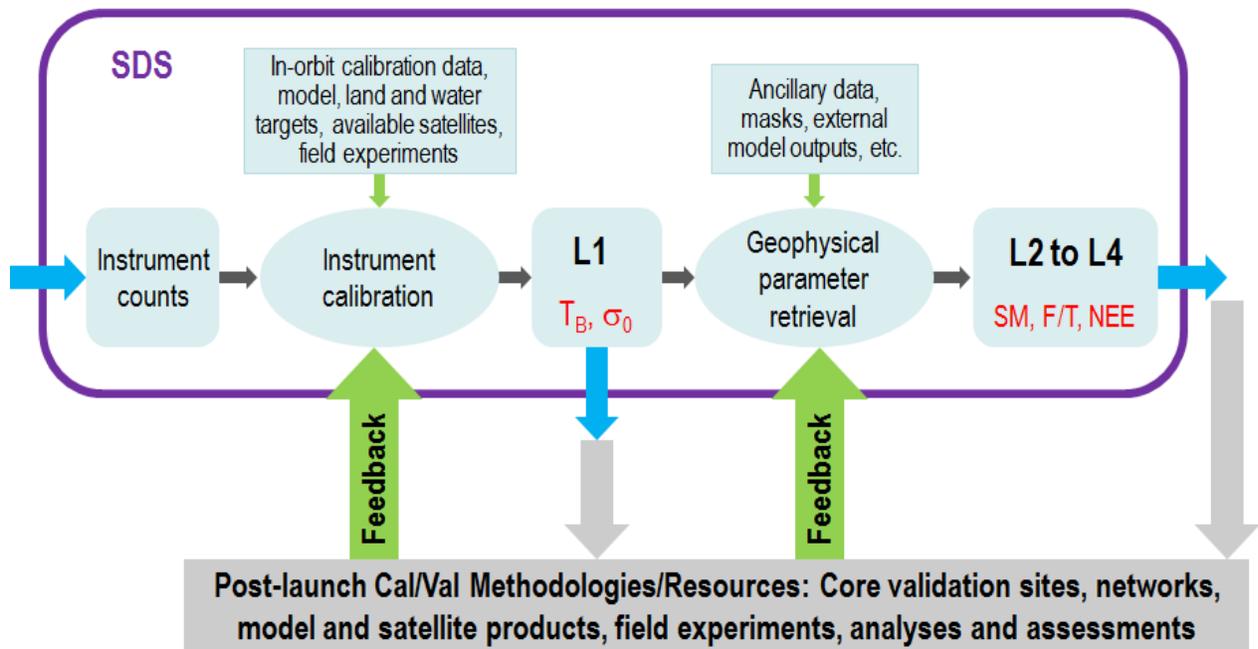


Figure 2.1. Overview of the SMAP Cal/Val Process.

SMAP established a unified definition base in order to effectively address the mission requirements. These are documented in the SMAP Handbook/ Science Terms and Definitions [3], where Calibration and Validation are defined as follows:

- *Calibration*: The set of operations that establish, under specified conditions, the relationship between sets of values or quantities indicated by a measuring instrument or measuring system and the corresponding values realized by standards.
- *Validation*: The process of assessing by independent means the quality of the data products derived from the system outputs.

The SMAP Mission L1 Requirements Document [4] states: the baseline science mission shall provide estimates of surface binary freeze/thaw state for the region north of 45° N latitude, which includes the boreal forest zone, with a spatial classification accuracy of 80% at 3 km spatial resolution and 2-day average intervals. Following the loss of the SMAP radar in July 2016, the Freeze/Thaw Team adopted the same domain, retrieval accuracy, and latency requirement for the radiometer derived L3\_FT\_P product as for the radar derived L3\_FT\_A. The spatial resolution baseline requirement could not be met with radiometer inputs, although the 9 km resolution L3\_FT\_P\_E does meet the minimum SMAP mission requirement of 10 km. The minimum mission requirement of 70% classification accuracy (versus baseline of 80%) also applies to both resolutions of the FT\_P products.

In order to ensure the public's timely access to SMAP data, before releasing validated products the mission is required to release beta-quality products. The maturity of the products in the beta release is defined as follows:

- Early release is used to gain familiarity with data formats.
- Intended as a testbed to discover and correct errors.
- Minimally validated and still may contain significant errors.
- General research community is encouraged to participate in the quality assessment and validation, but need to be aware that product validation and quality assessment are ongoing.
- Data may be used in publications as long as the fact that the data are beta quality is indicated by the authors. Drawing quantitative scientific conclusions is discouraged. Users are urged to contact science team representatives prior to use of the data in publications, and to recommend members of the instrument teams as reviewers.
- The estimated uncertainties will be documented.
- May be replaced in the archive when an upgraded (provisional or validated) product becomes available.

In assessing the maturity of the L3\_FT\_A product, the L3\_FT\_A team also considered the guidance provided by the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) [5]:

- Stage 1: Product accuracy is assessed from a small (typically < 30) set of locations and time periods by comparison with *in situ* or other suitable reference data.
- Stage 2: 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: 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 results for stage 3 are systematically updated when new product versions are released and as the time-series expands.

With this release, the L3\_FT\_P team has completed Stage 2 (see also [6]).

### 3 L3\_FT\_P ALGORITHM

**Figure 3.1** shows the data sets and processing chain associated with SMAP freeze/thaw algorithm implementation and product generation, including input and output data. The L3\_FT\_P product consists of daily composite landscape freeze/thaw state derived from the AM (descending) and PM (ascending) overpass radiometer data (L1C\_TB half-orbits) north of 45°N. The L3\_FT\_P product is gridded and provided on a 36 km Equal Area Scalable Earth grid version 2 (EASE-grid) in both global and north polar projections. The same data flow applies to the enhanced resolution product (L3\_FT\_P\_E) with L1C\_TB at 36 km replaced by L1C\_TB\_E at 9 km resolution. The L3\_FT\_P(E) algorithm is applied to L1C\_TB(E) granules for unmasked land regions. The resulting intermediate freeze/thaw products (Figure 2) serve two purposes: (1) these data are assembled into global daily composites in production of the L3\_FT\_P product, and (2) the freeze/thaw product derived from global AM L1C\_TB granules provide the binary freeze/thaw state flag supporting generation of the L2 and L3 soil moisture passive products.

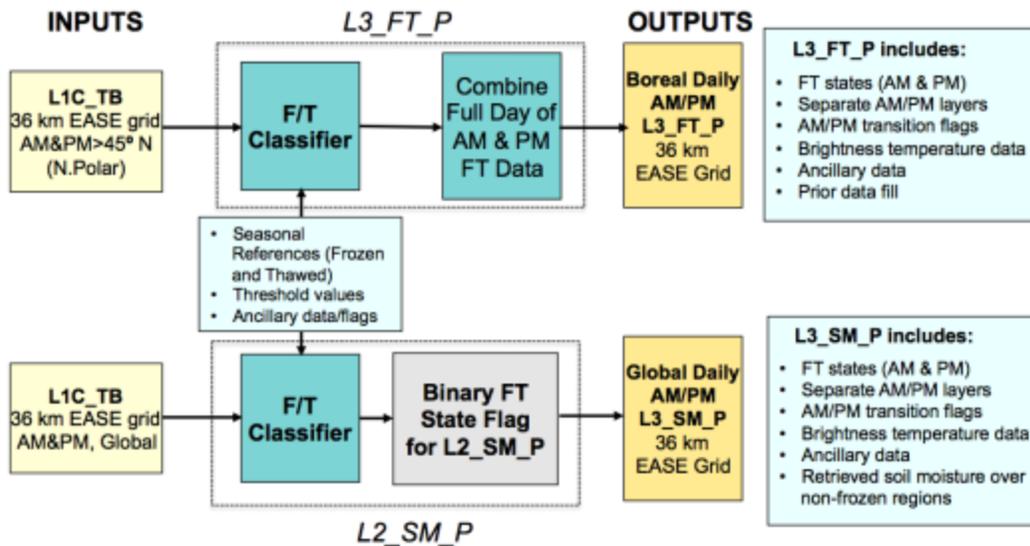


Figure 3.1. Processing sequence for generation of the L3\_FT\_P product and the binary freeze/thaw state flag for use in L2\_SM\_P.

The SMAP L3\_FT\_P freeze/thaw algorithm is based on a seasonal threshold approach. While other freeze/thaw algorithmic approaches are possible (for example, moving window; temporal edge detection) these techniques do not fulfill the SMAP data latency requirement, and so are not discussed further in this document.

The seasonal threshold (baseline) algorithm examines the time series progression of the remote sensing signature relative to signatures acquired during seasonal reference frozen and thawed states. The algorithm is applied to the normalized polarization ratio (NPR) of SMAP radiometer measurements:

$$NPR = \frac{TBV - TBH}{TBV + TBH} \quad (4)$$

Decreases and increases in NPR are associated with landscape freezing and thawing transitions, respectively. The decrease in NPR under frozen conditions is a result of small increases in the V-pol brightness temperature combined with larger increases at H-pol [7][8][9]. Various studies have shown the NPR to be preferred over other approaches as it minimizes sensitivity to physical temperature and outperforms other L-band brightness temperature based approaches [8][10].

A seasonal scale factor  $\Delta(t)$  is defined for an observation acquired at time  $t$  as:

$$\Delta t = \frac{NPR(t) - NPR(fr)}{NPR(th) - NPR(fr)} \quad (5)$$

where  $NPR(t)$  is the normalized polarization ratio calculated at time  $t$ , for which a freeze/thaw classification is sought, and  $NPR(th)$  and  $NPR(fr)$  are normalized polarization ratios corresponding to the frozen and thawed reference states, respectively.

A threshold level  $T$  is then defined such that:

$$\begin{aligned} \Delta(t) &> T \\ \Delta(t) &\leq T \end{aligned} \quad (6)$$

defines the thawed and frozen landscape states, respectively. This series of equations (4-6) are run on a grid cell-by-cell basis for unmasked portions of the FT domain. The output from Equation (6) is a dimensionless binary state variable designating either frozen or thawed conditions for each unmasked grid cell. The parameter  $T$  is fixed at 0.5 across the entire FT domain; optimization are presently under development.

## 4 L3\_FT\_P VALIDATION METHODOLOGY

The L3\_FT\_P freeze/thaw product provides estimates of land surface freeze/thaw state expressed as a categorical (frozen, thawed, or [inverse] transitional) condition. The SMAP Level 1 baseline mission requirement is that the L3 freeze/thaw product will be provided for land areas north of 45 degrees north latitude with a mean spatial classification accuracy of 80% at 3 km spatial resolution and 2-day average temporal sampling. The accuracy of the L3 product will be determined by comparison of the SMAP freeze/thaw retrievals with in situ measurements from sites within northern latitude ( $\geq 45^\circ\text{N}$ ) land areas (see Section 7.2).

The in situ validation data include all core validation sites, selected sites from the sparse networks using criteria based on site representativeness (uniform and representative terrain and land cover), and a global air temperature measurement network. The validation is based on reference freeze/thaw flags derived from co-located air temperature and soil temperature corresponding to the local time of the descending and ascending satellite overpasses.

The computation of the classification accuracy proceeds as follows: Let  $s_{AM/PM}(i,t) = 1$  if the L3\_FT\_A product at grid cell  $i$  (on the SMAP 3 km EASE grid) and time  $t$  indicates frozen conditions for AM (descending) or PM (ascending) overpass, respectively, and let  $s_{AM/PM}(i,t) = 0$  if the L3\_FT\_A product indicates thawed conditions for AM or PM overpass, respectively. Likewise, let  $v_{AM/PM}(i,t) = 1$  if the corresponding reference flag indicates frozen conditions at the AM or PM overpass, and  $v(i,t) = 0$  for thawed conditions at the AM or PM overpass. Next, the error flag  $\delta$  is set by comparing the SMAP product to the validating observations:

$$d_{AM/PM}(i,t) = \begin{cases} 0 & \text{if } s_{AM/PM}(i,t) = v_{AM/PM}(i,t) \\ 1 & \text{if } s_{AM/PM}(i,t) \neq v_{AM/PM}(i,t) \end{cases} \quad (5.1)$$

Note that a single L3\_FT\_A flag is produced each day, but is derived from separate descending (AM) and ascending (PM) overpasses. The L3\_FT\_A flags will therefore be separated back into binary freeze/thaw classes for the AM and PM orbits, producing two retrieval match-ups each day.

The mission Level 1 requirement will be satisfied if (for both AM and PM overpasses together):

$$1 - \frac{\sum_{i=1}^{N_i} \sum_{t=1}^{N_t(i)} d(i,t)}{\sum_{i=1}^{N_i} N_t(i)} \leq 0.8 \quad (5.2)$$

Equation 5.1 will be solved daily, to provide instantaneous determinations of freeze/thaw spatial accuracy, using the available reference sites. The baseline mission requirement of 80% accuracy will be assessed cumulatively (in a running manner with each new day of data added to the previous days). Assessment with multiple reference FT flags (air temperature, soil temperature) allows algorithm performance metrics to be computed for various surface conditions (i.e. wet snow versus dry snow), and assist in determining the landscape components driving the radar response. Retrieval performance is also summarized monthly to reduce sensitivity to prolonged periods of consistent frozen and thawed states in the winter and summer, respectively. In addition to overall flag agreement, counts of freeze and thaw omission and commission errors ('false freeze' retrievals vs. 'false thaw' retrievals) are also tabulated.

Comparisons between SMAP L3\_FT\_P, other satellite derived FT products from Aquarius [6] and AMSR-2 [7], and soil temperature fields from NASA GMAO are also performed in order to evaluate spatial agreement, and changes in continental-scale frozen area over time.

## 5 FINAL RELEASE PROCESS

This section describes refinement of the L3\_FT\_P product following the failure of the SMAP radar in July 2015 to release in December 2016. The primary activities were deriving the freeze and thaw references, and developing approaches to mitigate false freeze retrievals during the summer.

