Seasonal cycle of sea surface water characteristics in climate models

A thesis submitted to the School of Environmental Sciences at the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Yanxin Wang

July 2021

Abstract

The seasonal cycle of sea surface water characteristics is important for the global climate system. Seasonal extrema of sea surface temperature (SST) and sea surface salinity (SSS) determine water mass properties below the surface. Evaluation of climate models typically focuses on annual or long-term mean state, not on seasonal extrema. In this thesis, the seasonal cycles of SST and SSS in HiGEM and SST seasonal extrema in 20 CMIP6 models are assessed globally.

Sparse sampling leads to large differences between observational climatologies in both SST and SSS in polar regions. There are also large SST differences in regions with strong SST horizontal gradient, likely because gridding on coarse resolution can smooth the gradient. To exclude regions with large differences between climatologies, masks are proposed for global model assessments.

The results demonstrate the importance of evaluating model performance not simply against annual mean properties. Although the biases in SST and SSS seasonal extrema are largely consistent with their annual means, the amplitude of SST and SSS biases has large seasonal variations in specific regions. Large seasonal variations of SST bias in CMIP6 models occur in eastern boundary upwelling regions, polar regions, the North Pacific and eastern equatorial Atlantic. Large seasonal variations of SSS bias in HiGEM occur in equatorial and polar regions. SST biases in some CMIP6 models have seasonal spatial patterns. Models with greater vertical resolution in the ocean typically demonstrate better representation of SST extrema, particularly seasonal maximum SST. However, no significant relationship is found with ocean model horizontal resolution.

List of Acronyms

AABW Antarctic Bottom Water.

AAIW Antarctic Intermediate Water.

ACC Antarctic Circumpolar Current.

AMIP Atmospheric Model Intercomparison Project Phase.

AMOC Atlantic meridional overturning circulation.

APB Autonomous Pinniped Bathythermograph.

CGCM coupled general circulation model.

CMIP Coupled Model Intercomparison Project.

CTD conductivity-temperature-depth.

ENSO El Niño - Southern Oscillation.

GOF goodness of fit.

HighResMIP High Resolution Model Intercomparison Project.

IPCC Intergovernmental Panel on Climate Change.

ITCZ Intertropical Convergence Zone.

ITP Ice-Tethered Profile.

MBT Mechanical Bathythermograph data.

MLD mixed layer depth.

NAC North Atlantic Current.

NADW North Atlantic Deep Water.

NPIW North Pacific Intermediate Water.

OSD Ocean Station Data.

PFL Profiling float data.

RMSE root mean square error.

SPCZ South Pacific Convergence Zone.

SSD sea surface density.

SSS sea surface salinity.

SST sea surface temperature.

XBT Expendable data.

List of Symbols

T_{max} maximum SST of the monthly mean climatology

T_{min} minimum SST of the monthly mean climatology

T_{mean} annual mean SST of the monthly mean climatology

T_{cycle} range of the SST seasonal cycle (T_{max} - T_{min})

S_{max} maximum SSS of the monthly mean climatology

S_{min} minimum SSS of the monthly mean climatology

S_{cycle} range of the SSS seasonal cycle (S_{max} - S_{min})

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my friends and colleagues in COAS, for providing a supportive and enjoyable working environment. Working from home during COVID made me miss you all. Thank you to my first-year flatmates at the UEA Ziggurats, who helped me settle in the new environment and cured my loneliness. I would also like to thank my friends at 95 Friends Road and 2 Salvin Court. Thank you for your support and company during the pandemic, without you I could not go through that hard time.

A big thank you to my parents, who always had my back and gave me encouragement. Thank you to my partner, who always cheered me up when I felt down and forced me to start regular running to cope with stress. The pandemic taught me how much you all to me.

Finally, the most thanks to my supervisors - Karen, Dave and Gillian, for meeting with me every week and giving me constructive suggestions, commenting on my paper draft over and over again, and carefully reading my thesis. You were always willing and enthusiastic to help me. Without you, I would not have succeeded in finishing this thesis.

1

Introduction

1.1   Climate and climate modelling

Climate has a profound influence on human and it is important for water, agriculture, energy, human health and ecosystems. However, since the beginning of the Industrial Revolution, human activities, particular those producing greenhouse gases, have begun to have a large impact on global climate. Excessive amounts of greenhouse gases are released into the atmosphere and the heat from the warming atmosphere can be transferred into the ocean. Ocean warming can cause extreme weather which impacts people living in coastal areas. As sea surface temperature (SST) is related to ocean heat content, it is an essential variable for quantifying climate change.

Climate modelling is a key tool for understanding and predicting the response of climate system to human-induced forcing. Many climate system models (CSMs) have been extended into earth system models. Earth system models include physical processes in other climate models, but also include the biogeochemical processes. Representing the interaction of biogeochemical processes and physical climate alters the response of physical climate to forcing such as that associated with anthropogenic emissions of greenhouse gases (Flato, 2011).

As running the most comprehensive and highest resolution models requires high computational costs, model family, which contains a range of models, is developed for different applications (Pope et al., 2007). A model family shares a common physical framework, but has different complexity and resolution. For example, as models in the HadGEM family, HadGEM3-GC31-MM was running at higher resolution

(Andrews et al., 2020); E3SM-1-0 (an earth system model) was running at lower resolution with increased complexity (Golaz et al., 2019).

The quality of model simulations is assessed by comparing simulated climate with observations, and then analyse the differences. If climate models are able to simulate past climate variables, they are more likely to provide good forecasts of the future. Historical runs from around 1850 to near-present allow scientists to compare model predictions of the past climate to recorded observations (Eyring et al., 2016). Therefore, it is crucial to evaluate the performance of climate models using historical runs.

1.2 Sea surface water characteristics

The ocean and atmosphere interact at the sea surface. Sea surface water characteristics reflect the coupling of ocean-atmosphere and they are important variables to better understand interactions between the ocean and atmosphere. SST and sea surface salinity (SSS) are two important physical characteristics of sea surface water.

SST exerts a major influence on the exchange of energy between ocean and atmosphere. Longwave radiation, sensible and latent heat fluxes all depend on SST. Air-sea heat fluxes are important components of the climate system, which enable energy exchange between ocean and atmosphere. Much of the small-scale variability of the wind is attributable to SST: cool SST stabilizes marine atmospheric boundary layer and decouples the surface winds from winds aloft, thus decreases the wind speed; warm SST destabilizes the boundary layer and decrease the vertical shear of the wind, thus increases the wind speed (Chelton et al., 2004). SST gradient therefore influences the wind stress divergence and curl (Chelton et al., 2004). In the tropical oceans, when SST increases, atmospheric deep convection tends to occur more frequently and with larger intensity (Bjerknes, 1966). This influences the development of tropical cyclones (hurricanes and typhoons), which draw energy from warm ocean waters to the atmosphere. SST seasonal maximum in the tropical region determines formation and intensity of tropical cyclones (Evans, 1993; Tory and Dare, 2015).

There exists a SST threshold, below which tropical cyclones do not form. Palmen (1948) suggested that a SST threshold of about 26°-27°C may be necessary for tropical cyclone formation, while Dare and McBride (2011) concluded that the SST threshold was 25.5°C.

SST has strong controls on precipitation. Warmer SST increases atmospheric moisture, which favours precipitation. Roxy (2014) quantified the SST-precipitation relationship in the tropical oceans and indicated that precipitation increases at 2 mm/day for an 1°C increase of SST. SST can be used to monitor the tropical Pacific. Niño SST indices are commonly used to define El Niño and La El Niña events (Trenberth and Stepaniak, 2001). SST along the equator in the Pacific Ocean gets warmer during El Niño, which has global impact. When El Niño occurs, there are heavy rainfall in the southern United States and severe drought in Australia, Indonesia and Southern Asia. Changes in SST can shift storm tracks, contributing to precipitation in some areas. Brayshaw et al. (2011) found that SST fronts related to Gulf Stream and North Atlantic Drift impact the midlatitude storm track and enhance precipitation on the warm side of SST front.

SST is also important in chemical and biological oceanography. The variation of SST can threaten sensitive ocean life such as coral. Temperature is fundamental to determining coral health and survival, and temperature anomalies can cause coral bleaching. SST seasonal maximum coincides with coral bleaching event both in onset and duration (Brown et al., 1996). Therefore, SST is used to predict coral bleaching and it also has important application for reef restoration (Liu et al., 2006; Foo and Asner, 2020). Above average SST caused large-scale coral bleaching in nearly every major coral reef ecosystem (Maynard et al., 2008). SST also changes the frequency and intensity of harmful algal bloom such as "red tide". Higher SST under climate warming will be a favourable condition for algal blooms and probably lead to an intensification of algal blooms (Lürling et al., 2013).

Both atmospheric and oceanic processes govern SST. On the atmospheric side, wind speed, air temperature, cloudiness and humidity are influencing factors of the energy exchange at sea surface; on the oceanic side, currents, vertical mixing and mixed layer

depth influence heat transport and thus SST. Following Deser et al. (2010), the equation for the mixed layer layer temperature (equal to SST) is written as

$$\partial T/\partial t = Q_{net}/(\rho C_p H) + (\vec{V}_{geo} + \vec{V}_{ek}) \cdot \vec{\nabla} T + (W_e + W_{ek})(T - T_b)/H, \qquad (1.1)$$

where T is mixed layer temperature or SST, Q_{net} is net surface energy flux, ρ is density of seawater, C_p is specific heat of seawater, V_{geo} is geostrophic current velocity, V_{ek} is Ekman current velocity, W_e is vertical entrainment rate, W_{ek} is Ekman pumping velocity, and T_b is the temperature of the water at depth that is entrained into the mixed layer. Q_{net} is defined as

$$Q_{net} = Q_{sw} + Q_{lw} + Q_{sh} + Q_{lh}, \qquad (1.2)$$

where Q_{sw} is downward solar radiative flux, Q_{lw} is longwave radiative flux, Q_{sh} is sensible heat flux, Q_{lh} is latent heat flux. The radiative fluxes ($Q_{sw}+Q_{lw}$) are functions of air, temperature, humidity and cloudiness; the turbulent energy flux ($Q_{sh} + Q_{lh}$) is linearly proportional to the wind speed and difference of air-sea temperature or humidity.

Solar radiation passes through atmosphere and clouds, reaching the ocean surface. The incoming solar radiation is partially absorbed and reflected by clouds and water vapor, and thus only part of the solar radiation reaches the ocean and is converted into heat energy (Q_{sw}). Q_{sw} is the largest term in the four heat flux terms. The path of the sunlight through the atmosphere and the length of daylight combined determines the seasonality of solar radiation. Therefore, Q_{sw} is largest in summer and smallest in winter; the higher the latitude, the larger the seasonal variation in Q_{sw}. Q_{lw} is the electromagnetic energy that is radiated outward by the ocean. Q_{lw} mainly depends on SST (the warmer the SST, the larger the Q_{lw}). Q_{sh} is the heat conduction between the sea surface and atmosphere due to the air-sea temperature difference. When air is cooler than sea surface, heat will be conducted away from the sea, resulting in a negative Q_{sh}; when air is warmer than sea surface, heat will be conducted into the sea,

resulting in a positive Q_{sh}. Q_{lh} is usually the second largest of the four heat flux terms and it is associated with evaporation. Wind speed, SST and the humidity in the air determine evaporation rate and thus Q_{lh} (Talley, 2011).

Sea surface salinity (SSS) plays an important role in ocean stratification and influences the processes controlling mixed layer temperature through changes in mixed layer depth (MLD). Near-surface freshening in the tropics contributes to the formation of barrier layer, and barrier layers limit the mixing of cold thermocline waters into the near surface layer as the mixed layer is shallower than isothermal layer (Sprintall and Tomczak, 1992; Scannell and McPhaden, 2018). By modulating SST, SSS indirectly affects air-sea interaction.

At low-mid latitudes, SSS can be a proxy for the impact of the hydrologic cycle, or the flux of freshwater across the air-sea interface (Bingham et al., 2012). At high latitudes, SSS can be a proxy for the sea ice melting-freezing cycle (Garcia-Eidell et al., 2019). High SSS indicates strong evaporation or brine rejection process, while low SSS is linked to strong precipitation or sea ice melt.

SSS is governed by evaporation, precipitation, sea-ice melting and formation, and salinity transport by currents, vertical mixing and mixed layer depth. Similar to the heat budget equation 1.1, the equation for the mixed layer salinity (equal to SSS) can be written as

$$\partial S/\partial t = S(E + F - P - M)/H + (\vec{V}_{geo} + \vec{V}_{ek}) \cdot \vec{\nabla} S + (W_e + W_{ek})(S - S_b)/H, \quad (1.3)$$

where S is mixed layer salinity or SSS, E is evaporation, F is sea ice formation, P is precipitation, M is sea ice melting, S_b is the salinity of the water at depth that is entrained into the mixed layer.

Evaporation and precipitation are the two main components of hydrological cycle. Water moves from the ocean surface to the atmosphere in form of evaporation and back in form of precipitation. Evaporation rate depends on wind speed, SST and the

humidity in the air. Subtropical regions are dominated by high evaporation and have a high SSS. Evaporation is lower over the equator than the subtropical regions because of the combination of higher moisture and lower wind speed, even though the SST is warmer. Precipitation occurs when the air becomes saturated with water vapor, and the heaviest precipitation occurs within the tropical regions, especially along ITCZ and SPCZ.

Sea ice formation and melting are the main processes controlling the SSS seasonal cycle in the polar regions. Moreover, the exchange of salt between sea ice and ocean can lead to changes of sea water density. When sea ice forms, brine rejection occurs and SSS increases, and therefore density of surface water increases and the water sinks. When sea ice melts, freshwater enters the sea and SSS decreases. Sévellec et al. (2017) connected the Arctic sea-ice decline with AMOC slow-down, as the freshwater fluxes due to sea ice melting can result in less dense water and weaken the circulation.

Sea water properties within the mixed layer is well mixed vertically. Temperature and salinity are vertically uniform within the mixed layer. MLD varies seasonally. It can be less than 20 m in the summer hemisphere, while reaching over 500 m in the winter hemisphere in the subpolar regions (de Boyer Montégut et al., 2004). Vertical mixing can deepen the mixed layer and entrain water with different temperature/salinity from below. Vertical mixing occurs due to buoyancy loss and/or wind. Cooling or evaporation at the sea surface can deepen the mixed layer to over several hundred meters, or even over 1000 m in deep convection locations in winter (Talley, 2011). However, wind-stirred mixing cannot extend that deep. It cannot be deeper than 100 or 150 m and it can only reach this depth in winter (Talley, 2011). Other than the vertical entrainment, horizontal advection associated with geostrophic current and Ekman transport can also cause SST/SSS change by bringing waters with different temperature/salinity from other regions.

As temperature and salinity determine density, seasonal extrema of SST and SSS are important for water mass formation and have a direct effect on ocean circulation. Fresh and warm waters with low density will remain at the surface. Salty and cold surface waters with large density can lead to strong vertical mixing, creating dense deep

and intermediate waters (Fig. 1.1). Therefore, seasonal minimum SST and seasonal maximum SSS are especially important in areas where dense water forms, such as the Arctic. The Gulf Stream brings high-salinity waters northwards into the high latitudes, where they cool and sink, forming dense water masses and push the thermohaline ocean circulation. The evolution of surface water masses to intermediate or deep water masses is important for the global redistribution of heat and salt. It also provides a pathway for the transport of anthropogenic carbon dioxide to greater depths and slows down climate warming (Bopp et al., 2015).
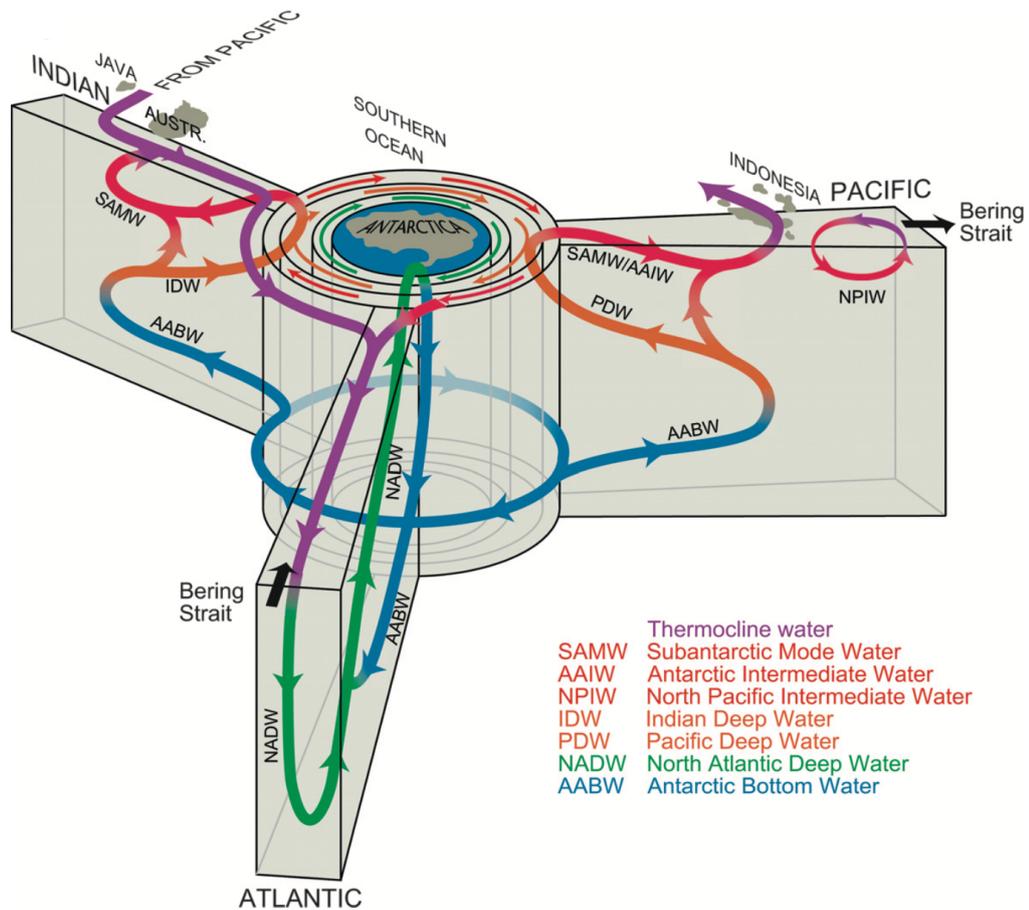
Figure 1.1: Schematic of intermediate water and deep water formation, adapted from Talley (2011).

Realistic model simulation of seasonal extrema of SST and SSS is important for prediction of intermediate and deep waters and hence large scale ocean circulation. Seasonal maximum SST is also important for model prediction of tropical cyclones and coral bleaching. Therefore, a good representation of seasonal extrema of SST and SSS is an essential aspect of model skill for future climate projections.

1.3 The Coupled Model Intercomparison Project

The Coupled Model Intercomparison Project (CMIP) is an international collaboration providing a multi-model context for climate simulations. The objective of CMIP is to better understand past, present and future climate changes arising from natural, unforced variability or as a response to changes in radiative forcing. CMIP promotes a standard set of model simulations in order to evaluate how realistic the models are in simulating the recent past and provide projections of future climate change on near term (out to about 2035) and long term (out to 2100 and beyond). The results of CMIP model runs are used extensively in the Intergovernmental Panel on Climate Change (IPCC) reports (Solomon et al., 2007; Stocker, 2014), which provides policy makers with scientific assessments of climate change.

U.K.'s High-Resolution Global Environmental Model (HiGEM) was developed by the Natural Environment Research Council, the Met Office and the academic community in UK. HiGEM is based on HadGEM1 and it is a CMIP5 model (Shaffrey et al., 2009). The horizontal resolution is $0.83°$ latitude $\times$ $1.25°$ longitude for the atmosphere, and $1/3° \times 1/3°$ globally for the ocean and sea ice. In the vertical, the atmosphere model has 38 levels with a top at 39 km; the ocean model has 40 unevenly spaced levels, which is about 10 m resolution close to the surface, and increases gradually to about 300 m at depth. The ocean component is formulated on a spherical latitude-longitude grid, with a singularity at the North Pole.

CMIP6 (CMIP Phase 6) is the most current of the CMIPs. CMIP6 models with different characteristics allow investigation of the factors related to differences in model performance. Models used in this thesis vary in ocean grid type, ocean vertical coordinate, ocean horizontal and vertical resolution, atmosphere horizontal and vertical resolution, Earth system or not (more details can be seen in table. 4.2).

The desire for better climate modelling motivated an increase of model resolution. Higher model resolution generally leads to mathematically more accurate models (Flato et al., 2013). In any model, processes which occur over too small a scale to be resolved must be parameterised. For example, the CMIP6 models with higher

horizontal resolution (1/3° or 1/4°) are eddy permitting, whereas in the lower resolution models (1°) eddies are parameterised (Flato et al., 2013). Previous studies have mostly emphasised the benefits of increasing the horizontal resolution. The representation of boundary currents, ocean fronts, eddies and air-sea fluxes can be significantly improved as resolution increases (Hewitt et al., 2017; Kirtman et al., 2012; Roberts et al., 2016). Therefore, as horizontal resolution is increased, pronounced SST bias reduction occurs in the Southern Ocean, in the Agulhas retroflection region, and along the Gulf Stream extension in the North Atlantic (Skákala et al., 2019; de la Vara et al., 2020; Chassignet et al., 2020). However, higher model resolution does not necessarily lead to more reliable simulations (Flato et al., 2013). Richter and Tokinaga (2020) showed that the CMIP6 models with the smallest SST bias in the tropical Atlantic all have relatively high horizontal resolution, but there are also models with low-resolution that perform well. Chassignet et al. (2020) used four pairs of matched low-resolution and high-resolution ocean simulations from global ocean-sea-ice models to isolate the effect of ocean horizontal resolution. They found that increased horizontal resolution does not improve SST bias unambiguous in all regions for all models.

The models used here have four different horizontal grid type: regular latitude-longitude, tripolar, displaced pole and unstructured meshes. Regular latitude-longitude grid lines converging towards the pole can be a source of numerical difficulties in ocean modelling. To leave a smooth, singularity-free grid in the Arctic, a displaced pole grid (Jones et al., 2005) has the North Pole displaced over Greenland, and a tripolar horizontal grid (Murray, 1996) has poles over Eurasia, North America and Antarctica. By using an unstructured-mesh, it is possible to put a focus on dynamically active regions such as the North Atlantic Current (NAC), the Southern Ocean and the tropics while using relatively coarse resolution elsewhere (Semmler et al., 2020).

Compared with ocean horizontal resolution, ocean vertical resolution has drawn much less attention. Xavier et al. (2008) revealed that coarse vertical resolution of oceanic general circulation models limits their ability to represent intraseasonal processes, such as the formation of diurnal warm layer. Ge et al. (2017) conducted numerical

experiments using an oceanic general circulation model with 1 m and 10 m vertical resolution for the upper ocean and found that 1 m vertical resolution had a more realistic representation of the vertical temperature structure in the upper ocean.

For ocean vertical coordinate, most of the models considered here use z-level or some variation of it (z^*-level), but there are also z-isopycnal and sigma level. All layers of traditional z-level models have fixed thickness, and layers of z^*-level models are rescaled for more accurate representation of free-surface variations. Z-level models are prone to high spurious diapycnal mixing, which is related to spurious heat uptake and modified water masses (Griffies et al., 2000; Willebrand et al., 2001; Legg et al., 2006). Isopycnal models have no diapycnal mixing and produce no spurious mixing, but they can suffer from insufficient resolution in the mixed layer (Bleck, 1978). Hybrid coordinate (e.g. z-isopycnal coordinate) combine the advantages of z-level in the upper ocean for higher resolution and isopycnal coordinate in the deep ocean to reduce diapycnal mixing (Bleck, 2002; Chassignet et al., 2003; Legg et al., 2009). Sigma coordinate (σ) models have all layers contract to follow bottom bathymetry and are particularly useful for coastal modeling, but errors exist in the pressure gradient when layers have large slopes along steep topography (Shchepetkin and McWilliams, 2005).

1.4 Representation of sea surface water in climate models

1.4.1 Annual mean

Previous model evaluation mainly focuses on the simulation of ocean mean state (annual or longer-term mean SST). Flato et al. (2013) analysed the long-term mean differences of zonal SST distribution between Coupled Model Intercomparison Project Phase 5 (CMIP5) models and observations, indicating larger biases at mid to high latitudes than at other latitudes. Wang et al. (2014) showed long-term mean SST biases in CMIP5 multi-model mean (Fig. 1.2), and stated that SST biases in specific regions (e.g. the cold biases in the North Atlantic and North Pacific, and the warm bias in the Southern Ocean) are commonly linked to a too weak Atlantic meridional overturning

circulation (AMOC).
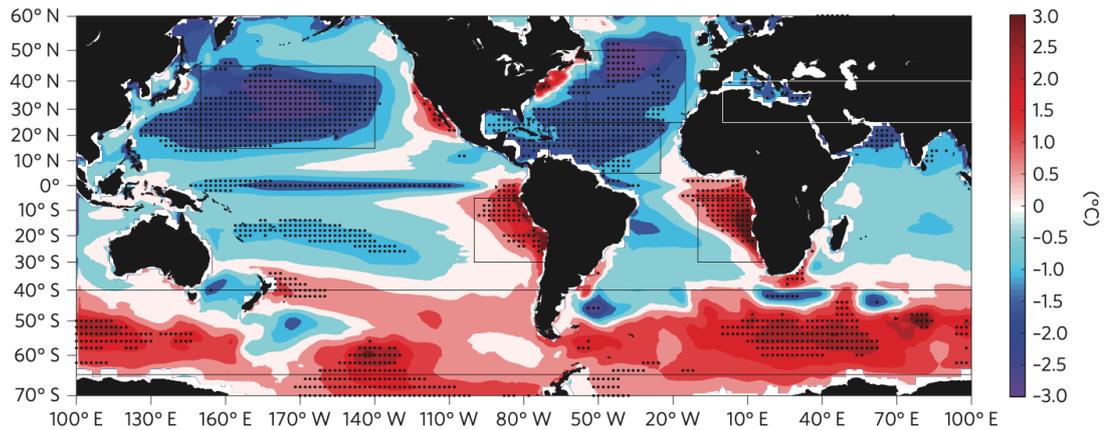
Figure 1.2: The annual mean SST bias averaged in 22 CMIP5 models. The dots denote where at least 18 of 22 models have the same sign in the SST bias. Adapted from Wang et al. (2014).

Models show a common cold bias in the North Atlantic SST and a warm bias near the North American coast, due to the poor representation of Gulf Stream and NAC in the models (Willebrand et al., 2001; Eden et al., 2004; Keeley et al., 2012). The poor representation of the NAC, which is too zonal in lower-resolution models, leads to a significant SST cold bias (up to 6°) in the northwest Atlantic (Kuhlbrodt et al., 2018), and this is common in 1° ocean models (Danabasoglu et al., 2014). This SST cold bias can be dependent on the ocean resolution, as there are significant improvements in the NAC with ocean-only simulations at higher resolutions (Storkey et al., 2018; Marzocchi et al., 2015). A finer horizontal ocean resolution has been shown to enable a more realistic SST over the North Atlantic in CMIP6 models via a better representation of the Atlantic Ocean heat transport (Docquier et al., 2019; Roberts et al., 2020).

Most coupled climate models have substantial SST warm biases in the Southern Ocean (Sallée et al., 2013; Hyder et al., 2018). These warm biases have been linked to insufficient cloud which causes excessive downward surface short-wave radiation (Bodas-Salcedo et al., 2012). Analysing CMIP5 and Atmospheric Model Intercomparison Project Phase 5 (AMIP5) in combination enabled Hyder et al. (2018) to separate the influences of atmospheric model errors and coupled feedbacks, and they were therefore able to demonstrate that in the Southern Ocean the variations of SST biases across the CMIP5 ensemble are primarily caused by the variations of net surface flux bias in AMIP5.

A pronounced SST cold bias exists in the subtropical North Pacific (Griffies et al., 2009; Burls et al., 2017; Zhu et al., 2020). Griffies et al. (2009) used ocean-only models driven by prescribed-atmospheric forcing and produced a cold bias over the subtropical regions, which indicates that ocean model processes may be significant for this cold bias. Using CMIP5 models, Burls et al. (2017) suggested that the cold SST bias is linked to cloud albedo errors, which leads to insufficient surface short-wave fluxes.

Models suffer biases in the annual mean equatorial SST (Davey et al., 2002). In the tropics, the SST biases could be classified into two types: one exhibiting broad meridional structures that are due to cloud bias, and one associated with Pacific and Atlantic cold tongue bias that are due to thermocline depth bias (Li and Xie, 2012). Models have a diversity in representing thermocline depth. A shallower thermocline facilitates the equatorial cold tongue, contributing to a cold bias (Richter and Xie, 2008; Li and Xie, 2012). Harlaß et al. (2018) showed that models with high horizontal and vertical atmospheric resolution have better simulation of the equatorial cold tongue, as these models can improve the wind structure which is associated with the thermocline depth.

From the fouth assessment report of IPCC, warm biases exist in the upwelling zones off the Peruvian, Namibian and Californian coasts in most coupled models, as the prevalent subtropical stratocumulus cloud decks are poorly represented in models (Solomon et al., 2007). Most CMIP5 models still show substantial SST warm biases in the eastern boundary upwelling systems, especially in the southeast Atlantic region (Wang et al., 2014; Richter, 2015). Richter (2015) attributed the warm SST bias of the eastern boundary upwelling regions to underestimated cloud and insufficient upwelling due to overly weak winds. The SST bias in the eastern boundary upwelling region has been shown to be sensitive to ocean horizontal resolution (de la Vara et al., 2020; Small et al., 2015).

There are only few studies that evaluate climate models in representing SSS. SSS has a fresh bias in the subpolar North Atlantic and the Arctic in the CMIP5 multi-model mean (Flato et al., 2013), which can increase stratification, hamper convection and lead to excessive sea ice. The fresh bias also has the potential to weaken AMOC through

reducing density in the sinking regions (Park et al., 2016). Sallée et al. (2013) found a fresh bias of SSS in the Southern Ocean in CMIP5 models, which may constrain the transport of anthropogenic carbon from surface to subsurface (Terhaar et al., 2021). Volodin et al. (2017) showed that INM-CM5 (a CMIP5 model) has an overall fresh bias in the global ocean except the Arctic, which has a saline bias up to 1-5. However, the considerable saline bias in the Arctic might be from the uncertainty of observational climatology they used, which will be discussed in this thesis. Fathrio et al. (2017) found that the CMIP5 multi-model mean has a saline SSS bias of 1.5 in the Bay of Bengal and a fresh SSS bias of 0.4 in the western Indian Ocean and southeastern Indian Ocean. They attributed the biases in the Bay of Bengal and western Indian Ocean to precipitation bias, while the fresh bias in southeastern Indian Ocean was attributed to salt advection bias.

1.4.2 Seasonal cycle

Another important aspect of model skill is how well it can simulate SST and SSS seasonal extrema. Accurate mean does not guarantee accurate seasonal extrema or annual cycles. Identifying the errors in simulating seasonality of water masses and their sources are important to improve our understanding of the physical processes related to SST and SSS seasonal extrema and thus improve the simulation of SST and SSS in climate models.

Only a few previous studies have assessed the performance of models in simulating the annual cycle of SST. Wang et al. (2014) concluded that the spatial patterns of SST biases are largely independent of season, but the amplitudes of SST biases vary with season in some locations (Fig. 1.3). For example, there is a warm SST bias in the Southern Ocean throughout the year, but the amplitude of the warm bias is much larger in summer and autumn than in winter and spring.

In the equatorial Pacific, 23 coupled general circulation models (CGCMs) (many are in CMIP) were evaluated by Davey et al. (2002), while only few models have a seasonal cycle of upper ocean temperature similar to that observed in real ocean. Braconnot et al. (2007) used a CGCM with various different convection and cloud schemes to
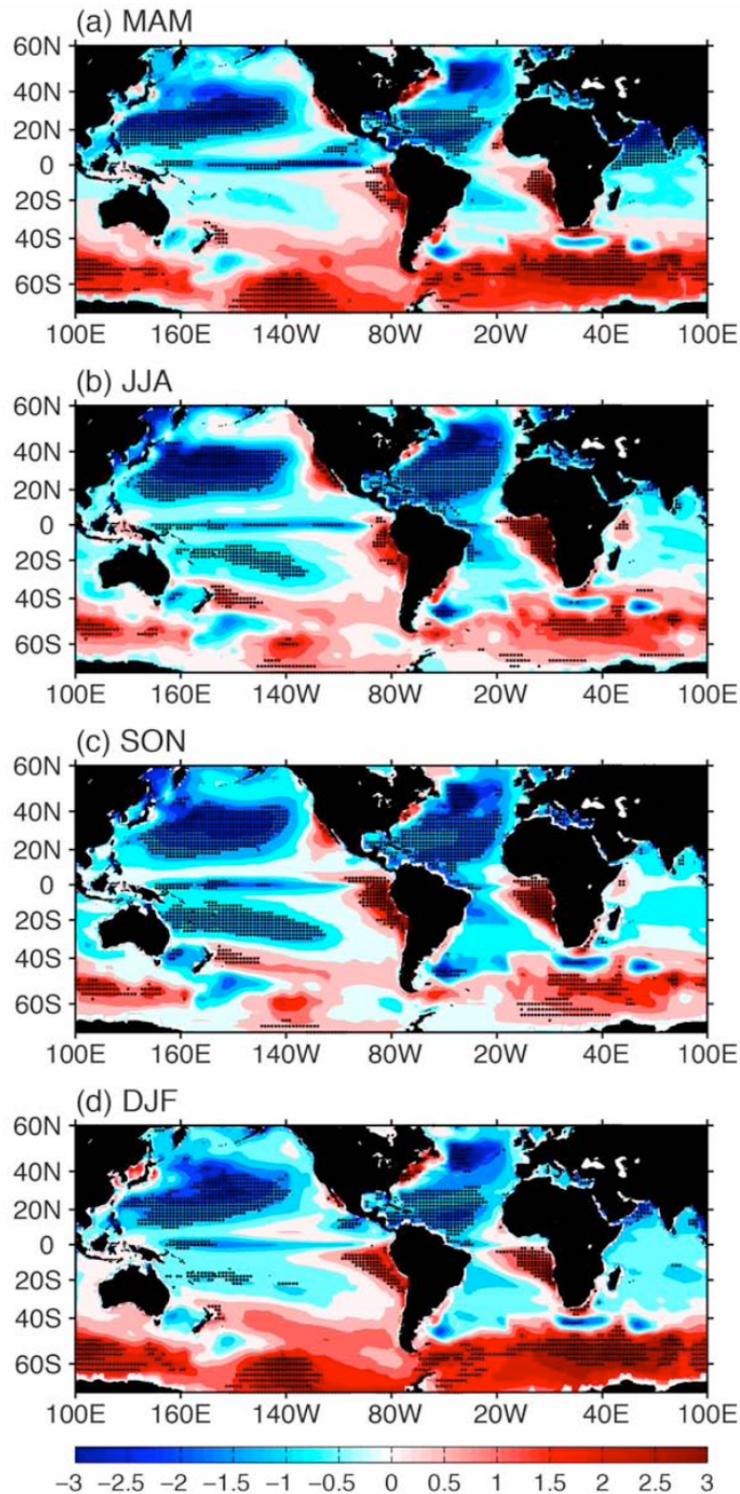
Figure 1.3: The seasonal variation of SST biases in 22 CMIP5 models, adapted from Wang et al. (2014). SST bias during (a) spring (March-April-May), (b) summer (June-July-August), (c) autumn (September-October-November) and (d) winter (December-January-February). The dots denote where at least 18 of 22 models have the same sign in the SST bias.

demonstrate that clouds and convection can affect the seasonal variations of equatorial SST. Mechoso et al. (1995) examined the seasonal cycle of SST over the equatorial Pacific and stated that the seasonal equatorial SST biases in CGCMs resulted from Intertropical Convergence Zone (ITCZ) errors, and these biases can be alleviated by improved simulation of the seasonal cycle of meridional wind (De Szoeke and Xie, 2008). AchutaRao and Sperber (2002) indicated that in 17 CMIP models, flux-corrected models tend to have a better representation of SST annual cycle in the equatorial Pacific. All the factors mentioned before could play a part in SST seasonal cycle in the equatorial Pacific.

SST warm biases in the eastern tropical Atlantic vary seasonally in CGCMs and reach a maximum during June-July-August (Prodhomme et al., 2019; Richter and Xie, 2008; Richter et al., 2014). Richter et al. (2012) used the GFDL coupled GCM to investigate those SST biases in the tropical Atlantic and concluded that a large portion of that biases is due to too weak easterlies during boreal spring. Due to the weak easterlies, the thermocline of the eastern equatorial Atlantic deepens, which prevents cold tongue formation. In the equatorial Atlantic Ocean, a CGCM: CFS (climate forecast system) can capture the annual cycle of zonal SST gradients, although with warm biases and too little variability of SST over the southeastern ocean (Hu et al., 2008).

Over the northeastern Pacific ITCZ, CMIP5 models have seasonally dependent SST biases: warm bias in summer and cold bias in winter (Song and Zhang, 2020). Song and Zhang (2020) suggested this bias is caused by the poor simulation of the North American monsoon in models. In the CMIP5 models they considered, an easterly wind bias exists all year around. When the easterly wind dominates in winter, the too strong wind in the models enhances surface evaporation and leads to a cold SST bias, while in summer, the wind becomes westerly and the easterly wind bias manifests as weaker westerly winds, with the consequent reduced evaporation leading to a warm SST bias. The easterly wind bias can be improved by increasing the atmospheric horizontal resolution Song and Zhang (2020). Based on coupled ocean-atmosphere model simulations, Liang and Wu (2013) indicated that the seasonal cycle of SST in the extratropical North Pacific models can be better represented with the help of improved representatin of solar penetration into the ocean. In summer, the SST warm bias

in models can be partly reduced when shortwave radiation penetrates further into the thermocline, while in winter, MLD deepens and the subsurface warm water is entrained into the surface, reducing the SST cold bias in model (Liang and Wu, 2013). Zhu et al. (2020) showed that SST cold biases of the North Pacific subtropics vary seasonally in CMIP6 models, and the seasonality of the upper ocean cold bias is linked to vertical diffusivity. As discussed above, better representation of the monsoon, solar penetration and vertical diffusion could improve SST simulation in the subtropical North Pacific.

In the southwestern Indian Ocean, seasonal SST forecasts are better in ocean models with upwelling and Rossby wave dynamics than ones using a slab ocean mixed layer (Xie et al., 2002). Yokoi et al. (2009) suggested that CMIP3 models underestimated the magnitude of semiannual variability in the southern tropical Indian Ocean, which was related to the overestimation of the magnitude of annual variability in thermocline depth. CMIP5 models still perform poorly in simulating the shallow climatological thermocline in the southwestern Indian Ocean, which is essential for successful seasonal predictions in the Indian Ocean (Nagura et al., 2013).

Fathrio et al. (2017) illustrated that in CMIP5 multi-model mean the seasonal variation of SSS in the tropical Indian Ocean is poorly simulated. Using fixed depth salt budget analysis, they found that the seasonal biases of precipitation and horizontal advection are mainly responsible for the seasonal SSS bias in the tropical Indian Ocean. Parekh et al. (2016) attributed the seasonal SSS bias over Bay of Bengal in CFSv1 model to poorly simulated precipitation and river runoff during summer and autumn, and the seasonal SSS bias in Arabian Sea to poorly simulated East Indian coastal current in winter and Somali jet in summer.

1.5 Overview of this thesis

It is important that climate models give accurate projections of seasonal extrema of sea surface water characteristics, because of their impact on the world's climate system through heat and freshwater fluxes at the ocean surface, and through water mass formation and subsequent impact on ocean circulation. Previous model

evaluations have mainly focused on the simulation of annual mean, not on seasonal cycle. Although there have been a few studies assessing the performance of models in simulating SST seasonal cycle, they commonly used specific months to represent different seasons or focused on specific regions. As will be discussed in chapter 2, this study uses the local maximum and minimum SSTs regardless of the month in which they occur. To our knowledge, there has been no assessment of biases in seasonal SST extrema on a global scale for the CMIP6 models.

The questions addressed in this thesis are as follows: (1) How well do the climate models simulate the seasonal extrema of sea surface water characteristics? (2) Is there any systematic dependence of model performance on the basic model characteristics (e.g. model resolution)?

Chapter 2 introduces the definition of seasonal extrema and compares different observational climatologies. Some regions of the ocean are found to display considerable differences between different climatologies, and these regions are excluded from the model evaluations in later chapters. Chapter 3 assesses the seasonal cycle of sea surface water characteristics in HiGEM. The significant bias of SST seasonal cycle in HiGEM leads us to the study of SST seasonal extrema in the latest state-of-the-art climate models, CMIP6. Chapter 4 focuses on the representation of SST seasonal cycle in 20 CMIP6 models and discusses the impact of various model characteristics. Chapter 5 contains the conclusions, and also discusses the limitations of this work and provides suggestions for future work.

2

Observational climatologies and methodology

2.1 Introduction

Observational reference is the basis for model evaluation. Model evaluation work assesses the performance of models by comparison against observations and hence the quality of model evaluation relies on the underlying observations. Uncertainties in the observational reference can cause uncertainties of model evaluation results (Kotlarski et al., 2019). Observational uncertainties can be from instrumental error and quality control. Since climate model evaluation of the global ocean relies on gridded reference data sets, problems can also arise from spatial interpolation of measurements especially in regions with sparse sampling or high spatial variability (Wagner et al., 2007). To address observational uncertainties, previous works employed multiple reference data sources for model evaluation (Kotlarski et al., 2005; Cheneka et al., 2016; Haslinger et al., 2013).

In this chapter, we will assess five observational climatologies to investigate the uncertainty of observational SST and SSS. By comparing the climatologies, we will indicate the regions with large SST differences between climatologies and discuss the possible reasons for these differences. To avoid the regions with high observational uncertainties, we will suggest an uncertainty mask and the grid points under the uncertainty mask will be excluded from the following model evaluation in Chapters 3 and 4.

2.2 Observational climatologies

2.2.1 Climatologies

The observational monthly climatologies used in this thesis are the World Ocean Atlas 2013 (WOA13), the World Ocean Atlas 2018 (WOA18), the Monthly Isopycnal and Mixed-layer Ocean Climatology (MIMOC), the WOCE/Argo Global Hydrographic Climatology (WAGHC) and the Hadley Centre sea ice and SST data (HadISST) (Table 2.1). To make the time frames of the 5 climatologies consistent, the time period we picked in each climatology is 1981-2010 or most close to 1981-2010, as a climatological standard normal.

	Monthly Climatology Name				
	WOA13	WOA18	MIMOC	WAGHC	HadISST
Horizontal resolution	$0.25° \times 0.25°$	$0.25° \times 0.25°$	$0.5° \times 0.5°$	$0.25° \times 0.25°$	$1° \times 1°$
Vertical levels	57	57	81	38	
Max. depth	1500 m	1500 m	1950 dbar	1900 m	surface
Spatial interpolation	isobaric	isobaric	isopycnal	isopycnal	
Time span	1955-2012	1981-2010	mainly 2007-2011	mainly 1985-2016	1981-2010
Variables	T and S	T and S	θ and S	T and S	SST
References	Locarnini et al. (2013)	Locarnini et al. (2018)	Schmidtko et al. (2013)	Gouretski (2018)	Rayner et al. (2003)

Table 2.1: Parameters of the monthly climatologies we used in this chapter. T, S, and θ refer to temperature, salinity and potential temperature respectively.

SSTs in WOA13, WOA18, MIMOC and WAGHC are bulk temperatures, characterising the water of top few meters and measured with in-situ instruments. HadISST includes in-situ measurements and satellite-based SST adjusted from skin temperature. Wentz et al. (2000) compared SSTs from satellite and ocean buoys. The bulk temperature is measured by buoys at 1 to 1.5 m depth, while the skin temperature measured by infrared sensor represents the upper few microns of the ocean. Their root mean square difference is about 0.6 °C in the equatorial regions (Wentz et al., 2000).

Despite the differences in time-period, analysis method, ingested datasets, analysis

resolution in space (see the details in descriptions of each climatology below), the main data source for the climatologies used in this chapter (except for HadISST) is the World Ocean Database (WOD). WOD is the world's largest collection of ocean profile data that are publicly available, uniformly formatted and quality controlled (Boyer et al., 2013). A list of datasets in WOD13 (World Ocean Database 2013) is shown in Table 2.2. WOD13 includes temperature and salinity fields at observed depth levels as well as interpolated to a set of 138 standard depth levels.

DATASET	SOURCE
OSD	Bottle, low-resolution Conductivity-Temperature-Depth (CTD), low-resolution XCTD data, and plankton data
CTD	High-resolution Conductivity-Temperature-Depth (CTD) data and high-resolution XCTD data
MBT	Mechanical Bathythermograph (MBT) data, DBT, micro-BT
XBT	Expendable (XBT) data
SUR	Surface only data (bucket, thermosalinograph)
APB	Autonomous Pinniped Bathythermograph - Time-Temperature-Depth recorders attached to elephant seals
MRB	Moored buoy data from TAO (Tropical Atmosphere-Ocean), PIRATA (moored array in the tropical Atlantic), MARNET, and TRITON (Japan-JAMSTEC)
PFL	Profiling float data
DRB	Drifting buoy data from surface drifting buoys with thermistor chains
UOR	Undulating Oceanographic Recorder data from a Conductivity/Temperature/Depth probe mounted on a towed undulating vehicle
GLD	Glider data

Table 2.2: Instrument types in WOD13, from Boyer et al. (2013).

The standard depth levels for the climatologies are shown in Fig. 2.1. The output of WAGHC and MIMOC climatologies we used is on isobaric surfaces, although their spatial interpolation is done on isopycnal surfaces. WOA13, WOA18, WAGHC and MIMOC have the same standard depth levels in the upper 50 m, with 5 m vertical resolution nearest to the surface. WAGHC begins to deviate from the others at 50 m, indicating it has larger depth intervals. MIMOC has the highest vertical resolution and WAGHC has the lowest vertical resolution.

The Argo profiling float project since 2003 provides continuous global observations of temperature and salinity and highly improved the data quality and global coverage (Gould et al., 2004). In historically data poor areas like the Southern Ocean (Fan et al., 2014), if each observation was given equal weight, the data distribution can be biased toward the Argo period. To address this issue, climatologies (e.g. WOA13 and WOA18) are averaged for every decade, and then equally contribute to the final climatological mean fields for all the decades (Boyer et al., 2014).
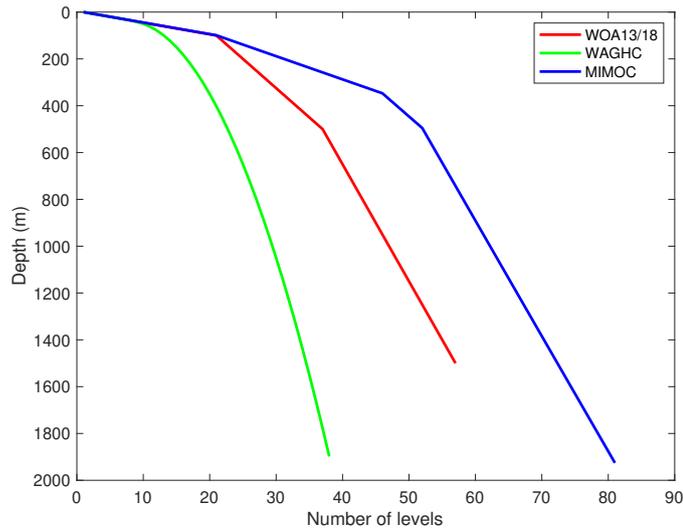
Figure 2.1: Depths against number of levels for WOA13, WOA18, MIMOC and WAGHC. X-axis is the number of levels above the certain depth.

The monthly climatologies we use in this chapter are all gridded products. However, the primary limitation of gridded climatologies based on in-situ observation is data coverage in space and time (Locarnini et al., 2013; Zweng et al., 2013). Fields are based on limited data in some regions such as the Southern Ocean and at deeper levels (Locarnini et al., 2013; Zweng et al., 2013). Or data may exist for only one season, for example, in the Southern Ocean (Fig. 2.2), thus precluding any annual analysis (Locarnini et al., 2013; Zweng et al., 2013).
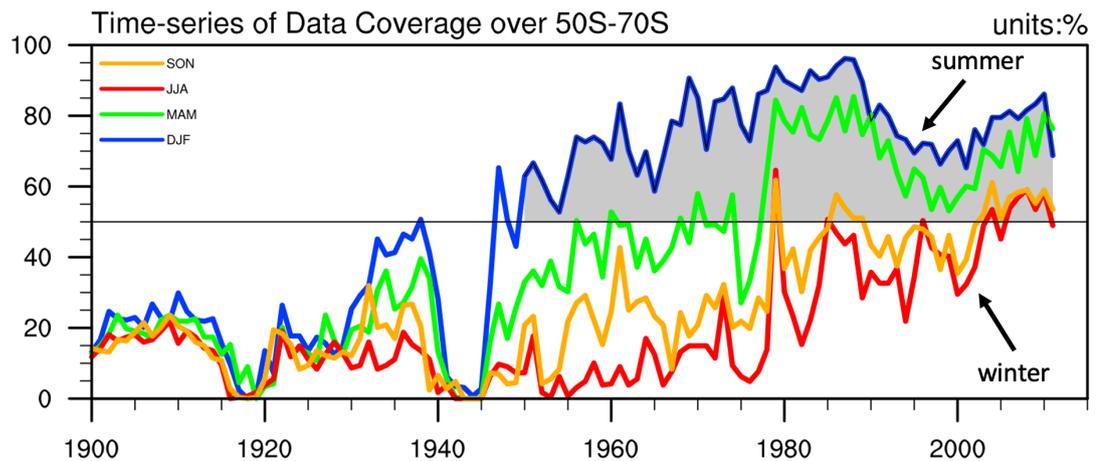
Figure 2.2: Data coverage of HadSST3 over the Southern Ocean (50°-70°S), expressed as a percentage of $5° \times 5°$ grid boxes with at least one observation per season. Continuous data coverage > 50% is shaded. Adapted from Fan et al. (2014).

The spatial interpolations of the climatologies are usually performed on isobaric or

isopycnal surfaces. Sea water density has a nonlinear dependency on temperature and salinity, thus averaging of temperature or salinity on isobaric surfaces results in the production of water masses different from those of the observed data (Lozier et al., 1994). Large gradients of temperature and salinity can also be smoothed when averaging on isobaric surfaces. Isopycnal maps better follow water parcels both laterally and vertically. Since mixing in the ocean interior takes place predominantly along isopycnal surfaces, the interpolation performed on isopycnals can minimize production of artificial water masses (Gouretski, 2018). However, isopycnal maps can be biased near their surface outcrops where data are only available on one side of the mapped grid point, which may result in small temperature inversions or other discontinuities when matching mixed layer and isopycnal properties at outcrop locations, especially in regions of large surface density gradients and sparse data distributions (Schmidtko et al., 2013).

For WAGHC, the largest difference between the isopycnally and isobarically averaged climatologies occurs in regions of strong spatial temperature and salinity gradients, for example, the Gulf Stream, Kuroshio, and ACC (Fig. 2.3), because averaging on isobars can smooth water properties (temperature and salinity) at frontal regions. In those regions, the absolute difference in temperature can exceed 1°C. The differences peak at about 150 m and then start to diminish with increasing depth.

2.2.2   WOA13 and WOA18

The World Ocean Atlas 2013 (WOA13) provides a long-term set of analysed monthly climatologies for temperature and salinity (Locarnini et al., 2013; Zweng et al., 2013). It is the gridded objective analysis of WOD13. Generally, the observational data diminish in number with increasing depth. In the upper ocean, it is reasonable to illustrate large-scale ocean features. However in some deep ocean areas, the distribution of observations can be inadequate for describing characteristics of ocean properties (Locarnini et al., 2013; Zweng et al., 2013). For temperature and salinity, WOA13 includes analysis with quarter-degree horizontal resolution. The quarter degree monthly climatologies for temperature and salinity are available for 1955-2012 and the 2005-2012. Depths for
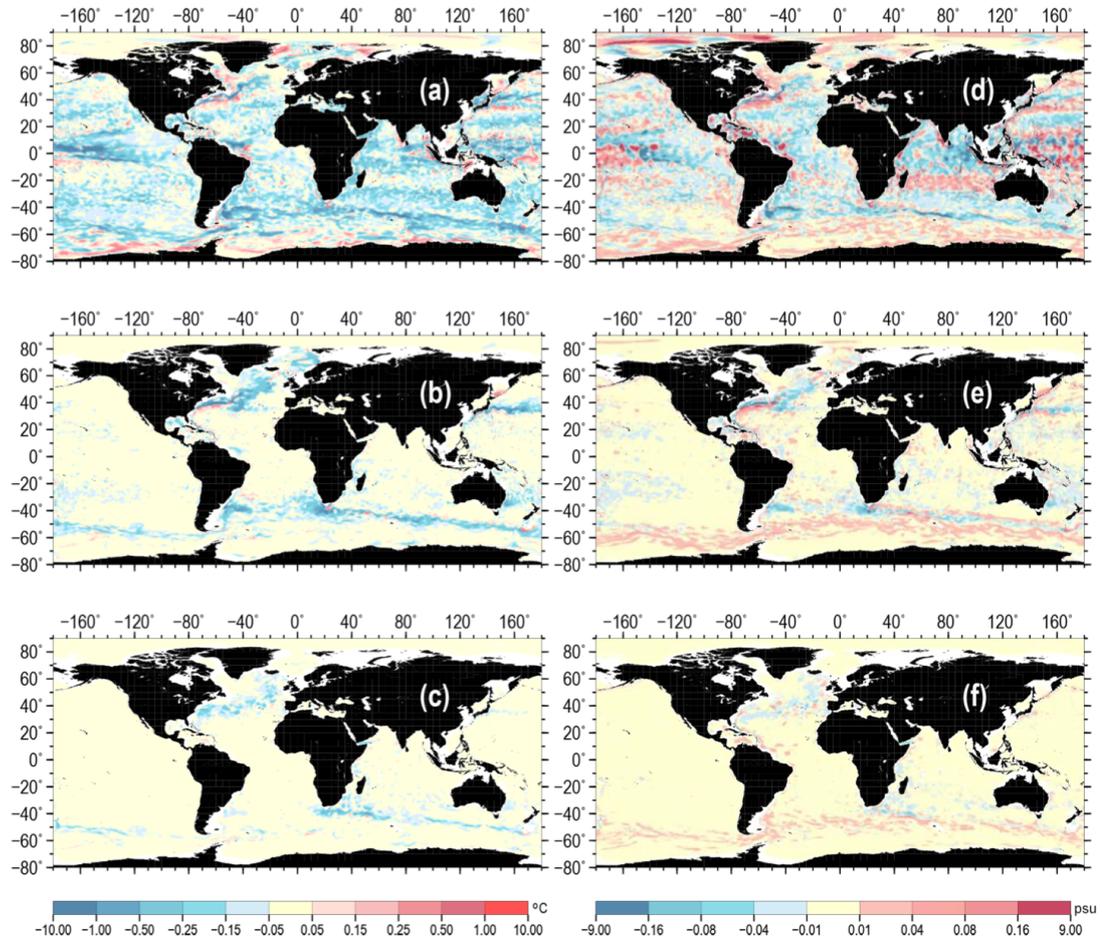
Figure 2.3: Differences between the isopycnally averaged and isobarically averaged WAGHC climatologies of (a-c) temperature and (d-f) salinity in January at (a,d) 150 m, (b,e) 518 m, (c,f) 1050 m. Adapted from Gouretski (2018).

the monthly climatology we used is 0-1500 m with 57 levels, decreasing resolution with depth. The gridded products have been used widely for climate studies, ocean model initialization and validation. For example, Seidov et al. (2017) analysed the decadal variability of ocean heat content and temperature trends in the North Atlantic Ocean using WOA13. Graham et al. (2016) used WOA13 as lateral boundaries for a regional model to study on-shelf heat transport along the West Antarctic Peninsula. WOA13 was also used for model evaluation in CMIP5 (Flato et al., 2013).

