Remote Sensing of Environment 237 (2020) 111545

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

# Improving SMAP freeze-thaw retrievals for pavements using effective soil temperature from GEOS-5: Evaluation against in situ road temperature data over the U.S

Simon Kraatz[a,*], Jennifer M. Jacobs[a,b], Ronny Schröder[a], Eunsang Cho[a], Heather J. Miller[c], Carrie M. Vuyovich[d]

[a] Department of Civil and Environmental Engineering, University of New Hampshire, Durham, NH 03824, USA
[b] Earth System Research Center, University of New Hampshire, Durham, NH 03824, USA
[c] Department of Civil and Environmental Engineering, University of Massachusetts Dartmouth, North Dartmouth, MA 02747, USA
[d] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

ARTICLE INFO

Edited by: Jing M. Chen

*Keywords:*
SMAP
Passive microwave
Freeze/thaw
Roads
RWIS

ABSTRACT

Seasonal freeze-thaw (FT) affects over half the northern hemisphere and impacts many key processes of the Earth System such as energy exchange, hydrology and vegetation. Nearly all past studies using spaceborne FT retrievals have focused on characterizing FT specifically for natural environments. FT in the built environment is also routinely studied and a topic of great interest, especially with regards to transportation infrastructure. Whereas natural FT process are frequently investigated using spaceborne observations, FT studies of roads are often limited to local scales, using in situ or nearby weather station data only. Comparisons between FT retrievals obtained from NASA's Soil Moisture Active Passive (SMAP) satellite and roads in Alaska (AK) and the Contiguous United States (CONUS) showed that spaceborne FT retrievals had good agreement with road data. But those results also indicated that NASA FT retrievals in CONUS were relatively too warm compared to road data. If SMAP FT retrievals were to be used for identifying FT transition timing for applications by the transportation community, it is also important for frozen conditions to be identified more accurately. This work is primarily concerned with improving frozen retrievals made in CONUS by calculating new Normalized Polarization Ratio (NPR) thresholds as compared to those currently used in SMAP FT. We found that focusing on a temporal subset of October through May for comparisons greatly improved the correlation between NPR and effective soil temperature ($T_{eff}$, one of SMAP's ancillary datasets), often from about zero to 0.6. We then applied linear regression between NPR and $T_{eff}$ to obtain new NPR thresholds resulting in the FT-Roads (FT-R) product. NASA FT and FT-R were evaluated against road data at about 1000 locations in CONUS and a battery of different tests indicated that FT-R performed better under nearly all conditions compared to NASA FT. Overall, NASA FT accuracies were 69% and 80% for 6 am and 6 pm SMAP retrievals, while FT-R achieved accuracies of 79% and 82%. We also investigated the potential for using $T_{eff}$ for road FT (6 am, only) and found that those comparisons were even more accurate (84%). We've also quantified inter- and intraregional differences of SMAP FT performance and found that accuracy metrics vary over twice as much between geographic subdivisions (9%) as compared to between the states within a subdivision (4%). Most importantly, the main goal of improving the detection of in situ frozen conditions in CONUS was realized, with FT-R accurately detecting frozen conditions > 50% more frequently than NASA FT.

## 1. Introduction

Seasonal freeze-thaw (FT) impacts more than half of the global land area and is a major control on natural processes (Kim et al., 2011; Zhang, 2003). Seasonal FT cycles are also very important for the built environment, especially with respect to roads. Roads are much more susceptible to damage during the spring thaw period than during other times of the year (Andersland and Ladanyi, 2004; Kestler et al., 2011; Simonsen and Isacsson, 1999). Department of Transportation (DOT) agencies often attempt to reduce damage to roads by placing load

* Corresponding author at: Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA, 01003, USA.
*E-mail address:* skraatz@umass.edu (S. Kraatz).

https://doi.org/10.1016/j.rse.2019.111545
Received 30 April 2019; Received in revised form 4 November 2019; Accepted 14 November 2019

restrictions during the spring thaw period (Mahoney et al., 1986; MNDOT, 2014; Van Deusen et al., 1998). However, it is extremely difficult to determine when thawing occurs over the vast network of roads. This makes operational decisions about when, where and for how long to implement load restrictions very challenging.

Owing to the large spatial scales involved and ability to frequently monitor FT states, spaceborne observations using passive microwave instruments have potential value for this task (Kraatz et al., 2017, 2019c, 2019b). Microwave instruments are highly sensitive to water and can be used to estimate soil moisture (SM) during non-frozen seasons or the physical state of water (Judge et al., 1997; Kalantari et al., 2014; Mätzler, 2006). L-band (~1 GHz) microwave observations are especially useful for retrieving soil properties because they correspond to a relatively greater emission depth in soils (~5 cm) (Entekhabi et al., 2014). L-band observations are also less impacted by the atmosphere, vegetation, and snow compared to observations made at higher frequencies. Therefore, retrievals of soil properties are generally more accurate when made at L-band compared to higher frequency bands (Kim et al., 2018).

Passive microwave observations from the Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS) platforms are especially useful for spaceborne FT retrievals. The SMAP mission is administered by the National Aeronautics and Space Administration (NASA) while SMOS is administered by the European Space Agency (ESA). SMAP has a relatively shorter period of record since it was launched in 2015 compared to 2009 for SMOS. But SMAP observations have a higher spatial resolution ($36^2$ km$^2$ vs. $50^2$ km$^2$), lower revisit time (2 days vs. 3 days) and are made at a constant incidence angle of 40° (Entekhabi et al., 2014; Kerr et al., 2010). NASA also provided a FT product (NASA FT) from launch (Xu et al., 2018), while a comparable product is still under development for SMOS. NASA FT contains two FT values per day (0 for thaw, 1 for frozen), corresponding to local equator crossing times of 6 am (AM) and 6 pm (PM). In the following, SMAP FT is used to refer to FT retrievals made using SMAP observations while NASA FT refers to the NASA FT products and its algorithms.

So far, relatively few assessments of SMAP FT retrieval accuracy have been made, and they have mainly focused on the densely instrumented SMAP Core Validation Sites (CVS). SMAP FT retrievals performed well for eight high-latitude CVSs (> 50°N) with accuracy usually exceeding 80% between 1 April 2015 to 7 July 2015 (Derksen et al., 2017). Kraatz et al. (2018) also evaluated SMAP FT retrievals, but at sites south of 45°N and using a slightly different retrieval algorithm because NASA FT originally did not extend south of 45°N. That study compared SMAP FT retrievals against in situ data at three CVSs (near ~40 °N) and indicated that SMAP FT retrievals may also have a good accuracy (> 70%) below 45 °N. A global scale evaluation of SMAP FT correspondence with air temperature data for annual temporal subsets of 2016 and 2017 showed that the $36^2$ km$^2$ product was about 4% more accurate than the $9^2$ km$^2$ product (Kim et al., 2019). For the coarser FT product, retrieval accuracies were 78% and 90% for AM and PM retrievals, respectively. The study also showed that SMAP FT accuracies were about 1 to 3% better south of 45°N compared to north of 45°N.

However, these and most other prior assessments of SMAP FT have been restricted to data collected in natural soils and have therefore not taken advantage of other available datasets pertaining to the built environment, such as road temperature (RT) data collected by Road Weather Information Systems (RWIS). A lack of such comparisons is not surprising, because it would not be expected that SMAP FT retrievals have any significant correspondence with the road surface FT state. One reason is that these comparisons are made based on two different means of inferring FT states: SMAP FT is based on changes in permittivity whereas road FT states are inferred from temperature readings. Another key concern is the substantial mismatch of spatial scales: SMAP retrievals are made on the order of $40^2$ km$^2$, whereas roadsdimensions are on the order of meters and even the areal coverage of all roads contained in the 39 km × 47 km SMAP footprint is negligible compared to other landcover classes. Also, SMAP is responding to decreases in real permittivity, which occurs when liquid water changes to ice. Even if the SMAP observations were influenced by the pavement's changing FT state, the resulting changes in permittivity would require additional interpretation in light of pavement permittivity having different responses to variations in density, moisture and salinity than a natural soil (Jaselskis et al., 2003). Additionally, roads have very different physical attributes compared to natural soils. Roads are impervious, darker in colour, treated for ice and snow and are engineered to transport water efficiently. Roads have lower albedo compared to surrounding landscapes, especially during winter when roads are routinely cleared of snow. While roads are in direct contact with the atmosphere, surface road temperatures may often be quite different from air and soil temperatures (Andersland and Ladanyi, 2004; Eftekhari et al., 2018; Kraatz et al., 2017, 2019a). Furthermore, comparisons between roads and natural sites showed that roads had much shorter periods of isothermal conditions during spring thaw, compared to natural areas (Kraatz et al., 2017). And unlike natural sites, roads also experience freezing point depression due to the application of road salts and frictional heating. Although there are many substantial differences between roads and natural sites, comparisons made between SMAP landscape FT states and roads may begin to provide insights that are also useful for the built environment.

Nonetheless, recent comparisons of NASA FT Version 1 retrievals against RTs in Alaskan roads indicated high correspondence for thaw timing (Kraatz et al., 2017, 2019c). NASA FT version 2 has also been evaluated against RT data at about 100 Road Weather Information System (RWIS) stations located in CONUS (Kraatz et al., 2019b). That study also indicated average accuracies of 65% and 76% for AM and PM FT retrievals, respectively, and was similar to the accuracies obtained for the few natural sites studied in CONUS (Kraatz et al., 2018). However, CONUS accuracy metrics were usually substantially lower compared to results obtained for Alaska. While somewhat lower accuracies are expected for locations further south, NASA FT retrievals for CONUS were relatively warm, meaning that in situ frozen conditions were frequently missed. Nearly 80% of all in situ frozen conditions were correctly detected in the Alaska study (Kraatz et al., 2019c), but only about 40% were detected in CONUS (Kraatz et al., 2019b). Johnston et al. (2019) also showed that SMAP FT retrievals currently are relatively too warm in CONUS from comparisons to GEOS-5 air, surface and 0–10 soil temperatures.

Inaccurate detections of frozen conditions can be a problem if SMAP FT were to be used e.g. an ancillary dataset to assist transportation decision making. For example, a statewide study for Minnesota (Kraatz et al., 2020) found that NASA FT retrieved most of the northern TDP sites as thawed for early to mid-January 2017 although TDP data showed temperatures below freezing to depths of over 90 cm. While this does not necessarily mean that the SMAP retrieval was incorrect (e.g. there was snowmelt or flooding) this causes issues for using SMAP FT to for example determine dates for applying SLRs.

Therefore, the goal of this work is to improve SMAP FT frozen state retrieval rates for roads in CONUS such that the performance approaches that found in the Alaskan roads study. The large difference in road frozen state detection rates between the two regions indicates that there may be opportunities for improving frozen detection rates in CONUS. For this purpose, we develop an algorithm for SMAP FT road retrievals (‘FT-R’). This work also extends previous comparisons between NASA FT and RWIS FT observations (Kraatz et al., 2019b) by about 900 stations across the CONUS and Alaska and now includes approximately 1000 RWIS sites. This expanded investigation will provide more comprehensive information regarding the spatial and temporal variations of SMAP FT accuracy compared to roads throughout CONUS.

The hypotheses for this study are that retrievals from FT-R compared to the extended road dataset, would usually: (1) exceed 70% accuracy; (2) perform better for PM than AM; (3) perform better for

Alaska than in CONUS; and (4) substantially improve the detection rate of in situ frozen conditions in CONUS over the NASA FT product.

## 2. Data and data processing

### 2.1. SMAP data (L3_SM_P, version 4 and 5; L3_FT_P, version 2)

The SMAP data record started on 31 March 2015 and provides global coverage every 2–3 days. The radiometer has an ellipsoidal instantaneous field of view of 38 km by 49 km. SMAP products are gridded on $36^2$ $km^2$ Equal Area Scalable Earth (EASE 2.0) grids for the standard products, and on $9^2$ $km^2$ grids for the enhanced products (Brodzik et al., 2014, 2012). This study only uses the standard $36^2$ $km^2$ resolution products, because it was shown to be more accurate than the $9^2$ $km^2$ product (Derksen et al., 2017; Kim et al., 2019). SMAP datasets used in this study are freely available from the National Snow and Ice Data Center[1] (NSIDC).

NASA FT Version 1 was originally produced only on the northern EASE grid (azimuthal equal-area projection), covering latitudes north of 45°N. The original product only used the Normalized Polarization Ratio (NPR) in the algorithm for delineating frozen from thawed conditions. In the following, this approach is referred to as FT-NPR. When SMAP does not have an overpass on a given day, FT values are filled according to the most recently determined FT state by FT-NPR. FT is represented with integer values of 0 and 1 for thaw and freeze, respectively. The version 2 product now also includes a FT product on a global EASE grid (Cylindrical Equal-Area Projection).[2] One important change is that the FT-NPR algorithm has been modified, and newly updated and calibrated Tb values are used (Chan et al., 2018).[3] Version 2 also now uses a second algorithm for FT detections – mostly at southern grids - that is based on thresholds developed from the vertically polarized component of the brightness temperature ($Tb_V$) and modelled surface temperatures, referred to as FT-SCV (Dunbar et al., 2018; Kim et al., 2017). Because FT-NPR is not computed at those grids that freeze fewer than 20 times, FT-SCV may be used instead of FT-NPR to determine the FT state. The SMAP FT product also employs additional error mitigation via a weekly AMSR-E climatology mask to avoid false freeze and thaw detections and additionally also sets the FT states to 'thawed' for when a Tb value exceeds 273 K (Dunbar et al., 2018). Furthermore, the NASA FT algorithm also requires that the difference between NPR reference values for thawed ($NPR_{th}$) and frozen conditions ($NPR_{fr}$) is positive and exceeds a certain magnitude ($> 0.1$) in order for FT-NPR to be computed (Dunbar et al., 2018). The SMAP FT product features a variable named 'retrieval_algorithm_flag', which is set to 0, 1 or 2 for no retrieval, FT-NPR and FT-SCV, respectively.

In this work, we only use data corresponding to those dates and times for which SMAP overpasses occurred. For comparisons against road data, we use the global NASA FT version 2 product (R16010), which consists of the combined results using the FT-NPR and FT-SCV algorithms. For developing FT-R, we use the same NPR values used by FT-NPR.

The SMAP soil moisture product mainly provides estimates of volumetric soil moisture in $m^3/m^3$, but also includes ancillary data that is needed to calculate SM values (O'Neill et al., 2016b, 2018b). $T_{eff}$, an estimate for the physical temperature of soils corresponding to L-band retrieval depths (Choudhury et al., 1982), provides ancillary data used to compute SM values. The general idea is that $T_{eff}$ can be estimated by combining the temperature values of different soil layers (e.g. 0 cm, 10 cm and/or 20 cm) using soil temperatures from the Goddard Earth Observing Model version 5 (GEOS-5) analysis (Chan et al., 2016; SMAP, 2015).

