Northern Hemisphere Surface Freeze/Thaw Product from Aquarius L-band Radiometers

Michael Prince¹,², Alexandre Roy³,²,¹, Ludovic Brucker⁴,⁵, Alain Royer¹,², Youngwook Kim⁶, Tianjie Zhao⁷

¹ Centre d’Applications et de Recherches en Télédétection (CARTEL), Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada
² Centre d’Étude Nordique, Université Laval, Québec, Canada
³ Université de Montréal, Département de Géographie, Montréal, QC, H2B 2V8, Canada
⁴ NASA Goddard Space Flight Center, Cryospheric Sciences Laboratory, Code 615, Greenbelt, MD 20771, USA
⁵ Universities Space Research Association, Goddard Earth Sciences Technology and Research Studies and investigations, Columbia, MD 21044, USA
⁶ Numerical Terradynamic Simulation Group, College of Forestry & Conservation, The University of Montana, Missoula, MT 59812, USA
⁷ State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China

Correspondence to: Michael Prince (michael.prince@usherbrooke.ca)

Abstract. In the Northern Hemisphere, seasonal changes in surface freeze/thaw (FT) cycle are an important component of surface energy, hydrological and eco-biogeochemical processes that must be accurately monitored. This paper presents the weekly polar-gridded Aquarius passive L-Band surface freeze/thaw product (FT-AP) distributed on the Equal-Area Scalable Earth Grid version 2.0, above the parallel 50° N, with a spatial resolution of 36 km x 36 km. The FT-AP classification algorithm is based on a seasonal threshold approach using the normalized polarization ration, references for frozen and thawed conditions and optimized thresholds. To evaluate the uncertainties of the product, we compared it with another satellite FT product also derived from passive microwave observations but at higher frequency: the resampled 37 GHz FT Earth Science Data Record (FT-ESDR). The assessment was carried out during the overlapping period between 2011 and 2014. Results show that 77.1% of their common grid cells have an agreement better than 80%. Their differences vary with land cover type (tundra, forest and open land) and freezing and thawing periods. The best agreement is obtained during the thawing transition and over forest areas, with differences between product mean freeze or thaw onsets of under 0.4 weeks. Over tundra, FT-AP tends to detect freeze onset 2–5 weeks earlier than FT-ESDR, likely due to FT sensitivity to the different frequencies used. Analysis with mean surface air temperature time series from six in situ meteorological stations shows that the main discrepancies between FT-AP and FT-ESDR are related to false frozen retrievals in summer for some regions with FT-AP. The Aquarius product is distributed by the U.S. National Snow and Ice Data Center (NSIDC) at https://nsidc.org/data/nsidc-0736/versions/1 with the DOI 10.5067/OV4R18NL3BQR.
1 Introduction

Seasonal freezing and thawing affect over half of the Northern Hemisphere. Landscape freeze/thaw (FT) state transitions show highly variable spatial and temporal patterns, with measurable influences to climate (IPCC, 2014; Peng et al., 2017; Poutou et al., 2004), hydrological (Gouttevin et al., 2012; Gray et al., 1985), ecological (Kumar et al., 2013; Black et al., 2000) and biogeochemical processes (Selvann et al., 2016; Xu et al., 2013; Schaefer et al., 2011). The surface FT state affects the latent heat exchange and the energy balance at the interface between soil surface and the overlying medium. The vegetation growing season is sensitive to the annual non-frozen period (Kim et al., 2012), while vegetation net primary production and net ecosystem CO₂ exchange (NEE) with the atmosphere is impacted by FT timing variability (Barr et al., 2009; Kurganova et al., 2007). Comprehensive in situ observational long-term data sets for soil state characteristics across terrestrial environments are still limited or inadequate, mostly for northern remote regions. Remote sensing in the thermal emission domain offers great potential for detecting changes in land-surface temperature, but is strongly limited by clouds, vegetation and snow cover (e.g. Langer et al., 2013). Spatially and temporally continuous information on soil freeze/thaw changes is lacking for the regions of both seasonal frozen ground and permafrost.

Passive microwave remote sensing has been proven sensitive to the surface FT state due to large changes in surface dielectric properties between predominantly frozen and non-frozen conditions, and it offers global coverage. The remotely sensed FT detection capability at L-band (1.4 GHz) has been developed and validated in several studies (Zheng et al., 2017; Roy et al., 2017b; Rautiainen et al., 2012; Schwank et al., 2004). At L-band, the shallow depth contributing to the radiation (around 5 cm for an unfrozen soil) and the strong permittivity difference between water and ice (Δε_water) make it favourable for FT retrieval (Rautiainen et al., 2012; 2014). In recent years, passive L-band FT algorithms were created for NASA’s Aquarius (Roy et al., 2015), ESA’s soil moisture and ocean salinity (SMOS) (Rautiainen et al., 2016), and NASA’s soil moisture active/passive (SMAP) (Derksen et al., 2017) missions. An FT Earth Science Data Record (FT-ESDR) was also produced using a higher microwave frequency at Ka-band (37 GHz) (Kim et al., 2017a). This product offers consistent and continuous global daily information on the FT state for several decades (1979-2016; Kim et al., 2017b). Observations were recorded by the scanning multi-channel microwave radiometer (SMMR), the special sensor microwave/imager (SSM/I), and the SSM/I Sounder (SSMIS).

