,
Manuscript prepared for Geosci. Model Dev.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 21 August 2015

The Arctic Predictability and Prediction on Seasonal-to-Interannual TimEscales (APPOSITE) data set

Jonathan J. Day[1], Steffen Tietsche[2], Mat Collins[3], Virginie Guemas[4,5], Anabelle Guilliory[6], Helge Gößling[7], William W. J. Hurlin[8], Masaoshi Ishii[9], Sarah P. E Keeley[2], Daniela Matei[10], Rym Msadek[5], Michael Sigmond[11], Hiroaki Tatebe[12], and Ed Hawkins[1]

[1]NCAS-Climate, Department of Meteorology, University of Reading, Reading, UK
[2]European Centre for Medium-Range Weather Forecasts, Reading, UK
[3]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK
[4]Institut Català de Ciències del Clima, Barcelona, Spain
[5]CNRM/GAME, Toulouse, France
[6]British Atmospheric Data Centre, Rutherford Appleton Laboratory, Chilton, UK
[7]Alfred Wegener Institute, Bussestraße, Bremerhaven, Germany
[8]Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA
[9]Meteorological Research Institute, Tsukuba, Japan
[10]Max Planck Institute for Meteorology, Hamburg, Germany
[11]Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada
[12]Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

Correspondence to: J. J. Day (j.j.day@reading.ac.uk)

Abstract. Recent decades have seen significant developments in seasonal-to-interannual timescale climate prediction capabilities. However, until recently the potential of such systems to predict Arctic climate had not been assessed. This paper describes a multi-model predictability experiment which was run as part of the Arctic Predictability and Prediction On Seasonal to Inter-annual Timescales (APPOSITE) project. The main goal of APPOSITE was to quantify the timescales on which Arctic climate is predictable. In order to achieve this, a coordinated set of idealised initial condition predictability experiments, with seven general circulation models, was conducted. This was the first model intercomparison project designed to quantify the predictability of Arctic climate on seasonal to inter-annual timescales. Here we present a description of the archived data set (which is available at the British Atmospheric Data Centre) and an update of the project's results, taking the latest data into account. Although designed to address Arctic predictability, this data set could also be used to assess the predictability of other regions and modes of climate variability on these timescales, such as the El Niño Southern Oscillation.

1 Introduction

The rapid reduction in Arctic summer sea ice has increased demand for Arctic sea ice forecasts at seasonal-to-interannual time scales (Eicken, 2013). This information is crucial for end users in marine industries as well as local communities (Stephenson et al., 2013). 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; Peterson et al., 2014) 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 (e.g. melt ponds, Schröder et al., 2014) and incomplete knowledge of important initial state variables, such as sea ice thickness and subsurface ocean properties (Day et al., 2014a), which are not well observed. There is also an inherent limit to predictability in the Arctic climate system due to chaotic atmospheric variability (Blanchard-Wrigglesworth et al., 2011b; Holland et al., 2010). If the skill of a given forecast system is already 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. The "perfect-model" approach to estimating predictability involves producing initial-condition ensemble-predictions with a General Circulation Model (GCM), which are verified against the model itself rather than against observations of the real world (following Griffies and Bryan, 1997b). It is therefore not hampered by changes to the observational network over time or changes in predictability due to secular climate change, which hampers this kind of analysis in the real world. It therefore provides an upper bound for the predictive skill obtainable in a world governed by the same physical equations as the model (Hawkins et al., 2015), though may not necessarily be an upper bound for the limit of predictability in the real world (Eade et al., 2014; Shi et al., 2015).

The perfect model approach has previously been used to quantify and understand predictability of coupled modes of climate variability, such as the Atlantic Meridional-Overturning Circulation (AMOC) (e.g. Griffies and Bryan, 1997a; Collins, 2002; Pohlmann et al., 2004) and the El Niño Southern Oscillation (ENSO) (Collins et al., 2002), leading to the development of operational seasonal-to-decadal prediction systems based on atmosphere-ocean climate models (e.g. Smith et al., 2007; Jin et al., 2008).

Using this approach Collins et al. (2006) demonstrated that the timsecale on which the AMOC is predictable, in a GCM, is dependent on the GCM used. These inter-model differences in predictability arise because different GCMs have different representations of the underlying physical equations and parameters. It is therefore likely that there will be inter-model differences in predictability for other climate variables, so in order to robustly estimate the limits of predictability in the real world it is important to conduct such analyses in multiple GCMs. The APPOSITE MIP was designed to diagnose the limits of initial-value predictability of Arctic sea ice in multiple GCMs. Previous studies had estimated this limit in individual climate models, but with different experiment designs. All these experiments demonstrated initial condition predictability on seasonal-to-interannual timescales but with significant differences in the details (Blanchard-Wrigglesworth et al., 2011b; Holland et al., 2010; Koenigk and Mikolajewicz, 2009; Tietsche et al., 2013; Guemas et al., 2014). However, because the experimental protocol was inconsistent between the studies, it was not clear whether differences in predictability were inherent in the models themselves or due to differences in the experimental set-up. For the APPOSITE ensemble a consistent protocol was followed so that differences in predictability were only the result of differences in the models themselves. The first results of this project were presented in Tietsche et al. (2014).

Here we present a detailed description of the APPOSITE experiment, as archived at the British Atmospheric Data Centre (BADC) as well as an update on the results of Tietsche et al. (2014), including more models than available at the time of publication. The paper is outlined as follows: Section 2 describes the experiment in detail as well as the mean state of the models used, Section 3 includes and update of the perfect-model results of Tietsche et al. (2014) followed by the conclusions in Section 4. Details of the data set archived at the BADC are included as an appendix.

2 Description of the simulations

Seven different coupled climate models performed simulations for APPOSITE (see Table 1). Six of these models followed the same experimental protocol, which is described in Sections 2.1 & 2.2. One model, CanCM4 followed a slightly different protocol which is described in Section 2.3.

