September 2011 Sea Ice Outlook
June Report

K.A. Peterson, A. McLaren, A. Arribas, M. Gordon, H. Hewitt and R. Wood
GloSea4 Seasonal Forecast System, UK Met Office, Exeter, UK.

*Caveat: This is an experimental projection and **does not** represent an official UK Met Office forecast.*

1) Extent Projection

4.0 ± 1.2 million square kilometres.

2) Method/Techniques (Coupled Atmosphere-ice-ocean-land surface model ensemble runs)

This projection is an experimental prediction from the UK Met Office seasonal forecast system, GloSea4 (Arribas et al., 2011). GloSea4 is an ensemble prediction system using the HadGEM3 coupled climate model (Hewitt et al., 2011). The particular realization of the flexible resolution HadGEM3 model used in the current GloSea4 system updated in September of 2010 is:

- Atmosphere/Soil Moisture: 85 level N96 (~120km) version of the UM (Met Office Unified Model; Davies et al., 2005). Top level is 85.0km.
- Ocean: 75 level ORCA1 (~1 deg tripolar ocean grid) version of NEMO (Nucleus for European Modelling of the Ocean) Ocean Model (Madic, 2008). Surface level is 1m with 8 levels in the top 10m.
- Sea Ice: CICE (Los Alamos sea ice model; Hunke and Lipscomb, 2010) with same horizontal resolution as the ocean ORCA1 grid and 5 ice categories (single thermodynamic layer).
- Land Surface: MOSES scheme (Met Office Surface Exchange Scheme; Essery et al., 2003) with same resolution as the atmospheric grid.
- Coupling: Every 3 hours.

As is standard for any seasonal prediction system, the system is composed of a seven month ensemble forecast initialized at 00Z daily. Every day, 2 ensemble members are generated with differing stochastic physics (Bowler et al., 2009), from which a set (42) of ensemble members are generated by combining many (21) days of lagged start dates. In turn, this ensemble forecast is calibrated, and bias corrected by a 14 year set of ensemble hindcasts. The hindcasts are initiated on the 1st, 9th, 17th, and 25th of each month, with each initialization date generating 3 different ensemble members via the stochastic physics parametrization.

Before coupling, the atmosphere and land surface were initialized to a re-gridded atmospheric analysis (NWP 4D-Var analysis for the forecast; Rawlins et al, 2007, and ERAI for the hindcast; Dee et al, 2009), while the ocean and sea ice were initialized with a version of the Met Office Optimal Interpolation (OI) scheme used for short term ocean forecasting (Storkey et al., 2010; Stark et al., 2008), but here adapted for the ORCA1 resolution. A description of the sea ice initialization and the performance of GloSea4 in the prediction of sea ice will be forthcoming (Peterson et al., 2011). Summarizing the initialization strategy:

- Atmosphere/Soil Moisture: NWP (N512; ~25km) 00Z analysis (Forecast); ERAI (T255; ~80km) 00Z analysis (Hindcast)
- Ocean: Optimal Interpolation Assimilation of Sea Surface Temperature, as well as temperature and salinity profiles using the same ORCA1 ocean model used in the forecast. Hindcast – Ocean Driven by Bulk Forcing derived from ERAI atmospheric data. Forecast – Ocean driven by flux forcing derived from NWP analysis atmospheric data.
- Sea Ice: Optimal Interpolation Assimilation of Ice concentration from satellite data. No assimilation of ice thickness was performed. Forcing: See ocean above.
- Frequency: Forecast -- Daily; Hindcast -- 4 times monthly (1^{st}, 9^{th}, 17^{th}, 25^{th})

