

Benefits of Increasing the Model Resolution for the Seasonal Forecast Quality in EC-Earth[Ⓞ]

C. PRODHOMME,^a L. BATTÉ,^b F. MASSONNET,^{a,c,d} P. DAVINI,^{e,f} O. BELLPRAT,^a V. GUEMAS,^{a,b}
AND F. J. DOBLAS-REYES,^{a,c,g}

^a *Barcelona Supercomputing Center, Barcelona, Spain*

^b *CNRM-GAME, Météo-France, Toulouse, France*

^c *Institut Català de Ciències del Clima (IC3), Barcelona, Spain*

^d *Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute, Université Catholique de Louvain, Louvain-La-Neuve, Belgium*

^e *Laboratoire de Météorologie Dynamique/IPSL, Ecole Normale Supérieure, Paris, France*

^f *Institute of Atmospheric Sciences and Climate, ISAC-CNR, Torino, Italy*

^g *Institució Catalana de Recerca I Estudis Avançats (ICREA), Barcelona, Spain*

(Manuscript received 8 February 2016, in final form 15 June 2016)

ABSTRACT

Resolution in climate models is thought to be an important factor for advancing seasonal prediction capability. To test this hypothesis, seasonal ensemble reforecasts are conducted over 1993–2009 with the European community model EC-Earth in three configurations: standard resolution ($\sim 1^\circ$ and ~ 60 km in the ocean and atmosphere models, respectively), intermediate resolution ($\sim 0.25^\circ$ and ~ 60 km), and high resolution ($\sim 0.25^\circ$ and ~ 39 km), the two latter configurations being used without any specific tuning. The model systematic biases of 2-m temperature, sea surface temperature (SST), and wind speed are generally reduced. Notably, the tropical Pacific cold tongue bias is significantly reduced, the Somali upwelling is better represented, and excessive precipitation over the Indian Ocean and over the Maritime Continent is decreased. In terms of skill, tropical SSTs and precipitation are better reforecasted in the Pacific and the Indian Oceans at higher resolutions. In particular, the Indian monsoon is better predicted. Improvements are more difficult to detect at middle and high latitudes. Still, a slight improvement is found in the prediction of the winter North Atlantic Oscillation (NAO) along with a more realistic representation of atmospheric blocking. The sea ice extent bias is unchanged, but the skill of the reforecasts increases in some cases, such as in summer for the pan-Arctic sea ice. All these results emphasize the idea that the resolution increase is an essential feature for forecast system development. At the same time, resolution alone cannot tackle all the forecast system deficiencies and will have to be implemented alongside new physical improvements to significantly push the boundaries of seasonal prediction.

1. Introduction

Climate forecasting at a subseasonal to interannual time range is now done routinely and operationally by an increasing number of research centers and institutions. Although forecasting systems using numerical general circulation models (GCMs) have made substantial progress in the last decades (Doblas-Reyes et al.

2013), systematic errors and the misrepresentations of key processes still hinder forecast quality and limit the value of dynamical prediction in certain areas of the globe (Lin 2007; Guemas et al. 2012; Vannièrè et al. 2013; Voldoire et al. 2014). The desire to better capture physical processes in the ocean and atmosphere, alongside a continued development of the computational efficiency of high-performance clusters used to run climate models, has motivated an increasing number of studies using higher-resolution components of the climate system for historical simulations and climate change projections (Gent et al. 2010; Delworth et al. 2012; Sakamoto et al. 2012; Hourdin et al. 2013), as well as for seasonal forecasting (MacLachlan et al. 2015; Jia et al. 2015).

[Ⓞ] Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-16-01117.s1>.

Corresponding author e-mail: Chloé Prodhomme, chloe.prodhomme@bsc.es

DOI: 10.1175/JCLI-D-16-01117.1

Positive impacts of increased resolution have been reported in many studies using fully coupled or atmosphere-only GCMs (CGCMs and AGCMs, respectively), as documented in Table 1. From this table, it appears that increasing the resolution of atmosphere and/or ocean has beneficial impacts on the representation of the Asian monsoon independently of the baseline resolution in both AGCMs (Sperber et al. 1994; Lal et al. 1997; Branković and Gregory 2001; Mizielinski et al. 2014; Johnson et al. 2016) and CGCMs (Gent et al. 2010; Delworth et al. 2012). Despite those improvements, many studies acknowledge that the main biased features of the monsoon are not corrected by increasing resolution (Martin 1999; Johnson et al. 2016). For other phenomena, such as atmospheric blockings, the impact of increased resolution exhibits a dependence on the baseline resolution: improvements are only noted in studies considering atmospheric resolution higher than 100 km (AGCM experiments: Berckmans et al. 2013; Matsueda et al. 2009; Jung et al. 2012; Dawson and Palmer 2014). The well-known wide spread among representations of the CGCM cold tongue bias (Vannière et al. 2013) has been shown to be reduced in three different studies using high-resolution ocean component (less than 35 km; Shaffrey et al. 2009; Roberts et al. 2009; Sakamoto et al. 2012), thanks to a better representation of tropical instability waves (Roberts et al. 2009). The improvement of the representation of orographic winds and their effects on the ocean at high resolution (atmosphere resolution higher than 50 km) improves the representation of the Somali jet (AGCM study; Johnson et al. 2016) and Pacific upwelling (CGCM studies; Gent et al. 2010; Sakamoto et al. 2012). The characteristics of El Niño–Southern Oscillation (ENSO) have shown to be sensitive to atmospheric resolution (Guilyardi et al. 2004; seasonal forecast: Jia et al. 2015), oceanic resolution (Kirtman et al. 2012), and both oceanic and atmospheric resolutions (Shaffrey et al. 2009; Sakamoto et al. 2012; Delworth et al. 2012). However, in seasonal forecast experiments using resolution and grid spacing ratios similar to these studies, MacLachlan et al. (2015) conclude that the ENSO skill was not affected by an increased resolution. Similarly, Zhu et al. (2015) do not show any change in the ENSO skill when increasing atmospheric resolution to 15 km coupled with an oceanic resolution of 100 km. Several other improvements have also been reported as listed in Table 1.

Most of the previously cited studies concerned pre-industrial, historical, and climate projection experiments. The impact of increasing resolution in the atmosphere on seasonal forecast quality has been comparatively much less documented. GFDL operates a high-resolution atmosphere (0.5°) and standard

resolution ocean (1°) versions of their CM2.5 coupled climate model for seasonal forecasts (Jia et al. 2015; Yang et al. 2015; see Table 1 for more details) as a contribution to the North American Multimodel Ensemble (NMME) coordinated experiment. The UK Met Office GloSea5 system now uses high resolution in the ocean (0.25° and 75 vertical levels) and in the atmosphere (0.83° longitude \times 0.55° latitude, \sim 50 km at midlatitudes), and exhibits encouraging improvements in extratropical forecasting skill, including surface North Atlantic Oscillation (NAO), winter storminess, and near-surface temperature and wind speed over Europe and North America (Scaife et al. 2014; MacLachlan et al. 2015).

The aim of the present study is to answer the following two questions: 1) What is the impact of increasing oceanic and atmospheric resolutions on seasonal forecast quality? 2) Could an increase of resolution result in an improvement of key phenomena at the seasonal time scale, such as the NAO, ENSO, blocking, and Indian monsoon? In particular, how model-dependent are the results of Scaife et al. (2014) and MacLachlan et al. (2015), especially concerning the NAO? To answer those questions, we performed seasonal reforecasts over a 17-yr period with the EC-Earth coupled model (ocean, atmosphere, land, and sea ice components) at three different resolutions: standard (SRes), increased in the ocean (IRes), and increased in both the ocean and the atmosphere (HRes). The standard horizontal resolution settings for EC-Earth are T255 for the atmosphere (approximately 0.7° in latitude and longitude and 91 vertical levels) and a 1° oceanic grid (with refinements around the equator and at the poles) counting 46 vertical layers. At the intermediate resolution the oceanic resolution is increased to a 0.25° horizontal grid with 75 vertical layers and in the high resolution both oceanic and atmospheric resolutions are enhanced, with 0.25° and 75 vertical layers in the ocean and T511 in the atmosphere (\sim 0.35° in latitude and longitude). These resolutions have been chosen because they are comparable with the resolutions of GloSea5 (slightly higher in the atmosphere) and thus we could expect to see similar improvements especially for the NAO skill. Following Table 1, we might also expect improvement of the Asian monsoon, midlatitude blocking frequency, Pacific cold tongue bias, oceanic upwelling, and the representation of ENSO. Analyzing IRes and HRes separately will help in understanding the importance of enhanced resolution in the ocean rather than in atmosphere. The separate effects of oceanic and atmospheric resolutions have been poorly documented up to now (Kirtman et al. 2012).

The paper is organized as follows. Section 2 describes the EC-Earth coupled model in detail, the experimental

TABLE 1. Brief summary on the existing literature on the impact of resolution increase on different climate aspects (++) mark that the author of the considered study conclude on improvements, (==) when no change is observed, and (-) when the increase of resolution leads to a degradation. The resolution (in km) is estimated approximatively from the information available, considering averaged longitude and latitude resolution. SF stands for seasonal forecast and SI for seasonal integrations (initialized atmosphere-only simulations). This table was filled with all the information available to the best of our knowledge but is probably not exhaustive.

Reference	AGCM	CGCM	Ocean res	Atmos res	Asian monsoon	European blocking	Cold tongue	ENSO	NAO	Ocean upwelling	Others
Sperber et al. (1994)	X			600 → 210	(++)						
Lal et al. (1997)	X			600 → 210	(++)						
Hourdin et al. (2013)	X	X	200	300 → 180							Improve cloud cover, reduce cold SST bias, polar jet
Martin (1999)	X			600 → 100	(==)						
Shaffrey et al. (2009)		X	100 → 35	150 → 100			(++)	(++)			
Roberts et al. (2009)		X	100 → 35	135 → 90			(++)				
Scaife et al. (2014)		SF	25	70					(++)		Extratropical forecasting skill (T2M, wind), winter storm Arctic oscillation
MacLachlan et al. (2015)			100 → 25	150 → 70				(==) skill	(++)		
Branković and Gregory (2001)	SI			210 → 60	T159: (++) T319: (--)						geopotential at 500 hPa, wind, temperature at 200 hPa
Jia et al. (2015)		SF	100	220 → 50				(++) skill			T2M skill, ENSO teleconnection; storm track
Yang et al. (2015)										(++) Pacific upwelling	Ice thickness
Gent et al. (2010)		X	100	100 → 50	(++)		(==)				
Kirtman et al. (2012)		X	70 → 10	50							Change in precipitation
Sakamoto et al. (2012)		X	140 → 25	110 → 50			(++)	change (++)		(++) Pacific upwelling	Bias reduction for precip and T2M
Delworth et al. (2012)		X	100 → 25	200 → 50	(++)		(==)	(++)			Double ITCZ
Berekmans et al. (2013)	X			130 → 60		(++)					
Johnson et al. (2016)	X			200 → 40	(++)					(++)	Somalian jet
Matsueda et al. (2009)	X			180 → 20		(++)					
Zhu et al. (2015)		X	100	60 → 15				(==)			
Jung et al. (2012)	SI			125 → 10		(++)					

setup, and the methods used to assess the changes due to the resolution. Global results in terms of forecast quality are presented in [section 3](#). Then, the results of the resolution changes will be further described in [section 4](#), in three different subsections focused first on the tropics, then on the midlatitudes and more specifically Europe, and finally on high latitudes. Discussion and conclusions can be found in [section 5](#).

2. Experiments and methods

a. The EC-Earth 3.0.1 coupled model

The EC-Earth Earth system model (ESM), used to perform the reforecast of this study, is developed by the EC-Earth consortium, counting close to 20 European institutions. EC-Earth consists in the coupling of different models representing the components of the Earth system: atmosphere, land, ocean, sea ice, vegetation, glaciers, and atmospheric chemistry. In the present study, the version 3.0.1 of the coupled model, including the ocean, atmosphere, land, and sea ice components, is used to reforecast seasonal climate over a 17-yr period. An earlier version of the EC-Earth ESM is described in detail in [Hazeleger et al. \(2012\)](#). The main differences between these two versions are an improved radiation scheme ([Morcrette et al. 2008](#)) and a new cloud microphysics scheme ([Forbes et al. 2011](#)).

The atmosphere model is the Integrated Forecasting System (IFS) model cy36r4. In line with the objectives of this study, we use two different horizontal atmospheric resolutions: T255 (linear triangular truncation at wavenumber 255, corresponding to approximately 0.7° in latitude and longitude) and T511 (linear triangular truncation at wavenumber 511, corresponding to approximately 0.35° in latitude and longitude). At both resolutions, the number of vertical layers is identical (91, up to 0.01 Pa). The time steps for IFS are 2700 and 900 s for T255 and T511 configurations, respectively.

EC-Earth embeds the NEMO version 3.3.1 ocean model ([Madec 2008](#)). Again, in order to address the scientific question raised in the introduction, we use two different horizontal resolutions: the ORCA-1 ($\sim 1^\circ$) and ORCA-025 ($\sim 0.25^\circ$) grids with, respectively, 46 and 75 layers. The grid has higher horizontal resolution (of about one-third of a degree in the ORCA-1 configuration and one-tenth of a degree in the ORCA-025 configuration) near the equator to resolve equatorial planetary waves. The grid has three poles, one on the South Pole and two near the North Pole (one in Canada and one in Siberia), but at different longitudes and latitudes. As a result, the horizontal resolution is increased in the vicinity of the North Pole. The sea ice component

TABLE 2. Settings for atmospheric and ocean resolution for the three sets of experiments presented in this manuscript.