[1] SMAP data available from NSIDC at https://nsidc.org/

[2] SMAP FT Version 2 available at https://doi.org/10.5067/YN94K53QM061

[3] SMAP L1C brightness temperature V4 available at https://doi.org/10.5067/ZVILG0PS6CTI

$T_{eff}$ calculations were changed between L3_SM_P version 4 to version 5[4] to reduce a dry bias during non-frozen conditions for certain land cover types (O'Neill et al., 2018a). This change decreased the number of frozen days per year for large areas of the U.S. for version 5 as compared to version 4 (Fig. S1). For this study, $T_{eff}$ version 4 R15180 data (O'Neill et al., 2016a)[5] are used because their distribution of frozen counts more closely resembled observations at SMAP CVSs (Kraatz et al., 2018). Walker et al. (2019) also reported that $T_{eff}$ Version 5 temperatures exceeded physical measurements by more than 5 °C in their study at the Iowa SMAP CVS (Walker et al., 2019). Further details regarding the algorithm used to calculate $T_{eff}$ and how it evolved over time are provided in S1.

### 2.2. RWIS data

RWIS RTs are obtained from the Meteorological Assimilation Data Ingest System (MADIS).[6] In total, MADIS includes data records from about 3000 RWIS stations throughout the U.S., with records being archived since 2004. More than 33 state DOTs contribute their data. Each station may report data at different times (e.g. 5 minutes, hourly, daily) with most providing data at hourly or shorter time intervals. For comparisons with SMAP AM and PM observations, we aggregated MADIS data to hourly time intervals. A single station includes up to four road surface temperatures, four road subsurface temperatures and one air temperature observation at each RWIS site. Near surface road temperatures are measured within the road structure, most likely within the top few centimeters of the asphalt concrete layer (e.g. in the extensive Alaskan RT network,[7] surface temperatures are measured by a hockey puck shaped sensor with its top placed at 3 mm below the road surface).

Road surface temperatures may provide the most appropriate measures for comparison to SMAP FT, because SMAP FT retrievals had been shown to be highly sensitive to FT at the surface (Rowlandson et al., 2018; Zheng et al., 2019). While SMAP's FT response results from changes in permittivity when liquid water changes to ice, the RWIS temperature profiles provide a proxy measure of changing water state. The use of temperature to indicate water state change is supported by Kestler et al.'s (2007) study. Their study used coincident observations of unfrozen moisture content in roads as measured by subsurface instrumentation for temperature (thermistors) and moisture (time domain reflectometer and radio frequency moisture sensors) and pavement strength (portable or lightweight falling weight deflectometers) to show that subsurface temperature measured via thermistors provide a reasonable metric to identify the onset of thaw and indicated that SLRs can be placed when near-surface temperatures rise to 0 °C.

RWIS temperature data were pre-processed to ensure data quality (Supplement 2). Some states did not have road data for all years. For example, MADIS only included data for Illinois in 2016. We only consider TDP sites that froze at least 30 times over the three-year study period. We also require that each site had at least 120 AM or PM in situ measurements for the period examined, which was 1 October through 31 May for water years 2016, 2017 and 2018. Because of these requirements, AM comparisons were made at 1072 and 29 stations in CONUS and Alaska. PM comparisons were made at 958 stations in CONUS and 29 stations in Alaska.

[4] $T_{eff}$ Version 5 was taken from SMAP SM Version 5, available at https://doi.org/10.5067/ZX7YX2Y2LHEB

[5] $T_{eff}$ Version 4 was taken from SMAP SM Version 4, available at https://doi.org/10.5067/OBBHQ5W22HME

[6] MADIS data available at https://madis.ncep.noaa.gov/

[7] http://www.roadweather.alaska.gov/iways/roadweather/forms/Glossary.html

## 3. Methods

### 3.1. NASA baseline FT algorithm (FT-NPR)

FT-NPR uses a seasonal threshold approach (STA) to categorize landscape states as frozen or thawed. The STA first step computes NPR values for each grid cell, date and overpass

$$NPR = 100 * (Tb_V - Tb_H)/(Tb_V + Tb_H) \quad (1)$$

where $Tb_V$ and $Tb_H$ are the vertical and horizontal polarization Tb values, respectively.

The second step is to determine the most representative values of NPR for frozen and thawed conditions. Freeze reference values ($NPR_{fr}$) are the average of the 20 smallest NPR values occurring during the coldest months of the year (January and February). For NASA FT version 2, thaw references ($NPR_{th}$) are the average of all NPR values occurring during the hottest two months of the year (July and August) (Dunbar et al., 2018). The version 2 averaging scheme was found to provide better accuracy metrics compared to early approaches using the average of the 20 highest summer NPR values (in version 1).

The STA third step calculates a seasonal scaling factor ($\Delta(t)$) for each date and overpass at SMAP grids where both $NPR_{th}$ and $NPR_{fr}$ could be computed

$$\Delta(t) = (NPR(t) - NPR_{fr})/(NPR_{th} - NPR_{fr}) \quad (2)$$

The denominator of Eq. (2) is also referred to as the dynamic range ($\Delta NPR$). Finally, calculated $\Delta(t)$ values are compared to a threshold value to determine whether a landscape is frozen or not. For FT version 1 and 2 algorithms, if $\Delta(t)$ exceeds 0.5, the landscape is considered thawed; otherwise it is considered frozen.

### 3.2. Alternate FT retrievals (FT-R)

The FT-R algorithm consists of three parts: (1) cold period mask (CPM) development, (2) $NPR_{thr}$ value determination and (3) FT state determination by grid cell (Fig. 1).

The goal of the CPM is to determine, for each grid cell, the periods during which frozen retrieval should be attempted. To be consistent with the FT-NPR approach the CPM is limited to those SMAP grid cells that have a valid $NPR_{fr}$ value, meaning cells were not assigned a masked value and that cell values are within the valid range as given in the data product. A grid cell's CPM has a freeze start month (fr_start) and a freeze stop month (fr_stop) that are determined based on the monthly mean effective temperature ($\overline{T_{eff}}$), as shown in the upper part of Fig. 2. For FT-R, frozen retrievals are not attempted for periods outside the fr_start to fr_stop months, instead the FT state is set to thawed. To allow for the most frozen retrievals possible, the CPM uses the coldest $T_{eff}$ values that have occurred since SMAP launched, namely water year 2018 AM values (Kraatz et al., 2018). Because this study mainly focuses on frozen conditions in CONUS, only $T_{eff}$ values occurring between 1 October and 31 May were used to calculate the monthly $\overline{T_{eff}}$ values. Subsequent analyses only were applied to the 'FT-R domain'. The FT-R domain is defined by the intersection of the CPM as defined in (Fig. 1, 'cold period mask') and $NPR_{fr}$ values (Fig. S3). The CPM is defined by grid cells that have a value for both fr_start and fr_stop and where fr_start occurs before fr_stop (Fig. 1, 'cold period mask').

$NPR_{thr}$ values are determined from linear regression following Kim et al. (2011), but instead of using the entire years' time series, we only performed linear regression for observations between fr_start to fr_stop months, inclusive. Additionally, the regression in our study is between NPR versus $T_{eff}$ rather than $\Delta(t)$ versus weather station temperature used in Kim et al. (2011). The linear regression was conducted at every grid in the FT-R domain and the y-intercept (the NPR value) that corresponds to $T_{eff}$ = 0 °C was used as the threshold for delineating frozen and thawed conditions (Fig. 1, '$NPR_{thr}$'). We determined two sets of $NPR_{thr}$ values, based on AM and PM retrievals, respectively.

Finally, NPR(t) values for each retrieval inside the FT-R domain are compared to $NPR_{thr}$ to determine whether conditions are thawed or frozen. If NPR is greater than the grid cell's $NPR_{thr}$, then the road is denoted as thawed. Otherwise, if NPR is less than or equal to the grid cell's $NPR_{thr}$, the road is frozen (Fig. 1, 'FT').

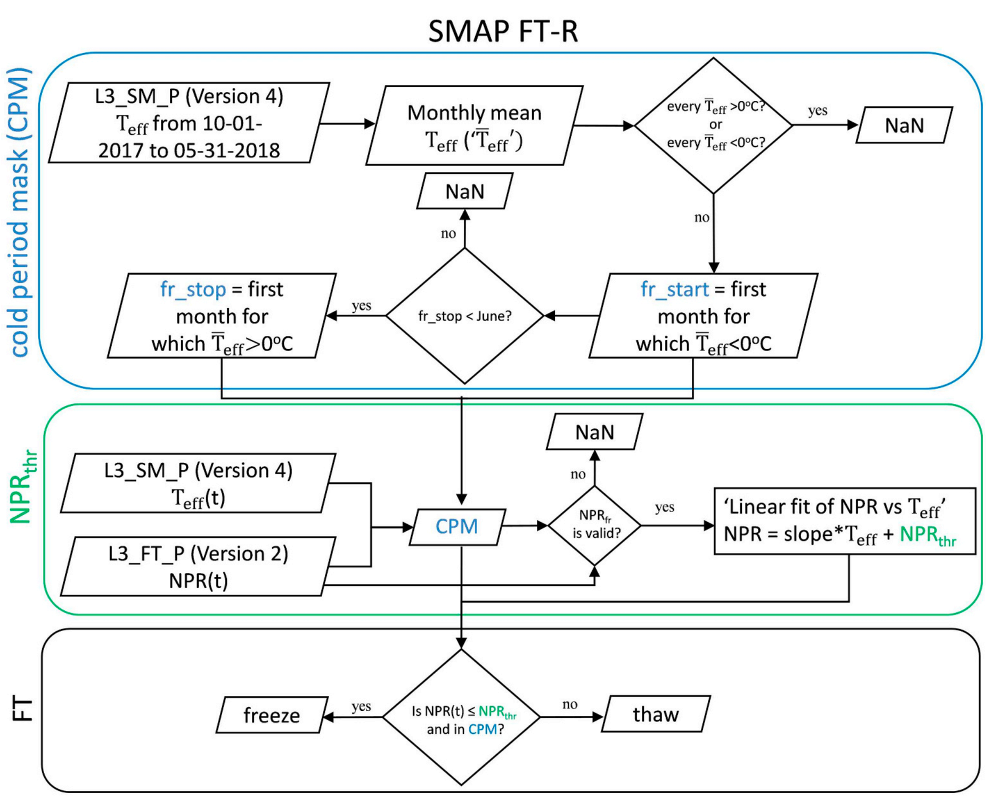


**Fig. 1.** Flowchart showing steps to develop alternate freeze-thaw retrievals (FT-R), and consists of (1) cold period mask (CPM) development, (2) NPR threshold values ($NPR_{thr}$) determining based on a pixel-wise linear fit of the NPR and $T_{eff}$ values (Kim et al., 2011) and (3) NPR values to $NPR_{thr}$ and the CPM comparison in order to determine whether a SMAP grid is frozen or thawed. The CPM only uses morning overpass (AM) data values. $NPR_{thr}$ and freeze-thaw (FT) values are found separately for AM and PM data.

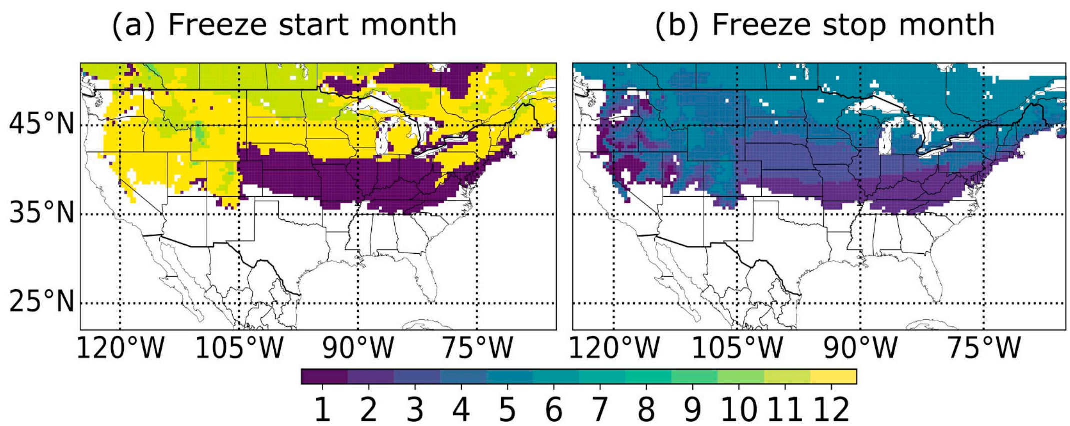


**Fig. 2.** The cold period mask by grid cell based on monthly (October – May) effective soil temperature values for water year 2018 where (a) the start month (first month of subzero average temperature) and (b) the stop month (first month of above-zero average temperature) occur within the October to May time period.

### *3.3. Comparison of NASA FT, FT-R, and road data*

Several approaches are used to compare NASA FT, FT-R and in situ road observations. First, to better understand how the algorithms differ spatially, we mapped each result and conducted qualitative comparisons. Second, to evaluate algorithm performance over time for the same locations, we plotted satellite FT results along with RT time series for sample grid cells. Third, qualitative spatiotemporal comparisons are conducted for CONUS via mapping 5-day averages for FT and RT, for several 5-day periods on the same figure. Fourth, RT values were examined to determine how they differ by FT category (frozen or thawed). RTs values were classified by SMAP FT state and summarized using boxplots. Fifth, probability density plots about 0 °C using for 2 °C RT bins were used to determine the distribution of the retrieval errors. Sixth, to quantitatively evaluate strengths and weaknesses of the two SMAP FT algorithms, we calculated performance metrics for every RT using a confusion matrix, and aggregated these metrics averaged by state. A confusion matrix summarizes the four possible outcomes for any coincident pair of satellite FT and in situ RT data. The two correct outcomes occur if both are frozen (SMAP = 1, RT = 1) or both are thawed (SMAP = 0, RT = 0). Errors occur when satellite and road FT states disagree. If the satellite misses an observed freeze (SMAP = 0, RT = 1), then it is noted as a "Missed Freeze" event. If the satellite indicates frozen conditions when the observed values indicate that the road is thawed (SMAP = 1, RT = 0), then it is noted as a "False Freeze" event. We calculate the Missed Freeze, False Freeze, Freeze Accuracy, Thaw Accuracy and Overall Accuracy metrics as follows

$$\text{Missed Freeze} = 100 * N_{SMAP=0,RT=1}/(N_{total}) \tag{3}$$

$$\text{False Freeze} = 100 * N_{SMAP=1,RT=0}/(N_{total}) \tag{4}$$

$$\text{Freeze Accuracy} = 100 * N_{SMAP=1,RT=1}/(N_{total,RT=1}) \tag{5}$$

$$\text{Thaw Accuracy} = 100 * N_{SMAP=0,RT=0}/(N_{total,RT=0}) \tag{6}$$

$$\text{Overall Accuracy} = 100 * (N_{SMAP=1,RT=1} + N_{SMAP=0,RT=0})/(N_{total}) \tag{7}$$

where N are counts of the combined SMAP and RT states, $N_{total}$ is the total number of events and $N_{total,\ RT=1}$ and $N_{total,\ RT=0}$ are the total number of RT frozen and thawed occurrences, respectively.

