


Multi-model seasonal forecast of Arctic sea-ice: forecast uncertainty at pan-Arctic and regional scales

E. Blanchard-Wrigglesworth¹  · A. Barthélemy² · M. Chevallier³ · R. Cullather⁴ · N. Fučkar⁵ · F. Massonnet^{2,5} · P. Posey⁶ · W. Wang⁷ · J. Zhang⁸ · C. Ardilouze³ · C. M. Bitz¹ · G. Vernieres^{4,9} · A. Wallcraft⁶ · M. Wang^{10,11}

Received: 1 July 2016 / Accepted: 6 October 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Dynamical model forecasts in the Sea Ice Outlook (SIO) of September Arctic sea-ice extent over the last decade have shown lower skill than that found in both idealized model experiments and hindcasts of previous decades. Additionally, it is unclear how different model physics, initial conditions or forecast post-processing (bias correction) techniques contribute to SIO forecast uncertainty. In this work, we have produced a seasonal forecast of 2015 Arctic summer sea ice using SIO dynamical models initialized with identical sea-ice thickness in the central Arctic. Our goals are to calculate the relative contribution of model uncertainty and irreducible error growth to forecast uncertainty and assess the importance of post-processing, and to contrast pan-Arctic forecast uncertainty with regional forecast uncertainty. We find that prior to forecast post-processing, model uncertainty is the main contributor to forecast uncertainty, whereas after forecast post-processing forecast uncertainty is reduced

overall, model uncertainty is reduced by an order of magnitude, and irreducible error growth becomes the main contributor to forecast uncertainty. While all models generally agree in their post-processed forecasts of September sea-ice volume and extent, this is not the case for sea-ice concentration. Additionally, forecast uncertainty of sea-ice thickness grows at a much higher rate along Arctic coastlines relative to the central Arctic ocean. Potential ways of offering spatial forecast information based on the timescale over which the forecast signal beats the noise are also explored.

Keywords Sea ice · Seasonal forecast · Arctic · Forecast uncertainty

1 Introduction

Interest in seasonal forecasts of Arctic sea ice has grown in recent years as a result of growing socio-economic activity in the region and as a science-motivated effort to test our understanding and ability to predict seasonal changes

Electronic supplementary material The online version of this article (doi:[10.1007/s00382-016-3388-9](https://doi.org/10.1007/s00382-016-3388-9)) contains supplementary material, which is available to authorized users.

✉ E. Blanchard-Wrigglesworth
ed@atmos.washington.edu

¹ Department of Atmospheric Sciences, University of Washington, Seattle, WA 98195-1640, USA

² Georges Lemaître Centre for Earth and Climate Research, Université catholique de Louvain, Louvain-La-Neuve, Belgium

³ Centre National de Recherches Météorologiques, UMR 3589, Météo France, Toulouse, France

⁴ Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA

⁵ Barcelona Supercomputing Center, Barcelona, Spain

⁶ Naval Research Laboratory, Stennis Space Center, Hancock County, MS, USA

⁷ NOAA/NWS/NCEP/Climate Prediction Center, College Park, MD, USA

⁸ Applied Physics Laboratory, Polar Science Center, University of Washington, Seattle, WA, USA

⁹ Science Systems and Applications, Inc., Greenbelt, MD, USA

¹⁰ Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA 98195-1640, USA

¹¹ Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration, Seattle, WA, USA

in the Arctic. Since 2008, seasonal forecasts of September sea-ice extent—collectively called the September Sea Ice Outlook (SIO)—have been annually produced, collected and discussed by the Arctic research community under the auspices of the Arctic Research Consortium of the United States (ARCUS) and since 2013 by the Sea Ice Prediction Network (SIPN, see <http://www.arcus.org/sipn/sea-ice-outlook>).

An initial assessment of forecast skill in SIO over 2008–2013 by Stroeve et al. (2014) found that overall skill was only marginally better than a linear trend forecast, while dynamical model predictions slightly underperformed relative to statistical predictions. Unexpectedly, forecast skill did not significantly improve as the forecast lead time decreased from 4 to 2 months. Blanchard-Wrigglesworth et al. (2015) compared forecast skill in SIO dynamical models to the skill in perfect models and hindcasts (retrospective forecasts) and found that SIO skill was markedly low, not even beating a simple damped anomaly persistence forecast. In addition, it was found that SIO models were equally poor in predicting each other, an indication that differences in the initial conditions and/or the model physics across the models must play a role in the large forecast spread (and thus large forecast uncertainty). Blanchard-Wrigglesworth et al. (2015) also found that when initialized with identical perturbation anomalies, the responses of four SIO models' forecasts were significantly different, suggesting that different model physics are a key contributor to forecast uncertainty.

Another factor that may lead to large forecast spread in SIO is the application of different post-processing adjustments or bias correction methods on raw forecasts. Given the large trend in September sea-ice extent during recent decades, removing climatological means or trends from different periods for bias correction may result in significantly different forecasts, all other factors being equal. It is noteworthy that for the 2015 SIO, just 11 out of 37 submissions by dynamical models documented a bias correction procedure in their forecasts, and among these forecasts, bias correction techniques were inconsistent, both in the hindcast time-period used for bias correction, and which quantities were corrected (mean, trend, or both). Indeed, the skill of seasonal and annual forecasts of sea ice extent has been found to be particularly sensitive to the choice of bias correction method on the raw forecast (Fučkar et al. 2014).

Studies using both idealized 'perfect-model' setups (e.g., Day et al. 2014a) and statistical models (e.g., Lindsay et al. 2008) have highlighted the importance of sea-ice thickness as a predictor of summer sea-ice area. Unfortunately, sea-ice thickness is a complex, difficult-to-observe state variable (e.g., Haas 2003). Large scale, temporally infrequent satellite-based observations are only recently becoming more readily available (e.g., Laxon et al. 2013;

Tilling et al. 2016), while small scale but high resolution airborne observations such as those from NASA Operation IceBridge (Kurtz et al. 2013) may provide accurate sea-ice thickness estimates for a certain region at a certain time. Given the lack of spatially-and-temporally rich observations of sea-ice thickness, modeling centers have either not assimilated sea-ice thickness into seasonal forecasts' initial conditions (e.g., Sigmond et al. 2013) or used reconstructions of sea-ice thickness in order to initialize seasonal forecasts (e.g., Guemas et al. 2014; Collow et al. 2015). The latter study showed that the skill of seasonal forecasts of sea ice is dependent on the quality of the sea-ice thickness initial conditions while, importantly, it has also been shown that the different ocean–sea ice reanalyses that are used to initialize forecasts with different dynamical models can be vastly different (Chevallier et al. 2016). Ultimately, it is not known to what degree different initial conditions across models affect overall SIO dynamical model forecast skill and forecast uncertainty.

To date, no multi-model SIO sea ice forecast initialized with identical sea ice has been produced. In this work, we show results from a multi-model seasonal forecast of Arctic sea ice in which all models are initialized with identical sea-ice thickness within an Arctic basin domain. We use two sets of forecast ensembles that are initialized with a recent 1 May climatology and a 2015 1 May sea-ice thickness, which allows us to investigate both the growth of different sources of forecast uncertainty and the impact that forecast post-processing has on forecast uncertainty. Additionally, we are able to investigate differences between pan-Arctic and regional forecasts.

The paper is organized as follows: we first describe the forecast design in Sect. 2 and show results in Sect. 3. We end with a discussion of results and conclusions in Sect. 4.

