# Soil Moisture Active Passive (SMAP) Project
# Calibration and Validation for the L3_FT_P and L3_FT_P_E Data Product (Version 3)

**Citation:**

Xiaolan Xu[1], R. Scott Dunbar[1], Andreas Colliander[1], Youngwook Kim[2], John Kimball[2], Chris Derksen[3]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA
[2]University of Montana, Missoula, MT
[3]Environment and Climate Change Canada (ECCC), Toronto ON M3H ST4, CA



**October 30, 2020**

**JPL D-56296**

National Aeronautics and Space Administration

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

# TABLE OF CONTENTS

# 1 EXECUTIVE SUMMARY

During the post-launch Cal/Val Phase of SMAP there were 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. The SMAP mission now encompasses more than five years of global observations and with ongoing operations under its extended mission phase. 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 and gridded in both north polar grid and global grid projection formats. The complete daily coverage is achieved by combining up to 3 days of SMAP orbital swath data. This report provides analysis and assessment of the latest (Version 3) validated SMAP Level 3 Landscape Freeze/Thaw (L3_FT_P) and Enhanced resolution Freeze/Thaw (L3_FT_P_E) passive products. The Version 3 product is relatively mature, incorporating calibration and validation assessments, and refinements gained from detailed studies of prior product releases and a longer data record.

Assessment methodologies are separated for the polar and global grid products due to the different coverage and predominant underlying algorithms associated with each product domain. The polar product assessment includes comparisons of SMAP freeze/thaw (FT) retrievals with *in situ* observations from core validation sites (CVS) and sparse networks where both air temperature and surface soil temperature have been used to indicate the ground condition. The core sites have multiple in-situ measurements and the data have been upscaled to SMAP grid resolution; whereas the sparse network data only have one station per grid cell. The global grid product is validated through comparisons between the SMAP FT retrievals and corresponding surrogate FT observations derived from WMO in situ daily surface air temperature measurements, along with assessments of potential degradation in FT performance due to increases fractional open water cover, tree cover and terrain heterogeneity. The resulting FT product validation and performance assessment meets the criteria established by the Committee on Earth Observing Satellites (CEOS) for Stage 2 validation.

The pre-launch SMAP baseline science mission objective for the freeze/thaw product 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 mean spatial classification accuracy of 80% and 2-day average temporal precision separated by AM and PM overpasses and 3km spatial resolution; a lower mean spatial classification accuracy of 70% was established as the minimum mission requirement [1]. Given the failure of the SMAP radar in July 2015, the original mission baseline requirement will continue to be addressed, albeit at reduced spatial resolutions of 36 km and 9 km using SMAP radiometer inputs. The switch to SMAP passive microwave brightness temperatures with a coarser (~40-km) native footprint as primary inputs for the FT classification introduces fundamental differences in algorithm performance and product specifications for L3_FT_P[E] compared to the SMAP radar derived product (L3_FT_A; April – July 2015). However, due to the high correlation between the brightness temperature retrievals and freeze/thaw related dielectric shifts and physical temperature changes at lower latitudes, the released passive products are able to provide complete global coverage and up to twice-daily observations of freeze/thaw status, including all land areas where frozen temperatures are a significant constraint to surface water mobility and ecosystem processes.

This document provides an evaluation of the SMAP passive FT product accuracy and performance for the period of record extending from 1 April 2015 to 1 April 2020. The assessment 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 product spatial resolution was relaxed from 3 km to 36 km due to the change from SMAP radar to radiometer inputs, but this evaluation shows that the enhanced 9-km grid resolution L3_FT_P_E product provides very similar accuracy to the standard 36 km resolution L3_FT_P product. The primary sources of uncertainty are:

(1) Differences (typically less than 1 week) in the identification of the primary freeze and thaw transition dates in the fall and spring, respectively, which are expected due to challenges in determining the landscape FT state at 36 km and 9 km resolution in the presence of horizontal and vertical (soil/snow/vegetation) variability in FT state. Specific processes which influence radiometer retrieval of FT state include the influence of wet snow cover (which triggers thaw flags in the presence of frozen soil), and the insulative effects of snow on soil temperature (freeze retrievals in the presence of thawed soil).

(2) Obviously false freeze retrievals during summer months and false thaw retrievals during winter. In the higher latitudes, anomalous freeze and thaw retrievals (i.e. speckles) are mainly caused by unstable algorithm FT reference conditions; sub-grid scale variability in microclimate and brightness temperatures due to inhomogeneous landcover and rugged terrain; and open water cover contamination of the brightness temperature signal over adjacent land areas within the satellite footprint. At lower latitudes, additional factors influence temporal changes in landscape dielectric properties and can degrade FT signal-to-noise. The low correlation of the physical temperature and brightness temperature increase the uncertainty of the FT prediction from the extended algorithm.

# 2 OBJECTIVES OF CAL/VAL

During the post-launch Cal/Val (Calibration/Validation) Phase of the SMAP mission 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 grid resolution of 9 km. The approaches and procedures follow those described in the SMAP Cal/Val Plan [2]; full details on the Level 3 radiometer derived freeze/thaw algorithm and products are provided in the Algorithm Theoretical Basis Document [3].

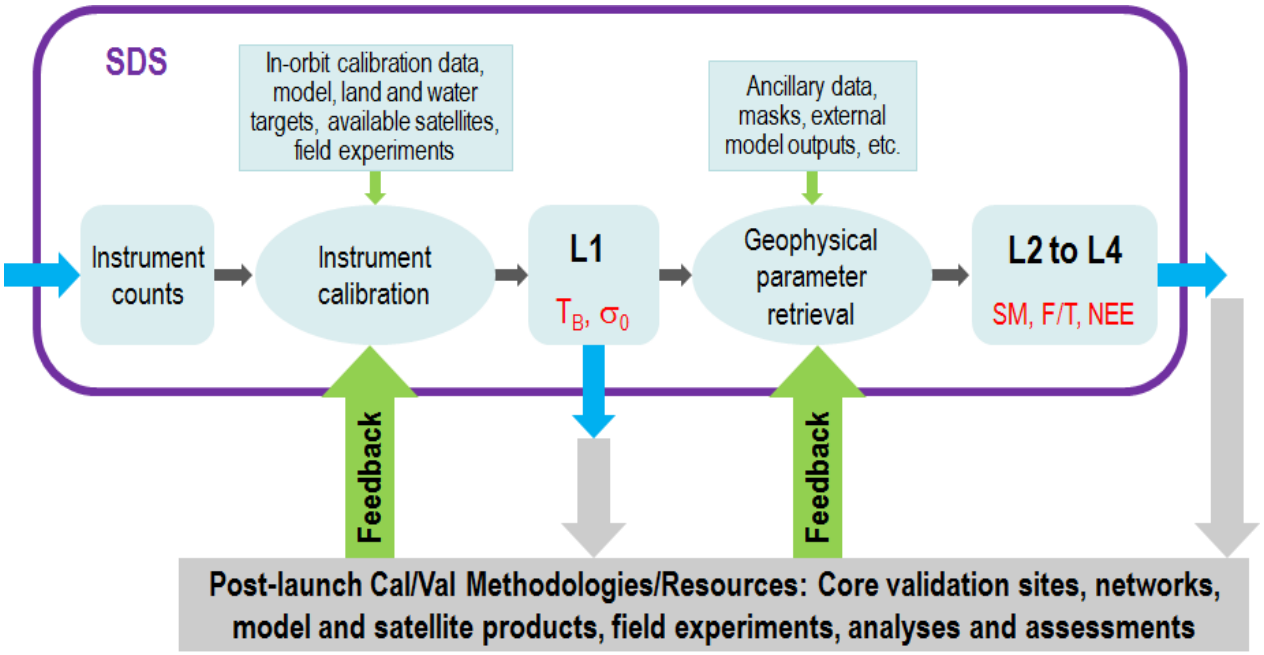



Figure 2.1. Overview of the SMAP Cal/Val Process.

The SMAP mission established a unified definition base in order to effectively address the mission requirements. These are documented in the SMAP Handbook/ Science Terms and Definitions [4], 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.

In assessing the maturity of the L3_FT_P[E] product, the Freeze/Thaw team 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 product release, the freeze/thaw team has completed Stage 4.

# 3 L3_FT_P[E] ALGORITHMS

The SMAP L3_FT_P[E] products are based a combination of two different freeze/thaw classification algorithms. A normalized polarization ratio (NPR) of SMAP V- and H-polarized brightness temperatures is used as the baseline algorithm, while a modified single channel V-polarized brightness temperature threshold algorithm (SCV) is used as an extended algorithm. Both the NPR and SCV algorithms use on a seasonal threshold approach for the FT classification. 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.

An overview of the L3_FT_P processing sequence is provided in Figure 3.1. Research using SSM/I radiometer and SeaWinds-on-QuikSCAT scatterometer data indicate substantial variability of freeze/thaw spatial and temporal dynamics derived from AM and PM overpass data, with important linkages to surface energy balance and carbon cycle dynamics[6][7]. L3_FT_P algorithm products generated utilizing both ascending (PM) and descending (AM) orbit radiometer data streams enable regional assessment and monitoring of diurnal variability in terrestrial freeze/thaw state dynamics.

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 (in orange blocks). The L3_FT_P product consists of daily composite landscape freeze/thaw state derived from the AM (descending) and PM (ascending) overpass calibrated brightness temperature (L1C_TB half-orbits). The L1C_TB data product contains gridded TB data on three 36-km EASE Grid projections: global projection, north polar projection, and south polar projection. The north polar projection and global projection L1C_TB records are used as inputs for freeze/thaw (FT) processing. The polar grid FT product is limited to land areas above 45°N latitude consistent with the original L3_FT_A product, but with an extended FT classification domain encompassing permanent snow/ice (classified by IGBP) affected land areas. The freeze/thaw polar grid classification domain covers regions of Earth's land mass where low temperatures are a significant constraint to vegetation productivity and terrestrial carbon exchange [8]. The global grid L3_FT_P product is limited to the global freeze/thaw domain defined from the heritage FT-ESDR products [9], which include all land areas where seasonally frozen temperatures influence ecological processes and land surface water mobility. The L3_FT_P product is provided on a 36 km Equal Area Scalable Earth grid version 2 (EASE-

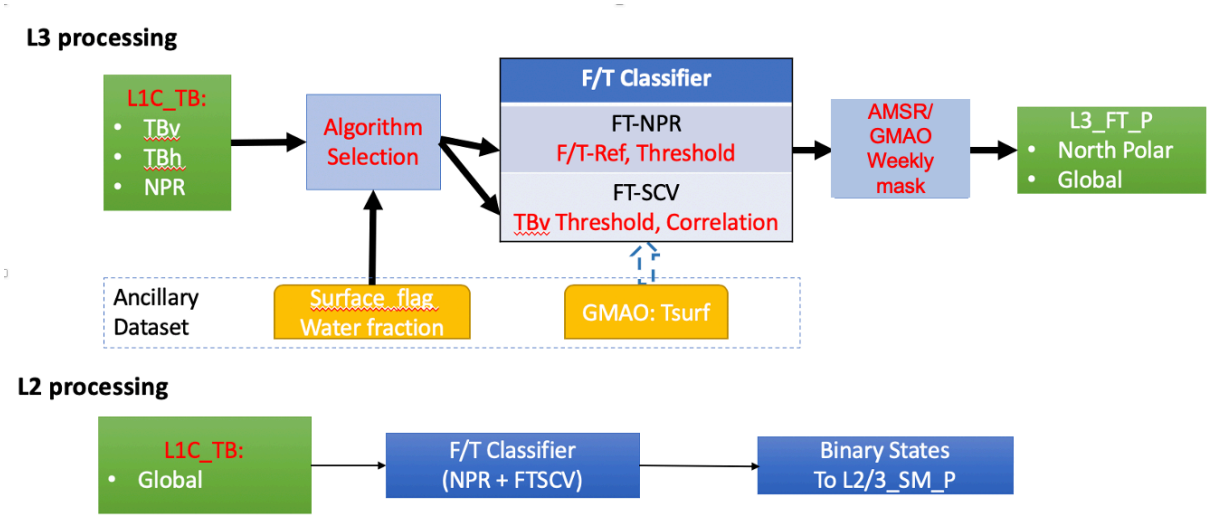



Figure 3.1. Processing sequence for generation of the L3_FT_P product and the binary freeze/thaw state flag for use in L2_SM_P

grid) format in both global and north polar projections, and daily time step. 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 (Urban is masked out). The resulting intermediate freeze/thaw products serve two purposes: (1) these data are assembled into global and north polar daily composites in production of the L3_FT_P[E] product, and (2) the freeze/thaw product derived from global L1C_TB granules provide the binary freeze/thaw state flag supporting generation of the SMAP L2 and L3 soil moisture passive products.

The L3_FT_P[E] baseline algorithm is a seasonal threshold method using the brightness temperature normalized polarization ratio (NPR). 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 [10][11]. 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 [12][13][14]. The NPR is most effective in areas with low to moderate vegetation cover and stable NPR reference states for seasonal frozen and non-frozen conditions. The NPR signal-to-noise is lower in dense vegetation areas (e.g. forests) due to more diffuse scattering of microwave emissions and reduced V and H polarization TB differences.

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:

$$NPR = \frac{TBV - TBH}{TBV + TBH} \tag{3.1}$$

A seasonal scale factor Δ(t) is defined for an observation acquired at time *t* as:

$$\Delta t = \frac{NPR(t) - NPR(fr)}{NPR(th) - NPR(fr)} \tag{3.2}$$

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:

$$\begin{aligned} \Delta t > T,\ Thaw \\ \Delta t \leq T, Freeze \end{aligned} \tag{3.3}$$

defines the thawed and frozen landscape states, respectively. This series of equations (3.1-3.3) are run on a grid cell-by-cell basis for unmasked portions of the FT domain. The output from Equation (3.3) 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.

There are two limitations for applying the NPR baseline algorithm. First, the algorithm relies on the proper references generally defined from winter frozen and summer non-frozen conditions. In the current scheme, the freezing references require at least 20 days of frozen conditions to set up. This requirement limits the southern boundary of the freezing reference. Secondly, the reference difference needs to be large enough to perform the algorithm (NPR>0.1), which excludes the dry areas when there is less changes in the soil dielectric constant. The extended algorithm FT-SCV (Freeze/Thaw algorithm using Single Channel TBV) is introduced to fill in the gaps where the baseline algorithm (NPR) is not valid in the global F/T

domain. The extended algorithm is a modified seasonal threshold algorithm that was first used in the latest (v4) FT-ESDR products [9]. It relies on the sensitivity of vertical (V) polarized brightness temperatures to freeze/thaw related shifts in land surface dielectric properties, which tend to dominate the seasonal TB signature in areas with a significant frozen season. The TB based FT transition is also correlated with surface temperature changes near the 0.0°C freezing point of pure water, so that surface air temperatures from global weather stations and reanalysis data have been used to identify frozen and non-frozen reference TB or dB conditions for the satellite microwave FT retrievals [15][16]. The threshold in the FT-SCV algorithm does not depend on the freeze and thaw reference derived from winter and summer periods. Instead, it exploits the pixel-wise linear relationship between TBV and ancillary surface air temperatures to define the TBV freeze/thaw transition point for each grid cell, which makes it suitable as an extension to the baseline NPR algorithm.

Therefore, the SMAP extended freeze/thaw single channel algorithm (FT-SCV) is defined as follows:

$$R > 0.5, \quad FT = \begin{cases} Thaw\,, \; if\; Tbv > threshold \\ Freeze,\; if\; Tbv \leq threshold \end{cases} \tag{3.4}$$

$$R < -0.5, \quad FT = \begin{cases} Thaw\,, \; if\; Tbv < threshold \\ Freeze,\; if\; Tbv \geq threshold \end{cases} \tag{3.5}$$

The freeze/thaw threshold in the FT-SCV algorithm is determined a priori on a per grid cell basis based on the empirical linear relationship between the SMAP TBV retrievals and daily surface air temperatures obtained from the GMAO global reanalysis; here, the GMAO was selected over other available global reanalysis products because it uses a similar GEOS-5 land model assimilation system as used in the SMAP Level 2/3 soil moisture products and Level 4 (L4) model enhanced products. The SMAP brightness temperature value that crosses the 0.0°C surface air temperature is defined as the freeze/thaw threshold for the selected grid cell in the FT-SCV algorithm. A global map of the correlation between the SMAP TBV and GMAO surface temperatures (Figure 3.2) is used to define the potential domain for the FT-SCV retrievals. Areas with stronger positive TBV and temperature correlations are assumed to have more reliable freeze/thaw thresholds and FT-SCV accuracy even though the brightness temperature retrievals are more directly sensitive to land surface FT dielectric shifts rather than temperature during the freeze/thaw transition period. The extended FT-SCV algorithm is only used for grid cells where the absolute value of the TBV and temperature correlation is larger than 0.5 and when the baseline NPR algorithm is not valid. Negative TBV and temperature correlation areas are associated with lake ice and surface water inundation

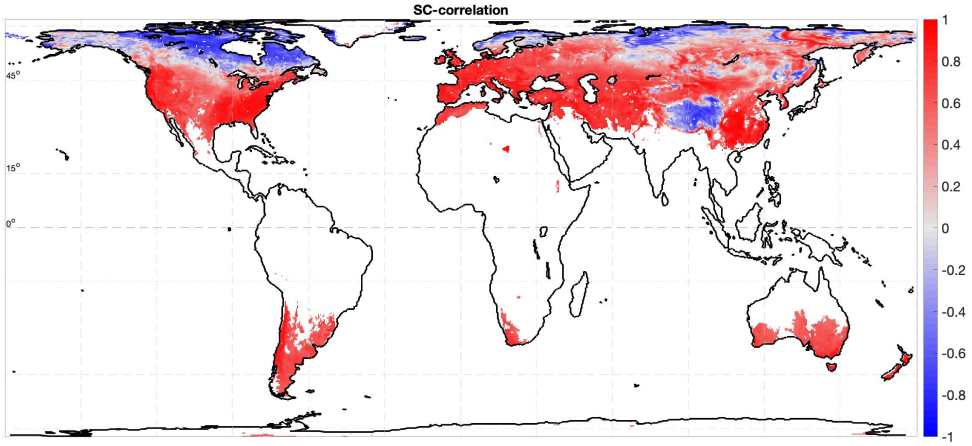



Figure 3.2. Correlation between the SMAP Tbv and GMAO surface temperature

effects located predominantly in areas with barren or low to moderate vegetation cover. In Figure 3.3, we illustrate the freeze/thaw domain for the L3_FT_P[E] baseline and extended algorithms.

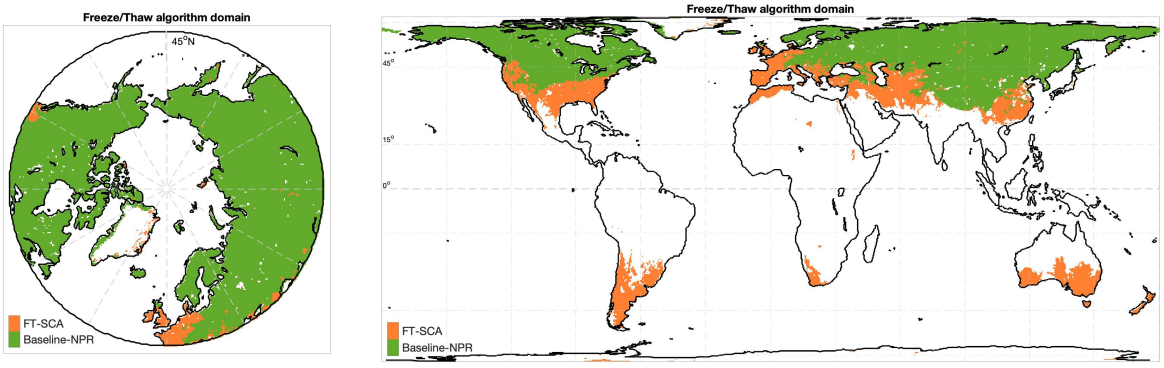



Figure 3.3. Freeze/thaw algorithm Polar domain (Left) and Global domain (Right) showing areas where FT-SCA and NPR algorithms are applied.

# 4 L3_FT_P[E] VALIDATION METHODOLOGY

The L3_FT_P[E] freeze/thaw product provides estimates of land surface freeze/thaw state expressed as a categorical (frozen, thawed, or [inverse] transitional) condition in north polar grid and global grid formats. Due to the different spatial coverage and shape distortion of the grid projections, we validate the polar grid product using all core validation sites and selected sites from the sparse networks using criteria based on site representativeness (uniform and representative terrain and land cover). The validation is based on reference freeze/thaw flags derived from both co-located air temperatures and soil temperatures corresponding to the local time of the descending and ascending satellite overpasses. The global grid product is primarily validated through grid cell-to-point comparisons with in situ daily surface air temperature measurement based FT observations from the WMO global weather station network; here, a single representative weather station is selected to represent FT conditions within the larger surrounding SMAP grid cell. Both the polar and global grid FT products are also used used for cross-product comparisons.

The computation of the classification accuracy proceeds as follows: Let $s_{AM/PM}(i,t) = 1$ if the L3_FT_P[E] product at grid cell $i$ (on the SMAP 36km / 9km 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_P[E] 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:

$$\delta_{AM/PM}(i,t) = \begin{cases} 0 & \text{if } s_{AM/PM}(i,t) = v_{AM/PM}(i,t) \\ 1 & \text{if } s_{AM/PM}(i,t) \neq v_{AM/PM}(i,t) \end{cases} \tag{4.1}$$

Note that a single L3_FT_P[E] flag is produced each day, but is derived from separate descending (AM) and ascending (PM) overpasses. The L3_FT_P[E] flags are therefore 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 is satisfied if (for both AM and PM overpasses together):

$$1 - \left( \sum_{i=1}^{N_i} \sum_{t=1}^{N_t(i)} \delta(i,t) \Big/ \sum_{i=1}^{N_i} N_t(i) \right) \geq 0.8 \tag{4.2}$$

Equation 4.1 is solved daily to provide instantaneous determinations of freeze/thaw spatial accuracy, using the available reference sites. The baseline mission requirement of 80% accuracy is 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 (e.g. wet snow versus dry snow), and to assist in determining the landscape components driving the radiometer 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.

