Seasonal Forecasts of the Pan-Arctic Sea Ice Extent Using a GCM-Based Seasonal Prediction System

MATTHIEU CHEVALLIER, DAVID SALAS Y MÉLIA, AURORE VOLDOIRE, AND MICHEL DÉQUÉ

Centre National de Recherches Météorologiques/Groupe d'Etude de l'Atmosphère Météorologique, Météo-France, CNRS, Toulouse, France

GILLES GARRIC

Mercator-Océan, Ramonville Saint-Agne, France

(Manuscript received 8 August 2012, in final form 14 February 2013)

ABSTRACT

An ocean–sea ice model reconstruction spanning the period 1990–2009 is used to initialize ensemble seasonal forecasts with the Centre National de Recherches Météorologiques Coupled Global Climate Model version 5.1 (CNRM-CM5.1) coupled atmosphere–ocean general circulation model. The aim of this study is to assess the skill of fully initialized September and March pan-Arctic sea ice forecasts in terms of climatology and interannual anomalies. The predictions are initialized using "full field initialization" of each component of the system. In spite of a drift due to radiative biases in the coupled model during the melt season, the full initialization of the sea ice cover on 1 May leads to skillful forecasts of the September sea ice extent (SIE) anomalies. The skill of the prediction is also significantly high when considering anomalies of the SIE relative to the long-term linear trend. It confirms that the anomaly of spring sea ice cover in itself plays a role in preconditioning a September SIE anomaly. The skill of predictions for March SIE initialized on 1 November is also encouraging, and it can be partly attributed to persistent features of the fall sea ice cover. The present study gives insight into the current ability of state-of-the-art coupled climate systems to perform operational seasonal forecasts of the Arctic sea ice cover up to 5 months in advance.

1. Introduction

Within the last few years, the shrinking summer Arctic sea ice cover has awakened interest in obtaining seasonal outlooks of the sea ice cover. Such outlooks are intended to give valuable information, for example, on marine accessibility of maritime routes or on the duration of the ice-free season in the marginal ice zones. Only a few institutions produce sea ice predictions using a coupled atmosphere–ocean general circulation model (AOGCM), although such models are becoming the standard in operational seasonal forecasts of air temperature and precipitation [e.g., the European Seasonal to Interannual Prediction Project, version 4 (EUROSIP4) system (http://www.ecmwf.int/products/catalogue/VIII.html) or the National Centers for Environmental Prediction Climate Forecast System, version 4].

Previous studies have explored various aspects of Arctic sea ice predictability at seasonal time scales, mainly focusing on predictions of the minimum September sea ice extent (SIE). Some providers of sea ice predictions currently use statistical techniques based on observations or model reconstructions. Lindsay et al. (2008) found that over the last few decades most of the predictive potential of such techniques lay in the long-term downward trend of September sea ice extent. Others used ocean–sea ice models driven by historical atmospheric forcing to produce ensemble predictions (e.g., Zhang et al. 2008). Contributors of the Study of Environmental Arctic Change (SEARCH)/Sea Ice Outlook project (http://www.arcus.org/search/seaiceoutlook/index.php) pointed out the role of long-term preconditioning of the sea ice cover: a thinner ice cover is more vulnerable to melt during the summer, resulting in less extensive September sea ice cover. Nevertheless, sea ice predictability is also affected by

Denotes Open Access content.

Corresponding author address: Matthieu Chevallier, CNRM-GAME/GMGEC/ASTER, Météo-France, 42, avenue Gaspard Coriolis, 31057 Toulouse CEDEX, France.
E-mail: matthieu.chevallier@meteo.fr

DOI: 10.1175/JCLI-D-12-00612.1

its interaction with other climate components: Kauker et al. (2009) showed that 66% of the September 2007 anomaly was already determined in late June from the winter sea ice thickness (SIT) and May–June wind stresses. Anomalous winds can advect SIT anomalies in regions where sea ice will be more prone to melting, leading to further loss than predictable using initial SIT anomaly alone.

Successful sea ice seasonal predictions not only rely on a comprehensive initialization of relevant sea ice variables, they also depend on the ability of the forecast system to capture high-frequency climate variability and associated feedbacks, even if the chaotic nature of the atmosphere strongly reduces the atmospheric-driven predictability. Thus, fully coupled AOGCMs seem more suitable to investigate Arctic sea ice predictability. Moreover, owing to changing relationships between September SIE and some sea ice predictors typically used in statistical forecast techniques, Holland and Stroeve (2011) claim that sea ice forecast systems need to include a coupling to the atmosphere.

Blanchard-Wrigglesworth et al. (2011a) used coupled AOGCM simulations and observations to explore the inherent predictability of the Arctic sea ice area (SIA). They show that the time scale of the Arctic SIA persistence is 2–4 months, although a reemergence of memory may occur at certain seasons because of coupled processes. In simulations with another AOGCM, Chevallier and Salas y Mélia (2012, hereafter CSM2012) highlighted the role of the ice thickness distribution (ITD) and found that the September SIA is potentially predictable up to 6 months in advance, given the area covered by thick ice (>1.5 m) is known in March. They named this a ''memory regime'' since only the thickest ice present in May survives at the end of the melt season. They also highlighted a ''persistence regime'' for winter predictability: persistence of a fall SIA anomaly is the best way to potentially predict March SIA, but only up to 3 months in advance.

In the present paper, we assess the skill of historical seasonal predictions of the pan-Arctic SIE performed with the Centre National de Recherches Météorologiques Coupled Global Climate Model version 5.1 (CNRM-CM5.1) coupled AOGCM over the period 1990–2008. In section 2, we describe the model, the initialization techniques, as well as the forecast protocol. Seasonal predictions of September and March pan-Arctic SIE are then evaluated in section 3. A discussion follows in section 4, and section 5 draws some conclusions.

2. Model description and experiments setup

a. Models

Simulations are performed with the CNRM-CM5.1 coupled atmosphere–ocean general circulation model. CNRM-CM has been developed jointly by the Centre National de Recherches Météorologiques-Groupe d'Etudes de l'Atmosphère Météorologique (CNRM-GAME) and the Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS) so as to contribute to phase 5 of the Coupled Model Intercomparison Project (CMIP5). A full description and a basic evaluation of the system can be found in Voldoire et al. (2013).

CNRM-CM5.1 includes the global spectral atmospheric model Action de Recherche Petite Echelle Grande Echelle, version 5.2 (ARPEGE-Climat v5.2), operated on a T127 spectral truncation (roughly 1.4° resolution in both longitude and latitude). Land surface processes and air–sea turbulent exchanges are simulated through the Surface Externalisée (SURFEX) platform. The ocean component of CNRM-CM5.1 is based on the ocean part of the Nucleus for European Modeling of the Ocean (NEMO) model. The OGCM ORCA 1° global tripolar quasi-isotropic grid is used: its nominal horizontal resolution is 1°, with a latitudinal refinement of 0.5° in the Arctic Ocean. A total of 42 vertical levels are used (10 in the uppermost 100 m). The Global Experimental Leads and Sea Ice Model for Atmosphere and Ocean (GELATO5) dynamic–thermodynamic sea ice model (Voldoire et al. 2013) is directly embedded in the ocean component and uses the same horizontal grid. It includes elastic–viscous–plastic rheology, semi-Lagrangian advection of ice slabs, ridging and rafting, parameterization of lead processes, snow–ice formation, prognostic salinity, and an advanced snow cover scheme that represents the effect of snow ageing on snow density. In the present study, eight thickness categories are used: 0–0.2, 0.2–0.5, 0.5–0.9, 0.9–1.5, 1.5–2.5, 2.5–4, 4–6, and over 6 m. For the treatment of ice vertical heat diffusion, every slab of ice is divided into nine vertical layers and may be covered with one snow layer.

