Sentinel-1 detection of perennial firn aquifers in the Antarctic Peninsula

Lena G. Buth1,*, Valeria Di Biase2,*, Peter Kuipers Munneke2, Stef Lhermitte3, Sanne B. M. Veldhuijsen2, Sophie de Roda Husman3, Michiel R. van den Broeke2, and Bert Wouters3

*The authors contributed equally to the work
1Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
2Institute for Marine and Atmospheric Research, Department of Physics, Utrecht University, Utrecht, the Netherlands
3Department of Geoscience & Remote Sensing, Delft University of Technology, Delft, the Netherlands

Correspondence: Valeria Di Biase (v.dibiase@uu.nl)

Abstract. In recent years, the existence of perennial firn aquifers in the Antarctic Peninsula (AP) has been confirmed by in situ observations. Previous studies have suggested that these subsurface aquifers, together with meltwater ponds at the surface, provide a reservoir of liquid water to feed propagating fractures, promoting hydrofracture-driven ice-shelf disintegration. This study maps perennial firn aquifers in the AP from space using C-band Synthetic Aperture Radar imagery from ESA's Sentinel-1 (S1) mission. With these observations, we detect firn aquifers at 1 km x 1 km resolution, for the period 2017 to 2020. Existing methods, that use S1 data and rely on a backscatter intensity difference threshold approach, are prone to misclassify late-melt events as aquifers, when applied to the AP. Therefore, we have developed and evaluated a new approach that is better suited to the Antarctic environment. The new method exploits the characteristic, gradual backscatter increase during the (partial) refreezing of the liquid water in the firn layer after the peak of the melt season. Most firn aquifers are detected in the north and northwest of the AP, as well as on the Wilkins and George VI ice shelves. Aquifer locations detected with the present methodology agree with in situ observations and with model simulations of firn water content.

1 Introduction

The AP (Fig. 1) is the mildest and wettest region in Antarctica. Accumulation rates are high, up to several metres of water equivalent per year (w.e. yr−1) on its western side (Fig. 1(a)). Summer melt rates are the highest of the continent, mainly at lower elevations (Fig. 1(b); van Wessem et al. (2017)). Situated between two ocean basins that are covered by sea ice for extended periods each year, the AP experiences a large natural variability in temperature. At the beginning of this century, the rapid warming observed during the twentieth century (Hansen et al., 2010) reversed into a weak cooling trend (Turner et al., 2016; van Wessem et al., 2015). However, enhanced anthropogenic forcing is expected to further warm the AP in the long term. The expected concurrent increase in surface melt (Trusel et al., 2015; Siegert et al., 2019) makes it essential to understand its pathways, as it is known to play a role in ice shelf disintegration (Scambos et al., 2000; van den Broeke, 2005; Scambos et al., 2009; Gilbert and Kittel, 2021). On the Antarctic ice sheet, the fate of surface meltwater is either refreezing, runoff (Bell et al., 2018), supraglacial storage in melt ponds (Banwell et al., 2019) or englacial storage in buried meltwater lakes (Lenaerts et al.,...
2017; Bell et al., 2018). Under specific conditions, it is also possible that the meltwater percolates into the pore space of the firn without refreezing, to form firn aquifers, areas of subsurface liquid water storage. Perennial firn aquifers, which exist year-round (hereafter we shorten this terminology to firn aquifers), develop when a favorable combination of high summer melt and high winter accumulation rates prevent part of the stored meltwater to refreeze during winter (Kuipers Munneke et al., 2014), after which the aquifer can be recharged in the following summer. Extensive firn aquifers have been observed on the Greenland Ice Sheet, both in situ (Forster et al., 2014) and with ground and airborne radar, most prominently relying on data from the NASA Operation Ice Bridge (OIB) mission (Forster et al., 2014; Miège et al., 2016; Chu et al., 2018), and satellite remote sensing, using S1 C-band radar imagery (Brangers et al., 2020) and L-band brightness temperature measurements from the Soil Moisture Active Passive (SMAP) satellite mission (Miller et al., 2020a). The longevity of these firn aquifers differs widely (Horlings et al., 2022). Seasonal expansion and retreat of perennial firn aquifers can occur due to fluctuating meltwater availability throughout the year (Horlings et al., 2022). The aquifers in the Greenland ice sheet store and transmit liquid water within the ice sheet on decadal time scales (Miller et al., 2020a). Apart from Greenland, aquifers have been observed in smaller mountain glaciers around the world (Fountain, 1989; Kawashima, 1997; Schneider, 1999). Kuipers Munneke et al. (2014) states that the best conditions for perennial firn aquifer formation are when snow accumulates at rates of 1000 to 3000 mm w.e. y\(^{-1}\). These conditions are found in the AP, the northernmost and warmest part of Antarctica, which experiences significant surface melting during summer (Van Wessem et al., 2016). This climate resembles that of southeastern Greenland. In the AP, firn aquifers have also been observed, in situ and using ground penetrating radar (GPR), on the Wilkins Ice Shelf (Montgomery et al., 2020b), on the Müller Ice Shelf (MacDonell et al., 2021) and on King George Island (Simões et al., 2004; da Rosa et al., 2014). Occurrence of firn aquifers has been corroborated by firn model simulations (van Wessem et al., 2021). However, to our knowledge, no spatially comprehensive observational studies of firn aquifers in Antarctica have been published so far. This study aims to map firn aquifers from space using S1 radar backscatter measurements, with a focus on the AP. We analyze an existing method by Brangers et al. (2020) (hereafter referred to as the backscatter intensity difference threshold approach) that uses S1 data, and demonstrate its tendency to interpret late-season melt events as aquifers. Therefore we develop a new method (hereafter referred to as the recovery time threshold approach) that exploits the dependence of the C-band backscatter signal on the wetness of the firn (Stiles and Ulaby, 1980) and the resulting characteristic delayed backscatter increase after the peak melt season (here referring to the summer months December, January and February). This method acknowledges that melt can occur throughout the year in the AP (Kuipers Munneke et al., 2018) but is not sensitive to its occurrence.

