Analyzing the impact of CryoSat-2 ice thickness initialization on seasonal Arctic Sea Ice prediction

Richard Allard[1] , E. Joseph Metzger[1], Neil Barton[2], Li Li[3], Nathan Kurtz[4], Michael Phelps[5], Deborah Franklin[5], Ole Martin Smedstad[5], Julia Crout[5] and Pamela Posey[5]

[1]Ocean Science Division, U.S. Naval Research Laboratory, Stennis Space Center, MS, USA; [2]Marine Meteorology Division, U.S. Naval Research Laboratory, Monterey, CA, USA; [3]Remote Sensing Division, U.S. Naval Research Laboratory, Washington, DC, USA; [4]NASA, Cryospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA and [5]Perspecta, Inc., Stennis Space Center, MS, USA

Abstract

Twin 5-month seasonal forecast experiments are performed to predict the September 2018 mean and minimum ice extent using the fully coupled Navy Earth System Prediction Capability (ESPC). In the control run, ensemble forecasts are initialized from the operational US Navy Global Ocean Forecasting System (GOFS) 3.1 but do not assimilate ice thickness data. Another set of forecasts are initialized from the same GOFS 3.1 fields but with sea ice thickness derived from CryoSat-2 (CS2). The Navy ESPC ensemble mean September 2018 minimum sea ice extent initialized with GOFS 3.1 ice thickness was over-predicted by 0.68 M km^2 (5.27 M km^2) vs the ensemble forecasts initialized with CS2 ice thickness that had an error of 0.40 M km^2 (4.99 M km^2), a 43% reduction in error. The September mean integrated ice edge error shows a 18% improvement for the Pan-Arctic with the CS2 data vs the control forecasts. Comparison against upward looking sonar ice thickness in the Beaufort Sea reveals a lower bias and RMSE with the CS2 forecasts at all three moorings. Ice concentration at these locations is also improved, but neither set of forecasts show ice free conditions as observed at moorings A and D.

1. Introduction

Seasonal sea ice prediction (Merryfield and others, 2013; Stroeve and others, 2014; Blockley and Peterson, 2018) is gaining in importance as operational ice production centers (e.g., National Ice Center (NIC), Environment and Climate Change Canada) are fielding requests to provide extended-range forecasts of sea ice conditions to support navigation, in some cases months in advance. The NIC has been using the U.S. Navy's pre-operational 45-day fully-coupled modeling system to support ICEX 2018, and the McMurdo Resupply missions in the Ross Sea in 2018 and 2019. The Navy ESPC has provided model output (both Arctic and Antarctic) in support of the Year of Polar Prediction (YOPP, Allard and others, 2018a) and are regular contributors to the Sea Ice Prediction Network (SIPN2, https://www.arcus. org/sipn). In addition to ice thickness and ice concentration forecasts, ice drift forecasts can be used to monitor the expected movement of the ice edge within the Marginal Ice Zone. We are regular contributors to Sea Ice Drift Forecast Experiment (https://www.polarpredic-tion.net/yopp-activities/sidfex/) and the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC, https://www.mosaic-expedition.org/) projects.

With the availability of new sources of satellite-derived ice thickness (derived from freeboard) products, it is prudent to incorporate observed ice thickness measurements into operational modeling systems to improve both the short-term (e.g., days) and long-term (months) forecasts. In this paper, we describe the U.S. Navy's efforts to improve the prediction of the September Arctic sea ice mean and minimum by reinitializing a fully-coupled global modeling system with CryoSat-2 ice thickness data.

A global coupled climate model was used by Day and others (2014) to examine the impact of Arctic sea ice thickness initialization on seasonal forecast skill. A series of experiments were performed by initializing the sea ice thickness on 1 January and 1 July vs control simulations initialized with climatological ice thickness. They found a significant improvement in forecast skill for the September sea ice extent and lower RMSE for the 2 m air temperature vs the climatological experiments. Blanchard-Wrigglesworth and others (2016) examined model uncertainty and the predictability of the September 2015 sea ice extent by initializing eight dynamical models (both global and regional) with the same ice thickness field from the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) (Schweiger and others, 2011) for 1 May 2015. They found a large divergence in the September sea ice extent for all the models that varied by several million km^2. Performing post-processing of the model forecasts with a bias correction using each model's climatology resulted in much improved forecasts for sea ice extent and sea ice volume.

cambridge.org/aog

Williams and others (2016) examined seasonal forecasting of the Arctic sea ice extent with a dynamical mechanism based on winter preconditioning. A Lagrangian trajectory model was used to backtrack the September sea ice edge to any date from the previous winter. They found a 38% reduction in error when using a multivariate regression model of the September sea ice extent based on ice export from the peripheral Arctic seas and Fram Strait ice export. Using the CICE model, Schröder and others (2014) found skill in predicting the September Arctic sea ice minimum based on the spring melt pond fraction.

Seasonal prediction using the Climate Forecast System Reanalysis (CFSR) was examined by Collow and others (2015) where they performed a set of 9-month ensemble predictions for the period of 2005–2014. Two sets of experiments were performed in which initial March ice thickness was provided by PIOMAS and CFSR. They found that the use of PIOMAS as a proxy for satellite-derived ice thickness (e.g., CryoSat-2) performed the best in predicting the September sea ice extent minimum as demonstrated by increased correlation and lower RMSE. Dirkson and others (2017) utilized the Canadian Climate Model version 3 (CanCM3) and three statistical models used to derive initial sea ice thickness estimate to initialize a real-time forecasting system for the period of 1981–2012. They found that the combination of sea ice thickness fields that represented the thinning of the ice pack over many years and interannual variability led to the best predictive skill for pan-Arctic ice area and regional sea ice concentration.

The predictability of the September sea ice minimum by assimilating CryoSat-2 (CS2) ice thickness data in the UK Met Office's coupled seasonal prediction system GloSea was examined by Blockley and Peterson (2018). Seasonal predictions were initialized on three different spring start dates for an eight-member ensemble prediction system for the years 2011–2015. They found a significant improvement vs a control set of ensembles for the same time periods (without using CS2 data) for sea ice extent and ice edge location for September forecasts as well as improvements to near-surface air temperature and pressure fields.