In the coupled mode, all components are coupled through the Ocean Atmosphere Sea Ice and Soil, version 3 (OASISv3) coupler. The ocean–sea ice component of CNRM-CM5.1 (denoted NEMO-GLT) can also be used in forced mode with prescribed atmospheric fields. Air–ocean and air–ice turbulent fluxes are computed from meteorological fields using the Large and Yeager (2004) bulk formulas. Incident radiative fluxes and precipitation are also prescribed. In both forced and coupled modes, the parameter values of all model components (including sea ice) are rigorously the same.

b. Experimental setup

1) Ocean–sea ice reconstruction

The NEMO-GLT model is used in forced mode to produce an ocean–sea ice reconstruction over the period

1990–2009. The model is driven at the surface by the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (ERA-Interim) fields (Dee et al. 2011) with corrections applied to the downwelling part of the shortwave (SW) and longwave (LW) radiative fluxes southward from 65°N. The method for corrections consists of applying a monthly local climatological coefficient to the daily ERA-Interim fluxes. The coefficient is calculated by the ratio between the monthly climatological Global Energy and Water Cycle Experiment (GEWEX) Surface Radiation Budget, version 3.1 (SRB3.1) satellite data and the ERA-Interim fields. A low-pass Shapiro filter is also applied to the climatological maps in order to retain the original synoptic variability. By construction, this correction has no interannual impacts and can be applied throughout the ERA-Interim period. Based on in situ measurements, a recent work done by Lüpkes et al. (2010) highlighted the ERA-Interim surface biases over the Arctic Ocean. Although the ERA-Interim performs best compared to other reanalyses, Jakobson et al. (2012), with different and independent in situ measurements, it showed similar coherent warm and wet biases at the ERA-Interim central Arctic sea ice surface. As suggested by Lüpkes et al. (2010), we then reduced the surface air temperature by 2°C and the humidity by 15% northward of 80°N over the sea ice cover only. There is neither data assimilation nor nudging in the NEMO-GLT ocean–sea ice reconstruction. A sea surface salinity restoring toward the Levitus et al. (1998) climatology is applied to all surface grid cells not covered with sea ice. Before running the ERA-Interim forced simulation, the model has been spun up over the period 1958–89 using Drakkar Forcing Set version 4.2 (DFS4.2) forcing fields (Brodeau et al. 2010), which is a forcing set, based on the 40-yr ECMWF Re-Analysis (ERA-40), designed for global ocean studies.

The model performs very well in simulating the annual cycle, trends, and interannual variability of the Arctic SIE, compared to Arctic SIE obtained from the National Snow and Ice Data Center (NSIDC), Fetterer et al. (2002). The correlation of reconstructed monthly SIE anomalies with NSIDC SIE anomalies over the period reaches 0.87 (detrended SIE: 0.81). The interannual variability of sea ice volume (SIV) anomalies compares well with estimates from the Panarctic Ice-Ocean Model and Assimilation System (PIOMAS) (e.g., Zhang et al. 2008; Schweiger et al. 2011), which stands as a reference in Arctic SIV reconstruction (correlation of monthly SIV anomalies reconstruction/PIOMAS: 0.85; detrended SIV: 0.80).

September sea ice cover compares well with observations (Fig. 1a), in spite of a lack of sea ice in the Beaufort Sea and along the Siberian coasts. March sea ice concentration (SIC) is generally well simulated (Fig. 2a), especially in the Barents Sea, the Bering Sea, and the Sea of Okhotsk. In the Greenland Sea the ice edge lies too far off the southeastern coast of Greenland. This eastward spread anomaly is also present in other NEMO configurations at higher resolution (Drakkar Group, 2012, personal communication) sharing the same ERA-Interim forcing. This anomaly may be due to biases in both the oceanic circulation and the atmospheric forcing. SIT fields generally exhibit realistic structures (Figs. 1d, 2d), although a comparison with draft observations from the sea ice thickness climate data record (Lindsay 2010) suggests an overall underestimation of SIT (bias: −0.23 m; RMSE: 0.5 m). March mean SIT reaches 4 m off Greenland and in the Canadian Arctic Archipelago (CAA), which is almost 1.5 m less than observational data from the Ice, Cloud and Land Elevation Satellite (ICESat) (e.g., Kwok and Cunningham 2008). These negative SIT biases results in an underestimation of the ice volume flux through Fram Strait by 20% compared to observations. A more in-depth validation of the NEMO-GLT reconstruction can be found in Chevallier (2012).

2) COUPLED FORECASTS

For the ensemble experiments, the coupled model CNRM-CM5.1 was used in the same configuration as in CMIP5 climate simulations (Voldoire et al. 2013). The prediction start dates are 1 May and 1 November: these start dates are standard for summer and winter atmosphere seasonal forecasts. From each start date, a nine-member ensemble is run for five months. Atmospheric greenhouse gases and aerosols are prescribed as in CMIP5 historical simulations, but no other data are used to drive the coupled model during the forecasts.

The experiments initialized on 1 May are investigated as tentative seasonal forecasts of the September annual sea ice minimum. As the Arctic SIV reaches its maximum in April, 1 May sea ice cover must well reflect the winter preconditioning. It has also been shown that the April–May sea ice thickness distribution contributes to September sea ice potential predictability (e.g., CSM2012). Experiments initialized on 1 November are tentative seasonal forecasts of the March annual maximum SIE. It is a test of the ability of the coupled model to realistically shape the sea ice edge in the marginal ice zone (MIZ), where the influence of the ocean is strong (e.g., Bitz et al. 2005).

For each start date, ocean and sea ice are initialized according to the ocean–sea ice state obtained from the NEMO-GLT forced reconstruction at the same date. Using the ocean–sea ice reconstruction allows us to fully

