

Investigating the Relationship Between Satellite-Based Freeze/Thaw Products and Land Surface Temperature

Jeremy Johnston¹, Student Member, IEEE, Viviana Maggioni¹, and Paul Houser¹, Member, IEEE

Abstract—This paper investigates surface temperature variables as they relate to passive microwave-derived surface freeze/thaw (FT) state and assesses the accuracy of such FT products relative to surface temperature. Utilizing retrievals from the soil moisture active/passive (SMAP), advanced microwave scanning radiometer, and special sensor microwave imager instruments, surface FT records have previously been derived globally. Moderate Resolution Imaging Spectroradiometer skin temperature, North American Land Data Assimilation System (NLDAS) skin, 0–10-cm soil layer, and 2-m air temperatures are compared to the various FT state products (FTSPs) by defining the threshold for FT state transitions at 0 °C. This paper utilizes the 2015–2016 overlap period in FT records within the NLDAS domain. Spatial variability of classification accuracy (CA) is then investigated over the study area. A proportional differencing method also enables the identification of biases between FTSPs and surface temperature variables. Additionally, by analyzing probability distribution functions of FTSPs associated with temperature values, we assess the distribution of temperature variables as they relate to FT classifications. Classification agreement is shown to vary with sensor configuration, seasonality, and retrieval time. Air temperature is found to have the highest CAs across FTSPs (81%–91%), while NLDAS soil exhibits a close relationship to FTSPs, especially in regard to SMAP products. Finally, ascending (P.M.) retrievals are shown to be increasingly linked to the selected temperature parameters as compared to descending (A.M.) observations. This paper contributes to an improved understanding of current FTSPs and will benefit efforts to enhance future FT products.

Index Terms—Earth, geoscience North America, land surface temperature, microwave radiometry, passive microwave (P-MW) remote sensing, remote sensing (RS), satellite applications.

I. INTRODUCTION

AS WATER in various components of the landscape freezes, its movement is largely curtailed, as are its contributions

Manuscript received January 22, 2019; revised June 4, 2019; accepted June 27, 2019. Date of publication August 7, 2019; date of current version September 29, 2019. This work was supported by the National Aeronautics and Space Administration under award number NNX17AL38G. (Corresponding author: Jeremy Johnston.)

J. Johnston and V. Maggioni are with the Department of Civil, Environmental, and Infrastructure Engineering, George Mason University, Fairfax, VA 22030 USA (e-mail: jjohns60@gmu.edu; vmaggioni@gmu.edu).

P. Houser is with the Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA 22030 USA (e-mail: phouser@gmu.edu).

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2019.2926942

to hydrology, vegetation productivity, and biogeochemical processes. Extensive research has been conducted to identify the impact of frozen terrain on carbon cycling and vegetation growth, especially as it pertains to northern regions with well-defined frozen and thawed seasonality [1]–[3]. The role of frozen ground in inhibiting the movement of water both in terms of evaporation and percolation deep into the soil has also been shown to impact surface energy fluxes [4]. Freeze/thaw (FT) state is also linked to socio-economic impacts through its connections to flooding and water availability [5], [6]. Because of the fundamental role and impact of FT on climate, hydrology, and ecology, several satellite remote sensing (RS) techniques have been developed to observe and quantify these processes at global scales. These methods primarily use passive microwave (P-MW) radiometry in the L-, K-, and Ka-bands to infer surface FT state through their sensitivity to changes in the soil dielectric constant. Increases in observed brightness temperatures (T_b) have shown to be coincident with surface freezing, as the loss factor related to microwave emissions is reduced relative to that of liquid water [7], [8]. These changes in sensor measured emissions, brought on by the phase change of water, have been proven useful in classifying a binary FT state.

As satellite retrievals are acquired at various P-MW frequencies, times, and spatial resolutions, significant variations in FT classifications can occur. These P-MW observations are sensitive to the holistic landscape FT state, which can include a heterogeneous combination of different vegetation covers, soils, snow, and terrain. This landscape complexity is not well understood, as defined FT state can represent canopy, snow surface, or soil FT state depending on observation band, location, and time [9]. Similarly, P-MW-based FT state has been shown to vary from air temperature measurements exhibiting sensitivity to land surface FT properties. This occurs especially in times of ephemeral FT events in which snow cover and soil thermal inertia can contribute to prolonged frozen periods even as air temperatures have climbed above freezing [10], [11]. Lower frequency L-band (1.4 GHz) T_b measurements, as are utilized in the soil moisture active passive (SMAP) [12] FT products, have been shown to have an increased emission depth sensitivity as compared to Ka-band retrievals (36.5, 37 GHz) [13]. These Ka-band retrievals as are observed by the Advanced Microwave Scanning Radiometer (AMSR) [14], [15] and the Special Sensor Microwave/Imager (SSMI) [16], [17] radiometers are characterized by wavelengths (λ) of around 1 cm. As a result, Ka-band

TABLE I
STUDY SELECTED F/T PRODUCTS

Freeze/Thaw (FT) Product (Sensors)	Product Name	MW band	Spatial Resolution	Period of Record	Domain	Overpass Times (LST) Descending (AM) Ascending (PM)
SMAPe - FT (SMAP)	SMAP_L3_FT_P_E	L-Band (1.41 GHz)	9 km	April 2015 - Present	> 45° N latitude	6:00 18:00
SMAP - FT (SMAP)	SMAP_L3_FT_P	L-Band (1.41 GHz)	36 km	April 2015 - Present	> 45° N latitude	6:00 18:00
AMSR - FT (AMSR-E, AMSR2)	AMSR_36V_AM/PM_FT	Ka-Band (36.5 GHz)	25 km	2002 - 2016*	Global	1:30 13:30
SSMI - FT (SSMR, SSM/I, SSMIS)	SSMI_37V_AM/PM_FT	Ka-Band (37 GHz)	25 km	1987 - 2016*	Global	6:30 18:30

*At time of study completion.

retrievals exhibit increased sensitivity to surface features such as vegetation compared to SMAP *L*-band (λ , 15–30 cm) retrievals [13], [18], [19].

Several prior investigations into the use of *L*-band P-MW to estimate surface FT state have been undertaken. Ground-based radiometers (ELBARA-II), airborne, and soil moisture ocean salinity (SMOS) *L*-band retrievals, all demonstrated the ability to detect soil frost through changes in the measured P-MW signatures [18]. Change detection algorithms are able to differentiate between frozen and thawed states by utilizing the distinct *L*-band microwave signatures of frozen and thawed soil states across polarizations [13]. Distinct differences in ground-based radiometer Tb retrievals across varied land covers and soil types were also observed, however, these algorithms were able to detect soil freezing at 2–10-cm depth with >93% accuracy [13].

Recent studies have bolstered these efforts by applying change detection techniques to estimate surface FT state using SMOS satellite observations on a hemispheric scale [20]. SMOS along with field validation experiments, such as the Scanning *L*-band Active Passive experiment (SLAPex; [21]), have continued to demonstrate the effectiveness of *L*-band P-MW at estimating surface FT state. Clear decreases in the difference between vertical and horizontal polarizations have been repeatedly observed as a result of soil freezing [13], [21]. Performed over crop land in Manitoba during diurnal FT cycles, SLAPex utilized *in situ*, airborne, and SMAP observations to demonstrate the prevalence of SMAP frozen classifications when >60% of *in situ* measurements indicated surface temperatures below 0 °C [21]. Similar studies utilizing ground-based radiometers in the *Ka*-band have also proved sensitive to soil FT state in the 0–5-cm soil layer [19].

These differences in P-MW frequencies have resulted in two unique types of FT state products (FTSPs). The *L*-band is represented by the SMAP passive FT products (Version 1) [22], produced at both 36 km × 36 km (SMAP-FT) [23] and at an enhanced 9 km × 9 km (SMAP_E-FT) [24] resolution (see Table I). The *Ka*-band-based FTSPs, produced as part of the Freeze/Thaw Earth System Data Record (FT-ESDR) [25], are derived from AMSR and SSMI instruments on a 25 km × 25 km grid [26]. Both SMAP and FT-ESDR derived products utilize a seasonal thresholding approach, in which retrievals are compared to Tb-derived FT reference states in order to classify areas as frozen or thawed. FT product validation remains an ongoing process; however, *in situ* air and soil temperature measurements have been the predominant evaluation dataset in meeting

classification accuracy (CA) goals of both SMAP and ESDR FT records [23], [24], [25]. In validating SSMI and AMSR-based FTSPs, the use of *in situ* validation sites has in many cases been limited to areas of primarily homogenous terrain and vegetation. This is done in order to best represent landscape state through point measurements, which tend to be unrepresentative of large areas. Moreover, the sparse availability of *in situ* observations, especially in high-latitude regions, has made validation through these methods incomprehensive. The use of optical and thermal infrared RS techniques has been proposed to bolster validation efforts, but inconsistent coverage due to clouds, effects of snow cover, limited emission depth, and low solar illumination during FT transitions limit their use as effective methods [25].

Previous studies have found high correlation of FT-ESDR products to air and soil temperature, as thresholding algorithms in many cases have been optimized based on these observations [8], [25]. SMAP-FT validation efforts have largely followed a similar path by utilizing *in situ* air and soil temperatures to evaluate and improve FT classification algorithms. This is done through the use of flags marking disagreement between FT state and *in situ* observations [22], [27]. Evaluation and improvement of SMAP FTSPs is ongoing, which considers the effects of terrain, surface water, snow cover, and vegetation through sensitivity analysis and masking procedures. Validation efforts, in a limited capacity, have been performed using modeled skin and soil temperatures as well as multiscale measurements intended to reduce representation errors incurred through the use of point observations [9], [22]. Previous studies have also applied SMAP-based FT algorithms to latitudes south of 45 °N and validated these classifications utilizing air, 5-cm soil temperatures, and modeling approaches [28]. Historically, FT validation efforts have relied on imperfect measures, due to the difficulty in representing aggregate landscape FT state as is measured by P-MW RS techniques. SLAPex demonstrated the effects of spatial heterogeneity on how it relates to large-scale FT classifications [21], however, the majority of validation efforts continue to rely on point observations. As ground-based radiometer studies have shown varied retrievals depending on surface vegetation, soil, and moisture characteristics [13], [19], satellite-based FT observations tend to only capture the aggregate emissions of these components making it difficult to accurately characterize an entire area at the kilometer scale.

The primary goals of this paper are as follows. First, to determine which land surface temperatures act as the best surrogates

TABLE II
STUDY SELECTED TEMPERATURE PRODUCTS

Temperature Variable	Product Name	Spatial Resolution	Temporal Resolution	Period of Record	Domain
NLDAS 2m Air Temperature	NLDAS_FORA0125 V2	0.125° (~11 km)	Hourly	1979 - Present	25 - 53° N, 67 - 125° W
NLDAS 0-10cm Averaged Soil Temperature	NLDAS_NOAH0125 V2	0.125° (~11 km)	Hourly	1979 - Present	25 - 53° N, 67 - 125° W
NLDAS Surface Skin Temperature	NLDAS_NOAH0125 V2	0.125° (~11 km)	Hourly	1979 - Present	25 - 53° N, 67 - 125° W
MODIS-Aqua Land Surface Temperature	MYD11C1 V006	0.05° (~4 km)	1:30 AM/PM LST	July 2002 - Present	Global

to satellite FT product defined surface states and second, to investigate the differences and uncertainties among various FTSPs. Furthermore, we assess the accuracy and limitations of four FTSPs over a study area encompassing varied terrain, climate, and land cover. Through a spatial analysis, we examine the factors contributing to FT classification differences between FTSPs and temperature-threshold-based FT classifications. This paper aims to improve our understanding of what current satellite FT products represent and ways in which they can be enhanced through a regional assessment of landscape FT complexity. Comparisons across existing global P-MW FT products have not been previously performed at this scale utilizing both modeled and satellite observed temperature variables. This assessment of spatial variability and uncertainty between FT products and surface temperature variables will contribute to the development of improved FTSPs moving forward. We first present the methodologies used to compare FTSPs to each other and to various temperature products (see Section II), followed by descriptions and discussion of the resulting comparison metrics (see Section III). Study limitations, relevance, future research directions, and conclusions are presented in Sections IV and V.

II. METHODOLOGY

This paper compares the following four temperature datasets to SMAP, AMSR, and SSMI FT products (see Tables I and II):

- 1) National Land Data Assimilation System Version 2 (NLDAS) Noah land surface modeled skin temperature ($T_{N\text{-skin}}$) [29];
- 2) NLDAS Noah 0–10-cm average soil temperature ($T_{soil0\text{-}10cm}$);
- 3) NLDAS forcing 2-m air temperature ($T_{2m\text{-air}}$) [30];
- 4) Moderate Resolution Imaging Spectrometer (MODIS) [31] infrared-based skin temperature ($T_{M\text{-skin}}$) on board the Aqua satellite platform.

These comparisons are completely based upon product overlapping temporal (April 2015–2016) and spatial domains (45–53° N, 67–125° W), which includes the northern conterminous United States and southern Canada (see Table II). This time series was sufficient to capture multiple seasonal and ephemeral

FT transitions occurring within a 21-month period of FT product temporal overlap. The Phase 2 of NLDAS project (NLDAS-2) dataset has been improved over the previous version of NLDAS ([32]; [33]; <http://ldas.gsfc.nasa.gov>) based on output from the atmospheric data assimilation system at the NASA Global Modeling and Assimilation Office [34], [35], [36]. NLDAS forcing variables such as 2-m air temperature are derived from North American regional reanalysis (NARR) datasets, which are based on *in situ* station, radiosonde, aircraft, and satellite observations [37]. Significant positive biases in shortwave radiation have been observed in NARR datasets, however, this has been improved upon in NLDAS-2 forcing by performing bias correction with Geostationary Operational Environmental Satellite (GOES) observations [35]. Several studies have been performed validating NLDAS $T_{2m\text{-air}}$ against observations from the Oklahoma Mesonet, the atmospheric radiation measurement (ARM) program, and cloud and radiation test bed (CART) sites. High correlation of NLDAS to air temperature observations was shown across Mesonet, ARM, and CART sites with a small negative bias [32], [38]. Modeled surface temperature variables highly dependent on model forcing, including $T_{N\text{-skin}}$ and $T_{soil0\text{-}10cm}$, have also performed well at reproducing *in situ* observations [35]. NLDAS skin temperature has been shown to tend toward a warm bias during the mid-day period as well as a 3–5 K cold bias during the night time hours [36]. These skin temperature biases are also expected to contribute to error in soil temperatures, still top layer soil temperatures have been shown to compare best with observations relative to lower layers. This is reflected by a 2.5 K annual negative bias with an absolute error of 0.3 K when compared to soil temperature observations at 10 cm, best among modeled soil layers [36]. This cold bias in $T_{soil0\text{-}10cm}$ is most pronounced during the summer and fall, while $T_{soil0\text{-}10cm}$ tends to be lower than observed most notably during the night time hours of the winter months. Multiple assimilation inputs along with imperfect model forcing have been shown to introduce small errors into NLDAS temperature data [36]. This must be considered when assessing study results, especially as it pertains to uncertainty around the freezing point.

A total of 16 comparisons across all FT and temperature products are performed. NLDAS Coordinated Universal Time (UTC) hourly temperature values are divided into multiple regions

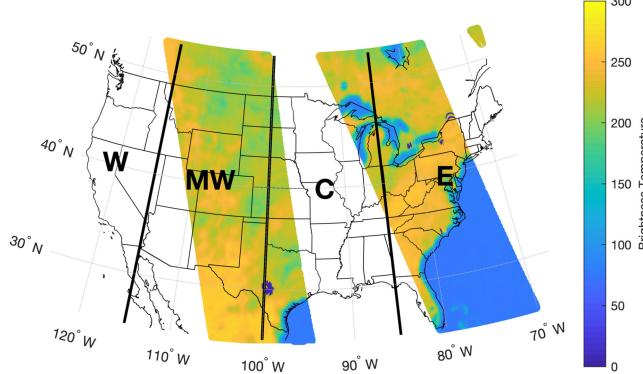


Fig. 1. Splice boundary zones for NLDAS data time matching overlaid with example SMAP Tb swaths (11/06/2015).

over the domain (see Fig. 1). This is done in order to provide temperature data more representative of satellite overpass times, as opposed to utilizing a single UTC hour over the entire domain. For example, the latter approach could identify temperature values from the eastern zone at 8:00 A.M. local solar time (LST), while temperature values in the western zone would be representative of 5:00 A.M. LST, inducing unnecessary errors due to parts of the domain being warmed by sunlight. Sensor overpass retrievals are acquired at specific LST as opposed to UTC timing making the use of multiple regions necessary. To account for this, temperature comparisons were completed by combining all zones into a single temperature snapshot representative of four different hours of NLDAS instantaneous temperature states. Daily A.M./P.M. MODIS Aqua retrievals were used as provided by the data producers.