# 5 R17 RELEASE PROCESS

This section describes refinement of the L3_FT_P[E] product following the final release of the passive product in May 2018 to the new R17 release in August 2020. The primary activities were updating the freeze and thaw references for the baseline algorithm and improving mitigation flags.

## 5.1 Freeze and Thaw References Updates

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. These pre-launch references (*NPR(th)*) were replaced with SMAP radiometer measurements. Originally, the 20 highest (lowest) NPR values from these periods were retained and averaged to create the thaw (freeze) reference based on the Aquarius analysis. Further analysis of the SMAP three-year dataset, enabled additional refinements to the FT reference generation to improve product accuracy and performance. In winter, there are places that only have a few days of freezing that do not meet the 20 days requirement for establishing a reliable freeze reference. In Figure 4, the light green area is the region where the freezing period is long enough to perform the reference threshold method using the NPR as an indicator. The boundary is set up by the freezing reference. For the thaw reference, we reassess three methods to reduce potential summer false flag situations, where the thaw reference is generated from: A) the highest 20 NPR values from the summer period (July to August); B) the highest 20 NPR values from the whole year; and C) the NPR average over the months of July and August. In Table 6, we demonstrate the false freeze and false thaw error estimates from the three methods. The false thaw error is defined as: 1. False Thaw (False T) where the NPR algorithm predicts thaw, while GMAO model surface temperature is less than -5°C; 2. False Freeze (False F) where the NPR algorithm predicts freezing, while GMAO model surface temperature is greater than 5°C.

Table 5.1. F/T false flag error estimate for three thawing references.

| AM | H20(summer) | H20(year) | Avg(summer) |
|---|---|---|---|
| | False T False F | False T False F | False T False F |
| NPR with weekly AMSR-mask | 1.11% 3% | 2.55% 2.2% | *1.98% 1.56%* |
| **PM** | **H20(summer)** | **H20(year)** | **Avg(summer)** |
| | False T False F | False T False F | False T False F |
| NPR with weekly AMSR-mask | 1.45% 2.36% | 2.57% 1.54% | *2.15% 1.3%* |

Overall, the NPR averaging over the two (July, August) summer months gives the best performance for both AM and PM freeze/thaw retrievals. In the current release, the two months averaging over the five years is used for the NPR thaw reference.

Data were separated by ascending and descending orbit. Because of differences in the seasonal evolution of L-band brightness temperature compared to radar, which has generally greater temporal variability and sensitivity to parameters such as soil moisture and vegetation phenology, the methodological

approach to NPR freeze and thaw references will be refined in future product releases. In addition, the reference values will be updated following each transition season. The current SMAP freeze and thaw NPR references are shown in Figure 5.1.

The NPR freeze/thaw retrieval threshold ($T$) is still fixed at 0.5. Threshold ($T$) optimization experiments are also in progress. Unique optimized thresholds can be determined for ascending and descending overpasses, and freeze-to-thaw and thaw-to-freeze transitions, by applying a linear fit to values of $\Delta t$ ($0.1 < \Delta t < 0.9$; see equation 3.3). The value of $\Delta t$ at the intersect of reanalysis surface soil temperature = 0 represents the optimized threshold. Optimization approaches will be evaluated using in situ measurements from the cal/val network in advance of future product releases.

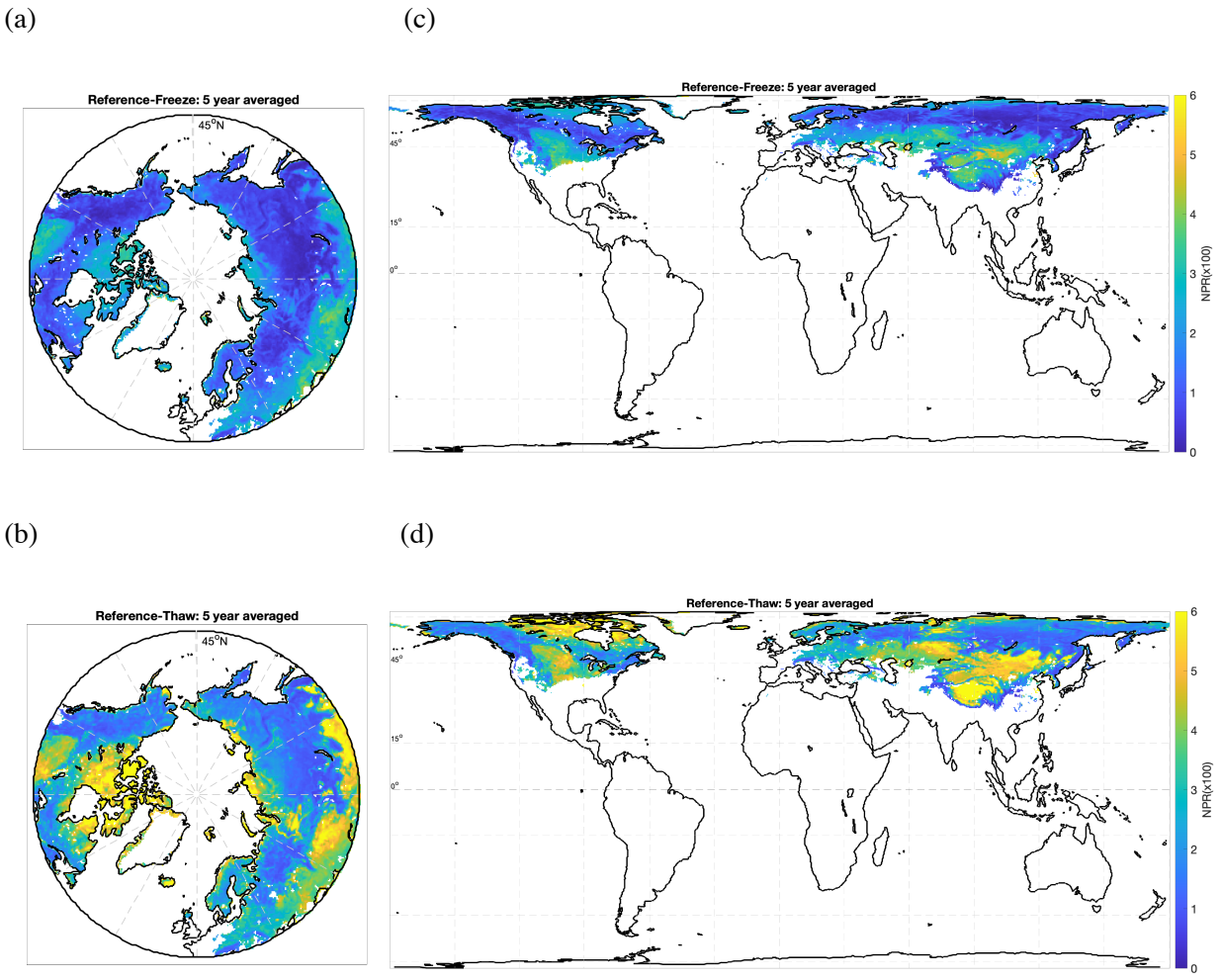



Figure 5.1. SMAP NPR algorithm (a) freeze and (b) thaw references in polar grid; NPR (c) freeze and (d) thaw references in global grid. Units are NPR scaled by 100 and averaged over five years (2015-2020).

## 5.2 False Flags Mitigation

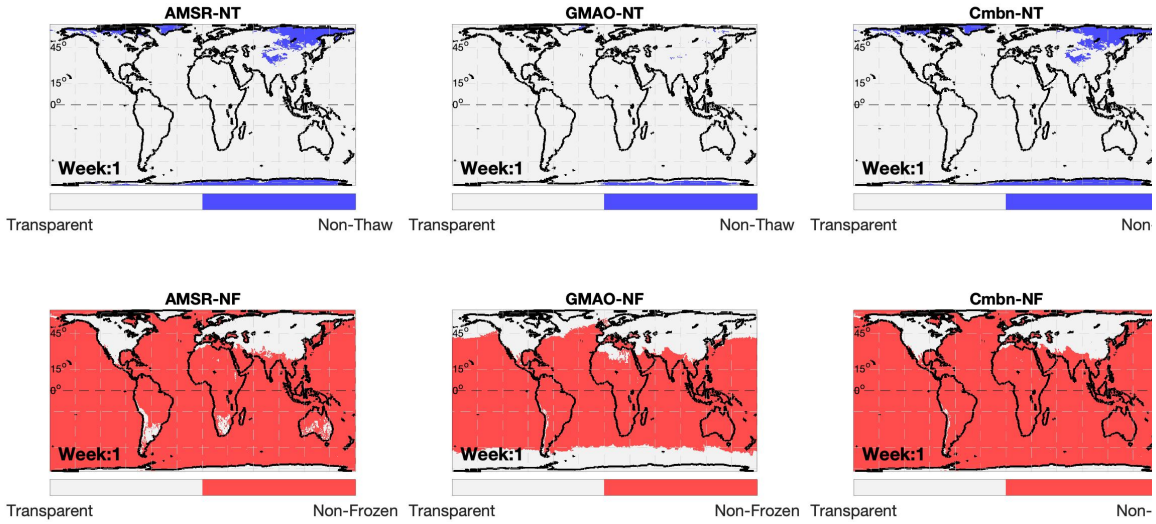
Following the pixel-wise determination of freeze/thaw state, two additional processing steps are applied to mitigate summer season false freeze and winter season false thaw retrievals. First, if the brightness temperature magnitude at either V or H pol is greater than 273K, the pixel is set to thaw regardless of the retrieval. Second, 'never frozen' (NF) and 'never thawed' (NT) masks (Figure 7) were calculated based on a long-term (2002-2019) daily AMSR (AMSR-E and AMSR2 36.5V GHz TB) derived freeze/thaw global record (using the approach described in [11]) over the 2002-2015 period and GEOS-FP surface temperature (2015-2020). These masks were then applied using a 31-day moving window approach to fix the retrieval state for each day of year (doy) for pixels that never changed freeze/thaw state during the AMSR record:

$$NeverFrozen(doy) = \sum_{i=doy-15}^{doy+15} Freeze\_AMSR\_flag(i)$$
$$NF\_mask = (NF == 0) \tag{5.1}$$

$$NeverThawed(doy) = \sum_{i=doy-15}^{doy+15} Thaw\_AMSR\_flag(i)$$
$$NT\_mask = (NT == 0) \tag{5.2}$$

While these additional processing steps do not remove all false flags, they substantially reduce obviously false flags without relying on other ancillary information. If we compare the example between the winter never thaw mask and summer never frozen mask in Figure 5.2, it's obvious that the summer mask almost covers the whole northern polar domain while the winter one only covers portion of the polar domain. This is due to the moving monthly average applied in the mask generation. To minimize the mask influence over the transition period, the monthly moving window is applied to generate the mask, which means that for a pixel to be masked out on certain day, requires ±15 days of the past 10 year AMSR FT prediction to be consistently frozen or thawed. The freezing duration over years has larger variability than the summer thawed season. Overall, the summer false flags have been more effectively removed than the winter false flags. On the other hand, we have brightness temperature as an additional screening during the summer. Therefore, higher accuracy for summer than winter is shown in the overall statistics for both polar grid and global grid validations in Section 6.

(a) Week 1 (Winter)




(b) Week 26 (Summer)

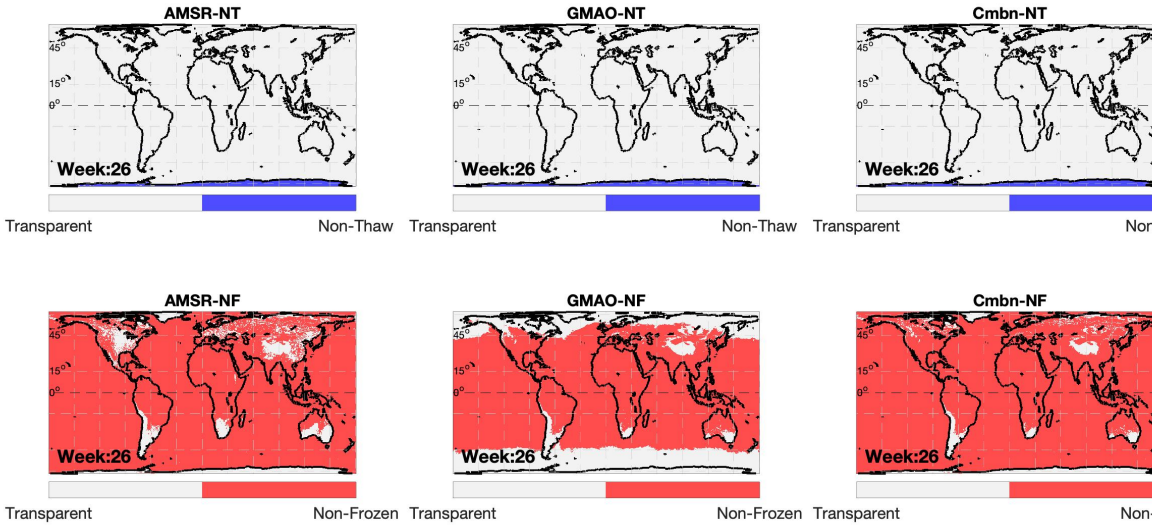



Figure 5.2. Example never thawed (left) and never frozen (right) masks for selected winter (calendar week 1) and summer (week 26) periods.

## 5.3 Extended FT-SCV Algorithm Setup

According to Equations (3.4) and (3.5), application of the extended SCV algorithm depends solely on defining a reference brightness temperature freeze/thaw (FT) threshold for each grid cell. The previous version of the FT thresholds was defined for each grid cell using an empirical linear regression relationship between the SMAP TBV retrievals and collocated GMAO surface temperatures obtained from the SMAP passive soil moisture products from the full year of 2016 data. The current version combined all the available data from 2015 to 2020. The FT threshold is the 0°C intersect between the SMAP TBV and GMAO surface temperature linear regression line. The AM/PM GMAO temperatures and corresponding descending/ascending overpass of the SMAP TBV retrievals are used in the regression. The threshold map is the same for AM and PM processing, but the retrieval algorithm is applied separately to generate the AM/PM freeze/thaw retrievals. The selection of the GMAO surface temperature is used for consistency with the SMAP soil moisture and L4 products, including the same gridding process. The FT thresholds are shown in Figure 5.3. There are alternative thresholds that can be defined from other reanalysis datasets (e.g. ERA-5) or other satellite land surface temperature products. The inter-comparison between different versions of thresholds requires a longer data record and is still under analysis. Threshold optimizations and updates will be included in future product releases.

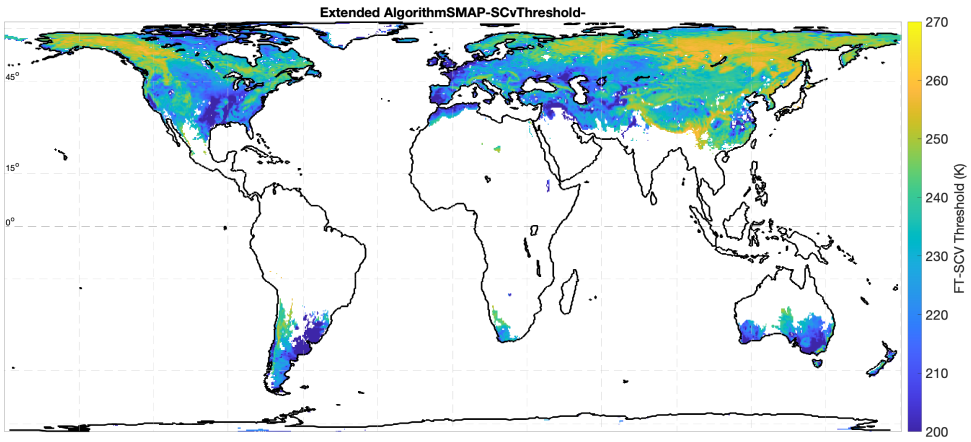



Figure 5.3. Freeze/thaw brightness temperature thresholds used for the FT-SCV extended algorithm; areas outside of the FT-SCV domain are shown in white.

## 5.4 Validated Release Testing

Testing of the L3_FT_P[E] algorithm code was conducted using the final release L1C_TB(_E) T16700 (April 1st 2015 through April 1st 2020). All of the analyses described in Section 6 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 binary FT image derived from SMAP radiometer measurements and the L3_FT_P polar-grid product is shown in Figure 6.1.

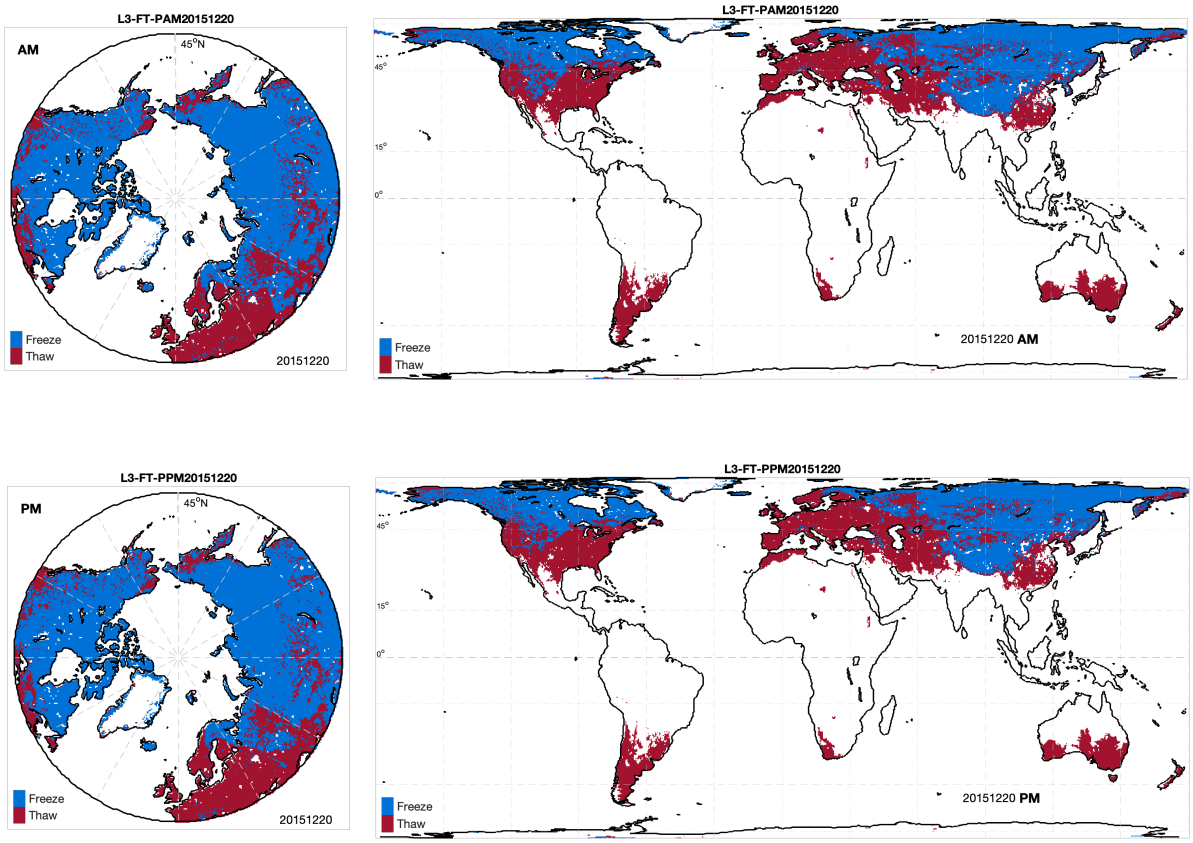



Figure 6.1 . Binary freeze/thaw retrievals from SMAP radiometer measurements for 20 December 2015 in polar (left panel) and global(right panel). Top panel for AM and bottom panel for PM.

## 6.2 Polar Grid Validation

As outlined in the SMAP mission Calibration and Validation (Cal/Val) plan the accuracy of SMAP products are primarily determined through comparisons with in situ measurements. For FT, validation sites are located across northern latitude (≥ 45°N) land areas and include core (multiple sub-grid cell measurements) and sparse (single measurement point) in situ station observation networks. Here, the polar grid product is used for the core sites and sparse network validation.

### *6.2.1 Core Sites Validation (CVS)*

The same format is followed for the presentation of results from each core validation site (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 same SMAP grid cell. Corresponding time series of FT flags are also provided, with in situ flags determined from surface air temperature (<= 0°C, frozen) measurements. For soil temperature, the forest sites (Berms sites, Sodankyle, Saariselka, Baie-James) usually have a surface organic layer and surface soil temperature from the probe does not drop below zero Celsius even when the landscape is effectively frozen. To avoid the uncertainty of the soil probe, we set the FT threshold for forest sites to be 1°C (<= 1°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 (light blue for frozen and light red for thaw) false freeze (SMAP = freeze, reference flags = thaw; bright blue), and false thaw (SMAP = thaw, reference flags = freeze, bright red).

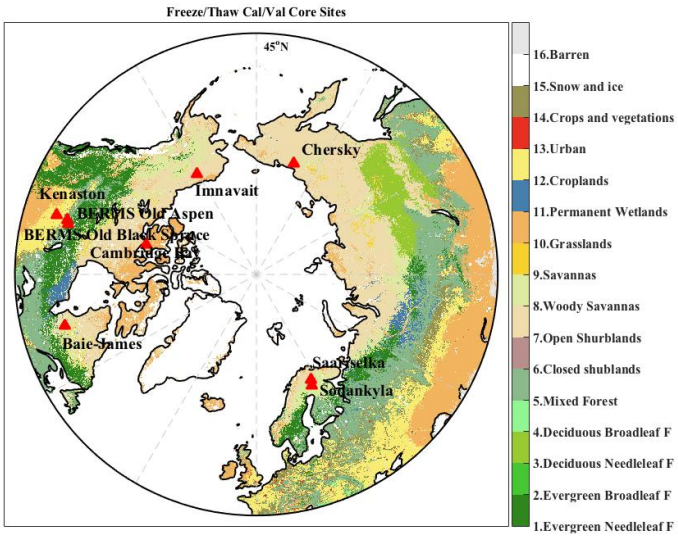



Figure 6.2. Core sites used for Freeze/Thaw validation.

