

# Skill assessment of a set of retrospective decadal climate predictions with EC-Earth

Author: JAUME RUIZ DE MORALES CÉSPEDES

Supervisors: FROILA PALMEIRO NUÑEZ<sup>1</sup>, fm.palmeiro@meteo.ub.edu, ROBERTO BILBAO<sup>2</sup>, roberto.bilbao@bsc.es, PABLO ORTEGA<sup>2</sup>, pablo.ortega@bsc.es

<sup>1</sup> *Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain. and*

<sup>2</sup> *Earth Sciences Department, Barcelona Supercomputing Center (BSC), Barcelona, Spain,*

**Abstract:** The climate system is changing with unprecedented consequences for the environment and many socioeconomic sectors. Hence the importance of predicting these changes. This study aims to produce an evaluation of the predictive skill in a decadal prediction system performed with EC-Earth. It specifically targets three variables of high relevance for human activities, such as sea surface temperature, the sea surface height anomaly (which quantifies sea level rise) and the total cloud cover (which is critical for storm development). The evaluation has mostly focused on two major ocean basins (Pacific and Atlantic), where important modes of variability like the El Niño-Southern Oscillation and the Atlantic Multidecadal Variability take place, and also on the Equatorial stratosphere, where the Quasi-Biennial Oscillation, a highly predictable mode, occurs. Concerning the results, we have shown high prediction skill for all variables in the first forecast year. In the following years, we note a general reduction of the predictive skill, particularly in the southeastern Tropical Pacific, which might point to deficiencies in the model to simulate ENSO periodicity and/or regionality. Furthermore, a general lack of skill in the North Atlantic, may imply that the Atlantic Multidecadal Variability, at least in EC-Earth, is not a source of sea level predictability. Regarding the QBO, results have shown a high prediction skill, especially in the first 29 months. However, the QBO cycle periodicity is not well represented by EC-Earth, which degrades the credibility of the predictions in the subsequent forecast years.

## I. INTRODUCTION

The Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) clearly states that “The warming of the climate system is undeniable, each of the last three decades has been successively warmer than any previous one since 1850” (IPCC 2014). This global warming is one of the most clear expressions of climate change, but not the only one. Important changes have been observed in several climate subsystems: the atmosphere, the hydrosphere, the cryosphere, and the biosphere.

While both the atmosphere and the ocean have been warming, following the continuous increase in CO<sub>2</sub> concentrations in the atmosphere, other important impacts have also been documented (Stocker *et al.* 2013). For example, the rate of ocean sea-level rise since 1850 has been greater than for the past two millennia, putting coastal regions at risk. Regarding the cryosphere, during the last two decades Greenland, Antarctica and the continental glaciers have shown a significant reduction of ice mass, which could partly explain the observed sea level rise (IPCC 2014).

Therefore, in this context in which the climate system is changing, it is crucial to anticipate these variations that can potentially have unprecedented consequences in the environmental system and also in many socio-economic sectors (Yeager *et al.* 2018). Hence the importance of performing and evaluating decadal predictions, which can provide critical climatic information up to 10 years ahead. These are performed with numerical global climate models, very similar to those used to per-

form the weather forecasts, but cover radically different timescales (Meehl *et al.* 2021).

In the weather forecasts, it is particularly important to constrain well the present state of the Earth system from observations, because an imprecise definition (or error) of the observed initial state can lead to hugely different results due to the Lorentz butterfly effect. This “initial value problem” is central to weather prediction, as the growth of these initial errors is what limits the accuracy of the forecasts to only a few days. Interestingly, components of the climate system, like the ocean and cryosphere, present slowly-varying processes that can provide predictive capacity at the climate timescales (seasons, years and decades). Initializing accurately the observed state of those components is thus very important for climate prediction, in particular for the first forecasts months and years. Another source of predictive skill, which is more important at longer timescales (several years to decades), derives from the radiative changes caused by variations in external forcings, such as solar irradiance or the concentrations of greenhouse gases and aerosols, which control the amount of energy that enters and remain in the climate system. This is thus a “forced boundary condition problem”, and is very important to make well-informed climate change projections for the next future.

Having accurate decadal predictions thus relies on resolving well both the initial value and the forced boundary condition problems. The latter is addressed by including educated estimates of future changes in **radiating forcing** and the former by ensuring that the low-frequency **internal variability** modes of the climate are

well initialised with observations, and realistically simulated by the climate models (Meehl *et al.* 2014, 2021, Yeager *et al.* 2018).

This work will assess the skill in a decadal prediction system performed with the global climate model EC-Earth, paying particular attention to the contributions of internal climate variability to forecast skill.

## II. BACKGROUND

This section provides (I) an overview of the main modes of internal climate variability that operate at interannual to decadal timescales and can thus provide predictive skill, and (II) the list of variables whose skill will be assessed, and an explanation of their relevance.

### A. Internal climate variability modes

We focus on the two major ocean basins (Pacific and Atlantic) and a purely stratospheric mode with important worldwide impacts.

**PACIFIC OCEAN:** In this region, El Niño-Southern Oscillation (ENSO) is a climate pattern that repeats every few years, with teleconnections and repercussions all across the globe. The chain of events behind the ENSO phenomenon is explained in the following. In the Tropical Pacific, trade winds (i.e. easterlies) prevail, pushing the generally warm ocean surface masses to the west. At the eastern side, the wind causes ocean upwelling of relatively cold water masses from the deeper ocean, which increases the supply of nutrients into the surface, with positive effects for the fishing sector in the area. The relatively warmer waters in the western Tropical Pacific lead to increased evaporation and atmospheric convection, creating a zone of low pressure, which contrasts with the conditions in the cold eastern Tropical Pacific, where a semi-permanent subtropical anticyclone is established (i.e. the South Pacific High). In a given year, this dipole of high and low pressures between the eastern and western Tropical Pacific can be strengthened or weakened with respect to the climatological state, giving rise to the two opposite phases of ENSO: El Niño and La Niña (Enfield *et al.* 2001, Luo *et al.* 2008, Zhang and Zhao 2015).

*El Niño*, the warm phase, is characterised by warmer than average surface waters in the equatorial eastern Pacific, caused by a weakening of the trade winds along the equator. This phase has negative effects on fisheries, agriculture, and the local climate (Enfield *et al.* 2001).

*La Niña*, the cold phase, consists of the opposite phenomenon, i.e. the presence of colder than average SST in the equatorial eastern Pacific, caused by an intensification of the trade winds. In this phase, the upwelling on the eastern side is intensified, bringing more nutrients to the surface and creating good conditions for fishing.

ENSO oscillates irregularly between its two phases with a periodicity between 2-7 years (Zhang and Zhao 2015). These changes can be predicted from months to

even years in advance (Luo *et al.* 2008). ENSO can also present decadal modulations, typically referred to as the Pacific Decadal Oscillation (PDO), which has a very active center of action in the subpolar North Pacific.

**ATLANTIC OCEAN:** The major mode of internal variability in this ocean is the Atlantic Multidecadal Variability (AMV), which is characterised by a persistent and coherent pattern of SST anomalies (either positive or negative) occupying the whole North Atlantic basin (Enfield *et al.* 2001). As a consequence of this phenomenon, the North Atlantic SST experiences prolonged periods of warm and cool conditions, referred to as 'positive' and 'negative' AMV phases.

