Reply to Anonymous Referee #2 from 21 Oct 2017

Note: Author responses are in plain text following the original referee comment shown in italicized text.

Line 50. I suggest adding Fetterer and Untersteiner (1998) and Arntsen et al. (2015) to the reference list.

We have added both references to this section. Although the Arntsen et al. 2015 paper uses the method presented in Miao et al. 2015.

Line 66. This isn't quite right. Several previous works have demonstrated surface feature retrievals from high-resolution imagery throughout the seasonal evolution of ice surface conditions.

We have reviewed a large number of previous works detailing the classification of high resolution optical imagery of sea ice (e.g. Arntsen et al., 2015; Fetterer and Untersteiner, 1998; Inoue et al., 2008; Kwok, 2014; Lu et al., 2010; Miao et al., 2015; Perovich et al., 2002b; Renner et al., 2014; Webster et al., 2015), and have not come across any that are demonstrated on a complete seasonal melt cycle. To the best of our knowledge no such work exists. We will gladly incorporate further information into this section if the reviewer can point us to the references they are referring to.

Lines 136+. How does the algorithm differ from that in Miao et al? Please describe any differences.

While this algorithm is inspired by the work of Miao et al. 2015 to use image segmentation followed by classification with a random forest algorithm, the implementation of that workflow is quite different. To convey this, we have added line 132: "Our implementation of the segmentation and classification, however, were custom-built using well known image processing tools (Pedregosa et al., 2011, van der Walt et al., 2014) in an open source format".

The algorithm presented by Miao et al. (2015) uses the ENVI GIS software package. As such, there are some specifics that remain proprietary to ENVI. Where we know how the Miao et al. algorithm behaves, we have stated the similarities and differences. We use a custom-built segmentation technique (section 3.2) that is different than the Miao et al. method. In the random forest machine learning technique, we use some attributes that were developed by Miao et al. (2015) (attributed in line 267 and 303), as well as attributes new to our method (lines 302-313).

Line 236. Does this melt pond definition include or exclude melt ponds that are melted through? Relative to previous works, is it typical to include submerged ice in the melt pond class or is it unique to this approach?

Our melt pond definition, which is provided in lines 239-248. excludes the area of melt ponds that has melted through completely (see Figure 5). Our approach to the surface classifications was to consider primarily shortwave optical properties. Submerged ice and melt ponds have similar optical properties and impact the solar energy balance in the same way. Thus it makes sense to group them into a single category. Previous works have taken both approaches. Miao et al. (2015), for example, presented a method for distinguishing general submerged ice from contained melt ponds by analyzing their proximity to open water. However, separating these classes is not necessary for all applications. Spectral unmixing algorithms, such as those presented by Rosel et al. 2012, to determine melt pond fraction on a larger scale consider only aggregate optical properties, and melt pond fraction would necessarily include the general submerged ice category as well.

Line 241. Submerged ice isn't a type of melt pond. Please clarify this point. It would be helpful to comment on the effects of submerged ice on melt pond statistics of area and geometry, especially for scenes of advanced melt in the marginal ice zone.

Lines 361+. It would be helpful to explicitly include submerged ice in the melt pond class throughout the text and figures. For example, instead of "Melt pond," please state "Melt pond and submerged ice" or "Melt pond + submerged ice."

To be more clear on our definition, we have changed this category to be "melt pond and submerged ice" throughout the figures and text.

Our hope is to spur community discussion with these surface type definitions, and so we have presented what we feel is the most widely applicable way to standardize 'ponded ice'. We acknowledge in line 236 that there are different opinions. We found that many experts in the sea ice community have subtly different definitions of the surface types, even beyond the distinction being made here. (as we discovered when producing the data for Figure 4.) From a shortwave optical stand point, submerged ice and melt ponds are functionally the same, and since radiative balance is a primary reason to study ponds, we argue in line 244-245 that it makes sense to group them as a single category. For a study concerned with pond geometry this is obviously not the case, and there are methods (such as those discussed in Miao et al. 2015) to separate general submerged ice from melt ponds. These could easily be applied to our output by an interested user.

Line 245. Would this criterion also include sea ice darkened by sediment and algae during the melt season?

Yes, though we have not seen either of these features in the images we processed for this paper.

Lines 252 and 307. Please provide more details on the shadow detection step for panchromatic and multispectral imagery. Does it differ from previous works?

We did not use the shadow category for multispectral imagery. There is not a separate step for shadow detection, per se, rather an additional training category for the machine learning algorithm. We have edited the discussion of the shadow category in section 3.3.2 to reflect that we are not presenting shadows as a classification category. We have also added lines 258-261 to illustrate the differences to previous approaches to handling shadows.

This step does differ from previous work. In Webster et al. (2015), for example, ridge shadows are directly masked and set to the maximum pixel values. Our approach also differs from that in Miao et al. (2015), as our shadow class is not an independent classification in the output and it is only used for images prior to melt onset. Miao et al. (2016) details a more sophisticated ridge and shadow detection scheme

Line 290. Please describe how image dates are used in the classification scheme.

