

## Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system

C. MacLachlan,<sup>a\*</sup> A. Arribas,<sup>a</sup> K. A. Peterson,<sup>a</sup> A. Maidens,<sup>a</sup> D. Fereday,<sup>a</sup> A. A. Scaife,<sup>a</sup> M. Gordon,<sup>a</sup>  
M. Vellinga,<sup>a</sup> A. Williams,<sup>a</sup> R. E. Comer,<sup>a</sup> J. Camp,<sup>a</sup> P. Xavier<sup>a</sup> and G. Madec<sup>b</sup>

<sup>a</sup>Met Office Hadley Centre, Exeter, UK

<sup>b</sup>French National Centre for Scientific Research, LOCEAN-IPSL, Paris, France

\*Correspondence to: C. MacLachlan, Met Office Hadley Centre, Fitzroy Road, Exeter, UK, E-mail:  
craig.maclachlan@metoffice.gov.uk

This article is published with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

This article describes the UK Met Office Global Seasonal forecast system version 5 (GloSea5). GloSea5 upgrades include an increase in horizontal resolution in the atmosphere (N216–0.7°) and the ocean (0.25°), and implementation of a 3D-Var assimilation system for ocean and sea-ice conditions. GloSea5 shows improved year-to-year predictions of the major modes of variability. In the Tropics, predictions of the El Niño–Southern Oscillation are improved with reduced errors in the West Pacific. In the Extratropics, GloSea5 shows unprecedented levels of forecast skill and reliability for both the North Atlantic Oscillation and the Arctic Oscillation. We also find useful levels of skill for the western North Pacific Subtropical High which largely determines summer precipitation over East Asia.

*Key Words:* seasonal forecasting; ensembles; Arctic Oscillation; ENSO; WNPSTH; MJO

Received 31 January 2014; Revised 7 May 2014; Accepted 8 May 2014; Published online in Wiley Online Library 27 June 2014

### 1. Introduction

As a result of a change of strategy aimed to improve the skill of long-range forecasts more rapidly, a new generation of the Met Office seasonal forecasting system was implemented in 2009 (Arribas *et al.*, 2011).

The central idea in our new strategy was to fully integrate the development of the Met Office coupled model and the seasonal forecasting system. This way we could rapidly implement all scientific upgrades into the forecasting system. To achieve this, various technical challenges need to be solved—for example, the forecast and hindcast components need to be run concurrently to avoid the need to complete all hindcast runs in advance—but our strategy has proved to be successful. While most centres around the world introduce improvements to their forecasting systems only every 5–10 years, we have been able to introduce major improvements every year since 2009 and this has been reflected in the skill of the forecasts.

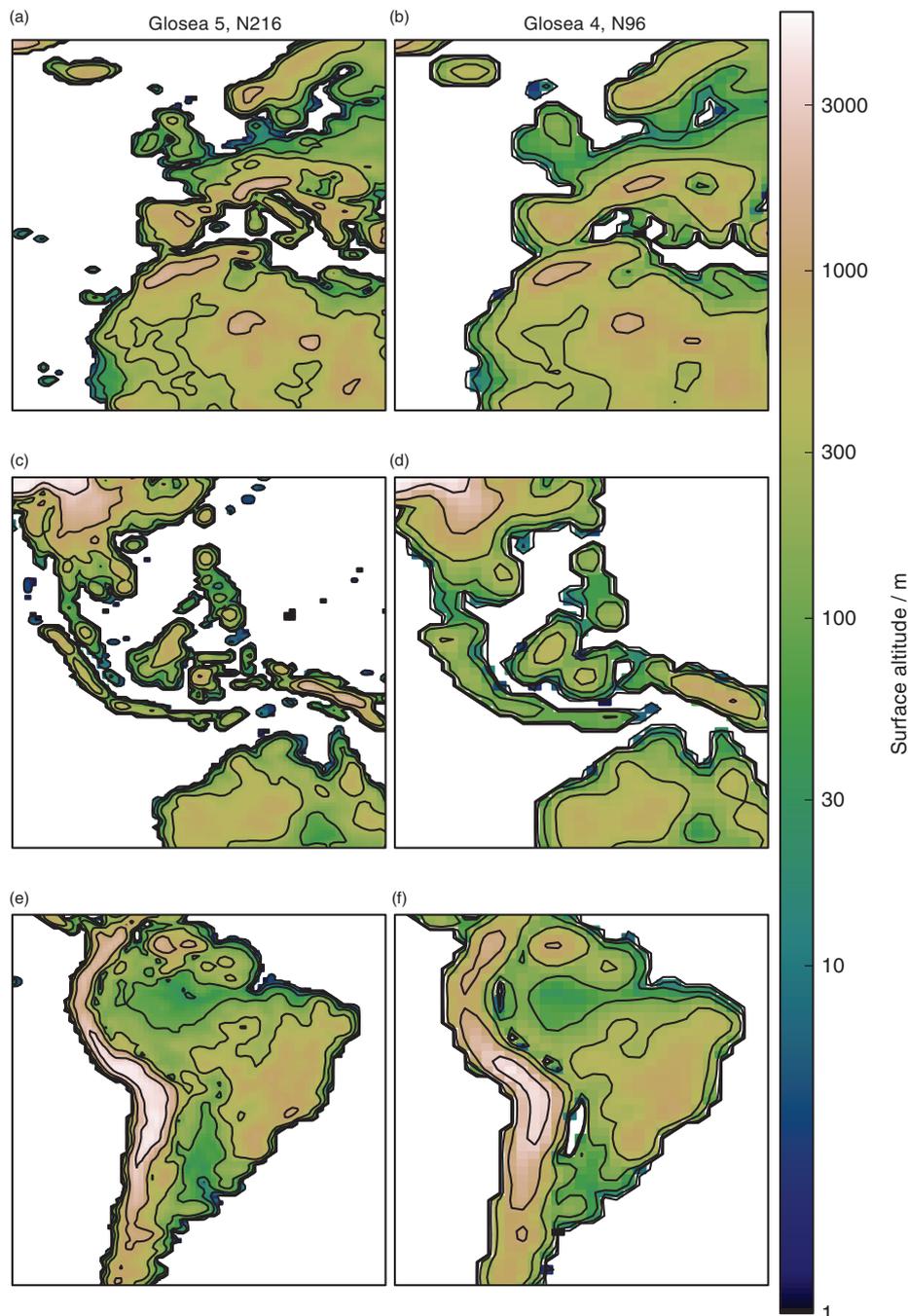
Between the implementation of GloSea4 in September 2009 and GloSea5 in January 2013, the system has had several significant upgrades in four packages: (i) fully resolved stratosphere, 3 hourly atmosphere–ocean coupling and assimilated sea-ice; (ii) daily initialisation of forecast members; (iii) an upgrade to the physical parametrizations in the model; and (iv) an increase in horizontal resolution. The frequency of these enhancements is reliant on the concurrent running of the hindcast.

In March 2011 the system was expanded to cover sub-seasonal time-scales, specifically the period from 2 to 6 weeks. To facilitate this change, the number of forecast ensemble members was increased and the member initialization was changed from weekly to daily. In November 2011 the scientific configuration of the model was upgraded to the Met Office Unified Model (MetUM) Global Atmosphere 3.0 (Walters *et al.*, 2011).

In January 2013 a new version of the seasonal forecast system was implemented operationally. This system, named GloSea5, features a higher horizontal resolution model than GloSea4. Each of the component model grids have been refined. The atmosphere and land surface resolution has been increased from 1.875°×1.25° to 0.833°×0.556°. The grid spacing in the ocean and sea-ice models has been reduced from 1° to 0.25°.

The ocean and sea-ice in the model are now initialised using the NEMO (Nucleus for European Modelling of the Ocean) three-dimensional variational ocean data assimilation based on the multi-institution NEMOVAR project (Mogensen *et al.*, 2009, 2012). Previously an analysis correction method was used. The new data assimilation system uses the same resolution and physical parametrizations as the ocean model in GloSea5 and includes sea-ice assimilation.

Increasing the horizontal resolution involves a factor ~10 increase in computing cost but improves the performance of the coupled model. The model used in GloSea4 exhibited a cold sea-surface temperature bias in the Northwest Atlantic.



**Figure 1.** Contour plots showing the (a, c, e) GloSea5 N216 and (b, d, f) GloSea4 N96 orography over the (a, b) European, (c, d) Maritime Continent, and (e, f) South American regions. Contours are at 1, 10, 30, 100, 300, 1000 and 3000 m.

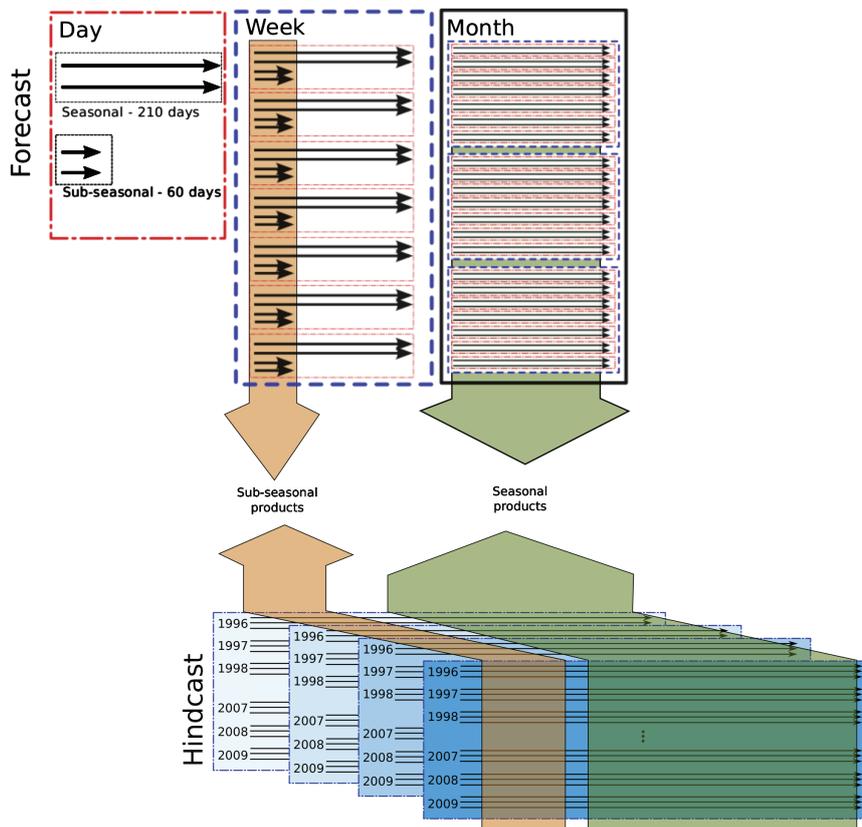
Similar biases have been reported in other models with similar resolutions (Smith *et al.*, 2000; Danabasoglu *et al.*, 2010). This bias is caused by a poor representation of the location of the North Atlantic Current. Increasing the resolution in the ocean model greatly improves the path of the North Atlantic Current, removing the cold bias. Removing this bias improves the frequency of blocking events in Northern Europe (Scaife *et al.*, 2011). Increased atmospheric model horizontal resolution, specifically the orography, also improves the occurrence of blocking (Berckmans *et al.*, 2013).

At coarser ocean resolutions than  $1/3^\circ$ , tropical instability waves (TIWs) are poorly resolved (Roberts *et al.*, 2009). With the higher-resolution ocean model, TIWs are better resolved (Graham, 2014), improving the spatial representation of sea-surface temperature anomalies in the Tropical Pacific. Increases in both the atmospheric and oceanic resolutions ensure the greatest improvement in El Niño teleconnections (Dawson *et al.*, 2012).

