The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2018-151-AC1, 2018

The Cryosphere Discussions

Interactive comment on "Seasonal sea ice prediction based on regional indices" *by* John E. Walsh et al.

John E. Walsh et al.

jewalsh@alaska.edu

Received and published: 30 November 2018

Responses to Reviewers

First of all, we wish to thank both reviewers for their careful reading of the paper and for the insightful and constructive suggestions. In the following pages, we provide point-by-point responses following every comment from the reviewer.

Anonymous Referee #1 Received and published: 4 October 2018

Overview and broad comments

This study by Walsh et al uses a long historical record of sea ice coverage data to estimate the amount of variance explained in two (mainly September) quantities, the

Beaufort Sea Index and pan-Arctic sea ice extent, by the sea ice extent in different Arctic sub-regions. They consider both concurrent correlations and lagged correlations of these quantities with sea ice conditions for previous months, and separate these correlations into a total and interannual estimation. They argue that a piece-wise linear trend is most appropriate for detrending the different time series for estimating the latter.

The noteworthy pieces of information to come out of this study are (1) a piece-wise linear trend is the best option for detrending the time series considered in the study, (2) the break-point year emergent in that analysis which can be interpreted as an acceleration of the ongoing negative trend in sea ice cover is in the mid- to late- 1990's, (3) that interannual variability in the BSI can be explained by June sea ice coverage in the Beaufort Sea and by July coverage in the Chukchi Sea, (4) consistent with other studies the September pan-Arctic ice extent has significant autocorrelation back to July (about two months), and (5) that the Laptev and East Siberian Seas explain the most concurrent correlation in September pan-Arctic SIE.

As it stands though, in my opinion the paper requires substantial revisions and additional analyses before being re-submitted.

The key goal of the study seems to be to provide baseline metrics against which sea ice prediction studies can be evaluated. However, that baseline has already been established for one of the two predictands considered in this study, pan-Arctic SIE, in several other studies based on autocorrelation (i.e. persistence). No reference to these other studies is ever made. The one thing that separates the result about the pan-Arctic SIE predictand in this study from others is that a very long historical record was used here, but that point should be emphasized when motivating the study (it's currently not mentioned). The rest of the lagged correlation analysis for pan-Arctic SIE, which was between it and the SIE for the various subregions, didn't yield higher correlations than lagged pan-Arctic sea ice extent itself. So, for evaluating prediction skill, is it not the autocorrelation of pan-Arctic SIE that is the important baseline to beat? What therefore is the motivation to consider the lagged correlation analysis with the different sub-regions?

We agree that the baseline for persistence-based predictions have been established in previous studies (e.g., Blanchard-Wrigglesworth et al., 2011; Day et al., 2014; Bushuk et al., 2017, 2018), and the revision places added emphasis on the use of a longer record length (back to 1953 rather than 1979) in this regard. The use of autocorrelations and cross-correlations was essentially a vehicle for illustrating the issues associated with detrending in a predictive framework. The main intent of the paper is to show how detrending is a key step in the depiction of persistence-based statistical predictions. We illustrate the effect of detrending for both pan-Arctic ice extent and regional metrics in order to show that predictive applications on both scales must address detrending in a rigorous way, and that there are various alternatives for detrending. While these alternative detrending strategies are known, the relative effectiveness of the various alternatives has not been addressed in previous studies. (Goldstein et al., 2016, The Cryos. Disc.; 2018, Sci. Rep.) come closest by comparing representations based on linear trends and discontinuities in the mean. An addition novel outcome of the present study is the synthesis of break-point information across the various Arctic sub-regions and (in a reviewer-based revision) across seasons.

With respect to the BSI predictand, that analysis seems okay and is fine to report on, but I'm not sure all of the information presented (particularly in the tables; see specific comments below) is needed to reach the conclusions made.

As indicated in the detailed responses below, we are relegating two of the the tables to Supplementary Material.

I was also a bit surprised (and disappointed) to see that only these two predictands were considered if the goal is to provide baseline skill numbers, especially considering the first predictand has already received considerable attention in other studies. Like in the supplementary material of Bushuk et al 2018 (reference below), it would be good

to extend the analysis to treat all of the different regions as separate predictands and compute the autocorrelations for each of those regions. This would help put the results of the Bushuk et al study into context, as they used a shorter and different observational dataset than is used here. As of now, and considering some of my suggestions to remove certain parts of the results section (see specific comments below), this additional analysis would help strengthen that section and would provide useful baseline numbers for future studies.

