Pan-Arctic and regional sea ice predictability: initialization month dependence

Article

Published Version

Open Access

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Pan-Arctic and Regional Sea Ice Predictability: Initialization Month Dependence

J. J. DAY, S. TIETSCHE, AND E. HAWKINS

NCAS-Climate, Department of Meteorology, University of Reading, Reading, United Kingdom

ABSTRACT

Seasonal-to-interannual predictions of Arctic sea ice may be important for Arctic communities and industries alike. Previous studies have suggested that Arctic sea ice is potentially predictable but that the skill of predictions of the September extent minimum, initialized in early summer, may be low. The authors demonstrate that a melt season ''predictability barrier'' and two predictability reemergence mechanisms, suggested by a previous study, are robust features of five global climate models. Analysis of idealized predictions with one of these models [Hadley Centre Global Environment Model, version 1.2 (HadGEM1.2)], initialized in January, May and July, demonstrates that this predictability barrier exists in initialized forecasts as well. As a result, the skill of sea ice extent and volume forecasts are strongly start date dependent and those that are initialized in May lose skill much faster than those initialized in January or July. Thus, in an operational setting, initializing predictions of extent and volume in July has strong advantages for the prediction of the September minimum when compared to predictions initialized in May.

Furthermore, a regional analysis of sea ice predictability indicates that extent is predictable for longer in the seasonal ice zones of the North Atlantic and North Pacific than in the regions dominated by perennial ice in the central Arctic and marginal seas. In a number of the Eurasian shelf seas, which are important for Arctic shipping, only the forecasts initialized in July have continuous skill during the first summer. In contrast, predictability of ice volume persists for over 2 yr in the central Arctic but less in other regions.

1. Introduction

The rapid reduction in Arctic summer sea ice has led to a large increase in demand for forecasts of sea ice conditions at seasonal to interannual time scales (Eicken 2013). This is important information for end users, including those interested in marine accessibility for routing ships (e.g., Stephenson et al. 2011). This interest has led to the development of a number of operational seasonal sea ice prediction systems (e.g., Sigmond et al. 2013; Chevallier et al. 2013; Wang et al. 2013), which are initialized from observations.

These operational prediction systems show some skill in predicting summer sea ice conditions, but diagnosing the source of forecast errors is problematic. Such forecast errors may be due to both inadequate representation of important physical processes in the model and incomplete knowledge of the initial state of key variables such as sea ice thickness and subsurface ocean properties, which are not well observed. There is also an inherent limit to predictability in the Arctic climate system due to chaotic atmospheric variability (e.g., Holland et al. 2010; Blanchard-Wrigglesworth et al. 2011b). If a given forecast system is close to this inherent limit, then any attempt to improve sea ice predictions would be futile.

To address the key question of whether there is the potential to improve the operational prediction systems, we consider a more idealized situation. Analysis of ''perfect model'' experiments with coupled global climate models (GCMs) provides a setting where perfect knowledge of the initial model state exists and there are no model biases. Such an approach has been used to quantify and understand predictability of the ocean state, especially in the Atlantic (e.g., Collins 2002; Pohlmann et al. 2004), leading to the development of operational decadal prediction systems (e.g., Smith et al. 2007).

This technique has also previously been adopted for the Arctic region by Koenigk and Mikolajewicz (2008), finding that Arctic sea ice thickness showed signs of high predictability at interannual time scales. Using the Community Climate System Model, version 4 (CCSM4),

Blanchard-Wrigglesworth et al. (2011b) found that Arctic sea ice area and volume are continuously potentially predictable for 1–2 yr and 2–4 yr, respectively. However, the mean climate state is thought to affect the limit of sea ice extent predictability, which is expected to be lower in a future thinner sea ice environment (Holland et al. 2010).

A lagged correlation analysis performed with the CCSM3 indicates that sea ice extent predictability may be dependent on the start month (Blanchard-Wrigglesworth et al. 2011a, hereafter BW2011a). BW2011a found that memory of sea ice area, indicated by lagged correlation, is generally lost in the first 2–5 months for each start month. After the initial decline in memory, reemergence of significant correlations is found in later months. BW2011a highlight two mechanisms for this. A melt-to-freeze season reemergence occurs between pairs of months where the ice edge is in the same position (e.g., May and December) and is due to persistence of SST anomalies. A summer-to-summer reemergence mechanism is due to the persistence of sea ice thickness anomalies in the Arctic basin. Perhaps most importantly, observations of September sea ice extent are only significantly correlated with sea ice extent from the previous August and July (hence no evidence of summer-to-summer reemergence). Anomaly correlations between September extent and earlier months are not significant, indicating that forecast skill may drop off rapidly. This spring drop off in skill is thought to be caused by both the rapid motion of the ice edge (hence persistence of conditions at the ice edge in one particular month does not lead to predictability at the ice edge the next month) and because reemergence mechanisms that act to prolong skill during other months of the year are not effective in spring (BW2011a).

However, it is important to determine if this type of lagged correlation is relevant for learning about the skill of initialized predictions. The skill of the Canadian Seasonal to Interannual Prediction System (CanSIPS) in predicting sea ice extent seems to have start date dependence consistent with BW2011a, where forecasts of detrended September extent anomalies initialized before June show no significant skill (Sigmond et al. 2013), but this is unlikely to be the limit of predictability because of the lack of sea ice thickness initialization in these forecasts. Chevallier et al. (2013), who do include thickness information in their initialization, predict detrended September extent from May initialization with significant skill (anomaly correlation of 0.6). Perfect model forecasts of September extent, initialized before June, also retain significant amounts of skill in the MPI model (e.g., Fig. 10 of Tietsche et al. 2013).

Open questions in this area which will be addressed in this study include the following:

(i) Are the reemergence mechanisms discussed by BW2011a consistent across models?
(ii) Are the predictability properties implied by the lagged correlation reemergence, such as a strong start date dependence, a feature of initialized forecasts?
(iii) What month should forecasts of the September sea ice extent minimum be initialized from?

In this study, a similar analysis to BW2011a is extended to five different GCMs to assess the robustness of the persistence properties and lagged correlation reemergence mechanisms. However, it is also important to determine if these properties affect initialized predictions. To achieve this, several sets of idealized perfect model forecasts with one of these GCMs, a version of the Hadley Centre Global Environment Model, version 1.2 (HadGEM1.2), were initialized in January, May, and July, which are before, during, and after this melt season "predictability barrier," respectively. This set of ensembles includes more start months than previous potential predictability studies, allowing the start month dependence of the potential skill in pan-Arctic and regional sea ice extent and volume to be investigated.

Although measures such as pan-Arctic sea ice extent and volume are important, it is likely that predictions of regional ice properties are most important to end users. Few studies have looked at predictability of ice cover at the basin scale. One exception is Koenigk and Mikolajewicz (2008), who find that sea ice concentration predictability is significantly higher in the basins adjacent to the Atlantic (e.g., Barents and Labrador Seas) than those in the central Arctic. We also explore regional sea ice extent and volume predictability in this analysis.

This paper is structured as follows: In section 2, we describe the GCMs utilized and the experimental design. Section 3 considers diagnostic predictability of Arctic conditions in a range of GCMs, and section 4 discusses the prognostic predictability using perfect model simulations. We conclude and discuss the implications of our findings in section 5.

2. Global climate models used

a. Multimodel analysis

To examine sea ice predictability, multicentennial simulations with five fully coupled atmosphere–ice–ocean GCMs are used. The simulations with HadGEM1.2 (Shaffrey et al. 2009); Geophysical Fluid Dynamics Laboratory Climate Model, version 3 (GFDL CM3; Griffies et al. 2011); European Consortium Earth System Model version 2.3 (EC-Earth2.3; Hazeleger et al. 2012); and Max Planck Institute Earth System Model

(MPI-ESM; Jungclaus et al. 2013) were run with greenhouse gases and other radiative forcings fixed at "present day" levels (1990 for CCSM3, HadGEM1.2, and GFDL CM3; 2005 for MPI-ESM and EC-Earth2.3). Apart from CCSM3, the other models are an ensemble of opportunity, run as part of the Arctic Potential Predictability on Seasonal to Interannual Time Scales (APPOSITE) project (Tietsche et al. 2014; http://arp.arctic.ac.uk/projects/arctic-predictability-and-prediction-seasonal-inte/). The present-day control simulation of CCSM3 from phase 3 of the Coupled Model Intercomparison Project (CMIP3) multimodel dataset is also included in a number of the multimodel comparison figures to facilitate direct comparison with BW2011a. In the analysis of diagnostic predictability that follows (section 3), each of the control runs were linearly detrended to remove any model drift. For this analysis, 200 yr from each of the control runs were used, apart from HadGEM1.2 and CCSM3, from which 249 and 300 yr were used, respectively. The models included in this analysis cover a diversity of mean sea ice states (see Fig. S1 of Tietsche et al. 2014).

b. HadGEM1.2

A more in-depth description of HadGEM1.2 is provided here since the majority of this analysis, including the perfect model study, is performed using this model. HadGEM1.2 is similar to the CMIP3 version of HadGEM1, which is fully described in Johns et al. (2006). The atmosphere component has a resolution of 1.25° latitude by 1.875° longitude with 38 layers in the vertical. The ocean component has a zonal resolution of 1° and a meridional resolution of 1° between the poles and 30° latitude, increasing smoothly to $\frac{1}{3}$° at the equator with 40 unevenly spaced levels in the vertical.

A number of improvements to HadGEM1 are included in HadGEM1.2, including changes to the snow-free sea ice albedo, runoff into frozen soil, and the calculation of surface fluxes. Each of these changes improved the HadGEM1.2 mean state compared to HadGEM1 (for full details, see Shaffrey et al. 2009).

The sea ice component of HadGEM1.2 is identical to HadGEM1 and was fully described and evaluated by McLaren et al. (2006). The sea ice component shares much of its code with the CICE sea ice model (Hunke and Lipscomb 2004). Mean sea ice extent in the reference simulation used in this study is higher than mean observations during the satellite era (1979–2012) with sea ice volume also significantly higher than observed estimates (see Fig. S1 of Tietsche et al. 2014). This is in part due to a cold bias in the North Pacific (McLaren et al. 2006). However, predictability metrics indicate

that sea ice extent and volume predictability in this model is fairly typical when compared to other GCMs (see Fig. 1 of Tietsche et al. 2014). This and the strong performance of this model in reproducing many other climate indices (Johns et al. 2006) indicates that this is a useful model with which to investigate sea ice predictability.

3. Analysis of present-day control simulations

a. Lagged correlations

One simple measure of potential predictability is lagged correlation, which is a measure of the skill in a system from this particular form of persistence. To determine if the lagged correlation properties of monthly mean pan-Arctic sea ice extent discussed by BW2011a are robust, the analysis is repeated for the present-day control simulations of four other GCMs. We calculate lagged correlations of sea ice extent to examine whether the same persistence/reemergence properties exist and compare with observations. Further, we extend this analysis to monthly mean sea ice volume in both the GCMs and the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) reanalysis (Zhang and Rothrock 2003). We also perform correlations between sea ice extent, SST, and sea ice thickness in the GCMs to examine whether the reemergence mechanisms are consistent with BW2011a.