The description of GloSea5 and analysis of its performance will be the focus of this article, particularly the performance at the seasonal time-scale. In section 2, we describe the GloSea5 system and the improvements introduced above in more detail. The representation of specific phenomena—the El Niño Southern Oscillation (ENSO), the North Atlantic Oscillation (NAO), the Arctic Oscillation (AO), the western North Pacific Subtropical High (WNPSH), the Madden–Julian Oscillation (MJO), and tropical storms—is investigated in section 3.1. Standard verification scores for seasonal prediction are presented in section 3.2. Finally, we summarize the results and discuss future plans for the GloSea5 system in section 4.

## 2. Description of upgrades to the model and the ensemble prediction system

In this section we describe the main technical features of the Met Office seasonal forecast system: the model configuration, how



**Figure 2.** Schematic showing the difference between weekly and daily initialization and the additional members used for the sub-seasonal forecast. The diagram shows how the four forecast members initialized each day are combined in a lagged ensemble. Sub-seasonal products are generated from 7 days of forecast members. Seasonal products use 3 weeks of forecast members in the ensemble. Each week a hindcast set for a given initialization date is completed. The same hindcast is used to bias correct both seasonal and sub-seasonal products.

it is initialised, and the construction of the ensemble used to generate products issued by the Met Office. The previous system was described in Arribas *et al.* (2011) and many of the details are still relevant.

## 2.1. Model configuration

The coupled HadGEM3 model used in the seasonal forecast system consists of the following components:

- Atmosphere: MetUM (Walters *et al.*, 2011; Brown *et al.*, 2012), Global Atmosphere 3.0
- Land surface: Joint UK Land Environment Simulator (JULES; Best *et al.*, 2011), Global Land 3.0
- Ocean: NEMO (Madec, 2008), Global Ocean 3.0
- Sea-ice: The Los Alamos Sea Ice Model (CICE; Hunke and Lipscomb, 2010), Global Sea-Ice 3.0

The dynamical core of the UM (called NewDynamics) uses a semi-implicit semi-Lagrangian discretization to solve the fully compressible, non-hydrostatic atmospheric equations of motion. The stochastic physics scheme Stochastic Kinetic Energy Backscatter v2 (SKEB2; Bowler *et al.*, 2009) is included to represent unresolved processes and provide small grid-level perturbations during the model integration. Climate forcings (e.g. methane, CO<sub>2</sub>, etc.) are set to observed values up to the year 2005; after this point the emissions follow the Intergovernmental Panel on Climate Change (IPCC) RCP4.5 scenario. Climatologies with a seasonal variation are used for other aerosols (biogenic aerosols, biomass burning, black-carbon, sea salt, sulphates, dust, and organic carbon fossil fuels). These climatologies have been generated from a climate simulation using HadGEM2 (except dust which is from a HadGEM1a run). The Stratosphere–troposphere Processes And their Role in Climate (SPARC; Cionni *et al.*, 2011) observational climatology is used for ozone, which includes a

seasonal cycle. The solar forcing is the same in the forecast and hindcast, with an interannual variation.

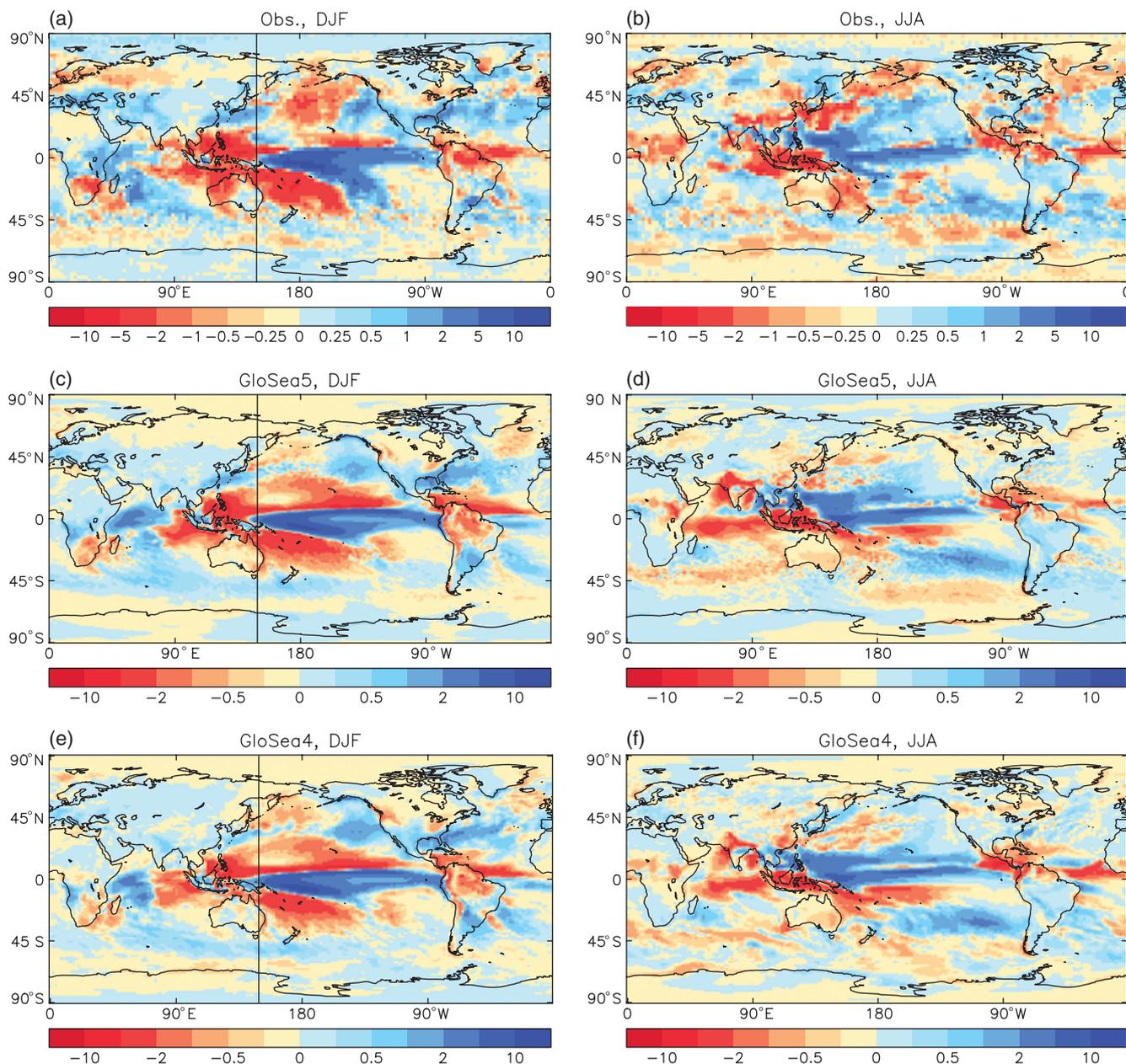
### 2.1.1. Global Atmosphere 3.0

A detailed description of the Global Atmosphere 3.0 configuration is given in Walters *et al.* (2011) where the developments between version 2.0 and 3.0 are also discussed. The basis of this science configuration has been adopted by all the operational global models used in the Met Office (although the configurations are not exactly the same due to unavoidable temporal and spatial resolution differences). There have been numerous changes to the physical parametrizations used in the coupled model since Global Atmosphere 2.0: introduction of cloud inhomogeneity, reduction of spurious drizzle, reduction of spurious deep convection, introduction of the JULES land surface model (Blyth *et al.*, 2006), and the facility to read iceberg calving ancillary data.

### 2.1.2. High-resolution model

The higher-resolution version of HadGEM3 used in the GloSea5 system uses the Global Atmosphere 3.0 configuration. Most of the physical parametrizations remain the same between the two resolutions. The high-resolution model requires a reduced time step and altered diffusion settings to increase stability. In the ocean model with the ORCA 0.25 grid, some of the major closed seas (Great Lakes, Lake Victoria, Caspian Sea and the Aral Sea) are included.

The resolution of the HadGEM3 model used in GloSea4 was N96L85 ORCA 1 L75; in GloSea5 it has been increased to N216L85 ORCA 0.25 L75. This means that the horizontal resolution in the atmosphere has increased from  $1.88^\circ \times 1.25^\circ$  to  $0.83^\circ \times 0.56^\circ$  (i.e. approximately 120 km in midlatitudes to 50 km). Figure 1 compares the orography used in the GloSea4 and GloSea5



**Figure 3.** Precipitation difference ( $\text{mm day}^{-1}$ ) between El Niño and La Niña years for (a, b) observations, (c, d) GloSea5 N216, and (e, f) GloSea4 N96 in (a, c, e) DJF and (b, d, f) JJA. For the DJF plots, a vertical line corresponding to the westward tip of the observed wet anomaly has been added to show the improvement between GloSea4 and GloSea5. The DJF means are derived from the November starts and the JJA means are derived from the May starts. In (c, d, e, f), the pattern correlations with observations are 0.80, 0.61, 0.76 and 0.59 (all  $\pm 0.01$ ), respectively.

systems, illustrating the improvement in resolution. Both ocean resolutions use a tripolar grid with enhanced resolution in the polar regions. The lower resolution NEMO model has a grid spacing of  $1^\circ$ , except between  $20^\circ\text{S}$  and  $20^\circ\text{N}$  where the grid was refined to  $0.33^\circ$ . This refinement in resolution was included to give a better representation of tropical waves and is a standard part of the ORCA 1 configuration. In the high-resolution model, the ocean has a grid spacing of  $0.25^\circ$  globally with no refinement in the equatorial region. This is equivalent to 27 km on the Equator.

## 2.2. Initialization

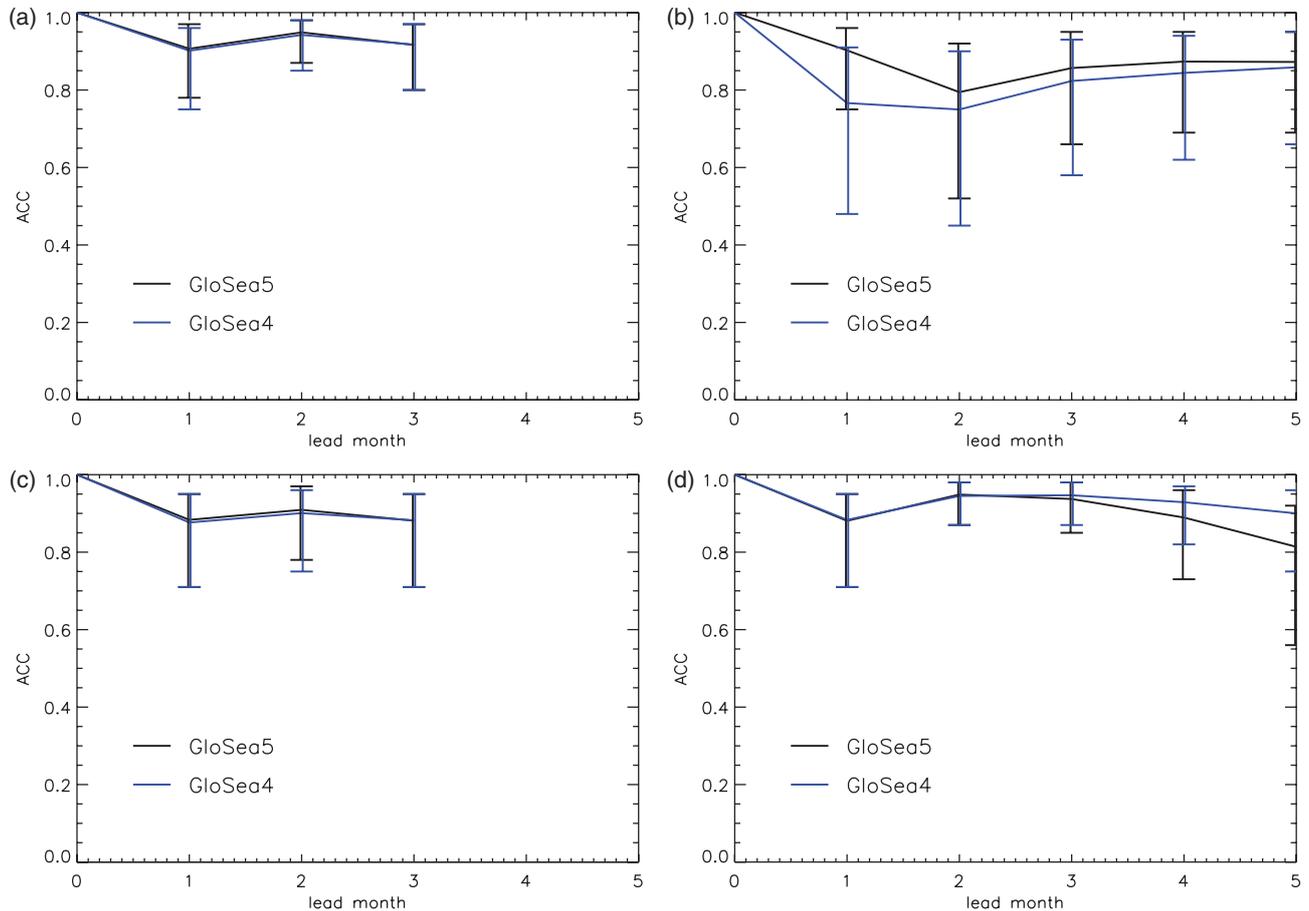
### 2.2.1. Daily forecast initialization