In this scheme, Missed Freeze, False Freeze and Overall Accuracy add up to 100%. Freeze Accuracy and Thaw Accuracy indicate the proportion of in situ frozen and thawed conditions detected by SMAP FT, out of the total possible.

All the analyses noted above were conducted for the CONUS RWIS stations. Confusion matrices were also calculated for the 29 Alaskan stations (step six). These results were used to assess SMAP's performance in CONUS versus Alaska.

## 4. Results and discussion

### *4.1. Cold period mask (CPM)*

The CPM shows that freeze start and stop months in CONUS were delimited to latitudes north of 35°N and that most regions south of 37°N did not have below-freezing monthly average temperatures (Fig. 2). Regions at the southern boundary only froze for a short duration except when located at higher altitude. As expected, southern regions freeze later (December or January) and thaw earlier (January or February) than northern regions and elevated areas (i.e. the Rocky Mountains) which typically freeze earlier (October and November) and thaw later (April or May).

The CPM includes some features that are more difficult to explain. Despite water masks in the Great Lakes region, land between James Bay and the Great Lakes appears to have a two-month delay in its freeze onset as compared to the surrounding areas. Similarly, the Lake Superior region's January freeze onset was much later than nearby areas, which were frozen by November. While not explored in detail, we note that these region's numerous smaller waterbodies and wetlands rather than a single, contiguous waterbody, may cause delay the freeze onset for similar reasons as to those that caused delays in each of the Great Lakes.

Monthly $T_{eff}$ values along the West Coast did not fall below freezing except for regions north of 50°N. Kraatz et al. (2018) also found that $T_{eff}$ values along the West Coast were warmer than the same latitudes for East Coast by nearly 10 °C for 2016, 2017 and 2018. Furthermore, NASA FT could not calculate $NPR_{fr}$ values for the same West Coast region (Fig. S3).

We expect that these $T_{eff}$ features would not significantly impact accuracy metrics, because only roads intersecting both the CPM and $NPR_{fr}$ masks were evaluated and few, if any, out of the 44 RWIS stations in Michigan occur in the regions indicating a January freeze onset. If these relatively late freeze onsets are erroneous, those errors would show as those sites having a relatively larger proportion of missed frozen conditions relative to others in the area.

### *4.2. Correlation between NPR and $T_{eff}$ for different temporal subsets*

$NPR_{thr}$ value uncertainty, derived from the linear fit between NPR and $T_{eff}$, will decrease with increasing correlation between NPR and $T_{eff}$. Fig. 3 shows that coefficient of correlation substantially improved nearly everywhere in CONUS when only cold periods were considered as compared to year-round temperatures. The biggest improvements occur south of 48°N and in the Great Plains region.

These CPM relationships are statistically significant with nearly all regions in CONUS having *p*-values were $< 0.05$. Locations that did not

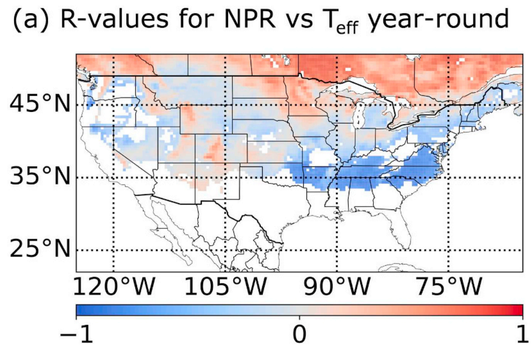


**Fig. 3.** Coefficient of correlation values for the morning (AM) comparisons between the normalized polarization ratio (NPR) and effective soil temperature ($T_{eff}$) for year-round values (a), and for when the temporal subset corresponding to the cold period mask (CPM) is used (b). Results for the PM overpass were similar (not shown).

have statistically significant relationships ($p > .1$, Fig. S4) typically occurred near boundaries of the FT-R domain, especially along the southern edge east of 105°W. These results are to be expected because those transitional regions will be thawed in some of the years where conditions were warmer (e.g., 2016 and 2017). A few SMAP grids located in the Northeast (NH, ME, VT, NY and PA) did not have statistically significant relationships, which might be due to a difficulty in obtaining a strong NPR signal over needleleaf forests (Kraatz et al., 2018).

Where CPM relationships were strong, $NPR_{fr}$ values typically exceeded one (Fig. S3). Most locations that did not have statistically significant relationships had the smallest $NPR_{fr}$ values (near zero). We lateron noticed that $NPR_{th}$ values were not populated for those grids in the NASA FT product, meaning that $NPR_{thr}$ values could not be computed at those locations (Eq. 2). At these locations NASA FT states were instead determined using FT-SCV. Therefore, in the few locations not having $NPR_{th}$ values, comparisons are between FT-SCV and FT-R. Overall, nearly all comparisons correspond to the two NPR-based approaches (FT-NPR and FT-R).

The strengthened relationship between NPR and $T_{eff}$ when the CPM was used is consistent with previous findings (Kraatz et al., 2018). The degree of correlation between NPR and physical temperature had also been used by Derksen et al. (2017) to evaluate the performance of SMAP FT algorithms. But this type of comparison makes the inherent assumption that NPR and temperature should usually be directly related in that low NPR values correspond to dry and colder conditions (i.e. frozen, relative permittivity < 5) and high NPR values correspond to wet and warmer conditions (i.e. thawed, relative permittivity > 5) conditions (Lemmetyinen et al., 2016). However, factors not directly related to soil temperature also impact the magnitude of NPR throughout the year. For example, NPR values are impacted by snow, vegetation, precipitation events, ephemeral water (e.g., floods) and droughts (Entekhabi et al., 2014; Kraatz et al., 2018; Lemmetyinen et al., 2016; Schroeder et al., 2015; Schwank et al., 2015). Probably the most important impact is that small NPR values, which would typically indicate frozen conditions, can temporally coincide with high soil temperatures during summer (Kraatz et al., 2018). Omitting these hot and dry data by limiting the analysis period using the CPM improves the linear regression fit between NPR versus $T_{eff}$ and also improves the NPR intercept estimates needed to differentiate frozen from thawed conditions.

### 4.3. FT-NPR versus FT-R NPR threshold values

$NPR_{thr}$ values for NASA FT-NPR were calculated by using the provided $NPR_{th}$ and $NPR_{fr}$ values and solving Eq. (2) for NPR(t) where $\Delta(t)$ was set to 0.5. Fig. 4 maps the difference between the two $NPR_{thr}$ values, FT-R minus FT-NPR. The $NPR_{thr}$ values for FT-R are higher (by 1 to 2) for CONUS between about 120°W to 80°W. Positive differences mean

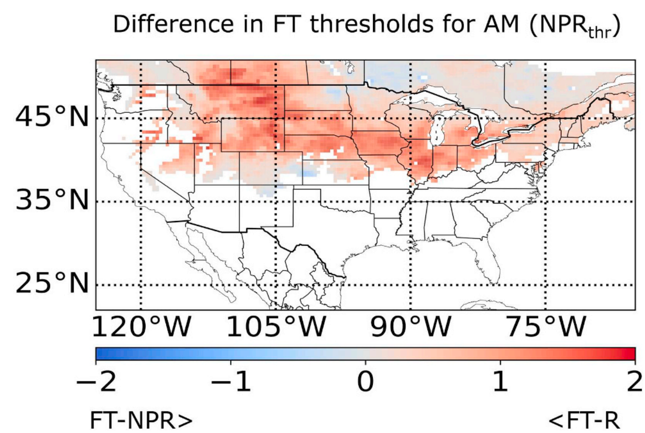


**Fig. 4.** Normalized Polarization Ratio threshold ($NPR_{thr}$) differences between FT-R and $NPR_{thr}$ values obtained from NASA FT (specifically FT-NPR) during the morning (AM) overpass. Positive values indicate that FT-R exceeds FT-NPR.

that NPR time series would more often be classified as frozen using FT-R compared to FT-NPR. In regions with limited differences such as most of Canada and the Northeast, the algorithms ‘could’ perform similarly. ‘Could’ is used, because FT retrievals are binary and even small differences in $NPR_{thr}$ values may produce substantial differences between NASA FT and FT-R classifications, depending on the distribution of NPR (t) values about $NPR_{thr}$ at each SMAP grid.

### 4.4. Basic differences between NASA FT and FT-R in CONUS

Comparisons were made between NASA FT and FT-R in 5-day windows (Fig. 5). For these comparisons, FT averages were calculated from available FT retrievals where FT values are set to 0 and 1 for thawed and frozen conditions, respectively. The 5-day mean FT state is a value between 0 and 1 inclusive. The spatial extent of the FT-R product differs from NASA FT because NASA FT also includes FT-SCV retrievals, it has better coverage in CONUS. Overall, the qualitative differences between NASA FT or FT-R retrievals are relatively minor: there are no major fundamental differences with respect to the large-scale structure and distribution of FT. This is expected, because FT states are determined from the same NPR time series and only differ by the $NPR_{thr}$ values. FT-R frozen conditions are more contiguous compared to NASA FT. Also, frozen conditions extend further south for FT-R, especially near the Great Lakes and the Rockies.

Seasonal FT transitions in roads often have diurnal variations that can be used as indicators for thaw onset. In the first 5-day period shown, NASA FT PM retrievals are as cold or even colder than NASA FT AM retrievals while FT-R results show PM thawing in many regions. In the second 5-day period shown, NASA FT PM retrievals also show

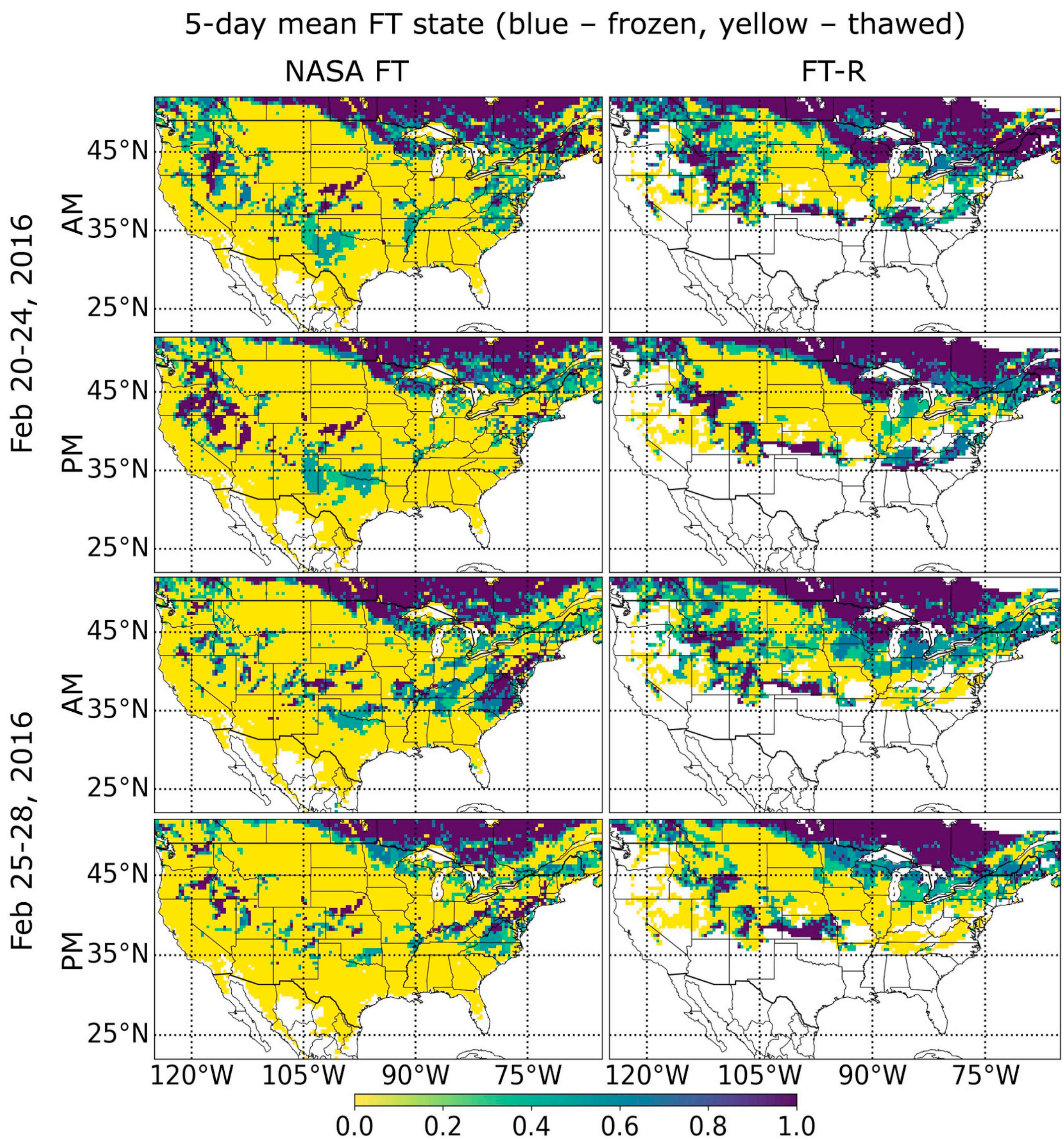


**Fig. 5.** Mean values of NASA (left) and FT-R (right) retrievals for two subsequent 5-day periods for morning (AM) and evening (PM) data (rows). The background grid is colour-coded using SMAP AM or PM retrievals. The 5-day mean values are calculated from the daily values for thaw (zero) and frozen (one). The 5-day mean values range from consistently thawed (yellow) to consistently frozen (dark blue). White areas are locations where NASA FT is not produced. State, country and coastlines are indicated by a black line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

thawing but to a lesser extent than the FT-R results show PM thawing in many regions. In the Rockies, NASA FT appears to be much colder during both 5-day periods for PM compared to AM, while FT-R clearly shows warmer conditions for PM than for AM. Part of the difference could be explained by FT-R having two different $NPR_{thr}$ values one each for AM and PM retrievals as compared to the single $NPR_{thr}$ value used by NASA FT.

### 4.5. Comparisons between SMAP FT and in situ road FT

#### 4.5.1. Sample road stations

The SMAP FT algorithms performance was analyzed at about 1000 RWIS stations. Representative results for two RWIS stations also show that FT-R retrievals indicate frozen conditions more frequently than NASA FT (Fig. 6). Compared to NASA FT, frozen retrievals for FT-R occur more frequently when RTs also indicate frozen conditions. Whereas NASA FT frequently indicates thawed conditions while RTs are below freezing. NASA FT also often indicates thawed conditions when RTs are clearly below −5 °C.

