Response to review by reviewer #2

We thank the reviewer for constructive criticism, allowing us to improve our manuscript. A majority of the reviewer's requests have been met. Below follow our answers and comments, highlighted in blue after each of the reviewer's comments. Blue page (P) and line (L) numbers refers to the manuscript published in TCD and red numbers refers to the marked-up revised version.

General comments

The authors have studied melt pond fraction (MPF) estimation using TerraSAR-X dual-polarization SAR imagery acquired over drift ice north of Svalbard, and presented empirical models for MPF estimation in two different wind speed conditions (low speed and intermediate speed).

In the Introduction Section authors give good overview on importance of melt ponds on the Arctic sea ice heat budget and in the Arctic climate system, and on previous studies on melt pond fraction estimation with optical and SAR imagery. The number of previous SAR studies on melt pond detection and fraction estimation is quite large, and so far a generic method for the estimation has not been presented/developed. There has been some success for the melt pond fraction (MPF) estimation over smooth landfast ice using C-band co-polarization ratio (HH and VV pol SAR images needed) or HH-pol backscattering coefficient (σ). For MPF estimation over drift ice only few studies has been conducted. At least it seems that MPF estimation over drift ice with C-band single pol imagery is not possible. Over drift ice sea ice deformation features like ice ridges and make MPF estimation in theory much more difficult than over smooth landfast ice. Other frequencies than C-band have been used only in few case studies. Likely (to my opinion) accurate MPF estimation is only possible with high resolution (<5-10 m) SAR imagery. So far time series of MPF maps over the Arctic have been produced only with optical imagery (MODIS, MERIS). These charts are limited by accuracy of automatic cloud masking and persistent cloud cover during the Arctic summer. Accurate MPF charts from SAR imagery would supplement greatly the MPF charts from optical imagery.

Section 2 discusses nicely about melt pond signatures in SAR imagery, but it could also include overview of observed MPF behavior during melt ponding season (ponding, drainage etc.), see e.g. D. G. Barber, J. J. Yackel, and J. M. Hanesiak, "Sea ice, RADARSAT-1 and arctic climate processes: A review and update," Can. J. Remote Sens., vol. 27, no. 1, pp. 51–61, 2001.

We appreciate this suggestion, and have included a brief comment on this in the manuscript.

- The following sentences were added (Melt ponds in SAR imagery, P4L113, P4L127-130): The coverage of melt ponds varies during the melt season, starting out with a high fractional cover, and reducing as the ponds drain. At the end of the melt season, the melt ponds refreeze. This evolution is mirrored by a seasonal variation in the sea ice SAR signature (Barber et. al 2001)."

Dual-polarization TerraSAR-X imagery acquired over drift ice for MPF estimation have been previously studied by Kim et al. (2013) and Han et al. (2016). Kim et al. (2013) used only one TSX image acquired in Aug 2011 over East Siberian Sea, but they have large amount of co-incident airborne very fine resolution X-band (single pol) images. Comparison MPF data was from airborne photography. Han et al. (2016) used the same datasets and also one additional TSX image acquired in July 2011 over the Chukchi Sea. Kim et al. (2013) estimated MPF with "We first delineate the ice and melt pond features using image processing software (ENVIA EX), based on the combination of multiscale segmentation and aggregation methods."; not discussed in more details, but Han et al. (2016) studied various polarimetric parameters and their textural features in MPF estimation by machine learning approaches.

The authors have used here four TSX dual-polarized StripMap images acquired during ICE2012 campaign in north of Svalbard. Comparison MPF data was from helicopter-borne optical imagery. In addition, surface roughness data was calculated from stereo camera imagery, and weather data was measured by R/V Lance at the ICE2012 campaign site. They have studied MPF estimation with different polarimetric parameters calculated from the dual-pol TSX imagery, as was done also by Han et al. (2016). The main questions now are: 1) Does this study give new scientific results/information compared to Kim et al. (2013) and Han et al. (2016)? 2) In what ways it is different to them, data and/or methods?