### 5.1 Freeze and Thaw References

Various techniques were tested pre-launch using Aquarius data for isolating measurements characteristic of frozen and thawed conditions, including temporal averages (i.e. during January/February for freeze; July/August for thaw) and averages of a fixed number of lowest/highest seasonal backscatter values. Post-launch, it was determined that the optimal reference difference was achieved using the 20 highest NPR values from SMAP radiometer measurements during July and August 2015 (thaw) and the 20 lowest NPR during January and February 2016 (freeze) for the northern ( $\geq 45^\circ\text{N}$ ) domain. Data were separated by ascending and descending orbit. The methodological approach to defining NPR freeze and thaw references will continue to be refined for future product releases. In addition, the reference values will be updated following each transition season. To better illustrate the NPR values throughout this report, we scale the number by 100. The initial SMAP freeze and thaw scaled NPR references are shown in **Figure 5.1a and b**, with the reference difference in **Figure 5.1c**. Of particular importance are the white regions in **Figure 5.1c** which correspond to areas with a scaled NPR difference of less than 0.1, which is insufficient for FT retrievals. In **Figure 5.1d**, we show the reference difference in the global grid, which shows the reference difference is problematically small over many mid-latitude regions. Note that an AMSR-E derived ‘never frozen’ mask is also applied to the reference difference map in **Figure 5.1d**.

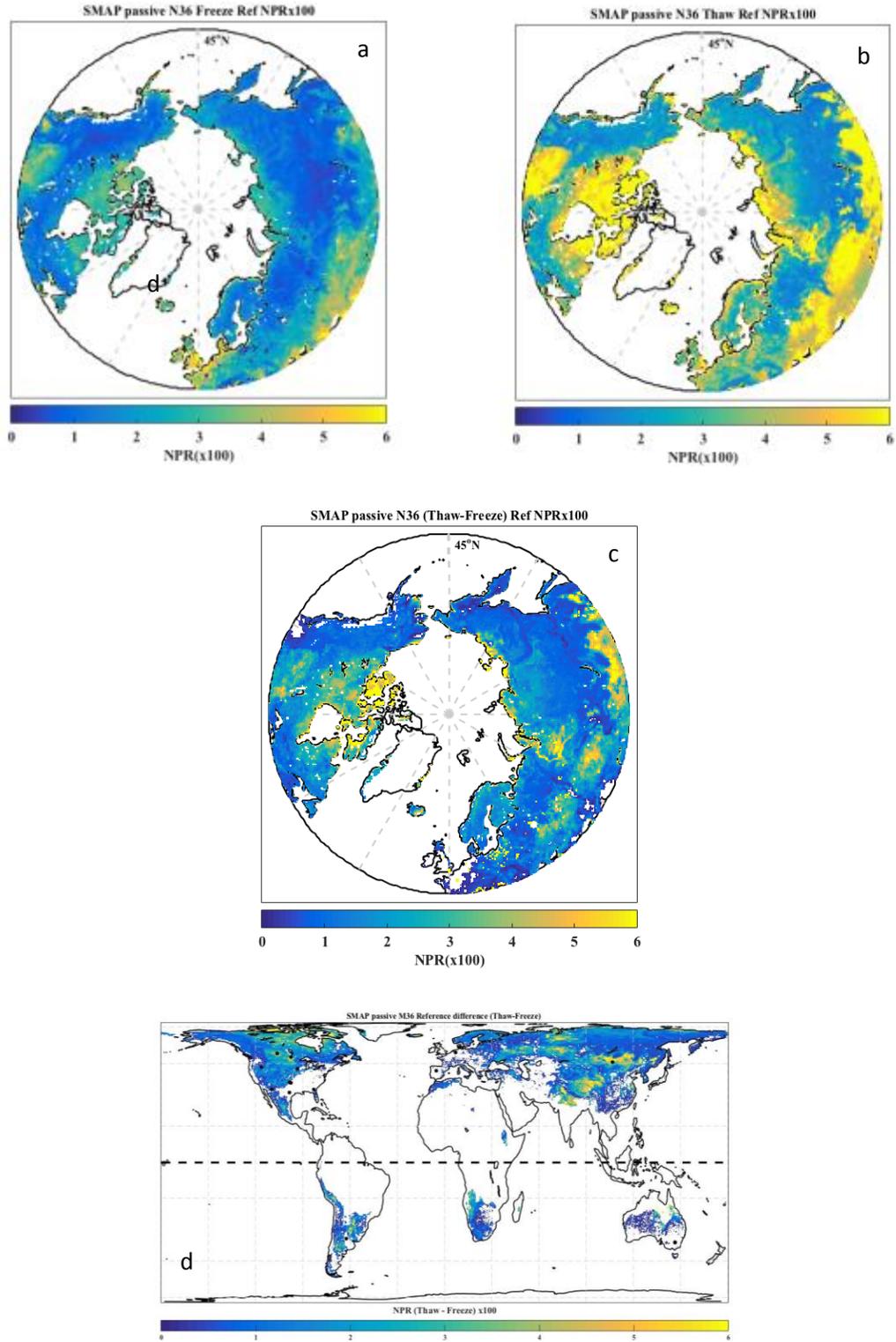


Figure 5.1. (a) SMAP freeze and (b) thaw references; reference difference on the (c) polar grid and (d) global grid. Blank areas either fall into the AMSRE climatology mask or the reference difference is  $< 0.1$ .

## 5.2 False Freeze Mitigation

Following the pixel wise determination of freeze/thaw state, two additional processing steps are applied to mitigate summer season false freeze and winter season false thaw retrievals. First, if the brightness temperature magnitude at either V or H pol is greater than 273, the pixel is set to thaw regardless of the retrieval. Second, ‘never frozen’ and ‘never thawed’ masks (**Figure 5.2**) were calculated from daily AMSR-E and AMSR2 derived freeze/thaw maps (using the approach described in [11]) over the 2002-2015 period. These masks were then applied using a 31-day moving window approach to fix the retrieval state each day for pixels that never changed freeze/thaw state during the AMSR record:

$$NeverFrozen(doy) = \sum_{i=doy-15}^{doy+15} Freeze\_AMSR\_flag(i) \quad (7)$$

$$NF\_mask = (NF == 0)$$

$$NeverThawed(doy) = \sum_{i=doy-15}^{doy+15} Thaw\_AMSR\_flag(i) \quad (8)$$

$$NT\_mask = (NT == 0)$$

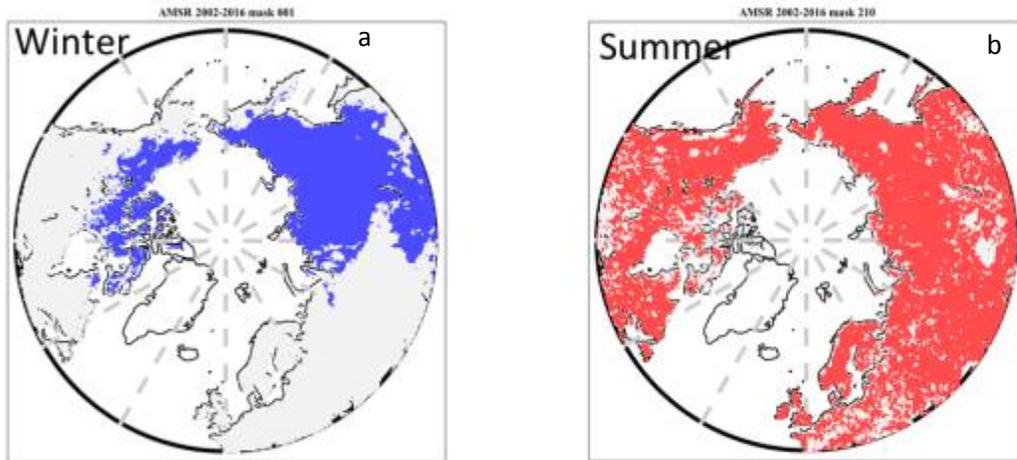


Figure 5.2. Example never thawed (a) and never frozen (b) masks for 1 January and 29 July.

While these additional processing steps do not remove all false flags, they substantially reduce obviously false flags without relying on ancillary surface temperature information.

*These false freeze mitigation efforts were developed and tested during the L3\_FT\_P algorithm development period. However, only the 273 K threshold test conservative and a fixed AMSR-E ‘never frozen’ mask were applied in the initial dataset release. The 31-day moving window masks will be applied in a subsequent update.*

### **5.3 Validated Release Testing**

Testing of the L3\_FT\_P(\_E) algorithm code was conducted using the final release L1C\_TB(\_E) T13210 (31 March 2015 through 31 December 2015) and T13240 (1 January 2016 – 31 October 2016) data sets. All of the analyses described in Section 7 are based on this dataset, which forms the basis of the product release assessment.

## 6 ASSESSMENTS

### 6.1 Large Scale Patterns and Features

An example of seven coincident FT estimates for 20 April 2015 is shown in **Figure 6.1** (note the Aquarius data cover a week centered on 20 April). While there are resolution differences (3 km for SMAP radar; 100 km for Aquarius; 25 km for AMSR2), all seven datasets capture the same general FT pattern, with some regional differences in areas of complex terrain, and along freeze/thaw transition areas. In general, the passive sensor FT products (Aquarius and AMSR2) retrieve less frozen area than the active sensor derived FT products.

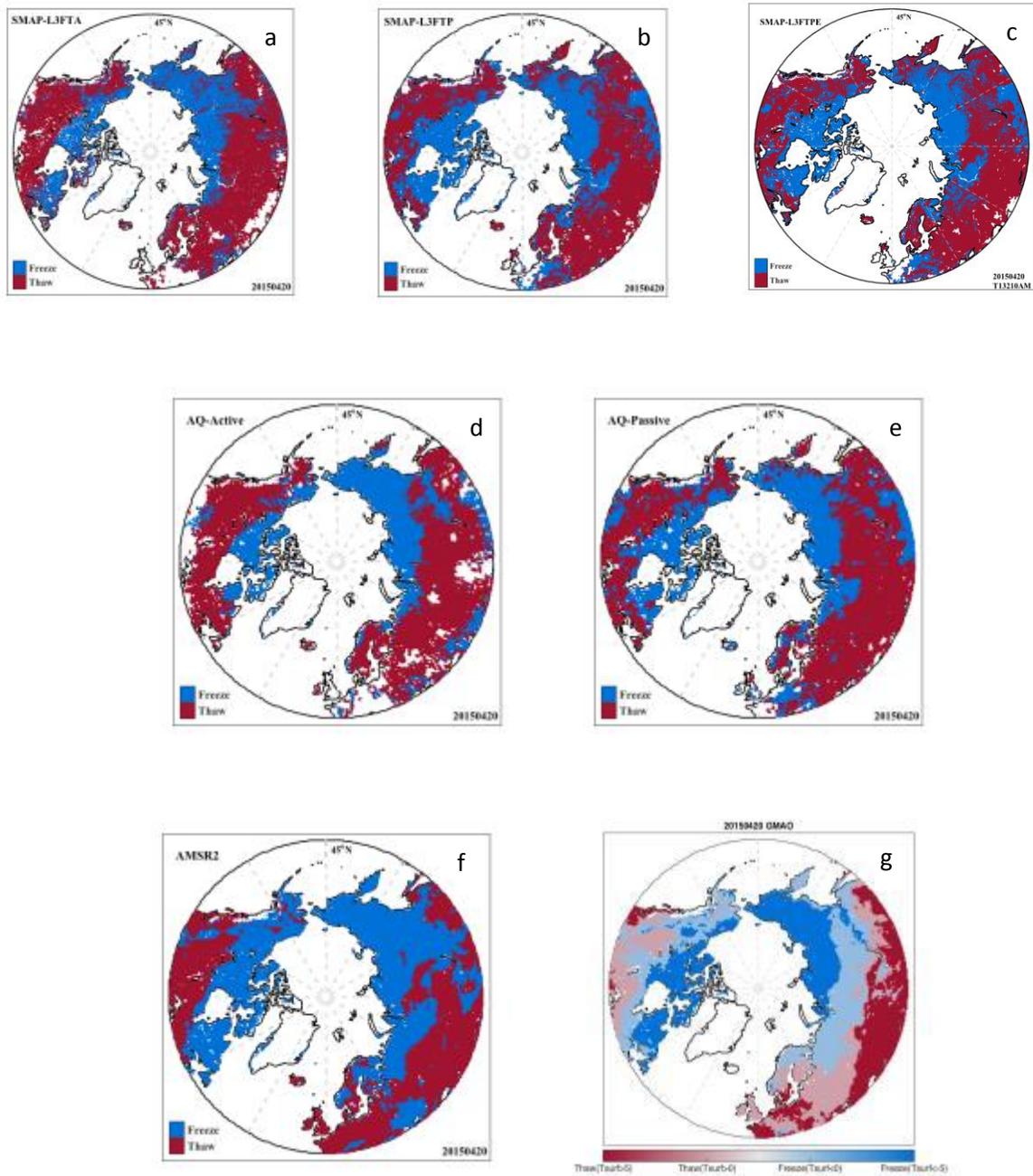


Figure 6.1 Snapshot comparison (20 April 2015) of six satellite derived FT retrievals: (a) SMAP L3\_FT\_A; (b) SMAP L3\_FT\_P; (c) SMAP L3\_FT\_P\_E; (d) Aquarius active [12]; (e) Aquarius passive; (f) AMSR2 [11]. Daily averaged FT information from GMAO Tsurf simulations are shown in (g).

## 6.2 Core Validation Sites (CVS)

The same format is followed for the presentation of results from each core validation site (**Table 6.1; Figure 6.2**). Time series of SMAP derived NPR values with coincident soil and air temperature measurements were separated for descending (~6 am local) and ascending (~6 pm local) overpass times. Surface measurements were averaged at sites with multiple soil temperature probes within the SMAP grid cell. Corresponding time series of FT flags are also provided, with in situ flags determined from soil and air temperature ( $\leq 0^{\circ}\text{C}$  = frozen). Flag agreement was calculated for the entire study period and on a monthly basis for SMAP derived FT state versus soil and air temperature derived FT state (1 represents perfect flag agreement through each available time series). These statistics simply illustrate the proportion of days with the same FT state; they do not consider the nature of any classification errors. To address this, an error matrix for each site was constructed showing the total absolute occurrence of flag agreement (green cells) false freeze (SMAP = freeze, reference flags = thaw; blue cells), and false thaw (SMAP = thaw, reference flags = freeze, yellow cells).