Two important differences between the forecast initialization and the hindcast initialization should be highlighted. First the hindcast atmosphere and soil moisture initial conditions come from the ECMWF interim re-analysis (ERAI), whereas the forecast atmosphere and soil moisture initial conditions come from the UK Met Office real time Numerical Weather Production (NWP) analysis. The ERAI analysis does not fully resolve the stratosphere and has significantly different soil moisture characteristics to the NWP analysis. Both of these may influence the atmospheric general circulation that develops during the coupled forecast period, and in turn the sea ice circulation. More important for the purposes of this outlook is the different external forcing presented to the ocean and sea ice system during their assimilation: During the hindcast the ocean and sea ice are externally forced by (CORE) bulk forcing (Large and Yeager, 2009; Brodeau et al, 2010) as determined by atmospheric conditions supplied by the ERAI analysis. During the forecast, the ocean and sea ice are externally forced by fluxes derived from the NWP analysis. In particular, the top and bottom melt forcing, as well as other fluxes above and below the ice are derived under the assumption of 2m thick ice throughout the Arctic ice covered region. Since only the ice concentration is controlled by the sea ice assimilation strategy, this may have considerable consequences for the sea ice thickness. Indeed, we believe the sea ice thickness, particularly during and after the spring melt may be too thin in the forecast initial fields as compared to the hindcast initial fields. A convergence of forecast and hindcast strategy to having both use bulk forcing is planned, but is not available for the 2011 sea ice outlook – and may not be available for a 2012 outlook either – depending on upgrade schedules (dictated in turn by super-computer upgrade schedules).

Besides being used as a check on the skill of the forecast, the 1996-2009 hindcast that runs in parallel with the forecast is also used to calibrate the forecast for systematic model error. In particular, if the climatological ice extent from the hindcast is biased high or low compared to the observations over the 1996-2009 period, then the forecast is adjusted upward or downward accordingly. Unfortunately, this does mean our forecast ice concentrations will also be biased. Due to the very non-gaussian nature of ice concentration, particularly near the ice edge, it is virtually impossible to correct the ice concentration bias. Thus figures of ice concentration from the model that are given below will have a known (but not correctable) systematic basis. In these plots, the forecast and observed climatological ice extents will be displayed, and the viewer will have to qualitatively adjust for this bias.

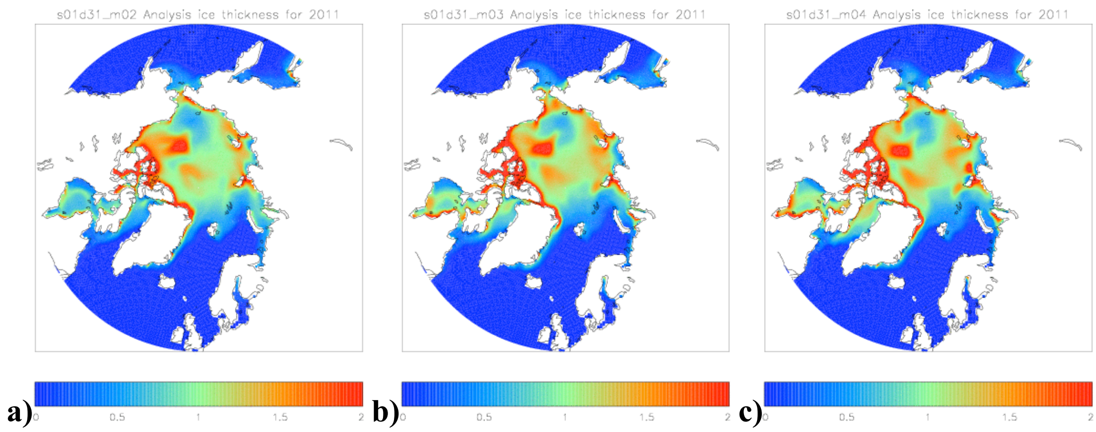


Figure 1. Plots of sea ice thickness from the sea ice analysis for a) February, b) March, and c) April. These would be the approximate thicknesses that the coupled model would be initiated with during these months – they would actually be initiated with the instantaneous values, not the monthly average values shown here. Note that the ice thickness does not increase much during this period.