Table 6.1 Summary of Freeze/Thaw 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)<br>53.99N; 105.12W (OBS) |
| Sodankyla | J. Pulliainen | Finland | Coniferous Forest | 67.36N; 26.64E |
| Saariselka | J. Pullianen | 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 |

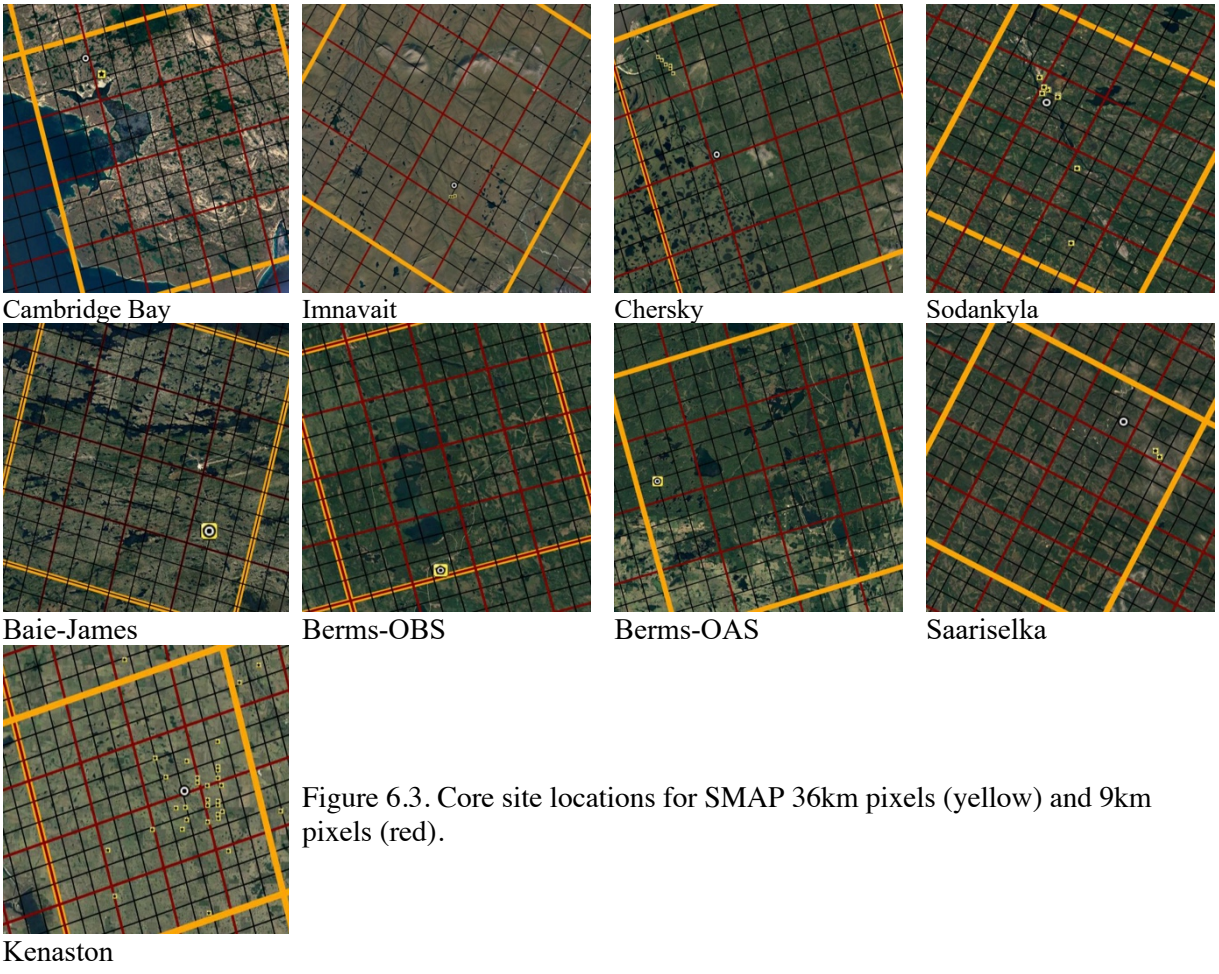
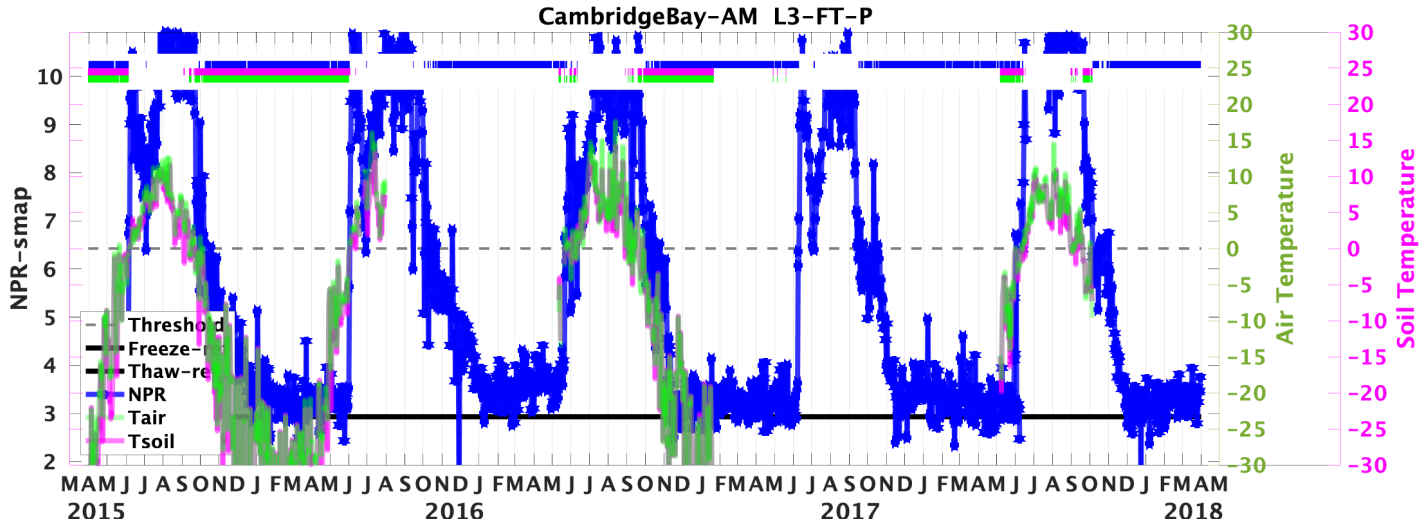


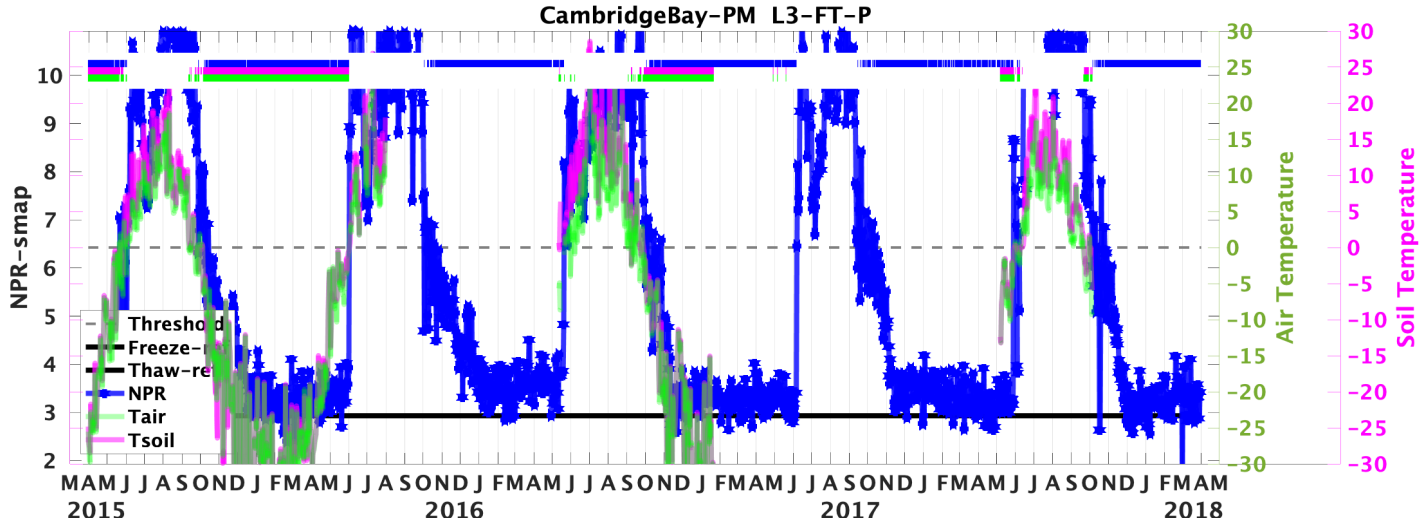

Figure 6.3. Core site locations for SMAP 36km pixels (yellow) and 9km pixels (red).

# I. Cambridge Bay

(a)




(b)



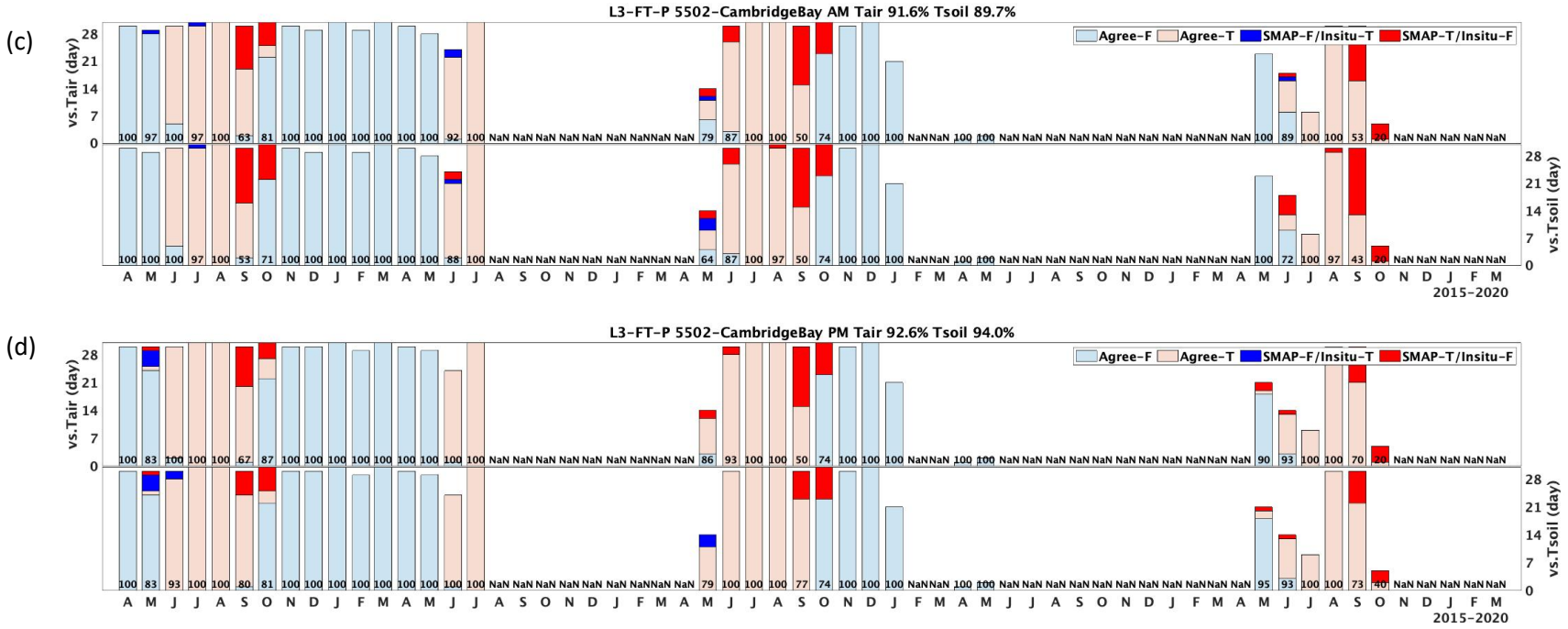
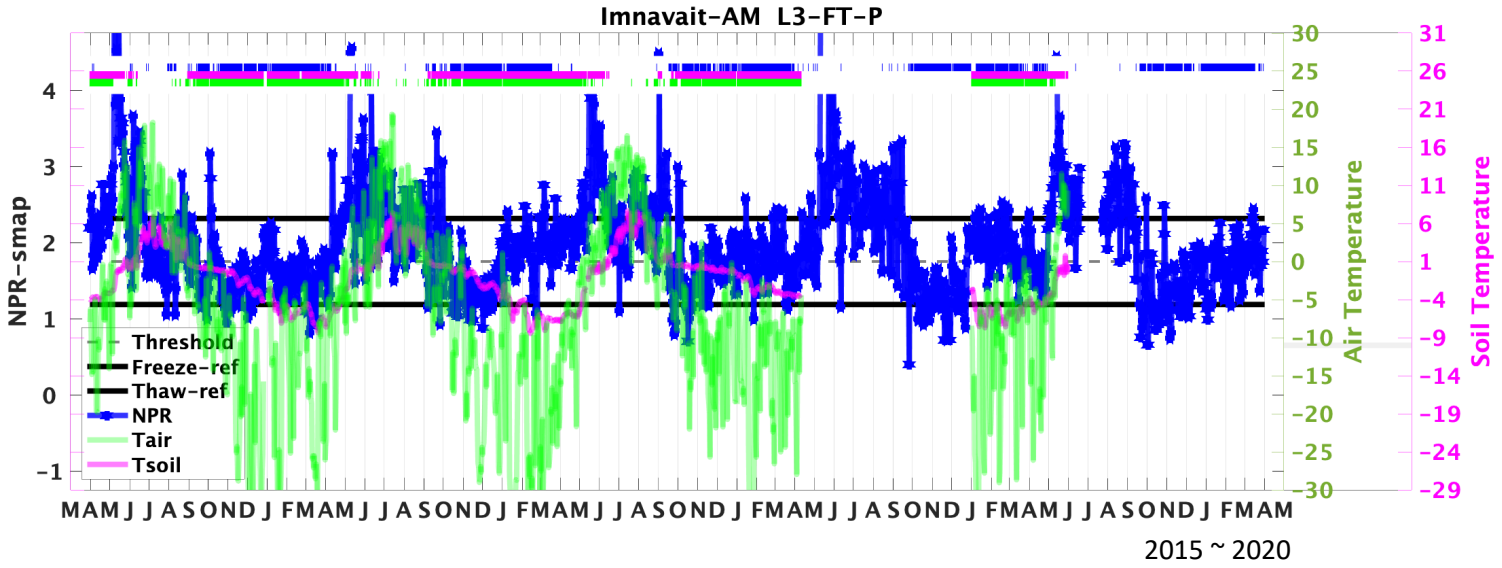



Figure 6.4. (a) Descending (AM) and (b) ascending (PM) overpass time series of NPR (blue), air temperature (green), and soil temperature (magenta) for Cambridge Bay. Horizontal lines on the top note the occurrence of frozen flags for SMAP derived (blue), air temperature derived (green) and soil temperature derived (magenta) conditions. (c) AM (d) PM Monthly summary of freeze/thaw flag agreement and error. The bars show the agreement and error days. The number in the bottom of each bar is the accuracy in percentage for each month.

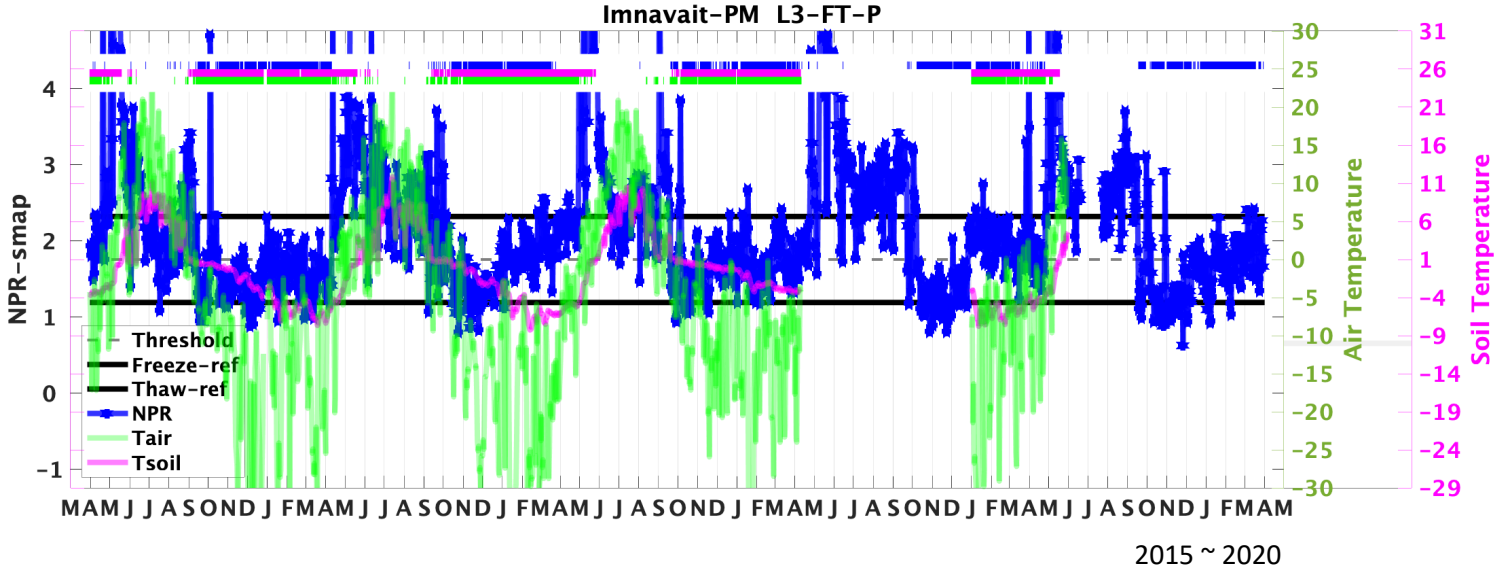
Cambridge Bay is a high latitude, open tundra site with shallow snow, and a long frozen season. As such, it presents a fairly straightforward case for FT detection. As shown in Figure 6.4 (a,b), there is a strong and clean NPR response to both freeze and thaw onset. Soil and air temperature are strongly 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 6.4c). Flag agreement is over 80% for most months and for both soil and air temperature reference flags except for the two fall freezing season conditions. The SMAP derived frozen condition is lagging behind the ground observations. This may be due to the relaxation of the pixel size in the passive product. The Cambridge Bay site is close to the coast. The 36km pixel used in the validation includes some portion of the ocean (Figure 6.3). The inhomogeneity of the land cover may delay the SMAP NPR signal in the fall since the ocean requires lower temperature and longer time to get frozen. The overall classification (shown in the title of Figure 6.4c,d) reflects the strong agreement at this site.

