Comparing the seasonal predictability of the Tropical Pacific variability in EC-Earth3 at two horizontal resolutions

Aude Carréric[1], Pablo Ortega[1], Roberto Bilbao[1], Carlos Delgado-Torres[1], Vladimir Lapin[1], Ferran Lopez-Marti[3,4], Markus Donat[1,2], and Francisco Doblas-Reyes[1,2]

[1]Barcelona Supercomputing Center (BSC), Barcelona, Spain
[2]ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain
[3]Department of Earth Sciences, Uppsala University, Uppsala, Sweden
[4]Centre of Natural Hazards and Disaster Science (CNDS), Uppsala, Sweden

Correspondence: Aude Carréric (aude.carreric@bsc.es)

Abstract. Seasonal predictability is an active field of research given its strong potential to guide decision-making in many societal and economic sectors. In this study, we compare the predictive skill of the climate model EC-Earth3 at two different horizontal resolutions. The standard resolution – SR – (high resolution – HR) is of around 70 (40) km in the atmosphere and 100 (25) km in the ocean. Both forecast systems are initialised in the same way in May and cover the period 1990-2015, with a forecast period of 8 months. We focus on the ~~Tropics~~Tropical Pacific, and particularly on El ~~Niño~~ Niño Southern Oscillation (ENSO), the main source of predictability at seasonal timescales. ~~Statistically~~ Small but statistically significant improvements are found in HR with respect to SR for predicting ENSO~~. However, the predictive~~ , skill improvements that cannot be generalized to all regions and initialization times. In the May initialised predictions, skill drops quickly in the Western Equatorial Pacific (WEP) in both configurations, more pronouncedly in SR. The poor skill in the WEP is directly linked to a misrepresentation of its relationship with the ENSO region, which is ultimately associated with an overly strong westward extension of ENSO-related variability, a model error more pronounced in SR. This erroneous spatial simulation of ENSO is related to the mean cold bias of the cold tongue, which progressively extends westwards with the forecast time. ~~We show that an~~ An overly weak air-sea coupling, more pronounced in SR, prevents the model from simulating the correct ENSO development. We also show that a ~~better~~ more realistic simulation of the Atlantic ~~Niño~~ Niño teleconnection with the tropical Pacific in HR compared to SR leads to better ENSO prediction. ~~Improving~~ Increasing model resolution can ~~increase~~ improve the predictive skill of forecast systems in certain regions and for certain seasons by improving the simulation of the mean state and atmospheric teleconnections. However, ENSO simulation errors and mean state biases need to be better understood to improve forecasts, in particular in the WEP, a region of convection particularly important for teleconnections to extratropics.

1 Introduction

Seasonal prediction is an area of research attracting growing interest beyond the scientific community due to its strong potential for informing stakeholders on timescales ranging from months to seasons in many sectors e.g. agriculture and food security, health, energy production, water management, disaster risk reduction (e.g., **??**). The usability of these predictions rely on their

accuracy and skill. Seasonal predictability can arise from two major sources. The first is related to the representation of the externally forced signals (such as from volcanic eruptions, solar activity or anthropogenic greenhouse gases), which have caused important climate trends in recent decades and will shape the near-term future. The second is internal variability, usually associated with oceanic, sea ice and land processes operating from monthly to multi-decadal timescales (**?**, fig.2). The premise of seasonal prediction is that such internal variability processes, when adequately ~~modeled~~ modelled and initialized, can improve our predictive capacity by shaping the atmospheric evolution (**??**). However, seasonal prediction is still a challenging field, both in terms of meeting users' needs for reliable and accurate forecasts (**??**) and in terms of better understanding the sources of predictability in order to improve model forecasts, particularly dynamical forecasts based on General Circulation Models (GCMs) whose components are becoming increasingly complex (**??**).

Seasonal predictions derive most of their predictive skill from the tropical regions, especially from El Niño Southern Oscillation (ENSO) dynamics (e.g., **???**). ENSO has significant weather impacts worldwide mediated via different atmospheric teleconnections (**?**) that could lead to economic impacts and sovereign risks (**??**). Significant progress has been made in ENSO prediction over the last decades, with current dynamical prediction systems providing skilful predictions of ENSO events up to 6-12 months ahead (**????**). However, ENSO predictive skill shows important variations depending on the initial ENSO state of the prediction (**??**), on the targeted event (**??**), on the initialisation month due to the ENSO "spring predictability barrier" (**??**), and on the phase of the equatorial Atlantic variability (**?**), among other factors. Also, ENSO skill is modulated by decadal variability (**???**). The ENSO phase in turn has an impact on the predictability of near-surface temperature and precipitation over remote regions like Europe and Africa, including for monsoon systems, multiple seasons to years in advance (**??**).

Dynamical seasonal forecasts, based on GCMs, are affected by ~~important~~ model and initialisation errors, which can limit their applicability. Appropriate simulation of the observed mean state is essential to correctly represent the inter-annual variability (**??**) that affects the skill of forecast systems (**????**). In GCMs, the ENSO mean state is affected by two common biases: (1) a basin-wide cold bias and (2) a too westward extension of the cold tongue region in the tropical Pacific compared to observations, the so-called "cold tongue bias" (**??**). These common mean state biases are reflected in the misrepresentation of the spatial structure of El Niño events, typically characterized by positive SST anomalies that extend too far into the west and ~~that take overly low values~~ whose amplitudes are underestimated compared to observations in the cold tongue region (**???**). Furthermore, these two biases are intertwined, with the magnitude of the cold bias favouring a more westward extension of the cold tongue via oceanic zonal advection feedbacks via oceanic zonal advection feedbacks (**?**). Dynamical seasonal forecasts are also hindered by these GCMs errors in the tropical Pacific (**????**), which alter the pattern and amplitude of ENSO events (**?**), but also their transition (**??**), and the associated wintertime atmospheric teleconnections (**?**). Using 11 different operational seasonal forecast systems, **?** showed that these ENSO-related forecast errors develop in a few days after initialisation and are mainly driven by the seasonal cycle rather than the forecast lead time. **?** argued that the strong cold bias in the cold tongue region simulated in the CESM1 seasonal forecast system emerged as a consequence of an initialisation adjustment, which produced an imbalance between the ocean subsurface and the model dynamics.

Among the various approaches being considered to improve the skill of dynamical seasonal prediction systems, increasing the horizontal resolution of the climate models they are based on is a promising avenue. Enhancing the resolution can help

to explicitly represent fundamental processes, such as ocean mesoscale eddies, that are parametrized at the standard model configurations that use grid spacings of around 100 km. The better simulation of the ocean mesoscale eddies, which in the Tropics requires grid spacings in the ocean of 25 km or finer, can improve the representation of key air-sea teleconnection mechanisms (**??**), leading to reduced common mean-state biases of global models. Current high resolution modelling efforts are particularly focused on understanding the role of these oceanic mesoscale processes in the simulated climate mean state and variability (**??????**). Regarding the tropical Pacific region, several recent studies using high ($\sim$25 km in the ocean) to very-high ($\sim$10 km in the ocean) resolution climate models showed an improved tropical Pacific mean state with a reduced cold tongue bias (**?????**). Focusing specifically on ENSO, **?** showed that higher ocean horizontal resolution configurations in a multi-model ensemble had a reduced bias in the equatorial Pacific mean state which led to a better simulation of the spatial patterns of ENSO-related anomalies. However, increasing the atmospheric horizontal resolution alone was found to deteriorate ENSO asymmetry. From the reduction of mean-state biases with higher resolution, we can expect improved skill in predicting certain regions at seasonal scales, e.g. in tropical sea surface temperature (**?**). **?** analysed the performance of the new European Centre of Medium-Range Weather Forecasts (ECMWF) operational seasonal forecast system, SEAS5, which in addition to upgraded versions of the atmospheric and oceanic components, also featured a finer resolution compared to the previous version SEAS4 (100 versus 25 km in the ocean, 80 versus 36 km in the atmosphere). They showed an improvement in the cold tongue bias in SEAS5 that led to better air and sea surface temperature predictive skill over the Tropics, although given the various upgrades between systems, it is not possible to know whether the better performance arises solely from the increase in resolution.

In this study, we explore the sensitivity of the forecasts to horizontal resolution, both in terms of skill and forecast biases, using two different seasonal prediction systems based on the climate model EC-Earth3 (**?**). ~~Both systems are initialized following exactly the same strategy, and only differ in the horizontal grid spacing considered:~~ We use a Standard Resolution (SR) configuration based on an eddy-parametrised ocean, and a High Resolution (HR) version based on an eddy-permitting ocean, their atmospheric configurations being set up to match the respective ocean resolutions as closely as possible. Both systems are initialized following exactly the same strategy, so that forecast differences can be directly attributed to the change in resolution, which involves both changes to the grid spacing and parametrization choices. Differences between systems can be attributed to the change in process parametrization adapted for each resolution, to the tuning of the model and to the initial conditions. We focus on the Tropical Pacific region where prominent improvements are found for predicting ENSO in the HR version, trying to shed light on the reasons behind the improvements. The paper is structured as follows: Section 2 describes the seasonal forecast systems, the observational data used as references and the methods. In Section 3, we assess the prediction skill differences between the SR and HR forecast systems in the tropical Pacific and investigate two different physical mechanisms that can explain the skill differences. The main conclusions of the study and a discussion of the results are provided in Section 4.

2 Methods and data

2.1 EC-Earth3 climate model

We use the EC-Earth v3.3 GCM (**?**) in its atmosphere-ocean-sea ice configuration. Two resolutions are considered: a SR configuration, with a mean resolution of around 70 km and 100 km in the atmosphere and the ocean, respectively, and a HR configuration, with a resolution of around 40 km and 25 km in the atmosphere and the ocean, respectively. The EC-Earth3.3 model includes three major components: the atmospheric model Integrated Forecasting System (IFS) cy36r4 (https://confluence.ecmwf.int/display/FCST/Implementation+of+IFS+Cycle+36r4) with the associated land surface module called Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL) (**?**), the Nucleus for European Modelling of the Ocean (NEMO) ocean model in its version 3.6 (**?**) and the OASIS3-MCT module (**?**) that couples the main components. IFS is an operational global meteorological forecasting model developed and maintained by ECMWF. The dynamical core of IFS is hydrostatic, two-time-level, semi-implicit, semi-Lagrangian and applies spectral transformations between grid-point space and spectral space. Vertically, the model is discretized using a finite-element scheme. NEMO is a state-of-the-art modelling framework for the ocean, based on the Navier-Stokes equations, used for oceanographic research, operational oceanography, seasonal forecasting and climate research studies. The NEMO version used in EC-Earth3.3 is known as "v3.6 stable". NEMO is composed by Océan PArallélisé (OPA), for the ocean dynamics and thermodynamics, a primitive equation model adapted to regional and global ocean circulation problems down to kilometric scale, and by Louvain-la-Neuve sea Ice Model (LIM) in its version 3 for the sea-ice dynamics and thermodynamics (**??**). EC-Earth3-SR uses T255L91, a spectral Gaussian grid, in the atmosphere, and ORCA1L75, a tripolar ocean grid, in the ocean. EC-Earth3-HR uses T511L91 in the atmosphere and ORCA025L75 in the ocean. ORCA1L75 has a latitudinal refinement of the grid in the tropics to $\sim$40 km. In comparison, the latitudinal spacing of ORCA025L75 is about 27 km in the tropics. Both ocean configurations have 75 vertical levels with thickness increasing from 1 m below surface up to 500 m in the deep ocean and uses a partial step z-coordinate.