My answers: Study area is different (Chukchi Sea vs. Svalbard) which could have influence on the results if sea ice conditions (FYI vs. MYI) where different; authors should discuss this in the paper. Wind speed is taken into account here unlike in Kim et al. (2013) and Han et al. (2016). Wind speed has large effect on the backscattering from melt ponds (not frozen). Somewhat more SAR images have used, four here compared to two in Han et al. (2016), making the results here more reliable. The developed MPF algorithms are linear functions between MPF and one polarimetric parameter. I favor this kind of simple approach as the results can be related easily to theoretical backscattering models. Han et al. (2016) utilized machine learning approaches where relations between polarimetric parameters (and scattering theory) and an estimated parameter may not be very clear. However, I think that the paper in its current form gives quite little new information/findings compared to previous studies on the MPF estimation with SAR.

We appreciate the papers by Kim et al. (2013) and Han et al. (2016), which as our study focus on melt ponds and SAR. Kim et al. (2013) focuses on airborne SAR and barely discuss satellite imaging and is hence very different from our study. Han et al. (2016) presents a machine learning approach for melt pond fraction retrieval from X-band SAR based on two satellite SAR scenes, one that has corresponding information about actual melt pond fraction.

As the reviewer has commented on, there are several differences between the study of Han et al. (2016) and our study. Han et al. (2016) focuses on MYI, and explicit request more studies on other sea ice types in their conclusion. Our manuscript focuses on level FYI, which has different microwave signature than MYI, as commented on in Han et al. (2016). The effect of sea ice surface roughness, satellite noise floor, wind speed, and incidence angle are not discussed in Han et al. (2016), but wind speed and incidence angle are suggested to be investigated in future studies. These are important factors in understanding the melt pond signature of X-band SAR and are therefore included in our study, adding essential information. Only one scene with corresponding melt pond information was employed by Han et al. (2016. Increasing the number of scenes, as done in our manuscript, is of large importance to improve the understanding of melt ponds signature in SAR imagery.

While Han et al. (2016) focuses on combining several polarimetric signatures in machine learning algorithms for melt pond retrieval, our manuscript concentrates about the individual polarimetric features, and the influence of melt ponds on these. This could form a basis for more advanced methods later on, guiding which features to use under specific wind speed conditions and incidence angles. Melt ponds and sea ice are treated as different classes in the study of Han et al. (2016). Our study on the other hand, focuses on the signature of mixtures of melt ponds and sea ice.

When it comes to the results, different polarimetric features are found sensitive to melt pond fraction in Han et al. (2016) and our study. While co-polarisation ratio and HH intensity were the most promising features in our study, co-pol phase difference, alpha angle, and HH intensity were found to be the most important ones in Han et al (2016). This could be due to e.g. difference in wind conditions, which there is little information about in the study by Han et al. (2016).

In summary, we find that Han et. al (2016) and our study does not present much overlapping information and findings, hence our study are complementary. The sea ice types, methods and results are different, and our study includes several factors (more scenes, wind speed, incidence angle, surface roughness etc.) not considered in Han et al. (2016). To evolve in the understanding of melt ponds signatures and impact on X-band SAR, and SAR in general, a variety of studies is necessary, and we believe our study contributes with enough new insight, and is worthy of publication.

To clarify the differences in the study of Han et al. (2016) and our study for the readers of our manuscript, the following changes have been applied:

- The following sentences were rephrased (Introduction, P3L82, P3L85-91): "Han et al. (2016) combined multiple polarimetric SAR features in MPF estimation by machine learning methods, employing the co-polarisation channels of the MYI X-band SAR scene explored in Kim et al. (2013). An additional scene was also included in the study, though without melt pond information. The study showed promising results, but the authors claim that more scenes with various sea ice types and incidence angles are needed to develop a general propose MPF model. Lack of wind information is also limiting the relevance of the study."
- The following paragraph was added (Discussion, P15L490, P18L592-604): "The findings in our study deviate from the findings of Han et al. (2016) where σ^{o}_{HH}, $<\rho$, and, α_1 were found to be the most prominent polarimetric features in separating melt ponds, sea ice and open water in high resolution X-band SAR imagery. Differences in sea ice type, sea ice surface roughness, wind conditions, and SAR incidence angle could possibly explain why different polarimetric features are sensitive to MPF in the two studies. The methods of the two studies are also slightly different, as Han et al. (2016) classify each pixel into melt pond, sea ice or open water, while our study focuses on mixtures of melt ponds and sea ice. Exact wind information lacks in Han et al. (2016) but the wind speed is expected to be low. This could explain why σ^{o}_{HH} contributes strongly in MPF estimation, and is then in accordance to our findings. The diverging results in the two studies emphasize the need of investigating melt ponds impact on SAR imagery under different conditions and for a variety of sea ice types. It also stresses the importance of supplementary measurements of parameters like wind speed and sea ice surface roughness."