This study presents the new Aquarius passive FT product for the Northern Hemisphere, distributed by the US National Snow and Ice Data Center (NSIDC) at nsidc.org/data/nsidc-0736/versions/1. The product precision and uncertainties are addressed by comparing Aquarius FT retrievals with the FT-ESDR product for the overlapping period (2011 - 2014). The Aquarius passive FT product (referred to as FT-AP hereinafter) is based on the Aquarius weekly Level-3 L-Band brightness temperature (TB) product (Brucker et al., 2015; NSIDC: http://nsidc.org/data/AQ3_TB/versions/5). The algorithm uses a relative frost factor (FFrel; see e.g. Rautiainen et al., 2014) based on normalized polarization ratio (NPR) temporal change detection (Roy...
et al., 2015). To our knowledge, few intercomparisons between L- and Ka-band FT products exist (Derksen et al., 2017), and none evaluated inter-annual variability differences. However, it is well established that different frequencies interact differently with ground components (vegetation, soil, snow), canopy, etc. For instance, observations at L-band are less sensitive than at Ka-band to snow, plant biomass and surface roughness (Ulaby et al., 1986). Being less prone to disturbances above the ground, the L-band emission should give better information on the ground state in forested and snow-covered areas. In addition, since $\Delta \varepsilon_{\text{ice/water}}$ is larger at L-band ($\Delta \varepsilon_{\text{ice/water}} \approx 83$) than at Ka-band ($\Delta \varepsilon_{\text{ice/water}} \approx 10$) (Artemov and Volkov, 2014), there should be a higher sensitivity to the ground phase transition at L-band. Hence, because differences between products can be attributed to the microwave frequency and the algorithm used, the FT-AP is also compared with surface air temperature (SAT) observations.

The main objective of this study is to present and evaluate the weekly FT-AP by comparing it to the FT-ESDR and to SAT observations across the Northern Hemisphere. First, we describe the new FT-AP product, designed by the algorithm developed by Roy et al. (2015), but applied across the Northern hemisphere. Then, we investigate the spatial and temporal FT variations from both FT-AP and FT-ESDR products over the Northern hemisphere. We then investigate the cause of the main differences between products from in situ information. The comparison aims to identify the similarities and differences between L-band and Ka-Band FT products for further improvements of FT monitoring across the Northern hemisphere.

2. Method

2.1 Aquarius passive FT product (FT-AP)

The Aquarius FT product was generated using the Aquarius weekly averaged polar gridded L-band TB product distributed on the EASE-Grid 2.0, above the parallel 50° N, with a spatial resolution of 36 km x 36 km (Brucker et al., 2014). This formatted TB was specially designed for the study of northern regions. The FT classification algorithm is based on a seasonal threshold approach (STA) using frost factor index (FFrel) (Eq. 1), introduced by Rautiainen et al. (2014), where $FF_{\text{NPR}}$ is the frost factor based on the normalized polarization ratio between TB at vertical and horizontal polarizations ($TB_V$ and $TB_H$; Eq. 2). $FF_{\text{fr}}$ and $FF_{\text{th}}$ are reference frozen and thawed frost factors obtained respectively by averaging the five minimum and maximum $FF_{\text{NPR}}$ values of 2012 and 2013.

$$FF_{\text{rel}} = \frac{FF_{\text{NPR}} - FF_{\text{fr}}}{FF_{\text{th}} - FF_{\text{fr}}}$$

$$FF_{\text{NPR}} = \frac{TB_V - TB_H}{TB_V + TB_H}$$

A threshold ($\tau$) was determined by optimization to classify the surface as frozen or thawed if the FFrel is lower or higher than the threshold (Eq. 3).
If $FF_{rel} < \tau \rightarrow \text{freeze}$
or if $FF_{rel} > \tau \rightarrow \text{thaw}$                            (3)

The thresholds optimized in Roy et al. (2015) over Northern America for three basic land covers (tundra, forest, open land) were applied over the Northern Hemisphere using the Land Cover Classifications Derived from Boston University MODIS / Terra Land Cover Data (LCC$_{BU}$) (see Sect. 2.4).