While FT-R frozen retrievals have better correspondence with RTs compared to NASA FT, it cannot detect frozen conditions occurring ahead of the CPM start date. For example, the frozen conditions ahead of the CPM December start date in November at the 010NH site. Because the CPM at this site did not start until December, those retrievals were missed. The advantage of the CPM is that it can prevent some false frozen retrievals such as those occurring during October for NASA FT. Although some early season frozen conditions were missed by FT-R, its flags overall have better correspondence with RTs and are able to capture many more frozen conditions.

Furthermore, it is important to note that FT classifications at this SMAP grid appear to be quite sensitive to $NPR_{thr}$: Figs. 4 and 6 shows that the $NPR_{thr}$ value for FT-R is only slightly larger than that for FT-NPR, but that such a small change resulted in considerably more frozen classifications. $NPR_{thr}$ values are indicated in Fig. 6 by the thin dashed lines which are plotted at 1.55, 1.94 and 1.58, 1.78 for NASA FT, FT-R and sites 010NH and CO056, respectively. Results presented in this

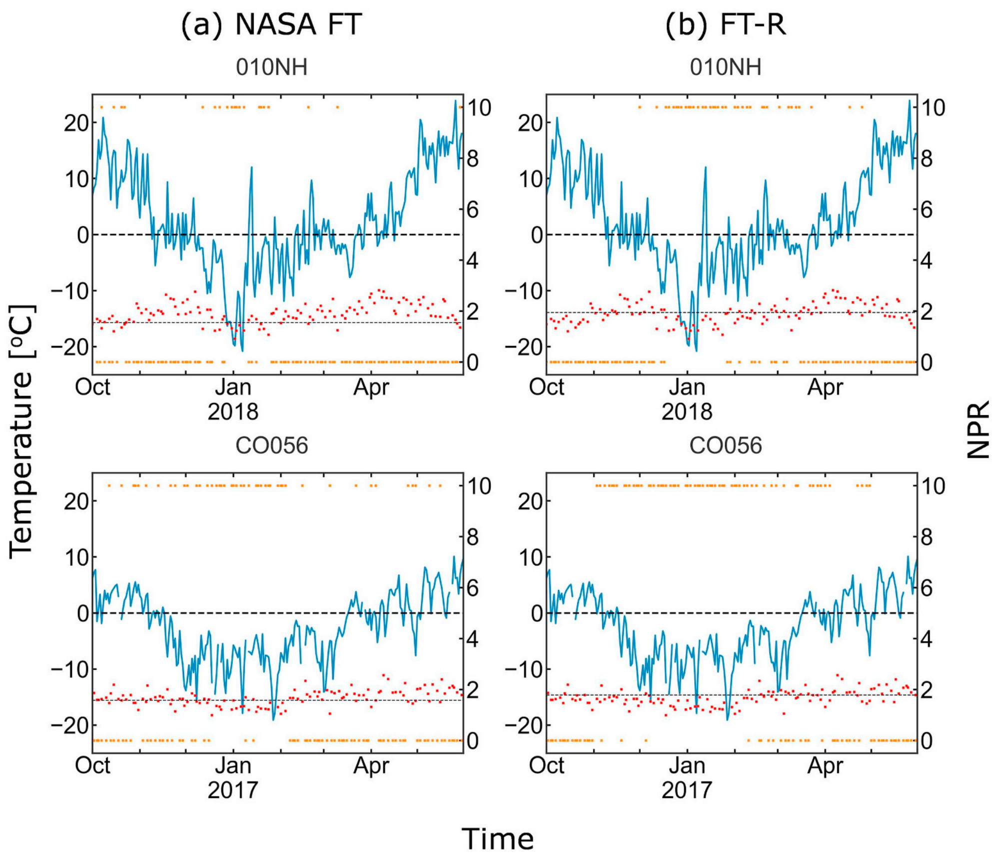


**Fig. 6.** Comparison of AM (morning) road temperatures to time series of NASA FT (left column) and FT-R (right column) for the RWIS 010NH site located in New Hampshire (top row) and the RWIS CO056 site in Colorado (bottom row). Orange and red markers are the FT classifications [0 = thawed, 1 = frozen] × 10 and NPR time series, respectively. The thin dashed black indicates the freeze-thaw delineating value ($NPR_{thr}$), which depends on location and whether NASA FT or FT-R are used. Freeze start and freeze stop values for RWIS sites 010NH and CO056 are December and April, and November and April, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

work can also be used to better understand FT sensitivity to $NPR_{thr}$ throughout CONUS by considering the difference between FT-NPR and FT-R $NPR_{thr}$ values (Fig. 4) in context with the effects they have on the accuracy metrics presented in Section 4.5.5.

#### 4.5.2. National comparison of SMAP FT states and in situ road FT

SMAP FT retrievals were mapped as described in Section 4.4 and the in situ road FT states were overlain on this map. The road FT state was determined based on its 5-day mean RT value where a 5-day mean RT value below 0 °C it was considered to be frozen otherwise it was considered to be thawed. Comparisons use the same 5-day periods as Fig. 5, but for 2018 instead of 2016.

Frozen extents were much greater in 2018 (Fig. 7) than in 2016 (Fig. 5) for the same periods. Again in 2018, FT-R results are consistently colder than NASA FT. For the 2016 SMAP FT retrievals, frozen conditions were mainly restricted to mountainous regions or areas north of 45°N with some isolated frozen regions further south (Fig. 5). Whereas for 2018, SMAP FT retrievals show that frozen conditions are widespread over much of central and western CONUS. Unlike for 2016, the 2018 results show continuous, connected frozen conditions extending as far south as 35°N for both NASA FT and FT-R. The larger frozen extents observed for 2018 as compared to 2016 are consistent with $T_{eff}$ values that Kraatz et al. (2018) mapped for CONUS.

Fig. 7 shows that both FT retrievals have good correspondence with the observed road states. The main difference between the FT products is that NASA FT AM retrievals indicate considerably more thawed regions than the RTs. The NASA FT product has many "islands" of thawed areas, surrounded by otherwise frozen areas, especially between 90°W to 115°W. For those areas, roads observations typically indicate frozen conditions. In principle, the coverage of frozen regions is reasonable for NASA FT, but it is too small compared to that indicated by RTs. In comparison, FT-R retrievals are in very good agreement with road FT states for the 20–24 February period. Beyond 25 February 2018 as spring thaw progresses, road observations indicate strong diurnal temperature swings about 0 °C with the result that FT states oscillate between frozen in AM and thawed in PM. Strong diurnal FT signals in roads were noted in past studies of low volume roadways in Alaska and CONUS (Kraatz et al., 2019a, 2019b).

Both SMAP FT algorithms are temporally consistent with each other and the RTs. They were both able to capture relative cool periods observed near the end of the spring thaw period (Fig. S5). Overall, both algorithms are able to capture substantial differences in FT over time (years or 5-day periods), and SMAP FT trends appear to be very consistent with $T_{eff}$ values and/or in situ road conditions.

Kraatz et al. (2019c) suggested using the SMAP FT Version 1 AM frozen and PM FT thawed retrievals as an early indicator for impending road thawing that may be useful to identify those roads that would likely also soon be consistently thawed. At CONUS scale, neither SMAP FT retrieval appears to produce a substantial difference in FT between AM and PM observations. The diurnal range of SMAP FT is relatively smaller than that found in roads. While SMAP FT may perform relatively better for roads in late spring for AM (FT-R) or PM (NASA FT) retrievals, neither approach is able to capture the late spring AM frozen to PM thawed transitions occurring in roads. As the thaw season progresses, an increasing number of roads experience the AM frozen to PM thawed transition until nearly all roads become thawed for PM (Fig. 7). At the same time, many roads show in situ frozen conditions for AM. While NASA FT results appear to correspond better with PM road thaw later during spring thaw, they do not accurately indicate frozen conditions for AM FT retrievals (Fig. 7).

In general, it should be expected that SMAP AM and PM retrievals would not be able to match the road FT states during late spring thaw (e.g. for 2018, 25 February and later). The SMAP field of view (FOV) is approximately 39 km × 47 km and will incorporate all landscape elements well beyond the built environment - including snow-covered and frozen vegetation. Roads are usually cleared of snow and have a

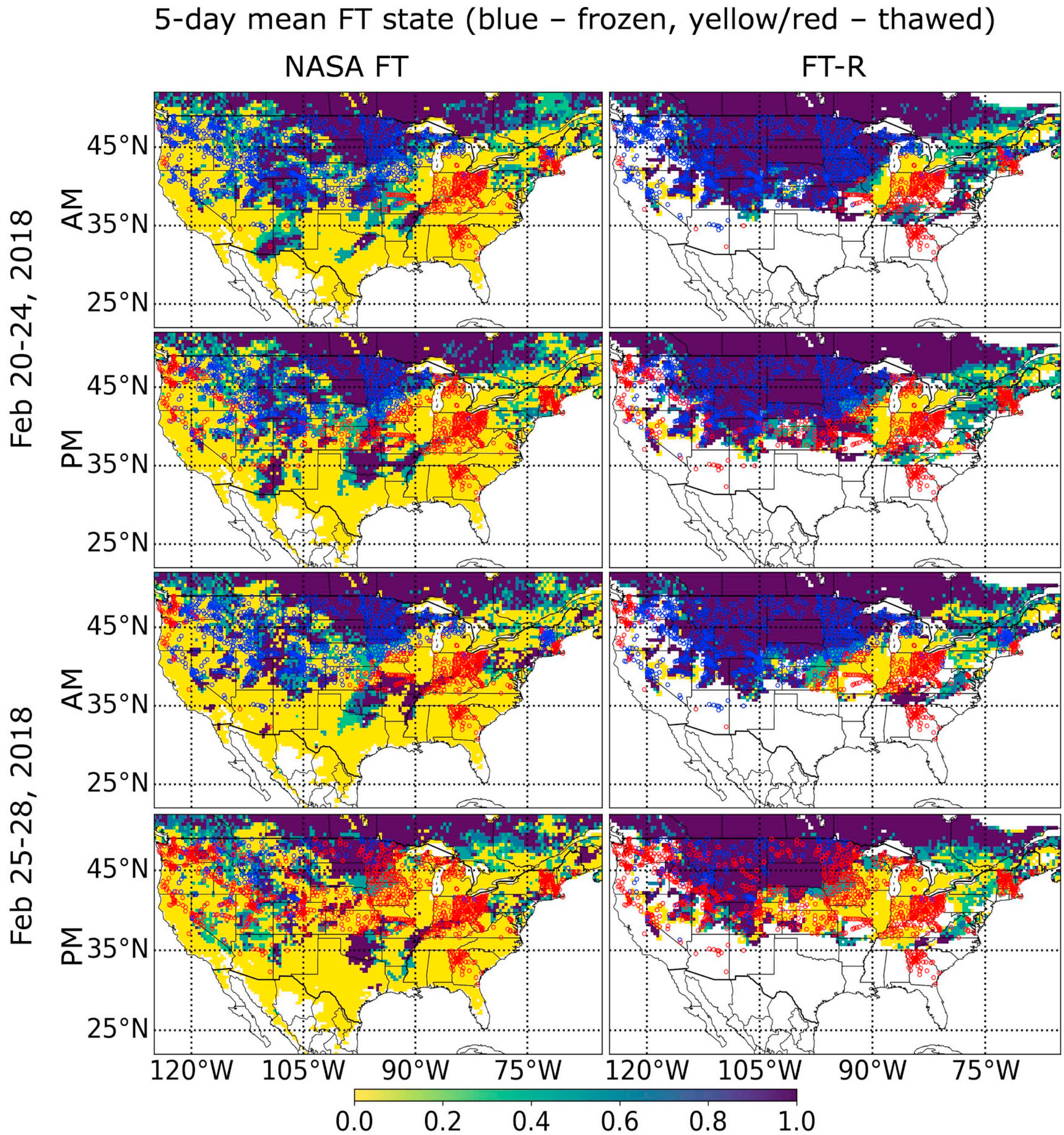


**Fig. 7.** Mean values of NASA FT (left) and FT-R (right) retrievals for two subsequent 5-day periods for morning (AM) and evening (PM) data (rows). The background grid is colour-coded using SMAP AM or PM retrievals. The 5-day mean values are calculated from the daily values for thaw (zero) and frozen (one). The 5-day mean values range from consistently thawed (yellow) to consistently frozen (dark blue). White areas are locations where SMAP FT is not produced. State, country and coastlines are indicated by a black line. The blue and red circles indicate locations of Road Weather Information System (RWIS) stations and whether they are frozen or thawed, respectively. RWIS stations' FT state is set to frozen if the 5-day mean RT is below 0 °C, otherwise the FT state is set to thawed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

relatively lower albedo and higher absorption solar radiation as compared to natural sites that dominate the SMAP FOV. Furthermore, roads are mainly impervious to water and road substructures are specifically designed to drain water efficiently. For these reasons, the diurnal temperature variations in roads should be substantially greater than what is possible in natural landscapes. SMAP FT retrievals are sensitive to liquid water, and if PM thawing of the landscape does not produce a strong wet signal then probably either the AM or the PM SMAP FT state would likely agree with road observations, not both.

#### 4.5.3. Road temperatures by SMAP FT state

Overall, SMAP FT can clearly distinguish between frozen and thawed roads in CONUS based on the individual RT values (Fig. 8). The 25th through 75th percentile RT values are clearly separated with little, if any, overlap indicating that SMAP FT is able to discriminate between frozen and thawed roads irrespective of whether comparisons are made for NASA FT, FT-R and AM or PM retrievals. In line with previous observations regarding the strong diurnal FT swings in roads (Fig. 7), PM RTs are clearly warmer than for AM, by approximately 5 °C. Overall, NASA FT retrieved 79% and 83% of AM and PM roads as thawed, respectively. Whereas FT-R retrieved 68% and 70% of AM and PM data as thawed, respectively.

Fig. 8 also shows that FT-R is relatively better at discriminating frozen from thawed roads compared to NASA FT. For thawed conditions, the median RT values for FT-R are relatively warmer than for NASA FT, resulting in a better separation between frozen and thawed roads. For frozen conditions, the RT interquartile range (IQR) is relatively smaller for FT-R than NASA FT. While the PM median RTs corresponding to frozen conditions are very similar, FT-R achieved a comparable median value and IQR with about 34,000 more data points

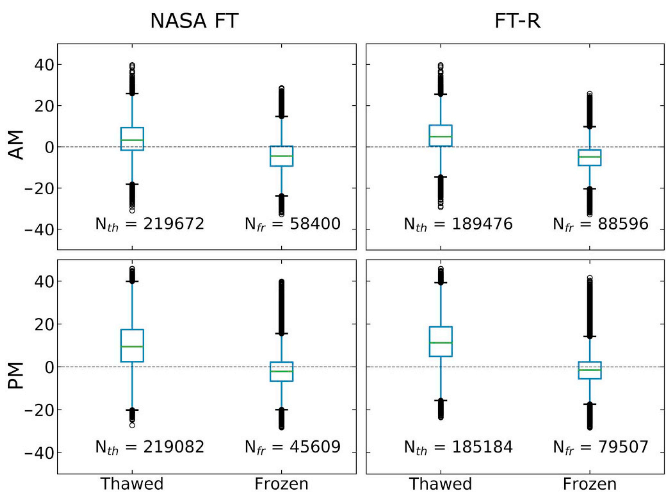


**Fig. 8.** Boxplots of October – May road temperatures (RT) obtained from Road Weather Information System stations in CONUS for when SMAP FT (left column) and FT-R (right column) retrieved a landscape as frozen and thawed. The numbers of data points for SMAP thawed ($N_{th}$) and frozen ($N_{fr}$) retrievals appear below each box. Each box shows the 25th to 75th percentile range of RTs. Whiskers extend to the last data point that is inside the 75th percentile value +1.5*IQR, where IQR is the interquartile range (the 75th percentile – 25th percentile value). Only points that lie outside this range are plotted. The green line is plotted at the median value of the dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

identified as frozen.