When GloSea4 was first implemented, members were initialised once a week (on Monday) and three consecutive weeks of forecast members would be combined in a lagged ensemble. Each week, 14 forecasts were completed with the HadGEM3 coupled model; the

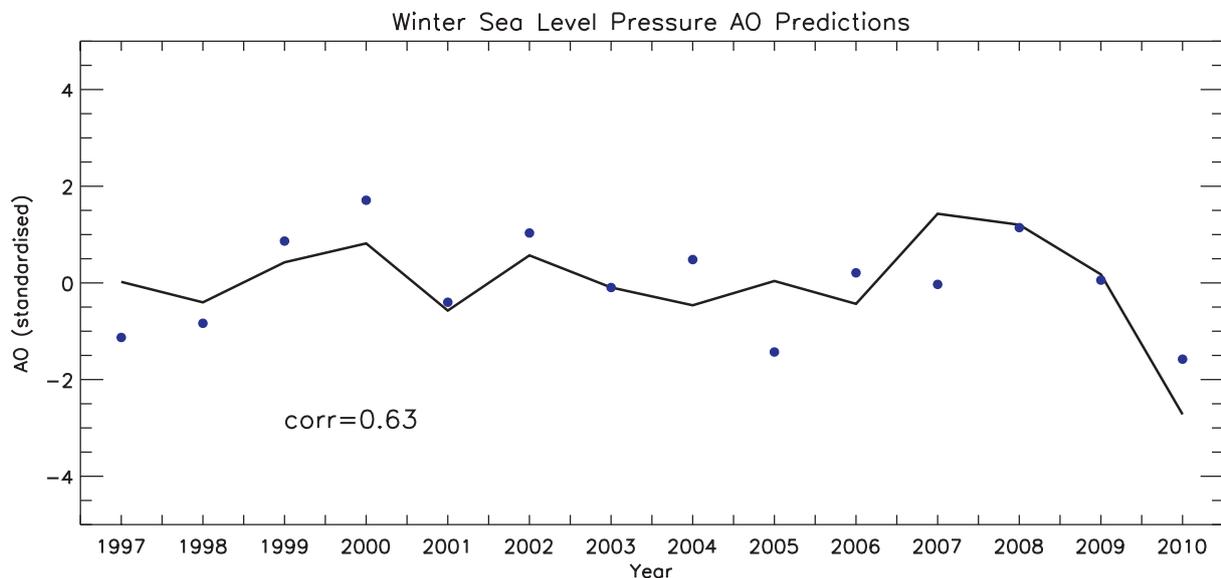
members were initialized from analyses valid for 0000 UTC on the Monday of that week. A stochastic physics scheme SKEB2 (Bowler *et al.*, 2009) was used to generate spread between members initialized from the same analysis. For operational load balancing, these members were spread across the computational week, two forecast members being computed each day. Since March 2011 the forecast members have been initialized using analyses valid for that day. Figure 2 shows the daily initialization schematically.

The initial atmospheric conditions for the forecast members are generated by the Met Office operational numerical weather prediction (NWP) 4D-Var data assimilation system (Rawlins *et al.*, 2007). The ocean and sea-ice initial conditions are taken from the short-range ocean forecasting and data assimilation system. An overview of the ocean and sea-ice data assimilation system is given in section 2.2.2.

Reanalyses from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim project are used to initialize the atmosphere and land surface in the hindcast members. The soil moisture in both the forecast and hindcast



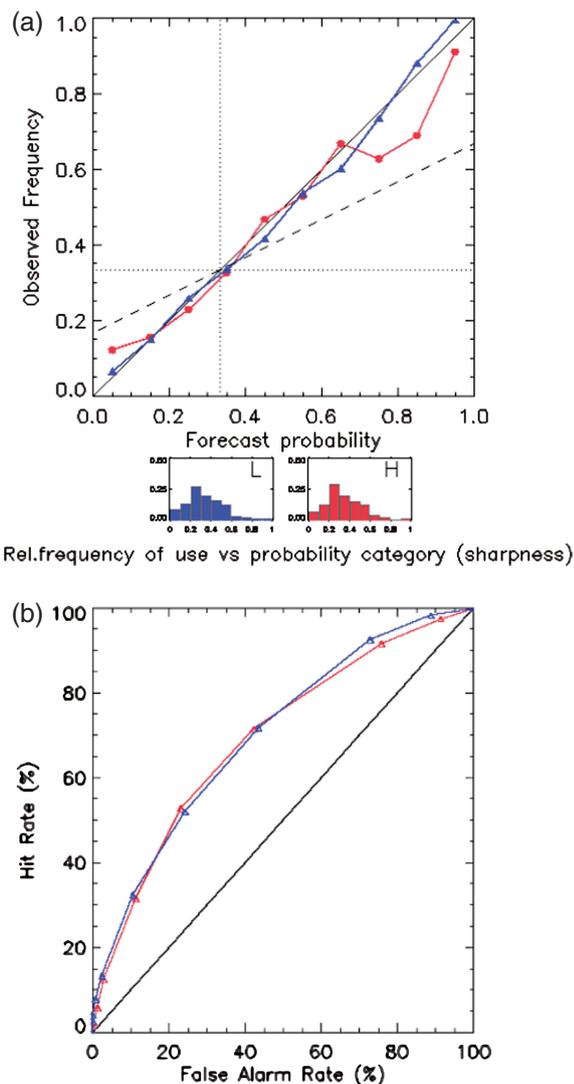
**Figure 4.** Anomaly correlation coefficients with observations in the Niño 3.4 region for GloSea5 and GloSea4 for varying lead times for starts in (a) February, (b) May, (c) August and (d) November. The error bars show the 95% confidence level for each correlation. The error bars have been computed using the Fisher transformation (Snedecor and Cochran, 1989) and displaced slightly horizontally for visual clarity. The May and November starts have been run out to a 5-month lead time, but those in the February and August sets only to a 3-month lead time.



**Figure 5.** The winter (DJF) Arctic Oscillation from observations (black line) and ensemble mean hindcasts (points) for the winters 1996/1997–2009/2010. The ensemble mean and observations are standardized by dividing by their respective standard deviations; the ensemble means standard deviation is 1.3 hPa, and the observed and ensemble member standard deviations are 3.2 and 3.5 hPa respectively. The vertical axis shows units of interannual standard deviation. The anomalies are calculated from the November start date hindcast set.

are initialised from a land surface reanalysis using the JULES land surface model forced with the Integrated Project Water and Global Change (WATCH) Forcing Data methodology applied to ERA-Interim data (WFDEI; Weedon *et al.*, 2011). The forecast soil moisture is initialized from the interannually varying mean climatology of the reanalysis. The hindcast soil moisture uses the time series from the reanalysis.

The forecast soil moisture is not initialized from the NWP data assimilation as the soil moisture climatologies from the Met Office data assimilation and ERA-Interim are significantly different. When initializing the forecast with this discrepancy, the bias correction would introduce an error in the surface temperatures which are closely linked to the soil moisture. Until a consistent way of initialising the hindcast and forecast is found,



**Figure 6.** Statistical scores for mean sea-level pressure in GloSea5 for the Northern Atlantic region for the winter season (DJF). (a) Reliability diagram. The red line shows the upper tercile and the blue line the lower tercile, the diagonal line perfect reliability, the dashed line is the line of no skill, and the dotted line gives climatology (Wilks, 1995). (b) Relative operating characteristics (ROC) curve. The red line shows the upper tercile and the blue line is the lower tercile.

the forecast soil moisture in GloSea5 will continue to be initialized with climatological values.

The system used to initialize the forecast ocean and sea-ice has been run with historical forcing to produce a reanalysis. It is used to initialize the ocean and sea-ice components in the hindcast. The system is described in the next section.

### 2.2.2. Ocean and sea-ice initialization

GloSea5 uses the Forecast Ocean Assimilation Model (FOAM) Ocean Analysis (Blockley *et al.*, 2013) to initialize the ocean and sea-ice components of the coupled forecast model. An equivalent product, the GloSea5 Ocean and Sea Ice Analysis, a 23 year (1989–2011) reanalysis, supplies initial conditions for the hindcasts. Both ocean analyses use the new NEMOVAR (Mogensen *et al.*, 2012; Waters *et al.*, 2014) assimilation scheme developed jointly by the UK Met Office, Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), ECMWF, and Institut National de Recherche en Informatique et en Automatique/Laboratoire Jean Kuntzmann (INRIA/LJK).

The assimilation system (based on the NEMOVAR scheme) used to create the forecast and hindcast analyses has the same ocean and sea-ice model (NEMO/CICE ORCA 0.25 L75) as the coupled model used in GloSea5. In the ocean–sea-ice data

assimilation system, the surface boundary forcing is calculated using the Coordinated Ocean Research Experiments (CORE) bulk formula formulation of Large and Yeager (2009). The Met Office NWP atmospheric analysis is used to force the FOAM analysis, and the ERA-Interim atmospheric analysis (Dee *et al.*, 2009) is used for the hindcast ocean reanalysis (GloSea5 Ocean and Sea Ice Analysis).

NEMOVAR is a multivariate, incremental 3D-Var first guess at appropriate time (FGAT) system. The system implemented at the Met Office operates on a daily cycle with a 1 day time window and uses an incremental analysis step. It assimilates both satellite and *in situ* observations of sea-surface temperature (SST), sea-level anomaly satellite data, sub-surface temperature and salinity profiles, and satellite observations of sea-ice concentration. The temperature, salinity and sea-level observations are assimilated in a multivariate fashion using balance relationships between the variables (hydrostatic and geostrophic balance, plus preservation of the temperature–salinity relationship in density), while the sea-ice concentration is assimilated as a univariate field. Furthermore, bias correction schemes are implemented to reduce the bias inherent in satellite measurements of SST (Martin *et al.*, 2007; Donlon *et al.*, 2012) and to reduce the bias in the (supplied) mean dynamic topography correction required to convert measurements of sea-level anomaly into sea-surface height (Lea *et al.*, 2008).