The World Ocean Atlas 2018 (WOA18) release updates previous versions of WOA13 to include approximately three million new oceanographic casts added to the WOD since previous release as well as renewed and updated quality control (Locarnini et al., 2018; Zweng et al., 2018). However, even with additional data, WOA18 is still hampered by a lack of data like WOA13 (Locarnini et al., 2018; Zweng et al., 2018). For example,

there is limited sampling in the Southern Ocean in winter (Fig. 2.4), even though animal mounted temperature profiles have been added in WOA18 and improved the data coverage in some high latitude areas (Locarnini et al., 2018).
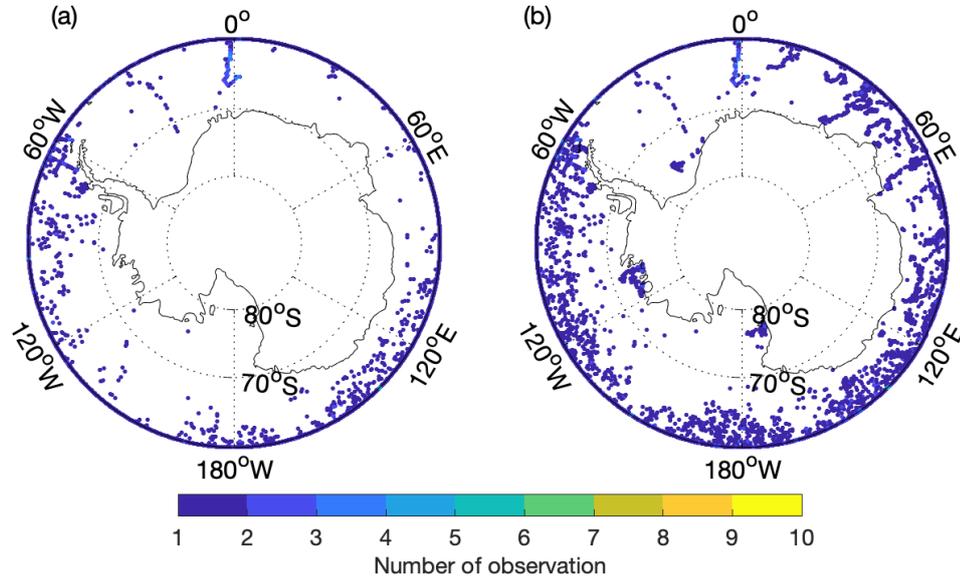
Figure 2.4: Data distribution of SST in July (austral winter) from (a) WOA13 and (b) WOA18.

In WOA18, the standard error of the statistical mean of SST in a quarter-degree grid box (Locarnini et al., 2018) was calculated using

$$s = \sqrt{\frac{\sum_{n=1}^{N}(x_n - \bar{x})^2}{(N-1) \cdot N}}, \tag{2.1}$$

where x_n is the nth SST value in the gridbox, $\bar{x}$ is mean of all SST values in the gridbox, and N is the total number of SST values in the gridbox.

For 1981-2010 decadal average, the standard error of the mean SST in each month over most of the world ocean is below $0.1°C$, while in some coastal regions the standard error is over $0.5°C$ (Fig. 2.5). The errors estimated in WOA18 contain errors from sampling and measurement, but not from the creation of the gridded data sets (e.g. area-averages from sparsely distributed observations).
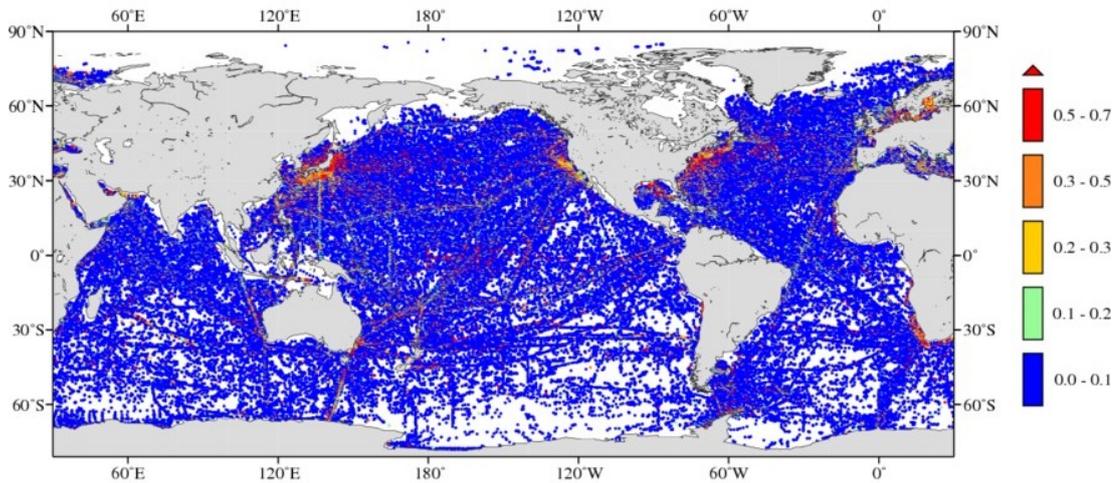
Figure 2.5: Standard error (unit:°C) of the mean for 1981-2010 averaged January SST for quarter-degree grid in WOA18, adapted from WOA 2018 Figures (https://www.ncei.noaa.gov/access/world-ocean-atlas-2018f/bin/woa18f.pl).

2.2.3 Monthly Isopycnal & Mixed-layer Ocean Climatology (MIMOC)

MIMOC is a global monthly, isopycnal and mixed layer ocean climatology (Schmidtko et al., 2013). MIMOC provides three products: (1) mapped mixed layer properties, (2) mapped water properties on isopycnal surfaces, and (3) water properties from the first two products merged onto a regular pressure grid, from 0 to 1950 dbar. MIMOC includes conductivity-temperature-depth (CTD) profiles from Argo floats, shipboard data from the World Ocean Database 2009 (WOD2009) and Ice-Tethered Profile (ITP) data (Fig. 2.6). Argo CTD data are the main data contributor in the open ocean, and ITPs provide data under Arctic sea ice. MIMOC mostly reflects the 2007-2011 modern ocean state, although old data as early as the 1970s are used when no data during 2007-2011 are available. Compared with other climatologies in widespread use at the time, MIMOC is better or as good as at preserving features observed in a synoptic survey and minimizing the influences of eddies, planetary waves, internal waves and tides, and other transient phenomena (Schmidtko et al., 2013).

In MIMOC, Argo (Argo, 2000) data are the main data contributor in the open ocean (Schmidtko et al., 2013). But since Argo just covered the open ocean but not continental shelves, some marginal seas and most ice-covered regions, shipboard data are used in these regions (Schmidtko et al., 2013). The sampling periods of shipboard, Argo and
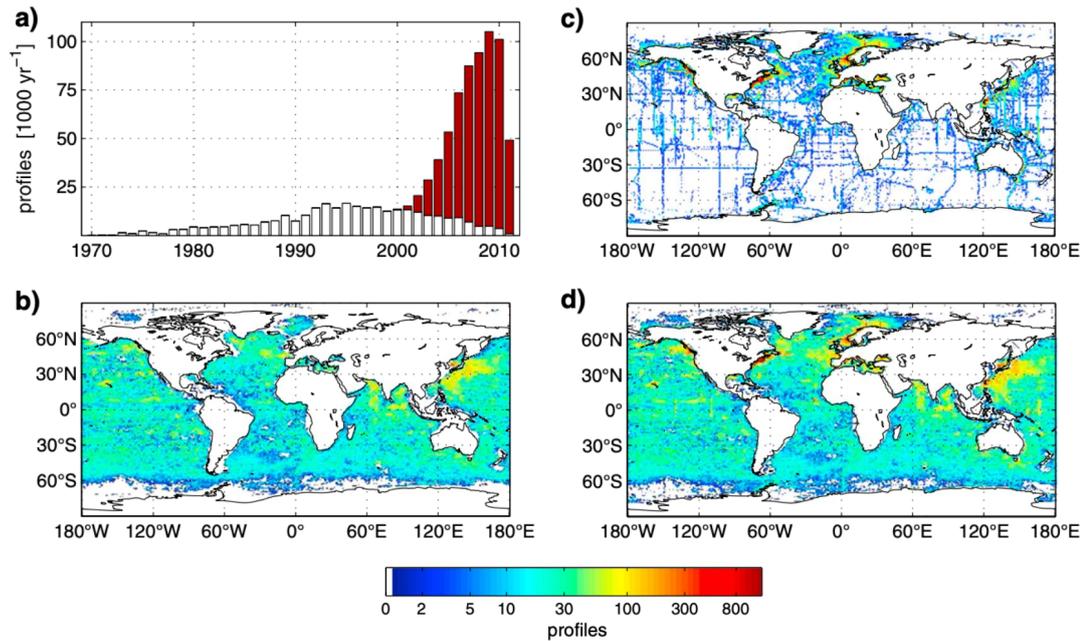
Figure 2.6: MIMOC data distribution (a) Temporal distribution of CTD profiles from WOD2009 (white), Argo and ITP profiles (red). Spatial distribution for each grid box of (b) Argo and ITP profiles, (c) WOD2009 profiles and (d) Argo, ITP and WOD2009 profiles combined. Adapted from Schmidtko et al. (2013).

ITP are very different. Argo provides a large amount of data for the open ocean during 2007-2011, while ITP data in shelf regions and high-latitudes are limited during 2007-2011 and more representative of the ocean state before 2000 (Schmidtko et al., 2013).

2.2.4 World Ocean Circulation Experiment-Argo Global Hydrographic climatology (WAGHC)

WAGHC is a climatology with a 1/4° spatial resolution resolving the annual cycle of temperature and salinity on a monthly basis (Gouretski, 2018). The nominal climatological period is 1985-2016, although they relax back to the older data when no data from 1985-2016 window (mostly in high latitudess and several marginal seas) are available (Gouretski, 2018). WAGHC is an update of the WGHC climatology (WOCE Global Hydrographic Climatology) (Gouretski and Koltermann, 2004), using the improved data from the Argo programme. It has two versions of the climatology, with spatial interpolation performed on isobaric and isopycnal surface respectively. As the mixing of water mass in the real ocean takes place along isopycnals, averaging along isobaric surface does not reflect the process and can produce artificial water

	Number of profiles	% all
Instrumentation type		
OSD	2098823	44.452
CTD	971222	20.570
PFL	1368880	28.992
APB	282593	5.985
Data source		
WOD13	4665330	98.810
AWI, Bremerhaven, Germany	50848	1.077
Canadian institutions	5340	0.113

Table 2.3: Instrumentation types and data sources of the WAGHC climatology. Adapted from Gouretski (2018), their table.1.

masses. We compared SST and SSS in isobarically and isopycnally averaged versions and found that different version does not make any difference to the sea surface water characteristics. In this Chapter we use the isopycnal version of WAGHC.

WOD13 and its update in January 2017 serve as the main data source for WAGHC (Fig. 2.3). OSD, CTD, PFL and APB were used. The APB data were only used in the Southern Hemisphere where data coverage is poor. XBT and MBT were not used as temperature and salinity are required for the spatial interpolation on isopycnal surfaces (Gouretski, 2018). WAGHC added 50848 profiles from Alfred Wegener Institute (AWI), Bremerhaven and 5340 profiles from Canadian institutions to the profiles from WOD13 (Boyer et al., 2013), which improves the database for northern polar regions significantly.

The WAGHC differs from WOA13 in the interpolation method (isopycnal versus isobaric averaging) and database (WAGHC includes additional 4 years new Argo profile and hydrographic data from the North Polar regions). The WAGHC climatology has better representation of the thermohaline structure both in the data poor polar region and some data abundant regions (e.g. Baltic sea, Caspian sea, Gulf of California, Caribbean Sea and the Weddell sea) compared with WOA13, which produces unrealistic salinity and temperature in these regions (Gouretski, 2018).

2.2.5 Hadley Centre sea ice and SST data (HadISST)

The Hadley Centre sea ice and SST data set version 1 (HadISST1) is a combination of monthly global fields of SST and sea ice concentration on a 1° latitude-longitude grid from 1870 to date (Rayner et al., 2003). HadISST1 was developed at the Met Office Hadley Centre and improves upon previous global sea ice and SST (GISST) dataset. The SST fields in HadISST1 have more uniform variance through time and better month-to-month persistence than those in GISST (Rayner et al., 2003). HadISST includes both in situ SSTs and satellite-based SSTs. In situ data are taken from individual ship observations from the Met Office Marine Data Bank (MDB) and the Comprehensive Ocean-Atmosphere Data Set (COADS). Satellite based SSTs are from the advanced very high resolution radiometer (AVHRR) (Rayner et al., 2003). Satellite SST data are based on measuring electromagnetic radiation that left the ocean surface and transmitted through the atmosphere. Such data have to be calibrated using in situ observations. The primary purpose of HadISST1 is to force atmospheric models in the simulation of recent climate and to evaluate coupled atmosphere-ocean models (Rayner et al., 2003). The spatial resolution of HadISST1 is not high enough to resolve very localized SST features or the meanderings of the Gulf Stream (Rayner et al., 2003).

2.3 How to define seasonal extrema?

This thesis is concerned with the seasonal cycle of water properties. This section will introduce two methods to describe the seasonal cycle. One is using a sinusoidal annual cycle, and the other is using maximum and minimum values. Here, using SST in WOA13 data as an example, we will compare the two methods and explain why we choose the monthly maximum and minimum values to define seasonal extrema.

2.3.1 The sinusoidal annual cycle

The annual cycle can be assessed by fitting the monthly properties to a sine curve. The equation is written as:

$$E(i) = A\cos(\omega(i - \varnothing)) + \beta, \tag{2.2}$$

where E is the expected value of the monthly property, i is the month, β is the annual mean value, ω is the wave frequency (here, $\omega = 2\pi/12$), and A and $\varnothing$ are the amplitude and phase of the fitted sinusoid. Phase indicates the month when value reaches maximum.

To see how well the monthly time series can be fitted to a sinusoidal annual cycle, here we also calculated the goodness of fit (GOF) χ^2:

$$\chi^2 = 1 - \frac{\sum_{i=1}^{N}(O_i - E_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O}_i)^2}, \tag{2.3}$$

where O_i is an observed value in month i, and N is the sample size. $\bar{O}_i$ is the mean value of O_i. The value of GOF ranges from 0 to 1. A large value of GOF means the observation fits the estimation well, while a small value means a poor fit.

The monthly time series is fitted into a sinusoidal annual cycle by using harmonic analysis (Bloomfield, 2004). Harmonic analysis decomposes the time series into a sum of sinusoidal components, and the sinusoidal annual cycle obtained here is the sinusoidal component with a 12 months length period. The original 12 months time series might be not long enough to obtain a component with 12 months period. To obtain a better fitting, we repeated the monthly climatological time series (N=12) 10 times to get a longer time series (N=120) for harmonic analysis. It shows the difference between time series with different lengths in Fig. 2.7. After repeating the SST time series 10 times, the fitting of the sine curve in Fig. 2.7b (GOF=0.64) is better than that in Fig 2.7a (GOF=0.60).

Examples of sine fitting for the monthly climatological SSTs at different locations are shown in Fig. 2.8. In Figs. 2.8a-b, two SST time series at two points in the North Pacific both have good fit to sine curves. SST has a larger amplitude of annual cycle in the subtropical zone (at 153°E, 36°N) than in the tropical zone (at 166°E, 22°N) (Figs. 2.8a-b). The maximum observed SST in Figs. 2.8a-b occurs in September, very
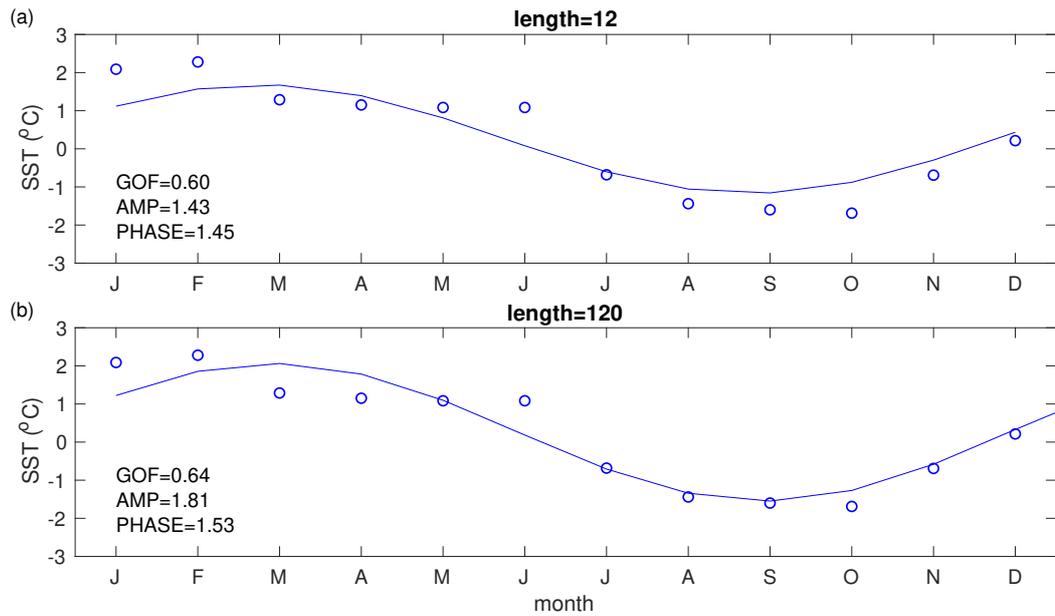
Figure 2.7: Monthly climatological SSTs (circles) and their fitted sine curves (lines) for (a) the original time series (length=12) and (b) the time series repeating the original time series 10 times (length=120) at point 60°S, 90°E from WOA13

close to the phase of their fitted sine curves (the beginning of October). The timings of observed minimum SST in Fig. 2.8a-b are April and March respectively. Here the GOF in Fig. 2.8b (0.86) is about the same as in Fig. 2.8a (0.85) even though the monthly SST in Fig. 2.8b seems better fitted to the sinusoidal annual cycle. That is because larger standard deviation makes smaller GOF according to Equation 2.3.

The sea water at 179°E, 79°N in the Arctic is almost frozen throughout the year except in April when the SST peaks at -0.4 °C (Fig. 2.8c). The SST time series at 62°E, 16°N in the Arabian Sea has a semi-annual cycle, resulting in the poor fit for the one-year period sine curve (Fig. 2.8d). In Figs. 2.8c-d, the months of maximum observational SST are April and June respectively, while their phases are both about one month later. The coldest month in Figs. 2.8c-d is March. For Fig. 2.8c in the Arctic, the coldest timing is just one month ahead of the warmest month. Figs. 2.8c-d indicates that when time series do not have an obvious annual cycle, the one-year period sine fitting will be poor and the phase which is expected to represent the month of maximum observational SST can be incorrect.

It should be noticed that in some cases the observational uncertainty can lead to errors in finding T_{max} and T_{min}. When the standard error is large at or near the timing of
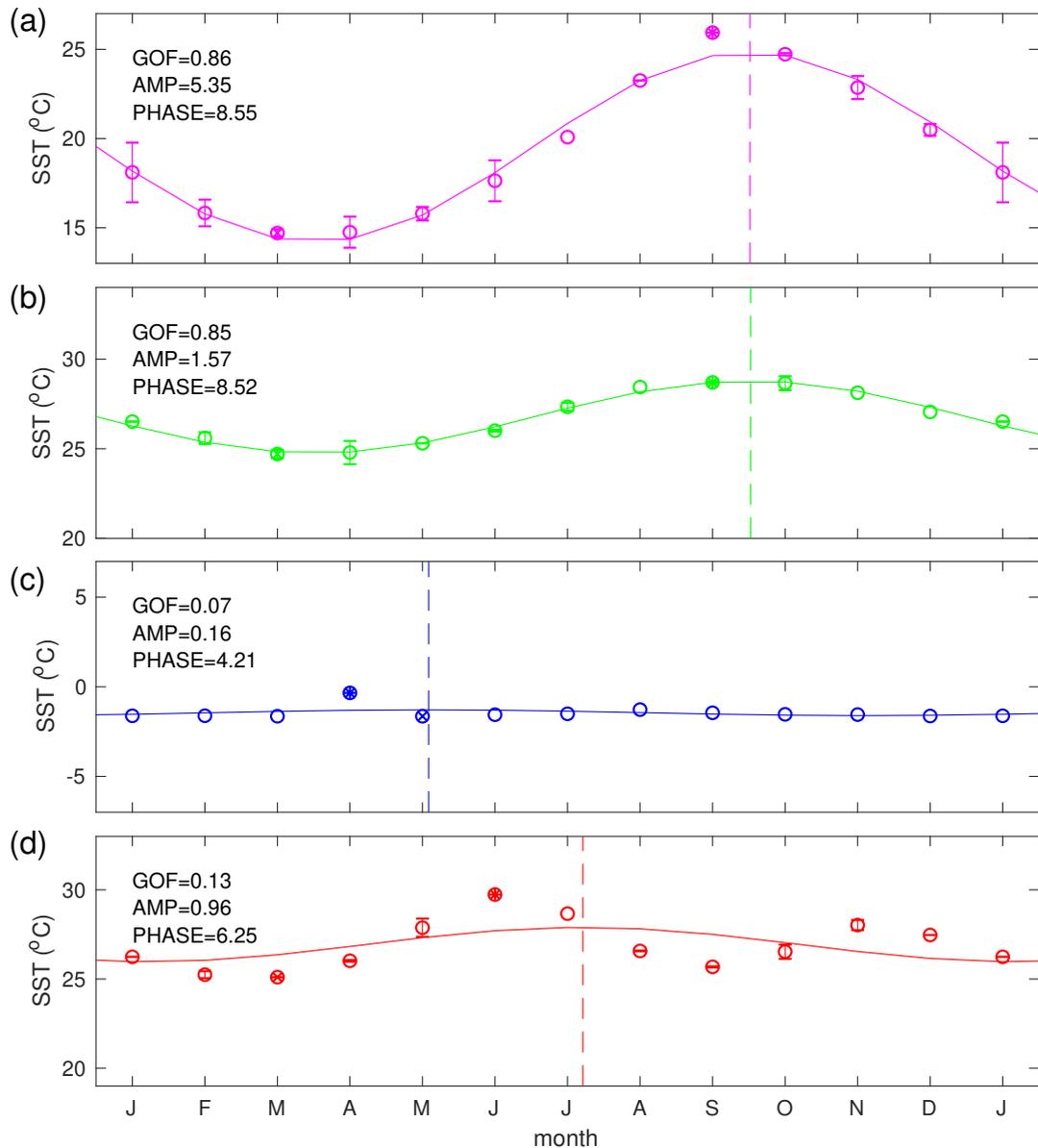
Figure 2.8: Monthly climatological SSTs (circles) and their fitted sine curves (lines) at (a) 153°E, 36°N, (b) 166°E, 22°N, (c) 179°E, 79°N (d) 62°E, 16°N (star points in Fig. 2.10b). Bars on the circles are standard errors of SSTs, as described in Section 2.2.2 (circles without bars indicate no observation in the grid box of the point). Colors here match colors of stars in Fig. 2.10. Ticks on X-axes represent the middle of each month. Four time series have the same range for their Y-axes. The value of GOF, amplitude and phase for each sinusoid are shown at the left upper corner. The circles with star and cross represent maximum and minimum observed SST in 12 months respectively, while the dashed lines are the maxima of the sinusoidal annual cycles.

T_{max}/T_{min}, there can be errors in finding the timing of T_{max}/T_{min}, and hence errors in T_{max}/T_{min}. For example, in Figs. 2.8a-b T_{min} occurs in March, however, as the SST in April has a large standard error, the timing of T_{min} will be April if the SST in April reaches its lowest value within the error bar. Consequently, the T_{min} can be smaller. When the standard error is small compared to the difference between T_{max}/T_{min} and the second maximum/minimum SST, it is very unlikely to have errors in finding T_{max}/T_{min} (Figs. 2.9a,b,d,e,f,i).

To define a threshold for GOF, we randomly picked 9 examples with GOF value varying from 0.1 (very bad fitting) to 0.9 (very good fitting): GOF=0.1 at 150.125°E, 0.125°S; GOF=0.2 at 150.125°E, 1.125°N; GOF=0.3 at 147.875°E, 5.125°N; GOF=0.4 at 155.875°E, 5.125°S; GOF=0.5 at 155.875°E, 7.375°N; GOF=0.6 at 156.375°E, 10.125°N; GOF=0.7 at 65.125°E, 1.125°N; GOF=0.8 at 150.125°E, 10.125°S; GOF=0.9 at 150.125°E, 29.125°N (Fig. 2.9). Here, we choose the GOF=0.5 as the threshold. The areas with GOF larger than 0.5 have well fitted annual cycle of SST, and their amplitudes and phases can be trusted to study seasonal extrema. In areas where GOF<0.5, the difference between the timings of observed maximum SST and fitted sinusoid is larger than one month, and therefore sinusoidal annual cycle is not applicable to be used for detecting the SST seasonal extrema and their timings.

The maps for amplitude and phase of the fitted sinusoids, and its GOF in SST are shown in Fig. 2.10. The amplitude of SST sinusoids is largest in the western North Pacific and Atlantic, which can be over 10°C near the western boundary (Fig. 2.10a). Compared with the northern hemisphere, the amplitude in the southern hemisphere is much smaller (generally less than 5°C), due to less land in the southern hemisphere. SST of the northern hemisphere mainly reaches maximum in September and SST of the southern hemisphere mainly reaches maximum in March (Fig. 2.10b).

SST seasonal variabilities in global ocean do not all comply with sinusoidal annual cycles. It has been illustrated that the mid latitudes is dominated by sinusoidal annual cycle, while in equatorial and polar region sinusoidal annual cycle can only explain a low percentage to the total SST seasonal variance (Trenberth, 1983; Yashayaev and Zveryaev, 2001). Here we investigate how well observed monthly SST time series can
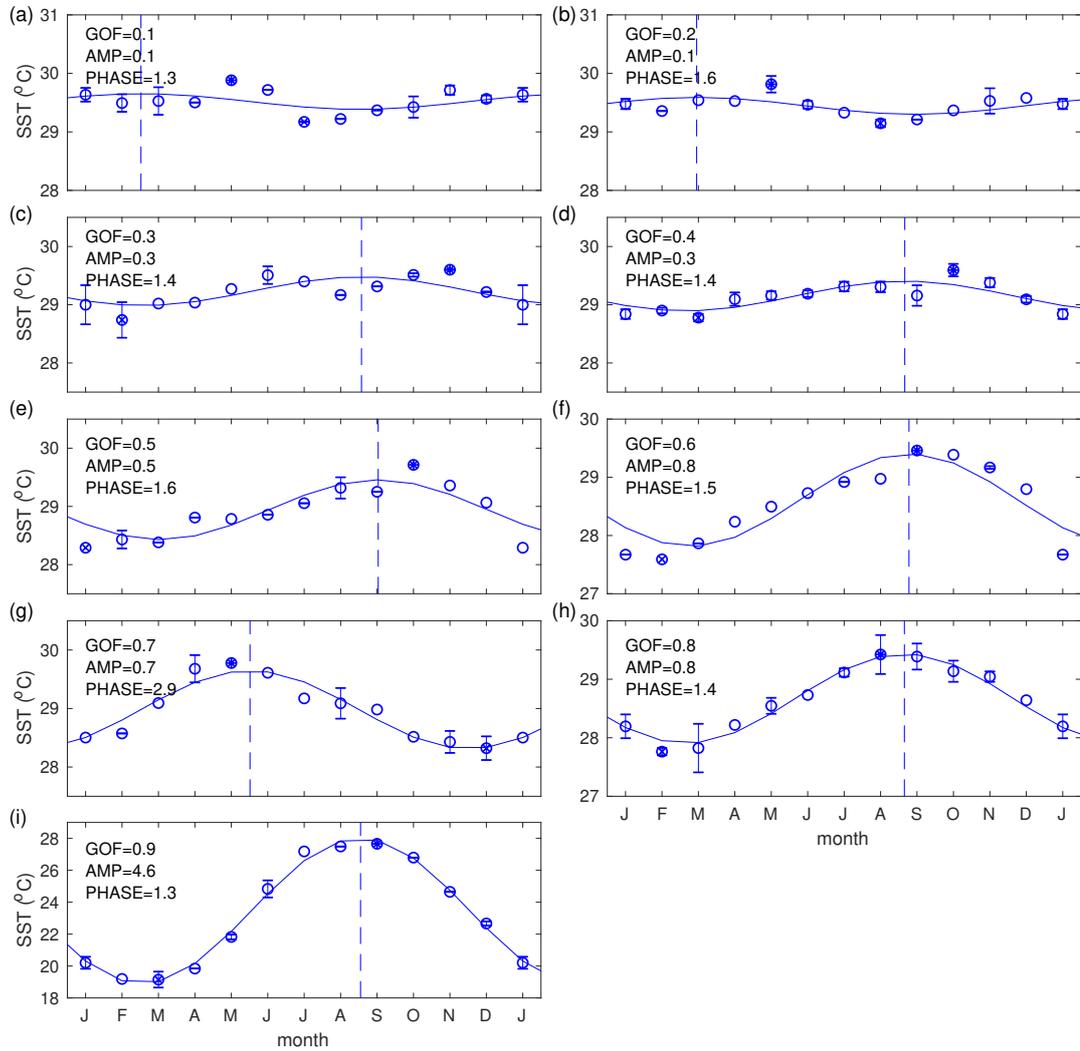
Figure 2.9: Monthly climatological SSTs (circles) and their fitted sine curves (lines) with different GOF value. (a) GOF=0.1, at 150.125°E, 0.125°S, (b) GOF=0.2, at 150.125°E, 1.125°N, (c) GOF=0.3, at 147.875°E, 5.125°N, (d) GOF=0.4, at 155.875°E, 5.125°S, (e) GOF=0.5, at 155.875°E, 7.375°N, (f) GOF=0.6, at 156.375°E, 10.125°N, (g) GOF=0.7, at 65.125°E, 1.125°N, (h) GOF=0.8, at 150.125°E, 10.125°S and (i) GOF=0.9, at 150.125°E, 29.125°N. Bars on the circles are standard errors of SSTs, as described in Section 2.2.2 (circles without bars indicate no observation in the grid box of the point). Ticks on X-axes represent the middle of each month. Time series (a-h) have the same range for their Y-axes. The value of GOF, amplitude and phase for each fitted sinusoid are shown on each panel. The circles with star and cross represent maximum and minimum observed SST in 12 months respectively. The dashed lines indicate the maxima of the sinusoids.
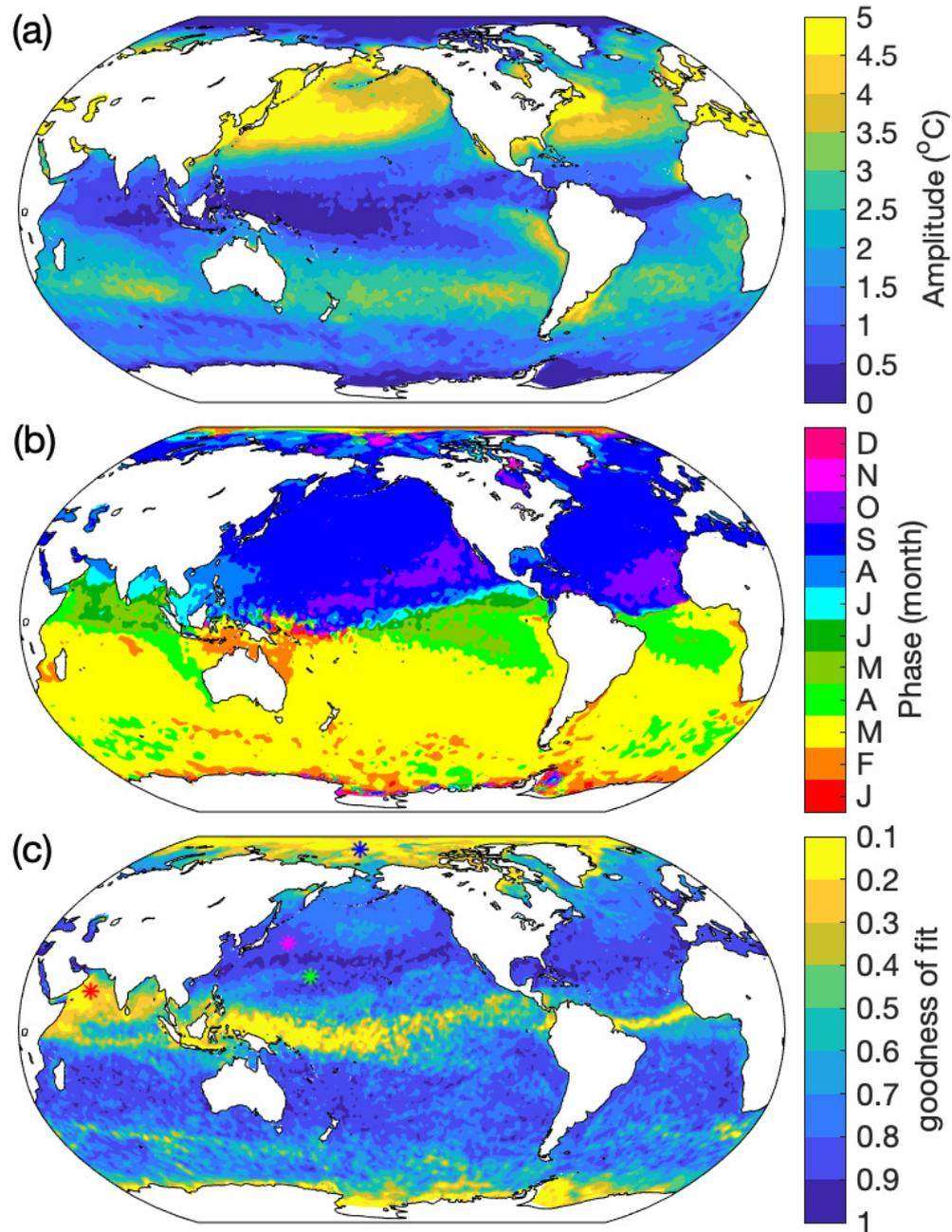
Figure 2.10: (a) Amplitude, (b) Phase, and (c) GOF of fitted sinusoids for monthly climatological SST from WOA13. The stars are the points for SST time series in Fig. 2.8.

be fitted to a sinusoidal annual cycle and calculated the GOF. It is found that the locations with bad fit (Fig. 2.10c) are roughly consistent with locations of Yashayaev and Zveryaev (2001) with low contribution of the sinusoidal annual cycle to the total variance of SST (Fig. 2.11). The examples of bad fitting can be seen in Figs. 2.8c-d. In the polar regions, North Indian Ocean and areas near the equator the value of GOF is poor (less than 0.5) (Fig. 2.10c).
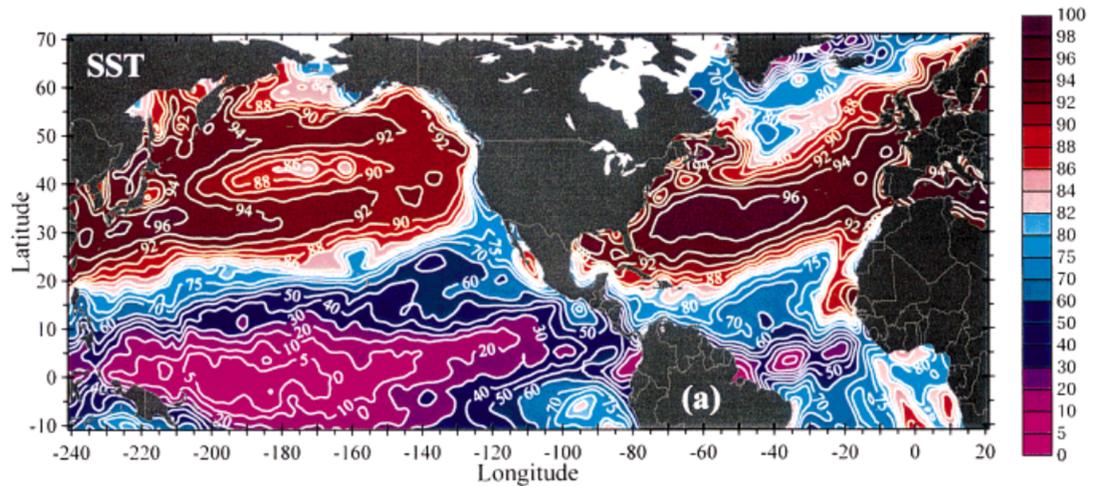
Figure 2.11: Contribution of the annual cycle to the total variance of SST, in percent. Adapted from Yashayaev and Zveryaev (2001).

Examples of sine fitting for the monthly climatological SSSs at different locations are shown in Fig. 2.12. In Fig. 2.12a, the SSS time series at 143.375°E, 31.875°N in the North Pacific have good fit to a sine curve. The timing of SSS maximum is close to the phase of the fitted sine curve, with a difference less than one month. At the other point (166.125°E, 22.125°N) in the North Pacific, the SSS time series (Fig. 2.12b) does not have an annual cycle similar to the one in Fig. 2.12a. It looks more random and has a bad fit to the one-year period sine curve. At 65.125°E, 14.125°N in the Arabian Sea, SSS has semi-annual cycle, which leads to a poor fit for the one-year period sine curve (Fig. 2.12c). The phase is about 2 months earlier than the timing of maximum. At 179°E, 79°N in the Arctic, SSS in August is only about 24, much less than the SSS (about 30) in other months. That may indicate large uncertainty of SSS in the Arctic, although the standard error is unavailable at that point. The phase there is three months earlier than the timing of SSS maximum.

2.3.2 Maximum and minimum values

Since not everywhere in the world ocean can fit a yearly-period sine curve very well, we used another method to define seasonal extrema: finding the maximum or minimum value in the climatological monthly time series, and then the month of maximum or minimum value is the timing of seasonal extrema. This method can be applied everywhere in the world ocean even in areas with bad fitting for sinusoidal
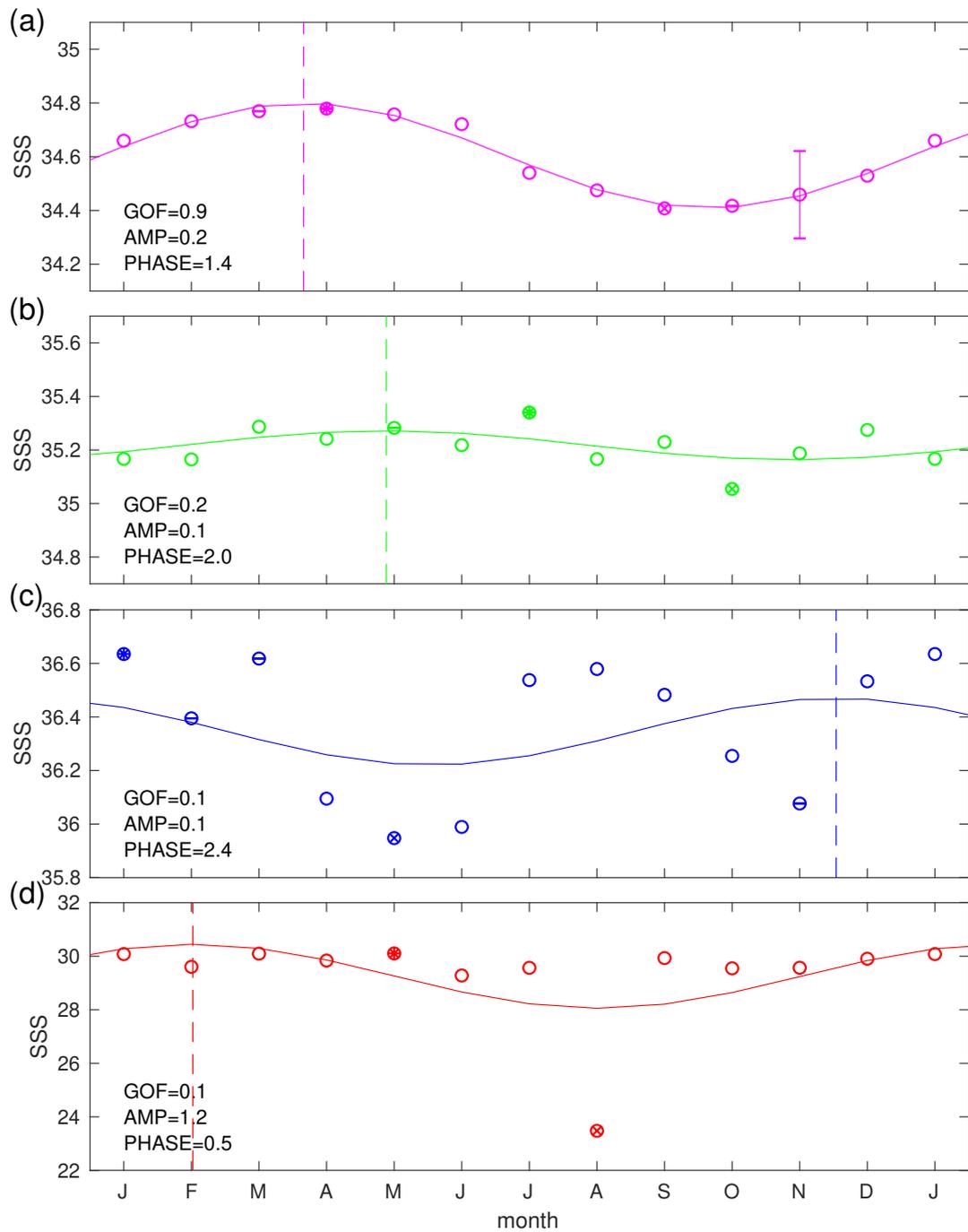
Figure 2.12: Monthly climatological SSSs (circles) at (a) 143.375°E, 31.875°N, (b) 143.375°E, 25.625°N, (c) 65.125°E, 14.125°N (d) 179°E, 79°N. Bars on the circles are standard errors of SSSs.

annual cycles.

Hereafter, T_{max} refers to maximum temperature from the climatological monthly temperature, T_{min} refers to the minimum temperature. The timings of T_{max} and T_{min} represent summer and winter respectively. Similarly, S_{max} and S_{min} refer to maximum and minimum salinity. The timing of S_{max} is the saltiest season, while the timing of S_{min} is the freshest season.

The peak-to-peak annual amplitude is calculated by T_{max} minus T_{min}, hereafter T_{cycle}. Fig. 2.13 shows T_{cycle} at the surface, 100 m and 700 m. Here 100 m represents the subsurface and 700 m below the thermocline. T_{cycle} decreases with depth. At surface, SST has the largest T_{cycle} in the subtropical area and the smallest seasonal cycle in polar areas and equatorial area. T_{cycle} in the western-central Pacific and western Atlantic can be larger than 10°C (see Fig. 2.13a), while the range there decreases to less than 5°C at 100 m (Fig. 2.13b). At 100 m, the equatorial areas in Indian Ocean, eastern Pacific and western Atlantic have larger T_{cycle} than higher latitudes (Fig. 2.13b). That may be related to the seasonal cycle of MLD. The mixed layer in these areas is deeper than 100 m in winter (Kara et al., 2003), which allows the surface heat to penetrate to more than 100 m. In summer, the global mixed layer is shallower than 100 m (Kara et al., 2003), the water at 100 m is below the mixed layer thus it is hard to be reached by the surface heating. At 700 m, T_{cycle} becomes much smaller and most of the world ocean has T_{cycle} less than 1 °C (Fig. 2.13c). This is due to the fact that sea water below thermocline can hardly be influenced by seasonal atmospheric processes. In this thesis, we will focus on the sea surface water which has the most significant seasonal cycle.

The global spatial pattern is similar for T_{cycle} (Fig. 2.13a) and for the amplitude of the fitted sinusoidal annual cycles (Fig. 2.10). If the annual cycle can be well fitted to a sinusoid, T_{cycle} is twice as large as the amplitude of the sinusoidal annual cycle. Therefore, to make the figures comparable, the range of color bar in Fig. 2.13 is twice as large as that in Fig. 2.10a. It can be seen that in the polar regions and equatorial regions where GOF is poor (Fig. 2.10c), the amplitude based on sinusoidal annual cycle (Fig. 2.10a) is underestimated compared with the real amplitude: $1/2*(T_{max}-T_{min})$ (Fig. 2.13a).
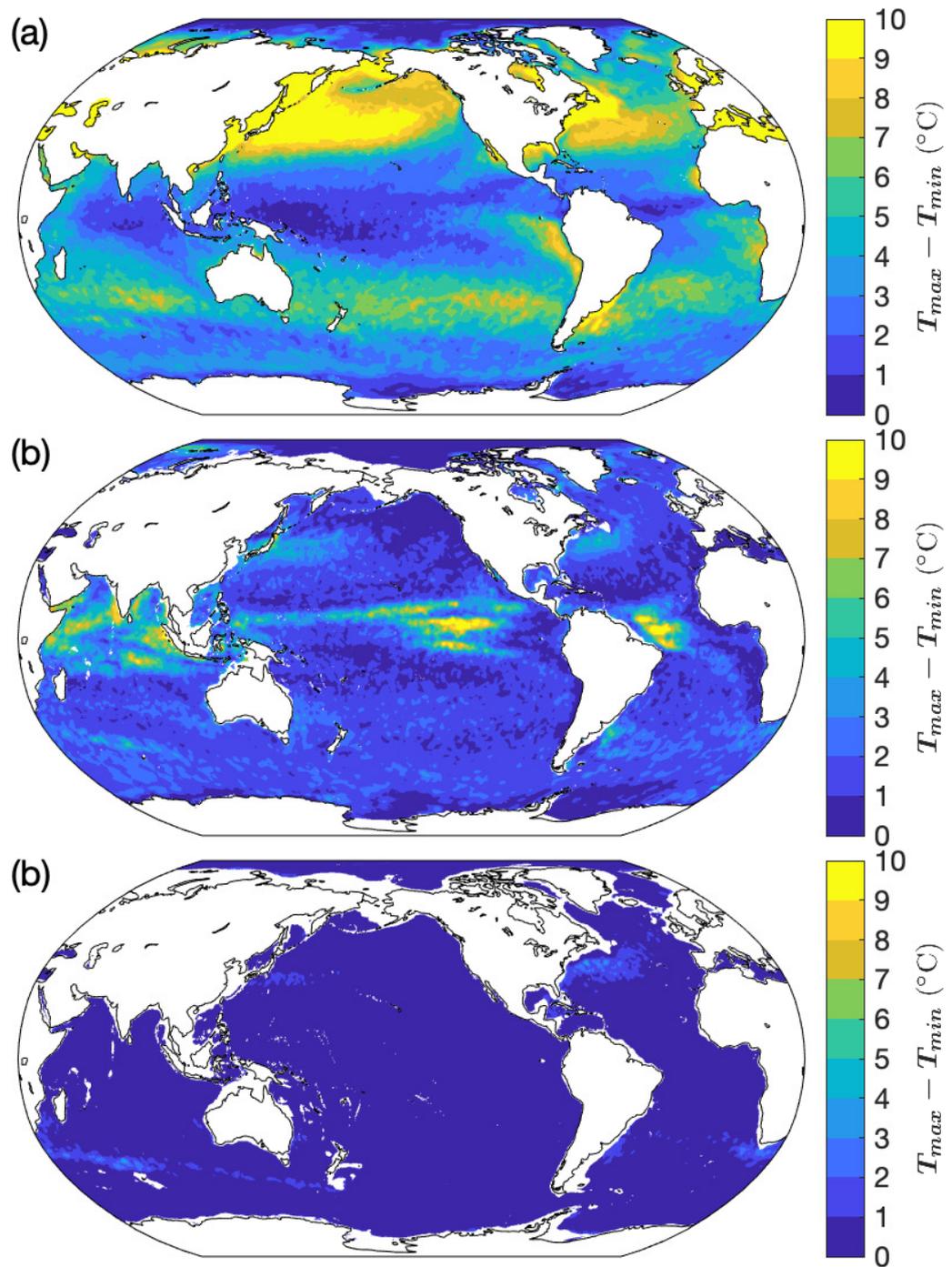
Figure 2.13: T_{cycle} of monthly climatological sea water temperature from WOA18 (a) at surface, (b) at 100 m, and (c) at 700 m. The range of the color bar is twice as large as that in Fig. 2.10a.

The peak-to-peak annual amplitude of SSS is calculated by S_{max}-S_{min}, hereafter S_{cycle}. Large S_{cycle} can be seen in the polar regions, the tropical regions under the ITCZ and SPCZ, and near coastal regions affected by river runoff (Fig. 2.14). Sea ice formation and melting can be responsible for the large S_{cycle} in the polar regions. Freshwater sources from rainfall and river runoff are the main causes of the seasonal variation of SSS in the tropical and coastal regions (Yu et al., 2021).
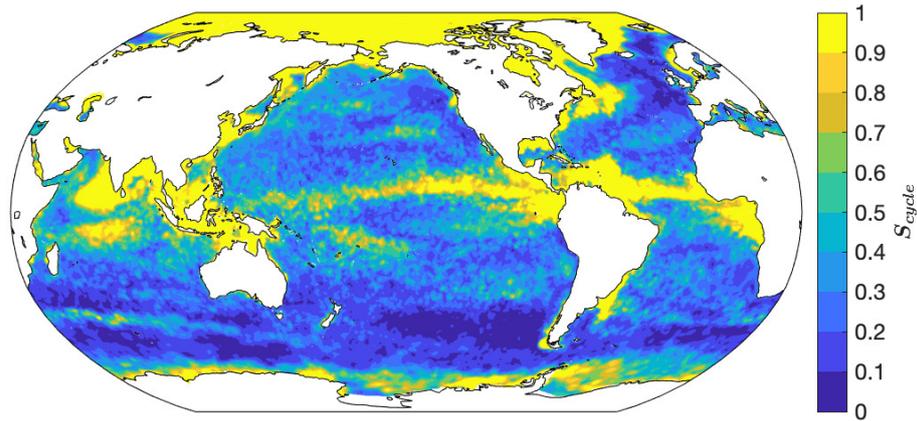
Figure 2.14: S_{cycle} of monthly climatological sea surface water in WOA18.

Compared with using the sinusoidal annual cycle to describe seasonal cycle, using the seasonal maximum and minimum values does not show limitations in regions where an annual sinusoid explains little of the total seasonal variance. The seasonal maximum and minimum can be found whenever they occur. Therefore this thesis will use seasonal maximum and minimum values to study seasonal cycle, and the range of seasonal cycle will be represented by T_{cycle} and S_{cycle}.

2.4 Seasonal extrema

2.4.1 Time periods to choose for WOA13 and WOA18

WOA13 provides monthly climatologies with $1/4°$ spatial resolution for time periods 1955-2012 and 2005-2012. T_{max} from the 2005-2012 climatology is generally warmer than that in the 1955-2012 climatology around the world, but the warming is not uniform (Fig. 2.15). The warming trend of T_{max} in the Atlantic Ocean is most obvious,

compared with the Pacific Ocean and Indian Ocean. The warm difference for T_{max} larger than 1 °C can be seen in most subpolar North Atlantic and the western subpolar North Pacific, while at lower latitudes it is much smaller. Using Hydrostation data, Hallam et al. (2021) compared August-September-October SSTs (summer SSTs) over the North Pacific in the periods of 2000-2019 and 1980-1999 and showed the largest warm trend in the western subpolar North Pacific, which is consistent with our results in Fig. 2.15c. It can also be noticed that at high latitudes regions in the southern hemisphere and in the eastern North Pacific and eastern equatorial Pacific, T_{max} for 2005-2012 climatology is colder than that for 1955-2012 climatology (a cooling trend). Similarly, Turkington et al. (2019) also revealed a cooling SST trend between 1962 and 2011 in these regions. Solomon and Newman (2012) indicated a warming trend for the 1900-2010 period in the equatorial Pacific warm pool and a weak cooling in the cold tongue, which is consistent with our analyses (Fig. 2.15c). The eastern Pacific cooling associated with intensification of the atmospheric Walker circulation can be driven by the warming trend in Atlantic SST (McGregor et al., 2014).

Central-eastern Pacific SST vary with the phase of the El Niño - Southern Oscillation (ENSO). Thus climatology based on recent short period 2005-2012 can be sensitive to the global warming trend or strong El Niño and therefore do not in general reflect long-term climate. To let the WOA13 climatology we use better represent the "climate normal" instead of global warming, the time period for WOA13 was chosen to be 1955-2012. It can also make the time frames of the 5 climatologies more consistent.

