

We appreciate the comments by Anonymous Referee #2, who has highlighted important areas with additional detail would clarify the work. Since we acknowledge Anonymous Referee #2 has not read the review by and response to Anonymous Referee #1, and hence there is some repetition in our response here to our previously posted response, due to similar recommendations being made. That said, Anonymous Referee #2 has highlighted important areas with additional detail would clarify the work. In the response below, we address each of the main comments by Anonymous Referee #2, and where noted added additional detail to the revised paper.

### Main Comments

1. As noted in the response to Anonymous Referee #1, we apologize for the brevity of the description of the exploratory phase of the analysis. We have documented in the revised text the three methods that were used to identify statistically significant breakpoints in the timeseries described in section 2, as follows:

- We split the record into sub-periods of three years' duration and ranked the differences in the moving averages.
- Using the top ranked differences for non-overlapping time periods, we tested a series of models for significance as follows:

$$y = \alpha + \beta T + \varepsilon \quad (1)$$

$$y = \alpha + \beta \pi_s + \varepsilon \quad (2)$$

$$y = \alpha + \beta_1 T + \beta_2 \pi_s + \varepsilon \quad (3)$$

where  $\alpha$  is a constant,  $T$  is the linear trend, and  $\pi$  is the sub-period mean. The free parameters are the  $\beta$ , and  $\varepsilon$  is minimized.

- In our revision, we implemented the breakpoint detection approach suggested by Rodionov (2004) as an additional check on our results.

As suggested by the referee, we have added these equations and a more detailed description of the analysis process to the Method description, and we agree that this makes the results much easier to follow. As is evident, it was a nested series of decisions that was required to fully convince us fully that a shift in the mean was the appropriate model for the open water time series over a simple linear trend.

2. The self-organizing map is an unsupervised classification technique based on neural networks. There is an extensive literature on this approach, which does not vary appreciably from application to application. Since this classification approach is analyzed in great detail in Lynch et al. [2016] and is not the main focus of the paper, and is analyzed in great detail in Lynch et al. [2016] and hence we chose not to describe the process in great detail in this publication. However, Lynch et al. [2016] was not included in our original reference list. We apologize for this omission and have corrected it in the revised manuscript. In addition, we have provided some more detail regarding the technique, as follows:

“The allocation is achieved by minimizing the Euclidean distance between a vector representing the sea level pressure matrix and the vector representing that node. After each allocation, the entire array of nodes is adjusted to maximize the

Euclidean distance between nodes. A 5x4 array was selected to provide a balance between a practical minimum of nodes and a desirable maximum of variability represented, as described in Lynch et al. [2016].”

Further, we have tried to clarify the clause that the referee found confusing as follows:

“In this analysis, summer daily sea level pressure anomalies (that is, the differences between a July, August or September day and the entire period average for July, August and September)...”

We have now also added the following to the reference list:

Lynch, A.H., M.C. Serreze, E.N. Cassano, A.D. Crawford and J. Stroeve, 2016: Linkages between Arctic summer circulation regimes and regional sea ice anomalies. *J. Geophys. Res.* 121 (13), 7868-7880. DOI: 10.1002/2016JD025164.

3. Because the third regression has two degrees of freedom, the referee is correct in noting that using  $r^2$  alone would result in the selection of the model that has the most free parameters. Hence the use of an adjusted correlation coefficient was required to account for the increase in explanatory variables. Since adjusted  $r^2$  adjusts for the number of predictors in the model, it is smaller than  $r^2$  alone and can become negative. While  $r^2$  is a measure of fit, the adjusted  $r^2$  is a measure of the suitability of alternative models.

### Minor Comments

Page 1, lines 23-24. This is an extremely strange sentence, the meaning of which is unclear.

We appreciate this suggestion for clarity and have rewritten the sentence and provided an appropriate citation.

Page 1, lines 30-31. Krupnik & Jolly, AMAP, and Liu & Kronback are not listed in the References. Page 5 line 14, Meier 2005 is not listed in the References.

We apologize for the missing references. All have been added to the reference list.

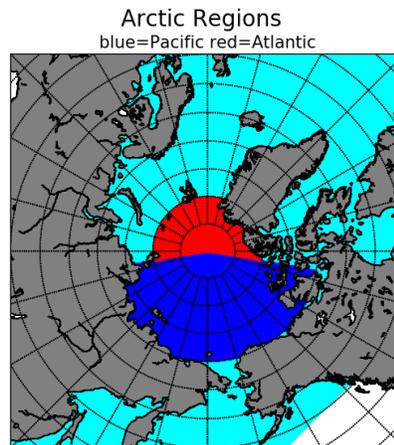
Page 2, line 15. What does "artifacts" refer to here?

While in common usage the word “artifact” tends to mean a human-made object, in this case, we are using the term “artifact” using the second meaning of the word, i.e., detection of a spurious signal arising from extrinsic characteristics of the record, such as the length of a record or the dates on which it starts or ends. (ref, Merriam-Webster dictionary, definition 2.)