Both standard and enhanced resolutions SMAP FTSPs were utilized to investigate the degree to which temperature parameter association to FT classification is dependent on resolution. This is particularly useful as temperature data are available at finer spatial resolutions, 0.05° (~ 5 km) and 0.125° (~ 12 km), as compared to the moderate resolution of FT-ESDR and standard resolution SMAP FT products of 25 and 36 km, respectively. Additionally, both *Ka*-band-based FTSPs included in the FT-ESDR were used to evaluate temporal differences in retrieval times (see Table I) when using the similar sensing frequencies of AMSR (36.5 GHz) and SSMI (37 GHz). Comparisons between SMAP-based *L*-band and FT-ESDR-based *Ka*-band FTSPs was an additional objective of this paper.

A. CA Thresholding

A simple method used to compare binary FT state and temperature is by thresholding the temperature values at 0°C . The continuous temperature datasets are converted into a binary one with values either greater than 0°C (thawed state) or lower than/equal to 0°C (frozen state). Realistically, the ground surface may remain thawed at temperatures below 0°C or frozen at temperatures above 0°C due to freezing point depression and thermal inertia [40], [41]. However, this threshold provides a

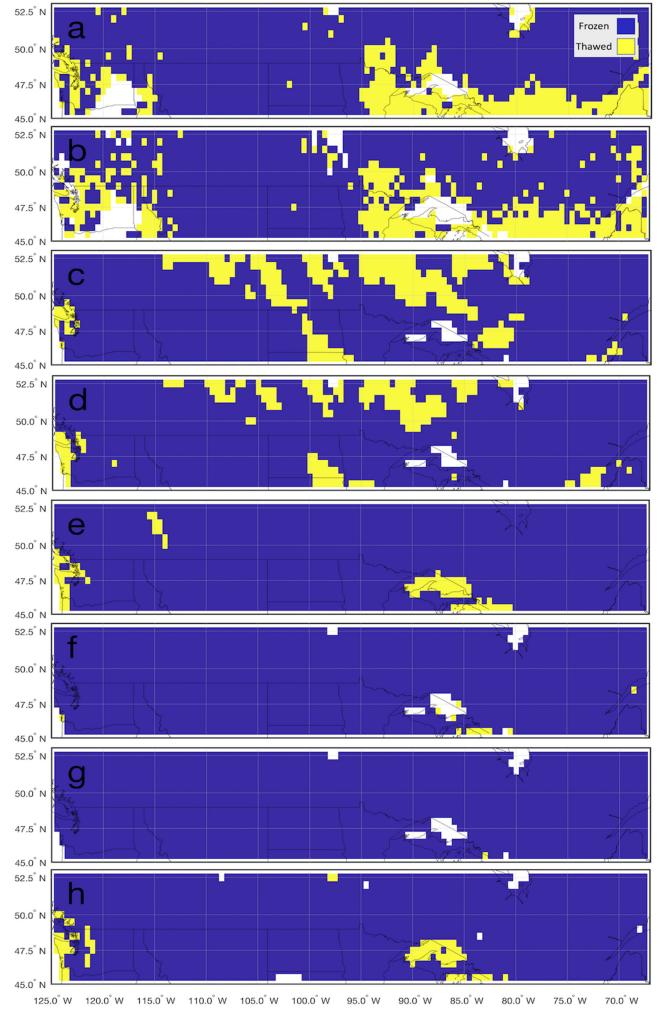


Fig. 2. Maps of FTSPs and 0°C threshold-based temperature products fit to a 0.5° grid (11/29/2015 A.M. Overpass): (a) SMAP-EFT (9 km), (b) SMAP-FT (36 km), (c) AMSR-FT (25 km), (d) SSMI-FT, (e) NLDAS $T_{2m-\text{air}}\text{-FT}$, (f) $T_{\text{soil}0-10\text{cm}}\text{-FT}$, (g) $T_{N-\text{skin}}\text{-FT}$, and (h) $T_{M-\text{skin}}\text{-FT}$.

reasonable proxy for examining FT state under the assumption that phase changes at the surface will begin to occur at this temperature. After some investigation into threshold optimization, we decided to avoid setting this value above or below 0°C as to avoid the introduction of an additional warm/cold bias.

First, temperature data and FTSPs are matched to a common 0.5° grid. Fig. 2 shows a FT snapshot from the A.M. overpass on 11/29/2015 in which snow cover was prevalent over much of the domain. Variations in the FT state are apparent, most noticeably during seasonal transition periods when diurnal and ephemeral freeze events occur frequently. As stated, complexities related to sensor emission depth, wet/dry snow, surface water, and varied land covers can exhibit significantly different retrieval responses dependent on time, resolution, and sensing frequency. Overall CAs are calculated for all overpasses within the study period by comparing the two binary datasets. CA quantifies the agreement between the temperature products (NLDAS and MODIS) and the FTSPs across the study region on a 0.5° grid using a cell-by-cell

comparison

$$CA = \frac{FTy}{FTy + FTn} \quad (1)$$

where FTy indicates agreement and FTn disagreement between data pairs. Large water bodies and grid cells missing frequent temperature or FT state classifications were excluded. Additionally, MODIS Aqua retrievals are limited in coverage due to cloud cover, therefore, all regions are not equally represented. Cells containing no data in either comparison member are excluded from overall CA calculations. Previously, similar CA-based approaches have been performed over China utilizing both AMSR and SSMI-based FT classifications [39].

B. Proportional Comparison Methodology

In fitting data to a coarser resolution grid, the use of simple averaging can result in mischaracterizing the true ground state, especially in regions with complex topography. Alternately, a grid-by-grid proportional analysis can be adopted. Frozen proportions are calculated as the number of defined frozen values divided by the total number of data points within each 0.5° grid cell. After these values are calculated for both temperature-based FT classes and FTSPs, a difference in frozen proportion (D_{FTp}) is defined by subtracting temperature product frozen proportion (F_{Tp}) from the corresponding FTSP frozen proportion (F_{FTp})

$$D_{FTp} = F_{FTp} - F_{Tp}. \quad (2)$$

For example, if all NLDAS air temperatures ($1/8^\circ$) within a 0.5° grid cell (16 values) are below the freezing point (0°C) and only half of the FTSP pixels (containing 2–25 values, dependent on cell location and FTSP resolution) are classified as frozen, D_{FTp} would be -0.5 , signifying 50% less frozen area from the FTSP compared to the temperature-based frozen area within the specified cell. This method avoids effects of significant temperature variations and rounding errors incurred when upscaling FT state variables to a relatively coarse 0.5° grid. The proportional comparison methodology also enables a spatial representation in which positive values represent greater FTSP defined frozen area as compared to temperature products, whereas negative values define a lesser FTSP defined frozen area in relation to temperature threshold-based products. This captures FTSP spatial tendencies to either overestimate or underestimate frozen proportion relative to MODIS and NLDAS temperature data that are not captured by CA comparisons. This methodology moves beyond the binary agreement approach outlined in Section II-A and contributes to the identification of the causes of FT classification disagreement.

C. F/T and Temperature Distributions

To determine the temperature variables with the strongest relationship to FTSP states, multiple methods were employed. These utilize FT classifications as matched with the corresponding temperatures on a grid cell-by-cell basis over the entire study period. As a result, all corresponding temperature values deemed frozen and thawed by each of the FTSPs were extracted and used to define two unique probability distributions Functions (PDFs).

These distributions are summarized by boxplots employed in Section III-C. The PDF overlap proportions were analyzed for each comparison by quantifying the proportional overlay area between the two distributions. This methodology is used to enumerate the degree to which land surface temperature products distinguish FT states similar to those derived from P-MW emissions, where perfectly unique FT state distributions are equal to 0, and identical distributions equal to 1. Additionally, the value ranges at which land surface temperatures diverge in FT classification are also identified. These temperature values are determined as those in which 1% of the data has been classified below a given temperature as thawed, and above as frozen (see Table V). In an effort to summarize critical PDF characteristics, the 1% threshold was employed to remove outliers in which values at low or high temperatures are classified as both frozen and thawed. Similarly, the two-sample Kolmogorov–Smirnov test is also employed to verify the relationship between the FT distributions. The test statistic is calculated using the greatest difference between the frozen and thawed cumulative distribution functions, where larger values indicate increasingly different distributions [42], [43]. In summary, these methodologies do not depend on cell-by-cell agreement between temperature threshold-based FT classifications and FT product-based classifications, but they rather focus on temperature product values as directly associated with FTSP classifications.

III. RESULTS AND DISCUSSION

Differences between FTSPs and modeled/observed temperature (T_{2m-air} , $T_{soil0-10cm}$, T_{N-skin} , and T_{M-skin}) are investigated in an effort to determine, which products act as the best surrogates to FTSP classifications spatially, seasonally, and in regard to temperature biases. Variations between morning (A.M.) and afternoon (P.M.)-based FT retrievals are also analyzed.

A. Classification Accuracy

First, we examined the CA (1) of the FT products with respect to various temperature products during April 2015–December 2016 (see Table III). Overall, FTSPs appear to agree best with T_{2m-air} (81%–91%) and show the least agreement with T_{N-skin} (76%–85%). However, $T_{soil0-10cm}$ shows a comparatively stronger relationship with SMAP FTSPs than with FT-ESDR records, likely due to the deeper emission depth of SMAP P-MW retrievals. MODIS T_{M-skin} -based FT classifications show intermediate agreement with all products and are similar to soil temperature for AMSR and SSMI-FT products. An improvement in CA (+2.0%) of T_{M-skin} was observed when utilizing the higher resolution SMAP_E – FT (9 km) product compared to SMAP-FT (36 km). This is potentially due to the improved ability to resolve smaller scale variability captured by the higher resolution MODIS product ($\sim 5\text{ km}$). Increases in CA from SMAP-FT to SMAP_E – FT ranged from 1.7% to 2.4% across all comparisons. Noticeably, all SMAP FTSPs showed a clear reduction in CA as compared to FT-ESDR products (see Table III). This is investigated further by examining proportional frozen percentage and corresponding frozen and thawed temperature distributions in the following sections.

TABLE III
CA RESULTS FOR ALL PRODUCTS (APRIL 2015–2016), OVERALL AND DIVIDED SEASONALLY, ORDERED BY OVERALL CA PER PRODUCT

Freeze/Thaw Product	Comparison (NLDAS Domain >45° latitude)	Classification Accuracy (0°C threshold) April 2015 - 2016	2015			2016		
			AMJ	JAS	OND	JFM	AMJ	JAS
SMAPe-FT	NLDAS 2m Air Temperature	0.835	0.84	0.88	0.77	0.80	0.85	0.90
	NLDAS 0-10cm Avg Soil Temperature	0.822	0.81	0.88	0.78	0.75	0.83	0.90
	MODIS Skin Temperature	0.786	0.76	0.86	0.70	0.80	0.78	0.87
	NLDAS Skin Temperature	0.777	0.80	0.86	0.67	0.72	0.82	0.88
SMAP-FT	NLDAS 2m Air Temperature	0.811	0.81	0.83	0.75	0.80	0.83	0.85
	NLDAS 0-10cm Avg Soil Temperature	0.800	0.79	0.83	0.76	0.75	0.81	0.85
	MODIS Skin Temperature	0.766	0.74	0.81	0.69	0.80	0.77	0.83
	NLDAS Skin Temperature	0.760	0.78	0.82	0.67	0.72	0.80	0.84
AMSR-FT	NLDAS 2m Air Temperature	0.913	0.94	1.00	0.80	0.86	0.94	1.00
	NLDAS 0-10cm Avg Soil Temperature	0.874	0.89	1.00	0.72	0.79	0.91	1.00
	MODIS Skin Temperature	0.860	0.84	0.98	0.69	0.86	0.86	0.97
	NLDAS Skin Temperature	0.849	0.84	0.97	0.71	0.82	0.86	0.97
SSMI-FT	NLDAS 2m Air Temperature	0.898	0.94	0.99	0.75	0.84	0.94	1.00
	NLDAS 0-10cm Avg Soil Temperature	0.874	0.90	1.00	0.71	0.80	0.91	1.00
	MODIS Skin Temperature	0.864	0.83	0.98	0.71	0.87	0.85	0.97
	NLDAS Skin Temperature	0.841	0.90	0.99	0.63	0.78	0.90	0.99

Seasonal analyses provide information on the dynamics of FTSPs during FT onset and during the winter frozen and summer thaw periods. As anticipated, during summer (JAS), when temperatures are largely well above freezing, all temperature products perform the best at classifying the FT state (see Table III). As the fall transitions (OND) to winter freeze (JFM), NLDAS air (CA 75%–81%) and top layer soil temperatures (CA 76%–82%) become the best surrogates to SMAP FT states, as characterized by a CA increase relative to skin temperatures of 6%–13% during the fall transition period. While this relationship also holds for *Ka*-band FT-ESDR products, it is less pronounced (0%–12%). Additionally, during the spring thaw period (AMJ), slight increases on the order of 3%–12% in CA are observed across most members compared to winter period while a reduction in CA was observed when using $T_{M\text{-skin}}$ (0% to –3%). All comparison members show a relative increase in CA when comparing the 2015 freeze onset to the 2016 spring period (5%–17%). This indicates improved agreement between temperature products and FTSPs when identifying FT conditions during the thaw period (CA 74%–94%) as compared to the freeze onset (CA 63%–86%). Additionally, FTSP defined state across all products is in comparatively poor agreement with NLDAS skin temperature (see Table III). This is probably a result of sensor sensitivity deeper than the surface skin temperatures, in addition to potential NLDAS model deficiencies in representing temperatures of complex snow-covered surfaces. In support of this, FTSPs are shown to have better agreement with MODIS $T_{M\text{-skin}}$ as the measured (MODIS) and modeled (NLDAS) skin temperatures exhibited a comparative increase during JFM of 4%–9%. Small variations in CA between AMSR and SSMI-FT can likely be attributed to the varied overpass times (see Section III-B) and slight variations in P-MW depth sensitivity of SSMI (36.5 GHz) as compared to AMSR (37 GHz).

Large increases in CA are also observed during the primary freeze onset period (OND) from 2015 (CA 63%–80%) to 2016 (CA 69%–86%). The presence of near record high temperatures in December 2015 across the domain, likely resulting in more

ephemeral FT-events, may be to blame [44], [45]. Depending on the characteristics of freeze and thaw onset as well as snow cover variability, CA is expected to vary annually. An uptick in accuracy over nearly all CA proportions occurs during the mid-winter period (JFM), as classified FT state and surface temperatures are continuously frozen across much of the domain. This is true for all but SMAP FTSP comparisons with soil temperature, in which the thermal insulation properties of snow cover play a role. This is investigated in Section III-B as SMAP exhibits a reduced frozen area relative to FT-ESDR in March. This indicates SMAP defining a smaller frozen area as compared to ESDR FT products during the spring melt onset. Overall, high CA values for both $T_{M\text{-skin}}$ and $T_{2m\text{-air}}$ best characterized FT products during the JFM period. Limited spatial coverage of MODIS-based temperatures may play a role in this result, though our results suggest MODIS observed skin temperature as a reasonable proxy to FTSPs. During spring thaw (AMJ), $T_{2m\text{-air}}$ again appears to be the best surrogate to P-MW measured FT state. As P-MW-based RS techniques begin to observe water in the landscape, FT products appear to closely resemble $T_{2m\text{-air}}$, since air temperature tends to define surface melt onset. Soil and skin temperatures tend to continue to reflect the temperatures of snow insulated soil and surface snow temperatures making these products less likely to indicate the presence of ephemeral thawed conditions as air temperatures creep above freezing [9].