WOA18 provides monthly climatologies for 1955-2017, 1981-2010 and 2005-2017. The T_{max} for climatologies of 2005-2017 and 1955-2017 (Fig. 2.16a) show a similar global warming trend to that for climatologies of 2005-2012 and 1955-2012 in WOA13 (Fig. 2.15c), except that the warming trend in the Pacific is larger in the former. The reason might be that time periods used for calculation of T_{max} difference are different. The warming trend dominates the central-eastern Pacific in Fig. 2.16a but not in Fig. 2.15c, which might be related to the El Niño events in 2014-2016. As 2005-2012 climatology from WOA13 and 2005-2017 climatology from WOA18 are based on short records, they can be largely influenced by natural variability like El Niño and lead to different trends. Therefore T_{max} in those climatologies is sensitive to natural variability because of their
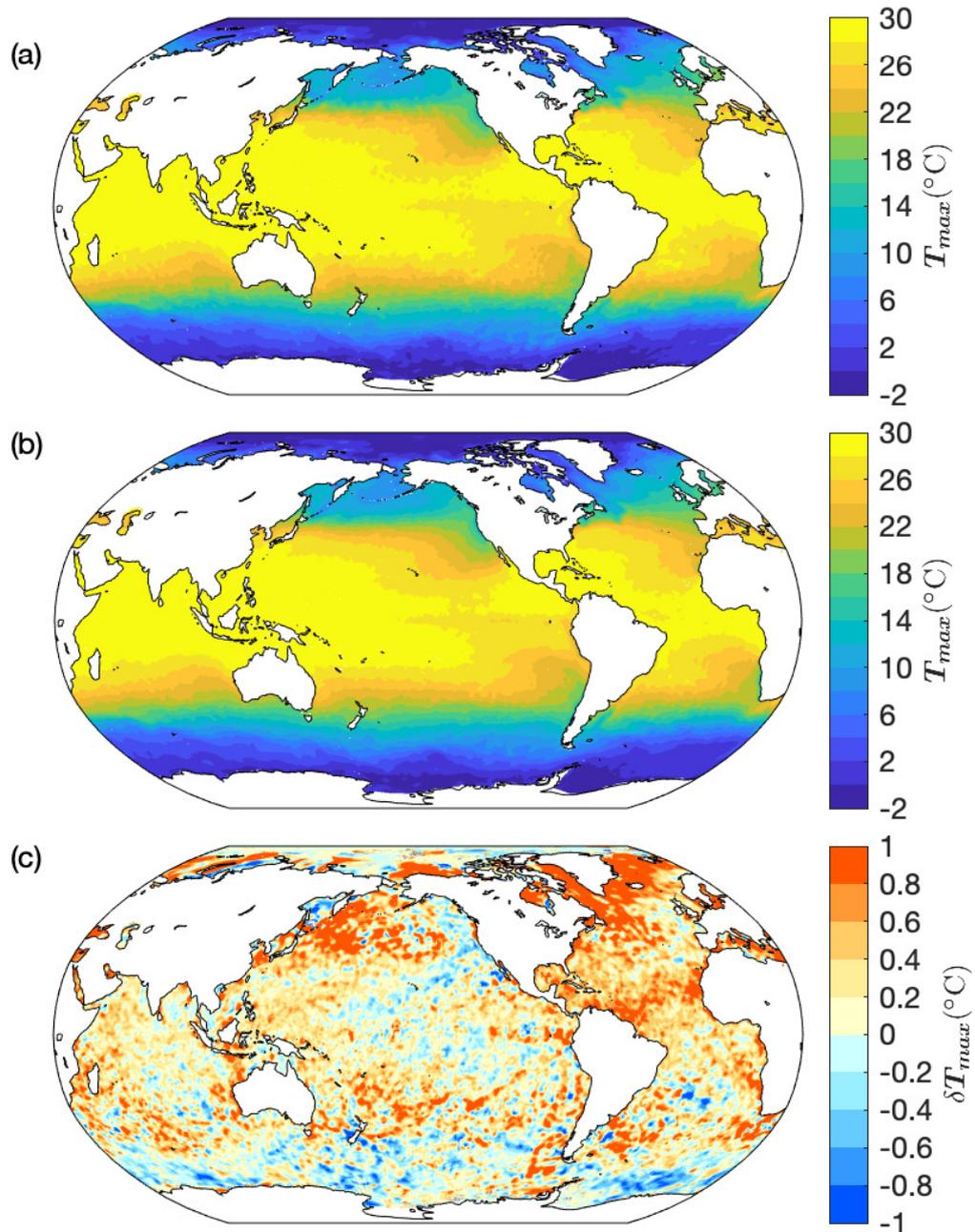
Figure 2.15: (a) T_{max} at sea surface from the WOA13 climatology for 2005-2012 (b) T_{max} at sea surface from the WOA13 climatology for 1955-2012 (c) Difference between T_{max} for 2005-2012 and 1955-2012, which is (a) minus (b).

short time periods.

The difference between T_{max} for climatologies of 1981-2010 and 1955-2017 (Fig. 2.16b) is generally less than 0.2°C, which means their different time periods do not make a big difference in T_{max}. That might due to less significant warming in 1981-2010 compared to 2005-2017. Considering both the time consistent with other observations and the representation of "climate normal", 1981-2010 was chosen for WOA18.
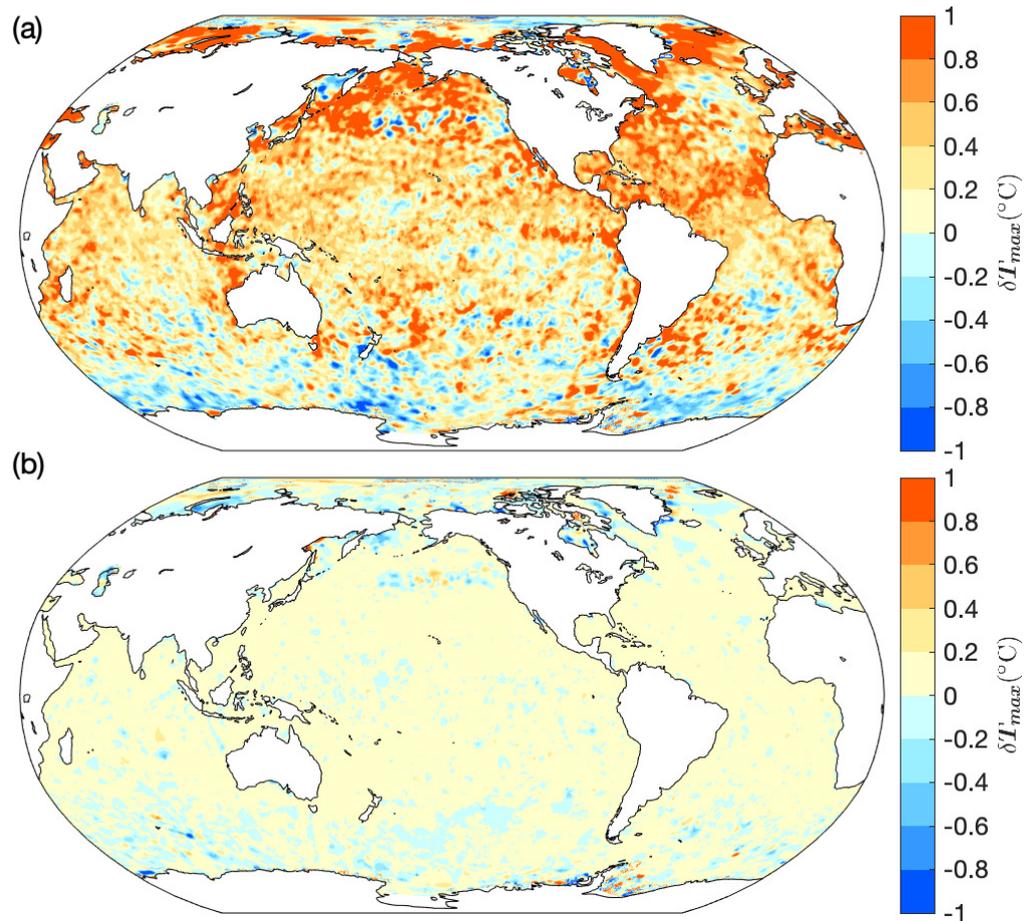
Figure 2.16: (a) Difference between T_{max} at sea surface from the WOA18 climatology for 2005-2017 and 1955-2017, (b) Difference between T_{max} for 1981-2010 and 1955-2017.

2.4.2   Seasonal extrema of 5 observational climatologies

Sea surface temperature

Fig. 2.17 shows T_{max} maps at the sea surface from the five climatologies and the corresponding month of T_{max}. High T_{max} in the tropics is due to net heating, and low T_{max} at high latitudess is due to net cooling. The T_{max} at surface on the western side is generally warmer than the T_{max} on the eastern side because of the boundary currents. The western boundary current carries warm water from lower latitudes to higher latitudes, while the eastern boundary current carries cold water from higher latitudes to lower latitudes. Upwelling in the eastern boundary current regions can also contribute to the lower temperature there. The cold tongue in the eastern equatorial Pacific can be explained by upwelling due to the divergence of Ekman transport associated with trade winds.
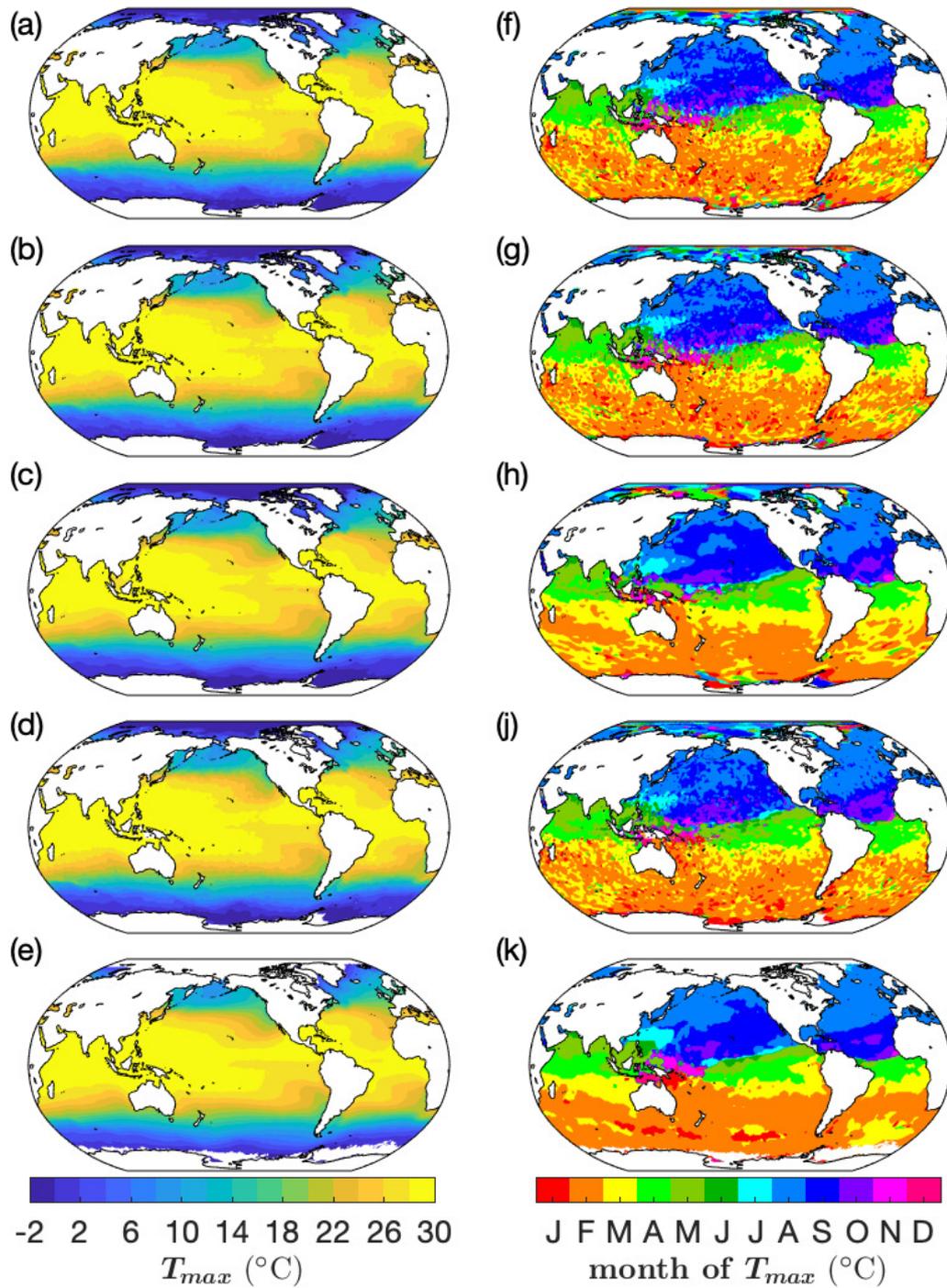
Figure 2.17: (a-e) T_{max} and (f-k) timing of T_{max} at sea surface in 5 observational climatologies: (a, f) WOA13, (b, g) WOA18, (c, h) MIMOC, (d, j) WAGHC and (e, k) HadISST.

T_{max} at the surface occurs at the end of the warming season. SST mainly peaks in August or September in the northern hemisphere, while in the southern hemisphere it mainly peaks in February or March (Fig. 2.17). In areas where SST cannot be well represented by a sinusoidal annual cycle (poor GOF) (Fig. 2.10c), the timing of T_{max} is more varied (Fig. 2.17). In the equatorial region and northern Indian Ocean, timing of T_{max} varies from March to July. In the north of Australia, T_{max} occurs in November, December and January. In the Indian Ocean, the timing of T_{max} changes from February to June further north until the Arabian Sea.

As seen in Figs. 2.17g, 2.10b, the timing of T_{max} is one month earlier than the phase of SST annual cycle in many regions at mid-high latitudes. However, the real timing difference between the max-min method and the sinusoid method should be half a month instead of one month, and this half a month systematic difference might be due to the limited temporal resolution of SST data. Taking two SST monthly time series in those regions (Figs. 2.8a-b) as examples, we found that the phase of SST annual cycle is early October while the timing of T_{max} is mid-September (ticks on X-axes represent the middle of each month). Because of the one-month interval of the color bar, the one-month difference exists in Figs. 2.17g, 2.10b. If the temporal resolution of SST data can be improved from one-month to half a month, there will be no systematic timing differences between the two methods, because the phase of SST annual cycle and the timing of T_{max} will both be early October in Figs. 2.8a-b.

Fig. 2.18 shows the maps of T_{min}. The equatorial cold tongue is obvious in the Pacific, extending westward from the eastern boundary. There are sharp fronts with Antarctic Circumpolar Current (ACC) and western boundary currents (Gulf Stream, Kuroshio current and Brazil current). In most of the northern hemisphere, T_{min} at surface occurs in February or March. In the tropical Pacific and Indian Ocean, the timing of T_{min} can be as early as January and December. In the southern hemisphere, T_{min} occurs mainly in August or September. At the cold tongue in the equatorial Pacific, timing of T_{min} is different from the surrounding regions due to the upwelling. That is because below the surface the timing of seasonal extrema are delayed as much as two months relative to the surface (Talley, 2011).
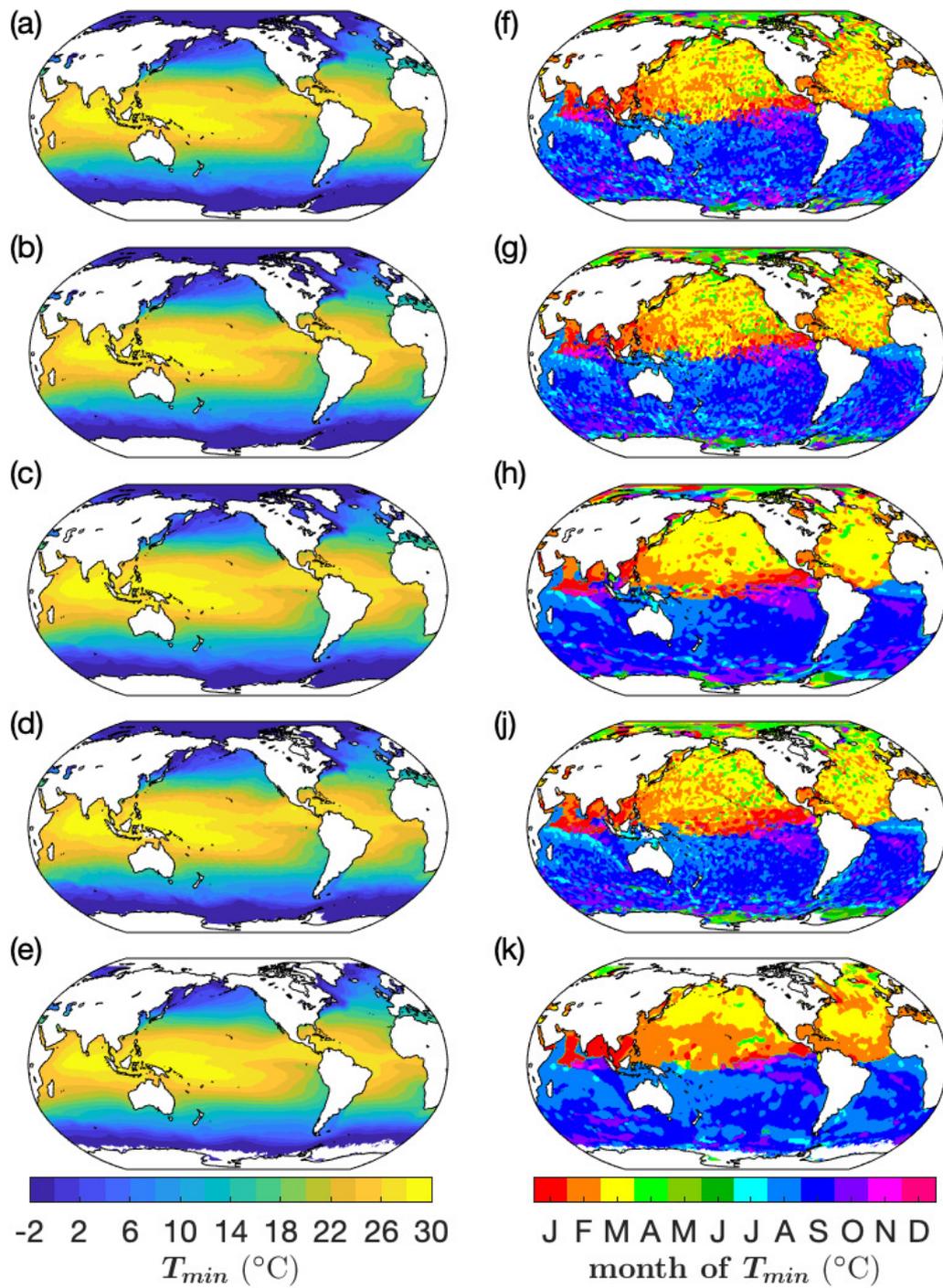
Figure 2.18: (a-e) T_{min} and (f-k) timing of T_{min} at sea surface in 5 observational climatologies (WOA13, WOA18, MIMOC, WAGHC and HadISST).

The time series of SST can be unrealistic from the climatologies in the polar regions, here two points were taken in the Weddell Sea as examples. At the point near the tip of Antarctic Peninsula (Fig. 2.19a), SST from WOA18 and WOA13 peaks in July (the austral winter) and SST in July from WOA18 and WOA13 is over 2 $°C$ warmer than its adjacent months (June and August) (Fig. 2.19b), both of which are very unlikely to occur in reality. From the map of data distribution from WOA18, there are no data in the coastal Weddell Sea in July (Fig. 2.19a). The areas marked in white contain no sampling data to construct the SST field, and thus the values there can be unrealistic. At the more southern point (Fig. 2.19a), some climatologies do not have complete 12 month data which may be related to the existence of sea ice. HadISST only has SST in December-March, while SST in July-December is missing from WAGHC (Fig. 2.19c).

The SST time series in the Weddell Sea (Fig. 2.19b-c) remind us that extra attention should be paid when using data in polar areas from observational climatologies. Therefore, when calculating seasonal extrema of SST and their timings, points without complete 12 months data are eliminated. Because HadISST does not have winter SST data in the polar regions (Fig. 2.20), there is blank space in the maps of T_{max} and T_{min} (Figs. 2.17e, 2.18e).

Sea surface salinity

The spatial distribution of S_{max} (Figs. 2.21a-d) is influenced by precipitation, evaporation, runoff and ice freezing and melting. The salinity of the surface water is at a maximum in subtropical latitudes, where evaporation exceeds precipitation. Due to the difference of evaporation minus precipitation, the North Atlantic is the most saline ocean at the surface, and the North Pacific is the least saline, which explains why deep water formation occurs in the North Atlantic rather than in the North Pacific (Craig et al., 2017). S_{max} in the western Indian Ocean is larger than that in the eastern Indian Ocean, which results from strong evaporation in the Arabian Sea and considerable river run-off in the Bay of Bengal (Talley, 2011). In most areas of the northern hemisphere, SSS peaks in March or April, and in the southern hemisphere peaks in September or October (Figs. 2.21e-h). In the Antarctic S_{max}
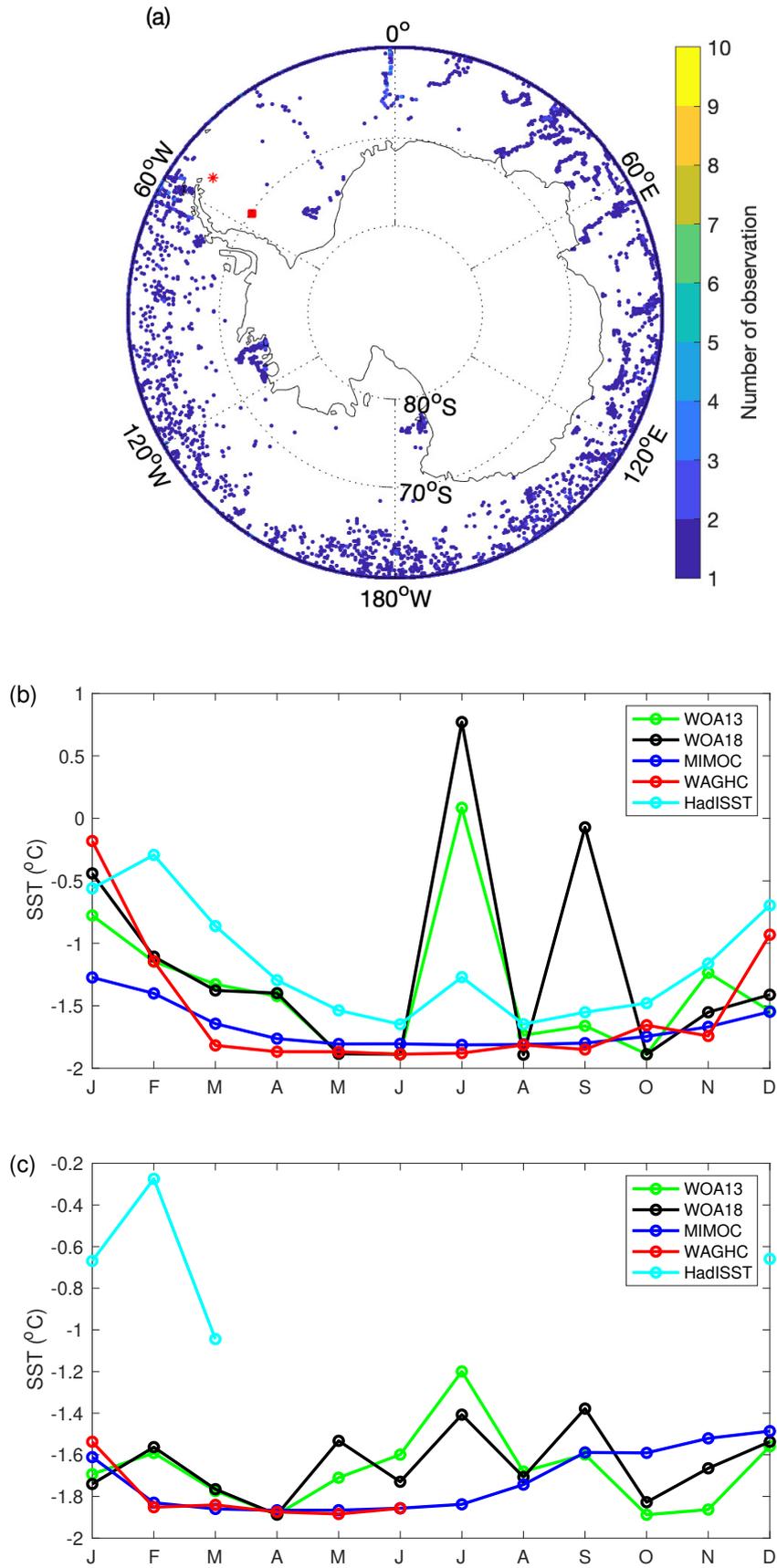
Figure 2.19: (a) WOA18 data distribution of SST in July in the Weddell Sea. The time series of SST from 5 observational climatologies (WOA13, WOA18, MIMOC, WAGHC and HadISST) at (b) the red asterisk and (c) the red square.
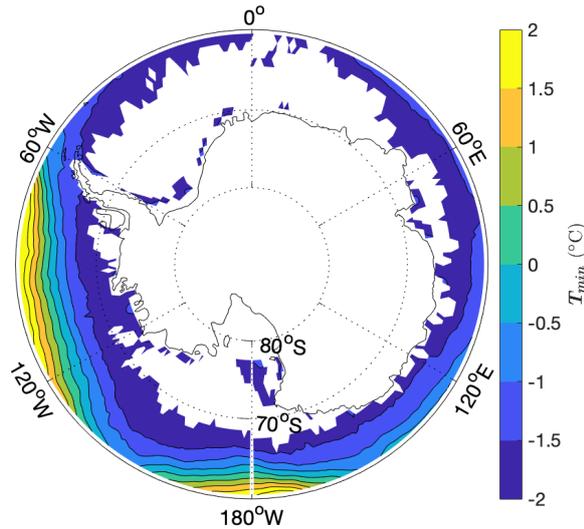
Figure 2.20: T_{min} at sea surface in HadISST in the Antarctic.

occurs during August-November, in the Arctic S_{max} occurs during April-June. In the equatorial Pacific and Atlantic, timing of S_{max} is dominated by April-July when evaporation minus precipitation peaks.

S_{min} has the same global pattern with S_{max} (Figs. 2.21, 2.22), and they are both consistent with the pattern of annual mean evaporation minus precipitation (Trenberth et al., 2007). S_{min} is lower at the ITCZ, which is due to the high precipitation. Like SST, SSS has strong gradient across the ACC and western boundary currents (Gulf Stream, Kuroshio current and Brazil current). S_{min} in the Southern Ocean mainly occurs in January-February when sea ice melts in the austral summer. The timing distribution of S_{max} and S_{min} is not as zonal as T_{max} and T_{min}. It varies in the Pacific and Atlantic Ocean from west to east.

There are differences between climatologies for T_{max}, T_{min}, S_{max}, S_{min} and their timings. The spatial patterns and amplitudes are similar but not identical between climatologies. The data coverage is not complete in some climatologies, for example, there are no data from WAGHC in Hudson Bay and no data from HadISST in Hudson Bay and the polar regions.
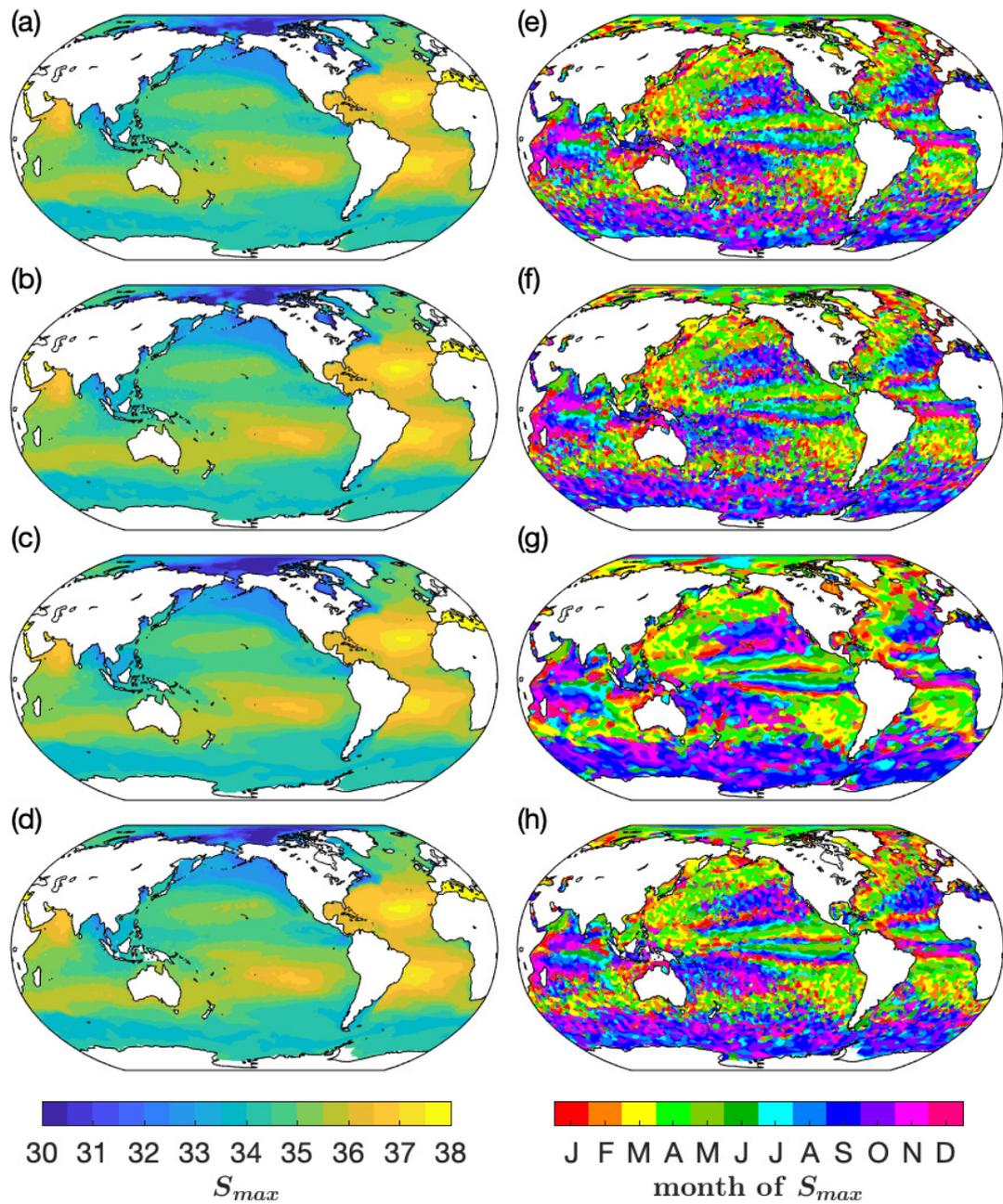
Figure 2.21: (a-d) S_{max} and (e-h) timing of S_{max} at sea surface in (a,e) WOA13, (b,f) WOA18, (c,g) MIMOC and (d,h) WAGHC.

2.4.3 Difference of the seasonal extrema between observational climatologies

Sea surface temperature

The maximum difference between any two of the five observational climatologies (WOA13, WOA18, MIMOC, HadISST and WAGHC) for T_{max} and T_{min} is typically
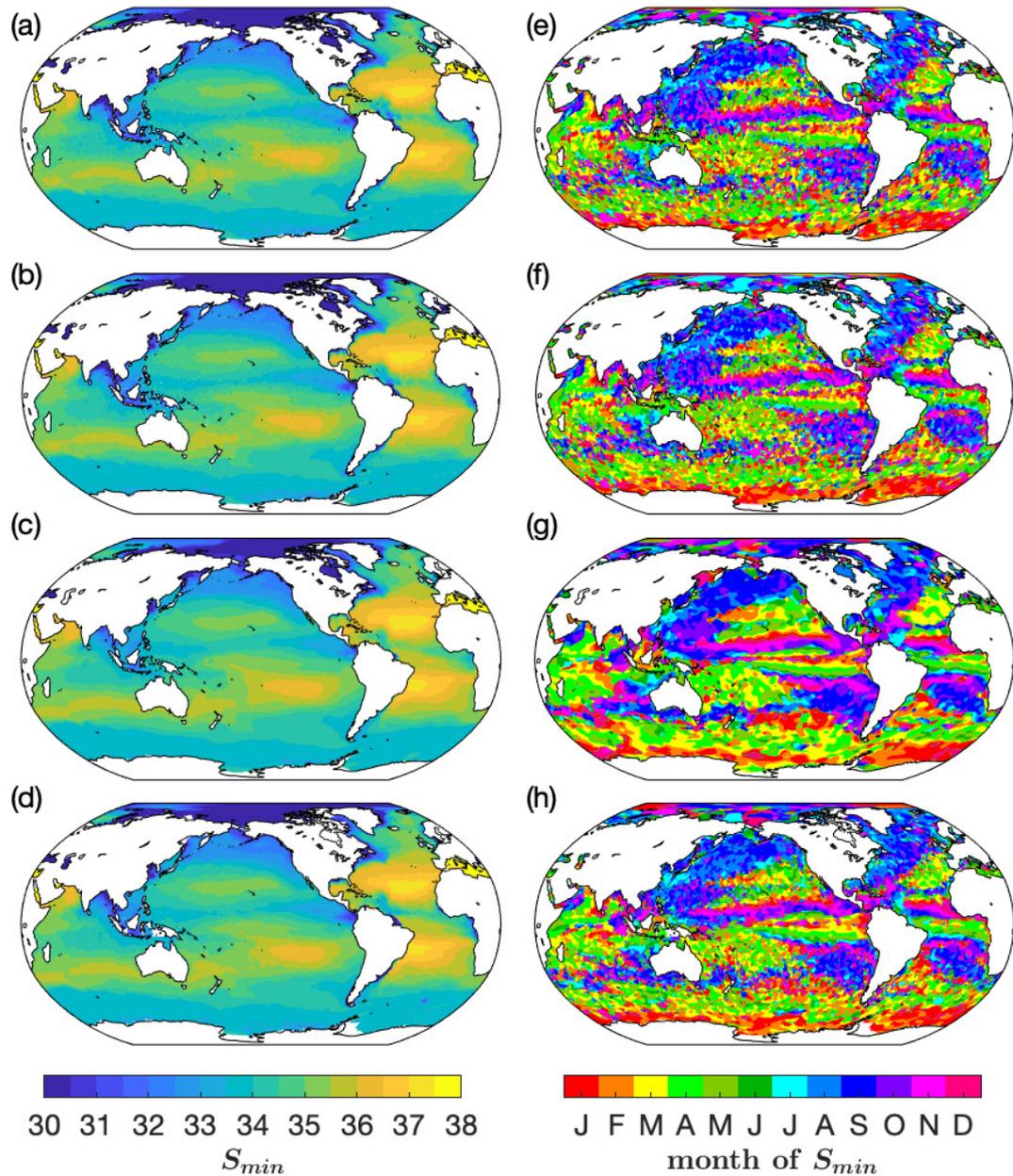
Figure 2.22: (a-d) S_{min} and (e-h) timing of S_{min} at sea surface in (a,e) WOA13, (b,f) WOA18, (c,g) MIMOC and (d,h) WAGHC.

about 1°C (Fig. 2.23). The maximum difference in T_{max} is larger in the northern hemisphere than in the southern hemisphere, and the difference of more than 2°C occurs in the Arctic, North Pacific and North Atlantic. One possible reason for these differences is the different time periods between climatologies. Consistent with the larger T_{max} difference between climatologies in the northern hemisphere, the warming trend of T_{max} in WOA13 is also stronger in the northern hemisphere than in the southern hemisphere (Fig. 2.15). The large T_{max} difference in the Arctic may be due

to limited observations there (Fig. 2.19). The maximum difference between climatologies in the Arctic (except for the Russia coast) is smaller in T_{min} than in T_{max} (Fig. 2.23a-b), which may be explained that T_{min} there in all the climatologies is close to the freezing point (Fig. 2.18). The five observational climatologies have larger differences in T_{cycle} than in T_{max} and T_{min} (Fig. 2.23), because the differences in T_{cycle} can be the sum of the differences in T_{max} and T_{min}.

In regions with strong temperature gradient, both coarse horizontal resolution and averaging on isobars can smooth the gradient. The areas with large differences between climatologies (e.g. Gulf Stream, Brazil Current and ACC) (Fig. 2.23) are consistent with the areas with strong SST gradient (Fig. 2.24), which may indicate that the differences between 5 observational climatologies are caused by the different horizontal resolutions.

To avoid the differences from different time periods, we chose the three climatologies: WOA18, HadISST and WAGHC and calculated their differences for T_{max}, T_{min} and T_{cycle} (Fig. 2.25). These climatologies have time periods of 1981-2010 or close to 1981-2010 (table. 2.1), whereas MIMOC only reflects the 2007-2011 modern ocean state and WOA13 is an old version of WOA18 with a time period of 1955-2012. Therefore, we decided to use only WOA18, HadISST and WAGHC to study the uncertainty of SST in observational climatologies.

The maximum differences between the three climatologies (WOA18, HadISST and WAGHC) (Fig. 2.25) are smaller than between the 5 climatologies (Fig. 2.23), but there are still large differences in the coastal regions, Weddell Sea, Hudson Bay, ACC, Gulf stream and Humboldt Current. Here, we set two requirements to find regions with high uncertainty. The first requirement is that the maximum difference in T_{max} or T_{min} between the three climatologies is larger than 2°C (2°C contribute to about 0.5 kg/m^3 density); the second requirement is that the values for all 12 months are not complete. Grid points meet either of the two requirements are considered uncertain (Figs. 2.25a-b). All the uncertain grid points for T_{max} and T_{min} are masked by red dots in Fig. 2.25a-b. For T_{cycle}, the grid points are considered uncertain when they are uncertain in either T_{max} or T_{min}, and hence the mask for T_{cycle} is the
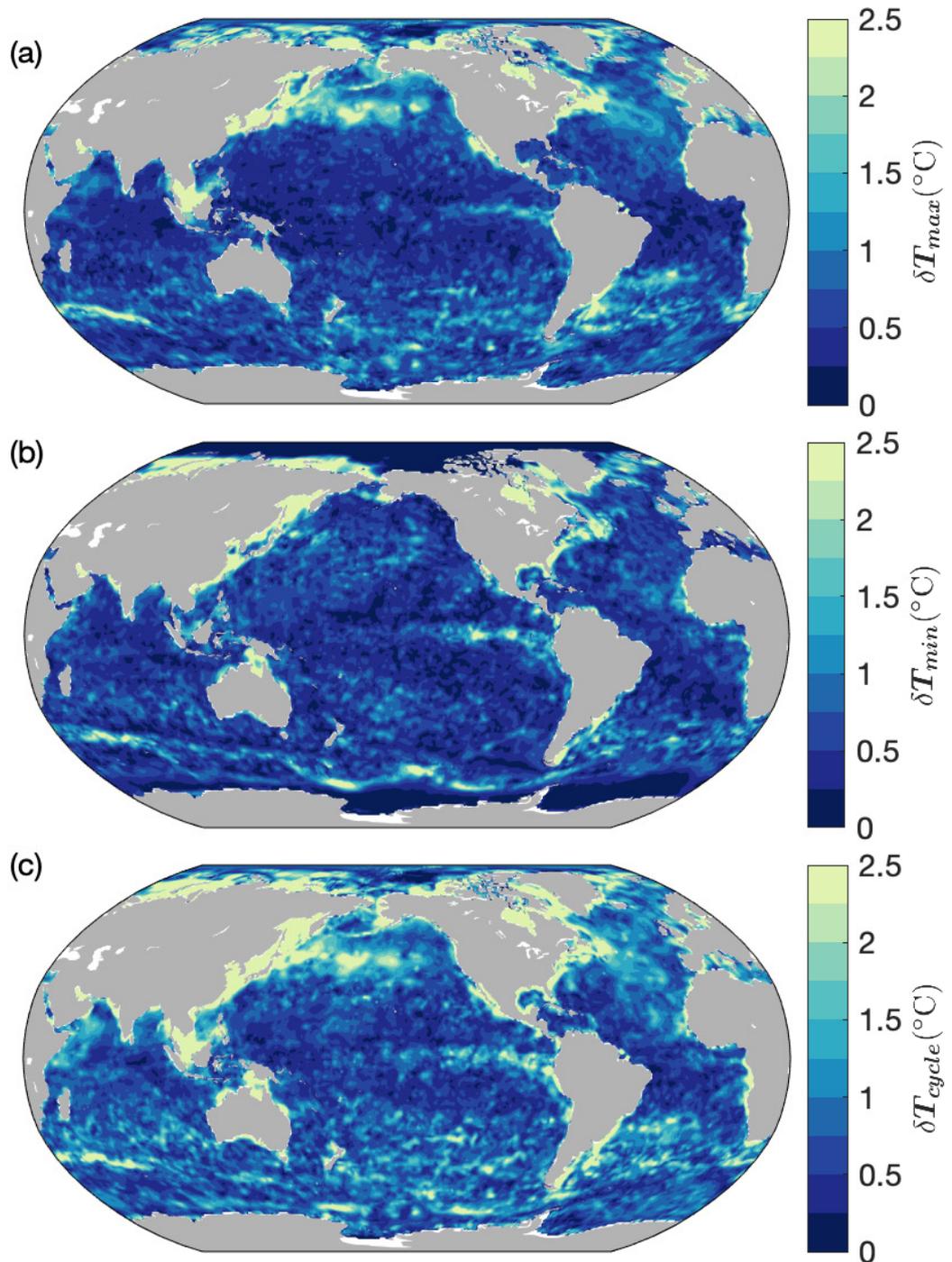
Figure 2.23: The maximum difference between any two of the five observational climatologies (WOA13, WOA18, MIMOC, HadISST and WAGHC) for (a) T_{max}, (b) T_{min} and (c) T_{cycle} at surface.
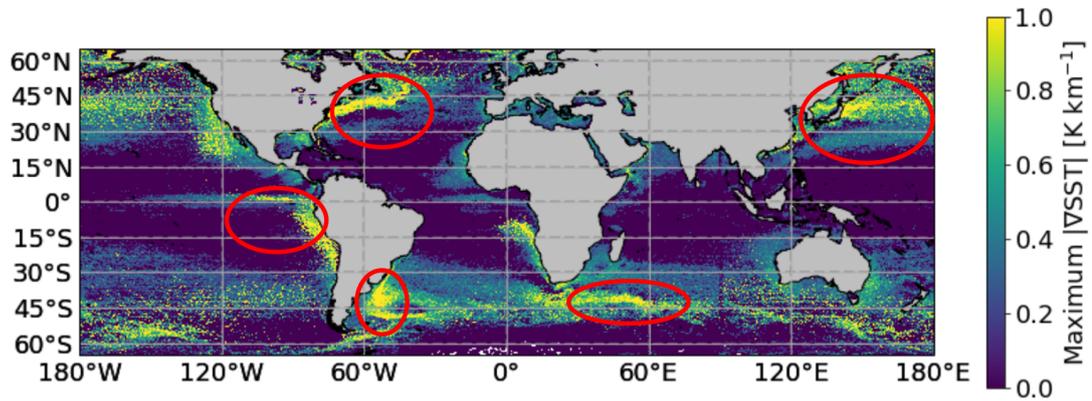
Figure 2.24: Maximum of the annual mean maximum spatial gradient of SST. Red circles indicate areas with large SST gradient. Adapted from González-Haro et al. (2019).

.

combination of masks in T_{max} and T_{min} (Fig. 2.25). All these grid points with large uncertainty will be excluded from the following model assessment. 4%, 3% and 4% of the ocean's surface area is excluded for T_{max}, T_{min}, and T_{cycle} respectively.

The observational climatologies with the maximum and minimum SST values are shown in Fig. 2.26. It can be seen that in most of the world ocean WOA18 has the coldest T_{max} and T_{min}. HadISST has the warmest T_{max} and T_{min}.

Sea surface salinity

The amplitude of salinity difference between WOA18 and WAGHC is shown in Fig. 2.27. The color bar for salinity is comparable to the color bar for temperature in Figs. 2.23, 2.25, considering their contributions to density of sea water. Here we only compare salinity in WOA18 and WAGHC but not in MIMOC and WOA13, because the time periods of MIMOC and WOA13 are inconsistent with other climatologies. In most of the ocean, the maximum salinity difference is below 0.5, while in some areas of the Arctic it is larger than 0.5. Similar to the masks for T_{max} and T_{min}, there are also masks for S_{max} and S_{min}. Grid points where the maximum differences in S_{max} or S_{min} between the two climatologies (WOA18 and WAGHC) is larger than 0.5, and grid points which do not have values for all 12 months in both climatologies are excluded for the following model assessment. Here, 0.5 is used as the threshold because the contribution of 0.5 salinity to density is comparable to that of 2°C temperature,
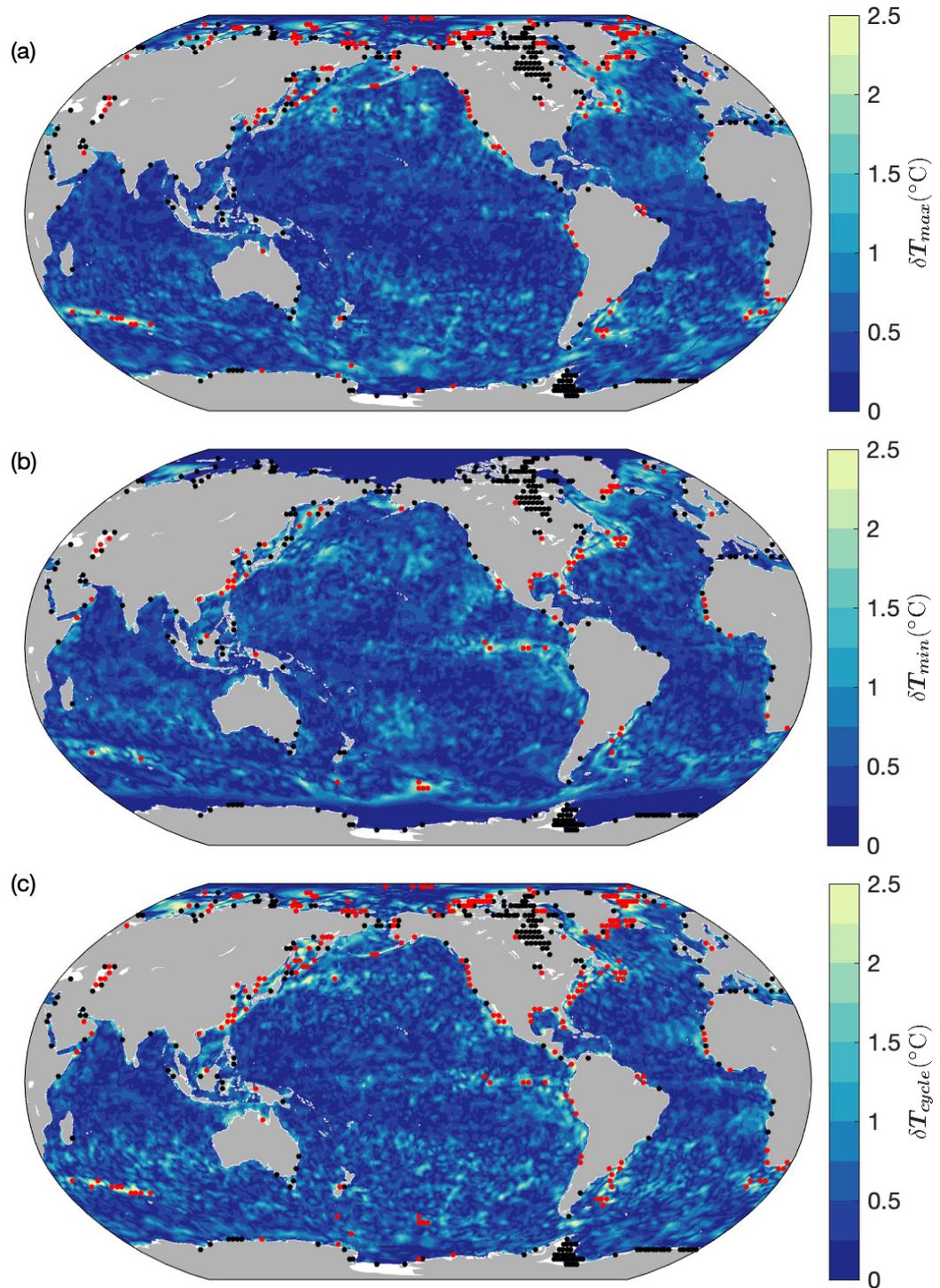
Figure 2.25: The maximum difference between any two of WOA18, HadISST and WAGHC for (a) T_{max}, (b) T_{min} and (c) T_{cycle} at surface. Red dots indicate where the maximum differences are larger than 2°C. Black dots indicate where there is no values for all 12 months for at least two climatologies.

.
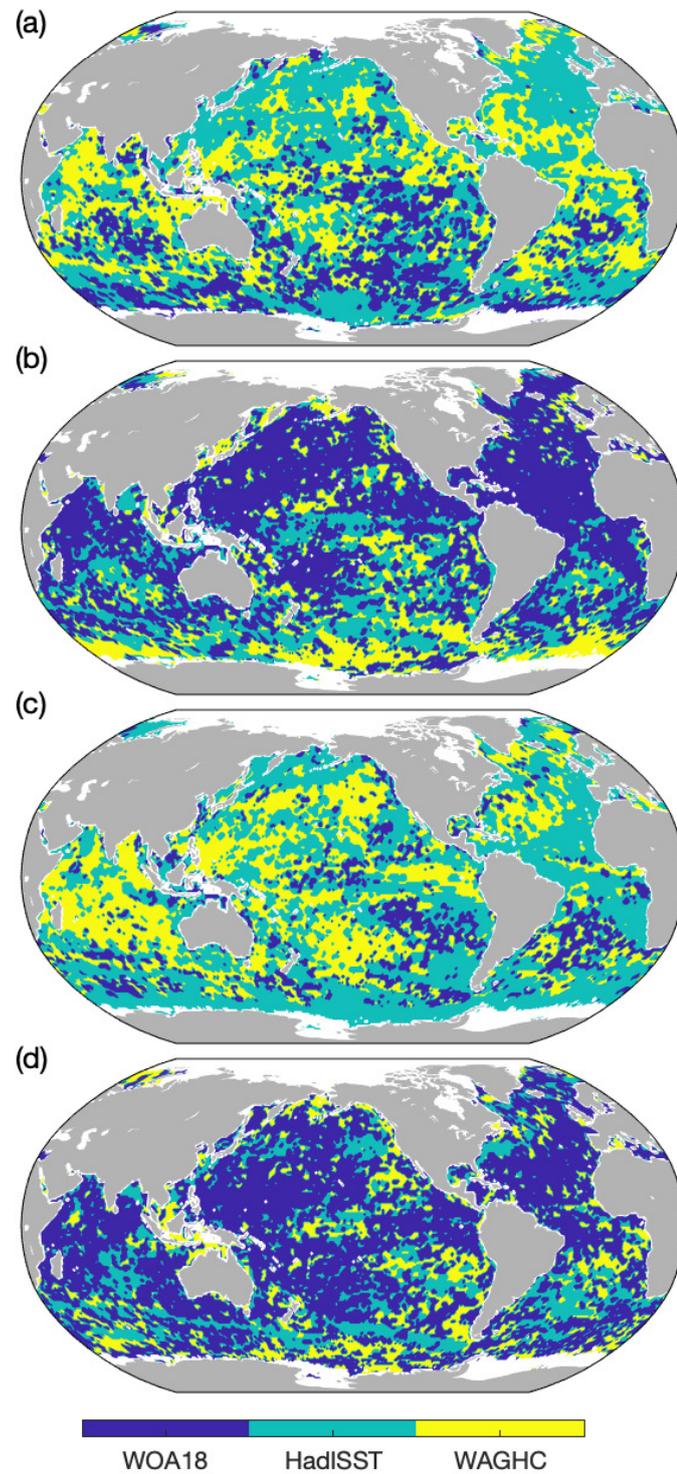
Figure 2.26: The observation product among WOA18, HadISST and WAGHC with (a) maximum value for T_{max}, (b) minimum value for T_{max} (c) maximum value for T_{min} and (d) minimum value for T_{min} at surface.

which was used as the threshold for the SST uncertainty mask. The difference for S_{min} between the two observational climatologies is larger than that for S_{max}, especially in some coastal areas, such as the Maritime Continent, Caribbean, Bay of Bengal and Gulf of Guinea.
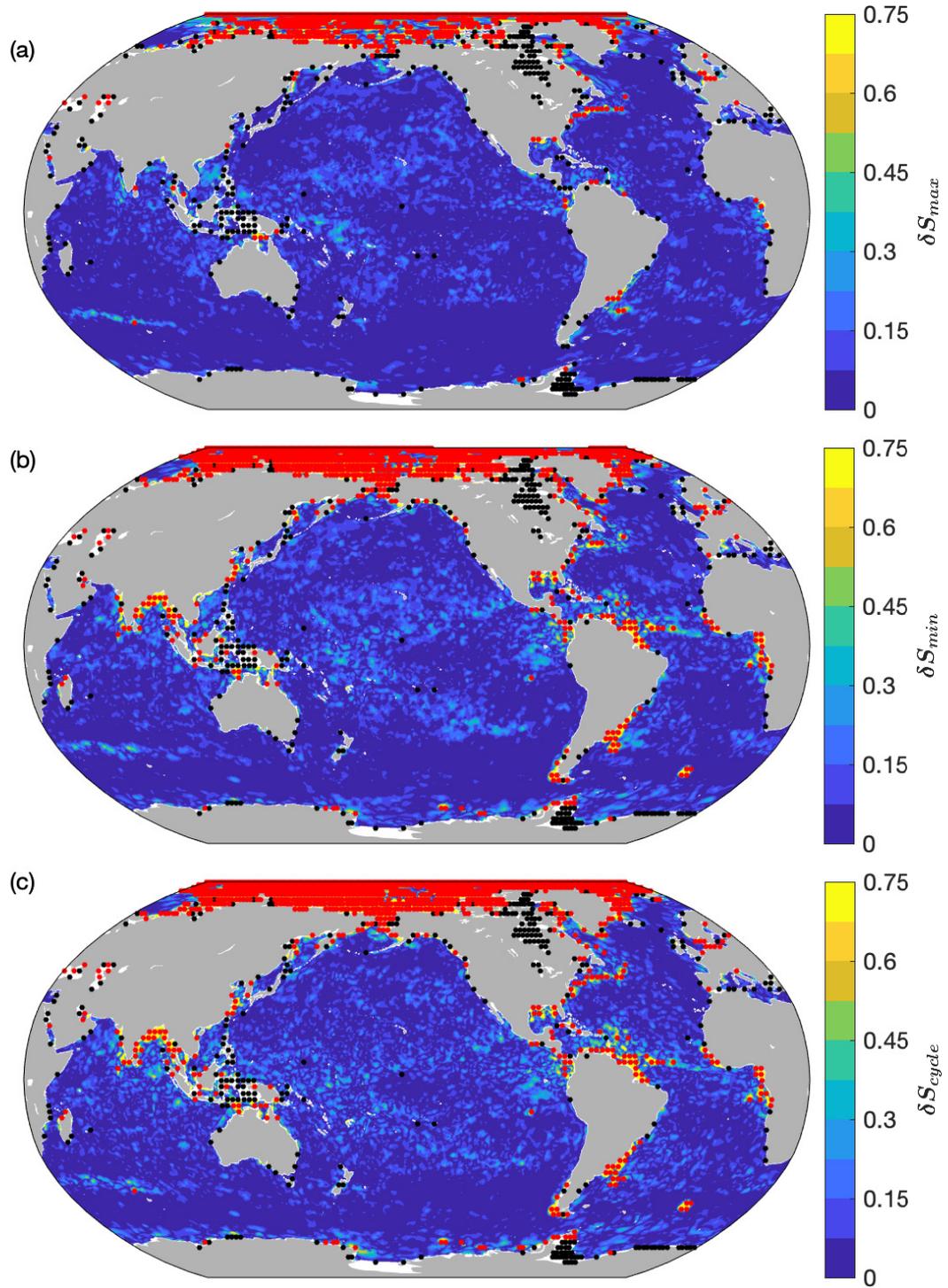
Figure 2.27: The amplitude of differences between WOA18 and WAGHC for (a) S_{max} and (b) S_{min}. Red dots indicate where the maximum differences are larger than 0.5. Black dots indicate where there are no values for all 12 months for the two climatologies.