Aquarius operated three non-scanning radiometers at different incidence angles (29.2°, 38.4° and 46.3°) and with different 3 dB footprint sizes (respectively 76 km x 94 km, 84 km x 120 km and 97 km x 156 km). Based on the LCC$_{BU}$, the thresholds found in Roy et al., (2015) were used to create FT maps for each radiometer. The three FT maps were then blended to create a fourth map, which offers more complete spatial coverage. For every grid cell, radiometer 2 (38.4°) is prioritized, while radiometer 3 is only used if data from the other radiometers is not available for the given grid cell. This blended algorithm was chosen based on the performance given for each radiometer in Roy et al. (2015) (radiometer 2 gave the best results, while radiometer 3 gave the worst results). Due to the width of Aquarius’ swath and its revisit time, 16.5% of the terrestrial 36-km grid cells have less than 95% observations over the period and 16% were not measured at all. Thus, the intercomparison with the FT-ESDR product (Sect. 2.2) was only made when FT-AP data were available for a given date. The time span for this analysis runs from August 2011 with the first Aquarius observations to 31 December 2014 with the latest FT-ESDR data available at the time of our analysis.

### 2.2 FT-ESDR product

The first version of the FT-ESDR product (Kim et al., 2011) was based on an STA similar to the FF$_{rel}$ but applied exclusively to the TB$_{V}$ at 37 GHz instead of the NPR. In the new extended product (Kim et al., 2017b; NSIDC: https://nsidc.org/data/nsidc-0477/versions/4), a modified seasonal threshold algorithm (MSTA) was used to determine thresholds for each grid cell to obtain better accuracy. It consists of a grid-cell-wise weighted empirical linear regression relationship between the 37 GHz TB$_{V}$ measurements and daily surface air temperature estimates from the ERA-Interim global reanalysis.

The extended FT-ESDR product used in this study is derived from the SSM/I 37 GHz brightness temperatures (footprint of 38 km x 30 km) and resampled at a grid cell resolution of 25 km on the global EASE-Grid v1.0. The observations were recorded twice per day, which gives the possibility of attributing discrete frozen or thawed states for morning and afternoon. The final classification offers four discrete surface states: “frozen all day”, “thawed all day”, “frozen in AM and thawed in PM” (transitional) and “thawed in AM and frozen in PM” (inverse-transitional). In this study, the latter two classes were combined into a single transitional class. In order to compare the two products, the FT-ESDR was first spatially resampled to the EASE-Grid 2.0 with the nearest neighbor method choosing the smallest distance between pixel centers. Then, FT-ESDR was temporally resampled based on the rule that the most frequently occurring class over the seven days of a week is adopted as
the value for the entire week. In cases where the frozen and thawed classes occurred with equal frequency during a single week (e.g., two days frozen, two days thawed and three days transitional), the transitional class was attributed. This latter class occurs mainly during the transition seasons of spring and fall. Thus, we assigned the transitional class to thawed class during spring and summer since it indicates the beginning of the thawing process and we assigned the transitional class to the frozen class during fall and winter since it indicates the beginning of the freezing process. This FT-ESDR resampling procedure ensured that the two products were at the same temporal and spatial resolutions with only the frozen and thawed categories, making comparison possible.

2.3 Land cover classification

The land cover information (Fig. 1) comes from the EASE-Grid 2.0 LCCBU (Brodzik and Knowles, 2011; NSIDC: nsidc.org/data/nsidc-0610/versions/1), using the same grid as the FT-AP product. The seventeen land cover classes were grouped to obtain four classes: tundra, forest, open land (savanna, cropland and grassland) and water (see Roy et al., 2015). Each grid cell was assigned its single most prominent class of land cover. All grid cells with more than 20% of water and ice indicated by the LCCBU were masked.
Figure 1: Land cover classes: tundra (blue), forest (green), open land (yellow) and water/ice mask (white). Red dots show weather station locations.

2.4 Weather stations

Six weather stations (Table 1) were selected for validation from the National Climatic Data Center (NCDC) Climate Data Online website (CDO, https://www.ncdc.noaa.gov/cdo-web/datasets). Two tundra, two forest and two open land sites were chosen for a comparison between the product classifications and the in situ SAT. All of the sites are more than 200 km from a coast, except the Kamchatka site; its distance of about 85 km from the sea may have an influence on the large L-band field of view. The average SAT for each day (TAVG$_{day}$) were used to create a time series for each site. For statistical purposes, the weekly resampling method used on the FT-ESDR product was also applied to the SAT daily values, using 0°C as the threshold between frozen and thawed states (TAVG$_{week}$) (see Roy and al., 2015).
Ruggedness values from a 30 arc-second resolution elevation map (Gruber 2012; University of Zurich: http://www.geo.uzh.ch/microsite/cryodata/pf_global/) were resampled to the EASE-Grid 2.0 with the drop in the bucket approach. In order to represent a ruggedness value at the Aquarius footprint scale, the mean value of a 3 x 3 grid cell window centered on each weather station pixel was calculated (Rug_Mean). To each value was attributed a class according to the Gruber (2012) classification.