CA maps follow similar patterns across temperature product types (see Appendix A and Fig. 3). FT-ESDR products show comparable accuracy across the domain with higher CA in homogenous regions without prevalent surface water features and complex topography. FT-ESDR products show highest CA in regions downwind of the rocky mountains and slightly reduced CA in eastern portions of the domain with complex vegetation. However, comparisons of FT-ESDR products to MODIS skin temperature show marked improvements in CA over mountainous regions as compared to NLDAS parameters ($T_{N\text{-skin}}$ and $T_{\text{soil}0-10\text{cm}}$). SMAP FTSPs illustrate similar dynamics to an exaggerated degree, as decreased CA values are prevalent in the

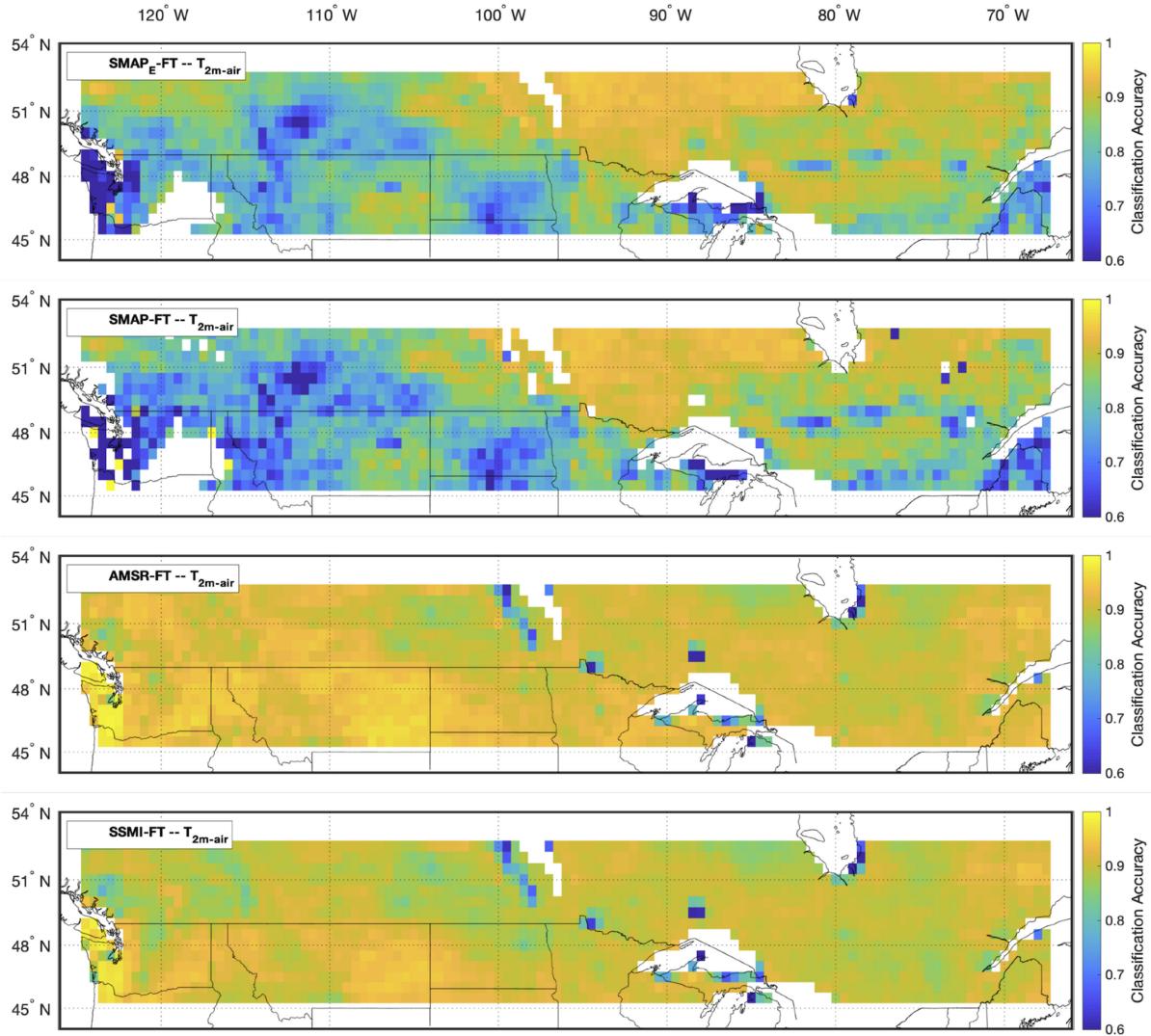


Fig. 3. Maps of CA for SMAP-FT and FT-ESDR products as compared to NLDAS 2-m air temperature ($T_{2m\text{-air}}$) [April 2015–2016]. White areas signify no comparable data/surface water. Refer to Appendix A for all CA spatial distribution maps.

rocky and cascade mountain regions. In contrast to FT-ESDR products, SMAP FTSPs exhibit the strongest relationship to temperature variables in central Canada. The reductions in CA across portions of western and eastern Canada, central North Dakota, and Maine are primarily attributed to false freeze flags (see Section III-C). Notably, these areas have a similar high prevalence of surface water features. Continued analysis into the regional dynamics driving these results will be assessed through a detailed analysis of differences in grid cell frozen proportion.

B. Proportional Differencing Analysis

By utilizing the proportional differencing methodology outlined in (2), comparisons between FT and temperature-based products were completed. The analysis was broken down to examine trends encompassing the entire study period as well as timeseries trends and through a separation of A.M. and P.M. FT classifications. These results provide insight into biases between

temperature-based and FTSP classifications both regionally and temporally.

1) Proportional Summary Results: Two statistical measures are computed for each FT/temperature comparison as follows: first, mean, representing FTSPs tending to overestimate (+) or underestimate (−) the temperature threshold-based frozen proportion and second, absolute mean, used to represent the magnitude of overall average difference. The absolute average results mirror what was found in the CA analysis, with all FT products most closely comparable to $T_{2m\text{-air}}$ (see Table IV). However, there is decreased separation between MODIS and NLDAS skin temperature metrics. Overall, the results indicate an average frozen proportion mismatch across all grid cells ranging from 8.9% (AMSR-FT, $T_{2m\text{-air}}$) to 24.7% (SMAP_E – FT, $T_{N/M\text{-skin}}$). In contrast to CA results, the 9-km SMAP_E – FT has an increased average magnitude of disagreement as compared to the 36-km SMAP-FT. Presumably, the use of significantly more data points introduced increased variability in

TABLE IV

PROPORTIONAL DIFFERENCING SUMMARY VALUES, COMPUTED AS AVERAGE ACROSS ALL PIXELS AND COMPLETE TIME PERIOD [APRIL 2015–2016];
 PROPORTIONAL MEAN (D_{FTP}) AND ABSOLUTE PROPORTIONAL MEAN $D_{FTP-abs}$ ORDERED FROM LOWEST TO HIGHEST
 $D_{FTP-abs}$ AND SEPARATED BY EACH F/T PRODUCT

		Proportional Differencing Results (D_{FTP})	
Freeze/Thaw Product [minus]		Temperature Product	
		D_{FTP}	$D_{FTP-abs}$
SMAP _E -FT	NLDAS 2m Air Temperature	0.007	0.189
	NLDAS 0-10cm Avg Soil Temperature	-0.026	0.204
	MODIS Skin Temperature	-0.087	0.247
	NLDAS Skin Temperature	-0.088	0.247
SMAP-FT	NLDAS 2m Air Temperature	-0.004	0.179
	NLDAS 0-10cm Avg Soil Temperature	-0.036	0.195
	NLDAS Skin Temperature	-0.099	0.240
	MODIS Skin Temperature	-0.098	0.243
AMSR-FT	NLDAS 2m Air Temperature	-0.035	0.089
	NLDAS 0-10cm Avg Soil Temperature	-0.083	0.131
	NLDAS Skin Temperature	-0.129	0.155
	MODIS Skin Temperature	-0.141	0.163
SSMI-FT	NLDAS 2m Air Temperature	-0.050	0.104
	NLDAS 0-10cm Avg Soil Temperature	-0.082	0.130
	MODIS Skin Temperature	-0.132	0.159
	NLDAS Skin Temperature	-0.130	0.159

D_{FTP} between overpasses. Nearly all D_{FTP} comparison values shown in Table IV are negative, indicating that during the study period most temperature products defined more frozen area as compared to FTSPs (2). The use of a lower temperature threshold ($<0^{\circ}\text{C}$) would result in an increased D_{FTP} by reducing the frozen extent accordingly. However, when using a threshold at 0°C , the only case in which D_{FTP} was positive occurred when comparing SMAP_E – FT to T_{2m-air} (see Fig. 4). These results support the ability of P-MW RS techniques to capture phase changes in the surface that do not generally occur until temperatures drop below 0°C for an extended period. The documented NLDAS cold biases are also a likely contributor (see Section II) [32], [35], [36], [38]. The prevalence of summer false freeze classifications (see Section III-C) in SMAP FTSPs is likely to affect this tendency, as overestimated regions skew the mean result. Even so, SMAP_E – FT only defined a larger frozen extent by less than 1% when averaged across cells. Thus, FTSPs used in this paper are drastically more likely to define a lesser frozen extent than temperature threshold-based FT classifications, especially in the eastern portion of the domain. $T_{M\text{-skin}}$ is shown to perform the poorest at mirroring satellite products, as $T_{M\text{-skin}}$ defined frozen proportions estimated a greater frozen area than FTSPs by 8.8%–14.1% on average. This may be due in part to higher resolution retrievals and sensitivity of P-MW RS methods to deeper surface state. AMSR- and SSMI-based FTSPs again closely mirrored one another (see Fig. 4 and Appendix B). With FT-ESDR products tending toward underestimation of temperature-based frozen proportions across a wider area of the domain (compared to SMAP FTSPs), with the closest FT to temperature agreement occurring in the western plains.

FT products visually follow similar patterns of variation in D_{FTP} regionally, as compared to CA results when averaged over the study period (see Fig. 4 and Appendix B). However,

regional trends are highlighted by FTSPs exhibiting less or more frozen area than temperature products. In central Canada, FTSPs show increased variability as compared to surface temperatures with D_{FTP} ranging from around -0.1 (FT-ESDR) to -0.05 to 0.05 (SMAP-FT). Eastern regions tend toward underestimation of temperature defined frozen area, whereas more mountainous regions display increased heterogeneity in classification D_{FTP} (see Fig. 4). This result can be expected as a byproduct of imperfect modeled temperature parameters, limited spatial resolution affecting the ability to resolve smaller scale topographic features, and known deficiencies in FTSPs at observing FT dynamics across complex topography [17], [27], [46]. A reduction in D_{FTP} from SMAP-FT to SMAP_E – FT in portions of the domain indicate improvements in the enhanced SMAP product's ability to resolve complex features (coastal areas notwithstanding). Interestingly, there appears to be extensive regions where F_{FTP} exceeds F_{Tp} focused on the leeward side of the mountains, which tend to be inherently drier. While they do share similarities, these results verify clear spatial inconsistency across FTSPs as compared to various land surface temperature variables. However, the small magnitude of proportional difference suggests that the satellite observed FT state is closely linked to the variables, especially air temperature.

A time-series-based analysis was performed utilizing monthly T_{2m-air} , $T_{soil0-10cm}$, $T_{N\text{-skin}}$, and $T_{M\text{-skin}}$ mean D_{FTP} values (see Fig. 5). Similar to CA, temperature-based frozen proportion values show the strongest relationship with FTSPs May through October. FT-ESDR products closely resemble NLDAS air and soil temperatures during this period as D_{FTP} values fall well within the -0.1 to 0.1 range signifying very similar temperature-based and FT product classifications (see Fig. 5). Skin temperatures tend to lead air and soil temperature in estimating increased frozen extent relative to FTSPs by September due to

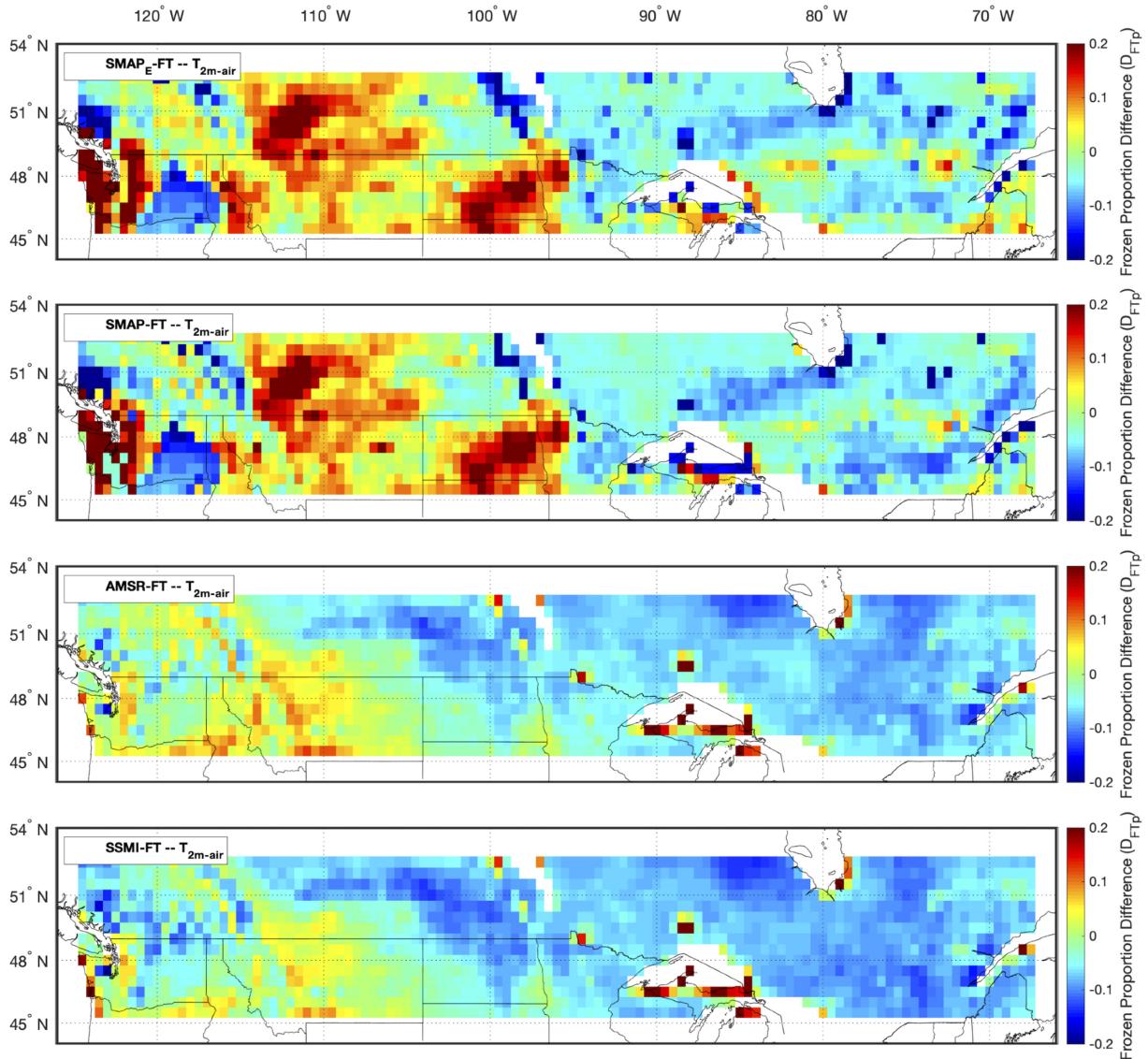


Fig. 4. Overall (2015–2016) mean proportional differences (D_{FTp}) maps of FT products to NLDAS 2-m air temperature for SMAP-E-FT, SMAP-FT, AMSR-FT, SSMI-FT. See Appendix B for all D_{FTp} maps.

the onset of freeze events. Additionally, FTSP defined frozen extent is exceeded by NLDAS $T_{N\text{-skin}}$ throughout much of the study period as MODIS skin temperature shows comparatively increased agreement with FTSPs through transitional periods and the January–February extended frozen period (see Fig. 5). For SSMI-FT, MODIS skin temperature shows a closer relationship than air temperature to FTSPs during the mid-winter period. As discussed, FTSPs exhibited reduced agreement to temperature products during freeze and thaw onset tending to underestimate temperature derived frozen proportions. Potential false freeze flags in SMAP FTSPs exhibit the opposite, as mean D_{FTp} values show SMAP overestimating temperature-based frozen proportions even in mid-summer. Summer frozen areas in SMAP products are contributors to the increased absolute mean D_{FTp} values and reduced CA over the summer period relative to ESDR FT products. However, agreement between temperature and FT products continues to be at its lowest during the shoulder

seasons. This is likely due to increased ephemeral and diurnal FT events which tend to occur across the domain in spring and fall.

2) *Proportional A.M./P.M. Overpass Variations:* All FTSPs evaluated as part of this paper are derived from both descending (A.M.) and ascending (P.M.) overpasses (see Table I). Previous studies have identified significant variations in the relationship between FT classifications and temperature variables dependent on overpass time [8], [9]. Characteristically, morning overpasses (completed at either 1:30 A.M. or ~6:00 A.M. LST) have cooler temperatures and little solar exposure as compared to afternoon overpasses at 1:30 P.M. or 6:00 P.M. As a result, A.M. FT classifications have proven more susceptible to mismatches with surface temperature due to diurnal freeze events in which surface temperatures may creep below freezing while satellite FT retrievals continue to measure thawed surface emissions, or vice versa.