A detailed catalogue of the observations used in the real-time and reanalysis assimilation can be found in Appendix 4.

### 2.3. Ensemble prediction system

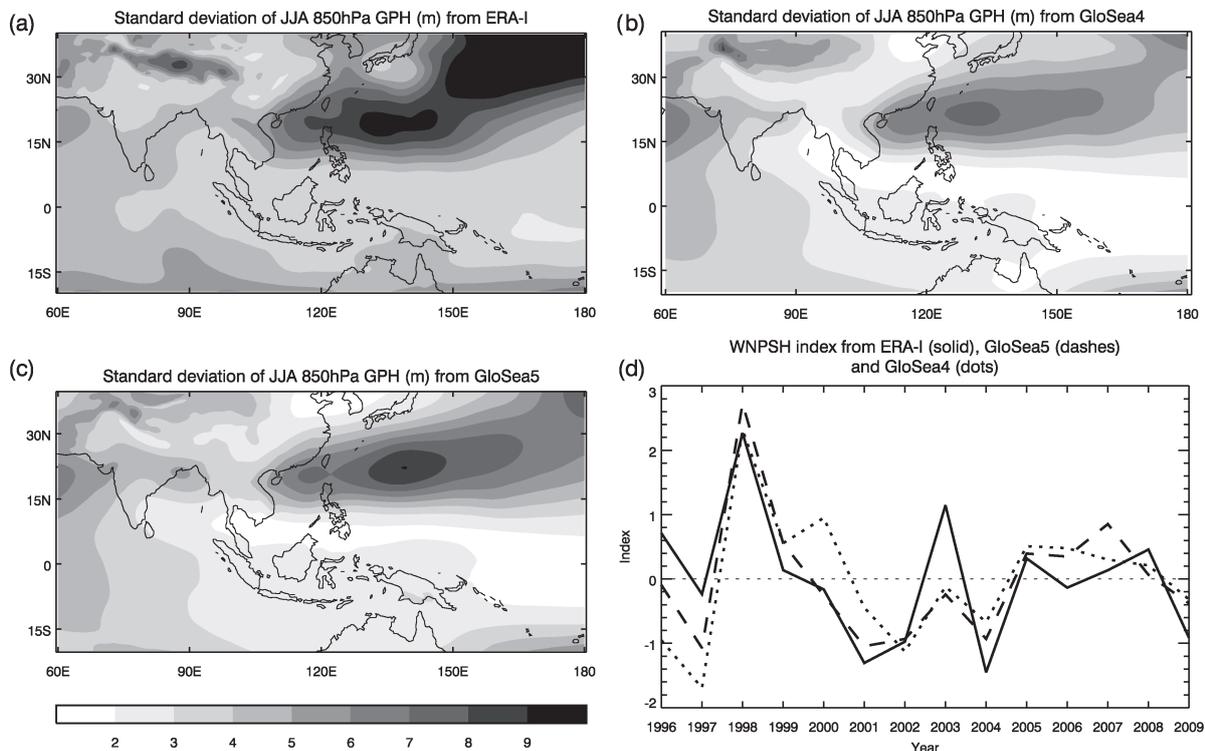
GloSea5 is a seamless monthly to seasonal forecast system comprising three parts: an intraseasonal forecast, a seasonal forecast, and a hindcast. Each day we complete four members initialized with 0000 UTC analyses from the NWP global data assimilation and the ocean–sea-ice data assimilation system. Two of these members are run out to 210 days (seasonal forecast members) and two are run out to 60 days (intraseasonal members). The first 60 days of all forecast members are used in the creation of intraseasonal products.

The hindcast set is used to bias correct all forecast members. Each week we complete a ‘hindcast week’ consisting of 14 years with three members per start date per year. The hindcasts are initialized from the following set of start dates: (1, 9, 17, and 25 in each calendar month). The same model is used for both hindcast and forecast members. Spread between members initialized on the same date is achieved through the use of a stochastic physics scheme. The number of years in the operational hindcast is sufficient for establishing the model climatology for bias correction (Arribas *et al.*, 2011).

The bias correction of the forecast members, including the weighting strategy, use the method described in Arribas *et al.* (2011). Previously all forecast members initialized in a week could be bias corrected with the same set of hindcasts, but now each forecast initialization date is bias corrected using the relevant hindcast set. The hindcast set is constructed from the four closest hindcast weeks. For example, the hindcast set corresponding to the forecast for the 13 April is: 1, 9, 17 and 25 April. Each hindcast week is weighted with the function  $w = e^{-d^2/100}$  where  $d$  is the lag/lead in days between the forecast and hindcast start dates.

For subseasonal forecasts, the bias corrected seasonal and subseasonal members from the last seven days are combined to form a lagged ensemble containing 28 forecast members. The subseasonal forecast products are produced on a daily basis and are used in the Met Office’s operational monthly forecasts.

Seasonal forecast members from the previous three weeks are combined, resulting in a 42-member ensemble for the next six months. These products are updated on a weekly basis. Generating the forecast on a weekly cycle means that frequent updates can be



**Figure 7.** Interannual variability of JJA 850 hPa geopotential height for 1996–2009 from (a) ERA-Interim, (b) GloSea4 and (c) GloSea5. (d) shows a comparison of the WNP5H index from ERA-Interim (solid), GloSea5 (dashed) and GloSea4 (dotted).

given when required. A monthly update for the UK Government is freely available on the Met Office website.\*

As a designated Global Producing Centre by the World Meteorological Organisation, skill scores and real-time forecasts from GloSea5 are freely available (updated monthly) through the Met Office website.†

### 3. System performance and evaluation

To evaluate the performance of the GloSea5 system a set of hindcasts have been completed. These are similar to the operational hindcasts which are run in real time. A subset of the operational hindcast initialization dates were chosen:

- 25 January, 1 February, 9 February ('February start'); covering March–April–May (MAM).
- 25 April, 1 May, 9 May ('May start'); covering June–July–August (JJA).
- 25 July, 1 August, 9 August ('August start'); covering September–October–November (SON).
- 25 October, 1 November, 9 November ('November start'); covering December–January–February (DJF).

There are four ensemble members for each individual date in the 'February' and 'August' sets and eight ensemble members for each of the initialisation dates in the 'May' and 'November' sets. Where possible, the results from the GloSea5 system are compared to GloSea4. Three members per initialization date are available for the GloSea4 hindcast set. Additional members are included in the May and November sets as these cover the key summer and winter seasons for our seasonal forecasts.

For the purpose of these analyses, we have expanded the hindcast set for the GloSea5 system to more accurately represent the skill of the GloSea5 seasonal forecast system. The operational set

of hindcasts (14 years, three members per year, and four start dates) is enough to establish a model climatology for bias correction of the forecast members (Arribas *et al.*, 2011). In the real-time forecast, we use 42 forecast members (from the most recent three weeks). Therefore, when investigating system performance with a reduced number of ensemble members, we do not achieve an exact representation of operational forecast skill. We have increased the number of ensemble members here in order to get a closer estimation of the skill of the real-time forecast. There is a larger number of members in the summer and winter seasons as these are key periods for many of the extreme events we would like to capture in operational forecasts. A subset of this hindcast data is available through the WMO World Climate Research Program Climate-system Historical Forecast Project (CHFP; <http://www.wcrp-climate.org/wgsip/chfp/>; accessed 25 May 2014).

#### 3.1. Synoptic performance

We first examine the leading patterns of interannual variability in the Tropics (ENSO) and Northern Hemisphere Extratropics (AO). Predictability of other tropical modes of variability such as the WNP5H and MJO are also examined. Finally, we examine the ability of GloSea5 to represent tropical storms, an important example of an extreme weather pattern.

##### 3.1.1. ENSO

The ENSO is the single largest source of interannual variability in the Tropics and has been known to show high levels of seasonal predictability (Barnston *et al.*, 2011). Thus, a good representation of ENSO and its teleconnections are required for regional climate prediction skill on seasonal time-scales. Figure 3 shows the average El Niño–La Niña precipitation differences in boreal winter (DJF) and summer (JJA) for GloSea4, GloSea5 and observations. GloSea4 overestimates the penetration of the ENSO anomalies into the West Pacific. This is a common coupled climate model error which appears in most seasonal forecast systems and coupled climate models (Luo *et al.*, 2005; Guilyardi *et al.*, 2009). The westward extension

\*Met Office three-month outlook; <http://www.metoffice.gov.uk/publicsector/contingency-planners>; accessed 25 May 2014.

†<http://www.metoffice.gov.uk/science/specialist/seasonal>; accessed 25 May 2014.

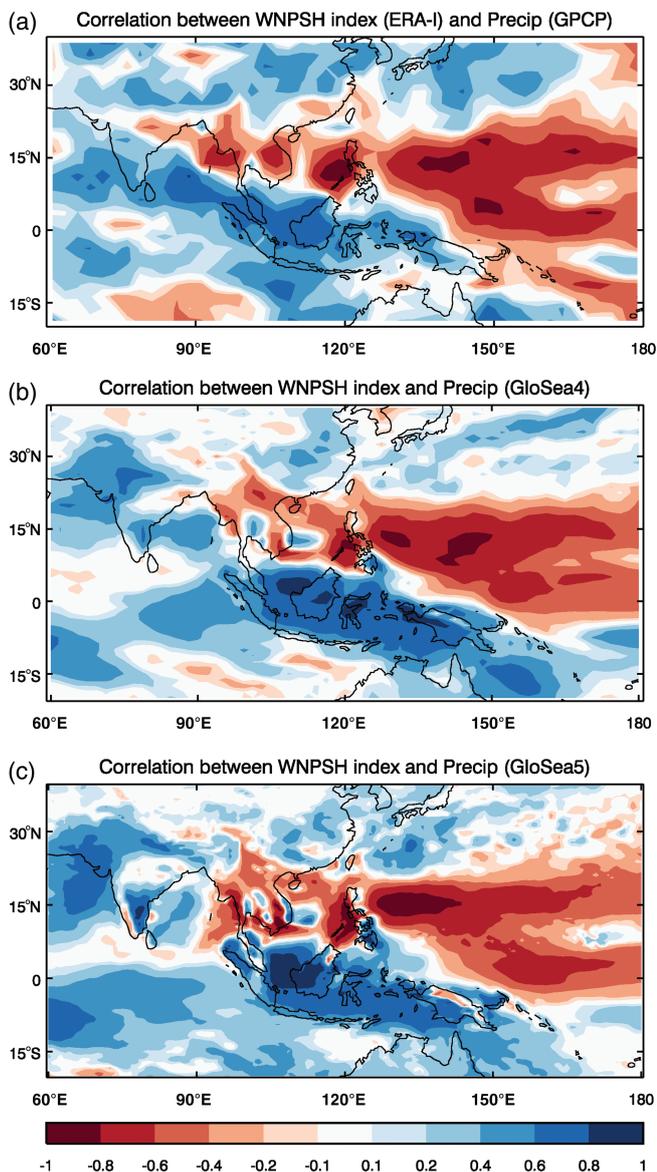


Figure 8. Correlation between WNP SH index and JJA precipitation from (a) observations, (b) GloSea4 and (c) GloSea5. For the forecast systems, the correlation is taken between the ensemble mean WNP SH index and precipitation from individual hindcast members.

of ENSO is much improved in GloSea5 compared to GloSea4 (Figure 3(a, c, e)). This results in an increase from 0.76 to 0.80 in the global pattern correlation of this teleconnection with the observed pattern. The pattern correlations are given for Figure 3(c–f)). The confidence intervals for significance at the 95% level are calculated using Fisher’s test and model results are regridded to the coarser observational grid. The improvement between GloSea4 and GloSea5 is statistically significant.