In Fig. 1, the lagged correlation of sea ice extent in observations and five climate models is shown for each start month against lead time [i.e., in each panel, from left to right, the top row shows January correlated with January (lag 0), January correlated with February (lag 1), etc.]. The sea ice extent time series of each month of the observations ([from the National Snow and Ice Data Centre (NSIDC); Meier et al. 2012] and each GCM were linearly detrended before the correlations were calculated. The exact correlation values obtained from the observations depend on the detrending method used; however, the shape of the contours in 1-month lead-time space do not (not shown).

The patterns of lagged correlation in Fig. 1 are similar to those described for the CCSM3 by BW2011a: namely, there is evidence of the melt-to-freeze season and September–September reemergence of correlation in all the models, despite large differences in the mean state of the sea ice (see Fig. 2). However, there is large intermodel spread in both the relative magnitude of the reemergence and the absolute magnitude of correlations with, for example, GFDL CM3 having much higher correlations at all lead times than MPI-ESM. Intermodel differences in sea ice extent
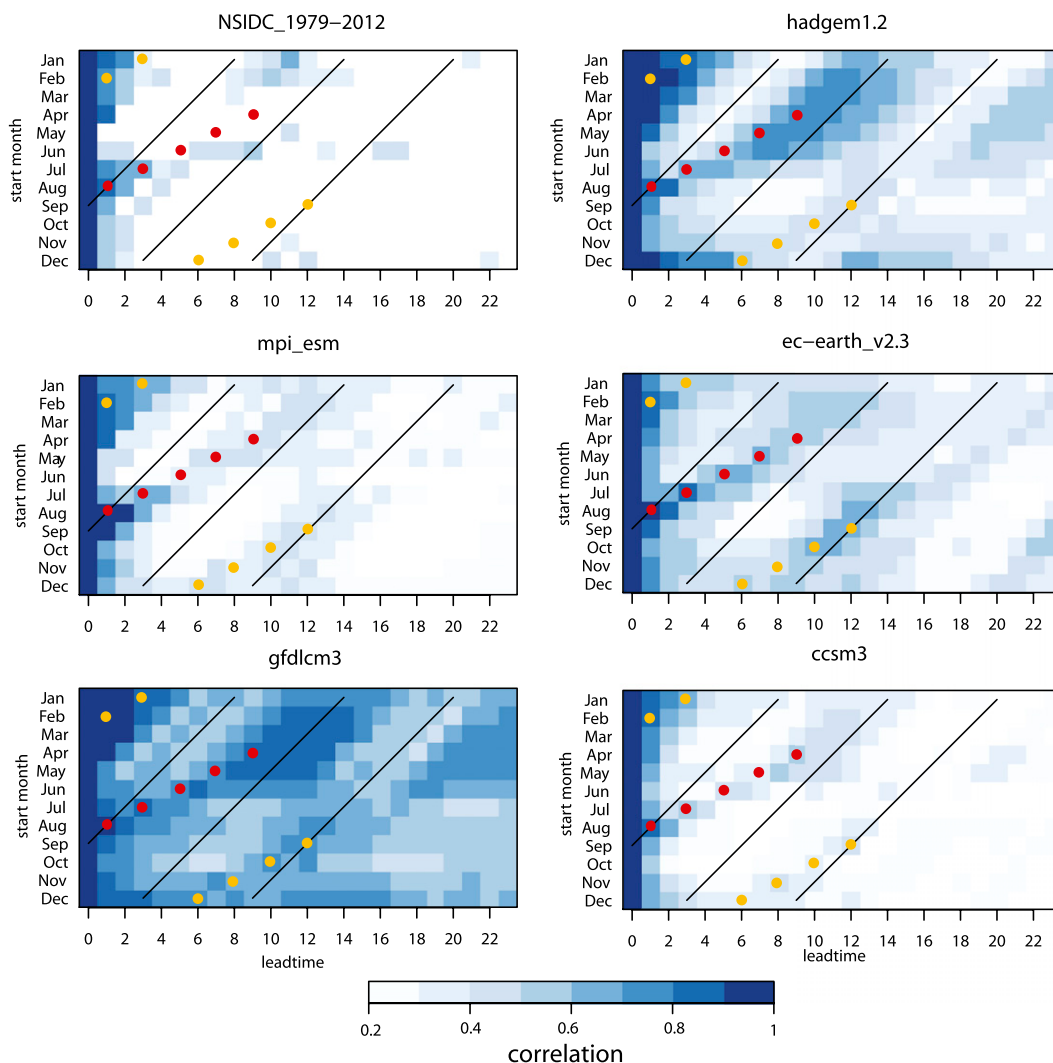
Extent

FIG. 1. Lagged correlation of NSIDC pan-Arctic sea ice extent derived from passive microwave observations (Meier et al. 2012) and GCM control simulations, for each start month, against lead time [i.e., from left to right, the top row shows January correlated with January (lag 0), January correlated with February (lag 1), etc.]. Black lines indicate values for September and March target months. Red dots indicate months that have increased correlation as a result of melt-to-freeze season reemergence. Similarly, orange dots indicate increased correlation as a result of the thickness related summer-to-summer reemergence mechanism.

year-to-year autocorrelation appear to correspond with differences in sea ice volume variability in the models (see Figs. 2a,b), rather than differences in the mean ice volume (see Figs. 2c,d). This appears to contradict Holland et al. (2010), who find that initial condition predictability is lower in a low sea ice state (using the same model in both climates). However, lagged correlation only measures one type of predictability and this is not a controlled experiment, since the models have many differences other than just their mean sea ice cover.

The lagged correlations from all the models indicate that forecasts initialized in January and July will lose skill slowly compared to other months, in part because of the thickness and melt-to-freeze season reemergence mechanism (see Fig. 5 of BW2011a). Correlations with October–December as the target month are also high in models and observations because of the melt-to-freeze season reemergence. The latter also seems to be the case in the initialized forecast system of Sigmond et al. (2013) (see their Figs. 3 and S1), in which forecasts of October sea ice extent are predicted skillfully at longer lead times
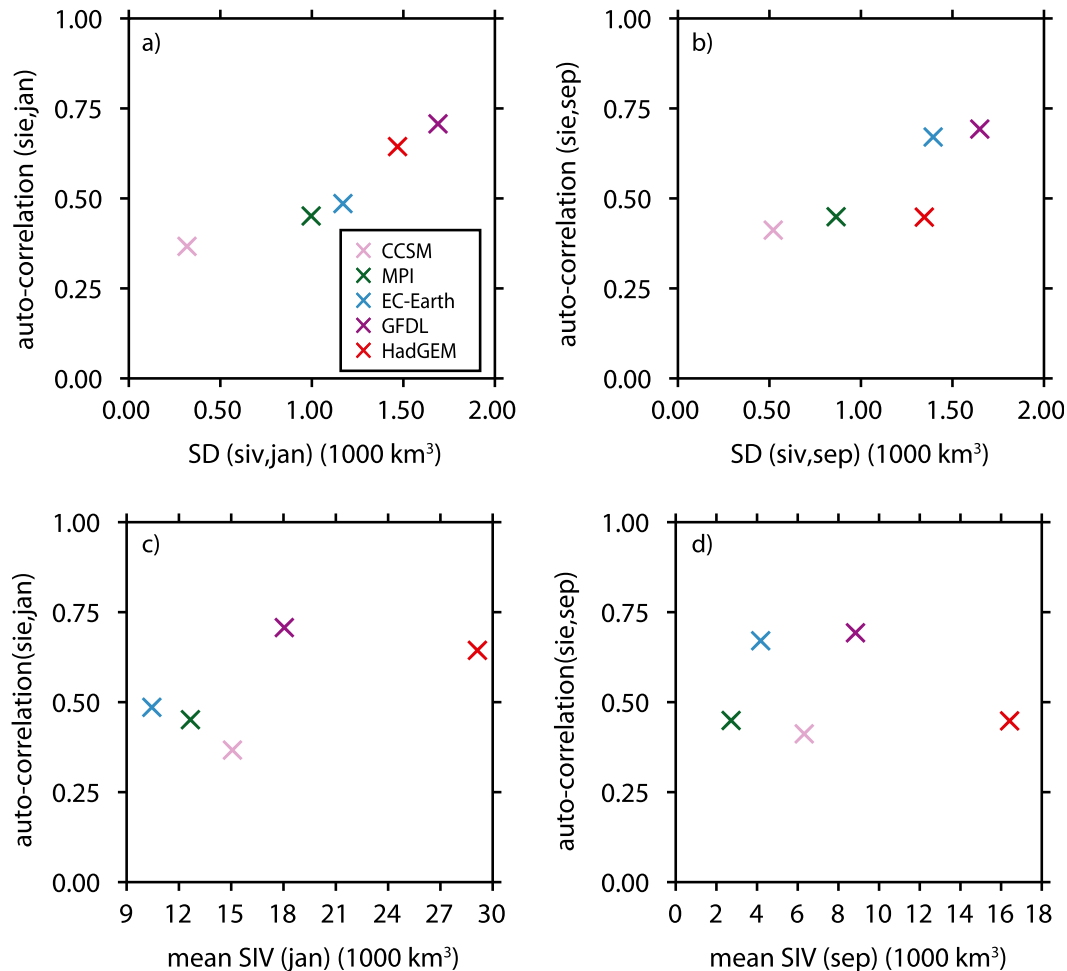
FIG. 2. Scatter diagrams of sea ice extent autocorrelation and volume standard deviation for (a) January and (b) September and sea ice extent autocorrelation and mean volume for (c) January and (d) September.

than other months because of this mechanism. The lack of sea ice thickness assimilation in this forecast system may be the reason why there is a lack of summer-to-summer reemergence in skill.

Lagged correlations in the GCMs are generally higher than in the observations at all (except lag 0) lead times (as also found by BW2011a). The reemergence of correlations, ubiquitous among the models, is also far smaller in the observations. This could be explained by the following:

(i) Sampling error: Only 34 yr of satellite observations exist, compared to over 200 simulated years for each model. It is possible that the observed period is an outlier.

(ii) Detrending: By detrending the observations (to remove the impact of anthropogenic climate change) it is possible that some of the low frequency internal variability is removed, therefore reducing the autocorrelation of the time series. Day

et al. (2012) estimate that between 5% and 30% of the observed decline since the 1970s could be caused by internal variability related to the Atlantic multidecadal oscillation (AMO).

(iii) Autocorrelation: The models could have a greater proportion of variability at low frequencies than the real world.

(iv) Inadequate representation of processes in models: It is possible that the processes associated with the reemergence are different in these models compared to the real system.