In the Arctic, the difference of SSS between WOA18 and WAGHC is more than 5 in some areas (Fig. 2.28). The difference is much larger in S_{min} than in S_{max}. The large salinity difference between the two climatologies in the Arctic is likely to due to the unrealistic salinity in WOA18. As an old version of WOA18, WOA13 produced unrealistically high salinities exceeding 36 and much lower salinities compared with WAGHC in the Arctic Ocean (Gouretski, 2018). Gouretski (2018) attributed part of the salinity difference between WOA13 and WAGHC to the much poorer WOA13 data base in the Arctic compared with WAGHC.
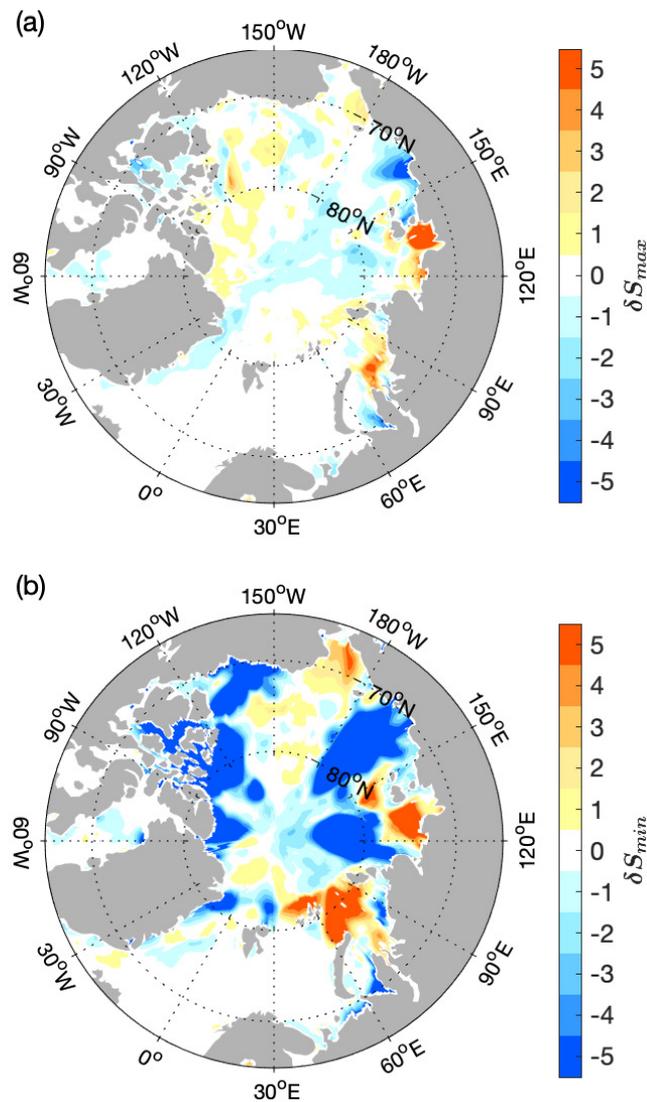
Figure 2.28: The difference between WOA18 and WAGHC (WOA18-WAGHC) for (a) S_{max} and (b) S_{min}.

Similar to Fig. 2.26, Fig. 2.29 shows the climatologies with maximum or minimum values for S_{max} and S_{min}. For most of the Atlantic Ocean and equatorial Pacific,

SSS is saltier in WAGHC than in WOA18 for both S_{max} and S_{min}. For most of the North Pacific, SSS is saltier in WOA18 than in WAGHC for both S_{max} and S_{min}. One possible reason is that WAGHC and WOA18 cover different time periods (see table. 2.1), and there is SSS changes in response to global change (Durack et al., 2012). Based on WOA01, Durack et al. (2012) revealed S_{mean} change for 1950-2000 (Fig. 2.30). The regions where WAGHC has saltier S_{max} and S_{min} (Fig. 2.29) are consistent with the regions with intensifying SSS over 1950-2000 (Fig. 2.30).
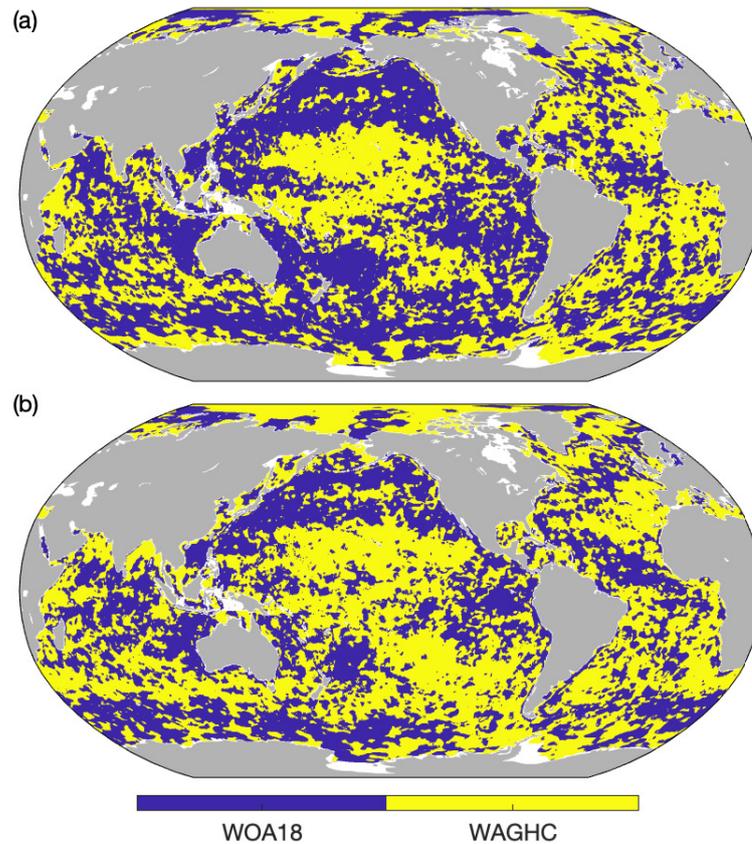
Figure 2.29: The climatologies between WOA18 and WAGHC with (a) maximum value for S_{max}, (b) maximum value for S_{min}.

2.4.4 Difference in the timing of seasonal extrema between observational climatologies

The timings for T_{max} and T_{min} are different in the five observational climatologies. The maximum timing differences are less than one month in most of the world ocean, except for the polar regions and the equatorial regions, where the phase differences can be as large as 6 months (Fig. 2.31). The grid points are considered as uncertain (masked in
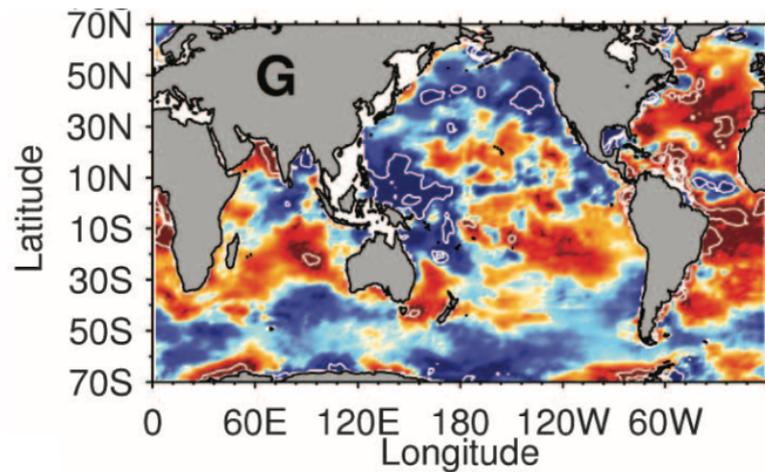
Figure 2.30: Observed S_{mean} change over 1950-2000, adapted from Durack et al. (2012).

Fig. 2.31) when their timing difference is larger than 2 months or the values for all 12 months are not complete for at least two climatologies.

In the polar regions, the large timing differences may be due to data scarcity there. Lack of observations means differences in gridding and/or interpolation method may produce disproportionately large changes in calculated values compared with regions which are better observed. In addition, T_{min} is approximately constant for several months, thus any small change in calculated T_{min} (due to differences in gridding, interpolation method, or inclusion of different datasets) may lead to large changes in the month of T_{min}. In the equatorial regions, there is no strong driver of seasonality and thus little difference between T_{max} and T_{min}. Any small changes (due to differences in gridding, interpolation method, or inclusion of different datasets) may alter the timing considerably, which may explain the large timing differences in the equatorial regions.

The timing differences for S_{max} and S_{min} between the climatologies are much larger than those for T_{max} and T_{min}, and the patterns of the timing differences for SSS (Fig. 2.32) are much more complicated than for SST (Fig. 2.31). Most of the world ocean is covered by the uncertainty mask, which indicates that the timing of the seasonal cycle of SSS from climatologies is highly uncertain. The few regions not covered by uncertainty mask are mostly consistent with regions where there is a large standard deviation (more than 0.4) of monthly mean SSS in Yu et al. (2021), suggesting that the climatologies are likely unable to capture the small seasonal variability of SSS.
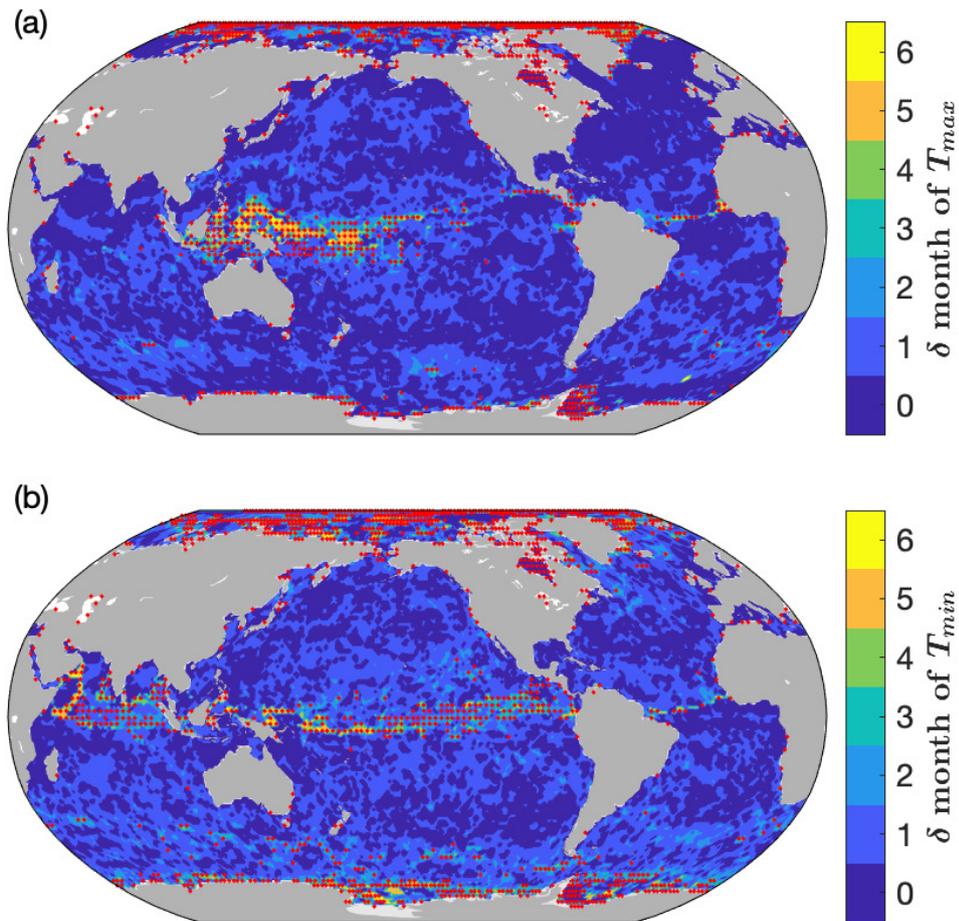
Figure 2.31: The maximum difference between any two of WOA18, HadISST and WAGHC for (a) timing of T_{max} and (b) timing of T_{min}. The dots indicate area where the maximum differences are larger than 2 months.

2.5 Conclusion

This chapter is a preparation for the model evaluation in the following two chapters. In this chapter, we compared two different methods to define seasonal extrema: fitting a sinusoidal annual cycle and finding the monthly maximum and minimum values. Fitting a sinusoidal annual cycle cannot be applied everywhere in the world ocean, as some regions are not dominated by an annual sinusoid (for example, the polar regions and monsoon regions). However, the monthly maximum and minimum values can always be defined in the world ocean, even in regions with non-sinusiodal annual cycle. Therefore, in this thesis seasonal extrema were defined as the maximum and minimum
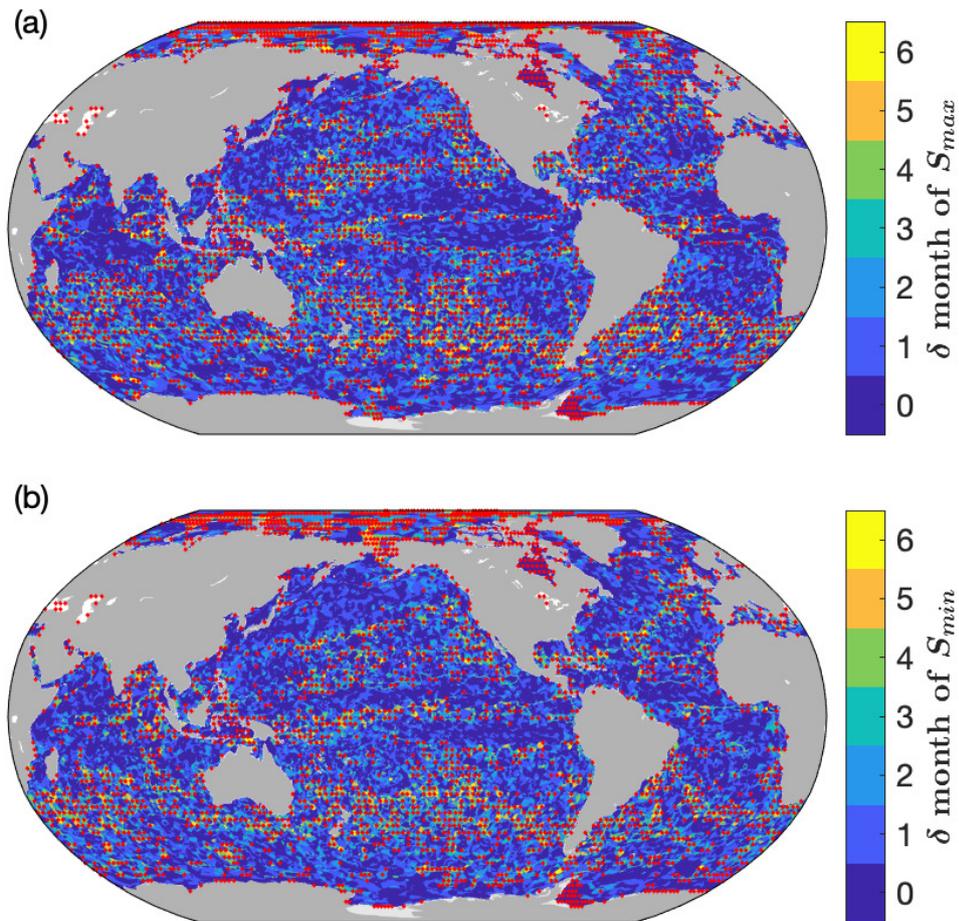
Figure 2.32: The maximum difference between WOA18 and WAGHC for (a) timing of S_{max} and (b) timing of S_{min}. The dots indicate area where the maximum differences are larger than 2 months.

values from the monthly climatology, whichever months they occur in.

Five observational climatologies (WOA13, WOA18, MIMOC, WAGHC and HadISST) were compared in this chapter. By calculating the maximum differences between the climatologies, the regions with large differences were found. For those regions, extra attention should be paid when choosing a climatology for model evaluation. The uncertainty of the climatological data can affect the model evaluation when a climatology is chosen as a reference to assess model output. In these areas with large difference between climatologies, it is hard to say whether the difference between model and climatology represents a real bias in the model, or it is just from the uncertainty of the climatology.

For T_{max} and T_{min}, there are more than 2°C differences between climatologies in regions with strong SST horizontal gradient (e.g. Gulf Stream, Kuroshio Extension, Brazil Current, ACC and coastal upwelling regions), which may be due to different horizontal resolution in climatologies. Large differences also exist in the polar regions, most likely because of limited observation in ice-covered regions. In the Antarctic, some climatologies do not have values for all 12 months or have unrealistic SST values in austral winter.

To exclude the regions with high uncertainty, masks of high uncertainty were suggested, which cover regions with large differences (more than 2°C for SST and 0.5 for SSS) between climatologies or do not have values for all 12 months. The masks will be used to exclude grid points with high uncertainty for model assessment in the following two chapters.

In this thesis, we choose WOA18 with high uncertainty regions excluded as a standard for the following model evaluations. World Ocean Atlas is a widely used observation dataset. WOA13 was used for climate studies, ocean model initialization and model evaluation in CMIP5. As the newest version of World Ocean Atlas, WOA18 has a fine horizontal resolution ($0.25° \times 0.25°$) and a time period of 1981-2010 which represents the "climate normal" well. WOA18 is also the latest released product and with the largest number of profiles, among the five climatologies. In the next chapter, the seasonal cycle of sea surface water characteristics in HiGEM will be assessed against the WOA18 climatology, and the masks will be used to exclude regions of high uncertainty.

3

Seasonal cycle of sea surface water characteristics in HiGEM

3.1 Introduction

SST and SSS set the water mass properties below the surface when subduction brings surface water into the ocean interior in winter. To have accurate projection of intermediate and deep water, it is necessary for climate models to deliver accurate projections of seasonal cycle of SST and SSS. This chapter will discuss SST, SSS, sea surface density (SSD) and MLD together, as SST and SSS together determine SSD, and then SSD together with vertical stratification determine MLD (that is the depth to which surface water will sink).

When starting the model evaluation work, the CMIP6 model output was not available, and hence here we start by comparing WOA18 to a CMIP5 model - HiGEM. HiGEM has a horizontal resolution of 0.83° latitude × 1.25° longitude in the atmosphere, and 1/3° × 1/3° in the ocean (Shaffrey et al., 2009). The high horizontal resolution in the ocean means that eddies can be represented (at least partially). The Gent and McWilliams parameterization (Gent and Mcwilliams, 1990) is not used in HiGEM because it could cause low eddy variability and front erosion in the model. The ocean model has 40 unevenly spaced levels in the vertical with enhanced resolution near the surface, which enables the mixed layer and air-sea interaction processes to be better resolved. The evaluation of HiGEM in Shaffrey et al. (2009) is about its control run. The ocean initial conditions used the potential temperature and salinity in September from the 1/4° World Ocean Atlas 2001 (WOA2001). In Shaffrey et al. (2009), HiGEM were

run for 70 years from the initial conditions and the spin-up took 20 years. In this chapter, we will use the HiGEM historical run over 1981-2010 and do the evaluation by comparing with WOA18 observational climatology averaged over 1981-2010.

In this chapter, HiGEM will be used as an example to develop techniques for further evaluation of CMIP6 models. The evaluation of HiGEM in this chapter is a feasibility study before we start assessing SST seasonal extrema in CMIP6 models in Chapter 4. Here, we choose HiGEM because it is a high resolution model and it was readily available and well documented (Shaffrey et al. (2009) had assessed the mean state of SST and SSS in HiGEM). We will evaluate the seasonal extrema of SST and SSS, and their timings in HiGEM, from a global perspective.

We will then focus on specific regions and investigate monthly time series of SST, SSS, SSD and MLD for individual points which are chosen as representative of the specific regions. The points A-N include the regions where there are significant SST or SSS biases and some of them are related to water mass formation. The point A is linked to NPIW formation; the point I is linked to Subantarctic Mode Water; the points L-M are linked to Antarctic Bottom Water (AABW); the point K is linked to North Atlantic Deep Water (NADW); the point N is linked to Antarctic Intermediate Water (AAIW). There are some water mass formation regions missing from points A-M, but there are no significant SST or SSS biases in those regions (e.g. Labrador Sea where NADW forms, Mediterranean Sea where Mediterranean Deep water forms). Except performing analysis for individual points, we also did monthly time series for $10° \times 10°$ boxes. The time series for individual points and $10° \times 10°$ boxes are similar, which indicates that the individual points are representative of larger regions and the monthly time series is not dominated by specific local effects.

3.2 Seasonality of sea surface water - a global perspective

3.2.1 SST and SSS seasonal extrema

Seasonal extrema of SST and SSS are obtained from WOA18 monthly climatology for the time period 1981-2010 as described in chapter 2. T_{max} and T_{min} in HiGEM range

from more than 30°C in the tropics to freezing temperature (about -1.8°C) in the polar regions where ice forms (Figs. 3.1a, 3.1b). The seasonal variability of SST is obvious at mid-high latitudes, especially in the Northern Hemisphere (e.g. Kuroshio and Gulf Stream) where the SST annual range can be larger than 11°C (Fig. 3.1c). T_{cycle} rises from 0-1°C at the equator to over 10°C at mid-high latitudes, then decreases toward the polar regions (Fig. 3.1c).
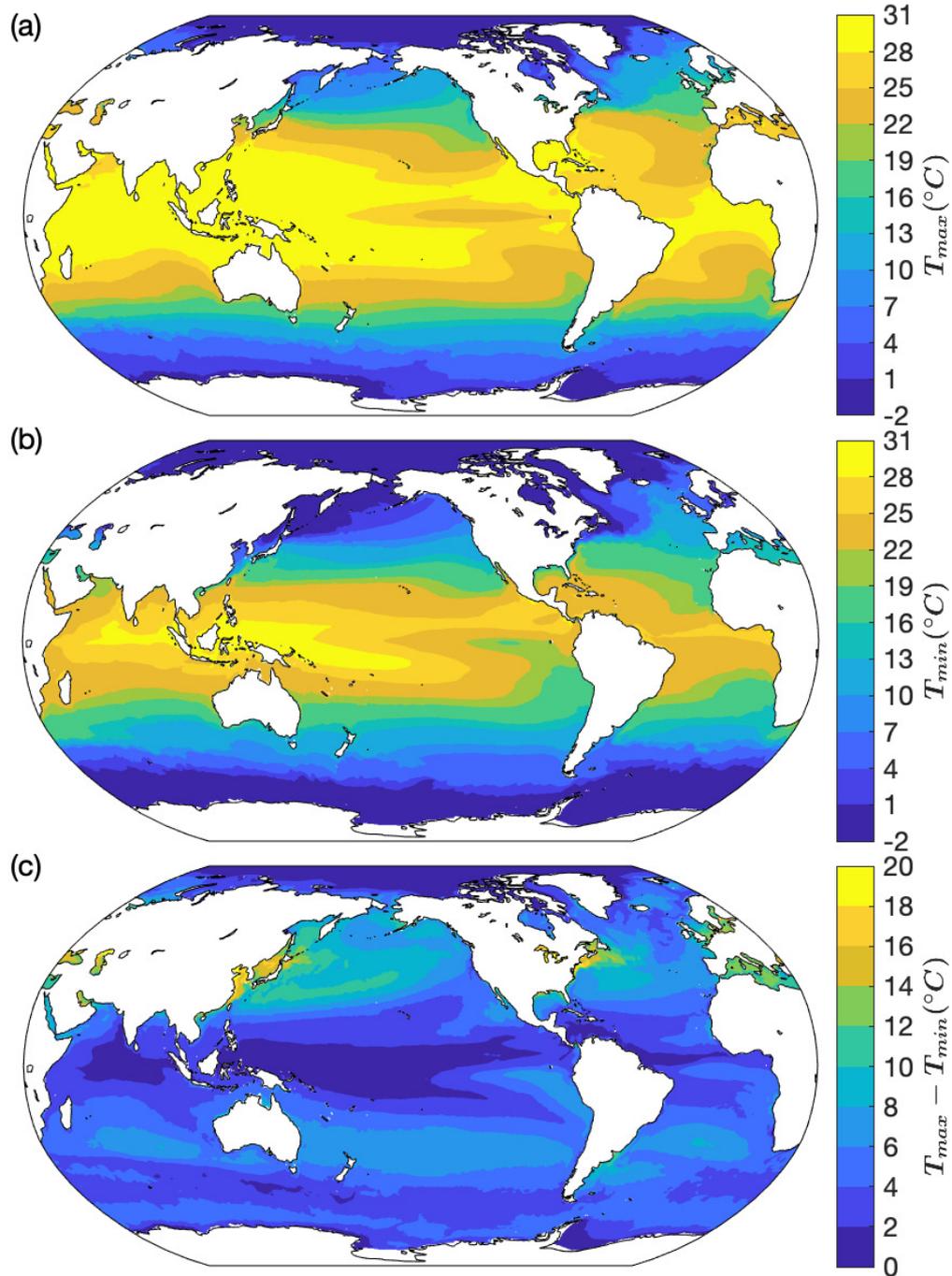
Figure 3.1: (a) T_{max} (b) T_{min} and (c) T_{cycle} in HiGEM.

By comparing T_{max}, T_{min} and T_{cycle} in HiGEM and in WOA18, we obtained their

biases in HiGEM (Fig. 3.2). We noticed that those biases are significant compared to the uncertainty in observation (Fig. 3.2). Specifically, most of the SST bias larger than 0.5°C is robust, not arising from the uncertainty in WOA18.

The spatial patterns of biases are similar in T_{max} and T_{min} (Figs. 3.2a, 3.2b) in some regions, which means some of the bias is seasonally independent, and can therefore be seen as biases in T_{mean} (Fig. 3.3). For the North Pacific, South Atlantic, Arabian Sea and Greenland Sea, although they have obvious bias in T_{max} and T_{min}, this is largely because of biases in T_{mean} (Fig. 3.3); HiGEM is simulating the seasonality reasonably well.

Even in locations where T_{max} and T_{min} biases have broadly similar spatial patterns, the magnitude of biases in T_{max} and T_{min} can be different, so there is still a seasonal component to the bias. For example, the warm bias in the Southern Ocean is up to 4°C larger in T_{max} than in T_{min}, which may be due to biases in cloud cover (Shaffrey et al., 2009). Underestimated cloud can lead to overestimated Q_{sw} in equation 1.2 and thus cause an SST warm bias. In winter, cloud biases make little difference to SST since there is little incoming solar heat flux. In summer, reduced cloud cover over the Southern Ocean leads to enhanced incoming solar heat flux, and thus considerably warmer SSTs.

Different magnitudes of T_{max} and T_{min} biases lead to biases in T_{cycle} (Fig. 3.2c). T_{cycle} is typically 1-2°C larger in HiGEM than in WOA18 over most of the world ocean. One contributing factor is likely to be the overly shallow summer MLD in HiGEM (Fig. 3.4a), which will tend to intensify the increase in SST due to the summer surface heat flux. As indicated by equation 1.1, overly small H will cause too much SST increase. Although biases exist in winter MLD as well (Fig. 3.4b), the overall greater depth of the mixed layer (too large H in equation 1.1) in winter will diffuse the effect of surface heat fluxes (Q_{net} in equation 1.1) over a much greater volume of water, reducing the contribution to any winter SST biases.

The spatial patterns of SSS bias are broadly similar in S_{max} and S_{min} (Figs. 3.5a, 3.5b), and thus are mostly associated with bias in S_{mean} rather than seasonal bias (Fig. 3.6). This SSS bias pattern at low latitudes corresponds well with the precipitation bias in
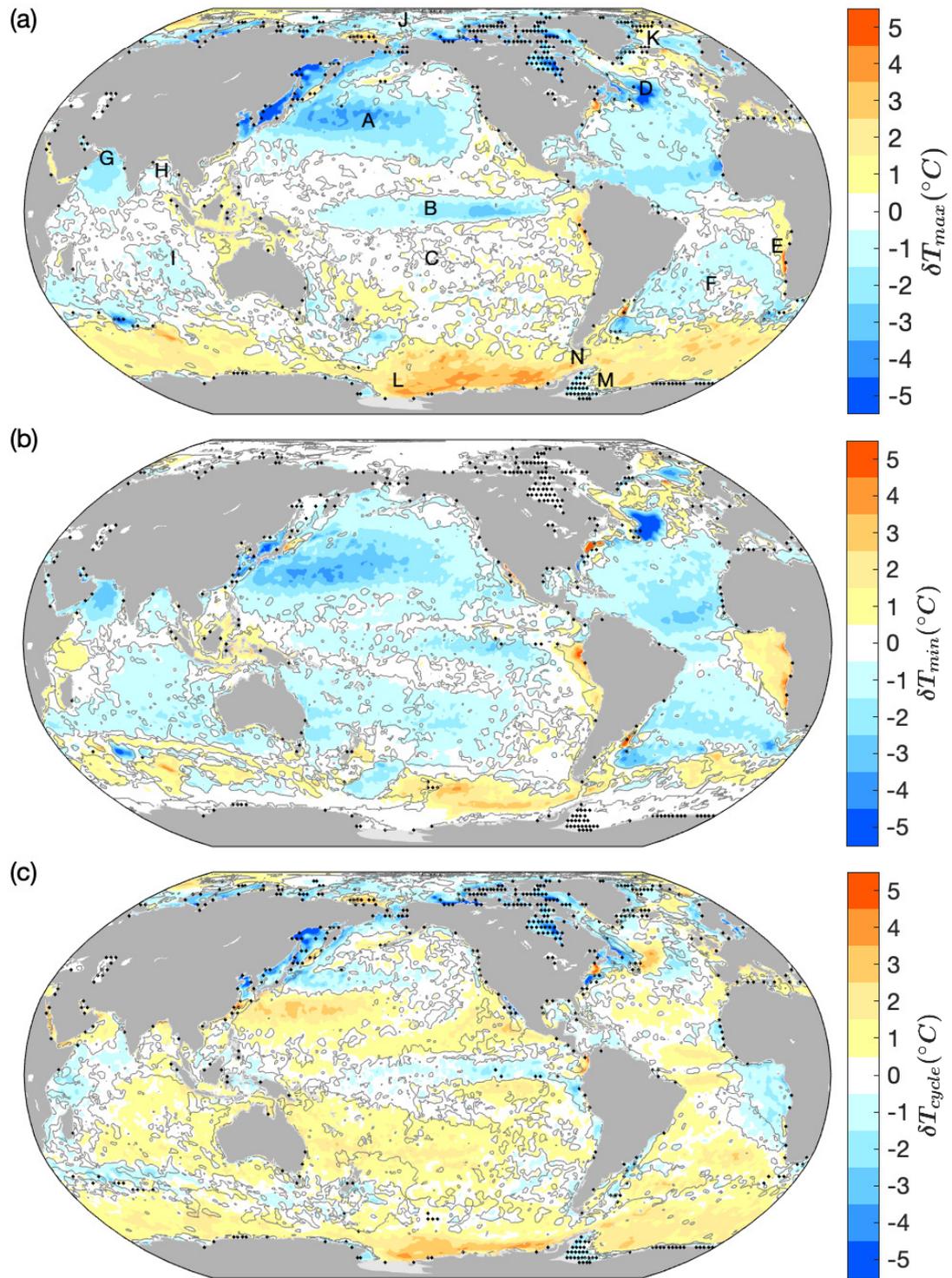
Figure 3.2: The differences between HiGEM and WOA18 (HiGEM minus WOA18), for (a) T_{max} (b) T_{min} and (c) T_{cycle}. Black dots mark grid points excluded from this analysis, as described in Chapter 2. Grey contours show where the bias is equal to the uncertainty in WOA18 (the maximum difference between WOA18, WAGHC and HadISST). Letters indicate the points for time series shown below.
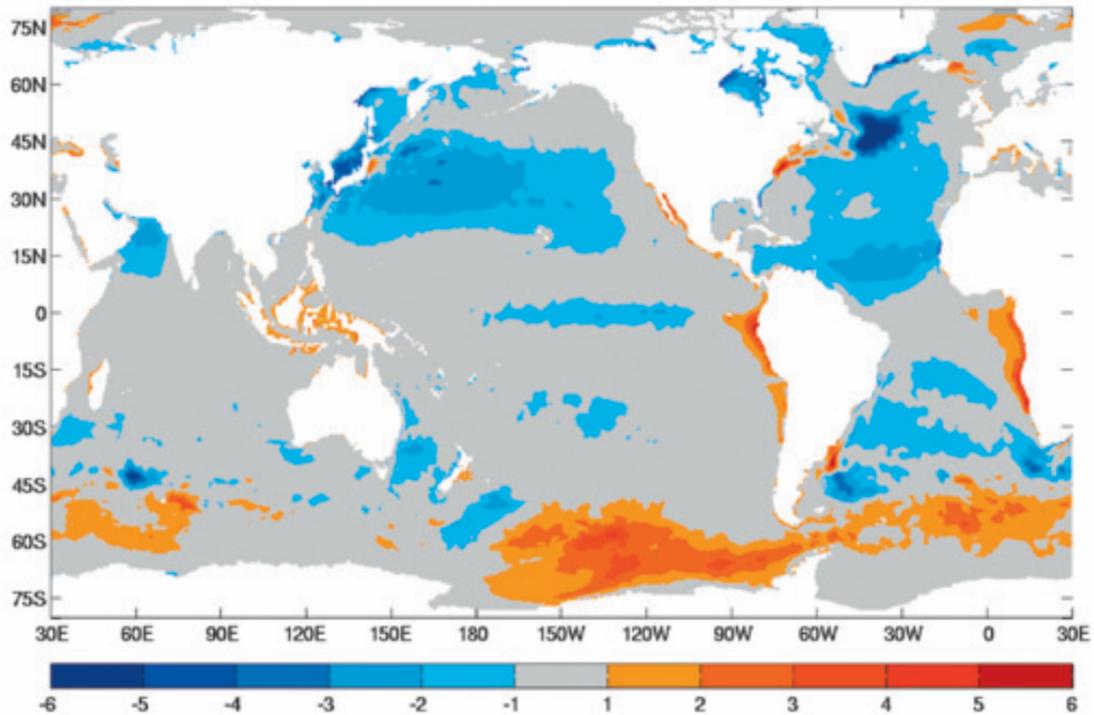
Figure 3.3: The differences between HiGEM and WOA2001 for T_{mean} (unit: °C). Adapted from Shaffrey et al. (2009).

HiGEM (Shaffrey et al., 2009). As indicated by equation 1.3, excessive precipitation will cause too large SSS decrease. Specifically, the fresh biases over Northern Hemisphere ITCZ, South Pacific Convergence Zone (SPCZ), Maritime Continent and equatorial Indian Ocean may be linked to excessive precipitation in these regions, while the saline bias over the equatorial Pacific (Fig. 3.5) is likely related to the insufficient precipitation here.

As seen above for SST, the magnitude of the bias can be different for S_{max} and S_{min}, leading to biases in S_{cycle} (Fig. 3.5c). However, the difference between S_{max} and S_{min} biases is relatively small, considering their contributions to density of sea water (Figs. 3.2, 3.5). Large seasonal variations of SSS biases are mostly found in the tropics (perhaps due to seasonal biases in the quantity and/or geographical extent of the tropical precipitation, and the Weddell and Ross Seas (perhaps due to poor simulation of sea ice formation and melting).

Here the seasonal bias of MLD in HiGEM was studied. To define MLD, we followed the method of (De Boyer Montégut et al., 2004), using 10 m as the reference depth and use a density threshold of 0.03 kg/m^3. SSD was calculated from SST and SSS using
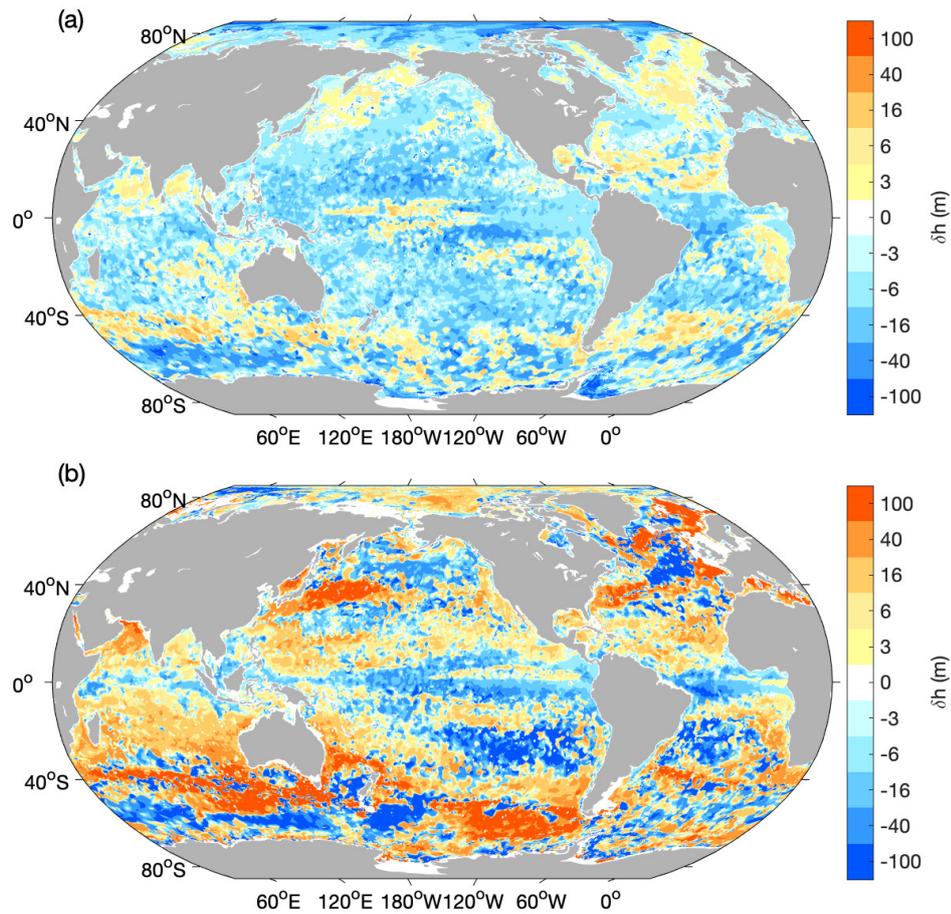
Figure 3.4: MLD difference between HiGEM and WOA18 (HiGEM minus WOA18) at the time of (a) T_{max} and (b) T_{min}.

TEOS-10 (McDougall and Barker, 2011). It is shown that seasonality of bias is also significant in MLD, with an overall too shallow MLD in summer and an overall too deep MLD in winter (Fig. 3.4). The too deep winter MLD suggests intermediate and deep waters being too deep. SST and SSS is easy to be influenced by surface fluxes of heat and freshwater when MLD is shallow. Therefore seasonal MLD bias contributes to the seasonal biases in SST and SSS.

3.2.2 Timing of seasonal extrema

The timing of T_{max} and T_{min} is well simulated in HiGEM. In most of the world ocean, the month biases are within 1 month (Fig. 3.7). Although the timing differences between HiGEM and WOA18 exist in the polar and equatorial regions, it is hard to
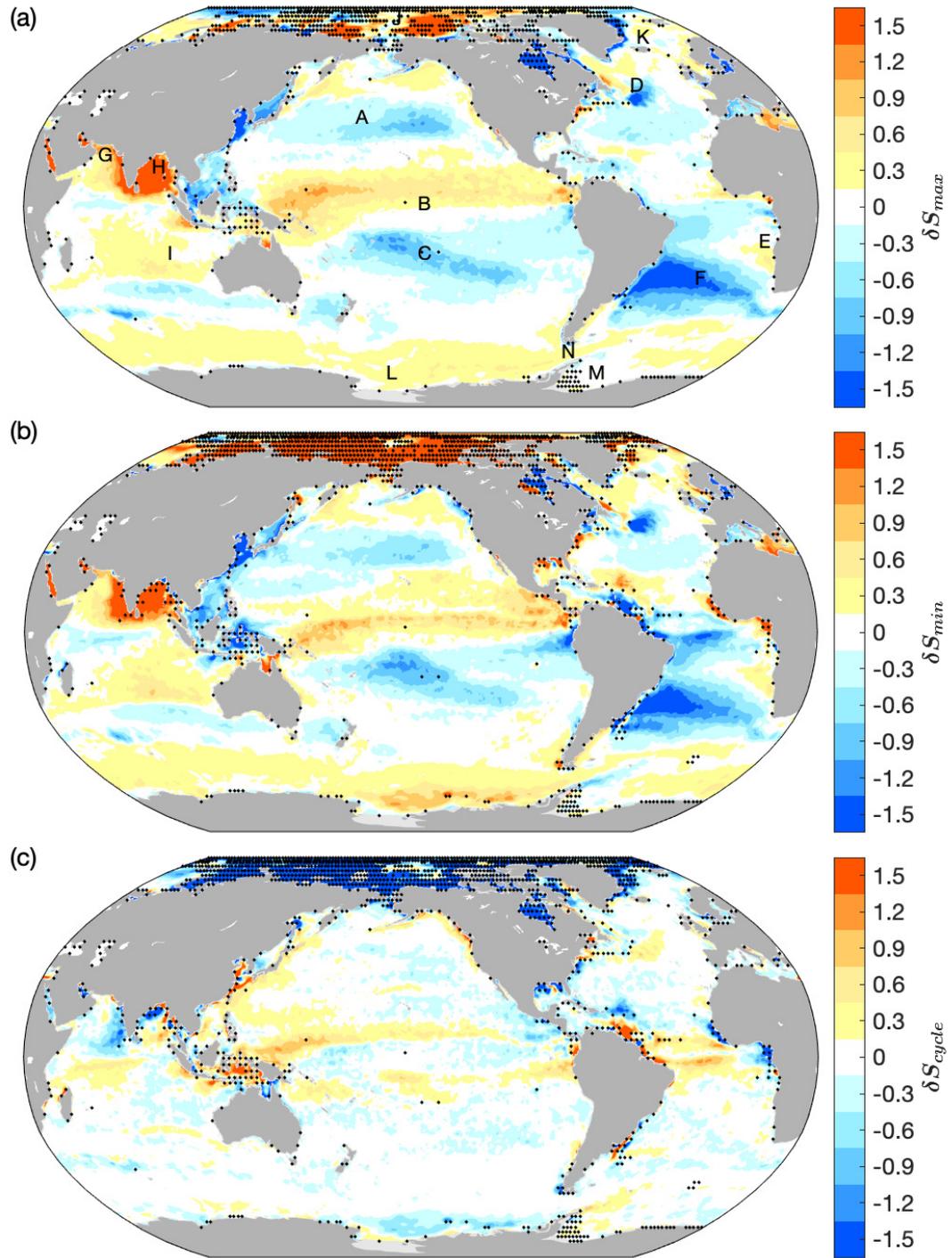
Figure 3.5: Differences between HiGEM and WOA18 for (a) S_{max}, (b) S_{min} and (c) S_{cycle}. Black dots mask grid points excluded from this analysis, as described in Chapter 2.

say whether these differences are from model biases or from the observation as these regions are mostly covered by the masks (Fig. 3.7).

Most of the world ocean is covered by the SSS uncertainty masks (Fig. 3.8), which means the bias in the timing of S_{max} and S_{min} in these regions may origin from the
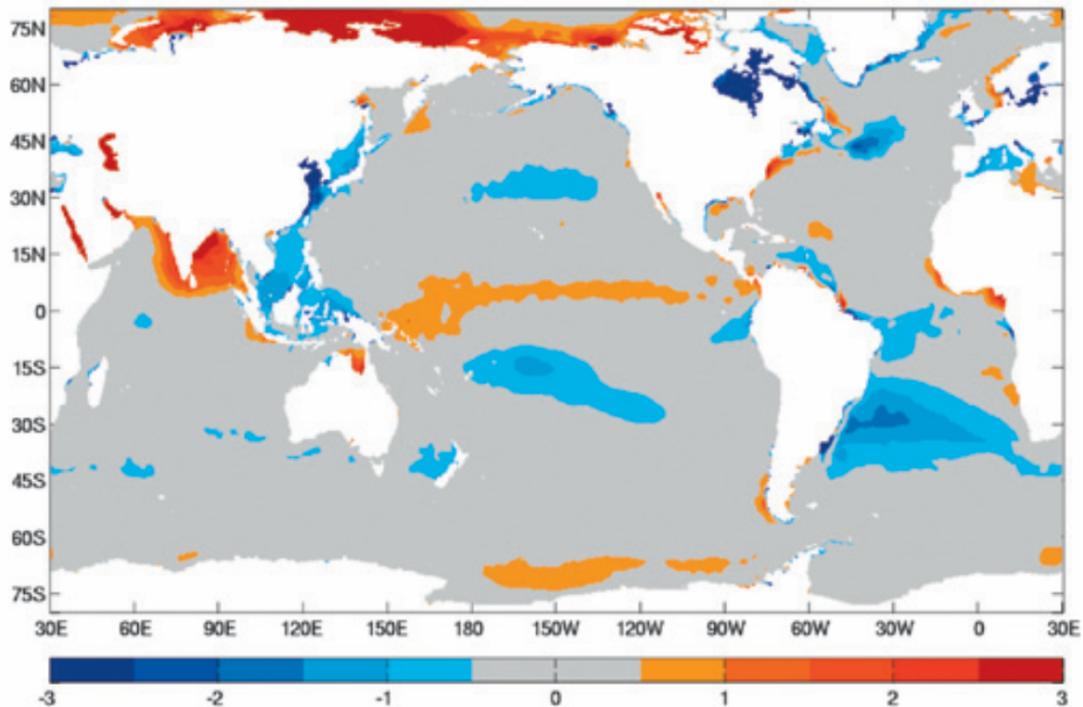
Figure 3.6: The differences between HiGEM and WOA2001 for S_{mean}. Adapted from Shaffrey et al. (2009).

SSS observation uncertainty. There are only a few regions which are not masked and show large biases, for example, tropical Indian Ocean and tropical Atlantic, with their timing biases up to 6 months. Using a lower-order mixed layer salinity model with satellite-derived data sets and World Ocean Atlas 2005, Yu (2011) stated that E-P controls seasonal SSS variability in the tropical convergence zones featuring heavy rainfall. Therefore, poor simulation of seasonal precipitation in HiGEM may cause poor simulated SSS seasonality and thus bias in the timing of S_{max} and S_{min}.

3.3 Seasonality of sea surface water in various regions

3.3.1 North Pacific

In North Pacific, HiGEM has a SST cold bias throughout the year and there is no large difference between T_{max} bias and T_{min} bias (Figs. 3.2a, 3.2b, 3.9a). Consistent with our analysis, the cold bias can also be seen in CMIP5 and CMIP6 multi-model means for the annual mean SST (Wang et al., 2014; Zhu et al., 2020). Burls et al. (2017) linked
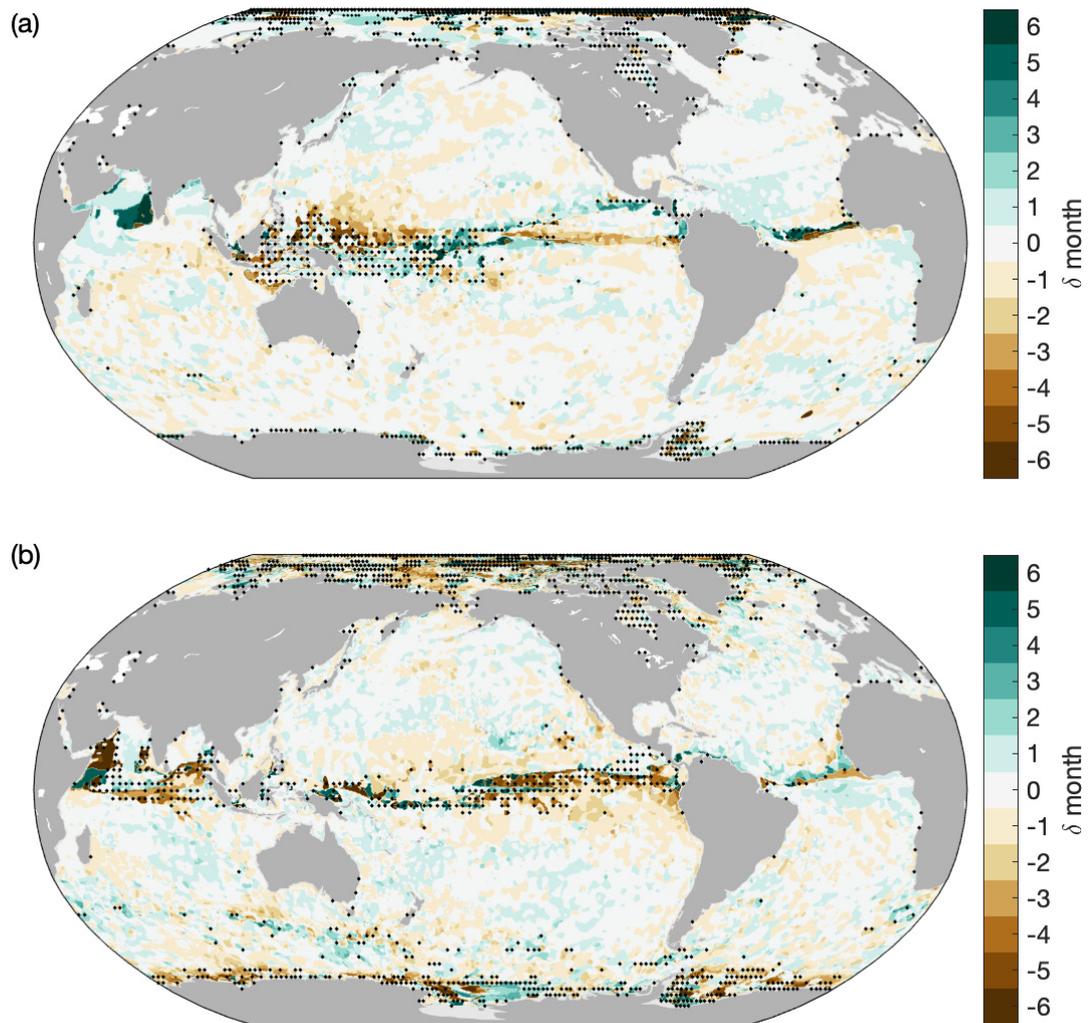
Figure 3.7: Biases in the timing of (a)T_{max} and (b)T_{min} in HiGEM. Black dots mask grid points excluded from this analysis, as described in Chapter 2.

the cold bias in CMIP5 models to too small shortwave fluxes related to a positive cloud albedo bias. Wang et al. (2014) linked the cold SST bias in CMIP5 models to intensified westerly winds over the North Pacific, which cools the SST through enhanced latent heat flux and southward ocean advection associated with Ekman transport. However, this cannot explain the cold bias in HiGEM because the wind over the North Pacific Ocean is too weak in summer (Fig. 3.10).

HiGEM has a fresh SSS bias over the North Pacific throughout the year (Figs. 3.5a, 3.5b, 3.9b). That fresh bias may be linked to the wind bias suggested by Shaffrey et al. (2009). The easterly wind bias (too weak westerly winds in HiGEM) leads to too weak southward ocean advection associated with Ekman transport. Too weak transport of
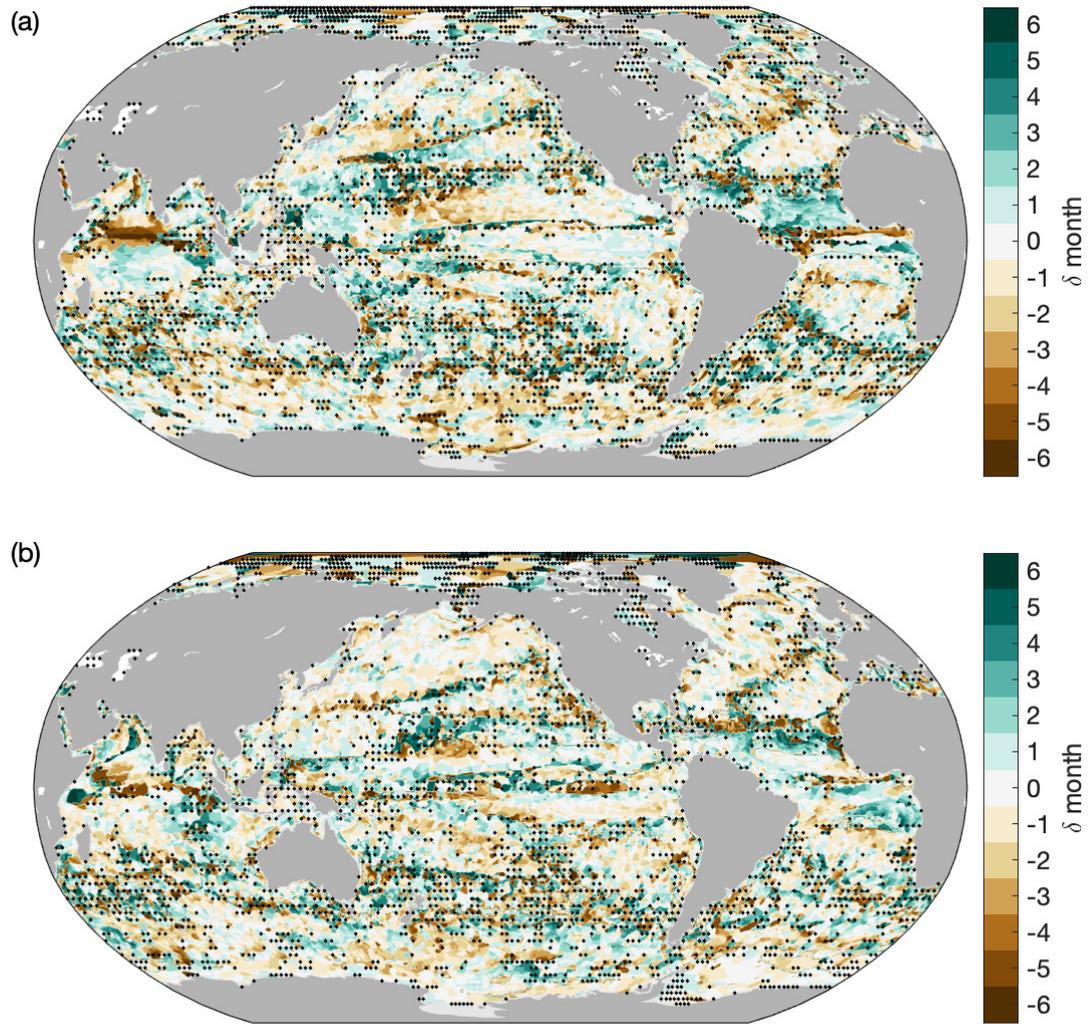
Figure 3.8: Biases in the timing of (a) S_{max} and (b) S_{min} in HiGEM. Black dots mask grid points excluded from this analysis, as described in Chapter 2.

saline water from higher latitudes contributes to a SSS fresh bias (too weak V_{ek} leads to too small SSS increase in equation 1.3). In addition, too weak wind over North Pacific in HiGEM results in insufficient evaporation, which also contributes to the too fresh SSS (too small E leads to too small SSS increase in equation 1.3).