Figure 4 shows the lead time dependence of the anomaly correlation coefficient (ACC) for SST in the Niño 3.4 region (120°W–170°W, 5°S–5°N) with the HadISST observational dataset (Rayner *et al.*, 2006) for each of the hindcast start dates. The skill of GloSea5 for the ENSO forecast is very high (the ACC is over 0.8 for a 5 month lead time). The performance of GloSea5 is slightly higher than GloSea4, although this difference is not statistically significant.

3.1.2. Winter Arctic Oscillation

Outside the Tropics, the AO (Thompson and Wallace, 1998) is the single most important pattern of Northern Hemisphere interannual variability. It occurs throughout the seasons and is

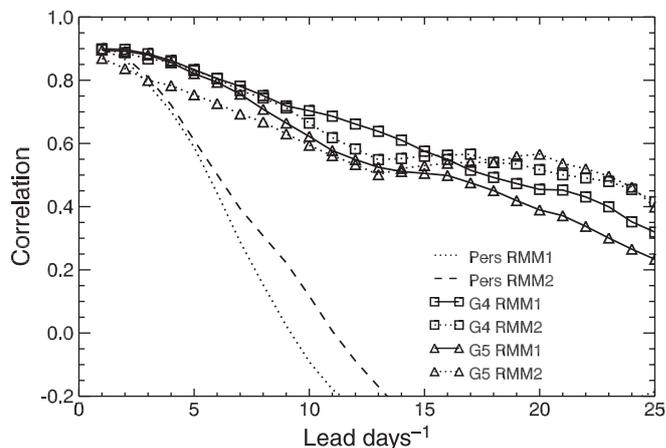


Figure 9. Correlation coefficients at different lead times of hindcast RMM1 and 2 (solid and dotted lines respectively) against observed RMMs. Open squares indicate the GloSea4 system and open triangles GloSea5. The correlations of persisted RMMs are also shown. All of the hindcast start dates have been used here.

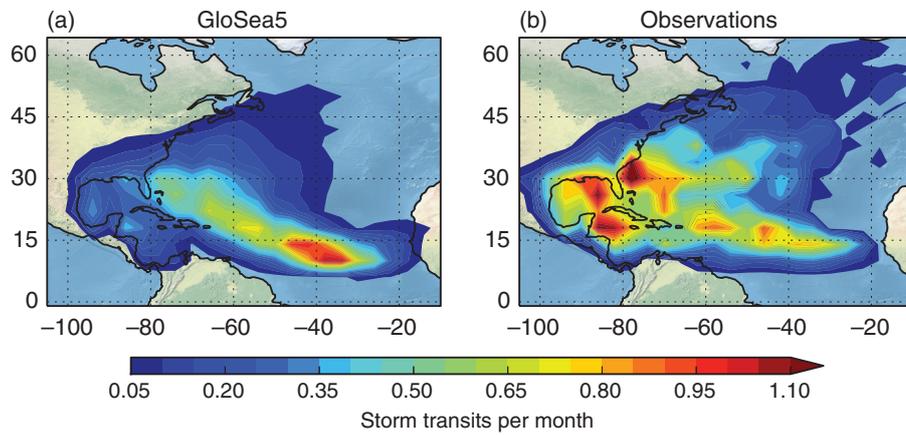
key to explaining year-to-year fluctuations in surface temperature, rainfall and circulation for many regions across the extratropical Northern Hemisphere. Predictability of the surface AO and its regional Atlantic equivalent the NAO on seasonal time-scales has until now been very low in dynamical forecast systems, with leading seasonal forecast centres reporting skill levels that are low and statistically indistinguishable from zero for winter forecasts starting in early November (Müller *et al.*, 2005; Johansson, 2007; Kim *et al.*, 2012). Forecasts starting at the onset of winter show slightly more encouraging skill levels, but this includes predictability from initial conditions which affect the first month (Derome *et al.*, 2005). However, observed relationships hint at higher potential levels of predictability (Cohen and Jones, 2011; Folland *et al.*, 2012).

Figure 5 shows that the AO in GloSea5 has highly significant levels of forecast skill for predictions starting in early November.

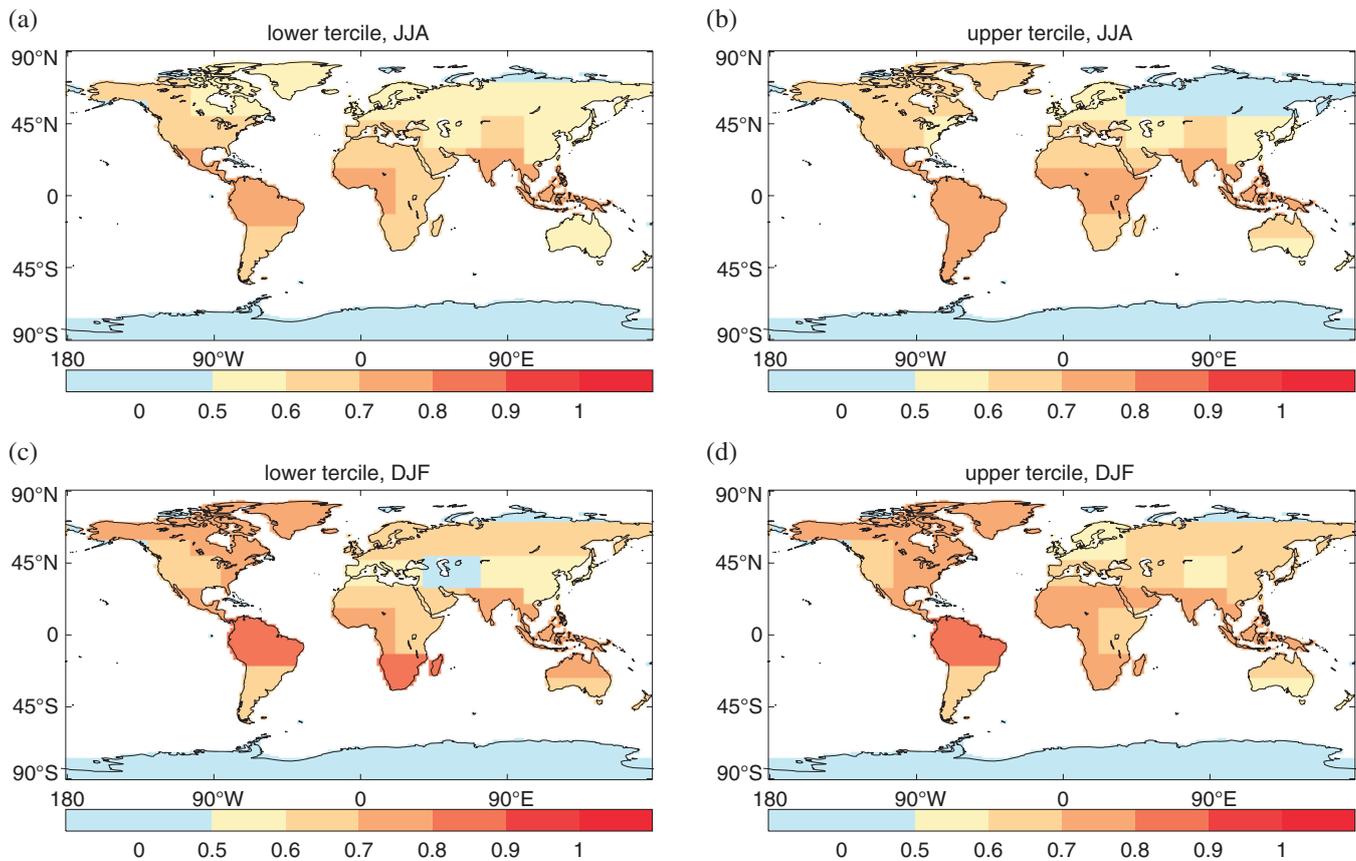
The correlation score between the ensemble mean AO and the observed AO is 0.63 in the hindcast which is significant at the 99% level. A similar level of predictability is found for the NAO (correlation 0.62; Scaife *et al.*, 2014) and is higher than the level of skill reported from other state-of-the-art forecast systems (Kim *et al.*, 2012). This also explains the increased levels of skill in winter surface temperature and rainfall in the extratropical regions in GloSea5 (section 3.2) and suggests that skilful seasonal forecast products can be made for the Extratropics from GloSea5. The increased skill of the AO in GloSea5 is also accompanied by good reliability for extratropical sea-level pressure predictions in GloSea5 (Figure 6(a)).

3.1.3. Western North Pacific Subtropical High

Improvements in skill are not confined to the Northern Atlantic region. We also see that the WNP SH is better represented in GloSea5 than in GloSea4. This feature is an important driver of Asian summer climate. Figure 7(a)–(c) compares the interannual variability of JJA 850 hPa geopotential height (GPH850) from the ERA-Interim reanalysis with that from the ‘May start’ in GloSea4 and GloSea5 hindcasts. The variability is larger in GloSea5 than in GloSea4, bringing it closer to that found in the observational reanalysis. The WNP SH index (Figure 7(d)) is defined as the normalized anomaly of JJA GPH850 over the region of maximum variability (115–150°E, 15–25°N; e.g. Wang *et al.*, 2013a). The correlation score for this index in GloSea5 is 0.8, compared with 0.61 for GloSea4. This improvement cannot be attributed to GloSea5’s larger ensemble size: random subsamples of the GloSea5 members with nine members per year produce a 98% confidence interval for the correlation score of (0.75, 0.84). The



**Figure 10.** Monthly mean tropical storm track density for (a) GloSea5 and (b) observations from HURDAT2 during the period June–November 1996–2009. Model results are normalised with respect to ensemble size.



**Figure 11.** Skill maps for 1.5 m temperature in GloSea5 for (a, b) JJA and (c, d) DJF: (a, c) ROC lower tercile and (b, d) ROC upper tercile.

sources of the WNPSH predictability in GloSea5 will be a subject of further research.

Summer rainfall over much of East Asia is strongly influenced by the WNPSH (e.g. Lu and Dong, 2001). Figure 8(a) shows the correlation between JJA 1996–2009 precipitation from the Global Precipitation Climatology Project (GPCP; Adler *et al.*, 2003) dataset and the ERA-Interim WNPSH index. Positive correlations over Eastern China are reproduced by both GloSea4 and GloSea5 (Figure 8(b, c)). These patterns, together with the improved WNPSH index skill in GloSea5, show promise for the development of seasonal rainfall predictions for this region and there are useful levels of skill for user-relevant variables such as precipitation over the Yangtze river basin (not shown).

### 3.1.4. Madden–Julian Oscillation

Intraseasonal variability in the Tropics is dominated by the MJO (Madden and Julian, 1994). It is characterized as

eastward-propagating, equatorially trapped, baroclinic oscillations in the tropical wind field and causes significant anomalies of convection and rainfall during its passage from the equatorial Indian Ocean to the Pacific. The MJO is known to be a major source of extended-range predictability in the Tropics (Jones *et al.*, 2000; Waliser *et al.*, 2003; Liess *et al.*, 2005). In recent years, the role of the MJO in the global climate system has been well recognised due to its interactions with other components of the climate system and because it represents an important link between the weather and seasonal-to-interannual climate variations. Recent studies suggest that the tropical dynamics associated with certain phases of MJO could influence the midlatitude weather regimes and hence contribute to improved predictability of NAO at extended range (Cassou, 2008; Lin *et al.*, 2009, 2010).