It is a challenge to distinguish between these potential causes, especially as it could be a combination of all four. For example, BW2011a show that lagged correlations calculated from different 30-yr instances of the CCSM3, exhibit a large spread, indicating the importance of point i. Although the satellite-era sea ice concentration data only cover the 34-yr period used in this study, one could

conceive of using additional historical observations of the sea ice (e.g., Rayner et al. 2003; Divine and Dick 2006) to distinguish between points i, ii, and iii, and this is planned further work.

b. Mechanisms for predictability

BW2011a argue that the melt-to-freeze season correlation reemergence (between pairs of months indicated by red dots in Fig. 1) is caused by the persistence of SST anomalies in the vicinity of the ice edge. This reemergence acts between pairs of months, one month in the melt season and one in the freeze season, when the ice edge is in approximately the same position in both months (e.g., May and December). They hypothesize that in the melt season month, an anomaly in the ice edge imparts an SST anomaly of the opposite sign. As the ice edge retreats, the SST anomaly persists in its original position but, when the ice edge returns to the same position, the SST anomaly imparts an ice edge anomaly of the same sign as the original sea ice anomaly.

When correlating May sea ice extent with both May and December SSTs for each GCM control simulation (Fig. 3), we find that significant correlations persist between May and December in the vicinity of the ice edge and marginal ice zone (indicated by the purple line). This supports the analysis of BW2011a by demonstrating that anomalies in May SST associated with sea ice edge anomalies persist through to December in all four models. This is a necessary condition for the melt-to-freeze season reemergence mechanism to exist and is ubiquitous across the models.

It is difficult to determine why some models have stronger melt-to-freeze season reemergence than others. It is likely that the duration of SST persistence in the models is important, but Fig. 3 indicates that the areas where SSTs cause reemergence differ between the models (i.e., correlations between sea ice extent and SST) are highest in the Bering Sea for HadGEM1.2 and highest in the Barents Sea for GFDL CM3.

The summer-to-summer reemergence is thought to be associated with persistence of thickness anomalies. BW2011a suggest that an anomaly in sea ice extent imparts a sea ice thickness anomaly, of the same sign and in the same region, adjacent to the ice edge. If this thickness anomaly persists until the following September then, as the ice edge returns to the position of the thickness anomaly at the end of the following melt season, the thickness anomaly imparts an ice edge/extent anomaly of the same sign as the previous winter. Correlating September sea ice extent and September thickness, it is obvious that anomalies in thickness are associated with anomalies in extent for each model (see Fig. 4, left). In each of the models this thickness anomaly persists

through to the September of the following year, acting as a potential reemergence mechanism (see Fig. 4, right).

We hypothesize that the strength of the ice volume related summer-to-summer reemergence is proportional to the size of the standard deviation of sea ice volume. Those models with larger sea ice volume and thickness variability will have larger thickness anomalies, which are more likely to persist through the year and cause reemergence. This relationship appears to be supported by Fig. 2b, but it is impossible to prove with only five models.

Both mechanisms of reemergence discussed in BW2011a therefore appear to be supported by analysis with four additional models presented here. However, extending this correlation analysis to a detrended observational sea ice concentration and SST dataset [Rayner et al. 2003; Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST)] does not reveal such high correlations (not shown). This is consistent with the lack of reemergence in the lagged correlation in observed sea ice extent (see Fig. 1a).

c. Sea ice volume predictability

Lagged correlations indicate that sea ice volume is much more persistent than extent. They also show that sea ice volume also has a melt-to-freeze season reemergence, similar to sea ice extent (Fig. 5). This is a feature of PIOMAS and all models apart from CCSM3. Again, correlations between May sea ice volume and May SST at the ice edge persist through until December (not shown), indicating that the mechanism for melt-to-freeze season sea ice volume reemergence is the same as that of extent.

A summer-to-summer reemergence, such as that exhibited by sea ice extent, is not exhibited by sea ice volume. However, one would not necessarily expect sea ice volume to have this property. For sea ice extent, this reemergence is associated with the persistence of thickness anomalies in the central Arctic that decouple from extent when the ice edge moves south and recouple (and reemerge) when it moves back northward to the location of this anomaly. No such decoupling would be expected for volume.

d. Seasonal cycle of predictability

Importantly, Figs. 1 and 5 show that the lagged correlation of both extent and volume decrease with time at different rates, depending on the start month. This is most pronounced for extent; the lag 1 correlation has a strong seasonal cycle (see Fig. 6a) as a result of both reemergence mechanisms, which increase the lag 1 correlation in the late winter and late summer, respectively. Start months in May–July and October–December have
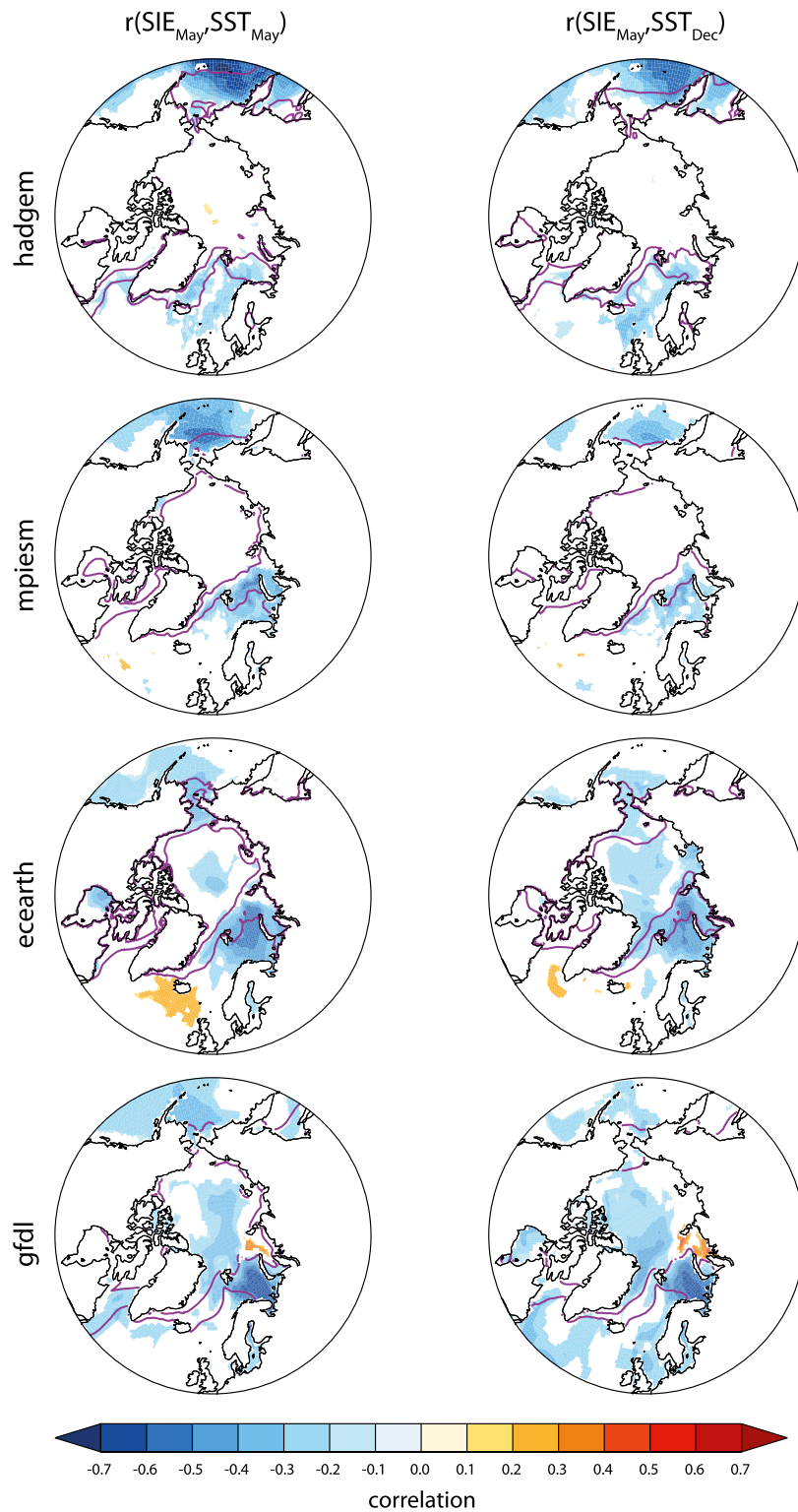
FIG. 3. Maps of the correlation between May Arctic sea ice extent and gridpoint SST for (left) May and (right) December. Each row corresponds to a different GCM, as labeled. Only correlations significant at the 99% level are plotted. The purple line is the 10% contour of standard deviation of (left) May and (right) December monthly mean sea ice concentration, indicating the approximate region of ice edge variability for each month.
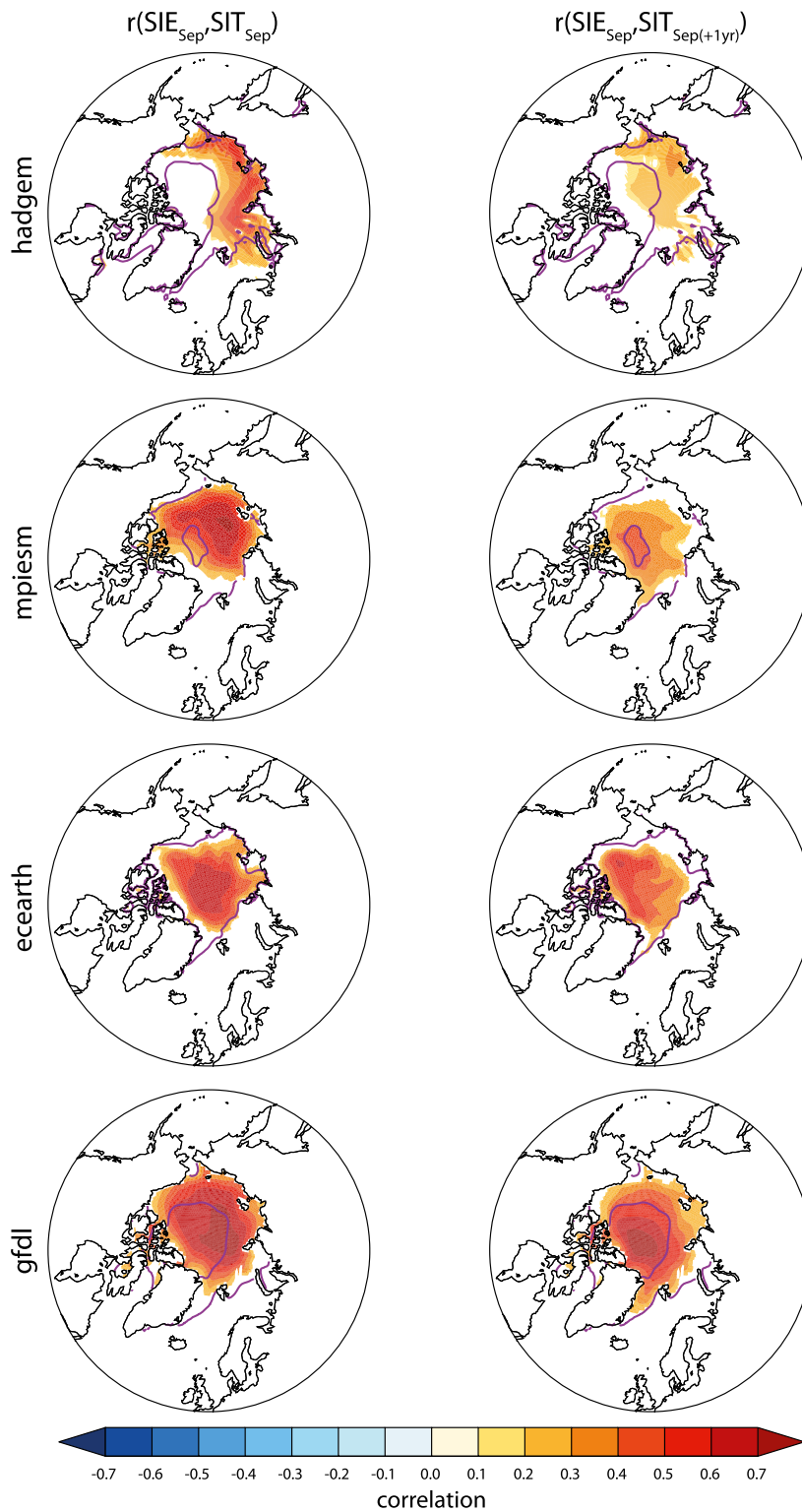
FIG. 4. Maps of the correlation between September Arctic sea ice extent and monthly mean gridpoint September sea ice thickness at (left) zero lag and (right) 1-yr lag. Each row corresponds to a different GCM as labeled. Only correlations significant at the 99% level are plotted. The purple line is the 0.1 contour of standard deviation of September monthly mean sea ice concentration; this indicates the approximate region of ice edge variability.
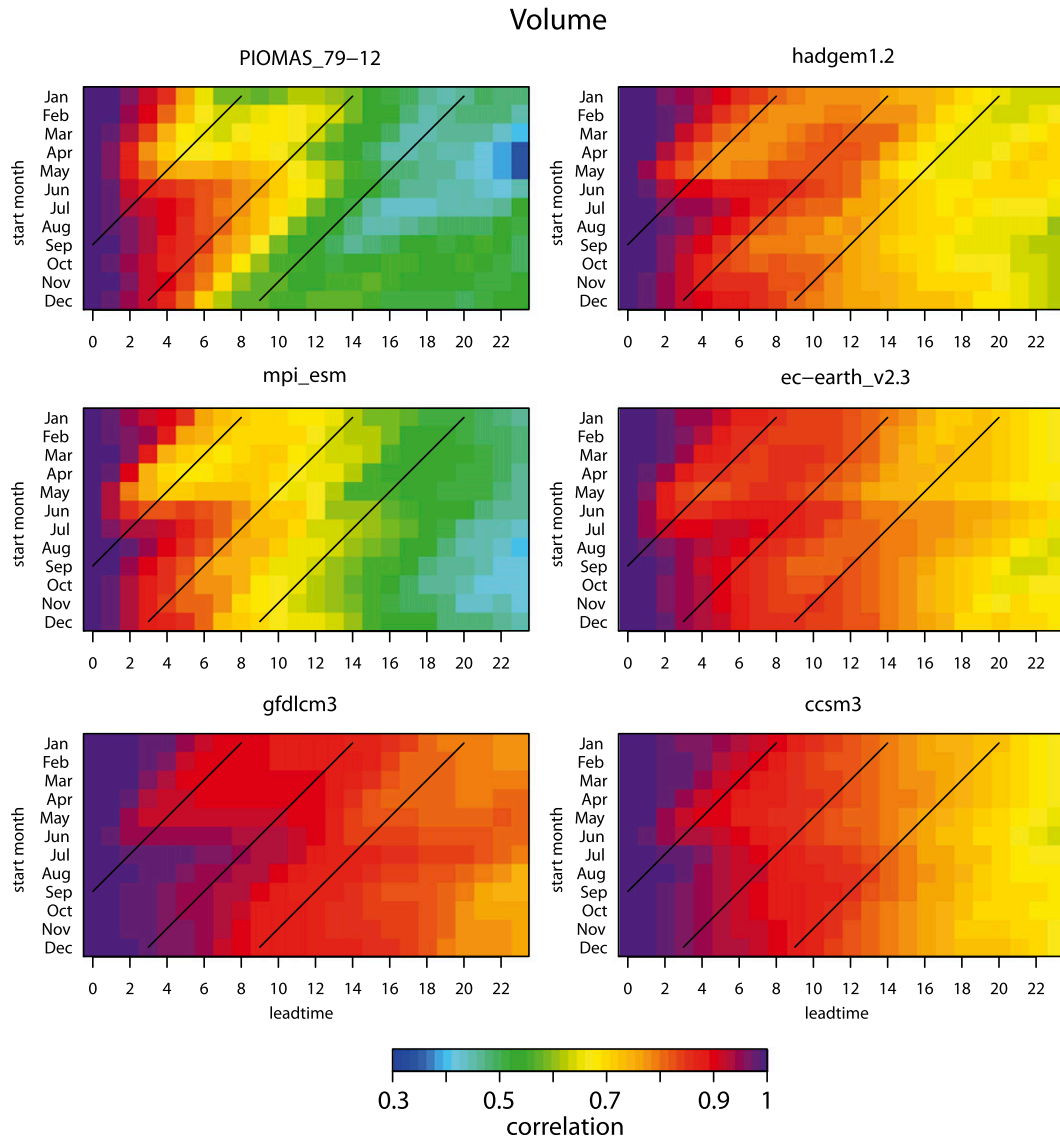
Volume