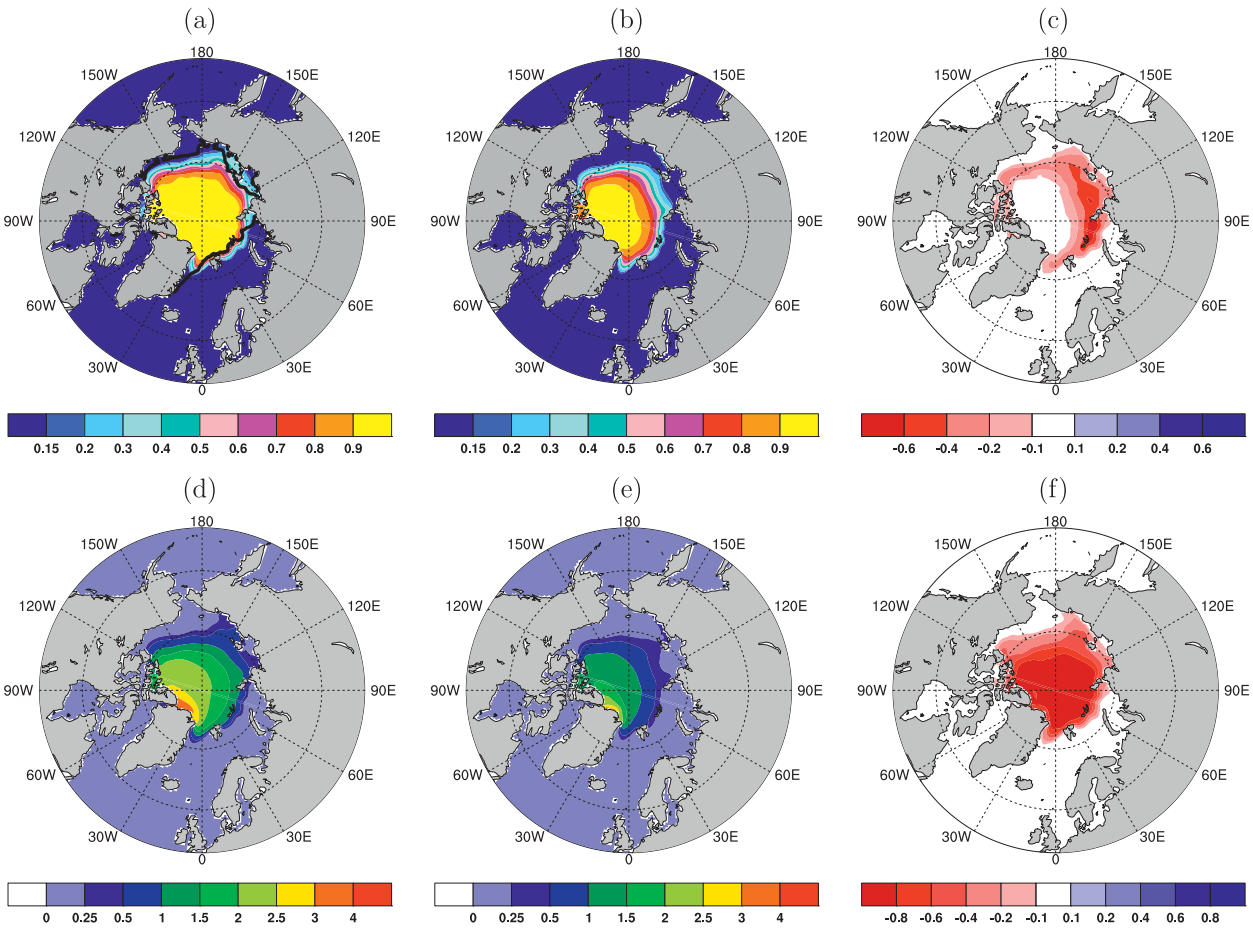

FIG. 1. Sea ice climatology in September 1990–2008. Sea ice concentration: (a) NEMO-GLT reconstruction (black line, 1990–2008 NSIDC sea ice edge), (b) forecast, and (c) difference (b) minus (a). Sea ice thickness (m): (d) reconstruction, (e) forecast, and (f) difference (e) minus (d).

initialize sea ice dynamically and thermodynamically. In particular, each of the eight ice categories is thermodynamically initialized, as the reconstruction was built with the same ocean–sea ice model. The atmospheric initial states are taken from the ERA-Interim at the given initialization date and eight days around this date, in order to generate nine-member ensembles (''atmosphere lag average'').

3. Ensemble predictions

a. Assessment method

Our predictions with the coupled model are assessed as ''deterministic'' forecasts: for example, we primarily look at the nine-member ensemble means as the forecasts. Monthly outputs are considered. We focus on the predicted Arctic SIE rather than Arctic SIA since extent is thought to be more accurately observed than area (Parkinson and Cavalieri 2008). In the present paper, performance of the system is evaluated quantitatively by calculating the anomaly correlation coefficient (ACC) skill score against observed SIE from NSIDC.

Anomalies are computed as follows. When the overall quality of the system is assessed, ''raw anomalies'' are calculated, that is, anomalies relative to the average. Raw forecast anomalies are anomalies relative to the average of all forecasts of all members over the period 1990–2008 for September forecasts and 1991–2009 for March forecasts (denoted as the ''forecast climatology''). Raw observed anomalies are computed relative to the climatology of NSIDC SIE. When the predictability from initial conditions is assessed, anomalies are calculated relative to the linear long-term trend and denoted as ''detrended anomalies.'' In that case, the ACC skill score is called ''detrended ACC.'' Since 19 values may not be sufficient to compute robust trends, the significance (95% level) of the linear fit has always been checked before calculating detrended anomalies.

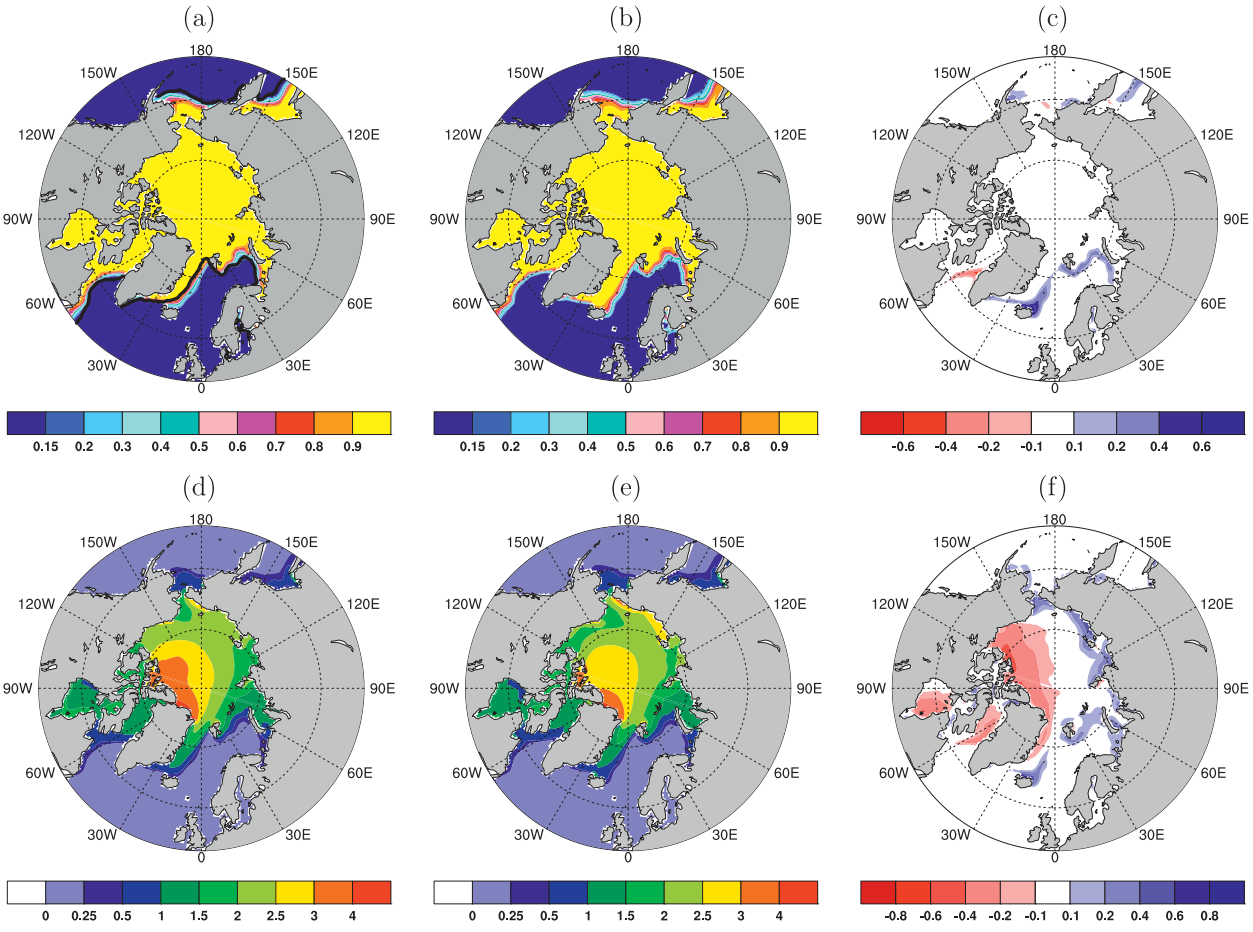

FIG. 2. As in Fig. 1, but for March 1991–2009 sea ice climatology.