#### 4.5.4. Errors between road temperatures and SMAP freeze-thaw

For all cases considered, the most errors occurred when RTs were close to 0 °C, which is also the FT delineating threshold used in this study (Fig. 9). These results support the idea that SMAP FT retrievals have good sensitivity and relatively low bias with respect to RTs. However, the error distribution is very uneven between positive and negative RTs, especially for NASA FT. The distribution shows that NASA FT retrievals are biased towards being too warm resulting in under-detects of frozen conditions. While FT-R AM retrievals also have a warm bias, the bias is smaller compared to NASA FT. The error distribution for FT-R PM is nearly even about 0 °C, indicating an even smaller bias.

#### 4.5.5. Confusion matrix summary of SMAP and road FT results

SMAP FT Freeze Accuracy and Overall Accuracy results were summarized for CONUS (Figs. 10 and 11). FT-R usually captures at least 40% and often 60% or more of in situ frozen conditions. Except for Minnesota and North Dakota, NASA FT only captured 20 to 40% of in situ frozen conditions. Locations where the FT-R approach most improved Freeze Accuracy coincide with locations of Overall Accuracy improvement (Fig. 10 and to a lesser degree Fig. 11). Because there are relatively fewer PM frozen conditions and NASA FT is biased towards detecting thawed conditions, it has comparable Overall Accuracy to FT-R for PM comparisons.

Fig. 10 shows that there are well-defined clusters for which AM the SMAP FT retrieval accuracies were comparable. NASA FT and FT-R both indicate relatively lower performance in the West and Northeast, while performance in the Midwest and South is greatly improved. Overall accuracy varies gradually in space with performance smoothly transitioning between lower and greater values: there are virtually no low performing stations (‘blue’) directly adjacent to high performing stations (‘red’). SMAP FT retrievals accuracy does not vary substantially for nearby stations and at state level.

For PM retrievals, while Freeze Accuracy is substantially improved for both algorithms, the Overall Accuracy does not increase to the same degree for FT-R as it did for NASA FT (Fig. 11). While more in situ frozen conditions are captured, the relatively higher rate of False Freeze detections limits the level of improvement in Overall Accuracy (Fig. 12). False Freeze detections during PM can be attributed to the roads' strong diurnal temperature swing that cannot be captured by

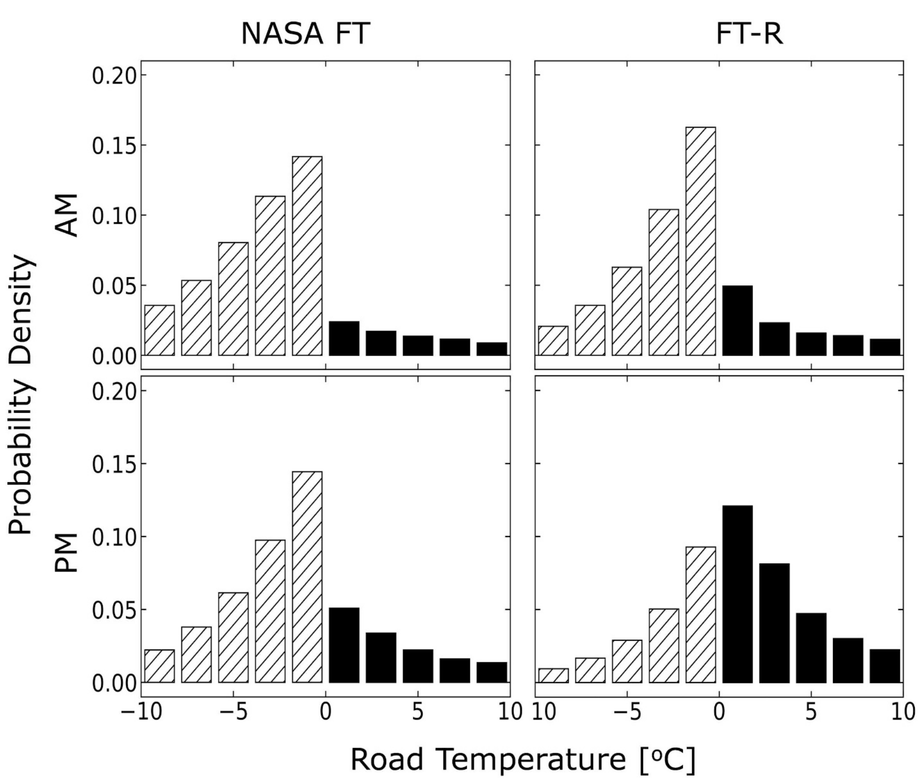


**Fig. 9.** Histograms (2 °C bins) of road temperatures (RT) for the contiguous United States (CONUS) from October to May for erroneous morning (AM) and evening (PM) retrievals for NASA FT (left column) and FT-R (right column). The figure shows two types of errors: (1) RT frozen (≤0 °C) with SMAP indicating a thawed state (hatched); (2) RT thawed (> 0 °C) with SMAP indicating a frozen state (filled).

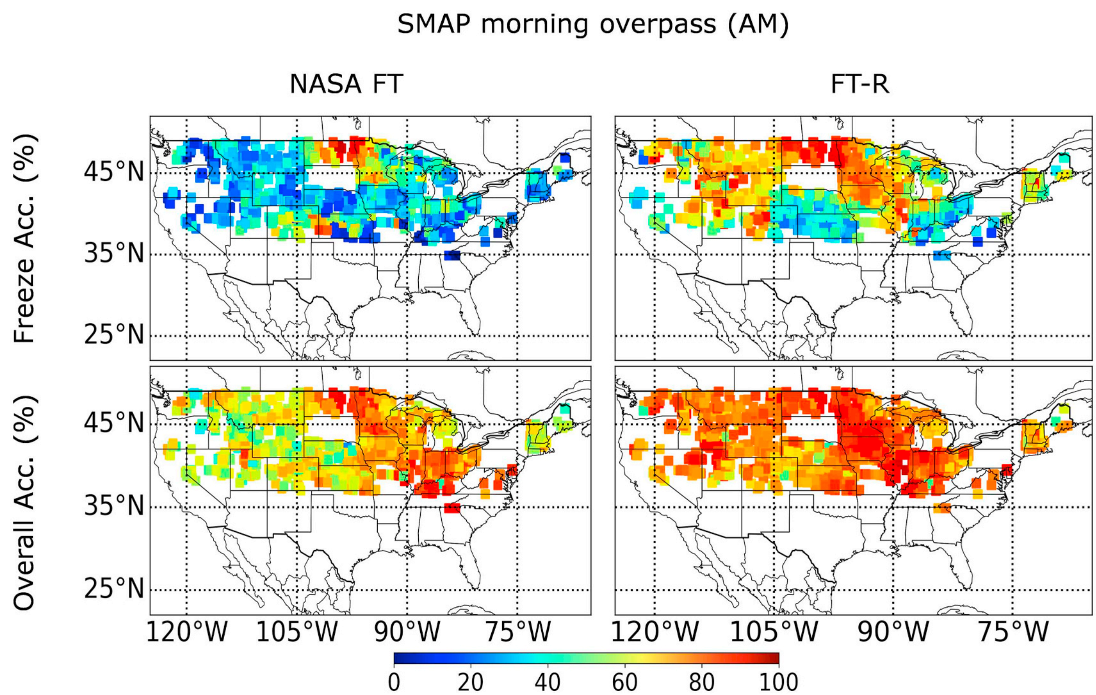


**Fig. 10.** Map of accuracy statistics for road temperatures (RT) for morning (AM) comparisons with RTs data for NASA FT (left column) and FT-R (right column). Shown are statistics for SMAP correctly identifying frozen conditions (top row) and SMAP overall accuracy (bottom row) with respect to roads. Coast, state and country boundaries are shown in black lines.

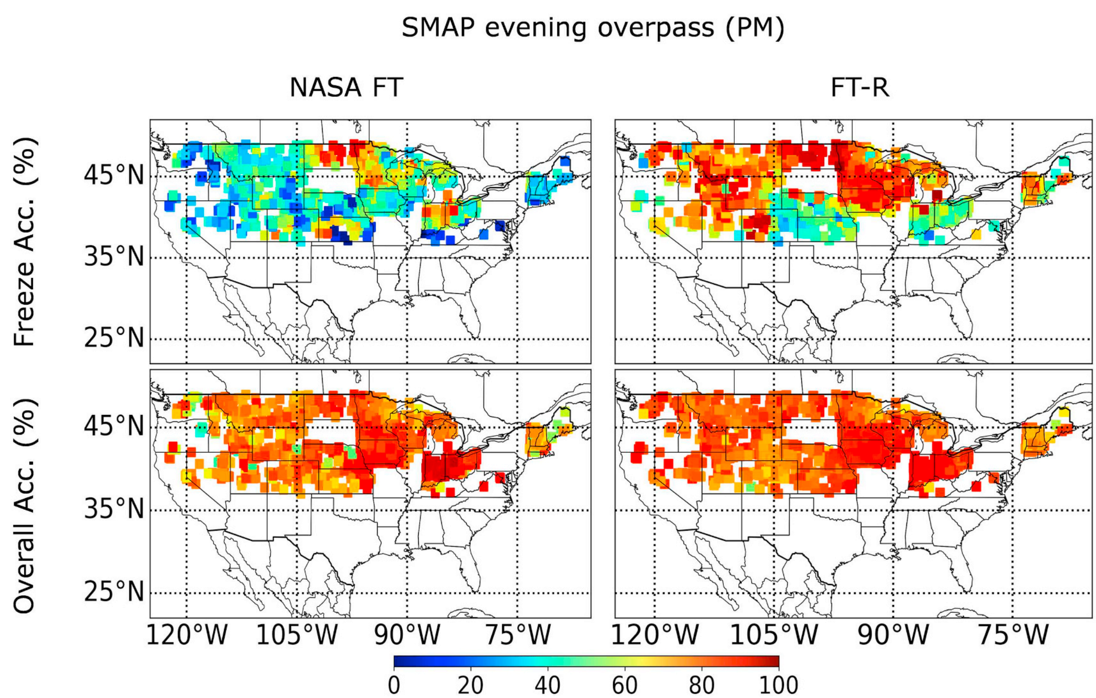


**Fig. 11.** Same as Fig. 10 but for evening comparisons (PM).

either SMAP FT product. The Overall Accuracy for NASA FT and FT-R are comparable for PM retrievals, but FT-R appears to perform slightly better in most states, except for Indiana and Ohio where NASA FT has better Freeze and Overall Accuracies. The strength of FT-R is that it usually performs at least comparable to NASA FT, irrespective of whether AM or PM comparisons are made, and it has greatly improved ability to detect in situ frozen conditions.

Fig. 13 shows that while FT-R also misses some frozen conditions (< 20%), NASA FT retrievals are especially poor west of 105°W where about half of in situ frozen conditions are not detected. This result is consistent with the relatively poor Freeze Accuracy and explains the reduced Overall Accuracy for locations west of 105°W by NASA FT (Fig. 10).

It was noted in Section 4.1 that areas with water bodies (i.e. Great Lakes and surrounding regions) had a delayed freeze onset compared to their surrounding areas (Fig. 2). The FT-R results show that these areas have relatively few missed frozen conditions (Fig. 13), few false freeze detections (Fig. 12); and relatively high Freeze Accuracy and Overall

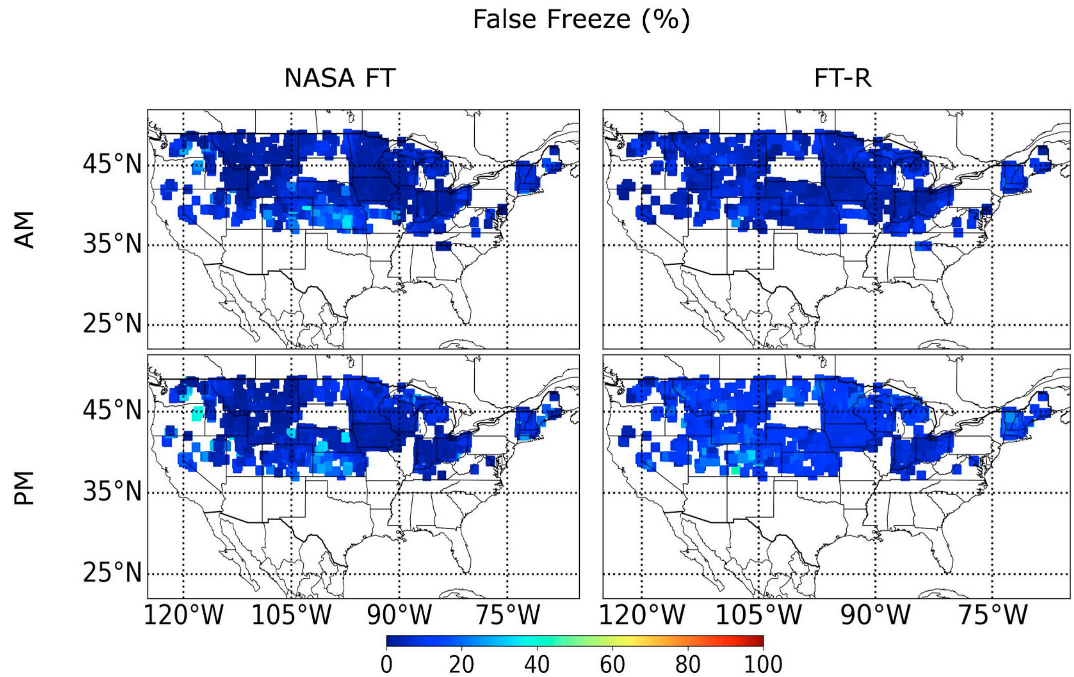


**Fig. 12.** Same as Fig. 10 but for morning and afternoon comparisons of SMAP retrieving frozen conditions while roads are thawed ("False Freeze").