Table 6.1 Summary of L3\_FT\_P core validation sites.

Site Name	Site PI	Area	IGBP Land Cover	Location
Kenaston	A. Berg	Saskatchewan, Canada	Croplands	51.41N 106.50W
Boreal Ecosystem Research and Monitoring Sites	A. Black	Saskatchewan, Canada	Coniferous Forest	53.63N; 106.20W (OA) 53.99N; 105.12W (OBS)
Sodankyla	J. Pulliainen	Finland	Coniferous Forest	67.36N; 26.64E
Saariselka	J. Pulliainen	Finland	Grasslands	68.38N; 27.42E
Chersky	M. Loranty	Eastern Siberia	Deciduous Needleleaf	68.65N; 161.65E
Imnavait	E. Eukirchen	Alaska, USA	Barren/Sparse	68.62N; 149.30W
Baie-James	A. Langlois	Quebec, Canada	Coniferous Forest	53.41N; 75.013W
Cambridge Bay	A. Langlois	Northwest Territories, Canada	Barren/Sparse	69.15N; 105.11W

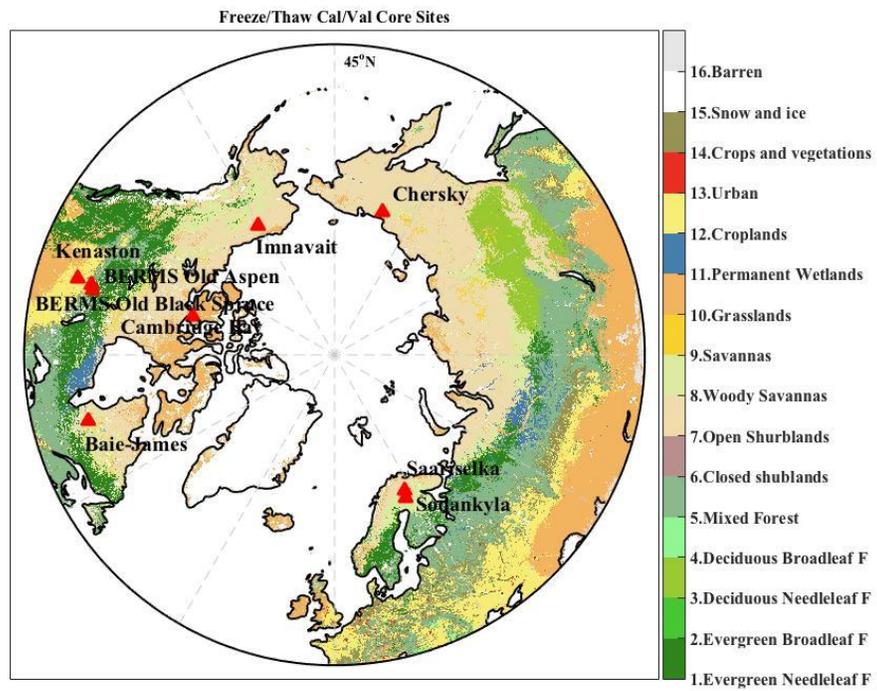


Figure 6.2. Core sites used for L3\_FT\_P validation.

## 6.2.1 L3\_FT\_P

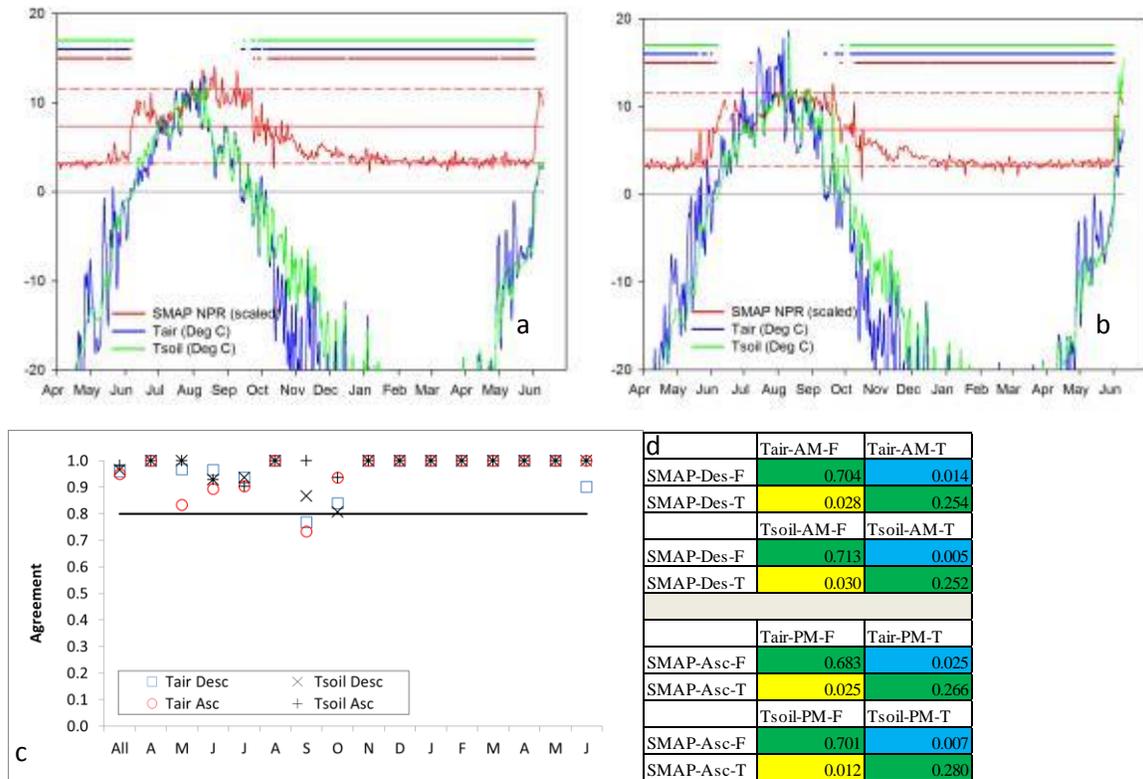


Figure 6.3. (a) Descending and (b) ascending overpass time series of NPR, air temperature, and soil temperature for Cambridge Bay. Horizontal lines note the occurrence of frozen flags. Dashed horizontal red lines note SMAP NPR freeze and thaw reference values. (c) Monthly summary of freeze/thaw flag agreement. (d) Classification error matrix: flag agreement (green cells) false freeze (SMAP = freeze, reference flags = thaw; blue cells), and false thaw (SMAP = thaw, reference flags = freeze, yellow cells).

Cambridge Bay is a high latitude, open tundra site with shallow snow, and a long frozen season. As such, it presents a fairly straightforward case for FT detection. As shown in **Figure 6.3a and b**, there is a strong and clean NPR response to both freeze and thaw onset. Soil and air temperature are strongly coupled at this site because the snowpack is thin and dense and hence has a lower insulative effect on soil temperatures than observed at boreal sites. Because of this, the NPR derived FT transitions agree closely with both air and soil temperature time series (**Figure 6.3c**). Flag agreement is over 80% for all months and for both soil and air temperature reference flags except for air temperature in September 2015. This is because of a short period of below freezing air temperatures (soil temperatures remained above zero) during which the SMAP retrievals remained thawed. The classification error matrix (**Figure 6.3d**) reflects the strong overall agreement at this site.

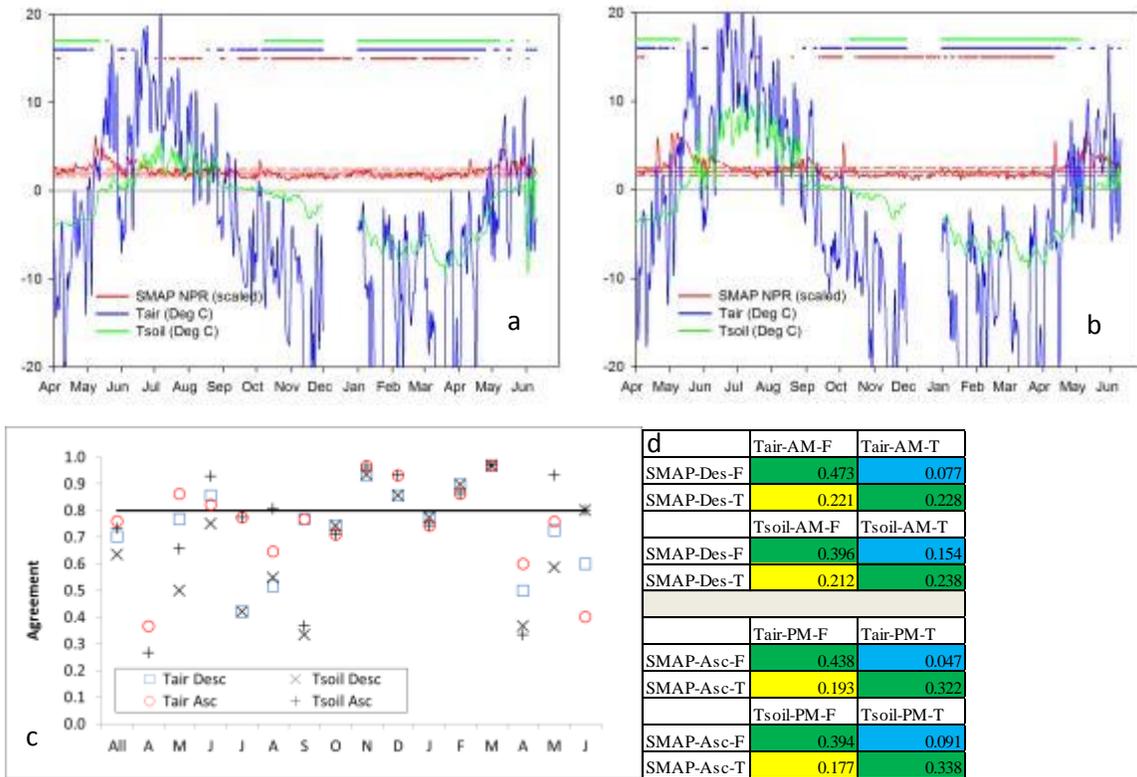


Figure 6.4. As in Figure 6.2 but for Innavait.

Innavait is a cold climate, open tundra site similar to Cambridge Bay, but the reference difference between freeze and thaw values is very small and the NPR time series indicates a weak signal to noise ratio (**Figure 6.4a and b**). It's unclear if that is due to high lake fraction or some other physiographic component of the Alaska north slope, or the proximity to the complex topography of the Brooks Range. Despite the small reference difference, there is a strong NPR response to spring thaw, but the NPR values drop again very quickly after the initial thaw transition. This decrease in NPR magnitude combined with a highly variable NPR time series results in periods of false freeze events during the summer. While there is generally good flag agreement with Tair and Tsoil measurements between October and March, there is overall better thaw transition agreement with air temperature because of the wet snow influence in spring. During autumn, soil freeze lags the air temperature transition, leading to a period in September where air temperatures are below zero, NPR response indicates frozen conditions, but Tsoil remains above zero.

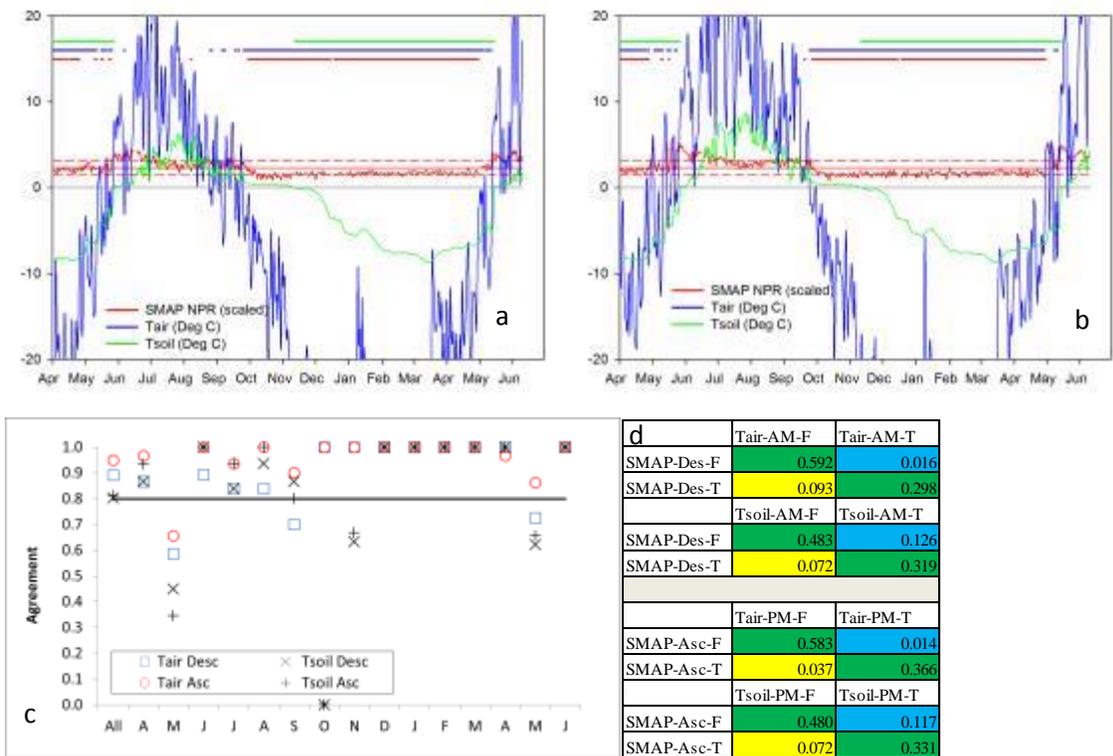


Figure 6.5. As in Figure 6.2 but for Chersky.