## II. Imnavait

(a)




(b)



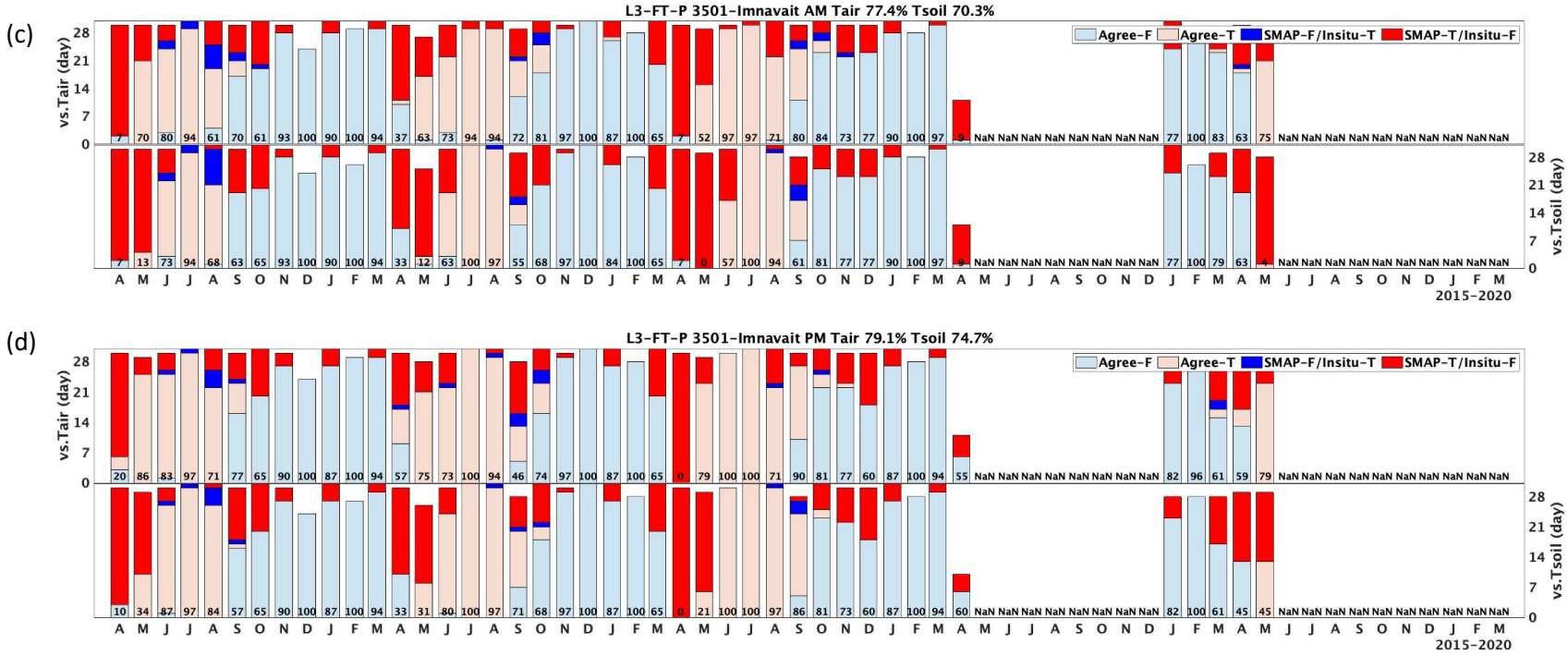
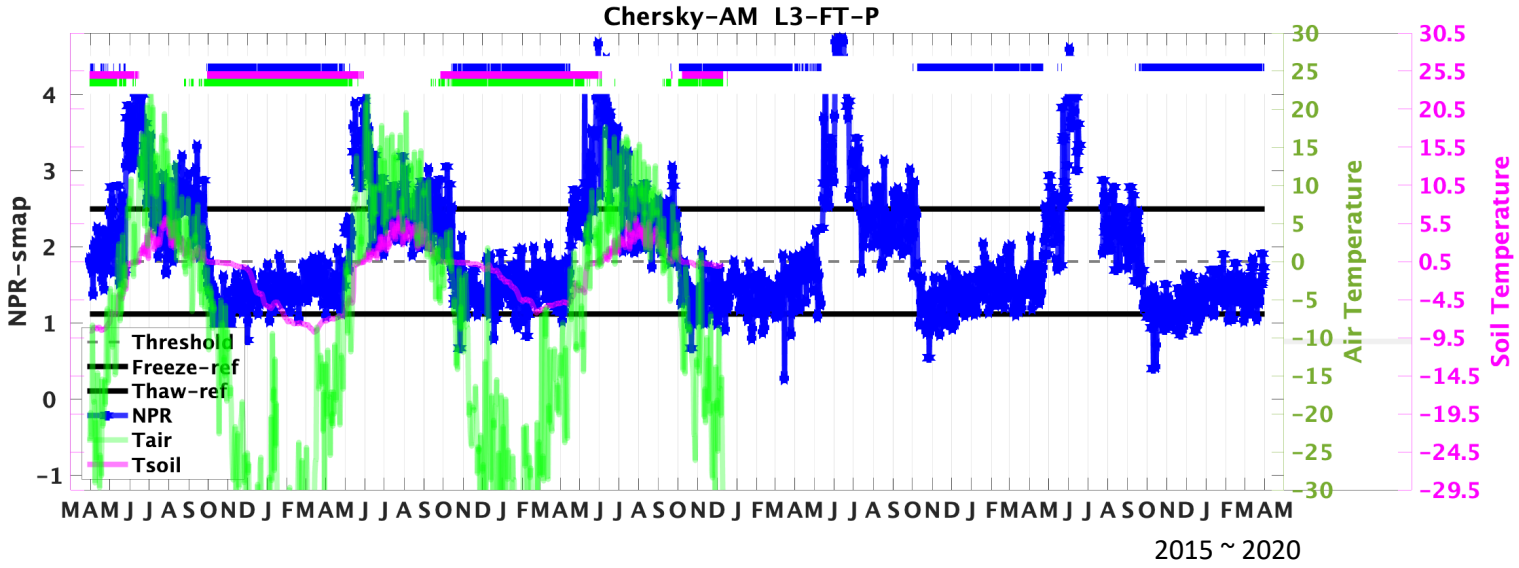



Figure 6.5. As in Figure 6.4 but for Imnavait, AK.

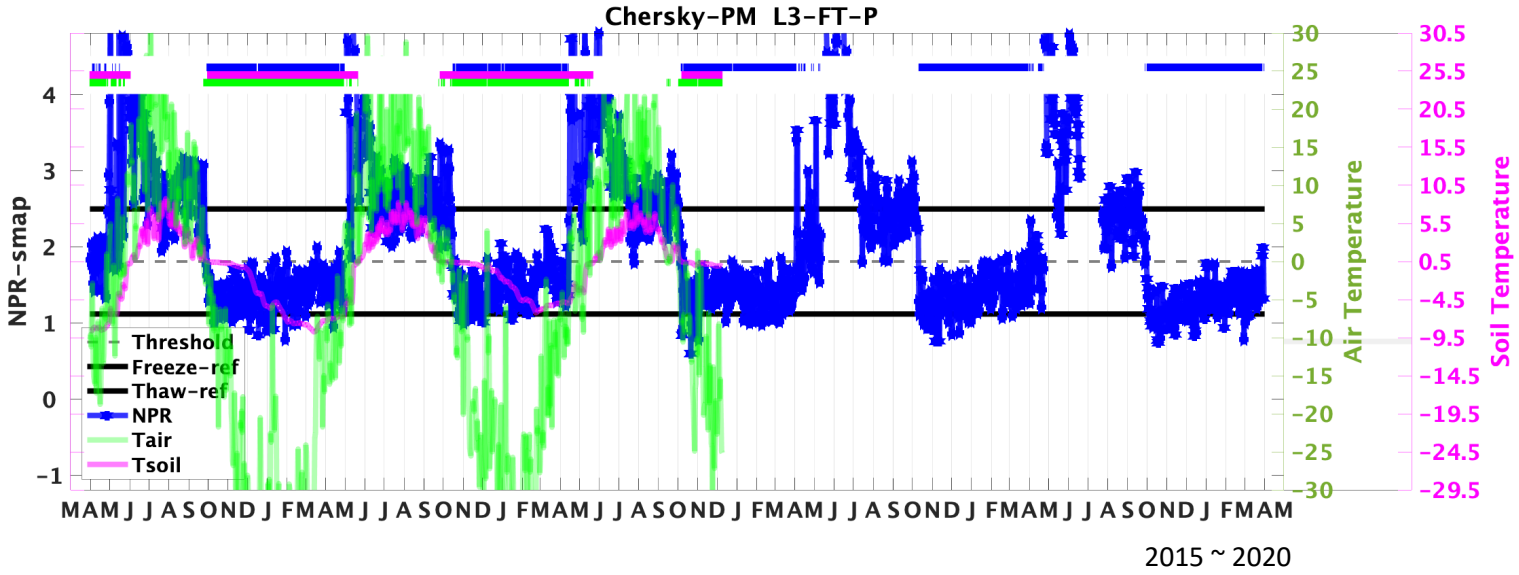
Imnavait is a cold climate, open tundra site similar to Cambridge Bay, but the reference difference between freeze and thaw values is very small and the NPR time series indicates a weak signal-to-noise ratio (Figure 6.5a and b). It's unclear if the weaker signal is due to complex terrain, a high lake fraction or other factor of this north slope site located in the foothills of the Alaskan Brooks Range. Despite the small reference difference, there is a strong NPR response to spring thaw, usually a few weeks ahead of both air temperature and soil temperature observations. This may be due to the lower value of the thaw reference derived from the summer months. The NPR values drop again very quickly after the initial thaw transition and lead to a fairly low thaw reference. This decrease in NPR magnitude combined with a highly variable NPR time series also results in periods of false flag events during the summer and winter. However, false thaw flags in the winter are effectively masked out by the climatology mask due to its higher latitude location but not the false freeze. During the fall, soil temperature lags the air temperature transition, leading to a period in September where air temperatures are below zero, the NPR response indicates frozen conditions, but Tsoil remains above zero.

## III. Chersky

(a)




(b)



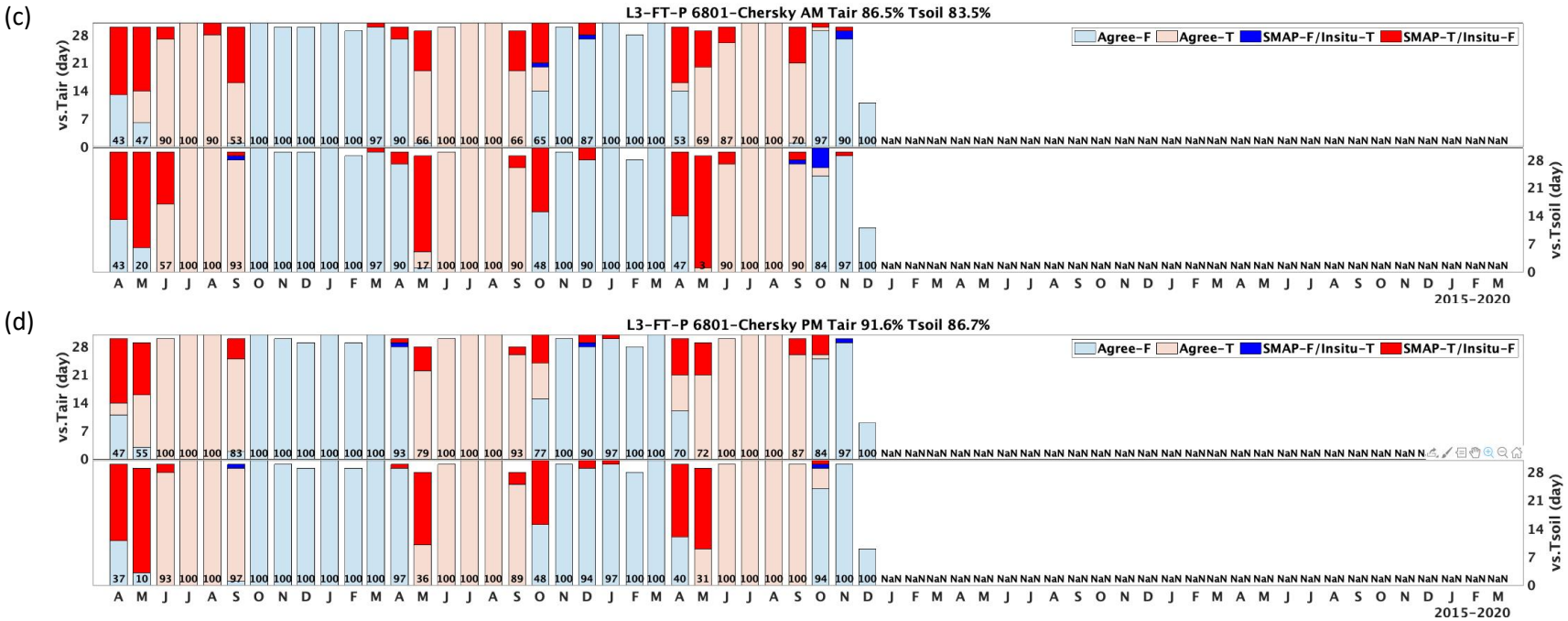
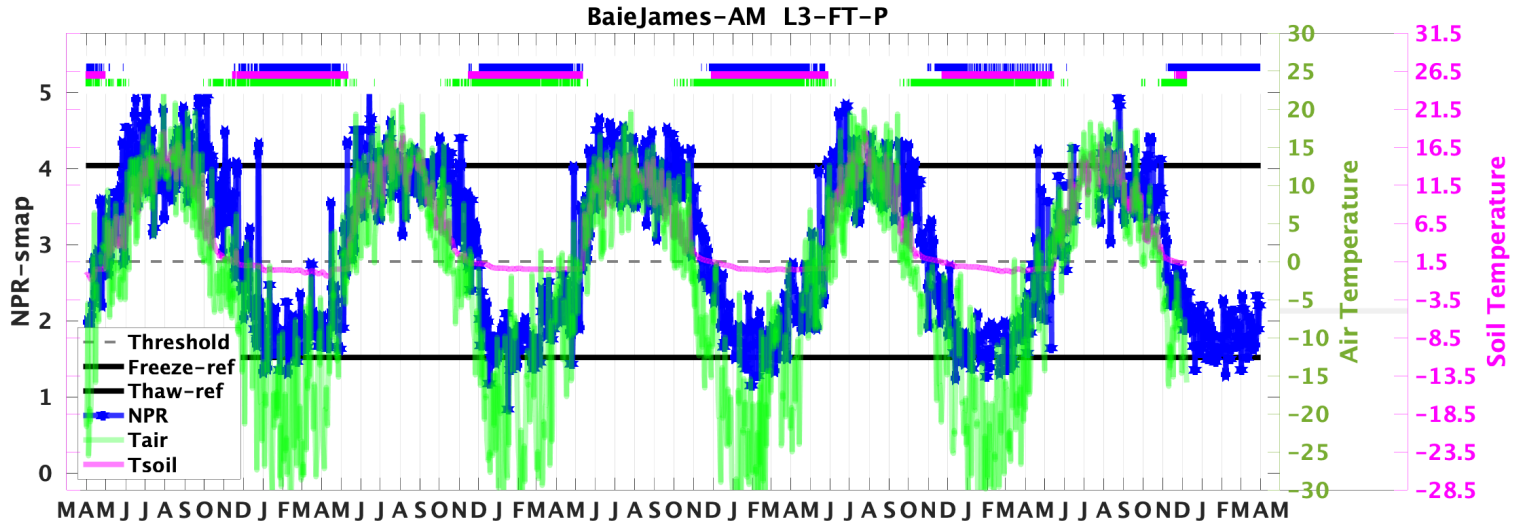



Figure 6.6. As in Figure 6.4 but for Chersky.

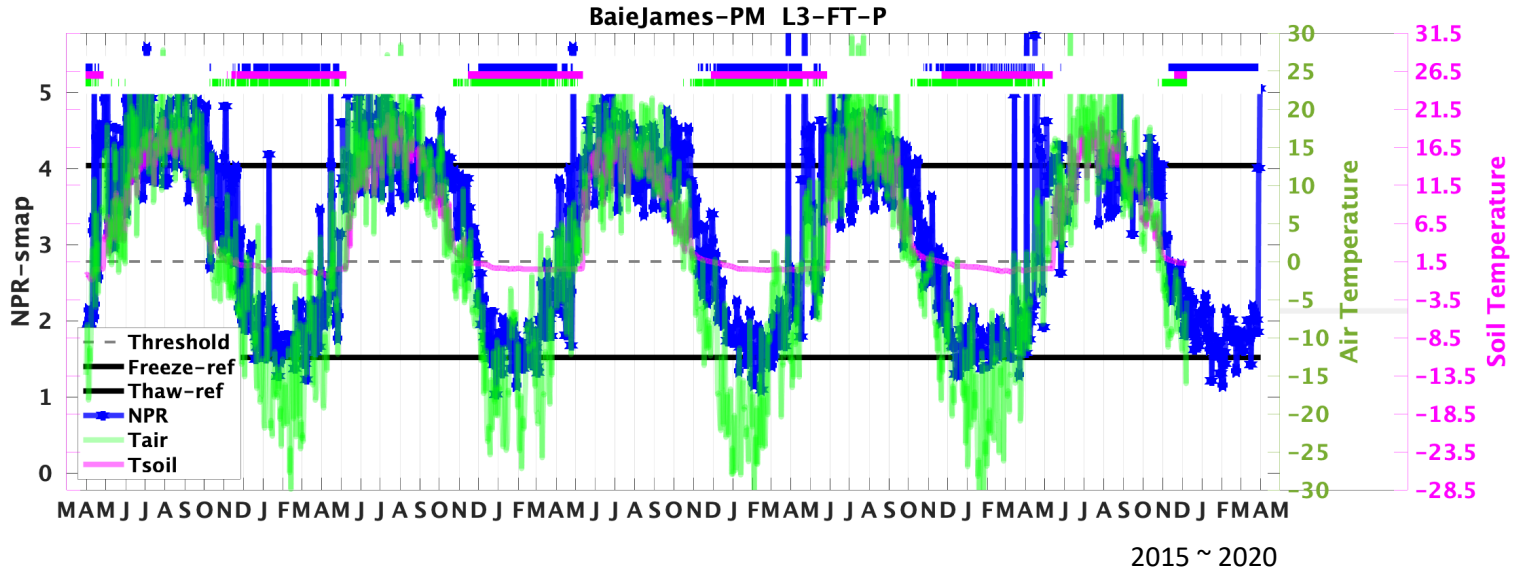
Chersky is a cold climate, deciduous needleleaf forest site in eastern Siberia. As will be seen at subsequent forested sites, Tsoil lags behind the Tair transition by a short period during the spring thaw, (Figure 6.5a and b). During fall, soil freeze lags behind freezing air temperatures much longer because of the insulative effect of snow and forest cover. At Chersky, the soil temperature stays around 0°C for about two months before dropping to colder temperatures. This uncertainty for FT flags derived from soil temperature is much larger than for surface temperature. Therefore, the SMAP FT retrieval agreement with air temperature is about 10% higher than for soil temperature.

## IV. Baie-James

(a)




(b)



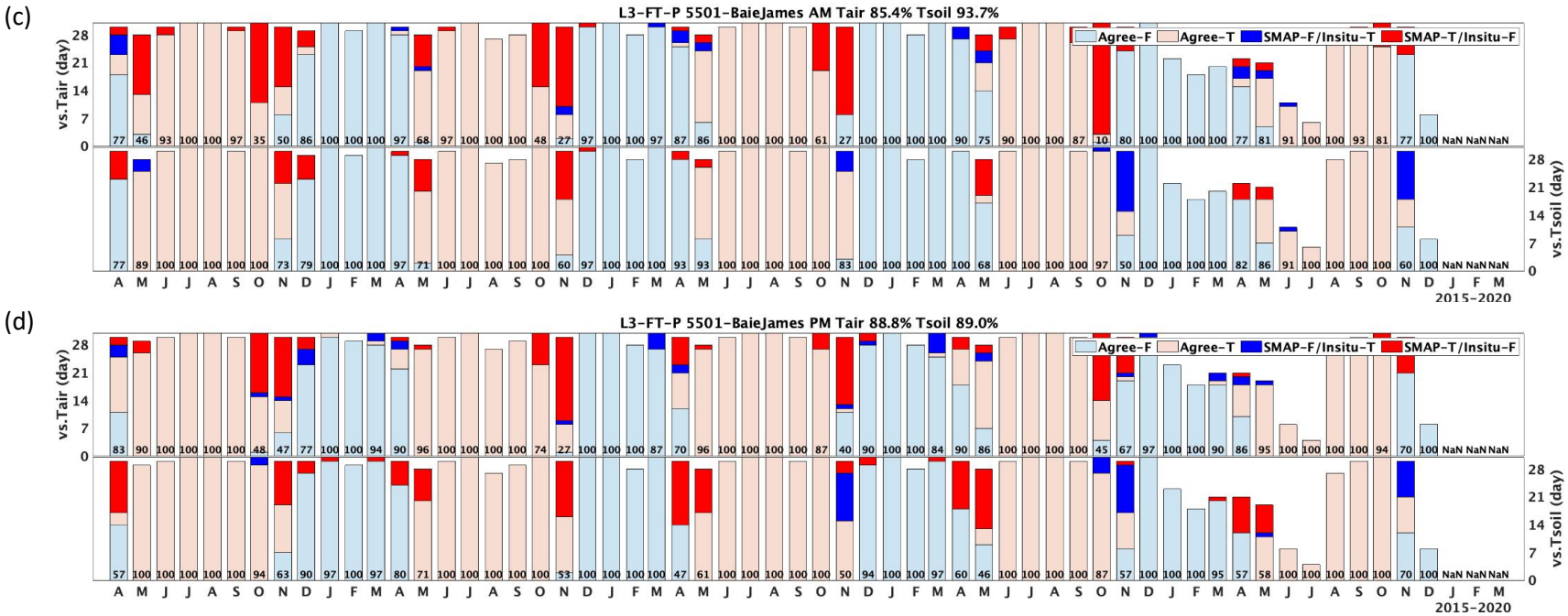
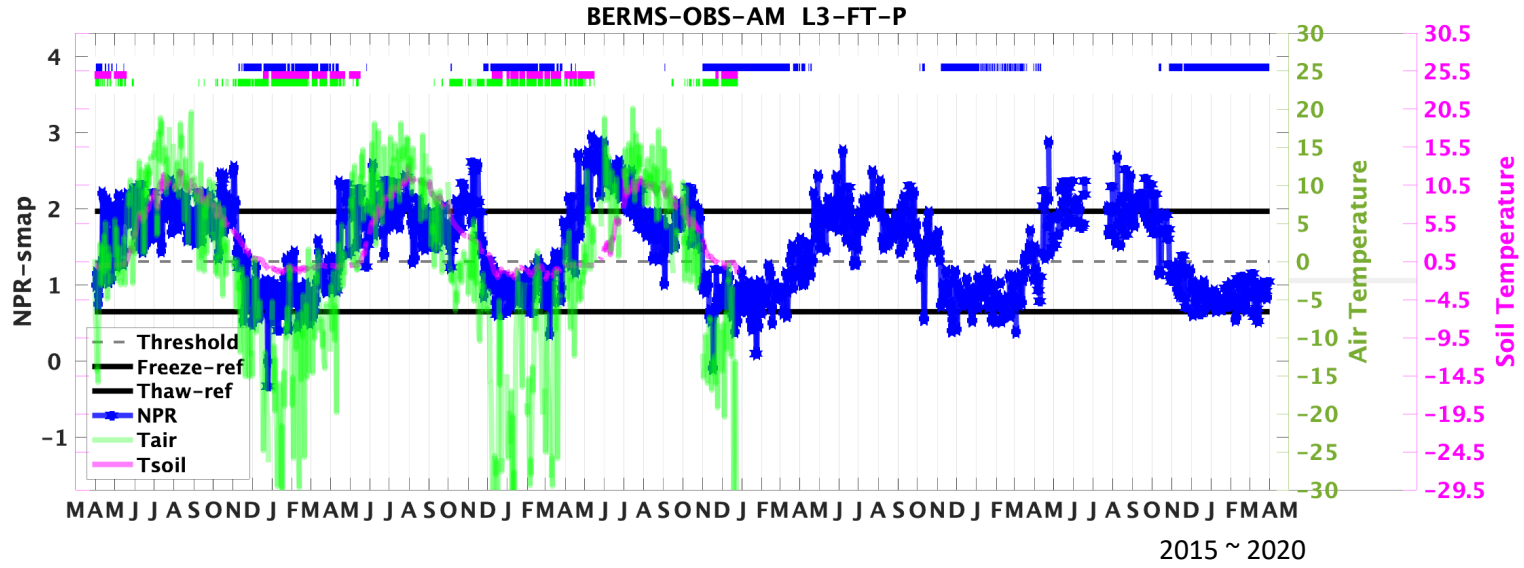



Figure 6.7. As in Figure 6.4 but for Baie-James.