Even if forecast and observed trends do not have the same magnitude, ACCs (between anomalies relative to the climatologies) tend to be high. That is why we also investigate detrended ACC. Removing the long-term linear trend is a way to isolate predictability from initial conditions ("first kind") from predictability due to external forcing ("second kind"; e.g., Branstator and Teng 2010), which makes up the predictability associated to increasing greenhouse gases.

b. 1 May initialization

Figure 1b shows the climatology of September SIC as predicted by the system. This climatology is calculated by averaging the predictions of all nine members of all 19 ensemble forecasts initialized on 1 May. We compare this climatology to the climatology of September SIC in the NEMO-GLT reconstruction (Fig. 1a). September forecasts clearly underestimate the SIC in the north of the Kara, Laptev, and East Siberian Seas and to a lesser extent in the CAA. The forecast experiences a drift during the melt season. The drift in SIE (difference between observed and predicted SIE along the forecast period, averaged over all ensembles) is most intense in June and July ($-0.45 \times 10^6\,\mathrm{km^2\,month^{-1}}$) and stabilizes by mid-August. This drift results in a systematic underestimation of the predicted September SIE by nearly $1.4 \times 10^6\,\mathrm{km^2}$. September SIE is also underestimated in the uninitialized historical run performed with CNRM-CM5.1 (Voldoire et al. 2013). Most of this bias is due to an overestimation of the surface downwelling shortwave radiation by $\sim 30\,\mathrm{W\,m^{-2}}$ in summer during the forecasts, which enhances sea ice melting.

We show in Fig. 3a the predicted September SIE raw anomalies and the raw observed September SIE anomalies. In 14 of the 19 ensemble forecasts, the correct raw anomaly sign is predicted. Ensembles have a standard deviation (STD) around the ensemble mean ranging from 180 000 (1997) to $460\,000\,\mathrm{km^2}$ (1995). The ensemble STD seems to depend on neither the predicted nor the initial SIE (nearly zero correlation in both cases). Using the STD as a metric of the spread, we found that,

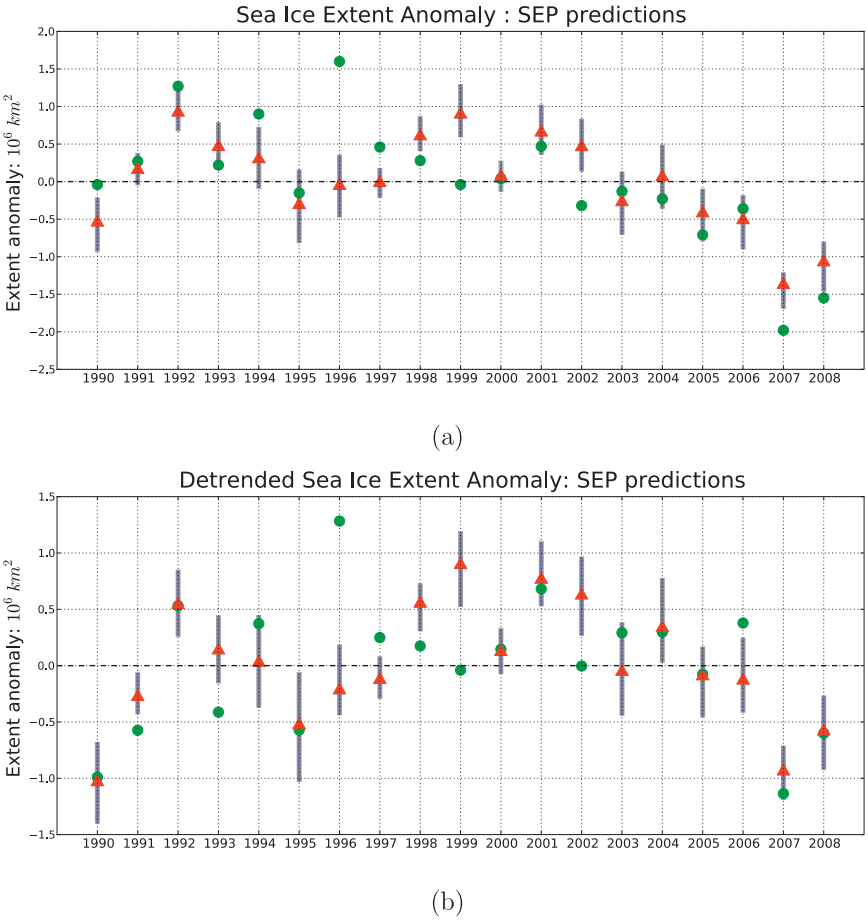

FIG. 3. September SIE forecast anomalies relative to (a) the climatologies and (b) the long-term trends (linear fit): observation anomalies (green dots), ensemble forecast means (red triangles), and ensemble forecast spreads (±1 std dev, gray bars).

in 10 of the 19 ensembles, the observed raw anomaly lies within one STD from the forecasted raw anomaly, which is about half of the years.

In Fig. 3b, detrended anomalies are shown. Here, 11 of the 19 ensemble forecasts give the right anomaly sign, and in 11 years the observed detrended anomaly lies within one STD from the forecasted detrended anomaly.

A few predictions are conspicuous for bad or good reasons. Our system misses the highly positive SIE anomaly of September 1996. This positive anomaly occurred after several months of low ice extent (March 1996 SIE was March record low of the period 1979–2000). To our knowledge, this record high has not been well documented in the literature. Since the initialization alone fails in driving this record high, we speculate that particular weather patterns may have played a significant role in shaping this anomaly. On the contrary, observed lows of September 1995, 2007, and 2008 are correctly captured by the system. The 2007 predictions can also be highlighted since all predicted SIEs in that ensemble lie below 1990–2006 predictions.

The ACC skill scores of the model forecasts, given in Table 1, are significantly high, even when the long-term linear trend is removed. It shows that the skill of prediction lies not only in the long-term trend but also in the state of the May sea ice cover.

To assess the added value of our model forecasts, we compare the model forecasts to forecasts produced with

TABLE 1. The ACC skill of the seasonal predictions of the September SIE over the period 1990–2008: "model forecast" denotes prediction with the coupled model. All ACCs are computed relative to SIE provided by NSIDC (Fetterer et al. 2002). Bold font indicates 95% significant correlations (bootstrap test).