Chersky is a cold climate, deciduous needleleaf forest site in eastern Siberia. As will be seen at subsequent forested sites, Tsoil lags behind the Tair transition by a short period during the spring thaw, (**Figure 6.5a and b**). During fall, soil freeze lags behind air temperatures because of the insulative effect of snow cover. At Chersky, the NPR continues to decrease after air temperatures fall below freezing, but soil freeze is not required to decrease values below the freeze threshold (it takes nearly 2 months for the near surface soil to actually freeze). With the exception of May 2015, there is excellent agreement between Tair, Tsoil, and NPR derived FT flags from spring 2015 through June 2016.

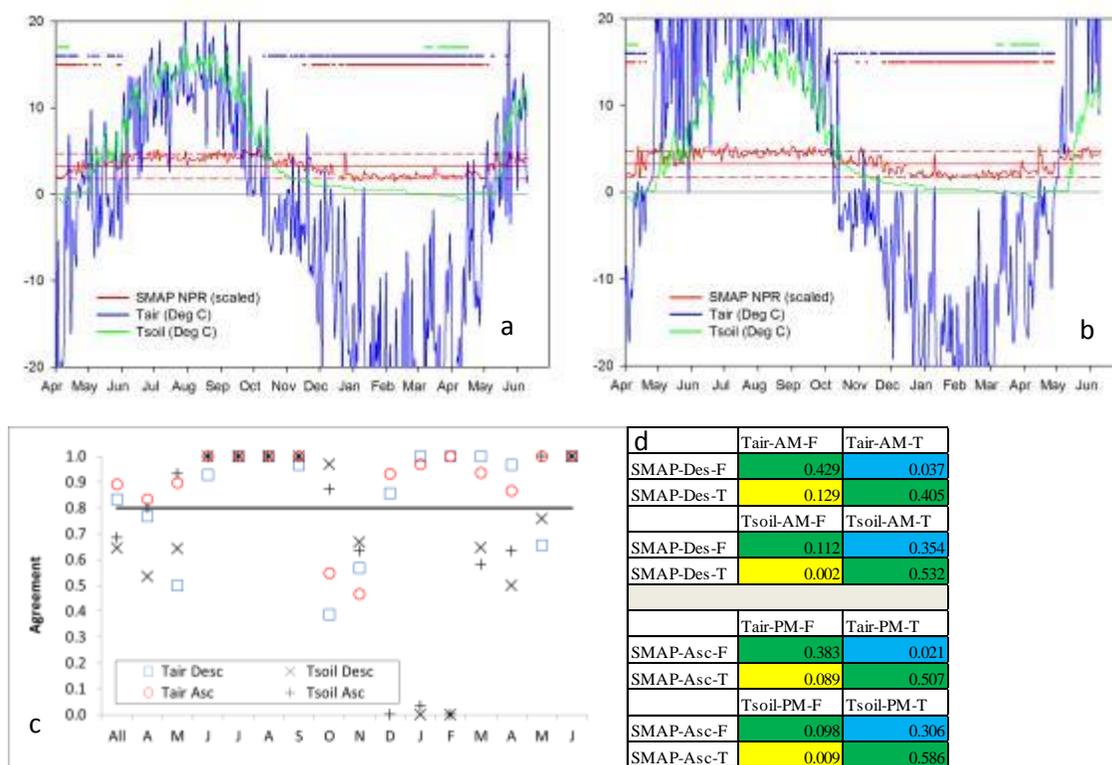


Figure 6.6. As in Figure 6.2 but for Baie-James.

As was observed at Chersky, the seasonal evolution of NPR at Baie-James responds closely to Tair, with significant lags for Tsoil (**Figure 6.6a and b**). With the exception of October and November 2015, flag agreement with air temperature was stronger than soil temperature (**Figure 6.6c**). Fall 2015 is a difficult time period to interpret, however, because air temperatures were largely below zero, but the decrease in NPR was insufficient to transition below the threshold value. Retrievals remained thawed (despite the cold air temperatures) in agreement with Tsoil which remained above zero because of the insulating effect of snow cover. The NPR values decreased below the threshold by December, triggering frozen flags more than three months in advance of the soil temperature measurements falling below zero. This suggests a significant insulative effect from snow at this site, likely enhanced by a thick organic layer (~20 cm) above the mineral soil. The overall warm soil temperatures at this site will be investigated further as they may indicate instrument uncertainty.

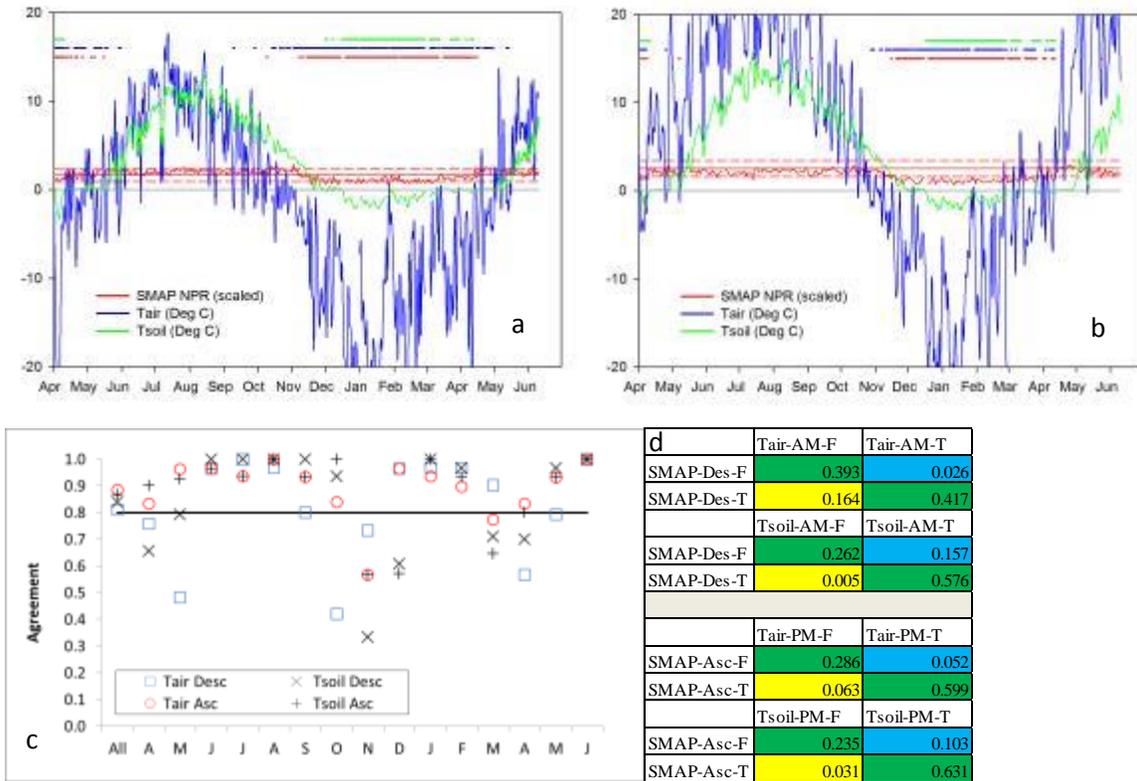


Figure 6.7. As in Figure 6.2 but for BERMS – Old Black Spruce.

The BERMS Old Black Spruce (OBS) site has a small reference difference and small amplitude seasonal cycle in the NPR time series (**Figure 6.7a and b**). Despite this apparently weak FT signal, retrieval performance exceeds the 80% accuracy target at this site (**Figure 6.7c**), and summer false freeze events are minimal. The comparison of NPR with Tsoil and Tair is typical of forest environments with close agreement during the thaw transition but long periods following the freeze transition when Tsoil remains above zero (~3 months) long after the onset of continuous air temperatures below zero. The primary classification error at this site is therefore false freeze compared to a Tsoil reference (**Figure 6.7d**).

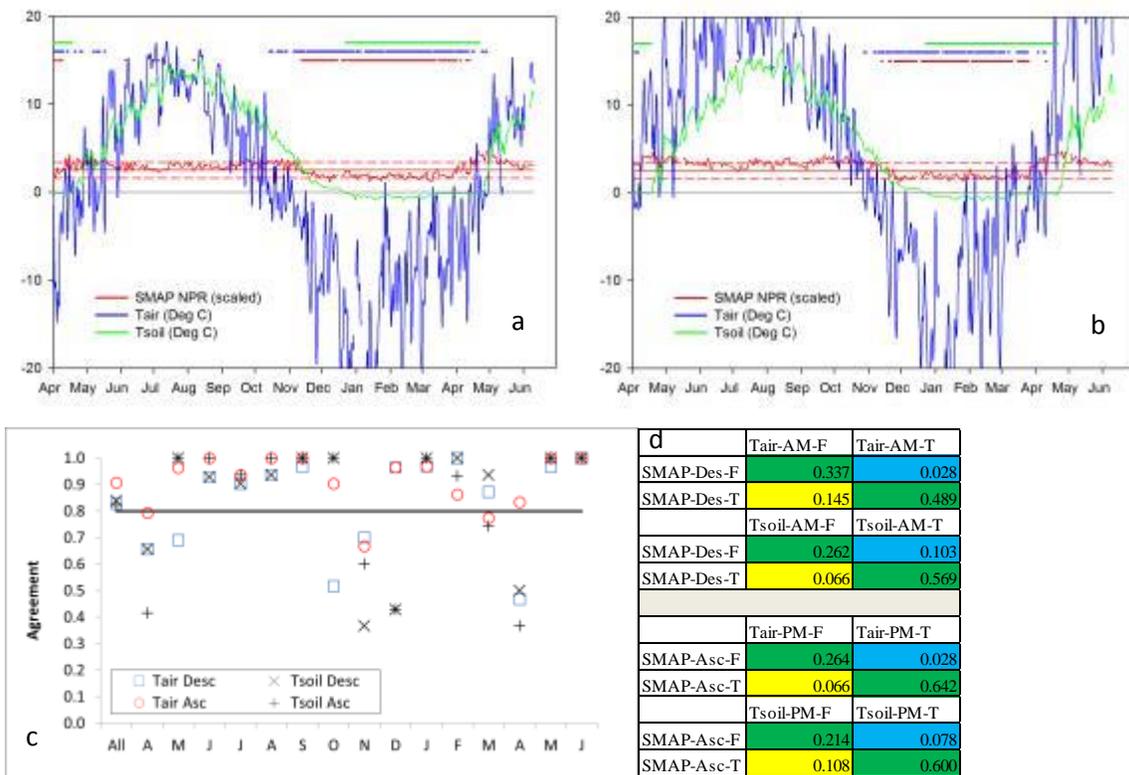


Figure 6.8. As in Figure 6.2 but for BERMS – Old Aspen.

The BERMS old aspen (OA) site is in geographic proximity to the BERMS OBS site. The conditions are similar between sites with the exception of the forest type. Like OBS, the OA site has a small reference difference, low seasonal amplitude in NPR, but a limited number of summer false freeze events (**Figure 6.8a and b**). Consistent with the other forest sites, there is a 2-3 month lag in Tsoil dropping below zero after the freeze onset, which accounts for the majority of the classification errors. Flag agreement with Tair is generally strong.

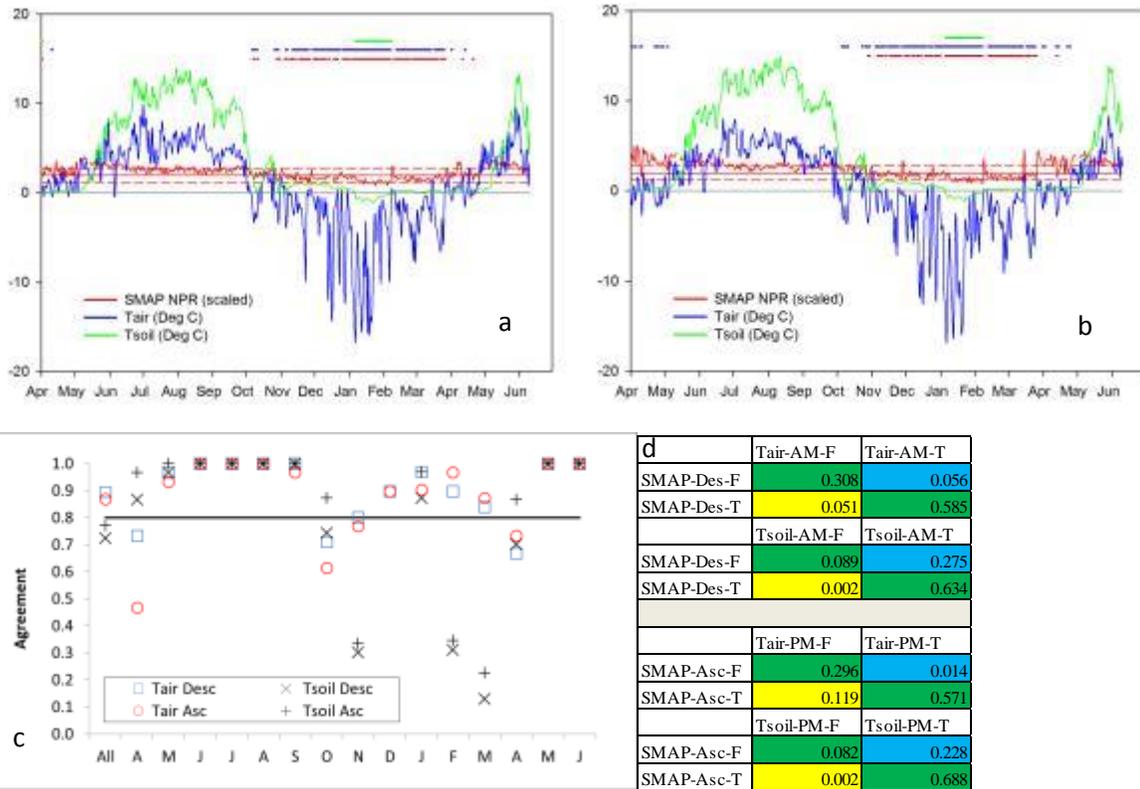


Figure 6.9. As in Figure 6.2 but for Sodankyla.

The Sodankyla site is located in a forested region with a high fraction of wetlands, hence there is a very short frozen soil period despite cold air temperatures (**Figure 6.9a and b**). This results in strong overall agreement with Tair derived reference flags, but poor agreement with Tsoil during the early and late frozen periods.