Yang and others (2019) performed Arctic seasonal ice forecasts for the period of 25 May–1 October 2016 using the 11-member ensemble based Sea Ice Seasonal Prediction System comprised the MITgcm ice-ocean modeling system with 18 km resolution. Initial conditions were provided by restart fields from an operational ice-ocean modeling system which assimilates sea ice concentration and CS2/SMOS ice thickness. The study showed a significant improvement in predicting the Arctic sea ice extent for the Pan-Arctic. The best performance was found for the Central Arctic, Canadian Archipelago and Baffin Bay/Davis Strait/Labrador Seas. An evaluation of the ice thickness prediction at the Beaufort Gyre Exploration Project (BGEP) Upward Looking Sonar (ULS) moorings in the Beaufort Sea showed a significant reduction in the RMSE at all three moorings compared to the operational Climate Forecast System Version 2 (CFSv2) ensembles which were not initialized with satellite-derived ice thickness data.

2. Modeling system used

In this study, we use the Navy Earth System Prediction Capability (Navy ESPC) (Barton and others, 2019), which is a global coupled atmosphere – ocean – sea ice prediction system developed by the Naval Research Laboratory. The component models include: atmosphere – NAVy Global Environmental Model (NAVGEM); ocean – HYbrid Coordinate Ocean Model (HYCOM) and sea ice – Community Ice CodE (CICE, Hunke and Lipscomb, 2008). To couple the models, we use the National Unified Operational Prediction Capability (NUOPC) tools built on top of the Earth System Modeling Framework (ESMF). The sea ice

Table 1. Model fields exported to atmosphere, ocean and ice models for coupling

NAVGEM	HYCOM	CICE
u-wind @10 m	Sea sfc temp	Ice concentration
v-wind @10 m	Sea sfc salinity	Ice stress x-dir
Air temp @2 m	Sea sfc u-ocean	Ice stress y-dir
Specific humidity @2 m	Sea sfc v-ocean	Shortwave flux at ice bottom
Air density @2 m		Longwave flux at ice bottom
sfc air temp		Salt flux at bottom
Downward shortwave		Water flux at bottom
Downward longwave		u-comp ice velocity
Precipitation		v-comp ice velocity
Mean sea level pressure		Ice thickness
		Ice albedo
		Ice sfc temp

and ocean models are configured on the same tripole grid with ∼3.5 km resolution at the North Pole. HYCOM employs 41 vertical levels in the water column. NAVGEM has a 37 km horizontal resolution with 60 vertical levels. Navy ESPC is presently undergoing operational testing to support the generation of deterministic (0 to 16 days), and probabilistic (0 to 45 days) forecasts. For the coupling between HYCOM to CICE, HYCOM exports sea surface temperature (SST), sea surface salinity (SSS) and the u- and v-components of surface ocean current (SSU, SSV) to the model coupler, which makes it available to both the atmosphere and sea ice model. Table 1 summarizes the model fields exchanged for coupling.

In the control runs (Pegion and others, 2019), ensemble forecasts are initialized from the operational US Navy Global Ocean Forecasting System (GOFS) 3.1 (Metzger and others, 2014) for the ocean and sea ice using the Navy Coupled Ocean Data Assimilation (NCODA, Cummings and Smedstad, 2014) system that assimilates Special Sensor Microwave Imager/Sounder (SSMIS, Maslanik and Stroeve, 1999) and JAXA Advanced Microwave Scanning Radiometer2 (AMSR2, Comiso and others, 2003, Comiso and Nishio, 2008) sea ice concentration products for the cryosphere. Atmospheric initial conditions are from operational NAVGEM (Hogan and others, 2014) using the Naval Research Laboratory Atmospheric Variational Data Assimilation System (NAVDAS) (Baker and others, 2007). The ensemble-based forecasts are generated using a time-lagged approach in which the first ensemble member is initialized on 1 May 2018 and run through the end of September. This is repeated for 2–10 May 2018, in which a separate set of forecasts are run through the end of September. The time-lagged initialization is short enough (10 days) that it does not have an impact on the results presented in this study.

Another set of experiments are initialized from the same sea ice restart files above, but with ice thickness derived from CryoSat-2 (Kurtz and Harbeck, 2017) for the 10-day period beginning 1 May 2018. We use the technique to reinitialize the CICE ice thickness field from CS2 as described by Allard and others (2018b). We calculate the anomaly between the observed ice thickness field (on the model grid) from the control experiment for a particular date (e.g., 1 May 2018). The anomaly is added to the ice thickness field for the CS2 experiment. If the observed ice concentration indicates ice is present, while the CS2 ice thickness field indicates no ice is present, the data assimilation system adds 0.5 m of ice at these locations. For areas outside the Central Arctic (e.g., Barents Sea), the mean snow depth of first-year ice from climatology (see Tilling and others, 2015) is used in the CS2 retrievals.

Both sets of experiments are performed with 10 time-lagged ensemble members for the period of 1–10 May 2018 run through 30 September 2018. We present results for the predicted vs
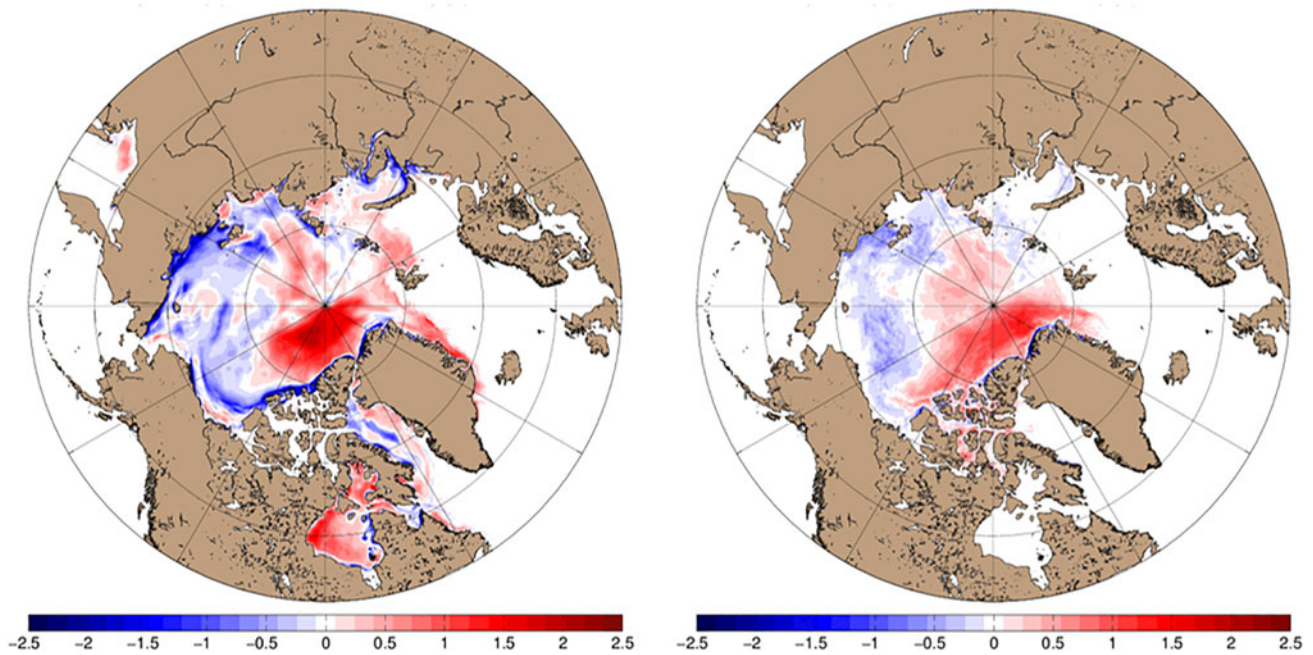
Fig. 1. Ensemble mean ice thickness difference (*m*) between forecasts initialized with CS2 and control forecasts for (left) 15 May 2018 and (right) 15 September 2018. Blue shades indicate that CS2 has less ice than control.
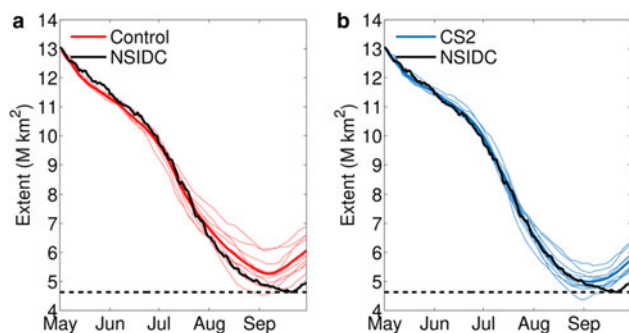