2.1 Control simulations

The presence of strong secular trends in the observed Arctic sea ice cover complicates the analysis of ensemble prediction studies. Predictability of the climate system changes with mean climate (DelSole et al., 2013; Holland et al., 2010) complicating the assessment of predictability in a transient climate. The experimental protocol therefore asked for both control simulations and ensemble predictions to be conducted in GCMs with forcing fixed at present-day values.

Since the perfect-model approach uses initial conditions generated by the model itself, in order to perform the experiments it is necesary to run a control simulation first. Long present-day control simulations with each model were run under fixed present-day radiative forcings. For practical reasons the year that the forcings correspond to differ, but by no more than a decade or two, between the different simulations. After a spin-up phase of about 100 years, each model is integrated for at least 100 more years (archived lengths listed in Table 1) to get a good estimate of the mean state, the remaining drift, and natural variability. It is worth noting that some of these simulations have significant drifts in the mean sea ice climatology (see Figures 1 & 2).

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 and E6F, we use exactly the same model versions that have been used for the Coupled Model Intercomparison Project Phase 5 (CMIP5). These models have all been well evaluated against observations during the model development phase, and their weaknesses and strengths are well-documented (see references in Table 1). However, in order to understand differences in sea ice predictability, we focus on differences in their mean state and variability.

The modeled present-day sea ice mean state and variability in the control runs differ considerably between the models but encompass the observed state between 1983 and 2012 (see Figures 3, 4 & 5). Before calculating the standard deviation, shown in Fig. 5, a linear trend was removed from sea ice extent and volume timeseries for each model. Interannual variability of summer sea ice extent appears to be negatively correlated to it's mean, in line with previous studies (Goosse et al., 2009; Holland et al., 2008). This does not appear to be the case for winter.

2.2 Ensemble predictions

To diagnose the inherent predictability in each of these models, 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). Depending on the model, start dates were selected for between 8 and 18 years of the control run. These were chosen to sample a range of high, low and medium sea ice states, while keeping start dates suficiently spaced in time to consider them independent (see Figure. 1). 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. Differences in the evolution of each ensemble member are solely determined by the chaotic nature of the simulated climate system. Ensembles have between 7 and 16 members, and each ensemble was run for 3 years, with the exception of MIROC5, which was run for 3.5 years.

A minimum contribution for models to be included in the APPOSITE experiment was to submit a control run and predictability experiments started on the 1st 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 (Blanchard-Wrigglesworth et al., 2011a; Day et al., 2014b). Although we restrict our analysis to the simulations started in July, some groups have also submitted simulations started in January, May and November. Refer to Table 1 for details on the integrations performed.

2.3 CanCM4 transient experiments

The set of simulations with the CanCM4 model was run using a different protocol, in order to facilitate direct comparison of these simulations with the CanSIPS operational seasonal prediction system, which uses the same climate model (Sigmond et al., 2013).

The CanCM4 simulations were different in two key respects. Firstly, they were run under a transient climate, with observed historical forcing agents prescribed. Secondly, initial condition ensembles were generated every year and only run for 1 year. In all other regards, such as the method of ensemble generation, these simulations are the same as the other APPOSITE perfect model simulations.

3 Perfect model intercomparison

An intermodel comparison of Arctic sea ice predictability, using four climate models, was published in Tietsche et al. (2014). Here we present an update of this study, including the MIROC5, E6F and CanCM4 climate models.

3.1 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 as the truth in turn, and comparing it with every other member as the forecast. The ensemble Normalised Root Mean Squared Error (NRMSE) compares forecast RMSE to the climatological variability:

$$\mathrm{NRMSE} = \frac{\sqrt{\langle (x_{kj} - x_{ij})^2 \rangle_{i,j,k\neq i}}}{\sqrt{2\sigma^2}} \tag{1}$$

where $\langle \cdot \rangle_i$ denotes the expectation value, to be calculated by summing over the specified index with appropriate normalization, $x_{ij}(t)$ is the sea ice extent at lead time t for the ith member of the jth ensemble. The denominator is the climatological RMSE between two independent realisations. Significance of this is calculated using an f-Test, following Collins (2002).

The second metric is the anomaly correlation coefficient (ACC). This is defined as:

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

where μ_j is the climatological mean at the time of the j-th ensemble prediction.

At some lead-time, both of these metrics become insignificantly different from their asymptotic limit (0 for ACC and 1 for NRMSE), and the lead-time at which this happens can be used to define the limit of predictability. However, the ACC metric is more conservative than the NRMSE metric and so differences in the lead-time at which these metrics become insignificant gives some spread in the estimate of the time when the limit of predictability is actually reached.

3.2 Fixed forcing experiments

Although sea ice extent predictability decreases rapidly during the first year, with the exception of EC-Earth, all models (and both metrics) show significant levels of predictability for the first year. After the first year of simulation, two of the models, MIROC5 and GFDL-CM3, show significant levels of predictability at all later lead times. At the other end of the predictability spectrum, E6F is only intermittently predictable after the first year. Predictability in E6F (and to a lesser extent HadGEM1.2) has a strong seasonal cycle with months surrounding the winter extent maximum significantly predictable through till the end of the simulation and no significant summer predictability after the first year.

Sea ice volume is much more predictable than sea ice extent in all models. Apart from E6F all models exhibit significant predictability in all 3 years of the simulations. In a prognostic predictability analysis with decadal simulations, Germe et al. (2014) similarly found that winter sea ice extent was predictable out to seven years in their model, compared to three years in summer and found that volume was predictable out to nine years ahead.

3.3 CanCM4 transient experiments

Both the NRMSE and ACC metrics indicate lower levels of predictability in CanCM4 for sea ice extent and sea ice volume. It is possible that the CanCM4 model actually has inherently lower levels of initial condition predictability than the other models. However, there are reasons to expect that both metrics will be more conservative using the transient protocol.

