Reviewer #2
The manuscript uses RADARSAT-2 data to estimate melt pond fraction within the Canadian Arctic. The manuscript is clear and well written with figures clearly supporting the presented results and the discussion.

Howell et al.
We thank this reviewer for her/his comments that have improved this manuscript. We have incorporated almost all this reviewer's suggestions.

Reviewer #2
I found the investigation into the correlation between the different regions and the melt pond fraction one of the most important findings of this study. Maybe this finding could be more explicitly stated in the abstract and also in the conclusion? "Static/stable sea ice regions showed a higher detrended correlation." The mentioning of several regions is a bit vague.

Howell et al.
Agreed.

Revised Abstract:
Dynamically stable sea ice regions within the CAA exhibited higher detrended correlations between RADARSAT-2 f_{pk} summer sea ice area.

Revised Conclusions:
The results presented in this study indicate that dynamically stable sea ice regions within the CAA exhibit a higher detrended correlation between RADARSAT-2 f_{pk} and summer sea ice area.

Reviewer #2
Single pol RADARSAT-2 data was used, why is that? Was the combination of HH + HV lacking? Or did the HH-channel contribute sufficient information? This may have been covered in earlier work by e.g. Scharien et al., but would then be worth reiterating.

Howell et al.
Single pol RADARAT-2 was used for two reasons. The first is that Scharien et al. (2017) found the HV data produced noisy results and the second there is not sufficient HV imagery in the early of the RADARSAT-2 to cover CAA. The latter is because only in the recent years has HH+HV been ordered operationally throughout the CAA.

Revised Data:
We limited our analysis to only RADARAT-2 at HH polarization because Scharien et al. (2017) found HV produced noisy results in addition to there not being sufficient imagery at HV polarization in the early of the RADARSAT-2 record to cover CAA.

Reviewer #2
The in-situ area only covers areas with a relatively high proportion of melt ponds, were any other in-situ data available that could be used for the validation with a smaller proportion of melt ponds? Moreover, the area covered for the in-situ data is rather small compared to the pixel size

of the RADARSAT-2 images. Are there larger datasets, either more locations or covering a larger area that could be used to strengthen the argument?

Howell et al.

Yes, we do have aerial photograph estimates of melt pond fraction obtained over and adjacent to the LiDAR site in 2012 from Scharien et al. (2014), which we have made use of to compare with RADARSAT-2 f_{pk} estimates. We have added a new Figure 7b with the aerial photograph data and revised the following sections:

Revised Section 3.2

Figure 7a compares the time series of the entire 100 m LiDAR melt pond fraction coincident with the f_{pk} determined from RADARSAT-2 at the coinciding pixels. For 2011, RADARSAT-2 f_{pk} corresponds to the end of stage I and beginning of stage II thus providing a very good representation of the seasonal peak of the f_p, when the melt pond control on heat uptake and ice decay, through the ice-albedo feedback, is greatest. For 2012, RADARSAT-2 f_{pk} also corresponds to the end of stage I and beginning of stage II but is ~0.1 lower than *in situ* f_p values. This is likely due to the short duration but very high maximum f_p of 0.78 in 2012 as Scharien et al. (2017) found that equation (1) sometimes underestimates very high f_p due to the low $\gamma°$ signal associated with very smooth FYI.

Figure 7b shows the distribution of RADARSAT-2 f_{pk} and the f_p determined from aerial photo observations on June 22nd, 2012 near Resolute. The aerial photographs were acquired within 1 week of f_{pk} coverage being observed at the LiDAR site. The comparison was done by averaging all RADARSAT-2 pixels within each aerial photo. The mean aerial photograph f_p was 0.54 and RADARSAT-2 f_{pk} was 0.53 with an the RMSE of 0.10 and bias of 0. The distributions are in reasonably good agreement but RADARSAT-2 values are slightly narrower than the distribution of f_p from the aerial photographs. It is likely the RADARSAT-2 distribution is narrow on the left tail because our method captures peak pond coverage and some of the regions photographed were before or after their seasonal peak. We attribute the narrow right tail to the documented underestimation of equation (1) from Scharien et al. (2017). However, it is notable that both RADARSAT-2 and the aerial photograph datasets capture the same bimodal f_p distribution, with the first mode around 0.4-0.5 characterizing rougher sea ice areas and the second mode around 0.7 capturing smooth flooded sea ice.