May initialization	Model forecast	Volume regression
SIE (raw)	**0.72** ($p < 0.001$)	**0.54** ($p < 0.001$)
SIE (detrended)	**0.60** ($p = 0.002$)	0.24 ($p = 0.13$)

an empirical technique. Persistence of initial SIE is a challenging technique typically used in atmospheric seasonal forecasting. However, the persistence of April and May SIE anomalies decreases quickly with time in the observations, as SIE varies rapidly during that time of the year (Blanchard-Wrigglesworth et al. 2011a). As a result, September SIE anomalies correlate weakly with previous April and May SIE anomalies. Rather than persistence, we compare model forecast skill scores with those of a linear regression of September SIE anomalies based on previous April monthly SIV anomaly estimates using the NEMO-GLT reconstruction ("SIV regression"). Actually, September SIEs correlate significantly with April and May SIVs in the reconstruction (0.63 and 0.69 correlation, respectively, significant at the 95% level). This linear regression is built in cross-validation mode: for a particular year, the empirical forecast is deduced from a linear fit built with all but this year. The "detrended SIV regression" is the regression built from anomalies relative to the long-term trends.

Table 1 shows that SIV regression is skillful using raw data. This relative high skill scores seems to be mainly due to the long-term trend: skill score of detrended SIV regression falls below the 95% significance threshold. Table 1 also shows that in both cases model forecasts beat SIV regression. This emphasizes that the nonlinear model physics allows for a better forecast of September SIE than a simple linear model.

c. 1 November initialization

The forecast climatology of March SIC is presented in Fig. 2b, with differences relative to the climatology of the NEMO-GLT reconstruction (Fig. 2c). The forecast climatology matches very well the reconstruction in spite of a slight excess of ice in the Nordic and Barents Seas. This excess results in a slight overestimation of the SIE by 230 000 km^2. Drift of SIE forecast is only significant during January. Notably the SIC is underestimated at the ice edge in the Labrador Sea south of Davis Strait. In Hudson Bay as well as in Baffin Bay, the sea ice cover is 30 cm too thin (Fig. 2f). This is due to a late freeze-up onset in both regions. March SIT is clearly underestimated along the northern coasts of Greenland and the CAA. By contrast, it is slightly overestimated along the Siberian coast. Further analyses (not shown) have revealed that both biases have a dynamical origin.

Figure 4a shows that predicted and observed March SIE raw anomalies correlate well over the period. In 13 of the 19 ensemble forecasts, the right anomaly sign is predicted, which is of the same order as in September predictions. Ensembles have STD between 80 000 (2008) and 340 000 km^2 (2006). STDs are almost half the spread of September predictions. In 10 cases, the observed raw anomaly falls within one STD of the ensemble forecasted raw anomaly.

Figure 4b shows that the right anomaly sign relative to the trend is predicted in 15 of the 19 forecasts and that the gray bar overlaps the observed SIE detrended anomaly in 11 cases. We found that using the STD metric both September and March predictions are comparable.

ACC skill scores of the system (Table 2) are of the same order as those of September predictions when using raw SIE. Detrended ACC is slightly lower but still highly significant.

We compare model forecasts ACC skill scores to those of SIE anomaly persistence using the reconstruction. As both October and November SIE and SIV anomalies weakly correlate with March SIE anomalies, it is difficult to design an alternative empirical method to compare our model forecasts with, based only on sea ice quantities. Forecasts based on persistence are typical benchmarks for assessing more complex model forecasts. The ACC for persistence is significantly high when considering raw forecasts and observations. After removing long-term trends, persistence ACC is only statistically significant at the 80% level. In both cases, the ACC of model forecasts is higher than ACC for persistence, which confirms the relevance of model forecasts.

4. Discussion

a. Predictability regimes

High values of ACC skill scores of both September and March seasonal forecasts lead us to investigate what in the initial states drives the model predictability. In CSM2012, predictability of the September sea ice cover was found to be related to the ice thickness distribution as early as in spring, whereas persistence of early winter sea ice conditions seems to be the main source of March sea ice predictability. We investigate the relevance of the concepts defined in CSM2012 for our system forecasts.

For September predictions, high ACC skill scores of both model forecasts and SIV regression forecasts show that the long-term trend of September SIE predictability is directly related to the variability of spring SIV, hence highlighting the role of the preconditioning in shaping the September sea ice cover. Thus, September predictions illustrate the memory regime related to SIV and thickness distribution. As in CSM2012, we calculated the variance explained by anomalies of thick ($h > 1.5$ m) or thin ($h < 1.5$ m) ice-covered areas using the ice thickness distribution. In May the area covered with thick ice explains 50% of September SIE anomalies. By contrast, the area covered with thin ice in May

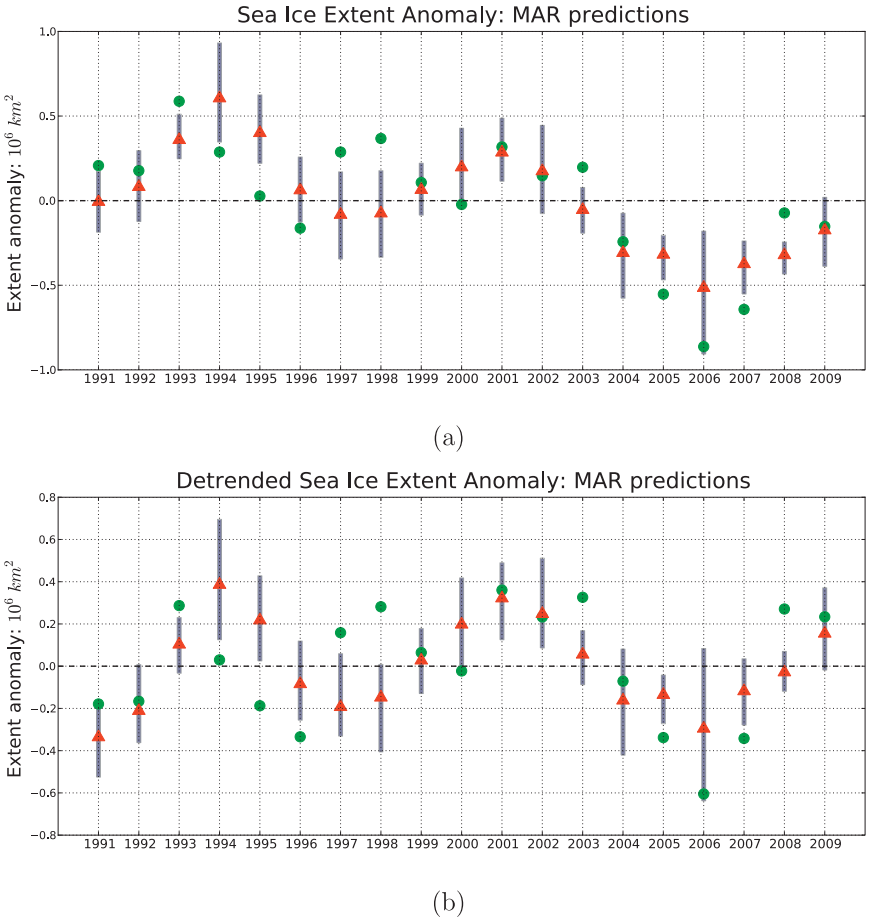

FIG. 4. As in Fig. 3, but for March SIE forecasts.

does not explain any significant fraction of the variance of September SIE anomalies.

Regarding March prediction skills, SIE persistence provides significant skills at the 95% level, which is not the case for volume regression (not shown). The skill of the prediction lies in the ability of the system to keep memory of an existing SIE anomaly on 1 November. Such SIE anomaly may arise from an early or delayed freeze-up onset in the marginal seas. This property is similar to the persistence regime described in CSM2012. Nevertheless, neither ITD-based predictor explains a significant fraction of forecast variability. This is not surprising: in CSM2012, inherent predictability of March sea ice cover originating from ITD only extends 3 months in advance.

