To the editor and reviewers of The Cryosphere regarding manuscript tc-2019-40:

Please note that this same document has been supplied to each reviewer. Our comments often include more than one reviewer, so it was best for all to see the same response.

General comments and request for extension to implement an update plan:

Thank you for obtaining four reviews for our paper. These will improve the quality of our paper.

Per the editor's request, we have not prepared a revised manuscript at this time. In order to respond to all of the issue raised by referees, we will significantly revise the manuscript. We generally classify the referees' comments as major and minor. The minor comments will be addressed by appropriate additions and clarifications to the paper. The major comments require additional analysis on our part. We will do this new work and incorporate the findings into the manuscript.

We will need a bit of time to submit this revised paper. Given that there was a long delay in receiving comments, our planned revision schedule shifted to later in the year. We propose to complete our update to the manuscript by the end of September.

This document outlines our plan to respond to all reviewers' comments, both major and minor. If this plan is acceptable to you, we will implement it by the end of September.

Authors' response to the range of referee recommendations:

The reviewers have a wide range of recommendations: one requested minor revisions, two requested major revisions, and one recommended rejection.

In addition to validation issues noted by Referees #1 and #2, the rejection recommended by Referee #3 seems to be based primarily on a lack of peer-reviewed documentation of earlier versions. We recognize this issue and agree that it would be better if fuller peer-reviewed documentation had been published earlier, but that was the province of the original PIs of the product.

Now that we—the authors of the present manuscript—have taken over the production, we have worked to improve documentation of the product and by submitting this manuscript, we are seeking a peer-reviewed publication that provides both a comprehensive description of the data set and a specific description of the changes in the product as it was updated from version 3 to version 4. Several specific concerns in the literature – e.g., the problem noted by Szanyi et al. regarding the non-optimal weighting of the buoys that led to "bulls-eye" patterns in the daily motions – are addressed with the version 4 update. With the additional validation and description of the data set that we outline below, it is our intention—also noted by Referees #1 and #4—that this paper shall become the definitive peer-reviewed reference for this data set.

Referee #3 is also critical of the level of documentation for this product. We note that both Referee #1 and #4 disagree with this assessment and are complimentary of the documentation that we included in the manuscript. We note that further details on the basic processing are included in the online NSIDC documentation for the products. We feel the documentation is on par with many other published data sets and is generally sufficient. We also note that several peer-reviewed papers have been published

using the products to investigate science products and that the ice age product has regularly been included in the NOAA Arctic Report Card, which includes a peer review.

If there are some specific areas that you feel are lacking in documentation, we can certainly try to address them. If there are any other specific issues raised by Referee #3 that you feel we need to address, please let us know.

Overview of referee comments:

We have included all comments by all reviewers at the end of this document. Those reviews are color-coded by reviewer. Major comments are highlighted. We judge all other comments to be minor comments, all of which we will address.

Considering all reviewers' comments together, we observe three primary major concerns in common to most of the reviews:

- **Major Concern 1**: The first is the lack of quantitative validation of the motion and age estimate.
- **Major Concern 2**: The second is in how the difference between version 3 and version 4 changes over time.
- **Major Concern 3**: The third is the use of a 1% rule for deriving ice motion from wind speed.

Here, we outline how we plan to address each of these major comments.

Major Concern 1 specifically refers to issues of quantitative validation.

The impetus for this manuscript is primarily to demonstrate that we've improved the processing – particularly in regards to the interpolation of the buoys -- and thus our focus was simply to show the differences between the versions. However, we recognize that the reviewers have legitimate points on showing some quantitative validation. For instance, it will be helpful to compare both sea ice motion and sea ice age to independent observations or proxies.

We plan to add a section on quantitative validation to the paper. For ice motion, we will use independent buoys to derive motion uncertainty statistics (bias, RMSE, etc.) to quantify the difference in uncertainties in V3 and V4. This will be focused on the Arctic. The Antarctic ice motion estimates are considered to be more experimental, and motions there are known to be of lower quality. If we find suitable buoy observations to provide some error quantification with a reasonable amount of effort, we will consider doing that. However, this is lower priority than other issues in our view.

As Referee #1 notes, quantitative validation of sea ice age is not possible. However, we may be able to at least compare first-year vs. multi-year sea ice coverage. This at least quantifies the reduction in FYI bias due to unrealistic convergence noted in Szanyi et al. We will do this via comparison with U.S. National Ice Center sea ice charts.