FIG. 5. Lagged correlation of pan-Arctic sea ice volume in PIOMAS (Zhang and Rothrock 2003) and GCM control simulations is plotted, for each start month, against lead time [i.e., from left to right, the top row shows January correlated with January (lag 0), January correlated with February (lag 1), etc.]. The diagonal black lines indicate September and March values.

low lag 1 correlation since the ice edge is moving most rapidly (see Fig. 6c). The lag 1 correlation seasonal cycle for volume is simpler than extent, having only one minimum between May and July, corresponding to the months with the largest change in volume (Figs. 6b,d). Using sea ice extent from previous months as a predictor for September sea ice extent, Fig. 6e indicates that the skill of a September forecast will increase rapidly as the predictor month approaches September. This is not the case for volume, which does not have this sharp drop off in skill as the predictor is moved away from September (Fig. 6f).

These diagnostic analyses of control simulations indicate that potential forecast skill might be very dependent on start month. We next explore whether this is also the case in perfect model predictions.

4. Perfect model predictability experiments

To investigate the start date dependence of initial condition predictability, a set of perfect model ensemble prediction experiments were run using HadGEM1.2. These simulations were started from initial states of 1 January, 1 May, and 1 July in each of 10 yr chosen from
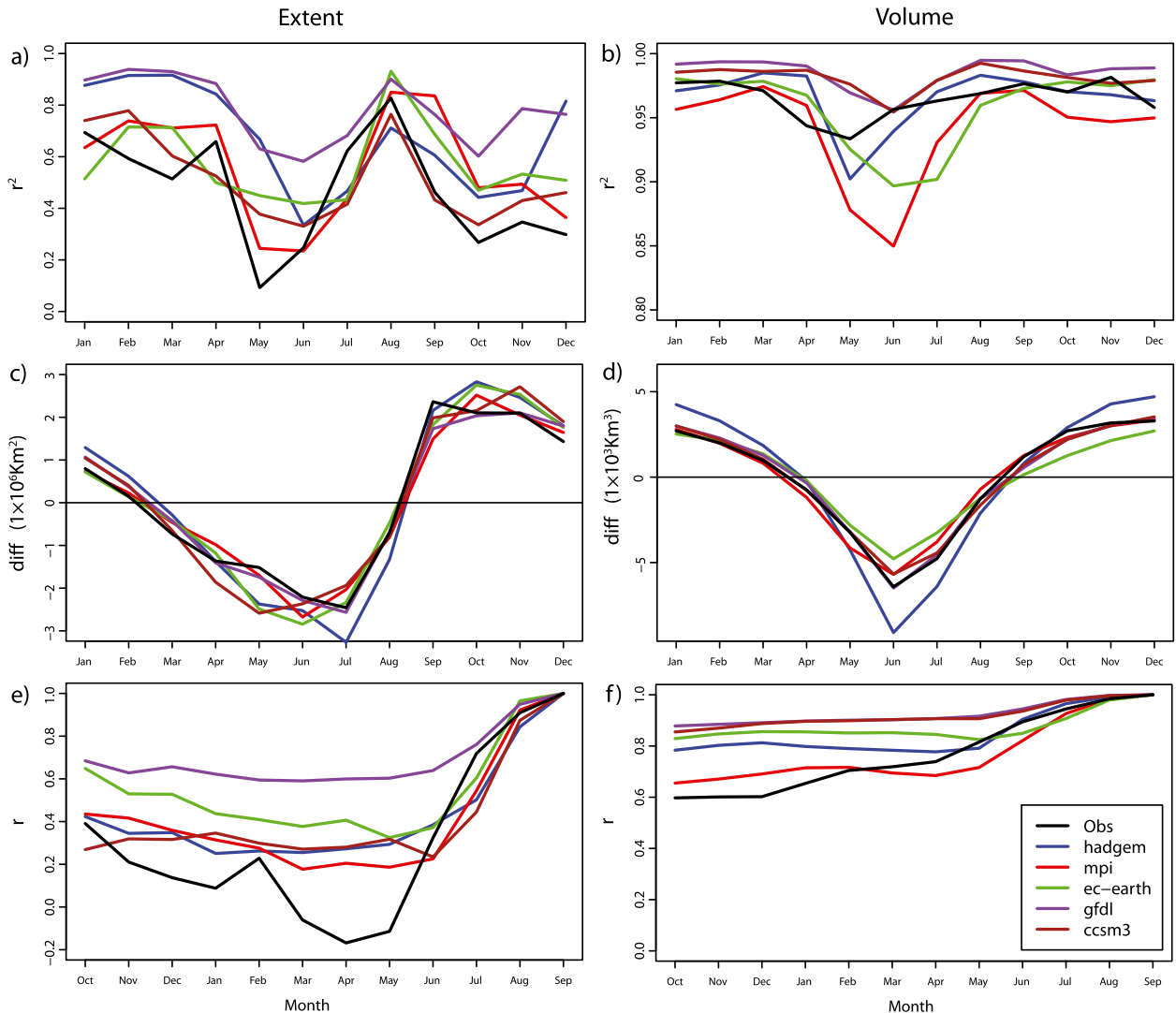
FIG. 6. Lag 1 autocorrelation for (a) extent and (b) volume; month-to-month change in (c) extent and (d) volume; and correlation between monthly mean (predictor) and September mean (predictand) (e) extent and (f) volume.

the control simulation to sample a range of sea ice states. Each ensemble contains 16 members with each member having identical initial conditions to the reference run, except for a tiny, spatially varying, Gaussian white noise perturbation (with $\sigma = 10^{-4}$ K) to the SST field. This is similar to the methodology used to assess sea ice predictability in Koenigk and Mikolajewicz (2008) and Blanchard-Wrigglesworth et al. (2011b).