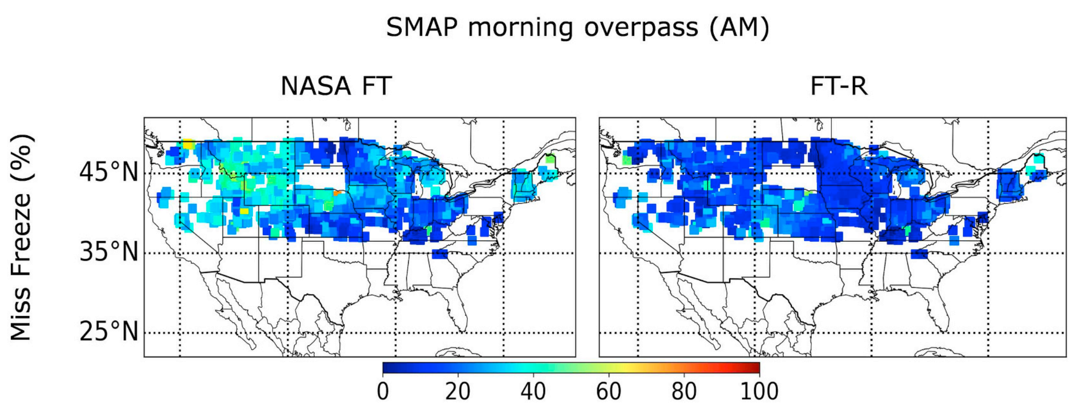


**Fig. 13.** Same as Fig. 10 but for not detecting frozen conditions and morning comparisons. Missed Freeze rates for evening retrievals are comparable and not shown.

Accuracy. Therefore, $T_{eff}$ and the resulting CPM did not appear to negatively impact results near the Great Lakes.

Accuracies for NASA FT and FT-R performance for the 26 CONUS states plus Alaska and their regional performance metrics[8] are summarized in Tables 1 and S6 for AM and PM, respectively. FT-R performed better than NASA FT with overall accuracies in CONUS of 79% and 82% compared to 69% and 80% for AM and PM retrievals, respectively. FT-R Overall Accuracy was substantially better (here, by ≥10%) for 16 (4) out of the 26 CONUS states for AM (PM). While NASA FT may occasionally have greater Overall Accuracy during PM, it is usually comparable to or below the values obtained for FT-R. There were no states or overpass times for which NASA FT performed substantially better than FT-R.

FT-R was substantially better at detecting frozen conditions compared to NASA FT at 22 (22) out of 26 states for AM (PM) retrievals. There were only few states for which NASA FT had better Freeze Accuracy than FT-R, but those improvements did not exceed 10%. FT-R also substantially improved Thaw Accuracy for 2 (1) states for AM (PM) retrievals. NASA FT performed substantially better for 0 (6) states for AM (PM) retrievals, making it somewhat better than FT-R for detecting thawed conditions. This is consistent with the idea that NASA FT is biased towards detecting thawed conditions and would usually have better Thawed Accuracy than FT-R (Section 4.5.4).

When using FT-R, errors for Missed Freeze decreased substantially for 15 (7) states for AM (PM) retrievals. Also, False Freeze errors decreased substantially at 2 (2) states for AM (PM) retrievals for FT-R. NASA FT has a lower False Freeze at many states, but there were no states for which either False Freeze or Missed Freeze was substantially lower for NASA FT. Overall, FT-R performed substantially better than NASA FT and excels at detecting frozen conditions.

SMAP FT retrieval performance also varied strongly across geographic regions (Tables 1, Figs. 10–13). For example, NASA FT performed relatively poorer in the Northeast (59%) and West (62%) compared to retrievals made in the Midwest (74%) or South (82%). But for nearly all states within a region, performance metrics were within ± 6% of the mean. Even in the Midwest and West, encompassing 10 states each, only two states each fell outside a ± 6% range. Whereas regional differences varied by about 9% about their mean (69%), states within the Northeast, Midwest, South and West only varied by 3%, 5%,

[8] United States Census Bureau, Geography Division. "Census Regions and Divisions of the United States". Available at https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

**Table 1**
Summary of accuracy metrics in percent by state and region for morning overpass retrievals for NASA FT and FT-R compared to road temperatures. Results for FT-R are shown in brackets. Statistics were only computed for states having more than four stations ($N \geq 4$). Cells in for which FT-R exceeded (fell below) SMAP FT by 10% or more are colored in green (red) for increases (decreases). Changes for afternoon comparisons are less substantial (Table S6). MF: missed freeze, the rate of SMAP identifying in situ frozen conditions as thawed; FF: false freeze, the rate of SMAP identifying in situ thawed conditions as frozen; FA: freeze accuracy, the proportion of all in situ frozen conditions detected; TA: thaw accuracy, the proportion of all in situ thawed conditions detected; Overall: overall accuracy, the proportion of correct retrievals out of all retrievals; Total: for N the sum of stations in region, for accuracy metrics it is the state-average weighted by N.

| | | | Errors (%) | | Accuracy (%) | | |
|---|---|---|---|---|---|---|---|
| Region | State | N | MF | FF | FA | TA | Overall |
| Northeast | ME | 7 | 36 (24) | 10 (6) | 31 (54) | 79 (87) | 54 (69) |
| | MA | 9 | 27 (15) | 10 (7) | 23 (56) | 85 (90) | 64 (78) |
| | NH | 22 | 35 (22) | 7 (9) | 27 (53) | 87 (83) | 58 (69) |
| | VT | 30 | 31 (17) | 9 (11) | 36 (65) | 81 (77) | 60 (72) |
| Total | Northeast | 68 | 32 (19) | 9 (9) | 31 (59) | 83 (82) | 59 (72) |
| Midwest | IA | 64 | 25 (12) | 1 (3) | 31 (68) | 98 (95) | 73 (85) |
| | IL | 25 | 21 (10) | 1 (2) | 33 (70) | 99 (97) | 78 (88) |
| | IN | 24 | 17 (14) | 3 (3) | 33 (44) | 96 (96) | 80 (82) |
| | KS | 83 | 16 (18) | 16 (5) | 36 (30) | 79 (94) | 68 (78) |
| | MI | 44 | 29 (18) | 6 (7) | 39 (61) | 89 (87) | 65 (75) |
| | MN | 85 | 19 (11) | 2 (3) | 63 (78) | 95 (93) | 79 (85) |
| | ND | 28 | 14 (8) | 7 (7) | 74 (86) | 85 (84) | 79 (85) |
| | NE | 50 | 29 (22) | 6 (3) | 31 (46) | 91 (95) | 66 (75) |
| | OH | 101 | 17 (14) | 3 (5) | 35 (43) | 96 (93) | 80 (80) |
| | WI | 38 | 25 (13) | 2 (6) | 43 (69) | 96 (90) | 73 (81) |
| Total | Midwest | 542 | 21 (14) | 5 (4) | 41 (57) | 92 (93) | 74 (81) |
| South | KY | 17 | 11 (9) | 6 (7) | 21 (42) | 93 (92) | 83 (84) |
| | VA | 7 | 12 (13) | 8 (8) | 32 (27) | 90 (91) | 80 (80) |
| Total | South | 24 | 11 (10) | 7 (8) | 24 (38) | 92 (91) | 82 (83) |
| West | CA | 10 | 28 (21) | 9 (5) | 24 (43) | 86 (90) | 63 (73) |
| | CO | 79 | 24 (20) | 10 (8) | 46 (54) | 82 (86) | 66 (73) |
| | ID | 73 | 33 (16) | 4 (4) | 26 (63) | 93 (93) | 63 (80) |
| | MO | 16 | 13 (8) | 13 (10) | 37 (62) | 84 (87) | 74 (82) |
| | MT | 69 | 36 (16) | 2 (5) | 30 (69) | 96 (89) | 62 (79) |
| | NV | 45 | 33 (18) | 7 (5) | 25 (61) | 87 (91) | 59 (77) |
| | OR | 9 | 35 (19) | 15 (3) | 21 (57) | 72 (95) | 50 (78) |
| | UT | 33 | 30 (15) | 6 (5) | 32 (69) | 89 (90) | 65 (80) |
| | WA | 27 | 29 (14) | 8 (6) | 24 (66) | 88 (91) | 63 (80) |
| | WY | 73 | 38 (17) | 3 (6) | 30 (69) | 93 (86) | 59 (77) |
| Total | West | 434 | 31 (17) | 6 (6) | 32 (63) | 89 (89) | 62 (78) |
| Total | CONUS | 1068 | 25 (16) | 6 (5) | 36 (59) | 90 (90) | 69 (79) |
| Alaska | AK | 29 | 22 (12) | 5 (8) | 66 (81) | 91 (82) | 74 (81) |

2% and 4%, respectively.

Further studies would help to more clearly attribute these differences, but most likely they can be explained by differences in land cover and climate. Both can greatly impact retrieval performance, for different reasons. Spatial heterogeneity is related to issues of representativeness of point measurements of the greater SMAP footprint (Alemohammad et al., 2018; Kim et al., 2018, 2019; Kraatz et al., 2018; Peng et al., 2017). Climate and weather can also play a role, because of how NPR references for frozen and thawed conditions are determined in the NASA FT-NPR approach. For example, if $NPR_{th}$ is determined for locally very dry conditions during summer algorithm performance should be decreased, because $NPR_{th}$ is supposed to be a reference for reasonably 'wet' conditions, and the dynamic range of NPR (see Eq. (2)) will decrease (Kraatz et al., 2018).

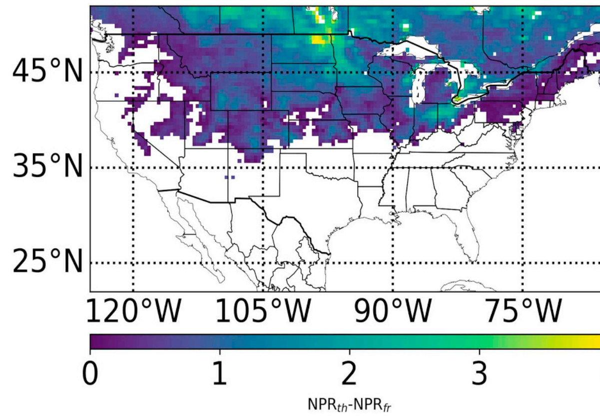


**Fig. 14.** The difference between NASA FT-NPR reference thresholds for thawed ($NPR_{th}$) and frozen ($NPR_{fr}$) conditions.

#### 4.6. Impacts of CPM, evaluation window and ΔNPR on accuracy metrics

Fig. 14 shows that most areas in CONUS have $\Delta NPR < 2$ (NASA FT-NPR AM and PM retrievals use the same NPR thresholds, so ΔNPR is the same for both). ΔNPR values are especially small in the Northeast, West and near the southern edge of the NASA FT-NPR domain, but considerably greater (2 or more) near the Great Lakes the Midwest. Comparisons of Fig. 14 and Figs. 10–13) clearly demonstrate the importance of ΔNPR for FT retrievals. Figs. 10 and 11 show that locations where the highest accuracies were found, closely mirror locations having ΔNPR values approaching 2 or more. While comparisons with Fig. 13 show that low ΔNPR areas tend to have relatively high rates of missed frozen conditions.

It is important for ΔNPR to exceed a minimum threshold value to be able to accurately delineate between frozen and thawed conditions. NASA's requirement of needing ΔNPR to exceed 0.1 for FT-NPR retrievals being made is good for maximizing the FT-NPR domain, but accuracies for areas having such small ΔNPR are most likely going to be poor. Results presented in this study and Kraatz et al. (2018) both hint at sites needing ΔNPR values of about 2 or greater for achieving reasonably good accuracy (e.g. in the 70–80% range). Kraatz et al. (2018) also examined the importance of ΔNPR and Δ(t) (ergo also $NPR_{thr}$) and found that SMAP FT did not perform well at the Idaho CVS, having ΔNPR of only about 1.5, but results at CVSs having $\Delta NPR > 2$ were good. Although the overall accuracy for the Idaho site was in the 70% range, other tests such as evaluating the correlation between NPR and situ temperatures indicated that FT retrievals still performed poorly there. R-values were negative ($R \sim -0.6$) for year-round comparisons and even when a more suitable temporal subset was used for comparisons (October – March) R values were close to zero. That study did not use any error mitigation techniques, so that all the computed FT states corresponded to a retrieval according to Eqs. (1) and (2). Whereas this study used the CPM period for comparing R-values between NPR and $T_{eff}$ and R-values further improved and became positive at most locations, including at Idaho ($R \sim 0.5$). For this reason, regression could be used to adjust $NPR_{thr}$ values to be more accurate during the CPM period.

It is also important to consider the effects of reporting accuracy metrics for different temporal windows. The advantage for using sub-annual temporal windows is to give relatively more weight to those periods for which FT states are computed by Eqs. (1) and (2), as opposed to when FT values are due to a mask. NASA FT-NPR uses a never frozen/never thawed mask, and the FT-R approach uses a never frozen CPM. These masks are important for preventing false detections of frozen conditions during summer. If the evaluation window were focused only on summer or year-round retrievals, accuracy metrics may be more heavily informed by the performance of the masks as opposed to the retrievals and associated NPR threshold values.

While this study includes regions where the evaluation window of October – May is considerably greater than the CPM mask (e.g. for some areas in Idaho the CPM duration is only 2 months), CPM windows constitute half or more of the evaluation window for most areas (e.g. Fig. 2). The value of using short evaluation windows is also reflected in the SPLFTP/_E Assessment Report Version 2 (Xu et al., 2018),[9] which shows comparisons by month at FT CVS sites and by day-of-year for the global assessment. It could also be useful for future studies to provide a breakdown for when FT states were assigned instead of calculated or separate accuracy metrics only for when retrievals are made.

Kraatz et al. (2018) also evaluated SMAP FT accuracy metrics as function of Δ(t) (ergo, $NPR_{thr}$) and showed that this parameter can be viewed as determining the trade-offs between accurately detecting frozen conditions, thawed conditions and overall accuracy. And that while the default value of Δ(t) of 0.5 was shown to be reasonably good, performance usually further improved when selecting grid specific Δ(t) values, with values usually ranging between 0.3 and 1.0. Our evaluation of NASA FT-NPR is consistent with the idea that the current approach for calculating $NPR_{thr}$ values appears to optimize for overall accuracy but at the cost of accurately detecting frozen conditions in CONUS. Fig. 6 showed that our CPM mask is probably relatively more heavy-handed for assigning thawed states (only two locations shown), but this also allows us to better optimize $NPR_{thr}$ values for accurately detecting frozen conditions.

### 4.7. A conceptual model for the correspondence between SMAP and road FT state

Snow cover may impact FT state performance differently for roads as compared to natural sites. In natural sites, snow is known to be a major source of errors with respect to accurately detecting soil FT states. For example, studies by Schwank et al. (2014, 2015) clearly showed that even dry snow impacts L-band FT retrievals, due to refraction and impedance matching. Wet snow also impacts SMAP FT retrievals because the high emissivity of liquid water results in greater NPR values.