2 Model forecast design

In this study, we use the Pan-Arctic Ice Ocean Model and Assimilation System (PIOMAS, Zhang and Rothrock 2003) sea-ice thickness fields to initialize seasonal forecasts in eight different dynamical models. PIOMAS has been shown to simulate observed pan-Arctic sea-ice volume with high fidelity (Schweiger et al. 2011), while in a recent study a prediction system initialized with PIOMAS was found to show higher skill than when initialized with other sea-ice thickness datasets (Collow et al. 2015). Participating models in our study are listed in Table 1 and range from regional ice-ocean models to global fully-coupled atmosphere-ice-ocean models. All models have submitted forecasts to the SIO in past years. We run two sets of simulations that are initialized on 1 May and are 5 months in length. The first set of simulations, the climatology

Table 1 List of participating models

Model	Type	Ensemble size	Resolution (sea ice model)
CNRM	Fully coupled	15	1 lon × 1 lat
EC-EARTH2.3	Fully coupled	20	1 lon × 1 lat
NASA GEOS-5	Fully coupled	10	1 lon × 1 lat
NCAR CESM1	Fully coupled	9	0.9 lon × 1.25 lat
NOAA CFSv2	Fully coupled	16	0.5 lon × 0.5 lat
NRL GOFS3.1	Global ice-ocean	10	0.1 lon × 0.1 lat
PIOMAS	Regional ice-ocean	7	0.8 lon × 0.8 lat
UCL NEMO-LIM3	Global ice-ocean	10	1 lon × 1 lat

See table S1 for further details

forecast, is initialized with mean 2007–2014 1 May sea-ice thickness that represents a recent climatology of 1 May sea ice thickness. The second set of simulations, the 2015 forecast, is initialized with the 2015 1 May sea-ice thickness. Initial conditions for all other variables are taken from 1 May 2015 in both sets (table S1). Given that most of the spring predictability of summer sea ice extent in dynamical models is borne from sea ice thickness (Day et al. 2014b), the climatology forecast will evolve toward the equivalent of a 2007–2014 climatology of summer sea ice conditions, enabling us to use this simulation as an approximation of a recent climatology against which we can post-process the experiment simulation. Given the strong biases that GCMs tend to exhibit in their simulations of Arctic sea ice (e.g., Massonnet et al. 2012), we anticipate that both sets of simulations will exhibit model drift, and forecast post-processing allows us to remove this drift. In the results, we show two forecasts: the absolute 2015 forecasts, and ‘anomaly’ 2015 forecasts, which are computed as a difference between the 2015 and climatology simulations. All models produce ensembles for both sets of simulations, with the number of ensemble members ranging from 7 to 20 for each individual model (see Table 1). Details of each model’s components and initialization techniques are given in the supplementary material (table S1) with further information listed in the individual reports associated with the SIO summer 2015 reports (see the SIO url listed above).

When initializing the sea-ice thickness in each set of simulations, we only modify sea-ice thickness in the Arctic basin. We apply a geographic weighting mask as follows:

$$hi_{IC} = hi_{nat} + w(hi_{PIO} - hi_{nat}), \quad (1)$$

where hi_{IC} is the initial condition sea-ice thickness used in the simulations, hi_{nat} is the native sea-ice thickness of each model on 1 May, hi_{PIO} is the PIOMAS sea-ice thickness on 1 May, and w is a weighting mask that varies between 0 and 1 (see Fig. 1 and the supplementary material). The reasons for the design and application of the weighting mask are: one, September sea-ice extent is influenced by spring sea-ice thickness in the central Arctic but not by spring sea-ice

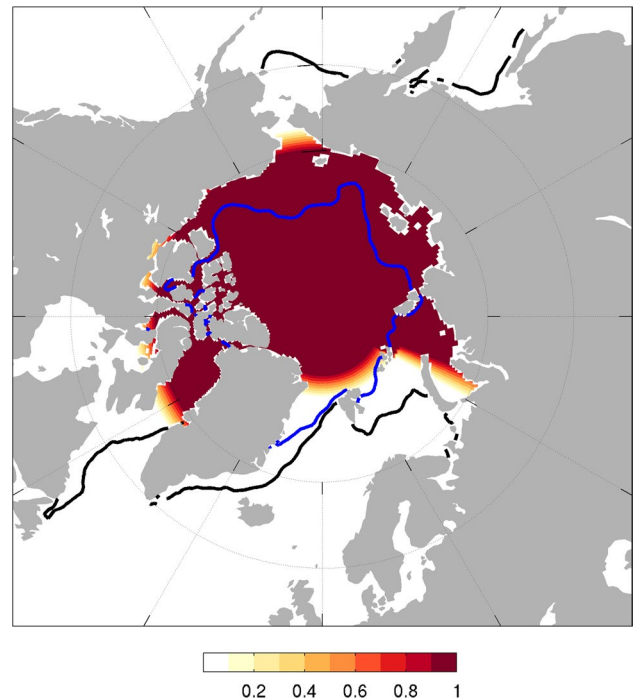
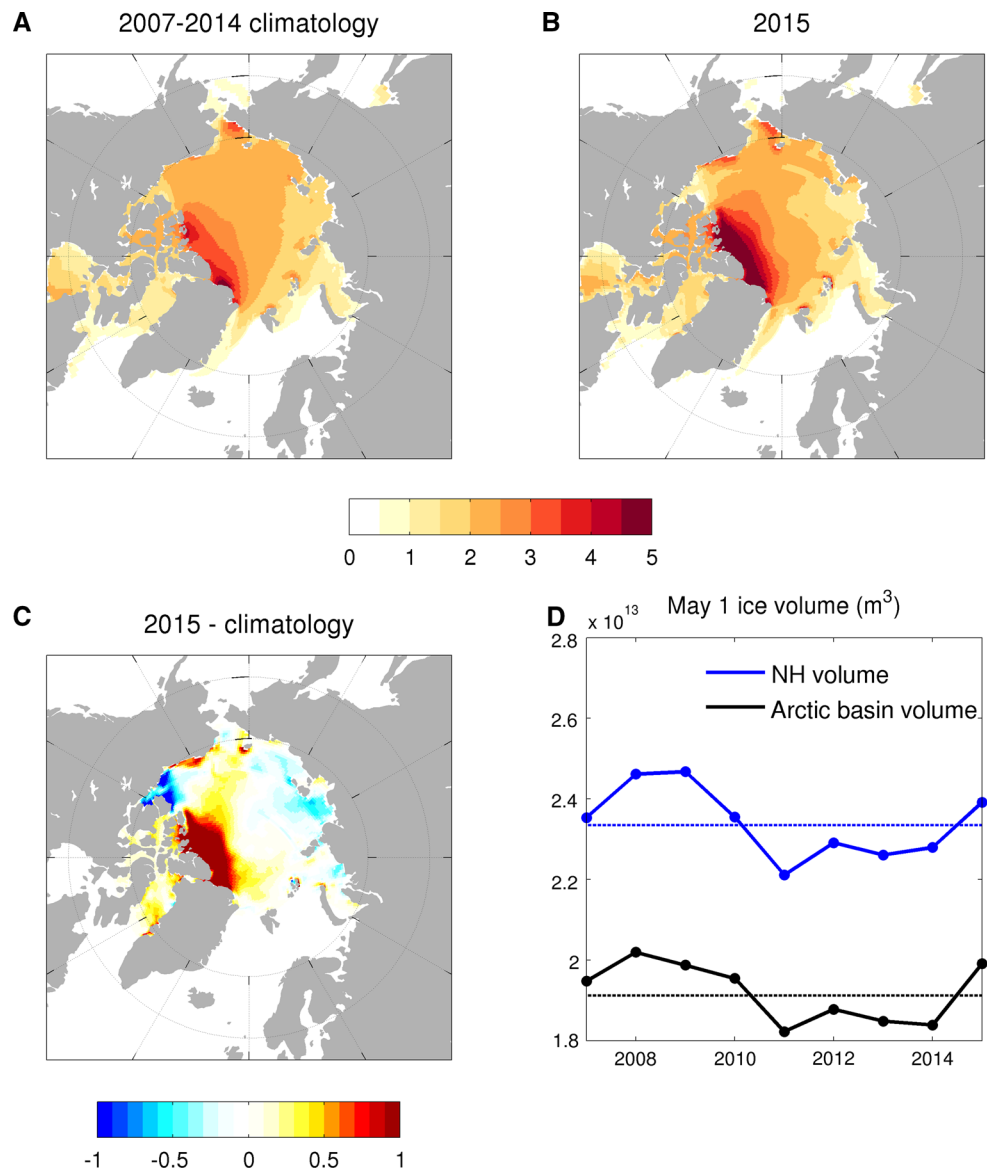