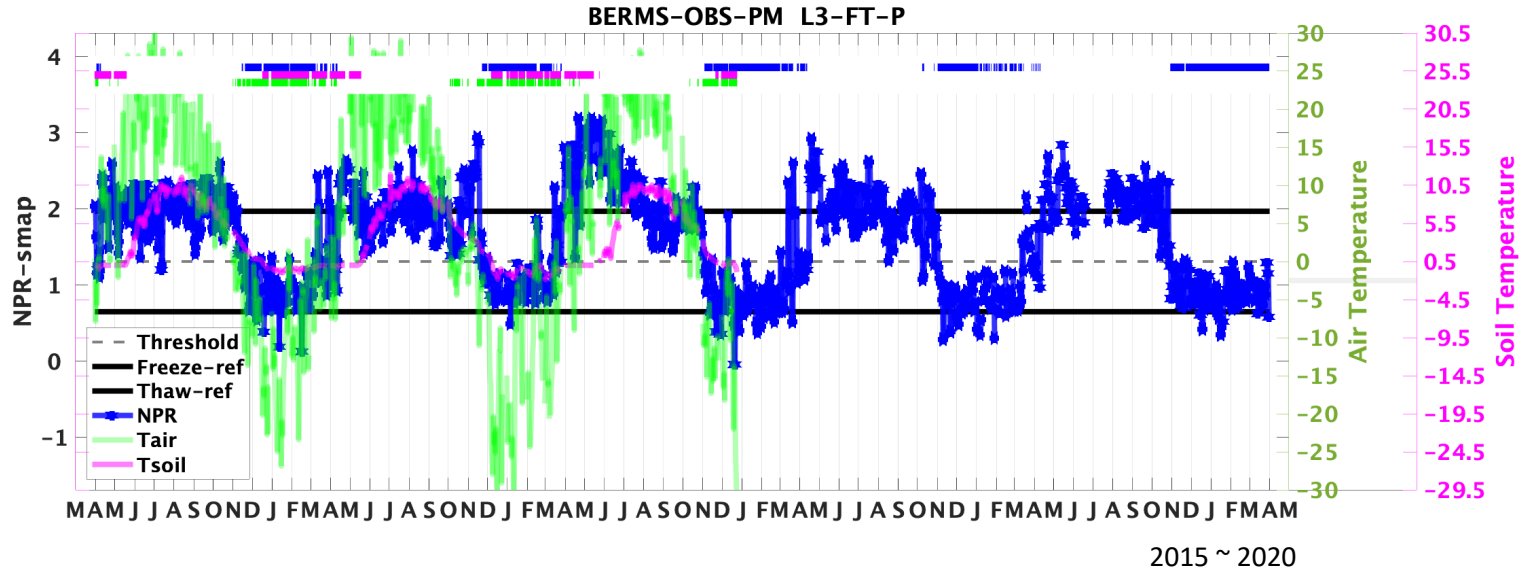
The seasonal evolution of NPR at Baie-James responds closely to Tair in the thawing season (Apr. Mar.) (Figure 6.7a and b). It reflects the NPR values being very sensitive to wet snow, while the soil is still frozen. In the fall, the soil temperature lags about three months behind surface air temperature freeze onset even with the Tsoil FT threshold adjusted to 1°C. 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. Here, SMAP FT results that are intermediate between the air temperature and soil temperature based FT observations may provide a more accurate condition overall.

## V. Berms-OBS

(a)




(b)



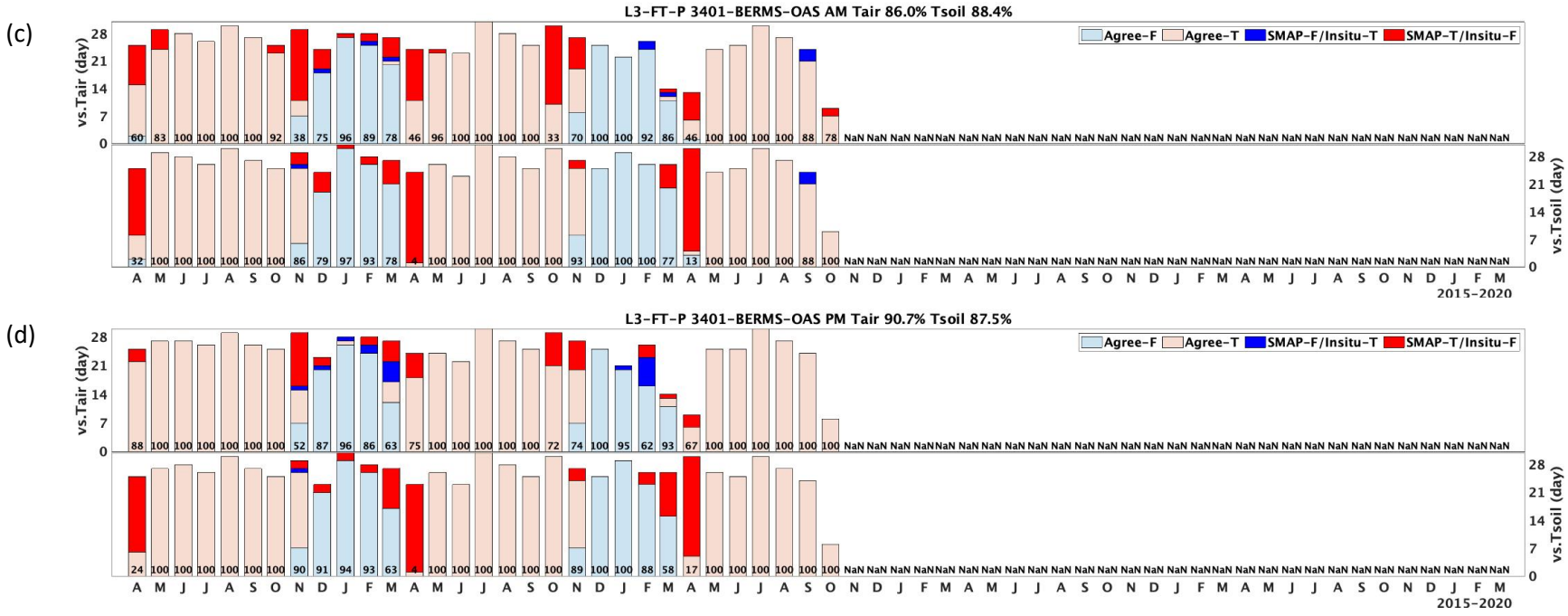
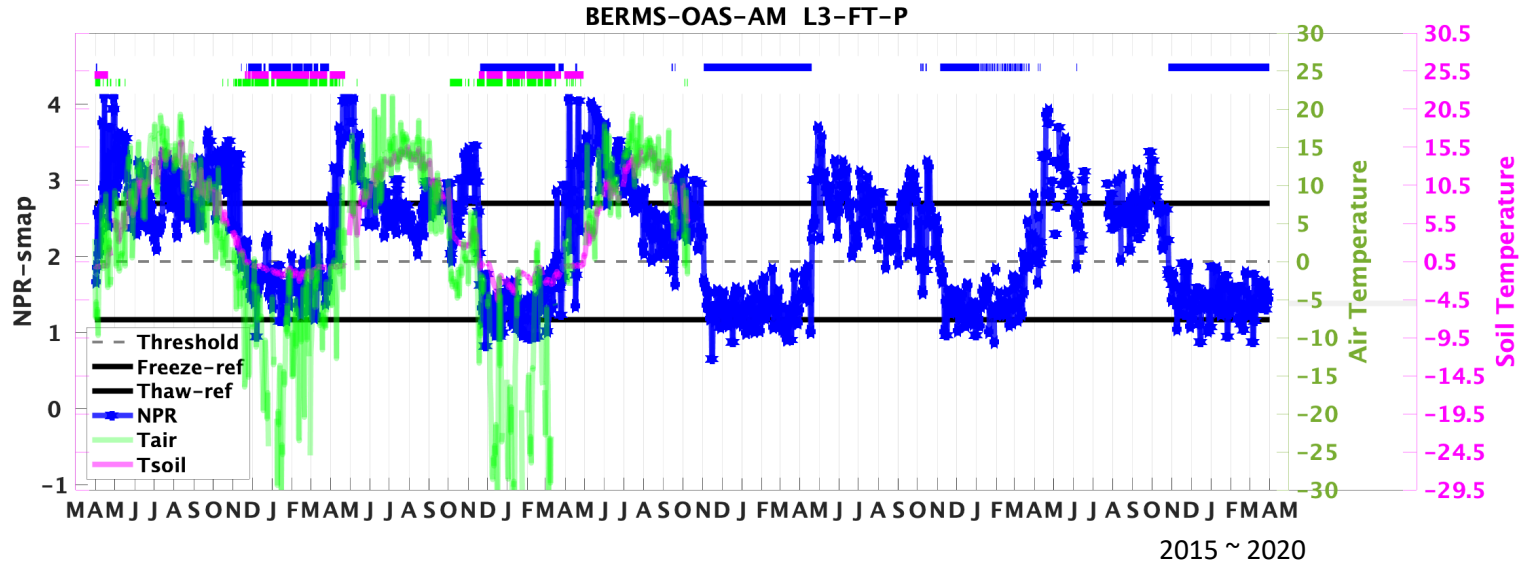



Figure 6.8. As in Figure 6.4 but for Berms-OAS.

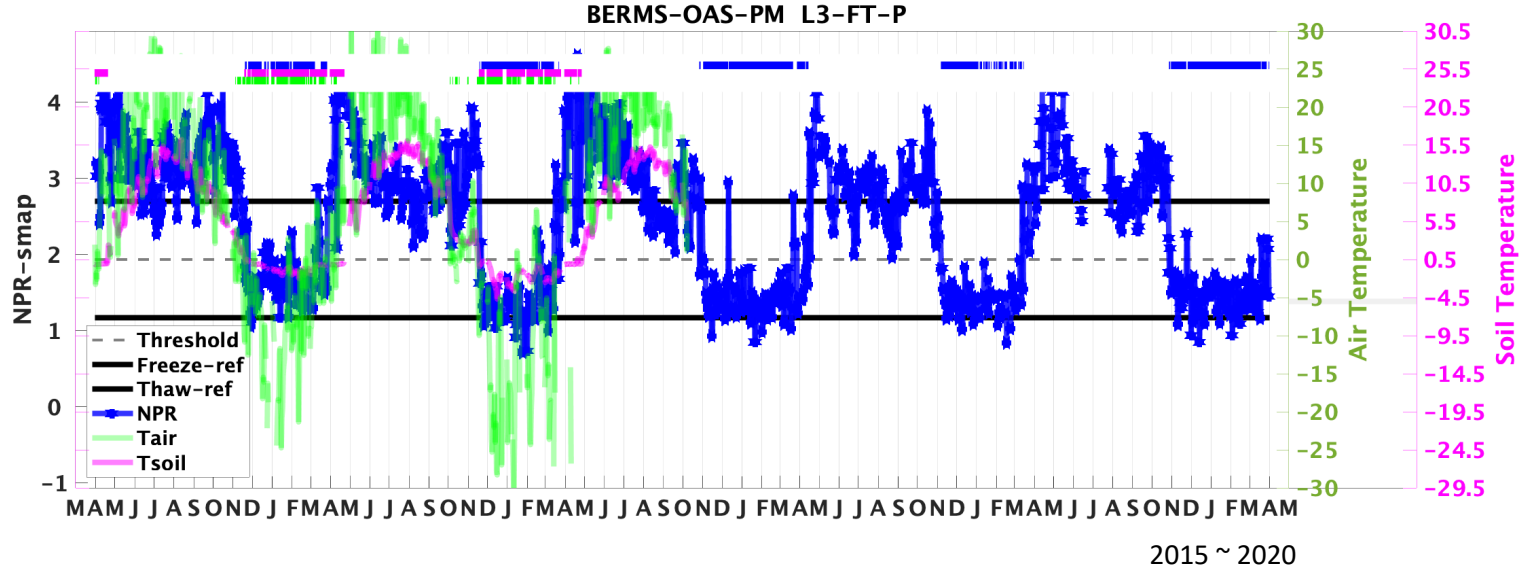
The BERMS Old Black Spruce (OBS) site has a small NPR reference difference and seasonal amplitude (Figure 6.8a and b). Despite this apparently weak FT signal, retrieval performance exceeds the 80% accuracy target at this site (Figure 6.8c), and summer false freeze events are minimal, though there are anomalous speckles over the winter that show up as a false thaw flag. The comparison of NPR with Tsoil and Tair observations at this site is typical of forest environments with snow cover. The insulative effect in the freezing transition and wet snow effect in the thawing transition create most of the difference between the three flags.

## VI. Berms-OAS

(a)




(b)



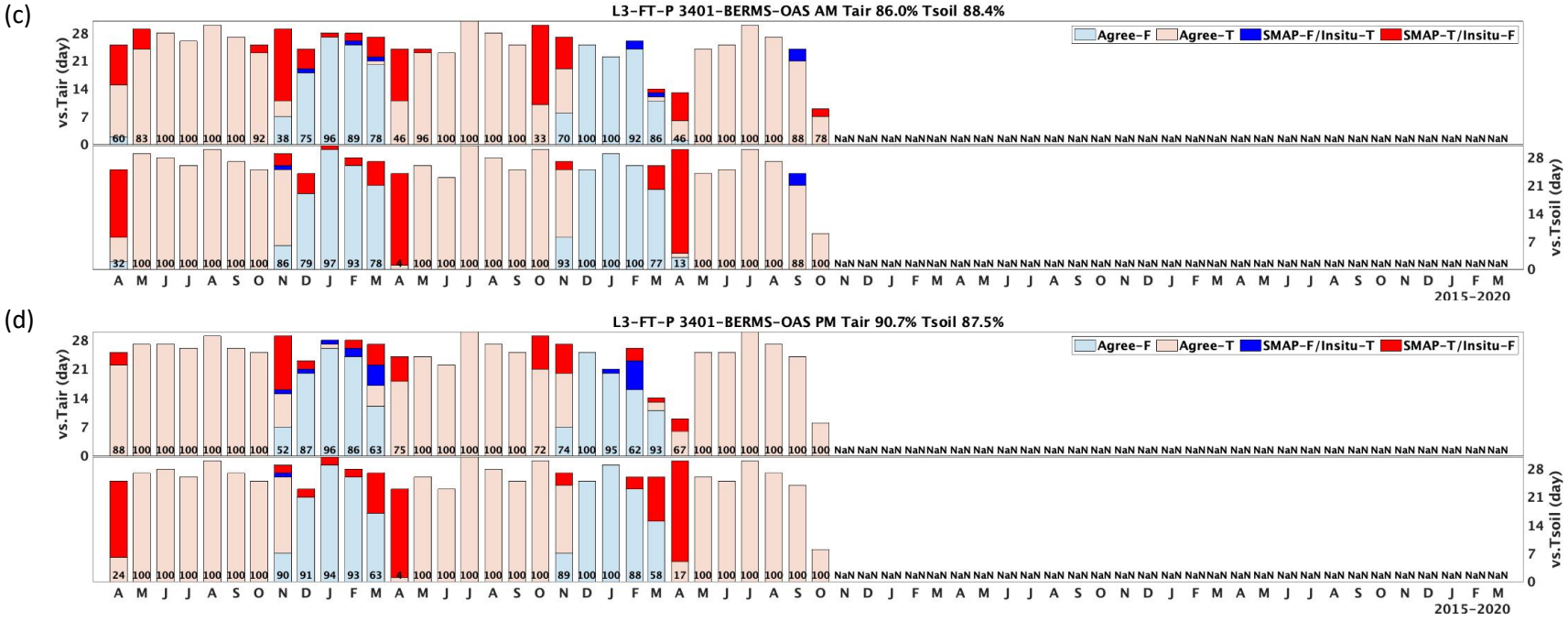
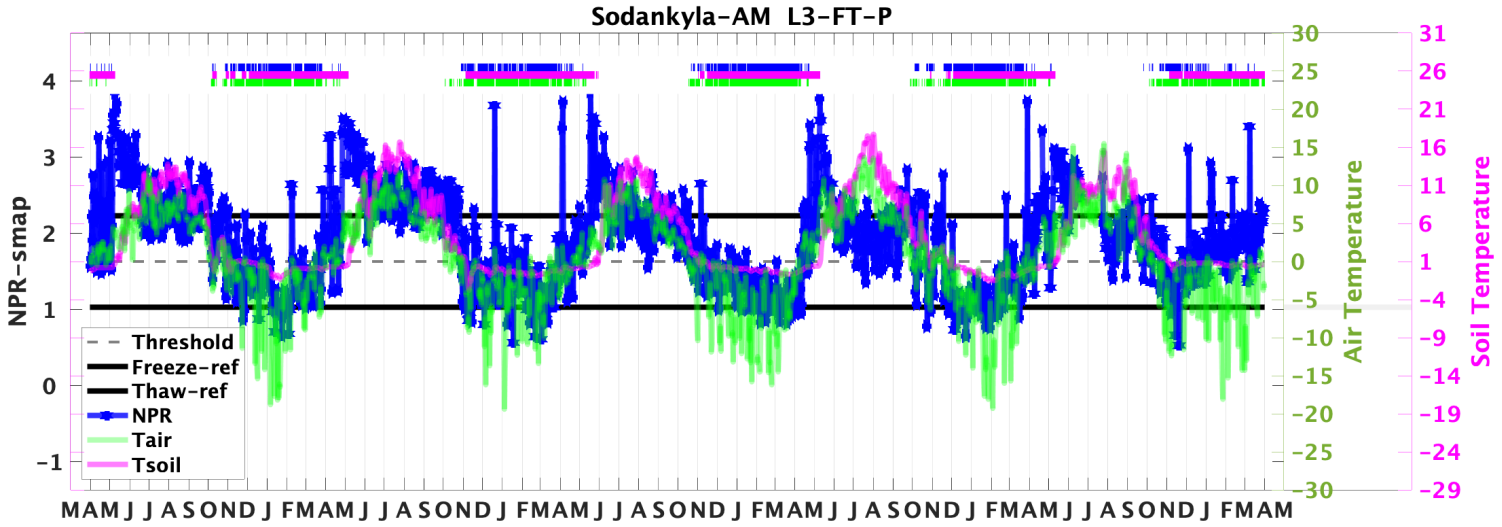



Figure 6.9. As in Figure 6.4 but for Berms-OAS.

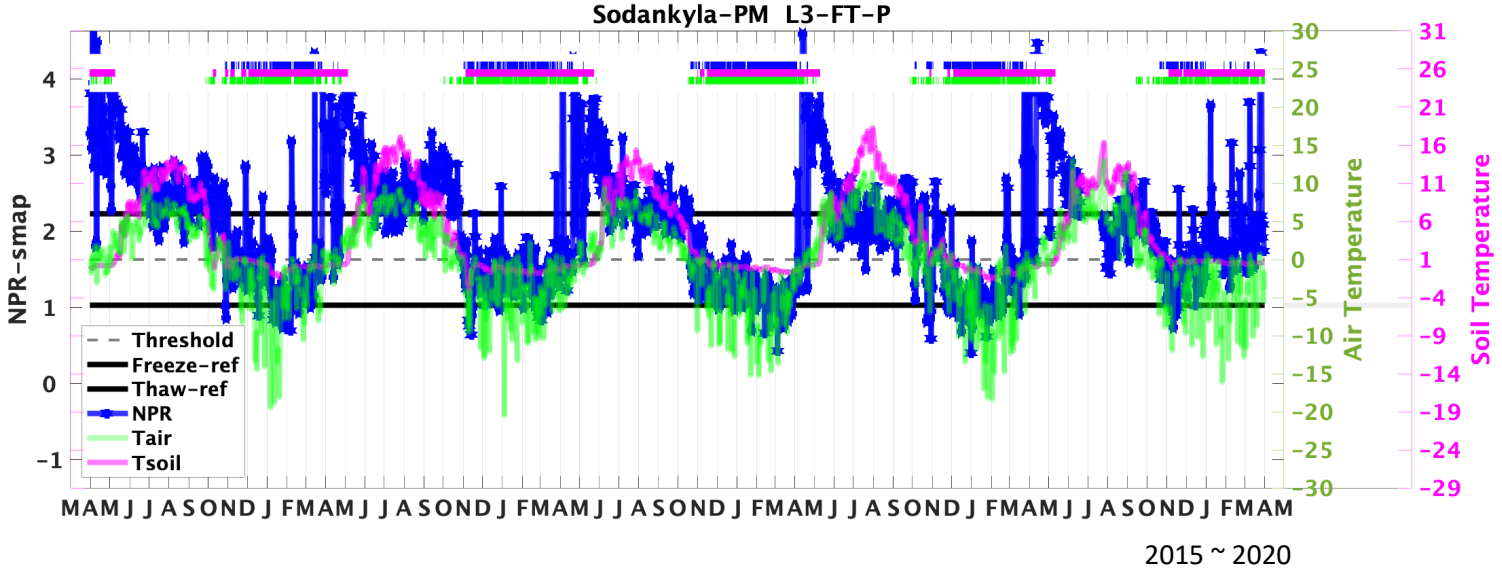
The BERMS old aspen (OA) site is in geographic proximity to the BERMS OBS site. The conditions are similar between sites with the exception of the forest type. Like OBS, the OA site has a small NPR reference difference and low seasonal amplitude, but a limited number of summer false freeze events (Figure 6.9a and b). Consistent with the other forest sites, there is a temporal lag of a few weeks in Tsoil going above zero Celsius following spring snow melt, which accounts for the majority of the classification errors. However, compared to the OBS sites, the OA site has more open area, therefore the onset of freezing for the three signals are aligned better due to the lower deciduous forest cover at this site.