Experiment name	Atmosphere resolution	Ocean resolution
SRes	T255L91	ORCA1L46
IRes	T255L91	ORCA025L75
HRes	T511L91	ORCA025L75

used in this study is the second version of the Louvain-la-Neuve (LIM2) sea ice model ([Fichefet and Morales Maqueda 1997](#)). The time steps are 3600 and 1200 s for ORCA-1 and ORCA-025, respectively, and the sea ice model is called every ocean time step at both resolutions.

EC-Earth uses the H-TESEL (TESEL for Tiled ECMWF Scheme for Surface Exchanges over Land) scheme for the land surface ([van den Hurk et al. 2000](#)), which includes an improved representation of hydrology over the TESEL scheme, in agreement with more recent IFS cycles ([Balsamo et al. 2009](#)).

The atmosphere and ocean/ice components are coupled with the Ocean Atmosphere Sea Ice Soil version 3 (OASIS3; [Valcke 2013](#)) coupler. The coupling frequency between the ocean and atmosphere is every 3 h.

b. Experimental setup and initialization strategy

Three sets of 4-month 10-member seasonal reforecasts were carried out with different configurations of the atmosphere and ocean components ([Table 2](#)). Forecasts start on 1 May and 1 November every year from 1993 to 2009. Three experiments are considered: SRes (standard resolution: T255-ORCA1L46), IRes (intermediate resolution: T255-ORCA025L75), and HRes (high resolution: T511-ORCA025L75). For all configurations the default version of EC-Earth 3.0.1 as released by the EC-Earth consortium has been used. The simulations have been run using the autosubmit workflow manager ([Manubens-Gil et al. 2016](#)).

It is worth noting, and remembering for the remainder of the paper, that only the standard-resolution version has undergone extensive tuning. However, a number of parameters have been modified in the high-resolution configuration. The high-resolution ocean model uses increased albedo parameters for sea surface temperatures and sea ice, increased nonlinear bottom drag, modified eddy diffusivities and viscosities, and increased surface input of the turbulent kinetic energy. Furthermore, the Langmuir parameterization is switched off and the advection scheme for tracers has been changed for individual members due to numerical instabilities. In the atmospheric model the parameters related to the momentum flux of gravity waves and a limiter for wind tendencies in the upper atmosphere is changed. All

parameter and their respective values are summarized in the supplementary information except for the numerical parameters associated to the solver, which change according to the model resolution. The main characteristics of these simulations are summarized in Table 1.

The atmospheric initial conditions are generated from ERA-Interim. As the number of vertical levels is different in EC-Earth and ERA-Interim, a vertical interpolation of the model-level variables is performed. The 10 atmospheric initial conditions used to create the ensemble are generated using atmospheric singular vectors (Du et al. 2012). Ocean and sea ice initial conditions provided by the Global Ocean Reanalysis and Simulations (GLORYS2v1), produced at the ORCA-025 resolution (Ferry et al. 2010), have been used for IRes and HRes. For the SRes experiment, these initial conditions have been interpolated to the ORCA1 resolution and smoothed to avoid initial shocks.

c. Methods

1) BIAS AND BIAS CORRECTION

In our study, we analyze reforecast skill and bias for the May and boreal summer [June–August (JJA)] predictions (for May initialization), and for the November and boreal winter [December–February (DJF)] predictions (for November initialization). When initialized from an estimate of the observed climate state, reforecasts typically drift toward their attractor (i.e., toward the own stationary model climate). The drift can be understood as a systematic error that depends on the forecast time. Because model bias is a long-standing issue in climate science, we first address the impact of enhanced resolution on the development of this bias in our system. We define the forecast climatology for a specific forecast time as the forecast values averaged over all the members and all the start dates. The bias is therefore defined as the difference between the forecast and the observed climatologies over the same period. We apply a simple per-pair bias correction (García-Serrano et al. 2013) to our seasonal reforecasts before estimating the forecast skill and reliability. This method of bias correction consists in subtracting the climatology of the forecast to the different member and start dates. The same procedure is applied on the observations. This way the mean bias of the model is removed and the reforecasts can be compared with the observations. This method has been used in cross-validation mode; this way the climatology is calculated excluding the year that is forecasted.

2) SKILL ASSESSMENT

The skill is assessed using the anomaly correlation coefficient (ACC), the root-mean-square error (RMSE), and the Brier skill score (BS; Brier 1950) as defined below:

$$\text{ACC}(f) = \frac{\sum_i [m_i(f) - \bar{m}(f)][r_i(f) - \bar{r}(f)]}{\sqrt{\sum_i [m_i(f) - \bar{m}(f)]^2 \sum_i [r_i(f) - \bar{r}(f)]^2}}, \quad (1)$$

$$\text{RMSE}(f) = \sqrt{\frac{1}{N} \sum_{i=1}^N [m_i(f) - r_i(f)]^2}, \quad (2)$$

where f is the forecast time, $m_i(f)$ the bias-corrected model ensemble mean forecast for year i , $\bar{m}(f)$ the forecast averaged over the reforecast period, $r_i(f)$ and $\bar{r}(f)$ the corresponding reference data at forecast time f , and N the number of years in the reforecast period. Also,

$$\text{BS}(f) = \frac{1}{N} \sum_{i=1}^N [y_i(f) - o_i(f)]^2, \quad (3)$$

where $y_i(f)$ is the probabilistic forecast of a given event (e.g., temperature reaches over the second tercile of reference data in year i at forecast time f) based on the fraction of ensemble members predicting this event; $o_i(f)$ is the corresponding “observation” in the reference data and has a value of 1 if the event happens, 0 otherwise. The Brier score is a distance in probability space and should be as small as possible.

We use the standard reliability-resolution-uncertainty decomposition of the Brier score as in Toth et al. (2003), by binning forecast–observation pairs according to $J = 11$ since our ensemble size is 10 members. If N_j is the number of forecasts worth y_j , the decomposition is written as follows:

$$\begin{aligned} \text{BS}(f) &= \frac{1}{N} \sum_{j=1}^J N_j [y_j(f) - \bar{o}(f)]^2 \\ &\quad - \frac{1}{N} \sum_{j=1}^J N_j [o_j(f) - \bar{o}(f)]^2 + \bar{o}(f)[1 - \bar{o}(f)], \quad (4) \\ \text{BS}(f) &= \text{Rel}(f) - \text{Res}(f) + \text{Unc}(f). \quad (5) \end{aligned}$$

Reliability (Rel) is an estimate of the conditional bias of the probabilistic forecast, whereas resolution (Res) evaluates how well the model separates probabilistic events. Uncertainty (Unc) is independent from the model and depends only on the probabilistic event we choose to evaluate the model on.

We provide uncertainty estimates and confidence intervals with all the skill assessment. We use the Student

distribution with N degrees of freedom to estimate the significance level of correlation, N being the effective number of independent data calculated following the method of [von Storch and Zwiers \(2001\)](#). The significance of the difference between two correlations is estimated using the methodology of [Siegert et al. \(2016\)](#), which takes into account the dependence from sharing the same observations in both correlation coefficients. This method to assess the significance of the difference of two correlations also takes into account the independent number of data, which is necessary given the serial correlation typical of the time series considered.

Different observational and reanalysis datasets are used in the verification process to assess the robustness of the results. The ERA-Interim reanalysis ([Dee et al. 2011](#)) is used for the two-meter air temperature (T2M), sea level pressure (SLP), winds, and daily precipitations. For sea surface temperature (SST), the ERSST v3b ([Smith et al. 2008](#)) and European Space Agency Climate Change Initiative (ESA CCI) datasets are used. For precipitation, in addition to the daily precipitation from ERA-Interim, the Global Precipitation Climatology Project (GPCP) v2.2 ([Adler et al. 2003](#)) and the Global Precipitation Climatology Centre (GPCC) data product version 4 ([Schneider et al. 2008](#)) are used. For sea ice, the bootstrap sea ice concentrations version 2 ([Comiso 2000](#)) were used as a reference, although the evaluation was also conducted on alternative products: ESA CCI ([Ivanova et al. 2015](#)), Ocean and Sea Ice Satellite Application Facility (OSI-SAF; [Eastwood et al. 2014](#)), HadISST ([Rayner et al. 2003](#)), and Centennial In Situ Observation-Based Estimates of the Variability of SST and Marine Meteorological Variables, version 2 (COBE-2; [Hirahara et al. 2014](#)) to gauge the sensitivity of results to observational error. The reference datasets will be referred to as observation in the following. All the verification, as well as part of the plotting, have been done using the version 2.1.1 of the R-based s2dverification package (<http://cran.r-project.org/web/packages/s2dverification/index.html>). All the area-averaged indices have been computed directly on the native grids. For maps data have been interpolated on a $1^\circ \times 1^\circ$ grid. Some tests have been performed to assess the sensitivity of the results to this interpolation showing that interpolation does not affect the score representation (not shown).

3) DEFINITION

The monsoon onset date has been estimated from [Wang and LinHo \(2002\)](#) criteria. The onset date is defined as the first day for which precipitation from ERA-Interim averaged over India exceeds 5 mm day^{-1} . To estimate this date, we first compute the quantile corresponding to the 5 mm day^{-1} threshold, in cross-validation

mode, from observed daily data for the months of May and June. Then, the forecasted onset date is estimated as the first day exceeding this quantile in the model.

Several methods can be used to compute the NAO index, such as EOF analysis of sea level pressure or 500-hPa geopotential height fields (e.g., [Doblas-Reyes et al. 2003](#)), sea level pressure point value differences between Reykjavik and Gibraltar (e.g., [Maidens et al. 2013](#)), or area-averaged sea level pressure differences between subpolar regions and midlatitudes (e.g., [Stephenson et al. 2006](#)). In this study, we choose to compute the NAO with two different indices based on an EOF analysis leading to an NAO index for the model (Pmod) and the observations (Pobs). The Pobs (Pmod) NAO index computation method consists in computing the NAO pattern as the leading EOF of sea level pressure in the reference dataset (the reforecast data) in cross-validation mode and projecting the model sea level pressure anomalies for the given month/season onto the pattern to compute the index. These results have also been compared with the method of [Stephenson et al. \(2006\)](#).

The instantaneous blocking index computed as in [Davini et al. \(2012\)](#) is used. This index is a 2D extension of the [Tibaldi and Molteni \(1990\)](#) index and provides a measure of the Rossby wave breaking activity in the midlatitudes (between 30° and 75°N). It is based on the reversal of the daily 500-hPa geopotential height gradient: data are interpolated on a common $2.5^\circ \times 2.5^\circ$ grid before the index is computed. The GHGS and GHGN gradients are thus computed as follows:

$$\text{GHGS}(\lambda_0, \phi_0) = \frac{Z500(\lambda_0, \phi_0) - Z500(\lambda_0, \phi_S)}{\phi_0 - \phi_S}, \quad (6)$$

$$\text{GHGN}(\lambda_0, \phi_0) = \frac{Z500(\lambda_0, \phi_N) - Z500(\lambda_0, \phi_0)}{\phi_0 - \phi_S}, \quad (7)$$

where ϕ_0 ranges from 30° to 75°N and λ_0 ranges from 0° to 360° ; $\phi_S = \phi_0 - 15^\circ$ and $\phi_N = \phi_0 + 15^\circ$. Instantaneous blocking (IB) is identified when

$$\text{GHGS}(\lambda_0, \phi_0) > 0 \quad \text{GHGS}(\lambda_0, \phi_0) < -\frac{10 \text{ m}}{^\circ\text{lat}}. \quad (8)$$

ERA-Interim reanalysis data are used to assess the monsoon onset, NAO, and blocking results.

3. Results

a. Bias reduction

The impact of increasing resolution on the model mean climate is first assessed by considering how the

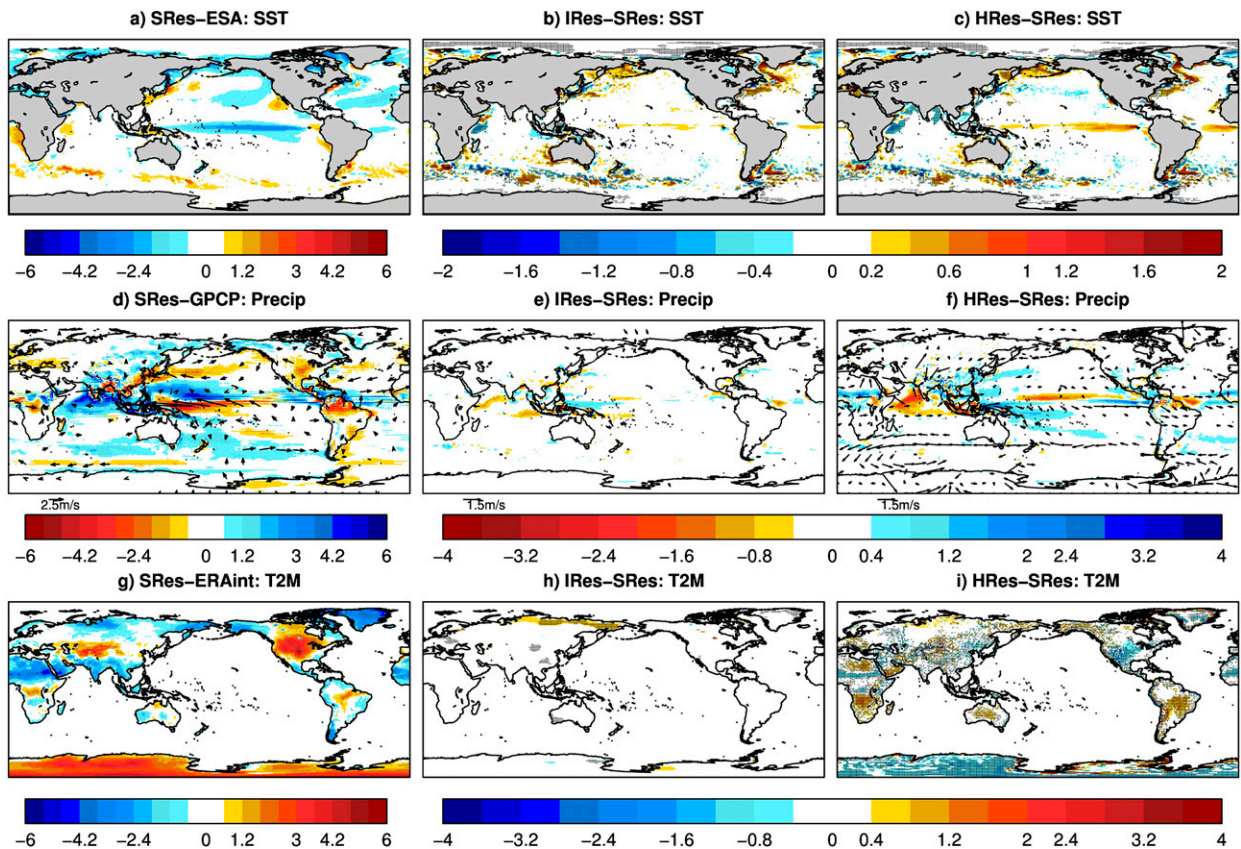


FIG. 1. (a) Mean bias of SST (unit: K) in JJA (months 2 to 4 of the forecasts initialized in May) in the SRes experiment, with respect to ESA. (b) Difference of SST for the IRes experiment with respect to the SRes experiments. (c) Difference of SST for the HRes experiment with respect to the SRes experiments. (d)–(f) As in (a)–(c), but for precipitation (colors, unit: mm day^{-1}) and 850-hPa winds (arrows, unit: m s^{-1}). (g)–(i) As in (a)–(c), but for near-surface atmosphere temperature (unit: K). For (b), (c), (e), and (f), dots show the points where the difference is significant at the 95% confidence level according to the Student's t test. For (e) and (f) only the values of wind differences significant at 95% confidence level are shown.