The paper is structured as follow: Sect. 2 presents data and methods, Sect. 3 the results, Sect. 4.1 and 4.2, respectively, a comparison with ground truth and with firn models simulated data; the robustness of the thresholding is addressed in Sect. 5, followed by a discussion in Sect. 6 and by the conclusions in Sect. 7. A sensitivity analysis of the backscatter intensity difference threshold approach is presented in the Appendix A.
Figure 1. AP 2017-2020: average annual accumulation (a) and surface melt (b) as modeled in RACMO2.3p2 (updated from van Wessem et al. (2017)). The locations of the four backscatter time series mentioned below (Fig. 10) are indicated by the numbered white circles in (a): Larsen C (1), George VI (2), Wilkins (3) and Müller (4) ice shelves.

2 Data and Methods

2.1 Preprocessing of Sentinel-1 data

Synthetic Aperture Radar (SAR) Level-1 Ground Range Detected (GRD) images from the Copernicus S1 mission were used to detect the presence of firn aquifers in the AP. S1 measurement of backscatter strength have been available since the Sentinel-1A satellite was launched in April 2014. Sufficiently dense data coverage for our purpose was only reached after the launch of Sentinel-1B in April 2016. Consequently, this analysis uses data from 2016 (September) to 2020, as it was the most current and comprehensive dataset available at the time of our research. A lack of S1 data in 2017 does not allow a complete analysis across the entire AP (e.g., missing data on King George Island and part of the northern regions) for that year. The radar instruments operate at a center frequency of 5.405 GHz (C-band), so that clouds and solar illumination conditions do not influence the radar signal. We used acquisitions from the Extra-Wide Swath Mode, with a swath of 400 km and a spatial resolution of 20
m x 40 m. A preprocessed version of the data was accessed via the Python API of the Google Earth Engine (GEE) (Gorelick et al., 2017), where GRD border noise removal, thermal noise removal, radiometric calibration and terrain correction was already applied. The data are available in dual polarization: similar results were obtained by analyzing the co-polarized and cross-polarized data separately. Therefore, only the results of co-polarized (HH) backscatter intensity measurements ($\sigma^0_{HH}$) are presented here. Data were aggregated by linear scale averaging and projected onto the global cylindrical equal-area scaleable Earth Grid resolution version 2 (EASE-2) (Brodzik et al., 2012). The aggregation to 1 km improves the signal-to-noise ratio and is considered sufficient for an initial mapping of firn aquifers (Brangers et al., 2020). The S1-A/B satellites are each in a 12 day repeat cycle with 175 orbits per cycle, therefore the constellation offers an exact 6-day repeat cycle. For each satellite pass within one cycle, identified by its relative orbit number (RON), the geometry of the orbiting satellite with respect to the Earth’s surface is different, which influences backscatter measurements and results in a RON-specific bias. This bias was determined pixel-wise by taking the average $\sigma^0_{HH}$ of each RON and subtracting the average $\sigma^0_{HH}$ of all orbit numbers. Subsequently, the bias was removed for all RONs with at least 50 data points. Data points from remaining RONs were discarded, because we argue that the bias cannot be accurately determined for RONs with less data points, as the average might be influenced by seasonal variability.

### 2.2 Field observations

In recent years, the presence of firn aquifers in the AP has been confirmed through in situ observations and GPR surveys. Specifically, aquifers were identified in three distinct locations, visually represented by colored dots in Fig. 5. These findings served as crucial validation data to assess the effectiveness of the S1 aquifer detection method, based on the recovery time threshold approach.

In December 2018, fieldwork was conducted on the Wilkins Ice Shelf (70°48’00.0”S, 71°42’36.0”W) using a combination of borehole drilling, hydrological measurements, and GPR profiles (Montgomery et al., 2020b). During this survey, an average depth to water table of 13.4 m was measured (Fig. 5, green dot). The existence of a firm aquifer in this area has been substantiated through an analysis of radio-echo-sounding profiles from NASA Operation IceBridge (2014) (Montgomery et al., 2020a). These profiles were acquired over the Wilkins Ice Shelf utilizing the Multichannel Coherent Radar Depth Sounder (MCoRDS). The analysis revealed a distinct bright reflector within the upper firn layer, resembling the high-amplitude reflections typically associated with firm aquifers observed in the Greenland Ice Sheet.