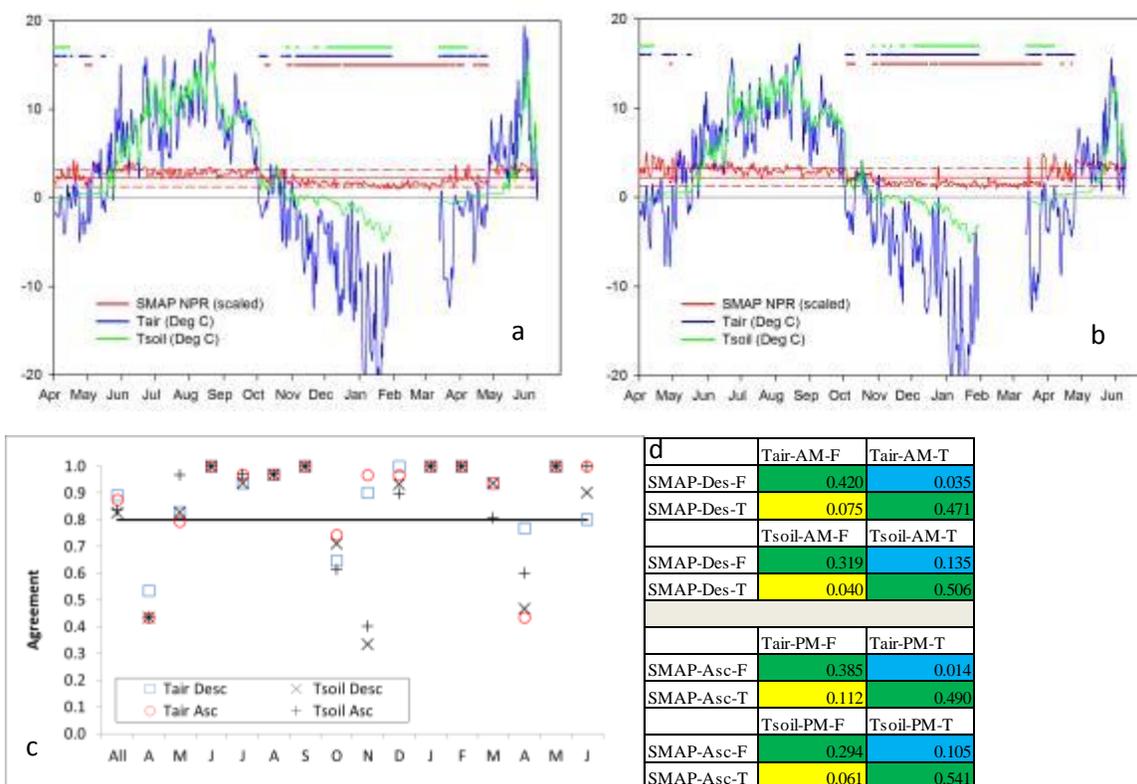


Figure 6.10. As in Figure 6.2 but for Saariselka.

The Saariselka site is located in an upland tundra environment surrounded by boreal forest, creating heterogeneous sub-grid land cover conditions. Despite this, the NPR time series exhibits strong response to FT transitions (**Figure 6.10a and b**), and overall flag agreement is above 80% (**Figure 6.10c**). Disagreement with Tsoil flags is limited to November 2015 because of the lag in soil freezing, consistent with other core sites. Lower agreement during April of 2015 and 2016 is related to NPR fluctuations around the threshold during periods of diurnal freeze and thaw.

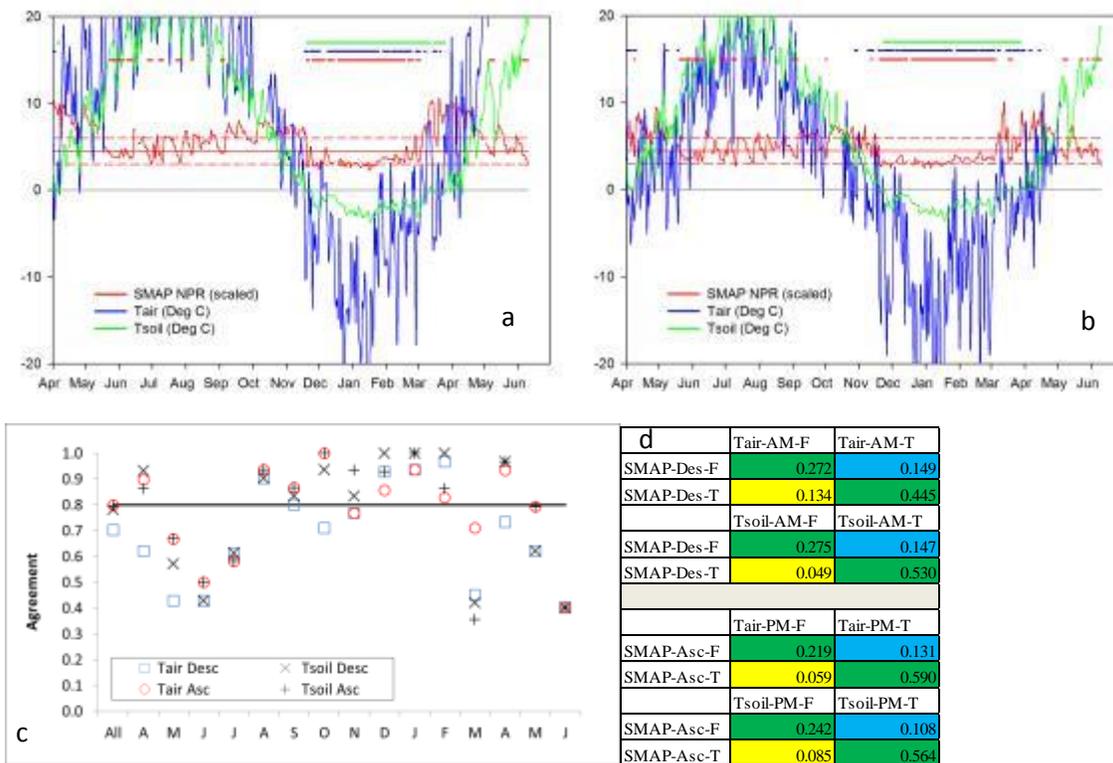


Figure 6.11. As in Figure 6.2 but for Kenaston.

Kenaston is unique because it is the only open prairie core site located near the southern margin of the FT domain. NPR response during the freeze and thaw transitions is strong and is in close agreement with Tair and Tsoil measurements during the fall freeze transition (**Figure 6.11a and b**). There was an early thaw transition in NPR compared to both Tair and Tsoil. Unlike all other core sites, the largest errors occur during the summer months due to false freeze periods due to oscillations in the NPR above and below the threshold. This false freeze problem is pervasive across some sub-regions of the FT domain (see Section 6.1) but time series over Kenaston shows that the primary transitions are still accurately determined despite the summer false freeze issue.

## 6.2.2 L3\_FT\_P\_E

The L3\_FT\_P (36 km) assessment (presented in **Figures 6.3** through **6.11**) was repeated for L3\_FT\_P\_E (9 km). Overall, the assessment indicated very similar results, both in term of the seasonal cycle of NPR with respect to air temperature, soil, and snow state, and the validation statistics. Given the similarity in results, we will not present the full assessment in this report. Instead, **Figure 6.12** shows the difference in FT flag agreement relative to air and soil temperature for L3\_FT\_P\_E minus L3\_FT\_P. When computed in this manner, positive values correspond to improved agreement for the enhanced resolution product; negative values correspond to degraded agreement. At most sites, the change in flag agreement is less than +/- 5%. The exceptions are Chersky (reduced flag agreement at 9 km for both Tair and Tsoil) and Baie-James (reduced flag agreement for Tsoil only). In order to explore why these larger differences occurred at these sites, both 36 and 9 km time series are shown in **Figure 6.13** (descending orbits only).

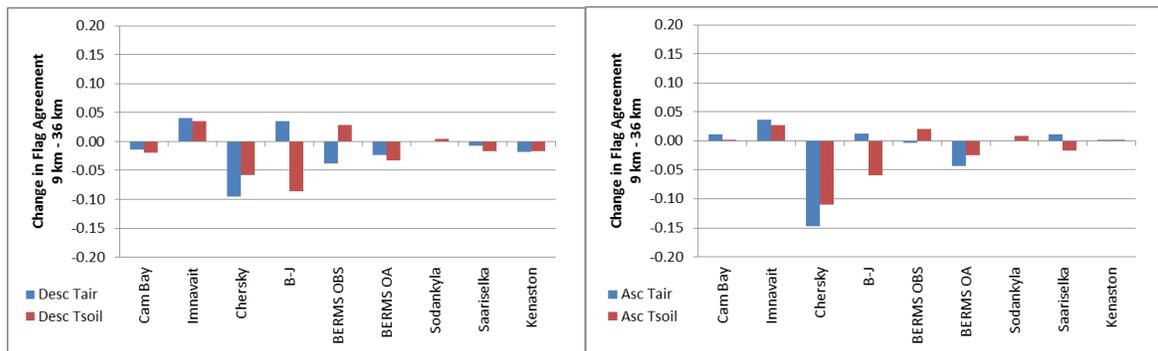


Figure 6.12. Difference in FT flag agreement, L3\_FT\_P\_E minus L3\_FT\_P for descending (left) and ascending (right) overpasses. Positive values correspond to improved agreement for the enhanced resolution product; negative values correspond to degraded agreement.

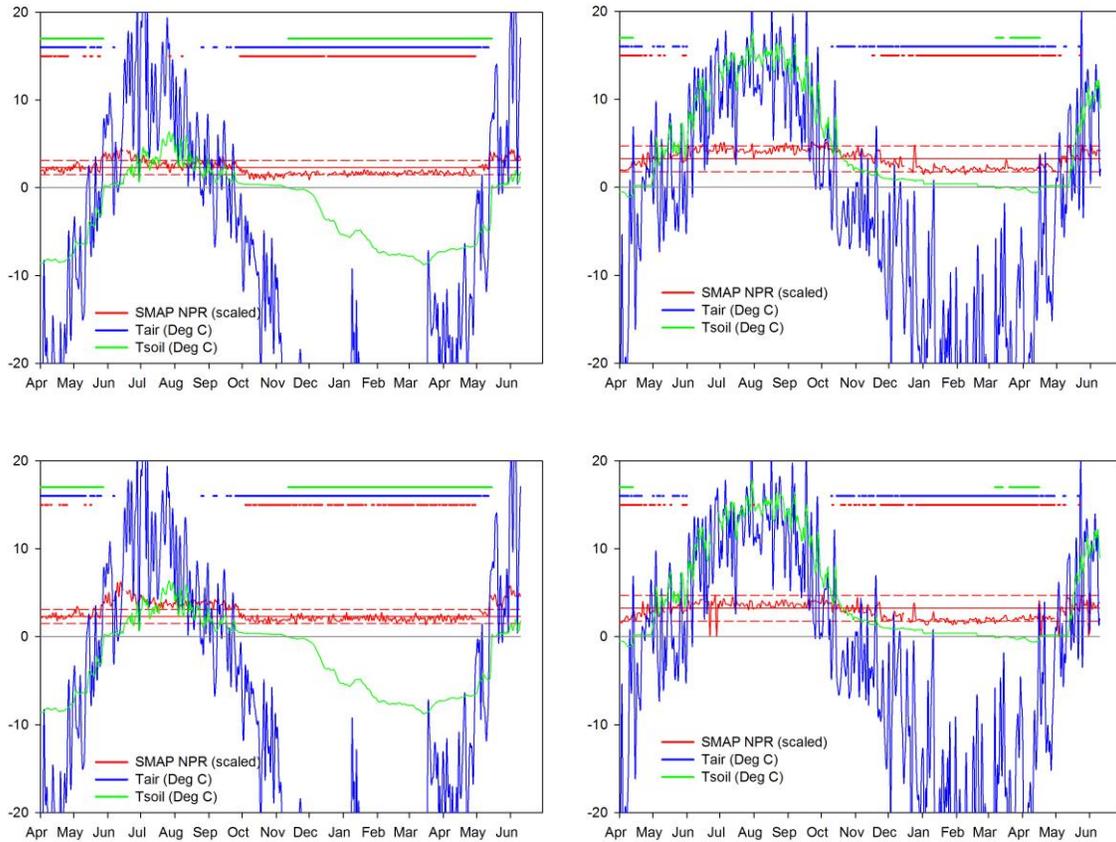


Figure 6.13. Descending overpass time series of NPR, air temperature, and soil temperature for Chersky (left) and Baie-James (right) at 36 km (top) and 9 km (bottom) resolutions. Horizontal lines note the occurrence of frozen flags. Dashed horizontal red lines note SMAP NPR freeze and thaw reference values.

At Chersky, the reduced FT flag agreement is due to a slight increase in NPR at 9 km resolution. High frequency variability during the frozen season results in the NPR magnitude fluctuating above and below the FT threshold. This results in short duration false thaw periods during the mid-winter. The same NPR variability is evident at 36 km, but the magnitude consistently remains below the threshold. At Baie-James, an earlier freeze onset in October 2015 is derived at 9 km resolution. This produces better overall flag agreement with Tair (note the positive values for this metric in **Figure 6.12**). However, since the timing of soil freeze lags so far behind the air temperature transition at this site, earlier freeze detection at 9 km results in weaker agreement with Tsoil flags.

### 6.2.3 Core Site Summary

The number of core sites available for validation of SMAP FT products is limited in number, but they yield important insights into the behavior of NPR time series for tundra, forested, and prairie environments. There are consistent land cover and climatic controls on the NPR response to FT state as summarized below.

*Tundra sites* are characterized by a strong NPR response with clean transitions during spring and fall driven by relatively low vegetative biomass cover, a strong seasonal FT cycle driven by very cold winter temperatures, and thin snow cover which melts quickly in spring and is a poor insulator of the underlying soil.