The seasonal cycle of SST is well simulated in HiGEM (Fig. 3.9a). However, there are biases in the seasonal cycle of SSS in HiGEM (Fig. 3.9b). The range of SSS seasonal cycle is too large in HiGEM and the saltiest surface water occurs in March rather than May in WOA18 (Fig. 3.9b). The bias of SSS seasonality may be caused by the seasonality of biases in precipitation and evaporation related to wind.
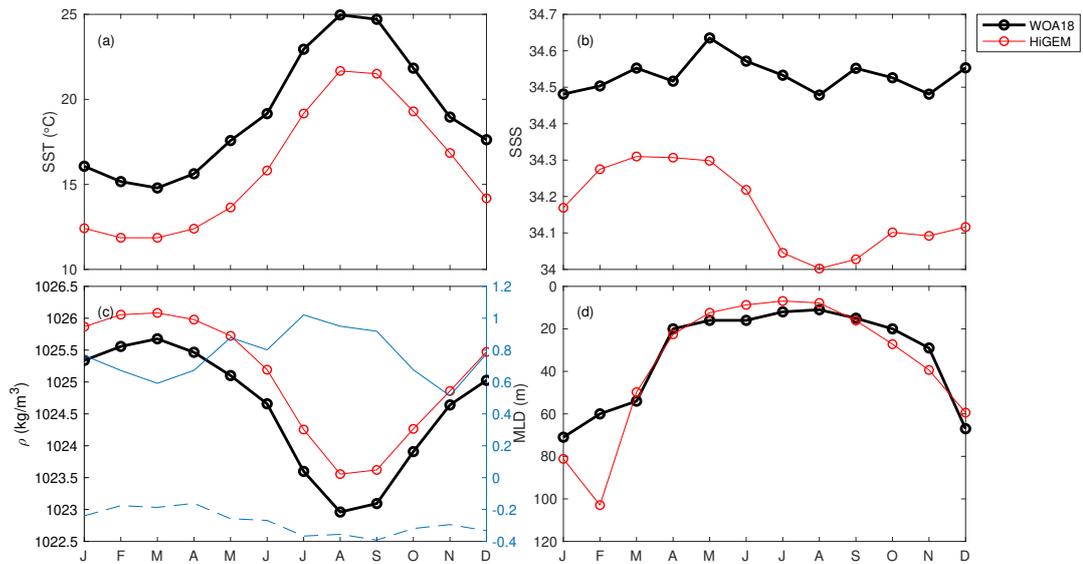
Figure 3.9: Time series for HiGEM and WOA18 at Point A in Fig. 3.2a (the North Pacific) for (a) SST, (b) SSS, (c) SSD and (d) MLD. Blue lines and blue dashed lines indicate the contribution of SST difference (HiGEM minus WOA18) and SSS difference to the density difference respectively.
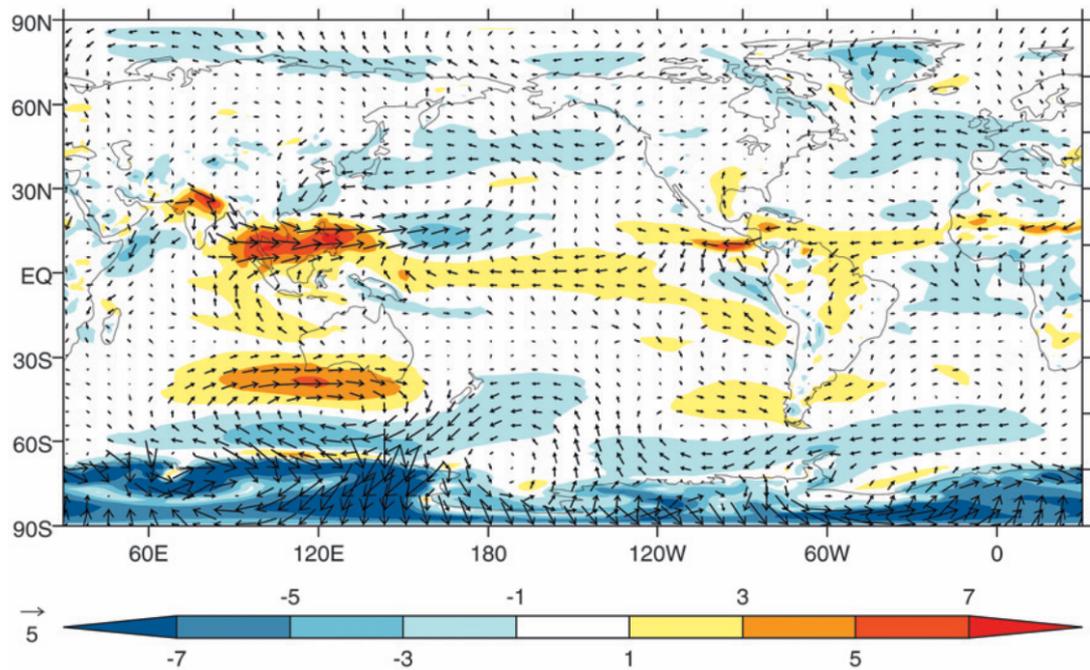
Figure 3.10: The June-July-August 850 hPa wind vector and wind speed differences (m/s) for HiGEM minus ERA-40. Adapted from Shaffrey et al. (2009).

Variations in SSD at Point A are dominated by variations in SST. The contribution of SST bias to SSD bias is around 0.6-1.0 kg·m^{-3}, while the contribution of SSS bias ranges from -0.4 to -0.2 kg·m^{-3}. The combination of cold bias and fresh bias results in too large surface density in HiGEM over the North Pacific and the density bias is around 0.5 kg·m^{-3} during December-January-February (Fig. 3.9c). The North Pacific Intermediate Water (NPIW) originates at the sea surface in the northwest Pacific (Talley, 2011). As the vertical movement of water masses based on density, too dense surface water in winter may lead to too deep NPIW in HiGEM. Consistent with the too dense surface water, MLD in February is over 40 m deeper in HiGEM than in WOA18 (Fig. 3.9d).

3.3.2 Equatorial Pacific

HiGEM has an over-extensive and too-strong equatorial Pacific cold tongue (Figs. 3.2a, 3.2b, 3.11a), which is common in CGCMs (Mechoso et al., 1995; Latif et al., 2001; Davey et al., 2002; Meehl et al., 2005). A saline bias also exists over the equatorial Pacific (Figs. 3.2b, 3.11b). In HiGEM, the excessively strong easterly wind stresses induce too much upwelling of cold and salty water (Shaffrey et al., 2009), which is associated with the cold bias and saline bias in equatorial Pacific (too large V_{ek} leads to too large SST decrease in equation 1.1 and too large SSS decrease in equation 1.3). The cold bias of equatorial Pacific is slightly larger in summer than in winter (Figs. 3.2a, 3.2b, 3.11a), which may be due to the stronger cooling effect of upwelling in summer. The too strong upwelling leads to more steeply tilted thermocline, with the west equatorial Pacific becoming deeper and eastern equatorial Pacific getting shallower. This might be the reason for too shallow MLD (Fig. 3.11d). The saline bias over the equatorial Pacific (Fig. 3.5) also coincides with precipitation low bias in HiGEM (Fig. 3.12). The equatorial precipitation bias could be a response to double-ITCZ problem in CGCMs, which is characterized by excessive precipitation over Northern Hemisphere ITCZ, South Pacific convergence zone, Maritime Continent and equatorial Indian Ocean (Lin, 2007). The cold bias (Fig. 3.11a) and saline bias (Fig. 3.11b) of surface water both lead to too dense surface water (Fig. 3.11c).
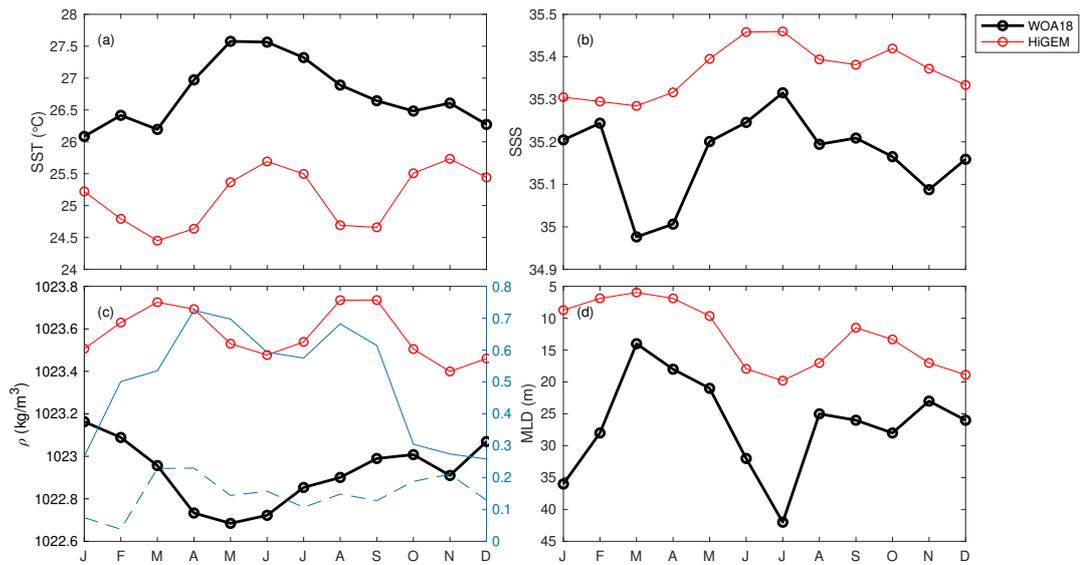
Figure 3.11: As Figure 3.9, but for Point B in the equatorial Pacific.
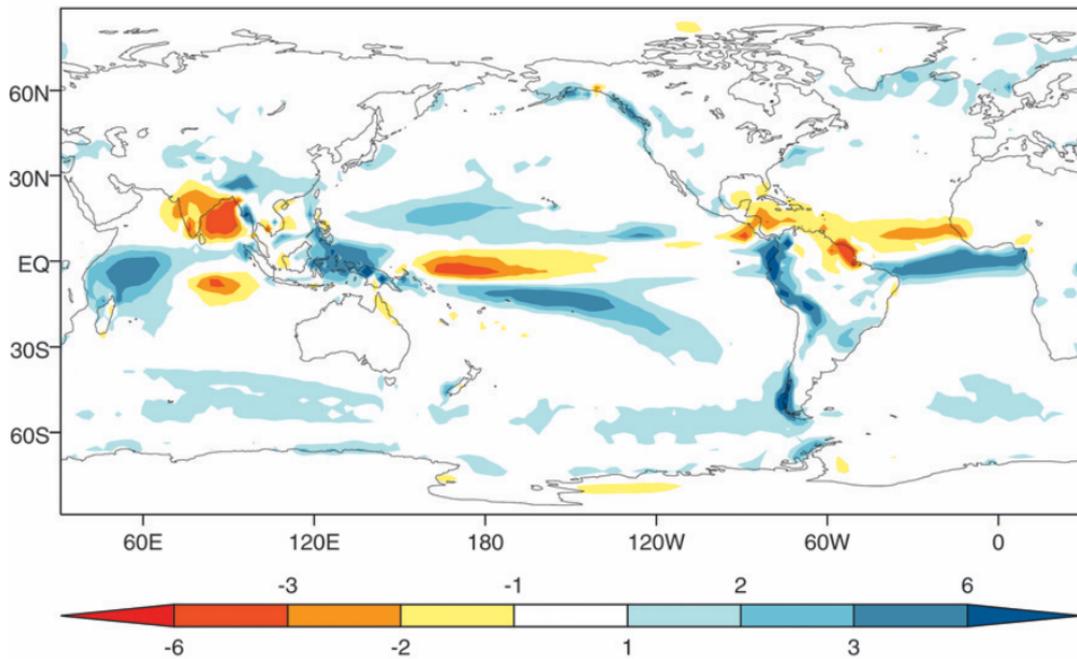
Figure 3.12: The annual mean precipitation difference for HiGEM minus CMAP. Adapted from Shaffrey et al. (2009).

The saline bias over the equatorial Pacific can be related to the cold bias (Fig. 3.2), as the precipitation response is closely tied to the SST (Xie et al., 2010; Keeley et al., 2012). Xie et al. (2010) indicated that tropical precipitation is positively correlated with SST as the moist instability dominated by surface humidity follows SST change. On the other hand, the too saline SSS over equatorial Pacific (Figs. 3.5a, 3.5b) contributes to a weaker stratification and a deeper mixed layer, which makes the SST harder to be

heated by the solar radiation and hence leads to a cold bias.

The seasonal cycle of SST is poorly simulated at Point B, where a semi-annual cycle of SST exists in HiGEM but not in WOA18 (Fig. 3.11a). Although both SST and SSS contribute to the surface density, the seasonal cycle of surface density is dominated by the seasonal cycle of SST, as can be seen from the strong anti-correlation of SST and density. The seasonal dense bias is mainly attributed to the seasonal cold bias (Fig. 3.11c). The poor simulation of the SST seasonal cycle may be caused by bias in seasonal surface radiation or seasonal easterly winds over the equatorial Pacific. The poorly simulated SSS seasonal cycle may be attributed to the bias of seasonal precipitation and evaporation due to the easterly wind.

3.3.3   South Pacific

In SPCZ, HiGEM has a large SSS bias (about 1) throughout the year (Figs. 3.2a, 3.2b, 3.13b). This fresh bias can be linked to the overestimated precipitation in HiGEM (Shaffrey et al., 2009), which might be related to the bias of SPCZ. Most CMIP3 and CMIP4 models simulate an overly zonal band of precipitation, rather than a diagonal band extending into the southeast Pacific (Brown et al., 2011, 2012). When HiGEM simulates a too zonal SPCZ, precipitation at Point C will be overestimated. The fresh SSS bias at Point C shows a seasonal cycle, with larger bias in December-January-February (Fig. 3.13b), which may be attributable to the seasonality of SPCZ intensity. As indicated by Brown et al. (2012), SPCZ is most intense in December-January-February, which is consistent with the larger SSS bias during December-January-February in HiGEM (Figs. 3.13b). The too fresh SSS (Fig. 3.13b) leads to too light surface water (Fig. 3.13c).

At point C there is a cold SST bias in winter and a slight warm SST bias in summer, which leads to too large seasonal cycle (Figs. 3.2a, 3.2b, 3.13a). Too strong precipitation can be associated with overestimated cloud, which blocks too much shortwave radiation, contributing to a cold bias. This is consistent with the cold SST bias during June-November (Fig. 3.13a). In HiGEM, too shallow MLD at Point C (Fig. 3.13d) makes upper ocean easier to be heated by summer solar radiation, which
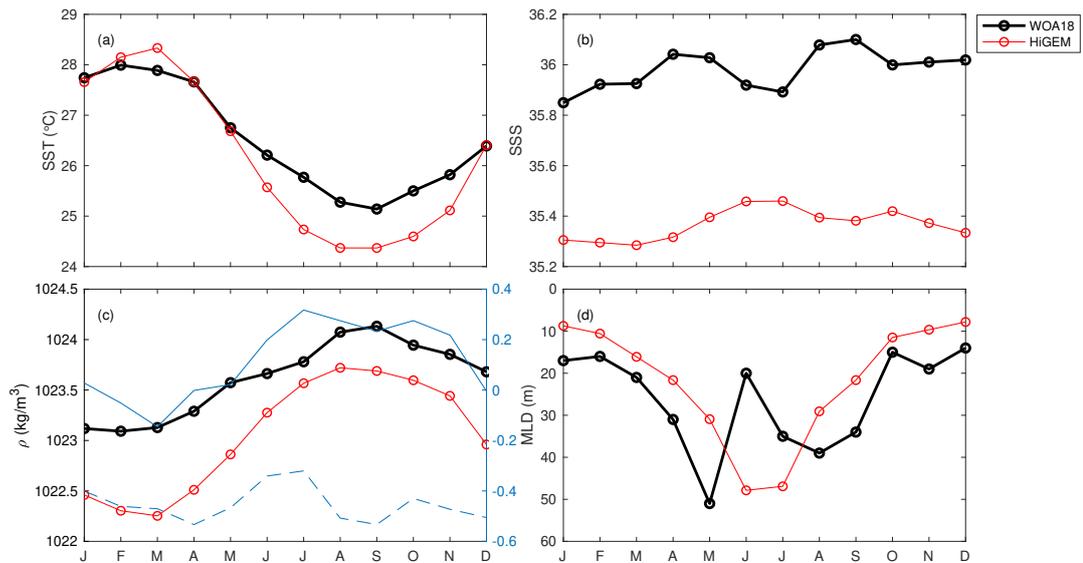
Figure 3.13: As Figure 3.9, but for Point C in the South Pacific.

can be linked to the too warm SST during February-March (Fig. 3.13a). Too shallow MLD in HiGEM turns into too deep MLD in June (Fig. 3.13d), which is consistent with the timing when SST cold bias begins (Fig. 3.13a).

3.3.4 North Atlantic

The largest cold bias in the North Atlantic SST is seen to be over 5°C in HiGEM, and it is accompanied by a fresh SSS bias. The contribution to density bias of cold bias is larger than that of salinity bias, leading to a dense bias in SSD (Fig. 3.2c, 3.14c). The cold/fresh bias is attributed to the poor performance in simulating NAC in models (Willebrand et al., 2001; Eden et al., 2004; Keeley et al., 2012). NAC has warm/salty water of subtropical origin to the right/south and cold/fresh subpolar water to the left/north. In models, the NAC is too zonal, turning northward near the mid-Atlantic ridge region in model instead of the Grand Banks in reality, which leads to the cold/fresh bias east of the Grand Banks (Figs. 3.2a, 3.2b).

There is also a warm bias close to the North American coast, owing to the poor performance in simulating the Gulf Stream in models (Willebrand et al., 2001; Eden et al., 2004; Keeley et al., 2012). The Gulf Stream separation in models is too far north from the coast of the US, which causes the warm bias near the coast (Figs. 3.2a, 3.2b).
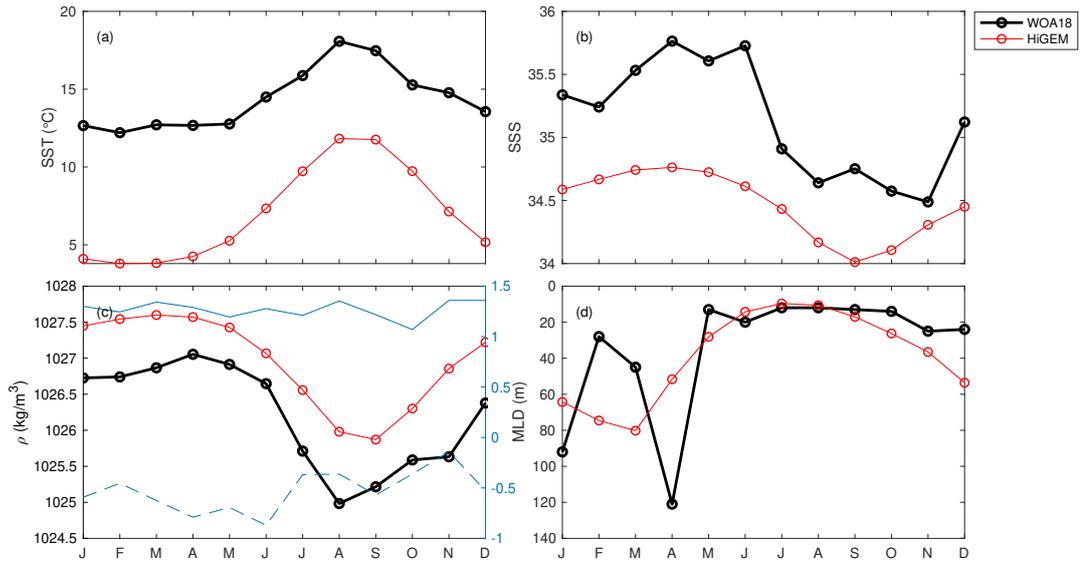
Figure 3.14: As Figure 3.9, but for Point D in the North Atlantic.

The cold SST bias associated with NAC and the warm SST bias associated with the Gulf Stream are both larger in winter than in summer (Figs. 3.2a, 3.2b, 3.14a), which can be explained by the asymmetric effect of ocean heat advection on SST between winter and summer (Liu et al., 2005). The poorly simulated paths of Gulf Stream and NAC can cause larger bias in winter than in summer, because the ocean heat advection of western boundary currents (Gulf Stream and NAC) in winter is more important to balance the heat flux to the atmosphere (Liu et al., 2005).

3.3.5 Benguela upwelling region

In the Benguela upwelling region, SST warm biases exist in HiGEM and vary seasonally (Figs. 3.2a, 3.2b). The year-round warm bias is related to the under-representation of the stratocumulus cloud and coastal upwelling (Shaffrey et al., 2009), which is common in CGCMs (Solomon et al., 2007; Huang et al., 2007). The lack of stratocumulus cloud causes a warm SST bias through excessive shortwave radiation (too large Q_{sw} in equation 1.2), and less cooling from insufficient upwelling (too small V_{ek} in equation 1.1) also contributes to a warm SST bias.

In the Benguela upwelling region, HiGEM has too small SST seasonal variance (Figs. 3.2a, 3.2c), and this may be linked to the too deep mixed layer (Fig. 3.4). Too deep MLD makes SST harder to be cooled by surface heat loss and warmed by surface heat
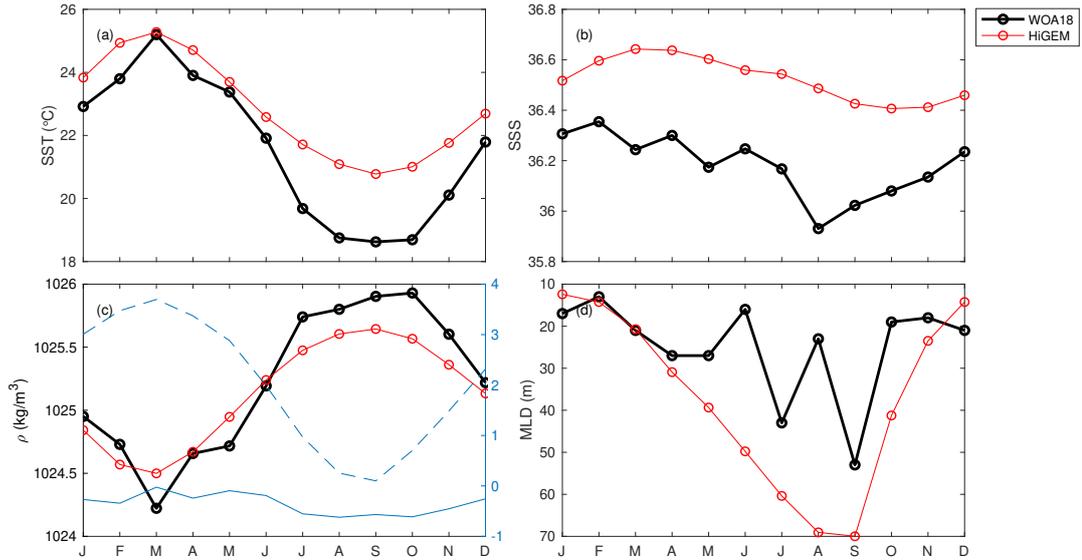
Figure 3.15: As Figure 3.9, but for Point E in the Benguela upwelling region.

gain, which leads to too small seasonal amplitude.

The warm SST biases in the Benguela upwelling region is smaller in summer than in winter, with the difference being 1-2°C (Fig. 3.2c). The warm bias in CGCM is suggested to be caused by underestimated cloud and insufficient upwelling (Richter, 2015). The warm bias due to underestimation of stratocumulus can be amplified by shallow mixed layer in summer (smaller H will cause larger SST increase in equation 1.1). Using satellite data, (Letelier et al., 2009) found that upwelling has the strongest cooling effect on SST in summer, which is relevant to the peak of upwelling-favorable wind in December and January. SSS has a seasonal saline bias, with the largest bias occurring in August (Fig. 3.15b). The seasonal density bias is dominated by the seasonal saline bias (Fig. 3.15c). The saline bias may be attributed to the underestimated cloud, which leads to excessive shortwave radiation and strengthens evaporation. Any bias in seasonal processes related to cloud (which impacts Q_{sw} in equation 1.1 and E in equation 1.3), upwelling (V_{ek} in equations 1.1 and 1.3) and MLD (H in equations 1.1 and 1.3) will contribute to the seasonal SST and SSS biases.

3.3.6 South Atlantic

The South Atlantic is dominated by fresh SSS bias in HiGEM (Fig. 3.5). The fresh SSS bias in the tropical South Atlantic is due to the precipitation bias in HiGEM (Shaffrey

et al., 2009). This fresh bias located in the tropical South Atlantic is common in coupled models and it is caused by the southward shift of the Atlantic ITCZ (Harrison et al., 2014; Liu et al., 2014). The maximum fresh SSS bias (more than 1.5) in the South Atlantic exists in midlatitudes (Fig. 3.5). (Sato and Polito, 2014) showed that South Atlantic subtropical mode water forms in the Brazil Current recirculation gyre on the western side of the basin, using Argo profiles, and hence the fresh SSS bias in HiGEM may spread into the mode water.

The SSS bias may be related to the Agulhas leakage. Using drifting buoy and float data, Richardson (2007) demonstrated that Agulhas leakage brings saltier water from the Indian Ocean into the Atlantic. In CCSM4 model, there is a good correlation between upper ocean salinity in the south Atlantic and salt transport of Agulhas leakage (Weijer and Van Sebille, 2014). When the salt transport associated with Agulhas leakage is much weaker in models than in observations, there will exist a significant salinity bias in midlatitudes of the South Atlantic. Holton et al. (2017) diagnosed the variability of Agulhas leakage in 6 ocean model simulations of varying resolution, and found that high resolution ($< 1/10°$) models can capture the spatio-temporal characteristics of Agulhas leakage, while coarser resolution models ($> 1/4°$) have difficulties in representing the Agulhas rings properties. With its $1/3°$ resolution, HiGEM may also have problems simulating the Agulhas leakage, which may lead to the fresh bias in the South Atlantic. Due to the large fresh bias, surface water is too light in HiGEM (Fig. 3.16c), which contributes to a more stratified upper ocean and thus causes a too shallow mixed layer (Fig. 3.16d).

In the South Atlantic, there is a cold SST bias of about 2°C during July-August-September (Fig. 3.16a), which may be related to too shallow MLD (Fig. 3.16d), as a shallower mixed layer is easier to be cooled in winter (smaller H can lead to larger SST decrease in equation 1.1). Eddies shed from the Agulhas Retroflection carry warm water in Indian Ocean and move northwest-ward into the Atlantic (Richardson, 2007). Sea surface height variability in the South Atlantic is too small in HiGEM (Shaffrey et al., 2009), which may indicate insufficient Agulhas eddies, contributing to the cold SST bias.
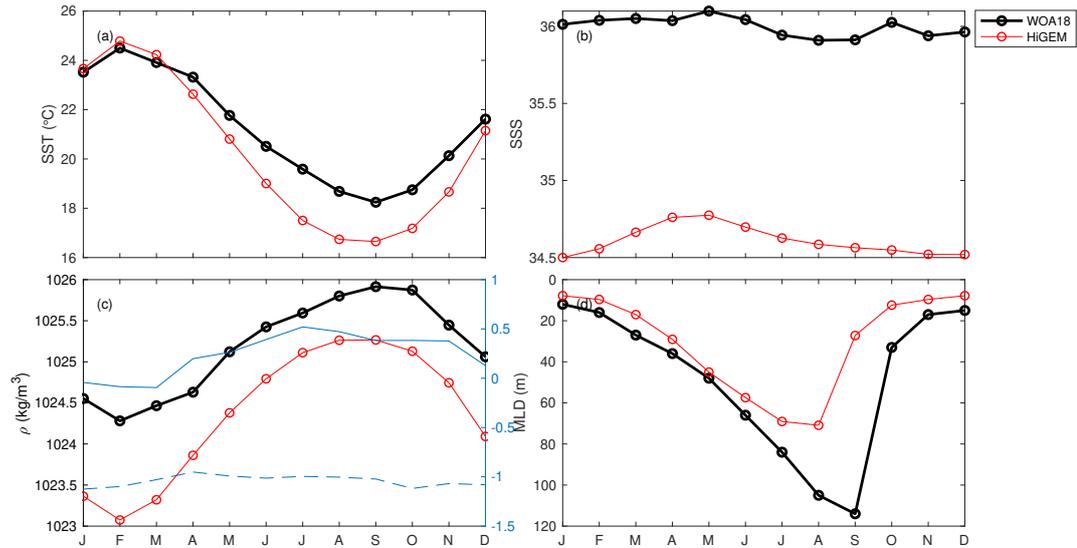
Figure 3.16: As Figure 3.9, but for Point F in the South Atlantic.

3.3.7 North Indian Ocean

A semi-annual cycle of SST in the Arabian Sea and Bay of Bengal is due to the effects of the monsoon (Figs. 3.17a, 3.18a). During summer monsoon, the southwest wind brings humid maritime air into the sea; during winter monsoon, the northeast wind brings cold continental air into the sea; during spring and fall intermonsoons the wind is weak. Wind cools the Arabian Sea and Bay of Bengal through enhanced latent heat flux (Q_{lh} in equation 1.2), resulting in a SST semi-annual cycle. SST seasonality in the north Indian Ocean is linked to the seasonal cycle of tropical cyclone intensity (Gilford et al., 2017). SSTs in Arabian Sea and Bay of Bengal are important components that cause surface moisture convergence, which is closely related to the onset of summer monsoon (Sijikumar and Rajeev, 2012; Jiang and Li, 2011). Since the Indian summer monsoon onset date is crucial for agriculture, accurate representation of the SST annual cycle of Arabian Sea and Bay of Bengal in models has significant societal applications (Prodhomme et al., 2015).

A cold SST bias exists in the Arabian Sea in both T_{max} and T_{min}, and it is larger in T_{min} than in T_{max} (Figs. 3.2a, 3.2b, 3.17a). The cold bias largely contributes to the dense bias (Fig. 3.17c). Winter cold biases are known to be detrimental to rainfall simulation in CGCMs (Marathayil et al., 2013). Marathayil et al. (2013) suggested that during December-January-February, excessive north-easterly monsoon winds in
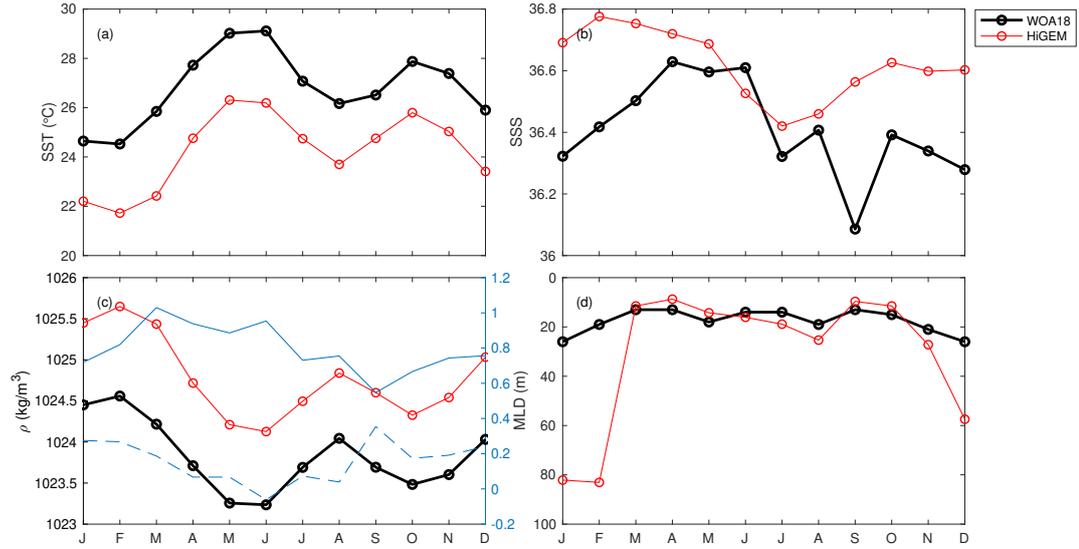
Figure 3.17: As Figure 3.9, but for Point G in the Arabian Sea.

HiGEM lead to the cold SST bias in the Arabian Sea. That also explains why there is a larger cold bias in T_{min} than in T_{max} (Fig. 3.2) (strong wind leads to large loss of Q_{sh} and Q_{lh} in equation 1.2). The excessive north-easterly monsoon winds (could cause excessive E in equation 1.3) may also contribute to the large bias of SSS and MLD during December-January-February (Fig. 3.17b, 3.17d). The poor simulated monsoon in model may also cause a phase bias in the northwest Indian Ocean. The month of T_{max} in HiGEM is up to 6 months later than in WOA18 (Fig. 3.7a), depending on whether the first or second peak of the semi-annual signal reaches higher temperatures.

In the Arabian Sea and Bay of Bengal, the lack of precipitation in the summer Indian monsoon leads to too saline surface water in HiGEM (Figs. 3.5a, 3.5b) (Shaffrey et al., 2009). These saline biases in HiGEM are consistent with that in the multi-model mean of CMIP5 models, which is also mainly attributed to precipitation bias (Fathrio et al., 2017). The Bay of Bengal receives high fresh water flux through river runoff, while this feature cannot be captured in CFSv1 and CFSv2 models (Parekh et al., 2016). If it is the same case in HiGEM, poorly simulated river runoff could be another reason for the too saline surface water in the Bay of Bengal.

The salinity bias over the Bay of Bengal has a seasonal cycle, with a larger bias in S_{min} than in S_{max} (Figs. 3.5, 3.17b). The saline bias reaches 3 and leads to a dense bias which reaches 2.5 kg·m^{-3} (Fig. 3.17c). Specifically, the saline bias is largest during
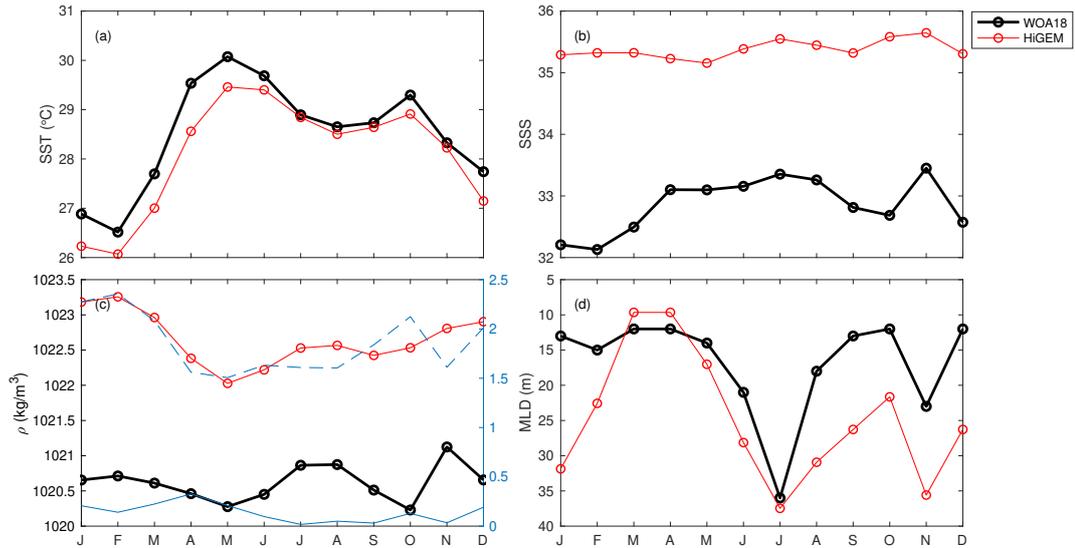
Figure 3.18: As Figure 3.9, but for Point H in the Bay of Bengal.

January-February (Fig. 3.17b). As shown by Parekh et al. (2016), CFSv1 and CFSv2 models also show a saline bias with seasonal cycle over the Bay of Bengal. Parekh et al. (2016) attributed this saline bias to weaker model precipitation and improper river runoff, especially during summer and fall (Parekh et al., 2016). However, the saline bias in CFSv1 and CFSv2 models is largest during October-November-December, different from the result in HiGEM (January-February).

3.3.8 South Indian Ocean

The magnitude of the SST bias in the Indian Ocean is smaller than in other basins. In the South Indian Ocean, there is only a slight SST cold bias in HiGEM (Figs. 3.2a, 3.2b, 3.19a). A similar cold bias in the South Indian Ocean can be seen in the CMIP5 multi-model mean (Wang et al., 2014). The cold SST bias and saline SSS bias cause too dense surface water in the South Indian Ocean (Fig. 3.19c).

Winter MLD in HiGEM is too deep (Figs. 3.4b, 3.19d). Soares et al. (2019) revealed that in both model and observation the seasonal cycle of MLD across the tropical southern Indian Ocean is largely influenced by the monsoon winds in the eastern Indian Ocean, and hence too strong winds during June-July-August in HiGEM (Fig. 3.10) might be the reason for the too deep MLD.
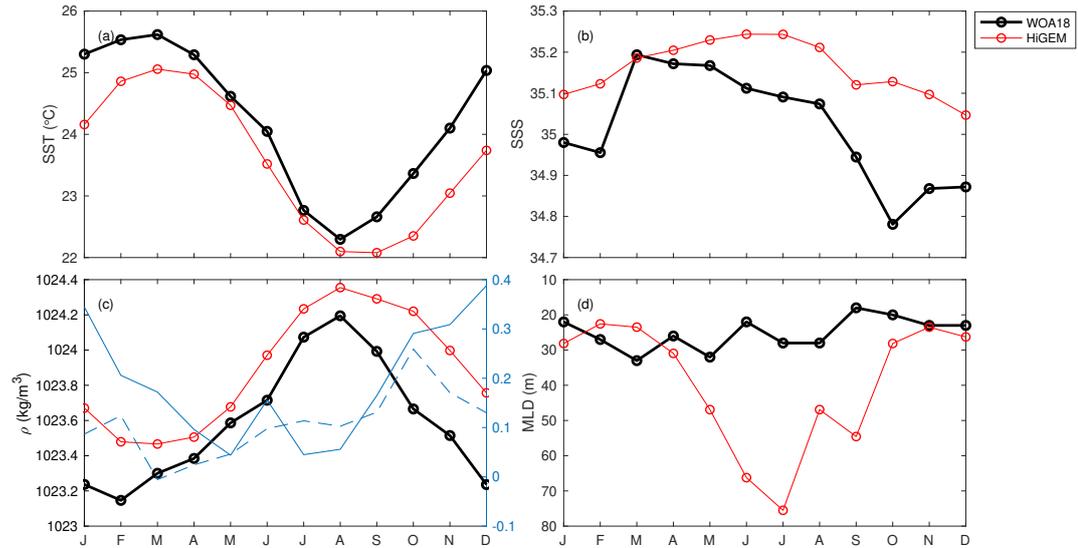
Figure 3.19: As Figure 3.9, but for Point I in the south Indian Ocean.

Lee et al. (2011) found that in HiGEM the rate of subduction of Subantarctic Mode Water in the Indian Ocean is set by the winter MLD. Therefore, the too deep winter MLD (Fig. 3.4) may contribute to overestimated subduction in the Southern Indian Ocean, which can lead to biases of the Subantarctic Mode Water.

3.3.9 Polar regions

The summer warm SST bias in the Southern Ocean (Figs. 3.2a, 3.2b, 3.20a, 3.21a), because there is too much shortwave radiation due to insufficient cloud (Shaffrey et al., 2009). The warm bias of the Southern Ocean is associated with smaller amounts of sea ice in HiGEM (Shaffrey et al., 2009), which in turn may contribute to warm T_{max} biases, as regions which are ice free in model in summer, but ice covered in reality, will absorb more solar radiation. The Southern Ocean warm bias has a seasonal cycle, with larger T_{max} bias than T_{min} bias (Fig. 3.2). In sea ice formation regions such as the Weddell Sea and Ross Sea (Figs. 3.20, 3.21), the winter SSTs both in HiGEM and WOA18 are close to the freezing point, leading to a small T_{min} bias in HiGEM. As solar radiation is negligible at high latitudes in winter, the winter warm bias due to insufficient cloud is much smaller than that in summer. The seasonal cycle of MLD may also contribute to the seasonal cycle of this warm SST bias. Shallower summer mixed layer have smaller heat content, thus the bias in solar radiation can result in a

larger bias in summer (smaller H can cause too large SST increase in equation 1.1).
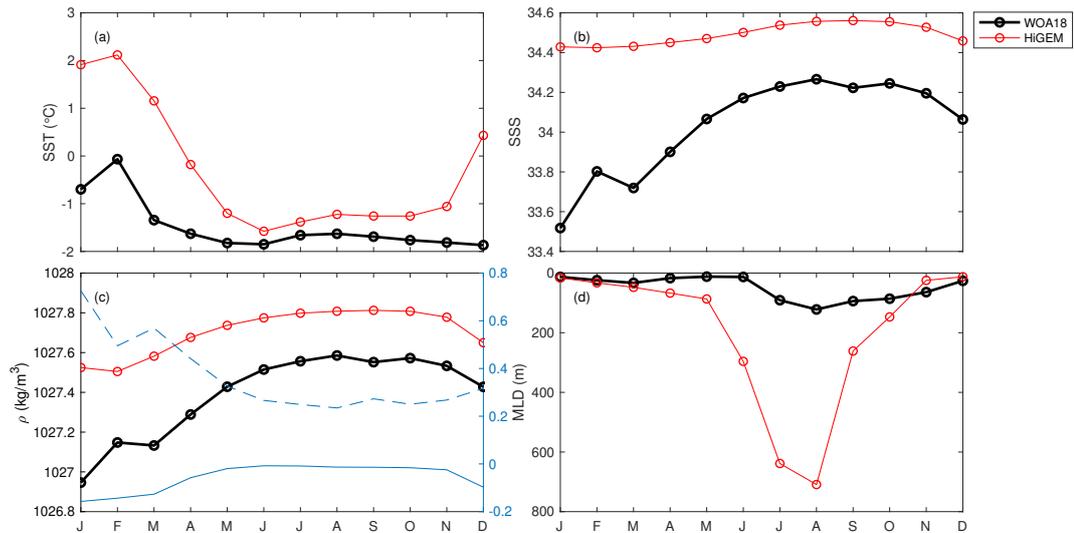
Figure 3.20: As Figure 3.9, but for Point L in the Ross Sea.
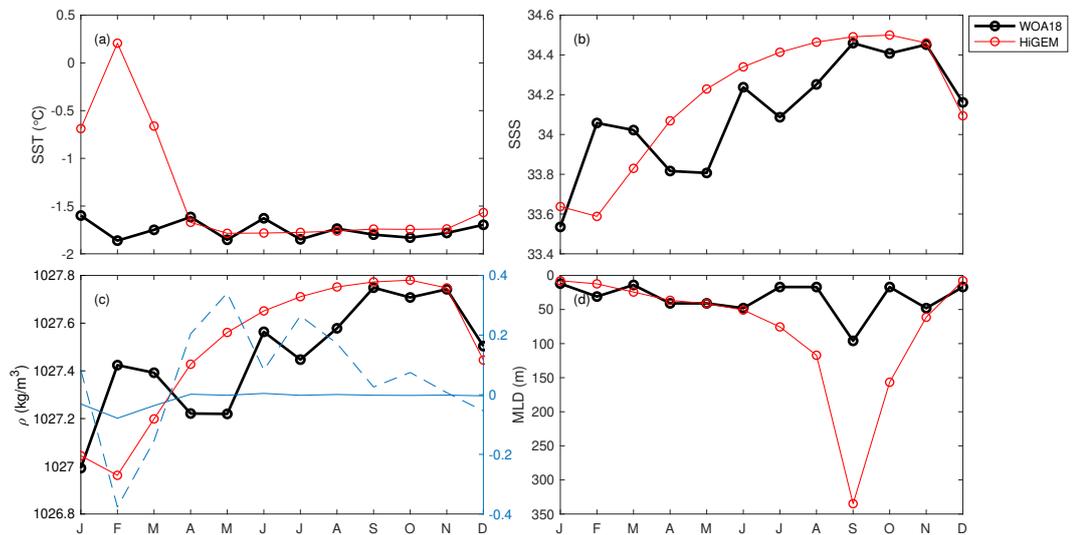
Figure 3.21: As Figure 3.9, but for Point M in the Weddell Sea.

Further north (though still within the Southern Ocean), the winter SST warm bias (Fig. 3.2b) is likely to be related to open ocean deep convection in HiGEM, which rarely occurs in reality (Heuzé et al., 2013). Open ocean deep convection brings warm water from the ocean interior to the surface and cause warm SST bias, which is consistent with the winter SST warm bias.

In most of the Southern Ocean, there is a saline bias of surface water (Figs. 3.5a-b). Vertical mixing brings entrainment of relatively salty deep water into the upper ocean. Excessive vertical mixing which is linked to too deep mixed layer in winter (Fig. 3.4b) contributes to the saline bias. The seasonality of Antarctic sea ice is less in HiGEM

than in observations (Shaffrey et al., 2009), which means the melt of sea ice in summer is less in HiGEM. This contributes to the saline bias in S_{min}. In the Ross Sea and Weddell Sea, the SSS bias dominates the SSD bias (Figs. 3.20c, 3.21c).

Sallée et al. (2013) showed that in the Southern Ocean the surface warm and light bias can be propagated into the deep ocean, using CMIP5 models. Therefore, the warm and saline bias at the surface in HiGEM can also spread over the entire water column in the Southern Ocean. Specifically, the too warm and saline surface water at point N (Fig. 3.22) will lead to too warm and too saline AAIW; at point L in the Ross Sea, the too warm and saline surface water (Fig. 3.20) will lead to too warm and saline AABW.
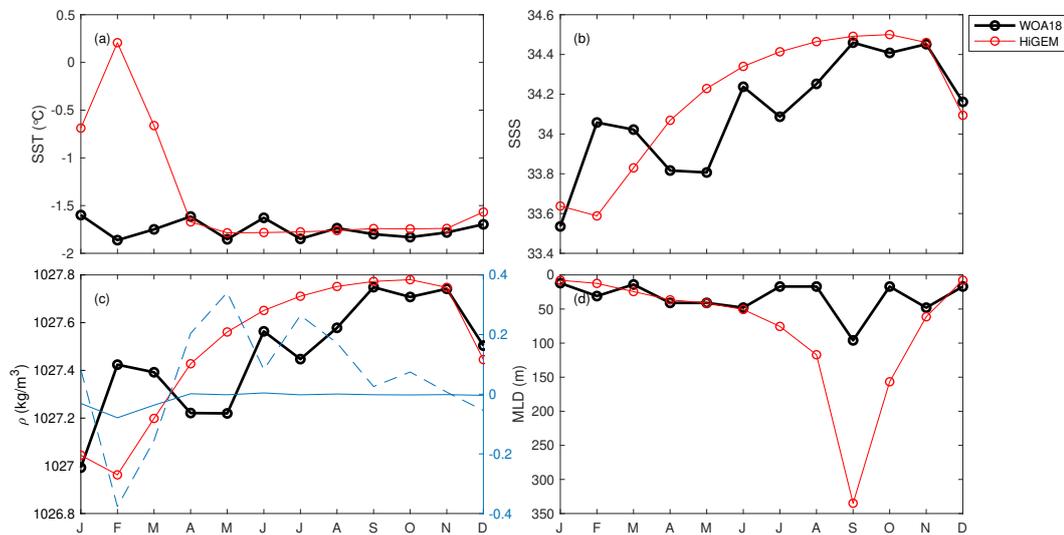
Figure 3.22: As Figure 3.22, but for Point N in the Southern Ocean.

Like the Antarctic, the Arctic has no significant biases in T_{min} (Fig. 3.2b). The winter SST at Point J in the Arctic has a cold bias of only 0.1°C, because the freezing point is about 0.1°C lower in HiGEM than in WOA18 (Fig. 3.23a). In summer, there is a cold SST bias in the Arctic (Fig. 3.2a), which is consistent with too much sea ice in HiGEM (Shaffrey et al., 2009). Bias in salinity was not considered because of the uncertainty in the climatological salinity (Fig. 3.5).

In the Greenland Sea, there is a cold SST bias, while it is larger in winter than in summer (Figs. 3.2a, 3.2b, 3.24a). That may be linked to the bias of Atlantic Ocean heat transport. The global peak poleward heat transport is weaker in HiGEM than the observations (Shaffrey et al., 2009), which may explain the cold SST bias. As Point K is a local convection source, the cold bias and the consequent dense bias of sea surface
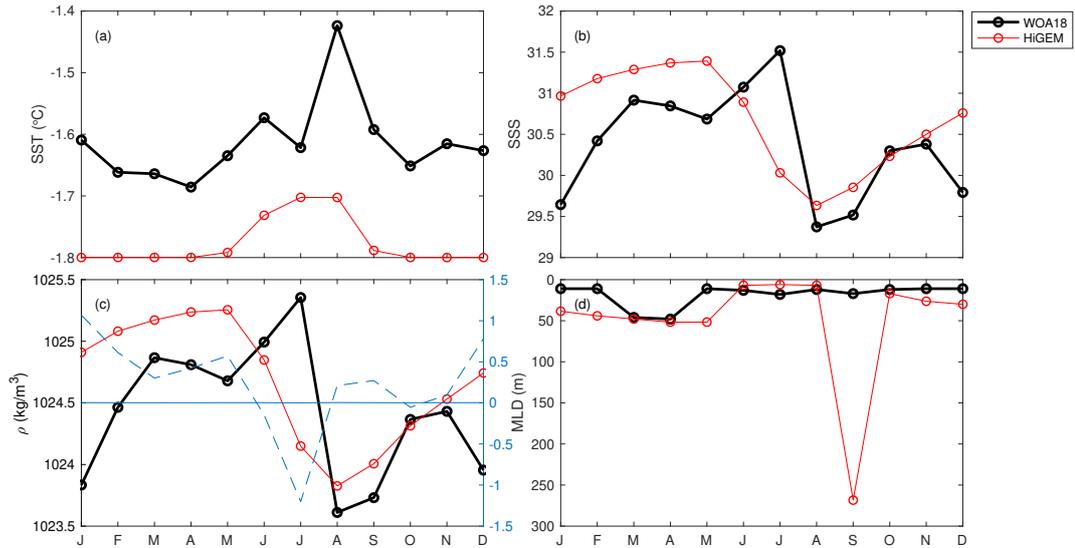
Figure 3.23: As Figure 3.9, but for Point J in the Arctic.

water will lead to a deeper convection, and North Atlantic Deep Water formed there will be too cold and too dense.
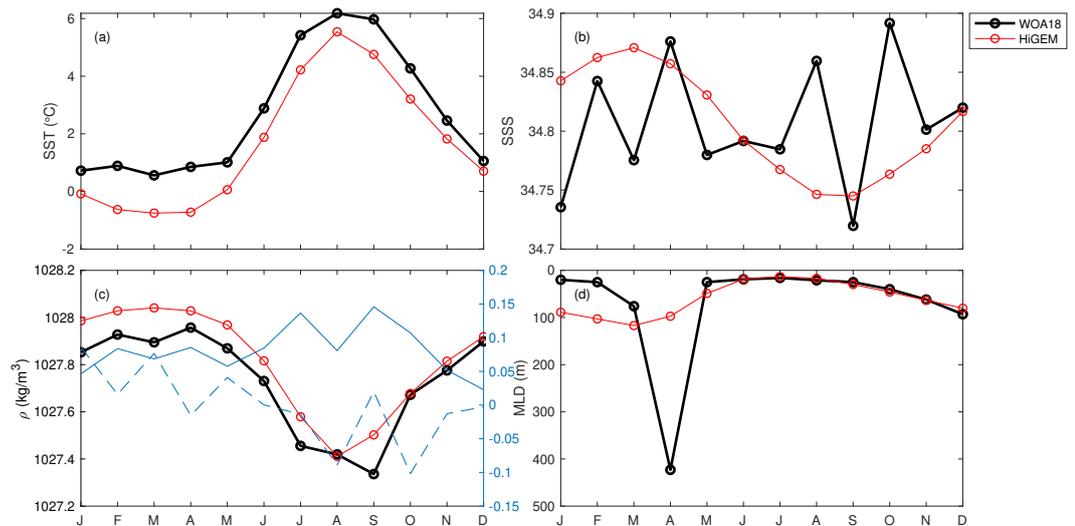
Figure 3.24: As Figure 3.9, but for Point K in the Greenland Sea.

The SSS seasonal cycle corresponds well with the SST seasonal cycle (Figs. 3.24a, 3.24b). At Point K in the Greenland Sea, SSS peaks in winter when sea ice forms with brine rejection, while SSS minimizes in summer when sea ice melts. Due to the lack of observation in ice-covered regions, SSS in WOA18 (Fig. 3.24b) may be unrealistic.

3.4   Conclusion

HiGEM have seasonally dependent SST and SSS biases. The typical value of biases for T_{max} and T_{min} is 1-2°C, while it is 0.3-0.6 for S_{max} and S_{min}. Although patterns of biases are similar in T_{max} and T_{min}, the amplitudes of T_{max} and T_{min} biases are different, leading to T_{cycle} biases. This is also the case for the SSS bias. Because of the seasonality of SST and SSS biases, there exists seasonal variation of surface density bias, which might contribute to the seasonality of MLD bias. In turn, the seasonal MLD bias may also affect the seasonal SST and SSS biases.