In the case of NRMSE, detrending a short timeseries reduces the climatological variance since without multiple ensemble members to estimate the forced trend, some internal variability is removed in attempting to remove the forced trend (see Hawkins et al., 2015).

In the case of ACC, the reference climate (which is a linear fit to the control run) is a much closer fit to the control in the case of the short CanCM4 transient control run than it is for the long fixed forcing control runs, which have large decadal anomalies, which will reduce the correlation This is exactly analogous to the way that the ACC between two timeseries is reduced by removing the trend from both.

4 Conclusions

In this paper we have presented the protocol for the APPOSITE Arctic sea ice predictability multi-model intercomparison. We have compared the mean state and variability of Arctic sea ice cover in the models with observed estimates and estimated the limit of initial condition Arctic sea ice extent and volume predictability, updating the results of Tietsche et al. (2014).

The results of this analysis can be summarised into the following points:

- The winter sea ice extent is predictable at interannual timescales (or possibly longer timescales) in all models.

- There is significant intermodel spread in the timescale at which summer sea ice extent is predictable, with some models not showing any interanual or longer timescale predictability, and others showing significant predictability throughout all months of the 3 year simulations.

- Sea ice volume is much more predictable than sea ice extent in all models. Apart from E6F all models exhibit significant predictability in all lead months up to 3 years.

All the data used in this study will be archived at the British Atmospheric Data Centre (http://catalogue.ceda.ac.uk/uuid/d330c7873c3f4880893bdedb547bea20, insert DOI when ready). As well as enabling the results of the APPOSITE project to be reproduced, this will also allow these predictability experiments to be utilised to improve understanding of predictability of other areas and variables, such as Antarctic sea ice cover (e.g. Holland et al., 2013) or even ENSO (e.g. Collins et al., 2002).

Appendix A: Database description

APPOSITE requested a specific set of variables from participants focused on sea ice analysis, but have archived many other variables besides. The file and directory naming convention, followed by the archived data set, is very similar to that followed by CMIP5 (http://cmip-pcmdi.llnl.gov/cmip5/output_req.html).

APPOSITE required participants to prepare their data files so that they meet the following constraints.

- The all data files are in netCDF binary file format and ideally conform to the CF (Climate and Forecast) metadata convention (outlined on the website http://cf-pcmdi.llnl.gov). In instances where it was not possible to produce fully CF complaint netCDF files, participants were required to follow the CMOR variable naming convention.

- There must be only one output variable per file.

- The file names have to follow the file naming convention outlined below.

Each variable is be contained in a single directory of a directory tree with the following structure:
<model>/<runtype>/<submodel&frequency>/<variable>
Where *runtype* is ‘ctrl’ or ‘pred’ for the control run or ensemble predictions respectively, *model* is

the name of the climate model (e.g. hadgem1_2, mpiesm, ...), *variable* is the CMOR name for a given climate variable and *submodel&frequency* indicates the model sub-component and frequency (e.g. Amon, Aday, Omon and Oday).

Files are named using the following convention:

<variable>_<submode&frequency>_<model>_<runtype>_<run>_<time>.nc

Where *run* is a concatenated string including the start year, prediction start month and ensemble member number for ensemble predictions (e.g. 2005Jul3); or simply contains “r1” for a control run.

For example,

tas_Amon_hadgem1_2_ctrl_r1_200501-200512.nc for control runs,

or

tas_Amon_hadgem1_2_pred_2005Jul3_200507-200806.nc for the 3rd ensemble member of an ensemble started on the 1st July 2005.

Acknowledgements. This work was supported by the Natural Environment Research Council (grant NE/I029447/1). Data storage and processing capacity was kindly provided by the British Atmospheric Data Centre (BADC). Thanks to Yanjun Jiao (CCCma) for his assistance with the CanCM4 simulations.

References

Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C. M., and DeWeaver, E.: Persistence and Inherent Predictability of Arctic Sea Ice in a GCM Ensemble and Observations, Journal of Climate, 24, 231–250, doi:10.1175/2010JCLI3775.1, http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3775.1, 2011a.

Blanchard-Wrigglesworth, E., Bitz, C., Holland, M., McCarthy, C., Takei, Y., Purcell, A., Dehecq, A., Tregoning, P., Potter, E., McClusky, S., et al.: Influence of initial conditions and climate forcing on predicting Arctic sea ice, Geophys. Res. Lett, 38, L18 503, 2011b.

Chevallier, M., Salas y Mélia, D., Voldoire, A., Déqué, M., and Garric, G.: Seasonal Forecasts of the Pan-Arctic Sea Ice Extent Using a GCM-Based Seasonal Prediction System, Journal of Climate, 26, 6092–6104, doi:10.1175/JCLI-D-12-00612.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00612.1, 2013.

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

Collins, M., Frame, D., Sinha, B., and Wilson, C.: How far ahead could we predict El Niño?, Geophysical Research Letters, 29, 130–1–130–4, doi:10.1029/2001GL013919, http://doi.wiley.com/10.1029/2001GL013919, 2002.

Collins, M., Botzet, M., Carril, A. F., Drange, H., Jouzeau, A., Latif, M., Masina, S., Otteraa, O. H., Pohlmann, H., Sorteberg, A., et al.: Interannual to decadal climate predictability in the North Atlantic: a multimodel-ensemble study, Journal of climate, 19, 1195–1203, 2006.

Day, J. J., Hawkins, E., and Tietsche, S.: Will Arctic sea ice thickness initialization improve seasonal forecast skill?, Geophys. Res. Lett., 41, 7566–7575, doi:10.1002/2014GL061694, http://onlinelibrary.wiley.com/doi/10.1002/2014GL061694/abstract, 2014a.