*Boreal sites* also exhibit a clear FT signal during seasonal transitions, but have a weak annual cycle in NPR amplitude because of the influence of forest vegetation. Although winter air temperatures are cold, the insulative effect of a deep/low density snowpack is significant, and can delay soil freeze for 1 to 5 months after air temperatures drop below zero (although measurement uncertainty in Tsoil, and the additional insulative effects of a thick organic layer overlying the mineral soil needs to be explored further at some sites). The timing/magnitude of initial snow accumulation event(s), and soil temperature at the time of initial snow accumulation are likely important factors in controlling the duration of the lag before boreal soils freeze in winter.

The *prairie site* was characterized by a strong NPR response to freeze and thaw transitions, but summer false freeze events are very problematic due 1 to 3 week cyclical variability in NPR.

There are also land cover independent influences on the NPR signal. In all cases, regardless of vegetation and climatic zone, wet snow cover induced a thaw response in the NPR time series when air temperatures increased above zero in spring even over frozen soil. This creates a consistent tendency for SMAP derived spring thaw flags to lead soil thaw.

Some care must be taken in interpreting the FT flag agreement results with Tair and Tsoil measurements – both of these variables are imperfect at characterizing the landscape FT state. There are long periods in the early winter, particularly at boreal forest sites, when flag agreement with Tsoil measurements are weak because the soil at depth remains unfrozen due to the insulating effect of snow cover. This is despite air temperatures being consistently below zero continuously for weeks and snow lying on the surface. The landscape is effectively ‘frozen’ even if this is not captured by the Tsoil measurements. Conversely in spring, soil remains frozen after air temperatures increase above zero and snowmelt is underway. Because the SMAP NPR responds to wet snow, flag agreement is better with Tair during the onset of spring melt, even when the soil is still frozen. Overpass time also has an influence on the evaluation statistics. Agreement is generally better for the ascending (PM) orbits because the NPR response is not always sensitive to ephemeral and transitional freeze events during which the in situ measurements indicate frozen conditions in the morning and thawed conditions in the afternoon. When the SMAP retrievals remain ‘thawed’ during these events, they disagree with the in situ measurements at the time of morning overpass (when in situ measurements are frozen), but agree at the time of afternoon overpass (when in situ measurements are thawed). These tendencies are reflected in the overall flag agreement statistics, which are stronger for Tair relative to Tsoil, and for ascending (PM) versus descending (AM) overpasses (**Table 6.2**). Resolution enhancement from 36 to 9 km in the brightness temperature inputs to the FT retrieval resulted in minor overall changes to NPR behavior over the seasonal FT cycle, and resulting accuracy in FT retrievals (**Table 6.2**). Flag agreement was slightly lower for both Tair and Tsoil metrics at both overpass times, but remained above the baseline mission objection for Tair, and above the minimum mission requirement for Tsoil.

Table 6.2 Overall flag core site flag agreement statistics for L3\_FT\_P.

<b>FT_P</b>			
<b>Orbit</b>	<b>Cases</b>	<b>Flag Agreement Tair</b>	<b>Flag Agreement Tsoil</b>
Des	3852	0.834	0.783
Asc	3852	0.875	0.811
<b>FT_P_E</b>			
Des	3823	0.821	0.766
Asc	3829	0.862	0.794

### 6.3 Sparse Networks

The sparse networks consist of selected stations from the SnoTel networks. Compared with the core site stations, there is only one in-situ measurement in each satellite grid. To ensure the in-situ measurements from the single station represents the whole grid, the criteria for the sparse network stations are 1) relative uniform land cover 2) relative smooth terrain over the satellite grid 3) valid soil temperature at 5 cm and air temperature measurement at 2 m. In addition, the reference difference in the grid cells must exceed 0.5, so that a robust retrieval algorithm can be performed. In **Figure 6.14** and **Table 6.3**, we summarize the sparse network stations in Alaska. A threshold value of 0°C was used to determine FT state from both soil and air temperature measurements. Hourly data are available from all selected stations in SnoTel network; to match with the daily ascending and descending signals, we average the hourly measurements for AM (1:00~12:00) and PM (13:00 ~ 24:00).

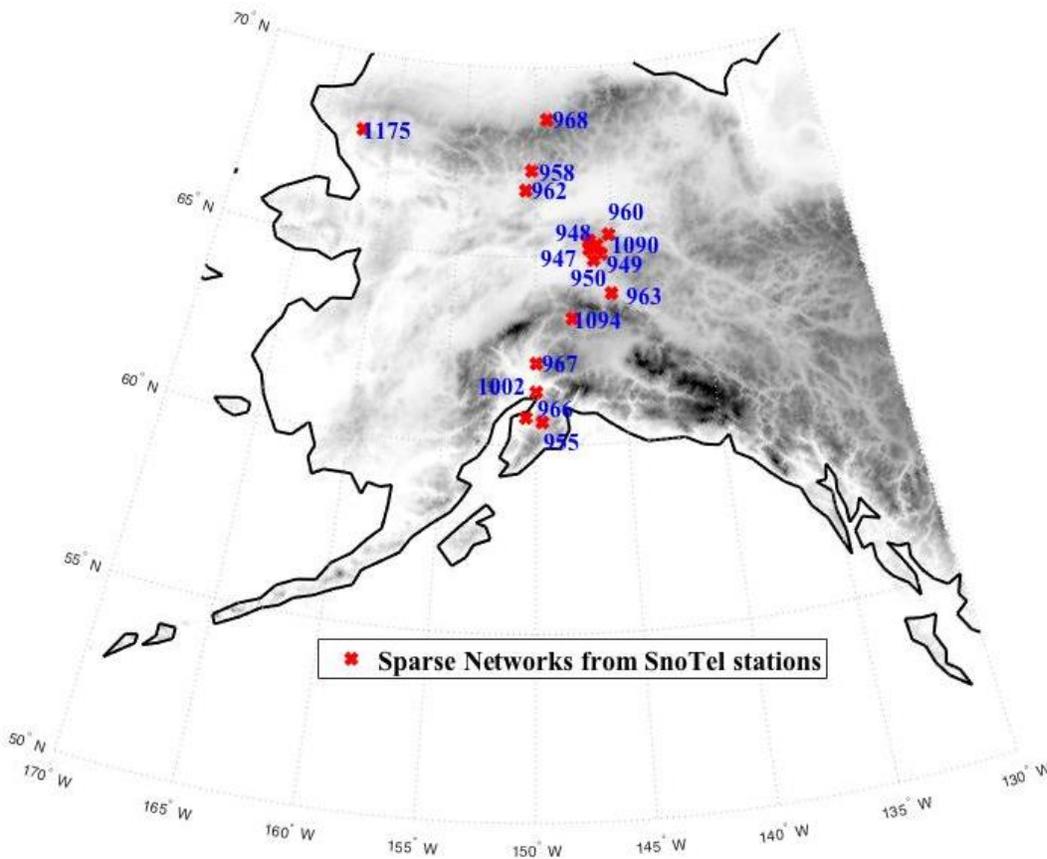


Figure 6.14. SnoTel sites used for L3\_FT\_P evaluation.

Table 6.3 List of sparse network stations

ID	LAT	LON	Name	State	Land Cover	Elevation
1002	61.398	-149.94	Point Mackenzie	AK	Grassland	250
1090	65.416	-146.59	Upper Nome Creek	AK	Grassland	2520
1094	63.345	-147.77	Monahan Flat	AK	Shrub open	2710
1175	67.963	-162.21	Kelly Station	AK	Shrub open	310
947	65.153	-146.61	Little Chena Ridge	AK	Shrub open	2000
948	65.277	-146.17	Mt Ryan	AK	Shrub open	2800
949	65.01	-145.9	Monument Creek	AK	Savannah woody	1850
950	64.888	-146.34	Munson Ridge	AK	Savannah woody	3100
955	60.63	-149.64	Summit Creek	AK	Shrub open	1400
958	67.267	-150.22	Coldfoot	AK	Savannah woody	1040
960	65.517	-145.28	Eagle Summit	AK	Savannah woody	3650
962	66.773	-150.67	Gobblers Knob	AK	Shrub open	2030
963	63.958	-145.42	Granite Creek	AK	Grassland	1240
966	60.744	-150.52	Kenai Moose Pens	AK	Forest mixed	300
967	62.173	-149.99	Susitna Valley High	AK	Forest mixed	375
968	68.629	-149.21	Imnavait Creek	AK	Shrub open	3050

In **Figure 6.15**, we compare the retrieved FT flag accuracy against the reference FT state determined from soil and air temperatures respectively, for both ascending and descending passes over the SNOTEL sites. From November to February, the air temperature and soil temperature flags yield very similar performance metrics, both of which are above the 80% baseline mission requirement threshold set for SMAP. This indicates freeze retrievals coincident to both air and soil temperature falling below zero. However, in October (both 2015 and 2016), the accuracy decreases to approximately 65% for both air temperature and soil temperature. This indicates some timing errors in the SMAP derived freeze onset, which may be improved through threshold optimization in future releases. During the thaw season (April and May), the SMAP FT flags are in better accordance with the air temperature rather than soil temperature. The stronger retrieval agreement with the air temperature flag indicates the NPR value is responding to wet snow, consistent with the core validation site analysis.

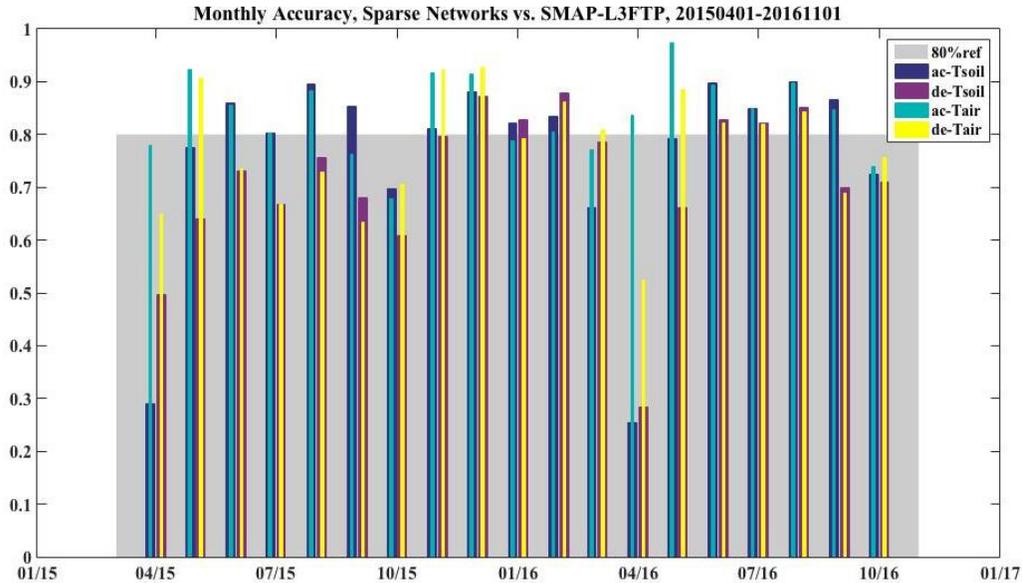


Figure 6.15 Monthly FT flag agreement for all validated SnoTel sites, 2015-2016. Narrow columns indicate use of air temperature measurements for validation; wide columns indicate use of 5 cm soil temperature. Gray shading indicates 80% flag agreement requirement set by the SMAP mission.

A detailed summary of the evaluation of the SMAP passive derived FT product using observations from the SnoTel network is shown in **Table 6.4** and **Table 6.5**. The statistics were computed separately for ascending/descending overpasses (the first three columns for each month corresponds to ascending/PM overpass and the last three columns are descending/AM). It is evident that the primary source of uncertainty in the SMAP derived FT retrievals is during the spring transition when the reference data indicate frozen ground and the SMAP retrieval indicate a thawed state. Again, this is consistent with the influence of wet snow on the NPR value.

Table 6.4. Sparse network flag agreement statistics with respect to measured soil temperature.

Month	Accuracy (PM)	SMAP-F Temp-T	SMAP-T Temp-F	Accuracy (AM)	SMAP-F Temp-T	SMAP-T Temp-F
Apr-15	0.289	0.000	0.711	0.498	0.013	0.489
May-15	0.774	0.052	0.173	0.640	0.033	0.327
Jun-15	0.860	0.136	0.004	0.731	0.261	0.008
Jul-15	0.802	0.198	0.000	0.667	0.333	0.000
Aug-15	0.896	0.104	0.000	0.756	0.244	0.000
Sep-15	0.853	0.107	0.040	0.680	0.298	0.022
Oct-15	0.698	0.228	0.073	0.609	0.318	0.073
Nov-15	0.811	0.164	0.024	0.796	0.176	0.029
Dec-15	0.881	0.082	0.037	0.871	0.082	0.047
Jan-16	0.822	0.030	0.148	0.828	0.032	0.140
Feb-16	0.834	0.000	0.166	0.878	0.000	0.122
Mar-16	0.662	0.004	0.333	0.785	0.009	0.206
Yearly Statistic	<b>0.765</b>	<b>0.092</b>	<b>0.143</b>	<b>0.728</b>	<b>0.150</b>	<b>0.122</b>
Apr-16	0.253	0.000	0.747	0.284	0.004	0.711
May-16	0.793	0.005	0.202	0.662	0.040	0.298
Jun-16	0.898	0.093	0.009	0.829	0.167	0.004
Jul-16	0.849	0.151	0.000	0.821	0.179	0.000
Aug-16	0.899	0.101	0.000	0.852	0.148	0.000
Sep-16	0.866	0.111	0.023	0.699	0.280	0.021
Oct-16	0.725	0.146	0.129	0.710	0.247	0.043

Table 6.5 Sparse network flag agreement statistics with respect to measured air temperature.