Between 20 and 22 February 2019, three short firm cores (M1; M2; M3) were collected from the Müller Ice Shelf (M1: 67°15’15.5” S, 66°53’44.5” W; M2: 67°15’16.8” S, 66°50’59.8” W; M3: 67°15’37.6” S, 66°48’22.3” W) (MacDonell et al., 2021). At site M3, the recovered core was saturated to the point that the drill was filled with liquid water; this saturated section was found closer (3.7 m) to the surface than the other two sites (M1: 16.7 m and M2: 18.5 m). A water table was also present on King George Island ice cap in several studies (e.g., Simões et al. (2004); da Rosa et al. (2014)).

During GPR surveys carried out in January 2011 at Wanda Glacier, King George Island, a strong reflection marked a water table at an average depth of 35 m (Fig. 5, yellow dot)(da Rosa et al., 2014).
2.3 Firn model and regional climate model

AP firn aquifers have been simulated with a firn model forced with the output of a regional climate model. In this study we compare the S1-detected spatial extent of aquifers with results from the latest version of the IMAU Firn Densification Model (IMAU-FDM v1.2A) (Ligtenberg et al., 2011; van Wessem et al., 2021; Veldhuijsen et al., 2023). Using layers of 2 to 15 cm thickness, the semi-empirical IMAU-FDM simulates the transient evolution of a vertical firn column using a Lagrangian layer administration. The model includes parameterizations for heat conduction, dry snow densification, meltwater percolation and refreezing. Meltwater percolation is simulated with the bucket method, whereby each firn layer has a maximum irreducible water content that decreases with increasing density following Coléou et al. (1999). Any remaining meltwater moves downwards through the layers in a single model timestep, until it refreezes in a layer with sufficient available pore space and if sufficient latent heat of refreezing can be stored in the layer. Otherwise, remaining liquid water can be retained up to the maximum irreducible water content. Percolation continues down to the pore close-off depth, where the firn layer transitions to impermeable glacier ice. There, instantaneous runoff is assumed. IMAU-FDM does not simulate standing water or lateral water flow, so only a qualitative comparison can be made with S1 detected aquifers on the basis of the presence of irreducible liquid water content (LWC). When applied over the Greenland ice sheet, a previous version of IMAU-FDM correctly predicted firn aquifers observed using in situ data and airborne radar (Forster et al., 2014; Kuipers Munneke et al., 2014; Steger et al., 2017). IMAU-FDM is forced at its upper boundary by three-hourly fields of instantaneous surface temperature, 10-m wind speed, cumulative accumulation (total precipitation minus sublimation minus drifting snow erosion) and cumulative surface liquid water input (melt plus rain) from the regional atmospheric climate model RACMO2.3p2 (Fig. 1, Van Wessem et al. (2016)). RACMO2 has a horizontal resolution over the AP of 5.5 km x 5.5 km, which also determines the horizontal resolution of IMAU-FDM.

2.4 SMRT radiative transfer model

To compare IMAU-FDM and S1 time series of firn aquifers for selected locations (see Fig. 1(a)), we translated the IMAU-FDM time-series of density, temperature and liquid water content into a simulated S1 HH-backscatter time-series using the Snow Microwave Radiative Transfer (SMRT) radiative transfer model (Picard et al., 2018). Ten-day IMAU-FDM depth profiles of density, temperature and liquid water content with a vertical resolution of 4 cm were converted into a SMRT snowpack, where we assumed an exponential microstructure model with correlation length of 1 mm throughout the profile. This SMRT snowpack was subsequently translated into modelled S1 backscatter using SMRT with S1 settings for frequency (5.405 GHz) and typical incidence angle (30°) and using the Discrete Ordinate and Eigenvalue Solver (DORT).

2.5 Aquifer detection algorithms

The backscatter of a microwave radar signal decreases with increasing wetness of the snow, due to the higher absorption and difference in the scattering of the radar signal by liquid water (Stiles and Ulaby, 1980; Shi and Dozier, 1992; Tsai et al., 2019). Using the SMRT model (Picard et al., 2018, see Sect. 2.4), Fig. 2 shows the simulated backscatter for four snowpacks with...
a 1 m simulated aquifer layer (density of 600 kg m\(^{-3}\), 5 % liquid water content) introduced at different depths, compared to the simulated backscatter of an identical snowpack without the aquifer layer. Therefore, we used the IMAU-FDM snow/firn profile at four locations (see also Fig. 10) on 8 Oct 2019, when no liquid water was present in the profile. When the aquifer layer is introduced at shallow depths, it causes a clear reduction in backscatter compared to the dry scenario. However, with increasing depths of the aquifer layer, this contrast decreases. It means that gradual burial or refreezing of a wet layer leads to a characteristic backscatter increase. An important limitation is that the backscatter reduction can only be observed if the wet layer is in the upper few meters of the firn. In the present setting, at a depth of 8-10 m the difference becomes smaller than the S1 radiometric accuracy (3 standard deviations) of 1 dB. Unlike the observations of OIB flights, the S1 data therefore do not allow for a direct measurement of the depth of the water table. Consequently, aquifer detection using S1 data can only be based on the temporal evolution of the S1 radar backscatter after the peak melt season. During melt episodes, the backscatter signal is low for all locations with snow melt. At locations without significant liquid melt water retention, the refreezing after the melt season results in an increase in \(\sigma^0_{HH}\) up to pre-melt season values typically within a few days or weeks. In contrast, persisting liquid water results in a delayed increase of the radar backscatter signal after the peak melt season. In the presence of an underlying firn aquifer, latent heat from partial refreezing and high snow accumulation slow down the refreezing of the upper layers (Phillips et al., 2010; Humphrey et al., 2012; Kuipers Munneke et al., 2014; Brangers et al., 2020). This also results in an increase of radar backscatter, but considerably delayed compared to non-aquifer locations. The difference in the \(\sigma^0_{HH}\) curve over time can therefore be used to differentiate firn aquifers from non-firn aquifers. A variety of approaches have been developed to quantify such a delay in signatures for the detection of firn aquifers in Greenland. Brangers et al. (2020) used a backscatter intensity difference threshold approach at fixed times determined by the difference of S1 radar backscatter near the end of winter to that of early autumn to deduce the presence of firn aquifers, while Miller et al. (2020a) fitted a sigmoid curve to L-Band brightness temperature curves. In this study, we present an alternative recovery time threshold approach based on the time it takes for the backscatter to recover to pre-melt values. To assess the performance of this alternative method, we compare the recovery time threshold approach (Sect. 2.5.1) to the existing backscatter intensity difference threshold of Brangers et al. (2020) (Sect. 2.5.2).