## VII. Sodankyla

(a)




(b)



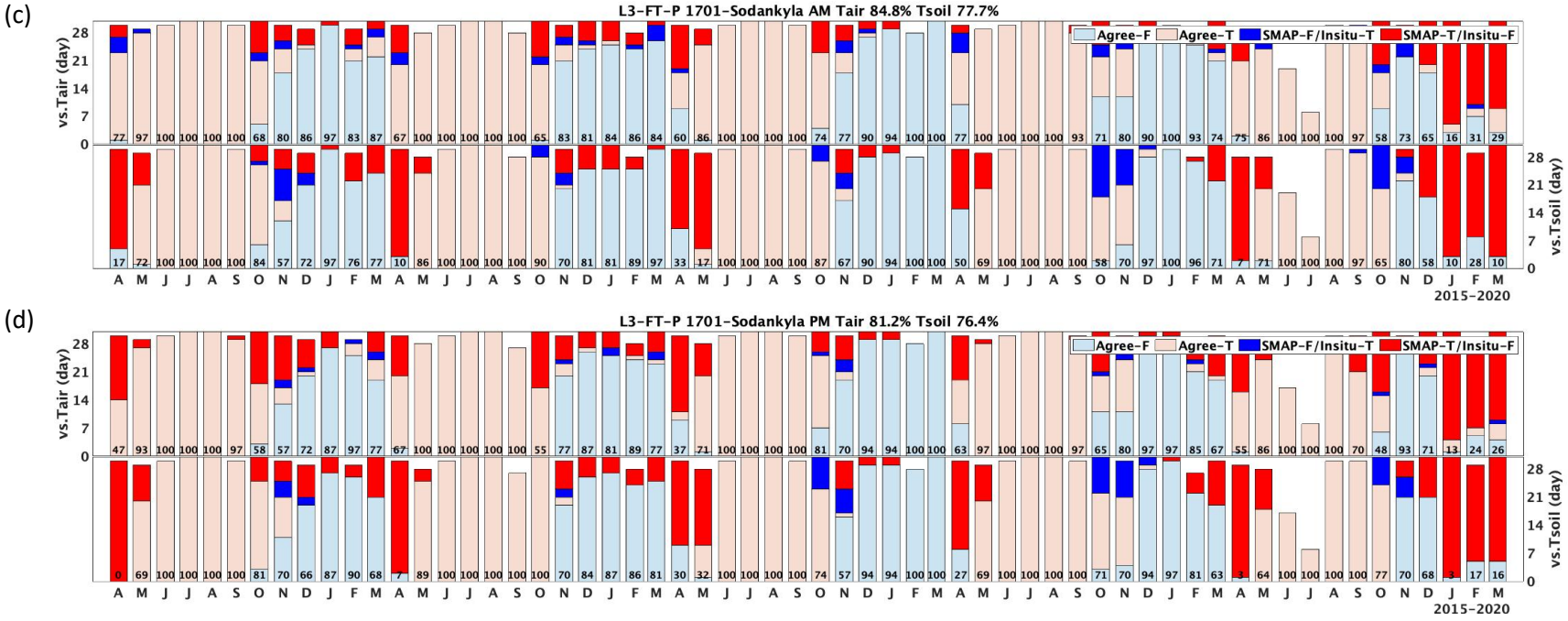
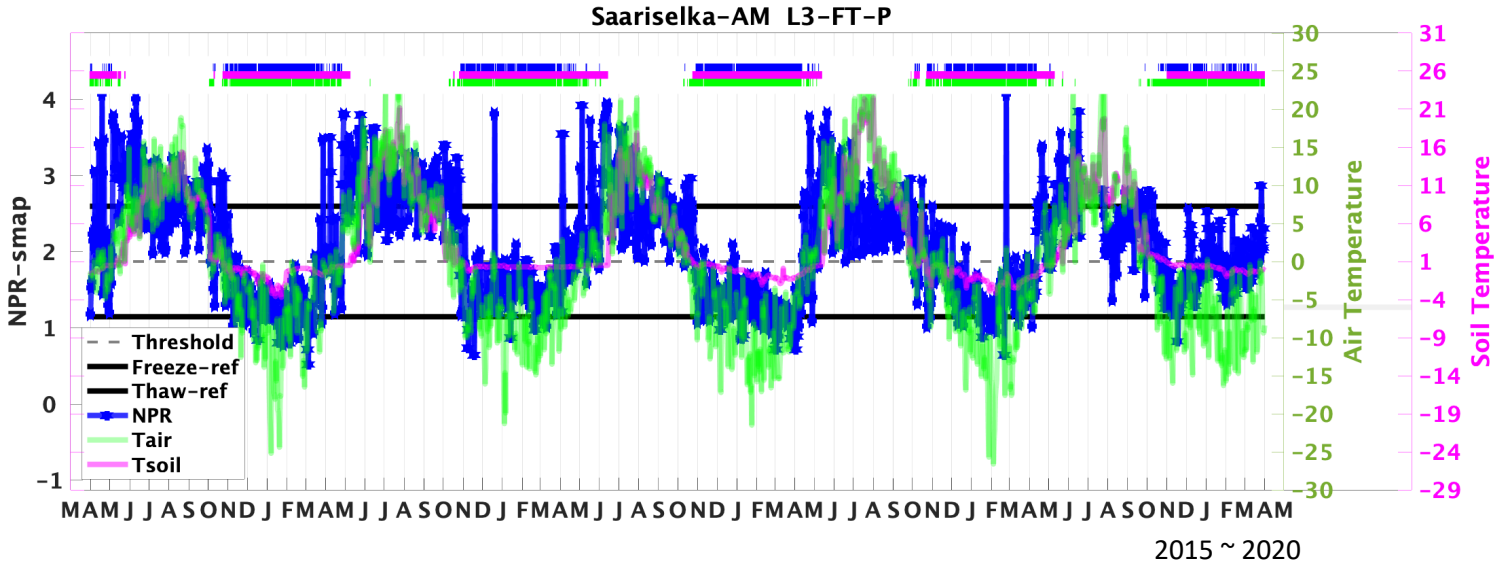



Figure 6.10. As in Figure 6.4 but for Sodankyla.

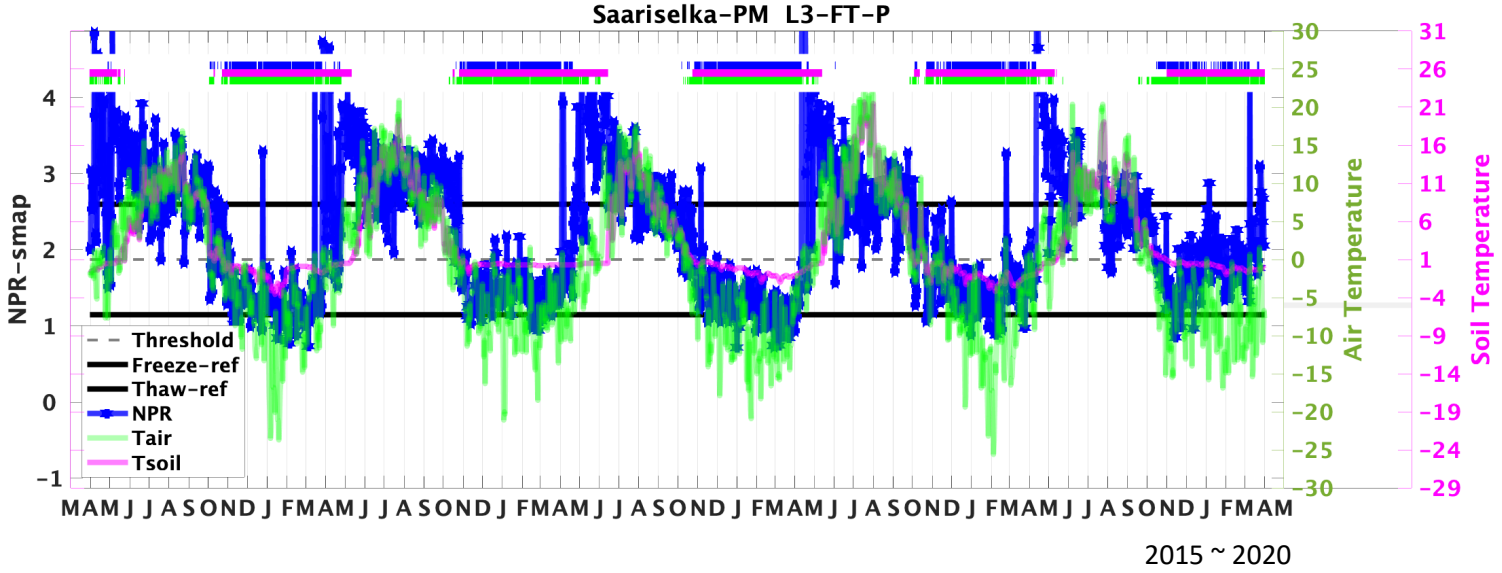
The Sodankyla site is located in a forested region with a high fraction of wetlands. (Figure 6.10a and b). These results are very similar to the BERMS sites and reflect strong overall FT agreement with Tair derived reference flags, but poor agreement with Tsoil during the early and late frozen periods.

## VIII. Saariselka

(a)




(b)



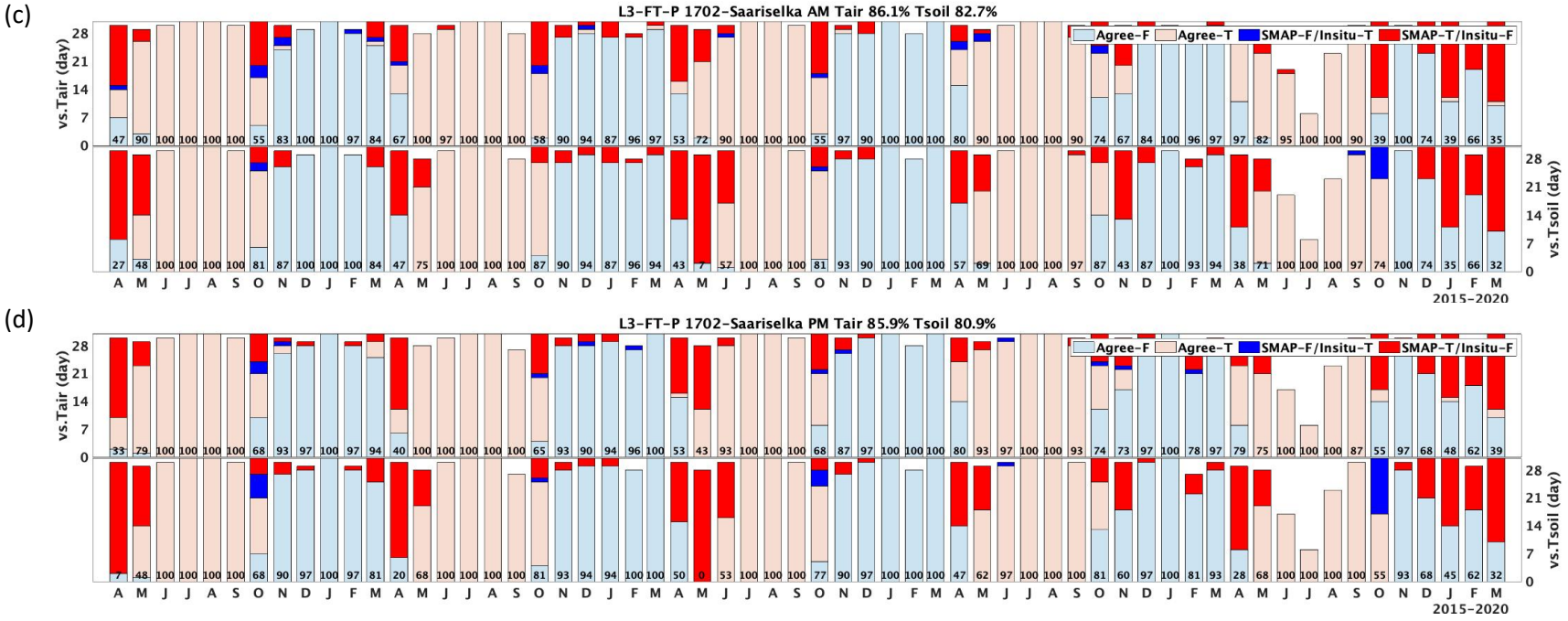
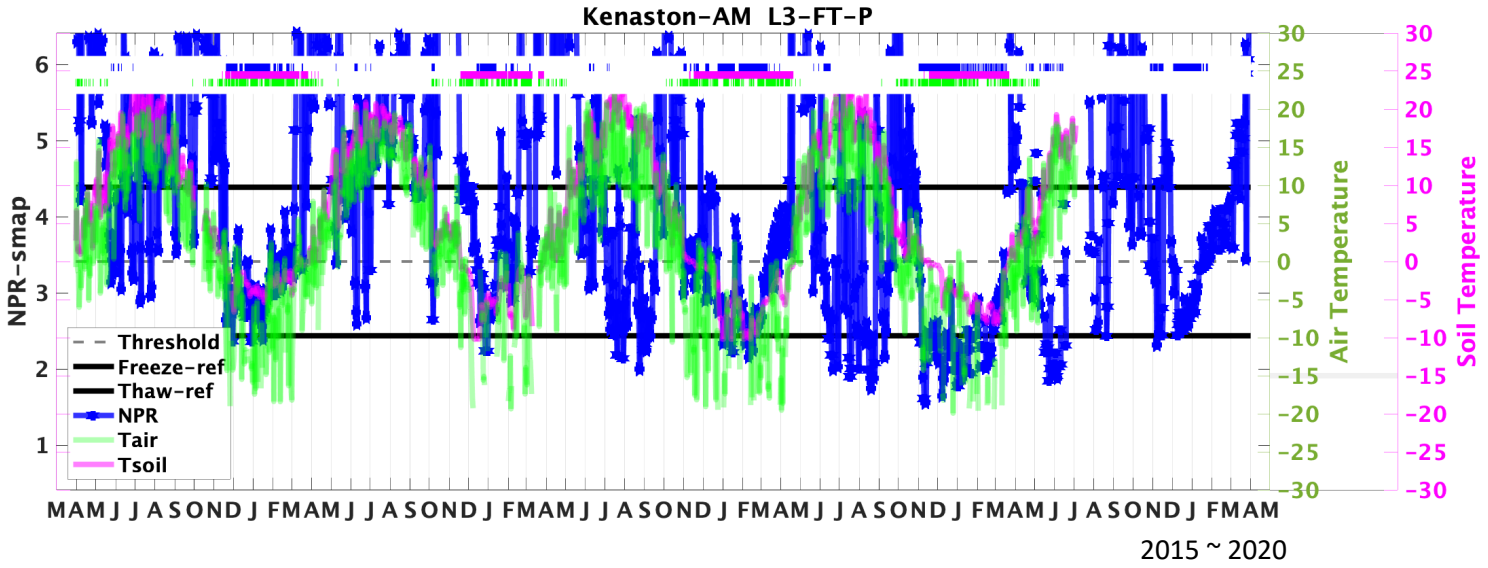



Figure 6.11. As in Figure 6.4 but for Saariselka.

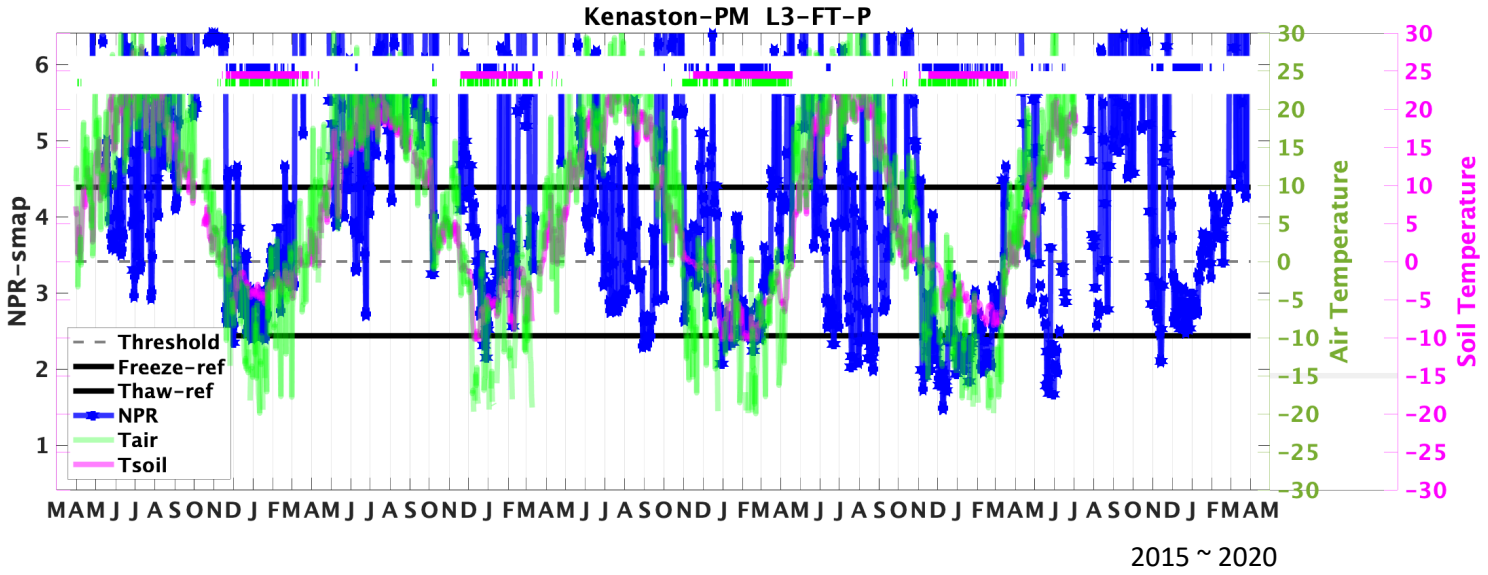
The Saariselka sites are located in an upland tundra environment surrounded by boreal forest, creating heterogeneous sub-grid land cover conditions. Despite this landscape complexity, the NPR time series exhibit a strong response to FT transitions (Figure 6.11a and b), and overall flag agreement is above 80% (Figure 6.11c). Lower agreement during Apr./May. is related to NPR fluctuations around the FT threshold during periods of wet snow and diurnal freeze and thaw transitions.

## IX. Kenaston

(a)




(b)



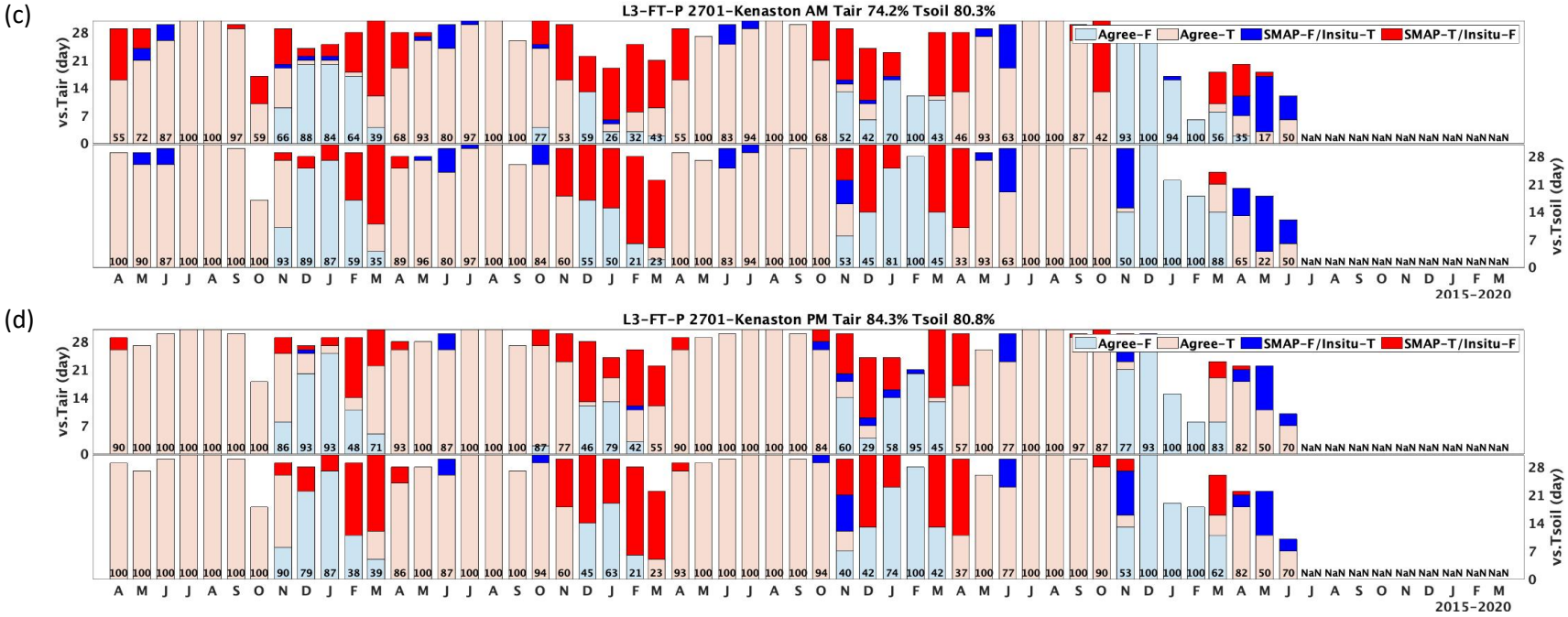



Figure 6.12. As in Figure 6.4 but for Kenaston.

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 (Figure 6.12a and b). There was an early NPR thaw transition compared to both Tair and Tsoil. Unlike all other core sites, during the summer months, there were large NPR oscillations above and below the FT threshold. The apparent false freeze flags are effectively removed by the never-frozen climatology mask. Therefore, the monthly and overall FT accuracy statistics vs. Tair and Tsoil over Kenaston are still above 80 for PM.

### *6.2.2 L3_FT_P vs. L3_FT_P_E*

The L3_FT_P (36 km) assessment (presented in Figures 6.4 through 6.12) was repeated for L3_FT_P_E (9 km). Overall, the assessment indicated very similar results, including the seasonal cycle of NPR with respect to air temperature, soil, and the validation statistics. Given the similarity in results, we do not present the full assessment in this report. Instead, Figure 6.13 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. The change in flag agreement is less than +/- 5%. The large difference occurs in Chersky (improved 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 6.14 (descending orbits only).