### Table 1: Latitude, longitude and land cover of each weather station

<table>
<thead>
<tr>
<th>Country/Region</th>
<th>Land Cover</th>
<th>Lat.</th>
<th>Long.</th>
<th>Ruggedness</th>
<th>Rug_Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamchatka/Kljuchi</td>
<td>Tundra</td>
<td>56.3167</td>
<td>160.8331</td>
<td>undulating</td>
<td>hilly</td>
</tr>
<tr>
<td>Canada/Québec</td>
<td>Tundra</td>
<td>57.9167</td>
<td>-72.9833</td>
<td>undulating</td>
<td>undulating</td>
</tr>
<tr>
<td>US/Alaska</td>
<td>Forest</td>
<td>64.7761</td>
<td>-141.162</td>
<td>mountainous</td>
<td>mountainous</td>
</tr>
<tr>
<td>Russia/Siberia</td>
<td>Forest</td>
<td>63.7831</td>
<td>121.6166</td>
<td>flat</td>
<td>flat</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>Open land</td>
<td>53.2166</td>
<td>63.6166</td>
<td>flat</td>
<td>flat</td>
</tr>
<tr>
<td>Canada/Saskatchewan</td>
<td>Open land</td>
<td>50.2666</td>
<td>-107.733</td>
<td>undulating</td>
<td>undulating</td>
</tr>
</tbody>
</table>

### 3 Results

#### 3.1 Spatial FT analysis

Figure 2a shows the percentage of concordant classifications between the two products for the 3.7 year overlapping period. Overall, the results show that there is good agreement between the two products. In general, forest areas have a better percentage of concordance than other land covers. However, some regions show important discrepancies, especially along coastal margins and in mountainous and open areas (such as in northern Europe, Kazakhstan (and surroundings) and the Canadian Prairies). Those lower percentages correspond to regions for which difficulties to detect the FT were already noted in Roy et al. (2015) and Kim et al. (2017a) (see Sect. 4). Figure 2b shows that 77.1% of the common grid cells have more than 80% agreement. More specifically, 41.6% of the grid cells have more than 90% agreement over 3.7 years, with 10.0% of them having more than 95%. About 35.5% of the grid cells have an agreement between 80% and 90%; only 22.8% of the cells have an agreement lower than 80%.
Figure 2: (a) Map of the percentage agreement between FT-AP and FT-ESDR classification for the whole period studied and (b) derived frequency distribution of the mean percentage agreement over the whole study area (Lat. > 50° N).
3.2 Temporal analysis

An analysis was made to identify similarities and differences between the two products used for retrieving surface FT state during the freezing (fall) and thawing (spring) periods. For each land cover type, Figure 3 shows the time series of the fraction of land frozen (for all land at latitudes greater than 50° N). To reduce the effect of obvious false retrievals on the analysis and to focus on the differences primarily related to the physics of the measurements (i.e., L-band vs. Ka-band), only grid cells with an agreement percentage between FT-AP and FT-ESDR higher than 80% (from Fig. 2a) were considered. Light blue zones indicate periods for which the FT-ESDR transitional class is set to the frozen class (see Sect. 2.2).

Figure 3: Time series of percentage of frozen grid cells for FT-AP and FT-ESDR for the three land covers (tundra, forest and open lands).

---

Figure 3: Time series of percentage of frozen grid cells for FT-AP and FT-ESDR for the three land covers (tundra, forest and open lands).
Figure 3 gives information on temporal differences between the products. The difference between FT-AP and FT-ESDR in terms of the percentage of frozen grid cells for a given day ($\Delta%_{\text{frozen}}$) is greatest during the fall in tundra, at $10\text{–}27\%$. In forest, $\Delta%_{\text{frozen}}$ is much lower than in tundra, with differences of $0\text{–}12\%$. For these two land covers (tundra and forest), the agreement between the products varies by year. The horizontal shift between the curves indicates time delays ($\Delta t_{\text{time}}$) for the two products to reach the same percentage of frozen grid cells. In tundra, $\Delta t_{\text{time}}$ ranges from 1–3 weeks. In forest, $\Delta t_{\text{time}}$ is always less than one week. This result demonstrates an excellent overall consistency between the products. However, FT-AP shows the percentage of frozen land increasing every summer to a peak that is not perceived with FT-ESDR. In tundra, those maximum values vary between 17% (2014) and 28% (2013) and are lower in forest at 7% (2013) and 10% (2012). Even if some of those detections represent the real state of the surface, the FT-AP peaks may be mainly caused by false frozen detections, which were noticed in the SMAP product (Derksen et al., 2017). False frozen detections are identified in our analysis using observations from the weather stations (Fig. 6, Sect. 3.3). In open land, FT-AP retrievals tend to vary frequently by showing noticeable unexpected frozen retrievals in summer and thawed retrievals in winter (blue lines in Fig. 3). FT-ESDR shows almost no frozen regions in summer, but unfrozen regions in winter, evidence that the open land regions are at the southern limits of the freeze regions. This in turn makes retrieval more difficult due to the higher temporal variability in FT events in winter.