This suggestion is incorporated into the revision, and it is especially helpful in emphasizing that a novel feature is the use of longer time series than have been used on other recent studies that used only the post-1978 satellite data. In the revision, we present comparisons of the regional autocorrelations for two periods, 1953 onward and 1979 onward, with explicit reference to the Bushuk et al. (2018) regional autocorrelations.

I was missing from the introduction a sufficient argument for the need for this study in the context of current literature on sea ice prediction. The additional references to, and discussion of some key studies listed below are needed to place the goals of this study and its findings into perspective. Specifically here are studies on persistence/autocorrelation of observed and modeled pan-Arctic and regional sea ice extent:

Blanchard-Wrigglesworth, Edward, et al. "Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations." Journal of Climate 24.1 (2011): 231-250. – Includes estimates of autocorrelation for pan-Arctic SIE and SIA in both observations and models. Also includes contributions of lagged Beaufort Sea sea ice thickness to september SIA and SIE variability. – This paper was referenced in the study, but it's odd that only the SST re-emergence result was mentioned.

Sigmond, M., et al. "Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system." Geophysical Research Letters 40.3 (2013): 529-534. – Their Fig. 3 shows persistence forecasts of SIE from two observational datasets.

Bushuk, Mitchell, et al. "Regional Arctic sea–ice prediction: potential versus operational seasonal forecast skill." Climate Dynamics (2018): 1-23. – In their supplementary material, they show autocorrelations for individual regions from observations and compare against models.

These references are indeed relevant. We have carefully gone through each one and identified points of comparison for inclusion (with references) in the revision. As noted above, we also include quantitative comparisons of the post-1953 and post-1979 autocorrelations in order to broaden the temporal context relative to the previous studies. We distinguish between the persistence information derived from models and from observational data, as studies such as Blanchard-Wrigglesworth et al. (2011), Sigmund et al. (2013) and Bushuk et al. (2018) include both. We also add information (including explicit references) to other relevant papers that have addressed trends and autocorrelation issues in a predictive context: Day et al. (2014, J. Climate), Bushuk et al. (2017, J. GRL), Goldstein et al. (2016, The Cryos. Disc.).

Throughout the paper, the word "predictability" is used somewhat misleadingly. It should be made clear early on that what is being referred to as predictability throughout the rest of the paper is a very specific estimate of predictability – that is, the predictability of a certain sea ice coverage quantity based on the lagged cross-correlation or auto-correlation (often referred to as memory or persistence) between that quantity and other sea ice coverage quantities. Otherwise, it sounds like the authors are referring to predictability in general, which encompasses all sources of predictability and estimates of theoretical predictability limits.

Agreed. We have drafted a paragraph, to be inserted following line 82, clarifying that we are addressing only persistence-based statistical predictability based on the observational record, and we will state clearly that there are other sources of predictability, especially those captured by physical-dynamical models of the coupled atmosphere-ocean-ice system. We will insert persistence-based in front of predictability, beginning with line 12 in the abstract. We do note that the original manuscript was "up front" in referring to "forecast skill achieved by other methods such as more sophisticated

statistical formulations, numerical models, and heuristic approaches" (Abstract, first sentence).

Guemas, Virginie, et al. "A review on Arctic seaice predictability and prediction on seasonal to decadal timescales". Quarterly Journal of the Royal Meteorological Society 142.695 (2016): 546-561. – Provides a review on sources of predictability, including memory/persistence of sea ice coverage and thickness. See references therein.

Bushuk, Mitchell, et al. "Regional Arctic sea–ice prediction: potential versus operational seasonal forecast skill." Climate Dynamics (2018): 1-23. – Focuses on regional predictability.

The revised Introduction summaries the conclusions in Guemas et al. (2016) concerning contributions of the atmosphere and ocean to sea ice predictability, and it highlights Bushuk et al.'s (2017) finding that the ocean surface and subsurface state contributes to predictability in the North Atlantic subarctic.

The de-trending analysis could be expanded on. The 1979-onward period is what is commonly used in prediction studies, and there is an ongoing debate in the literature about what de-trending method is most appropriate for pan-Arctic SIE, particularly when the most recent few years are included. Extending the de-trending analysis to focus also on this shorter period would be helpful for future prediction studies. While most authors choose to fit a linear trend over that period, Dirkson et al 2017 suggested a quadratic fit; others suggest de-trending with a high-pass filter but this has the unfortunate effect of removing the first or last sample from the data over an already short period. It would be worthwhile, and would add to the paper, if the authors could make a case for the piece-wise linear trend over this shorter period if indeed it is much better than a linear or quadratic fit. Also, as prediction studies are beginning to focus more on regional sea ice prediction than on pan-Arctic SIE/SIA, it would be helpful to know what choice is most appropriate for the different regions and whether this choice depends on the month or season. For instance, Bushuk et al 2017 and Dirkson et al 2017 argued that linear de-trending was sufficient for regional SIE and local SIC over a similar period considered here.