Other processes may also be responsible for the year-to-year variability of the predicted ice edge, for example, ocean heat inflows. The ability of the model to correctly capture these ocean inflows is beyond the scope of the present paper. However, we performed a quick analysis of March SIE forecast variance using the reconstructed sea surface temperature (SST) averaged over the North Pacific and North Atlantic Oceans (north of 30°N). We found that October North Pacific and North Atlantic SST correlate well with March SIE in the reconstruction (−0.65 and −0.43, respectively) and that October North Pacific SST explains 22% of the March SIE forecast variability. This simple analysis shows that there may be a fall preconditioning of the upper ocean in the marginal ice zone (here, the Pacific MIZ).

b. *Model drift and predictability*

September forecast biases arise from systematic radiative biases in the atmospheric component of CNRM-CM5.1. These atmospheric biases drive a significant drift of forecasted SIE during June and July.

TABLE 2. As in Table 1, but for seasonal predictions of the March SIE over the period 1991–2009.

Nov initialization	Model forecast	Persistence
SIE (raw)	**0.74** ($p < 0.001$)	**0.51** ($p = 0.01$)
SIE (detrended)	**0.53** ($p = 0.008$)	0.20 ($p = 0.20$)

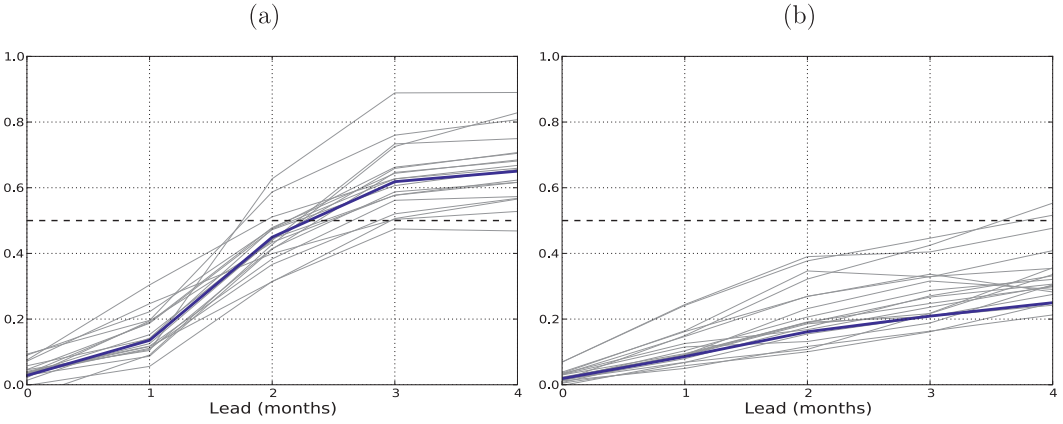

FIG. 5. Metric μ, described in the appendix, calculated for SIT in (a) 1 May forecasts and (b) 1 Nov forecasts: individual forecasts (gray lines) and forecast climatology (blue lines). At every lag, the metric value below 0.5 means that the 2D structure of forecasted SIT is closer to that of SIT in the reconstruction than to that of the "coupled model attractor" (defined as the climatology of SIT in an uninitialized simulation with CNRM-CM5.1).

September predicted SIT exhibits also a sizeable negative bias (Fig. 1f). For March predictions, minor SIT biases occur because of dynamic biases in the atmospheric and ocean circulations. The forecast drift in SIE is small, although present in January.

Is there a loss of predictability due to the coupled model biases? Are the biases so strong that in a few months all Arctic SIT structure is adjusted to resemble that of the uninitialized coupled model (the "attractor")? We defined a metric quantifying the loss of memory in SIT structure during the forecast lead time. This metric is denoted μ and is described in full detail in the appendix. The metric μ estimates the drift of the forecast system from its initial state (based on the reconstruction) toward the attractor of the uninitialized coupled model (here, the 1990–2008 mean state of historical ensemble simulations with CNRM-CM5.1 performed in the framework of CMIP5).

Figure 5a shows that, for 1 May SIT forecasts, μ grows quickly in May and June and exceeds 0.5 in July and August. This means that in the second half of the forecast periods, the SIT structure becomes "closer" to the SIT structure from the attractor. However, the drift of μ stabilizes at the end of the forecast when incoming shortwave starts to decline and the radiative bias decreases.

Conversely, in the case for 1 November initialization (Fig. 5b), μ rarely exceeds 0.5 after 4 months of forecasts. It shows that the SIT structure still reflects that of the reconstruction, even after such a rather long lead time. It suggests that March thickness forecasts initialized on 1 November may have a value, conversely to September thickness forecasts initialized on 1 May, in which too much information is lost due to the spring–summer atmospheric biases.

c. *Initial value versus external predictability*

As an attempt to only consider predictability from initial conditions, we removed a linear trend. The hypothesis here is that this long-term linear trend is associated with the response to external climate forcing. That ACC skill scores remain highly significant for detrended model forecasts shows that there is a predictability in our system directly related to the initial values and that the predictive skill does not only rely on the long-term trend for such a lead time. By contrast, Lindsay et al. (2008) stated that most of the skill of their empirical technique lies in the trend for lead time greater than 3 months. However, detrended lagged correlations between September SIE and simulated ice quantities suggest some initial value predictability at least for September SIE (their Fig. 9). We also addressed initial value predictability by looking at the growth of the cross-ensemble standard deviation, using the root-mean-square deviation (RMSD); see, for example, Blanchard-Wrigglesworth et al. (2011b). This metric helps comparing the growth of the nine-member ensemble spread to the intrinsic spread of the model.

Figure 6 shows the RMSD for SIE and SIV for the 1 May and 1 November initialization, averaged over all 19 ensemble forecasts. As in Blanchard-Wrigglesworth et al., the reference RMSD is built using historical ensemble runs with CNRM-CM5.1, performed in the framework of CMIP5. The RMSD of this uninitialized ensemble represents the limit above which we consider there is no predictability. For both start dates, the RMSD for the forecast SIE and SIV is less than that of the reference, which means that initial value

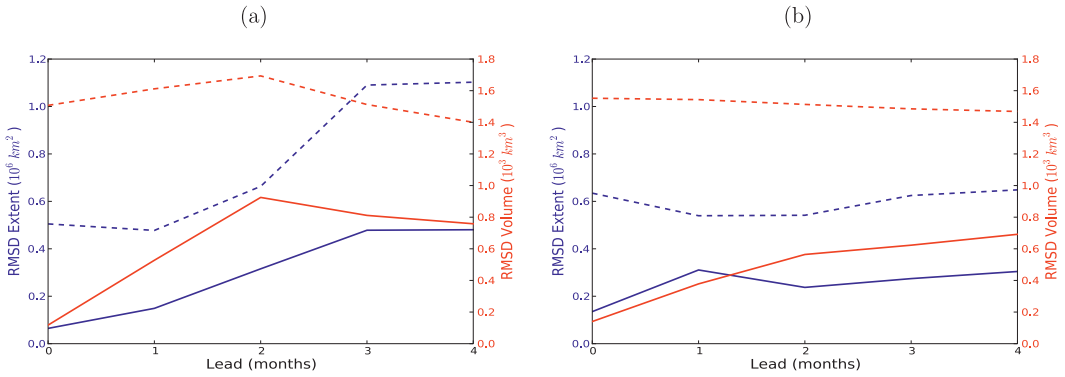

FIG. 6. RMSD of Arctic sea ice extent (blue lines) and volume (red lines) for (a) 1 May and (b) 1 Nov forecasts. Estimates of RMSD from the reference integration (dashed lines) indicate the limit of no predictability.