To spatially represent the information provided by $\Delta t_{\text{time}}$, maps in Fig. 4 indicate the week of the year of the freeze onset for each product (top and middle maps). The freeze onset is defined as the first week of the year when the state changed from thawed to frozen and stayed frozen for two more consecutive weeks. This variable can only be identified for grid cells that contain observations over several weeks in a row and have good agreement (>80%) according to Fig. 2a. Figure 4 also shows the difference in freeze onset between the two products (bottom maps), defined as FT-AP minus FT-ESDR. A negative value means that FT-AP detects the freeze onset earlier than FT-ESDR (represented by cold colors) and inversely for a positive value (represented by warm colors).

Comparing FT-AP and FT-ESDR maps shows a global tendency of FT-AP to reach the freeze onset 2–5 weeks earlier than FT-ESDR in the tundra regions (blue zones in Fig. 4). In 2013 and 2014 (Fig. 4c–d), this tendency is stronger, with more regions experiencing an earlier freeze onset by 3–5 weeks according to FT-AP. While these differences are less noticeable in the forest, some local discrepancies are observable with noticeable inter-annual variabilities.
Table 2 gives freeze onset means (μ) and standard deviations (σ) in weeks of the year for each land cover and year. Over tundra, it shows the greatest freeze onset mean difference (Δμ = μ_{FT-AP} - μ_{FT-ESDR}) between the two products in 2013, with Δμ = 2.4 weeks, and the smallest difference in 2011, with Δμ = 1.3 weeks. Over forest, the differences are much smaller; the greatest occurs in 2011, with Δμ = 0.7 weeks, and the smallest in 2012, with Δμ = 0.0 weeks. As noted for Fig. 4, FT-AP tends to detect freeze onset earlier than FT-ESDR. These freeze onset differences suggest that there is a divergence in the FT signal at L and Ka bands, and that there might be complementary information in the two signals (this is further addressed in the discussion).

<table>
<thead>
<tr>
<th>Land Cover</th>
<th>2011 μ</th>
<th>2012 σ</th>
<th>2013 μ</th>
<th>2014 σ</th>
<th>All years μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tundra</td>
<td>40.6</td>
<td>2.5</td>
<td>39.7</td>
<td>3.5</td>
<td>38.6</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>4.7</td>
<td>38.6</td>
<td>3.5</td>
<td>39.2</td>
</tr>
<tr>
<td>Δμ</td>
<td>-1.3</td>
<td>-1.4</td>
<td>-2.4</td>
<td>-2</td>
<td>-1.8</td>
</tr>
<tr>
<td>Forest</td>
<td>43.8</td>
<td>2.9</td>
<td>43.5</td>
<td>3.6</td>
<td>43.3</td>
</tr>
<tr>
<td></td>
<td>43.5</td>
<td>3.6</td>
<td>43.3</td>
<td>3.6</td>
<td>43.3</td>
</tr>
<tr>
<td>Δμ</td>
<td>-0.7</td>
<td>0</td>
<td>-0.7</td>
<td>0</td>
<td>0.3</td>
</tr>
<tr>
<td>Open lands</td>
<td>45.8</td>
<td>2.8</td>
<td>46.4</td>
<td>4.8</td>
<td>45.2</td>
</tr>
<tr>
<td></td>
<td>45.2</td>
<td>4.7</td>
<td>44.9</td>
<td>3.6</td>
<td>45.5</td>
</tr>
<tr>
<td>Δμ</td>
<td>-1.3</td>
<td>0.3</td>
<td>-3.4</td>
<td>0.3</td>
<td>-1</td>
</tr>
</tbody>
</table>

For the thawing period, differences between the products according to Fig. 3 and Table 3 are small for all land covers, meaning that globally the two products respond similarly to landscape thaw. This result is consistent across land covers and for the three spring seasons available for this analysis with a stronger variability for open lands. The sensitivity of passive microwave frequencies to the water present in the snow at the beginning of the thaw explains the similarity between the products in spring (Roy et al., 2017a; Hallikainen et al., 1986). Thaw onset maps created from the difference of thaw onset between the products (bottom maps), defined as FT-AP minus FT-ESDR, illustrate the consistency between products, but highlight some local differences (Fig. 5). More specifically, FT-ESDR tends to retrieve thaw earlier in North America, while FT-AP retrieves thaw earlier in Eurasia. These differences are more pronounced in 2014 (Fig. 5c).
Figure 5: Thaw onset maps, where colors indicate the week of year, for a) 2012, b) 2013 and c) 2013 with FT-AP (top), FT-ESDR (middle) and difference between the products (FT-AP minus FT-ESDR; bottom)

Table 3: Means (μ), standard deviation (σ) and means difference between products of thaw onset date (week of the year) for each land cover