bias in the SRes reforecasts changes in the IRes and HRes experiments. The winter (DJF) and summer (JJA) biases are computed for forecast months 2 to 4 of the forecast initialized on 1 November and 1 May respectively following the methodology described in section 2c. These are the fast growing systematic errors of the coupled model, as some biases take several years to develop. The biases described here do not necessarily correspond to the stationary biases of the coupled model (Toniazio and Woolnough 2014). Figures 1 and 2 show results for JJA and DJF reforecasts, respectively, for SST, precipitation, and 850-hPa winds. Figures 1a and 2a show the classical cold tongue bias in the equatorial Pacific (Vanni ere et al. 2013) of up to 3°C in the SRes. A similar bias is also visible in boreal summer in the tropical Atlantic (Wahl et al. 2011; Liu et al. 2012). SRes simulates too-warm SST in the western boundary currents, Kuroshio, and Gulf Stream (as was the case in EC-Earth 2.2; Hazeleger et al. 2012; Gent et al. 2010)

and in the Antarctic circumpolar region. All the previously mentioned biases are present in both seasons but are stronger in DJF.

In summer, in the Indian Ocean, SRes exhibits a warm bias in the Somali upwelling that is common to other coupled models (Prodhomme et al. 2014). In the summer hemisphere (JJA in the Northern Hemisphere and DJF in the Southern Hemisphere) SRes exhibits a large cold bias in the midlatitudes.

Precipitation biases are mainly visible in the tropics, in part because the precipitation is stronger there. Precipitation is excessive in the Indian Ocean in both winter and summer, especially over the Maritime Continent [consistent with Neale and Slingo (2003)] and for the monsoon precipitations [as in Prodhomme et al. (2014)]. In the Pacific, the double ITCZ bias is also found for precipitation in winter, although in the summer the ITCZ is shifted northward (Lin 2007; Bellucci et al. 2010).

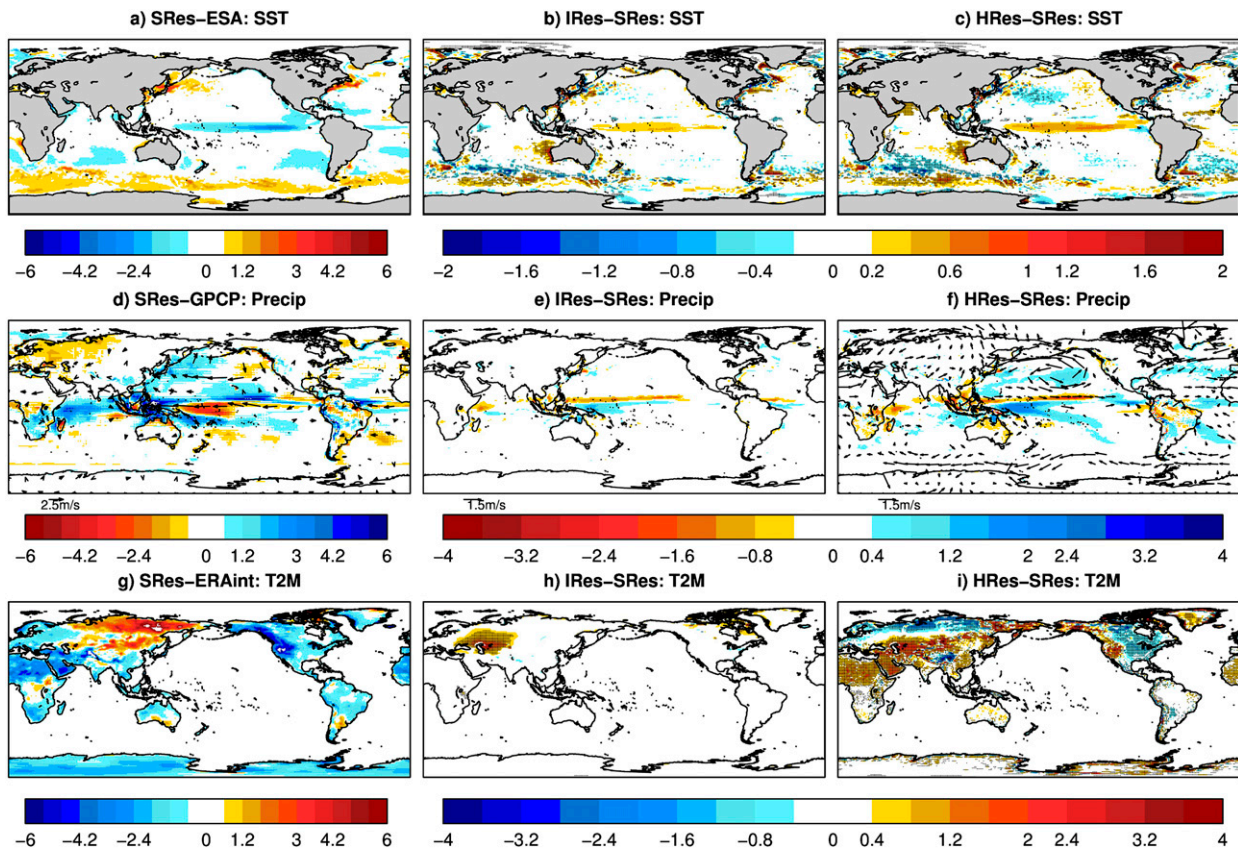


FIG. 2. As in Fig. 1, but for DJF (months 2 to 4 of the forecast initialized in November).

Figures 1b, 1c, 2b, and 2c show the bias differences between the SRes and the IRes and HRes experiments, respectively, for the SST. Increasing the resolution is in general beneficial for the SST bias. IRes and HRes show some important and similar improvements with respect to SRes: 1) the cold bias in summer in the North Pacific and North Atlantic basins, 2) the warm bias in the Somali upwelling, and 3) the cold tongue bias are all reduced, especially in boreal winter. This last improvement could be related to a better representation of tropical instability waves (Roberts et al. 2009). The higher-resolution atmosphere in HRes further reduces the cold tongue bias and the warm bias in the Somali upwelling, due to the better representation of the air–sea coupling and Ekman pumping (Chowdary et al. 2016).

For wind and precipitation the main changes occur when both oceanic and atmospheric resolutions are increased. In the Indian Ocean, in agreement with Prodhomme et al. (2014), the SST cooling, associated with the increase of oceanic resolution, leads to a decrease of the excessive oceanic precipitation (Figs. 1c,f). Probably because of the improved orography, the bias over the Maritime Continent is also reduced, which is consistent with Love et al. (2011) and Schiemann et al. (2014).

To quantify the change in terms of mean state when the resolution is increased, Table 3 shows the percentage of Earth's surface where the bias is reduced in HRes, with respect to SRes for SST, T2M, and precipitation. This table shows that for all variables the bias is reduced for more than half of Earth's surface, so according to this metric the mean state is improved when the resolution is increased, which confirms the results of Figs. 1 and 2 described above. To conclude, the increase of oceanic resolution leads to an improvement of the SST mean state but does not affect the precipitation, winds, and temperature. The combination of both oceanic and atmospheric resolution increases improves the representation of the mean state of all the considered variables. This encouraging result is consistent with the literature (Jung et al. 2012; Sakamoto et al. 2012) and needs to be further investigated from a physical point of view in order to clearly attribute processes that could be responsible for these improvements. While this will be the topic of future studies, we want here to stress that these improvements have been obtained with minimal tuning of the IRes and HRes configurations, opening promising perspectives when it will come to calibrate high-resolution GCMs in a more thorough way. The next

TABLE 3. Surface of the globe where the bias in HRes is reduced with respect to SRes is shown in bold. The improvement is further decomposed for the area where the bias in SRes is negative (in italics) and positive (normal font).

	SST	Land 2-m temperature	Precipitation	Land precipitation
May start date	37.1% + 24.4% = 61.5%	34.0% + 23.6% = 57.6%	29.8% + 20.6% = 50.5%	33.3% + 26.2% = 59.5%
November start date	39.1% + 21.1% = 60.2%	50.4% + 1% = 60.7%	32.1% + 23.2% = 55.3%	34.5% + 25.6% = 60.2%

section will investigate how the prediction skill of the system is affected by the resolution changes.

b. Impact on the forecast quality

The left columns of Figs. 3 and 4 show the ACC time correlation for the SRes experiment for SST, T2M, and precipitation with respect to the observation for JJA and DJF reforecasts, respectively. The correlation is computed for each grid point of the reference observed dataset so as to give a fair comparison of the skill for all the resolutions after applying a conservative interpolation of the reforecasts. EC-Earth3 exhibits standard correlation levels, comparable to other state-of-the-art seasonal forecasting systems (Doblas-Reyes

et al. 2013). For all variables, scores are the highest over the tropical area, especially in the tropical Pacific basin, related to the El Niño–Southern Oscillation (ENSO) predictability (Figs. S1 and S2; Phelps et al. 2004; Landman and Beraki 2012; Doblas-Reyes et al. 2013). The skill of precipitation is lower than for SST, a common feature of dynamical seasonal forecast systems (Figs. 3 and 4; Doblas-Reyes et al. 2013). In summer, the correlation is generally lower than for DJF, associated with a weaker ENSO predictability (Figs. 4 and 5), due to the well-known spring predictability barrier (Chen et al. 2004; Duan and Wei 2013).

The right columns of Figs. 3 and 4 show for the two seasons the difference in ACC time correlation between

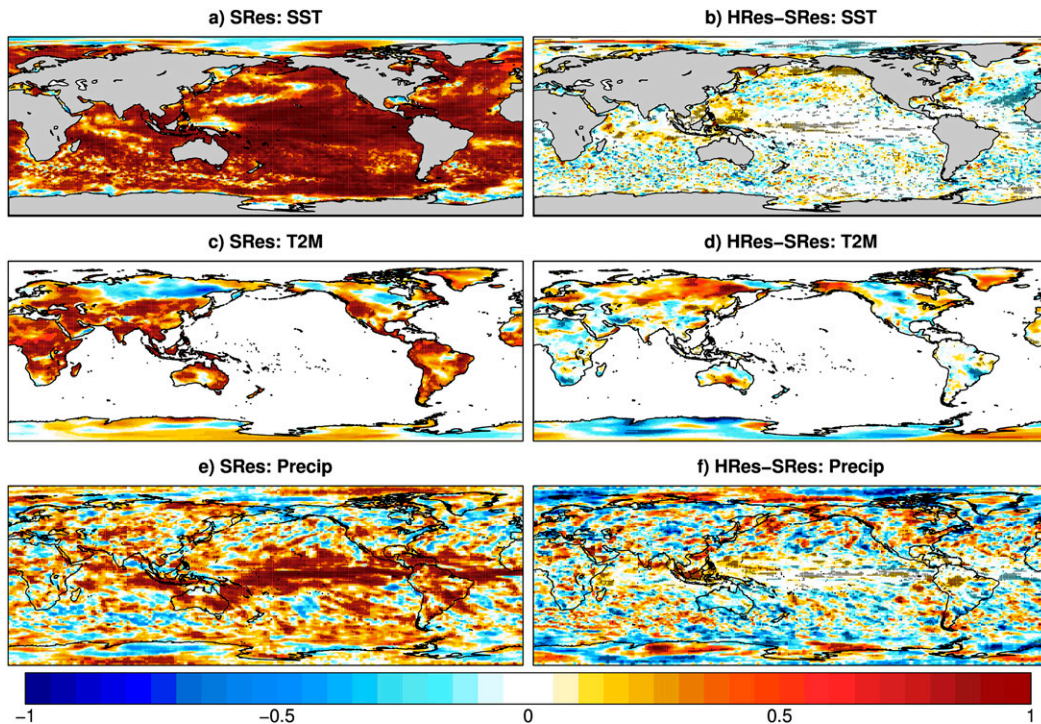


FIG. 3. (a) Correlation of JJA (months 2 to 4 of the forecasts initialized in May) ensemble mean SST in SRes with respect to ESA SST. Dots show the area where the correlation is significant at the 95% confidence level. (b) Difference of correlation of the ensemble mean between the HRes and SRes experiments for the JJA SST. The correlations are computed with respect to the ESA SST. Dots show the area where the difference of correlation is significant at 95% confidence level. (c),(d) As in (a),(b), but for near-surface atmosphere temperature. The correlations are computed with respect to the ERA-Interim 2-m temperature. (e),(f) As in (a),(b), but for precipitation. The correlations are computed with respect to the GPCP precipitation.