Day, J. J., Tietsche, S., and Hawkins, E.: Pan-Arctic and Regional Sea Ice Predictability: Initialization Month Dependence, Journal of Climate, 27, 4371–4390, doi:10.1175/JCLI-D-13-00614.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-13-00614.1, 2014b.

DelSole, T., Yan, X., Dirmeyer, P. A., Fennessy, M., and Altshuler, E.: Changes in Seasonal Predictability Due to Global Warming, Journal of Climate, p. 130820142556003, doi:10.1175/JCLI-D-13-00026.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-13-00026.1, 2013.

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., Golaz, J.-C., Ginoux, P., Lin, S.-J., Schwarzkopf, M. D., Austin, J., Alaka, G., Cooke, W. F., Delworth, T. L., Freidenreich, S. M., Gordon, C. T., Griffies, S. M., Held, I. M., Hurlin, W. J., Klein, S. A., Knutson, T. R., Langenhorst, A. R., Lee, H.-C., Lin, Y., Magi, B. I., Malyshev, S. L., Milly, P. C. D., Naik, V., Nath, M. J., Pincus, R., Ploshay, J. J., Ramaswamy, V., Seman, C. J., Shevliakova, E., Sirutis, J. J., Stern, W. F., Stouffer, R. J., Wilson, R. J., Winton, M., Wittenberg, A. T., and Zeng, F.: The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3, Journal of Climate, 24, 3484–3519, doi:10.1175/2011JCLI3955.1, http://journals.ametsoc.org/doi/abs/10.1175/2011JCLI3955.1, 2011.

Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L., and Robinson, N.: Do seasonal to decadal climate predictions underestimate the predictability of the real world?: SEASONAL TO DECADAL PREDICTABILITY, Geophysical Research Letters, pp. n/a–n/a, doi:10.1002/2014GL061146, http://doi.wiley.com/10.1002/2014GL061146, 2014.

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

Germe, A., Chevallier, M., Salas y Mélia, D., Sanchez-Gomez, E., and Cassou, C.: Interannual predictability of Arctic sea ice in a global climate model: regional contrasts and temporal evolution, Climate Dynamics, doi:10.1007/s00382-014-2071-2, http://link.springer.com/10.1007/s00382-014-2071-2, 2014.

Goosse, H., Arzel, O., Bitz, C. M., de Montety, A., and Vancoppenolle, M.: Increased variability of the Arctic summer ice extent in a warmer climate, Geophysical Research Letters, 36, doi:10.1029/2009GL040546, http://doi.wiley.com/10.1029/2009GL040546, 2009.

Griffies, S. and Bryan, K.: A predictability study of simulated North Atlantic multidecadal variability, Climate dynamics, 13, 459–487, 1997a.

Griffies, S. M. and Bryan, K.: Predictability of North Atlantic Multidecadal Climate Variability, Science, 275, 181–184, doi:10.1126/science.275.5297.181, http://www.sciencemag.org/cgi/doi/10.1126/science.275.5297.181, 1997b.

Griffies, S. M., Winton, M., Donner, L. J., Horowitz, L. W., Downes, S. M., Farneti, R., Gnanadesikan, A., Hurlin, W. J., Lee, H.-C., Liang, Z., Palter, J. B., Samuels, B. L., Wittenberg, A. T., Wyman, B. L., Yin, J., and Zadeh, N.: The GFDL CM3 Coupled Climate Model: Characteristics of the Ocean and Sea Ice Simulations, Journal of Climate, 24, 3520–3544, doi:10.1175/2011JCLI3964.1, http://journals.ametsoc.org/doi/abs/10.1175/2011JCLI3964.1, 2011.

Guemas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., Fučkar, N. S., Germe, A., Hawkins, E., Keeley, S., Koenigk, T., Salas y Mélia, D., and Tietsche, S.: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales: Arctic Sea-Ice Predictability and Prediction, Quarterly Journal of the Royal Meteorological Society, pp. n/a–n/a, doi:10.1002/qj.2401, http://doi.wiley.com/10.1002/qj.2401, 2014.

Hawkins, E., Tietsche, S., Day, J. J., Melia, N., Haines, K., and Keeley, S.: Aspects of designing and evaluating seasonal-to-interannual Arctic sea-ice prediction systems, Q.J.R. Meteorol. Soc., pp. n/a–n/a, doi:10.1002/qj.2643, http://onlinelibrary.wiley.com/doi/10.1002/qj.2643/abstract, 2015.

Hazeleger, W., Wang, X., Severijns, C., Ştefănescu, S., Bintanja, R., Sterl, A., Wyser, K., Semmler, T., Yang, S., Hurk, B. v. d., Noije, T. v., Linden, E. v. d., and Wiel, K. v. d.: EC-Earth V2.2: description and validation of a new seamless earth system prediction model, Clim Dyn, 39, 2611–2629, doi:10.1007/s00382-011-1228-5, http://link.springer.com/article/10.1007/s00382-011-1228-5, 2012.

Holland, M. M., Bitz, C. M., Tremblay, B., and Bailey, D. A.: The role of natural versus forced change in future rapid summer Arctic ice loss, in: Arctic sea ice decline: observations, projections, mechanisms, and implications, edited by DeWeaver, E., Bitz, C., and Tremblay, B., vol. 180 of *Geophys Monogr Ser*, AGU, Washington, 2008.

Holland, M. M., Bailey, D. A., and Vavrus, S.: Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3, Climate Dynamics, 36, 1239–1253, http://www.springerlink.com/index/g357731164vn69n1.pdf, 2010.

Holland, M. M., Blanchard-Wrigglesworth, E., Kay, J., and Vavrus, S.: Initial-value predictability of Antarctic sea ice in the Community Climate System Model 3, Geophysical Research Letters, pp. n/a–n/a, doi:10.1002/grl.50410, http://doi.wiley.com/10.1002/grl.50410, 2013.