A comparison of the performance of the new GloSea5 system against the GloSea4 in forecasting the MJO indices is presented in Figure 9. The methodology of defining MJO indices and characterising the MJO phases was discussed in section 3c

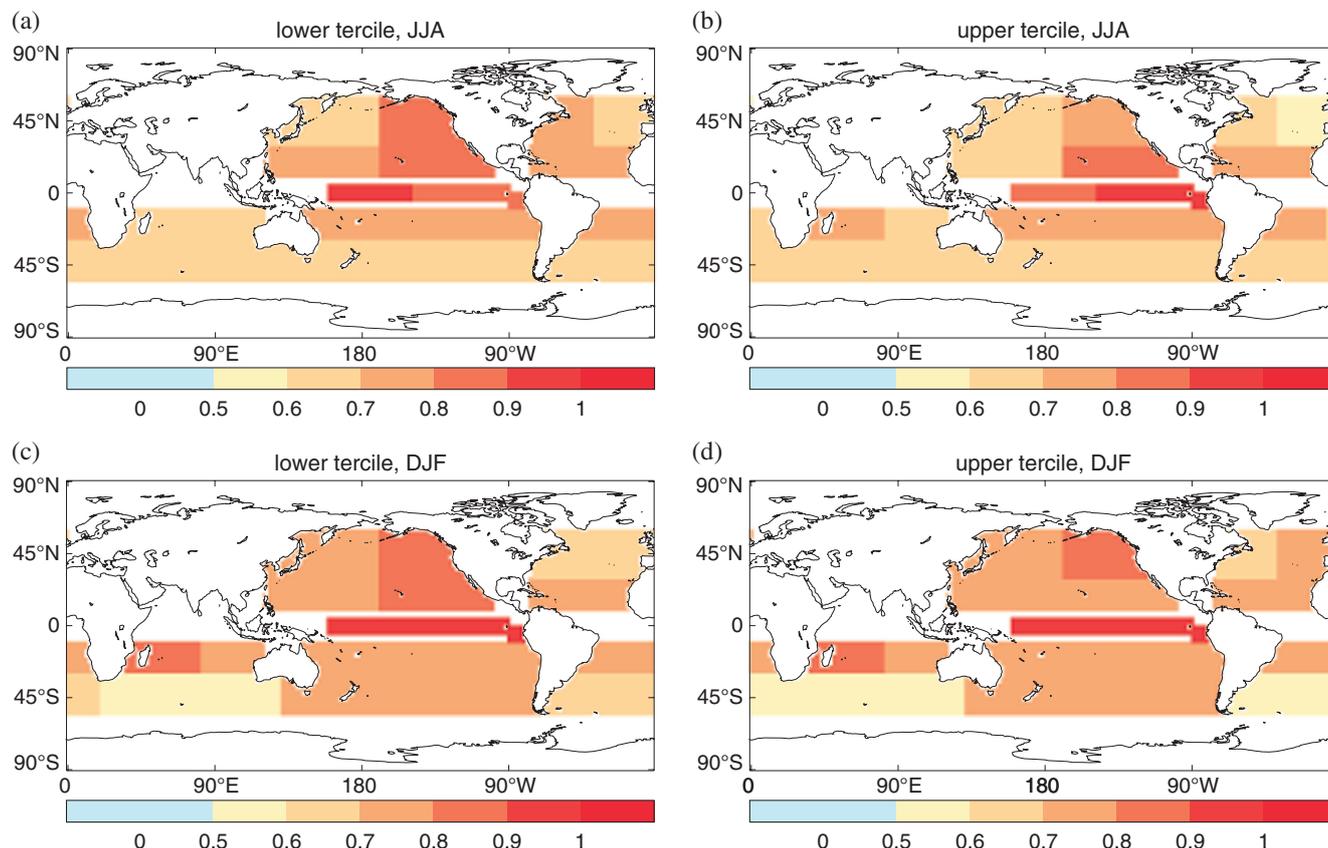


Figure 12. As Figure 11, but for SST.

of Arribas *et al.* (2011). It is based on combined empirical orthogonal function (CEOF) analysis of anomalies of NOAA outgoing long-wave radiation (OLR) and 850 and 200 hPa zonal winds from NCEP/NCAR<sup>‡</sup> reanalyses averaged over the latitude band 15°S–15°N. Anomalies of forecasted parameters are projected onto the first two observed CEOFs and the resulting time coefficients are named as Real-time Multivariate MJO indices (RMM1 and RMM2) as discussed by Wheeler and Hendon (2004).

Figure 9 shows the correlations between the forecast RMMs and the observed RMMs calculated at each forecast lead time. The forecasts are considered to be skilful when the correlations are greater than 0.5. Correlation values of persistence are shown for comparison. The MJO forecast skill of GloSea5 and GloSea4 are similar. GloSea5 RMM2 has a slightly better skill than GloSea4 for lead times of 15–20 days. However there is a small reduction in correlation skills of RMM1 in GloSea5 for lead times of 10–15 days. Both systems produce skilful MJO forecasts up to 15–20 days for both RMMs, which is well above the skills given by the persistence of initial conditions. Overall, GloSea5 shows significant skill in predicting the MJO out to around 20 days, which is similar to that in other monthly forecast systems (Vitart *et al.*, 2007b). Despite initializing the hindcasts from a different analysis system (ERA-Interim), GloSea5 produces high skill comparable to models which use consistent atmospheric analysis as initial conditions (e.g. Wang *et al.*, 2013b) for hindcasts and forecasts. This implies that the possible error derived from calculating model climatologies from initial conditions from a different analysis system is smaller than the forecast signal.

### 3.1.5. Tropical storms

Predictions of tropical storms are of huge interest to society and industry. On seasonal time-scales, climate models show skill in

predicting numbers of Atlantic tropical storms at lead times of up to 1 year (Vitart *et al.*, 2007a); more recently, multi-year forecasts of tropical storm frequency have been developed (Smith *et al.*, 2010). However, probabilistic forecasts of landfall risk have not been demonstrated beyond the medium-range time-scale.

We examine the tracks of model tropical storms in GloSea5 over the North Atlantic basin for June–November 1996–2009 using TRACK (Strachan *et al.*, 2013, and references therein). Model storms are compared with observations from the revised Atlantic best-track Hurricane Database (HURDAT2; Landsea and Franklin, 2013) in Figure 10. Overall the observed spatial distribution of storm tracks in the North Atlantic is well captured by GloSea5. However, the basin-wide frequency of storms is generally too low, in agreement with other studies (Bengtsson *et al.*, 2007; Strachan *et al.*, 2013). The lack of storms in the Caribbean Sea and Gulf of Mexico (Figure 10) is not seen in higher-resolution versions of HadGEM3 (Roberts *et al.*, 2014), but has been seen in other models (e.g. GFDL-HiRAM and FSU-COAPS; Strazzo *et al.*, 2013).<sup>§</sup> In GloSea5, a low SST bias and/or unfavourable atmospheric conditions (e.g. high vertical wind shear) may be inhibiting genesis in these regions. There are also too few storms passing through the region from the eastern Atlantic; this may be due to a weakness (or break) in the North Atlantic subtropical high causing the early recurvature of storms to the north instead of steering them westward. Further analysis is being carried out to test these hypotheses which are beyond the scope of this article.

The majority of tropical storms in GloSea5 have genesis in the central and eastern tropical Atlantic, a region particularly important for landfall studies since these storms (so-called ‘Cape Verde hurricanes’) have historically comprised the majority of the most destructive US landfalling hurricanes (Kossin *et al.*, 2010).

<sup>‡</sup>(US) National Centers for Environmental Prediction/National Center for Atmospheric Research

<sup>§</sup>Geophysical Fluid Dynamics Laboratory High Resolution Atmospheric Model; Florida State University Center for Ocean–Atmospheric Prediction Studies.

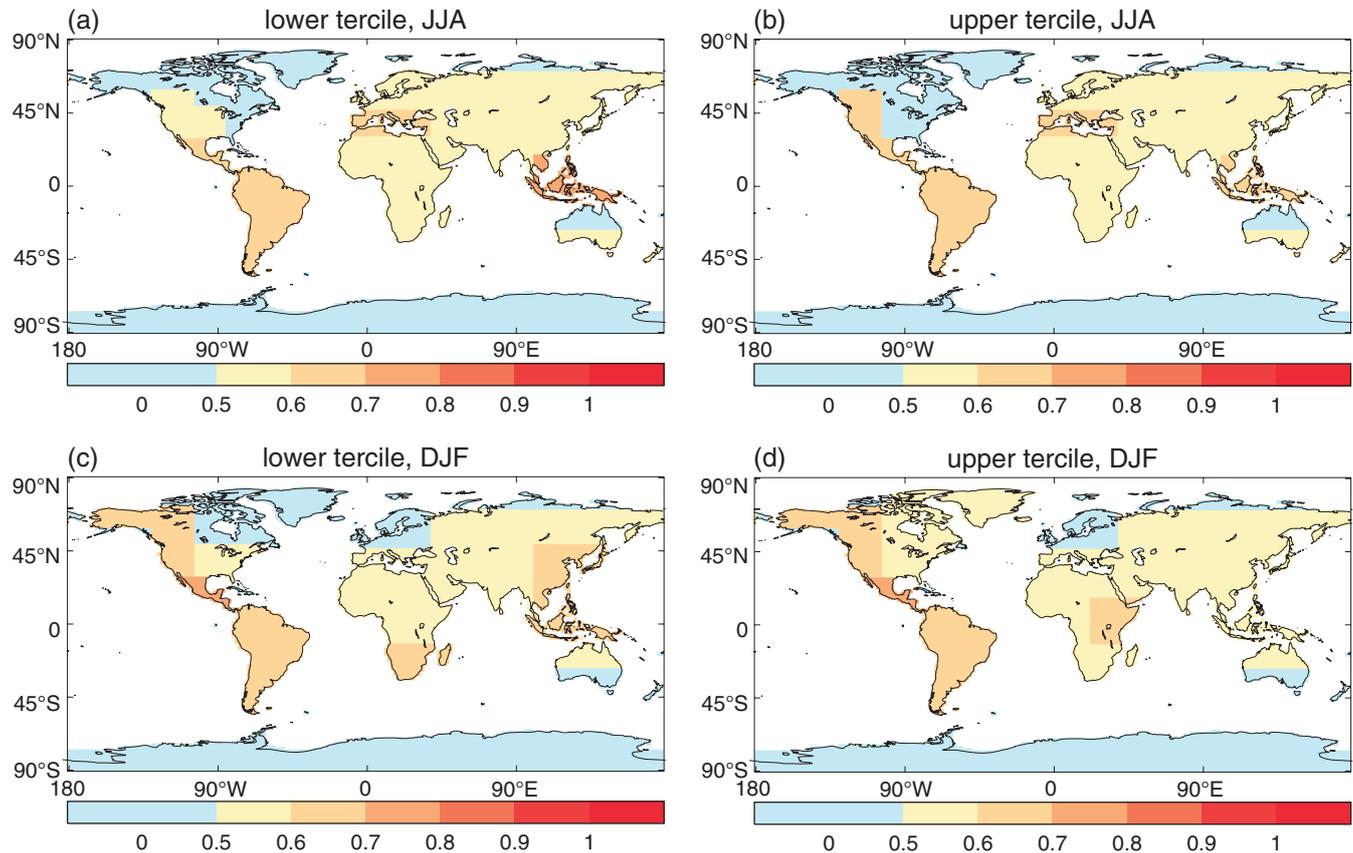


Figure 13. As Figure 11, but for precipitation.