Dirkson, Arlan, William J. Merryfield, and Adam Monahan. “Impacts of sea ice thickness initialization on seasonal Arctic sea ice predictions." Journal of Climate 30.3 (2017): 1001-1017.

Bushuk, Mitchell, et al. “Skillful regional prediction of Arctic sea ice on seasonal timescales." Geophysical Research Letters 44.10 (2017): 4953-4964.

Following this suggestion, we have expanded on our evaluation of the benefits of a piecewise linear fit relative to a single linear trend line. Specifically, we have identified all cases (across regions and the winter, spring and summer seasons) in which the piecewise fit reduces the residual variance by more than 5% relative to a single linear trend. Only these cases are included in our (revised) summary figure containing – for each season separately – the temporal histograms of the break-point years. The revised figure, containing plots for the three seasons preceding September, will replace the original Figure 5, which combined the results for the three seasons into a single bar plot.

Specific Comments

âĂć It should be stated in the abstract early on (before L11-13) the prediction method that will be used in this paper.

The text beginning on line 11 of the Abstract is modified to: “In this study, we use observational data to evaluate the contribution of the trend to the skill of persistence-based statistical forecasts of monthly and seasonal ice extent on the pan-Arctic and regional scales.”

âĂć L19: “statistical predictability” should be replaced with “statistical skill...” based on said approach. It’s not statistical predictability in the sense of a theoretical limit, which isn’t directly possible to estimate.

Agreed. Will use “statistical skill” here, and in other instances where “predicatability” had been incorrectly used. The previous item (revision of lines 9-11 in the Abstract) is an example of the rewording to “skill” rather than “predicatability”.

âĂć L17-18; Which month(s) are being referred to here?

Revised to say “September trend”.

âĂć L31-32; Statement requires a reference.

The AMAP (2017) reference in the preceding sentence will be inserted here. Can also add a reference to a new paper on Arctic indicators: Box et al. (2018, “Key indicators of Arctic climate change, 1971-2017”, Env. Res. Lett., in press).

âĂć L37-39; Should cite: Maslanik, James, et al. “Distribution and trends in Arctic sea ice age through spring 2011." Geophysical Research Letters 38.13 (2011).

Revised to cite Maslanik et al. (2011) and also the more recent AMAP (2017).

âĂć L47-50; This positive trend in the Bering Sea in winter is disappearing if more recent years are taken into account... Ah, this is stated in the conclusions, but is probably more appropriate to place here instead.

New statement added at end of paragraph (line 51): “The only region in which sea ice shows a weak positive trend has been the Bering Sea (Parkinson, 2014), although the extreme negative sea ice anomalies in the Bering Sea during the past few years have essentially eliminated any positive trend.”

âĂć The departures are also affected by antecedent sea ice conditions themselves. Please see the references given in the general comments above related to persistence.

We assume this comment refers to lines 74-76, where we will add ”„,in addition to antecedent sea ice conditions themselves” to the sentence ending on line 76.

âĂć L79-82; This is not correct. “Ice-ocean model” implies that both the ice and ocean evolve freely as determined by the model (only the atmospheric forcing is re- quired). Also, the statement ignores the fact that fully coupled models, which de- termine both the atmospheric and ocean/ice conditions prognostically, are now used more often for sea ice prediction. These models are also limited by the chaotic nature of the climate system, but this is typically accounted for by running ensembles. There is important literature on predictability using said models that isn't referenced or discussed here. Please see general comment above highlighting predictability literature.

Text on on lines 79-80 is modified as follows: "Even ice-ocean models, which are initialized to current sea ice and ocean conditions, require atmospheric forcing in order to predict future ocean states. Moreover, fully coupled models, which determine both the atmospheric and ocean/ice conditions prognostically, are now used increasing often for sea ice predictions. Ensembles of coupled simulations are generally run because of the chaotic nature of the climate system. These models can be run for much longer time periods than the observational sea ice record, so they can provide statistics of sea ice persistence (autocorrelations) subject to the "perfect model" assumption. Examples of studies employing the "perfect model" approach are Holland et al. (2010. Clim. Dyn.), Blanchard-Wrigglesworth et al. (2011), Day et al. (2014), Bushuk et al. (2018) and Bushuk et al. (2018). In these model simulations, autocorrelation of sea ice anomalies tends to be greater in the model results than in observational data (e.g., Blanchard-Wrigglesworth et al., 2011, their Fig. 2; Day et al., 2014, their Fig. 1)."