Perfect model experiments such as those described above do not suffer from model error because the model is being used to predict itself. Neither do they suffer from a lack of information about the initial state, since the full atmosphere–ice–ocean state of the reference simulation is known precisely for each of the start times. For this reason, the predictability estimates obtained give an upper bound

to the predictability obtainable with the HadGEM1.2 simulator (e.g., Collins 2002; Latif et al. 2006).

a. Predictability metrics

To define predictability in this study, we use two predictability metrics as defined by Collins (2002). In such a perfect model study, any ensemble member may be chosen as the "truth" and the effective sample size can be increased by taking each member in turn. The ensemble RMSE is then defined as

$$\text{RMSE}(t) = \sqrt{\frac{1}{n} \sum_{j=1}^{N_s} \sum_{i=1}^{N_m} \sum_{k \neq i} [x_{kj}(t) - x_{ij}(t)]^2}, \quad (1)$$

where $x_{ij}(t)$ is the sea ice extent at lead time t for the ith member of the jth ensemble and $n = N_s N_m (N_m - 1) - 1 = 2399$, where $N_s(=10)$ is the number of start dates and $N_m(=16)$ is the number of ensemble members. The state is predictable at lead time t when $\mathrm{RMSE}(t) < \sqrt{2}\mathrm{RMSE}_{\mathrm{ctrl}}(t)$, where $\mathrm{RMSE}_{\mathrm{ctrl}}(t)$ is the RMSE of a climatological reference forecast [calculated by replacing x_{ij} with the climatology of the reference simulation in Eq. (1)] and the significance of this inequality is calculated using an f test with $159(=N_m N_s - 1)$ degrees of freedom.

To compare the predictability of initialized forecasts with the lagged correlation properties of the control simulation, the anomaly correlation coefficient (ACC) is used. This is defined as

$$\mathrm{ACC}(t) = \frac{\langle [x_{kj}(t) - \overline{x}(t)][x_{ij}(t) - \overline{x}(t)] \rangle_{i,j,k \neq i}}{\langle [x_{ij}(t) - \overline{x}(t)]^2 \rangle_{i,j}}, \quad (2)$$

where $\langle \cdot \rangle_i$ denotes the expectation value, to be calculated by summing over the specified index, and $\overline{x}(t)$ is the monthly mean climatology of x at lead time t (calculated from the control simulation) and also treats each ensemble member in turn as the truth (Collins 2002).

b. Lagged correlation compared to perfect model skill in HadGEM1.2

The lagged correlation (a diagnostic measure of predictability) is compared with the ACC (a prognostic forecast metric) of the initialized perfect model runs in Fig. 7. Each of the lagged correlation time series (dashed lines) consists of the values along a row of the HadGEM1.2 panel of Fig. 1 for extent and Fig. 5 for volume.

Figure 7a clearly shows that the reemergence of extent seen in the lagged correlation metric is also a feature of the initialized runs (solid lines) for each start month. Furthermore, even though the volume lagged correlation time series does not show very strong reemergence, the ACC of the initialized forecasts do (Fig. 7b). Peculiarly, reemergence of the May ensembles' ACC skill in predicting extent results in higher values of ACC the following winter than the July ensemble, even though the July ensemble is initialized closer to the verification time. This property is a feature of the lagged correlation measure but is not seen in the RMSE score; thus, this feature may be due to having a relatively small sample size (see Fig. 11a). Surprisingly, the volume lagged correlations in Fig. 7b are higher than the perfect model ACC, indicating that lagged correlation statistical forecast outperforms the perfect model forecast. We can only assume that the particular start dates chosen for

initializing the GCM ensembles were relatively unpredictable for volume.

Focusing on the first forecast year and lining up the time series as a function of lead time, rather than verification month, highlights the start date dependence of skill (Figs. 8a,b). Skill in the May ensemble forecast of extent declines much more rapidly over the first 4 months than the January and July ensembles, before skill reemerges over the next 5 months. There is also a reemergence of skill in the May and January forecasts, approximately mirroring the timing of the lagged correlation reemergence. Figure 8b shows that volume may also have some initialization month dependence, with the May ACC and lagged correlation declining faster than those for January and July, out to a lead time of 5 months.

This comparison between lagged correlation in the control run and anomaly correlation in the initialized forecasts indicates that lagged correlation is an informative measure of sea ice extent and volume predictability in HadGEM1.2. The intermodel similarity of the lagged correlation properties and their reemergence mechanisms gives confidence that these mechanisms are not model dependent.

It is also possible to construct a simple statistical forecast for extent and volume from the lagged correlation for each start month,

$$x(t_0 + \tau) = \beta(\tau)x(t_0), \quad (3)$$

where x is the extent/volume anomaly, t_0 is the forecast start month, τ is the lead time in months, and $\beta(\tau)$ is the autocorrelation of x at a lag of τ (e.g., Hawkins et al. 2011). Figure 9 shows the RMSE of the perfect model ensemble and the RMSE of the lagged correlation forecast, which was calculated by replacing the second term in Eq. (1) with the lagged correlation statistical forecast for each start year. It is noticeable that the lagged correlation forecasts behave similarly to the perfect ensembles in a number of ways: May and July forecast error grows faster than January and volume is much more predictable than extent.

Comparison of the lagged correlation forecast with the initialized ensemble also indicates that, for extent, there is much more skill in the initialized forecasts. This indicates that, while understanding the mechanisms that cause the lagged correlation patterns are important, there are sources of skill that are not included in that measure. For sea ice volume, this result is more intuitive than that shown in Fig. 7, where overall skill (lagged correlation of the whole control run) is compared with the skill of certain start dates. This also highlights the importance of considering multiple predictability metrics.
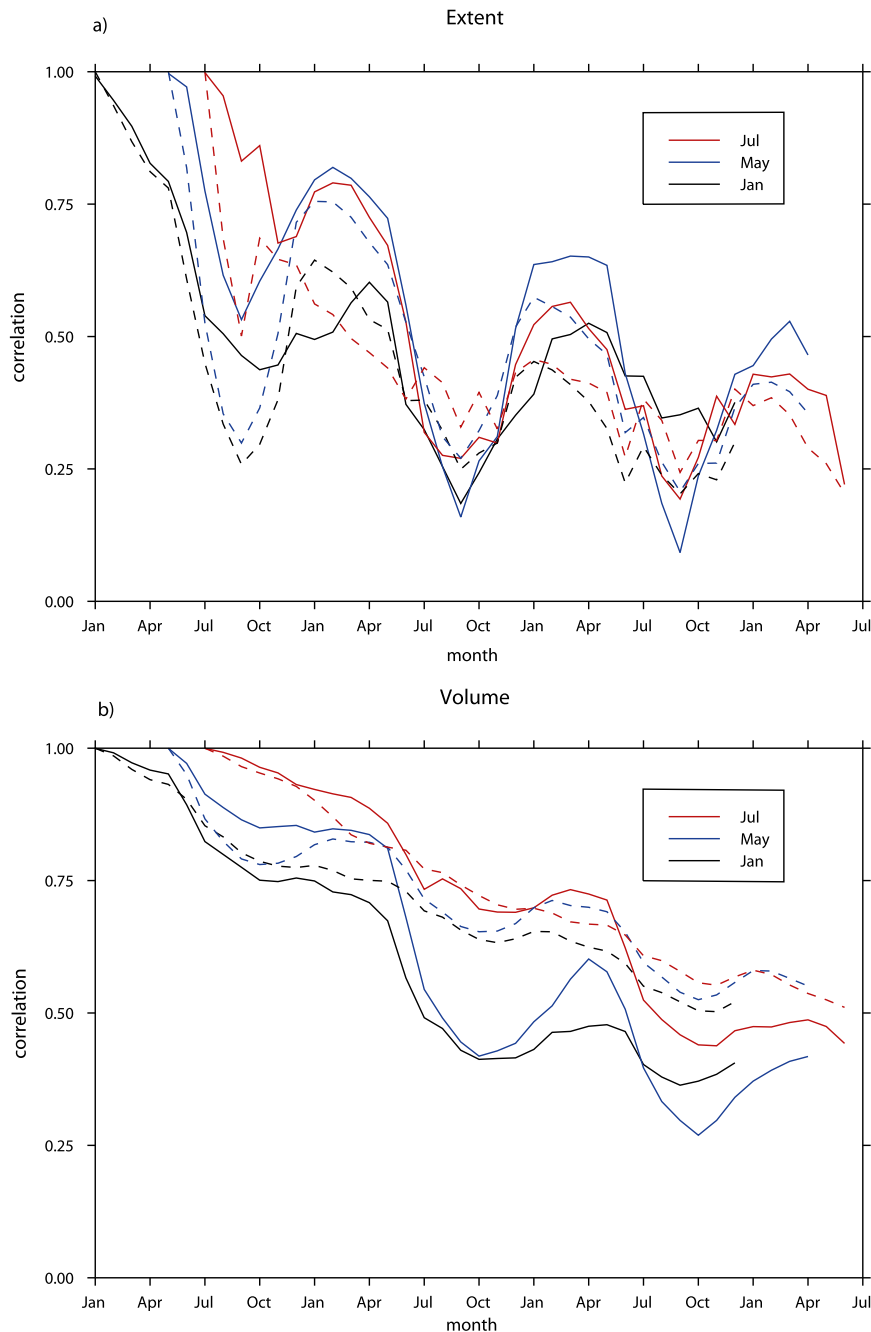
FIG. 7. Ensemble ACC for HadGEM1.2 perfect model predictions (solid line) and
HadGEM1.2 control simulation lagged correlation (dashed line) for (a) extent and (b) volume.

On the other hand, the lagged correlation volume forecasts perform relatively well compared to the perfect model ensemble, indicating that this type statistical forecast may work well for predicting volume. The initialized forecasts have significantly more skill than the lagged correlation forecast in lead months 1–4 for forecasts initialized in January, 0–7 for forecasts initialized in May, and 0–4 and 7–10 when initialized in July. Interestingly,

the lagged correlation forecast performs well at lead times of over a year for all start months.