3) <u>Rationale</u>

Our forecast of 4.0±1.2 million square kilometres was based on the seasonal forecast data for September 2011 from a start date centred on 31/03/2011, and included 7 month forecasts initiated between 21/03/2011 and 10/04/2011. The hindcast dates used for calibration were therefore hindcasts initiated on 25/03, 01/04, and 09/04 of each of the hindcast years. These **were not** the latest start dates available to us, and therefore neglect any useful initialization data available from April and May. Furthermore, September represents the sixth or seventh month of these seasonal forecast integrations, with little expectation of skill in the atmospheric circulation at these lead times, or for that matter, any expectation of skill in atmospheric circulation over the three summer months (June/July/August) instrumental for the development of ice in September. The reasons for this decision were based on the following:

1. The subsequent April and May start dates developed a suspected large negative bias in the September 2011 ice extent compared with the hindcast period of 1996 through 2009.
   - The ice thickness coming out of the ocean and ice analysis for April and May appears to be unrealistically thin, giving too much ice melt during the summer (Figure 1).
   - This bias is suspected to be a result of the different external forcing presented to the ocean/sea ice system during the assimilation period. In particular, there appears to be too little freezing of ice from below.

2. The 31/03/2011 start date represented one of the smallest biases in the climatological September ice extent between model forecast and observations. It also represented a transition between a positive (too much ice during the hindcast) bias and a negative (too little ice during the hindcast) bias as compared with observations.

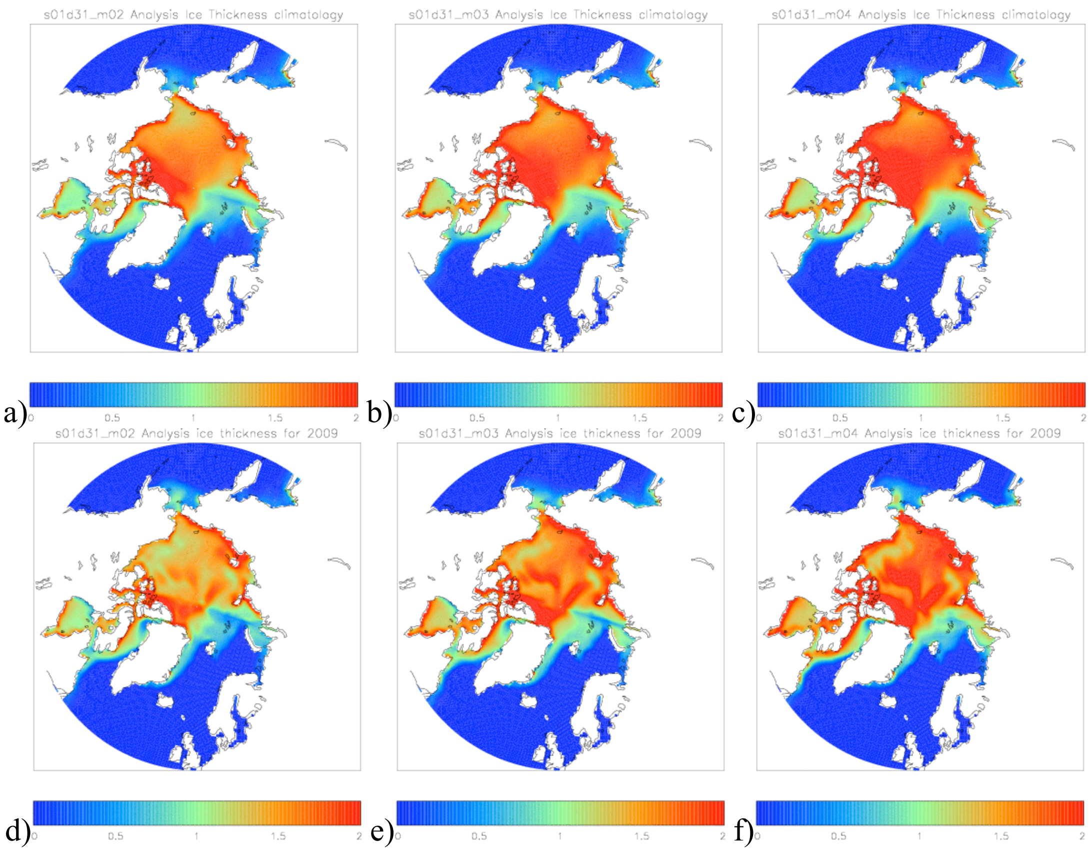


Figure 2. Sea ice thickness over the 1996-2009 period in the sea ice analysis for a) February, b) March, and c) April. Sea ice thickness in 2009 from the sea ice analysis for d) February, e) March, f) April. Again, these would be the approximate thicknesses that the coupled model would be initiated with during these months. Note that in both the climatology and 2009, the sea ice thickness increase during this period.

