Revision of the first submission of
Divine et al:
Regional albedo of Arctic first-year drift ice in advanced stages of melt from the combination of in situ measurements and aerial imagery

Summary
The authors use observations from a drift experiment north of Svalbard in summer 2012 (ICE12) to derive regional sea ice albedo from aerial imagery. The method is based on an image analysis that quantifies areal fractions of different surface types, which are then merged with in-situ albedo measurements to calculate aggregate albedo. Statistical methods are applied to obtain regional albedo values and their uncertainties, as well as to generate Gaussian distribution functions of surface albedo.

General comments
The manuscript is well structured and good to read. The authors use state of the art methods and put their work into context with relevant literature in this field. Graphs, figures and tables are mostly clear and illustrate the presented work well.
Also the scientific content is timely, as improvements of large-scale estimates of surface albedo are strongly needed for various other applications and studies. Surface albedo and its spatial and temporal variability is certainly an important issue to improve our general understanding of polar processes and observed changes; and finally to improve our modelling capabilities. Hence, the topic is highly relevant for "The Cryosphere".

However, I found two aspects that need major revisions and improvements throughout the manuscript:

1) Why do the authors chose to develop and use another / new method to retrieve aerial fractions of different surface types from the images. It is not clear what the advantages are in comparison to the algorithm of Renner et al (2013), a method that was developed in the same group and has been used successfully on almost identical photographs. In contrast, I see a number of reasons to better use the Renner et al approach, because it has different advantages: 1) The direct determination and use of bright and dark pond fractions, 2) the direct treatment of submerged ice, 3) it seems more straightforward to me, and 4) I do not understand the need and advantages of additional simplifications.

2) The authors motivate their work with advances for numerical (coupled) general circulation models (GCMs) and the need of such generalized data sets for different applications. While I agree to this motivation, I do not see how these results effectively contribute to reaching this aim.

- The derived aggregate albedo values for the ice camp region and the marginal ice zone are presented as results for one given time and place. But what can we learn from these numbers for improving our general understanding of albedo feedbacks? How can those numbers explicitly be used in GCM or other large-scale applications?
- The 95% of uncertainty result from in-situ measurements is derived, but how does this result advance albedo studies?

- The dependency of aggregate albedo on observational scales is an interesting and important aspect, but conclusions from this are not drawn or need to be presented more explicitly.
- The application of the bootstrap method is a new aspect with respect to albedo data. How does this help to improve upcoming studies and analyses? What is the real benefit from this or how does this advance also for other applications?

In summary, in the current form, the merit of those results seems rather small, or it has to be presented in a better way. Additional questions, which I suggest to discussed and highlighted more, are e.g.:

- What does this study and their results mean for models and other studies?
- What are the results that advance this field and how can they be used?
- Why should this method be used instead of existing ones?

More specific comments

Title:

- I suggest shortening the title, because the in-situ measurements are only a minor part of the study. Instead the manuscript has a strong focus on developing a new method. I suggest using "sea ice" instead of "drift ice", e.g.: Deriving regional albedo of Arctic summer sea ice aerial imagery

Abstract:

- Reading the abstract, I was uncertain what are the (main) results of this study. They are somehow included, but mixed into methodological or general statements. The wording could be improved to highlight the key results.
- P3700 L18 refers to the "relevance" of the results. But what in which way are they relevant? The last sentence is a general statement, but not specific to this study, it would be good to conclude what the overall benefit of this study / development is. (see also general comment above)

Introduction

- Since the importance of melt pond fractions and their seasonality is highlighted, a short passage that considers remote sensing of melt ponds (seasonality, trends, variability, e.g. Roesel et al) would be a good addition.
- The term of "aggregate albedo" (Perovich et al) could be introduced earlier, because it is explicitly used by various other studies already (e.g. around P 3702, L6). In the same context: What is the difference between regional and aggregate albedo? Does the regional albedo of Arctic FYI (P3702, L18) include open water? It might be useful to give definitions of how these terms are used at some place (methods or introduction) and then use them in this way.
- The last paragraph of the introduction should better highlight the aim and expected results than giving a directory of the coming sections.

Data and methods

- What is the reason to use 2 cameras (instead of only one) and take every second frame of each? This seems to complicate the procedure and increase the effort.
- The distribution of dark and bright ponds is a function of ice type and ice thickness distributions, as well as melt progress into the season. I do not see the advantage of the simplification and using the ratio (r=2.8). In contrast, I consider

the method by Renner et al as better suited. It also solves the problem of submerged ice, which comes up as uncertainties here. Is computing time a critical aspect here?
- I do not feel confident to review details of the statistical methods (Sect. 2.4 and 3.2), hence there are no specific comments on this.

Results and Discussion
- The differences of Flight 6 are related to the position 60km further south in the MIZ. Is it that simple? Having an ice concentration map (see suggested modifications on Fig. 1 below) might make it clearer.

Conclusions
- P3717 L10: "regional aggregate albedo" and "bootstrap albedo" are compared. I understand that both are the same quantities. Hence, they should have the same name (see comment above).
- Comparing albedo observations or results with other studies is certainly difficult and very limited through different regions, seasons, and years. However, here it would be useful to give some more details. How much lower are these values, how are those compared? One more aspect is, that also other studies differ in a way that some provide albedos from single measurements or single surface types, while others cover larger areas or more variability.
- The comparison with Lu et al (2010) refers to "similar latitudes". It has to be considered that the study of Lu et al was performed in the Beaufort / Chuckchi Sea, which has very different climatic conditions in similar latitudes.

Figures and tables
- I suggest to replace the two plates in Figure 1 with one single figure: Combine both plates into one larger map and use an ice concentration map as a background (there is no need for topography information) for the flight tracks. It is also hard to read the current figure, at least on a standard print-out (labels, yellow line, …)
- Fig. 3: Light grey appears as white, at least on my print out
- Fig 4: Replace the abbreviation MIZ with the full term
- If Figure 1 is improved (see above), the maps in Fig. 3 and 4 become obsolete