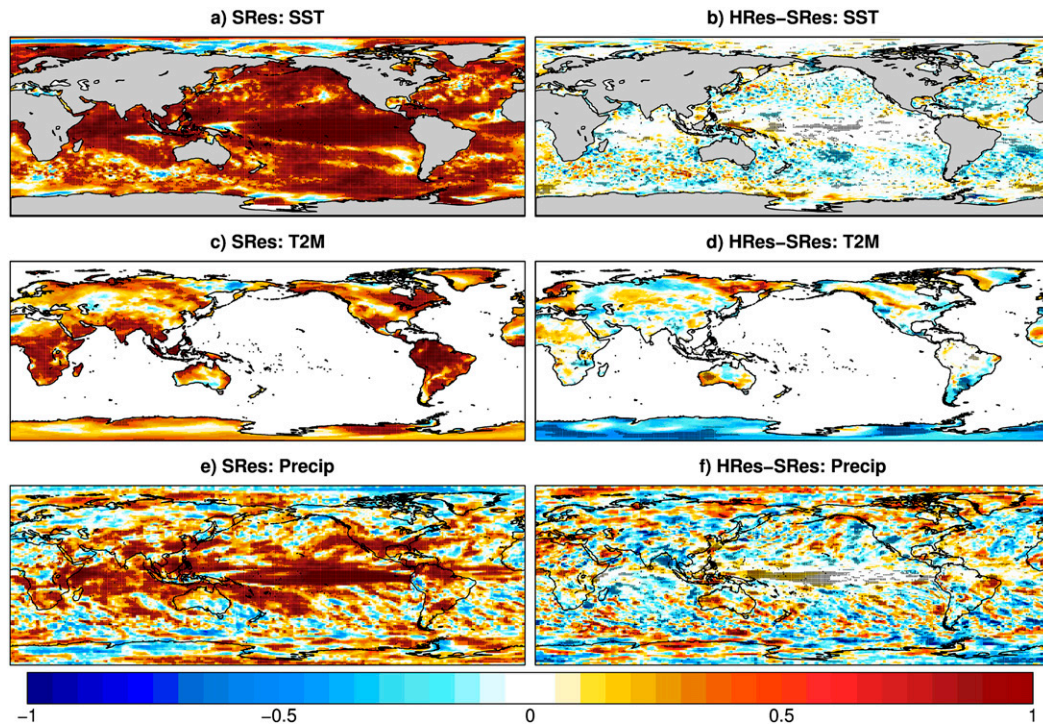


FIG. 4. As in Fig. 3, but for DJF (months 2 to 4 of the forecast initialized in November).

HRes and SRes. To have an objective metric of the changes, Table 4 shows the percentage of the globe for which the skill is increased in HRes with respect to SRes and the percentage of the globe with correlation significant at 95% confidence level in SRes and HRes. In general, the changes are patchy, for approximately half of Earth's surface, the skill is higher in HRes than SRes and the area with significant skill remains equivalent in the two reforecasts. For both seasons, we find few areas of significant change in correlation by increasing the resolution (Figs. 3b,d,f and 4b,d,f). The ACC decreases in the Southern Ocean, especially in boreal winter, as well as in the Atlantic Ocean and in the polar regions in the summer hemisphere (Figs. 3b, d,f and 4b,d,f). In boreal summer, the main areas of improvement common for the three considered variables, are the equatorial Indian Ocean, Maritime Continent, and western equatorial Pacific (Figs. 3b,d,f), which is consistent with the observed bias reduction in the region (see section 3a). This suggests that the improved resolution in the Indo-Pacific leads to a better representation of physical processes in the region, for example the upwelling (Akuetevi et al. 2016) and gravity waves in the Maritime Continent (Love et al. 2011). Large correlation increases are also noticed for T2M, for both JJA and DJF, in the north of Russia, Europe, and Alaska in the high-resolution

experiment (Figs. 3b,d,f and 4b,d,f), where the biases also decrease in HRes compared to SRes (Fig. 1). It is worth restating here that despite the significant correlation increase the ACC might remain insignificant in some regions. This correlation increase could be associated with a better representation of the main variability modes in these regions, such as blocking or the NAO. This will be investigated in further detail in section 4b.

To conclude, it appears that increasing the resolution does not affect substantially skill at the gridpoint level. This conclusion holds with a different skill metric such as RMSE (see supplementary Figs. S3 and S4). In some areas, including large impact areas such as the Indian Ocean and the Indian monsoon region, some skill increase is visible. However, it is worth taking into account that the comparison between standard and high-resolution experiments is not completely fair in the sense that the standard resolution has undergone substantial tuning, while the same settings with no further tuning are used for the high-resolution configurations.

To better understand the skill changes that occur when the resolution is increased and investigate these changes at a subseasonal scale, the next section will provide a deeper assessment of skill and forecast quality for selected regions of interest.

TABLE 4. In bold, surface of the globe where the ACC time correlation in HRes is increased with respect to SRes. In each cell, line 2 (line 3) shows the relative surface of the planet for which the correlation is significant at the 95% confidence level in SRes (HRes).

	SST	Land 2-m temperature	Precipitation	Land precipitation
May	47.8%	53.7%	50.6%	50.3%
start date	SRes: 70.4% HRes: 69.3%	SRes: 33.5% HRes: 35.7%	SRes: 20.9% HRes: 21.8%	SRes: 16.2% HRes: 18.3%
November	43.2%	48.3%	51.6%	55.4%
start date	SRes: 64.3% HRes: 59.8%	SRes: 36.8% HRes: 36.1%	SRes: 27.6% HRes: 27.6%	SRes: 20.1% HRes: 22.7%

4. Regional skill and forecast quality assessment

a. Tropical variability

1) ENSO AND THE ATLANTIC NIÑO

The main source of skill at seasonal time scales is the ENSO phenomenon. From the maps presented in the previous section, the tropical Pacific does not stand out as a region where skill is improved. Still, dots showing significant differences are visible in the equatorial Pacific (Figs. 3 and 4). To evaluate the changes occurring in this highly relevant region, we look more specifically at the skill of the Niño-3.4 index (SST averaged over 5°S–5°N, 190°–240°E), which is generally used to assess the skill of a system to forecast ENSO. Figure 5 shows that the resolution increase does improve the ENSO skill. The skill increase occurs for both summer and winter seasons and during the whole forecast length, although the increase is stronger in summer. Increasing only the oceanic resolution already improves the ENSO skill and the benefit is reinforced when the resolution of both components is increased, which is consistent with the strong coupled nature of the ENSO phenomenon. This improvement is robust when considering both deterministic scores such as correlation with different observational products (Figs. 5a,e), RMSE (Figs. 5b,f), and probabilistic skill scores such as the Brier score (Figs. 5c,d,g,h). The Brier score is smaller (therefore better) in all cases considered in HRes and/or IRes with respect to SRes, due to improvements in both resolution and reliability components, except for the DJF Niño-3.4 index reaching above the second tercile. In this case (Fig. 5h) the reliability is better and the reliability curve is closer to the perfect reliability diagonal, but the resolution score decreases. As shown in Table 1, the increase of atmospheric and/or oceanic resolution is not always associated with the improved representation of ENSO (MacLachlan et al. 2015; Jia et al. 2015; Zhu et al. 2015). From this small number of studies using different atmospheric and oceanic resolutions and grid ratios (Table 1), it is hard to extract one common argument explaining why the skill increases in some models and

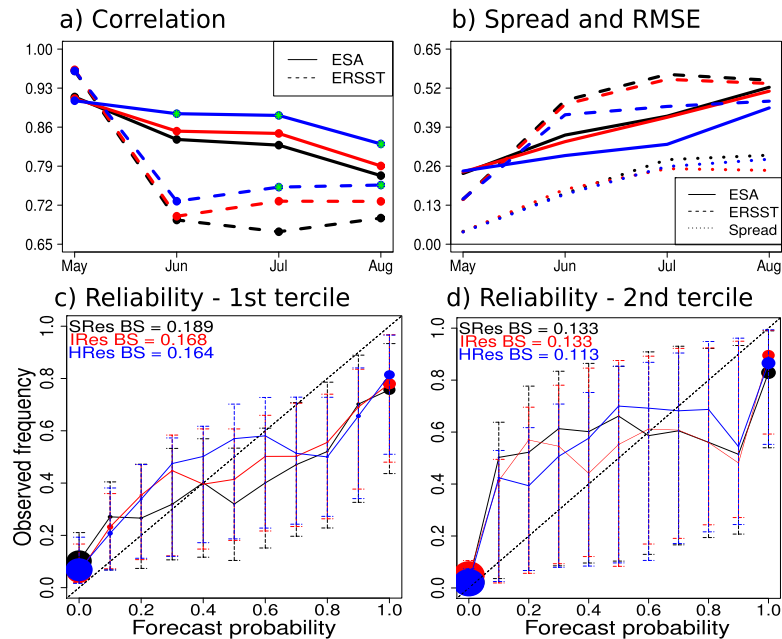
not in others. These improvements might be linked to the mean state improvement in the tropical Pacific (Figs. 1 and 2) and also to the better representation of high-frequency and small-scale coupled processes associated to increased horizontal but also vertical resolution (Masson et al. 2012). In order to better understand why ENSO is affected by resolution changes, other sensitivity experiments, including seasonal forecast and preindustrial control simulations with different horizontal and vertical resolutions in the ocean and atmosphere, in particular an experiment with resolution increased only in the atmosphere component, would be needed. The robustness of these results must also be assessed with larger simulations, since 17 start dates and 10 members is relatively limited.

The skill of the Atlantic Niño and the West African monsoon does not increase with model resolution (Fig. S5). This result might be expected for several reasons. First, in the literature the tropical Atlantic and surrounding precipitation have not been highlighted as affected by resolution changes (see Table 1). Then, the strongly biased tropical Atlantic mean state (Fig. 1) is not affected by resolution increase, which suggests that no change in the representation of this region occurs when the resolution increases. Exarchou et al. (2016, manuscript submitted to *Climate Dyn.*) investigate in detail the mechanisms leading to the formation of the biases in the EC-Earth 3 model in this region and conclude that biases are associated with the misrepresentation of cloud cover in the Angola–Benguela region and around the equator due to an erroneous subtropical overturning cell. Those two processes are not expected to be strongly impacted by resolution changes. As expected in the absence of improvement in the tropical Atlantic, the skill of the West African monsoon average precipitation is not clearly improved when the resolution is increased (Fig. S5).

2) INDIAN MONSOON AND INDIAN OCEAN

Several studies reported an improved representation of the Indian monsoon with an increase of resolution (Table 1; Sperber et al. 1994; Lal et al. 1997; Gent et al.

May start dates



November start dates

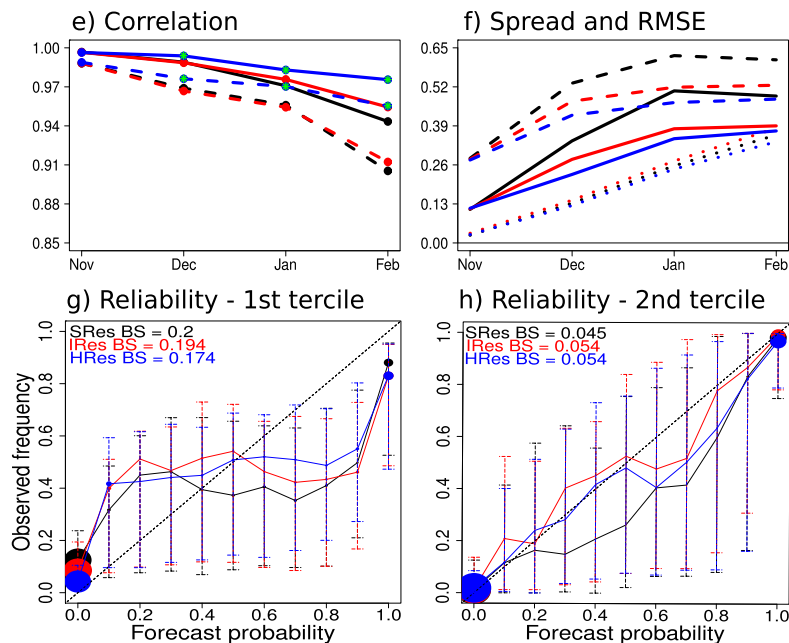


FIG. 5. Comparison of Niño-3.4 skill between SRes (black), IRes (red), and HRes (blue). (a) Correlation in the Niño-3.4 region between SRes, IRes, and HRes and the observation (ESA, plain line and ERSST, dashed line) for May start dates. Dots show the correlation significant at 95% confidence level and green stars show the correlation in IRes and HRes significantly different from SRes at 95% confidence level. (b) RMSE in the Niño-3.4 region of SRes, IRes, and HRes with respect to the observation (ESA, plain line and ERSST, dashed line) for May start dates. The dotted lines show the spread of the different experiment with the same color as the RMSE. (c) Reliability diagram for Niño-3.4 SST below the first tercile (La Niña-like events) for JJA 1993–2009; the corresponding Brier score is shown in the top left corner. Error bars indicate significance intervals at 95% computed with a bootstrapping method. The size of the dots is proportional to the size of the population bin for each forecast probability. (d) As in (c), but for the SST above the second tercile (El Niño-like events). (e)–(h) As in (a)–(d), but for November start dates.