### 2.5.1 Backscatter intensity difference threshold approach

A backscatter intensity difference threshold approach based on Brangers et al. (2020) was applied for the detection of aquifers in the AP in the years 2018 to 2020 only as not sufficient S1 data were available for 2017. The detection is based on the reduction in \(\sigma^0_{HH}\) at fixed dates in early autumn (i.e., March) relative to that near the end of winter (i.e., October). Subsequently, this reduction is compared to a predefined threshold, and if it surpasses this threshold, aquifers are assumed to be present. Given the limited extent of aquifer mapping in the AP, calibration of a dedicated threshold proved to be unfeasible. Instead, we adopted a threshold value based on Brangers et al. (2020) in Greenland, who used a difference of 9.4 dB between the end of winter and early autumn.
2.5.2 Sentinel-1 aquifer detection: recovery time threshold approach algorithm

Alternatively, a new recovery time threshold approach is introduced as illustrated in Fig. 3. The S1 data in Fig. 3 represent an example of a backscatter time-series in an aquifer location with reduced backscatter intensities $\sigma_{0,\text{Jan}}$ in the peak melt season between 1 and 15 January and high, stable backscatter intensities in September $\sigma_{0,\text{Sep}}$ that correspond again to pre-melt backscatter intensities. Our novel recovery time threshold method is now based on the time it takes for the backscatter to recover to pre-melt values instead of using the backscatter difference in backscatter between autumn and end of winter. The recovery time threshold approach involves the following steps, as illustrated in the flowchart of Fig. 4.

First, we masked out locations without a distinct increase in backscatter after the melt season ($\sigma_{0,\text{Sep}} - \sigma_{0,\text{Jan}} \leq 3$ dB): they are considered dry-snow regions without summer melt. Next, we excluded blue-ice areas using the (Hui et al., 2014) blue-ice data, as they can yield backscatter responses similar to aquifer locations. To smooth the $\sigma_{0,\text{HH}}$ time-series, we applied a 15-day centered moving average as the third step. Subsequently, the recovery date was determined as the first day of the year at which the smoothed time-series recovers to 80% relative to the pre-melt backscatter intensity (DOY$_{80}$). To calculate DOY$_{80}$, we compared the value of $\sigma_{0,\text{Sep}}$ from the previous year with $\sigma_{0,\text{Jan}}$ of the current year. Pixels with DOY$_{80} \geq 97$ were considered firn aquifer locations. Finally, for pixels that did not meet the aquifer criteria from the fifth step, we applied a sixth step based on the...
Figure 3. Visual illustration of the firn aquifer detection method. The blue dots represent an exemplum, showcasing a backscatter time-series in a real aquifer location; the periods 1-15 January and 1-30 September are highlighted as bars in salmon pink. The corresponding backscatter values and the 80% threshold are indicated by horizontal lines. The vertical blue line shows the first day on which the threshold 80% is exceeded (DOY$_{80}$), coinciding with the aquifer detection threshold DOY$_{80} \geq 97$.

criterion $\bar{\sigma}^0_{\text{Sep}} \leq -12$ dB, where $\bar{\sigma}^0_{\text{Sep}}$ is the mean $\sigma^0$ value in September. This step aimed to identify perennial aquifer locations where substantial amounts of near-surface liquid water do not refreeze during winter. By including these pixels, we accounted for areas that did not recover a stable September value consistently across years. To generate the final map of firn aquifers, we applied a temporal and a spatial filter. Spatially, we eliminated connected areas with 3 or fewer pixels, corresponding to an area of 3 km$^2$. By enforcing the requirement of contiguous positive pixels, we could reduce the risk of considering individual noisy points that might not truly represent an aquifer. Temporally, we excluded locations that were detected for only one year during the 4-year period under consideration. The use of a fixed DOY$_{80} \geq 97$ threshold does not account for variations in melt season length from year to year. It implies that we focus solely on locations that demonstrated a consistent presence as aquifers for at least two years during the study period.
Applying the threshold \( \text{DOY}_{80} > 97 \)