<table>
<thead>
<tr>
<th></th>
<th>2011 μ</th>
<th>σ</th>
<th>2012 μ</th>
<th>σ</th>
<th>2013 μ</th>
<th>σ</th>
<th>All years μ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tundra</td>
<td>FT-AP</td>
<td>19.1</td>
<td>3</td>
<td>19.1</td>
<td>3.2</td>
<td>18.8</td>
<td>3.6</td>
<td>19.0</td>
</tr>
<tr>
<td></td>
<td>FT-ESDR</td>
<td>18.7</td>
<td>2.3</td>
<td>18.7</td>
<td>2.4</td>
<td>18.8</td>
<td>2.5</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td>Δμ</td>
<td>0.4</td>
<td>0.4</td>
<td>0</td>
<td>0.3</td>
<td>0</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>Forest</td>
<td>FT-AP</td>
<td>14.7</td>
<td>2.3</td>
<td>15.4</td>
<td>2</td>
<td>13.7</td>
<td>3</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>FT-ESDR</td>
<td>14.3</td>
<td>2</td>
<td>15.3</td>
<td>1.5</td>
<td>14</td>
<td>2.4</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>Δμ</td>
<td>0.4</td>
<td>0.1</td>
<td>-0.3</td>
<td>0.1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>Open lands</td>
<td>FT-AP</td>
<td>11.9</td>
<td>2.4</td>
<td>13.3</td>
<td>3</td>
<td>11.2</td>
<td>3.5</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>FT-ESDR</td>
<td>12.1</td>
<td>1.9</td>
<td>14</td>
<td>1.7</td>
<td>11.6</td>
<td>2.9</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>Δμ</td>
<td>-0.2</td>
<td>-0.7</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
</tbody>
</table>

3.3 Comparison with weather stations

In Sect. 3.1, it was shown that there were some regions where both products show significant discrepancies. In order to better assess the observed variabilities, we looked at six different sites (Fig. 1) to evaluate the temporal evolution of both FT products and compared them to SAT measurements. The objective was to identify any difficulties the products may have monitoring FT in particular conditions. SAT was chosen as in situ reference since Roy et al. 2015 showed that SAT was the best proxy to validate satellite FT products. Table 4 shows the percentages of agreement of weekly FT detection over the entire period between FT-AP, FT-ESDR and TAVG\_week. The mean agreement between the satellite products and in situ measurement is 81.6% for FT-AP and 92.0% for FT-ESDR.
Table 4: Agreement (%) of weekly FT detections between FT-AP and FT-ESDR and between satellite products and in situ data (TAVG week) for each site over the entire period. The sites are defined in Table 1.

<table>
<thead>
<tr>
<th>Country/Region</th>
<th>Land Cover</th>
<th>FT-AP / FT-ESDR (%)</th>
<th>FT-AP / TAVG week (%)</th>
<th>FT-ESDR / TAVG week (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamchatka/Kljuchi</td>
<td>Tundra</td>
<td>68.7</td>
<td>67.9</td>
<td>94.3</td>
</tr>
<tr>
<td>Canada/Québec</td>
<td>Tundra</td>
<td>83.8</td>
<td>90.8</td>
<td>89.7</td>
</tr>
<tr>
<td>US/Alaska</td>
<td>Forest</td>
<td>87.7</td>
<td>88.9</td>
<td>94.3</td>
</tr>
<tr>
<td>Russia/Siberia</td>
<td>Forest</td>
<td>97.1</td>
<td>97.7</td>
<td>97.1</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>Open lands</td>
<td>66.3</td>
<td>70.9</td>
<td>92.0</td>
</tr>
<tr>
<td>Canada/Saskatchewan</td>
<td>Open lands</td>
<td>76.2</td>
<td>73.3</td>
<td>84.6</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td></td>
<td><strong>80.0</strong></td>
<td><strong>81.6</strong></td>
<td><strong>92.0</strong></td>
</tr>
</tbody>
</table>

5

(a)
At the Kamchatka site (Fig. 6a, Table 4), FT-AP has a low agreement with TAVGweek at 67.9%. The error occurs mostly in summers with obvious false frozen misclassifications, since SAT is over 0°C during that period. In contrast, there is a strong agreement of 94.3% between FT-ESDR and TAVGweek, with differences occurring in the transitional period with no specific pattern between the years. The difficulty in the retrieval could be due to the fact that the Kamchatka site’s grid cell has a very low difference between the minimum and maximum NPR values (ΔNPR) used to create FF før and FFfter, with ΔNPR = 0.015 and ΔNPR = 0.021 for radiometers 2 and 3, respectively. This low difference may lead to a lower sensitivity to FT. Moreover, there is a change of ruggedness classification (Table 1) from the one grid cell ruggedness (SSM/I footprint scale) to the Rug_M (Aquarius footprint scale) from undulating to mountainous. With a coastline at about 85 km, a major difference of spatial variability exists between SSM/I and Aquarius measurements over the Kamchatka site.