predictability is substantial over the forecast lead periods. RMSD for SIV in the case of 1 May start dates grows quickly and peaks in July, which is coincident with the drift mentioned above. For all individual start dates (not shown), the ensemble RMSD is substantially less than the reference RMSD.

d. September 2007

The 1 May initialization forecasts show that the record low sea ice cover of September 2007 (relative to the period 1990–2006) was actually predictable as early as May 2007 (Fig. 3a). Figure 7a shows the model reconstruction of September 2007 sea ice cover, which is reasonably close to observations. Colored areas in Fig. 7 show grid cells where at least 10%, 50%, and 100% of the 2007 ensemble predicts SIC above 15%. Compared to the forecast climatology (climatology of the ensemble fractions that predict SIC to be above 15%, Fig. 7b), the September 2007 sea ice prediction corresponds to an anomalous low sea ice cover. We also see that some members capture some aspects of the historical September 2007 sea ice cover: for example, the sea ice bridge between the central Arctic and northern Siberia. The predicted pattern is shifted compared to observations.

Nevertheless, some major features of the September 2007 sea ice cover are missed, especially in the Chukchi–East Siberian Seas where a large open water area appeared, owing to the combination of extreme atmospheric and oceanic forcings. Because of the chaotic nature of the atmosphere it is impossible to predict the whole sequence of weather patterns partly responsible

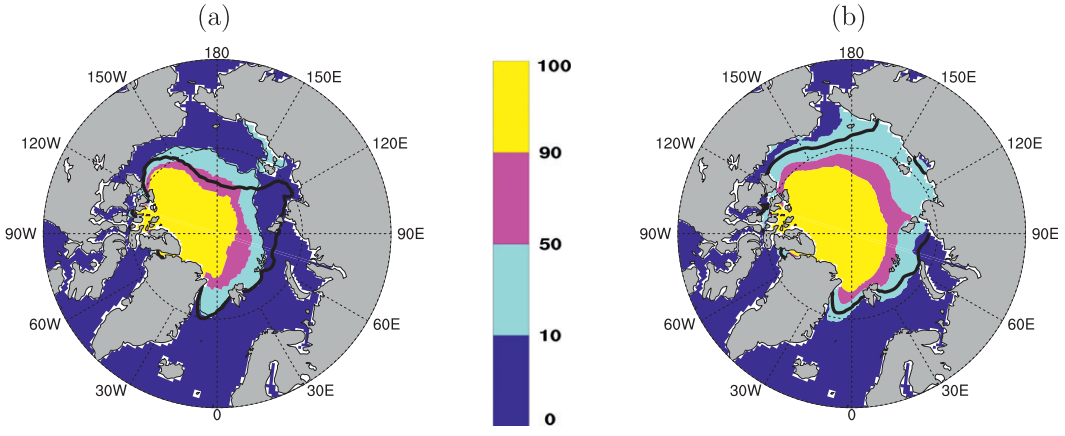

FIG. 7. (a) September 2007 predictions of sea ice presence, compared to (b) the 1990–2008 forecast climatology of sea ice presence: sea ice presence corresponds to SIC above 15%. Colors indicate the fraction of the ensemble (%) that predict sea ice presence in September 2007 in (a) and the climatological fraction (%) averaged over all September predictions over the period 1990–2008 in (b). Black lines denote the simulated ice edge for September 2007 in (a) and the 1990–2008 climatological September ice edge in (b) in the NEMO-GLT reconstruction.

for the 2007 minimum between May and September. Forecasts initialized later during the melt season (1 August; not shown) produce better forecasts as some oceanic forcing processes are taken into account (ocean heat advection through Bering Strait).

5. Conclusions

Hindcast seasonal forecast of the Arctic sea ice cover have been performed with the CNRM-CM5.1 AOGCM. For every year between 1990 and 2008, start dates are 1 May and 1 November for seasonal forecasts of the September and March SIE. Ocean and sea ice are initialized according to their respective states in a forced simulation performed with the ocean–sea ice component of CNRM-CM5.1. Nine-member ensembles are produced by only perturbing the atmosphere.

We addressed the performance of the system as well as insights into patterns of predictability of the Arctic sea ice cover deduced from these hindcasts. In spite of a drift due to radiative biases, skill scores of September SIE anomaly forecasts are significantly high. The skill score is still high after removing the linear fit on September data over the period 1990–2008. This suggests that the predictability of the September sea ice cover is due to the initial conditions. Further analyses show that the initial states of Arctic SIV as well as predictors deduced from the ITD significantly explain the variance of the SIE forecasts. This is an illustration of the ''memory regime'' of predictability described in CSM2012.

CNRM-CM5.1 performs also well in March SIE forecasting. Although the long-term linear trend in SIE seems to play a more important role in these predictions, the ACC skill score is also still highly significant after removing the trends. The initial SIE explains part of the variance of March SIE forecasts, but the analogy with the ''persistence regime'' from CSM2012 can barely be made since this regime is not meant to apply as early as October–November.

Our study confirms the potential for AOGCMs to produce quantitative seasonal forecasts of the Arctic sea ice cover, in spite of the high degree of unpredictability of the Arctic atmospheric circulation. Our analysis shows in both cases that there is a predictability from initial conditions and that the predictive skill does not rely only on the long-term trend (conversely to, e.g., Lindsay et al. 2008). As our initialization run does not involve any data assimilation (that is why we do not call it a reanalysis), one may consider the skills of our forecasts surprising. This is all the more the case as the ACC is calculated with respect to observed data, which are not assimilated in our reconstruction. Using a simple assimilation technique (SST nudging) improves greatly the reconstruction, for example, by reducing the winter SIC bias in the Greenland Sea (not shown). We may wonder if using this alternative reconstruction to initialize would improve the skill of seasonal forecasts with CNRM-CM5.1. The skill of our predictions, however, shows that a reconstruction only forced by atmospheric forcing can serve as a source of predictability for the pan-Arctic SIE.

In section 3, we show that for 50%–60% of September and March forecasts the observed SIE anomalies fall within the forecast spread, measured as the ensemble STD. This metric may suggest that the forecast spread is generally insufficient. We choose to build nine-member ensembles using an ''atmosphere lag average'' technique to remain consistent with the standard of operational seasonal forecasts at Météo-France. Larger ensembles must certainly be considered. However, we think that a nine-member ensemble may be enough to capture some valuable information on the forecast uncertainties. Future work would be needed to assess critical ensemble sizes for sea ice prediction. The method used to generate the ensembles must also be improved in order to increase the spread of the forecasts. An immediate way forward would be to apply perturbations to the ocean or sea ice initial states. This may help the model so as to explore a larger spectrum of trajectories and certainly improve the forecasts at some particular dates.

Winter forecasts cannot be assessed in a comprehensive way by addressing only pan-Arctic integrated quantities. Past studies have documented the winter dipolelike seesaw variability of the SIC in the MIZ (e.g., Deser and Teng 2008). Additionally, predictability in different MIZs may arise for very different reasons (persistence, ocean preconditioning). A more systematic study of forecast skills over subbasins is beyond the scope of the present work and will be addressed in a forthcoming paper. A regional investigation may also be of interest for summer forecasts: for example, to get more information on the date of clearance of maritime routes in the Canadian Arctic Archipelago.