S1 GRD images (2016-2020), HH pol.;
GEE preprocessing
EASE-2 (1 km x 1 km) projection

Masking locations with \( \delta^0_{\text{Sep}} - \delta^0_{\text{Jan}} \leq 3 \text{ dB} \)

Excluding blue ice areas

Smoothing \( \delta^0_{\text{Jan}} \) time series

Determining recovery data (DOY_{80})

Applying the threshold \( \delta^0_{\text{Sep}} \leq -12 \text{ dB} \)

Positive detection

Noise removal:
remove connected areas with 3 or fewer pixels

Reliability evaluation:
exclude pixels detected for only one year

Aquifer detection

Figure 4. Sentinel-1 Aquifer detection algorithm: flowchart using the recovery time threshold approach.

3 Results

3.1 Aquifer extent using Sentinel-1 remote sensing data

3.1.1 Aquifer extent using backscatter intensity difference threshold approach

Using the backscatter intensity difference threshold approach, aquifers are detected in the north, northwest and west coastal regions (Fig. 5). On the George VI Ice Shelf, aquifers are mostly detected for one year, except for a small area in the north where aquifers are detected for three years. The three colored dots in Fig. 5 represent the locations where in situ data are available: i) on the Wilkins Ice Shelf (green dot), aquifers are detected for three years; ii) on the Müller Ice Shelf (blue dot), aquifers are only detected for one year; iii) on most of the King George island, aquifers are detected for two years, but on the glacier of interest (Wanda Glacier, yellow dot), no aquifers are detected. Furthermore, a large aquifer area is detected on the Larsen C Ice Shelf, where no aquifers have been observed (Hubbard et al., 2016). It is a region with high snow melt but relatively low accumulation rates, conditions under which aquifers generally do not form (Kuipers Munneke et al., 2014).
The red box in Fig. 5 shows the backscatter intensity time-series within a defined region of Larsen-C during the year 2020. The analysis of the timeseries underscores a limitation of the method, which relies on a fixed threshold based on two distinct points in time. Specifically, when the average value during one of these periods (in this instance, early autumn, i.e., March) experiences a significant reduction due to a specific phenomenon (e.g., a late-season melt event), it creates a false positive classification. The large-scale average backscatter intensity of March 2020 significantly differs from the seasonal average values of preceding years. Consequently, in the year 2020, the observed backscatter difference with respect to October exceeds the 9.4 dB threshold, leading to the classification of the entire area as a potential aquifer. This region is comparable to the Larsen C area that experienced the latest melting in the study by Luckman et al. (2014). This suggests that this methodology exhibits the tendency to misinterpret late-season melting as an aquifer. Further analysis of the applied threshold’s value is provided in Appendix A.

3.1.2 Aquifer extent using the new recovery time threshold approach

Here, we present the detection of aquifers using the new method presented in Sect. 2.5.2. Firn aquifers are predominantly detected in the coastal regions located in the north, northwest, and west of the AP (Fig. 6, aquifer locations color-coded by number of years detected). Moreover, in 2017 and 2020, additional firn aquifer locations were identified on the Wilkins Ice Shelf, as well as on the northern and southern extremities of the George VI Ice Shelf. For a more detailed analysis of the yearly variations, Fig. 6 includes lateral panels illustrating the results for each individual year. To provide a comprehensive overview, Table 1 summarizes the total aquifer areas detected for each year analyzed, as well as the combined findings for the entire study period (all years).
Figure 5. S1 aquifers detected with the backscatter intensity difference threshold approach (2018-2020). Shades of grey indicate the number of years each pixel has been identified as an aquifer. Dark grey represents pixels detected as aquifers for three years, while lighter grey shades correspond to those detected for two or one year. The three colored dots represent the locations where field observations have been conducted confirming the presence of firm aquifers: Wilkins Ice Shelf (green dot), in December 2018; Müller Ice Shelf (blue dot) in February 2019; King George Island (yellow dot) where a perennial aquifer is observed using GPR observations starting in 2004. In the red box, the backscatter time-serie (from June 2019 to June 2020) of the Larsen C area located in the red square. The AP maps in the current and subsequent sections of this study make use of the AP basemap sourced from the Quantarctica database (Matsuoka et al., 2021).
Figure 6. Main panel: S1 aquifers detected with the recovery time threshold approach. Shades of red represent the number of years in which each pixel has been detected as an aquifer: dark red represents the pixels which have been detected as aquifers for four years, lighter reds represent the pixel detected for three or two years. The colored dots represent the locations where field observations have been conducted confirming the presence of firn aquifers: Wilkins Ice Shelf (green dot), in December 2018; Müller Ice Shelf (blue dot) in February 2019; King George Island (yellow dot) where a perennial aquifer is observed using GPR observations starting in 2004.

The lateral panels illustrate the firn aquifer areas detected for each individual year: 2017, 2018, 2019, and 2020. The light yellow color indicates detections that occurred only for one year and are not considered as aquifer detection in the present study. In 2017, the lack of S1 data prevented aquifer detection across the entire AP, resulting in missing data on King George Island and parts of the northern regions. Firn aquifers are predominantly detected in the coastal regions in the north, northwest, and west of the AP. Additionally, there were notable firn aquifer locations on the Wilkins and George VI ice shelves in the years 2017 and 2020.
Table 1. Total aquifer area detected with S1 (recovery time threshold approach) in each of the four study years; in the last line, the area continuously detected as aquifer between 2017 and 2020.

