Response to Reviewers

Anonymous Referee #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).

Thank you for the useful comments.

We added a section where we validated the combined motion field via comparison with CRREL buoys. This provides an independent quantitative assessment of the combined motion field and the (minor) improvement from Version 3 to Version 4. As the reviewer notes, validation of ice age is not really feasible. We have added a paragraph further discussing validation and also discuss a recent paper (Lee et al., 2017) that included the ice age in an intercomparison with other ice age/type estimates, finding good consistency between them.

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 "patternmatching" is preferable in these two cases.

We changed the phrasing from "feature-tracking" to "pattern-matching" in the manuscript.

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

The 4X method was selected empirically by trying different values and weighing the improvement in motion accuracy versus complexity and computational expense and 4X was found to be the best choice. Text has been added to clarify this.

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

The spatial filtering requires at least two other neighboring grid cells to have displacements within two grid cells of each vector. The two-grid cell limit has been added to the text for clarification.

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

The threshold was the same in both versions, but a bug in the code was discovered that reduced the number of vectors used in the neighborhood filtering process. This effectively increased the number of vectors required for a "good" match to get a valid motion vector, and thus reduced the overall number of retrieved vectors. We've added a sentence to clarify this.

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.

Yes, 1% is likely low. As noted in the text, in the original algorithm development, 1% was chosen based on Thorndike and Colony (1982). But we recognize that this source is likely out of date, particularly in light of observed increases in ice speed relative to wind speed (Spreen et al., 2011). However, as noted in the text, in Version 4, we did not change the wind-ice relationship, though we plan to investigate this further for a future version and will likely change it. We did run the product with a 2% relationship and found that there was very little change in the combined motion product. This is not unexpected since winds are given a low weight compared to the other sources. We have added some text here to acknowledge that 2% is likely more accurate.

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.

We added that the “largely unbiased” assessment comes from previous studies. And we have added a validation section of the combined motions to provide a quantitative assessment.

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?

The values of C and D were derived during the original development of the product and, as noted in the text, were derived empirically. C was derived via comparisons with buoys, while D was derived based on spatial cross-correlations between estimates at different separation distances. In this paper, we effectively examined the sensitivity to C and D by changing the weighting from the closest 15 vectors to the 15 highest weighted. This effectively increased the distance scale for high C observations (i.e., buoys). As described later in the paper, this did have a noticeable effect on the visual appearance of the fields (removing much of the “bulls-eye” seen in Figure 3), but did not change the motion statistics much, as shown in the statistical results in the new validation section. The sensitivity to C is also seen in the test of doubling the wind speed forcing, which had little effect on the overall combined motions because the weight of the winds is relatively smaller.

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?

There can be more than one parcel in a grid cell. Each grid cell is initialized with one parcel, but as parcels advect, two (or more) parcels can merge into the same grid cell.

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

As we’re describing the visual “speckling”, it is a qualitative assessment, not quantitative. We’ve changed the text to note this and removed “substantially” since we don’t quantify the speckling effect. The quantitative effect is shown in Figure 9, showing increased 4+ year old ice in Version 4.

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?

This is a very good point. We've added to the discussion to explicitly note these changes, which are due to the passive microwave sensor changes and their different temporal and spatial sampling, which affects the discretization of the motion estimates. We've added text to further discuss this. And we have added a validation section that quantitative compares the motions from the two versions to independent buoy estimates.

Anonymous Referee #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.

Edits have been made to the beginning of the abstract to remove repetitiveness and to flow better.

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

Quantitative results have been added to the abstract.

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.

We have added a validation section and expanded discussion of the results in various places throughout the paper.

Specific comments:

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

This wording has been changed in response to the above comment about the abstract. The words have been replaced by more precise and quantitative terminology.

1/18: "recent years" or "recent decades"?

We rewrote this as “over the 30+ year time series”

1/26: Suggest to rephrase "it is more difficult to draw solid quantitative".

We replaced “draw” with “determine”.

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).

The text has been updated to reflect the launch of ICESat-2.

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

We added a mention of IceBridge, including a reference to sea ice estimates from it.

3/8-9: Redundant?

Rewrote to eliminate the redundancy.

3/12-14: Shorten.

We removed a sentence to shorten this section.

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

Changed.

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

Added sentence that the typical interval is 1 to 3 days.

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

Sentence added to provide information on motivation/rationale for the 4X oversampling.

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

Rewritten to clarify that the AVHRR discontinued being used in the motion product after 2000.

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

A sentence was added to address this. Different thresholds were tried and 0.4 was found to be best in terms of not being overly stringent (removing "good" matches) while still eliminating most "bad" matches.

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?

No swath data were used. All Tbs are gridded daily composites, including the near-real-time source. We added "gridded" in this section to make that clear.

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.

We added text and the Rampal reference (as well Spreen et al., 2011) to explain the use of 0.01 and to note that this is likely not optimal. We will investigate changing this, but as this paper is documenting Version 4 in which 1% is still used. We conducted an analysis to test the sensitivity to the wind speed and found the changes in the combined motion field to be small. This is not surprising since winds are weighted much less than other sources.

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

Changed.

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

Added.

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.

The motions are combined via optimal interpolation, which uses a weighted distance method, as described at the beginning of the paper. The weighting takes account of the distance and distribution of each vector within the valid distance range of the target interpolation point. Weights are also dependent on the quality of the motion source, but for the Antarctic, all motions are passive microwave. As discussed in this section, if there are too few motions, even the weighted method may not produce realistic fields, which was the case in Version 3 due to incorrectly removing "good" vectors. This was fixed in Version 4 and yields more realistic Antarctic motion fields. We've added text to better explain this.