Page 3, lines 23-25. A map of the regions would be helpful.

Since the regions are simply longitudinally bound sectors we did not include a map for space reasons, but we have created one (shown below), and will include it in the final revision if the editor requests it. The regional definitions follow the conventions defined by *Gloersen et al.* [1992] and are in use by operational centers as formal analysis areas. We have added this citation:

Gloersen, P., W. J. Campbell, D. J. Cavalieri, J. C. Comiso, C. L. Parkinson, and H. J. Zwally, (1992). Arctic and Antarctic sea ice, 1978-1987: Satellite passive-microwave observations and analysis, NASA Spec. Publ., SP-511.



Page 4, line 3. IABP needs a reference.

Thank you for that point of clarification which we missed. We have preplaced the undefined acronym for the International Drifting Buoy Program with the reference:

Pfirman, S.L., R. Colony, D. Nürnberg, H. Eicken, I. Rigor (1997),  
Reconstructing the origin and trajectory of drifting Arctic sea ice. *J. Geophys. Res. Ocean.*, 102, 12575–12586

Page 5, line 2. "The open water fraction area..." Which one, fraction or area?

We apologize for the error, the correct phrase is "open water area".

Page 5, line 30. I know that  $R^2$  is the squared correlation of the fit, but what is the "adjusted"  $R^2$ ? Page 7 line 16: "had a negative adjusted  $R^2$ " – strange that squared correlation can be negative; what sort of adjustment is done to  $R^2$ , and why?

The adjusted  $R^2$  formulation is commonly used when considering alternative model specifications as in this paper. The specific formulation of adjusted  $R^2$  and the associated discussion of its use can be found in standard statistics or

econometric textbooks. It is always lower than  $R^2$ , and decreases when a predictor improves a model by less than expected by chance. It can be negative.

Page 7, line 2. What does it mean for a time series to be "temporally uniform"?

A time series is temporally uniform when the procedure for creating the time series is constant over time, as for, for example, a re-analysis. In the text, we have clarified our overly efficient terminology.

Page 7, lines 26-27. Concerning possible errors in the SSM/I data, the authors cite personal communication and an article in the Washington Post. Aren't the errors actually documented somewhere?

We have removed the Washington Post article at the request of Anonymous Referee #1. The errors were discovered in the course of the research conducted in this paper. National Snow and Ice Data Center released a corrected version on July 6<sup>th</sup>, after the submission of this manuscript, but we felt it was important to document this problem as it affects the literature that uses the existing data set.

Page 8, lines 1-2. Regarding the ice age anomalies in the first 6 years of the record, "this is likely an artifact of the data product's initialization" – isn't this documented somewhere?

This is documented in the Data section, which describes how the ice age product is initialized (and hence why it may be problematic), and provides the reference for this product (Tschudi et al. 2016).

Page 8, line 33. "A long-term, high quality and temporally consistent record of Arctic open water remains an elusive goal." This is a strange sentence that needs further discussion.

This manuscript and Lynch et al. (2016) describes the ways in which we do not possess a long-term, high-quality and temporally consistent record of ice and open water in the Arctic. We are not sure why this sentence would be considered strange considering the foregoing, but have provided some references at the end of the sentence in the hope that this addresses the referee's concern (Lynch et al. 2016 and Eisenman et al. 2014b).

Page 9, lines 9-10. "the processes associated with these shifts can only be tested independently using a modeling approach" – hasn't this (modeling approach) been done before?

These shifts have not been documented before. The closest effort identified is the Holland et al. (2006) paper, but the time scale of these shifts is much longer than being discussed here (decadal rather than annual). But we are presently starting this modeling work in collaboration with NCAR CSEM researchers.

Figure 1. The scale bar is too small to read.

This has been corrected in the high resolution versions of the figures.

Figure 2. The caption says "open water (%)" but the vertical scale runs from 0 to 0.8, suggesting that it is fraction rather than percent. The legends are too small to read. Why is some data plotted as lines and other data plotted as points? The colors are difficult to distinguish.

The caption been corrected. Some data are plotted as points because they are very similar to the other versions of the record (but not identical). Hence we have included these time series to highlight the differences but as points for clarity.

Figure 3. The axis labels are too small to read easily. The vertical axis is labelled "Open Water Area [%]" but the vertical scale runs from 0 to 0.6, suggesting that it is fraction rather than percent.

This has been corrected.

Figure 5. I can't figure out what the message is here. What is the reader supposed to notice? The colors are difficult to distinguish.

Figure 5 demonstrates the correlation between open water and ice age, as a part of the discussion of likely mechanism for these abrupt shifts, a mechanism associated with a threshold in ice thickness, for which ice age is an imperfect proxy. The colors were selected to avoid problems for the red-green colorblind reader and are the same as Figure 3.