Finally, we are aware that the lead time of these forecasts is quite long. Operational centers would probably initialize their forecasts later in time (e.g., September outlooks based on June or July data), knowing that persistence provides enhanced potential predictability at these shorter time scales (e.g., Blanchard-Wrigglesworth et al. 2011a). As expected, skill scores of September forecasts initialized on 1 August (not shown) are very high and beat those of all empirical predictions. Our results, however, show that skillful forecasts are possible up to five months in

advance. Forecasts at such a term may be beneficial for many planning needs.

Acknowledgments. This work formed part of a Ph.D. thesis at the Centre National de Recherches Météorologiques, Toulouse (France), funded by Ecole des Ponts ParisTech, Marne-la-Vallée (France). Financial travel support granted by Total R&D is acknowledged. Computing support was provided by the Météo-France/DSI supercomputing center. The authors wish to thank Agathe Germe and Aurélien Ribes for insightful discussions. The authors finally offer sincere thanks to Cecilia Bitz, Francisco Doblas-Reyes, two anonymous reviewers, and editor Anand Gnanadesikan for constructive comments that led to improvements in the manuscript.

APPENDIX

Metric Quantifying Forecast Drift for 2D Spatial Fields

This appendix contains details on the metric used in section 4b to quantify the drift of the prediction toward the attractor of the coupled AOGCM. For a 2D variable X (SIC, SIT), we defined a metric quantifying the forecast drift over time. Given a start date D and a lead time T, the metric is given by

$$\mu(D,T)=\frac{\langle F(D,T)-H(D,T),C(T)-H(D,T)\rangle}{\|C(T)-H(D,T)\|^2},$$

where $\langle A,B\rangle \equiv \int_R A(x)B(x)\,dx$ and $\|A\|^2 \equiv \langle A,A\rangle$. The integration is performed over the region R in which A and B are simultaneously nonzero; dx denotes the area of the grid cell x.

In the above equation $F(D, T)$ is the forecast of X at lead time T; $H(D, T)$ is the value of X in the reconstruction (the ''reference'') at the date of the forecast; and $C(T)$ is the climatology of the variable X at the corresponding month in CMIP5 historical simulations with CNRM-CM5.1 (independent of the start date). By construction, $F(D, 0) = H(D, 0)$; thus, $\mu(D, 0) = 0$.

Using a geometrical analogy, this metric gives a measure of the distance between the forecast and the reference at each time, scaled by the distance between the coupled model climatology and the reference: $\mu(D, T) < 0.5$ means that the forecast field is ''closer'' to the reference than to $C(T)$. As the forecast initial state is taken from the reference, it means that the system still remembers its initial conditions at T.

Here $C(T)$ could be seen as the attractor of the coupled system. If all initial conditions are lost within T, the final state is close to the climatology of the uninitialized coupled model, resulting in $\mu(D, T) \approx 1$.

REFERENCES

Bitz, C. M., M. M. Holland, E. C. Hunke, and R. E. Moritz, 2005: Maintenance of the sea-ice edge. *J. Climate,* **18,** 2903–2921.

Blanchard-Wrigglesworth, E., K. 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.

——, 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.

Branstator, G., and H. Teng, 2010: Two limits of initial-value decadal predictability in a CGCM. *J. Climate,* **23,** 6292–6311.

Brodeau, L., B. Barnier, A.-M. Treguier, T. Penduff, and S. Gulev, 2010: An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modell.,* **31,** 88–104, doi:10.1016/j.ocemod.2009.10.005.

Chevallier, M., 2012: On the seasonal predictability of the Arctic sea ice. Ph.D. thesis, Université Paris-Est, 202 pp.

——, and D. Salas y Mélia, 2012: The role of sea ice thickness distribution in the Arctic sea ice potential predictability: A diagnostic approach with a coupled GCM. *J. Climate,* **25,** 3025–3038.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.,* **137,** 553–597.

Deser, C., and H. Teng, 2008: Recent trends in Arctic sea ice and the evolving role of atmospheric circulation forcing, 1979-2007. *Arctic Sea Ice Decline: Observations, Projections, Mechanisms, and Implications, Geophys. Monogr.,* Vol. 180, Amer. Geophys. Union, 133–150.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie, 2002: Sea ice index. National Snow and Ice Data Center, Boulder, CO, digital media. [Available online at http://nsidc.org/data/g02135.html.]

Holland, M. M., and J. Stroeve, 2011: Changing seasonal sea ice predictor relationships in a changing Arctic climate. *Geophys. Res. Lett.,* **38,** L18501, doi:10.1029/2011GL049303.

Jakobson, E., T. Vihma, T. Palo, L. Jakobson, H. Keernik, and J. Jaagus, 2012: Validation of atmospheric reanalyses over the central Arctic Ocean. *Geophys. Res. Lett.,* **39,** L10802, doi:10.1029/2012GL051591.

Kauker, F., T. Kaminski, M. Karcher, R. Giering, R. Gerdes, and M. Voßbeck, 2009: Adjoint analysis of the 2007 all time Arctic sea-ice minimum. *Geophys. Res. Lett.,* **36,** L03707, doi:10.1029/2008GL036323.

Kwok, R., and G. F. Cunningham, 2008: ICESat over Arctic sea ice: Estimation of snow depth and ice thickness. *J. Geophys. Res.,* **113,** C08010, doi:10.1029/2008JC004753.

Large, W. G., and S. G. Yeager, 2004: Diurnal to decadal global forcing for ocean and sea-ice models. NCAR Tech. Note NCAR/TN-460+STR, 22 pp.

Levitus, S., and Coauthors, 1998: *Introduction.* Vol. 1, *World Ocean Atlas 1998,* NOAA Atlas NESDIS 18, 346 pp.

Lindsay, R., 2010: New unified sea ice thickness climate data record. *Eos, Trans. Amer. Geophys. Union,* **91,** 405–406.

——, J. Zhang, A. J. Schweiger, and M. A. Steele, 2008: Seasonal predictions of ice extent in the Arctic Ocean. *J. Geophys. Res.,* **113,** C02023, doi:10.1029/2007JC004259.

Lüpkes, C., T. Vihma, E. Jakobson, G. Koenig-Langlo, and A. Tetzlaff, 2010: Meteorological observations from ship cruises during summer to the central Arctic: A comparison with reanalysis data. *Geophys. Res. Lett.,* **37,** L09810, doi:10.1029/2010GL042724.

Parkinson, C. L., and D. J. Cavalieri, 2008: Arctic sea ice variability and trends, 1979–2006. *J. Geophys. Res.,* **113,** C07003, doi:10.1029/2007JC004558.

Schweiger, A., R. Lindsay, J. Zhang, M. Steele, H. Stern, and R. Kwok, 2011: Uncertainty in modeled Arctic sea ice volume. *J. Geophys. Res.,* **116,** C00D06, doi:10.1029/2011JC007084.

Voldoire, A., and Coauthors, 2013: The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dyn.,* doi:10.1007/s00382-011-1259-y, in press.

Zhang, J., M. Steele, R. Lindsay, A. Schweiger, and J. Morison, 2008: Ensemble 1-year predictions of Arctic sea ice for the spring and summer of 2008. *Geophys. Res. Lett.,* **35,** L08502, doi:10.1029/2008GL033244.