Fig. 2. (a) Arctic sea ice extent (M km^2) for the control run (red) vs (b) the runs performed with CS2 initialization (blue). The thick dark red/blue lines represent the ensemble mean, while the thin lines are the individual ensemble members. NSIDC data are shown in black. Dashed black line represents observed 4.63 M km^2 minimum extent on 23 September 2018.

Fig. 3. (a) Arctic sea ice volume (1000 km^3) for the control run (red) vs (b) the runs performed with CS2 initialization (blue). The thick dark red/blue lines represent the ensemble mean, while the thin lines are the individual ensemble members. PIOMAS results are shown in black.

observed September sea ice mean and minimum extent to examine the impact of satellite-derived ice thickness initialization vs the set of experiments which did not utilize this data. Figure 1 presents the difference between the forecasts with and without CS2 initialization on 15 May and 15 September. Albeit early in the seasonal forecast period, the model's ice thickness bias compared to CS2 is evident on 15 May. In both figures, the CS2 forecasts show thicker ice is evident along the Canadian Archipelago with thinner ice, especially on 15 May, along the Canadian, U.S. and Russian coasts.

3. Results

Figure 2 depicts the sea ice extent for all ten ensemble members for both sets of forecasts. The sea ice extent ensemble mean shown in Figure 2a closely matches the observed extent through 29 July; afterwards it gradually shows a positive bias that remains for the duration of the forecast period. Figure 2b shows that the ensemble mean is closer to the observed minimum ice extent, but ~18 days early. The CS2 ensemble mean forecast sea ice
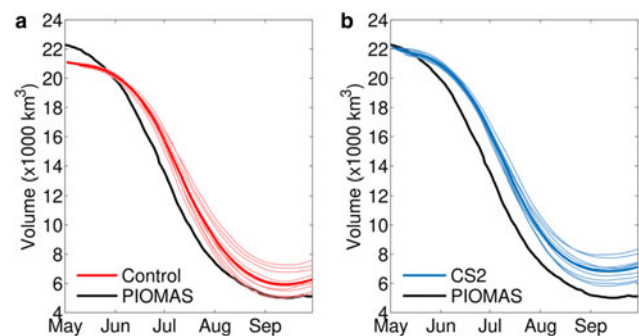
extent shown in Figure 2b closely mirrors observations through 31 August (~120 days) before showing a positive bias for the remainder of the forecast period. The Navy ESPC ensemble mean September 2018 minimum sea ice extent initialized with GOFS 3.1 ice thickness (control) was over-predicted by 0.64 M km^2 vs the ensemble set of forecasts initialized with CS2 ice thickness which had an error of 0.36 M km^2 (observed mean extent was 4.63 M km^2), a 43% reduction in error. The National Snow and Ice Data Center (NSIDC) minimum (Fetterer and others, 2017) was observed on 23 September, while the Navy ESPC minima occurred on 5 September for both sets of forecasts.

Figure 3 depicts the Pan-Arctic sea ice volume (1000 km^3) for the 5-month forecast period vs the PIOMAS sea ice volume estimates. Although actual Pan-Arctic sea ice volume data does not exist, PIOMAS is generally used for comparison and is widely used by the sea ice modeling community. The control run shown on the left begins with a lower volume vs PIOMAS but is higher beginning in early June and maintains a positive bias for the remainder of the forecast period. The CS2 ensemble forecasts initial volume shown in Figure 3b is in excellent agreement with PIOMAS but by mid-May begins to trend higher, with an
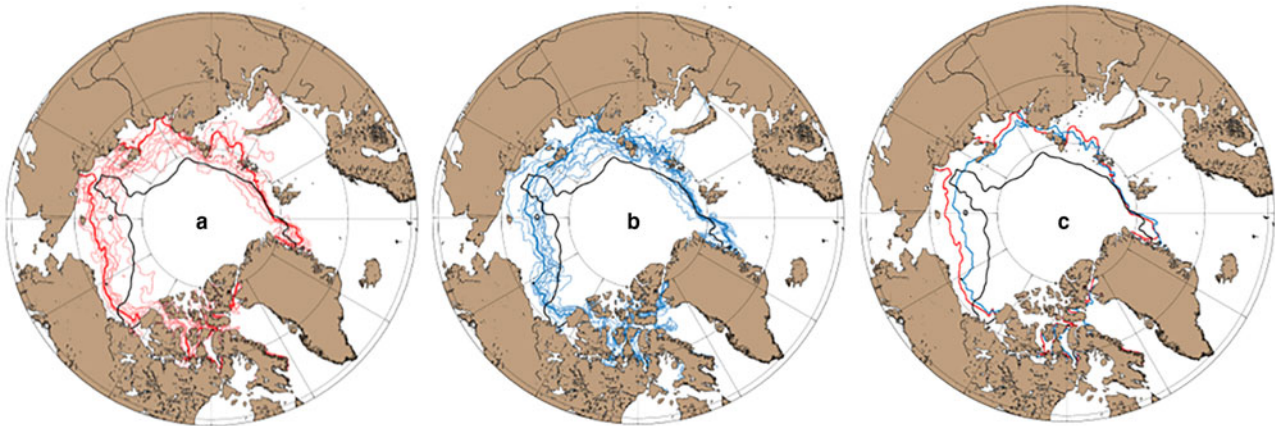
Fig. 4. (a) September mean sea ice extent prediction for ten ensemble members from control run; dark red line denotes ensemble mean. (b) Same as (a) but based on CS2 initialization; dark blue line represents ensemble mean. (c) Ensemble mean for control (red) and CS2 (blue). Black lines denote NSIDC observed mean September extent.