Previous studies showed the importance of using a correctly tuned climate model for prediction (**?**). For the EC-Earth3-SR, we use the Climate Model Intercomparison Project 6 (CMIP6; **?**) version which has been intensively tuned by the EC-Earth consortium prior to the production of simulations for CMIP6 (**?**). EC-Earth3-HR (**?**) is the equivalent configuration to the SR one in terms of component versions, but with additional parameters tuned to improve more specifically process representation and mean biases in the equatorial Pacific and the North Atlantic, two key regions for near-term climate prediction skill.

One particularly interesting feature of the HR configuration is the improvement in the simulated variability of the deep convection in the Labrador Sea and the Atlantic Meridional Overturning Circulation (AMOC) compared to EC-Earth3-SR, where deep convection occurred too intermittently due to an overly strong local density stratification. The biases in the Labrador Sea caused a problem in the decadal prediction system based on EC-Earth3-SR, for which an initialization shock occurred leading to a collapse of the Labrador Sea convection (**?**), inducing a quick degradation of the predictive skill in the Subpolar North Atlantic, a source region of decadal variability and predictability (**?**). This problem is not present in EC-Earth3-HR, for which the Labrador Sea convection remains active and stable, with a sustained impact on the strength of the AMOC.

2.2 Seasonal prediction system protocol

The ability of climate models to make accurate predictions is evaluated by performing retrospective predictions, also referred to as hindcasts. These are ensembles of predictions with forecast horizons of several months that start from initial states in the past (or start dates) evenly distributed over the hindcast period. We performed two sets of seasonal hindcasts with each resolution configuration of EC-Earth3. The hindcast period covers 1990-2015, i.e. 26 start-dates. In this study, we focus on the hindcast initialized on the 1st of May but we note that a November-initialized hindcast system has also been produced for the two resolution configurations. Each hindcast system has 20 ensemble members and has a forecast length of 8 months, which for the case of the May initialized forecast, allows us to cover boreal summer through early boreal winter.

The hindcast systems have been initialized using full-field initialisation. The initial conditions for the atmosphere and land components are taken from the ERA5 reanalysis dataset (**?**) and subsequently interpolated to the same grid as the atmospheric component for each configuration, EC-Earth3-SR and EC-Earth3-HR. The ensemble of the atmospheric initial conditions is generated by introducing random infinitesimal perturbations (to the order of 10^5 K) in the 3D air temperature field, that aim to represent our best estimate of the observational uncertainty. This process is repeated for all start dates. The ocean and sea ice initial conditions come from ocean-sea ice forced simulations, called hereafter ocean reconstructions, with NEMO3.6-LIM3 in stand-alone mode. The NEMO3.6-LIM3 stand-alone configurations are tuned in the same way as the corresponding coupled configuration except for the snow conductivity of the HR configuration for which we are using the default value (rn_cdsn = 0.27 in the HR stand-alone configuration, rn_cdsn=0.15 in the HR coupled configuration). We introduced this change in snow conductivity to have comparable sea ice volumes in both configurations, as the ocean-sea ice stand-alone version tends to produce thicker ice than the coupled one.

The ocean reconstructions are driven by surface atmospheric fluxes of ERA5 HRES and assimilate sea surface temperature (SST) and salinity (SSS) from the ECMWF Ocean Reanalysis System 5 (ORAS5; **?**) and 3D ocean temperature and salinity below the mixed layer from the EN4 v4.2.1 ocean objective analysis ~~(?), with g10 bias corrections (?)~~(**?**, we used the version with the g10 bias . The atmospheric variables used to force the ocean are air temperature at 10 m, latent and sensible heat fluxes, specific humidity at 10 m, precipitation (including rain and snow) and surface winds at 10 m. SST and SSS from ORAS5 are assimilated using relatively strong temperature and salinity restoring coefficients of -200 $\mathrm{W/m^2/K}$ and -750 $\mathrm{kg/m^2/s/psu}$, respectively. The reason to assimilate EN4 v4.2.1 in the subsurface is to avoid the non-stationary bias reported in the North Atlantic for the ECMWF seasonal forecast system SEAS5 (**?**) that arises from problems in the ORAS5 subsurface fields (**?**). A Newtonian relaxation term is applied to assimilate the 3D temperature and salinity of EN4 v4.2.1, with the weak relaxation timescale going from 30 days at the surface to 3650 days at the bottom, increasing monotonically and with a reduction around the coastlines. There is a 10 times weaker subsurface damping between 15°S-15°N to avoid spurious vertical velocity effects (**?**). The methodological choices on the products to assimilate and on the strength of the restoring coefficients to apply have been tested in small ensembles of retrospective seasonal forecasts at standard resolution, and those given an overall better performance in terms of forecast biases and skill were finally used.

We produced five different members of these SR and HR historical ocean reconstructions, covering the period 1959-2021. The ensemble members have been generated by applying infinitesimal perturbations to some of the atmospheric forcings from ERA5 (10 m air temperature and sensible and latent heat fluxes) and by using five different ocean initial conditions, corresponding to the restarts on December 31st of the last five years of an ocean-only spin-up. The ocean-only spin-up has been run with constant 1959 ERA5 forcing and assimilation of 1959 ocean observations (again from ORAS5 at the surface and EN4-v4.2.1 at the subsurface), and covers a total period of 20 years to allow for the model to equilibrate.

The simulations were performed using a workflow manager developed at the Barcelona Supercomputing Center (BSC), called Autosubmit (**?**), and its associated graphical interface (**?**). Autosubmit ensures an optimal use of the computing resources by handling dependencies between jobs in an automatic way and packing multiple tasks in the same job execution. It can then easily handle simulations with complex workflows with different members and start dates. The simulation's configurations have been optimized in terms of computing resources, so that simulations run using the least amount of computing resources possible, while maintaining optimum execution times (**?**).

2.3 Observational data and reanalysis

As references to assess the forecast systems, we use different observational and reanalysis products covering the hindcast period 1990-2015. SST observations are taken from the ESA SST CCI v2.1 level 4 analysis (**?**) to benefit from their high horizontal resolution (0.05°latitude-longitude resolution). To address the impact of observational uncertainty in the results, we also compare the forecast systems against the ECMWF ERA5 reanalysis (**?**) and the ECMWF ORAS5 reanalysis (**?**). We also initially used the Hadley Centre Sea Ice and SST dataset (HadISST) in its version 1.1 (**?**) as verification product, but it should be noted that while using ERA5 and ORAS5 reanalyses give similar results than ESA SST CCI, the variability of the tropical Pacific in the HadISST v1.1 product is different and changes the values of skill metrics, especially the anomaly correlation coefficient (not shown). This feature can be linked to zonal and meridional discontinuities issues documented in the technical report (**?**). The HadISST v1.1 dataset has therefore not been included in the presentation of the results of this study. For the oceanic mixed layer depth, we use the ORAS5 reanalysis. Similar results are obtained when using the EN4 v4.2.2 dataset (**?**) as verification product (not shown). For all atmospheric variables (winds, precipitation and velocity potential), the ERA5 reanalysis is used for verification. Similar results for winds and precipitation are obtained when using the JRA-55 reanalysis (**?**) as verification product (not shown).

2.4 Analysis methodology

All the analysis based on spatial fields has been done by interpolating first the outputs of the models onto a regular grid at 1 degree resolution to compare with the observational and reanalysis reference products. The computation of the spatial SST over the different regions of interest (ENSO, Atlantic-Niño and Western Equatorial Pacific regions) has been performed as the area-weighted averaged SST in the original grids of the forecast systems and the verification products. Data post-processing, statistical analyses and the significance assessments have been performed using R4.1.2 and the startR (**??**, v2.3.1), s2dv (**??**, v2.0.0) and ClimProjDiags (**?**, v0.3.3) R-packages.

2.4.1 Drift correction

The initialized predictions cannot be compared with observed data at face value due to a phenomenon known as forecast drift, which manifests as the models transition from an observation-constrained initial state to their own attractor. It is generally assumed that the drift manifests in a consistent manner independent of the initial climate state. It is hence a standard approach to compare only the deviations of the simulations and observations from their lead-time dependent mean state, or climatology. In this study, the forecast anomalies, or drift-corrected anomalies, are computed by subtracting the lead-time dependent monthly climatology over the entire hindcast period 1990-2015.

The mean bias by lead months gives an indication of the impact that the forecast drift can have (Supplementary Figure 1). Both systems exhibit certain non-stationary characteristics in the forecast anomalies for all ENSO indices, which appear to be more pronounced for the longest forecast lead-times during the first half of the hindcast period. This non-stationary behaviour introduces some spurious signals in the drift-corrected anomalies that are expected to degrade their forecast skill. However, since the non-stationary effects are similar in both systems, we can assume that they do not explain the differences in skill.

2.4.2 Anomaly correlation coefficient

Prediction skill is assessed with the Anomaly Correlation Coefficient (ACC) against the reference datasets, computed using a pointwise Pearson correlation between the ensemble mean forecast anomaly and the corresponding observed anomaly over the entire re-forecast period. The significance of the correlations is determined by a two-sided Student-t test, at the 95 % confidence level. The statistical significance of the correlation differences is computed with a two-sided test to assess whether the skill of the two forecast systems are significantly different (**?**) at the 95 % confidence level. The time series autocorrelation has been taken into account when estimating the statistical significance following **?**.

3 Results

3.1 Skill assessment

The ACC maps of SST anomalies in the tropics (Fig. 1, left and middle column) show significant skill in most of the tropical Pacific and Atlantic ocean, in both EC-Earth3 forecast systems, with the exception of the most western equatorial Pacific region and, to a lesser extent, some parts of the southern tropical Atlantic Ocean. EC-Earth3 in its HR configuration shows higher predictive skill (statistically significant) than its SR configuration in the central and western equatorial Pacific for all lead times considered.