c. Potential skill in regional sea ice prediction

1) EXTENT

Analysis of the predictability properties of sea ice extent and volume is next performed for the regional
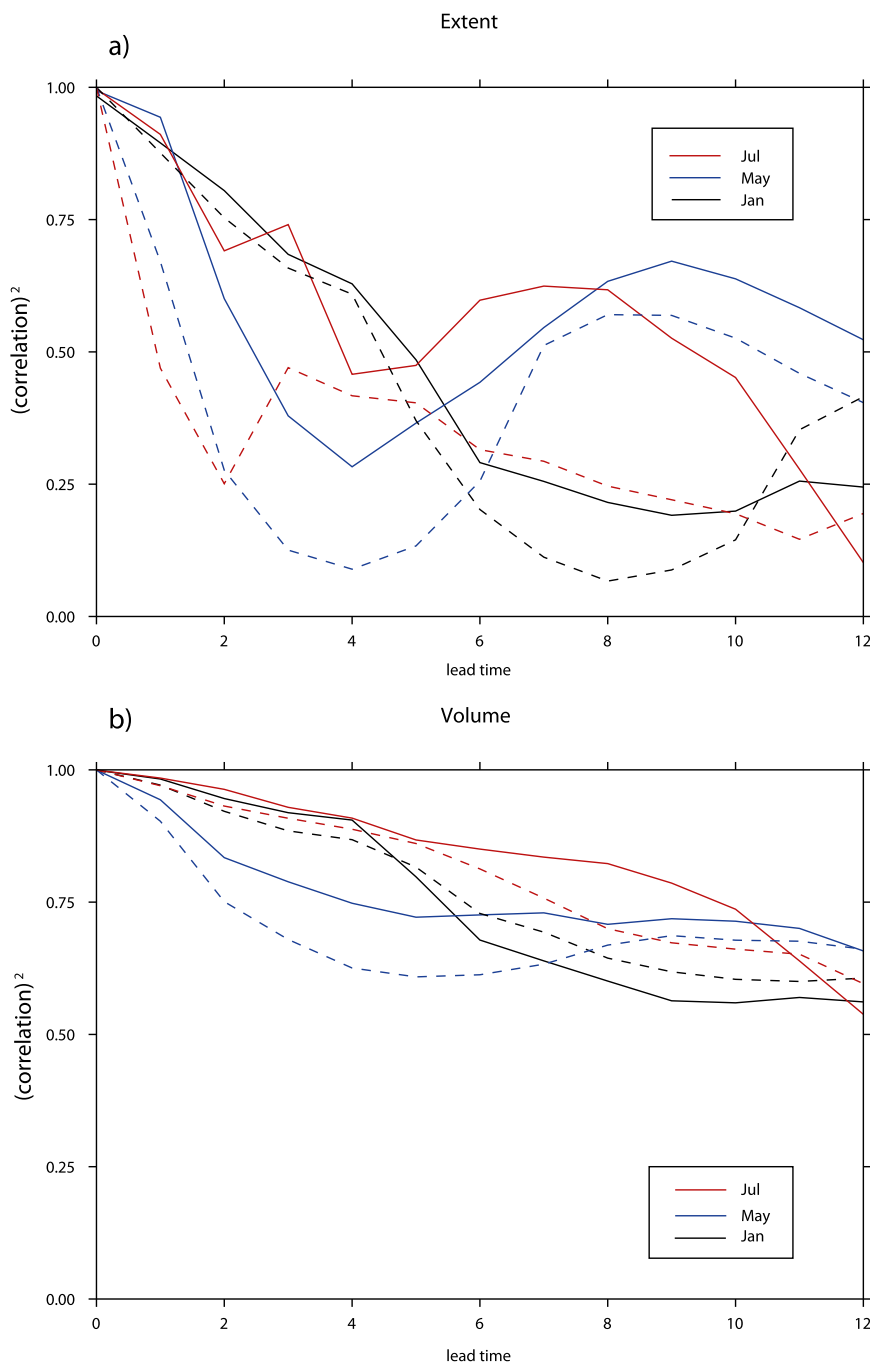
FIG. 8. Squared ensemble ACC for HadGEM1.2 perfect model predictions (full line) and HadGEM1.2 control simulation lagged correlation (dashed line) for (a) extent and (b) volume. These have been plotted as a function of lead time and squared to highlight the differences between ensembles.

areas defined in Fig. 10. To quantify the predictability we calculated the RMSE for each basin as shown in Eq. (1). The Canadian Archipelago is excluded from the analysis; one would not expect this bathymetrically complex area to be well simulated by a GCM of this spatial resolution.

The significant skill in pan-Arctic extent prediction, as measured by RMSE, is continuous in the January ensemble, out to the second June (see Figs. 9a, 11a). Extent is not significantly predictable during the subsequent summer [July–November (JASON)], but skill returns through the winter until the subsequent JASON. The May
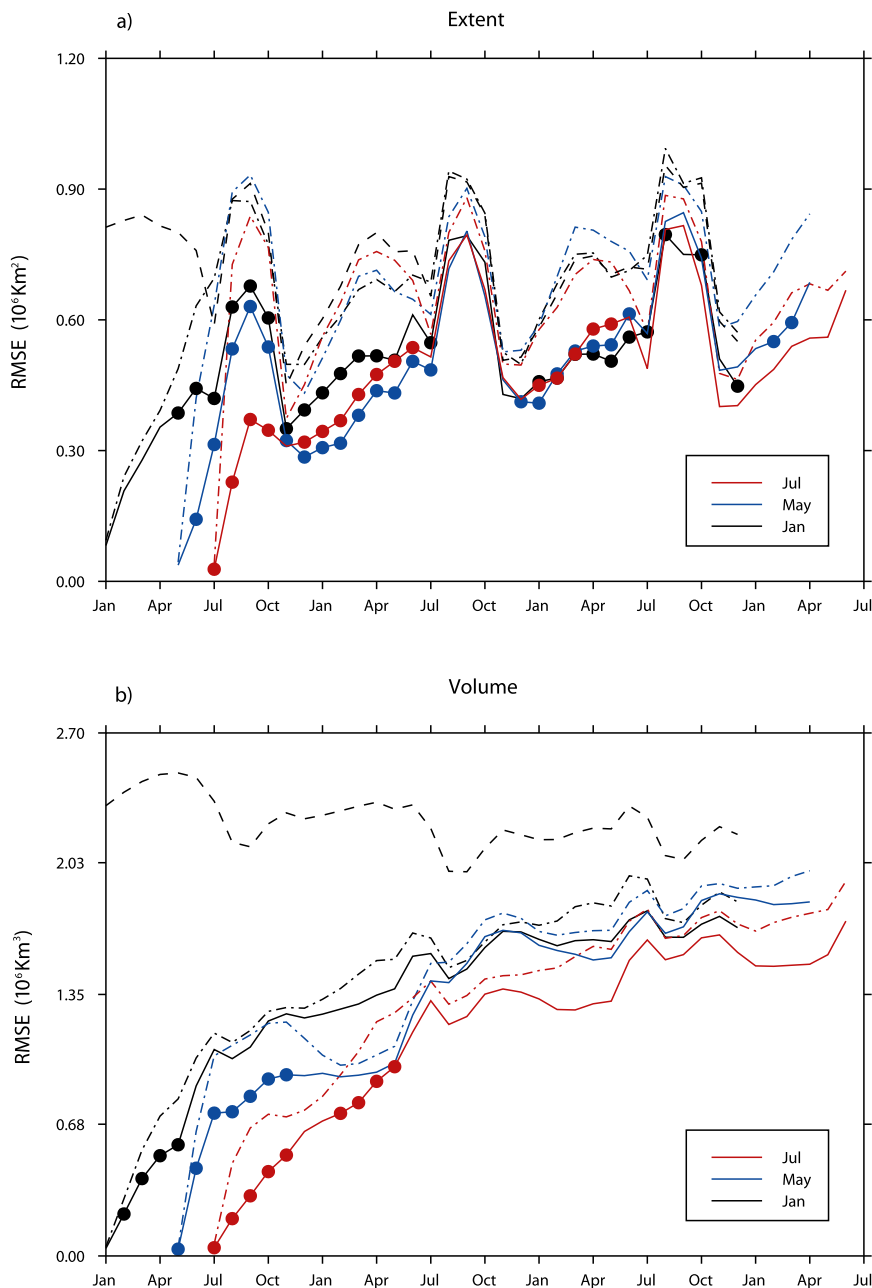
FIG. 9. Ensemble RMSE for HadGEM1.2 perfect model predictions (solid line) and the RMSE of a lagged correlation statistical forecast, derived from control simulation lagged correlations (dotted–dashed line) for (a) extent and (b) volume. The limit no predictability for the January ensembles is shown by the dashed line ($\sqrt{2} \times \mathrm{RMSE}_{\mathrm{ref}}$). Dots on the solid lines indicate where the perfect model RMSE is less than that of the statistical forecast.

ensemble is significantly more skilful than the January ensemble for the first 3 months, after which the skill is not significantly different from the January ensemble through ASON. After this, significant skill returns for 4 months, after which time the ensembles are indistinguishable. The July ensemble is only significantly more skilful than May ensemble for the first 4 months (JASO),

after which point it is slightly less skilful than May, presumably because of sampling issues.

In terms of predictability, the basins can be roughly split into two groups: those in the central Arctic with perennial sea ice cover and those in the peripheral basins with seasonal ice cover. The first of these groups includes the central Arctic, Kara, Laptev and Siberian,
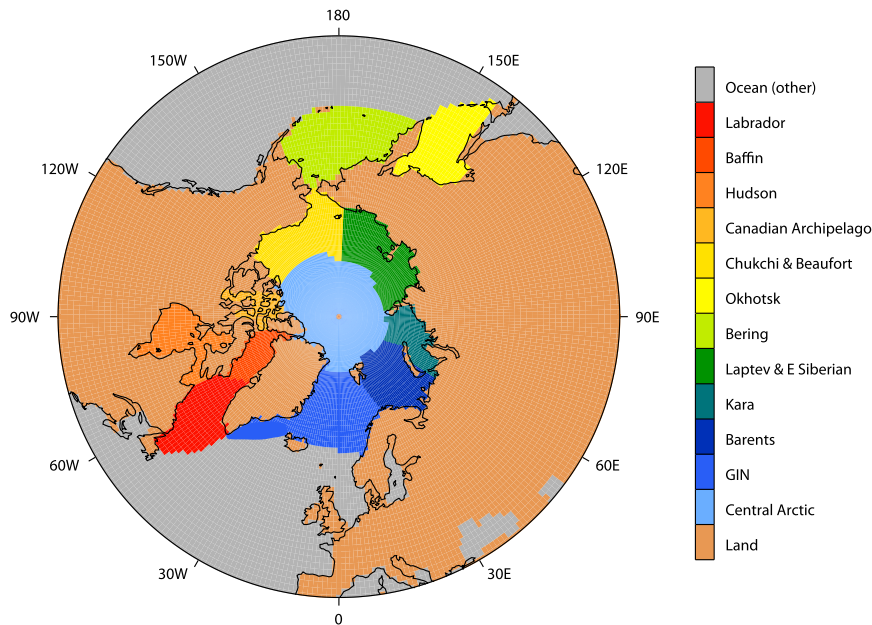
FIG. 10. Map of the Arctic Ocean and basins.

Chukchi and Beaufort, Hudson, and Baffin regions (Figs. 11b,e,f,h,j,k). Apart from the central Arctic, extent in these regions is trivially predictable in the non-JJASON months because of extremely low climate variability. During the JJASON period, the January ensemble only has skill for a very limited number of months in some of the basins and none in others. Similarly, for the May ensembles, only in the Chukchi and Beaufort region does the May ensemble have skill for all of JJASON. Only from the July ensemble are the summer conditions in these basins simulated for this period with significant skill.

The second set, the peripheral regions, including the Greenland–Iceland–Norwegian (GIN), Barents, Bering, Okhotsk, and Labrador Seas (Figs. 11c,d,g,i,l), have quite different characteristics. The duration of predictive skill is much higher in these basins, with the January ensemble exhibiting continuous skill for 16 months in the GIN and Barents Seas and longer in the Labrador Sea. In these three basins, the May ensembles have smaller RMSE than the January ensembles at all verification times and so too the July ensembles with respect to May. The duration of skill in the Bering and Okhotsk Seas is less than the basins on the Atlantic side but still longer than the central basins. The relatively long duration of skill in the peripheral seas, compared to the central regions, is the reason the pan-Arctic RMSE shows skill for longer in winter than summer (see Figs. 11a,l). This is because during the winter the ice edge is in the regions where extent is predictable at long lead times.