Fig. 1 Weighting mask w (see Eq. 1 in main text and supplementary material) used for application of initial condition sea-ice thickness. The *black contour* shows the mean 2007–2014 sea ice edge in NSIDC observations on 1 May and the *blue contour* shows the mean September sea ice edge for 2007–2014 in NSIDC observations

thickness in sub-Arctic seas (see e.g., Figure 4 in Day et al. 2014b)—in other words, the seasonal sea-ice zone in sub-Arctic seas melts back due to the seasonal cycle by late summer irrespective of spring sea-ice thickness in sub-Arctic seas—and two, in dynamical models it is much more complex to remove or add sea ice completely near the ice edge than to simply edit sea-ice thickness in grid cells that are already ice covered, as initial shock can lead to numerical instabilities and strong, unrealistic adjustment of upper ocean properties where the sea ice has been removed/added. Importantly, in PIOMAS on 1 May, the Arctic basin (as defined by our weighting mask) contains 80 % of total northern

Fig. 2 **a** Mean PIOMAS 1 May sea-ice thickness in meters over the 2007–2014 period (referred to as climatology), **b** PIOMAS 2015 1 May sea-ice thickness, **c** 2015—climatology difference within the masked region shown in Fig. 1, and **d** time series of 1 May pan-Arctic (blue) and Arctic basin (as defined by w in Fig. 1, in black) sea-ice volume over 2007–2015. The dashed lines are the 2007–2014 mean for both domains



hemisphere sea-ice volume and in addition both volume time series are strongly correlated ($r = 0.95$, see Fig. 2d). Over most of the Arctic basin, sea-ice thickness is modified north of the mean 1 May sea-ice edge, but south of the mean September sea-ice edge (Fig. 1). The only exception occurs in Svalbard's vicinity, where the mean 1 May and September sea-ice edges are almost collocated. This is the only region where the initial sea-ice thickness is modified north of the September sea ice edge. However, given the climatological sea-ice drift from the north, which advects spring thickness anomalies into the region, and the small impact that September sea-ice anomalies in this region have on total September sea-ice area variability (which is dictated mostly by variability in the Siberian and Alaskan sectors, see e.g., Figure 14 in Deser et al. 2000), we do not expect this design feature to have a significant impact on our results.

PIOMAS uses an ice-thickness distribution (ITD) scheme whereby sea ice is distributed into different ice-thickness categories at each grid cell. For those models that do not have an ITD scheme, the mean PIOMAS sea-ice thickness is simply specified at each grid cell. For models that do have an ITD scheme, the 12-category PIOMAS ITD is linearly interpolated into the model's n -category ITD (for example, the CICE sea-ice model has a 5-category ITD) in order to both conserve mean sea-ice thickness and the distribution of sea-ice thickness. Snow and ice enthalpy per unit volume are conserved, and all other sea ice variables (sea-ice concentration, snow depth, etc.) and other component variables, including ocean component variables, are left unchanged from the different models' native initial conditions (table S1).

The 1 May sea-ice thickness initial conditions used in both sets of simulations are shown in Fig. 2. Compared to

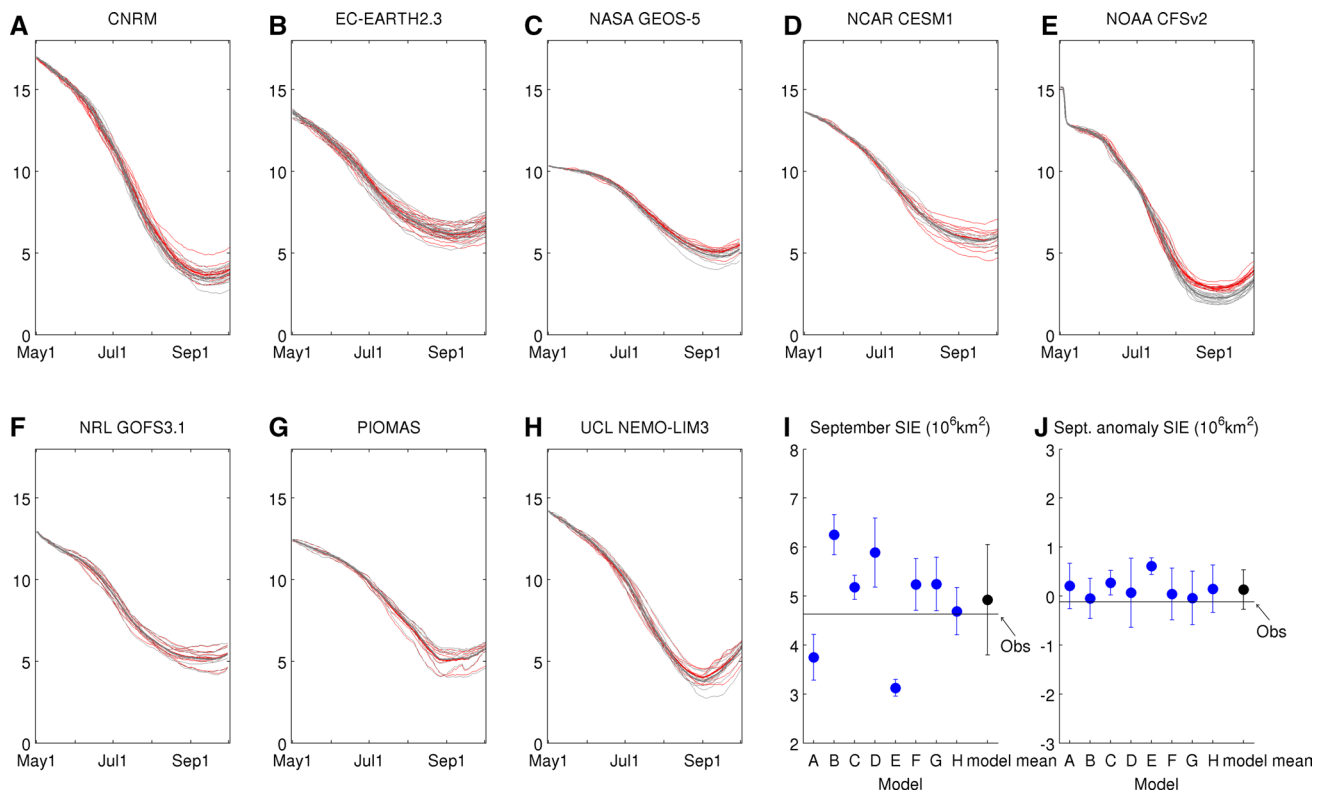


Fig. 3 Daily 1 May to October 1 total northern hemisphere sea-ice extent (SIE, in 10^6 km^2) for all 8 models in the 2015 forecast (red) and climatology forecast (black) simulations (panels a–h). Thin lines represent individual model runs, bold lines represent model means. Panel i (j) shows mean September SIE forecasts (SIE anomaly forecasts) for all 8 models in blue, the multi-model mean in black and

the observed 2015 September sea-ice extent (anomaly relative to 2007–2014) as a black line. The model mean September SIE anomaly is also shown in black. The error bars indicate one standard deviation for each model, and the model mean error bar the standard deviation of the all-model ensemble

the 2007–2014 climatology, 2015 had thicker ice north of the Canadian Arctic Archipelago and Greenland, north of Alaska and in the Baffin Bay, and thinner ice over the Laptev and East Siberian seas and the McKenzie River Delta. Overall, total sea-ice volume in the northern hemisphere and within the Arctic basin was greater in 2015 relative to the 2007–2014 climatology. For observations of sea ice extent and concentration, we use the National Snow and Ice Data Center (NSIDC) Sea Ice Index data (Fetterer et al. 2002). We have found that our results are not significantly sensitive to the choice of observational product by also using the NSIDC bootstrap algorithm data (Comiso et al. 2000) (not shown).

3 Results

3.1 Sea-ice extent, concentration and volume

We begin by showing values of daily sea-ice extent and mean September sea-ice extent in Fig. 3. Initial sea-ice extent on 1 May differs markedly across all models. This is not unexpected: since the forecast design does not modify

the sea ice edge of each model's original initial conditions (as the May 1 sea-ice edge lies to the south of the weighting mask area in Fig. 1), there is an initial difference in sea-ice extent across models. These inter-model differences in sea-ice extent increase throughout the forecast period, resulting in a significant model spread of the ensemble-mean September sea-ice extent, which ranges from $3 \times 10^6 \text{ km}^2$ to over $6 \times 10^6 \text{ km}^2$. The model-mean September sea-ice extent of $4.92 \times 10^6 \text{ km}^2$ compares well with the observed 2015 value of $4.63 \times 10^6 \text{ km}^2$ (Fig. 3i), but has a large uncertainty associated with it as reflected by the standard deviation of the all-model ensemble.