<table>
<thead>
<tr>
<th>Year</th>
<th>S1 aquifer area [km$^2$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>97761</td>
</tr>
<tr>
<td>2018</td>
<td>54370</td>
</tr>
<tr>
<td>2019</td>
<td>35543</td>
</tr>
<tr>
<td>2020</td>
<td>111739</td>
</tr>
<tr>
<td>2017-20</td>
<td>18956</td>
</tr>
</tbody>
</table>

4 Validation and evaluation

4.1 Validation with ground truth data

The S1 detected aquifers with the recovery time threshold have been compared with field observations (Sect. 2.2), represented with colored dots in the main panel of Fig. 6. Full agreement is found on King George Island and on the Müller Ice Shelf, where S1 detects aquifers throughout the entire observation period (as mentioned in Sect. 3.1.2, S1 data are missing in 2017 on part of King George Island). On the Wilkins Ice Shelf, aquifers at the field observation site are detected in 2017 and 2020.

4.2 Comparison with IMAU-FDM LWC

Considering the lack of extensive aquifer mapping in the AP and the small number of field observations, the S1 results have been compared with the output of IMAU-FDM. Figure 6 shows that aquifers are detected mainly in regions where both the snow accumulation and snow melt rates are high, in agreement with previous studies (Kuipers Munneke et al., 2014). In regions with high snowmelt but relatively low accumulation rates, such as the Larsen C Ice Shelf, aquifers are not detected with the new recovery time threshold, in contrast to the results obtained using the backscatter intensity different approach.

Figure 7 further quantifies this by comparing the S1 detected aquifer pixels with the average fields of accumulation and surface snowmelt of RACMO2.3p2 for 2017-2020. Detected aquifers are preferentially located in regions where both surface melt and accumulation rates are high. Figure 7 also reveals that the fraction of detected aquifers generally does not decline with increasing accumulation for a given melt rate. This indicates that the S1 recovery time threshold approach does not fail to detect aquifers in regions with high accumulation rates and the associated fast burial.

Figure 8 presents a visual comparison between the areas identified as aquifers using S1 and IMAU-FDM data. It compares the occurrence of LWC $> 0$ mm w.e. y$^{-1}$ in the IMAU-FDM dataset with the locations detected as aquifers in the S1 data. This comparison offers valuable insights into the concurrence of aquifer detections between the two datasets: the aquifer locations are in agreement on the north and west coasts and islands and partially on the Wilkins Ice Shelf. Discrepancies are found on the George VI Ice Shelf, where S1 observes more aquifers, and on the northern part of the Wilkins Ice Shelf, where IMAU-FDM
Figure 7. Percentage of S1 detected firn aquifer occurrences in all years (2017-2020) as a function of annual surface melt (y-axis) and accumulation (x-axis) averaged over the period 2017-2020. RACMO2 grid cells are grouped in melt bins of 50 mm w.e. y\(^{-1}\) and accumulation bins of 200 mm w.e. y\(^{-1}\).

detects more aquifers. The relationship between the IMAU-FDM LWC probability and the detection frequency of S1 aquifers is represented in Fig. 9. It is evident that pixels with S1 positive detections of only a single year exhibit significantly lower LWC probabilities compared to pixels with multi-annual detections. On the other hand, pixels identified as aquifers by S1 throughout the entire study period demonstrate LWC > 0 mm w.e. y\(^{-1}\) in approximately 70\% of cases. Furthermore, the upper value of the interquartile range extends up to 100\%.
Figure 8. (a) IMAU-FDM LWC heat map. The colorbar represents the percentage of times each location is characterized by LWC > 0 mm w.e. yr⁻¹ in the years 2017-2020. (b) Firn aquifer area detected by S1 with the new recovery time threshold approach.
Figure 9. IMAU-FDM LWC probability, defined as frequency of LWC > 0 mm w.e. y⁻¹, versus frequency of S1 aquifer detections, showing that LWC probabilities increase for more frequent S1 aquifer detections. The median value is represented by the red lines, while the upper and lower boundaries of the blue box indicate the 75th and 25th percentiles, respectively.

5 Analysis: threshold robustness

5.1 Sentinel-1 and SRTM backscatter time series at four locations and rationale for fixed $\bar{\sigma}_0^{\text{Sep}}$ threshold