- ○ While this bias, which represents a systematic model bias, should not be equated with forecast skill – indeed later start dates do give better correlation of September ice extent with observations: Combined with the suspected thin sea ice bias of the 2011 Arctic sea ice in the forecast, there is a positive feedback towards decreasing Arctic sea ice extent. For reference, the forecast initiated on 12/05 gave an ice extent of only 2.7±0.3 million square kilometres, even after bias correcting this ice extent.
- ○ During the hindcast analysis, the ice thickness appears to thicken considerably between February and April (figure 2), which is not observed in the forecast analysis (Figure 1). Since the coupled model bias is to melt too much ice, this lack of increasing thickness in the forecast (figure 1) produces a positive feedback leading to the smaller ice extents at later lead times.

3. Because of the small bias between observed and forecast climatology, it will be possible to show ice concentrations with little known systematic bias. **This however,** should not be equated with accurate forecast skill, as the climatological bias **only** corrects for systematic model error, and **does not** correct forecast error resulting from non-linearly evolving model and initialization error.

4. For our particular hindcast set, the skill of persistence also degraded for later start dates. By persistence, we refer to the use of the sea ice extent anomalies from the start date as a proxy for September ice extent anomalies. For instance, the correlation between Febuary/March/April sea ice extent anomalies and September sea ice extent anomalies is 0.73/0.63/0.60 respectively. Much of the skill of persistence seems to relate to the linear (negative) trend in the sea ice extent during the hindcast period, which is probably more dominant against interannual variability during the sea ice minimum and maximum. Nevertheless, there is some indication of enhanced initialization skill for the earlier start dates, which must be weighed against (the considerable) forecast error at longer lead times.