We now consider the extent anomaly forecasts (i.e., the 2015 minus climatology forecasts). Compared to the forecasts of absolute extent, the spread in the forecast of the September extent anomaly is much smaller across models, with anomaly forecasts ranging from ~ 0 to $0.6 \times 10^6 \text{ km}^2$. The model-mean anomaly forecast is $0.2 \times 10^6 \text{ km}^2$. The observed 2015 anomaly sea-ice extent value was $-0.1 \times 10^6 \text{ km}^2$ (with respect to the observed 2007–2014 climatology), well within one standard deviation of the models' forecasts (Fig. 3j).

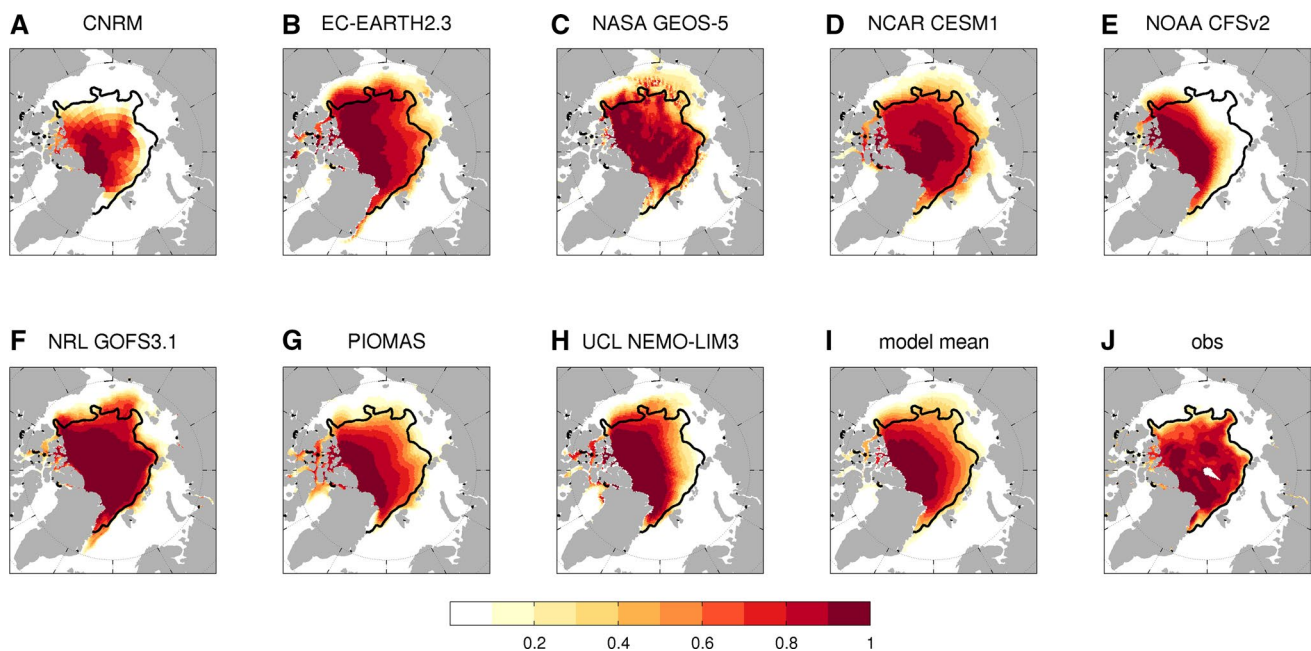


Fig. 4 Ensemble-mean 2015 September sea-ice concentration forecasts for all 8 models. The *black line* represents the 15 % concentration contour in the NSIDC observations, which is usually taken to define the sea ice edge

Forecasts of 2015 September sea-ice concentration (Fig. 4) have a large spread across the models, mirroring the large spread in extent in Fig. 3i. This model spread is also seen in the climatology simulations (not shown). As with pan-Arctic extent, the model-mean 2015 forecast shows a reasonable agreement with observations (Fig. 4i). Forecasts of September sea-ice concentration anomaly (Fig. 5) show large differences across models, unlike the forecast of September sea-ice extent anomaly (Fig. 3i). As a result, the model-mean concentration anomaly forecast shows no consistent patterns of positive or negative anomalies, unlike the observed 2015 anomaly with respect to the observed 2007–2014 climatology. In Fig. 5, individual model ensemble-mean forecasts also tend to forecast weaker anomalies than observed, but this partly results from the ensemble-averaging which acts to reduce anomaly amplitudes by filtering out chaotic components.

Forecasts of daily sea-ice volume (Fig. 6) show a large divergence in absolute forecasts across models. As is the case for sea-ice area, the spread in forecasted September sea-ice volume anomalies (Fig. 6j) is much smaller than the total volume forecast spread (Fig. 6i). All models simulate a positive anomaly in September sea-ice volume, the positive anomaly signal in the initial conditions (see Fig. 2 above) persisting throughout the summer. Unfortunately, there are no observational data of sea-ice volume against which the forecasts can be validated, but we use the PIOMAS reconstructed sea-ice volume values as a proxy for observations (black lines in Fig. 6j, i). Results from PIOMAS suggest

that September 2015 sea-ice volume was close to the 2007–2014 volume, hinting that forecasts slightly overestimate sea-ice volume and the persistence of the positive initial condition volume anomaly.

3.2 Forecast uncertainty

We now quantify forecast uncertainty and its partition into model uncertainty (also known as structural uncertainty) and irreducible error growth. When assessing or making use of a multi-model forecast, uncertainty in the forecast originates from model uncertainty and the irreducible forecast uncertainty that arises from the chaoticity of the climate system present within each model's forecasts (for ice-ocean models that do not have an atmospheric model component, irreducible forecast uncertainty will also arise from the uncertainty of forecast forcing). Additional forecast uncertainty will originate from uncertainty in the initial conditions (e.g., the difference between the PIOMAS sea-ice thickness fields and the unknown, real-world sea-ice thickness fields)—our forecast design however precludes us from estimating this source of forecast uncertainty.

Variables of interest to address forecast uncertainty are the sea-ice volume and area within the masked region of Fig. 1. To avoid differences among models due merely to their differing land masks (especially in the Canadian Arctic Archipelago region), each model's output is regridded to the PIOMAS grid and subject to the PIOMAS land mask. With this processing, the models have nearly the same

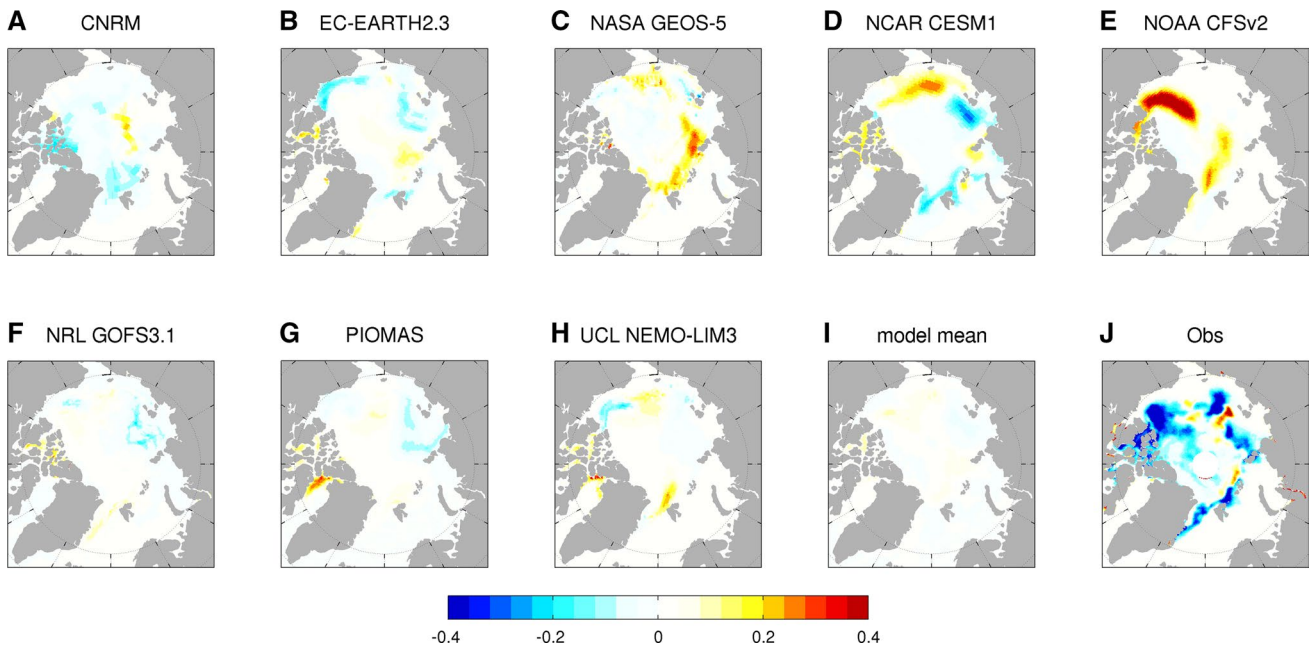


Fig. 5 Mean September sea-ice concentration anomaly (2015—climatology) forecasts for all 8 models, model mean, and the observed September concentration anomaly (2015—climatology)