During a positive AMV phase, the warmer waters have more available energy for developing storm systems, resulting in an increase in the frequency of hurricane activity. Other reported impacts for this phase are increased summer rainfall in Northern Europe, heavier rainfall in the African Sahel region, droughts in northeastern Brazil, drier conditions in North America, reduced sea ice in the Arctic and stronger summer monsoon in India (McCarthy *et al.* 2015, Sutton *et al.* 2018). Negative AMV phases are associated with the opposite impacts.

The processes controlling the AMV are still under debate, although it is generally assumed that its low-frequency variations arise from changes in the strength of the Atlantic Meridional Overturning Circulation (Knight *et al.* 2005), ultimately driven by deep water formation processes in the Labrador and Greenland Seas. Thanks to the decadal nature of the AMV fluctuations, the North Atlantic is a region of a high predictive skill at decadal timescales (Collins *et al.* 2006).

**THE QUASI-BIENNIAL OSCILLATION (QBO):** The QBO is a periodical fluctuation of the zonal wind direction in the equatorial stratosphere. Strong winds over the equator change regularly their direction every 14 months, leading to an alternation of easterly and westerly winds. Therefore, every 28 months there is a complete cycle. The driving forces of the QBO are the atmospheric waves that rise from the troposphere, produced by the intense convection of tropical systems. These waves propagate and disturb the stratosphere, providing a force to move the wind that will be descending over time (Baldwin *et al.* 2001, Palmeiro *et al.* 2020).

Studying and understanding the QBO is essential, as it has effects on both the subtropical jet and the polar jet. The speed of the jet winds is intensified or attenuated depending on the direction of the QBO. It should be emphasised that the polar jet, the subtropical jet and the Intertropical Convergence Zone (ITCZ) are the planetary weather drivers, with repercussions across the globe. In general, when the QBO is easterly, the jet streams weaken, and there's more chance of a sudden stratospheric warming event and colder winters in Northern Europe. On the other hand, when the QBO is westerly,

the jets intensifies, with moderated winter conditions, although with more winter storms and heavy rainfall (*Baldwin et al. 2001*). Due to its strong regularity and its influence on the Northern Hemisphere mid-latitude jet and the Arctic polar vortex, the QBO is a phenomenon of particular interest for climate prediction.

### B. Choice of variables for the skill assessment

The typical protocol to assess the predictive skill of a given decadal prediction system is to perform a set of retrospective predictions, so that it can be contrasted against the observed variability to determine where and for which variables the predictions are useful. The choice of variables assessed is important, and priority is usually given to those that have been observed for a long period and with a good spatial distribution. Satellite data is particularly useful to this end, given the continuous global coverage they provide. We will focus on three variables for which at least three decades of satellite observations are available, and that are known to be sensitive to internal climate variability. Since the QBO is not directly observed, we will evaluate the ability of our system to predict its against a selection of atmospheric reanalyses.

**SEA SURFACE TEMPERATURE (SST):** SST is defined at the interface between the ocean and the overlying atmosphere, and it is thus able to capture the interactions between both realms. It has always been a variable of interest for the scientific community, as “it controls the heat, momentum, salt and gas fluxes between the ocean and the atmosphere” (*Emery 2015*).

The SST is the easiest oceanographic variable to observe and measure. Over the last 150 years, it has been measured by ships and by both moored and drifting buoys. In recent decades there has been an important change in the methods for measuring the SST, as satellites are now also being used to that end. This has revolutionized its spatial resolution as well as coverage, allowing to systematically measure over regions of difficult access, like the poles (*Emery 2015, Matthews 2013*).

**SEA SURFACE HEIGHT ANOMALY (SSHA):** Sea level rise, typically measured as a sea surface height anomaly with respect to a given temporal period, is one of the most direct and serious impacts of climate change. Contemporary sea level rise results principally from (I) the thermal expansion of the ocean as it stores the excess of energy entering the Earth system and (II) the transfer of freshwater from land into the ocean, particularly from glacier and ice-sheet melting (*Stocker et al. 2013*).

Regional changes in sea level follow local changes in ocean density which can obey to changes in temperature and salinity (referred to as thermohaline and halohaline sea level change, respectively) and are controlled by changes in the heat, freshwater and wind-stress surface fluxes, resulting in changes in density and ocean circulation (*Gregory et al. 2019*). Water and ice redistribution between the land and ocean may also affect the regional

sea level by changing the Earth’s gravitational field and rotation (*Stocker et al. 2013*). Changes in atmospheric pressure also affect regional sea level through the inverse barometer effect.

Sea-level rise is expected to have massive worldwide impacts. Associated impacts include contamination of drinking water due to saline intrusion, an increase in the power of storms and coastal recession due to marine erosion, enhancing the risk of flooding (*Bosello et al. 2007*)

**TOTAL CLOUD COVER (TCC):** Cloud cover is an essential component of the Earth system due to its role in climate regulation. Its ability to affect radiative forcing (short and long waves) makes it a key contributor to climate feedback mechanisms (*Costa-Surós et al. 2013, Ramanathan et al. 1989*). However, determining how clouds contribute to climate change is difficult, due to the complexity of the processes involved, the big quantity of information needed, the spatial distribution and the uncertainty associated with the available data (*Costa-Surós et al. 2013, Solomon et al. 2007*). Furthermore, not all clouds have the same effect on radiation. Clouds come in a variety of dimensions, opacities, and properties, which have a very different effect on solar short-wave and longwave radiation (*Tapakis and Charalambides 2013*). Clouds reflect part of the sun’s radiation back into space, a process that tends to cool the climate, but they also absorb and emit infrared radiation towards the earth’s surface (greenhouse effect), a process that tends to warm it. Satellite data suggest that in general cloud cover represents a cooling effect on our planet. This means that without cloud cover, our atmosphere would be warmer (*Ahrens 2009*).

Currently, we are facing a global warming mainly caused by anthropogenic greenhouse gases. As the atmosphere is warming, it is expected that more water vapour will be added to the air and total cloudiness could increase as well (*Costa-Surós et al. 2013, Enríquez-Alonso et al. 2017*). Given the important role of TCC both on the global and local climate, and its close link to SSTs (i.e. warm waters promote evaporation and cloud formation), we will evaluate the ability of our model to predict its changes.

## III. DATA & METHODOLOGY

### A. Observational Data

In this study, several reanalyses and observational data from different sources have been used, to evaluate the consistency across the products before the prediction skill of the model is evaluated.

Gridded observational data has been used for analysing SST, SSHA and TCC (see Table 1 for further details). Specifically, for SST, the L4 product of ESA (*Merchant and Embury 2020*), HadISSTv1.1 (*Rayner et al. 2003*) and ERSST (*Huang et al. 2017*) were used. C3S vDT2018

(*Taburet et al.* 2019), and CMEMS L4 were used for SSHA. Finally, EUMETSAT CLARA-A2.1 (*Karlsson et al.* 2017) and ESA AVHRR-PM v3.0 (*Stengel et al.* 2020) were used for the evaluation of TCC. The QBO was analysed using zonal wind data between 10-200hPa from the ERA5 (*Hersbach et al.* 2020), ERAinterim (*Dee et al.* 2011) and JRA-55 (*Kobayashi et al.* 2015) atmospheric reanalyses. For all variables and reanalyses a merged product, computed as the average across datasets on each time step over their overlap period, has been produced to isolate the common variability. Prior to this, all products were interpolated into a regular  $1^\circ \times 1^\circ$  grid. Forecasts were evaluated against the merged products.