Jin, E. K., Kinter, J. L., Wang, B., Park, C.-K., Kang, I.-S., Kirtman, B. P., Kug, J.-S., Kumar, A., Luo, J.-J., Schemm, J., Shukla, J., and Yamagata, T.: Current status of ENSO prediction skill in coupled ocean–atmosphere models, Climate Dynamics, 31, 647–664, doi:10.1007/s00382-008-0397-3, http://link.springer.com/10.1007/s00382-008-0397-3, 2008.

Johns, T., Durman, C., Banks, H., Roberts, M., McLaren, A., Ridley, J., Senior, C., Williams, K., Jones, A., Rickard, G., et al.: The new Hadley Centre climate model (HadGEM1): Evaluation of coupled simulations, Journal of Climate, 19, 1327–1353, 2006.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J. S.: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model, Journal of Advances in Modeling Earth Systems, 5, 422–446, doi:10.1002/jame.20023, http://onlinelibrary.wiley.com/doi/10.1002/jame.20023/abstract, 2013.

Koenigk, T. and Mikolajewicz, U.: Seasonal to interannual climate predictability in mid and high northern latitudes in a global coupled model, Climate Dynamics, 32, 783–798, doi:10.1007/s00382-008-0419-1, http://www.springerlink.com/content/x015q91107575um8/, 2009.

Notz, D., Haumann, F. A., Haak, H., Jungclaus, J. H., and Marotzke, J.: Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's Earth system model, Journal of Advances in Modeling Earth Systems, 5, 173–194, doi:10.1002/jame.20016, http://doi.wiley.com/10.1002/jame.20016, 2013.

Peterson, K. A., Arribas, A., Hewitt, H. T., Keen, A. B., Lea, D. J., and McLaren, A. J.: Assessing the forecast skill of Arctic sea ice extent in the GloSea4 seasonal prediction system, Climate Dynamics, doi:10.1007/s00382-014-2190-9, http://link.springer.com/10.1007/s00382-014-2190-9, 2014.

Pohlmann, H., Botzet, M., Latif, M., Roesch, A., Wild, M., and Tschuck, P.: Estimating the decadal predictability of a coupled AOGCM, Journal of climate, 17, 4463–4472, http://journals.ametsoc.org/doi/abs/10.1175/3209.1, 2004.

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

Schröder, D., Feltham, D. L., Flocco, D., and Tsamados, M.: September Arctic sea-ice minimum predicted by spring melt-pond fraction, Nature Climate Change, doi:10.1038/nclimate2203, http://www.nature.com/doifinder/10.1038/nclimate2203, 2014.

Schweiger, A., Lindsay, R., Zhang, J., Steele, M., Stern, H., and Kwok, R.: Uncertainty in modeled Arctic sea ice volume, Journal of Geophysical Research, 116, doi:10.1029/2011JC007084, http://doi.wiley.com/10.1029/2011JC007084, 2011.

Shaffrey, L. C., Stevens, I., Norton, W. A., Roberts, M. J., Vidale, P. L., Harle, J. D., Jrrar, A., Stevens, D. P., Woodage, M. J., Demory, M. E., Donners, J., Clark, D. B., Clayton, A., Cole, J. W., Wilson, S. S., Connolley, W. M., Davies, T. M., Iwi, A. M., Johns, T. C., King, J. C., New, A. L., Slingo, J. M., Slingo, A., Steenman-Clark, L., and Martin, G. M.: U.K. HiGEM: The New U.K. High-Resolution Global Environment Model—Model Description and Basic Evaluation, Journal of Climate, 22, 1861–1896, doi:10.1175/2008JCLI2508.1, http://journals.ametsoc.org/doi/abs/10.1175/2008JCLI2508.1, 2009.

Shi, W., Schaller, N., MacLeod, D., Palmer, T. N., and Weisheimer, A.: Impact of hindcast length on estimates of seasonal climate predictability, Geophysical Research Letters, 42, 1554–1559, doi:10.1002/2014GL062829, http://doi.wiley.com/10.1002/2014GL062829, 2015.

Sidorenko, D., Rackow, T., Jung, T., Semmler, T., Barbi, D., Danilov, S., Dethloff, K., Dorn, W., Fieg, K., Goessling, H. F., Handorf, D., Harig, S., Hiller, W., Juricke, S., Losch, M., Schröter, J., Sein, D. V., and Wang, Q.: Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: model formulation and mean climate, Climate Dynamics, doi:10.1007/s00382-014-2290-6, http://link.springer.com/10.1007/s00382-014-2290-6, 2014.

Sigmond, M., Fyfe, J. C., Flato, G. M., Kharin, V. V., and Merryfield, W. J.: Seasonal forecast skill of Arctic sea ice area in a dynamical forecast system, Geophysical Research Letters, 40, 529–534, doi:10.1002/grl.50129, http://doi.wiley.com/10.1002/grl.50129, 2013.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., and Murphy, J. M.: Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model, Science, 317, 796 –799, doi:10.1126/science.1139540, http://www.sciencemag.org/content/317/5839/796.abstract, 2007.

Stephenson, S. R., Smith, L. C., Brigham, L. W., and Agnew, J. A.: Projected 21st-century changes to Arctic marine access, Climatic Change, doi:10.1007/s10584-012-0685-0, http://link.springer.com/10.1007/s10584-012-0685-0, 2013.

Tietsche, S., Notz, D., Jungclaus, J. H., and Marotzke, J.: Predictability of large interannual Arctic sea-ice anomalies, Climate Dynamics, doi:10.1007/s00382-013-1698-8, http://link.springer.com/10.1007/s00382-013-1698-8, 2013.

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

Wang, W., Chen, M., and Kumar, A.: Seasonal Prediction of Arctic Sea Ice Extent from a Coupled Dynamical Forecast System, Monthly Weather Review, 141, 1375–1394, doi:10.1175/MWR-D-12-00057.1, http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-12-00057.1, 2013.