Fig. 5. Arctic regions used for the IIEE analysis.

overall higher volume compared to the control run shown in Figure 3a. We attribute this higher volume to the thicker initial ice shown in the Central Arctic between the Canadian Archipelago and the North Pole shown in Figure 1 where CS2 ice thickness is 2–2.5 m thicker than the control run. An examination of NASA Operation IceBridge (OIB) ice thickness overlaid on CS2 thickness for early May (not shown) reveals the CS2 ensemble forecasts have significantly thicker ice compared to the OIB data. We attribute this to the complex scattering behavior from CS2 that can lead to anomalous features in the thickness fields. Therefore we have more confidence the thinner OIB ice thickness is more realistic than what we see from CS2 ensemble forecasts in this region.

The predicted September mean sea ice extent is shown in Figure 4. The solid red and blue lines represent the ensemble mean and the black solid line depicts the NSIDC mean ice extent. Figure 4c depicts the ensemble mean ice extent for the control and CS2 forecasts. The most noticeable improvement is found in the Beaufort/Chukchi Sea. Overall, some improvement in shown in all regions except off the coast of northeast Greenland

and near Banks Island (73.08°N, 120.72°W, see Fig. 7 for more detail) where there are significant differences between the observed and predicted ice edge.

We calculate the Integrated Ice Edge Error (IIEE, Goessling and others, 2016) to assess the skill in the control and CS2 forecasts for regions shown in Figure 5 to determine how well the models agree on the 15% sea ice concentration isoline. Model data are interpolated to the 25 km NSIDC ice concentration grid. Figure 6a shows the IIEE for the Pan-Arctic domain. Both model forecasts easily beat anomaly persistence (based on 20-year climatology from the period of 1998–2017) indicated by the black dashed line. The control run shows a slightly lower error for the period of ∼ 15 June–31 July, while the CS2 forecasts show a more significant reduction in error from 1 August–30 September. Minor improvement is shown for the Beaufort/ Chukchi/Bering Seas (Fig. 6b), but the CS2 and persistence values are very similar. Little difference is shown for the Barents/Kara Seas (Fig. 6c) and Canadian Archipelago (Fig. 6d) where both forecasts are clear improvements over persistence. Table 2 presents the mean IIEE for September for Arctic regions shown in Figure 5. An 18% improvement over the control run is exhibited by the CS2 forecasts for the entire Arctic domain, with the Bering/ Chukchi/Beaufort region exhibiting the best regional improvement over the control run at 30%, while the Greenland/ Norwegian Sea showed the only decline at 6%.