**B. Model Data & experiments**

In this study we have analysed the retrospective decadal predictions performed within the Coupled Model Intercomparison Project Phase 6 (CMIP6) with the global Atmosphere-Ocean General Circulation Model (AOGCM) EC-Earth3, which has a nominal resolution both in the ocean and the atmosphere of  $1^\circ$ . Further information about the climate model used and its components can be found in *Bilbao et al.* (2021).

CMIP6 is coordinated by WCRP’s Working Group on Coupled Modelling (WGCM). The CMIP6 supports 23 Model Intercomparison Projects (MIPs), that have common protocols and standards, including the use of fair data and free access. The retrospective decadal predictions herein analysed follow the experimental protocol recommended in the Decadal Climate Prediction Project (DCPP; *Boer et al.* 2016), that is one of the MIPs aforementioned. DCPP focuses its research on the ability to predict climate variations with forecast times of up to one decade ahead. In order to achieve skilful forecasts at these timescales, DCPP includes idealised predictability studies and retrospective forecasts using climate models and statistical methods.

The decadal predictions with EC-Earth include yearly start dates for the period 1960-2020, each producing 10-member predictions for the next 10 years after initialization.

**C. Methodology**

An important issue in climate prediction is the forecast drift, which stems from the fact that models are initialised from an observed state that is not necessarily compatible with the model preferred state, which is generally biased. As the predictions progress, they thus drift towards the preferred model state. In order to avoid forecast errors associated to this drift, it is very important to correct its effects. This can be partly accomplished by evaluating the predictions against observations in anomaly space (e.g. *Bilbao et al.* 2021, *Meehl et al.* 2014). For this so-called “mean drift correction” method to be effected, the model anomalies need to be computed with respect to a forecast time dependent climatology.

The skill of the predictions has been evaluated with Anomaly Correlation Coefficients (ACCs), that is just a linear correlation between the observed and predicted anomalies after the mean drift correction (see Appendix for its actual definition).

Table I. Observation-based products used in the study

Variable	Products	Time Coverage	Resolution
SST (°C)	ESA L4	01-1982/12-2020	$0.05^\circ \times 0.05^\circ$
	HadISSTv1.1	01-1870/12-2020	$1^\circ \times 1^\circ$
	ERSSTv5	01-1854/12-2020	$2^\circ \times 2^\circ$
	Merged product	01-1982/12-2020	$1^\circ \times 1^\circ$
SSHA (m)	C3S vDT2018	01-1993/10-2019	$0.25^\circ \times 0.25^\circ$
	CMEMS L4	01-1993/02-2020	$0.25^\circ \times 0.25^\circ$
	Merged product	01-1993/12-2018	$1^\circ \times 1^\circ$
TCC (%)	EUMETSAT CLARA-A2.1	01-1982/05-2019	$0.25^\circ \times 0.25^\circ$
	ESA AVHRR-PM v3.0	01-1982/12-2016	$0.5^\circ \times 0.5^\circ$
	Merged product	01-1982/12-2016	$1^\circ \times 1^\circ$
Ua (m/s)	ERA5	01-1979/12-2020	$0.25^\circ \times 0.25^\circ$
	ERA-Interim	01-1979/08-2019	$0.25^\circ \times 0.25^\circ$
	JRA	01-1958/12-2020	$1.25^\circ \times 1.25^\circ$
	Merged product	01-1979/08-2019	$1^\circ \times 1^\circ$

**IV. RESULTS & DISCUSSION**

**A. Consistency across observations**

For increasing the confidence in the skill assessment, it is important to make a first assessment of the consistency (or degree of agreement) between the different observational products. The temporal consistency across the products has been measured with temporal correlations between each pair of datasets. Also, to have a more reliable observational reference, we have computed for each time step the average between the different products available, so that observational errors in each dataset can, at least partly, cancel with each other. The long-term mean of each merged observational dataset and its inter-annual standard deviation are shown in Figures 1 and 2 for illustrative purposes.

SEA SURFACE TEMPERATURE: The climatology of SSTs (see Fig. 1) shows the typical pattern of maximum temperatures ( $\sim 30^\circ\text{C}$ ) at the equator and in tropical areas (where the absorption of solar radiation is also maximum), and minimum temperatures ( $\sim 0^\circ\text{C}$ ) close to the poles. The regions showing largest interannual variability in SST, as indicated by their high standard deviation, are the Equatorial Pacific Ocean, where ENSO takes place, as well as the Labrador Sea and Gulf Stream region in the North Atlantic. Large standard deviations are also seen in the Northern Pacific ocean, both in the region of the Kuroshio current and the West Coast of the United States. These regions of high variability are the ones in which having high predictive skill is potentially more important.

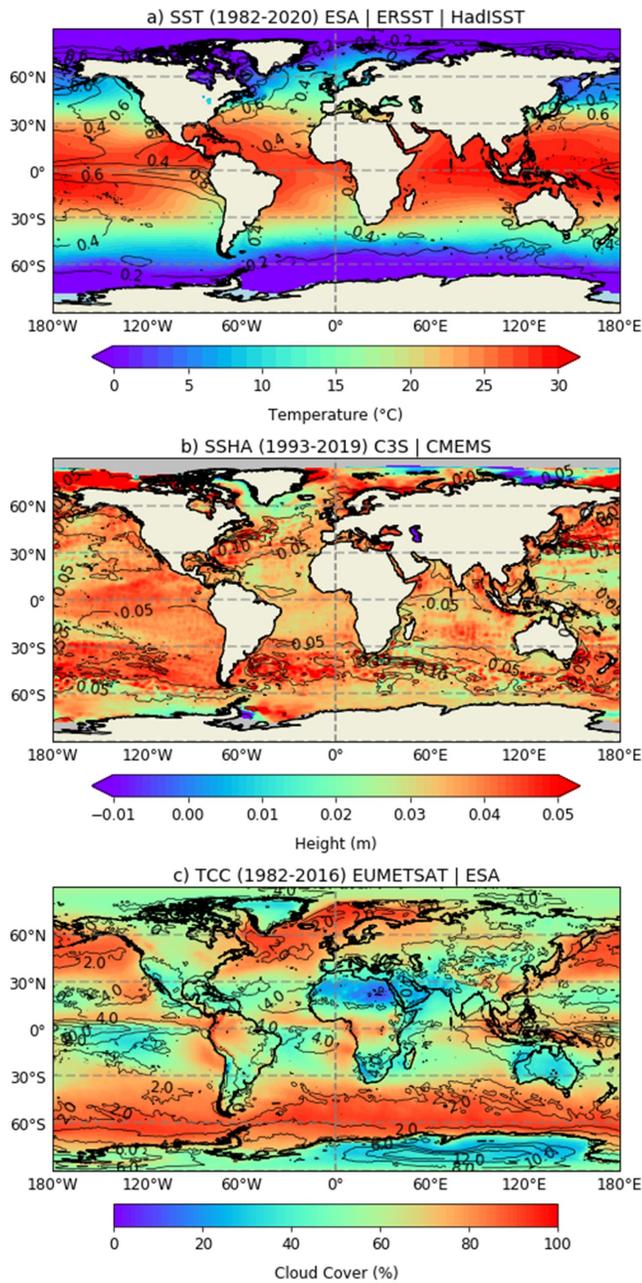


Figure 1. Temporal climatology (shaded contours) and Standard Deviation (black contours) of the merged observational datasets (see Table 1) of a) SST (period 1982-2020), b) SSHA (period 1993-2019) and c) TCC (period 1982-2016).