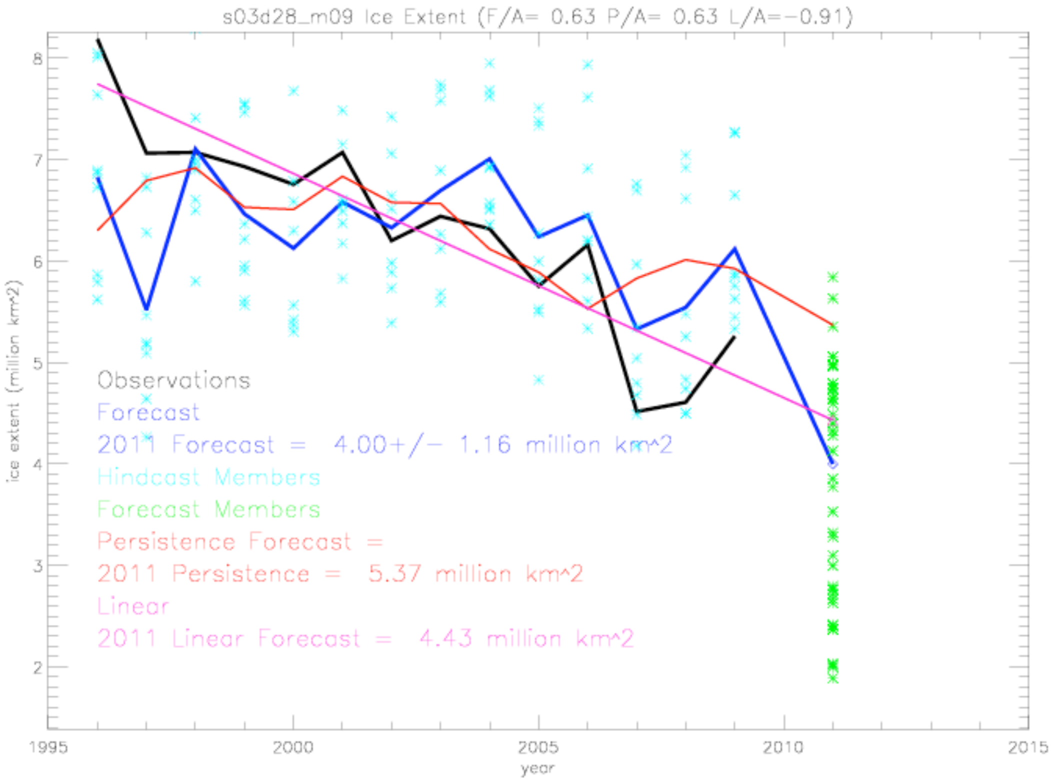


Figure 3: Time series of ice extents for the 1996-2009 hindcast, plus the 2011 forecast. The black line is the observations, and the blue line is the forecast values. The cyan and green x's are the hindcast and forecast ensemble members respectively. Also added to the graph is the persistence forecast (March anomalies added to September climatology) in red and the linear trend in the observations in magenta.

With all these considerations in mind, Figure 3 is a plot of September sea ice extent for the hindcast period of 1996 through 2009 plus the forecast for September 2011 (blue line, ending in a blue diamond – the 2011 forecast). **Note** that there is no hindcast value for 2010. Also included on the plot are the observations in black, the persistence forecast (adding the March anomalies onto the September climatology) in red, and the linear trend in the observations (over the 1996-2009 period) in magenta. The correlation between the hindcast ice extents for September and observed ice extents was 0.63. This is significantly different from zero at the 94% confidence level, the number of effective degrees of freedom being lowered due to serial correlation (Zwiers and von Storch, 1995). For comparison, March persistence correlates with September ice extent again at 0.63, but this is significantly different from zero only at the 74% confidence level owing to a very high degree of serial correlation (trend) in the two time series. The detrended correlation between hindcast and observation is 0.73, while the detrended correlation between March persistence and observation is -0.24. The detrended correlation of the forecast with observations is non-zero at the 99% confidence level, and fairly obviously, the correlation between February persistence and September observations is completely related to the trend in both values.

Figure 4 shows a plot of forecast ice concentration. The thick black line in the plot represents the ice extent. The green line is the model ice extent climatology over the hindcast period of 1996-2009 and blue line is the observed ice extent climatology over the same period. The overall ice extent area is fairly well modelled in the hindcast with an observed ice extent of 6.3 million square kilometres versus a hindcast of 6.1 million square kilometres, for a bias of 0.2 million square kilometres. Nevertheless, there are regional differences in the climatological ice edge that one should take into account when viewing the ice concentration.

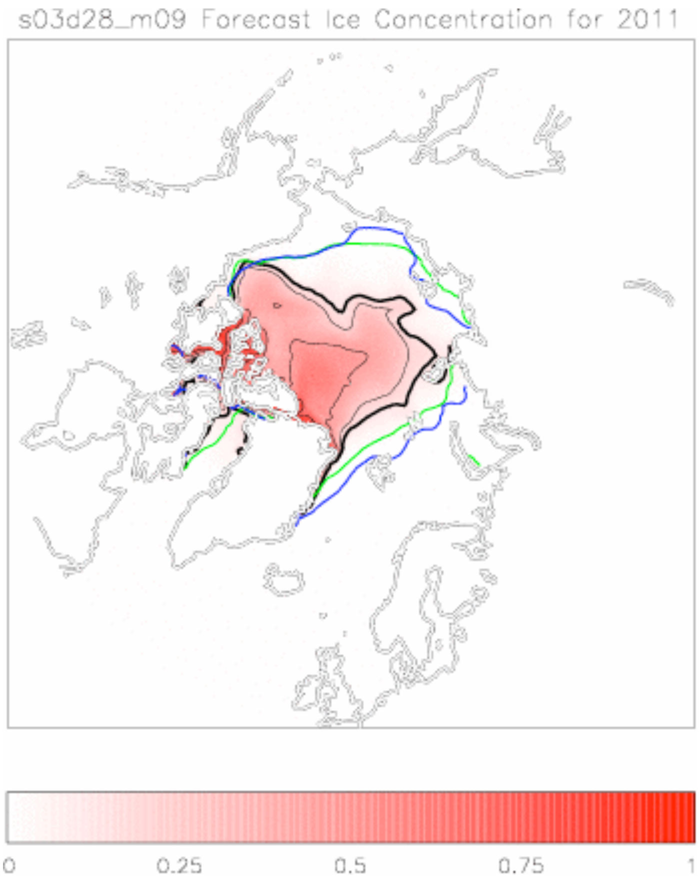


Figure 4. Forecast 2011 sea ice concentration. The thick black line is the forecast ice extent (ice concentration of 0.15) and the thinner black lines are contour intervals of 0.25. The green line is the model ice extent climatology during the hindcast period of 1996-2009, while the blue line is the observed climatological ice extent over the same period.

