

# *Seasonal to interannual Arctic sea-ice predictability in current GCMs*

Article

Published Version

Creative Commons: Attribution 3.0 (CC-BY)

Tietsche, S., Day, J.J., Guemas, V., Hurlin, W. J., Keeley, S.P.E., Matei, D., Msadek, R., Collins, M. and Hawkins, E.  
ORCID: <https://orcid.org/0000-0001-9477-3677> (2014)  
Seasonal to interannual Arctic sea-ice predictability in current GCMs. *Geophysical Research Letters*, 41 (3). pp. 1035-1043. ISSN 0094-8276 doi: 10.1002/2013GL058755 Available at <https://reading-pure-test.eprints-hosting.org/35894/>

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

Published version at: <http://onlinelibrary.wiley.com/doi/10.1002/2013GL058755/abstract>

To link to this article DOI: <http://dx.doi.org/10.1002/2013GL058755>

Publisher: American Geophysical Union

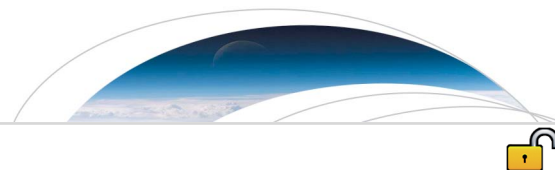
All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online



## RESEARCH LETTER

10.1002/2013GL058755

## Key Points:

- Arctic sea ice is potentially predictable for up to 3 years in current GCMs
- Potential prediction errors are amplified at the coasts of the Arctic ocean
- Advection processes are very important for spatial error patterns

## Supporting Information:

- Readme
- Text S1
- Text S2
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6
- Figure S7
- Figure S8
- Figure S9

## Correspondence to:

S. Tietsche,  
s.tietsche@reading.ac.uk

## Citation:

Tietsche, S., J. J. Day, V. Guemas, W. J. Hurlin, S. P. E. Keeley, D. Matei, R. Msadek, M. Collins, and E. Hawkins (2014), Seasonal to interannual Arctic sea ice predictability in current global climate models, *Geophys. Res. Lett.*, *41*, 1035–1043, doi:10.1002/2013GL058755.

Received 19 NOV 2013

Accepted 20 JAN 2014

Accepted article online 23 JAN 2014

Published online 15 FEB 2014

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Seasonal to interannual Arctic sea ice predictability in current global climate models

S. Tietsche<sup>1</sup>, J. J. Day<sup>1</sup>, V. Guemas<sup>2,3</sup>, W. J. Hurlin<sup>4</sup>, S. P. E. Keeley<sup>5</sup>, D. Matei<sup>6</sup>, R. Msadek<sup>4</sup>, M. Collins<sup>7</sup>, and E. Hawkins<sup>1</sup>

<sup>1</sup>NCAS-Climate, Department of Meteorology, University of Reading, Reading, UK, <sup>2</sup>Institut Català de Ciències del Clima, Barcelona, Spain, <sup>3</sup>CNRM/GAME, Toulouse, France, <sup>4</sup>Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA, <sup>5</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK, <sup>6</sup>Max Planck Institute for Meteorology, Hamburg, Germany, <sup>7</sup>College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

**Abstract** We establish the first intermodel comparison of seasonal to interannual predictability of present-day Arctic climate by performing coordinated sets of idealized ensemble predictions with four state-of-the-art global climate models. For Arctic sea ice extent and volume, there is potential predictive skill for lead times of up to 3 years, and potential prediction errors have similar growth rates and magnitudes across the models. Spatial patterns of potential prediction errors differ substantially between the models, but some features are robust. Sea ice concentration errors are largest in the marginal ice zone, and in winter they are almost zero away from the ice edge. Sea ice thickness errors are amplified along the coasts of the Arctic Ocean, an effect that is dominated by sea ice advection. These results give an upper bound on the ability of current global climate models to predict important aspects of Arctic climate.

### 1. Introduction

The Arctic environment has been experiencing rapid changes in the last few decades, despite the recent slowdown in increasing global surface temperatures. However, natural variability in the Arctic is potentially strong enough to mask or enhance these changes both on multiannual to decadal time scales [Kay *et al.*, 2011; Day *et al.*, 2012] and on seasonal to interannual time scales. Reliable initial-value predictions of Arctic sea ice on these shorter time scales are important for stakeholders in the Arctic [e.g., Smith and Stephenson, 2013] and have the potential to improve seasonal predictions in northern midlatitudes, as there is growing evidence for the influence of Arctic sea ice anomalies on atmospheric circulation and precipitation patterns [e.g., Balmaseda *et al.*, 2010; Screen, 2013]. Here we conduct the first coordinated, multimodel suite of so-called “perfect-model” ensemble prediction experiments to diagnose the inherent potential seasonal to interannual predictability of Arctic sea ice in current global climate models (GCM).