âĂć L83-84; It's difficult to see how this paper is an "extension" of Drobot 2003; while this study also focuses on the Beaufort and Chukchi Seas, and extends an analy- sis of statistical prediction skill in that region based on more recent observational data, the analysis done in Drobot 2003 (statistical prediction using multiple linear regression with many more predictors) is not repeated here.

We will delete this sentence, but will retain the subsequent summary (lines 84-92; 95-96) of Drobot's (2003) relevant study.

âĂć L88-92; The caution raised in the Drobot 2003 study I think definitely deserves attention given Arctic sea ice and Arctic climate have changed, as stated. Al- though this study distinguishes between prediction of sea ice coverage with and without the trend, as written it sounds like the authors will carry out an analysis similar to Drobot 2003, and determine the impact of the changing conditions on the statistical relationships found in that study, which is not what is done here.

Lines 91-92 revised to "While the present study will not include the type of multiple-predictor evaluation carried out by Drobot (2003), it will provide a more regionally comprehensive and updated assessment of sea ice anomaly persistence in a predictive context".

âĂć L95-96; This is incorrect. While the Drobot 2003 study didn't consider the effects of detrending as stated, Blanchard-Wrigglesworth et al 2011 did. Specifically they detrended the model projections of SIE by subtracting the ensemble mean at each point in time (removing the forced signal). Additionally, they detrended the SIE observations by subtracting the long-term linear trend.

Sentence in lines 95-96 will be deleted.

âĂć L97-100; I find this overview too vague and is broader in scope than what is actually done in the paper. I think it would be more accurate and helpful to the reader if the authors provided, in order of appearance in the paper, what will specifically be done.

Lines 97-100 will be replaced by the following paragraph containing a more specific summary of the following sections: "In the present paper, we use the autocorrelation statistic to quantify the skill of persistence as a control forecast of pan-Arctic and regional sea ice extent. In addition to utilizing the more conventional metric of ice extent in regional and pan-Arctic domains, we include a regional sea ice index developed in the 1970s to capture interannual variations of marine access in the Beaufort Sea.. A primary focus of the evaluation is the method of detrending the data, as various alternative methods have not been fully explored in the literature. We show that the piecewise linear method generally results in the smallest residual variance about the trend line, and we then perform an across-region synthesis of information on the break-points in the trend lines in different seasons. Our period of analysis extends back to 1953, which results in a considerably larger sample of years than the more commonly used satellite period (1979 onward). Finally, we examine lagged cross-correlations to determine whether pan-Arctic ice extent or Beaufort Sea summer ice conditions are foreshadowed in a statistical sense by antecedent ice conditions in particular subregions of the Arctic."

âĂć L102; "seasonal climatologies, persistence, and trend" of what? Should clarify that you are referring to these quantities for sea ice coverage itself.

Revised to "arising from climatological sea ice coverage, sea ice persistence, and sea ice trends".

âĂć L104; "ice-ocean models" should be replaced with "dynamical models". Ice- ocean models are a specific subset of these, but fully-coupled models are used more often.

Revised to "dynamical models" to include fully coupled models as well as ice-ocean models.

âĂć L105; The Sea Ice Outlook was managed by the Sea Ice Prediction Network, and starting in 2018 became managed by the Sea Ice Prediction Network–Phase 2. The Sea Ice Outlook is not a previous name for Sea Ice Prediction Network. See https://www.arcus.org/sipn/sea-ice-outlook.

Revised to "The Sea Ice Outlook, coordinated by the Sea Ice Prediction Network now in its Phase 2 (https://www.arcus.org/sipn/sea-ice-outlook), provides an annual compilation. . ."

âĂć L118; Suggest changing "maps" to "gridded records". Presumably what is being referred to here is data in digitized form and not the presentation of the data literally as maps.

Will change "maps" to "digitized records"

âĂć L124-127; Say the product name here. The product is described in the next sentence without actually saying explicitly that it is used.