The higher predictive skill in EC-Earth3-HR than in EC-Earth-SR over the central tropical Pacific is confirmed by examining the ACC as a function of forecast time for the spatially averaged SST in three different ENSO regions (Fig. 2). In both the Niño3.4 (5°S-5°N, 170°W-120°W) and Niño3 (5°S-5°N, 150°W-90°W) regions, skill is higher in the HR configuration than in the SR configuration, an increase that is statistically significant for forecast months 3-5, i.e. the boreal summer (July-August-September). The increase in ENSO predictive skill in HR with respect to SR is, however, lower and not significant in its

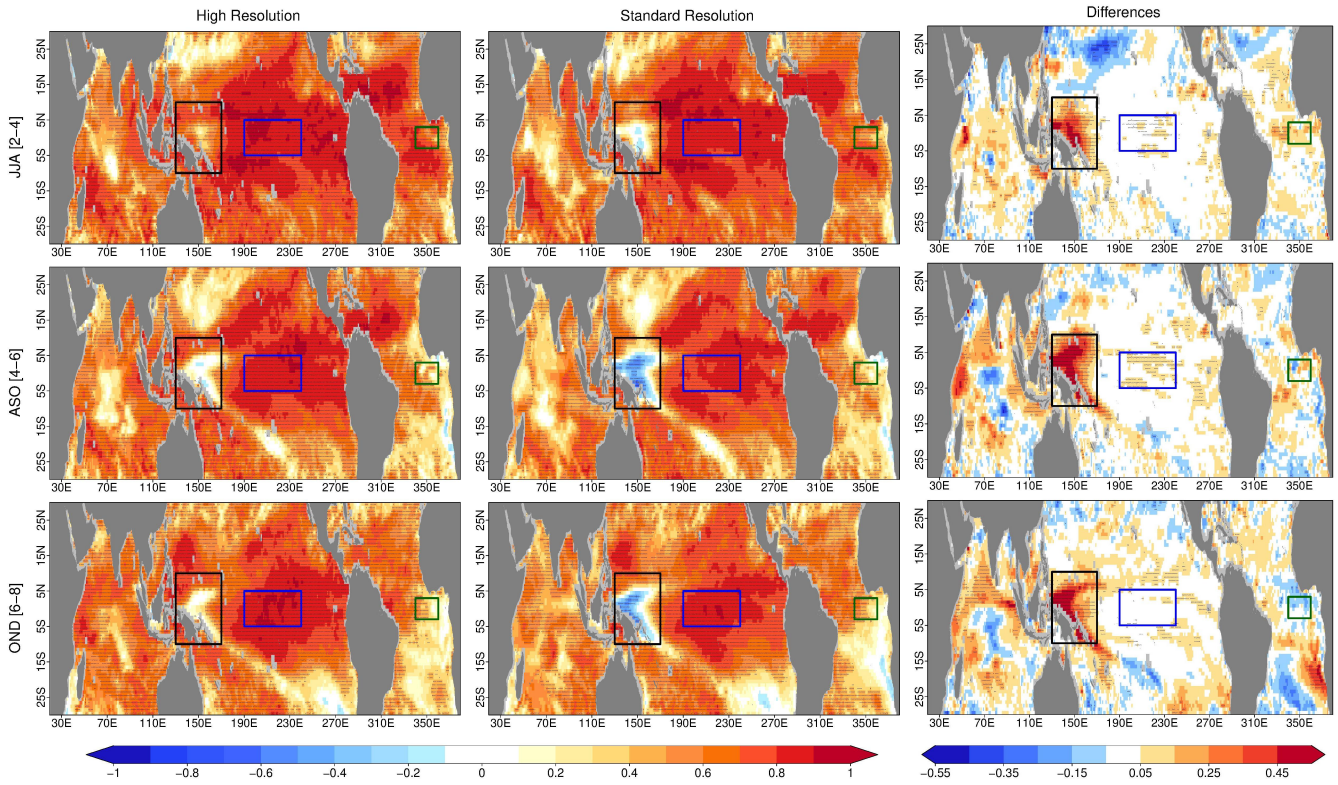

Figure 1. Maps of ACC of sea surface temperatures (SST) for the forecasted season (top row) June-July-August (JJA, lead-months 2-4), (middle row) August-September-October (ASO, lead-months 4-6) and (bottom row) October-November-December (OND, lead-months 6-8) initialized in May. The left column corresponds to the HR configuration, the middle one to the SR configuration and the right column shows the difference in correlation between HR and SR. In the right column, red indicates an improvement in the skill with increased resolution, while blue indicates a degradation of the skill. The reference dataset is the ESA SST CCI dataset. The ACC has been computed over the 1990-2015 period for each individual grid point after interpolation to a regular 1º grid and using the ensemble mean of the 20 members of each system. Dashes indicate that the values are statistically significant at the 95 % level. The blue box is the Niño 3.4 region (5°S-5°N, 170°W-120°W), the black box is the Western Equatorial Pacific (WEP) region (10°S-10°N, 130°E-170°E) and the green box is the Atlantic Niño region (3°S-3°N, 20°W-0).

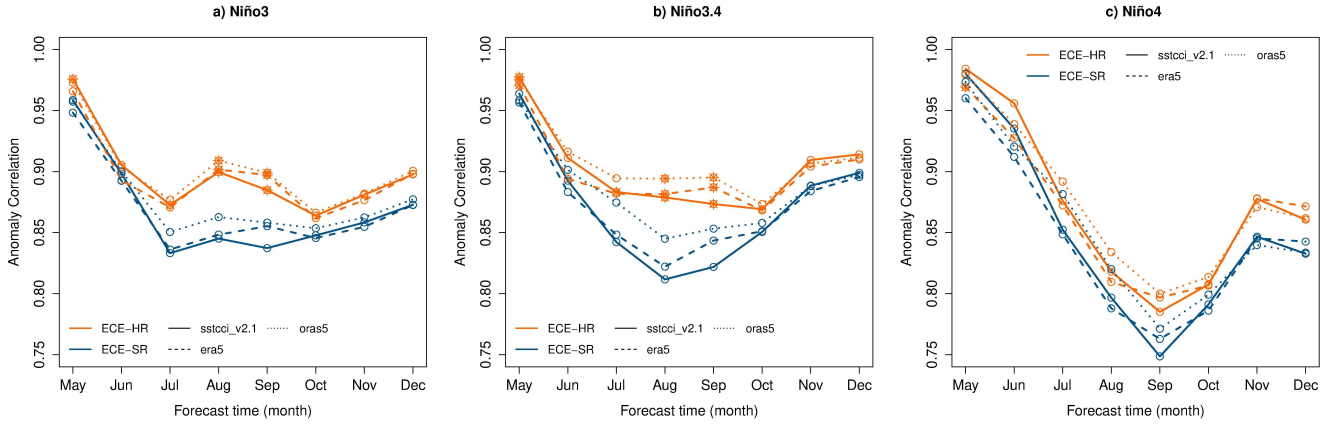

Figure 2. ACC in the (a) Niño3 (5°S-5°N, 150°W-90°W), (b) Niño3.4 (5°S-5°N, 170°W-120°W) and (c) Niño4 (5°S-5°N, 160°E-150°W) regions as a function of forecasted month for (dark blue) EC-Earth3-SR and (~~light blue~~orange) EC-Earth3-HR. The reference datasets are ESA SST CCI dataset (plain line), ERA5 dataset (dashed line) and ORAS5 dataset (dotted line). The hindcast period is 1990-2015 and the 20 members of each system are used. Open circles mean that the ACC is statistically significant at 95 % level of confidence and stars mean that the differences of correlations between the SR and HR configurations of EC-Earth are statistically significant.

westernmost part, as shown in Figure 2c for the Niño4 region (5°S-5°N, 160°E-150°W). Further west in the equatorial Pacific, next to the Maritime Continent, EC-Earth3-HR shows statistically significant higher predictive skill than EC-Earth3-SR, which has no predictive skill in this region (Fig. 1). However, the skill, even if substantially improved, is still not statistically significant for all lead times in EC-Earth-HR. This Western Equatorial Pacific (WEP) region is further analysed in Section 3.2.

The higher predictive skill in the equatorial Pacific in EC-Earth3-HR can arise for various reasons. One of these could be a better representation of the local physical processes that drive the tropical Pacific variability and in particular ENSO. This improvement could be linked to the higher horizontal resolution of the model, via for instance a ~~better~~ more realistic simulation of intrinsic variability (**?**) or a better orographic effect on the land-atmosphere-ocean interactions over the Maritime Continent (**?**). Another possible reason could be a ~~better~~ more realistic simulation of key teleconnections that affect the region, like the influence of Atlantic El Niño/La Niña events on the Tropical Pacific, as dynamical prediction systems that represent it more accurately have been shown to have higher ENSO predictive skill (**?**).

In the next sections, we focus on processes that can explain why the predictive skill in the ENSO region is higher in the HR configuration than in the SR configuration. We first focus on the western equatorial Pacific to analyse why this region is poorly predicted in EC-Earth3-SR, the potential reasons for the enhanced skill in EC-Earth-HR, and whether the model's predictive skill in this region is linked to ENSO. Secondly, we analyse the potential link between the equatorial Atlantic variability and ENSO skill.