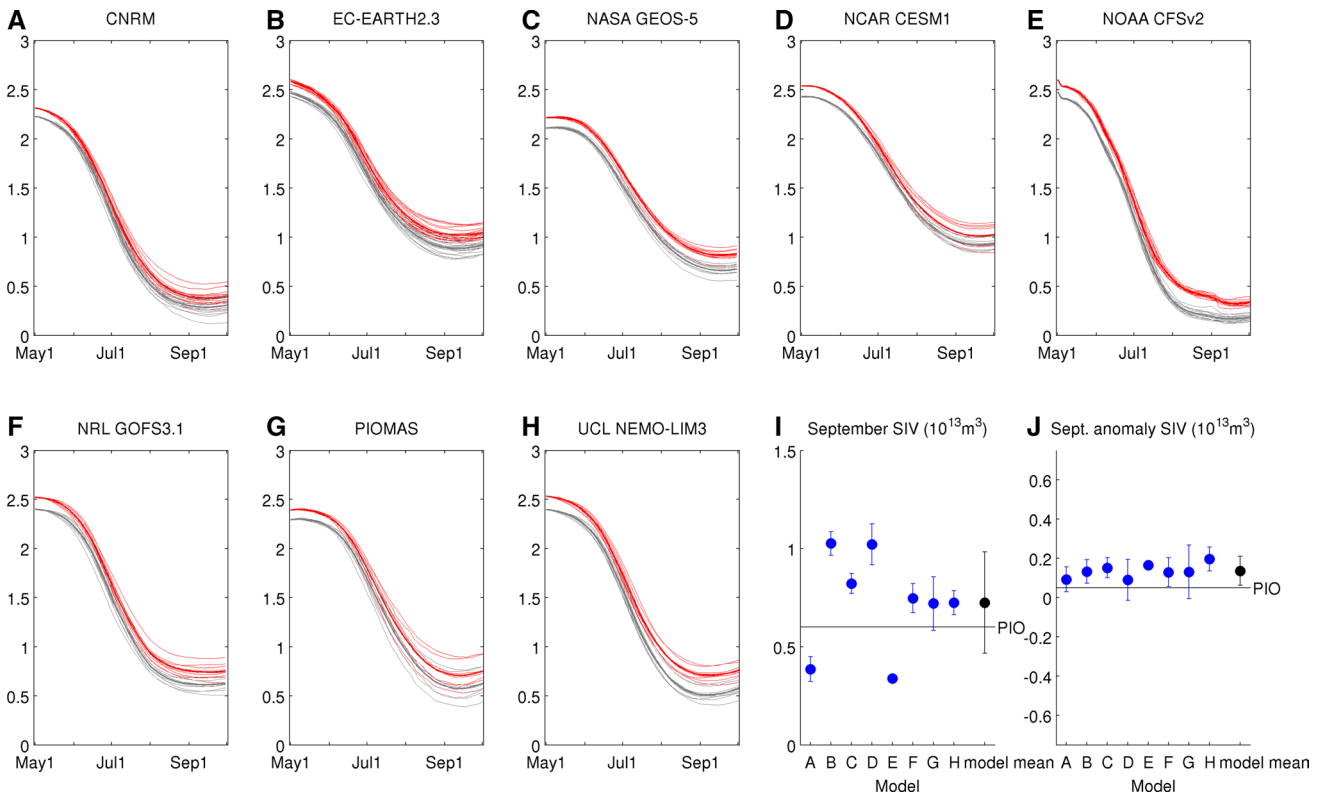
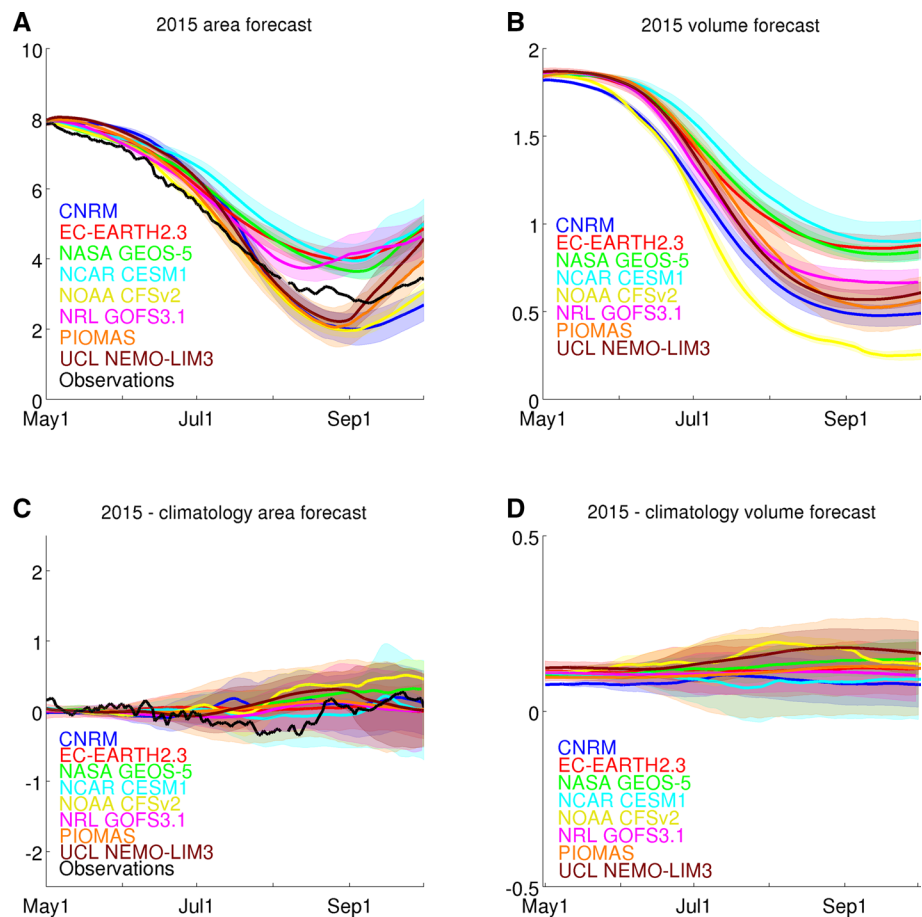


Fig. 6 As Fig. 3, but for sea-ice volume (SIV, in 10^{13} m^3). Panel i (j) shows mean September SIV forecasts (SIV anomaly forecasts) for all 8 models in blue, the multi-model mean in black and the PIOMAS (in assimilation mode) 2015 September SIV (anomaly relative to 2007–

2014), labeled 'PIO'. The model mean September SIV anomaly is also shown in black. The error bars indicate one standard deviation for each model, and the model mean error bar the standard deviation of the all-model ensemble

Fig. 7 Daily 1 May to October 1 Arctic basin sea-ice area (in 10^6 km^2) and volume (in 10^{13} m^3) 2015 forecasts and anomaly (2015—climatology) forecasts for sea-ice area (c) and volume (d) for all 8 models. **Bold lines** indicate the model mean value for each model, the **shaded area** around the bold line indicate one standard deviation for each model. The **black lines** in a and c show daily observed Arctic basin sea-ice area



sea ice conditions in the masked region on 1 May, and the growth in their forecast spread henceforth is easily seen (Fig. 7).