Similar idealized prediction studies with individual GCMs have shown that there is little hope to predict decadal pan-Arctic sea ice anomalies [Tietsche *et al.*, 2013]. However, there might be potential to predict decadal anomalies along the ice edges of the North Atlantic [Koenigk *et al.*, 2012]. On seasonal to interannual time scales, there is growing evidence from idealized studies that skillful sea ice predictions are possible [Koenigk and Mikolajewicz, 2008; Blanchard-Wrigglesworth *et al.*, 2011; Holland *et al.*, 2011]. Recently, progress has also been made in predicting observed sea ice extent/area on seasonal time scales: Chevallier *et al.* [2013], Wang *et al.* [2013], and Sigmond *et al.* [2013] demonstrate skillful predictions up to a few months ahead.

However, two questions remain unanswered: first, for lead times longer than a few months the seasonal prediction studies mentioned above mostly show little skill beyond the secular trend. Has the limit of initial-value predictability been reached already, or are insufficient knowledge of Arctic initial conditions and model deficiencies impeding skillful predictions? Second, previous studies did not explicitly assess spatial predictions of monthly resolved sea ice thickness and concentration, which are much more relevant both for practical applications and theoretical understanding. Furthermore, the predictability of sea ice volume and thickness has been discussed much less than the predictability of sea ice area and extent, due to a lack of observations to date.

Here we address the questions raised above by performing a suite of idealized ensemble prediction experiments with four state-of-the-art GCMs. The experimental design allows us to diagnose the predictability

**Table 1.** Overview of Control (CTRL) and Prediction Ensemble Runs Performed for This Study

Model	HadGEM1.2	MPI-ESM-LR	GFDL CM3	EC-Earth V2
CTRL years	249	200	200	200
Start dates	10	12	8	9
Ensemble size	16	9	16	7

inherent to the models. We analyze potential prediction error growth of the integrated sea ice properties volume and extent as well as the spatial distribution of sea ice concentration and thickness prediction errors and examine the physical processes that cause error growth.

## 2. Models and Methods

We use four different coupled climate models: GFDL CM3 [Donner *et al.*, 2011; Griffies *et al.*, 2011], HadGEM1.2 [Johns *et al.*, 2006; Shaffrey *et al.*, 2009], MPI-ESM-LR [Notz *et al.*, 2013; Jungclaus *et al.*, 2013], and EC-Earth V2 [Hazeleger *et al.*, 2012; Sterl *et al.*, 2012]. All of the models have a fully prognostic sea ice component, which accounts for changes in sea ice from both thermodynamic and advective processes that occur in interaction with the atmosphere above and the ocean below. The sea ice model components have conceptual differences in their treatment of important aspects of sea ice dynamics, like the local ice thickness distribution, vertical heat flux through the ice, and heat exchange at the ice-ocean interface.

Except for HadGEM1.2, we use exactly the same model versions that have been used for the Coupled Model Intercomparison Project Phase 5 (CMIP5). They have all been thoroughly evaluated against observations during the model development phase, and their weaknesses and strengths are well-documented (see above references).

Our intent is to diagnose the limits of *initial-value* predictability of Arctic sea ice inherent to the GCMs. This provides an upper bound for the predictive skill obtainable from predictions of observed climate with the same models, because initial conditions are observed imperfectly and models are biased. To diagnose the inherent predictability, we perform a suite of ensemble predictions and calculate skill measures by treating each ensemble member in turn as hypothetical observations, following the methodology of Collins [2002].

The presence of strong secular trends in the observed Arctic sea ice cover complicates the analysis of ensemble prediction studies for two reasons: (i) a significant fraction of the skill of any prediction system is due to predicting the trend more or less correctly [Sigmond *et al.*, 2013], which we are not interested in here, and (ii) the properties of the system change over time, so that it is difficult to sample enough predictions with the same baseline climate.

Therefore, as a baseline for the predictability experiments, we perform long present-day control simulations. The radiative forcing is representative of the year 1990 for HadGEM1.2 and GFDL CM3, and for the year 2005 for MPI-ESM-LR and EC-Earth V2. After a spin-up phase of about 100 years, each model is integrated for at least 200 more years to get a good estimate of the mean state, the remaining drift, and natural variability.

The modeled present-day sea ice mean state and variability in the control run differ considerably between the models but encompass the observed state between 1983 and 2012 (see Figure S1 in the supporting information).

During the 200 years of control simulation, we choose 8–12 individual years as start years for ensemble predictions, deliberately sampling the range of natural variability in terms of pan-Arctic sea ice extent and volume and ocean heat transport into the Arctic, while trying to have start years separated far enough to consider them independent. Ensembles have 7–16 members, and the maximum lead time is 3 years. Refer to Table 1 for details on the integrations performed.