The relatively long duration of skill in the peripheral seas is likely to be caused, at least in part, by the interaction of the sea ice with slowly evolving modes of climate variability, such as the AMO (e.g., Mahajan et al. 2011; Day et al. 2012). Such modes exhibit predictability on interannual to decadal time scales (e.g., Meehl et al. 2009; Matei et al. 2012). Sea ice in the Labrador Sea, which exhibits the longest duration of predictability, is close to the subpolar gyre, an area where SSTs shows high levels of potential predictability (e.g., Collins 2002; Boer 2004).

2) VOLUME

Similarly to Blanchard-Wrigglesworth et al. (2011b), pan-Arctic volume is continuously predictable for all months of the 3-yr integrations (see Figs. 9b, 12a). The May ensembles exhibit significantly more skill than the January ensembles for over 1 yr, and the July ensembles only more skill than the May ensembles for the first 8 months, after which the ensembles are statistically indistinguishable. At the regional scale a similar pattern is observed with volume generally predictable for as long as an extent or longer in the same region (see Fig. 12). In contrast to sea ice extent, both May and January ensembles exhibit skill in predicting volume in the central Arctic basins during the first summer (Figs. 12b,e,f,h,j,k).

It is clear from Fig. 12b that the May ensembles lose skill in predicting central Arctic sea ice volume more rapidly than the January and July ensembles. This is also the case for the Kara, Laptev and east Siberian, Chukchi
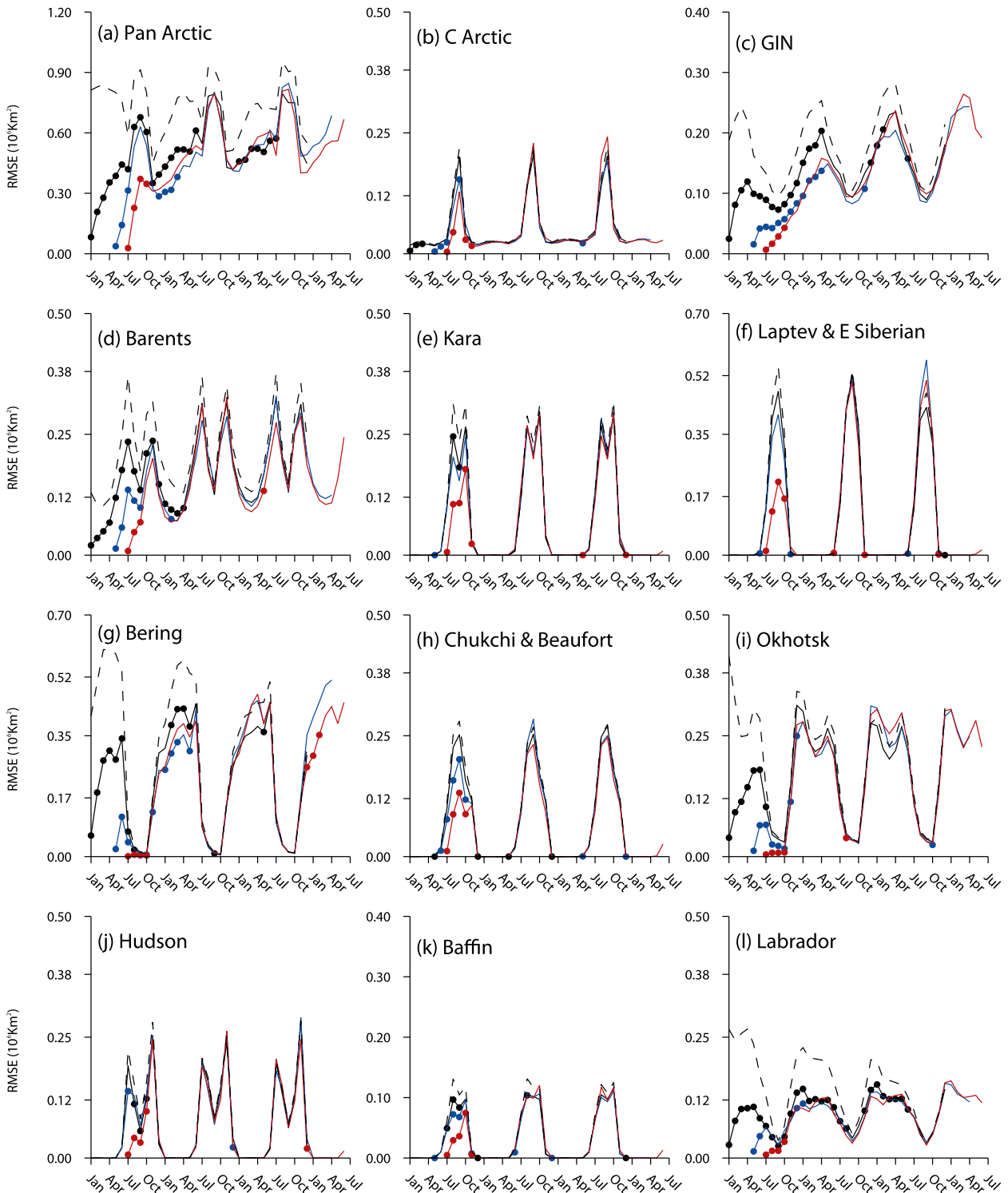
Fig. 11. RMSE of pan-Arctic and basin sea ice extent for January (black line), May (blue line), and July (red line) ensembles. The limit of no predictability is indicated by the dashed line ($\sqrt{2} \times \mathrm{RMSE_{ref}}$). Dots indicate times when the following relations are significant at the 99% level: $\mathrm{RMSE_{Jan}} < \mathrm{RMSE_{ref}}$, $\mathrm{RMSE_{May}} < \mathrm{RMSE_{Jan}}$, and $\mathrm{RMSE_{Jul}} < \mathrm{RMSE_{May}}$.
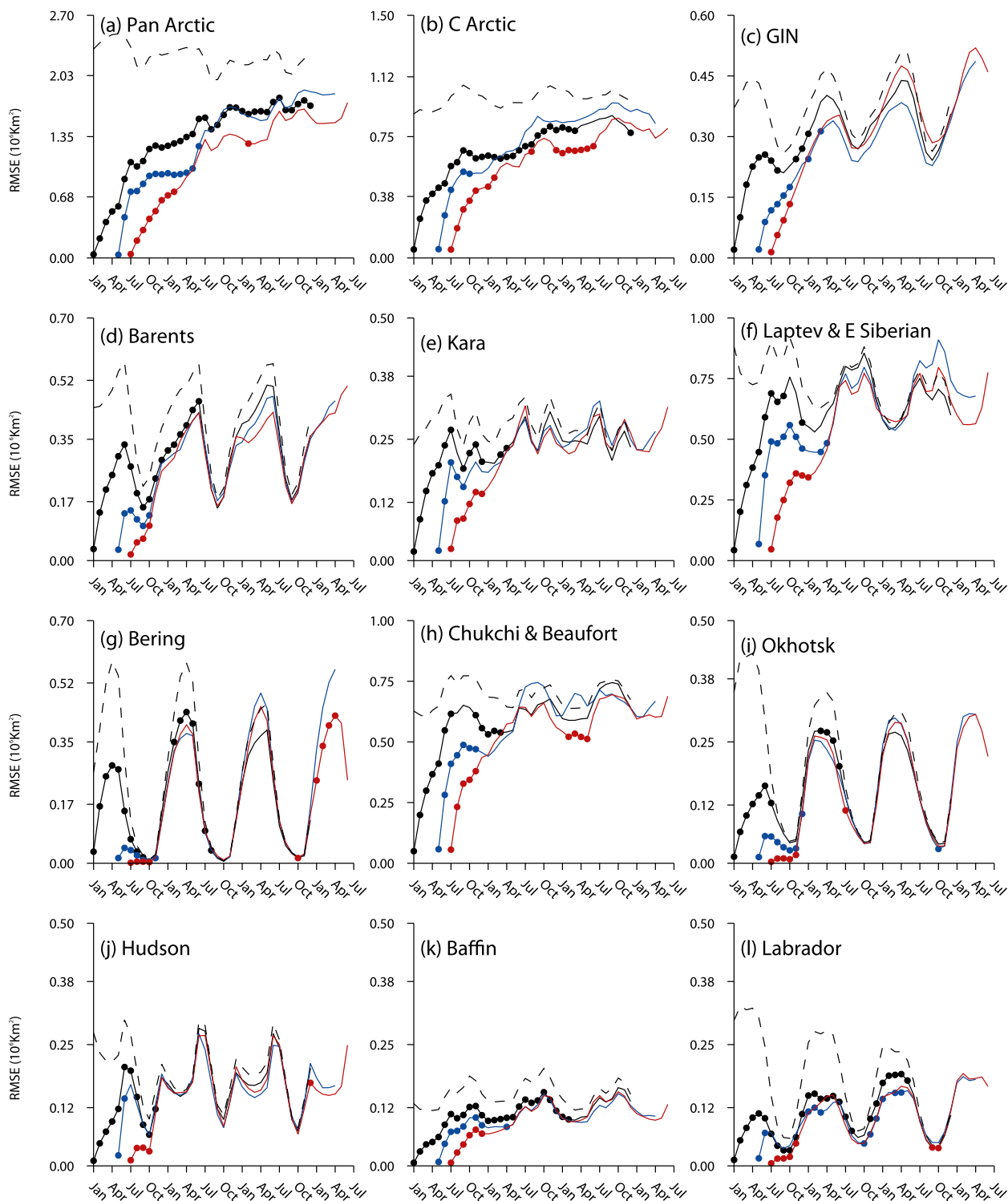
FIG. 12. RMSE of pan-Arctic and basin sea ice volume for January (black line), May (blue line), and July (red line) ensembles. The limit of no predictability is indicated by the dashed line ($\sqrt{2} \times \mathrm{RMSE_{ref}}$). Dots indicate times when the following relations are significant at the 99% level: $\mathrm{RMSE_{Jan}} < \mathrm{RMSE_{ref}}$, $\mathrm{RMSE_{May}} < \mathrm{RMSE_{Jan}}$, and $\mathrm{RMSE_{Jul}} < \mathrm{RMSE_{May}}$.

and Beaufort, Hudson, and Baffin regions. Ensemble RMSE is affected by the seasonal cycle of variance (i.e., low variance in the reference simulation implies low RMSE in the ensemble); in some regions that have low September volume variance, such as the Sea of Okhotsk, ensembles are indistinguishable during September. This corresponds with low ice volumes in these regions during September. In each of the regions (excluding Hudson Bay), the May ensembles have additional skill over the January ensembles for the first 5–9 months and the July ensembles have additional skill over the May ensembles for the first 4–7 months.