The statistical reliability of the developed empirical MPF estimation models seems quite low, r2 was at best only 0.21 and RMSE is high.

We agree that the R^2 values for the regression fits are low. For our comment on this, see point 2f) in the reply to Reviewer #1.

The value of the paper could be improved by following changes and additions:

The empirical models for the MPF estimation were developed using datasets over a large ice floe. Why were not all co-incident SAR imagery vs. airborne photography used? How results would change if they were?
As this study was performed on drifting sea ice, co-location between SAR scenes and helicopter data is very challenging. This is also commented on in the conclusion of the manuscript. Most of the airborne photos were not possible to be co-located with the satellite observations exact enough to meet our demands of a high quality study. However, for the investigated floe, we managed to do a reliable co-location, and the floe was also the only floe appearing in two of the scenes. We therefore chose to focus on this specific floe in our investigations to secure the quality of the study.

Both wind speed and SAR incidence angle have large effect on the MPF estimation. Wind speed is now taken into account by MPF models for two different wind conditions. I suggest you developed

MPF models which include incidence angle or compensate σ incidence angle variation before MPF estimation. The study should include more variable wind speed conditions, but in the current dataset these are not present

We agree that it is desirable to study the melt pond signature under as many wind speed situations as possible to make a robust melt pond estimation. Our data set consists of four scenes, and it is therefore not possible to make such a robust model from this small dataset. However, our study highlights the importance of wind speed, and can serve as a starting point for future studies. This is already commented on in the conclusion of the manuscript.

Incidence angle correction has been introduced to the manuscript, to improve the understanding of Fig. 9. The following changes have been made in the manuscript:

- The following sentences were added (Method, P10L295, P11L340-346): " Incidence angle correction was applied to the scenes for a better comparison, employing the following equation (Kellndorfer et al., 1998)

  $$\sigma^\circ_{corr} = \sigma^\circ\,(\sin\theta / \sin\theta_{ref})$$

  where σ° is the original backscatter coefficient, θ is the center incidence angle of the scene to be corrected, and θ_{ref} is the reference incidence angle of scene T4. The correction was only applied in the low-wind case, as it canceled in the intermediate wind case due to the use of a co-polarisation ratio."
- The following reference was added: Kellndorfer, J., L.E., P., M.C., 715 D., and Ulaby, F. T.: Toward consistent regional-to-global-scale vegetation characterization using orbital SAR systems, IEEE Transactions on Geoscience and Remote Sensing, 36, 1396-1411, doi:10.1109/36.718844, 1998.
- The following sentence was rewritten (Results, P13L413, P15L501-502): "Incidence angle correction according to Eq. 16 is applied to the figure, accounting for σ°_{VV} decrease with incidence angle."
- The following sentence was rewritten (Discussion, P14L449, P17L548-550): "The underestimation of FMP in scenes T1-T3 is likely related to higher wind speeds at the time of acquisition."

You could study effect of sea ice type in the MPF estimation, e.g. by first segmenting the SAR images to level ice and deformed ice categories (with the help aerial photography if possible). In best case we could have also sea ice type taken into account in the MPF estimation. You have also surface roughness data which could be utilized here.

We agree that the effect of surface roughness could be better presented in the manuscript. We have now introduced two classes of sea ice in the scatter plots (Fig. 4 and 7), representing totally level ice and partly deformed sea ice. The classification is based on visual interpretation of the helicopter images. The following changes have been introduced in the text in relation with the classification:

- The following sentences were added (Results, P11L355, P13L408-411): Grey dots correspond to areas with some degree of sea ice deformation, while blue dots correspond to areas with completely level ice. Deformation information is extracted from visual inspection of the helicopter images."
- The following sentence was rewritten (Results, P12L358, P13L414-416): "A majority of the lowest $R_{VV/HH}$ values are appearing in partly deformed areas. Areas with some degree of deformation also represent the lowest f_{MP}."
- The following sentence was added (Results, P13L395, P15L470-471): "Grey dots correspond to partly deformed areas, while blue dots represent level ice."

Show MPF maps from some SAR images and discuss spatial variation present, does it make sense? You have four SAR images, how does estimated MPF behave temporally? Now Table 5 shows MPF averages over the full scenes, but these are not much discussed in the text, and temporal variation does not seem right (36.2->45.7->31.2->53.3).