Initial conditions for the ensembles are created by perturbing the sea surface temperature field in the control run state randomly by a very small amount (spatially uncorrelated Gaussian noise with a standard deviation of  $10^{-4}$  K). The perturbation is so small that it is equivalent to assuming perfect knowledge of the initial conditions, and the evolution of the ensemble is solely determined by the chaotic nature of the climate system. Start date is always 1 July, which allows an assessment of seasonal predictions of the late-summer sea ice minimum relevant for applications like operational shipping forecasts. Note that other seasonal predictions are more commonly started in May, which might lead to a sharply decreased skill in predicting the late-summer minimum (J. J. Day *et al.*, Pan-Arctic and regional sea ice prediction: initialisation month dependence, submitted to *Journal of Climate*, 2013).

We characterize the ensemble predictions by their root-mean-square error (RMSE) and their anomaly correlation coefficient (ACC), treating each ensemble member in turn as hypothetical observations. While the former is a measure of forecast accuracy directly relevant to applications, the latter is a measure of predictive skill. If  $x_{ij}$  is the value of the climate variable  $x$  for the  $i$ th ensemble member of the  $j$ th prediction ensemble, then we calculate the RMSE as

$$\text{RMSE} = \sqrt{\langle (x_{ij} - x_{kj})^2 \rangle_{i,j,k \neq i}} \quad (1)$$

where  $\langle \cdot \rangle_i$  denotes the expectation value, to be calculated by summing over the specified index with appropriate normalization. The ACC between any two ensemble members is calculated as

$$\text{ACC} = \frac{\langle (x_{ij} - \mu_j)(x_{kj} - \mu_j) \rangle_{i,j,k \neq i}}{\langle (x_{ij} - \mu_j)^2 \rangle_{i,j}}, \quad (2)$$

where  $\mu_j$  is the climatological mean at the time of the  $j$ th ensemble prediction, represented by the linear fit to the control run time series (see Figures S2 and S3 in the supporting information). The ACC compares the forecast accuracy of the ensemble predictions with the forecast accuracy of climatology as a trivial reference forecast. When ACC is above zero, there is at least some potential for skillful predictions. Note that, in our idealized setup, prediction ensembles have neither unconditional nor conditional biases, and hence, the ACC gives the same information as the mean-square-error skill score (see *Goddard et al.* [2013] and supporting information).

It is important to keep in mind that the forecast errors defined here (RMSE) are *potential* forecast errors, i.e., a lower bound for forecast errors achievable when predicting observations with the same GCMs. In the same way, the predictive skill defined here (ACC) is potential predictive skill, i.e., an upper bound for predictive skill achievable when predicting observations with the same GCMs.

### 3. Predictability Estimates

#### 3.1. Growth of Ensemble Prediction Errors

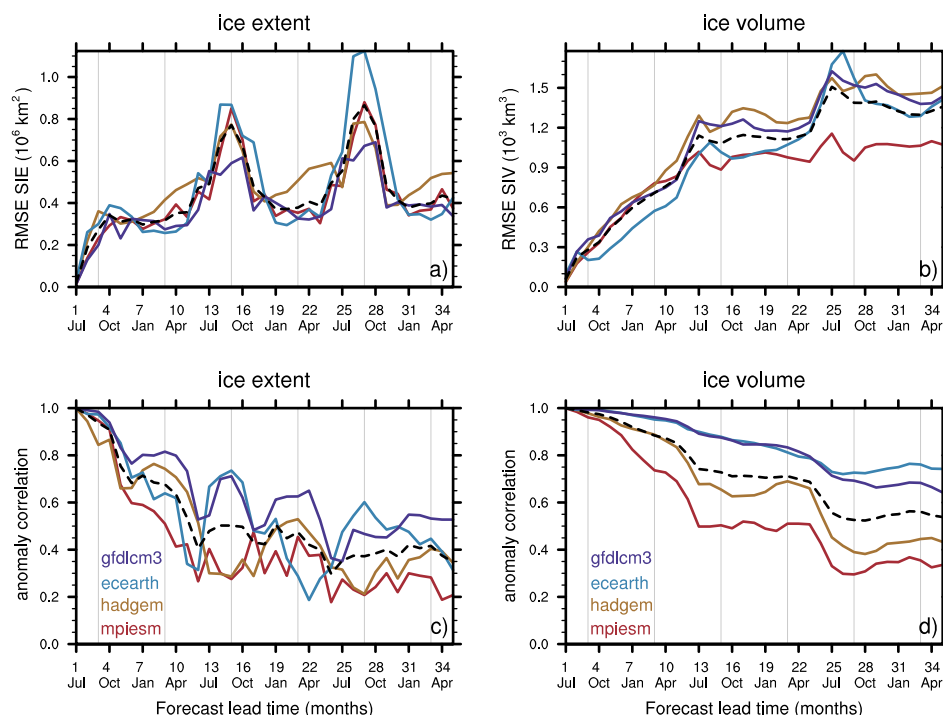
We calculate RMSE and ACC of pan-Arctic sea ice extent (SIE) and sea ice volume (SIV) from monthly data for lead times between 1 and 36 months, to characterize the growth of potential prediction errors and the decay of potential predictability.