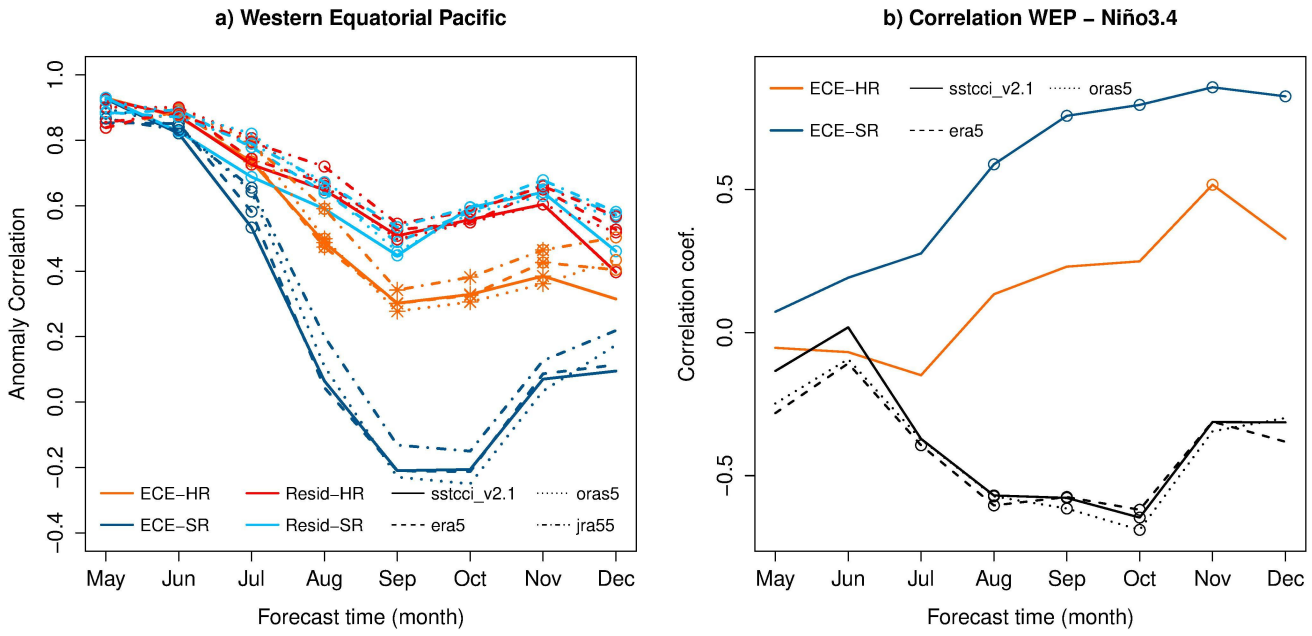

Figure 3. (a) ACC ~~(blue lines)~~ in the western equatorial Pacific (WEP - 10°S-10°N, 130°E-170°E) SST anomalies as a function of forecasted month for EC-Earth3-SR (dark blue) and EC-Earth3-HR (~~light blue~~orange). The reference datasets are ESA SST CCI dataset (plain line), ERA5 dataset (dashed line) and ORAS5 dataset (dotted line). The correlation coefficients, compared to these 3 reference datasets, of the residuals of the regression of the WEP SST anomalies onto Niño3.4 SST anomalies are shown for EC-Earth3-SR (~~red~~light blue) and EC-Earth3-HR (~~orange~~red). Open circles mean that the ACC is statistically significant at 95 % level of confidence and stars mean that the differences of correlations between the SR and HR configurations of EC-Earth are statistically significant at 95 % level of confidence. (b) Correlation coefficients between SST anomalies in the western equatorial Pacific and Niño3.4 region (5°S-5°N, 170°W-120°W). The black lines are for the same observation products as used in Fig.3a as references. Open circles mean that the correlation is statistically significant at 95% level of confidence.

3.2 Exploring the prediction skill in the western equatorial Pacific

As highlighted in Figure 1, the Western Equatorial Pacific (WEP; herein defined between 10°S-10°N and 130°E-170°E) shows a quick drop in predictive skill, particularly in EC-Earth3-SR. Figure 3a ~~(blue lines)~~ shows the ACC time series of the SST in this region for both forecast systems. In both configurations, the skill in predicting SST quickly drops after the second month of forecast, reaching very low skill levels between September and October. However, the skill is higher in the EC-Earth3-HR configuration and remains statistically significant for another month (until August) compared to EC-Earth3-SR. The differences in the skill between the two configurations are also statistically significant from August to November.

When estimating the link between the WEP and Niño3.4 regions, represented by the correlation coefficients between the associated SST time-series (Fig. 3b), the observations consistently show negative values for all lead times, thus indicating that

both regions vary in opposite ~~phase~~phases: when the ENSO region gets cold, the WEP region tends to get warm. However, in both initialized predictions, the WEP-Niño3.4 relationship is the opposite of that observed, displaying predominantly positive correlation coefficients. This effect is especially pronounced in EC-Earth3-SR where correlation coefficients are positive since the first month after initialisation and remain substantially higher than for EC-Earth3-HR. The poor predictive skill in the WEP region and the erroneous positive WEP-Niño3.4 relationship is also present in other forecast systems, as for instance the operational North American Multi-Model Ensemble (NMME) (**?**, Fig. S3c).

We now explore whether this erroneous co-variability between ENSO and the WEP region affects the predictive skill of the WEP region. To do this, we compute the prediction skill over that region after regressing out the influence of ENSO. In practice, this is done by computing the linear regression of the WEP SST time-series onto that of Niño3.4, both for the model and for the reference datasets, retaining the residuals in each case and calculating their correlation with each other (Fig 3a~~, red and orange lines~~). When removing the ENSO signal from the WEP region, the correlation coefficients are higher and similar in both EC-Earth3 configurations, which indicates that the misrepresented link with ENSO is behind the poor prediction skill in the WEP region.

3.2.1 ENSO-related errors

This erroneous feature is at least partly due to the common ENSO-related SST westward extension bias, which is present in many climate models (**???**). Figure 4 shows the ENSO-related errors spatial patterns, which are diagnosed following **?** as the regression at each grid point of the forecast errors (forecast minus observation) onto the observed standardized Niño3.4 index. The ENSO-related error patterns are thus an indicator of how close the spatial representation of ENSO in the forecast is to what it should be based on observations. As a reference, we also include in contours the regression pattern of the observed SST anomalies onto their respective normalized Niño3.4 index, to compare the simulated ENSO-related errors to the observed ENSO-related variability. At the end of the forecast period (October-November-December, OND), EC-Earth3 predictions in both configurations exhibit a cold ENSO-related error in the central-to-eastern equatorial Pacific, which extends farther west in the EC-Earth3-SR (Fig. 4). This error induces ENSO events of lower amplitude in this central region, a feature that is more pronounced in EC-Earth3-SR. This tendency to underestimate the magnitude of ENSO is accompanied by a warm ENSO-related error on the westernmost part of the basin, which is due to the aforementioned common bias in climate models of extending ENSO-related SST anomalies westward. In EC-Earth3-SR in particular, the spatial pattern associated with ENSO variability is of the same sign across the entire equatorial Pacific while in observations the change of signs in SST anomalies occurs around 160°E (Fig. 4 contours for observation). This same variability phase throughout the entire equatorial Pacific in the predictions lowers the predictive skill in this WEP region and has implications on the ENSO-Western Pacific teleconnections as shown in Figure 3.

Given the strong air-sea interactions in the tropics, these ENSO-related errors are also found in other variables, such as zonal surface winds and mixed layer depth, and develop further with lead time. Figure 5 shows the time evolution of the ENSO-related errors averaged between 5°S and 5°N from the first month after initialisation (May) to the last forecast month (December) for the SST, the zonal surface winds and the mixed layer depth. In both prediction systems, the ENSO-related SST

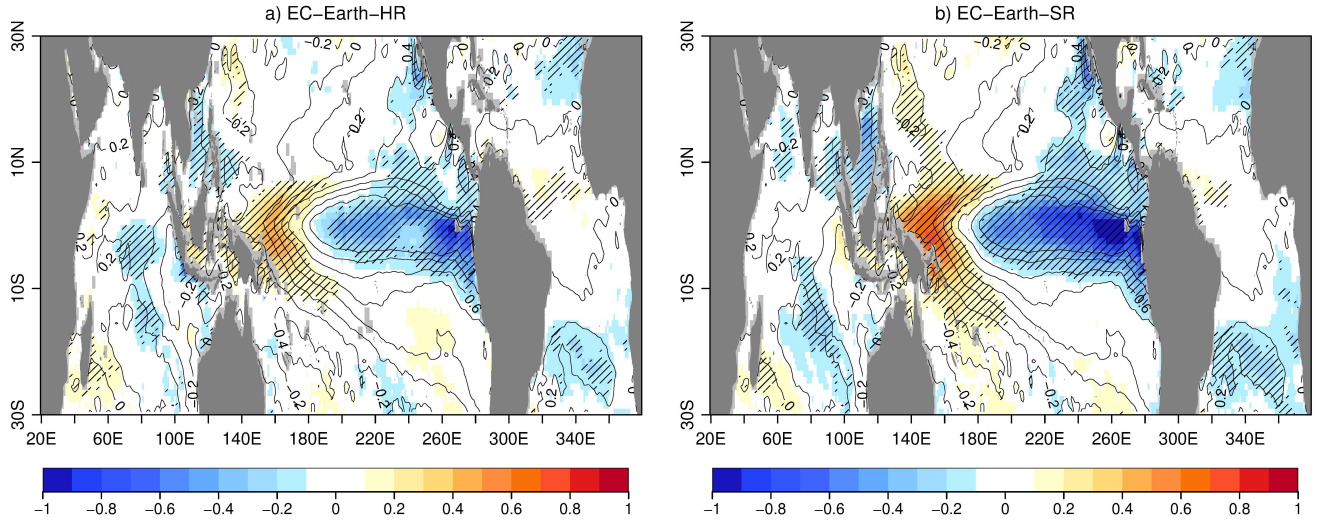

Figure 4. Maps of (colours) ENSO-related SST errors as defined in **?** and (contours, every 0.2) regression coefficients of Niño 3.4 SST anomalies time-series onto SST anomalies fields in October-November-December (OND) for (a) EC-Earth-HR and (b) EC-Earth-SR. The hindcast period used is 1990-2015. The reference product is the ESA CCI v2.4 dataset. Dashes indicate that the values are statistically significant at the 95 % level.

errors start in the first month after initialisation, with a small warm bias forming around 160°E (the first sign of the erroneous westward extension of ENSO) which then grows with time. While the ENSO-regressed SST anomalies in observations start to be positive from around 160°E to cover the entire central and eastern tropical Pacific, the longitude where this happens in both EC-Earth3 configurations is around 140°E. This leads to warmer ENSO-regressed SST in the models than in observations in the western Pacific, which results in the positive ENSO-related errors seen in Figure 5a,b, with a maximum error in September-October. In the eastern part of the tropical Pacific, the negative (cold) ENSO-related errors start in the easternmost part of the equatorial Pacific and then spread to the dateline, in particular because positive SST anomalies are not high enough while El Niño events are developing. The westward overextension of ENSO-related SST anomalies is associated with ENSO-related errors in the winds, particularly during ENSO's development phase which starts in August. During the positive phase of ENSO, i.e. El Niño events, the westerly winds are too weak and too constrained around 150°E-160°E while the westerly winds spread between 160°E-200°E and at higher speed in observations (Fig. 5, arrows). Since the central equatorial Pacific is a key region for wind-mixed layer interaction processes, the ENSO-related errors are also evident in the mixed layer dynamics. During El Niño events, the mixed layer depth (MLD) anomalies change sign, between 200°E-220°E in observations, from positive anomalies (i.e. deepening) in the east to negative anomalies (i.e. shallowing) in the west. This transitional region of the MLD is shifted westward in the model following the errors in the wind, which induces positive ENSO-related MLD errors (Fig. 5c,d), particularly between 180°E-200°E.