We quantify forecast uncertainty as follows: model uncertainty (MU hereafter) is defined as the across-model standard deviation of each models' ensemble-mean forecasts, or $\sigma(\bar{f})$, where \bar{f} is a model's ensemble-mean forecast; the uncertainty associated with the irreducible error growth (IE hereafter) is defined as the mean of the standard deviation of each model's forecast ensemble, or $\overline{\sigma(f)}$. We also compute the model uncertainty of the anomaly forecast (aMU hereafter), or $\sigma(\bar{f}')$, where \bar{f}' is the anomaly of the model's ensemble-mean forecast (bold lines in Fig. 7c, d). Figure 8 shows the values for all 3 uncertainty metrics for sea-ice volume and sea-ice area. For both quantities, MU dominates forecast uncertainty, and is about 2–3 times greater than IE. Interestingly, error growth is not linear with time, and IE shows a rapid period of error growth from early June to mid July and may be related to the previously reported 'predictability barrier' (Blanchard-Wrigglesworth et al. 2011; Day et al. 2014b), a period in early summer when positive albedo-ice melt feedbacks initiated mainly by unpredictably atmospheric conditions result in a rapid loss of predictive skill. In contrast, aMU is about

an order of magnitude lower than MU, and only about a third or half of IE. We note however that a fraction of MU and aMU will derive from the sampling uncertainty of the ensemble means given the finite ensemble sizes, rather than model uncertainty *per se*. This fraction can be estimated by the standard error of the mean, which given the ensemble sizes in the experiment will be $\sim 0.28 \times \text{IE}$. The implication of anomaly forecasting (i.e., *post-processing* the forecast) for the skill and uncertainty of seasonal sea-ice forecasts is discussed further below.

Our forecast design allows us to investigate how forecast uncertainty varies regionally. We have applied the same forecast uncertainty analysis as above but using sea-ice thickness at each grid-cell rather than pan-Arctic metrics. Figure 9 shows spatial patterns of MU, IE and aMU for four different forecast lead times (forecast days 30, 60, 90 and 120). During the first 2 months of the forecast, MU and IE grow rapidly along the Arctic coastlines, indicating higher forecast uncertainty in those regions relative to the central Arctic. At longer forecast times, MU also grows significantly in the central Arctic, unlike IE which tends to maintain the high coastal / low central Arctic pattern seen in earlier months. As is the case for pan-Arctic volume, aMU at each grid cell is much reduced relative to

Fig. 8 Daily 1 May to October 1 model uncertainty (MU), irreducible error growth (IE), and anomaly forecast model uncertainty (aMU)—see main text for definitions—for **a** sea-ice area in 10^{11} m^2 and **b** sea-ice volume in 10^{12} m^3

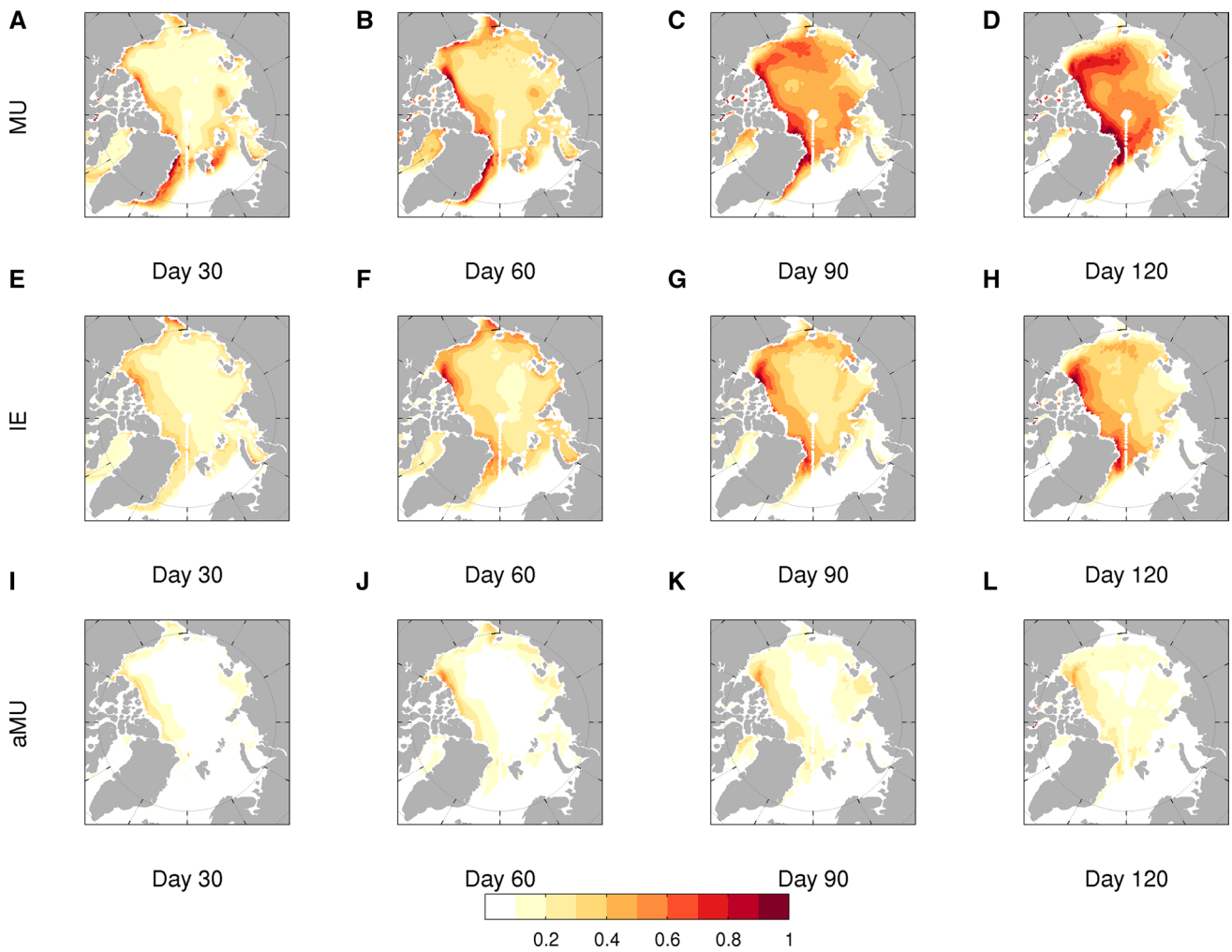
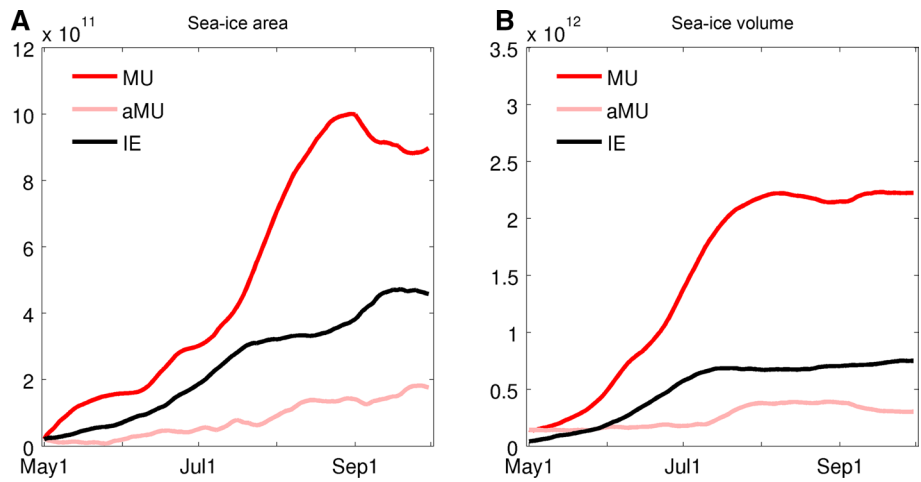


Fig. 9 Spatial model uncertainty (MU, *top row*), irreducible error (IE, *middle row*) and anomaly model uncertainty (aMU, *bottom row*) for sea-ice thickness at forecast lead times of 30 days (**a, e, i**), 60 days (**b, f, j**), 90 days (**c, g, k**) and 120 days (**d, h, l**). All units are meters