The SIE RMSE (Figure 1a) shows several distinct phases: during the first four lead months (July–October), there is rapid error growth. With the onset of the first freezing season in October, the RMSE growth levels off. RMSE then stays fairly constant throughout all freezing seasons (lead months 4–11, 16–23, and 28–36) with the average across models only increasing from  $0.3 \cdot 10^6 \text{ km}^2$  in the first freezing season to  $0.4 \cdot 10^6 \text{ km}^2$  in the third freezing season. During the melting seasons, SIE RMSE shows very pronounced peaks up to  $0.8 \cdot 10^6 \text{ km}^2$ . This reflects the increased natural variability of SIE during summer as diagnosed from the control runs (see Figure S1 in the supporting information; note that the RMSE saturates at  $\sqrt{2}$  of the control run standard deviation, as pointed out by *Collins* [2002]).

The SIV RMSE (Figure 1b) shows a fairly linear initial growth throughout the first 10 lead months (until May) up to values of  $1000 \text{ km}^3$ . From May to July, error growth accelerates, and then levels off until the following May, when the pattern repeats.

Despite obvious intermodel differences, the following features of RMSE growth for SIE and SIV seem to be robust across models: (i) fast initial error growth for the first 4 (SIE) to 10 (SIV) lead months, (ii) subsequent slow error growth with an initial drop of errors during freezing season, (iii) fast error growth during the next melting season. A possible simple explanation for the faster error growth during the melting season is the dominance of the positive ice-albedo feedback, which acts to amplify initial perturbations of the sea ice conditions. Similarly, a possible simple explanation for the slower error growth during the freezing season is the dominance of the negative ice thickness-growth rate feedback, which acts to dampen initial perturbations of the sea ice conditions [cf. *Tietsche et al.*, 2011].

As a measure for potential predictability, we now discuss the anomaly correlation coefficient (ACC). We test the statistical significance of nonzero ACC at the 95% confidence level following *Collins* [2002] and find that, except for the lowest three data points for SIE ACC, all ACC values shown in Figures 1c and 1d are statistically



**Figure 1.** (a and b) Lead-time dependence of SIE RMSE and SIV RMSE for all models. (c and d) Lead-time dependence of SIE ACC and SIV ACC for all models. September and March are marked by thin gray vertical lines. Dashed lines represent the averages across models.

significant. For SIE (Figure 1c), the average across models decreases quickly during the first 12 lead months to a value of about 0.5, with a subsequent slower decrease to 0.4 at 36 months lead time. However, there are large intermodel differences beyond lead month 5, in both the magnitude and temporal evolution of ACC.

Comparing our findings of potential SIE ACC with the actual detrended SIE ACC reported by *Sigmond et al.* [2013] suggests that there is a high potential to improve sea ice predictions in operational forecast systems. For instance, predictions started in July shown by *Sigmond et al.* [2013] have no significant ACC after the first October, whereas in our experiments at that time potential ACC still has a high value of 0.6 to 0.8 consistently in all models.

The SIV ACC (Figure 1d) shows a much smoother temporal evolution. A generic stepwise decrease in skill over lead time is evident, as already discussed for the SIV RMSE (Figure 1b). At the maximum lead time of 36 months, all models have statistically significant ACC. However, differences between the models are considerable: while HadGEM1.2 and MPI-ESM-LR have clearly expressed seasonal variations, and an ACC of around 0.4 at the maximum lead time, GFDL CM3 and EC-Earth V2 show only a weak seasonal signal, and even at the maximum lead time have high ACC of around 0.7.

The considerable intermodel differences in SIV ACC may be related to different amounts of low-frequency SIV variability simulated by the models. The SIV control run time series has a larger fraction of the total variability on decadal time scales for EC-Earth V2 and GFDL CM3 when compared to HadGEM1.2 and MPI-ESM-LR (not shown). Therefore, once SIV in EC-Earth V2 or GFDL CM3 is initialized above or below the climatological mean, it will tend to stay above or below the mean for longer, which would explain the higher anomaly correlation between ensemble members. We note that for other possible measures of potential predictability the intermodel differences are smaller. For instance, potential predictability based on RMSE [Collins, 2002] gives a range of 0.2 to 0.4 for all models at 36 months lead time (not shown).

### 3.2. Spatial Patterns of Ensemble Prediction Errors

For practical applications of Arctic sea ice predictions, the spatial patterns of sea ice thickness (SIT) and sea ice concentration (SIC), rather than the aggregated quantities SIV and SIE are relevant. Looking at spatial patterns also avoids overconfidence in the predictions, as it excludes compensation of errors of opposite sign in different locations. In this section, we concentrate on discussing the spatial patterns of potential errors for

predictions of SIT and SIC. These spatial potential forecast errors, together with an estimate whether there is potential skill, are essential information for forecast applications.

