

## ***Interactive comment on “Seasonal sea ice forecast skills and predictability of the KMA’s GloSea5” by Byoung Woong An et al.***

**Anonymous Referee #2**

Received and published: 7 January 2019

This work investigates the seasonal prediction skill of Arctic sea-ice extent (SIE) and sea-ice thickness in the Korean Meteorological Administration’s version of GloSea5. The authors use a variety of skill metrics to attempt to quantify the seasonal prediction skill of this system. They also claim that prediction skill is determined primarily by sea ice drift, freshwater fluxes, and sea ice thickness. The topic of this manuscript is appropriate for the Cryosphere and has potential to be of broad interest within the sea ice community.

This manuscript is generally poorly written and the scientific argumentation is difficult to follow. While having the potential to be an important study, the manuscript in present form suffers from a number of significant flaws, which I expand upon below. These flaws are sufficiently important that my recommendation for this manuscript is rejection.

C1

Note: I will use the convention p.l throughout this review to refer to page number p and line number l of the discussion paper.

Major Comments:

1) The Authors claim that “Hindcast covers for the period from 1991 to 2010 and three ensemble members are initialized on fixed calendar dates, i.e., 1st, 9th, 17th and 25th.” However, from what I can gather it appears that they only consider hindcasts covering the period of January 2016 to February 2017. This implies that the authors only have one, or at most two, verification dates with which to evaluate the skill of this system. It is simply not possible to evaluate a prediction system’s forecast skill with so few verification dates. The authors make seemingly generic statements about the system’s prediction skill throughout the manuscript (and in the title), however these statements are invalid given the extremely short hindcast period considered. The authors must either (i) expand the number of hindcast years so that statistically robust statements about prediction skill can be made; or (ii) recast this work as a seasonal prediction case study focusing on the year 2016. If they opt for (ii) they should strictly avoid using the term “forecast skill” in the manuscript.

2) Given that the results presented in this manuscript do not represent a robust prediction skill assessment, these findings add little to the existing literature on seasonal Arctic sea ice prediction. The authors rightfully mention a number of studies that have used hindcasts to examine prediction skill (each based on a set of hindcast experiments spanning at least 20 years), but make little to no effort to place their results in the context of these earlier works.

3) Many aspects of the methodology and the results of the manuscript are unclear, including:

3.17: It is unclear what initialization dates, what ensemble sizes, and what time period are used for these hindcasts.

C2

5.5: It is unclear what is being shown in Fig. 1. Is this the ensemble spread from forecasts initialized on Jan 1, 2016? The forecasts only run for 240 days, so are multiple initialization dates shown in this figure? This is not clear from the text or caption.

5.17: As with Figure 1, it is unclear what is being plotted in Fig. 2. What initialization dates are being used here? Also, what units are used in Fig. 2b? RMSE should be in  $\text{km}^2$ .

8.3: Ice that is thicker than 30cm is commonly seen near the Siberian coastline in winter, which contradicts the authors' statement here.

8.13: Where is this snow depth verification data coming from? My understanding is that CryoSat-2 uses the Warren climatology for snow depth, which by definition does not have any interannual variability.

11.3: It is unclear how this correlation is being computed. There is only one verification period available for each of these forecasts. How are the authors defining a correlation here? What quantities are being correlated?

11.7: What SIT data is being used for the months of May and June (no CryoSat-2 data is available in this period)?

13.5: Why is the CRPSS so drastically different than the correlation and generalized discrimination scores?

14.23: It is unclear what figure in the manuscript supports this statement.

14.28: This statement is unsupported by the results of the manuscript.

15.2: It is unclear what this statement means.

15.5: It is unclear how the manuscript has shown this.

15.6-7: This has also not been shown.

15.18: The results do not show this.

C3

4) The manuscript makes no connection with previous work on sea-ice prediction performed with the GloSea system. In particular, Peterson et al. (2015) assess Arctic sea ice prediction skill using the GloSea4 system. Comparisons should be made to the model biases and forecast skill results of Peterson et al (2015). Also, information should be provided on the key differences (in terms of both initialization and model formulation) between GloSea4 and GloSea5. The authors should attempt to relate any skill differences to the differences between these systems.

References:

Peterson KA, Arribas A, Hewitt H, Keen A, Lea D, McLaren A (2015) Assessing the forecast skill of Arctic sea ice extent in the GloSea4 seasonal prediction system. *Clim Dyn* 44(1–2): 147–162

---

Interactive comment on The Cryosphere Discuss., <https://doi.org/10.5194/tc-2018-217>, 2018.

C4