SST biases in HiGEM have large seasonal variations in polar regions and eastern boundary upwelling regions. In polar regions, T_{max} bias is larger than T_{min}. T_{min} bias is close to 0°C in polar regions because SSTs in HiGEM are at or close to freezing. As solar radiation is negligible at high latitudes in winter, the SST bias due to cloud error is much smaller than that in summer. The seasonal cycle of MLD also impacts seasonal SST bias. Deep winter mixed layers will diffuse the effect of surface heat fluxes, thus an error in heat fluxes or mixing processes leads to a small T_{min} bias. In eastern boundary upwelling regions (especially the Benguela Current and Humboldt Current), the warm biases are smaller in T_{max} than T_{min}. The seasonal upwelling cooling effect associated with seasonal upwelling-favourable wind contributes to the seasonality of the SST bias.

Large seasonal variations of SSS biases mainly exist in equatorial regions and polar regions, which is likely to be related to the poor simulated seasonal precipitation and sea ice. In other regions, the biases in S_{max} and S_{min} are largely accounted for by S_{mean}. The seasonality of SSS bias is relatively small compared to the seasonality of SST bias, considering to their contributions to density.

In regions where water mass forms, the bias in SST and SSS can be propagated into deeper ocean. In the Southern Ocean, the significant surface cold and saline biases can be transferred into ocean interior and cause too cold and too saline AAIW and AABW. The cold bias in the Greenland Sea can cause a cold bias in the NADW.

The seasonal cycle of SST is well simulated in HiGEM. The bias of T_{max} and T_{min}

timing is within 1 month in most of the world ocean, whereas in polar regions, equatorial regions and monsoon regions the timing bias is up to 6 months. Compared with SST, the timing of S_{max} and S_{min} is poorly simulated in specific regions.

Previous model evaluation studies focus on T_{mean}, however a good representation of T_{mean} does not guarantee accurate T_{max} and T_{min}. The magnitude of differences between T_{max} and T_{min} biases is typically 1-2°C. By comparing with the T_{mean} bias in HiGEM (Shaffrey et al., 2009), we found a larger T_{max} bias in polar regions, and a larger T_{min} bias in the eastern boundary upwelling regions.

As a feasibility study for Chapter 4, this chapter shows that SST and SSS biases in HiGEM have seasonal variations. SST is an essential parameter in weather prediction and atmospheric model simulations, and it is also important for identifying El Niño, formation of tropical cyclones and coral bleaching. Since SST bias attracts more interest from climate scientists and meteorologists, in next chapter we will focus on SST biases in CMIP6 models. Here, questions arise for Chapter 4: How is the performance of other state-of-the-art climate models in simulating T_{max} and T_{min}? Which models have the best performance? Do models have common features in the seasonal variations of SST biases?

4

Seasonal extrema of sea surface temperature in CMIP6 models

In this chapter, the performance of CMIP6 models in simulating SST seasonal extrema is assessed, and the possible causes are explored. This study has been submitted to Environmental Research Letters in May 2021 with the same title. The text in this chapter is unchanged but has been reformatted for the thesis. The results which are not shown in the paper due to limited space are provided in supplementary material at the end of this chapter.

4.1 abstract

CMIP6 model sea surface temperature (SST) seasonal extrema averaged over 1980-2010 are assessed against the World Ocean Atlas (WOA18) observational climatology. We propose a mask to identify and exclude regions of large differences between three commonly-used climatologies. The biases in SST seasonal extrema are largely consistent with the annual mean SST biases. However, the amplitude and spatial pattern of SST bias vary seasonally in the 20 CMIP6 models assessed. Large seasonal variations in the SST bias occur in eastern boundary upwelling regions, polar regions, the North Pacific and eastern equatorial Atlantic. These results demonstrate the importance of evaluating model performance not simply against annual mean properties. Models with greater vertical resolution in their ocean component typically demonstrate better representation of SST extrema, particularly seasonal maximum SST. No significant relationship with horizontal ocean model resolution is found.

4.2 Introduction

Seasonal extrema of sea surface temperature (SST) are important for the global climate system. SST seasonal maxima influence the formation and intensity of tropical cyclones (Palmen, 1948; Dare and McBride, 2011; Holland, 1997; Sun et al., 2017). SST seasonal maxima may also be associated with marine heatwaves, which can cause damage to marine ecosystems worldwide, including biomass decrease, bleaching of coral reefs, and deaths of marine animals (Cheung and Frölicher, 2020; Hughes et al., 2018; Jones et al., 2018). Heatwaves have been observed in all major ocean basins over the recent decade (Frölicher and Laufkötter, 2018) and are projected to increase in intensity and frequency over the 21st century (Oliver et al., 2019). SST seasonal minima are closely linked to the freezing and melting of sea ice and determine the properties of intermediate and deep water. Heat loss in winter allows surface water to subduct into the deep ocean, important for thermohaline circulation. Therefore, future projections of tropical cyclones, heatwaves, water mass formation or sea ice extent require our climate or earth system models to have a realistic representation of SST seasonal extrema.

Typically, however, evaluations of climate model historical runs focus on annual or long-term mean SST, revealing common biases across many models (Wang et al., 2014; Flato et al., 2013). Assessments of model performance in simulating SST seasonal cycles are less common, and are often only regional. For example, a marked seasonal variability of SST warm bias in the eastern tropical Atlantic has been documented in CMIP5 (Coupled Model Intercomparison Project Phase 5) and CMIP6 (CMIP Phase 6) models (Prodhomme et al., 2019; Richter et al., 2014; Richter and Tokinaga, 2020). In these models, the eastern tropical Atlantic warm bias is maximum in boreal summer (June-July-August), which has been attributed to the largest wind biases occurring during spring (Richter et al., 2012; Richter and Tokinaga, 2020). Similarly, CMIP6 model SST cold biases in the North Pacific subtropics vary seasonally (Zhu et al., 2020). Song and Zhang (2020) suggested that the CMIP5 multi-model mean has seasonally dependent SST biases in the northeastern Pacific Ocean, with a warm bias during summer and a cold bias during winter, which they argued was caused by poorly simulated North American monsoon winds. Wang et al. (2014) showed that the amplitude of CMIP5

multi-model mean SST biases varies seasonally and therefore an accurate annual mean SST does not guarantee accurate seasonal extrema or seasonal cycle. Here we evaluate the seasonal cycle globally in 20 state-of-the-art CMIP6 climate models, to provide a foundation for model SST bias identification and future reduction. By presenting maps of SST bias in seasonal extrema for each model, we highlight the care needed in selecting these models for future climate projections in particular regions. Section 2 introduces the models and the analysis techniques, including evaluation of uncertainties in global observational climatologies. Section 3 presents and discusses the biases in SST maxima and minima, and explores possible causes.

4.3   Data and Methods

The historical runs of 20 models (table 4.1) were averaged over 1981-2010 to create monthly mean climatologies for each model. The first ensemble member (r1i1p1f1) is used, except when r1i1p1f1 is not available; we choose r1i1p1f3 for HadGEM3-GC3-LL and HadGEM3-GC3-MM; r1i1p1f2 for UKESM1-0-LL. The models include those incorporating biogeochemical cycling (earth system models) as well as conventional climate models. The ocean vertical coordinate is typically z-level (or the related z^*) but some models use isopycnal, sigma or hybrid coordinates (table 4.1). The total number of levels and thickness of top grid cell are used as proxies for ocean vertical resolution. The global averaged thickness of top grid cell in INM-CM5-0 was calculated using the sigma coordinates and bottom topography obtained from E.M.Volodin (personal communication). The thickness of the top grid cell in other models was obtained from the references cited in table 4.1.

To examine the seasonal cycle of SST, most studies picked specific months to represent summer and winter (for example, Zhang and Zhao (2015); Liu et al. (2020)). However, model seasonal cycles may be out of phase with observations and real-ocean maxima and minima occur in different months in different regions. Instead, here we simply take the maximum and minimum SST of the monthly mean climatologies (T_{max} and T_{min}) at each grid point, whichever months they occur in, for both model and observation. T_{max} and T_{min}, plus the annual mean SST (T_{mean}) and the range of the seasonal cycle

Table 4.1: The 20 CMIP6 models used in this study; the horizontal resolution of their ocean component; ocean vertical coordinate (z - z traditional height coordinate; z^* - rescaled height coordinate for more accurate representation of free-surface variations; ρ - isopycnic coordinate; σ - terrain following sigma coordinate; several symbols refer to a hybrid coordinate) and total number of ocean vertical levels; thickness of the ocean top grid; and references.

Model	Horizontal resolution	Vertical coordinate	Total levels	Top grid thickness	References
ACCESS-CM2	100 km	z^*	50	10 m	Bi et al. (2020)
ACCESS-ESM1-5	100 km	z^*	50	10 m	Law et al. (2017)
AWI-CM-1-1-MR	25 km	z-σ	46	5 m	Semmler et al. (2020)
BCC-CSM2-MR	50 km	z	40	10 m	Wu et al. (2019)
BCC-ESM1	50 km	z	40	10 m	Wu et al. (2020)
CESM2	100 km	z	60	10 m	Danabasoglu et al. (2020)
CanESM5	100 km	z	45	6 m	Swart et al. (2019)
E3SM-1-0	50 km	z^*	60	10 m	Golaz et al. (2019)
GFDL-CM4	25 km	z^*-ρ	75	2 m	Held et al. (2019)
GISS-E2-1-G	100 km	z	40	10 m	Kelley et al. (2020)
GISS-E2-1-H	100 km	z-ρ-σ	32	10 m	Kelley et al. (2020)
HadGEM3-GC31-LL	100 km	z^*	75	1 m	Andrews et al. (2020)
HadGEM3-GC31-MM	25 km	z^*	75	1 m	Andrews et al. (2020)
INM-CM5-0	50 km	σ	40	7.3 m	Volodin et al. (2017)
IPSL-CM6A-LR	100 km	z^*	75	2 m	Boucher et al. (2020)
MIROC6	100 km	z-σ	62	2 m	Tatebe et al. (2019)
MPI-ESM1-2-HR	50 km	z	40	12 m	Müller et al. (2018)
NorESM2-MM	100 km	ρ	53	2.5 m	Seland et al. (2020)
SAM0-UNICON	100 km	z	60	10 m	Park et al. (2019)
UKESM1-0-LL	100 km	z^*	75	1 m	Sellar et al. (2019)

($T_{cycle} = T_{max} - T_{min}$) from the model climatologies are compared with the World Ocean Atlas 2018 (WOA18) observational climatology on a grid spacing of $0.25° \times 0.25°$ (Locarnini et al., 2018), which covers the period from 1981 to 2010. The model fields were interpolated to the same grid as WOA18. Biases are defined as model values minus WOA18 values. For the multi-model mean, at each grid point we average T_{max}, T_{min}, T_{mean} and T_{cycle} across the 20 CMIP6 models. To quantify the performance of CMIP6 models, we calculated the area-weighted root mean square error of the model against WOA18 (henceforth RMSE) for global SST.

Since there is some uncertainty in observational climatologies because of sparse sampling, instrumental error, quality control or gridding techniques, we compared

three recent climatologies: WOA18, WOCE-Argo Global Hydrographic Climatology (WAGHC)(Gouretski, 2018) (covering the time period 1985-2016), and HadISST (Rayner et al., 2003) (covering the time period 1981-2010). Any grid points where the maximum difference in T_{max} or T_{min} between the three climatologies is larger than 2°C are considered uncertain for that variable, and these grid points are excluded from our assessment. Any grid points which did not have values for all 12 months for at least two climatologies are also excluded. For T_{mean} and T_{cycle}, we exclude any points where either T_{max} or T_{min} is excluded. 4%, 3%, 4% and 4% of the ocean's surface area is excluded for T_{max}, T_{min}, T_{mean} and T_{cycle} respectively. In our global maps, these points are masked, and in calculations of global and regional metrics, these points are excluded.

4.4 Results and Discussion

4.4.1 Model representation of SST extrema

For the multi-model mean, T_{max} and T_{min} have larger global RMSEs than T_{mean} (Fig. 4.1), as SST biases with opposite signs in different seasons compensate each other when calculating the annual mean. Similarly, the T_{max} and T_{min} global RMSEs of the multi-model mean are smaller than the RMSEs of individual models (Figs. 4.1b-c, 4.2, 4.3). Therefore, a small bias in T_{mean} does not guarantee a realistic T_{max} or T_{min}; a small bias in a multi-model mean does not guarantee good performance of individual models.

The magnitudes of biases in T_{max} and T_{min} vary from model to model (Figs. 4.2, 4.3). The multi-model mean has RMSE less than 1°C in both T_{max} and T_{min} (0.89°C and 0.87°C respectively). Most models have T_{max} and T_{min} RMSEs between 1°C and 2°C. Only HadGEM3-GC31-LL and GFDL-CM4 have T_{max} RMSE less than 1°C (0.94°C and 0.93°C respectively). GISS-E2-1-H has the largest T_{max} RMSE of 1.89°C and MIROC6 has the largest T_{min} RMSE of 1.62°C (Figs. 4.2, 4.3).

T_{max} and T_{min} biases vary with latitude (Figs. 4.1b-c, 4.2, 4.3, 4.4g-h). Typically, the RMSE of T_{max} at 30°-80° is 1-2°C larger than at low latitudes (latitudes between 30°N and 30°S) (Fig. 4.4g). For GISS-E2-1-H, GISS-E2-1-G, BCC-CSM2-MR, BCC-ESM1
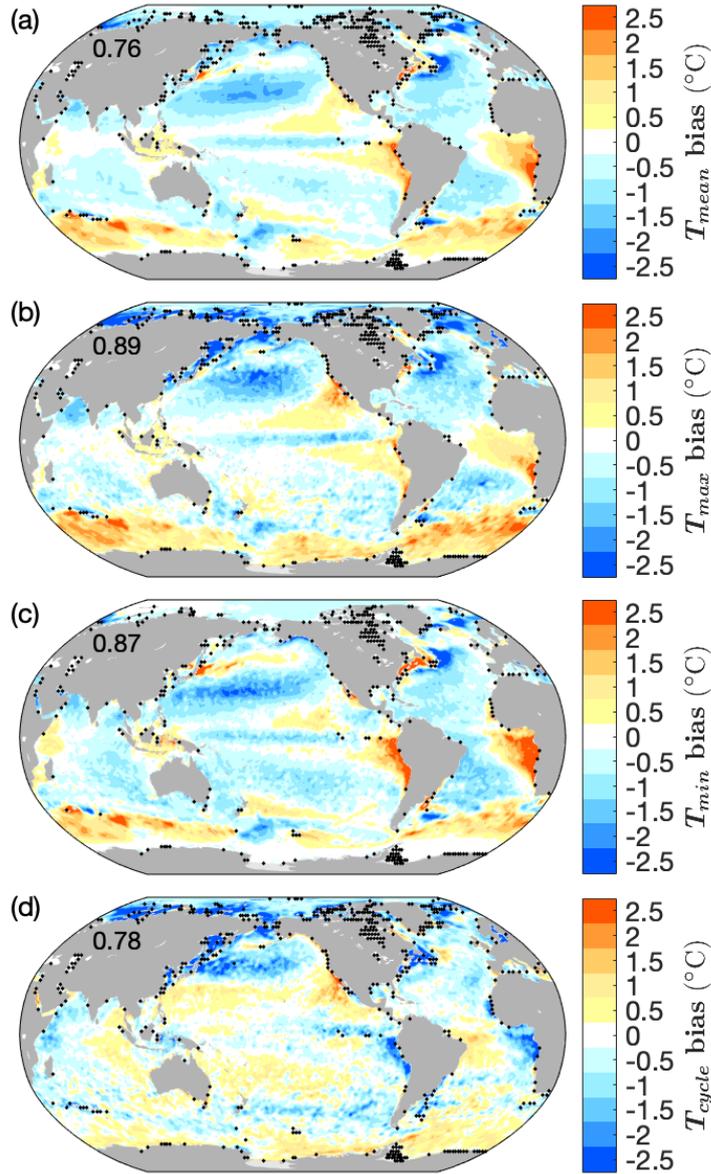
Figure 4.1: Biases (model minus climatology) of multi-model mean in (a) T_{mean} (b) T_{max} (c) T_{min} (d) T_{cycle}. Black dots mark grid points excluded from our analysis, as described in section 4.3. The numbers indicate the global RMSE (°C).

and IPSL-CM6A-LR, T_{max} RMSEs at 30°N-80°N are about 3°C larger than at low latitudes. A similar pattern can be seen for T_{min}, but the variation of biases with latitude is much smaller than for T_{max} (Fig. 4.1c, 4.4h). Flato et al. (2013) found a similar result for some CMIP5 models, with larger zonal mean biases in T_{mean} at latitudes between 30° and 70° than at other latitudes. The larger biases, and greater difference between T_{max} and T_{min}, at mid-high latitudes (latitudes greater than 30 degrees in both hemispheres) may be explained by the large seasonal cycle of mixed layer depth there. Shallower summer mixed layers have smaller heat capacity, thus a small
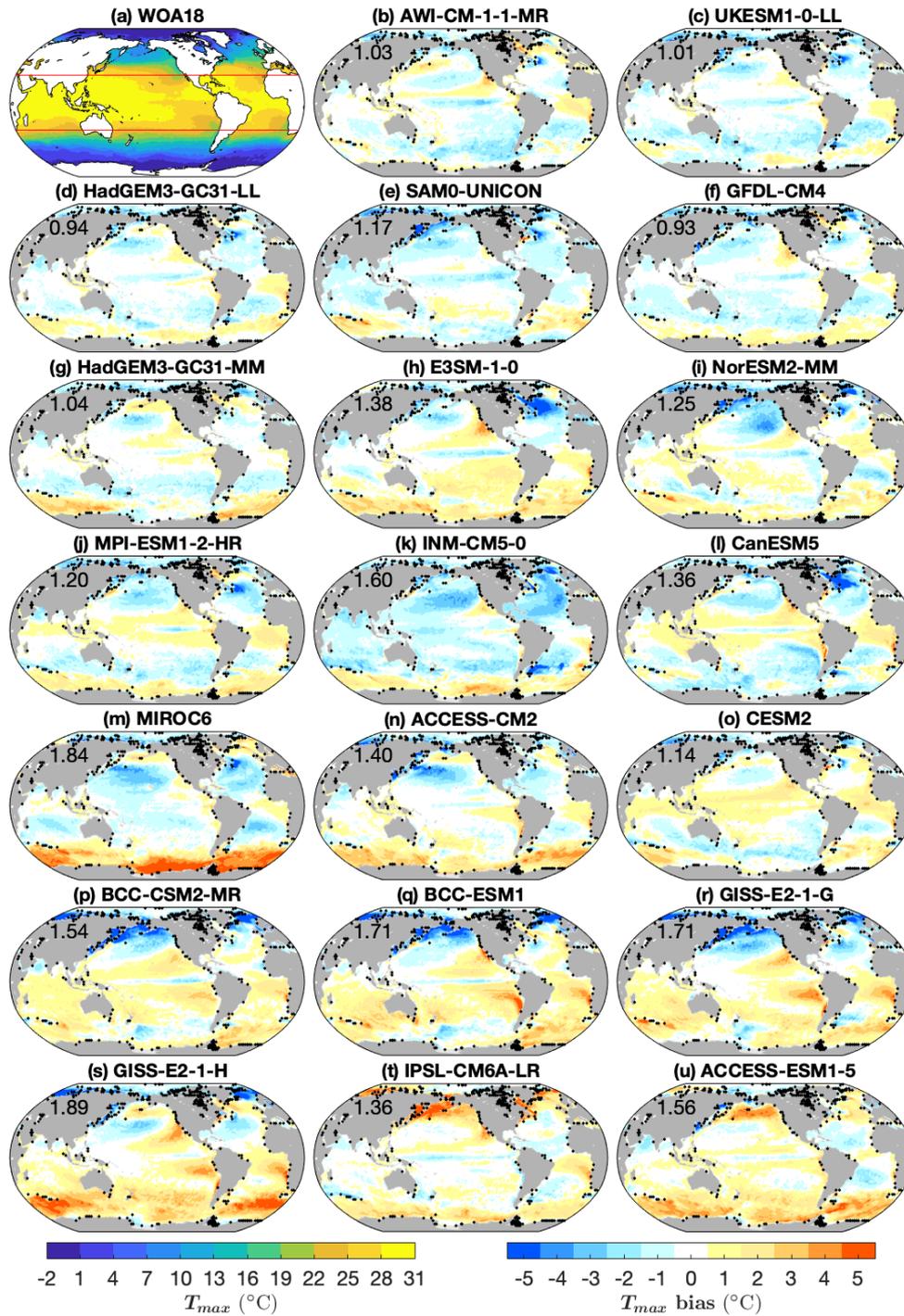
Figure 4.2: (a) T_{max} in WOA18 and (b-u) T_{max} model biases. Black dots mark grid points excluded from our analysis, as described in section 4.3. The numbers on (b-u) indicate the global RMSE of T_{max}. Red lines in (a) are 30°N and 30°S. Note that the range of bias color bar is twice as much as in Fig. 4.1.
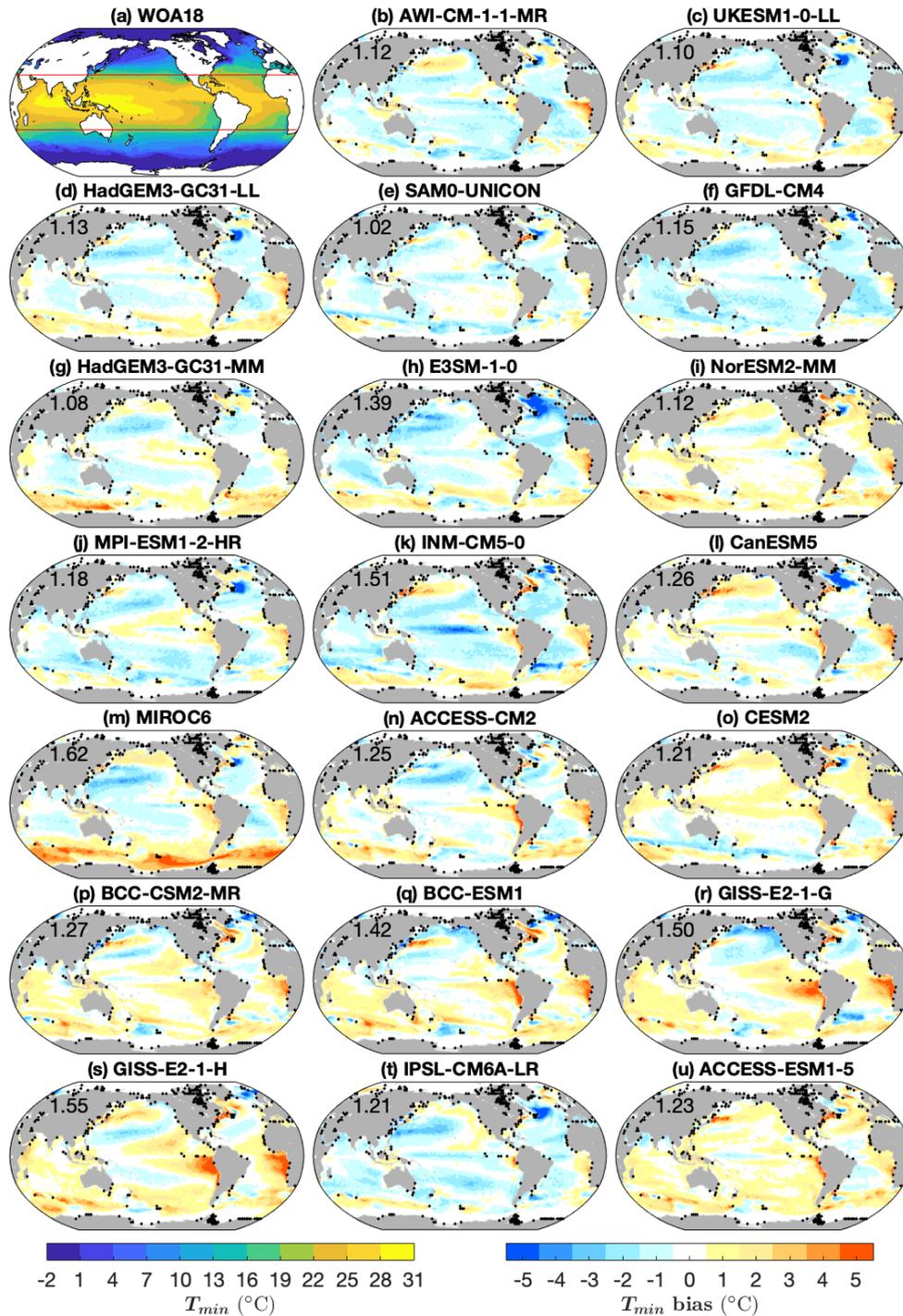
Figure 4.3: As Fig. 4.2, but for T_{min}.

error in heat fluxes or mixing processes can result in a large bias for T_{max}, though this will be modulated by any seasonal biases in mixed layer depth. The difference between biases in T_{max} and T_{min} leads to biases in T_{cycle} (Fig. 4.1d). The RMSE of T_{cycle} at low latitudes is typically 1°C, whereas at mid-high latitudes it is larger, particularly in the Northern Hemisphere (Fig. 4.4i). The T_{cycle} RMSE in IPSL-CM6A-LR and MIROC6 reaches 4°C at high latitudes (Fig. 4.4i).

In polar regions, there are very small T_{min} biases (Figs. 4.1c, 4.3, 4.4h) except for MIROC6 in the Antarctic. Winter SSTs in models are at, or close to, freezing, but cannot go below freezing because sea ice forms instead. As long as the models have realistic freezing points, T_{min} biases will be low. Some models have salinity-dependent freezing points (Beaumet et al., 2019) in which case a salinity bias can cause a bias in temperature. T_{min} biases in the Arctic are larger than in the Antarctic (Figs. 4.1c, 4.4e-f). The larger T_{min} cold biases in the Arctic may suggest larger saline biases or too high sea ice extent in models (Shu et al., 2020).

In the subtropical North Pacific, the SST cold bias is typically 0.5-1°C smaller in T_{max} than in T_{min}, which leads to a too large T_{cycle} (Figs. 4.1b-d, 4.2, 4.3). Zhu et al. (2020) showed a similar seasonal SST cold bias in the CMIP6 multi-model mean, but not in the CMIP5 multi-model mean. Too strong westerly winds (Wang et al., 2014) and biased cloud albedo (Burls et al., 2017) are possible reasons for the year round cold bias. The westerly winds cool the surface through latent heat flux and southward ocean advection due to Ekman transport. The latent heat loss is larger in summer when evaporation is stronger (Yu, 2007), while the ocean heat advection is larger in winter when meridional SST gradient is greater. For the biased cloud albedo which causes insufficient surface shortwave fluxes, its associated cold bias can be smaller in winter when there is less solar radiation.

SST biases are seasonally dependent in the northeastern Pacific Inter Tropical Convergence Zone (ITCZ) (Figs. 4.1b-c, 4.2, 4.3). For the multi-model mean, there is an evident warm bias which exceeds 2°C in T_{max} and a cold bias of 0.5-1.5°C in T_{min}. Consistent with our results, Song and Zhang (2020) showed in CMIP5 models a warm bias during summer-autumn and a cold bias during winter-spring in the northeastern
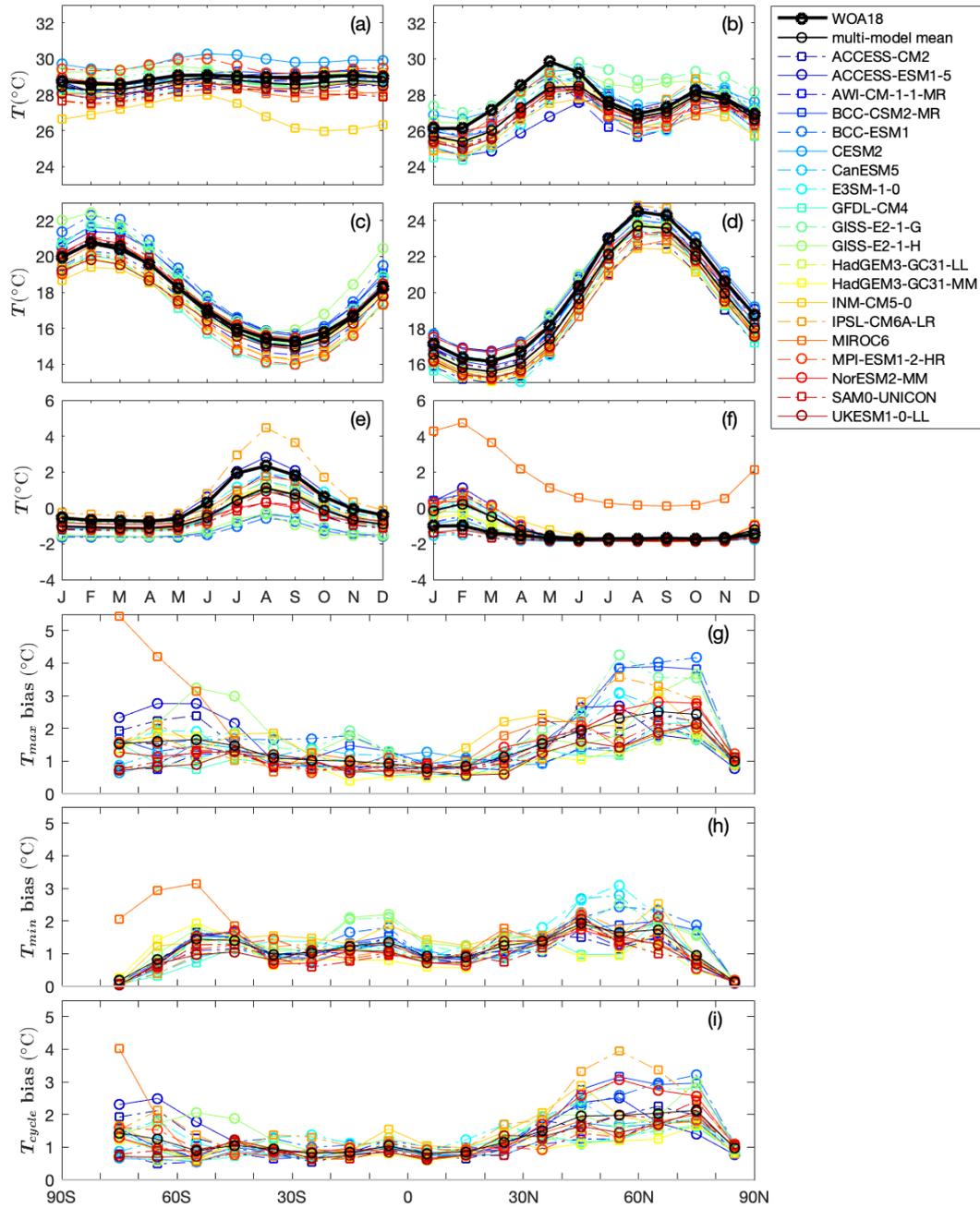
Figure 4.4: Monthly time series of area-weighted mean SST over (a) western equatorial Pacific (5°S - 5°N, 140°E - 160°W), (b) northwestern Indian Ocean (60 - 70°E, 10 - 20°N), (c) subtropical Southern Hemisphere (30° - 40°S), (d) subtropical Northern Hemisphere (30 - 40°N), (e) Arctic (70 - 80°N), (f) Antarctic (70 - 80°S), and area-weighted RMSE in 10° bands for (g) T_{max}, (h) T_{min}, (i) T_{cycle}. Y-axis range is same for (a-f).

Pacific ITCZ. They suggested that this seasonal bias is caused by an easterly wind bias throughout the year in that monsoon region. During winter-spring, the northeastern Pacific ITCZ is dominated by easterly winds, so overly strong easterly winds enhance surface evaporation and lead to cold biases. In contrast, during summer-autumn when westerly winds dominate, the simulated wind is too weak, which causes the warm bias. The northeastern Pacific is a region where tropical cyclones and heatwaves occur (Gilford et al., 2017; Frölicher and Laufkötter, 2018), and hence a warm bias of over 2°C in T_{max} may lead to overprediction of tropical cyclones and heatwaves by these models.

The multi-model mean has a cold bias in T_{max} and a warm bias in T_{min} over the Northwest Pacific, leading to a too small T_{cycle} (bias of more than 2°C) (Figs. 4.1b-d). The warm bias in winter can be seen in many models, especially in ACCESS-ESM1-5, BCC-ESM1, CanESM5 and INM-CM5-0 (Fig. 4.3). Models with a warm bias in T_{min} are likely to have overly intense winter storms, as warm SSTs will increase the storm energy source. Greeves et al. (2007) demonstrated that there was a clear link in the Hadley Centre models between winter SST warm bias to the east of Japan and increased storm intensity in the region, and this is likely to be the case in other models as well. The winter warm bias to the east of Japan was also found in a CMIP5 multi-model mean (Wang et al., 2018), but from our results the warm bias extends further east (Figs. 4.1c).

The large cold biases at northern hemisphere high latitudes in BCC-CSM2-MR, BCC-ESM1, GISS-E2-1-G and GISS-E2-1-H, are typically 2-5°C smaller in T_{min} than in T_{max} (Figs. 4.2, 4.3, 4.4g-h). These cold biases have been previously linked to cloud biases. The negative cloud radiative forcing is excessive in BCC-CSM2-MR (Wu et al., 2019) and BCC-ESM1 (cloud simulation likely to be similar to BCC-CSM2-MR), while overestimated cloud in GISS-E2-1-G and GISS-E2-1-H (Kelley et al., 2020) blocks more of the incoming solar radiation. As solar radiation is negligible at high latitudes in winter, the SST cold bias due to cloud bias is much smaller in winter than in summer, consistent with our results. Deep winter mixed layer depths and SSTs close to freezing likely also contribute to the smaller cold biases in T_{min} than in T_{max} at high latitudes.

In most models there is a warm T_{mean} bias in the Southern Ocean, commonly attributed to excessive short wave radiation linked to underestimated cloud (Hyder et al., 2018). The warm bias is larger for T_{max} than T_{min} (Figs. 4.1b-c, 4.2, 4.3, 4.4g-h), because the lack of incoming solar radiation in winter means cloud biases have minimal effect on surface solar insolation. Shallower mixed layer depths in summer will also tend to enhance any bias in incoming solar insolation. The larger warm bias in T_{max} than T_{min} results in a sea ice extent that is too small in most CMIP6 models, especially in summer (Beadling et al., 2020; Shu et al., 2020). As mode water and intermediate water primarily form within the winter mixed layer on the northern edge of the ACC (Talley, 1999), the T_{min} warm bias can be transferred into ventilated layers of the Southern Ocean. Sallée et al. (2013) indicated that in CMIP5 models the surface warm bias of the Southern Ocean spread over the entire water column.

MIROC6 stands out with the largest warm bias in the Southern Ocean (Figs. 4.2m, 4.3m), with a T_{max} RMSE between 3 and 5°C and a T_{min} RMSE between 2 and 3°C at 50-80° S (Fig. 4.4g). Therefore, the largest biases in MIROC6 occur in the regions where there should be sea ice and where the deep ocean is ventilated. Beadling et al. (2020) found that MIROC6 has the lowest sea ice extent in the Southern Ocean among the CMIP6 models in both summer and winter, and Tatebe et al. (2019) revealed annual warm biases which can exceed 2°C in the intermediate and deep layers in MIROC6.

In eastern boundary upwelling regions (especially the Benguela and Humboldt Currents), most models have a seasonal warm bias that is 1-5°C smaller in T_{max} than in T_{min} (Figs. 4.1b-c, 4.2, 4.3). The warm bias we found in CMIP6 models may lead to excessive precipitation. Rouault et al. (2003) indicated that in the South East Atlantic Ocean off Angola and Namibia the warm events during the T_{max} period increase precipitation along the coast of those countries (sometimes extending inland). Underestimation of cloud, and insufficient upwelling due to overly weak winds, are suggested causes for the warm SST bias (Richter, 2015). Letelier et al. (2009) showed that in the Humboldt Current coastal region the cooling effect of upwelling is strongest in austral summer, which is consistent with the peak of upwelling-favourable wind in December and January. A poor simulation of the seasonal cloud and upwelling processes will contribute to the seasonality of the SST bias in eastern

upwelling boundary regions.

Most models have a seasonal warm SST bias in the eastern equatorial Atlantic (Figs. 4.1b-c, 4.2 and 4.3). The T_{min} multi-model mean bias can be more than 2°C larger than the T_{max} multi-model mean bias. Richter and Tokinaga (2020) showed a similar seasonal warm bias in the CMIP6 multi-model mean, which is about 1-2°C larger during June-July-August than March-April-May. Richter et al. (2012) argued that the warm SST bias in eastern equatorial Atlantic during June-July-August is linked to wind stress errors during March-April-May. GISS-E2-1-G and GISS-E2-1-H have the largest seasonality of SST warm bias in the eastern equatorial Atlantic, with T_{min} biases up to 5°C. Richter and Tokinaga (2020) illustrated that warmer than observed SSTs in the equatorial Atlantic lead to excessive precipitation. Roxy (2014) quantified SST-precipitation relationship: a 1°C SST increase corresponds to a 2 mm/day precipitation increase. Therefore, the 5°C T_{min} warm bias in GISS-E2-1-G and GISS-E2-1-H could cause a 10 mm/day increase in precipitation.

Although the amplitudes of biases are different in T_{max} and T_{min}, the global patterns and signs of T_{max} bias and of T_{min} bias are similar to each other in most models (Figs. 4.2, 4.3). Wang et al. (2014) also indicated that the SST bias of the CMIP5 multi-model mean has a pattern independent of season but did not analyse the seasonality in bias in individual models. Our results show two exceptions: E3SM-1-0 and IPSL-CM6A-LR, which both have an overall warm bias in T_{max}, but an overall cold bias in T_{min} (Figs. 4.2h,t 4.3h,t), which tend to cancel out in the annual means. The T_{max} RMSE is 1.38°C for E3SM-1-0 and 1.36°C for IPSL-CM6A-LR, the T_{min} RMSE is 1.39°C for E3SM-1-0 and 1.21°C for IPSL-CM6A-LR, whereas the T_{mean} RMSE is only 1.17 °C for E3SM-1-0 and 0.94°C for IPSL-CM6A-LR. In E3SM-1-0, the global annual average mixed layer depth is generally too shallow (Golaz et al., 2019), which can contribute to the summer SST warm bias and winter SST cold bias, and a similar process may be affecting IPSL-CM6A-LR. These results illustrate the risks involved in assessing only the annual mean values, as a model may have greater biases than assumed, which may mean that tropical cyclone formation, for example, is overpredicted in these models.

In mid-latitudes the SST seasonal cycle is well represented by an annual sinusoid

whereas in equatorial and polar regions an annual sinusoid explains little of the total SST seasonal variance (Trenberth, 1983; Yashayaev and Zveryaev, 2001). In regions with fairly sinusoidal SST annual cycles such as the subtropics, models have realistic SST seasonal cycles with well simulated amplitude and phase of the annual cycle (Figs. 4.4c-d). Phase biases are mainly within 1 month. In subtropical regions, the seasonal SST biases are consistent with biases in T_{mean}. Differences between the T_{max} and T_{min} biases are smaller than those in non-sinusoidal regions (Fig. 4.4). In regions with non-sinusoidal SST seasonal cycles such as the western equatorial Pacific, northwestern Indian Ocean, the Arctic and the Antarctic, models tend to have biases in amplitudes or phases of their SST seasonal cycles (Figs. 4.4a-b,e-f).

In the western equatorial Pacific, the SST seasonal cycle in WOA18 is modest (less than 1°C), whereas in some models such as MPI-ESM1-2-HR, GISS-E2-1-G, GISS-E2-1-H and especially INM-CM5-0 the seasonal cycle is much larger (Fig. 4.4a). In INM-CM5-0, the range of SST seasonal cycle is about 2°C and there is a cold SST bias throughout the year, reaching 3°C during September-October-November (Fig. 4.4a). Similar to our analysis, Volodin et al. (2017) noted that INM-CM5-0 has a cold bias of more than 4°C in annual mean temperature in the upper 700 m of the western equatorial Pacific. The cold bias could limit the skills of models in simulation of ENSO and ENSO induced teleconnections, resulting in a too westward located rising branch of the Walker circulation with a weak convective response (Bayr et al., 2018) and a significant underestimation of ENSO-related precipitation anomaly over California (Bayr et al., 2019). Jiang et al. (2021) showed that in the equatorial western Pacific INM-CM5-0 has an overly weak ENSO-related SST bias, whereas the most other CMIP6 models have an overly strong SST bias, which may be linked to our finding that INM-CM5-0 has the largest SST cold bias.

In the northwestern Indian Ocean where the monsoon system prevails, SST has a semi-annual cycle, but most models are unable to reproduce this with the correct amplitude and phase (Fig. 4.4b). Most CMIP6 models have SST cold biases in this region throughout the year, while the biases are generally larger during March-April-May than other months and the multi-model mean fails to simulate the primary maximum SST (Fig. 4.4b). As the northwestern Indian Ocean is an important moisture source

for Indian monsoon precipitation (Prodhomme et al., 2014; Levine and Turner, 2012), the cold biases in the CMIP6 models could lead to overly weak monsoon precipitation. Consistent with our result, McKenna et al. (2020) found a cold SST bias over the northwestern Indian Ocean in the CMIP6 multi-model mean. Fathrio et al. (2017) showed that the SST cold bias over the western Indian Ocean in the CMIP5 multi-model mean has a seasonal cycle with the coldest SST bias occurring in April, whereas the coldest SST bias in our CMIP6 multi-model mean occurs in May. GISS-E2-1-G and GISS-E2-1-H fail to simulate a realistic second minimum SST in August (Fig. 4.4b), which would lead to overly intense tropical cyclones. SST in the northwestern Indian Ocean is closely related to the onset of the summer monsoon (Sijikumar and Rajeev, 2012; Jiang and Li, 2011). The timing of the primary maximum SST is two months later in ACCESS-ESM1-5 than in WOA18 (Fig. 4.4b), which may suggest a delay of the summer monsoon onset date in projections using that model.

4.4.2 Impact of model characteristics on SST seasonal extrema

We have shown that biases in T_{max}, T_{min} and T_{cycle} are different between models. We now use the diversity in the 20 CMIP6 models to explore the effects of different model characteristics on the magnitude of these biases as quantified by global area weighted RMSE for T_{max}, T_{min}, T_{cycle} and T_{mean}.

No significant correlation was found between the models' seasonal biases and horizontal ocean resolution, demonstrated by the lack of a relationship with the symbol size in Figs. 4.5 and 4.6. Chassignet et al. (2020) used four pairs of matched low-resolution and high-resolution ocean simulations from FSU-HYCOM, AWI-FESOM, NCAR-POP and IAP-LICOM to isolate the effect of ocean horizontal resolution, and compared their representation of global SST. They found that enhanced horizontal resolution does not deliver unambiguous SST bias improvement in all regions for all models, which is consistent with our finding. Nor did we find any correlation of seasonal biases with atmospheric resolution, ocean grid type, ocean vertical coordinate, and inclusion (or not) of biogeochemical processes (circles or squares in Figs. 4.5 and 4.6).

The only characteristic yielding a statistically significant relationship was the ocean
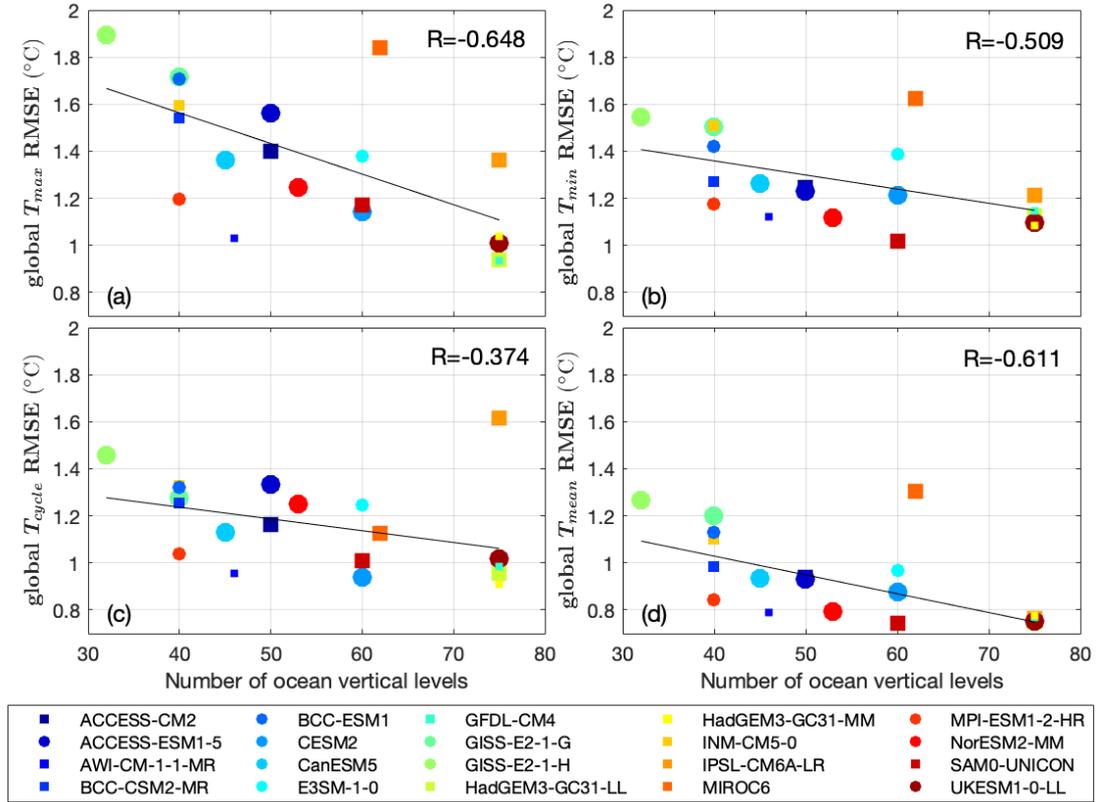
Figure 4.5: Global RMSE of (a) T_{max}, (b) T_{min}, (c) T_{cycle} and (d) T_{mean}, all against the total number of vertical levels in ocean. Circles represent earth system models, while squares represent non earth system models. The size of the markers represents the ocean horizontal resolution for that model, with larger markers for models with lower horizontal resolution. The black line is the line of best fit (with the least sum of squared errors). The inter-model correlation R is shown on each panel.

vertical resolution, for which we use as proxies the total number of vertical levels and top grid cell thickness (table. 4.1). The number of vertical levels in the upper ocean (e.g. upper 200 m) cannot be unambiguously determined for models using an isopycnal or sigma vertical coordinate (6 out of 20 in our study) as their level depths vary with location and time (Bleck, 2002; Shchepetkin and McWilliams, 2005). Excluding the isopycnal and sigma models, the remaining high vertical resolution models are mainly from the Met Office Hadley Centre family, and hence any relationship between SST biases and vertical resolution in the upper ocean might have been overly influenced by that particular family.

For the 20 models, there is a decrease in bias with increasing total number of vertical levels (Fig. 4.5). Linear regression was performed for T_{max}, T_{min}, T_{cycle} and T_{mean}. We also calculated the inter-model correlation between global RMSE and total number
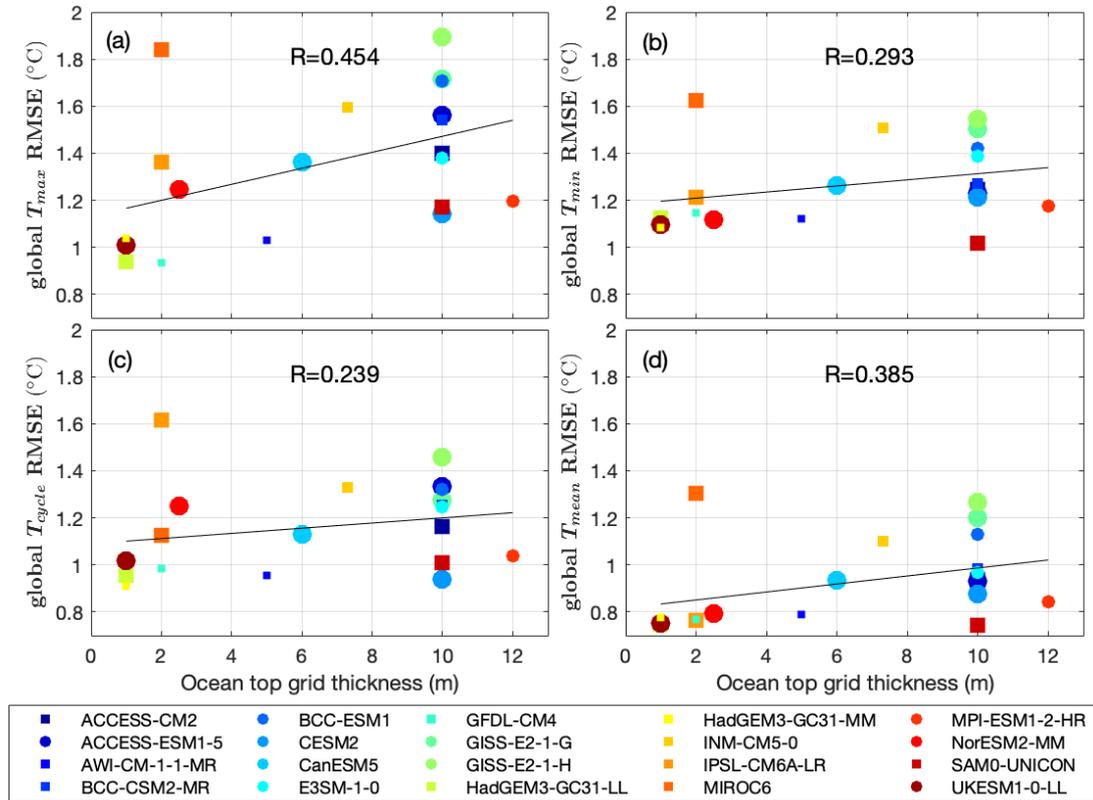
Figure 4.6: As Fig. 4.5, but against the thickness of top grid in ocean.

of vertical levels following the method of Wang et al. (2014). The correlations are significant in T_{max}, T_{min}, and T_{mean}, with the highest correlation of -0.648 in T_{max}. The higher correlation between global T_{max} RMSE and ocean vertical resolution is likely linked to a shallower mixed layer depth in summer than in winter. The impact of ocean vertical resolution can also be seen from RMSE against top grid thickness (but with lower inter-model correlation relative to total number of vertical levels): models with a smaller top grid thickness tend to have smaller biases (Fig. 4.6).

The importance of vertical resolution for reducing seasonal biases is not unexpected; SST is influenced by ocean stratification and ocean vertical mixing processes, whose representation depends upon the vertical resolution. It has been found that high resolution in the upper ocean is important for the representation of diurnal and intraseasonal SST variability in ocean general circulation models (Misra et al., 2008; Xavier et al., 2008; Ge et al., 2017). Our study emphasises the importance of vertical resolution for simulating seasonal extreme SST and annual mean SST.

4.5 Conclusions

Using the newly-released CMIP6 models, this study provides a global view of the biases in SST extrema, identifies regions with large seasonal bias, and suggests a future direction to reduce these biases. Global area-weighted T_{max}, T_{min} and T_{cycle} RMSEs are typically 1-2°C. Most models have T_{max} and T_{min} biases of the same sign at most grid points, apart from IPSL-CM6A-LR and E3SM-1-0 which have an overall warm bias in T_{max} and an overall cold bias in T_{min}. When averaged across the whole globe, the bias in T_{mean} is typically consistent with T_{max} and T_{min} biases, but certain regions (eastern boundary upwelling regions, polar regions, the eastern equatorial Atlantic, the North Pacific) show significant differences between winter and summer biases. In regions with non-sinusoidal SST seasonal cycles, models tend to have biases in amplitudes and/or phases of their SST seasonal cycles. For the models we examined, those with increased vertical resolution in the ocean generally had a better representation of SST extrema, particularly T_{max}. This is likely related to the ability of the higher resolution models to better represent the surface mixed layer, and particularly shallow mixed layers in summer. For improving the accuracy of future climate projections, we suggest that as much priority (or possibly more) should be given to increasing vertical ocean model resolution as is given to increasing horizontal resolution.

4.6 Supplementary material

This section presents supplementary material, including some figures and detailed methodology that was not shown in the paper submitted to Environmental Research Letters.

4.6.1 Methodology

To quantify the performance of CMIP6 models for SST seasonal extrema, the global area-weighted average of SST biases $\overline{T}$ is calculated using equation 4.1 for each model.

$$\overline{T} = \frac{\sum\limits_{i=1}^{n} T_i \cdot S_i}{\sum\limits_{i=1}^{n} S_i}, \tag{4.1}$$

where T_i is the SST bias (that is model minus observation) at grid point. S_i is the area for the grid box. n is the total number of grids.

From the global area-weighted average of SST biases, we can tell if a model has an overall warm bias or cold bias. However, it is hard to quantify the amplitude of the global bias as the warm and cold biases in different regions may compensate for each other. To solve this problem, we also calculated the global area-weighted root mean square (RMS) of the SST biases T_{rms} using equation 4.2.

$$T_{rms} = \sqrt{\frac{\sum\limits_{i=1}^{n} T_i^2 \cdot S_i}{\sum\limits_{i=1}^{n} S_i}}, \tag{4.2}$$

To study the relationship between SST bias and model characteristics, linear regression was performed. The best linear fit was obtained using least-squares. The correlation coefficient R is calculated to study the inter-model correlation between SST bias and the model characteristics, following the method of Wang et al. (2014).

$$R = \frac{\sum(T_{rms_j} - \overline{T_{rms}})(C_j - \bar{C})}{\sqrt{\sum(T_{rms_j} - \overline{T_{rms}})^2 \sum(C_j - \bar{C})^2}}, \tag{4.3}$$

where R is the correlation coefficient, T_{rms_j} is the RMS bias for model j, C_j is the value of the chosen model characteristic for model j. $\overline{T_{rms}}$ is the multi-model mean of T_{rms_j}, $\bar{C}$ is the multi-model mean of C_j.

4.6.2 Seasonal cycle of SST in CMIP6 models

The SST seasonal cycle T_{cycle} in models is compared with WOA18 to obtain the map of T_{cycle} bias for each model (Fig. 4.7). The difference between bias in T_{max} and T_{min}

causes T_{cycle} bias (Fig. 4.7). Large T_{cycle} biases mean that T_{max} and T_{min} biases are very different from T_{mean}, and hence we cannot evaluate the simulation of SST seasonal extrema just based on T_{mean}. IPSL-CM6A-LR has the largest global area-weighted T_{cycle} RMSE (1.62°C) as T_{max} has an overall warm bias and T_{min} has an overall cold bias (Figs. 4.2t, 4.3t, 4.7t). Therefore, when using IPSL-CM6A-LR for projection, precipitation and storms could be overestimated in winter and underestimated in summer. In the subpolar regions of the Northern Hemisphere, IPSL-CM6A-LR has a T_{cycle} bias of more than 4°C (Fig. 4.7t). T_{cycle} biases at northern hemisphere high latitudes are large in BCC-CSM2-MR, BCC-ESM1, GISS-E2-1-G and GISS-E2-1-H, as the cold bias there is large in T_{max} but small in T_{min} (Figs. 4.2p-s, 4.3p-s, 4.7p-s). Therefore, the projected precipitation can be too small compared to observation, especially in summer. In the eastern boundary upwelling regions, all models have seasonal cycles that are too small because the warm bias is smaller in T_{min} than in T_{max}.