The Quebec site (Fig. 6b), also over tundra land cover, has better product agreements with TAVGweek than the Kamchatka site, with percentages around 90%. FT-AP has generally a better agreement with TAVGweek during the fall freezing periods. There are only minor exceptions due to a few false frozen retrievals in summer. These exceptions show a typical situation in which FT-AP detects the freeze onset earlier than FT-ESDR, as mentioned in Sect. 3.2. The relatively higher ΔNPR (ΔNPR = 0.024 and 0.032 for radiometers 1 and 2, respectively) could be a factor generating fewer false flag retrievals than in the Kamchatka site grid cell.
For forest sites (Fig. 6c–d), both products have good agreement with \( \text{TAVG}_{\text{week}} \). The statistics for the Siberia site highlight the highest agreement: 97.7% for FT-AP and 97.1% for FT-ESDR. Interestingly, the forest sites have \( \Delta_{\text{NPR}} \) values comparable to those of the tundra sites, with \( \Delta = 0.022 \) and 0.029 for radiometers 2 and 3 respectively in Siberia, and a unique \( \Delta = 0.010 \) for radiometer 1 in Alaska. The latter value is the lowest of all the sites in this study. Since Alaska has relatively good FT-AP agreements (87.7% with \( \text{TAVG}_{\text{week}} \) and 88.9% with FT-ESDR), clearly small differences between FF_{fr} and FF_{th} alone cannot explain the false frozen retrieval problem at L-band.
At the open land sites, the low agreement (Fig. 6e–f) between FT-AP and TAVG\textsubscript{week} (70.9\% in Kazakhstan and 73.3\% in Saskatchewan) is mainly due to the false frozen retrieval in summer. During the transitional period, the FT-AP is in good agreement with TAVG\textsubscript{week}, sometimes better than FT-ESDR, especially in the fall of 2012, 2013 and 2014 in Kazakhstan. Nevertheless, FT-ESDR agrees relatively well, with 92.0\% in Kazakhstan and 84.6\% in Saskatchewan. The winter of 2011 in Saskatchewan was particularly warm, and the products reacted differently to a succession of over \textcircled{0}\textdegree C events, which affected the overall agreement percentage. The \(\Delta_{\text{NPR}}\) of the open land sites are 0.088 for radiometer 2 in Kazakhstan and 0.029 and 0.095 for radiometers 1 and 3 in Saskatchewan. Consequently, since these are the highest values of all sites, in this case, the false frozen retrievals cannot be explained by a small value of \(\Delta_{\text{NPR}}\).
Figure 6. FT detection for each reference site (see Table 1), with FT-ESDR (red dots) and FT-AP (blue dots) against surface air temperature (black dots and blue line) in a) Kamchatka, b) Quebec, c) Alaska, d) Siberia, e) Kazakhstan and f) Saskatchewan.
Comparing both products to SAT at different sites shows that FT-AP tends to identify false frozen retrieval in summer periods. It is out of the scope of this paper to explain why these misclassifications occur, but some hypotheses will be given in the discussion section.

4. Discussion

This study shows that overall FT-AP agrees well with weekly averaged SAT and with the Ka-Band FT-ESDR. Despite its being a weekly product, FT-AP has good sensitivity to the FT state of the landscape. Despite some regional discrepancies in forested landscape, very good agreements between FT-AP and FT-ESDR were found in this land cover, suggesting that the sensitivity of L- and Ka-bands to FT are more similar in forested landscape.

However, the study reveals that in certain regions, FT-AP seems to give false identifications of freezing surface in summer. These findings concord with other L-band FT analyses using SMAP and SMOS (Derksen et al., 2017; Rautiainen et al., 2016). Some regions like the coastlines, Kamchatka, Kazakhstan, Scandinavia, northern Europe, Alaska, the Canadian Rockies and the Canadian Prairies show agreement below 80% between FT-AP and FT-ESDR. An attempt was made to explain the false frozen retrievals occurring in the Kamchatka site and the two open land sites by looking at the \( \Delta_{NPR} \) values, but no direct relationship was observable. Relatively small \( \Delta_{NPR} \) are found for Kamchatka, but they are similar to those of Siberia, which has agreement higher than 95% with TAVG\(_{\text{week}}\). The Alaska site has the smallest \( \Delta_{NPR} \) of all the sites but does not possess the false frozen retrieval problem. To the contrary, the open land sites have the highest \( \Delta_{NPR} \) values and both have frozen retrievals during summer. Hence, \( \Delta_{NPR} \) can explain some of the weak classifications, but not all of them.