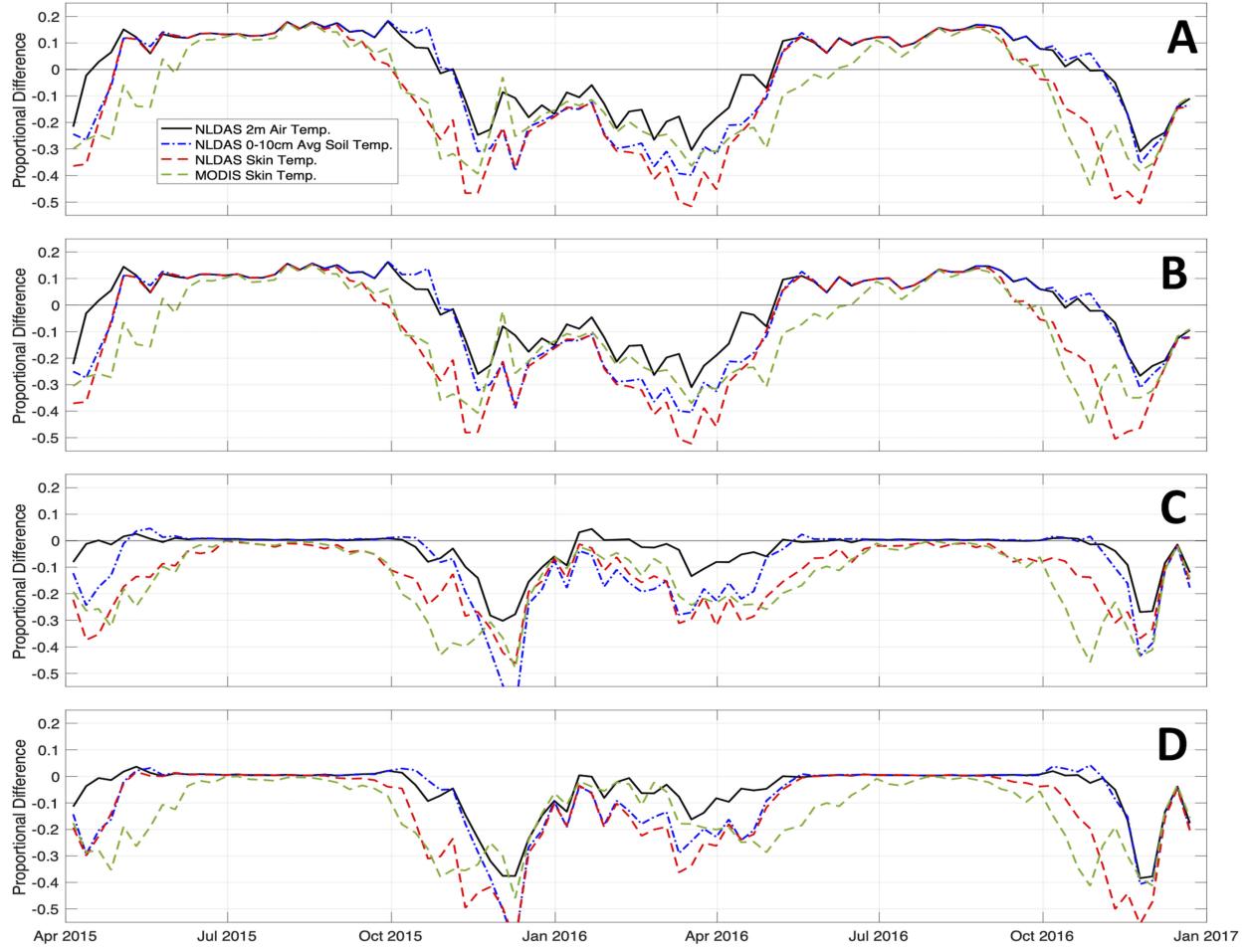


Fig. 5. Timeseries of mean frozen proportion differences (D_{FTp}) of MODIS skin, NLDAS 2-m air, NLDAS 0–10-cm soil, and NLDAS skin temperatures as compared to FT products averaged over complete domain. (a) SMAPE-FT, (b) SMAP-FT, (c) AMSR-FT, and (d) SSMI-FT.

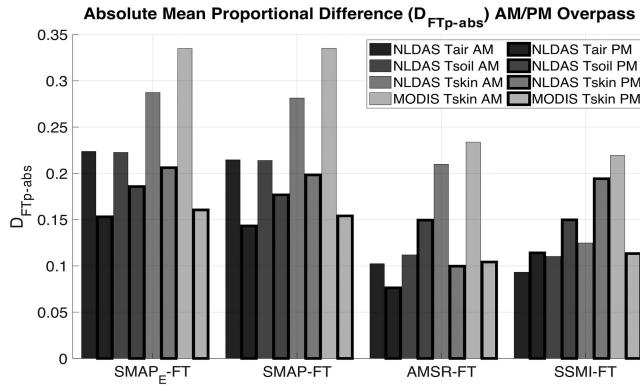


Fig. 6. Absolute mean proportional differences ($D_{FTp} - \text{abs}$) separated into A.M. and P.M. periods for all comparisons. Unbounded bars indicate A.M. $D_{FTp} - \text{abs}$, and bounded P.M.

As presented in Fig. 6, the general rule holds true in our analysis as absolute mean proportional differences are significantly reduced from A.M. to P.M. retrievals across nearly all comparison members. This is most notable in SMAP FTSPs and comparisons involving $T_{M-\text{skin}}$. The exceptions include SSMI to $T_{2m-\text{air}}$,

$T_{\text{soil}0-10\text{cm}}$, $T_{N-\text{skin}}$ and AMSR to $T_{\text{soil}0-10\text{cm}}$. As shown across all FTSPs to $T_{\text{soil}0-10\text{cm}}$ comparisons, the relatively small reductions or increases in absolute D_{FTp} can be attributed to the ability of soil to insulate itself. As a result, soil temperature tends to be less susceptible to slight temperature variations and avoids falling below freezing during ephemeral freeze events in which FT classifications remain thawed. Conversely, soil can also retain colder temperatures as surface thaws, as the presence of wet surfaces may result in thawed P-MW-based FT classifications while soil remains frozen at depth. The most significant reduction in absolute D_{FTp} occurs in all FTSP comparisons with MODIS skin temperature. This occurs as a result of MODIS defining far more frozen extent in the A.M. as compared to P.M. observations relative to FTSPs.

Additionally, variations in A.M./P.M. performance in FT-ESDR products can be partially attributed to varied overpass times in which early morning AMSR retrievals tend to define comparatively more frozen area relative to SSMI-FT. Opposite to other FTSPs, SSMI-FT compared to NLDAS air, skin, and soil temperatures presents an increase in absolute D_{FTp} from A.M. to P.M. There are several complex dynamics driving the FT classifications of FTSPs, including surface characteristics, sensor

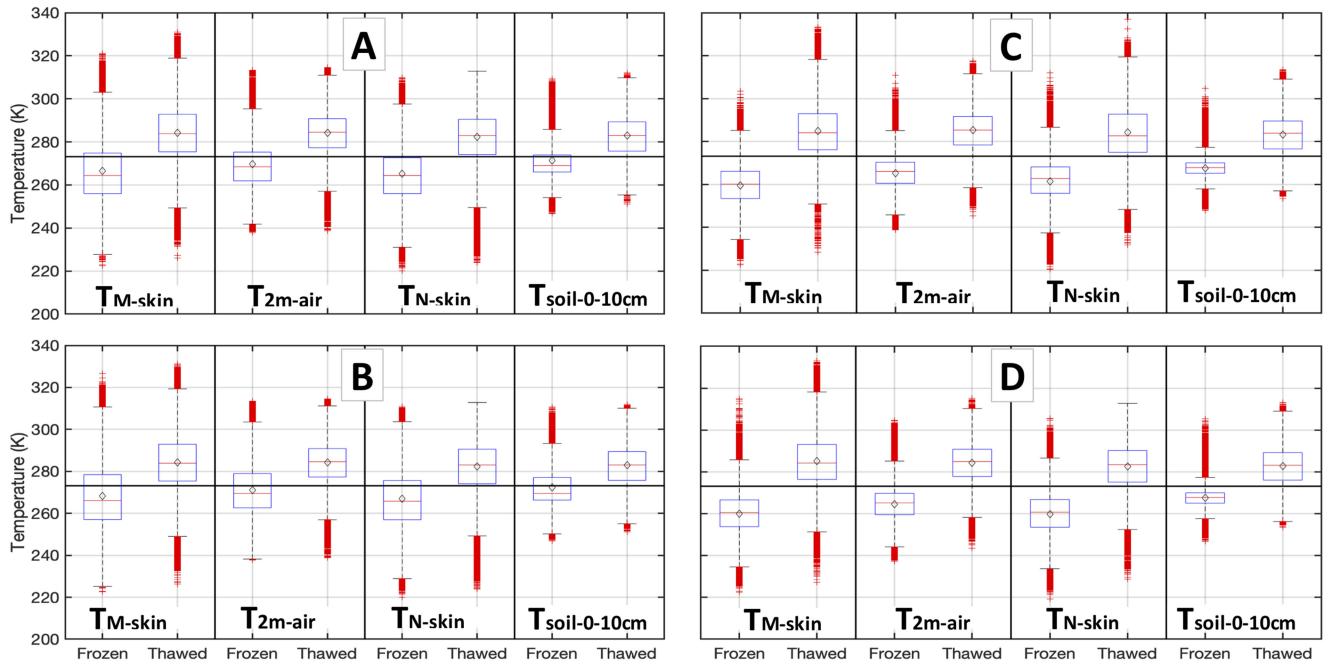


Fig. 7. Boxplot distributions of select MODIS and NLDAS temperature datasets ($T_{M\text{-skin}}$, $T_{2m\text{-air}}$, $T_{N\text{-skin}}$, and $T_{\text{soil}0\text{-}10cm}$) compared to (A) SMAPE-FT, (B) SMAP-FT, (C) AMSR-FT, and (D) SSMI-FT. Line at freezing point for reference, diamond marks mean.

configurations, and overpass times. During transitional periods where ephemeral FT events are common, slight temperature differences can lead to large differences in temperature threshold-based FT classifications. This is especially pronounced when dealing with varied retrieval times of different FTSPs. A precise time-targeted approach will need to be conducted to draw improved conclusions across different satellite-based FT products.

C. F/T Class Identification

This section investigates which temperature datasets most clearly define unique frozen and thawed temperature distributions. In contrast to comparing FT state to temperature at the grid scale through thresholding, this approach separates frozen and thawed states as defined by FTSPs and associates them with the corresponding $T_{2m\text{-air}}$, $T_{\text{soil}0\text{-}10cm}$, $T_{N\text{-skin}}$, and $T_{M\text{-skin}}$ temperatures.

1) *Distribution Boxplots:* Following the association of NLDAS and MODIS observed temperatures with FT states, distributions are created for each temperature variable (see Fig. 7). The plots in Fig. 7 show clear variations between frozen and thawed value ranges across all temperature variables. In most cases, a clear overlap is observed between FT distributions well above and below the freezing point of 273 K. However, temperature distributions derived from FT-ESDR do not show overlap between the central 50% of the data in any case, whereas SMAP products frequently do so (see Fig. 7). Much of this can be attributed to frequent SMAP FT classifications of regions as frozen while surface temperature variables indicate values >300 K. Slight reductions in these high temperature frozen classifications were observed when using SMAP_E – FT (9 km) as compared to

SMAP-FT (36 km). As a result, slight increases were observed in separation between FT distribution values as compared to SMAP-FT reflected by a shift of frozen distribution median values by 3–5 K. This increased separation between distributions supports the idea that SMAP_E – FT shows improvement at accurately characterizing surface FT state relative to SMAP-FT. The histogram-based analysis in the following section further investigates the cause of such discrepancies between FTSPs and temperature products.

2) *F/T State PDFs:* PDF histograms are used here to visualize FT classification dynamics. By calculating the proportion of overlap between PDFs, we are able to quantify which associated temperature values defined the most unique frozen and thawed distributions relative to various FTSPs. A minimal overlap signifies a comparison in which the frozen and thawed states were most clearly discerned by temperature variables (see Table V). As expected, all PDFs show highest values of overlap and associated classification uncertainty at the freezing point generally decreasing logarithmically away from this point in both directions (see Fig. 8). The resulting overlap proportions help to verify that none of the temperature variables defined particularly unique FT distributions when calculated with SMAP products (>20.3% overlap). This is compared to overlap values as low as 9% when comparing NLDAS $T_{2m\text{-air}}$ to associate AMSR FT classifications (see Fig. 8). Additionally, in order to summarize the characteristics of each PDF, the temperature at which 1% of data is classified below as thawed (Low) and above as frozen (High) are presented in Table V. A Two-sample Kolmogorov–Smirnov test is performed as an additional method to determine the relationship between the two states temperature probability distributions. A larger test statistic indicates that the

TABLE V

MEAN TEMPERATURE OF FROZEN AND THAWED PDFS, OVERLAP PROPORTION VALUES OF PDFS, DISTRIBUTION BOUNDS SUMMARY CHARACTERISTICS, AND TEST STATISTIC FROM TWO-SAMPLE KOLMOGOROV–SMIRNOV TEST COMPARING FROZEN AND THAWED TEMPERATURE DISTRIBUTIONS

Comparison Members	Distribution Mean Temperature (K)		PDF Overlap Proportion	Temperature (K) of Uncertain FT Classification: 1% of data classified below Low (as thawed) and above High (as frozen)		Two Sample Kolmogorov– Smirnov Test Statistic	
	Frozen	Thawed		Low	High		
Freeze/Thaw Product	NLDAS-ST 0-10cm	271.4	282.9	0.203	266.0	296.0	0.596
	NLDAS-2m	269.8	284.2	0.204	263.8	297.4	0.593
	NLDAS-SKT	265.3	282.4	0.213	256.3	297.4	0.574
	MODIS-SKT	266.4	284.3	0.231	257.2	306.6	0.539
	NLDAS-ST 0-10cm	272.4	282.9	0.230	265.8	297.2	0.541
	NLDAS-2m	271.0	284.3	0.230	262.9	298.9	0.541
	NLDAS-SKT	266.9	282.4	0.239	255.3	298.6	0.523
	MODIS-SKT	268.2	284.3	0.257	256.4	307.5	0.485
	NLDAS-2m	265.2	285.4	0.090	267.8	280.7	0.820
	MODIS-SKT	259.6	284.9	0.099	259.8	279.4	0.803
AMSR-FT	NLDAS-SKT	261.3	284.4	0.103	263.2	278.4	0.794
	NLDAS-ST 0-10cm	267.7	283.2	0.111	265.8	279.4	0.779
	NLDAS-2m	264.5	284.4	0.099	266.0	281.0	0.802
	MODIS-SKT	260.0	285.1	0.103	260.1	281.0	0.795
	NLDAS-ST 0-10	267.5	282.9	0.108	265.6	279.3	0.785
SSMI-FT	NLDAS-SKT	259.7	282.6	0.110	257.9	279.3	0.779

Larger KS test statistic indicates larger variation between cumulative distribution functions.

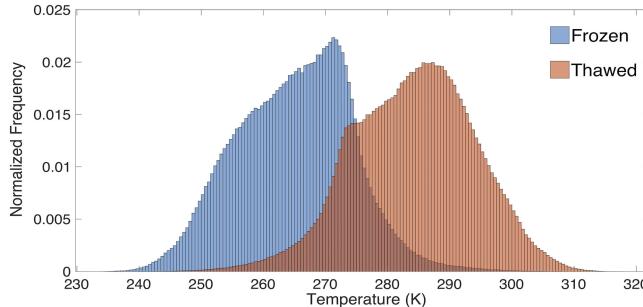


Fig. 8. PDF histogram of AMSR-FT defined frozen and thawed states matched to corresponding NLDAS air temperatures; frozen distribution to the left, overlap shown in intermediate shade.

two distributions are increasingly varied. The resulting outputs (see Table V) perfectly mirror the PDF overlap proportions, providing verification in identifying which temperature/FTSP combinations provide the most unique frozen and thawed temperature classifications.