In general, the correlations between the different products (ESA, ERSST, HadISST) are very high (above 0.7). The minimum correlation value between any given pair of products is shown in Fig. 2. Interestingly, all the regions that exhibited high standard deviation values also show high correlation values across products (above 0.9), except for the Gulf Stream region. The Southern Ocean is the region of poorest agreement across datasets, which is a consequence of the reduced availability of observations

in the region, which increases the uncertainty.

In order to provide a temporal illustration of the level of agreement between the different products, we have analysed the temporal evolution (see Figure 3) of two main regions: the North Atlantic Ocean (10-60°N and 80-15°W) and the Tropical Pacific (10°S-20°N and 180-100°W). These two regions were chosen for their importance in climate variability and prediction and their global teleconnections, mainly mediated via ENSO and the AMV.

We can see that while a clear warming trend is seen in the North Atlantic (Fig. 4a), supported by the three products, the Tropical Pacific remains rather flat and exhibit pronounced interannual variability (Fig. 4b), related to the phenomenon of El Niño and La Niña. For this later region it can also be seen that there is a strong agreement among the products for the whole time period. In the North Atlantic, there is also a strong agreement across products in terms of the variability, but seem to exhibit systematic differences in the mean state. This is perhaps caused by the fact that the products originally have different resolutions and have been interpolated into a common grid.

**SEA SURFACE HEIGHT ANOMALY:** The SSHA climatology of the merged product (see Fig. 1), shows a global sea-level rise of about 0.03-0.04 m between 1993-2019 (we note that the anomaly is computed with respect to the reference period 1993-2012). The areas with the highest positive SSHA are the Caribbean Sea, the Arctic and Antarctic Seas, the Kuroshio region, and the Sub-polar Southern Hemisphere (i.e. between 30°-60°S). We highlight also some areas showing a negative SSHA (indicative of a mean sea level reduction during 2013-2019) which include northeast Siberia, and a small part of the Weddell Sea, both of which might be due to local sea ice growth. The highest standard deviation values are achieved in the regions of the western boundary currents, both in the Northern and Southern Hemisphere. This includes the Gulf Stream and Kuroshio, Agulhas and East Australia Currents.

We highlight that the two products considered also show very strong correlations of more than 99(%) everywhere excepts in the regions of sea ice (Fig. 2). This might be due to the fact that both products derive from the same raw satellite information, and only differ in the calibration algorithm (Cazenave *et al.* 2019), which implies that they might be subject to similar observational uncertainties.

The SSHA from 1993 to 2018 shows an increase in both the North Atlantic Ocean and the Tropical Pacific regions (Figure 3). On the one hand, the North Atlantic Ocean shows a steady increase, with small year to year variations. This trend seems to have intensified in the last 8 years (2010-2018). On the other hand, the Tropical Pacific also shows a long-term SSHA rise, but in this case modulated by important year-to-year variations related

to ENSO (e.g. associated with the strong 1998 el Niño and 1999 la Niña events. The two products show very similar trends and variability in both regions.

**TOTAL CLOUD COVER:** cloudiness is a variable with large regional and temporal variability, as it is can influenced by several factors, as SSTs, the general atmospheric circulation, orography, frontal masses, and convection. The climatology of the merged product of TCC (Fig. 1) shows maximum values in the Intertropical Convergence Zone (ITCZ), and in the subpolar latitudes of both hemispheres, in particular over the ocean. In contrast, the minimum values take place in the subtropical latitudes over continental areas, especially where large deserts such as the Sahara, the Kalahari Desert and the Australian deserts are located. There is also a cloud cover minimum over Antarctica and North of Greenland.

TCC is very heterogeneous, as the standard deviation across the world map ranges from 0-14, with minimum standard deviations in Antarctica and maximum values in the central ENSO zone. Large correlations (typical of 0.7 or higher) and therefore strong agreement in the variability of the two observational products are found everywhere, in particular over the ocean regions (Fig. 2).

As far as cloudiness is concerned, no clear trend is seen in the period studied in any of the regions, the North Atlantic Ocean or Tropical Pacific, due to its high variability (Figure 3). Both regions, and the Tropical Pacific in particular, show important year-to-year oscillations, with some notable differences between the two observational products, higher than for the two previous variables.

**QUASI-BIENNIAL OSCILLATION:** An altitude-time cross-section of the zonally-averaged zonal wind at the Equator shows 17 westerly and easterly phases of the QBO between 1979-2019 (Fig. 4).

It should be noted that the standard deviation between the products (which measures their degree of disagreement) is greater in the 10-20 hPa levels, and especially in the early years. This is expected as the accuracy of observational data, including from satellites, assimilated by the reanalyses have been improving over time.

The QBO index is typically computed as the zonally-averaged zonal wind at the Equator and 50 hPa. Figure 4b shows its evolution for the individual and the merged reanalyses. They all show a large degree of agreement, both in terms of the intensity and duration of the easterly/westerly phases, although JRA-55 generally produces slightly higher values, while ERAinterim slightly lower ones. No clear trend can be seen for the QBO.

**B. Predictive skill assessment**

This subsection evaluates the skill of the decadal predictions for the three chosen variables against the respective merged products. Different forecast times have been considered: the first forecast year, the average of the