We evaluate the ice thickness at the three mooring locations shown in Figure 7. ULS ice draft from the Woods Hole Oceanographic Institution BGEP (Krishfield and Proshutinsky, 2006) are used in this study. Raw ice draft data sampled at 2 s intervals is converted to ice thickness by dividing the draft by 0.89 (Rothrock and others, 2003). A 5-day moving average of the ULS was calculated to filter noisy data. Figure 8a, c and e show the comparison of observed ice thickness at moorings A, B and D (respectively) vs the control and CS2 5-month predictions. While the ULS indicates open water beginning in August at mooring A and little or no ice in September at mooring D, the forecasts indicate ice throughout the period at all three mooring locations. The CS2 results at mooring A highlight the best agreement with observation with a minimum ice thickness of ∼0.25–0.30 m in September. Table 3 presents summary statistics for ice thickness showing the CS2 forecasts performed best with the lower bias and RMSE at all three moorings. An examination of the mean ice extent shown in Figure 7 illustrates that mooring B is poleward of the mean sea ice extent, but mooring A is closer to the mean extent for the CS2 set of experiments. Figure 8b, d and f present ice concentration from both SSMIS and AMSR2
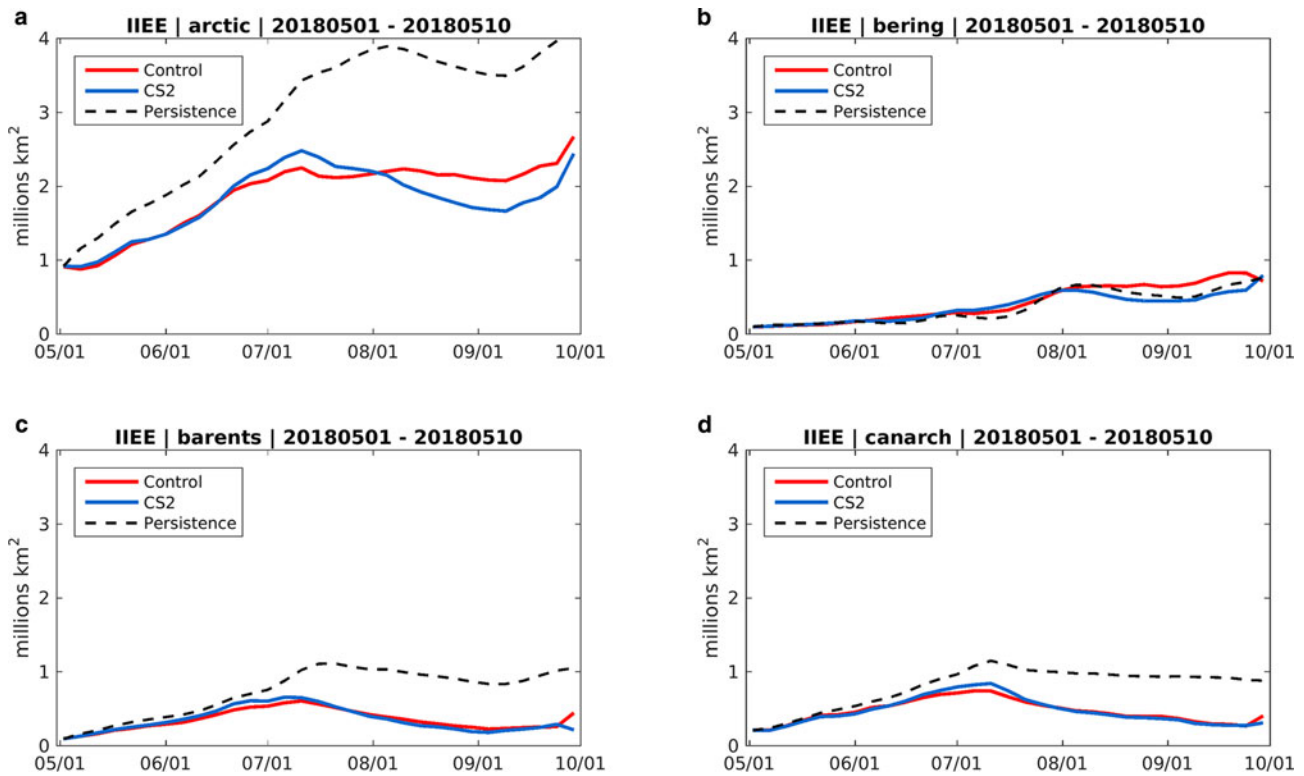
Fig. 6. Integrated IIEE (million km²) for control (red) and CS2 (blue) forecasts for (a) Pan Arctic, (b) Bering/Beaufort/Chukchi Sea, (c) Barents/Kara Seas and (d) Canadian Archipelago. Anomaly persistence is indicated by the black dashed line.
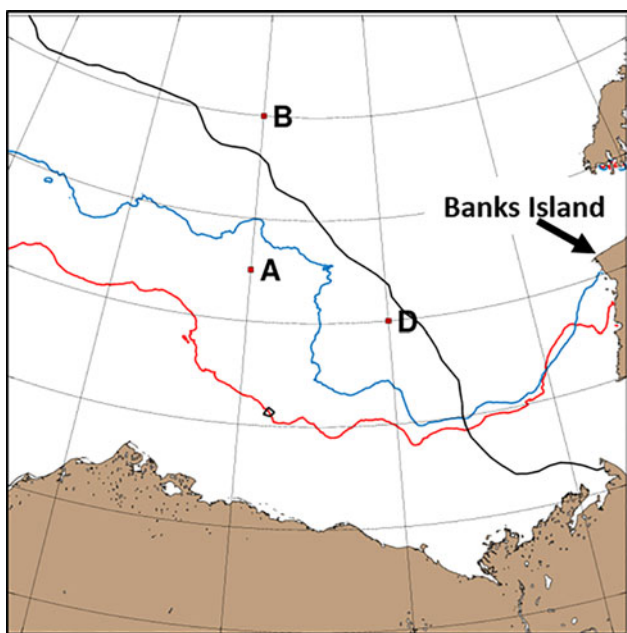

Fig. 7. Location of ULS moorings in the Beaufort Sea. Lines represent September mean ice extent for control (red), CS2 (blue) and observations (black) minimum ice extent as shown in Figure 4c.

Table 2. September 2018 mean IIEE for the control and CS2 experiments

Region	Control	CS2	% Improvement	Persistence	% Improvement
Arctic	2.18	**1.79**	18	3.68	**51**
Greenland/ Norwegian Seas	0.12	0.13	−6	0.28	**54**
Barents/Kara Seas	0.24	**0.23**	5	0.90	**74**
Laptev Sea/E. Siberian Sea	**0.75**	0.61	19	0.62	**2**
Bering/Chukchi/ Beaufort Seas	0.77	**0.52**	30	0.59	**12**
Canadian Archipelago	0.31	**0.30**	4	0.92	**67**

Grey boxes indicate best performance.

interpolated to the ULS mooring locations compared against the control and CS2 forecasts. CS2 results at mooring A show the best qualitative performance with reduced ice concentration vs the control run. However, since the models did not predict open water, concentration values are higher than observed.

To better understand the impact of CS2 initialization on the seasonal prediction of the September Arctic mean and minimum sea ice extent we analyzed the September mean 2 m air temperature. Figure 9a depicts the September mean 2 m air temperature

difference from NAVGEM between the CS2 and control forecasts. Yellow/red shading indicates that the CS2 run is warmer by 2–4°C in the Chukchi Sea as shown by the red rectangle. Figure 9b shows the air temperature averaged for this rectangle between 72 and 78° N and 180 to 140°E. The ensemble spread is indicated by the red (control) and blue (CS2) shading. The most significant difference appears in September when the air temperature difference is on the order of 2–3°C as shown in Figure 9c. The CS2 initialization led to thinner/no ice in this region as shown in Figure 1 resulting in open water and a generally warmer air temperature in May and again in late August through September.