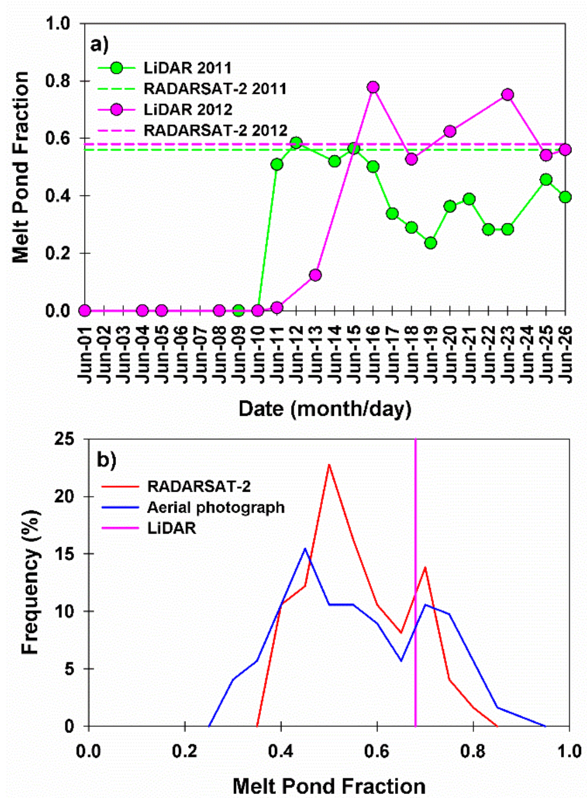

Figure 7. a) Temporal evolution of observed melt pond fraction (f_p) and RADARSAT-2 peak melt pond fraction (f_{pk}) at *in situ* observations sites for 2011 (74.7229°N; 95.1763°W) and 2012 (74.7264°N; 95.5772°W). b) Frequency distribution of RADARSAT-2 f_{pk} and aerial photograph f_p observations in Resolute Passage on June 22, 2012; the pink vertical link represents the mean LiDAR f_p on June 22, 2012.

Revised Data

Aerial photographs of estimated f_p directly over the LiDAR site and the adjacent sea ice area away from land and open water were also obtained on June 22, 2012. The aerial photographs have a pixel resolution 0.22 m resolution, cover 750 m by 750 m. In total, 123 aerial photographs of f_p were used and a complete description of the dataset is provided in Scharien et al. (2014).

Added Reference

Scharien, R. K., Hochheim, K., Landy, J., and Barber, D. G.: First-year sea ice melt pond fraction estimation from dual-polarisation C-band SAR – Part 2: Scaling in situ to Radarsat-2, The Cryosphere, 8, 2163–2176, https://doi.org/10.5194/tc-8-2163-2014, 2014.

Reviewer #2

The comparison between the results using Sentinel-1 and RADARSAT-2 imagery was interesting, but a discussion about why the results are different (e.g. Fig 6) is missing. Both of the images being C-band SAR one would expect the results to align quite well. Please discuss this. The comparison between the RADARSAT-2 and MODIS data, particularly figure 8, seems to suggest large differences between the two sensors, where even the maximum fp is significantly lower than the RADARSAT-2 estimates.

Howell et al.

We should have provided more discussion between and Sentinel-1 and RADARSAT-2 as also suggested by Reviewer #1.

Revised Section 3.2

Frequency distributions of RADARSAT-2 f_{pk} and Sentinel-1 f_{pk} from Scharien et al. (2017) in the CAA for 2016 and 2017 are shown in Figure 6. Sentinel-1 appears to estimate more regions of lower f_{pk} compared to RADARSAT-2 which are typically associated with MYI. Whereas, RADARSAT-2 estimates more regions of higher f_{peak} which are typically associated with smooth FYI. We consider these subtle differences to be primarily the result of taking the mean of all available April RADARSAT-2 imagery (Table 1) over all incidence angles in the CAA compared to only using images from Sentinel-1 within the CAA constrained to a certain incident angle range. As shown in Figure 2, the uncertainty in RADARSAT-2 f_{pk} varies depending on the number of pixel overlaps (images). Overall, the f_{pk} distributions are in good agreement between both sensors.

As for the MODIS product, it underestimates peak pond melt fraction in the CAA and is more representative of pond coverage at synoptic timescales. Even the maximum f_p from MODIS is from an 8-day running mean of daily pond fraction estimates, so will underestimate the seasonal peak f_p if the duration of peak ponding is <8 days. As also suggested by Reviewer #1 we firm up the wording here to point this out and have revised the text in 3.2 as follows:

Revised Section 3.2

The seasonal time series of the 8-day composite MODIS f_p, the maximum seasonal MODIS f_p and the predicted RADARSAT-2 f_{pk} for 2009-2011 is shown in Figure 8. MODIS f_p observations within the CAA indicate initial pond formation occurred in May for all years with f_{pk} reached in mid-July for 2009 and in early June for 2010 and 2011. Compared to the RADARSAT-2 f_{pk} values, the peak MODIS f_p is ~0.20 smaller. RADARSAT-2 f_{pk} is higher on average than MODIS because the MODIS 8-day product does not represent f_{pk}. The MODIS f_p observations are determined weekly using 8-day composite image products that would include some melt pond formation and drainage processes prior-to, and after, the seasonal peak. Moreover, MODIS f_p observations give the time series of f_p therefore even the highest seasonal estimated MODIS f_p is reduced because while some regions of the CAA are at their seasonal peak but others are behind or ahead. To that end, we also calculated the maximum f_p from MODIS regardless of timing during the melt season, for each pixel, also shown in Figure 7. These values more closely compare with the RADARSAT-2 f_{pk} but are still ~0.05 smaller on average. Even the maximum f_p from MODIS is from an 8-day running mean of daily pond fraction estimates, so will underestimate the seasonal peak f_p if the duration of peak ponding is <8 days. However, the top whisker of the box plot of the maximum f_p from MODIS indicates that MODIS does capture some regions at peak during the 8-day time series. Although we are using MODIS f_p product to compare against out RADARSAT-2 f_{pk} estimates, Rösel et al. (2012) found that the MODIS f_p product also has errors up to ~0.1. Overall, MODIS f_p estimates are more representative of the seasonal mean f_p rather than f_{pk} within the CAA.

Reviewer #2

Were there regions in the CAA that showed better agreement between the MODIS and RADARSAT-2 estimates?

Howell et al.
Not really. We produced spatial maps but decided not to include them because the do not provide as much information as the boxplots.

Specific comments
Reviewer #2
Consider moving the information about stages of lake evolution on page 6 to the information about data or similar instead. Readers unfamiliar with melt pond development would be aided by an earlier introduction to the different stages. On P3 it is stated that the evolution stages covered by the field work covers 3 out of 4 stages, but on P 6 R177-179 it states that stage I and II was captured. Please clarify.

Howell et al.
We assume the reviewer means pond evolution. This seems the ideal place to describe these stages in accordance with the Figure 7 showing that the LiDAR site captures stages 1 to 3. Since the site is over first-year ice stage 4 will not occur and requires no discussion. We have removed references to the melt pond stages in the data description. Re-reading our text, it seems clear that the RADARSAT-2 f_{pk} values fall within end of stage I and beginning of stage II at the LiDAR site.

Reviewer #2
Is it expected that the environmental conditions remain reasonably stable in CAA during the month of April? If so maybe that could be added to strengthen the argument for combining RADARSAT-2 data for the analysis?

Howell et al.
Yes, it is expected. We have already explicitly stated this in the methodology: "…together with the fact that the majority of the sea ice in the CAA is landfast (immobile) during April which results in a temporally stable f_{pk} for all April images."

Reviewer #2
Minor comments
The use of the words excellent and good in the abstract are slightly abstract. Maybe it would be possible to provide some statistical measure?

Howell et al.
We added a statistical measure the temporal linkage but the spatial needs to be visual.

Revised Abstract
The temporal variability of RADARSAT-2 f_{pk} over the 10-year record was found to be strongly linked to the variability of mean April multi-year ice area with a statistically significant detrended correlation (R) of R=-0.89. The spatial distribution of RADARSAT-2 f_{pk} was found to be in excellent agreement with the sea ice stage of development prior to the melt season.

Reviewer #2
P2 L41. What is the difference between sea ice area and extent? Should it possibly say sea ice type and sea ice extent?

Howell et al.
No area and extent are the correct terms and the ones most commonly used. Sea ice area is ice concentration multiped by the area of the region. Extent is also calculated as area multiplied by ice concentration but it this assumes that the area is 100% provided it is greater than a certain threshold (i.e. typically 15%). A great explanation is found on the NSIDC website "A simplified way to think of extent versus area is to imagine a slice of swiss cheese. Extent would be a measure of the edges of the slice of cheese and all of the space inside it. Area would be the measure of where there is cheese only, not including the holes. That is why if you compare extent and area in the same time period, extent is always bigger. A more precise explanation of extent versus area gets more complicated."
http://nsidc.org/arcticseaicenews/faq/#:~:text=The%20most%20common%20threshold%20(and,said%20to%20be%20ice%20free.

Reviewer #2
P2 L43. Does fp here relate to maximum/mean values? Please clarify

Howell et al.
As suggested by Reviewer #1 we have modified the notation throughout the manuscript to denote melt pond fraction as f_p and peak melt pond fraction as f_{pk}.

Reviewer #2
P6. L169. Should it be : : : allows us to place the: : :?

Howell et al.
Yes. Inserted "to".

Reviewer #2
P6. R192. Should this be Figure 8?

Howell et al.
Yes.

Reviewer #2
Fig 1. Please state what the green star indicates in the figure text.

Howell et al.
New Figure caption as follows:
Figure 1. Map of the Canadian Arctic Archipelago region (red shading). The green star indicates the location of the LiDAR and aerial photograph observations.

Reviewer #2
Fig 7. Should it be -W in the coordinates.

Howell et al.
Removed the '-'