Revised to ". . .we compute ice extent using the gridded Arctic-wide sea ice concentration product known as "Gridded Monthly Sea Ice Extent and Concentration, 1850 Onward (Walsh et al., 2015), referred to in the National Snow and Ice Data Center (NSIDC) catalog as G10010. This dataset is based on observations from approximately 15 historical sources. . .".

âĂć L154-155; While this is true, previous studies have shown that there are shorter persistence time scales for pan-Arctic SIE than for pan-Arctic SIA due to high frequency dynamical influences that change SIE, but not SIA (Blanchard- Wrigglesworth et al 2011; their Fig. 14). This should be mentioned here.

Will insert the following after line 156: "It should be noted, however, that persistence time-scales of pan-Arctic sea ice area have been shown in previous studies (e.g., Blanchard-Wrigglesworth et al., 2011) to be longer than those of pan-Arctic sea ice extent because high-frequency forcing can change ice extent more than it changes ice area (i.e., by converging or diverging ice floes in the absence of ridging or melt)."

âĂć L157ff; Was there any interpolation done to get the G10010 dataset onto the same grid used to define the MASIE sub-regions?

No. The MASIE region mask defines the regions on its grid. For our comparisons, we used the region code of the closest MASIE grid cell for each grid cell in the G10010 grid. No data needed to be interpolated to re-project the MASIE region mask onto the G10010 grid.

âĂć L182; Somewhere early on in the methods section it should be stated what years are considered in the study.

Will change the first sentence of Section 3 to say "As shown in Figure 3, Arctic ice extents have generally been decreasing over the post-1953 period of this study."

âĂć L191-192; Please clarify. I think what is meant is that the presence of a trend in a time series inflates forecast skill when using the anomaly correlation coefficient to assess that skill. It should be noted though that for distance metrics like the root mean square error, skill can actually be inflated by detrending the data if the two time series have different magnitudes of trends.

This statement referred to the use of simple persistence (autocorrelation) of the time series to be predicted, in which case there is no second time series. However, we see the reviewer's point for cases in which a second variable (e.g., another region's ice extent) is the predictor. The reword Lines 191-192: "However, a trend can inflate persistence-based forecast skill when a variable is used to predict itself (assuming the historical trend continues into the future)".

âĂć L198-202; please refer to general comment on de-trending, but this paragraph would be a good place to reference choices made in other sea ice prediction studies.

While the revision includes new references and descriptions of their relevance to the present study (cf. response to this reviewer's earlier paragraph beginning with "The key goal of this study…"), we will also replace the sentence in Lines 202-204 with the following summary: "The previous studies cited earlier (e.g., Blanchard-Wrigglesworth et al., 2011; Sigmund et al., 2013; Day et al., 2014; Bushuk et al., 2017, 2017) have generally relied on least-squares linear fits for detrending. Goldstein et al. (2016, 2018), by contrast, showed that discontinuous changes in the mean better captured time series (such as open water area) characterized by abrupt changes. In the spirit of the Goldstein et al. studies, we explore various options for detrending a time series such as those in Figures 2 and 3, for which the changes are more pronounced in recent decades than in earlier decades. In such cases, a single multi-decadal trend line cannot be expected to optimally represent the historical evolution."

âĂć L209ff; As of now, I find the description of the fitting methodology a bit hard to

follow. It could be more straight-forward both in terms of its description and implementation. If one thinks of the function as a piece-wise linear trend defined by

y=a1x+b1, x < xb (1) and y=a2x+b2, x > xb (2)

where xb is the breakpoint year, continuity of y at xb can be ensured by letting b2= (a1−a2)xb+b1. Substituting b2 into the equation yields 4 parameters to be estimated (a1, b1, a2,xb), which can be done using the 'curve_fit' function. This avoids having to also use both the 'curve_fit' and 'lin_regress' function, as the parameters found by 'curve_fit' can be used to describe the two lines completely. Regardless of which method is used, the equation(s) used in 'curve_fit' should be shown in the paper.

The procedure outlined by the reviewer essentially captures the procedure we used. The "curve_fit" function is defined in lines 504-794 of the file https://github.com/scipy/scipy/blob/master/scipy/optimize/minpack.py, which we will add to the description of our methodology after Line 216. This function performs a least-squares fit to the function by modifying the function's parameters. A starting "guess" of the function parameters is provided by the user. The linear algebra methods of the scipy numerical library is then used. The slopes and break points that emerge will be the same as those obtained by the use of (1) and (2). .

âĂć L227-229; This was just stated on L207-208.