Lastly, we examine the impact of CS2 initialization on the predicted sea surface temperature. Figure 10 depicts the mean September 2018 HYCOM sea surface temperature difference between the control and CS2 ensemble forecasts. The red colors indicate where the CS2 forecasts are warmer due to a lack of ice cover. Qualitatively, there is a strong correlation between the warmer SSTs off the coast of Alaska, Bering Sea and East Siberian Sea and the poleward position of the CS2 minimum
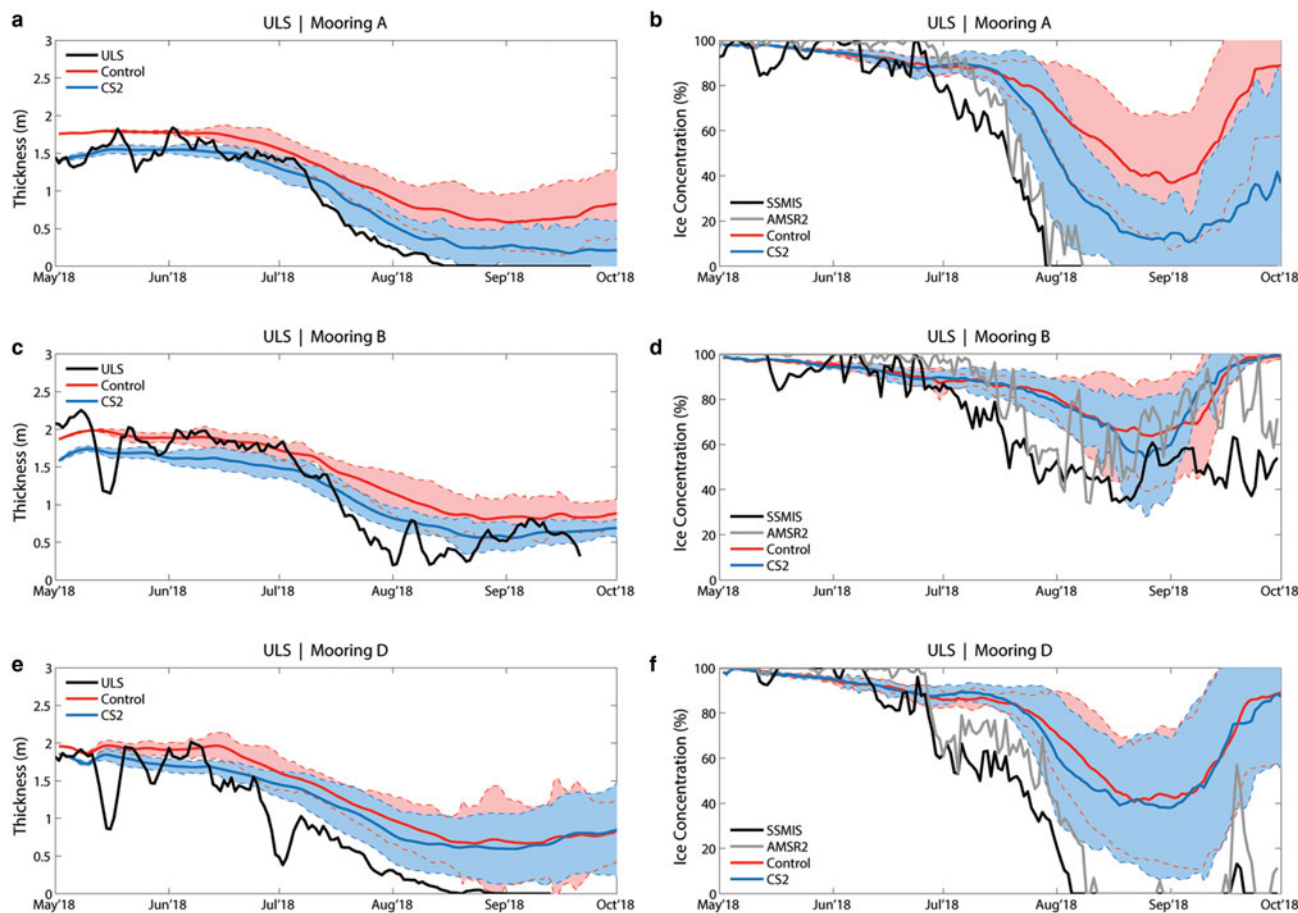
Fig. 8. (Left) 5-month ice thickness forecast at the ULS Moorings A, B and D shown in Figure 7 for the period of 1 May–30 September, 2018. Black lines on left denote 5-day moving average of observed ice thickness vs control (red) and CS2 (blue) ensemble forecasts. (Right) Predicted ice concentration at the same locations vs AMSR2 (grey) and SSMIS (black). Shading and dashed lines indicate ensemble spread.

Table 3. Ice thickness statistics at 3 ULS locations for control and CS2 5-month predictions

1 May–30 Sept 2018	ULS A	Control	CS2
Mean (m)	0.81	1.24	**0.93**
Bias		0.43	**0.12**
RMSE		0.49	**0.21**
R^2		**0.99**	0.99
1 May–30 Sept 2018	ULS B	Control	CS2
Mean (m)	1.19	1.44	**1.18**
Bias		0.24	**−0.01**
RMSE		0.38	**0.28**
R^2		0.92	**0.93**
1 May–30 Sept 2018	ULS D	Control	CS2
Mean (m)	0.87	1.39	**1.24**
Bias		0.52	**0.42**
RMSE		0.59	**0.49**
R^2		**0.94**	**0.94**

Bold text indicate best performance

extent shown in Figure 5c. We note colder SSTs in the Fram Strait, an area where we observed a southern retreat of the sea ice extent minimum also shown in Figure 4c.

4. Summary and Discussion

The reinitialization of the Navy ESPC with CS2-derived ice thickness data shows a noticeable difference at the start of the 5-month forecasts (Fig. 1a). There is generally less ice in the Beaufort and Chukchi Seas (compared to the GOFS 3.1 ice thickness) but significantly more ice between the northern Greenland and

Ellesmere Island coasts to the North Pole. That pattern is still evident on 15 September (Fig. 1b) with open water seen in the marginal seas, but with a somewhat reduced magnitude.

We find a 43% reduction in the September minimum sea ice extent error when initializing with CS2. Examining tabular data associated with Figure 2, the predictive skill of the Pan-Arctic sea ice extent is extended by 33 days (31 August vs 29 July) using CS2 data. The ice volume predictions compared to PIOMAS show a positive bias vs the control run. We attribute this to thicker ice in portions of the central Arctic from CS2 that does not melt as much or advect out of the region compared to control run. An examination of NASA IceBridge ice thickness data for April 2018 reveals that the CS2 ice thickness appears to be anomalously high in this region. The September mean sea ice extent showed the most significant improvement in the Beaufort/Chukchi Sea, with a September distance error reduction of 35 km, a 20% improvement. Marginal improvements are also seen in the East Siberian, Laptev, Kara, Barents Seas and the Canadian Archipelago. We observe a slight degradation in performance in the Greenland Sea. The IIEE for the Pan Arctic showed a reduction in error of ~0.33 m km^2 during the period of 1 August–15 September. Some improvement is also seen for the Bering/Beaufort/Chukchi and Laptev Seas. The mean September IIEE error is reduced by 18% in the Arctic using CS2 data, with a 30% improvement shown for the Bering/Chukchi/Beaufort Sea region.