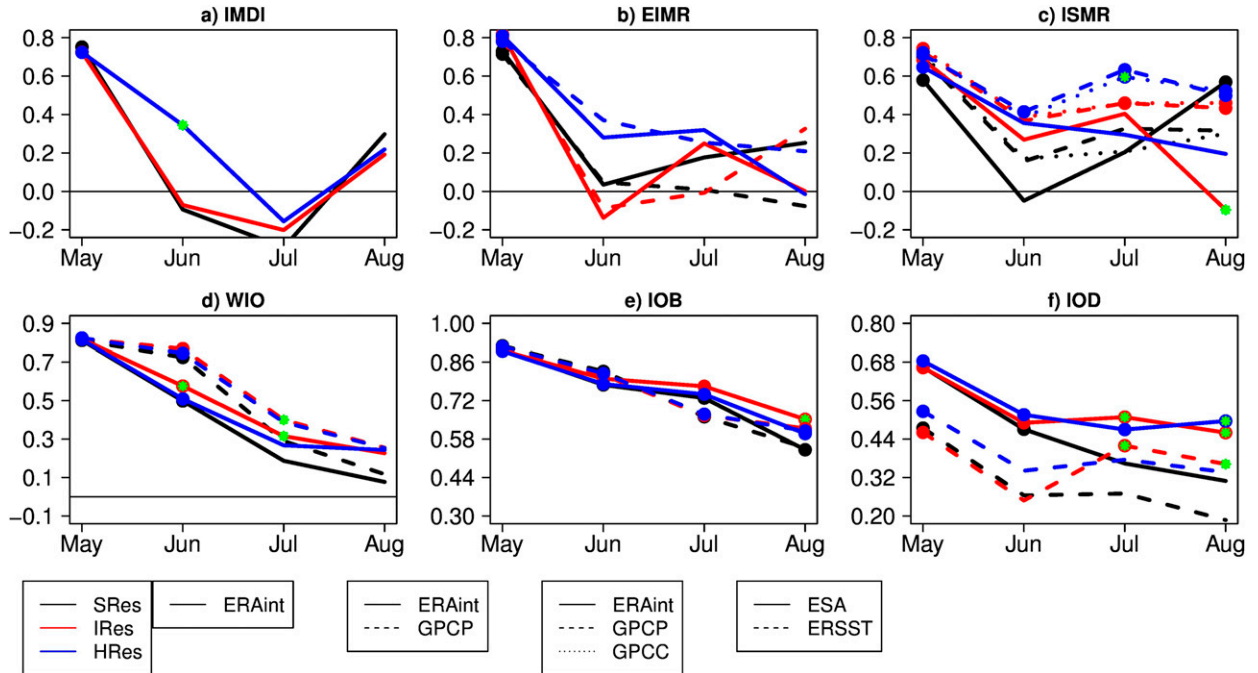


FIG. 6. Correlation of different indices as a function of forecast time. For all figures, dots show the correlation significant at 95% confidence level and green stars show the correlation in IRes and HRes significantly different from SRes at 95% confidence level. (a) Indian monsoon dynamical index (IMDI): zonal wind at 850 hPa averaged in the region 5°–15°N, 40°–80°E minus the wind at averaged in the region 20°–30°N, 70°–90°E. (b) Extended Indian monsoon rainfall (EIMR): precipitation averaged in the region 10°–30°N, 70°–110°E. (c) Indian summer monsoon rainfall (ISMR): precipitation averaged over land in the region 5°–30°N, 70°–95°E. (d) Western Indian Ocean (WIO): SST averaged in the region 10°S–10°N, 60°–80°E. (e) Indian Ocean basin (IOB): SST averaged in 20°S–20°N, 40°–110°E. (f) SST averaged in the Arabian Sea (10°–30°N, 40°–80°E).

2010; Delworth et al. 2012; Johnson et al. 2016). In addition, the monsoon and the monsoon onset date are known to be strongly influenced by ENSO (Boschat et al. 2011; Prodhomme et al. 2015); therefore improvement of ENSO representation might lead to improvement in the Indian monsoon. Figures 6a–c show three different monsoon indices: the Indian summer monsoon dynamical index (IMDI; zonal wind at 850 hPa averaged in the region 5°–15°N, 40°–80°E minus the zonal wind averaged in the region 20°–30°N, 70°–90°E; Wang et al. 2001), the extended Indian monsoon rainfall (5°–30°N, 70°–95°E), and the Indian summer monsoon rainfall (continental precipitation over the region 5°–30°N, 70°–95°E). All these indices show an improvement of skill with the increase of resolution at the beginning of the monsoon season. The improvement of early monsoon skill could be associated with an improvement of the forecast of the monsoon onset date, one of the most relevant variables of the monsoon. To confirm this hypothesis we have estimated the monsoon onset date based on the criteria of Wang and LinHo (2002), following the methodology described in section 2c. We obtain the results described in Table 5, which shows the estimated mean threshold, the mean, the standard deviation, and the ACC of the onset

date in the different simulations. The three simulations show a relatively high correlation for this precipitation-based monsoon onset date, close to other studies of monsoon onset predictability (Alessandri et al. 2014). The skill of the onset date is increased in IRes and HRes compared to SRes, and the highest skill is obtained in the IRes reforecast; however, the changes are not significant at the 95% confidence level.

To show the increase of monsoon skill, Figs. 6d–f show the skill of three indices known to influence the

TABLE 5. Onset date characteristics estimated from the criteria of Wang and LinHo (2002), which is the day for which precipitation from ERA-Interim averaged over India exceeds 5 mm day⁻¹ (in the ISMR region, continental precipitation over the region 5°–30°N, 70°–95°E). The quantile corresponding to this value is estimated with cross-validation for the month of May and July in the different simulations, the average corresponding threshold is displayed in the table. Then the onset date in the model is estimated as the first day exceeding this quantile in the model.

	ERA-Interim	SRes	IRes	HRes
Mean	26 May	17 May	17 May	18 May
Threshold (mm day ⁻¹)	5	4.46	4.32	4.97
Std dev (days)	10.9	6.9	7.5	6.39
Correlation		0.61	0.69	0.65

monsoon and especially the monsoon onset: the SST in the Arabian Sea (Levine and Turner 2012), in the western Indian Ocean (WIO; Prodhomme et al. 2014), and in the Indian Ocean basin (IOB; Boschat et al. 2011; Prodhomme et al. 2015). The skill of these three indices is slightly increased in IRes and HRes; as for the monsoon onset date, the skill is the highest in IRes. This last point requires deeper investigation. This increase of skill in the Indian Ocean and the increase of ENSO skill could explain the improvement of the monsoon predictability in IRes and HRes.

The results suggest that higher resolution in the ocean and/or atmosphere can improve the simulation of the Indo-Pacific SST, which seems to have a positive impact on the Indian monsoon predictability. To better understand the sources of skill coming from the resolution increase and disentangle the effects of large-scale and local processes, an assessment of the large-scale dynamics of the monsoon, vertical and horizontal wind shear in the Indian Ocean, north–south tropospheric temperature gradient, dynamical onset date (Xavier et al. 2007), and the local processes in the Indian Ocean, such as the intraseasonal oscillation (Goswami 2005), would be needed. The robustness of the small improvements of the onset date skill should also be tested by comparing different indices (Fasullo and Webster 2003; Xavier et al. 2007; Wang et al. 2009; Bombardi and Carvalho 2009).

b. Atmospheric circulation at midlatitudes

1) NORTH ATLANTIC OSCILLATION

Impacts of increasing ocean and atmosphere resolution are assessed over the Northern Hemisphere midlatitude both in terms of variability (Northern Hemisphere blocking index) and skill in forecasting interannual variations of the main climate indices (e.g., NAO).

The NAO accounts for a substantial part of the large-scale atmospheric circulation over the North Atlantic region at seasonal-to-interannual time scales. Several studies have evaluated the (generally limited) forecast skill of GCMs of the seasonal NAO index (e.g., Doblas-Reyes et al. 2003; Arribas et al. 2011; Kim et al. 2012). Recently, higher-resolution GCMs were found to have significant skill in forecasting the NAO (Scaife et al. 2014). Butler et al. (2016) showed by studying the Climate-System Historical Forecast Project database that models with a well-resolved stratosphere tend to better represent responses to ENSO and the quasi-biennial oscillation (QBO) in upper levels of the atmosphere, but that this does not necessarily translate into boreal winter NAO skill improvements. The NAO skill assessment is highly uncertain when using limited ensemble

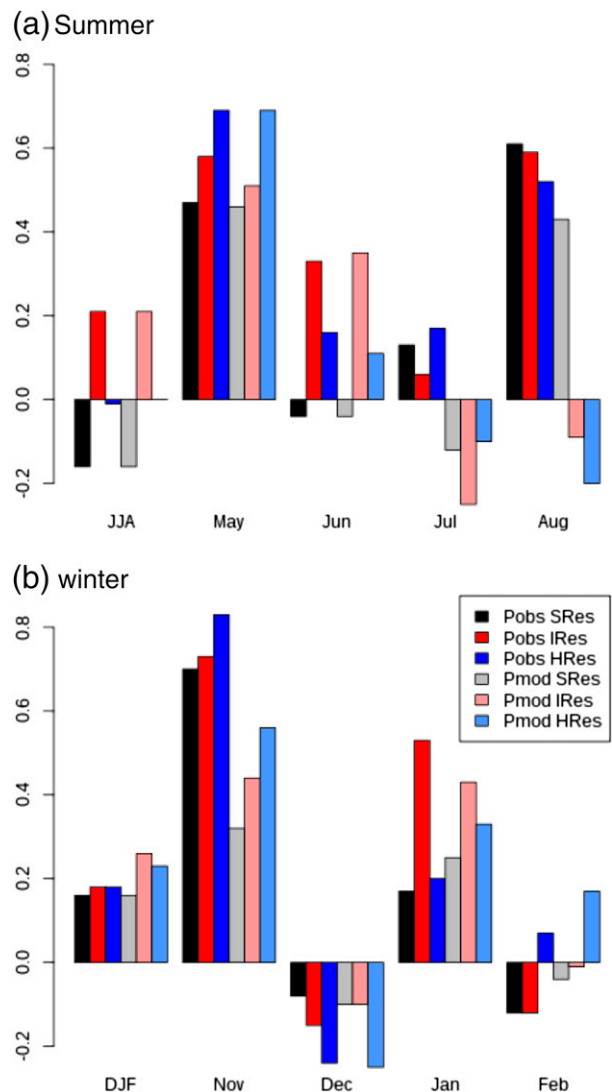


FIG. 7. Correlation of the NAO index in boreal (a) summer and (b) winter reforecasts for months 2–4 and computed month by month for SRes (black), IRes (red), and HRes (blue), using both Pobs (dark colors) and Pmod (pale colors) computations.

sizes and short reforecast periods (e.g., Shi et al. 2015). We therefore chose to evaluate the NAO reforecast skill using several NAO indices (pressure difference, Pmod and Pobs; see section 2c for more detail), and focused on both seasonal (months 2–4 average) skill and monthly skill up to four months lead.

NAO index correlation coefficients with respect to ERA-Interim data for boreal winter and summer in the three experiments studied here are shown in Fig. 7 for Pobs and Pmod NAO index calculations. Results from Fig. 7 can be summarized as follows. The NAO prediction skill in EC-Earth is quite poor with standard resolution over the 1993–2009 reforecast period for both

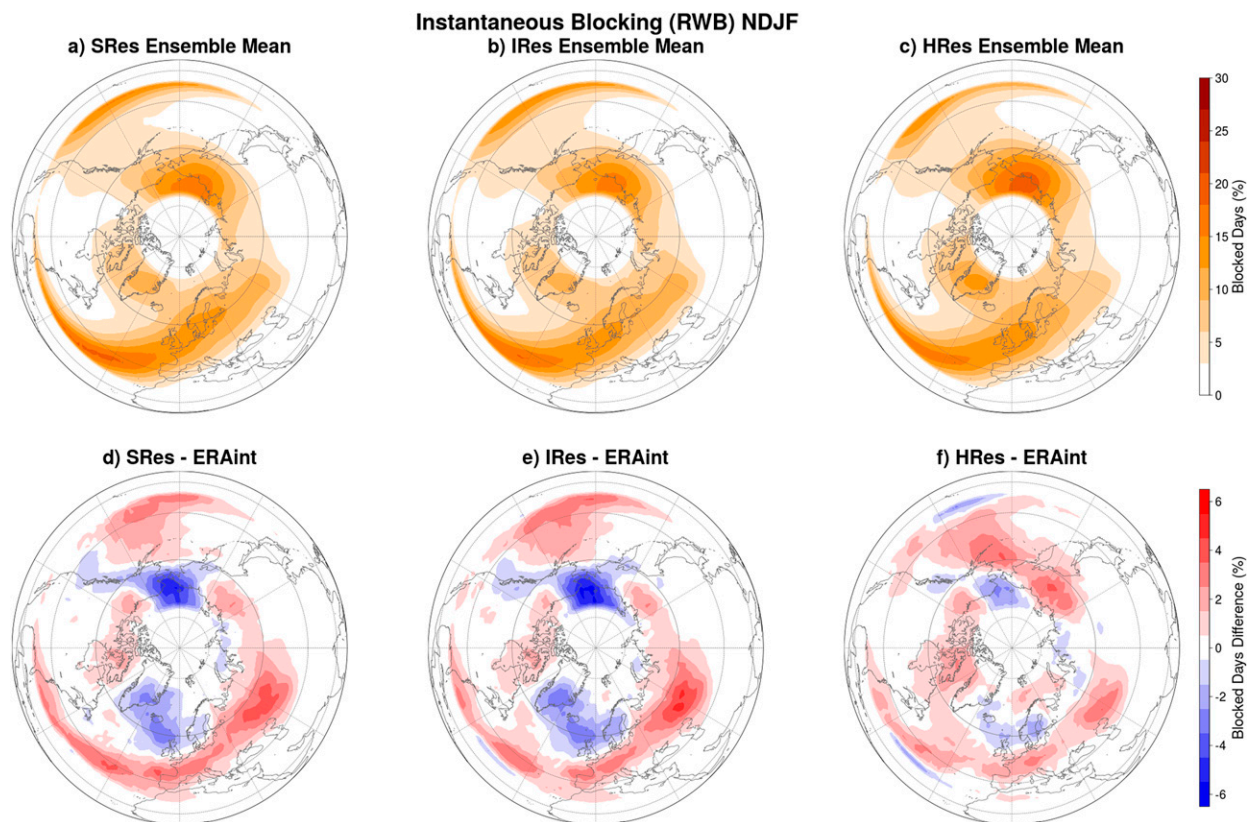


FIG. 8. (a) Instantaneous blocking frequency averaged over the full period (1993–2009) expressed as a percentage of blocked days in the SRes experiment. (b) As in (c), but for IRes. (c) As in (a), but for HRes. (d) Differences of blocking frequency expressed as a percentage of blocked days in the SRes experiment with respect to months; interim reanalysis. (e) As in (d), but for IRes. (f) As in (d), but for HRes.

boreal winter and summer. Increasing resolution leads to some improvements in NAO correlation, especially for boreal winter (Fig. 7b). For boreal summer (Fig. 7a) results are more contrasted, since for higher forecast times (those exceeding month 3) the NAO correlation drops when increasing resolution.