Notably, the comparisons showed very distinctive temperature distributions between both T_{2m-air} and T_{M-skin} when compared to FT-ESDR products. For FT-ESDR, with overlap proportions ranged from 9.0% to 9.9% (T_{2m-air}) to 9.9%–10.3% (T_{M-skin}). MODIS distributions were derived utilizing around 300 000 frozen and 900 000 thawed classifications. This amounted to about 1/3rd the total data-points as compared to NLDAS parameters, due to MODIS temperature data availability. Even so, the ability of T_{M-skin} to define clearly separate temperature distributions when associated with ESDR FT classifications can help justify the use of satellite-based thermal products in estimating and evaluating the accuracy of FT states, especially in remote regions. Using MODIS skin temperature in efforts to validate SMAP-FT products, however, may not be as suitable due to a sensor observation deeper than the surface skin layer. As the highest overlap proportions for SMAP products ranged from 21.3% to 25.7% when associated with skin temperatures.

Characteristics of each temperature/FTSP PDF are uniquely representative of variations in FTSP product over the domain. These characteristics are especially useful when evaluating whether skin temperature, soil temperature, or air temperature is most representative of current FTSPs along with associated measurement dynamics. In the case of T_{N-skin} , we were able to identify differences in representation of ground FT state as compared to FTSPs. Having a considerably higher frequency of thawed classifications at the freezing point, T_{N-skin} appears to over represent the frequency of temperatures around 273 K compared to other temperature variables (see Fig. 9). NLDAS modeled T_{N-skin} has minimal thermal inertia and is very sensitive to the longwave radiation balance [35], [36]. Therefore, it may cool quickly and not be represented in the broader landscape FT state as observed using P-MW emissions. As $T_{soil0-10cm}$ is related to surface skin temperature in NLDAS assimilation systems, the freezing point bias can also be identified in the resulting $T_{soil0-10cm}$ PDFs (see Fig. 10). These peaks can also be associated with the increased energy required in phase changes, in which a surface can remain at the freezing point for an extended period of time. Due to NLDAS model physics [32], [35], [36], it is possible for the modeled temperatures to hover slightly above freezing when in reality much of a surface area has fallen below freezing and may be deemed frozen by FTSPs. Since SMAP FTSPs show a visible skew toward high temperature freeze classifications [see Figs. 9(a), 10(a)], we will focus on discussing the characteristics of FT-ESDR derived temperature PDFs. Further analysis into SMAP products should be completed as potential false freeze classifications are verified and removed.

As mentioned, T_{2m-air} (9.0%–9.9%) and T_{M-skin} (9.9%–10.3%) values as associated with FT product classifications exhibit minimal overlap proportions. These temperature variables along with T_{N-skin} are also largely characterized by overlap symmetry [Figs. 8, 9(b), 10(b)]. In contrast, $T_{soil0-10cm}$ exhibits a slight bias toward increased thaw classifications around 270 K when compared to all FTSPs (see Fig. 10). This supports the tendency of P-MW FT classifications to observe the top few

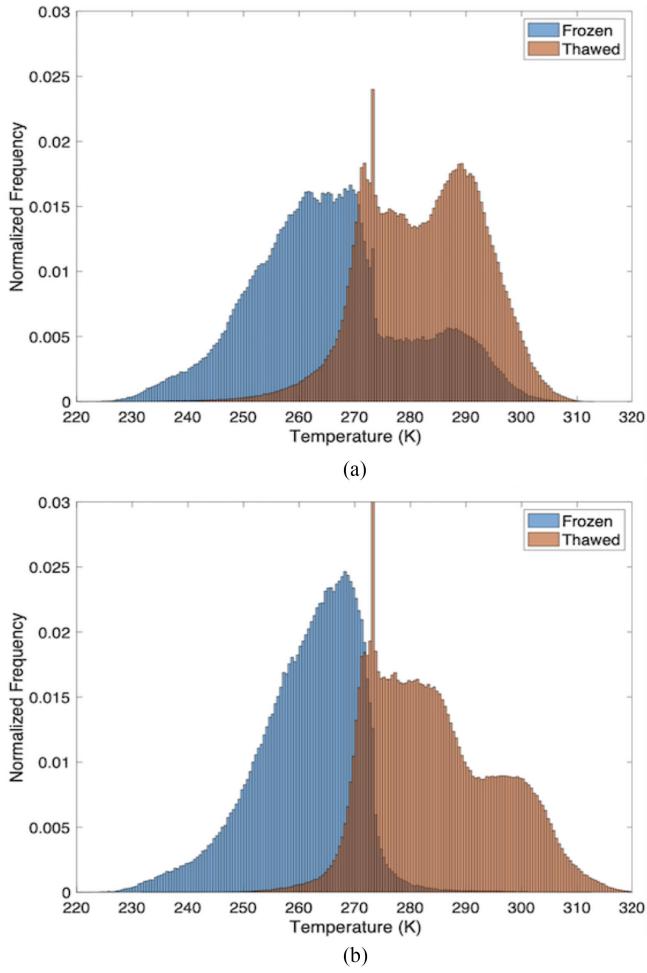


Fig. 9. PDF histograms of SMAPE – FT (a) and AMSR-FT (b) defined frozen and thawed states matched to corresponding NLDAS skin temperatures; frozen distribution to the left, overlap area shown in intermediate shade.

centimeters of the surface, as soil temperatures can remain below freezing, while the P-MW observation indicates surface thaw. This increase in thawed classifications below the freezing point for $T_{\text{soil}0-10\text{cm}}$ can also be related to periods in which wet snow is prevalent at the surface resulting in a thawed P-MW Tb observation, even as soil temperatures are modeled to be frozen. In reality, this can also occur in the opposite direction as soils remain thawed even as surface temperatures indicate frozen conditions. This illustrates a weakness of using air temperature as a direct indicator of FT as it can both ignore the deeper surface state in addition to P-MW RS methods not being directly sensitive to air temperature. Reductions in the range of PDF temperature values is also prevalent as soil temperature is less susceptible to large temperature swings due to thermal inertia and increased insulation of the soil (250–305 K). Finally, while $T_{N-\text{skin}}$ displays symmetric characteristics, the overlap region is shown to extend far above and below the freezing point with increased uncertainty (258–279 K) (Table V). Higher temperature variability (225–320 K) also highlights deficiencies in using NLDAS skin temperature to accurately and precisely estimate ground FT state.

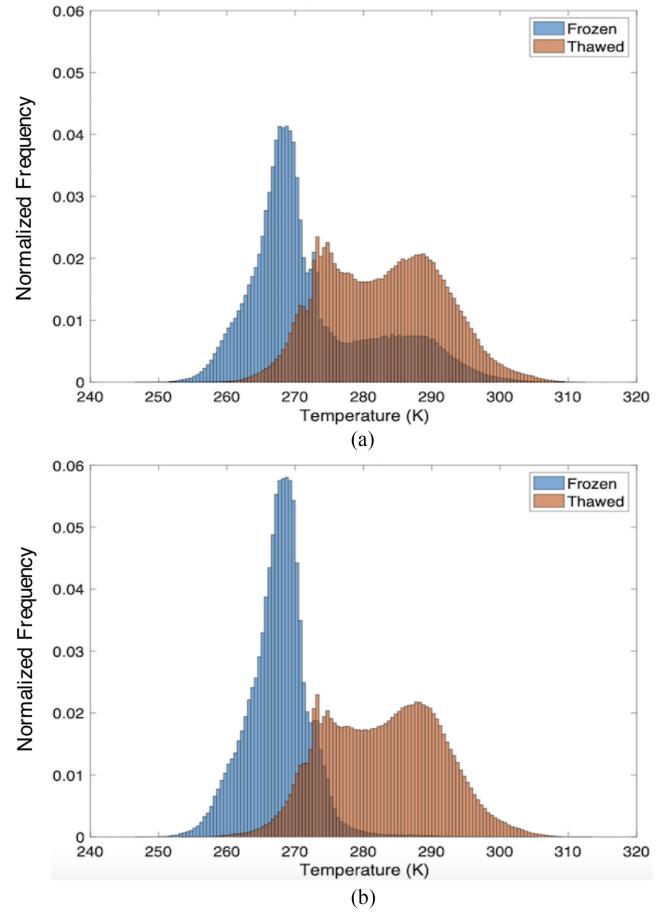


Fig. 10. PDF histograms of SMAP-FT (a) and SSMI-FT (b) defined frozen and thawed states matched to corresponding NLDAS 0–10-cm layer soil temperatures; frozen distribution to the left, overlap area shown in intermediate shade.

3) False Freeze Classifications: While FT-ESDR derived PDF histograms are largely symmetric in overlap region (see Fig. 8), asymmetric characteristics are prevalent in SMAP FTSP-based temperature distributions. This illustrates the prevalence of classifications in which higher temperatures are defined as frozen when observed with SMAP, while low temperature thaw classifications are less variable across products (see Table V). For example, a well-defined spike in frozen classifications occurs around 290 K when comparing SMAP-FT to NLDAS skin and soil temperatures [see Figs. 9(a), 10(a)]. These spikes are nonexistent when examining corresponding FT-ESDR distributions and occur at similar temperatures across all SMAP PDFs. This is reflected by the higher temperatures at which 1% of classifications remain frozen above across SMAP comparisons (296.0–307.5 K) as compared to FT-ESDR records (279.3–281.0 K) (see Table V). This difference in excess of 20 K is not prevalent when examining the low end of temperatures at which 1% of classifications occur below as thawed. These range from 255.3–266 K (SMAP) and 257.8–267.9 K (FT-ESDR), exhibiting a difference on the order of 2 K. Errors are expected to occur when high temperatures are associated with frozen classifications due to topographic effects, temporal or spatial resolution mismatch, temperature uncertainties, and FT algorithm

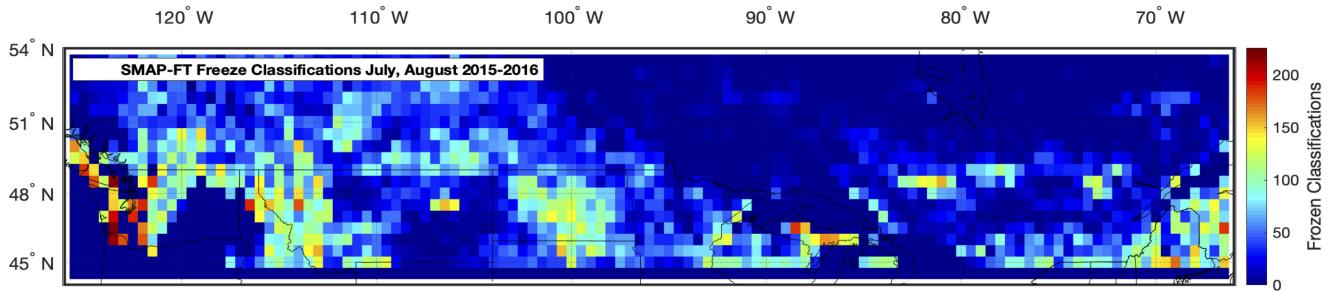


Fig. 11. SMAP-FT summer freeze classifications, July–August 2015 and 2016.

limitations. However, the excessive occurrence of these errors in SMAP-temperature associated PDFs can most likely be linked to false freeze classifications by the SMAP-FT algorithm.

To investigate this, the location and time of occurrence in which high temperature freeze classifications are examined. In July and August, characterized by minimum temperatures well above the freezing point across the vast majority of the domain, we identify regions in which freeze classifications were likely false (see Fig. 11). These regions correspond to areas with poor agreement both in CA (see Fig. 3) and in which D_{FTp} values were positive (see Fig. 4), indicating that SMAP FTSPs define an increased frozen extent as compared to temperature datasets. It is likely these false summer freeze identifiers significantly reduced CA and resulted in positive D_{FTp} values during the summers of 2015 and 2016 for SMAP comparisons (see Fig. 5). Outside of high-elevation regions with persistent snow cover (many classifications west of 110°W), areas indicating frequent frozen classifications, including areas in the Upper Midwest, Maine, Central and Eastern Canada, are almost certainly SMAP-FT false freeze classifications. It is likely that dry conditions and the relatively dense presence of surface water in some of these regions result in similar FT reference states [8], [22]. FTSPs in such regions will need to be improved moving forward to correct for false freeze classifications, especially in the mid-latitudes. Algorithm updates have been made in the recent SMAP Version 2 FT products to improve upon this and extend to global product coverage [47], [48].

IV. LIMITATIONS AND STUDY RELEVANCE

There are a number of assumptions and limitations that must be kept in mind when comparing FT products to NLDAS and MODIS temperature variables. Current FT products have shown an improved agreement as compared to *in situ* temperatures over more homogenous northern regions than in mountainous regions or southern regions with frequent FT transitions. The presence of different vegetation types and frequent large precipitation events has also been shown to reduce FT product reliability [25]. Various landscape characteristics also contribute to erroneous FT classifications derived from both SMAP and ESDR especially in regions with small differences between frozen and thawed reference states [8], [22]. This can be due to high densities of water bodies or over dry regions in certain regions as substantiated by this study.

Prior studies have also identified differences in P-MW derived FT product accuracy between A.M. and P.M. overpasses, especially when compared to *in situ* air temperature [8], [9]. This is likely due to ephemeral or diurnal FT transitions that are not captured by soil temperatures as a result of increased thermal inertia and snowpack insulation in the A.M. period. However, P.M. overpasses tend to capture surface state during significantly warmer air temperatures, which may lead to improved afternoon agreement during transitional periods. The presence of wet snow can lead to apparent thawed Tb emissions in SMAP and ESDR that match more closely to air temperature in the afternoon. It is important to note that radiometer-based FT products tend to retrieve less frozen area as compared to active products such as SMAP and Aquarius radars [27]. While active FT RS methods are not assessed in this paper, comparisons have been performed over Alaska utilizing L-, C-, and Ku-band radars showing a similar ability to detect changes in the surface FT state [49]. Similar to P-MW approaches, variability has been identified in backscatter due to changes in vegetation, moisture content, and topography.

Additional variations in P-MW retrieval times across FTSPs will inherently lead to differences in FT classifications as early afternoon thaw emissions vary in comparison to those in the early evening. Also, mid-night time retrievals (1:30 A.M.) tend to give more thawed classifications as compared to early morning (~6:00 A.M.) retrievals due to a reduced diurnal freezing period. The methodology outlined in Section II attempts to match corresponding NLDAS temperature values to various product overpass times, however, there are certainly errors induced by slight mismatches in P-MW retrieval and model times. This is especially true for the comparison of MODIS (1:30 A.M./P.M. LST) products to FT classifications derived from SMAP and SSMI (6 and 6:30 A.M./P.M. LST). Sparse coverage due to cloud cover of MODIS Aqua skin temperature retrievals also increases comparison uncertainty by reducing evaluation data points. Furthermore, as FT and temperature products are matched to a lower resolution grid, distortions of FT classifications due to classification rounding and temperature averaging can induce further error. This use of different resolution products can increase uncertainty especially in mountainous regions, where temperatures can vary significantly within a single grid pixel.

While limited, this methodology has the potential to be improved through the removal of clearly false classifications,

examination of an extended spatial and temporal record, the utilization of additional temperature datasets (GOES, *in situ* networks), the assessment of microwave backscatter observations, and precise time matching techniques over an extended spatial and temporal domain. Still, by utilizing regional NLDAS and MODIS temperature data, we are able to identify spatial dynamics across FT products that are difficult to examine through the use of limited point measurements and core validation sites. Moving forward, it will be critical to further assess SMAP *L*-band FT products as aspects of this analysis showed variations in the relationship to surface temperatures compared to *Ka*-band retrievals. As likely false freeze classifications are removed in the future, relationships of SMAP-FT to land surface variables should become clearer. The need to further investigate thermal profiles as they relate to regions with snow cover, precipitation, and varied landcover types will help to improve understanding of FT dynamics under various conditions as they relate to microwave RS techniques. Additionally, intergrid cell heterogeneity in FT state is often not well captured at the resolution at which these products are developed. The implementation of a fractional or higher-resolution FT products will likely improve the representation of FT state across the global domain, especially during freeze and thaw onset. While there are limitations to the methodology utilized in this paper, resulting CA, proportional differencing, and FT-temperature associated PDF analyses have proven insightful in assessing the relationship between P-MW-based FTSPs and surface temperature. These methods provide a comprehensive assessment of these relationships through space and time, resulting in the ability to compare AMSR and SSMI FTSPs, examine differences between SMAP 36 km and enhanced 9 km products, as well as between *K*- and *L*-band-based P-MW FT classifications across an extended domain. These results will also inform future efforts to improve satellite-based FTSPs by helping to identify regions in which accurate FT classifications have proven difficult.