Figure 2a shows the SIC RMSE for all models at three different lead times. There is reasonable agreement between the models that for the first September (lead month 3), highest forecast errors are to be expected in the marginal ice zone in the Arctic Ocean, where natural variability is high (cf. Figures S4 and S5 in the supporting information), whereas in the Arctic Basin forecast errors are lower.

In the first March after initialization (lead month 9), SIC RMSE is large in the vicinity of the climatological ice edge in the North Atlantic / North Pacific, but very small (below 4%) in the Arctic Ocean. In the second September (lead month 15), SIC RMSE is substantially larger than in the first September, and its pattern is already similar to the pattern of SIC variability in the control run, which indicates low potential predictive skill (stippling in Figure 2 and Figure S6 in the supporting information). Here potential predictive skill is calculated as  $1 - \text{RMSE}/(\sqrt{2}\sigma_{\text{ref}})$ , with  $\sigma_{\text{ref}}$  the standard deviation of the detrended control run [see Collins, 2002].

It is instructive to compare Figure 2a to Figure 1a: analyzing only SIE errors might give an overconfidence in how well the models agree. For instance, although GFDL CM3 and MPI-ESM-LR have virtually the same SIE RMSE for the first September, the corresponding SIC RMSE fields are quite different.

The spatial distribution of SIT RMSE (Figure 2b) shows that, for the first September, the largest potential errors tend to occur in the marginal ice zone of the Arctic Ocean, but there are considerable intermodel differences. Again, it is interesting to compare the SIT RMSE shown in Figure 2b with the SIV RMSE shown in Figure 1b. Converting the SIV RMSE for the first September to an expected SIT RMSE by dividing over the ice-covered surface area yields a value of about 3 cm. This is far lower than the actual average SIT RMSE at a specific grid point, which is more than 30 cm. This discrepancy indicates that spatial error compensation plays an important role and is certainly an important point to keep in mind when interpreting the large-scale integrated quantity SIV.

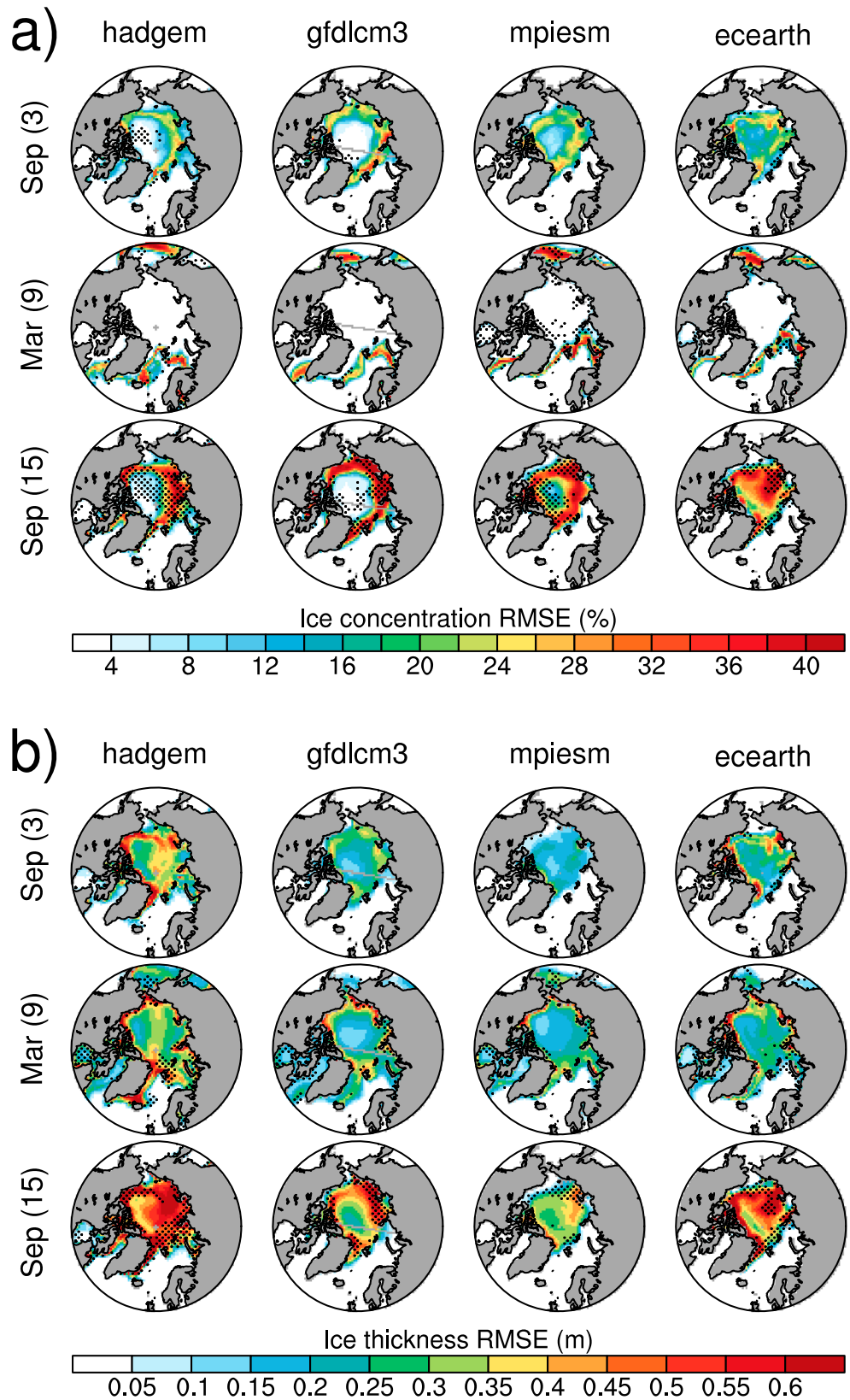
Moving on to SIT prediction errors for the first March after initialization, we see that, despite obvious differences between the models, some consistent features emerge: while SIT errors are mainly low in the interior of the Arctic and close to the ice edge (between 0.2 and 0.3 m), they are amplified along the Arctic coasts (0.6 m and higher). This is simulated by all models, and it is hence likely that this will be a robust feature of any future prediction system. It is not a trivial ice edge effect, because the errors are not amplified where the ice edge is away from coasts (Labrador, Barents, and Kara Seas, and North Pacific). In the second September (lead month 15), SIT errors are again quite model dependent, but much larger than in the first September, by a factor of 2 to 3. For comparison of the SIV RMSE patterns in Figure 2b with the control run variability, and for an estimate of potential predictive skill, refer to Figures S7–S9 in the supporting information.