The ESPC system demonstrates a positive ice thickness bias as shown in Figure S1, which covers the period of 2017–2019 for the 150-day SIPN2 ensemble forecasts initialized 1 May but without
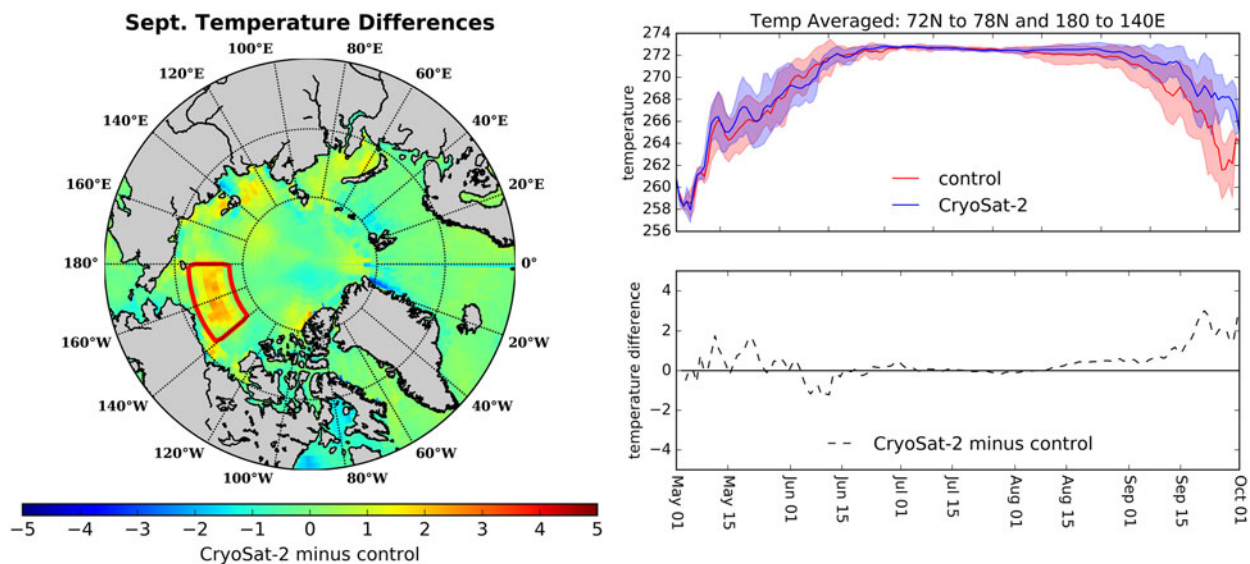
Fig. 9. (a) Mean September 2 m air temperature difference from NAVGEM between control and CS2 forecasts. Red rectangle denotes region of significantly warmer temperatures. (b) 2 m air temperatures for the red rectangular box averaged from 72–78°N, 180–140°E. Spread between both sets of ensemble simulations is shown. (c) CS2 minus control run 2 m air temperature for 5-month forecast period. Note significantly warmer temperatures in late September where open water occurs.
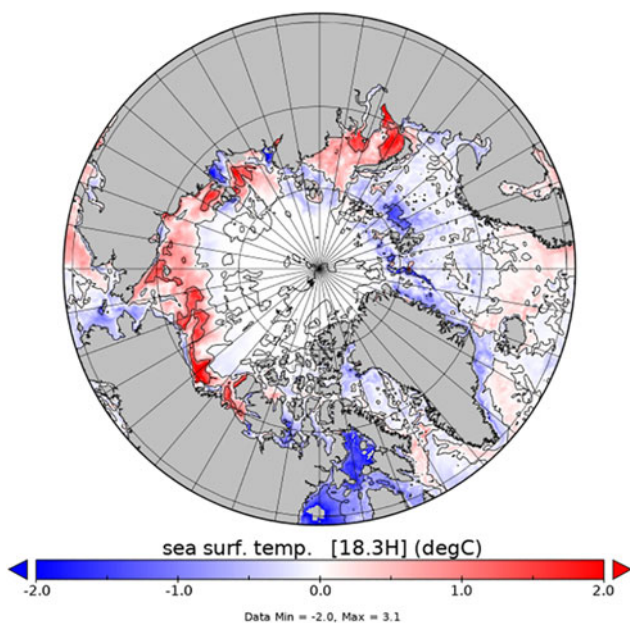

Fig. 10. HYCOM ensemble mean September 2018 sea surface temperature difference (°C) between CS2 and control forecasts. Red colors indicate the CS2 runs resulted in warmer SST.

CS2 data. Studies underway are investigating the use of CS2 2-day along-track data in our real-time prediction systems. We anticipate the use of this data will reduce this thickness bias.

The ensemble mean predictions for the September sea ice extent minimum show an improvement using CS2 data, however the predicted sea extent minimum occurs 18 days earlier than observed for either of the forecast experiments. The sea ice volume vs PIOMAS (Fig. 3) shows that the initial volume using CS2 is consistent, however the volume difference actually increases during the forecast period compared to the control run. We attribute the role of the thicker ice mentioned in the previous paragraph, which does not melt enough as a contributing cause of the higher volumes. We hypothesize that augmenting the CS2 dataset with SMOS ice thickness (e.g., Yang and others, 2019) could lead to thinner ice in some of the marginal seas

possibly resulting in a more realistic Pan Arctic sea ice volume when considering the impacts of ice advection.

The clear impact of CS2 initialization on the September sea ice extent is shown in Figure 4c. The most noticeable improvement compared to the NSIDC mean September ice extent is in the Beaufort, Chukchi and portions of the East Siberian Seas. While a marginal improvement is shown for the Laptev, Kara and Barents Seas, there is still a significant difference compared to the observed mean extent. The IIEE (Fig. 6a) for the entire Pan Arctic domain shows a slightly higher error with the CS2 experiments from days 50–90, but a much improved (lower error) score is evident during days 90–150. The CS2 experiments show a decrease in error from days 70–130, while the control experiments show an increase for most of the 150-day forecast.

The CS2 set of ensembles showed a consistent improvement in skill in forecasting ice thickness at all 3 BGEP ULS mooring site sites (Fig. 8a, c and e) in the Beaufort Seas for the 150-day period with a reduction in bias and RMSE at all three locations. These finding are consistent with those of Yang and others (2019) where they performed similar experiments for a 4-month period. The coupled modeling system was unable to predict open water at all mooring locations for either sets of experiments; however, the ice concentration time series (Fig. 8b) shows a significant improvement vs the control run. Again, we hypothesize that the use of a blended CS2/SMOS ice thickness could lead to improved forecasts.

We examined the impact of CS2 initialization on the 2 m air temperature, especially during September and showed a significant warming (Fig. 9a) in the Chukchi and Laptev Seas vs the control run, an area where we observed a much improved sea ice extent. This warming is also correlated with the warmer SST modeled for this region (Fig. 10). Future work may investigate how warmer predicted September SST's may delay the onset of freezing in the Autumn.

References

Allard RA and 10 others (2018*b*) Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice–ocean modeling system. *Advances in Space Research* **62**, 1265–1280. doi: 10.1016/j.asr.2017.12.030.