Results are similar for both methods shown in the figure, and confirmed in the case of boreal winter by computing area-averaged sea level pressure differences as in Stephenson et al. (2006) (not shown). However it is crucial to note that uncertainty intervals for these scores are very large and differences are often not significant.

2) ATMOSPHERIC BLOCKING

Atmospheric blocking simulation is a common issue in state-of-the-art GCMs. The most recent generation of GCMs still exhibits large biases, especially over the European region (Anstey et al. 2013). To evaluate the winter atmospheric blocking variability, the instantaneous blocking index computed as in Davini et al. (2012) is hereafter used.

Figure 8 shows the average bias for the full winter season (November–February) for the SRes, IRes, and HRes

experiments. SRes shows limited bias over both North Atlantic and Pacific; however, the classic negative bias over Europe is clearly present. Increasing the oceanic resolution does not lead to any evident improvement; conversely, increasing the atmospheric resolution leads to a reduction of the bias over Europe and the North Pacific. This is in agreement with experiments showing a notable improvement of blocking over the Euro-Atlantic region following an increase in horizontal resolution (e.g., Matsueda et al. 2009; Jung et al. 2012; Dawson et al. 2012). Indeed, there is mounting evidence that reduced Euro-Atlantic blocking bias is associated with better-resolved transient eddy activity, which can sustain blocking persistence (Berckmans et al. 2013), and with a higher orography variance (especially relevant over the Rocky Mountains), which affects the mean tilt of the Atlantic eddy-driven jet, favoring higher geopotential height values over Europe (Jung et al. 2012; Berckmans et al. 2013).

On the other side, the lack of improvements following an increase in oceanic resolution is likely associated with the persistent SST biases in the North Atlantic region, which can be seen in both SRes and IRes (Fig. 2). Although minor improvements can be observed, there are

small changes in the SST midlatitude frontal zone (and in general over North Atlantic) between the low and high oceanic resolution experiments. This partially contrasts with the work of Scaife et al. (2011), where increased oceanic resolution leads to improved blocking. However, we must keep in mind that the November initializations and the short-term duration of the current simulations certainly play a significant role. They indeed keep both the models far from their own attractor, potentially reducing the benefits of a better resolved oceanic circulation.

Regarding the predictive skill, an extremely weak signal is found (not shown). Greenland blocking shows weak skill in agreement with results for the NAO, but nothing significant emerges over Europe. Considering the improvement in the mean state for the HRes experiments, it is likely that a larger number of ensemble members will be needed to achieve forecast skill for blocking over Europe.

c. Polar regions

Seasonal sea ice prediction is a growing field of research (Chevallier et al. 2013; Wang et al. 2013). In the Arctic, forecasting the sea ice conditions a few months in advance is of great strategic relevance (Emerson and Lahn 2012; U.S. Navy 2012). As an example of response to these challenges, the Sea Ice Prediction Network (SIPN; www.arcus.org/sipn) has collected, each year since 2008, submissions from various groups using a range of methods in order to predict the upcoming summer sea ice conditions in the Arctic. Predicting the seasonal development of Antarctic sea ice is arguably less interesting from an end-user perspective, especially over the winter season, but is still challenging from a scientific point of view given the recent chain of high records of sea ice extent in 2012 (Turner et al. 2013), 2013 (Reid et al. 2015), and 2014 (Massonnet et al. 2015b). The present section attempts to assess systematically how resolution can impact the skill and biases of sea ice seasonal prediction.

We focus here on the fourth month of prediction (August and February) as the sea ice cover exhibits strong seasonality reaching its maximum and minimum extensions in late local winter or summer, respectively. (Note that September is conventionally preferred to August, but our simulations do not extend that far). We define sea ice extent as the cumulative area of ocean grid cells where sea ice concentration exceeds 15%, and compute this diagnostic from the monthly-mean sea ice concentration field, for various sectors of the planet.

We note first that changing the resolution does not have an obvious impact (positive or negative) on model bias (Fig. S6). This bias is systematically larger between any experiment and the observations than the difference between any two experiments, for all sectors and start dates

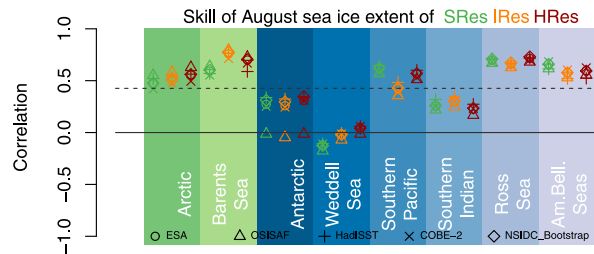


FIG. 9. Correlation of 1993–2008 August forecasts (initialized in May) for the three experiments SRes (green), IRes (orange), and HRes (red), against five observational products (symbols) in seven sectors: Arctic [0° – 90° N], Barents Sea [68° – 82° N, 18° – 70° E], Antarctic [0° – 90° S], Weddell Sea [0° – 90° S, 60° W– 20° E], southern Pacific [0° – 90° S, 90° – 160° E], southern Indian Ocean [0° – 90° S, 20° – 90° E], Ross Sea [0° – 90° S, 130° W– 160° E] and Amundsen–Bellinghousen Seas [0° – 90° S, 60° – 13° W]. The plain horizontal line is the zero line, and the dashed line the threshold above which correlations are significant at 5% confidence level.

considered. The results should thus be viewed as “conservative” in the sense that high-resolution does not degrade existing model biases; it remains to be established whether fine tuning can eventually reduce them. In either case, the message is straightforward: resolution alone does not spontaneously improve the mean state of simulated Arctic and Antarctic sea ice, at least at the seasonal time scale. Improving the estimation of ocean–sea ice initial conditions would likely play a more important role to reduce basinwide and regional biases, as suggested by recent studies (Guemas et al. 2016; Massonnet et al. 2015a).

We then consider the ability of the model to predict interannual variations of August sea ice extent. Because of secular negative (positive) trends in sea ice extent in the Arctic (Antarctic), we focus on detrended time series to avoid overinterpretation of model skill. Figure 9 reports correlation of detrended time series of sea ice extent in August in several sectors of the Arctic and Southern Oceans for the three experiments SRes, IRes, and HRes, and using five sea ice concentration products as observational references. This makes a total of 40 triplets (120) of correlations. We find the HRes experiment to rank highest 19 times. If all tests were independent from each other (which is questionable since observations and sea ice extent in sectors are well correlated with one another), a result as extreme as this one could happen 4.5% of the time just by chance. An inspection of actual time series of sea ice extent (Fig. S6) reveals that correlation changes are the consequence of hardly noticeable changes in the time series. Using alternative observational products alters the estimated correlations, by amounts sometimes as large as differences across experiments themselves, but preserves rankings anyway (Fig. 9). If not spectacular, the overall improvement of the scores, consistent among observational products, when resolution increases is a reassuring

TABLE 6. Summary of the changes associated with an increase of resolution in the EC-Earth seasonal predictions. Improvements are marked with (++), degradations with (--), and results showing no specific changes are marked with (==).

	Mean state	Skill
Global maps	(++) For all variables the percentage of Earth's surface where the bias is reduced is higher than 50%	(--) The skill is not improved at the grid point level
ENSO and Pacific SST	(++) Reduction of the cold tongue bias (++) Better representation of the Kuroshio current	(++) Increase of the ENSO skill
Indian Ocean SST	(++) Improvement of the Somalian upwelling	(++) Slight improvement of skill of western Indian Ocean, Arabian Sea SST, and IOB (stronger in IRes)
Atlantic SST	(++) Reduction of the warm bias in summer in the North Atlantic (++) Better representation of the Gulf Stream (==) No improvement in the tropical Atlantic	(++) Decrease of skill in the northeastern Atlantic (--) Decrease of the skill in the equatorial Atlantic
Indian monsoon	(++) Reduction of the excessive oceanic precipitation over the Indian Ocean. (==) Seasonal cycle of the monsoon is not improved (not shown)	(++) Improvement of the skill of dynamical and precipitation monsoon index and slight improvement of the prediction of the monsoon onset date (especially in IRes)
African monsoon	(==) No improvement (not shown)	(==) No improvement
Atmospheric blocking	(++) Improvement of the blocking representation	(==) No skill
North Atlantic Oscillation		(++) Improvement for winter and beginning of summer (--) Degradation for August
Sea ice	(==) No difference. Model biases remain unchanged.	(++) Slight improvement for boreal summer prediction. (--) Slight degradation for boreal winter prediction

result. However, the difference of correlation between HRes and SRes are not significant at 95% confidence level in most of the cases.

A similar analysis conducted for the target month of February (not shown here) reveals a totally different picture. The SRes run ranks highest in 20 cases out of 40, which would have happened 2.1% of the time by chance. By a symmetrical argument, we conclude that using a higher-resolution, but untuned configuration degrades the ability of the model to forecast sea ice extent variations. Sea ice extent in the Arctic in winter is largely controlled by oceanic heat convergence (Bitz et al. 2005), and we hypothesize that a simple change in resolution is not sufficient to transport the adequate amounts of heat northward—at least when the same parameterizations as the ones for the SRes configuration are used.

When considering both the mean state and variability, it is apparent that our three simulations resemble more each other than they resemble observations. This indicates that resolution alone cannot fully address long-standing biases, and that other ongoing developments (modeling, new parameterizations, generation of better initial conditions) have to come into play.

5. Conclusions

In the present study, we have compared three different simulations run with the same version of the EC-Earth

coupled model. The simulations differ only in their oceanic and atmospheric resolutions. This unique set of simulations offers the possibility to investigate systematically the benefits of resolution for seasonal climate prediction. It is not possible to extract a common conclusion for all regions, seasons, variables, and lead times considered in this paper, as results vary highly depending on the process investigated. We first summarize cases where resolution is clearly beneficial, then examples where it leaves baseline prediction capabilities unchanged (or even deteriorates them), and finally cases for which the assessment is uncertain.

The increase of oceanic and/or atmospheric resolution in EC-Earth 3 generally leads to an improved representation of the mean state of the reforecast. The SST is more sensitive to resolution increase in the ocean model, whereas atmospheric variables (precipitation, near-surface wind, T2M over land) are improved only if atmospheric resolution is also increased. Some improvement of the mean state, especially in the tropical Indo-Pacific, is associated with an improvement of skill in this region for both SST (Niño-3.4, IOB, western Indian Ocean, and Arabian Sea) and monsoon simulation (onset date, dynamical index, and rainfall). Some minor improvements are also visible in the sea ice and NAO predictability, as well as the representation of the blocking indices. All those improvements, summarized in Table 6, suggest that, despite the absence of specific

tuning, increasing the resolution is a way forward to improve both the simulation of the mean climate and the predictability of high-impact phenomena. Other phenomena have been shown to be improved with resolution increase and have not been assessed in this study, such as tropical cyclones (Caron et al. 2011; Manganello et al. 2016) or polar jet (Guemas and Codron 2011; Hourdin et al. 2013).

However, despite these improvements, the SST skill is decreased in the high-resolution simulation over more than half of the ocean surface. This is especially true in the tropical Atlantic, where both mean state and skill are deteriorated in the high-resolution runs compared to the standard resolution. Most coupled models are strongly biased in the Atlantic (Tonizzo and Woolnough 2014); the PREFACE (Enhancing Prediction of Tropical Atlantic climate and its Impacts) EU FP7 project, which partly funded this study, aims to better understand the reasons underlying the formation of those biases. Our results illustrate the limitations to increasing resolution as a means of improving seasonal forecasts, particularly in the case of the tropical Atlantic. In the scope of the project, different studies are in preparation tackling these biases with different approaches, such as wind stress and flux corrections, assessment of the daily evolution of the model drift in seasonal reforecasts, and CMIP5 multi-model analysis. It is essential to study together seasonal forecast and long-term simulations to better understand how both systematic error and variability of the long run could be related to seasonal forecast drift and skill.