The false freeze classification in open land regions could be related to the crop growth (LAI increase) cycle. The growing vegetation leads to a stronger emission from the vegetation in both horizontal and vertical polarization (Gherboudj et al., 2012), causing a depolarization of the signal that decreases the NPR. This creates a similar effect to the FT signal and could lead to false freeze identifications in summer (Roy et al., 2015; Rautiainen et al., 2016). Another important factor that could influence the precision of L-band FT retrieval is the possibility of low soil moisture before freezing. Since the FT retrieval is based on \( \Delta_{\text{Ice/water}} \), low soil moisture will lead to a low FT signal. Hence, in dry regions and where there is irrigation only during the growing season like in Kazakhstan and the Canadian Prairies, dry soil could be misclassified as frozen soil as it has low permittivity.

Moreover, the different initial footprints of the analysed datasets could also explain some differences between them. For example, coastline proximity likely played a role in the Kamchatka results. Even if the products were resampled at the same scale, the surface heterogeneity such as the fraction of water (lakes and water near coastlines) within the initial footprint could...
generate changes in FT signals. In mountainous regions, it is possible that intra-pixel freeze onset date variability exists caused by colder temperatures at higher altitudes in contrast to warmer temperatures at lower altitudes. In this case, the frozen detections in some summer periods could be concurred with real freezing.

Putting aside those areas, the intercomparison shows recurrent patterns in the global annual freezing and thawing periods. A two to five week freeze onset delay is observed in tundra regions every year. Since this pattern is not as clearly seen as in forested area, it is unlikely that the differences come from the algorithm (i.e. STA vs MSTA methods). The causes are likely related to the physical behavior of microwave emissions at different frequencies, such as differences in emission and sensing depth, vegetation effects (as discussed previously) and ice/snow cover. Rowlandson et al. (in review) and Roy et al. (2017b) showed that L-band is sensitive to the freezing of the very surface related to the strong dielectric discontinuity, while the 37 GHz TB sensitivity is more related to the land surface temperature variation (Kim et al., 2017a). Hence, it is possible that the higher contrast of ice/water permittivity of L-band would make it more sensitive to the water-ice transition over the large landscape of a pixel (Artemov and Volkov, 2014; Roy et al. 2015; Rautianen et al., 2012, 2014). However, it remains that both FT products have different algorithms that could also lead to discrepancies. The FT-AP product looks at polarization ratio information, and its calibration is based on the land cover type. On the other hand, FT-ESDR is optimized pixel by pixel using single polarization observations at 37 GHz based primarily on the temperature information. Hence, further detailed ground-based measurements of soil state with radiometric emission at both frequencies could help to better differentiate these effects.

Soil heterogeneity makes the comparison with a single punctual in situ SAT limited (McColl et al., 2016; Lyu et al., 2018). SAT is a more homogenous reference than soil temperature, which influences the emission (by Planck’s law) of landscape elements such as soil, snow and vegetation. However, microwave emissions detected by a satellite radiometer with all the spatial variability of the environment within a pixel cannot be solely validated by SAT, since it does not consider phenomena like thermal inertia and latent heat exchange.

5. Data availability

The FT-AP is archived and distributed by the NASA Distributed Active Archive Center of the National Snow and Ice Data Center (NSIDC DAAC). The FT-AP can be accessed through the NSIDC online public data server (nsidc.org/data/nsidc-0736/versions/1). Table 5 summarizes all the data sets used in this study and lists where they are available for download.
6. Conclusion

In recent years, more attention has focussed on the use of satellite observations to retrieve surface freeze/thaw state. The new FT product derived from L-Band Aquarius passive observations (FT-AP) ensures, with the still operating SMAP mission, an L-band passive FT monitoring continuum with NASA’s spaceborne radiometers, for a period beginning in August 2011. In this study, we evaluated the FT-AP and compared it with a product based on 37 GHz measurements (FT-ESDR). This investigation has shown that FT-AP was generally good at retrieving the FT state of the surface for the given time of Aquarius mission, as 77.1% of the common grid cells have more than 80% agreement with FT-ESDR. Differences between the FT-AP and FT-ESDR occur during the complex transitional freezing and thawing periods. The comparison with in situ daily surface air temperature (SAT) showed cases of good concordance with FT-AP and station measurements during those periods. It was also shown that false frozen retrievals in summer with FT-AP also lead to discrepancies between both products. The problem can be caused by surface properties such as vegetation and low soil moisture that influence the L-band NPR.

The study showed that differences between FT products can be caused by the response of frequency to the component in a pixel like vegetation, soil, snow, footprint size. Deeper analysis of multi-frequency differences in relation to FT retrieval is needed. Hence, our results open the path to look at the fusion of multi-frequency algorithms for FT retrievals from passive microwave satellite observations and upcoming missions like the WaterCycle Observation Mission (WCOM; Shi et al., 2016).

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments

The authors would like to thank the Canadian Space Agency (CSA) and the National Sciences and Engineering Research Council of Canada (NSERC) for their financial support. L. Brucker was supported by NASA Interdisciplinary Research in Earth Science (IDS).
References