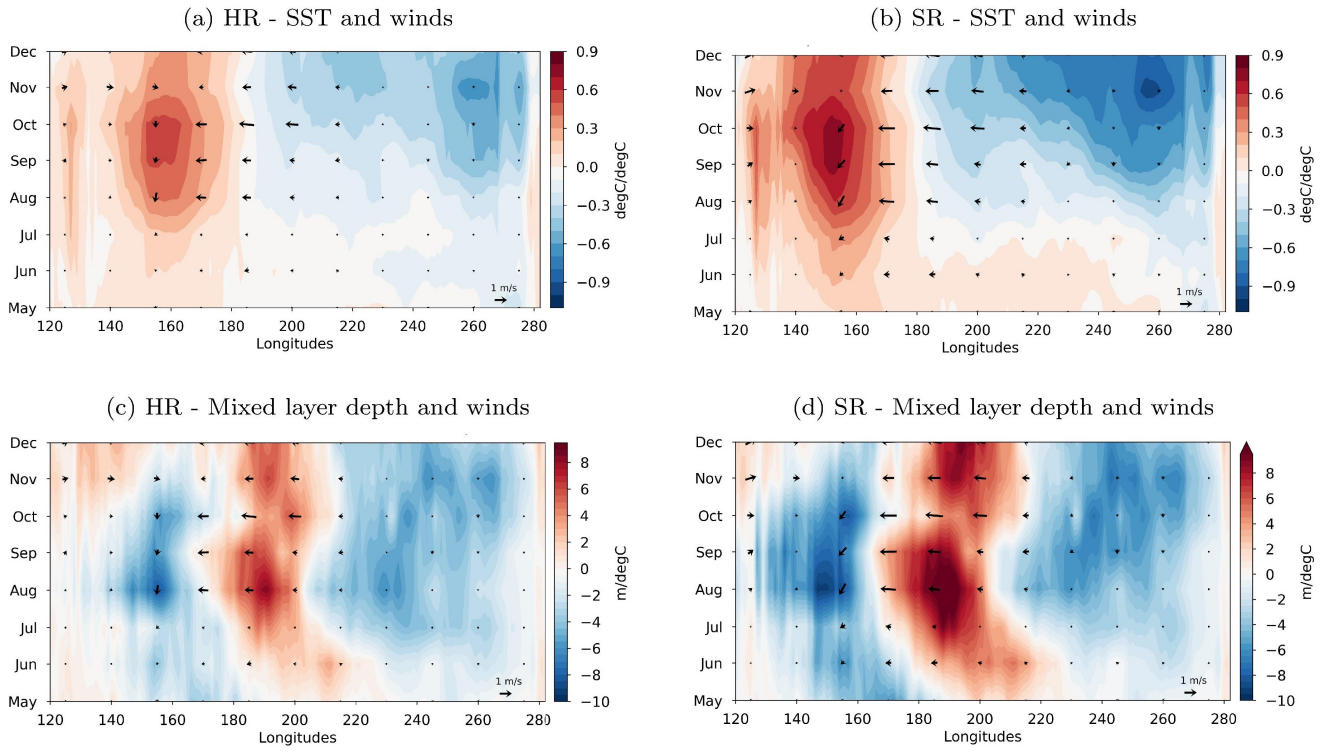

Figure 5. Hovmoller from May to December of the ENSO-related errors for (a,b) SST (shaded) and surface winds (arrays) and (c,d) mixed layer depth (shaded) and surface winds (arrays) in EC-Earth3-HR (a,c) and EC-Earth3-SR (b,d) forecast systems. The mixed layer depth is defined as the depth at which there is a change of the density of 0.03 $\mathrm{kg/m^3}$ with respect to the density at 10 m. The reference datasets are ESA SST CCI for the SST, ERA5 for the surface winds and ORAS5 for the mixed layer depth.

3.2.2 Linking to mean state biases

The erroneous westward extension of the ENSO-related variability in the model can potentially arise from mean state biases, which appear already in the first month after initialisation (Fig. 6). In both EC-Earth3 forecast systems, the SST show the common cold bias in the central-to-eastern equatorial Pacific, particularly in the SR configuration, which worsens with forecast time as it extends westwards (Fig. 6a,b). The mixed layer depth also shows a mean state bias since the first forecast month, characterized by too deep mixing between 160°E-180°E, which corresponds to a westward shift in the normal deepening of the mixed layer depth (Fig. 6e,f), a shift that intensifies with time. By contrast, east of 180°E-200°E, the mixed layer is too shallow, which matches the SST cold bias in this region due to too intense equatorial upwelling. Associated with this westward shift in the deepening of the mean mixed layer, the mean surface zonal winds show a westward displacement of the easterly winds, peaking between 180°E-200°E in the model, compared to around 220°E in the observations. The intensity of the easterlies is stronger in the models and the temporality is also shifted, with the peak intensity taking place in July in the models rather than

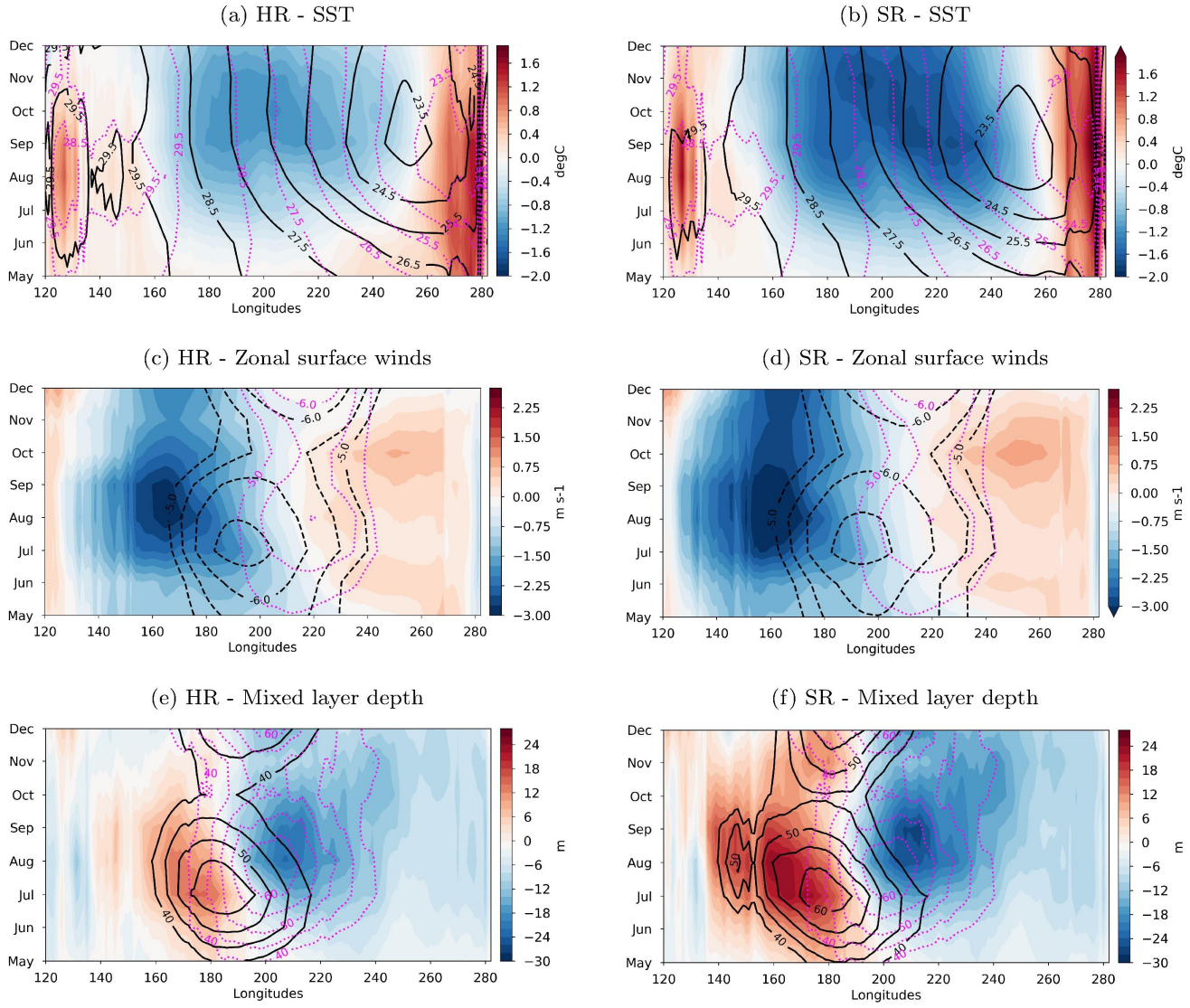

Figure 6. Hovmoller from May to December of the mean state biases for (a,b) SST, (c,d) zonal surface winds and (e,f) mixed layer depth in EC-Earth3-HR (a,c,e) and EC-Earth3-SR (b,d,f) forecast systems. The black lines represent the mean state of the forecast systems, the purple lines are the mean states of the corresponding reference dataset. The reference datasets are ESA SST CCI for the SST, ERA5 for the surface winds and ORAS5 for the mixed layer depth.

in August as in the observations, a shift in the peak that is also found in the mixed layer depth. Even if they are similar in nature, ~~ENSO-related errors~~ the mean state biases and the erroneous westward extension develop more rapidly and are more pronounced over time in EC-Earth3-SR than in EC-Earth3-HR (Fig. ~~5~~6), which also happens for the ~~mean state biases and the erroneous westward extension~~ ENSO-related errors (Fig. ~~6~~5).