Major Concern 2 refers to questions of how the data set in version 3 and 4 change over time

The validations outlined above will also help us address the second main comment– how the differences between V3 and V4 change over time. We believe this is primarily due to the change to the buoy interpolation since that was the most significant change. The uncertainty quantification approach noted above will indicate improvement from V3 to V4. So, differences indicate an improvement in V4. If independent buoys exist during requisite periods, we can also compare different years at different periods in the record to quantify how the changes from V3 to V4.

Major Concern 3 refers to parameters used in one of the algorithms used to compute a component of the data set.

The final point is in regards to the use of the 1% rule for wind-derived ice motions. We recognize that this is likely not optimal, particularly with increasing speed. The 1% factor is based on peer-reviewed literature (albeit rather old) and was defined in the product by the earlier PIs. We kept that the same for simplicity while focusing on higher priority issues. The wind-forced motions have low weighting in the OI scheme and generally have a limited influence on the overall motion product. We will do a simple sensitivity test to show the effect of using 1% vs. a more physically realistic 1.5-2%.

Response summary:

We will incorporate our responses to each major comment by either adding new sections to the paper, updating existing sections, or providing an appropriately detailed supplement to the manuscript.

As noted above, we will also address all reviewers' minor comments by clarifying or improving the noted portions of the original manuscript.

Please let us know if you find our plan satisfactory and if we have missed any major concerns that you feel we need to address. Also, please let us know if our proposed time line for submitting a revised manuscript by the end of September is acceptable.

Regards –
Mark Tschudi, Walt Meier, and J. Scott Stewart

For reference, the comments of the four referees for The Cryosphere manuscript “tc-2019-40” are provided here.

Comments are color-coded per referee, and those we identify as major are highlighted. Our authors’ responses below are preceded by “AR:”.

Reviewer #1:

> The paper provides a detailed and very well written description of the new ice motion and ice age product to be delivered by NSIDC. Importance of these products is justified by a comprehensive introduction. Changes in the production chain at all stages - from individual drift components to ice age computation - are properly documented. It is illustrated that the changes at the lower level (new optimal interpolation scheme) have impact at the higher level products (larger extent of older ice) but predominantly in the beginning of the observation period (before 1996).
>
> Notwithstanding the high quality of the paper, in my opinion it fails to quantitatively prove that the ice motion and age products have been actually enhanced. The only evidence that ice motion was improved is qualitative - visual comparison of drift components on figures 3 and 4. Improvement of the ice age product is also illustrated only visually - a more homogeneous ice age distribution is presented on figure 6. Given the high demand for these products a proper quantitative validation is of vital importance. A section needs to be added where the ice motion is compared with other existing independent ice motion products including, for example, AMSR2 derived drift, SAR derived drift, drift of the buoys that were excluded from the optimal interpolation, etc. Although a direct validation of ice age product is probably impossible due to absence of a similar independent product, it can be validated indirectly by comparison of multi-year ice extent with products derived from passive microwave sensors or scatterometers in March - April. It is required to include in this section the widely used product quality metrics such as RMSE, bias, Pearson correlation coefficient, etc. (and preferably both for version 3 and version 4) in order quantitatively prove the enhancement of the products and illustrate applicability in different scientific domains (trend computation, assimilation in numerical models, etc).

AR: Our additional validation work and consequent comparisons between versions 3 and 4 will address these concerns.

> Minor comments No grammar mistakes or typos were identified and the minor comments only concern few clarifications / corrections that are needed in the text.
>
> P3, L23 and L25. Some authors distinguish feature-tracking (detection of individual keypoints on two images => description of keypoints by a binary vector based on => brute-force matching of keypoints, eg. SURG, ORB, etc) from pattern-matching (maximum cross correlation continuously applied to every n-th pixel) [e.g. Rublee et al., 2011, Berg et al., 2014, Korosov et al., 2017]. Maybe a consistent use of “pattern-matching” is preferable in these two cases.

AR: We will reference different feature-tracking methods.

P4, L18. How was the effectiveness of 4X oversampling estimated?

P5, L5. What is the criterion for omitting rogue vectors? Difference from median of vectors in the vicinity? What is the threshold for screening?

P5, L12. What were the thresholds used in V3 and V4 for filtering PMW vectors?

AR: We will update the text to address these issues.

P5, L22. 1% seem to be quite an underestimation of ice drift speed. In addition, this relation cannot be constant in space and time. With the available large amount of collocated data on wind speed and observed ice drift it should be quite simple to illustrate validity of this 'constant 1%' assumption. It would be important to justify it, e.g. in the Discussion section where the relationship between ice drift and wind speed is illustrated spatially and temporally.