FT classification errors over natural sites can be broken down into errors due to dry and wet snow. Dry snow has the effect of decreasing NPR values, and thus may decrease the correspondence of SMAP FT retrievals with soil FT states when thawed soils are covered by dry snow. This type of error should be more prevalent during the freeze-up period while soils may still be thawed, and the snowpack is relatively dry. In contrast, wet snow elevates NPR values, potentially causing the FT delineating threshold to be exceeded and resulting in thawed retrieval state even when soils are frozen. This error should occur during spring thaw and midwinter thaw events.

Snow states that result in FT classification errors for natural sites do not necessarily cause errors in road FT classification because there are important differences in natural and road surfaces. With regards to dry snow errors, comparisons between SMAP and road FT states may be relatively better as compared to natural sites because roads freeze earlier than natural sites (Kraatz et al., 2017) due to the lack of vegetation, litter, and snow that insulate the natural sites. There may also be fewer wet snow errors for roads than natural sites. Kraatz et al. (2017) found that roads have shorter periods of isothermal conditions during the spring transition and that they thaw earlier as compared to natural sites. This difference can be attributed to roads lacking insulating properties, having lower albedo and having increased surface warming capacity as compared to natural sites. Again, SMAP thaw signals may more accurately indicate road FT state than soil FT.

These qualitative comparisons hint at the potential of SMAP FT being more accurate for roads than for natural sites during spring and midwinter thawing. This idea is also supported by SMAP FT comparisons made between nearby natural sites and roads (Kraatz et al., 2017) which clearly showed that road thaw dates closely corresponded to thaw timings from SMAP, both of which occurred about 2–3 weeks ahead thaw at natural sites. Relatively earlier thaw timing detected by SMAP had also been noted in Derksen et al. (2017) and Zheng et al. (2019).

These comparisons are not entirely straightforward because FT transition periods may have a strong diurnal signal in both snowpack ripening and road temperatures (Fig. 7). Differences between road warming and snow ripening processes may result in their respective FT state not aligning. For example, if the snowpack does not become sufficiently wet by the time of SMAP PM overpasses, then retrievals will not reflect the roads thawed state. Because transition periods may persist for a month or longer over the course of winter, roads also have strong diurnal swings that may cause substantial errors.

While a better understanding of wet and dry snow errors as well as diurnal effects during spring thaw periods are needed, it is entirely possible that SMAP FT may perform even better against road temperatures than natural soil temperatures. Results presented in Kraatz et al. (2017) support this idea, but comparisons were mainly focused on transition timing dates and inter-comparisons of temperature time series rather than direct comparisons of FT states between natural sites and roads. Therefore, the conceptual framework presented here needs to be further developed and verified using more direct intercomparisons among SMAP, road and natural FT states.

One limitation of this conceptual model is that it requires snow in the SMAP pixel. When snow is not present, the performance of SMAP FT relative to natural and road locations should be different. Nonetheless, this model is a useful starting point for developing hypotheses and conducting follow-up studies comparing SMAP, natural and road FT states.

### 4.8. Potential of using $T_{eff}$ to determine road FT

$T_{eff}$ was important to this work since it was needed to create the CPM and to determine the new $NPR_{thr}$ values via the regression of NPR versus $T_{eff}$. This study only used the $36^2$ $km^2$ $T_{eff}$ values, but they are also available at much higher spatial resolutions ($9^2$ $km^2$, found in the enhanced SMAP SM product[10]), and can be scaled to even higher resolutions to meet needs of the transportation community. Furthermore, $T_{eff}$ being specifically matched for L-band radiometry for soil moisture retrievals, GEOS-5 soil temperatures computed at other depths may be even more relevant and accurate for transportation applications. Based on the substantially improved results we reported earlier, it may be possible that $T_{eff}$ is even be more accurate for determining FT than SMAP at roads (or natural sites).

Overall, $T_{eff}$ values accurately track the variations of the road temperature (Fig. 15). For both sites, the onset of frozen conditions lags slightly behind those shown by the TDPs, but the timing of thawed conditions occurs at about the same time. However, $T_{eff}$ is substantially warmer (often by about 10 °C) than the TDPs. This may be due to $T_{eff}$ being modelled according to natural soils and for somewhat deeper depths than what these TDPs measure. According to results at these two sites, $T_{eff}$ appears to be well-suited for binary FT classifications and accurately identifying thaw dates.

Comparisons of $T_{eff}$ AM (Fig. 16) to NASA FT (Figs. 10–13) show that $T_{eff}$ performs substantially better against TDPs than NASA FT. Overall accuracy improved nearly everywhere, especially in the Northeast and West. This appears to be mainly due to improved detections of frozen conditions at those locations. Frozen conditions are also more accurately identified, except for a few locations, for example

[9] Available at https://nsidc.org/data/SPL3FTP/versions/2

[10] SMAP Enhanced SM product, available at https://doi.org/10.5067/7KKNQ5UURM2W

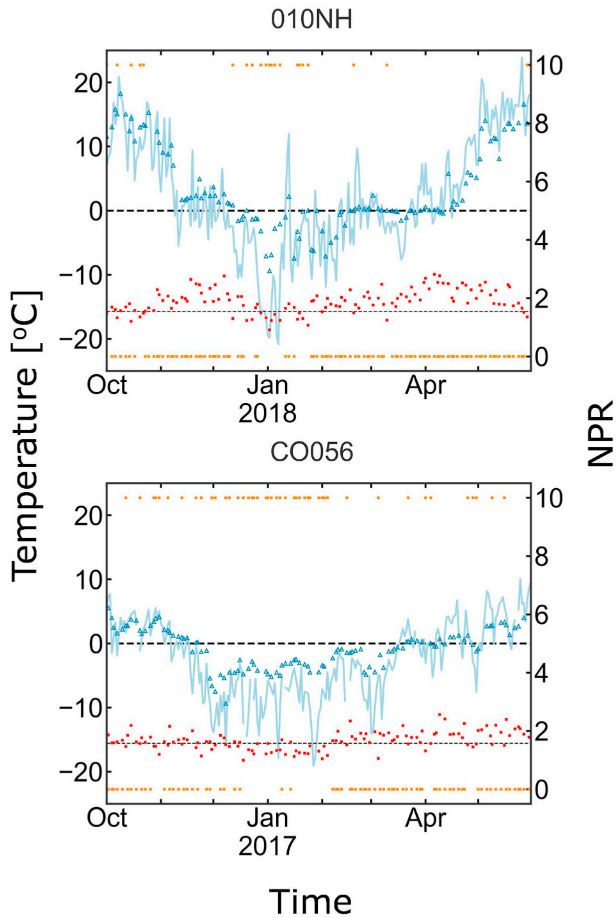


**Fig. 15.** Same as Fig. 6, but only showing results for NASA FT and with the $T_{eff}$ time series plotted as triangles and also using a lighter shade for the line showing the TDP data.

North Dakota and western Kansas. While $T_{eff}$ retrievals suffer from comparable false detections rates, they are found in different locations: $T_{eff}$ has relatively more false detections in the West, whereas NASA FT has relatively more false detections in the south of the FT-R domain, especially in Kansas. Frozen conditions are also far more rarely missed when $T_{eff}$ is used over NASA FT, especially in the West.

FT-R retrievals improved over NASA FT corresponding to the above, showing the important role $T_{eff}$ plays for improving FT retrievals. But as FT-R is informed by both SMAP retrievals and model results, there are some notable differences. Frozen conditions were detected somewhat more accurately by $T_{eff}$ in the West and Northeast, but FT-R performed somewhat better in several states in the Midwest. Overall Accuracy for $T_{eff}$ is also slightly better than FT-R in those regions, but elsewhere, it is more difficult to discern any significant improvements from visual comparisons. Occurrences of False Freeze are rare for $T_{eff}$, only occurring at rates of about ~20% or lower in the West and Northeast, whereas comparably low rates are found in FT-R but more evenly distributed throughout CONUS except for a few states such as Iowa. The occurrence and distribution of missed frozen conditions are nearly the same for $T_{eff}$-based retrievals and FT-R, with occasional differences. For example, FT-R has fewer missed frozen conditions in Minnesota and North Dakota, whereas it had relatively more in the Northeast.

$T_{eff}$ performed 5 and 15% better with respect to FT-R and NASA FT in CONUS (Table 1). It also performed better in Alaska, exceeding overall accuracies for FT-R and NASA FT by 5 and 12%. Since the model performed substantially better for road FT than NASA FT, it is reasonable that FT-R accuracy metrics fall somewhere in-between those reported for NASA FT and $T_{eff}$ (Table 2).

At least for the AM comparisons and criteria examined here, there seem to be few advantages in using NASA FT or FT-R over $T_{eff}$ for determining road FT states. Although $T_{eff}$ would be a better choice in nearly all cases, there are rare cases for which FT-R may have a benefit. For example, we noted that FT-R on occasionally outperformed $T_{eff}$ such as in North Dakota and was that its performance was comparable in many other states. Additional differences may be revealed in more exhaustive comparisons such as for other regions or metrics (e.g. identifying timing of road freeze and thaw) are made. Overall, FT-R performance is still impressive given that $T_{eff}$ values were only used once for informing new $NPR_{thr}$ values and that FT classifications only depended on SMAP NPR time series. Going forward, it would also be valuable to examine how FT-R performs in years that were not used for calibrating $NPR_{thr}$.

## 5. Conclusions

The major contribution of this work was to introduce a new approach to improve SMAP FT retrievals for use in CONUS roads, referred to as FT-R. The method is straightforward and can be implemented using data that are already part of the SMAP soil moisture product. Strengths and weaknesses of NASA FT and FT-R were characterized using a number of different analyses and in situ data from approximately 1000 RWIS sites. Overall, the standard NASA FT product agreed well with the RWIS observations showing an accuracy of 69% for AM and 80% for PM comparisons, but it suffered from persistent and nearly universal under-detections of in situ frozen conditions. While NASA FT performed substantially worse than FT-R for AM comparisons, its accuracy was only 2% less than FT-R for PM comparisons. Performance comparisons showed that FT-R performed better than NASA FT against road data collected in CONUS and Alaska. Median RT values for FT-R thawed conditions were relatively greater than for NASA FT, resulting in a better separation between frozen and thawed roads. We also showed that FT-R improved retrievals of frozen conditions in CONUS by 50–75% over NASA FT, and that this was the predominant reason explaining why Overall Accuracy improved. The fact that we were able to achieve such substantial improvements with the present method suggests that there are further opportunities for improving SMAP FT retrievals in CONUS.

SMAP FT performance was also tabulated by state and geographic region. Results showed that SMAP FT retrievals varied much more between regions (9%) than among states (4%) located in the same region. The state-level performance metrics should be of value to DOTs for its potential inclusion in the decision-making process at state level.

While we were not able to identify any particular advantages of using the standard NASA FT product instead of FT-R, we expect NASA FT will improve if $NPR_{thr}$ values (or equivalently, the seasonal threshold values) are further adjusted as planned in future releases of the product. The value of this work is that the introduced approach allows for easy and automatic retrievals of improved thresholds at each SMAP grid. Also, the threshold values determined in this study could be directly applied to the NASA FT product by adjusting the seasonal threshold values at each grid accordingly.

It is important to note that while this study mainly concerned itself with making comparisons against road temperatures, we used effective soil temperatures that were already included in the SMAP soil moisture product. Therefore, FT-R retrievals will resemble modelled soil temperature values more closely than NASA FT. FT-R might also perform relatively better against natural site data, and further studies are needed to determine if the approach presented here would also improve FT correspondence with natural systems.

It was also interesting to note that even when $T_{eff}$ – a modelled quantity estimating the temperature of natural soils rather than roads - was used to calibrate FT-R, FT-R performance also substantially improved with respect to roads. We also found that $T_{eff}$ values compared even better than NASA FT and FT-R to roads. However, $T_{eff}$ is an

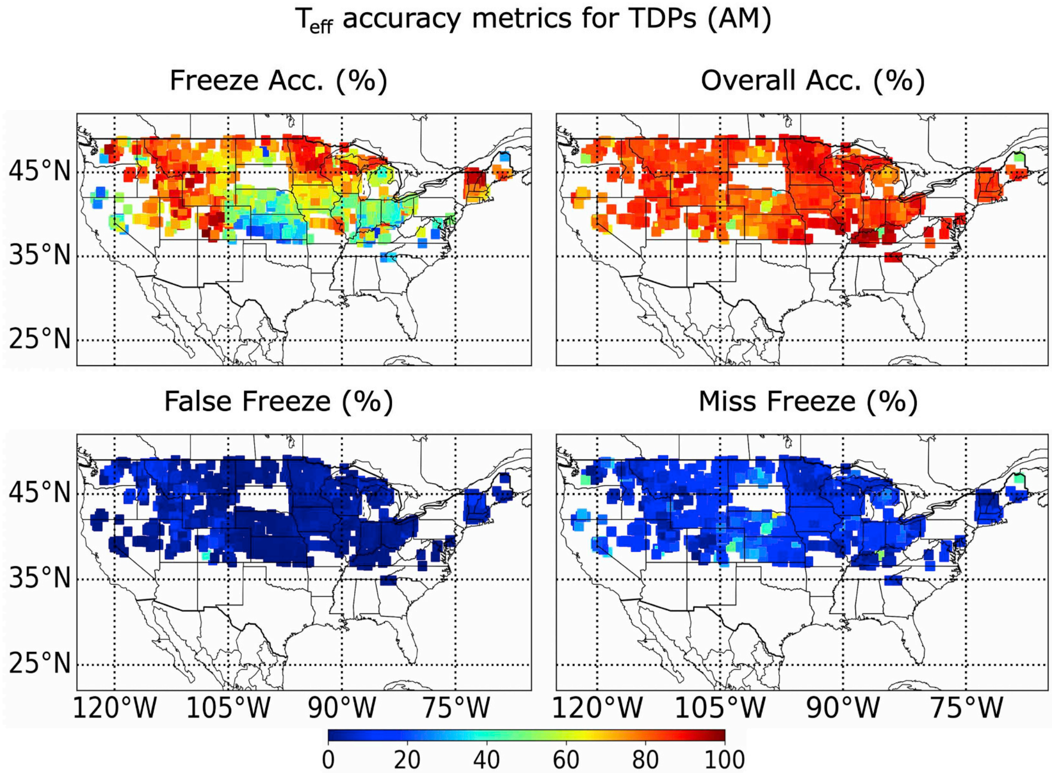


**Fig. 16.** Accuracy metrics for morning comparisons (AM) directly using modelled soil temperatures ($T_{eff}$) instead of SMAP FT retrievals.