Spatial MPF maps have been introduced to the manuscript. See comments to point 3) in the reply to Reviewer#1.

During the campaign, the MPF was stable with a mean of 33.2%. The variation seen in Table 5 is due to differences in wind speed conditions and incidence angle. This is thoroughly commented on and discussed in the last paragraph of section 4.1 (New 3.1) and in the third paragraph in section 5 (Discussion).

Can you compare your estimates with those from optical imagery? See http://icdc.zmaw.de/1/daten/cryosphere/arctic-meltponds.html

The suggested data set lasts until 2011, and can therefore not be used in our study. We agree that it would be interesting to compare our data to optical data, and investigated this opportunity early in the manuscript process. However, there were cloudy conditions during the campaign, and optical data were therefore not accessible. This underlines the advantage of a possible melt pond fraction estimation from SAR.

The study would benefit greatly from a larger SAR dataset. Are there any co-incident TSX vs. in-situ / airborne data from NICE2015 campaign you could use? You really need more wind speed conditions for the MPF estimation development. Even including more TSX images without corresponding comparison data is possible, you could study spatial and temporal trends. In addition, any fine resolution C-band images available? Comparison between C- and X-band would be nice addition.

We agree that one should aim to use as many SAR scenes as possible in a study like this. But this has to be balanced by the actual access to scenes with high quality information about melt pond fraction and sea ice conditions retrieved from in situ measurements and helicopter photos. Such data sets are very rare, and it is therefore important to publish results from existing data sets, even if they as in our case have a limited number of scenes.

The N-ICE-2015 campaign was finished in mid-June 2015, before the onset of intense surface melt, and data from this campaign is therefore not appropriate for our study. C-band scenes were planned collected during the campaign, but acquisition priorities hampered collection of such scenes. We have also attended other campaigns with intention of increasing the data amount. The lack of success with these efforts emphasizes the value of the presented data set.

In general, the paper is well written and structured, and easy to read and understand. The data processing and analysis methods are scientifically sound and discussed in needed detail. I am afraid in the current form the paper gives quite few new scientific findings compared to previous studies.

We appreciate that the reviewer likes our manuscript. The novelty of this study is already discussed above. We find that our study brings in new and different findings compared to the only other existing study on the topic (Han et al. 2016), and we believe it is worthy of publication.

From Conclusions: “Future studies should aim to include a larger number of satellite scenes acquired during various sea ice conditions, melt pond evolution stages, wind speeds and incidence angles. The effect and limitations of sea ice surface roughness and dependency on filtering size and scale should also be further investigated.”

You should consider taking some of these topics to this paper!

Wind speed, incidence angle, surface roughness, and filtering size are all discussed in our manuscript. This sentence simply states that more data is needed to make a robust algorithm for melt pond fraction retrieval in X-band. To make this absolutely clear, we have rewritten the sentence:

- The following sentences were rewritten (Conclusion, P16L525, P19L651-654): "For development of a robust operational method, future studies should aim to include a larger number of satellite scenes acquired with various sea ice conditions, melt pond evolution stages, wind speeds, and incidence angles. The effect and limitations of sea ice surface roughness and dependency on filtering size and scale should also be further investigated."

Finally, Yackel and Barber (2000) speculated that MPF may be more closely related to the albedo than to melt pond fraction due to the fact that albedo results from the integration of all surface types (snow, saturated snow, melt ponds) which contribute to the measured MPF. What's the authors' view on this; would it be better to investigate the relationship between SAR data and albedo than SAR and melt pond fraction? Please, discuss this in Introduction Section.

This question might have a typing error, and we find it slightly unclear. We interpret the question to ask whether estimated melt pond fraction should be compared to albedo instead of observed melt pond fraction. Albedo refers to the average reflection of waves in the visible range of the microwave spectrum. As SAR uses microwaves to evaluate the sea ice surface, we find it more credible to utilize differences in the microwave signature between melt ponds and sea ice, or methods that employ statistical features describing fractional mixtures of surfaces.

In our study, albedo is not measured, and would have to be estimated by upscaling from melt pond fraction measurements and in situ measured albedo values of different surface types. This method would inevitably introduce additional uncertainty to the results (see f.ex. Divine et al., 2015 for estimates made for the study area), and is therefore not advisable for our data set. The relationship between albedo and polarimetric features will therefore not be discussed in detail in our manuscript.