There is a slightly different message in the two statements, with the first (Lines 207-208) referring to single time series while the second (Lines 227-229) refers the across-region comparison of the break points. The revision clarifies this distinction by changing Lines 227-229 to: "Because the break-points are computed separately for each region, the use of the two-piece linear fit allows comparisons of the timing of the break-points across the various subregions".

âĂć L239ff; This is an interesting result, but is it necessary to include the break-point year for all calendar months from January-September in Figure 5? The piece-wise equation seems like a sensible approach to determine a breakpoint year in the summer months (except maybe in the central Arctic). However, for winter and spring months is it really the case that a two-piece linear trend yields a better mean squared error than a simple linear trend? If not, then the breakpoint year for those months won't carry any significant meaning other than that it's an emergent parameter you get from fitting the two-piece linear trend. Is that perhaps what is being seen in Fig. 5 for years before the 1990's and after the early 2000's? Breaking down the information in Fig. 5 for each month and region would make the results easier to interpret and more informative.

The reviewer raises an important point concerning the seasonality of the break-points. In response, we have expanded this section by constructing histograms of break-points for the three antecedent seasons: Jan-Mar, Apr-Jun, and Jul-Sep. The revised Figure 5 now contains three panels, one for each season. (We limited the break-points included in the plots to those for which the two-piece fit reduced the variance about the trend lines by at least 5% relative to a single linear trend line). The result is that the clustering of break-points in the 1990s is more apparent in summer and winter than in spring, although the number of plots satisfying the variance-reduction criterion is greater in the summer than in the winter.

âĂć L250; Is the reason for showing Figure 6 not also to see what regions contribute the most explained variance to pan-Arctic SIE?

Yes. Line 250 is revised to "In order to illustrate the effect of detrending and to show which regions contribute the most explained variance to pan-Arctic sea ice extent, Figure 6 shows. . .".

âĂć L256-260; Why say the significance level for this specific autocorrelation value? Is it the maximum autocorrelation value for all of the samples considered?

We are not sure which "specific autocorrelation" value the reviewer is referring to, as lines 255-256 provide the ranges of our correlations. The confusion may arise from the significance thresholds on lines 256-258, for which we also provide a range that en-

compasses the corresponding range of autocorrelations (from 0 to 0.40). We will insert a statement in Line 258 that "None of the regional or pan-Arctic ice extent correlations exceeded 0.40)."

âĂć L260-262; This is the significance value when the sample has an autocorrelation of zero. Do the detrended time series have no autocorrelation? Also, this statement is true for most regions, but not all according to Fig. 6.

Figure 6 shows cross-correlations between regional and pan-Arctic ice extent, not autocorrelations. The detrended time series have very small autocorrelations (0.15 or less), which results in effective sample sizes greater than 50 (following Bretherton, 1999), in which case the 95% significance threshold increases only from 0.26 to 0.28 – with no changes to our conclusions about statistical significance.

âĂć L263-265; This is true when the trend is included, but not when the time series are detrended according to Fig 6. Please clarify which is being referred to here. Can the authors speculate at all why it is the East Siberian Sea and Laptev Sea explain the most synchronous interannual variance in pan-Arctic SIE? Is it the region that has the most interannual variability in September?

Yes, this statement needed to be revised for clarification, as noted by the reviewer. To be precise: "According to Figure 6, the regions contributing most strongly to September pan-Arctic sea ice variations (including trends) are the Beaufort, Chukchi and East Siberian Seas. After the data are detrended, the regions contributing most to September pan-Arctic sea ice variations are the East Siberian and Laptev Seas". The somewhat surprisingly large contribution of the Laptev Sea is consistent with the "dynamical preconditioning" hypothesis of Williams et al. (2016), which we cite in the revision. The variances of the detrended September extents of East Siberian and Laptev Seas are indeed among the largest of all the regions, although the Chukchi Sea's interannual variance is essentially as large".

âĂć L266-275; Is Figure 7 really necessary? It seems reasonable to consider the contributions of variability in different regions to variability in pan-Arctic SIE, but is there any physical basis for thinking that regions far away from the Beaufort Sea would show any explained variance in the BSI, apart from the trend? It's worthwhile to know how far from the Beaufort Sea variations matter, but those regions are shown in Fig. 9 already for lag 0.

While we are open to removing material (see following comment), we would prefer to retain Figure 7 if possible because it shows that regions of significant explained variance include the Canadian Archipelago to the east, as well as the Chukchi Sea to the west. There is a difference in the "scale of influence" in Figures 6 and 7 that is worth noting, and we will add a statement to that effect (at Line 275).

