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

Page 6, line 33: Write "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.