Some hints of improvement of the representation of the North Atlantic Oscillation and blocking indices have been found when the resolution is increased; however the limited ensemble size and number of start dates considered here are not sufficient to assess robustly the changes in the forecasting skill of these phenomena. These considerations open a common debate in the seasonal forecasting community: whether resources should be used to increase resolution or ensemble size. Running the HRes experiment is approximately 19 times more expensive than the SRes experiments. In addition, other strategies, such as introducing stochastically perturbed parameterization tendencies (SPPT) perturbations in the atmosphere could lead to comparable increase of ENSO skill to the one obtained in HRes (Batté and Doblas-Reyes 2015). It is legitimate to ask if such an expensive experiment is justified given the amplitude of the observed improvements. To answer this question it is worth investigating if an equivalent tuning to the one performed for the low-resolution configuration would help the HRes version significantly improve its performances. A substantial effort should be done for the tuning of the high-resolution

version of EC-Earth to answer this question, ideally following an objective framework applied to both model resolutions (Bellprat et al. 2012).

As the horizontal resolution increases in both the ocean and the atmosphere, GCMs move closer to the so-called “grey zone” of spatial scales for which classic physical parameterizations may no longer be adapted. As shown by Jung et al. (2012), an upper limit to improvements due to increasing resolution in current-generation GCMs is to be expected, and research is now underway to design scale-aware parameterizations that could help push back limitations in climate forecasting, provided that appropriate computational resources are available (e.g., Stan and Xu 2014; Campin et al. 2011). At the spatial resolutions considered in this study (from $\sim 1^\circ$ to $\sim 0.25^\circ$), model biases and forecast skill are in general moderately affected by the sole changes in resolution, indicating that further process-based tuning will be necessary for the newly developed high-resolution versions of the model. Yet, in some particular cases (ENSO prediction, Indian Ocean), these simple changes in resolution are sufficient to significantly advance the forecast system capabilities. Developing and tuning global forecast systems at higher resolutions is thus of utmost importance for pushing the boundaries of seasonal prediction.

Acknowledgments. The research leading to these results has received funding from the EU Seventh Framework Programme FP7 (2007–2013) under Grant Agreements 308378 (SPECS), 603521 (PREFACE), and 607085 (EUCLEIA), the Horizon 2020 EU program under Grant Agreements 641727 (PRIMAVERA) and 641811 (IMPRES), and the ESA Climate Change Initiative (CCI) Living Planet Fellowship VERITAS-CCI. We acknowledge PRACE for awarding access to Marenostrum3 based in Spain at the Barcelona Supercomputing Center through the HiResClim project. We acknowledge the work of the developers of the s2dverification R-based package (<http://cran.r-project.org/web/packages/s2dverification/index.html>) and autosubmit workflow manager (<https://pypi.python.org/pypi/autosubmit/3.5.0>). Paolo Davini acknowledges the funding from the European Union’s Horizon 2020 research and innovation programme COGNAC under the European Union Marie Skłodowska-Curie Grant Agreement 654942.

REFERENCES

- Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, **4**, 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

- Akuetevi, C. Q. C., B. Barnier, J. Verron, J.-M. Molines, and A. Lecointre, 2016: Interactions between the Somali Current eddies during the summer monsoon: Insights from a numerical study. *Ocean Sci.*, **12**, 185–205, doi:10.5194/osd-12-735-2015.
- Alessandri, A., A. Borrelli, A. Cherchi, S. Materia, A. Navarra, J.-Y. Lee, and B. Wang, 2014: Prediction of Indian summer monsoon onset using dynamical subseasonal forecasts: Effects of realistic initialization of the atmosphere. *Mon. Wea. Rev.*, **143**, 778–793, doi:10.1175/MWR-D-14-00187.1.
- Anstey, J. A., and Coauthors, 2013: Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution. *J. Geophys. Res. Atmos.*, **118**, 3956–3971, doi:10.1002/jgrd.50231.
- Arribas, A., and Coauthors, 2011: The GloSea4 ensemble prediction system for seasonal forecasting. *Mon. Wea. Rev.*, **139**, 1891–1910, doi:10.1175/2010MWR3615.1.
- Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. K. Betts, 2009: A revised hydrology for the ECMWF Model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeorol.*, **10**, 623–643, doi:10.1175/2008JHM1068.1.
- Batté, L., and F. J. Doblas-Reyes, 2015: Stochastic atmospheric perturbations in the EC-Earth3 global coupled model: Impact of SPPT on seasonal forecast quality. *Climate Dyn.*, **45**, 3419–3439, doi:10.1007/s00382-015-2548-7.
- Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär, 2012: Objective calibration of regional climate models. *J. Geophys. Res.*, **117**, D23115, doi:10.1029/2012JD018262.
- Bellucci, A., S. Gualdi, and A. Navarra, 2010: The double-ITCZ syndrome in coupled general circulation models: The role of large-scale vertical circulation regimes. *J. Climate*, **23**, 1127–1145, doi:10.1175/2009JCLI3002.1.
- Berkmans, J., T. Woollings, M. E. Demory, P. L. Vidale, and M. Roberts, 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.
- Bitz, C. M., M. M. Holland, E. C. Hunke, and R. E. Moritz, 2005: Maintenance of the sea-ice edge. *J. Climate*, **18**, 2903–2921, doi:10.1175/JCLI3428.1.
- Bombardi, R. J., and L. M. V. Carvalho, 2009: IPCC global coupled model simulations of the South America monsoon system. *Climate Dyn.*, **33**, 893–916, doi:10.1007/s00382-008-0488-1.
- Boschat, G., P. Terray, and S. Masson, 2011: Interannual relationships between Indian summer monsoon and Indo-Pacific coupled modes of variability during recent decades. *Climate Dyn.*, **37**, 1019–1043, doi:10.1007/s00382-010-0887-y.
- Branković, Č., and D. Gregory, 2001: Impact of horizontal resolution on seasonal integrations. *Climate Dyn.*, **18**, 123–143, doi:10.1007/s003820100165.
- Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3, doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.
- Butler, A. H., and Coauthors, 2016: The Climate-system Historical Forecast Project: Do stratosphere-resolving models make better seasonal climate predictions in boreal winter? *Quart. J. Roy. Meteor. Soc.*, **142**, 1413–1427, doi:10.1002/qj.2743.
- Campin, J.-M., C. Hill, H. Jones, and J. Marshall, 2011: Superparameterization in ocean modeling: Application to deep convection. *Ocean Modell.*, **36**, 90–101, doi:10.1016/j.ocemod.2010.10.003.
- Caron, L.-P., C. G. Jones, and K. Winger, 2011: Impact of resolution and downscaling technique in simulating recent Atlantic tropical cyclone activity. *Climate Dyn.*, **37**, 869–892, doi:10.1007/s00382-010-0846-7.
- Chen, D., M. A. Cane, A. Kaplan, S. E. Zebiak, and D. Huang, 2004: Predictability of El Niño over the past 148 years. *Nature*, **428**, 733–736, doi:10.1038/nature02439.
- Chevallier, M., D. Salas y Méliá, A. Voldoire, M. Déqué, and G. Garric, 2013: Seasonal Forecasts of the pan-Arctic sea ice extent using a GCM-based seasonal prediction system. *J. Climate*, **26**, 6092–6104, doi:10.1175/JCLI-D-12-00612.1.
- Chowdary, J. S., A. Parekh, S. Ojha, C. Gnanaseelan, and R. Kakatkar, 2016: Impact of upper ocean processes and air-sea fluxes on seasonal SST biases over the tropical Indian Ocean in the NCEP Climate Forecasting System. *Int. J. Climatol.*, **36**, 188–207, doi:10.1002/joc.4336.
- Comiso, J. C., 2000: Variability and trends in Antarctic surface temperatures from in situ and satellite infrared measurements. *J. Climate*, **13**, 1674–1696, doi:10.1175/1520-0442(2000)013<1674:VATIAS>2.0.CO;2.
- Davini, P., C. Gualdi, S. Gualdi, and A. Navarra, 2012: Bidimensional diagnostics, variability, and trends of Northern Hemisphere blocking. *J. Climate*, **25**, 6496–6509, doi:10.1175/JCLI-D-12-00032.1.
- Dawson, A., and T. N. Palmer, 2014: Simulating weather regimes: Impact of model resolution and stochastic parameterization. *Climate Dyn.*, 2177–2193, doi:10.1007/s00382-014-2238-x.
- , —, and S. Corti, 2012: Simulating regime structures in weather and climate prediction models. *Geophys. Res. Lett.*, **39**, L21805, doi:10.1029/2012GL053284.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.
- Delworth, T. L., and Coauthors, 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *J. Climate*, **25**, 2755–2781, doi:10.1175/JCLI-D-11-00316.1.
- Doblas-Reyes, F. J., V. Pavan, and D. B. Stephenson, 2003: The skill of multi-model seasonal forecasts of the wintertime North Atlantic Oscillation. *Climate Dyn.*, **21**, 501–514, doi:10.1007/s00382-003-0350-4.
- , J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: Status and prospects. *Wiley Interdiscip. Rev.: Climate Change*, **4**, 245–268, doi:10.1002/wcc.217.
- Du, H., F. J. Doblas-Reyes, J. García-Serrano, V. Guemas, Y. Soufflet, and B. Wouters, 2012: Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations. *Climate Dyn.*, **39**, 2013–2023, doi:10.1007/s00382-011-1285-9.
- Duan, W., and C. Wei, 2013: The “spring predictability barrier” for ENSO predictions and its possible mechanism: Results from a fully coupled model. *Int. J. Climatol.*, **33**, 1280–1292, doi:10.1002/joc.3513.
- Eastwood, S., T. Lavergne, and R. Tonboe, 2014: Algorithm theoretical basis document for the OSI SAF global reprocessed sea ice concentration product. EUMETSAT Network Satellite Application Facilities, 28 pp. [Available online at http://osisaf.met.no/docs/osisaf_cdop2_ss2_atbd_sea-ice-conc-reproc_v1p1.pdf.]
- Emerson, C., and G. Lahn, 2012: Arctic opening: Opportunity and risk in the high north. Chatham House and Lloyds, 60 pp. [Available online at <http://www.chathamhouse.org/publications/papers/view/182839>.]
- Fasullo, J., and P. J. Webster, 2003: A hydrological definition of Indian monsoon onset and withdrawal. *J. Climate*, **16**, 3200–3211, doi:10.1175/1520-0442(2003)016<3200a:AHDOIM>2.0.CO;2.
- Ferry, N., L. Parent, G. Garric, B. Barnier, and N. Jourdain, 2010: Mercator Global Eddy Permitting Ocean Reanalysis