### 3.3. Physical Processes

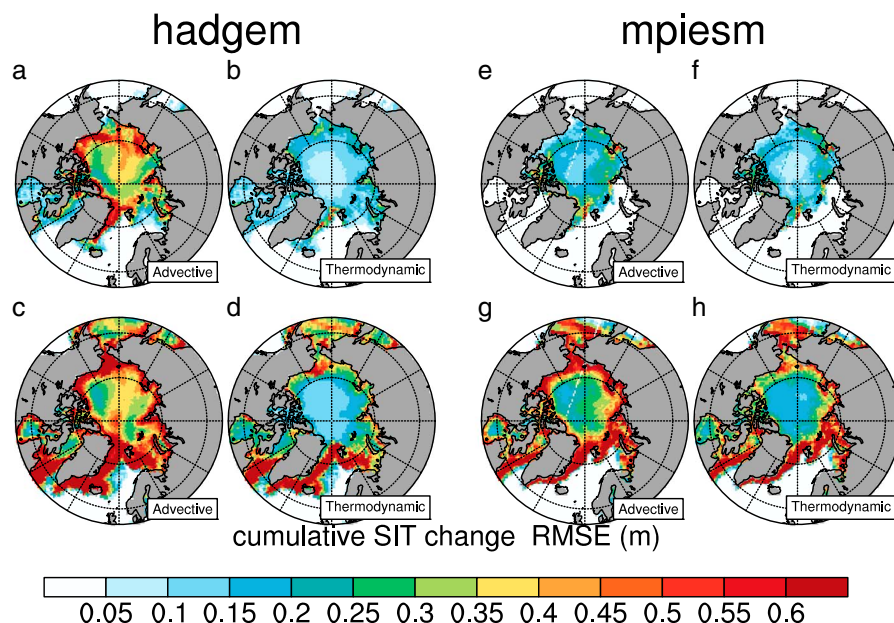
Sea ice thickness at a particular location can be changed by two classes of physical processes: (i) advective change from divergent thickness transport, and (ii) thermodynamic change from heat fluxes at the interfaces of the ice that lead to freezing or melting. Both have different spatial and temporal characteristics, so differentiating to what degree the potential prediction errors discussed above are caused by either of the two process classes provides insight into pathways through which SIT predictability is lost. Furthermore, because advective processes conserve ice volume, predictions of pan-Arctic SIV are only affected by thermodynamic processes. Thus, even if two given models have similar predictability of SIV, differences in simulating advective processes could still lead to large differences in predictability of SIT.

We separate the SIT errors shown in Figure 2b into a thermodynamic and an advective contribution by diagnosing the respective cumulative thickness changes since initialization (also see Holland *et al.* [2011]) and calculating their RMSE as in equation (1). We have this diagnostic available for only two models, HadGEM1.2 and MPI-ESM-LR. Note that the total (squared) RMSE shown in Figure 2b is not always the sum of thermodynamic and advective (squared) RMSE, since the two often have a strong negative covariance (for details, see supporting information).