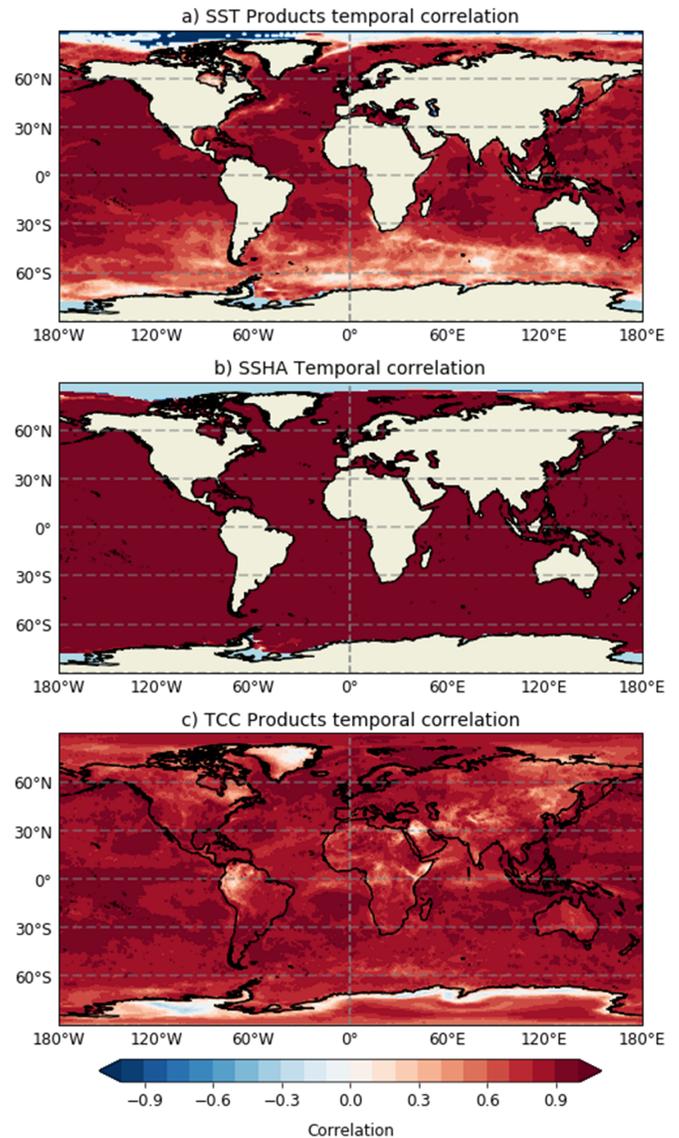


Figure 2. Temporal correlation across the observational products of (a) SST, (b) SSHA and (c) TCC. For every grid-point the minimum correlation value between products is shown.

first 5 forecast years, and the average of the last 5 forecast years (Fig. 6). We have also computed the skill for the QBO index, in this case as a function of the forecast month (Fig. 6).

**SEA SURFACE TEMPERATURE:** Overall, from all of the three variables that have been evaluated, SST has the highest forecast skill. This is somewhat expected, as SST is a variable that is directly assimilated when producing the initial conditions from which the predictions are started. None of the other 2 variables is directly assimilated. During the first year of the prediction, there are large and significant ACC values in almost all oceans basins. We remind that this skill comes from internal variability, through initialization, but also from the repre-

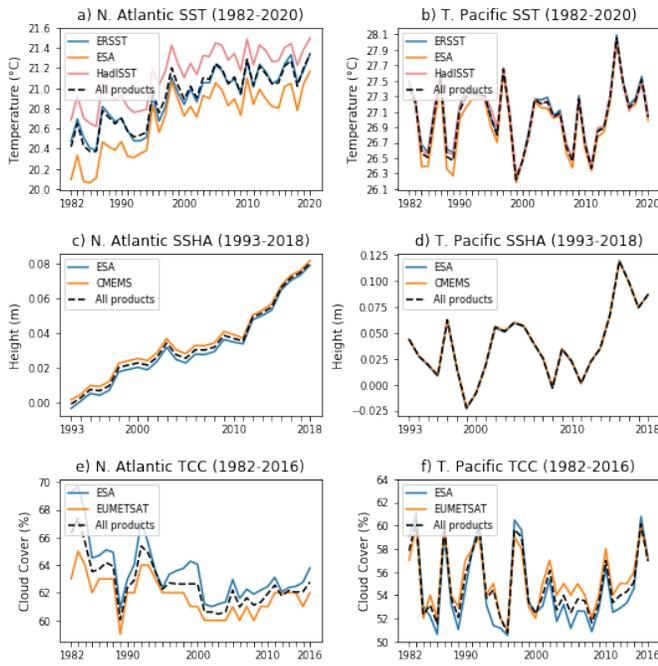


Figure 3. Regional averages of SST, SSHA and TCC in the North Atlantic and Tropical Pacific for the individual and merged observational datasets.

sensation of the externally forced trends. When taking an average of the first 5 forecast years, ACC values remain significant in most ocean regions, except in the Central region of the North Atlantic Subpolar Gyre, the Southern Ocean all along the Antarctic Circumpolar Current, and the southeastern Tropical Pacific. ACC values are higher than for forecast year 1, but this simply reflects the fact that in this case skill is computed for 5-year averaged timeseries, which screens out the interannual variability. Regarding the forecast skill for the average of the 6th to 10th forecast years, ACC values are practically the same as for the first 5, although some significant skill is lost in the central and eastern Tropical Pacific. The fact that this region only exhibits significant skill in the first forecast year suggest that the benefits of initializing the correct ENSO phase are lost quickly, pointing to deficiencies in the model when simulating ENSO periodicity.

**SEA SURFACE HEIGHT ANOMALY:** As previously mentioned, changes in sea level are driven by different processes, like the thermosteric and halosteric changes in sea water density, and changes in the ocean circulation and atmospheric pressure. However, AOGCMs like EC-Earth can simulate some but not all of these contributions. The governing physical equations included in EC-Earth explicitly resolve the changes in surface wind stress and the heat and freshwater surface fluxes, as well as their influence on the density fields and the ocean circulation. These are the main causes of regional sea level. EC-Earth, however, does not include glacier and ice-sheet models and cannot therefore simulate the role of glacier

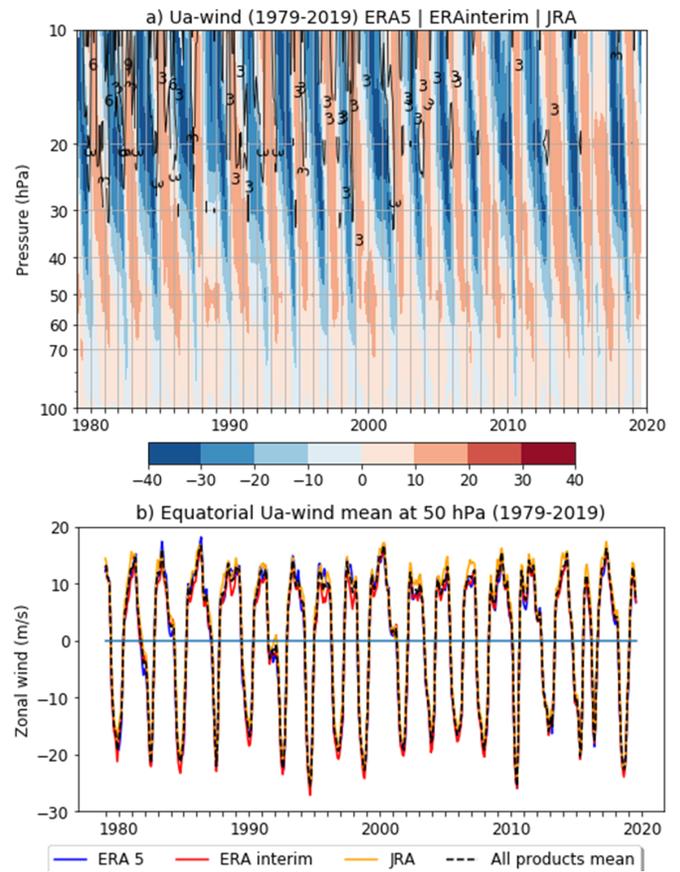


Figure 4. a) Monthly evolution of the zonally averaged zonal wind at the Equator as a function of altitude in the merged reanalysis dataset combining ERA5, ERAinterim and JRA-55 during the period 1979-2019 (shading). Black contours represent the standard deviation across the reanalyses. (b) Monthly evolution of the zonally averaged zonal wind at the Equator and at 50 hPa in the merged and individual reanalyses.

and ice-sheet melting on sea level change, which is especially relevant to predict SSHA in the polar regions. The predictive skill over these regions is therefore expected to be poor. Because, in a global sense, sea level changes are primarily controlled by variations in the external radiative forcing and by the meltwater fluxes (which are misrepresented in the model) we have decided to remove the global average of SSHA both from the model and the observations, to thus focus on the regional variations with respect to the ocean geoid, variations that are expected to be mostly controlled by internal variability processes.