âĂć Are tables 1-4 necessary?

No, the tables are not necessary except for supporting the finding that a larger fraction of September pan-Arctic variance is explained by antecedent pan-Arctic extent than by antecedent regional extent (see comment below on Figure 8). If the journal allows Supplementary Material, we will relegate Tables 1-4 to the SM; otherwise we will omit Tables 3 and 4, but would prefer to retain Tables 1 and 2 to support the above finding. The text in Lines 276-292 will be modified to summarize the results presently therein without citing Tables 3 and 4 unless those tables are included as SM.

For Tables 1 and 3, when the full time series are compared (non-detrended), the fact that there will be correlation between the different regions and the 2 predictands when both predictor and predictand contain trends is not really a surprise, is it? For Tables 2 and 4, the only relevant information to prediction of the BSI is limited to a few (expectantly nearby) regions, and for the most part to a short number of lag months. For those areas and lags where there is any explained variance over 10% , that information is already plotted in Figs. 8 and 9. Why not just say when describing Figs 8 and 9 that no explained variance values greater than 10% were found in other regions?

It is not surprising that the trends inflate the correlations, although the extent to which

the correlations decrease with detrending might not have been anticipated without actually evaluating the cross-correlations. There is indeed a limited spatial scale of coherence in the BSI results in Tables 2 and 4. We will follow the reviewer's suggestion and shorten the text by deleting Lines 314-320 and replacing it with "The BSI variance explained by all other regions is less than 10%".

If an argument to keep the tables is made, the shading scheme should be explained.

As noted above, Tables 3 and 4 will be deleted from the main text. The two levels of shading simply denote explained variances that exceed 10% and 40%.

âĂć L284ff and Figure 8; I'm surprised to see the values of 0.41 and 0.05 for the Beaufort Sea in Fig. 8 (and identical values for the Canadian Archipelago in the tables 1, 2) for lag months January-April. How can this be if sea ice extent is not variable in the Beaufort Sea and the Canadian Archipelago in the winter?

These values are actually misleading, and they are consequences of very small digitization artifacts. In one of the earlier sources of the pre-satellite era, the winter sea ice concentrations were digitized with a slightly different land mask. This biased digitization impacted both the regional (Beaufort) and Canadian Archipelago) as well as the pan-Arctic ice extents. The revised will contain correlations based on corrected values, and the revised correlations will be zero.

âĂć Figure 8; I think it makes sense to also show the autocorrelation for September pan-Arctic SIE (recognizing it is in Tables 1 and 2; see comment above). It would put the contributions of these other regions into context for prediction purposes. According to Table 2, Pan-Arctic SIE contributes at different lags contributes more to September SIE than any individual region. Should that result itself be mentioned as the most significant (albeit arguably expected) result in terms of prediction? This result should be compared against the studies mentioned in the general comments, but it seems to agree with them. Also, the lagged Laptev Sea SIE contributes more to September Pan-Arctic SIE than the Beaufort Sea, so why not show it too?

As suggested, we are adding pan-Arctic and Laptev Sea panels to Figure 8. The revised text also highlights the result that the pan-Arctic extent of July and August indeed correlates more highly than any regional extent with September pan-Arctic ice extent in both the non-detrended and the detrended data (Tables 1 and 2). This is indeed the rationale for retaining Tables 1 and 2. The finding that the lagged pan-Arctic correlations exceed the lagged regional vs. pan-Arctic correlations is consistent with the perfect-model results in Bushuk et al.'s (2017) Figure 2, which we will cite, although this comparison is not apples-vs.-apples: Bushuk et al. show the skill pf predictions of regional extent (not pan-Arctic extent) in their regional panels. The same is true for Day et al.'s (2014) Fig. 11 and for Bushuk et al.'s (2018) Figs. 6/9/10/11.

A physical explanation for Laptev Sea can be found in Williams et al 2016: Williams, James, et al. "Dynamic preconditioning of the minimum September sea-ice extent." Journal of Climate 29.16 (2016): 5879-5891.

The revision will cite this paper and its explanation (cf. response to comment on Lines 263-265).

âĂć L288-291; "...reaching zero by...". Near zero, not zero exactly. It's also odd that these are not zero exactly when ice covers these regions completely through April

Revision of Line 289: ". . .approaching zero by 3-4 months". The non-zero values are another consequence of the land-mask difference noted above (cf. response to comment on L284 and ff), and the revision will contain correlations based on corrected (temporally consistent) values of ice extent.