AR: We will describe the effect of using a larger wind-ice speed relationship.

P7, L12. Is there a proof that the motion is "largely unbiased"? It is important to add a validation section (as explained in the general comment section above) to prove this statement.

P8, L5 and L6. What is the impact of values of C and D parameters on the drift speed quality (visual appearance) and accuracy (as can be retrieved from validation)? How sensitive are the motion and age products?

P9, L26. I'm confused by the phrase "...all parcels in the 12.5 km ice age grid are initialized with an age-class...". Does it mean that there are several parcels per grid cell? How many?

P10, L19. How much "substantially"? It would be nice to have a numerical characteristics to compare V3 and V4.

P11, L7. I don't quite agree that the difference between V3 and V4 is "fairly consistent over time". It grows from almost 0 (between 1980 and 1986) to almost 1 cm/s (between 2012 and 2017)! It clearly contributes to the difference in drift speed trends between v3 and v4. But which one is more correct? It is very unfortunate that proper quantitative validation is not provided. Maybe this difference is an indication of uncertainty of the motion product and the observed trends are actually statistically insignificant?

AR: We will update the text to address these issues.

Reviewer #2:

Review of "An enhancement to sea ice motion and age products" by Mark A. Tschudi et al. TC-2019-40

General comments:

1) 7-10: Stale opening sentences in the abstract. Reads boring, repeats phrases.

2) Abstract: Suggest to provide more "scientific" results/summary. And drop the first few sentences.

3) The ms is a bit plain and could be lifted by addition of further investigation of the ice-motion and ice-age data sets and discussion of the results.

AR: We will update the text to address these issues.

Specific comments:

1/12-13: Pls specify/give example on how they "are not substantially different between the versions."

1/18: "recent years" or "recent decades"?
1/26: Suggest to rephrase "it is more difficult to draw solid quantitative".

2/26: Correct "will expand greatly with the launch of the NASA ICESat-2 in September 2018" as all this is happened (i.e., it is not longer in the past).

2/27: Could mention Op IceBridge in this paragraph.

3/8-9: Redundant?

3/12-14: Shorten.

3/17: Change "ice motion" to "sea ice motion".

3/25-26: Provide info on typical repeat frequency of "Two geolocated, spatially-coincident, temporally-consecutive satellite images".

4/18: How is the oversampling rate of "4" motivated?

4/34: This statement is not correct as is: "AVHRR was discontinued after 2000." Please qualify or remove.

5/3: How is the threshold of "0.4" motivated?

5/8-15: It is not clear how exactly previous versions dealt with input PM data. Can you separate into composite versus swath or similar?

AR: We will update the text to address these issues.

5/19-24: The assumption that sea ice moves at 0.01 of the wind speed (for the Arctic) needs to be reviewed, especially in an environment of highly variable and increasing wind speeds. –> Underestimate of the ice speed. I.e., Rampal et al. [2009], Positive trend in the mean speed and

deformation rate of Arctic sea ice, 1979–2007, J. Geophys. Res., 114, C05013, doi:10.1029/2008JC005066.

AR: We will present an analysis of overall motion response to an increase in the wind-ice speed relationship.

5/29: Replace "data" in "These buoys monitor meteorological and oceanographic data", i.e., to read "conditions" or "states".

5/33: Mention explicitly that there are too few sea-ice buoys in the Southern Ocean.

8/23ff: It is not clear how the few PM (or combined) motion vectors are treated to derive a broad map of sea-ice motion (on EASE grid)? It appears as if severe extrapolation is taking place.

AR: We will update the text to address these issues.

8/24: There are several experiments with decent buoy arrays available for some parts of the Antarctic sea-ice zone. Why not use some of those to at least assess the skill of the product.. and to possibly explore the suitability of Antarctic ice-buoy data to provide information into the ice-motion product discussed here.

AR: We will use available Antarctic buoys to provide additional validation of the Southern Hemisphere motions.

9/2: The netCDF file should include an additional mask (0/1) where one can mask all gridded ice motion that is "too far" from an actual observation, where the value of "too far" needs to be discussed.

9/32: How is the limit of "16 years" for the maximum ice age set? Physical motivation?o

10/19: There is not quantitative measure of how V4 ice age as improved relative to V3: "there is less “speckling”".

10/21ff: In discussing the relative "ageing" of Arctic ice from V3 to V4 there are no physical details provided as to what process would be the main driver of this change.

11/1ff: The discussion of trends and variability in ice motion & age between V3 and V4 should be more quantitative. – Also, regional contributions should be explored.