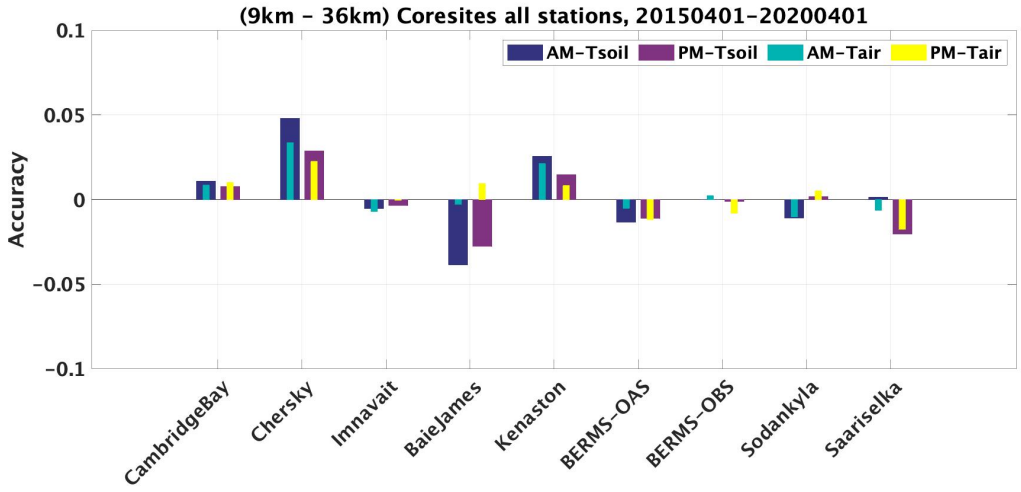



Figure 6.13. Difference in FT flag agreement, L3_FT_P_E minus L3_FT_P for descending and ascending overpasses. Positive values correspond to improved agreement for the enhanced resolution product; negative values correspond to degraded agreement.

### *6.2.3 Core Site Summary*

The number of core sites available for validation of SMAP FT products is limited, but they yield important insights into the behavior of NPR time series for tundra, boreal forest, 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 less effective insulator of the underlying soil relative to a deeper snowpack.

***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 and vegetation is significant, and can delay soil freeze for 1 to 5 months after air temperatures drop below zero Celsius (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 and heavily rely on the climatology masks.

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 occurs despite air temperatures being consistently below zero continuously for weeks with 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 Celsius 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 slightly 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 indicate frozen conditions), but agree at the time of afternoon overpass (when in situ measurements indicate thawed conditions). 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. 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 (Figure 6.14).

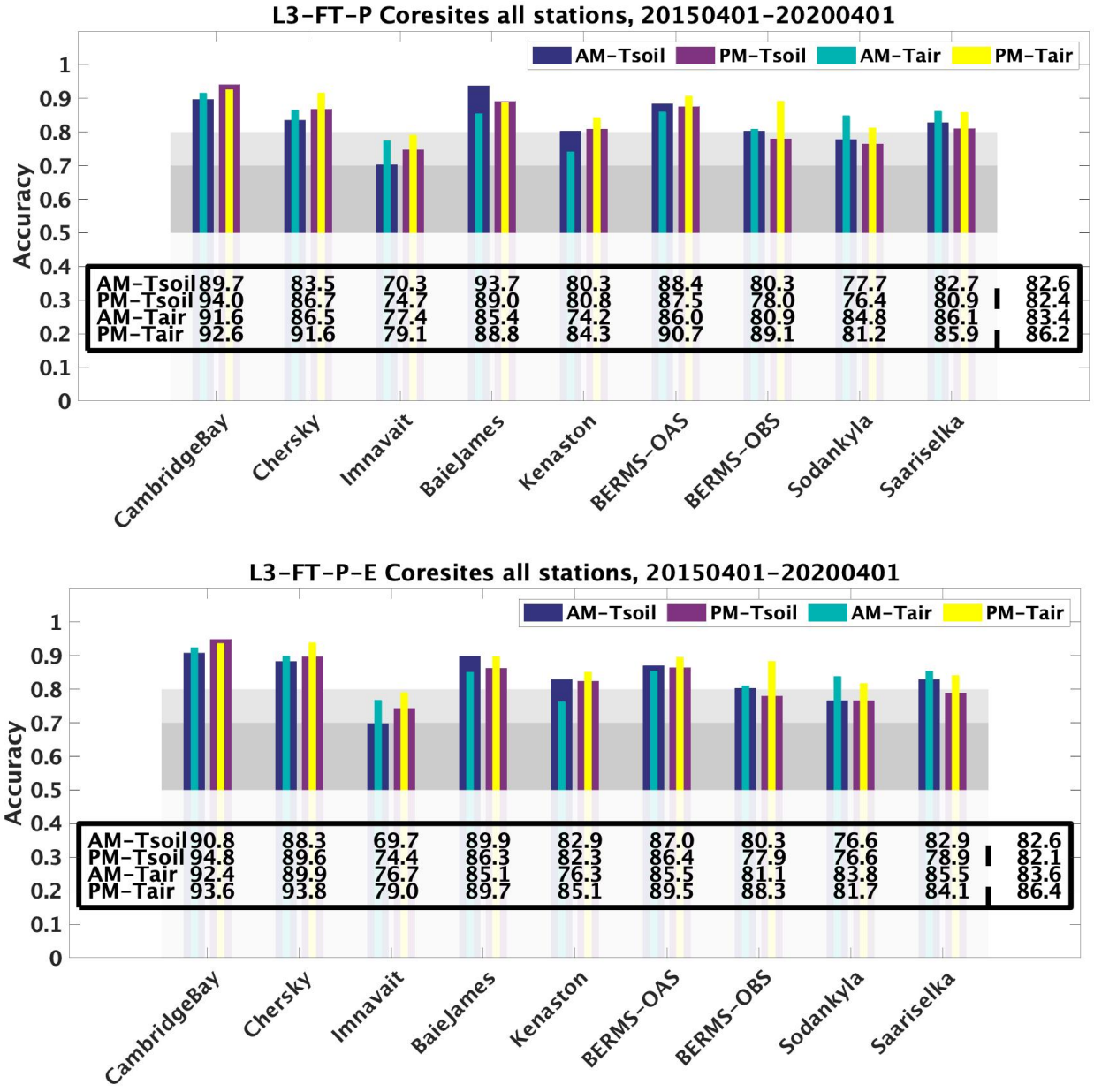



Figure 6.14. Three-year statistics for each core validation station compared with in-situ flags derived from air temperature and soil temperature with the overall statistic in the last column. Top: accuracy over standard 36km L3-FT-P; Bottom: accuracy over enhanced 9km L3-FT-P-E.

### *6.2.4 Sparse Networks*

The sparse networks consist of selected stations from the NRCS SNOTEL network. Compared with the core site stations, there is only one in-situ measurement within a satellite grid cell. To help ensure that the in-situ measurements from the single station represent the larger footprint indicated from the SMAP grid cell, the criteria for selecting the sparse network stations are: 1) relative uniform land cover; 2) relative smooth terrain within the satellite grid cell; 3) valid soil temperature at 5cm and air temperature measurement at 2m. In addition, the NPR reference difference in the grid cell needs to be larger than 0.5, so that the algorithm can be performed with adequate signal-to-noise. In Figure 6.15, we illustrate all of the selected NPR sparse network stations, which are limited to the Alaska region. 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 SMAP daily ascending and descending signals, we average the hourly measurements into AM (1:00~12:00) and PM (13:00 ~ 24:00) local time.

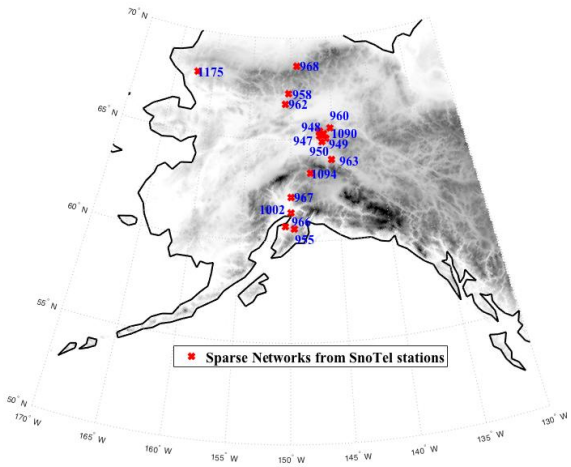



Figure 6.15. SNOTEL sites used for SMAP NPR Freeze/Thaw product evaluation.

Table 6.2 List of SNOTEL sparse network stations

| ID | LAT | LON | Name | State | Land Cover | Elevation(m) |
|---|---|---|---|---|---|---|
| **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 |

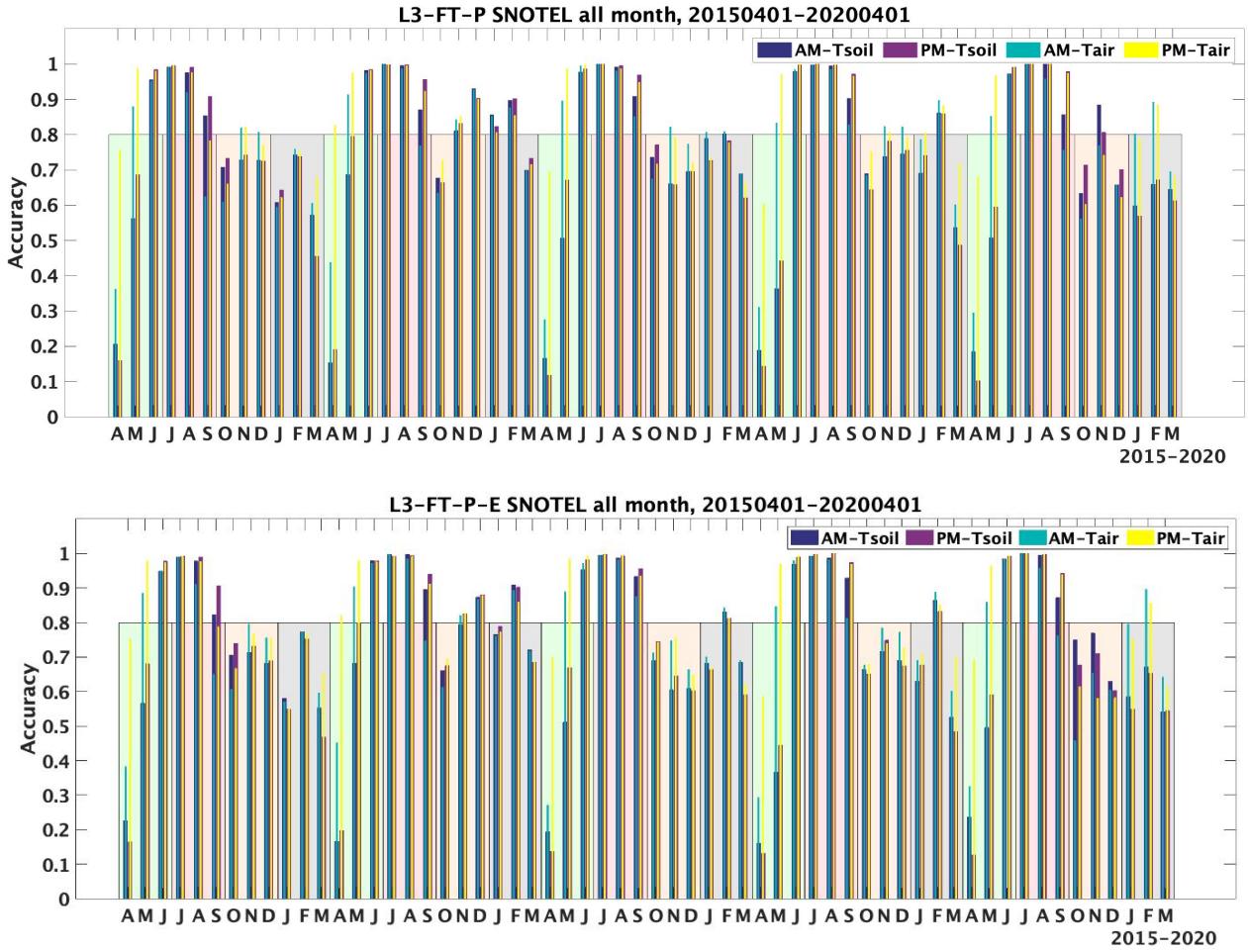



Figure 6.16 Monthly FT flag agreement for all validated SNOTEL sites from 2015-2020. Narrow columns indicate use of air temperature measurements for validation; wide columns indicate use of 5 cm soil temperature. Green/red/yellow/grey shading indicates 80% flag agreement in terms of the four season (spring, summer, fall and winter) requirement set by the SMAP mission.

In Figure 6.16, we compare the NPR retrieved FT flag monthly accuracy against the reference FT state determined from in situ soil and air temperature measurements respectively, for both ascending and descending passes over the SNOTEL sites. The seasonal pattern over three years is very similar. The months of June, July, and August have the highest accuracy of nearly 100%. Compared with the initial L3_FT product version, the new installed weekly climatology mask effectively removed most of the summer false freeze anomalies. However, during the winter (December to March), the accuracy drops to around 80% due to the high frequency variability of the NPR signal and ineffectively mask over the winter. In October, the accuracy drops to around 70% for both air temperature and soil temperature reference. This indicates that the SMAP derived FT flag has some delay or advance in freezing in the transition period, resulting in a larger error estimate relative to the ground observations, and which may require further threshold optimization. On the other hand, in the thawing season (April and May), the SMAP FT flag has much better correspondence with air temperature than soil temperature. The stronger retrieval agreement with the air temperature flag in spring indicates that the NPR value is responding to wet snow.

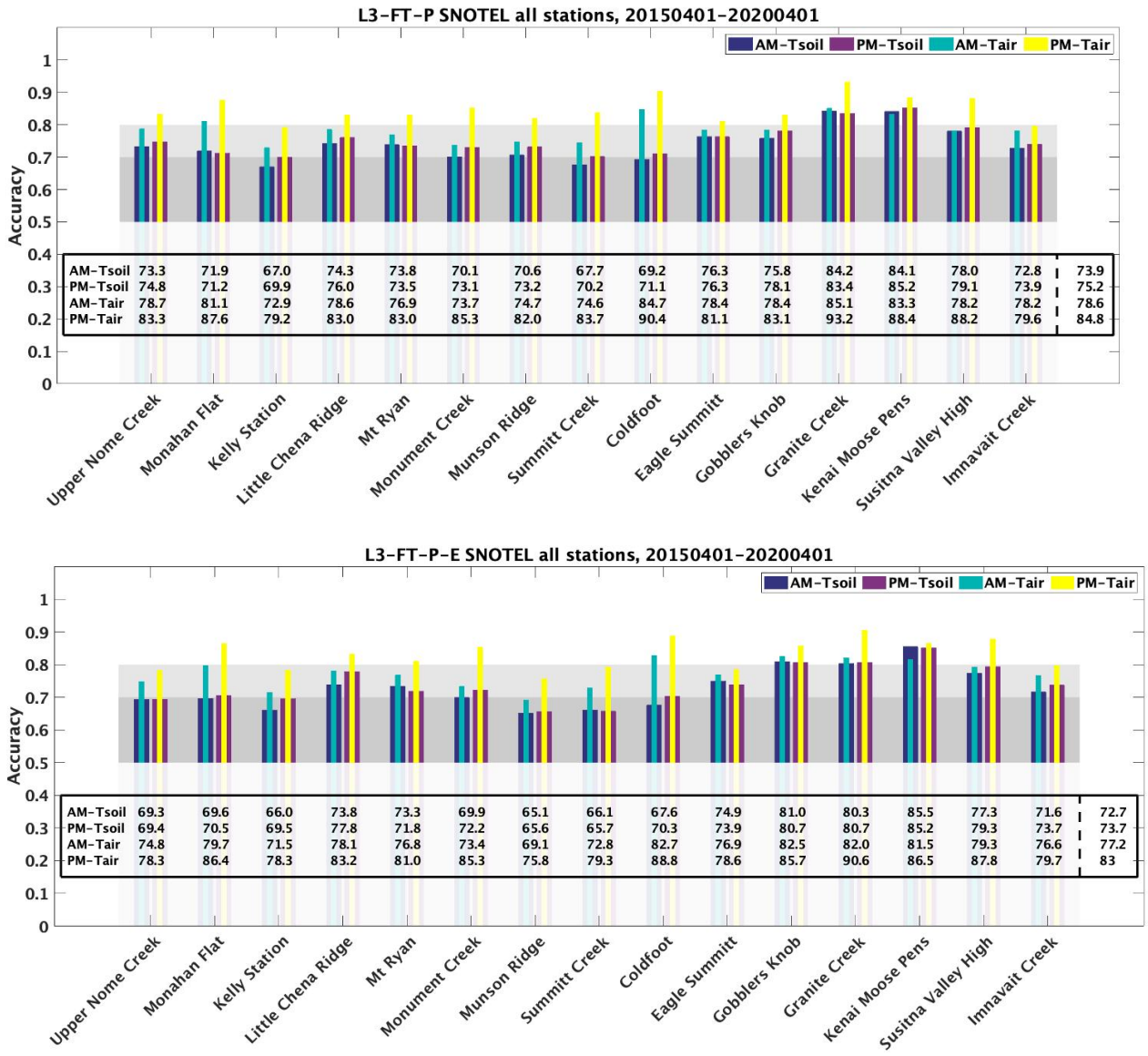



Figure 6.17 Summary of SMAP NPR FT three-year accuracy over each SNOTEL validation site for both ascending and descending pass results. The overall statistic is in the last column.

The three-year statistics summary of the evaluation of the SMAP passive derived FT product using observations from each SNOTEL station is shown in Figure 6.17. The statistics were computed separately for ascending/descending overpass FT retrievals. It is evident that the primary source of uncertainty in the SMAP derived FT retrievals occurs during the transition when the reference data indicate frozen ground while the SMAP FT retrievals show a thawed state. Again, this is consistent with the influence of wet snow on the NPR value. The larger uncertainty may also reflect greater spatial heterogeneity in microclimate, snow cover and FT conditions during the seasonal transitions. Sub-grid spatial heterogeneity within the relatively coarse (~40-km) SMAP footprint is also larger over complex terrain where SNOTEL sites are generally located.

We repeat the sparse network assessment for the SMAP enhanced FT product (L3_FT_P_E). In Figure 6.18, we plot the accuracy difference between the enhanced product and standard product over three years of data. Negative (positive) values indicate lower (higher) accuracy in the enhanced (9km) grid product relative to the standard (36km) grid product. The finer grid product has more fluctuation over time which introduces more false FT prediction in the quiet season (summer and winter) and degrades overall accuracy. Overall, the enhanced and standard grid FT products show similar performance; these results also indicate no clear accuracy improvement from the enhanced grid processing over the standard 36-km grid FT product.

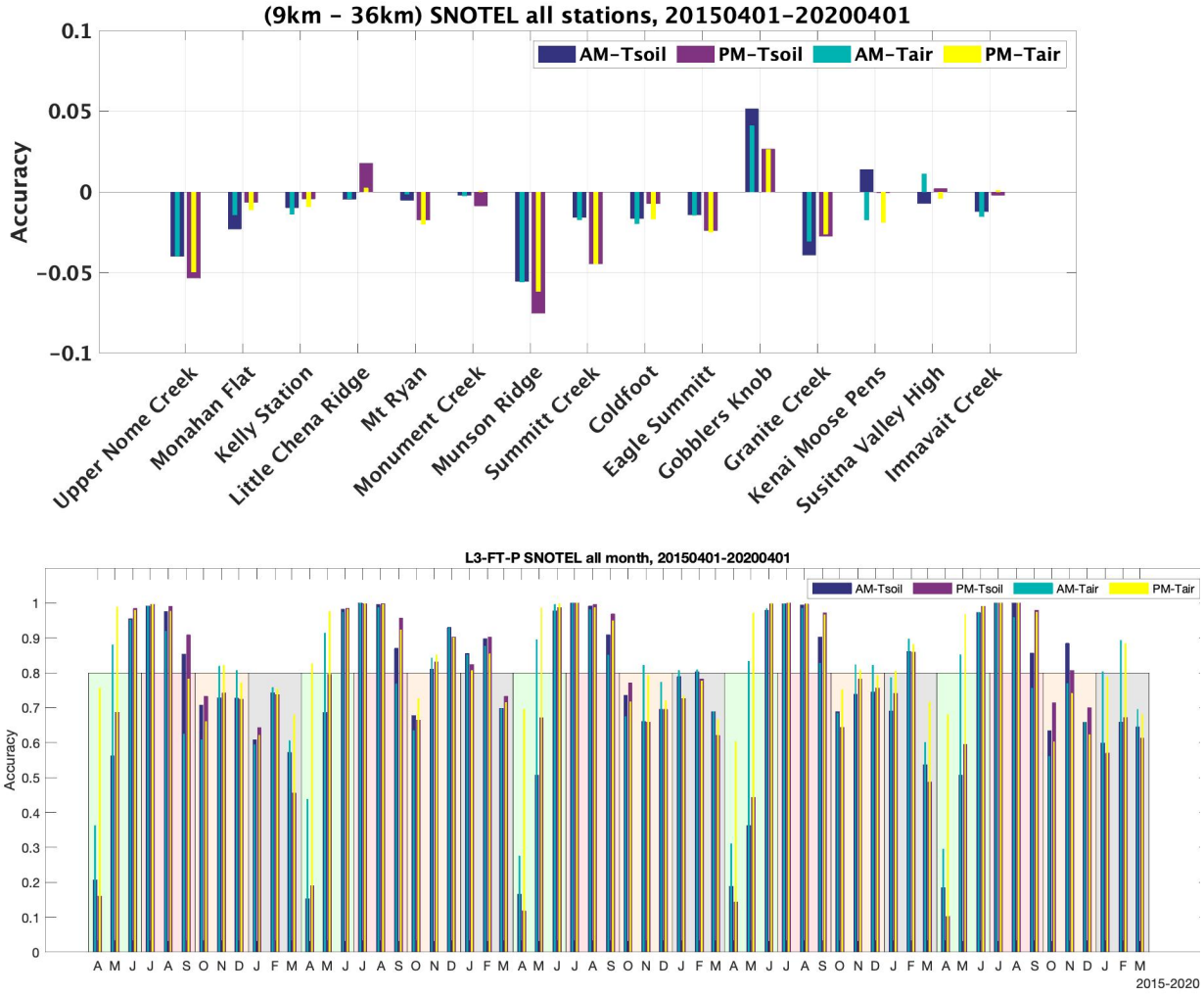



Figure 6.18 Monthly NPR FT flag agreement difference between Enhanced product (9km) and the standard product (36km) for all SNOTEL validation sites from 2015-2016. Narrow columns indicate use of air temperature measurements for validation; wide columns indicate use of 5 cm soil temperature; negative (positive) values indicate lower (higher) accuracy in the 9km product relative to the 36km product.