V. CONCLUSION

In this paper, SMAP-, AMSR-, and SSMI-based P-MW FT products are compared to NLDAS air, skin, soil, and MODIS skin temperatures, respectively. These comparisons were performed with the intent of better understanding P-MW FT classifications as they relate to temperature and yielded extensive results regarding regional and temporal FT trends. As anticipated, results show that frozen definition changes with varied sensor configurations and retrieval timings as these products define different frozen extents especially in transitional periods. It is important to keep these differences in mind as changes in surface moisture states and observed emission depth cause differences between FT products and associated classifications. Through CA, air temperature (T_{2m-air}) is shown as the best proxy for all satellite FTSPs included in this paper. With CA values over the study period (April 2015–December 2017) of 91% (AMSR-FT, 25 km), 90% (SSMI-FT, 25 km), 84% (SMAP_E – FT, 9 km), and 81% (SMAP-FT, 36 km) when

using a threshold of 0 °C. Across all comparisons, skin temperature (MODIS and NLDAS) exhibit the lowest CA values compared to FTSPs ranging from 84.1%–86.4% (FT-ESDR) to 76.0%–78.6% (SMAP). Additionally, SMAP FTSPs show a stronger relative relationship with $T_{soil0-10cm}$ than FT-ESDR products. This result supports the enhanced sensitivity of *L*-band measurements to subsurface soil FT state. Regionally, variability in classification was observed across comparison members, most notably between SMAP (*L*-band) and FT-ESDR (*Ka*-band) products. Improvements in multiple comparison metrics, including CA, are also observed when comparing temperature products to the higher resolution SMAP_E – FT (84%) product to SMAP-FT (81%). This reflects an improvement in resolving smaller scale features in the enhanced product.

While results suggest P-MW derived FT states are most closely represented by air temperature, this relationship is proven stronger during spring thaw as compared to freeze onset for FT-ESDR defined states. The relationship is less clear in regard to SMAP-FT product comparisons. Overall, FTSPs are shown to frequently define less frozen area as compared to threshold-based temperature products when employing a proportional comparison technique (D_{FTP}) ranging from 0.7% to –14.1% when averaged across the period. Decreases in absolute proportional difference ($D_{FTP-abs}$) are observed from descending (A.M.) to ascending (P.M.) overpass classifications indicating improved temperature-based estimation during the afternoon period for SMAP and AMSR FTSPs. This improvement is less defined when examining differences in soil temperature as it is less susceptible to signaling frozen conditions during ephemeral FT events. The improved agreement between ESDR FTSPs and temperature datasets is less pronounced A.M. to P.M. as compared to SMAP FTSPs, even having the opposite relationship when comparing SSMI-FT to NLDAS air, soil, and skin temperatures.

When examining temperature values as associated with FT classifications, FT-ESDR-based products exhibit more clearly defined FT distributions. The observed overlap of FT distributions ranged from 9.0% to 11.1% (FT-ESDR) as compared to 23.0%–25.7% (SMAP-FT) and 20.3%–23.1% (SMAP_E – FT). NLDAS air temperature defined the most unique FT states when utilizing AMSR and SSMI distributions (9.0% and 9.9%), while SMAP-based distributions were shown to be best distinguished by soil temperature (20.3% and 23.0%). These results were verified through a Two-sample Kolmogorov–Smirnov Test and through an examination of PDF characteristics. Due to the prevalence of potential false freeze classifications in SMAP FTSPs, we believe these metrics should be revisited when freeze classifications with associated temperature values >300 K can be resolved. In conclusion, observed FT dynamics are proven to be dependent on sensor platform (i.e., band, spatial, and temporal resolution), season, and region as reflected across comparison members. The ability to improve our understanding of complex FT processes through the combination of satellite observations, models, and *in situ* measurements appears promising.

APPENDIX A

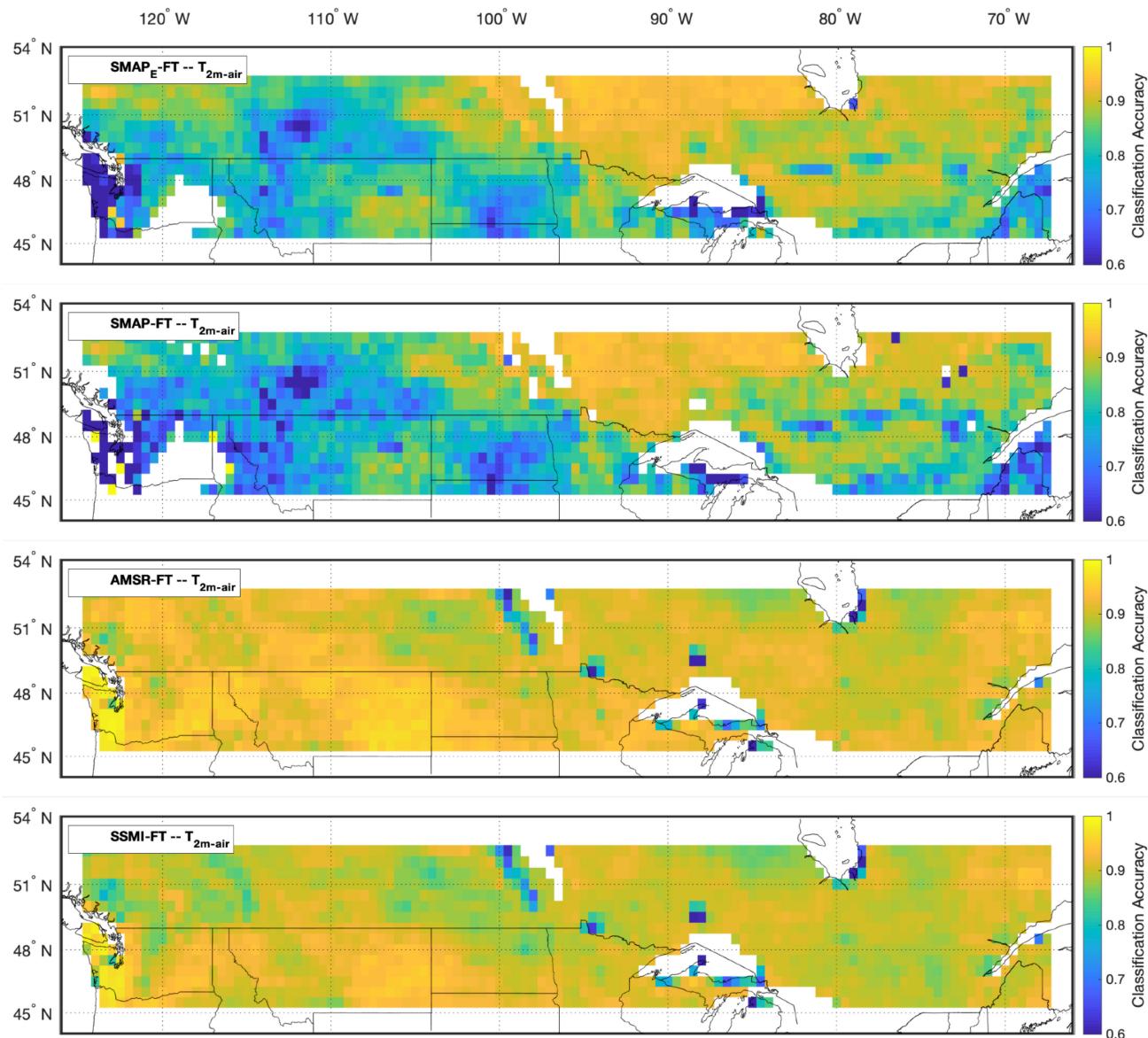


Fig. 12. Grid wise CA of FT products to NLDAS-2 2-meter air temperature.

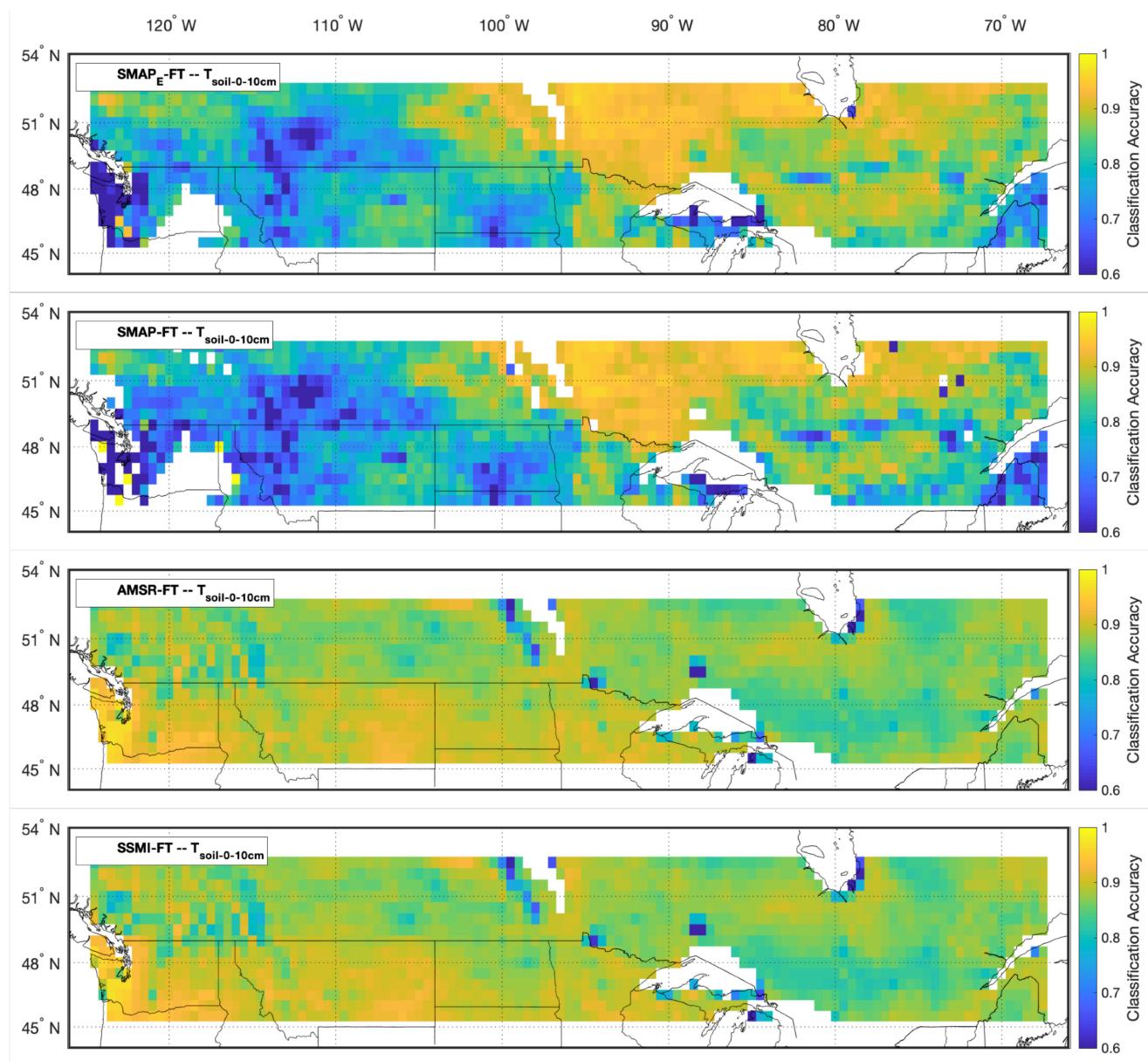


Fig. 13. Grid wise CA of FT products to NLDAS-2 0-10-cm soil layer temperature.

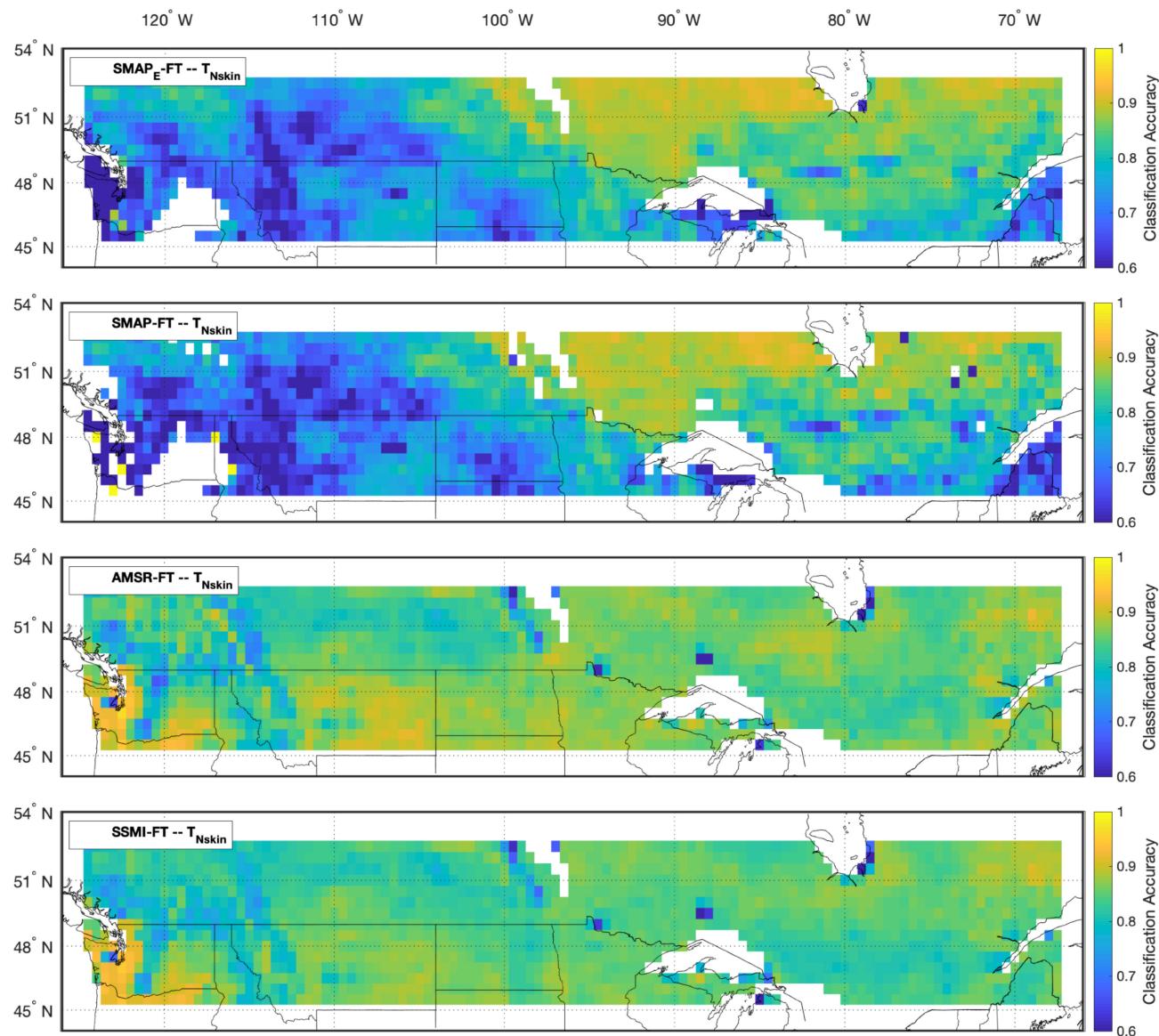


Fig. 14. Grid wise CA of FT products to NLDAS-2 skin temperature.