12/7: Correct "Fennoscandian peninsula." to "Fennoscandian Peninsula." (upper case)

Fig.7: There seems to be a cyclical signal in the ice-speed difference between V3 and V4. Decadal or perhaps 11 - 12 years. Can different PM sensors be the reason for this? Or the speed magnitude??

Fig.9: The version difference in ice age for 4yr+ is not well explained.

Fig.10 & 11 are not well explained/discussed.

AR: We will update the text to address these issues.

Reviewer #3:

> Review of An enhancement to sea ice motion and age products by Tschudi, M., et al.
>
> Summary: This contribution attempts to illustrate the enhancements - to be understood mainly as extension - of two sea-ice products issued by the National Snow and Ice Data Center (NSIDC), namely the NSIDC sea-ice motion data set and the NSIDC sea-ice age data set. The latter is based on the former. The manuscript advertizes the data sets, informs a bit about the history of these two data sets and describes briefly changes made to the processing which potentially led to an enhancement in quality of both products.
>
> My overall impression is that this paper is not suitable for publication and should be rejected.
>
> It lacks essential information about the retrieval procedure, and the retrieval uncertainties. It further lacks results of an evaluation. It presents trends which seem artificial. It is incomplete in terms of geographical coverage. It contains errors. There are many open technical questions which are not answered in the manuscript and also not in the respective documentation of the data set(s) on the NSIDC web pages.
>
> This paper is written as if it extends a reference benchmark paper where all the required missing details could be found. But this is not the case. Such a benchmark paper does not yet exist for the sea-ice motion product. As the authors stress, the sea-ice motion data set is unique, it has a unique length, and it allows unique applications. And therefore it requires a unique extensive high-quality paper first, in which the reader and the data users can learn about all details and limitations associated with the data set and its generation and evaluation.

AR: We expect that our additional validation work will improve the manuscript so that it serves as the benchmark paper for this data set.

> General Concerns:
>
> GC1: No systematic evaluation of the products has been undertaken - neither for version 3 nor for version 4 of the sea-ice motion product. Also the associated newest sea-ice age data set is not evaluated. In your case, it is not sufficient to just compare version 3 and version 4 of the product because a systematic, detailed evaluation of version 3 products is missing in the scientific literature. There is hence no benchmark against which this new version 4 can be quantitatively referenced. Section 2.2 does not provide new results. There is no indication of a useful sea-ice motion retrieval uncertainty provided along with the product, like is done for sea-ice concentration and thickness data sets. The authors do not present results of an evaluation neither of the newly derived components of the sea-ice motion entering the gridded product nor of the gridded product.

AR: Our additional validation work will permit this evaluation.

GC2: The reader and data set user is informed about user statistics, the importance of the two data sets, some selected bits of the history of the retrievals, and a relatively unspecific description of the changes made to the methods which leaves many open questions. This is, however, potentially not what a reader of this paper and user of this data set would have expected for the following reason: There is no specific paper in which the various retrieval processes, their uncertainties, the caveats of the differentspatio-temporal resolution of the input data sets, a detailed description of the merging (optimal interpolation) approach and its uncertainties have been published so that the full package of detailed, high-quality information is visible at a glance. The retrieval, the input data, the pre-processing steps all these are not transparently described. In other words: A benchmark reference paper containing all bits and pieces is missing so far. And in this context this paper about an "enhancement" seems of doubtful value.

AR: Our additional validation work will improve the manuscript so that it serves as the benchmark paper for this data set.

GC3: The introduction is a nice compilation of recent work dedicated to changes in Arctic sea-ice area / extent, multiyear ice fraction and thickness and in Antarctic sea-ice area and extent. But: during the past two decades or so various other approaches for sea-ice motion retrieval have been developed and the respective data sets are also in use. This paper lacks a review of this work. The retrieval method is not put into context of the current research landscape in this field. This applies to new algorithm developments (both method and input satellite data) as well as evaluation studies. What I, in this context, understand the least, is that despite evidence exists in the literature from various groups using predecessors of the ice motion data set (mainly version 2 and 3), that the inter-sensor inconsistencies cause artificial trends computed from the ice motion product and render parts of the product not useful, you do not comment about this.

AR: The artificial divergence issues raised by Szyani et al. will be specifically addressed.

GC4: I also miss an evaluation of the ice-age data set and/or more quantitative statements about its reliability and potential uncertainty. I do not rate a comparison to the previous version of the data set as providing enough evidence for a proven enhancement. Such a comparison provides only qualitative information about the potential sign of the enhancement. Any quantitative information which would go beyond the comparison to the previous version is lacking.