Figure 3 suggests that—maybe somewhat surprisingly, cf. Holland *et al.* [2011]—advective processes are more important for the potential SIT prediction errors than thermodynamic processes: nowhere is the thermodynamic RMSE significantly larger than the advective RMSE. For the first September, thermodynamic SIT RMSE is very similar in HadGEM1.2 and MPI-ESM-LR, whereas advective SIT RMSE is much larger



**Figure 2.** (a) SIC RMSE and (b) SIT RMSE in lead months 3 (September), 9 (March), and 15 (September) for all models. Stippling indicates areas where potential predictive skill is below 10% (see Figures S6 and S9 in the supporting information).



**Figure 3.** RMSE of accumulated thickness change since initialization for (a–d) HadGEM1.2 and (e–h) MPI-ESM-LR. Thickness change due to advective processes (Figures 3a, 3c, 3e, and 3g). Thickness change due to thermodynamic processes (Figures 3b, 3d, 3f, and 3h). Upper row is for the end of September (lead month 3), lower row is for the end of March (lead month 9).

in HadGEM1.2. Thus, the larger total SIT RMSE of HadGEM1.2 shown in Figure 2b is dominated by advective processes. The higher advective SIT RMSE in HadGEM1.2 is readily explained by its thicker climatological ice cover (see Figure S7 in the supporting information).

The coastal error amplification in March is dominated by advective processes in both models. Scale analysis of the sea ice velocity equation reveals that the dominating force balance is between atmospheric drag and internal sea ice forces [Tietsche, 2012, p.64]. Therefore, it is highly probable that the coastal amplification of forecast errors is caused by forecast errors in surface winds. Due to the ubiquitous tendency of advective and thermodynamic changes to counteract each other, the thermodynamic RMSE is also amplified at the coasts, but not enough to fully compensate the advective RMSE. Around the winter ice edge in the North Atlantic sector, where advected sea ice quickly melts, both thermodynamic and advective RMSE are each much larger than the total RMSE, but compensate each other to give low total RMSE.

In summary, we find that (i) advective SIT errors are more important than thermodynamic SIT errors in determining the SIT RMSE pattern for the first September, (ii) the coastal error amplification is caused by advective processes, and (iii) the importance of model-dependent advective errors might explain why models look more similar in terms of SIV RMSE than in terms of SIT RMSE.

#### 4. Summary and Conclusions

We perform idealized ensemble predictions with four state-of-the-art GCMs and establish the first inter-model comparison of potential predictability of Arctic sea ice on seasonal to interannual time scales. The simulated present-day mean state and variability of Arctic sea ice show considerable differences between the models. Nevertheless, the models broadly agree on the growth rate and magnitude of potential forecast errors. Initially, there is fast growth of forecast error and fast decline of potential predictive skill as measured by anomaly correlation (4 months for sea ice extent, 10 months for SIV). Over the remaining lead time, forecast errors are close to the climatological saturation value and converge slowly to saturation, whereas anomaly correlation is significantly larger than zero throughout. This demonstrates that, in these GCMs, there is much potential for skillful initial-value predictions of Arctic sea ice on seasonal time scales, and some potential on interannual time scales.

In general, spatial patterns of sea ice thickness and concentration are more difficult to predict than the often used aggregated quantities of sea ice extent and sea ice volume: for each model, spatial error compensation

causes local ice thickness to be less predictable than the average pan-Arctic ice thickness. Additionally, there are large discrepancies between the patterns of sea ice thickness error between the models at least in some months. However, all models consistently simulate an amplification of forecast error close to the coasts of the Arctic Ocean in winter, a feature that we suggest to be related to advective sea ice processes. This finding implies operational forecasts of ice thickness are particularly difficult in places where societal benefits would be largest.

The results presented here represent the potential forecast errors and the potential predictive skill of Arctic sea ice that are inherent to the GCMs used, by evaluating forecasts with model output instead of observations. This approach provides a lower bound of forecast errors and an upper bound of predictive skill that is achievable with these GCMs and is therefore useful to determine where there is room for improvement in recently performed predictions of observed Arctic sea ice. The approach also allows us to analyze potential forecast errors and potential predictive skill of Arctic sea ice thickness, which receives much less attention than sea ice concentration at present, simply because there are no reliable large-scale observations of it. Nevertheless, sea ice thickness is directly relevant to shipping applications [Smith and Stephenson, 2013] and arguably at least as important as sea ice concentration for understanding and predicting the dynamics of the sea ice cover.

We see our results as a starting point and a resource for discussion about Arctic biases in coupled models, and about how to best improve real forecasts of Arctic climate. A direct comparison of the potential skill demonstrated here with the actual skill that the same models demonstrate when forecasting the observed climate is currently the subject of further research.

#### Acknowledgments

This research was supported by the APPOSITE project (grant NE/I029447/1) funded by the UK Natural Environment Research Council. Data storage and processing capacity was kindly provided by the British Atmospheric Data Centre (BADC). We thank Chris Ferro, James Screen, and Stefan Siebert for helpful discussions.

The Editor thanks Matthieu Chevallier and an anonymous reviewer for their assistance in evaluating this paper.