The global area-weighted average of bias in most models has the same sign (warm or cold) in both T_{max} and T_{min}, where as IPSL-CM6A-LR and E3SM-1-0 have opposite signs in T_{max} and in T_{min}, leading to a too large T_{cycle} (Fig. 4.8). In models except IPSL-CM6A-LR and E3SM-1-0, the sign of the bias in T_{max} and T_{min} is the same as that in T_{mean}.

In most of the models, the global RMSE is larger in T_{max} than in T_{min} (Fig. 4.9). T_{mean} has a smaller RMSE than T_{max} and T_{min}, as SST biases with opposite signs in summer and winter compensate each other when calculating the annual mean. As the bias in T_{max} and T_{min} is largely consistent with T_{mean} bias (Figs. 4.10a-b), T_{cycle} RMSE is small compared with T_{max} and T_{min} RMSEs. Unlike T_{max} and T_{min}, T_{cycle} does not have RMSE that largely consistent with T_{mean} RMSE (Figs. 4.10c). For T_{max} and T_{min}, GISS-E2-1-H and MIROC6 have the largest RMSE among models, while GFDL-CM4 and HadGEM3-GC3-1-LL have the smallest RMSE. For T_{cycle}, IPSL-CM6A-LR has the largest RMSE among models, although its RMSE in T_{max} and T_{min} is relatively small compared with other models.

T_{cycle} biases (Fig. 4.7) indicate differences between Tmax biases and Tmin biases.
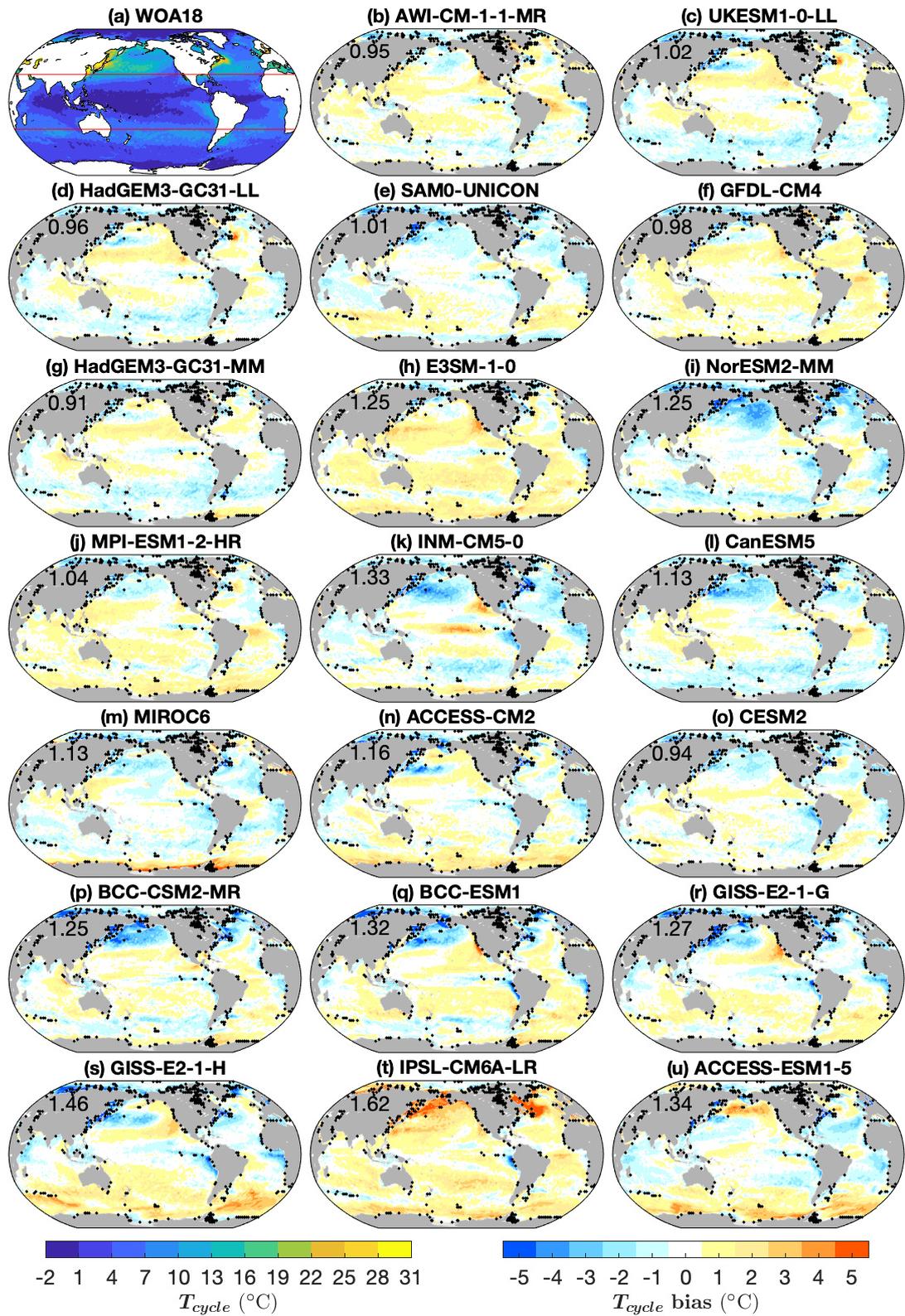
Figure 4.7: As Fig. 4.2 , but for T_{cycle}.
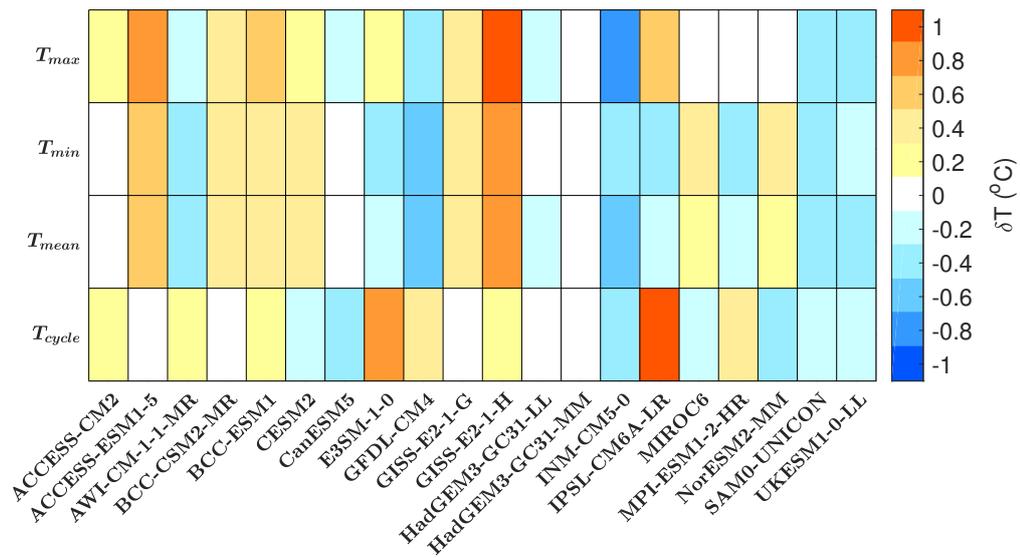
Figure 4.8: The global area-weighted average of the biases in T_{max}, T_{min}, T_{mean} and T_{cycle}.
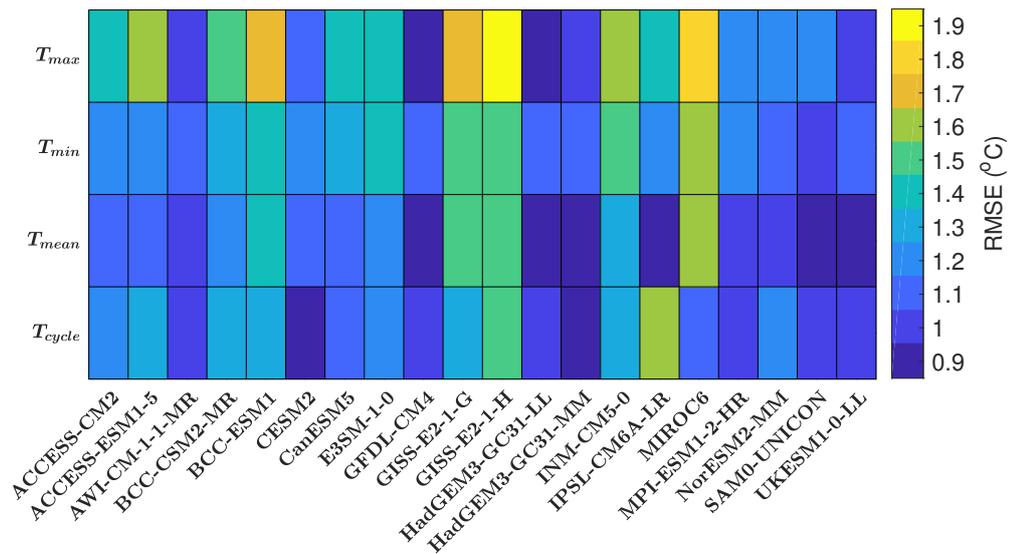
Figure 4.9: The global area-weighted RMSE of the biases in T_{max}, T_{min}, T_{mean} and T_{cycle}.

For example, IPSL-CM6A-LR has warm SST biases in summer and cold SST biases in winter (Figs. 4.2, 4.3), and its global RMSE in Tcycle is the largest (1.62°C) among the models (Fig. 4.7). For models with large Tcycle biases, the biases in their projections of climate variables (e.g. heatwave, storm, precipitation) can be very different between in summer and in winter.

Different biases in T_{max}, T_{min}, T_{cycle} and T_{mean} (Figs. 4.8, 4.9) suggest that models
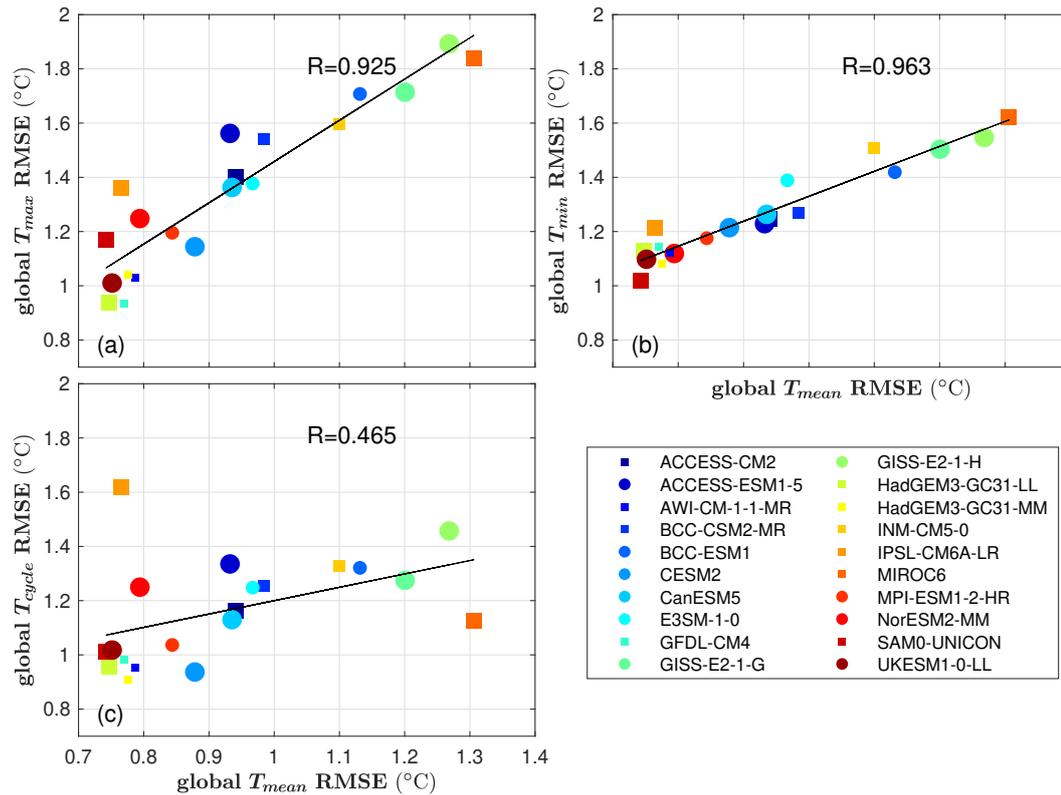
Figure 4.10: The global area-weighted RMSE of the biases in (a) T_{max}, (b) T_{min}, (c) T_{cycle}, all against the global area-weighted T_{mean}. The correlation R is shown on each panel.

have different performance in simulating SST seasonal variation and annual mean. The "best" and "worst" models depend on whether you choose SST seasonal extrema or annual mean as your metric. The best choice of models depends on whether you want to simulate phenomenons related to T_{max} or T_{min} or T_{mean}. For example, GFDL-CM4 and HadGEM3-GC31-MM are best for simulating tropical cyclones and heatwaves; SAM0-UNICON is best for simulating the properties of intermediate and deep waters.

4.6.3 The impact of model characteristics on SST seasonal extrema

The impact of oceanic vertical resolution at low latitudes and mid-high latitudes

According to the variation with latitudes of SST biases (Fig. 4.4) and maximum mixed layer depth (Fig. 4.11), the globe can be broken down into low latitudes and mid-high latitudes, using 30° as the boundary. At low latitudes, the amplitude of SST bias is

small and the maximum MLD is less than 200 m; at mid-high latitudes, the amplitude of SST bias becomes larger and the maximum MLD is greater than 200 m in the North Atlantic, north western Pacific and along the ACC.
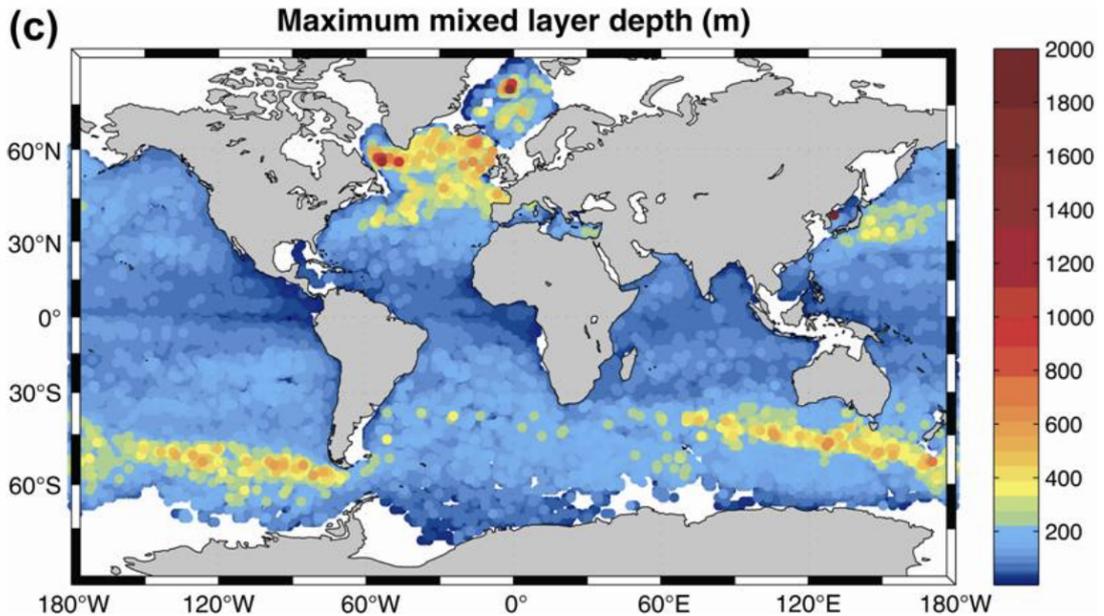
Figure 4.11: Averaged maximum MLD from the Argo profiling float data set (2000–2009), adapted from Talley (2011).

The significant relationship between the RMSE and ocean vertical levels exists both at mid-high latitudes and low latitudes (Fig. 4.12). The area-weighted RMSE of T_{max} is larger at mid-high latitudes than at low latitudes (Fig. 4.12a,b), as shown in Fig. 4.4a.

Unlike at mid-high latitudes, the sensitivity of bias to vertical resolution is similar for T_{max} (-0.012°C per level) and T_{min} (-0.008°C per level) at low latitudes (Fig. 4.12b,d). The reason might be that low latitudes are ice free and the MLD there is less seasonal. Furthermore, the impact of number of ocean vertical levels on T_{cycle} bias at low latitudes is weak (only -0.002°C per level), as the amplitude of SST seasonal cycle is small in equatorial regions (Figs. 4.4a).

As shown in Figs. 4.5, 4.12, the better vertical resolution could lead to a better simulation. To see the geographical structure to the improvements, the analysis is expanded to each grid point. We calculated inter-model correlations between local SST biases and ocean vertical resolutions.

T_{max} biases over most of the world ocean tend to be reduced when increasing the ocean
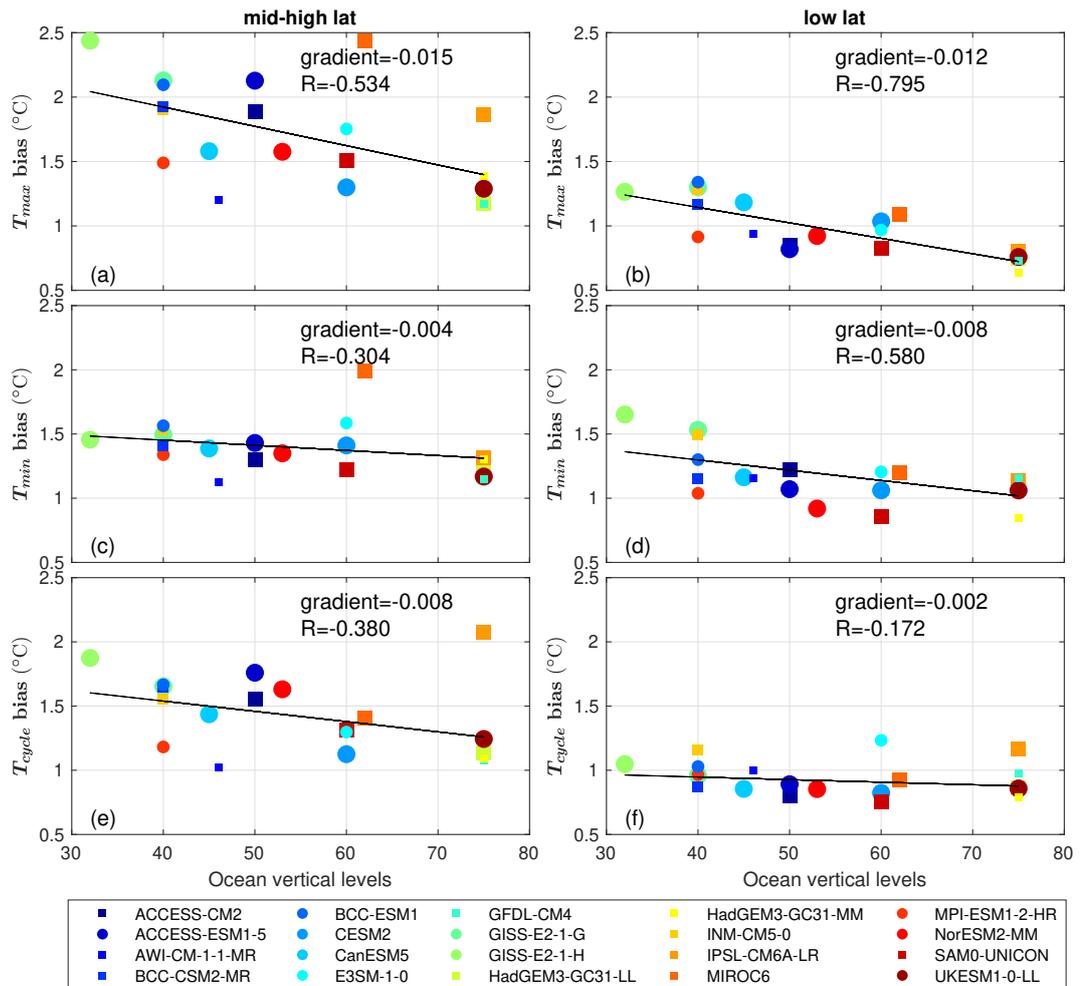
Figure 4.12: The area-weighted SST RMSE against total vertical levels for (a-b) mid-high latitudes (30°-90°) and (c-d) low latitudes 30°S-30°N. Black line is the best linear fit for the scatters. The gradient of the linear regression and inter-model correlation R are shown on each panel.

vertical resolutions (Fig. 4.13a). The inter-model correlation is largest in the eastern Pacific and Atlantic, likely due to that the mixed layer is shallower in those regions (Fig. 4.14). When the mixed layer is shallower, a realistic representation of MLD is more important for the simulation of SST.

In the eastern South Atlantic and Pacific, and tropical Pacific, T_{min} biases is sensitive to ocean vertical resolution as the inter-model correlation between the T_{min} bias and total number of ocean vertical levels is smaller than -0.6 (Fig. 4.13b). However, there are some regions (e.g. subtropical North and South Pacific, South Atlantic, eastern Indian Ocean) with positive inter-model correlation, where T_{min} bias decreases with vertical resolution.
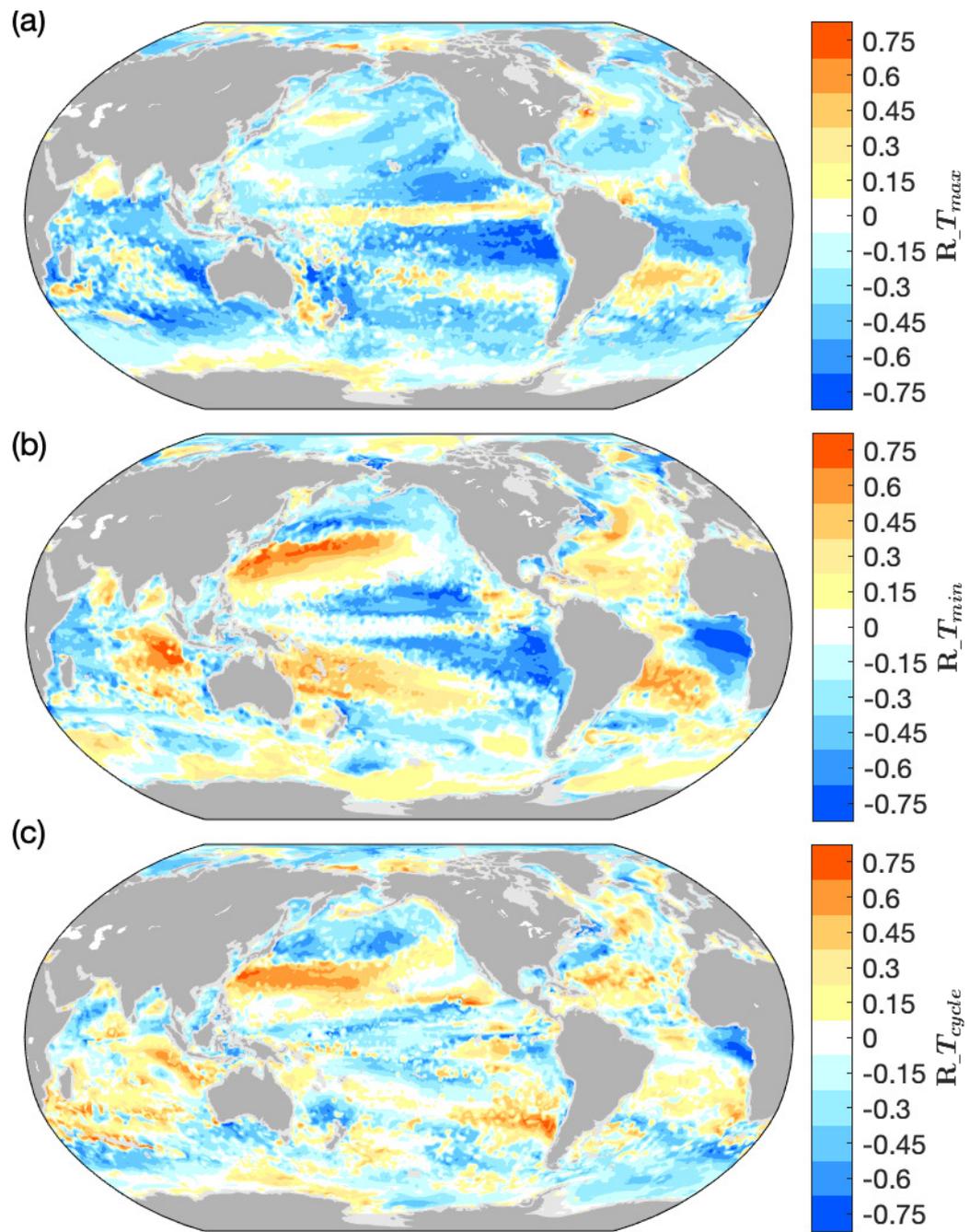
Figure 4.13: Inter-model correlation between the local bias and the total number of ocean vertical levels for (a)T_{max}, (b)T_{min} (c)T_{cycle}.

As for T_{cycle}, the geographical pattern of inter-model correlation is more complicated (Fig. 4.13c). The T_{cycle} bias in the subtropical North Pacific increases with vertical resolution, while the T_{cycle} bias in the tropical eastern Atlantic decreases with vertical resolution. The inter-model correlation is small over most of the world ocean.
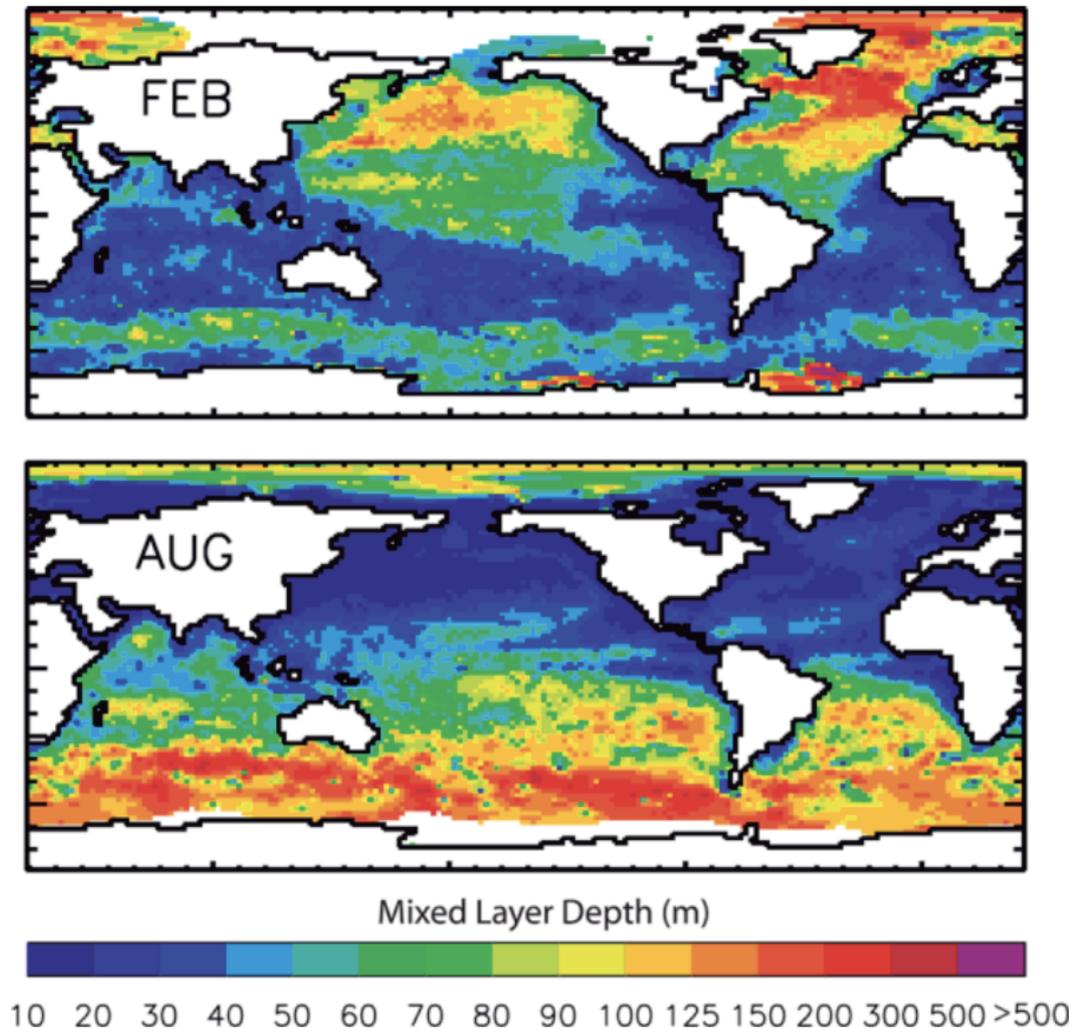
Figure 4.14: Global mixed layer depth climatology in February and August. Adapted from de Boyer Montégut et al. (2004).

The impact of other model characteristics on SST biases

In addition to ocean vertical levels, we also examined the impact of other model characteristics (table. 4.2) on the magnitude of SST biases. No significant correlation was found between those model characteristics (ocean horizontal resolution, atmosphere horizontal and vertical resolutions) and RMSE for T_{max}, T_{min}, T_{mean} and T_{cycle} (Figs. 4.15, 4.17, 4.16).

Model	Ocean Horizontal resolution	Ocean grid	Ocean vertical coordinate	Ocean total levels	Top grid thickness	Atmosphere horizontal resolution	Atmosphere vertical levels	Earth system model	References
ACCESS-CM2	100 km	tripolar	z^*	50	10 m	250 km	85 levels	N	Bi et al. (2020)
ACCESS-ESM1-5	100 km	tripolar	z^*	50	10 m	250 km	38 levels	Y	Law et al. (2017)
AWI-CM-1-1-MR	25 km	unstructured meshes	z-σ	46	5 m	100 km	95 levels	N	Semmler et al. (2020)
BCC-CSM2-MR	50 km	tripolar	z	40	10 m	100 km	46 levels	N	Wu et al. (2019)
BCC-ESM1	50 km	tripolar	z	40	10 m	250 km	26 levels	Y	Wu et al. (2020)
CESM2	100 km	displaced pole	z	60	10 m	100 km	32 levels	Y	Danabasoglu et al. (2020)
CanESM5	100 km	tripolar	z	45	6 m	100 km	49 levels	Y	Swart et al. (2019)
E3SM-1-0	50 km	unstructured meshes	z^*	60	10 m	100 km	72 levels	Y	Golaz et al. (2019)
GFDL-CM4	25 km	tripolar	z^*-ρ	75	2 m	100 km	33 levels	N	Held et al. (2019)
GISS-E2-1-G	100 km	regular lat-lon	z	40	10 m	250 km	40 levels	Y	Kelley et al. (2020)
GISS-E2-1-H	100 km	tripolar	z-ρ-σ	32	10 m	250 km	40 levels	Y	Kelley et al. (2020)
HadGEM3-GC31-LL	100 km	tripolar	z^*	75	1 m	250 km	85 levels	N	Andrews et al. (2020)
HadGEM3-GC31-MM	25 km	tripolar	z^*	75	1 m	100 km	85 levels	N	Andrews et al. (2020)
INM-CM5-0	50 km	displaced pole	σ	40	7.3 m	100 km	73 levels	N	Volodin et al. (2017)
IPSL-CM6A-LR	100 km	tripolar	z^*	75	2 m	250 km	79 levels	N	Boucher et al. (2020)
MIROC6	100 km	tripolar	z-σ	62	2 m	250 km	81 levels	N	Tatebe et al. (2019)
MPI-ESM1-2-HR	50 km	tripolar	z	40	12 m	100 km	95 levels	Y	Müller et al. (2018)
NorESM2-MM	100 km	tripolar	ρ	53	2.5 m	100 km	32 levels	Y	Seland et al. (2020)
SAM0-UNICON	100 km	displaced pole	z	60	10 m	100 km	30 levels	N	Park et al. (2019)
UKESM1-0-LL	100 km	tripolar	z^*	75	1 m	250 km	85 levels	Y	Sellar et al. (2019)

Table 4.2: The 20 CMIP6 models used in this study; the horizontal resolution of their ocean component; ocean grid; ocean vertical coordinate (symbols are the same as table 4.1); total number of ocean vertical levels; thickness of the ocean top grid; the horizontal resolution of their atmosphere component; total number of atmosphere vertical levels; earth system model or non earth system model; and references.
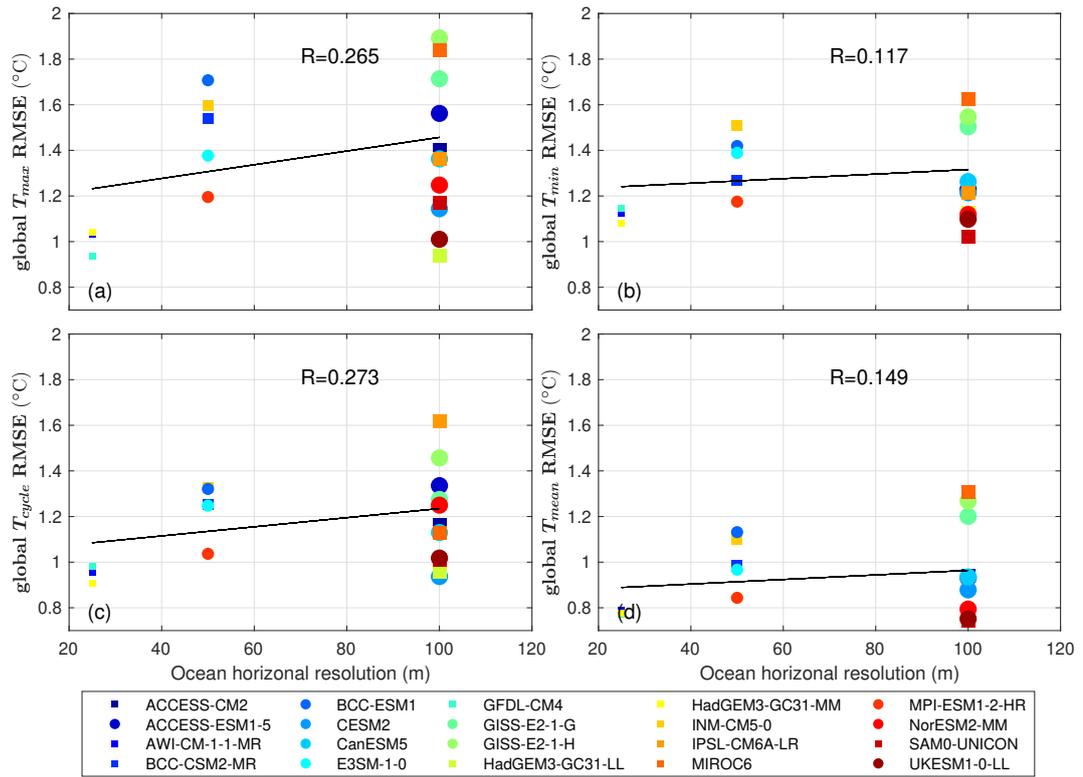
Figure 4.15: As Fig. 4.5, but for ocean horizontal resolution.
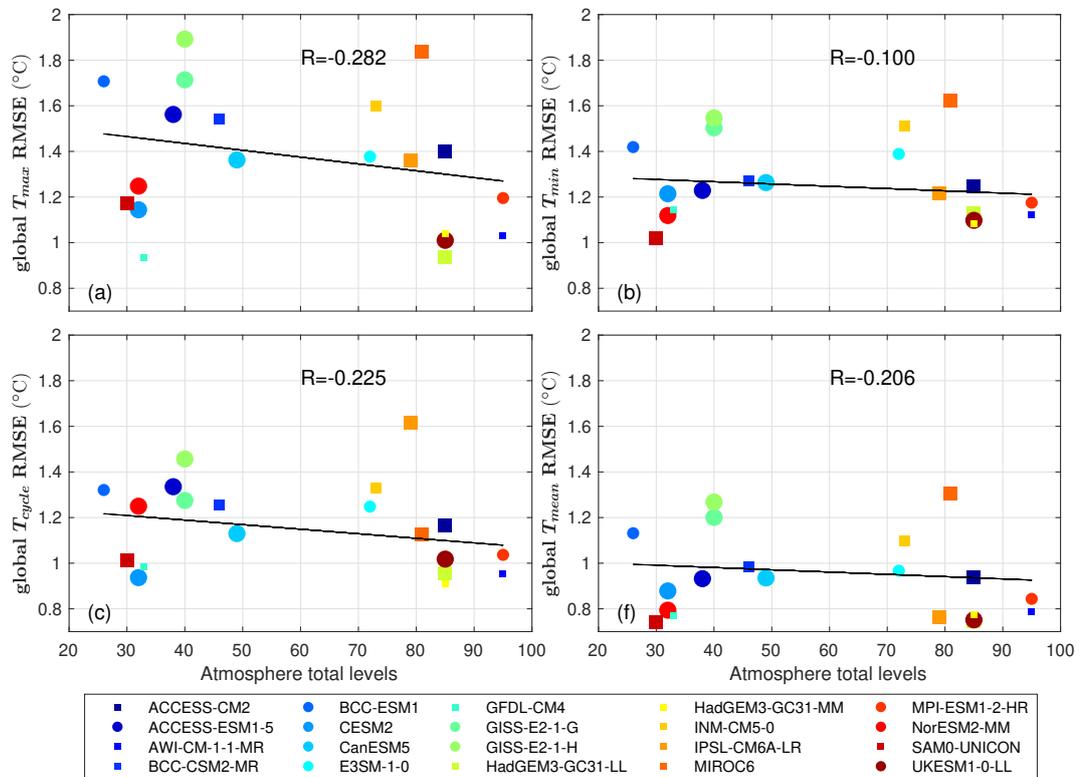
Figure 4.16: As Fig. 4.5, but for atmosphere vertical levels.
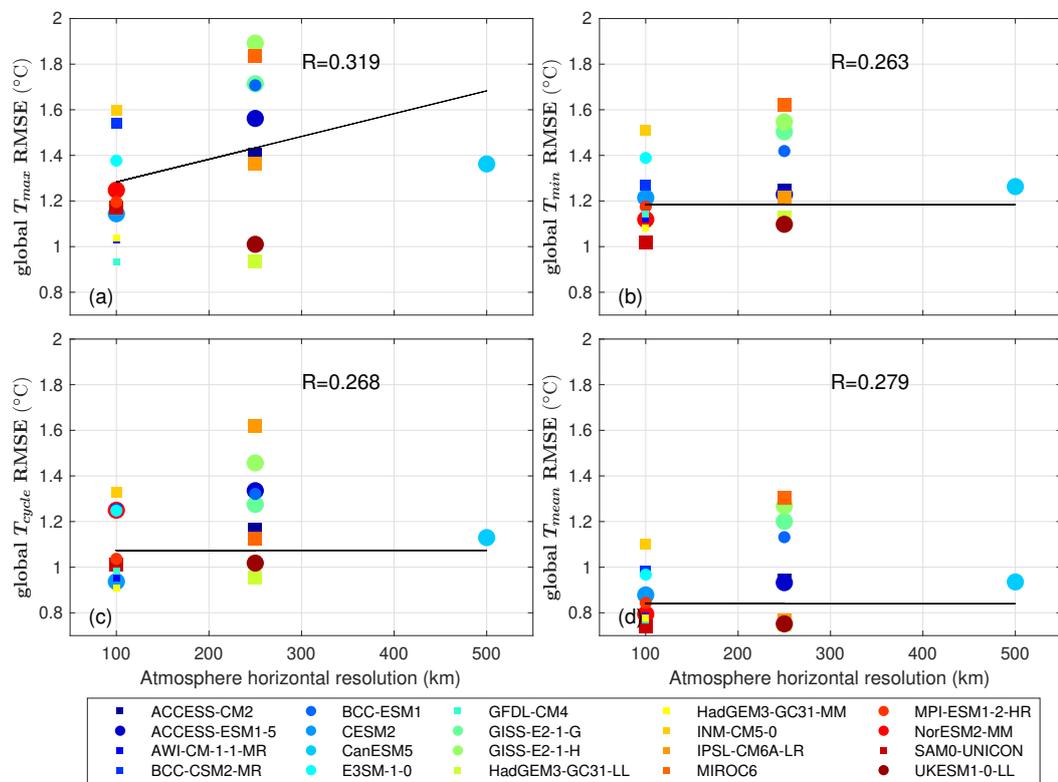
Figure 4.17: As Fig. 4.5, but for atmosphere horizontal resolution.

5

Discussion and conclusions

5.1 Why does the definition of seasonal extrema matter?

To examine the seasonal cycle of SST, we picked the month when local seasonal SST maxima/minima occur to represent summer and winter, and the monthly maxima/minima are defined as the seasonal extrema. However, our definition is different from most studies. Some pick specific months (e.g. August and February) or number of months (e.g. June-July-August and December-January-February) to represent summer and winter, while some use a yearly-period sinusoid and its associated amplitude and phase to describe the SST seasonal cycle.

In many regions (e.g. mid latitudes) the SST seasonal cycle is well represented by a sinusoid, however this is not the case in equatorial regions, monsoon regions and polar regions. Dwyer et al. (2012) and Stine et al. (2009) described seasonal cycle of surface temperature by the phase and amplitude of a sinusoid with a period of 1 year, but they had to exclude the regions where an annual sinusoid explains little of the total seasonal variance. However, the definition of seasonal extrema in this thesis can be applied to the whole world ocean, even in areas that are not dominated by an annual sinusoid.

SST seasonal extrema tend to occur in August and February in most of the world ocean but there are also exceptions. In monsoon regions, SST has a semi-annual cycle and the extrema do not occur in August or February. In polar regions, winter SST is often at or close to freezing, thus seasonal minimum may last for many months. Even at mid-high latitudes where the SST seasonal cycle is well represented by an annual sinusoid, the seasonal extrema may not occur exactly in August/February. For

example, the SST seasonal minimum occurs in September instead of August in the subtropical Southern Hemisphere (30-40°S). Hieronymus et al. (2014) studied global water mass formation rates in an ocean model using monthly averaged SST and SSS in August and February. However, August/February are not always the coldest months in southern/northern hemispheres when the strongest formation occurs. Wang et al. (2014) calculated the CMIP5 multi-model mean SST biases in March-April-May, June-July-August, September-October-November and December-January-February, and demonstrated that the SST biases have seasonally independent patterns and seasonal varying amplitudes. If using the definition of SST seasonal extrema in this thesis, Wang et al. (2014) may found that the bias patterns vary seasonally and the seasonality of SST biases amplitude is actually larger. Similarly, Liu et al. (2020) and Chen and Wang (2015) used January-February-March mean and July-August-September mean SSTs to study the seasonal variance of SST, which may be larger when using the monthly maximum/minimum SSTs as in this thesis.

However, it should also be noticed that uncertainty in observation may lead to errors in finding the timing of SST maxima/minima. To be specific, when the observation uncertainty is larger than the difference between SST maxima/minima and the second maximum/minimum SST, errors are likely to exist in the timing of SST maxima/minima we found.

For the future, we suggest using monthly maxima/minima to describe the SST seasonal cycle and using the month when local seasonal SST maxima/minima occur to represent summer and winter, especially in the equatorial region, polar region and monsoon region. This definition can help us better study the range of SST seasonal cycle and seasonal extrema, which are important to water mass formation, tropical cyclone formation, heatwave and sea ice extent. For example, in northwestern Indian Ocean where tropical cyclone occurs, SST in August or SST averaged in July-August-September is 26-27°C, whereas the T_{max} reaches 30°C in May (Fig. 4.4b). SST threshold below which tropical cyclones do not form was commonly proposed to be 26-27°C (Tory and Dare, 2015). Therefore, using SST in August or SST averaged in July-August-September rather than T_{max} can lead to underpredicted

tropical cyclones formation.

5.2 Uncertainty of observational climatologies

By comparing observational climatologies, we obtained SST differences larger than 2°C and SSS differences larger than 0.5 in some regions between climatologies. It demonstrates that there is some uncertainty of SST/SSS in climatologies due to sparse sampling, instrumental error, quality control or gridding techniques. Therefore, we need to pay extra attention when using observational climatologies in the regions with high uncertainty.

The SST difference between climatologies is largest (more than 2°C) in regions with strong SST gradient, such as the Gulf Stream, Kuroshio Extension, Brazil Current, ACC and coastal upwelling regions. The techniques used to construct climatologies, such as interpolation of sparse observational data onto a regular grid, and averaging over many years, will tend to smooth SST gradients. It seems resonable to infer that the differences between climatologies with high SST gradients may be due to differences in their interpolation/gridding/averaging methodologies. Most model evaluation works suggested model biases based on only one observational climatology. Lauer and Hamilton (2013) calculated SST biases in CMIP5 models in comparison with the National Oceanic and Atmospheric Administration (NOAA) optimum interpolation daily SST analysis, and found warm SST biases of 1-3°C in eastern boundary upwelling regions. Considering the uncertainty of observation, in some locations part of these biases may be attributed to biased reference data.

There also exist large differences between climatologies in polar regions, likely due to the lack of observations in ice-covered regions and different time periods. The SST field constructed in areas without observational data can be unrealistic. At some locations in the Antarctic, the SST of WOA13/WOA18 peaks in July - austral winter, which rarely happens in reality. Therefore, in those areas the SST biases of HiGEM and CMIP6 models based on WOA18 cannot be trusted.

The SSS in climatologies has differences of more than 0.5 in much of the Arctic, which

indicates high uncertainty of SSS in the Arctic. Dong et al. (2020) suggested a saline SSS bias exceeding 1.5 in the Beaufort Sea of the Arctic in CAS-ESM2.0, using WOA13. Shaffrey et al. (2009) showed a saline bias of up to 3 in the Arctic in HiGEM, using World Ocean Atlas 2001 (WOA2001). Both of those saline biases could be from the uncertainty of the reference climatologies. As most of the Arctic is much more fresher in WOA18 than in WAGHC (Figs. 2.27, 2.29), and the previous versions of WOA18 (WOA2013 and WOA2001) are likely to have similar values with WOA18, the saline biases in Shaffrey et al. (2009) and Dong et al. (2020) could be reduced or even convert into fresh biases when using WAGHC as a reference.

For the future, we suggest that not taking observational climatologies as the "truth" and thinking carefully before picking an observational climatology as reference. Extra attention needs to be paid when using a climatology in regions with large uncertainty. To avoid the uncertainty of climatologies, it might be necessary to compare several climatologies and mask areas with large difference between climatologies.

In this thesis, we also compared climatologies with different time periods and found that due to climate change the choice of time period can make a difference. Therefore, we recommend that when doing model evaluation, make sure the model simulations and observational climatologies have the same (or nearly the same) time period.

5.3 Model biases in the seasonal cycle of sea surface water characteristics

The seasonal extrema of SST and SSS affect the density extrema of sea surface water, which determines the occurrence of subduction. The evolution of surface water masses to intermediate and deep water masses is important for global heat, freshwater, carbon budgets and thermohaline ocean circulation. However, most climate model historical run evaluations focus on annual or longer-term mean SST/SSS, and the SST/SSS seasonal extrema have not been evaluated globally. In this thesis we evaluated SST seasonal extrema in 20 CMIP6 models, plus SST and SSS seasonal extrema in HiGEM, and demonstrated the importance of evaluating

model performance not simply against annual mean properties.

We found that T_{max} and T_{min} biases are largely consistent with the annual mean SST biases, but in certain regions in some models T_{max} and T_{min} biases are different in both amplitude and spatial patterns. T_{max} biases are generally larger than T_{min} biases, especially at mid-high latitudes. Wang et al. (2014) showed that CMIP5 multi-model mean SST biases have spatial patterns independent of seasons. However, a multi-model mean does not reflect performance of individual models, my study found that spatial pattern of SST biases varies seasonally in some models. Specifically, IPSL-CM6A-LR and E3SM-1-0 have an overall warm bias in T_{max} but an overall cold bias in T_{min}. For these models which have biases of different sign, the T_{max} and T_{min} biases can be much larger than T_{mean} bias. Therefore, it can be risky if we only assess models based on T_{mean}, which may lead to biased prediction of tropical cyclone, heatwaves, water mass formation and sea ice extent. For example, Oliver et al. (2019) used CMIP5 models including IPSL-CM5A-LR to estimate future changes in marine heatwaves to the end of the 21st century, and indicated significant increase in the intensity and duration of marine heatwaves. However, according to my study, even IPSL-CM6A-LR, as the CMIP6 version of IPSL-CM5A-LR, have an overall warm bias in summer, which may suggest overpredicted marine heatwave in the future projection in Oliver et al. (2019).

In regions with sinusoidal SST seasonal cycles, the biases in SST seasonal extrema are largely consistent with the annual mean SST biases. In regions with non-sinusoidal SST seasonal cycles, models tend to have biases in their seasonal cycles. In addition to the seasonal extrema of SST, the timing of seasonal extrema is also an important parameter to evaluate seasonal cycle. In the monsoon regions, phase biases are up to 6 months in some models, which may be linked to biased onset time of monsoon. Large seasonal variations in the SST bias occur in eastern boundary upwelling region, polar region and the eastern equatorial Atlantic. Large bias of SSS seasonal variation exists in equatorial and polar regions. The large seasonality of SST/SSS biases means that there is a poor simulation of the seasonal processes. It suggests that in these regions model evaluation should not only focus on annual mean values, seasonal extreme values also needs to be considered.

In this thesis, we compared 20 CMIP6 models using RMSE. Here we use RMSE rather than averaged bias for global SST because biases with opposite signs in different regions compensate each other when calculating global averaged bias. On a global scale, HadGEM3-GC31-LL and GFDL-CM4 have the best performance in simulating summer SST, while HadGEM3-GC31-MM and SAM0-UNICON have the best performance in simulating winter SST. These models can be useful tools for understanding and predicting SST seasonal variability. Models with best performance on annual mean SST are HadGEM3-GC31-LL and SAM0-UNICON, different from those for SST seasonal extrema. However, the above recommendation of models only relies on global RMSE. Your selection of the best model should also depend on the location you care about, as the model performance differs in locations.

Previous studies have emphasised the benefits of increasing ocean model horizontal resolution in SST simulation, whereas ocean vertical resolution has drawn much less attention. The High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6 was proposed to determine the robust benefits of increased horizontal model resolution based on multi-model ensemble simulations, but vertical resolution was not considered in the project (Haarsma et al., 2016). This thesis concludes that increased ocean vertical resolution also needs to be valued. By comparing 20 CMIP6 models, we found that SST is better simulated in models with higher ocean vertical resolution. The sensitivity of SST bias to ocean vertical resolution is larger in summer than in winter, especially at mid-high latitudes. The improvement of SST simulation shows geographical structures. T_{max} and T_{min} biases are especially sensitive to ocean vertical resolution in the eastern Pacific and eastern South Atlantic, where mixed layer is shallower than other regions. This is likely related to the ability of the higher resolution models to better represent the surface mixed layer, and particularly shallow mixed layers in summer. Models with coarse vertical resolution (few vertical levels or thick upper layer) are likely not able to realistically simulate vertical stratification of the upper ocean and results in biased SST.

Our results suggest that global SST biases are sensitive to ocean vertical resolution and there is a geographical pattern for the improvement of SST simulation. Over most of the world ocean SST biases can be reduced when increasing the ocean vertical resolution.

However, in some specific regions (e.g. subtropical North Pacific, South Atlantic, South Pacific and eastern Indian Ocean) enhancing ocean vertical resolution could increase T_{min} biases, which still needs to be explained.

In this thesis, we investigated the impact of model characteristics on global SST biases. The model characteristics include: ocean grid type, ocean vertical coordinate, ocean horizontal/vertical resolutions, atmosphere horizontal/vertical resolution and earth system model or not. Only ocean vertical resolution shows a significant relationship with SST bias. No clear relationship was found with any other model characteristics considered here. Displaced pole grid and tripolar grid were came up to avoid the pole problem; unstructured-mesh grid enables higher resolution in dynamically active regions. However, there is no significant impact of ocean grid type on RMSE of global SST, which may due to that ocean grid type only influences specific regions (e.g. the polar regions, the NAC, the Southern Ocean), not the whole world ocean. Higher atmosphere-ocean resolution leads to smaller SST biases in specific regions, for example, the eastern boundary upwelling regions (Kuhlbrodt et al., 2018) and the NAC (Andrews et al., 2019), but my study shows no significant impact of atmosphere-ocean resolution on the SST bias on a global scale.

Previous works indicated that higher ocean horizontal resolution improves the representation of boundary currents, ocean fronts and eddies, and hence reduces SST biases in these regions. Skákala et al. (2019) showed that the representation of SST in the Southern Ocean is substantially improved by increasing horizontal resolution from $1°$ to $1/12°$, using NEMO global model. de la Vara et al. (2020) used four different configurations of AWI-CM and demonstrated that the Agulhas Current, Benguela Current and coastal upwelling are better simulated with increased horizontal resolution. Chassignet et al. (2020) investigated the impact of horizontal resolution on SST bias based on four pairs of matched low and high resolution simulations and indicated that SST bias in western boundary currents, equatorial currents and ACC are significantly improved in the high-resolution models. Docquier et al. (2019) and Roberts et al. (2020) showed that finer ocean horizontal resolution results in a better represented Atlantic Ocean heat transport, resulting in a more realistic SST over the North Atlantic. In line with the studies mentioned above, my results also

demonstrated that higher ocean horizontal resolution reduces SST bias in regions with boundary currents, ocean fronts and eddies. However, in this thesis, higher horizontal resolution does not lead to an unambiguous reduced bias in all regions (consistent with Chassignet et al. (2020)), and therefore there is not a significant relationship between global SST RMSE and ocean horizontal resolution.

In this thesis, we also evaluated SSS seasonal extrema in HiGEM. It is found that the biases in S_{max} and S_{min} are largely accounted for by S_{mean}. Seasonal variation of SSS bias is more than 1 in the equatorial regions, likely due to poor simulated seasonal precipitation. In the polar regions (most of the Arctic and part of the Antarctic), there is also a significant seasonal variation of SSS bias in HiGEM based on WOA18. However, the seasonal SSS bias in the Arctic might be from the high uncertainty of SSS observation. Seasonal SSS bias accompanied with seasonal SST bias lead to seasonal surface density bias, which might contribute to the seasonality of MLD bias. The seasonal variation of SSS biases is relatively small compared to that of SST bias, considering their contributions to density of sea water.