There is a list of specific comments which I can provide on request if need be.

AR: We will compare both version 3 and version 4 fields against independent observations, as described above.

Reviewer #4:

This paper present a complete documentation of the Polar Pathfinder sea ice drift and age dataset hosted on the National Snow and Ice Data Center as well as the improvement made from Version 3 to Version 4. The main improvement in the drift dataset is in the optimal interpolation scheme used to merge the satellite, buoy and free-drift estimates into one single dataset. The new scheme now use a weighted average of different drift products, with the

weight calculated from their respective errors, and a radius of influence that is based on the decorrelation spatial scale derived from observations. The main improvement in the age dataset is the use of the new ice drift dataset discussed in this paper. Main results include a significant reduction in large spatial gradients at the junction where buoy and satellite products were used in the merged product (mainly in the earlier part of the record where buoy data is more scarce), faster sea ice drift speed and more older ice again in the earlier part of the record (when compared with Version 3).

This paper was long awaited. As stated by the authors, the NSIDC drift and age data is used by several groups, but was lacking a single source in the scientific literature that can be cited describing in details the method used to create the dataset. The paper is mostly well written except in places where some editing is required (see example below).

I recommend that the paper be published after addressing the comments below.

Comments:

Title: Mention polar pathfinder and/or NSIDC in the title.

Page 3, line 13: "...along as well as nothing the changes...". This sentence needs to be rephrased.

Page 3, line 16: Include a table listing all data sources, the time period when the data is used in the merged product and the spatial resolution of the data product.

Page 3, line 19: It should be mentioned here that the optimal interpolation scheme is described in more details below, so the reader knows there is more than just Figure 1 describing the new scheme.

Page 4, line 19: 7.23 cm/sec is still larger than the mean sea ice drift in the Beaufort Gyre. It also means that the satellite drift estimates are not continuous. Later the authors mention that temporal interpolation into a weekly product and spatial (optimal) interpolation smooths the derived velocity field. A sentence should added to tell the reader that this is discussed later on in the paper.

Page 4, line 19: The over-sampling procedure need to be described. Are the images linearly interpolated to get this sub-pixel resolution? If so, what kind of interpolation is used? I would think that the error in the over-sampled images would be a function of the interpolation scheme (linear versus non-linera), etc. Please discuss. Also, how large is the window that is translated to get at the maximum correlation?

Page 4, line 33: Please rephrase in the active form for clarity.

Page 5, line 11: "... and there was NO provenance..." instead?

Page 5, line 19: The spatial resolution of the wind product needs to be stated.

Page 6, line 3: This corresponds to a weighted average drift speed for a 36-hour period with weight 1, 2, 1 for the midnight-noon, noon-midnight, and midnight-noon (next day) drift

estimates. Or are these calculated as (2*Drift(midnight-midnight) + 1*Drift(Noon-Noon)) / 3 for a true 1-day average? This should be clarified “cancel each other” instead.

Page 6, line 4: Give the bias estimates for all products quoted in this paragraph, in a similar manner as for the SSMI daily velocity component (Meier et al, 2000.

Page 7, line 14: Is the SHEBA GPS data part of the IABP buoy dataset? If so, we do not expect a large error because it is used in the polar path finder dataset.

Page 8, line 1: With fixed C values, the sum of all weights used for a given estimate will not add up to 1. An additional formula used to calculate the final (optimally interpolated) drift velocity (including the division by the sum of the weights) should be added.

Page 8, line 1: The error for each drift product (and used to calculate C) should be included in a Table.

Page 8, line 2: I am guessing that D will depend on the satellite product used to calculate the spatial cross-correlation, when in reality, the de-correlation length scale would be only location dependent. This should be clarified. Please state whether D is a constant or if it varies spatially.

Page 8, line 13: Interpolating or averaging? If interpolating, what interpolation scheme (i.e. bi-linear interpolation, kriging, spline) is used?

Page 9, line 5: This is the first time mentioned. What is the resolution of the daily product if not the same as the weekly product?

Page 9, line 6: It is desirable to use the weekly product for all applications, no? I.e. not just for trend estimates.

Page 9, line 21: Put “in year” in parenthesis at the end of sentence for better flow.

Page 10, line 9: Error in drift vectors can lead to convergence of different track into the same grid cell. I believe that the oldest of the tracks is retained in the algorithm. This is another mechanism by which younger ice is lost. It should be mentioned here.

Page 10, line 10: “to spin-up to obtain”. Please rephrase.

AR: We will update the text to address these issues.