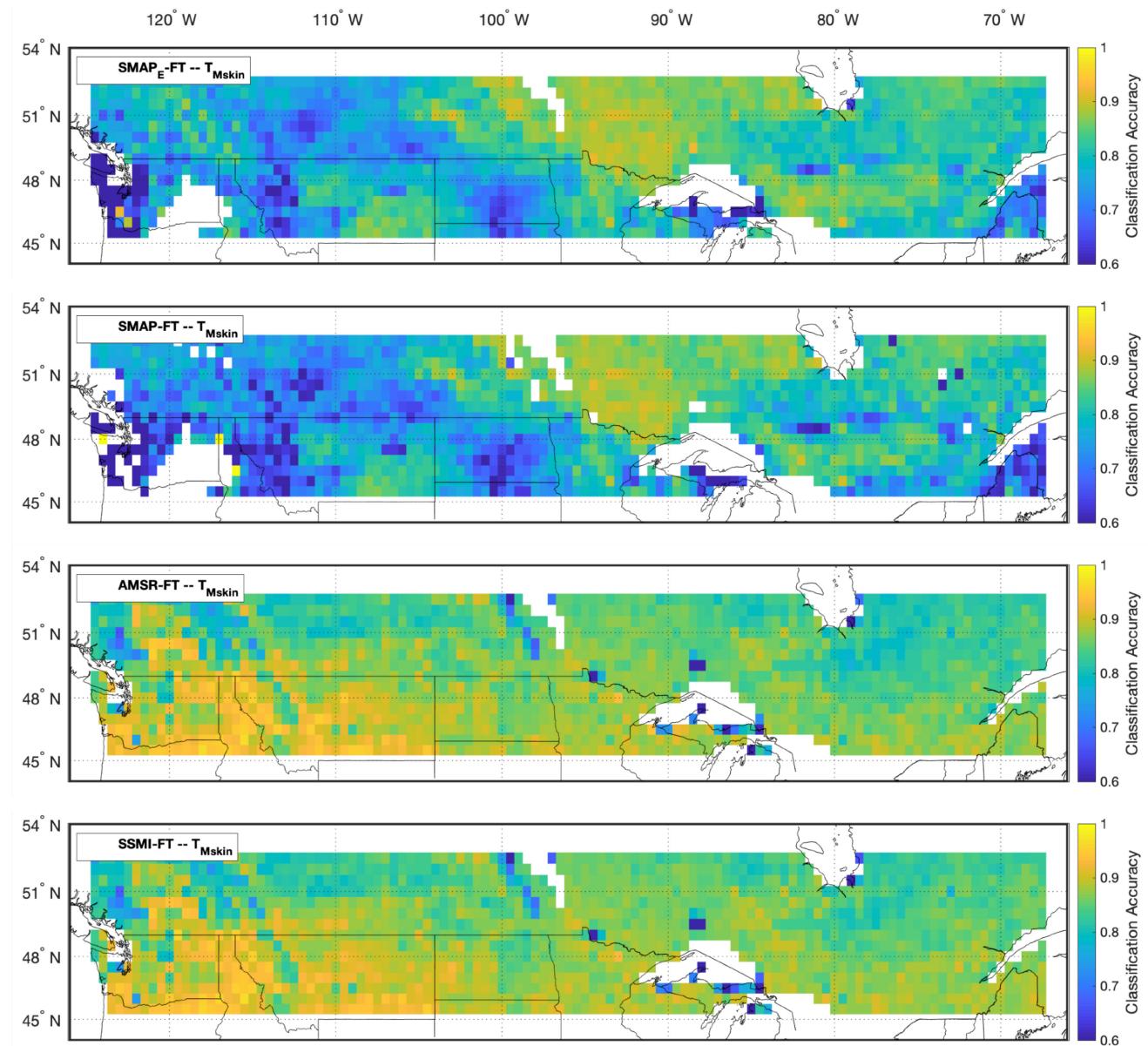


Fig. 15. Grid wise CA of FT products to MODIS-Aqua skin temperature.

APPENDIX B

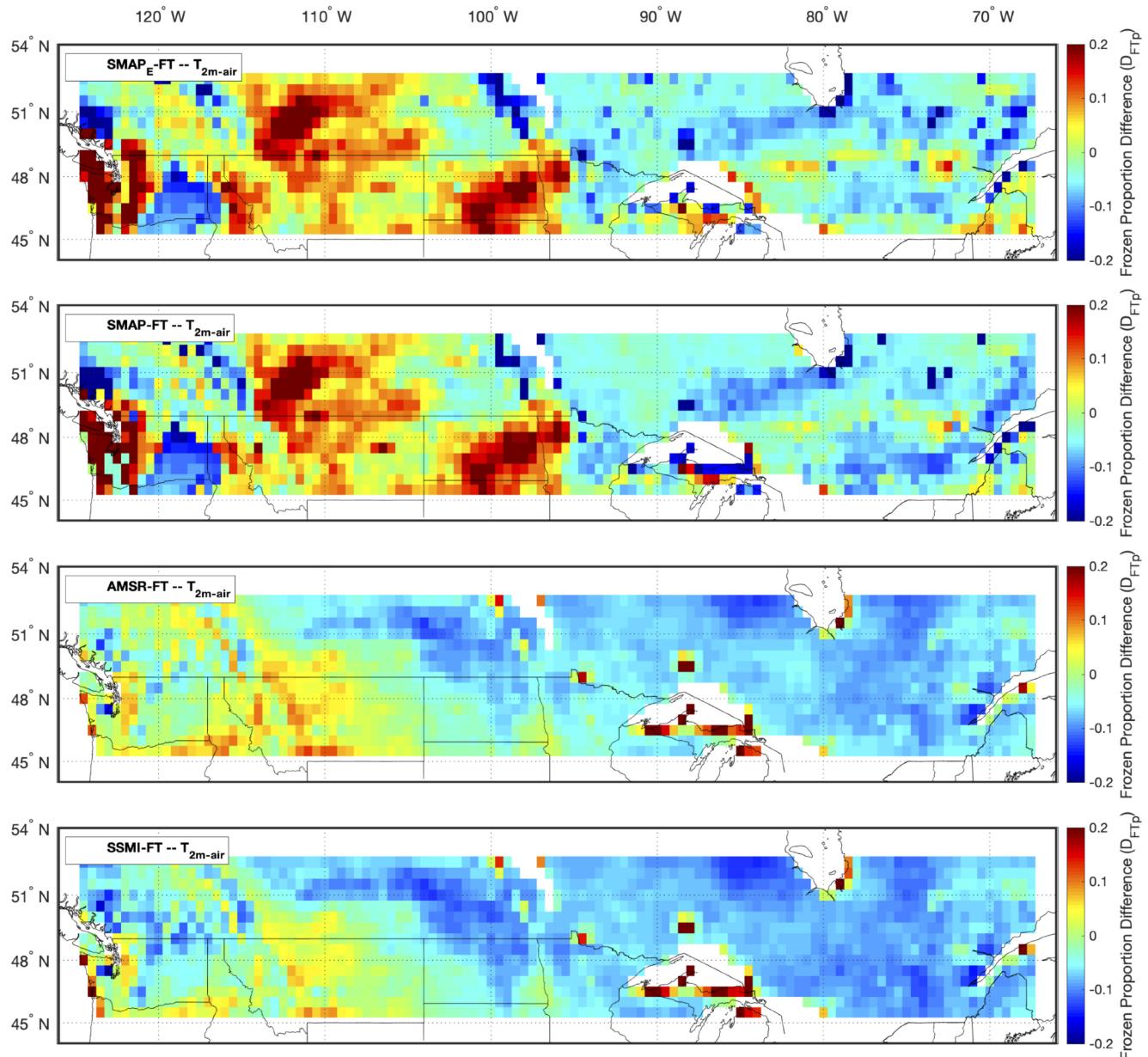


Fig. 16. Grid wise frozen proportion difference maps (DFTp), FT product minus NLDAS-2 2-meter air temperature-based frozen proportion.

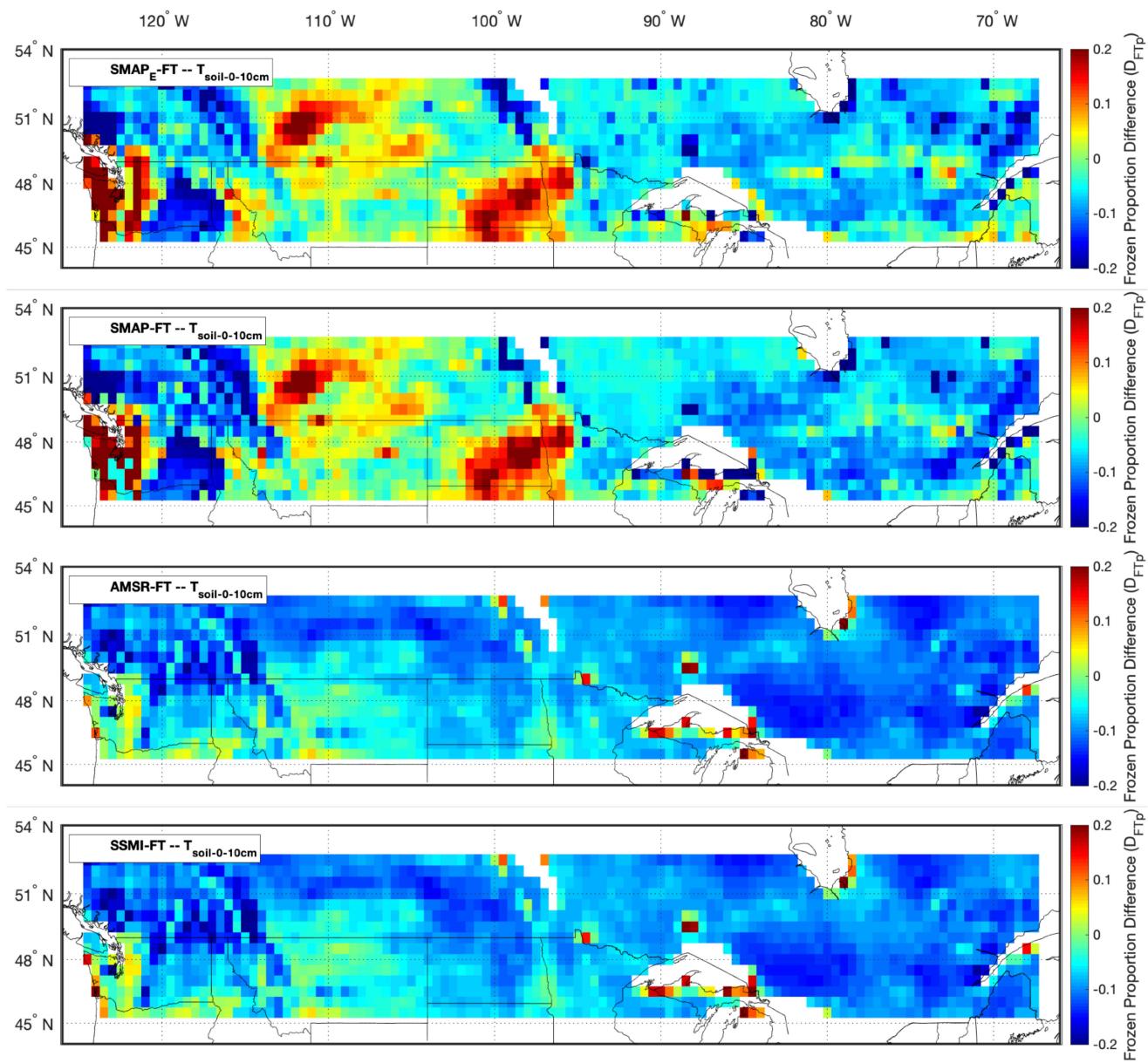


Fig. 17. Grid wise frozen proportion difference maps (D_{FTp}), FT product minus NLDAS-2 0-10-cm soil temperature-based frozen proportion.

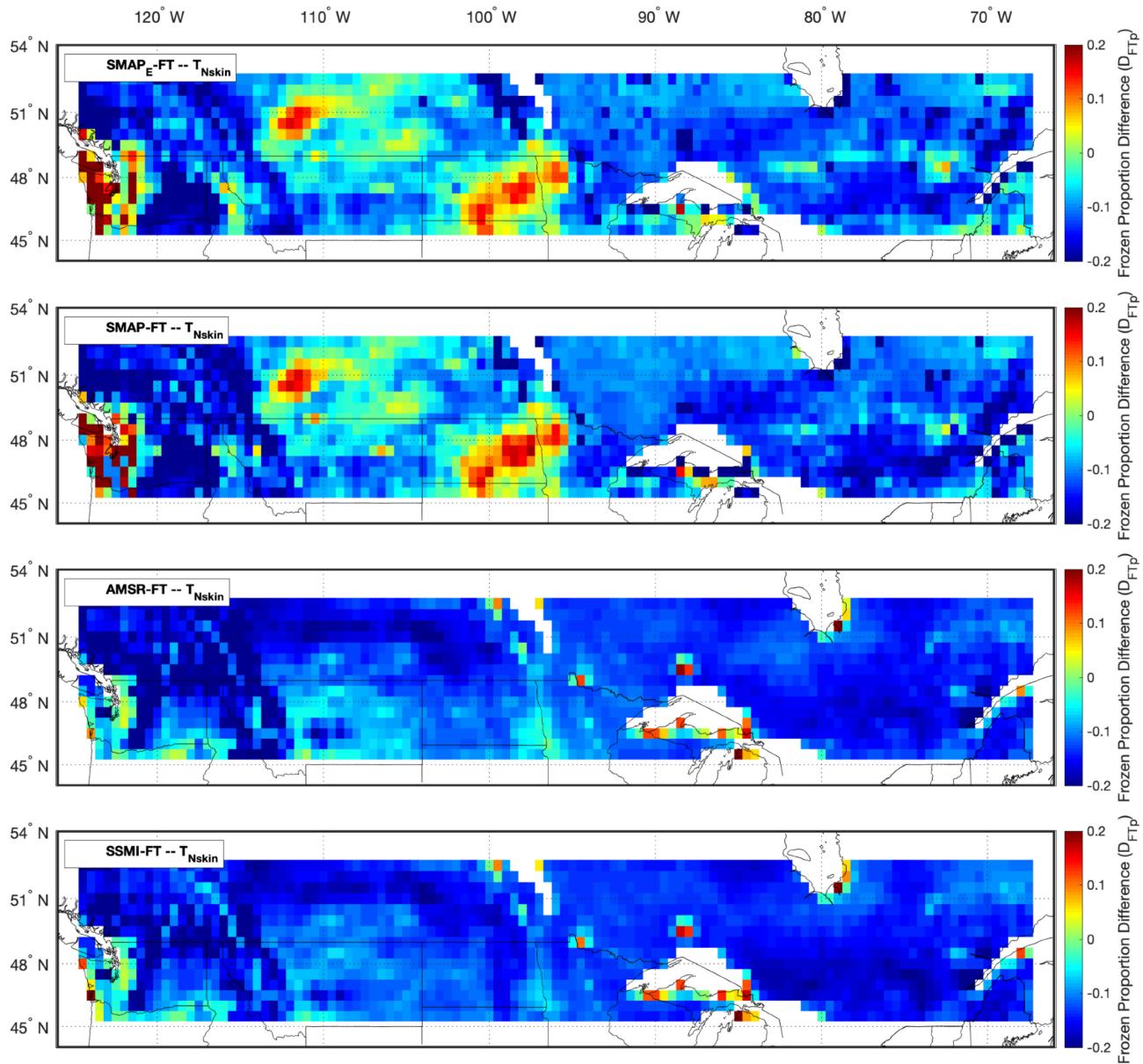


Fig. 18. Grid wise frozen proportion difference maps (D_{FTp}), FT product minus NLDAS-2 skin temperature-based frozen proportion.