We have edited lines 296-300 to clarify how image dates are used for classification. The image acquisition dates are an attribute that the random forest can use to make a prediction. Image date is a simple means of estimating melt state, which improves the ability of the classifier to correctly predict surface conditions.

Lines 305/491. How does this step distinguish a neighboring ridge from snow-covered ice? It's not clear, does the algorithm identify ridges as a separate class?

Line 491 incorrectly implied that we are detecting ridges directly. We've edited line 491 to clear up this point. While we have methods for indirect detection of ridges (i.e. their shadows – see revised lines 258-261) we do not distinguish ridges from snow covered ice. We have reworked lines 309-318 to illustrate that bright ridge pixels are an example of the benefit of looking at object neighbors, and not a method for creating a ridge class.

Line 310. Here and elsewhere, trade-offs in computational expenses are mentioned. It would be helpful to give a ballpark estimate of the computational expense if possible, e.g., is it O(N) or O(N^2)?

The algorithm is roughly O(N), but it is difficult to quantify the computational expense in big-O terms for this application. High resolution satellite images are quite large, and can easily have millions of image objects. Therefore, any small increase in the time required to evaluate each image object (such as a more complex neighbor analysis) dramatically increases the total processing time.

We have edited lines and 197, 274 to refine our meaning behind 'computational expense'.

Line 313. It's surprising that the Literal Image Derived Products from the Global Fiducials Library have been excluded from this analysis, as these publicly available images have been the data source for several analyses of high-resolution sea ice imagery (Arntsen, Fetterer, Kwok, Webster). Do the authors anticipate that users will find the algorithm suitable for processing this imagery given its radiometric inconsistencies? Why or why not?

We do not anticipate any issues with the NTM imagery from the Global Fiducials Library, and 1m resolution is high enough to get good results (see figures 11 and 12). From an image processing standpoint, the NTM imagery is very similar to panchromatic WorldView imagery, and we therefore do not believe processing the NTM imagery would change the discussions of this paper. In unpublished work Arntsen has tested this algorithm on the NTM imagery with success.

We have added line 151: "The imagery sources chosen for this analysis were selected to be representative of the variation that exists in optical imagery of sea ice, but there is an abundance of image data that can be processed with this technique."

Line 319. Are the different results between experienced and inexperienced users a matter of definition? For example, how do experienced and inexperienced users classify submerged ice near floe edges?

The experience and inexperienced users had the same classification definitions in front of them as they worked their way through the training sets. Though some users might have had different opinions of the surface types on their own, the lack of standard definition is not the reason for disagreement. The definitions of the ice types, including for example submerged ice, were set in advance and provided to all users. The differences arise from the user's ability to interpret the definitions and apply them. We have added a sentence in the paper clarifying this point (line 329). As we established the definition of melt pond to include submerged ice on the edge of a floe ahead of time, users were consistent in their classification of these categories.

Lines 406+. Is this an Eulerian or Lagrangian site? How do the authors distinguish changes due to spatial heterogeneity from seasonal melt progression?

Line 786/Figure 10. Please state whether this site was Eulerian or Lagrangian in the caption.

The site is Eulerian. We have clarified this in the relevant figure captions and added line 505: "The site is Eulerian; it observes a single location in space and does not follow a single ice floe through its lifecycle as it drifts".

Lines 508+. For the aggregate-scale analysis, what type of ice was present in the analyzed scenes? How might the results change based on the presence of different sea ice types?

The images used in the aggregate scale analysis contained primarily first year ice in various stages of melt, and we have noted this in line 522. We have noted in the manuscript in lines 585/591 that this method applies only to melt pond fraction, as we discovered that the images were not large enough to accurately capture ice fraction. Within the melt pond fraction category, we do not believe a different ice type would substantially change the results (lines 594-596), as this analysis is at its core a statistical problem (how to estimate a population based on a sample).

Lines 760/Figure 4 & 779/Figure 8. I suggest presenting the pixel counts as percentages of the total pixels evaluated and providing the total pixel count in the caption for ease of reading.

That is a good suggestion. We have added that information to the figure.

Line 782/Figure 9. In the caption or text, please provide the average scene size.

Another good suggestion, which we have also added to the figure caption.

Line 791/Figure 11. I suggest adding the resolution size as a secondary x-axis on the top of the plots for ease of reading.

We have changed the x-axis to be in units of resolution in meters. The axis is still on a log scale, but as you suggested, it is much easier to read in this format.

Line 52. Morphology seems like the wrong word here.

Changed 'morphology of surface conditions' to 'morphology of surface features'. The morphology of a feature is its structure or form, and here we are discussing the difficulty of lower resolution optical sensors in directly observing the structure of surface features.