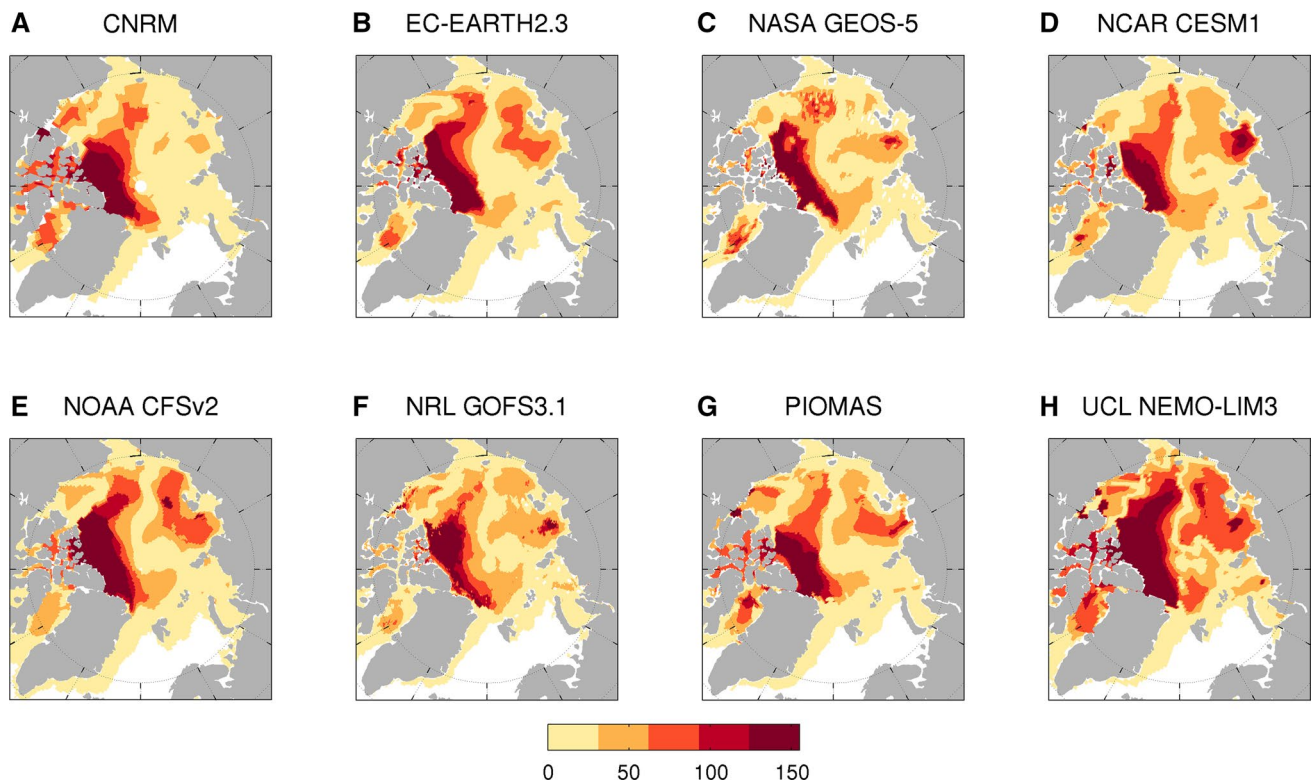


Fig. 10 Forecast period ('signal time' in the text) in days from initialization over which the 2015 ensemble mean ice thickness forecast anomaly (the signal) is greater than an estimate of ensemble spread

(the noise). The *color scale* represents month with May, June, July, August and September in increasingly *darker red coloring*

MU. Interestingly, an area of high initial growth of aMU in the first 2 months north of the CAA and Greenland tends to advect with the climatological ice drift toward the Beaufort Gyre by the end of the forecast period.

3.3 Skillful forecast lead time

The potential value of a forecast to stakeholders greatly depends on the geographic characteristics of the forecast. Integrated metrics such as pan-Arctic sea-ice extent or volume, while being of interest from an academic point of view, are not of great practical use to stakeholders, who are more interested in regional metrics. Unfortunately, regional metrics tend to have lower predictive skill than integrated metrics (e.g., Day et al. 2014b; Goessling et al. 2016), while the physical mechanisms that lead to loss of predictability with forecast lead time are different across pan-Arctic and regional scales (Tietsche et al. 2016). These differences between pan-Arctic and regional predictability can also be seen in our results when considering the multi-model forecast: models uniformly forecast 2015 September pan-Arctic extent and volume anomalies (Figs. 3j, 6j), yet there is no agreement across models on the spatial patterns of September sea-ice concentration anomalies (Fig. 5).

An alternative way to communicate useful forecast information is to show the length of time over which the forecast has skill relative to using climatology as a forecast. To achieve this we compute the daily signal strength of the sea ice thickness forecast by calculating the signal-to-noise ratio as follows:

$$ss(t) = |(\overline{hi_e(t)} - \overline{hi_c(t)})/\sigma(hi_e)|, \quad (2)$$

where ss is the signal strength, hi_e and hi_c are the mean ice thickness at each grid cell in the 2015 and climatology forecasts respectively, and σ is the standard deviation operator. When ss is greater than 1, the signal of the forecast is greater than the noise. We calculate ss for each day, grid cell, and model. Finally, we compute the last day on which ss is greater than one, and call this the 'signal time'. Figure 10 shows the spatial patterns of signal time for all eight models' forecasts. Overall, the pattern bears resemblance to the pattern of initial condition anomaly in Fig. 2c, which is unsurprising: the larger the initial condition anomaly (whether it be positive or negative), or signal (i.e., the numerator in Eq. 2 at time = 0), the longer the forecast signal tends to last. However, there are notable exceptions. In most models, a tongue of long signal time extends toward the Chukchi Sea from the north of the CAA, despite the

initial condition anomaly being weak north of the Chukchi Sea. By contrast, in many models, the positive initial condition anomaly along the north coast of Alaska does not translate into a long forecast signal. This likely results from the large growth in IE along coastlines shown in Fig. 9: the fast growth in $\sigma(hi_e)$ means that ss drops below one at an earlier date. Importantly, only a relatively small area of the Arctic shows signal times that last until September.

4 Discussion and conclusions

In this work, we have produced a multi-model seasonal sea-ice forecast using identical sea-ice thickness initial conditions in the Arctic basin to explore the relative contribution of model uncertainty and irreducible error growth to forecast uncertainty. Furthermore, we have explored differences in skill between more standard pan-Arctic metric forecasts and regional forecasts that may be of more use to stakeholders.

Forecasts of September 2015 using 1 May 2015 sea-ice thickness initial conditions show that prior to forecast post-processing, model uncertainty dominates forecast uncertainty. It is indeed remarkable how fast different models diverge in their forecasts, and by September the range in sea-ice extent is several million km², amply above the observed variability in September sea-ice extent. The fast divergence documented here is impressive, especially in light of the long, annual persistence timescales that sea-ice thickness anomalies tend to exhibit in dynamical coupled models (Blanchard-Wrigglesworth and Bitz 2014).

In contrast, when post processing the forecast by subtracting each model's climatology forecasts, models agree much better in their forecasts of sea-ice extent and volume anomalies. Model uncertainty decreases by an order of magnitude, and the irreducible error growth becomes the dominant source of forecast uncertainty. This result suggests that effective post-processing adjustment (bias correction) plays an important role in seasonal forecasting of September sea ice, a problem that is exacerbated not only by the large model biases with respect to climatology but also by the large trends in the observations and the sensitivity of the trend to the time period chosen for its calculation. It is likely that improvements in the application of bias correction methods across SIO models would result in reduced forecast uncertainty and higher forecast skill. Unlike sea-ice extent and volume, forecasts of sea-ice concentration still show large inter-model differences after post-processing, indicating that regional forecasts of sea ice are more uncertain, and thus likely to show less skill, than pan-Arctic sea-ice extent forecasts.