#### References

- Balmaseda, M. A., L. Ferranti, F. Molteni, and T. N. Palmer (2010), Impact of 2007 and 2008 Arctic ice anomalies on the atmospheric circulation: Implications for long-range predictions, *Q. J. R. Meteorol. Soc.*, *136*(652), 1655–1664, doi:10.1002/qj.661.
- Blanchard-Wrigglesworth, E., C. M. Bitz, and M. M. Holland (2011), Influence of initial conditions and climate forcing on predicting Arctic sea ice, *Geophys. Res. Lett.*, *38*, L18503, doi:10.1029/2011GL048807.
- 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. Clim.*, *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, *Clim. Dyn.*, *19*(8), 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*(3), 034011, doi:10.1088/1748-9326/7/3/034011.
- Donner, L. J., et al. (2011), The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3, *J. Clim.*, *24*(13), 3484–3519, doi:10.1175/2011JCLI3955.1.
- Goddard, L., et al. (2013), A verification framework for interannual-to-decadal predictions experiments, *Clim. Dyn.*, *40*(2), 245–272, doi:10.1007/s00382-012-1481-2.
- Griffes, S. M., et al. (2011), The GFDL CM3 Coupled Climate Model: Characteristics of the ocean and sea ice simulations, *J. Clim.*, *24*(13), 3520–3544, doi:10.1175/2011JCLI3964.1.
- Hazeleger, W., et al. (2012), EC-Earth V2.2: Description and validation of a new seamless earth system prediction model, *Clim. Dyn.*, *39*(11), 2611–2629, doi:10.1007/s00382-011-1228-5.
- Holland, M. M., D. A. Bailey, and S. Vavrus (2011), Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3, *Clim. Dyn.*, *36*, 1239–1253, doi:10.1007/s00382-010-0792-4.
- Johns, T. C., et al. (2006), The new Hadley Centre climate model (HadGEM1): Evaluation of coupled simulations, *J. Clim.*, *19*(7), 1327–1353, doi:10.1175/JCLI3712.1.
- Jungclaus, J. H., N. Fischer, H. Haak, K. Lohmann, J. Marotzke, D. Matei, U. Mikolajewicz, D. Notz, and J. S. von Storch (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*(2), 422–446, doi:10.1002/jame.20023.
- Kay, J. E., M. M. Holland, and A. Jahn (2011), Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world, *Geophys. Res. Lett.*, *38*, L15708, doi:10.1029/2011GL048008.
- Koenig, T., and U. Mikolajewicz (2008), Seasonal to interannual climate predictability in mid and high northern latitudes in a global coupled model, *Clim. Dyn.*, *32*(6), 783–798, doi:10.1007/s00382-008-0419-1.
- Koenig, T., C. König Beatty, M. Caian, R. Döscher, and K. Wyser (2012), Potential decadal predictability and its sensitivity to sea ice albedo parameterization in a global coupled model, *Clim. Dyn.*, *38*(11–12), 2389–2408, doi:10.1007/s00382-011-1132-z.
- Notz, D., F. A. Haumann, H. Haak, J. H. Jungclaus, and J. Marotzke (2013), Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's Earth system model, *J. Adv. Model. Earth Syst.*, *5*, 173–194, doi:10.1002/jame.20016.
- Screen, J. A. (2013), Influence of Arctic sea ice on European summer precipitation, *Environ. Res. Lett.*, *8*(4), 044015, doi:10.1088/1748-9326/8/4/044015.
- Shaffrey, L. C., et al. (2009), U.K. HiGEM: The new U.K. high-resolution global environment model – Model description and basic evaluation, *J. Clim.*, *22*(8), 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, L. C., and S. R. Stephenson (2013), New Trans-Arctic shipping routes navigable by midcentury, *Proc. Nat. Acad. Sci.*, *110*(13), E1191–E1195, doi:10.1073/pnas.1214212110.
- Sterl, A., et al. (2012), A look at the ocean in the EC-Earth climate model, *Clim. Dyn.*, *39*(11), 2631–2657, doi:10.1007/s00382-011-1239-2.

- Tietsche, S. (2012), Initialization and predictability of Arctic sea ice in a global climate model, Ph.D. thesis, University of Hamburg.
- Tietsche, S., D. Notz, J. H. Jungclaus, and J. Marotzke (2011), Recovery mechanisms of Arctic summer sea ice, *Geophys. Res. Lett.*, *38*, L02707, doi:10.1029/2010GL045698.
- Tietsche, S., D. Notz, J. H. Jungclaus, and J. Marotzke (2013), Predictability of large interannual Arctic sea-ice anomalies, *Clim. Dyn.*, *41*(9), 2511–2526, doi:10.1007/s00382-013-1698-8.
- Wang, W., M. Chen, and A. Kumar (2013), Seasonal prediction of Arctic sea ice extent from a coupled dynamical forecast system, *Mon. Weather Rev.*, *141*, 1375–1394, doi:10.1175/MWR-D-12-00057.1.