For the future, we recommend more efforts on the evaluation of SST and SSS extrema, which are closely related to water mass formation and propagating climate signal into the deep ocean. Deep water and intermediate water form within the mixed layer in winter when heat loss and/or brine rejection result in dense surface water and destabilizes the water column. Therefore, bias in SST and SSS can be transferred into ventilated layers. By studying SST/SSS seasonal extrema and the related water mass formation, we can have a better understanding of global thermohaline ocean circulation and ocean's role in climate system.

5.4 Limitations and future work

Global SST RMSE were assessed against ocean vertical resolution, yielding a statistically significant relationship. However it can be difficult to assess the impact of ocean vertical resolution alone, because there are additional model characteristics differences. To isolate the impact of horizontal resolution, experiments at different

horizontal resolution (ranging from an atmosphere-ocean resolution of 130 km-1° to 25 km-1/12°) were performed based on HadGEM3-GC3.1, using the same forcings and initial conditions (Roberts et al., 2019). Modellers could do similar things (compare high resolution models with their low-resolution counterparts) to vertical resolution in the future, in order to determine the benefits of increased ocean model resolution on model simulations. Here we suggest modellers carry out an equivalent to HighResMIP for ocean vertical resolution.

In this thesis, we used only one ensemble member for each model. Model simulation can vary in ensemble members due to internal variability, given the initial condition uncertainty. For example, using coupled CESM over the historical period of 1920-2005, Murphy et al. (2021) found the mean variance over 41 ensemble members of AMV indices (average SST in 0°-60°N, 80°W-0°) is 0.011°C^2. The monthly ensemble spread of SST over 60°S-60°N is around 0.1°C during 1981-2010 in Feng et al. (2018). Considering the SST difference between ensemble members, our results can be changed if we pick another ensemble member. However, the results are still essentially robust as the ensemble spread is small compared to the SST biases in the models. If possible, we could use multi-ensemble mean instead of choosing only one ensemble member.

Our work focused on sea surface water, whereas intermediate water or deep water was not discussed in this thesis. By analysing seasonality of temperature bias at 30°W section in HiGEM (Fig. 5.1), we found that the seasonality of temperature bias mainly exists in the upper ocean (0-200 m), but there is also seasonal variation of temperature bias below the upper ocean in specific regions (e.g. 100-500 m at 40°S), which suggests bias in seasonal cycle of intermediate water. Intermediate waters take part of the Meridional Overturning Circulation and are of key importance to the transport of global heat, nutrients and carbon dioxide (Sallée et al., 2012). The characteristics of intermediate waters are strongly tied to the characteristics of the deep mixed layer where they developed. It could be interesting to evaluate seasonal cycle of intermediate water in CMIP6 models, so that we can see if the SST bias has penetrated into the intermediate layers.

It would also be useful to examine the seasonal cycle of sea surface water in CMIP5
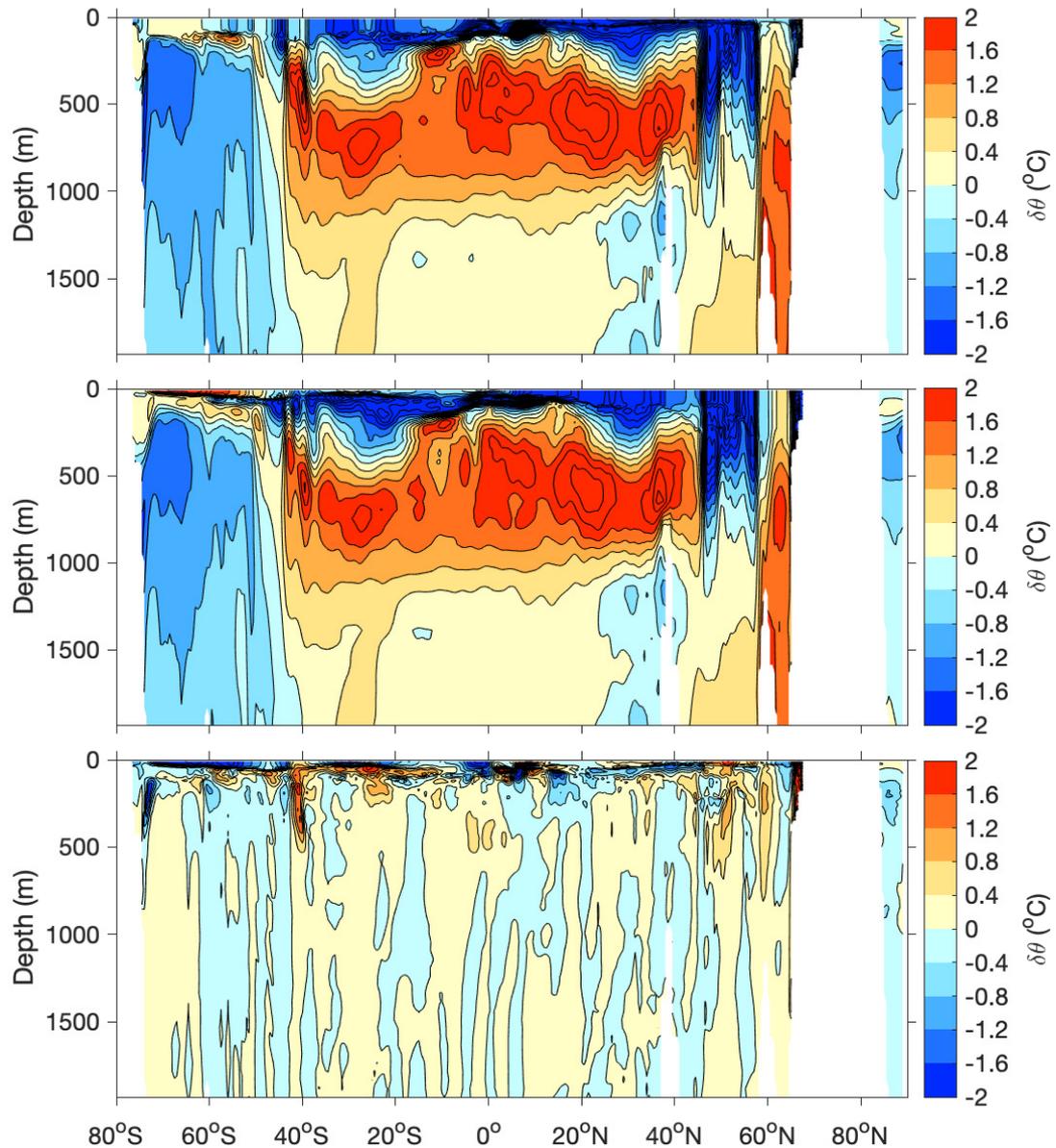
Figure 5.1: Potential temperature difference between HiGEM and MIMOC at 30°W (in the Atlantic Ocean) in (a) August (b) February and (c) the difference between (a) and (b).

models and compare with CMIP6 models, so we can see if there is a robust improvement from CMIP5 to CMIP6. We could also compare CMIP5 models and their CMIP6 counterparts with finer ocean vertical resolution to verify whether the increase in vertical resolution has a positive impact on the fidelity of the simulation of SST. Comparing the CMIP5 and CMIP6 versions of each model may help us find other model characteristics that benefit the simulation of seasonal cycle of sea surface water.

We have found that in most CMIP6 models and HiGEM, the warm biases in eastern

boundary upwelling regions are 1-5°C larger in winter than in summer. In this thesis, we have not explained why the warm biases are larger in winter than in summer. Underestimation of stratocumulus and insufficient upwelling are suggested causes for these warm SST biases (Richter, 2015). Letelier et al. (2009) suggested that the cooling effect of upwelling is the strongest in summer. The warm bias due to stratocumulus or upwelling is supposed to be amplified by the shallow mixed layer in summer. However, these seasonal processes contribute to a larger summer SST warm bias rather than a larger winter SST warm bias as seen in our results (Fig. 4.1). To explain the seasonality of SST bias in eastern boundary upwelling regions, further work (e.g. investigation of seasonal upwelling intensity and seasonal MLD in models) still needs to be done.

It is found that in the Southern Ocean, MIROC6 stands out as having an exceptionally large warm bias, especially in summer (more than 5°C). It also stands out from 21 other CMIP6 models with the largest warm bias of annual mean 0-100 m averaged temperature (Beadling et al., 2020). One possible reason is that the area of open ocean deep convection (which brings deep warm water to the surface) is larger in MIROC6 than in other CMIP6 models (Heuzé, 2021). However, more information (e.g. cloud, MLD, sea water temperature below the surface) is required to explain why MIROC6 has a much larger surface warm bias than other models.

Unlike most CMIP6 models that have similar SST bias patterns in summer and winter, IPSL-CM6A-LR and E3SM-1-0 have an overall warm bias in summer but an overall cold bias in winter. The biases in E3SM-1-0 may be attributed to its too shallow global annual average MLD (Golaz et al., 2019). As shallower mixed layer will intensify the effect of surface flux on SST, SST in E3SM-1-0 can be overly warmed in summer and overly cooled in winter. However the possible reasons for the biases in IPSL-CM6A-LR are still unclear. Much work is still to be done to explain why SST biases in these two models have different sign at most grid points.

This study only examined CMIP6 simulations. Analyses of biases in CMIP6 and AMIP6 (Atmospheric Model Intercomparison Project phase 6) simulation in combination could be done to separate the influences of atmospheric model errors and coupled feedback due to SST biases (Hyder et al., 2018). AMIP6 uses the same atmospheric models

as in CMIP6, but with a prescribed lower boundary condition of observed SST. To further examine the mechanisms behind SST bias, mixed layer heat budget could also be employed.

5.5   Summary of my thesis

In this thesis, we assessed the seasonal cycle of sea surface water characteristics in HiGEM and CMIP6 models based on an observational climatology. Because of the uncertainty of the observational climatology, we compared several climatologies and finally used WOA18 with the high uncertainty areas excluded. On a global scale, seasonal SST biases are consistent with biases in annual mean SST. However, in particular areas in some models, the amplitude and spatial pattern of SST bias vary seasonally. Large seasonal variations in amplitude of SST bias occur in eastern boundary upwelling regions, polar regions and eastern equatorial Atlantic. IPSL-CM6A-LR and E3SM-1-0 have an overall warm bias in summer and an overall cold bias in winter. These results demonstrate the importance of evaluating model performance not simply against annual mean properties. The impact of model characteristics on SST biases were investigated for the models we examined. It is found that models with increased ocean vertical resolution have a better representation of SST, particular summer SST. No significant relationship with ocean horizontal resolution is found.

Bibliography

AchutaRao, K. and Sperber, K. (2002). Simulation of the El Niño Southern Oscillation: Results from the coupled model intercomparison project. *Climate Dynamics*, 19(3-4):191–209.

Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., Burke, E., Dittus, A. J., Florek, P., and Gray, L. J. (2020). Historical simulations with HadGEM3-GC3. 1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, page e2019MS001995.

Andrews, T., Andrews, M. B., Bodas-Salcedo, A., Jones, G. S., Kuhlbrodt, T., Manners, J., Menary, M. B., Ridley, J., Ringer, M. A., Sellar, A. A., et al. (2019). Forcings, feedbacks, and climate sensitivity in HadGEM3-GC3. 1 and UKESM1. *Journal of Advances in Modeling Earth Systems*, 11(12):4377–4394.

Argo (2000). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). *SEANOE*.

Bayr, T., Domeisen, D. I., and Wengel, C. (2019). The effect of the equatorial Pacific cold SST bias on simulated ENSO teleconnections to the North Pacific and California. *Climate Dynamics*, 53(7):3771–3789.

Bayr, T., Latif, M., Dommenget, D., Wengel, C., Harlaß, J., and Park, W. (2018). Mean-state dependence of ENSO atmospheric feedbacks in climate models. *Climate Dynamics*, 50(9):3171–3194.

Beadling, R., Russell, J., Stouffer, R., Mazloff, M., Talley, L., Goodman, P., Sallée, J., Hewitt, H., Hyder, P., and Pandde, A. (2020). Representation of Southern Ocean

Properties across Coupled Model Intercomparison Project Generations: CMIP3 to CMIP6. *Journal of Climate*, 33(15):6555–6581.

Beaumet, J., Krinner, G., Déqué, M., Haarsma, R., and Li, L. (2019). Assessing bias corrections of oceanic surface conditions for atmospheric models. *Geoscientific Model Development*, 12(1):321–342.

Bi, D., Dix, M., Marsland, S., O'Farrell, S., Sullivan, A., Bodman, R., Law, R., Harman, I., Srbinovsky, J., Rashid, H. A., et al. (2020). Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model. *Journal of Southern Hemisphere Earth Systems Science*, 70(1):225–251.

Bingham, F. M., Foltz, G., and McPhaden, M. (2012). Characteristics of the seasonal cycle of surface layer salinity in the global ocean. *Ocean Science*, 8(5):915.

Bjerknes, J. (1966). A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus*, 18(4):820–829.

Bleck, R. (1978). On the use of hybrid vertical coordinates in numerical weather prediction models. *Monthly Weather Review*, 106(9):1233–1244.

Bleck, R. (2002). An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modelling*, 4(1):55–88.

Bloomfield, P. (2004). *Fourier analysis of time series: an introduction*. John Wiley & Sons.

Bodas-Salcedo, A., Williams, K., Field, P., and Lock, A. (2012). The surface downwelling solar radiation surplus over the Southern Ocean in the Met Office model: The role of midlatitude cyclone clouds. *Journal of Climate*, 25(21):7467–7486.

Bopp, L., Lévy, M., Resplandy, L., and Sallée, J.-B. (2015). Pathways of anthropogenic carbon subduction in the global ocean. *Geophysical Research Letters*, 42(15):6416–6423.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., and Bopp, L. (2020). Presentation and evaluation

of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, pages e2019MS002010–in.

Boyer, T. P., Antonov, J. I., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., Johnson, D. R., Locarnini, R. A., Mishonov, A. V., O'Brien, T. D., et al. (2013). World Ocean Database 2013.

Boyer, T. P., Garcia, H. E., Locarnini, R. A., Zweng, M. M., Mishonov, A. V., Reagan, J. R., Antonov, J. I., Baranova, O. K., Biddle, M. M., Johnson, D. R., et al. (2014). 2013 World Ocean Atlas Aids High-Resolution Climate Studies. *Eos, Transactions American Geophysical Union*, 95(41):369–370.

Braconnot, P., Hourdin, F., Bony, S., Dufresne, J. L., Grandpeix, J. Y., and Marti, O. (2007). Impact of different convective cloud schemes on the simulation of the tropical seasonal cycle in a coupled ocean-atmosphere model. *Climate Dynamics*, 29(5):501–520.

Brayshaw, D. J., Hoskins, B., and Blackburn, M. (2011). The basic ingredients of the North Atlantic storm track. Part II: Sea surface temperatures. *Journal of the Atmospheric Sciences*, 68(8):1784–1805.

Brown, B., Dunne, R., and Chansang, H. (1996). Coral bleaching relative to elevated seawater temperature in the Andaman Sea (Indian Ocean) over the last 50 years. *Coral Reefs*, 15(3):151–152.

Brown, J. R., Moise, A. F., and Delage, F. P. (2012). Changes in the South Pacific Convergence Zone in IPCC AR4 future climate projections. *Climate Dynamics*, 39(1-2):1–19.

Brown, J. R., Power, S. B., Delage, F. P., Colman, R. A., Moise, A. F., and Murphy, B. F. (2011). Evaluation of the South Pacific Convergence Zone in IPCC AR4 climate model simulations of the twentieth century. *Journal of Climate*, 24(6):1565–1582.

Burls, N. J., Muir, L., Vincent, E. M., and Fedorov, A. (2017). Extra-tropical origin of equatorial pacific cold bias in climate models with links to cloud albedo. *Climate Dynamics*, 49(5):2093–2113.

Chassignet, E., Yeager, S., Fox-Kemper, B., Bozec, A., Castruccio, F., Danabasoglu, G., Kim, W., Koldunov, N., Li, Y., Lin, P., Liu, H., Sein, D., Sidorenko, D., Wang, Q., and Xu, X. (2020). Impact of horizontal resolution on global ocean-sea-ice model simulations based on the experimental protocols of the Ocean Model Intercomparison Project phase 2 (OMIP-2). *Geoscientific Model Development Discussions*.

Chassignet, E. P., Smith, L. T., Halliwell, G. R., and Bleck, R. (2003). North Atlantic simulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference pressure, and thermobaricity. *Journal of Physical Oceanography*, 33(12):2504–2526.

Chelton, D. B., Schlax, M. G., Freilich, M. H., and Milliff, R. F. (2004). Satellite measurements reveal persistent small-scale features in ocean winds. *Science*, 303(5660):978–983.

Chen, C. and Wang, G. (2015). Role of North Pacific mixed layer in the response of SST annual cycle to global warming. *Journal of Climate*, 28(23):9451–9458.

Cheneka, B. R., Brienen, S., Fröhlich, K., Asharaf, S., and Früh, B. (2016). Searching for an added value of precipitation in downscaled seasonal hindcasts over East Africa: COSMO-CLM forced by MPI-ESM. *Advances in Meteorology*, 2016.

Cheung, W. W. and Frölicher, T. L. (2020). Marine heatwaves exacerbate climate change impacts for fisheries in the northeast Pacific. *Scientific Reports*, 10(1):1–10.

Craig, P. M., Ferreira, D., and Methven, J. (2017). The contrast between Atlantic and Pacific surface water fluxes. *Tellus A: Dynamic Meteorology and Oceanography*, 69(1):1330454.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., Emmons, L., Fasullo, J., Garcia, R., and Gettelman, A. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, 12(2):e2019MS001916.

Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Böning, C., Bozec, A., Canuto, V. M., et al. (2014). North Atlantic simulations in coordinated ocean-ice reference experiments phase II (CORE-II). Part I: mean states. *Ocean Modelling*, 73:76–107.

Dare, R. A. and McBride, J. L. (2011). The threshold sea surface temperature condition for tropical cyclogenesis. *Journal of Climate*, 24(17):4570–4576.

Davey, M., Huddleston, M., Sperber, K., Braconnot, P., Bryan, F., Chen, D., Colman, R., Cooper, C., Cubasch, U., Delecluse, P., et al. (2002). STOIC: a study of coupled model climatology and variability in tropical ocean regions. *Climate Dynamics*, 18(5):403–420.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, 109(C12).

De Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research C: Oceans*, 109:1–20.

de la Vara, A., Cabos, W., Sein, D. V., Sidorenko, D., Koldunov, N. V., Koseki, S., Soares, P. M., and Danilov, S. (2020). On the impact of atmospheric vs oceanic resolutions on the representation of the sea surface temperature in the South Eastern Tropical Atlantic. *Climate Dynamics*, pages 1–25.

De Szoeke, S. P. and Xie, S. P. (2008). The tropical eastern Pacific seasonal cycle: Assessment of errors and mechanisms in IPCC AR4 coupled ocean-atmosphere general circulation models. *Journal of Climate*, 21(11):2573–2590.

Deser, C., Alexander, M. A., Xie, S.-P., and Phillips, A. S. (2010). Sea surface temperature variability: Patterns and mechanisms. *Annual review of marine science*, 2:115–143.

Docquier, D., Grist, J. P., Roberts, M. J., Roberts, C. D., Semmler, T., Ponsoni, L., Massonnet, F., Sidorenko, D., Sein, D. V., Iovino, D., et al. (2019). Impact

of model resolution on Arctic sea ice and North Atlantic Ocean heat transport. *Climate Dynamics*, 53(7-8):4989–5017.

Dong, X., Jin, J., Liu, H., Zhang, H., Zhang, M., Lin, P., Zeng, Q., Zhou, G., Yu, Y., Song, M., et al. (2020). CAS-ESM2. 0 Model Datasets for the CMIP6 Ocean Model Intercomparison Project Phase 1 (OMIP1). *Advances in Atmospheric Sciences*, pages 1–10.

Durack, P. J., Wijffels, S. E., and Matear, R. J. (2012). Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science*, 336(6080):455–458.

Dwyer, J. G., Biasutti, M., and Sobel, A. H. (2012). Projected changes in the seasonal cycle of surface temperature. *Journal of Climate*, 25(18):6359–6374.

Eden, C., Greatbatch, R. J., and Böning, C. W. (2004). Adiabatically correcting an eddy-permitting model using large-scale hydrographic data: application to the Gulf Stream and the North Atlantic Current. *Journal of Physical Oceanography*, 34(4):701–719.

Evans, J. L. (1993). Sensitivity of tropical cyclone intensity to sea surface temperature. *Journal of Climate*, 6(6):1133–1140.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958.

Fan, T., Deser, C., and Schneider, D. P. (2014). Recent Antarctic sea ice trends in the context of Southern Ocean surface climate variations since 1950. *Geophysical Research Letters*, 41(7):2419–2426.

Fathrio, I., Manda, A., Iizuka, S., Kodama, Y.-M., and Ishida, S. (2017). Evaluation of CMIP5 models on sea surface salinity in the Indian Ocean. In *IOP Conference Series: Earth and Environmental Science*, volume 54, page 012039. IOP Publishing.

Feng, X., Haines, K., and de Boisséson, E. (2018). Coupling of surface air and sea

surface temperatures in the CERA-20C reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 144(710):195–207.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M. (2013). Evaluation of Climate Models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 741–866.

Flato, G. M. (2011). Earth system models: an overview. *Wiley Interdisciplinary Reviews: Climate Change*, 2(6):783–800.

Foo, S. A. and Asner, G. P. (2020). Sea surface temperature in coral reef restoration outcomes. *Environmental Research Letters*, 15(7):074045.

Frölicher, T. L. and Laufkötter, C. (2018). Emerging risks from marine heat waves. *Nature Communications*, 9(1):1–4.

Garcia-Eidell, C., Comiso, J. C., Dinnat, E., and Brucker, L. (2019). Sea surface salinity distribution in the Southern Ocean as observed from space. *Journal of Geophysical Research: Oceans*, 124(5):3186–3205.

Ge, X., Wang, W., Kumar, A., and Zhang, Y. (2017). Importance of the vertical resolution in simulating SST diurnal and intraseasonal variability in an oceanic general circulation model. *Journal of Climate*, 30(11):3963–3978.

Gent, P. R. and Mcwilliams, J. C. (1990). Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, 20(1):150–155.

Gilford, D. M., Solomon, S., and Emanuel, K. A. (2017). On the seasonal cycles of tropical cyclone potential intensity. *Journal of Climate*, 30(16):6085–6096.

Golaz, J. C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., and Bader, D. C. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, 11(7):2089–2129.

González-Haro, C., Ponte, A., and Autret, E. (2019). Quantifying tidal fluctuations in remote sensing infrared SST observations. *Remote Sensing*, 11(19):2313.

Gould, J., Roemmich, D., Wijffels, S., Freeland, H., Ignaszewsky, M., Jianping, X., Pouliquen, S., Desaubies, Y., Send, U., Radhakrishnan, K., et al. (2004). Argo profiling floats bring new era of in situ ocean observations. *Eos, Transactions American Geophysical Union*, 85(19):185–191.

Gouretski, V. (2018). World ocean circulation experiment-argo global hydrographic climatology. *Ocean Science*, 14:1127–1146.

Gouretski, V. and Koltermann, K. P. (2004). WOCE global hydrographic climatology. *Berichte des BSH*, 35:1–52.

Graham, J. A., Dinniman, M. S., and Klinck, J. M. (2016). Impact of model resolution for on-shelf heat transport along the West Antarctic Peninsula. *Journal of Geophysical Research: Oceans*, 121(10):7880–7897.

Greeves, C., Pope, V., Stratton, R., and Martin, G. (2007). Representation of Northern Hemisphere winter storm tracks in climate models. *Climate Dynamics*, 28(7-8):683–702.

Griffies, S. M., Biastoch, A., Böning, C., Bryan, F., Danabasoglu, G., Chassignet, E. P., England, M. H., Gerdes, R., Haak, H., Hallberg, R. W., et al. (2009). Coordinated ocean-ice reference experiments (COREs). *Ocean Modelling*, 26(1-2):1–46.

Griffies, S. M., Böning, C., Bryan, F. O., Chassignet, E. P., Gerdes, R., Hasumi, H., Hirst, A., Treguier, A.-M., and Webb, D. (2000). Developments in ocean climate modelling. *Ocean Modelling*, 2(3-4):123–192.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., et al. (2016). High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6. *Geoscientific Model Development*, 9(11):4185–4208.

Hallam, S., Guishard, M., Josey, S. A., Hyder, P., and Hirschi, J. (2021). Increasing tropical cyclone intensity and potential intensity in the subtropical Atlantic around

Bermuda from an ocean heat content perspective 1955–2019. *Environmental Research Letters*, 16(3):034052.

Harlaß, J., Latif, M., and Park, W. (2018). Alleviating tropical Atlantic sector biases in the Kiel climate model by enhancing horizontal and vertical atmosphere model resolution: climatology and interannual variability. *Climate Dynamics*, 50(7-8):2605–2635.

Harrison, M., Adcroft, A., and Hallberg, R. (2014). Atlantic watermass and circulation response to persistent freshwater forcing in two coupled general circulation models. *Climate Dynamics*, 42(1-2):59–68.

Haslinger, K., Anders, I., and Hofstätter, M. (2013). Regional climate modelling over complex terrain: an evaluation study of COSMO-CLM hindcast model runs for the Greater Alpine Region. *Climate Dynamics*, 40(1-2):511–529.

Held, I., Guo, H., Adcroft, A., Dunne, J., Horowitz, L., Krasting, J., Shevliakova, E., Winton, M., Zhao, M., and Bushuk, M. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, 11(11):3691–3727.

Heuzé, C. (2021). Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models. *Ocean Science*, 17(1):59–90.

Heuzé, C., Heywood, K. J., Stevens, D. P., and Ridley, J. K. (2013). Southern Ocean bottom water characteristics in CMIP5 models. *Geophysical Research Letters*, 40(7):1409–1414.

Hewitt, H. T., Bell, M. J., Chassignet, E. P., Czaja, A., Ferreira, D., Griffies, S. M., Hyder, P., McClean, J. L., New, A. L., and Roberts, M. J. (2017). Will high-resolution global ocean models benefit coupled predictions on short-range to climate timescales? *Ocean Modelling*, 120:120–136.

Hieronymus, M., Nilsson, J., and Nycander, J. (2014). Water mass transformation in salinity–temperature space. *Journal of Physical Oceanography*, 44(9):2547–2568.

Holland, G. J. (1997). The maximum potential intensity of tropical cyclones. *Journal of the Atmospheric Sciences*, 54(21):2519–2541.

Holton, L., Deshayes, J., Backeberg, B., Loveday, B., Hermes, J., and Reason, C. (2017). Spatio-temporal characteristics of Agulhas leakage: A model inter-comparison study. *Climate Dynamics*, 48(7-8):2107–2121.

Hu, Z. Z., Huang, B., and Pegion, K. (2008). Leading patterns of the tropical Atlantic variability in a coupled general circulation model. *Climate Dynamics*, 30(7-8):703–726.

Huang, B., Hu, Z. Z., and Jha, B. (2007). Evolution of model systematic errors in the tropical Atlantic basin from coupled climate hindcasts. *Climate Dynamics*, 28(7-8):661–682.

Hughes, T. P., Anderson, K. D., Connolly, S. R., Heron, S. F., Kerry, J. T., Lough, J. M., Baird, A. H., Baum, J. K., Berumen, M. L., Bridge, T. C., et al. (2018). Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science*, 359(6371):80–83.

Hyder, P., Edwards, J. M., Allan, R. P., Hewitt, H. T., Bracegirdle, T. J., Gregory, J. M., Wood, R. A., Meijers, A. J., Mulcahy, J., Field, P., et al. (2018). Critical Southern Ocean climate model biases traced to atmospheric model cloud errors. *Nature Communications*, 9(1):1–17.

Jiang, W., Huang, P., Huang, G., and Ying, J. (2021). Origins of the excessive westward extension of ENSO SST simulated in CMIP5 and CMIP6 models. *Journal of Climate*, 34(8):2839–2851.

Jiang, X. and Li, J. (2011). Influence of the annual cycle of sea surface temperature on the monsoon onset. *Journal of Geophysical Research: Atmospheres*, 116(D10).

Jones, P. W., Worley, P. H., Yoshida, Y., White III, J. B., and Levesque, J. (2005). Practical performance portability in the Parallel Ocean Program (POP). *Concurrency and Computation: Practice and Experience*, 17(10):1317–1327.

Jones, T., Parrish, J. K., Peterson, W. T., Bjorkstedt, E. P., Bond, N. A., Ballance, L. T., Bowes, V., Hipfner, J. M., Burgess, H. K., Dolliver, J. E., et al. (2018). Massive mortality of a planktivorous seabird in response to a marine heatwave. *Geophysical Research Letters*, 45(7):3193–3202.

Kara, A. B., Rochford, P. A., and Hurlburt, H. E. (2003). Mixed layer depth variability over the global ocean. *Journal of Geophysical Research: Oceans*, 108(C3).

Keeley, S. P., Sutton, R. T., and Shaffrey, L. C. (2012). The impact of North Atlantic sea surface temperature errors on the simulation of North Atlantic European region climate. *Quarterly Journal of the Royal Meteorological Society*, 138(668):1774–1783.

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., Ackerman, A. S., Aleinov, I., Bauer, M., and Bleck, R. (2020). GISS-E2.1: Configurations and climatology. *Journal of Advances in Modeling Earth Systems*, page e2019MS002025.

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., Kinter, J. L., Loft, R., Rousset, C., Siqueira, L., et al. (2012). Impact of ocean model resolution on CCSM climate simulations. *Climate Dynamics*, 39(6):1303–1328.

Kotlarski, S., Block, A., Böhm, U., Jacob, D., Keuler, K., Knoche, R., Rechid, D., and Walter, A. (2005). Regional climate model simulations as input for hydrological applications: evaluation of uncertainties. *Advances in Geosciences*, 5:119–125.

Kotlarski, S., Szabó, P., Herrera, S., Räty, O., Keuler, K., Soares, P. M., Cardoso, R. M., Bosshard, T., Pagé, C., Boberg, F., et al. (2019). Observational uncertainty and regional climate model evaluation: A pan-European perspective. *International Journal of Climatology*, 39(9):3730–3749.

Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., Hill, R., Graham, T., Ridley, J., Blaker, A., et al. (2018). The low-resolution version of HadGEM3 GC3.1: Development and evaluation for global climate. *Journal of Advances in Modeling Earth Systems*, 10(11):2865–2888.

Latif, M., Sperber, K., Arblaster, J., Braconnot, P., Chen, D., Colman, A., Cubasch, U., Cooper, C., Delecluse, P., DeWitt, D., et al. (2001). ENSIP: the El Nino simulation intercomparison project. *Climate Dynamics*, 18(3-4):255–276.

Lauer, A. and Hamilton, K. (2013). Simulating clouds with global climate models: A

comparison of CMIP5 results with CMIP3 and satellite data. *Journal of Climate*, 26(11):3823–3845.

Law, R. M., Ziehn, T., Matear, R. J., Lenton, A., Chamberlain, M. A., Stevens, L. E., Ying-Ping, W., Srbinovsky, J., Bi, D., and Yan, H. (2017). The carbon cycle in the Australian Community Climate and Earth System Simulator (ACCESS-ESM1)–Part 1: Model description and pre-industrial simulation. *Geoscientific Model Development*, 10(7):2567.

Lee, M. M., Nurser, A. G., Stevens, I., and Sallée, J.-B. (2011). Subduction over the Southern Indian Ocean in a high-resolution atmosphere–ocean coupled model. *Journal of Climate*, 24(15):3830–3849.

Legg, S., Briegleb, B., Chang, Y., Chassignet, E. P., Danabasoglu, G., Ezer, T., Gordon, A. L., Griffies, S., Hallberg, R., Jackson, L., et al. (2009). Improving oceanic overflow representation in climate models: the gravity current entrainment climate process team. *Bulletin of the American Meteorological Society*, 90(5):657–670.

Legg, S., Hallberg, R. W., and Girton, J. B. (2006). Comparison of entrainment in overflows simulated by z-coordinate, isopycnal and non-hydrostatic models. *Ocean Modelling*, 11(1-2):69–97.

Letelier, J., Pizarro, O., and Nuñez, S. (2009). Seasonal variability of coastal upwelling and the upwelling front off central Chile. *Journal of Geophysical Research: Oceans*, 114(C12).

Levine, R. C. and Turner, A. G. (2012). Dependence of Indian monsoon rainfall on moisture fluxes across the Arabian Sea and the impact of coupled model sea surface temperature biases. *Climate Dynamics*, 38(11-12):2167–2190.

Li, G. and Xie, S. P. (2012). Origins of tropical-wide SST biases in CMIP multi-model ensembles. *Geophysical Research Letters*, 39(22).

Liang, X. and Wu, L. (2013). Effects of solar penetration on the annual cycle of sea surface temperature in the North Pacific. *Journal of Geophysical Research-Oceans*, 118:2793–2801.

Lin, J. (2007). The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean–atmosphere feedback analysis. *Journal of Climate*, 20(18):4497–4525.

Liu, F., Lu, J., Luo, Y., Huang, Y., and Song, F. (2020). On the oceanic origin for the enhanced seasonal cycle of SST in the midlatitudes under global warming. *Journal of Climate*, 33(19):8401–8413.

Liu, G., Strong, A. E., Skirving, W., and Arzayus, L. F. (2006). Overview of NOAA coral reef watch program's near-real time satellite global coral bleaching monitoring activities. In *Proceedings of the 10th international coral reef symposium*, volume 1793, pages 1783–1793. Okinawa Gurugram.

Liu, Q., Xie, S.-P., Li, L., and Maximenko, N. A. (2005). Ocean thermal advective effect on the annual range of sea surface temperature. *Geophysical Research Letters*, 32(24).

Liu, W., Liu, Z., and Brady, E. C. (2014). Why is the AMOC monostable in coupled general circulation models? *Journal of Climate*, 27(6):2427–2443.

Locarnini, R. A., Mishonov, A. V., Baranova, O. K., Boyer, T. P., Zweng, M. M., Garcia, H. E., Reagan, J. R., Seidov, D., Weathers, K., Paver, C. R., and Smolyar, I. (2018). *World Ocean Atlas 2018, Volume 1: Temperature*.

Locarnini, R. A., Misnov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., and Seidov, D. (2013). *World Ocean Atlas 2013, Volume 1: Temperature*.

Lozier, M. S., McCartney, M. S., and Owens, W. B. (1994). Anomalous anomalies in averaged hydrographic data. *Journal of Physical Oceanography*, 24(12):2624–2638.

Lürling, M., Eshetu, F., Faassen, E. J., Kosten, S., and Huszar, V. L. (2013). Comparison of cyanobacterial and green algal growth rates at different temperatures. *Freshwater Biology*, 58(3):552–559.

Marathayil, D., Turner, A., Shaffrey, L., and Levine, R. (2013). Systematic winter sea-surface temperature biases in the northern Arabian Sea in HiGEM and the CMIP3 models. *Environmental Research Letters*, 8(1):014028.

Marzocchi, A., Hirschi, J. J.-M., Holliday, N. P., Cunningham, S. A., Blaker, A. T., and Coward, A. C. (2015). The North Atlantic subpolar circulation in an eddy-resolving global ocean model. *Journal of Marine Systems*, 142:126–143.

Maynard, J. A., Turner, P. J., Anthony, K. R., Baird, A. H., Berkelmans, R., Eakin, C. M., Johnson, J., Marshall, P. A., Packer, G. R., Rea, A., et al. (2008). ReefTemp: An interactive monitoring system for coral bleaching using high-resolution SST and improved stress predictors. *Geophysical Research Letters*, 35(5).

McDougall, T. J. and Barker, P. M. (2011). Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox. *SCOR/IAPSO WG*, 127:1–28.

McGregor, S., Timmermann, A., Stuecker, M. F., England, M. H., Merrifield, M., Jin, F.-F., and Chikamoto, Y. (2014). Recent Walker circulation strengthening and Pacific cooling amplified by Atlantic warming. *Nature Climate Change*, 4(10):888.

McKenna, S., Santoso, A., Gupta, A. S., Taschetto, A. S., and Cai, W. (2020). Indian Ocean Dipole in CMIP5 and CMIP6: characteristics, biases, and links to ENSO. *Scientific Reports*, 10(1):1–13.

Mechoso, C. R., Robertson, A. W., Barth, N., Davey, M., Delecluse, P., Gent, P., Ineson, S., Kirtman, B., Latif, M., Treut, H. L., et al. (1995). The seasonal cycle over the tropical Pacific in coupled ocean-atmosphere general circulation models. *Monthly Weather Review*, 123(9):2825–2838.

Meehl, G. A., Covey, C., McAvaney, B., Latif, M., and Stouffer, R. J. (2005). Overview of the coupled model intercomparison project. *Bulletin of the American Meteorological Society*, 86(1):89–93.

Misra, V., Marx, L., Brunke, M., and Zeng, X. (2008). The equatorial Pacific cold tongue bias in a coupled climate model. *Journal of Climate*, 21(22):5852–5869.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., and Haak, H. (2018). A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1. 2-HR). *Journal of Advances in Modeling Earth Systems*, 10(7):1383–1413.

Murphy, L. N., Klavans, J. M., Clement, A. C., and Cane, M. A. (2021). Investigating the roles of external forcing and ocean circulation on the Atlantic multidecadal SST variability in a large ensemble climate model hierarchy. *Journal of Climate*, 34(12):4835–4849.

Murray, R. J. (1996). Explicit generation of orthogonal grids for ocean models. *Journal of Computational Physics*, 126(2):251–273.

Nagura, M., Sasaki, W., Tozuka, T., Luo, J. J., Behera, S. K., and Yamagata, T. (2013). Longitudinal biases in the Seychelles Dome simulated by 35 ocean-atmosphere coupled general circulation models. *Journal of Geophysical Research: Oceans*, 118(2):831–846.

Oliver, E. C., Burrows, M. T., Donat, M. G., Sen Gupta, A., Alexander, L. V., Perkins-Kirkpatrick, S. E., Benthuysen, J. A., Hobday, A. J., Holbrook, N. J., Moore, P. J., et al. (2019). Projected marine heatwaves in the 21st century and the potential for ecological impact. *Frontiers in Marine Science*, 6:734.

Palmen, E. (1948). On the formation and structure of tropical hurricanes. *Geophysica*, 3(1):26–38.

Parekh, A., Chowdary, J. S., Sayantani, O., Fousiya, T., and Gnanaseelan, C. (2016). Tropical Indian Ocean surface salinity bias in Climate Forecasting System coupled models and the role of upper ocean processes. *Climate Dynamics*, 46(7-8):2403–2422.

Park, S., Shin, J., Kim, S., Oh, E., and Kim, Y. (2019). Global climate simulated by the Seoul National University atmosphere model version 0 with a unified convection scheme (SAM0-UNICON). *Journal of Climate*, 32(10):2917–2949.

Park, T., Park, W., and Latif, M. (2016). Correcting North Atlantic sea surface salinity biases in the Kiel Climate Model: influences on ocean circulation and Atlantic Multidecadal Variability. *Climate Dynamics*, 47(7):2543–2560.

Pope, V., Brown, S., Clark, R., Collins, M., Collins, W., Dearden, C., Gunson, J., Harris, G., Jones, C., Keen, A., et al. (2007). The Met Office Hadley Centre climate modelling capability: the competing requirements for improved resolution,

complexity and dealing with uncertainty. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1860):2635–2657.

Prodhomme, C., Terray, P., Masson, S., Boschat, G., and Izumo, T. (2015). Oceanic factors controlling the Indian summer monsoon onset in a coupled model. *Climate Dynamics*, 44(3-4):977–1002.

Prodhomme, C., Terray, P., Masson, S., Izumo, T., Tozuka, T., and Yamagata, T. (2014). Impacts of Indian Ocean SST biases on the Indian Monsoon: as simulated in a global coupled model. *Climate Dynamics*, 42(1-2):271–290.

Prodhomme, C., Voldoire, A., Exarchou, E., Deppenmeier, A.-L., García-Serrano, J., and Guemas, V. (2019). How does the seasonal cycle control equatorial Atlantic interannual variability? *Geophysical Research Letters*, 46(2):916–922.

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

Richardson, P. L. (2007). Agulhas leakage into the Atlantic estimated with subsurface floats and surface drifters. *Deep Sea Research Part I: Oceanographic Research Papers*, 54(8):1361–1389.

Richter, I. (2015). Climate model biases in the eastern tropical oceans: causes, impacts and ways forward. *Wiley Interdisciplinary Reviews: Climate Change*, 6(3):345–358.

Richter, I. and Tokinaga, H. (2020). An overview of the performance of CMIP6 models in the tropical Atlantic: mean state, variability, and remote impacts. *Climate Dynamics*, 55(9):2579–2601.

Richter, I. and Xie, S. P. (2008). On the origin of equatorial Atlantic biases in coupled general circulation models. *Climate Dynamics*, 31(5):587–598.

Richter, I., Xie, S.-P., Behera, S. K., Doi, T., and Masumoto, Y. (2014). Equatorial Atlantic variability and its relation to mean state biases in CMIP5. *Climate Dynamics*, 42(1-2):171–188.

Richter, I., Xie, S.-P., Wittenberg, A. T., and Masumoto, Y. (2012). Tropical Atlantic biases and their relation to surface wind stress and terrestrial precipitation. *Climate Dynamics*, 38(5-6):985–1001.

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., Jackson, L. C., Kuhlbrodt, T., Mathiot, P., Roberts, C. D., et al. (2019). Description of the resolution hierarchy of the global coupled HadGEM3-GC3. 1 model as used in CMIP6 HighResMIP experiments. *Geoscientific Model Development*, 12(12):4999–5028.

Roberts, M. J., Hewitt, H. T., Hyder, P., Ferreira, D., Josey, S. A., Mizielinski, M., and Shelly, A. (2016). Impact of ocean resolution on coupled air-sea fluxes and large-scale climate. *Geophysical Research Letters*, 43(19):10–430.

Roberts, M. J., Jackson, L. C., Roberts, C. D., Meccia, V., Docquier, D., Koenigk, T., Ortega, P., Moreno-Chamarro, E., Bellucci, A., Coward, A., Drijfhout, S., Exarchou, E., Gutjahr, O., Hewitt, H., Iovino, D., Lohmann, K., Putrasahan, D., Schiemann, R., Seddon, J., Terray, L., Xu, X., Zhang, Q., Chang, P., Yeager, S. G., Castruccio, F. S., Zhang, S., and Wu, L. (2020). Sensitivity of the Atlantic Meridional Overturning Circulation to Model Resolution in CMIP6 HighResMIP Simulations and Implications for Future Changes. *Journal of Advances in Modeling Earth Systems*.

Rouault, M., Florenchie, P., Fauchereau, N., and Reason, C. J. (2003). South East tropical Atlantic warm events and southern African rainfall. *Geophysical Research Letters*, 30(5).

Roxy, M. (2014). Sensitivity of precipitation to sea surface temperature over the tropical summer monsoon region and its quantification. *Climate Dynamics*, 43(5-6):1159–1169.

Sallée, J.-B., Matear, R. J., Rintoul, S. R., and Lenton, A. (2012). Localized subduction of anthropogenic carbon dioxide in the Southern Hemisphere oceans. *Nature Geoscience*, 5(8):579–584.

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J., Bracegirdle, T. J., and Wang, Z. (2013). Assessment of Southern Ocean mixed-layer depths in CMIP5 models:

Historical bias and forcing response. *Journal of Geophysical Research: Oceans*, 118(4):1845–1862.

Sato, O. and Polito, P. (2014). Observation of South Atlantic subtropical mode waters with Argo profiling float data. *Journal of Geophysical Research: Oceans*, 119(5):2860–2881.

Scannell, H. A. and McPhaden, M. J. (2018). Seasonal mixed layer temperature balance in the southeastern tropical Atlantic. *Journal of Geophysical Research: Oceans*, 123(8):5557–5570.

Schmidtko, S., Johnson, G. C., and Lyman, J. M. (2013). MIMOC: A global monthly isopycnal upper-ocean climatology with mixed layers. *Journal of Geophysical Research: Oceans*, 118(4):1658–1672.

Seidov, D., Mishonov, A., Reagan, J., and Parsons, R. (2017). Multidecadal variability and climate shift in the North Atlantic Ocean. *Geophysical Research Letters*, 44(10):4985–4993.

Seland, Ø., Bensten, M., Graff, L., Olivié, D., Toniazzo, T., and Gjermundsen, A. (2020). The Norwegian Earth System Model, NorESM2–Evaluation of the CMIP6 DECK and historical simulations. *Geoscientific Model Development Discussions*.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'connor, F. M., Stringer, M., Hill, R., and Palmieri, J. (2019). UKESM1: Description and evaluation of the UK Earth System Model. *Journal of Advances in Modeling Earth Systems*, 11(12):4513–4558.

Semmler, T., Danilov, S., Gierz, P., Goessling, H. F., Hegewald, J., Hinrichs, C., Koldunov, N., Khosravi, N., Mu, L., Rackow, T., et al. (2020). Simulations for CMIP6 with the AWI climate model AWI-CM-1-1. *Journal of Advances in Modeling Earth Systems*, 12(9):e2019MS002009.

Sévellec, F., Fedorov, A. V., and Liu, W. (2017). Arctic sea-ice decline weakens the atlantic meridional overturning circulation. *Nature Climate Change*, 7(8):604–610.

Shaffrey, L. C., Stevens, I., Norton, W. A., Roberts, M. J., Vidale, P. L., Harle, J. D., Jrrar, A., Stevens, D. P., Woodage, M. J., Demory, M. E., Donners, J., Clark,

D. B., Clayton, A., Cole, J. W., Wilson, S. S., Connolley, W. M., Davies, T. M., Iwi, A. M., Johns, T. C., King, J. C., New, A. L., Slingo, J. M., Slingo, A., Steenman-Clark, L., and Martin, G. M. (2009). U.K. HiGEM: The new U.K. high-resolution global environment model - Model description and basic evaluation. *Journal of Climate*, 22:1861–1896.

Shchepetkin, A. F. and McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, 9(4):347–404.

Shu, Q., Wang, Q., Song, Z., Qiao, F., Zhao, J., Chu, M., and Li, X. (2020). Assessment of sea ice extent in CMIP6 with comparison to observations and CMIP5. *Geophysical Research Letters*, 47(9):e2020GL087965.

Sijikumar, S. and Rajeev, K. (2012). Role of the Arabian Sea warm pool on the precipitation characteristics during the monsoon onset period. *Journal of Climate*, 25(6):1890–1899.

Skákala, J., Smyth, T. J., Torres, R., Buckingham, C. E., Brearley, A., Hyder, P., and Coward, A. C. (2019). SST dynamics at different scales: Evaluating the oceanographic model resolution skill to represent SST processes in the Southern Ocean. *Journal of Geophysical Research: Oceans*, 124(4):2546–2570.

Small, R. J., Curchitser, E., Hedstrom, K., Kauffman, B., and Large, W. G. (2015). The Benguela upwelling system: Quantifying the sensitivity to resolution and coastal wind representation in a global climate model. *Journal of Climate*, 28(23):9409–9432.

Soares, S. M., Richards, K. J., Bryan, F. O., and Yoneyama, K. (2019). On the Seasonal Cycle of the Tropical South Indian Ocean. Part I: Mixed Layer Heat and Salt Budgets. *Journal of Climate*, 32(6):1951–1972.

Solomon, A. and Newman, M. (2012). Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record. *Nature Climate Change*, 2(9):691.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., Miller, H., et al. (2007). The physical science basis. *Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*, 2007:235–337.

Song, F. and Zhang, G. J. (2020). The Impacts of Horizontal Resolution on the Seasonally Dependent Biases of the Northeastern Pacific ITCZ in Coupled Climate Models. *Journal of Climate*, 33(3):941–957.

Sprintall, J. and Tomczak, M. (1992). Evidence of the barrier layer in the surface layer of the tropics. *Journal of Geophysical Research: Oceans*, 97(C5):7305–7316.

Stine, A. R., Huybers, P., and Fung, I. Y. (2009). Changes in the phase of the annual cycle of surface temperature. *Nature*, 457:435–440.

Stocker, T. (2014). *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., et al. (2018). UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions. *Geoscientific Model Development*, 11(8):3187–3213.

Sun, Y., Zhong, Z., Li, T., Yi, L., Hu, Y., Wan, H., Chen, H., Liao, Q., Ma, C., and Li, Q. (2017). Impact of ocean warming on tropical cyclone size and its destructiveness. *Scientific Reports*, 7(1):1–10.

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., and Hanna, S. (2019). The Canadian Earth System Model version 5 (CanESM5.0.3). *Geoscientific Model Development*, 12(11):4823–4873.

Talley, L. D. (1999). Some aspects of ocean heat transport by the shallow, intermediate and deep overturning circulations. *Geophysical Monograph-American Geophysical Union*, 112:1–22.

Talley, L. D. (2011). *Descriptive physical oceanography: an introduction*. Academic Press.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., et al. (2019). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, 12(7):2727–2765.

Terhaar, J., Frölicher, T. L., and Joos, F. (2021). Southern ocean anthropogenic carbon sink constrained by sea surface salinity. *Science Advances*, 7(18):eabd5964.

Tory, K. J. and Dare, R. A. (2015). Sea surface temperature thresholds for tropical cyclone formation. *Journal of Climate*, 28(20):8171–8183.

Trenberth, K. E. (1983). What are the seasons? *Bulletin of the American Meteorological Society*, 64(11):1276–1282.

Trenberth, K. E., Smith, L., Qian, T., Dai, A., and Fasullo, J. (2007). Estimates of the global water budget and its annual cycle using observational and model data. *Journal of Hydrometeorology*, 8(4):758–769.

Trenberth, K. E. and Stepaniak, D. P. (2001). Indices of el niño evolution. *Journal of Climate*, 14(8):1697–1701.

Turkington, T., Timbal, B., and Rahmat, R. (2019). The impact of global warming on sea surface temperature based El Niño–Southern Oscillation monitoring indices. *International Journal of Climatology*, 39(2):1092–1103.

Volodin, E., Mortikov, E., Kostrykin, S., Galin, V. Y., Lykossov, V., Gritsun, A., Diansky, N., Gusev, A., and Iakovlev, N. (2017). Simulation of the present-day climate with the climate model INMCM5. *Climate Dynamics*, 49(11-12):3715–3734.

Wagner, S., Kunstmann, H., Bárdossy, A., et al. (2007). Uncertainties in water balance estimations due to scarce meteorological information: a case study for the White Volta catchment in West Africa. *Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management (Proceedings of Symposim HS2004 at IUGG2007, Perugia, July 2007)*, pages 86–97.

Wang, C., Zhang, L., Lee, S.-K., Wu, L., and Mechoso, C. R. (2014). A global perspective on CMIP5 climate model biases. *Nature Climate Change*, 4(3):201–205.

Wang, C., Zou, L., and Zhou, T. (2018). SST biases over the Northwest Pacific and possible causes in CMIP5 models. *Science China Earth Sciences*, 61(6):792–803.

Weijer, W. and Van Sebille, E. (2014). Impact of agulhas leakage on the atlantic overturning circulation in the ccsm4. *Journal of Climate*, 27(1):101–110.

Wentz, F. J., Gentemann, C., Smith, D., and Chelton, D. (2000). Satellite measurements of sea surface temperature through clouds. *Science*, 288(5467):847–850.

Willebrand, J., Barnier, B., Böning, C., Dieterich, C., Killworth, P. D., Le Provost, C., Jia, Y., Molines, J. M., and New, A. L. (2001). Circulation characteristics in three eddy-permitting models of the North Atlantic. *Progress in Oceanography*, 48(2-3):123–161.

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., et al. (2019). The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6. *Geoscientific Model Development*, 12:1573–1600.

Wu, T., Zhang, F., Zhang, J., Jie, W., Zhang, Y., Wu, F., Li, L., Yan, J., Liu, X., Lu, X., et al. (2020). Beijing Climate Center Earth System Model version 1 (BCC-ESM1): model description and evaluation of aerosol simulations. *Geoscientific Model Development*, 13(3).

Xavier, P. K., Duvel, J.-P., and Doblas-Reyes, F. J. (2008). Boreal summer intraseasonal variability in coupled seasonal hindcasts. *Journal of Climate*, 21(17):4477–4497.

Xie, S.-P., Annamalai, H., Schott, F. A., and McCreary, J. P. J. (2002). Structure and mechanisms of south Indian Ocean climate variability. *Journal of Climate*, 15(8):864–878.

Xie, S.-P., Deser, C., Vecchi, G. A., Ma, J., Teng, H., and Wittenberg, A. T. (2010).

Global warming pattern formation: Sea surface temperature and rainfall. *Journal of Climate*, 23(4):966–986.

Yashayaev, I. M. and Zveryaev, I. I. (2001). Climate of the seasonal cycle in the North Pacific and the North Atlantic oceans. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 21(4):401–417.

Yokoi, T., Tozuka, T., and Yamagata, T. (2009). Seasonal variations of the Seychelles Dome simulated in the CMIP3 models. *Journal of Physical Oceanography*, 39(2):449–457.

Yu, L. (2007). Global variations in oceanic evaporation (1958–2005): The role of the changing wind speed. *Journal of Climate*, 20(21):5376–5390.

Yu, L. (2011). A global relationship between the ocean water cycle and near-surface salinity. *Journal of Geophysical Research: Oceans*, 116(C10).

Yu, L., Bingham, F. M., Lee, T., Dinnat, E. P., Fournier, S., Melnichenko, O., Tang, W., and Yueh, S. H. (2021). Revisiting the global patterns of seasonal cycle in sea surface salinity. *Journal of Geophysical Research: Oceans*, 126(4):e2020JC016789.

Zhang, L. and Zhao, C. (2015). Processes and mechanisms for the model sst biases in the North Atlantic and North Pacific: a link with the atlantic meridional overturning circulation. *Journal of Advances in Modeling Earth Systems*, 7(2):739–758.

Zhu, Y., Zhang, R.-H., and Sun, J. (2020). North Pacific upper-ocean cold temperature biases in CMIP6 simulations and the role of regional vertical mixing. *Journal of Climate*, 33(17):7523–7538.

Zweng, M. M., Reagan, J. R., Antonov, J. I., Locarini, R. A., Mishonov, A. V., Boyer, T. P., Garcia, H. E., Baranova, O. K., Johnson, D. R., Seidov, D., and Biddle, M. M. (2013). World Ocean Atlas 2013 Volume 2: Salinity. *World Ocean Atlas 2013*.

Zweng, M. M., Reagan, J. R., Seidov, D., Boyer, T. P., Locarnini, R. A., Garcia, H. E., Mishonov, A. V., Baranova, O. K., Weathers, K., Paver, C. R., and Smolyar, I. (2018). *World Ocean Atlas 2018, Volume 2: Salinity*.