## 6.3 Global Grid Validation

The global grid FT retrievals were evaluated against independent FT estimates derived from in situ daily surface (screen-height) air temperature measurements from 5020 weather stations which report daily observations to the World Meteorological Organization (WMO) located across the global FT domain (Figure 6.19). 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°C) temperature threshold between frozen and non-frozen conditions, and compared with respective SMAP descending (AM) and ascending (PM) overpass FT retrievals following previously developed methods[7][9]. 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 selection of WMO validation stations was done separately for the L3_FT_P[E] 36-km and 9-km grid products. The surface air temperature daily minimum ($SAT_{mn}$) and maximum ($SAT_{mx}$) records for the selected stations were used to define daily frozen ($T \leq 0°C$) and non-frozen ($T > 0°C$) 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 $SAT_{mn}$ and $SAT_{mx}$ 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.

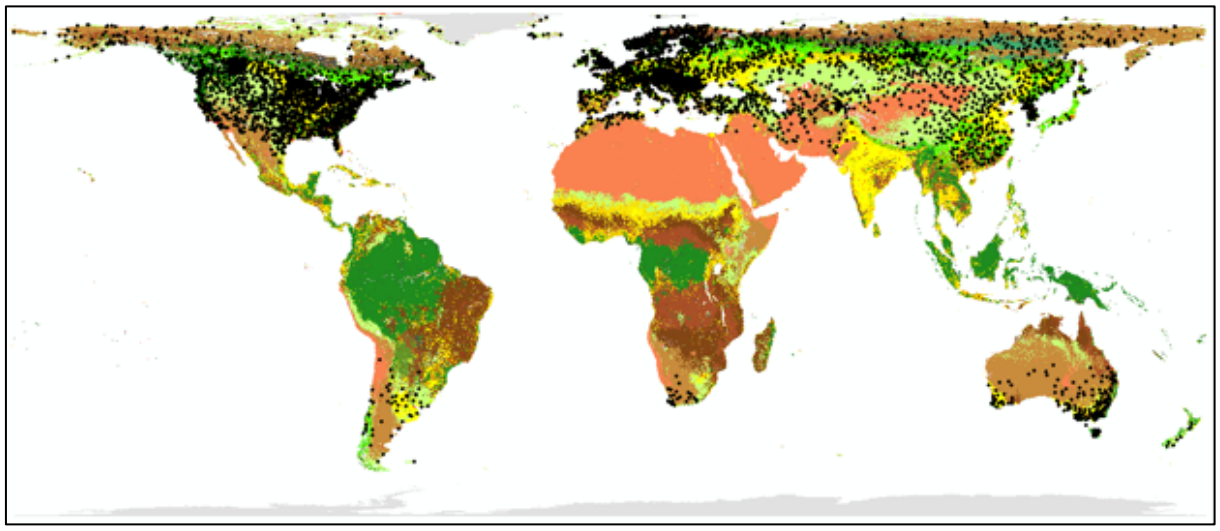


Figure 6.19. MODIS MCD12Q1 IGBP Land cover map showing WMO weather stations (black dots) used for SMAP FT validation assessment over the global domain.

The SMAP FT data records from the L3_FT_P[E] products were used for the WMO station based validation assessment. The daily (AM and PM) FT outputs are gridded to 36-km and 9-km resolution global 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 in the L3_FT_P[E] products.

A plot of the resulting daily mean AM and PM overpass FT spatial classification accuracies for the global domain from the L3_FT_P[E] records is shown in Figure 6.20. The SMAP FT agreement with the WMO SAT derived FT records is generally higher during the summer when the pattern and persistence of non-frozen conditions is more homogeneous and stable, more variable during the more heterogeneous spring and fall transition periods, and lowest during winter. For 36-km and 9-km products the PM overpass FT results generally exceed the 80% mean spatial classification accuracy threshold of the baseline mission

requirement. The FT classification accuracy from the AM overpass results is significantly lower than the PM results, but is generally above 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, and inconsistent observation time between the SMAP AM overpass and daily WMO minimum SAT measurement.

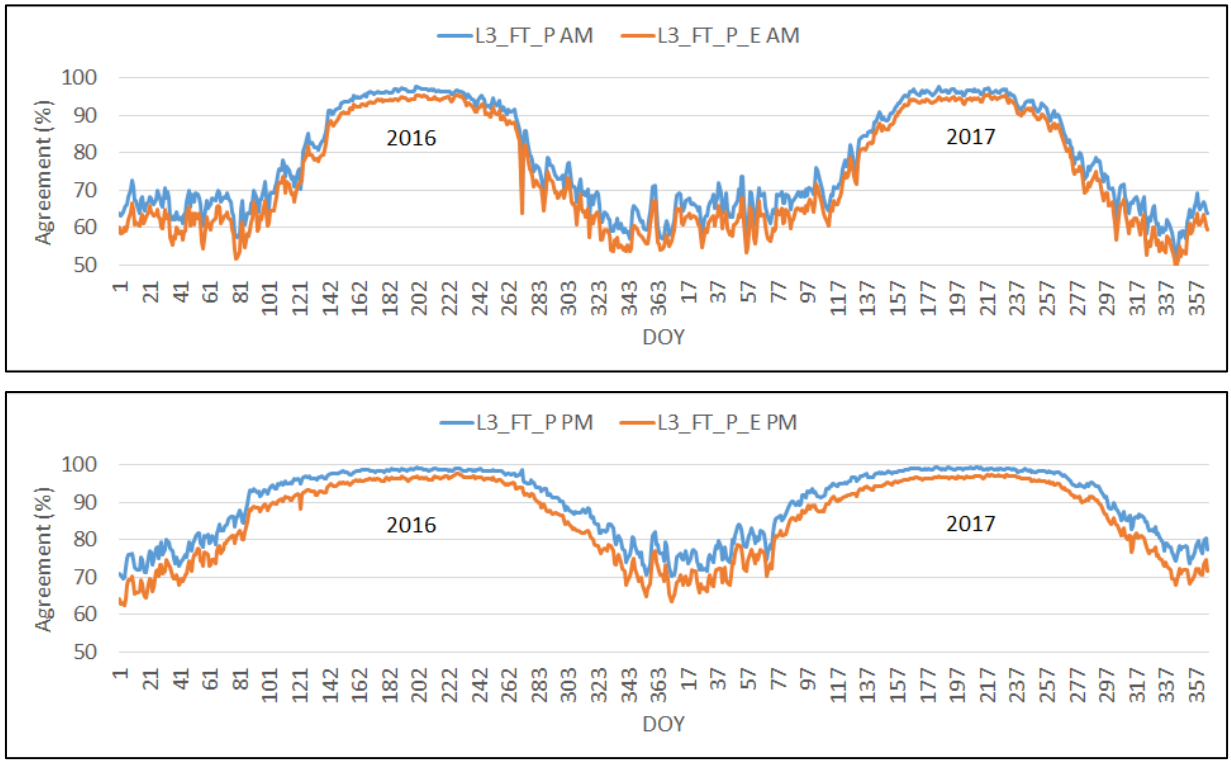



Figure 6.20. Seasonal pattern of SMAP daily mean FT classification agreement (%) in relation to WMO in-situ surface air temperature measurements for AM (top) and PM (bottom) overpass results; blue and red lines denote respective 36-km and 9-km grid results.

The validation results over the global domain show relatively small (3-4%) FT accuracy differences between the 9-km and 36-km gridded products, similar to the core and sparse site validation results. The accuracy of the 36-km product was slightly better than the 9-km product for the PM and 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 closer to the native footprint (~40-km) of the SMAP radiometer rather than the 9-km product grid.

The spatial distribution of mean annual FT classification agreement (%) for the 36-km and 9-km products, and AM and PM overpass results in relation to grid cell-to-point comparisons with the WMO stations is shown in Figure 6.21. The spatial pattern in FT accuracy is generally similar between the 9-km and 36-km results. The AM overpass FT accuracy is lower than PM accuracy, particularly near coastlines and in more transient FT areas. The spatial pattern in FT classification accuracy indicated from the WMO station comparisons also reflects lower $T_b$ signal-to-noise in higher vegetation cover areas (e.g. forests) and over complex topography.

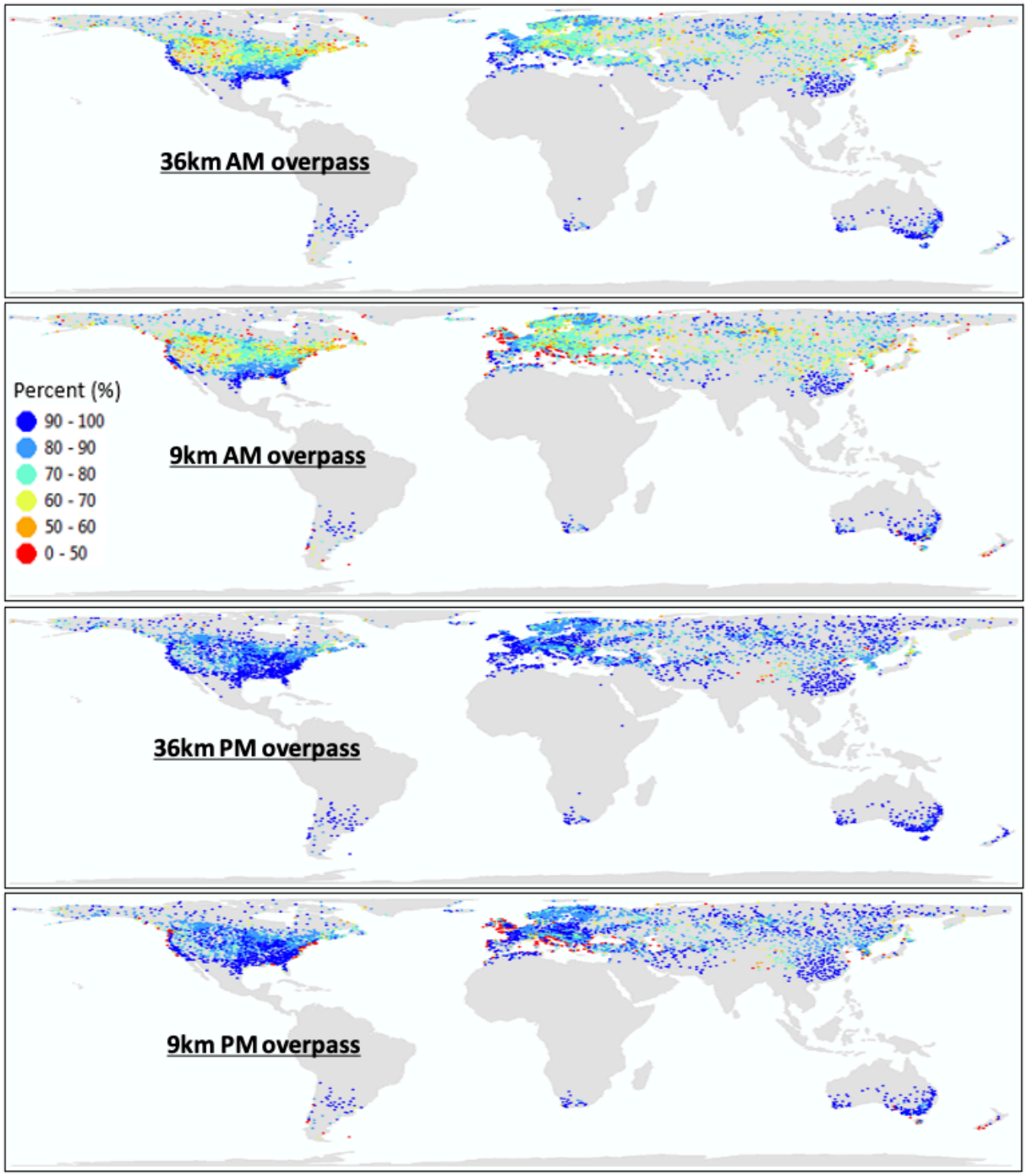



Figure 6.21. The spatial distribution of SMAP FT classification agreement (%) in relation to global WMO station sites from the 36-km and 9-km, and AM and PM overpass results for 2016.

The mean annual FT classification agreement was analyzed for 2016-2017 SMAP PM FT data records from both the 36-km and 9-km grid products. We found generally lower FT accuracy in areas with higher fractional water cover (Fw), where Fw represents the spatial proportion of open water within a grid cell derived from a 300m resolution ENVISAT land cover map (Figure 6.22). The FT classification agreement was also inversely proportional to the amount of forest cover and terrain complexity within a grid cell, as derived from respective MODIS (MOD44B) 250m vegetation continuous fields and 30m ASTER GDEM data. However, the relative impact of vegetation and terrain factors on FT classification accuracy was generally less than the influence of Fw cover. The 36-km grid product also showed less sensitivity to vegetation cover, terrain heterogeneity and Fw constraints than the 9-km grid product.

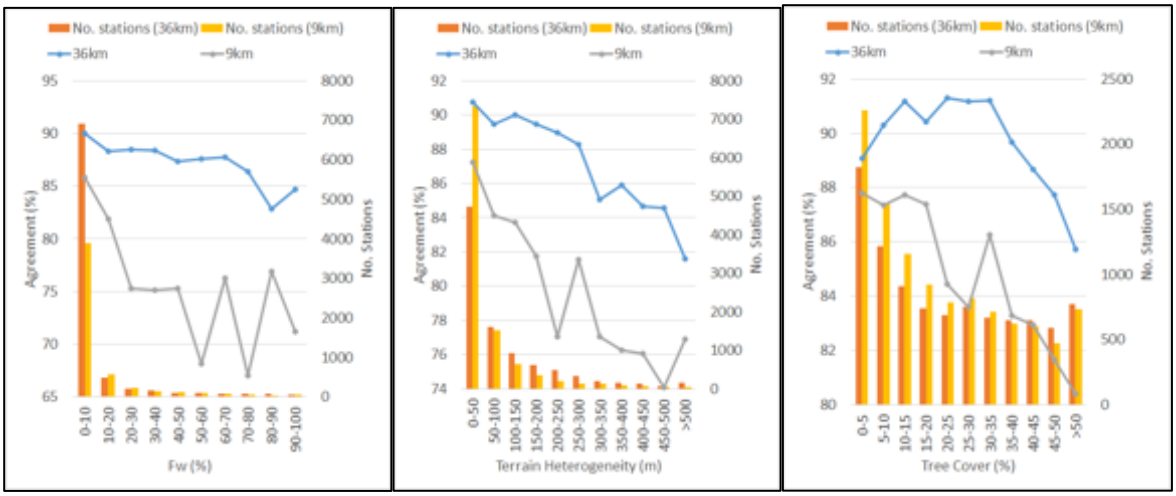



Figure 6.22. Dominant spatial heterogeneity indicators affecting 2016-2017 PM overpass FT classification agreement for individual grid cells, including open water fraction (Fw), terrain heterogeneity, and percent tree cover.

The overall FT classification accuracy from the L3_FT_P daily FT data records is presented in Table 6.3. These results summarize the SMAP FT validation assessments against WMO weather station observations for global domains during 2015-2020. The SMAP results are compared against WMO average daily temperature and AM vs. minimum temperature and PM vs. maximum temperature. Overall, the PM overpass FT classification results exceed the targeted 80% accuracy threshold for the baseline mission requirement.

Table 6.3. Mean annual FT classification accuracy (%) of the L3_FT_P (36km) & L3_FT_P[E] (9km) product daily FT frozen flags in relation to WMO air temperature based FT station observations for daily minimum temperature and maximum temperature and daily average temperature.

| 2016-2019 (full calendar year) | FT-AM vs. Tmin | FT-PM vs Tmax | FT-AM vs. Tavg | FT-PM vs. Tavg |
|---|---|---|---|---|
| L3-FT-P | 79.5% | 92.9% | 88.36% | 88.38% |
| L3-FT-P-E | 73.2% | 87.9% | 81.91% | 81.92% |

## 6.4 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 and sparse validation sites showed FT flag agreement relative to in situ air temperature and soil temperature measurements to be at or near the 80% mission accuracy requirement, with generally better agreement for ascending (PM) versus descending (AM) orbits, and air temperature versus soil temperature derived reference flags. 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 the 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.

Validation was also performed using in situ daily surface (screen-height) air temperature measurements from 5020 weather stations located across the global and northern (≥45°N) portions of the FT domain. These results were generally consistent with the core and sparse network validation results, and confirmed that the PM overpass FT classification results from the 9-km and 36-km products exceed the targeted 80% accuracy threshold for the baseline mission requirement. The AM FT accuracy is significantly lower than the PM FT results for both the 36-km and 9-km products, but is generally above the 70% accuracy threshold of the minimum mission.

The FT retrieval patterns showed no obvious artifacts between the NPR and SCA derived domains. The apparent FT accuracy and performance was generally consistent between the NPR and SCA derived results despite the different areas and conditions represented by the two algorithms. Overall, FT classification accuracy was generally lower during active FT transition periods when the spatial and temporal heterogeneity in landscape freezing and thawing is larger; whereas, FT accuracy was higher during more stable frozen and non-frozen seasons (e.g. winter and summer). The spatial pattern in FT accuracy is also influenced by landscape heterogeneity within the sensor footprint, where FT accuracy was degraded in areas with higher fraction of surface water cover, greater terrain variability and vegetation cover. Thus, variations in the global pattern of FT accuracy reflect landscape heterogeneity, climate variability and the duration and stability of seasonal frozen and non-frozen conditions.

The combined validation results show relatively small (3-4%) FT accuracy differences between the 9-km and 36-km gridded products. However, the overall FT accuracy of the 36-km product was slightly better than the 9-km product for both the PM and AM results. The 9-km grid product showed stronger negative impacts from landscape heterogeneity than the 36-km product. Potential noise introduced from the Backus Gilbert $T_b$ downscaling method used for the L3_FT_P_E retrievals may also contribute to lower than expected performance in this product relative to the L3_FT_P standard 36-km product.

# 7 OUTLOOK

This report describes the validated release of the SMAP radiometer derived landscape freeze/thaw product (Version 3) at standard (L3_FT_P) and enhanced (L3_FT_P_E) spatial resolutions. Priorities for further enhancement to these products will focus on the following areas:

- *Optimization of algorithm parameters*. The plan for L3_FT_P[E] 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[E] product development and validation.
- *False flag mitigation*. False freeze mitigation efforts were developed and tested during the L3_FT_P[E] algorithm development period. By implementing monthly 'never frozen' mask, the false freeze flags in the summer were greatly reduced. However false flags in the winter and transition period are still present and need further investigation, benefitting from an extended mission data record.
- *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. Other ongoing or planned field campaigns directly relevant to SMAP FT product refinements and applications include the NASA Arctic Boreal Vulnerability Experiment (ABoVE) and SnowEx campaigns. Analyses of various data from these activities will primarily support L3_FT_P[E] refinement and SMAP FT science applications pertaining to boreal-Arctic hydrology, snow and permafrost related processes, and terrestrial carbon cycle.
- *Refinement of global FT flags*. In the current release, the freeze/thaw products include a stand-alone global grid product with the FT classification domain defined from climatological temperatures from global reanalysis data. The inter-comparison between products is on-going, and includes assessments of the influence of spatial distortion effects between global and polar grid FT retrievals relative to in situ ground truth observations.
- *Science development*. To date, emphasis has been placed on algorithm and product refinements to verify and enhance product accuracy, coverage and performance. Moving forward, there is also a need to further evaluate the product performance and utility in addressing SMAP science objectives and related FT applications. Targeted application areas include: quantifying FT impacts on land surface phenology, hydrology, permafrost, and boreal carbon sink strength; clarifying the value of the L-band FT retrieval in characterizing soil processes; clarifying FT impacts on ET, snow cover properties and cold season flood risk. The above areas have direct traceability to SMAP science and application objectives, including improving understanding of water-carbon-energy cycle linkages, the boreal carbon sink, and enhancing regional flood risk assessments.

# 8 ACKNOWLEDGEMENTS

This document resulted from analyses and discussion among the Freeze/Thaw 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] 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.

[2] Science Data Calibration and Validation Plan Release A, March 14, 2014 JPL D-52544

[3] 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 C, August 31, 2020. Available at http://smap.jpl.nasa.gov/science/dataproducts/ATBD/.

[4] 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.

[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] McDonald, K.C, and J.S. Kimball. (2005) Hydrological application of remote sensing: Freeze-thaw states using both active and passive microwave sensors. In *Encyclopedia of Hydrological Sciences*. Vol. 5., M.G. Anderson and J.J. McDonnell (Eds.), John Wiley & Sons Ltd.

[7] Kim, Y., J.S. Kimball, K.C. McDonald and J. Glassy. (2011) Developing a global data record of daily landscape freeze/thaw status using satellite passive microwave remote sensing. *IEEE Transactions on Geoscience and Remote Sensing,* **49**, 949-960.

[8] Nemani, R.R., C. Keeling, H. Hashimoto, W. Jolly, S. Piper, C. Tucker, R. Myneni, and S. Running. (2003) Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science,* **300**, 1560-1563.

[9] Kim Y, Kimball JS, Glassy J, Du J. An extended global Earth system data record on daily landscape freeze–thaw status determined from satellite passive microwave remote sensing. Earth System Science Data. 2017 Feb 16;9(1):133-47.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] Podest, E., K.C. McDonald, and J.S. Kimball. (2014) Multi-sensor microwave sensitivity to freeze-thaw dynamics across a complex boreal landscape. *Transactions in Geoscience and Remote Sensing,* **52**, 6818-6828.

[16] 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.