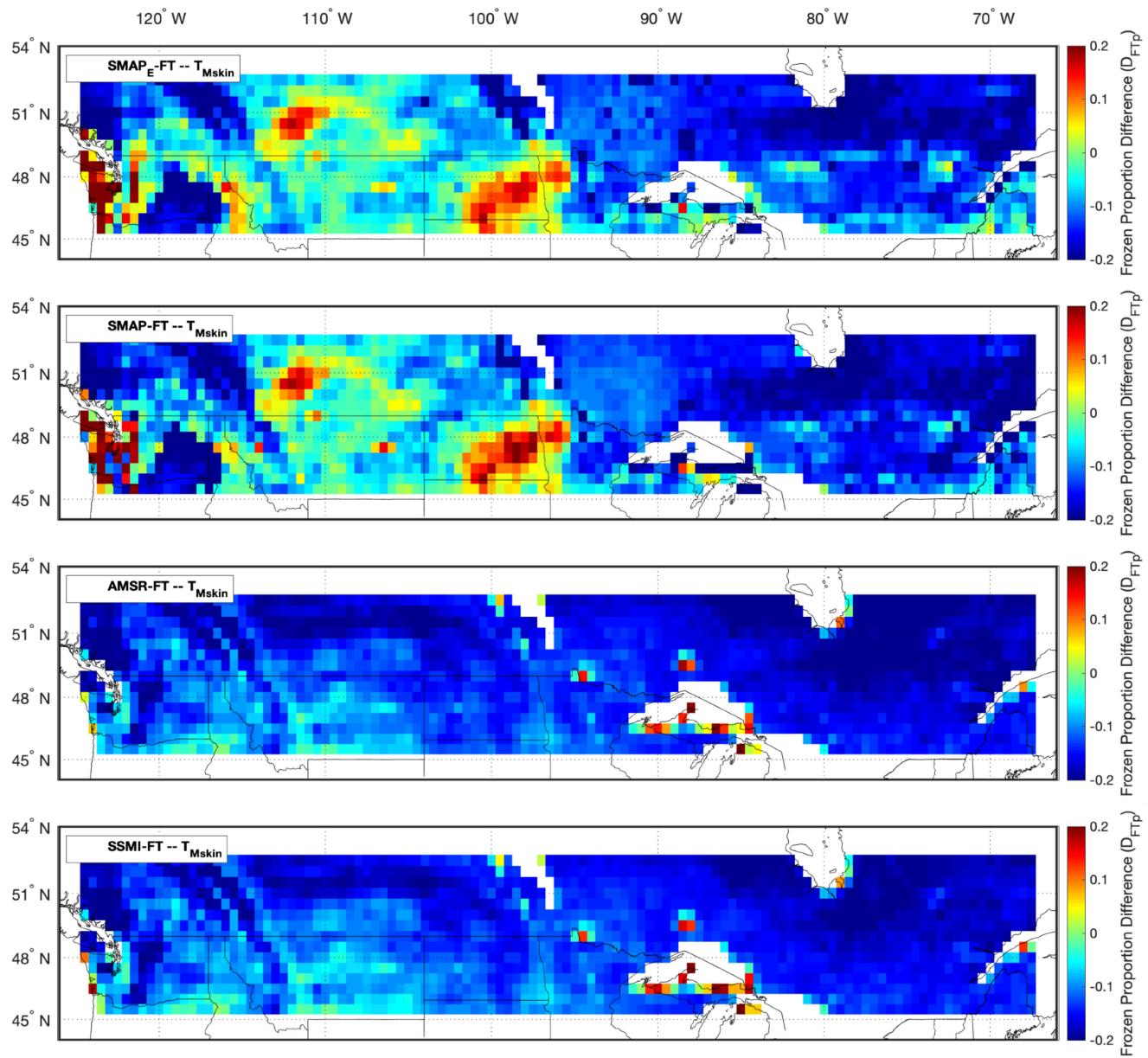


Fig. 19. Grid wise frozen proportion difference maps (D_{FTP}), FT product minus MODIS-Aqua skin temperature-based frozen proportion.

REFERENCES

- [1] J. S. Kimball, K. C. McDonald, S. W. Running, and S. E. Frolking, "Satellite radar remote sensing of seasonal growing seasons for boreal and subalpine evergreen forests," *Remote Sens. Environ.*, vol. 90, no. 2, pp. 243–258, Mar. 2004.
- [2] S. Ullah, R. Frasier, L. Pelletier, and T. R. Moore, "Greenhouse gas fluxes from boreal forest soils during the snow-free period in Quebec, Canada," *Can. J. Forest Res.*, vol. 39, no. 3, pp. 666–680, Mar. 2009.
- [3] J. Kreyling, C. Beierkuhnlein, K. Pritsch, M. Schloter, and A. Jentsch, "Recurrent soil freeze-thaw cycles enhance grassland productivity," *New Phytol.*, vol. 177, no. 4, pp. 938–945, 2008.
- [4] X. Zhang and S. Sun, "The impact of soil freezing/thawing processes on water and energy balances," *Adv. Atmos. Sci.*, vol. 28, no. 1, pp. 169–177, Jan. 2011.
- [5] B. Teufel *et al.*, "Investigation of the 2013 Alberta flood from weather and climate perspectives," *Climate Dyn.*, vol. 48, no. 9–10, pp. 2881–2899, May 2017.
- [6] A. Kalra, T. Piechota, R. Davies, and G. Tootle, "Changes in U.S. streamflow and western U.S. snowpack," *J. Hydrol. Eng.*, vol. 13, Mar. 2008, Art. no. 156.
- [7] T. Zhang and R. L. Armstrong, "Soil freeze/thaw cycles over snow-free land detected by passive microwave remote sensing," *Geophys. Res. Lett.*, vol. 28, pp. 763–766, 2001.
- [8] Y. Kim, J. S. Kimball, K. C. McDonald, and J. Glassy, "Developing a global data record of daily landscape freeze/thaw status using satellite passive microwave remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 949–960, Mar. 2011.
- [9] H. Lyu *et al.*, "Validation of the SMAP freeze/thaw product using categorical triple collocation," *Remote Sens. Environ.*, vol. 205, pp. 329–337, Feb. 2018.
- [10] Y. Kim, J. S. Kimball, D. A. Robinson, and C. Derksen, "New satellite climate data records indicate strong coupling between recent frozen season changes and snow cover over high northern latitudes," *Environ. Res. Lett.*, vol. 10, no. 8, 2015, Art. no. 084004.
- [11] L. Wang, P. Toose, R. Brown, and C. Derksen, "Frequency and distribution of winter melt events from passive microwave satellite data in the pan-Arctic, 1988–2013," *Cryosphere*, vol. 10, pp. 2589–2602, 2016.
- [12] D. Entekhabi *et al.*, *SMAP Handbook—Soil Moisture Active Passive: Mapping Soil Moisture and Freeze/Thaw from Space*. Washington, DC, USA: NASA, 2014.
- [13] K. Rautiainen *et al.*, "Detection of soil freezing from L-band passive microwave observations," *Remote Sens. Environ.*, vol. 147, pp. 206–218, May 2014.
- [14] T. Kawanishi *et al.*, "The advanced microwave scanning radiometer for the earth observing system (AMSR-E), NASA's contribution to the EOS for global energy and water cycle studies," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 2, pp. 184–194, Feb. 2003.
- [15] T. Zhao, L. Zhang, L. Jiang, S. Zhao, L. Chai, and R. Jin, "A new soil freeze/thaw discriminant algorithm using AMSR-E passive microwave imagery," *Hydrol. Process.*, vol. 25, no. 11, pp. 1704–1716, May 2011.
- [16] N. C. Grody, "Classification of snow cover and precipitation using the special sensor microwave imager," *J. Geophys. Res., Atmos.*, vol. 96, no. D4, pp. 7423–7435, 1991.
- [17] Y. Kim, J. S. Kimball, J. Glassy, and J. Du, "An extended global earth system data record on daily landscape freeze-thaw status determined from satellite passive microwave remote sensing," *Earth Syst. Sci. Data*, vol. 9, no. 1, pp. 133–147, Feb. 2017.
- [18] K. Rautiainen *et al.*, "L-band radiometer observations of soil processes in boreal and subarctic environments," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1483–1497, May 2012.
- [19] L. Zhang, T. Zhao, L. Jiang, and S. Zhao, "Estimate of phase transition water content in freeze–Thaw process using microwave radiometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4248–4255, Dec. 2010.
- [20] K. Rautiainen *et al.*, "SMOS prototype algorithm for detecting autumn soil freezing," *Remote Sens. Environ.*, vol. 180, pp. 346–360, Jul. 2016.
- [21] T. L. Rowlandson *et al.*, "Capturing agricultural soil freeze/thaw state through remote sensing and ground observations: A soil freeze/thaw validation campaign," *Remote Sens. Environ.*, vol. 211, pp. 59–70, Jun. 2018.
- [22] S. Dunbar *et al.*, "SMAP—algorithm theoretical basis document (ATBD)—level 3 radiometer freeze/thaw data products," California Inst. Technol., Pasadena, CA, Rep. JPL D-56288, 2016.
- [23] X. Xu, R. S. Dunbar, C. Derksen, A. Colliander, Y. Kim, and J. S. Kimball, "SMAP L3 Radiometer Global and Northern Hemisphere Daily 36 km EASE-Grid Freeze/Thaw State, Version 1," NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, CO, USA, 2016. [Online]. Available: <https://doi.org/10.5067/YN94K53QM061>, Accessed on: July 16, 2018.
- [24] X. Xu, R. S. Dunbar, C. Derksen, A. Colliander, Y. Kim, and J. S. Kimball, "SMAP Enhanced L3 Radiometer Global and Northern Hemisphere Daily 9 km EASE-Grid Freeze/Thaw State, Version 1," NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, CO, USA, 2016. [Online]. Available: <https://doi.org/10.5067/1M7OJC7R7VKI>, Accessed on: July 16, 2018.
- [25] J. S. Kimball, K. C. McDonald, Y. Kim, and J. Glassy, "An earth system data record for land surface freeze/thaw state. Algorithm theoretical basis document (ATBD), version 1," Numerical Terradynamic Simul. Group Publ., Missoula, MT, USA, Rep. 7-2009, Jul. 2009.
- [26] Y. Kim, J. S. Kimball, J. Glassy, and K. C. McDonald, "Measures global record of daily landscape freeze/thaw status, version 4," NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, CO, USA, 2017. [Online]. Available: <https://doi.org/10.5067/MEASURES/CRYOSPHERE/nsidc-0477.004>, Accessed on: Jul. 16, 2018.
- [27] C. Derksen *et al.*, "Retrieving landscape freeze/thaw state from Soil Moisture Active Passive (SMAP) radar and radiometer measurements," *Remote Sens. Environ.*, vol. 194, no. Suppl. C, pp. 48–62, Jun. 2017.
- [28] S. Kraatz *et al.*, "Evaluation of SMAP freeze/thaw retrieval accuracy at core validation sites in the contiguous united states," *Remote Sens.*, vol. 10, no. 9, Sep. 2018, Art. no. 1483.
- [29] Y. Xia *et al.*, "NLDAS Noah Land Surface Model L4 Hourly 0.125 × 0.125 degree V002," Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA, Rep. NASA/GSFC/HSL, 2012. [Online]. Available: <https://doi.org/10.5067/47Z13FNQODKV>, Accessed on: May 15, 2018.
- [30] Y. Xia *et al.*, "NLDAS Primary Forcing Data L4 Hourly 0.125 × 0.125 degree V002," Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA, Rep. NASA/GSFC/HSL, 2009. [Online]. Available: <https://doi.org/10.5067/615LHHOHZHN4>, Accessed on: May 15, 2018.
- [31] Z. Wan, S. Hook, and G. Hulley, "MYD11C1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 0.05Deg CMG V006," NASA EOSDIS LP DAAC, Washington, DC, USA, 2015. [Online]. Available: <https://doi.org/10.5067/MODIS/MYD11C1.006>, Accessed on: May 15, 2018.
- [32] B. A. Cosgrove *et al.*, "Real-time and retrospective forcing in the north american land data assimilation system (NLDAS) project," *J. Geophys. Res., Atmos.*, vol. 108, no. D22, 2003, Art. no. 8842.
- [33] M. Rodell *et al.*, "The global land data assimilation system," *Bull. Amer. Meteor. Soc.*, vol. 85, no. 3, pp. 381–394, Mar. 2004.
- [34] M. J. Suarez *et al.*, "Documentation and validation of the goddard earth observing system (GEOS) data assimilation system, version 4," NASA Technical Report Series on Global Modeling and Data Assimilation, Rep. NASA/TM-2005-104606, Vol. 26, Apr. 2005.
- [35] Y. Xia *et al.*, "Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): I. Intercomparison and application of model products," *J. Geophys. Res., Atmos.*, vol. 117, no. D3, pp. 1–27, Feb. 2012.
- [36] Y. Xia *et al.*, "Validation of noah-simulated soil temperature in the north american land data assimilation system phase 2," *J. Appl. Meteorol. Climatol.*, vol. 52, no. 2, pp. 455–471, 2013.
- [37] F. Mesinger *et al.*, "North american regional reanalysis," *Bull. Amer. Meteor. Soc.*, vol. 87, no. 3, pp. 343–360, Mar. 2006.
- [38] L. Luo *et al.*, "Validation of the north american land data assimilation system (NLDAS) retrospective forcing over the southern Great Plains," *J. Geophys. Res., Atmos.*, vol. 108, no. D22, 2018, Art. no. 8843.
- [39] L. Chai, L. Zhang, Y. Zhang, Z. Hao, L. Jiang, and S. Zhao, "Comparison of the classification accuracy of three soil freeze–thaw discrimination algorithms in China using SSMIS and AMSR-E passive microwave imagery," *Int. J. Remote Sens.*, vol. 35, no. 22, pp. 7631–7649, Nov. 2014.
- [40] R. B. Campbell, "The freezing point of water in puddled and unpuddled soils at different soil moisture tension values," All Graduate Theses and Dissertations, Utah State University, Logan, UT, USA, 1951.

- [41] S. Suzuki, "Verification of freezing point depression method for measuring matric potential of soil water," *Soil Sci. Plant Nutrition*, vol. 50, no. 8, pp. 1277–1280, Feb. 2004.
- [42] F. J. Massey, "The Kolmogorov-Smirnov test for goodness of fit," *J. Amer. Statist. Assoc.*, vol. 46, no. 253, pp. 68–78, Mar. 1951.
- [43] L. H. Miller, "Table of percentage points of Kolmogorov statistics," *J. Amer. Statist. Assoc.*, vol. 51, no. 273, pp. 111–121, Mar. 1956.
- [44] NCEI, "December 2015 Temperature Anomaly—Climate at a Glance," National Centers for Environmental Information (NCEI), Asheville, NC, USA, 2019. [Online]. Available: <https://www.ncdc.noaa.gov/cag/divisional/mapping/110/tavg/201512/1/anomaly>. Accessed on: May 24, 2019.
- [45] NCEI, "December 2016 temperature anomaly—Climate at a Glance," National Centers for Environmental Information (NCEI), Asheville, NC, USA, 2019. [Online]. Available: <https://www.ncdc.noaa.gov/cag/divisional/mapping/110/tavg/201612/1/anomaly>. Accessed on: May 24, 2019.
- [46] T. Zhao *et al.*, "Estimation of high-resolution near-surface freeze/thaw state by the integration of microwave and thermal infrared remote sensing data on the Tibetan Plateau," *Earth Space Sci.*, vol. 4, no. 8, Aug. 2017, Art. no. 2017EA000277.
- [47] S. Dunbar, X. Xu, A. Colliander, C. Derksen, J. S. Kimball, and Y. Kim, "SMAP - Algorithm theoretical basis document (ATBD) - level 3 radiometer freeze/thaw data products version B," California Inst. Technol., Pasadena, CA, USA, Rep. JPL D-56288, 2018.
- [48] X. Xu, R. S. Dunbar, A. Colliander, Y. Kim, J. Kimball, and C. Derksen, "Soil moisture active passive (SMAP) project calibration and validation for the L3_FT_P and L3_FT_P_E data product (version 2)," Jet Propulsion Laboratory, Pasadena, CA, USA, Rep. JPL D-102646, 2018.
- [49] E. Podest, K. C. McDonald, and J. S. Kimball, "Multisensor microwave sensitivity to freeze/thaw dynamics across a complex boreal landscape," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6818–6828, Nov. 2014.



Jeremy Johnston (S'19) received the B.S. and M.S. degrees in 2017 and 2019, respectively, from George Mason University, Fairfax, VA, USA, where he is currently working toward the Ph.D. degree in environmental and water resources engineering with the Department of Civil, Environmental and Infrastructure Engineering.

He is currently a Graduate Research Assistant with the George Mason University. He hopes to apply remote sensing data to aid in global water resource management and to utilize environmental monitoring data to support and develop more efficient human systems at the city scale. His research interests include the use of various remote sensing techniques (satellite, *in situ*, and UAV) to contribute to the understanding of global hydrologic and energy cycling as well as toward monitoring a changing climate.



Viviana Maggioni received the B.S. and M.S. degrees from the Politecnico of Milan, Milan, Italy, in 2003 and 2006, respectively, and the Ph.D. degree from the University of Connecticut, Storrs, CT, USA, in 2012, all in environmental engineering.

She is currently an Assistant Professor of Environmental and Water Resources Engineering with George Mason University, Fairfax, VA, USA. In particular, she is interested in the application of remote sensing techniques to estimate and monitor hydrological variables. Her work has direct applications in water resources management, weather and climate prediction, as well as agriculture and irrigation practices. Her research interests include the intersection of hydrology and remote sensing.

Dr. Maggioni is currently a Deputy Chair of the AGU Precipitation Committee and an Associate Editor of the *Journal of Hydrometeorology* and *Journal of Hydrology*.



Paul Houser received the B.S. and Ph.D. degrees in hydrology and water resources from the University of Arizona, Tucson, AZ, USA, in 1992 and 1996, respectively.

He is currently a Professor with the Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA, USA, where he is the Codirector of the Center for Energy Science and Policy. He is an expert in local to global land surface-atmospheric remote sensing, *in situ* observation and numerical simulation, development and application of data assimilation methods, and global water and energy cycling. His research interests include integrating energy and water research across traditional disciplines (e.g., nexuses) that transition theoretical research to academic/public education and real-world application.