ACC values for SSHA in the first year of prediction are large and significant, especially in the Pacific Ocean and the Indian Ocean, where ENSO-driven changes in the winds, precipitation and the atmospheric circulation provide high levels of predictability. Large significant ACC values are also seen in the Eurasian and Alaskan sectors of the Arctic and in some areas of the Southern Ocean. The Atlantic has comparatively lower levels of

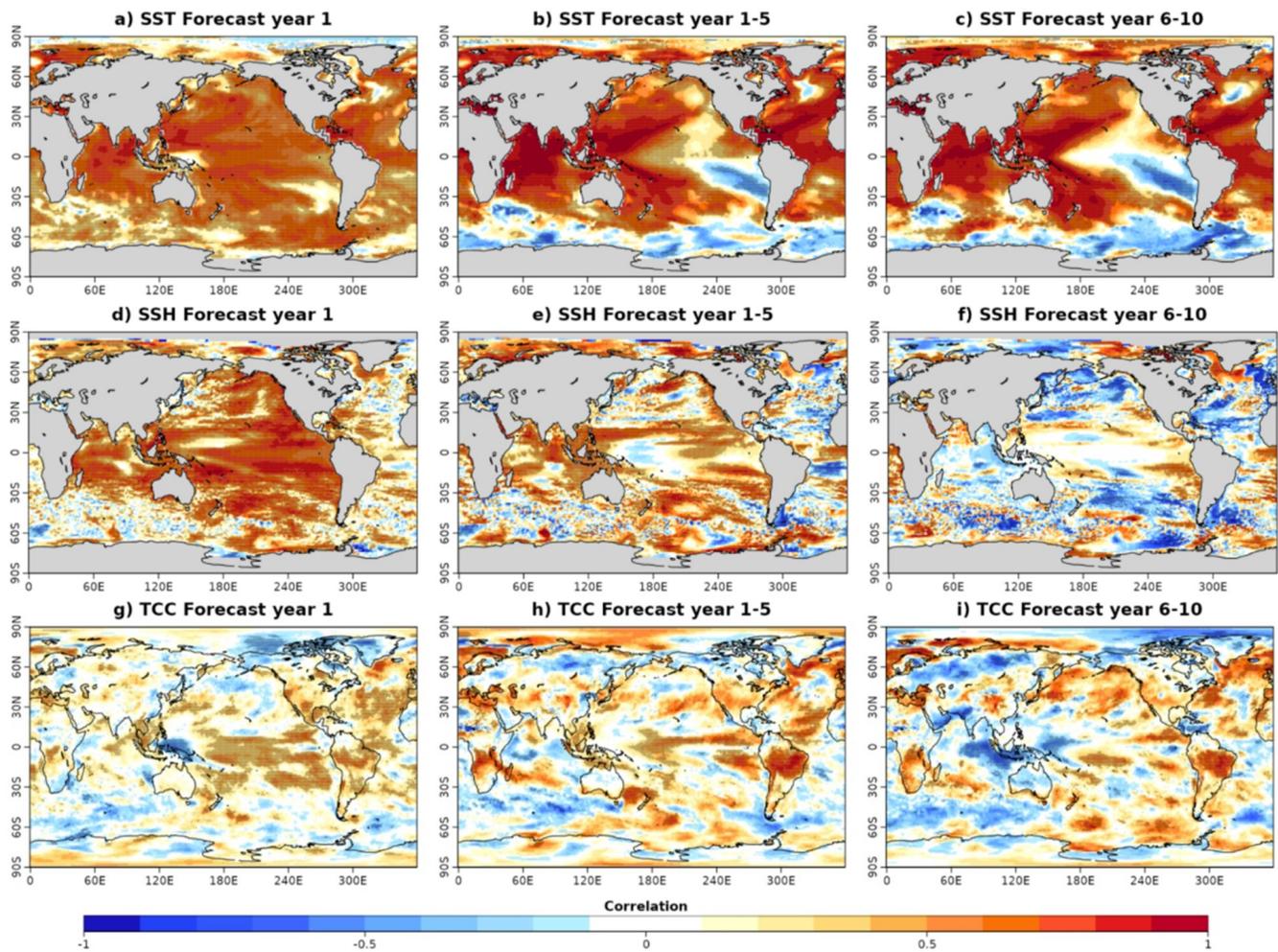


Figure 5. Forecast skill maps (as characterised by the anomaly correlation coefficient) for the first year of prediction, the average between the first and fifth years of prediction and the average between the sixth and the tenth years of prediction, for the variables: (a-c) SST; (d-f) SSH, and (g-h) TCC. Statistically significant values at the 95% confidence level are indicated with dots.

skill, with larger values confined to the Labrador region and the western side of the basin. During the first 5 forecast years, there is a significant reduction in forecast skill, especially in the Pacific Ocean and in the Indian Ocean, which could be related with the previously noticed problems in predicting ENSO. High levels of skill are still seen in the Arctic and the Southern Ocean. From the 6th to the 10th forecast year, most of the skill in SSH is lost. Only some isolated areas maintain significant ACC values in the Labrador Sea, the South Atlantic Ocean, and the southeast Pacific Ocean. The general lack of skill in the North Atlantic suggests that the AMV, at least in EC-Earth, is not a source of sea level predictability.

**TOTAL CLOUD COVER:** TCC is a key variable for climate modeling and prediction due to its effect on the radiative forcing. However, models tend to differ in the representation of its related feedbacks and climate impacts, which has promoted the implementation of a dedicated MIP, the Cloud Feedback Model Intercomparison

Project (CFMIP; *Webb et al. (2017)*) to bring light on their final contribution to climate change and variability (and thus by extension to climate predictability). In the following we evaluate the skill to predict it for the decadal predictions with EC-Earth.

The first forecast year presents significant ACC values in the ENSO region and some of its areas of influence, like the South American continent and Indonesia, and also in the North Atlantic, from which skill expands into the mediterranean region. From the 1st to the 5th forecast year, high levels of predictability are maintained over the previous regions, and in particular in the Tropical Pacific, the Amazon basin, the Subpolar North Atlantic and the Tasman Sea. The fact that both the Eastern Tropical Pacific and its areas of influences have high significant skill suggests that not all aspects of ENSO are misrepresented (and therefore badly predicted) by the model. This could imply that some ENSO flavours (either the central or eastern Niños and Niñas) might indeed be re-

alistically simulated. From the 6th to the 10th forecast year, there are still some areas with large significant correlations, namely the Amazon basin and the Subpolar North Atlantic, although other regions that previously exhibit positive skill now show very poor correlation values, the most prominent case being Indonesia.