Month	Accuracy (PM)	SMAP-F Temp-T	SMAP-T Temp-F	Accuracy (AM)	SMAP-F Temp-T	SMAP-T Temp-F
Apr-15	0.780	0.069	0.150	0.650	0.036	0.315
May-15	0.924	0.071	0.005	0.907	0.045	0.048
Jun-15	0.855	0.136	0.009	0.735	0.257	0.008
Jul-15	0.802	0.198	0.000	0.667	0.333	0.000
Aug-15	0.883	0.103	0.013	0.730	0.218	0.052
Sep-15	0.762	0.092	0.145	0.634	0.154	0.212
Oct-15	0.679	0.171	0.150	0.705	0.115	0.180
Nov-15	0.917	0.050	0.033	0.924	0.043	0.033
Dec-15	0.915	0.043	0.043	0.927	0.035	0.038
Jan-16	0.788	0.071	0.141	0.793	0.062	0.145
Feb-16	0.805	0.086	0.108	0.862	0.044	0.094
Mar-16	0.771	0.097	0.132	0.808	0.026	0.165
Yearly Statistic	<b>0.824</b>	<b>0.099</b>	<b>0.077</b>	<b>0.778</b>	<b>0.114</b>	<b>0.107</b>
Apr-16	0.836	0.033	0.131	0.524	0.029	0.447
May-16	0.974	0.005	0.021	0.886	0.045	0.069
Jun-16	0.893	0.093	0.013	0.822	0.164	0.013
Jul-16	0.849	0.151	0.000	0.819	0.181	0.000
Aug-16	0.897	0.101	0.002	0.843	0.146	0.011
Sep-16	0.847	0.104	0.049	0.690	0.185	0.125
Oct-16	0.740	0.174	0.086	0.757	0.153	0.090

We repeated the sparse network assessment for the enhanced resolution product (L3\_FT\_P\_E). In **Figure 6.16**, we plot the accuracy difference between the enhanced resolution product and standard product over a year. A positive number means the enhanced resolution product shows improvement in flag agreement. The finer resolution retrievals are slightly noisier over time (as shown in **Figure 6.13** for Chersky), which introduces more errors in the stable seasons (summer and winter). This tends to slightly degrade the overall accuracy. However, as was the case at the core sites (see **Figure 6.12**), the enhanced resolution product has very similar performance statistics compared to the standard resolution product.

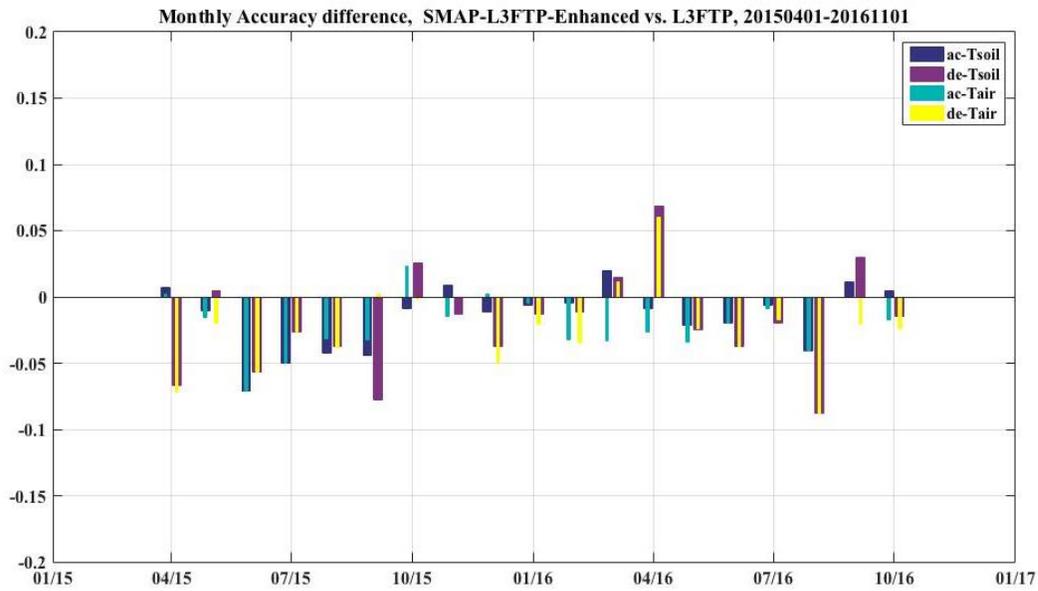


Figure 6.16 Monthly FT flag agreement difference between Enhance product (9km) and regular product (36km) for all validated SnoTel sites, 2015-2016. Narrow columns indicate use of air temperature measurements for validation; wide columns indicate use of 5 cm soil temperature.

## 6.4 Air Temperature Validation

The SMAP FT retrievals were evaluated against independent FT estimates derived from in situ daily surface (screen-height) air temperature measurements from 1,920 weather stations which report hourly observations to the World Meteorological Organization (WMO) located across the FT ( $\geq 45^\circ\text{N}$ ) domain (**Figure 6.17**). These northern stations represented a subset of a larger set of 4,178 stations distributed across the global FT classification domain that were used for validating FT frozen flags in the SMAP L3\_SM\_P[E] products (validation across the hemispheric domain is not reported here). The WMO station records were obtained from the NCDC Global Summary of the Day (NWS, 1988). The station daily minimum and maximum air temperatures were converted to FT estimates assuming a fixed ( $0.0^\circ\text{C}$ ) temperature threshold between frozen and non-frozen conditions, and compared with respective SMAP descending and ascending overpass FT retrievals following previously developed methods (Kim et al. 2012). The Euclidian distance between each grid cell centroid and the WMO station locations was computed to select a single representative station closest to the center of a grid cell when two or more stations were located within the same cell. The surface air temperature daily minimum ( $\text{SAT}_{\text{mn}}$ ) and maximum ( $\text{SAT}_{\text{mx}}$ ) records for the selected stations were used to define daily frozen ( $T \leq 0^\circ\text{C}$ ) and non-frozen ( $T > 0^\circ\text{C}$ ) temperature conditions, and compared with corresponding FT classification results from the overlying grid cells and respective AM and PM overpass periods, assuming that the local timing of daily  $\text{SAT}_{\text{mn}}$  and  $\text{SAT}_{\text{mx}}$  occurs near the SMAP equatorial crossing times [11]. The FT classification agreement was assessed through grid cell-to-point comparisons between the WMO daily SAT measurements and overlying SMAP FT results.

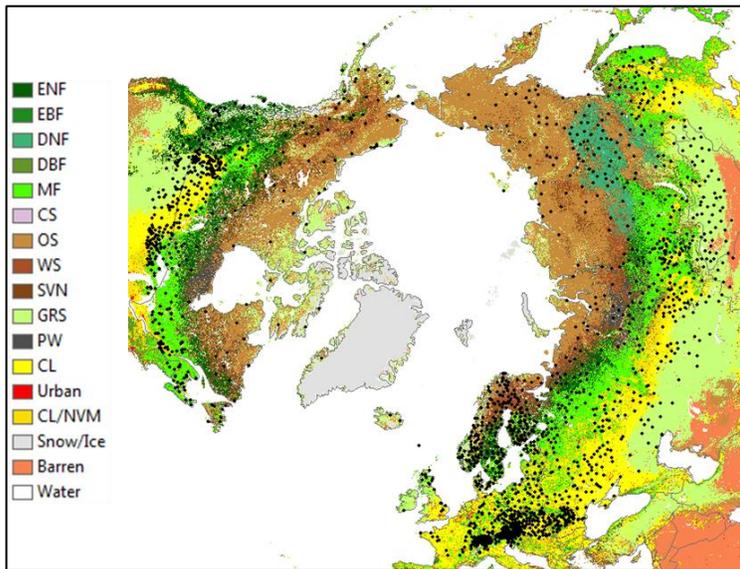


Figure 6.17. MODIS MCD12Q1 IGBP Land cover map showing WMO weather stations (black dots) used for SMAP FT validation assessment over the northern domain.

The SMAP FT frozen flag from the L3\_SM\_P[E] products was used for the WMO station based validation assessment. The daily (AM and PM) FT frozen flags embedded in the soil moisture products are derived using the same FT classification algorithm as the SMAP L3\_FT\_P[E] products, but the underlying  $T_b$  inputs and FT outputs are gridded to 36-km and 9-km resolution global EASE-grid (V2) formats rather than polar EASE-grid (V2) formats. Grid cells dominated by permanent snow/ice cover and large water bodies or where seasonal FT conditions have an insignificant impact on ecosystem

processes were excluded from the validation assessment. Grid cells with minimal (<0.1) NPR difference between reference frozen and non-frozen conditions were also excluded from the validation assessment.

A plot of the resulting daily mean AM and PM overpass FT spatial classification accuracies for the northern domain from the L3\_SM\_P[E] records is shown in **Figure 6.18**. The FT accuracy is generally higher during summer when the pattern and persistence of non-frozen conditions is more homogeneous and stable, and lower during the more heterogeneous spring and fall transition periods. For both products the PM overpass FT results generally exceed the 70% mean spatial classification accuracy threshold of the minimum mission requirement. The 9-km PM overpass FT results also generally meet or exceed the targeted 80% accuracy threshold of the baseline mission. However, the FT classification accuracy from the AM overpass results is significantly lower than the PM results and is generally at or below the 70% accuracy target for the minimum mission. The lower AM FT accuracy relative to the PM results is consistent with the SMAP FT core validation site comparisons and previous FT global assessment studies based on other satellite microwave sensors [11]. The lower AM FT classification accuracy may reflect one or more factors including relatively larger SAT and FT heterogeneity during the morning observations.

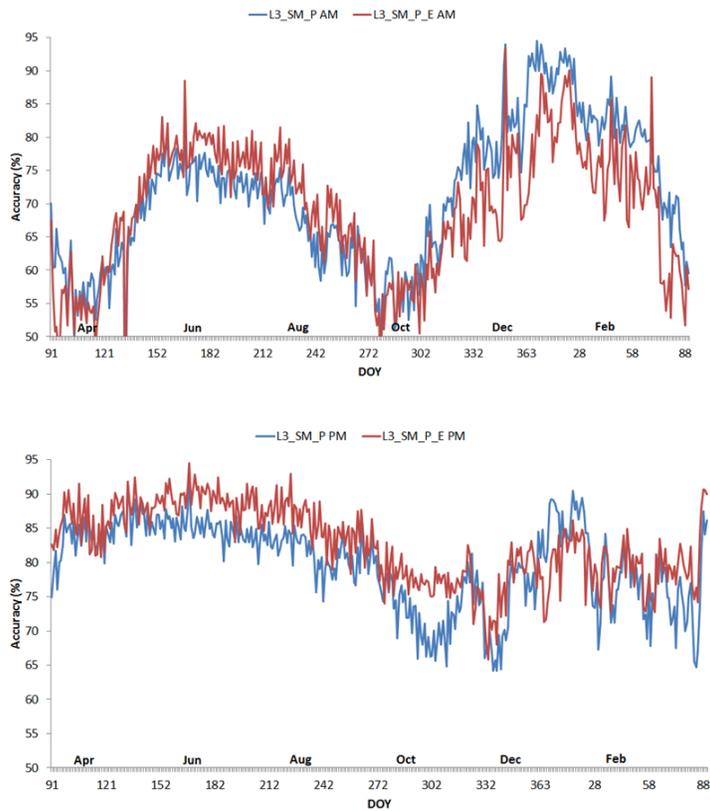


Figure 6.18. Seasonal pattern of SMAP daily mean FT classification accuracy (%) in relation to WMO in-situ surface air temperature measurements for AM (top) and PM (bottom) overpass results; blue and red lines denote respective 36-km and 9-km grid results.

The validation results over the northern domain show relatively small (2-3%) FT accuracy differences between the 9-km and 36-km gridded products. Accuracy for the 9-km product was slightly better than the 36-km product for the PM results, but was lower than the 36-km product for the AM results. The relatively minor differences in FT accuracy between the 9-km and 36-km grids may reflect potential noise

introduced from the Backus Gilbert  $T_b$  downscaling method and an effective resolution that is coarser than the final 9-km product gridding.

The spatial distribution of mean annual FT classification agreement (%) for the 9-km AM and PM overpass results in relation to grid cell-to-point comparisons with the WMO stations is shown in **Figure 6.19**. The spatial pattern in FT accuracy is generally similar between the 9-km and 36-km results (not shown). Accuracy is generally higher over northern boreal, Arctic and high elevation areas with relatively longer and consistent frozen seasons. Accuracy is lower over the southern portions of the domain and along the boundaries of major air masses where the ability to identify consistent NPR frozen and non-frozen reference conditions is constrained by more transient FT conditions. The AM overpass FT accuracy is lower than PM accuracy, particularly in the more transient FT areas. The spatial pattern in FT classification accuracy indicated from the WMO station comparisons also reflects lower NPR signal-to-noise in higher vegetation cover areas (e.g. forests). Coastal areas also include lower  $T_b$  and NPR signal-to-noise due to open water contamination, though the 9-km results show relatively improved delineation of complex coastal and mountain areas than the 36-km results.

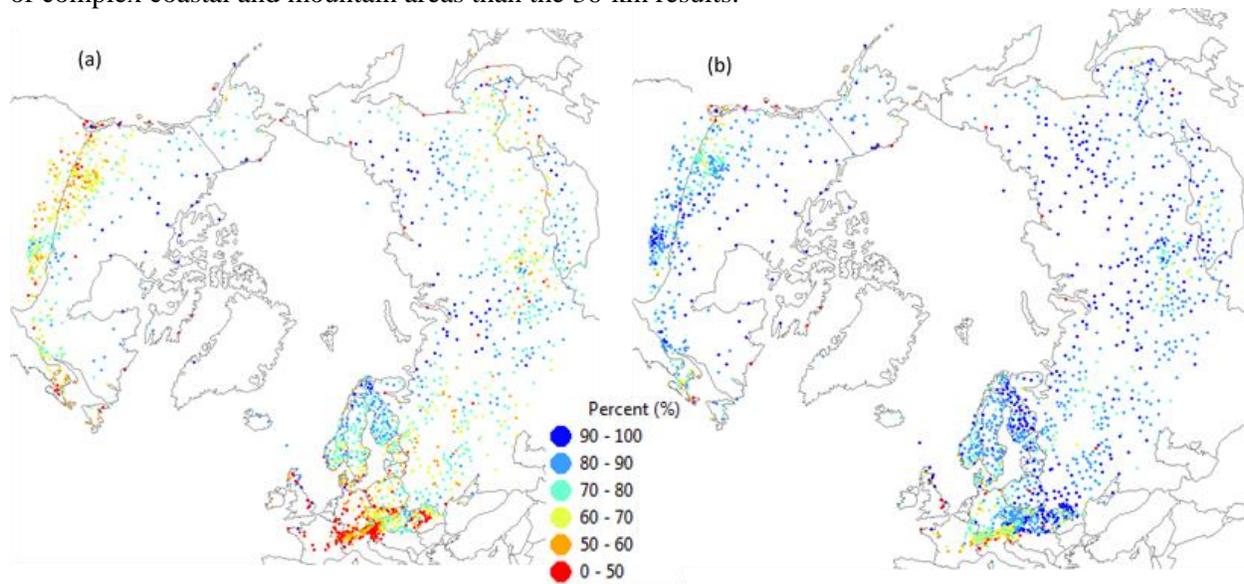


Figure 6.19. The spatial distribution of SMAP FT classification accuracy for northern WMO station sites from the 9-km AM (a) and PM (b) overpass results.