Watanabe, M., Suzuki, T., O'ishi, R., Komuro, Y., Watanabe, S., Emori, S., Takemura, T., Chikira, M., Ogura, T., Sekiguchi, M., Takata, K., Yamazaki, D., Yokohata, T., Nozawa, T., Hasumi, H., Tatebe, H., and Kimoto, M.: Improved Climate Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity, J. Climate, 23, 6312–6335, doi:10.1175/2010JCLI3679.1, http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3679.1, 2010.

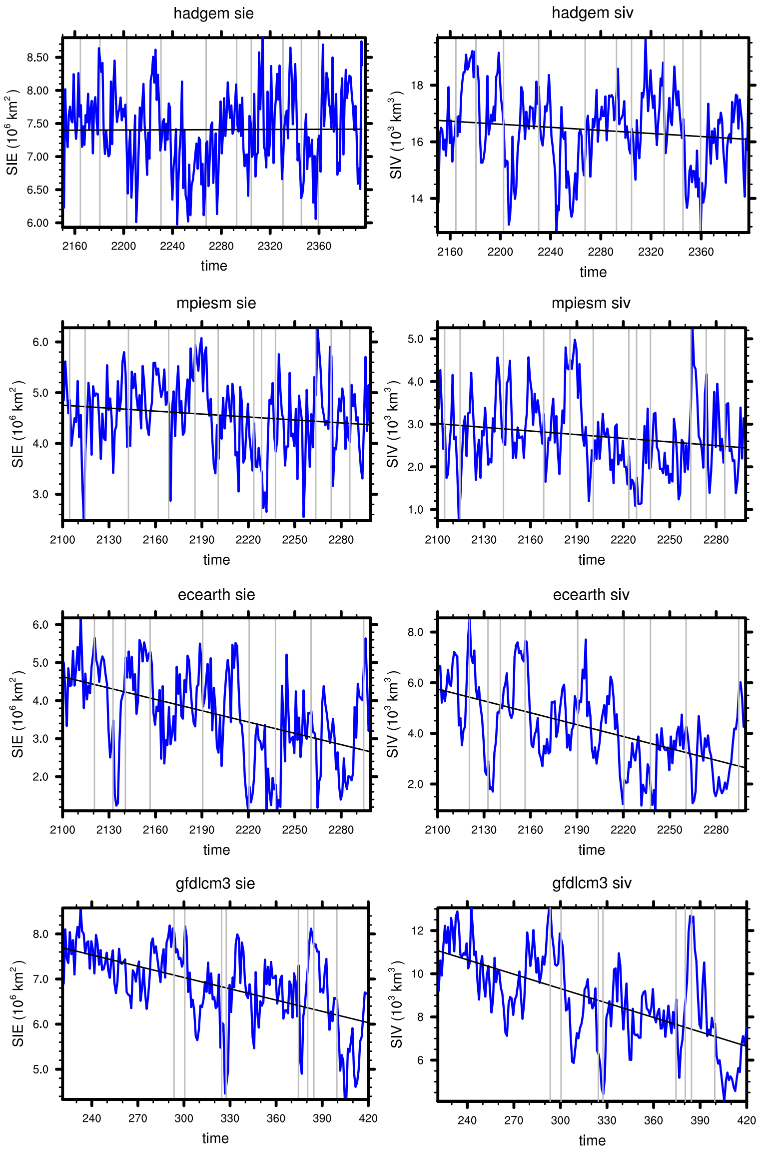

Figure 1. Timeseries of monthly mean September sea ice extent (sie, left column) and sea ice volume (siv, right column) in each model's control simulation (blue) with the line of best fit to data (black). Vertical grey lines indicate start years used to initialise simulations.

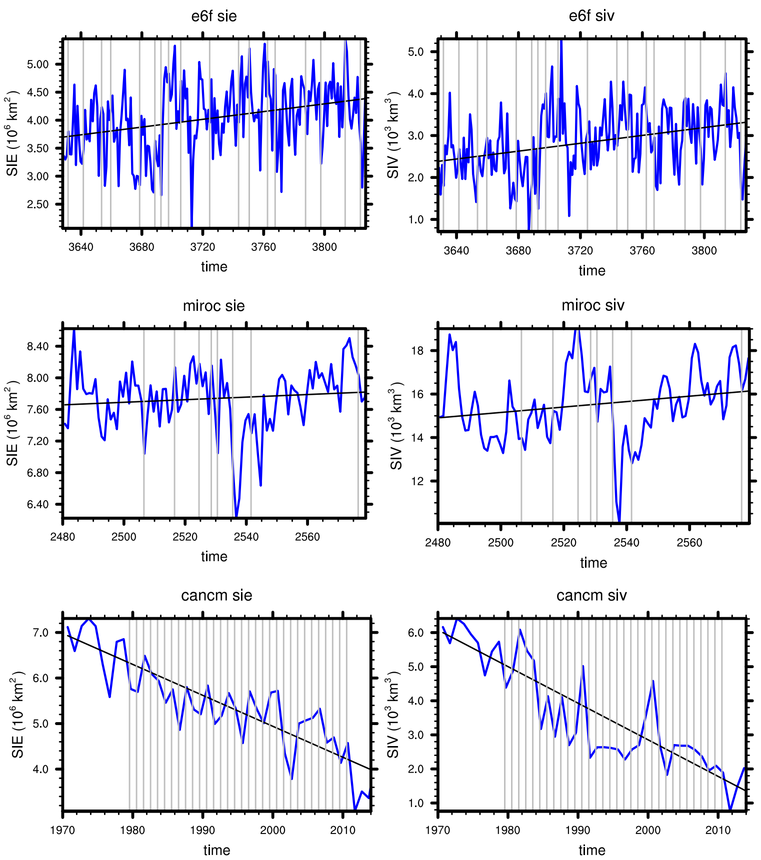

Figure 2. continuation of previous figure for the other GCMs used.