Specific comments

1.Introduction

page 3, lines 90-92: terms 'dual polarimetric' and 'dual-polarisation' used, confusing…I think it should be 'dual-polarisation' for SAR imagery with two polarizations.

We agree in this, and have changed the term:

- The following sentence was rewritten (Introduction, P3L90-92, P4L101-103): "The objective of this study is to investigate the potential of melt pond fraction retrieval from level drifting FYI with dual-polarisation X-band satellite SAR. A data set consisting of four high resolution dual-polarisation TerraSAR-X satellite scenes…"

2. Melt ponds in SAR imagery

p. 4, l. 118: "Observed surface roughness increases with increasing frequency, making X-band more sensitive to small-scale surface roughness than C-band."
I think surface roughness is physical property of a surface, and its effect on backscattering depends on radar wavelength.

We agree that this sentence was unprecise, and have rewritten it:

- The following sentence was rewritten, (Melt ponds in SAR imagery, P4L118, P5L135-137): "X-band is more sensitive to small-scale surface roughness than C-band, as the effect of surface roughness depends on radar wavelength."

l. 124: “Six of these features are included in our study and are described in the following subsection.” In Table 4 there are eight features.

This is correct, and we have corrected the sentence:

- The following sentence was rewritten (Melt ponds in SAR imagery, P4L124, P5L142): “Eight of these features...”

3.2 Data set

Give absolute calibration accuracy and equivalent number of looks (ENL) (i.e. effect of radar fading) in TSX images. Do they have any significant effect on your data analysis results?

The absolute radiometric calibration accuracy is 0.6 dB. This parameter is an image-wide measure and includes temporal drift. Hence, it probably varies more from near-range to far-range than in a local region. As we focus on small regions in our main analysis, this accuracy will probably not play a significant role. However, for development of a future operational algorithm across many images, the calibration accuracy might be of importance. We therefore include the accuracy in our manuscript, and note that it should be further explored in future studies. The following changes were introduced in the manuscript:

- The following sentence was added (Method, P8L231, P9L272): “The absolute radiometric calibration accuracy of TSX is 0.6 dB (Airbus Defense and Space, 2013).”
- The following sentence was added (Discussion, P15L490, P18L611-614) “The absolute radiometric accuracy of TSX scenes could also influence the results of our study, but this influence is expected to be very small compared to other uncertainties.”

ENL is a far more complicated issue, and we do not see how including this would strengthen the manuscript. We are not statistically modelling the speckle distributions, and only look at the mean values after smoothing. ENL does have some bearing on the degree of variation, but this is directly evident form our observations without the complication of trying to interpret the effect of ENL. Also, radar texture, or non-Gaussianity, which is observed to be high for TSX imagery, is a far more influential factor. All simple ENL measures are Gaussian-based estimates, which do not correctly capture the texture aspect, and could lead to incorrect interpretation. Based on these reflections, we have decided not to include ENL in the manuscript.

4.1 Sea ice conditions

Was there any nighttime re-freezing on sea ice and melt pond surfaces which could have influenced backscattering signatures in T1 TSX image acquired at 06:52 UTC on 28 July?

Nighttime refreezing was not observed during the campaign, and is therefore unlikely to have influenced any of the SAR signatures. This is now clarified in the manuscript:

- The following sentences were changed/added (Method, P10L314, P8L256-257): “Air temperatures were varied little between -1 to 1.5C. Combined with the oceanic heat flux, the ice was therefore in continuous melt even at nighttime.”

4.3 Intermediate-wind case

p. 12, l. 358: “From visual inspection of the helicopter images, some of the lowest RVV/HH values origin from slightly deformed areas with a surface roughness possibly exceeding the Bragg criterion.” Please discuss how this sea ice condition leads to low RVV/HH.

We agree that this should be better described. We have therefore expanded the sentence with a possible explanation of the low values.

- The following sentence was rewritten (Results, P12L358, P13L411-414): “The partly negative values of RVV/HH imply that $\sigma^0_{HH} > \sigma^0_{HH}$. This might be a result of multiple scattering events in the sea ice volume or sea ice surface, possibly connected with sea ice deformation.”
- The following sentence was rewritten (Discussion, P14L460, P17L559-560): “Multiple scattering events in the sea ice surface and sea ice volume may also have contributed to the large sample variations observed in Figs. 4 and 8.