The overall FT classification accuracy from the L3\_SM\_P[E] daily FT frozen flags is presented in **Table 6.6**. These results summarize the SMAP FT validation assessments against WMO station observations for the northern ( $\geq 45^\circ\text{N}$ ) and domain. Overall, the PM overpass FT classification results from both the 9-km and 36-km products exceed the targeted 70% accuracy threshold for the minimum mission. The PM FT accuracy for the 9-km product is enhanced by 3-6% relative to the 36-km product and exceeds the 80% accuracy threshold for the baseline mission. However, the AM FT accuracy is significantly lower than the PM FT results, and is generally at or below the 70% accuracy threshold of the minimum mission. The SMAP FT validation assessment against the WMO station network observations is based on an assessment of the FT frozen flag from the L3\_SM\_P[E] products, which are produced in a global EASE-grid format. A similar FT accuracy assessment using the L3\_FT\_P[E] polar EASE-grid products is expected to produce higher FT accuracy over the northern domain owing to less polar grid cell distortion at higher latitudes in relation to WMO ground station locations.

Table 6.6. Mean annual FT classification accuracy (%) of the L3\_SM\_P\_E (9km) and L3\_SM\_P (36km) product daily FT frozen flags in relation to WMO air temperature based FT station observations for northern ( $\geq 45^\circ\text{N}$ ) and global domains.

<b>FT product</b>	<b>AM overpass</b>	<b>PM overpass</b>
<b>9km</b>	69.2 %	82.6 %
<b>36km</b>	71.4 %	79.7 %

## 6.5 Satellite Time Series Inter-comparison

A comparison of the time series of SMAP (L3\_FT\_A and L3\_FT\_P) derived frozen area across land areas north of 45° with AMSR2 [11] and GMAO surface temperature derived estimates are shown in **Figure 6.20** for both overpasses. There are inter-dataset differences in the rate of areal thaw due to differences in frequency (i.e. L-band for SMAP; Ka-band for AMSR2), spatial resolution (i.e. 3 km for SMAP L3\_FT\_A; 25 km for AMSR2) and the sensitivity of the active versus passive measurements. An important difference between products is the timing and duration of complete summer thaw conditions. The GMAO Tsurf product shows complete thaw by early June, in close agreement with the pm overpasses of the AMSR2 product. The SMAP and Aquarius estimates retain a higher amount of frozen area due to false freeze flags. These false flags, which are more predominant in the AM overpasses, will be removed to a large extent when the AMSR derived daily masks (as described in Section 5.2) are applied in future product releases.

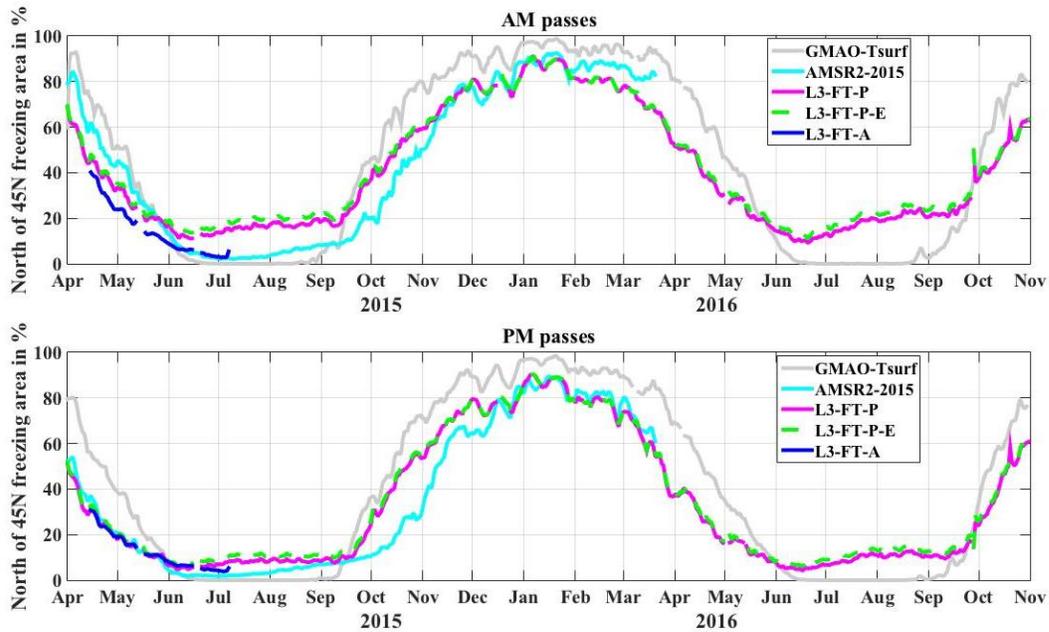


Figure 6.20. Time series of % frozen area across the FT domain indicated from AMSR2, SMAP radar and radiometer datasets.

## 6.6 Summary

In general, FT retrievals are challenging to validate because the spaceborne measurement must be related to the FT state of a horizontally (land cover; topography) and vertically (soil/snow/vegetation) heterogeneous scene. There are also a limited number of sites with comprehensive measurements available for validation.

Assessment at high-latitude core validation sites showed stronger flag agreement statistics for Tair relative to Tsoil, and for ascending (PM) versus descending (AM) overpasses. Resolution enhancement from 36 to 9 km in the brightness temperature inputs to the FT retrieval resulted in minor overall changes to NPR behavior over the seasonal FT cycle, and resulting accuracy in FT retrievals. There is a tendency for the SMAP spring thaw signal to lead the soil temperature based FT reference flag, which is attributed to the influence of wet snow cover on the radiometer signal and subsequent delay in soil thawing until snow melt is at an advance state. There are long periods in the early winter, particularly at boreal forest sites, when flag agreement with Tsoil measurements are weak because the soil at depth remains unfrozen due to the insulating effect of snow cover. This is despite air temperatures being consistently below zero continuously for weeks and snow lying on the surface. The landscape is effectively ‘frozen’ even if this is not captured by the Tsoil measurements. It is therefore important for users to understand that the satellite FT retrievals represent an integrated landscape state, not simply the near-surface soil layer.

The results from the core validation sites were confirmed using a network of sparse measurements across Alaska. Flag agreement using air temperature and soil temperature measurements was at or near the 80% mission accuracy requirement, with generally better agreement for ascending versus descending orbits, and air temperature versus soil temperature derived reference flags.

Validation was performed using in situ daily surface (screen-height) air temperature measurements from 1,920 weather stations which report hourly observations to the World Meteorological Organization (WMO) located across the FT ( $\geq 45^\circ\text{N}$ ) domain. FT classification accuracy from the AM overpass results were lower (generally at or below the 70% accuracy target for the minimum mission) than the PM results (generally at the 80% baseline mission accuracy target). Retrieval accuracy was generally higher over northern boreal, Arctic and high elevation areas with relatively longer and consistent frozen seasons. Accuracy was lower over the southern portions of the domain, spatially coincident to regions with a small reference difference between thawed and frozen states. Validation results show relatively small (2-3%) FT accuracy differences between the 9-km and 36-km gridded products.

Time series of frozen area across the FT domain show differences between multi-frequency active and passive microwave derived datasets, and highlight the degree to which summer season false freeze retrievals are present in the SMAP FT products. Implementation of daily AMSR-E derived ‘never frozen’ and ‘never thawed’ masks in future product releases will address this issue.

## 7 OUTLOOK

This report describes the validated release of the SMAP radiometer derived landscape freeze/thaw product at standard (L3\_FT\_P) and enhanced (L3\_FT\_P\_E) spatial resolutions. Priorities for enhancement to these products will focus on the following areas:

*Optimization of algorithm parameters.* The plan for L3\_FT\_P product updates is based on a rotating schedule of reference updates, threshold optimization, and re-processing. Optimization experiments will be conducted at core and sparse network sites as part of L3\_FT\_P product development and validation.

*False freeze mitigation.* False freeze mitigation efforts were developed and tested during the L3\_FT\_P algorithm development period. However, only the 273 K threshold test and a conservative and fixed AMSR-E ‘never frozen’ mask were applied in the initial dataset release. More robust false freeze flag removal is anticipated when the 31-day moving window AMSR-E masks are applied in a subsequent update.

*Incorporating Field Campaign results.* Unlike soil moisture, there is no legacy of airborne L-band remote sensing campaigns to support process studies, scaling, and algorithm development for FT. An active/passive L-band airborne freeze-thaw campaign (collaboration between NASA, Environment Canada, and Agriculture and Agri-Food Canada) was conducted during transient FT events over agricultural land in Manitoba, Canada during the first two weeks of November 2015. Analysis of this dataset will primarily support L3\_FT\_P[E] refinement.

*Validation of global FT flags.* This report is focused on the assessment of flags in the L3\_FT\_P products, which are limited to land areas north of 45N. The FT flags are also, however, produced on the global grid for eventual use in L2-L4 soil moisture and carbon flux products. FT flag validation across the global grid is therefore still required, particularly given the prevalence of summer false freeze flags in regions with a weak FT signal.

*Science development.* To date, emphasis has been placed on algorithm and product development, particularly given the need to switch the FT processing stream from radar to radiometer inputs. Moving forward, there is a need to develop science applications focused on FT state and carbon, water, and energy budgets, and the use of FT retrievals to evaluate land surface model simulations of soil thermal state.

## **8 ACKNOWLEDGEMENTS**

This document resulted from analyses and discussion among the L3\_FT\_P Team, Cal/Val Partners, and other members of the SMAP Project Team. The contributions of the following individuals are noted (alphabetically): Aaron Berg, Andy Black, Eugenie Euskirchen, Alexandre Langlois, Mike Loranty, Kimmo Rautiainen, Tracy Rowlandson, Alexandre Roy, Alain Royer, and Jilmarie Stephens.

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, in support of the SMAP mission.

Canadian contributions to SMAP FT product development and validation are supported by Environment and Climate Change Canada and the Canadian Space Agency.

## 9 REFERENCES

- [1] Science Data Calibration and Validation Plan Release A, March 14, 2014 JPL D-52544
- [2] Dunbar, S., X. Xu, A. Colliander, C. Derksen, J. Kimball, and Y. Kim. “Algorithm Theoretical Basis Document (ATBD): L3\_FT\_P and L3\_FT\_P\_E,” Revision A, October 15, 2016. Available at <http://smap.jpl.nasa.gov/science/dataproducts/ATBD/>.
- [3] Entekhabi, D., S. Yueh, P. O’Neill, K. Kellogg et al., SMAP Handbook, JPL Publication JPL 400-1567, Jet Propulsion Laboratory, Pasadena, California, 182 pages, July, 2014.
- [4] SMAP Level 1 Mission Requirements and Success Criteria. (Appendix O to the Earth Systematic Missions Program Plan: Program-Level Requirements on the Soil Moisture Active Passive Project.). NASA Headquarters/Earth Science Division, Washington, DC, version 5, 2013.
- [5] Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV): <http://calvalportal.ceos.org/CalValPortal/welcome.do> and WWW: Land Products Sub-Group of Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV): <http://lpvs.gsfc.nasa.gov>
- [6] Chris Derksen, C., X. Xu, R. S. Dunbar, A. Colliander, Y. Kim, J. Kimball, A. Black, E. Euskirchen, A. Langlois, M. Loranty, P. Marsh, K. Rautiainen, A. Roy, A. Royer, and J. Stephens, “Retrieving landscape freeze/thaw state from Soil Moisture Active Passive (SMAP) radar and radiometer measurements”, Remote Sensing of Environment, in review.
- [7] Rautiainen, K., J. Lemmetyinen, J. Pulliainen, J. Vehviläinen, M. Drusch, A. Kontu, J. Kainulainen and J. Seppänen, “L-Band radiometer observations of soil processes in boreal and subarctic environments”, IEEE Transactions on Geoscience and Remote Sensing, 50(5), 1483-1497, 2012.
- [8] Rautiainen, K., J. Lemmetyinen, M. Schwank, A. Kontu, C. Ménard, C. Mätzler, M. Drusch, A. Wiesmann, J. Ikonen, and J. Pulliainen, “Detection of soil freezing from L-band passive microwave observations”, Remote Sensing of Environment, 147, 206–218, 2014.
- [9] Rautiainen, K., T. Parkkinen, J. Lemmetyinen, M. Schwank, A. Wiesmann, J. Ikonen, C. Derksen, S. Davydov, A. Davydova, J. Boike, M. Langer, M. Drusch and J. Pulliainen, “SMOS prototype algorithm for detecting autumn soil freezing”, Remote Sensing of Environment, 180, 346–360, 2016.
- [10] Roy, A., A. Royer, C. Derksen, L. Brucker, A. Langlois, A. Mialon and Y. Kerr. “Evaluation of spaceborne L-band radiometer measurements for terrestrial freeze/thaw retrievals in Canada”, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(9), 4442 – 4459, 2015.
- [11] Kim, Y., J. Kimball, K. Zhang, and K. McDonald, “Satellite detection of increasing Northern Hemisphere non-frozen seasons from 1979 to 2008: Implications for regional vegetation growth”, Remote Sensing of Environment, 121: 472–487, 2012.
- [12] Xu, X., C. Derksen, S. Yueh, R. S. Dunbar, and A. Colliander “Freeze/thaw detection and validation using Aquarius’ L-band backscattering data”, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(4), 1370-1381, 2016.