Table 1. Details of simulations submitted to the APPOSITE database.

Model	CTRL length	Forcing year	Start Dates	Start Months	Ensemble Size	References
HadGEM1.2	249	1990	10	Jan, May, Jul	16	Johns et al. (2006)
						Shaffrey et al. (2009)
MPI-ESM	200	2005	12(Jul), 16(Nov)	Jul, Nov	9(Jul), 16(Nov)	Notz et al. (2013)
						Jungclaus et al. (2013)
GFDL-CM3	200	1990	8	Jan, Jul	16	Donner et al. (2011)
						Griffies et al. (2011)
EC-Earth2.2	200	2005	9	Jul	7	Hazeleger et al. (2012)
MIROC5	100	1990	8	Jan, Jul	8	Watanabe et al. (2010)
E6F	200	1990	18	Jan, Jul	9	Sidorenko et al. (2014)
CanCM4	45	transient (1970-2014)	32	Jan, Jul,	10	Sigmond et al. (2013)

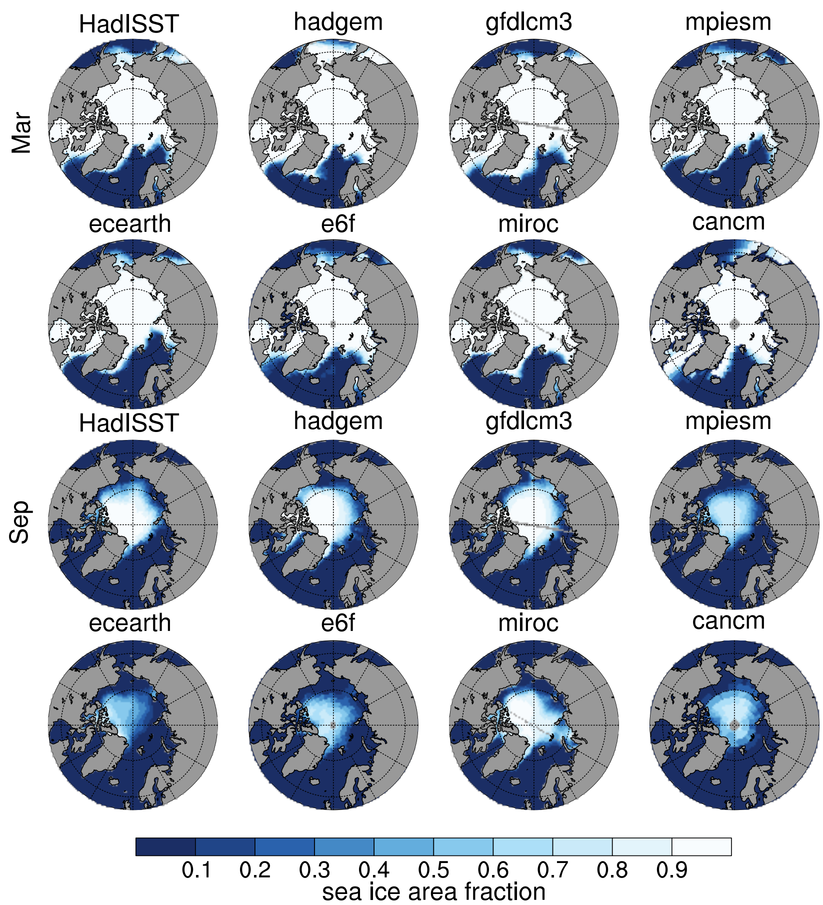

Figure 3. Average sea-ice concentration in present-day model control simulations and from HadISST (1983-2012) (Rayner et al., 2003).

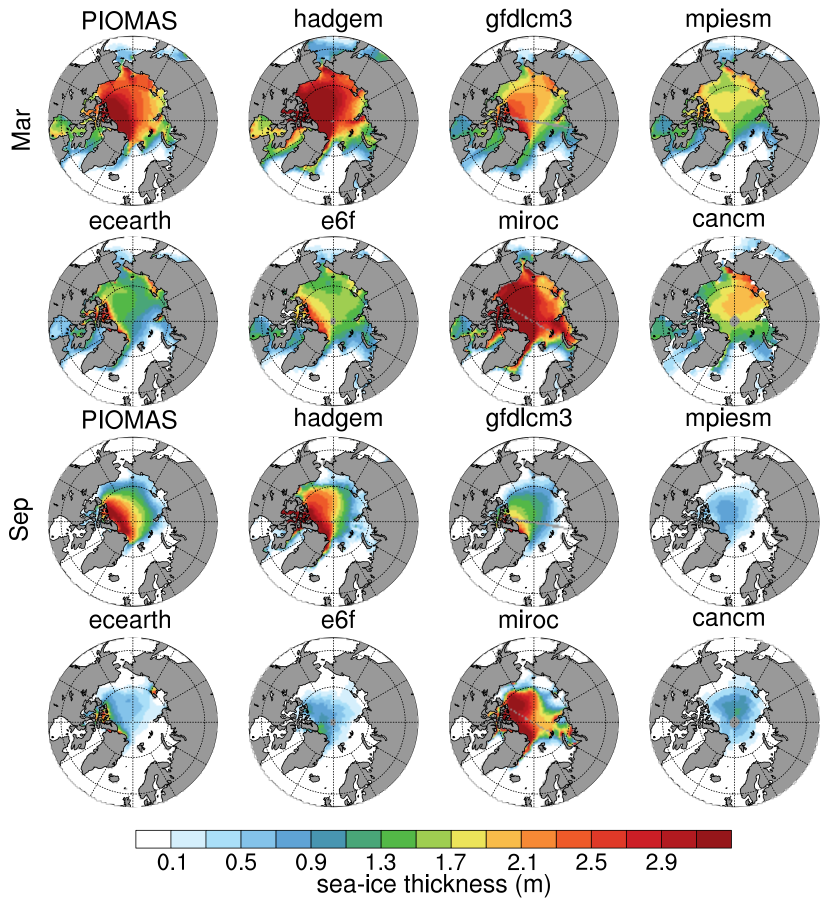

Figure 4. Average sea-ice thickness in present-day model control simulations and from PIOMAS (1983-2012) (Schweiger et al., 2011).