It is encouraging to note that, in GloSea5, some storms that originate from this region also have tracks that cross the US coast. The ability to simulate storms from the eastern tropical Atlantic has greatly improved compared to GloSea4 (not shown), most likely due to the increase in horizontal resolution, which may provide a better simulation of both the frequency and intensity of African easterly waves from which these storms develop (Bain *et al.*, 2013).

### 3.2. Statistical performance

Figures 11–13 show the relative operating characteristics (ROC; Stanski *et al.*, 1989) for near-surface temperature, SST and precipitation for the GloSea5 system in the periods JJA and DJF. The near-surface temperature and precipitation scores were calculated for the Giorgi regions (Giorgi and Francisco, 2000) (the Australian region has been divided in two), and the SST scores have been calculated for the Niño and Enhanced Ocean Data Assimilation and Climate Prediction (ENACT; ECMWF, 2005) regions (regions are cut at 55°S to prevent sea-ice being included in the analysis). The SST results were verified against the Hadley Centre Global Sea Ice and Coverage and Sea Surface Temperature dataset (HadISST; Rayner *et al.*, 2003). ERA-Interim is used to verify the near-surface temperatures and the Global Precipitation Climatology Project (GPCP; Adler *et al.*, 2003) dataset to verify precipitation. Values above 0.5 denote useful skill compared to climatology.

Figure 11 shows the near-surface temperature ROC for upper and lower terciles in JJA and DJF. The plots show useful skill across the globe. There are useful levels of skill over Northern Europe and North America in winter, resulting from the AO/NAO performance. We do find a decrease in skill for the Northern Europe region for the summer period compared to GloSea4. This is due to a low SST bias which develops in the North Atlantic in the new model during the summer months because of a reduced amount of mixing in the upper layers of the ocean. Figure 12

shows the SST ROC for upper and lower terciles in JJA and DJF. Again we find useful skill over the globe. The values are either similar to GloSea4 or are improved. The increase in ROC in the North Atlantic is  $\sim 0.1$ . Figure 13 shows the precipitation ROC for upper and lower terciles in JJA and DJF.

## 4. Summary and future work

This article describes changes to the Met Office seasonal forecasting system since November 2011. The forecast ensemble has been extended to cover sub-seasonal time-scales and ensemble members are initialized daily, the physical parametrizations have been improved, and finally, the horizontal resolution has been increased to unprecedented levels for an operational seasonal forecast system, particularly for the ocean model component. We have described the updates to the forecast ensemble and the skill improvements in GloSea5. This upgrade has delivered improved ENSO patterns and useful skill for predicting the Arctic Oscillation and North Atlantic Oscillation. Improvements in predicting the Western North Pacific Subtropical High and in representing tropical cyclones have also been found. A more detailed investigation of tropical storms in GloSea5 can be found in (J. Camp, 2014; personal communication). Statistically significant prediction skill for tropical storm number, accumulated cyclone energy index, and landfall frequency is shown.

The seasonal forecast system has had a programme of frequent updates which has only been possible due to the companion hindcast being run in real time. We intend to continue this approach in the future. In late 2013 we will begin to test a new scientific configuration of the model, Global Atmosphere 6.0, which has a new dynamical core and other improvements to the physical parametrizations.

We are currently investigating the coupled model configuration described here for medium-range and decadal time-scales, with the aim of building a seamless medium-range–monthly–seasonal–decadal prediction system.

## Appendix

## Catalogue of observations used in NEMOVAR assimilation

SST: Sea-Surface Temperature satellite and *in situ* data

- *In situ* (ICOADS; Woodruff *et al.*, 2011).
- National Oceanic and Atmospheric Agency Pathfinder Advanced Very High Resolution Radiometer (AVHRR; Casey *et al.*, 2010).
- European Space Agency (Advance) Along-Track Scanning Radiometer (ATSR-1, ATSR-2, AATSR) on board the ERS-1 (ATSR-1), ERS-2 (ATSR-2) and Envisat (AATSR) satellites.
- ARCLAKE ATSR dataset for lakes and inland seas.
- After 2008 and for the forecast (FOAM) analysis, the Group for High-Resolution Sea Surface Temperature (GHRSSST) near-real-time dataset is used. In addition to the AVHRR and (A)ATSR instruments above, this also includes:
  - AVHRR observations from the MetOp-A and MetOp-B satellites.
  - Advanced Microwave Scanning Radiometer (AMSR-E) Instrument aboard the NASA AQUA satellite (no longer transmitting data).

## SLA: Sea-Level Anomaly altimeter satellite data

- Data from CLS/AVISO along-track delayed-time product (near-real-time data for forecast).
- Eight different instruments, eleven different orbits: ERS-1, ERS-2, Envisat, Topex/Poseidon, Jason 1, Jason 2, Geosat F/O, Cryostat.

## Sea-ice concentration data:

- EUMETSAT Ocean Sea Ice Satellite Application Facility (OSI-SAF) special sensor microwave imager (SSM/I) reanalysis (OSI-SAF, 2011).
- Reanalysis to 2007 and real-time product from 2008.

## Subsurface temperature and salinity data:

- GloSea5 Ocean Analysis uses quality-controlled EN3 profiles dataset (Ingleby and Huddleston, 2007).
- Forecast uses real-time equivalent which includes ARGO profiling floats, TRITON-TAO array and other moored buoys, expendable bathythermographs (XBTs), mechanical bathythermographs (MBTs), and conductivity–temperature–depth (CTD) sensors.

## Acknowledgements

This work was supported by the Joint DECC and Defra Integrated Climate Programme, DECC/Defra (GA01101) and received funding from the EU-funded SPECS programme (FP7-308378). NOAA AVHRR data were provided by GHRSSST and the US National Oceanographic Data Center. This project was supported in part by a grant from the NOAA Climate Data Record (CDR) Program for satellites. The authors wish to thank the numerous Met Office research staff who have contributed to make GloSea5 possible.

## References

Adler R, Huffman G, Chang A, Ferraro R, Xie P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P, Nelkin E. 2003. The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **4**: 1147–1167.

Arribas A, Glover M, Maidens A, Peterson K, Gordon M, MacLachlan C, Graham R, Fereday D, Camp J, Scaife AA, Xavier P, McLean P, Colman A, Cusack S. 2011. The GloSea4 ensemble prediction system for seasonal forecasting. *Mon. Weather Rev.* **139**: 1891–1910, doi: 10.1175/2010MWR3615.1.

Bain CL, Williams KD, Milton SF, Heming JT. 2013. Objective tracking of African Easterly Waves in Met Office models. *Q. J. R. Meteorol. Soc.* **140**: 47–57, doi: 10.1002/qj.2110.

Barnston AG, Tippett MK, L'Heureux ML, Li S, DeWitt DG. 2011. Skill of real-time seasonal ENSO model predictions during 2002–2011: Is our capability increasing? *Bull. Am. Meteorol. Soc.* **93**: 631–651, doi: 10.1175/BAMS-D-11-00111.1.

Bengtsson L, Hodges KI, Esch M. 2007. Tropical cyclones in a T159 resolution global climate model: Comparison with observations and re-analyses. *Tellus A* **59**: 396–416, doi: 10.1111/j.1600-0870.2007.00236.x.

Berckmans J, Woollings T, Demory ME, Vidale PL, Roberts M. 2013. Atmospheric blocking in a high resolution climate model: Influences of mean state, orography and eddy forcing. *Atmos. Sci. Lett.* **14**: 34–40, doi: 10.1002/asl2.412.

Best MJ, Pryor M, Clark DB, Rooney GG, Essery RLH, Ménard CB, Edwards JM, Hendry MA, Porson A, Gedney N, Mercado LM, Sitch S, Blyth E, Boucher O, Cox PM, Grimmond CSB, Harding RJ. 2011. The Joint UK Land Environment Simulator (JULES), model description—Part 1: Energy and water fluxes. *Geosci. Model Dev.* **4**: 677–699, doi: 10.5194/gmd-4-677-2011.

Blockley EW, Martin MJ, McLaren AJ, Ryan AG, Waters J, Lea DJ, Mirouze I, Peterson KA, Sellar A, Storkey D. 2013. Recent development of the Met Office operational ocean forecasting system: An overview and assessment of the new global foam forecasts. *Geosci. Model Dev. Discuss.* **6**: 6219–6278, doi: 10.5194/gmdd-6-6219-2013.

Blyth E, Best MJ, Cox PM, Essery RLH, Boucher O, Harding RJ, Prentice C, Vidale P, Woodward I. 2006. JULES: A new community land surface model. *IGBP News.* **66**: 9–11.

Bowler N, Arribas A, Beare S, Mylne KE, Shutts G. 2009. The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.* **135**: 767–776.

Brown A, Milton S, Cullen M, Golding B, Mitchell J, Shelly A. 2012. Unified modeling and prediction of weather and climate: A 25-year journey. *Bull. Am. Meteorol. Soc.* **93**: 1865–1877, doi: 10.1175/BAMS-D-12-00018.1.

Casey K, Brandon T, Cornillon P, Evans R. 2010. The past, present and future of the AVHRR Pathfinder SST Program. In *Oceanography from Space: Revisited*, Barale V, Gower J, Alberotanza L. (eds.) Springer: Berlin, doi: 10.1007/978-90-481-8681-5\_16.

Cassou C. 2008. Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature* **455**: 523–527, doi: 10.1038/nature07286.

Cionni I, Eyring V, Lamarque JF, Randel WJ, Stevenson DS, Wu F, Bodeker GE, Shepherd TG, Shindell DT, Waugh DW. 2011. Ozone database in support of cmip5 simulations: Results and corresponding radiative forcing. *Atmos. Chem. Phys. Discuss.* **11**: 10875–10933, doi: 10.5194/acpd-11-10875-2011.

Cohen J, Jones J. 2011. A new index for more accurate winter predictions. *Geophys. Res. Lett.* **38**: L21701, doi: 10.1029/2011GL049626.

Danabasoglu G, Large WG, Briegleb BP. 2010. Climate impacts of parameterized Nordic Sea overflows. *J. Geophys. Res.: Oceans* **115**: C11005, doi: 10.1029/2010JC006243.

Dawson A, Matthews AJ, Stevens DP, Roberts MJ, Vidale P. 2012. Importance of oceanic resolution and mean state on the extra-tropical response to El Niño in a matrix of coupled models. *Clim. Dyn.* **41**: 1439–1452, doi: 10.1007/s00382-012-1518-6.

Dee DP, Berrisford P, Poli P, Fuentes M. 2009. ERA-Interim for climate monitoring. *ECMWF Newsl.* **119**: 5–6.

Derome J, Lin H, Brunet G. 2005. Seasonal forecasting with a simple general circulation model: Predictive skill in the AO and PNA. *J. Clim.* **18**: 597–609, doi: 10.1175/JCLI-3289.1.

Donlon CJ, Martin M, Stark J, Roberts-Jones J, Fiedler E, Wimmer W. 2012. The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.* AATSR Special Issue. **116**: 140–158, doi: 10.1016/j.rse.2010.10.017.

ECMWF. 2005. 'ENACT: Enhanced ocean data assimilation and climate prediction'. A European Commission Framework 5 project: Contract no. EVK2-CT2001-00117, 1 January 2002 to 31 December 2004, Technical report, ECMWF: Reading, UK.

Folland CK, Scaife AA, Lindesay J, Stephenson DB. 2012. How potentially predictable is northern European winter climate a season ahead? *Int. J. Climatol.* **32**: 801–818, doi: 10.1002/joc.2314.