**QUASI-BIENNIAL OSCILLATION:** During the first forecast months, the predictive skill of the QBO is very high (close to 1 ACC), and shows a progressive linear skill decrease during the first 6 years, remaining significant at the 95% confidence level during the first 3. After the sixth forecast year all information from initialization is lost, the QBO predictive skill shows marked oscillations from high to low correlation values. This is a result of an underestimation of the periodicity of the QBO in EC-Earth model with respect to observations. In the reanalysis, we have seen that a complete cycle of QBO takes 28 months (about 14 months per phase), whereas in the climate model a full cycle lasts 24 months (not shown). As in both cases, the QBO oscillations are very sharp, the difference in periodicity introduces lags between them that generates constructive or destructive interferences when correlating them at different forecast times.

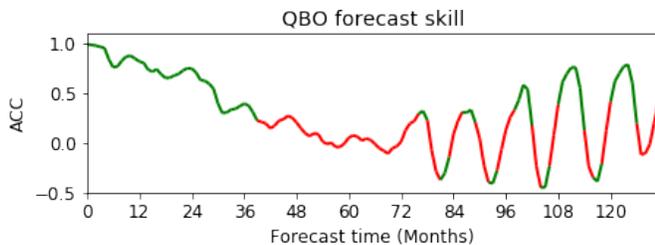


Figure 6. Anomaly Correlation Coefficient for the QBO index as a function of the forecast month. Statistically significant values at the 95% confidence level are colored in green, whereas non-significant in red.

The significant correlation values after the 6th year you should be interpreted with caution, as they result from an spurious constructive interference with no physical basis.

## V. CONCLUSIONS

This study has aimed to produce a robust evaluation of the model predictive skill (of the decadal hindcasts), focused on climate variables with high societal relevance for which several long observational products with good spatial and temporal coverage are available to be used as reference for validation. Those are sea surface temperature, sea level and the total cloud cover. The representation (and predictability) of the Quasi-biennial Oscillation has also been assessed.

In a first step we have shown a generally high degree of agreement between the different observational products available for each variable. All the products have shown an overall correlation of 0.7 (or higher) and very similar

temporal evolution across two regions of interest selected for their contribution to interannual climate variability: the North Atlantic Ocean and the Tropical Pacific.

From the three chosen variables we have identified the regions that have better (and significant) predictive skills as a function of forecast time. As it was expected, from all the variables analysed, SST has the highest forecast skill because it is directly assimilated when producing the initial conditions from which the predictions are started. During the first forecast year, there are large and significant ACC values for SST in almost all oceans. However, when taking an average of the first 5 forecast years, there is a reduction of the predictive skill on the southeastern Tropical Pacific. This reduction of the predictive skill during the first forecast years has also been seen for the SSHA, which suggests that the predictability related to ENSO phase is lost very quickly after initialization, potentially linked to deficiencies in the representation of ENSO, e.g. its periodicity. Interestingly, for TCC the Eastern Tropical Pacific and its neighboring areas of influence show high significant skill for all forecast years, which could indicate that not all aspects of ENSO are misrepresented by the model. The North Atlantic has shown high level of skill for SST and TCC, but a general lack of skill for SSH, may imply that the AMV, at least in EC-Earth, is not a source of sea level predictability. It might also be related to an initialization shock affecting the Labrador Sea that has been reported in *Bilbao et al.* (2021). Due to its strong periodicity, the QBO is highly and significantly predictable in the EC-Earth forecasts in the first three years after initialization, a very encouraging result given the important driving role of the QBO on the weather at planetary scales. However, we have also seen that small differences in the simulated periodicity with respect to observations eventually lead to a loss of skill, and create an artificial pattern of alternating periods of high and low predictability.

Throughout this work, we have shown that despite the fact that all the variables studied are influenced and interconnected with each other, not all of them agree in terms of predictive skill. Each of them provides vital information to comprehend the characteristics of the atmosphere and the oceans, which is essential for climate research. It is therefore very important to extend the analysis to other variables, and also to prediction systems based on other climate models with different errors. Identifying and understanding is indeed a very necessary step to eventually correct them (*Marvel et al.* 2015).

## ACKNOWLEDGMENTS

I would like to thank all the people that contributed to this project, without whom this would not have been possible. In particular, the climate prediction group from BSC, who have treated me as one of the team, and have provided all the help, opportunities and tools needed. Starting from Pablo Ortega the most senior supervisor of this paper, for his guidance, dedication and predispo-

sition at all times to solve doubts and for being a great inspiration. Also to Roberto Bilbao for his valuable suggestions, and help with data handling issues. And to Froila Palmeiro, the tutor of this thesis, for her expertise and thought-provoking discussion. Finally, I am deeply appreciative of my parents and sister, for always being the most constant supporters.

REFERENCES

Ahrens, C. D., *Meteorology today : an introduction to weather, climate, and the environment*, 2009.

Baldwin, M. P., et al., The Quasi-Biennial Oscillation, *Rev Geophys* 39:179–229, p. 51, 2001.

Bilbao, R., et al., Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth, *Earth System Dynamics*, 12(1), 173–196, doi:10.5194/ESD-12-173-2021, 2021.

Boer, G. J., et al., The Decadal Climate Prediction Project (DCPP) contribution to CMIP6, *Geoscientific Model Development*, 9(10), 3751–3777, doi:10.5194/GMD-9-3751-2016, 2016.

Bosello, F., R. Roson, and R. S. J. Tol, Economy-wide Estimates of the Implications of Climate Change: Sea Level Rise, *Environmental and Resource Economics* 2006 37:3, 37(3), 549–571, doi: 10.1007/S10640-006-9048-5, 2007.

Cazenave, A., et al., Observational Requirements for Long-Term Monitoring of the Global Mean Sea Level and Its Components Over the Altimetry Era, *Frontiers in Marine Science*, 0(SEP), 582, doi: 10.3389/FMARS.2019.00582, 2019.

Collins, M., et al., Interannual to Decadal Climate Predictability in the North Atlantic: A Multimodel-Ensemble Study, *Journal of Climate*, 19(7), 1195–1203, doi:10.1175/JCLI3654.1, 2006.

Costa-Surós, M., J. Calbó, J. A. González, and J. Martin-Vide, Behavior of cloud base height from ceilometer measurements, *Atmospheric Research*, 127, 64–76, doi:10.1016/j.atmosres.2013.02.005, 2013.

Dee, D. P., et al., The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597, doi: 10.1002/QJ.828, 2011.

Emery, W. J., Air Sea Interactions: Sea Surface Temperature, in *Encyclopedia of Atmospheric Sciences: Second Edition*, pp. 136–143, Elsevier Inc., doi:10.1016/B978-0-12-382225-3.00065-7, 2015.

Enfield, D. B., A. M. Mestas-Núñez, and P. J. Trimble, The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S., *Geophysical Research Letters*, 28(10), 2077–2080, doi:10.1029/2000GL012745, 2001.