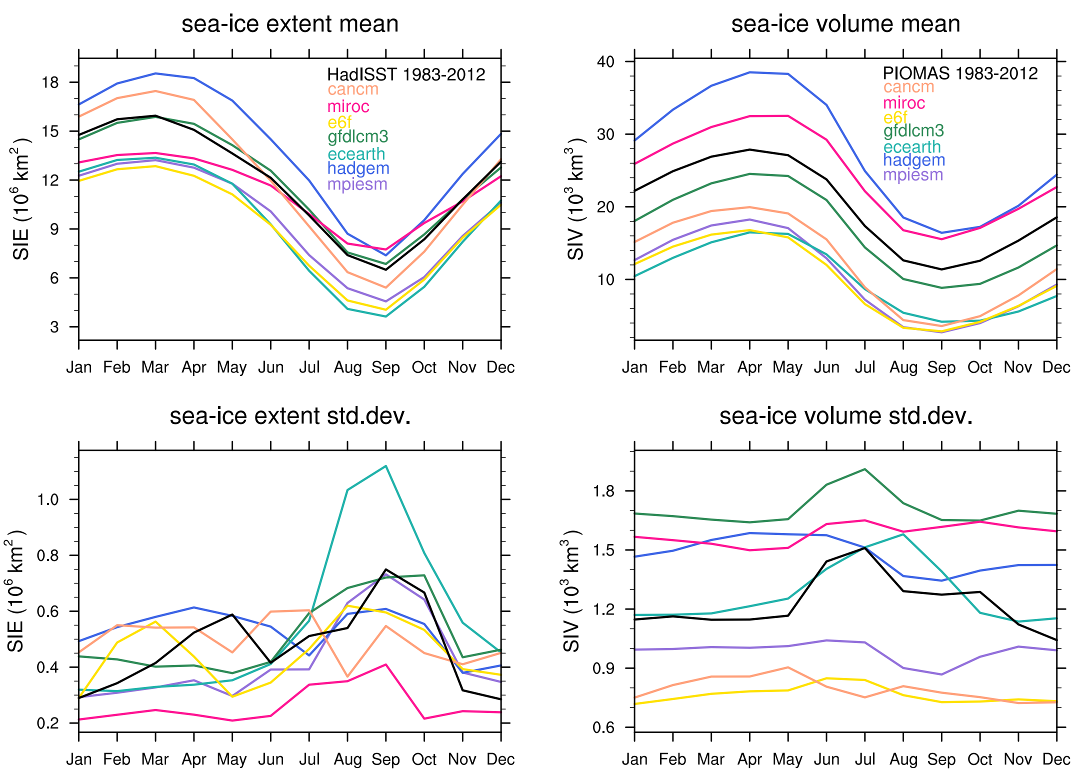

Figure 5. Seasonal cycle of monthly mean sea-ice extent (a), volume (b) and standard deviation of sea ice extent (c) and volume (d) in present-day model control simulations. The HadISST observations of sea ice extent and PIOMAS reconstruction of ice volume are included as a reference.

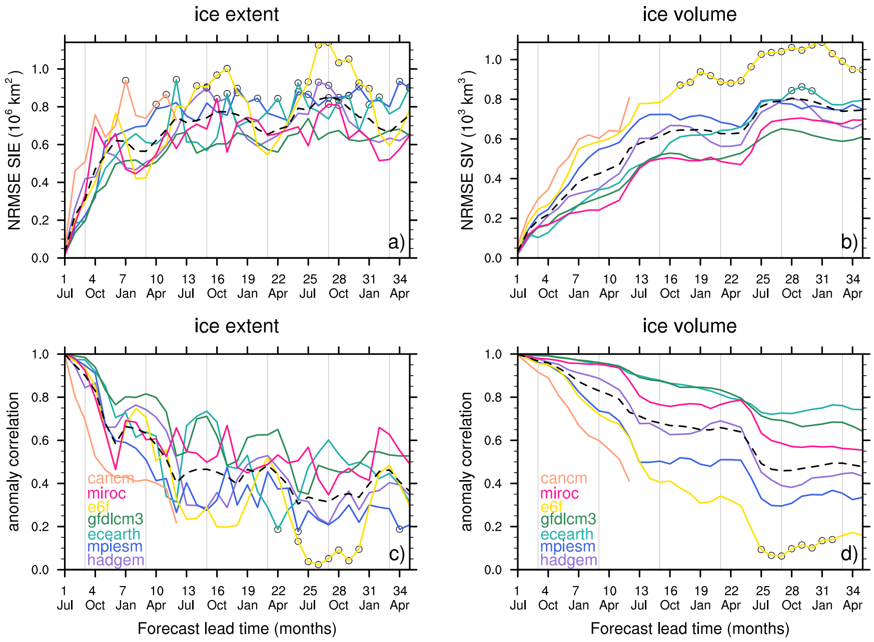

Figure 6. (a and b) Lead-time dependence of SIE NRMSE and SIV NRMSE for all models. (c and d) Lead-time dependence of SIE ACC and SIV ACC for all models. September and March are marked by thin gray vertical lines. Dashed lines represent the averages across models. Circles indicate where metrics do not indicate significant predictability (at 95%). Updated from Tietsche et al. (2014).