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.

Yes, there are some Antarctic buoys, such as the IPAB buoys. But the coverage in space and time is much less than in the Arctic. We may consider adding these and other available buoys to the Antarctic fields at some point. Regarding validation, we've chosen not to so for now. We plan to look at these as part of a more comprehensive enhancement for a future major version update when resources allow. The focus of the product has been the Arctic, and age is only produced for the Arctic.

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.

We agree that this would be a good idea and will be considered in the next version of the product. We do note that a flag value is included in the error estimate in the daily combined fields to indicate where all vectors are far from the interpolated grid cell. A sentence has been added to note this.

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

This was designed in the original implementation under the rationale that there is little or no ice of this age. Most of the small fraction of 16-year ice has other ice ages associated with that grid cell. For these cells, the algorithm reassigns the ice age to the oldest age of the remaining ice in the cell that is less than 16 years. Eliminating the tracking of parcels older than 16 years also increases computational efficiency. Note that all ice older than 5 years is treated as one group in the ice age browse imagery, and in analyses we perform, such as the annual NOAA Arctic Report Card.

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

The quantitative assessment is in the change in the proportion of older ice versus younger ice discussed in Section 4 and shown in Figures 9 and 10. We’ve added text to explicitly link these sections.

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.

As noted earlier, the issue in Version 3 is the discontinuity in the buoy estimates and other source motions due to sub-optimal spatial weighting, as found by Szanyi et al. (2016). The discontinuity is largely removed with the improved weighting. We added text here to more clearly describe this.

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.

We’ve added more quantitative discussion on the motion and age products. The main focus of the paper is to present the new version of the products, compare and validate their differences, and briefly show long-term trends as an example of their application. While regional variability is certainly of interest and worth investigating, we feel that it is beyond the scope of this paper.

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

Corrected.

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??

Yes, the change in sensors had an effect on the magnitude due to differing temporal or spatial sampling. We've added text at the beginning of Section 4 to discuss this further.

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

We've added more to the discussion of Figure 9.

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

We have added more discussion of both figures.

Anonymous Referee #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 seaice 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.

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.

There has been significant validation done on the basic algorithm, particularly the MCC approach (Kwok et al., 1998; Kwok, 2008; Meier et al., 2000), as well as on a previous version of the specific product (Sumata et al., 2014; Sumata et al., 2015).

Uncertainty estimates are included in the daily combined motions, based on the optimal interpolation and the relative weights of the source data. We've added mention of this to the text.

Sea ice age is difficult to directly validate as there is not validation with sufficient accuracy and coverage. We have added a reference (Lee et al., 2017) to an intercomparison study that shows good overall agreement between the ice age data and other ice type/age products.

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 different spatio-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.

Such information is provided is the User Guide for the product, provided by NSIDC. We chose not to include this information in the manuscript in the interest of brevity. We've added reference to the NSIDC User Guide. We agree that a peer-reviewed document on the original development of the product would have been useful. However, the original product developer chose not to submit such a paper. As such, our purpose with this manuscript is not to try to recreate a history of which we do not know all of the details, but to document the changes and improvements the current team has made to the newest product as well as giving an overall summary of the processing that is described in the NSIDC User Guide.

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.

We've added references to other ice motion products in the Introduction section. We apologize for this oversight.

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.

Quantitative assessment of the ice age product is difficult because there are no high-quality validation data sets to compare to. Since age is directly derived from the motion product, improvements in motion estimates should carry on to the age product. We also show qualitatively that the spatial distribution of age is more realistic. Finally, we have added a reference that compares the ice age product to other ice type/age products and demonstrates that our product is reasonably consistent with the other estimates.

Reviewer #4

This paper presents 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 uses 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.

Thank you for the useful comments.

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

Added NSIDC to the title.

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

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.

We've added a table, Table 2.

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.

Added.

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.

A sentence has been added.

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-linear), etc. Please discuss. Also, how large is the window that is translated to get at the maximum correlation?

Further details on the oversampling are provided and the window size is noted.

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

This is addressed in response to a comment from another reviewer.

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

Changed.

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

The spatial resolution was added and reference made to the added table.

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

There are two observations, both encompassing 24 hours: midnight to midnight, and noon to noon. These are averaged to get one 24-hour estimate. However, you are correct, the four observations combined span 36 hours.

Page 6, line 33: Write "cancel each other" instead.

Changed.

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.

We've added bias estimates based on Meier et al., 2000 and similar studies.

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.

The SHEBA data are not part of the IABP 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.

We added a note to indicate that the weights are normalized so that they sum to one.

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

These values were calculated early in the development of the original product. There is not a table of errors for each drift product. Errors are discussed for each product in the NSIDC user guide for the motion product: https://nsidc.org/data/nsidc-0116.

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.

D *may depend to some degree on the satellite product, but it's really the correlation length scale of the motions themselves, which is source independent. In any event, it is treated as a constant in the product processing.*

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

It is optimally interpolating, referring back to the method noted in the previous sentence. We added "optimally" and "kriging" for clarity.

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?

The daily and weekly motions have the same spatial resolution (25 km). The weekly average smooths out day-to-day variability due to the coarseness of the resolution. We added some text for clarity.

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

Yes. We've changed the text to note this.

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

Edited to flow better.

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.

We've added text in the previous paragraph to note this.

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

Rephrased for clarity.