Enriquez-Alonso, A., J. Calbó, A. Sanchez-Lorenzo, and E. Tan, Discrepancies in the Climatology and Trends of Cloud Cover in Global and Regional Climate Models for the Mediterranean Region, *Journal of Geophysical Research: Atmospheres*, 122(21), 11,664–11,677, doi:10.1002/2017JD027147, 2017.

Gregory, J. M., et al., Concepts and Terminology for Sea Level: Mean, Variability and Change, Both Local and Global, *Surveys in Geophysics* 2019 40:6, 40(6), 1251–1289, doi:10.1007/S10712-019-09525-Z, 2019.

Hersbach, H., et al., The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049, doi: 10.1002/QJ.3803, 2020.

Huang, B., et al., Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, *Journal of Climate*, 30(20), 8179–8205, doi:10.1175/JCLI-D-16-0836.1, 2017.

IPCC, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 151 pp., 2014.

Karlsson, K. G., et al., CLARA-A2: The second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data, *Atmospheric Chemistry and Physics*, 17(9), 5809–5828, doi: 10.5194/ACP-17-5809-2017, 2017.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann, A signature of persistent natural thermohaline circulation cycles in

observed climate, *Geophysical Research Letters*, 32(20), 1–4, doi: 10.1029/2005GL024233, 2005.

Kobayashi, S., et al., The JRA-55 Reanalysis: General Specifications and Basic Characteristics, *Journal of the Meteorological Society of Japan. Ser. II*, 93(1), 5–48, doi:10.2151/JMSJ.2015-001, 2015.

Luo, J.-J., S. Masson, S. K. Behera, and T. Yamagata, Extended ENSO Predictions Using a Fully Coupled Ocean–Atmosphere Model, *Journal of Climate*, 21(1), 84–93, doi:10.1175/2007JCLI1412.1, 2008.

Marvel, K., M. Zelinka, S. A. Klein, C. Bonfils, P. Caldwell, C. Doutriaux, B. D. Santer, and K. E. Taylor, External Influences on Modeled and Observed Cloud Trends, *Journal of Climate*, 28(12), 4820–4840, doi:10.1175/JCLI-D-14-00734.1, 2015.

Matthews, J. B. R., Comparing historical and modern methods of sea surface temperature measurement – Part 1: Review of methods, field comparisons and dataset adjustments, *Ocean Science*, 9(4), 683–694, doi:10.5194/os-9-683-2013, 2013.

McCarthy, G. D., I. D. Haigh, J. J. Hirschi, J. P. Grist, and D. A. Smeed, Ocean impact on decadal Atlantic climate variability revealed by sea-level observations, *Nature*, 521(7553), 508–510, doi: 10.1038/NATURE14491, 2015.

Meehl, G. A., et al., Decadal climate prediction an update from the trenches, *Bulletin of the American Meteorological Society*, 95(2), 243–267, doi:10.1175/BAMS-D-12-00241.1, 2014.

Meehl, G. A., et al., Initialized Earth System prediction from subseasonal to decadal timescales, *Nature Reviews Earth & Environment*, 0123456789, doi:10.1038/s43017-021-00155-x, 2021.

Merchant, C. J., and O. Embury, Adjusting for Desert-Dust-Related Biases in a Climate Data Record of Sea Surface Temperature, *Remote Sensing* 2020, Vol. 12, Page 2554, 12(16), 2554, doi: 10.3390/RS12162554, 2020.

Palmeiro, F. M., J. García-Serrano, O. Bellprat, P. A. Bretonnière, and F. J. Doblas-Reyes, Boreal winter stratospheric variability in EC-EARTH: High-Top versus Low-Top, *Climate Dynamics*, 54(5-6), 3135–3150, doi:10.1007/s00382-020-05162-0, 2020.

Ramanathan, V., R. D. Cess, E. F. Harrison, P. Minnis, B. R. Barkstrom, E. Ahmad, and D. Hartmann, Cloud-radiative forcing and climate: Results from the earth radiation budget experiment, *Science*, 243(4887), 57–63, doi:10.1126/science.243.4887.57, 1989.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *Journal of Geophysical Research: Atmospheres*, 108(14), doi:10.1029/2002JD002670, 2003.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Avery, M. Tignor, and H. Miller, *Climate Change 2007: The Physical Science Basis. Working Group I Contribution to the Fourth Assessment Report of the IPCC*, vol. 1, 2007.

Stengel, M., et al., Cloud\_cci Advanced Very High Resolution Radiometer post meridiem (AVHRR-PM) dataset version 3: 35-year climatology of global cloud and radiation properties, *Earth System Science Data*, 12(1), 41–60, doi:10.5194/ESSD-12-41-2020, 2020.

Stocker, T. F., et al., *Technical summary*, pp. 33–115, Cambridge University Press, Cambridge, UK, doi:10.1017/CBO9781107415324.005, 2013.

Sutton, R. T., G. D. McCarthy, J. Robson, B. Sinha, A. T. Archibald, and L. J. Gray, Atlantic Multidecadal Variability and the U.K. ACSIS Program, *Bulletin of the American Meteorological Society*, 99(2), 415–425, doi:10.1175/BAMS-D-16-0266.1, 2018.

Taburet, G., A. Sanchez-Roman, M. Ballarotta, M. I. Pujol, J. F. Legéais, F. Fournier, Y. Faugere, and G. Dibarboure, DUACS DT2018: 25 years of reprocessed sea level altimetry products, *Ocean Science*, 15(5), 1207–1224, doi:10.5194/OS-15-1207-2019, 2019.

Tapakis, R., and A. G. Charalambides, Equipment and methodologies for cloud detection and classification: A review, doi: 10.1016/j.solener.2012.11.015, 2013.

Webb, M. J., et al., The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6, *Geoscientific Model Development*, 10(1), 359–384, doi:10.5194/GMD-10-359-2017, 2017.

Yeager, S. G., et al., Predicting near-term changes in the earth system: A large ensemble of initialized decadal prediction simulations using the community earth system model, *Bulletin of the American Meteorological Society*, 99(9), 1867–1886, doi:10.1175/BAMS-D-17-0098.1, 2018.

Zhang, W. F., and Q. Zhao, The quasi-periodicity of the annual-cycle forced ENSO recharge oscillator model, *Communications in Nonlinear Science and Numerical Simulation*, 22(1-3), 472–477, doi: 10.1016/J.CNSNS.2014.08.013, 2015.

## VI. APPENDIX

Anomaly Correlation (ACC) The Anomaly Correlation Coefficient (ACC) is used to measure the degree of agreement for the same variables between the model's forecast and the observations. It compares the anomalies (differences) of forecasts and observations, with the corresponding climatological values ( $c_i$ ) at each gridpoint:

$$f'_i = f_i - c_i \quad (1)$$

$$o'_i = o_i - c_i \quad (2)$$

A standard definition of ACC is:

$$ACC = \frac{\sum_{i=1}^M (f'_i - \bar{f}') (o'_i - \bar{o}')}{\sqrt{\sum_{i=1}^M (f'_i - \bar{f}')^2 (o'_i - \bar{o}')^2}} \quad (3)$$