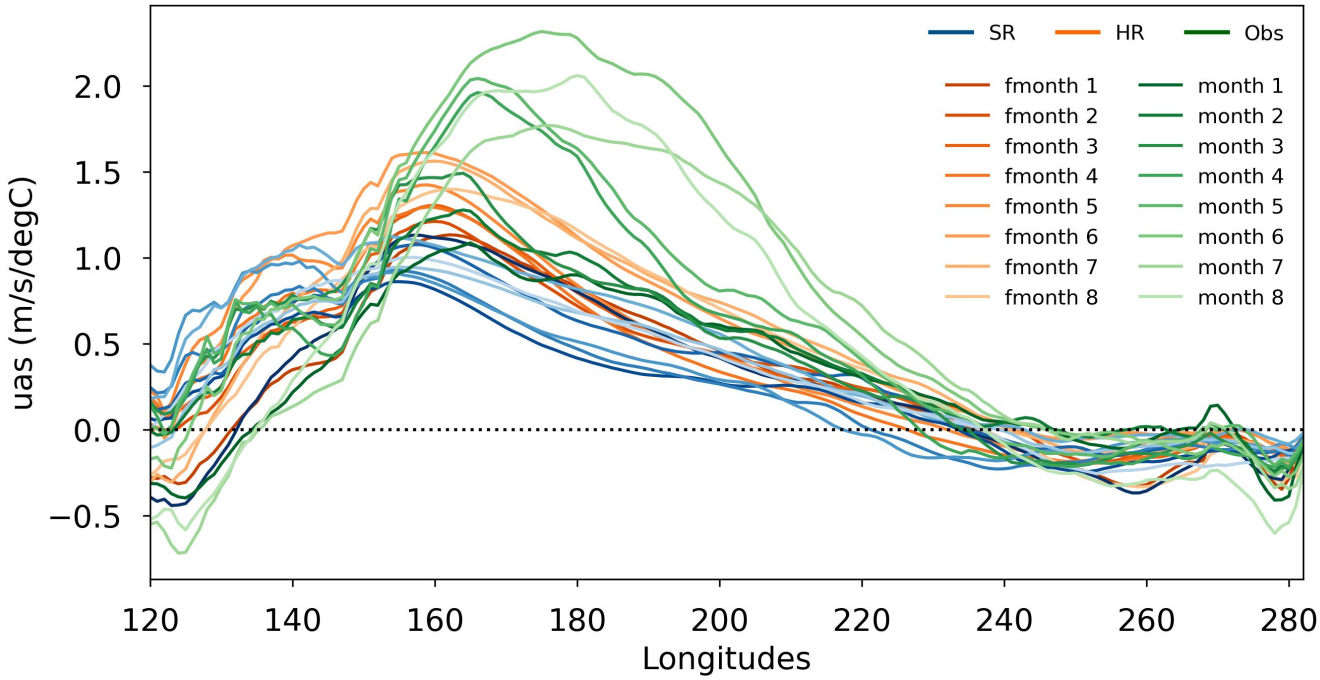

Figure 7. Longitudinal profiles (5°S-5°N) of the time-evolution of the ENSO-regressed surface zonal winds for ERA5 and the EC-Earth3-SR (~~plain~~ blue lines) and EC-Earth3-HR (~~dotted~~ orange lines) forecast systems. The different shades of colours represent each month along the forecast time, from the first forecast month, May, in ~~red~~ dark green for the observations, in dark blue for ~~the forecast systems~~ EC-Earth3-SR and in dark orange for EC-Earth3-HR to the eighth forecast month, December, in ~~yellow for~~ the ~~observations~~respective same colours, ~~in green for the forecast systems~~but lighter.

3.2.3 Winds variability deficiency

The ENSO-related winds differ between EC-Earth3-SR and EC-Earth3-HR ~~particularly during the development of~~ as El Niño events evolve with forecast time (Fig. 7). In EC-Earth3-SR, the westerly winds in the western Pacific are too weak and ~~do not grow~~ are not strengthening as El Niño events ~~develop~~grow, reaching only around 1 $\mathrm{m/s/degC}$ in December, values which are comparable to observations at the beginning of the events development phase (May). The westerly winds in EC-Earth3-HR reach higher peak values ($\sim$ 1.5 $\mathrm{m/s/degC}$) in December even if they remain lower than in the observations ($\sim$ 2 $\mathrm{m/s/degC}$). The EC-Earth3-SR model deficiency to simulate strong ENSO events, linked to too low winds variability and a weak air-sea coupling in the Tropical Pacific, is documented in **?** and is therefore also present in the initialized simulations. It is less pronounced in EC-Earth3-HR, which may be related to the improved mean state, which ~~reduces~~ would reduce the impact of the coupled feedback bias due to the weak coupling.

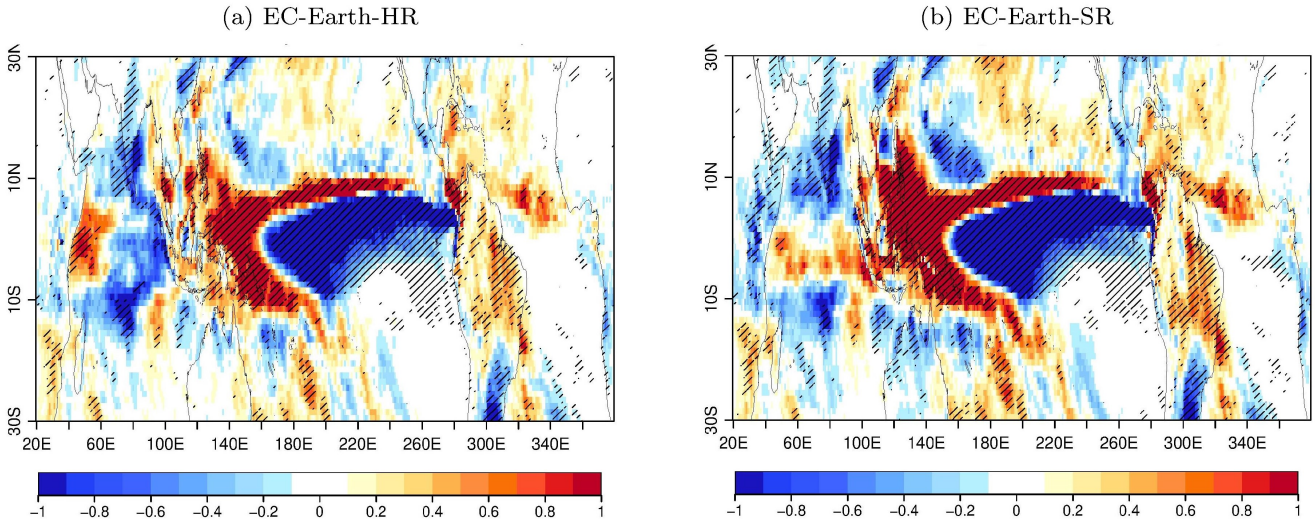

Figure 8. Maps of (colours) ENSO-related precipitation errors as defined in **?** in October-November-December (OND) for (a) EC-Earth-HR and (b) EC-Earth-SR. The hindcast period used is 1990-2015. The reference product is the ERA5 reanalysis.

3.2.4 Summary of results on prediction skill in the WEP region

To summarize, ENSO variability is shifted westwards and weakened in the forecasts, these two errors developing rapidly from the first forecasted month, persisting over time and being more pronounced in EC-Earth3-SR than in EC-Earth3-HR. These ENSO-related errors, which induce low SST predictive skill in the WEP regions, are linked to the erroneous spatial simulation of ENSO, which is itself likely linked to the mean state representation. EC-Earth3-SR shows particularly poor predictive skill in the WEP region compared to EC-Earth3-HR due to higher ENSO-related errors and possibly due to the mean state bias. Overall, a weak air-sea coupling, weaker in EC-Earth3-SR, ~~prevents~~ may prevent the model from simulating the correct ENSO development.

~~The westward shift in ENSO SST and wind zones of influence directly impacts precipitation, shifting the location of the convergence zone~~ Precipitation associated with ENSO events also shows a deficiency in model simulations, as the convergence zone does not move correctly toward the centre of the equatorial Pacific during El Niño events (Fig. 8). ENSO-related precipitation ~~moves~~ is located too far west reaching the first islands of the Maritime continent, particularly in EC-Earth3-SR. The penetration of precipitation over the maritime continent (red colour) does not allow to correctly simulate the precipitation deficits that these islands, and the Philippines in particular, are likely to experience during El Niño events. At the same time, too little precipitation is simulated along the equatorial Pacific (blue colour), an error that grows with time reinforcing the separation of the Inter-tropical Convergence Zone (ITCZ) into two branches, particularly in EC-Earth-SR.

Improving the spatial simulation of ENSO and the tropical Pacific mean state in the model is particularly important for improving predictions of rainfall and SST in the region and over the maritime continent, a region of convection particularly important for teleconnections to extra-tropics (**?**).

3.3 Influence of Atlantic teleconnections on ENSO prediction differences

One of the most prominent modes of variability in the tropical Atlantic is the Atlantic-Niño (**??**). Atlantic Niño events are usually defined by the SST anomalies in the equatorial region, west to the Gulf of Guinea (3°S-3°N, 20°W-0, see green box in Fig. 1), hereafter called the ATL3 region. The Atlantic-Niño mode of variability exhibits higher variability during the boreal summer (**?**). It has been shown that summer Atlantic Niño events favour the development of Pacific La Niña events the following winter (**??**) by modulating the Walker circulation. During an Atlantic Niño event in boreal summer, the warmer SST lead to anomalous heating, altering the Walker cell by increasing upward motions over the Atlantic ocean and anomalous subsidence over the central Pacific. The latter then enhances easterly wind anomalies over the central and western equatorial Pacific, which triggers an upwelling oceanic Kelvin wave that propagates eastward. This subsequently cools the surface, triggering the Pacific Bjerknes feedback by further enhancing equatorial easterly winds and promoting the development of a Pacific La Niña event 6 months later in boreal winter. This teleconnection has been shown to enhance the prediction skill of ENSO the following winter (**??**), particularly during the late twentieth century, 1980-2015, which covers the hindcast period of this study, an influence that is subject to a decadal modulation via the South American low-level jet (**?**).

We now explore whether the improvement in the predictive skill in ENSO in EC-Earth3-HR with respect to EC-Earth3-SR can be directly related to this existing teleconnection. The improvement can come either from (1) better skill in ATL3 which impacts the ENSO region without any change needed in the teleconnection itself, (2) a ~~better~~ more realistic teleconnection that transfers the predictability from ATL3 into the ENSO region even without changes in skill in ATL3, or (3) both. Figure 9a shows the correlation coefficients between the ATL3 index in summer (JJA), when the Atlantic-Niño event peaks, and the temporal evolution of Niño3.4 from summer to early winter, to estimate how strong their co-variability is and how it evolves in time. This correlation is negative in the observations, as warm (cold) summer Atlantic Niño (La Niña) events are linked with the later occurrence of cold (warm) winter ENSO events. The teleconnection is established between JJA ATL3 and ENSO at subsequent lead times. Both models underestimate the teleconnection strength between the Atlantic Niños and ENSO, but it is only statistically significant in EC-Earth3-HR, which also has a more realistic strength.

In the ATL3 region itself, the EC-Earth3-HR forecast system does not present significantly higher predictive skill than EC-Earth3-SR (Fig. 9b). In particular, both configurations present a strong drop in the predictive skill in this region from July to September, which is consistent with other forecast systems (**??**).