Allard RA, Barton N, Metzger EJ and Phelps MW (2018*a*) CICE outputs from Navy Earth System Model (NESM) and Global Ocean Forecast System (GOFS) 3.1, 2018. Naval Research Laboratory, Stennis Space Center, PANGAEA. doi: 10.1594/PANGAEA.890343.

Baker NL and 10 others (2007) An overview of the NRL Atmospheric Variational Data Assimilation (NAVDAS) and NAVDAS-AR (Accelerated Representer) Systems. In: *18th AMS Conference on Numerical Weather Prediction*, 25–29 June, Park City, UT, Paper 2B.1, 6pp. Available at http://ams.confex.com/ams/pdfpapers/124031.pdf.

Barton N and 11 others (2019) *Earth System Prediction Capability (ESPC) Initial Operational Capability (IOC) ensemble system*. Validation Test Report, Naval Research Laboratory, NRL Memorandum Report 7532-19-9928, 66 pp.

Blanchard-Wrigglesworth E and 13 others (2016) Multi-model seasonal forecast of Arctic sea–ice: forecast uncertainty at pan-Arctic and regional scales. *Climate Dynamics* **49**(4), 1399–1410. doi: 10.1007/s00382-016-3388-9.

Blockley EW and Peterson KA (2018) Improving Met Office seasonal predictions of Arctic sea ice using assimilation of CryoSat-2 thickness. *The Cryosphere* **12**, 3419–3438. doi: 10.5194/tc-12-3419-2018.

Collow TW, Wang W, Kumar A and Zhang J (2015) Improving Arctic sea ice prediction using PIOMAS initial sea ice thickness in a coupled ocean–atmosphere model. *Monthly Weather Review* **143**(11), 4618–4630.

Comiso JC, Cavalieri DJ and Markus T (2003) Sea ice concentration, ice temperature, and snow depth using AMSR-E data. *IEEE Transactions on Geoscience and Remote Sensing* **41**(2), 243–252. doi: 10.1109/TGRS.2002.808317.

Comiso JC and Nishio F (2008) Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data. *Journal of Geophysical Research* **113**, C02S07. doi: 10.1029/2007JC0043257.

Cummings JA and Smedstad OM (2014) Ocean data impacts in global HYCOM. *Journal of Atmospheric and Oceanic Technology* **31**, 1771–1791. doi: 10.1175/JTECH-D-14-00011.1.

Day JJ, Hawkins E and Tietsche S (2014) Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophysical Research Letters* **41**, 7566–7575. doi: 10.1002/2014GL061694.

Dirkson A, Merryfield WJ and Monahan A (2017) Impacts of sea ice thickness initialization on seasonal Arctic sea ice predictions. *Journal of Climate* **30**(3), 1001–1017. doi: 10.1175/JCLI-D-16-0437.1.

Fetterer F, Knowles K, Meier WN, Savoie M and Windnagel AK (2017, updated daily) *Sea Ice Index*, Version 3. [NH-Daily-Extent]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi: 10.7265/N5K072F8. (Accessed 29 August 2019).

Goessling HF, Tietsche S, Day JJ, Hawkins E and Jung T (2016) Predictability of the Arctic sea ice edge. *Geophysical Research Letters* **43**, 1642–1650. doi: 10.1002/2015GL067232.

Hogan TF and 16 others (2014) The navy global environmental model. *Oceanography* **27**(3), 116–125. doi: 10.5670/oceanog.2014.73.

Hunke E and Lipscomb W (2008) CICE: the los alamos sea ice model documentation and software user's manual Version 4.0. Tech. Rep. LA-CC-06-012. Los Alamos Natl. Lab., Los Alamos, NM.

Krishfield R and Proshutinsky A (2006) BGOS ULS data processing procedure. Woods Hole Oceanographic Institution, March 2006, 14 pp.

Kurtz N and Harbeck J (2017) CryoSat-2 Level 4 sea ice Elevation, freeboard and thickness. Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center.

Maslanik J and Stroeve J (1999) Near-real-time DMSP SSMIS daily polar gridded sea ice concentrations, Version 1. [May–September 2018]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: 10.5067/U8C09DWVX9LM. [4–5 October 2018].

Merryfield WJ and 7 others (2013) The Canadian seasonal to interannual prediction system: part 1: models and initialization. *Monthly Weather Review* **141**, 2910–2945. doi:10.1175/MWR-D-12-00216.1.

Metzger EJ and 12 others (2014) US Navy operational global ocean and Arctic ice prediction systems. *Oceanography* **27**(3), 32–43. doi: 10.5670/oceanog.2014.66.

Pegion K and 28 others (2019) The Subseasonal Experiment (SubX): A multi-model subseasonal prediction experiment. *Bulletin of the American Meteorology Society* **100**, 2043–2060. doi: 10.1175/BAMS-D-18-0270.1.

Rothrock DA, Zhang J and Yu Y (2003) The Arctic ice thickness anomaly of the 1990s: a consistent view from observations and models. *Journal of Geophysical Research* **108**, 3083. doi: 10.1029/2001JC001208.

Schröder D, Feltham DL, Flocco D and Tsamados M (2014) September Arctic sea-ice minimum predicted by spring melt-pond fraction. *Nature Climate Change* **4**, 353–357. doi: 10.1038/nclimate2203.

Schweiger A, Lindsay R, Zhang J, Steele M and Stern H (2011) Uncertainty in modeled Arctic sea ice volume. *Journal of Geophysical Research* **116**, 1–21. doi: 10.1029/2011JC007084.

Stroeve J, Hamilton LC, Bitz CM and Blanchard-Wrigglesworth E (2014) Predicting September sea ice: ensemble skill of the SEARCH Sea Ice Outlook 2008–2013. *Geophysical Research Letters* **41**, 2411–2418. doi: 10.1002/2014GL059388.

Tilling R, Ridout A, Shepherd A and Wingham D (2015) Increased Arctic sea ice volume after anomalously low melting in 2013. *Nature Geoscience* **8**, 643–648. doi: 10.1038/NGEO2489.

Williams J, Tremblay B, Newton R and Allard R (2016) Dynamic preconditioning of the minimum September sea-ice extent. *Journal of Climate* **29**, 5879–5891. doi: 10.1175/JCLI-D-15-0515.1.

Yang Q and 6 others (2019) Improving Arctic sea ice seasonal outlook by ensemble prediction using an ice-ocean model. *Atmospheric Research* **227**, 14–23. doi: 10.1016/j.atmosres.2019.04.021.