To further clarify the advantage, but also the complexity, of the S1 recovery time threshold approach, Fig. 10 shows the time series of relevant observed and modeled parameters at four sample locations (see Fig. 1(a)) with progressively higher backscatter reduction and modeled LWC. There is a clear correspondence between the presence of liquid water in IMAU-FDM and the corresponding forward-modeled SMRT S1 backscatter time-serie, on the one hand, and the observed S1 backscatter time series on the other hand. For instance, in the summer of 2020, IMAU-FDM simulates a higher LWC, and the modeled SMRT and observed S1 backscatter time series show good correspondence with a more gradual increase in backscatter. However, in summer 2019, IMAU-FDM simulates a lower liquid water content, and the modeled SMRT backscatter increases rapidly after the melt season, whereas the observed S1 backscatter time-serie increases only gradually. This indicates that IMAU-FDM in combination with SMRT potentially underestimates aquifer presence for some years. On the Larsen C Ice Shelf (67°06’36.0”S, 64°20’42.0”W, Fig. 10(a)), a relatively dry location with rapid meltwater refreezing and therefore small amounts of retained liquid water, every year the observed backscatter increases rapidly to $\sigma_0^{\text{Sep}}$ values at the end of the melt season and DOY$_{80}$ remains below the threshold value of 97. As a result, no aquifers are detected when using the DOY$_{80}$ detection method. On
Figure 10. Time series for four example locations on (a) Larsen C, (b) George VI, (c) Wilkins, and (d) Müller ice shelves (for locations, see Fig. 1(a) of observed and smoothed Sentinel-1 backscatter signal (red), modeled vertically (0-15 m depth) integrated IMAU-FDM LWC (green) and resulting modeled SRTM backscatter (blue). The colored vertical lines represent the DOY$_{80}$ parameter (purple) and the threshold DOY$_{80} = 97$ (pink) for the respective location and year.

George VI (70°18′36.0″S, 68°39′00.0″W), Wilkins (70°48′00.0″S, 71°42′36.0″W) and Müller (67°15′00.0″S, 66°52′12.0″W) ice shelves in Figs. 10(b)-(d), the increase in observed S1 backscatter is more gradual with the regular occurrence of DOY$_{80} \geq 97$, meaning that a firn aquifer is detected by S1. In the case of the location on Müller Ice Shelf (Fig. 10(d)), higher amounts of retained liquid water result in lower backscatter in September, leading to non-uniform increases in backscatter. This sometimes causes the DOY$_{80}$ threshold to underestimate the presence of aquifers, especially strong melt years like 2020. The Müller Ice Shelf time series demonstrate the need for the inclusion of an additional detection condition, namely that of the mean $\sigma^0$ value in September ($\bar{\sigma}^0_{\text{Sep}}$) to be $\bar{\sigma}^0_{\text{Sep}} \leq -12$ dB.
5.2 DOY$_{80}$ threshold

To investigate the sensitivity of the DOY$_{80}$ threshold, we varied its value and assessed the impact on the aquifer detection. Lowering the DOY$_{80}$ threshold below 90 (Fig. 11, panel (c)), led to the detection of aquifers in the southeastern region of Larsen C Ice Shelf, where aquifers are not expected nor observed (Kuipers-Munneke et al., 2014). The region identified as an aquifer on Larsen C using DOY$_{80} < 90$ is comparable to the area where surface melt is known to occur very late in the season (Luckman et al., 2014). Consequently, we have adopted this value as the lowest threshold since lower values seem to trigger the detection of late-season melting. Conversely, increasing the DOY$_{80}$ threshold beyond 103 (Fig. 11, panel (c)) resulted in the exclusion of aquifer detections, affecting observations on the Wilkins Ice Shelf which is no longer detected as an aquifer. We consider the uncertainty of our method to be bracketed by these two values of 90 and 103 (Fig. 11, panel (a) and (b)).

To consolidate our findings into a singular output, we selected the mean value within this DOY$_{80}$ range as a fixed threshold, DOY$_{80} = 97$.

6 Discussion

The S1 firn aquifer detection method presented here offers a good estimation of the extent of firn aquifers in the AP, supported by in situ observations and combined firn model and regional climate model output. Compared to previous methods based on backscatter intensity differences, the new method demonstrates more skill in the AP, particularly in terms of false detections: none of the analyzed years recorded evident cases of late melt detection being classified as firn aquifers. However, it is important to note that the S1 method does not directly estimate the liquid water content, making it challenging to solely rely on Sentinel-1 data to discriminate aquifers from the presence of near-surface meltwater, especially during years with long melt seasons. The threshold used for firn aquifer detection is a fixed date range. However, it is important to note that all DOY$_{80}$ thresholds within the range 90 and 103 are scientifically viable and, given the constraints of our current validation dataset, further refinement of this parameter can be achieved with more S1 observations and/or more in situ data. Choosing a later DOY$_{80}$ threshold may lead to underdetecting firn aquifers in years with short melt seasons or deep aquifers that no longer influence the signal after several months. Conversely, an earlier threshold may include more false detections due to surface meltwater or late melt. The selected range finds a balance between these scenarios, leading to plausible results without requiring additional datasets.