It is thus the ~~better~~ more realistic simulation of the teleconnection between the Atlantic Niño/a and the following ENSO event in EC-Earth3-HR that can partly explain the improvement in ENSO predictive skill in this model (**?**). To investigate other aspects of the ATL3 teleconnection, we compute the correlation between the JJA ATL3 SST anomalies time-series and different atmospheric variables: precipitation and zonal winds at 850 hPa and 200 hPa (Fig. 10). Compared to ERA5, EC-Earth3-HR shows a similar representation of convection over the tropical Atlantic and the Amazon basin, as illustrated by the significant

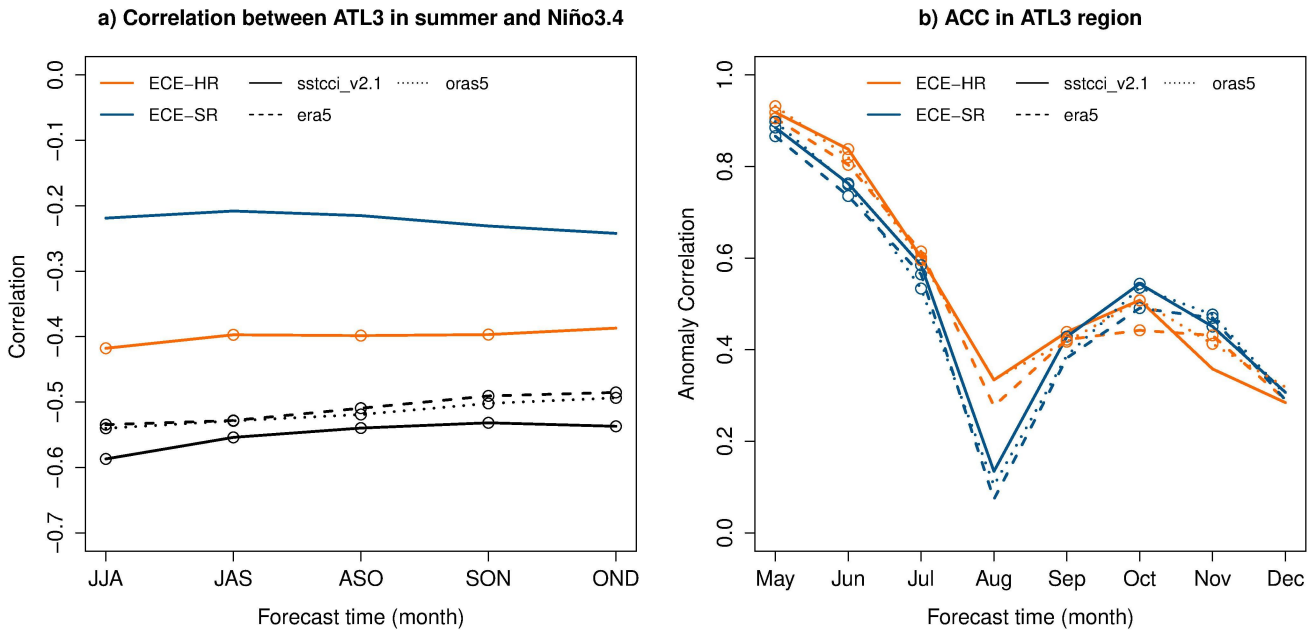

Figure 9. (a) Correlation coefficients between SST anomalies in the ATL3 region (20°W-0, 3°S3°N) in JJA and Niño3.4 region for moving 3-month means. The black lines are for reference datasets: (plain line) ESA SST CCI , (dashed line) ERA5 and (dotted line) ORAS5. (b) ACC in the ATL3 SST anomalies as a function of forecast month for (dark blue) EC-Earth3-SR, (~~light blue~~orange) EC-Earth3-HR compared to the same reference datasets as in a). Open circles mean that the ACC is statistically significant at 95 % level of confidence and plain circles mean that the differences of correlation between the SR and HR configurations of EC-Earth are statistically significant.

positive correlations in precipitation and zonal winds at 850 hPa and significant negative correlations in the zonal winds at 200 hPa which correspond to a divergence zone. This convection area is a prerequisite for the connection between summer Atlantic Niño/a and ENSO (**?**). EC-Earth3-SR also shows the same behaviour but the correlation values are clearly underestimated. The ~~better~~ more realistic representation of the teleconnection pattern in EC-Earth3-HR is also shown in the western-to-central tropical Pacific, with a more realistic shape and statistical significance of the subsidence zone over the Maritime continent seen in the precipitation and zonal winds.

This ~~better simulation~~ simulation closer to observations of the teleconnection between the Atlantic Niño/a and the following ENSO event in EC-Earth3-HR may be related to the higher horizontal resolution allowing orographic effect over the Maritime continent, contributing to the generation of a local Walker cell (**?**). This potential effect of the horizontal resolution on the convection over the Maritime continent linked to the Atlantic-Pacific teleconnection requires further investigation and is beyond the scope of the study.

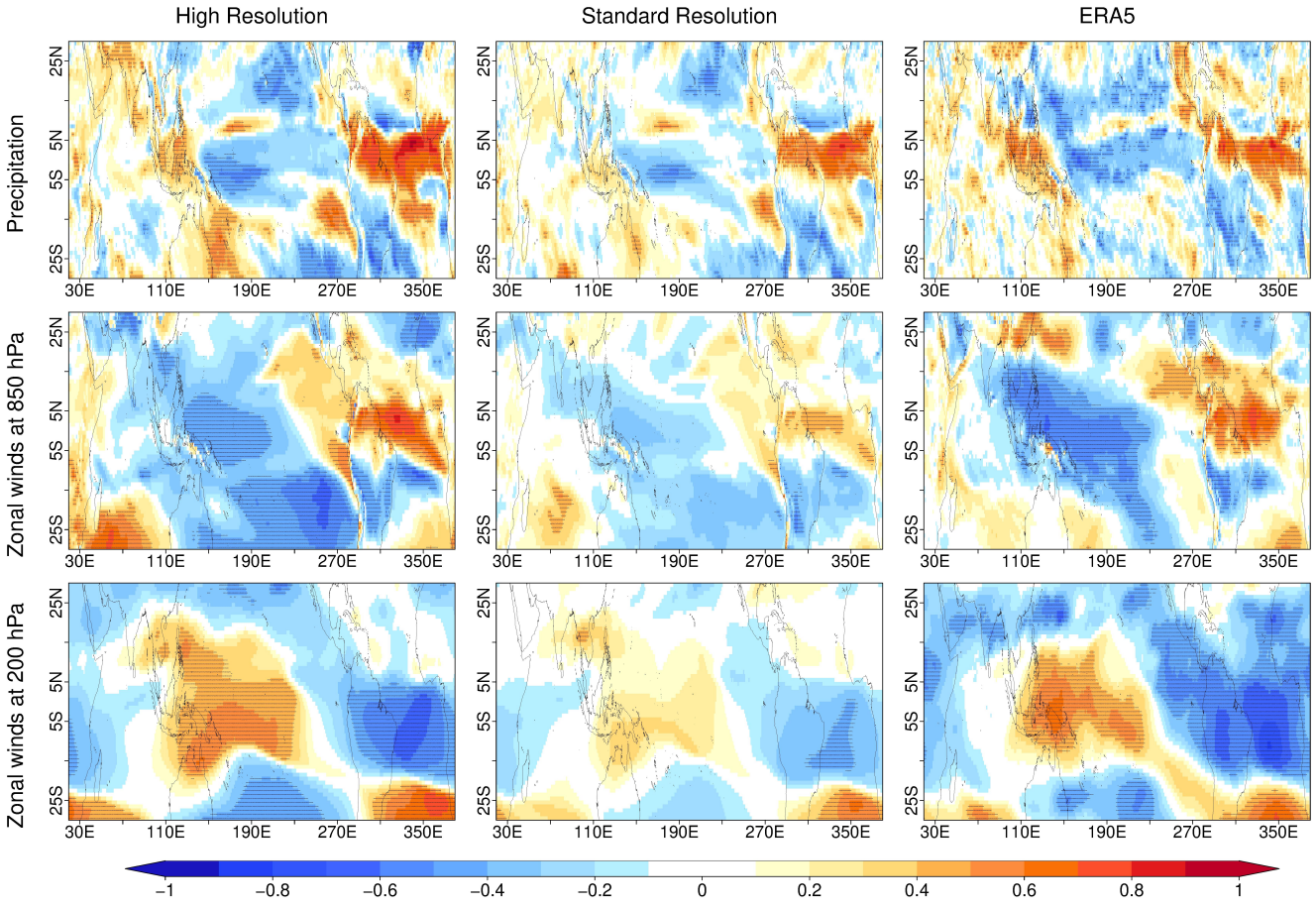

Figure 10. Teleconnection between the JJA ATL3 SSTA anomalies and the JJA precipitation (first row), zonal winds at 850 hPa (middle row) and zonal winds at 200 hPa (bottom row) for EC-Earth3-HR (left), EC-Earth-SR (middle) and ERA5 (right). Dashes indicate that the values are statistically significant at the 95 % level.

4 Summary and discussion

This study investigates and compares seasonal predictions with a high-resolution (HR) and a standard-resolution (SR) configuration of EC-Earth3. It focuses on the Tropics, and particularly on the predictive skill associated with ENSO, the main source of predictability at seasonal timescales. We analyze the differences in the predictive skill of the two different configurations, initialized in the same way but using two different horizontal resolutions. The SR and HR configurations are based on an atmospheric component, IFS, of $\sim$100 km and $\sim$40 km of resolution, respectively, and on an ocean component, NEMO3.6, of $\sim$100 km and $\sim$25 km respectively. We focus on the hindcast initialized in May of each year over the period 1990-2015 using 20 members.

The EC-Earth3-HR forecast system shows ~~higher~~ a small but statistically significant increase in predictive skill in the equatorial Pacific ~~than~~ relative to EC-Earth3-SR~~,~~. This modest improvement is consistent across regions (from the Western Equatorial Pacific ~~(WEP) to ENSO regions, not only in SSTbut also in~~ to the canonical ENSO regions) and across variables (SST, zonal surface winds, precipitation, air surface temperature and mixed layer depth). We investigate in this paper several processes that can explain this higher predictive skill up to 8 months ahead. The differences between the two forecast systems can arise, among other things, from (1) the representation of physical processes that are no longer parametrized but resolved at the higher resolution, (2) differences in their initial conditions (IC) or (3) the refined orography resulting from the resolution enhancement, that can subsequently interact with atmospheric circulation and induce moisture convergence (**?**). This last hypothesis cannot be investigated further, as it requires conducting sensitivity experiments in which the orography remains identical for both resolutions, which is beyond the scope of this paper. Regarding the differences in initialization, we use the same ERA5 IC for the atmosphere, but interpolated to the respective horizontal resolutions, which should have a minimum impact. The ocean IC are produced using the same ocean model but at different resolutions, nudged towards and forced by the same observation datasets but interpolated to the respective horizontal resolutions. The spread between the IC members is higher in the HR reconstructions than in the SR reconstruction and closer to the equivalent spread between members of ORAS5 (not shown), which can impact the prediction skill of the forecast systems. Finally, regarding the differences in the simulation of physical processes, certain processes that are no longer parametrized at the higher resolution can potentially improve the mean state and key atmospheric teleconnections, such as mesoscale ocean eddies and their interactions. In this perspective, we focus on two main aspects. The first one is the spatial representation of ENSO, which can be related to the mean state representation. The second one is related to atmospheric teleconnections with the Atlantic Niño/a region.