5. Summary and discussion

Information relevant to the development of seasonal-to-interannual Arctic sea ice prediction systems has been presented. In particular, the intermodel robustness of sea ice cover persistence properties and mechanisms has been investigated through diagnostic analysis of control simulations with five different GCMs. The relevance of these properties for initialized prediction was then investigated through a series of idealized, "perfect model" runs with HadGEM1.2.

a. Conclusions

The key conclusions regarding Arctic sea ice predictions are as follows:

- The models considered exhibit similar sea ice extent predictability properties (measured by lagged correlation, as discussed in BW2011a), with all models exhibiting correlation reemergence in certain months because of the persistence of SST and sea ice thickness anomalies; however, the strength of correlation varies between the models.
- The lagged correlation of sea ice extent observations and volume reanalysis are smaller than all models considered in this study. However, the reasons for this are not clear.
- The lagged correlation of sea ice volume is generally higher than extent, with larger differences between the models. However, some of the processes governing the temporal patterns of lagged correlation appear to be similar. For example, most of the models exhibit a melt-to-freeze season reemergence of lagged correlation similar to that seen in extent, which our analysis indicates is driven by the same SST persistence mechanism as extent reemergence and occurs in the same months.
- Lagged correlations with all models indicate that the skill of predictions for both the extent and volume summer minima improve sharply when the initialization time is after May for a September verification

time. This is also true for initialized perfect model predictions, which show that forecasts initialized in May lose skill more rapidly in the first 4 months of the forecast than those initialized in January or July. This is because reemergence mechanisms act in the months after July and January to reduce the rate of decay in correlation compared to May.
- Idealized forecasts of pan-Arctic extent during summer months are not significantly more skilful when initialized from May compared to January. This is also indicated by the lagged correlations which decay much more rapidly during early summer than during winter months.
- Regionally, September extent in the central Arctic and marginal basins is generally only predictable from the ensembles initialized in July, but are trivially predictable outside of summer as these regions are entirely ice covered (in the model). Volume is predictable for longer than extent in all regions.
- Sea ice extent in the seasonal ice zone of the North Atlantic regions considered are predictable 1.5–2.5 yr ahead.

b. Discussion and implications

These conclusions are relatively consistent with the analysis of the CanSIPS operational initialized forecast system. Sigmond et al. (2013) find that July–October pan-Arctic sea ice extent variability is not skillfully hindcast in ensembles initialized in May or before. In their analysis, winter conditions are predictable from ensembles initialized as much as 1 yr before the verification time. Similar findings for the National Centers for Environmental Prediction (NCEP) Climate Forecast System, version 2 (CFSv2) are found by Wang et al. (2013). The similarity of conclusions suggests that the spring time rapid drop off in predictability reported in the operational systems is inherent to the model's climate, rather than due to uncertainty in the initial conditions. However, one would expect skill in these operational systems will decrease more rapidly with lead time than in the perfect model forecasts because of the imperfect initialization of key fields such as sea ice thickness.

It is also clear from studying the regional properties of sea ice cover predictability, although perhaps trivial, that the duration of skill in the pan-Arctic extent and volume is derived from the properties of the basins that make up the ice edge at a given month of year. For example, the ensembles are only skilful in predicting pan-Arctic extent for the first summer and the regions in the Arctic basin. During winter months, however, when the ice edge is in the peripheral seas, the extent in the peripheral seas and pan-Arctic is predictable for all winters of the simulations. Conversely, because the largest volume

of ice is in the central Arctic, the pan-Arctic volume predictability characteristics are inherited from this region.

The conclusion that summer extent in the central Arctic and marginal seas, at least in HadGEM1.2, is much less predictable from ensembles initialized on (or before) 1 May than those after could be important for operational forecast centers and end users such as for Arctic shipping (Khon et al. 2009; Stephenson et al. 2013). This is particularly relevant as May is currently widely used to initialize operational summer seasonal forecasts. It is also important to note that the predictability properties of sea ice cover may be lower in a warmer climate with less sea ice (Holland et al. 2010).

REFERENCES

Blanchard-Wrigglesworth, E., K. C. Armour, C. M. Bitz, and E. DeWeaver, 2011a: Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations. *J. Climate,* **24,** 231–250, doi:10.1175/2010JCLI3775.1.

——, C. M. Bitz, and M. M. Holland, 2011b: Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophys. Res. Lett.,* **38,** L18503, doi:10.1029/2011GL048807.

Boer, G., 2004: Long time-scale potential predictability in an ensemble of coupled climate models. *Climate Dyn.,* **23,** 29–44, doi:10.1007/s00382-004-0419-8.

Chevallier, M., D. Salas y Mélia, A. Voldoire, M. Déqué, and G. Garric, 2013: Seasonal forecasts of the pan-Arctic sea ice extent using a GCM-based seasonal prediction system. *J. Climate,* **26,** 6092–6104, doi:10.1175/JCLI-D-12-00612.1.

Collins, M., 2002: Climate predictability on interannual to decadal time scales: The initial value problem. *Climate Dyn.,* **19,** 671–692, doi:10.1007/s00382-002-0254-8.

Day, J. J., J. C. Hargreaves, J. D. Annan, and A. Abe-Ouchi, 2012: Sources of multi-decadal variability in Arctic sea ice extent. *Environ. Res. Lett.,* **7,** 034011, doi:10.1088/1748-9326/7/3/034011.

Divine, D. V., and C. Dick, 2006: Historical variability of sea ice edge position in the Nordic Seas. *J. Geophys. Res.,* **111,** C01001, doi:10.1029/2004JC002851.

Eicken, H., 2013: Ocean science: Arctic sea ice needs better forecasts. *Nature,* **497,** 431–433, doi:10.1038/497431a.

Griffies, S. M., and Coauthors, 2011: The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations. *J. Climate,* **24,** 3520–3544, doi:10.1175/2011JCLI3964.1.

Hawkins, E., J. Robson, R. Sutton, D. Smith, and N. Keenlyside, 2011: Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach. *Climate Dyn.,* **37,** 2495–2509, doi:10.1007/s00382-011-1023-3.

Hazeleger, W., and Coauthors, 2012: EC-Earth v2.2: Description and validation of a new seamless Earth system prediction model. *Climate Dyn.,* **39,** 2611–2629, doi:10.1007/s00382-011-1228-5.

Holland, M. M., D. A. Bailey, and S. Vavrus, 2010: Inherent sea ice predictability in the rapidly changing arctic environment of the Community Climate System Model, version 3. *Climate Dyn.,* **36,** 1239–1253, doi:10.1007/s00382-010-0792-4.

Hunke, E. C., and W. H. Lipscomb, 2004: CICE: The Los Alamos Sea Ice Model, documentation and software, version 3.1. Los Alamos National Laboratory Rep. LA-CC-98-16, 56 pp.

Johns, T. C., and Coauthors, 2006: The new Hadley Centre climate model (HadGEM1): Evaluation of coupled simulations. *J. Climate,* **19,** 1327–1353, doi:10.1175/JCLI3712.1

Jungclaus, J. H., and Coauthors, 2013: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.,* **5,** 422–446, doi:10.1002/jame.20023.

Khon, V. C., I. I. Mokhov, M. Latif, V. A. Semenov, and W. Park, 2009: Perspectives of northern sea route and northwest passage in the twenty-first century. *Climatic Change,* **100** (3–4), 757–768, doi:10.1007/s10584-009-9683-2.

Koenigk, T., and U. Mikolajewicz, 2008: Seasonal to interannual climate predictability in mid and high northern latitudes in a global coupled model. *Climate Dyn.,* **32,** 783–798, doi:10.1007/s00382-008-0419-1.

Latif, M., M. Collins, H. Pohlmann, and N. Keenlyside, 2006: A review of predictability studies of Atlantic sector climate on decadal time scales. *J. Climate,* **19,** 5971–5987, doi:10.1175/JCLI3945.1.

Mahajan, S., R. Zhang, and T. L. Delworth, 2011: Impact of the Atlantic meridional overturning circulation (AMOC) on Arctic surface air temperature and sea ice variability. *J. Climate,* **24,** 6573–6581, doi:10.1175/2011JCLI4002.1.

Matei, D., J. Baehr, J. H. Jungclaus, H. Haak, W. A. Mller, and J. Marotzke, 2012: Multiyear prediction of monthly mean Atlantic meridional overturning circulation at 26.5°N. *Science,* **335,** 76–79, doi:10.1126/science.1210299.

McLaren, A. J., and Coauthors, 2006: Evaluation of the sea ice simulation in a new coupled atmosphere-ocean climate model (HadGEM1). *J. Geophys. Res.,* **111,** C12014, doi:10.1029/2005JC003033.

Meehl, G. A., and Coauthors, 2009: Decadal prediction: Can it be skillful? *Bull. Amer. Meteor. Soc.,* **90,** 1467–1485, doi:10.1175/2009BAMS2778.1.

Meier, W. N., J. Stroeve, A. Barrett, and F. Fetterer, 2012: A simple approach to providing a more consistent arctic sea ice extent time series from the 1950s to present. *Cryosphere,* **6,** 1359–1368, doi:10.5194/tc-6-1359-2012.

Pohlmann, H., M. Botzet, M. Latif, A. Roesch, M. Wild, and P. Tschuck, 2004: Estimating the decadal predictability of a coupled AOGCM. *J. Climate,* **17,** 4463–4472, doi:10.1175/3209.1.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.,* **108,** 4407, doi:10.1029/2002JD002670.

Shaffrey, L. C., and Coauthors, 2009: U.K. HiGEM: The new U.K. high-resolution global environment model—Model description and basic evaluation. *J. Climate,* **22,** 1861–1896, doi:10.1175/2008JCLI2508.1.

Sigmond, M., J. C. Fyfe, G. M. Flato, V. V. Kharin, and W. J. Merryfield, 2013: Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system. *Geophys. Res. Lett.,* **40,** 529–534, doi:10.1002/grl.50129.

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science,* **317,** 796–799, doi:10.1126/science.1139540.

Stephenson, S. R., L. C. Smith, and J. A. Agnew, 2011: Divergent long-term trajectories of human access to the Arctic. *Nat. Climate Change,* **1,** 156–160, doi:10.1038/nclimate1120.

——, ——, L. W. Brigham, and J. A. Agnew, 2013: Projected 21st-century changes to arctic marine access. *Climatic Change,* **118,** 885–899, doi:10.1007/s10584-012-0685-0.

Tietsche, S., D. Notz, J. H. Jungclaus, and J. Marotzke, 2013: Predictability of large interannual Arctic sea-ice anomalies. *Climate Dyn.,* **41,** 2511–2526, doi:10.1007/s00382-013-1698-8.

——, and Coauthors, 2014: Seasonal to interannual Arctic sea-ice predictability in current GCMs. *Geophys. Res. Lett.,* **41,** 1035–1043, doi:10.1002/2013GL058755.

Wang, W., M. Chen, and A. Kumar, 2013: Seasonal prediction of Arctic sea ice extent from a coupled dynamical forecast system. *Mon. Wea. Rev.,* **141,** 1375–1394, doi:10.1175/MWR-D-12-00057.1.

Zhang, J., and D. A. Rothrock, 2003: Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Mon. Wea. Rev.,* **131,** 845–861, doi:10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2.