**Table 2**
Summary of accuracy metrics in percent by region for morning comparisons between NASA FT, FT-R, $T_{eff}$ and road temperatures data. MF: missed freeze, the rate of SMAP identifying in situ frozen conditions as thawed; FF: false freeze, the rate of SMAP identifying in situ thawed conditions as frozen; FA: freeze accuracy, the proportion of all in situ frozen conditions detected; TA: thaw accuracy, the proportion of all in situ thawed conditions detected; Overall: overall accuracy, the proportion of correct retrievals out of all retrievals.

| | | FA (%) | TA (%) | FF (%) | MF (%) | Overall |
|---|---|---|---|---|---|---|
| CONUS | NASA FT | 36 | 90 | 6 | 25 | 69 |
| | FT-R | 59 | 90 | 5 | 16 | 79 |
| | $T_{eff}$ | 66 | 93 | 4 | 13 | 84 |
| Alaska | NASA FT | 66 | 91 | 5 | 22 | 74 |
| | FT-R | 81 | 82 | 8 | 12 | 81 |
| | $T_{eff}$ | 93 | 71 | 11 | 3 | 86 |

ancillary dataset and is needed for retrieving soil moisture, and is adjusted as needed for improving SMAP SM retrievals. For example, we've shown that version 5 $T_{eff}$ values were substantially warmer over western states and neither FT-R nor $T_{eff}$ would likely perform well there had version 5 values been used.

There are many potential sources of error impacting the performance of SMAP FT retrievals, especially when making comparisons with roads that usually make up a negligible portion of the land cover found in a SMAP footprint. Despite the mismatched scales and many sources of error, the 39 × 47 km SMAP FT retrievals are surprisingly accurate with respect to highly localized temperature measurements in roads. While never frozen/never thawed masks can be very important for improving accuracies, SMAP FT retrievals were usually also accurate during the time periods where neither of the masks applied.

Work in this field would further benefit from developing an improved understanding of the mechanisms and/or teleconnections that allow for these spaceborne FT retrievals to be so strongly linked to local measurements in roads and to explore whether FT retrievals perform similarly well for other elements of the built environment.

## Author responsibilities

S·K., J.J. and R.S. designed the study. S·K. obtained and processed the datasets. S.K., J.J., E.C., analyzed and interpreted the results. S.K wrote the manuscript with support of J.J., R.S., E.C., H.M. and C.V.

## Declaration of competing interest

None.

## Acknowledgements

This work was supported by NASA Research Opportunities in Earth and Space Sciences (ROSES), NNX16AN34G. We would also like to thank Dr. Jo Sias and Dr. Eshan V. Dave for helpful discussions regarding seasonal frost in roads. We would also like to thank the three anonymous reviewers who also provided helpful comments and suggestions for improving the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2019.111545.

## References


Alemohammad, S.H., Kolassa, J., Prigent, C., Aires, F., Gentine, P., 2018. Global downscaling of remotely sensed soil moisture using neural networks. Hydrol. Earth Syst. Sci. 22, 5341–5356.

Andersland, O.B., Ladanyi, B., 2004. Frozen Ground Engineering. John Wiley & Sons.

Brodzik, M., Billingsley, B., Haran, T., Raup, B., Savoie, M., 2014. Correction: Brodzik, M.J., et al. EASE-Grid 2.0: incremental but significant improvements for earth-gridded data sets. ISPRS International Journal of Geo-Information 2012, 1, 32–45. ISPRS Int. J. Geo-Information 3, 1154–1156.

Brodzik, M.J., Billingsley, B., Haran, T., Raup, B., Savoie, M.H., 2012. EASE-Grid 2.0: incremental but significant improvements for earth-gridded data sets. ISPRS Int. J. Geo-Information 1, 32–45.

Chan, S., Njoku, E.G., Colliander, A., 2018. SMAP L1C Radiometer Half-orbit 36 km

EASE-Grid Brightness Temperatures, Version 4. NASA Natl. Snow Ice Data Cent. Distrib. Act. Arch. Cent, Boulder, Colorado USA. https://doi.org/10.5067/ZVILG0PS6CTI. WWW Document. (Accessed 7.4.19).

Chan, S.K., Bindlish, R., O'Neill, P.E., Njoku, E., Jachson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Yueh, S., 2016. Assessment of the SMAP passive soil moisture product. Transactions on Geoscience and Remote Sensing 54 (8), 4994–5007.

Choudhury, B.J., Schmugge, T.J., Mo, T., 1982. A parameterization of effective soil temperature for microwave emission. J. Geophys. Res. Ocean. 87, 1301–1304.

Derksen, C., Xu, X., Scott Dunbar, R., Colliander, A., Kim, Y., Kimball, J.S., Black, T.A., Euskirchen, E., Langlois, A., Loranty, M.M., Marsh, P., Rautiainen, K., Roy, A., Royer, A., Stephens, J., 2017. Retrieving landscape freeze/thaw state from Soil Moisture Active Passive (SMAP) radar and radiometer measurements. Remote Sens. Environ. 194, 48–62. https://doi.org/10.1016/j.rse.2017.03.007.

Dunbar, S., Xu, X., Colliander, A., Derksen, C., Kimball, J., Kim, Y., 2018. SPL3FTP/_E Algorithm Theoretical Basis Document [2018] [WWW Document]. https://nsidc.org/sites/nsidc.org/files/technical-references/L3_FT_P_ATBD_RevisionB.docx (accessed 7.4.19).

Eftekhari, S., Miller, H., Kraatz, S., Jacobs, J., 2018. Use of a pavement surface-temperature prediction model for SLR application (No. 18-01481). In: 97th Annu. Meet. Transp. Res. Board.

Entekhabi, D., Yueh, S., O'Neill, P.E., Kellogg, K.H., Allen, A., Bindlish, R., Brown, M., Chan, S., Colliander, A., Crow, W.T., 2014. SMAP Handbook–soil Moisture Active Passive: Mapping Soil Moisture and Freeze/Thaw from Space.

Jaselskis, E.J., Grigas, J., Brilingas, A., 2003. Dielectric properties of asphalt pavement. J. Mater. Civ. Eng. 15, 427–434.

Johnston, J., Maggioni, V., Houser, P., 2019. Investigating the relationship between satellite-based freeze/thaw products and land surface temperature. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12 (9), 3247–3271.

Judge, J., Member, S., Galantowicz, J.F., England, A.W., Dahl, P., 1997. Freeze/thaw classification for prairie soils using SSM/I radiobrightnesses. IEEE Trans. Geosci. Remote Sens. 35, 827–832.

Kalantari, P., Bernier, M., McDonal, K.C., Poulin, J., 2014. Using available time series of passive and active microwave to develop SMAP freeze/thaw algorithms adapted for the canadian subarctic. In: Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International. IEEE, pp. 2550–2553.

Kerr, Y.H., Waldteufel, P., Wigneron, J.P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.J., Font, J., Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martin-Neira, M., Mecklenburg, S., 2010. The SMOS L: new tool for monitoring key elements ofthe global water cycle. Proc. IEEE 98, 666–687. https://doi.org/10.1109/JPROC.2010.2043032.

Kestler, M., Berg, R., Bigl, S., 2011. Statistical significance of freeze-thaw related factors on cumulative damage to flexible pavements. Transp. Res. Rec. J. Transp. Res. Board 55–67.

Kestler, M.A., Berg, R.L., Steinert, B.C., Hanek, G.L., Truebe, M.A., Humphrey, D.N., 2007. Determining when to place and remove spring load restrictions on low-volume roads: three low-cost techniques. Transp. Res. Rec. 1989 (1), 219–229.

Kim, H., Parinussa, R., Konings, A.G., Wagner, W., Cosh, M.H., Lakshmi, V., Zohaib, M., Choi, M., 2018. Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products. Remote Sens. Environ. 204, 260–275. https://doi.org/10.1016/j.rse.2017.10.026.

Kim, Y., Kimball, J.S., Mcdonald, K.C., Glassy, J., 2011. Developing a global data record of daily landscape freeze/thaw status using satellite passive microwave remote sensing. IEEE Trans. Geosci. Remote Sens. 49, 949–960.

Kim, Y., Kimball, J.S., Glassy, J., Du, J., 2017. An extended global earth system data record on daily landscape freeze–thaw status determined from satellite passive microwave remote sensing. Earth Syst. Sci. Data 9, 133–147.

Kim, Y., Kimball, J.S., Xu, X., Dunbar, R.S., Colliander, A., Derksen, C., 2019. Global assessment of the SMAP freeze/thaw data record and regional applications for detecting spring onset and frost events. Remote Sens. 11, 1317.

Kraatz, S., Jacobs, J.M., Miller, H., Daniel, J., 2017. A comparison of freeze-thaw in roads with passive microwave satellite observations from SMAP. In: in: AGU Fall Meeting Abstracts.

Kraatz, S., Jacobs, J., Schröder, R., Cho, E., Cosh, M., Seyfried, M., Prueger, J., Livingston, S., 2018. Evaluation of SMAP freeze/thaw retrieval accuracy at core validation sites in the contiguous United States. Remote Sens. 10, 1483.

Kraatz, S., Jacobs, J.M., Miller, H.J., 2019a. Spatial and temporal freeze-thaw variations in Alaskan roads. Cold Reg. Sci. Technol. 157, 149–162. https://doi.org/10.1016/j.coldregions.2018.10.006.

Kraatz, S., Miller, H.J., Jacobs, J.M., Dave, E.V., Sias, J., 2019b. Accuracy Assessment of Satellite-Based Freeze-Thaw Retrievals on Low-Volume Roads in the United States. Transp. Res. Rec. https://doi.org/10.1177/0361198119854093.

Kraatz, S., Miller, H.J., Jacobs, J.M., 2019c. Remotely sensed freeze-thaw from the soil moisture active passive instrument to inform the timing of seasonal load restrictions in Alaska. Transp. Res. Rec. https://doi.org/10.1177/0361198119834558.

Kraatz, S., Miller, H., Poirier, B., Moradi, M., Jacobs, J., 2020. Potential use of NASA's SMAP freeze-thaw tool to assist in seasonal load restriction timing in the northern United States. Transp. Res. Board Annu. Meet. 2020.

Lemmetyinen, J., Schwank, M., Rautiainen, K., Kontu, A., Parkkinen, T., Mätzler, C., Wiesmann, A., Wegmüller, U., Derksen, C., Toose, P., Roy, A., Pulliainen, J., 2016. Snow density and ground permittivity retrieved from L-band radiometry: application to experimental data. Remote Sens. Environ. 180, 377–391. https://doi.org/10.1016/j.rse.2016.02.002.

Mahoney, J.P., Rutherford, M.S., Hicks, R.G., 1986. Guidelines for spring highway use restrictions. In: Summary Report. No. WA-RD-80.2.

Mätzler, C., 2006. Thermal microwave radiation: applications for remote sensing. Iet.

MnDOT, 2014. Process for seasonal load limit starting and ending dates. In: Technical Memorandum No. 14-10-MAT-02. Minnesota Department of Transportation (MnDOT).

O'Neill, P., Chan, S., Njoku, E.G., Jackson, T., Bindlish, R., 2016a. SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 4 [WWW Document]. NASA Natl. Snow Ice Data Cent. Distrib. Act. Arch. Cent, Boulder, Color. USA. https://doi.org/10.5067/OBBHQ5W22HME.

O'Neill, P., Chan, S., Njoku, E.G., Jackson, T.J., Bindlish, R., 2016b. SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 4. NASA Natl. Snow Ice Data Cent. Distrib. Act. Arch. Cent, Boulder, Colorado USA. https://doi.org/10.5067/OBBHQ5W22HME. WWW Document. (accessed 7.4.19).

O'Neill, P., Bindlish, R., Chan, S., Njoku, E., Jackson, T., 2018a. Soil Moisture Active Passive (SMAP) Algorithm Theoretical Basis Document SMAP L2 & L3 Radar Soil Moisture (Active) Data Products, Revision D. pp. 1–82.

O'Neill, P., Chan, S., Njoku, E.G., Jackson, T.J., Bindlish, R., 2018b. SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 5. NASA Natl. Snow Ice Data Cent. Distrib. Act. Arch. Cent, Boulder, Colorado USA. https://doi.org/10.5067/ZX7YX2Y2LHEB. WWW Document. (accessed 7.4.19).

Peng, J., Loew, A., Merlin, O., Verhoest, N.E.C., 2017. A review of spatial downscaling of satellite remotely sensed soil moisture. Rev. Geophys. 55, 341–366.

Rowlandson, T.L., Berg, A.A., Roy, A., Kim, E., Lara, R.P., Powers, J., Lewis, K., Houser, P., McDonald, K., Toose, P., 2018. Capturing agricultural soil freeze/thaw state through remote sensing and ground observations: a soil freeze/thaw validation campaign. Remote Sens. Environ. 211, 59–70.

Schroeder, R., McDonald, K.C., Chapman, B.D., Jensen, K., Podest, E., Tessler, Z.D., Bohn, T.J., Zimmermann, R., 2015. Development and evaluation of a multi-year fractional surface water data set derived from active/passive microwave remote sensing data. Remote Sens. 7, 16688–16732.

Schwank, M., Rautiainen, K., Mätzler, C., Stähli, M., Lemmetyinen, J., Pulliainen, J., Vehviläinen, J., Kontu, A., Ikonen, J., Ménard, C.B., 2014. Model for microwave emission of a snow-covered ground with focus on L band. Remote Sens. Environ. 154, 180–191.

Schwank, M., Mätzler, C., Wiesmann, A., Wegmüller, U., Pulliainen, J., Lemmetyinen, J., Rautiainen, K., Derksen, C., Toose, P., Drusch, M., 2015. Snow density and ground permittivity retrieved from L-band radiometry: a synthetic analysis. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 3833–3845.

Simonsen, E., Isacsson, U., 1999. Thaw weakening of pavement structures in cold regions. Cold Reg. Sci. Technol. 29, 135–151.

SMAP, 2015. Ancillary Data Report Surface Temperature [WWW Document].

Van Deusen, D., Schrader, C., Bullock, D., Worel, B., 1998. Springtime thaw weakening and load restrictions in Minnesota. Transp. Res. Rec. J. Transp. Res. Board 21–28.

Walker, V.A., Hornbuckle, B.K., Cosh, M.H., Prueger, J.H., 2019. Seasonal evaluation of SMAP soil moisture in the U.S. Corn Belt. Remote Sens. 11, 2488.

Xu, X., Dunbar, R.S., Derksen, C., Colliander, A., Kim, Y., Kimball, S., 2018. SMAP L3 radiometer northern hemisphere daily 36 km EASE-grid freeze/thaw state, version 2 [WWW Document]. NASA Natl. Snow Ice Data Cent. Distrib. Act. Arch. Cent. https://doi.org/10.5067/YN94K53QM061.

Zhang, T., 2003. Investigation of the near-surface soil freeze-thaw cycle in the contiguous United States: algorithm development and validation. J. Geophys. Res. 108, 8860. https://doi.org/10.1029/2003JD003530.

Zheng, D., Li, X., Wang, X., Wang, Z., Wen, J., van der Velde, R., Schwank, M., Su, Z., 2019. Sampling depth of L-band radiometer measurements of soil moisture and freeze-thaw dynamics on the Tibetan Plateau. Remote Sens. Environ. 226, 16–25.