First, we show that in the WEP region, both systems lose their SST predictive skill rapidly but more drastically in EC-Earth3-SR. The poor skill in the WEP region is directly linked to a misrepresentation of ENSO influence, defined as the Niño3.4 region influence, as when its signal is removed from the WEP's SST variability, both forecast systems show significant skill levels in the region. This misrepresentation of ENSO's influence is linked to the overextended westward location of ENSO-related SST anomalies, a model error that is more pronounced in EC-Earth3-SR. Although these ENSO-related errors are present in other forecast systems and well documented (**??**), it is difficult to disentangle their ultimate cause, and in particular if it is an ocean, atmospheric or coupled systematic error of the model or just an initialisation shock response in one of the components. **?**, using

11 different operational models, argued that forecast errors develop within days after initialisation and seem to develop first in the atmosphere before subsequently influencing SST and precipitation errors. In both configurations of the EC-Earth3 forecast system, the errors appear rapidly in the surface zonal winds and the mixed layer depth. Determining the origin of the errors would require further investigations using sensitivity experiments, which is beyond the scope of this study.

The poorest skill in EC-Earth3-SR compared to EC-Earth3-HR in the WEP region, associated with larger ENSO-related errors, seems linked to larger mean state biases, and in particular the "cold tongue bias", a common bias in GCMs (**??**)~~, which leads to larger ENSO-related errors~~. Several studies showed that using higher resolution models could reduce the cold bias of the cold tongue (**?????**). In particular, **?** showed that the ENSO-related SST anomalies differences between SR and HR configurations are caused by differences in atmospheric convection related to surface zonal wind differences. The differences in atmospheric convection have an impact on important ENSO feedbacks, such as the Ekman feedback, the zonal advection feedback and the thermocline feedback, which play a key role in the growth rate of ENSO events (e.g., **?**). They suggested that these differences in ENSO-related variability, in SST, surface wind and convection, can arise from the mean state bias more pronounced in the SR models, with a westward extension of a colder bias in the tropical Pacific. In the case of the more pronounced cold mean state bias, the associated convection has to be located more to the west to react to the warm SST anomalies starting ENSO development. Indeed, the convection process requires an absolute SST threshold to happen, instead of an anomaly threshold (**?**). The lower (higher) mean biases that improves (degrades) ENSO variability may explain the higher (lower) ENSO predictions in EC-Earth3-HR (EC-Earth3-SR). It remains to be assessed if similar improvements are consistently identified for HR versions in other forecast systems.

Finally, the stronger decrease of skill in EC-Earth3-SR is also likely related to a weaker air-sea coupling in EC-Earth3-SR than in EC-Earth3-HR, which prevents the model from simulating the correct ENSO development. Analysing the impact of the weak air-sea coupling alone in the predictions is beyond the scope of this study but is of particular importance not only for climate predictions but also for future scenario experiments.

The second studied mechanism is the Atlantic Niño teleconnection with the tropical Pacific. In EC-Earth3-HR, a ~~better~~ more realistic simulation of the teleconnection leads to an increase in the predictive skill of ENSO compared to EC-Earth3-SR. We also note that the predictive skill in the Atlantic Niño region itself appears to be insensitive to model resolution, decreases rapidly after initialisation in both systems, and cannot explain the improved ENSO skill for EC-Earth3-HR. Improving the Atlantic variability and removing mean state biases is particularly important to improve Equatorial Atlantic (**??**) and ENSO prediction (**?**). However, **?** argued that the Atlantic Niño teleconnection with ENSO could be just a statistical artifact due to ENSO's autocorrelation. The consistency and impact of the teleconnection signal require further investigation.

The low-frequency variability also impacts the predictability of different variables and regions. For instance, **?** showed that, since 1950, ENSO forecast skill monotonically increased to reach high values in the late 20th century and started to decline in the last decades. **?** showed that the central tropical Pacific is particularly sensitive to decadal internal variability linked to low-frequency oscillations such as the Tropical Pacific Decadal Variability (TPDV). The impact of these low-frequency variability features on the predictive skill cannot be evaluated in our relatively short seasonal forecasts but could be addressed in the new multi-year climate predictions that are being considered for the next phase of DCPP. These envisage substantially higher

hindcast periods than the ones we considered in this study, long enough to include several phases of low frequency modes, as well as longer forecast horizons to assess if the added predictive value of high resolution persists after the first forecast year.

These new multi-annual DCPP predictions, with several start dates per year and covering a longer period than usual seasonal forecasts, will also enable a more detailed analysis of the impact of the mean state bias and its seasonal variation on the predicted development of ENSO events. **?** showed that, in the operational forecast systems they analysed, the ENSO-related errors are mainly driven by the seasonal cycle rather than the forecast lead time, a feature that can be analysed further with the proposed new-DCPP. Moreover, with the longer hindcast period, this analysis can be carried out by composites, for each type of ENSO event (El Niño, La Niña and neutral events), thanks to a larger number of samples. Indeed, one of the limitations of this study is that, through the use of linear regression, El Niño and La Niña events are considered as linearly opposed, a limitation that could be overcome by increasing the number of simulated events for each phase.

All in all, using higher resolution model components can improve the predictive capabilities of forecasting systems, as this is the case for the EC-Earth3 model for predictive skill of the tropical Pacific on a seasonal scale when initialized in May. However, the ~~contribution is less noticeable~~ added value of resolution is less evident in other parts of the globe or when predictions are initialized in November~~(not shown), which~~ . Indeed, for the November-initialized predictions, ENSO skill shows no statistically significant difference between the two configurations (Supplementary Figure 2). The results of higher predictive skill with high resolution but depending on the season, region, and variables analysed, raises the question of the balance between benefits and challenges. Using a previous version of EC-Earth3, **?** showed that increasing horizontal resolution does not result in a general improvement of the climate variability. However, the HR configuration of EC-Earth3 is 16 to 20 times more expensive in terms of computing resources than the SR configuration, to which must be added the storage and post-processing capacities of the associated heavy outputs. Conversely, standard resolution simulations are more economical to produce and store, which facilitates and accelerates the production and analysis of long and large ensembles of simulations. These can be leveraged to test model developments, refine the model tuning, produce sensitivity simulations designed to address specific scientific questions, constrain more accurately the predictable signals, and increase model complexity at a lower cost, all with the aim of improving our understanding of the climate system. The choice of modelling strategy remains dependent on the scientific question and the intended application, as all strategies are complementary (**???**).

In conclusion, this study highlights the key role of model errors, both in terms of variability and mean state, in predictive skill in the equatorial Pacific of both forecast systems. Improving the relationship between ENSO and the WEP regions, whose erroneous simulation is related to the erroneous westward extension of ENSO-related variability, ~~itself linked to SST and winds mean state and variability biases in the models, could lead to a better~~ could improve predictive skill in the western Pacific and over the maritime continent, a key region for ENSO-related air-sea teleconnections (**???**). At the same time, improving teleconnection between the Atlantic Niño and the equatorial Pacific could improve the prediction of ENSO events. A better understanding of the sources and development of the models' errors is essential to improving predictive systems, and this regardless of the horizontal resolution, since most of the biases persist in high resolution, even if some have been improved. The new multi-annual DCPP hindcasts could contribute to fill this gap.

Data availability. Data availability

Data from the EC-Earth3-SR and EC-Earth3-HR seasonal prediction simulations, as well as the ocean and sea-ice reconstructions used to initialise the prediction simulations, are available upon request.

ERA5 data are available via https://cds.climate.copernicus.eu/datasets/reanalysis-era5-complete. ORAS5 data can be downloaded from https://cds.climate.copernicus.eu/datasets/reanalysis-oras5. EN.4.2.2 data were obtained from https://www.metoffice.gov.uk/hadobs/en4/ and are © British Crown Copyright, Met Office, 2013, provided under a Non-Commercial Government Licence http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/. ESA SST CCI v2.1 level 4 data can be downloaded via https://catalogue.ceda.ac.uk/uuid/62c0f97b1eac4e0197a674870afe1ee6/.

Code availability. Code availability

EC-Earth: The EC-Earth model is restricted to institutes that have signed a memorandum of understanding or letter of intent with the EC-Earth consortium and a software license agreement with the ECMWF. Confidential access to the code and to the data used to produce the simulations described in this paper can be granted for editors and reviewers; please use the contact form at http://www.ec-earth.org/about/contact.

Post-processing analysis: Ocean diagnostics have been computed using Earthdiagnostics, a Python-based package developed at the BSC (https://earth.bsc.es/gitlab/es/earthdiagnostics, BSC-CNS and Vegas-Regidor, 2020). For data loading, post-processing and calculation of verification metrics, the startR (https://cran.r-project.org/web/packages/startR/index.html, **??**, v2.3.1), s2dv (https://cran.r-project.org/web/packages/s2dv/index.html, **??**, v2.0.0) and ClimProjDiags (https://earth.bsc.es/gitlab/es/ClimProjDiags, **?**, v0.3.3) R libraries have been used. The codes used for the results and figures are available upon request.

Author contributions. Author contributions

AC led the analysis. AC, PO, RB and FDR designed the experiments and AC carried them out. AC, VL, RB and FLM contributed to creating the initial conditions for the seasonal forecast systems. AC, CDT and RB developed the codes of analysis. PO, MD and FDR contributed to the interpretation of the results. AC prepared the manuscript with contributions from all co-authors.

Competing interests. Competing interests

The authors declare that they have no conflict of interest.

Financial support. Financial support

This study was supported by the Horizon Europe project ASPECT “Adaptation-oriented Seamless Predictions of European ClimaTe” (grant agreement number: 101081460) and the "SupercompuTing And Related applicationS Fellows Program" (STARS) programme, part of the COFUND call of the Marie Skłodowska Curie Actions (MSCA) within the Horizon 2020 Programme (grant Agreement number: 754433). We are also grateful for support by the Departament de Recerca i Universitats de la Generalitat de Catalunya for the Climate Variability and Change (CVC) Research Group (reference: 2021 SGR 00786).

Acknowledgements. Acknowledgements

Acknowledgements. The authors acknowledge Eric Ferrer, Miguel Castrillo, Pierre-Antoine Bretonnière for their technical support with the model and the workflow manager; Juan Acosta-Navarro and Etienne Tourigny for their help in the development of the initialisation protocol; An-Chi Ho, Victòria Agudetse Roures and Núria Pérez-Zanon for their help and support in the use of the R packages; Eneko Martín and Alba Santos Espeso for their help in the development of the mixed layer depth script; Saskia Loosveldt-Tomas for her support in using the Earthdiagnostic tool; Margarida Samso Cabre for her help to retrieve the data; Paloma Trascasa-Castro, Yohan Ruprich-Robert, Verónica Martín Gómez and Ángel G. Muñoz for their scientific contribution to the development of this study. The EC-Earth3 simulations and data handling were performed using the internal computing resources available at MareNostrum and additional resources from the Red Española de Supercomputación (RES-AECT-2021-2-0028 and RES-AECT-2022-2-0005 projects). Acknowledgement is also made for the use of ECMWF’s computing and archive facilities in this research via special projects. The authors acknowledge the technical support provided by Barcelona Supercomputing Center and the Computational Earth Sciences group of the Earth Sciences department. We also want to thank the reviewers for their constructive comments.