4) <u>Executive Summary</u>

Our 2011 September ice extent forecast is 4.0±1.2 million square kilometres. This is based upon a forecast from the GloSea4 seasonal forecast system using a coupled atmosphere/ocean/sea ice model initialized from observations between 21 March and 10 April. The quoted error is based on the standard deviation of the ensemble members' sea ice extent from their average value. A bias correction of 0.2 million square kilometres has been added to the forecast to account for the climatological bias of lower ice extents forecast in the model over the hindcast period. However, a further bias toward lower ice thicknesses in the actual forecast as compared to the hindcast initialization is also suspected. Therefore we suspect that our forecast may be biased towards a smaller ice extent from the ultimate reality. Furthermore, this bias appears to become even more exaggerated with later start dates, hindering our ability to update the forecast at a later time.

5) <u>Estimate of Forecast Skill</u>

GloSea4 is an ensemble forecast system. The spread of the forecast members allows us to place an uncertainty error on the forecast. The standard deviation of the spread of 2011 ensemble members is 1.2 million square kilometres. The skill of the forecast has been calibrated against a hindcast set done over the years 1996-2009. The correlation between the hindcast set and observations is 0.63, which is significantly different from zero at the 94% confidence level and is comparable with the persistence forecast (correlation of 0.63, but significantly different from zero only at the 74% confidence level) over the same period. The detrended correlation between observations and forecast is 0.73, which is significantly different from 0 at the 99% confidence level. It would appear that the hindcast is accurately representing the year to year variability, but does not get the observed trend in the ice extent **despite** that being represented in the initialization.

Due to the ensemble nature of the seasonal forecast system, we can also make probabilistic statements. 100% of the forecast ensemble members give a September ice extent below 6.2 million square kilometres, the threshold for the lowest tercile during the hindcast period, and 88% of the forecast ensemble members give a September ice extent below 5.3 million square kilometres, which represents the threshold for the lowest quintile in the hindcast period. Hindcast calibration of these probabilistic forecasts have proven skilful with relative operating characteristic (ROC; http://www.metoffice.gov.uk/research/areas/seasonal-to-decadal/gpc-outlooks/user-guide/interpret-roc) scores of 0.93 and 0.96 respectively. A ROC score of 0.5 and below represents skill no better or worse than climatology, while a ROC score of 1.0 represents perfect predictability. Thus there is a good (88%) probability that September ice extent should be below the 2009 minimum of 5.3 million square kilometres (which happen to be equivalent to the threshold for the lowest quintile in our hindcast set). This probability can be addressed with a fair amount of skill in that highly probable scores give a good ratio of correct forecasts to missed forecast (the basis of the ROC score).

Final Caveat: Unfortunately, there would appear to be an additional bias between the ice extent in the forecast versus the ice extents during the hindcast caused by differences in the forcing of the ocean and sea ice model during the assimilation. This in turn has probably led to the forecast model being initialized with too thin ice -- albeit with the correct ice concentration. Figure 5 shows a plot of forecast September ice thickness versus the hindcast climatology. A reminder that no constraint is put on the ice thickness by the initialization process, and therefore we have no expectation that these thicknesses represent the thickness observations in any meaningful way. As a better comparison of the 2011 ice thickness with the ice thickness in the hindcast in the more recent past, figure 6 shows a plot of forecast September ice thickness for the years 2007, 2008 and 2009. Given the systematic model error, we would expect neither the hindcast climatology nor the more recent year hindcasts to represent true thickness observations – and indeed we encourage readers to comment on our ability to accurately model thickness. However, even less confidence is entailed in the 2011 forecast. Since the GloSea4 sea ice forecast is an experimental forecast in which the current system was only implemented last September, this gives us no basis beyond the hindcast on which to base the forecast. Possible biases between the forecast and the hindcast, which appear likely to exist, cannot be quantified in any meaningful way. We hope to address some of these issues by applying both hindcast and forecast methods to the 2010 sea ice extent – if resources allow. However, what is ultimately required is a real time estimate of sea ice thickness which could then be incorporated into the assimilation system and better account for errors in the forcing of the sea ice.