Although we find good agreement between the observed sea-ice area anomaly and the forecast anomaly (Fig. 7c),

we are unable to make any assessment of the reliability of the forecast system, since such an assessment requires a large sample size of forecasts and observations (e.g., Hamill 2001). However, we note that the magnitude of IE in Fig. 8a is close to the observed standard deviation of (detrended) September sea-ice area (~ 0.5 million km²), an indication that the dispersion component of forecast skill (the predictability of the ensemble spread) is close to saturation. Given that additional uncertainty from the aMU and unknown uncertainty associated with initial conditions will add to total forecast uncertainty, this means that, if the 2015 forecast is representative of the models' mean state predictability (i.e., if 2015 is not an anomalously unpredictable year that led to faster-than-usual IE growth), any overall forecast skill of September extent in the model system discussed in this work could only originate from the signal component of predictability (the predictability of the ensemble mean, see e.g., Branstator and Teng 2010; Blanchard-Wrigglesworth et al. 2011 for a more detailed discussion of both sources of initial-value predictability). Unfortunately, a long dataset of annual forecasts (or hindcasts) is needed to estimate this component of predictability, though we note that previous idealized experiments have shown that the signal component of sea-ice predictability can be higher than the dispersion component over the first two seasons of forecast lead time (Blanchard-Wrigglesworth et al. 2011).

We have also explored regional aspects of forecast uncertainty and a regional forecast metric based on the forecast lead time over which the signal beats the noise that may be of use to stakeholders. In terms of regional forecast uncertainty, there is a stark contrast between the Arctic coastlines and the central Arctic basin, particularly during the initial 2–3 months of the forecast period. Forecast uncertainty grows rapidly along the Arctic coastlines, and in some regions contributes to shorten the timescale over which the forecast signal beats the noise. A similar geographic pattern of error growth has been shown in idealized 'perfect-model' experiments (Tietsche et al. 2014), which hinted that sea ice dynamics play a key role in this degradation of forecast skill. Unfortunately not all of the participating models output thermodynamic and dynamic ice volume tendencies, precluding us from investigating the partition of error growth further. Considering the regional needs of stakeholders in the Arctic, our results have significant consequences for the future design and plausibility of stakeholder-focused seasonal sea-ice forecasts.

There are some limitations imposed by the design of our forecast simulations. It is important to note that forecast uncertainty associated with inter-model differences in initial conditions in other (non sea-ice) components of the system that affect the forecast evolution of Arctic sea ice will manifest itself in our results as part of the model

uncertainty component. While ultimately it would be desirable to diagnose model uncertainty in a forecast set-up in which all initial conditions across all components are identical, the challenges involved in designing such a forecast are considerable. Our effort represents a step in that direction.

Acknowledgments We wish to thank the Sea Ice Prediction Network, at whose annual workshop the model experiment was originally discussed. EBW, CMB and MW were supported by ONR Grant N0014-13-1-0793. JZ was supported by ONR Grant N00014-12-1-0112. This publication is partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement No. NA10OAR4320148, Contribution No. 2016-01-17, and Pacific Marine Environmental Laboratory Contribution Number 4515.

References

- Blanchard-Wrigglesworth E, Bitz CM (2014) Characteristics of Arctic sea-ice thickness variability in GCMs. *J Clim* 27(21):8244–8258
- Blanchard-Wrigglesworth E, Bitz CM, Holland MH (2011) Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophys Res Lett*. doi:10.1029/2011GL048807
- Blanchard-Wrigglesworth E, Cullather R, Wang W, Zhang J, Bitz C (2015) Model forecast skill and sensitivity to initial conditions in the seasonal Sea Ice Outlook. *Geophys Res Lett* 42(19):8042–8048
- Branstator G, Teng H (2010) Two limits of initial-value decadal predictability in a CGCM. *J Clim*. doi:10.1175/2010JCLI3678.1
- Chevallier M, Smith GC, Dupont F, Lemieux JF, Forget G, Fujii Y, Hernandez F, Msadek R, Peterson KA, Storto A, Toyoda T et al (2016) Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project. *Clim Dyn* 1–30. doi:10.1007/s00382-016-2985-y
- Collow TW, Wang W, Kumar A, Zhang J (2015) Improving Arctic sea ice prediction using PIOMAS initial sea ice thickness in a coupled ocean–atmosphere model. *Mon Weather Rev* 143(11):4618–4630
- Comiso J (2000, updated 2015) Bootstrap sea ice concentrations for nimbus-7 smmr and dmsp ssm/i, version 2. Digital Media, National Snow and Ice Data Center
- Day J, Hawkins E, Tietsche S (2014a) Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophys Res Lett* 41(21):7566–7575
- Day J, Tietsche S, Hawkins E (2014b) Pan-Arctic and regional sea ice predictability: initialisation month dependence. *J Clim* 27(12):4371–4390
- Deser C, Walsh JE, Timlin MS (2000) Arctic sea ice variability in the context of recent atmospheric circulation trends. *J Clim* 13(3):617–633
- Fetterer F, Knowles K, Meier W, Savoie M (2002, updated 2015) Sea ice index. <http://nsidc.org/data/go2135.html>
- Fučkar NS, Volpi D, Guemas V, Doblas-Reyes FJ (2014) A posteriori adjustment of near-term climate predictions: accounting for the drift dependence on the initial conditions. *Geophys Res Lett* 41(14):5200–5207
- Goessling HF, Tietsche S, Day JJ, Hawkins E, Jung T (2016) Predictability of the Arctic sea ice edge. *Geophys Res Lett* 43:1642–1650. doi:10.1002/2015GL067232
- Guemas V, Doblas-Reyes FJ, Mogensen K, Keeley S, Tang Y (2014) Ensemble of sea ice initial conditions for interannual climate predictions. *Clim Dyn* 43(9–10):2813–2829
- Haas C (2003) Dynamics versus thermodynamics: the sea ice thickness distribution. In: *Sea Ice: an Introduction to its physics, chemistry, biology, and geology*. Wiley-Blackwell, Oxford, p 402
- Hamill TM (2001) Interpretation of rank histograms for verifying ensemble forecasts. *Mon Weather Rev* 129(3):550–560
- Kurtz N, Farrell S, Studinger M, Galin N, Harbeck J, Lindsay R, Onana V, Panzer B, Sonntag J (2013) Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data. doi:10.5194/tc-7-1035-2013
- Laxon SW, Giles KA, Ridout AL, Wingham DJ, Willatt R, Cullen R, Kwok R, Schweiger A, Zhang J, Haas C et al (2013) Cryosat-2 estimates of Arctic sea ice thickness and volume. *Geophys Res Lett* 40(4):732–737
- Lindsay RW, Zhang J, Schweiger AJ, Steele MA (2008) Seasonal predictions of ice extent in the Arctic Ocean. *J Geophys Res Oceans*. doi:10.1029/2007jc004259
- Massonnet F, Fichetef T, Goosse H, Bitz CM, Philippon-Berthier G, Holland MM, Barriat PY (2012) Constraining projections of summer arctic sea ice. *The Cryosphere* 6(6):1383–1394
- Schweiger A, Lindsay R, Zhang J, Steele M, Stern H, Kwok R (2011) Uncertainty in modeled Arctic sea ice volume. *J Geophys Res* 116:C00D06. doi:10.1029/2011JC007084
- Sigmond M, Fyfe J, Flato G, Kharin V, Merryfield W (2013) Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system. *Geophys Res Lett* 40(3):529–534
- Stroeve JC, Hamilton L, Bitz CM, Blanchard-Wrigglesworth E (2014) Predicting September sea ice: ensemble skill of the SEARCH sea ice outlook 2008–2013. *Geophys Res Lett* 41(7):2411–2418
- Tietsche S, Day JJ, Guemas V, Hurlin WJ, Keeley SPE, Matei D, Msadek R, Collins M, Hawkins E (2014) Seasonal to interannual Arctic sea ice predictability in current global climate models. *Geophys Res Lett* 41(3):1035–1043
- Tietsche S, Hawkins E, Day JJ (2016) Atmospheric and oceanic contributions to irreducible forecast uncertainty of arctic surface climate. *J Clim* 29(1):331–346
- Tilling RL, Ridout A, Shepherd A (2016) Near real time Arctic sea ice thickness and volume from CryoSat-2. *The Cryosphere* (in review). doi:10.5194/tc-2016-21
- Zhang J, Rothrock D (2003) Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Mon Weather Rev* 131(5):845–861