- GLORYS1V1: Description and results. *Mercator Ocean Quarterly Newsletter*, Vol. 36, International CLIVAR Project Office, Southampton, United Kingdom, 15–27.
- Fichefet, T., and M. A. Morales Maqueda, 1997: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *J. Geophys. Res.*, **102**, 12 609–12 646, doi:10.1029/97JC00480.
- Forbes, R., A. M. Tompkins, and A. Untch, 2011: A new prognostic bulk microphysics scheme for the IFS. ECMWF Tech. Memo. 649, 28 pp. [Available online at [http://www.ecmwf.int/en/elibrary/9441-new-prognostic-bulk-microphysics-scheme-ifs.](http://www.ecmwf.int/en/elibrary/9441-new-prognostic-bulk-microphysics-scheme-ifs)]
- García-Serrano, J., F. J. Doblas-Reyes, J. R. Haarsma, and I. Polo, 2013: Decadal prediction of the dominant West African monsoon rainfall modes. *J. Geophys. Res. Atmos.*, **118**, 5260–5279, doi:10.1002/jgrd.50465.
- Gent, P. R., S. G. Yeager, R. B. Neale, S. Levis, and D. A. Bailey, 2010: Improvements in a half degree atmosphere/land version of the CCSM. *Climate Dyn.*, **34**, 819–833, doi:10.1007/s00382-009-0614-8.
- Goswami, B. N., 2005: South Asian monsoon. *Intraseasonal Variability of the Atmosphere–Ocean Climate System*, K. M. Lau and D. E. Waliser, Eds., Springer Praxis, 19–61.
- Guemas, V., and F. Codron, 2011: Differing impacts of resolution changes in latitude and longitude on the midlatitudes in the LMDZ atmospheric GCM. *J. Climate*, **24**, 5831–5849, doi:10.1175/2011JCLI4093.1.
- , F. J. Doblas-Reyes, F. Lienert, Y. Soufflet, and H. Du, 2012: Identifying the causes of the poor decadal climate prediction skill over the North Pacific. *J. Geophys. Res.*, **117**, D20111, doi:10.1029/2012JD018004.
- , and Coauthors, 2016: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **142**, 546–561, doi:10.1002/qj.2401.
- Guilyardi, E., and Coauthors, 2004: Representing El Niño in coupled ocean–atmosphere GCMs: The dominant role of the atmospheric component. *J. Climate*, **17**, 4623–4629, doi:10.1175/JCLI-3260.1.
- Hazeleger, W., and Coauthors, 2012: EC-Earth V2.2: Description and validation of a new seamless Earth system prediction model. *Climate Dyn.*, **39**, 2611–2629, doi:10.1007/s00382-011-1228-5.
- Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75, doi:10.1175/JCLI-D-12-00837.1.
- Hourdin, F., and Coauthors, 2013: Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Climate Dyn.*, **40**, 2167–2192, doi:10.1007/s00382-012-1411-3.
- Ivanova, N., and Coauthors, 2015: Inter-comparison and evaluation of sea ice algorithms: Towards further identification of challenges and optimal approach using passive microwave observations. *Cryosphere*, **9**, 1797–1817, doi:10.5194/tc-9-1797-2015.
- Jia, L., and Coauthors, 2015: Improved seasonal prediction of temperature and precipitation over land in a high-resolution GFDL climate model. *J. Climate*, **28**, 2044–2062, doi:10.1175/JCLI-D-14-00112.1.
- Johnson, S. J., and Coauthors, 2016: The resolution sensitivity of the South Asian monsoon and Indo-Pacific in a global 0.35° AGCM. *Climate Dyn.*, **46**, 807–831, doi:10.1007/s00382-015-2614-1.
- Jung, T., and Coauthors, 2012: High-resolution global climate simulations with the ECMWF model in Project Athena: Experimental design, model climate, and seasonal forecast skill. *J. Climate*, **25**, 3155–3172, doi:10.1175/JCLI-D-11-00265.1.
- Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere winter. *Climate Dyn.*, **39**, 2957–2973, doi:10.1007/s00382-012-1364-6.
- Kirtman, B. P., and Coauthors, 2012: Impact of ocean model resolution on CCSM climate simulations. *Climate Dyn.*, **39**, 1303–1328, doi:10.1007/s00382-012-1500-3.
- Lal, M., U. Cubasch, J. Perlwitz, and J. Waszkewitz, 1997: Simulation of the Indian monsoon climatology in ECHAM3 climate model: Sensitivity to horizontal resolution. *Int. J. Climatol.*, **17**, 847–858, doi:10.1002/(SICI)1097-0088(19970630)17:8<847::AID-JOC165>3.0.CO;2-6.
- Landman, W. A., and A. Beraki, 2012: Multi-model forecast skill for mid-summer rainfall over southern Africa. *Int. J. Climatol.*, **32**, 303–314, doi:10.1002/joc.2273.
- Levine, R. C., and A. G. Turner, 2012: Dependence of Indian monsoon rainfall on moisture fluxes across the Arabian Sea and the impact of coupled model sea surface temperature biases. *Climate Dyn.*, **38**, 2167–2190, doi:10.1007/s00382-011-1096-z.
- Lin, J. L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean–atmosphere feedback analysis. *J. Climate*, **20**, 4497–4525, doi:10.1175/JCLI4272.1.
- Liu, H., C. Wang, S.-K. Lee, and D. Enfield, 2012: Atlantic warm-pool variability in the IPCC AR4 CGCM simulations. *J. Climate*, **25**, 5612–5628, doi:10.1175/JCLI-D-11-00376.1.
- Love, B. S., A. J. Matthews, and G. M. S. Lister, 2011: The diurnal cycle of precipitation over the Maritime Continent in a high-resolution atmospheric model. *Quart. J. Roy. Meteor. Soc.*, **137**, 934–947, doi:10.1002/qj.809.
- MacLachlan, C., and Coauthors, 2015: Global Seasonal Forecast System version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, doi:10.1002/qj.2396.
- Madec, G., 2008: NEMO ocean engine. Note du Pole de modélisation 27, Institut Pierre-Simon Laplace, 209 pp.
- Maidens, A., A. Arribas, A. A. Scaife, C. MacLachlan, D. Peterson, and J. Knight, 2013: The influence of surface forcings on prediction of the North Atlantic Oscillation regime of winter 2010/11. *Mon. Wea. Rev.*, **141**, 3801–3813, doi:10.1175/MWR-D-13-00033.1.
- Manganello, J. V., and Coauthors, 2016: Seasonal forecasts of tropical cyclone activity in a high-atmospheric-resolution coupled prediction system. *J. Climate*, **29**, 1179–1200, doi:10.1175/JCLI-D-15-0531.1.
- Manubens-Gil, D., J. Vegas-Regidor, C. Prodhomme, O. Mula-Valls, and F. J. Doblas-Reyes, 2016: Seamless management of ensemble climate prediction experiments on HPC platforms. *2016 Int. Conf. on High Performance Computing & Simulation (HPCS)*, Innsbruck, Austria, IEEE, 895–900, doi:10.1109/HPCSim.2016.7568429.
- Martin, G. M., 1999: The simulation of the Asian summer monsoon, and its sensitivity to horizontal resolution, in the UK Meteorological Office Unified Model. *Quart. J. Roy. Meteor. Soc.*, **125**, 1499–1525, doi:10.1002/qj.49712555703.
- Masson, S., P. Terray, G. Madec, J. J. Luo, T. Yamagata, and K. Takahashi, 2012: Impact of intra-daily SST variability on ENSO characteristics in a coupled model. *Climate Dyn.*, **39**, 681–707, doi:10.1007/s00382-011-1247-2.
- Massonnet, F., T. Fichefet, and H. Goosse, 2015a: Prospects for improved seasonal Arctic sea ice predictions from

- multivariate data assimilation. *Ocean Modell.*, **88**, 16–25, doi:10.1016/j.ocemod.2014.12.013.
- , V. Guemas, N. S. Fučkar, and F. J. Doblas-Reyes, 2015b: The 2014 high record of Antarctic sea ice extent. *Bull. Amer. Meteor. Soc.*, **96**, S163–S167, doi:10.1175/BAMS-D-15-00093.1.
- Matsueda, M., R. Mizuta, and S. Kusunoki, 2009: Future change in wintertime atmospheric blocking simulated using a 20-km-mesh atmospheric global circulation model. *J. Geophys. Res.*, **114**, D12114, doi:10.1029/2009JD011919.
- Mizielinski, M. S., and Coauthors, 2014: High-resolution global climate modelling: The UPSCALE project, a large-simulation campaign. *Geosci. Model Dev.*, **7**, 1629–1640, doi:10.5194/gmd-7-1629-2014.
- Morcrette, J.-J., H. W. Barker, J. N. S. Cole, M. J. Iacono, and R. Pincus, 2008: Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System. *Mon. Wea. Rev.*, **136**, 4773–4798, doi:10.1175/2008MWR2363.1.
- Neale, R., and J. Slingo, 2003: The Maritime Continent and its role in the global climate: A GCM study. *J. Climate*, **16**, 834–848, doi:10.1175/1520-0442(2003)016<0834:TMCAIR>2.0.CO;2.
- Phelps, M., A. Kumar, and J. O'Brien, 2004: Potential predictability in the NCEP CPC dynamical seasonal forecast system. *J. Climate*, **17**, 3775–3785, doi:10.1175/1520-0442(2004)017<3775:PPITNC>2.0.CO;2.
- Prodhomme, C., P. Terray, S. Masson, T. Izumo, T. Tozuka, and T. Yamagata, 2014: Impacts of Indian Ocean SST biases on the Indian monsoon: As simulated in a global coupled model. *Climate Dyn.*, **42**, 271–290, doi:10.1007/s00382-013-1671-6.
- , —, G. Boschat, and T. Izumo, 2015: Oceanic factors controlling the Indian summer monsoon onset in a coupled model. *Climate Dyn.*, **44**, 977–1002, doi:10.1007/s00382-014-2200-y.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- Reid, P., S. Stammerjohn, R. Massom, T. Scambos, and J. Lieser, 2015: The record 2013 Southern Hemisphere sea-ice extent maximum. *Ann. Glaciol.*, **56**, 99–106, doi:10.3189/2015AoG69A892.
- Roberts, M. J., and Coauthors, 2009: Impact of resolution on the tropical Pacific circulation in a matrix of coupled models. *J. Climate*, **22**, 2541–2556, doi:10.1175/2008JCLI2537.1.
- Sakamoto, T. T., and Coauthors, 2012: MIROC4h—A new high-resolution atmosphere–ocean coupled general circulation model. *J. Meteor. Soc. Japan*, **90**, 325–359, doi:10.2151/jmsj.2012-301.
- Scaife, A. A., and Coauthors, 2011: Improved Atlantic winter blocking in a climate model. *Geophys. Res. Lett.*, **38**, L23703, doi:10.1029/2011GL049573.
- , and Coauthors, 2014: Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.*, **41**, 2514–2519, doi:10.1002/2014GL059637.
- Schiemann, R., M. E. Demory, M. S. Mizielinski, M. J. Roberts, L. C. Shaffrey, J. Strachan, and P. L. Vidale, 2014: The sensitivity of the tropical circulation and Maritime Continent precipitation to climate model resolution. *Climate Dyn.*, **42**, 2455–2468, doi:10.1007/s00382-013-1997-0.
- Schneider, U., T. Fuchs, A. Meyer-Christoffer, and B. Rudolf, 2008: Global precipitation analysis products of the GPCC. Global Precipitation Climatology Centre, 12 pp.
- Shaffrey, L. C., and Coauthors, 2009: U.K. HiGEM: The new U.K. high-resolution global environment model—Model description and basic evaluation. *J. Climate*, **22**, 1861–1896, doi:10.1175/2008JCLI2508.1.
- Shi, W., N. Schaller, D. Macleod, T. N. Palmer, and A. Weisheimer, 2015: Impact of hindcast length on estimates of seasonal climate predictability. *Geophys. Res. Lett.*, **42**, 1554–1559, doi:10.1002/2014GL062829.
- Siegert, S., D. B. Stephenson, O. Bellprat, M. Ménégoz, and F. J. Doblas-Reyes, 2016: Detecting improvements in forecast correlation skill: Statistical testing and power analysis. *Mon. Wea. Rev.*, doi:10.1175/MWR-D-16-0037.1, in press.
- Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006). *J. Climate*, **21**, 2283–2296, doi:10.1175/2007JCLI2100.1.
- Sperber, K. R., S. Hameed, G. L. Potter, and S. J. Boyle, 1994: Simulation of the Northern summer monsoon in the ECMWF model: Sensitivity to horizontal resolution. *Mon. Wea. Rev.*, **122**, 2461–2481, doi:10.1175/1520-0493(1994)122<2461:SOTNSM>2.0.CO;2.
- Stan, C., and L. Xu, 2014: Climate simulations and projections with a super-parameterized climate model. *Environ. Modell. Software*, **60**, 134–152, doi:10.1016/j.envsoft.2014.06.013.
- Stephenson, D. B., and Coauthors, 2006: North Atlantic Oscillation response to transient greenhouse gas forcing and the impact on European winter climate: A CMIP2 multi-model assessment. *Climate Dyn.*, **27**, 401–420, doi:10.1007/s00382-006-0140-x.
- Tibaldi, S., and F. Molteni, 1990: On the operational predictability of blocking. *Tellus*, **42A**, 343–365, doi:10.1034/j.1600-0870.1990.t01-2-00003.x.
- Toniazzo, T., and S. Woolnough, 2014: Development of warm SST errors in the southern tropical Atlantic in CMIP5 decadal hindcasts. *Climate Dyn.*, **43**, 2889–2913, doi:10.1007/s00382-013-1691-2.
- Toth, Z., O. Talagrand, G. Candille, and Y. Zhu, 2003: Probability and ensemble forecasts. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. Joliffe and D. Stephenson, Eds., Wiley, 137–163.
- Turner, J., J. S. Hosking, T. Phillips, and G. J. Marshall, 2013: Temporal and spatial evolution of the Antarctic sea ice prior to the September 2012 record maximum extent. *Geophys. Res. Lett.*, **40**, 5894–5898, doi:10.1002/2013GL058371.
- U.S. Navy, 2012: Arctic Roadmap 2014–2030. Navy Task Force Climate Change, 43 pp. [Available online at http://www.navy.mil/docs/USN_arctic_roadmap.pdf.]
- Valcke, S., 2013: The OASIS3 coupler: A European climate modelling community software. *Geosci. Model Dev.*, **6**, 373–388, doi:10.5194/gmdd-5-2139-2012.
- van den Hurk, B. J. J. M., A. C. M. Beljaars, and A. K. Betts, 2000: Offline validation of the ERA-40 surface scheme. ECMWF Tech. Memo. 295, 42 pp.
- Vannière, B., E. Guilyardi, G. Madec, F. J. Doblas-Reyes, and S. Woolnough, 2013: Using seasonal hindcasts to understand the origin of the equatorial cold tongue bias in CGCMs and its impact on ENSO. *Climate Dyn.*, **40**, 963–981, doi:10.1007/s00382-012-1429-6.
- Voldoire, A., M. Claudon, G. Caniaux, H. Giordani, and R. Roehrig, 2014: Are atmospheric biases responsible for the tropical Atlantic SST biases in the CNRM-CM5 coupled model? *Climate Dyn.*, **43**, 2963–2984, doi:10.1007/s00382-013-2036-x.

- von Storch, H., and F. W. Zwiers, 2001: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Wahl, S., M. Latif, W. Park, and N. Keenlyside, 2011: On the tropical Atlantic SST warm bias in the Kiel Climate Model. *Climate Dyn.*, **36**, 891–906, doi:10.1007/s00382-009-0690-9.
- Wang, B., and LinHo, 2002: Rainy season of the Asian–Pacific summer monsoon. *J. Climate*, **15**, 386–398, doi:10.1175/1520-0442(2002)015<0386:RSOTAP>2.0.CO;2.
- , R. Wu, and K.-M. Lau, 2001: Interannual variability of the Asian summer monsoon: Contrasts between the Indian and the western North Pacific–East Asian monsoon. *J. Climate*, **14**, 4073–4090, doi:10.1175/1520-0442(2001)014<4073:IVOTAS>2.0.CO;2.
- , Q. Ding, and P. V. Joseph, 2009: Objective definition of the Indian summer monsoon onset. *J. Climate*, **22**, 3303–3316, doi:10.1175/2008JCLI2675.1.
- Wang, W., M. Chen, and A. Kumar, 2013: Seasonal prediction of arctic sea ice extent from a coupled dynamical forecast system. *Mon. Wea. Rev.*, **141**, 1375–1394, doi:10.1175/MWR-D-12-00057.1.
- Xavier, P. K., C. Marzina, and B. N. Goswami, 2007: An objective definition of the Indian summer monsoon season and a new perspective on the ENSO–monsoon relationship. *Quart. J. Roy. Meteor. Soc.*, **133**, 749–764, doi:10.1002/qj.45.
- Yang, X., and Coauthors, 2015: Seasonal predictability of extra-tropical storm tracks in GFDL’s high-resolution climate prediction model. *J. Climate*, **28**, 3592–3611, doi:10.1175/JCLI-D-14-00517.1.
- Zhu, J., and Coauthors, 2015: ENSO prediction in Project Minerva: Sensitivity to atmospheric horizontal resolution and ensemble size. *J. Climate*, **28**, 2080–2095, doi:10.1175/JCLI-D-14-00302.1.