âĂć L319-320; Doesn't the BSI contain information from earlier months than September? That is, it's not a strictly September prediction metric.

Yes, the BSI includes pre-September information such as the length of the navigation season in the Beaufort. In the revision, we add the following to Line 320: "This percentage of explained variance is even less than one might have anticipated, given that

the BSI includes information on the length of the navigation season, which can begin well before September, i.e., as early as July in some years".

âĂć L336-338; Specify that this is a general finding; it's certainly not true for all "ice extents" considered.

Revision of Line 336: "Based on the raw (not detrended) time series, the antecedent ice extents in a substantial fraction of the Arctic regional seas provide significant predictive skill. . .".

âĂć L338-340; "(and even years, given the multidecadal scales of the trends)"... While probably true, this isn't a direct conclusion/finding of this study.

Will delete the parenthetical phrase, so the revised statement is limited to ". . .the regional extents of prior seasons".

âĂć 342-345; The study being referenced here is based on six years of hindcasts which were created by averaging over many different techniques/models, and the results are not reflective of the predictive skill shown by several other prediction studies. To put the results of this paper into context, it should be stated that numerous other studies have shown higher skill than anomaly persistence forecasts (which is essentially the method used here). See references in the introduction of Bushuk et al 2018 (reference in general comments).

We will grant that other sea ice prediction efforts have outperformed persistence and will state this explicitly, citing Bushuk et al. (2018) and studies such as Tivy et al. (2011), Shroeder et al. (2014), Yuan et al. (2016) and Petty et al. (2017). However, in this context, perfect-model studies do not seem to fit the point being made; persistence-derived predictability is greater in perfect models than in corresponding operational forecasts, as even some of the perfect-model studies show. In this respect, the SIPN is the acid test of the current state of sea ice prediction (at least for September pan-Arctic ice extent). A compilation of SIPN results for the past eight years (the "state of the art"), completed just in the past two weeks by Larry Hamilton, shows that, on balance, the SIPN consensus forecasts outperform detrended anomaly persistence by only a small amount. While that persistence metric is based on year-to-year September variations, the SIPN forecasts for September are made in June, July and August – less than a season ahead, and on the favorable side of the springtime "prediction barrier". So, while the revised text will affirm the reviewer's point, there are some caveats to the claim that numerous studies have shown higher skill than anomaly persistence.

âĂć L355; The influence of trend predictive skill is well recognized in the sea ice prediction community, and skill results based on detrending are commonly presented in sea ice prediction studies, so I'm not sure I see how challenge (1) follows from this particular study.

The reviewer makes a valid point here. The revised text rephrases this "challenge" so that it is merely a call to make optimum use of persistence as a control forecast. In that respect, we mean going beyond simple period-of-record linear trends in in deriving persistence forecasts and in defining anomalies relative to trends.

Technical corrections

âĂć Figure 3: I think it would be easier to see and compare magnitudes of the anomalies and trends if Fig 3 were split into two sub-panels: one for the Beaufort and one for pan-Arctic SIE. The sub-panel for pan-Arctic SIE could have double y-axis (one for March, one for September) so that there isn't such a large vertical space between lines.

Agreed. We have remade Figure 3 into two panels as suggested.

âĂć L194-197; Sentence is worded awkwardly, specifically the "Because..., so ...".

Revision into two sentences: "One of our main interests in this study is whether or not interannual variations of preceding regional ice extents correlate with later BSI values. In order to exclude the effect of the overall trends in the correlation of these time series, we detrend the data and explore various methods for doing so".

âĂć Figure 4: It looks like there is a kink in the piece-wise linear functions near the breakpoint. This shouldn't be there.

Yes, we see that in the lower panel and have corrected the two-piece linear trend line. Thank you for catching the glitch.

âĂć L255; "accrual".. what is accruing here? Why not just say "corresponding to correlations..."

That was a typo. Should have been "actual", but "corresponding" is even better.

âĂć L257 and 258; typo? "a 60-year samples" should be "a 60-year sample"

Revised by deleting "s" at end of "samples".

âĂć L296; Change to ..."when sea ice extent" for the Beaufort and East Siberian Seas are predictors. The seas themselves aren't predictors.

You are right. We have revised to say that sea ice extents in those seas are the predictors..

âĂć L319-320; incomplete sentence

"that" (second word) needed to be deleted.

âĂć L333-335; "(increases of trends)" not needed as it is explained what is meant by break-points in the second part of the sentence.

"(increases of trends)" deleted in revision.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2018-151, 2018.