Giorgi F, Francisco R. 2000. Uncertainties in regional climate change prediction: A regional analysis of ensemble simulations with the HADCM2 coupled AOGCM. *Clim. Dyn.* **16**: 169–182.

Graham T. 2014. The importance of eddy-permitting model resolution for simulation of the heat budget of tropical instability waves. *Ocean Model.* **79**: 21–32, doi: 10.1016/j.ocemod.2014.04.005.

Guilyardi E, Wittenberg A, Fedorov A, Collins M, Wang C, Capotondi A, van Oldenborgh GJ, Stockdale T. 2009. Understanding El Niño in ocean–atmosphere general circulation models: Progress and challenges. *Bull. Am. Meteorol. Soc.* **90**: 325–340, doi: 10.1175/2008BAMS2387.1.

Hunke EC, Lipscomb WH. 2010. 'CICE: The sea ice model documentation and software user's manual, version 4.1', Technical report LA-CC-06-012. Los Alamos National Laboratory: Los Alamos, NM.

Ingleby B, Huddleston M. 2007. Quality control of ocean temperature and salinity profiles—historical and real-time data. *J. Mar. Syst.* **65**: 158–175, doi: 10.1016/j.jmarsys.2005.11.019.

Johansson Å. 2007. Prediction skill of the NAO and PNA from daily to seasonal time scales. *J. Clim.* **20**: 1957–1975, doi: 10.1175/JCLI4072.1.

- Jones C, Waliser DE, Schemm JKE, Lau WKM. 2000. Prediction skill of the Madden and Julian oscillation in dynamical extended-range forecasts. *Clim. Dyn.* **16**: 273–289.
- Kim HM, Webster P, Curry J. 2012. Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere winter. *Clim. Dyn.* **39**: 2957–2973, doi: 10.1007/s00382-012-1364-6.
- Kossin JP, Camargo SJ, Sitkowski M. 2010. Climate modulation of North Atlantic hurricane tracks. *J. Clim.* **23**: 3057–3076, doi: 10.1175/2010JCLI3497.1.
- Landsea CW, Franklin JL. 2013. How ‘good’ are the best tracks?—Estimating uncertainty in the Atlantic Hurricane Database. *Mon. Weather Rev.* **141**: 3576–3592, doi: 10.1175/MWR-D-12-00254.1.
- Large W, Yeager S. 2009. The global climatology of an interannually varying air–sea flux data set. *Clim. Dyn.* **33**: 341–364, doi: 10.1007/s00382-008-0441-3.
- Lea DJ, Drecourt JP, Haines K, Martin MJ. 2008. Ocean altimeter assimilation with observational- and model-bias correction. *Q. J. R. Meteorol. Soc.* **134**: 1761–1774, doi: 10.1002/qj.320.
- Liess S, Waliser DE, Schubert S. 2005. Predictability studies of the intraseasonal oscillation with the ECHAM5 GCM. *J. Atmos. Sci.* **62**: 3320–3336.
- Lin H, Brunet G, Derome J. 2009. An observed connection between the North Atlantic Oscillation and the Madden–Julian Oscillation. *J. Clim.* **22**: 364–380.
- Lin H, Brunet G, Fontecilla JS. 2010. Impact of the Madden–Julian Oscillation on the intraseasonal forecast skill of the North Atlantic Oscillation. *Geophys. Res. Lett.* **37**: L19803, doi: 10.1029/2010GL044315.
- Lu R, Dong B. 2001. Westward extension of North Pacific subtropical high in summer. *J. Meteorol. Soc. Jpn. Ser. II* **79**: 1229–1241.
- Luo JJ, Masson S, Roeckner E, Madec G, Yamagata T. 2005. Reducing climatology bias in an ocean–atmosphere CGCM with improved coupling physics. *J. Clim.* **18**: 2344–2360, doi: 10.1175/JCLI3404.1.
- Madden RA, Julian PR. 1994. Observations of the 40–50 day tropical oscillation: A review. *Mon. Weather Rev.* **122**: 813–837.
- Madec G. 2008. NEMO Ocean Engine, Note du Pole de Modelisation. Institut Pierre-Simon Laplace (IPSL): Paris.
- Martin M, Hines A, Bell M. 2007. Data assimilation in the FOAM operational short-range ocean forecasting system: A description of the scheme and its impact. *Q. J. R. Meteorol. Soc.* **133**: 981–995.
- Mogensen K, Balmaseda M, Weaver AT, Martin M, Vidard A. 2009. NEMOVAR: A variational data assimilation system for the NEMO ocean model. In *ECMWF Newsletter*, Walter Z. (ed.) **120**: 17–21. ECMWF: Reading, UK.
- Mogensen K, Balmaseda MA, Weaver AT. 2012. ‘The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4’, Technical Report TR-CMGC-12-30. CERFACS: Toulouse, France.
- Müller WA, Appenzeller C, Schär C. 2005. Probabilistic seasonal prediction of the winter North Atlantic Oscillation and its impact on near surface temperature. *Clim. Dyn.* **24**: 213–226, doi: 10.1007/s00382-004-0492-z.
- OSI-SAF. 2011. Global Sea Ice Concentration Reprocessing Dataset 1978–2009 (v1.1). Ocean and Sea Ice Satellite Application Facility. EUMETSAT: Darmstadt, Germany. <http://osisaf.met.no> (accessed 25 May 2014).
- Rawlins F, Ballard SP, Bovis KJ, Clayton AM, Li D, Inverarity GW, Lorenc AC, Payne TJ. 2007. The Met Office global four-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.* **133**: 347–362, doi: 10.1002/qj.32.
- Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**: D14, doi: 10.1029/2002JD002670.
- Rayner N, Brohan P, Parker D, Folland CK, Kennedy J, Vanicek M, Ansell T, Tett S. 2006. Improved analyses of changes and uncertainties in sea surface temperature measured *in situ* since the mid-nineteenth century: The HadSST2 data set. *J. Clim.* **19**: 446–469.
- Roberts MJ, Clayton A, Demory ME, Donners J, Vidale PL, Norton W, Shaffrey L, Stevens DP, Stevens I, Wood RA, Slingo JM. 2009. Impact of resolution on the tropical Pacific circulation in a matrix of coupled models. *J. Clim.* **22**: 2541–2556, doi: 10.1175/2008JCLI2537.1.
- Roberts MJ, Vidale P, Mizielinski M, Demory ME, Schiemann R, Strachan J, Hodges K, Camp J. 2014. Tropical cyclones in the upscale ensemble of high-resolution global climate models. *J. Clim.* (in press).
- Scaife AA, Copsey D, Gordon C, Harris C, Hinton T, Keeley S, O’Neill A, Roberts M, Williams K. 2011. Improved Atlantic winter blocking in a climate model. *Geophys. Res. Lett.* **38**: L23703, doi: 10.1029/2011GL049573.
- Scaife AA, Arribas A, Blockley E, Brookshaw A, Clark RT, Dunstone N, Eade R, Fereday D, Folland CK, Gordon M, Hermanson L, Knight JR, Lea DJ, MacLachlan C, Maidens A, Martin M, Peterson AK, Smith D, Vellinga M, Wallace E, Waters J, Williams A. 2014. Skilful long-range prediction of European and North American winters. *Geophys. Res. Lett.* **41**: 2514–2519, doi: 10.1002/2014GL059637.
- Smith RD, Maltrud ME, Bryan FO, Hecht MW. 2000. Numerical simulation of the North Atlantic Ocean at 1/10°. *J. Phys. Oceanogr.* **30**: 1532–1561, doi: 10.1175/1520-0485(2000)030<1532:NSOTNA>2.0.CO;2.
- Smith DM, Eade R, Dunstone NJ, Fereday D, Murphy JM, Pohlmann H, Scaife AA. 2010. Skilful multi-year predictions of Atlantic hurricane frequency. *Nat. Geosci.* **3**: 846–849, doi: 10.1038/ngeo1004.
- Snedecor GW, Cochran WG. 1989. *Statistical Methods* (8th edn). Iowa State University Press: Iowa City, IA.
- Stanski HR, Wilson LJ, Burrows WR. 1989. ‘Survey of common verification methods in meteorology’, Technical report 8. WMO: Geneva, Switzerland.
- Strachan J, Vidale PL, Hodges K, Roberts M, Demory ME. 2013. Investigating global tropical cyclone activity with a hierarchy of AGCMs: The role of model resolution. *J. Clim.* **26**: 133–152, doi: 10.1175/JCLI-D-12-00012.1.
- Strazzo S, Elsner JB, LaRow T, Halperin DJ, Zhao M. 2013. Observed versus GCM-generated local tropical cyclone frequency: Comparisons using a spatial lattice. *J. Clim.* **26**: 8257–8268, doi: 10.1175/JCLI-D-12-00808.1.
- Thompson DWJ, Wallace JM. 1998. The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.* **25**: 1297–1300, doi: 10.1029/98GL00950.
- Vitart F, Huddleston MR, Déqué M, Peake D, Palmer TN, Stockdale TN, Davey MK, Ineson S, Weisheimer A. 2007a. Dynamically-based seasonal forecasts of Atlantic tropical storm activity issued in June by EURO-SIP. *Geophys. Res. Lett.* **34**: L16815, doi: 10.1029/2007GL030740.
- Vitart F, Woolnough S, Balmaseda MA, Tompkins AM. 2007b. Monthly forecast of the Madden–Julian Oscillation using a coupled GCM. *Mon. Weather Rev.* **135**: 2700–2715, doi: 10.1175/MWR3415.1.
- Waliser DE, Lau KM, Stern W, Jones C. 2003. Potential predictability of the Madden–Julian Oscillation. *Bull. Am. Meteorol. Soc.* **84**: 33–50.
- Walters DN, Best MJ, Bushell AC, Copsey D, Edwards JM, Falloon PD, Harris CM, Lock AP, Manners JC, Morcrette CJ, Roberts MJ, Stratton RA, Webster S, Wilkinson JM, Willett MR, Boutle IA, Earnshaw PD, Hill PG, MacLachlan C, Martin GM, Moufouma-Okia W, Palmer MD, Petch JC, Rooney GG, Scaife AA, Williams KD. 2011. The Met Office Unified Model Global Atmosphere 3.0/3.1 and JULES Global Land 3.0/3.1 configurations. *Geosci. Model Dev.* **4**: 919–941, doi: 10.5194/gmd-4-919-2011.
- Wang B, Xiang B, Lee JY. 2013a. Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. *Proc. Natl. Acad. Sci. U.S.A.* **110**, doi: 10.1073/pnas.1214626110.
- Wang W, Hung MP, Weaver S, Kumar A, Fu X. 2013b. MJO prediction in the NCEP climate forecast system version 2. *Clim. Dyn.* **42**: 2509–2520, doi: 10.1007/s00382-013-1806-9.
- Waters J, Lea DJ, Martin MJ, Mirouze I, Weaver AT, While J. 2014. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Q. J. R. Meteorol. Soc.*, doi: 10.1002/qj.2388.
- Weedon GP, Gomes S, Viterbo P, Shuttleworth WJ, Blyth E, Österle H, Adam JC, Bellouin N, Boucher O, Best M. 2011. Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J. Hydrometeorol.* **12**: 823–848, doi: 10.1175/2011JHM1369.1.
- Wheeler MC, Hendon HH. 2004. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Weather Rev.* **132**: 1917–1932.
- Wilks DS. 1995. *Statistical Methods in the Atmospheric Sciences*. Academic Press: London.
- Woodruff S, Worley S, Lubker S, Ji Z, Freeman J, Berry D, Brohan P, Kent EC, Reynolds R, Smith S, Wilkinson C. 2011. ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.* **31**: 951–967, doi: 10.1002/joc.2103.