Interpreting the causes of discontinuous aquifer detection (two or three years out of four) is challenging due to limited in situ data. The existence of both seasonal and perennial aquifers on the AP, suggested by IMAU-FDM data, indicates that the discontinuous S1 outputs might represent seasonal aquifers, which do not persist through the winter season and do not form again in the following melt season under unfavorable environmental conditions (Horlings et al., 2022). However, according to recent literature, most of the aquifers detected in Greenland have a semi-permanent (decades) nature (Miller et al., 2020b). The hypothesis is that firn aquifers in the AP are likely to exist for multiple years in a row, similar to observations on the Greenland Ice Sheet. The discontinuous detection of aquifers in certain areas, such as the southern part of the Wilkins Ice Shelf, could be attributed to the non-direct measurement of the water table, where top-down refreezing depends on fluctuating annual surface...
Figure 11. S1 aquifers detected with recovery time threshold approach using: (a) the lower threshold range (DOY$_{80}$ ≥ 90), and (b) the upper threshold range (DOY$_{80}$ ≥ 103); (c): the aquifer map in the bottom panel shows the output obtained using a threshold below the considered range (DOY$_{80}$ = 85), while the aquifers displayed in the top panel are detected using a threshold higher than the upper limit of the range (DOY$_{80}$ = 120)

melt rates. Additionally, variations over time in water table depth or thickness, as observed in Greenland aquifers (Miège et al.), may influence the backscatter signal.

The firn model IMAU-FDM, used for comparison of the S1 aquifer retrieval, has some limitations due to its simplified treatment of firn hydrology. For instance, lateral flow is not included in IMAU-FDM, and the model currently does not allow for the presence of liquid water beyond the capillary capacity of the firn. Although IMAU-FDM has been shown to capture the spatial extent of aquifers and ice slabs, it likely underestimate the volume of liquid water in firn aquifers. It makes a more quantitative comparison using modeled LWC challenging.

The current version of the detection algorithm does not eliminate outliers in the backscatter time series. While outliers are not expected to significantly affect the estimated total aquifer area, they might influence the parameter DOY$_{80}$ and consequently impact the classification of individual locations in specific years.

A limitation in the use of Sentinel data for the detection of firn aquifers is that for some years and some locations in the AP,
the number of data points is not sufficient for this analysis (e.g., 2016). This issue might become relevant in the current years as, despite longer time series that can be analyzed, the amount of data is affected by the end of the mission for the Sentinel-1B satellite after its failure in December 2021. The launch of the third Copernicus Sentinel-1C, scheduled in 2023, should restore the robustness of the algorithm.

Future studies incorporating OIB flight measurements could enhance the S1 detection algorithm.

7 Conclusions

This study presents a method for mapping AP firn aquifers from space using S1, based on a delayed increase in SAR backscatter after the peak melt season in the presence of a firn aquifer. Full agreement has been found between S1 and the three field observations considered. Good agreement between satellite and model is found in the spatial distribution: S1 firn aquifers are found in coastal regions in the northwest and north of the AP, on the Wilkins Ice Shelf, and in the north and south of the George VI Ice Shelf. According to the model, the firn aquifers along the AP are confined to narrow bands in the lowest-lying areas; discrepancies with S1 are partially observed on the Wilkins and George VI ice shelves. The locations of the firn aquifers coincide with regions of high snow accumulation and snow melt, supporting previous studies that these conditions are required for their formation. With temperatures in the AP rising in a future warming climate, increased snowfall/rainfall and surface melt and less refreezing could increase the amount of liquid water in the firn and thereby enhance firn aquifer formation, influencing the stability of ice shelves. It thus remains important to continue observing firn aquifers in the AP. Further studies are needed to optimally validate the presented S1 detection method. More in situ measurements of the firn liquid water content, ground radar data, and improved modeling of water transport in firn can help calibrate the demonstrated method and aid to better distinguish perennial aquifers from transient surface meltwater.

Appendix A: Sensitivity analysis of the backscatter intensity difference threshold approach

To explore the possibility of finding a less sensitive threshold value to late-melt events, simulations with different thresholds were conducted. The analysis aimed to strike a balance between mitigating late melt detection and preserving aquifer detection accuracy at in situ locations. Using a lower threshold value of 8 dB (see Fig. A1(a)), the output revealed a larger aquifer detection on the Larsen C. On the other side, empirical adjustments demonstrated that a higher threshold could reduce late melt detection, though it significantly affected aquifer detection at other locations. In fact, the melt area detected on Larsen C decreased as the threshold value increased. However, considering the detection of aquifers at field observation locations (as shown in Sect. 2.2), using a threshold value of 11 dB already had notable effects: aquifers were detected for three years on Wilkins Ice Shelf, for only one year on Müller Ice Shelf, and no detection on Wanda Glacier, whilst a relatively small aquifer area on Larsen C remained part of the output (Fig. A1(b)). In light of these observations, no further attempts were made to increase the threshold value beyond the current setting.
Figure A1. S1 aquifers detected using the backscatter intensity difference threshold method with two different threshold values: (a) a lower threshold of 8 dB, which resulted in a larger aquifer detection on the Larsen C; (b) a threshold of 11 dB, which led to a decrease in the melt area detected on Larsen C. However, this increase in the threshold value negatively affected the detection of aquifers at field observation locations (indicated by colored dots).
Code and data availability. Data and code will be made available in an open repository after acceptance.

Author contributions. LB, VD, BW and PKM designed the study and the Sentinel-1 experiments; LB and VD carried them out. SV ran the IMAU-FDM model; SL the SRTM model. LB and VD prepared the manuscript; all authors reviewed and commented on manuscript drafts.

Competing interests. MB, BW and SL are members of the editorial board of journal The Cryosphere.

Financial support. Sanne Veldhuijzen, Bert Wouters, Stef Lhermitte and Sophie de Roda Husman acknowledge funding by NWO grant OCENW.GROOT.2019.091.
References