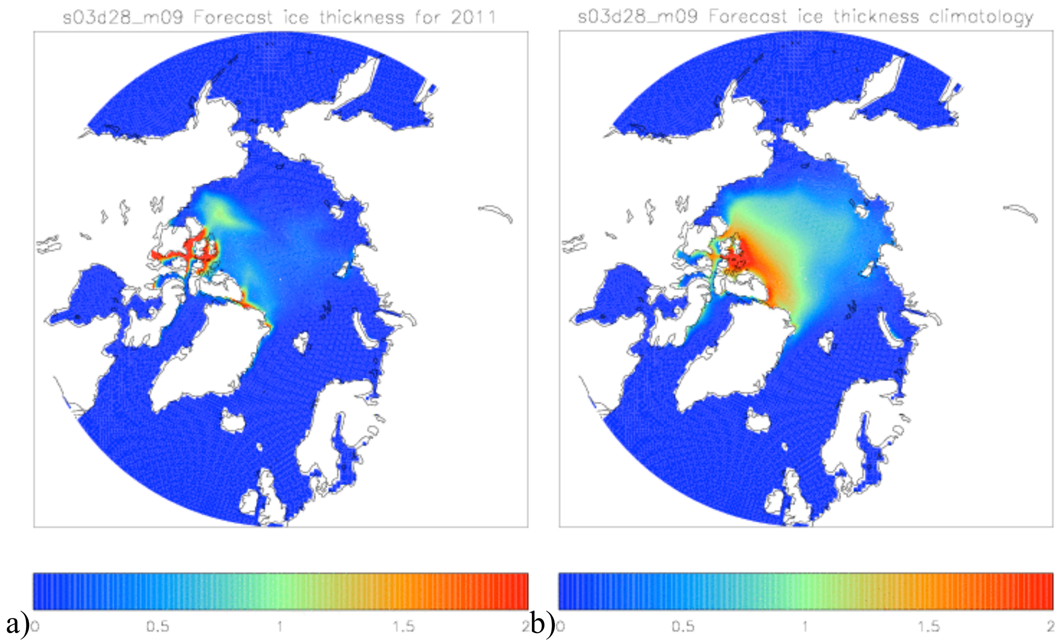
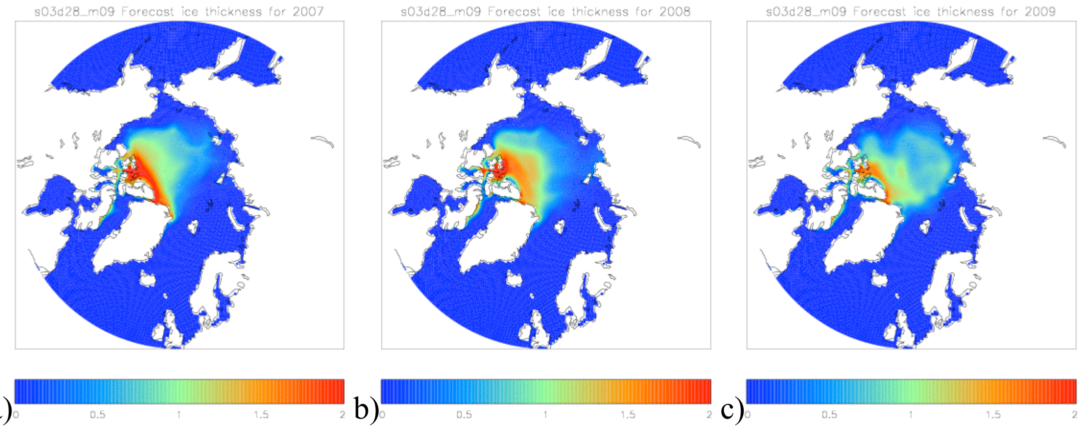


Figure 5. a) Forecast September 2011 ice thickness. b) Hindcast climatological ice thickness.



a) b) c)
Figure 6. Forecast September ice thickness for a) 2007, b) 2008, and c) 2009.

As it stands, it is likely that are forecast ice extent will be too small, and our estimates of forecast probabilities for low ice extents too high. This bias is further exaggerated for later start dates, making July and August updates to the outlook unlikely, or at the very least, hard to interpret.

References

Arribas , A., Glover, M., Maidens, A., Peterson, K., Gordon, M., MacLachlan, C., Graham, R., Fereday, D., Camp, J., Scaife, A.A., Xavier, P., McLean, P., Colman, A., and Cusack, S, 2011: The GloSea4 ensemble prediction system for seasonal forecasting, *MWR*, DOI: 10.1175/2011MWR3615.1. To be published in June 2011 issue *MWR*, **139**(6), pp. 1891-1910.

Bowler, N.E., A. Arribas, S. Beare, K. Mylne and G. Shutts, 2009: The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system, *Quart. Jour. Roy. Met. Soc.,* **135**, pp. 767-776

Brodeau L, Barnier B, Treguier AM, Penduff T, and Gulev S, 2010: An ERA40-based atmospheric forcing for global ocean circulation models, Ocean Modelling **31**(3-4), pp. 88-104.

Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A. and Wood, N., 2005: A new dynamical core for the Met Office's global and regional modelling of the atmosphere, *Quart. J. Roy. Meteor. Soc.,* **131**, pp. 1759-1782.

Dee, Dick, P. Berrisford, P. Poli, M. Fuentes, 2009: ERA-Interim for climate monitoring. ECMWF Newsletter, **119**, pp. 5-6.

Essery, R.L.H., M.J. Best, R.A. Betts, P.M. Cox, and C.M. Taylor, 2003: Explicit representation of subgrid heterogeneity in a GCM land-surface scheme. *J. Hydrometeor*., **4**, pp. 530–543.

Hewitt, H. T., D. Copsey, I. D. Culverwell, C. M. Harris, R. S. R. Hill, A. B. Keen, A. J. McLaren and E. C. Hunke, 2011: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, *Geosci. Model Dev.,* **4**, pp. 223-253, doi:10.5194/gmd-4-223-2011.

Hunke, E.C., and W.H. Lipscomb, 2010: CICE: The Los Alamos sea ice model documentation and software user's manual, version 4.1. LA-CC-06-012, Los Alamos National Lab, 75 pp.

Large, W.G., and S.G. Yeager, 2009: The global climatology of an interannually varying air-sea flux data set. *Clim. Dyn.* **33**(2-3), pp. 341-364, doi: 10.1007/s00382-008-0441-3.

Madec, G., 2008: NEMO ocean engine, Note du Pole de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27, ISSN No 1288-1619, 2008.

Peterson, K. Andrew, A. Arribas, A. McLaren, H. Hewitt and M. Gordon, 2011: Skill of Arctic Sea Ice Prediction in GloSea4 Seasonal Forecast System, *Journal of Climate,* in preparation.

Rawlins, F., S.P. Ballard, K.J. Bovis, A.M. Clayton, D. Li, G.W. Inverarity, A.C. Lorenc and T.J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quart. Jour. Roy. Met. Soc.,* **133**: pp. 347–362. doi: 10.1002/qj.32.

Stark, J. D., J. Ridley, M. Martin, and A. Hines, 2008. Sea ice concentration and motion assimilation in a sea ice−ocean model. *J. Geophys. Res.,* **113**, C05S91, doi:10.1029/2007JC004224.

Storkey, D., E.W. Blockley, R. Furner, C. Guiavarc'h, D. Lea, M.J. Martin, R.M. Barciela, A. Hines, P. Hyder, and J.R. Siddorn, 2010: *Journal of Operational Oceanography*, **3**(1), pp. 3-15.

Zwiers, F. W., and H. von Storch, 1995: Taking serial correlation into account in tests of the mean. *J. Climate* **8**, 336-351.